



Utilisation de l'ADN circulant pour l'étude de l'hétérogénéité tumorale et l'évolution clonale en oncologie pédiatrique

Mathieu Chicard

► To cite this version:

Mathieu Chicard. Utilisation de l'ADN circulant pour l'étude de l'hétérogénéité tumorale et l'évolution clonale en oncologie pédiatrique. Cancer. Université Paris-Saclay, 2020. Français. NNT : 2020UP-ASL002 . tel-03154451

HAL Id: tel-03154451

<https://theses.hal.science/tel-03154451>

Submitted on 1 Mar 2021

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.

Utilisation de l'ADN circulant pour l'étude de l'hétérogénéité tumorale et l'évolution clonale en oncologie pédiatrique

Thèse de doctorat de l'université Paris-Saclay

**École doctorale n° 582 : cancérologie :
biologie - médecine - santé (CBMS)**

Spécialité de doctorat: Biochimie et Biologie Moléculaire
Unité de recherche : U830 – Cancer, Hétérogénéité, Instabilité et Plasticité (CHIP)
Référent : Faculté de médecine

Thèse présentée et soutenue à Paris, le 2 septembre 2020, par

Mathieu CHICARD

Composition du Jury

Pr. Gilles VASSAL

Professeur, Institut Gustave Roussy

Président

Pr. Frédérique PENAULT-LLORCA

Professeure, Centre Jean Perrin

Rapporteur & Examinatrice

Dr. Pierre SAINTIGNY

Docteur, Centre Léon Bérard

Rapporteur & Examineur

Dr. Valérie COMBARET

Docteure, Centre Léon Bérard

Examinatrice

Dr. Ludovic LACROIX

Docteur, Institut Gustave Roussy

Examineur

Dr. Gudrun SCHLEIERMACHER

Docteure, Institut Curie

Directrice de thèse

Titre : Utilisation de l'ADN circulant pour l'étude de l'hétérogénéité tumorale et l'évolution clonale en oncologie pédiatrique

Mots clés : ADN circulant, Biomarqueurs, Hétérogénéité, Plasticité, Cancers Pédiatriques

Résumé : Les cancers pédiatriques représentent un défi thérapeutique, et une compréhension des mécanismes d'échappement aux traitements est nécessaire pour pouvoir développer de nouvelles approches thérapeutiques.

L'ADN circulant peut être libéré par une tumeur dans les fluides corporels et permet de détecter et suivre des altérations génétiques tumorales par des prélèvements successifs peu invasifs tels qu'une prise de sang.

Dans ce travail, une technique de séquençage de type « whole exome » (séquençage de l'ensemble des exons) de l'ADN circulant a été mis au point pour permettre l'étude des altérations génétiques tumorales dans le plasma chez des enfants atteints de cancer.

Ces analyses soulignent la grande hétérogénéité génétique spatiale et temporelle des cancers pédiatriques. De plus, un rôle important de l'évolution clonale dans la progression de la maladie a ainsi pu être mis en évidence.

Des approches utilisant les caractéristiques particuliers de l'ADN circulant ont également permis d'inférer le profil d'expression, basées sur l'empreinte des sites de début de transcription, ou le profil épigénétique de la tumeur. En plus d'une aide à une classification des tumeurs, ces particularités pourront aider à l'observation d'un changement d'identité cellulaire sous traitement.

L'ADN circulant est donc un formidable outil pour mieux comprendre l'échappement aux traitements d'une tumeur par son hétérogénéité spatiale et temporelle et sa plasticité.

Title : Use of circulating DNA for the study of tumor heterogeneity and clonal evolution in pediatric oncology

Keywords : cfDNA, Biomarkers, Heterogeneity, Plasticity, Pediatrics Cancers

Abstract : Pediatric cancers represent a therapeutic challenge, and an understanding of the mechanisms of escape from treatment is necessary in order to be able to develop new therapeutic approaches. Circulating DNA can be released by a tumor in bodily fluids and enable to detect and follow genetic alterations in tumors by successive minimally invasive samples such as a blood test.

In this work, a "whole exome" sequencing technique for circulating DNA was developed to allow the study of genetic tumor alterations in plasma in children with cancer. These analyzes highlight the frequent spatial and temporal genetic heterogeneity of pediatric cancers. In addition, an important role of clonal evolution in the progression of the disease was thus highlighted.

Approaches using the particular characteristics of circulating DNA have also made it possible to infer the expression profile, based on the imprint of the transcription start sites, or the epigenetic profile of the tumor. In addition to an aid in the classification of tumors, these features may help to observe a change in cellular identity under treatment.

Circulating DNA is therefore an important tool for better understanding mechanisms of escape from treatment of a tumor based on spatial and temporal heterogeneity and cellular plasticity.

*I've seen things you people wouldn't believe...
Attack ships on fire off the shoulder of Orion...
I watched C-beams glitter in the dark near the Tannhäuser Gate...
All those moments will be lost in time, like...
tears in rain*

Roy Batty, 2019

Remerciements

A Gudrun, pour m'avoir poussé dans cette folle aventure, pour sa confiance, sa bienveillance et de m'avoir toujours motivé notamment avec la participation à des congrès aux quatre coins du monde mais surtout pour son incroyable patience et tolérance particulièrement vis-à-vis de mon caractère de cochon auvergnat.

A ma petite Mylène, qui m'a appris que l'amour frappe quand on si attend le moins et pour avoir veillé sur notre fils pendant la rédaction de ma thèse.

A Eiji, petit Roi de Ryokoku, pour son oreille musicale et son goût prononcé pour les patates.

A mes parents qui ont toujours eu du mal à comprendre mon métier mais qui restent fiers.

A mon frère Anthony, mécanicien du Liban, à Darlène, la discrète et à ma nièce Lena, la tornade qui pensent que je peux "réparer" des enfants.

A ma sœur Emma, chercheuse de lori en freelance.

A ma grand-mère Odette qui aurait aimé voir son petit-fils devenir Docteur ainsi qu'à mon grand-père Daniel et à ma tante Noëlle, qui m'ont appris que l'honnêteté devait toujours prévaloir malgré les conséquences ainsi que le rôle.

A ma grand-mère Monique, mon grand-père Arsène et sa compagne Anne qui m'ont appris l'importance de la famille.

A mes cousins, Nicolas le Bugarachois, Jérémy le panda, Benjamin la biscotte et Julien... l'insaisissable, tous petits par leurs âges mais grands par leurs tailles (sauf Nicolas), leurs cœurs; leurs appétits et leurs sens de l'humour gras.

A mes tantes Françoise et Sylvie et le reste de la famille immense par le nombre et le sens de la fête.

A Nathalie, pour son humour aussi noir que son café et sa culture encyclopédique. Une véritable amie dont le soutien m'a permis de ne pas craquer.

A Yasmine, toujours disponible et bon public qui malgré ses doutes arrivent à de petits miracles de bio-info.

Aux bio-informaticiens partis trop tôt...trop loin... enfin pas trop n'ont plus...dans Paris, Léo, Wilfried et Adrien réduisant la part Geek du labo.

A Sandrina, à qui je dois beaucoup des Dr. Pepper.

A Angela, sans qui l'Australie se serait révélée beaucoup plus sauvage.

A Irène, ma petite "vieille" M2, devenue grande et la pédiatre officielle d'Eiji.

A Céline et Zhi-Yan que j'ai toujours pris plaisir à dépanner pour pallier leurs lacunes en informatique liée à l'âge.

A l'équipe RTOP, Alex (Super Secrétaire), Jaydutt (I hope you don't end up deaf), Mamy (réduit un peu le sucre), Maria (Eiji est plus mignon), Nada (Bouh), Amaury, Solène, Chiara, Claire, Alice, Rachida, Charles et Nadia.

A tous les stagiaires que j'ai eu plaisir à former et partager mon travail notamment Austin et son Français aussi irréprochable que ses manip.

A Olivier Delattre, pour m'avoir accueilli au sein de son labo et pour ces remarques incroyablement pertinentes.

A Franck Bourdeaut, pour son incroyable dynamisme et sa confiance dans la réalisation du projet LCR et qui m'a permis de ne pas me sentir trop perdu dans l'immensité de Kyoto.

Au département de transfert, notamment Sergio et Dominique (la retraite n'est pas possible).

A Valérie Combaret, sans qui le début de ce projet n'aurait pas été possible mais également pour ces analyses toujours minutieuses.

A l'U830 et en particulier Carole (toujours rapide et impeccable), Keila et Didier ainsi que l'UGS en particulier Eve.

A la plateforme NGS, Patricia, Virginie, Virginie, Sonia, Sylvain, Laura et Benoit.

A la Plateforme Genom'IC, Franck, qui n'a pas hésité à m'embaucher malgré un précédent emploi à Micromania, Seb, Elodie, Big Flo et Juliana.

Au "petit" boulot qui permettent de croiser des personnes incroyables et "vraies" comme mon gamin Fabrice, Jenny, Danou et Manivone à MacDonalds, Julie, Julien et Stéphane à Toys'R Us, Emy, Julien et Camille à Micromania.

A tous mes "bounats" de Clermont, Nico Z, Raph, Damien, Teddy, Sophie, Kty, Marion, JB, Rémy et Pikachu.

A Christelle et Cyril, pour se moquer de ma prononciation de certains mots tels que *rose*, *jaune*, *gauche* ou *pauvre*.

A la famille Bohec, pour les spaetzles aux morilles.

A Izia la vieille carne, Plume de Duvet, Neptune le Boursouflash et Wicket le Vouvou.

A vous, le Jury, d'avoir pris du temps pour lire et écouter ma thèse.

Abréviations

aCGH : Puces d'Hybridation Génomique Comparative

ADN (DNA) : Acide DesoxyriboNucléique

ARN (RNA) : Acide RiboNucléique

ATRT : Tumeur Rhabdoïde Tératoïde Atypique

BAM : Binary Alignment Map

BED : Browser Extensible Data

cfDNA : Cell-Free DNA (ADN circulant)

ChIP : ImmunoPrecipitation de CHromatine

CNS : Système Nerveux Central

CMTB : Clinical Molecular Tumor Board

CR/PR : Rémission Complète/Partielle

ctDNA : Circulating Tumor DNA

ddPCR : Digital Droplet PCR

EDTA : ÉthylèneDiamineTétraAcétique

ETMR : Tumeur Embryonnaire avec Rosettes Multicouche

EWS : Sarcome d'Ewing

FACS : Fluorescence Activated Cell Sorting

FFPE : Formalin-Fixed, Paraffin-Embedded

FISH : Hybridation In Situ en Fluorescence

HP : Hépatoblastome

IGV : Integrative Genomics Viewer

LCR (CSF) : Liquide CéphaloRachidien

lcWGS : Low Coverage Whole Genome Sequencing

MAF : Fréquence Allélique Mutante

Mb: MégaBase

MB : Médulloblastome

MPNST : Tumeur Maligne des Gainés des Nerfs Périphériques

NB : Neuroblastome

NGS : Next Generation Sequencing

NOS : Not Otherwise Specified

nt : NucléoTide

pb (bp) : Paire de Base

PBS : Phosphate Buffered Saline

PCR : Polymerase Chain Reaction

PNET : Tumeur NeuroEctodermique Primitive

qPCR : PCR quantitative

RB : RétinoBlastome

RCF : Force Centrifuge Relative

RMS : Rhabdomyosarcome

RRBS : Reduced Representation Bisulfite Sequencing

SD/PD : Maladie Stable/Progressive

SNP : Single Nucleotide Polymorphism

SNV : Single Nucleotide Variant

WES : Whole Exome Sequencing

WGS : Whole Genome Sequencing

WT : Tumeur de Wilms

TSS : Site d'Initiation de la Transcription

UDI : Unique Dual Indexes

UMI : Unique Molecular Identifier

Sommaire

1	INTRODUCTION :	- 1 -
1.1	LES TUMEURS PEDIATRIQUES :	- 1 -
1.1.1	<i>Neuroblastome :</i>	- 3 -
1.1.1.1	Les altérations génétiques du neuroblastome	- 6 -
1.1.2	<i>Caractérisation moléculaire dans d'autres cancers pédiatriques</i>	- 11 -
1.1.2.1	Médulloblastome :	- 11 -
1.1.2.2	Rhabdomyosarcome	- 14 -
1.1.2.3	Tumeur d'Ewing	- 15 -
1.1.2.4	Tumeur de Wilms	- 16 -
1.2	ROLE DE LA GENETIQUE DANS L'ECHAPPEMENT AUX TRAITEMENTS	- 17 -
1.2.1	<i>Hétérogénéité et évolution clonale</i>	- 17 -
1.2.1.1	Mise en évidence de l'évolution clonale	- 18 -
1.2.1.2	Évolution clonale dans les cancers	- 20 -
1.2.1.3	Hétérogénéité et évolution clonale dans les tumeurs pédiatriques	- 24 -
1.2.2	<i>Changement d'identité cellulaire</i>	- 28 -
1.2.2.1	Type de plasticité	- 30 -
1.2.2.2	Plasticité cellulaire dans les tumeurs pédiatriques	- 31 -
1.3	LES BIOPSIES LIQUIDES	- 34 -
1.3.1	<i>Les biopsies liquides dans le cancer</i>	- 35 -
1.3.2	<i>Origine de l'ADN circulant :</i>	- 37 -
1.3.3	<i>Découverte de l'ADN circulant :</i>	- 39 -
1.3.4	<i>Le Diagnostic PréNatal Non Invasif (DPNI)</i>	- 41 -
1.3.5	<i>Autres utilisations de l'ADN circulant</i>	- 43 -
1.3.6	<i>Utilisation de l'ADN circulant pour la caractérisation et le suivi du cancer</i>	- 44 -
1.3.7	<i>Caractérisation moléculaire des cancers et classification par l'ADN circulant par des approches génomiques</i>	- 47 -
1.3.8	<i>Etude de l'évolution clonale par l'ADN circulant</i>	- 51 -
1.3.9	<i>Transcriptome et épigénétique de la tumeur à partir de l'ADN circulant</i>	- 54 -
1.3.10	<i>Biopsies liquides chez les enfants</i>	- 57 -
2	HYPOTHESE ET OBJECTIF	- 64 -

3	RESULTATS	- 66 -
3.1	QUANTITE D'ADN CIRCULANT SELON LE TYPE DE TUMEUR PEDIATRIQUE.....	- 67 -
3.1.1	<i>Extraction de l'ADN circulant, quantification et qualification.....</i>	- 67 -
3.1.2	<i>Quantités différentes selon les types de tumeurs</i>	- 69 -
3.2	SEQUENÇAGE D'EXOME A PARTIR DE L'ADN CIRCULANT	- 71 -
3.2.1	<i>Construction des librairies et capture d'exome</i>	- 71 -
3.2.2	<i>Pipeline bio-informatique.....</i>	- 73 -
3.3	ETUDE DE L'EVOLUTION CLONALE DU NEUROBLASTOME PAR L'ADN CIRCULANT.....	- 74 -
3.3.1	<i>Comparaison entre la tumeur primaire et l'ADN circulant au diagnostic</i>	- 74 -
3.3.2	<i>Analyse des plasmas à la rechute ou lors de rémission.....</i>	- 77 -
3.3.3	<i>Séquençage ciblé sur l'ADN circulant.....</i>	- 79 -
3.3.3.1	<i>Design et mise au point du séquençage ciblé.....</i>	- 79 -
3.3.3.2	<i>Suivi des altérations dans le plasma.....</i>	- 80 -
3.3.3.3	<i>Modélisation de l'évolution clonale</i>	- 82 -
3.4	EVOLUTION CLONALE DANS LES CANCERS PEDIATRIQUES	- 84 -
3.4.1	<i>Utilisation du WGS low coverage pour déterminer le contenu tumoral dans l'ADN libre.....</i>	- 84 -
3.4.2	<i>Comparaison des altérations génétiques détectées par WES entre la tumeur et le plasma</i>	- 86 -
3.4.2.1	<i>Neuroblastome.....</i>	- 86 -
3.4.2.2	<i>Rhabdomyosarcome</i>	- 90 -
3.4.2.3	<i>Sarcome d'Ewing</i>	- 92 -
3.4.2.4	<i>Tumeur du foie</i>	- 93 -
3.4.2.5	<i>MPNST.....</i>	- 95 -
3.4.2.6	<i>Tumeurs cérébrales</i>	- 96 -
3.4.2.7	<i>Rechute</i>	- 99 -
3.4.3	<i>Identification des voies de signalisation impactées par les altérations génétiques détectées dans la tumeur et le plasma</i>	- 102 -
3.4.4	<i>Séquençage ciblé et modélisation de l'évolution clonale</i>	- 103 -
3.4.4.1	<i>Neuroblastome.....</i>	- 104 -
3.4.4.2	<i>Rhabdomyosarcome</i>	- 112 -
3.4.4.3	<i>Sarcome d'Ewing</i>	- 115 -
3.4.4.4	<i>Tumeur du foie</i>	- 116 -
3.4.4.5	<i>MPNST.....</i>	- 118 -

3.4.4.6	Tumeurs cérébrales	- 118 -
3.4.5	<i>Conclusion</i>	- 119 -
3.5	UTILISATION DE L'ADN CIRCULANT POUR L'ETUDE DU TRANSCRIPTOMIQUE, DE L'EPIGENETIQUE ET DE LA PLASTICITE	- 120 -
3.5.1	<i>Prédiction d'expression se basant sur l'empreinte des TSS dans l'ADN circulant</i>	- 120 -
3.5.1.1	Design d'un panel.....	- 121 -
3.5.1.2	Pipeline bio-informatique	- 122 -
3.5.1.3	Prédiction.....	- 123 -
3.5.2	<i>ChIPSeq sur l'ADN circulant</i>	- 128 -
3.5.2.1	Protocole de ChIPSeq	- 128 -
3.5.2.2	Peak Calling.....	- 130 -
3.5.3	<i>Conclusion</i>	- 133 -
3.6	APPLICATION CLINIQUE DE L'ADN CIRCULANT	- 134 -
3.6.1	<i>Recherche d'altérations ciblables à la rechute à des fins thérapeutiques : Etude MappyActs</i> -	134 -
3.6.1.1	Protocole de WES.....	- 135 -
3.6.1.2	Quantité de cfDNA et ctDNA.....	- 136 -
3.6.1.3	Comparaison des analyses en WES entre la tumeur et le plasma.....	- 137 -
3.6.2	<i>Caractérisation et classification des tumeurs cérébrales</i>	- 144 -
3.6.2.1	Extraction des liquides céphalo-rachidien	- 144 -
3.6.2.2	Construction des librairies	- 145 -
3.6.2.3	Analyse du séquençage.....	- 145 -
3.6.2.4	Rhabdomyosarcome	- 149 -
3.6.2.5	Design du panel "Tumeur Cérébrale Embryonnaire"	- 152 -
3.6.2.6	Classification des échantillons.....	- 153 -
3.6.3	<i>Classification par le méthylome</i>	- 156 -
3.6.3.1	Choix du RRBS	- 156 -
3.6.3.2	Classification des tumeurs pédiatriques	- 157 -
4	DISCUSSION, PERSPECTIVES ET CONCLUSION	- 159 -

BIBLIOGRAPHIE

ANNEXE

1 Introduction :

1.1 Les tumeurs pédiatriques :

Chaque année, environ 1750 cas de cancer pédiatrique sont diagnostiqués en France. Les cancers de l'enfant demeurent rares puisqu'ils ne représentent que 1 à 2% de l'ensemble des cancers. Les cancers pédiatriques les plus fréquentes sont les leucémies mais 65% des cancers sont des tumeurs solides telles que les tumeurs cérébrales ou le neuroblastome (Figure 1.1).

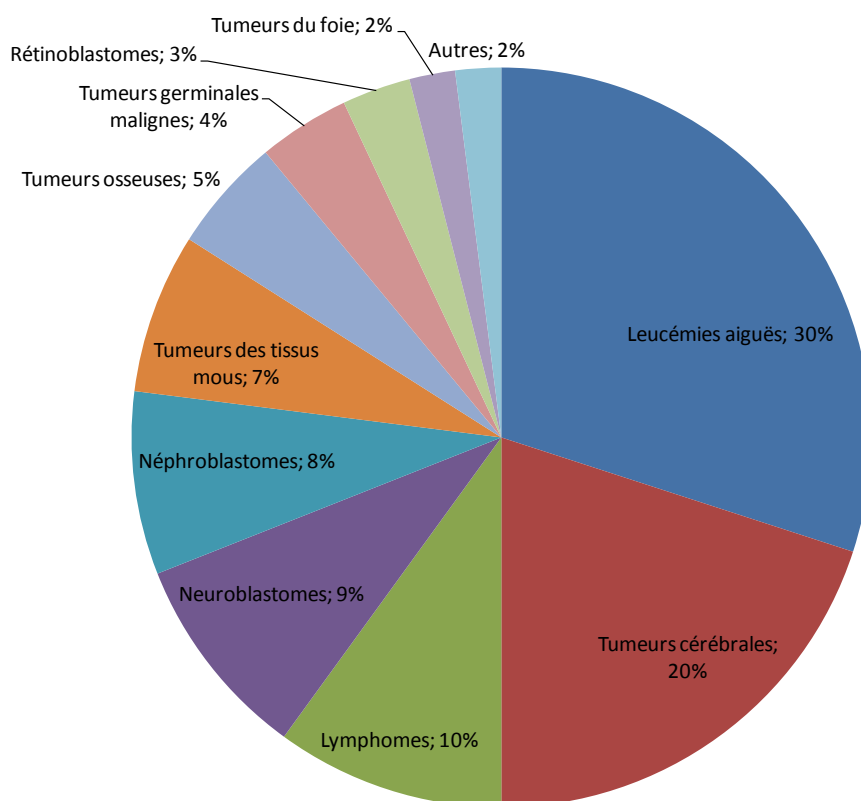


Figure 1.1 : Répartition des tumeurs pédiatriques en France (source : Collège National des Pédiatres Universitaires (CNPU)).

Les types histologiques de ces tumeurs sont très différents de ceux des cancers de l'adulte et ces tumeurs ne sont généralement présentes que chez l'enfant. La plupart des tumeurs solides sont dites embryonnaires et vont toucher des organes en développement. Ces tumeurs vont ainsi avoir une croissance souvent très rapide sans que ce facteur soit forcément

lié à la gravité de la maladie. Pour certains types de tumeurs, notamment chez les nourrissons, des régressions spontanées sont observées¹. La plupart des tumeurs solides ont pour origine des tissus embryonnaires et se développent dans une fenêtre de temps particulière, y compris pendant le développement embryonnaire.

En l'espace de quelques décennies, grâce aux traitements multimodaux (traitement local par chirurgie ou radiothérapie, chimiothérapie, chimiothérapie haute dose, et plus récemment les traitements ciblés et immunothérapies), la survie globale à 5 ans des enfants atteints de ces tumeurs est passée à plus de 80 %². Mais d'un côté, certaines tumeurs ne répondent pas aux traitements proposés, et parfois des rechutes très agressives peuvent être observées. D'autre part, certains traitements restent très lourds peuvent être associés à des séquelles à moyen ou à long terme.

Les approches de caractérisation moléculaire ont donc eu pour objectif soit d'identifier des biomarqueurs diagnostiques, pronostiques et prédictifs afin d'améliorer la classification de ces tumeurs, soit d'identifier des cibles thérapeutiques moléculaires. En règle générale, les tumeurs pédiatriques sont généralement peu mutées (entre 0,17 à 0,79 mutations somatiques par MégaBase (MB) contre 1 à 10 par MB pour les tumeurs de l'adulte)³ impliquant moins de la moitié des oncogènes récurrents dans les cancers de l'adulte tels que *NRAS*, *KRAS* ou *TP53*⁴ (Figure 1.2). Peu de signatures mutationnelles, c'est-à-dire de combinaisons de mutations résultant d'une mutagénèse spécifique telles que la réparation de l'ADN ou l'exposition aux UV, ont été observées en dehors de celles liées à l'âge mais aussi à la recombinaison homologue pour les ostéosarcomes^{3,5}. Dans la plupart des cancers pédiatriques, une augmentation significative du nombre d'altérations est observée en cas de rechute avec parfois des enrichissements tels que la voie RAS-MAPK dans le neuroblastome^{6,7}.

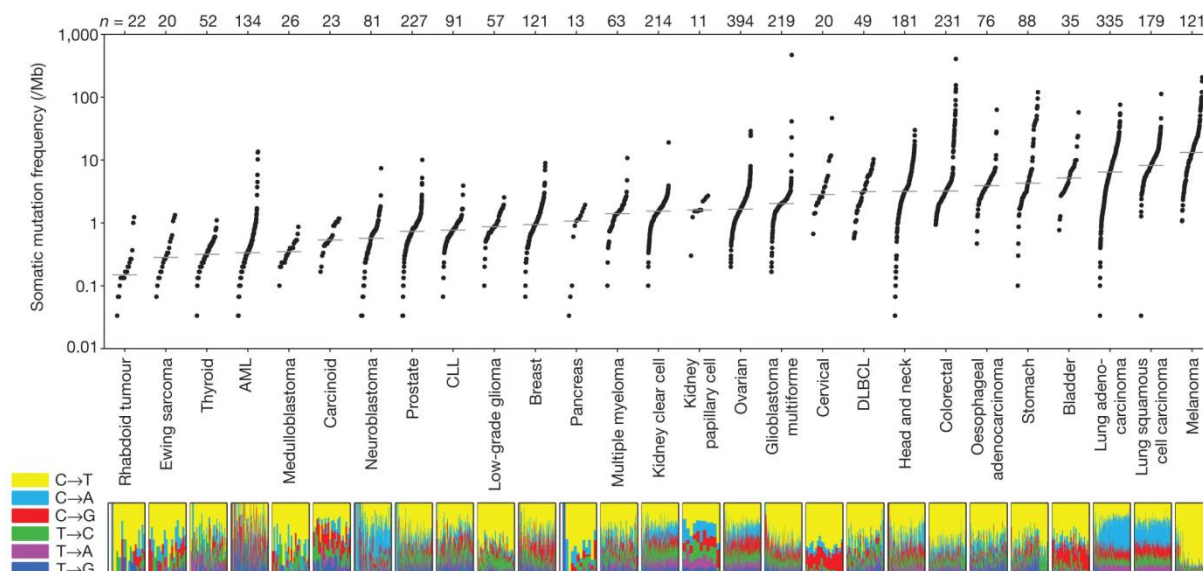


Figure 1.2 : Charges mutationnelles par type de tumeurs (Figure en accès libre).

Les tumeurs pédiatriques ont des fréquences de mutations somatiques très faibles comparées aux tumeurs de l'adulte: Tumeur rhabdoïde (0 à 37 mutations par cas), sarcome d'Ewing (2 à 40 mutations), médulloblastome (6 à 26 mutations) et neuroblastome (0 à 222 mutations)⁸.

1.1.1 Neuroblastome :

Le neuroblastome est la tumeur solide extracraniale la plus fréquente de l'enfant et touche environ 120 – 150 enfants par an en France. Le neuroblastome est une tumeur qui apparaît lors du développement du système nerveux sympathique et qui est issue de la crête neurale⁹. La tumeur se développe majoritairement dans la glande surrénale et/ou dans les ganglions sympathiques⁹.

L'origine du neuroblastome n'est pas connue mais serait due à la différenciation des progéniteurs sympathoadrénux en cellules de ganglion sympathique et en cellules chromaffines¹⁰ (Figure 1.3).

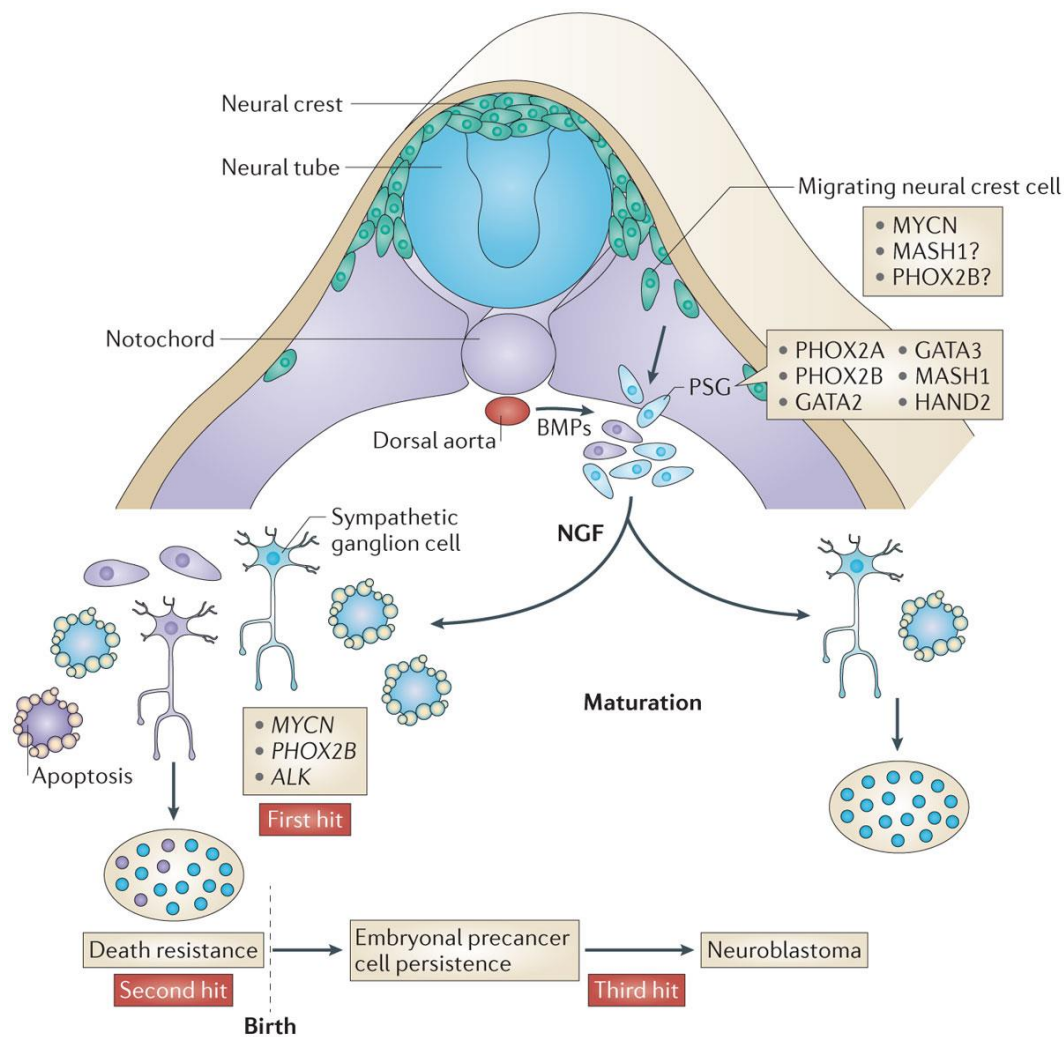


Figure 1.3 : Origine du neuroblastome¹⁰

Sous l'influence de *MYCN* et des protéines morphogénétiques de la moelle (BMP), les progéniteurs des neuroblastes vont migrer de la crête neurale au tube neural. Un premier événement touchant *MYCN* ou dû à des mutations constitutionnelles de *ALK* ou de *PHOX2B* empêcherait la différenciation des neurones sympathiques. Ces cellules embryonnaires resteraient persistantes après la naissance. Un autre événement pour le moment inconnu amènerait à la malignité de ces cellules et au développement de neuroblastomes.

Le neuroblastome est une tumeur extrêmement hétérogène d'un point de vue clinique et génétique.

Les tumeurs neuroblastiques périphériques présentent différents degrés de différenciations, distinguant ainsi le ganglioneurome, (les cellules tumorales présentent une différenciation terminale), le ganglioneuroblastome (certaines cellules tumorales présentent la différenciation neuronale, d'autres correspondent aux cellules ganglionnaires), et le neuroblastome (principalement des petites cellules tumorales immatures, rondes). La prise en compte de la composante des cellules Schwanniennes stromales permettra de classer les

tumeurs neuroblastiques en quatre catégories : neuroblastome (stroma pauvre), ganglioneuroblastome mélangé (stroma riche), ganglioneuroblastome nodulaire et ganglioneurome (stroma dominant)¹¹.

Les symptômes cliniques du neuroblastome sont directement liés à la localisation de la tumeur primaire et des sites métastatiques. Une maladie métastatique est détectée dans environ 50% des patients au moment du diagnostic, avec le plus souvent une atteinte des ganglions lymphatiques régionaux, une atteinte médullaire ou osseuse, alors qu'une atteinte hépatique ou cutanée peut être observée en particulier chez les nourrissons.

Les stades localisés de la maladie (stade 1, 2 ou 3 de l'ancienne classification INSS, stades L1 et L2 de la nouvelle classification INRG) et les stades 4s (Ms pour la classification INRG) auront de meilleures chances de guérison avec une survie à 5 ans généralement au-dessus de 80%^{12,13}. En revanche, les stades plus avancés et métastatiques (INSS stade 4, INRG Stade M) auront des pronostics plus sombres avec une survie globale à 5 ans à 40-50% pour les stades M de haut risque. Mais la classification en tumeurs de haut risque ne sera pas seulement dépendante du stade mais de l'âge du patient (les enfants de moins de 18 mois ayant un meilleur pronostic que les enfants âgés de plus de 18 mois au diagnostic) et de certaines altérations génétiques telles que l'amplification du gène *MYCN* et de l'âge du patient. Certains stades métastatiques très avancés touchant des enfants très jeunes (< 1 an) appelés 4s peuvent également être amenés à régresser spontanément.

Les traitements des patients atteints de neuroblastomes peuvent ainsi aller d'un simple suivi observatoire à des traitements plus lourds couplant chimiothérapie, chimiothérapie haute-dose, chirurgie, radiothérapie, immunothérapie et chimiothérapie de maintenance¹⁴.

1.1.1.1 Les altérations génétiques du neuroblastome

Le neuroblastome possède peu d'altérations génétiques récurrentes^{6,15-17}. Dans 20% des cas, une amplification du gène *MYCN* est observée. La présence d'une amplification de *MYCN* est associée à un pronostic sombre, ainsi la détermination du statut d'amplification de *MYCN* fait partie du bilan diagnostique de routine¹⁸⁻²⁰.

Le gène *MYCN* code un facteur de transcription impliqué dans de nombreux rôles en régulant certains gènes de la prolifération tels que *CDK4* ou *CHK1*, de la migration (ou métastase) en activant ou désactivant des gènes tels que les intégrines mais également dans la différenciation, la pluri-potence ou l'angiogenèse²¹.

Le gène *MYCN* a également un rôle important dans l'apoptose et son amplification amène à une surexpression de la voie MDM2-TP53, médiateur de l'apoptose^{22,23}. Les cellules *MYCN* amplifiées sont ainsi très proliférantes mais également très apoptotiques.

L'amplification du gène *MYCN* semble aussi influencer l'épigénétique de la cellule avec un rôle important dans la maintenance de la chromatine active²⁴. *MYCN* interagit avec EZH2, impliqué dans la triméthylation de l'histone 3 K27 mais le facteur de transcription est aussi associé à une augmentation de l'expression d'histones acétylases^{25,26}. Des interactions avec TETs ou DNMT3A impliqués dans la méthylation de l'ADN ont aussi été observées laissant penser que *MYCN* pourrait ainsi augmenter la répression de certains gènes suppresseurs de tumeurs en augmentant la méthylation de l'ADN ou la désacétylation des histones^{27,28}.

Le neuroblastome est également une tumeur très remaniée au niveau chromosomique et le pronostic peut être très différent en fonction du type d'altération. En effet, les tumeurs dont le profil chromosomique est dit "numérique", c'est-à-dire avec des pertes ou des gains de chromosomes entiers, sont plutôt de bon pronostic. Pour les tumeurs dont le profil est dit "segmentaire", c'est-à-dire avec des pertes ou des gains partiels de chromosomes, celles-ci

sont de pronostic sombre^{29,30}. Le gain du 17q et les pertes du 1p et 11q sont le plus souvent observés et sont liés dans la majorité des cas à une amplification du gène *MYCN*^{31,32}. D'autres altérations sont également observées telles que le gain du 1q et 2p, la perte 3p, 4p et 14q, cependant un pronostic pauvre ne semble pas être associé à la présence de ces altérations. Lors de la rechute, des pertes du 1p et 6q sont fréquentes^{6,16}.

Le pronostic sombre associé à ces gains et notamment à ces pertes partielles de chromosome pourrait être dû au fait de la présence de gènes suppresseurs de tumeurs tels que *CHD5*, *CAMTA1*, *KIF1B*, *CASZ1* et miR-34a présents notamment sur le 1p³³.

Dans 6 à 12% des cas de neuroblastomes, le gène *ALK* (Anaplastic Lymphoma Kinase) est muté³⁴⁻³⁶. Ce gène code un récepteur tyrosine kinase et est un membre de la famille des récepteurs à l'insuline³⁷. Sa dimérisation et son autotransphosphorylation jouent un rôle important dans les voies RAS/MAPK, PI3K/AKT et JAK/STAT³⁸⁻⁴⁰.

Les mutations, dans son domaine kinase (F1174, F1245 et R1275) représentent 85% des mutations somatiques trouvées dans le neuroblastome et vont empêcher l'auto-inhibition de *ALK*^{34,41}. En plus d'une dérégulation des voies précédemment citées, l'activation de *ALK* dans le neuroblastome aurait aussi un rôle dans la stimulation de la transcription du gène *MYCN* mais également dans la voie RET conduisant à une prolifération cellulaire importante lorsque ces deux altérations sont observées⁴²⁻⁴⁵. Une activation de *ALK* amène ainsi à la phosphorylation des protéines SHC et à la dérégulation de l'apoptose^{46,47}.

Dans 2% des cas, le gène *ALK* peut également être amplifié. L'amplification du gène *ALK* est souvent associée à une amplification du gène *MYCN*⁴⁸⁻⁵⁰. Cependant le lien entre ces deux amplifications n'a pas été défini.

Ces altérations de *ALK* peuvent également être présentes à un niveau sous clonal au moment d'un diagnostic et émerger au moment de la rechute laissant penser que des sous-

clones portant une mutation de *ALK* seraient plus résistants aux traitements et conduiraient à une rechute plus sévère^{36,51}.

D'autres altérations génétiques récurrentes sont observées dans le neuroblastome telles des pertes de fonction d'*ATRX* ou des réarrangements de *TERT* peuvent être observés, le plus souvent en l'absence d'amplification de *MYCN*^{15,52-56}. Bien qu'intervenant dans le remodelage de la chromatine, les modifications de *ATRX* tout comme celle de *TERT*, responsable du maintien de longueurs des télomères, conduisent à des longueurs alternatives des télomères (ALT)⁵⁷. Les cellules ainsi touchées ne rentreront pas en sénescence et continueront à se diviser rapidement^{58,59}.

D'autres gènes impliqués dans le remodelage de la chromatine et membres du complexes SWI/SNF ont été détecté comme étant mutés dans le neuroblastome notamment les gènes *ARID1A* et *ARID1B*^{60,61}. Cependant même si ces mutations semblent avoir un impact sur la fonction de la protéine, leurs effets sur la structure de la chromatine n'ont pas été définis. Il en est de même pour le gène *SMARCA4*, plutôt impliqué dans le développement des tumeurs rhabdoïdes, pour lequel des mutations ont été détectés mais leurs implications dans le neuroblastome restent inconnues^{60,62}.

Le neuroblastome est une tumeur très hétérogène et les patients atteints de tumeurs de haut risque ont tendance à rechuter avec l'apparition de nouvelles et nombreuses altérations au niveau de la rechute (Figure 1.4)^{6,30,51,63}. Ces altérations observées à la rechute sont très souvent des mutations dans la voie RAS-MAPK touchant des gènes tels que *ALK*, *PTPN11*, *NF1*, *BRAF*, *HRAS*, *NRAS* et *KRAS* mais aussi des délétions de gènes telles que *CDKN2A*.

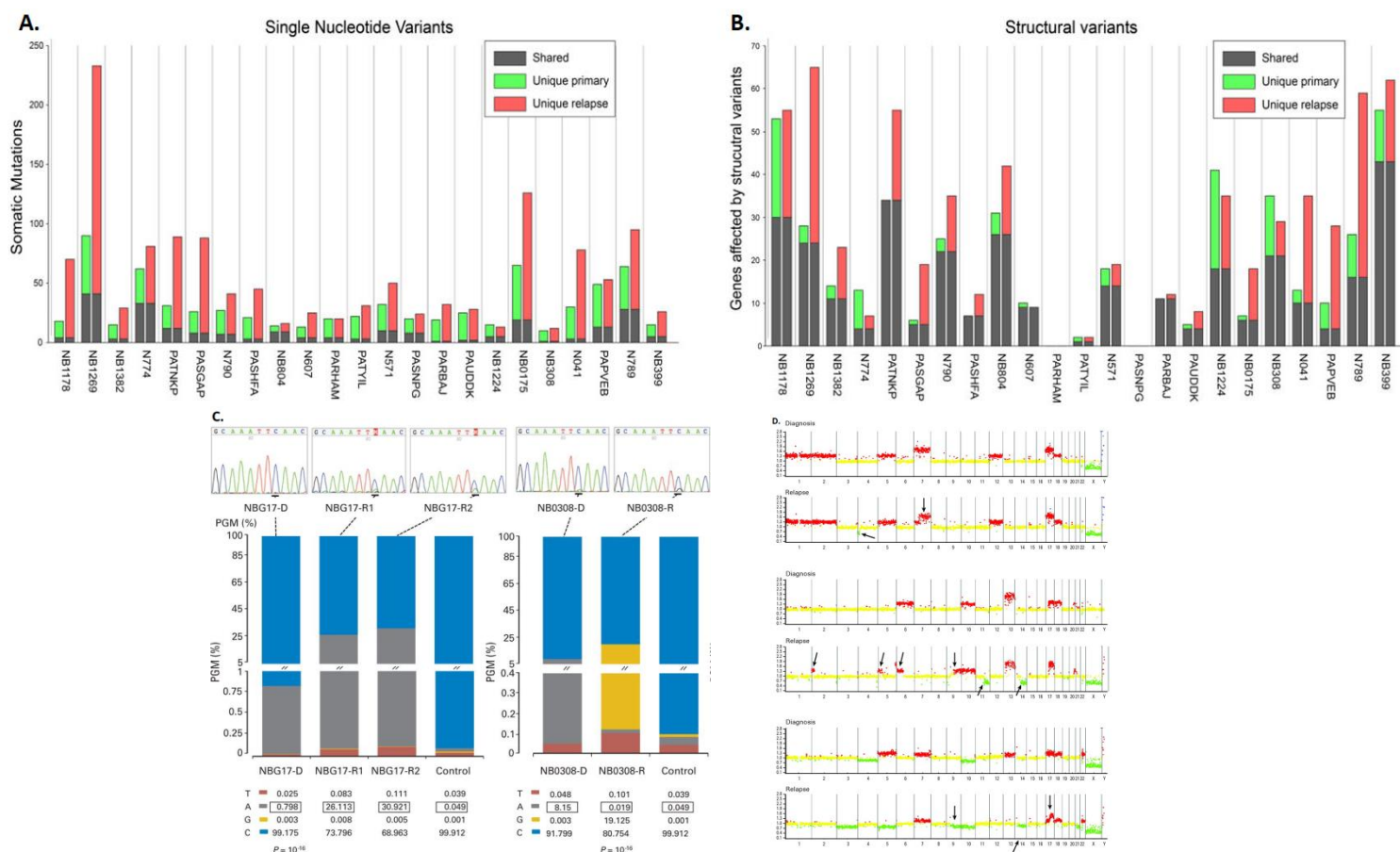


Figure 1.4 : Augmentation des altérations entre diagnostic et rechute pour le neuroblastome. Nombre de mutations non synonymes (A.) et variants structuraux (B.) détectés par WGS (Whole Genome Sequencing) de 23 tumeurs primaires et leurs rechutes. En gris, les altérations partagées par la tumeur primaire et la rechute, en vert les altérations seulement présentes dans la tumeur primaire et en rouge la tumeur rechute⁶.

Emergence de mutation *ALK* F1174L entre diagnostic et rechute (C.). Pour le patient NBG17, la mutation a seulement été détectée à la rechute par Séquençage Sanger mais était présente à un niveau sous clonal au diagnostic et détecté par un séquençage à très grande profondeur par NGS Ion torrent PGM. Pour le patient, NB0308 une nouvelle mutation *ALK* est détecté à la rechute et totalement absente dans la tumeur primaire⁵¹.

Comparaison de CGH array réalisées sur trois paires de tumeurs au diagnostic et à la rechute (D.). Les nouvelles altérations apparaissant à la rechute sont indiquées par les flèches noires (4p et 7q pour le premier cas, 2p, 5q, 6p, 9q, 11q et 14q pour le second et 9q, 14q et 17q pour le troisième)³⁰.

Ces récentes caractérisations dans le neuroblastome ont amené à une nouvelle proposition de classement des neuroblastomes de haut risque⁶⁴⁻⁶⁶. En plus de l'âge, stade et des altérations en nombre de copie ; la taille des télomères ou de l'activité de la télomérase ainsi que la présence de mutations dans les voies RAS/p53 sont pris en compte. Les neuroblastomes sont ainsi classés en haut risque et très haut risque.

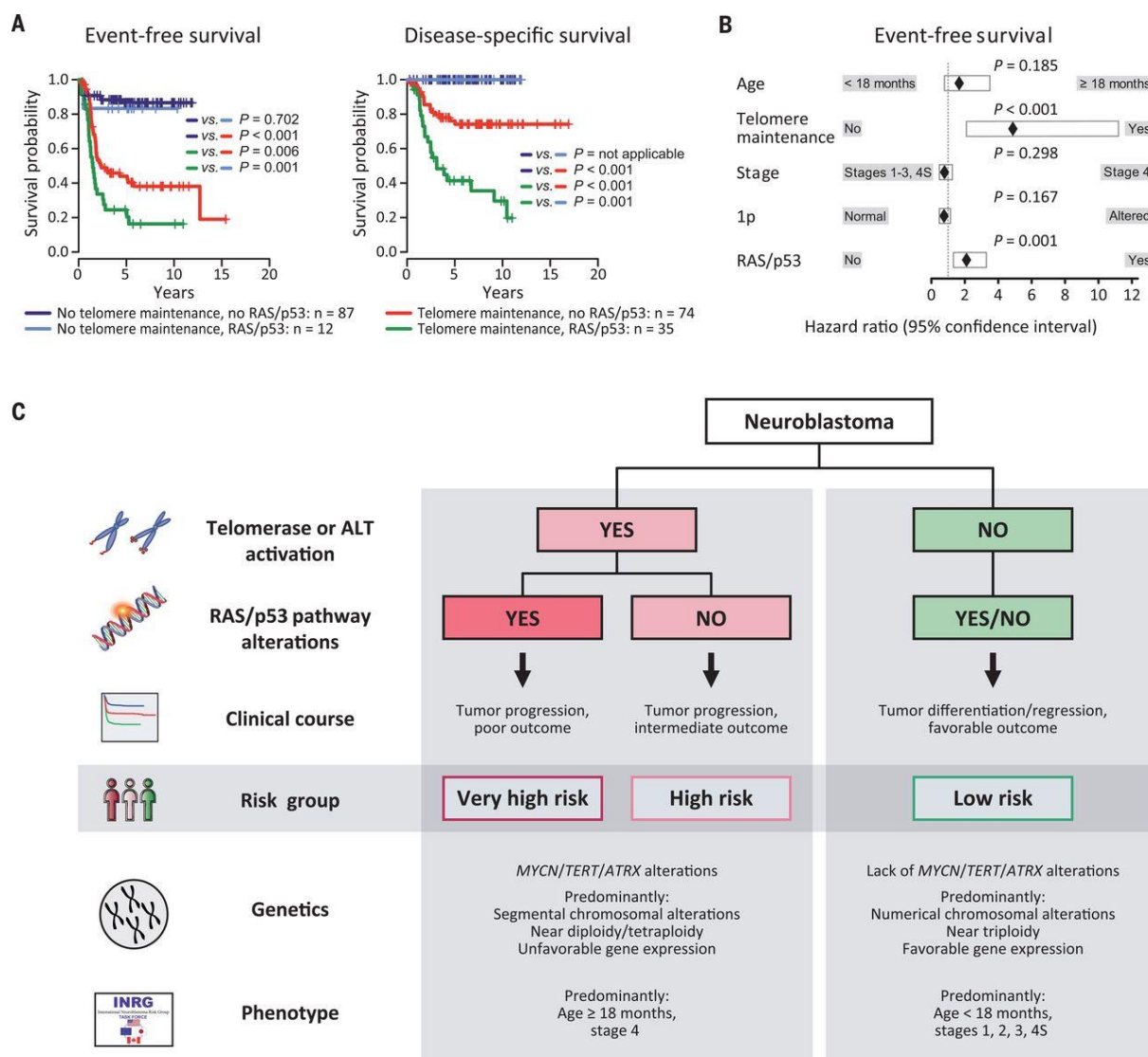


Figure 1.5. Nouvelle classification des neuroblastomes de haut-risque (Figure en accès libre)⁶⁵.
 Courbe de survie sans événement et maladie spécifique en présence ou absence de mutations dans les voies RAS/p53 et l'activité de maintenance de télomères (A.). Régression de Cox multivariée considérant l'âge, la maintenance des télomères, le stade, le statut du chromosome 1p et les mutations dans les voies RAS/p53 (B.). Schéma de la nouvelle classification clinique de sous-groupe de neuroblastome (C.).

La prédisposition héréditaire reste très rare avec seulement 1% des cas de neuroblastome possédant des mutations touchant *ALK* ou *PHOX2B*^{34,41,67-69}.

La caractérisation du neuroblastome reste incomplète. D'autres mécanismes pourraient expliquer son agressivité. Les nouvelles techniques de séquençage ont mis en lumière dans les tumeurs de haut risque une modification d'expression d'ADN longs non codants conduisant à une sur- ou sous-expression de gènes impliqués dans le développement du neuroblastome tels que *MYCN*, *ALK* ou *PHOX2B*⁷⁰.

L'analyse ChIP-Seq de MYCN a démontré que la dérégulation des transcriptions du gène *MYCN* suite à son amplification ne serait possible sans la présence du facteur de transcription TWIST1⁷¹.

Le séquençage long-read possible avec les technologies PacBio et Nanopore a pu décrire une structure circulaire extrachromosomique de l'amplification de *MYCN* déstabilisant et remodelant l'intégrité du génome⁷².

Les altérations génétiques du neuroblastome sont des informations importantes pour sa classification permettant d'établir un pronostic. En fonction de cette classification, un traitement adapté peut être proposé au patient. Cependant ces altérations génétiques peuvent évoluer au fil du temps rendant l'évolution de la maladie parfois imprévisible.

A l'instar du neuroblastome, la caractérisation des altérations génétiques est aussi une grande importance dans les cancers pédiatriques.

1.1.2 Caractérisation moléculaire dans d'autres cancers pédiatriques

1.1.2.1 Médulloblastome :

Le médulloblastome est un cancer embryonnaire du cervelet et des structures adjacentes. Ce sont les tumeurs cérébrales les plus fréquentes chez les enfants et ils sont classés par groupes moléculaires distincts : le groupe WNT, le groupe Sonic Hedgehog (SHH) où leurs noms sont due à des modifications dans des voies bien définies et les groupes 3 et 4 où les causes de la maladie sont plus pléiotropiques et dont les gènes et les voies modifiés ne sont pas clairement identifiés (Figure 1.6)¹⁰.

Les tumeurs du groupe WNT touchent principalement des enfants et des jeunes adultes (à partir de 4 ans avec une moyenne d'âge de 11 ans)⁷³. Le pronostic de ce groupe est très favorable avec un taux de survie à 5 ans de 95% avec peu de risque de métastases. Son histologie est dite classique caractérisée par des paquets denses de petites cellules rondes ou ovales avec un ratio noyau:cytoplasme très élevé. Le gène *CTNNB1* est muté somatiquement dans 85% des cas et une mutation constitutionnelle du gène *APC* est aussi observée de manière exclusive^{74,75}. Une perte du chromosome 6 est aussi caractéristique du groupe WNT et des mutations dans le complexe de remodelage de la chromatine SWI/SNF sont également présentes dans 33% des cas.

