



HAL
open science

Découverte de nouveaux transcrits de fusion dans des tumeurs pédiatriques en rechute et caractérisation fonctionnelle d'un nouvel oncogène LMO3-BORCS5

Célia Dupain Jourda

► **To cite this version:**

Célia Dupain Jourda. Découverte de nouveaux transcrits de fusion dans des tumeurs pédiatriques en rechute et caractérisation fonctionnelle d'un nouvel oncogène LMO3-BORCS5. Cancer. Université Paris Saclay (COMUE), 2018. Français. NNT : 2018SACLS310 . tel-02426230

HAL Id: tel-02426230

<https://theses.hal.science/tel-02426230>

Submitted on 2 Jan 2020

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

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

Découverte de nouveaux transcrits de fusion dans des tumeurs pédiatriques en rechute et caractérisation fonctionnelle d'un nouvel oncogène LMO3-BORCS5

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

École doctorale n°568 : **Signalisations et réseaux intégratifs en biologie (Biosigne)**

Spécialité de doctorat: Aspects moléculaires et cellulaires de la biologie

Thèse présentée et soutenue à Villejuif, le 16 octobre 2018, par :

Célia Dupain-Jourda

Composition du Jury :

Michaël SCHUMACHER Directeur de recherche, Université Paris-Saclay (UMR 1195)	Président du jury
Xavier COUMOUL Professeur des universités, Université Paris-Descartes (UMR-S 1124)	Rapporteur
Evelyne SEGAL-BENDIRDJIAN Directrice de recherche, Université Paris-Descartes (UMR-S 1007)	Rapporteuse
Corinne DUPUY Directrice de recherche, Université Paris-Saclay (UMR 8200)	Examinatrice
Jean-Louis MERLIN Professeur des universités - Praticien hospitalier, Université de Lorraine (UMR 7039)	Examineur
Liliane MASSADE Directrice de recherche, Université Paris-Saclay (UMR 8203)	Directrice de thèse

REMERCIEMENTS

Pendant ces années passées à Gustave Roussy, j'ai eu l'occasion de rencontrer de nombreuses personnes ayant contribué au déroulement de ma thèse et je souhaite profiter de cette occasion pour les remercier.

Je voudrais tout d'abord remercier ma directrice de thèse Liliane Massade, pour m'avoir permis d'effectuer ces travaux et de mener à bien ce projet. Merci pour son encadrement, sa disponibilité et le temps consacré. Je remercie également Lluis Mir de m'avoir accueillie dans l'unité UMR 8203.

Je tiens à remercier les membres du jury pour l'évaluation de cette thèse. Merci à Evelyne Ségal-Bendirdjian et Xavier Coumoul pour le suivi tout le long de la thèse et leurs conseils scientifiques. Je remercie Jean-Louis Merlin et Corinne Dupuy d'avoir accepté de m'offrir leur expérience et leur temps pour examiner mes travaux. Merci également à Michaël Schumacher.

Mes plus profonds remerciements vont à mon équipe pour leur contribution à ce travail, pour leur aide, leur soutien indéfectible et pour la bonne ambiance quotidienne. A Céline Gracia tout d'abord, pour son aide scientifique et technique majeure dans ce projet, merci pour ta curiosité scientifique, ton efficacité hors-pair et pour les connaissances que tu m'as transmises. Merci également pour ta bonne humeur, ton soutien moral et ton aide dans mes recherches professionnelles. J'ai été très chanceuse de pouvoir te rencontrer et travailler avec toi. Merci à Giorgia Urbinati, pour sa disponibilité, les nombreuses discussions scientifiques, pour m'avoir fait partager tes connaissances mais également ton expérience et pour m'avoir toujours redonné l'envie d'aller jusqu'au bout dans les mauvais moments. Merci à Suzan Boutary pour ta présence, ton soutien et ta bonne humeur à toute épreuve. Tu resteras l'une des plus belles rencontres humainement parlant lors de ces années. Un grand merci à Marie, pour son extrême gentillesse, sa joie, les rires et le soutien que tu m'as apporté à chaque instant depuis que tu es arrivée. Tu fais partie de ces personnes qui m'ont donné l'envie de persévérer. Cette thèse n'aurait pas été la même sans vous. Merci à Hafiz M. Ali pour m'avoir transmis ses connaissances techniques et scientifiques en Master 2 et qui m'ont servies pour poursuivre mes travaux de thèse. Merci également aux étudiants de passage dans l'équipe et que j'ai eu la chance de rencontrer : Médina, Maria, Tobias, Beatrice, Soukaïna, Audrey, Théo, Jérôme, Cristina et Valentina.

Je tiens également à remercier l'équipe de pédiatrie, sans qui cette thèse n'aurait pu être réalisable. Je remercie Birgit Geoerger, de m'avoir donné l'opportunité de travailler sur ce projet, d'avoir un accès privilégié aux données d'une cohorte unique de patients, de pouvoir collaborer avec autant de personnes impliquées dans la médecine de précision à Gustave Roussy et pour m'avoir fait partager son expertise. Mille mercis à Anne Harttrampf, pour son suivi tout au long du projet, sa contribution majeure dans ces travaux de thèse et son expertise. Merci de m'avoir énormément appris sur la pédiatrie, la clinique, la bioformatique et de m'avoir permis de garder en tête pourquoi notre recherche était capitale pour les enfants et leurs familles. Merci également pour ton soutien et pour nos discussions. Merci aussi à Estelle Daudigeos-Dubus, pour sa disponibilité

à chaque instant, ses conseils scientifiques et son aide pour mes recherches professionnelles. Merci également à Jenny Marques-Costa et Noémie Assoun pour leur gentillesse et leur aide.

Cette thèse n'aurait pas été la même sans la présence de tous les membres de l'UMR 8203 que je remercie infiniment. Un grand merci à Hanna, Julia, Florian, Adeline, Sandra, Emilie, Léna pour votre présence et pour la bonne humeur. Je remercie également Isabelle Croquison pour son aide et sa disponibilité pour la gestion de la partie administrative et Jean-Rémi Bertrand le « monsieur bricolage » de l'unité pour son aide à la résolution de nombreux problèmes. J'adresse un remerciement tout particulier à Linda, l'une des plus belles rencontres de cette thèse, devenue une amie plus qu'une collègue. Merci d'avoir été là chaque jour et de m'avoir tant aidée, toi qui pouvais comprendre mieux que personne la vie en doctorat, merci pour nos discussions et tes conseils mais aussi pour nos rires et nos sorties. Merci pour ta joie de vivre, ta gentillesse et ta générosité, comme il est rarement donné de rencontrer dans une vie.

Pendant ma thèse, j'ai pu compter sur l'expertise des membres de la plateforme de bioinformatique. Merci à Yannick Boursin et Marc Deloger pour la mise en place de ChimComp et pour votre savoir-faire. Merci à Guillaume Robert-Siegwald, Manuel Lebeurrier et Windy Rondof pour l'analyse des nombreuses données et sans qui ces travaux de thèse n'auraient pu voir le jour. Merci à Guillaume Meurice pour les analyses de microarray et pour sa disponibilité.

Je tiens à remercier nos collaborateurs, ayant permis la réalisation de ce travail : Pierre Khoueiry pour son aide précieuse, son implication et son expertise apportés dans la rédaction du premier article. Catherine Richon pour les échantillons issus de biopsies de patients. Merci à Julie Rivière de nous avoir transmis son savoir-faire sur la technique de CRISPR/Cas9 et pour son aide précieuse. Merci aux membres de la plateforme d'évaluation préclinique dirigée par Mr Patrick Gonin ainsi que Olivia Bawa et Paule Opolon pour les études d'histologie des tumeurs *in vivo*. Merci également à la plateforme de cytométrie en flux et particulièrement Yann Lécluse pour le tri cellulaire.

Un énorme merci bien sûr aux patients et à leurs familles, qui acceptent d'être prélevés pour nous fournir du matériel d'étude et nous permettre de faire avancer la recherche.

Merci au comité de l'École Doctorale Biosigne de m'avoir accordé la chance de mener un doctorat de Sciences en m'attribuant une bourse.

Enfin, je terminerai par un grand merci à mes proches, dont j'ai reçu le soutien indéfectible pendant ces années.

Merci à mes partenaires du groupe compétiteurs du club de boxe française de Thiais pour les entraînements dans la bonne humeur chaque semaine, et merci à mon coach Youcef de me repousser dans mes limites et croire en moi. Merci de m'avoir apporté un équilibre mental et physique précieux pendant mon doctorat.

Merci également à mes amis du groupe de M2 : Sabrina, Javier, Sarah, Sandra pour leur présence et les soirées aux 2bis. Merci à Camille pour son soutien et ses conseils.

J'adresse un grand merci à mes amis de longue date. Merci à Maxime pour nos discussions et pour ta disponibilité, même à des milliers de kilomètres. Merci à Cécile pour ces deux premières années d'études passées ensemble et son amitié. Merci à mes amies d'enfance : Mathilde, Manon et Jo pour votre soutien indéfectible, votre présence constante et vos mots. Vous faites parties des personnes m'ayant portée jusqu'au bout. Je remercie Justine d'avoir été à mes côtés tout au long de mes études, d'être toujours disponible et de m'avoir toujours poussée vers le haut. Merci de s'être toujours souciée de ma carrière et de mon bonheur, et d'être si compréhensive tu es une amie en or.

Mes plus grands remerciements vont aux membres de ma famille. Merci à Papi et Odette, pour leur aide et leur soutien, merci à Mamette pour m'avoir transmis cet intérêt pour la science et cette curiosité. Merci à Michelle et René. Merci également à ma belle-famille pour leurs encouragements. Un grand merci à Mamie, pour les longs appels téléphoniques, les lettres et pour le soutien. Vivre loin n'est pas facile, merci de toujours prendre le temps de m'écouter et de te soucier de mon quotidien. Merci à ma sœur Elodie, de m'avoir soutenue particulièrement sur cette fin de thèse et de m'avoir encouragée. Merci de m'apprendre énormément de choses et de m'inspirer par ta volonté et ta force.

Ma plus profonde gratitude va à mes deux parents. Merci pour votre soutien financier et moral durant toutes ces années d'études et toute ma vie. Il m'aurait été impossible d'arriver jusqu'ici sans vous et sans tout ce que vous avez fait pour moi. Merci pour toutes les bonnes valeurs que vous nous avez transmises. Je vous remercie d'être des parents si compréhensifs, présents et soudés. Je vous dois tout.

Enfin je tiens à remercier la personne qui a peut-être le plus contribué à ma réussite, mon compagnon Steeven. Merci d'avoir été là chaque jour de cette thèse et depuis toutes ces années le long de mes études, de m'apporter un équilibre ainsi qu'une aide précieuse au quotidien. Tu es l'une des personnes les plus fortes mentalement que je connaisse, merci de m'avoir transmis cette force, de constamment m'élever vers le haut et d'avoir été un coach à chaque moment difficile, tu m'as énormément appris. Merci de m'avoir écoutée, d'être si compréhensif, de m'avoir toujours poussée à réaliser ce qui me tenait à cœur.

Pour conclure, merci à toutes les personnes qui ont fait de ces années une expérience professionnelle et humaine enrichissante et merci à celles et ceux qui m'ont sincèrement soutenue tout au long. Merci à tous ceux qui m'ont encouragée et aidée.

TABLE DES MATIERES

ABREVIATIONS.....	10
PREFACE.....	12
Chapitre I : Etude de la présence de gènes de fusions dans des tumeurs solides pédiatriques en rechute.....	14
I. Epidémiologie, classification et traitements actuels des cancers pédiatriques.....	15
A. Généralités sur les tumeurs pédiatriques.....	15
B. Particularités des cancers pédiatriques par rapport aux cancers adultes : causes, caractéristiques histologiques et traitements des tumeurs pédiatriques.....	18
1. Etiologie des cancers de l'enfant et de l'adolescent.....	20
a) Prédispositions génétiques et immunodéficiences.....	20
b) Facteurs environnementaux.....	21
2. Caractéristiques histologiques et moléculaires des tumeurs pédiatriques.....	23
3. Traitement des enfants en oncologie et limites des thérapies classiques actuelles.....	28
a) Types de thérapies et effets secondaires.....	28
b) Utilisation des thérapies ciblées en oncologie pédiatrique.....	32
C. Classification des différents types de tumeurs solides.....	34
1. Tumeurs cérébrales et du système nerveux central (SNC).....	35
2. Tumeurs d'origine embryonnaire.....	36
a) Neuroblastome.....	37
b) Tumeur de Wilms ou néphroblastome.....	38
c) Retinoblastome.....	38
d) Tumeurs embryonnaires rares du système nerveux central.....	39
3. Sarcomes de l'os et des tissus mous.....	40
a) Ostéosarcome.....	43
b) Sarcome d'Ewing.....	45
(1) Incidence, épidémiologie et diagnostic.....	45
(2) Caractéristiques moléculaires.....	46
(3) Traitements et taux survie.....	47
c) Rhabdomyosarcomes.....	48
4. Autres types tumoraux.....	50
II. Apport de la génétique moléculaire dans la médecine de précision utilisée en cancérologie.....	51
A. Anomalies génétiques recherchées en cancérologie et techniques d'analyses de génétique moléculaire à haut-débit.....	53
1. Principe général du séquençage haut-débit.....	53
2. Types d'anomalies moléculaires génétiques recherchées par séquençage haut-débit.....	59
3. Cas des puces d'hybridation génomique comparative (CGH array).....	62

B.	Conduite d'essais cliniques guidés par l'analyse du profil moléculaire tumoral de patients, obtenu par séquençage à haut débit.....	63
1.	Généralités sur les essais cliniques de médecine de précision en oncologie et exemples.....	63
2.	Cas de l'essai moléculaire de précision MOSCATO-01 conduit à Gustave Roussy...67	67
3.	Limites de la médecine de précision en pédiatrie.....	69
a)	Limites liées aux cohortes pédiatriques et aux essais cliniques en cours....	69
b)	Limites de l'utilisation des thérapies ciblées pour le traitement des cancers pédiatriques en Europe.....	70
c)	Evolution clonale et hétérogénéité intratumorale.....	70
d)	Limites liées au NGS.....	72
C.	Conclusion de la partie II.....	73
III.	Transcrits de fusions et cancers.....	74
A.	Généralités.....	74
B.	Méthodes de détection des transcrits de fusion.....	77
C.	Article de synthèse décrivant les gènes de fusion et les cancers pédiatriques.....	81
	RESULTAT I : Mise au point d'un outil de détection de fusions à partir de données RNA-seq et découverte de nouveaux transcrits de fusion dans une cohorte de patients pédiatriques en rechute.....	82
A.	Introduction de l'article.....	82
B.	Article 1: "Discovery of new fusion transcripts in a cohort of pediatric solid cancers at relapse and relevance for personalized medicine".....	82
	Chapitre II : Caractérisation d'un nouvel oncogène retrouvé dans le sarcome d'Ewing : LMO3-BORCS5.....	129
I.	Cas de l'oncogène de fusion LMO3-BORCS5.....	130
A.	Les protéines à domaine LIM de type LMO : description, fonction et rôle dans le développement tumoral.....	130
1.	Généralités et fonctions.....	130
2.	LMO3.....	131
3.	Autres membres de la famille LMO.....	132
B.	Structure et fonctions connues du complexe BORC.....	134
	RESULTAT II : Découverte d'un nouveau transcrit de fusion oncogénique LMO3-BORCS5 et caractérisation fonctionnelle.....	136
A.	Introduction de l'article.....	136
B.	Article 2: "Functional characterization of a new oncogenic fusion transcript LMO3-BORCS5".....	136
C.	Résultats complémentaires : interaction entre EWS-FLI et LMO3-BORCS5.....	175
1.	Matériel et méthodes.....	175
2.	Résultats.....	175
	DISCUSSION GENERALE.....	178
	CONCLUSION GENERALE.....	187

REFERENCES.....	189
ANNEXE : Résultats des travaux de Master 2.....	205
A. Effets moléculaires et cellulaires de la réexpression des facteurs de transcription thyroïdiens TTF-1 et PAX-8.....	205
B. Article “Induction of TTF-1 or PAX-8 expression on proliferation and tumorigenicity in thyroid carcinomas”.....	206

ABBREVIATIONS

aCGH	array Comparative Genomic Hybridization
ALL	Acute Lymphoblastic Leukemia
AMM	Autorisation de Mise sur le Marché
aRMS	alveolar rhabdomyosarcoma
ATRT	Atypical Tteratoid Rhabdoid Tumor
ATU	Autorisation Temporaire d'Utilisation
BIOMEDE	Biological Medicine for Diffuse Intrinsic Pontine Glioma Eradication
BORC	BLOC-1 Related Complex
BORCS5	BLOC-1 Related Complex Subunit 5
CDS	Coding sequence
COG	Children's Oncology Group
DIPG	Diffuse Instrinsic Pontine Gliomas
dNTPs	désoxyriboNucléotides TriPhosphates
EBV	Epstein Barr Virus
ECIS	European Cancer Information System
eRMS	embryonal rhabdomyosarcoma
ESMART	European Proof-of-Concept Therapeutic Stratification Trial of Molecular Anomalies in Relapsed or Refractory Tumors in Children
ETMR	Embryonal Tumors with Multilayered Rosettes
FLHCC	Fibrolamellar hepatocellular carcinoma
GIST	Gastro Intestinal Stromal Tumors
HIV	Human Immunodeficiency Virus
HPV	Human Papilloma Virus
IARC	International Agency for Research on Cancer
ICCC	International Classification of Childhood Cancers
IMR	Imagerie à Résonnance Magnétique
INCa	Institut National du Cancer
LAL	Leucémie Aigüe Lymphoblastique
LAL-T	Leucémie Aigüe Lymphoblastique T
LAM	Leucémie Aigüe Myéloïde
LDB1	LIM domain-binding protein 1
LID	LIM interaction domain
LIM-HD	LIM homedomain
LMC	Leucémie Myéloïde Chronique

LMMJ	Leucémie MyéloMonocytaire Juvénile
LMO	LIM-domain DOnly
LNH	Lymphomes Non-Hodgkiniens
MAP	Mitogen-Activated Protein
MAPPYACTS	MoleculAr Profiling for Pediatric and Young Adult Cancer Stratification
MATCH	Molecular Analysis for Therapy Choice
MOSCATO	MOlecular Screening for CAncer Treatment Optimization»
NBCCS	Nevoid Basal Cell Carcinoma Syndrome
NCI	National Cancer Institute
NEPENTHE	Next Generation Personalized Neuroblastoma Therapy
NGS	Next Generation Sequencing
NIH	National Institute of Health
pNET	primitive NeuroEctodermal Tumors
RNAPII	RNA polymerase II
RNA-seq	RNA sequencing
SIDA	Syndrome d'ImmunoDéficiency Acquis
siRNA	Small-interfering RNA
SNC	Système Nerveux Central
SNS	Système Nerveux Sympathique
SNV	Single Nucleotide Variation
T-ALL	T-Acute Lymphoblastic Leukemia
TDM	Tomodensitométrie
TEP	Tomographie par émission de positons
TK	Tyrosine Kinase
UTR	Untranslated transcribed region
WES	Whole-Exome sequencing
WGS	Whole-Genome Sequencing

PREFACE

Les gènes de fusion (aussi appelés transcrits de fusion ou ARN chimériques) présentent un grand intérêt en clinique car ils constituent des cibles thérapeutiques potentielles idéales, justifié par la découverte d'un médicament, l'imatinib mésylate, capable d'inhiber l'oncoprotéine de fusion BCR-ABL caractéristique de la leucémie myéloïde chronique, et d'induire une rémission totale à 5 ans chez plus de 90% des patients (1,2). De nombreux autres succès de thérapies ciblées appliquées dans le cadre de produits de fusion ont été reportés jusqu'à aujourd'hui, mais également l'utilisation de fusions comme outils diagnostiques dans des groupes de tumeurs extrêmement rares (3-9). Par conséquent, la découverte de nouveaux transcrits de fusion dans les cancers reçoit de plus en plus d'attention et est en augmentation, notamment grâce au développement du séquençage de l'ARN à haut débit (voir article de revue, Chapitre I, partie III-C, Dupain *et al.*, 2017). Les cancers pédiatriques comprennent les cas de cancers diagnostiqués chez les enfants et les jeunes adultes. Soixante pourcent d'entre eux sont des tumeurs solides. Alors que la majorité des patients atteignent une rémission à long terme après traitement, 20% rechutent et décèdent de la maladie en raison d'un échec du traitement, faisant des cancers pédiatriques la première cause de décès par maladie chez les enfants. Les traitements actuels sont donc à leurs limites de toxicité et d'efficacité (10). De nouveaux traitements spécifiques sont nécessaires aussi bien pour améliorer la qualité de vie des patients en rémission que pour traiter les patients résistants.

Les études de séquençage ont montré que les cancers pédiatriques possèdent moins de mutations que les cancers adultes, probablement en raison de l'absence d'effets génotoxiques exogènes, mais aussi de leur origine embryonnaire (11). En parallèle, il a été rapporté que les tumeurs associées aux gènes de fusion présentent généralement un taux de mutation plus faible que les tumeurs n'étant pas associées à des fusions (12). **Nous avons émis l'hypothèse que les tumeurs pédiatriques seraient donc plus susceptibles d'exprimer des transcrits de fusion et que leur exploration, en particulier à la rechute où peu de données sont disponibles, pourrait contribuer à comprendre la particularité de ces néoplasies et fournir de nouvelles cibles thérapeutiques.**

En effet, la plupart des études ont été réalisées sur des patients au moment du diagnostic tandis que l'analyse des tumeurs à la rechute est plus rare mais nécessaire pour enrichir le répertoire des altérations, y compris les fusions, responsables de la progression tumorale (13).

Pour cela, nous avons mené une étude rétrospective sur les données de séquençage de l'ARN de 48 patients pédiatriques en rechute avec différentes tumeurs solides, inclus dans l'étude de screening moléculaire menée à Gustave Roussy MOSCATO-01 pour «Molecular Screening for Cancer Treatment Optimization» (NCT01566019)(14,15).

Le séquençage de l'ARN est un outil puissant, largement utilisé pour découvrir de nouveaux événements de fusion. Cependant, l'augmentation spectaculaire de la détection de gènes de fusion, inclue aussi bien des évènements initiateurs responsables de l'oncogénicité (appelés « drivers »), que des évènements passagers non-impliqués dans l'initiation et la progression tumorale. De plus, la détection d'évènements non récurrents retrouvés individuellement chez les patients augmente également proportionnellement, bien que leur signification biologique soit actuellement peu définie (11). Pour répondre aux besoins cliniques et de recherche, il est donc nécessaire de disposer d'outils bioinformatiques à la fois spécifiques et sensibles pour la détection des fusions à partir du séquençage de l'ARN, mais également intégrant des prédicteurs du potentiel oncogénique des fusions.

Le premier but de mes travaux a donc été **d'établir une méthode robuste de détection de fusions *in silico*, et d'identifier de nouveaux transcrits de fusion oncogéniques dans les cancers pédiatriques en rechute (Article 1 Dupain et. *al*, en révision, Molecular Therapy).**

Parmi les fusions potentiellement oncogéniques détectées, la fusion LMO3-BORCS5 retrouvée chez un patient atteint de sarcome d'Ewing, aussi bien à la rechute qu'au diagnostic, a particulièrement retenu notre attention. Par conséquent, le deuxième objectif du projet a été de **comprendre le rôle et la fonction de LMO3-BORCS5 dans le développement et/ou la résistance des tumeurs (Article 2 Dupain et. *al*, soumis).**

Dans ce manuscrit, j'exposerai les connaissances actuelles sur le sujet, les résultats obtenus seront intégrés sous forme d'articles suivis d'une conclusion qui ouvrira les perspectives.

Chapitre I

Etude de la présence de gènes de fusions dans des tumeurs solides pédiatriques en rechute

I. Epidémiologie, classification et traitements actuels des cancers pédiatriques

Les cas de cancers surviennent très rarement avant l'âge de 20 ans, et lorsque c'est le cas, ils soulèvent un éventail de problèmes médicaux, éthiques et sociétaux. Ces types distincts de néoplasmes nécessitent donc des types de traitement spécifiques (16).

L'oncologie pédiatrique est une discipline médicale relativement récente, considérée comme une entité différente des tumeurs adultes dans les années 1960. L'existence de l'oncologie pédiatrique est donc fondée sur la compréhension des différences entre les tumeurs des enfants et des adultes. Au regard de son existence récente, des progrès spectaculaires et significatifs concernant le traitement des patients ont été réalisés sur un temps très court en oncologie pédiatrique et il est maintenant observé plus de 85% de rémission chez les enfants et adolescents atteints de cancer (17). Néanmoins, l'efficacité des traitements actuels ainsi que leurs fortes toxicités limitent les progrès et les taux de survie stagnent. Dans cette première partie, j'exposerai les différents types de cancers pédiatriques solides, qui correspondent aux différents diagnostics des patients inclus dans la cohorte étudiée dans mon projet de thèse. Ces cancers représentent des pathologies rares et peu connues, très différents de l'adulte sur de nombreux aspects qui seront évoqués dans cette partie. L'épidémiologie des cancers pédiatriques, les connaissances actuelles et les défis auxquels sont confrontés la recherche et la clinique au regard de ces pathologies seront également explicités ci-dessous.

A. Généralités sur les tumeurs pédiatriques

Les cancers dits « pédiatriques » regroupent les cas de cancers diagnostiqués chez les enfants entre 0 et 14 ans et chez les adolescents, entre 15 et 19 ans. Chaque année en France, 2500 nouveaux cas environ sont diagnostiqués chez les enfants (1 750 cas) et les adolescents (800 cas). Bien que considérés comme maladies rares puisqu'ils représentent environ 1% des cancers totaux, le cancer est sur le point de dépasser les maladies infectieuses en étant l'une des causes les plus élevées au monde de mortalité liée à la maladie chez les enfants de plus de 1 an (18,19). Dans les pays économiquement développés, le cancer pédiatrique représente d'ailleurs la maladie la plus fréquente entraînant la mort chez les moins de 15 ans, malgré le fait que la plupart des tumeurs pédiatriques soient associées à des taux de survie élevés (17). Selon l'Agence Internationale pour la Recherche contre le Cancer (IACR) 3,317 décès dû à des cancers pédiatriques ont été reportés en Europe en 2012 et environ 80 000 cas dans le monde (20).

De plus, une étude internationale également coordonnée par l'IACR a montré que leur incidence mondiale est en forte hausse avec une augmentation de la fréquence des cancers pédiatriques de

13% en 2001-2010, comparativement aux années 1980, atteignant un taux d'incidence annuel de 140 par million enfants âgés de 0-14 ans dans le monde entier (16). L'augmentation de cette incidence au cours du temps est également visible au niveau européen, pour les deux sexes (**Figure I A-1**) (21).

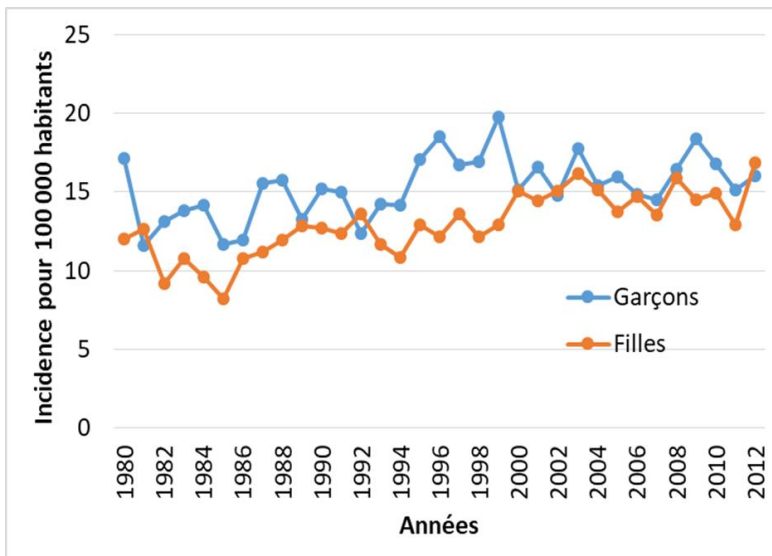


Figure I A-1 : Evolution de l'incidence des cancers pédiatriques en Europe entre 1980 et 2012. Source: ECIS - European Cancer Information System. A partir de: <https://ecis.jrc.ec.europa.eu>.

Une partie de cette augmentation pourrait être attribuable à une détection meilleure ou plus précoce de ces cancers et à l'amélioration de l'accès aux soins médicaux (22) et beaucoup de cas restent encore non comptabilisés et non déclarés en raison d'un manque de registres des cancers pédiatriques dans un grand nombre de pays (18).

Sur l'ensemble des cancers pédiatriques diagnostiqués entre 0 et 19 ans dans le monde, les leucémies concernent le plus grand nombre de cas (26%), suivies des tumeurs cérébrales et du système nerveux central (18%) puis des lymphomes (14%). Néanmoins, il est important de noter que les types de cancers pédiatriques les plus souvent diagnostiqués varient selon les tranches d'âge. Ainsi, les leucémies constituent le cancer le plus fréquent chez les enfants de moins de 15 ans, représentant près d'un tiers des cas de cancer infantile. Les tumeurs du système nerveux central arrivent en deuxième position (21% des cas) suivies des lymphomes (12%). Chez les enfants de moins de 5 ans, un tiers des cas concernent les tumeurs embryonnaires, telles que le neuroblastome, le rétinoblastome, le néphroblastome ou l'hépatoblastome. Chez les adolescents, les cancers les plus fréquents étaient les lymphomes (23%), suivis des tumeurs de type épithéliales et des mélanomes (21%), puis des leucémies (15,4%) (16). (**Figure I A-2**).

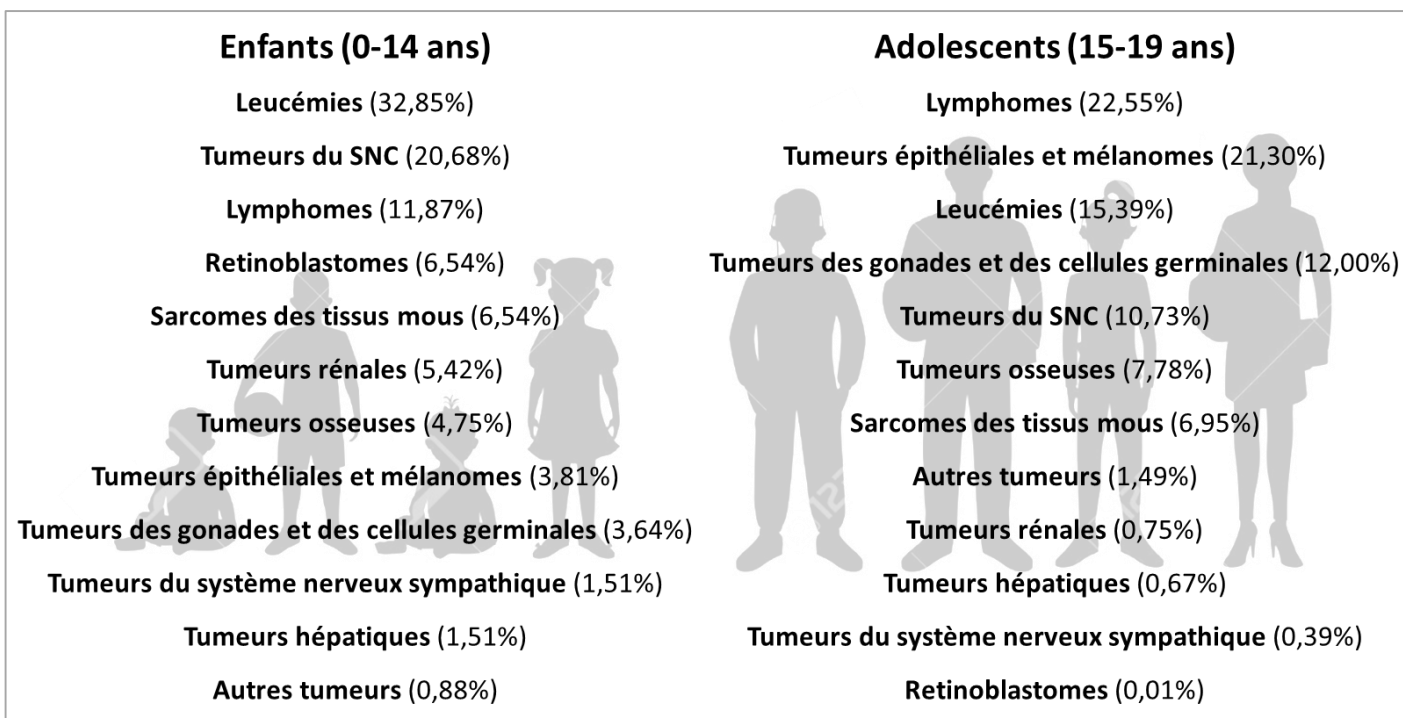


Figure I A-2: Répartition de l'incidence des différents types cancers pédiatriques selon l'âge et dans le monde. Adapté d'après les résultats publiés par l'IARC (International Agency for Research on Cancer), par Steliarova-Foucher et. *al* (16). SNC = Système nerveux central.

En France et selon les registres nationaux recueillis par l'Institut National du Cancer (INCa) entre 2007 et 2011, cette répartition diffère légèrement chez les adolescents. Ainsi, les lymphomes restent en tête des diagnostics les plus fréquents, suivis des tumeurs du système nerveux central en deuxième position devant les leucémies et les tumeurs de type épithéliales et mélanomes (19). De manière intéressante, la moitié des cancers de l'enfant survient avant l'âge de 5 ans (23). Des différences entre les deux sexes sont également observées : chez les garçons, la majorité des diagnostics concernent des lymphomes, des médulloblastomes et des sarcomes alors que chez les filles, les tumeurs des gonades et des cellules germinales ainsi que les tumeurs de type épithéliales (mélanomes, carcinomes thyroïdiens) sont plus fréquentes. (23).

Les hémopathies malignes regroupent les leucémies et les lymphomes et constituent la grande majorité des cas de cancers pédiatriques (34%) et sont classées selon l'organisation mondiale de la santé en fonction du type cellulaire affecté.

Les leucémies englobent les hémopathies malignes issues de progéniteurs des cellules hématopoïétiques de la moelle osseuse, de type lymphoïde ou myéloïde. Les lymphomes quant à eux concernent des hémopathies malignes prenant naissance dans les organes lymphoïdes secondaires à partir des cellules lymphocytaires.

Les leucémies comprennent 80% de leucémies aiguës lymphoblastiques (LAL), 15% de leucémies

aiguës myéloïdes (LAM) et moins de 5% de leucémies myéloïdes chroniques (LMC). Parmi les LAM se trouvent différents sous-types hétérogènes, mais les avancées technologiques de ces dernières décennies dans les techniques d'analyses moléculaires ont permis d'améliorer leur classification, avec l'exemple de la translocation t(8; 21) ou des mutations du gène FLT3 présents dans certains groupes de patients. Les LMC sont caractérisées par la translocation t(9; 22), le chromosome de Philadelphie donnant lieu au gène et à la protéine de fusion BCR-ABL. La LMC comprend également un sous-type spécifique trouvé uniquement chez les jeunes enfants, la leucémie myélomonocitaire juvénile (LMMJ) qui est négative pour BCR-ABL. D'autres types de LMC sont également présents et identifiés et comprennent les leucémies myélomonocytaires, monocytaires et lymphocytaires chroniques.

Les lymphomes représentent environ 10% des cas de cancers pédiatriques et sont de deux types : non-hodgkinien (62% des lymphomes de l'enfant) et Hodgkinien (38% des cas). Les lymphomes non-hodgkiniens (LNH) regroupent le lymphome de Burkitt (environ 40% des LNH soit 2 à 3% de tous les cancers infantiles), le lymphome lymphoblastique et le lymphome à grandes cellules (représentant ensemble 30% des LNH infantiles totaux) (24).

Bien que les leucémies et les lymphomes représentent la majorité des cancers pédiatriques, mon travail de thèse s'est porté uniquement sur l'analyse des tumeurs pédiatriques « solides ». Ainsi, je ne les mentionnerai pas dans les parties suivantes, en revanche je décrirai de manière plus complète les tumeurs de type « solide » que je détaillerai dans le **paragraphe C**.

B. Particularités des cancers pédiatriques par rapport aux cancers adultes : causes, caractéristiques histologiques et traitements des tumeurs pédiatriques

L'incidence des cancers est plus élevée à l'âge adulte que chez les enfants et les adolescents, avec en tête les carcinomes, notamment les carcinomes du poumon, colorectaux, du sein et de la prostate. En revanche, les cancers pédiatriques ne concernent qu'environ un enfant sur 300 et les tumeurs prédominantes sont les leucémies et lymphomes, les tumeurs cérébrales et les sarcomes, alors que les carcinomes sont relativement rares (seulement 6% environ des diagnostics de cancers pédiatriques (25).

Il est communément considéré que les cancers à l'âge adulte résultent d'un processus multi-étapes impliquant une succession d'évènements mutationnels (**Figure I B-1**), souvent provoqués par l'exposition à des agents toxiques tels que le tabagisme, l'alcool, les infections virales et d'autres facteurs environnementaux. En revanche, les cancers pédiatriques ne peuvent pas être justifiés par l'exposition chronique à de tels facteurs favorisant l'accumulation de mutations au cours du temps et par conséquent le développement de tumeurs. Ils sont alors plus susceptibles de résulter d'évènements génétiques moins nombreux mais aux conséquences plus dramatiques,

liés à un syndrome génétique héréditaire ou non (25). Ainsi, l'étude des cancers pédiatriques est particulièrement importante puisque l'identification d'altérations génétiques à l'origine du développement tumoral peut fournir des indications sur les mécanismes tumorigènes et sur la biologie des cancers, comme ce fut par exemple le cas pour le rétinoblastome entre autres (26,27).

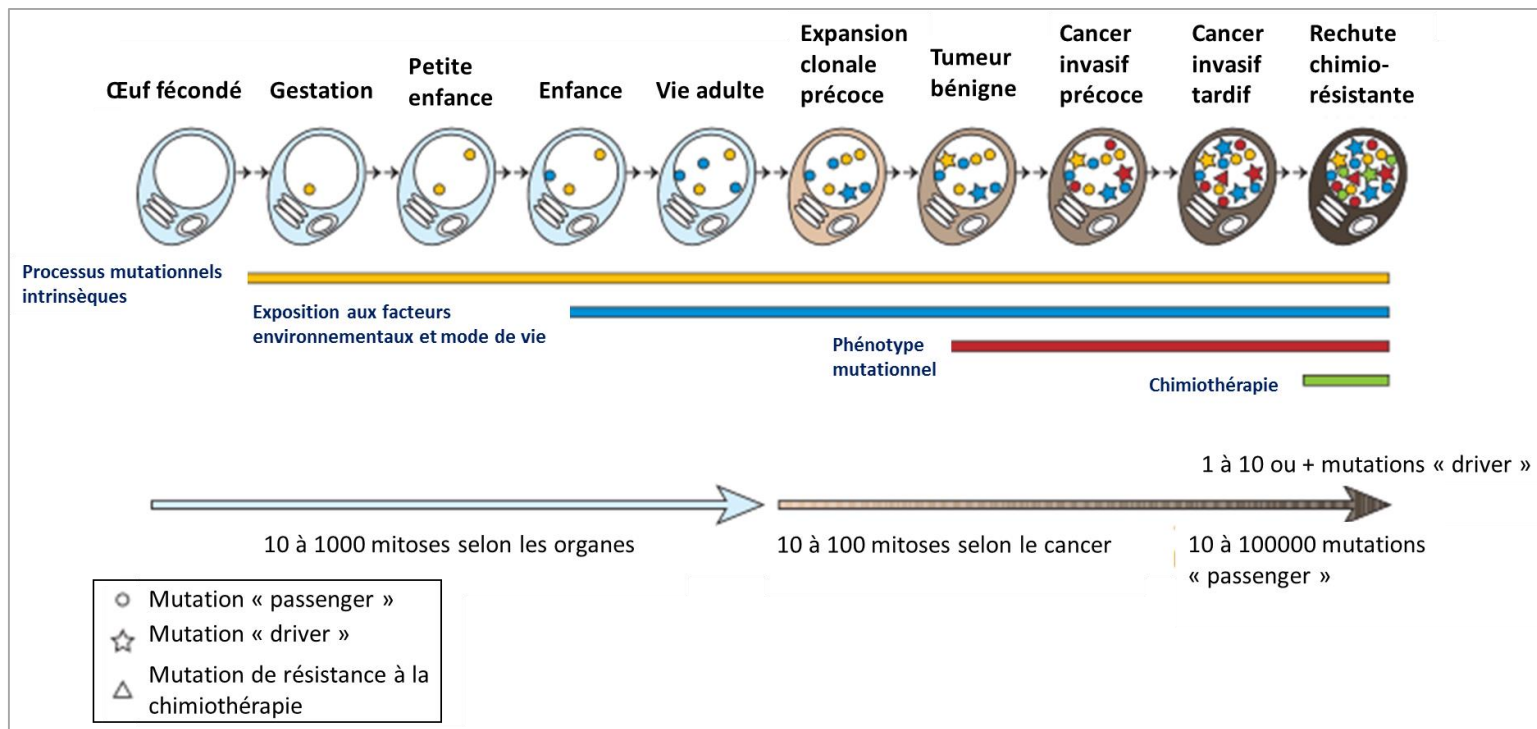


Figure I B-1 : Mutations somatiques acquises au cours des divisions cellulaires mitotiques, de l'œuf fécondé jusqu'à une cellule unique cancéreuse et les processus qui y contribuent. Adapté d'après Stratton et. al (7).

Mutation « passager » = mutation n'ayant aucun effet sur la cellule cancéreuse.

Mutation « driver » = mutation à l'origine de l'expansion clonale tumorale.

Au vu de ces observations, il apparaît donc logique que les cancers pédiatriques soient considérés comme une entité à part entière et différente des cancers de l'adulte, ayant des caractéristiques moléculaires et histologiques propres. De ce fait et comme il a été mentionné dans les sous-parties précédentes, la classification des cancers de l'enfant est basée non pas sur le site anatomique de la tumeur primaire comme chez l'adulte, mais sur le type histologique. Certains types histologiques sont notamment retrouvés exclusivement chez les enfants et adolescents, en particulier les types embryonnaires détaillés dans la **partie I.B.2**. L'identification de sous-groupes histologiques ayant un impact direct sur le pronostic représente l'un des plus gros défis en oncologie pédiatrique (22,28-30).

Enfin, au regard des traitements, les schémas appliqués chez l'adulte sont souvent transposés chez les patients pédiatriques à tort, considérant ainsi les patients pédiatriques comme « adultes

miniatures ». Les différents organes et tissus de ces patients contrairement aux adultes sont encore en phase de croissance et d'évolution ce qui limite considérablement le choix des traitements anticancéreux, pouvant avoir des effets à court et long termes importants et dans certains cas irréversibles. L'aspect de prise en considération des impacts sur le futur des patients est une caractéristique impactant les décisions thérapeutiques en oncologie pédiatrique et constitue l'une des raisons principales à l'impossibilité de transposer des traitements adultes directement chez l'enfant. D'un autre côté, l'état de croissance présent dans les organismes de ces patients présente l'avantage d'avoir des processus de régénération nettement plus rapides et ainsi les enfants et adolescents traités tolèrent beaucoup mieux les chimiothérapies intensives (17). L'application de chimiothérapies à hautes doses a d'ailleurs contribué en majeure partie à l'amélioration considérable des taux de survie en oncologie pédiatrique depuis les années 1980 (24).

1. Etiologie des cancers de l'enfant et de l'adolescent

Contrairement aux cancers chez l'adulte, seule une faible proportion de cancers de l'enfant ont des causes génétiques ou environnementales connues et dans 80-90% des cas ils surviennent de manière sporadique. L'hétérogénéité de ces tumeurs aux niveaux biologiques et cliniques ainsi que leur très faible incidence rend l'identification de causes spécifiques aux cancers pédiatriques très délicate (22). Cependant, quelques facteurs de risque ont été identifiés et regroupent :

- les facteurs génétiques et maladies d'immunodéficience
- les facteurs environnementaux

a) *Prédispositions génétiques et immunodéficiences*

Environ 8 à 10% des cancers pédiatriques sont associés à des mutations germinales conduisant à une prédisposition cancéreuse (31). Différents troubles génétiques peuvent être associés à un risque accru de développement de cancers pédiatriques et peuvent être de différents types (32):

- Des syndromes génétiques
- Des prédispositions génétiques autosomales dominantes
- Des troubles à transmission autosomale récessive
- Des insuffisances médullaires
- Des immunodéficiences

Les différentes prédispositions génétiques ou syndromes identifiés, ainsi que les gènes et les cancers associés connus sont répertoriés dans le **tableau I B-1 a. 1**. Environ 75 syndromes héréditaires de prédisposition associés aux cancers pédiatriques ont été décrits, la plupart d'entre eux étant considérés comme rares et avec une transmission autosomique dominante. La connaissance des facteurs génétiques de prédisposition est importante non seulement pour l'enfant mais également pour les membres de la famille apparentés. L'étude de ces altérations contribue également à mieux isoler et comprendre les voies de signalisation responsables de l'initiation et/ou de la progression tumorale (17,24).

	Prédisposition génétique ou syndrome	Gènes affectés connus	Cancer(s) associé(s)
Maladies génétiques	Syndrome familial de Wilms	WT1/2	Tumeur de Wilms, néphroblastome
	Syndrome de Down	Trisomie du chromosome 21	LAM, LAL
	Syndrome de Turner	Rétention du chromosome Y	Gonadoblastome
	Syndrome de Beckwith-Wiedemann	Principalement IGF2	Hépatoblastome, tumeur de Wilms
	Syndrome de Bloom	BLM	Nombreux cancers
Prédispositions génétiques cancéreuses autosomales dominantes	Li-Fraumeni	TP53/CHK2/SNF5	Nombreux cancers
	Rétinoblastome familial	RB1	Rétinoblastome
	Mutations du gène APC	APC	Cancer colorectal, hépatoblastome
	Cancer colorectal héréditaire non polyposique	MSH2/MLH1/PMS2	Cancer colorectal
	Syndrome de néoplasies endocrines multiples	MEN2	Thyroïde, parathyroïde, pancréas, glande pituitaire et glande surrénale
	Neurofibromatose de type 1	NF1	Tumeurs du système nerveux central et périphérique
	Neurofibromatose de type 2	NF2	Shwannome, méningiomes, astrocytomes
	Syndrome de Cowden	PTEN	Multiple visceral cancers
	Syndrome de Gorlin (NBCCS)	PTCH	Nombreux cancers des cellules basales cutanées
	Sclérose tubéreuse	TSC1/2	Tumeurs bénignes
	Syndrome de Von Hippel-Lindau	VHL	Carcinome rénal, hémangioblastomes rétiens et cérébelleux, phéochromocytome
	Polypose adénomateuse familiale	APC	Cancer colorectal
	Troubles à transmission autosomale récessive	Xeroderma pigmentosum	ERCC2
Ataxia telangectasia		ATM	Tumeurs lymphoïdes
Insuffisances médullaires	Anémie de Fanconi	Gènes du complexe FANC	Nombreux cancers
Immunodéficiences	Déficiences IgA	IGAD1	Cancer gastro-intestinal
	Syndrome de Wiskott-Aldrich	WAS	Lymphomes non-Hodgkiniens
	Déficit immunitaire combiné sévère (DICS)		Tumeurs lymphoïdes
	Infections par HIV ou EBV	-	Nombreux cancers

Tableau I B-1 a. 1: Syndromes et prédispositions génétiques à l'origine de certains cancers pédiatriques. Adapté d'après <https://emedicine.medscape.com/article/989841-overview/> et Dupain et. al (9). NBCCS = Nevoid Basal Cell Carcinoma Syndrome; HIV = Human Immunodeficiency Virus; EBV = Epstein Barr Virus; LAM = Leucémie Aigüe Myéloïde ; LAL = Leucémie Aigüe Lymphoïde.

b) Facteurs environnementaux

L'implication de facteurs environnementaux dans l'étiologie des cancers de l'enfant est encore discutée, cependant elle reste mineure comparativement aux cas adultes. En effet des facteurs tels que le surpoids, l'alcool, la sédentarité et le tabac par exemple ne sont pas transposables aux cas

de cancers pédiatriques. Certains facteurs environnementaux sont tout de même associés à certains types de tumeurs comme les rayonnements ionisants ou certains produits chimiques mais de manière générale et jusqu'à présent, les causes de la plupart des cancers pédiatriques demeurent principalement inconnues (9,24,32).

L'effet des facteurs environnementaux externes est également corrélé à la durée de l'exposition et n'explique pas le développement de tumeurs chez les nourrissons et les enfants de moins de 5 ans. Les principaux facteurs associés qui ont été décrits jusqu'ici sont (17,22,24):

- L'exposition aux radiations : l'utilisation de la radiothérapie pour traiter une première tumeur a été associée à un risque plus élevé de développer des cancers secondaires (tumeurs osseuses, thyroïdiennes et cérébrales). Les données recueillies provenant des conséquences des expositions à la bombe atomique à Hiroshima et Nagasaki constituent sans doute les évidences les plus solides. Un lien a également été établi entre les examens radiologiques du troisième trimestre de grossesse et certaines leucémies.
- L'exposition à certains produits chimiques : une forte relation a été suggérée entre l'exposition des parents et le cancer de l'enfant associé. Les agents responsables du développement de tumeurs comprennent les pesticides (tumeurs du SNC), les solvants (tumeurs du SNC, leucémies, neuroblastome, hépatoblastome), les métaux (hépatoblastome), les produits dérivés du pétrole (tumeur de Wilms, leucémies, hépatoblastome), le plomb (tumeur de Wilms), le bore (tumeur de Wilms) et également certaines chimiothérapies (leucémies). La dioxine a également été associée au cancer de la thyroïde, à la leucémie myéloïde aiguë (LAM) et au lymphome de Hodgkin et enfin le trichloroéthane a été impliqué dans un cas de leucémie.
- L'infection par des virus:
 - Hépatite C associée au carcinome hépatocellulaire
 - Virus d'Epstein Barr associé au lymphome de Burkitt en Afrique, au lymphome de Hodgkin et au carcinome nasopharyngé
 - Papillomavirus humain (HPV pour « Human Papilloma Virus ») de type 16 associé aux cancers du col de l'utérus, de la vulve, du vagin, de l'anus, à des cas de papillomatose et aux carcinomes oro- et naso-pharyngés.
- La consommation de tabac, une obésité et d'autres facteurs associés à la mère.

Le faible nombre de patients atteints de cancers pédiatriques et donc inclus dans des cohortes d'étude a rendu l'évaluation de l'implication des facteurs environnementaux difficiles. Cependant, l'analyse des facteurs héréditaires est de plus en plus fructueuse, compte tenu de l'explosion des avancées technologiques de séquençage du génome et de la disponibilité des bases de données génétiques.

2. Caractéristiques histologiques et moléculaires des tumeurs pédiatriques

Les cancers pédiatriques diffèrent fondamentalement des cancers de l'adulte, non seulement au niveau de leur étiologie mais également au niveau biologique (histologique et moléculaire), même au sein d'une même pathologie.

L'une des premières différences frappantes évoquée précédemment, est la faible occurrence de carcinomes comparés aux adultes, avec une majorité de leucémies, reflétant une caractéristique spécifique de la maladie. Les principales tumeurs cérébrales et autres tumeurs solides qui surviennent chez les enfants comme le médulloblastome, le neuroblastome, le rhabdomyosarcome, le sarcome d'Ewing, l'ostéosarcome ou la tumeur de Wilms, sont extrêmement rares chez les adultes (**Figure I B-2**) (33).

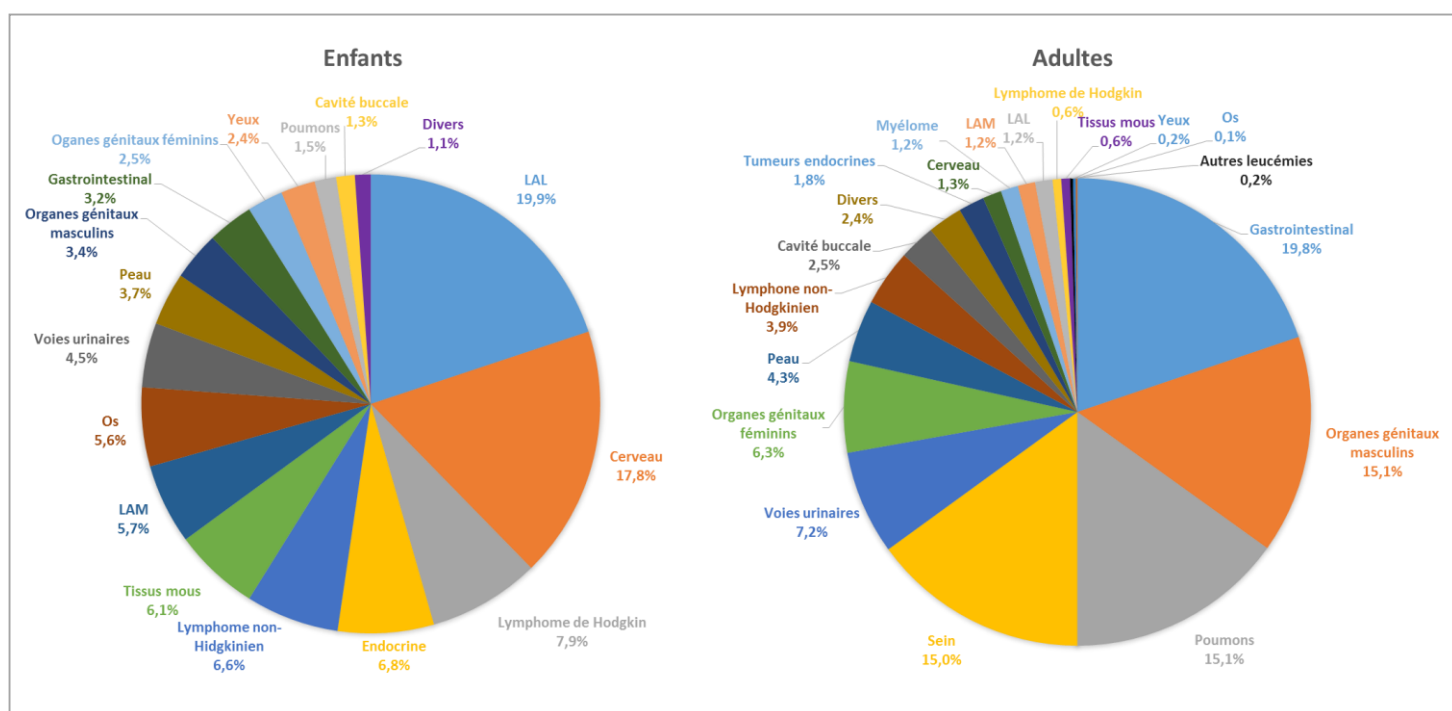


Figure I B-2 1: Fréquence des différents cancers diagnostiqués chez les enfants et les adultes. LAM = leucémie aigüe myéloïde ; LAL = leucémie aigüe lymphoblastique. Adapté d'après Downing et. al (33).

Il a été décrit que de nombreux cancers pédiatriques se développeraient à partir des tissus en développement qui subissent une expansion importante au cours de la formation et de la maturation précoces des organes, constituant une autre caractéristique propre aux cancers pédiatriques (partie I.C.2.) (33,34).

Ainsi les tumeurs solides, excepté les tumeurs cérébrales et du système nerveux central, peuvent survenir à partir des trois lignées germinales: mésoderme, endoderme ou ectoderme (**Figure I-B-2 2**) ayant pour conséquence une difficulté à identifier les cellules d'origine des tumeurs pédiatriques et une diversité histologique remarquable (35). Par exemple, les cellules tumorales

du rétinoblastome s'organisent en sphères contenant des vésicules synaptiques, forment des jonctions synaptiques et sécrètent des neurotransmetteurs alors que les cellules tumorales de rhabdomyosarcome ont une morphologie organisée en myofibrille avec présence abondante de glycogène, rappelant les propriétés du muscle squelettique (35).

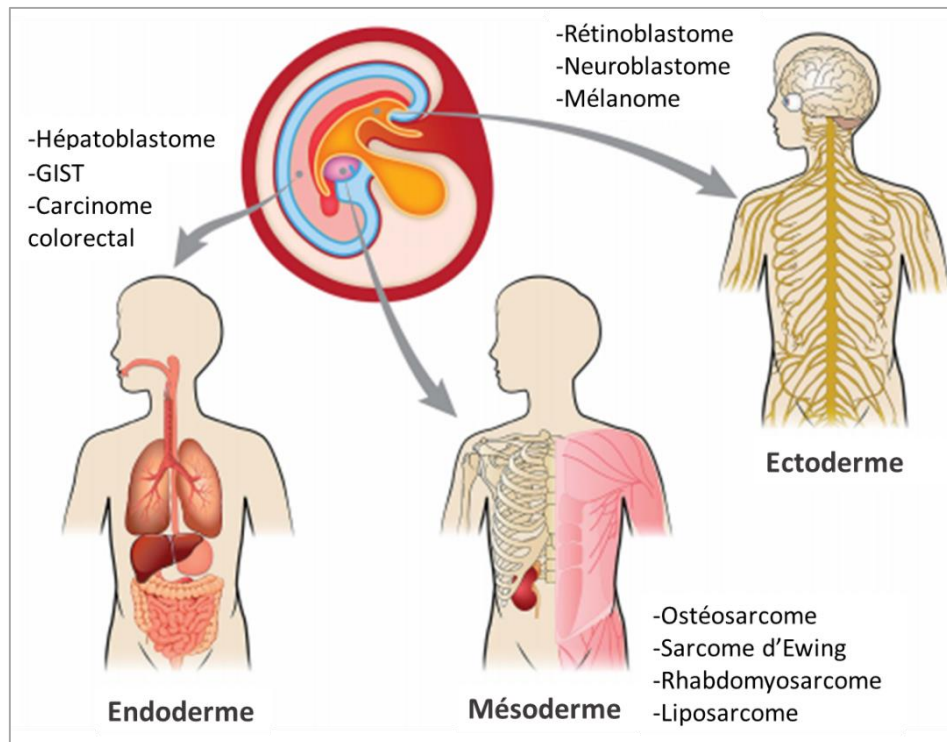


Figure I B-2 2: Origines développementales diverses des tumeurs pédiatriques solides. En dehors des tumeurs cérébrales, les tumeurs solides peuvent émerger à partir des trois feuillets embryonnaires, comme c'est le cas pour les différents exemples cités ci-dessus. D'après Chen et al, (35). GIST = Tumeurs stromales gastro-intestinales, dérivé de l'anglais «Gastro Intestinal Stromal Tumors».

En plus de leur diversité, les différents types histologiques sont généralement spécifiques à une certaine tranche d'âge, comme il a été évoqué dans la partie I.A (Figure 1). La **Figure I B-2 3** illustre ces différences de manière plus détaillée, avec des tranches d'âge beaucoup plus définies et ne se restreignant pas à deux groupes (enfant versus adolescents). Ainsi ces tranches se découpent de la manière suivante :

- tumeurs embryonnaires, caractéristiques de la petite enfance 0 – 5 ans.
- les sarcomes sont généralement propres aux adolescents, qu'ils proviennent des os ou des tissus mous.
- les carcinomes très rares en pédiatrie, représentent 2% des cancers chez les enfants et 12% chez les adolescents.

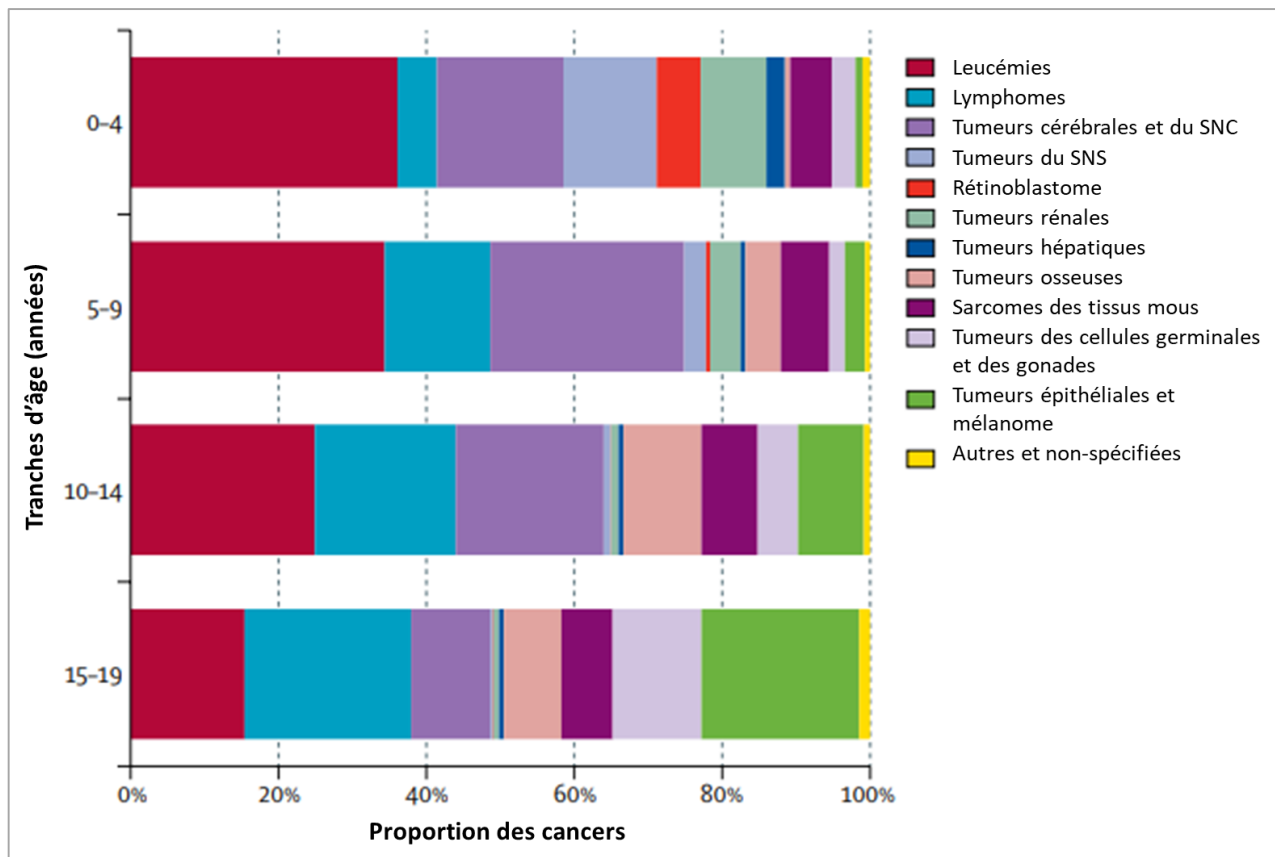


Figure I B-2 3: Distribution proportionnelle mondiale des différents types de cancers en fonction des tranches d'âge et selon la classification ICC-3. Source : Steliarova-Foucher et. *al* (16). SNC = système nerveux central ; SNS= système nerveux sympathique.

Les cancers de l'enfant et de l'adolescent sont également caractérisés par un comportement biologique particulier. Comme décrit précédemment, ces tumeurs sont généralement hautement agressives, avec un indice de prolifération cellulaire élevé (temps de doublement se comptant en heures et jours plutôt qu'en semaines et en mois comme il est généralement le cas chez l'adulte) et accompagnées d'une dissémination métastatique précoce. De plus, le stade de tumeur *in situ* n'est pas reconnaissable dans les tumeurs pédiatriques contrairement aux adultes et par conséquent, de manière globale les tumeurs pédiatriques sont considérées comme une pathologie aiguë menaçant particulièrement le pronostic vital des patients (17).

Au niveau moléculaire, les cancers pédiatriques sont caractérisés par un nombre moindre de mutations par rapport aux tumeurs adultes, comme l'ont souligné plusieurs études (71,72). En comparant le nombre de mutations non synonymes par tumeur, il apparaît que les tumeurs avec des défauts des systèmes de réparation de l'ADN (cas des tumeurs colorectales) et les cancers induits par des mutagènes (cancer du poumon et mélanome) sont les tumeurs adultes ayant le plus grand nombre de mutations. En revanche, les tumeurs ayant le plus petit nombre de

mutations restent les leucémies et les cancers pédiatriques avec une moyenne de 9,6 mutations/tumeur (variant de 0,02 mutation par mégabase pour l'hépatoblastome à 0,49 dans le lymphome de Burkitt ; **Figure I B-2 4**).

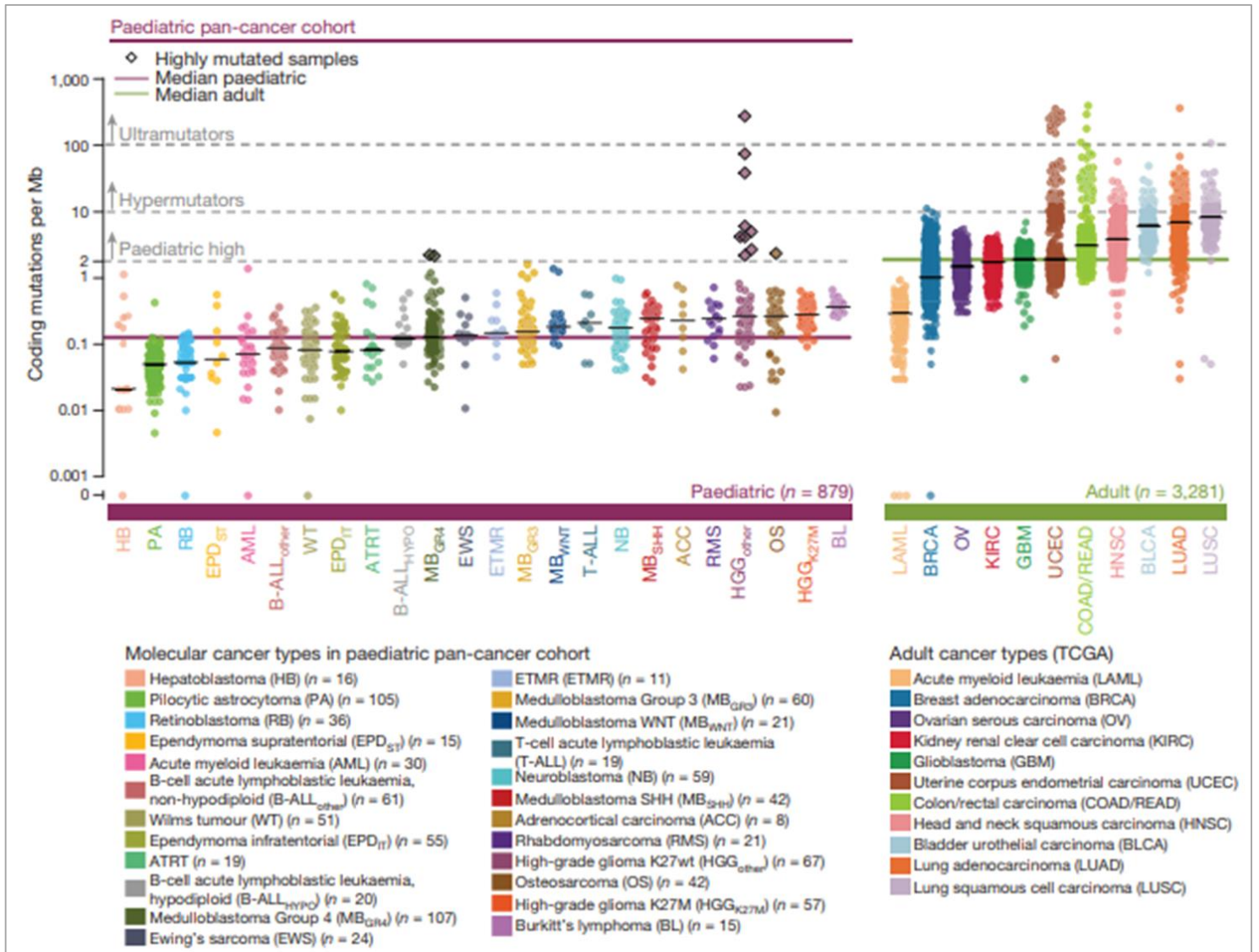


Figure I B-2 4 : Taux de mutations non-synonymes par mégabase (Mb) dans les cancers pédiatriques et adultes. Fréquence des mutations somatiques situées dans des zones codantes, parmi 24 types de cancers pédiatriques (n = 879 tumeurs primaires) et 11 adultes (n = 3 281 tumeurs). Les échantillons hypermutés et hautement mutés sont séparés par des lignes grises en pointillé et représentés par des losanges noirs. La médiane des charges mutationnelles de mutation médiane sont indiquées par des lignes continues (noir = types de cancer ; violet = cancers pédiatriques ; vert = cancers adultes). Source : Gröbner et. al, (36).

De plus, plusieurs études ont montré que le profil de mutations au sein d'une même pathologie était différent selon leur apparition chez les enfants ou les adultes. Quelques exemples peuvent

être cités comme les cas de leucémie aigüe myéloïde et les gliomes de haut grade qui ont des profils différents chez les enfants et les adultes (72, 73). En effet, chez les enfants, certains gliomes apparaissent au niveau du tronc cérébral sous forme de gliomes infiltrants du tronc cérébral (DIPGs pour « Diffuse Intrinsic Pontine Gliomas » en anglais). Les DIPGs présentent des mutations uniques dans des gènes codant pour 2 des 16 isoformes de l'histone H3 (37). Ces mutations ne sont détectées que dans les DIPGs et, à une fréquence plus faible, dans les glioblastomes pédiatriques survenant en dehors du tronc cérébral, mais ne sont détectées dans aucun autre type tumoral, que ce soit chez les enfants et adolescents ou chez les adultes (33).

Ces différences dans les profils et les charges mutationnels pourraient s'expliquer en partie par les étiologies différentes des tumeurs pédiatriques et adultes mais aussi par le fait que les cancers pédiatriques proviennent de progéniteurs cellulaires. En effet, ces cellules n'ont pas encore eu le temps, contrairement à chez l'adulte, de se renouveler et d'accumuler des erreurs et par conséquent présentent un faible nombre de mutations (11).

Gröbner et. *al* ont montré que *TP53* était le gène le plus fréquemment muté (4% des cas), cependant ce taux reste dix fois moins important que dans les cancers adultes. De manière intéressante, l'ensemble des gènes les plus fréquemment mutés concernent les gènes liés au contrôle des modifications épigénétiques (25% des cas) suivi des régulateurs transcriptionnels et des gènes associés aux kinases MAP (« Mitogen-Activated Protein ») (12-15%) (36). En effet, l'épigénétique joue un rôle majeur dans l'initiation et la progression des tumeurs pédiatriques et est devenue un domaine majeur de recherche dans les tumeurs solides pédiatriques, en raison de nombreux exemples comme le changement épigénétique observé dans les rétinoblastomes après inactivation de *RB1*, ou encore l'identification de mutations récurrentes dans des régulateurs épigénétiques tels que *ASXL1/2*, *BCOR*, *ATRX*, etc. (35,38).

Enfin, au niveau moléculaire, les cancers pédiatriques sont également caractérisés par la récurrence d'altérations (mutations et altérations du nombre de copies) de gènes impliqués dans le développement du tissu d'origine, rejoignant la particularité histologique de certaines tumeurs à se développer à partir de tissus embryonnaires (39-41).

En conclusion de cette sous partie I.C.2, nous pouvons résumer les principales différences histologiques et moléculaires des cancers pédiatriques par rapport aux tumeurs adultes selon les points suivants :

CARACTERISTIQUES HISTOLOGIQUES	<ul style="list-style-type: none"> ▪ Types tumoraux différents de l'adulte, très peu de carcinomes parmi les cancers pédiatriques ▪ Grande diversité histologique ▪ Origine embryonnaire ▪ Types de cancers spécifiques à certaines tranches d'âge ▪ Tumeurs particulièrement agressives à fort indice de prolifération
CARACTERISTIQUES MOLECULAIRES	<ul style="list-style-type: none"> ▪ Peu de mutations ▪ Différents profils mutationnels au sein d'une même pathologie entre les cancers pédiatriques et adultes ▪ Rôle majeur de l'épigénétique dans l'initiation et la progression des cancers pédiatriques ▪ Récurrence d'altérations génétiques des gènes liés au développement dans les tumeurs pédiatriques

3. Traitement des enfants en oncologie et limites des thérapies classiques actuelles

a) *Types de thérapies et effets secondaires*

Malgré le manque de nouveaux médicaments pour le traitement des cancers de l'enfant au cours des 20 dernières années, les taux de survie n'ont cessé de s'améliorer pour atteindre, dans les pays développés, environ 80% (22,33). Le **tableau I B-3 a. 1** détaille les taux de survie des différents types de cancers pédiatriques. Ce succès repose sur l'utilisation de chimiothérapies multimodales et à hautes doses combinées à de la radiothérapie. En effet, de par leur caractère hautement agressif, ces tumeurs sont généralement particulièrement chimio- et radiosensibles et ont donc une bonne réponse aux thérapies (22,33).

Bien que les taux de guérison des cancers pédiatriques soient parmi les plus impressionnants par rapport à ceux des tumeurs adultes, le cancer reste la première cause de mortalité par maladie chez les enfants de plus de 1 an dans les pays développés.

	1975-79 (%)	2003-2009* (%)
Tous types confondus	63	83
Leucémies	48	84
LAL	57	90
LAM	21	64
Lymphomes et néoplasmes réticuloendothéliaux	72	91
Lymphome de Hodgkin	87	97
Lymphome non-Hodgkinien	47	85
Tumeurs cérébrales et du SNC	59	75
Ependymome	37	81
Astrocytome	69	85
Médulloblastome	47	70
Neuroblastome et ganglioneuroblastome	54	79
Rétinoblastome	92	99
Tumeur de Wilms	75	90
Tumeurs hépatiques	25	74
Tumeurs osseuses	49	73
Ostéosarcome	45	71
Sarcome d'Ewing	42	72
Rhabdomyosarcome	49	64
Tumeurs des cellules germinales du testicule	74	96
Tumeurs des cellules germinales de l'ovaire	75	94
Carcinomes thyroïdiens	99	98
Mélanome	83	95

Tableau I B-3 a. 1 : Taux de survie observés chez les patients atteints de cancer entre 0-19 ans, entre 1975-1979 comparativement aux taux observés entre 2003-2009, selon les différents types de cancers. LAL=leucémie aigüe lymphoblastique ; LAM=leucémie aigüe myéloïde. * certains cas ont été suivis jusqu'en 2010. D'après la société américaine du cancer, 2014 (22).

Les traitements classiques principaux utilisés en oncologie pédiatriques consistent en (17):

- La chirurgie : elle occupe une place encore très importante dans la prise en charge des tumeurs pédiatriques et pour certaines tumeurs elle représente un traitement curatif. La chirurgie ne permet qu'un contrôle local de la tumeur. Lorsque la résection tumorale radicale n'est pas possible, cela signifie généralement un risque plus élevé de récurrence locale ou de rechute.
- Diverses chimiothérapies à hautes doses : La chimiothérapie est la thérapie systémique la plus efficace contre les cancers pédiatriques. Contrairement aux cellules saines, les cellules cancéreuses ont la faculté de se développer continuellement sans réponse aux signaux d'inhibition de la croissance cellulaire. La chimiothérapie se base sur l'utilisation de molécules ciblant les cellules avec un fort index de prolifération, par un effet cytostatique et cytotoxique. Dans les protocoles de traitement en oncologie pédiatrique, différentes molécules sont généralement utilisées en combinaison (au moins deux médicaments anticancéreux), permettant de réduire les risques de résistances. Une chimiothérapie adjuvante est destinée à augmenter ou stimuler les effets positifs d'une autre forme de traitement, comme la chirurgie ou la radiothérapie. Une chimiothérapie néo-adjuvante est un prétraitement administré afin de réduire la taille d'une tumeur cancéreuse avant une intervention chirurgicale (et donc limiter l'exérèse) ou une

radiothérapie (augmenter l'efficacité).

▪ La radiothérapie : La radiothérapie représente un autre type de thérapie locale. La radiothérapie englobe l'utilisation de rayonnements à haute énergie provenant des rayons X, des rayons gamma, des neutrons, des protons et d'autres sources afin de tuer les cellules cancéreuses. Ces radiations provoquent des dommages à l'ADN dans les cellules cancéreuses exposées. Les lésions deviennent, à doses élevées, trop nombreuses pour être réparées par les cellules avant le cycle de réplication suivant, et conduisent donc à l'arrêt des divisions cellulaires et à l'apoptose des cellules tumorales. En oncologie pédiatrique, les indications pour la radiothérapie ont été considérablement réduites au cours des années en raison du risque important de complications à long-terme. Les facteurs limitants à l'utilisation de la radiothérapie chez les enfants comprennent : la radiosensibilité tumorale, la radiorésistance des organes environnants, la taille du champ de rayonnement et la dose cumulée utilisée, l'âge de l'enfant, la combinaison avec la chimiothérapie (certains agents sont utilisés comme « radio-sensibilisateurs »).

Dans la plupart des cas, ces différents traitements se font en association (exemple : radiochimiothérapie simultanée préopératoire). Les limites majeures de l'utilisation de ces thérapies classiques restent leur toxicité, avec des risques de complication et des effets indésirables associés (**Tableau I B-3 a. 2**). L'apparition des différents effets secondaires dépend de la localisation tumorale, des doses reçues, des traitements associés et de l'état de santé globale du patient.

Cependant ces effets secondaires sont dits à « court-terme » et sont à distinguer cliniquement des effets à « long-terme », qui eux ont tendance à être irréversibles et apparaissent plusieurs mois à plusieurs années après le traitement. Ces effets sont variés mais peuvent être cités :

- Des altérations de la croissance et du développement
- Des altérations de la fonction de certains organes
- Diverses maladies chroniques (42)
- L'apparition de cancers secondaires, avec environ 8% de risques de développer un cancer secondaire dans les vingt années après le diagnostic initial (24,43,44)
- Des accidents vasculaires-cérébraux
- Des risques accrus d'obésité
- Des troubles liés à la reproduction : infertilité, anomalies congénitales, ménopause précoce etc.
- Des troubles psychologiques : troubles du comportement, détresse psychologique, etc.

Type	Local/Systémique	Effets secondaires court-terme
Chirurgie	Local	Douleurs, fatigue, perte d'appétit, gonflement, engourdissement, saignement, infection, lymphœdème, etc.
Chimiothérapies (Agents alkylants; antimitotiques; antimétabolites; inhibiteurs de topoisomérases)	Systémique	Fatigue Toxicités hématologiques: destruction des CSH, neutropénie, thrombopénie, anémie, lymphopénie, etc. Nausées, vomissements Mucite, stomatite Troubles du transit Toxicités rénales Toxicités vésicales: cystites hémorragiques Toxicités dermatologiques: alopecie, hyperpigmentation, syndrome main/pied, extravasation
Radiothérapie	Local	Toxicités dermatologiques Fatigue Dépression Selon la localisation tumorale: - Tête et cou: sécheresse buccale, difficulté à la déglutition, nausées, lymphœdème, caries etc. - Poitrine: difficulté à la déglutition, essoufflement, malformation du sein, pneumonie, fibrose. - Estomac et abdomen: nausées, vomissements, diarrhées. - Bassin: diarrhées, saignements, incontinence, irritation de la vessie, problèmes liés à la sexualité et à la fertilité

Tableau I B-3 a. 2 : Différents types de traitements classiques utilisés en oncologie pédiatrique et effets secondaires indésirables à court-terme.

Différents facteurs de risques sont associés à ces différentes conditions et dépendent des traitements reçus, de l'âge au diagnostic ou encore de l'état de santé général du patient (45). Les effets secondaires à court et long-terme limitent l'utilisation de doses trop importantes et peuvent irrémédiablement affecter la qualité de vie à long terme des patients et représentent donc un aspect extrêmement important dans la prise de décision thérapeutique, plus que dans les cas adultes. Cependant, étant donné que l'incidence des cancers pédiatriques est en augmentation ces dernières années et que le taux de survie global s'améliore, il y aura à l'avenir de plus en plus d'adultes ayant été traités pour un cancer dans leur enfance. La mise au point de nouvelles thérapies afin de limiter les toxicités et la qualité de vie des patients à l'âge adulte représente donc l'un des défis futurs les plus importants en oncologie pédiatrique (24,46-48).

b) Utilisation des thérapies ciblées en oncologie pédiatrique

Le but dans le futur, en oncologie pédiatrique et aussi adulte, est donc d'introduire de nouvelles thérapies moléculaires ciblées qui permettent i) une amélioration de la survie globale des patients, ii) une réduction de la morbidité et de la mortalité liées à la maladie et/ou aux traitements classiques et iii) une réduction des effets secondaires à long-terme des thérapies anticancéreuses. Les thérapies ciblées ont pour but la destruction des cellules cancéreuses de manière plus sélective, en épargnant les cellules et les tissus sains, provoquant par conséquent moins d'effets secondaires chez l'hôte par rapport à la chimiothérapie cytotoxique traditionnelle (49).

Les thérapies ciblées font partie d'une « médecine de précision », un terme faisant référence à une médecine qui repose sur des traitements développés sur les bases d'une meilleure connaissance des mécanismes biologiques conduisant à l'apparition et au développement des tumeurs (50). En effet, leur mécanisme passe par une modulation des cibles moléculaires impliquées dans la cancérogenèse et la croissance tumorale.

Ces thérapies peuvent être divisées en deux groupes : les inhibiteurs intracellulaires (petites molécules chimiques) et les inhibiteurs extracellulaires (biomédicaments) décrits dans la **Figure I B-3 b. 1**. Il n'existe actuellement pas de consensus sur la définition des thérapies ciblées. Certaines classes d'agents anticancéreux partageant des caractéristiques communes avec les thérapies ciblées, ne sont pas officiellement classées comme telles (50). Il s'agit de:

- L'hormonothérapie : qui a pour but de bloquer l'action pro-oncogénique des stéroïdes dans les cancers hormonaux-dépendants. À ce titre, elles peuvent être rapprochées de la classe des inhibiteurs de mécanismes oncogéniques.
- L'immunothérapie.

En 2015 en France et pour le traitement du cancer, 41 thérapies ciblées disposaient d'une AMM (Autorisation de Mise sur le Marché) et 2 d'une ATU (Autorisation Temporaire d'Utilisation) de cohorte. Parmi ces 43 thérapies, 34 appartiennent à la classe des inhibiteurs intracellulaires et 9 à celle des inhibiteurs extracellulaires (50). Néanmoins, la majorité de ces médicaments sont recommandés uniquement pour le traitement des cancers chez l'adulte. De plus, les médicaments qui possèdent une autorisation pour le traitement de cancers pédiatriques ont tous en premier lieu été développés chez l'adulte (51).

Ces thérapies ciblées représentent donc le futur de l'oncologie pédiatrique, où il est nécessaire d'éviter les toxicités sur le long-terme et d'augmenter l'efficacité des traitements pour permettre aux taux de rémission, actuellement à un plateau, d'augmenter. Le besoin d'identification de nouvelles cibles thérapeutiques, passant par une meilleure connaissance des mécanismes

biologiques d'oncogénicité est en pleine expansion. Cet aspect est rendu en partie possible au cours de la dernière décennie, grâce à :

- L'émergence et l'application clinique des techniques de séquençage à haut débit, générant des bases de données moléculaires de patients à grande échelle.
- L'étude des tumeurs pédiatriques à la rechute afin de déchiffrer les mécanismes responsables des résistances observées chez certains patients et pouvoir mettre en évidence des cibles thérapeutiques pouvant être neutralisées avant l'apparition de rechute.

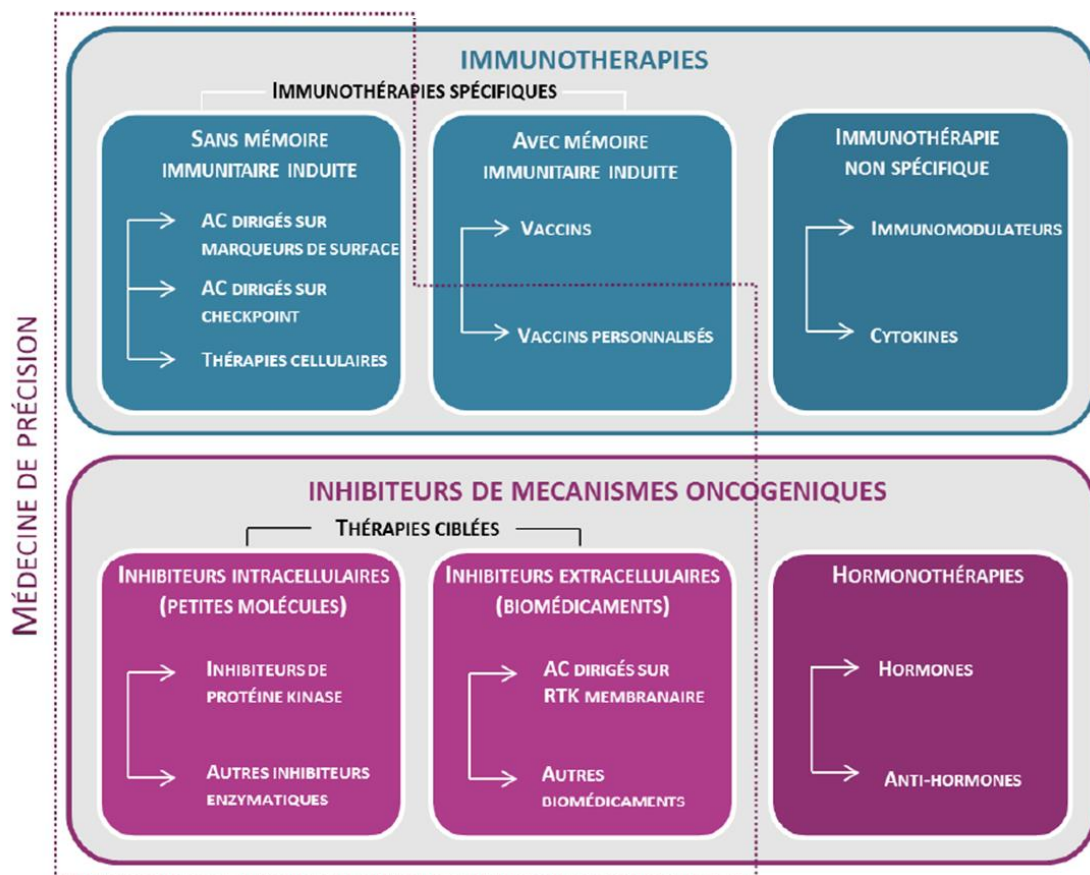


Figure I B-3 b. 1 : Classification des thérapies dites « ciblées » et leur utilisation en médecine de précision. Source: « Les thérapies ciblées dans le traitement du cancer en 2015 - État des lieux et enjeux », Institut National du Cancer (INCa) <http://www.e-cancer.fr/>.

Pour résumer, les cancers de l'enfant sont en plusieurs points différents des cancers de l'adulte (**Figure I B-3 b. 2**). Les cancers pédiatriques possèdent donc une classification, une prise en charge thérapeutique et une recherche fondamentale et pré-clinique spécifiques et distinctes de chez l'adulte.

Ces pathologies nécessitent une attention particulière afin de proposer de nouvelles alternatives thérapeutiques et mieux comprendre l'étiologie de ces cancers mais peuvent également, comme cela a été le cas pour d'autres maladies rares (26), être à l'origine d'importantes découvertes des mécanismes d'oncogenèse encore méconnus.

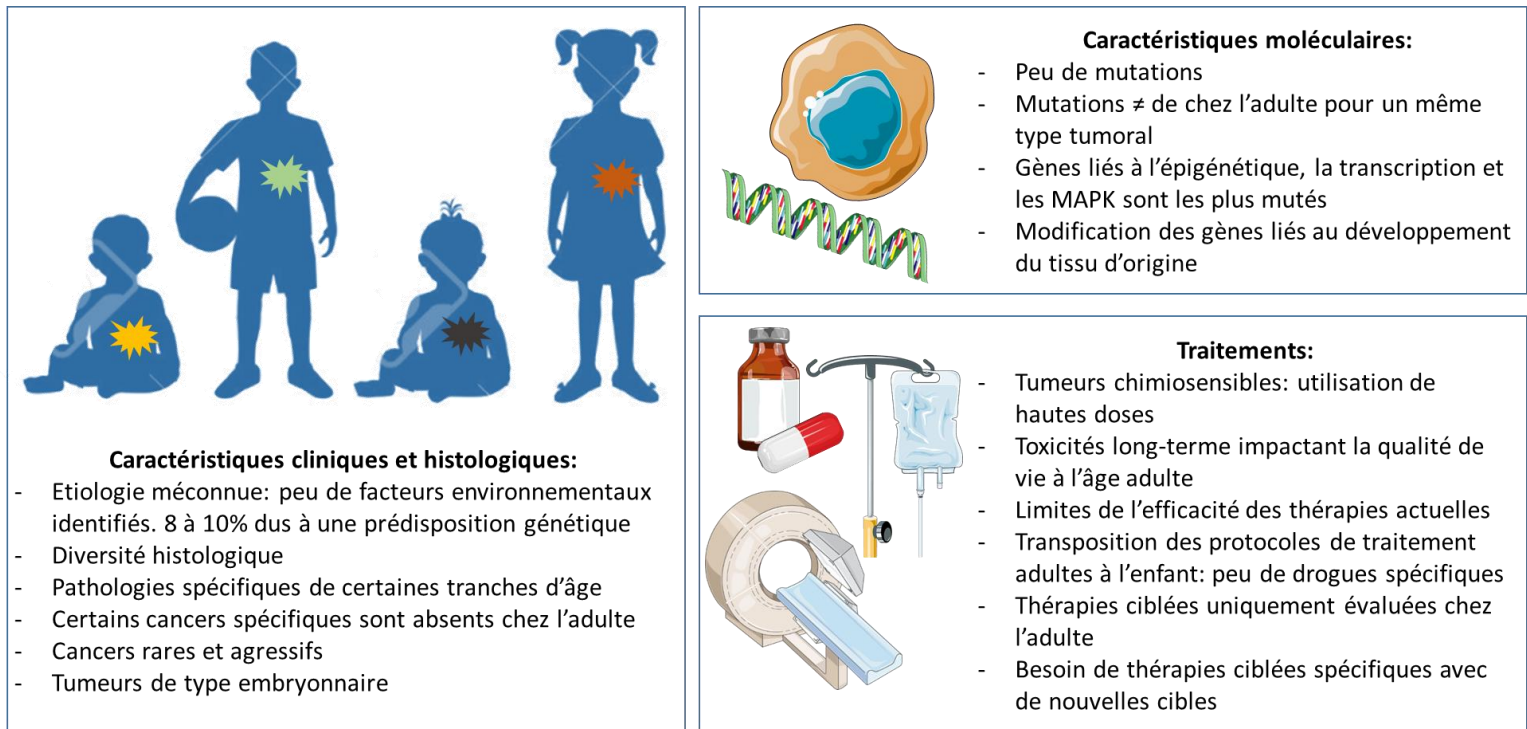


Figure I B-3 b. 2: Caractéristiques cliniques, histologiques, moléculaires et thérapeutiques spécifiques aux cancers pédiatriques.

C. Classification des différents types de tumeurs solides

Les cancers chez les enfants et les adolescents sont classés par type histologique selon 12 groupes principaux établis par la classification internationale des cancers infantiles (ICCC) (52). Certains des cancers qui se développent chez les enfants sont rarement vus chez les personnes plus âgées, notamment les cancers de type « embryonnaires » qui surviennent à partir de cellules embryonnaires et proviennent de tissus et d'organes en développement (*e.g.* neuroblastome, medulloblastome, rhabdomyosarcome ou encore rétinoblastome). Certains cancers pédiatriques, en particulier ceux qui sont les plus fréquents chez les adolescents, sont à l'inverse plus semblables à ceux rencontrés chez l'adulte (*e.g.* leucémie myéloïde aiguë, lymphome de Hodgkin, carcinome de la thyroïde et mélanome).

Afin de faciliter la description des tumeurs de type « solide », celles-ci seront classées en 4 groupes histologiques: les tumeurs cérébrales et du système nerveux central (SNC), les tumeurs d'origine embryonnaire, les sarcomes de l'os et des tissus mous et les autres types de tumeurs (22).

1. Tumeurs cérébrales et du système nerveux central (SNC)

Les tumeurs cérébrales et du système nerveux central représentent la deuxième catégorie de cancers pédiatriques majoritairement diagnostiqués et la première catégorie de cancers pédiatriques solides. L'introduction de l'imagerie à résonance magnétique (IRM) et l'utilisation accrue de biopsies assistées par ordinateur ont notamment contribué à une forte augmentation de l'incidence globale au milieu des années 1980. Avec environ 21% des cas chez les enfants et 11% chez les adolescents, cette catégorie très hétérogène regroupe une grande variété de diagnostics différents (**Tableau I C-1 1**), classés selon : les types cellulaires et tissulaires affectés, la localisation (**Figure I C-1 1**) et le grade, allant de I (faible) à IV (élevé) (22). Les astrocytomes, les médulloblastomes et les ependymomes constituent les entités les plus fréquentes avec environ 35%, 10-20% et 10% des cas respectivement.

Tumeurs astrocytaires		Tumeurs oligodendrocytaires		Ependymomes	
Astrocytome pilocytique	(I)	Oligodendrogliome	(II)	Sub-ependymome	(I)
Astrocytome pilomyxoïde	(II)	Oligodendrogliome anaplasique	(III)	Ependymome myxopapillaire	(I)
Astrocytome diffus	(II)			Ependymome	(II)
Xanthroastrocytome pléiomorphe	(II)			Ependymome anaplasique	(III)
Astrocytome anaplasique	(III)				
Glioblastome	(IV)				
Tumeurs du plexus choroïde		Tumeurs neuronales et neuro-gliales		Tumeurs de la région pinéale	
Papillome du plexus choroïde	(I)	Gangliogliome	(I)	Pinéocytome	(I)
Papillome atypique du plexus choroïde	(II)	Gangliogliome anaplasique	(III)	Pinéoblastome	(IV)
Carcinome du plexus	(III)	Tumeur dysembryoplasique neuroépithéliale	(I)	Tumeur papillaire de la région pinéale	(II/III)
		Neurocytome	(II)		
Tumeurs embryonnaires		Tumeurs de la région sellaïre		Tumeurs des nerfs périphériques	
Medulloblastome	(IV)	Craniopharyngiome	(I)	Schwannome	(I)
Tumeur primitive neuroectodermique supratentorielle	(IV)			Neurofibrome	(I)
Tumeur atypique tératoïde/rhabdoïde	(IV)			Tumeurs malignes de la gaine des nerfs périphériques	(II-IV)
Tumeurs des méninges		Tumeurs germinales			
Méningiome	(I)	Tératome			
Méningiome atypique	(II)	Choriocarcinome			
Méningiome anaplasique	(III)	Germinome			
Hémangiopéricytome	(II)	Carcinome embryonnaire			
Hémangiopéricytome anaplasique	(III)	Tumeur du sac vitellin			
Hémangioblastome	(I)				

Tableau I C-1 1 : Classification des tumeurs du système nerveux central retrouvées chez les enfants et les adolescents et leurs grades (de I à IV). Adapté de Frühwald et. al (53), d'après la classification de l'Organisation Mondiale de la Santé.

La morbidité est l'une des plus élevées dans ce groupe, principalement en raison de la localisation des tumeurs qui rend la chirurgie et les traitements difficiles (22).

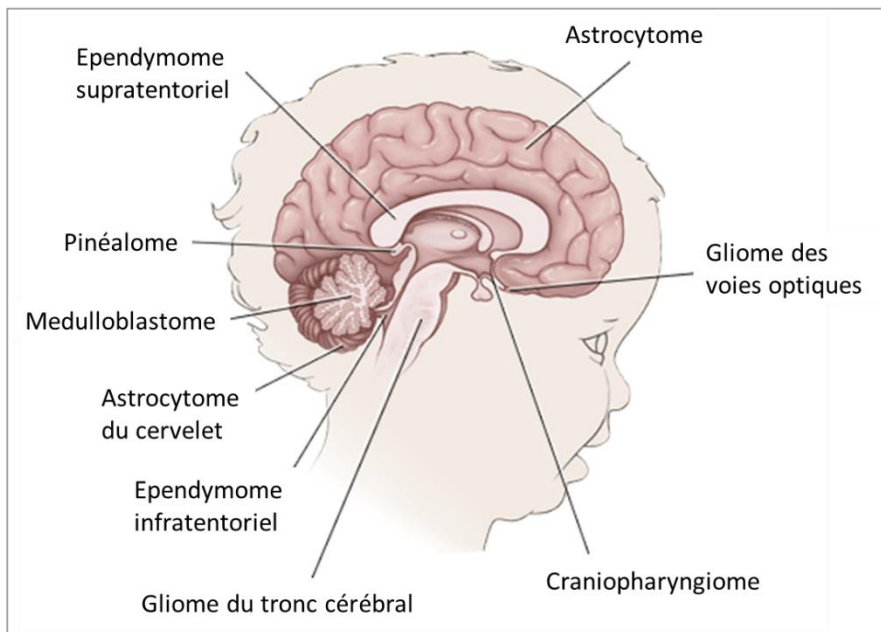


Figure I C-1 1 : Localisation des principaux types de tumeurs du système nerveux central retrouvées chez les enfants et les adolescents. Source : <https://brain-surgery.com/brain-tumors-in-children-be-aware/>

Le traitement appliqué dépend du type de cancer, du grade, de l'emplacement de la tumeur, de sa taille et d'autres facteurs pronostiques. Quand elle est possible, une intervention chirurgicale est effectuée pour retirer la tumeur tout en évitant d'endommager les tissus sains et celle-ci est généralement complétée par une chimiothérapie et / ou une radiothérapie (24).