Le groupe SHH possède une distribution d'âge assez particulière puisqu'il va toucher majoritairement des enfants de moins de 3 ans et de jeunes adultes de plus de 17 ans et plutôt de sexe masculin. Le taux de survie est bon avec un risque modéré de métastases Son histologie peut être classique ou desmoplastique caractérisée par des îlots de cellules de différenciation neurocytique. Les gènes de la voie Sonic Hedgedog sont majoritairement touchés avec des mutations ou insertions/délétions dans *PTCH1*, des mutations de *SUFU* ou *SMO*. Des amplifications de *GLI2* et des mutations dans le promoteur *TERT* sont également observées⁷⁶⁻⁷⁸. Des altérations en nombre de copie telles que la perte du 9q, 10q, 14q, 17q et des gains du chromosomes 2 et 9p sont fréquentes.

Subgroup		WNT	SHH			
Subtype			α	β	γ	δ
Demographics	Frequency (%)	100	29	16	21	34
	Age (bar height corresponds with percentage)					
	Gender (%)	45 ♂ 55 ♀	63 ♂ 37 ♀	47 ♂ 53 ♀	55 ♂ 45 ♀	69 ♂ 31 ♀
Clinical features	Histology	Classic	Classic > desmoplastic > LCA	Desmoplastic > classic	Desmoplastic > MBEN classic	Classic > desmoplastic
	Metastasis (%)	12	20	33	9	9
	5-year OS (%)	98	70	67	88	89
Molecular features	Cytogenetics					
	Driver events	CTNNB1, DDX3X or SMARCA4 mutation	- MYCN or GLI2 amplification - TP53 mutation - PTCH1 mutation (less)	- PTCH1 or KMT2D mutation - SUFU mutation/deletion - PTEN deletion	- PTCH1, SMO or BCOR mutation - PTEN deletion	- PTCH1 mutation - TERT promoter mutation

Subgroup		Group 3							Group 4
Subtype		I	II	III	IV	V	VI	VII	VIII
Demographics	Frequency (%)	4	13	9	10	8	9	22	25
	Age (bar height corresponds with percentage)								
	Gender (%)	60 ♂ 40 ♀	77 ♂ 23 ♀	78 ♂ 22 ♀	68 ♂ 32 ♀	71 ♂ 29 ♀	67 ♂ 33 ♀	66 ♂ 34 ♀	75 ♂ 25 ♀
Clinical features	Histology	Classic > desmoplastic	LCA, classic	Classic > LCA	Classic	Classic	Classic	Classic	Classic
	Metastasis (%)	35	57	56	58	62	45	45	50
	5-year OS (%)	77	50	43	80	59	81	85	81
Molecular features	Cytogenetics								
	Driver events	- GFI1 and GFI1B activation - OTX2 amplification	- MYC amplification - GFI1 and GFI1B activation - KBTBD4, SMARCA4, CTDNEP1 or KMT2D mutation	MYC amplification (less)	No common driver events	MYC or MYCN amplification	- PRDM6 activation - MYCN amplification (less)	KBTBD4 mutation	- PRDM6 activation - KDM6A, ZMYM3 or KMT2C mutation

Figure 1.6 : Classification des médulloblastomes (Figure en accès libre)⁷³.

Les groupes 3 et 4 sont plus difficilement caractérisables et touchent rarement de jeunes adultes. La classification des tumeurs dans ces groupes est plutôt due à leur non-classification dans les deux précédents groupes. Malheureusement les médulloblastomes de groupe 3 et 4 sont associés à un pronostic plutôt défavorable et un risque élevé de présenter des métastases.

Le groupe 3 serait donc caractérisé plutôt par une amplification de *MYC*, une expression *GFI1* ou *GFI1B* ou des mutations dans *SMARCA4*. Le groupe 4 serait surtout décrit par une activation et une forte expression de *PRMD6* et des mutations *KDM6A*, *ZMYM3* ou *KMT2C*⁷⁹.

La caractérisation de ces deux derniers groupes reste sujette à débat mais l'utilisation de puces à méthylation semble plus robuste que les techniques utilisées jusqu'à ce jour pour la classification des médulloblastomes. Ces méthodes sont de plus en plus utilisées pour orienter le diagnostic⁸⁰⁻⁸².

1.1.2.2 Rhabdomyosarcome

Le rhabdomyosarcome (RMS) est le sarcome des tissus mous d'origine mésenchymateuse le plus fréquent chez les enfants⁸³. Cette tumeur peut être très agressive et nécessite des traitements lourds⁸⁴. Les rhabdomyosarcomes sont décrits par deux types :

- les rhabdomyosarcomes alvéolaires (aRMS) sont constitués de petites cellules rondes et denses rappelant des alvéoles de poumon. Le type alvéolaire touche plutôt des adolescents et est très agressif. Ces tumeurs sont définies majoritairement par des fusions entre les gènes *PAX3* ou *PAX7* et *FOXO1*^{85,86}. Des autres fusions avec le gène *PAX* et des amplifications du gène *MYCN* et *CDK4* ont été observées mais aucune mutation récurrente n'a été identifiée dans ce type de RMS⁸⁷.

- Les rhabdomyosarcomes embryonnaires (eRMS) sont constitués de cellules ressemblant à des squelettes immatures de myoblastes. Le type embryonnaire touche de jeunes enfants avec un pronostic plus favorable (en dehors de forme métastatique)⁸⁸. Le profil génétique de ces tumeurs est plus hétérogène avec des pertes d'hétérozygotie au niveau du 11p15.5, des amplifications du gène *MDM2* et des mutations dans les gènes *TP53*, *NF1*, *PIK3CA*, *CTNNB1*, *DICER1*, *FGFR4* ou dans les voies RAS⁸⁹⁻⁹⁵. La plupart

de ces mutations semblent exclusives entre elles et ne vont toucher chacune que 5% des cas de eRMS⁸⁷.

1.1.2.3 Tumeur d'Ewing

Les tumeurs d'Ewing sont les secondes tumeurs osseuses après l'ostéosarcome chez les enfants et adolescents et ces tumeurs peuvent également envahir les tissus mous⁹⁶. Ce sont des cancers très agressifs avec un fort taux de guérison pour les formes localisées (70 à 80% de guérison) et très faible pour les formes métastatiques (30% de guérison). L'origine de ces tumeurs reste une énigme, celles-ci ayant des profils de méthylation très atypiques et hétérogènes mais elles pourraient provenir de cellule mésenchymateuse dérivée de la moelle osseuse^{97,98}.

Le sarcome d'Ewing est caractérisé par une translocation chromosomique récurrente le plus souvent entre les gènes de la famille *FET* tels que *EWSR1* et ceux de *ETS* tels que *FLI1*^{99,100}. Dans 85% des cas, la fusion se fait entre *EWSR1-FLI1* mais une fusion peut aussi être observée entre *EWSR1-ERG* (10%) et dans de rares cas entre d'autres gènes de la famille *ETS* voire *FET* (<1%)¹⁰¹⁻¹⁰⁴.

En dehors de la fusion et d'altérations en nombre de copie telles que les gains du 8, 1q, 12 et des pertes du 16q et 19q, très peu d'aberrations notamment de mutations sont observées^{105,106}.

Des mutations de *STAG2*, *TP53* et des délétions de *CDKN2A* (semblant être exclusives à ces deux mutations) ont été observés entre 17 à 7% des cas^{107,108}.

1.1.2.4 Tumeur de Wilms

Les tumeurs de Wilms sont les tumeurs rénales les plus fréquentes chez les enfants (90%). Ces tumeurs proviendraient d'un développement rénal aberrant du notamment à des mutations dans des gènes impliqués dans la néphrogénèse fœtale¹⁰⁹. Des cellules souches embryonnaires avec une hyperméthylation de H19 peuvent éventuellement être responsables de l'expansion de la tumeur¹¹⁰.

Environ une quarantaine de gènes mutés seraient impliqués dans la génétique de tumeurs de Wilms tels que *TP53*, *WT1*, *CTNNB1*, *MYCN* ou *DROSHA*¹¹¹⁻¹¹³. Des altérations en nombre de copie sont également observées telles que le gain du gène *MYCN*, des pertes du 1p, 4q, 11q, 14q et 16q¹¹⁴⁻¹¹⁶. De nombreuses altérations constitutionnelles sont aussi responsables du développement de ces tumeurs telles que les délétions ou les mutations dans le gène *WT1* ou *DICER1*¹¹⁷⁻¹¹⁹.

1.2 Rôle de la génétique dans l'échappement aux traitements

Dans la plupart des cancers pédiatriques, les traitements de première ligne permettent d'obtenir des réponses aux traitements, au niveau de la tumeur primaire ainsi qu'au niveau des métastases, le cas échéant. Mais des progressions ou des rechutes peuvent survenir notamment dans les formes de haut risque. Dans ces situations, des techniques de séquençage haut débit ont pour objectif une caractérisation moléculaire complète pour établir un profil moléculaire à la rechute. Ce type d'analyse permet de proposer aux patients de plus en plus des traitements ciblés dans le cadre d'essai ou de phases précoces¹²⁰⁻¹²³.

La comparaison des profils moléculaires entre le diagnostic et la rechute souligne l'importance de l'évolution clonale dans la résistance et l'échappement aux traitements mais également la plasticité cellulaire avec des modifications épigénétiques de cellules.

1.2.1 Hétérogénéité et évolution clonale

Le cancer a pour origine une cellule somatique qui a accumulé des altérations génétiques (mutations, modifications épigénétiques, altérations et remaniements chromosomiques). Ces mutations modifient le phénotype des cellules leur permettant d'échapper à la régulation homéostatique et de maintenir un nombre de cellules défini.

Au point de vue biologique, cette transformation d'une cellule normale en une cellule maligne est très proche du concept d'évolution par Darwin¹²⁴⁻¹²⁶. L'évolution se poursuit au long de la croissance, de la progression, de la résistance au traitement et de la rechute du cancer entraînée par l'adaptation des cellules cancéreuses. L'hétérogénéité tumorale qu'elle soit spatiale ou temporelle est une conséquence inévitable du processus évolutif.

Grâce aux avancées de la génomique, il est possible de mieux caractériser les tumeurs et de permettre des mesures quantitatives de l'évolution des sous-populations cellulaires, ou clones, dans le temps et l'espace¹²⁷⁻¹²⁹.

1.2.1.1 Mise en évidence de l'évolution clonale

En 1953, de précédentes études estimaient que le cancer était dû à une seule mutation, cette mutation serait dû en majeure partie à l'âge et donc au nombre de divisions qu'ont subies les cellules ou une exposition prolongée à un composé mutagène. Cette hypothèse signifierait que les individus seraient "égaux" face au cancer et, de ce fait, une augmentation linéaire nette du nombre de cas aurait dû être observée selon l'âge des patients. Cette hypothèse ne permettait pas d'expliquer les cancers pédiatriques, notamment par leurs âges et représentés par des tumeurs très agressives au développement rapide. Ce développement rapide étant dû à une division dérégulée des cellules, de nombreuses mutations spontanées devaient donc être possible suite au cycle cellulaire accéléré. Il fut donc proposé que le cancer était dû à plusieurs mutations pouvant s'acquérir dans le temps¹³⁰.

Ainsi, une étude du rétinoblastome suggéra que les types bilatéraux (touchant les deux yeux) non héréditaires étaient dus à différentes mutations par tumeurs (certains cas ayant plus de 2 tumeurs). Un cas de régression spontanée fut également observé sans signe clinique distinctif. Les mutations présentes dans ce cancer seraient donc différentes d'un individu à un autre mais également au sein d'un individu atteint de plusieurs tumeurs¹³¹. Les avancées de la biologie moléculaire et des analyses généalogiques ont par la suite démontré la prédisposition pour ce type de tumeur mais cette étude a mis en mal l'homogénéité du cancer¹³².

Mais ce n'est qu'en 1976 grâce une étude sur les leucémies que fut proposé le terme "évolution clonale" (Figure 1.7). Une cellule deviendrait néoplasique suite à une exposition

d'un carcinogène (radiation, produit chimique, virus) ou une altération génétique spontanée. Ces cellules pourraient devenir plus instables suite à une nouvelle exposition ou à l'apparition de nouvelle mutation mais également en fonction de la nutrition des cellules ou de la thérapie. Sous l'influence de ces facteurs ainsi que du statut immunitaire d'un patient ou de la présence d'une infection, des sous-clones pourraient ainsi émerger.

Le modèle d'évolution clonale propose qu'une cellule subisse un changement génétique induit ou spontané subissant ainsi une "prolifération néoplasique". Par la suite, des altérations génétiques aléatoires au sein de ces cellules néoplasiques créent des cellules avec de nouvelles mutations ainsi que des phénotypes variables et ces cellules vont subir une pression de sélection. La plupart de ces altérations génétiques étant néfastes, une partie de ces cellules seront éliminées soit par compétition pour les ressources, soit détruites par le système immunitaire de l'hôte agissant comme des prédateurs décrits dans la sélection naturelle des espèces¹²⁴. Parfois, ces changements génétiques offrent un avantage sélectif à la cellule tumorale conduisant à l'émergence d'une sous-population dominante.

Ces cycles de diversification et sélection entraînent la progression tumorale vers des niveaux de malignité plus élevée¹³³.

Fig. 1. Model of clonal evolution in neoplasia. Carcinogen-induced change in progenitor normal cell (N) produces a diploid tumor cell (T_1 , 46 chromosomes) with growth advantage permitting clonal expansion to begin. Genetic instability of T_1 cells leads to production of variants (illustrated by changes in chromosome number, T_2 to T_6). Most variants die, due to metabolic or immunologic disadvantage (hatched circles); occasionally one has an additional selective advantage (for example, T_2 , 47 chromosomes), and its progeny become the predominant subpopulation until an even more favorable variant appears (for example, T_4). The stepwise sequence in each tumor differs (being partially determined by environmental pressures on selection), and results in a different, aneuploid karyotype in each fully developed malignancy (T_6). Biological characteristics of tumor progression (for example, morphological and metabolic loss of differentiation, invasion and metastasis, resistance to therapy) parallel the stages of genetic evolution. Human tumors with minimal chromosome change (diploid acute leukemia, chronic granulocytic leukemia) are considered to be early in clonal evolution; human solid cancers, typically highly aneuploid, are viewed as late in the developmental process.

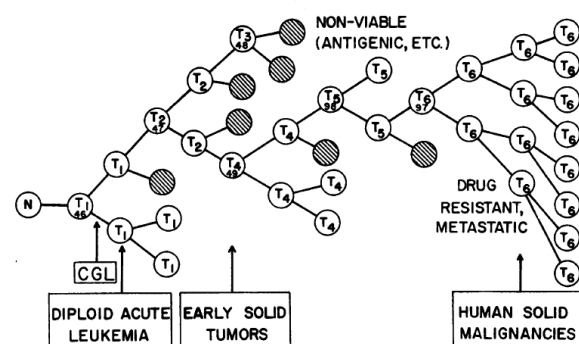


Figure 1.7 : Modèle d'évolution clonal dans la néoplasie (Figure en accès libre)¹³³.

Observée initialement dans les leucémies, l'évolution clonale a été ensuite décrite dans de nombreux autres types de tumeurs et confirmée grâce aux progrès de la génétique et de la biologie moléculaire tels que les puces à ADN ou le séquençage haut débit.

1.2.1.2 Évolution clonale dans les cancers

Le principe d'évolution clonale repose sur l'hétérogénéité clonale d'une tumeur au niveau génétique.

Des premières analyses en FISH (Hybridation In Situ en Fluorescence) ont permis d'observer cette hétérogénéité dans le glioblastome. Ces tumeurs sont connues pour avoir des amplifications des gènes *EGFR*, *PDGFRA* et *MET* mais également des délétions des gènes *CDKN2A*. L'analyse en FISH révéla que la plupart des cellules possédaient plusieurs de ces anomalies et que des sous-populations avec des altérations exclusives étaient observées et, pour certaines tumeurs, les trois amplifications étaient totalement exclusives. L'analyse de trois de ces tumeurs par aCGH permit d'observer l'inter-hétérogénéité entre tumeurs mais la faible de cette technique sensibilité ne permit pas de détecter les populations mineurs¹³⁴.

Dans le cancer du sein, des tumeurs ont été découpés en parties égales¹³⁵. Les cellules en fonction de leur ploïdie ont été triées par FACS puis leurs ADN hybridés sur aCGH. Certaines parties des tumeurs ont été considérées comme monogéniques avec une partie aneuploïde et une partie diploïde certainement due au stroma ou aux cellules immunitaires présentes et d'autres parties de tumeurs semblent être polygénomiques avec des populations ayant des aneuploïdies différentes. En fonction des gains ou des pertes, des arbres phylogénétiques de ces tumeurs ont pu être proposés et pour deux tumeurs, une topographie de l'évolution clonale en partant d'un bout de la tumeur jusqu'à l'autre fut observée avec une augmentation des altérations en nombre de copie. L'observation de niveau différent de gain de certains gènes tels que *KRAS* pu être corrélée par FISH avec la présence de plus en plus de

cellules présentant une amplification forte du gène selon la position du fragment dans la tumeur.

Ce même schéma fut reproduit mais cette fois-ci en couplant FACS et séquençage du génome entier (Whole Genome Sequencing, WGS). Le WGS réalisé étant à faible couverture, seules les altérations en nombre de copie ont pu être analysées. En plus de la tumeur, une métastase a pu être analysée en WGS. Cependant, aucune nouvelle phylogénie n'a été observée par rapport à l'étude précédente. Les cassures chromosomiques dans le génome semblent donc apparaître très tôt dans l'oncogenèse de la tumeur du sein. Comme aucune altération en nombre de copie supplémentaire n'a été observé dans la métastase, d'autres facteurs tels que des mutations pourraient être à l'origine de développement¹³⁶.

Dans le cancer du pancréas, l'analyse de la découpe d'une tumeur et des métastases par séquençages de certains gènes a permis de modéliser plusieurs phases de métastases avec l'apparition de premiers sous-clones envahissant le péritoine puis beaucoup plus tard dans la vie de la tumeur, de nouveaux sous-clones, certainement plus agressifs capables d'envahir le foie puis les poumons¹³⁷.

En conséquence, l'analyse en WGS à plus forte couverture pour détecter des altérations en nombre de copie, des mutations ou des translocations couplée à des modèles bioinformatiques plus puissants ont permis de proposer une "histoire" de la tumeur avec la caractérisation des altérations et leurs apparitions dans le temps^{138,139}. Les altérations classées en "driver " (ayant un impact sur l'oncogenèse) ou " passenger " doivent ainsi être clustérisées puis inférer à un clone et en fonction du nombre d'altérations et de leurs fréquences, une phylogénie peut être proposée.

De surcroît, la modélisation est encore plus poussée dans le cadre du séquençage de tumeur primaire et de leur rechute. Un clone ancestral peut être ainsi modélisé qui donnera un

clone primaire et des sous-clones. Suite au traitement, des clones seront éliminés et des sous-clones pourront émerger ainsi que de nouveaux clones^{140,141}. Pour chaque étude, des altérations communes sont observées entre la tumeur primaire et sa rechute. Si bien que le clone ancestral, portant des altérations partagées par tous les clones, ne semble jamais être éliminé dans le cadre d'une rechute pouvant expliquer la résistance aux traitements.

Dans le carcinome à cellule claire du rein, le consortium TRACERx a ainsi proposé des analyses très poussées couplant séquençage d'exomes (WES), puces SNP et puces d'expression sur des tumeurs découpées en plusieurs parties mais également sur des métastases¹⁴². Alors qu'une mutation dans le gène *VHL* était présente dans le clone ancestral pour quelques patients, la majorité des mutations avec un effet délétère apparaissait plus tard dans l'histoire de la maladie. Des mutations de *SETD2* et de *KDM5C* étaient détectées mais différentes selon la partie de la tumeur ou de la métastase où elles ont été observées. Les clones tumoraux et métastatiques étaient très vite indépendants et semblaient évoluer vers le même type de phylogénie avec des mutations drivers communes.

Ensuite, l'étude a été élargi son analyse pour inclure 100 patients pour lesquels 1206 parties de tumeurs (médiane de 7, allant de 3 à 75 parties par patient) ont été séquencées majoritairement par WES mais également par WGS et panel dédié au cancer du rein¹⁴³⁻¹⁴⁵. Parmi ces patients, 46 avaient des métastases qui ont pu être aussi séquencées de la même façon que la tumeur. Cette expérience à plus grande échelle a confirmé l'étude précédente à savoir des mutations dans le gène *VHL* apparaissant très tôt et mais également la possibilité de phylogénie commune avec des altérations différentes mais dans des gènes drivers communs. En fonction des mutations présentes dans les sous-clones métastatiques, la prolifération serait plus ou moins rapide touchant des organes différents. Ainsi des clones avec une évolution rapide contenant des altérations du gène *BAP1* sont retrouvés dans le poumon alors que des

sous-clones avec une prolifération atténuée avec des altérations *PI3K* sont plutôt présents dans les os.

Les sous-clones métastatiques d'évolutions rapides sont très remaniés au niveau génétique et peu hétérogènes mais leurs disséminations fulgurantes et leurs fixations à de multiples sites conduisent à une résistance aux traitements tandis que les sous-clones avec une évolution plus lente et avec peu de site de fixation possèdent un pronostic plus favorable. Ainsi pour ces tumeurs, l'apparition de mutation sous-clonale de certains gènes drivers, plus ou moins tôt, dans la maladie serait responsable de sa gravité.

Dès lors, le consortium a proposé une étude sur 100 patients atteints de cancer du poumon non à petites cellules dans laquelle l'exome de 327 parties de tumeur a été séquencé (de 2 à 8 parties de tumeur séquencées par patients)¹⁴⁶. Les mutations drivers telles que *TP53* ou *BRAF* étaient acquises très tôt et certainement par le clone ancestral. De nombreuses mutations sous-clonales de gènes tels que *NF1*, *PICK3CA* ou d'autres gènes impliqués dans le remodelage de la chromatine ou la réparation de l'ADN ont été aussi détectés dans la majorité des tumeurs. La ploïdie de certains clones présentant des altérations en nombre de copie telles que des amplifications de *CDK4* ou *FOXA1* semblait jouer aussi un rôle dans l'agressivité de la tumeur.

Ainsi plus la tumeur aura des altérations sous-clonales avec des mutations et des amplifications de gènes "driver", plus elle sera à même de s'adapter au traitement et d'avoir un impact sur la survie du patient.

La présence d'un sous-clone résistant à un traitement n'est pas forcément responsable d'une rechute. Dans une étude sur le cancer du sein, deux clones ont été identifiés, un clone basal muté dans le gène *HRAS* certainement responsable de l'expansion de la tumeur et un

clone luminale avec une expression du gène *WNT*. Le gène *HRAS* pouvant être ciblé, une chimiothérapie aurait pu détruire la tumeur. Des souris transgéniques avec une mutation dans le gène *HRAS* et une expression activable du gène *WNT* ont montré que la tumeur devenait résistante quand le gène *WNT* était exprimé suggérant non pas une compétition entre les clones mais une collaboration¹⁴⁷.

Une autre expérience, utilisant des xenogreffes de lignées cellulaires issues de cancer du sein, a mis en évidence des sous-clones contenant des gènes drivers responsables de l'oncogenèse mais avec un développement de la tumeur lent. Cependant la vitesse de croissance de développement s'accélérait avec l'apport d'Interleukine11. Un autre sous-clone fut ainsi mis en évidence répondant favorablement à l'IL11. Sa prolifération devenait beaucoup plus importante en présence des autres sous-clones¹⁴⁸. Une collaboration entre différentes sous populations cellulaires tumorales, ou sous-clones, est donc nécessaire pour une expansion rapide de la tumeur.

1.2.1.3 Hétérogénéité et évolution clonale dans les tumeurs pédiatriques

Dans les cancers solides pédiatriques, une hétérogénéité et une évolution clonale ont pu être documentées dans plusieurs pathologies

Dans le neuroblastome, l'amplification du gène *MYCN* observée par FISH, responsable du développement du neuroblastome, peut être plus ou moins forte d'une cellule à une autre¹⁴⁹. Ces observations suggèrent une hétérogénéité de l'amplification intra-tumorale de *MYCN* mais l'origine de ces différents taux d'amplification reste une énigme.

Lors de la rechute chez des patients atteints de neuroblastome, les mutations dans le gène *ALK* peuvent souvent émerger. Par des séquençages à très grande profondeur ces

mutations peuvent être détectées dans la tumeur au diagnostic à un niveau sous-clonale parfois très faible (autour de 0,5%)^{36,51}.

Ainsi, une étude sur deux lignées cellulaires issues d'un même patient atteint de neuroblastome dont une lignée établie à partir de la tumeur primaire et une autre partie sur la métastase a permis d'observer des résultats surprenant sur la dynamique des mutations d'*ALK* au cours des passages de lignées¹⁵⁰. Initialement, la plupart des cellules issues de la tumeur primaire arborent deux mutations dans le gène *ALK*, F1174 et L1196. La fréquence des mutations augmente tout d'abord en même temps. Puis la mutation L1196 devient plus importante en termes de fraction allélique mutée pour ensuite être complètement effacée par la F1174 avec des cellules très prolifératives. La mutation L1196 est seulement détectée dans la lignée issue de la tumeur primaire et pas dans celle issue des métastases possédant seulement la F1174. Cette expérience laisse penser que les clones L1196 et F1174 ont besoin de collaborer pour "construire" la tumeur puis une compétition a lieu entre les clones. Les clones F1174 dominants et proliférant plus vite pourraient métastaser. Ces résultats ont également un impact sur l'utilisation thérapeutique de médicaments ciblant *ALK* (par exemple le crizotinib) avec une sensibilité différentielle selon les clones et conduisant donc à un effet partiel au niveau de la tumeur¹⁵¹.

Chez des patients atteints de sarcome d'Ewing, l'expression du gène *STAG2*, observé par marquage immunohistochimique, était réprimée lors de la rechute comparée au diagnostic. Le séquençage en WGS a ainsi permis de détecter une mutation de *STAG2* entraînant une perte de fonction. Cette altération passait ainsi d'un niveau sous-clonal à un niveau clonal entre les tumeurs au diagnostic et les tumeurs à la rechute¹⁰⁷.

Dans le médulloblastome, la classification moléculaire ne semble pas changer en cas de rechute¹⁵². Cependant, le WGS sur des paires de tumeurs primaires et de tumeurs à la rechute a permis d'observer la disparition de clones avec une altération dans *PTCH1* et l'émergence de clone avec perte de *CDKN2A*. Le groupe de tumeur ne change pas mais un des ponts principaux de la classification en SHH est rompu¹⁵³.

Chez patients atteints de tumeurs de Wilms unilatérales, les tumeurs peuvent présenter des parties différentes visibles au scanner et avec des nodules répondant différemment à la chimiothérapie. Après exérèse de la tumeur, l'analyse des profils génomiques par puces Illumina CytoSNP a pu montrer des altérations supplémentaires dans certains nodules. Ainsi, un arbre phylogénétique fut proposé et représentant l'évolution de certaines tumeurs. De plus, un nodule résistant aux chimiothérapies présentait des altérations supplémentaires plus agressives telles que le gain du gène *MYCN*¹⁵⁴.

Dans le glioblastome de l'enfant, une sous-population de cellule de type embryonnaire avec un cycle cellulaire lent serait responsable d'une rechute très agressive¹⁵⁵. Une coopération clonale fut aussi mise en évidence pour ce type de tumeurs pédiatriques. Des analyses en WES ont été réalisées pour caractériser les altérations génétiques de la tumeur mais également les populations clonales. Les cellules ont été triées en FACS puis identifiées par TargetSequencing pour isoler les clones. En mettant en culture deux clones différents ensemble, la vitesse de croissance n'était pas impactée mais l'invasion ou la migration dans un matrigel était plus important que lorsque les clones étaient cultivés individuellement suggérant la coopération entre clones dans l'expansion de la tumeur¹⁵⁶.

L'analyse par puces SNP et par FISH de tumeurs pédiatriques (neuroblastome, néphroblastome, tumeur rhabdoïde et hépatoblastome) découpées en 2 à 4 parties a montré des altérations en nombre de copie différentes et des altérations supplémentaires dans les métastases et les tumeurs à la rechute par rapport celles présentes dans la tumeur primaire. De plus, des lignées issues de néphroblastome et de neuroblastome ont été utilisées comme xenogreffes sur des souris. Après sacrifices des souris, l'analyse des tumeurs a révélé des altérations supplémentaires à celles détectées dans les lignées. Certaines altérations étaient également exclusives entre xenogreffes issues d'une même lignée¹⁵⁷. Une hétérogénéité spatiale mais également temporelle des tumeurs pédiatriques a ainsi été mise en évidence.

L'étude a ensuite été étendue à 250 fragments provenant de 54 tumeurs pédiatriques (neuroblastomes, tumeurs de Wilms et rhabdomyosarcomes) avec une analyse par puce SNP pour la totalité des tumeurs et en WES pour 54 régions de 19 tumeurs. Quatre types d'évolution clonales ont été décrits :

- la coexistence de 2 clones majoritaires, le plus souvent, qui évoluent, collaborent et restent stable,
- la variation sous clonale avec l'apparition de nombreuses altérations mais non délétères,
- le "balai" où un clone devient dominant voire unique,
- l'explosion où un gène driver va subir une altération dérégulant la stabilité du clone et conduisant à une explosion d'altération.

Le dernier type d'évolution semblait plus fréquent dans les tumeurs de Wilms avec des pronostics très sombres pour les patients atteints de celles-ci¹⁵⁸.

L'hétérogénéité clonale et son évolution permettent d'expliquer en grande partie l'échappement aux traitements. De plus en plus de types de cancers semblent contenir des

altérations dans des voies clefs leur permettant de s'adapter mais ces résultats ne peuvent être obtenus que sous certaines conditions : analyser sur des larges cohortes nécessitant des moyens de séquençages importants, l'accès à l'intégralité de la tumeur primaire et si possible aux métastases. Ce type d'étude est donc plus difficile à réaliser pour les cancers pédiatriques du fait de leur rareté (100 cas de neuroblastome par an en France contre 60 000 cas de cancer du sein ou 45 000 cas de cancer du poumon) et de la difficulté voire l'impossibilité d'accéder à la totalité de la tumeur en pédiatrie.

1.2.2 Changement d'identité cellulaire

En plus de l'évolution clonale basée sur des événements génétiques, le changement d'identité cellulaire ou la plasticité cellulaire pourrait également contribuer à l'échappement thérapeutique¹⁵⁹. La résistance d'une tumeur ne serait pas due à un arsenal de cellules avec différents types d'altérations génétiques mais à une reprogrammation de ces cellules en fonction d'une exposition ou à son microenvironnement jusqu'à un stade indifférencié proche de l'embryonnaire¹⁶⁰⁻¹⁶³. Effectivement suite à l'établissement de greffes de tumeur de patients atteints de cancer du sein sur des souris immunodéprimées, seules des sous-populations de cellules indifférenciées permettaient la repousse de la tumeur. Une fois développées, les tumeurs étaient composées de cellules tumorales mais également de cellules indifférenciées¹⁶⁴.

La théorie de plasticité cellulaire s'opposait tout d'abord à l'évolution clonale et était plutôt basée sur des non-compréhensions cliniques. Finalement les deux théories sont maintenant complémentaires pour expliquer l'échappement tumoral¹⁵⁹.

Cette plasticité de la cellule est surtout connue pour répondre à un stress tel qu'une inflammation, une blessure voire un stress oncogénique. Une altération dans le contrôle de cette plasticité peut ainsi conduire à une transformation maligne. De nombreux cancers sont due à des mutations drivers dans des gènes impliqués dans le développement ou dans la maintenance de la chromatine¹⁶⁵.

Dans le glioblastome, des tumeurs porteuses de mutations dans le gène *TP53* deviennent résistant à la radiothérapie grâce à un changement d'identité vers des cellules de type mésenchymateux¹⁶⁶. La perte de *p53* dans les cellules du foie de souris inductibles nestin conduit au développement des cellules progénitrices malignes puis selon les mutations dues à l'instabilité de la cellule à des carcinomes hépatocellulaires ou des cholangiocarcinomes¹⁶⁷. La réexpression de *p53* empêche les cellules progénitrices de proliférer. Le gène *p53* a donc un rôle majeur dans la restriction de la plasticité cellulaire.

Les cellules squameuses de l'estomac suite à des stress dus à *Helicobacter pylori* mais aussi sous l'influence du tabac ou de l'alcool semblent passer par un état embryonnaire non déterminé puis en cellule souche de l'intestin. Ces dernières seraient responsables d'adénocarcinome gastrique¹⁶⁸.

La perte du gène *STK11* transforme des cellules d'adénocarcinome du poumon muté *KRAS* en cellule de carcinome cellulaire squameux dans des modèles de souris. Ces cellules ont également une surexpression de la méthyltransférase H3K27 et des remodelages importants de la chromatine suggérant que le complexe Polycomb est touché¹⁶⁹. Mais aucune interaction n'a été observée entre ce complexe et le gène *STK11*.

1.2.2.1 Type de plasticité

Trois types de plasticité, ou métaplasia, ont été décrites :

- Transition épithéliale-endothéliale (ou mimétisme vasculogénique) : la tumeur va former des canaux micro-vasculaires. Ces derniers peuvent aller jusqu'à communiquer avec les vaisseaux sanguins. Cette transition s'observe surtout dans les tumeurs très agressives et métastiques¹⁷⁰.
- Transition épithéliale-hématopoïétique : les cellules présentes dans les artères peuvent donner de multiples précurseurs de cellules hématopoïétiques. Cette transition pourrait jouer un rôle dans les leucémies mais n'a jamais été décrite¹⁷¹.
- Transition épithélio-mésenchymateuse (EMT) : cette transition va permettre aux cellules de se libérer de leur jonction. Cette étape peut avoir lieu de nombreuses fois dans la vie d'une cellule notamment lors du développement de la crête neurale. De nombreux cancers sont associés à ce type de transition et ce processus peut être à l'origine de leur dissémination.

Effectivement il a été observé dans le cancer du sein, des lignées de cellules inductibles EMT deviennent mésenchymateuses en augmentant de l'expression des transporteurs d'ATP en présence de drogues. Les cellules deviennent alors plus mobiles, plus envahissantes et plus résistantes¹⁷².

Dans le cancer du pancréas, l'analyse de croisement de souris altérées ou non au niveau des cellules cancéreuses type endothéliales et dans la voie de transition (EMT) ne montrent aucune invasion ou grossissement de la tumeur quand les cellules sont EMT+. Mais ces cellules sont résistantes aux drogues anti-prolifération¹⁷³.

La transition EMT pourrait être majoritaire dans l'échappement aux traitements. Cette voie permet à des tumeurs de devenir résistante à de nombreux types de chimiothérapie mais aussi à la radiothérapie et à l'immunothérapie. La forme des cellules mésenchymateuses leur permet également d'éviter les interactions avec des cellules immunitaires¹⁷⁴. Cette forme permet à la tumeur d'éviter une "agression" mais pas de proliférer. Une fois passée, la tumeur peut reprendre une forme épithéliale plus propice à son développement.

1.2.2.2 Plasticité cellulaire dans les tumeurs pédiatriques

L'observation de la plasticité des tumeurs reste compliquée nécessitant des études sur des lignées ou souris inductibles dans la voie de métaplasia. Cependant des études récentes sur certaines tumeurs pédiatriques ont mis en évidence un rôle important de la plasticité cellulaire.

Dans le sarcome d'Ewing, la fusion *EWSR1-FLI1* peut être exprimée de manière hétérogène au sein d'une même lignée cellulaire. Les cellules à forte expression *EWSR1-FLI1* profilèrent très vite tandis que celles à faible expression migrent et envahissent en produisant des sphéroïdes. En inversant l'expression de la fusion, les phénotypes sont également inversés. Ainsi l'expression d'*EWSR1-FLI1* diminuant dans le temps, une tumeur une fois établie peut alors métastaser. Ces métastases vont ensuite proliférer suivant une augmentation de l'expression du gène chimère¹⁷⁵.

Concernant le neuroblastome, l'analyse de lignées cellulaires par RNA-Seq et Chip-Seq a mis en évidence que co-existaient deux populations de cellules neuroblastiques dérivées d'une tumeur d'un même patient : une population adrénérergique formant des petites sphères semi-attachées et immobiles et une population mésenchymateuse avec une large membrane et

très mobile. La première population de type adrénergique exprime des gènes tels que *PHOX2A*, *PHOX2B* et *DBH* et la seconde des marqueurs mésenchymateux tels que *SNAI2*, *VIM* et *FN1*^{176,177}.

L'expression de la forme mésenchymateuse semble très proche des cellules de la crête neurale et possède donc une forme indifférenciée.

Deux super-enhancers spécifiques semblent régir ces deux populations avec une activation différentielle de gènes pour chaque population tels que *PRRX1* pour le type mésenchymateux. Ainsi, l'introduction du facteur de transcription *PRRX1* dans des cellules de type adrénergique permet d'observer une transition à un état mésenchymateux. Toutefois, son inactivation dans des cellules de type mésenchymateux n'a pas permis un retour à l'état adrénergique

L'analyse sur des tumeurs primaires et des rechutes n'ont pas permis de mettre en évidence le type mésenchymateux dans celles-ci mais des analyses par immunohistochimie ont permis d'observer des cellules *PRRX1* positive. Ces résultats laissent penser que ce type de populations est présente à un niveau sous-clonal.

D'autre part, les cellules mésenchymateuses sont plus résistantes aux chimiothérapies standards contre le neuroblastome telles que la cisplatine. De plus, certaines lignées adrénergiques traitées à la cisplatine prennent une signature mésenchymateuse et résistent par la suite au traitement. Selon ces observations, les neuroblastomes pourraient ainsi échapper aux traitements en remodelant leurs cellules vers l'identité cellulaire de type mésenchymateuse, plus résistant aux traitements.

Le microenvironnement semble également avoir un rôle important et notamment les cellules de Schwann. Le peu de présence de cellules de Schwann autour d'un neuroblastome est souvent synonyme de mauvais pronostic. L'analyse par RNA Seq de cellules cultivées

dans différentes conditions de milieux a permis de mettre en évidence l'expression d'une protéine de sécrétion peu décrite, EGFL8. L'utilisation de cette "hormone" sur des lignées cellulaires de neuroblastome peu différenciées et très prolifératives a permis de les transformer en cellules neurales différenciées et peu prolifératives. Cette molécule pourrait ainsi être utilisé pour "jouer" sur l'identité cellulaire du neuroblastome et le rendre moins agressif¹⁷⁸.

L'évolution clonale et la plasticité ont d'abord été opposées dans le mécanisme d'échappement aux traitements, tout comme la sélection naturelle et la dérive génétique, mais comme ces derniers, ces mécanismes semblent être complémentaires et faire partie de la vie d'une tumeur¹⁷⁹.

Grâce à l'incroyable évolution de la biologie moléculaire et l'apport du séquençage haut-débit, la complexité de la tumeur du point de vue son hétérogénéité et de sa plasticité a pu être mieux comprise. Mais ces études sont souvent lourdes que ce soit dans leurs mises en place (création de cohorte) mais également en coût de séquençage. Certaines de ces analyses n'ont été observées qu'in situ voire au mieux in vivo dans des modèles murins. Les études les plus complètes nécessitant d'avoir une vue globale de la maladie n'ont été possibles qu'après la mort d'un patient. Les technologies de Single Cell notamment grâce à des systèmes de microfluidiques de plus en plus avancées permettent de mieux cartographier la tumeur. Cependant malgré la baisse des couts de séquençages, ces analyses restent très couteuses demandant du matériel frais et les protocoles fonctionnels se limitent pour le moment l'analyse de l'ARN ou de la chromatine¹⁸⁰.

Afin d'avoir des informations plus précises et en temps réel sur le développement de la maladie et l'apparition ou la disparition d'altérations au sein de la tumeur, l'utilisation des

biopsies liquides pour les analyses et notamment de l'ADN circulant présent dans les fluides corporels semble plus adaptée.

1.3 Les biopsies liquides

Les biopsies liquides, connus de longues dates, sont devenus des outils révolutionnaires qui permettent l'accès à des informations biologiques ou pathologiques.

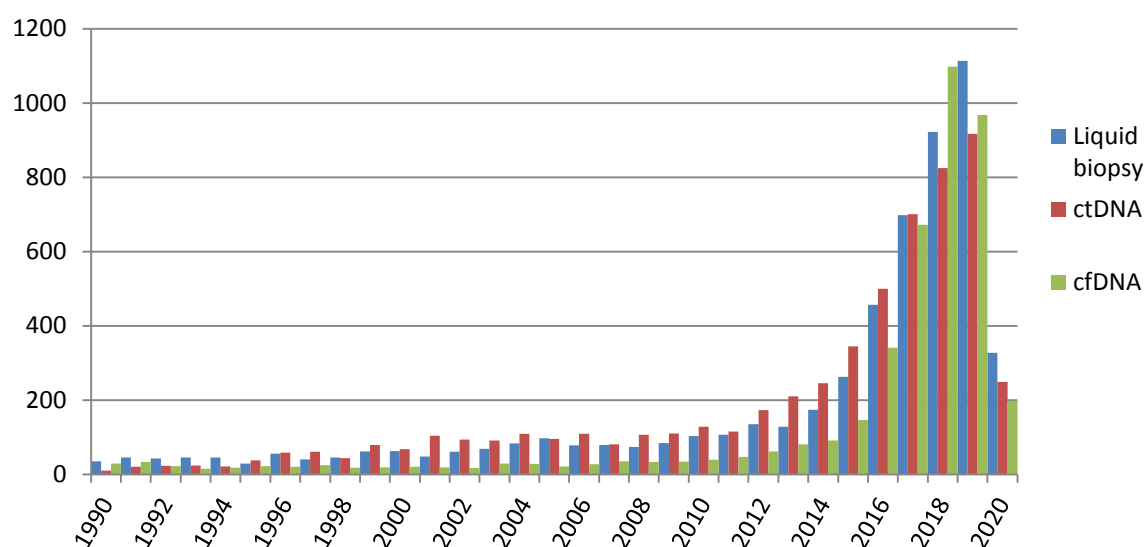


Figure 1.8 Nombre de citation des termes "Liquid Biopsy", "ctDNA" et "cfDNA" dans les résumés Pubmed par an.

Effectivement des fluides du corps tels que le sang, le liquide céphalorachidien, la salive, l'urine, les selles, le liquide pleural ou le liquide amniotique permettent d'isoler différents composants, cellulaires ou moléculaires. Certaines cellules peuvent libérer leurs ADN nommé ADN circulant (cfDNA) dans ces fluides suite à leurs apoptoses, nécroses ou par sécrétions. Si ces cellules ont une origine tumorale, le cfDNA sera nommé ADN tumoral circulant (ctDNA).

La promesse d'avoir accès à la tumeur avec une méthode peu invasive et de façon successive a ainsi fait exploser le nombre de recherches ou d'études cliniques et ce d'autant plus que les techniques de biologie moléculaire telles que le NGS se sont démocratisées grâce à une réduction des coûts des protocoles et des instruments ces dernières années (Figure 1.8).

1.3.1 Les biopsies liquides dans le cancer

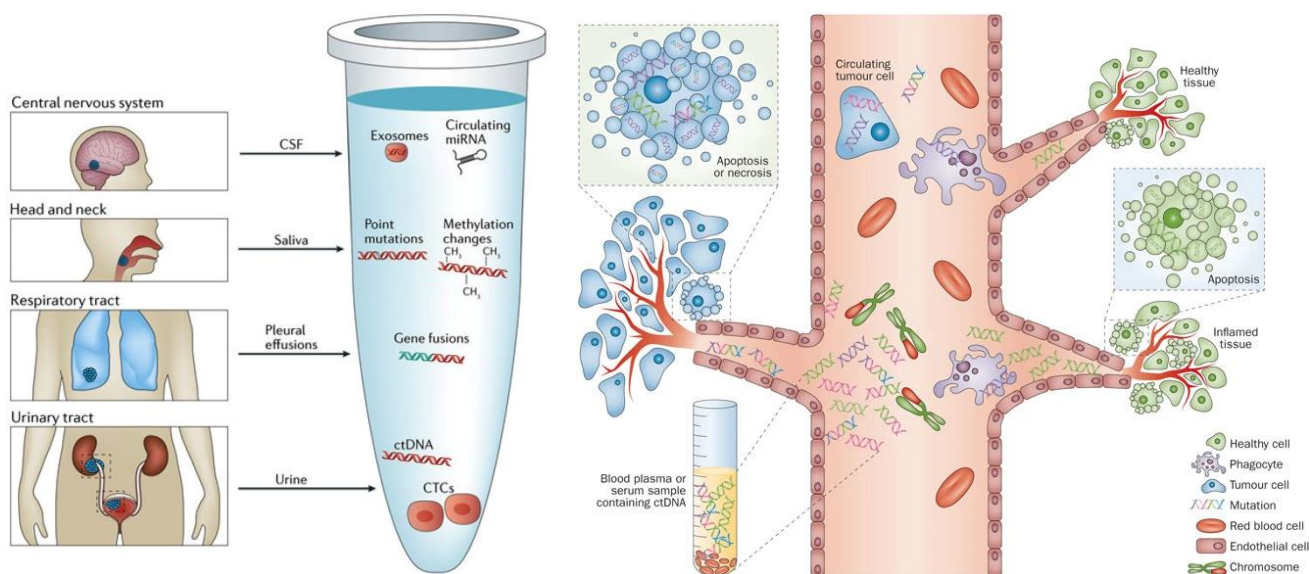


Figure 1.10 : Type de biopsie liquide et contenu (Figures en accès libre)^{181,182}.

L'ADN circulant n'est qu'une fraction du contenu d'une biopsie liquide (Figure 1.10). Une tumeur peut également libérer des cellules circulantes (CTC) et leur présence a un impact direct sur le pronostic dans plusieurs pathologies telles que le cancer du sein¹⁸³⁻¹⁸⁵. Les cellules circulantes sont considérées comme la source majeure des métastases et leur nombre dans le sang est corrélé avec le taux de survie¹⁸⁶⁻¹⁸⁹.

Leur isolation tout d'abord par centrifugation et FACS puis par des systèmes de plus en plus complexes tels que la microfluidique ou un collecteur de cellules tumorales avec une aiguille recouverte d'anticorps fixés ont permis de mieux conserver et d'étudier ces cellules¹⁹⁰. L'intégrité des CTCs étant préservée leur mise en culture ou leur greffe est devenue possible permettant ainsi la recherche de traitement plus ciblée^{191,192}.

Les exosomes sont des vésicules extracellulaires permettant une communication entre les cellules. Leur sécrétion par les cellules permet le transfert de lipides, de protéines ou des ARN présents dans le cytosol¹⁹³. Les tumeurs peuvent libérer ces exosomes notamment dans les cancers avancés¹⁹⁴. La présence d'exosomes laisse supposer une "communication" entre la tumeur et un destinataire inconnu, notamment par la production d'intégrines. Le nombre élevé d'exosomes est synonyme de mauvais pronostic^{195,196}.

En plus d'être un marqueur direct de l'avancée de la maladie, l'isolation des exosomes par centrifugation par gradient ou par anticorps permet d'avoir accès à une majeure partie de l'ARN de la tumeur (messenger, non-codant ou miRNA), à certaines protéines (ubiquitaires, du cytosquelette) mais également à une partie du contenu du noyau (ADN, facteur de transcription). Les exosomes peuvent ainsi être très utiles dans la caractérisation d'une tumeur et son suivi dans le temps. Ce système de communication pourrait également être utilisé en défaveur de la tumeur en utilisant les exosomes comme véhicule de drogues¹⁹⁷.

Dans les exosomes, des miRNA peuvent être détectés et leur rôle dans le cancer n'est plus à prouver^{198,199}. Ces régulateurs peuvent être également facilement détectables dans le sang pouvant être présents dans des microvésicules ou des complexes propres^{200,201}. La présence et la quantité de certains miRNA sont un excellent outil de diagnostic.

En dehors des exosomes, l'ARN messenger circulant reste plus difficile à observer. L'analyse du profil d'expression tumorale à partir de biopsie liquide reste majoritairement faite sur une extraction d'ARN du plasma total incluant les exosomes voire les CTCs²⁰²⁻²⁰⁵. L'isolation des plaquettes sanguines montre un contenu de messagers provenant des tumeurs tels que la présence de messagers de *EGFR* et *PCA3* mutés dans gliomes et des cancers de la prostate ou de messagers de *TP53* mutés dans le cancer du sein²⁰⁶⁻²⁰⁸. Les plaquettes ou

d'autres microvésicules contenant des messagers seraient donc un autre moyen de communication entre la tumeur et les métastases.

1.3.2 Origine de l'ADN circulant :

L'ADN circulant (cfDNA) et sa fraction tumorale, l'ADN tumoral circulant (ctDNA) est un outil important dans le contexte des biopsies liquides. L'ADN circulant est constitué par de petits fragments d'ADN d'environ 160-170 paires de base correspondant à l'enroulement de l'ADN autour d'un nucléosome mais également par des fragments multiples de 160 provenant de di-, tri voire et même poly-nucléosomes (Figure 1.9)^{209,210}.

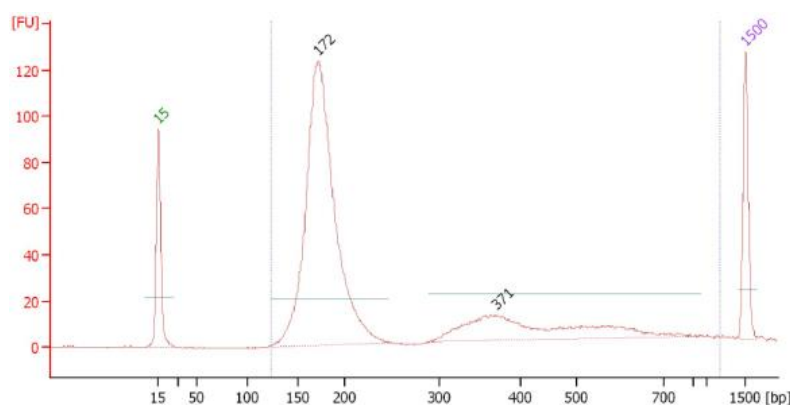


Figure 1.9 Electrophorèse capillaire d'un ADN circulant sur puces High Sensitivity passées sur un Bionanalyzer 2100 d'Agilent technologies. Un pic majoritaire de 170 pb ainsi que des multiples de ce pic sont observés.

Des fragments plus courts d'environ 90 paires de bases ont également été observés notamment chez des patients atteints de cancer²¹¹⁻²¹³. Ces tailles plus petites restent incomprises et marginales pouvant être dues soit à des différences d'enroulage de l'ADN suite à des altérations de la chromatine, soit à la fixation de l'ADN circulant à des facteurs de transcription au lieu d'un nucléosome ou soit tout simplement à un biais dans la préparation des échantillons.

L'ADN circulant peut provenir d'une sécrétion active par les cellules mais provient majoritairement de l'apoptose et de la nécrose cellulaire²¹⁴⁻²¹⁶. Les quantités d'ADN circulant peuvent varier chez des individus sains en fonction de l'âge, du genre, de l'indice de masse corporelle, de l'hygiène, des menstruations ou d'une grossesse mais elles peuvent également être impactées par des maladies telles que des infections ou des cancers²¹⁷⁻²¹⁹.

La demi-vie de l'ADN circulant est de 16 minutes à 3 heures avec un renouvellement complet pouvant aller de 2 jours à 2 semaines²²⁰⁻²²³. Sa dégradation n'est pas encore comprise mais pourrait provenir de la découpe par une DNase du sang, ou d'une filtration par le rein et une sortie dans l'urine ou d'une dégradation par les macrophages dans le foie²²⁴⁻²²⁷.

L'étude *in vitro* de macrophages en apoptose ou nécrose pouvant relâcher de l'ADN laissait supposer que l'ADN circulant proviendrait majoritairement des globules blancs chez un individu sain²²⁸.

La comparaison en WGS low coverage entre le profil d'ADN circulant et celui de l'ADN de certains tissus découpés par nucléosome de l'ADN avec la MNase a permis d'observer des "empreintes" similaires entre le cfDNA et l'ADN de certains tissus. La majorité de ces tissus étant des lignées lymphoïdes, myéloïdes ou issues de la moelle osseuse, l'origine majoritaire des cellules hématopoïétiques de l'ADN circulant a été ainsi confirmée^{229,230}. Le découpage par MNase de lignées cancéreuses générerait des empreintes également spécifiques. Dans le plasma des patients atteints de cancers, ces "empreintes" cancer étaient observés mais n'étaient pas directement liés au type de cancer. Par exemple, le profil d'un ADN circulant pouvait contenir des profils issus de lignées du cancer du sein et de la prostate alors que le patient était atteint de carcinome hépatocellulaire.

1.3.3 Découverte de l'ADN circulant :

L'ADN circulant fut suggéré en 1931 avant même la découverte de la structure de l'ADN par Watson & Crick^{231,232}. Le dosage du phosphate nucléique dans le sang d'individu sain puis de personnes atteintes de leucémie montrait des différences significatives de phosphate nucléique dans le sang mais la présence de leucocytes devait certainement perturber ces résultats.

Ce n'est qu'en 1948 que le dosage des phosphates nucléiques a été réutilisé non pas dans le sang mais sur le plasma sanguin pour s'affranchir de la présence des leucocytes²³³. Avec la technique de Schmidt & Thannhauser, les phosphates ribonucléiques pouvaient être dosés séparément des phosphates désoxyribonucléiques mais également de ceux provenant de phosphoprotéine²³⁴. Dix sujets "sains" et 15 avec une pathologie (diabète, insuffisance cardiaque, cirrhose, "grossesses",...) ont été prélevés. Peu de variations furent observées entre les individus hormis un sujet avec un niveau bas ayant un régime spécifique et un taux élevé d'acides nucléiques pour les femmes enceintes. Des niveaux importants d'acides nucléiques ont été également observés dans des individus atteints de cancer mais non significatifs.

En 1973, les taux d'ADN circulant ont pu être analysés au cours temps sur des patients atteints de différentes maladies telles que le lupus²³⁵. Les quantités d'ADN circulant étaient extrêmement élevées chez les patients atteints de lupus ou de polyarthrite rhumatoïde et élevées chez les patients atteints de cancers notamment de leucémie. Lors d'épisode de glomérulonéphrite pour les patients atteints de lupus, la quantité d'ADN circulant augmentait fortement pour redevenir ensuite indétectable. Ces quantités fluctuaient également en fonction des épisodes d'exacerbation.

En 1977, le dosage de l'ADN circulant fut réalisé chez 173 patients atteints de 11 différents types de cancer tels que le cancer du poumon, du sein et des organes génitaux féminins avant et après radiothérapie²³⁶. Quelques différences furent observées entre cancers mais elles étaient non significatives. Cependant, les cas métastatiques présentaient des niveaux beaucoup plus élevés d'ADN circulant et ce quel que soit le type de tumeur. La quantité d'ADN circulant baissait chez les patients répondant favorablement aux traitements et augmentait ou demeurait inchangée chez les patients pour lesquels le traitement n'avait pas fonctionné. Pour la première fois et même si la source de l'ADN circulant chez les individus sains restait controversée, l'ADN tumoral circulant (ctDNA) fut mise en évidence avec rélargage direct des tumeurs et proportionnel à sa malignité.

Par la suite, ces résultats ont été confirmés dans de nombreuses études notamment avec des niveaux très élevés d'ADN circulant dans les cas métastatiques mais également dans certains types de cancer avec un pronostic très défavorable tels que le carcinome du pancréas²³⁷⁻²³⁹. L'utilisation de l'ADN circulant comme outil de diagnostic commençait à germer.

A partir de 1993, des mutations dans le gène *NRAS* ou *KRAS* ont pu être détectés par PCR dans le sang et ces mutations étaient communes à celles retrouvées dans les tumeurs de patients atteints de cancer colorectal ou de leucémie²⁴⁰⁻²⁴². Dans des cancers des voies aériennes supérieures, l'analyse de microsatellite par PCR permis d'observer des pertes d'hétérozygotie communes dans le sang et leurs tumeurs associés²⁴³. Le ctDNA est ainsi validé comme étant un moyen d'accéder à l'information génétique de la tumeur par une simple prise de sang.

À partir des années 2000 avec la démocratisation du séquençage et l'augmentation de la sensibilité des méthodes de biologie moléculaire, l'ADN circulant a permis de mieux comprendre les tumeurs et de pouvoir suivre l'évolution de la maladie dans le temps. Mais ce sont finalement des recherches hors oncologie qui ont permis de mieux exploiter cet outil.

1.3.4 Le Diagnostic PréNatal Non Invasif (DPNI)

Lors de la grossesse, le développement du fœtus peut amener au transfert de lymphocytes entre fœtus et mère, ce qui permet de supposer la présence d'ADN circulant dans le sang de la mère²⁴⁴.

Les premiers tests ont été basés sur la détection du chromosome Y dans le sang de la mère attendant des enfants de sexe masculin^{245,246}. Dans des études initiales, le chromosome Y était détecté dans 80% des cas pour le plasma et 70% dans le sérum.

La confirmation d'ADN fœtal circulant (cffDNA) pouvait être utile dans le diagnostic pré natal. L'ADN fœtal a pour origine les cellules apoptotiques du placenta et des trophoblastes. Il peut représenter 3 à 13% de l'ADN circulant total de la mère²⁴⁷⁻²⁴⁹. L'utilisation du cffDNA constitue ainsi une excellente alternative à l'amniocentèse.

La recherche du statut rhésus de l'enfant, de maladies liées aux chromosomes sexuels ou de maladie liée à des mutations présentes chez le père a pu être proposée par l'analyse du cffDNA²⁵⁰⁻²⁵². Dans un premier temps, la recherche d'aneuploïdie de certains chromosomes telle que la trisomie 21 provoquant le syndrome de Down, la trisomie 18 pour le syndrome d'Edwards ou la trisomie 13 pour le syndrome de Patau restait initialement plus compliquée.

Cependant, l'analyse de la méthylation de 22 îlots CpG a permis de différencier la trisomie 21 du fœtus et celle du gène *SERPINB5* de différencier la trisomie 18 par PCR bisulfite n'amplifiant que les nucléotides convertis par bisulfite^{253,254}. Puis, la sensibilité de la

droplet PCR a permis de mieux détecter ces trisomies avec des protocoles plus simples et moins coûteux²⁵⁵.

Finalement, le séquençage de nouvelle génération a permis de franchir une nouvelle étape avec la possibilité de regarder plus qu'un seul chromosome à savoir la totalité du génome^{256,257}.

Ainsi, des kits commerciaux de DPNI ont été rapidement proposé avec une sensibilité de détection de 99% pour les trisomies 21 et 18 et 80% pour la trisomie 13²⁵⁸.

En 2013, une femme de 37 ans n'ayant aucun antécédent médical se vue proposer ce type de test en raison de son âge avancé pendant sa grossesse²⁵⁹. Une trisomie 13 et une monosomie 18 furent détectées à partir de deux aliquots différents et une amniocentèse a été alors proposée. Un caryotype, un marquage FISH et une puce SNP ne détectèrent aucune anomalie chromosomique chez le fœtus. L'hôpital conclut à une erreur de prélèvement.

Un nouveau test fut réalisé pendant la grossesse selon le protocole établi par l'hôpital amenant au même résultat que le précédent. Lors de l'échographie du fœtus, aucune anomalie ou retard de développement n'ont été observés. Au troisième trimestre, la patiente se présenta à l'hôpital avec des saignements vaginaux et la présence d'une masse dans le cervix correspondant à un léiomyome mais l'opération était trop risquée en fin de grossesse.

Deux semaines après l'accouchement, la masse dans le cervix avait disparu mais une autre masse était présente dans le vagin. Cette tumeur fut classée en carcinome neuroendocrine de haut grade avec de nombreuses métastases. L'analyse par FISH de la tumeur et de ces métastases permit d'observer une trisomie 13 et une monosomie 18 et aucune de ces altérations ne fut trouvée dans le placenta.

Le test est maintenant très utilisé aux Etats-Unis et en cas de suspicion d'aneuploïdie chez le fœtus et suite à cette analyse, une amniocentèse est réalisée pour confirmation. Ainsi, une relecture de 450 000 tests effectués sur 3 ans a permis de détecter 55 femmes avec des génomes remaniés non attribués au fœtus. Douze femmes n'avaient pas été suivies mais les 43 autres avaient un suivi régulier dont 18 ont développé des tumeurs malignes telles que des lymphomes de Hodgkin ou des cancers du sein²⁶⁰.

Ces cas particuliers consentirent à l'utilisation des méthodes de DPNI pour la caractérisation de tumeur, notamment très remaniées²⁶¹.

1.3.5 Autres utilisations de l'ADN circulant

Lors d'une greffe d'organe, la présence d'ADN circulant du donneur peut être détectée²⁶². La détection de chromosome Y dans le sang de femmes ayant reçu un transplant d'un donneur masculin comme le foie, le rein ou le cœur souligne que l'organe greffé libère de l'ADN circulant^{263,264}. Ce largage d'ADN pourrait provenir de la mort des cellules de la greffe et donc pourrait être utilisé comme marqueur de rejet.

Le suivi de l'ADN circulant donneur spécifique (cfdDNA) par PCR sur le chromosome Y puis par des panels dans le temps a ainsi permis d'observer l'augmentation de ces quantités d'ADN lors d'un rejet^{265,266}.

Les traitements par immunosuppresseurs pour éviter le rejet rendent les patients vulnérables aux infections bactériennes ou virales. Le suivi en NGS de l'ADN circulant de patients greffés a pu ainsi détecter la présence d'ADN viral ou bactérien et leur évolution a été confirmée par qPCR²⁶⁷. L'analyse du virome pour détecter des virus tels que les hépatites, le VIH ou la grippe ou du microbiome en cas d'infection est possible avec l'ADN circulant^{268,269}.

Les études sur le transplant et sur le virome peuvent être appliquées au cancer avec des dynamiques de libérations rapides de l'ADN circulant en cas de réponse positive à un traitement mais également diagnostiquer la présence de certains virus oncogène tels que HpV^{270,271}.

1.3.6 Utilisation de l'ADN circulant pour la caractérisation et le suivi du cancer

Même s'il a été mis très tôt en évidence, la proportion d'ADN tumoral circulant dans le sang étant très faible, une augmentation de la sensibilité de détection des méthodes de biologie moléculaire était nécessaire.

Des mutations détectées dans les tumeurs par séquençage Sanger purent être suivies dans des plasmas provenant de patients atteints de cancers colorectaux²²². La mise en place de BEAMing (PCR à émulsion couplé à des billes magnétiques) avec une sensibilité de 0,01 % et validés ensuite par qPCR a permis de détecter les niveaux d'ADN tumoral circulant ainsi que leur dynamique selon le traitement. Les quantités d'ADN étaient corrélées avec les données cliniques par une diminution des quantités d'ADN tumoral circulant lors des chimiothérapies ou chirurgies et une augmentation lors de l'apparition de rechute.

Par la suite, la droplet PCR fit chuter le seuil de sensibilité à 0,001%²⁷²⁻²⁷⁵. Cette augmentation de la sensibilité a rendu possible le suivi de la maladie à un niveau résiduel minimal (MRD) et de prévenir ainsi une éventuelle rechute.

Le plasma sanguin ne contient pas forcément les informations génétiques de tous les types de tumeurs. Le ctDNA des tumeurs cérébrales ou des métastases dans le système nerveux est vraisemblablement bloqué par la barrière hématoencéphalique. L'utilisation du

liquide céphalo-rachidien (LCR) prélevé par ponction lombaire ou directement dans un réservoir posé en cas d'hypertension intracrânienne est plus adaptée pour ce type de tumeur²⁷⁶.

Par conséquent, certaines mutations de gènes propres aux gliomes tels que *IDH1*, *IDH2*, *TP53*, *ATRX*, *TERT*, *H3F3A* et *HIST1H3B* ont pu être détectées dans le LCR dans 80% des cas^{277,278}.

Par ailleurs, l'urine peut également être plus adaptée à certains types de cancer. Une étude de la population iranienne a permis de détecter dans l'urine deux hotspots de mutations dans le gène *TERT* impliqués dans les tumeurs urothéliales^{279,280}. L'immensité de la cohorte et son suivi strict malheureusement dû au régime du pays ont permis de supposer la présence de cellules tumorales dans la vessie 10 ans avant le diagnostic du cancer. Ces cellules pré-cancer pourraient ainsi être en attente d'un choc environnementale ou d'erreur de réplication pour devenir malignes.

Cependant, malgré une sensibilité et une sensibilité très forte, les différents types de PCR à émulsions ou en gouttelettes ne sont limités qu'à 2 ou quelques mutations par réaction et demandent un biomarqueur récurrent de la maladie ou une connaissance préalable de la tumeur^{188,281–284}.

L'utilisation de panels de gènes avec des identifiants moléculaires uniques (UDI) permettant d'identifier une molécule d'ADN avant amplification et couplé au séquençage haut-débit a conduit à suivre l'évolution d'une tumeur avec une sensibilité similaire à la droplet PCR^{285–287}.

Des panels propres à certaines maladies telles que le cancer du poumon à non petites cellules ont ainsi été dessinés suite à l'évolution des bases de données de séquençage telles que le TCGA²⁸⁸. Le panel CAPP-seq (CAncer Personalized Profiling by deep SEQuencing)

couvrait ainsi 139 altérations soit 95% des mutations observées dans les bases de données donc 407 WES issus de TCGA. Le panel réduit à 125kb pour diminuer les coûts de séquençages avait une sensibilité théorique de 0,01% avec une couverture de 30 000X. Pour chaque patient, le panel détecta des mutations et pour la plupart les quantités de ctDNA théoriques étaient corrélées avec la taille de la tumeur et sa réponse au traitement. Pour un cas, le niveau de ctDNA augmentait dans le temps alors que cliniquement une diminution de la taille de la tumeur était observée. Le patient décéda des suites du cancer laissant à penser à la progression de métastases microscopiques de la maladie.

D'autre part, un panel de séquençage haut débit incluant 354 gènes et permettant également l'analyse d'altérations en nombre de copie, ou CNA (Copy Number Alteration), a été utilisé sur le ctDNA présent dans le LCR de patients atteints de tumeurs cérébrales ou de cancer avec métastases dans le système nerveux central²⁸⁹. Des altérations en nombre de copie ont été détectées dans 50% des cas de tumeurs cérébrales et 63% des cas métastatiques. Certains types de tumeurs cérébrales tels que le gliome semblaient relâcher peu ou pas de ctDNA. Dans la plupart des cas, le ctDNA était en accord avec les données cliniques : une détection de copies mutées d'ADN circulant lors de la rechute ou un envahissement du système nerveux par des métastases et également l'observation de nouvelles altérations telles que des mutations dans le gène *KRAS* qui n'étaient pas présentes dans la tumeur primaire mais dans la tumeur à la rechute.

Il est maintenant proposé que ces méthodes de suivi de l'ADN tumoral circulant permettent de détecter une rechute voire même de prédire le développement d'un cancer, avant tout signe clinique. Cependant, elles demandent une connaissance préalable de la maladie ou une caractérisation de la tumeur ne permettant pas de s'affranchir d'une biopsie.

1.3.7 Caractérisation moléculaire des cancers et classification par l'ADN circulant par des approches génomiques

En fonction du stade de la maladie, une tumeur peut plus ou moins libérer de l'ADN tumoral circulant rendant la détection plus complexe pour les stades précoces. Cette fréquence de détection et de libération de ctDNA diffère également entre les types de cancer (Figure 1.11)²⁹⁰.

Pour certains types de cancer ou pour des stades avancés, le nombre de molécules mutantes est plus important et pourrait être détecté avec des sensibilités moindres. Des analyses de types WES ou WGS avec des couvertures acceptables pour limiter les coûts de séquençages sont donc possibles.

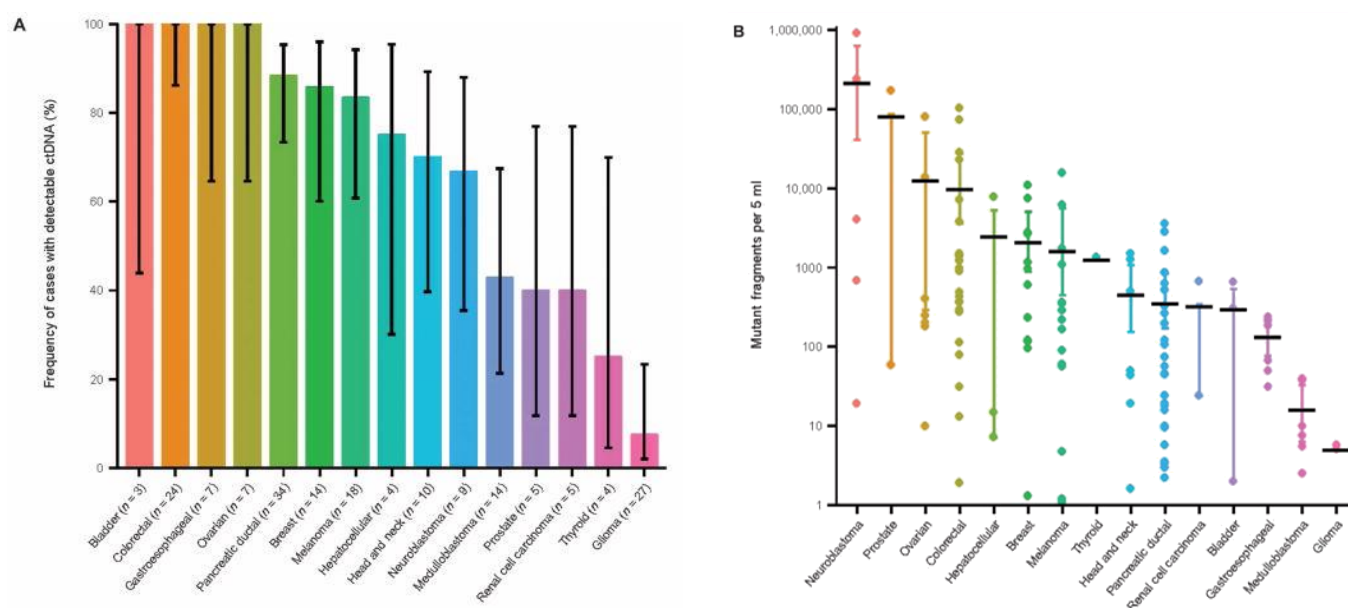


Figure 1.11 : Fréquences de détections de ctDNA (A) et quantité d'ADN tumorale circulant (B) selon le type de cancer (Figure en accès libre)²⁹⁰.

Des WGS avec une couverture de 10x ont pu être réalisés sur le plasma de patient atteint de cancer métastatique colorectal ou du sein²⁹¹. La faible couverture permettait seulement d'observer des grandes altérations en nombre de copie (1Mb) ou de translocation mais sur la totalité du génome. Dans le plasma, des profils similaires à la tumeur furent observés pour

chaque cas. Le WGS permis également de calculer le contenu tumoral présent dans l'ADN circulant avec des fractions de ctDNA allant de 12 à 44% dans les plasmas prélevés au moment du diagnostic et au moment de la rechute.

Avec des protocoles adaptés aux faibles quantités de départ d'ADN circulant, les mutations pouvaient être détectées par WES dans le plasma de patients atteints de maladie avancée telle que le cancer du poumon à non-petites cellules²⁹²⁻²⁹⁴. De nombreuses mutations communes et également ciblables par chimiothérapie notamment dans les gènes *ALK*, *PIK3CA* ou *PTEN* étaient observés dans la tumeur et dans le plasma. Toutes les mutations détectées dans la tumeur n'étaient pas observées dans le plasma. Cependant certaines mutations étaient observées uniquement dans le plasma touchant des gènes avec un fort potentiel thérapeutique tels que *MTOR*. Le séquençage supplémentaire par WES de métastases mis en avant la présence de mutations communes entre l'ADN circulant et les métastases mais absente dans la tumeur telles qu'une mutation dans le gène *ESR1*²⁹⁵. Cette différence entre tumeur et plasma pourrait être due à l'hétérogénéité spatiale de la maladie.

Dès lors, l'analyse des plasmas chez des patients atteints de tumeurs résistantes à des traitements a permis d'observer une évolution avec l'apparition de nouvelles mutations²⁹⁶. Une augmentation de la fraction d'allèle muté fut observée dans les gènes *PIK3CA*, *BMI*, *SMC4* ou *MTOR* dans des cas de cancer du sein et de cancer des ovaires suite à un traitement au Paclitaxel. Suite au traitement par Gefitinib, des mutations des gènes *EGFR*, *TP53* ou *NFKB1* sont apparues dans un cas de cancer du poumon. La plupart de ces mutations sont connues pour amener à une résistance dans des traitements affiliés. L'ADN circulant a ainsi permis d'observer l'évolution clonale sans information préalable de la tumeur.

Les informations données par le WES peuvent être précieuses dans la décision d'un traitement mais ne permettent pas toujours de classer une tumeur telle que les différents groupes du médulloblastome. La méthylation de l'ADN joue un rôle important dans la différenciation des cellules et reste spécifique à un type de tissu^{297,298}. L'utilisation des données présentes dans TCGA comme les profils d'altérations en nombre de copie, d'expression, de méthylation ou protéique pour 10 000 tumeurs correspondant à 33 types de cancers a démontré que la méthylation des tumeurs conserve effectivement une empreinte de leurs tissus d'origine²⁹⁹.

Ainsi, l'étude de la méthylation de l'ADN de la tumeur peut permettre de classer certaines tumeurs dont les observations anatomopathologiques n'ont pas rendu possible à classer.

Lors d'une étude d'une cohorte de 1000 tumeurs, la fusion de *EWS* n'a pas pu être mise en évidence dans 30 tumeurs à forte suspicion de cancers indifférenciés "Ewing Like" et ce malgré l'utilisation de FISH, RNA Seq sur tissus fixés en FFPE et marquage immunohistochimique³⁰⁰. L'utilisation de puces de méthylation a permis de classer 14 de ces tumeurs en sarcome d'Ewing. Ainsi, une nouvelle analyse étendue par RNA Seq total couplé à une FISH a permis de caractériser des fusions rares.

Les bases de données de méthylation étant de plus en plus importantes permettent aujourd'hui de classer certains types de tumeurs notamment les tumeurs cérébrales^{301,302}. La base de données du DKFZ (Deutsches KrebsForschungsZentrum) contient maintenant près de 30 000 méthylomes de tumeurs obtenues par puces de méthylation Illumina 450k ou Epic. Le clustering de 401 tumeurs cérébrales avec ces bases de données a ainsi permis une nouvelle classification pour 50 cas (12%) mais 89 cas ne purent être classifiés suite à une mauvaise qualité d'échantillons.

La méthylation de l'ADN circulant peut également être un outil dans la caractérisation mais surtout la classification des tumeurs. Des outils de machine learning tels que CancerLocator classifient des tumeurs avec une précision entre 50 et 80% à partir de contenus de ctDNA très faibles (>5%) dans le plasma mais utilisant des données issues de WGS traité au bisulfite, technique demeurant très coûteuse³⁰³.

Le méthylome de l'ADN circulant a permis une nouvelle fois de mettre en évidence son origine hématopoïétique mais également par des outils de déconvolution de mettre en évidence la présence de greffes, de choc septique et de cancer³⁰⁴⁻³⁰⁶.

En fonction de son organe d'origine et de son épigénétique, l'ADN circulant aurait une fragmentation différente due à des condensations autres de certaines régions de la chromatine²³⁰. De plus, un millilitre de plasma contient environ 2000 équivalents génomes pouvant rendre la détection de mutations compliquée pour les stades précoces de certains cancers mais les altérations plus importantes des fragments de tailles différentes d'ADN circulant ont possiblement plus de chance d'être détectées³⁰⁷. Ainsi, le WGS low coverage sur l'ADN circulant de 208 patients atteints de cancer du sein, colorectal, du poumon, ovarien, du pancréas ou gastrique comparé à celui de 215 ADN circulant d'individu sain a permis de révéler des profils très différents de découpage de l'ADN circulant³⁰⁸. Après enrichissement par TargetSequencing de certains locus de gènes impliqués dans ce type de cancer tels que *TP53* ou *BRAF*, des fragments montraient des tailles différentes de 6,5 pb de moyenne pouvant être plus courtes jusqu'à 30 pb ou plus longues jusqu'à 40 pb. Le suivi de ces fragmentations pendant un traitement aux inhibiteurs de tyrosine kinase pour 19 patients atteints de cancer du poumon montrait une corrélation avec la fraction allélique des mutations détectées dans le plasma.

L'utilisation de "Machine Learning" sur la fragmentation ainsi que les mutations dans l'ADN circulant a permis de détecter un cancer dans 91% des cas, même précoce. L'outil nommé DELFI (DNA EvaLuation of Fragments for Early interception) pourrait ainsi être utilisé pour la caractérisation précoce de certains types de cancer et demanderait cependant un enrichissement préalable de WGS low coverage sur l'ADN circulant pour d'autres cancers et des biomarqueurs récurrents permettant à l'algorithme d'augmenter sa sensibilité et sa spécificité.

1.3.8 Etude de l'évolution clonale par l'ADN circulant

L'ADN tumoral circulant, représentant une partie de l'hétérogénéité clonale de la tumeur, a également permis de mieux comprendre l'évolution clonale. Le séquençage de type WES ou WGS permet donc d'identifier l'ensemble des altérations génétiques somatiques d'un cancer à partir de l'ADN circulant. Ces approches peuvent donc être utilisées non seulement pour une caractérisation moléculaire à un moment donné, mais pour le suivi dans le cadre d'échantillons séquentiels.

Cependant le WES sur l'ADN circulant, s'il peut être réalisé à partir de faibles quantités d'ADN (<10 ng), son analyse sera possible en fonction d'une couverture de séquençage adaptée et d'un contenu tumoral suffisant dans l'ADN circulant. Par exemple, un contenu tumoral supérieur à 10% permettra de détecter des mutations avec fraction allélique de 5% à condition de réaliser un WES avec une couverture supérieure à 150x^{292,309}.

Pour connaître le contenu tumoral dans le plasma soit la fraction ctDNA dans l'ADN circulant, une présélection des plasmas est possible avec l'utilisation de WGS low coverage de 0,1x à 1x. Le séquençage du génome entier à faible couverture permet d'observer les grandes altérations en nombre de copies et leur ratio allélique d'inférer une fraction tumorale. En fonction du contenu tumoral dans l'ADN circulant, les WES des plasmas ont ainsi plus de

chances d'être contributifs avec la détection d'altération. Les cas avec de hauts contenus tumoraux peuvent quant à eux permettre la détection d'événements sous-clonaux³¹⁰.

Chez une patiente atteinte d'un cancer du sein, le séquençage par WES de deux parties de la tumeur primaire, d'un nodule lymphatique, des 5 métastases présentes dans le cerveau, le poumon, le foie, la moelle et les ovaires ainsi que 9 plasmas sanguins à différents temps de la maladie a permis de décrire un clone primaire, un métastatique et 8 privés³¹¹. Le suivi de ces clones dans le temps a permis de voir une certaine stabilité de la tumeur malgré le traitement avec des variations plutôt sous clonales laissant supposer une synergie entre clone majoritaire et sous clones pour l'adaptation. Par exemple, suite au traitement par le lapatinib et capecitabine, il a été mis en évidence dans le plasma l'explosion d'un clone très proche de la métastase du poumon et fut corrélé à la présence d'un nouveau nodule dans cet organe. Malheureusement la patiente est décédée 4 mois après le dernier prélèvement de plasma d'une aggravation générale avec une progression très importante de la maladie dans les poumons (nouveaux nodules, effusion pleurale, masse dans la paroi thoracique).

Pour des patients atteints de tumeurs cérébrales, l'ajout du séquençage par panel ou WES conjugué à la droplet PCR du liquide céphalo-rachidien permis de suivre la maladie et l'évolution de son hétérogénéité dans le système nerveux central³¹². L'analyse de plasma de patients de gliomes ne détecta aucune altération tumorale contrairement à l'analyse du LCR de mêmes patients. Chez les patients atteints de cancers du sein ou de cancers du poumon métastatiques avec présence de métastases dans le cerveau, l'analyse du LCR a permis de montrer qu'ils présentaient des altérations communes aux méninges non retrouvées dans le plasma et d'autres sites de fixations. La modélisation phylogénétique de la tumeur laissait supposer que les clones originaires de l'envahissement du système nerveux central étaient

arrivés très tard dans le développement de la maladie accumulant de nombreuses altérations supplémentaires. De plus, l'évolution des altérations dans le LCR et le plasma était directement corrélée aux données cliniques. Par exemple, une augmentation de la fraction allélique mutante dans le LCR était corrélée avec le développement des métastases cérébrales alors que la fragmentation allélique mutante diminuait dans le plasma. En revanche, la diminution de la fréquence allélique des mutations dans le LCR était corrélée avec une régression des métastases cérébrales alors que la fréquence des mutations dans le plasma augmentait, indice de prolifération de la maladie en dehors du système nerveux central.

En somme, l'analyse des LCR est un très bon outil de diagnostic pour les tumeurs cérébrales et peut-être également complémentaire à celle du plasma pour suivre l'évolution d'un cancer métastatique.

Cependant, pour la détection de la maladie résiduelle, le WES n'est pas le meilleur outil. Une augmentation de sa couverture pour augmenter la sensibilité est théoriquement possible mais le coût et la compétition possible entre les sondes de capture rendraient sa mise en œuvre très compliquée.

Grâce aux découpages préalables des tumeurs, le consortium TracerX put développer un panel à partir de 100 patients atteints de cancer du poumon. Le séquençage de ces cancers à l'aide du panel a permis de détecter plus de 2 SNVs (Single Nucleotide Variant) dans le plasma chez 46 patients. Dans la majorité des cas, des SNVs sous clonaux provenant que d'une seule partie de la tumeur étaient observés dans le plasma. La fréquence allélique de ces mutations dans le plasma était fortement corrélée à la taille de la tumeur. Grâce à ce panel utilisé avec un séquençage à très haute profondeur (40 000x), la dynamique des clones a pu être suivie. Ainsi, l'émergence de certains clones a permis de prédire une rechute.

Le séquençage de l'exome à partir de l'ADN circulant permet d'avoir une vue plus globale de la maladie et d'identifier de nouvelles cibles en cas de rechute mais sa faible sensibilité ne permet pas pour le moment de prédire l'émergence d'un clone. La baisse continue des coûts de séquençage et la facilité de construire un panel pourraient permettre de proposer des panels personnalisés à des patients avec des séquençages plus larges en cas de prévention de rechute.

Cependant le séquençage d'ADN circulant, issu de quantité infime d'ADN, doit être couplé avec des filtres bio-informatiques strictes pour éviter les faux positifs et pour exclure le mosaïcisme des cellules hématopoïétiques pouvant être détecté comme une mutation tumorale^{313,314}. Ces filtres peuvent être notamment utilisés avec des bases de données de séquençage publiques. Associé à du "Machine Learning", il est ainsi possible de classer une tumeur en fonction de critères d'agressivité d'un cancer notamment dans les stades précoces en fonction du séquençage du plasma. En effet, un programme "Lung-CLiP" pour classer le stade précoce soit en stade à risque a pu ainsi être proposé avec une sensibilité théorique de 98% suite à un apprentissage à partir du séquençage de 104 tumeurs et plasmas de patients atteints de cancer du poumon et 56 plasmas sains contrôles³¹⁵.

1.3.9 Transcriptome et épigénétique de la tumeur à partir de l'ADN circulant

En plus de l'étude des altérations génétiques, une analyse de l'expression des gènes et/ou épigénétique est importante dans de multiples contextes. Ceci peut en théorie contribuer à une classification pathologique des cancers. De plus, le profil d'expression et d'épigénétique peut être un indicateur de l'identité cellulaire et ainsi permettre une étude de plasticité cellulaire.

Néanmoins, en dehors de la présence de certains marqueurs mésenchymateux dans les CTCs, les biopsies liquides n'ont jamais mis en évidence la plasticité de la tumeur¹⁹⁰. Cependant certaines propriétés de l'ADN circulant pourraient être utilisées pour observer ces transitions.

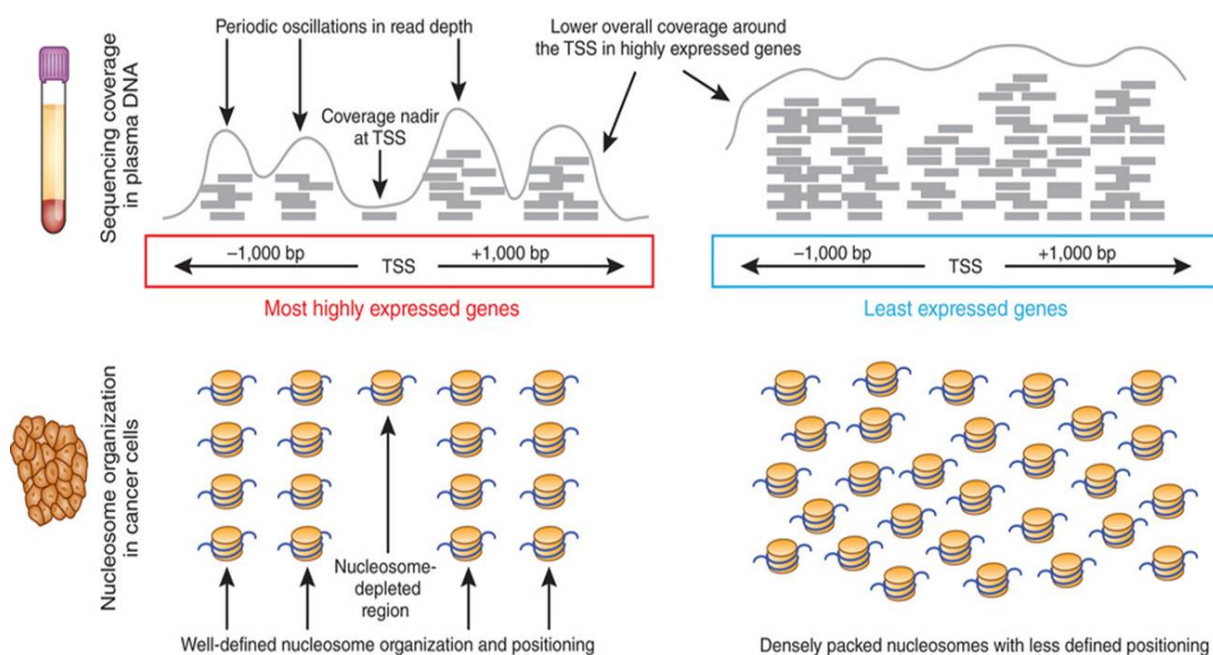


Figure 1.12 : Corrélation entre expression, couverture et positionnement des nucléosomes (Figure en accès libre)³¹⁶.

La difficulté d'isolation de l'ARNm circulant ne permet pas de suivre l'évolution de l'expression d'une tumeur dans le sang. Cependant le découpage de l'ADN circulant selon les nucléosomes permet de voir certains motifs en WGS²³⁰. Au niveau du TSS (Transcription Start Site) de certains gènes, une diminution de la couverture était observée et cette observation était corrélée avec une forte expression du gène dans le sang³¹⁶. À l'inverse des gènes exprimés, les gènes silencieux ou peu exprimés avaient une couverture homogène de séquençage (Figure 1.12). En effet, ce "trou" au niveau de la couverture proviendrait certainement d'une région déplétée en nucléosomes suite à l'ouverture de la chromatine pour faciliter la fixation d'un facteur de transcription. Cette région serait ensuite perdue et ceci

serait dû à son exposition aux endonucléases sanguines telles que la DNaseI entraînant sa dégradation ou lors de l'extraction des ADN circulants.

Avec des outils informatiques et du "Machine Learning", la prédiction d'expression d'une tumeur est possible seulement avec l'analyse de couverture de séquençage au niveau des régions TSS. Cependant, une quantité de plus de 75% de ctDNA dans le plasma est préférable pour ce type d'analyse. La prédiction est correcte à plus de 78% avec notamment la prédiction d'expression d'isoformes de gènes ayant une origine de transcription différente mais ces résultats ne se limitent qu'à certains gènes présents dans des régions amplifiées du génome.

Ces modifications d'expressions sont dues à des variations épigénétiques dans la tumeur. Si l'analyse de la méthylation de l'ADN circulant est plutôt utilisée dans la caractérisation d'une tumeur, l'immunoprécipitation de la chromatine (ChIP) semble possible. En effet, l'ADN circulant est enroulé autour d'un nucléosome^{306,317-319}. La méthylation d'histone (H3K4me3, H3K4me2, H3K4me1, H3K36me3) pour des marques activatrices de gènes ou d'enhancer a pu être détectée par ChipSeq³²⁰. Mais les résultats étaient très bruités et même si des différences étaient observées entre des individus sains et cancéreux, le parasitage par les cellules saines n'a pas permis de mettre en évidence des oncogènes.

L'optimisation de ces premières études pourrait permettre la mise en évidence de la plasticité cellulaire avec l'activation ou l'inactivation de certains gènes de types mésenchymateux ou de la méthylation/acétylation de leur promoteur ou enhancer.

1.3.10 Biopsies liquides chez les enfants

L'utilisation de biopsies liquides est une promesse particulière pour les cancers pédiatriques. En effet, ces cancers pouvant toucher des enfants très jeunes voire des nourrissons, dans certaines situations un accès direct à la tumeur peut s'avérer difficile. De plus, certaines de ces tumeurs très nécrosées contiennent du matériel génétique très dégradé. L'agressivité et le développement vélocé de certaines tumeurs demandent également une caractérisation rapide de la tumeur.

Des prélèvements successifs de la tumeur ne sont que très rarement possibles en oncologie pédiatrique et concernant avant tout les leucémies.

Ainsi l'étude de l'ADN circulant est un outil puissant pour la caractérisation moléculaire et l'étude des altérations génétiques des cancers pédiatriques. Cependant, les volumes de sang pouvant être prélevés sont parfois limités (en fonction de l'âge des patients) mais sont en partie compensés par une quantité d'ADN circulant plus importante que chez l'adulte²¹⁸.

Le neuroblastome est la tumeur solide extra-crânienne la plus fréquente et le plasma des patients atteints de cette tumeur contient une quantité importante d'ADN circulant^{9,290}. Dès le début de l'utilisation de l'ADN circulant comme outil contre le cancer, le Centre Léon Bérard de Lyon (CLB) proposa de détecter l'amplification du gène *MYCN* dans le sang³²¹. Dans une première étude, le *MYCN* circulant fut détecté par PCR et PCR en temps réel dans le sang de 31 patients sur 32 avec un neuroblastome *MYCN* amplifié identifié en CGH et/ou FISH sur la tumeur et seulement dans le sang d'un patient sur 70 n'ayant pas d'amplification identifiée en CGH et/ou en FISH sur la tumeur. Huit patients dont le diagnostic initial avait pu mettre en évidence une amplification de *MYCN* ont également eu un plasma supplémentaire, prélevé après la chimiothérapie. Pour la plupart, le signal de *MYCN* détecté dans le plasma diminuait

et pour un patient, le signal augmentait et ce, deux mois avant le diagnostic d'une rechute. L'étude a pu être étendue et a permis d'inclure tous les stades de neuroblastome. L'amplification de *MYCN* dans le sang était détectée par qPCR chez plus de 82% des stades avancés mais seulement chez 10% chez des patients atteints de stades peu avancés et localisés (1 et 2). Le gain du 17q pu également être détecté dans le sang avec une sensibilité de plus de 50% par qPCR³²².

Avec la sensibilité de la PCR en gouttelettes, 2 hotspots de mutation dans le gène *ALK* ont pu être détectés avec un taux de détection de 0,15%³²³.

Ces premières études avec des quantités infimes de plasma (200 µL) permirent d'observer de fortes quantités d'ADN circulant dans le neuroblastome et d'ouvrir la voie à des études plus larges.

Par ailleurs, le profil génomique est un marqueur important dans le pronostic du neuroblastome^{29,30}. Une analyse du profil génomique dans le sang fut possible avec des protocoles adaptés tels que les puces Affymetrix Oncoscan™ compatibles avec des échantillons fixés en FFPE ou des dérivés de kit DPNI par WGS à faible couverture^{324,325}.

Dans la plupart des cas, les puces Oncoscan™ ont permis de montrer que les profils génomiques de la tumeur et du plasma étaient similaires mais majoritairement dans les cas métastatiques où la quantité d'ADN circulant était significativement plus importante (Figure 1.13).

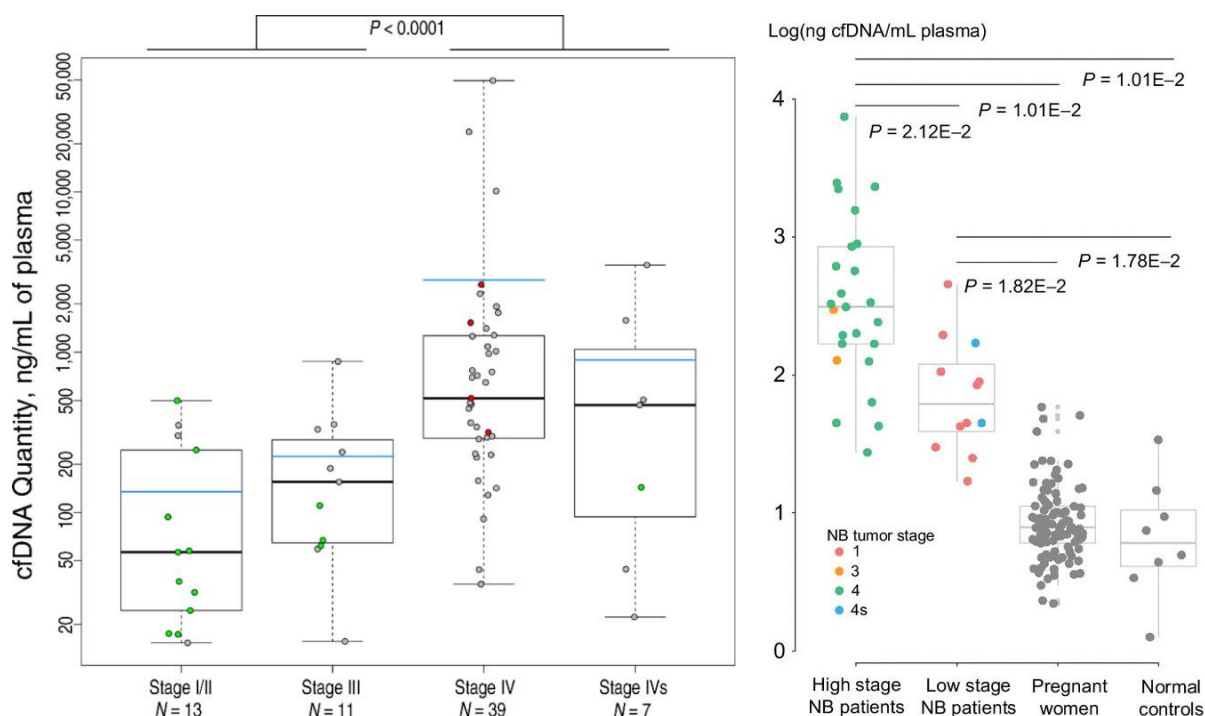


Figure 1.13 : Quantité d'ADN circulant dans le neuroblastome (ng/mL de plasma) selon les stades (Figure en accès libre)^{324,325}.

Certaines altérations n'étaient trouvées que dans le plasma telles que des gains ou amplification de *TERT*, *IGF1R* ou *LIN28B*. La comparaison des profils de la tumeur, du plasma et de la moelle osseuse envahie laissait penser que le ctDNA provenait majoritairement des métastases. Pour certains cas où le profil de la tumeur ne présentait aucune altération, le profil du plasma était dynamique démontrant ainsi l'intérêt de ces méthodes comme aide au diagnostic moléculaire dans des situations où un prélèvement tumoral ne serait pas informatif.

Le WGS à faible couverture permet également de détecter des altérations caractéristiques et pronostiques dans de nombreux cancers pédiatriques en plus du neuroblastome tels que les tumeurs de Wilms, ostéosarcome ou rhabdomyosarcome³²⁶.

L'isolation d'exosome par FACS à partir du sang de patients atteints de neuroblastome a permis de caractériser trois miR-342-3p, -29c, et let-7b par qPCR avec une baisse d'expression significative pour des patients ne répondant pas à la chimiothérapie³²⁷. Deux de

ces miRNA (miR-29c et let-7b) auraient une action directe sur le gène *MYCN* et le miR let-7b induirait l'apoptose ainsi que l'inhibition de l'expansion de la cellule. Ces miRNA pourraient être des marqueurs importants dans la réponse aux traitements.

La méthylation du promoteur du gène *RASSF1A* dans le sang semble corrélée au pronostic du patient atteint de neuroblastome³²⁸. Alors que ce promoteur est en majorité hyperméthylé dans la tumeur (94%), sa méthylation est moindre dans le plasma (25%). Sa déméthylation est liée à un pronostic plutôt sombre. De nombreuses études sont en cours avec comme objectif le suivi de la méthylation du promoteur dans le temps et pourrait, à terme, être utilisé comme marqueur.

Pour d'autres types de tumeurs résultant d'un facteur de transcription chimère suite à la fusion de 2 gènes, l'observation de cette altération est plus complexe dans le plasma et demande une connaissance préalable des points de cassure de la translocation.

Ainsi, dans le sarcome d'Ewing, le suivi de la fusion *EWS-FLI1* dans le plasma n'a été possible qu'après sa caractérisation à l'aide d'un panel d'amplicons séquencés sur la tumeur pour pouvoir proposer des primers spécifiques pour détecter cette fusion chez 20 patients^{329,330}.

La fusion *PAX3-FOXO1*, liée au rhabdomyosarcome alvéolaire, a dû d'abord être séquencée en Sanger après une RT-PCR à partir de l'ARN puis de l'ADN de la tumeur pour caractériser le breakpoint³³¹. La maladie a pu ainsi être suivie dans le sang et également dans la moelle osseuse. Les quantités de cette fusion détectée par qPCR dans le sang étaient corrélées avec le statut de la maladie tel qu'une diminution de fragments d'ADN portant la fusion dans le sang lorsque la maladie régressait. L'analyse, de par ses nombreuses étapes et sa complexité, n'a été ainsi possible que pour un seul patient.

De façon surprenante, les tumeurs osseuses pédiatriques comme les ostéosarcomes peuvent parfois contenir des taux d'ADN tumoral circulant élevés. Cependant la détection d'altérations à des fréquences élevées telles que des translocations de *TP53* ou des mutations dans le gène *ATRX* dans le sang serait plutôt possible grâce à des nodules issus de métastases de la tumeur^{332,333}.

Les tumeurs du rein sont composées à 90% de tumeurs de Wilms et ces tumeurs sont traitées par chimiothérapie néo-adjuvante après confirmation clinique et radiologique mais leur histologie n'est réalisée qu'après néphrectomie, une biopsie étant trop risquée. Chez les patients atteints de ces tumeurs, la caractérisation par WES de l'ADN circulant permis de détecter des mutations associées aux tumeurs de Wilms dans des gènes tels que *WT1*, *DICER1* ou *CTNNB1* ou des altérations en nombre de copies telles qu'un gain du 7q dans le plasma de l'ensemble des cas étudiés³³⁴. Certains patients avec une maladie pourtant localisée possédaient également des altérations uniquement dans le plasma dans des gènes tels qu'*ARID1B*, *SIX1* et *CTNNB1* suggérant que l'analyse du plasma reflète mieux de l'hétérogénéité clonale que celle de la tumeur. Un cas de sarcome à cellules claires du rein et un de carcinome à cellules rénales présentait également des altérations communes entre la tumeur et le plasma. La caractérisation par WES pourrait être ainsi un outil intéressant dans le diagnostic moléculaire des tumeurs du rein.

Pour le rétinoblastome, 92% des altérations en nombre de copies détectées dans des tumeurs d'yeux énucléés étaient retrouvées dans l'ADN circulant de l'humeur aqueuse. Aucune analyse n'a été faite dans le sang mais la présence de ctDNA reste envisageable au vu de la vascularisation de ces tumeurs.

D'autre part, l'analyse avec un panel commercial Ion AmpliSeq™ Cancer Hotspot Panel v2 couvrant 50 oncogènes chez 44 patients atteints de 14 cancers pédiatriques différents (neuroblastome, tumeur rhabdoïde du rein, néphroblastome, hépatoblastome, ostéosarcome, Fibrosarcome,...) a permis de détecter dans le plasma de nombreuses altérations caractérisant le type de cancer dont les patients étaient atteints telles que des délétions de *SMARCB1* pour les tumeurs rhabdoïdes, de *WT1* pour les néphroblastomes (certainement caractérisé par qPCR car ce gène n'est pas inclus dans le panel) ou de mutations de *ALK* pour le neuroblastome³³⁵. Pour 3 cas, des mutations ont continué à être détectées après chirurgie de la tumeur primaire et des lésions métastatiques distantes ont pu être mises en évidence. L'ADN circulant est donc un excellent représentant de l'hétérogénéité de la maladie.

Cependant, les tumeurs cérébrales sont encore un défi. Si la maladie reste dans le système nerveux central, le ctDNA ne pourra franchir la barrière hématoencéphalique et ainsi ne sera pas détectée dans le sang. Certaines tumeurs cérébrales telles que le médulloblastome ou le gliome semblent diffuser du ctDNA dans le sang mais aucune donnée clinique n'a été rapportée pour ces cas (métastase extra-crânienne ?)²⁹⁰. L'utilisation du liquide céphalorachidien est donc recommandée malgré sa difficulté de prélèvement et sa faible quantité. Des mutations dans des gènes propres au gliome tels que *IDH1* ou *H3F3A* (mutation K27M) ont pu être détectées par droplet PCR chez deux jeunes enfants dont un pour lequel aucune information génétique de la tumeur n'était disponible²⁷⁸. La mutation K27M dans le gène codant l'histone H3F3A a également pu être détectée dans le LCR en séquençage Sanger chez 2 enfants atteints de gliome. Cependant, des volumes de LCR permettant cette détection étaient élevés (de 0,4 à 2 mL) et ceci avait été possible suite à la pause de réservoir pour le traitement chez ces patients³³⁶.

Même si les études sur l'ADN circulant dans les cancers pédiatriques sont peu nombreuses, l'ADN circulant commence à être utilisé comme outil de diagnostic clinique (recherche du *MYCN* circulant, en Oncoscan™ ou WGSlc sur le plasma pour les cas de neuroblastome).

En recherche, la plupart des études se limitent à un panel restreint de gènes et ne permettent pas de suivre la maladie dans sa totalité et mieux comprendre l'échappement aux traitements. Au vu du contenu tumoral des plasmas de patients atteints de tumeurs pédiatriques, des séquençages couvrant une majeure partie du génome sont possibles tels que le WES voir le WGS avec des couvertures standards. Certaines métastases semblent avoir acquis de nouvelles altérations qui peuvent être détectées grâce à l'ADN circulant. Ces altérations supplémentaires par rapport à la tumeur sont une aide précieuse dans la compréhension de l'évolution de la maladie et de son hétérogénéité. De plus, le mosaïsme dû aux cellules hématopoïétiques étant lié à l'âge, il est donc très peu présent chez l'enfant contrairement à l'adulte et il ne pourra donc pas parasiter la détection d'altérations sous-clonales^{313,337}.

En somme, l'hétérogénéité moléculaire et son évolution parfois imprévisible et agressive des tumeurs pédiatriques ainsi que leurs difficultés d'accessibilité pourraient rendre l'ADN circulant indispensable comme aide au diagnostic et au suivi clinique.

2 Hypothèse et objectif

Chez les enfants atteints de cancer, l'ADN circulant et ses nombreuses propriétés semblent être un outil adéquat pour permettre l'analyse des altérations génétiques d'une tumeur. De plus, l'ADN circulant pourrait être utilisé comme outil diagnostique en l'absence de biopsie permettant d'accéder à la tumeur.