La plupart des tumeurs du SNC sont des tumeurs gliales classées selon leur emplacement (supratentoriel, cérébelleux ou du tronc cérébral). Les astrocytomes supratentoriels représentent de 30 à 40% des cas, tandis que le reste des tumeurs gliales englobent les astrocytomes cérébelleux et les gliomes du tronc cérébral (15% chacun). La mise en évidence de sous-groupes dans les différentes catégories de cette classification a permis d'apporter de forts outils pronostiques comme c'est le cas pour les gliomes exophytiques qui ont un excellent pronostic tandis que les tumeurs infiltrantes diffuses sont de mauvais pronostic (22).

2. Tumeurs d'origine embryonnaire

Les tumeurs dites embryonnaires regroupent des tumeurs se développant à partir des cellules fœtales embryonnaires. Ces tumeurs sont généralement diagnostiquées chez les enfants avant l'âge de 5 ans et sont donc exclusivement spécifiques aux cancers pédiatriques (22). Les différents types de tumeurs embryonnaires incluent: le neuroblastome, les tumeurs de Wilms, le rétinoblastome et d'autres types de tumeurs telles que les tumeurs embryonnaires avec rosettes multicouche (abrégées « ETMRs », selon l'anglais « embryonal tumors with multilayered

rosettes »), les médulloépithéliomes et les tumeurs tératoïdes/rhabdoïdes atypiques (abrégées « ATRT », selon l'anglais « atypical teratoid rhabdoid tumor »). Le médulloblastome fait également partie des tumeurs ayant une origine embryonnaire, néanmoins il est également classé dans la catégorie des tumeurs du système nerveux central et explicité dans le paragraphe précédent.

a) Neuroblastome

Le neuroblastome est une tumeur embryonnaire du système nerveux sympathique (SNS), représentant la tumeur extra-crânienne solide la plus courante chez les enfants (entre 7 et 10% des cancers pédiatriques) et responsable de 15% de tous les décès de cancers pédiatriques. En France, 130 à 150 nouveaux cas sont observés par an, avec 90% des cas diagnostiqués avant l'âge de 5 ans (54).

Dérivant de la lignée sympathique surrenalienne de la crête neurale, le site primitif peut se localiser aussi bien dans l'abdomen, au niveau du sympathique rétro-péritonéal ou de la surrenale (70 % des cas), que dans toute autre localisation anatomique normale du système sympathique (cou, thorax ou pelvis). La propagation de la maladie se faisant vers les ganglions, la moelle osseuse et le squelette, la recherche d'une atteinte à distance du site initial est capitale au moment du diagnostic (55). Les neuroblastomes sont classés en différents stades allant de 1 à 4 et 4s. Les stades 1, 2 et 3 correspondent aux formes localisées alors que le stade 4 correspond aux formes métastatiques. Les pronostics associés à des différents groupes sont variables et dépendent de plusieurs paramètres (28):

- l'âge du patient au diagnostic : Moroz et. *al* ont montré qu'au-delà de 18 mois le risque de rechute était plus élevé (56)
- la ploïdie cellulaire
- la présence d'une amplification du gène MYCN, associée à un mauvais pronostic (57)
- la présence de délétions des chromosomes 1p et/ou 11q, associées à un mauvais pronostic (58)
- la présence de mutations activatrices du gène ALK, signe de mauvais pronostic (59).

Sur la base de ces pronostics, les patients sont catégorisés en 3 groupes de risque prédictif de rechute : faible, intermédiaire et à haut risque, chaque groupe ayant des significations thérapeutiques et pronostiques propres. La probabilité de prolongation de la survie sans maladie pour les patients de chaque groupe est de 95% à 100% pour le groupe à faible risque, de 85% à 90% pour le groupe intermédiaire et de moins de 30% pour le groupe à haut risque (28,59).

Selon le stade et les facteurs pronostiques, les enfants atteints de neuroblastome sont le plus souvent traités par chirurgie et/ou chimiothérapies multimodales; les patients présentant une maladie à haut risque peuvent recevoir une chimiothérapie et une radiothérapie à haute dose suivie d'une greffe de cellules souches (22).

b) Tumeur de Wilms ou nephroblastome

Aussi appelé néphroblastome, la tumeur de Wilms est une tumeur rénale embryonnaire qui survient habituellement chez les enfants de moins de 5 ans. C'est la tumeur rénale pédiatrique la plus fréquente (92% des tumeurs rénales dans ce groupe et 5-6% des cancers infantiles totaux). Les tumeurs de Wilms peuvent également être diagnostiquées chez les adolescents et les adultes de manière beaucoup plus rare (moins de 1% des tumeurs rénales), faisant de cette entité une particularité d'oncologie pédiatrique (60).

Environ 10% des cas sont associés à des malformations congénitales, telles que des anomalies du tractus urogénital. La tumeur de Wilms est également fortement associée à de nombreux syndromes génétiques (voir section I. C. 1.). Au cours de la vie fœtale, le futur rein se développe via l'interaction de deux structures fœtales : le bourgeon urétéral (structure épithéliale qui formera les tubules rénaux) et le blastème métanéphrogène (tissu mésenchymateux donnant le néphron). Dans 1% des cas, le blastème peut persister à la naissance et alors dériver vers une transformation tumorale en tumeur de Wilms (60).

Comme dans le cas du neuroblastome, les tumeurs diagnostiquées dans la petite enfance sont de meilleur pronostic. Les autres facteurs pronostiques associés à la survie à long terme comprennent un stade peu avancé et une histologie favorable (22,24). Dans la majorité des cas, le traitement consiste en une combinaison de chirurgie et de chimiothérapie, avec en plus une radiothérapie pour les patients à haut risque. Le néphroblastome ou tumeur de Wilms constitue l'un des plus grands succès en oncologie pédiatrique, avec un taux de survie proche de 90% en absence de métastases et de 75% en présence de celles-ci (60).

c) Retinoblastome

Avec une incidence globale d'environ 2%, le rétinoblastome est une tumeur solide de la rétine relativement rare, qui se développe à partir des progéniteurs rétinien de l'œil en développement. Ce type de tumeur survient habituellement chez les enfants de moins de 5 ans et représente 6% des cancers dans ce groupe. Les rétinoblastomes peuvent être unilatéraux (un seul œil touché) ou bilatéraux (les deux yeux sont affectés) et 6% sont familiaux contre 94% sont sporadiques.

Tous les cas de rétinoblastomes bilatéraux sont dus à des mutations germinales du gène suppresseur de tumeur *RB1* contre 15% des rétinoblastomes unilatéraux sporadiques. En 1971, Knudson a proposé l'hypothèse de deux « hits », suggérant que deux mutations chromosomiques du gène *RB1* étaient nécessaires pour le développement du rétinoblastome (61). Dans le cas des rétinoblastomes héréditaires, la mutation initiale du gène *RB1* est de type germinale tandis que la

deuxième mutation allélique est somatique et apparaît dans les cellules rétiniennes au cours du développement. Par conséquent, les cas héréditaires sont prédisposés au développement de tumeurs monoculaires (62).

La survie des patients atteints de rétinoblastome s'est considérablement améliorée avec un taux de survie à cinq ans de plus de 90%, cependant le taux de mortalité reste élevé pour cette pathologie (50%) (62). Le type de traitement appliqué dépend en grande partie de l'étendue de la maladie et de l'état de propagation aux tissus adjacents, tout en prenant en compte la balance efficacité de traitement/préservation de la vue. Les petites tumeurs peuvent parfois être traitées par cryothérapie, photocoagulation au laser ou thérapie thermique. Les patients présentant une maladie unilatérale plus avancée, sans invasion des tissus voisins, sont souvent traités par chirurgie (énucléation), qui peut être le seul traitement nécessaire. Les enfants atteints de rétinoblastome bilatéral et certains cas de rétinoblastome unilatéral peuvent être traités par chimiothérapie pour arriver à une taille tumorale réduite où le traitement local est efficace. Les patients atteints d'une maladie plus avancée sont traités par chimiothérapie et parfois par chirurgie et/ou irradiation (22,63).

d) Tumeurs embryonnaires rares du système nerveux central

Ce groupe de tumeur appartient également aux tumeurs du SNC (**section C.1. Tableau 2**), mais il est important de les mentionner dans cette partie étant donnée leur origine embryonnaire claire, caractère spécifique aux cancers pédiatriques. On y retrouve les tumeurs embryonnaires à rosettes pluristratifiées (ETMR) et les tumeurs térétoïdes rhabdoïdes atypiques (ATRT). Au niveau histologique, ces tumeurs sont caractérisées par la présence de rosettes, c'est-à-dire une structure multicouche constituée de cellules tumorales autour d'un lumen.

L'ETMR appartient au groupe des tumeurs neuroectodermiques primitives du système nerveux central (PNET) et survient typiquement chez les nourrissons et les jeunes enfants (en particulier chez les moins de 2 ans), dans la partie supratentorielle du cerveau dans 66% des cas. Ces tumeurs sont particulièrement agressives et associées à un mauvais pronostic mais extrêmement rares avec seulement une centaine de cas reportés dans la littérature, bien que la prévalence soit probablement sous-estimée étant donnée la difficulté à les diagnostiquer (29,64–66).

Le traitement des ETMRs consiste en une résection chirurgicale suivie parfois d'un traitement adjuvant avec chimiothérapie et/ou radiothérapie maximale, dont le rôle n'a pas clairement été établi (64).

Les tumeurs atypiques térétoïdes/rhabdoïdes (ATRT) sont des tumeurs embryonnaires du système nerveux central hautement malignes affectant majoritairement les enfants de moins

de 3 ans. Elles constituent environ 1 à 2% des tumeurs pédiatriques du système nerveux central. Tout comme les ETMR, l'incidence précise reste méconnue, certainement dû à la reconnaissance récente de ces tumeurs comme entité pathologique distincte et non plus comme tumeur neuro-ectodermique primitive comme ce fut le cas pendant des décennies, et l'absence d'études prospectives chez les patients atteints de ce type de tumeur (67). Sur le plan histologique, la présence d'une combinaison de cellules rhabdoïdes et d'éléments neuro-ectodermiques, épithéliaux et mésenchymateux ainsi qu'une perte d'expression de SMARCB1/INI et SMARCA4/BRG1 permet d'aider à poser le diagnostic d'ATRTR (67-70).

Jusqu'à présent, il n'y a pas de consensus concernant la prise en charge optimale de cette tumeur. La localisation généralement critique empêche la résection complète chirurgicale de ce type de tumeur et le jeune âge des patients restreint l'intégration optimale de la radiothérapie dans la prise en charge de cette tumeur agressive (67).

3. Sarcomes de l'os et des tissus mous

Les sarcomes sont des tumeurs supposées d'origine mésenchymateuse, constitués de deux groupes distincts que sont : i) les tumeurs osseuses et ii) les sarcomes des tissus mous. Comme leur nom l'indique, les tumeurs osseuses sont retrouvées à différentes localisations le long des os, mais peuvent également se développer à partir du cartilage. Les sarcomes des tissus mous se développent à partir des tissus conjonctifs du corps, des muscles, des articulations, de tissus adipeux, des nerfs, des tissus cutanés profonds et des vaisseaux sanguins (22).

Les sarcomes se produisent également plus fréquemment chez les jeunes adultes et les adolescents et représentent près de 21% de la totalité des cancers pédiatriques solides contre moins de 1% de tous les cancers solides adultes. La grande majorité des sarcomes diagnostiqués (tous âges confondus) sont des sarcomes des tissus mous (87,3% *versus* 12,7% pour les tumeurs osseuses, **Figure I C-3 1**).

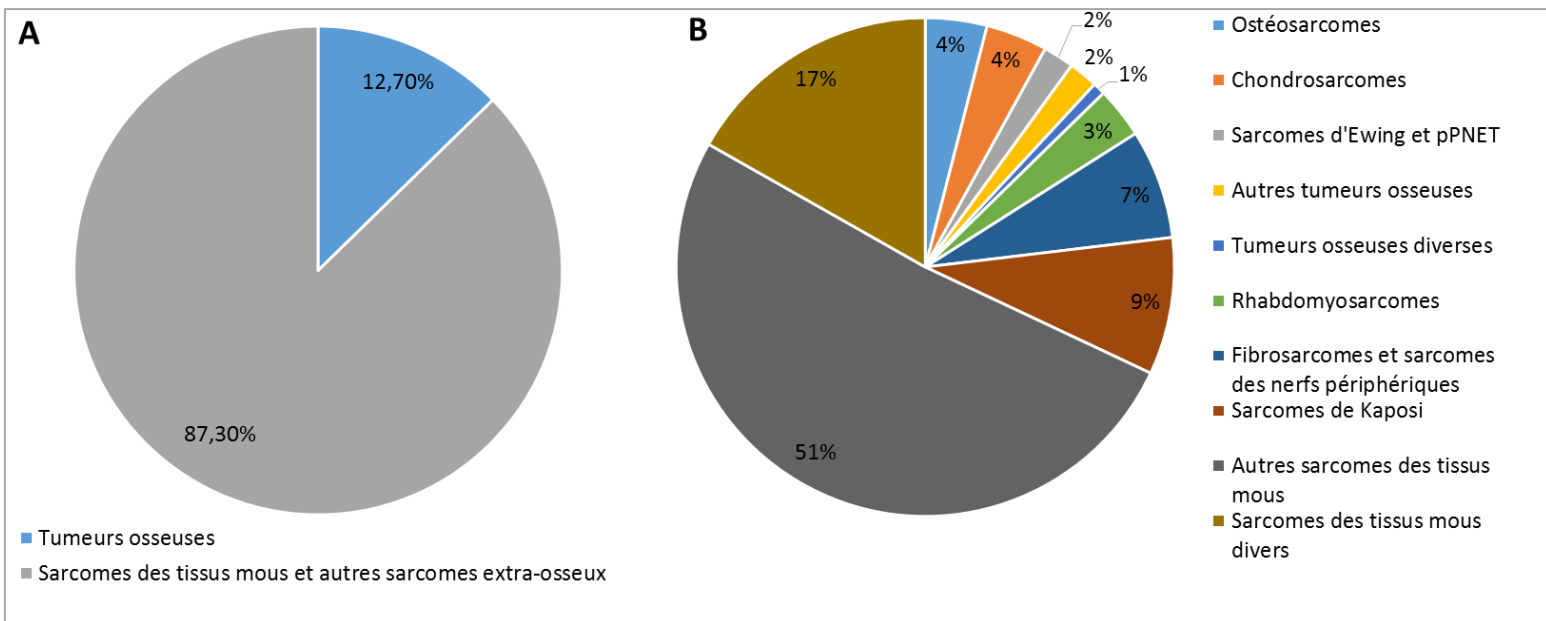


Figure Figure I C-3 1 : Répartition de l'incidence des cas de sarcomes, en fonction des deux familles principales (A) et selon les groupes histologiques de la classification ICC-3 (B).
Adapté d'après : Burningham et. al, 2012

En pédiatrie, l'ostéosarcome et le sarcome d'Ewing sont les deux types de tumeurs osseuses les plus courantes chez les enfants et les adolescents tandis que le type le plus fréquent de sarcome des tissus mous est le rhabdomyosarcome (71).

Bien que relativement rares, les sarcomes sont des pathologies avec de forts taux de mortalité, en particulier les sarcomes des tissus mous. La raison principale se trouve dans le manque de symptômes distincts aux stades précoces, conduisant à un diagnostic retardé, à un stade avancé de la maladie avec présence de métastases dans une grande partie des cas. La recherche de nouvelles options thérapeutiques pour les patients pédiatriques avec ces diagnostics est donc une priorité. De plus, les sarcomes sont constitués de différentes pathologies hétérogènes, avec plus de 50 sous-types histologiques distincts et demeurent des maladies très rares ce qui peut les rendre très difficiles à étudier (71). En effet, la classification internationale des cancers de l'enfant, établie par l'agence internationale pour la recherche contre le cancer (IARC pour « International Agency for Cancer Research ») subdivise les deux grands groupes (les tumeurs osseuses et les sarcomes des tissus mous) selon les différents types histologiques des pathologies retrouvées (Tableau I C-3 1).

Tumeurs osseuses malignes
Osteosarcome
Chondrosarcome
Sarcome d'Ewing et autres sarcomes osseux
Sarcome d'Ewing et tumeur d'Askin
pPNET osseux
Autres tumeurs osseuses spécifiques
Néoplasmes malins fibrogéniques osseux
Chordomes malins
Tumeurs malignes odontogènes
Tumeurs malignes osseuses diverses
Tumeurs osseuses non-spécifiques
Sarcomes des tissus mous et autres sarcomes extra-osseux
Rhabdomyosarcomes
Fibrosarcomes, tumeurs des nerfs périphériques et autres néoplasmes fibromateux
Tumeurs fibroblastiques et myofibroblastiques
Tumeurs de la gaine de myéline
Autres néoplasmes fibromateux
Sarcome de Kaposi
Autres sarcomes des tissus mous spécifiques
Tumeurs d'Ewing et tumeur d'Askin des tissus mous
pPNET des tissus mous
Tumeur rhabdoïde extra-rénale
Liposarcomes
Tumeurs fibrohistiocytiques
Léiomyosarcomes
Sarcomes synoviaux
Tumeurs des vasculaires
Néoplasmes osseux et chondromateux des tissus mous
Sarcome alvéolaire des tissus mous (ASPS)
Sarcomes des tissus mous divers
Sarcomes des tissus mous non-spécifiques

Tableau I C-3 1 : Classification internationale des tumeurs osseuses et des sarcomes des tissus pédiatriques. D'après la Classification Internationale des Cancers de l'Enfant, troisième édition (ICCC-3) (10).

Le groupe des sarcomes d'Ewing et autres sarcomes osseux comprend les tumeurs apparentées histogénétiquement, dont les tumeurs primitives neuroectodermiques périphériques (pPNET). Les autres tumeurs osseuses sont composées des tumeurs du cartilage (chondrosarcomes) et d'autres tumeurs osseuses malignes spécifiées (comprenant diverses tumeurs, dont certaines sont extrêmement rares chez les enfants) et les tumeurs osseuses non-spécifiées. Le groupe des sarcomes des tissus mous est subdivisé en 5 sous-groupes. Le sarcome de Kaposi, lié à l'infection par le virus de l'herpès humain 8 (HHV8), se développe particulièrement chez les individus co-infectés par le virus de l'immunodéficience humaine (VIH) et se distingue des autres tumeurs vasculaires en raison de sa grande incidence dans de nombreux pays africains, touchés par l'épidémie du syndrome d'immunodéficience acquise (SIDA).

Dans cette sous partie, je détaillerai les trois types de tumeurs les plus fréquemment diagnostiqués chez les enfants et les adolescents dans le groupe des sarcomes : l'ostéosarcome, le sarcome d'Ewing et le rhabdomyosarcome.

a) Ostéosarcome

L'ostéosarcome est une tumeur squelettique primaire de haut-grade caractérisée par la présence de cellules fusiformes malignes d'origine mésenchymateuse et de dépôts de matrice ostéoïde immature (72). Ces tumeurs sont relativement rares (moins de 1% des cancers totaux), mais avec un pic de l'incidence chez les adolescents. Ainsi les ostéosarcomes sont le type de tumeurs osseuses le plus fréquent en pédiatrie, représentant environ 3,4% des cas de cancers dans cette population et 56% des tumeurs osseuses malignes (72).

Les ostéosarcomes se développent habituellement dans les zones à croissance osseuse rapide, de ce fait ils sont localisées dans la partie métaphysaire des os longs, les sites les plus communs étant le fémur distal (30%), le tibia proximal (15%) et l'humérus proximal (10%) (73) (**Figure I C-3 a. 1**). Les patients présentent généralement une douleur localisée et un gonflement de la zone affectée. L'exposition aux radiations, comme la radiothérapie qui aurait été donnée pour traiter un premier cancer par exemple, et les agents alkylants ont été identifiés comme facteurs favorisant l'apparition d'ostéosarcome, tout comme certains syndromes de prédisposition génétique (ex : rétinoblastome familial et syndrome de Li-Fraumeni) (24).

Au moment du diagnostic, 20% des patients présentent des métastases dont 75% dans les poumons. Aujourd'hui, la majorité des patients est traitée avec la combinaison d'une chimiothérapie néo-adjuvante, suivie par une résection chirurgicale et d'une chimiothérapie adjuvante postopératoire. Indépendamment du régime de chimiothérapie, la résection chirurgicale de la tumeur et des métastases détectables reste une étape essentielle pour améliorer la survie du patient et consiste en une chirurgie d'amputation ou de conservation/sauvetage de membre (72). De ce fait, le taux de survie à 5 ans de l'ostéosarcome est passé de 45% en 1975-79, à 71% en 2003-2009 grâce aux stratégies de chimiothérapie multimodale modernes. Néanmoins la survie globale demeure relativement stable depuis plus de deux décennies et il est très probable que les améliorations futures proviendront de l'addition de thérapies ciblées et du traitement conventionnel (22,74).

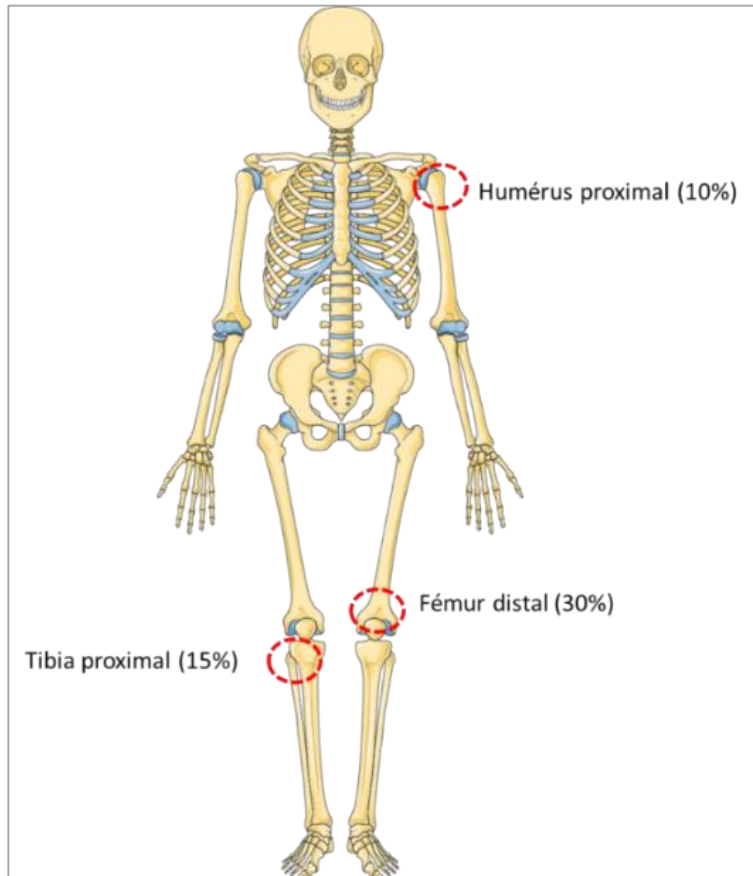


Figure I C-3 a. 1 : Principaux sites de manifestation des ostéosarcomes chez les enfants et les adolescents.

Dans cette catégorie de tumeurs, il est important de mentionner leur aspect génomique particulier, caractérisé par un nombre élevé de variants structuraux mais peu de variants nucléotidiques uniques (SNV pour « Single Nucleotide Variation »). D'autres caractéristiques génomiques propres des ostéosarcomes sont nécessaires d'être mentionnées (75–77):

- Le nombre de mutations faux-sens est d'environ 25 par génome, un nombre plus élevé que la majorité des cancers pédiatriques mais qui reste en dessous des scores observés chez les cancers adultes
- Des altérations du gène *TP53* sont présentes dans la majorité des ostéosarcomes, sous forme de mutations, délétions et réarrangements génomiques
- Quelques cas arborent une amplification du gène *MDM2*, un inhibiteur de P53, et constitue ainsi un autre moyen indirect d'inactivation de ce gène suppresseur de tumeur.
- Le gène *RB1* est communément inactivé
- D'autres gènes fréquemment altérés incluent *ATRX*, *DLG2*, *PTEN* et l'axe PI3K/mTOR en général

b) Sarcome d'Ewing

(1) Incidence, épidémiologie et diagnostic

Le sarcome d'Ewing représente la deuxième tumeur osseuse la plus fréquente (34%) après l'ostéosarcome (78), avec un pic d'incidence dans la deuxième décennie de la vie. Environ 80% des patients diagnostiqués ont moins de 20 ans, avec une médiane d'âge située à 15 ans (79,80). Il a été observé de manière intéressante que les populations caucasiennes étaient plus souvent affectées que les populations asiatiques et africaines (79). De même, les hommes sont plus touchés par cette maladie que les femmes (59% versus 41%)(81). Il a été établi qu'entre 1973 et 2004, un nombre de 2,93 cas/1000000 de sarcome d'Ewing ont été diagnostiqués, le classant ainsi parmi les cancers rares (82).

Des études utilisant des techniques d'immunohistochimie, de cytogénétique, de la génétique moléculaire et de culture de tissus indiquent que le sarcome d'Ewing serait dérivé de cellules souches mésenchymateuses, même si des études réalisées *in vivo* ont remis en cause cette idée (81,83). D'autres nomenclatures anciennement utilisées telles que « tumeur neuro-ectodermique primitive périphérique », « tumeur d'Askin » (sarcome d'Ewing de la paroi thoracique) et « sarcome d'Ewing extra-osseux », souvent regroupées ensemble sous le terme de « tumeurs de la famille Ewing » se réfèrent à la même pathologie du sarcome d'Ewing (81).

La plupart des sarcomes d'Ewing surviennent dans les os (85%) mais peuvent également provenir des tissus mous (15%). Les sites d'apparition les plus communs du sarcome d'Ewing sont les os pelviens, les os longs des membres inférieurs et les os de la paroi thoracique (**Figure 8**) (84). Environ un quart des patients ont des métastases détectables au moment du diagnostic, situées dans la majeure partie des cas dans les poumons et les os ou la moelle osseuse (10% chacun, **Figure I C-3 b. 1-1**) (82)

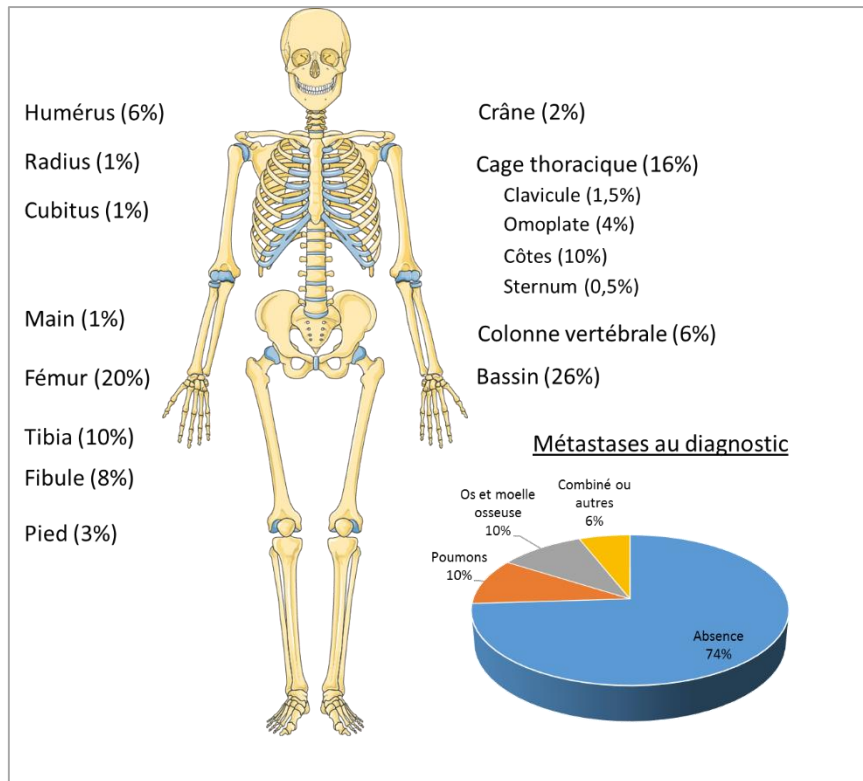


Figure I C-3 b. 1-1 : Différents sites de localisation des tumeurs primaires et des métastases du sarcome d'Ewing. La tumeur primitive survient dans 85% des cas dans les os, préférentiellement dans la cage thoracique, le bassin et les os longs de la partie inférieure, et dans 15% des cas dans les tissus mous. Adapté de Bernstein et. *al* (84).

Les symptômes sont habituellement caractérisés par une douleur osseuse extrême, la présence d'une masse le long de l'os, une fatigue, une perte de poids inexpliquée et une fièvre intermittente. Les tests et procédures suivants peuvent alors être utilisés pour établir le diagnostic (81):

- Imagerie par résonance magnétique (IRM).
- Scanner TDM (tomodensitométrie, afin de détecter mes métastases pulmonaires)
- Tomographie par émission de positons (TEP).
- Scintigraphie osseuse, afin de détecter les métastases osseuses
- Aspiration de la moelle osseuse et biopsie.
- Rayons X
- Numération sanguine complète.
- Des études biochimiques, comme le dosage de la lactate déshydrogénase (LDH).

(2) Caractéristiques moléculaires

Le sarcome d'Ewing est caractérisé dans plus de 85% par la translocation $t(11 ; 22) (q24 ; q12)$ à l'origine du gène de fusion *EWS-FLI1* (85) (**Figure 9**). Le diagnostic moléculaire, utilisant soit

l'hybridation in situ fluorescente (FISH) pour détecter le gène de fusion, soit la PCR après reverse transcription (RT-PCR) font maintenant partie de l'examen pathologique de routine (82). La fusion entre la partie N-terminale de EWS (codant pour une protéine de liaison à l'ARN fortement exprimée) avec le domaine de liaison à l'ADN de FLI1, un membre de la famille des facteurs de transcription ETS (de l'anglais « Erythroblastalis Transforming-Virus-1 ») conduit à une protéine chimérique à activité transcriptionnelle aberrante et responsable de la prolifération cellulaire et de la tumorigénicité (85) (**Figure I C-3 b. 2-1**). Plus précisément, Kovar et. al (86) ont montré que la protéine de fusion EWS-FLI1 avait une double fonction dans le processus de transformation oncogénique :

- Une inhibition des protéines P21 et P53 conduisant à une inhibition de l'apoptose dans les cellules tumorales.
- Une activation de la voie de signalisation d'IGF1, contribuant à une augmentation de la prolifération cellulaire.

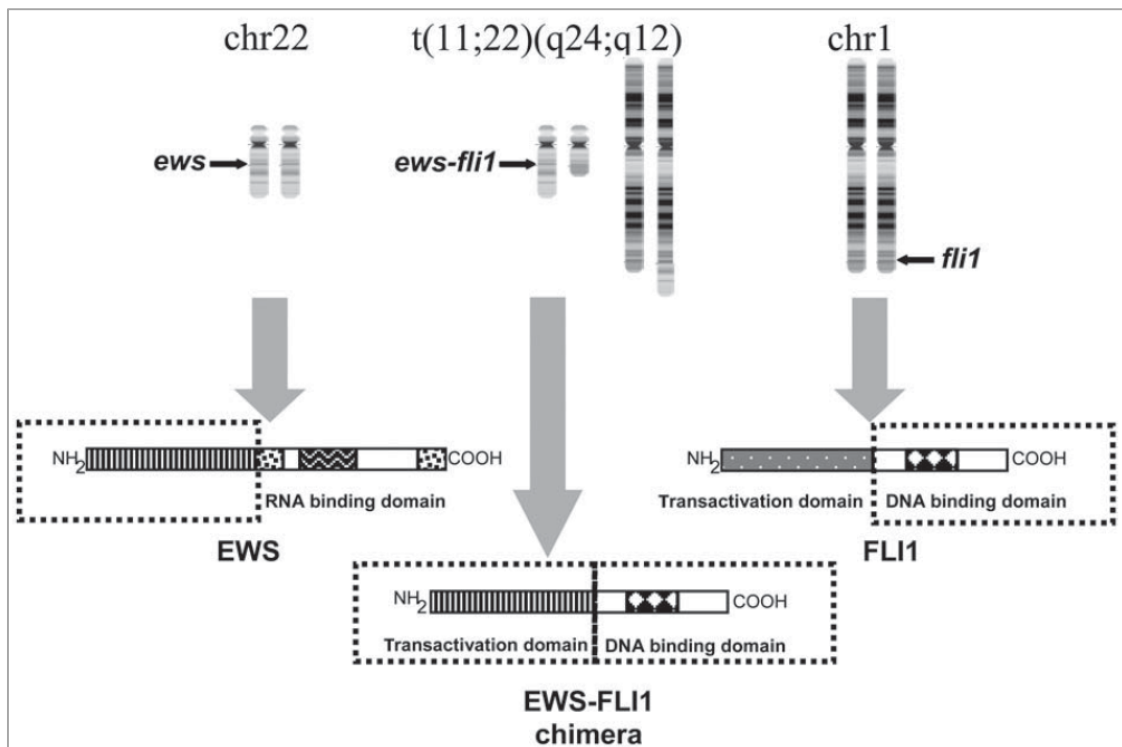


Figure I C-3 b. 2-1 : La translocation réciproque entre les chromosomes 11 et 22 conduisant à la formation du gène de fusion *EWS-FLI1* et, après transcription et traduction, à une protéine chimérique EWS-FLI1 comprenant la partie N-terminale de EWS responsable de la régulation transcriptionnelle et la partie C-terminale de FLI1 contenant le domaine de liaison à l'ADN. D'après Bernstein et. al (84)

(3) Traitements et taux survie

Le protocole de traitement utilisé comprend généralement une chimiothérapie multimodale néo-adjuvante avec chirurgie et/ou radiothérapie afin de traiter la tumeur localement puis d'une

chimiothérapie adjuvante supplémentaire pour prévenir des rechutes (81). La radiothérapie peut parfois être administrée seule dans le cas de tumeurs situées dans des endroits difficilement accessibles ou trop grandes pour être réséquées par une intervention chirurgicale. Cependant, il a été montré que ces patients avaient un mauvais pronostic et que la thérapie échouait dans 26% des cas, un nombre supérieur au taux de rechute observé après chirurgie suivie ou non de radiothérapie (4-10%) (87).

L'introduction de chimiothérapies multimodales en complément de la radiothérapie a considérablement amélioré le taux de survie, passant de 59% à 78% chez les enfants de moins de 15 ans et de 20% à 60% chez les adolescents de 15 à 19 ans entre 1975 et 2010 (78,81). Des études internationales coopératives ont conduit à l'établissement d'un protocole de chimiothérapie utilisant une combinaison d'agents anti-cancéreux : vincristine, doxorubicine, cyclophosphamide, ifosfamide et étoposide, bien que les protocoles thérapeutiques diffèrent entre l'Europe et l'Amérique du Nord (82). Cependant, les traitements conventionnels de cette maladie n'ont pas évolué depuis trois dernières décennies et induisent de nombreux effets indésirables à long terme. La plupart des études ont alors tenté de comprendre la pathogenèse des sarcomes d'Ewing ainsi que les voies de signalisation activées afin de trouver de nouvelles cibles thérapeutiques. La cible préférentielle étant l'oncogène de fusion EWS-FLI1, mais les techniques de thérapies ciblées bloquant son produit n'ont abouti à aucun résultat clinique significatif (78), par conséquent de nouvelles cibles sont nécessaires pour cette pathologie.

Différents facteurs peuvent influencer le pronostic des patients atteints de sarcome d'Ewing. La localisation de la tumeur primaire au niveau des extrémités distales, une tumeur extra-osseuse, un diagnostic à un jeune âge et enfin l'absence de métastases constituent des facteurs de bon pronostic (81). La présence de métastase reste le prédicteur le plus puissant puisque le taux de survie à 5 ans des patients non-métastatiques au diagnostic est de 70% contre 20-30% en présence de métastases (79).

c) Rhabdomyosarcomes

Le rhabdomyosarcome (RMS) est un cancer se développant habituellement dans les muscles squelettiques. Ce cancer représente environ 3,5% des cancers infantiles (0-14 ans) et 2% des cancers de l'adolescent (15-19 ans), avec un pic de l'incidence chez les enfants de moins de 6 ans et au début de l'adolescence (22). Cette incidence est plus ou moins corrélée avec le type de rhabdomyosarcome (24). En effet, il existe deux sous-types majeurs: le rhabdomyosarcome embryonnaire (eRMS ; environ 75% des cas), plus fréquent chez les enfants de moins de 5 ans et chez les garçons, et le rhabdomyosarcome alvéolaire (aRMS ; environ 23% des cas), dont l'incidence ne varie pas selon l'âge (22). D'autres rhabdomyosarcomes de types mixtes,

pléiomorphiques ou anaplasiques et à cellules fusiformes représentent moins de 2% des cas de rhabdomyosarcomes pédiatriques (88).

Les sites anatomiques majeurs de localisation des RMS sont les voies génito-urinaires (24%), la zone paraméningée (16%), les extrémités (19%), les orbites (9%), la tête et le cou (10%) et divers autres sites constitués de tissus mous (22%) (89). Le plus souvent, les patients souffrant de RMS présentent une masse à l'endroit de la tumeur, cependant les signes et les symptômes associés dépendant principalement du site de la tumeur primaire (89).

La majorité des cas de rhabdomyosarcomes alvéolaires est caractérisée, comme dans le cas des sarcomes d'Ewing, par la présence d'une translocation spécifique donnant lieu à un gène de fusion. Il s'agit des fusion de type PAX3/7-FKHR caractérisées par la translocation $t(2;13)(q35;q14)$ ou, moins fréquemment, $t(1;13)(p36;p14)$ donnant lieu respectivement aux fusions PAX3-FOXO1 (55-70% des cas) et PAX7-FOXO1 (10-22% des cas) (90). La protéine de fusion qui en résulte présente une activité transcriptionnelle aberrante due à la proximité entre le domaine de liaison à l'ADN de PAX et le domaine d'activation de FKHR (52) (**Figure I C-3 c. 1**). La détection de ces fusions est de ce fait utilisée en routine pour le diagnostic du rhabdomyosarcome alvéolaire.

Généralement, les patients atteints de rhabdomyosarcome reçoivent une chimiothérapie multimodale en plus d'une chirurgie, une radiothérapie ou une combinaison des deux. Bien que le taux de survie du rhabdomyosarcome se soit amélioré (de 49% en 1975-1979 à 64% en 2003-09), il reste inférieur à celui observé pour de nombreux autres cancers pédiatriques et de nouvelles études sont nécessaires (22).

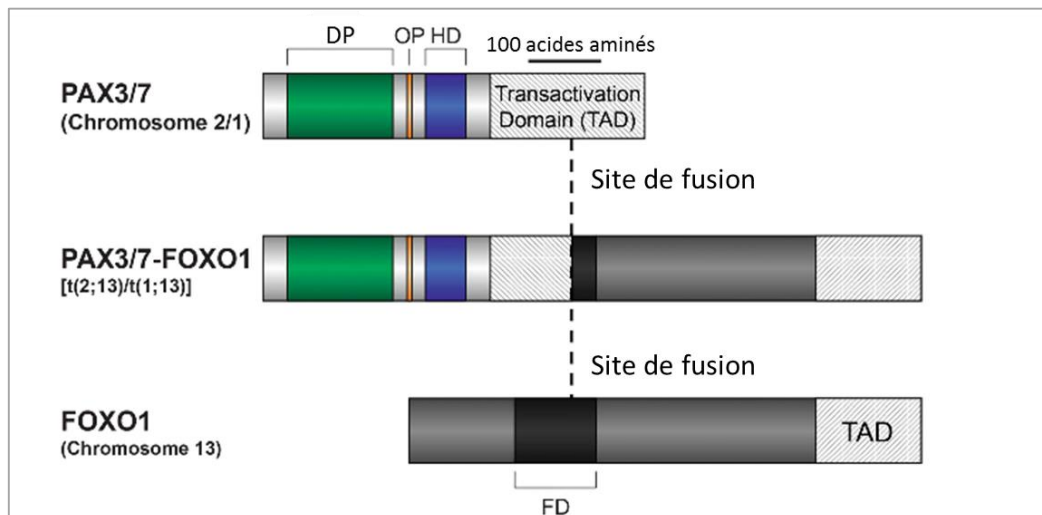


Figure I C-3 c. 1 : Structure des protéines de fusions dérivées des gènes de fusion PAX-FOXO1, uniques au rhabdomyosarcome alvéolaire. Ces fusions résultent de translocations chromosomiques rapprochant les gènes PAX3 ou PAX7 situés sur les chromosomes 2 et 1 respectivement, et le gène FOXO1 est situé sur le chromosome 13. La protéine chimère obtenue contient alors le domaine de liaison à l'ADN de PAX et le domaine de transactivation de FOXO1. DP = domaine « paired » ; OP = octapeptide ; HD = homéodomaine ; « Transactivation domain » = domaine de transactivation. Source : Mahajan et. *al* (91).

4. Autres types tumoraux

Les autres types tumoraux retrouvés chez les enfants et les adolescents de manière plus rare comprennent :

- Les tumeurs rénales
- Les tumeurs hépatiques
- Les tumeurs des cellules germinales et des gonades
- Les mélanomes
- D'autres localisations tumorales diverses

Ces derniers sont détaillés et classés dans le **tableau I C-4 1**. Dans ce groupe on retrouve notamment les tumeurs de type « carcinomes ». Ce terme définit les cancers se développant à partir d'un tissu épithélial et peuvent être de 3 types : épidermoïdes (peau, poumons, œsophage par exemple), glandulaires (sein, reins, thyroïde, etc.) ou indifférenciés (aucune structure tissulaire reconnaissable). Alors que leur incidence est très haute chez les adultes et constitue le type tumoral majoritaire, ceux-ci sont beaucoup plus rares chez les enfants, avec une incidence inférieure à 1 cas par million chez les moins de 20 ans, et représentant approximativement 6% des cancers pédiatriques (25). Chez les jeunes patients, les tumeurs montrant une différenciation épithéliale tendent particulièrement à défier les limites du terme «carcinome». A titre d'exemple,

les tumeurs de Wilms et les hépatoblastomes sont souvent en grande partie épithéliaux mais ne sont généralement pas considérés comme carcinomes.

Parmi les tumeurs les plus fréquemment étudiées et retrouvées en oncologie pédiatrique de ce groupe peuvent être cités le carcinome des cellules rénales, le carcinome hépatocellulaire et les tumeurs malignes des cellules germinales gonadiques.

Tumeurs rénales	Autres néoplasmes épithéliaux et mélanomes malins
Néphroblastomes et autres tumeurs rénales non épithéliales	Carcinomes adrénocorticaux
Carcinomes rénaux	Carcinomes thyroïdiens
Tumeurs malignes rénales non spécifiées	Carcinomes naso-pharyngés
Tumeurs hépatiques	Mélanomes malins
Hépatoblastome	Carcinomes cutanés
Carcinomes hépatiques	Autres carcinomes et types non spécifiés
Tumeurs malignes hépatiques non spécifiées	Carcinomes des glandes salivaires
Tumeurs des cellules germinales, tumeurs trophoblastiques et néoplasmes des gonades	Carcinomes colorectaux
Tumeurs germinales intracrânielles et intraspinales	Carcinomes de l'appendice
Tumeurs germinales malignes intracrânielles et extragonadiques	Carcinomes pulmonnaires
Tumeurs malignes germinales gonadiques	Carcinomes du thymus
Germinomes gonadiques malins	Carcinomes mammaires
Tératomes gonadiques malins	Carcinomes du col de l'utérus
Carcinomes embryonnaires gonadiques	Carcinomes de la vessie
Tumeurs du sac vitellin	Carcinomes oculaires
Choriocarcinome gonadique	Carcinomes d'autres sites spécifiés
Tumeurs des gonades malignes mixtes	Autres néoplasmes malins et types non spécifiés
Gonadoblastome gonadique malin	Autres tumeurs malignes spécifiées
Carcinomes des gonades	Tumeur stromale gastrointestinale
Autres tumeurs malignes des gonades non spécifiées	Pancréoblastome
	Blastome pulmonnaire et pleuropulmonnaire
	Autres néoplasmes complexes mixtes et néoplasmes stromaux
	Mésothéliome
	Autres tumeurs malignes spécifiées
	Autres tumeurs malignes non spécifiées

Tableau I C-4 1: Autres types tumoraux retrouvés parmi les cancers pédiatriques et classification selon la Classification Internationale des Cancers de l'Enfant, troisième édition (« International Classification of Childhood Cancers » ICC-3) (52).

II. Apport de la génétique moléculaire dans la médecine de précision utilisée en cancérologie

La découverte de nouvelles cibles ainsi que la meilleure compréhension des cancers pédiatriques et des cancers en général, sont aujourd'hui accélérées par l'arrivée des nouvelles techniques de séquençage à haut-débit. Ce type de séquençage est également appelé séquençage nouvelle génération, séquençage massif, séquençage deuxième génération ou séquençage NGS (de l'anglais « Next Generation Sequencing »).

Ces technologies ont été introduites sur le marché dans les années 2000 pour répondre à la limitation que présente la méthode standard Sanger. Celle-ci permettait de lire des zones comprenant jusqu'à 1000 paires de bases en une seule réaction rendant le séquençage d'un génome ou même d'un exome humain inenvisageable en un temps limité. Au lieu de séquencer un seul fragment d'ADN, cette nouvelle génération de séquençage permet de séquencer en parallèle des millions de fragments d'ADN différents en une seule réaction, la rendant significativement plus rapide, moins chère, plus précise et plus sensible que la méthode Sanger. L'arrivée de ces techniques au milieu des années 2000 a donc permis le séquençage du génome entier, de l'exome et du transcriptome, entraînant une baisse de 50 000 fois du coût du séquençage du génome humain tout en augmentant les capacités de séquençage (**Figure II-1**). Le développement du NGS permet aujourd'hui à la fois de mieux comprendre les complexités de certains génomes et d'être utilisé comme outil clinique (92,93).

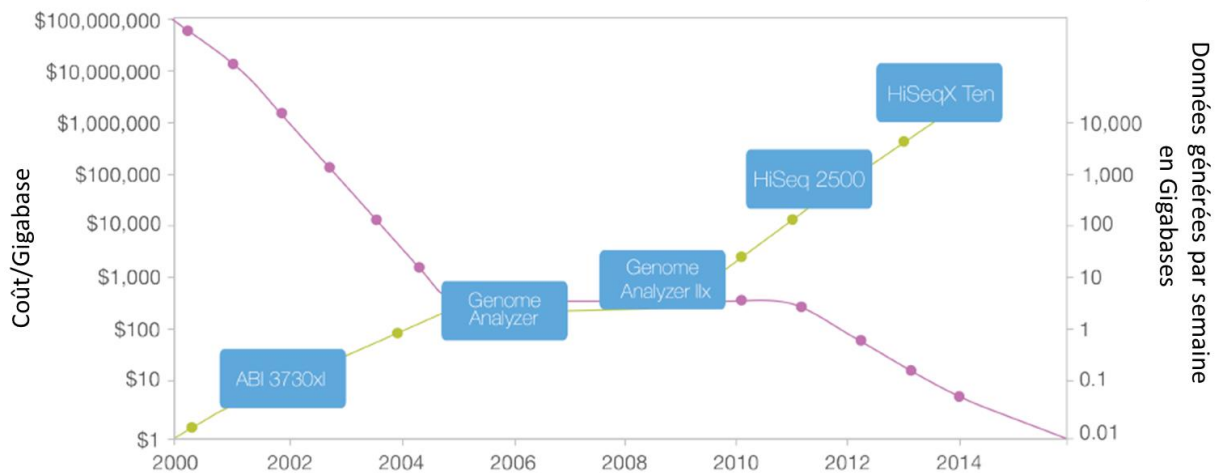


Figure II-1 : Evolution du coût du séquençage et des données générées depuis 2000. Les valeurs des axes des ordonnées des deux côtés du graphique sont logarithmiques. L'axe violet représente le coût et l'axe vert le nombre de données générées par semaine. Adapté d'après « An introduction to Next-Generation Sequencing Technology », Illumina (94).

Depuis, un nombre croissant de mutations, transcriptomes et méthylomes est répertorié dans le monde, dans de nombreuses formes de cancer (95-97) et contribue à l'effort commun pour valider de nouvelles altérations oncogéniques, cibles médicamenteuses et biomarqueurs cliniques (98). Nos connaissances sur le génome cancéreux se sont considérablement améliorées et ces avancées ont permis de comprendre son extrême variabilité biologique, avec une hétérogénéité inter-diagnostique, interindividuelle et intra-tumorale (99,100). Il a entre autres été constaté que cette hétérogénéité pourrait expliquer les différences de réponses aux traitements observées entre les patients, qui ont chacun une biologie tumorale unique (101). Ces variations doivent alors

se refléter dans le choix des thérapies affiliées à chaque patient pour permettre une efficacité maximale avec des effets secondaires minimales, ce qui a été regroupé sous le terme de « médecine personnalisée » ou « de précision ». La médecine de précision est basée sur la connaissance du répertoire des altérations génétiques oncogéniques et la disponibilité de molécules anti-cancéreuses ciblant ces altérations ou les voies de signalisation associées (98). Ainsi, l'émergence de programmes de criblages massifs des altérations moléculaires chez les patients, d'essais cliniques de médecine de précision et de nouvelles thérapies ciblées est aujourd'hui en augmentation, avec pour but de traiter et classer les patients directement au regard de leurs altérations moléculaires (14,15,102,103).

A. Anomalies génétiques recherchées en cancérologie et techniques d'analyses de génétique moléculaire à haut-débit

1. Principe général du séquençage haut-débit

Différentes techniques sont actuellement disponibles sur le marché pour le séquençage haut-débit, dont les trois principales sont : 454 Life-sciences/Roche (104), Illumina/Solexa (105,106) et Life Technologies (92). Elles reposent sur des principes différents et sont sujettes à des types d'erreurs spécifiques. Une description et comparaison des différentes plateformes de NGS est fournie par Goodwin et. *al* (92). Ces différentes technologies consistent cependant en quatre étapes communes :

1 / Préparation des bibliothèques d'ADN:

Etant donné que le séquençage haut débit se fait uniquement sur des molécules d'ADN, chaque séquençage débute par l'extraction du matériel à séquencer et la préparation des bibliothèques (matrices d'ADN qui seront lues) qui varie selon la source de ce matériel [ADN génomique (ADNg), ARN rétrotranscrit aussi appelé ADN complémentaire (ADNc) ou ADN immunoprécipité].

Quelle que soit la méthode utilisée, la préparation d'une bibliothèque a pour but de casser l'ADN en fragments et d'y ajouter des adaptateurs aux extrémités. Ces adaptateurs vont servir d'amorces pour les réactions d'amplification et/ou de séquençage en aval (107). La bibliothèque peut ensuite être séquencée directement (une seule molécule d'ADN séquencée) ou amplifiée avant séquençage (méthode utilisée dans mon étude) (108).

Pour l'analyse d'exome et les séquençages de zones ciblées, une étape d'enrichissement est nécessaire, qui peut être effectué par hybridation à des séquences complémentaires (capture de séquence) ou par PCR (107).

2/ Amplification des molécules par PCR :

La plupart des plates-formes de séquençage ne peuvent aujourd’hui pas séquencer une molécule unique d’ADN. Une amplification est donc nécessaire pour produire un signal suffisant pour détecter l’addition des nucléotides au cours du séquençage. Les stratégies d’amplification varient selon les plates-formes (92) et inclue généralement des amplifications en pontage (**Figure 18**). Grâce à la présence d’adaptateurs à la fois sur les fragments d’ADN et sur une matrice solide, l’étape d’amplification permet également de séparer et immobiliser les différentes populations de fragments d’ADN sur cette surface solide, paramètre nécessaire pour que les millions de micro-réactions de séquençage puissent s’effectuer en parallèle sur chaque matrice spatialement distincte (**Figure II A-1 1**) (108).

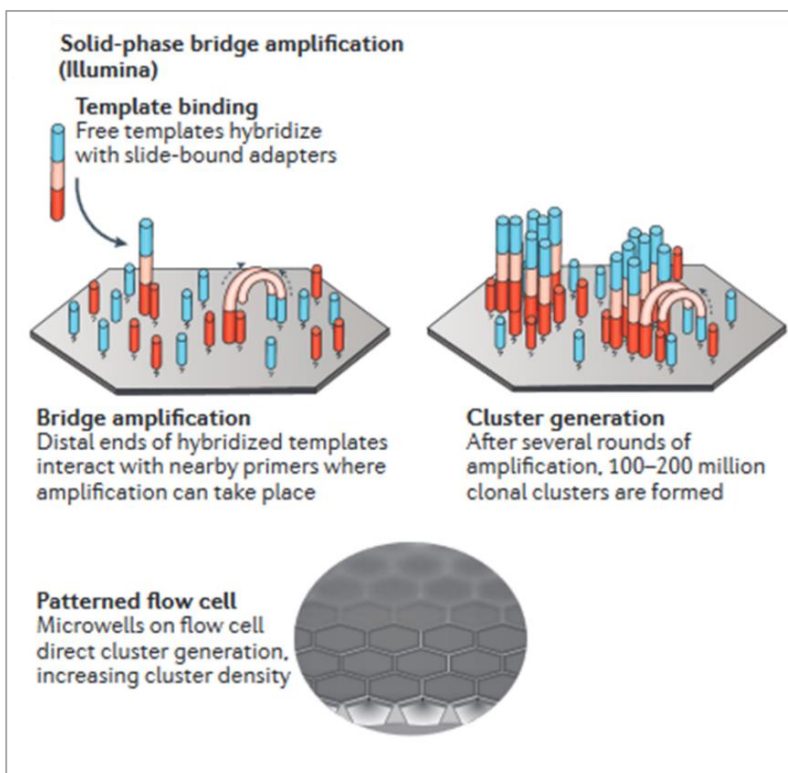


Figure II A-1 1: Stratégie d’amplification des matrices d’ADN par pontage sur surface solide (technique Illumina). L’ADN fragmenté est hybridé à une amorce immobilisée sur support solide grâce aux séquences des adaptateurs flanqués aux extrémités. L’extrémité libre non hybridée peut ensuite interagir avec d’autres amorces proches, formant une structure en pont. Une réaction de PCR est utilisée pour synthétiser le deuxième brin. Source : Goodwin et. al, (91).

3/ Réaction de séquençage (par « ligation » ou par « synthèse ») :

Les données de séquençage utilisées au cours de mon projet de thèse ont été générées par séquençage de synthèse de la technologie Illumina, à l’origine d’environ 90% des données de séquençages générées (94). Dans cette technique, l’ADN polymérase catalyse l’incorporation de désoxyribonucléotides triphosphates (dNTPs) marqués par 4 fluorochromes différents et un terminateur labile, dans un brin de matrice d’ADN fixé sur une surface solide après amplification. A chaque cycle séquentiel de synthèse d’ADN, les dNTPs sont incorporés, le terminateur empêchant l’ajout d’une nouvelle base. Le milieu réactionnel est ensuite lavé pour éliminer les nucléotides non incorporés et les fluorochromes liés aux dNTPs fixés sont excités. Les émissions lumineuses émises sont alors captées puis traduites en une séquence informatique. A la fin de

chaque cycle, le fluorochrome et le terminateur sont clivés (**Figure II A-1 2**). Chaque séquence provenant d'un seul fragment d'ADN lu est appelée « lecture » ou « read ».

Ce séquençage peut être de type simple (« single-end ») ou apparié (pour « paired-end »). En « single-end », une seule extrémité des matrices d'ADN est séquencée, tous les fragments d'ADN ayant une probabilité égale d'être lus dans le sens direct ou inverse. En « paired-end », les lectures directes et inverses sont appariées et les deux extrémités des fragments d'ADN linéaires sont lues (108). Les reads dans le sens direct ou inverse peuvent ou non se chevaucher. Une anomalie dans l'alignement attendu sur le génome de référence entre deux extrémités de reads appariés peut indiquer une variation structurale. La taille des reads obtenue varie globalement entre 30-400 pb pour la technique « single end » et entre 300-500 pb pour le séquençage « paired-end »(108).

Le choix du séquençage simple ou apparié dépend de l'application clinique et du type d'anomalies génétiques recherchées. Le séquençage non-apparié fournit une large quantité de données, rapidement et économiquement. Peu de matériel génétique est nécessaire à ce séquençage, et l'analyse est simplifiée mais cependant moins précise. En plus de permettre une meilleure détection des réarrangements génomiques, le séquençage apparié permet un meilleur alignement sur les régions d'ADN contenant des séquences répétées et produit des lectures de séquences continues plus longues, utiles pour les séquençages de novo. Ce type de séquençage peut avoir des applications multiples puisqu'il ne nécessite ni méthylation de l'ARN ni digestion par des enzymes de restrictions. De plus, la petite taille des « reads » rend l'analyse des données plus précise et de meilleure qualité.

Le séquençage par paire a été utilisé pour générer les données de séquençage analysées dans mon projet de thèse et sera discuté dans la sous-partie suivante.

Deux notions sont également importantes en NGS et sont choisies en fonction de l'analyse réalisée : la **couverture** et la **profondeur**. La couverture représente le pourcentage du génome de référence couvert par au moins un read. La profondeur représente le nombre unique de reads incluant un nucléotide donné dans la séquence reconstituée (109). La plupart des utilisateurs déterminent le niveau de couverture nécessaire en fonction de l'application, de la taille du génome de référence, du niveau d'expression et de la littérature. Plus le matériel génétique à séquencer a une taille importante, plus il sera coûteux de séquencer à une grande profondeur. En revanche, plus le matériel génétique à séquencer est restreint, plus il sera facile de procéder à un séquençage profond. Augmenter la profondeur permet d'augmenter les chances de détecter des altérations rares (109).

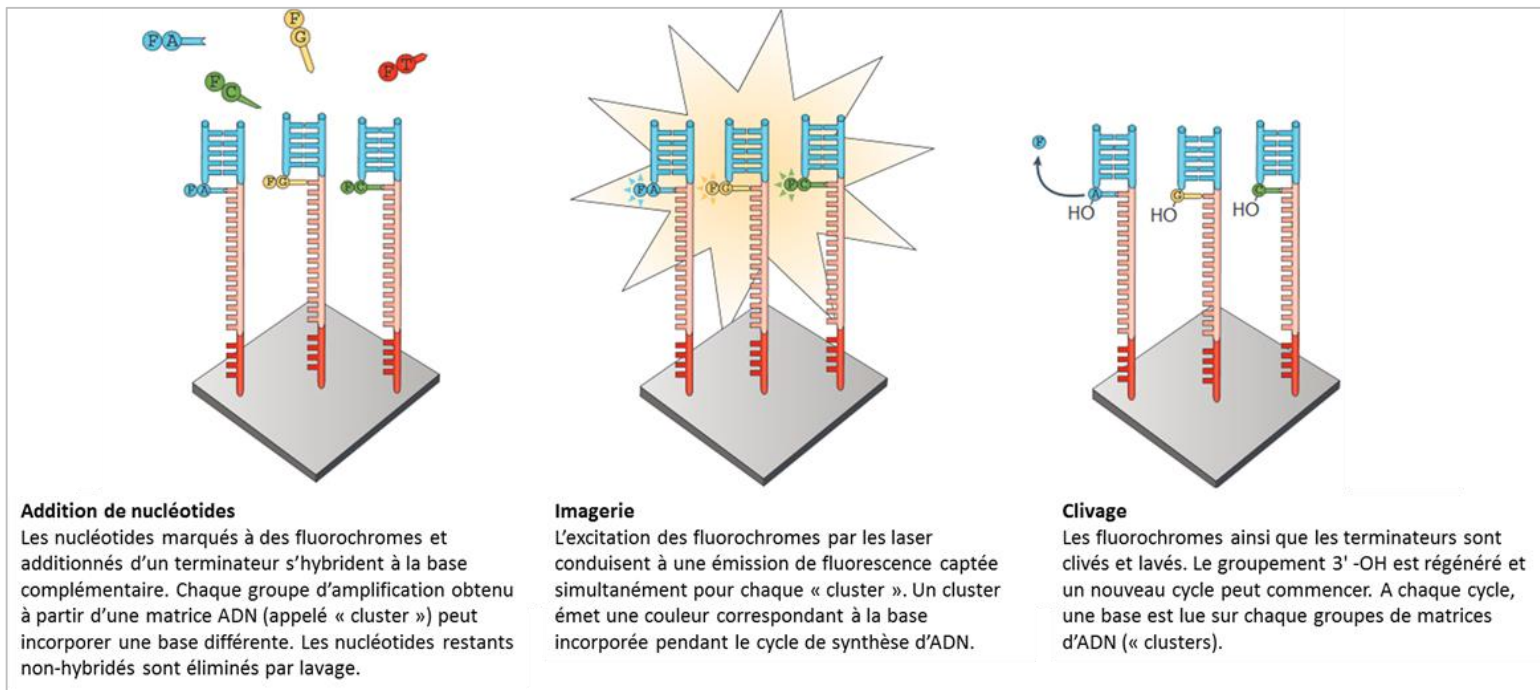


Figure II A-1 2: Séquençage par synthèse: technologie Illumina par terminaison réversible.
Adapté d'après Goodwin et. al, (92).

4/ Analyse des données générées.

Les données brutes lues subissent une série de processus bioinformatiques (regroupés sur le terme de « pipeline ») afin de délivrer un fichier final des altérations génétiques détectées. Ces processus comprennent le démultiplexage (tri de toutes les lectures par code barre/échantillon), l'analyse de la qualité des lectures, la comparaison des lectures à un génome de référence et l'identification/annotation des altérations (107). Lorsqu'un génome de référence n'est pas disponible, les données de séquençage sont assemblées *de novo* (**Figure II A-1 3**). L'alignement de séquence et l'assemblage des reads représentent un domaine actif de la recherche computationnelle avec de nouvelles méthodes constamment développées (108)

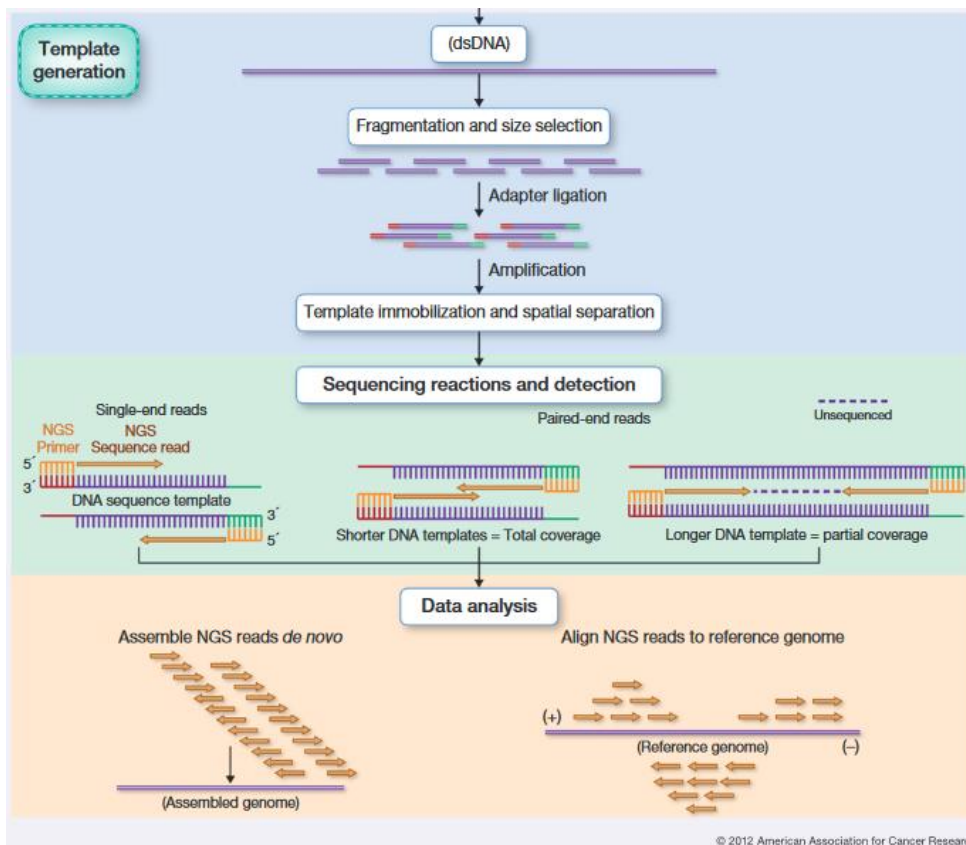


Figure II A-1 3 : Résumé des étapes de préparation de la librairie, séquençage et analyse des données. L'analyse des données peut être fait soit par assemblage du génome obtenu *de novo* ou soit par alignement sur un génome de référence connu. Source : Rizzo et. al, (107).

Les quatre étapes détaillées ci-dessus sont résumées dans la **figure II A-1 4** où sont exposées les méthodes utilisées dans le cadre de la technologie Illumina. Sur cette figure est représenté un exemple de séquençage génomique (panel A) pour rechercher par exemple des mutations chez les patients (panel D). Dans mon projet de thèse, les mêmes étapes ont été réalisées utilisant la même technique Illumina, seules la préparation de la librairie (séquençage effectué sur l'ARNm après conversion en ADNc) et l'analyse des données (recherche de fusions et non de mutations) diffèrent de cette figure II A-1 4 :

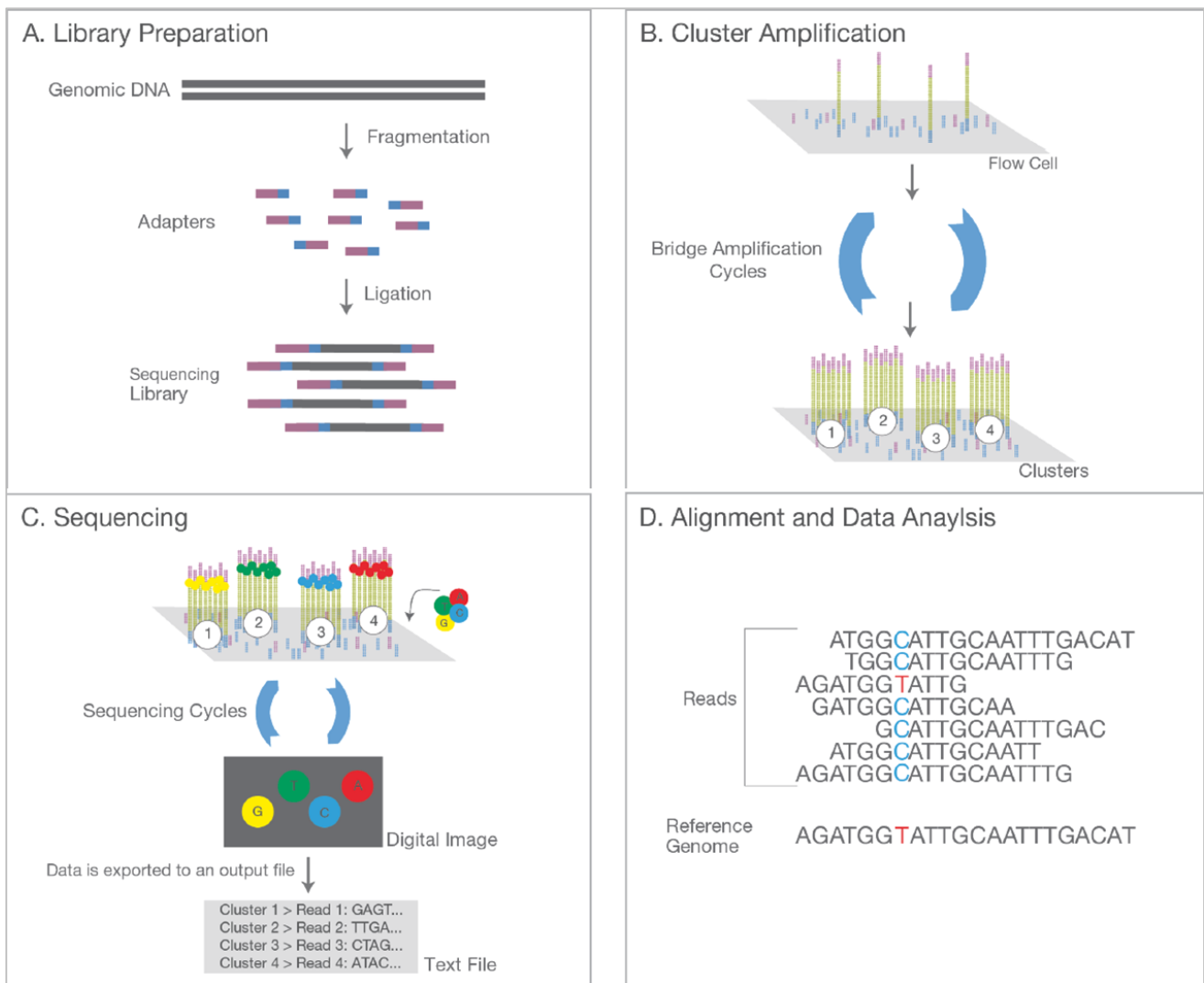


Figure II A-1 4 : Vue d'ensemble de la chimie du séquençage de nouvelle génération de type Illumina, exemple d'un séquençage génomique pour la recherche de mutations. (A) préparation de la librairie, (B) amplification des matrices d'ADN, (C) séquençage, et (D) alignement et analyse des données. Source : « An introduction to Next-Generation Sequencing Technology », Illumina (94).

2. Types d'anomalies moléculaires génétiques recherchées par séquençage haut-débit

Les plateformes de NGS regroupent une grande variété de méthodes, permettant aux chercheurs et aux cliniciens de poser quasiment pratiquement toutes les questions liées au génome, transcriptome ou à l'épigénome. Les méthodes de séquençage diffèrent principalement par la façon dont les échantillons d'ADN ou d'ARN sont obtenus (organisme, type de tissu, cellules normales ou pathologiques, conditions expérimentales, etc.) et par les options d'analyse des données (94,108). Bien que le nombre de méthodes NGS ne cesse de croître, les techniques les plus couramment utilisées et leurs applications en cancérologie sont présentées ici et résumées dans le **tableau II A-2**.

Tableau II A-2 : Résumé des différents types de séquençages possibles et des anomalies génétiques pouvant être détectées par ces technologies.

Génome
Séquençage du génome entier (WGS) <ul style="list-style-type: none">• SNVs• Variants structuraux (y compris les gènes de fusion)• Variations du nombre de copies• Perte d'hétérozygotie
Séquençage de l'exome (WES) <ul style="list-style-type: none">• SNVs• Petites insertions et délétions• Variations du nombre de copies• Perte d'hétérozygotie
Séquençage ciblé

Transcriptome
Séquençage de l'ARN (total et ARNm) <ul style="list-style-type: none">• Quantification de l'expression de gènes• Variants d'épissage• Identification de transcrits de fusion
Séquençage des petits ARN et ARN non-codants <ul style="list-style-type: none">• Détection de miRNA

Epigénome
Séquençage des zones méthylées
Séquençage par immunoprécipitation de la chromatine (ChIP-seq)
Séquençage des fragments associés aux ribosomes

- Le séquençage du génome entier (WGS pour « Whole-Genome Sequencing »), permet un séquençage complet du génome tumoral, incluant les régions intragéniques et les introns. Cette technique permet d'identifier un large type d'anomalies génomiques cancéreuses comprenant les SNVs (« single nucleotide variations » donc les mutations ponctuelles), les variations structurelles, des variations du nombre de copies et des pertes d'hétérozygotie. Généralement et afin de différencier les altérations somatiques des altérations germinales, le génome tumoral ainsi que le génome sain de l'individu correspondant, classiquement obtenu à partir des cellules sanguines,

sont séquencés en parallèle. Les variations structurelles retrouvées en cancérologie comprennent les délétions, les duplications, les insertions, les inversions et les translocations (incluant les gènes de fusion). Elles peuvent être détectées grâce à l'utilisation du séquençage « paired-end », en pointant les paires anormalement alignées sur un génome de référence. Le séquençage du génome entier est généralement utilisé en recherche et offre le maximum de possibilités d'identifier de nouvelles anomalies génétiques liées à l'oncogenèse. Néanmoins, grâce à la baisse significative du coût du séquençage par la méthode NGS, le WGS sera dans le futur de plus en plus appliqué en routine en clinique (108).

- Le séquençage de l'exome (WES pour « Whole-Exome Sequencing»), rendu possible grâce à une étape de capture des exons avant séquençage (110) permet de séquencer les exons d'environ 20000 gènes connus. Le WES donne des informations sur les SNVs, les insertions, les délétions et les altérations du nombre de copies retrouvées dans les parties codantes du génome. Le WES est généralement plus fréquemment utilisé dans le cadre de travaux de recherche mais est de plus en plus appliqué en clinique pour la médecine de précision.

- Le séquençage ciblé, permet de séquencer uniquement un petit panel de gènes connus (entre 10 et 500), généralement associés à des anomalies pouvant être ciblées par des thérapies disponibles. Il donne des informations sur les SNVs, les insertions, les délétions et les variations du nombre de copies. Malgré le fait que très peu d'informations soient obtenues avec ce type de séquençage, celui-ci est très utilisé en routine clinique de par son faible coût (94,108).

- Le séquençage de l'ARN est utilisé afin de détecter des variants d'épissage, des transcrits de fusion et de quantifier l'expression de gènes. Selon si les ARN totaux, messagers, ou non-codants sont extraits, différentes informations peuvent être obtenues. Le séquençage du transcriptome fournit un aperçu complet du profil transcriptionnel cellulaire à un moment biologique donné et améliore la puissance des méthodes de découverte de nouveaux transcrits.

- Le séquençage de l'épigénome est l'une des dernières techniques de séquençage et permet de séquencer les zones génomiques méthylées ou associés à de la chromatine (94).

L'identification des anomalies génétiques est réalisée par l'utilisation de logiciels bioinformatiques spécifiques et dépend du profil d'alignement des reads au génome, au transcriptome ou à l'épigénome de référence choisi (111). La **figure II A-2** illustre les différents profils de répartition des reads obtenus pour les altérations du génome. Il est important de souligner que l'identification de translocations au niveau génomique ainsi que de transcrits de

fusions au niveau transcriptomique n'est possible que par la technique de séquençage apparié et sera explicitée ci-après.

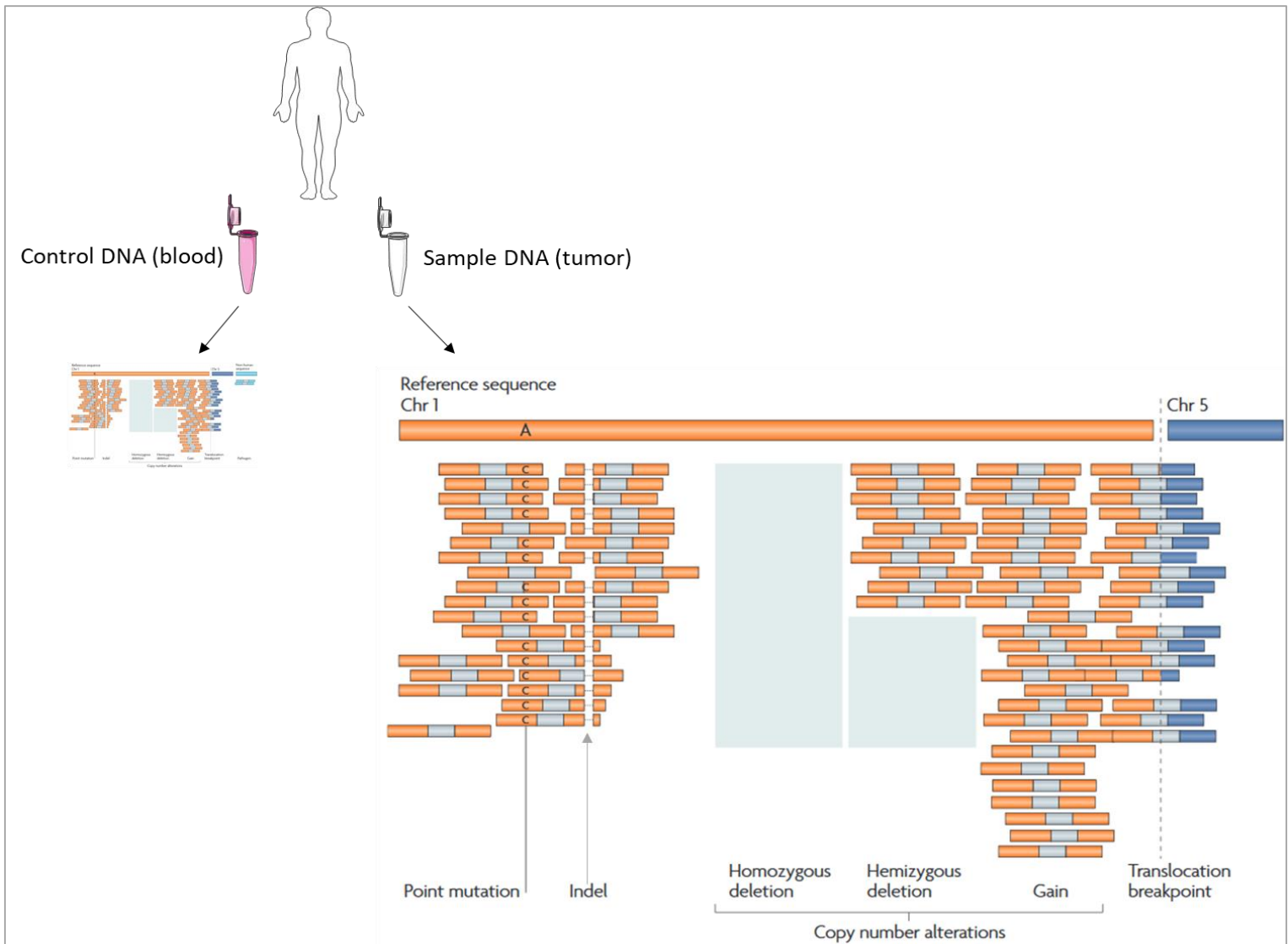


Figure II A-2 : Types d'altérations du génome tumoral pouvant être détectées par séquençage haut-débit. L'ADN contrôle « sain » du patient ainsi que l'ADN tumoral sont extraits et séquencés. La comparaison des deux profils permet d'établir les altérations génétiques tumorales présentes. Sur cette figure, les fragments d'ADN de la librairie sont représentés sous forme de barres dont les extrémités colorées (en orange ou bleu) correspondent aux extrémités séquencées (les « reads »). La partie grise de chaque barre représente la partie non séquencée du fragment. Les lectures sont alignées sur le génome de référence (principalement le chromosome 1 dans cet exemple) et les couleurs des reads indiquent sur quel chromosome ils s'alignent (ici orange = chromosome 1 ; bleu = chromosome 5). Les mutations ponctuelles ici représentées par une mutation d'une adénosine à une cytosine, et les petites insertions et délétions (indels) indiquées par une flèche, sont identifiées par plusieurs reads indiquant une séquence non-référencée. Les changements dans la profondeur de

séquençage (par rapport au contrôle sain) sont utilisés pour identifier les changements de nombre de copies : les carrés gris représentent des lectures absentes (délétion homozygote ou hemizygotique) ou gagnées (gain/amplification) dans l'échantillon tumoral. Les paires de reads s'alignant sur différents loci génomiques (dans ce cas sur les chromosomes 1 et 5) indiquent un réarrangement plus large (ici une translocation). Source : Meyerson et al, (111).

3. Cas des puces d'hybridation génomique comparative (CGH array)

La technique des puces d'hybridation génomique comparative (aCGH pour « array comparative genomic hybridization ») est conçue pour identifier les régions génomiques avec altérations du nombre de copies. Dans cette méthode, l'ADN provenant à la fois des génomes de référence (sain) et testé (tumoral) sont marqués différemment avec des fluorochromes et hybridés de manière compétitive à des cibles d'ADN disposées sur une puce (**Figure II A-3**). La puce hybridée est ensuite lue par excitation puis émission des fluorochromes et le ratio d'intensité de signal sur chaque cible d'ADN entre la condition saine et tumorale reflète l'état du nombre de copies des gènes à un segment du génome donné (**Figure II A-3**). De nombreuses avancées dans la technologie CGH ont été faites depuis son développement au milieu des années 1990 avec une meilleure couverture du génome, améliorant sa résolution.

L'aCGH est généralement utilisé en clinique de par son faible coût, et permet de fournir une information sur les altérations du nombre de copies lorsque le WGS n'est pas disponible. Toutefois cette technique ne permet pas de détecter les échanges de matériel génétique équilibrés comme les translocations réciproques.

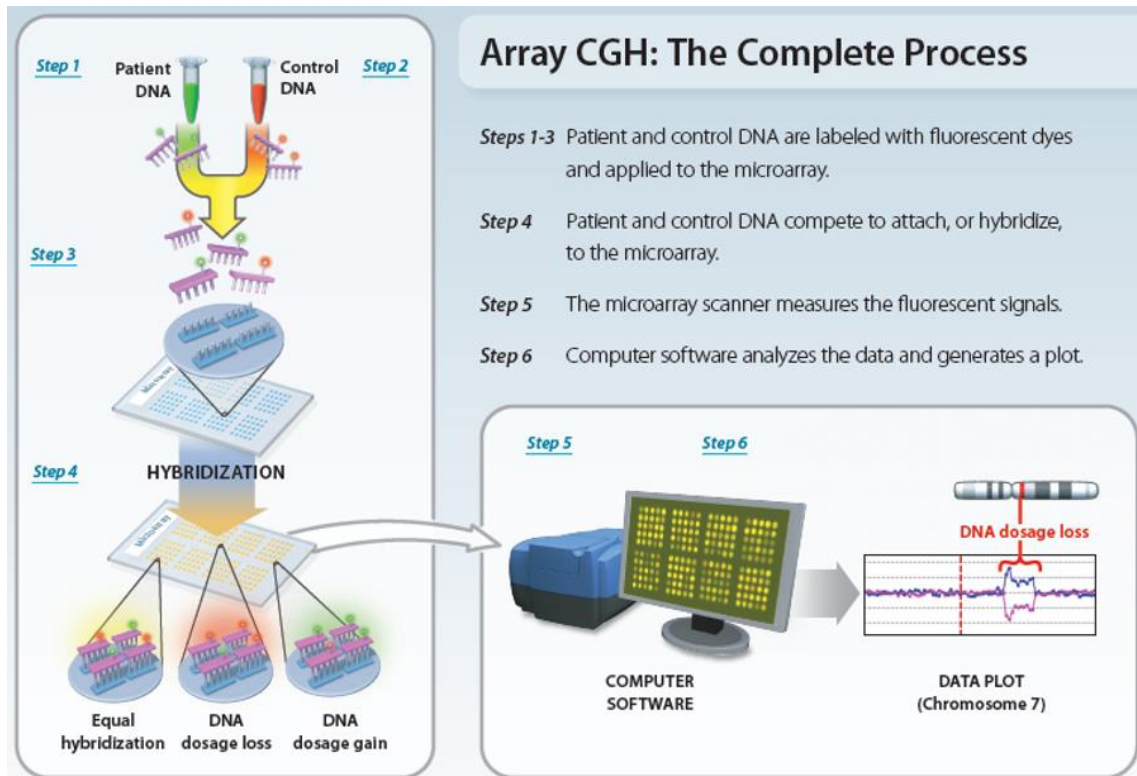


Figure II A-3: Schéma des étapes de l'utilisation des puces d'hybridation génomique comparative (aCGH). L'ADN tumoral et de référence sont marqués de manière différentielle avec de la cyanine-5 et de la cyanine-3 respectivement puis sont hybridés de manière compétitive à une puce d'ADN génomique. La puce est constituée de cibles d'ADN sélectionnées pour couvrir certaines régions chromosomiques ou le génome entier. On a alors plusieurs cas possibles, soit une hybridation équivalente des deux échantillons, soit une perte d'hybridation dans l'échantillon tumoral (plus d'ADN contrôle hybridé), soit une augmentation de l'hybridation dans l'échantillon tumoral (plus d'ADN tumoral hybridé). La puce est ensuite lue par un scanner afin de mesurer les signaux de fluorescence, puis le rapport des deux intensités de fluorescence est calculé bioinformatiquement. Pour chaque cible, ce rapport indique le nombre de copies relatives. La cible est ensuite retracée par rapport à sa position dans le génome humain pour générer un profil par chromosome. *Source: Theisen, A. (2008), Nature Education.*

B. Conduite d'essais cliniques guidés par l'analyse du profil moléculaire tumoral de patients, obtenu par séquençage à haut débit

1. Généralités sur les essais cliniques de médecine de précision en oncologie et exemples

La médecine de précision est définie par l'Institut National de la Santé (NIH « National Institute of Health ») comme une approche prenant en compte la variabilité des gènes, des protéines, de l'environnement et du style de vie de chaque patient, utilisée pour la prévention, le diagnostic et le traitement de maladies (**Figure II B-1**). Elle combine la pharmacologie et l'étude des gènes, des protéines, des métabolites et de leurs fonctions pour développer des molécules plus efficaces et moins toxiques (112). Les étapes nécessaires pour l'application et l'évaluation des données génomiques en clinique incluent généralement : 1) le diagnostic, 2) l'analyse génomique, 3) la

décision médicale d'orientation de traitement au regard des altérations lorsque le patient présente des anomalies pouvant être ciblées, 4) l'administration de la thérapie correspondante, lorsqu'elle est disponible et 5) le suivi de la réponse du patient (**Figure II B-1 1**).

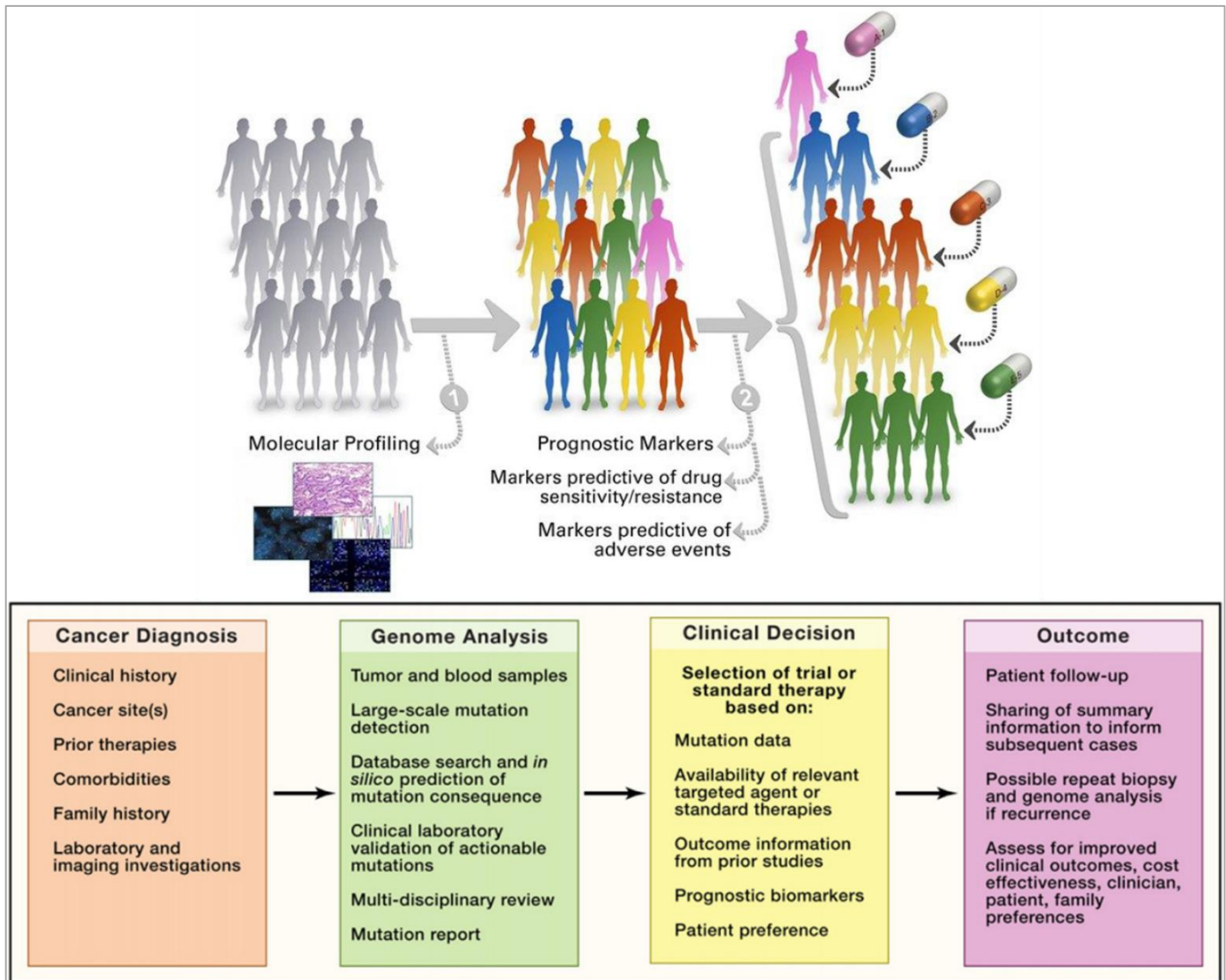


Figure II B-1 1: Concept de médecine “de précision” ou “personnalisée” utilisée dans le traitement des cancers. Les étapes clés sont les suivantes : 1) le recrutement des patients et l'acquisition d'informations cliniques pertinentes, 2) la collecte d'échantillons et l'analyse génomique, 3) l'interprétation des résultats de l'analyse génomique basée sur les conséquences fonctionnelles et cliniques connues liées à une altération donnée et 4) l'apport de ces informations aux cliniciens et aux patients afin d'adapter la prise en charge. Généralement, les patients ayant peu de chances de bénéficier des traitements classiques et ont donc un pronostic relativement mauvais et/ou sont plus susceptibles de bénéficier d'un nouveau traitement en raison de la présence d'anomalies génétiques tumorales prédisant la sensibilité, l'absence de résistance ou la toxicité d'un traitement sont inclus dans les essais cliniques correspondants. Sources : présentation orale du Dr Marc Deloger dans le cadre de la formation intitulée « Genetic alterations in cancer and therapeutic strategies » en 2017, et Dancey et. al, (98).

Quelques exemples de thérapies ciblées les plus connues en cancérologie regroupent : 1) l'imatinib mesylate ciblant la protéine chimère BCR-ABL causée par le chromosome de Philadelphie retrouvé dans plus de 90% de leucémie myéloïdes chroniques (2), 2) le cetuximab un anticorps monoclonal bloquant le récepteur de l'EGF dans le cas de patients mutés pour *KRAS* dans le cas de cancer colorectal (113) ou encore 3) le trastuzumab, un anticorps monoclonal inhibant le récepteur HER2 surexprimé à la surface des cellules de cancers du sein (114).

Les données issues d'analyses rétrospectives et d'essais cliniques précoces ont démontré que la stratégie d'association d'agents ciblés avec des altérations génomiques donnait lieu à des résultats encourageants (115). De ce fait, l'intégration significative de la génomique dans les soins aux patients est de plus en plus accentuée et a donné lieu à l'arrivée d'essais de séquençage « clinique » aussi appelés essais cliniques basés sur l'analyse du génome. Ces études ont pour but d'évaluer la faisabilité d'un profilage moléculaire en routine ainsi que la capacité à analyser et interpréter chaque résultat individuel de séquençage pour permettre d'orienter le traitement du patient et son inclusion dans les essais cliniques correspondants (116).

Contrairement à un essai clinique randomisé classique dans lequel une nouvelle stratégie thérapeutique est évaluée par rapport à la pratique standard actuelle, un essai clinique basé sur la génomique offre la possibilité de sélectionner de nombreuses options thérapeutiques sur la base du profilage moléculaire (98).

Au cours des deux dernières années, sept études de séquençage « clinique » en oncologie pédiatrique ont été publiées et ont abouti à des taux similaires (30-60%) d'altérations pouvant potentiellement être ciblées (**Tableau II B-1 1**) (116).

Etude	Institution	Types de tumeurs	Types de patients	Patients analysés	Altérations ciblables (%)	Année de publication	PMID
Peds-MiOncoSeq	University of Michigan	Tumeurs solides (y compris du SNC) et hématologiques	En rechute/résistants ou à risque élevé au diagnostic	91	46	2015	26325560
iCat	Dana-Farber Institute et autres	Tumeurs solides extracrâniennes	En rechute/résistants ou à risque élevé au diagnostic	89	34	2016	26822149
BASIC3	Baylor College of Medicine	Tumeurs solides (y compris du SNC)	Au diagnostic	150	39	2016	26822237
INFORM	German Cancer Research Center et autres	Tumeurs solides (y compris du SNC) et hématologiques	En rechute/résistants	52	50	2016	27479119
MBB Program	Institut Curie	Tumeurs solides (y compris du SNC)	En rechute/résistants ou à risque élevé au diagnostic	58	40	2016	27896933
PIPseq	Columbia University Medical Center	Tumeurs solides (y compris du SNC) et hématologiques	En rechute/résistants ou à risque élevé au diagnostic	101	38	2016	28007021
MOSCATO-01	Gustave Roussy	Tumeurs solides (y compris du SNC)	En rechute/résistants	69	61	2017	28733441

Tableau II B-1 1: Etudes de séquençage clinique publiées en oncologie pédiatrique. Source : Forrest et. al, (116).

Ces différentes études de séquençage cliniques permettent ensuite de pouvoir orienter les patients vers les essais cliniques de médecine de précision offrant les thérapies ciblées adaptées afin de valider l'apport clinique de ces études.

Un certain nombre d'essais cliniques de médecine de précision en pédiatrie ont été récemment terminés, sont en cours ou bien en développement et peuvent être groupées en 4 catégories (**Figure II B-1 2**) (116) :

- **Les essais cliniques de type “parapluie” (“umbrella”) chez des patients en maladie progressive avec un seul type de diagnostic** : évaluent l'efficacité de plusieurs traitements avec différentes cibles dans un seul type de cancer. Ces essais sont préférables et envisagés lorsqu'une bonne compréhension des altérations pouvant être ciblées par thérapies dans un certain type tumoral est connue et lorsque la fréquence de ces altérations est élevée.
- **Les essais cliniques de type “panier” (“basket”) chez des patients résistants ou en rechute avec différents diagnostics** : testent un seul traitement dans un grand nombre de types tumoraux présentant toutes une ou plusieurs altérations ciblées par ce traitement. Ils permettent d'évaluer l'efficacité d'un même traitement dans différents types de cancers.
- **Les essais à agent unique de thérapie ciblée chez des patients à des stades avancés** : lorsque les critères d'éligibilité comprennent un biomarqueur génomique, ces essais contribuent à notre compréhension de l'efficacité des agents ciblés dans les cancers pédiatriques. Ils sont généralement sponsorisés par des laboratoires pharmaceutiques et utilisent bien souvent des médicaments principalement développés chez l'adulte.
- **Essais cliniques de précision spécifiques d'un certain type de cancer chez des patients au diagnostic** : comme pour les essais « parapluie », ces essais nécessitent une compréhension des altérations génétiques retrouvées dans un type tumoral donné. De plus, il est nécessaire dans ces essais d'avoir des preuves de la relation entre les altérations génétiques, le pronostic et/ou l'activité de thérapies ciblées au sein du type tumoral traité.

	<p>« Umbrella » diagnostic spécifique Chez des patients en maladie progressive</p>	<p>NCI–COG Pediatric MATCH AcSe´-ESMART</p>
	<p>Basket Chez des patients résistants ou en rechute</p>	<p>Ruxolitinib ou Dasatinib avec chimiothérapie avec LAL Philadelphia-Like NEPENTHE (Neuroblastome)</p>
	<p>Agent unique de thérapie ciblée Chez des patients à des stades avancés</p>	<p>Larotrectinib : tumeurs avec fusions NTRK Tazemetostat : tumeurs INI-1 neg Crizotinib : tumeurs avec alterations de ALK, MET ou ROS1 Ceritinib : tumeurs avec activation de ALK Dabrafenib et Trametinib : tumeurs avec la mutation BRAF V600 Afatinib : tumeurs avec deregulation de la voie ErbB</p>
	<p>Spécifiques d'un type de cancer Chez des patients nouvellement diagnostiqués</p>	<p>Total Therapy study XVII: Mutations JAK/STAT dans les lymphomes et LAL Dasatinib en combinaison : LAL avec fusions ciblables par inhibiteurs de TK Combinaison thérapeutique + Dasatinib : LAL-B Philadelphia-Like Thérapie orientée en fonction des risques cliniques et moléculaires : médulloblastome BIOMEDE : Gliomes infiltrants du tronc cérébral</p>

Figure II B-1 2: Types et exemples d'essais cliniques de précision en oncologie pédiatrique.

Source: Forrest et. al, (116). NCI: National Cancer Institute ; COG: Children's Oncology Group ; MATCH: Molecular Analysis for Therapy Choice; AcSé-ESMART: Essai thérapeutique du programme lancé par l'Institut National du Cancer (INCa); ESMART: European Proof-of-Concept Therapeutic Stratification Trial of Molecular Anomalies in Relapsed or Refractory Tumors in Children; LAL: Leucémie aigüe lymphoblastique; NEPENTHE: Next Generation Personalized Neuroblastoma Therapy; TK: tyrosine kinase; LAL: leucémie aigüe lymphoblastique; BIOMEDE: Biological Medicine for Diffuse Intrinsic Pontine Glioma Eradication.

2. Cas de l'essai moléculaire de précision MOSCATO-01 conduit à Gustave Roussy

Les études de séquençage « clinique » ont avant tout démontré la faisabilité du séquençage sur un seul ou plusieurs sites et ouvrent la voie à d'autres essais cliniques ultérieurs conçus cette fois pour quantifier l'impact des thérapies ciblées en oncologie pédiatrique. A l'institut Gustave Roussy notamment, l'essai clinique prospectif monocentrique MOSCATO-01 (Molecular Screening for Cancer Treatment Optimization) incluant des patients pédiatriques a évalué le bénéfice clinique

de tester un large éventail de gènes chez des patients atteints de cancers «difficiles à traiter». Les patients étaient admissibles s'ils présentaient une tumeur locale à un stade avancé, non résecable ou métastatique et ayant progressé à la suite d'au moins une ligne de traitement de chimiothérapie et s'ils étaient considérés comme « incurables » par un conseil médical multidisciplinaire désigné. Les patients présentant une altération génomique les rendant éligibles à un traitement ciblé approuvé dans la pathologie associée ne pouvaient être inclus dans cet essai. Après consentement éclairé (numéro de l'essai clinique : NCT01566019), les acides nucléiques ont été extraits à partir des biopsies tumorales fraîchement congelées et analysés par hybridation génomique comparative (aCGH) et séquençage haut débit (WES, RNA-seq et séquençage ciblé). Après ces étapes, le rapport d'analyse génomique est discuté pour chaque patient au cours d'un conseil hebdomadaire multidisciplinaire, et une thérapie adaptée est attribuée en conséquence (**Figure II B-2 1**). La capacité d'action d'une drogue sur une cible a été définie par ce conseil et selon la disponibilité de médicaments pouvant atteindre directement ou indirectement cette cible. Les patients auxquels une thérapie ciblée a été proposée et l'ayant reçue ont constitué la population où l'efficacité de la méthode globale a été évaluée. Au total, 1035 adultes ont été inclus et cette étude a démontré que le séquençage des tumeurs améliorait les résultats chez 33% des patients atteints de cancers avancés (14).

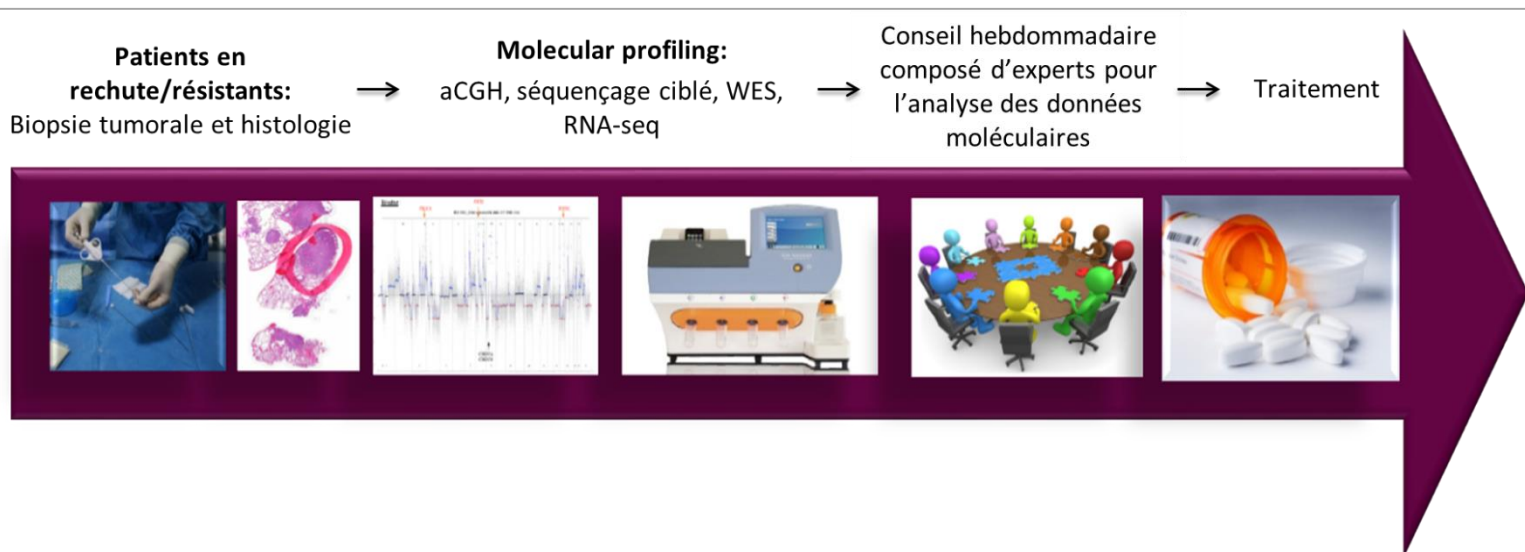


Figure II B-2 1: Processus de prise en charge des patients inclus dans l'essai clinique MOSCATO-01 mené à Gustave Roussy, de la biopsie tumorale en passant par l'analyse génomique jusqu'à l'attribution d'un traitement ciblé adapté.

L'essai a été adapté en 2012 pour inclure des patients pédiatriques à partir de 6 mois. Pour le groupe des adultes, l'objectif principal était d'évaluer le bénéfice alors que pour la cohorte pédiatrique, l'étude avait un objectif plus descriptif et de preuve de concept, étant donné la faible taille de l'échantillon et l'hétérogénéité des diagnostics. Soixante-treize patients pédiatriques ont

été biopsiés dont 69 ayant eu au moins une analyse moléculaire. Des altérations potentiellement ciblables ont été identifiées chez 60,9% des patients et ont même conduit à un changement de diagnostic chez trois patients (15).

Les données générées chez les patients pédiatriques inclus dans cet essai clinique ont servi de base à mon projet de thèse, me permettant de travailler sur les données RNA-seq (réalisées chez 48 patients) pour la détection de fusions.

3. Limites de la médecine de précision en pédiatrie

a) Limites liées aux cohortes pédiatriques et aux essais cliniques en cours

En pédiatrie, en dépit de l'augmentation de la mise en place d'essais de thérapies ciblées basés sur l'analyse moléculaire de chaque patient indépendamment, la rareté des tumeurs et par conséquent, la difficulté à rassembler des cohortes d'études ayant assez de puissance statistique constitue l'une des difficultés principales empêchant l'avancée de la connaissance de ces pathologies et l'amélioration des taux de survie (26,98). Les patients atteints de ces cancers rares sont, comparativement à certaines tumeurs très répandues de l'adulte, désavantagés en raison du défi de la conduite d'essais cliniques dans ce cadre. Cependant, il est attendu que le nombre de patients atteints de cancers rares augmente, à mesure que l'identification de sous-ensembles moléculaires de cancers augmente. Ceci nécessite alors un changement dans la conduite des essais cliniques et de la recherche sur ces types de cancer (26). De plus et de manière plus globale, la fréquence des gènes mutés identifiés comme étant liés au cancer est souvent inférieure à 10%, ainsi la taille de l'échantillon requis pour les essais en génomique doit être suffisamment grande pour permettre d'évaluer l'effet d'un traitement ciblé. En outre, ces problèmes peuvent être surmontés par la mise en place d'études multicentriques collaboratives, au niveau international (98).

D'autre part, la majorité des données génomique en cancérologie disponibles sont basées sur des tumeurs primaires plutôt que sur des cancers métastatiques ou avancés. Ces derniers peuvent pourtant fournir des informations précieuses étant donné qu'ils acquièrent généralement des mutations supplémentaires responsables de la progression tumorale, de résistances et/ou de l'agressivité (117). Ces tumeurs en pédiatrie sont considérées comme des maladies « ultra-orphelines » exceptionnellement rares et nécessitent également la mise en place d'études collaboratives. A Gustave Roussy, l'ouverture récente de l'essai clinique de précision MAPPYACTS (MoleculAr Profiling for Pediatric and Young Adult Cancer Treatment Stratification,

NCT02813135), multicentrique et international incluant uniquement des patients pédiatriques en rechute ou résistants contribue aux efforts mondiaux pour mieux comprendre les cancers de l'enfant et permettre d'augmenter les taux de guérison.

b) Limites de l'utilisation des thérapies ciblées pour le traitement des cancers pédiatriques en Europe

Jusqu'à présent, la recherche de cibles dans les tumeurs adultes a concentré la plupart des efforts pharmaceutiques. Les patients atteints de cancers pédiatriques bénéficient rarement de ces découvertes en raison du faible recoupement des anomalies génétiques survenant dans les tumeurs adultes. De plus, comme évoqué précédemment, les cancers pédiatriques comportent moins de mutations et par conséquent moins de cibles potentielles associées (118). En réalité et de manière générale, la majorité des altérations génomiques ont une fonction encore inconnue, les empêchant donc d'être ciblées (98). Le recueil de données génomiques ainsi que des études fonctionnelles des altérations retrouvées dans les cas de cancers pédiatriques (au diagnostic et à la rechute) sont donc nécessaires pour permettre d'augmenter le répertoire d'anomalies pouvant être ciblées (118).

En plus d'un manque de connaissance des altérations responsables des cancers pédiatriques, il y a actuellement relativement peu de thérapies ciblées disponibles pour les enfants (51,118). L'accès à de nouveaux médicaments reste une des limites majeures. En effet, pour des raisons éthiques il est généralement admis que les nouveaux médicaments doivent d'abord subir des essais chez l'adulte avant d'être administrés aux enfants.

Un des inconvénients de la thérapie ciblée, est l'acquisition rapide de résistances (119). Les thérapies ciblées sont plus efficaces lorsqu'elles sont combinées à : 1) d'autres thérapies ciblées 2) des molécules antinéoplasiques, 3) de la radiothérapie et/ou 4) de l'immunothérapie, mais la manière optimale de combiner ces traitements n'est pas toujours claire. D'autres tests sont nécessaires pour cibler une anomalie moléculaire (118).

c) Evolution clonale et hétérogénéité intratumorale

L'approche personnalisée est basée sur l'observation que le cancer est une maladie extrêmement hétérogène, avec des différences importantes inter-diagnostic, mais aussi interindividuelles pour un même diagnostic. A cela se rajoute l'hétérogénéité intra-tumorale (117) ; en effet, une tumeur est composée de plusieurs sous-clones, divergents dans leur répertoire d'altérations moléculaires et conduisant à différentes proportions d'allèles mutées dans un même échantillon de tumeur

(Figure II B-3 c. 1). Les techniques actuelles de séquençage étant encore majoritairement incapables d'individualiser chaque sous-clone, de nombreuses informations moléculaires sont absentes des analyses actuelles.

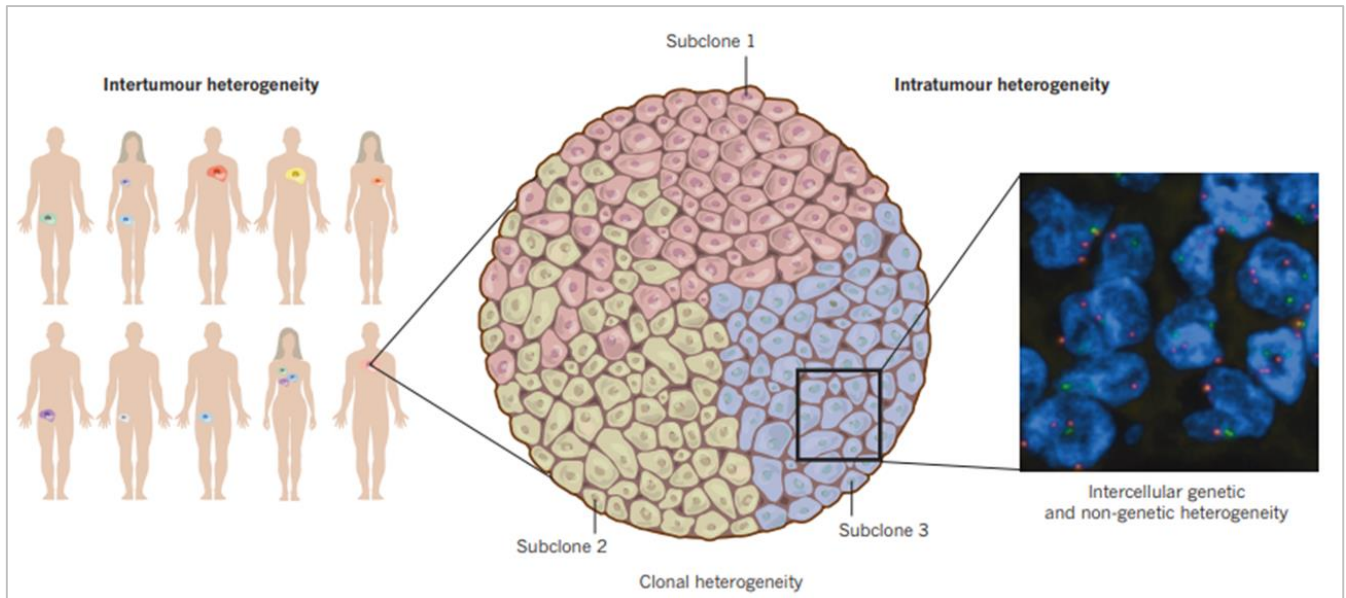


Figure II B-3 c. 1: Hétérogénéité inter- et intra-tumorale. Des variations génétiques et phénotypiques sont observées entre les tumeurs issues de différents tissus et cellules, ainsi qu'entre des individus ayant le même type de tumeur (hétérogénéité intertumorale). Au sein d'une tumeur, il existe une diversité sous-clonale (hétérogénéité intratumorale). Les sous-clones tumoraux peuvent présenter une hétérogénéité génétique et épigénétique. L'hétérogénéité génétique intercellulaire est exacerbée par l'instabilité génomique et peut favoriser l'émergence de sous-clones tumoraux. L'architecture sous-clonale des tumeurs peuvent varier avec le temps si elles sont soumises à certains facteurs comme par exemple un traitement. Source: Burrell et. al, (120).

Cette évolution clonale de diverses populations au cours de la progression tumorale contribue à l'hétérogénéité et est responsable de l'échec des traitements chez certains patients décrits comme résistants. En effet, les sous-clones résistants à certaines molécules présents initialement dans la tumeur ou apparaissant au cours des thérapies ciblées représentent l'un des défis majeurs pour le succès des thérapies personnalisées (Figure II B-3 c. 2) (101). Les nouvelles techniques de séquençage à l'échelle de cellules uniques sont actuellement en développement et permettront de mieux comprendre cette architecture clonale chez les patients atteints de pathologie cancéreuse (121).

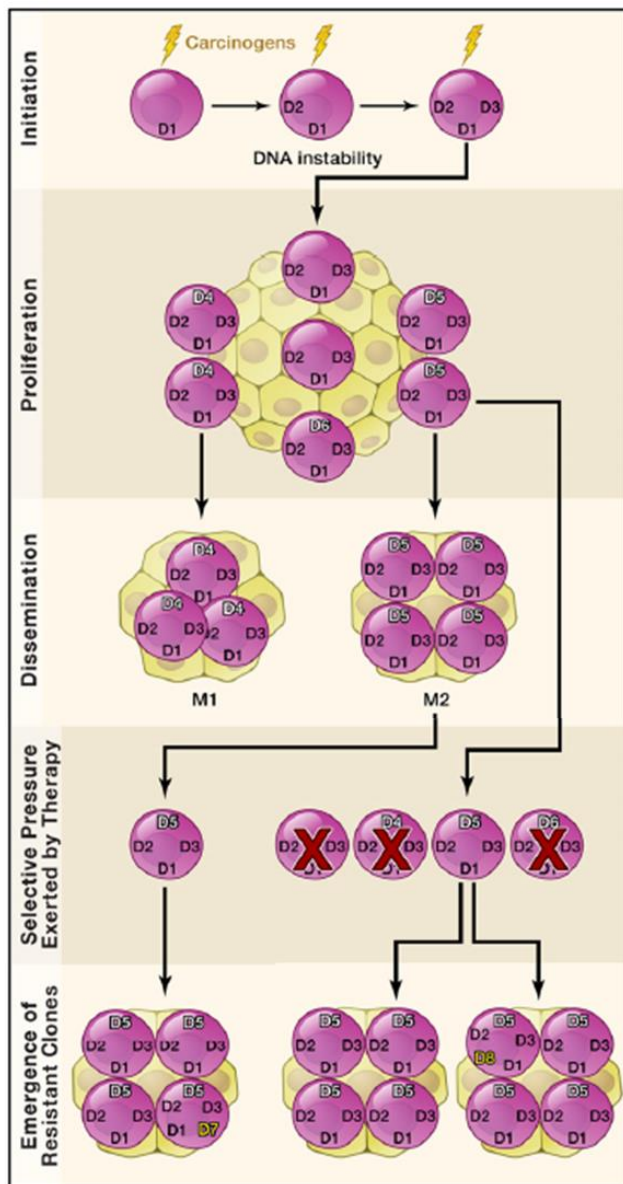


Figure II B-3 c. 2: Accumulation de mutations « drivers » au cours de l'évolution et de l'histoire d'une tumeur. L'accumulation de mutations conduit à l'apparition de la tumeur primaire, puis à la croissance et aux métastases. Les sous-clones avec de nouvelles mutations peuvent devenir dominants au sein des métastases ou des tumeurs résistantes ou en rechute, dû à des pressions sélectives exercées par des chimiothérapies cytotoxiques ou ciblées. Source : Dancy et al, (98).

d) *Limites liées au NGS*

De manière paradoxale, l'une des principales limites du NGS constitue également son atout puisqu'elle réside dans la quantité de données produites : à l'échelle de mégabases (millions) à gigabases (milliards), dont l'interprétation est une tâche fastidieuse et dont le stockage reste un paramètre limitant (108). De plus, chaque étape de génération des données nécessite un suivi des données, un contrôle qualité et des approches analytiques complexes qui reposent sur du personnel spécialisé.

Des limites liées aux techniques de séquençage haut-débit actuellement utilisées peuvent également être pointées comme : 1) la difficulté à détecter les variants structuraux de grande taille avec des lectures (« reads ») de seulement quelques centaines de paires de bases, 2) l'alignement

sur des génomes connus (toujours en raison de la petite taille des lectures) ou encore 3) la difficulté à séquencer les zones avec des séquences répétées (107).

C. Conclusion de la partie II

L'accumulation d'analyses de séquençage et de données issues des essais cliniques guidés par la génétique moléculaire, ainsi que les résultats des études biologiques associées et le développement continu de médicaments permettent d'approfondir les découvertes et d'élargir les possibilités de traitement ciblés pour les enfants atteints de cancer (116). Les défis principaux pour l'approche de médecine personnalisée/de précision en pédiatrie comprennent la sélection de cibles optimales, l'analyse des données de NGS, la compréhension des profils génétiques, la détermination de combinaisons appropriées de thérapies et la résolution des défis organisationnels, commerciaux, réglementaires et sociétaux (101).

Cette médecine de précision est en constante évolution et dans le futur, de nouvelles approches seront attendues. Parmi elles, les biopsies tumorales liquides permettant d'identifier l'ADN provenant de cellules cancéreuses dans le sang seront optimisées grâce à la diminution des coûts de séquençage et du développement du séquençage de cellule unique. Ces biopsies, en plus d'être non-invasives, permettent un suivi fréquent de l'évolution tumorale pendant et après traitement pour détecter l'apparition de résistances (118,122). La gamme d'altérations génomiques analysées ne sera certainement plus restreinte à l'analyse du génome mais aussi au séquençage des altérations épigénétiques (épigénomique), des néo-antigènes tumoraux et du répertoire des lymphocytes T, des petits ARNs et micro-ARNs mais également à l'application de la métabolomique pour comprendre et prévoir rationnellement les réponses aux traitements et les profils de toxicité (**Figure II C**) (118).

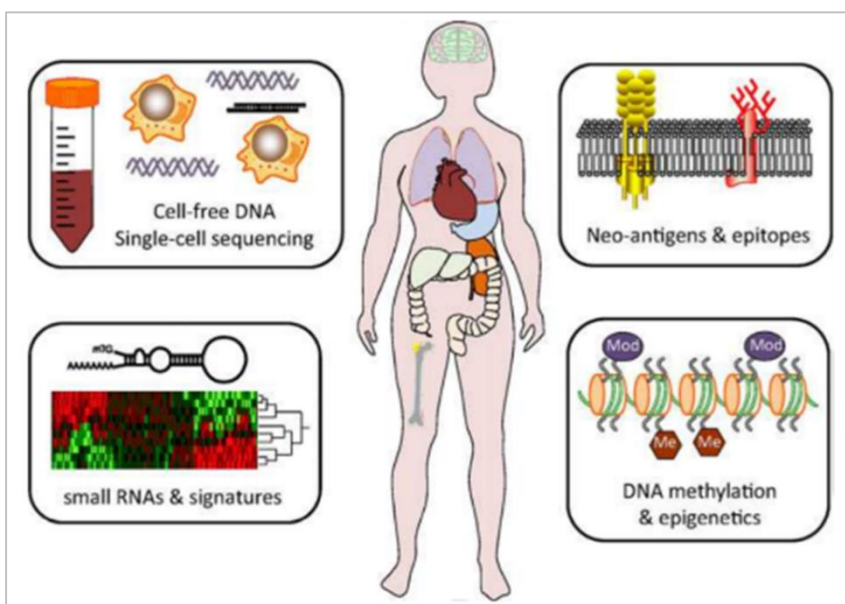


Figure II C : Orientations futures en médecine de précision. De nouvelles études pourraient définir les rôles cliniques de plusieurs nouveaux domaines de la médecine de précision, dont quatre comprenant: l'épigénomique, le profilage des petits ARNs, les néoantigènes et le répertoire d'épitopes ainsi que le séquençage à cellules uniques et l'ADN circulant. Source: Mody et. al, (118)

III. Transcrits de fusions et cancers

A. Généralités

La capacité à établir le portrait génomique complet d'un patient représente l'une des conditions préalable indispensable pour faire de la médecine personnalisée une réalité. En cancérologie, de nombreuses études ont porté sur la recherche des mutations, des insertions, des délétions et des variations du nombre de copies (11,123–126). Bien que ces altérations constituent une grande partie des modifications génétiques tumorales, les gènes de fusions jouent également un rôle essentiel dans l'oncogenèse (127).

Les progrès des technologies NGS ont permis une meilleure compréhension de ces événements, au regard de leur prévalence et leur rôle et un nombre croissant de nouvelles fusions sont découvertes d'années en années. Cependant, leur utilisation en clinique reste principalement limitée par le manque d'études fonctionnelles, afin de déterminer les conséquences biologiques à la fois dans un contexte cancéreux mais également dans des cellules normales (12,128).

Aussi appelés « transcrits de fusion », « oncogènes de fusion » ou encore « oncogènes de jonction », ils peuvent être provoqués par des événements au niveau génomique (translocation, inversion, insertion, délétion) ou transcriptionnel (« readthrough » ou événements d'épissage) (12,129).

Au niveau génomique, les points de cassures peuvent aussi bien se situer dans la partie 5'UTR, la région codante (CDS), une région non-codante ou dans la partie 3'UTR (**Figure III-A 1**) et ainsi avoir différentes conséquences : la formation d'une oncoprotéine chimérique, une surexpression aberrante (généralement d'un oncogène) ou la disruption d'un gène (généralement un suppresseur de tumeur) (**Figure III-A 2**).

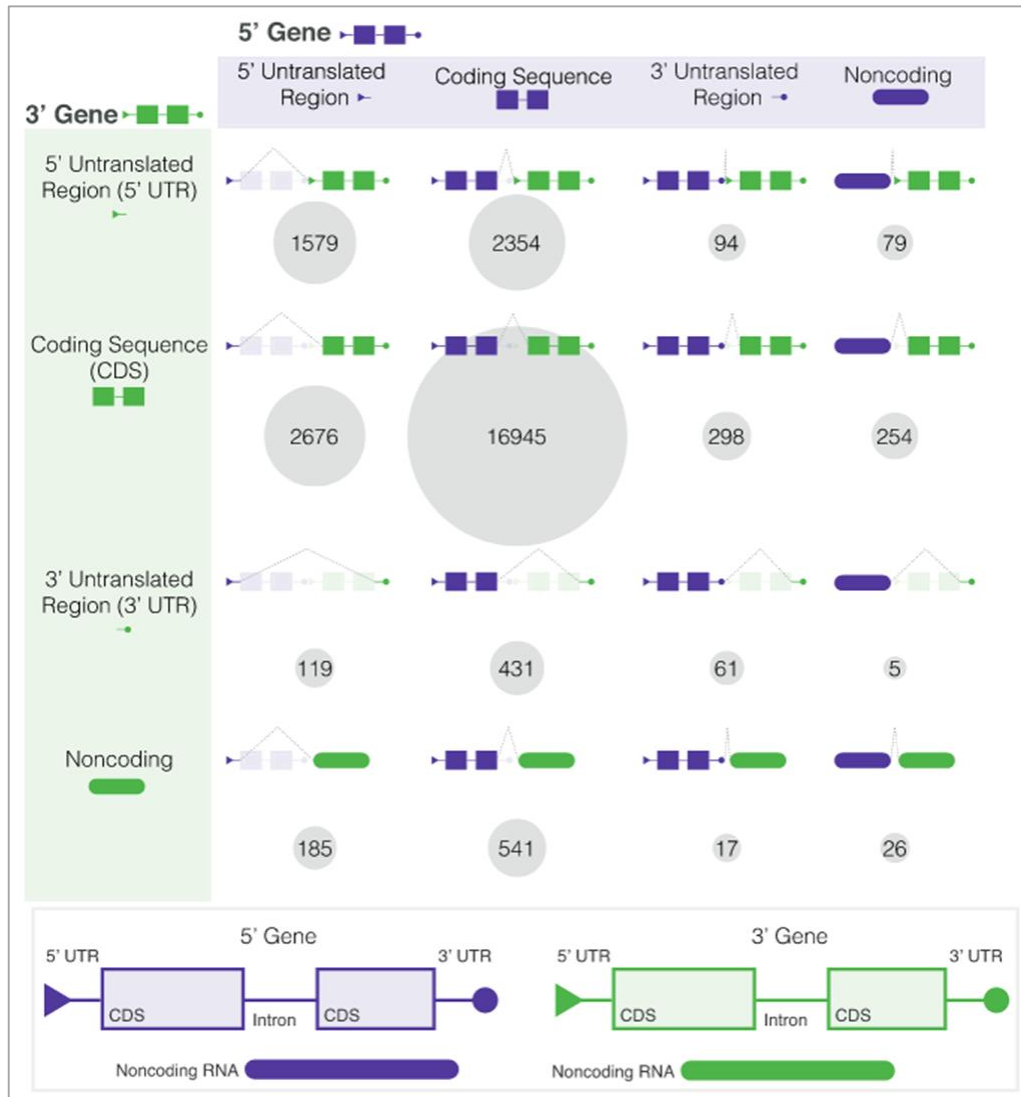


Figure III-A 1 : Schéma des points de cassure des partenaires impliqués dans une fusion de deux gènes. Le violet indique le partenaire 5', et le vert indique le partenaire 3'. Les points de cassure peuvent se produire dans les régions génomiques suivantes: 5'UTR (triangle), séquence codante (CDS, rectangle), 3'UTR (cercle) et région non codante (bâtonnets arrondis). Pour chaque événement de fusion, une ligne pointillée relie les points de rupture dans les partenaires 5' et 3' pour illustrer la fusion prédite. La taille du cercle et les nombres (en gris clair) représentent le nombre total de fusions détectées dans la chaque catégorie. Source : Gao et. al, (130).

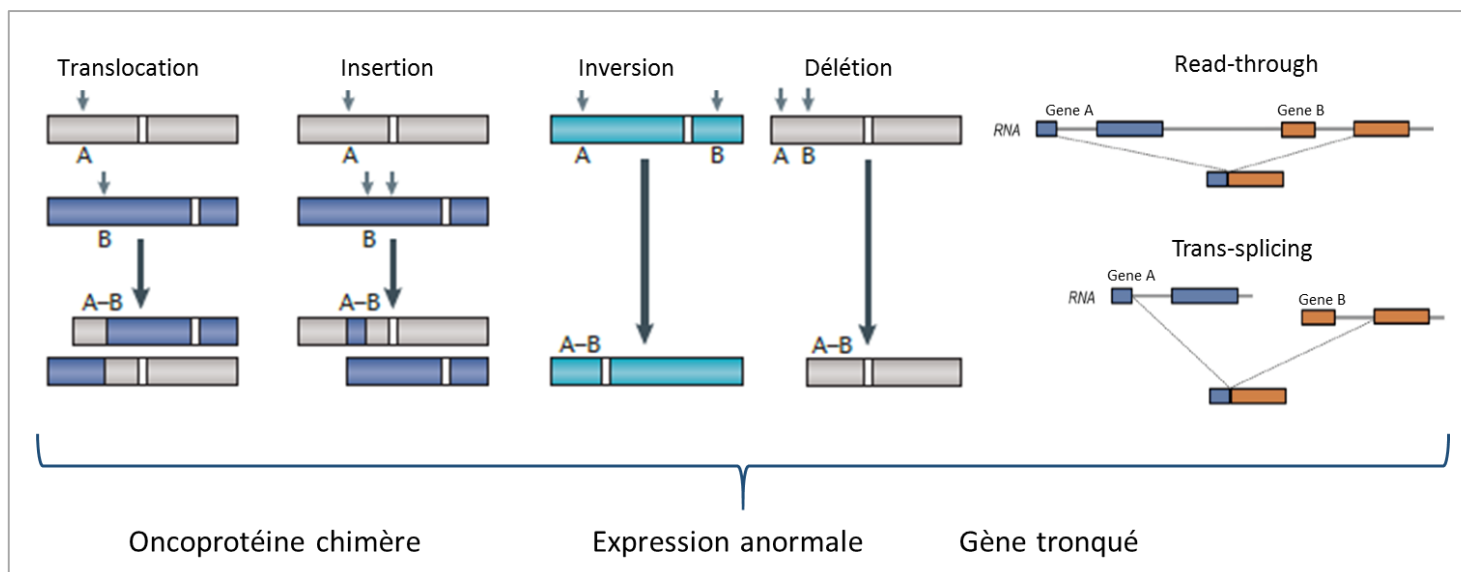


Figure III-A 2 : Evènements génomiques et transcriptionnels responsables de la formation de transcrits de fusions et conséquences fonctionnelles possibles. Les « read-through » sont provoqués suite à de défauts dans la terminaison de la transcription, conduisant à une lecture au-delà de la limite du gène 3'. Adapté d'après Mertens et. al, Zhang et. al et Kuehner et. al (12,129,131).

L'étude des fusions spécifiques aux cellules cancéreuses peut contribuer à la prise en charge en clinique de plusieurs manières. Premièrement, elles peuvent être utilisées comme cibles thérapeutiques dans différents types de cancer lorsqu'elles sont spécifiques aux cellules cancéreuses et responsables de la progression tumorale. Quelques exemples comprennent les fusions impliquant les gènes *ALK*, *RET*, *BRAF*, *RAF1*, *FGFR1-4*, *TFE3*, *NTRK* et *NOTCH1-3* (123). Deuxièmement, elles peuvent également servir de biomarqueurs diagnostiques ou pronostiques comme c'est le cas pour *EWS-FLI1* dans le sarcome d'Ewing, *PAX3/7-FOXO1* dans le rhabdomyosarcome alvéolaire, *DNAJB1-PRKCA* dans le carcinome fibrolamellaire hépatocellulaire (FLHCC) ou *TMPRSS2-ERG* dans le cancer de la prostate (9,123). Certaines fusions peuvent être à la fois des cibles thérapeutiques et des biomarqueurs diagnostiques, comme *BCR-ABL* retrouvée dans la leucémie myéloïde chronique (LMC) et ciblée par l'imatinib mesylate (2) ou *PML-RAR α* retrouvée chez les patients atteints de leucémie aiguë promyélocytaire et ciblée par un traitement à l'acide rétinoïque (132).

Les fusions peuvent également contribuer à classer les patients en sous-groupes moléculaires et à mieux comprendre la biologie de malignités rares telles que *C11orf95-RELA* retrouvée dans l'épendymome supratentorial, *DNAJB1-PRKCA* dans le FLHCC ou *HEY1-NCOA2* dans le chondrosarcome mésoenchymateux (133–135). De plus, le suivi de la maladie résiduelle minimale par quantification des transcrits de fusion associés au cancer chez les patients peut fournir une indication de la réponse au traitement comme c'est le cas pour *BCR-ABL* ou *NPM-ALK* dans le

lymphome anaplasique à grandes cellules (136,137). Enfin, et de manière intéressante, très peu d'études ont établi une relation entre l'expression des transcrits de fusion et la résistance aux traitements anticancéreux, évoquant la nécessité d'explorer cette composante. Par exemple, la fusion TMPRSS2-ERG conduit à une augmentation de l'efficacité de réparation de l'ADN et confère une résistance des cellules cancéreuses de la prostate à la radiothérapie (138). Dans les leucémies, les fusions BCR-ABL, ETV6-RUNX1 ou impliquant MLL ont, de leur côté, été associées à une résistance à l'apoptose, suggérant un impact potentiel sur la sensibilité aux traitements (139). Plus récemment, Hartmaier et. *al* ont révélé que les fusions avec le gène *ESR1* étaient prépondérantes dans les cancers du sein métastatiques ER-positifs, et pouvaient être à l'origine des résistances aux traitements endocriniens (140).

B. Méthodes de détection des transcrits de fusion

Les nouvelles méthodes de séquençage haut-débit ont largement contribué à l'identification de nouvelles fusions. Comme décrit dans la partie II-A 2, les gènes ou transcrits de fusions peuvent être détectés à partir des données de séquençage du génome entier (WGS) ou de l'ARN (RNA-seq).

L'utilisation des données de RNA-seq comporte cependant plusieurs avantages et est, de ce fait la technique la plus employée pour la recherche de fusions (141). Ceci est dû :

- à son faible coût par rapport à la taille des données générées et du temps d'analyses moindres comparativement au WGS,
- à une détection au niveau transcriptionnel mettant en évidence la transcription du gène de fusion,
- à la capacité de détecter des transcrits de fusion dus à des événements transcriptionnels tels que les « readthrough » et les événements d'épissage.

L'analyse de la présence de fusions à partir de données de WGS ou RNA-seq est réalisée à l'aide de logiciels bioinformatiques spécialisés (141). A ce jour, il existe des dizaines d'outils bioinformatiques pour la détection de fusion, chacun différant en termes de sensibilité et de spécificité, en fonction de leurs algorithmes et des différents filtres appliqués. Plus précisément, ces logiciels diffèrent selon le format de reads pris en compte (non-apparié « single-end » ou apparié « paired-end »), le type de séquençage, le génome de référence sur lesquels l'alignement est réalisé, ou encore selon les critères de sélection d'une fusion (distance minimale entre deux « reads » d'une paire, nombre de nucléotides séquencé de chaque côté d'un point de fusion, présence de « read-through » ou encore nombre de reads correspondant à des régions homologues ou répétées) (142). Les principaux outils pouvant être utilisés et leurs caractéristiques sont récapitulés dans le **Tableau III-B 1**.

Logiciels	Format de reads		Type de séquençage		Référence		Filtre sur le nombre de nucléotides séquencés de chaque côté du point de fusion	Filtre les transcrits "read-through"	Filtre le nombre de reads minimum supportant une fusion	Filtre les produits de PCR lors de la génération des librairies	Filtre les reads correspondants à des séquences homologues	Outil bioinformatique utilisé pour la première étape d'alignement
	Non apparié	Apparié	WGS	RNA-seq	Génome	Transcriptome						
Bellerophon		✓		✓	✓		N	O	O	O	O	TopHat
BreakFusion		✓		✓	✓	✓	N	N	N	N	N	BWA/BLAT
ChimeraScan		✓		✓	✓	✓	10	O	4	N	N	Bowtie/BWA
EricScript		✓		✓	✓	✓	N	O	3/1	O	O	BWA/BLAT
FusionAnalyser		✓		✓	✓	✓	O	O	O	N	O	BWA
FusionCatcher	✓	✓		✓	✓	✓	10	O	3/1	N	O	Bowtie/STAR/BLAT/Bowtie2
FusionFinder	✓	✓		✓	✓	✓	N	O	N	N	O	Bowtie
FusionHunter	✓	✓		✓	✓	✓	10	O	3/1	O	O	Bowtie
FusionMap	✓	✓	✓	✓	✓	✓	O	O	O	O	O	GSNAP
FusionQ	✓	✓		✓	✓	✓	10	N	3/1	N	O	Bowtie
FusionSeq	✓	✓		✓	✓	✓	N	O	O	O	O	ELAND
JAFFA	✓	✓		✓	✓	✓	N	O	3/1	N	O	Bowtie/BLAT
MapSplice	✓	✓		✓	✓	✓	N	N	N	N	N	Bowtie
DeFuse	✓	✓		✓	✓	✓	10	O	3/1	N	O	Bowtie/BLAT
SOAPFuse	✓	✓		✓	✓	✓	10	N	3/1	N	N	Soap2/BWA/BLAT
TopHat-Fusion	✓	✓		✓	✓	✓	10	O	3/1	N	O	Bowtie
PRADA	✓	✓		✓	✓	✓	N	N	N	N	N	BWA/BLAT
ShortFuse		✓		✓	✓	✓	N	N	O	N	N	Bowtie
SnowShoes-FTD		✓		✓	✓	✓	N	O	2/N	O	O	Bowtie/BWA

Tableau III-B 1 : Caractéristiques des principaux logiciels de détection de fusions et paramètres par défaut. D'après Kumar et al, (141).

Quels que soient les outils utilisés, la méthodologie reste identique et se compose de trois étapes:

1. Alignement des reads et filtrage

Initialement, quelques études ont évaluées la présence de fusions en utilisant un séquençage non-apparié, cependant aujourd'hui le séquençage apparié (« paired-end »), plus sensible est majoritairement utilisé.

Un « read » contenant une séquence de jonction est appelé « split-read » ou « junction read » et ne s'aligne pas directement sur le génome de référence (**Figure III-B 1, encadré en violet**). Lorsque deux « reads » d'une même paire s'alignent sur deux gènes différents sur le génome de référence mais la séquence de jonction n'est pas contenue dans un des deux « reads », il s'agit d'une paire de « spanning reads » (**Figure III-B 1, encadré en orange**). L'analyse à la fois des « split reads » et des « spanning reads » permet l'identification de fusions (**Figure III-B 1**).

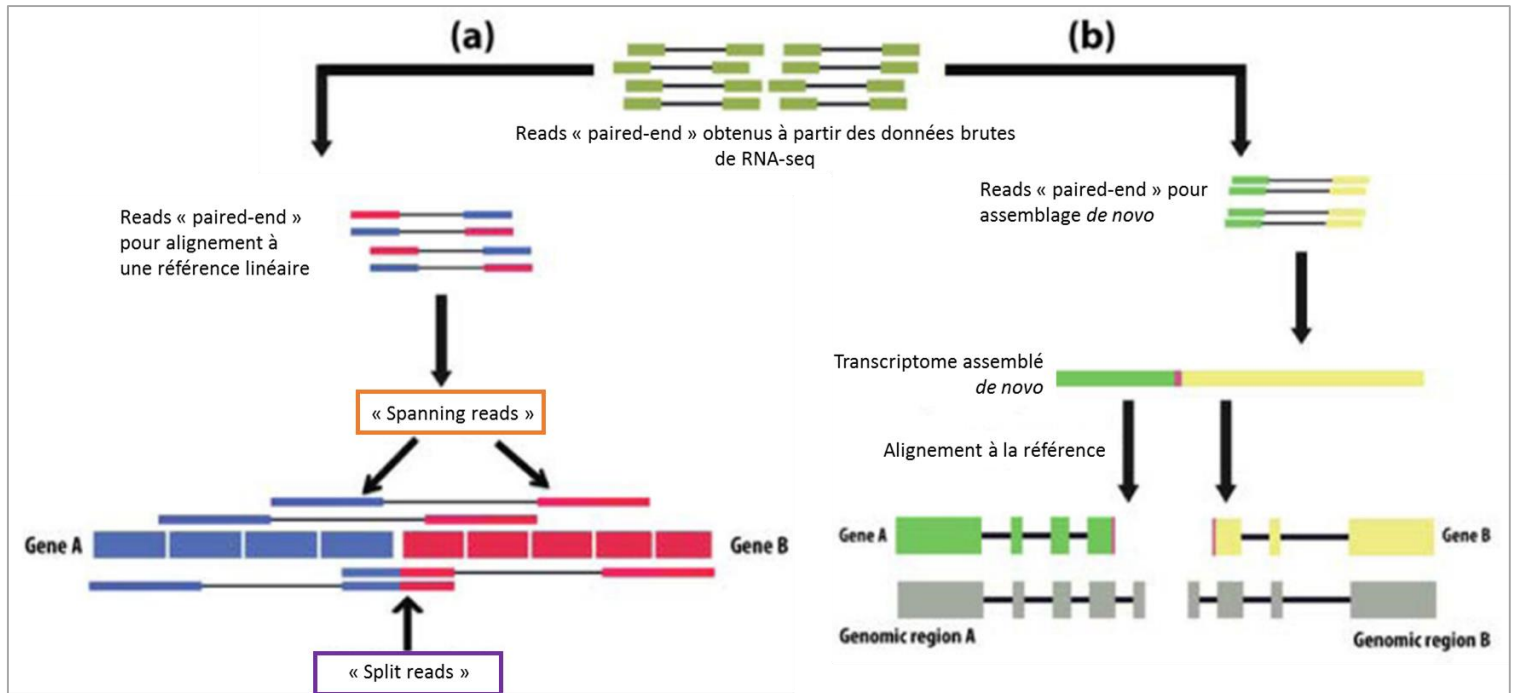


Figure III-B 1: Deux approches pour les étapes initiales de la détection de fusions. Les fragments d'ADN de la librairie sont représentés par des bâtonnets, ayant une partie non colorée correspondant à la partie non séquencée et une partie colorée (vert, rouge, bleu ou jaune) correspondant à la partie séquencée (les « reads »). (a) Approche de « l'alignement en premier ». Les lectures appariées sont alignées en premier pour déterminer le point de jonction grâce aux « spanning reads » et « split reads ». (b) Approche de « l'assemblage en premier », les lectures appariées sont d'abord assemblées de novo pour obtenir un transcriptome ayant une jonction de fusion qui est ensuite aligné sur des séquences de référence. L'approche (a) est la plus rapide et la plus couramment utilisée. Adapté d'après Kumar et. al, (141)

2. Détection du point de jonction de la fusion

Après l'alignement, les reads non-pertinents pour la détection de fusions sont éliminés de l'analyse. En revanche, les reads ne correspondant pas à des régions connues du génome (représentant des potentiels "split reads") sont mémorisés [Figure III-B 2 (ii)]. Les locus de jonction et les transcrits de fusion putatifs sont ensuite prédits. Les reads qui ne correspondaient à aucune région sont alors alignés sur les transcrits de fusion prédits.

3. Assemblage et sélection du transcrit de fusion

Après avoir identifié les points de jonction des fusions, les séquences sont générées en joignant celles correspondant aux deux gènes partenaires [Figure III-B 2 (iii)]. Les fusions candidates appuyées par le plus de « reads » peuvent représenter les candidats les plus probables, cependant en éliminant les fusions ayant peu de « reads », le risque de rejeter des fusions réelles peu exprimées augmente.

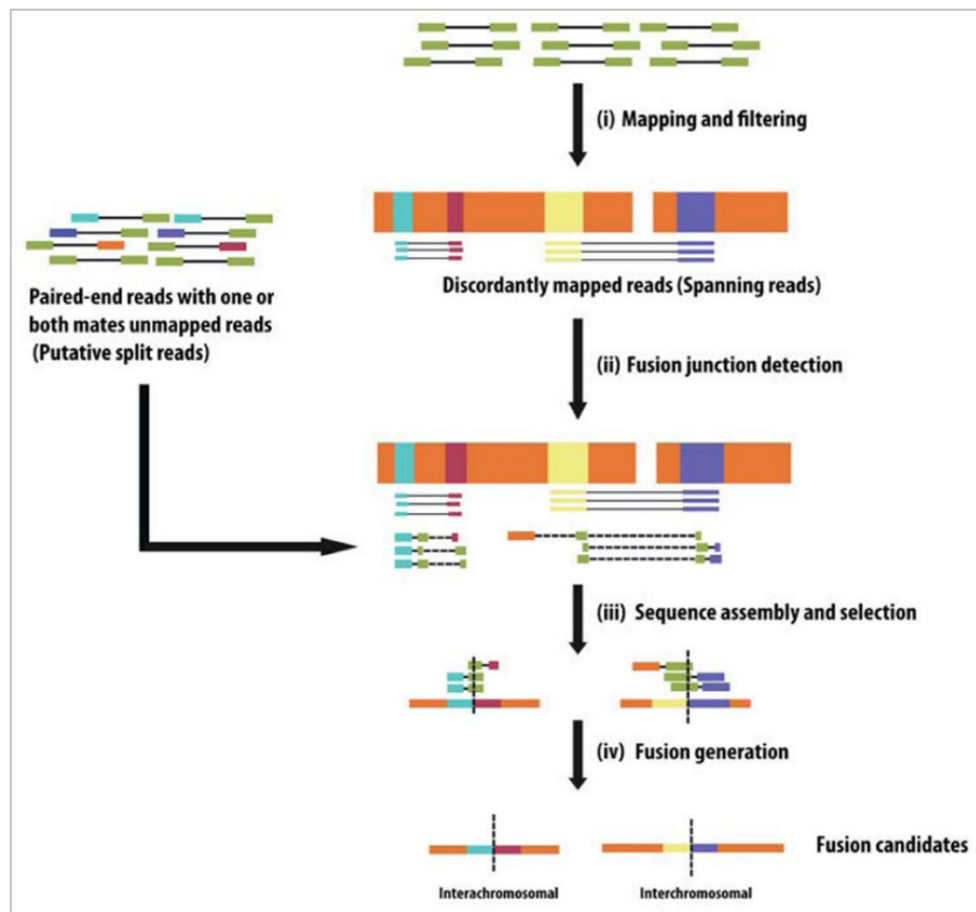


Figure III-B 1 : Etapes de détection de fusions à partir de donnée de séquençage haut-débit. Dans cette figure également, les fragments d'ADN de la librairie sont représentés par des bâtonnets, ayant une partie non colorée correspondant à la partie non séquencée et une partie colorée (vert, bleu, violet ou jaune) correspondant à la partie séquencée (les « reads »). (i) les paires de reads sont alignées sur des séquences de référence et des reads discordants sont directement alignés sur les gènes de fusions potentiellement détectés. (ii) Pour les paires dont au

moins l'un des deux reads ne s'aligne pas à la séquence de référence (potentiels « split reads »), la paire non alignée est découpée virtuellement en segments à aligner avec les fusions potentielles estimées. Ici, les deux morceaux d'un même read sont reliés par une ligne pointillée. (iii) Les séquences des fusions candidates sont assemblées. (iv) Séquences de fusion avec la probabilité la plus élevée d'être des fusions réelles. La ligne pointillée verticale représente le point de jonction de la fusion.

Le séquençage de l'ARN permet de détecter de nouvelles fusions non découvertes à grande échelle mais cependant, comme en témoigne l'augmentation spectaculaire des fusions de gènes reportées depuis son introduction, ces analyses comprennent également des événements «passagers» (secondaires et non responsables de la progression tumorale). Ces événements passagers sont présents en raison de la grande instabilité génomique des cancers et du caractère non-récurrent (dans des patients uniques) de la large majorité des fusions détectées et dont la signification est actuellement peu claire (12). Une approche utilisant plusieurs logiciels simultanément (méthode appelée de « meta-caller ») a été montrée comme permettant de combiner ces différents paramètres et ainsi d'augmenter la sensibilité et la spécificité (142). Cette approche a été utilisée lors de mes travaux de thèse où trois logiciels ont été combinés pour la détection de fusions à partir de données de RNA-seq.

C. Article de synthèse décrivant les gènes de fusion et les cancers pédiatriques

Revue « Relevance of fusion genes in pediatric cancers: toward precision medicine »

Dans cet article, nous décrivons d'une manière assez exhaustive les gènes de fusion connus et retrouvés dans les cancers pédiatriques ainsi que l'importance de leur détection dans le développement de la cancérogenèse et dans le traitement de ces cancers.

Relevance of Fusion Genes in Pediatric Cancers: Toward Precision Medicine

Célia Dupain,¹ Anne Catherine Harttrampf,¹ Giorgia Urbinati,¹ Birgit Georger,¹ and Liliane Massaad-Massade¹

¹Vectorology and Anticancer Therapies, UMR 8203 CNRS, University Paris-Sud, Gustave Roussy, Université Paris-Saclay, 94805 Villejuif, France

Pediatric cancers differ from adult tumors, especially by their very low mutational rate. Therefore, their etiology could be explained in part by other oncogenic mechanisms such as chromosomal rearrangements, supporting the possible implication of fusion genes in the development of pediatric cancers. Fusion genes result from chromosomal rearrangements leading to the juxtaposition of two genes. Consequently, an abnormal activation of one or both genes is observed. The detection of fusion genes has generated great interest in basic cancer research and in the clinical setting, since these genes can lead to better comprehension of the biological mechanisms of tumorigenesis and they can also be used as therapeutic targets and diagnostic or prognostic biomarkers. In this review, we discuss the molecular mechanisms of fusion genes and their particularities in pediatric cancers, as well as their relevance in murine models and in the clinical setting. We also point out the difficulties encountered in the discovery of fusion genes. Finally, we discuss future perspectives and priorities for finding new innovative therapies in childhood cancer.

Pediatric cancers represent 1% of all cancers¹ and comprise cancer cases diagnosed in children younger than 14 years² and adolescents and young adults aged 15–19 years.³ Pediatric cancers include more than 60 different types of cancer derived from different tissues. More than 250,000 cases are diagnosed worldwide each year among children and adolescents younger than 20 years (<http://www.childhoodcancerinternational.org>). Although 80% of pediatric patients reach long-term remission after treatment, 20% die from recurrence of the malignancy; therefore, more therapeutic targets are needed to improve survival rates.¹ Moreover, current treatments are based on multimodal aggressive chemotherapies, and it is well known that antineoplastic treatments have long-term side effects that could impact patients' quality of life.⁴

Pediatric cancers differ from adult cancers. In fact, most pediatric cancers arise from embryonal rather than epithelial cells;⁵ consequently, the etiology of pediatric cancers is different. Zhang et al.⁶ previously reported that 8%–10% of pediatric cancers are associated with germline alterations leading to a cancer predisposition (Table 1). The involvement of environmental factors in childhood cancer etiology is still discussed, but there is clearly a lower implication of exogenous toxic effects in children than in adults (e.g., smoking, alcohol consumption, sun exposure, and overweight and sedentary lifestyle). Very few environmental factors have been related to pediatric cancers

to date: ionizing radiation and electromagnetic fields are some of the effects that remain a source of controversy, whereas some chemicals such as dioxin, trichloroethane, pesticides, solvents, metals, petroleum products, boron, and pollution have been associated with specific cancer types.^{7,8} Other factors such as epigenetics and immune system deregulation have also been identified as being responsible for tumorigenesis.^{9,10}

Vogelstein et al.¹¹ highlighted that pediatric cancers usually harbor fewer genetic mutations than adult cancers. They explained that the lower mutation rate found in pediatric cancers may be due to the embryonal origin of these cancers, which did not have enough time to renew. As a consequence, the tumors harbor very few mutations at the basal state.¹¹ These differences in mutation patterns could also be explained in part by the difference in cancer initiation, such as the implication of exogenous toxic effects.

A possible hypothesis of their etiology could be the presence of chromosome rearrangements, which is one of the first mechanisms described to be responsible for carcinogenesis.¹² These rearrangements can lead, in some cases, to fusion genes, which are the juxtaposition of two previously separate genes localized on the same (intra-chromosomal) or two different (inter-chromosomal) chromosomes. This event can activate proto-oncogenes or inactivate tumor suppressor genes. It has also been shown in adult cancers that some tumors with driving fusions have a much lower mutational burden compared to tumors without fusions.¹³ Thus, pediatric cancers appear to be the consequence of chromosomal rearrangements rather than mutation events.

We should keep in mind that a recurrence of alterations (mutations and copy-number alterations) of genes involved in embryogenesis¹⁴ and epigenetic regulation are also described in pediatric cancers⁹ but will not be discussed here.

In this review, we describe recent molecular knowledge on fusion genes in pediatric cancers. We then outline the potential therapeutic utility of fusion genes, their relevance in murine models, and we

Received 2 November 2016; accepted 24 January 2017;
<http://dx.doi.org/10.1016/j.omtn.2017.01.005>.

Correspondence: Liliane Massaad-Massade, Vectorology and Anticancer Therapies, UMR 8203 CNRS, Université Paris-Sud, Gustave Roussy, Université Paris-Saclay, 114 rue Edouard Vaillant, 94805 Villejuif, France.

E-mail: liliane.massade@gustaveroussy.fr



Table 1. Syndromes of Genetic Predisposition Found among Children with Cancer and the Corresponding Affected Genes

Genetic Predisposition	Affected Gene ^a
Li-Fraumeni syndrome	TP53/CHK2/SNF5
Familial Wilms tumor	FTW1/2
Familial retinoblastoma	RB1
Fanconi anemia	FANCA
IgA deficiency	IGAD1
Blood syndrome	BLM
Wiskott-Aldrich syndrome	WAS
Ataxia telangiectasia	ATM
Familial adenomatous polyposis	APC
Hereditary non-polyposis colon cancer	MSH2/MLH1/PMS2
Tuberous sclerosis	TSC1/2
Beckwith-Wiedemann syndrome	Complex
Xeroderma pigmentosum	ERCC2

^aChildren's Health and the Environment, WHO Training Package for the Health Sector (World Health Organization, <http://www.who.int/ceh>).

describe recent findings in the clinical setting as well as challenges associated with their discovery.

History of Fusion Genes

Fusion genes were primarily discovered in leukemia and other hematological diseases. In 1962, Nowell¹⁵ described the first specific chromosomal rearrangement in chronic myeloid leukemia (CML). In 1973, Rowley¹⁶ highlighted the existence of a reciprocal translocation between the long arms of chromosomes 9 and 22, t(9q; 22q), named the "Philadelphia chromosome." Twenty years later in the early 1980s, molecular studies of the translocation revealed a fusion between the 3' part of the *ABL1* gene in chromosome 9 and the 5' part of the *BCR1* gene in chromosome 22.¹⁷ This translocation resulting in BCR-ABL was one of the first specific alterations found in human neoplasm. Ten years later, a tyrosine kinase inhibitor (TKI) targeting BCR-ABL, imatinib mesylate (Glivec; Novartis), was discovered and approved by the Food and Drug Administration (FDA) in 2001 as a cancer treatment for CML.¹⁸ This TKI was one of the first targeted therapies used for cancer treatment and led to a major improvement of CML prognosis, with a remission in 80% of cases. After this time, it became clear that fusions can drive cancer development and are potential therapeutic targets in anti-cancer treatment in a very specific manner. Consequently to imatinib mesylate's history, fusion genes in hematological neoplasia and sarcoma were discovered¹⁹ and, more recently, high-resolution sequencing technologies enabled exploration of more fusion genes in other tumors.^{20,21} It should be noted that most of the studies and discoveries to date were made among adult patients; nevertheless, some pediatric cancers have been described to also harbor fusion genes that are involved in patients' diagnosis and/or targeted treatments (Table 2). However, more explorations are needed to not only identify new targets

but also to understand the function of fusion genes and their correlation to tumor initiation and progression.²² Indeed, of all of the multiple fusions identified by next-generation sequencing in each pediatric pathology (Table 2), only a few of them received attention for extensive functional studies, thus precluding the investigation and discovery of new druggable targets.¹⁹

Molecular Mechanisms Underlying Fusion Genes

Four different structural chromosomal rearrangements are observed in gene fusion: translocation, insertion, inversion, and deletion.²³ According to the coding or regulatory sequences affected, the physiological consequences are different.

Abnormal mRNA Expression

One of the consequences of fusion genes is the abnormal expression of one of the genes involved in the fusion (usually the 3' part of a gene). More precisely, the 5' part of a gene (gene 1), including the regulatory sequences and the 5' UTR part, is fused to the 3' part of a second gene (gene 2), which often only includes the coding sequence and the 3' UTR. The consequence of such fusion is that if gene 1 is constitutionally expressed or highly expressed and gene 2 is a proto-oncogene, then gene 2 will be upregulated by the promoter and regulatory sequences of gene 1. In this case, the result is a strong transcriptional activation of proto-oncogene 2, whose own regulatory elements are lost and is put under the complete control of the first gene.²² Fusion genes leading to abnormal protein expression were first described in adult tumors. Among the best-known examples are the fusions involving immunoglobulin genes and the *MYC* proto-oncogene, found in various hematological malignancies. Several translocations were found to be responsible for the fusion of *MYC* with other genes: t(8;14) led to IGH-*MYC* fusion, t(2;8) gave IGH-*MYC*, and t(8;22) resulted in IGL-*MYC*. *MYC* is an important transcription factor that binds to DNA in a non-specific manner to activate genes involved in cell cycle progression and inhibition of apoptosis. In these cases, the *MYC* gene is upregulated due to its 5' fusion with immunoglobulin regulatory sequences (promoter and enhancers).^{22,24} Indeed, the fusion results in constitutive activation of the *MYC* oncogene, increasing cell proliferation and tumorigenicity.

Another often-mentioned example is the fusion *TMPRSS2-ERG* found in 50% of prostate cancers, affecting chromosome 21.²⁰ This rearrangement puts *ERG* under the complete dependence of *TMPRSS2* regulatory sequences, upregulated by androgens present in the tissue. This leads to an overexpression of *ERG*.²⁰

Chimeric Oncoproteins

Chromosome rearrangements can also generate chimeric oncoproteins. In this case, both coding sequences of respective functional domains are conserved in the fusion. Its translation results in a protein with functional domains derived from both of the fused genes. Most of the chimeric oncoproteins characterized act as aberrant transcription factors, with a strong activation of unspecific target genes resulting in cell transformation.²⁵

**Table 2. Main Fusions Identified in Pediatric Cancers**

Pathology (Total No. of Recurrent Fusions)	Fusion Transcript (Most Frequent and Recurrent) ^a	Chromosome Abnormality	Diagnosis	Reference	Prognosis
Leukemia (113)					
	NUP98-NSD1	t(5;11)	acute monoblastic, myeloblastic, and myelomonocytic leukemia	90	poor
	SET-NUP214	del(9)	acute monoblastic, myeloblastic, and myelomonocytic leukemia; T-ALL	91	poor
	IgH-MYC/CEBPD/BCL2/CRLF2/ID4	t(8;14), t(8;14), t(14;18), t(X;14), t(6;14)	ALL	92-96	poor
	MYC rearranged	t(8;14), t(2;8), t(8;22)	various hematologic malignancies	97-99	poor
	TCRA/B/D rearranged		ALL	100-102	no prognostic value
	KMT2A-AFF1	t(1;11)	ALL, biphenotypic leukemia, AML	103	poor
	E2A-PBX1	t(1;19)	B-ALL	104,105	poor
	MLL-AF4	t(4;11)	B-ALL	106	poor
	PAX5-ETV6	t(9;12)	B-ALL	107	good
	ETV6-RUNX1	t(12;21)	B-ALL, AML, acute myeloblastic leukemia	108	good
	CALM-AF10	t(10;11)	T-ALL	109	poor
	TEL-AML1/RUNX1	t(8;12), t(12;21)	T-ALL, B-ALL	108,110	good
	AML1-ETO	t(8;21)	AML	111	good
	RPN1-PRDM16/MECOM	t(1;3), inv(3)	AML	112,113	poor
	NPM-MLF1	t(3;5)	AML, acute erythroleukemia	114	poor
	DEK-CAN	t(6;9)	AML, acute myeloblastic leukemia	115	poor
	ETV6-MDS-1/CHIC2	t(3;12), t(4;12)	AML, acute myeloblastic leukemia	116,117	poor
	RUNX1-CBFA2T3	t(16;21)	AML, acute myeloblastic leukemia, acute myelomonocytic leukemia	118	good
	CBFB-MYH11	inv(16)	AML, acute monoblastic leukemia, acute myelocytic leukemia, CML	119	good
	PML-RARA	t(15;17)	AML, biphenotypic leukemia, acute promyelocytic leukemia, CML	120,121	good
	USP16-RUNX1	del(21)	CML	122	poor
	BCR-ABL	t(9;22)	CML; pediatric ALL; undifferentiated and biphenotypic leukemia	123,124	good
	IgH-MYC	t(8;14)	various hematologic malignancies	125	good
	KMT2A-ELL/MLLT1	t(11;19)	various hematologic malignancies	126	poor
	MLL rearranged		ALL, biphenotypic leukemia, AML	127	poor
	ZMYM2-FGFR1	t(8;13)	various hematologic malignancies	128	poor
Lymphomas (41)					
	IGH-BCL10/BCL11A/FOXP1	t(1;14), t(2;14), t(3;14)	extranodal marginal zone B cell lymphoma, CLL, diffuse large B cell lymphoma, MALT lymphoma	129-131	poor for t(1;14); no information available for t(2;14) and t(3;14)
	IGK-BCL2/KDSR/CDK6	t(2;18), t(2;7)	follicular lymphoma, splenic marginal zone B cell lymphoma	132,133	no information available
	NPM1-ALK	t(2;5)	mature B cell neoplasms, mature T cell neoplasms	52	good
	IGL-BCL2	t(18;22)	mature B cell neoplasms, CLL	134	poor
	IGH/IGK-MYC	t(2;8), t(8;14)	various B cell neoplasms	135	poor
	IGH/HSP90AA1/IGL-BCL6	t(3;14), t(3;22)	various lymphoma types	136,137	good
	IGH-CCND1/BCL2/CEBPA	t(11;14), t(14;18), t(14;19)	various lymphoma types	138	no prognostic value

(Continued on next page)



Table 2. Continued

Pathology (Total No. of Recurrent Fusions)	Fusion Transcript (Most Frequent and Recurrent) ^a	Chromosome Abnormality	Diagnosis	Reference	Prognosis
	PAX5-IGH	t(9;14)	various lymphoma types	139	no prognostic value
	IGK/IGH-BCL6/CCND1	t(2;11), t(2;3)	various lymphoma types	136,140	no prognostic value
	IGH-TRA	inv(14)	various T cell neoplasms	141	no information available
Brain Tumors (4)					
	GOPC-ROS1	del(6)	astrocytoma grade III-IV	142	no prognostic value
	FGFR1/3-TACC1/3	del(4)	glioblastoma	58	no prognostic value
	IER3IP1-HDHD2	del(18)	meningioma	143	no prognostic value
	SLC44A1-PRKCA	t(9;17)	neuroglial neoplasm	144	no prognostic value
	KIAA1549-BRAF	dup(7)	pilocytic astrocytoma, astrocytoma grade II, pediatric oligodendroglioma	50	good
Kidney Tumors (4)					
	MALAT1-TFEB	t(6;11)	adenocarcinoma	145	good
	SFPQ/CLTC/ASPSCR1-TFE3	t(X;1), t(X;17)	adenocarcinoma, renal cell carcinoma	146-148	poor
Bone and Soft-Tissue Tumors (37)					
	PAX3/7-FKHR	t(2;13), t(1;13)	alveolar rhabdomyosarcoma	28	poor
	ASPSCR1-TFE3	t(X;17)	alveolar soft-part sarcoma	46	poor
	THRAP3-USP6	t(1;17)	aneurysmal bone cyst	149	no prognostic value
	EWSR1-ATF1	t(12;22)	CCS	150	poor (unconfirmed data)
	EWSR1-WT1	t(11;22)	desmoplastic round cell tumor	151	poor
	EWS-WT1/CHN	t(11;22), t(12;22), t(9;22)	DSCRT; CCS of tendons and aponeuroses; chondrosarcoma	152	poor
	WWTR1-CAMTA1	t(1;3)	epithelioid hemangioendothelioma	153	no bad prognosis
	EWSR1-FLI1/ERG	t(11;22)	Ewing's sarcoma	154,155	no prognostic value
	FUS-ERG	t(16;21)	Ewing's sarcoma	156	poor
	ETV6-NTRK3	t(12;15)	infantile fibrosarcoma	157	good
	FUS-DDIT3	t(12;16)	liposarcoma	158	poor
	VGLL2 rearranged		scRMS; SRMS	159	no prognostic value
	AHRR-NCOA2	t(5;8)	soft-tissue tumor, special type	160	no prognostic value
	MLL4-GPS2	t(17;19)	spindle cell sarcoma	161	poor
	SS18-SSX1	t(X;18)	synovial sarcoma	162	poor
	CIC-DUX4	t(4;19)	undifferentiated round cell sarcoma	163	poor
Liver Tumors (1)					
	DNAJB1-PRKACA	t(19;19)	fibrolamellar hepatocellular carcinoma	143	poor

ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; B-ALL, B-cell acute lymphoblastic leukemia; CCS, clear cell sarcoma; CLL, chronic lymphocytic leukemia; CML, chronic myeloid leukemia; DSCRT, desmoplastic small round cell tumor; MALT, mucosa-associated lymphoid tissue; ScRMS, sclerosing rhabdomyosarcoma; SRMS, spindle cell rhabdomyosarcoma; T-ALL, T-lineage acute lymphoblastic leukemia.

^aMost of the fusions are listed in the Mitelman database (<http://cgap.nci.nih.gov/Chromosomes/RecurrentAberrations>).

To date, several chimeric oncoproteins have been identified in blood malignancies but very few are described in carcinomas. One well-known example is the fusion RET-PTC found in papillary thyroid carcinomas, where RET (a gene involved in the regulation of cell survival, growth, differentiation, and migration) is constitutionally activated.²⁶ Concerning pediatric cancers, EWS-FLI1 fusion is found in 85% of Ewing's sarcoma, which results in a chimeric oncoprotein

between the amino terminus of EWS and the carboxy terminus of FLI, giving aberrant transcriptional activity.²⁷ In rhabdomyosarcomas, PAX3/7-FKHR fusions are found in 80%–85% of alveolar rhabdomyosarcomas and contain the PAX3/PAX7 DNA binding domain and the FKHR transcriptional activation domain. Therefore, PAX target genes are highly transcriptionally activated.²⁸ Occasionally, cryptic fusion partners occur in these fusion-driven tumors.



Gene Function Truncation

More recently, studies have shown a third impact of fusion genes resulting in the entire cutout of gene domains. The truncated gene is generally a tumor suppressor gene and is therefore inactivated by the truncation, leading to cell transformation. Inactivation of the tumor suppressor gene can occur through different mechanisms: the fusion can either act as the second hit (e.g., *CDKN2A* or *NF1* genes) or lead to a dominant effect on the wild-type protein (e.g., *PAX5* fusions) or even provoke haploinsufficiency of the disrupted protein (e.g., *RUNX1* fusions).^{29–32}

Another category that does not involve chromosomal rearrangements is the occurrence of read-through transcripts. This type of modification results in transcription-induced gene fusions (TIGFs) that are not induced by a modification in the DNA sequence but are a consequence of alternative splicing. Thus, the genes stay located far apart in the same chromosome or on different chromosomes, but mRNA fusion transcripts are formed as a consequence of *cis*-TIGFs or *trans*-TIGFs.³³ Most TIGFs probably represent random *in vivo* events with no impact on the cell's functions and some are also artifacts of deep sequencing. Nevertheless, some *cis*-TIGFs have been associated with specific organ localization, such as *SLC45A3-ELK4* and *MSMB-NCOA4* detected in normal and neoplastic prostate tissue^{33,34} and *SCNN1A-TNFRSF1A* and *CTSD-IFITM10* identified in normal and neoplastic breast tissue.³⁵

Relevance of Fusion Genes in Pediatric Cancers

Leukemia is the most common type of pediatric cancer (around 30% of total cases) and most of the recurrent identified fusions were found in this pathology, as detailed in Table 2.

Concerning pediatric solid tumors, several fusions have been described and can be grouped into two categories. The first category concerns fusions that were considered until now only as diagnostic markers, and the second category includes fusions providing new therapeutic targets.

In the first category, the most described example is Ewing's sarcoma. Ewing's sarcoma is characterized by a reciprocal translocation t(11;22) (q24;q12) leading to the fusion between *EWS* and *FLI1* genes.³⁶ The corresponding chimeric protein acts as an aberrant transcription factor. Different translocations have been found in patients suffering from Ewing's sarcoma involving TET/ETS partners and the list is still growing.³⁷ Another example of fusions belonging to the first category are *PAX3/7-FKHR* fusions found in alveolar rhabdomyosarcomas due to the translocation t(2;13)(q35;q14) or, less commonly, t(1;13)(p36;p14), leading to *PAX3-FOXO1* (55%–70%) or *PAX7-FOXO1* (10%–22%) fusion genes, respectively.³⁸ The resulting fusion proteins harbor high transcriptional activity due to the juxtaposition of the PAX protein DNA binding domain and the FKHR activation domain.²⁸ In another example, Honeyman et al.³⁹ identified the *DNAJB1-PRKCA* fusion in 100% of patients with fibrolamellar hepatocellular carcinoma (FLHCC), a rare subset of pediatric hepatocarcinoma. Although its function is not yet elucidated, this fusion provides

a new and unique diagnostic biomarker for patients with FLHCC. Another fusion in this category is *C11orf95-RELA*, which was identified recently in ~70% of supratentorial ependymomas.⁴⁰ *RELA* is a gene encoding for the RelA p65 subunit of the nuclear factor κ B (NF- κ B) complex. When fused, *RELA* loses its upstream regulator sequences (5' UTR) but the coding sequence is conserved. The function of *C11orf95* is still unknown. *C11orf95* has been found fused with different partners, suggesting a possible role of its zinc-finger domains.⁴¹ The *C11orf95-RELA* fusion seems to arise through chromothripsis and leads to a constitutive activation of NF- κ B signaling.⁴⁰ Ependymomas are classified according to their location in the supratentorial, infratentorial region of the brain or the spinal cord.⁴² It is important to highlight that *C11orf95-RELA* is detected only in supratentorial ependymomas,⁴⁰ thus providing a new molecular diagnostic biomarker and possible new therapeutic target. Finally, the reciprocal translocation t(X;18; p11,q11), which is found in all cases of synovial sarcoma, leads to fusions involving the *SYT* gene (also named *SSX18*; namely, *SYT-SSX1* or *SYT-SSX2* fusions).^{43,44} The resulting chimeric oncoproteins contain the activation domain of *SYT* and the regulatory domain of *SSX1/2* genes. The newly protein appears to be a transcriptional co-regulator via protein-protein interactions and it has been described to be able to regulate SWI-SNF chromatin remodeling complexes.⁴⁵ This fusion is not found in any other tumor types which makes *SYT-SSX* an important diagnostic biomarker, and also relevant for prognostic purposes.⁴⁴

The second category includes fusions with therapeutic outcomes, such as *ASPSCR1-TFE3*,⁴⁶ *FGFR1-TACC1* and *FGFR3-TACC3*,^{47–49} *KIAA1549-BRAF*,^{50,51} *NPM-ALK*,^{52,53} and *NTRK* fusions (R. Nagasubramanian et al., 2016, *J. Clin. Oncol.*, abstract).⁵⁴

ASPSCR1-TFE3 [t(X;17)(p11.2;q25)] and, more generally, *TFE3* fusions are found in more than 95% of patients with alveolar soft-part sarcoma and in a subset of patients with renal cell carcinoma.^{46,55} This oncoprotein acts also as an aberrant transactivator, stronger than the native *TFE3* protein, and has a nuclear localization. *MET* is among the genes strongly transactivated by *ASPSCR1-TFE3* protein.⁵⁶ This may lead to sensitivity to *MET* inhibitors *in vitro*,⁵⁷ which is currently being explored in a clinical trial with crizotinib (NCT01524926).

The *FGFR1-TACC1* and *FGFR3-TACC3* fusions were first identified in glioblastoma multiforme (GBM; also known as grade IV astrocytoma). They have since been found in various cancer types, such as bladder cancer, lung adenocarcinoma, head and neck cancer, nasopharyngeal carcinoma, and esophageal squamous cell carcinoma.^{47–49} The rearrangement that fuses the tyrosine kinase domains of fibroblast growth factor receptor (FGFR) to the transforming acidic coiled-coil (TACC) harbors a constitutive kinase activity. Interestingly, the inhibition of FGFR kinase by pharmacological agents counteracts the oncogenicity of the fusion *in vitro* and *in vivo* in glioblastoma.⁵⁸ Therefore, the discovery of *FGFR-TACC* fusions is of great interest and the role of FGFR inhibitors is currently being explored in clinical trials with multiple FGFR kinase inhibitors, such as BGJ398 and AZD4547 (NCT01975701 and NCT02824133).



KIAA1549-BRAF also belongs to the second category of fusions and is found in 65%–75% of sporadic pilocytic astrocytomas^{50,51} but is also detected in other pediatric brain tumors.^{59,60} Importantly, KIAA1549-BRAF leads to an over-activation of mitogen-activated protein kinase (MAPK) signaling,⁵⁰ resulting in an oncogenic addiction to this pathway.⁶⁰ The first experiments showing its therapeutic potential involved targeting cells transfected with the KIAA1549-BRAF fusion using the BRAF inhibitor vemurafenib (PLX4032), a selective BRAF^{V600E} inhibitor,⁶¹ as *BRAF* mutations also result in MAPK signaling activation. Although the anti-tumor effects of vemurafenib in *BRAF* (V600E)-mutated cells are well demonstrated, the treatment of cell lines expressing KIAA1549-BRAF with PLX4032 resulted in activation of the MAPK pathway, leading to increased cell proliferation in vitro and tumor growth in vivo.^{62,63} The mechanism underlying this paradoxical activation was explained by Poulikos et al.,⁶⁴ who described that PLX4032 is able to bind on RAF isoforms and induce interaction with RAS-GTP, which leads to the pathway activation. In *BRAF* (V600E)-mutated tumors, the level of RAS may not be sufficient to transactivate wild-type BRAF. This RAS activation is notably discussed as a mechanism responsible for the development of squamous cell carcinoma in normal skin tissue after BRAF^{V600E}-inhibiting treatment.⁶⁵

Second-generation BRAF inhibitors such as PLX PB-3 are currently being developed, which lead to inhibition of KIAA15-49-BRAF with decreased proliferation and tumorigenicity and may not induce the paradoxical MAPK activation. As a result, they could be more efficient for patients carrying KIAA1549-BRAF fusions (S.-S. Lang et al., 2012, Neurosurgery, abstract).

Because MEK is a downstream effector of RAF in the MAPK signaling pathway, KIAA1549-BRAF-positive cells may consequently be sensitive to MEK inhibitors (e.g., selumetinib and trametinib) that are able to block the pathway downstream of RAF and thus independently from the RAF alteration underlying the pathway activation. In a pre-clinical study, selumetinib (AZD6244) showed evidence of anti-tumor activity in a juvenile pilocytic astrocytoma xenograft model.⁶⁶ This is being currently explored in patients carrying KIAA1549-BRAF fusions (NCT01089101 and NCT02124772). Taken together, these findings demonstrate the importance of testing the suitability of a drug in a relevant setting. Their effectiveness is dependent on the cellular and the molecular context of the targeted tumors, and the “general” use of kinase inhibitors may be considered with caution.

Another example is the fusion NPM-ALK, caused by the translocation t(2;5)(p23;q35),⁵² which fuses the dimerization domain of NPM with the tyrosine kinase domain of ALK. ALK is a tyrosine kinase receptor; therefore, the fusion gives a chimeric oncoprotein with a constitutively activated kinase.⁵² The translocation is found in 30%–50% of advanced-stage anaplastic large cell lymphoma cases in adults;⁶⁷ 90% of anaplastic large cell lymphoma cases among children are ALK positive.⁶⁸ This fusion can be targeted via ALK inhibitors such as crizotinib or ceritinib, which have been registered for patients carrying non-small cell lung cancer (NSCLC) with ALK translocations

and gave promising results with high rates of response.⁶⁹ Promising results have also been reported in pediatric patients with anaplastic large cell lymphoma⁵³ (NCT00939770 and NCT01742286) (B. Georger et al., 2015, J. Clin. Oncol., abstract).

Finally, and more recently, NTRK fusions involving either NTRK1, 2, or 3 located on chromosome 1 q21-q22 have raised great interest in the clinical setting, as they are found in different neoplasms in adults and children. A total of 22 different 5' partners were found to be fused with *NTRK* genes.⁵⁴ *NTRK* codes for the family of tropomyosin receptors, which, once fused, harbor a constitutively activated kinase function. This leads to cancer progression by activating different oncogenic pathways such as the MAPK or AKT pathways.⁷⁰ Among childhood neoplasms harboring NTRK fusions are found soft-tissue sarcoma, congenital infantile fibrosarcoma, glioblastoma, low-grade glioma, pilocytic astrocytoma, congenital mesoblastic nephroma, acute myeloid leukemia, and various other tumor types. Therefore, for these patients, targeting NTRK is of great clinical interest and a novel inhibitor (LOXO-101) is currently being evaluated in the clinical setting (NCT02637687) (R. Nagasubramanian et al., 2016, J. Clin. Oncol., abstract).

Taken together, all of these studies emphasize the notion that fusions can be of great clinical interest not only as new therapeutic targets but also as biomarkers. Indeed, fusions specific to a cancerous disease are ideal for diagnostic purposes and for subgroup classifications; therefore, they can also be used as prognostic biomarkers. Some fusions, such as BCR-ABL in CML⁷¹ or NPM-ALK in ALCL,⁷² are even used to monitor the tumor load and treatment response (as minimal residual disease markers). Hence, even individual or non-recurrent fusions could have clinical value in an individual patient.⁷³ It should be kept in mind that current treatments are at their limits of toxicity and efficacy. Accordingly, specific treatments are needed to treat non-responsive patients and also to improve the quality of life of cured patients.

Issues Facing Discovery of Fusion Genes in Childhood Cancers

Analyses of fusions have a strong interest in the field of childhood cancers. The discovery step is still in its infancy for these malignancies, probably because studies addressing the discovery of fusion genes encounter important difficulties. The first difficulty comes from the biopsy and the way in which the DNA or RNA was extracted. The purity and quantity of the extracted material will impact the quality of further analysis. Most of the studies involved in fusion discovery use next-generation sequencing (NGS) techniques and new bioinformatics analysis methods. It is necessary to combine the strengths of both sequencing technologies and computational strategies, together with the discrimination of events tightly correlated with cancer development.⁷⁴ One difficulty encountered in gene fusion detection is the ability of discrimination between existent new discoveries and computer analysis artifacts. However, too stringent selection parameters applied to avoid false positives can preclude the detection of effective oncogenic fusions. Being able to cut out the artifacts from the analysis is one of the biggest challenges, since current techniques



of deep sequencing are more sensitive than before and allow the detection of rare events; however, the risk of false positives is still present.⁷⁵ It is also important to note that each bioinformatics tool used for fusion detection and its subsequent algorithms differ in terms of sensitivity and specificity, depending on the different filters and criteria applied. Therefore, combining several tools and even different types of analysis could be highly relevant.^{19,76}

Moreover, cancer and carcinomas especially often have highly rearranged genomes, and many of the gene fusions that are detected can represent “passenger” events that are caused by chromosomal instability.^{13,77} These passenger rearrangements are secondary events present in the tumor, which are different from the primary oncogenic event responsible for the development of the disease. They are particularly frequent in relapsing patients, where treatments by radiotherapy and ionization can lead to “passenger fusions” not involved in oncogenic processes.^{74,78} Among these patients, it is more difficult to distinguish the fusion gene capable of an oncogenic effect from the passenger fusions that can be constitutively present in normal and cancerous cells.

Regarding pediatric cancers, the emergence of NGS and new informatics tools to detect fusions have aimed to detect recurrent fusions with expected targetable domains, such as fusions with tyrosine kinase activities. However, considering the rarity of these types of neoplasia, the probability of finding rearranged genes is much lower than for adult cohorts.⁷⁸ In addition, the major difficulty is linked to the intratumoral heterogeneity, where subclonal fusion events can be weakly detectable, making their identification more complex.^{19,74}

Pertinence of Fusion-Based Murine Models in Childhood Cancers

Most *in vivo* studies have assessed the tumorigenic potential of a fusion gene in xenografted tumors on immunocompromised animals. Despite the validity of this experimental approach, some questions (e.g., the role of a fusion gene in early stages of the development of a pediatric cancer) remain difficult to answer using xenografted tumors. Therefore, several mouse models were developed in order to better understand pediatric tumors’ oncogenesis. These models helped to demonstrate the importance of (1) the cell of origin, (2) the differentiation stage, and (3) a second hit in complement to the first one (fusion oncogene) necessary for disease development. With the first desire to provide new models to study pediatric cancers, four striking examples of Ewing’s sarcoma, myxoid liposarcoma, alveolar rhabdomyosarcoma, and synovial sarcoma models have provided major information on the tumorigenesis mechanisms of fusion-harboring pediatric cancers.

Lin et al.⁷⁹ explored the role of EWS/FLI1 in Ewing’s sarcoma and showed that its specific expression in mesenchymal cells of the embryonic limb buds impaired the development of mouse limbs but did not induce tumors. This suggested that the cell of origin may not be a mesenchymal stem cell. Interestingly, when *TP53* was simul-

taneously deleted, EWS-FLI1 promoted tumor formation.⁷⁸ Therefore, the presence of additional mutations notably in *TP53* seems to be required for the transformation process by EWS-FLI1 in Ewing’s sarcoma.

In myxoid liposarcoma, previous studies demonstrated that the introduction of the *FUS-CHOP* transgene into the mouse genome led to the development of liposarcomas,^{80,81} but its specific expression in differentiated aP2-expressing adipocytes did not result in tumor formation,⁸² suggesting the importance of the “stage of differentiation” in driving the tumors.

In alveolar rhabdomyosarcoma, the expression of PAX3-FKHR fusion in mouse primary mesenchymal stem cells did not lead to tumor formation,⁸³ while its expression in differentiated muscle cells at a late stage of embryogenesis induced tumors but at a low frequency. The addition of *TP53* or *INK4A/ARF* disruption increased the frequency of alveolar rhabdomyosarcoma (RMS), emphasizing the fact that additional mutations to PAX3-FKHR fusion might be necessary to generate alveolar RMS.⁸⁴

Finally, in synovial sarcoma, tumors were observed in mice models only when the fusion gene *SYT-SSX2* was expressed in immature myoblasts.⁸⁵ This observation again highlights the importance of both the cell of origin and the differentiation stage factors in fusion-driven tumorigenesis.

What Is Next?

Despite all of the emerging challenges related to fusion detection, identifying new fusions remains one of the most important priorities of research in pediatric cancers where new therapeutic targets are needed. The difficulty in discovering new fusion genes is minimal compared to the vast clinical consequences in terms of prognosis, diagnosis, and new therapeutic opportunities and to the importance of understanding the molecular mechanisms underlying a fusion gene.

Characterization of genomic alterations found in pediatric cancers would clarify their relevance, help to classify patients into subgroups, and aid in the discovery of new therapeutic, prognostic, diagnostic, or predictive biomarkers. Moreover, functional studies are strongly needed to better understand the relevance of the biological pathways involving fusion genes in the development of pediatric malignancies. Indeed, well-characterized fusion genes could reveal new targets in pediatric cancers for which there are currently no specific therapies available.⁸⁶

It is well known that only a small number of relevant molecular subsets have been characterized (e.g., the four subtypes of medulloblastoma); therefore, particular attention should be given to the rare histological subgroups.⁸⁷ These subsets could be of strong interest for diagnosis determination, understanding new mechanisms of tumor progression or resistance, and could help to bring forward new unexpected therapeutic targets.



With small histological subgroups, high priority should be given to the future study of genomic alterations at relapse. In fact, several studies have reported the importance of analyzing tissues at relapse,^{88,89} and clinical trials such as MAPPYACTS (molecular profiling for pediatric and young adult cancer treatment stratification) conducted in our institute will contribute to this investigation (NCT02613962). The study of patients' genomic patterns at relapse could lead to better knowledge of cancer-driving and resistance mechanisms but could also allow better classification of patients.

Moreover, knowing that pediatric cancers are rare diseases that differ from one patient to another, an isolated fusion found only in a single patient should not be discarded and "case-by-case" studies as well as genome-based clinical trials are required. To counteract the hurdles due to the rarity of pediatric cancers, data sharing of high-throughput sequencing analysis would offer a personalized-medicine approach based on targeting junction oncogenes.⁸⁶ More specifically, new therapeutic technologies are emerging regarding the targeting of junction oncogenes, such as the use of small interfering RNA (siRNA). siRNAs allow inhibition of gene expression at the mRNA level, avoiding their translation. This type of strategy offers promising opportunities for targeted therapy, since it specifically targets fusion transcripts only expressed in tumor cells, without affecting other genes. Our laboratory showed previously that the inhibition of fusion oncogenes responsible for tumor progression of thyroid cancer (RET/PTC) or prostate cancer (TMPRSS2-ERG) by the injection of vectorized siRNAs led to a decrease in tumor growth.^{73–77}

An increased number of new fusion genes are expected to be discovered, owing to worldwide projects (The Cancer Genome Atlas [TCGA], <https://tcga-data.nci.nih.gov/docs/publications/tcga/>) and to data-sharing platforms (e.g., the International Cancer Genome Consortium [ICGC] platform, <https://icgc.org/>; and the European Genome-phenome Archive [EGA], <https://ega.crg.eu/>).¹⁹ This will provide new opportunities of treatments, such as therapies based on siRNA technology for children carrying neoplasia.

ACKNOWLEDGMENTS

This study is supported by funding from "the Fondation ARC pour la recherche sur le cancer" within the Joint Translational Call 2014 "Translational research on human tumour heterogeneity to understand and overcome recurrence and resistance to therapy" TRANSCAN-2 JTC2014, ERA-NET ("Targeting Of Resistance in PEDIatric Oncology" [TORPEDO]).

REFERENCES

- American Cancer Society (2016). Key statistics for childhood cancers. <https://www.cancer.org/cancer/cancerinchildren/detailedguide/cancer-in-children-key-statistics>.
- Bahadur, G., and Hindmarsh, P. (2000). Age definitions, childhood and adolescent cancers in relation to reproductive issues. *Hum. Reprod.* 15, 227.
- Minnesota Department of Health (2016). Childhood cancers: facts and figures. https://apps.health.state.mn.us/mndata/cancer_child#childbraincancer_year.
- Fidler, M.M., Reulen, R.C., Winter, D.L., Kelly, J., Jenkinson, H.C., Skinner, R., Frobisher, C., and Hawkins, M.M.; British Childhood Cancer Survivor Study Steering Group (2016). Long term cause specific mortality among 34 489 five year survivors of childhood cancer in Great Britain: population based cohort study. *BMJ* 354, i4351.
- Marshall, G.M., Carter, D.R., Cheung, B.B., Liu, T., Mateos, M.K., Meyerowitz, J.G., and Weiss, W.A. (2014). The prenatal origins of cancer. *Nat. Rev. Cancer* 14, 277–289.
- Zhang, J., Walsh, M.F., Wu, G., Edmonson, M.N., Gruber, T.A., Easton, J., Hedges, D., Ma, X., Zhou, X., Yergeau, D.A., et al. (2015). Germline mutations in predisposition genes in pediatric cancer. *N. Engl. J. Med.* 373, 2336–2346.
- Kaatsch, P. (2010). Epidemiology of childhood cancer. *Cancer Treat. Rev.* 36, 277–285.
- Kupfer, G.M. (2015). Childhood cancer epidemiology: overview, tools of study, cancer incidence. <http://emedicine.medscape.com/article/989841-overview>.
- Huether, R., Dong, L., Chen, X., Wu, G., Parker, M., Wei, L., Ma, J., Edmonson, M.N., Hedlund, E.K., Rusch, M.C., et al. (2014). The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. *Nat. Commun.* 5, 3630.
- Orentas, R.J., Lee, D.W., and Mackall, C. (2012). Immunotherapy targets in pediatric cancer. *Front. Oncol.* 2, 3.
- Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., Jr., and Kinzler, K.W. (2013). Cancer genome landscapes. *Science* 339, 1546–1558.
- Duesberg, P.H. (1987). Cancer genes generated by rare chromosomal rearrangements rather than activation of oncogenes. *Med. Oncol. Tumor Pharmacother.* 4, 163–175.
- Yoshihara, K., Wang, Q., Torres-Garcia, W., Zheng, S., Vegesna, R., Kim, H., and Verhaak, R.G. (2015). The landscape and therapeutic relevance of cancer-associated transcript fusions. *Oncogene* 34, 4845–4854.
- Walz, A.L., Ooms, A., Gadd, S., Gerhard, D.S., Smith, M.A., Guidry Auvil, J.M., Meerzaman, D., Chen, Q.R., Hsu, C.H., Yan, C., et al. (2015). Recurrent DGCR8, DROSHA, and SIX homeodomain mutations in favorable histology Wilms tumors. *Cancer Cell* 27, 286–297.
- Nowell, P.C. (1962). The minute chromosome (Ph1) in chronic granulocytic leukaemia. *Blut* 8, 65–66.
- Rowley, J.D. (1973). Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature* 243, 290–293.
- de Klein, A., van Kessel, A.G., Grosveld, G., Bartram, C.R., Hagemeyer, A., Bootsma, D., Spurr, N.K., Heisterkamp, N., Groffen, J., and Stephenson, J.R. (1982). A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. *Nature* 300, 765–767.
- Basel, N. (2001) FDA approves Novartis' unique cancer medication Glivec®. <http://www.evaluategroup.com/Universal/View.aspx?type=Story&id=5838>.
- Mertens, F., Johansson, B., Fioretto, T., and Mitelman, F. (2015). The emerging complexity of gene fusions in cancer. *Nat. Rev. Cancer* 15, 371–381.
- Tomlins, S.A., Rhodes, D.R., Perner, S., Dhanasekaran, S.M., Mehra, R., Sun, X.-W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R., et al. (2005). Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 310, 644–648.
- Ju, Y.S., Lee, W.-C., Shin, J.-Y., Lee, S., Bleazard, T., Won, J.-K., Kim, Y.T., Kim, J.I., Kang, J.H., and Seo, J.S. (2012). A transforming KIF5B and RET gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. *Genome Res.* 22, 436–445.
- Mitelman, F., Johansson, B., and Mertens, F. (2007). The impact of translocations and gene fusions on cancer causation. *Nat. Rev. Cancer* 7, 233–245.
- Griffiths, A.J.F., Gelbart, W.M., Miller, J.H., and Lewontin, R.C. (1999). Chromosomal rearrangements. In *Modern Genetic Analysis*, A.J.F. Griffiths, W.M. Gelbart, J.H. Miller, and R.C. Lewontin, eds. (W.H. Freeman), <http://www.ncbi.nlm.nih.gov/books/NBK21367/>.
- Küppers, R. (2005). Mechanisms of B-cell lymphoma pathogenesis. *Nat. Rev. Cancer* 5, 251–262.
- Schwab, M., ed. (2012). *Encyclopedia of Cancer* (Springer), <http://link.springer.com/10.1007/978-3-642-16483-5>.