En outre, cet outil permettrait de suivre l'évolution de la tumeur dans le temps (pendant le traitement, au cours du suivi) afin de détecter une modification des altérations génétiques pendant le traitement et en cas de rechute.

Avec la possibilité de prélèvements successifs peu invasifs, l'ADN circulant semble un bon moyen pour étudier l'hétérogénéité clonale et la plasticité des tumeurs.

Pour cette étude, les tumeurs pédiatriques ont été étudiées au diagnostic et dans leurs évolutions pendant le traitement à travers la caractérisation par l'ADN circulant.

A partir de protocoles de recherche clinique impliquant l'Institut Curie, de nombreux échantillons de plasmas ou de LCR provenant de patients atteints de tumeurs pédiatriques ont pu être collectés.

Ainsi, deux stratégies basées sur l'analyse de l'ADN circulant et ses propriétés ont pu être proposées et basées sur :

- L'étude de la génétique des tumeurs par l'ADN circulant reposant sur des séquençages par WES et des séquençages ciblés.
- L'étude pour inférer l'expression génique et l'épigénétique des tumeurs par l'ADN circulant par l'analyse des TSS, de la chromatine et du méthylome.

Finalement, dans certains cas, ces approches pourraient être utilisées comme aide au diagnostic moléculaire clinique avec une description d'altérations génétiques pouvant être des cibles thérapeutiques ou comme aide à la classification lors du diagnostic.

3 Résultats

Différentes cohortes d'enfants atteints de cancers ont permis de collecter un nombre importants d'échantillons de plasmas et de LCR (n=488).

Ces patients ont été inclus dans des études menées par l'Institut Curie ou en partenariat après signature du consentement selon la réglementation en vigueur :

- L'étude multicentrique non interventionnelle PHRC IC2007-09 a permis de récolter de nombreux plasmas prélevés chez des patients atteints de neuroblastome.
- Le programme multicentrique MappyAct coordonné par l'Institut Gustave Roussy a proposé une étude ancillaire permettant de récolter les plasmas de patients lors de leur rechute (ClinicalTrials.gov Identifier : NCT02613962).
- L'étude multicentrique non interventionnelle PNET 5 incluant des patients atteints de médulloblastome a permis de récolter de nombreux échantillons de LCR (ClinicalTrials.gov Identifier : NCT02066220).
- L'étude interventionnelle NGS Kids unicentrique à l'Institut Curie a été basée sur la détection d'altérations génétiques dans les fluides corporels chez des patients atteints de tumeurs pédiatriques de haut risque (ClinicalTrials.gov Identifier: NCT02546453).

Dans un premier temps, les quantités d'ADN circulant extraits de plasma et de LCR ont pu être déterminées.

3.1 Quantité d'ADN circulant selon le type de tumeur pédiatrique

Le plasma a été séparé du sang dans un tube EDTA standard par centrifugation à 2000 rcf pendant 10 minutes puis congelé à -80°C dans les 1 à 24 heures suivant la séparation.

Pour la majorité des cas, le plasma a été séparé dans les 2 à 4 heures suivant le prélèvement pour une qualité optimale de l'ADN circulant³³⁸⁻³⁴⁰. Au-delà de ce temps, les globules blancs présents dans le sang risquent d'être dégradés et de libérer leur ADN génomique.

3.1.1 Extraction de l'ADN circulant, quantification et qualification

L'ADN circulant a été extrait en utilisant le QIAamp Circulating Nucleic Acid Kit™ de Qiagen® avec le QiaVac24s system™ couplé à une pompe à vide. Les quantités de plasma étant variables (50 µL à 4 mL), les volumes les plus faibles ont été complétés avec du PBS (Phosphate-Buffered Saline) 1X pour obtenir des volumes similaires d'échantillons lors d'une même série d'extractions. Le protocole du constructeur a été suivi avec une élution dans 36 µL qui est le volume minimal pour recouvrir les membranes présentes sur les colonnes d'extractions.

A partir de 2018, une adaptation de protocole avec l'utilisation d'un nouveau kit QIAamp MinElute ccfDNA kit™ a été initiée.

Ce kit fonctionnant avec des billes magnétiques pour l'isolation de l'ADN circulant, seulement 1 mL de plasma a été extrait pour l'extraction avec ce kit pour une meilleure fixation sur support magnétique. Un meilleur rendement et des extractions plus rapides ont été obtenus avec ce kit.

L'ADN circulant a ensuite été dosé par Qubit™ avec le kit High Sensitivity (HS). En fonction de la taille des séries et de la concentration des échantillons, l'ADN circulant a été qualifié sur puces High Sensitivity DNA au bioanalyzer 2100 d'Agilent® (avec une sensibilité de 500 pg/μL d'échantillon, Figure 3.1.1) ou avec le kit HS Small Fragment du Fragment Analyzer d'Agilent® (avec une sensibilité de 50 pg/μL). Un pic majoritaire de 160 à 200 bp devait être observé et sa concentration a été retenue.

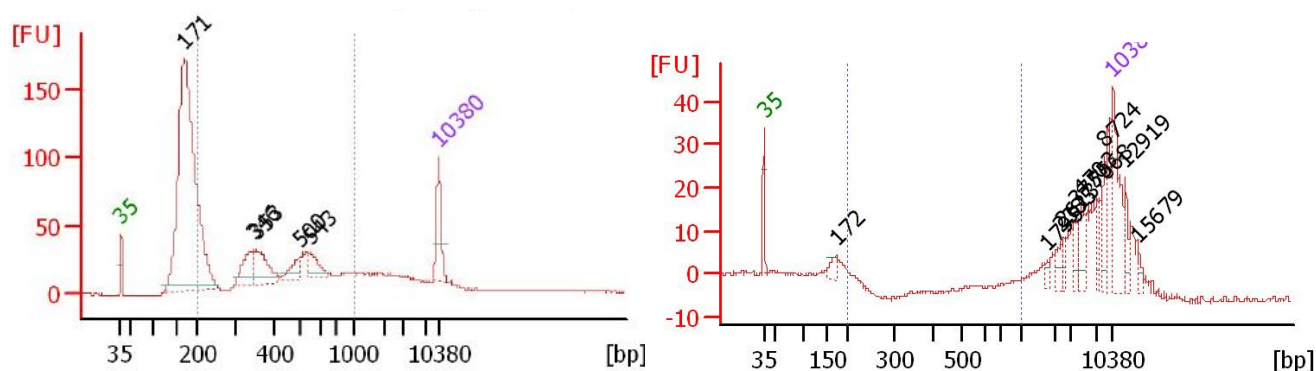


Figure 3.1.1 : Electrophorèse d'ADN circulant sur puces HS DNA au Bioanalyzer 2100. A gauche un profil propre avec un pic majoritaire à 170 pb correspondant à un nucléosome puis des multiples de ce pic pour des di- et tri- nucléomes. A droite, un échantillon contaminé avec de l'ADN en haut poids moléculaire provenant de l'ADNg de globules blancs dégradés.

3.1.2 Quantités différentes selon les types de tumeurs

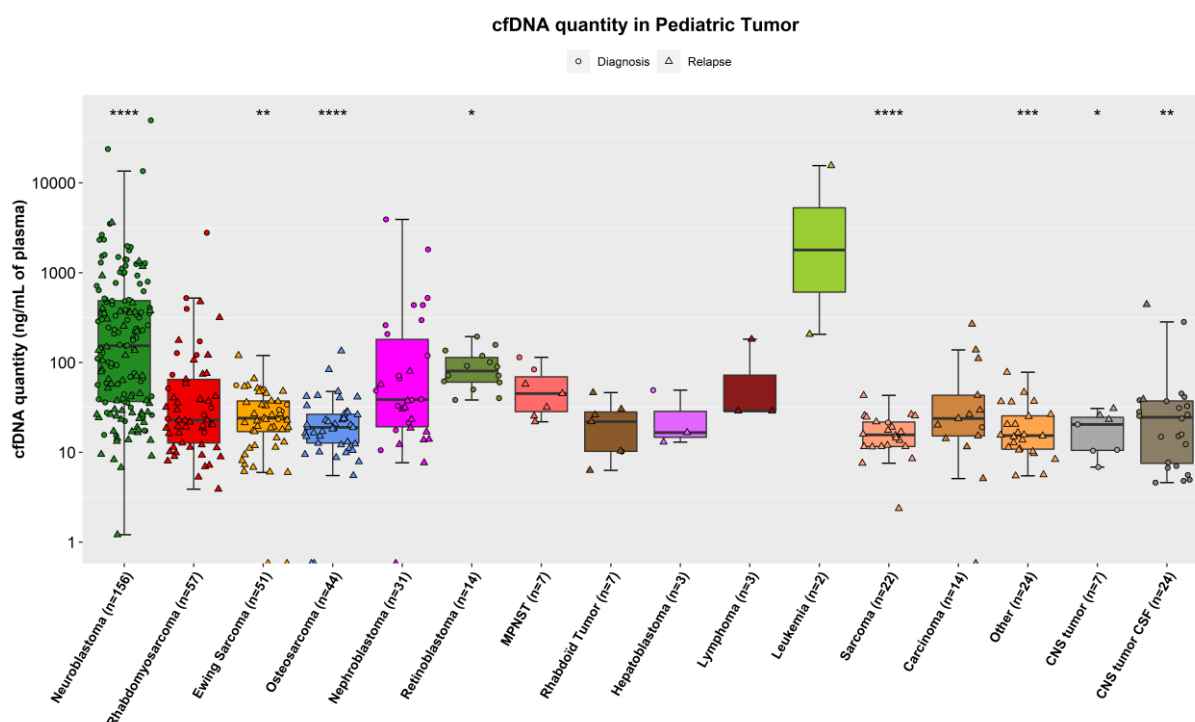


Figure 3.1.2 : Quantité d'ADN circulant par pathologie au diagnostic (○) ou la rechute (Δ) avec une échelle log10.

Près de 500 échantillons de plasmas et de LCR au moment du diagnostic (n=180) ou de la rechute (n=286) ont pu être extraits au sein de l'équipe RTOP. Pour être comparables, les quantités d'ADN circulant ont été extrapolées à 1mL de plasma (Figure 3.1.2).

Selon le type de tumeur, une quantité différente d'ADN circulant est observée (Anova p-value 3,21e-02) ainsi qu'une diminution des quantités d'ADN circulant très significative entre diagnostic et rechute (Wilcoxon p-value 2,56e-24). Cependant cette différence peut être considérée comme biaisée car seulement une trentaine de cas sont paires avec le diagnostic et la rechute d'un même patient et un nombre important de patients ont eu des rechutes peu volumineuses telles que des nodules ou micronodules pulmonaires n'émettant peut-être pas la même quantité d'ADN circulant qu'une rechute localisée sur le site primaire de la tumeur.

Les neuroblastomes sont connus pour relâcher des quantités parfois gargantuesques d'ADN circulant^{290,324,325}. Dans cette pathologie, les extractions contiennent des quantités très importantes d'ADN circulant (médiane 91,60 ng/mL de plasma, quantité de 1,21 à 49 700 ng/ml) en comparaison avec les autres tumeurs (Wilcoxon p-value 3,78e-14). Ces taux peuvent peut-être s'expliquer par une vascularisation importante des tumeurs avec des aspects souvent nécrotiques et atteintes métastatiques fréquentes.

Les ostéosarcomes semblent avoir des quantités mineurs d'ADN circulant (Wilcoxon p-value 3,65e-5 ; médiane 23,67 ng/ml, quantité de 0 à 134 ng/ml). Ces tumeurs touchant l'os sont très peu vascularisées sans forcément de contact avec les tissus mous pouvant expliquer une circulation moindre de son ADN circulant. Ces cancers touchent également plus souvent des adolescents qui possèdent naturellement moins d'ADN circulant que les enfants^{216,217}.

Même si aucune différence significative n'a été observée, comparés à l'ensemble des quantités d'ADN circulant des autres tumeurs, les néphroblastomes ainsi que les rétinoblastomes possèdent également de fortes quantités d'ADN circulant pouvant certainement s'expliquer par une taille importante de la tumeur pour la première et une vascularisation importante pour les deux.

Concernant les tumeurs cérébrales, aucune différence significative n'a été observée entre les quantités observées dans le plasma (médiane 18,31 ng/ml, quantité de 6,84 à 30,75 ng/mL) et le LCR (médiane 49,71 de 4,60 à 442,75 ng/mL). Cependant et contrairement au LCR, l'ADN circulant est toujours présent dans le sang chez un individu même sain.

A cette étape, la quantité d'ADN tumoral circulant, ctDNA dans chaque ADN circulant ne peut être établie. Cependant les fortes quantités d'ADN circulant pour certains types de tumeurs telles que le neuroblastome laissent suggérer un ratio important de contenu tumoral.

3.2 Séquençage d'exome à partir de l'ADN circulant

Pour mieux répondre à la complexité, sur le plan génétique, des tumeurs pédiatriques et la faible fréquence de marqueurs récurrents, un séquençage de l'exome complet est le plus adapté⁹. Malgré les fortes quantités d'ADN circulant relarguées par dans certains types de tumeurs tels que le neuroblastome, les quantités d'ADN circulant réellement utilisables restaient tout de même faibles (pour 100 µL de plasma, seul 10 ng avait été extrait). Les protocoles de WES existants ont dû être adaptés pour permettre l'étude d'une faible quantité de départ.

Parmi les trois principaux protocoles de capture d'exomes commercialisés que sont les kits Illumina®, Agilent® et Roche®, le kit Roche® semblait le plus adapté notamment avec la technologie Kapa™ permettant la construction de librairies avec une faible quantité d'ADN (en partant d'1 ng d'ADN). La technologie Illumina® basée sur tagmentation et donc la nécessité d'une fragmentation enzymatique a été très vite écartée, ne répondant pas au découpage de 160 pb de l'ADN circulant. Le kit Agilent® a également été éliminé car il nécessitait une quantité d'ADN de 200 ng pour un résultat interprétable.

3.2.1 Construction des librairies et capture d'exome

Suite à l'extraction de l'ADN circulant, les librairies ont été construites avec le Kapa Library Preparation Kit Illumina platforms™. L'ADN circulant n'avait pas besoin de fragmentation et pour augmenter le rendement de librairies, l'étape de ligation des adaptateurs

pour le séquençage Illumina et l'identification des échantillons furent modifiées avec un ratio de 10:1 et laissée sur la nuit (au lieu de 15 minutes). L'étape de sélection de taille fut conservée pour retirer les fragments de moins de 250 bp et de plus de 450 pb (ADN circulant avec adaptateurs : $160 + 130 = 290$ bp). Une PCR de 9 cycles, soit 2 cycles additionnels en comparaison avec le protocole standard, fut réalisée. Le reste du protocole fut suivi selon les recommandations du fabricant.

L'étape de sélection de taille permettait de seulement conserver le pic mononucléosome de l'ADN circulant retirant ainsi l'ADNg en cas de contamination mais également les pics de di-, tri-nucléosome, etc (Figure 3.2.1). Un test de séquençage par WES sans sélection de taille d'un échantillon d'ADN circulant ne montra aucune information supplémentaire dans les ADN circulants issus de multi-nucléosome et la capture de l'exome fut moins efficace.

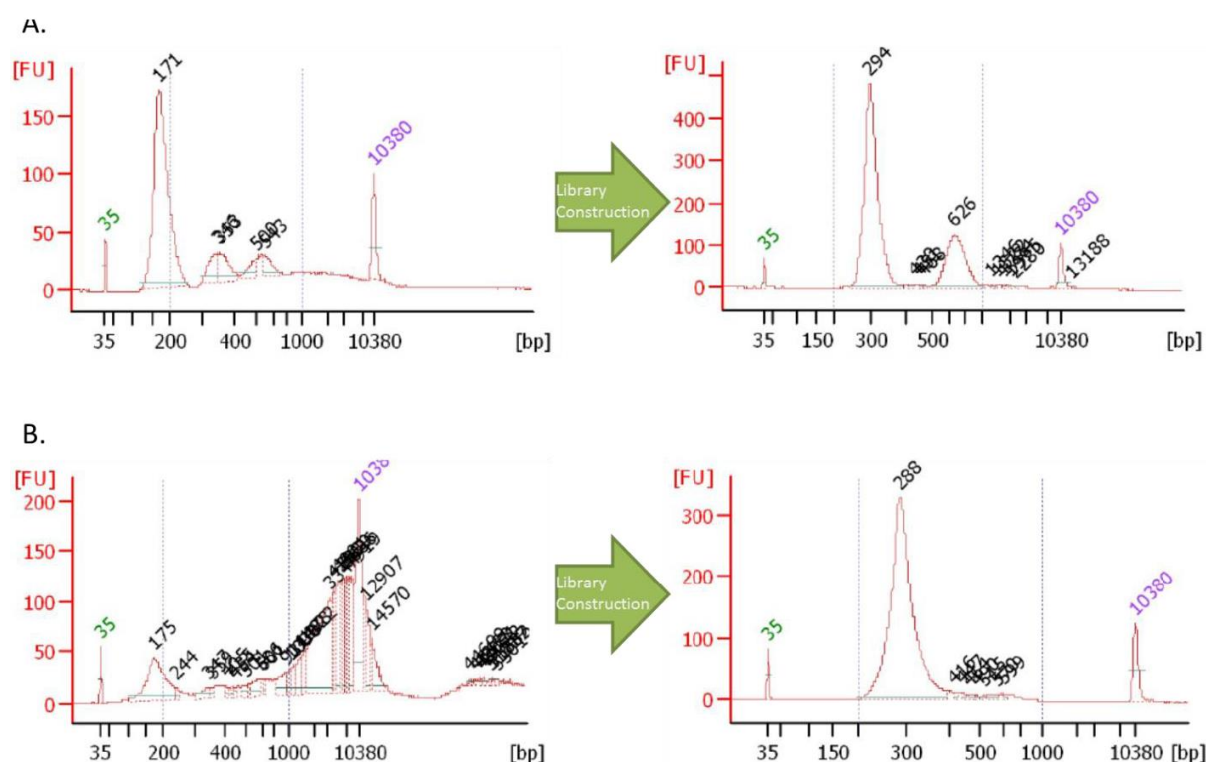


Figure 3.2.1 : Electrophorégramme du cfDNA extrait puis utilisé pour la librairie passée sur Bioanalyzer 2100 avec une puce HS DNA. Avec un plasma de bonne qualité contenant un pic majoritaire de 170bp pour le cfDNA (A) et un de mauvaise qualité avec une contamination d'ADNg en haut poids moléculaire (B). La construction des librairies permet de conserver seulement le pic de 170 pb avec ajout d'adaptateurs de 130 pb.

La capture d'exome a été réalisée avec SeqCap EZ Human Exome Kit v3.0 (Nimblegen Roche Sequencing™) selon le protocole du fabricant. Huit échantillons ont été multiplexés au sein d'une même capture puis séquencés sur 3 lignes de flowcell High-Output d'HiSeq2500™ en Paired-End 100 pour une couverture théorique de 100x. Ce protocole a également été utilisé sur des échantillons fixés en FFPE avec seulement 10 ng d'ADNg et avec une fragmentation de l'ADN par Covaris® ME220.

3.2.2 Pipeline bio-informatique

Après le séquençage, l'application de pipelines bio-informatique a pour objectif d'identifier des altérations génétiques (SNV, altération de nombre de copies) dans les différents échantillons. Pour cela, après alignement par Bowtie2 permettant les mismatch à 4%, les fichiers BAM ont été nettoyés selon les recommandations Genome Analysis Toolkit en enlevant les duplicats^{341,342}.

La recherche de variants de nucléotide unique (SNV : Single Nucleotide Variant) dans les WES a été réalisée avec 3 outils : GenomeAnalysisTK-3.5 UnifiedGenotyper, HaplotypeCaller, and Samtools-0.1.18³⁴³. L'outil Annovar-v2013-07-29 avec les bases de données cosmic-v64, dbsnp-v137 et RefSeq a été utilisé pour l'annotation et la prédiction fonctionnelle a été réalisée avec Polyphen2, LRT, MutationTaster et Sift³⁴⁴⁻³⁴⁸. Les SNVs avec une qualité <30, une couverture <20 dans la tumeur ou le plasma et avec moins de 2 reads supportés par le variant ont été éliminés. Seuls les SNVs non synonymes ou synonymes avec un identifiant COSMIC dans les exons des gènes codants ou dans un site d'épissage ont été gardés. Les variants reportés avec une fréquence de plus de 1% dans les bases de données 1000 genomes (1000gAprl_2012) ou Exome Sequencing Project (ESP6500) ont été considérés comme des polymorphismes et évincés.

En parallèle, un séquençage par WES sur l'ADN génomique de la tumeur primaire et sur l'ADN constitutionnel a également été réalisé. Ainsi, l'ADN constitutionnel du patient a été utilisé comme filtre avec une couverture supérieure à 10 reads et un variant présent dans moins de 2 reads pour que l'altération soit considérée comme somatique. Les altérations étaient ensuite validées ou rejetées une à une après visualisation par Integrative Genomics Viewer (IGV)³⁴⁹.

Les altérations de nombre de copies ou ACN, ont été analysées avec VarScan.v2.3.5 et DNACopy-1.42.0.^{350,351}. Une estimation du contenu tumoral dans la tumeur primaire ou l'ADN circulant a été réalisée avec Sequenza³⁵².

3.3 Etude de l'évolution clonale du neuroblastome par l'ADN circulant

Pour étudier l'évolution clonale, l'ADN circulant de patients atteints de neuroblastome a été séquençé par WES (Publication en annexe)³⁵³. Dix-neuf patients atteints de neuroblastome avec des plasmas collectés pendant le suivi ont été sélectionnés. Pour ces patients, un séquençage de la tumeur et de l'ADN constitutionnel avait été réalisé par WES (n=9) ou par WGS (n=4) (Tableau 1).

3.3.1 Comparaison entre la tumeur primaire et l'ADN circulant au diagnostic

Une première étape de l'analyse a porté sur une détermination du contenu du matériel tumoral dans chaque échantillon. Pour cela, l'analyse du contenu tumoral avec l'outil Sequenza a révélé une moyenne de 73% de cellule tumorale dans la tumeur (de 15% à 98%) et de 60% de ctDNA dans l'ADN circulant (de 3% à 99%). Dans chaque cas, des SNVs ont

été détectés dans le plasma au diagnostic et dans un seul cas, les altérations en nombre de copies n'ont pas été observées en raison d'un contenu tumoral trop faible.

Dans un premier temps, les altérations en nombre de copie ont été analysées. Sur un total de 162 altérations en nombre de copie détectées, 151 altérations sont communes à la tumeur et au plasma. Trois sont spécifiques à la tumeur et 8 au plasma (Figure 3.3.1). Pour 11 cas pour lesquels l'ADN circulant avait aussi pu être analysé avec les puces Oncoscan™, les mêmes altérations ont été observées³²⁴. Une amplification du gène IGF1R a ainsi été confirmée par ces deux techniques mais également par qPCR (Figure 3.3.2 B).

Ensuite, les SNVs ont été analysés et une moyenne de 19 SNVs (de 9 à 69) communs a été observée entre la tumeur et l'ADN circulant incluant des gènes connus pour être impliqués dans le neuroblastome tels que *ALK* ou d'autres gènes de la voie MAPK mais également des oncogènes tels que *ARHGAP* pour 3 cas, *NOTCH* pour 1 cas ou *MLL* pour 4 cas. En moyenne, 6 SNVs ont été observés uniquement dans la tumeur et 22 uniquement dans l'ADN circulant (Figure 3.3.1).

Case Number	INSS Stage	Diagnosis		2nd Timepoint		Last News		Number of Plasmas	Primary Tumor Genetic Alteration		Pathological Diagnosis		Plasma volume (μL)		Plasma : Quantity of DNA per mL(ng)		Quality of CtDNA (200bp peak ratio)	
		Age (M)	delay diag- 2nd Timepoint (M)	Status	delay diag- last news (M)	Status	MYCN status		Genomic Copy Number Profile	Histology	% tumor cell ¹	Diag	Relapse/2nd Timepoint	Diag	Relapse/2nd Timepoint	Diag	Relapse/2nd Timepoint	
1	4	20	7	SD	14	Dead	3	A	Numerical		NOS NB	95	300	300	13533	71	96%	97%
2*	3	4	53	CR	47	Alive	6	A	Segmental		NOS NB	60	200	400	406	104	>10%	63%
3	4	25	7	CR	50	Alive	7	A	Segmental		poorly differentiated NB	90	400	400	2170	193	81%	95%
4	3	19	6	CR	51	Alive	6	A	Segmental		poorly differentiated NB	90	300	500	76	262	95%	95%
5	4	47	6	CR	40	Alive	4	NA	Segmental		poorly differentiated NB	90	400	100	1295	1267	70%	30%
6	4	101	1	PR	7	Dead	4	A	Segmental		indifferentiated NB	70	300	300	794	371	94%	73%
7	4s	3	7	PD	16	Dead	4	A	Segmental		poorly differentiated NB	98	300	300	1248	3617	21%	15%
8	4	19	12	CR	27	Alive	6	A	Segmental		poorly differentiated NB	MD	500	300	1113	232	95%	56%
9	3	20	4	CR	36	Alive	6	A	Segmental		poorly differentiated NB	50	400	300	2328	73	87%	95%
10	4	32	27	relapse	31	Dead	2	NA	Segmental		poorly differentiated NB	50	800	1750	387	368	97%	>10%
11	4	11	23	relapse	31	Dead	2	NA	Segmental		poorly differentiated NB	90	1500	1000	639	105	35%	70%
12	1	58	8	relapse	69	Alive	2	NA	Segmental		ganglioneuroblastoma	80	200	1500	210	137	>10%	73%
13	4	31	22	relapse	24	Dead	2	NA	Segmental		poorly differentiated NB	90	800	200	1278	364	77%	83%
14	4	14	9	relapse	11	Dead	1	A	Segmental		poorly differentiated NB	60	700	/	930	/	84%	/
15	4	30	17	relapse	18	Dead	2	NA	Segmental		NOS NB	60	200	1500	1530	376	89%	95%
16	4	66	26	relapse	49	Alive	2	NA	Segmental		poorly differentiated NB	MD	800	200	39	65	76%	93%
17	2	30	14	relapse	21	Dead	2	NA	Segmental		poorly differentiated NB	MD	2000	3300	26	30	86%	83%
18	2	9	/	/	12	Dead	1	NA	Segmental		poorly differentiated NB	90	200	/	350	/	72%	/
19	4	18	14	relapse	21	Dead	2	A	Segmental		poorly differentiated NB	MD	2200	500	42	1210	78%	>10%

Tableau 1 : Résumé des caractéristiques des patients.

Abbreviations: SD, stable disease; M, months; NC, noncontributive.

*Cas diagnostiqué stade IV avec seulement un sérum disponible. L'étude de ce cas a été basée sur sa rechute localisée et son suivi avec des plasma.

3.3.2 Analyse des plasmas à la rechute ou lors de rémission

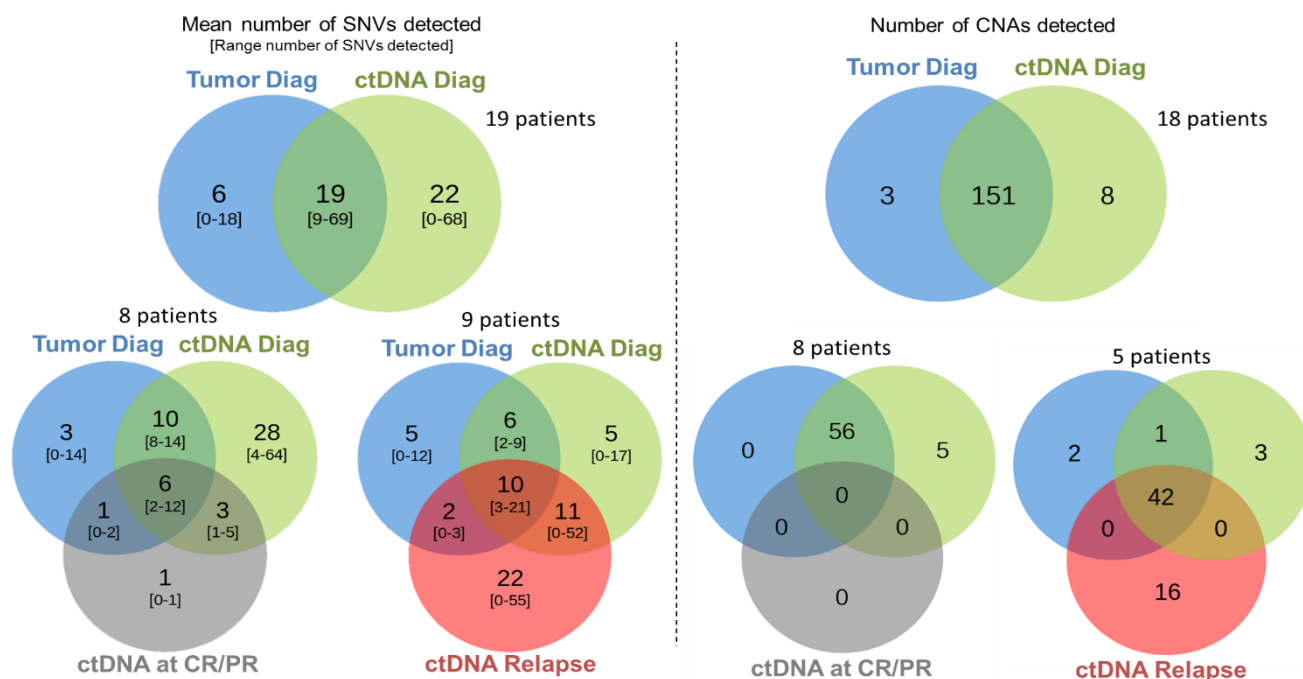


Figure 3.3.1 :A gauche, Diagramme de Venn des moyennes de mutations détectées au diagnostic en WES entre dans la tumeur et le plasma puis pour les patients en rémission partielle (PR) ou complète (CR) en gris ou avec une rechute en rouge. A droite, diagramme de Venn pour les altérations en nombre de copies.

Un séquençage par WES a également été réalisé sur un plasma recueilli dans un second temps dans le traitement du patient. Pour 8 patients, cela correspondait à une rémission partielle ou complète (PR/CR) tandis que pour 9 autres, une rechute avait été confirmée (PD).

Pour les patients en PR/CR, une moyenne de 6 SNVs (de 2 à 12) a été observée entre la tumeur primaire, l'ADN circulant au diagnostic et l'ADN circulant après le traitement (Figure 3.3.1). La majorité des mutations détectées au moment du diagnostic disparaît après le traitement et notamment les mutations *ALK* (Figure 3.3.2 D). Pour 3 cas, un nouveau SNV a émergé détecté par ce second WES (mutations dans les gènes *ZNF814*, *ILDR2*, *RREB1*).

Pour les patients avec une rechute, les fractions tumorales dans le plasma étaient de 67% en moyenne (de 43% à 89%) permettant de caractériser des altérations en nombre de copies pour 5 cas. Pour ces patients en PD, 10 SNVs en moyenne (de 3 à 21) étaient communs entre le plasma au diagnostic et le plasma à la rechute tels qu'une mutation dans *ALK*. Une

moyenne de 22 nouveaux SNVs (de 0 à 55) est apparue à la rechute dans le plasma avec des mutations dans des gènes décrits comme jouant un rôle dans le cancer comme *KRAS*, *ADCY7* ou *NCOR1* (Figure 3.3.2 E). Une analyse de voies de signalisations ciblées par ces altérations a permis d'observer un enrichissement dans les voies MAPK précédemment décrites dans le neuroblastome et également dans la voie de signalisation de la protéine A (Ingenuity Pathways Analysis (IPA))⁶.

Pour les 5 cas avec un profil d'altérations en nombre de copies interprétables, 16 nouvelles altérations ont été observées dans le plasma à la rechute en comparaison à celui au diagnostic (Figure 3.3.1) incluant une amplification du gène *CDK4* (Figure 3.3.2 C) ou du gène *CDK6* associée à une délétion de *BCL11A*.

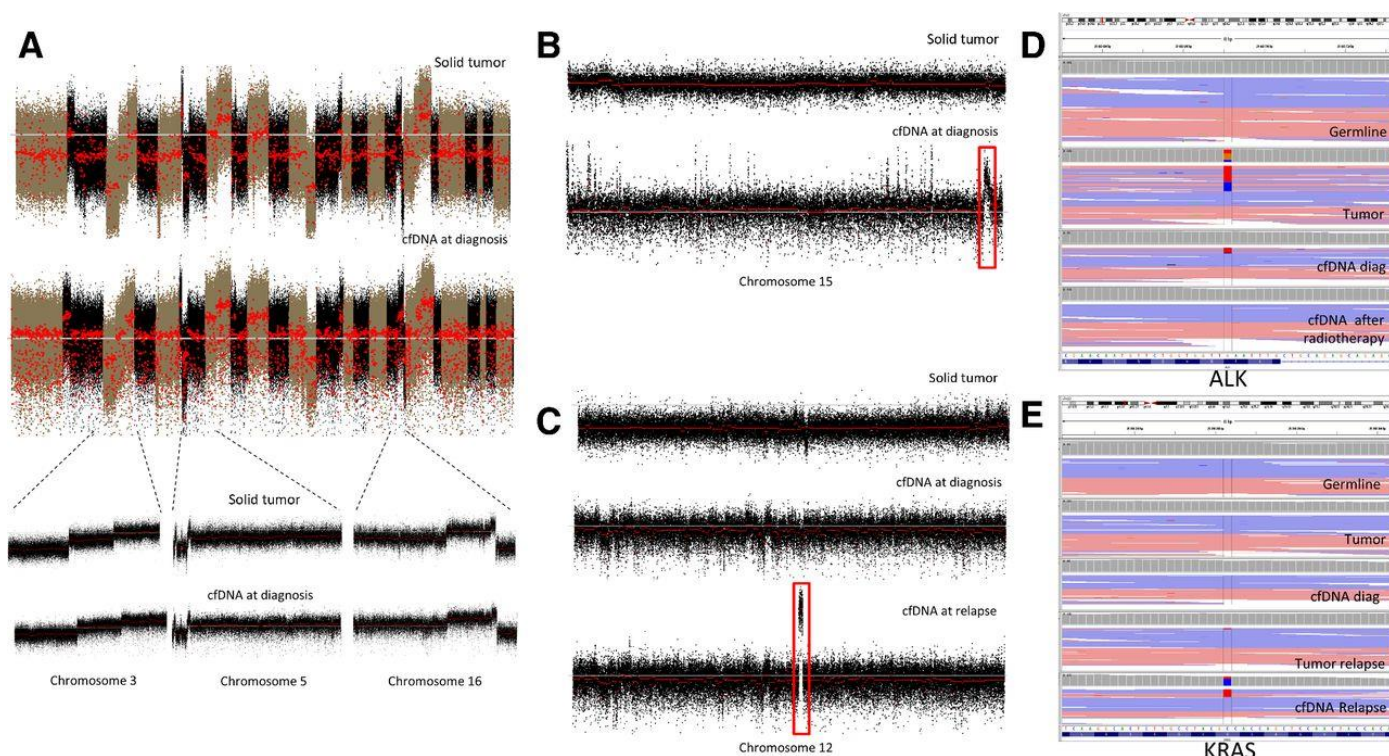


Figure 3.3.2 : Altérations génétiques observées entre la tumeur primaire et le cfDNA (A). Gain observé du gène *IGF1R* sur le chromosome 15 dans le cfDNA (B). Amplification du gène *CDK4* détecté dans le cfDNA à la rechute, l'ADN de la tumeur à la rechute n'était pas disponible(C). Double mutation dans le gène *ALK* présente au diagnostic dans la tumeur (30%-17%) et le cfDNA (17%-2%) et absente du cfDNA après le traitement (G>T-G>C, p.F1174L, chr2: 29443695) (D). Mutation *KRAS* absente au diagnostic dans la tumeur et le cfDNA et à la rechute à un niveau sous-clonal dans la tumeur (3%) et à un niveau clonal dans le cfDNA (28%) (C>T p.G13D, chr12: 2539828)(E).

Ces résultats en WES montrent une vue plus globale de la maladie dans sa spatialité mais également dans sa temporalité.

3.3.3 Séquençage ciblé sur l'ADN circulant

Au total, l'analyse des tumeurs et des ADN circulants au diagnostic par WES de patients atteints de neuroblastome a permis de détecter 1211 SNVs tumeurs (et ADN circulant) spécifiques. Pour permettre de modéliser l'évolution clonale, un panel contenant l'ensemble de ces SNVs a été dessiné. Puis, l'ADN circulant issu de l'ensemble des plasmas recueillis à différents temps du traitement a été séquencé à une profondeur de 1000x (Figure 3.3.3).

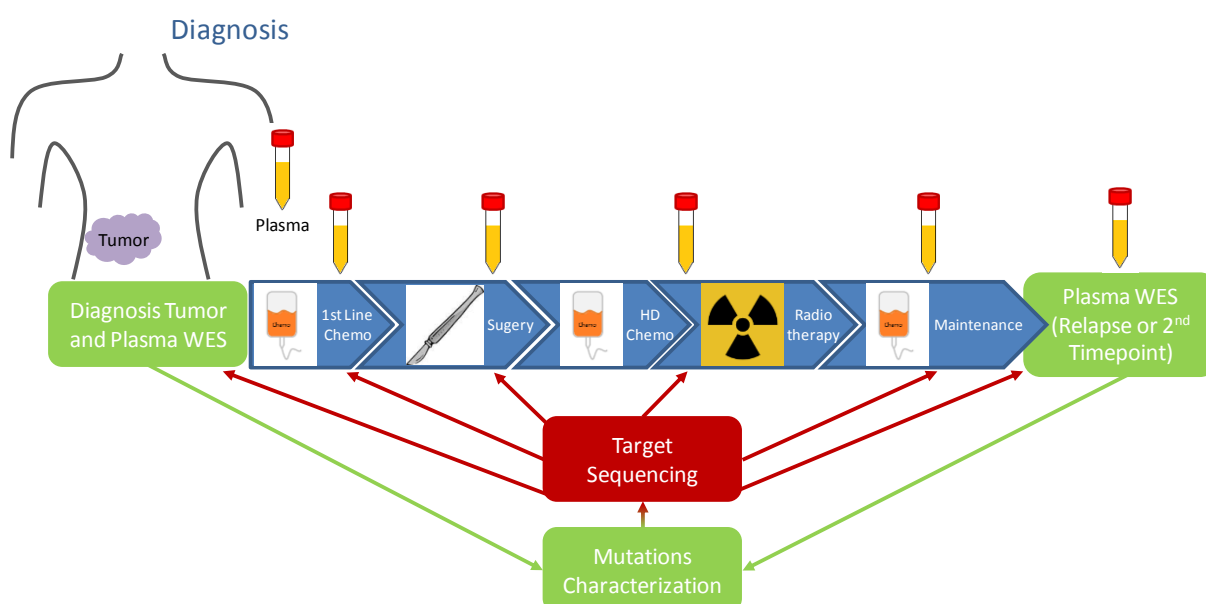


Figure 3.3.3 : Plan d'analyse de l'évolution clonale dans le neuroblastome.

3.3.3.1 Design et mise au point du séquençage ciblé

Le panel a été dessiné avec des régions de 100 pb et en leurs centres les SNVs identifiés en WES puis soumis au logiciel Nimbledesign™ de Roche®. Le panel couvrait théoriquement 99,4% de régions contenant les mutations pour une taille de 110kb. Le maintien de la technologie Roche® a permis d'utiliser le même protocole pour la construction des bibliothèques. Cependant la taille réduite du panel a nécessité une double capture contrairement à l'exome pour augmenter l'enrichissement de nos cibles³⁵⁴. Après une

incubation de 16 h et une purification, une PCR de 5 cycles a été appliquée aux échantillons post capture et une nouvelle capture a été réalisée sur la nuit. Après purification, une PCR de 14 cycles a été réalisée.

La double capture a été faite sur des pools de 12 échantillons puis séquencée sur MiSeq™ avec des kits V3 pour une couverture espérée de 1000x.

Aucune différence significative (T test pvalue = 0,897) et une très forte corrélation ($R^2 = 0,899$; Figure 3.1) ont été observées entre les fréquences alléliques du WES et celles du séquençage ciblé indiquant une concordance entre les deux techniques. Une sensibilité de 1 % pour le séquençage ciblé contre 5 % pour le WES a pu être établie avec une série de dilutions d'un ADN circulant à forte fraction tumorale (Figure 3.3.4).

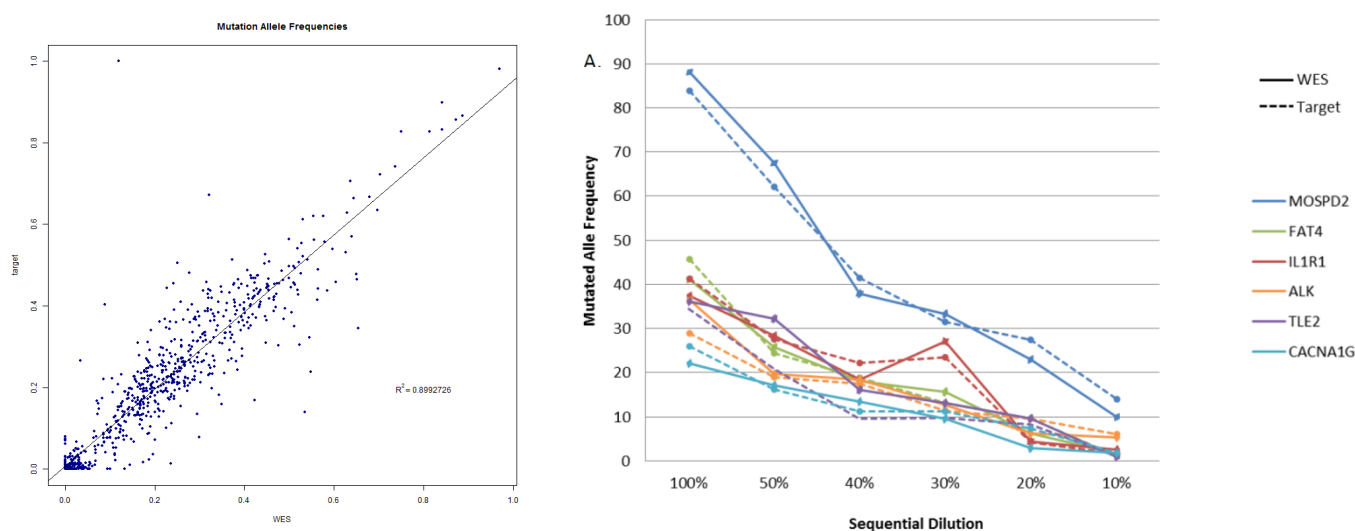


Figure 3.3.4 : Sensibilité et spécificité du TargetSequencing. Comparaison entre fraction allélique des SNVs dans le WES et le TargetSeq (A gauche). Une forte corrélation est observée entre les deux techniques ($R^2=0.899$). Dilutions en série du cfDNA d'un patient (avec un fort contenu de ctDNA dans le plasma > 90%) (A droite) avec l'ajout de cfDNA d'un donneur sain à 50%, 40%, 30%, 20% et 10%. Six mutations détectées par le WES dans la tumeur primaire et dans le cfDNA et suivies : *MOSPD2* (88,68% ; C>A ; p.Q487K ; chrX:14937833), *FAT4* (43,10% ; C>A ; p.P248Q ; chr4:126238309), *IL1R1* (32% ; G>T ; p.E276X ; chr2:102788365), *ALK* (29,61% ; C>A ; p.L1196M ; chr2:29443631), *TLE2* (42% ; C>A ; p.R385S ; chr19:3005948), *CACNA1G* (30.26% C>A ; p.R2195L ; chr17:48703943).

3.3.3.2 Suivi des altérations dans le plasma

Le séquençage haute profondeur par panel ciblé a permis d'analyser le plasma obtenu au diagnostic et celui après le traitement de tous les patients mais également des points

intermédiaires de plasma pour 9 patients avec une couverture de séquençage plus élevée pour la recherche d'évènements sous-clonaux.

Pour 9 cas, les fréquences alléliques de leurs SNVs ont pu être suivies pendant le traitement avec 2 à 6 prélèvements de plasma intermédiaires. Pour chaque cas, la fréquence allélique était en corrélation avec le statut clinique. Par exemple, pour un patient avec une maladie en rémission, une diminution de la fréquence allélique des mutations détectées au diagnostic était observée et pour un autre, une diminution lors de la première chimiothérapie était observée puis une augmentation suite à une progression de la maladie (Figure 3.3.5).

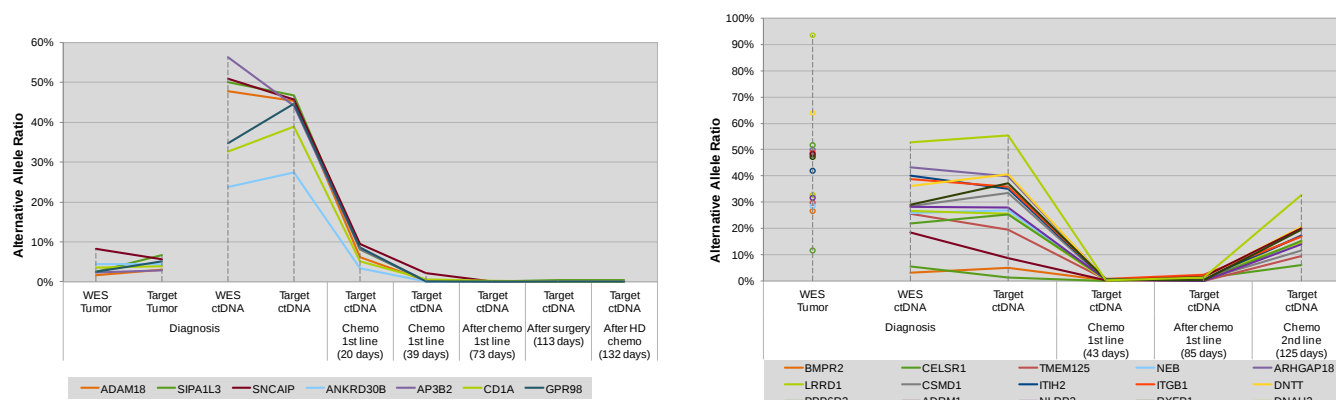


Figure 3.3.5 : Evolution de la fréquence allélique des mutations dans les échantillons séquentiels de cfDNA. Les fréquences obtenues grâce au WES sont comparées avec celles obtenues en Target dans le plasma et la tumeur au diagnostic (si cette dernière était disponible) puis le TargetSequencing est utilisé pour suivre ces mutations dans le temps dans les plasmas successifs. Le premier cas montre une disparition des mutations suite à la chimiothérapie (à gauche). Le second montre une résurgence des mutations corrélée à une progression métastatique de la maladie mais avec des fréquences différentes comparées à celles du diagnostic (à droite).

Pour 17 SNVs détectés par WES au moment de la rechute et répartis dans le plasma de 5 patients, la forte couverture du panel a permis de détecter également ces altérations à un niveau sous clonal au diagnostic (moyenne de fréquence allélique de 0,84% allant de 0,33 à 2,91%) suggérant que ces altérations pouvaient être contenues dans des clones mineurs au diagnostic ayant émergé pour devenir majeurs à la rechute (Figure 3.3.6). Parmi ces gènes altérés, 10 étaient impliqués dans la neuritogénèse et 5 dans le cycle cellulaire (IPA network analysis).

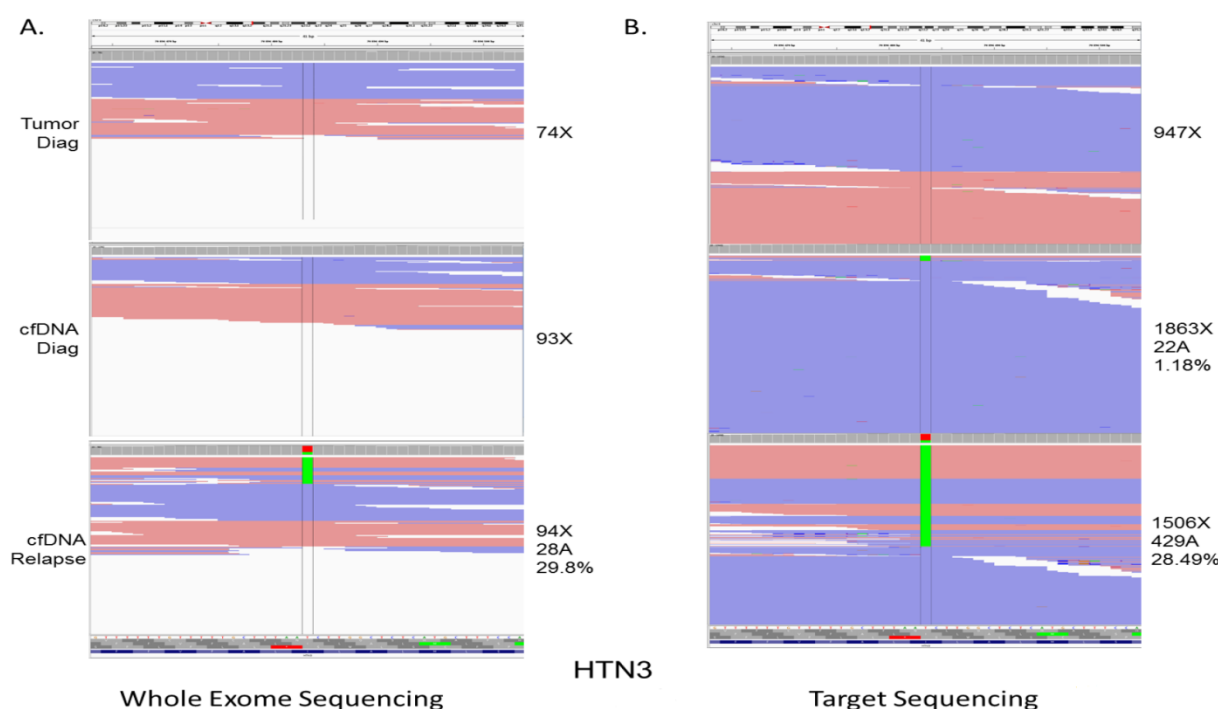


Figure 3.3.6 : Emergence d'une mutation observée par TargetSequencing sur IGV. Une mutation dans le gène *HTN3* a été détectée seulement à la rechute dans le plasma par WES. Cette mutation est confirmée avec la même fréquence allélique par le séquençage ciblé mais la couverture plus forte permet de détecter cette mutation à un niveau sous clonal dans le cfDNA au diagnostic.

3.3.3.3 Modélisation de l'évolution clonale

Le suivi des altérations dans les plasmas a pu être utilisé pour inférer à des clones en fonction des SNVs, de leurs fréquences et leurs dynamiques et ainsi permettre de modéliser une évolution clonale avec l'outil Fishplot^{355,356}. La modélisation a ainsi été possible pour 2 patients.

Par exemple, un patient a présenté une progression métastatique puis une maladie stable et par la suite sur le plan génétique, un clone persistant et résistant avec une concomitance de la disparition des autres clones a été identifié. Ce clone est caractérisé par une mutation dans le gène *HERC2* codant une protéine ligase de l'ubiquitine (Figure 3.4.5 A). Un second cas présentait 2 clones mineurs dans la tumeur au moment du diagnostic avec ensuite une augmentation de la fréquence allélique des mutations dans les gènes *ADRM1*, *BMPR2* et *CELSR1* lors du traitement (de 10% à 20% et de 1% à 5%) (Figure 3.4.5 B).

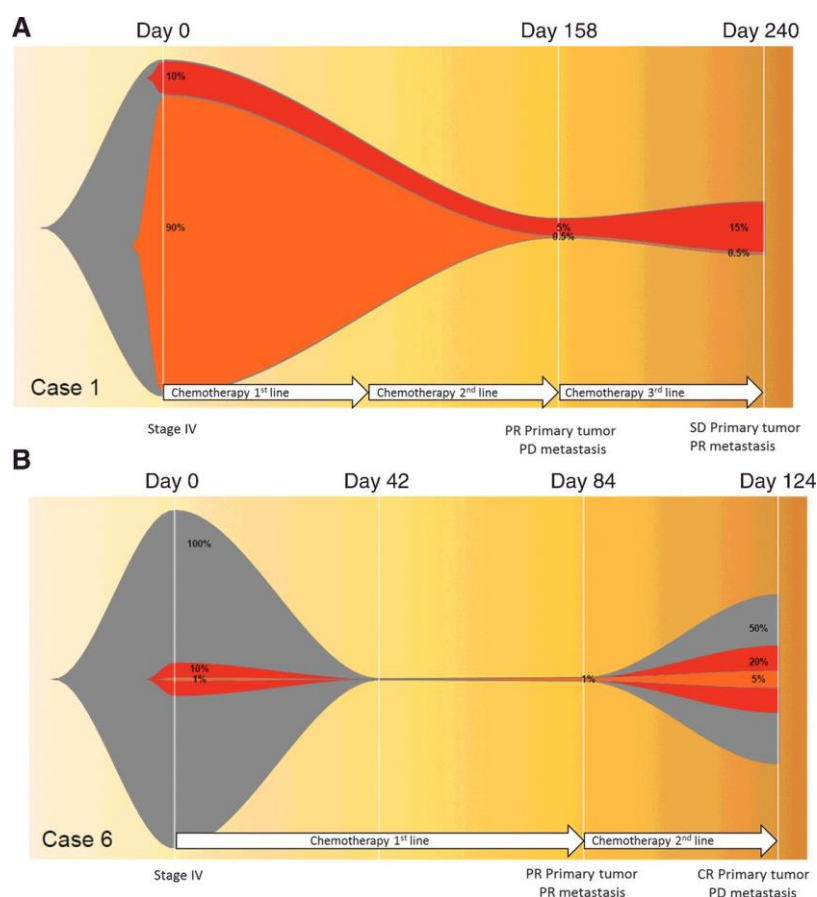


Figure 3.4.5 : Hypothèse d'évolution clonale basée sur la fréquence allélique des mutations détectées dans des plasmas séquentiels. Le clone ancestral est représenté en gris.

Pour le premier cas, un sous clone au diagnostic (en rouge) devient résistant à la chimiothérapie et prend la place du clone majoritaire (en orange) entraînant la rechute (A).

Pour le second cas, deux sous clones dérivés du clone ancestral sont présents au diagnostic (en rouge et en orange). Après la chimiothérapie, ces sous-clones semblent plus résistants à la chimiothérapie et leur collaboration avec le clone primaire pourrait être responsable de la rechute(B).

Cette étude a été publiée dans Clinical Cancer Research : DOI: 10.1158/1078-0432.CCR-17-1586³⁵³. Article en annexe

3.4 Evolution clonale dans les cancers pédiatriques

Suite à ces résultats pour des patients atteints de neuroblastome, une étude chez des patients atteints de tumeurs de hauts risques a été proposée (n=31). Cette étude nommée NGS Kids, était une étude prospective interventionnelle non thérapeutique institutionnelle (ClinicalTrials.gov Identifier: NCT02546453). Pour élargir l'étude, des échantillons provenant d'une étude rétrospective sur des patients atteints de neuroblastome ont également été ajoutés (n=20).

Type de Cancer	Nombre de patients	Age au diagnostic(mois)	Nombre de plasmas	Durée du suivi (mois)
Neuroblastome	12+20	6 - 125	1 - 9	0 - 30
Rhabdomyosarcome	7	24 - 172	2 - 7	2 - 26
Sarcome d'Ewing	3	44 - 132	4 - 6	3 - 20
MPNST	2	84 - 114	2 - 3	0,2 - 1
Gliome	2	80 - 218	2 - 3	0,4 - 2
Tumeur du foie	2	19 - 49	1 - 7	0 - 9
Tumeur cérébrale NOS	1	38	1	0
ATRT	1	47	1	0

Tableau 2 : Récapitulatif des types de tumeurs donc le plasma a été étudié par WES. Pour les neuroblastomes, 12 cas sont issus de l'étude prospective NGS Kids et 20 cas proviennent d'une étude rétrospective. MPNST : Malign Peripheral Nerve Sheath Tumors, NOS : Not Otherwise Specified, ATRT : Atypical Teratoid Rhabdoid Tumor.

3.4.1 Utilisation du WGS low coverage pour déterminer le contenu tumoral dans l'ADN libre

Sur chaque plasma au diagnostic, un WGS low coverage (0,1-2x) a été réalisé avec les librairies précaptures pour connaître les quantités de ctDNA dans le plasma. Les WES permettent également de connaître le contenu tumoral mais avec une sensibilité moindre >20% pour le WES contre >5% pour le WGS low coverage.

Pour chaque échantillon de plasma au diagnostic, les prédictions données par des outils WES ont pu ainsi être comparées tels que Facets ou Sequenza avec ceux du WGS low coverage comme ichorCNA (Figure 3.4.1)^{310,352,357}.

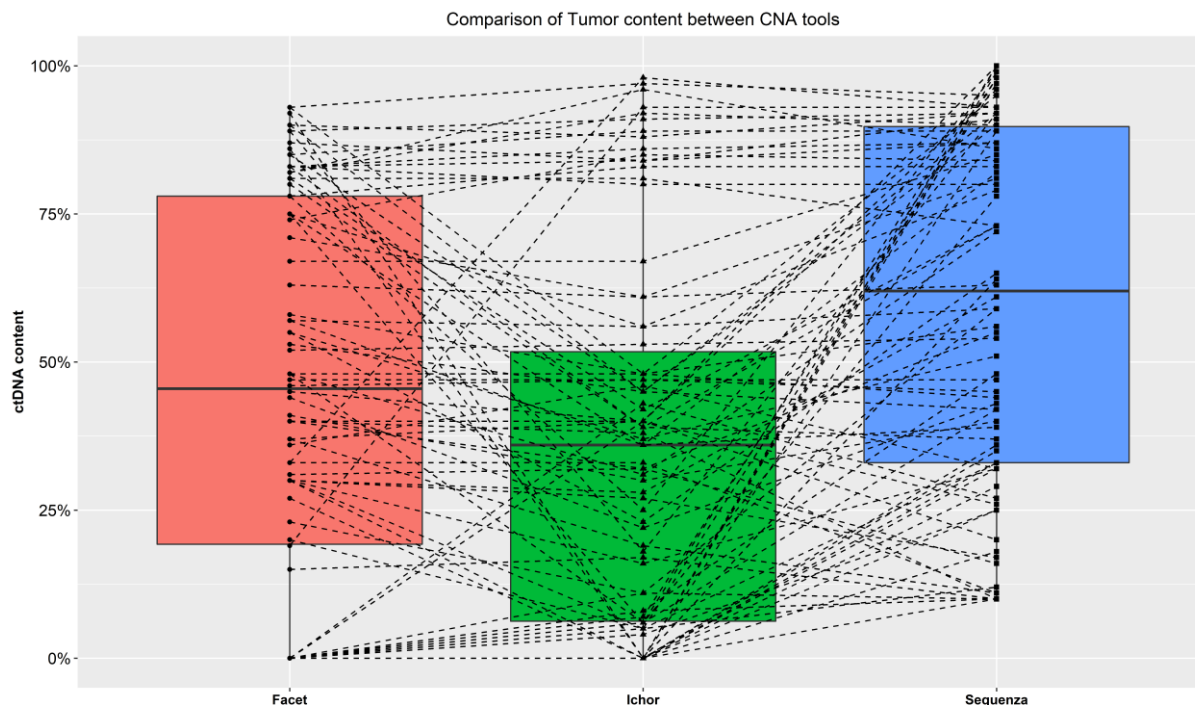


Figure 3.4.1 : Boxplot de la prédiction du contenu ctDNA pour les WES avec Facets et Sequenza et le WGS Low Coverage avec ichorCNA. Chaque cas a été calculé par les différents outils et relié par une ligne en pointillés.

Des disparités sont observées entre les outils, ichorCNA est plus sensible et prédit des contenus tumoraux jusqu'à 4 %. Sequenza semble également plus sensible que Facets pour certains cas avec 12 % de contenu tumoral détecté par Sequenza contre 19 % avec Facets mais Sequenza demande une relecture car son algorithme cherche à prédire un clone majoritaire qui peut être l'ADN sain (100 % pour Sequenza contre 0 % pour les autres outils). Même si Facets est légèrement moins sensible, l'outil ne demande pas une vérification systématique et possède de nombreux avantages comme un temps CPU très réduit comparé à Sequenza mais également la possibilité de fonctionner sans le germline du patient.

Pour le reste de l'étude, ichor a été privilégié et Facets a été utilisé comme alternative en cas d'absence de WGS low coverage. L'association d'ichor avec le WGS low coverage et

Facets avec le WES a également permis une meilleure caractérisation des positions des altérations en nombre de copies.

3.4.2 Comparaison des altérations génétiques détectées par WES entre la tumeur et le plasma

3.4.2.1 Neuroblastome

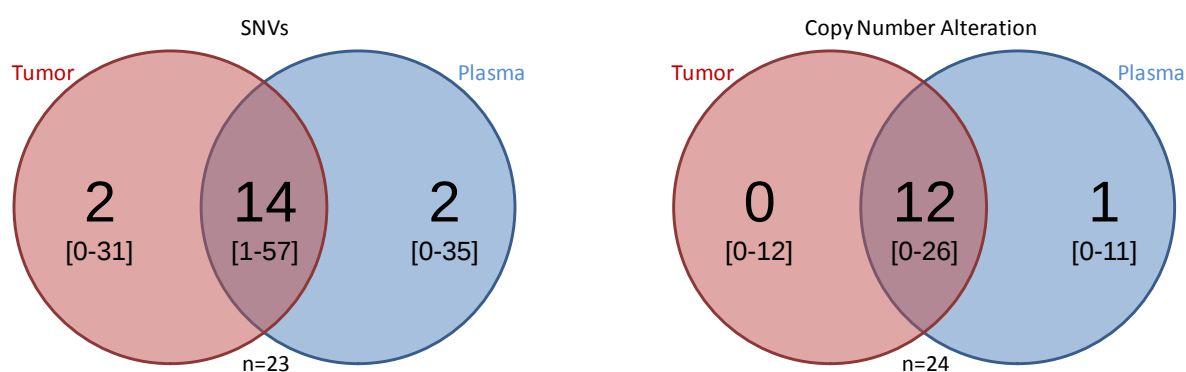


Figure 3.4.2: Diagramme de Venn des SNVs détectés dans les neuroblastomes par WES au diagnostic dans la tumeur et le plasma (à gauche) et pour les altérations en nombre de copies par WGS Low coverage (à droite).

Le contenu tumoral dans le plasma des patients atteints de neuroblastome était assez élevé avec une médiane de 41% calculée par ichorCNA (et Facets pour 2 patients sans WGS low coverage). La proportion du ctDNA dans l'ADN circulant allait de 0 % à 98 %. Pour deux échantillons, le contenu en ctDNA a été estimé à 0%, l'un provenait d'un patient atteint d'un neuroblastome de stade 2 soit localisé et de bas risques et pour l'autre cas, l'échantillon était issu d'un prélèvement de plasma réalisé 10 jours après le début du traitement.

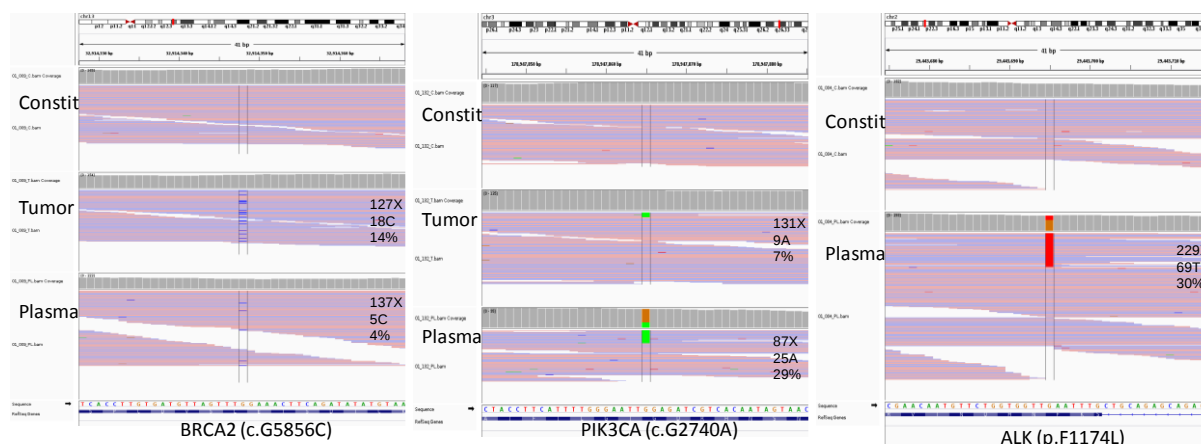


Figure 3.4.3 : SNVs communs au diagnostic entre la tumeur et le plasma tels *BRCA2* (cas 36) et *PIK3CA* (cas 44) et détectés dans le plasma en absence de tumeur tels que *ALK* (cas 33).

Pour 7 patients, le matériel tumoral n'était plus disponible au diagnostic et les résultats de séquençage du plasma au diagnostic n'ont donc pas pu être comparés à ceux de la tumeur. Pour 2 patients, l'ADN constitutionnel des patients n'était pas disponible et seules les altérations en nombre de copies ont été analysées et comparées à celles de la tumeur pour un cas.

Pour les 23 cas avec un trio de WES sur l'ADN constitutionnel, l'ADN tumorale et l'ADN circulant, une médiane de 14 SNVs allant de 1 à 57 a été observée dans la tumeur et le plasma (Figure 3.4.2), SNVs incluant notamment des oncogènes tels que *BRCA2* ou *PIK3CA* (Figure 3.4.3). Une médiane de 2 SNVs allant de 0 à 31 a été détectée spécifiquement dans la tumeur, SNVs tels que *KDM4D* et une médiane de 2 SNVs allant de 0 à 35 a été détectée dans le plasma dont des SNVs dans les gènes *CDH9* ou *TET2*.

Pour les 7 cas sans WES sur la tumeur, une médiane de 29 SNVs a été observée allant de 10 à 54 SNVs avec des variants dans le gène *ALK* (F1174L) pour de 2 cas (28, 33) ou le gène *TP73* ou *STAG2* (Figure 3.4.3).

Pour 24 cas, l'analyse des altérations en nombre de copies était possible, ne nécessitant pas l'ADN constitutionnel du patient. Une médiane de 12 altérations détectées allant de 0 à 26

altérations ont été observées dans la tumeur et le plasma, altérations telles que des amplifications récurrentes dans le neuroblastome du gène *MYCN* dans 11 cas, *ALK*, *TERT*, tous les deux dans 2 cas ou *ITGA1* observé dans d'autres types de cancers comme le cancer du pancréas³⁵⁸.

Le cas 21 est un bon reflet de l'hétérogénéité tumorale. L'ADN tumorale et l'ADN circulant possèdent de nombreuses altérations communes, 57 SNVs tels qu'une mutation dans le gène *TERT* (Figure 3.4.4) et 24 altérations en nombre de copies telles qu'une perte du 1p, un gain du 17 et également un chromothripsis sur le chromosome 15 (Figure 3.4.5).

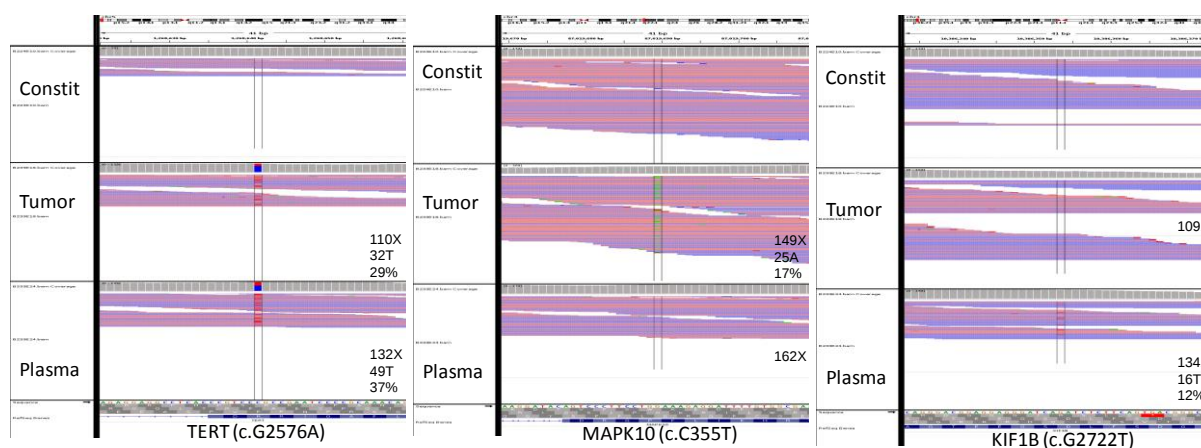


Figure 3.4.4 : SNVs observés pour le cas 21 au diagnostic : dans *TERT*, variant commun à la tumeur et au plasma ; dans *MAPK10*, variant spécifique à la tumeur et dans *KIF1B*, variant spécifique au plasma.

La tumeur semble posséder ses propres altérations telles qu'un variant dans le gène *MAPK10* (Figure 3.4.4) mais également des amplifications sur le chromosome 12 contenant les gènes *CACNA1C* (chr12: 2566733-2743578), *CMAS* (chr12: 21918863-22440262) (gène surexprimé dans des cancers du sein triple négatifs), *KRAS* (chr12: 24968390-25587246) et *CDK4* (chr12: 57944177-60098792) (Figure 3.4.5)³⁵⁹.

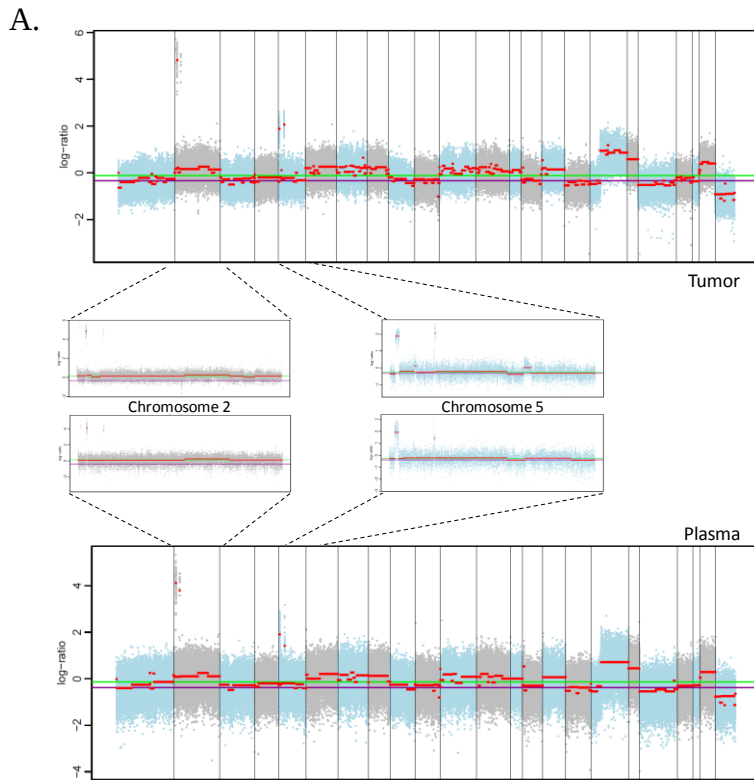
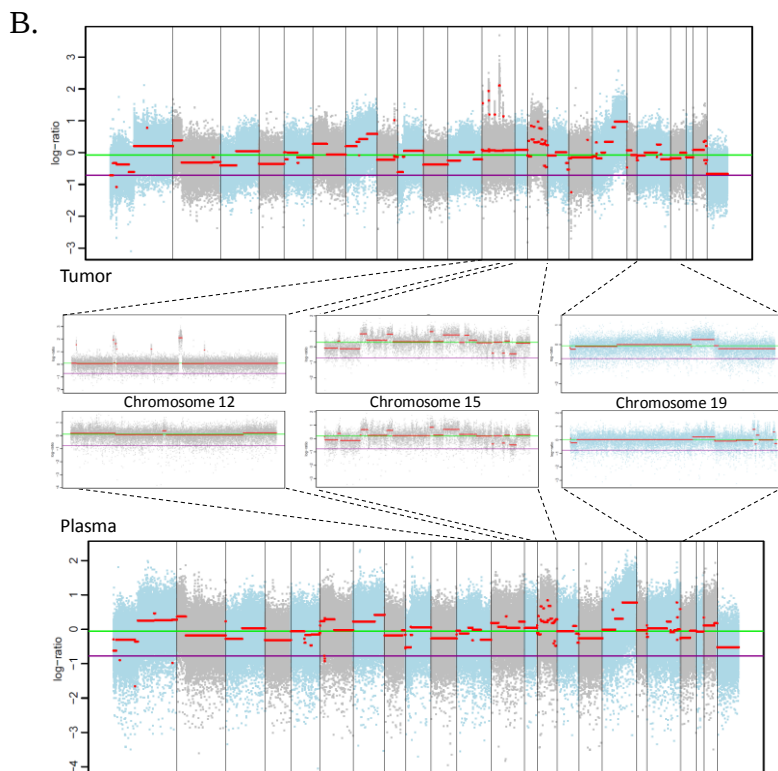


Figure 3.4.5 : Profil d'altération en nombre de copies de deux cas de neuroblastome représentés par Facets avec les WES sur la tumeur (en haut) et le plasma (en bas) au diagnostic.

A. Pour le premier cas (Cas 32), les mêmes altérations sont observées dans la tumeur et le plasma telles que des amplifications sur le chromosome 2 de *MYCN* (chr2: 15555651-16769308), du promoteur de *ALK* (chr2:30143176-30775367) ou sur le chromosome 5 de *TERT* (chr5: 1200946-5182295) ou *ITGA1* (chr 5: 51307439-52322786).



B. Pour le second cas (cas 21), la majorité des altérations sont communes entre la tumeur et le plasma notamment un chromothripsis sur le chromosome 15. Certaines altérations sont cependant spécifiques telles que des amplifications sur le chromosome 12 dans la tumeur contenant les gènes *CACNA1C* (chr12: 2566733-2743578), *CMAS* (chr12: 21918863-22440262) *KRAS* (chr12: 24968390-25587246), *CDK4* (chr12: 57944177-60098792) et pour le plasma (chr12: 54765170-56117811) contenant les gènes *ITGA5* et *ITGA6*.

Des amplifications sont également observées sur le chromosome 19 uniquement dans le plasma telles que le cluster de miRNA *C19MC* (chr19: 54175287:54611734).

La présence de nombreux gènes dans ces amplifications a permis à Facets de mieux caractériser ses amplifications contrairement à ichorCNA.

Des altérations sont également spécifiques au plasma comme une mutation dans le gène *KIFB* déjà observée dans des cas de neuroblastomes (Figure 3.4.4)³⁶⁰. Le chromosome 12 possède également une région gagnée (chr12:54765170-5611781) contenant les gènes *GTSF1*, *ITGA5*, *ITGA7*, *TESPA1* et *NEUROD4* dont la surexpression a été décrite dans des cas de germinomes, cancer du poumon et du pancréas³⁶¹⁻³⁶⁴. Des amplifications sur le chromosome 19 sont aussi présentes uniquement dans le plasma notamment le cluster de miRNA C19MC (chr19: 54175287:54611734), amplifié dans certains cancers notamment les tumeurs cérébrales de types ETMR (Figure 3.4.5). L'amplification de ce cluster pourrait jouer un rôle dans la reprogrammation des cellules pour les transitions EMT³⁶⁵⁻³⁶⁷.

3.4.2.2 Rhabdomyosarcome

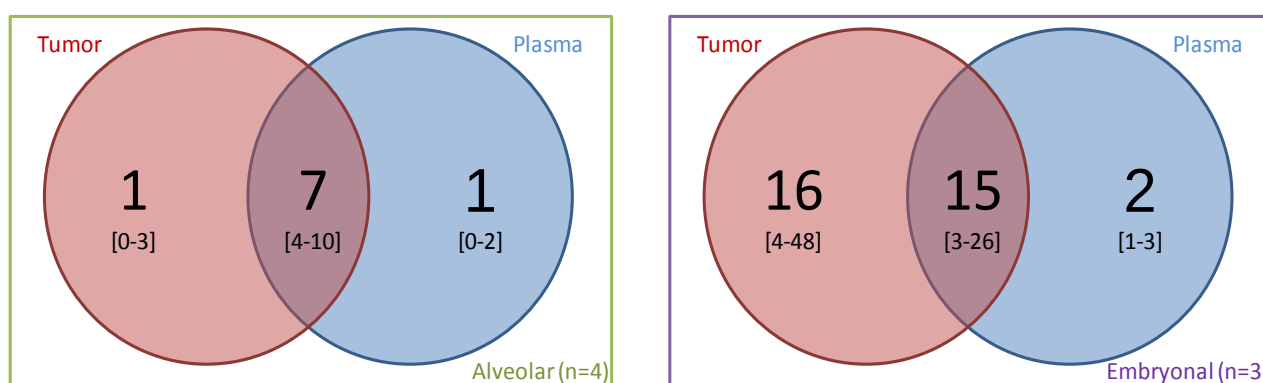


Figure 3.4.6 : Diagramme de Venn des SNVs détectés au diagnostic dans les rhabdomyosarcomes de type alvéolaire ou embryonnaire par WES sur la tumeur et le plasma.

Le contenu tumoral dans le plasma semble assez proche entre les deux types de rhabdomyosarcome avec une médiane de 5,13% allant de 0 à 33% pour les formes alvéolaires et une médiane de 5,79% allant de 0 à 67,3% pour les formes embryonnaires.

Un nombre plus important de variants a été détecté dans le type embryonnaire que dans le type alvéolaire qui est plutôt dû à la fusion *PAX-FOX*⁸³. Cependant le WES réalisé sur

l'ADN circulant ne permet pas de détecter cette fusion ayant lieu dans un intron des deux gènes pour la majorité des cas.

Toutefois, une médiane de 7 SNVs communs entre la tumeur et le plasma a été détectée avec une mutation dans le gène *CTC1* impliquée dans la maintenance des télomères. Cependant et en dehors de cette mutation, la plupart des variants détectés ne semblent pas être impliqués dans l'oncogénese (Figure 3.4.6)³⁶⁸.

Chez les patients atteints de rhabdomyosarcome embryonnaire, une médiane de 15 SNVs communs a été observée entre la tumeur et le plasma avec de nombreuses altérations dans des oncogènes tels que *KRAS* ou *BCOR* (Figure 3.4.7). Cependant, les altérations spécifiques semblent être moins présentes dans l'ADN circulant (médiane de 2 SNVs) que dans l'ADN tumoral (médiane de 16 SNVs). Certaines des mutations spécifiques à la tumeur touchent également des oncogènes tels que *BRAF* ou *BCOR*.

Selon ces observations, certaines cellules du rhabdomysarcome paraissent relâcher peu ou pas d'ADN circulant.

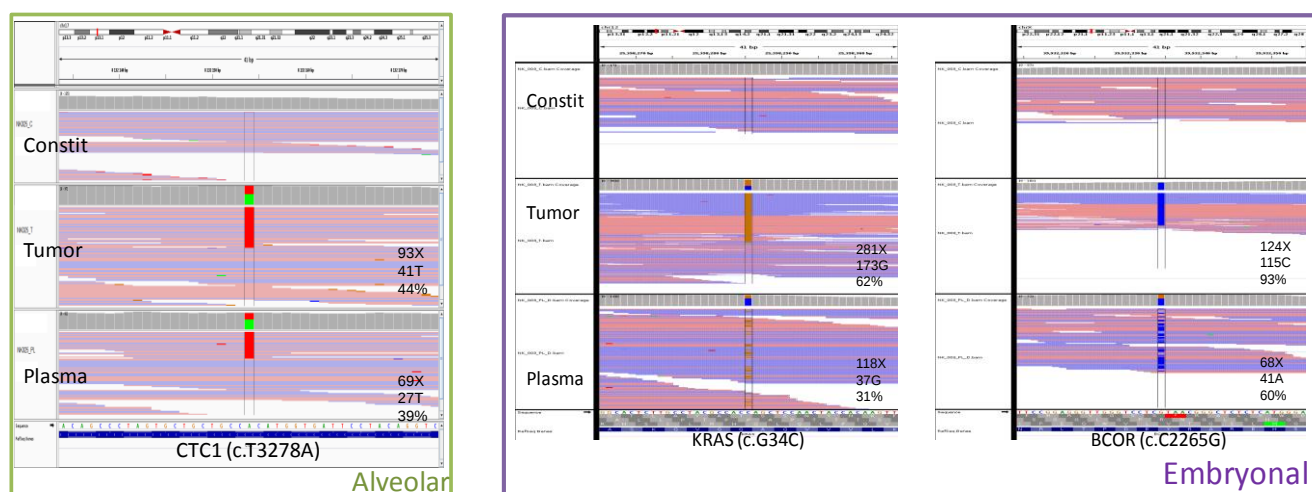


Figure 3.4.7 : SNVs observés au diagnostic dans le gène *CTC1* pour un cas de RMS alvéolaire et dans les gènes *KRAS* et *BCOR* pour un cas de RMS embryonnaire dans la tumeur et le plasma avec l'outil IGV.

Malgré un contenu tumoral peu élevé pour certains cas, le plasma présentait les mêmes altérations en nombre de copies que la tumeur hormis pour deux cas (1 aRMS et 1 eRMS) où aucune altération n'a été détectée dans l'ADN circulant.

Une médiane de 16 altérations (de 3 à 21 altérations) est observée dans le plasma contenant du ctDNA et la tumeur. De 1 à 9 altérations ont été observées seulement dans la tumeur contre 0 à 1 altération seulement dans le plasma.

Pour un cas de RMS de type alvéolaire, une amplification sur le chromosome 9 et une sur le chromosome 12 ont été observées dans la tumeur et dans le plasma (Figure 3.4.8). Ces altérations ont déjà été décrites dans certains cas de RMS alvéolaires notamment pour l'amplification 12q13-q14 contenant de nombreux oncogènes tels que *GLI1* et *CDK4*^{369–371}.

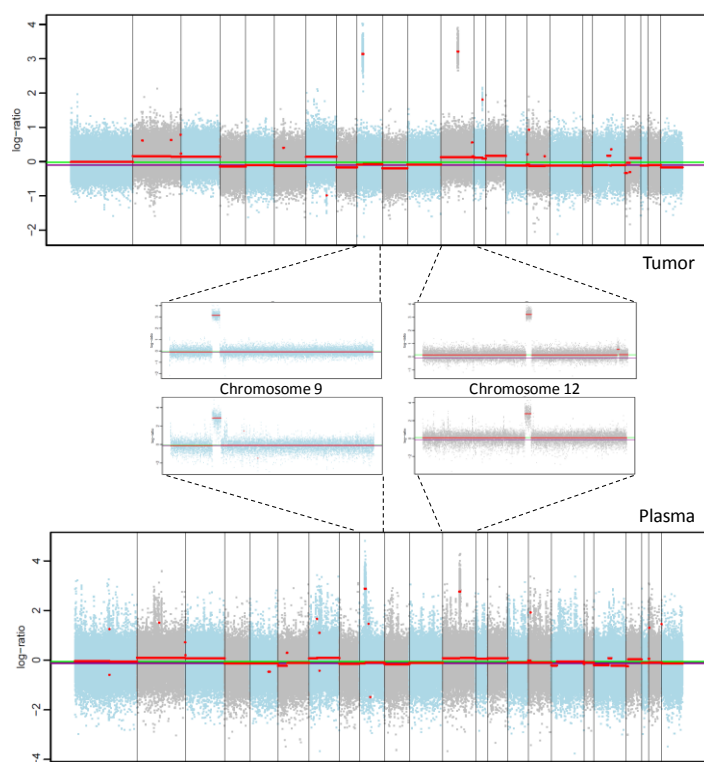


Figure 3.4.8 : Profils d'altérations en nombre de copies d'un cas de rhabdomyosarcome alvéolaire (Cas 25) représentés par Facets avec les WES sur la tumeur (en haut) et le plasma (en bas).

Deux amplifications sont observées dans le chromosome 9 (chr9: 35228155-36930771) et le chromosome 12 (chr12: 57566929-58378521). La deuxième amplification contient de nombreux oncogènes tels que *GLI1* ou *CDK4*.

La présence de nombreux gènes dans ces amplifications a permis à Facets de mieux caractériser ces amplifications contrairement à ichorCNA.

3.4.2.3 Sarcome d'Ewing

Le contenu de ctDNA dans le plasma a pu être prédit pour un seul cas de sarcome d'Ewing à 8,71% par ichorCNA avec 3 altérations communes à la tumeur et au plasma et 3 uniquement observées dans la tumeur. Les autres cas avaient des profils sans altération en nombre de copies y compris pour la tumeur ne permettant pas de prédire le contenu en ctDNA.

A l'instar des RMS alvéolaires, la fusion caractéristique de ces tumeurs n'a pas pu être détectée par WES mais une médiane de 4 SNVs communs à la tumeur et à l'ADN circulant a été observée. Néanmoins, une médiane de 0 a été détectée dans le plasma contre une médiane de 2 SNVs allant de 0 à 22 pour la tumeur (Figure 3.4.9). Pour le cas présentant une forme localisée, les résultats ont montré un nombre important d'altérations dans l'ADN circulant.

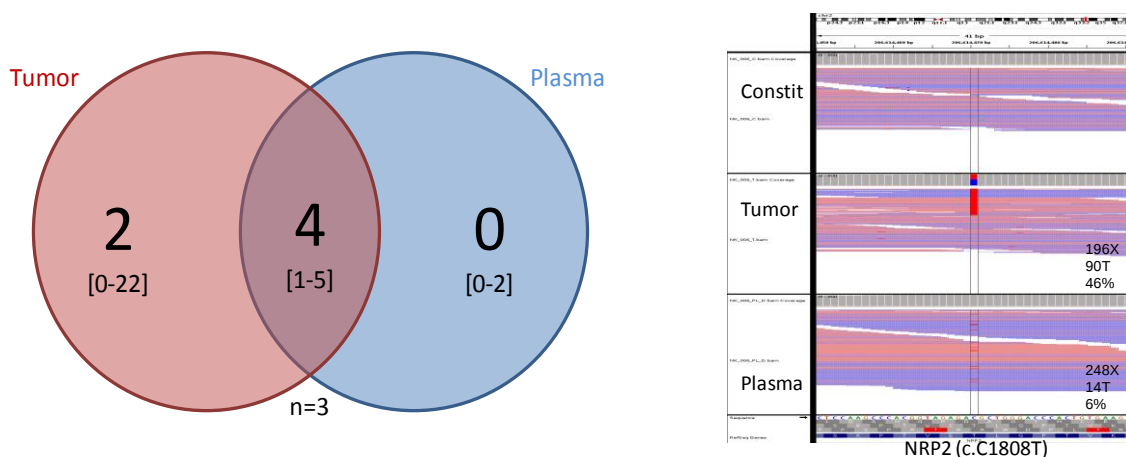


Figure 3.4.9 : Diagramme de Venn des SNVs détectés par WES au diagnostic dans les sarcomes d'Ewing entre la tumeur et le plasma (à gauche). Un SNV détecté sur le gène *NRP2* dans la tumeur et le plasma (à droite).

Ces tumeurs étant peu mutées et avec une libération de ctDNA limité, le WES n'est peut-être pas la meilleure option^{8,107}. La recherche de la fusion avec le WGS low coverage est sans doute possible mais demanderait une optimisation des outils bio-informatiques dans la recherche de breakpoint^{333,372,373}.

3.4.2.4 Tumeur du foie

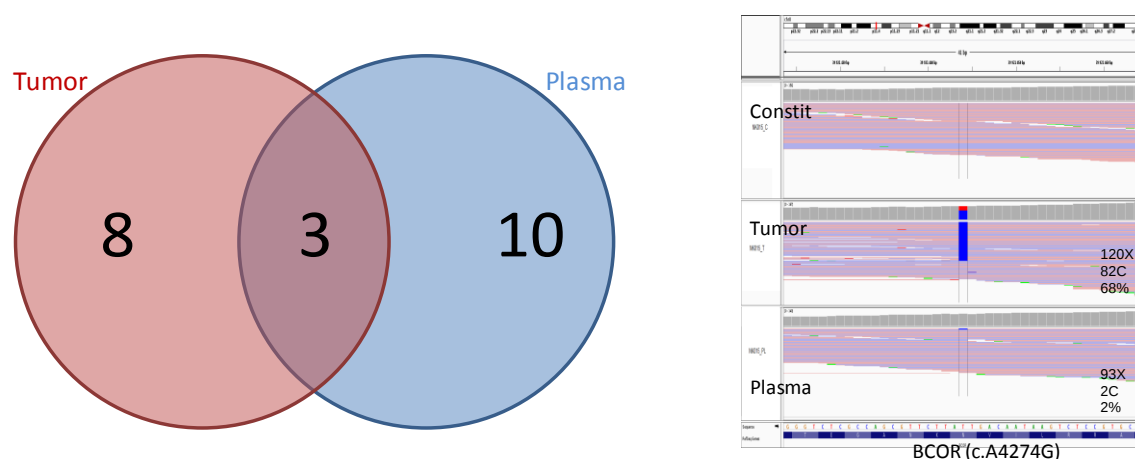


Figure 3.4.10 : Diagramme de Wenn des SNVs détectés par WES au diagnostic dans un hépatoblastome entre la tumeur et le plasma (à gauche). Un SNV détecté sur le gène *BCOR* dans la tumeur et le plasma (à droite).

Pour deux patients atteints de tumeur du foie (un hépatoblastome métastatique et un carcinome hépatique), un séquençage par WES de la tumeur et de l'ADN circulant a été réalisé au diagnostic. Le plasma du patient atteint de carcinome n'a été séquençé que par WES et contenait 19% de ctDNA selon Facets. Pour ce patient, aucun des 18 SNVs ou altérations en nombre de copies identifiées dans la tumeur n'ont été retrouvés dans le plasma laissant suggérer une erreur de Facets. Au contraire le plasma du patient atteint d'hépatoblastome contenait 22,15% de ctDNA selon ichorCNA.

Pour le patient atteint d'hépatoblastome, seulement 3 SNVs identifiés dans la tumeur sont retrouvés dans le plasma et souvent à des fractions alléliques très faibles comme pour *BCOR* (Figure 3.4.10, 68% dans la tumeur contre 2% dans le plasma). Huit SNVs étaient spécifiques de la tumeur et 10 du plasma. Ces différences s'observent également au niveau chromosomique avec un profil très remanié pour le plasma (13 altérations) dont seulement un gain du 6p qui est commun à la tumeur, seule et unique altération identifiée dans la tumeur (Figure 3.4.11).

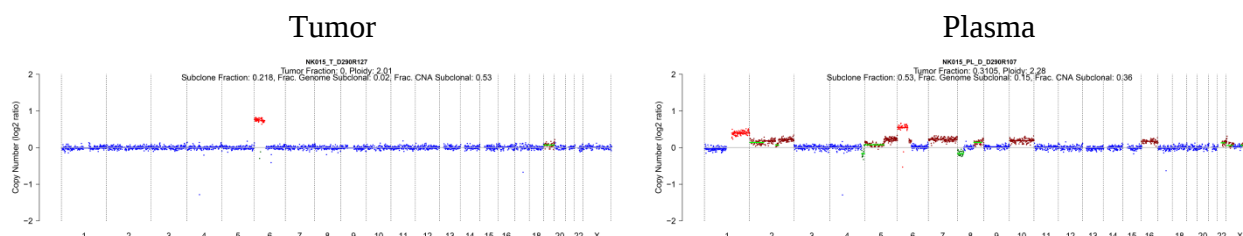


Figure 3.4.11 : Profils d'altérations en nombre de copies obtenu avec ichorCNA par WGS low coverage pour un cas d'hépatoblastome de la tumeur (à gauche) et de l'ADN circulant (à droite)

La présence de métastases peut expliquer ces différences, l'ADN circulant reflétant plutôt les sites métastatiques que la tumeur primaire. En effet, un clone ancestral a pu apparaître avec gain du 6p puis très tôt un clone localisé et un clone métastatique se seraient différenciés et auraient évolué différemment.

Contrairement à ceux que l'on pourrait penser, bien qu'étant un organe de grande taille et d'une importance majeure dans l'équilibre de la circulation sanguine, les tumeurs localisées

du foie ne semblent pas libérer des quantités importantes de ctDNA dans le sang. Au vu de sa fonction de filtre, le foie pourrait être la dernière étape dans la circulation d'ADN circulant. Ainsi l'analyse d'urine serait peut-être plus adaptée que celles du sang pour certaines tumeurs hépatiques^{216,374}.

3.4.2.5 MPNST

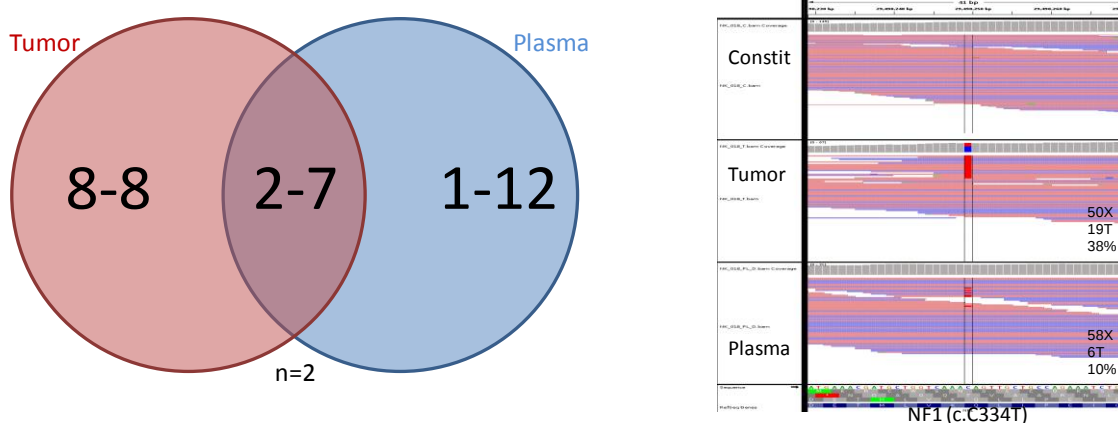


Figure 3.4.12 Diagramme de Venn des SNVs détectés par WES au diagnostic dans les MPNST entre la tumeur et le plasma (à gauche). Un SNV détecté sur le gène *NF1* dans la tumeur et le plasma (à droite).

Les MPNST (Malignant Peripheral Nerve Sheath Tumours) sont des tumeurs de l'enveloppe des nerfs et contrairement aux tumeurs cérébrales, elles semblent relarguer une quantité importante de ctDNA (16 à 17% selon ichorCNA). Deux à 7 SNVs communs entre la tumeur ont été détectés avec notamment une mutation perte de fonction dans le gène *NF1* connue pour être impliquée dans ce type de tumeur (Figure 3.4.12) mais également dans *PTCH1* plutôt impliqué habituellement dans les médulloblastomes^{75,375}. Huit SNVs sont spécifiques de la tumeur tels qu'un variant du gène *KDM2B*. Un à 12 SNVs sont spécifiques du plasma tels que des variants de *CNTNAP5* ou *PRKCA*.

La perte de fonction de *CDKN2A* est aussi récurrente dans ce type de tumeurs et pourrait être possible dans les 2 suite à une perte observée d'une partie du chromosome 9 (Figure 3.4.13).

L'ADN circulant possède également des altérations segmentaires supplémentaires par rapport à la tumeur notamment un gain du 1q dans les deux cas. Le second cas a un profil d'ADN circulant qui semble également plus remanié (21 altérations supplémentaires) que la tumeur au diagnostic. Cette observation laisse supposer que le ctDNA aurait une origine plutôt métastatique pour ces deux tumeurs^{295,353}.

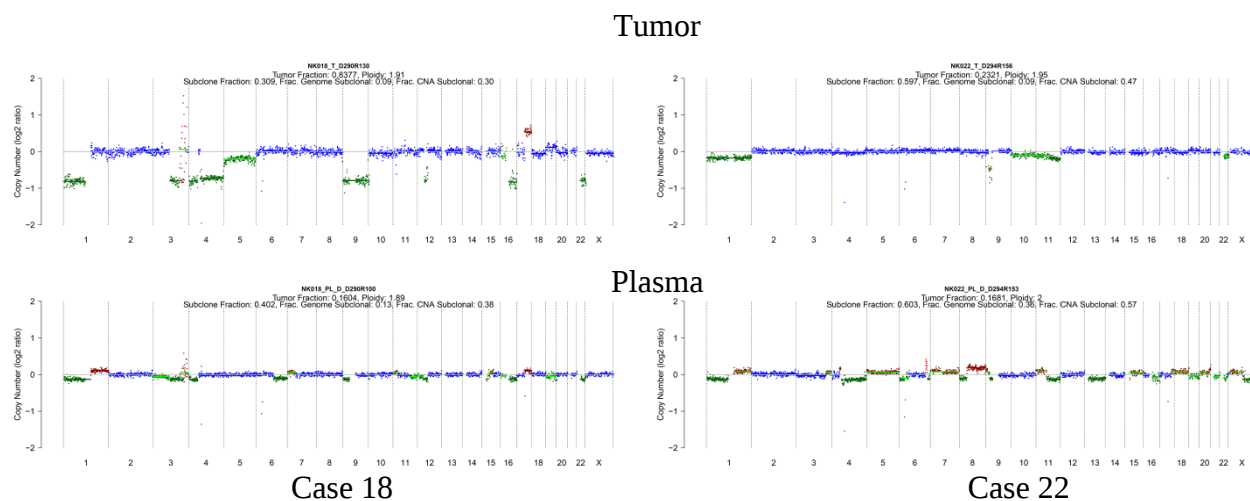


Figure 3.4.13 : Profils d'altérations en nombre de copies obtenu avec ichorCNA par WGS low coverage au diagnostic pour les deux cas de MPNST sur la tumeur (en haut) et le cfDNA (en bas). La majorité des altérations de la tumeur sont observées dans le plasma notamment une perte sur le chromosome 9 contenant le gène *CDKN2A*.

3.4.2.6 Tumeurs cérébrales

Pour 5 patients atteints de tumeurs cérébrales, les tumeurs primaires ont été séquencées par WES mais aussi les plasmas au diagnostic malgré la possibilité de non-circulation de ctDNA due à la barrière hématoïétique.

Le cas de tumeur rhabdoïde térétoïde atypique (ATRT) présentait 9 SNVs dans la tumeur et 5 dans le plasma. Les SNVs détectés dans le plasma pourraient être de faux positifs car supportés par quelques reads (2 à 5) et le contenu tumoral du plasma était estimé à 0% selon ichorCNA. La perte d'une partie du chromosome 22 contenant le gène *SMARCB1* typique des tumeurs rhabdoïdes n'était observée que dans la tumeur³⁷⁶.

Pour la tumeur embryonnaire "NOS", 23 SNVs ont été détectés seulement dans la tumeur. Trois SNVs ont été détectés dans le plasma dont 1 commun à la tumeur mais le contenu tumoral dans le plasma était de 0%. De nombreuses altérations sont présentes dans des oncogènes telles que des mutations dans *TP53*, *NF1* et des amplifications de *MYCN* ou *IGF1R* mais n'ont pas été trouvées dans le plasma.

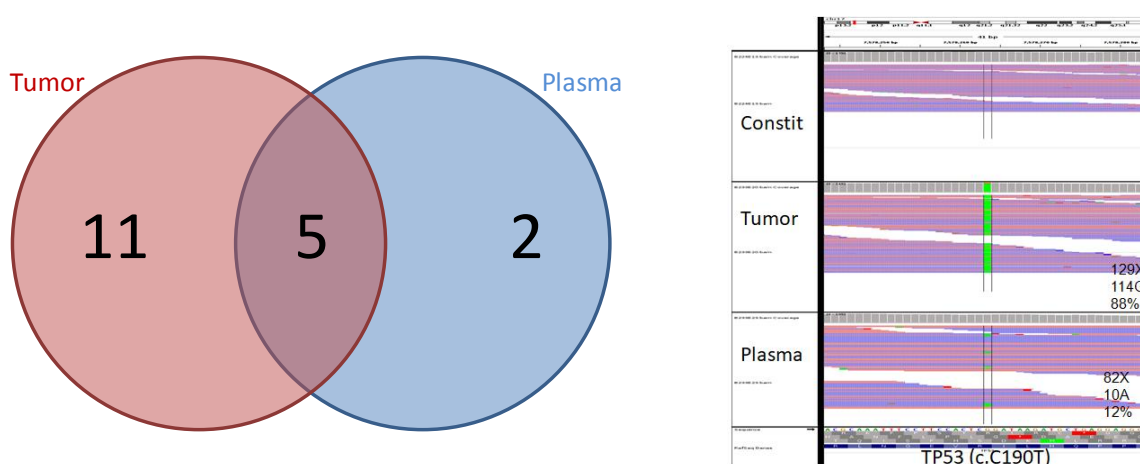


Figure 2.3.14 Diagramme de Venn des SNVs détectés par WES au diagnostic dans un gliome entre la tumeur et le plasma (à gauche). Un SNV détecté sur le gène *TP53* dans la tumeur et le plasma (à droite).

Un des deux cas de gliome de haut grade présentait un contenu de ctDNA de 11% dans le plasma. 5 SNVs étaient communs entre la tumeur et l'ADN circulant notamment une mutation dans le gène *TP53* (Figure 3.4.14). Onze SNVs étaient spécifiques de la tumeur contre 2 pour le plasma pour ces cas.

Malgré un faible contenu de ctDNA, quelques altérations de la tumeur sont retrouvées telles que les amplifications dans le chromosome 1, 2, 5, 8, 19 et 20 mais également de pertes dans le chromosome 8, 9 et 17 (Figure 3.4.15).

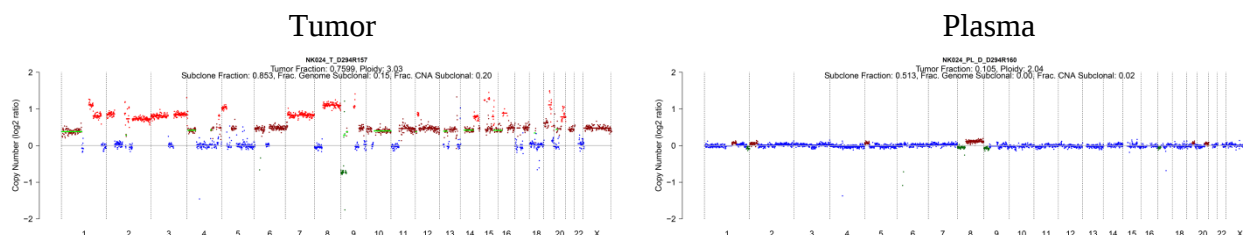


Figure 3.4.15 : Profils d'altérations en nombre de copies pour un cas de gliome (cas 24) obtenu avec ichorCNA par WGS low coverage au diagnostic sur la tumeur (à gauche) et le cfDNA (à droite). Malgré un faible contenu de ctDNA, quelques altérations ont pu être détectées.

La tumeur étant située dans les voies optiques et la présence de certaines mutations dans des régions amplifiées telles que *TP53* sur le chromosome 17 pourrait être une explication de la détection de certaines altérations dans le plasma.

Pour le second cas de gliome de haut grade, le contenu en ctDNA dans le plasma était prédit à 0%. Vingt SNVs étaient identifiés dans la tumeur et spécifiques à celle-ci. Un seul SNV commun a pu être détecté dans la tumeur et le plasma et 2 SNVs présents seulement dans le plasma. Comme pour la majorité des autres tumeurs cérébrales, le plasma ne semble pas contenir d'information génétique provenant de la tumeur.

Le plasma ne permet donc pas de caractériser les tumeurs exclusivement cérébrales par WES et l'utilisation du LCR est certainement plus adaptée mais reste plus compliquée à obtenir et *a fortiori* en faible volume (3 à 6 gouttes de 100 μ L).