26. Santoro, M., Melillo, R.M., and Fusco, A. (2006). RET/PTC activation in papillary thyroid carcinoma: European Journal of Endocrinology Prize Lecture. *Eur. J. Endocrinol.* *155*, 645–653.
27. Ohno, T., Rao, V.N., and Reddy, E.S. (1993). EWS/Flt-1 chimeric protein is a transcriptional activator. *Cancer Res.* *53*, 5859–5863.
28. Barr, F.G. (2001). Gene fusions involving PAX and FOX family members in alveolar rhabdomyosarcoma. *Oncogene* *20*, 5736–5746.
29. Duro, D., Bernard, O., Della Valle, V., Leblanc, T., Berger, R., and Larsen, C.J. (1996). Inactivation of the P16INK4/MTS1 gene by a chromosome translocation t(9;14)(p21-22;q11) in an acute lymphoblastic leukemia of B-cell type. *Cancer Res.* *56*, 848–854.
30. Storlazzi, C.T., Von Steyern, F.V., Domanski, H.A., Mandahl, N., and Mertens, F. (2005). Biallelic somatic inactivation of the NF1 gene through chromosomal translocations in a sporadic neurofibroma. *Int. J. Cancer* *117*, 1055–1057.
31. Coyau, E., Struski, S., Prade, N., Familiades, J., Eichner, R., Quelen, C., Bousquet, M., Mugneret, F., Talmant, P., Pages, M.P., et al. (2010). Wide diversity of PAX5 alterations in B-ALL: a Groupe Francophone de Cyto-genetique Hematologique study. *Blood* *115*, 3089–3097.
32. Agerstam, H., Lilljebjörn, H., Lassen, C., Swedin, A., Richter, J., Vandenberghe, P., Johansson, B., and Fioretos, T. (2007). Fusion gene-mediated truncation of RUNX1 as a potential mechanism underlying disease progression in the 8p11 myeloproliferative syndrome. *Genes Chromosomes Cancer* *46*, 635–643.
33. Nacu, S., Yuan, W., Kan, Z., Bhatt, D., Rivers, C.S., Stinson, J., Peters, B.A., Modrusan, Z., Jung, K., Seshagiri, S., and Wu, T.D. (2011). Deep RNA sequencing analysis of readthrough gene fusions in human prostate adenocarcinoma and reference samples. *BMC Med. Genomics* *4*, 11.
34. Rickman, D.S., Pflueger, D., Moss, B., VanDoren, V.E., Chen, C.X., de la Taille, A., Kuefer, R., Tewari, A.K., Setlur, S.R., Demichelis, F., and Rubin, M.A. (2009). SLC45A3-ELK4 is a novel and frequent erythroblast transformation-specific fusion transcript in prostate cancer. *Cancer Res.* *69*, 2734–2738.
35. Varley, K.E., Gertz, J., Roberts, B.S., Davis, N.S., Bowling, K.M., Kirby, M.K., Nesmith, A.S., Oliver, P.G., Grizzle, W.E., Forero, A., et al. (2014). Recurrent read-through fusion transcripts in breast cancer. *Breast Cancer Res. Treat.* *146*, 287–297.
36. Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., de Jong, P., Rouleau, G., et al. (1992). Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* *359*, 162–165.
37. Sankar, S., and Lessnick, S.L. (2011). Promiscuous partnerships in Ewing's sarcoma. *Cancer Genet.* *204*, 351–365.
38. de Souza, R.R., Oliveira, I.D., del Giudice Paniago, M., Yaoita, F.H.K., Caran, E.M.M., Macedo, C.R.P.D., Petrilli, A.S., Abib, Sde.C., de Seixas Alves, M.T., and de Toledo, S.R. (2014). Investigation of IGF2, Hedgehog and fusion gene expression profiles in pediatric sarcomas. *Growth Horm. IGF Res.* *24*, 130–136.
39. Honeyman, J.N., Simon, E.P., Robine, N., Chiaroni-Clarke, R., Darcy, D.G., Lim, I.I.P., Gleason, C.E., Murphy, J.M., Rosenberg, B.R., Teegan, L., et al. (2014). Detection of a recurrent DNAB1-PRKACA chimeric transcript in fibrolamellar hepatocellular carcinoma. *Science* *343*, 1010–1014.
40. Parker, M., Mohankumar, K.M., Punchihewa, C., Weinlich, R., Dalton, J.D., Li, Y., Lee, R., Tatevossian, R.G., Phoenix, T.N., Thiruvengadam, R., et al. (2014). C11orf95-RELA fusions drive oncogenic NF- κ B signalling in ependymoma. *Nature* *506*, 451–455.
41. Pietsch, T., Wohlers, I., Goschzik, T., Dreschmann, V., Denkhau, D., Dörner, E., Rahmann, S., and Klein-Hitpass, L. (2014). Supratentorial ependymomas of childhood carry C11orf95-RELA fusions leading to pathological activation of the NF- κ B signaling pathway. *Acta Neuropathol.* *127*, 609–611.
42. Wu, J., Armstrong, T.S., and Gilbert, M.R. (2016). Biology and management of ependymomas. *Neuro-oncol.* *18*, 902–913.
43. Clark, J., Rocques, P.J., Crew, A.J., Gill, S., Shipley, J., Chan, A.M., Gusterson, B.A., and Cooper, C.S. (1994). Identification of novel genes, SYT and SXX, involved in the t(X;18)(p11.2;q11.2) translocation found in human synovial sarcoma. *Nat. Genet.* *7*, 502–508.
44. Ladanyi, M., Antonescu, C.R., Leung, D.H., Woodruff, J.M., Kawai, A., Healey, J.H., Brennan, M.F., Bridge, J.A., Neff, J.R., Barr, F.G., et al. (2002). Impact of SYT-SSX fusion type on the clinical behavior of synovial sarcoma: a multi-institutional retrospective study of 243 patients. *Cancer Res.* *62*, 135–140.
45. Nielsen, T.O., Poulin, N.M., and Ladanyi, M. (2015). Synovial sarcoma: recent discoveries as a roadmap to new avenues for therapy. *Cancer Discov.* *5*, 124–134.
46. Ladanyi, M., Lui, M.Y., Antonescu, C.R., Krause-Boehm, A., Meindl, A., Argani, P., Healey, J.H., Ueda, T., Yoshikawa, H., Meloni-Ehrig, A., et al. (2001). The der(17)t(X;17)(p11;q25) of human alveolar soft part sarcoma fuses the TFE3 transcription factor gene to ASPL, a novel gene at 17q25. *Oncogene* *20*, 48–57.
47. Yuan, L., Liu, Z.-H., Lin, Z.-R., Xu, L.-H., Zhong, Q., and Zeng, M.-S. (2014). Recurrent FGFR3-TACC3 fusion gene in nasopharyngeal carcinoma. *Cancer Biol. Ther.* *15*, 1613–1621.
48. Capelletti, M., Dodge, M.E., Ercan, D., Hammerman, P.S., Park, S.-I., Kim, J., Sasaki, H., Jablons, D.M., Lipson, D., Young, L., et al. (2014). Identification of recurrent FGFR3-TACC3 fusion oncogenes from lung adenocarcinoma. *Clin. Cancer Res.* *20*, 6551–6558.
49. Carneiro, B.A., Elvin, J.A., Kamath, S.D., Ali, S.M., Paintal, A.S., Restrepo, A., Berry, E., Giles, F.J., and Johnson, M.L. (2015). FGFR3-TACC3: A novel gene fusion in cervical cancer. *Gynecol Oncol Rep* *13*, 53–56.
50. Jones, D.T.W., Kocialkowski, S., Liu, L., Pearson, D.M., Bäcklund, L.M., Ichimura, K., and Collins, V.P. (2008). Tandem duplication producing a novel oncogenic BRAF fusion gene defines the majority of pilocytic astrocytomas. *Cancer Res.* *68*, 8673–8677.
51. Sievert, A.J., Jackson, E.M., Gai, X., Hakonarson, H., Judkins, A.R., Resnick, A.C., Sutton, L.N., Storm, P.B., Shaikh, T.H., and Biegel, J.A. (2009). Duplication of 7q34 in pediatric low-grade astrocytomas detected by high-density single-nucleotide polymorphism-based genotype arrays results in a novel BRAF fusion gene. *Brain Pathol.* *19*, 449–458.
52. Morris, S.W., Kirstein, M.N., Valentine, M.B., Dittmer, K., Shapiro, D.N., Look, A.T., and Saltman, D.L. (1995). Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science* *267*, 316–317.
53. Mossé, Y.P., Lim, M.S., Voss, S.D., Wilner, K., Ruffner, K., Laliberte, J., Rolland, D., Balis, F.M., Maris, J.M., Weigel, B.J., et al. (2013). Safety and activity of crizotinib for paediatric patients with refractory solid tumours or anaplastic large-cell lymphoma: a Children's Oncology Group phase 1 consortium study. *Lancet Oncol.* *14*, 472–480.
54. Amatu, A., Sartore-Bianchi, A., and Siena, S. (2016). NTRK gene fusions as novel targets of cancer therapy across multiple tumour types. *ESMO Open* *1*, e000023.
55. Argani, P., Antonescu, C.R., Illei, P.B., Lui, M.Y., Timmons, C.F., Newbury, R., Reuter, V.E., Garvin, A.J., Perez-Atayde, A.R., Fletcher, J.A., et al. (2001). Primary renal neoplasms with the ASPL-TFE3 gene fusion of alveolar soft part sarcoma: a distinctive tumor entity previously included among renal cell carcinomas of children and adolescents. *Am. J. Pathol.* *159*, 179–192.
56. Kobos, R., Nagai, M., Tsuda, M., Merl, M.Y., Saito, T., Laé, M., Mo, Q., Olshen, A., Lianoglou, S., Leslie, C., et al. (2013). Combining integrated genomics and functional genomics to dissect the biology of a cancer-associated, aberrant transcription factor, the ASPSCR1-TFE3 fusion oncoprotein. *J. Pathol.* *229*, 743–754.
57. Tsuda, M., Davis, I.J., Argani, P., Shukla, N., McGill, G.G., Nagai, M., Saito, T., Laé, M., Fisher, D.E., and Ladanyi, M. (2007). TFE3 fusions activate MET signaling by transcriptional up-regulation, defining another class of tumors as candidates for therapeutic MET inhibition. *Cancer Res.* *67*, 919–929.
58. Singh, D., Chan, J.M., Zoppoli, P., Niola, F., Sullivan, R., Castano, A., Liu, E.M., Reichel, J., Porrati, P., Pellegatta, S., et al. (2012). Transforming fusions of FGFR and TACC genes in human glioblastoma. *Science* *337*, 1231–1235.
59. Antonelli, M., Badiali, M., Moi, L., Buttarelli, F.R., Baldi, C., Massimino, M., Sanson, M., and Giangaspero, F. (2015). KIAA1549-BRAF fusion gene in pediatric brain tumors of various histogenesis. *Pediatr. Blood Cancer* *62*, 724–727.
60. Kumar, A., Pathak, P., Purkait, S., Faruq, M., Jha, P., Mallick, S., Suri, V., Sharma, M.C., Suri, A., and Sarkar, C. (2015). Oncogenic KIAA1549-BRAF fusion with activation of the MAPK/ERK pathway in pediatric oligodendrogliomas. *Cancer Genet.* *208*, 91–95.
61. Joseph, E.W., Pratilas, C.A., Poulikakos, P.I., Tadi, M., Wang, W., Taylor, B.S., Halilovic, E., Persaud, Y., Xing, F., Viale, A., et al. (2010). The RAF inhibitor



- PLX4032 inhibits ERK signaling and tumor cell proliferation in a V600E BRAF-selective manner. *Proc. Natl. Acad. Sci. USA* 107, 14903–14908.
62. Hatzivassiliou, G., Song, K., Yen, I., Brandhuber, B.J., Anderson, D.J., Alvarado, R., Ludlam, M.J., Stokoe, D., Gloor, S.L., Vigers, G., et al. (2010). RAF inhibitors prime wild-type RAF to activate the MAPK pathway and enhance growth. *Nature* 464, 431–435.
 63. Sievert, A.J., Lang, S.-S., Boucher, K.L., Madsen, P.J., Slaunwhite, E., Choudhari, N., Kellet, M., Storm, P.B., and Resnick, A.C. (2013). Paradoxical activation and RAF inhibitor resistance of BRAF protein kinase fusions characterizing pediatric astrocytomas. *Proc. Natl. Acad. Sci. USA* 110, 5957–5962.
 64. Poulidakos, P.I., Zhang, C., Bollag, G., Shokat, K.M., and Rosen, N. (2010). RAF inhibitors transactivate RAF dimers and ERK signalling in cells with wild-type BRAF. *Nature* 464, 427–430.
 65. Holderfield, M., Nagel, T.E., and Stuart, D.D. (2014). Mechanism and consequences of RAF kinase activation by small-molecule inhibitors. *Br. J. Cancer* 111, 640–645.
 66. Kolb, E.A., Gorlick, R., Houghton, P.J., Morton, C.L., Neale, G., Keir, S.T., Carol, H., Lock, R., Phelps, D., Kang, M.H., et al. (2010). Initial testing (stage 1) of AZD6244 (ARRY-142886) by the pediatric preclinical testing program. *Pediatr. Blood Cancer* 55, 668–677.
 67. Gu, T.-L., Tothova, Z., Scheijen, B., Griffin, J.D., Gilliland, D.G., and Sternberg, D.W. (2004). NPM-ALK fusion kinase of anaplastic large-cell lymphoma regulates survival and proliferative signaling through modulation of FOXO3a. *Blood* 103, 4622–4629.
 68. Turner, S.D., Lamant, L., Kenner, L., and Brugières, L. (2016). Anaplastic large cell lymphoma in paediatric and young adult patients. *Br. J. Haematol.* 173, 560–572.
 69. Shaw, A.T., Yeap, B.Y., Solomon, B.J., Riely, G.J., Gainor, J., Engelman, J.A., Shapiro, G.L., Costa, D.B., Ou, S.H., Butaney, M., et al. (2011). Effect of crizotinib on overall survival in patients with advanced non-small-cell lung cancer harbouring ALK gene rearrangement: a retrospective analysis. *Lancet Oncol.* 12, 1004–1012.
 70. Nakagawara, A. (2001). Trk receptor tyrosine kinases: a bridge between cancer and neural development. *Cancer Lett.* 169, 107–114.
 71. Arpinati, M., Tolomelli, G., Bochicchio, M.T., Castagnetti, F., Amabile, M., Bandini, G., Bonifazi, F., Stanzani, M., Rosti, G., Martinelli, G., and Baccarani, M. (2013). Molecular monitoring of BCR-ABL transcripts after allogeneic stem cell transplantation for chronic myeloid leukemia. *Biol. Blood Marrow Transplant.* 19, 735–740.
 72. Kalinova, M., Krskova, L., Brizova, H., Kabickova, E., Kepak, T., and Kodet, R. (2008). Quantitative PCR detection of NPM/ALK fusion gene and CD30 gene expression in patients with anaplastic large cell lymphoma—residual disease monitoring and a correlation with the disease status. *Leuk. Res.* 32, 25–32.
 73. Wang, H.-P., Li, G.-X., Qiao, Z.-H., and Wang, H.-W. (2007). [Detection of tumor load in chronic myeloid leukemia during treatment with transplantation by conventional cytogenetics, nested-RT-PCR and FISH]. *J. Exp. Hematol. Chin. Assoc. Pathophysiol.* 15, 237–241.
 74. Wang, Q., Xia, J., Jia, P., Pao, W., and Zhao, Z. (2013). Application of next generation sequencing to human gene fusion detection: computational tools, features and perspectives. *Brief. Bioinform.* 14, 506–519.
 75. Lou, D.I., Hussmann, J.A., McBee, R.M., Acevedo, A., Andino, R., Press, W.H., and Sawyer, S.L. (2013). High-throughput DNA sequencing errors are reduced by orders of magnitude using circle sequencing. *Proc. Natl. Acad. Sci. USA* 110, 19872–19877.
 76. Carrara, M., Beccuti, M., Lazzarato, F., Cavallo, F., Cordero, F., Donatelli, S., and Calogero, R.A. (2013). State-of-the-art fusion-finder algorithms sensitivity and specificity. *BioMed Res. Int.* 2013, 340620.
 77. Kalyana-Sundaram, S., Shankar, S., Deroo, S., Iyer, M.K., Palanisamy, N., Chinnaiyan, A.M., and Kumar-Sinha, C. (2012). Gene fusions associated with recurrent amplicons represent a class of passenger aberrations in breast cancer. *Neoplasia* 14, 702–708.
 78. Stratton, M.R., Campbell, P.J., and Futreal, P.A. (2009). The cancer genome. *Nature* 458, 719–724.
 79. Lin, P.P., Pandey, M.K., Jin, F., Xiong, S., Deavers, M., Parant, J.M., and Lozano, G. (2008). EWS-FLI1 induces developmental abnormalities and accelerates sarcoma formation in a transgenic mouse model. *Cancer Res.* 68, 8968–8975.
 80. Pérez-Losada, J., Pintado, B., Gutiérrez-Adán, A., Flores, T., Bañares-González, B., del Campo, J.C., Martín-Martín, J.F., Battaner, E., and Sánchez-García, I. (2000). The chimeric FUS/TLS-CHOP fusion protein specifically induces liposarcomas in transgenic mice. *Oncogene* 19, 2413–2422.
 81. Pérez-Losada, J., Sánchez-Martín, M., Rodríguez-García, M.A., Pérez-Mancera, P.A., Pintado, B., Flores, T., Battaner, E., and Sánchez-García, I. (2000). Liposarcoma initiated by FUS/TLS-CHOP: the FUS/TLS domain plays a critical role in the pathogenesis of liposarcoma. *Oncogene* 19, 6015–6022.
 82. Pérez-Mancera, P.A., Vicente-Dueñas, C., González-Herrero, I., Sánchez-Martín, M., Flores-Corral, T., and Sánchez-García, I. (2007). Fat-specific FUS-DDIT3-transgenic mice establish PPARgamma inactivation is required to liposarcoma development. *Carcinogenesis* 28, 2069–2073.
 83. 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.
 84. 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.
 85. Haldar, M., Hancock, J.D., Coffin, C.M., Lessnick, S.L., and Capecchi, M.R. (2007). A conditional mouse model of synovial sarcoma: insights into a myogenic origin. *Cancer Cell* 11, 375–388.
 86. NIH (2015). NIH Childhood Genomics Workshop. <https://www.cancer.gov/types/childhood-cancers/research/childhood-genomics-workshop-summary.pdf>.
 87. Kool, M., Korshunov, A., Remke, M., Jones, D.T.W., Schlanstein, M., Northcott, P.A., Cho, Y.J., Koster, J., Schouten-van Meeteren, A., van Vuurden, D., et al. (2012). Molecular subgroups of medulloblastoma: an international meta-analysis of transcriptome, genetic aberrations, and clinical data of WNT, SHH, Group 3, and Group 4 medulloblastomas. *Acta Neuropathol.* 123, 473–484.
 88. Mi, S., Lin, M., Brouwer-Visser, J., Heim, J., Smotkin, D., Hebert, T., Gunter, M.J., Goldberg, G.L., Zheng, D., and Huang, G.S. (2016). RNA-seq identification of RACGAP1 as a metastatic driver in uterine carcinosarcoma. *Clin. Cancer Res.* 22, 4676–4686.
 89. Eleveld, T.F., Oldridge, D.A., Bernard, V., Koster, J., Daage, L.C., Diskin, S.J., Schild, L., Bentahar, N.B., Bellini, A., Chicard, M., et al. (2015). Relapsed neuroblastomas show frequent RAS-MAPK pathway mutations. *Nat. Genet.* 47, 864–871.
 90. Jaju, R.J., Fidler, C., Haas, O.A., Strickson, A.J., Watkins, F., Clark, K., Cross, N.C., Cheng, J.F., Aplan, P.D., Kearney, L., et al. (2001). A novel gene, NSD1, is fused to NUP98 in the t(5;11)(q35;p15.5) in de novo childhood acute myeloid leukemia. *Blood* 98, 1264–1267.
 91. Van Vlierberghe, P., Pieters, R., Beverloo, H.B., and Meijerink, J.P. (2008). Molecular-genetic insights in paediatric T-cell acute lymphoblastic leukaemia. *Br. J. Haematol.* 143, 153–168.
 92. Akasaka, T., Balasas, T., Russell, L.J., Sugimoto, K.J., Majid, A., Walewska, R., Karran, E.L., Brown, D.G., Cain, K., Harder, L., et al. (2007). Five members of the CEBP transcription factor family are targeted by recurrent IGH translocations in B-cell precursor acute lymphoblastic leukemia (BCP-ALL). *Blood* 109, 3451–3461.
 93. Nacheva, E., Kearney, L., Bower, M., Chaplin, T., Douek, E., Das, S., and Young, B.D. (1993). In situ hybridisation analysis of a homogeneously staining region at 11q23-24 in an acute myeloid leukaemia (M5) using yeast artificial chromosomes. *Genes Chromosomes Cancer* 7, 123–127.
 94. Russell, L.J., Akasaka, T., Majid, A., Sugimoto, K.J., Loraine Karran, E., Nagel, I., Harder, L., Claviez, A., Gesk, S., Moorman, A.V., et al. (2008). t(6;14)(p22;q32): a new recurrent IGH@ translocation involving ID4 in B-cell precursor acute lymphoblastic leukemia (BCP-ALL). *Blood* 111, 387–391.
 95. Rüssel, J., Dutta, U., Wand, D., Schlote, D., and Hansmann, I. (2009). The 9p24.3 breakpoint of a constitutional t(6;9)(p12;p24) in a patient with chronic lymphocytic leukemia maps close to the putative promoter region of the DMRT2 gene. *Cytogenet. Genome Res.* 125, 81–86.
 96. Vaishampayan, U.N., Mohamed, A.N., Dugan, M.C., Bloom, R.E., and Palutke, M. (2001). Blastic mantle cell lymphoma associated with Burkitt-type translocation and hypodiploidy. *Br. J. Haematol.* 115, 66–68.



97. Lin, P., and Medeiros, L.J. (2013). The impact of MYC rearrangements and “double hit” abnormalities in diffuse large B-cell lymphoma. *Curr. Hematol. Malig. Rep.* 8, 243–252.
98. Sewastianik, T., Prochorec-Sobieszek, M., Chapuy, B., and Juszczynski, P. (2014). MYC deregulation in lymphoid tumors: molecular mechanisms, clinical consequences and therapeutic implications. *Biochim. Biophys. Acta* 1846, 457–467.
99. van Krieken, J.H., and Kluin, P.M. (1992). [Non-Hodgkin lymphoma of the stomach as an entity. Concept formation and current developments]. *Ned. Tijdschr. Geneesk.* 136, 1200–1203.
100. Adriaansen, H.J., Soeting, P.W., Wolvers-Tettero, I.L., and van Dongen, J.J. (1991). Immunoglobulin and T-cell receptor gene rearrangements in acute non-lymphocytic leukemias. Analysis of 54 cases and a review of the literature. *Leukemia* 5, 744–751.
101. Allam, A., and Kabelitz, D. (2006). TCR trans-rearrangements: biological significance in antigen recognition vs the role as lymphoma biomarker. *J. Immunol.* 176, 5707–5712.
102. Schmidt, C.A., Przybylski, G., Seeger, K., and Siegert, W. (1995). TCR delta gene rearrangements in acute myeloid leukemia with T-lymphoid antigen expression. *Leuk. Lymphoma* 20, 45–49.
103. Shih, L.Y., Liang, D.C., Fu, J.F., Wu, J.H., Wang, P.N., Lin, T.L., Dunn, P., Kuo, M.C., Tang, T.C., Lin, T.H., and Lai, C.L. (2006). Characterization of fusion partner genes in 114 patients with de novo acute myeloid leukemia and MLL rearrangement. *Leukemia* 20, 218–223.
104. Hunger, S.P., Galili, N., Carroll, A.J., Crist, W.M., Link, M.P., and Cleary, M.L. (1991). The t(1;19)(q23;p13) results in consistent fusion of E2A and PBX1 coding sequences in acute lymphoblastic leukemias. *Blood* 77, 687–693.
105. Troussard, X., Charpentier, F., Haioun, C., Cordonnier, C., Vernant, J.P., and Rochant, H. (1985). [Bone marrow necrosis disclosing multiple myeloma of bone]. *Presse Med.* 14, 1888.
106. Domer, P.H., Fakhrazadeh, S.S., Chen, C.S., Jockel, J., Johansen, L., Silverman, G.A., Kersey, J.H., and Korsmeyer, S.J. (1993). Acute mixed-lineage leukemia t(4;11)(q21;q23) generates an MLL-AF4 fusion product. *Proc. Natl. Acad. Sci. USA* 90, 7884–7888.
107. Strehl, S., König, M., Dworzak, M.N., Kalwak, K., and Haas, O.A. (2003). PAX5/ETV6 fusion defines cytogenetic entity dic(9;12)(p13;p13). *Leukemia* 17, 1121–1123.
108. Golub, T.R., Barker, G.F., Bohlander, S.K., Hiebert, S.W., Ward, D.C., Bray-Ward, P., Morgan, E., Raimondi, S.C., Rowley, J.D., and Gilliland, D.G. (1995). Fusion of the TEL gene on 12p13 to the AML1 gene on 21q22 in acute lymphoblastic leukemia. *Proc. Natl. Acad. Sci. USA* 92, 4917–4921.
109. Dreyling, M.H., Martinez-Climent, J.A., Zheng, M., Mao, J., Rowley, J.D., and Bohlander, S.K. (1996). The t(10;11)(p13;q14) in the U937 cell line results in the fusion of the AF10 gene and CALM, encoding a new member of the AP-3 clathrin assembly protein family. *Proc. Natl. Acad. Sci. USA* 93, 4804–4809.
110. Shurtleff, S.A., Buijs, A., Behm, F.G., Rubnitz, J.E., Raimondi, S.C., Hancock, M.L., Chan, G.C., Pui, C.H., Grosveld, G., and Downing, J.R. (1995). TEL/AML1 fusion resulting from a cryptic t(12;21) is the most common genetic lesion in pediatric ALL and defines a subgroup of patients with an excellent prognosis. *Leukemia* 9, 1985–1989.
111. Erickson, P., Gao, J., Chang, K.S., Look, T., Whisenant, E., Raimondi, S., Lasher, R., Trujillo, J., Rowley, J., and Drabkin, H. (1992). Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, AML1/ETO, with similarity to Drosophila segmentation gene, runt. *Blood* 80, 1825–1831.
112. Bernstein, M.L., Esseltine, D.W., Emond, J., and Vekemans, M. (1986). Acute lymphoblastic leukemia at relapse in a child with acute myeloblastic leukemia. *Am. J. Pediatr. Hematol. Oncol.* 8, 153–157.
113. Viguie, F., Marie, J.P., Poler, F., and Bernadou, A. (1986). Three cases of preleukemic myelodysplastic disorders with the same translocation t(1;3). *Cancer Genet. Cytogenet.* 19, 213–218.
114. Yoneda-Kato, N., Look, A.T., Kirstein, M.N., Valentine, M.B., Raimondi, S.C., Cohen, K.J., Carroll, A.J., and Morris, S.W. (1996). The t(3;5)(q25.1;q34) of myelodysplastic syndrome and acute myeloid leukemia produces a novel fusion gene, NPM-MLF1. *Oncogene* 12, 265–275.
115. von Lindern, M., Fornerod, M., van Baal, S., Jaegle, M., de Wit, T., Buijs, A., and Grosveld, G. (1992). The translocation (6;9), associated with a specific subtype of acute myeloid leukemia, results in the fusion of two genes, dek and can, and the expression of a chimeric, leukemia-specific dek-can mRNA. *Mol. Cell. Biol.* 12, 1687–1697.
116. Kawamura-Saito, M., Yamazaki, Y., Kaneko, K., Kawaguchi, N., Kanda, H., Mukai, H., Gotoh, T., Motoi, T., Fukayama, M., Aburatani, H., et al. (2006). Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t(4;19)(q35;q13) translocation. *Hum. Mol. Genet.* 15, 2125–2137.
117. Peeters, P., Raynaud, S.D., Cools, J., Wlodarska, I., Grosgeorge, J., Philip, P., Monpoux, F., Van Rompaey, L., Baens, M., Van den Berghe, H., and Marynen, P. (1997). Fusion of TEL, the ETS-variant gene 6 (ETV6), to the receptor-associated kinase JAK2 as a result of t(9;12) in a lymphoid and t(9;15;12) in a myeloid leukemia. *Blood* 90, 2535–2540.
118. Gamou, T., Kitamura, E., Hosoda, F., Shimizu, K., Shinohara, K., Hayashi, Y., Nagase, T., Yokoyama, Y., and Ohki, M. (1998). The partner gene of AML1 in t(16;21) myeloid malignancies is a novel member of the MTG8(ETO) family. *Blood* 91, 4028–4037.
119. Claxton, D.F., Liu, P., Hsu, H.B., Marlton, P., Hester, J., Collins, F., Deisseroth, A.B., Rowley, J.D., and Siciliano, M.J. (1994). Detection of fusion transcripts generated by the inversion 16 chromosome in acute myelogenous leukemia. *Blood* 83, 1750–1756.
120. Borrow, J., Goddard, A.D., Sheer, D., and Solomon, E. (1990). Molecular analysis of acute promyelocytic leukemia breakpoint cluster region on chromosome 17. *Science* 249, 1577–1580.
121. Longo, L., Donti, E., Mencarelli, A., Avanzi, G., Pegoraro, L., Alimena, G., Tabilio, A., Venti, G., Grignani, F., and Pelicci, P.G. (1990). Mapping of chromosome 17 breakpoints in acute myeloid leukemias. *Oncogene* 5, 1557–1563.
122. Gelsi-Boyer, V., Trouplin, V., Adélaïde, J., Aceto, N., Remy, V., Pinson, S., Houdayer, C., Arnoulet, C., Sainty, D., Bentires-Alj, M., et al. (2008). Genome profiling of chronic myelomonocytic leukemia: frequent alterations of RAS and RUNX1 genes. *BMC Cancer* 8, 299.
123. Shtivelman, E., Lifshitz, B., Gale, R.P., and Canaan, E. (1985). Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. *Nature* 315, 550–554.
124. Stam, K., Heisterkamp, N., Grosveld, G., de Klein, A., Verma, R.S., Coleman, M., Dosik, H., and Groffen, J. (1985). Evidence of a new chimeric bcr/c-abl mRNA in patients with chronic myelocytic leukemia and the Philadelphia chromosome. *N. Engl. J. Med.* 313, 1429–1433.
125. Rabbitts, T.H. (1991). Translocations, master genes, and differences between the origins of acute and chronic leukemias. *Cell* 67, 641–644.
126. Huret, J.L., Brizard, A., Slater, R., Charrin, C., Bertheas, M.F., Guilhot, F., Hählen, K., Kroes, W., van Leeuwen, E., Schoot, E.V., et al. (1993). Cytogenetic heterogeneity in t(11;19) acute leukemia: clinical, hematological and cytogenetic analyses of 48 patients—updated published cases and 16 new observations. *Leukemia* 7, 152–160.
127. Meyer, C., Hofmann, J., Burmeister, T., Gröger, D., Park, T.S., Emerenciano, M., Pombó de Oliveira, M., Renneville, A., Villarese, P., Macintyre, E., et al. (2013). The MLL recombinome of acute leukemias in 2013. *Leukemia* 27, 2165–2176.
128. Xiao, S., Nalabolu, S.R., Aster, J.C., Ma, J., Abruzzo, L., Jaffe, E.S., Stone, R., Weissman, S.M., Hudson, T.J., and Fletcher, J.A. (1998). FGFR1 is fused with a novel zinc-finger gene, ZNF198, in the t(8;13) leukaemia/lymphoma syndrome. *Nat. Genet.* 18, 84–87.
129. Sonnier, J.A., Buchanan, G.R., Howard-Peebles, P.N., Rutledge, J., and Smith, R.G. (1983). Chromosomal translocation involving the immunoglobulin kappa-chain and heavy-chain loci in a child with chronic lymphocytic leukemia. *N. Engl. J. Med.* 309, 590–594.
130. Willis, T.G., Jayadel, D.M., Du, M.Q., Peng, H., Perry, A.R., Abdul-Rauf, M., Price, H., Karran, L., Majekodunmi, O., Wlodarska, I., et al. (1999). Bcl10 is involved in t(1;14)(p22;q32) of MALT B cell lymphoma and mutated in multiple tumor types. *Cell* 96, 35–45.
131. Wlodarska, I., Veyt, E., De Paepe, P., Vandenberghe, P., Nooijen, P., Theate, L., Michaux, L., Sagaert, X., Marynen, P., Hagemeijer, A., and De Wolf-Peters, C. (2005). FOXP1, a gene highly expressed in a subset of diffuse large B-cell lymphoma, is recurrently targeted by genomic aberrations. *Leukemia* 19, 1299–1305.



132. Corcoran, M.M., Mould, S.J., Orchard, J.A., Ibbotson, R.E., Chapman, R.M., Boright, A.P., Platt, C., Tsui, L.C., Scherer, S.W., and Oscier, D.G. (1999). Dysregulation of cyclin dependent kinase 6 expression in splenic marginal zone lymphoma through chromosome 7q translocations. *Oncogene* 18, 6271–6277.
133. Hillion, J., Mecucci, C., Aventin, A., Leroux, D., Wlodarska, I., Van Den Berghe, H., and Larsen, C.J. (1991). A variant translocation t(2;18) in follicular lymphoma involves the 5' end of bcl-2 and Ig kappa light chain gene. *Oncogene* 6, 169–172.
134. Seité, P., Leroux, D., Hillion, J., Monteil, M., Berger, R., Mathieu-Mahul, D., and Larsen, C.J. (1993). Molecular analysis of a variant 18;22 translocation in a case of lymphocytic lymphoma. *Genes Chromosomes Cancer* 6, 39–44.
135. Sümegi, J., Spira, J., Bazin, H., Szpirer, J., Levan, G., and Klein, G. (1983). Rat c-myc oncogene is located on chromosome 7 and rearranges in immunocytomas with t(6;7) chromosomal translocation. *Nature* 306, 497–498.
136. Baron, B.W., Nucifora, G., McCabe, N., Espinosa, R., 3rd, Le Beau, M.M., and McKeithan, T.W. (1993). Identification of the gene associated with the recurring chromosomal translocations t(3;14)(q27;q32) and t(3;22)(q27;q11) in B-cell lymphomas. *Proc. Natl. Acad. Sci. USA* 90, 5262–5266.
137. Xu, J., Chernos, J.E., Bernier, F., and Lowry, R.B. (2000). Characterization of an interstitial deletion del(13)(q22q32) using microdissection and sequential FISH and G-banding. *Genet. Test.* 4, 279–282.
138. Brito-Babapulle, V., Ellis, J., Matutes, E., Oscier, D., Khokhar, T., MacLennan, K., and Catovsky, D. (1992). Translocation t(11;14)(q13;q32) in chronic lymphoid disorders. *Genes Chromosomes Cancer* 5, 158–165.
139. Iida, S., Rao, P.H., Nallasivam, P., Hibshoosh, H., Butler, M., Louie, D.C., Dyomin, V., Ohno, H., Chaganti, R.S., and Dalla-Favera, R. (1996). The t(9;14)(p13;q32) chromosomal translocation associated with lymphoplasmacytoid lymphoma involves the PAX-5 gene. *Blood* 88, 4110–4117.
140. Ye, B.H., Chaganti, S., Chang, C.C., Niu, H., Corradini, P., Chaganti, R.S., and Dalla-Favera, R. (1995). Chromosomal translocations cause deregulated BCL6 expression by promoter substitution in B cell lymphoma. *EMBO J.* 14, 6209–6217.
141. Zech, L., Gahrton, G., Hammarström, L., Juliusson, G., Mellstedt, H., Robèrt, K.H., and Smith, C.I. (1984). Inversion of chromosome 14 marks human T-cell chronic lymphocytic leukaemia. *Nature* 308, 858–860.
142. Charest, A., Lane, K., McMahon, K., Park, J., Preisinger, E., Conroy, H., and Housman, D. (2003). Fusion of FIG to the receptor tyrosine kinase ROS in a glioblastoma with an interstitial del(6)(q21q21). *Genes Chromosomes Cancer* 37, 58–71.
143. Honeyman, J.N., Simon, E.P., Robine, N., Chiaroni-Clarke, R., Lim, I.L.P., Gleason, C.E., Murphy, J.M., Rosenberg, B.R., Teegan, L., et al. (2014). Detection of a recurrent *DNAJB1-PRKACA* chimeric transcript in fibrolamellar hepatocellular carcinoma. *Science* 343, 1010–1014.
144. Bridge, J.A., Liu, X.-Q., Sumegi, J., Nelson, M., Reyes, C., Bruch, L.A., Rosenblum, M., Puccioni, M.J., Bowdino, B.S., and McComb, R.D. (2013). Identification of a novel, recurrent SLC44A1-PRKCA fusion in papillary glioneuronal tumor. *Brain Pathol* 23, 121–128.
145. Kuiper, R.P., Schepens, M., Thijssen, J., van Asseldonk, M., van den Berg, E., Bridge, J., Schuurin, E., Schoenmakers, E.F.P.M., and van Kessel, A.G. (2013). Upregulation of the transcription factor TFEB in t(6;11)(p21;q13)-positive renal cell carcinomas due to promoter substitution. *Hum Mol Genet.* 12, 1661–1669.
146. Argani, P., Lui, M.Y., Couturier, J., Bouvier, R., Fournet, J.-C., and Ladanyi, M. (2003). A novel CLTC-TFE3 gene fusion in pediatric renal adenocarcinoma with t(X;17)(p11.2;q23). *Oncogene* 22, 5374–5378.
147. Clark, J., Lu, Y.-J., Sidhar, S.K., Parker, C., Gill, S., Smedley, D., Hamoudi, R., Linehan, W.M., Shipley, J., and Cooper, C.S. (1997). Fusion of splicing factor genes PSF and NonO (p54nrb) to the TFE3 gene in papillary renal cell carcinoma. *Oncogene* 15, 2233–2239.
148. Heimann, P., El Housni, H., Ogur, G., Weterman, M.A.J., Petty, E.M., and Vassart, G. (2001). Fusion of a novel gene, RCC17, to the TFE3 gene in t(X;17)(p11.2;q25.3)-bearing papillary renal cell carcinomas. *Cancer Res.* 61, 4130–4135.
149. Oliveira, A.M., Perez-Atayde, A.R., Dal Cin, P., Gebhardt, M.C., Chen, C.J., Neff, J.R., Demetri, G.D., Rosenberg, A.E., Bridge, J.A., and Fletcher, J.A. (2005). Aneurysmal bone cyst variant translocations upregulate USP6 transcription by promoter swapping with the ZNF9, COL1A1, TRAP150, and OMD genes. *Oncogene* 24, 3419–3426.
150. Antonescu, C.R., Zhang, L., Nielsen, G.P., Rosenberg, A.E., Dal Cin, P., and Fletcher, C.D. (2011). Consistent t(1;10) with rearrangements of TGFBR3 and MGEA5 in both myxoinflammatory fibroblastic sarcoma and hemosiderotic fibrolipomatous tumor. *Genes Chromosomes Cancer* 50, 757–764.
151. La Starza, R., Barba, G., Nofrini, V., Pierini, T., Pierini, V., Marcomigni, L., Perruccio, K., Matteucci, C., Storlazzi, C.T., Daniele, G., Crescenzi, B., et al. Multiple EWSR1-WT1 and WT1-EWSR1 copies in two cases of desmoplastic round cell tumor. *Cancer Genet.* 206, 387–392.
152. Hibshoosh, H., and Lattes, R. (1997). Immunohistochemical and molecular genetic approaches to soft tissue tumor diagnosis: a primer. *Semin. Oncol.* 24, 515–525.
153. Mendlick, M.R., Nelson, M., Pickering, D., Johansson, S.L., Seemayer, T.A., Neff, J.R., Vergara, G., Rosenthal, H., and Bridge, J.A. (2001). Translocation t(1;3)(p36.3;q25) is a nonrandom aberration in epithelioid hemangioendothelioma. *Am J Surg Pathol.* 25, 684–687.
154. Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., de Jong, P., Rouleau, G., et al. (1992). Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* 359, 162–165.
155. Ohno, T., Ouchida, M., Lee, L., Gatalica, Z., Rao, V.N., and Reddy, E.S. (1994). The EWS gene, involved in Ewing family of tumors, malignant melanoma of soft parts and desmoplastic small round cell tumors, codes for an RNA binding protein with novel regulatory domains. *Oncogene* 9, 3087–3097.
156. Ichikawa, H., Shimizu, K., Hayashi, Y., and Ohki, M. (1994). An RNA-binding protein gene, TLS/FUS, is fused to ERG in human myeloid leukemia with t(16;21) chromosomal translocation. *Cancer Res.* 54, 2865–2868.
157. Knezevich, S.R., McFadden, D.E., Tao, W., Lim, J.F., and Sorensen, P.H. (1998). A novel ETV6-NTRK3 gene fusion in congenital fibrosarcoma. *Nat Genet.* 18, 184–187.
158. Rabbitts, T.H., Forster, A., Larson, R., and Nathan, P. (1993). Fusion of the dominant negative transcription regulator CHOP with a novel gene FUS by translocation t(12;16) in malignant liposarcoma. *Nat Genet.* 4, 175–180.
159. Alaggio, R., Zhang, L., Sung, Y.S., Huang, S.C., Chen, C.L., Bisogno, G., Zin, A., Agaram, N.P., LaQuaglia, M.P., Wexler, L.H., and Antonescu, C.R. (2016). A molecular study of pediatric spindle and sclerosing rhabdomyosarcoma: identification of novel and recurrent VGLL2-related fusions in infantile cases. *Am J Surg Pathol.* 40, 224–235.
160. Jin, Y., Möller, E., Nord, K.H., Mandahl, N., Von Steyern, F.V., Domanski, H.A., Mariño-Enríquez, A., Magnusson, L., Nilsson, J., Sciot, R., et al. (2012). Fusion of the AHRR and NCOA2 genes through a recurrent translocation t(5;8)(p15;q13) in soft tissue angiofibroma results in upregulation of aryl hydrocarbon receptor target genes. *Genes Chromosomes Cancer* 51, 510–520.
161. O'Meara, E., Stack, D., Phelan, S., McDonagh, N., Kelly, L., Sciot, R., Debiec-Rychter, M., Morris, T., Cochrane, D., Sorensen, P., and O'Sullivan, M.J. (2014). Identification of an MLL4-GPS2 fusion as an oncogenic driver of undifferentiated spindle cell sarcoma in a child. *Genes Chromosomes Cancer* 53, 991–998.
162. Storlazzi, C.T., Anelli, L., Albano, F., Zagaria, A., Ventura, M., Rocchi, M., Panagopoulos, I., Pannunzio, A., Ottaviani, E., Liso, V., and Specchia, G. (2004). A novel chromosomal translocation t(3;7)(q26;q21) in myeloid leukemia resulting in overexpression of EVI1. *Ann Hematol.* 83, 78–83.
163. Kawamura-Saito, M., Yamazaki, Y., Kaneko, K., Kawaguchi, N., Kanda, H., Mukai, H., Gotoh, T., Motoi, T., Fukayama, M., Aburatani, H., et al. (2006). Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t(4;19)(q35;q13) translocation. *Hum Mol Genet.* 15, 2125–2137.

RESULTAT I : Mise au point d'un outil de détection de fusions à partir de données RNA-seq et découverte de nouveaux transcrits de fusion dans une cohorte de patients pédiatriques en rechute

A. Introduction de l'article

Etant données leurs particularités biologiques, nous avons supposé que les cancers pédiatriques seraient plus probablement dus à la présence de transcrits de fusion oncogéniques. Les répercussions cliniques majeures de la découverte de fusions oncogéniques, combinées avec le manque de données chez les patients pédiatriques en rechute, le premier but de mes travaux de thèse était d'établir une méthode robuste de détection de fusions *in silico*, et d'identifier de nouveaux transcrits de fusion oncogéniques dans des cancers pédiatriques en rechute.

Pour cela, nous avons mené une étude sur 48 patients pédiatriques présentant différentes tumeurs malignes solides en rechute ou résistantes, inclus dans l'essai moléculaire de précision MOSCATO-01 mené à Gustave Roussy. En utilisant les données de séquençage de l'ARN (RNA-seq) et en mettant au point une nouvelle pipeline bioinformatique pour la détection de fusions, suivi d'une validation par PCR, nous avons identifié et classé les fusions candidates les plus pertinentes pour une application clinique.

Grâce à cette analyse, nous avons pu observer que la majorité des fusions détectées affectaient les principales voies de signalisation impliquées dans les processus d'oncogénèse. Contrairement aux études précédentes chez l'adulte, aucune corrélation significative entre le nombre de fusions et le nombre de points de cassures génomiques n'a été observée, soulignant la particularité d'étudier des patients pédiatriques prétraités en rechute. Une proportion considérable de transcrits de fusions contenant des gènes suppresseurs de tumeurs a été détectée, reflétant leur importance dans les cancers pédiatriques. Peu de transcrits de fusions contenant des tyrosines kinases ont été retrouvés, et majoritairement dans les tumeurs cérébrales. Enfin et de manière remarquable, plus de 30% des patients présentaient une fusion pertinente ayant potentiellement un impact thérapeutique.

En conclusion, grâce à un nouvel outil robuste, nous avons détecté de nouveaux transcrits de fusions potentiellement oncogéniques chez des patients atteints de cancers pédiatriques en rechute. Cet outil peut être appliqué à d'autres tumeurs malignes, pour détecter et prioriser les études de validation expérimentale menant au développement de nouvelles options thérapeutiques pour les patients présentant des oncogènes de fusion.

B. Article 1: "Discovery of new fusion transcripts in a cohort of pediatric solid cancers at relapse and relevance for personalized medicine"

Discovery of new fusion transcripts in a cohort of pediatric solid cancers at relapse and relevance for personalized medicine

Célia Dupain^{1,2,3&}, Anne C. Harttrampf^{1,2,3&}, Yannick Boursin⁴, Manuel Lebeurrier⁴, Windy Rondof^{1,2,3,4}, Guillaume Robert-Siegwald⁴, Pierre Khoueiry⁵, Birgit Georger^{1,2,3,6} and Liliane Massaad-Massade^{1,2,3*}

¹ Université Paris-Sud 11, Laboratoire de Vectorologie et Thérapeutiques Anticancéreuses, UMR 8203, Villejuif, France-94805

² CNRS, Villejuif, Laboratoire de Vectorologie et Thérapeutiques Anticancéreuses, UMR 8203, France-94805.

³ Gustave Roussy, Laboratoire de Vectorologie et Thérapeutiques Anticancéreuses, UMR 8203, Villejuif, France-94805.

⁴ Gustave Roussy, Bioinformatics Platform, AMMICA, INSERM US23/CNRS UMS3655, Villejuif, France.

⁵ American University of Beirut, Faculty of Medicine, Dept. of Biochemistry and Molecular

⁶ Gustave Roussy, Department of Pediatric and Adolescent Oncology, Villejuif, France
Genetics, P.O.BOX 11-0236 DTS 419-B, Bliss Street, Beirut-Lebanon

&: both authors contributed equally to this work

*: correspondance to Liliane MASSADE, PhD, Laboratoire de Vectorologie et Thérapeutiques Anticancéreuses, UMR 8203, 114 rue Edouard Vaillant, Villejuif, France-94805

Running title: Fusion transcripts in pediatric solid cancers at relapse.

Key words: fusion transcripts, pediatric solid cancers, personalized medicine, RNA-sequencing, relapsed patients, molecular profiling.

Abstract

We hypothesized that pediatric cancers would be more likely due to fusion transcripts. To dissect the complexity of the fusion landscape in recurrent solid pediatric cancers, we conducted a study on 48 pediatric patients with different relapsing/resistant malignancies. By analysing RNA-sequencing data with a new in-house pipeline for fusions detection named ChimComp, followed by verification by real time PCR of the fusion transcripts (FT), we identified and classified the most confident candidates according to their biological function prediction and potential druggability. The majority of detected FT were predicted to affect key cancer pathways and described to be involved in oncogenesis. Contrary to previous descriptions in adults, we found no significant correlation between the number of fusions and genomic breakpoints, emphasizing the particularity to study resistant, pre-treated pediatric patients. A considerable proportion of FT containing tumor suppressor genes was detected, reflecting their importance in pediatric cancers. FT containing non-receptor tyrosine kinases occurred at low incidence and were predominant in brain tumors. Remarkably, more than 30% of patients presented a potentially druggable high-confidence fusion.

In conclusion, we detected new oncogenic FT in relapsing pediatric cancer patients by establishing a robust pipeline and that can be applied to other malignancies, to detect and prioritize experimental validation studies leading to the development of new therapeutic options.

Introduction

Fusion transcripts, or chimeric RNA, result from the juxtaposition of two genes, previously located separately from each other, due to chromosomal or non-chromosomal events. They can be the consequence of structural chromosomal rearrangements or be the product of alternative (cis-, trans-) splicing or transcriptional readthroughs [1]. Fusion transcripts can lead to the activation of proto-oncogenes or inactivation of tumor suppressor genes (TSGs) and are considered as one of the leading mechanism responsible of carcinogenesis [2].

They present a great interest in clinics as they are increasingly explored as potential therapeutic targets fueled by the discovery of a drug, imatinib mesylate, able to target BCR-ABL fusion oncoprotein [3]. Several studies aimed at targetting fusion products in various neoplasms including i) inhibitors of anaplastic lymphoma kinase (ALK) in EML4-ALK positive lung cancer or NPM-ALK positive anaplastic large-cell lymphomas; ii) larotrectinib in NTRK-containing fusions in various solid tumors [4] and iii) FGFR inhibitors to target FGFR-TACC fusion detected in 3% of IDH1/2 wild-type glioblastoma and other malignancies [5].

Fusion transcripts are considered as strong diagnostic biomarkers with several examples in this regard. Indeed, EWS-FLI1 found in Ewing's sarcoma or PAX3-FOXO1 responsible for 55–70% of alveolar rhabdomyosarcoma are characteristic disease-defining fusion transcripts. The detection of DNAJB1-PRKCA gene fusion in fibrolamellar hepatocellular carcinoma (FLHCC) serves as illustrative example for both the powerful detection at unprecedented resolution by RNA-sequencing (RNA-seq) and its successful diagnostic use in extremely rare tumors [6]. Therefore, the discovery of additional actionable fusion transcripts in cancers is attracting more attention.

High throughput sequencing of RNA (RNA-seq) is a powerful tool to detect so far undiscovered fusion events. On the other hand, as reflected by the dramatic increase of reported gene fusions since its introduction, this will likely include rearrangements which represent secondary, "passenger events", due to the genomically unstable nature of cancer, and non-recurrent events in individual patients whose biological significance is currently unclear [1]. Hence, it is important to keep the balance between stringent selection parameters to avoid false positives without missing relevant oncogenic fusions [7,8].

Up to date, there exist more than 30 different bioinformatics tools for fusion detection [9], each differing in terms of sensitivity and specificity. They majorly differ by their filtering approaches and algorithms regarding, among others, the distance between pair-end reads, the number of nucleotides mapping each side of a fusion breakpoint, the presence of read-through transcripts, the presence of duplicated reads and the number of reads mapping to homologous or repetitive regions. In numerous benchmarking studies, none of the existing tools has clearly outperformed [7,10,11], therefore, a meta-caller approach by combining different tools has been shown to

increase both specificity and sensitivity [10].

Additionally, in order to translate bioinformatics results into meaningful biology, predictive algorithms such as Oncofuse [12] have been developed to predict the oncogenicity of the detected fusions. This can help to select candidates with higher confidence for further experimental exploration, thus reducing financial and human burdens.

Solid pediatric cancers represent around 60% of all pediatric neoplasms and include the cases of malignancies diagnosed under 14 years-old. Around 80% of pediatric patients reach long-term remission after treatment while 20% of them die due to treatment failure. It is well-known that pediatric cancers differ from adults as many arise from embryonal rather than epithelial cells [13] and as they particularly harbor low mutation rates, mostly due to the lower implication of exogenous toxic effects [6]. The exploration of the fusions landscape in these pathologies could therefore help to unravel the particularity of these malignancies and provide new therapeutic targets.

Given the limited data in solid pediatric cancer patients at relapse, we conducted our study using previously published RNA-seq data of 48 pediatric patients with different malignancies, generated as part of the “Molecular Screening for Cancer Treatment Optimization” (MOSCATO-01, NCT01566019) trial conducted at Gustave Roussy [14]. Our aim was to identify new fusion transcripts in recurrent pediatric cancers and to predict their probable oncogenic potential and targetability. For this, we established a reliable *in silico* detection method named “ChimComp” that integrates information from three different fusion detection algorithms, along with their corresponding oncogenic effects and targetability. We then confirmed 91% of the tested candidates and identified new potential targetable oncogenic fusions.

Results

Patients characteristics

Clinical characteristics of patients are provided in **Table S1**. Briefly, the transcriptome of 48 patients included in MOSCATO-01 clinical trial with 22 different tumor entities at relapse or resistance was previously sequenced using RNA-seq [14]. The median age was 10.9 years (range, 0.8 - 23.3y) and the sex ratio was equal to 0.71 (28 males). Patients were classified into 3 categories based on their tumor types: sarcomas (40%), brain tumors (35%) and other types (25%). All patients had undergone antineoplastic treatment (median number of lines 2; range, 1 - 8). The median number of tumor cells in the analyzed specimen was high (around 80%).

ChimComp is a powerful pipeline to detect fusions from RNA-Seq data

Pediatric cancers are known to harbor low mutational rate which suggests that other genomic alterations, including among others fusion transcripts and structural variants, may be causative. In order to increase ability to detect true positive events from RNA-seq data, we devised a computational approach called “ChimComp” (**Figure 1**) that combines fusion transcript calls from three vastly used algorithms, Defuse [15], FusionCatcher [16] and TopHat [17] (Material and methods). All three tools have been used previously and showed diverse specificity and sensitivity in calling fusion transcript from RNA-seq data.

Using our ChimComp approach, we detected a total of 1,374 fusions in the 48 patients analyzed. The median number of fusions per patient detected was of 49.5 fusions (range, 3 - 223) (**Table 1**). Of those 1,374 fusions, 787 (57.2%) were intra-chromosomal and 588 (42.8%) inter-chromosomal. Only 8% (107/1,374) were already known and annotated in at least one of the databases, suggesting that the majority of fusion transcripts called here are de-novo and potentially positive.

We then assessed the presence of known oncogenes and tumor suppressor genes (TSG) publicly available database [18–20] (**Figure 1**). A total of 91/1,374 (7%) fusions were found to contain an oncogene (**Table S2**) and 137/1,374 (11%) a TSG (**Table S3**). Interestingly, every patient had at least one fusion containing either a known oncogene or TSG (**Table 1**), except patient 45 which had very few fusions detected. Nervous system tumors were the one with the highest rate of oncogenes (ependymoma, HGG, neuroblastoma, medulloblastoma, DIPG) while tumors with the highest number of TSGs were of various origins. A strong correlation between the total number of fusions and the number of fusions containing an oncogene or a TSG per patient was found (respectively: Spearman coefficient $r=0.4637$, $p\text{-val}=0.001$ and $r=0.6948$, $p\text{-val}<0.0001$).

When considering the output of each software independently, “Defuse” predicted the largest number of fusions followed by “TopHat” whereas “Fusion Catcher” software was the most conservative. A large number of fusions was exclusively detected by either “Defuse” or “TopHat”

(29.8% and 9.5% respectively) and when combined, these 2 tools share the highest number of fusion genes overlap (46.4%) (**Figure 2A**).

In order to define a high-confidence group of fusion transcripts, we combined the calls of the three algorithms by considering intersecting events. This identified a set of 118 fusion (9%; 118/1,374) of which 107 (90%) have not been previously reported (**Table S4**). The repartition of fusions in intra or inter-chromosomal for the common set followed the same tendency observed in the total fusions with 66/118 (55.9%) intra-chromosomal fusions and 52/118 (44.1%) inter-chromosomal fusions. The median spanning reads number (representing the reads aligning to the junction sequence) is 12.7 per fusion, emphasizing the high confidence of call in this group as they are mainly supported by high read numbers. Forty five fusions (38%) were predicted in-frame by Fusion Catcher, and 47/118 (39%) fusions had one or both genes known to be involved in tumorigenicity suggesting a fundamental role on oncogenicity and tumor progression (**Table S4**). Interestingly, among the cancer pathologies included in the cohort, medulloblastoma and glioblastoma were highly represented in the high-confidence fusion candidates detected which comprise 2 patients with highly rearranged transcriptome (16 and 48). Moreover, all the known fusions characterizing a cancer pathology (e.g. EWS/FLI-1 in EWS, patients 11 and 46 or HEY1-NCOA2 in mesenchymal chondrosarcoma, patient 33) were detected by our pipeline, reflecting the robustness of the ChimComp approach.

To identify fusions predicted to have a potential driving role, we scored detected fusions for their oncogenic potential using “Oncofuse” [12] and categorized all 1,374 fusions into three groups: fusions with high oncoscore (oncoscore > 0.7; 157/1,374 or 11.5%), fusions with a mild oncoscore ($0.35 < \text{oncoscore} < 0.7$; 45/1,374 or 3.3%) and fusions with a low oncoscore (oncoscore < 0.35; 258/1,374 or 18.8%) (**Figure 2B**). A total of 66.5% fusions (914/1,374) did not have an oncoscore, most likely due to limitations of “Oncofuse”. Interestingly, the percentage of fusions with oncoscores significantly increased (z-score < 1.96) when we considered the set of 118 high-confidence fusions (**Figure 2B**). For instance, 36.4% (43/118) of fusions jointly identified by the 3 tools exhibit a high oncoscore compared to 11.5% only (157/1,374), when all fusions were considered (**Figure 2B**). Similarly, 30.5% and 9.3% of fusions from the joint set belonged to categories of low and mild oncoscore respectively, compared to 18.7% (258/1,374) and 3.3% (45/1,374) respectively when all fusions were considered. (**Figure 2B**). Regarding fusions containing an oncogene, 29/91 (31.8%) had a high oncoscore (>0.7) and 47/91 (51.6%) had no oncoscore. For the TSG-containing fusions, 30/148 (20.3%) had a high oncoscore and 76/148 (51.4%) had no oncoscore.

Finally, in order to get functional insights for the identified fusions, we predicted the biological function of each fusion and classified them according to the “hallmarks of cancer” described by Hanahan et al., 2011 [21]. The major biological alterations identified were “sustainability to proliferative signaling” (18.1%), “resistance to cell death” (10.5%) and “evasion to growth suppressors” (10.5%) (**Figure 2C**). This suggests the involvement of these fusions in oncogenesis. Interestingly, 6.8% of fusions could affect several categories of biological functions, leading to possible multiple consequences. For a large proportion of fusions (17.3%) no classification could be attributed. Remarkably, those fusions were found to have a low or no oncoscore.

The above analysis showed that combining different fusion calling tools associated with a valid scoring scheme (Oncoscore) increases our chances in calling potentially active and actionable fusion transcripts.

The occurrence of fusions in relapsed pediatric cancer patients is not related to the number of genomic breakpoints or mutations

Previous studies reported that the number of fusion events and genomic instability were positively correlated while the number of mutations and fusion events were negatively correlated [1,22,23]. Taking advantage of the RNA-seq, WES and CGH data available for our samples [14], we asked whether our cohort of pediatric cancer patients recall known correlations or not.

The mutational burden (number of somatic non-synonymous mutations) as well as the number of genomic breakpoints were calculated from WES and CGH arrays respectively [14] (**Material and Methods**). Consistent with previous findings in pediatric tumor patients [24], the majority of our samples harbored a low mutation rate with a median of somatic non-synonymous mutations of 81 per patient (1.8 mutations/Mb) [14] (**Table 1**) and a mean number of chromosomal breakpoints of 670 per patient. In order to link the mutational rate and breakpoints number to the number of fusions, we used Spearman correlation coefficient at $\alpha = 5\%$. No correlation was found between the number of somatic mutations and fusions (**Figure 3A**) (Spearman coefficient $r = 0.027$ and $p\text{-value} = 0.857$) or between the number of breakpoints and fusions (when discarding the outlier patient number 48, Spearman coefficient $r = 0.246$, $p\text{-value} = 0.099$) (**Figure 3B**).

In-vivo validation identified a set of de-novo fusion transcripts

In order to validate our approach for fusion transcription identification, we aimed at checking by RT-qPCR a list of 42 fusion transcripts chosen among the highly confident set (detected by the 3 tools) and among fusions of lower confidence (**Table 2**). For this, we randomly chose a set of 35 predicted fusions transcripts in addition to 7 fusion transcripts previously identified [25–30]. Interestingly, 38/42 (90.5%) fusion transcripts were successfully validated in patients biopsies

including the 7 known cases. Among the transcripts identified, only 14.4% displayed a low mean spanning reads number (<5) strengthening the reliability of the high-confidence group defined. Remarkably, when we considered the high confidence set tested cases (29 fusions tested in total, 13 with high and intermediate oncoscore and 16 with low or no oncoscore), 100% of fusions with high and intermediate oncoscore (oncoscore > 0.35) were validated (13/13) whereas in the group with low or no oncoscore (<0.35), 3/16 (18%) could not be confirmed by RT-qPCR (*SAP30BP-CDRT4*, *ABR-DNAH2* and *C15orf57-CBX3* (**Table 2**, higher panel). *C15orf57-CBX3*, observed in 21 patients with a low oncoscore, was not validated thereby most likely representing a false-positive. The other 2 transcripts that could not be confirmed had >30 spanning reads but a low or no oncoscore (*ABR-DNAH2* and *SAP30BP-CDRT4*, respectively; **Table 2**).

In addition to the high-confidence candidates, 8 fusions detected by less than 3 tools and harboring a high oncoscore including one control healthy fusion (*LPP-TPRG1*) were also tested by RT-qPCR. Seven out of the eight tested fusions (88%) were successfully validated (**Table 2**, intermediate panel).

The fact that 100% of high confident fusions exhibiting a high oncoscore were validated positively indicates that our approach of intersecting call from different fusion transcripts algorithms coupled with a filtering based on oncoscore is a successful procedure to detect valid fusion transcripts.

Fusion oncogenes involving tyrosine kinases domains are mainly detected in brain tumors

“Targetable” and “actionable” fusions generally include tyrosine kinases genes (TK). We checked for the presence of TK genes among the list of 1,374 fusions identified.

Only 5/ 1,374 fusions (0.36%) involved known tyrosine kinases (**Table S5**). They correspond to *FES-MAN2A2*, *CCM2L-HCK*, *JAK2-PTPRD*, *MANBAL-SRC* and *TYRO3-KLHL18* (**Figure 4**) and were mainly detected by 2 tools, except for *JAK2-PTPRD*, detected by 3 tools and belonging to the high confidence set. All those fusions were validated by RT-qPCR followed and Sanger sequencing (**Table 2**). Of interest, they were mainly detected in brain tumors and with a high oncoscore (> 0.7) except for *FES-MAN2A2* that was found in a neuroblastoma and had no oncoscore. Moreover, within the same patient, these fusion oncogenes were found to have the highest oncogenic score compared to the other fusions (**Table S5**).

Fusions of potential therapeutic interest with currently available drugs

Nineteen fusion belonging to the high-confidence list and affecting 33% (16/48) of patients were identified as potentially targetable or contributing to drug resistance, being thus important for drug choice in the individual patient (**Table 3**). Fifteen out of nineteen fusions had a very high

oncoscore (> 0.89). The total list contains four well-known disease-defining transcripts as well as two other that have been listed in fusion databases before. The remaining 13 transcripts constitute individual events, occurring in 8 patients.

Five transcripts (5/19) occurred in patient 16 with a highly rearranged medulloblastoma, three of which concerned the ERBB signaling axis through different mechanisms, including two *via* *MGMT*. Further two patients with neuroblastoma (patient 26) and alveolar rhabdomyosarcoma (patient 14), presented two potentially targetable transcripts respectively. Eight of nineteen transcripts (42%) were part of the validation study and were thus confirmed, including 6 newly identified fusions that were not described before.

Potentially druggable effects for these 19 transcripts include: i) changes in drug sensitivity (*e.g.* *MGMT*/temozolomide; transcriptional overexpression of *ABCB1*, leading to multidrug resistance), ii) overexpression/overactivation of receptor tyrosine kinases and downstream pathways activation (*e.g.* *MET* upregulation by chimeric transcript factor generated by *ASPSCR1-TFE3*; overexpression of ERBB receptor family ligands as *NRG2*) and iii) synthetic lethality (*e.g.* inhibition of *WEE1* tyrosine kinase in *TP53*-deficient cells). The classification according to the biological function of each oncogenic fusion transcript was done in regard to the literature [31–50].

Further, a patient with osteosarcoma (patient 17) was selected for validation of 3 potentially targetable fusions based on the presence of genes who with their related pathways are strongly linked to osteosarcoma biology (*TP53-ZNF565*, *UBE4B-CTNNBIP1*, *NF2-KIAA0368*). These fusion transcripts have an oncoscore >0.9 and interestingly found by three, two or just one tool. They were also successfully validated by RT-qPCR (**Table 2**) and are discussed in supplementary information document (**SI 3**).

Discussion

With the aim to discover new fusion oncogenes in pediatric patients we carried out a retrospective study in a cohort of relapsed/resistant patients included in the MOSCATO-01 clinical trial [14]. The in-house analysis package ChimComp was created to reliably detect fusion transcripts in a unique cohort of pediatric patients with solely solid tumors.

We primarily concentrated on fusion candidates from the “high confidence group” detected by three tools representing 118/ 1,374 fusion transcripts (8.6%). To differentiate potential driving from passenger fusions, they were classified into four groups according to their oncoscore values. Interestingly, among them a misbalance in favor of high or low oncoscore (53.8% and 37%, respectively) was observed when compared to the distribution of all detected fusions. Remarkably, 100% of the selected candidates with high (>0.7) and intermediate (>0.35-0.7) oncoscores could be validated by RT-qPCR whereas 19% of selected fusion transcripts with low or no oncoscore could not. This observation seems to hold also for the group of fusions found by less than three tools with a validation rate of 87.5%. However, it must be considered that the selection of these fusions was mainly driven by biological interest and did not include fusions with a low/no oncoscore except the LPP-TPRG1 transcript which is known in healthy tissues. We observed a tendency towards successful validation of candidates with a high oncoscore rather than high read numbers. This observation is in line with emerging evidence that a fusion’s true existence cannot solely rely on the number of reads supporting it, this number being dependent on mapping parameters and sequencing depth [51]. However, we noticed certain limits of Oncofuse: i) inability to score the majority of fusion transcripts (herein 914 fusions (66%), ii) variation of oncoscore according to the orientation of both genes and iii) undervaluation of affected tumor suppressor genes. Regarding possible reasons for not validating events, this might include alignment errors, showing that the use of “meta-caller” pipeline alone does not allow to reach a “perfectly reliable” analysis. A visualization tool appears to be one of the solutions to support a fusion’s real existence at the highest in silico confidence level. Our fusions detection method also reached its limits as i) results would need to be compared to the respective healthy tissue of the individual patient and ii) shifts in the open reading frames cannot be reliably predicted on the bioinformatic level.

Interestingly, as seen for SAP30BP-CDRT4 and ABR-DNAH2 fusions, a high read number (>50) did not guarantee biological validation. As highlighted by Gao et. al, the use of 3 algorithms therefore gives more confidence to fusions with lower levels of read support that might otherwise have been discarded [52].

The total number of fusions detected per patient was particularly high. Moreover, we found no significant correlation between the number of fusions and genomic breakpoints supporting the hypothesis that the observed fusions could be mainly due to balanced events or non-genomic

events rather than unbalanced events or copy number changes as it was described for adult tumors [1,23]. The same studies described a negative correlation between fusions and mutations, which was not observed in our cohort. All these observations could be explained in part by the particular status of our cohort, enrolling pre-treated patients with resistant tumors having acquired a succession of genomic events to become more aggressive and might harbor more chaotic genomes compared to diagnosis. In addition at such tumor stages, the clonal diversity and intratumoral heterogeneity is higher which can distort the analysis [53].

In this cohort, 57% of all detected transcripts resulted from intra-chromosomal events, with an almost identical ratio in the high confidence group whereas 75% of fusions newly discovered by deep sequencing techniques were found to be to be intra-chromosomal, often resulting from subtle inversions or deletions [1]. Possible explanation could be: i) the exclusion of leukemias in our cohort, ii) the rarity of studies searching for fusion oncogenes especially in pediatric solid tumors, iii) the unique biology of pediatric tumors. However, this result should be confirmed at the genomic level to exclude trans-splicing mechanisms.

Both oncogenes and tumor suppressor genes (TSGs) were frequently found in our analysis. Usually, fusion oncogenes are of great interest because of their potential targetability. In our study, more than one third of events for the “high confidence group” were predicted to be in-frame, potentially giving rise to proteins with newly acquired or enhanced oncogenic functions. Most of these fusions present a low oncoscore, nevertheless the function of newly created oncoproteins should be investigated by functional studies. Remarkably, a considerable proportion of fusion transcripts containing tumor suppressor genes was detected. This reflects the importance of tumor suppressor genes in pediatric cancers. In fact, it is estimated that at least 10% of pediatric cancers are caused by hereditary genetic events mostly tumor suppressor gene inactivation [54,55]. Furthermore, the contribution of fusions leading to tumor suppressor genes inactivation might be generally underestimated.

Fusion oncogenes containing tyrosine kinase occurred at low incidence and predominantly found in patients with brain tumors, involving mostly non-receptor tyrosine kinases. Whereas for certain molecular subgroups of high grade gliomas genetic aberrations affecting receptor tyrosine kinases such as PDGFRA, MET, EGFR, FGFR and NTRK have been well defined, this has not been found in the extensive molecular characterization of medulloblastomas [56]. For ependymomas, EGFR overexpression is common but only seldom related to copy number changes. EGFR-containing gene fusions have been reported in a subset of ependymomas [57].

Given the therapeutic potential by a tyrosine kinase-containing fusion we selected these 5 fusion oncogenes for validation and sequencing. The only in-frame transcript that we confirmed by sequencing was CCM2L-HCK transcript in patient 30 with medulloblastoma who lacked aberrations in TP53 and genes of WNT and SHH pathways. The complete coding sequence of HCK

is preserved in the chimeric transcript. HCK belongs to the SRC family of non-receptor protein tyrosine kinases and increased HCK activity is capable to increase proliferation and cell survival as well as associate with receptor tyrosine kinases such as EGFR, FGFR and PDGFR to amplify their oncogenic potential [58]. HCK itself is targetable by small molecule kinase inhibitors such as dasatinib, bosutinib and saracatinib.

The sequencing of the other 4 fusions containing TK proved that the MANBAL-SRC fusion (patient 48) was out-of-frame. Of note, this patient with glioblastoma had several driving pathogenic mutations affecting PI3KCA, NF1 and FGFR1 next to amplifications of MDM2 and CDK4 (already described by Harttrampf et al., 2017 [14]). The FES-MAN2A2 (patient 26) and the TYRO3-KLHL18 transcripts (patient 18) led to a disruption of TK domains therefore of no therapeutic interest. The JAK2-PTPRD transcript identified in patient 21 with ependymoma truncates of JAK2 in the regulatory pseudo-kinase domain with subsequent loss of the C-terminal tyrosine kinase domain. However, this leads to a shift of the reading frame of PTPRD, a gene whose role as tumor suppressor has been documented in many malignancies, including neuroblastoma and glioblastoma [45,59]. In these pathologies, loss of PTPRD activates AURKA and STAT3, respectively. Importantly, both AURKA and STAT3 inhibitors are currently investigated in clinical trials, rendering downstream effects of tumor suppressor inactivation as an attractive therapeutic target. Ortiz et al. further provided in-vivo proof that loss of PTPRD cooperates with CDKN2A/B deletion, which was also found in our patient, and promotes tumorigenesis even if PTPRD is heterozygously inactivated [59].

Next we extracted fusions from the “high-confidence group” that might be also therapeutically relevant. One third of analyzed patients displayed a fusion transcript with potential druggability or involved in drug resistance to certain agents, three of which presented more than one of those. To which extent these fusions might have been present at diagnosis or, have evolved over time under current treatment regimens was not investigated but they might be important contributors to disease progression. However, data correlated to putative effects exerted by the fusion are often preclinical and result from the analysis of other, often “adult tumors” as for example 2 transcripts that we confirmed in 2 different patients (MYO16-IRS2 in patient 32 and MYB-AHI1 in patient 13), the latter of which has already been described in acute lymphoblastic leukemia and breast cancer [23,60]. Strong amplification of these loci has been linked to in vitro sensitivity to IGF1R inhibition [34]. These fusions were associated with a high-level focal amplification further supporting a fusion event [1,23] and potentially sensitizing the tumor towards IGF1R inhibition that was not be applied to the patients in the absence of an available drug.

Further interesting regarding potential targetability was transcript CDKN2A-ITCH (Table S5) in patient 41 with a malignant peripheral nerve sheath tumor. Next to its causal relation to neurofibromatosis type 1, recurrent genetic aberrations affecting tumor suppressor CDKN2A

have been well documented in this malignancy [61]. This transcript, was validated and sequenced and corresponded to p14ARF which acts as a stabilizer of TP53 by sequestering MDM2. The breakpoint leads to the inclusion of exon 1 in the final putative protein which contains the MDM2-binding domain. It has been experimentally shown that this proportion of the gene alone is sufficient to sustain its full TP53-stabilizing capacity, questioning the driving potential of this particular transcript [62–64]. We also validated three fusions with high oncoscore (TP53-ZNF565, UBE4B-CTNNBIP1 and NF2-KIAA0368) in one patient with osteosarcoma because their involved genes were linked to the biology of the patient's cancer. To which extent events like these could be causative, occur due to antineoplastic treatment, progression, or are random bystanders of multiple copy number aberrations, still needs to be defined.

In conclusion, we established a robust pipeline called ChimComp to detect fusion transcripts in cancer patients with 90.5% reliability and close to 100% if fusions with a very low oncoscore were omitted. Hence, fusions which are found by three tools and are primarily provided with a high oncoscore do not, in our opinion, necessarily require experimental validation.

The knowledge of fusion genes' function and their involvement in different signaling pathways can provide new therapeutic options for patients and therefore add significantly to personalized treatments. Also, fusion transcripts occurring in a single patient, so-called private events, should not be primarily discarded.

Material and methods

Patients

The cohort comprised of 48 pediatric patients with 21 different relapsed or refractory solid neoplasms, which had undergone RNA sequencing as part of the “Molecular Screening for Cancer Treatment Optimization” genomic profiling program (MOSCATO-01, NCT01566019) run at Gustave Roussy [14,65].

Whole-Exome and RNA sequencing

Extraction of tumor DNA, RNA and germline DNA, as well as library preparation, capture and Whole-Exome sequencing (WES), RNA-sequencing (RNA-seq, detailed in supplementary information **SI 1**) and CGH array were performed as part of the MOSCATO-01 clinical trial as previously published [14]. In the same study, the mutational burden per patient was also calculated from WES data. Bioinformatics analysis was done by IntegraGen and Gustave Roussy Bioinformatics Core.

The total number of genomic breakpoints per patient was calculated from the results of CGH array which were converted into “.cbs” files containing the genome position of the single segments, the probes and Log2 ratio value.

Fusions detection by ChimComp pipeline

To optimize the detection of potential fusion transcripts, a new analysis pipeline, called “ChimComp” was developed by Gustave Roussy Bioinformatics Core (**Figure 1 details the single steps**). ChimComp relies on 3 tools: TopHat v2.0.14 [17] followed by TopHat post, Defuse v0.6.2 [15] and Fusion Catcher [16]. The strength of ChimComp is to reconcile fusions predicted by different algorithms from the same data in order to classify and annotate these predictions. The non-filtered output from Defuse (file “results.classify.tsv”) was used. The 3 tools give a listing of fusion genes obtained from the raw data of RNA-Seq. A fusion is listed if: i) the anchor length (number of nucleotides overlapping each side of the breakpoint in spanning reads) is ≥ 10 nucleotides, ii) when a fusion is detected by 1 tool, the minimum number of supported read pairs is 40; if found by >1 tool no read threshold, iii) at least one of the genes has a non-intergenic breakpoint. Then, for each fusion detected, an oncogenic potential score was computed by using the Oncofuse algorithm [12]. This results in a score between 0 (non-oncogenic) and 1 (highly oncogenic). In our study, it was calculated for both gene orientations and the highest oncoscore value was considered. The detected fusions were annotated with different fusion genes databases comprising: Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer (<http://cgap.nci.nih.gov/Chromosomes/Mitelman>), TICdb [66], COSMIC

(<http://www.sanger.ac.uk/cosmic>), ChimerDB [67] and ChiTaRS [68]. All the fusions identified among the cohort were tested for the presence of tyrosine kinases (TK), oncogenes or tumor suppressor genes (TSGs) by comparing our listing of genes to a set of 90 known tyrosine kinases genes, 803 known oncogenes and 1 017 known human protein-coding tumor suppressor genes [18–20] (**Figure 1**).

Selection and validation of detected fusions by ChimComp

To validate our approach, 42 fusions found by ChimComp underwent reverse transcribed real time quantitative polymerase chain reaction (RT-qPCR) assay. We primarily focused on the fusions that were found by three tools (referred to as “high confidence group”), then divided the list into three categories, based on oncoscore results: oncoscore < 0.35; 0.35-0.7 and > 0.7. Of those, 25% of fusions were selected for experimental validation by RT-qPCR in each category, resulting in 29 fusions in total. As positive control, some of the well-known fusion oncogenes detected by ChimComp were also included (n=3). Moreover, 8 fusions detected by less than 3 tools harboring a high oncoscore, except for one control healthy fusion (*LPP-TPRG1*), were tested by RT-qPCR. Finally, 5 potentially targetable fusions harboring tyrosine-kinases coding genes were also validated. It should be noticed that the fusions were selected according to the relevance of genes involved in the fusion and the availability of the biological material.

Reconstitution of the fusion sequence and primers design

The breakpoint position in mRNA sequences of both genes was predicted *in silico* (see supplementary information document **SI 2**). The sequence around the breakpoint was used to design the primers necessary for RT-qPCR validation. Primers were selected with a size of about 20 nucleotides, a melting temperature of 60°C and an amplicon size of 100-300 base pairs, using Oligo Explorer 1.1.0 program (Kuopio University, Kuopio, Finland).

Biological validation by reverse transcription (RT) real time PCR (qPCR)

The same RNA used for RNA-Seq was used to evaluate the presence of the different fusions detected. To verify the respective size of each amplicon, 10µl of the amplified products were mixed with 2µl of gel loading dye (New England Biolabs, Evry, France) and loaded into 2% agarose gel mixed with DNA/RNA stain Midori Green Advance (Nippon Genetics Europe, Dueren, Germany). Then the gel was visualized under UV using Ingenius bioimaging system (Syngene, Cambridge, England). PCR products were sequenced after DNA extraction with NucleoSpin® Gel and PCR clean-up extraction kit (Macherey-Nagel, Dueren, Germany) by Sanger technique (Eurofins, Cochin Institute, Paris, France).

Statistical analysis

The numbers of fusions transcripts found in patients are presented as mean \pm interquartile range. To compare the number of fusions with mutational load or genomic breakpoints number per patient Spearman correlation coefficient was determined. To compare the average number of fusions per patient between the different groups of tumor types, Kruskal-Wallis followed by Dunns tests were used. To compare the percentage of total fusion vs fusions found by 3 tools in the 4 categories of oncoscore the z-score test was used. A z-score < 1.96 was considered as significant. All tests were performed by using GraphPad Prism 4 software, p-values and adjusted p-values < 0.05 were considered as a statistically significant level.

Acknowledgments

We are grateful to all patients and their parents for participating in the trial. We acknowledge Catherine Richon (technical assistance), Damien Drubay (biostatistics platform of Gustave Roussy Institute) as well as Marc Gabriel and Bastien Job (bioinformatics core of Gustave Roussy). This study was supported by Fondation Gustave Roussy, Fondation Philantropia (ACH), Université Paris-Sud (CD) and CNRS.

Conflict of interest

There are no competing financial interests in relation to the following work described.

References

- [1] Mertens F, Johansson B, Fioretos T, Mitelman F. The emerging complexity of gene fusions in cancer. *Nat Rev Cancer* 2015;15:371–381.
- [2] Duesberg PH. Cancer genes generated by rare chromosomal rearrangements rather than activation of oncogenes. *Med Oncol Tumor Pharmacother* 1987;4:163–75.
- [3] Druker BJ, Talpaz M, Resta DJ, Peng B, Buchdunger E, Ford JM, et al. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N Engl J Med* 2001;344:1031–7. doi:10.1056/NEJM200104053441401.
- [4] Nagasubramanian R, Wei J, Gordon P, Rastatter JC, Cox MC, Pappo A. Infantile Fibrosarcoma With NTRK3-ETV6 Fusion Successfully Treated With the Tropomyosin-Related Kinase Inhibitor LOXO-101. *Pediatr Blood Cancer* 2016;63:1468–70. doi:10.1002/pbc.26026.
- [5] Singh D, Chan JM, Zoppoli P, Niola F, Sullivan R, Castano A, et al. Transforming Fusions of FGFR and TACC Genes in Human Glioblastoma. *Science* 2012;337:1231–5. doi:10.1126/science.1220834.
- [6] Dupain C, Harttrampf AC, Urbinati G, Georger B, Massaad-Massade L. Relevance of Fusion Genes in Pediatric Cancers: Toward Precision Medicine. *Mol Ther - Nucleic Acids* 2017;6:315–26. doi:10.1016/j.omtn.2017.01.005.
- [7] Carrara M, Beccuti M, Lazzarato F, Cavallo F, Cordero F, Donatelli S, et al. State-of-the-Art Fusion-Finder Algorithms Sensitivity and Specificity. *BioMed Res Int* 2013;2013. doi:10.1155/2013/340620.
- [8] Wang Q, Xia J, Jia P, Pao W, Zhao Z. Application of next generation sequencing to human gene fusion detection: computational tools, features and perspectives. *Brief Bioinform* 2013;14:506–19. doi:10.1093/bib/bbs044.
- [9] Kumar S, Razzaq SK, Vo AD, Gautam M, Li H. Identifying fusion transcripts using next generation sequencing. *Wiley Interdiscip Rev RNA* 2016;7:811–23. doi:10.1002/wrna.1382.
- [10] Liu S, Tsai W-H, Ding Y, Chen R, Fang Z, Huo Z, et al. Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data. *Nucleic Acids Res* 2016;44:e47. doi:10.1093/nar/gkv1234.
- [11] Fernandez-Cuesta L, Sun R, Menon R, George J, Lorenz S, Meza-Zepeda LA, et al. Identification of novel fusion genes in lung cancer using breakpoint assembly of transcriptome sequencing data. *Genome Biol* 2015;16:7. doi:10.1186/s13059-014-0558-0.
- [12] Shugay M, Ortiz de Mendíbil I, Vizmanos JL, Novo FJ. Oncofuse: a computational framework for the prediction of the oncogenic potential of gene fusions. *Bioinforma Oxf Engl* 2013;29:2539–46. doi:10.1093/bioinformatics/btt445.
- [13] Marshall GM, Carter DR, Cheung BB, Liu T, Mateos MK, Meyerowitz JG, et al. The prenatal

- origins of cancer. *Nat Rev Cancer* 2014;14:277–89. doi:10.1038/nrc3679.
- [14] Harttrampf AC, Lacroix L, Deloger M, Deschamps F, Puget S, Auger N, et al. MOlecular Screening for CAncer Treatment Optimization (MOSCATO-01) in pediatric patients: A single institutional prospective molecular stratification trial. *Clin Cancer Res Off J Am Assoc Cancer Res* 2017. doi:10.1158/1078-0432.CCR-17-0381.
- [15] McPherson A, Hormozdiari F, Zayed A, Giuliany R, Ha G, Sun MGF, et al. deFuse: An Algorithm for Gene Fusion Discovery in Tumor RNA-Seq Data. *PLoS Comput Biol* 2011;7. doi:10.1371/journal.pcbi.1001138.
- [16] Nicorici D, Satalan M, Edgren H, Kangaspeska S, Murumagi A, Kallioniemi O, et al. FusionCatcher-a tool for finding somatic fusion genes in paired-end RNA-sequencing data. *BioRxiv* 2014:011650.
- [17] Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 2013;14:R36. doi:10.1186/gb-2013-14-4-r36.
- [18] Robinson DR, Wu Y-M, Lin S-F. The protein tyrosine kinase family of the human genome. *Oncogene* 2000;19:5548–57. doi:10.1038/sj.onc.1203957.
- [19] Liu Y, Sun J, Zhao M. ONGene: A literature-based database for human oncogenes. *J Genet Genomics* 2017;44:119–21. doi:10.1016/j.jgg.2016.12.004.
- [20] Zhao M, Sun J, Zhao Z. TSGene: a web resource for tumor suppressor genes. *Nucleic Acids Res* 2013;41:D970-976. doi:10.1093/nar/gks937.
- [21] Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell* 2011;144:646–74. doi:10.1016/j.cell.2011.02.013.
- [22] Kalyana-Sundaram S, Shankar S, Deroo S, Iyer MK, Palanisamy N, Chinnaiyan AM, et al. Gene fusions associated with recurrent amplicons represent a class of passenger aberrations in breast cancer. *Neoplasia N Y N* 2012;14:702–8.
- [23] Yoshihara K, Wang Q, Torres-Garcia W, Zheng S, Vegesna R, Kim H, et al. The landscape and therapeutic relevance of cancer-associated transcript fusions. *Oncogene* 2015;34:4845–54. doi:10.1038/onc.2014.406.
- [24] Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer Genome Landscapes. *Science* 2013;339:1546–58. doi:10.1126/science.1235122.
- [25] Chase A, Ernst T, Fiebig A, Collins A, Gr F, Reiter A, et al. A polymorphic TFG-GPR128 fusion gene associated with a copy number variant in healthy individuals. 2009.
- [26] Cachia D, Wani K, Penas-Prado M, Olar A, McCutcheon IE, Benjamin RS, et al. C11orf95–RELA fusion present in a primary supratentorial ependymoma and recurrent sarcoma. *Brain Tumor Pathol* 2015;32:105–11. doi:10.1007/s10014-014-0205-1.
- [27] Lin PP, Pandey MK, Jin F, Xiong S, Deavers M, Parant JM, et al. EWS-FLI1 induces

developmental abnormalities and accelerates sarcoma formation in a transgenic mouse model. *Cancer Res* 2008;68:8968–8975.

[28] Zhu C, Lv Y, Wu L, Guan J, Bai X, Lin J, et al. The landscape of gene fusions in hepatocellular carcinoma. *BioRxiv* 2016:055376. doi:10.1101/055376.

[29] Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. *Nat Rev Cancer* 2007;7:233–45. doi:10.1038/nrc2091.

[30] Wang L, Motoi T, Khanin R, Olshen A, Mertens F, Bridge J, et al. Identification of a novel, recurrent HEY1-NCOA2 fusion in mesenchymal chondrosarcoma based on a genome-wide screen of exon-level expression data. *Genes Chromosomes Cancer* 2012;51:127–39. doi:10.1002/gcc.20937.

[31] Bell RJA, Rube HT, Xavier-Magalhães A, Costa BM, Mancini A, Song JS, et al. Understanding TERT Promoter Mutations: A Common Path to Immortality. *Mol Cancer Res MCR* 2016;14:315–23. doi:10.1158/1541-7786.MCR-16-0003.

[32] Pilsworth JA, Cochrane DR, Xia Z, Aubert G, Färkkilä AEM, Horlings HM, et al. TERT promoter mutation in adult granulosa cell tumor of the ovary. *Mod Pathol Off J U S Can Acad Pathol Inc* 2018. doi:10.1038/s41379-018-0007-9.

[33] Tsuda M, Davis IJ, Argani P, Shukla N, McGill GG, Nagai M, et al. TFE3 fusions activate MET signaling by transcriptional up-regulation, defining another class of tumors as candidates for therapeutic MET inhibition. *Cancer Res* 2007;67:919–29. doi:10.1158/0008-5472.CAN-06-2855.

[34] Pavlicek A, Lira ME, Lee NV, Ching KA, Ye J, Cao J, et al. Molecular predictors of sensitivity to the insulin-like growth factor 1 receptor inhibitor Figitumumab (CP-751,871). *Mol Cancer Ther* 2013;12:2929–39. doi:10.1158/1535-7163.MCT-13-0442-T.

[35] Zhang B, Groffen J, Heisterkamp N. Resistance to farnesyltransferase inhibitors in Bcr/Abl-positive lymphoblastic leukemia by increased expression of a novel ABC transporter homolog ATP11a. *Blood* 2005;106:1355–61. doi:10.1182/blood-2004-09-3655.

[36] Marshall AD, Grosveld GC. Alveolar rhabdomyosarcoma – The molecular drivers of PAX3/7-FOXO1-induced tumorigenesis. *Skelet Muscle* 2012;2:25. doi:10.1186/2044-5040-2-25.

[37] Beretta GL, Cassinelli G, Pennati M, Zuco V, Gatti L. Overcoming ABC transporter-mediated multidrug resistance: The dual role of tyrosine kinase inhibitors as multitargeting agents. *Eur J Med Chem* 2017;142:271–89. doi:10.1016/j.ejmech.2017.07.062.

[38] Borad MJ, Champion MD, Egan JB, Liang WS, Fonseca R, Bryce AH, et al. Integrated genomic characterization reveals novel, therapeutically relevant drug targets in FGFR and EGFR pathways in sporadic intrahepatic cholangiocarcinoma. *PLoS Genet* 2014;10:e1004135. doi:10.1371/journal.pgen.1004135.

[39] Zhang Z, Miao L, Xin X, Zhang J, Yang S, Miao M, et al. Underexpressed CNBP2 participates in gastric cancer growth inhibition through activating the MAPK signaling pathway. *Mol Med*

Camb Mass 2014;20:17–28. doi:10.2119/molmed.2013.00102.

[40] Wick W, Weller M, van den Bent M, Sanson M, Weiler M, von Deimling A, et al. MGMT testing--the challenges for biomarker-based glioma treatment. *Nat Rev Neurol* 2014;10:372–85. doi:10.1038/nrneurol.2014.100.

[41] Carraway KL, Weber JL, Unger MJ, Ledesma J, Yu N, Gassmann M, et al. Neuregulin-2, a new ligand of ErbB3/ErbB4-receptor tyrosine kinases. *Nature* 1997;387:512–6. doi:10.1038/387512a0.

[42] Hirai H, Iwasawa Y, Okada M, Arai T, Nishibata T, Kobayashi M, et al. Small-molecule inhibition of Wee1 kinase by MK-1775 selectively sensitizes p53-deficient tumor cells to DNA-damaging agents. *Mol Cancer Ther* 2009;8:2992–3000. doi:10.1158/1535-7163.MCT-09-0463.

[43] Aarts M, Sharpe R, Garcia-Murillas I, Gevensleben H, Hurd MS, Shumway SD, et al. Forced mitotic entry of S-phase cells as a therapeutic strategy induced by inhibition of WEE1. *Cancer Discov* 2012;2:524–39. doi:10.1158/2159-8290.CD-11-0320.

[44] Zhao S, Sedwick D, Wang Z. Genetic alterations of protein tyrosine phosphatases in human cancers. *Oncogene* 2015;34:3885–94. doi:10.1038/onc.2014.326.

[45] Meehan M, Parthasarathi L, Moran N, Jefferies CA, Foley N, Lazzari E, et al. Protein tyrosine phosphatase receptor delta acts as a neuroblastoma tumor suppressor by destabilizing the aurora kinase A oncogene. *Mol Cancer* 2012;11:6. doi:10.1186/1476-4598-11-6.

[46] Fazal-Salom J, Vinci M, Carvalho D, Pemberton H, Pettitt SJ, Lord CJ, et al. Abstract 1932: Mutations in ATRX increase genetic instability and sensitivity to PARP inhibitors in paediatric glioblastoma cells. *Cancer Res* 2017;77:1932–1932. doi:10.1158/1538-7445.AM2017-1932.

[47] Petrilli AM, Fernández-Valle C. Role of Merlin/NF2 inactivation in tumor biology. *Oncogene* 2016;35:537–48. doi:10.1038/onc.2015.125.

[48] Shapiro IM, Kolev VN, Vidal CM, Kadariya Y, Ring JE, Wright Q, et al. Merlin deficiency predicts FAK inhibitor sensitivity: a synthetic lethal relationship. *Sci Transl Med* 2014;6:237ra68. doi:10.1126/scitranslmed.3008639.

[49] Brenner JC, Feng FY, Han S, Patel S, Goyal SV, Bou-Maroun LM, et al. PARP-1 inhibition as a targeted strategy to treat Ewing's sarcoma. *Cancer Res* 2012;72:1608–13. doi:10.1158/0008-5472.CAN-11-3648.

[50] Loganathan SN, Tang N, Fleming JT, Ma Y, Guo Y, Borinstein SC, et al. BET bromodomain inhibitors suppress EWS-FLI1-dependent transcription and the IGF1 autocrine mechanism in Ewing sarcoma. *Oncotarget* 2016;7:43504–17. doi:10.18632/oncotarget.9762.

[51] Huang Z, Jones DTW, Wu Y, Lichter P, Zapatka M. confFuse: High-Confidence Fusion Gene Detection across Tumor Entities. *Front Genet* 2017;8:137. doi:10.3389/fgene.2017.00137.

[52] Gao Q, Liang W-W, Foltz SM, Mutharasu G, Jayasinghe RG, Cao S, et al. Driver Fusions and Their Implications in the Development and Treatment of Human Cancers. *Cell Rep* 2018;23:227-

238.e3. doi:10.1016/j.celrep.2018.03.050.

[53] Burrell RA, Swanton C. Tumour heterogeneity and the evolution of polyclonal drug resistance. *Mol Oncol* 2014;8:1095–111. doi:10.1016/j.molonc.2014.06.005.

[54] Zhang J, Walsh MF, Wu G, Edmonson MN, Gruber TA, Easton J, et al. Germline Mutations in Predisposition Genes in Pediatric Cancer. *N Engl J Med* 2015;373:2336–46. doi:10.1056/NEJMoa1508054.

[55] Multi-Dimensional ClinOmics for Precision Therapy of Children and Adolescent Young Adults with Relapsed and Refractory Cancer: A report from the Center for Cancer Research | Clinical Cancer Research n.d. <http://clincancerres.aacrjournals.org/content/early/2016/03/18/1078-0432.CCR-15-2717> (accessed February 12, 2018).

[56] Pajtler KW, Witt H, Sill M, Jones DTW, Hovestadt V, Kratochwil F, et al. Molecular Classification of Ependymal Tumors across All CNS Compartments, Histopathological Grades, and Age Groups. *Cancer Cell* 2015;27:728–43. doi:10.1016/j.ccell.2015.04.002.

[57] Servidei T, Meco D, Muto V, Bruselles A, Ciolfi A, Trivieri N, et al. Novel SEC61G-EGFR Fusion Gene in Pediatric Ependymomas Discovered by Clonal Expansion of Stem Cells in Absence of Exogenous Mitogens. *Cancer Res* 2017;77:5860–72. doi:10.1158/0008-5472.CAN-17-0790.

[58] Poh AR, O'Donoghue RJJ, Ernst M. Hematopoietic cell kinase (HCK) as a therapeutic target in immune and cancer cells. *Oncotarget* 2015;6:15752–71. doi:10.18632/oncotarget.4199.

[59] Ortiz B, Fabius AWM, Wu WH, Pedraza A, Brennan CW, Schultz N, et al. Loss of the tyrosine phosphatase PTPRD leads to aberrant STAT3 activation and promotes gliomagenesis. *Proc Natl Acad Sci U S A* 2014;111:8149–54. doi:10.1073/pnas.1401952111.

[60] Atak ZK, Gianfelici V, Hulselmans G, De Keersmaecker K, Devasia AG, Geerdens E, et al. Comprehensive analysis of transcriptome variation uncovers known and novel driver events in T-cell acute lymphoblastic leukemia. *PLoS Genet* 2013;9:e1003997. doi:10.1371/journal.pgen.1003997.

[61] Brohl AS, Kahen E, Yoder SJ, Teer JK, Reed DR. The genomic landscape of malignant peripheral nerve sheath tumors: diverse drivers of Ras pathway activation. *Sci Rep* 2017;7. doi:10.1038/s41598-017-15183-1.

[62] Midgley CA, Desterro JM, Saville MK, Howard S, Sparks A, Hay RT, et al. An N-terminal p14ARF peptide blocks Mdm2-dependent ubiquitination in vitro and can activate p53 in vivo. *Oncogene* 2000;19:2312–23. doi:10.1038/sj.onc.1203593.

[63] The Somatic Genomic Landscape of Glioblastoma: Cell n.d. [http://www.cell.com/abstract/S0092-8674\(13\)01208-7](http://www.cell.com/abstract/S0092-8674(13)01208-7) (accessed February 12, 2018).

[64] Sturm D, Bender S, Jones DTW, Lichter P, Grill J, Becher O, et al. Paediatric and adult glioblastoma: multiform (epi)genomic culprits emerge. *Nat Rev Cancer* 2014;14:92–107.

doi:10.1038/nrc3655.

[65] Massard C, Michiels S, Ferté C, Le Deley M-C, Lacroix L, Hollebecque A, et al. High-Throughput Genomics and Clinical Outcome in Hard-to-Treat Advanced Cancers: Results of the MOSCATO 01 Trial. *Cancer Discov* 2017;7:586–95. doi:10.1158/2159-8290.CD-16-1396.

[66] Novo FJ, de Mendíbil IO, Vizmanos JL. TICdb: a collection of gene-mapped translocation breakpoints in cancer. *BMC Genomics* 2007;8:33. doi:10.1186/1471-2164-8-33.

[67] Kim N, Kim P, Nam S, Shin S, Lee S. ChimerDB--a knowledgebase for fusion sequences. *Nucleic Acids Res* 2006;34:D21-24. doi:10.1093/nar/gkj019.

[68] Gorohovski A, Tagore S, Palande V, Malka A, Raviv-Shay D, Frenkel-Morgenstern M. ChiTaRS-3.1-the enhanced chimeric transcripts and RNA-seq database matched with protein-protein interactions. *Nucleic Acids Res* 2017;45:D790–5. doi:10.1093/nar/gkw1127.

Supplementary information

Supplementary information 1: RNA-Seq

RNA-Seq base calling was performed using the Real-Time Analysis software sequence pipeline (RTA2) with default parameters (1). Single Nucleotides Variants (SNVs) were detected using GATK 3.4 following GATK online published best practices for RNA variant calling (STAR two-pass alignment (2), Picardtools (3) Mark Duplicates and Sort, GATK Split'N'Trim, Indel Realignment, Base Recalibration, then GATK HaplotypeCaller (4)).

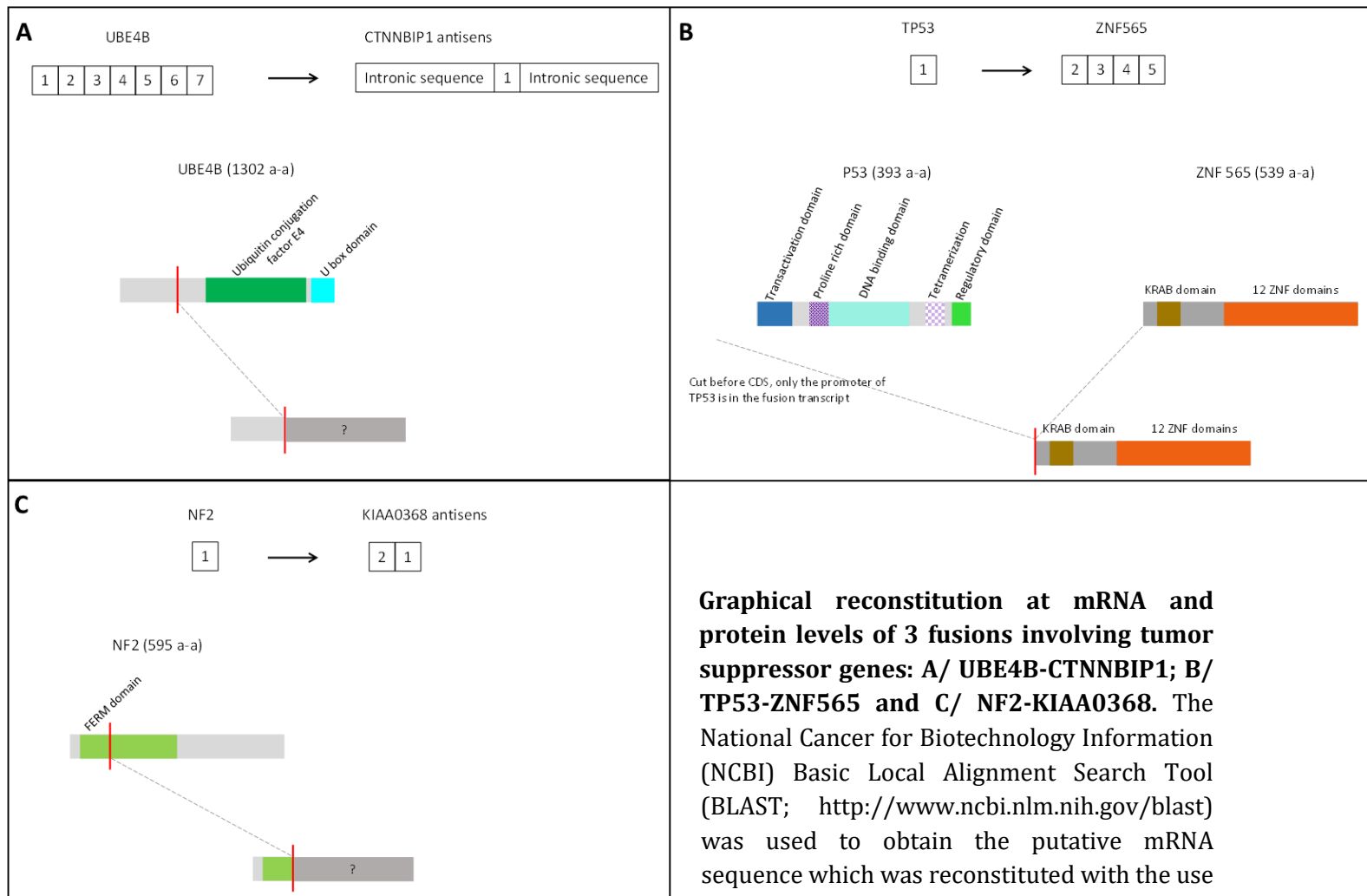
1. Real-Time Analysis (RTA) Support [Internet]. [cited 2018 Mar 12]. Available from: https://support.illumina.com/sequencing/sequencing_software/real-time_analysis_rta.html
2. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinforma Oxf Engl*. 2013 Jan 1;29(1):15–21.
3. Picard Tools - By Broad Institute [Internet]. [cited 2018 Mar 12]. Available from: <https://broadinstitute.github.io/picard/>
4. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010 Sep;20(9):1297–303.

Supplementary information 2: Breakpoint sequences reconstitutions

The global steps of processing in order to visualize the fusions at nucleotidic level and extract the breakpoint sequence for each fusion detected by ChimComp in one given patient are the followings:

- a) The breakpoint for which as many tools as possible are agreed is selected (consensus breakpoint), then the longest sequence of the consensus breakpoint is selected. If there is no breakpoint consensus, the longest sequence of fusion across all tools is selected (with its associated breakpoint).
- b) For each gene involved in the fusion, only transcripts overlapping the consensus breakpoint are kept. If the involved gene is an intergenic region, a region at -2000/+2000 bp from the breakpoint is built.
- c) The sequence corresponding to the union of the involved transcripts for each gene is extracted (according to the strand).
- d) The sequence of fusion is aligned to the involved transcripts of each gene with Blast in order to visualize any discordance on the orientation.
- e) Uniquely mapped reads for each gene are selected and only reads for which the identity is present in both genes are kept.
- f) The "read balance" is computed for each gene, in order to determine which part of the gene belongs to the fusion (left or right side of the breakpoint). The features (exons and introns) of each gene are trimmed according to the "read balance". Reads corresponding to these features outside of the range are removed.
- g) By rescaling the positions, the features of each gene and their associated reads are joined by the breakpoint.