Type de Cancer	Nombre de patients	Médiane de SNVs Communs	Médiane SNVs spécifiques à la tumeur	Médiane SNVs spécifiques au plasma
Neuroblastome	42	14	2	2
Rhabdomyosarcome	7	11	7	1
Sarcome d'Ewing	3	4	2	0
MPNST	2	4	8	6
Gliome	1	5	11	2
Tumeur du foie	1	3	8	10

Tableau 3: Médianes des SNVs détectés par WES selon le type de tumeurs pédiatriques.

3.4.2.7 Rechute

Parmi les 31 patients atteints de tumeurs de haut risque décrits précédemment, 6 cas ont rechuté lors de leur suivi : 2 patients atteints de neuroblastome, 1 patient atteint de sarcome d'Ewing, 1 patient atteint de rhabdomyosarcome, 1 patient atteint de MPNST et 1 patient atteint de gliome.

Pour le patient atteint de sarcome d'Ewing, le contenu tumoral n'a pu être calculé en l'absence d'altérations en nombre de copies sur la tumeur au diagnostic. Les nouveaux SNVs détectés à la rechute dans le plasma sont supportés par peu de reads. Ces derniers devront être confirmés par un séquençage ciblé à plus haute profondeur.

Pour le patient atteint de gliome, le contenu tumoral du plasma est passé de 11% à 0% entre le diagnostic et la rechute. La mutation dans *TP53* n'est plus observée et les quelques SNVs détectés dans le plasma à la rechute sont présents dans peu de reads.

Pour le patient atteint de rhabdomyosarcome, le contenu tumoral du plasma passe de 5,79% à 10% entre le diagnostic et la rechute. Deux nouvelles altérations ont été détectées à la rechute dans le plasma mais seulement supportées par quelques reads. Certaines altérations détectées seulement dans la tumeur primaire au diagnostic apparaissent à la rechute dans le plasma. D'autres altérations présentes au diagnostic dans le plasma et la tumeur telles qu'un variant dans *JAG2* disparaissent du plasma à la rechute. Les SNVs suivent une certaine dynamique pouvant être due à l'évolution clonale mais la sensibilité du WES est limitée nécessitant un séquençage à plus haute couverture.

Pour les deux cas de neuroblastome, le contenu tumoral dans le plasma est très élevé au diagnostic (84% et 88%) et reste élevé à la rechute (82% et 88%). Aucune altération en nombre de copies supplémentaire n'est détectée pour le premier cas à rechute (cas 1). Dans le second cas (cas 8), un gain du chromosome 7 est observé au diagnostic dans le plasma de

passant 2 copies à 3 copies. À la rechute, une des copies du chromosome 7 a perdu une partie de son bras court (7p).

De plus, une explosion du nombre de SNVs additionnels est observée à la rechute avec 42 mutations supplémentaires pour le premier cas et 101 pour le second. La majorité de ces SNVs est prédite comme ayant un impact délétère sur la protéine et touche certains oncogènes tels que *TP73* ou *ADCY3* (Figure 3.4.16).

A noter que le second est assez particulier car la tumeur primaire était une forme localisée de neuroblastome. Le patient a ensuite rechuté deux fois en neuroblastome de stade 4 et une localisation différente de la tumeur du diagnostic. Le patient a été inclus dans NGS Kids à partir de sa première rechute.

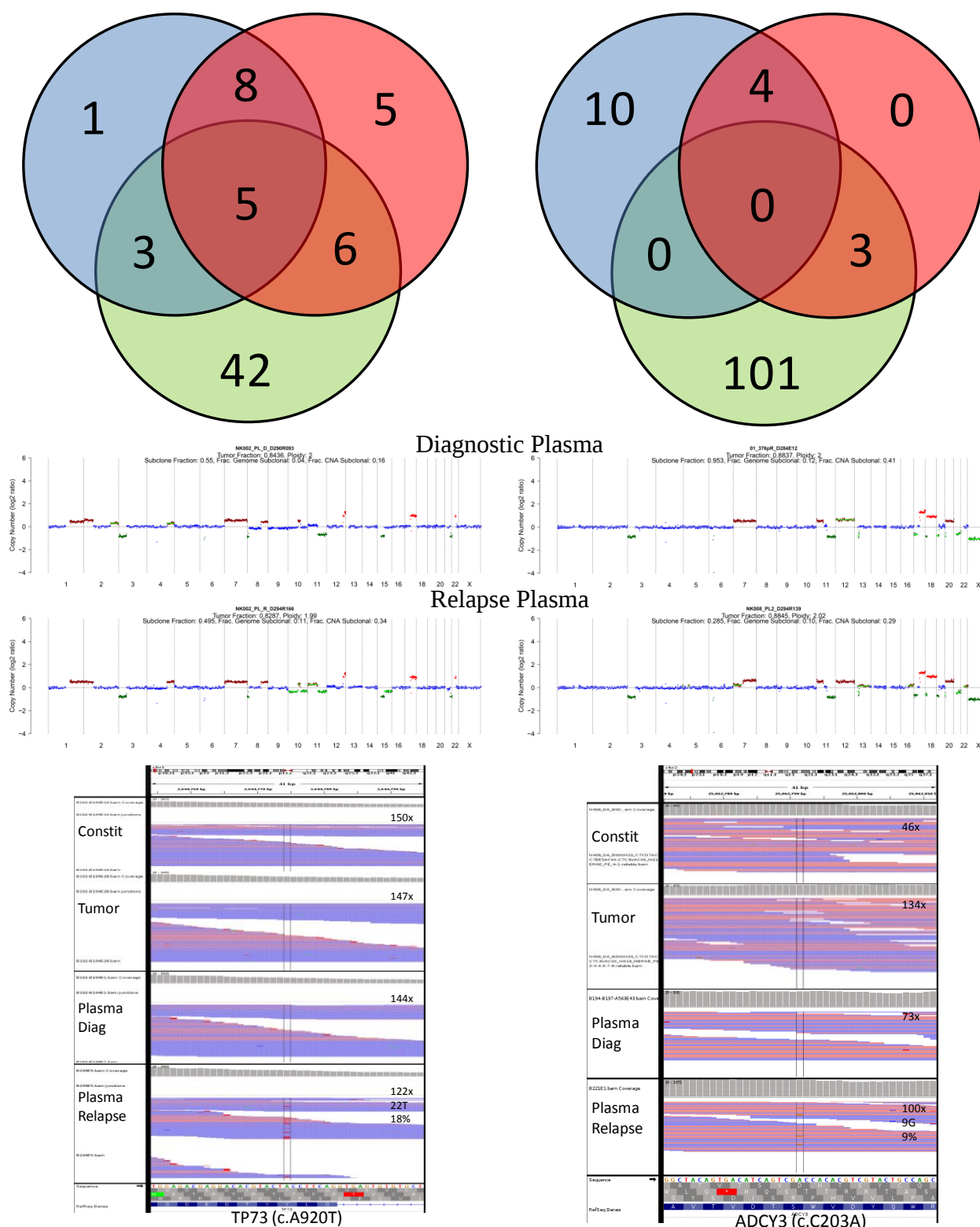


Figure 3.4.16 : En haut, diagramme de Venn des SNVs entre la tumeur au diagnostic (en bleu), le plasma diagnostic (en rouge) et le plasma rechute (en vert) pour le cas 2 (à gauche) et le cas 8 (à droite). Au centre, les profils d'altérations en nombre de copies entre le plasma au diagnostic et la rechute obtenus par WGS low coverage. Une seule altération supplémentaire est détectée à la rechute pour le cas 8 (perte du 7p). En bas, SNVs détectés à la rechute dans le gène *TP73* pour le cas 2 (à gauche) et *ADCY3* pour le cas 8 (à droite).

3.4.3 Identification des voies de signalisation impactées par les altérations génétiques détectées dans la tumeur et le plasma

Afin d'analyser les voies de signalisations impactées par les altérations génétiques vues dans les échantillons de tumeur ou de plasma, les gènes ciblés par des SNVs détectés par WES ont été soumis au logiciel Ingenuity Pathway Analysis (IPA). La totalité des gènes (n=1511) identifiés comme contenant au moins un SNV dans n'importe quel échantillon a été utilisée puis par groupes en fonction de la maladie ou de la nature du prélèvement (tumeur ou plasma) mais aucun enrichissement n'a été observé.

Ce set de gènes a ensuite été soumis à l'outil Reactome développé par l'European Molecular Biology Laboratory (EMBL) et qui propose une alternative gratuite à IPA³⁷⁷.

Ainsi, ces gènes détectés comme mutés dans l'ADN circulant, exclusifs ou communs à la tumeur, font partie des voies impliquant l'interaction entre cellules, l'organisation extracellulaire et également la communication cellule à cellule.

Pour le neuroblastome, un enrichissement dans la voie de SUMOylation est également observé pour les gènes détectés comme mutés dans le plasma.

La SUMOylation est un procédé très proche de l'ubiquitination dans la modification post-traductionnelle avec des enzymes communes telles que UPS22 ou ZMYM4³⁷⁸. Cependant, l'ubiquitination entraîne dans la majorité des cas à une dégradation des protéines, après une fixation d'ubiquitine, ces dernières sont détruites par le protéasome. La SUMOylation a plutôt une fonction de régulateur notamment des facteurs de transcription, des interactions entre protéines mais également du transport intra-extracellulaire³⁷⁹.

La modification de cette voie semble jouer un rôle important dans certains cancers avec des implications dans le cycle cellulaire, l'apoptose ou la prolifération³⁸⁰. Dans de nombreux types de cancers tels que les cancers du sein, de la prostate ou colorectal, la surexpression de certaines protéines du complexe SUMO tels que SENP1 ou Ubc9 amène à des formes de

cancers plus agressives voire métastatiques et la déplétion de ces protéines réduit la prolifération et l'invasion³⁸¹⁻³⁸⁵. Les interactions entre le complexe SUMO et les facteurs de transcription tels que STAT1 peuvent expliquer ces formes invasives de la tumeur avec des possibilités de changement d'identité vers des formes cellulaires mésenchymateuses plus mobiles^{386,387}.

Ces sets de gènes ont également été soumis à la base de données Gene Ontology (GO) par l'intermédiaire de l'outil Gene Set Enrichment Analysis (GSEA) bien que les données proviennent majoritairement de profil d'expression dans cette base de données³⁸⁸⁻³⁹¹. Une nouvelle fois, les voies d'interaction entre cellules et également de locomotion ont été retrouvées comme enrichies mais pour les gènes mutés exclusivement dans l'ADN circulant.

L'ensemble des outils utilisés montre que les enrichissements de voie dans le ctDNA, la partie tumorale de l'ADN circulant, concernent des gènes plutôt impliqués dans la migration cellulaire. Ceci amène à l'hypothèse que le ctDNA serait plutôt issu de clones ayant tendance à migrer et ainsi plus proches des métastases.

3.4.4 Séquençage ciblé et modélisation de l'évolution clonale

Une nouvelle fois, les SNVs détectés par WES sur la tumeur et sur le plasma ont été suivi dans le temps pour suivre l'évolution de la maladie.

Ainsi, 1151 SNVs ont été détectés en WES dans la tumeur et dans l'ADN circulant au diagnostic et pour certains patients, à la rechute. Avec une approche comparable à l'étude précédente sur le neuroblastome, un panel a ainsi été dessiné couvrant 200 bp autour de chaque SNV ainsi que 44 SNPs comme contrôle d'identité de chacun des échantillons pour une capture totale de 212kb.

Le séquençage ciblé a ensuite été réalisé sur les tumeurs primaires, le plasma au diagnostic ainsi que les plasmas de suivi (de 1 à 9 plasmas par patient) pour 28 patients.

Vingt-trois échantillons ont été poolés ensemble pour un total de 6 captures. Le protocole de double capture a été une nouvelle fois utilisé. Les captures ont ensuite été séquencées sur HiSeq 2500 Rapid Run™ avec une capture par ligne de flowcell permettant une couverture théorique de 2000 à 5000x.

Les fréquences d'allèles mutées détectées en séquençage ciblé sont fortement corrélées aux résultats obtenus en WES (Figure 3.4.17).

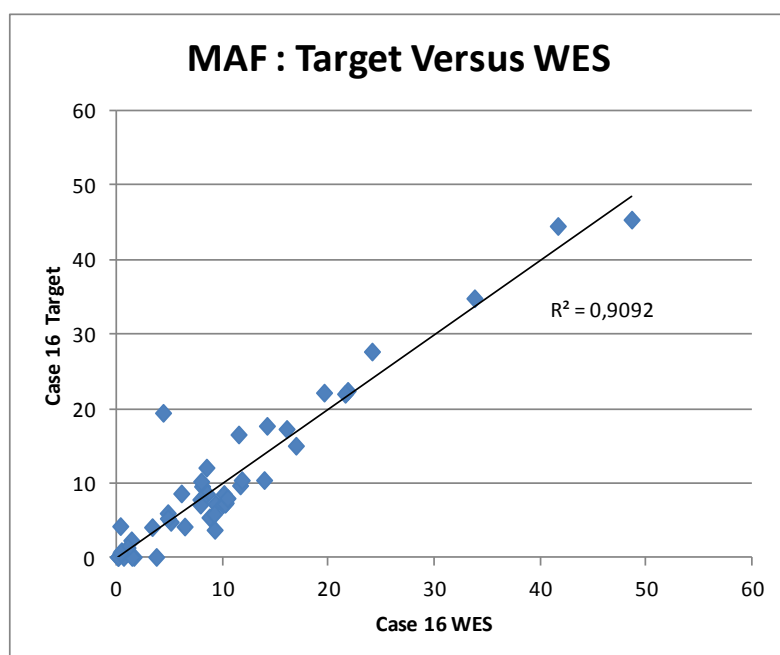


Figure 3.4.17 : Corrélation observée entre les fréquences des SNVs détectés en WES et en TargetSequencing pour le plasma au diagnostic du cas 16 ($R^2=0,9092$).

Les SNVs ont été ensuite suivies avec Samtools et seuls les SNVs supportés par 10 reads étaient considérés et permettant d'identifier les SNVs avec une sensibilité jusqu'à de 0,2%.

3.4.4.1 Neuroblastome

Sur les dix cas de neuroblastome avec un suivi de prélèvements de plasma, 9 possédaient une dynamique de fréquences alléliques de leurs SNVs dans les plasmas. Pour le cas 4, un patient atteint de neuroblastome de stade 2 (localisé), les résultats du séquençage ciblé n'ont pas permis d'analyser ce cas.

Pour les cas 13 et 14, les patients ont très vite répondu positivement à la chimiothérapie avec notamment une réduction de 75% du volume de la tumeur primaire pour le cas 13. Pour ce dernier, une augmentation du contenu en ctDNA dans le plasma ainsi de la fréquence allélique des SNVs identifiés par WES dans le plasma et la tumeur est observée après 4 jours de traitement (Figure 3.4.18). Ce surcroît de contenu tumoral serait plutôt dû à la destruction des cellules malignes par la chimiothérapie, et donc une libération de ctDNA dans le sang, qu'à une progression de la maladie^{392,393}.

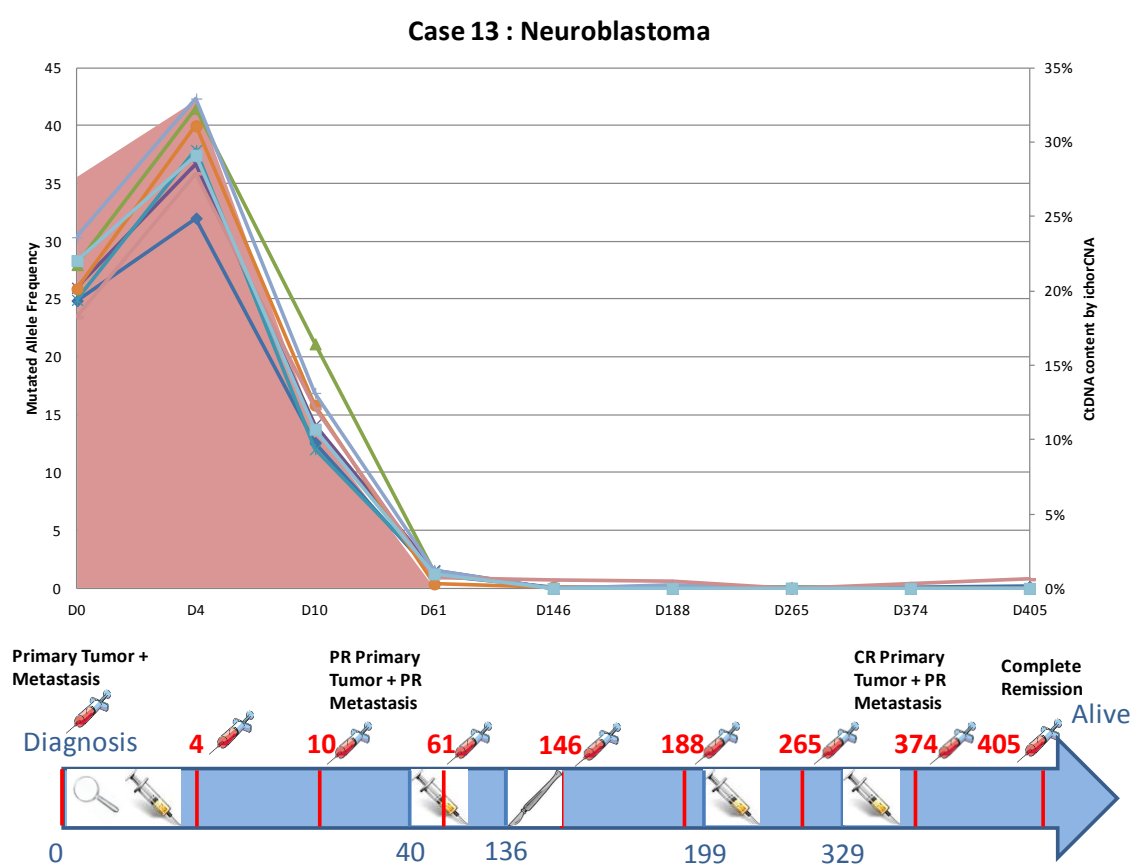
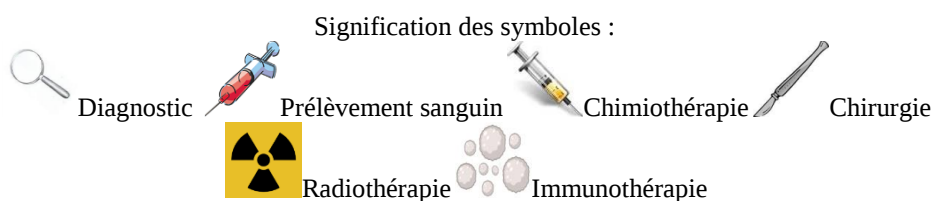


Figure 3.4.18 : Suivi du cas 13, patient atteint de neuroblastome avec une réponse au traitement. Les courbes représentent les fractions alléliques mutées et l'aire rose le contenu tumoral de l'ADN circulant calculé par ichorCNA.

La flèche sous le graphique correspond aux traitements reçus par le patient ainsi que le statut clinique de la maladie (CR : Complete Remission, PR : Partial Remission, SD : Stable Disease, PD : Progressive Disease).



Le suivi des quantités de ctDNA, reflétées par la fréquence allélique des SNVs et du contenu tumoral obtenu par WGS low coverage, sont en accord avec les statuts cliniques des patients. Cependant la couverture du Target reste trop faible pour détecter de la maladie résiduelle alors que la rémission des patients n'était pas complète.

Pour le cas 21, seulement une partie des SNVs identifiés dans la tumeur par WES est retrouvée dans le plasma et inversement mais avec des fréquences alléliques très proches du seuil de sensibilité du séquençage ciblé de 0,2% avec 11 reads alternatifs à 1,15% avec 38 reads alternatifs.

Pour ce cas métastatique, les métastases relargueraient donc une quantité importante de ctDNA et masqueraient en partie, le ctDNA provenant de la tumeur. Ces clones métastatiques pourraient être issus des sous-clones de la tumeur propices à la migration notamment avec la présence d'une amplification du cluster C19MC. Seul le cas 21 présentait des altérations sous-clonales dans la tumeur qui avait été détectées à un niveau clonal dans le plasma par WES.

Une nouvelle fois, une augmentation du ctDNA est observée après 5 jours de chimiothérapie alors que la maladie régresse, synonyme de bonne réponse à la chimiothérapie (Figure 3.4.19).

Malheureusement, le patient a rechuté 9 mois après l'arrêt du suivi, bien après le dernier plasma recueilli dans le cadre du protocole. Le délai trop important entre ce dernier plasma et la rechute ne permet pas d'étudier cette dernière.

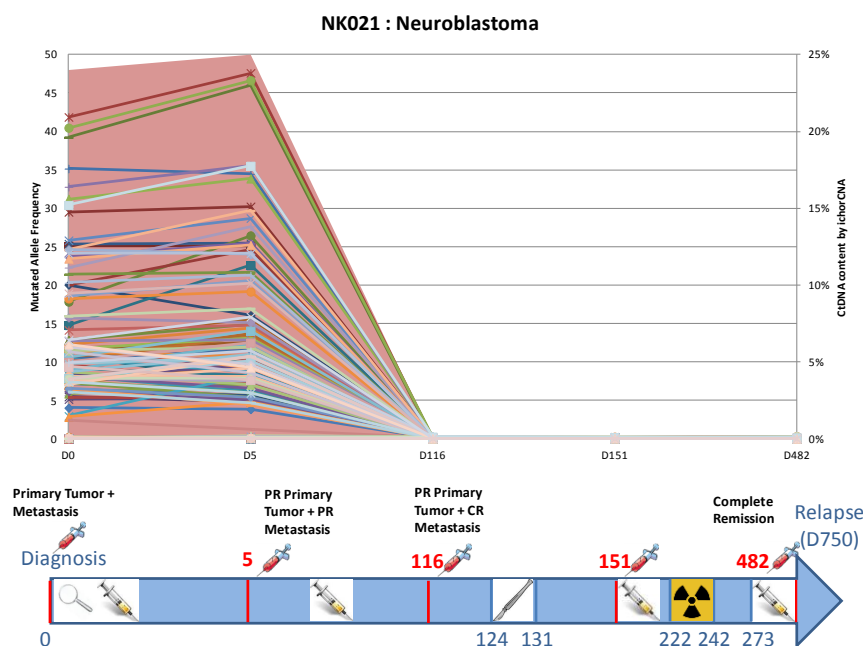


Figure 3.4.19 : Suivi du cas 21, patient atteint de neuroblastome de stade 4 possédant une forte hétérogénéité de variants dans la tumeur et le plasma.

Pour le cas 20, seul l'ADN de la tumeur était disponible au diagnostic. Les WES ont été réalisé sur des plasmas plus tardifs (J10 et J20) et n'ont pas permis de caractériser des altérations spécifiques à l'ADN circulant. Les SNVs de la tumeur ont pu tout même être utilisés pour un suivi dans les plasmas recueillis au cours du traitement. Une augmentation du niveau de ctDNA est ainsi observée à 20 jours du diagnostic mais est certainement dû à un relargage des cellules tumorales attaquées par la chimiothérapie, ce plasma étant prélevé le même jour qu'une prise de traitement (Figure 3.4.20).

Une progression osseuse a été observée 147 jours après le diagnostic et confirmé par une augmentation du ctDNA dans le plasma prélevé dans les 5 jours suivants. Malgré l'arrêt du suivi de plasma, le patient a bien répondu à son traitement (chimiothérapie, radiothérapie et immunothérapie anti-GD2) et est en rémission complète.

Pour ce cas, les fréquences alléliques des SNVs possédaient une dynamique similaire avec des fréquences de 1 à 5% à J20 et de 6 à 29% à J152 au moment de la rechute et n'ont pas permis de réaliser un modèle d'évolution clonale.

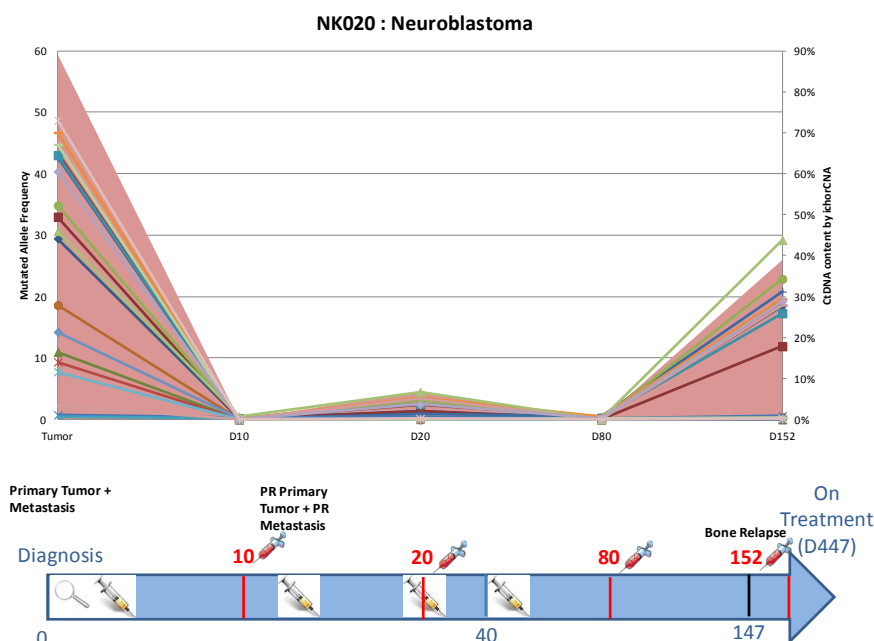


Figure 3.4.20 : Suivi du cas 20, patient atteint de neuroblastome ne possédant pas de plasma au diagnostic et avec un plasma au moment de la rechute.

La cohorte NGS Kids faisait partie d'une étude où les résultats de séquençage n'étaient pas utilisés dans la prise en charge du patient afin de proposer une thérapie ciblée. Cependant, un patient atteint de neuroblastome de stade 4 ne répondait pas aux traitements conventionnels malgré un séquençage de l'exome, de l'ARN et également l'utilisation d'un panel sur la tumeur, aucune cible thérapeutique n'a pu être détectée. La maladie n'a jamais répondu aux traitements nonobstant les changements de chimiothérapie, Rapid-Cojec, Cav à J53, Beacon à J90 et Pazopanib à J241 (Figure 3.4.21).

Les patients étant pseudonymisés pour l'étude NGS Kids, les rendus WES sur l'ADN circulant ont été demandés par la plateforme de diagnostic avec une demande de signalement de la présence d'une mutation d'*ALK* (F1174L), hotspot dans les neuroblastomes permettant d'inclure les patients dans un protocole visant à proposer une thérapie ciblée. L'ADN circulant a donc été utilisé avec un panel diagnostique afin de valider cette mutation selon les standards de détection de la plateforme de diagnostic (19% dans le WES et 20% dans le panel) et a permis de proposer un traitement adapté.

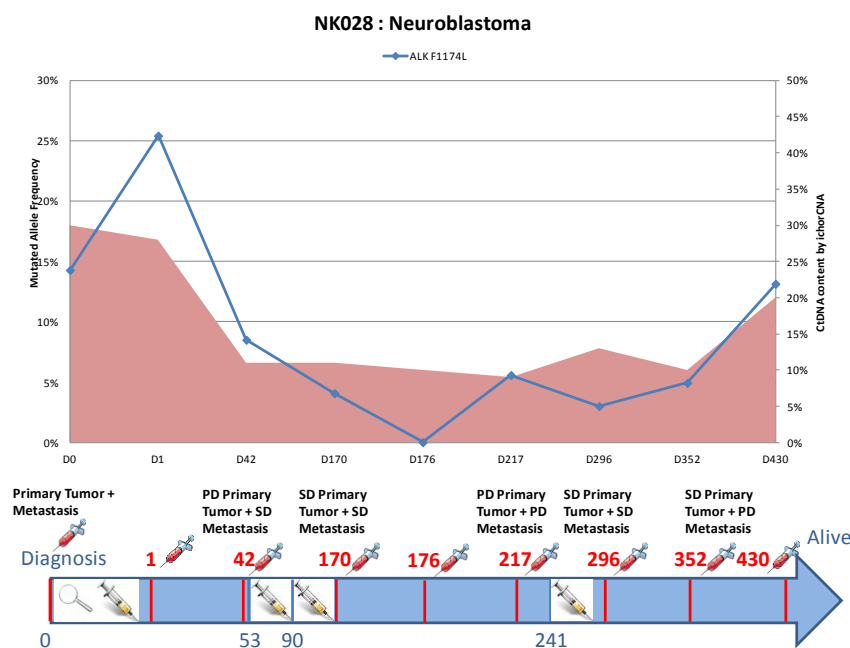


Figure 3.4.21 : Suivi du cas 20, patient atteint d'un neuroblastome présentant une mutation *ALK* non détectée dans la tumeur mais identifiée dans le plasma.

Le patient a subi plusieurs cycles de chimiothérapie sans réponse.

L'utilisation du cfDNA avec un panel diagnostique a permis de mettre en évidence une mutation *ALK* hotspot, le patient a pu être traité avec une chimiothérapie anti *ALK* (Lorlatinib).

Le patient a reçu un traitement Lorlatinib ciblé (anti *ALK*) et est en rémission partielle après 3 cycles de chimiothérapie et observation par scintigraphie. Le suivi de la mutation *ALK* F1174 dans l'ADN circulant avant l'administration de Lorlatinib a ainsi permis de voir des réponses seulement temporaires aux différentes chimiothérapies.

Pour le cas 2, un WES a été réalisé sur le plasma au diagnostic et à la rechute. Quatre dynamiques de mutations ont pu être ainsi observées (Figure 3.4.22) :

- des mutations présentes à une fréquence similaire entre le diagnostic et la rechute,
- des mutations apparaissant à la rechute,
- des mutations disparaissant à la rechute,
- et des mutations présentes à la rechute à un niveau sous-clonale révélées par le séquençage ciblé à haute profondeur.

Chacune de ces dynamiques peuvent être inférée à un clone. L'apparition de clone ou l'augmentation de la population pour certains pourraient être responsable de la rechute et de sa résistance aux traitements.

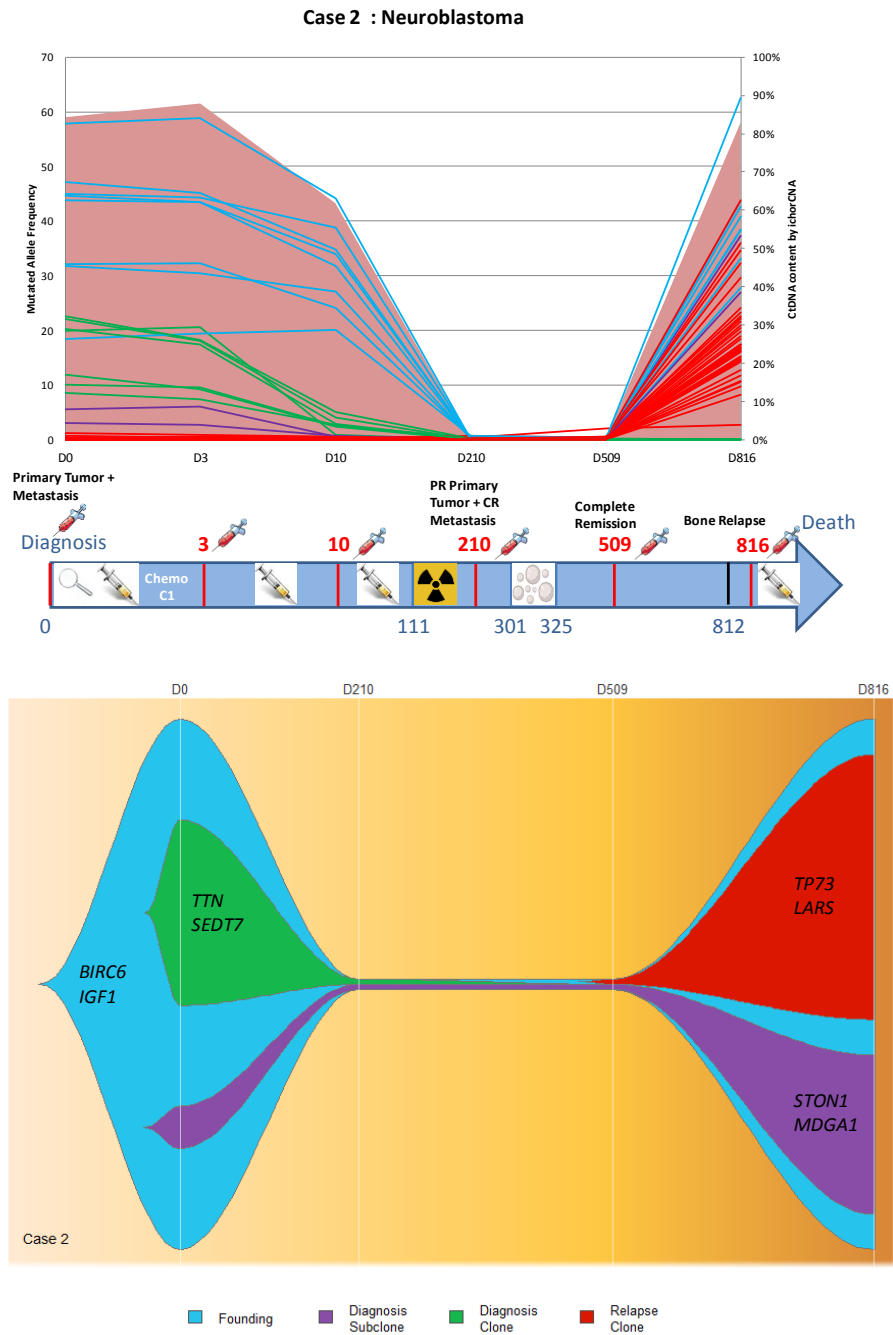


Figure 3.4.22 : Suivi du cas 2 avec une rechute (en haut) et modélisation de l'évolution clonale (en bas).

La dynamique des fréquences alléliques communes entre certains SNVs a permis de modéliser des clones et leurs évolutions.

En bleu, des SNVs présents au diagnostic et à la rechute certainement contenu dans le clone ancestrale (founding).

En vert, des SNVs présents seulement au diagnostic pouvant être inférés à un clone présent diagnostique.

En rouge, des SNVs apparaissent à la rechute et qui ont été associés à un clone présent spécifiquement à la rechute et dérivé du clone ancestral

En violet, des SNVs sont présents à un niveau sous clonal au diagnostic (<5%) et à un niveau clonal à la rechute.

Pour le cas 8, le diagnostic initial était un ganglioneuroblastome de stade INSS 2, avec par la suite une rechute métastatique un envahissement important de la moelle osseuse. Après un long traitement et une rémission partielle, la maladie a une nouvelle fois progressé conduisant au décès du patient.

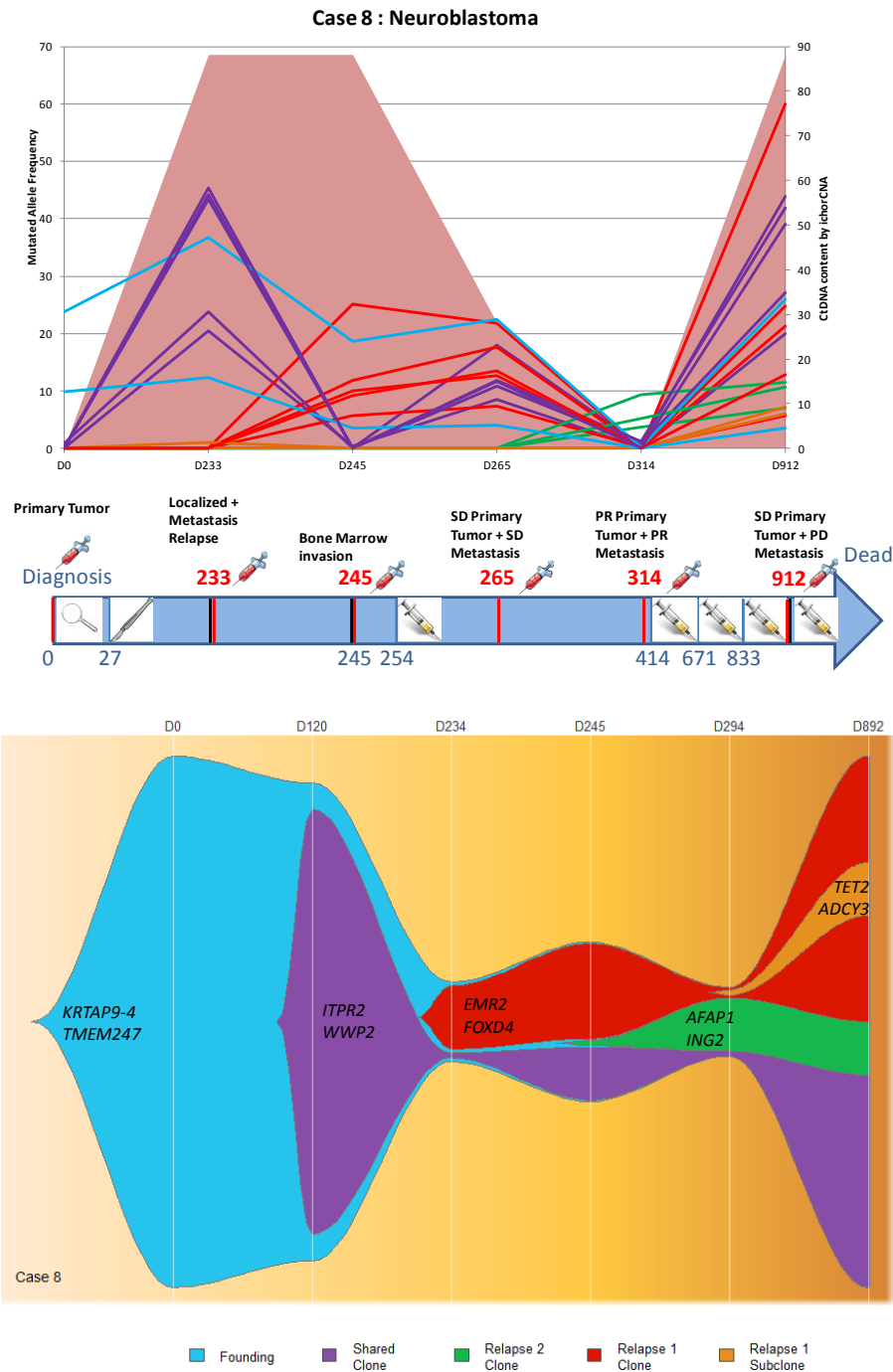


Figure 3.4.23 : Suivi du cas 8. Pour plus de lisibilité seulement quelques SNVs ont été sélectionnés mais pour chaque dynamique (couleur), une dizaine de SNVs étaient impliqués

Suivi du contenu tumoral (en haut) et modélisation de l'évolution clonale (en bas).

Chaque couleur est associée à une dynamique de SNVs et à un clone.

Cette résistance aux traitements semble être due la progression et l'apparition de nouveaux clones apparaissant à certains stades de progressions de la maladie. Ainsi les WES réalisés sur 2 plasmas recueillis lors de 2 rechutes couplés au séquençage ciblé ont permis d'observer l'apparition de mutation à différents moments du suivi et ont pu être inférées à différents clones (Figure 3.4.23).

Lors de la première rechute, un nouveau clone est apparu suivi d'un autre observé au moment de l'apparition des métastases osseuses. Ces clones semblent être sensibles aux traitements à J314 avec une rémission partielle de la maladie. Cependant, l'apparition d'un nouveau clone résistant aux traitements pourrait être responsable de la résurgence des autres clones et amener à la deuxième rechute. De plus, un autre clone est également présent et dérivé de la deuxième rechute. Ce dernier clone semble lié à une progression ostéomédullaire de la maladie conduisant au décès du patient.

En conclusion, le séquençage ciblé des plasmas de patients atteints de neuroblastome a permis de confirmer l'hétérogénéité spatiale et d'observer l'évolution clonale de la tumeur.

3.4.4.2 Rhabdomyosarcome

Seulement un cas de rhabdomyosarcome (type embryonnaire) n'a pas pu être suivi faute de détection d'altérations dans le plasma. Aucune donnée clinique (volume de la tumeur, âge du patient...) n'a permis d'expliquer cette absence de ctDNA dans le plasma.

Pour les 3 cas de rhabdomyosarcome alvéolaire, le ctDNA a pu être suivi dans les différents plasmas collectés au cours du suivi et avec une réponse positive aux traitements. Le séquençage ciblé a permis de suivre les mutations à des fréquences faibles jusqu'à 0,5% et d'observer une réponse rapide aux traitements (Figure 3.4.24 et Figure 1 Annexe). A l'instar des cas de neuroblastome, une augmentation des quantités de ctDNA ou de la fraction allélique de certaines mutations est observée dès le début du traitement (2 et 3 jours). Des dynamiques différentes sont observées entre les mutations mais leur faible nombre ne permet pas d'établir de modèles d'évolution clonale.

La totalité des cas alvéolaires est en rémission complète et pourraient être corrélées avec une réponse rapide aux traitements.

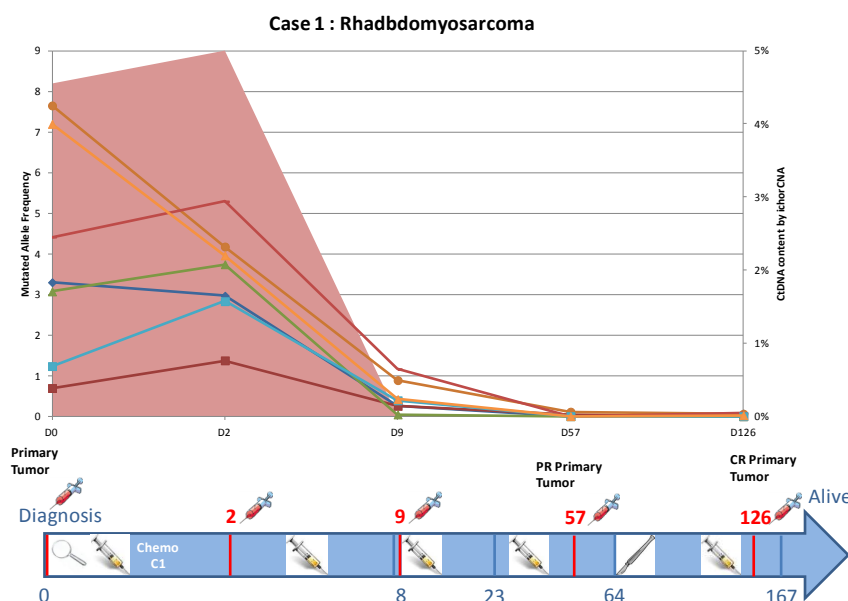


Figure 3.4.24 : Suivi d'un patient atteint rhabdomyosarcome alvéolaire (cas 1).

Pour le cas 3, les SNVs considérés comme exclusifs à la tumeur ou au plasma au diagnostic sont retrouvés dans les deux types d'échantillons avec des fréquences faibles. Par exemple, la mutation dans le gène *VRTN* (c.G481A) détectée à 25% dans le WES de la tumeur et 0% dans le plasma possède une fraction d'allèle muté (MAF) de 21,9% dans le séquençage ciblé de la tumeur et de 1,4% dans le plasma. De même pour une mutation exclusive au plasma dans le gène *GLP2R* (c.G967A) détectée à 14% est retrouvé à 15,6% dans le plasma et 0,29% dans la tumeur en séquençage ciblé. Cependant la mutation dans le gène *BRAF* (c.A1742G), propre à la tumeur, n'est pas retrouvée dans le plasma. En effet sa fréquence faible dans la tumeur (9% dans le WES et 5,8% dans le Target) peut laisser penser que la sensibilité du Target n'est pas suffisante pour sa détection. Cette mutation sous clonale aurait pu être responsable d'une éventuelle rechute.

Certaines altérations possèdent une dynamique différente entre elles avec des diminutions très précoces pour certaines (J1), d'autres avec des augmentations à J1 puis une diminution (Figure 3.4.25).

Ces dynamiques pourraient être dues à un effet clonal mais le faible nombre de mutations pour certaines dynamiques n'a pas permis de modéliser cette évolution.

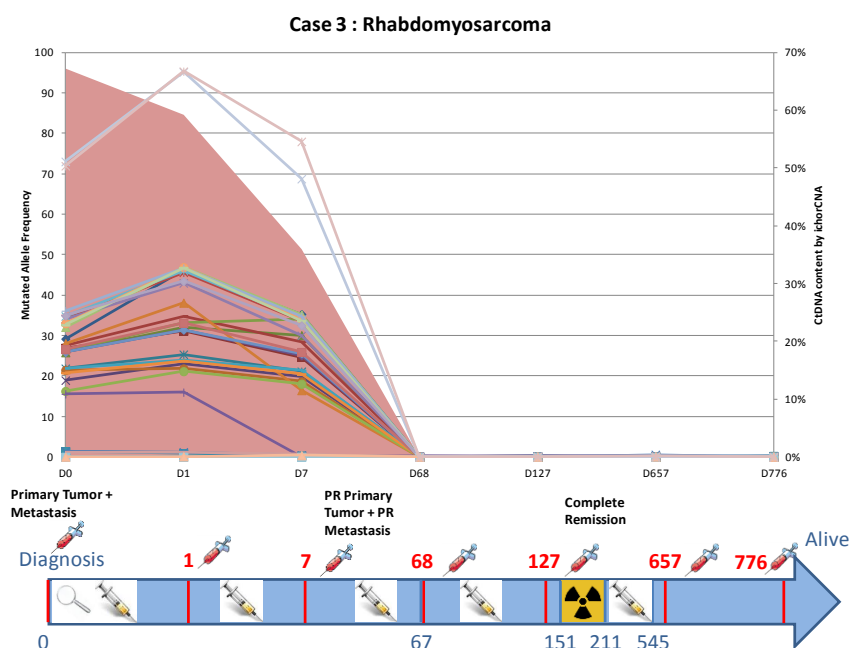


Figure 3.4.25 : Suivi du cas 3, patient atteint d'un rhabdomyosarcome embryonnaire avec réponse positive aux traitements.

Seul le cas 12 de rhabdomyosarcome embryonnaire a permis de modéliser une évolution clonale avec un WES réalisé sur le plasma au diagnostic et à la rechute. Certaines mutations sont ainsi présentes au diagnostic et à la rechute alors que d'autres étaient exclusives au diagnostic ou à la rechute (Figure 3.4.26). Ainsi le clone primaire reste présent lors de la rechute mais un clone spécifique au diagnostic est éliminé par les différents traitements. Un nouveau clone apparaît qui est probablement responsable de la rechute et malheureusement du décès du patient.

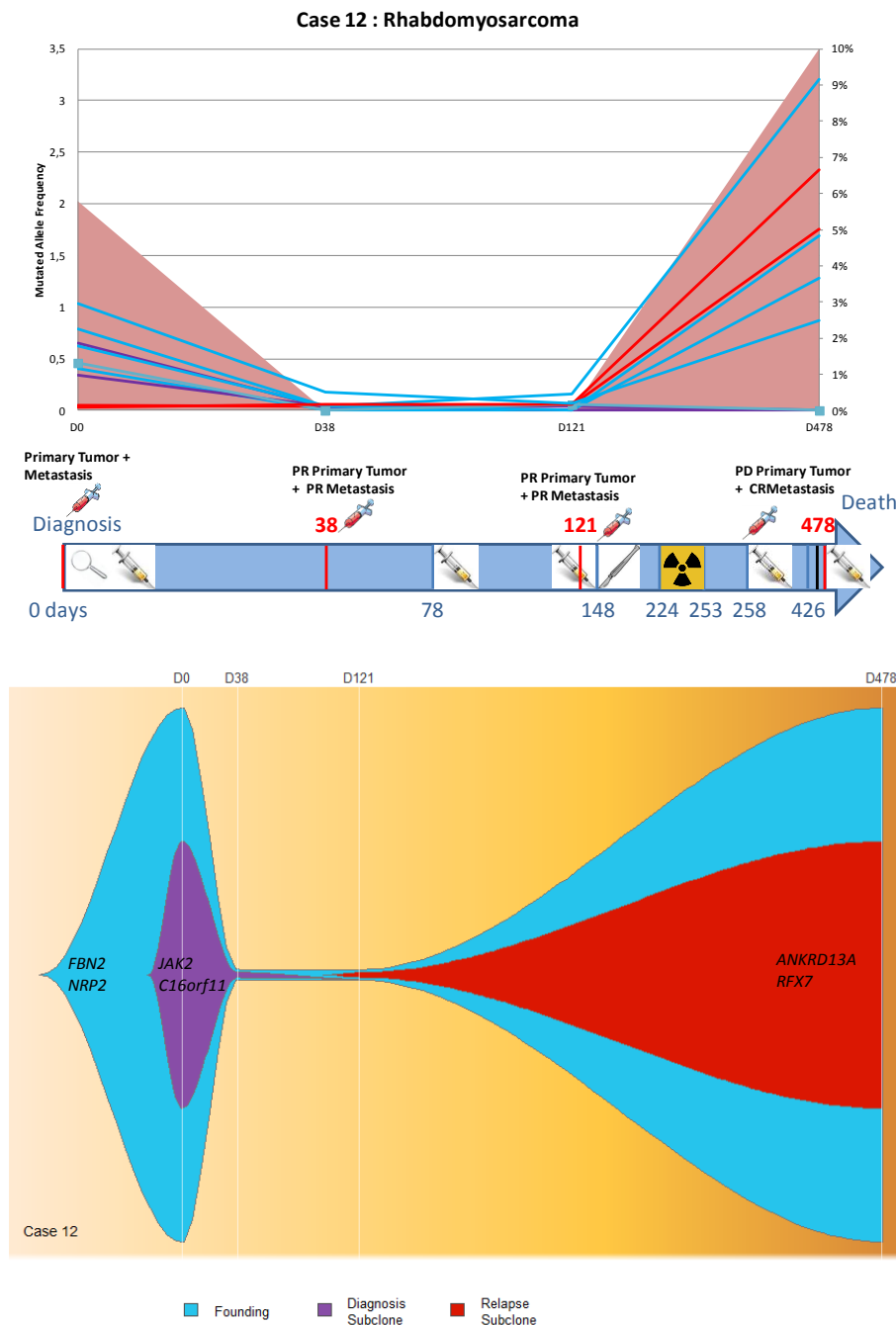


Figure 3.4.26 : Suivi du cas 12, patient atteint d'un rhabdomyosarcome embryonnaire avec une rechute (en haut) et modélisation de l'évolution clonale à partir des dynamiques mutationnelles (en bas).

3.4.4.3 Sarcome d'Ewing

Pour les deux cas d'Ewing métastatiques, le suivi n'a pas été possible car les SNVs pourraient être des SNPs malgré un filtre sur l'ADN constitutionnel des patients. Ces altérations restent à des fréquences stables dans le temps et le contenu tumoral n'a pas pu être calculé faute d'altérations en nombre de copie.

Le cas 6, patient atteint d'un sarcome d'Ewing localisé a très vite réagi aux traitements. Le patient était en rémission partielle au bout de 64 jours de traitement avec une diminution de la tumeur. Une diminution de la fréquence des mutations et du contenu tumoral a pu seulement être détectée à partir du deuxième jour pour disparaître au jour 64 pendant le quatrième cycle de chimiothérapie. Même après un traitement par radiothérapie, le contenu tumoral n'a pas été mis en évidence dans le plasma bien que le patient soit en rémission partielle (Figure 3.4.27).

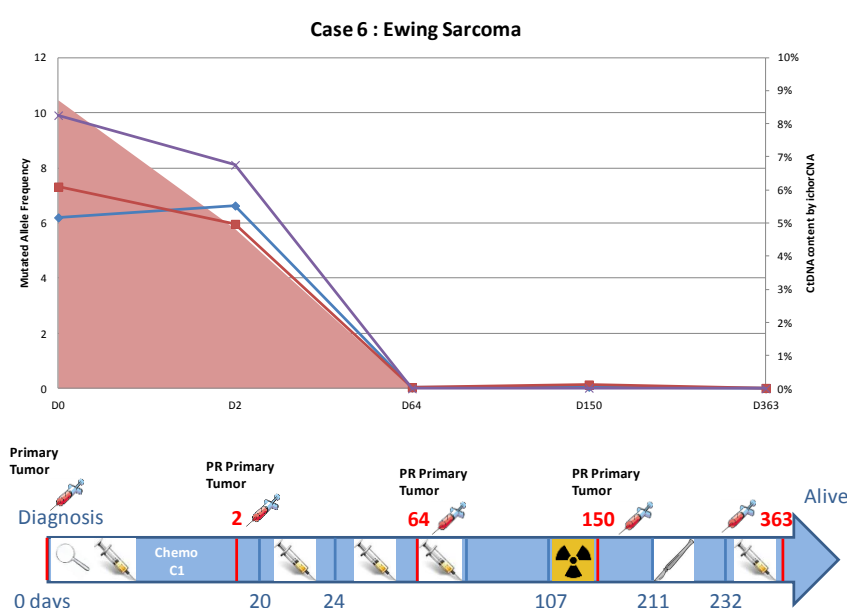


Figure 3.4.27 : Suivi du cas 6, patient atteint de sarcome d'Ewing.

3.4.4.4 Tumeur du foie

Pour le patient atteint d'un d'hépatoblastome, plusieurs clones ont été identifié dont un pour la tumeur primaire et deux pour les métastases. Les clones métastatiques semblent être affectés plus rapidement par la chimiothérapie avec une diminution de la fréquence de leur mutation dans le plasma tandis que le clone de la tumeur primaire voit ses fréquences de mutation augmenter puis diminuer. Cette augmentation des clones tumoraux dans le plasma est également observée par le WGS Low coverage. En effet, si le profil d'altérations en nombre de copie reste le même, le contenu sous-clonal déterminé par ichorCNA évolue. Avec un contenu tumoral de 31% dans le plasma, la part sous-clonale est de 53% et passe à 68%

pour un contenu tumoral de 22% pour le deuxième timepoint. Cette part sous-clonale augmentant alors que le ctDNA dans le plasma diminue, permet de confirmer la dynamique différente des clones tumoraux par rapport aux clones métastatiques (Figure 3.4.28).

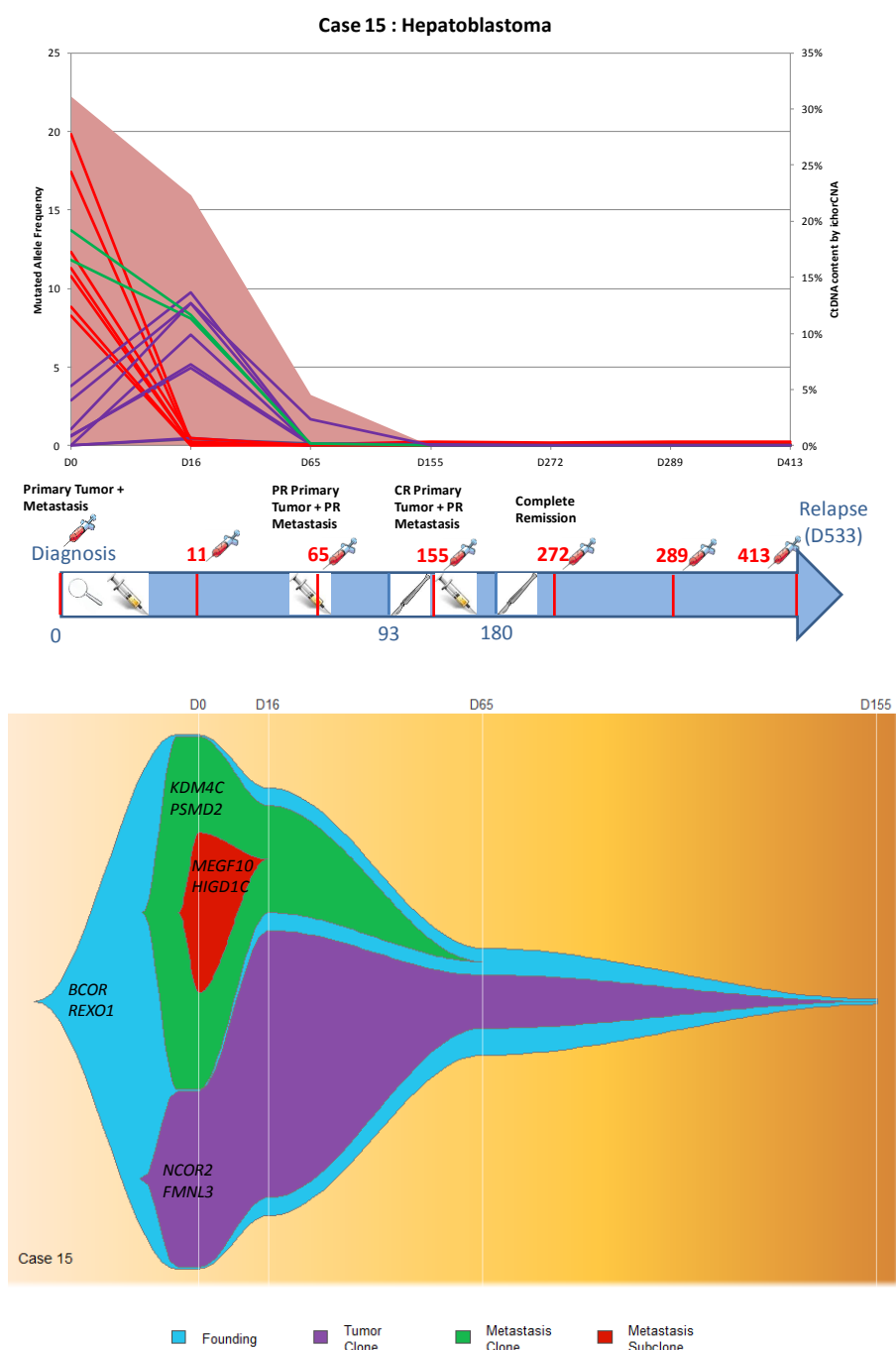


Figure 3.4.28 : Suivi du cas 15, patient atteint d'un hépatoblastome (en haut) et modélisation de l'évolution clonal à partir des dynamiques mutationnelles (en bas).

3.4.4.5 MPNST

Pour les deux cas de MPNST, le suivi a été court de 7 jours à 1 mois avec peu de plasma intermédiaire. La maladie était stable, le contenu de ctDNA dans le plasma ne descendant jamais en dessous de 11% et de même pour les mutations dont une seule varie de plus de 5%.

Le contenu de ctDNA n'a donc pas diminué pendant le traitement et les patients n'ont malheureusement jamais répondu malgré de nombreuses chimiothérapies, radiothérapies et une amputation de la jambe pour le cas 22, 15 jours après le diagnostic (Figure 3.4.29 et Figure Annexe 1).

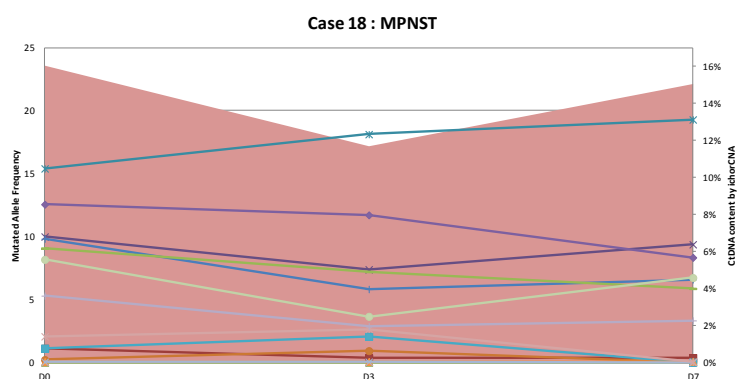


Figure 3.4.29 : Suivi de cas 18, patient atteint de MPNST corrélés avec des données cliniques de maladie stable ou progressive.

3.4.4.6 Tumeurs cérébrales

En dehors de la confirmation au diagnostic de la mutation *TP53* détecté dans le cas 24 atteint d'un gliome des voies optiques, le séquençage Target en haute profondeur n'a pas permis d'observer des altérations tumorales de tumeurs cérébrales dans le sang.

3.4.5 Conclusion

L'ADN circulant est un formidable outil pour suivre l'évolution clonale des tumeurs. Le fait de combiner les techniques de WES et de séquençage ciblé a permis de détecter et de suivre un maximum d'altérations et ainsi que leurs apparitions en cas de rechute.

Les outils de caractérisation clonale tels que ClonEvol, utilisant des WGS, sont très peu fonctionnels sur les données d'exomes^{141,394}. Un clone a donc été inféré "manuellement" selon une dynamique de mutations³⁵⁵. Une adaptation des outils existant prenant en compte le WES et le séquençage ciblé sera donc nécessaire pour une meilleure compréhension. Cette adaptation a été possible chez l'adulte pour caractériser la clonalité d'un prélèvement tumoral mais un nombre important d'altérations avaient été détectées par WES (574 mutation en moyenne)³⁹⁵.

Pour chaque patient avec une rechute, de nouvelles altérations ont été détectées synonyme d'apparition d'un ou de nouveaux clones. Le séquençage avec un large spectre type WES est donc indispensable lors d'une progression de la maladie afin de pouvoir détecter de nouvelles altérations. Ainsi, la dynamique des clones semble directement responsable de l'échappement aux traitements.

Le séquençage ciblé a également conduit à l'observation d'une augmentation des quantités de ctDNA dans le sang dans les premiers jours de chimiothérapie. Cette élévation pourrait être due à l'apoptose de cellules tumorales, synonyme d'efficacité du traitement.

Pour chaque cas, des différences sont observées dans la dynamique des variants tumoraux dès la première chimiothérapie. Ces dynamiques pourraient provenir de clones ayant des sensibilités différentielles aux traitements par chimiothérapie.

Cependant la caractérisation voire l'évolution d'une tumeur ne peut se limiter aux altérations génétiques. L'étude du profil d'expression et d'épigénétique d'une tumeur est également nécessaire et pouvant refléter la plasticité cellulaire.

3.5 Utilisation de l'ADN circulant pour l'étude du transcriptomique, de l'épigénétique et de la plasticité

Une plasticité des cellules tumorales pourrait être également à l'origine de l'échappement au traitement. Le passage d'une cellule différenciée à un état indifférencié leur permet de mieux résister à certaines chimiothérapies et affecte leur mobilité ou leur croissance¹⁷⁷.

Néanmoins, ces transitions ont été observées seulement *in vitro* à l'aide de profils d'expression ou de profils épigénétiques. La possibilité de prélèvements successifs de plasmas contenant de l'ADN circulant pourrait donc permettre de confirmer ces adaptations *in vivo*.

3.5.1 Prédiction d'expression se basant sur l'empreinte des TSS dans l'ADN circulant

Afin d'étudier l'expression des gènes d'une tumeur par l'ADN circulant, l'isolation de l'ARN circulant reste complexe et son analyse globale telle que le RNAseq est souvent très bruitée. Cependant, de nouvelles approches pourraient permettre l'utilisation de l'ADN circulant pour inférer l'expression des gènes des cellules tumorales. En effet, le découpage de l'ADN circulant est fait selon les histones et l'ouverture de la chromatine pour la fixation d'un facteur de transcription au niveau du site d'initiation (TSS) de la transcription d'un gène pourrait rendre cette région vulnérable lors de la circulation dans le sang.

Lors de séquençage Whole Genome de l'ADN circulant, les TSS des gènes exprimés sont ainsi déplétés en couverture. La prédiction de gène exprimé ou non est donc possible avec la recherche d'empreinte de couverture et la présence de ctDNA pourrait nous informer sur l'expression d'une tumeur.

3.5.1.1 Design d'un panel

L'étude des TSS devrait être possible à partir d'un séquençage par WGS à couverture élevé. Néanmoins ces méthodes demeurant coûteuse, une approche de séquençage ciblé sur les TSS a été choisie. Pour cela les positions des TSS ont été extraites de la base de données UCSC et seulement les gènes avec un numéro d'ARN codant ont été conservés pour un total de 21140 TSS pour 18211 gènes³⁹⁶.

Un panel couvrant 200 bp de part et d'autres des TSS a ensuite été dessiné sur NimbleDesign pour contenir la région possiblement déplétée en nucléosome précédemment décrite³¹⁶.

La capture TSS mesurant 8,78 Mb, le protocole développé pour l'exome a été utilisé, plus adapté aux captures de grande taille. Les captures ont été réalisées sur 23 échantillons avec au moins un contrôle provenant d'une prise de sang d'un individu sain réalisée par l'EFS (Etablissement Français du Sang) et d'un patient en rémission complète. Chaque capture a été déposée sur une ligne de Flowcell d'HiSeq 2500 Rapid Run™ pour une couverture théorique de 100x.

Quarante-quatre patients de la cohorte NGS Kids et d'études rétrospectives (Tableau annexe 1) ont été sélectionnés suite à la détection d'un fort contenu de ctDNA dans leur plasma au diagnostic (mediane de 45%, contenu allant de 0% à 99%). Une majorité des patients étaient atteints d'un neuroblastome (n=42), auxquels s'ajoutaient à un patient atteint d'hépatoblastome et un patient d'un sarcome d'Ewing. Six paires de plasmas au diagnostic et à la rechute ont également été analysés. Pour 18 patients, un profil d'expression a été établi sur l'échantillon tumoral était disponible, effectué soit en RNAseq, soit par puces Affymetrix Clariom D™.

3.5.1.2 Pipeline bio-informatique

Une première approche de Machine Learning a été réalisée sur les contrôles sains présents dans chaque capture. L'outil a été entraîné à partir de 1484 gènes exprimés dans le sang et 2405 non exprimés selon la base de données GTEx³⁹⁷.

Les profils de couvertures ont été générés par l'outil deepTools et le modèle d'apprentissage a été construit avec Scikit-learn en langage Python^{398,399}.

Les profils de couverture ont également été normalisés avec le profil de couverture sur de l'ADN génomique pour réduire les biais de capture pouvant être dû à la présence de régions répétées ou riches en AT.

La précision, soit le nombre de gènes pertinents, était de 87% pour les gènes exprimés et 86% pour les gènes non exprimés. Le rappel (ou la sensibilité), c'est-à-dire le nombre de gènes pertinents sélectionnés une nouvelle fois par l'outil, était de 86% pour les gènes exprimés et de 88% pour les non-exprimés.

L'outil a également été entraîné sur les gènes sélectionnés par l'équipe de Ulz *et al.*³¹⁶. Etabli sur la base de données Fantom5, le set de gènes contenant des gènes de ménages (n=3804) et des gènes non exprimés dans tous les tissus (n=670) n'a pas permis d'augmenter la sensibilité de l'outil (Figure 4.1.1)⁴⁰⁰.

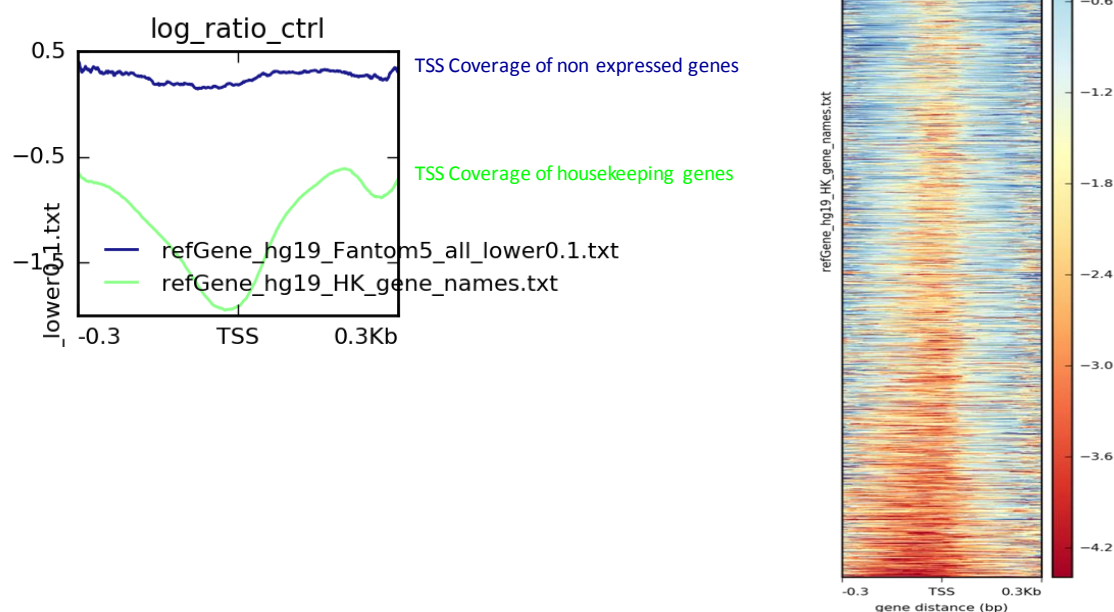


Figure 4.1.1 :Log ratio de la couverture autour du TSS de cfDNA issu d'un plasma sain pour les gènes de ménages (en vert) et non exprimés (en bleu) de la base de données Fantom5 (à gauche). Heatmap de la couverture autour du TSS pour les gènes de ménages (en bas) et non exprimés (en haut).

3.5.1.3 Prédiction

Le Machine Learning a été soumis à un premier échantillon de cfDNA avec un fort contenu en ctDNA et avec un profil d'expression disponible pour la tumeur (cas 66, 85% de ctDNA). La déplétion de couverture était observée pour les gènes prédits comme étant exprimés et, à l'inverse, était absente pour les gènes prédits comme non exprimés. En comparaison avec le RNAseq de la tumeur, les gènes prédits exprimés possédaient une expression plus forte que ceux prédits non exprimés validant ainsi cette approche (Figure 4.1.2).

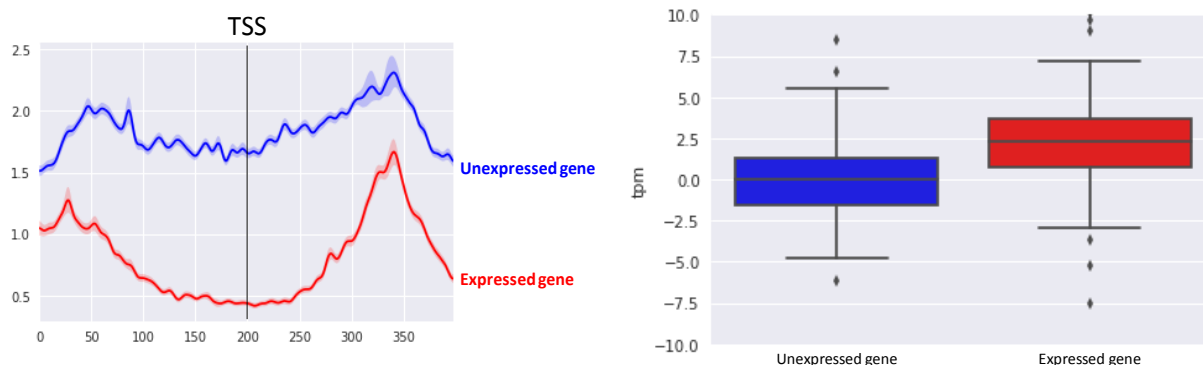


Figure 4.1.2 : Couverture autour du TSS des 2000 gènes les plus exprimés (en rouge) et des 2000 les moins exprimés (en bleu) dans le cfDNA pour le cas 66 (à gauche). TPM (Transcripts Per Kilobase Million) issus du RNASeq de la tumeur pour les gènes prédits exprimés (en rouge) et prédits non exprimés (en bleu).

Les sorties de l'outil de prédiction sont en cours d'analyse pour les autres échantillons. Les fichiers intermédiaires, notamment les profils de couverture générés par deepTools, ont permis d'observer de façon préliminaire la présence de déplétions ou non pour des gènes impliqués dans le neuroblastome et différents des empreintes des cellules hématopoïétique reflet du cfDNA.

Compte tenu de la présence d'une amplification de *MYCN* dans plus de la moitié de la cohorte, une couverture beaucoup plus forte pour le TSS du gène *MYCN* a donc été observée pour les cas "amplifiés". Les profils de couverture montrent la présence d'une déplétion quand le gène *MYCN* est amplifié suggérant une expression du gène (Figure 4.1.3). Les cas "non amplifiés" ne présentent pas cette déplétion nette et donc cette surexpression de *MYCN*. L'analyse individuelle des cas "non amplifiés" n'a pas permis d'observer une éventuelle surexpression du gène, la surexpression de *MYCN* est plutôt rare en absence d'amplicon^{401,402}.

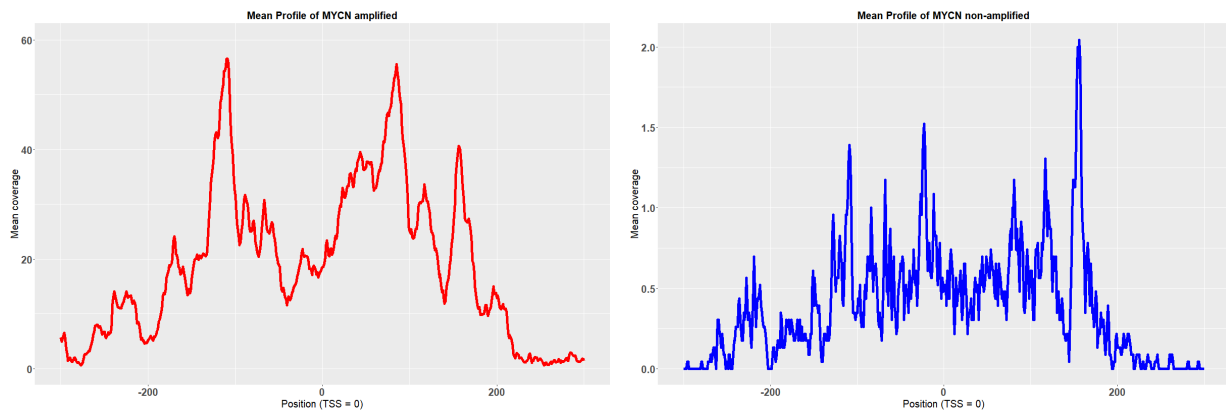


Figure 4.1.3 : Couverture autour du TSS de *MYCN* pour les cas avec amplification du gène (n=20, à gauche et en rouge) ou non (n=15, à droite et en bleu). Un profil différent est observé entre les échantillons *MYCN* amplifiés et *MYCN* non amplifiés.

La présence de l'amplification de *MYCN* conduit à une dérégulation du transcriptome de la tumeur. L'analyse en composante principale des profils de couvertures de l'intégralité des gènes entre *MYCN* amplifiés ou non amplifiés permet de séparer les deux groupes suggérant une variabilité significative (Figure 4.1.4). Cette variabilité pourrait être due aux gènes cibles de *MYCN* tels que *MDM2* ou *MYC* et également certains gènes responsables de la plasticité du neuroblastome tels que *GATA3* ou *NOTCH2*^{176,403}.

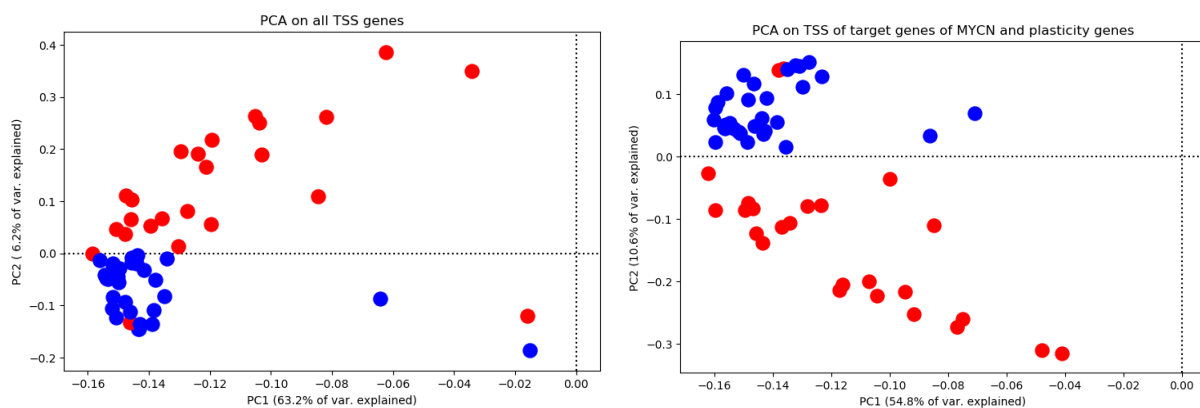


Figure 4.1.4 : Analyse en Composante Principale (ACP) des profils de couverture de TSS pour les cas "*MYCN* amplifié" (en rouge) et "*MYCN* non amplifiés" (en bleu). A gauche, l'ACP sur l'intégralité des TSS et à droite sur les TSS des gènes cibles de *MYCN* et de la plasticité.

Cependant l'analyse individuelle de ces gènes (2125 pour les gènes cibles de *MYCN*, 854 pour la plasticité) demeure inachevée et n'a pas permis d'identifier les gènes responsables de cette variabilité.

D'autres gènes tels que *ALK*, *ATRX* et *TERT* présentaient des profils différents entre patients. Après normalisation avec la couverture de la tumeur, *ALK* et *TERT* étaient prédits comme non exprimé dans le contrôle sain avec un pic au niveau du TSS suggérant la présence d'un nucléosome contrairement à *ATRX* où la zone du TSS était déplétée (Figure 4.1.5). Ces prédictions étaient en accord avec les données d'expressions du sang^{397,404}. Le gène *ALK* est prédit comme exprimé dans la majorité des cas de neuroblastome et prédit comme non exprimé pour certains cas de neuroblastome mais également pour le cas d'Ewing et d'hépatoblastome. De même pour le gène *ATRX* qui reste majoritairement exprimé.

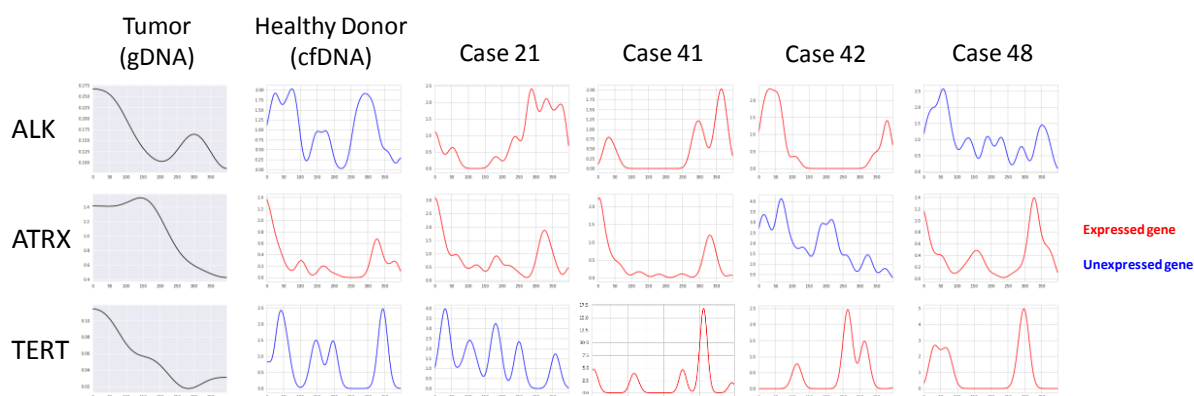


Figure 4.1.4 : Couverture autour des TSS de gènes impliqués dans le Neuroblastome tels qu'*ALK*, *ATRX* et *TERT*. Dans le sang, seulement *ATRX* est exprimé contrairement à *ALK* et *TERT* ; les profils de couverture reflètent ces conditions. Pour les 4 cas de neuroblastomes, des prédictions différentes sont observés selon les gènes. Ainsi *ALK* est prédit exprimé dans 3 cas mais pas le cas 48. De même pour *ATRX* et *TERT* qui ne sont pas prédits exprimés respectivement dans le cas 42 et 21.

L'outil d'analyse d'expression a rencontré plus de difficulté avec le gène *TERT* qui semble exprimé dans certaines tumeurs après analyse visuelle contrairement à sa prédiction et inversement. Comme pour la plupart des TSS, les régions de *TERT* sont très riches en régions répétées et pourraient être difficiles à couvrir.

Le gène *ALK* semble montrer une certaine dynamique dans les cas permettant une comparaison entre le diagnostic et la rechute. Pour le cas 46, malgré une prédiction d'expression de l'outil, le profil de couverture au diagnostic ressemble à un gène non exprimé contrairement à la rechute. Pour le cas 43, le gène *ALK* semble réprimé à la rechute selon

l'outil, cependant le profil de couverture pour la rechute observé reste tendancieux (Figure 4.1.6).

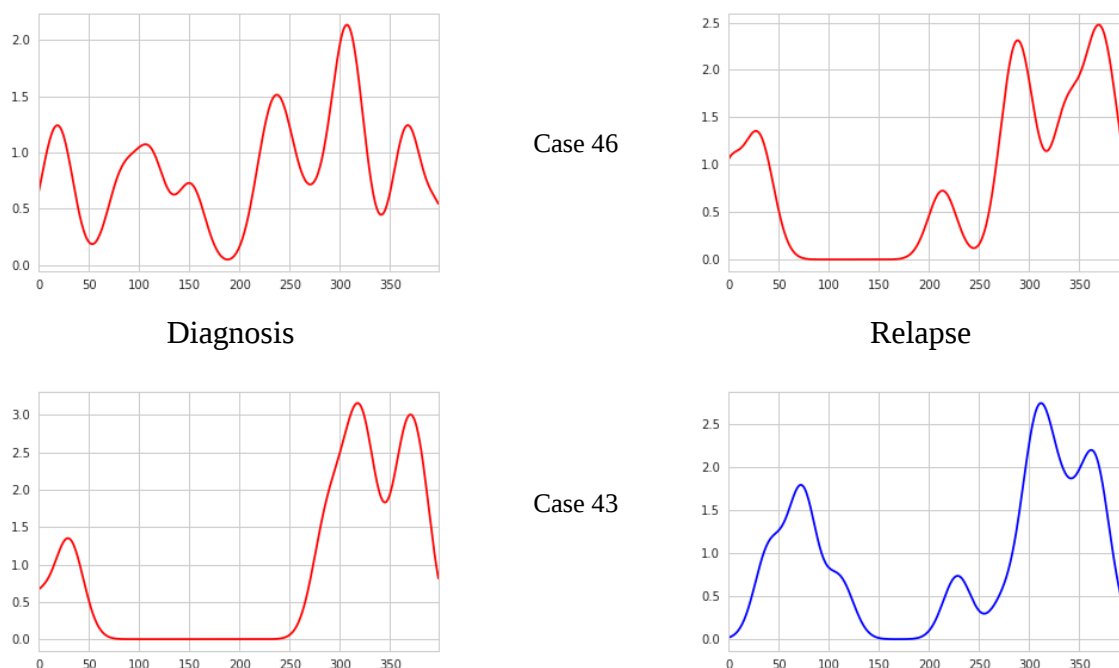


Figure 4.1.6 : Profils de couverture autour du TSS de *ALK* du cfDNA de plasma au diagnostic et à la rechute des cas 46 et 43.

L'observation manuelle de la totalité des profils étant impossible, l'optimisation du machine learning est nécessaire mais la prédiction d'expression de certains gènes tels que *ALK* impliqué dans le développement du neuroblastome prouve la possibilité d'avoir accès à l'expression de ces gènes d'intérêt dans ces tumeurs par l'ADN circulant. La recherche d'empreinte TSS pourrait donc être un formidable outil pour inférer l'expression des gènes d'une manière globale, et ainsi permettre l'identification d'une transition plastidique.

Par ailleurs, la plasticité des cellules tumorales peut également être observée par une étude épigénétique.

3.5.2 ChipSeq sur l'ADN circulant

Les transitions d'identité et de changement d'expression sont dues à des modifications épigénétiques par l'acétylation ou la méthylation des certaines histones responsables du compactage ou non de la chromatine¹⁶⁰.

L'hypothèse d'ADN circulant enroulé autour d'un nucléosome a été émise suite à des similarités entre le découpage de l'ADN circulant et de l'ADN par MNase^{229,230}. L'enroulement autour du nucléosome permettrait ainsi à l'ADN circulant d'être protégé des enzymes et de circuler dans les fluides. Cependant sa structure n'a jamais été isolée mais théoriquement une immunoprécipitation des histones pourrait permettre de le récupérer. Peu d'études de ChIP (Chromatone Immunoprecipitation) ont été réalisées sur le plasma et ont été validées seulement par PCR après marquage de certaines histones telles que H3K9me3⁴⁰⁵.

3.5.2.1 Protocole de ChipSeq

L'iDeal ChIP-seq kit for Histones™ de Diagenode® a été utilisé contenant les anticorps H3K4me3, marque activatrice de gène, auquel ont été ajoutés les anticorps H3K27me3, marque répressive, et H3K27ac, marque d'enhancer.

L'ADN circulant étant déjà découpé les étapes de pontage (cross-link) et de sonication ont été défalquées. Sept patients atteints de neuroblastome de la cohorte NGS Kids ont été sélectionnés sur la base du taux de ctDNA dans leur plasma (médiane de 80% allant de 24% à 90% selon l'outil Facets), de la quantité de cfDNA (médiane 1008 ng/mL de plasma allant 40 à 1497 ng/mL de plasma) et de la disponibilité de plasma (1 mL). Pour un patient, un plasma au diagnostic et un plasma à la rechute ont été utilisés. Pour réaliser le ChIP, 200 µL de plasma ont été utilisé pour chaque marque d'histone et pour l'input permettant de contrôler l'enrichissement. Le protocole standard du constructeur a été suivi avec 2 µg d'anticorps pour les 3 marques d'histone.

Les PCR de contrôle ont montré des enrichissements très aléatoires et étaient parasités par des dimères d'amorces notamment pour l'Input. Les échantillons ChIP ont ensuite été préparés par le Kit TruSeq Chip d'Illumina selon le protocole du constructeur et séquencé sur Novaseq avec 3 ou 4 ChIP par unité de flowcell S1.

Le ChIP-Seq a ensuite été aligné avec Bowtie2 et contrôlé par deepTools. Quatre échantillons sur 8 possédaient des contrôles de qualité suffisants tels qu'un espacement entre les courbes de couverture d'Input et des marques d'histones reflétant un enrichissement de ces dernières ou la présence de reads couvrant les TSS pour les marques H3K4me3 ou H3K27me3 (Figure 4.2.1).

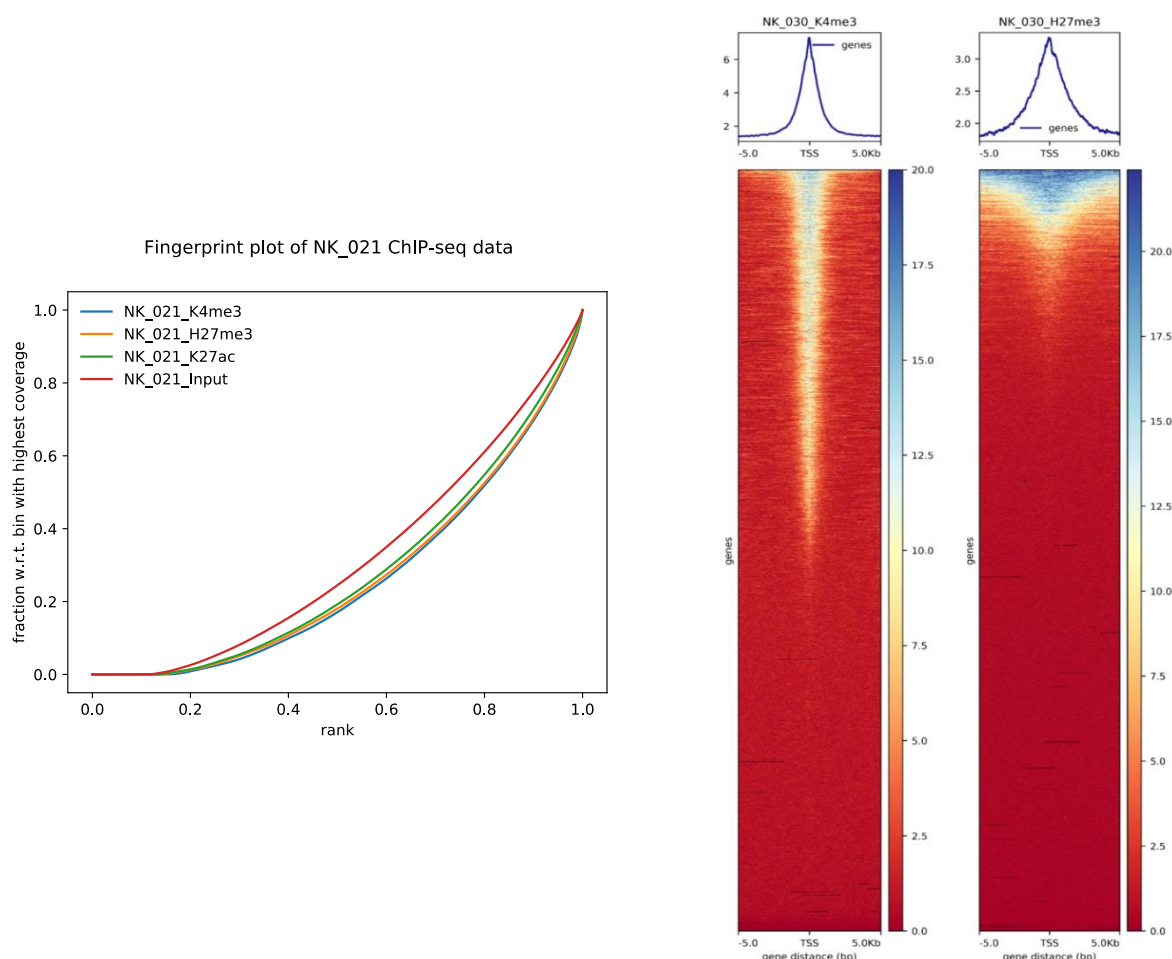


Figure 4.2.1 : Fingerprint plot des couvertures des différentes marques d'histones montrant un enrichissement de ces dernières par rapport à l'input du cas 21 (à gauche) et Heatmap des couvertures autour des TSS pour les marques H3K4me3 et H3K27me3.

Le "peak calling" a été peu informatif dû au bruit de fond très élevé mais certains gènes d'intérêt ont été détecté et confirmé par IGV.

3.5.2.2 Peak Calling

Le gène *CCND1*, considéré comme fortement exprimé dans le neuroblastome et non exprimé dans le sang (utilisé comme contrôle d'enrichissement de ChIP sur neuroblastome par PCR), possédait une marque activatrice, H3K4me3, pour 4 échantillons contrairement à l'antisens *CCND2-AS1* qui est totalement réprimé dans le neuroblastome et qui montrait une marque H3K27me3 (Figure 4.2.2)^{177,406}.

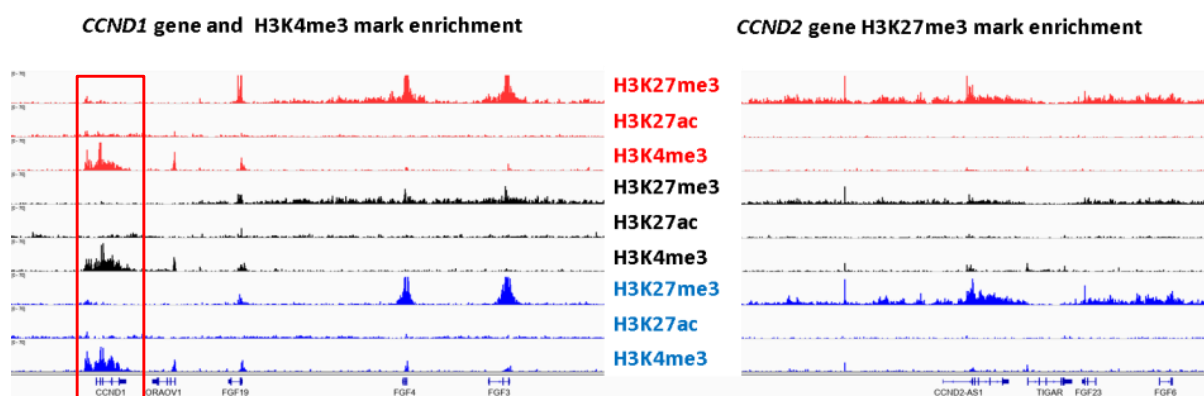


Figure 4.2.2 : Observation de "pic" selon les marques activatrices (H3K4me3) pour le gène *CCND1* (à gauche et encadré) et les marques répressives (H3K27me3) autour du gène antisens *CCND2-AS1* (à droite) dans le plasma de 3 cas de Neuroblastome (représenté par les différentes couleurs).

Contrairement aux résultats obtenus sur des lignées de cellules voire des tumeurs, la marque des enhancers H3K27ac n'était pas fonctionnelle sur les plasmas.

Une dizaine de gènes ont été décrit comme étant caractéristique des formes adrénérergiques tels que *GATA3*, *HAND2* ou *PHOX2B* et des formes mésenchymateuses tels que *PRRX1*, *NOTCH2* ou *SOX9*^{176,177}. Seules des marques activatrices représentatives des formes adrénérergiques ont pu être observées dans le plasma telles que *GATA3* et *HAND2* (Figure 4.2.3).

GATA3 and HAND2 genes and H3K4me3 mark enrichment

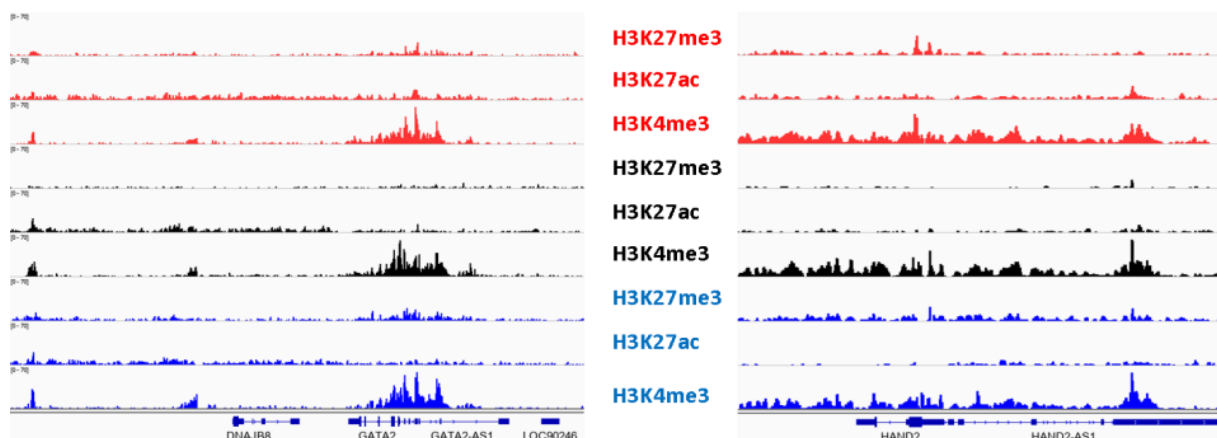


Figure 4.2.4 : Observation des "pic" pour les marques activatrices dans des gènes signatures de la forme adrénérergique tels que *GATA3* (à gauche) et *HAND3* (à droite).

Le gène *TWIST1* a été décrit avec un promoteur bivalent et comme étant une cible directe du facteur de transcription de MYCN^{407–409}. Cette bivalence semble jouer un rôle dans la reprogrammation de la cellule vers un état mésenchymateux. Elle a été observé dans 2 cas avec une marque H3K4me3 et H3K27me3 pour deux cas et seulement la marque activatrice pour un cas (Figure 4.2.4). La bivalence a notamment été décrite comme à bateau à voile ancré soumis au vent de par sa voile déployée contrairement à un promoteur réprimé où le bateau serait ancré et sa voile repliée⁴¹⁰. Les promoteurs bivalents seront ainsi prêts à partir pour transcription immédiate avec la levée (déméthylation) de l'ancre (H3K27).

TWIST1 showing both H3K4me3 & H3K27me3 mark enrichment

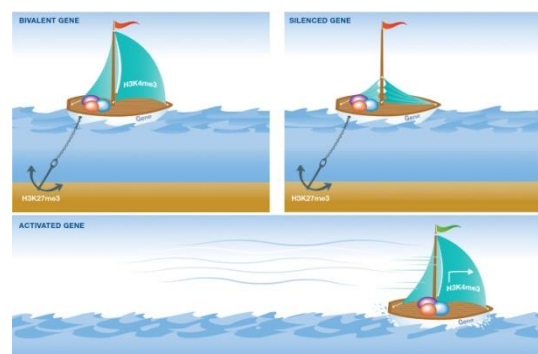
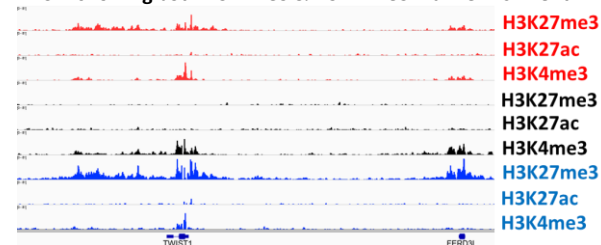


Figure 4.2.4 : Observation des "pic" dans le promoteur bivalent de gène *TWIST1* (à gauche) et concept du gène bivalent représenté par un bateau ancré sous le vent (à droite)⁴¹⁰.

Dans le neuroblastome, *TWIST1* co-occupe plus de 80% des enhancers de *MYCN* pour les cas amplifiés et conduit une augmentation de la prolifération^{71,407}. La totalité des cas sélectionnés ne possédaient pas d'amplification de *MYCN*, *TWIST1* pourrait peut-être jouer un rôle dans le neuroblastome en absence de dérégulation du gène *MYCN*.

L'analyse de marque épigénétique de la tumeur est donc possible à partir du plasma. La marque H3K27ac pour les enhancers n'a cependant pas fonctionné sur le plasma mais une seule quantité d'anticorps a testée. Les structures d'enhancer forment généralement une boucle de transcription permettant de l'accélérer et de souvent transcrire un cluster de gènes^{411,412}. Lors du découpage de l'ADN circulant, cette boucle doit certainement être détruite et donc la lysine 27 de l'histone 3 doit être désacétylée.

Les cas de neuroblastome présentaient des signatures plutôt adrénergiques avec des marques activatrices H3K4me3 sur des gènes tels que *GATA3* ou *HAND2*. Mais la présence de promoteurs bivalents impliqués dans la transition vers des formes mésenchymateuses tels que le promoteur *TWIST* pourrait amener à un changement d'identité éclair⁴¹³. Le cas avec seulement la marque activatrice pourrait donc être prémice au changement vers un état mésenchymateux. Cette rapidité de changement d'identité dans le neuroblastome pourrait lui permettre une adaptation véloce aux traitements tels que la chimiothérapie.

De plus, le gène *TWIST1* pourrait avoir un rôle dans la prolifération du neuroblastome en remplaçant l'activité de *MYCN* pour les cas "non amplifiés".

3.5.3 Conclusion

L'analyse de l'expression et de l'épigénétique de la tumeur par l'ADN circulant sont des méthodes très prometteuses.

La possibilité d'inférer l'expression d'un gène ou non selon l'empreinte présente au niveau de TSS pourrait permettre de classifier certains types de tumeurs dépendantes de l'activation/inactivation de certains gènes. Une fois le machine learning amélioré, cette méthode pourrait être un formidable outil dans le suivi de la plasticité de la tumeur et donc de sa possibilité d'échappement à un traitement suite à un changement d'identité.

Ces changements d'identité pourraient également être observés par une analyse épigénétique telle que le ChIPSeq réalisée sur l'ADN circulant. Ces premiers résultats montrent que non seulement le ChIPSeq est possible sur le plasma mais qu'il permet également d'avoir accès aux marques d'ouverture/fermeture de la chromatine de la tumeur.

3.6 Application Clinique de l'ADN circulant

3.6.1 Recherche d'altérations ciblables à la rechute à des fins thérapeutiques : Etude MappyActs

MappyActs (ClinicalTrials.gov Identifier: NCT02613962) est une étude internationale prospective de médecine de précision ayant pour objectif d'établir un profil moléculaire par WES et RNAseq des tumeurs réfractaires ou à la rechute chez les enfants et les adolescents, afin de permettre une orientation vers des essais cliniques précoces en fonction des altérations génétiques.

Jusqu'en 2019, 500 patients ont été recrutés et 433 tumeurs ont été analysées par WES et RNA-Seq et les résultats ont été soumis à un CMTB (Clinical Molecular Tumor Board) pour une discussion des altérations génétiques en vue d'inclusion dans des essais cliniques ciblés. Parmi ces patients inclus, 390 tumeurs étaient considérées contributives après une analyse WES et/ou RNAseq. Au total et chez 271 patients, au moins une altération génétique ciblable (548 altérations ciblables au total) a été identifiée par exemple dans des gènes impliqués dans les checkpoints du cycle cellulaire (n=97) ou dans la voie PI3K/AKT/mTOR (n=83).

En parallèle, une étude ancillaire a été réalisée sur le cfDNA des patients inclus dans MappyAct. Les plasmas des patients atteints de tumeurs solides extracérébrales prélevés à la rechute ont été sélectionnés pour analyse en WES permettant la recherche d'altérations ciblables (Figure 5.1.1).

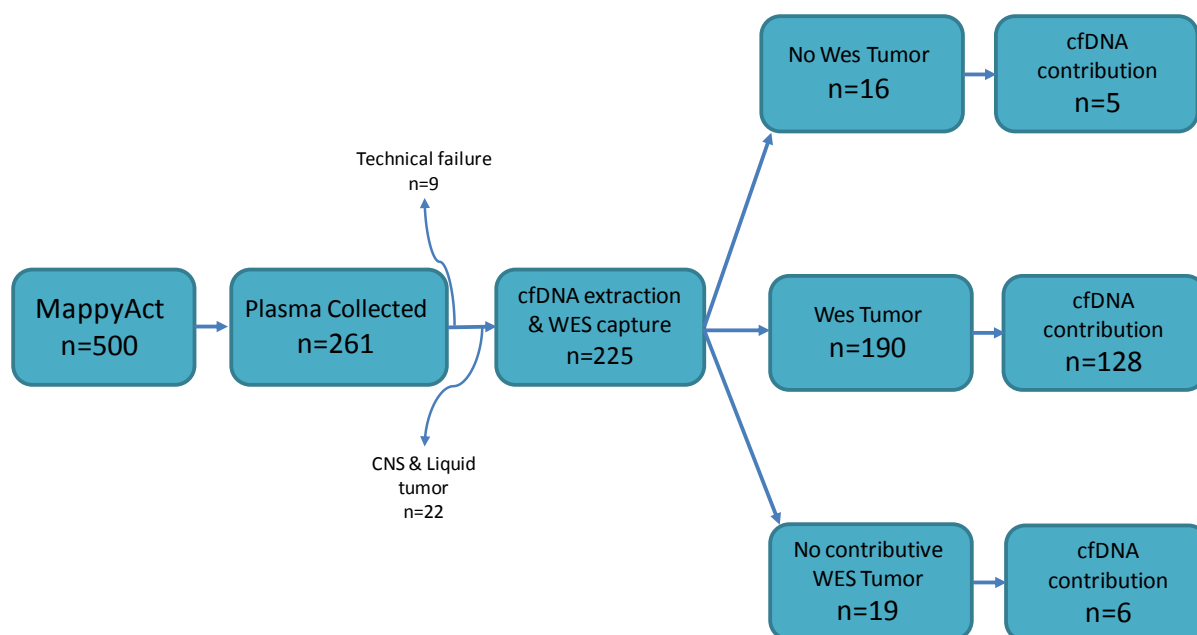


Figure 5.1.1 : Organigramme des résultats de WES sur l'ADN circulant des patients inclus dans MappyAct.