Supplementary information 3: Fusions detected and validated in patient 17 with osteosarcoma



Graphical reconstitution at mRNA and protein levels of 3 fusions involving tumor suppressor genes: A/ UBE4B-CTNNBIP1; B/ TP53-ZNF565 and C/ NF2-KIAA0368. The National Cancer for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST; <http://www.ncbi.nlm.nih.gov/blast>) was used to obtain the putative mRNA sequence which was reconstituted with the use of Nucleotide (NBCI). When the coding

sequence was obtained and predicted, Uniprot Knowledge Base (<http://www.uniprot.org/>) was used to reconstitute the predicted fusion protein domains.

Framed numbers represent the exons corresponding to the gene transcript written above each group of numbers. The arrows represent the junction point between the two transcripts to form the fusion. Colored boxes represent the two entire proteins involved in the fusion, with the different functional domains. The red bar shows the junction point and dotted lines illustrates the protein rearrangement after the fusion transcript formation. Intronic sequence = a transcript with a retained genomic intronic sequence of the corresponding gene.

This patient 17 with osteosarcoma presented with 92 putative fusions in accordance with a characteristically highly rearranged genome 3 of which were selected for biological validation by RT-qPCR based on their high biological interest in osteosarcoma. Concerning UBE4B-CTNNBIP1 fusion (A) (Oncoscore 0.9, found by DeFuse and TopHat) its transcript can potentially lead to altered expression of the CTNNBIP1 protein which inhibits the key effector of the WNT pathway, β -catenin. β -catenin is responsible for activation of a complex of transcription factors regulating WNT-responsive genes involved in developmental biology, tumor progression and importantly stem-cell renewal in osteosarcoma. The importance of the WNT pathway in osteoblast differentiation but also the misbalance between its activators and inhibitors in osteosarcoma have been demonstrated. Agents interfering with the WNT/ β -catenin pathway such as PORCN, i.e. LGK974, and WNT inhibitors such as CGX1321 which are in early phase clinical trials might be a strategy to pursue for these patients in the future. This fusion also leads to the disruption of UBE4B, a tumor suppressor identified in neuroblastoma that could be linked to development process. The TP53-ZNF565 fusion (B) (oncoscore of 0.96, found by 3 tools) could disrupt the coding sequence of TP53. TP53 inactivation is a hallmark finding in osteosarcoma, surprisingly not only by mutations but by structural variations in over 50% of cases investigated, and peculiarly, similar to our case, in the first intron. This might be a reflection of the gross genomic instability that characterizes osteosarcoma. Finally, NF2-KIAA0368 fusion (C) (Oncoscore 0.9, found only by TopHat) could lead to a frameshift of NF2, whose gene product has pleiotropic effects on several pathways, an important one of which is the negative regulation of the YAP/HIPPO pathway. YAP is overexpressed in a variety of tumors including osteosarcoma and NF2 inactivation has been shown to induce osteosarcoma in NF2 heterozygous mice and more specifically, NF2 is directly repressed by SOX2, a stem cell transcription factor which maintains cancer stem cells in osteosarcoma. Furthermore, NF2 is an inhibitor of mTORC1 and its deficiency has been also linked to sensitivity towards FAK inhibition (1–11).

1. Chen X, Bahrami A, Pappo A, Easton J, Dalton J, Hedlund E, et al. Recurrent Somatic Structural Variations Contribute to Tumorigenesis in Pediatric Osteosarcoma. *Cell Rep.* 2014 Apr 10;7(1):104–12.
2. Chen KS, Kwon WS, Kim J, Heo SJ, Kim HS, Kim HK, et al. A novel TP53-KPNA3 translocation defines a de novo treatment-resistant clone in osteosarcoma. *Cold Spring Harb Mol Case Stud.* 2016 Sep;2(5):a000992.
3. Martins-Neves SR, Corver WE, Paiva-Oliveira DI, van den Akker BEWM, Briaire-de-Bruijn IH, Bovée JVMG, et al. Osteosarcoma Stem Cells Have Active Wnt/ β -catenin and Overexpress SOX2 and KLF4. *J Cell Physiol.* 2016 Apr 1;231(4):876–86.

4. Martins-Neves SR, Paiva-Oliveira DI, Fontes-Ribeiro C, Bovée JVMG, Cleton-Jansen A-M, Gomes CMF. IWR-1, a tankyrase inhibitor, attenuates Wnt/ β -catenin signaling in cancer stem-like cells and inhibits in vivo the growth of a subcutaneous human osteosarcoma xenograft. *Cancer Lett.* 2018 Feb 1;414:1–15.
5. Rodda SJ, McMahon AP. Distinct roles for Hedgehog and canonical Wnt signaling in specification, differentiation and maintenance of osteoblast progenitors. *Dev Camb Engl.* 2006 Aug;133(16):3231–44.
6. Lin CH, Guo Y, Ghaffar S, McQueen P, Pourmorady J, Christ A, et al. Dkk-3, a secreted wnt antagonist, suppresses tumorigenic potential and pulmonary metastasis in osteosarcoma. *Sarcoma.* 2013;2013:147541.
7. Basu-Roy U, Bayin NS, Rattanakorn K, Han E, Placantonakis DG, Mansukhani A, et al. Sox2 antagonizes the Hippo pathway to maintain stemness in cancer cells. *Nat Commun.* 2015 Apr 2;6:6411.
8. McClatchey AI, Saotome I, Mercer K, Crowley D, Gusella JF, Bronson RT, et al. Mice heterozygous for a mutation at the Nf2 tumor suppressor locus develop a range of highly metastatic tumors. *Genes Dev.* 1998 Apr 15;12(8):1121–33.
9. Pai SG, Carneiro BA, Mota JM, Costa R, Leite CA, Barroso-Sousa R, et al. Wnt/beta-catenin pathway: modulating anticancer immune response. *J Hematol Oncol* *J Hematol Oncol.* 2017 May 5;10:101.
10. Li W, Cooper J, Karajannis MA, Giancotti FG. Merlin: a tumour suppressor with functions at the cell cortex and in the nucleus. *EMBO Rep.* 2012 Mar;13(3):204–15.
11. Shapiro IM, Kolev VN, Vidal CM, Kadariya Y, Ring JE, Wright Q, et al. Merlin deficiency predicts FAK inhibitor sensitivity: a synthetic lethal relationship. *Sci Transl Med.* 2014 May 21;6(237):237ra68.

Tables

Patient Number	Diagnosis ^a	Tumor type (S=Sarcoma, B=Brain tumors, O=Others)	Total fusions detected	Number of oncogenes detected among fusions	Number of TSGs detected among fusions	N° of fusions detected by 3 tools	N° of fusions detected by Defuse	N° of fusions detected by Fusion Catcher	N° of fusions detected by TopHat	N° of inter-chromosomal fusions	N° of intra-chromosomal fusions	Somatic non-synonymous mutations	Somatic non-synonymous mut rate/Mb	Genomic breakpoints
1	Nephroblastoma	O	53	0	3	2	46	9	30	33	20	600	13.04	n.a ^b
2	Sertoli Leydig granulosa tumor	O	41	2	2	2	34	6	27	27	14	91	1.98	77
3	Astroblastoma	B	40	2	4	1	35	9	19	21	19	45	0.98	14
4	ASPS	S	53	3	3	1	47	4	22	33	20	1030	22.39	22
5	Nephroblastoma	O	64	3	4	1	57	0	32	34	30	81	1.76	56
6	PNET	B	49	2	2	1	40	3	31	36	13	794	17.26	16
7	Hepatoblastoma	O	33	2	2	0	29	8	13	18	15	49	1.07	41
8	eRMS	S	49	1	4	3	44	12	27	27	22	66	1.43	48
9	FL-HCC	O	59	2	6	3	52	13	34	33	26	38	0.83	55
10	eRMS	S	23	1	1	0	19	4	12	16	7	50	1.09	31
11	Ewing sarcoma	S	46	4	2	3	40	6	22	18	18	66	1.43	18
12	Ewing-like sarcoma	S	31	1	2	1	28	6	18	21	10	78	1.70	23
13	eRMS	S	48	5	2	3	45	12	26	24	24	58	1.26	38
14	eRMS	S	52	2	3	3	47	13	25	28	24	65	1.41	29
15	Epithelioid Sarcoma	S	51	2	5	3	47	12	28	26	25	55	1.20	94
16	Medulloblastoma	B	153	6	19	24	144	33	115	78	75	142	3.09	398
17	Osteosarcoma	S	92	2	6	7	83	18	55	45	47	79	1.72	99
18	Ependymoma	B	65	5	6	3	62	2	35	32	33	39	0.85	22
19	SETTLLE	O	23	0	1	0	19	3	12	13	10	505	10.98	15
20	eRMS Li-Fraumeni	S	58	0	10	7	45	15	46	33	25	68	1.48	173
21	Ependymoma	B	93	3	6	14	84	37	58	40	53	80	1.74	213
22	HGG	B	51	5	7	1	44	10	23	25	26	470	10.22	19
23	DIPG	B	69	6	1	4	56	10	47	43	26	49	1.07	29
24	Hepatoblastoma	O	44	0	5	1	37	9	20	23	21	28	0.61	42
25	LGG	B	67	3	2	1	59	15	36	33	34	31	0.67	67
26	Neuroblastoma	O	90	5	8	5	85	23	41	47	43	352	7.65	218
27	Ependymoma	B	47	2	4	2	40	12	22	19	28	69	1.50	114
28	eRMS	S	62	3	2	2	56	16	25	29	33	n.w ^c	n.w ^c	135
29	eRMS	S	37	3	4	2	36	7	25	30	7	107	2.33	119
30	Medulloblastoma	B	58	2	4	1	52	17	26	24	34	n.w ^c	n.w ^c	121
31	HGG (Glioblastoma)	B	74	3	7	2	68	21	32	31	43	130	2.83	253
32	Medulloblastoma	B	48	2	5	3	44	17	23	22	26	115	2.50	163
33	Chondrosarcoma	S	34	2	1	0	29	6	16	18	16	184	4.00	129
34	Ependymoma	B	46	3	3	3	39	14	21	21	25	124	2.70	109
35	eRMS	S	33	4	3	0	29	8	16	12	21	128	2.78	164
36	Neuroblastoma	O	44	9	1	0	23	6	35	12	32	137	2.98	116
37	Medulloblastoma	B	50	3	4	1	48	15	20	19	31	182	3.96	169
38	eRMS	S	83	3	9	5	76	16	39	38	45	72	1.57	143
39	DSRCT	S	15	2	2	1	12	5	7	7	8	105	2.28	94
40	Osteosarcoma	S	82	3	7	12	70	25	59	40	42	126	2.74	227
41	MNST	O	25	0	3	2	20	5	14	18	7	128	2.78	136
42	Medulloblastoma	B	50	1	1	0	42	13	30	17	33	41	0.89	237
43	Abrikossoff tumor	O	36	2	2	1	31	7	11	15	21	81	1.76	35
44	Neuroblastoma	O	80	0	11	0	71	15	43	40	40	221	4.80	322
45	Medulloblastoma	B	3	0	0	0	2	0	2	2	1	70	1.52	342
46	Ewing's sarcoma	S	41	2	2	1	39	7	16	21	20	45	0.98	73
47	eRMS	S	25	1	1	3	19	5	19	17	8	92	2.00	261
48	Glioblastoma	B	223	15	23	18	199	61	159	102	121	125	2.72	423

Table 1. Fusions and genomic alterations metrics for the 48 pediatric patients studied.

Forty eight pediatric patients were enrolled in the previously described molecular profiling program MOSCATO-01 conducted at Gustave Roussy. Sequencing data and clinical record of these patients were used for our study. For each patient, identification number, associated diagnosis and tumor type is reported. Fusions metrics and type of chromosomal rearrangements were also assessed thanks to ChimComp pipeline. Finally, the associated mutational rate and breakpoints numbers were calculated from WES and CGH array respectively. ^a ASPS = alveolar soft part sarcoma; PNET = primitive neuroectodermal tumor; eRMS = embryonal rhabdomyosarcoma; FL-HCC = fibrolamellar hepatocellular carcinoma; aRMS = alveolar rhabdomyosarcoma; SETTLE = spindle epithelial tumor with thymus-like differentiation; HGG = high-grade glioma; DIPG = diffuse intrinsic pontine glioma; LGG = low-grade glioma; DSRCT = desmoplastic small-round-cell tumor; MNST = malignant nerve sheath tumor; TSGs = tumor suppressor genes.

^b n.a = not able to be calculated

^c n.w = no WES performed

Fusion gene 1	Fusion gene 2	Type of rearrangement ^c	Nb of tools	Nb of patients carrying fusion	Patient Number	Validated by PCR	Oncoscore	Nb of databases reporting the fusion	Read nb (mean)	Forward primer sequence	Reverse primer sequence	Oncogene(s)	Tumor suppressor gene(s)	Functional prediction
High-confidence candidate: found by 3 tools														
CDKN2A	ITCH	Inter	3	1	41	Yes	/	0	23	GGTTTGGTTCACATC	ATTGCTCCCTGGTCTCTAG	/	/	Inhibition of ubiquitination and loss of CDKN2A tumor suppressor gene
NF2	AK131325	Intra	3	1	40	Yes	/	0	74	GAGCTCAGCTCTCAAG	CCACAGATAACTCCAAC	/	NF2	Loss of NF2 tumor suppressor, negative regulator of cell cycle and apoptosis
SAP30BP	CDRT4	Intra	3	1	21	No	/	0	111	ACAGTAGACAGTCGGAAGA	TGAGAGGTATAGGTGACATA	/	/	Inhibition of cell death
SEMA3D	ABCBI	Intra	3	1	26	Yes	/	0	8,0	CCCTCTCCCTGAAAGG	CGGATTGACTGAATGCTG	/	/	Increased MDR (multidrug resistance)
TET1	TTC12	Inter	3	1	21	Yes	/	0	22	CAGAGTTTGGCTACAGATT	GTGTAGGGGTAACAAGTTT	/	/	Aberrant DNA methylation
ABR	DNAH2	Intra	3	1	40	No	0,00	0	94	CGGGAAACAAGGTAGAAA	AGCGGGGAAAATATGTCG	/	/	May have an influence on cell cycle and activation of protein kinases
FLVCR1	TATDN3	Intra	3	1	17	Yes	0,01	0	150	TGCAGCAGCATCTCTCTG	CGGTAGTGACATTTACAGG	/	/	May have an influence on cell metabolism or differentiation
VP53	NXN	Intra	3	1	30	Yes	0,03	0	168	ATCTCTGGGAAACATAGAC	GTTCCTGACACAACTCTC	/	VP53	Activation of Wnt signaling involved in cell growth and differentiation
FBX7	TERT	Intra	3	1	2	Yes	0,11	0	51	ATCACACCCCACTAAAGC	CGACAGGTACACACTCATC	/	/	Loss of cell cycle arrest by ubiquitination of Aurora kinases and overactivation of telomerase
MAML1	BEND2	Intra	3	1	209	Yes	0,11	0	209	GCCTGGCACTGTTTGC	GCCTCCACCACTCTCTG	/	/	Chromatin restructuring
TFG	GPRL28 ^a	Intra	3	1	18	Yes	0,15	6	152	ACTTGGGAGGATATGGG	TGTAGGGTGGTCTGTAG	TFG	/	Found in healthy individuals
KIAA1324L	TMEV243	Intra	3	1	16	Yes	0,19	0	17	CGGTGCTTTGATCTATGC	ACGGTTTTGGAGTAAATGG	/	/	May be involved in drug resistance
HMGCLL1	BAI3	Intra	3	1	21	Yes	0,21	0	88	ACCAGCTCCATCACTTG	GTGGTTTACCCTGATG	/	/	May be involved in differentiation in neurons
DNAI1	RERGL	Inter	3	1	47	Yes	0,26	0	7	SAGACTGGAGTTGTTGAM	CCCGGATCTGTAGATCAG	/	/	May have consequences on protein synthesis
HERC4	TMPRSS15	Inter	3	1	8	Yes	0,46	0	7,0	GTAATCCCTCTGAAAG	AGACCTTGGATACACTC	/	/	Inhibition of ubiquitination, stabilisation of proteins
IP6K1	RYBP	Intra	3	1	13	Yes	0,47	0	21,3	CGCAGCAAGAGGTTTTG	CACGAGGGTACTTGAAG	/	/	Increased kinase activity or apoptosis resistance
DMTF1	ABCBI	Intra	3	1	15	Yes	0,89	0	8,0	GGACCCAAACAGAAAGA	CGAAGACTCGAAACAGG	/	DMTF1	Increased MDR (multidrug resistance)
NRG2	ZNF208	Inter	3	1	16	Yes	0,92	0	25,0	CAACGGCAGAAAGAACT	AGGACACAGCAATFAAG	/	/	Activation of TK receptor
ERRF1	CUBN	Inter	3	1	16	Yes	0,95	0	4	GGCACAAATGCAATAGAG	GGAGGACACAAAGTCAAG	/	ERRF1	Activation of EGF
TP53	ZNF565	Inter	3	1	17	Yes	0,96	0	11,5	TCCATGTGCTCAAGACTG	GTTCCAGGCACTTCAATC	/	TP53	Inactivation of p53 function
EIF2AK1	OGDH	Intra	3	1	16	Yes	0,98	0	20,0	CTCTCCCGTTCTGTTG	GGGAGGGCACTTCTCTC	/	/	Illegitimate activation of kinase
CTSB	BACE2	Inter	3	1	11	Yes	0,98	0	129,7	CCACCATCAAGAGATCAG	CCACCCAGACAAGACTAC	/	/	Degradation of extracellular matrix and overexpression of BACE2
MYB	AH1	Intra	3	1	13	Yes	0,99	2	5,0	TCTGGAATTGTTGCTGATT	TGAGAAGAGGTTTGTGTTG	MYB, AH1	/	Aberrant proliferation and dedifferentiation of cells through MYB oncogene activation
ATRX	PASD1	Intra	3	1	27	Yes	1,00	0	20,0	TGGCAAGTAACAAAGCAG	TGGTCTGTGAGATAGCAG	/	/	Stabilisation of DNA replication via chromatin remodeling
LMO3	BORCS5	Intra	3	1	11	Yes	1,00	0	17,0	GCAACCAAGATCAAGG	CATTGTGGGGAAGTCTG	LMO3	/	Over-activation of oncogene LMO3
C11orf95	RELA ^a	Intra	3	2	21, 27	Yes	1,00	1	119,0	GTCCACCAAGAACTG	CCCTCGCACTGTAGC	/	/	Aberrant NF-kappa-B transcription program
EWSR1	FLI1 ^a	Inter	3	2	46,11	Yes	1,00	7	93,7	AGTTACCACCCCAAACTG6	CCAAGGGGAGGACTTGTGT	EWSR1, FLI1	/	Regulates the expression of a number of genes important for cancer progression
KANSL1	ARL17A ^a	Intra	3	9	40,47,56,13,18,20,21,2	Yes	/	2	72	GGCTCATGTTCTGACTTG	GACTTCTACTCTCTCCAC	/	/	Cancer predisposition. Linked to histone acetylation
C15orf57	CBX3 ^a	Inter	3	21	38,40,41,43,47,48,2,12; 14; 16,17,18; 20,22,23,25,27,28,31,3	No	0,01	3	16	TCTGTCTGCCAAATCTG	GGAAAGCCACCGACTCAC	/	/	Aberrant epigenetic gene repression
LPP	TPRG1 ^a	Intra	2	1	15	Yes	/	2	12,5	TCTCTGATTCCTGCTGTC	CCCTCTCTTGTCTCTGTC	/	/	Healthy
RASSF4	ZNF22	Intra	1	1	13	Yes	0,86	0	41,0	GAGTACCCGCTGATTTCC	CGCTTAGGCTTGTGTAAC	/	RASSF4	Disruption of RASSF4 tumor suppressor which promotes apoptosis and cell cycle arrest
CUBN	SMAP2	Intra	2	1	16	No	1,00	0	3,5	AGACTGTCTTGGTCAATC	CCCTTAGACTGGCAATCG	/	/	Probable de regulation in cell metabolism
MPP3	PYY	Intra	2	1	17	Yes	0,77	0	1,5	AGAACCACAGCACTTTG	TGACTTCACTCCCAATG	/	/	Increased cell proliferation
UBE4B	CTNBP1	Intra	2	1	17	Yes	0,90	1	5,3	GACCGAGTTGGAATAGAGG	GGAAAACGAGGAACTTAGG	/	UBE4B, CTNBP1	Loss of UBE4B and CTNBP1 tumor suppressor genes
NF2	KIAA0368	Inter	1	1	17	Yes	0,99	0	16,5	AACTCTTAGCACCCCAAGAM	TAAGCTGACACTGAAATCC	/	NF2	Loss of NF2 tumor suppressor gene
ARHGEF7	MYO16	Intra	2	1	32	Yes	1,00	0	4,0	CAGACTGTGCTTCAACG	AAGGGGAAAGGACTCTAG	/	/	Loss of ARHGEF7, a Ras-like family of rho proteins acting as a pro-apoptosis factor
HEV1	NCOA2 ^a	Intra	2	1	33	Yes	1,00	4	72,5	TGATTTGGTGGGAAAGTTC	TGATTTGGTGGGAAAGTTC	/	/	Aberrant expression of genes
TYRO3 ^b	KLHL18	Inter	2	1	18	Yes	0,78	0	7	COAAAGTGCCTGCAAGTG	GGGTAGCAGGGAATGCTTC	/	/	Disruption of TYRO3 domain
JAK2 ^b	PTPRD	Intra	3	1	21	Yes	1,00	0	22	AAAGAACCCTGTGAAAAGTC	ACGTCTCCCTGAGTTAGC	JAK2	JAK2, PTPRD	Disruption of PTPRD that regulates cell growth, differentiation and mitotic cycle
FES ^b	MAN2A2	Intra	2	1	26	Yes	/	0	11	SACAAGTCCCTGTGAAGTG	TTTCCGTGCTCTCTATG	FES	/	Disruption of the end of FES kinase domain
CCM2L	HCK ^b	Intra	2	1	30	Yes	1,00	0	7	GCCTATGATGCCGACTTAG	CCTGGTGTGTTGCTGTGTTG	/	/	Probably overexpression of HCK; increase of cell survival, proliferation and migration
MANBAL	SRC ^b	Intra	2	1	48	Yes	1,00	0	9	GCCTAGTGATCGGACTCTTC	CCCTGGCTCTGCTCTCTATAG	SRC	/	SRC kinase not in frame

Table 2. Description, recurrence, validation and function prediction of the fusion transcripts tested by RT-qPCR after detection by ChimComp on tumor RNA-seq.

To validate our approach, we first focused on the fusions from the “high confidence group” (detected by 3 tools), then divided the list into three categories, based on oncoscore results: oncoscore < 0.35; 0.35-0.7 and > 0.7. Of those, 25% of fusions were selected for validation by RT-qPCR in each category, resulting in 29 fusions in total (top panel). Additionally, 8 fusions detected by less than 3 tools harboring a high oncoscore, except for one control healthy fusion (*LPP-TPRG1*), were tested by RT-qPCR (middle panel). Finally, 5 potentially targetable fusions harboring tyrosine-kinases coding genes were also tested (lower panel). A total of 42 fusions found by ChimComp underwent RT-qPCR assay. As positive controls, some of the well-known fusion oncogenes detected by ChimComp were also included, namely *C11orf95-RELA*, *EWSR1-FLI1* and *HEY1-NCOA2*. A total of 38/42 fusions were successfully validated. For each fusion, the type of chromosomal rearrangement potentially involved, the number of tools detecting the fusion, the number of patients carrying the fusion and their identification numbers, the result of PCR, the oncoscore provided by ChimComp as well as the number of databases reporting the fusion also provided by ChimComp, the primers sequences used, the presence of oncogenes or tumor suppressor genes and the functional prediction are enlisted.

^a Known fusions, ^b with Tyrosine Kinase activity, ^c Intra = intrachromosomal, Inter = interchromosomal, Read nb mean = mean number of junction/spanning reads of all tools.

Patient Number	Gene1 ^a	Gene2 ^a	Fusion validated by RT-qPCR ^b	Oncoscore	Known fusions	Presumptive fusion reconstruction	Predicted biological effect	Additional genomic information from WES ^c	Possible drug	Mechanism	References
2	FBXL7	TERT	yes	0,109682436		Loss of cell cycle arrest mechanism by ubiquitination of Aurora kinases (FBXL7) + overactivation of telomerase (TERT)	Inactivation of FBXL7 and activation of TERT		TERT inhibitor; AURKA inhibitor	Maintenance of telomere ends + Cell cycle deregulation	Bell RJ et al., Understanding TERT Promoter Mutations: A Common Path to Immortality. <i>Mol Cancer Res.</i> 2016 Apr;14(4):315-23; Pilsworth JA, TERT promoter mutation in adult granulosa cell tumor of the ovary. <i>Mod Pathol.</i> 2018 Feb 15. doi:10.1038/s41379-018-
4	ASPCR1 (OG)	TFE3 (OG)	n.d.	0,954511534	x	Known fusion with generation of a chimeric transcription factor leading to MET upregulation	Activation of chimeric transcription factor		MET inhibitor	MET upregulation	Tsuda et al., TFE3 fusions activate MET signaling by transcriptional up-regulation, defining another class of tumors as candidates for therapeutic MET inhibition. <i>Cancer Res.</i> 2007 Feb 1;67(3):919-29.
13	MYB (OG)	AHL1 (TSG)	yes	0,988687802	x	Loss of the major AHL1 (SH3) domain, MYB coding sequence almost completely retained	Inactivation of AHL1	Strong focal amplification of MYB and AHL1	IGF1R inhibitor	Amplification of this locus has been linked to increased drug sensitivity	Pavlicek A et al., Molecular Predictors of Sensitivity to the Insulin-like Growth Factor 1 Receptor Inhibitor Figitumumab (CP-751,871). <i>Mol Cancer Ther.</i> 2013 Dec;12(12):2929-39.
14	ARHGEF7	ATP11A	n.d.	0,99049634	x	Loss of ATP11A regulation by truncation of Inc RNA	Activation of ATP11A	ATP11A amplification	Farnesyltransferase Inhibitors	Drug resistance	Zhang B et al., Resistance to farnesyltransferase inhibitors in Bcr/Abl-positive lymphoblastic leukemia by increased expression of A.D. Marshall, G.C. Grosveld, Alveolar rhabdomyosarcoma – The molecular
14	PAX7	FOXO1 (OG+TSG)	n.d.	0,97944093	x	Known fusion with generation of a chimeric transcription factor	Activation	PAX7, FOXO1 amplification	Inhibitors of FGFR2/4, ALK, MET, IGF1R, BET	Transcriptional activation of target genes	A.D. Marshall, G.C. Grosveld, Alveolar rhabdomyosarcoma – The molecular
15	DMTF1 (TSG)	ABCB1	yes	0,89356784		Loss of complete coding sequence of DMTF and fusion of DMTF promoter to ABCB1	Inactivation of DMTF and activation of ABCB1		Chemotherapy & receptor tyrosine kinase inhibitors	Drug resistance	Beretta et al., Overcoming ABC transporter-mediated multidrug resistance: The dual role of tyrosine kinase inhibitors as multitargeting agents. <i>Eur J Med Chem.</i> 2017 Dec 15;142:271-289.
16	ERRF1 (TSG)	CUBN	yes	0,951965411		Loss of complete coding sequence of ERRF1	Inactivation of ERRF1	ERF11 amplification	EGFR inhibitor	Increased drug sensitivity	Borad et al., Integrated genomic characterization reveals novel, therapeutically relevant drug targets in EGFR and EGFR pathways in
16	ESCO1	CNDP2 (TSG)	n.d.	0,087563145		Loss of only domain (dimerization domain) of CNDP2	Inactivation of CNDP2		MEK inhibitor	Activation of mitogen-activated protein kinase (MAPK) pathway	Zhang Z et al., Underexpressed CNDP2 participates in gastric cancer growth inhibition through activating the MAPK signaling
16	ASAP2	MGMT	n.d.	0,999883057		Truncation in DNA methyltransferase domain of MGMT	Inactivation of MGMT		Temozolomide	Increased drug sensitivity	Wick W et al., MGMT testing - the challenges for biomarker-based glioma treatment. <i>Nat Rev Neurol.</i> 2014 Jul;10(7):372-85.
16	MGMT	C8orf34	n.d.	0,933256552		Loss of 3/4 of MGMT coding sequence	Inactivation of MGMT		Temozolomide	Increased drug sensitivity	Wick W et al., MGMT testing - the challenges for biomarker-based glioma treatment. <i>Nat Rev Neurol.</i> 2014 Jul;10(7):372-85.
16	ZNF208	NRG2	n.d.	0,915020583		Complete coding sequence of NRG2 is kept resulting in overexpression of NRG2	Activation		Tyrosine kinase inhibitors targeting ErbB receptors	Activation of ErbB receptor tyrosine kinases	Carraway CL 3rd et al., Neuregulin-2, a new ligand of ErbB3/ErbB4-receptor tyrosine kinases. <i>Nature.</i> 1997 May 29;387(6652):512-6.
17	TP53 (TSG)	ZNF565	yes	0,956287477		Loss of complete coding sequence of TP53	Inactivation	Heterozygous deletion of TP53 and ZNF565	Inhibitors of WEE1	WEE1 sensitizes cells to DNA-damaging agents via G2 checkpoint inhibition in p53-deficient cells	Hirai H et al., Small-molecule inhibition of Wee1 kinase by WK-1775 selectively sensitizes p53-deficient tumor cells to DNA-damaging agents. <i>Mol Cancer Ther.</i> 2009 Nov;8(11):2992-3000; Aarts M et al., Forced mitotic entry of S-phase cells as a therapeutic strategy
21	JAK2 (OG)	PTPRD (TSG)	n.d.	0,999997433		Loss of PTPRD, out-of-frame	Inactivation of PTPRD	Heterozygous PTPRD deletion	STAT3 inhibitor; AURKA inhibitor; BET inhibitors	PTPRD dephosphorylates STAT3 and destabilizes AURKA and MYCN	Zhao S et al., Genetic alterations of protein phosphatases in human cancers. <i>Oncogene.</i> 2015 Jul 23; 34(30): 3885-3894; Meehan M et al., Protein tyrosine phosphatase receptor delta acts
26	ATRX	PASD1	n.d.	0,997390308		Truncation in helicase C domain	Inactivation	Heterozygous ATRX deletion	PARP inhibitors	PARP overexpression	Fazal-Salom et al., Mutations in ATRX increase genetic instability and sensitivity to PARP inhibitors in pediatric glioblastoma cells.
26	SEMA3D	ABCB1	yes	/		Probably overexpression of ABCB1 due to SEMA3D promoter activity or loss of regulatory sequences in 5' UTR of ABCB1	Activation	ABCB1 amplification	Chemotherapy & receptor tyrosine kinase inhibitors	Drug resistance	Beretta et al., Overcoming ABC transporter-mediated multidrug resistance: The dual role of tyrosine kinase inhibitors as multitargeting agents. <i>Eur J Med Chem.</i> 2017 Dec 15;142:271-289.
32	MYO16	IRS2	n.d.	0,999756812		Loss of all domains of IRS2 and of the last two domains of MYO16	Inactivation of IRS2	6 kb amplicon including both genes with high-level amplification	IGF1R inhibitor	Amplification of this locus has been linked to increased drug sensitivity	Pavlicek A et al., Molecular Predictors of Sensitivity to the Insulin-like Growth Factor 1 Receptor Inhibitor Figitumumab (CP-751,871). <i>Mol Cancer Ther.</i> 2013 Dec;12(12):2929-39.
40	NF2 (TSG)	AK131325	yes	/		Loss of NF2 which acts as negative regulator of HIPPO/YAP pathway and regulates activity of several receptor-tyrosine kinases	Inactivation		Inhibitors of pathways regulated by HIPPO/YAP- pathway including AKT/mTOR inhibitors and FAK inhibitors.	Pathway activation of major oncogenic pathways	Petrilli A, Fernandez-Valle C, Role of Merlin/NF2 inactivation in Tumor Biology. <i>Oncogene.</i> 2016 Feb 4; 35(5): 537-548; Shapiro IM et al., Merlin deficiency predicts FAK inhibitor sensitivity; a synthetic lethal relationship. <i>Sci Transl Med.</i> 2014;6:237ra68.
11, 46	EWSR1 (OG)	FLI1 (TSG)	yes	0,999925015	x	Known fusion with generation of a chimeric transcription factor	Activation of chimeric transcription factor		PARP inhibitor; BET inhibitor	Widespread transcriptional reprogramming of Ewing sarcoma cells required for proliferation and tumorigenesis	Brenner JC et al., PARP-1 inhibition as a targeted strategy to treat Ewing's sarcoma. <i>Cancer Res.</i> 2012 Apr 1;72(7):1608-13; Loganathan SN et al., BET bromodomain inhibitors suppress EWS-FLI1-dependent transcription and the IGF1 autocrine mechanism in
28, 29, 38	PAX3 (OG)	FOXO1 (OG+TSG)	n.d.	0,998640978	x	Known fusion with generation of a chimeric transcription factor	Activation	#29: PAX3 amplification	Inhibitors of FGFR2/4, ALK, MET, IGF1R, BET	Transcriptional activation of target genes	A.D. Marshall, G.C. Grosveld, Alveolar rhabdomyosarcoma – The molecular

Table 3. Potentially targetable fusion alterations with associated putative biological effect and predicted impact on drug sensitivity.

Fusion transcripts representing potential targetable alterations by existing anti-cancer drugs were found among the high-confidence list candidates and are enlisted with the associated patient number, the result of PCR validation, the associated oncoscore provided by ChimComp, the presumptive fusion reconstitution and biological effect, the possible drugs to target the fusion products, the mechanism involved, and the references in literature used. Additional relevant genomic information from WES of patients was also added when relevant regarding the associated potentially targetable fusion.

^a In bold: Gene whose alteration induced by the fusion has potential therapeutic implications ^bn.d. = not done, ^cGenomic results from concomitant WES if available (Harttrampf et al., Clin Can Res 2017).

Figures

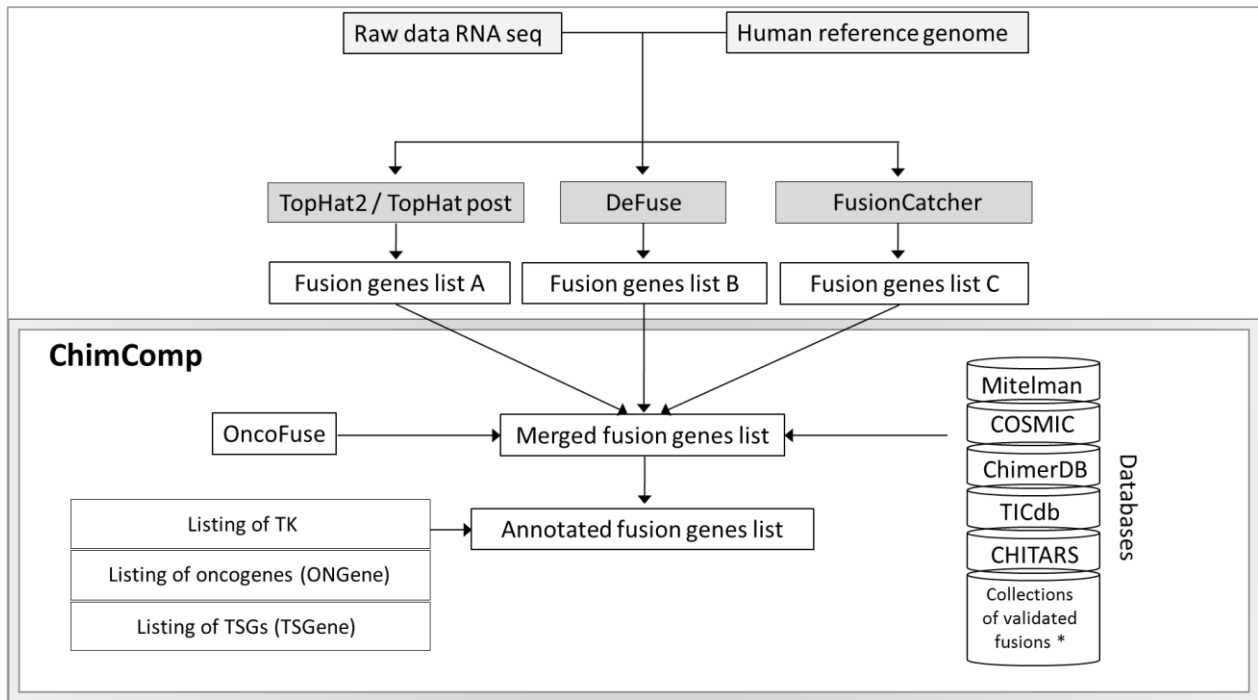


Figure 1. ChimComp pipeline workflow.

Raw data from RNAseq were mapped on the human reference genome GRCh37/hg19 (Genome Reference Consortium Human Build 37/human genome 19). Three software tools (TopHat2 followed by TopHat post, DeFuse and Fusion Catcher) were used to generate fusion genes lists (A, B and C). ChimComp processed the three lists in order to provide a unique merged fusion genes list per patient. The listing was analyzed by Oncofuse (left panel) to generate the oncoscores and compared to databases of known fusions (right panel). TK genes, oncogenes and tumor suppressor genes were searched in the final annotated fusion genes list. * Collection of fusions according to Babak Alaei-Mahabadi et al., 2016; TK tyrosine kinases; TSGs tumor suppressor genes.

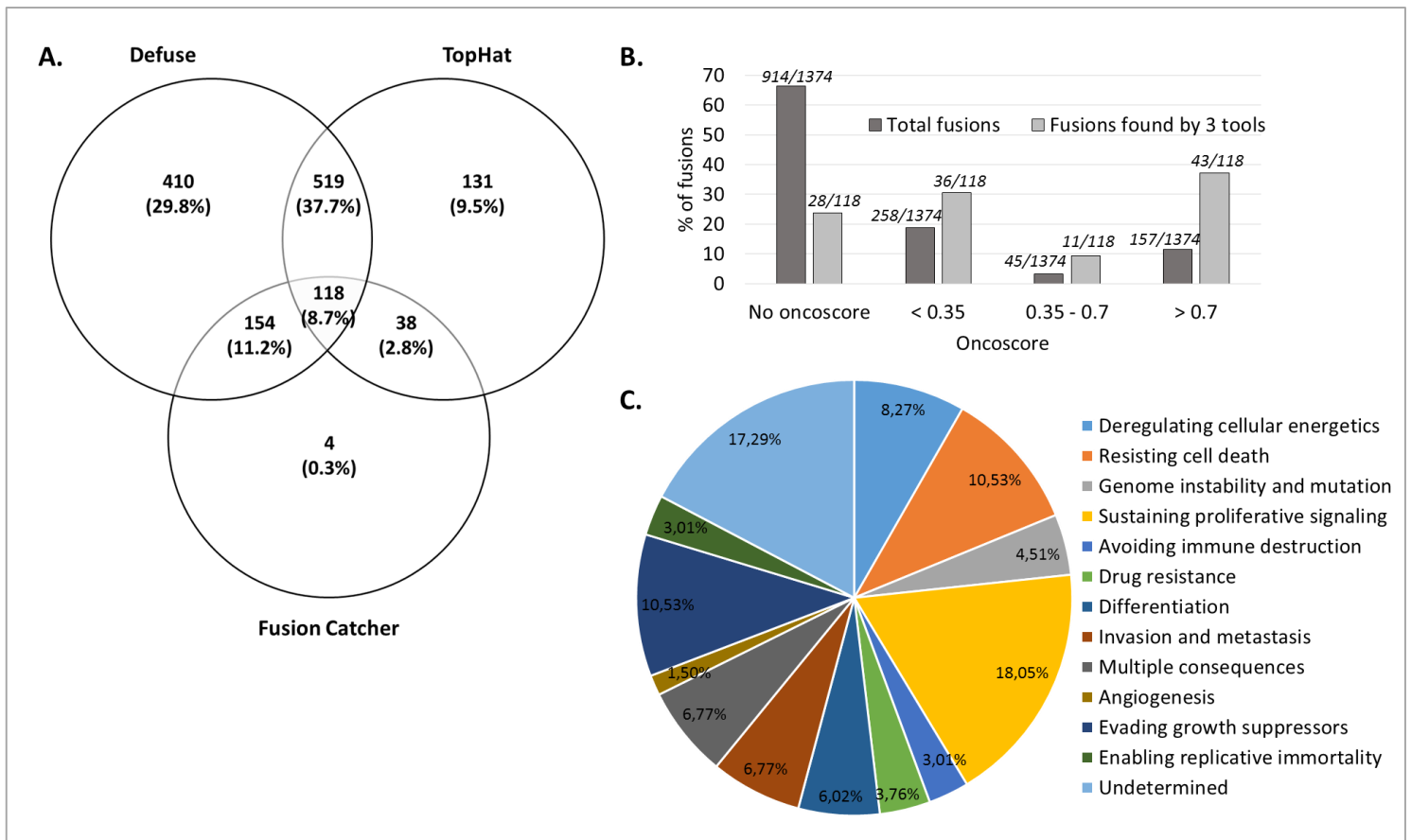


Figure 2. Analysis of the fusion transcripts found by ChimComp pipeline.

A. Venn diagram of the non-redundant 1 374 fusions detected from RNA-sequencing data of 48 pediatric patients analyzed by 3 tools (Defuse, TopHat and Fusion Catcher). B. Repartition of fusions according to oncoscore. The dark grey bars represent the total fusions (n = 1374) and the light grey bars represent the high-confidence group found by 3 tools (n = 118). C. Distribution of predicted biological alterations of the high confidence candidates group. Classification of the fusion transcripts was achieved according to the “hallmarks of cancer” described by Hanahan et al., 2011.

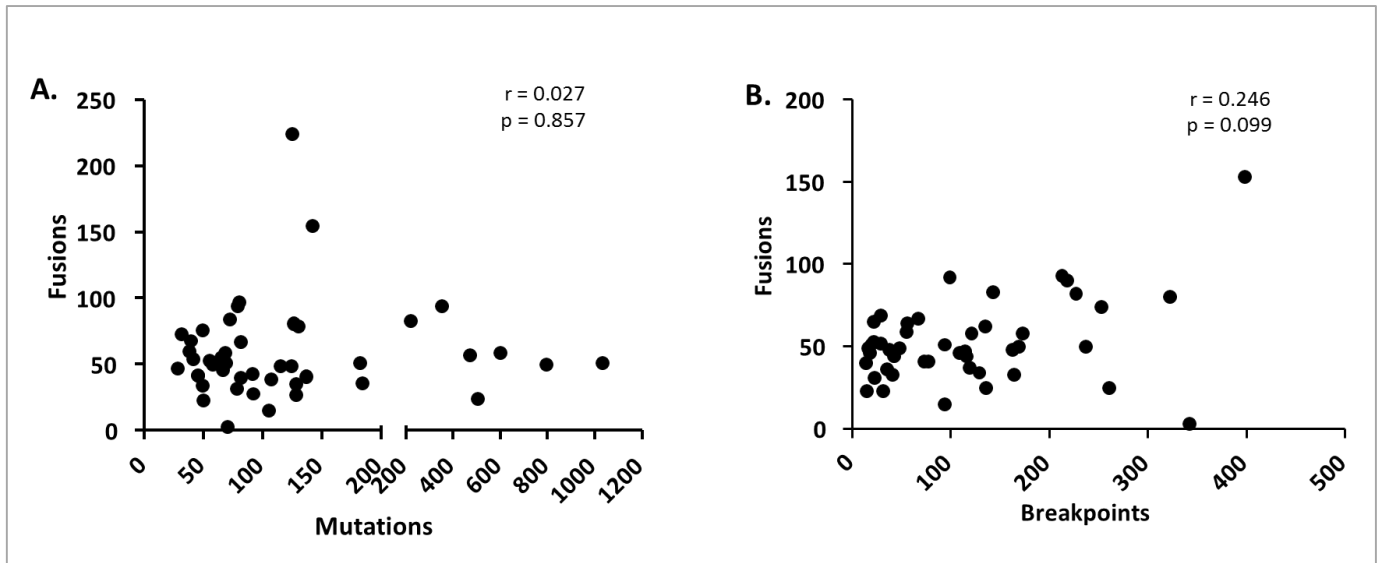


Figure 3. Relationship between fusions and mutations or genomic breakpoints in relapsed pediatric cancer patients.

The Spearman correlation coefficient was calculated to assess A) the relation between total fusions number (Y axis) and mutations (X axis) and B) the relation between total fusions number (Y axis) and genomic breakpoints (X axis). A weak correlation was found between the number of breakpoints and fusions (B).

r = Spearman correlation coefficient

p = p-value of Spearman test

r^2 = regression coefficient of linear regression curve with corresponding equation.

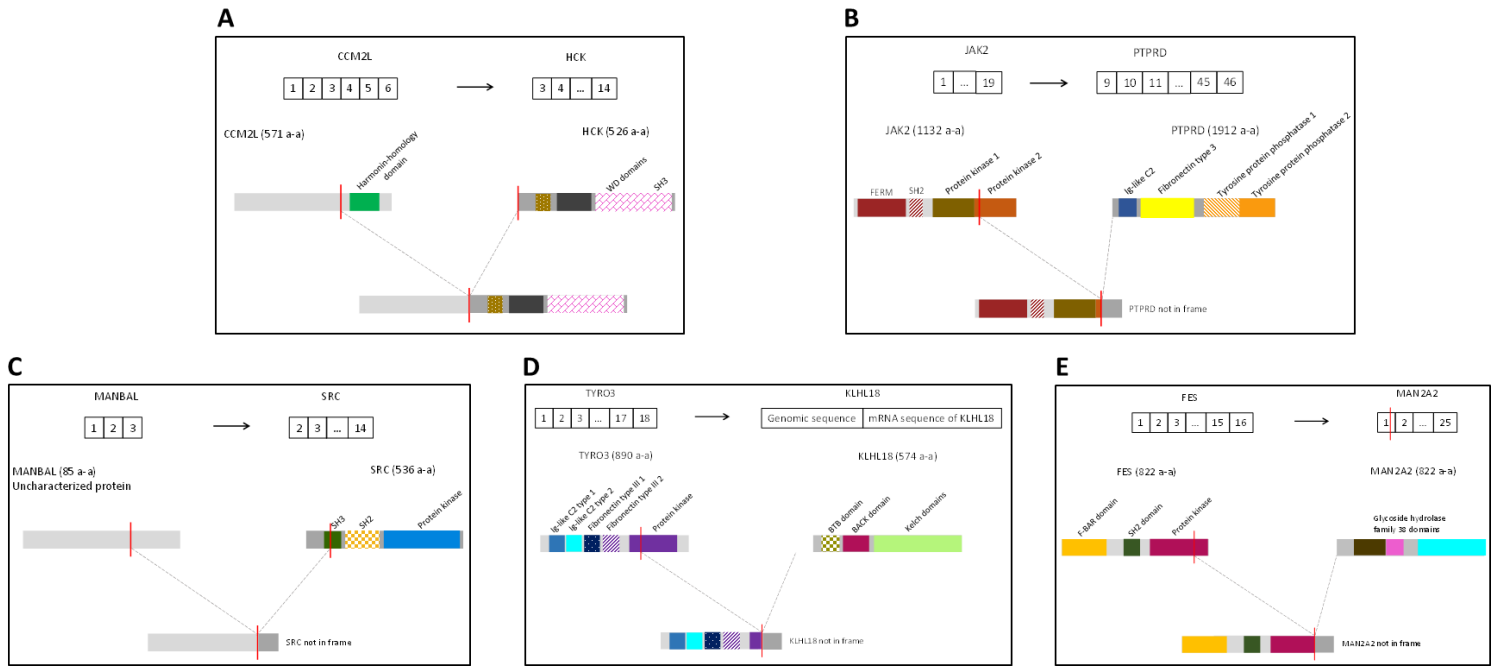


Figure 4. Graphical reconstitution of the mRNA and protein of 5 fusions involving tyrosine kinases genes.

The National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST; <http://www.ncbi.nlm.nih.gov/blast>) was used to obtain the putative mRNA sequence which was reconstituted with the use of Nucleotide (NCBI). When the coding sequence was obtained and predicted, Uniprot Knowledge Base (<http://www.uniprot.org/>) was used to reconstitute the predicted fusion protein domains. A. CCM2L-HCK; B. JAK2-PTPRD; C. MANBAL-SRC, D. TYRO3-KLHL18 and E. FES-MAN2A2. Framed numbers represent the exons corresponding to the gene transcript written above each group of numbers. The arrows represent the junction point between the two transcripts to form the fusion. Colored boxes represent the two entire proteins involved in the fusion, with the different functional domains. The red bar shows the junction point and dotted lines illustrates the protein rearrangement after the fusion transcript formation. Intronic sequence = a transcript with a retained genomic intronic sequence of the corresponding gene.

Supplementary tables

Patient No° MOSCATO [Harttrampf et al., Clin Can Res 2017]	Patient number used in this study	Diagnosis	Tumor type (S= Sarcoma, B= Brain tumors, O= Others)	Age at inclusion (years)	Sex	Number of previous treatments	Time between initial diagnosis and tumor sample collection (months)	Metastases at initial diagnosis	Metastasis at MOSCATO-01 study inclusion	Type of collected tumor tissues: Primary tumor (P)/Metastasis (M)	Tumor cells (%)
18	1	Nephroblastoma	O	5,1	M	3	49,4	no	yes	M	30
19	2	Sertoli Leydig granulosa cell tumor	O	19,9	F	5	51,1	no	yes	M	90
20	3	Astroblastoma	B	5,8	M	3	25,8	no	yes	M	60
21	4	Alveolar soft part sarcoma	S	20,1	F	2	36,4	yes	yes	M	30
23	5	Nephroblastoma	O	8,5	M	1	27,7	yes	yes	M	80
24	6	Primitive neuroectodermal tumor	B	6,5	F	4	56,1	no	yes	M	50
25	7	Hepatoblastoma	O	4,4	M	2	15,1	no	yes	M	70
26	8	Embryonal rhabdomyosarcoma	S	12,9	M	1	36,6	yes	yes	M	30
29	9	Fibrolamellar hepatocellular carcinoma	O	21,8	F	5	79,5	no	yes	M	60
30	10	Embryonal rhabdomyosarcoma	S	0,8	M	2	8,9	yes	yes	M	60
31	11	Ewing sarcoma	S	20,9	M	3	55,4	yes	yes	M	70
32	12	Ewing-like sarcoma	S	14,8	F	2	18,4	no	yes	M	30
33	13	Embryonal rhabdomyosarcoma	S	4,1	M	3	37,5	yes	yes	M	70
34	14	Alveolar rhabdomyosarcoma	S	16,1	F	1	25,5	yes	yes	M	60
35	15	Epithelioid sarcoma	S	17,9	F	2	11,7	no	yes	M	70
36	16	Medulloblastoma	B	11,4	M	2	32,1	no	yes	M	80
37	17	Osteosarcoma	S	5,0	M	3	5,7	no	yes	M	80
38	18	Anaplastic Ependymoma III*	B	6,0	M	2	43,3	no	yes	P	80
39	19	Spindle epithelial tumor with thymus-like differentiation	O	10,8	F	1	77,6	no	yes	M	90
40	20	Embryonal rhabdomyosarcoma; LFS*	S	1,3	F	2	2,4	no	no	P	80
41	21	Ependymoma	B	11,6	F	3	106,2	no	no	P	80
42	22	High grade glioma	B	8,5	F	1	30,4	no	yes	M	80
43	23	High grade glioma (DIPG) §	B	4,8	F	1	31,4	no	no	P	70
44	24	Hepatoblastoma	O	12,9	M	2	8,0	no	no	P	60
45	25	Low grade glioma (Ganglioglioma)	B	7,1	F	3	35,8	no	no	P	80
47	26	Neuroblastoma	O	18,0	F	8	54,0	no	yes	M	70
48	27	Ependymoma (revised from low grade glioma)	B	17,8	M	4	125,6	no	no	P	30
49	28	Alveolar rhabdomyosarcoma	S	14,3	F	3	15,3	no	yes	M	70
50	29	Alveolar rhabdomyosarcoma	S	4,7	F	1	25,2	yes	no	P	70
52	30	Medulloblastoma	B	12,8	F	4	94,3	no	yes	P	95
53	31	High grade glioma (GBM) #	B	20,8	F	1	23,5	no	yes	P	90
55	32	Medulloblastoma; LFS*; Williams-Beuren syndrome	B	10,3	F	1	21,0	yes	no	P	80
56	33	Mesenchymal chondrosarcoma	S	23,3	F	2	124,4	yes	yes	M	90
57	34	Ependymoma	B	8,3	M	3	80,5	no	no	P	90
58	35	Embryonal rhabdomyosarcoma	S	4,1	M	1	10,8	no	yes	M	80
59	36	Neuroblastoma	O	4,3	F	3	17,2	yes	yes	M	30
60	37	Medulloblastoma	B	10,8	F	2	60,9	no	no	P	100
61	38	Alveolar rhabdomyosarcoma	S	14,8	F	2	16,4	yes	yes	M	80
62	39	Desmoplastic small round cell tumor	S	17,7	F	1	31,3	no	yes	M	60
63	40	Osteosarcoma	S	17,4	M	2	5,8	no	yes	M	90
65	41	Malignant peripheral nerve sheath tumor; NF1*	O	13,3	M	3	25,1	no	yes	M	90
66	42	Medulloblastoma; Gorlin-Goltz syndrome	B	6,8	M	3	25,7	no	yes	M	90
67	43	Abrakosof tumor	O	17,0	F	n.a.	n.a.	yes	yes	P	90
68	44	Neuroblastoma	O	5,1	M	2	13,2	yes	yes	M	90
69	45	Medulloblastoma	B	9,8	F	1	80,3	no	yes	M	90
70	46	Ewing sarcoma	S	13,9	F	4	15,8	yes	yes	M	50
71	47	Embryonal rhabdomyosarcoma	S	1,3	M	2	4,5	no	no	P	80
72	48	High grade glioma (GBM) #	B	11,0	F	2	23,1	no	no	P	50

Table S1: Clinical characteristics of the pediatric patients at relapse enrolled in MOSCATO-01 clinical trial.

Identification numbers used for the initial MOSCATO-01 pediatrics study (from Harttrampf et al. 2017), identification used for the present fusion study, associated diagnoses, age, sex, number of previous treatments, time between initial diagnosis and biopsy, presence of metastasis as well as the type of tissue collected [*i.e* primary tumor (P) or metastasis (M)] were reported. Patients were classified in 3 groups according to tumor types: sarcoma (S), brain tumors (B) and others (O). The percentage of tumor cells content in each biopsy used for further studies was estimated by a pathologist as part of the clinical trial MOSCATO-01.

Diagnoses in italics represent constitutional conditions; * Li-Fraumeni syndrome; \$ Diffuse intrinsic pontine glioma; # Glioblastoma multiforme; ° Neurofibromatosis Type 1.

Oncogene	Fusion gene 1	Fusion gene 2	Number of patients carrying fusion	Pathology	Oncoscore
ACTN4	EIF3K	ACTN4	3	High grade glioma (GBM), High grade glioma (DIPG), Medulloblastoma	0,84132184
AFF3	RNF149	AFF3/LOC728323 intergenic region	1	Nephroblastoma	/
AFF3	AFF3	SEC14L2	1	Osteosarcoma	0,09362091
AGAP2	AGAP2	AVIL	1	Neuroblastoma	0,00511378
AGAP2	AGAP2	LOC100507250	1	Neuroblastoma	/
AH11	MYB	AH11	1	Embryonal rhabdomyosarcoma	0,9886878
AKAP13	KLHL25	AKAP13	1	Nephroblastoma	0,99225852
AKAP9	SLC46A3/CHAMP1 intergenic region	AKAP9	5	High grade glioma (GBM), High grade glioma (DIPG), Low grade glioma (Ganglioglioma), Mesenchymal chondrosarcoma, Ependymoma	/
ASPCR1	ASPCR1	TFE3	1	Alveolar soft part sarcoma	0,95451153
BAALC	BAALC	MFHAS1	1	High grade glioma	0,43591361
BCR	BCR	LRP5L	1	Osteosarcoma	0,74481894
CDK4	CDK4	LINC00486	1	Neuroblastoma	/
CSNK2A1	CSNK2A1	GALNT18	1	Medulloblastoma	0,99630797
CTBP2	CTBP2	OSTM1	1	Fibrolamellar hepatocellular carcinoma	/
CTNND2	CTNND2	LOC401177/BC036251 intergenic region	1	High grade glioma (GBM)	/
ECHS1	BC042451/MGC2752 intergenic region	ECHS1	1	High grade glioma (GBM)	/
ERAS	ERAS	HDAC6	1	Neuroblastoma	0,09033228
EWSR1	EWSR1	WT1	1	Desmoplastic small round cell tumor	0,43323828
EWSR1	EWSR1	FLJ1	2	Ewing sarcoma	0,99992502
FDP5	FDP5	PKLR/PGBD2 intergenic region	1	Hepatoblastoma	/
FES	FES	MAN2A2	1	Neuroblastoma	/
FLJ1	EWSR1	FLJ1	2	Ewing sarcoma	0,99992502
FOXO1	PAX7	FOXO1	1	Alveolar rhabdomyosarcoma	0,97944093
FOXO1	PAX3	FOXO1	3	Alveolar rhabdomyosarcoma	0,99864098
GNA12	GNA12	SUN1	1	Osteosarcoma	0,98231051
GNAS	INS-IGF2	GNAS	1	Embryonal rhabdomyosarcoma	/
GPM6A	VPS35	GPM6A	1	Anaplastic Ependymoma III*	0,22516328
GPM6B	GK	GPM6B	1	High grade glioma (DIPG)	0,01677199
GPM6B	GPM6B	ACTR3B/VIPR2 intergenic region	4	High grade glioma (GBM), Primitive neuroectodermal tumor	/
HMGA2	SRGAP1	HMGA2	1	High grade glioma (GBM)	0,2869805
HMGA2	HMGA2	AK055974/ANHX intergenic region	1	High grade glioma (GBM)	/
HMGA2	U3/VIPR2 intergenic region	HMGA2	1	High grade glioma (GBM)	/
IGF2	INS-IGF2	SEC31A	1	Sertoli Leydig granulosa cell tumor	/
IGF2	INS-IGF2	PPHLN1	1	Ependymoma (revised from low grade glioma)	/
IGF2	AK125040	INS-IGF2	1	Embryonal rhabdomyosarcoma	/
IGF2	INS-IGF2	GNAS	1	Embryonal rhabdomyosarcoma	/
IRF2	IRF2	LOC401164	1	Hepatoblastoma	/
IRS2	MYO16	IRS2	1	Medulloblastoma; <i>LFS*</i> ; <i>Williams-Beuren syndrome</i>	0,99975681
JAK2	JAK2	PTPRD	1	Ependymoma	0,99999743
JAK2	JB137814/PGBD2 intergenic region	JAK2	1	Neuroblastoma	/
JUND	JUND	LOC100506190/DQ577244 intergenic region	1	Anaplastic Ependymoma III*	/
KSR2	KSR2	CRMP1	1	Neuroblastoma	0,27557443
KSR2	KSR2	C8orf33/Chromosome_end intergenic region	1	Neuroblastoma	/
LMO1	LMO1	RIC3	3	High grade glioma (GBM), Ependymoma	0,99365086
LMO3	LMO3	LOH12CR1	1	Ewing sarcoma	0,99832794
MCC	LINC00266-1	MCC	1	Astroblastoma	/
MDM2	MDM2	NUP107	1	Neuroblastoma	0,85881908
MDM2	CPSF6	MDM2	1	Neuroblastoma	0,95873015
MDM2	FRS2	MDM2	1	Neuroblastoma	0,9737676
MDM2	CTDSP2	MDM2	1	Neuroblastoma	0,98908188
MFHAS1	BAALC	MFHAS1	1	High grade glioma	0,43591361
MIR100HG	MIR100HG	BC089451	5	Ependymoma, High grade glioma (DIPG), Low grade glioma (Ganglioglioma), Embryonal rhabdomyosarcoma, Medulloblastoma	/
MRE11A	MRE11A	ANKRD49/AK130852 intergenic region	1	Anaplastic Ependymoma III*	/
MTDH	MTDH	5S_rRNA/C8orf33 intergenic region	1	Mesenchymal chondrosarcoma	/
MYB	MYB	IL20RA	1	Embryonal rhabdomyosarcoma	0,96050213
MYB	MYB	AH11	1	Embryonal rhabdomyosarcoma	0,9886878
MYB	MYB	AK093525/LOC728323 intergenic region	1	Embryonal rhabdomyosarcoma	/
MYB	MYB	MYB/BC036251 intergenic region	1	Embryonal rhabdomyosarcoma	/
NEDD4	NEDD4	DQ601182/DDX11L9 intergenic region	1	Alveolar soft part sarcoma	/
NFIB	NFIB	TRNA/DQ577244 intergenic region	3	Medulloblastoma	/
NINL	NINL	PCDH18	1	High grade glioma (GBM)	0,73069713
PARK7	PARK7	MINOS1	1	High grade glioma (GBM)	6,06E-04
PARK7	PARK7	Metazoa_SRP/PGBD2 intergenic region	1	High grade glioma (GBM)	/
PAX3	PAX3	FOXO1	3	Alveolar rhabdomyosarcoma	0,99864098
PPARG	TUBB3	PPARGC1A	1	Anaplastic Ependymoma III*	0,10315707
PPARG	SYN2	PPARG	1	Alveolar rhabdomyosarcoma	/
PRKCA	BCYRN1	PRKCA	1	Nephroblastoma	/
PRKCI	ARMCX4/DDX11L family member intergenic region	PRKCI	1	High grade glioma (GBM)	/
RERE	RERE	ERI3	1	Medulloblastoma	0,90255114
RERE	RERE	PRKCZ	1	Medulloblastoma	0,99999744
RERE	RERE	LOC339442/PGBD2 intergenic region	1	Medulloblastoma	/
SET	SET	BC107094/BC036251 intergenic region	1	Alveolar rhabdomyosarcoma	/
SETBP1	SETBP1	AK095310	1	High grade glioma (DIPG)	/
SRC	MANBAL	SRC	1	High grade glioma (GBM)	0,99999645
SRSF3	SRSF3	LOC100500773	1	Alveolar rhabdomyosarcoma	/
SWAP70	SWAP70	BBOX1/AK130852 intergenic region	9	/	/
TCF3	TCF3	MIDN/MGC2752 intergenic region	1	Medulloblastoma; <i>LFS*</i> ; <i>Williams-Beuren syndrome</i>	/
TFE3	ASPCR1	TFE3	1	Alveolar soft part sarcoma	0,95451153
TFG	TFG	GPR128	1	Anaplastic Ependymoma III*	0,1481096
THRA	IKZF3	THRA	1	Alveolar rhabdomyosarcoma	0,80872298
TL1	ZC4H2/DDX11L family member intergenic region	TL1	3	Abrikossof tumor, High grade glioma (GBM), Osteosarcoma	/
TSPY1	TSPY1	CADPS	1	Medulloblastoma	/
UBE3C	CCDC132/VIPR2 intergenic region	UBE3C	1	Medulloblastoma	/
UBE3C	UBE3C	DHRS2/LOC90925 intergenic region	1	Medulloblastoma	/
VIM	VIM	LINC00486	1	Abrikossof tumor	/
VIM	OR51E2/AK130852 intergenic region	VIM	8	/	/
WT1	EWSR1	WT1	1	Desmoplastic small round cell tumor	0,43323828
YEATS4	FRS2	YEATS4/ANHX intergenic region	1	Neuroblastoma	/
YEATS4	YEATS4	CPM/ANHX intergenic region	1	Neuroblastoma	/
YWHAZ	FGF12	YWHAZ	1	Embryonal rhabdomyosarcoma, Fibrolamellar hepatocellular carcinoma	0,98987825
ZNF268	CAP1	ZNF268	1	Ependymoma (revised from low grade glioma)	0,85780377

Table S2: List of fusions involving at least one oncogene among the total non-redundant 1374 fusions detected by ChimComp. (source: Yining Liu, Jingchun Sun, Min Zhao (2017). *ONGene, a literature-based database for human oncogenes. Journal of Genetics and Genomics.*)

Tumor suppressor gene	Fusion gene 1	Fusion gene 2	Number of patients carrying fusion	Pathology	Oncoscore
ANAPC1	ANAPC1	RMND5A	1	Embryonal rhabdomyosarcoma	/
ARHGAP29	ARHGAP29	ARHGAP29/PGBD2 intergenic region	1	Neuroblastoma	/
AZGP1	AZGP1	GJC3	6	Several patients	/
BCR	BCR	LRP5L	1	Osteosarcoma	0.74
CACNA2D3	SELK/DQ575955 intergenic region	CACNA2D3	1	Neuroblastoma	/
CADM2	CADM2	XIRP1/DQ575955 intergenic region	1	Neuroblastoma	/
CADM4	CADM4	ZNF576	1	Ependymoma (revised from low grade glioma)	/
CAMTA1	CAMTA1	H6PD	2	Several patients	0.14
CAMTA1	RSU1	CAMTA1	1	Medulloblastoma	0.97
CAMTA1	CAMTA1	VAMP3	1	High grade glioma (GBM)	0.99
CD4	CD4	GPR162	1	Osteosarcoma	/
CD44	CD44	TRIM5	1	Neuroblastoma	/
CDH13	BTBD9	CDH13	1	Sertoli Leydig granulosa cell tumor	0.22
CNDP2	ESCO1	CNDP2	1	Medulloblastoma	0.08
CTNNA3	RBF0X2	CTNNA3	1	Ependymoma	0.21
CTNNBIP1	UBE4B	CTNNBIP1	1	Osteosarcoma	0.9
CUL5	CUL5	ACAT1	1	Alveolar rhabdomyosarcoma	0.99
DDX3X	DDX3X	SLC27A5	1	Anaplastic Ependymoma III*	/
DDX58	DDX58	TOPORS-AS1	1	Medulloblastoma	/
DMTF1	DMTF1	ABCB1	1	Epithelioid sarcoma	0.89
DNAJB1	DNAJB1	PRKACA	1	Fibrolamellar hepatocellular carcinoma	0.998
DNAJC11	RIMS3	DNAJC11/PGBD2 intergenic region	1	Medulloblastoma	/
DNAJC11	DNAJC11	Mir_584/DUX4L3 intergenic region	1	Medulloblastoma	/
DNAJC11	DNAJC11	ERRF1/PGBD2 intergenic region	1	Medulloblastoma	/
DND1	KANS1L1-AS1/FLJ43681 intergenic region	DND1	1	Neuroblastoma	/
EAF2	EAF2	EAF2/DQ575955 intergenic region	1	High grade glioma (GBM)	/
EDNRB	UGGT2	EDNRB	1	Embryonal rhabdomyosarcoma; LFS*	/
ERFR1	ERFR1	CUBN	1	Medulloblastoma	0.95
ERFR1	DNAJC11	ERFR1/PGBD2 intergenic region	1	Medulloblastoma	/
EYA4	ATP4A/MGC2752 intergenic region	EYA4	14	Several patients	/
FAM188A	FAM171A1	FAM188A/DUX4L3 intergenic region	1	Medulloblastoma	/
FAM188A	FAM188A	SIL1	1	Medulloblastoma	0.006
FAM188A	FAM188A	FAM171A1	1	Medulloblastoma	0.022
FAM188A	FAM188A/DUX4L3 intergenic region	PPIEL	1	Medulloblastoma	/
FBXO25	FBXO25	AK124970/PGBD2 intergenic region	1	Medulloblastoma	/
FBXO25	LOC728323	FBXO25	2	Several patients	/
FOXO1	PAX3/7	FOXO1	3	Several patients	0.99
GLIPR1	GLIPR1	GMPR	2	Several patients	0.06
GPCS5	GPCS5	GPCS-AS1/CHAMP1 intergenic region	1	Alveolar soft part sarcoma	/
GPCS5	GPCS5-AS1/CHAMP1 intergenic region	ATP11A	1	Alveolar rhabdomyosarcoma	/
GPCS5	ATP11A	GPCS5	1	Alveolar rhabdomyosarcoma	/
GPCS5	GPCS5	ABCC4/CHAMP1 intergenic region	1	Embryonal rhabdomyosarcoma; LFS*	/
HDAC3	HDAC3	DIAPH1	2	Several patients	0.87
HINT1	OSBP16/ANHX intergenic region	HINT1	1	High grade glioma (GBM)	/
IGFBP5	IGFBP5	LINC00486	1	Abrikossof tumor	/
IKZF3	IKZF3	THRA	1	Alveolar rhabdomyosarcoma	0.81
ING5	ING5	AK126180/LOC728323 intergenic region	1	Embryonal rhabdomyosarcoma	/
ITGAV	ITGAV	FAM171B	1	Epithelioid sarcoma	0.87
KISS1	GOLT1A	KISS1	1	Fibrolamellar hepatocellular carcinoma	/
KMT2C	FZD1/VIPR2 intergenic region	KMT2C	1	Nephroblastoma	/
KMT2C	DQ599872/VIPR2 intergenic region	KMT2C	1	Hepatoblastoma	/
KMT2C	PARP15/DQ575955 intergenic region	KMT2C	1	Neuroblastoma	/
KMT2C	CO9	KMT2C	4	Several patients	/
KMT2C	UG/DDX11L family member intergenic region	KMT2C	2	Several patients	/
LAT2	SNX18	LAT2	1	High grade glioma (GBM)	0.59
LIFR	LIFR	EGFLAM	1	High grade glioma	0.91
MARCKS	MARCKS	ANKS1B	1	Medulloblastoma	0.18
MAX	PPP2R5E	MAX	1	High grade glioma (GBM)	0.998
MCC	LINC00266-1	MCC	1	Astroblastoma	/
MEN1	MEN1	PGM2L1/AK130852 intergenic region	1	Malignant peripheral nerve sheath tumor; NF1*	/
MIA2	MIA2	CTAGE5	1	Hepatoblastoma	0.03
MTAP	MTAP	CDKN2B-AS1	1	Embryonal rhabdomyosarcoma	/
MYO1A	NAB2	MYO1A	1	High grade glioma (GBM)	0.06
NDN	POLR3E	NDN	1	Malignant peripheral nerve sheath tumor; NF1*	0.04
NEDD4	NEDD4	DQ601182/DDX11L9 intergenic region	1	Alveolar soft part sarcoma	/
NF1	NF1	RAB11FIP4	1	Embryonal rhabdomyosarcoma	0.92
NF2	KIAA0368	NF2	1	Osteosarcoma	0.99
NF2	NF2	AK131325	1	Osteosarcoma	/
NIT2	TBC1D23	NIT2	2	Several patients	0.12
NOTCH2	ATP4A/MGC2752 intergenic region	NOTCH2	2	Several patients	/
OPCML	OPCML	NTM	3	Several patients	0.007
OSGIN1	OSGIN1	NECAB2	1	Hepatoblastoma	0.003
PARK7	PARK7	MINOS1	1	Embryonal rhabdomyosarcoma	0
PARK7	PARK7	Metazoan_SRP/PGBD2 intergenic region	1	Embryonal rhabdomyosarcoma	/
PCDH17	CPE	PCDH17	1	High grade glioma (GBM)	0.12
PCDH9	PCDH9	SLC22A16	1	Embryonal rhabdomyosarcoma; LFS*	0.002
PCDH9	PCDH9	RTN4IP1	1	Embryonal rhabdomyosarcoma; LFS*	0.002
PDSSB	PDSSB	STARD13	1	Medulloblastoma	0.29
PDSS2	FNDC3A	PDSS2	1	Embryonal rhabdomyosarcoma; LFS*	0.81
PHLPP1	GAREM	PHLPP1	1	Medulloblastoma	0.12
PHLPP1	PHLPP1	RTTN	1	Medulloblastoma	0.05
PHLPP1	PHLPP1	GALR1/PARD6G intergenic region	1	Medulloblastoma	/
PLEKHO1	VPS45	PLEKHO1	14	Several patients	0.04
PNN	PNN	LINC00665	1	High grade glioma (GBM)	/
POU2F3	OAF	POU2F3	1	Abrikossof tumor	0.7
POU6F2	POU6F2	PPP1R1C	3	Several patients	0.002
PPARG	SYN2	PPARG	1	Alveolar rhabdomyosarcoma	/
PRICKLE1	KCNMB4	PRICKLE1	1	High grade glioma (GBM)	0.89
PRICKLE1	PRICKLE1	BCO42876/LOC728323 intergenic region	1	High grade glioma (GBM)	/
PRICKLE1	PRIM1	PRICKLE1/ANHX intergenic region	1	High grade glioma (GBM)	/
PRICKLE1	ZDHHC17	PRICKLE1	1	High grade glioma (GBM)	0.98
PRICKLE1	PRICKLE1	CHSY3/BC036251 intergenic region	1	High grade glioma (GBM)	/
PRICKLE1	KCNMB4	PRICKLE1	1	High grade glioma (GBM)	0.89
PRKAR1A	PRKAR1A	mir-108-1/FLJ43681 intergenic region	1	Hepatoblastoma	/
PTEN	PTEN	DM119543/PRMT2 intergenic region	1	Anaplastic Ependymoma III*	/
PTPRD	JAK2	PTPRD	1	Ependymoma	0.99
PTPR	PTPR	UG/LINC00266-1 intergenic region	1	Medulloblastoma	/
RASSF4	RASSF4	ZNF22	1	Embryonal rhabdomyosarcoma	0.86
RBMS3	RBMS3	LINC00616	1	Neuroblastoma	/
RBMS3	LUZP2	RBMS3	1	Neuroblastoma	0.06
RECK	SCML2	RECK	1	Medulloblastoma; LFS*; Williams-Beuren syndrome	0.7
RNF111	RNF111	ADAM10	1	Osteosarcoma	0.19
ROBO1	ULK4	ROBO1	1	Neuroblastoma	0.1
RPS6KA2	LINC00473	RPS6KA2	1	Fibrolamellar hepatocellular carcinoma	/
RTN4IP1	RTN4IP1	PCDH20/CHAMP1 intergenic region	1	Embryonal rhabdomyosarcoma; LFS*	/
RTN4IP1	PCDH9	RTN4IP1	1	Embryonal rhabdomyosarcoma; LFS*	0.002
RUNX1	ATP9B	RUNX1	1	Nephroblastoma	/
RUNX2	RUNX2	AK127888/PARD6G intergenic region	1	Osteosarcoma	/
SALL2	METTL3	SALL2	1	Ependymoma	0.73
SALL2	SALL2/LOC90925 intergenic region	RBBP4	1	Ependymoma (revised from low grade glioma)	/
SALL2	ANKH	SALL2/LOC90925 intergenic region	1	High grade glioma (GBM)	/
SDHA	SDHA	SDHAP3	1	Medulloblastoma	/
SDHA	SDHA	SDHAP2	1	High grade glioma (GBM)	/
SDHA	LOC728613	SDHA	2	Several patients	/
SEC14L2	AFF3	SEC14L2	1	Osteosarcoma	0.06
SOX7	KKR6	SOX7	1	High grade glioma (GBM)	/
SRGAP3	SRGAP3	CIAO1	1	Neuroblastoma	0.06
SSBP2	SSBP2	MIR1297/CHAMP1 intergenic region	2	Several patients	/
STAR13	PDSSB	STAR13	1	Medulloblastoma	0.29
TAGLN	SIDT2	TAGLN	21	Several patients	0.007
TCF3	TCF3	MIDN/MGC2752 intergenic region	1	Medulloblastoma; LFS*; Williams-Beuren syndrome	/
TCF7L2	TCF7L2	PPP6R3/AK130852 intergenic region	1	Ependymoma	/
TDGF1	TDGF1	TDGF1P3/DDX11L family member intergenic region	1	Anaplastic Ependymoma III*	/
THRA	IKZF3	THRA	1	Alveolar rhabdomyosarcoma	0.81
TP53	ZNF565	TP53	1	Osteosarcoma	0.96
TP53INP1	TP53INP1	ROBO2	1	Ependymoma (revised from low grade glioma)	/
TPTE2	TPTE2	ZNF770/DDX11L9 intergenic region	1	Mesenchymal chondrosarcoma	/
TRIM3	TRIM3	HPX	1	High grade glioma (GBM)	0.002
TSC22D1	TSC22D1	ACAP2	1	Embryonal rhabdomyosarcoma; LFS*	0.96
TSSC4	CD81	TSSC4	1	Ependymoma	0.006
UBE4B	UBE4B	CTNNBIP1	1	Osteosarcoma	0.9
UFL1	UFL1	TFB1M	1	Mesenchymal chondrosarcoma	0.004
UNC5D	GSR	UNC5D	1	High grade glioma (GBM)	0.88
VIM	OR51E2/AK130852 intergenic region	VIM	9	Several patients	/
VP53	VP53	NXN	1	Medulloblastoma	0.002
WT1	EW5R1	WT1	1	Alveolar rhabdomyosarcoma	0.43
XRCC5	XRCC5	FRMD4A	1	Alveolar rhabdomyosarcoma	0.41

Table S3: List of fusions involving tumor suppressor genes among the total non-redundant 1374 fusions detected by ChimComp. (source: tumor suppressor gene database, <http://bioinfo.mc.vanderbilt.edu/TSGene>)

Table S4: Fusions simultaneously detected by the 3 tools of ChimComp: Defuse, TopHat and Fusion Catcher referred to as “the high-confidence group”.

Genes names, patients associated, oncoscore, number of databases reporting the fusion, type of rearrangement, spanning reads mean representing the reads aligning to the junction sequence, pair reads mean representing discordant read pairs mapping to the two different genes of a fusion were provided by ChimComp. Diagnoses of patients were obtained from clinical data. The presence of oncogenes or tumor suppressor genes in the different fusion transcripts was sought in the following databases: ONGene by Liu et al. 2017 and TSGene (<http://bioinfo.mc.vanderbilt.edu/TSGene>) respectively. Fusion description, fusion breakpoints on each gene at the genomic level, predicted effect and identifier of the predicted fused transcripts were predicted by Fusion Catcher.

CML = chronic myeloid leukemia; T-ALL = T-cell acute lymphoblastic leukemia; ALL = acute lymphoblastic leukemia; AML = acute myeloid leukemia; PNET = primitive neuroectodermal tumor.

Fusion gene 1	Fusion gene 2	Pathology	Nb of tools	Name of tools	Predicted effect provided by "Fusion Catcher"	Domains gene 1	Domains gene 2	Ranking of the fusion regarding its oncoscore within the same patient	Spanning reads (mean) ¹	Predicted effect according to Fusion Catcher	Tumor entity known to be related to one or both genes
FES	MAN2A2	Neuroblastoma	2	Defuse / Fusion Catcher	CDS(truncated)/UTR	FBAR domain; SH2 domain; Protein kinase domain	/	/	5,5	CDS(truncated)/UTR	Acute promyelocytic leukemia; soft tissue sarcoma; Ewing sarcoma
CCM2L	HCK	Medulloblastoma	2	Defuse / Fusion Catcher	in-frame	CCM2 harmonin-homology domain	Serine-threonine/tyrosine-protein kinase catalytic domain; SH2 domain; Src homology-3 domain	1/59	3,5	in-frame	/
JAK2	PTPRD	Ependymoma	3	TopHat/ Defuse / Fusion Catcher	CDS(truncated)/UTR	Serine-threonine/tyrosine-protein kinase catalytic domain; SH2 domain; FERM domain	Tyrosine protein phosphatase domain	1/97	7,3	CDS(truncated)/UTR	Acute myeloid leukemia; non-meningothelial mesenchymal tumor
MANBAL	SRC	High grade glioma (GBM)	2	Defuse / Fusion Catcher	CDS(truncated)/UTR	Uncharacterized	Serine-threonine/tyrosine-protein kinase catalytic domain; SH2 domain; Src homology-3 domain	1/224	4,5	CDS(truncated)/UTR	Colorectal cancer; sarcoma; stomach adenocarcinoma
TYRO3	KLHL18	Anaplastic Ependymoma III*	2	Defuse / TopHat	Not found by fusion catcher	Serine-threonine/tyrosine-protein kinase catalytic domain; Immunoglobulin-like domain	Galactose oxidase; beta-propeller	3/70	3,5	Not found by fusion catcher	AML; CML; multiple myeloma; endometrial cancer; melanoma

Table S5: Characteristics of fusions involving tyrosine kinases genes among the total non-redundant 1374 fusions detected by ChimComp.

(source

<http://www.nature.com/onc/journal/v19/n49/full/1203957a.html>)

Caractérisation d'un nouvel oncogène retrouvé dans le sarcome d'Ewing: LM03-BORCS5

II. Cas de l'oncogène de fusion LMO3-BORCS5

A. Les protéines à domaine LIM de type LMO : description, fonction et rôle dans le développement tumoral

1. Généralités et fonctions

La sous-classe LMO (LIM-domain only) des protéines à domaine LIM est une famille de co-régulateurs nucléaires de la transcription, composée de quatre membres (LMO1-LMO4). Ils sont tous caractérisés par la présence exclusive de deux domaines LIM en tandem (LIM1 et LIM2) et aucun autre domaine fonctionnel (**Figure IV A-1 1**).

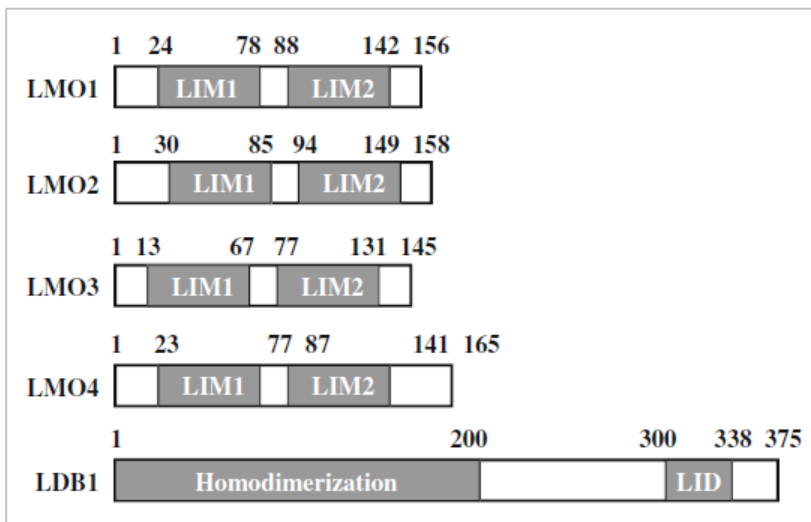


Figure IV A-1 1: Représentation schématique des protéines à domaines LIM humaines, comprenant les LMOs et LDB1. LID = LIM interaction domain. Source: Sang et. al, (142).

Le domaine LIM est un motif dit en « doigt de zinc », riche en cystéine, hautement conservé et retrouvé dans une grande variété de protéines nucléaires et cytoplasmiques. Ce domaine a été identifié à l'origine dans une autre famille de protéine à domaine LIM : les facteurs de transcriptions LIM-HD (LIM homedomain) (143). Ce domaine LIM ne peut se lier directement à l'ADN et agit comme un site d'accrochage pour l'assemblage de complexes multiprotéiques (**Figure IV A-1 2**). Deux mécanismes d'action des protéines LMO ont été reportés, le premier étant lié à la régulation de facteurs de transcription pour la régulation de l'expression de gènes cibles (**Figure IV A-1 2 panel A**) et le deuxième faisant appel aux protéines LIM-HD et LDB1 (**Figure IV A-1 2 panel B**). En effet, les deux domaines en tandem retrouvés chez les membres de la famille LIM (comprenant la famille LMO et LIM-HD) sont capables de se lier à la protéine LDB1 (LIM domain-binding protein 1), essentielle pour les processus de développement (143). La dimérisation de LDB1 permet ensuite la formation de complexes (tétramériques ou plus) avec les protéines LIM-HD et/ou LMO. Les protéines LMO peuvent également réguler l'activité transcriptionnelle des protéines LIM-HD *via* une compétition pour la liaison à LDB1 (**Figure IV A-1 2 panel B**).

Les protéines de la famille LMO fonctionnent en permettant la formation de nouveaux complexes transcriptionnels et/ou en perturbant des complexes transcriptionnels existants afin de moduler les programmes d'expression génique et d'agir sur différenciation cellulaire. Elles jouent de ce fait un rôle important dans les processus cellulaires responsable de l'oncogenèse comme la capacité d'auto-renouveaulement, la régulation du cycle cellulaire ou la formation de métastases (144).

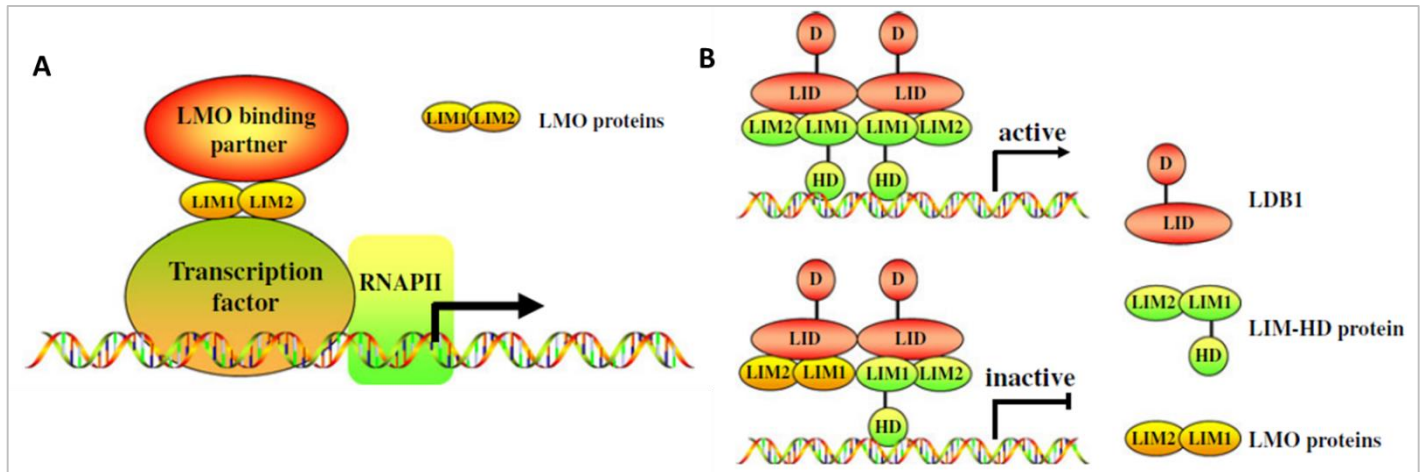


Figure IV A-1 2 : Mécanisme d'action des protéines à domaines LIM. A. Représentation schématique du rôle des protéines LMO dans la régulation transcriptionnelle de gènes cibles. Les protéines LMO régulent la transcription en agissant comme protéines de liaison entre des facteurs de transcriptions et leurs cofacteurs régulateurs. **B.** Mécanisme fonctionnel des protéines LMOs dans l'association avec LDB1 et les protéines LIM-HD. La dimérisation de LDB1 permet la formation de complexes multiprotéiques avec les protéines LIM-HD et/ou LMO qui fonctionnent ensemble pour déterminer le devenir cellulaire. RNAPII = RNA polymérase II; LIM-HD = LIM homedomain. Source: Sang et. al, (143).

2. LMO3

Le gène *LMO3* est situé sur le bras court du chromosome 12 (12p12.3) et sept variants d'épissage différents ont été identifiés pour ce gène. Des membres de la famille LMO, il est celui dont la fonction et le mode d'action moléculaire ont été le moins décrits. Son expression est tissu-spécifique puisqu'il est majoritairement exprimé au niveau cérébral. Une fois exprimé, la protéine est localisée de manière ubiquitaire dans la cellule, avec une prédominance dans le noyau (<http://www.genecards.com>). *LMO3* agit comme un régulateur critique de la neurogenèse et du caractère « souche » des cellules neurales en contrôlant l'activité des facteurs de transcription neuronaux (145).

L'activité oncogénique de *LMO3* a principalement été décrite dans le neuroblastome, où son expression est corrélée à une amplification du gène *MYCN* et associée à un mauvais pronostic chez les patients (146,147). Dans cette pathologie, les mêmes études ont également décrit la capacité

de LMO3 à former un complexe avec HEN2, un facteur de transcription neuronal et d'activer la transcription de Mash1 provoquant une induction de la prolifération cellulaire (146,147). Une étude menée par Hui *et. al*, a montré que LMO3 était capable d'interagir directement via son domaine LIM2 avec la protéine à motif « EF-hand » CIB1 (« calcium- and integrin-binding protein » aussi appelée calmyrine et KIP1) pour inhiber la prolifération dans des astrocytes (148). Larsen *et al.* ont également démontré la capacité de LMO3 à inhiber P53 par interaction directe, conduisant à une inhibition de la transcriptions des gènes pro-apoptotiques (149).

Les propriétés oncogéniques de LMO3 ont également été démontrées dans d'autre types tumoraux, comme l'adénocarcinome pulmonaire où il est associé à une amplification de l'oncogène NKX2-1 (150) ou plus récemment dans le cancer du pancréas où sa surexpression conduit à une activation des voies de signalisation d'AKT et par conséquent une augmentation de la prolifération et de la migration cellulaires (151).

3. Autres membres de la famille LMO

LMO1 et LMO2 ont initialement été identifiés comme des oncogènes associés à des translocations chromosomiques chez les patients avec des LAL-T (152). LMO1 a été initialement identifié comme étant associé à la translocation t(11;14) (p15;q11) (153,154). LMO2 est associé à plusieurs remaniements chromosomiques : les translocations t(11;14)(p13;q11), retrouvée dans 8-10% des cas, et t(7;11)(q35;p13) ainsi que la délétion del(11)(p12p13), majoritaire dans les LAL-T pédiatriques (143,155,156). Cependant, le mécanisme d'activation de LMO2 après translocation t(11;14) (p13;q11) dans les LAL-T est aujourd'hui encore peu compris. Il est intéressant de noter le rôle majeur de LMO1 et 2 dans cette pathologie, puisque même en l'absence de réarrangements chromosomiques impliquant ces gènes, leur surexpression est retrouvée dans ~50% des patients souffrant de LAL-T (144). Indépendamment des translocations, l'insertion rétrovirale dans la région promotrice de *LMO2* chez des patients atteints de déficit immunitaire combiné sévère (DICS) traités par thérapie génique a aussi provoqué la survenue de LAL-T (157).

LMO2 représente également un biomarqueur prédictif de la survie des patients atteints de lymphome diffus à grandes cellules B (143). Il a également été montré la capacité de LMO2 à se fixer sur le complexe protéique lié à l'ADN incluant le facteur GATA-1 dans les érythrocytes afin de contrôler leur différenciation (158). Terano *et. al* ont montré par ailleurs que la mutation du domaine LIM2 de LMO2, mais pas de LIM1, provoquait une inhibition de la fixation sur le facteur GATA-1 et ainsi une inhibition de l'érythropoïèse de progéniteurs foetaux *in vivo* (159).

Han *et. al* ont d'autre part suggéré le rôle de LMO2 dans la régulation de facteurs de transcriptions responsables du développement neural (143). Enfin LMO1, tout comme LMO3, jouent un rôle oncogénique dans un sous-groupe de patients atteints de neuroblastomes (160).

Enfin, LMO4 est le dernier membre de la famille identifié et a été décrit comme un auto-antigène dans le cancer du sein (143,161,162). LMO4 est très largement exprimé tout au long du développement embryonnaire et dans les tissus à l'âge adulte, contrairement aux profils d'expressions de LMO1, LMO2 et LMO3 qui sont restreints à certains tissus (143). Au niveau protéique, LMO4 ne partage que 50% d'homologie avec les domaines LIM des autres membres de la famille LMO. Pour ces raisons, LMO4 est considéré comme le membre le plus éloigné de la famille LMO (161). LMO4 est surexprimé dans plus de 50% des tumeurs mammaires primaires, et associé à une diminution de la différenciation des cellules épithéliales mammaires (143,163), une augmentation de la prolifération cellulaire et une inhibition de l'apoptose (164). LMO4 est notamment capable de réguler la voie de signalisation de ErbB-2 (165) ou encore d'agir comme un inhibiteur de BRCA1 (166), renforçant son caractère oncogénique.

Le rôle de LMO4 au sein du système nerveux et dans l'embryogenèse a également été décrit puisqu'il agit comme un inhibiteur direct de HEN1 (167) et est essentiel pour la survie, la formation du tube neural et la formation des os chez la souris (143,168,169). Des études récentes ont également montré sa fonction critique dans le développement normal et tumoral du tissu pancréatique (170,171).

Les protéines LMO jouent un rôle important dans l'initiation et/ou la progression de différents types de cancer, ainsi que dans la différenciation et le caractère souche des cellules où elles sont exprimées. La **figure IV A-3** récapitule l'impact des protéines LMO sur les différents cancers et processus cellulaires. Les LMO représentent donc des cibles thérapeutiques potentielles pour le traitement de différentes tumeurs. Actuellement, deux stratégies ciblant LMO2 ont été rapportées, utilisant soit un anticorps (172) soit un peptide inhibiteur (173) et confortent l'idée que l'inhibition de LMO2 pourrait être une stratégie efficace pour traiter ces types de LAL-T. Des inhibiteurs de LMO4 pourraient également être intéressants pour le traitement des cancers du sein de type basal, pour lesquels il existe actuellement peu de thérapies ciblées (143).

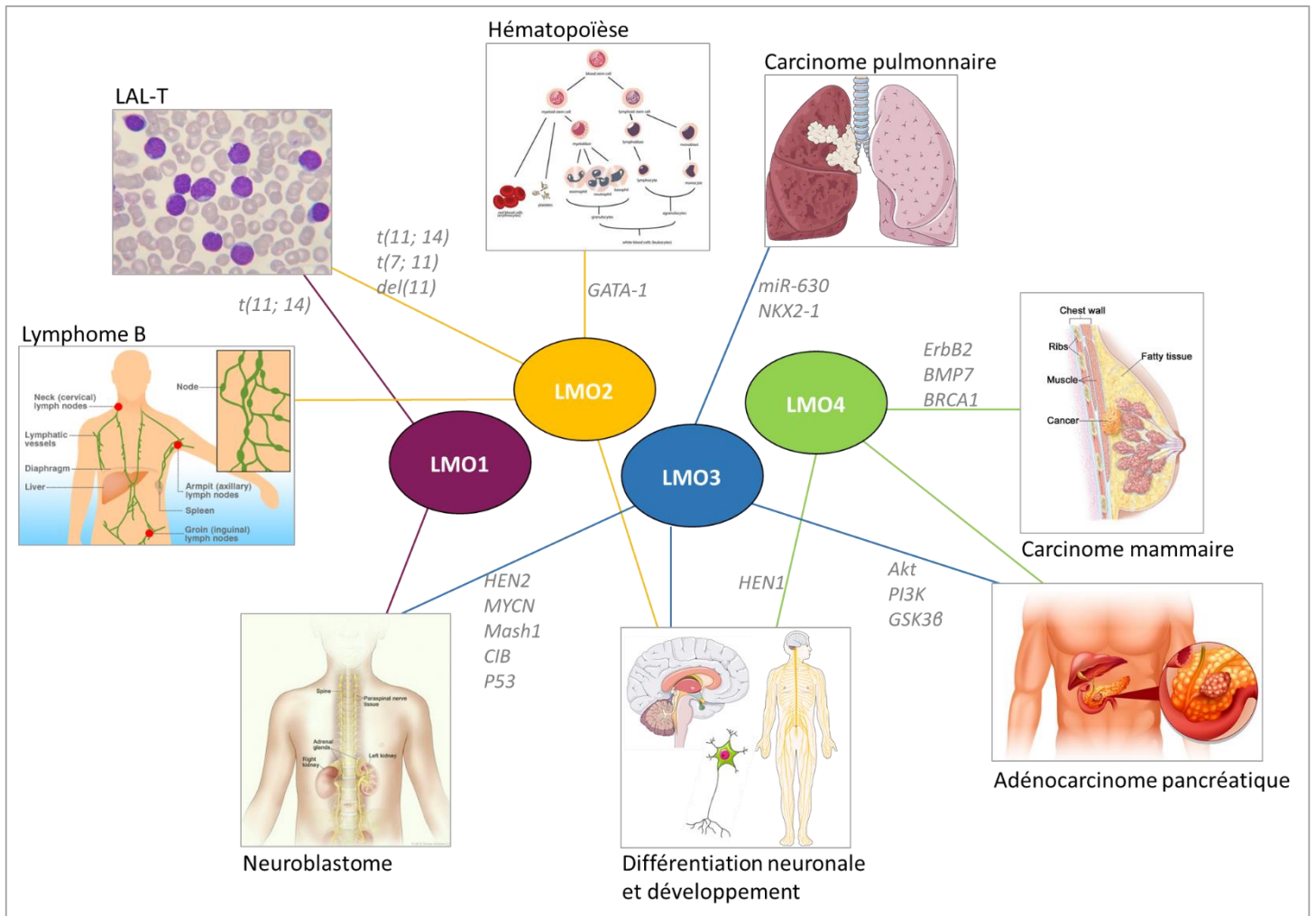


Figure IV A-3 : Résumé des fonctions cellulaires et impacts sur différents cancers des protéines LMO et exemples de voies de signalisation associées identifiées.

B. Structure et fonctions connues du complexe BORC

BORC (BLOC-1 Related Complex) est beaucoup moins connu et étudié que les protéines de la famille LMO. Les connaissances actuelles montrent que BORC est un complexe protéique composé de 8 sous-unités impliqué dans la régulation du transport lysosomal (174–177). Les différentes sous-unités consistent en trois sous-unités partagées avec le complexe BLOC-1 impliqué dans la biogenèse des organelles apparentées aux lysosomes (BLOS1, BLOS2 et la snapine), de la sous unité KXD1 et d'autres sous unités moins caractérisées : BORCS5/LOH12CR1/myrlyisine, C17orf59/lysperine, C10orf32/diaskedine et MEF2BNB (174). Au niveau moléculaire, BORC s'associe avec la membrane lysosomale, où il recrute la petite GTPase Arl8 (175). S'initie alors une chaîne d'interactions, notamment avec les kinésines, entraînant le déplacement des lysosomes vers le long des microtubules vers la périphérie de la cellule (**Figure IV B-1**).

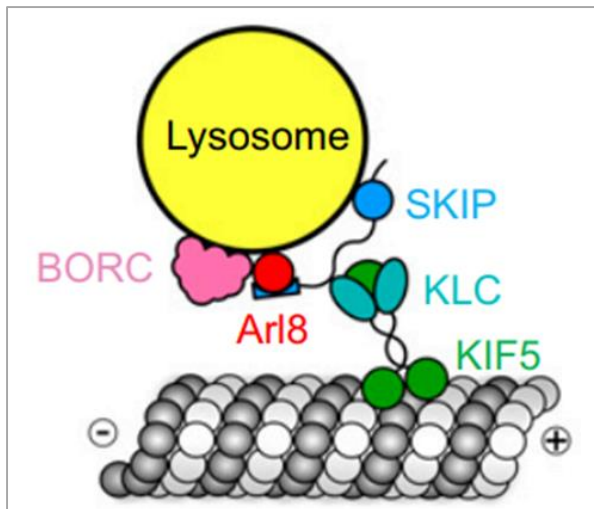


Figure IV B-1 : Représentation schématique du complexe BORG-Arl8-SKIP-kinésine-1 responsable du déplacement des lysosomes vers l'extrémité plus des microtubules. Source Farías et. al (175)

Pu et. al ont également montré l'implication de BORG dans la migration cellulaire soulignant l'importance du transport dépendant de BORG pour les autres fonctions des lysosomes. En effet, ces organelles membranaires ont pour fonction principale la dégradation des macromolécules après endocytose, phagocytose, autophagie ou transport de biosynthèse (178). Néanmoins les lysosomes participent aussi à divers autres processus cellulaires, comme la destruction de micro-organismes, la présentation d'antigènes, la détoxification, l'homéostasie du cholestérol, l'apoptose, les voies de signalisation métaboliques, la libération d'exosomes, la réparation des membranes plasmiques, la migration cellulaire et la formation de métastases (178).

De plus, d'autres études ont aussi démontré le rôle essentiel du complexe BORG dans le transport des vésicules synaptiques le long des axones neuronaux, à l'origine de processus axonaux critiques comme le maintien de la croissance et le renouvellement de l'autophagosome (176,179).

La sous-unité BORCS5 (BLOC-1 Related Complex Subunit 5), aussi appelée LOH12CR1 ou myrlysine, est codée par un gène situé sur le bras court du chromosome 12 (12p13.2). Aucun lien n'a été établi entre BORCS5 et cancer, cependant ce gène est situé dans un locus comprenant plusieurs gènes suppresseurs de tumeurs et dont la perte d'hétérozygotie a été fréquemment observée dans divers néoplasmes (180).

RESULTAT II : Découverte d'un nouveau transcrit de fusion oncogénique LMO3-BORCS5 et caractérisation fonctionnelle

A. Introduction de l'article

Comme exposé précédemment, les fusions sont considérées comme l'un des principaux mécanismes responsables de la cancérogenèse et constituent des outils précieux en clinique, aussi bien comme biomarqueurs que cibles pharmacologiques. Nous avons identifié un nouveau transcrit de fusion *LMO3-BORCS5*, présent sous deux isoformes, dans les biopsies tumorales au diagnostic et à la rechute d'un patient atteint de sarcome d'Ewing, ainsi que dans diverses lignées cellulaires. Après reconstitution de *LMO3-BORCS5*, ce transcrit est formé par une délétion de 4,2 MB sur le bras court du chromosome 12 entraînant la perte du domaine fonctionnel LIM2 de *LMO3* et un décalage du cadre de lecture du gène *BORCS5*. A partir de ces observations, notre objectif était donc de mieux **comprendre le rôle et la fonction de LMO3-BORCS5 dans le développement et/ou la résistance des tumeurs ainsi que d'évaluer** l'impact de la perte de *BORCS5*. Nous avons montré que l'isoforme le plus long de *LMO3-BORCS5* (i) a un fort potentiel oncogénique avec un impact sur la prolifération et la tumorigénicité, (ii) diminue l'expression des gènes liés à l'apoptose et la sensibilité à différents traitements anti-néoplasiques, (iii) conduit à l'inhibition de l'expression de gènes impliqués dans les processus de différenciation et à l'augmentation de l'expression de gènes impliqués dans la prolifération cellulaire et l'interaction avec la matrice extracellulaire.

De plus, nous avons démontré que *BORCS5* agit, *in vivo*, comme un suppresseur de tumeur. En conclusion, le nouveau transcrit de fusion oncogénique *LMO3-BORCS5* présente un rôle majeur dans la tumorigénèse et la sensibilité au traitement. C'est la première fois qu'une fusion découverte chez un patient unique à la rechute est également détectée dans plusieurs lignées cellulaires et présente un impact à la fois sur la tumorigénèse mais également la sensibilité au traitement. La détection et les études fonctionnelles des oncogènes de fusion sont donc importantes pour définir leur potentiel thérapeutique dans une approche de médecine personnalisée.

B. Article 2: "Functional characterization of a new oncogenic fusion transcript LMO3-BORCS5"

Functional characterization of a new oncogenic fusion transcript *LMO3-BORCS5*

Célia Dupain^{1,2,3}, Céline Gracia^{1,2,3}, Julie Rivière^{1,4}, Giorgia Urbinati^{1,2,3}, Karim Benihoud^{1,2,3}, Birgit Geoger^{1,2,3,5}, and Liliane Massaad-Massade^{1,2,3*}

¹ Université Paris-Sud 11, Laboratoire de Vectorologie et Thérapeutiques Anticancéreuses, UMR 8203, Villejuif, France-94805

² CNRS, Villejuif, Laboratoire de Vectorologie et Thérapeutiques Anticancéreuses, UMR 8203, France-94805.

³ Gustave Roussy, Laboratoire de Vectorologie et Thérapeutiques Anticancéreuses, UMR 8203, Villejuif, France-94805.

⁴ INSERM U1170, Gustave Roussy Cancer Center, Villejuif, France

⁵ Gustave Roussy, Department of Pediatric and Adolescent Oncology, Villejuif, France

*: Correspondance to Liliane MASSADE, PhD, Laboratoire de Vectorologie et Thérapeutiques Anticancéreuses, UMR 8203, 114 rue Edouard Vaillant, Villejuif, France-94805

Running title: Role of a new fusion transcript *LMO3-BORCS5* in tumorigenesis.

Key words: fusion transcripts, pediatric solid cancers, personalized medicine, relapsed patients.

Abstract

Fusion transcripts are considered as one of the leading mechanisms responsible for carcinogenesis and represent valuable tools in clinics as biomarkers or druggable targets. We identified a new fusion transcript *LMO3-BORCS5* in a patient with Ewing sarcoma at diagnosis, at relapse, as well as in various cell lines. Reconstitution of *LMO3-BORCS5* exhibited a 4.2 MB deletion within the short arm of chromosome 12 leading to the loss of the functional domain LIM2 in the *LMO3* gene and the disruption of the reading frame of *BORCS5* gene. Consequently, our aim was to investigate the functional role of *LMO3-BORCS5* transcript and the impact of *BORCS5* loss. We showed that *LMO3-BORCS5* (i) has a high oncogenic potential with impact on proliferation and tumorigenicity, (ii) decreases expression of apoptosis-related genes and treatments sensitivity and (iii) downregulates genes involved in differentiation and upregulates proliferative and extracellular matrix-related pathways. Additionally, we proved that *BORCS5* acts, *in vivo*, as a tumor suppressor gene. In conclusion, the newly discovered fusion oncogene *LMO3-BORCS5* in a unique patient and in several cell lines exhibits a major role in tumorigenesis and treatment sensitivity. Therefore, detection and functional studies of fusion oncogenes is of importance to define their therapeutic potential within a personalized treatment approach.

Introduction

Fusion transcripts are responsible for 20% of global cancer morbidity and have been recognized as major cancer drivers (1). They result from the juxtaposition of two genes, usually separated, due to chromosomal rearrangements or non-chromosomal events (2). They can affect coding or regulatory sequences which can lead to (i) the creation of a chimeric oncoprotein, (ii) an aberrant overexpression of the second partner involved in the fusion or (iii) truncation of a tumor-suppressor gene (TSG) (2).

Fusions can serve as diagnostic, prognostic, predictive biomarkers, as well as therapeutic targets (2–10). Apart from the recurrent known fusion genes, the majority of newly discovered gene fusions vary extremely in their expression and their recurrence is much lower than other somatic mutations (11). Even though it is commonly thought that recurrence represents the first criteria to consider an alteration as driver, the impact of an alteration is determined by patient and tumor's history (e.g. genetic background, tumor type or treatments received) and one mutation could have different consequences on drugs sensitivity depending on cancer types (12). Hence, non-recurrent fusions could have a high clinical value for the individual patient. In addition, most of studies have been performed on patients at diagnosis while the analysis of tumors at relapse remains one of the challenges to enrich the repertoire of evolving driver alterations, including fusions (13).

Previously, we conducted a retrospective study on the RNA-sequencing data of pediatric resistant or relapsing patients included in the molecular profiling trial MOSCATO-01 conducted at Gustave Roussy (14,15). Interestingly, we detected in the tumor of a patient with Ewing sarcoma (EWS) a new fusion transcript, LMO3-BORCS5 at diagnosis and at relapse (Dupain et al., paper under revision).

LMO3-BORCS5 results from a deletion of 4.2 megabases on chromosome 12 where both LMO3 and BORCS5 are located, respectively 12p12.3 and 12p13.2. LIM-domain-only proteins (LMO) are characterized by the presence of two functional domains: LIM1 and LIM2. The family of proteins is composed of four members (LMO1-4) mediating protein-protein interactions between transcription factors and other proteins, representing important transcription co-regulators acting on both tumorigenesis and differentiation (16).

The oncogenic activity of LMO3 has been described in neuroblastoma where its expression is correlated with amplified MYCN and activation of HEN2, leading to poor prognosis in patients (17). Larsen et al. demonstrated the ability of LMO3 to directly inhibit TP53, strengthening its oncogenic status (18). LMO3 was also found in lung carcinoma and gastric cancer (17,19,20). LMO3 is highly expressed in brain and acts as a regulator of neurogenesis and stemness by controlling neural-specific transcription factors, thus being critical for cell differentiation (21).

BORCS5 (BLOC-1 Related Complex Subunit 5) also named LOH12CR1, belongs to the BORC complex involved in the regulation of cell migration and lysosomal trafficking as well as in axonal transport of synaptic vesicles (22–24). It is situated on a locus containing several TSGs for which loss of heterozygosity has frequently been observed in various neoplasms (25). Yet, no study has described the different roles of BORC subunits or the possible role of BORCS5 loss on tumorigenicity.

The use of fusions as personalized medicine tools in clinics is often limited due to the absence of functional studies (26,27). Therefore, the aim of our study was to understand the role and function of LMO3-BORCS5 in the development and/or resistance of tumors and to evaluate the effects of BORCS5 knockout. Taking together, our results suggest that LMO3-BORCS5 plays an essential role in tumorigenesis and treatment sensitivity and could be a potential therapeutic target. However, BORCS5 appeared to be a double sword edge, with pro-oncogenic effects on proliferation and treatment sensitivity *in vitro* but acting as a tumor suppressor gene *in vivo* by inhibiting tumor invasiveness.

Material and methods

Patient history, detection and sequencing of LMO3-BORCS5 fusion transcript

LMO3-BORCS5 was identified together with *EWS-FLI1* after *in silico* RNA-seq analysis of a 20 years old male patient, included in the pediatric MOSCATO-01 cohort that only comprised two EWS (Dupain *et al.*, article under revision). He had relapsed 3.4 years after initial diagnosis treated with chemotherapy (vincristine, ifosfamide, doxorubicin, etoposide, dactinomycin, cyclophosphamide). The presence of *LMO3-BORCS5* fusion transcript was confirmed by RT-qPCR in tumor samples at diagnosis and relapse. The products were sequenced by Sanger technique (Eurofins, Cochin Institute, Paris, France).

Cell lines and cell culture.

To assess the presence of *LMO3-BORCS5* fusion transcript, 31 cell lines from various histological origins were explored including 28 pediatric cancer cell lines of the Innovative Therapies for Children with Cancer in Europe (ITCC) consortium cell line panel (**Table S1**) (28). For functional studies, A673 Ewing sarcoma cell line and the mouse embryonic fibroblast cells NIH-3T3 were used. DMEM or RPMI medium were supplemented with penicillin (100 U/ml), streptomycin (10 µg/ml) and 10% fetal calf serum (FCS) (Life Technologies™, Saint Aubin, France) for all cell lines or 10% new born calf serum (NBCS) for NIH-3T3. Cell cultures were maintained at 37°C in an atmosphere of 5% CO₂ and 95% humidity and were systematically tested by PCR analysis to be free of mycoplasma.

Construction and establishment of cell lines stably expressing LMO3-BORCS5

The sequences of *LMO3-BORCS5 v1* or *LMO3-BORCS5 v2* were obtained after reverse transcription of the patient's tumor RNA into cDNA followed by PCR amplification using the following primers: Forward: 5'-GCAACCGAAAGATCAAGG-3' and Reverse: 5'-CATTTGTTGGGGAAGTCTG-3'. After purification of the amplicons, the sequences were cloned under the control of the CMV promoter into the pcDNA3.1 expression vector. Two plasmids were obtained: pcDNA3.1-*LMO3-BORCS5v1* and pcDNA3.1-*LMO3-BORCS5 v2*.

Then, A673 and NIH-3T3 cells were transfected with empty vector (pcDNA3.1) or with either pcDNA3.1-*LMO3-BORCS5 v1* or pcDNA3.1-*LMO3-BORCS5 v2* to obtain stable clones as previously described and selected with neomycine (29).

CRISPR/Cas9 knock-out of BORCS5

BORCS5 gene was inactivated using the CRISPR/Cas9 system targeting the last exon containing the functional domain (30). Using CRISPOR software v4.2 (31), two couples of 20-base pairs

targeting sequences were designed (guide 1: 5'-AGAGAGACCGCCATCGTGTT-3' and guide 2: 5'-CTGGATCTGCTCGGCATACT-3') and synthesized (Eurofins Genomics, Ebersberg, Germany). Both guide 1 and 2 were respectively inserted into pLentiCRISPR V2 (Addgene 52961) containing either GFP or CFP respectively instead of Puromycine [pLentiCRISPR V2-GFP and pLentiCRISPR V2-CFP, provided by Julie Rivière (INSERM U1170, Gustave Roussy)].

A673 cells were co-transfected with both plasmids. After 27h, GFP and CFP-positive cells were selected using FACSaria™ Fusion cell sorter (BD Biosciences, Le Pont de Claix, France) and reseeded at a concentration of one cell/well to allow single colony formation. After 10 days, genomic DNA was extracted from individual colonies using QuickExtract™ DNA Extraction solution (Lucigen, Middleton, USA). Deletion of the target sequence was tested by PCR using the following primers: Forward 5'-GAGCTGTTTTCTCCTGGCCT-3' and Reverse 5'-TTCATGCTGAAGGGCTCCAG-3'.

Reverse transcription-quantitative PCR (RT-qPCR)

RNA extraction, reverse transcription and real-time PCR (qPCR) were performed as previously described (29). Primers were either purchased from Eurofins Genomics (Ebersberg, Germany) enlisted in **Table S2**, or included in a personalized Human qPCR SignArrays®(32) from AnyGenes (Paris, France).

RT-qPCR was performed for *LMO3-BORCS5* and the *LMO3* and *BORCS5* wild-type genes in: (i) patient's tumor sample at diagnosis and relapse, (ii) the 31 various cell lines and (iii) A673 or NIH-3T3 stably transfected with pcDNA3.1 (empty vector) or pcDNA3.1-*LMO3-BORCS5* v1 (referred to as "+*LMO3-BORCS5*"). For *LMO3-BORCS5*, the amplicon size was verified by electrophoresis in 2% agarose gel. PCR products were purified with DNA extraction with NucleoSpin® Gel and PCR clean-up extraction kit (Macherey-Nagel, Dueren, Germany) and sequenced by Sanger technique (Eurofins, Cochin Institute, Paris, France).

Genes involved in apoptosis *BAX*; *NOXA*; *P21* and *TP53* in A673 and those sorted out from microarray analysis for validation (*ABCC3*; *FGF1*; *IGF2*; *PAX7*; *PRKCA*; *WNT6*; *ALK*; *CUX2*; *FGF13*; *FRK*; *LZTS1*; *MMP10*; *MYOD1*; *NRP1*; *NTRK1*; *RHOJ* and *SEMA3C*) were also assessed.

For all experiments, gene expression was determined by quantification-comparative $2^{-\Delta\Delta C_t}$ method (33) and normalized to *GAPDH* levels. Results are expressed as relative mRNA levels and represent at least three independent experiments.

Doubling time and cell cycle analysis

Cell cycle analysis by flow cytometry and doubling time of the different clones were assessed as previously described (29).

Cell migration assays

Scratch test was performed to evaluate cell migration capacity as previously described (29). Results are presented as wound closure time in hour per cell line.

Cell survival assays

Cell line A673 was seeded in 96-well plates at 25000 cells/well. After 72h incubation with increased concentrations from 0 to 10 μ M of vincristine, cyclophosphamide, mafosfamide, etoposide, actinomycin, dasatinib, SN-38, doxorubicin, bortezomib, pazopanib or everolimus, MTT [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide] test was performed as previously described (34). Results are the mean \pm SD of at least two independent experiments containing 3 replicates for each condition.

Annexin V apoptosis assay

Apoptosis was determined by flow cytometry using Annexin V apoptosis detection kit (BD Biosciences, Le Pont de Claix, France) according to the manufacturer's instructions. Each cell line was seeded in 6-well plates (3x10⁵ cells/well) and treated with vincristine (10nM) or SN-38 (20 μ M). After 48h, cells were collected and analyzed as previously described (29). Experiments were performed in triplicate and data represent the percentage of apoptotic cells compared to non-treated cells.

Microarray analysis

The comparison of gene expressions between the two clones (C1 and C2) obtained from A673+empty vector and A673+LMO3-BORCS5 was assessed by microarray analysis as described previously (34) (supplementary Material and Methods section). Four independent RNA extractions were performed for each clone using RNeasy mini-kit (Qiagen, Courtaboeuf, France) following manufacturer's instructions. We defined up- or down-regulation as ratios greater than two-fold between A673+empty vector and A673+LMO3-BORCS5, and an adjusted p-value (FDR) < 0.05. To interpret the biological meaning genomic data, we used the ToppGene Suite (<https://toppgene.cchmc.org/enrichment.jsp>) and Ingenuity (www.ingenuity.com) software.

Immunoblotting

Total proteins were extracted from A673 WT and A673 KO BORCS5 cell lines as previously described (29). The following primary antibodies were used: monoclonal rabbit BORCS5 (1:1000, PAS-5821, Invitrogen, Carlsbad, CA, USA) and GAPDH-HRP as internal control (1:1,000; Sigma-Aldrich Chemicals Co., Saint Quentin Fallavier, France). Blots were washed and incubated with corresponding secondary antibodies: anti-rabbit-HRP (1:3,000; Cell Signaling technology, Saint

Quentin en Yvelines, France). Bands were revealed by enhanced chemiluminescence reagent (HRP system, Clarity™; Bio-Rad Laboratories, Marnes-la-Coquette, France) using ChemiDoc™ Imaging System (BioRad, Hercules, CA, USA).

Animal studies and tumorigenicity tests

All animal experiments were carried out according to French laws (Directive 2010/63/UE) and approved by the institutional Ethics Committee of Animal Experimentation (CEEA): authorization number CEEA IRCIV/IGR no. 26: 94-226, no. 2011-09.

Six-weeks-old nude nu/nu female mice were used to test the tumorigenicity of: (i) A673 empty vector *versus* A673+*LMO3-BORCS5* v1; (ii) NIH-3T3+empty vector *versus* NIH-3T3+*LMO3-BORCS5* v1 and (iii) A673 WT *versus* A673 KO BORCS5.

All cell lines were diluted in 100 µl PBS and injected subcutaneously into the flank of mice (n=5/group) at a rate of 10×10^6 cells/mouse. Mice were monitored for tumor growth and sacrificed when tumors reached a volume of ~ 1000 mm³. Half of each tumor was immediately frozen in liquid nitrogen for RT-qPCR, the other half was conserved in FineFix Formalin-Free Fixative (Fisher Scientific, Illkirch, France) for histochemistry analysis. For this latter analysis, tumor tissues were embedded in paraffin, sections (4 µm thick) were prepared and stained with hematoxylin-eosin-safranin (HES). Sections were examined with Zeiss-Axiophot microscope (Microscopy and Imaging center, Texas, USA).

Statistical analysis.

Data are presented as mean \pm SD. By using GraphPad Prism 4 software, Mann-Whitney test was used to compare two groups of treatments and Kruskal-Wallis test followed by Dunn's test for multiple treatments. Doubling time was calculated according to an exponential regression equation, given as: $y = axe^{bx}$, *via* the formula: $t_{1/2} = \ln(2)/b$. $p < 0.05$ was considered as statistically significant. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Results

LMO3-BORCS5 is detected in a patient with EWS and in various cell lines

By RT-qPCR, the *LMO3-BORCS5* fusion transcript in the EWS tumor samples at relapse showed two variant forms (*LMO3-BORCS5 v1* and *LMO3-BORCS5 v2*) while only one form was detected at diagnosis (*LMO3-BORCS5 v1*). *LMO3-BORCS5* was higher expressed at relapse compared to diagnosis (Ct: 22.6 vs 39.7) and co-expressed with *EWS-FLI1* (**Figure 1Aa**).

The *LMO3-BORCS5* fusion was also detected in 13 out of 31 (41%) cancer cell lines assessed (**Figure 1Ab**). As for the EWS tumor, the fusion transcript was present in the two variant forms (**Figure 1Bb**). However, *LMO3-BORCS5* expression was low (Ct \approx 30) as compared to the recurrent tumor, impeding the use of these cell lines as models for further studies.

The reconstitution of the sequence was performed after “Sanger” sequencing and showed the same result for EWS patient and for cell lines. *LMO3-BORCS5 v1* resulted in the breakpoint on *LMO3* at the end of exon 3 and, in the middle of the same exon for *LMO3-BORCS5 v2* (**Figure 1B**). In both cases, the breakpoint in *BORCS5* remains identical and is situated at the beginning of exon 3 (**Figure 1B**).

At the protein level on the N-terminal part, the variant 1 leads to a protein with complete conservation of LIM1 and loss of LIM2 domains of *LMO3*, while 24 amino-acids at the end of LIM1 domain are lost in variant 2 (**Figure 1C**). In both cases on the C-terminal part, the fusion transcripts lead to a frameshift in the *BORCS5* sequence giving a putative truncated out-of-frame *BORCS5* protein with the creation of 43 amino-acids associated with unknown protein domain/function (**Figure 1C**, in dark red).

Differences in LMO3 and BORCS5 gene expressions are observed among cell lines

The relative mRNA expression of *LMO3* and *BORCS5* wild-type has been assessed by RT-qPCR, by using primers designed within regions not retained in the fusion, among 31 cell lines previously tested for *LMO3-BORCS5* fusion. The expression was extremely variable among the different cell lines: 15/31 almost did not express *LMO3* gene while the neuroblastoma cell lines IGR-NB8, IGR-N91 and SH-SY5Y and the EWS cell lines STA-ET-1, TC71 and A673 presented the highest expression of *LMO3* (**Figure 2A**). Regarding expression of *BORCS5* wild-type, all 31 cell lines tested were positive with the highest expression found in the medulloblastoma cell line D293MED, followed by EWS (EW7 and SK-N-MC) and osteosarcoma (HOS-143B, IOR/OS18 and SAOS2) cell lines (**Figure 2B**).

Mutations and fusions in LMO3 and BORCS5 genes are found in diverse cancer types

To assess the presence of *LMO3* and *BORCS5* alterations according to cancer types, large-scale

cancer genomics data were investigated among samples included in 215 different studies (<http://www.cbioportal.org>). *LMO3* was altered on the genomic level in 1.5% of patients. These cases comprised 82% of missense mutations, 11% of mutations leading to truncation, 5% of fusions and 2% of in-frame mutations (**Figure S1A**). Five fusions were identified harboring *LMO3* and are detailed in **Table S3**. The function of *LMO3*-fused genes differed according to their position in the fusion. When *LMO3* is situated in 5', the second gene appears to systematically be related to cancer progression or involved in cell proliferation or migration. When situated in 3', *LMO3* fuses with genes related to metabolism (*B4GALNT3*, *FAM135A* and *AEBP2*) harboring high promoter activity (**Table S3, upper panel**).

BORCS5 gene was altered in 1.5% of cases. They included 81.7% of missense mutations, 11% of fusions, 6.1% of truncating mutations and 1.2% of in-frame mutations (**Figure S1B**). Nine fusions were detected containing *BORCS5* (**Table S3, lower panel**) and the function related to the partners fused with *BORCS5* were variable with genes of transportation regulation, membrane trafficking, cell adhesion or oncogenes (**Table S3, lower panel**).

An overview of *LMO3* and *BORCS5* genes alterations according to cancer types can be visualized on <http://www.cbioportal.org>. *LMO3* appeared to be mainly affected by amplification among the different cancer types, as well as *BORCS5* (**figures S1C and S1D**). We noticed that for *BORCS5*, certain cancers exclusively harbor mutations (colorectal adenocarcinoma) or deletions (non-Hodgkin lymphoma and leukemia) or amplifications (cutaneous melanoma) (**Figure S1D**).

The different clones derived from the same stable transfection showed the same profiles

For the human EWS cell line A673, two clones (C1 and C2) were obtained containing empty vector (**Figure S2 well 1**), two clones (C1 and C2) containing *LMO3-BORCS5* v1 (**Figure S2A, well 2**) and two clones (C1 and C2) containing *LMO3-BORCS5* v2 (**Figure S2A, well 3**). Then, we investigated: i) the proliferation rate of the two clones (C1 and C2) derived from A673, ii) their doubling time and the repartition of cells within the cell cycle as well as iii) their transcriptome profile. As shown in **Figure S2**, each two-to-two clones derived from the same transfection showed similar doubling time (**Figure S2B**), and almost the same cell-cycle repartition (**Figure S2C**). Moreover, similar transcriptome profiles were observed between both C1 and C2 clones derived from A673 cells stably transfected with *LMO3-BORCS5* v1 (**Figure S3**). Thus, we focused on the clones "C1" for further studies.

Concerning A673 transfected with *LMO3-BORCS5* v2, no significant difference in doubling time (**Figure S2B**) and cell cycle repartition was observed between A673+empty vector and A673+*LMO3-BORCS5* v2 (C1 or C2) (**Figure S2C**), therefore A673+*LMO3-BORCS5* variant 2 function was not more deeply investigated.

For the non-tumorigenic mouse fibroblast cell line NIH-3T3, one clone containing empty vector

and one clone containing pcDNA3.1-*LMO3-BORCS5* v1 were obtained (**Figure S2 wells 4 and 5 respectively**). One clone derived from NIH-3T3 containing pcDNA3.1-*LMO3-BORCS5* v2 was obtained (**Figure S2 well 6**), however the cells did not expand *in vitro*, precluding further studies.

LMO3-BORCS5 increases cell proliferation in vitro and tumorigenicity in vivo in both A673 and NIH-3T3 cells

The introduction of *LMO3-BORCS5* in both A673 and NIH-3T3 cell lines led to an increase of proliferation, reflected by a decrease in doubling time of about 20% in cells expressing the fusion compared to cells expressing the empty vector (**Figure 3A and B**). In A673 this was accompanied by a significant decrease of cells in G0/G1 phase and an increase in G2/M (**Figure 3C**), while this was not observed in NIH-3T3 (**Figure 3D**). No effect of *LMO3-BORCS5* expression on cell migration in both A673 and NIH-3T3 was detected (not shown).

After subcutaneous injection of A673 clones, an increase of tumor growth was observed in xenografts expressing *LMO3-BORCS5* compared to empty vector (**Figure 3E**). HES staining showed an increase of tumor infiltration in adipose (green arrows) and muscular (red arrows) tissues in presence of *LMO3-BORCS5* expressing tumors compared to controls (empty vector) (**Figure S4Aa**) accompanied with an increased number of cells in mitosis (yellow arrows) (**Figure S4Ab**). Importantly, *LMO3-BORCS5* expression in NIH-3T3 induced tumorigenicity (**Figure 3F**) and HES analysis showed a high number of mitotic cells in these tumors (**Figure S4B**). Of note, the presence of *LMO3-BORCS5* fusion transcript was confirmed by RT-qPCR in both A673+*LMO3-BORCS5* and NIH-3T3+*LMO3-BORCS5* tumors (Ct \approx 22 and 27 respectively).

LMO3-BORCS5 decreases sensitivity to antineoplastic agents and apoptosis

The effect of *LMO3-BORCS5* on drug sensitivities was assessed by a MTT test. Eleven anti-cancer drugs were screened, including 5 received by the patient (enlisted in Material and Methods). Only vincristine and SN-38 showed an increase of IC50 values in A673+*LMO3-BORCS5* (**Figure 4A**). The effects of vincristine and SN-38 treatment on early and late apoptosis of A673+*LMO3-BORCS5* was tested (**Figure 4B**). A reduction of cells in early apoptosis to vincristine treatment was observed in A673+*LMO3-BORCS5* cells compared to A673 empty vector (delta between treated and non-treated cells of 28.8% for A673+empty vector vs 13.7% for A673+*LMO3-BORCS5*). The same tendency was observed after treatment with SN-38 but was not statistically significant (delta of 31.1% for A673+empty vector vs 23.5% for A673+*LMO3-BORCS5*) (**Figure 4B**). Late apoptosis after vincristine or SN-38 treatment appeared to be unchanged in presence or absence of *LMO3-BORCS5*.

Expression of BAX, MDM2, NOXA, TP53 and P21 (also called CDKN1A) was assessed by RT-qPCR (**Figure 4C**). For A673, a significant decrease of all gene expressions was observed in

A673+*LMO3-BORCS5* compared to A673+empty vector except for TP53 (**Figure 4C, upper panels**). Concerning NIH-3T3, only P21 expression was decreased in NIH-3T3+*LMO3-BORCS5* (**Figure 4C, lower panels**).

LMO3-BORCS5 downregulates genes associated with cell differentiation pathways

Microarray analysis was performed on 2 clones of A673+*LMO3-BORCS5* and compared to A673 empty vector. A total of 295 genes were found differentially expressed [$-2 < \log_2(\text{fold change}) > 2$] (**Figure S3**). The heat map showed a clear clustering of the genes differentially regulated in A673+*LMO3-BORCS5* vs A673+empty vector clones (**Figure S3 A**), confirming that the differential in gene expressions observed is exclusively due to the presence of *LMO3-BORCS5*. The volcano plot showed that more than 80% of differentially expressed genes were downregulated (243 genes downregulated vs 52 genes upregulated; **Figure S3 B**).

Using ToppGene Suite, the pathways enriched in the 295 genes differentially expressed between A673+empty vector and A673+*LMO3-BORCS5* identified by microarray were analyzed (**Table 1**). When focusing on the 12 most significant enriched pathways in upregulated genes, they can be grouped into 3 different cellular functions: metabolism, extracellular matrix synthesis and proliferative signaling (**Table 1, upper panel**). Regarding the 12 most significant enriched pathways in downregulated genes, 60% (7/12) were related to cell differentiation, more precisely, to neural (e.g. “DCC signaling”, “NCAM1 interactions”, “TRKA signaling”), cardiac (e.g. “cardiac conduction signaling”, “muscle contraction pathway”) or renal (“aldosterone-regulated sodium reabsorption”) differentiations as well as cell pluripotency (transcriptional regulation of pluripotent stem cells pathway). Finally, apoptosis-related genes *BAX*, *NOXA*, *MDM2* and *P21*, found differentially expressed between A673+empty vector and A673+*LMO3-BORCS5* by RT-qPCR were also found to be downregulated in the microarray results (**Figure 4**). The microarray data related to the study have been submitted to the Array Express data repository at the European Bioinformatics Institute (<http://www.ebi.ac.uk/arrayexpress/>), under the accession number: E-MTAB-7021.

Then, we validated by RT-qPCR 18 genes highly linked to cancer pathways and found differentially expressed by the microarray analysis. All tested genes except *ABCC3* were validated, representing 95% of reliability (**Table 2**). The large majority of tested genes were found to be downregulated in presence of *LMO3-BORCS5*, while Fyn-related kinase (*FRK*) was the only one upregulated. Among these genes are found genes regulating cell fate determination (*FGF1* and *WNT6*), genes linked to neural differentiation (*ALK*, *CUX2*, *FGF13*, *NTRK1*, *PAX7* and *SEMA3C*), *IGF2*, *MYOD1* and *RHOJ* for muscle cell differentiation and *NRP1* regulating cardiac differentiation as well as genes linked to cell proliferation or migration (*P21*, *PRKCA*, *MMP10* and *LZTS1*; **Table 2**).

BORCS5 disruption decreases cell migration and resistance in vitro but increases tumorigenicity

First, we compared the two derived clones obtained after KO of BORCS5 by CRISPR-Cas9 system. We found that both C1 and C2 did not express BORCS5 protein testifying the knockdown outcome (**Figure S5A**), showed the *quasi* same doubling time (**Figure S5B**) as well as identical repartition of cells within the cell cycle (**Figure S5C**). Therefore, only KO BORCS5 C1 was used for the further studies, named "A673 KO BORCS5".

In order to understand the role of BORCS5 on oncogenesis and tumorigenicity, *BORCS5* gene was depleted using the CRISPR/Cas9 system in A673 cells. Sanger sequencing on clones harboring at least a heterozygous deletion of BORCS5 after PCR. We identified an homozygous KO BORCS5 clone with on one allele a deletion of 950 bp in the exon 4, and on the other allele an insertion of one nucleotide leading to a frameshift (**Figure S6 A et S6 B**). The corresponding protein loss was proved by Western Blot (**Figure S6 C**). Then, the A673 WT was compared to the A673 KO BORCS5 regarding proliferation, migration, sensitivity to anticancer drugs and tumorigenicity. A significant cell migration inhibition of 20% was noticed in A673 KO BORCS5 compared to A673 WT (**Figure 5A**).

The consequence of BORCS5 loss on drug sensitivities was assessed by MTT test with the eleven anti-cancer drugs previously tested. A significant decrease of IC50 values (60% vincristine and 40% for both SN-38 and mafosfamide) was observed in A673 KO BORCS5 compared to WT (**Figure 5B**). In parallel, the effects of vincristine and SN-38 on early and late apoptosis of A673 KO BORCS5 and A673WT cells were tested (**Figure 5C**). A significant increase of cells in late apoptosis after vincristine treatment was observed in A673 KO BORCS5 cells compared to A673 WT (delta between treated and non-treated cells of 18.4% for A673 WT vs 43% for A673 KO BORCS5). The same tendency was observed after treatment with SN-38 but was not significant (delta of 32.6% vs 40.2%) while early apoptosis in A673 was not affected by the KO BORCS5 (**Figure 5C**).

Interestingly, after xenografting of A673 WT and A673 KO BORCS5 cells, an increase of tumor growth was observed for A673 KO BORCS5 compared to A673 WT (**Figure 5D**). HES revealed a similar general architecture of collected tumors (**Figure S7 A**). However, an increase of invasion in striated muscle fibers was observed in A673 KO BORCS5 tumors compared to A673 WT tumors (**Figure S7 A, red arrows**). In contrast, no difference was observed regarding the number of cells in mitosis (**Figure S7 B**).

Discussion

This is the first study which reports the oncogenic function of a newly described fusion transcript LMO3-BORCS5. Initially found in a patient with EWS at relapse in addition to the tumor type characteristic EWS/FLI1 fusion, this fusion transcript was detectable also at diagnosis and in other cell lines albeit at lower levels and thus may not have been reported to date. According to cBioPortal cancer database, LMO3 and BORCS5 are altered in 1.5% of total cancers, and fusions represent the third and second alteration type respectively. Although this percentage appears to be low, it is probably underestimated as it comprises both adult and childhood cancers and several studies do not assess the presence of fusion genes.

LMO3-BORCS5 was detected in two variant forms (v1 and v2) both leading after translation to a loss of the LIM2 domain. Interestingly, the loss of interaction domains in fusions has been shown to be a sign of oncogenicity (35,36). The variant 1 of the fusion transcript LMO3-BORCS5 was abundant at relapse and barely detected at diagnosis, suggesting that the fusion transcript was not caused by prior anticancer treatments but probably initially present and selected during disease progression. LMO3-BORCS5 transcript was also detected in 41% of the screened cell lines and found to be positive for either LMO3-BORCS5 v1 and/or v2. Even if most of the cell lines tested derived from pediatric cancer types, the positive ones were of diverse histological origins. However, the expression level of LMO3-BORCS5 in these cell lines was relatively low as previously observed for other chimeras (37). This could suggest that, a second event might be required for LMO3-BORCS5 oncogenicity as reported for several fusions (38,39).

When assessing the expression of wild-type LMO3 and BORCS5 in 31 different tumor cell lines, LMO3 was mainly expressed in neuroblastoma and EWS. This result was concordant, as LMO3 is a brain-specific gene and both pathologies are thought to arise from precursors of the neural crest or neuroectoderm, respectively (39,40). Concerning BORCS5, it was expressed in all cell lines tested. The highest expression was found in cells of EWS and osteosarcoma bone tumors, as well as medulloblastoma. Therefore, LMO3 and BORCS5 should be explored in tumors derived from neural or bone precursors.

Our first aim was to investigate the function of LMO3-BORCS5 on oncogenesis. We showed that only LMO3-BORCS5 v1 plays a key role in oncogenesis: it increased proliferation in both cell lines (A673 and NIH-3T3) and increased or induced tumor growth and invasiveness in the A673 Ewing sarcoma and the non-tumorigenic NIH-3T3, respectively. This was strengthened by microarray analysis where LMO3-BORCS5 led to massive downregulation of the differentiation process. In fact, the LIM2 domain was shown to be critical for erythropoietic differentiation in LMO2 protein, highlighting the importance of LIM2 in cell differentiation (41). Loss of LIM2 within LMO3-BORCS5 could explain the dedifferentiation and consequently the oncogenic effects observed. Dedifferentiation is considered as one of the main characteristics of cancer cells, which are blocked at early stages of differentiation resulting in an over-proliferative state (42). This was

reflected by a shift of cells towards mitosis phase in presence of LMO3-BORCS5 observed by FACS and tumor histological study.

Additionally, decreased sensitivity to two antineoplastic agents (vincristine and SN-38) was observed in A673 cells expressing *LMO3-BORCS5*. We hypothesized that the inhibition of apoptotic gene expressions (BAX, MDM2, NOXA and P21) could favor resistance to cell death, and allow the selection through treatments of a minor clone expressing *LMO3-BORCS5*.

The absence of oncogenic effects for *LMO3-BORCS5* v2 would suggest that the complete conservation of LIM1 domain is required for oncogenicity.

The upregulated genes observed in microarray in presence of LMO3-BORCS5 appeared to act on proliferative signaling, consistent with results obtained *in vitro*. Some of these genes were also belonging to pathways related to extracellular matrix synthesis, an aspect not impacted when observing migration *in vitro* but important *in vivo* as tumors with LMO3-BORCS5 were more invasive. This could reflect the limits of the scratch test used *in vitro*, a model of study that could be inadequate to our cells. Of note, the Fyn-related kinase (FRK) was increased in presence of LMO3-BORCS5. This non-receptor tyrosine kinase leads to an activation of STAT3 pathway and inhibition of TP53, and is highly expressed in sarcomas (43). However, sensitivity to dasatinib with or without LMO3-BORCS5 showed no difference to treatment as described for hepatocellular carcinoma cell lines (43).

As LMO3-BORCS5 fusion leads to a frameshift of BORCS5 with the putative formation of 43 new amino-acids, our second aim was to understand the impact of BORCS5 loss on tumorigenicity. The disruption of BORCS5 in A673 cells appeared to have both pro- or anti-tumorigenic effects, which could depend on the tumor type and mainly on the microenvironment context. *In vivo* results would go in favor of a tumor suppressor role of BORCS5, consistent with the loss observed in the formation of LMO3-BORCS5 fusion in both patient's tumor and cell lines. Herein, the disruption of BORCS5 after LMO3-BORCS5 formation could have two roles (i) its loss on one allele could serve as "second hit" or (ii) the 43 new amino-acids created in the C-terminal part of the fusion protein after LIM1 domain could enable the stabilization of the new smaller LMO3 protein, preventing from its degradation (27).

Several studies previously reported that only in-frame transcripts needed to be considered as drivers (44,45), however we contrastingly show that an out-of-frame fusion can have highly oncogenic properties.

These results put a new insight on non-recurrent alterations and infirm previous studies suggesting that non-recurrent singleton fusions represent passenger mutations (27,46). As fusions appeared to be segregated according tumor type and given the rarity of EWS among population, the statistical power of our cohort might be insufficient to detect a recurrence (1,27). Additionally, it is possible that LMO3-BORCS5 was present in other patients but at lower levels,

undetectable with the coverage used for RNA-sequencing as clinically relevant somatic alterations might represent less than 5% of sequence reads (27).

In conclusion, we discovered a new fusion oncogene LMO3-BORCS5 that put the accent on the critical role of fusion transcripts in tumor initiation, progression and resistance even in the individual patient. Their deep study could explain some consequences missed by traditional analyses. This suggest a need of mind changing and favors the vision that cancer can only be treated by considering pharmacogenomics and precision medicine rather than looking for universality. Fusion transcripts being one of the most remarkable tools used in personalized medicine should be systematically integrated to new genome-based clinical trials.

Acknowledgements

This study was supported by Université Paris-Sud (CD) and CNRS and microarray analysis was supported by the "Taxe d'apprentissage Gustave Roussy - 2017 - CD".

We acknowledge Drs Nathalie Gaspar and Estelle Daudigeos-Dubus as well as the ITCC (Innovative Therapies for Children with Cancer) consortium for providing the pediatric cell lines and antineoplastic molecules used in the study, as well as Catherine Richon (technical assistance). We acknowledge Léna Zig and Linda Ramdani for their precious technical help for the cloning of *LMO3-BORCS5*.

The animal studies were performed within the animal facility of Gustave Roussy and authors gratefully thank the technical support of Olivia Bawa and expertise of Dr Paule Opolon (histocytopathology platform) for the preparation and analysis of histological slides.

References

1. Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. *Nat Rev Cancer*. 2007;7:233–45.
2. Dupain C, Harttrampf AC, Urbinati G, Geoerger B, Massaad-Massade L. Relevance of Fusion Genes in Pediatric Cancers: Toward Precision Medicine. *Mol Ther - Nucleic Acids*. 2017;6:315–26.
3. Kumar-Sinha C, Kalyana-Sundaram S, Chinnaiyan AM. Landscape of gene fusions in epithelial cancers: seq and ye shall find. *Genome Med*. 2015;7:129.
4. Cachia D, Wani K, Penas-Prado M, Olar A, McCutcheon IE, Benjamin RS, et al. C11orf95–RELA fusion present in a primary supratentorial ependymoma and recurrent sarcoma. *Brain Tumor Pathol*. 2015;32:105–11.
5. Graham RP, Jin L, Knutson DL, Kloft-Nelson SM, Greipp PT, Waldburger N, et al. DNAJB1-

PRKACA is specific for fibrolamellar carcinoma. *Mod Pathol Off J U S Can Acad Pathol Inc.* 2015;28:822-9.

6. PANAGOPOULOS I, GORUNOVA L, BJERKEHAGEN B, BOYE K, HEIM S. Chromosome aberrations and HEY1-NCOA2 fusion gene in a mesenchymal chondrosarcoma. *Oncol Rep.* 2014;32:40-4.

7. Arpinati M, Tolomelli G, Bochicchio MT, Castagnetti F, Amabile M, Bandini G, et al. Molecular monitoring of BCR-ABL transcripts after allogeneic stem cell transplantation for chronic myeloid leukemia. *Biol Blood Marrow Transplant J Am Soc Blood Marrow Transplant.* 2013;19:735-40.

8. Kalinova M, Krskova L, Brizova H, Kabickova E, Kepak T, Kodet R. Quantitative PCR detection of NPM/ALK fusion gene and CD30 gene expression in patients with anaplastic large cell lymphoma--residual disease monitoring and a correlation with the disease status. *Leuk Res.* 2008;32:25-32.

9. Han S, Brenner JC, Sabolch A, Jackson W, Speers C, Wilder-Romans K, et al. Targeted radiosensitization of ETS fusion-positive prostate cancer through PARP1 inhibition. *Neoplasia N Y N.* 2013;15:1207-17.

10. Hartmaier RJ, Trabucco SE, Priedigkeit N, Chung JH, Parachoniak CA, Vanden Borre P, et al. Recurrent hyperactive ESR1 fusion proteins in endocrine therapy-resistant breast cancer. *Ann Oncol Off J Eur Soc Med Oncol.* 2018;29:872-80.

11. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature.* 2013;499:214-8.

12. Dancy JE, Bedard PL, Onetto N, Hudson TJ. The Genetic Basis for Cancer Treatment Decisions. *Cell.* 2012;148:409-20.

13. Childhood Genomics Workshop NIH [Internet]. [cited 2015 Nov 20]. Available from: <http://www.ncbi.nlm.nih.gov/gate1.inist.fr/pmc/articles/PMC3403905/pdf/1471-2164-13-147.pdf>

14. Massard C, Michiels S, Ferté C, Le Deley M-C, Lacroix L, Hollebecque A, et al. High-Throughput Genomics and Clinical Outcome in Hard-to-Treat Advanced Cancers: Results of the MOSCATO 01 Trial. *Cancer Discov.* 2017;7:586-95.

15. Harttrampf AC, Lacroix L, Deloger M, Deschamps F, Puget S, Auger N, et al. Molecular Screening for Cancer Treatment Optimization (MOSCATO-01) in Pediatric Patients: A Single-Institutional Prospective Molecular Stratification Trial. *Clin Cancer Res Off J Am Assoc Cancer Res.* 2017;23:6101-12.

16. Sang M, Ma L, Sang M, Zhou X, Gao W, Geng C. LIM-domain-only proteins: multifunctional nuclear transcription coregulators that interacts with diverse proteins. *Mol Biol Rep.*

2014;41:1067–73.

17. Aoyama M, Ozaki T, Inuzuka H, Tomotsune D, Hirato J, Okamoto Y, et al. LMO3 interacts with neuronal transcription factor, HEN2, and acts as an oncogene in neuroblastoma. *Cancer Res.* 2005;65:4587–97.
18. Larsen S, Yokochi T, Isogai E, Nakamura Y, Ozaki T, Nakagawara A. LMO3 interacts with p53 and inhibits its transcriptional activity. *Biochem Biophys Res Commun.* 2010;392:252–7.
19. Song Y-F, Hong J-F, Liu D-L, Lin Q-A, Lan X-P, Lai G-X. miR-630 targets LMO3 to regulate cell growth and metastasis in lung cancer. *Am J Transl Res.* 2015;7:1271–9.
20. Qiu Y-S, Jiang N-N, Zhou Y, Yu K-Y, Gong H-Y, Liao G-J. LMO3 promotes gastric cancer cell invasion and proliferation through Akt-mTOR and Akt-GSK3 β signaling. *Int J Mol Med.* 2018;
21. Bao J, Talmage DA, Role LW, Gautier J. Regulation of neurogenesis by interactions between HEN1 and neuronal LMO proteins. *Dev Camb Engl.* 2000;127:425–35.
22. Pu J, Schindler C, Jia R, Jarnik M, Backlund P, Bonifacino JS. BORG, a Multisubunit Complex that Regulates Lysosome Positioning. *Dev Cell.* 2015;33:176–88.
23. Langemeyer L, Ungermann C. BORG and BLOC-1: Shared Subunits in Trafficking Complexes. *Dev Cell.* 2015;33:121–2.
24. Niwa S, Tao L, Lu SY, Liew GM, Feng W, Nachury MV, et al. BORG Regulates the Axonal Transport of Synaptic Vesicle Precursors by Activating ARL-8. *Curr Biol CB.* 2017;27:2569-2578.e4.
25. Montpetit A, Boily G, Sinnett D. A detailed transcriptional map of the chromosome 12p12 tumour suppressor locus. *Eur J Hum Genet EJHG.* 2002;10:62–71.
26. Mertens F, Johansson B, Fioretos T, Mitelman F. The emerging complexity of gene fusions in cancer. *Nat Rev Cancer.* 2015;15:371–381.
27. Latysheva NS, Babu MM. Discovering and understanding oncogenic gene fusions through data intensive computational approaches. *Nucleic Acids Res.* 2016;44:4487–503.
28. Daudigeos-Dubus E, Le Dret L, Lanvers-Kaminsky C, Bawa O, Opolon P, Vievard A, et al. Regorafenib: Antitumor Activity upon Mono and Combination Therapy in Preclinical Pediatric Malignancy Models. *PloS One.* 2015;10:e0142612.
29. Dupain C, Ali H, Mouhoub T, Urbinati G, Massade L. Induction of TTF-1 or PAX-8 expression on proliferation and tumorigenicity in thyroid carcinomas. *Int J Oncol [Internet].* 2016 [cited 2016 Jul 20]; Available from: <http://www.spandidos-publications.com/10.3892/ijo.2016.3617>
30. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science.* 2013;339:819–23.
31. Haeussler M, Schönig K, Eckert H, Eschstruth A, Mianné J, Renaud J-B, et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* 2016;17:148.

32. Buart S, Terry S, Noman MZ, Lanoy E, Boutros C, Fogel P, et al. Transcriptional response to hypoxic stress in melanoma and prognostic potential of GBE1 and BNIP3. *Oncotarget*. 2017;8:108786–801.
33. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods*. 2001;25:402–8.
34. Urbinati G, Ali HM, Rousseau Q, Chapuis H, Desmaële D, Couvreur P, et al. Antineoplastic Effects of siRNA against TMPRSS2-ERG Junction Oncogene in Prostate Cancer. *PLoS ONE*. 2015;10:e0125277.
35. Shugay M, Ortiz de Mendíbil I, Vizmanos JL, Novo FJ. Genomic hallmarks of genes involved in chromosomal translocations in hematological cancer. *PLoS Comput Biol*. 2012;8:e1002797.
36. Shugay M, Ortiz de Mendíbil I, Vizmanos JL, Novo FJ. Oncofuse: a computational framework for the prediction of the oncogenic potential of gene fusions. *Bioinforma Oxf Engl*. 2013;29:2539–46.
37. Frenkel-Morgenstern M, Lacroix V, Ezkurdia I, Levin Y, Gabashvili A, Prilusky J, et al. Chimeras taking shape: potential functions of proteins encoded by chimeric RNA transcripts. *Genome Res*. 2012;22:1231–42.
38. Keller C, Arenkiel BR, Coffin CM, El-Bardeesy N, DePinho RA, Capecchi MR. Alveolar rhabdomyosarcomas in conditional Pax3:Fkhr mice: cooperativity of Ink4a/ARF and Trp53 loss of function. *Genes Dev*. 2004;18:2614–26.
39. Lin PP, Pandey MK, Jin F, Xiong S, Deavers M, Parant JM, et al. EWS-FLI1 induces developmental abnormalities and accelerates sarcoma formation in a transgenic mouse model. *Cancer Res*. 2008;68:8968–8975.
40. Tsubota S, Kadomatsu K. Origin and mechanism of neuroblastoma. *Oncoscience*. 2017;4:70–2.
41. Terano T, Zhong Y, Toyokuni S, Hiai H, Yamada Y. Transcriptional control of fetal liver hematopoiesis: dominant negative effect of the overexpression of the LIM domain mutants of LMO2. *Exp Hematol*. 2005;33:641–51.
42. Cooper GM. *The Development and Causes of Cancer*. Cell Mol Approach 2nd Ed [Internet]. 2000 [cited 2018 May 20]; Available from: <https://www.ncbi.nlm.nih.gov/books/NBK9963/>
43. Goel RK, Lukong KE. Understanding the cellular roles of Fyn-related kinase (FRK): implications in cancer biology. *Cancer Metastasis Rev*. 2016;35:179–99.
44. Teles Alves I, Hartjes T, McClellan E, Hiltmann S, Böttcher R, Dits N, et al. Next-generation sequencing reveals novel rare fusion events with functional implication in prostate cancer. *Oncogene*. 2015;34:568–77.
45. Abate F, Zairis S, Ficarra E, Acquaviva A, Wiggins CH, Frattini V, et al. Pegasus: a comprehensive annotation and prediction tool for detection of driver gene fusions in cancer. *BMC*

Syst Biol. 2014;8:97.

46. Stransky N, Cerami E, Schalm S, Kim JL, Lengauer C. The landscape of kinase fusions in cancer. Nat Commun. 2014;5:4846.

Tables

	Pathway name	Source ^a	pValue	Nb of genes from input	Nb of genes in Annotation
Upregulated	Ensemble of genes encoding core extracellular matrix including ECM glycoproteins, collagens and proteoglycans	MSigDB C2 BIOCARTA (v6.0)	9.564E-4	5	275
	G-protein beta:gamma signalling	BioSystems: REACTOME	8.980E-3	2	51
	Ensemble of genes encoding extracellular matrix and extracellular matrix-associated proteins	MSigDB C2 BIOCARTA (v6.0)	6.614E-3	8	1028
	Defective CD320 causes methylmalonic aciduria	BioSystems: REACTOME	5.615E-3	1	2
	sulfate activation for sulfonation	BioSystems: BIOCYC	5.615E-3	1	2
	Sulfate assimilation	PantherDB	5.615E-3	1	2
	Sulfoglycolipids biosynthesis, ceramide/1-alkyl-2-acylglycerol => sulfatide/seminolipid	BioSystems: KEGG	5.615E-3	1	2
	inositol metabolic	Pathway Ontology	5.615E-3	1	2
	African trypanosomiasis	BioSystems: KEGG	4.310E-3	2	35
	Phase II conjugation	BioSystems: REACTOME	3.668E-3	3	111
	propanoate metabolic	Pathway Ontology	3.595E-2	1	13
	TNFR2 non-canonical NF-kB pathway	BioSystems: REACTOME	3.560E-2	2	106
Downregulated	Collagen degradation	BioSystems: REACTOME	9.841E-3	3	41
	DCC mediated attractive signaling	BioSystems: REACTOME	9.752E-3	2	14
	Breast cancer	BioSystems: KEGG	9.665E-4	7	144
	Cardiac conduction	BioSystems: REACTOME	8.902E-4	7	142
	POU5F1 (OCT4), SOX2, NANOG activate genes related to proliferation	BioSystems: REACTOME	8.418E-3	2	13
	Ensemble of genes encoding extracellular matrix and extracellular matrix-associated proteins	MSigDB C2 BIOCARTA (v6.0)	7.403E-3	20	1028
	Aldosterone-regulated sodium reabsorption	BioSystems: KEGG	7.402E-3	3	37
	NCAM1 (neural cell adhesion molecule) interactions	BioSystems: REACTOME	7.402E-3	3	37
	Trka Receptor Signaling Pathway (neuronal survival and differentiation)	MSigDB C2 BIOCARTA (v6.0)	7.174E-3	2	12
	insulin-like growth factor signaling	Pathway Ontology	7.174E-3	2	12
	Transcriptional regulation of pluripotent stem cells	BioSystems: REACTOME	6.855E-3	3	36
	Muscle contraction	BioSystems: REACTOME	6.783E-3	7	204

Table 1: Enrichment pathway analysis of the genes differentially expressed between A673 empty vector and A673+LMO3-BORCS5 identified by microarray.

A list of 295 differentially expressed genes in presence of *LMO3-BORCS5* in A673 cells was obtained from microarray analysis and analyzed for pathways enrichment using ToppGene Suite (<https://toppgene.cchmc.org>). The 12 most significant enriched pathways in both upregulated and downregulated genes are represented. The number of genes from input comprises the genes identified in the microarray analysis while the number of genes in annotation represents the total genes enlisted in the corresponding pathway database.

^a Source of pathways database associated. MSigDB C2 BIOCARTA (v6.0): Molecular Signatures Database, collection 2 “curated gene sets” from Broad Institute (<http://software.broadinstitute.org/gsea/msigdb/collections.jsp>); BioSystems: REACTOME: NCBI database of biological systems *accessible on Reactome* (<https://reactome.org/>); BioSystems: BIOCYC: NCBI database of biological systems *BioCyc* (<http://biocyc.org/>); PANTHER DB: Protein ANalysis THrough Evolutionary Relationships DataBase (<http://www.pantherdb.org/>); BioSystems: KEGG: Kyoto Encyclopedia of Genes and Genomes (<http://www.genome.jp/kegg/>); Pathway ontology: database available at <https://bioportal.bioontology.org/ontologies/PW>.

Genes tested in qPCR	Fold change	qPCR result	Function ^a	References in literature	Network (Ingenuity) ^b
ABCC3	2,01	=	ATP-binding cassette (ABC) transporters. Involved in drug resistance	Zhao et. al, 2013	Cancer, Organismal injury and abnormalities, cell injury
ALK	-3,129	↘	Receptor tyrosine kinase, role in proliferation, growth, transformation, apoptosis. Important role in the genesis and differentiation of the nervous system	Yau et. al, 2015	Cancer, Organismal injury and abnormalities, cell injury
CUX2	-2,788	↘	May be a transcription factor involved in neural	Duan et. al. 2018	Cancer. Organismal injury and abnormalities.
FGF1	-2,065	↘	Acts as a modifier of endothelial cell migration and proliferation and as an angiogenic factor. Thought to be also involved in organogenesis	Li et. al, 2015	Endocrine system disorders, gastrointestinal disease, metabolic disease
FGF13	-5,872	↘	Involved in embryonic development, cell growth, morphogenesis, tissue repair, tumor growth and invasion. Role in neuron polarization and migration in the cerebral cortex and hippocampus	Okada et. al, 2013	Endocrine system disorders, gastrointestinal disease, metabolic disease
FRK	2,202	↗	Regulates growth, cell differentiation and innate immune response. Can have either pro- or anti-oncogenic roles depending on cancer types	Goel et. al, 2016	Cancer, Organismal injury and abnormalities, cell injury
IGF2	-2,646	↘	Insulin family of polypeptide growth factors, involved in development and growth. Controls muscle terminal differentiation.	Brouwer-Visser et. al, 2015	Cellular compromise, neurological disease, organismal injury and abnormalities
LZTS1	-2,181	↘	Negative regulation of cell growth. Tumor suppressor	Zhou et. al, 2015	Cancer, cell cycle, organismal injury and abnormalities
MMP10	-2,532	↘	Breakdown of extracellular matrix in physiological and disease processes, such as metastasis	Justilien et. al, 2012	Cellular compromise, neurological disease, organismal injury and abnormalities
MYOD1	-2,411	↘	Myogenic bHLH transcription factor. Regulates muscle cell differentiation and muscle regeneration	Rekhi et. al, 2016	Cancer, Organismal injury and abnormalities, cell injury
NRP1	-2,206	↘	Role in development of the cardiovascular system, angiogenesis and in organogenesis outside the nervous system	Chaudhary et. al, 2014	Cancer, cell cycle, organismal injury and abnormalities
NTRK1	-2,019	↘	Regulation of proliferation, differentiation and survival of neurons	Vaishnavi et. al, 2013	Neurological diseases, organismal injury and abnormalities, cell death and survival
P21	-2,136	↘	Promotes cell cycle arrest	Parveen et. al, 2016	Cancer, cell cycle, organismal injury and abnormalities
PAX7	-3,163	↘	Brain development. Regulates transcription.	Barr et. al. 2001	Cancer, Organismal injury and abnormalities, cell injury
PRKCA	-2,114	↘	Roles in cell adhesion, transformation, cell cycle checkpoints and cell volume control	Bridge et. al, 2013	Cell-to-cell signaling, cellular assembly and organization, carbohydrate metabolism
RHOJ	-2,752	↘	May regulate angiogenesis and is involved in the regulation of fibroblasts cells morphology	Kim et. al, 2016	Endocrine system disorders, gastrointestinal disease, metabolic disease
SEMA3C	-2,528	↘	Axon guidance, migration and blood vessels remodeling	Man et. al, 2014	Endocrine system disorders, gastrointestinal disease, metabolic disease
WNT6	-2,035	↘	Implicated in oncogenesis and in several developmental processes, including regulation of cell fate and patterning during embryogenesis	Zhang et. al, 2015	Endocrine system disorders, gastrointestinal disease, metabolic disease

Table 2: RT-qPCR validation of 18 genes differentially expressed on microarray in presence of LMO3-BORCS5 in A673 cells.

Eighteen genes differentially expressed, as reflected by the fold change obtained from the microarray analysis, and highly linked to oncogenesis and associated to cancer pathways were validated by RT-qPCR. The expression of the genes in A673+LMO3-BORCS5 cells compared to A673 empty vector was determined by quantification-comparative $2^{-\Delta\Delta Ct}$ method and normalized to GAPDH levels. All genes were differentially expressed and validated, except for ABCC3.

^a Gene function according to the human gene databases GeneCards (<http://www.genecards.org/>).

^b Ingenuity network/pathway in which the gene is included (www.ingenuity.com).

Figures

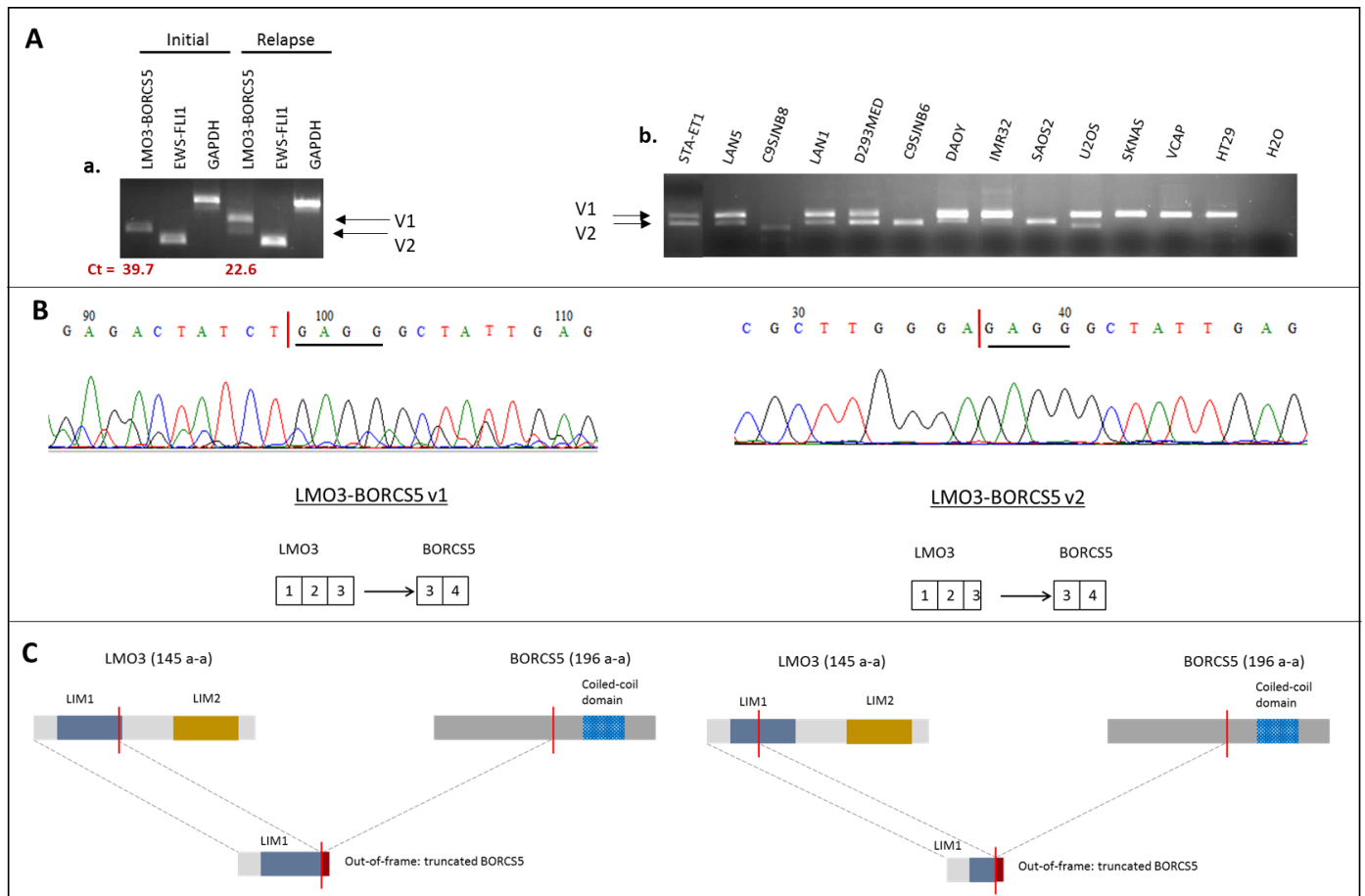


Figure 1: Detection of *LMO3-BORCS5* fusion transcript in a patient with EWS tumor samples at diagnosis and relapse and reconstitution of the different *LMO3-BORCS5* variants at mRNA and protein levels.

A. The presence of *LMO3-BORCS5* was tested by RT-qPCR in the mRNA obtained from tumor samples at diagnosis and at relapse of a patient with EWS (**a**) and from 31 cancer cell lines of different origins (**b**). After mRNA extraction and reverse transcription (RT), the presence of *LMO3-BORCS5* was assessed by real time PCR (qPCR) followed by migration of the products on 2% agarose gel. The cycle threshold (Ct) values of *LMO3-BORCS5* gene expression at diagnosis and at relapse is reported at the bottom of the figure. The expression of two variant transcripts named *LMO3-BORCS5* v1 (v1) and *LMO3-BORCS5* v2 (v2) was observed both in patient and cell lines. **B.** Sanger sequencing of *LMO3-BORCS5* v1 and v2 and graphical reconstitution at mRNA and protein level according to LMO3 sequence (NM_001243610.1).

C. Graphical reconstitution of *LMO3-BORCS5* v1 and *LMO3-BORCS5* v2 at protein level. The fusion is due to an interstitial deletion of 4.2 MB on chromosome 12 resulting in the disruption of the regulation domain of LMO3 (different regarding the variants) and disruption of BORCS5 in both variants.

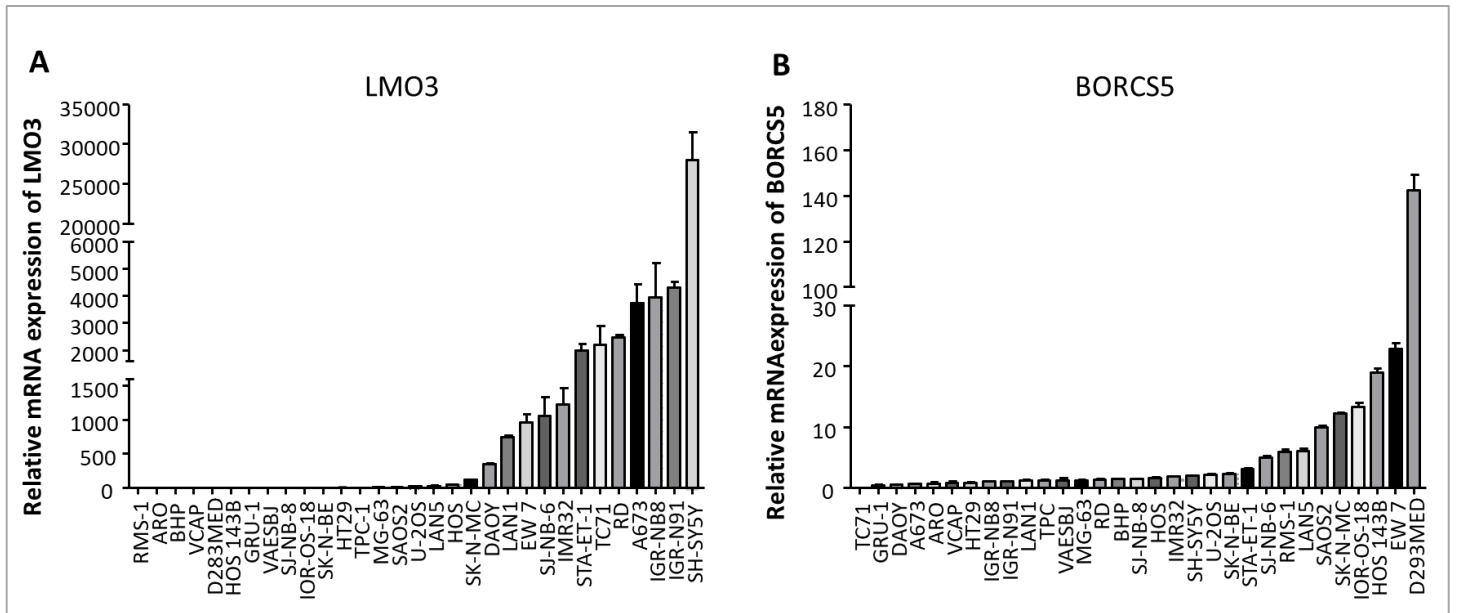


Figure 2: Profiles of *LMO3* and *BORCS5* in cell lines of various cancer types.

Expressions of *LMO3* (A) and *BORCS5* (B) genes in 31 cell lines analyzed by RT-qPCR. The results are expressed mean+SD of duplicates, as relative mRNA expression of *LMO3* or *BORCS5* according to *GAPDH*.

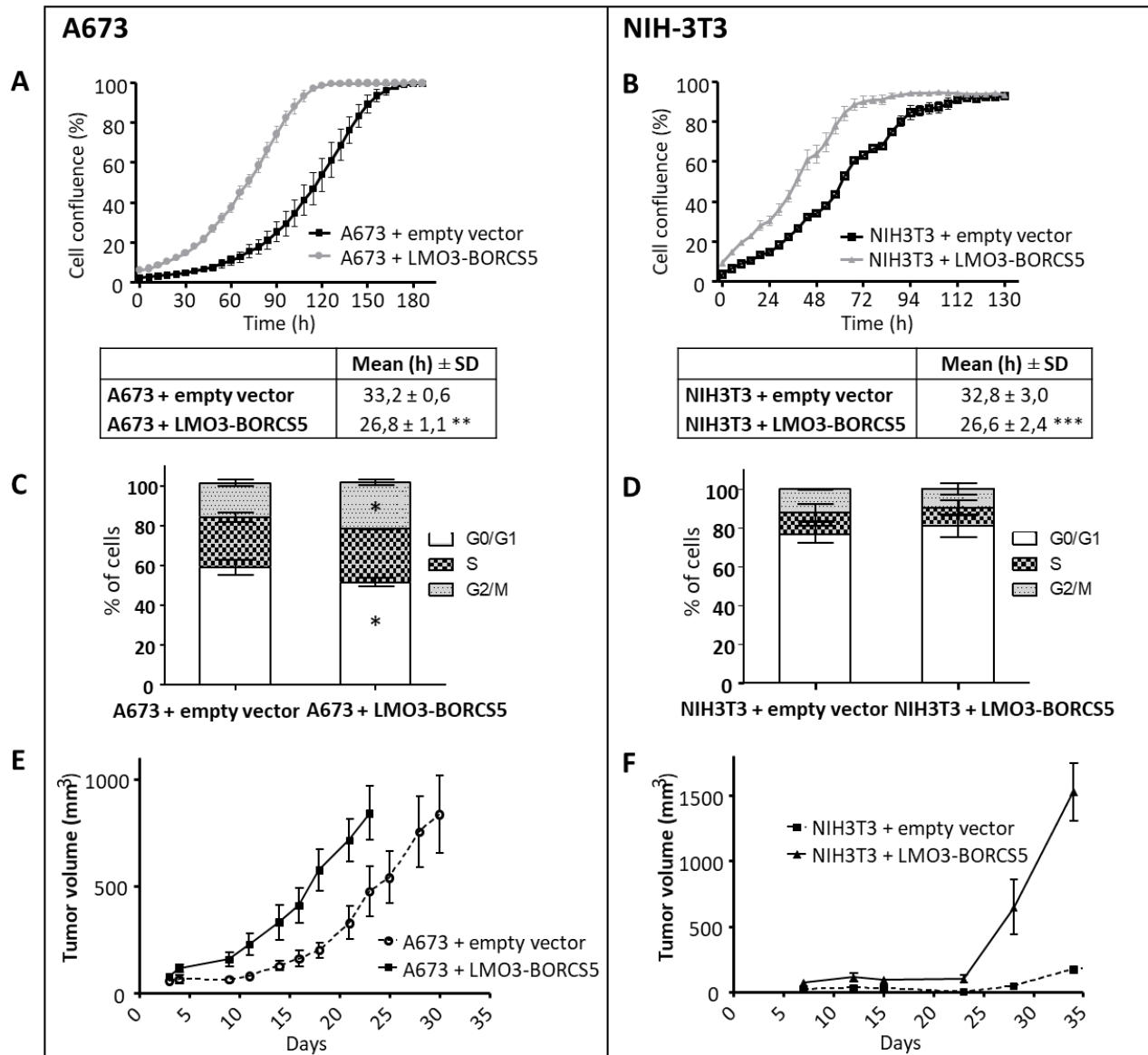


Figure 3: Effects of *LMO3-BORCS5* on cell proliferation, cell cycle and tumorigenicity in A673 and NIH-3T3 cell lines.

A673 (left) and NIH-3T3 (right) were stably transfected with the empty vector pcDNA3.1 or with pcDNA3.1 *LMO3-BORCS5*. Proliferation of A673 (A) or NIH-3T3 (B) cells using the IncuCyte™ video microscopy system was assessed. Doubling time of cells was calculated according to an exponential regression equation, given as: $y = ax^{bx}$, via the formula: $t_{1/2} = \ln(2)/b$. Results are presented as mean ± SD in hours of three independent experiments. Cell cycle repartition in A673 (C) or NIH-3T3 (D) cells obtained after DNA staining with propidium iodide-containing buffer followed by FACS analysis. The results are the mean of at least three independent experiments and represent the percentage of cells distributed in G0/G1, S and G2/M phases of the cell cycle. Tumor growth was monitored after inoculation of 10×10^6 A673 cells (E) or NIH-3T3 cells (F) in nude mice (n = 5/group). Results are presented as the tumor volume mean+SD in mm³ (n = 5) function of days.

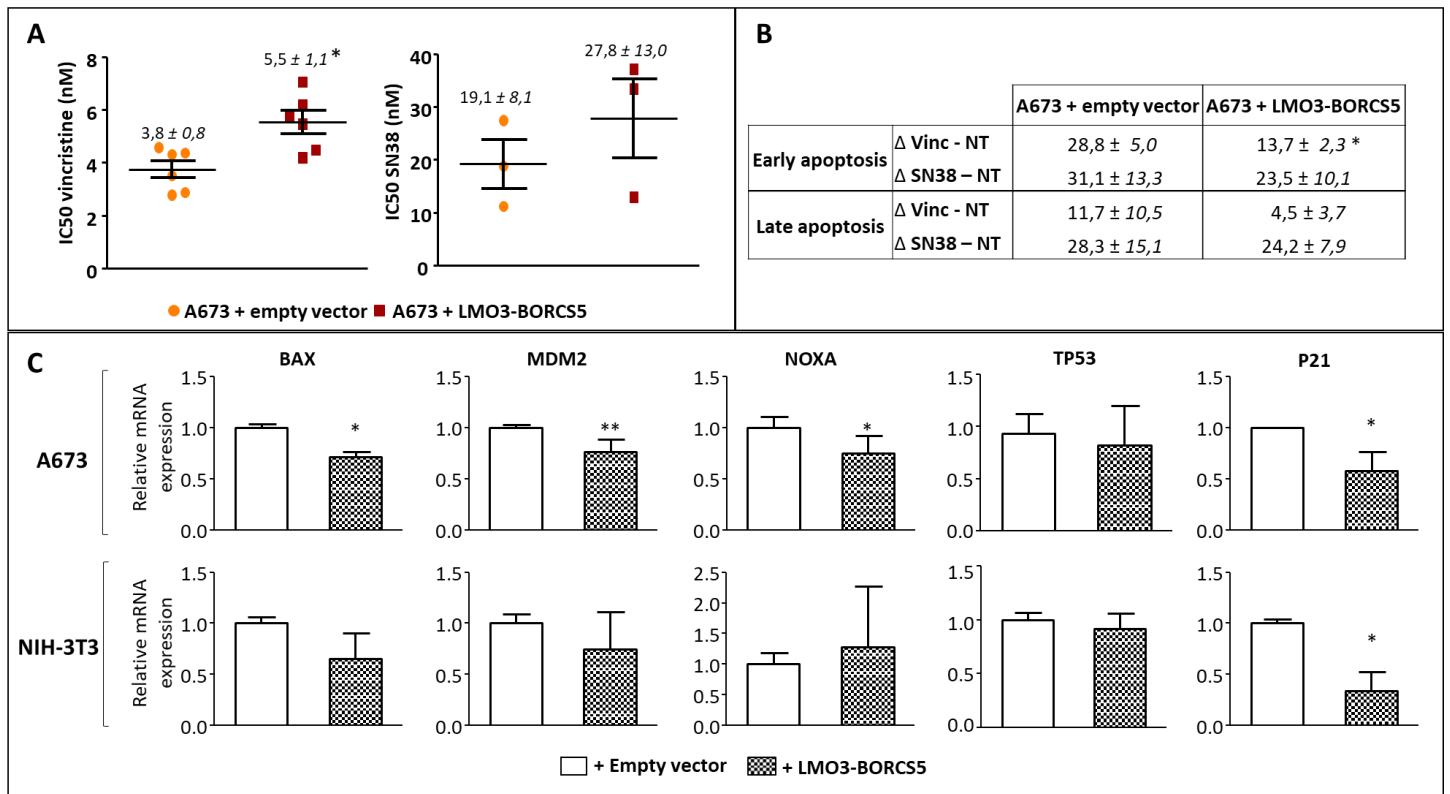


Figure 4: Effects of *LMO3-BORCS5* expression on the expression sensitivity to anti-cancer agents and of apoptosis-related genes.

A. IC₅₀ obtained after MTT assay of A673 stably transfected with pcDNA3.1 (A673+empty vector) or with pcDNA3.1 *LMO3-BORCS5* (A673+*LMO3-BORCS5*) cells treated with vincristine (Vinc) and SN-38. **B.** Early and late apoptosis detection by flow cytometry in A673 stably transfected with A673+empty vector or with A673+*LMO3-BORCS5* cells after treatment with Vinc or SN-38. The delta represents the difference between the percentage of cells in early or late apoptosis in cells treated with Vinc or SN-38 and the percentage of cells in early or late apoptosis in non-treated cells. **C.** Relative mRNA expressions of apoptosis-related genes BAX, MDM2, NOXA, TP53 and P21 by RT-qPCR in A673 and NIH3T3 cells stably transfected with empty vector pcDNA3.1 or with pcDNA3.1 *LMO3-BORCS5*. * $p < 0.05$: by using Mann-Whitney test, a statistical difference was found for Vinc treatment between A673+empty vector and A673+*LMO3-BORCS5* clones.

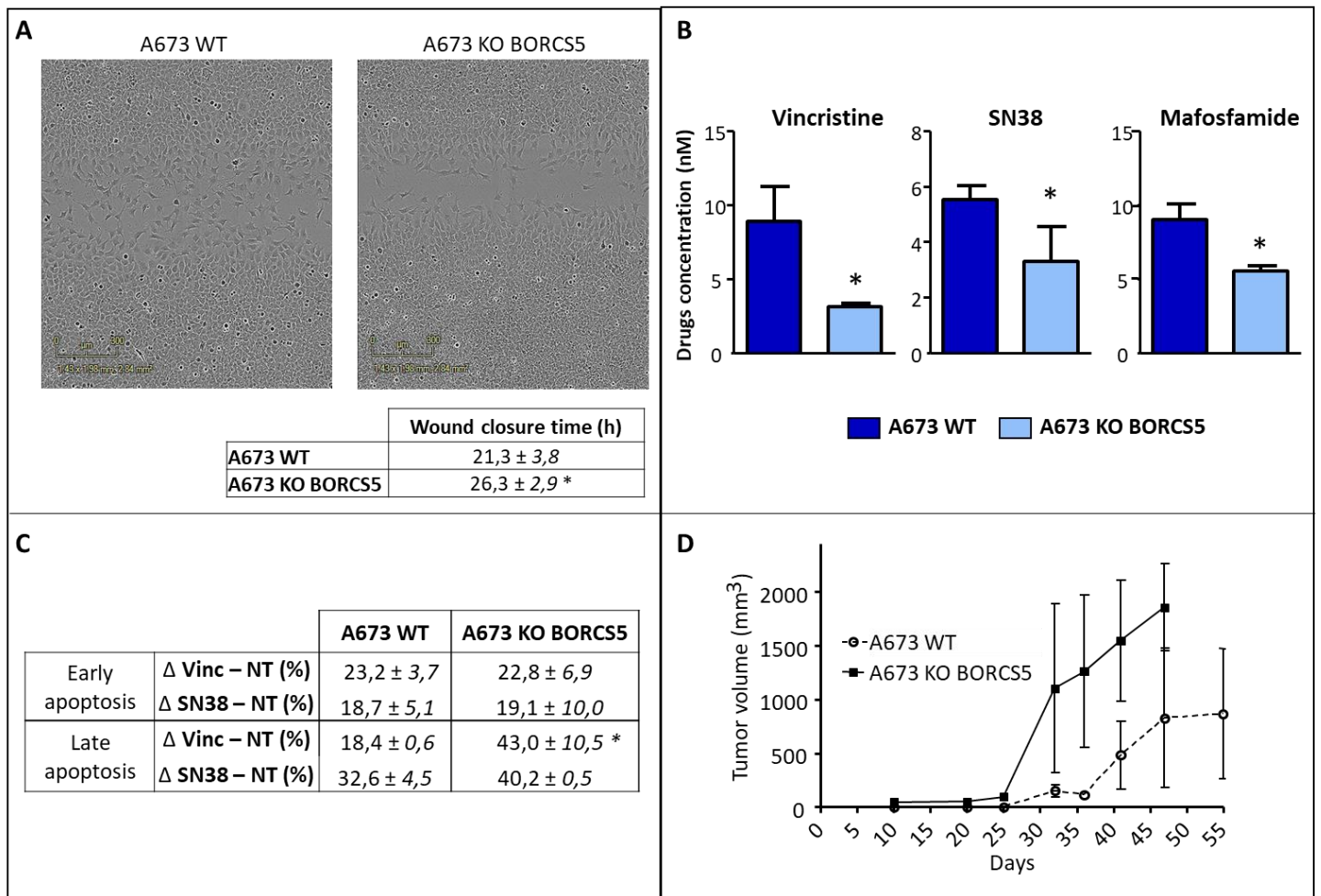


Figure 5: Impact of BORCS5 knockout on A673 cells *in vitro* and *in vivo*.

A. Effects of BORCS5 disruption on A673 cell migration. Migration ability of the cells was assessed using scratch wound test to compare A673 WT and A673 KO BORCS5. Cells in monolayer were seeded in Matrigel™ (BD) then the monolayer was scratched using a 96-pin Wound Maker and wound confluence was monitored every 4 hours IncuCyte™. Images (x10) representative of one experiment, taken by the Incucyte™ system of A673 WT or A673 KO BORCS5 cell migration after 15h of incubation at 37 °C and 5% CO₂. Results are also presented as the mean time to completely close the wound per cell line of three independent experiments. **B.** IC₅₀ obtained after MTT assay of A673 WT and A673 KO BORCS5 cells treated with vincristine, SN-38 and mafosfamide. **C.** Early and late apoptosis detection by flow cytometry in A673 WT and A673 KO BORCS5 cells after treatment with vincristine or SN-38. The delta represents the difference between the percentage of cells in early or late apoptosis in treated versus non-treated cells. **D.** 10x10⁶ cells of A673 WT and A673 KO BORCS5 cells were injected subcutaneously in Nude mice in order to monitor tumor growth and test the influence of BORCS5 disruption on tumorigenicity (n=5/group). Results are presented as the tumor volume mean in mm³ of 5 mice per group function of days.

Supplementary figures

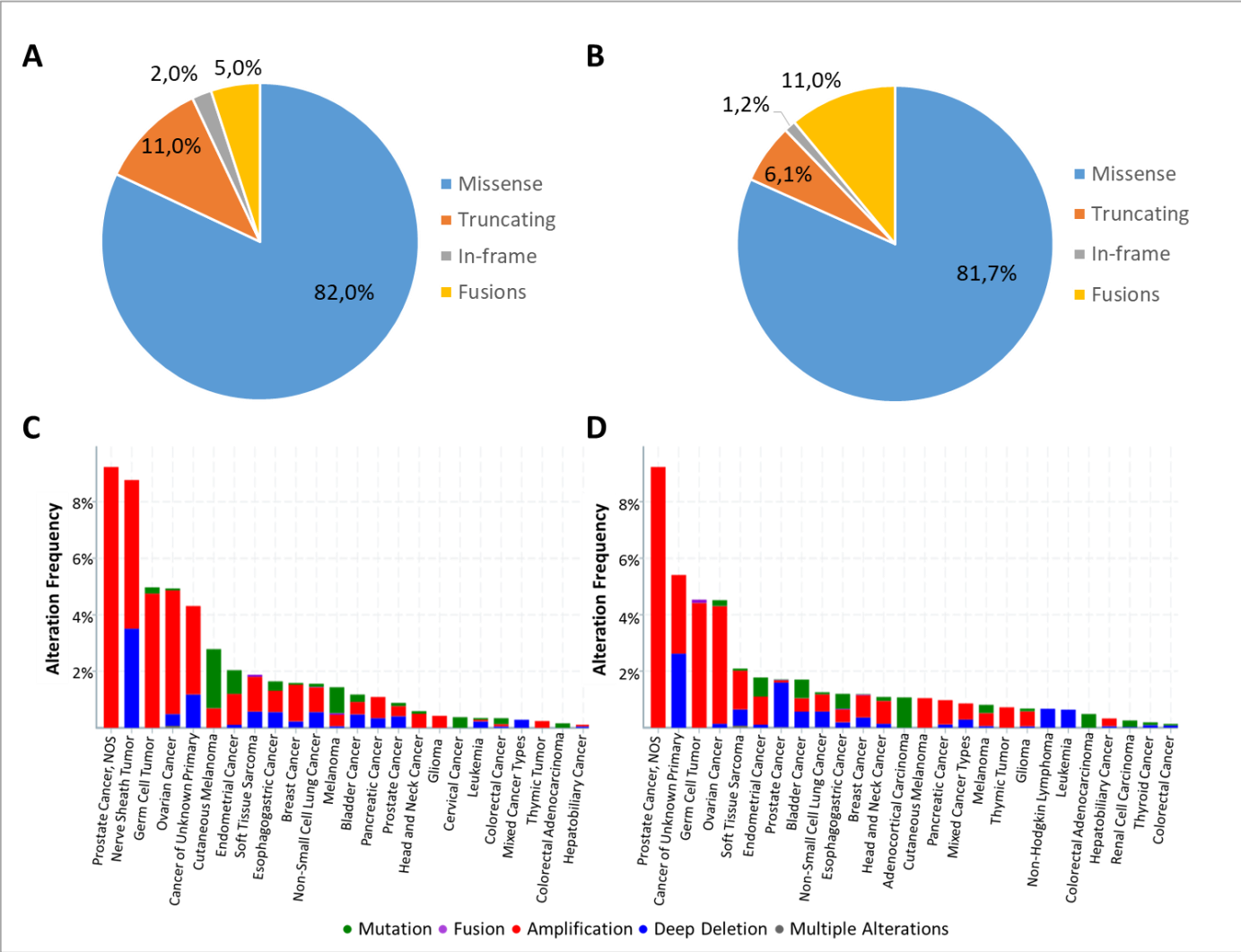


Figure S1: LMO3 and BORCS5 alterations across cancer types.

The different types of “mutations” affecting LMO3 (A) or BORCS5 (B) gene among 54510 samples analyzed included in 215 different studies, and alterations frequency in LMO3 (C) and BORCS5 (D) genes across cancer types. Results are presented as alterations frequency by cancer types. The proportion of mutations, fusions, amplifications, deep deletions and the presence of multiple alterations are represented in histograms per cancer type. Source: (<http://www.cbioportal.org>)

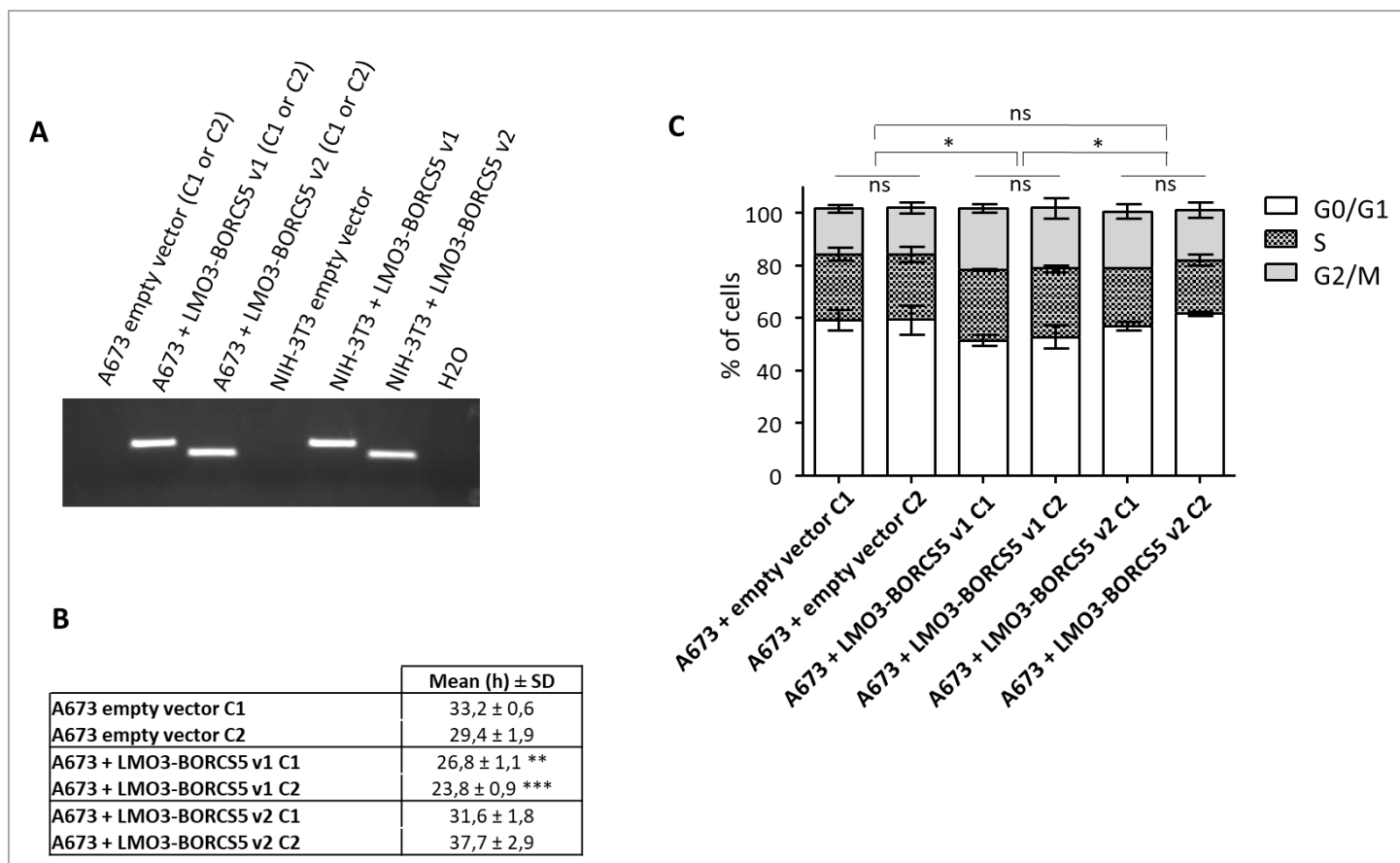


Figure S2: Validation of clones derived from A673 and NIH-3T3 cell lines stably transfected with empty vector, *LMO3-BORCS5 v1* or *LMO3-BORCS5 v2*. **A.** Empty vector pcDNA3.1, pcDNA3.1 containing *LMO3-BORCS5 v1* and pcDNA3.1 containing *LMO3-BORCS5 v2* were stably transfected into A673 (human EWS cell line) or NIH-3T3 (non-tumorigenic mouse fibroblasts cell line). Verification of *LMO3-BORCS5* presence in the clones was performed by RT-qPCR from total mRNA of each cell line followed by migration of the products on agarose gel. **B.** Doubling time of the different clones obtained after transfection of A673 cells with empty vector (clone 1 and 2), *LMO3-BORCS5 v1* (clone 1 and 2) and *LMO3-BORCS5 v2* (clone 1 and 2) cells was calculated according to an exponential regression equation, given as: $y = axe^{bx}$, via the formula: $t_{1/2} = \ln(2)/b$. Results are presented as Mean ± SD in hours of at least two independent experiments. **C.** Cell cycle repartition was obtained after DNA staining with propidium iodide-containing buffer followed by FACS analysis. Results are presented as Mean ± SD in % of cells in each cell cycle phase of at least two independent experiments. * $p > 0.05$, ** $p > 0.01$; *** $p > 0.001$ compared to A673 empty vector C1.

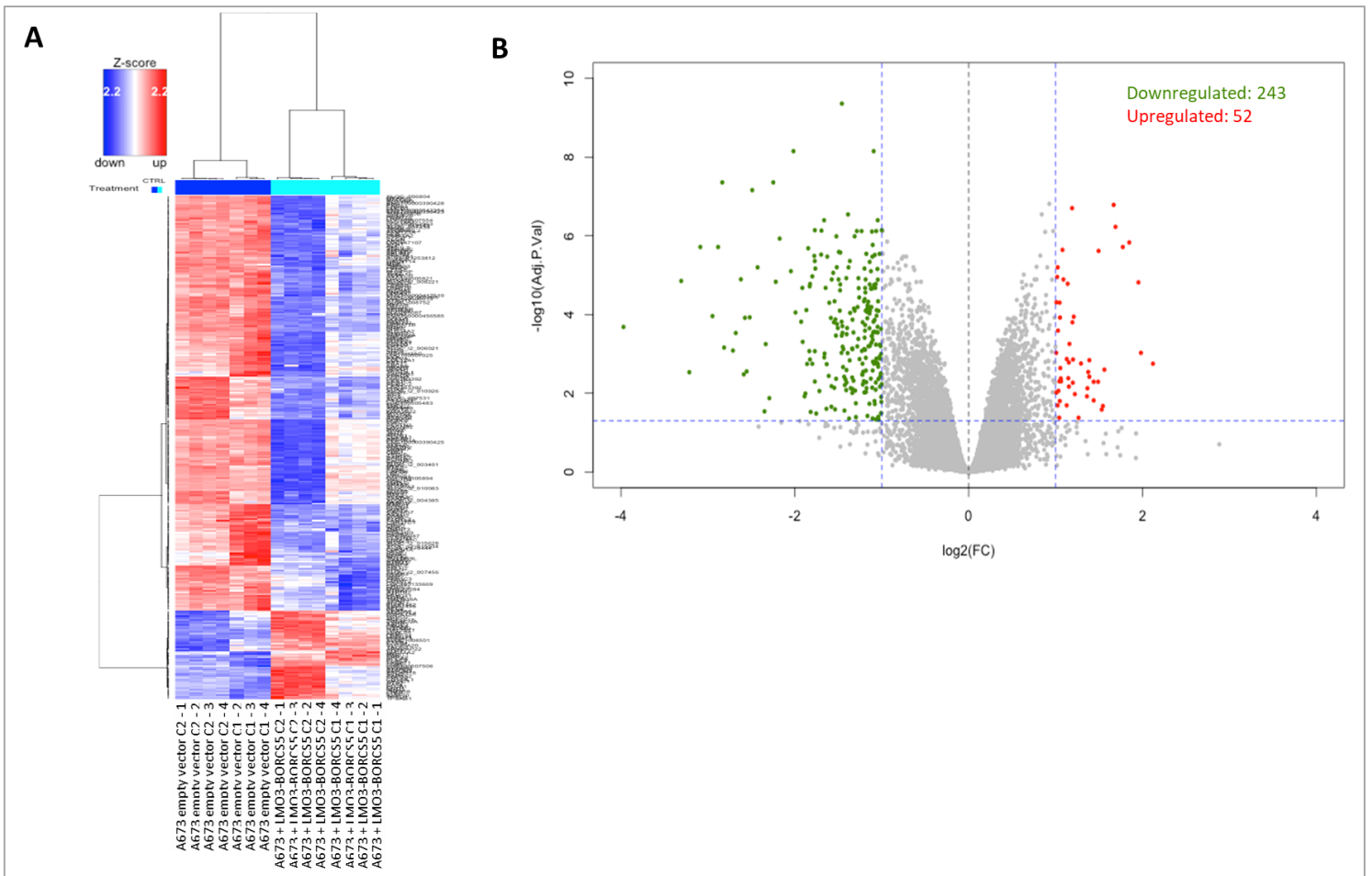


Figure S3: Differential analysis of gene expression by microarray in A673 cells stably transfected with pcDNA3.1 *LMO3-BORCS5* compared to A673 cells transfected with pcDNA3.1 empty vector.

Each clone obtained after stable transfection with empty vector pcDNA3.1 or with pcDNA3.1 *LMO3-BORCS5* were tested (C1 and C2). For each clone, four independent total RNAs extractions were performed (number 1 to 4). Agilent® SurePrint G3 Human GE 8x60K Microarray were used and images were analyzed by using Feature Extraction software with defaults settings. The top-ranked genes were selected with the following criteria: an absolute fold-change > 2 and an adjusted p-value (FDR) < 0.05. One sample (A673 empty vector C1 - 1) was discarded from the analysis due to bad quality. **A.** Heat map of the 295 genes differentially expressed in presence of *LMO3-BORCS5* presented as normalized \log_2 intensities. **B.** Volcano plot of the retrieved microarray results, plotting the negative \log_{10} of the adjusted P-value against the \log_2 of the fold change (FC). Green dots represent the significantly downregulated genes after stable transfection with pcDNA3.1 *LMO3-BORCS5* and red dots are those upregulated.

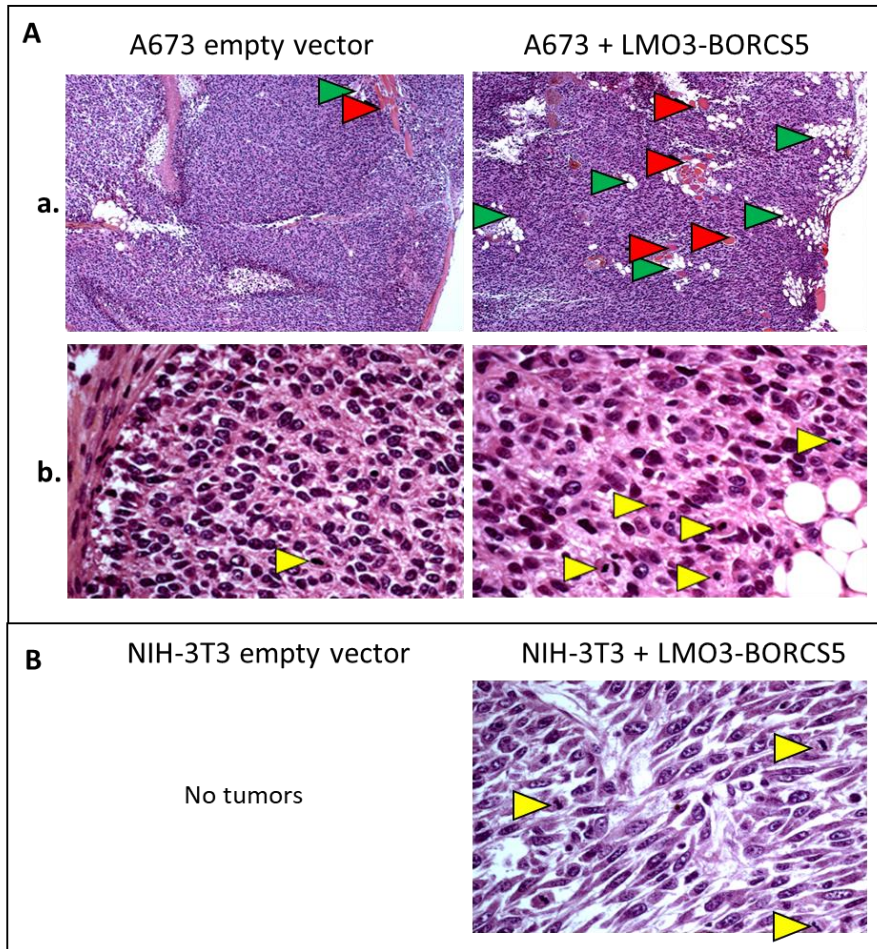


Figure S4: Histology of tumors obtained after subcutaneous injection of A673 and NIH-3T3 stably transfected with empty vector or *LMO3-BORCS5*.

Tumorigenicity tests were performed by injecting subcutaneously 10^6 cells into Nude mice. Tumors were measured until reaching a size of 1000 m^3 . **A.** Comparison of HES images in tumors obtained from A673 empty vector or A673+*LMO3-BORCS5* cell lines at x50 (**a**) or x400 (**b**) magnification. Histology revealed an increase of tumor infiltration in adipose (green arrows) and muscular (red arrows) tissues in presence of *LMO3-BORCS5* in A673 cells (upper panel). An increase in the number of cells in mitosis was also observed in presence of the fusion (yellow arrows) in A673 cells (lower panel). **B.** HES of tumors obtained from NIH-3T3+*LMO3-BORCS5* x400 magnification showing cells in mitotic division (yellow arrows).

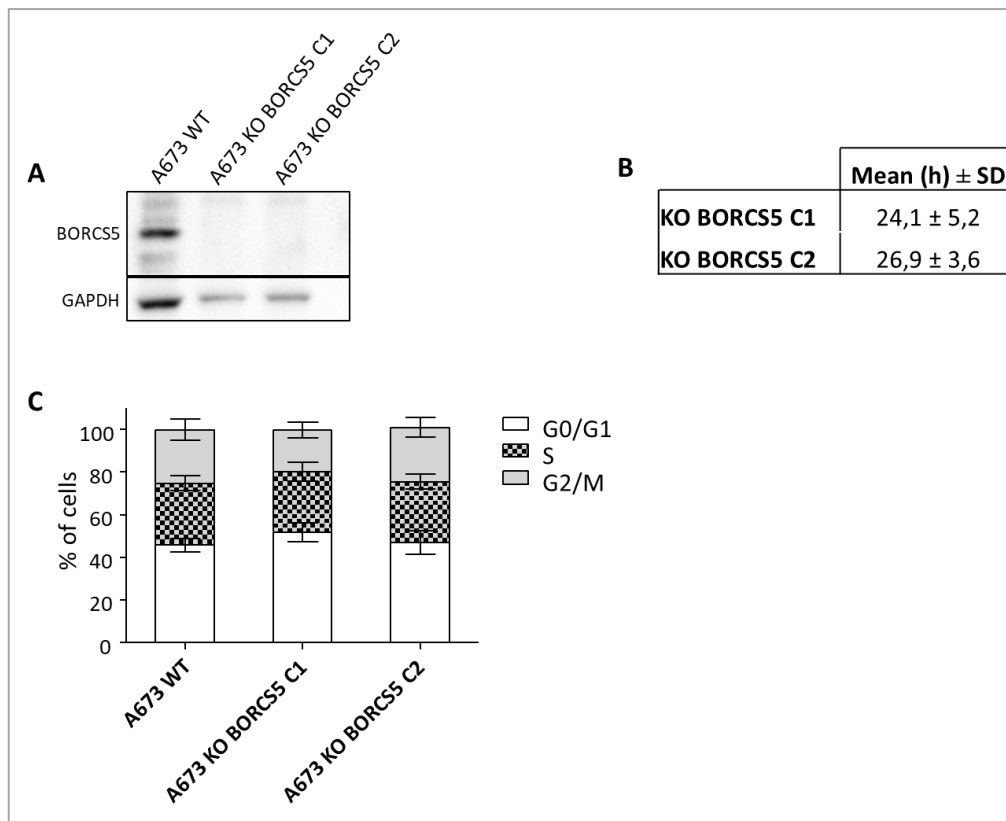


Figure S5. Comparison of the two clones obtained after KO of BORCS5 in A673 cells using CRIPR-Cas9 system. **A.** Western blot of BORCS5 protein in A673 KO BORCS5 clone 1 (C1) and A673 KO BORCS5 clone 2 (C2). **B.** Doubling time calculated according to an exponential regression equation, given as: $y = axe^{bx}$, via the formula: $t_{1/2} = \ln(2)/b$. Results are presented as Mean \pm SD in hours of at least two independent experiments. **C.** Cell cycle repartition after DNA staining with propidium iodide-containing buffer followed by FACS analysis. Results are presented as Mean \pm SD in % of cells in each cell cycle phase of at least two independent experiments.

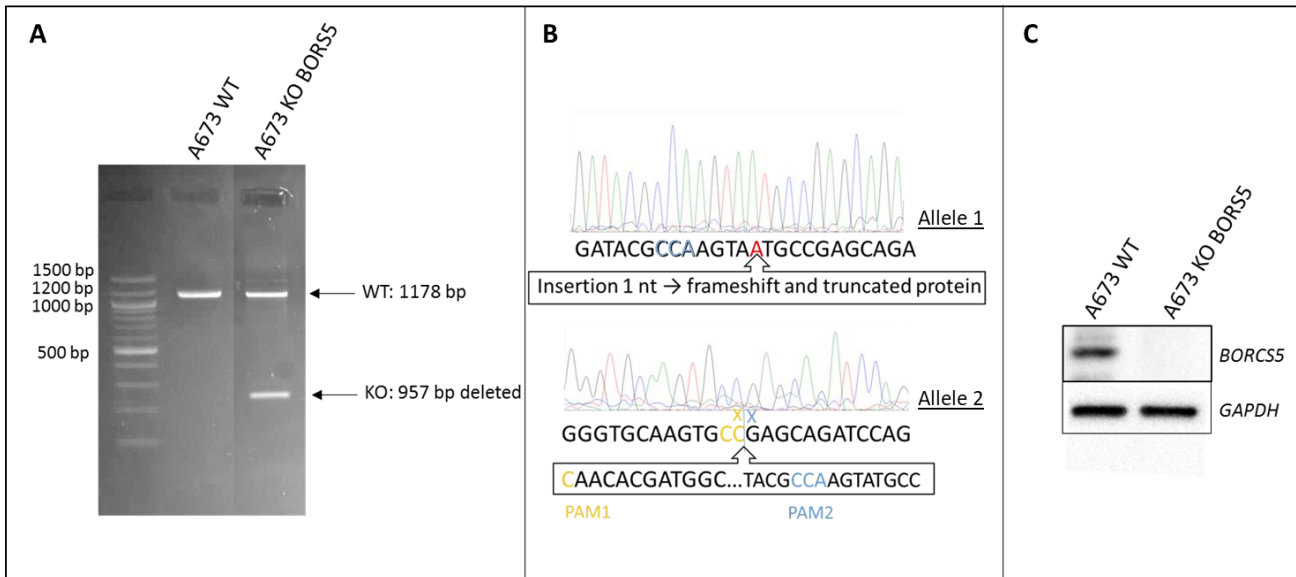


Figure S6: Knock-out of BORCS5 in A673 wild-type cells.

BORCS5 gene was inactivated using the CRISPR/Cas9 system targeting the last exon which contains the functional domain. After insertion of the 2 guides into pLentiCRISPR V2 (Addgene) containing either GFP or CFP respectively, A673 cells were co-transfected with both plasmids. After 27h, GFP and CFP-positive cells were selected using FACS Aria™ Fusion cell sorter and re-seeded to allow single colony formation. After 10 days, genomic DNA was extracted from individual colonies. **A.** Deletion of the target sequence was tested by PCR which resulted in a 950 bp smaller band in KO cells compared to WT cells. **B.** The KO was confirmed by Sanger sequencing which shows an insertion of one nucleotide on one allele (upper sequence) and a 957 bp deletion on the other allele (lower sequence) between two protospacer-adjacent motifs (PAM). **C.** Knock-out of BORCS5 was further confirmed by Western Blot on the total protein extracts of A673 WT and A673 KO BORCS5 cells. By using an anti-BORCS5 antibody, no BORCS5 protein was detected in the KO cells.

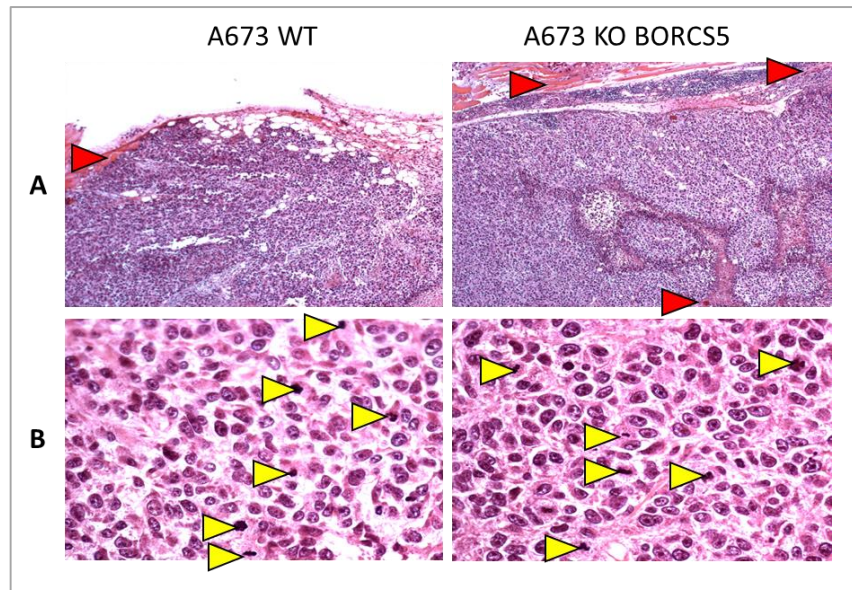


Figure S7: Histological slides of tumors derived from A673 cells +/- BORCS5.

Comparison of tumors obtained from A673 WT or A673 KO BORCS5 cell lines at x50 (**A.**) or x400 (**B.**) magnification. Histology of tumors revealed a slight increase of tumor infiltration in muscular (red arrows) tissues in absence of BORCS5 in A673 cells (upper panel). No difference regarding the number of cells in mitosis was observed between A673 WT and A673 KO BORCS5 cells (yellow arrows; lower panel).

Supplementary tables

Cell line	Pathology	Medium	Cell line provenance
A673	Ewing sarcoma	DMEM	ITCC
ARO	Anaplastic thyroid cancer	RPMI	ATCC
BHP 10-3	Papillary thyroid cancer	DMEM	ATCC
D283MED	Medulloblastoma	DMEM	ITCC
DAOY	Medulloblastoma	DMEM	ITCC
EW 7	Ewing sarcoma	RPMI	ITCC
GRU-1	Epitheloid sarcoma	DMEM + 1% non-essential amino acids	ITCC
HOS	Osteosarcoma	DMEM	ITCC
HOS 143B	Osteosarcoma	DMEM	ITCC
HT29	Colon adenocarcinoma	DMEM	ATCC
IGR-N91	Neuroblastoma	DMEM	Gustave Roussy
IGR-NB8	Neuroblastoma	DMEM	Gustave Roussy
IMR32	Neuroblastoma	RPMI	ITCC
IOR-OS-18	Osteosarcoma	DMEM	ITCC
LAN1	Neuroblastoma	RPMI	ITCC
LAN5	Neuroblastoma	RPMI	ITCC
MG-63	Osteosarcoma	DMEM	ITCC
RD	Embryonal rhabdomyosarcoma	RPMI	ITCC
RMS-1	Embryonal rhabdomyosarcoma	RPMI	ITCC
SAOS2	Osteosarcoma	DMEM	ITCC
SH-SY5Y	Neuroblastoma	RPMI	ITCC
SJ-NB-6	Neuroblastoma	DMEM	ITCC
SJ-NB-8	Neuroblastoma	DMEM	ITCC
SK-N-BE	Neuroblastoma	DMEM	ITCC
SK-N-MC	Ewing sarcoma	RPMI	ITCC
STA-ET-1	Ewing sarcoma	RPMI	ITCC
TC71	Ewing sarcoma	DMEM	ITCC
TPC-1	Papillary thyroid cancer	DMEM	ATCC
U-2OS	Osteosarcoma	DMEM	ITCC
VAESBJ	Epitheloid sarcoma	DMEM + 1% non-essential amino acids	ITCC
VCAP	Prostate cancer	DMEM	ATCC

Table S1: Cell lines used in the study with corresponding cancer types and culture medium.

All the following cell lines were tested for *LMO3-BORCS5* presence and for levels of expression of LMO3 WT and BORCS5 WT by RT-qPCR. RPMI = Roswell Park Memorial Institute medium; DMEM = Dulbecco's Modified Eagle's medium; ITCC = Innovative therapies for children with cancer; ATCC = American Type Culture Collection.

Genes	Organism	Forward (5'-3')	Reverse (5'-3')	Manufacturer
ABCC3	Human	CACCAACTCAGTCAAACGTGC	GCAAGACCATGAAAGCGACTC	Eurofins Genomics
FGF1	Human	ACACCGACGGGCTTTTATACG	CCCATTCTTCTTGAGGCCAAC	Eurofins Genomics
IGF2	Human	GTGGCATCGTTGAGGAGTG	CACGTCCCTCTCGGACTTG	Eurofins Genomics
PAX7	Human	TCCAAGATTCTTTGCCGCTAC	GGTCACAGTGCCCATCCTTC	Eurofins Genomics
PRKCA	Human	TGGACTTATCCATCAAGGGATGA	AGTGTGATCCATTCCGCAGAG	Eurofins Genomics
WNT6	Human	GGTGCAGAGTGCCAGTTC	CGTCTCCCGAATGTCCTGTT	Eurofins Genomics
BAX	Human	CCGAGAGGTCTTTTCCG	GCCTTGAGCACCAGTTTG	Eurofins Genomics
Bax	Mouse	GGAGCAGCTTGGGAGCG	AAAAGGCCCTGTCTTCATGA	Eurofins Genomics
NOXA	Human	CAGAGCTGGAAGTCGAGTGT	AGGAGTCCCCTCATGCAAGT	Eurofins Genomics
Noxa	Mouse	GGATTGTGATGTGATGAGAG	CCCCGTCTATTATTCTTGAC	Eurofins Genomics
P21	Human	GCCCCTCTACATCTTCTG	AATGCCCAGCACTCTTAGG	Eurofins Genomics
p21	Mouse	GTGGCCTTGTCGCTGTCTTG	CCCACGCCTATGGAATGGCT	Eurofins Genomics
P53	Human	CCTCCTCAGCATTTATCCG	CAAACACGCACCTCAAAGC	Eurofins Genomics
p53	Mouse	CAGCCAAGTCTGTTATGTGC	CGAAAAGTCTGCCTGTCTTC	Eurofins Genomics
GAPDH	Human / mouse	ATCCCATCACCATCTTCCAG	CCATCACGCCACAGTTCC	Eurofins Genomics
ALK	Human	n.a	n.a	AnyGenes
CUX2	Human	n.a	n.a	AnyGenes
FGF13	Human	n.a	n.a	AnyGenes
FRK	Human	n.a	n.a	AnyGenes
LZTS1	Human	n.a	n.a	AnyGenes
MMP10	Human	n.a	n.a	AnyGenes
NRP1	Human	n.a	n.a	AnyGenes
NTRK1	Human	n.a	n.a	AnyGenes
RHOJ	Human	n.a	n.a	AnyGenes
SEMA3C	Human	n.a	n.a	AnyGenes

Table S2: List of genes tested in RT-qPCR, sequences of primers used and corresponding manufacturer.

Primers provided from AnyGenes (Paris, France) were received as included in a personalized Human qPCR SignArrays® qPCR plate, therefore the sequence is only known by the manufacturer.

Fusion gene	Function of gene fused	Study	Sample ID	Cancer Type	Sex	Age (years)	Primary or Metastas	Primary tumor sit	TP53 status
LMO3-ETV6	Transcriptional repressor. Role in hematopoiesis and malignant transformation	MSK-IMPACT Clinical Sequencing Cohort (MSKCC, Nat Med 2017)	P-0008434-T01-1M5	Cutaneous Melanoma	F	Unspecified	Metastasis (lungs)	Skin	Mutation
B4GALNT3-LMO3	Transfers N-acetylgalactosamine (GaiNAC) onto glucosyl residues	Lung Adenocarcinoma (TCGA, PanCancer Atlas)	TCGA-55-A493-01	Lung Adenocarcinoma	F	54	Primary	Lungs	Mutation
LMO3-EP58	Regulation of axonal filopodia growth, stereocilia length, dendritic cell migration and cancer cell migration and invasion.	Breast Invasive Carcinoma (TCGA, PanCancer Atlas)	TCGA-A1-A05K-01	Breast Invasive Carcinoma, NOS	F	54	Primary	Breast	Mutation
FAM135A-LMO3	O-acyltransferase activity	Ovarian Serous Cystadenocarcinoma (TCGA, PanCancer Atlas)	TCGA-61-2102-01	Serous Ovarian Cancer	F	74	Primary	Ovary	Mutation
AEBP2-LMO3	DNA-binding transcriptional repressor, stimulates methylation of histone H3.	Sarcoma (TCGA, PanCancer Atlas)	TCGA-IV-A5VE-01	Leiomyosarcoma	F	60	Primary	Uterus	Deep deletion
BORCS5-ETV6	Transcriptional repressor. Role in hematopoiesis and malignant transformation	MSK-IMPACT Clinical Sequencing Cohort (MSKCC, Nat Med 2017)	P-0009946-T01-1M5	MSK-IMPACT Clinical Sequencing	F	-	Metastasis	Breast	Mutation
BORCS5-CCDC91	Regulation of membrane traffic through the trans-Golgi network	Breast Invasive Carcinoma (TCGA, PanCancer Atlas)	TCGA-A1-A05P-01	Breast Invasive Carcinoma (TCGA)	F	40	Primary	Breast	Mutation
CTNNA3-BORCS5	Involved in formation of stretch-resistant cell-cell adhesion complexes	Breast Invasive Carcinoma (TCGA, PanCancer Atlas)	TCGA-A7-A52X-01	Breast Invasive Carcinoma (TCGA)	F	48	Primary	Breast	WT
BORCS5-PTPRR	Histone deacetylase family	Lung Squamous Cell Carcinoma (TCGA, PanCancer Atlas)	TCGA-85-6175-01	Lung Squamous Cell Carcinoma (TCGA)	F	63	Primary	Lungs	WT but MDM2 amp
BORCS5-HDAC7	Binds to the GM-CSF promoter, to act as a repressor.	Prostate Adenocarcinoma (TCGA, PanCancer Atlas)	TCGA-V1-A8W5-01	Prostate Adenocarcinoma (TCGA)	M	56	Primary	Prostate	WT
YBX3-BORCS5	Cancer testis antigen gene family. Scaffold protein that structurally organizes MAPK and mediates JNK signaling	Prostate Adenocarcinoma (TCGA, PanCancer Atlas)	TCGA-G9-7525-01	Prostate Adenocarcinoma (TCGA)	M	64	Primary	Prostate	WT
BORCS5-SPAG9	Member of the cancer testis antigen gene family. Plays a role in tumor growth and development	Sarcoma (TCGA, PanCancer Atlas)	TCGA-FX-A2Q5-01	Sarcoma (TCGA, PanCancer Atlas)	F	61	Primary	Retroperitoneum	WT
KIAA1551-BORCS5	Unknown	Stomach Adenocarcinoma (TCGA, PanCancer Atlas)	TCGA-VG-A8PK-01	Stomach Adenocarcinoma (TCGA)	M	58	Primary	Stomach	Mutation
BICD1-BORCS5	Regulates Golgi-endoplasmic reticulum transport by recruiting the dynein-dynactin motor complex	Testicular Germ Cell Tumors (TCGA, PanCancer Atlas)	TCGA-ZG-AAKD-01	Testicular Germ Cell Tumors (TCGA)	M	18	Primary	Testes	WT

Table S3: Fusions involving LMO3 or BORCS5 genes with clinical characteristics of the corresponding patients.

Fusions detected in patients and involving LMO3 or BORCS5 genes, the corresponding study, cancer type, sex of the patient, age of the patient, tumor stage, primary tumor site and TP53 status were obtained from <http://www.cbioportal.org>.

TCGA = The Cancer Genome Atlas; MSKCC = Memorial Sloan Kettering Cancer Center; MSK-IMPACT = Memorial Sloan Kettering Integrated Mutation Profiling of Actionable Cancer Targets; NOS = Not Otherwise Specified; WT = Wild-Type; amp = amplification

C. Résultats complémentaires : interaction entre EWS-FLI et LMO3-BORCS5

1. Matériel et méthodes

Les cellules A673 exprimant stablement *LMO3-BORCS5* v1 ont été transfectées par des petits ARNs interférents (siRNA) afin d'évaluer l'effet respectif de l'inhibition de *EWS-FLI1* et de *LMO3-BORCS5* sur i) l'expression de *EWS-FLI1* et ii) l'expression de *LMO3-BORCS5* et ainsi déterminer l'interaction entre ces deux oncogènes de fusion.

Les cellules A673 précédemment décrites, issues de sarcome d'Ewing et stablement transfectées avec *LMO3-BORCS5* v1 ont été ensemencées à une concentration de 250000 cellules/puits dans des plaques six puits contenant le milieu DMEM supplémenté avec du sérum de veau fœtal 10%, de la pénicilline (100U/mL) et streptomycine (10µg/mL). Vingt-quatre heures plus tard, le milieu des cellules a été remplacé par 2mL de milieu OPTI-MEM (Life Technologies). La transfection a été réalisée en utilisant de la Lipofectamine 2000™ selon les recommandations du fabricant. Brièvement, dans du milieu OPTI-MEM seul (NT), le siRNA contrôle ne ciblant aucun gène, le siRNA dirigé contre *EWS-FLI1* ou le siRNA dirigé contre *LMO3-BORCS5* v1 (appelé ici LMO3-BORCS5) sont mélangés à 7,5µl de Lipofectamine 2000™ (1µG/mL) puis ajoutés aux cellules dans l'OPTI-MEM à une concentration finale de siRNA de 50nM/puits, et suivi d'une incubation à 37°C dans une atmosphère à 5% de CO₂ pendant 4h. A la fin de l'incubation, le milieu est aspiré et remplacé par du milieu DMEM complet.

Après 48h d'incubation à 37°C dans une atmosphère à 5% de CO₂, les cellules sont collectées et recueillies dans 350µl de tampon d'extraction d'ARN.

Chaque condition est effectuée en duplicat et l'expérience a été répétée trois fois indépendamment. Les séquences des siRNA utilisées sont répertoriées dans le tableau suivant :

siRNA	Séquence (5' → 3')
siRNA EWS-FLI1	AGCAGAACCCUUCUUAUGA[dT][dT]
siRNA LMO3-BORCS5	ACUAUCUGAGGGCUAUUGA[dT][dT]

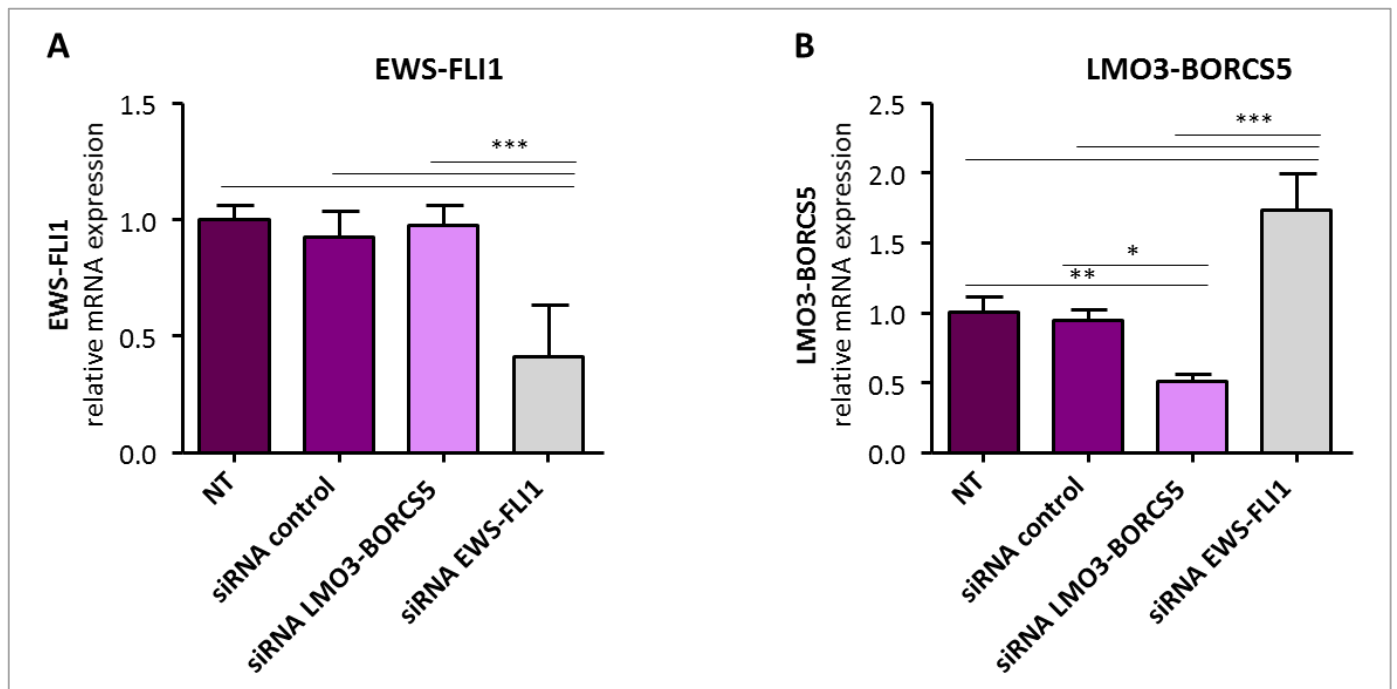
A partir des lysats cellulaires obtenus, les ARNs sont ensuite extraits puis convertis en ADNc et analysés par qPCR comme précédemment décrit dans la partie **Chapitre II, Résultats II-B**.

2. Résultats

Après analyse par RT-qPCR, nous avons pu observer que les deux siRNA utilisés dirigés respectivement contre *EWS-FLI1* et *LMO3-BORCS5* étaient spécifiques (**Figure 1**). Nous avons

également pu observer que l'inhibition de *LMO3-BORCS5* n'avait aucun impact sur l'expression de *EWS-FLI1* (**Figure 1A**). En revanche, l'inhibition de *EWS-FLI1* a conduit à une augmentation significative de l'expression de *LMO3-BORCS5* (**Figure 1B**).

Cependant, il est à noter que nous n'avons pas pu vérifier ce résultat au niveau protéique par Western Blot dû à l'absence d'anticorps pouvant reconnaître spécifiquement la fusion LMO3-BORCS5.



Evaluation de l'interaction entre LMO3-BORCS5 v1 et EWS-FLI1. Inhibition de l'expression des gènes *EWS-FLI1* ou *LMO3-BORCS5* v1 (*LMO3-BORCS5*) dans les cellules A673 exprimant stablement *LMO3-BORCS5* v1 par ARNs interférants (siRNA) et mesure de l'effet sur l'expression des gènes par RT-qPCR, L'expression a été déterminée par RT-qPCR, en utilisant la méthode de quantification comparative du $2^{-\Delta\Delta Ct}$ et après normalisation sur les niveaux d'expression de la GAPDH. Les résultats sont représentés par le niveau relatif d'ARNm de trois expériences indépendantes.

De manière étonnante, *EWS-FLI1* semble agir comme un inhibiteur de l'expression de *LMO3-BORCS5*, bien que les deux aient des fonctions oncogéniques. Comme évoqué précédemment, bien que *EWS-FLI1* soit l'altération à l'origine du développement et de l'apparition du sarcome d'Ewing, nous émettons l'hypothèse que *LMO3-BORCS5* pourrait être une deuxième altération oncogénique, dont le mode d'action moléculaire ne passe pas par une interaction avec *EWS-FLI1*, et permettant le maintien de la croissance tumorale et l'acquisition de résistances aux traitements conventionnels.

Dans un cadre plus général, et en se basant sur l'ensemble de nos résultats montrant que LMO3-BORCS5 augmente la prolifération, la résistance aux molécules antinéoplasiques et la tumorigénicité en présence ou en absence de EWS/FLI-1, il serait intéressant dans le futur d'évaluer l'impact de l'inhibition simultanée de ces deux oncogènes de fusion (EWS-FLI1 et LMO3-BORCS5) dans la progression tumorale.

DISCUSSION GENERALE

La découverte de gènes de fusion est d'un grand intérêt. Leur utilisation en clinique comme cibles thérapeutiques et comme biomarqueurs a été démontrée avec les exemples connus de BCR-ABL ou EWS-FLI1. L'étude des gènes des fusions participe également à l'amélioration de la compréhension des mécanismes biologiques menant au développement du cancer. Leur détection après séquençage de l'ARNm représente le premier pas vers leur découverte et il est donc essentiel de disposer d'outils fiables pour analyser les fusions, permettant de distinguer les événements réels des faux positifs, mais aussi les fusions responsables de l'oncogénicité de celles plus secondaires. Dans le but de découvrir de nouveaux transcrits de fusion oncogéniques, nous avons réalisé une étude rétrospective dans une cohorte de patients atteints de cancers pédiatriques solides, résistants ou en rechute inclus dans l'étude clinique MOSCATO-01 menée à Gustave Roussy (181).

La pipeline bioinformatique ChimComp développée au laboratoire a été créée pour détecter de manière fiable les transcrits de fusion dans une cohorte unique, en combinant trois outils bioinformatiques de détection de fusion. Après analyse, nous nous sommes principalement concentrés sur les candidats de fusion les plus « confiants », c'est à dire détectés par les trois outils simultanément. Ce groupe comprenait 118/1374 transcrits de fusion (8,6%). Afin de les classer selon leur oncogénicité, ces transcrits ont été analysés par le logiciel OncoFuse qui permet l'attribution d'un score oncogénique (oncoscore) selon la fonction connue des gènes et la littérature. Remarquablement, 100% des candidats sélectionnés avec des oncoscores élevés et intermédiaires ont été validés par RT-qPCR, alors que 19% des transcrits de fusion sélectionnés ayant un faible ou pas d'oncoscore ne l'ont pas été. Néanmoins, il est important de garder à l'esprit que la sélection de ces fusions pour la validation a été principalement motivée par des intérêts biologiques et il est très probable que d'autres fusions soient biologiquement existantes et validées, représentant des événements constitutifs, des réarrangements non-oncogéniques ou des fusions impliquant des gènes dont la fonction n'a pas encore été explorée. Cependant, le but de notre étude était de mettre en évidence des fusions pouvant représenter des cibles thérapeutiques qui pourraient être pertinentes pour la médecine personnalisée et contribuer à mieux comprendre les voies oncogéniques impliquées dans les cancers pédiatriques. Ce point souligne l'importance de connaître avec précision les questions scientifiques abordées et les objectifs d'une recherche de nouveaux transcrits de fusion.

Nous avons observé une tendance vers une validation préférentielle des fusions avec un haut oncoscore plutôt qu'avec un nombre de reads élevé. Cette observation est en ligne avec des

preuves émergentes que l'existence biologique d'un transcrit de fusion ne peut pas uniquement reposer sur le nombre de reads, ce nombre étant dépendant des paramètres d'alignement et de profondeur de séquençage (182,183)(183). L'utilisation de logiciels prédictifs tels que OncoFuse peut être bénéfique pour aider à sélectionner des fusions probablement intéressantes, mais seuls des tests fonctionnels valideront véritablement le potentiel oncogénique. De plus, nous avons pu noter certaines limites d'Oncofuse: i) l'incapacité à attribuer un score pour la majorité des transcrits de fusion (66%), ii) les variations de valeurs d'oncoscore en fonction de l'orientation des deux gènes dans la fusion et iii) la sous-évaluation des gènes suppresseurs de tumeurs.

Le séquençage de l'ARN et l'analyse bioinformatique complète des patients pédiatriques à la rechute inclus dans l'essai clinique MOSCATO-01 ont révélé des centaines de fusions au total. Aucune corrélation significative entre le nombre de fusions et de points de cassures génomiques n'a été trouvée, soutenant l'hypothèse que les fusions observées pourraient être principalement dues à des événements « équilibrés » (translocations réciproques par exemple) ou non-génomiques (au niveau transcriptionnel) (12,184). Ces études ont décrit une corrélation négative entre les fusions et les mutations, ce qui n'a pas été observé dans notre cohorte. Ces observations pourraient en partie s'expliquer par le statut particulier de notre cohorte, incluant des patients prétraités ayant acquis une succession d'événements génomiques et donc ayant des génomes plus chaotiques comparativement au diagnostic. De plus, à de tels stades tumoraux, la diversité clonale et l'hétérogénéité intratumorale sont plus élevées, ce qui peut biaiser l'analyse. (185).

Dans notre étude, plus d'un tiers des transcrits de fusion appartenant au groupe de « confiance » étaient prédits en phase de lecture, ce qui pourrait donner lieu à des protéines avec des fonctions oncogéniques nouvelles ou augmentées. La plupart de ces fusions présentent un faible oncoscore, néanmoins la fonction de nouvelles oncoprotéines chimères doit toujours être investiguée lors des études fonctionnelles. Des oncogènes et gènes suppresseurs de tumeurs ont été fréquemment retrouvés dans notre analyse. Une proportion considérable de transcrits de fusion contenant des gènes suppresseurs de tumeurs a été détectée, reflétant leur importance dans les cancers pédiatriques. En effet, il est estimé qu'environ 10% des cancers pédiatriques sont causés par des événements génétiques héréditaires, principalement l'inactivation de gènes suppresseurs de tumeurs (31,186). De manière générale, la contribution des fusions conduisant à l'inactivation des gènes suppresseurs de tumeurs pourrait être sous-estimée puisque la plupart des études se concentrent sur les fusions impliquant des oncogènes, plus facilement ciblables en thérapie.

Les fusions contenant des tyrosine-kinases ont été détectées à une faible incidence et majoritairement chez des patients atteints de tumeurs cérébrales. Alors que pour certains sous-groupes moléculaires de gliomes de haut grade, les aberrations génétiques affectant les récepteurs tyrosine kinases ont été bien définies, celles-ci n'ont pas été retrouvées lors de la caractérisation moléculaire des médulloblastomes (187). En ce qui concerne les épendymomes,

la surexpression d'EGFR est fréquente et des fusions impliquant l'EGFR ont été rapportées (188). Cependant, étant donné leur potentiel thérapeutique élevé, les fusions contenant des tyrosines kinases dans notre analyse ont été testées par RT-PCR suivi d'un séquençage de Sanger pour confirmer les points de jonction. Tous les transcrits ont été validés, mais seul *CCM2L-HCK* conserve le cadre de lecture. Ce transcrit a été retrouvé chez un patient atteint de médulloblastome, ne présentant pas d'altérations de *TP53* et des gènes des voies *WNT* et *SHH* au niveau tumoral. HCK appartient à la famille des protéines tyrosine kinases de type SRC et est capable d'augmenter la prolifération et la survie cellulaire notamment en s'associant à des récepteurs tyrosine kinases tels que EGFR, FGFR et PDGFR (189). La protéine HCK elle-même pourrait être directement ciblée par des inhibiteurs de kinases tels que le dasatinib, le bosutinib et le saracatinib. Le transcrit *JAK2-PTPRD* a été identifié chez un patient présentant un épépendymome mais conduit à une perte du domaine tyrosine kinase C-terminal de JAK2. Cependant, le transcrit conduit également à un changement du cadre de lecture de *PTPRD*, un gène dont le rôle en tant que suppresseur de tumeur a été reporté dans de nombreuses tumeurs, y compris le neuroblastome et le glioblastome. (190,191). Notre étude souligne à nouveau la faible implication des tyrosine-kinases dans les tumeurs pédiatriques.

En parallèle, nous avons extrait les fusions du groupe « de confiance » pouvant potentiellement être pertinentes sur le plan thérapeutique. Un tiers des patients analysés présentaient un transcrit de fusion avec potentiel d'association à une drogue connue ou impliqué dans la résistance aux agents antinéoplasiques. La présence de ces fusions au moment du diagnostic ou leur évolution au cours du temps et des traitements n'ont pas été étudiées, mais elles pourraient avoir un impact notable sur la progression de la maladie. De plus, les données appuyant les effets potentiels des fusions sont souvent précliniques et résultent de l'analyse d'autres tumeurs souvent « adultes » comme c'est le cas par exemple pour 2 transcrits impliquant le gène *MYB* confirmés chez 2 patients différents (*MYO16-IRS2* et *MYB-AHI1*). *MYB* a déjà été décrit dans la leucémie aiguë lymphoblastique et le cancer du sein (184,192) et son amplification a été associée *in vitro* à une sensibilité à l'inhibition d'IGF1R (193). Cependant l'inhibition d'IGF1R n'a pas pu être testée chez les patients correspondant en l'absence de drogue disponible.

En résumé, nous avons identifié très peu de fusions pouvant être associées directement à une molécule thérapeutique. Cette observation a également été faite par Gao et. *al*, où parmi les fusions provenant de plus de 9000 échantillons (33 types de cancers différents), seules 6% ont été identifiées comme ciblables par un agent thérapeutique existant (130). Le but ultime serait la génération d'agents spécifiques à chaque fusion. Ainsi, de nouvelles technologies émergent afin de cibler directement la jonction oncogénique, comme l'utilisation de petits ARNs interférents (siRNA). Ces siRNA permettent l'inhibition de gènes au niveau de l'ARNm, empêchant leur

traduction. Ce type de stratégie offre la possibilité de cibler spécifiquement les transcrits de fusion grâce à leur séquence de jonction unique, seulement exprimée dans les cellules tumorales, sans affecter l'expression d'autres gènes. Notre équipe a précédemment montré que l'inhibition des fusions responsables de la progression tumorale du cancer de la thyroïde (RET/PTC) ou de la prostate (TMPRSS2-ERG) par l'injection de siRNA vectorisés *in vivo* entraîne une diminution de la croissance tumorale (194–196). Dans le cadre de la fusion LMO3-BORCS5, son utilisation comme cible thérapeutique pourrait être envisageable au regard de nos résultats. Etant donné qu'aucun inhibiteur de LMO3 n'existe, les siRNA vectorisés pourraient être l'une des solutions afin de cibler cet oncogène. Etant donné l'expression et le rôle oncogénique de LMO3 dans les neuroblastomes, l'adénocarcinome pancréatique, il serait également intéressant de tester l'impact de certaines molécules antinéoplasiques utilisées dans le traitement de ces pathologies.

Dans la première partie de mon projet de thèse, nous avons donc établi une pipeline robuste appelée ChimComp pour détecter les transcrits de fusion chez les patients atteints de cancer, avec une fiabilité de 90,5% et proche de 100% lorsque des fusions avec un faible oncoscore sont écartées. Les raisons possibles de la non-validation de certains événements peuvent inclure des erreurs d'alignement, montrant que l'utilisation de plusieurs outils bioinformatiques simultanés (technique de « meta-caller ») ne permet pas d'atteindre une analyse parfaitement fiable. Un outil de visualisation de l'alignement des fusions semble être l'une des solutions pour appuyer l'existence réelle d'une fusion *in silico*. Notre méthode de détection de fusions a également atteint ses limites car i) les résultats doivent être comparés au tissu sain respectif de chaque patient et ii) les changements dans les cadres de lecture ne peuvent pas être prédits de manière fiable au niveau bioinformatique. Dans notre cohorte de patients, nous avons identifié plusieurs fusions pouvant potentiellement être des cibles thérapeutiques directes ou indirectes. La connaissance des différents effecteurs des voies des gènes impliqués dans ces fusions et la validation par des tests fonctionnels peuvent fournir de nouvelles options thérapeutiques pour les patients.

Dans le cas des tumeurs pédiatriques, la disponibilité des médicaments et l'effort de recherche principalement sur les tumeurs adultes peuvent limiter le traitement d'un patient au regard de son profil de gènes de fusion. Par conséquent, des essais cliniques basés sur la génomique tels que MOSCATO-01 (NCT01566019), couplés avec des tests fonctionnels en laboratoire sont nécessaires chez les patients pédiatriques et c'est ce qui est actuellement fait dans notre institut. Avec l'ouverture de MAPPYACTS (NCT02813135), un essai clinique multicentrique et international incluant seulement des enfants et des jeunes adultes en rechute, la méthode de détection développée pour identifier les fusions sera d'une grande aide. Nous espérons que l'élargissement de la cohorte améliorera notre étude préliminaire en augmentant notre base de données de fusions et permettra la détection d'événements récurrents responsables de la

progression tumorale et de la résistance dans les cancers pédiatriques. Etant donné que le cancer est une maladie très complexe et hétérogène, nous espérons qu'à l'avenir combiner la détection de la fusion avec la détection de variations du nombre de copies, le séquençage de l'exome entier (WES) et le séquençage du génome entier (WGS) permettra d'obtenir le profil moléculaire complet patient et par conséquent déterminer le traitement le plus approprié.

Parmi le groupe fusions « de confiance élevée », le transcrite de fusion *LMO3-BORCS5* a particulièrement retenu notre intérêt. Cette fusion a été détectée sous deux isoformes (*LMO3-BORCS5* v1 et v2), chez un patient atteint de sarcome d'Ewing, âgé de 20 ans et ayant rechuté après plusieurs traitements de chimiothérapie et de radiothérapie. Le transcrite *EWS-FLI1*, caractéristique du sarcome d'Ewing a également été détecté au niveau tumoral (85) tout comme des délétions des gènes *CDKN2A (P16)* et *CDKN2B (P15)*, impliqués dans la régulation du cycle cellulaire (197,198). En particulier, *CDKN2B* est connu pour promouvoir la stabilité de P53, qui contrairement à d'autres tumeurs, est rarement altéré chez les patients atteints de sarcome d'Ewing (~10%) (199).

Le variant 1 du transcrite de fusion *LMO3-BORCS5* a également été détecté au diagnostic chez le même patient, mais avec un niveau d'expression beaucoup plus faible suggérant que le transcrite de fusion i) n'est pas le résultat de réarrangements causés par la radiothérapie reçue par le patient et ii) était probablement exprimé dans un clone minoritaire, sélectionné au cours du temps et des traitements. En effet, il est connu que les cellules cancéreuses accumulent plusieurs modifications génétiques (y compris des fusions) dont la plupart sont des événements passagers. La sélection des clones conduisant à la transformation maligne dépend de multiples combinaisons d'événements oncogéniques (101). Au cours du processus d'évolution clonale de la tumeur, ces événements oncogéniques confèrent un avantage de sélection aux sous-clones résistants. Dans notre cas, il est possible que *EWS-FLI1* soit présent dans la cellule à l'origine de la tumeur initiale et donc exprimé dans tous les clones, mais que *LMO3-BORCS5* soit l'une des altérations complémentaires présente dans un sous-clone particulier, devenu dominant au cours des traitements par sa capacité de résistance.

Parallèlement, la présence du transcrite *LMO3-BORCS5* a été évaluée sur l'ARN de 31 lignées cellulaires différentes disponibles au laboratoire. 41% des lignées testées étaient positives pour le variant 1 et/ou le variant 2 de *LMO3-BORCS5*. Bien que la plupart des lignées cellulaires testées proviennent de cancers pédiatriques, les lignées positives présentent des origines histologiques diverses. Cependant, des informations sur la plupart de ces lignées cellulaires restent inconnues, notamment en ce qui concerne les traitements reçus et le stade d'évolution tumorale à partir duquel elles ont été établies, empêchant ainsi d'associer le statut de *LMO3-BORCS5* à un quelconque paramètre. Le niveau d'expression de *LMO3-BORCS5* dans ces lignées cellulaires était relativement faible en RT-qPCR, comme décrit précédemment dans la littérature pour la majorité

des nouvelles fusions découvertes (200). Ceci pourrait suggérer que *LMO3-BORCS5* est potentiellement présent à une plus forte incidence dans les différents tissus tumoraux et conférer un avantage à la résistance au traitement ou bien qu'il est faiblement exprimé dans des tissus sains et surexprimé au cours du processus de transformation. L'expression accrue de fusions constitutives apparemment «saines» au cours du développement tumoral du tissu associé a d'ailleurs déjà été décrite, comme *JAZF1-JJAZ1* dans l'endomètre ou *SLC45A3-ELK4* dans la prostate (128). L'évaluation de la présence d'un transcrit de fusion détecté dans la tumeur d'un patient, dans les tissus sains correspondants pourrait donc s'avérer d'une grande importance et fournir plus d'informations sur la biologie des fusions. L'expression faible de *LMO3-BORCS5* dans les lignées suggère également qu'une seconde altération inconnue pourrait être nécessaire pour l'expansion des cellules exprimant *LMO3-BORCS5* comme rapporté pour plusieurs autres fusions (201,202).

L'analyse du statut d'expression de *LMO3* et *BORCS5* « sauvages » a montré que *LMO3* était majoritairement exprimé dans les lignées de neuroblastome et de sarcome d'Ewing, un résultat concordant puisque *LMO3* est uniquement exprimé au niveau cérébral et que ces deux pathologies sont supposées provenir de précurseurs embryonnaires de la crête neurale ou neuroectodermique (202,203). Les lignées ayant le plus haut taux d'expression de *BORCS5* étaient issues de tumeurs osseuses (sarcome d'Ewing, ostéosarcome) et de médulloblastome. De ce fait, ces résultats renforcent l'idée que le sarcome d'Ewing ainsi que les tumeurs dérivées de précurseurs neurologiques ou osseux semblent être les pathologies dans lesquelles *LMO3* et *BORCS5* ont un rôle majeur, cependant d'autres études sont nécessaires.

LMO3 et *BORCS5* sont altérés au niveau génomique dans 1,5% des cas de cancer (source : www.cbioportal.org). Parmi ces altérations, les fusions représentent le troisième et le deuxième type d'altération des gènes *LMO3* et *BORCS5* respectivement. Bien que ce pourcentage soit faible, il est nécessaire de considérer que ces études regroupent des cas de cancer extrêmement différents, adultes et pédiatriques, et que la plupart des études n'analysent pas la présence de gènes de fusion. La quasi-totalité des fusions contenant *LMO3* (80%) sont également accompagnées d'une amplification de ce gène, et présentent des partenaires de fusion impliqués dans des voies de signalisation liées au cancer, renforçant le statut oncogénique de *LMO3* (143). De plus et de manière intéressante, 100% des patients présentant une fusion avec *LMO3* présentent également une délétion de *TP53*, dont la protéine est directement régulée par *LMO3* (149), et dont la voie de signalisation est également altérée dans la tumeur du patient inclus dans notre étude et exprimant *LMO3-BORCS5*. Un lien pourrait donc exister entre la voie P53 et *LMO3-BORCS5*.

L'analyse des séquences des deux isoformes du transcrit *LMO3-BORCS5* a montré différents points de jonction au niveau de *LMO3*, conduisant après traduction à une perte du domaine LIM2 et à une conservation totale (variant 1) ou à une perte de 24 acides aminés (variant 2) du domaine LIM1. De manière intéressante, la perte de domaines d'interaction protéiques dans des fusions est considérée comme un signe d'oncogénicité (204,205). Concernant *BORCS5*, le point de jonction est systématiquement le même entre les deux variants, situé avant l'exon 3 et provoque un décalage du cadre entraînant une rupture complète de *BORCS5* au niveau protéique, et la formation de 43 nouveaux acides aminés à la place. Plusieurs études ont rapporté l'existence de différents isoformes pour le même transcrit de fusion, mais leurs rôles respectifs distincts sont rarement évalués (206–208).

La compréhension des transcrits de fusion pour le développement de nouvelles thérapies est associée à la nécessité d'études de séquençage à grande échelle, à de nouvelles méthodes de détection efficaces *in silico*, combinées avec des études fonctionnelles (209,210). Pour cela, le deuxième objectif de mon projet doctoral était d'étudier le rôle de *LMO3-BORCS5* sur la tumorigénicité, incluant les différences entre les deux isoformes (v1 et v2), et l'impact de la perte de *BORCS5* sur la tumorigénicité. *LMO3-BORCS5* v2 n'a eu aucun effet sur la prolifération et le cycle cellulaire dans les lignées cellulaires A673 et NIH-3T3. En revanche, l'introduction de *LMO3-BORCS5* v1 (simplement appelé "*LMO3-BORCS5*" au cours de l'étude) dans ces mêmes lignées a conduit à une augmentation de la prolifération *in vitro* et à une augmentation (A673) ou induction (NIH-3T3) de la croissance tumorale et de l'invasion cellulaire *in vivo*.

Les effets de *LMO3-BORCS5* sur l'apoptose ont appuyé l'idée de l'existence d'un lien entre la voie *TP53* et *LMO3-BORCS5*. Ainsi, en présence de *LMO3-BORCS5* dans les cellules cancéreuses A673, l'expression des gènes liés à l'apoptose est inhibée et une diminution de la sensibilité à la vincristine et au SN-38 (métabolite actif de la camptothécine) est observée. Il convient de noter que la vincristine est l'une des molécules principales utilisée pour le traitement initial du sarcome d'Ewing et reçue par le patient exprimant *LMO3-BORCS5*. Ces résultats soulignent l'hypothèse antérieure de l'existence probable d'un clone minoritaire exprimant *LMO3-BORCS5* ayant résisté aux thérapies reçues et devenu le clone majoritaire à la rechute et mettent en avant le rôle critique de ce transcrit de fusion dans la résistance aux traitements, un aspect de l'impact des gènes de fusion rarement étudié et nécessitant plus d'investigations.

Les conséquences de l'introduction de *LMO3-BORCS5* sur le profil d'expression génique des cellules A673 ont également été évaluées par des analyses de microarray. La grande majorité des gènes différentiellement exprimés en présence de *LMO3-BORCS5* étaient inhibés. Parmi ceux-ci, la plupart étaient impliqués dans la différenciation cellulaire. Le processus de dédifférenciation est considéré comme l'une des caractéristiques principales des cellules cancéreuses, qui sont

bloquées aux premiers stades de la différenciation, ce qui entraîne un état de prolifération excessive (211). Les gènes surexprimés en présence de LMO3-BORCS5 étaient liés aux processus métaboliques et à la prolifération, en accord avec les résultats obtenus *in vitro*. Notamment, la kinase *FRK* est un des gènes dont l'expression était considérablement augmentée en présence de *LMO3-BORCS5*, ce qui a été validé par RT-qPCR. Cette tyrosine kinase conduit à une activation de la voie STAT3 et à l'inhibition de P53 et à des effets pro-tumorigènes dans certaines leucémies, ainsi que dans le carcinome hépatique et du pancréas. De plus, la *FRK* est surexprimée dans les cas de sarcomes (212). Cette protéine pourrait de ce fait représenter une cible potentielle indirecte dans les cellules exprimant *LMO3-BORCS5*.

Afin de comprendre l'impact du décalage du cadre de lecture de *BORCS5* après formation de la fusion *LMO3-BORCS5*, et plus généralement du rôle de *BORCS5* sur la tumorigénicité, nous l'avons inactivé dans les cellules A673 en utilisant le système CRISPR/Cas9. La perte de *BORCS5* a conduit à une diminution de la capacité de migration des cellules, comme décrit précédemment par Pu *et al.*, (174), et de la résistance aux traitements *in vitro*. Cependant, *in vivo*, une augmentation de la croissance tumorale et de la capacité d'invasion ont été observées. Ces derniers résultats suggèrent un rôle suppresseur de tumeur de *BORCS5*, compatible avec la perte observée dans la formation de la fusion *LMO3-BORCS5*, mais cependant ceci pourrait dépendre du type tumoral et du microenvironnement.

En conclusion à cette partie de l'étude, nous avons découvert un nouveau transcrit de fusion, *LMO3-BORCS5* jamais précédemment décrit et présent sous deux isoformes. La forme la plus longue de ce transcrit s'est avérée avoir un fort pouvoir oncogénique, agissant sur la prolifération cellulaire *in vitro* et *in vivo* et provoquant la transformation d'une lignée cellulaire non-tumorigène. Cette activité oncogénique semble être principalement due à une régulation négative des voies de différenciation, en adéquation avec le rôle des protéines LMO. En effet, la fusion *LMO3-BORCS5* conduit, après traduction, à une perte du domaine LIM2 de LMO3. Ce domaine s'avère important pour la fixation à d'autres protéines régulatrices, comme c'est le cas pour CIB, qui entraîne une relocalisation de LMO3 et une inhibition de la prolifération (148). De plus, Terano *et al.* ont montré que la mutation du domaine LIM2 de LMO2 entraîne l'inhibition de l'érythropoïèse dans les progéniteurs foetaux *in vivo*, en empêchant la liaison de LMO2 à GATA, soulignant ainsi l'importance de la préservation LIM2 pour la différenciation cellulaire (159). Etant donné que ce domaine LIM2 est perdu dans la fusion, cela pourrait expliquer les effets oncogéniques de *LMO3-BORCS5* et le profil génétique de dédifférenciation observé.

Plusieurs études ont précédemment rapporté que la préservation du cadre de lecture dans les fusions devaient être considérée comme un facteur déterminant pour l'oncogénicité (213,214), cependant nous montrons dans cette étude qu'une fusion dont le cadre de lecture n'est pas

conservé, peut avoir des propriétés fortement oncogéniques. Ici, la disruption de *BORCS5* après la formation de *LMO3-BORCS5* peut avoir deux rôles: i) la perte de *BORCS5* sur un allèle pourrait servir de second « hit » pour la perte du gène suppresseur de tumeur putatif ou ii) les 43 nouveaux acides aminés créés dans la partie C-terminale de la protéine de fusion après le domaine LIM1 pourraient permettre la stabilisation de la nouvelle protéine *LMO3* tronquée, empêchant sa dégradation (128).

Notre étude met l'accent sur le rôle critique des transcrits de fusion dans l'initiation tumorale, la progression et la résistance. Le cancer étant une maladie extrêmement variable, la présence récurrente d'une fusion semble être un critère insuffisant pour permettre de la considérer comme l'évènement responsable de la tumorigénicité. Ces résultats apportent un nouveau regard sur les altérations non-récurrentes et contredisent des études antérieures suggérant que les fusions uniques représentent uniquement des évènements secondaires (128,215). Cependant, *LMO3-BORCS5* a également été détecté à des faibles niveaux d'expression dans plusieurs lignées cellulaires. Par conséquent, plusieurs explications possibles pourraient expliquer l'absence de détection de cette fusion chez plusieurs patients. Premièrement, la prédominance d'une altération au sein d'une tumeur est liée à l'histoire du patient, au contexte génétique, au type de tumeur et aux combinaisons de traitements reçues. Dans notre étude, la combinaison des traitements précédemment reçus par le patient a probablement constitué le « cocktail » permettant l'évolution du clone exprimant *LMO3-BORCS5*. Deuxièmement, les cancers pédiatriques, et de manière encore plus importante les sarcomes d'Ewing, sont considérés comme des maladies rares avec une incidence très faible au sein de la population (216). Ainsi, la puissance statistique de notre cohorte pourrait être insuffisante pour pouvoir détecter une récurrence. Enfin, il est possible que *LMO3-BORCS5* soit présent dans d'autres tumeurs mais à des niveaux d'expression faibles, non détectables avec la couverture utilisée pour notre séquençage de l'ARN. Latysheva et al, ont par ailleurs suggéré que les altérations génomiques somatiques pertinentes pour la médecine personnalisée pourraient représenter moins de 5% des reads de séquençage, en raison de la contamination par des cellules non tumorales et la présence de plusieurs sous-populations de cellules tumorales. Afin d'augmenter la sensibilité, les protocoles doivent donc être adaptés avec une couverture de séquençage plus élevée (128).

Les fusions sont d'une importance capitale dans l'oncogenèse et leur étude approfondie pourrait expliquer certaines conséquences manquées par les analyses traditionnelles restreintes aux mutations et aux variations du nombre de copies (77). L'analyse des biopsies à la fois à la rechute et au diagnostic chez les patients peut contribuer à une meilleure compréhension de l'impact des transcrits de fusion dans l'évolution tumorale de chaque patient. L'arrivée de nouvelles techniques de séquençage améliorées à l'échelle de cellule unique contribuera à mieux

comprendre la place des fusions dans l'architecture clonale et à élargir considérablement les bases de données sur les altérations cancéreuses (121).

De manière générale, notre étude favorise la vision selon laquelle le cancer ne peut être traité qu'en considérant la pharmacogénomique et en appliquant une médecine personnalisée plutôt que de rechercher l'universalité, les transcrits de fusion représentant l'un des outils les plus remarquables en médecine de précision et leur détection devrait systématiquement être intégrée aux nouveaux essais cliniques basés sur l'analyse moléculaire des patients.

CONCLUSION GENERALE

Au cours de ma thèse j'ai pu :

- Contribuer au développement d'une « pipeline » robuste permettant de détecter avec une haute fiabilité des transcrits de fusions à partir des données RNA-seq.
- Analyser le paysage moléculaire des transcrits de fusion présents dans une cohorte de patients atteints de cancers pédiatriques solides en rechute
- Découvrir un nouveau transcrit de fusion *LMO3-BORCS5* jamais préalablement décrit et ayant un rôle central dans l'oncogenèse et dans la résistance à certains traitements antinéoplasiques.

La figure ci-dessous représente la conclusion des résultats obtenus lors de la préparation de ma thèse de sciences :

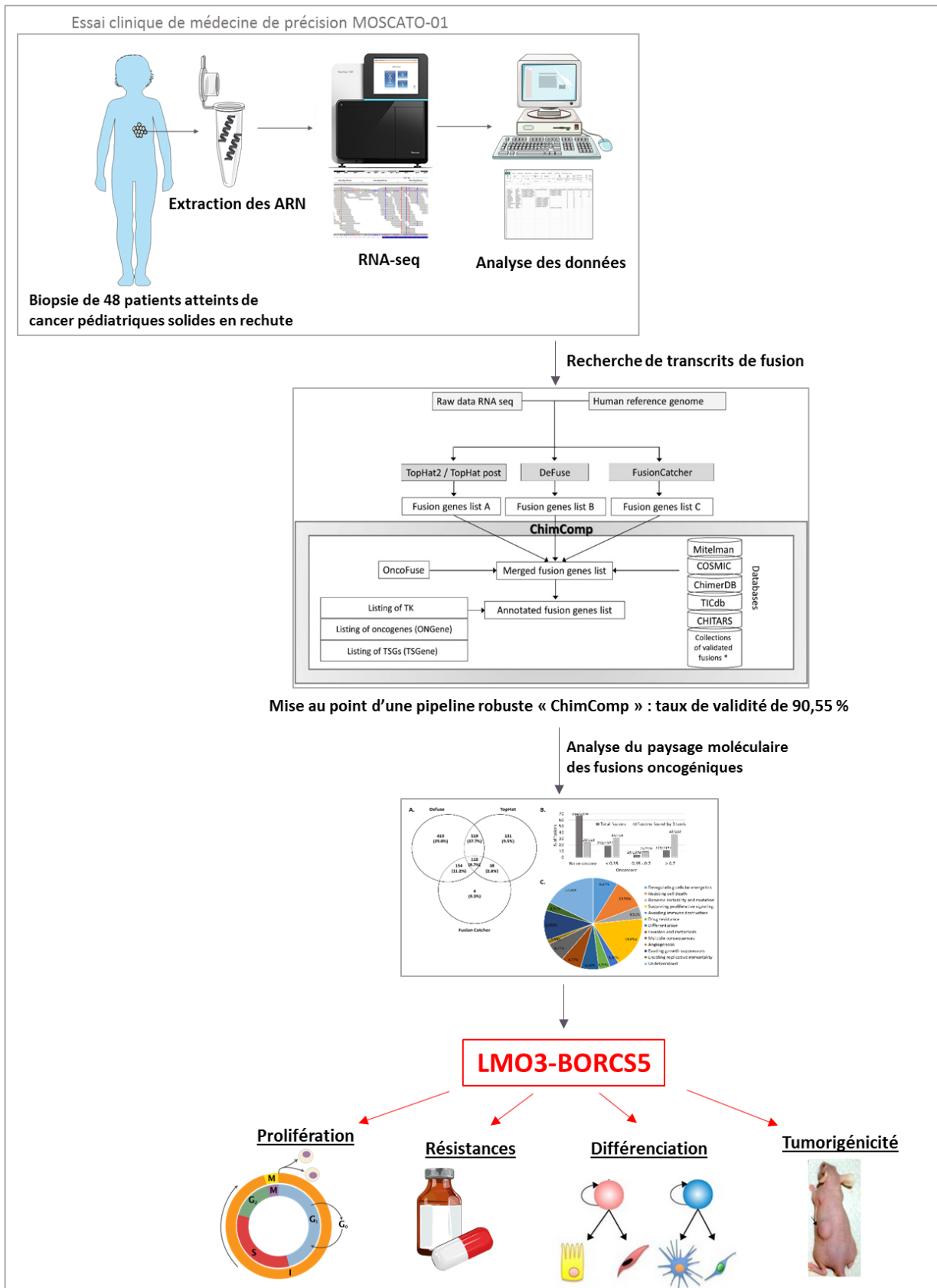


Schéma récapitulatif de l'ensemble des travaux de thèse.

REFERENCES

1. Druker BJ, Tamura S, Buchdunger E, Ohno S, Segal GM, Fanning S, et al. Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nat Med*. 1996 May;2(5):561–6.
2. Druker BJ, Talpaz M, Resta DJ, Peng B, Buchdunger E, Ford JM, et al. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N Engl J Med*. 2001 Apr 5;344(14):1031–7.
3. Gu T-L, Tothova Z, Scheijen B, Griffin JD, Gilliland DG, Sternberg DW. NPM-ALK fusion kinase of anaplastic large-cell lymphoma regulates survival and proliferative signaling through modulation of FOXO3a. *Blood*. 2004 Jun 15;103(12):4622–9.
4. Shaw AT, Yeap BY, Solomon BJ, Riely GJ, Gainor J, Engelman JA, et al. Effect of crizotinib on overall survival in patients with advanced non-small-cell lung cancer harbouring ALK gene rearrangement: a retrospective analysis. *Lancet Oncol*. 2011 Oct;12(11):1004–12.
5. Kumar A, Pathak P, Purkait S, Faruq M, Jha P, Mallick S, et al. Oncogenic KIAA1549-BRAF fusion with activation of the MAPK/ERK pathway in pediatric oligodendrogliomas. *Cancer Genet*. 2015 Mar;208(3):91–5.
6. Georger B, Schulte J, Zwaan CM, Casanova M, Fischer M, Moreno L, et al. Phase I study of ceritinib in pediatric patients (Pts) with malignancies harboring a genetic alteration in ALK (ALK+): Safety, pharmacokinetic (PK), and efficacy results. *J Clin Oncol* [Internet]. 2015 [cited 2016 Sep 28];33(suppl; abstr 10005). Available from: <http://meetinglibrary.asco.org/content/146923-156>
7. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature*. 2009 Apr 9;458(7239):719–24.
8. Di Stefano AL, Fucci A, Frattini V, Labussiere M, Mokhtari K, Zoppoli P, et al. Detection, Characterization, and Inhibition of FGFR-TACC Fusions in IDH Wild-type Glioma. *Clin Cancer Res*. 2015 Jul 15;21(14):3307–17.
9. Dupain C, Harttrampf AC, Urbinati G, Georger B, Massaad-Massade L. Relevance of Fusion Genes in Pediatric Cancers: Toward Precision Medicine. *Molecular Therapy - Nucleic Acids*. 2017 Mar 17;6:315–26.
10. Adamson PC. Improving the Outcome for Children With Cancer: Development of Targeted New Agents. *CA Cancer J Clin*. 2015;65(3):212–20.
11. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer Genome Landscapes. *Science*. 2013 Mar 29;339(6127):1546–58.
12. Mertens F, Johansson B, Fioretos T, Mitelman F. The emerging complexity of gene fusions in cancer. *Nature Reviews Cancer*. 2015;15(6):371–381.
13. Childhood Genomics Workshop NIH [Internet]. [cited 2015 Nov 20]. Available from:

<http://www.ncbi.nlm.nih.gov/gate1.inist.fr/pmc/articles/PMC3403905/pdf/1471-2164-13-147.pdf>

14. Massard C, Michiels S, Ferté C, Le Deley M-C, Lacroix L, Hollebecque A, et al. High-Throughput Genomics and Clinical Outcome in Hard-to-Treat Advanced Cancers: Results of the MOSCATO 01 Trial. *Cancer Discov.* 2017;7(6):586–95.
15. Harttrampf AC, Lacroix L, Deloger M, Deschamps F, Puget S, Auger N, et al. Molecular Screening for Cancer Treatment Optimization (MOSCATO-01) in Pediatric Patients: A Single-Institutional Prospective Molecular Stratification Trial. *Clin Cancer Res.* 2017 Oct 15;23(20):6101–12.
16. Steliarova-Foucher E, Colombet M, Ries LAG, Moreno F, Dolya A, Bray F, et al. International incidence of childhood cancer, 2001-10: a population-based registry study. *Lancet Oncol.* 2017 Jun;18(6):719–31.
17. Pediatric Oncology: General section [Internet]. [cited 2018 Jun 6]. Available from: <http://telemedicina.med.muni.cz/pediatric-oncology/index.php?pg=general-section>
18. International Childhood Cancer Day – 15 February - Resources - Resources [Internet]. [cited 2016 May 4]. Available from: <http://www.educationscotland.gov.uk/resources/i/internationalchildhoodcancerday.asp>
19. © Les cancers en France, Les Données INCa, édition 2015.
20. Globocan 2012 - Home [Internet]. [cited 2018 Jun 1]. Available from: <http://globocan.iarc.fr/Default.aspx>
21. European Cancer Information System [Internet]. [cited 2018 Jun 4]. Available from: <https://ecis.jrc.ec.europa.eu/>
22. Cancer Facts & Figures 2014, Special Section: Cancer in Children & Adolescents - acspc-041787.pdf [Internet]. [cited 2016 May 3]. Available from: <http://www.cancer.org/acs/groups/content/@research/documents/webcontent/acspc-041787.pdf>
23. Les spécificités des cancers des enfants et des adolescents - Cancérologie pédiatrique | Institut National Du Cancer [Internet]. [cited 2018 Jun 1]. Available from: <http://www.e-cancer.fr/Professionnels-de-sante/L-organisation-de-l-offre-de-soins/Cancerologie-pediatrique/Les-specificites-des-cancers-des-enfants-et-des-adolescents>
24. Childhood Cancer Epidemiology: Overview, Tools of Study, Cancer Incidence. 2015 Sep 25 [cited 2016 Jan 5]; Available from: <http://emedicine.medscape.com/article/989841-overview>
25. Vargas SO. Childhood Carcinoma. *Surgical Pathology Clinics.* 2010 Sep;3(3):689–710.
26. Boyd N, Dancey JE, Gilks CB, Huntsman DG. Rare cancers: a sea of opportunity. *Lancet Oncol.* 2016 Feb;17(2):e52–61.

27. Knudson AG. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci USA*. 1971 Apr;68(4):820-3.
28. Davidoff AM. Neuroblastoma. *Seminars in Pediatric Surgery*. 2012 Feb;21(1):2-14.
29. Gessi M, Giangaspero F, Lauriola L, Gardiman M, Scheithauer BW, Halliday W, et al. Embryonal tumors with abundant neuropil and true rosettes: a distinctive CNS primitive neuroectodermal tumor. *Am J Surg Pathol*. 2009 Feb;33(2):211-7.
30. Lee JS, Sanchez TR, Wootton-Gorges S. Malignant renal tumors in children. *Journal of Kidney Cancer and VHL*. 2015 May 9;2(3):84.
31. Zhang J, Walsh MF, Wu G, Edmonson MN, Gruber TA, Easton J, et al. Germline Mutations in Predisposition Genes in Pediatric Cancer. *N Engl J Med*. 2015 Dec 10;373(24):2336-46.
32. Kaatsch P. Epidemiology of childhood cancer. *Cancer Treat Rev*. 2010 Jun;36(4):277-85.
33. Downing JR, Wilson RK, Zhang J, Mardis ER, Pui C-H, Ding L, et al. The Pediatric Cancer Genome Project. *Nature Genetics*. 2012 Jun;44(6):619-22.
34. Marshall GM, Carter DR, Cheung BB, Liu T, Mateos MK, Meyerowitz JG, et al. The prenatal origins of cancer. *Nat Rev Cancer*. 2014 Apr;14(4):277-89.
35. Chen X, Pappo A, Dyer MA. Pediatric solid tumor genomics and developmental plasticity. *Oncogene*. 2015 Oct;34(41):5207-15.
36. Gröbner SN, Worst BC, Weischenfeldt J, Buchhalter I, Kleinheinz K, Rudneva VA, et al. The landscape of genomic alterations across childhood cancers. *Nature*. 2018 Mar;555(7696):321-7.
37. Wu G, Broniscer A, McEachron TA, Lu C, Paugh BS, Becksfort J, et al. Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nat Genet*. 2012 Jan 29;44(3):251-3.
38. Huether R, Dong L, Chen X, Wu G, Parker M, Wei L, et al. The landscape of somatic mutations in epigenetic regulators across 1000 pediatric cancer genomes. *Nat Commun*. 2014 Apr 8;5:3630.
39. Walz AL, Ooms A, Gadd S, Gerhard DS, Smith MA, Guidry Auvil JM, et al. Recurrent DGCR8, DROSHA, and SIX homeodomain mutations in favorable histology Wilms tumors. *Cancer Cell*. 2015 Feb 9;27(2):286-97.
40. Kohsaka S, Shukla N, Ameer N, Ito T, Ng CKY, Wang L, et al. A recurrent neomorphic mutation in MYOD1 defines a clinically aggressive subset of embryonal rhabdomyosarcoma associated with PI3K-AKT pathway mutations. *Nat Genet*. 2014 Jun;46(6):595-600.
41. Szuhai K, de Jong D, Leung WY, Fletcher CDM, Hogendoorn PCW. Transactivating mutation of the MYOD1 gene is a frequent event in adult spindle cell rhabdomyosarcoma. *J Pathol*. 2014 Feb;232(3):300-7.

42. Oeffinger KC, Mertens AC, Sklar CA, Kawashima T, Hudson MM, Meadows AT, et al. Chronic Health Conditions in Adult Survivors of Childhood Cancer. *New England Journal of Medicine*. 2006 Oct 12;355(15):1572–82.
43. Meadows AT, Friedman DL, Neglia JP, Mertens AC, Donaldson SS, Stovall M, et al. Second neoplasms in survivors of childhood cancer: findings from the Childhood Cancer Survivor Study cohort. *J Clin Oncol*. 2009 May 10;27(14):2356–62.
44. Robison LL. Treatment-associated subsequent neoplasms among long-term survivors of childhood cancer: the experience of the Childhood Cancer Survivor Study. *Pediatr Radiol*. 2009 Feb;39 Suppl 1:S32-37.
45. Late Effects of Treatment for Childhood Cancer [Internet]. National Cancer Institute. [cited 2018 Jun 26]. Available from: <https://www.cancer.gov/types/childhood-cancers/late-effects-hp-pdq>
46. Childhood Cancer Survivor Study [Internet]. [cited 2018 Jun 26]. Available from: <https://ccss.stjude.org/>
47. Childhood Cancer Survivor Study: An Overview [Internet]. National Cancer Institute. [cited 2018 Jun 26]. Available from: <https://www.cancer.gov/types/childhood-cancers/ccss>
48. Robison LL, Armstrong GT, Boice JD, Chow EJ, Davies SM, Donaldson SS, et al. The Childhood Cancer Survivor Study: a National Cancer Institute-supported resource for outcome and intervention research. *J Clin Oncol*. 2009 May 10;27(14):2308–18.
49. Bernstein ML. Targeted therapy in pediatric and adolescent oncology. *Cancer*. 2011 May 15;117(10 Suppl):2268–74.
50. Médecine de précision : les thérapies ciblées - Les thérapies ciblées | Institut National Du Cancer [Internet]. [cited 2018 Jun 26]. Available from: <http://www.e-cancer.fr/Professionnels-de-sante/Les-therapies-ciblees/Medecine-de-precision-les-therapies-ciblees>
51. Pearson ADJ, Pfister SM, Baruchel A, Bourquin J-P, Casanova M, Chesler L, et al. From class waivers to precision medicine in paediatric oncology. *Lancet Oncol*. 2017 Jul;18(7):e394–404.
52. Steliarova-Foucher E, Stiller C, Lacour B, Kaatsch P. International Classification of Childhood Cancer, third edition. *Cancer*. 2005 Apr 1;103(7):1457–67.
53. Frühwald MC, Rutkowski S. Tumors of the Central Nervous System in Children and Adolescents. *Dtsch Arztebl Int*. 2011 Jun;108(22):390–7.
54. Cancers de l'enfant : le neuroblastome [Internet]. Institut Curie. [cited 2018 Jun 9]. Available from: <https://curie.fr/dossier-pedagogique/cancers-de-lenfant-le-neuroblastome>
55. Michon J, Couturier J, Brisse H, Edeline V, Schleiermacher G, Vielh P, et al. Neuroblastomes : rôles respectifs de la clinique, de l'imagerie, de la pathologie et de la biologie dans les

décisions thérapeutiques. *Médecine thérapeutique / Pédiatrie*. 2003 May 1;6(3):172-84.

56. Moroz V, Machin D, Faldum A, Hero B, Iehara T, Mosseri V, et al. Changes over three decades in outcome and the prognostic influence of age-at-diagnosis in young patients with neuroblastoma: a report from the International Neuroblastoma Risk Group Project. *Eur J Cancer*. 2011 Mar;47(4):561-71.
57. Brodeur GM, Seeger RC, Schwab M, Varmus HE, Bishop JM. Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage. *Science*. 1984 Jun 8;224(4653):1121-4.
58. Diskin SJ, Hou C, Glessner JT, Attiyeh EF, Laudenslager M, Bosse K, et al. Copy number variation at 1q21.1 associated with neuroblastoma. *Nature*. 2009 Jun 18;459(7249):987-91.
59. Mossé YP, Laudenslager M, Longo L, Cole KA, Wood A, Attiyeh EF, et al. Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature*. 2008 Oct 16;455(7215):930-5.
60. Szychot E, Apps J, Pritchard-Jones K. Wilms' tumor: biology, diagnosis and treatment. *Translational Pediatrics*. 2014;3(1):13.
61. Knudson AG. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci USA*. 1971 Apr;68(4):820-3.
62. Pandey AN. Retinoblastoma: An overview. *Saudi J Ophthalmol*. 2014 Oct;28(4):310-5.
63. Chintagumpala M, Chevez-Barrios P, Paysse EA, Plon SE, Hurwitz R. Retinoblastoma: review of current management. *Oncologist*. 2007 Oct;12(10):1237-46.
64. Horwitz M, Dufour C, Leblond P, Bourdeaut F, Faure-Contier C, Bertozzi A-I, et al. Embryonal tumors with multilayered rosettes in children: the SFCE experience. *Childs Nerv Syst*. 2016 Feb;32(2):299-305.
65. Korshunov A, Sturm D, Ryzhova M, Hovestadt V, Gessi M, Jones DTW, et al. Embryonal tumor with abundant neuropil and true rosettes (ETANTR), ependymoblastoma, and medulloepithelioma share molecular similarity and comprise a single clinicopathological entity. *Acta Neuropathol*. 2014 Aug;128(2):279-89.
66. Eberhart CG, Brat DJ, Cohen KJ, Burger PC. Pediatric neuroblastic brain tumors containing abundant neuropil and true rosettes. *Pediatr Dev Pathol*. 2000 Aug;3(4):346-52.
67. Biswas A, Kashyap L, Kakkar A, Sarkar C, Julka P. Atypical teratoid/rhabdoid tumors: challenges and search for solutions. *Cancer Management and Research*. 2016 Sep;Volume 8:115-25.
68. Ho DM, Hsu CY, Wong TT, Ting LT, Chiang H. Atypical teratoid/rhabdoid tumor of the central nervous system: a comparative study with primitive neuroectodermal tumor/medulloblastoma. *Acta Neuropathol*. 2000 May;99(5):482-8.
69. Hasselblatt M, Gesk S, Oyen F, Rossi S, Viscardi E, Giangaspero F, et al. Nonsense mutation

- and inactivation of SMARCA4 (BRG1) in an atypical teratoid/rhabdoid tumor showing retained SMARCB1 (INI1) expression. *Am J Surg Pathol*. 2011 Jun;35(6):933–5.
70. Biswas A, Goyal S, Puri T, Das P, Sarkar C, Julka PK, et al. Atypical teratoid rhabdoid tumor of the brain: case series and review of literature. *Childs Nerv Syst*. 2009 Nov;25(11):1495–500.
 71. Burningham Z, Hashibe M, Spector L, Schiffman JD. The Epidemiology of Sarcoma. *Clinical Sarcoma Research*. 2012;2(1):14.
 72. Durfee RA, Mohammed M, Luu HH. Review of Osteosarcoma and Current Management. *Rheumatology and Therapy*. 2016 Dec;3(2):221–43.
 73. Osteosarcoma and MFH of Bone Treatment [Internet]. National Cancer Institute. [cited 2018 Jun 12]. Available from: <https://www.cancer.gov/types/bone/hp/osteosarcoma-treatment-pdq>
 74. Geller DS, Gorlick R. Osteosarcoma: A Review of Diagnosis, Management, and Treatment Strategies. *Osteosarcoma*. :14.
 75. Perry JA, Kiezun A, Tonzi P, Van Allen EM, Carter SL, Baca SC, et al. Complementary genomic approaches highlight the PI3K/mTOR pathway as a common vulnerability in osteosarcoma. *Proc Natl Acad Sci USA*. 2014 Dec 23;111(51):E5564-5573.
 76. Chen X, Bahrami A, Pappo A, Easton J, Dalton J, Hedlund E, et al. Recurrent Somatic Structural Variations Contribute to Tumorigenesis in Pediatric Osteosarcoma. *Cell Rep*. 2014 Apr 10;7(1):104–12.
 77. Lorenz S, Barøy T, Sun J, Nome T, Vodák D, Bryne J-C, et al. Unscrambling the genomic chaos of osteosarcoma reveals extensive transcript fusion, recurrent rearrangements and frequent novel TP53 aberrations. *Oncotarget*. 2015 Dec 11;7(5):5273–88.
 78. Biswas B, Bakhshi S. Management of Ewing sarcoma family of tumors: Current scenario and unmet need. *World J Orthop*. 2016 Sep 18;7(9):527–38.
 79. Lahl M, Fisher VL, Laschinger K. Ewing’s sarcoma family of tumors: an overview from diagnosis to survivorship. *Clin J Oncol Nurs*. 2008 Feb;12(1):89–97.
 80. Iwamoto Y. Diagnosis and treatment of Ewing’s sarcoma. *Jpn J Clin Oncol*. 2007 Feb;37(2):79–89.
 81. Ewing Sarcoma Treatment [Internet]. National Cancer Institute. [cited 2018 Jun 13]. Available from: <https://www.cancer.gov/types/bone/hp/ewing-treatment-pdq>
 82. Balamuth NJ, Womer RB. Ewing’s sarcoma. *Lancet Oncol*. 2010 Feb;11(2):184–92.
 83. Lin PP, Wang Y, Lozano G. Mesenchymal Stem Cells and the Origin of Ewing’s Sarcoma. *Sarcoma*. 2011;2011.
 84. Bernstein M, Kovar H, Paulussen M, Randall RL, Schuck A, Teot LA, et al. Ewing’s Sarcoma Family of Tumors: Current Management. *The Oncologist*. 2006 Jan 5;11(5):503–19.

85. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, et al. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature*. 1992 Sep 10;359(6391):162–5.
86. Kovar H. Downstream EWS/FLI1 - upstream Ewing's sarcoma. *Genome Medicine*. 2010 Jan 28;2:8.
87. Bernstein M, Kovar H, Paulussen M, Randall RL, Schuck A, Teot LA, et al. Ewing's sarcoma family of tumors: current management. *The oncologist*. 2006 May 1;11(5):503–19.
88. Childhood Rhabdomyosarcoma Treatment [Internet]. National Cancer Institute. [cited 2018 Jun 14]. Available from: <https://www.cancer.gov/types/soft-tissue-sarcoma/hp/rhabdomyosarcoma-treatment-pdq>
89. Meyer WH. Rhabdomyosarcoma. *Holland-Frei Cancer Medicine 6th edition* [Internet]. 2003 [cited 2018 Jun 14]; Available from: <https://www.ncbi.nlm.nih.gov/books/NBK13154/>
90. de Souza RR, Oliveira ID, del Giudice Paniago M, Yaoita FHK, Caran EMM, Macedo CRPD, et al. Investigation of IGF2, Hedgehog and fusion gene expression profiles in pediatric sarcomas. *Growth Hormone & IGF Research*. 2014 Aug;24(4):130–6.
91. Mahajan P, Leavey PJ, Galindo RL. PAX genes in childhood oncogenesis: developmental biology gone awry? *Oncogene*. 2015 May;34(21):2681–9.
92. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet*. 2016 17;17(6):333–51.
93. Tucker T, Marra M, Friedman JM. Massively Parallel Sequencing: The Next Big Thing in Genetic Medicine. *The American Journal of Human Genetics*. 2009 Aug;85(2):142–54.
94. An Introduction to Next-Generation Sequencing Technology. :16.
95. Joly Y, Dove ES, Knoppers BM, Bobrow M, Chalmers D. Data sharing in the post-genomic world: the experience of the International Cancer Genome Consortium (ICGC) Data Access Compliance Office (DACO). *PLoS Comput Biol*. 2012;8(7):e1002549.
96. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)*. 2015;19(1A):A68-77.
97. International Cancer Genome Consortium, Hudson TJ, Anderson W, Artez A, Barker AD, Bell C, et al. International network of cancer genome projects. *Nature*. 2010 Apr 15;464(7291):993–8.
98. Dancy JE, Bedard PL, Onetto N, Hudson TJ. The Genetic Basis for Cancer Treatment Decisions. *Cell*. 2012 Feb 3;148(3):409–20.
99. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med*. 2004 Aug;10(8):789–99.
100. Tsimberidou A-M. Targeted therapy in cancer. *Cancer Chemother Pharmacol*. 2015

Dec;76(6):1113–32.

101. Sabaawy HE. Genetic Heterogeneity and Clonal Evolution of Tumor Cells and their Impact on Precision Cancer Medicine. *J Leuk (Los Angel)*. 2013 Nov 18;1(4):1000124.
102. Siu LL, Conley BA, Boerner S, LoRusso PM. Next-Generation Sequencing to Guide Clinical Trials. *Clin Cancer Res*. 2015 Oct 15;21(20):4536–44.
103. Tsimberidou A-M. Initiative for Molecular Profiling and Advanced Cancer Therapy and challenges in the implementation of precision medicine. *Current Problems in Cancer*. 2017 May 1;41(3):176–81.
104. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005 Sep 15;437(7057):376–80.
105. Bennett ST, Barnes C, Cox A, Davies L, Brown C. Toward the 1,000 dollars human genome. *Pharmacogenomics*. 2005 Jun;6(4):373–82.
106. Bentley DR. Whole-genome re-sequencing. *Curr Opin Genet Dev*. 2006 Dec;16(6):545–52.
107. Yohe S, Thyagarajan B. Review of Clinical Next-Generation Sequencing. *Arch Pathol Lab Med*. 2017 Nov;141(11):1544–57.
108. Rizzo JM, Buck MJ. Key Principles and Clinical Applications of “Next-Generation” DNA Sequencing. *Cancer Prevention Research*. 2012 Jul 1;5(7):887–900.
109. Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP. Sequencing depth and coverage: key considerations in genomic analyses. *Nature Reviews Genetics*. 2014 Feb;15(2):121–32.
110. Warr A, Robert C, Hume D, Archibald A, Deeb N, Watson M. Exome Sequencing: Current and Future Perspectives. *G3: Genes|Genomes|Genetics*. 2015 Aug;5(8):1543–50.
111. Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. *Nature Reviews Genetics*. 2010 Oct;11(10):685–96.
112. NCI Dictionary of Cancer Terms [Internet]. National Cancer Institute. [cited 2018 Jul 10]. Available from: <https://www.cancer.gov/publications/dictionaries/cancer-terms>
113. Karapetis CS, Khambata-Ford S, Jonker DJ, O’Callaghan CJ, Tu D, Tebbutt NC, et al. K-ras mutations and benefit from cetuximab in advanced colorectal cancer. *N Engl J Med*. 2008 Oct 23;359(17):1757–65.
114. Romond EH, Perez EA, Bryant J, Suman VJ, Geyer CE, Davidson NE, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med*. 2005 Oct 20;353(16):1673–84.
115. Tsimberidou AM, Eggermont AMM, Schilsky RL. Precision cancer medicine: the future is now, only better. *Am Soc Clin Oncol Educ Book*. 2014;61–9.
116. Forrest SJ, Georger B, Janeway KA. Precision medicine in pediatric oncology. *Curr Opin*

- Pediatr. 2018 Feb;30(1):17–24.
117. Roychowdhury S, Chinnaiyan AM. Translating cancer genomes and transcriptomes for precision oncology. *CA Cancer J Clin.* 2016 Feb;66(1):75–88.
 118. Mody RJ, Prensner JR, Everett J, Parsons DW, Chinnaiyan AM. Precision medicine in pediatric oncology: Lessons learned and next steps. *Pediatr Blood Cancer.* 2017 Mar;64(3).
 119. Lackner MR, Wilson TR, Settleman J. Mechanisms of acquired resistance to targeted cancer therapies. *Future Oncol.* 2012 Aug;8(8):999–1014.
 120. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature.* 2013 Sep 19;501(7467):338–45.
 121. Hedlund E, Deng Q. Single-cell RNA sequencing: Technical advancements and biological applications. *Molecular Aspects of Medicine.* 2018 Feb 1;59:36–46.
 122. Kurihara S, Ueda Y, Onitake Y, Sueda T, Ohta E, Morihara N, et al. Circulating free DNA as non-invasive diagnostic biomarker for childhood solid tumors. *J Pediatr Surg.* 2015 Dec;50(12):2094–7.
 123. Kumar-Sinha C, Kalyana-Sundaram S, Chinnaiyan AM. Landscape of gene fusions in epithelial cancers: seq and ye shall find. *Genome Med.* 2015 Dec 18;7:129.
 124. Kanchi KL, Johnson KJ, Lu C, McLellan MD, Leiserson MDM, Wendl MC, et al. Integrated analysis of germline and somatic variants in ovarian cancer. *Nature Communications.* 2014 Jan 22;5:3156.
 125. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. *Nature.* 2013 Oct;502(7471):333–9.
 126. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature.* 2013 Jul 11;499(7457):214–8.
 127. Gao Q, Liang W-W, Foltz SM, Mutharasu G, Jayasinghe RG, Cao S, et al. Driver Fusions and Their Implications in the Development and Treatment of Human Cancers. *Cell Reports.* 2018 Apr 3;23(1):227–238.e3.
 128. Latysheva NS, Babu MM. Discovering and understanding oncogenic gene fusions through data intensive computational approaches. *Nucleic Acids Res.* 2016 Jun 2;44(10):4487–503.
 129. Zhang H, Lin W, Kannan K, Luo L, Li J, Chao P-W, et al. Aberrant chimeric RNA GOLM1-MAK10 encoding a secreted fusion protein as a molecular signature for human esophageal squamous cell carcinoma. *Oncotarget.* 2013 Nov 2;4(11):2135–43.
 130. Gao Q, Liang W-W, Foltz SM, Mutharasu G, Jayasinghe RG, Cao S, et al. Driver Fusions and Their Implications in the Development and Treatment of Human Cancers. *Cell Rep.* 2018 Apr 3;23(1):227–238.e3.

131. Kuehner JN, Pearson EL, Moore C. Unravelling the means to an end: RNA polymerase II transcription termination. *Nat Rev Mol Cell Biol.* 2011 May;12(5):283–94.
132. De Braekeleer E, Douet-Guilbert N, De Braekeleer M. RARA fusion genes in acute promyelocytic leukemia: a review. *Expert Rev Hematol.* 2014 Jun;7(3):347–57.
133. Cachia D, Wani K, Penas-Prado M, Olar A, McCutcheon IE, Benjamin RS, et al. C11orf95–RELA fusion present in a primary supratentorial ependymoma and recurrent sarcoma. *Brain Tumor Pathol.* 2015 Apr;32(2):105–11.
134. Graham RP, Jin L, Knutson DL, Kloft-Nelson SM, Greipp PT, Waldburger N, et al. DNAJB1-PRKACA is specific for fibrolamellar carcinoma. *Mod Pathol.* 2015 Jun;28(6):822–9.
135. PANAGOPOULOS I, GORUNOVA L, BJERKEHAGEN B, BOYE K, HEIM S. Chromosome aberrations and HEY1-NCOA2 fusion gene in a mesenchymal chondrosarcoma. *Oncol Rep.* 2014 Jul;32(1):40–4.
136. Arpinati M, Tolomelli G, Bochicchio MT, Castagnetti F, Amabile M, Bandini G, et al. Molecular monitoring of BCR-ABL transcripts after allogeneic stem cell transplantation for chronic myeloid leukemia. *Biol Blood Marrow Transplant.* 2013 May;19(5):735–40.
137. Kalinova M, Krskova L, Brizova H, Kabickova E, Kepak T, Kodet R. Quantitative PCR detection of NPM/ALK fusion gene and CD30 gene expression in patients with anaplastic large cell lymphoma--residual disease monitoring and a correlation with the disease status. *Leuk Res.* 2008 Jan;32(1):25–32.
138. Han S, Brenner JC, Sabolch A, Jackson W, Speers C, Wilder-Romans K, et al. Targeted radiosensitization of ETS fusion-positive prostate cancer through PARP1 inhibition. *Neoplasia.* 2013 Oct;15(10):1207–17.
139. Brown LM, Hanna DT, Khaw SL, Ekert PG. Dysregulation of BCL-2 family proteins by leukemia fusion genes. *Journal of Biological Chemistry.* 2017 Sep 1;292(35):14325–33.
140. Hartmaier RJ, Trabucco SE, Priedigkeit N, Chung JH, Parachoniak CA, Vanden Borre P, et al. Recurrent hyperactive ESR1 fusion proteins in endocrine therapy-resistant breast cancer. *Ann Oncol.* 2018 Apr 1;29(4):872–80.
141. Kumar S, Razzaq SK, Vo AD, Gautam M, Li H. Identifying fusion transcripts using next generation sequencing. *Wiley Interdiscip Rev RNA.* 2016;7(6):811–23.
142. Carrara M, Beccuti M, Lazzarato F, Cavallo F, Cordero F, Donatelli S, et al. State-of-the-Art Fusion-Finder Algorithms Sensitivity and Specificity. *Biomed Res Int [Internet].* 2013 [cited 2016 Mar 9];2013. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3595110/>
143. Sang M, Ma L, Sang M, Zhou X, Gao W, Geng C. LIM-domain-only proteins: multifunctional nuclear transcription coregulators that interacts with diverse proteins. *Mol Biol Rep.* 2014 Feb;41(2):1067–73.

144. Matthews JM, Lester K, Joseph S, Curtis DJ. LIM-domain-only proteins in cancer. *Nature Reviews Cancer*. 2013 Jan 10;13(2):111–22.
145. Bao J, Talmage DA, Role LW, Gautier J. Regulation of neurogenesis by interactions between HEN1 and neuronal LMO proteins. *Development*. 2000 Jan;127(2):425–35.
146. Isogai E, Ohira M, Ozaki T, Oba S, Nakamura Y, Nakagawara A. Oncogenic LMO3 collaborates with HEN2 to enhance neuroblastoma cell growth through transactivation of Mash1. *PLoS ONE*. 2011;6(5):e19297.
147. Aoyama M, Ozaki T, Inuzuka H, Tomotsune D, Hirato J, Okamoto Y, et al. LMO3 interacts with neuronal transcription factor, HEN2, and acts as an oncogene in neuroblastoma. *Cancer Res*. 2005 Jun 1;65(11):4587–97.
148. Hui L, Ji C, Hui B, Lv T, Ha X, Yang J, et al. The oncoprotein LMO3 interacts with calcium- and integrin-binding protein CIB. *Brain Research*. 2009 Apr 10;1265:24–9.
149. Larsen S, Yokochi T, Isogai E, Nakamura Y, Ozaki T, Nakagawara A. LMO3 interacts with p53 and inhibits its transcriptional activity. *Biochem Biophys Res Commun*. 2010 Feb 12;392(3):252–7.
150. Song Y-F, Hong J-F, Liu D-L, Lin Q-A, Lan X-P, Lai G-X. miR-630 targets LMO3 to regulate cell growth and metastasis in lung cancer. *Am J Transl Res*. 2015;7(7):1271–9.
151. Qiu Y-S, Jiang N-N, Zhou Y, Yu K-Y, Gong H-Y, Liao G-J. LMO3 promotes gastric cancer cell invasion and proliferation through Akt-mTOR and Akt-GSK3 β signaling. *Int J Mol Med*. 2018 Feb 8;
152. Valge-Archer V, Forster A, Rabbitts TH. The LMO1 and LDB1 proteins interact in human T cell acute leukaemia with the chromosomal translocation t(11;14)(p15;q11). *Oncogene*. 1998 Dec 17;17(24):3199–202.
153. Boehm T, Baer R, Lavenir I, Forster A, Waters JJ, Nacheva E, et al. The mechanism of chromosomal translocation t(11;14) involving the T-cell receptor C delta locus on human chromosome 14q11 and a transcribed region of chromosome 11p15. *EMBO J*. 1988 Feb;7(2):385–94.
154. Boehm T, Foroni L, Kennedy M, Rabbitts TH. The rhombotin gene belongs to a class of transcriptional regulators with a potential novel protein dimerisation motif. *Oncogene*. 1990 Jul;5(7):1103–5.
155. Van Vlierberghe P, van Grotel M, Beverloo HB, Lee C, Helgason T, Buijs-Gladdines J, et al. The cryptic chromosomal deletion del(11)(p12p13) as a new activation mechanism of LMO2 in pediatric T-cell acute lymphoblastic leukemia. *Blood*. 2006 Nov 15;108(10):3520–9.
156. Boehm T, Foroni L, Kaneko Y, Perutz MF, Rabbitts TH. The rhombotin family of cysteine-rich LIM-domain oncogenes: distinct members are involved in T-cell translocations to human chromosomes 11p15 and 11p13. *Proc Natl Acad Sci USA*. 1991 May 15;88(10):4367–71.

157. Hacein-Bey-Abina S, Von Kalle C, Schmidt M, McCormack MP, Wulffraat N, Leboulch P, et al. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science*. 2003 Oct 17;302(5644):415–9.
158. Wadman IA, Osada H, Grütz GG, Agulnick AD, Westphal H, Forster A, et al. The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J*. 1997 Jun 2;16(11):3145–57.
159. Terano T, Zhong Y, Toyokuni S, Hiai H, Yamada Y. Transcriptional control of fetal liver hematopoiesis: dominant negative effect of the overexpression of the LIM domain mutants of LMO2. *Exp Hematol*. 2005 Jun;33(6):641–51.
160. Wang K, Diskin SJ, Zhang H, Attiyeh EF, Winter C, Hou C, et al. Integrative genomics identifies LMO1 as a neuroblastoma oncogene. *Nature*. 2011 Jan 13;469(7329):216–20.
161. Kenny DA, Jurata LW, Saga Y, Gill GN. Identification and characterization of LMO4, an LMO gene with a novel pattern of expression during embryogenesis. *Proc Natl Acad Sci USA*. 1998 Sep 15;95(19):11257–62.
162. Racevskis J, Dill A, Sparano JA, Ruan H. Molecular cloning of LMO41, a new human LIM domain gene. *Biochim Biophys Acta*. 1999 Apr 14;1445(1):148–53.
163. Visvader JE, Venter D, Hahm K, Santamaria M, Sum EY, O'Reilly L, et al. The LIM domain gene LMO4 inhibits differentiation of mammary epithelial cells in vitro and is overexpressed in breast cancer. *Proc Natl Acad Sci USA*. 2001 Dec 4;98(25):14452–7.
164. Wang N, Lin KK, Lu Z, Lam KS, Newton R, Xu X, et al. The LIM-only factor LMO4 regulates expression of the BMP7 gene through an HDAC2-dependent mechanism, and controls cell proliferation and apoptosis of mammary epithelial cells. *Oncogene*. 2007 Sep 27;26(44):6431–41.
165. Montañez-Wiscovich ME, Seachrist DD, Landis MD, Visvader J, Andersen B, Keri RA. LMO4 is an essential mediator of ErbB2/HER2/Neu-induced breast cancer cell cycle progression. *Oncogene*. 2009 Oct 15;28(41):3608–18.
166. Sum EYM, Peng B, Yu X, Chen J, Byrne J, Lindeman GJ, et al. The LIM domain protein LMO4 interacts with the cofactor CtIP and the tumor suppressor BRCA1 and inhibits BRCA1 activity. *J Biol Chem*. 2002 Mar 8;277(10):7849–56.
167. Manetopoulos C, Hansson A, Karlsson J, Jönsson J-I, Axelson H. The LIM-only protein LMO4 modulates the transcriptional activity of HEN1. *Biochem Biophys Res Commun*. 2003 Aug 8;307(4):891–9.
168. Hahm K, Sum EYM, Fujiwara Y, Lindeman GJ, Visvader JE, Orkin SH. Defective neural tube closure and anteroposterior patterning in mice lacking the LIM protein LMO4 or its interacting partner Deaf-1. *Mol Cell Biol*. 2004 Mar;24(5):2074–82.
169. Lee S-K, Jurata LW, Nowak R, Lettieri K, Kenny DA, Pfaff SL, et al. The LIM domain-only

- protein LMO4 is required for neural tube closure. *Mol Cell Neurosci*. 2005 Feb;28(2):205–14.
170. Murphy NC, Scarlett CJ, Kench JG, Sum EYM, Segara D, Colvin EK, et al. Expression of LMO4 and outcome in pancreatic ductal adenocarcinoma. *Br J Cancer*. 2008 Feb 12;98(3):537–41.
171. Yu J, Ohuchida K, Nakata K, Mizumoto K, Cui L, Fujita H, et al. LIM only 4 is overexpressed in late stage pancreas cancer. *Mol Cancer*. 2008 Dec 22;7:93.
172. Nam C-H, Lobato MN, Appert A, Drynan LF, Tanaka T, Rabbitts TH. An antibody inhibitor of the LMO2-protein complex blocks its normal and tumorigenic functions. *Oncogene*. 2008 Aug 21;27(36):4962–8.
173. Appert A, Nam C-H, Lobato N, Priego E, Miguel RN, Blundell T, et al. Targeting LMO2 with a peptide aptamer establishes a necessary function in overt T-cell neoplasia. *Cancer Res*. 2009 Jun 1;69(11):4784–90.
174. Pu J, Schindler C, Jia R, Jarnik M, Backlund P, Bonifacino JS. BORC, a Multisubunit Complex that Regulates Lysosome Positioning. *Developmental Cell*. 2015 Apr 20;33(2):176–88.
175. Guardia CM, Farías GG, Jia R, Pu J, Bonifacino JS. BORC Functions Upstream of Kinesins 1 and 3 to Coordinate Regional Movement of Lysosomes along Different Microtubule Tracks. *Cell Rep*. 2016 Nov 15;17(8):1950–61.
176. Farías GG, Guardia CM, De Pace R, Britt DJ, Bonifacino JS. BORC/kinesin-1 ensemble drives polarized transport of lysosomes into the axon. *Proc Natl Acad Sci USA*. 2017 04;114(14):E2955–64.
177. Langemeyer L, Ungermann C. BORC and BLOC-1: Shared Subunits in Trafficking Complexes. *Developmental Cell*. 2015 Apr 20;33(2):121–2.
178. Saftig P, Klumperman J. Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function. *Nat Rev Mol Cell Biol*. 2009 Sep;10(9):623–35.
179. Niwa S, Tao L, Lu SY, Liew GM, Feng W, Nachury MV, et al. BORC Regulates the Axonal Transport of Synaptic Vesicle Precursors by Activating ARL-8. *Curr Biol*. 2017 Sep 11;27(17):2569-2578.e4.
180. Montpetit A, Boily G, Sinnett D. A detailed transcriptional map of the chromosome 12p12 tumour suppressor locus. *Eur J Hum Genet*. 2002 Jan;10(1):62–71.
181. Harttrampf AC, Lacroix L, Deloger M, Deschamps F, Puget S, Auger N, et al. MOlecular Screening for CAncer Treatment Optimization (MOSCATO-01) in pediatric patients: A single institutional prospective molecular stratification trial. *Clin Cancer Res*. 2017 Jul 21;
182. Huang Z, Jones DTW, Wu Y, Lichter P, Zapatka M. confFuse: High-Confidence Fusion Gene Detection across Tumor Entities. *Front Genet*. 2017;8:137.
183. Shah N, Lankerovich M, Lee H, Yoon J-G, Schroeder B, Foltz G. Exploration of the gene fusion landscape of glioblastoma using transcriptome sequencing and copy number data. *BMC*

genomics. 2013;14(1):818.

184. Yoshihara K, Wang Q, Torres-Garcia W, Zheng S, Vegesna R, Kim H, et al. The landscape and therapeutic relevance of cancer-associated transcript fusions. *Oncogene*. 2015 Sep 10;34(37):4845–54.
185. Burrell RA, Swanton C. Tumour heterogeneity and the evolution of polyclonal drug resistance. *Molecular Oncology*. 2014 Sep 12;8(6):1095–111.
186. Multi-Dimensional ClinOmics for Precision Therapy of Children and Adolescent Young Adults with Relapsed and Refractory Cancer: A report from the Center for Cancer Research | Clinical Cancer Research [Internet]. [cited 2018 Feb 12]. Available from: <http://clincancerres.aacrjournals.org/content/early/2016/03/18/1078-0432.CCR-15-2717>
187. Pajtler KW, Witt H, Sill M, Jones DTW, Hovestadt V, Kratochwil F, et al. Molecular Classification of Ependymal Tumors across All CNS Compartments, Histopathological Grades, and Age Groups. *Cancer Cell*. 2015 May 11;27(5):728–43.
188. Servidei T, Meco D, Muto V, Bruselles A, Ciolfi A, Trivieri N, et al. Novel SEC61G-EGFR Fusion Gene in Pediatric Ependymomas Discovered by Clonal Expansion of Stem Cells in Absence of Exogenous Mitogens. *Cancer Res*. 2017 Nov 1;77(21):5860–72.
189. Poh AR, O'Donoghue RJJ, Ernst M. Hematopoietic cell kinase (HCK) as a therapeutic target in immune and cancer cells. *Oncotarget*. 2015 Jun 30;6(18):15752–71.
190. Meehan M, Parthasarathi L, Moran N, Jefferies CA, Foley N, Lazzari E, et al. Protein tyrosine phosphatase receptor delta acts as a neuroblastoma tumor suppressor by destabilizing the aurora kinase A oncogene. *Mol Cancer*. 2012 Feb 5;11:6.
191. Ortiz B, Fabius AWM, Wu WH, Pedraza A, Brennan CW, Schultz N, et al. Loss of the tyrosine phosphatase PTPRD leads to aberrant STAT3 activation and promotes gliomagenesis. *Proc Natl Acad Sci USA*. 2014 Jun 3;111(22):8149–54.
192. Atak ZK, Gianfelici V, Hulselmans G, De Keersmaecker K, Devasia AG, Geerdens E, et al. Comprehensive analysis of transcriptome variation uncovers known and novel driver events in T-cell acute lymphoblastic leukemia. *PLoS Genet*. 2013;9(12):e1003997.
193. Pavlicek A, Lira ME, Lee NV, Ching KA, Ye J, Cao J, et al. Molecular predictors of sensitivity to the insulin-like growth factor 1 receptor inhibitor Figitumumab (CP-751,871). *Mol Cancer Ther*. 2013 Dec;12(12):2929–39.
194. Ali HM, Urbinati G, Chapuis H, Desmaele D, Bertrand J-R, Couvreur P, et al. Effects of siRNA on RET/PTC3 junction oncogene in papillary thyroid carcinoma: from molecular and cellular studies to preclinical investigations. *PLoS ONE*. 2014;9(4):e95964.
195. Ali HM, Maksimenko A, Urbinati G, Chapuis H, Raouane M, Desmaële D, et al. Effects of Silencing the *RET/PTC1* Oncogene in Papillary Thyroid Carcinoma by siRNA-Squalene Nanoparticles With and Without Fusogenic Companion GALA-Cholesterol. *Thyroid*. 2014

Feb;24(2):327–38.

196. Urbinati G, Ali HM, Rousseau Q, Chapuis H, Desmaële D, Couvreur P, et al. Antineoplastic Effects of siRNA against TMPRSS2-ERG Junction Oncogene in Prostate Cancer. *PLoS ONE*. 2015;10(5):e0125277.
197. Zhao R, Choi BY, Lee M-H, Bode AM, Dong Z. Implications of Genetic and Epigenetic Alterations of CDKN2A (p16INK4a) in Cancer. *EBioMedicine*. 2016 Jun 1;8:30–9.
198. Hannon GJ, Beach D. p15INK4B is a potential effector of TGF-beta-induced cell cycle arrest. *Nature*. 1994 Sep 15;371(6494):257–61.
199. Neilsen PM, Pishas KI, Callen DF, Thomas DM. Targeting the p53 Pathway in Ewing Sarcoma [Internet]. *Sarcoma*. 2011 [cited 2018 May 19]. Available from: <https://www.hindawi.com/journals/sarcoma/2011/746939/>
200. Frenkel-Morgenstern M, Lacroix V, Ezkurdia I, Levin Y, Gabashvili A, Prilusky J, et al. Chimeras taking shape: potential functions of proteins encoded by chimeric RNA transcripts. *Genome Res*. 2012 Jul;22(7):1231–42.
201. Keller C, Arenkiel BR, Coffin CM, El-Bardeesy N, DePinho RA, Capecchi MR. Alveolar rhabdomyosarcomas in conditional Pax3:Fkhr mice: cooperativity of Ink4a/ARF and Trp53 loss of function. *Genes Dev*. 2004 Nov 1;18(21):2614–26.
202. Lin PP, Pandey MK, Jin F, Xiong S, Deavers M, Parant JM, et al. EWS-FLI1 induces developmental abnormalities and accelerates sarcoma formation in a transgenic mouse model. *Cancer research*. 2008;68(21):8968–8975.
203. Tsubota S, Kadomatsu K. Origin and mechanism of neuroblastoma. *Oncoscience*. 2017 Sep 21;4(7–8):70–2.
204. Shugay M, Ortiz de Mendivil I, Vizmanos JL, Novo FJ. Genomic hallmarks of genes involved in chromosomal translocations in hematological cancer. *PLoS Comput Biol*. 2012;8(12):e1002797.
205. Shugay M, Ortiz de Mendivil I, Vizmanos JL, Novo FJ. Oncofuse: a computational framework for the prediction of the oncogenic potential of gene fusions. *Bioinformatics*. 2013 Oct 15;29(20):2539–46.
206. Ghasemian Sorbeni F, Montazersaheb S, Ansarin A, Esfahani A, Rezamand A, Sakhinia E. Molecular analysis of more than 140 gene fusion variants and aberrant activation of EVI1 and TLX1 in hematological malignancies. *Ann Hematol*. 2017 Oct;96(10):1605–23.
207. Hillion J, Mecucci C, Aventin A, Leroux D, Wlodarska I, Van Den Berghe H, et al. A variant translocation t(2;18) in follicular lymphoma involves the 5' end of bcl-2 and Ig kappa light chain gene. *Oncogene*. 1991 Jan;6(1):169–72.
208. Cai W, Li W, Ren S, Zheng L, Li X, Zhou C. Coexistence of three variants involving two different fusion partners of ROS1 including a novel variant of ROS1 fusions in lung adenocarcinoma:

- a case report. *J Thorac Oncol*. 2014 Jun;9(6):e43-46.
209. Robinson DR, Kalyana-Sundaram S, Wu Y-M, Shankar S, Cao X, Ateeq B, et al. Functionally recurrent rearrangements of the MAST kinase and Notch gene families in breast cancer. *Nat Med*. 2011 Nov 20;17(12):1646-51.
210. Tognon CE, Somasiri AM, Evdokimova VE, Trigo G, Uy EE, Melnyk N, et al. ETV6-NTRK3-mediated breast epithelial cell transformation is blocked by targeting the IGF1R signaling pathway. *Cancer Res*. 2011 Feb 1;71(3):1060-70.
211. Cooper GM. *The Development and Causes of Cancer. The Cell: A Molecular Approach* 2nd edition [Internet]. 2000 [cited 2018 May 20]; Available from: <https://www.ncbi.nlm.nih.gov/books/NBK9963/>
212. Goel RK, Lukong KE. Understanding the cellular roles of Fyn-related kinase (FRK): implications in cancer biology. *Cancer and Metastasis Reviews*. 2016 Jun;35(2):179-99.
213. Teles Alves I, Hartjes T, McClellan E, Hiltemann S, Böttcher R, Dits N, et al. Next-generation sequencing reveals novel rare fusion events with functional implication in prostate cancer. *Oncogene*. 2015 Jan 29;34(5):568-77.
214. Abate F, Zairis S, Ficarra E, Acquaviva A, Wiggins CH, Frattini V, et al. Pegasus: a comprehensive annotation and prediction tool for detection of driver gene fusions in cancer. *BMC Syst Biol*. 2014 Sep 4;8:97.
215. Stransky N, Cerami E, Schalm S, Kim JL, Lengauer C. The landscape of kinase fusions in cancer. *Nat Commun*. 2014 Sep 10;5:4846.
216. Key statistics for childhood cancer [Internet]. [cited 2016 Sep 14]. Available from: <http://www.cancer.org/cancer/cancerinchildren/detailedguide/cancer-in-children-key-statistics>

ANNEXE : Résultats des travaux de Master 2

A. Effets moléculaires et cellulaires de la réexpression des facteurs de transcription thyroïdiens TTF-1 et PAX-8

Les facteurs de transcription *TTF-1* (thyroid transcription factor) et *PAX-8* (paired box gene 8) sont responsables de l'organogénèse thyroïdienne et du maintien de la différenciation des thyrocytes. Nous avons émis l'hypothèse que l'induction de ces deux facteurs de transcription pourrait agir sur la différenciation et sur la tumorigénicité dans des carcinomes thyroïdiens.

Pour cette étude, nous avons utilisé des lignées cellulaires dérivant de carcinomes thyroïdiens papillaires (TPC-1 et BHP 10-3) et anaplasique (ARO) stablement transfectées par des vecteurs d'expression contenant les gènes *TTF-1* ou *PAX-8*. De plus, la capacité de divers agents pharmacologiques à moduler l'expression de TTF-1 et PAX-8 et leurs effets sur l'apoptose ont également été analysés. Les effets sur l'expression au niveau génique et protéique des deux facteurs, ainsi que sur la prolifération, le cycle cellulaire, la migration *in vitro* et la tumorigénicité dans les trois lignées de carcinomes thyroïdiens ont été évalués.

Nos résultats montrent que: i) TTF-1 induit réciproquement l'expression de PAX-8; ii) l'état basal de TTF-1 ou PAX-8 influence la prolifération, la migration et la tumorigénicité des cellules; iii) l'induction de TTF-1 a un effet plus important sur la prolifération cellulaire que PAX-8 et affecte principalement la tumorigénicité *in vivo*; et iv) TTF-1 s'est révélé plus sensible aux modulateurs épigénétiques que le PAX-8.

En conclusion, les résultats obtenus au cours de mon stage de M2 suggèrent que les facteurs de transcription *TTF-1* et *PAX-8* semblent être finement régulés, le déséquilibre de cette régulation pourrait conduire à un effet promoteur de la tumorigénèse et pro-tumorigène.

Par conséquent, il semblerait que TTF-1 et PAX-8 co-exprimés ont des propriétés anti-prolifératives et anti-tumorigènes jusqu'à un niveau d'expression seuil au-delà duquel ils sont capables d'induire des effets pro-tumorigènes. A l'avenir, il serait ainsi intéressant de prendre systématiquement en compte l'état basal d'expression de TTF-1 et de PAX-8 et de les étudier en tant que duo responsable du contrôle de la tumorigénicité thyroïdienne. Ceci pourrait, à long terme, ouvrir de nouvelles perspectives thérapeutiques pour les carcinomes thyroïdiens.

La figure ci-dessous représente la conclusion des résultats obtenus lors de la préparation de mon Master 2 effectué au sein du laboratoire :

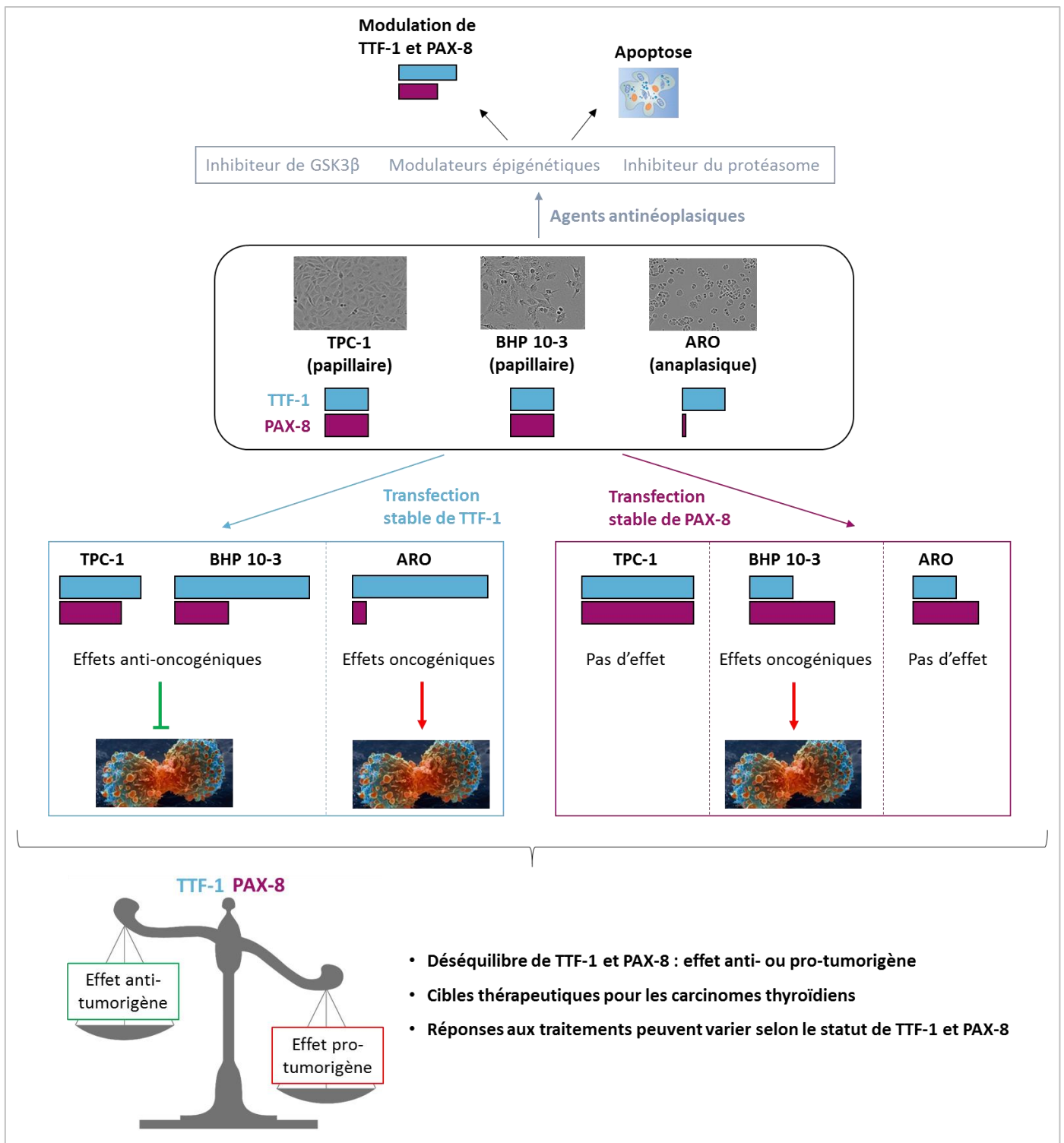


Schéma récapitulatif des résultats obtenus dans le cadre des travaux de Master 2

B. Article "Induction of TTF-1 or PAX-8 expression on proliferation and tumorigenicity in thyroid carcinomas"

Induction of TTF-1 or PAX-8 expression on proliferation and tumorigenicity in thyroid carcinomas

CÉLIA DUPAIN¹, HAFIZ M. ALI^{1,2}, TARIK AIT MOUHOUB¹,
GIORGIA URBINATI¹ and LILIANE MASSAAD-MASSADE^{1,2}

¹Vectorology and Anticancer Therapies, UMR 8203 CNRS, Université Paris-Sud, Gustave Roussy, Université Paris-Saclay, 94805 Villejuif, France; ²University College of Veterinary and Animal Sciences, IUB, 63100 Bahawalpur, Pakistan

Received May 19, 2016; Accepted July 4, 2016

DOI: 10.3892/ijo.2016.3617

Abstract. *TTF-1* and *PAX-8* are responsible for thyroid organogenesis and for maintenance of differentiation in thyrocytes. Thus, we hypothesized that the induction of these two transcription factors could affect proliferation and tumorigenicity. Moreover, the ability of various pharmacological agents to modulate expression of the *TTF-1* and *PAX-8* and their effects on apoptosis were also analysed. For this purpose, cell lines derived from papillary (TPC-1 and BHP 10-3) and anaplastic (ARO) thyroid carcinomas were stably transfected with expression vectors containing *TTF-1* or *PAX-8* genes. Subsequently, the effects on expression at gene and protein levels, as well as on cell growth, cell cycle, migration and *in vivo* tumorigenicity were studied. Our results showed that: i) *TTF-1* reciprocally induces *PAX-8* expression; ii) the basal state of *TTF-1* or *PAX-8* influences proliferation, migration and tumorigenicity; iii) the induction of *TTF-1* acts on cell proliferation more than *PAX-8* and mainly affects tumorigenicity; and iv) *TTF-1* was found to be more sensitive to epigenetic modulators than *PAX-8*. Therefore, we postulated that both *TTF-1* and *PAX-8* when co-expressed have anti-proliferative and anti-tumorigenic properties up to a threshold expression level and beyond that, are able to induce pro-tumorigenic effects. Hence in future, it will be quite interesting to systematically take into account the basal state of expression of *TTF-1* and *PAX-8*. It will also be important to study the two thyroid transcription factors as part of a duo. This could open in the long-term, new therapeutic perspectives for thyroid carcinomas.

Introduction

Genes involved in the thyroid differentiation are essentially thyroid transcription factor-1 (*TTF-1*) and human paired box-8 (*PAX-8*). Both *TTF-1* and *PAX-8* control the expression of

thyroglobulin (*Tg*), thyroperoxidase (*TPO*), thyroid-stimulating hormone receptor (*TSHr*) and sodium/iodide symporter (*NIS*) by binding to the promoters of these genes (1).

TTF-1 (also known as *NKX2-1*, *T/EBP* or *TITF-1*) is commonly expressed in thyroid gland and the central nervous system (2). It is considered as a marker of differentiation in thyroid and lung carcinomas and has been widely used to discern the tumours of thyroid and lung origin, in the patients with metastatic disease (3). Moreover, it is also a useful immunohistochemical marker in the diagnosis of these cancers (4,5). *TTF-1* mRNA has been detected in papillary thyroid carcinomas (PTC) but not in anaplastic cancers; therefore, *TTF-1* would be a marker to distinguish between these two types of thyroid neoplasms (6,7). Regarding its prognosis, *TTF-1* may be increased in PTC with aggressive clinical course (8). It is also reported that high or low expression of *TTF-1* compared to normal levels is a marker of worse prognosis in lung adenocarcinoma; suggesting its importance in disease progression (9). Notably, *NKX2-1* encoding *TTF-1* has been described as a double-edged sword gene due to its dual function, with both pro- and anti-oncogenic activities in lung cancer, depending on the co-existing oncogenic events and differentiation status (10-13).

PAX-8 is a transcription factor expressed in thyrocytes and renal cells (14). It was found to be important in the regulation of organogenesis of kidney and Müllerian system and expressed in several tumours including thyroid, kidney and ovarian carcinomas (14,15). It is also considered to be a diagnostic marker for renal, endometrial and ovarian cancers (16) but inconsistent results are reported concerning its expression in thyroid carcinomas. In fact, *PAX-8* was found to be expressed in differentiated thyroid tumours [PTC and follicular TC (FTC)] while, discrepancies were described in results for anaplastic thyroid carcinomas (6,17). Hence, the relevance of *PAX-8* in thyroid carcinogenesis still remains to be investigated.

Several authors pointed out the key roles of both *TTF-1* and *PAX-8* in the differentiation of thyrocytes. Presta *et al* (18) postulated that the induction of *PAX-8* induces re-differentiation of undifferentiated thyroid cancer cells (18). Mu *et al* (19) showed that co-expression of *TTF-1* and *PAX-8* in thyroid tumour cell lines, derived from papillary or follicular cancers and infected with recombinant adenoviruses (AdTTF-1 and AdPAX-8), leads to accumulation of *TPO* and *Tg* and organization and intracellular retention of iodine. Interestingly,

Correspondence to: Dr Liliane Massaad-Massade, Vectorology and Anticancer Therapies, UMR 8203 CNRS, Université Paris-Sud, Gustave Roussy, Université Paris-Saclay, 94805 Villejuif, France
E-mail: liliane.massade@gustaveroussy.fr

Key words: thyroid transcription factors, thyroid carcinomas, proliferation, apoptosis, tumorigenicity

transient expression of *TTF-1* and *PAX-8* in mouse embryonic stem cells leads to the development of functional follicular cells able to organize iodine (20). Therefore, *TTF-1* and *PAX-8* act together to promote the generation and differentiation of functional thyroid tissue.

Hence, the critical importance of *TTF-1* and *PAX-8* in thyroid organogenesis and their function directs the attention to their expression status which could be modulated during tumorigenesis by gene mutations, post-transcriptional regulations or by epigenetic modifications (21,22). The use of histone deacetylase inhibitors [HDACi; trichostatin A (TSA), depsipeptide (DEPSI), sodium butyrate (NaB), suberoylanilide hydroxamic acid (SAHA), LBH589 or valproic acid (VPA)] or demethylating agents [5-azacytidine (5'-AZA), generally used in the treatment of malignant hemopathies] showed dissimilar effects on expression of *TTF-1* or *PAX-8* (23-27). We noted only slight discrepancy in the published results, even if the authors used the same cell line and similar molecule, thus, making it quite difficult to conclude about the usefulness of these epigenetic modulators in cell differentiation and anticancer activity. Moreover, Kondo *et al* (28) correlated the absence of *TTF-1* expression with methylated status of *TTF-1* in several differentiated and undifferentiated thyroid carcinoma cell lines. They also showed a positive correlation between acetylation of histone H3-Lys9 and the absence of *TTF-1* expression (28) and thus, concluded that DNA demethylating agents could restore *TTF-1* gene expression in thyroid carcinoma cell lines. Our team showed that other molecules would impact on cell differentiation through expression of *TTF-1*. We reported that *TTF-1* is tightly regulated through Wnt/ β -catenin signaling pathway in PTC cells by several mechanisms including transcriptional regulation mediated by β -catenin-binding to a TCF/LEF-responsive element in the *TTF-1* promoter and by post-transcriptional modifications influencing *TTF-1* mRNA and protein expressions (29). We speculated that the administration of GSK-3 β inhibitors such as lithium chloride (LiCl) or other epigenetic *TTF-1* modulators could stimulate *TTF-1* expression and as a consequence, promote differentiation of thyroid cells in pathologies where *TTF-1* expression is low or absent (30). This hypothesis was recently validated when LiCl was shown to improve the efficacy of treatment for thyroid carcinoma by radioiodine (31). Moreover, bortezomib (Velcade[®]), a proteasome inhibitor used in multiple myeloma, showed antineoplastic effects on anaplastic thyroid carcinoma-derived cell lines and regulates *TTF-1* and *PAX-8* expressions (32). Furthermore, it has been described to modulate histone acetylation (33).

Thus, in the present study, we hypothesized that both *TTF-1* and *PAX-8* are tightly regulated and their over- or underexpression could influence the tissue differentiation. Therefore, we first investigated the role of *TTF-1* and *PAX-8* in proliferation and tumorigenicity, then we analyzed the efficiency of several pharmacological molecules able to modulate *TTF-1* and *PAX-8* expressions and studied their outcome in apoptosis.

Materials and methods

Chemicals. Dulbecco's modified Eagle's medium (DMEM), Opti-MEM, Roswell Park Memorial Institute medium

(RPMI), fetal calf serum (FCS), Lipofectamine 2000[™] (1 mg/ml), propidium iodide (PI) kit and PCR primers were purchased from Life Technologies[™] (Saint Aubin, France). Annexin-V Fluos kit was purchased from Roche (Neuilly-sur-Seine, France). BD Matrigel[™] (Basement Membrane Matrix Growth Factor Reduced) was purchased from BD Biosciences (Le Pont de Claix, France). Water was purified by Milli-Q system (Millipore, Saint Quentin en Yvelines, France). The chemicals used in the present study were of highest analytical grade.

Cell lines and cell culture. TPC-1, a human PTC cell line and ARO cells derived from anaplastic thyroid carcinoma were kindly provided by Dr C. Dupuy (Gustave Roussy, Villejuif, France). BHP 10-3^{SCmice} cells have the same genetic profile as TPC-1 cell line and were kindly provided by Dr G. Clayman (MD Anderson Cancer Center, Houston, TX, USA). TPC-1 and BHP 10-3 were grown in DMEM and ARO in RPMI medium. DMEM and RPMI were supplemented with 10% fetal calf serum (FCS), penicillin (100 U/ml) and streptomycin (10 μ g/ml) and maintained at 37°C in an atmosphere of 5% CO₂ and 95% humidity. Cells were systematically tested by PCR analysis to be free of mycoplasma.

Establishment of clones stably expressing *TTF-1* and *PAX-8*. Plasmids used in the present study are human expression vectors (cDNA *TTF-1* and cDNA *PAX-8*) cloned in pcDNA3 (34,35). pcDNA3.1-*TTF-1* and pcDNA3.1-*PAX-8* plasmids containing ampicillin and neomycin resistance genes were a generous gift from Dr M. Polak (INSERM UMR S1016 CNRS, UMR 8104 Institut Cochin, Paris, France). The pcDNA3 empty vector was used as a control.

To obtain the cell lines with *TTF-1* or *PAX-8* stable expressions, TPC-1, BHP 10-3 and ARO cells were seeded in 6-well plates and transfected with pcDNA-*TTF-1* or with pcDNA-*PAX-8* plasmids. In order to control the transfection effects, the three cell lines were also stably transfected with pcDNA3 empty vector. Transfections were realized by Lipofectamine 2000 as recommended by the supplier. Briefly, 6 μ g of plasmid and 7.5 μ l of Lipofectamine 2000 (1 μ g/ μ l) were mixed in 2 ml of serum free Opti-MEM culture medium. After 6 h of incubation at 37°C, the medium was replaced with complete DMEM culture medium containing FCS. Then, after 48 h of incubation, cell colonies were collected and expanded in the same growth medium. The selective pressure by neomycin (88 μ M) was maintained in cell culture during three weeks.

Pharmacological treatments. Cells were seeded in 6-well plates at a concentration of 4x10⁵ cells/well in 2 ml of DMEM. After 24 h, cells are treated by one of the pharmacological agents: trichostatin A (TSA; 300 nM or 1 μ M), lithium chloride (LiCl; 20 mM), valproic acid (VPA; 3 mM), 5'-azacytidin (5'-AZA; 500 nM or 1 μ M) or bortezomib (BOR; 100 nM). Cells were incubated for 48 h at 37°C in cell culture medium for hydrophilic (LiCl, VPA and 5'-AZA) or in ethanol (EtOH 100%) for hydrophobic (TSA and BOR) molecules. A calibration curve was performed by spectrophotometry to ensure the final TSA and BOR concentrations after dissolution. Each condition was tested in duplicate in three independent experiments.

Reverse transcription-quantitative PCR (RT-qPCR). Total RNA extraction and RT-qPCR were performed to compare *TTF-1* and *PAX-8* mRNA levels for: i) basal expression of wild-type cell lines; ii) WT cell lines vs. their stably transfected clones; and iii) treated cells with pharmacological molecules vs. untreated cells. RNA extraction was performed from 5×10^6 collected cells using RNeasy Mini kit (Qiagen, Courtaboeuf, France) as previously described (36). RNA purity and quantity were evaluated by NanoDrop® ND-1000 spectrophotometry (spectrophotometer; Thermo Fisher Scientific, Wilmington, DE, USA). First-strand cDNA was generated with M-MLV RT from Life Technologies and real-time PCR (qPCR) was carried out with StepOnePlus PCR System (AB Applied Biosystems, Villebon-sur-Yvette, France) using GoTaq® qPCR Master Mix (Promega, Charbonnières-les-Bains, France) according to the manufacturer's instructions. The following primers were used to amplify the target genes: i) *TTF-1* forward (F), 5'-CGCGTTTACCAAGGAAC-3' and *TTF-1* reverse (R), 5'-GAGTGTGCCAGAGTGAAG-3'; ii) *PAX-8* (F), 5'-AGGTGGTGGAGAAGATTGG-3' and *PAX-8* (R), 5'-ATAGGGAGGTTGAATGGTTG-3'. Gene expression was determined by quantification-comparative $2^{-\Delta\Delta C_t}$ method (37) and normalized to *GAPDH* levels using these sequences: [*GAPDH* (F), 5'-ATCCCATCACCATCTTCCAG-3' and *GAPDH* (R), 5'-CCATCACGCCACAGTTTCC-3']. Results are expressed as relative mRNA levels comparing the clones or treated cells to WT cell lines and represent at least three independent experiments realized in duplicate.

Immunoblotting. Cells were lysed and total proteins were extracted using mammalian protein extraction reagent (M-PER; Thermo Fisher Scientific, Rockford, IL, USA) in the presence of a protease inhibitor cocktail (Roche, Neuilly-sur-Seine, France) and quantified by Bio-RAD assay at 570 nm (36,38). The amount of 30 μ g of each sample were heated at 70°C for 10 min with 1X sample reducing buffer (NuPAGE sample reducing agent; Invitrogen) and 1X sample loading buffer (NuPAGE LDS sample agent; Invitrogen). Samples were then loaded on 10% polyacrylamide gel (NuPAGE Bis-Tris mini gels 10%; Life Technologies). Proteins were transferred on nitrocellulose membranes using iBlot™ Dry blotting system (Invitrogen). After saturation with either I-Block reagent (Tropix, Inc., Bedford, MA, USA) or 10% BSA (bovine serum albumin) solution, membranes were incubated overnight at 4°C under agitation with one of the following primary antibodies: monoclonal rabbit *TTF-1* (1:1,000, ab133737; Abcam Biochemicals, Paris, France); monoclonal rabbit *PAX-8* (1:1,000, ab53490; Abcam Biochemicals). β -actin-HRP (horseradish peroxidase) was used as an internal control (1:1,000; Sigma-Aldrich Chemicals Co., Saint Quentin Fallavier, France). Blots were then washed and incubated with corresponding secondary antibodies: anti-rabbit-AP (alkaline phosphatase 1:20,000; Tropix) or anti-rabbit-HRP (1:3,000; Cell Signaling Technology, Saint Quentin en Yvelines, France). Bands were revealed with CDP-Star Chemiluminescence reagent (alkaline phosphatase system; Perkin Elmer, Courtaboeuf, France) or by enhanced chemiluminescence reagent (HRP system, Clarity™; Bio-Rad Laboratories,

Marnes-la-Coquette, France). Experiments were repeated at least 3 times.

Doubling time and cell cycle analysis. In order to establish the doubling time of WT cell lines compared to their stably transfected clones by *TTF-1* or *PAX-8*, 10^4 cells/well were seeded in 200 μ l culture medium in 96-well plates. Cells were then incubated in the IncuCyte™ (Essen Instruments, Inc., Ann Arbor, MI, USA); that allows a non-invasive automated method to monitor viable cell growth and proliferation in culture. Each well was scanned by camera fitted in IncuCyte™ at 4-h intervals for three days, with 4 images per well. Thereby, time lapse cell confluence was obtained for each cell type and their doubling time was calculated according to an exponential regression equation, given as: $y = ax^{bx}$, via the formula: $t1/2 = \ln(2)/b$.

For cell cycle analysis, TPC-1, BHP 10-3 and ARO and their corresponding transfected clones were collected and incubated with DNA staining buffer (50 μ g/ml PI in 0.1% sodium citrate, 0.1% Triton X-100, 100 μ g/ml RNase A) in the dark, for 30 min at room temperature. Samples were then analysed by flow cytometry (Accuri C6 flow cytometer; BD Biosciences, San Jose, CA, USA). The results are of at least three independent experiments and represent the percentage of cells distributed in the cell cycle phases, in the stable clones vs. their corresponding WT cell lines or in treated cells vs. untreated ones.

Cell migration assays. Scratch test was performed to evaluate the effects of stable transfection either with *TTF-1* or *PAX-8* genes on cell migration, as previously described (39). Briefly, 100 μ l BD Matrigel™ was plated in 96-well ImageLock™ cell migration plates (Essen BioScience Inc., Ann Arbor, MI, USA) 24 h before seeding of TPC-1, BHP 10-3 and ARO cells and their corresponding stably transfected clones. Cells (1×10^4 /well) were then plated and when reached 90% confluence, the monolayer was scratched with a 96-pin WoundMaker (Essen BioScience). The cells were maintained in fresh culture medium until complete wound confluence and the cell mobility was monitored by IncuCyte™ every 4 h by 'scratch wound' scan type. Results are presented as scratch wound width in function of time.

Annexin V apoptosis assay. Apoptosis was determined by flow cytometric analysis using Annexin V kit. Each cell line was seeded in 6-well plates (3.5×10^5 cells/well) in 2 ml medium and pharmacologically treated with EtOH (20 μ l), TSA (300 nM or 1 μ M), LiCl (20 mM), VPA (3 mM), BOR (100 nM) or 5'-AZA (500 nM or 1 μ M). After 48 h of incubation, cells were collected and centrifuged and the pellets were stained with Annexin V Fluos kit (Roche, Neuilly-sur-Seine, France) according to the manufacturer's instructions. Experiments were performed in triplicate of independent experiments and data represent percentage of apoptotic cells compared to non-treated cells, normalized to EtOH condition.

Animal studies and tumorigenicity tests. All animal experiments and the use of cell lines were approved by the institutional Ethics Committee of Animal Experimentation (CEEA) and research council (Integrated Research Cancer

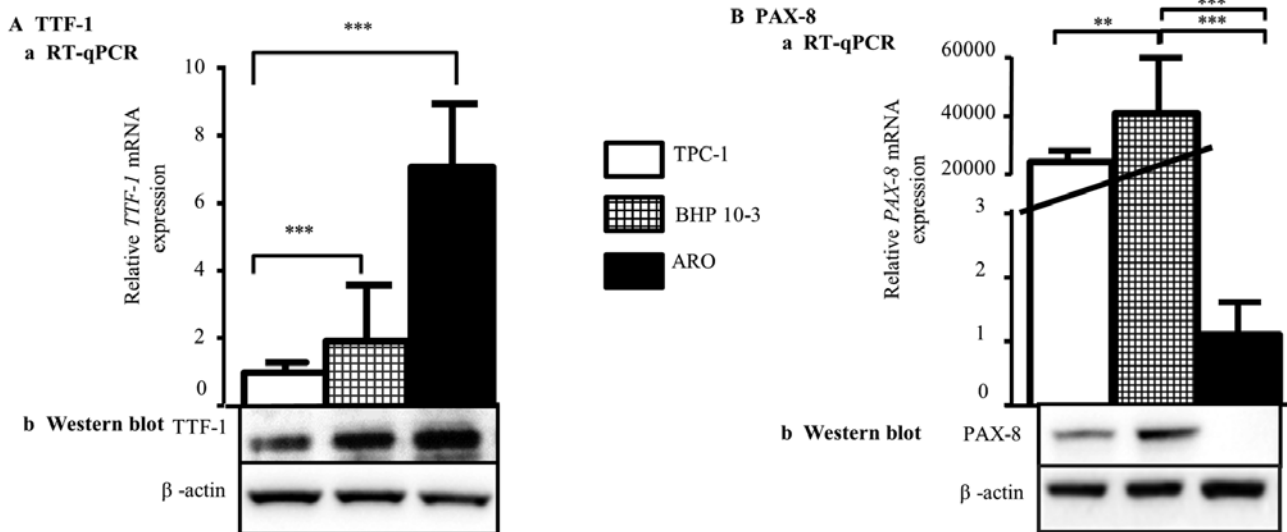


Figure 1. Basal expression of *TTF-1* and *PAX-8* in TPC-1, BHP 10-3 and ARO cells. Total mRNA and proteins were extracted from TPC-1, BHP 10-3 and ARO cells. (A-a) Relative *TTF-1* mRNA levels were analysed by RT-qPCR and compared to TPC-1 cell line. (B-a) Relative *PAX-8* mRNA levels were analysed by RT-qPCR and compared to TPC-1 cell line. All the results are normalised to *GAPDH* mRNA levels and are the mean \pm SD of three independent experiments. ** $P < 0.01$, *** $P < 0.001$. Western blot analysis of *TTF-1* (A-b) and *PAX-8* (B-b) proteins in TPC-1, BHP 10-3 and ARO cell lines. β -actin was used as an internal control. Results are representative of two independent experiments.

Institute in Villejuif; IRCIV), registered in the French Ministry of Higher Education and Research (Ministère de l'Enseignement Supérieur et de la Recherche; MESR) under the authorization number CEEA IRCIV/IGR no. 26: 94-226, no. 2011-09 and carried out according to French laws under the conditions established by the European Community (Directive 2010/63/UE). Six-week-old nude nu/nu and five-week-old NSG (NOD/SCID Gamma) female mice were purchased from the Animal Facility of the Institute Gustave Roussy and housed in a sterilised laminar flow caging system. Water and bedding were given *ad libitum* and autoclaved before being put in the cages. All efforts were made to minimize animal suffering and animals were sacrificed by CO_2 inhalation at the end of the experiments, before the collection of tumours.

Nude mice were used to test the tumorigenicity of BHP 10-3 and ARO cell lines and their respective clones. Since, TPC-1 cells are reported as non-tumorigenic in nude mice (40), NSG mice were used for this purpose. All cell lines and their corresponding clones diluted in 100 μl PBS were injected subcutaneously into the flank of mice ($n=5/\text{group}$) at a rate of 2.0×10^6 cells/mouse for BHP 10-3 and ARO or 10^7 cells/mouse for TPC-1. Mice were monitored every two or three days for tumour growth and then sacrificed when tumours reached a volume of 1,000 mm^3 . Tumours were immediately frozen in liquid nitrogen for western blot analysis.

Protein extractions from tumours and western blot experiments. Tumours were ground, total proteins were extracted and western blots were performed as described above, to evaluate expression of the *TTF-1* and *PAX-8* protein.

Statistical analysis. The data are presented as mean \pm SD (standard deviation). By using GraphPad Prism 4 software, Mann-Witney test was employed to compare between two groups of treatments and Kruskal-Wallis test was performed to compare among multiple treatments. Non-parametric analysis

of longitudinal data in factorial experiments was performed to compare the treatments *in vivo*, using 'nparLD' package from 'R' software. $P < 0.05$ was considered as the statistically significant level.

Results

TTF-1 and *PAX-8* have different expression profiles in wild-type cell lines. First, we tested expression of the *TTF-1* and *PAX-8* in TPC-1, BHP 10-3 and ARO cell lines by RT-qPCR (Fig. 1A) and western blot analysis (Fig. 1B). Regarding *TTF-1* basal expression, our results showed that *TTF-1* is more expressed in ARO compared to TPC-1 and BHP 10-3 cell lines, both at mRNA (Fig. 1A-a) and protein (Fig. 1B-a) levels. Concerning *PAX-8*, mRNA and protein were found to be highly expressed in BHP 10-3 cells compared to TPC-1 cells whereas it was detected slightly in anaplastic ARO cell line (Fig. 1A-b and B-b). These results confirmed the thyroid origin of the three cell lines, and the different expression profiles of *TTF-1* and *PAX-8* genes.

TTF-1 upregulates *PAX-8* in thyroid carcinoma cell lines. TPC-1, BHP 10-3 and ARO cell lines were stably transfected with pcDNA plasmids containing either *TTF-1* or *PAX-8* genes or with pcDNA3 empty vector. After 3 weeks of neomycin selection, the resulting clones were collected and characterized by RT-qPCR and western blot analysis.

Firstly, we assessed the effects of pcDNA3 empty vector on the expression of *TTF-1* and *PAX-8* mRNA by RT-qPCR. We found that *TTF-1* mRNA level was not affected by pcDNA3 stable transfection in either TPC-1 or ARO cells. Concerning BHP 10-3 cells, we noted a decrease in mRNA *TTF-1* expression that will not interfere with our following experiments, since no induction in *TTF-1* transcription factor was observed. Also, for *PAX-8*, the stable transfection did not affect the mRNA levels of the three cell lines tested (data not shown).

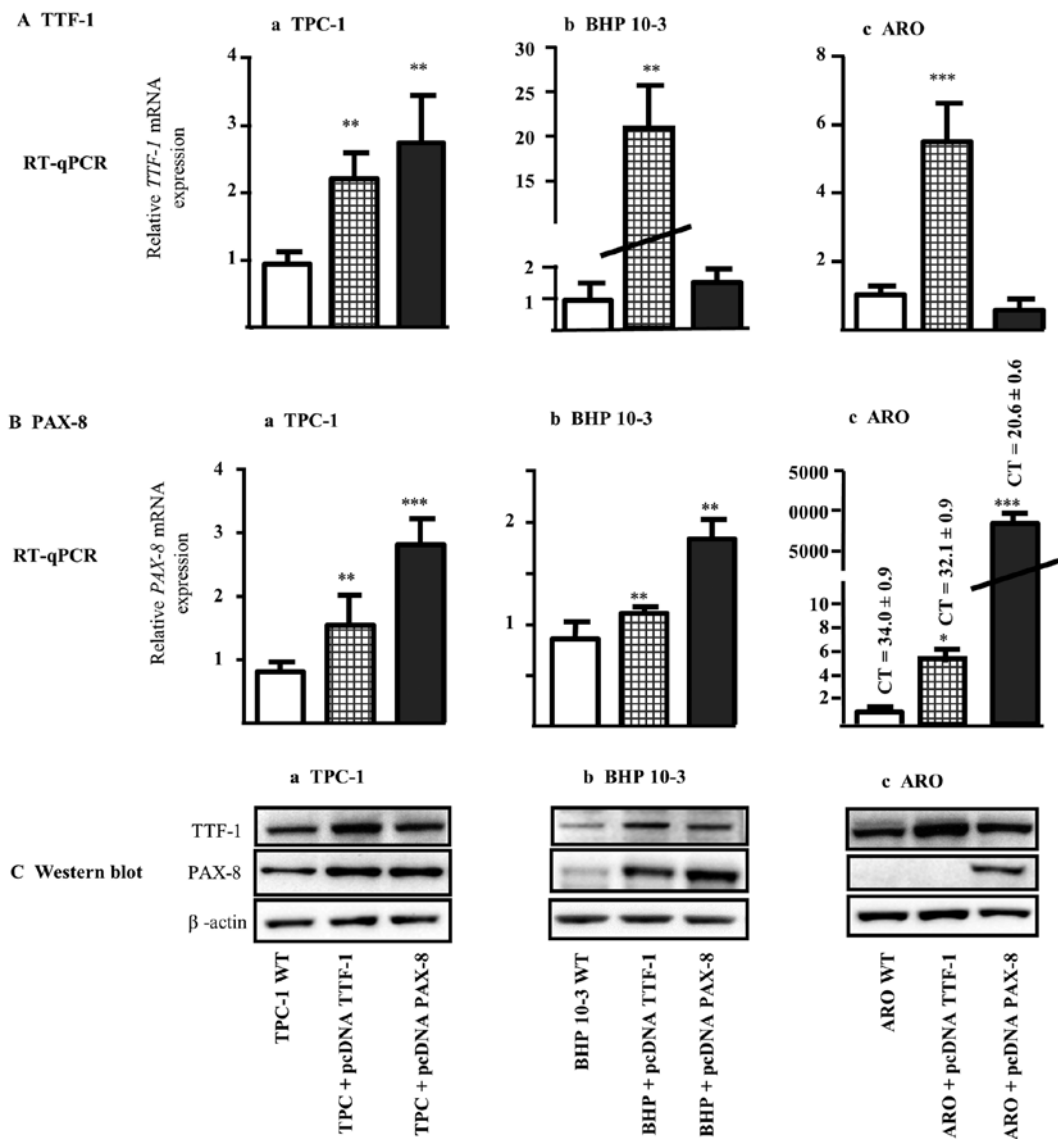


Figure 2. Induction of the expression of *TTF-1* and *PAX-8* in TPC-1, BHP 10-3 and ARO, after stable transfection. Total mRNA was extracted from TPC-1 (A-a and B-a; RT-qPCR), BHP 10-3 (A-b and B-b; RT-qPCR) and ARO cells (A-c and B-c; RT-qPCR). (A) *TTF-1* and (B) *PAX-8* mRNA levels were analysed by RT-qPCR and compared to their corresponding wild-type cell lines. Results are normalised to *GAPDH* mRNA levels. Bars represent mean \pm SD of three independent experiments. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. *TTF-1* and *PAX-8* protein levels were tested by western blot analysis in TPC-1 (C-a), BHP 10-3 (C-b) and ARO (C-c), clones and their respective WT cell lines. β -actin was used as an internal control. Results are representative of at least three independent experiments.

Then, clones derived from each cell line transfected with pcDNA plasmids containing either *TTF-1* or *PAX-8* genes were tested for the expression profiles of *TTF-1* and *PAX-8* mRNA and their corresponding proteins (Fig. 2). The relative mRNA *TTF-1* level in transfected clones with pcDNA *TTF-1* for TPC-1, BHP 10-3 and ARO cells was ~ 2.5 -, 20- and 5-fold, respectively more important compared to their corresponding wild type cell lines (Fig. 2A-a-c; RT-qPCR). The upregulation of *TTF-1* mRNA levels was paralleled by a similar increase in *TTF-1* protein in all pcDNA *TTF-1* derived clones (Fig. 2C-a-c; western blot analysis, upper panels). Similarly, the relative mRNA *PAX-8* expression was respectively 3-, 2- and 8000-fold higher in TPC-1, BHP 10-3 and ARO clones transfected with pcDNA *PAX-8* vector (Fig. 2B-a-c; RT-qPCR) and paralleled with higher *PAX-8* protein level (Fig. 2C-a-c; western blot analysis, mid-panels). It should be noticed that the strong increase in *PAX-8* level found in ARO clone transfected

with pcDNA *PAX-8* could be a consequence of low mRNA levels at basal state in ARO WT cells ($ct = 34 \pm 0.9$) (Fig. 2B-c, RT-qPCR and 2C-c, western blot analysis).

Therefore, in the three cell lines tested, *TTF-1* induction is able to enhance mRNA and protein *PAX-8* contents except when the basal state of *PAX-8* is absent or too low (Fig. 2B-c and 2C-c). Taking together our results showed that the upregulation of *TTF-1* is able to induce *PAX-8* expression.

Upregulation of TTF-1 acts more effectively on cell doubling time and cell cycle than PAX-8. Both cell doubling time and cell cycle of the clones were compared to their corresponding WT cells. In the three clones stably transfected with pcDNA *TTF-1*, our results showed that the doubling time calculated according to the confluence was significantly increased compared to their counterpart WT cells (Table I). The pcDNA *PAX-8* transfection had no effects on doubling time of TPC-1

Table I. Doubling time and cell cycle progression of TPC-1, BHP 10-3 and ARO cells and their corresponding clones transfected with pcDNA *TTF-1* or pcDNA *PAX-8*.

Cell lines and clones	Doubling time		Cell cycle	
	Mean \pm SD (h)	G0/G1 (%)	S (%)	G2/M (%)
TPC-1 WT	25.00\pm4.61	80.86\pm4.15	8.20\pm2.95	10.82\pm1.60
TPC-1 + pcDNA3	NT	80.2 \pm 13.15	5.45 \pm 1.48	13.90 \pm 13.44
TPC-1 + pcDNA <i>TTF-1</i>	43.12 \pm 5.76 ^c	93.70 \pm 2.63 ^c	3.34 \pm 2.28 ^a	2.30 \pm 0.57 ^c
TPC-1 + pcDNA <i>PAX-8</i>	25.23 \pm 2.87	85.02 \pm 4.62	6.32 \pm 3.30	8.48 \pm 1.77
BHP 10-3 WT	47.12\pm5.95	74.63\pm3.33	17.10\pm1.67	12.65\pm3.31
BHP 10-3 + pcDNA3	NT	77.95 \pm 0.78	12.00 \pm 1.93	15.65 \pm 1.06
BHP 10-3 + pcDNA <i>TTF-1</i>	57.75 \pm 5.57 ^b	77.93 \pm 2.74	11.00 \pm 1.41 ^b	11.05 \pm 2.20
BHP 10-3 + pcDNA <i>PAX-8</i>	33.36 \pm 3.68 ^c	58.30 \pm 13.23	18.9 \pm 7.3	18.98 \pm 6.92
ARO WT	28.73\pm2.14	54.00\pm2.26	32.90\pm0.28	14.20\pm1.41
ARO + pcDNA3	NT	59.25 \pm 0.49	29.30 \pm 1.41	11.05 \pm 2.05
ARO + pcDNA <i>TTF-1</i>	47.13 \pm 3.24 ^c	61.70 \pm 0.42 ^a	23.55 \pm 2.76 ^a	15.75 \pm 1.91
ARO + pcDNA <i>PAX-8</i>	25.74 \pm 5.54	56.55 \pm 2.90	22.15 \pm 0.35 ^c	22.10 \pm 1.98 ^a

Doubling time: WT cell lines and their corresponding stably transfected clones were incubated in IncuCyte™. The time lapse cell confluence was obtained for each cell type and their doubling time was calculated. NT, non tested for empty vector. Cell cycle: cells were incubated with propidium iodide (PI) after permeabilization and the repartition of cells within different phases of cell cycle was analysed by flow cytometry (FACS). All results are expressed as mean \pm SD, stars indicate the degree of significance observed between WT and their corresponding clones. *P<0.05, **P<0.01, ***P<0.001.

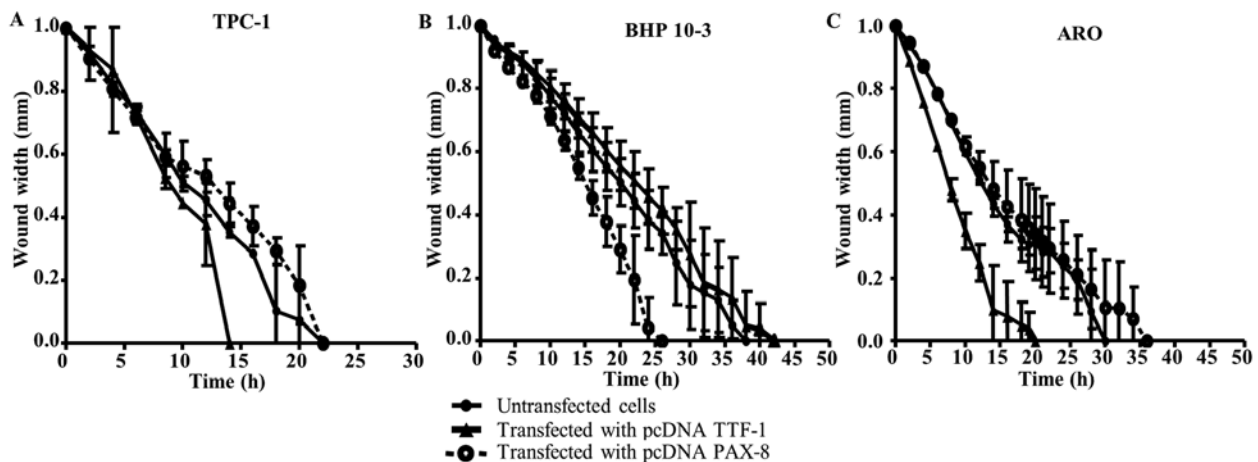


Figure 3. Migration ability of TPC-1, BHP 10-3 and ARO WT cell lines and their corresponding *TTF-1* and *PAX-8* clones. The cell motility was determined by the IncuCyte™. Results are presented as relative wound width in function of time of WT cell lines and their respective clones stably transfected with pcDNA *TTF-1* or pcDNA *PAX-8*. Migration ability of WT cells were compared to their corresponding clones. (A) TPC-1, (B) BHP 10-3 and (C) ARO. Bars represent mean \pm SD of two independent experiments.

and ARO derived clones but significantly increased the growth of BHP 10-3 clone; this might be due to high basal expression of *PAX-8* in BHP 10-3 WT cells compared to TPC-1 and ARO cells (Fig. 1B and C). Moreover, it should also be noted that the growth characteristics of each WT cell line is different, thus, making it difficult to compare their growth rate (TPC-1 and BHP 10-3 spread along the surface of the support while ARO grows in clusters).

These results were further supported by cell cycle studies, wherein we observed that *TTF-1* induction triggered an increase in cell percentage at G0/G1 phase, up to significant

level for TPC-1 and ARO clones (Table I). Concerning the empty vector, the results of cell cycle studies did not show any difference between WT cells and the corresponding cell clones transfected with pcDNA3 alone (Table I).

These results suggest that in the tested cell lines, only *TTF-1* could have anti-proliferative effects by delaying cell growth and halting cell cycle in G0/G1 phase.

A high basal state expression of PAX-8 influences the migration ability. Wild-type cell lines and their corresponding clones were assessed for their migration ability. We observed that

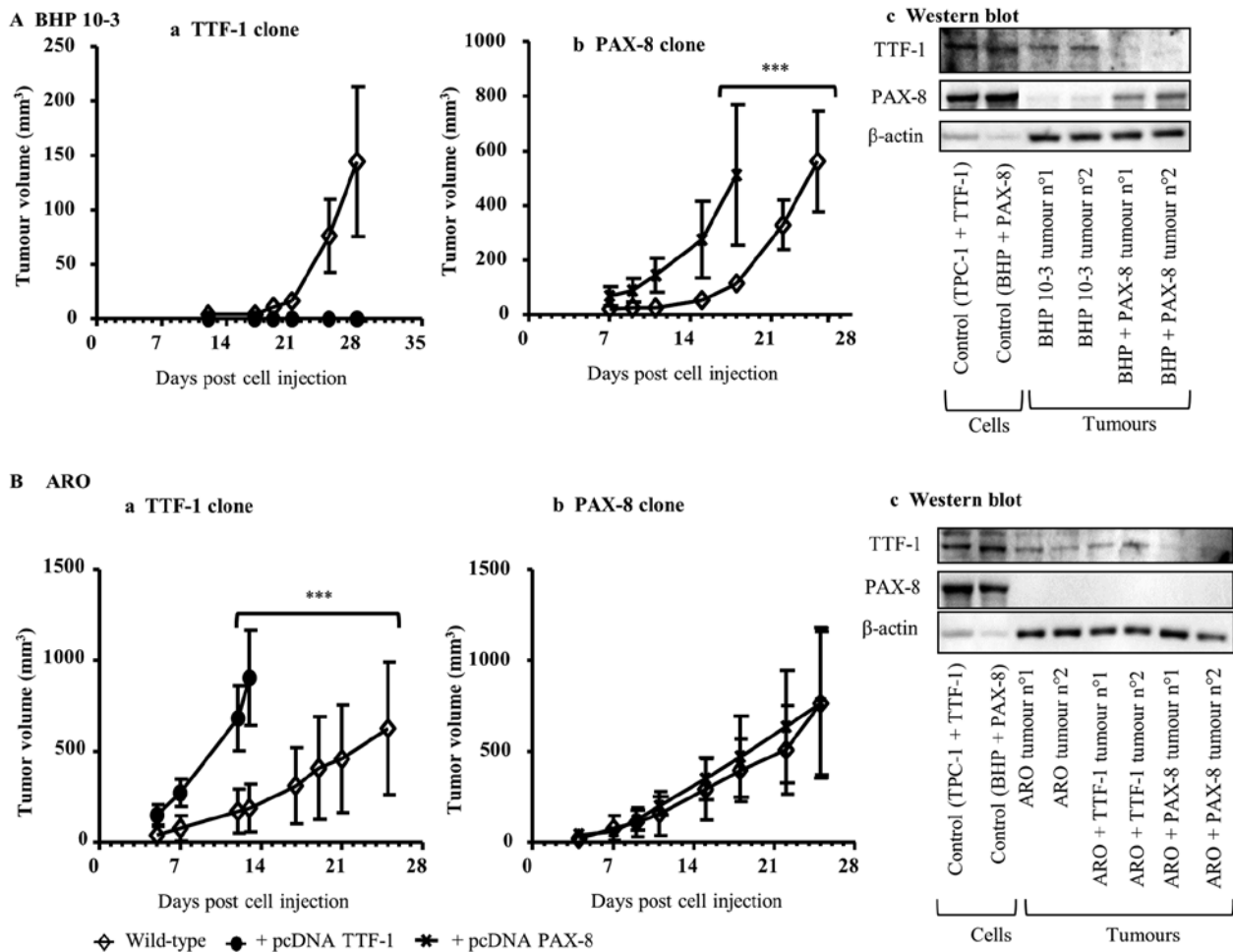


Figure 4. Tumorigenicity of BHP 10-3 and ARO and their corresponding stably transfected clones. Cell lines were tested for their tumorigenic ability in nude mice. Results are presented as tumour volume in function of time. Nude mice (n=5/group) were injected with BHP 10-3 cells (A) or with ARO cells (B). The corresponding clones transfected with pcDNA TTF-1 (A-a and B-a) or with pcDNA PAX-8 (A-b and B-b) were also xenografted. Tumour growth was followed during the course of the experiment until the tumour size reached 1 cm³. Mice were sacrificed and tumours were collected for western blot analysis using TTF-1 or PAX-8 antibodies (A-c and B-c). Positive controls were the clones derived from TPC-1 or BHP 10-3 cell lines maintained in culture medium. 'no. 1' and 'no. 2' represent the tumours collected from two different mice. β-actin was monitored as loading control. BHP, BHP 10-3; TTF-1, pcDNA TTF-1; PAX-8, pcDNA PAX-8.

TTF-1 transfection leads to increased cell migration in clones derived from TPC-1 and ARO WT cell lines (respectively for clones vs. WT: 14 vs. 22 h for TPC-1 and 14 vs. 30 h for ARO; Fig. 3A and C, respectively); no effects were observed for BHP 10-3 cells. While in this latter cell line, PAX-8 transfection enhanced the migration of its corresponding clone (respectively for clone vs. WT: 24 vs. 35 h; Fig. 3B). However, PAX-8 transfection did not affect the migration of TPC-1 and ARO cells. Herein, each transcription factor also seems to have an impact on the migration behavior that appeared to be more influenced when a high PAX-8 basal state level was detected.

TTF-1 expression mainly influences tumorigenicity. According to our previous results, we postulated that the induction of *TTF-1* or *PAX-8* could influence the tumorigenic potential. Concerning TPC-1 cells, known to be non-tumorigenic in nude mice (40), the enhancement in the expression of both transcription factors did not result in tumour formation in NSG mice, known to be more immuno-depressed than nude mice, injected with TPC-1 WT or with TPC-1 + pcDNA TTF-1 or TPC-1 + pcDNA PAX-8 clones (data not shown).

Regarding BHP 10-3 cell line, its injection in nude mice resulted in tumour development. This observation is in accordance with previous results showing that BHP 10-3 cells are tumorigenic in immuno-depressed mice (41). Notably, when BHP 10-3 + pcDNA TTF-1 clone highly expressing *TTF-1* was injected in nude mice, no tumours developed (Fig. 4A-a). Contrastingly, the inoculation of BHP 10-3 + pcDNA PAX-8 clone in nude mice increased tumour growth (Fig. 4A-b). These results were further confirmed by western blot analysis. Tumours collected at the end of the experiments and analyzed for TTF-1 and of PAX-8 protein contents showed an increase in PAX-8 levels in tumours derived from BHP 10-3 + pcDNA PAX-8 clone compared to BHP 10-3 WT tumours. Whereas, the TTF-1 protein contents were less expressed in the same samples (Fig. 4A-c).

In ARO cell line, also known to be tumorigenic (40), *TTF-1* induction (ARO + pcDNA TTF-1) increased tumour growth (Fig. 4B-a), whereas *PAX-8* upregulation (ARO + pcDNA PAX-8) did not influence tumorigenicity (Fig. 4B-b). Herein, western blot analysis showed that TTF-1 protein pattern seems to be slightly increased in the ARO + pcDNA TTF-1 derived

Table II. Effects of pharmacological molecules on expression of *TTF-1* and *PAX-8* in TPC-1, BHP10-3 and ARO thyroid cell lines.

	TPC-1		BHP 10-3		ARO	
	TTF-1	PAX-8	TTF-1	PAX-8	TTF-1	PAX-8
EtOH 1%	1.00±0.04	1.00±0.06	1.00±0.08	1.00±0.07	1.00±0.04	1.00±0.10
TSA 0.3 µM	1.12±0.54	1.56±0.28 ^a	0.95±0.16	0.71±0.10 ^a	2.80±0.37 ^b	1.09±0.10
TSA 1 µM	2.66±0.46 ^a	5.98±1.57 ^b	4.13±1.96 ^a	1.21±0.08 ^a	17.65±0.46 ^c	1.16±0.02
VPA 3 mM	5.20±1.47 ^a	1.35±0.20	1.99±0.22 ^b	1.55±0.05 ^a	2.70±0.53 ^a	5.23±1.24 ^b
5'-AZA 0.5 µM	0.30±0.11 ^b	0.31±0.13 ^c	0.40±0.10 ^b	0.56±0.17 ^a	14.92±1.37 ^b	2.56±0.60 ^a
5'-AZA 1.0 µM	0.27±0.09 ^b	0.30±0.09 ^c	0.36±0.09 ^a	0.72±0.32	15.25±0.56 ^b	3.27±0.08 ^c
LiCl 10 mM	3.54±1.11 ^a	2.37±1.32 ^a	0.65±0.18	0.97±0.08	0.23±0.05 ^b	0.76±0.15
BOR 100 nM	4.37±1.39 ^a	1.82±0.66	1.81±0.55 ^a	0.91±0.09	0.46±0.13 ^b	0.51±0.09 ^a

Variations in expression of *TTF-1* and *PAX-8* were observed in TPC-1, BHP 10-3 and ARO cells after 48-h treatment by trichostatin A (TSA), valproic acid (VPA), 5'-azacytidine (5'-AZA), lithium chloride (LiCl) and bortezomib (BOR). The control represents the cells treated with the vehicle (EtOH = 1%). *TTF-1* and *PAX-8* mRNA expressions were quantified by real-time qPCR and normalized with GAPDH mRNA level. Results of three independent experiments are reported as mean ± SD. Mann-Whitney statistical test was used to compare the difference between treated and untreated cells. Stars represent significant differences; *P<0.05, **P<0.01, ***P<0.001.

Table III. Effects of epigenetic modulators on apoptosis of TPC-1, BHP 10-3 and ARO cell lines.

	TPC-1		BHP 10-3		ARO	
	EA	LA	EA	LA	EA	LA
EtOH 1%	1.00±0.03	1.00±0.03	1.00±0.02	1.00±0.02	1.00±0.03	1.00±0.02
TSA 0.3 µM	1.42±0.02 ^c	1.50±0.02 ^a	1.18±0.02 ^c	1.00±0.02	2.02±0.02 ^b	1.10±0.03 ^c
TSA 1 µM	1.84±0.01 ^c	1.51±0.01 ^b	1.32±0.01 ^c	0.66±0.01 ^c	2.21±0.15 ^b	2.33±0.08 ^c
VPA 3 mM	1.10±0.01 ^b	5.00±0.01 ^c	1.10±0.01 ^a	1.00±0.01	1.10±0.01 ^b	1.43±0.07 ^b
5'-AZA 0.5 µM	1.48±0.01 ^b	4.00±0.01 ^c	0.88±0.02	2.50±0.02 ^b	1.87±0.03 ^c	0.84±0.04
5'-AZA 1.0 µM	1.55±0.01 ^b	3.00±0.02 ^c	1.02±0.01	2.50±0.02 ^b	1.53±0.06 ^c	1.13±0.07
LiCl 10 mM	1.40±0.05 ^c	4.00±0.05 ^c	0.89±0.05	2.00±0.05 ^b	1.29±0.05 ^b	0.91±0.04
BOR 100 nM	2.22±0.04 ^c	2.00±0.04 ^b	1.00±0.04	3.70±0.04 ^b	2.22±0.20 ^b	2.04±0.05 ^b

TPC-1, BHP 10-3 and ARO cells were treated for 48 h with trichostatin A (TSA), valproic acid (VPA), 5'-azacytidine (5'-AZA), lithium chloride (LiCl) and bortezomib (BOR). The cells treated with vehicle (EtOH = 1%) were used as an internal control. Cells were labelled with Annexin-V (AV) and propidium iodide (PI). Early apoptosis (EA; AV⁺PI⁻), late apoptosis (LA; AV⁺PI⁺). Results are expressed as mean ± SD. Mann-Whitney statistical test was used to compare the difference between treated and untreated cells. Stars represent significant differences; *P<0.05, **P<0.01, ***P<0.001.

clone. Concerning *PAX-8*, the protein was not expressed either in ARO derived clones (pcDNA *TTF-1* and pcDNA *PAX-8*) or in the corresponding WT tumours (Fig. 4B-c).

Taken together, these results demonstrated that the basal state of *TTF-1* or *PAX-8* influences tumorigenicity. Moreover, *TTF-1* seems to influence tumour development more than *PAX-8* expression.

TTF-1 is more sensitive to epigenetic modulators than *PAX-8*.

Then, the ability of some pharmacological agents to modulate expression of the *TTF-1* and *PAX-8* in TPC-1, BHP 10-3 and ARO cell lines was studied. It was found that *TTF-1* expression seems to be more affected than *PAX-8* by most of the molecules studied (Table II).

In TPC-1 cells, all molecules except 5'-AZA increased *TTF-1* expression whereas, only LiCl and TSA significantly induced the *PAX-8* level (Table II).

Regarding BHP 10-3 cells, all molecules enhanced *TTF-1* expression except for GSK-3β inhibitor (LiCl) and 5'-AZA which inhibited it. *PAX-8* expression was slightly induced only by VPA and reduced by both TSA and 5'-AZA (Table II).

In ARO cell line, only epigenetic modulators (HDACi and demethylating agents) significantly increased *TTF-1* expression, while LiCl and BOR lead to its inhibition (Table II). Concerning *PAX-8*, only VPA and 5'-AZA significantly increased its expression while BOR significantly decreased *PAX-8* level (Table II).

Epigenetic modulators provoke apoptosis. Flow cytometric analysis performed on TPC-1, BHP 10-3 and ARO cell lines showed that the used molecules could stimulate apoptosis (Table III). Regarding TPC-1 cell line, all the tested molecules induced early and late apoptosis (Table III). Notably, for BHP 10-3 cells, only HDACi were able to provoke early apoptosis while all the other molecules induced late apoptosis (Table III). Concerning ARO cell line, all the tested molecules induced early apoptosis however, only HDACi and BOR triggered late apoptosis (Table III).

Discussion

TTF-1 and *PAX-8* are the major transcription factors involved in the development and differentiation of thyroid gland and play a key role in regulating the expressions of thyroid genes. We hypothesized that an enhancement of expression of *TTF-1* and *PAX-8* could affect both aggressiveness and tumorigenic properties. This hypothesis was evaluated in three models of thyroid carcinoma, two cell lines derived from PTC (TPC-1, BHP 10-3) and one from anaplastic carcinoma (ARO cell line).

In these models, we studied the basal state of *TTF-1* and *PAX-8* expression. Our results are in accordance with previous studies showing the presence of both transcription factors in TPC-1 and BHP 10-3 cells (30,42). Concerning ARO anaplastic cell line, only *TTF-1* was detected, thus providing once more evidence of its thyroid origin (43,44). Moreover, our results showed that the basal state of the expression of *TTF-1* and *PAX-8* is a major parameter that must always be taken into consideration and could explain the divergence of results among various studies concerning the role of these factors (6,10,11,17). In fact, we found that the upregulation of either *TTF-1* or *PAX-8* could affect tumorigenic behavior independently of the histological origin of the cell lines used. We also confirmed previous observations showing that the overexpression of *TTF-1* induces *PAX-8* expression (45-47).

The functional studies depicted the importance of *TTF-1* in the regulation of cell cycle, migration and tumorigenicity. *TTF-1* overexpression led to anti-proliferative effects reflected by an increase in doubling time and raised cell percentage in G0/G1 phase leading to a decrease in cell migration and tumorigenic potential. Contrastingly, this inspection was not confirmed in ARO cells, wherein the upregulation of *TTF-1* led to an increase in cell migration and tumorigenicity. We hypothesized that the migratory and consequently the tumorigenic potential of enhanced *TTF-1* level are dependent on its background expression within the cell lines. Thus, the '*TTF-1* paradox' is again highlighted (10,48). In the same context, functional studies of *PAX-8* depicted that its basal state plays a key role in proliferation and tumorigenicity which are enhanced when *PAX-8* is highly expressed in tumor cells.

Taken together, our observations converge to the findings that the induction of *TTF-1* and *PAX-8* and their consequences on tumour growth depends on their basal states. Overexpression of one of these factors beyond a certain 'threshold' level could lead to pro-tumorigenic effects. Therefore, the modulation of these genes seems to be important to re-establish the differentiation balance.

To improve this hypothesis, we modulated the expression of *TTF-1* and *PAX-8* by several pharmacological agents. In the

case of *TTF-1*, our results support the observations reported in the literature showing that *TTF-1* is mainly regulated by epigenetic mechanisms such as methylation and acetylation but also by inhibitors of GSK-3 β and of proteasome (28,30). Whereas, *PAX-8* generally appeared to be less sensitive to different treatments; suggesting that other molecular mechanisms might also be involved in its regulation. Interestingly, demethylating agent (5'-AZA) was capable of inducing *PAX-8* only in ARO cell line, emphasizing that *PAX-8* regulation in anaplastic thyroid tumours is different from PTC.

The evaluation of these molecules on apoptosis showed that all the used molecules triggered early and/or late apoptosis or in some cases both. It is further important to highlight the fact that, even though some agents did not modulate gene expression, they succeeded in triggering apoptosis. Indeed, these pharmacological agents have a large spectrum and a wide range of targets (21,49,50). Epigenetic modulators such as HDACi or demethylating agents (TSA, VPA and 5'-AZA) used in our experiments are known to act directly on DNA conformation and therefore, on the regulation of several genes to induce inhibition of proliferation, cell cycle arrest, apoptosis and even senescence (49). In the same way, BOR or LiCl also act on several signaling pathways (51,52).

Taken together, our results suggest that different pathways could be involved in regulation of *TTF-1* and *PAX-8*. Therefore, the response to various treatments could be different according to the genetic background of the cell line studied. Nevertheless, we can speculate that an excessive induction of both transcription factors could lead to pro-tumorigenic effects e.g. treatment with 5'-AZA in anaplastic carcinomas could have opposite unintended effects. Indeed, 5'-AZA showed to increase, *in vitro*, iodine uptake and the expression of NIS (53) and has been tested in phase-I clinical trials. The positive effects of its use after clinical trials have not yet been published (<https://clinicaltrials.gov/ct2/show/NCT00004062?term=azacytidine+thyroid&rank=1>).

In conclusion, the present study improved our understandings regarding the role of thyroid transcription factors in tumorigenesis. We showed that both genetic and pharmacological approaches are able to induce expression of *TTF-1* and *PAX-8*. Thus, in future, it will be quite helpful to systematically take into account the basal state of these transcription factors while exploring tumour development. Moreover, it will also be important to consider *TTF-1* and *PAX-8* as part of a duo. These findings could open new therapeutic perspectives for the treatment of thyroid carcinomas.

Acknowledgements

We thank Professor Karim Benihoud and Dr Corinne Dupuy for their support and constructive discussions during the execution of the present study.

References

1. Civitareale D, Lonigro R, Sinclair AJ and Di Lauro R: A thyroid-specific nuclear protein essential for tissue-specific expression of the thyroglobulin promoter. *EMBO J* 8: 2537-2542, 1989.
2. Guazzi S, Price M, De Felice M, Damante G, Mattei MG and DiLauro R: Thyroid nuclear factor 1 (TTF-1) contains a homeodomain and displays a novel DNA binding specificity. *EMBO J* 9: 3631-3639, 1990.

3. Ordóñez NG: Value of thyroid transcription factor-1 immunostaining in distinguishing small cell lung carcinomas from other small cell carcinomas. *Am J Surg Pathol* 24: 1217-1223, 2000.
4. Chen J, Wall NR, Kocher K, Duclos N, Fabbro D, Neuberger D, Griffin JD, Shi Y and Gilliland DG: Stable expression of small interfering RNA sensitizes TEL-PDGFBetaR to inhibition with imatinib or rapamycin. *J Clin Invest* 113: 1784-1791, 2004.
5. Ziad A, Ruchala M, Breborowicz J, Gembicki M, Sowinski J and Grzymislawski M: Immunoeexpression of TTF-1 and Ki-67 in a coexistent anaplastic and follicular thyroid cancer with rare long-life surviving. *Folia Histochem Cytobiol* 46: 461-464, 2008.
6. Zhang P, Zuo H, Nakamura Y, Nakamura M, Wakasa T and Kakudo K: Immunohistochemical analysis of thyroid-specific transcription factors in thyroid tumors. *Pathol Int* 56: 240-245, 2006.
7. Ordóñez NG: Thyroid transcription factor-1 is a marker of lung and thyroid carcinomas. *Adv Anat Pathol* 7: 123-127, 2000.
8. Fenton CL, Patel A, Burch HB, Tuttle RM and Francis GL: Nuclear localization of thyroid transcription factor-1 correlates with serum thyrotropin activity and may be increased in differentiated thyroid carcinomas with aggressive clinical course. *Ann Clin Lab Sci* 31: 245-252, 2001.
9. Barletta JA, Perner S, Iafrate AJ, Yeap BY, Weir BA, Johnson LA, Johnson BE, Meyerson M, Rubin MA, Travis WD, *et al*: Clinical significance of TTF-1 protein expression and TTF-1 gene amplification in lung adenocarcinoma. *J Cell Mol Med* 13 (8B): 1977-1986, 2009.
10. Yamaguchi T, Hosono Y, Yanagisawa K and Takahashi T: NKX2-1/TTF-1: An enigmatic oncogene that functions as a double-edged sword for cancer cell survival and progression. *Cancer Cell* 23: 718-723, 2013.
11. Mu D: The complexity of thyroid transcription factor 1 with both pro- and anti-oncogenic activities. *J Biol Chem* 288: 24992-25000, 2013.
12. Hwang DH, Sholl LM, Rojas-Rudilla V, Hall DL, Shivdasani P, Garcia EP, MacConaill LE, Vivero M, Hornick JL, Kuo FC, *et al*: KRAS and NKX2-1 mutations in invasive mucinous adenocarcinoma of the lung. *J Thorac Oncol* 11: 496-503, 2016.
13. Takanashi Y, Tajima S, Hayakawa T, Neyatani H and Funai K: KRAS mutation-positive bronchial surface epithelium (BSE)-type lung adenocarcinoma with strong expression of TTF-1: A case providing a further insight as for the role of TTF-1 in the oncogenesis. *Int J Clin Exp Pathol* 8: 15338-15343, 2015.
14. Hu Y, Hartmann A, Stoehr C, Zhang S, Wang M, Tacha D, Montironi R, Lopez-Beltran A and Cheng L: PAX8 is expressed in the majority of renal epithelial neoplasms: An immunohistochemical study of 223 cases using a mouse monoclonal antibody. *J Clin Pathol* 65: 254-256, 2012.
15. Laury AR, Perets R, Piao H, Krane JF, Barletta JA, French C, Chiriac LR, Lis R, Loda M, Hornick JL, *et al*: A comprehensive analysis of PAX8 expression in human epithelial tumors. *Am J Surg Pathol* 35: 816-826, 2011.
16. Ordóñez NG: Value of PAX 8 immunostaining in tumor diagnosis: A review and update. *Adv Anat Pathol* 19: 140-151, 2012.
17. Nonaka D, Tang Y, Chiriboga L, Rivera M and Ghossein R: Diagnostic utility of thyroid transcription factors Pax8 and TTF-2 (FoxE1) in thyroid epithelial neoplasms. *Mod Pathol* 21: 192-200, 2008.
18. Presta I, Arturi F, Ferretti E, Mattei T, Scarpelli D, Tosi E, Scipioni A, Celano M, Gulino A, Filetti S, *et al*: Recovery of NIS expression in thyroid cancer cells by overexpression of Pax8 gene. *BMC Cancer* 5: 80, 2005.
19. Mu D, Huang R, Li S, Ma X, Lou C and Kuang A: Combining transfer of TTF-1 and Pax-8 gene: A potential strategy to promote radioiodine therapy of thyroid carcinoma. *Cancer Gene Ther* 19: 402-411, 2012.
20. Antonica F, Kasprzyk DF, Opitz R, Iacovino M, Liao XH, Dumitrescu AM, Refetoff S, Peremans K, Manto M, Kyba M, *et al*: Generation of functional thyroid from embryonic stem cells. *Nature* 491: 66-71, 2012.
21. Dokmanovic M, Clarke C and Marks PA: Histone deacetylase inhibitors: Overview and perspectives. *Mol Cancer Res* 5: 981-989, 2007.
22. Slingerland M, Guchelaar HJ and Gelderblom H: Histone deacetylase inhibitors: An overview of the clinical studies in solid tumors. *Anticancer Drugs* 25: 140-149, 2014.
23. Fortunati N, Catalano MG, Arena K, Brignardello E, Piovesan A and Boccuzzi G: Valproic acid induces the expression of the Na⁺/I⁻ symporter and iodine uptake in poorly differentiated thyroid cancer cells. *J Clin Endocrinol Metab* 89: 1006-1009, 2004.
24. Furuya F, Shimura H, Suzuki H, Taki K, Ohta K, Haraguchi K, Onaya T, Endo T and Kobayashi T: Histone deacetylase inhibitors restore radioiodide uptake and retention in poorly differentiated and anaplastic thyroid cancer cells by expression of the sodium/iodide symporter thyroperoxidase and thyroglobulin. *Endocrinology* 145: 2865-2875, 2004.
25. Pugliese M, Fortunati N, Germano A, Asioli S, Marano F, Palestini N, Frairia R, Boccuzzi G and Catalano MG: Histone deacetylase inhibition affects sodium iodide symporter expression and induces 131I cytotoxicity in anaplastic thyroid cancer cells. *Thyroid* 23: 838-846, 2013.
26. Puppini C, D'Aurizio F, D'Elia AV, Cesaratto L, Tell G, Russo D, Filetti S, Ferretti E, Tosi E, Mattei T, *et al*: Effects of histone acetylation on sodium iodide symporter promoter and expression of thyroid-specific transcription factors. *Endocrinology* 146: 3967-3974, 2005.
27. Vivaldi A, Miasaki FY, Ciampi R, Agate L, Collecchi P, Capodanno A, Pinchera A and Elisei R: Re-differentiation of thyroid carcinoma cell lines treated with 5-Aza-2'-deoxycytidine and retinoic acid. *Mol Cell Endocrinol* 307: 142-148, 2009.
28. Kondo T, Nakazawa T, Ma D, Niu D, Mochizuki K, Kawasaki T, Nakamura N, Yamane T, Kobayashi M and Katoh R: Epigenetic silencing of TTF-1/NKX2-1 through DNA hypermethylation and histone H3 modulation in thyroid carcinomas. *Lab Invest* 89: 791-799, 2009.
29. Raouane M, Desmaele D, Gilbert-Sirieix M, Gueutin C, Zouhiri F, Bourgaux C, Lepeltier E, Gref R, Ben Salah R, Clayman G, *et al*: Synthesis, characterization, and in vivo delivery of siRNA-squalene nanoparticles targeting fusion oncogene in papillary thyroid carcinoma. *J Med Chem* 54: 4067-4076, 2011.
30. Gilbert-Sirieix M, Makoukji J, Kimura S, Talbot M, Caillou B, Massaad C and Massaad-Massade L: Wnt/ β -catenin signaling pathway is a direct enhancer of thyroid transcription factor-1 in human papillary thyroid carcinoma cells. *PLoS One* 6: e22280, 2011.
31. Yamazaki CA, Padovani RP, Biscolla RP, Ikejiri ES, Marchetti RR, Castiglioni ML, Matsumura LK, Maciel RM and Furlanetto RP: Lithium as an adjuvant in the postoperative ablation of remnant tissue in low-risk thyroid carcinoma. *Thyroid* 22: 1002-1006, 2012.
32. Altmann A, Markert A, Askoxylakis V, Schöning T, Jesenofsky R, Eisenhut M and Haberkorn U: Antitumor effects of proteasome inhibition in anaplastic thyroid carcinoma. *J Nucl Med* 53: 1764-1771, 2012.
33. Oliva J, Dedes J, Li J, French SW and Bardag-Gorce F: Epigenetics of proteasome inhibition in the liver of rats fed ethanol chronically. *World J Gastroenterol* 15: 705-712, 2009.
34. Pohlenz J, Dumitrescu A, Zundel D, Martiné U, Schönberger W, Koo E, Weiss RE, Cohen RN, Kimura S and Refetoff S: Partial deficiency of thyroid transcription factor 1 produces predominantly neurological defects in humans and mice. *J Clin Invest* 109: 469-473, 2002.
35. Vilain C, Rydlewski C, Duprez L, Heinrichs C, Abramowicz M, Malvaux P, Renneboog B, Parma J, Costagliola S and Vassart G: Autosomal dominant transmission of congenital thyroid hypoplasia due to loss-of-function mutation of PAX8. *J Clin Endocrinol Metab* 86: 234-238, 2001.
36. Ali HM, Maksimenko A, Urbinati G, Chapuis H, Raouane M, Desmaële D, Yasuhiro H, Harashima H, Couvreur P and Massaad-Massade L: Effects of silencing the RET/PTC1 oncogene in papillary thyroid carcinoma by siRNA-squalene nanoparticles with and without fusogenic companion GALA-cholesterol. *Thyroid* 24: 327-338, 2014.
37. Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25: 402-408, 2001.
38. Urbinati G, Ali HM, Rousseau Q, Chapuis H, Desmaële D, Couvreur P and Massaad-Massade L: Antineoplastic effects of siRNA against TMPRSS2-ERG junction oncogene in prostate cancer. *PLoS One* 10: e0125277, 2015.
39. Ali HM, Urbinati G, Chapuis H, Desmaële D, Bertrand JR, Couvreur P and Massaad-Massade L: Effects of siRNA on RET/PTC3 junction oncogene in papillary thyroid carcinoma: From molecular and cellular studies to preclinical investigations. *PLoS One* 9: e95964, 2014.

40. Ouyang B, Knauf JA, Smith EP, Zhang L, Ramsey T, Yusuff N, Batt D and Fagin JA: Inhibitors of Raf kinase activity block growth of thyroid cancer cells with RET/PTC or BRAF mutations in vitro and in vivo. *Clin Cancer Res* 12: 1785-1793, 2006.
41. Ahn SH, Henderson Y, Kang Y, Chattopadhyay C, Holton P, Wang M, Briggs K and Clayman GL: An orthotopic model of papillary thyroid carcinoma in athymic nude mice. *Arch Otolaryngol Head Neck Surg* 134: 190-197, 2008.
42. van Staveren WC, Solís DW, Delys L, Duprez L, Andry G, Franc B, Thomas G, Libert F, Dumont JE, Detours V, *et al*: Human thyroid tumor cell lines derived from different tumor types present a common dedifferentiated phenotype. *Cancer Res* 67: 8113-8120, 2007.
43. Chen ST, Shieh HY, Lin JD, Chang KS and Lin KH: Overexpression of thyroid hormone receptor beta1 is associated with thyrotropin receptor gene expression and proliferation in a human thyroid carcinoma cell line. *J Endocrinol* 165: 379-389, 2000.
44. Zito G, Richiusa P, Bommarito A, Carissimi E, Russo L, Coppola A, Zerilli M, Rodolico V, Criscimanna A, Amato M, *et al*: In vitro identification and characterization of CD133(pos) cancer stem-like cells in anaplastic thyroid carcinoma cell lines. *PLoS One* 3: e3544, 2008.
45. Christophe-Hobertus C, Lefort A, Libert F and Christophe D: Functional inactivation of thyroid transcription factor-1 in PCC13 thyroid cells. *Mol Cell Endocrinol* 358: 36-45, 2012.
46. Miccadei S, De Leo R, Zammarchi E, Natali PG and Civitareale D: The synergistic activity of thyroid transcription factor 1 and Pax 8 relies on the promoter/enhancer interplay. *Mol Endocrinol* 16: 837-846, 2002.
47. Nitsch R, Di Dato V, di Gennaro A, de Cristofaro T, Abbondante S, De Felice M, Zannini M and Di Lauro R: Comparative genomics reveals a functional thyroid-specific element in the far upstream region of the PAX8 gene. *BMC Genomics* 11: 306, 2010.
48. Gilbert-Sirieix M and Massaad-Massade L: TTF-1: Neither angel nor demon. *Med Sci (Paris)* 27: 183-186, 2011 (In French).
49. Giannini G, Cabri W, Fattorusso C and Rodriguez M: Histone deacetylase inhibitors in the treatment of cancer: Overview and perspectives. *Future Med Chem* 4: 1439-1460, 2012.
50. Li KK, Li F, Li QS, Yang K and Jin B: DNA methylation as a target of epigenetic therapeutics in cancer. *Anticancer Agents Med Chem* 13: 242-247, 2013.
51. MacDonald BT, Tamai K and He X: Wnt/beta-catenin signaling: Components, mechanisms, and diseases. *Dev Cell* 17: 9-26, 2009.
52. Almond JB and Cohen GM: The proteasome: A novel target for cancer chemotherapy. *Leukemia* 16: 433-443, 2002.
53. Provenzano MJ, Fitzgerald MP, Krager K and Domann FE: Increased iodine uptake in thyroid carcinoma after treatment with sodium butyrate. *Otolaryngol Head Neck Surg* 137: 722-728, 2007.

Titre : Découverte de nouveaux transcrits de fusion dans des tumeurs pédiatriques en rechute et caractérisation fonctionnelle d'un nouvel oncogène LMO3-BORCS5

Mots clés : Transcrits de fusion, tumeurs solides pédiatriques en rechute, séquençage haut débit, oncogène de fusion LMO3-BORCS5

Résumé : Les cancers pédiatriques représentent la première cause de décès par maladie chez les enfants puisque 20% des patients présentent des résistances et rechutent. Les traitements actuels sont à leurs limites d'efficacité mais aussi de toxicité. L'amélioration des connaissances de la biologie des tumeurs résistantes et l'identification de nouvelles cibles moléculaires sont nécessaires pour ces pathologies rares, extrêmement diverses et mal comprises. Les oncogènes de fusion sont responsables de l'oncogenèse de ~ 20% des cancers. **En raison de la biologie particulière des cancers pédiatriques, nous avons émis l'hypothèse qu'ils seraient dus à des oncogènes de fusion.** Leur exploration, particulièrement à la rechute où les données actuelles sont limitées, permettrait de mieux comprendre ces pathologies et d'introduire de nouvelles cibles thérapeutiques. Notre premier objectif était **de détecter de nouveaux transcrits de fusion** parmi 48 patients pédiatriques en rechute inclus dans l'essai de médecine de précision MOSCATO-01 mené à Gustave Roussy. A partir de données de **RNA-seq**, nous avons détecté l'ensemble des fusions retrouvées dans ces tumeurs et classé les plus pertinentes selon leur fonction et leur potentiel thérapeutique. Un nouvel oncogène de fusion nommé **LMO3-BORCS5** retrouvé chez un patient atteint de sarcome d'Ewing dans les biopsies tumorales à la rechute et au diagnostic, et dans diverses lignées cellulaires a particulièrement retenu notre attention. Nos études fonctionnelles montrent que LMO3-BORCS5 a un rôle important dans la tumorigenèse et la sensibilité au traitement et que BORCS5 fonctionnerait, *in vivo*, comme un gène suppresseur de tumeur. **Ces résultats montrent le rôle critique des transcrits de fusion dans l'initiation et la progression tumorale ainsi que dans l'acquisition de résistance au traitement et que les fusions isolées ne doivent pas être ignorées.** A l'avenir, la prise en charge des cancers réside en une médecine moléculaire de précision, les oncogènes de fusions représentant l'un des outils les plus remarquables en clinique.

Title : Discovery of new fusion transcripts in pediatric tumors at relapse and functional characterization of a new oncogene LMO3-BORCS5

Keywords : Fusion transcripts, solid pediatric cancers at relapse, next-generation sequencing, LMO3-BORCS5 fusion oncogene.

Abstract: Pediatric cancers represent the first cause of death by disease in children as 20% of patients harbor resistances and die from recurrence of the malignancy. Nowadays treatment are at their limits of efficacy but also toxicity. Improving the knowledge on the biology of the relapsed/resistant tumors and identifying new molecular targets are needed for these rare, extremely diverse and misunderstood pathologies. Fusion oncogenes are known to be major drivers of oncogenesis, responsible for ~20% of cancers. **Due to their particular biology, we hypothesized that pediatric tumors would be more likely to harbor fusions** and that their exploration, especially at relapse where very limited data are available, could help to understand the particularity of these malignancies and consequently to provide new therapeutic targets. Our first aim was to **detect new fusion transcripts** in a cohort of 48 pediatric patients at relapse included in the precision medicine trial MOSCATO-01, previously conducted in Gustave Roussy. From the **RNA-seq** data, we described the landscape of fusions in these tumors and classified the most confident candidates according to their biological function and druggability predictions. A new fusion named **LMO3-BORCS5** found in biopsies of a patient with Ewing's sarcoma, at diagnosis and at relapse as well as in various cell lines particularly caught our attention. Functional studies showed that LMO3-BORCS5 has a high impact on tumorigenesis and treatment sensitivity and that BORCS5 would act, *in vivo*, as a tumor suppressor gene. **Our results strengthened the critical role of fusion transcripts in tumor initiation, progression and resistance and reinforces the idea that a private fusion oncogene should not be ignored.** Indeed, the future of cancer care lays in precision medicine, with fusion transcripts being one of the most remarkable tool for treatment decision