3.6.1.1 Protocole de WES

Le protocole de WES sur le cfDNA a été légèrement modifié par rapport aux études précédentes avec une utilisation du kit Kapa HyperPrep™ pour la construction des bibliothèques permettant une fusion des étapes de End Repair et de A-tailing et donc un gain de temps. La ligation étendue sur la nuit a été conservée mais avec des adaptateurs UDI et un ratio 20:1 permettant l'identification de 96 échantillons différents. Les sondes de capture ont été passées en Medexome™ qui malgré une taille de capture plus petite (48 Mb contre 64 Mb pour le SeqCap Exome V3™) donnait une couverture plus forte et une meilleure homogénéité. En fonction des séries, les exomes ont été séquencés sur Illumina Novaseq™ par des flowcell S1 pour 24 WES, des flowcell S2 pour 48 WES et des flowcell S4 pour 96 WES (Figure 5.1.1).

Le pipeline bio-informatique est resté similaire à celle précédemment utilisée. Cependant, les différentes technologies de captures d'exome Agilent Clinical Research Exome v2™ ou Twist Human Core Exome™ ont été utilisées lors du séquençage de la tumeur et l'ADN constitutionnel du patient et réalisées dans le cadre de la recherche de cibles

thérapeutiques contrairement à la technologie Roche Medexome™ pour le cfDNA et réalisée dans le cadre de l'étude ancillaire. De plus, les séquençages de ces captures ont été réalisés sur des machines différentes (Illumina HiSeq, NextSeq™ et NovaSeq™). Ces biais technologiques ont entraîné l'apparition de nombreux faux positifs. Ces erreurs apparaissant souvent dans les mêmes gènes voire les mêmes positions ont été exclues des analyses et inclus dans un fichier lorsqu'elles étaient présentes dans plus de 5 patients.

Par ailleurs, sept outils tels que SIFT, Polyphen ou MutationTaster ont été utilisés pour la prédiction des effets mutationnels. Un SNV a été considéré comme ciblable s'il était détecté dans un gène retenu par le CMTB et prédit comme délétère par 3 de ces outils et/ou avec un identifiant COSMIC^{345–348,414–417}.

3.6.1.2 Quantité de cfDNA et ctDNA

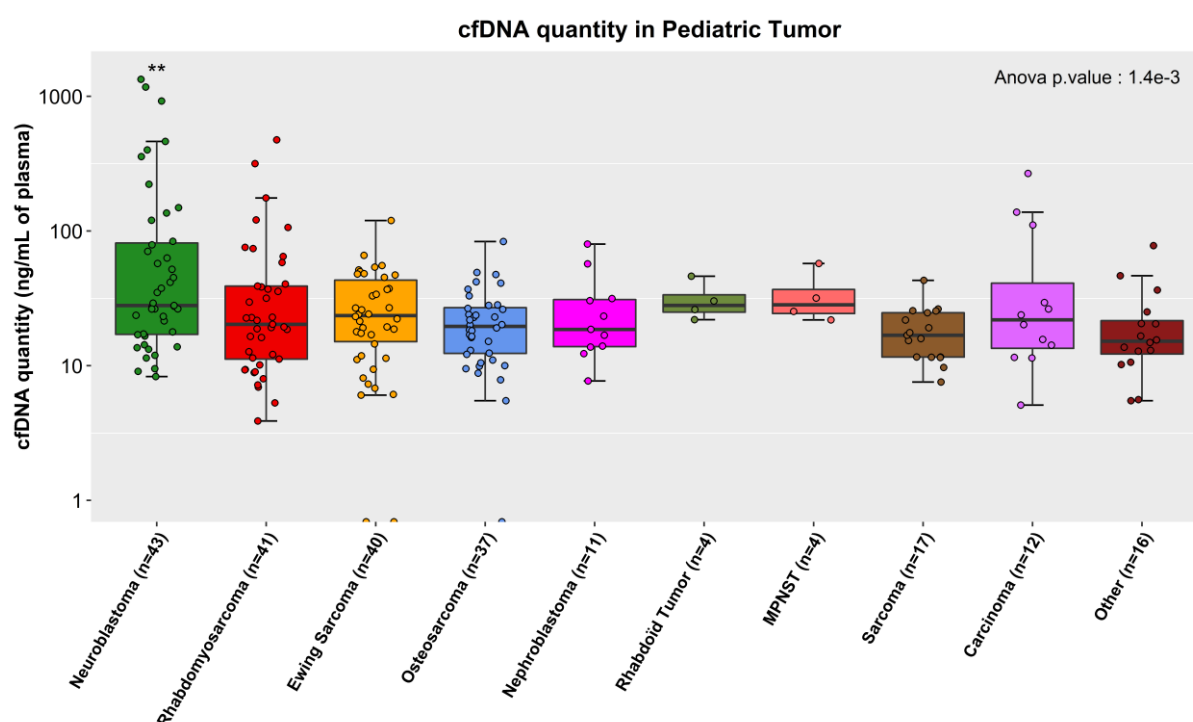


Figure 5.1.2: Quantité de cfDNA par type de tumeur (quantité en ng/mL de plasma, échelle log10). Les quantités d'ADN sont significativement différentes entre types de cancer (Test Anova : p.value = 1.4e-03), avec une forte différence pour les neuroblastomes (Test Wilcoxon = 4.1e-03).

L'ADN circulant a été extrait de 225 plasmas provenant de patients atteints de tumeurs solides extra-crâniennes lors de l'inclusion dans le protocole MappyActs (Tableau annexe 2).

Des différences significatives de quantité d'ADN circulant ont été observées entre chaque type de maladies (Anova p.value : $1,4e-03$). De plus, les quantités d'ADN circulant de patients atteints de neuroblastome (n=43) étaient significativement plus élevées que ceux des autres types de cancers pédiatriques (n=182) (Wilcoxon : $4,1e-03$) (Figure 5.1.2).

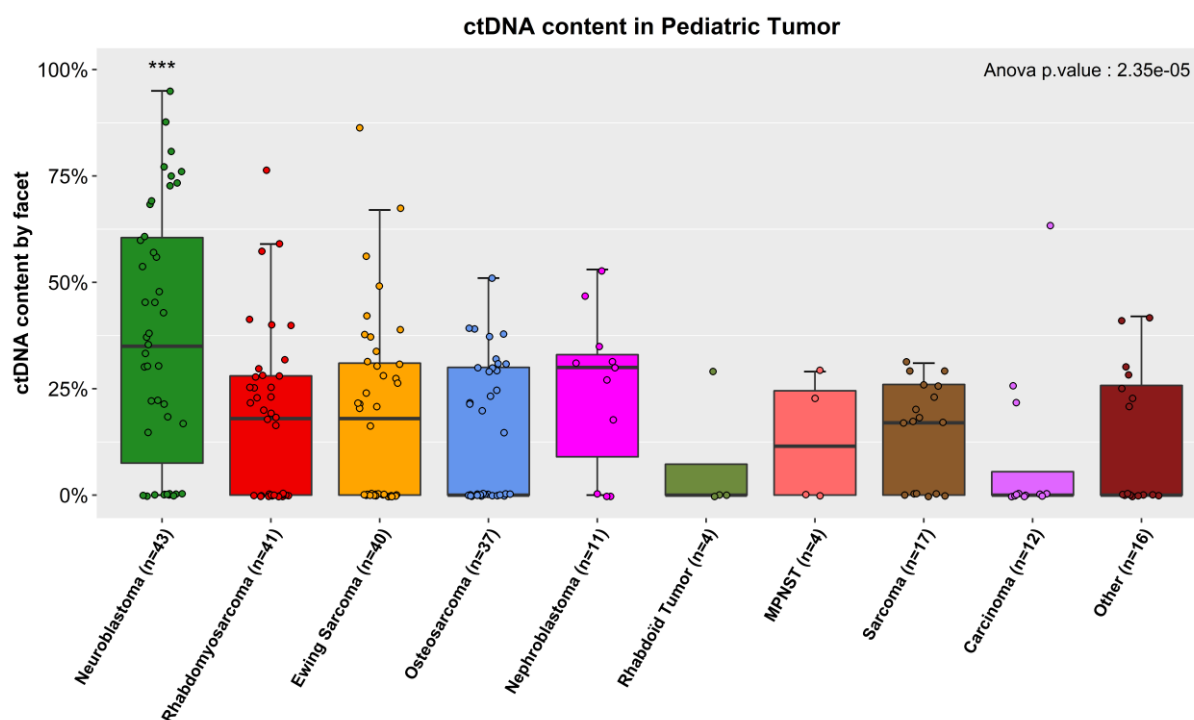


Figure 5.1.3: Contenu de ctDNA dans les plasmas calculé par Facets à partir WES pour chaque type de tumeur. Une forte différence est observée pour les neuroblastomes (test Wilcoxon = $9.686e-06$).

Le ctDNA contenu dans le plasma était en moyenne de 20% selon les calculs de l'outil Facets (de 0% à 95%)³⁵⁷. Une nouvelle fois, des différences significatives ont été observées entre chaque type de maladies (Anova p.value = $2,35e-05$) et surtout pour le neuroblastome (moyenne de 37% de 0 à 95%) comparé aux autres maladies (Wilcoxon = $3,01e-04$) (Figure 5.1.3).

3.6.1.3 Comparaison des analyses en WES entre la tumeur et le plasma

Dans une première étape, les taux de contributivité en WES de la tumeur et le plasma ont été analysés. Sur les 225 cas sélectionnés dans la cohorte MappyActs; 139 WES réalisés plasma ont été considéré comme contributif avec une observation d'altération en nombre de

copie et/ou au moins 3 mutations confirmées après visualisation par IGV (Figure 5.1.4). Pour la tumeur, 32 cas ont été considéré comme non contributifs suite à un contenu de cellules tumorales trop faible déterminé, soit par le compte rendu anapathologique, soit après séquençage de la tumeur par WES. Pour la contributivité des WES sur le plasma, une différence entre les maladies métastatiques et celles exclusivement localisées a été observée (66% de WES contributif pour les cas métastatiques et 47% pour les cas localisés ; χ^2 pvalue = 0,04).

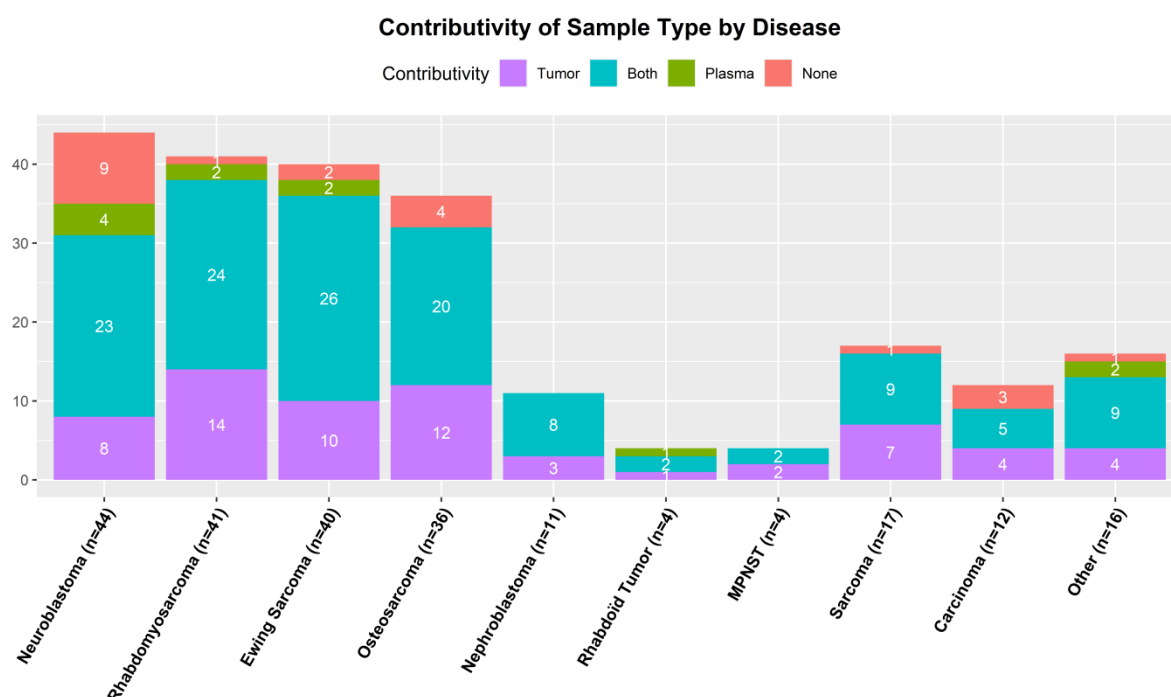


Figure 5.1.4 : Histogramme des WES contributifs par type de cancer. Les cas avec une tumeur non contributive étaient représentés par différents types de cancers : neuroblastome (n=13), sarcome d'Ewing (n=4), ostéosarcome (n=4), rhabdomyosarcome (n=3), carcinome (n=3), tumeur rhabdoïde hépatique (n=1), sarcome (n=1) et autres types de tumeurs (n=3). Pour ces cas, certains plasmas étaient contributifs (n=11) incluant divers types de cancers : neuroblastomes (n=4), sarcomes d'Ewing (n=2), rhabdomyosarcomes (n=2), tumeur rhabdoïde hépatique (n=1) et autres type de tumeurs (n=2).

Pour les 139 cas avec un WES sur le plasma contributif, 51 avait un profil d'altération en nombre de copie dynamique mais les captures d'exomes étant différentes entre l'ADN circulant et l'ADN constitutionnel, celles-ci rendaient les profils trop bruités pour permettre la recherche d'altérations focales telles que la perte du gène *CDKN2A*.

Par la suite, seuls les SNVs ont été considérés. Ainsi l'analyse des SNVs détectés dans la tumeur et dans l'ADN circulant d'un même patient a montré avec une médiane de 15 mutations communes entre la tumeur et le cfDNA (de 0 à 103 SNVs). Une médiane de 6 SNVs (de 0 à 127) a été détectée uniquement dans la tumeur et de 1 SNV (de 0 à 473) uniquement dans le plasma. Aucune différence significative n'a été observée pour la contributivité ou le nombre d'altérations détectées entre les types des cancers (Figure 5.1.5).

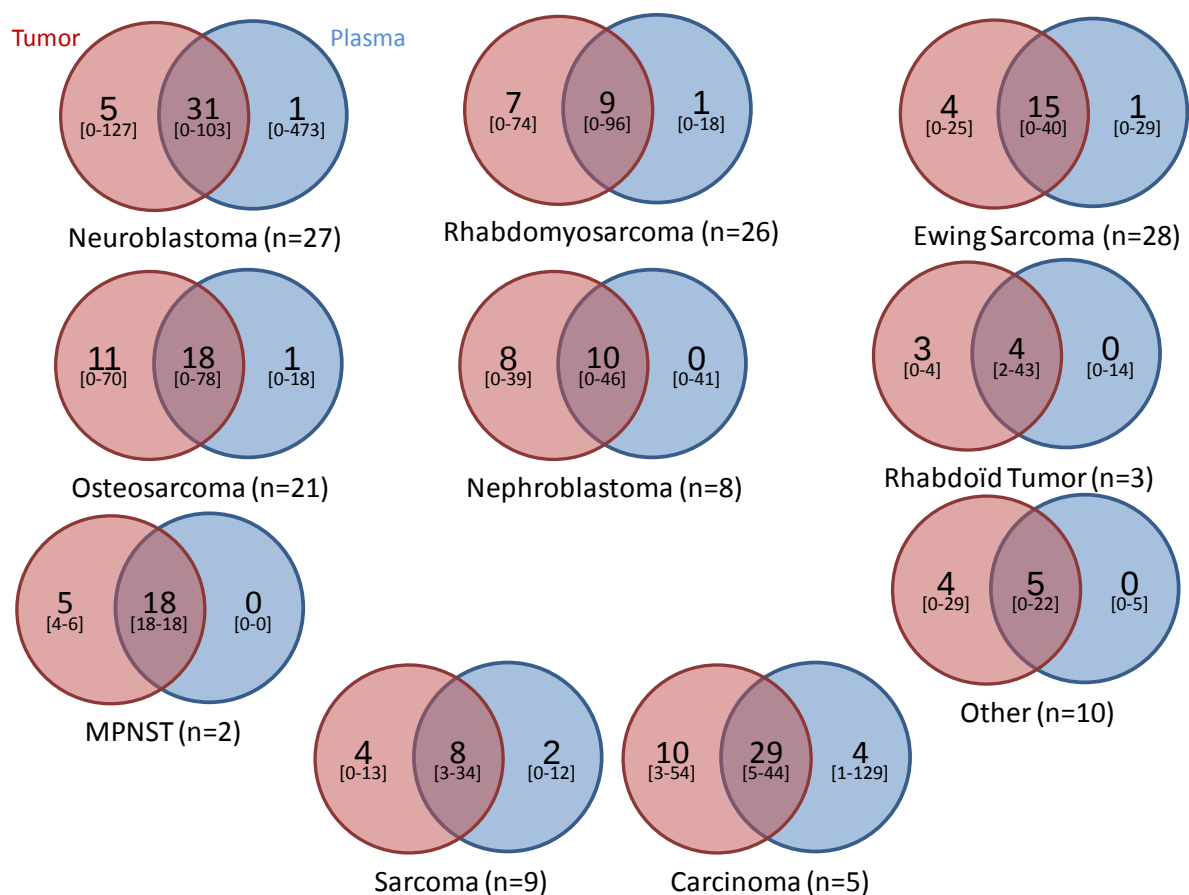


Figure 5.1.5 : Diagrammes de Venn des médianes de nombre de SNVs détectés dans la tumeur (en rouge) et dans le plasma (en bleu) pour les cas avec WES contributifs du plasma dans chaque type de cancer avec le nombre de cas entre parenthèses.

Dans un cas de rhabdomyosarcome, des altérations en nombre de copie étaient communes à la tumeur et au plasma telles qu'un gain à la fin du chromosome 2 et une perte du 6p et du 17p. Des pertes d'hétérozygotie sont également observées mais le profil de l'ADN circulant est très bruité dû aux technologies différentes de séquençages.

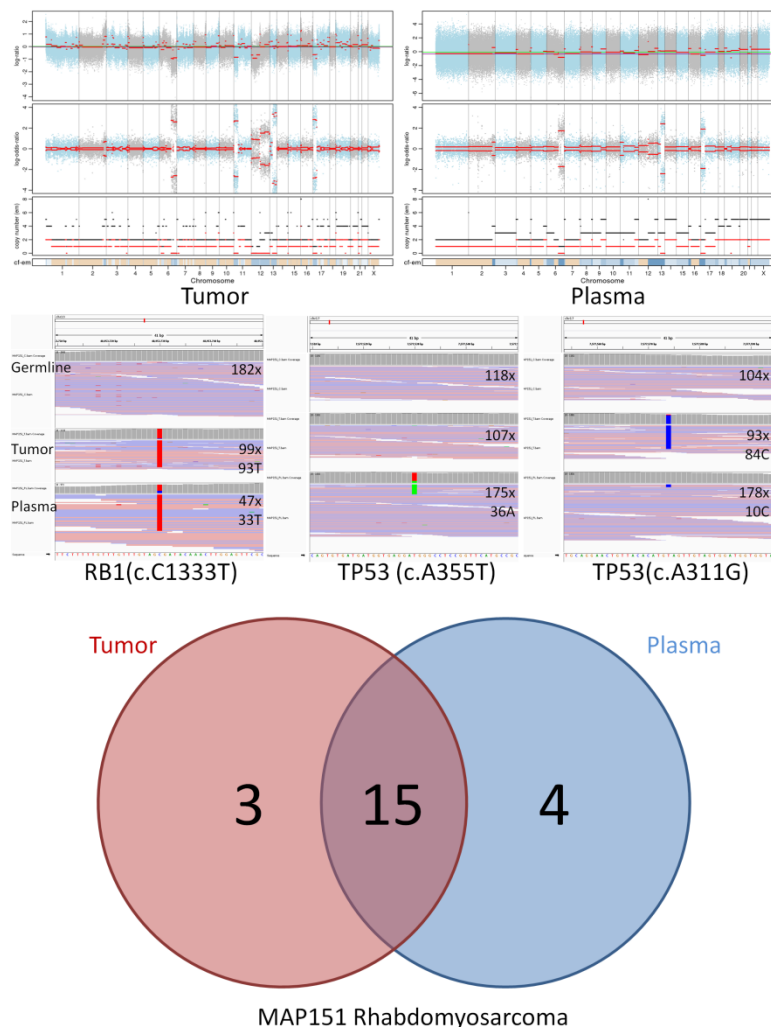
Quinze SNVs communs avec des mutations dans les gènes *RB1* et *TP53* ont été détectés (Figure 5.1.6). Trois SNVs étaient exclusifs à la tumeur et 4 au cfDNA avec deux mutations supplémentaires dans le gène *TP53*.

Figure 5.1.6 : Comparaison de résultats de WES de la tumeur et du plasma pour un cas de rhabdomyosarcome.

En haut, profils d'altérations en nombre de copie entre les deux types d'échantillons. Malgré un profil bruité pour le plasma, les altérations sont similaires.

Au milieu, SNVs détectés dans les gènes *RB1* et *TP53* dans la tumeur et dans le plasma représentés sous IGV. Au centre une des deux mutations trouvées uniquement dans le plasma pour *TP53*.

En bas, diagramme de Venn des SNVs trouvés dans la tumeur et le plasma.



Pour 4 patients, la rechute a eu lieu dans le système nerveux central et l'absence de détection de SNVs dans le sang pourrait s'expliquer par un blocage par la barrière hématoencéphalique. Mais pour un cas avec un diagnostic initial de pleuropneumoblastome, 22 SNVs étaient communs entre la rechute cérébrale et le plasma avec des mutations affectant les gènes *HRAS*, *DICER1* ou *KAT6A* (Figure 5.1.7). Cette détection dans le plasma pourrait s'expliquer par la présence d'autres métastases, effectivement ces tumeurs peuvent aussi présenter des récurrences au niveau hépatique⁴¹⁸.

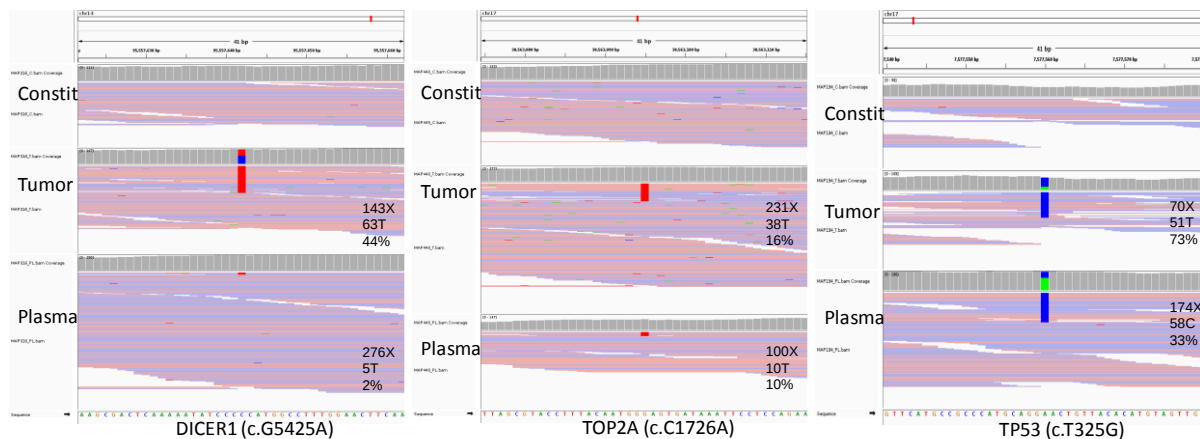


Figure 5.1.7 : SNVs communs à la tumeur et au plasma dans diverses pathologies telles qu'une mutation dans le gène *DICER1* pour une rechute de pleuropneumoblastome, dans le gène *TOP2A* pour un hépatoblastome et dans le gène *TP53* pour un ostéosarcome.

Pour 11 autres cas, plus d'un mois séparait le prélèvement du plasma et de celui de la tumeur dû à de multiples biopsies répétées pour obtenir un prélèvement avec un contenu en cellules tumorales suffisant. Seulement le WES d'un cas a pu être considéré comme contributif avec 94 SNVs communs à la tumeur et au plasma incluant une mutation dans *NOTCH1*.

Pour les 32 cas avec des prélèvements non contributifs de la tumeur, 11 plasmas ont été considérés comme contributifs avec une médiane de 18 SNVs (de 3 à 473). 12 SNVs ont été détectés dans des gènes ciblables par des drogues tels que *HRAS*, *ALK* ou *BRCA2* pour 5 patients (Figure 5.1.8).

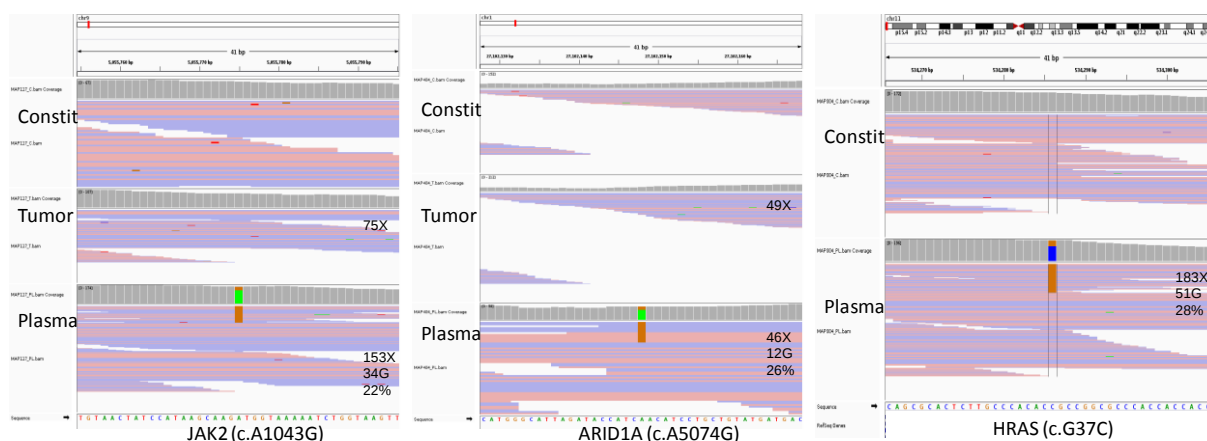


Figure 5.1.8 : SNVs ciblables seulement détectés dans le plasma tels qu'une mutation dans le gène *JAK2* pour un adénocarcinome lieberkuhnien, dans le gène *ARID1A* pour un sarcome d'Ewing et dans le gène *HRAS* pour un rhabdomyosarcome sans WES sur la tumeur (non contributif après analyse anatomopathologique).

Pour 81 cas de la cohorte "contributive" pour le plasma, 143 mutations dans des gènes constituant des cibles moléculaires (altérations génétiques ciblables ou actionnable) ont été détectées notamment dans *TP53* (35% des cas) ou *ALK* (7% des cas). En définitif, 63 % des altérations ciblables ont été détectées à la fois dans la tumeur et le plasma. De plus, 19 % des altérations étaient exclusives à la tumeur telles que *ARID1B* ou *MYCN* et 17 % des altérations étaient exclusives au plasma telles que *ATRX* ou *DICER1* (Figure 5.1.9).

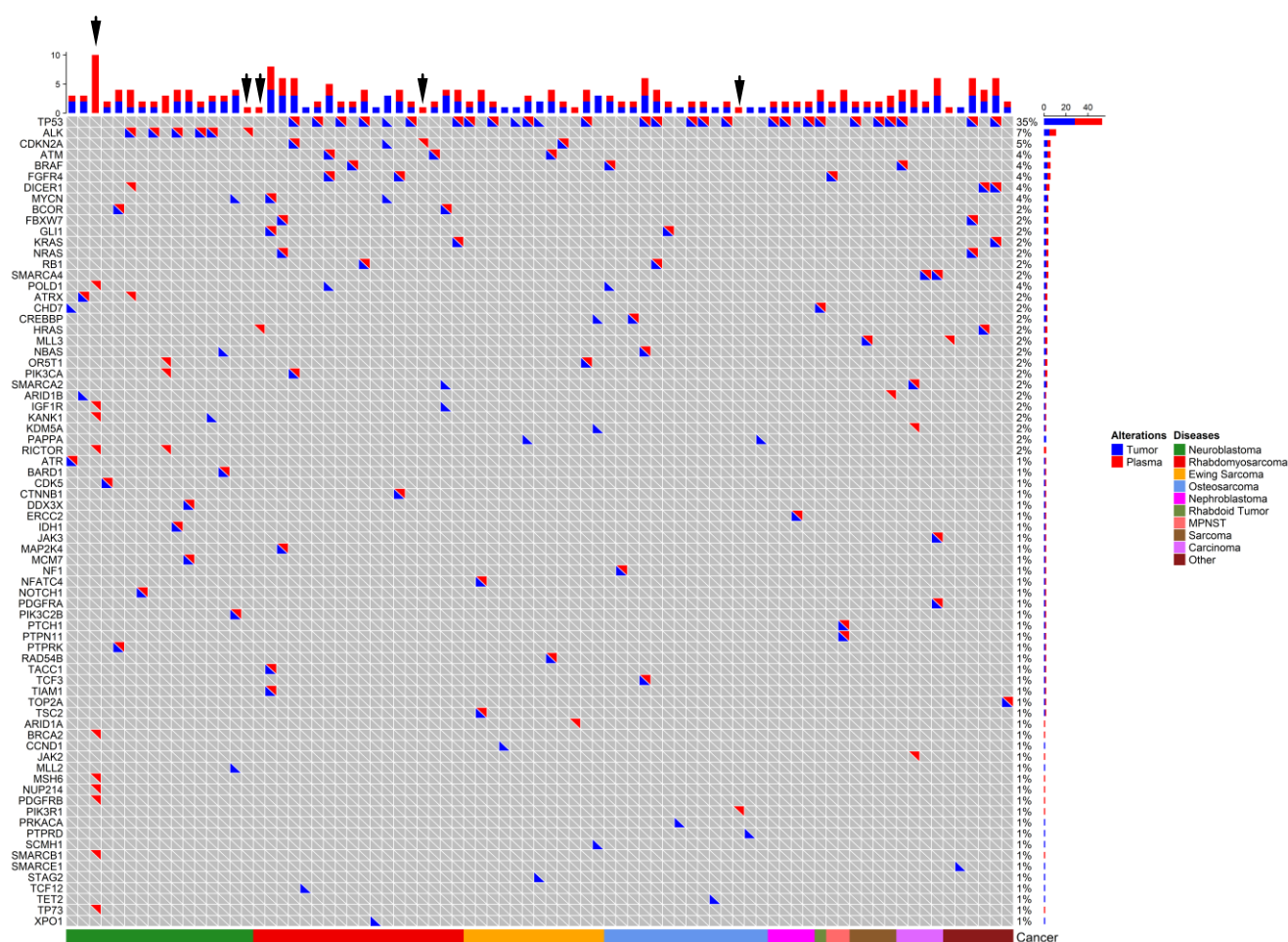


Figure 5.1.9 : Oncoprint des mutations ciblables détectées dans la tumeur (en bleu) et dans le plasma (en rouge). En haut, histogramme du nombre de SNVs par cas ; au centre, les gènes touchés par des mutations et en bas, le type de cancer. Les flèches représentent les cas sans WES sur la tumeur.

Au total, le WES sur ADN circulant a été contributif sur près de 62% des cas (139/225 cas) donc 34% où le WES de la tumeur n'avait pas été contributif (11/32 cas). La contributivité était assez similaire entre les différents types de maladie (62 % pour le

neuroblastome, 63 % le rhabdomyosarcome, 70 % pour les sarcomes d'Ewing ou 57% pour les ostéosarcomes). En conséquent, la non-contributivité peut s'expliquer par certains cas relâchant moins d'ADN circulant mais également par certaines rechutes sous forme de nodule pulmonaire centrimétrique unique. Cependant, les données cliniques sont en cours d'analyse pour pouvoir identifier une corrélation entre la contributivité d'une analyse de WES sur ADN circulant et taille de tumeur d'une tumeur ou des métastases à la rechute.

Pour les 139 cas avec un WES sur le plasma contributif, 58% des cas présentaient au moins une altération ciblable contre 70% pour l'ensemble des tumeurs de MappyAct. Cette différence peut s'expliquer par la non-considération des altérations en nombre de copie, l'absence de données de RNA-Seq et donc la détection de transcrit de fusion mais également par l'application de filtre plus stringent pour réduire les biais dus aux différentes technologies de séquençage.

De plus 17 et 19% des altérations étaient spécifiques à la tumeur ou au plasma pouvant s'expliquer par une hétérogénéité tumorale mais une nouvelle fois par un biais de séquençage. En effet et en plus de l'utilisation de technologies différentes entre le WES de la tumeur et du plasma, des différences pouvant aller jusqu'à 20% sont possibles entre exome d'un même échantillon ayant subi le même protocole⁴¹⁹.

En conclusion, cette étude a démontré le fort potentiel de l'ADN circulant comme outil complémentaire permettant de détecter des altérations de maladies réfractaires aux traitements conventionnels chez des enfants atteints de cancer et en situation de maladie en rechute.

L'étude MappyActs est en cours de soumission

3.6.2 Caractérisation et classification des tumeurs cérébrales

Les biopsies liquides peuvent être utiles en cas de difficultés d'accès de la tumeur par la biopsie. Cependant, les études sur le plasma de patients atteints de tumeurs cérébrales n'ont pas permis de caractériser clairement des altérations issues de tumeur dans l'ADN circulant. Le LCR, issues de ponctions lombaires, pourrait donc être une meilleure solution que le plasma.

Le WES sur l'ADN circulant du LCR pourrait caractériser certaines altérations propres aux tumeurs cérébrales permettant leur classification mais également d'éventuelles cibles thérapeutiques.

3.6.2.1 Extraction des liquides céphalo-rachidien

Environ 4 à 5 gouttes de Liquide Céphalo-Rachidien (LCR) ont été prélevées par ponction lombaire (400 à 500 μ L) chez 11 patients atteints de médulloblastome, 2 patients atteints d'ETMR (Embryonal Tumors with Multilayered Rosettes) et 1 patient atteint de rhabdomyosarcome avec des métastases cérébrales (Tableau Annexe 3). Les LCR ont ensuite été centrifugés à 1000g pendant 10 minutes. Par la suite, le surnageant a été utilisé pour une extraction d'ADN circulant dans la journée ou congelé à -20°C.

Le protocole d'extraction utilisé pour le plasma a été conservé pour le LCR.

La quantité médiane de cfDNA était de 33 ng par mL de plasma (de 8 à 443 ng par mL de plasma). De façon intéressante, le cas 11 (médulloblastome du groupe WNT) présentait un profil particulier de cfDNA (Figure 5.2.1). Les données cliniques ont révélé une méningite lors du prélèvement, le pic étalé observé pourrait donc être dû à des leucocytes malgré sa faible taille.

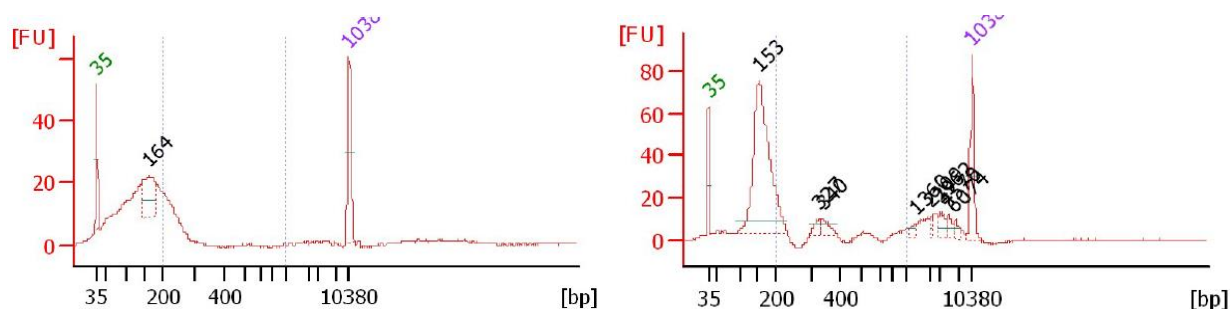


Figure 5.2.1 : Electrophorèses capillaires du cfDNA du cas 11 (à gauche) et du cas 13 (à droite) sur Bioanalyzer 2100.

3.6.2.2 Construction des librairies

Le protocole utilisé pour le cfDNA du plasma a donné des rendements aléatoires sur le LCR pour des quantités de départ similaires entre le plasma et le LCR. L'Accel-NGS 2S Hyb DNA Library Kit for Illumina™ de Swift Biosciences® a donc été utilisé selon le protocole du constructeur avec l'ajout d'une étape de sélection de taille par billes Ampure XP™ pour ne conserver que les fragments de 160 pb de cfDNA et une PCR de 9 cycles.

L'ADNg des tumeurs a également subi le même protocole avec une fragmentation mécanique préalable par Covaris™ ME220. L'ADN constitutionnel des patients n'était pas disponible.

Les échantillons ont ensuite été capturés par 8 avec le kit Nimblegen Medexome™ et séquencés sur 3 lignes de flowcell HighOutput avec l'HiSeq 2000™.

La pipeline bioinformatique était similaire aux études précédentes mais l'absence d'ADN constitutionnel et le bruit de fond important a conduit à utiliser l'outil CNVkit pour l'analyse des altérations en nombre de copie⁴²⁰.

3.6.2.3 Analyse du séquençage

Les rendements de librairie étant plus aléatoires que ceux du plasma, les contrôles de qualité du séquençage ont été renforcés. Onze des 14 WES ont ainsi été considérés comme

interprétables avec un nombre de reads mappés similaire au WES de la tumeur et un pourcentage de onTarget un peu plus faible mais correct (médiane de 82%) (Figure 5.2.2).

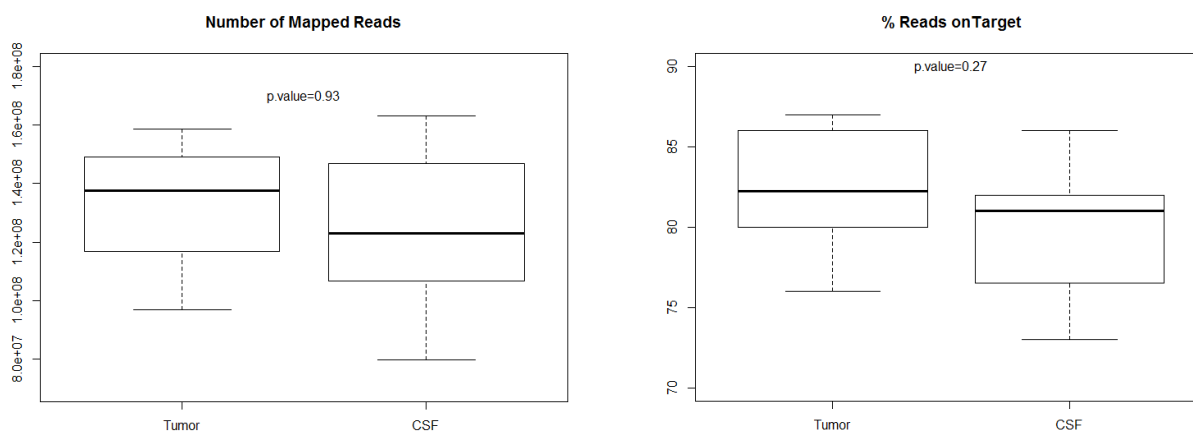


Figure 5.2.2 : Boxplots des reads mappés (à gauche) et du pourcentage onTarget (à droite) des WES de la tumeur et du LCR. Les t.test ne montrent aucune différence significative (p value sur le graphique).

Une médiane de 439 SNVs (de 18 à 872) communs a été observée entre la tumeur et le LCR. Quarante-six SNVs avec une médiane (de 3 à 132) ont été observés uniquement dans la tumeur et 66 (de 0 à 579) dans le LCR. Une forte corrélation est observée ($R^2 = 0,85$) entre la fréquence allélique des différents types d'échantillons laissant supposer une homogénéité tumorale ou la présence de nombreux polymorphismes malgré un filtrage grâce aux bases de données (Figure 5.2.2).

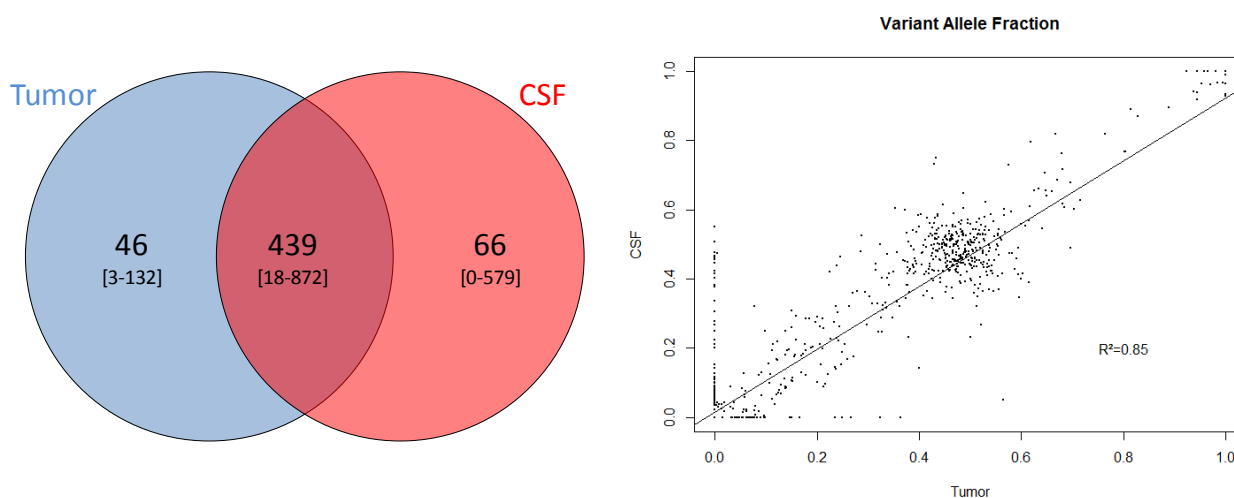


Figure 5.2.2 : Diagramme de Venn des SNVs détectés dans la tumeur et le LCR (à gauche) et nuage de points entre les fréquences alléliques des SNVs détecté dans la tumeur et le LCR (à droite) avec une forte corrélation ($R^2=0,85$).

Parmi ces SNVs, de nombreux variants ont été prédits comme délétères dans des gènes impliqués dans le développement des tumeurs cérébrales embryonnaires ou dans des oncogènes notamment impliqués dans les tumeurs cérébrales de l'enfant et de l'adulte tels que *EP400* ou *DICER1* (Figure 5.2.3)^{421–423}.

Une classification a ainsi pu être établie en fonction des SNVs détectés en accord avec les données cliniques. La présence de variants dans le gène *APC* pour la tumeur et dans le LCR a permis de classer 2 tumeurs de médulloblastome dans le groupe WNT. Un variant dans le gène *SMO* classifie une tumeur dans le groupe SHH (Figure 5.2.3). Les groupes 3 et 4 restent plus difficiles à caractériser mais certains SNVs propre à ces groupes sont tout de même détectés dans des gènes tels que *KTM2D* ou *NOTCH1*⁷⁵. De même pour les ETMR, la présence d'un SNV dans le gène *MYC* ou dans *SOX3* décrit dans ce type de tumeur ne permet pas de les classer^{424–426}.

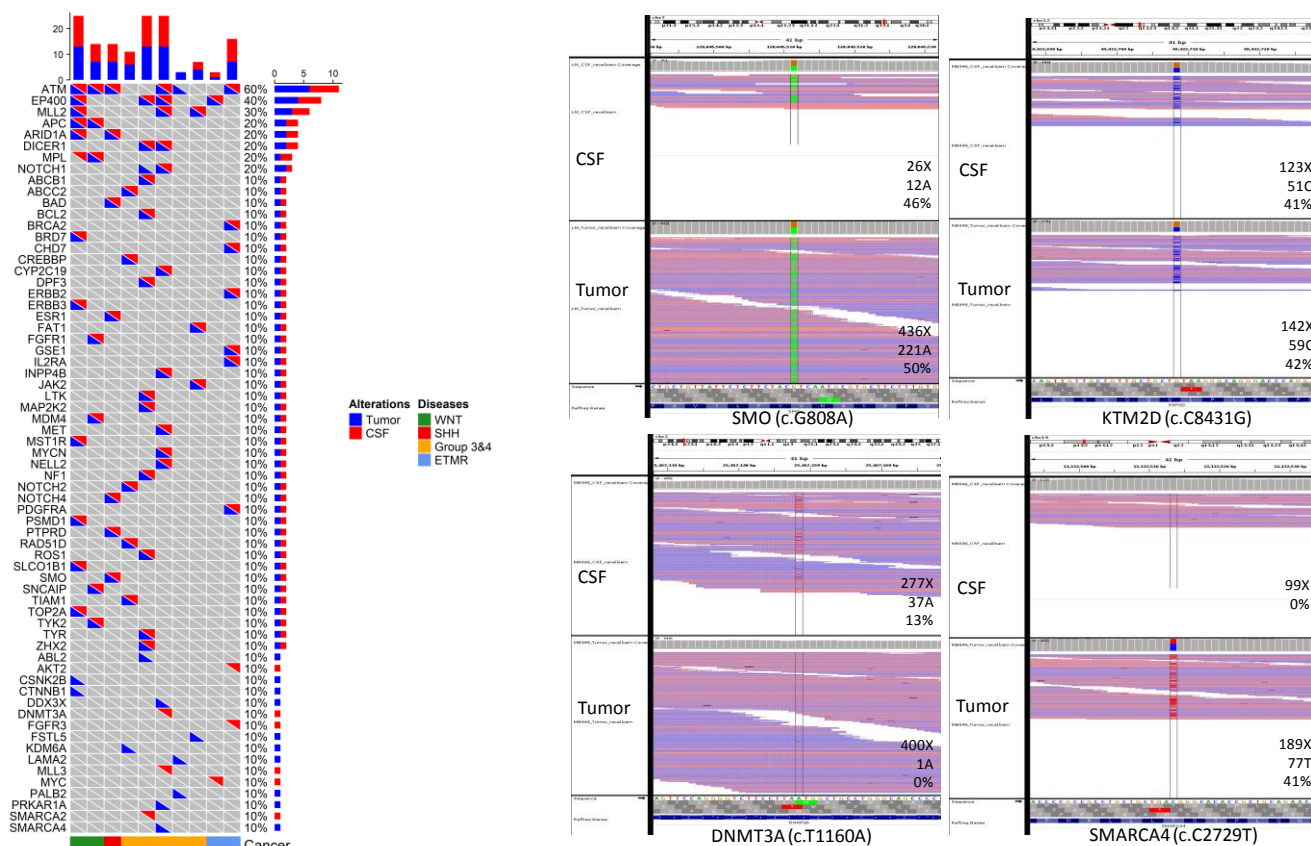


Figure 5.2.3 : Oncoprint des SNVs présents dans des oncogènes (à gauche) selon les groupes de médulloblastome et représentation de SNVs dans des gènes impliqués dans le médulloblastome commun aux deux types d'échantillons (*SMO* pour le cas 4 et *KTM2D* pour le cas 12) ou exclusif au LCR (*DNMT3A* pour le cas 12) ou à la tumeur (*SMARCA4* pour le cas 12).

Certains SNVs exclusifs à la tumeur tels que *SMARCA4* ou au LCR tels que *DNMT3A* au sein d'un même patient (cas 12) suggèrent une hétérogénéité clonale.

Cependant, certaines mutations somatiques telles que dans les gènes *APC*, *ATM* et *CTNNB1* ont toujours été observées comme exclusives entre elles pour le groupe WNT³⁶⁷. La plupart de ces SNVs pourrait être des polymorphismes. En effet, pour le cas 11, la détection d'une mutation *CTNNB1* seulement dans la tumeur et la présence d'un profil d'altérations en nombre de copie plat pour le LCR alors que des SNVs communs à la tumeur et le LCR ont pu être identifiés dans les gènes *ATM* ou *APC* mettent en lumière la difficulté de détection d'événement somatique sans la présence de WES sur l'ADN constitutionnel.

Toutefois, le cas 11 reste un cas particulier au vue de son profil d'ADN circulant et la possibilité de polymorphisme. Cependant, ces variants ont été prédit comme pathogène par les outils bio-informatiques et certains ont été décrits comme prévalent dans de nombreux cancers tels que *DICER1* voir pour certains types de médulloblastome tels que *APC*^{427,428}. Les altérations détectées par le WES du LCR, bien que non somatiques, pourraient donc jouer un rôle dans l'oncogenèse de la tumeur.

De plus, les profils d'altérations en nombre de copie ont été interprétables pour 6 des 11 WES dont 2 nécessitant le profil de la tumeur à cause d'un bruit de fond trop prononcé.

Ainsi, la présence d'une amplification de *MYCN* et également un gain du 17q permettent ainsi de classer le cas 7 dans le groupe 4 des médulloblastomes (Figure 5.2.4). De même, une amplification dans le cluster C19MC est observé dans le cas 13, celle-ci est souvent observé dans les ETMR (Figure 5.2.4)³⁶⁷.

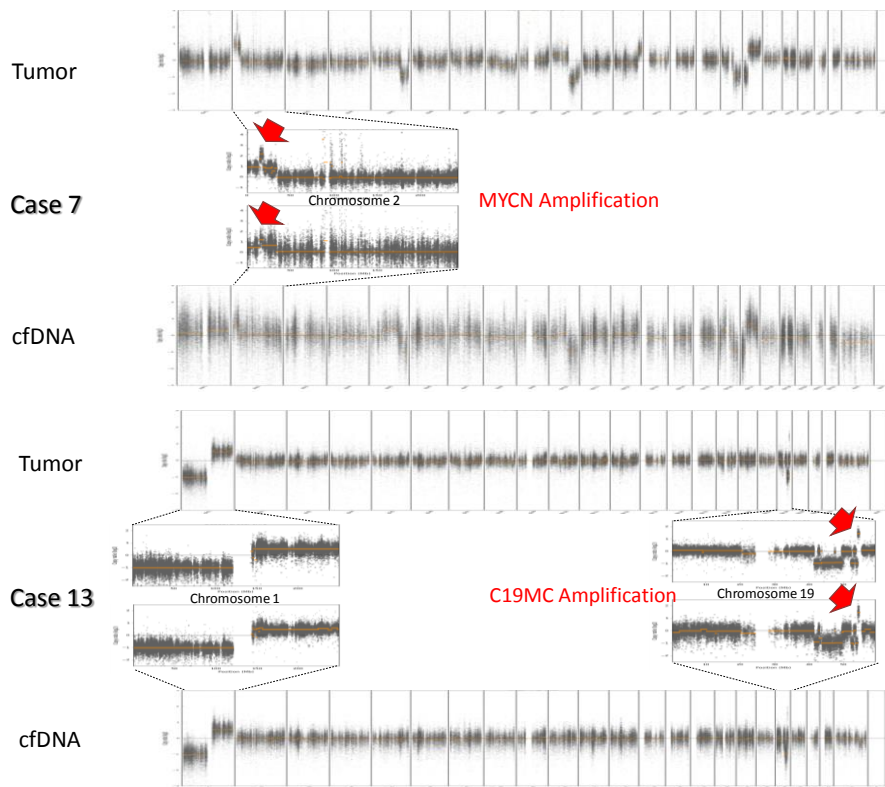


Figure 5.2.4 : Profils d'altérations en nombre de copie similaires entre la tumeur et le LCR avec une amplification du gène *MYCN* pour un médulloblastome de groupe 4 (Cas 7) et du cluster C19MC pour une ETMR (Cas 13).

3.6.2.4 Rhabdomyosarcome

L'ADN circulant d'un patient atteint de rhabdomyosarcome avec métastases cérébrales a également été extrait à partir du LCR à des fins diagnostiques en l'absence de biopsie tumorale. La quantité importante de cfDNA (96 ng/mL) a été utilisée pour réaliser une puce ADN Oncoscan™ permettant la visualisation des altérations en nombre de copie à partir d'ADN dégradé ainsi qu'un panel issue de l'étude SAFIR02 regroupant une soixantaine d'oncogène^{324,429}.

Le WES a pu être ensuite réalisé avec le reste du cfDNA soit seulement 7 ng.

439 SNVs étaient communs à la tumeur et au LCR. 132 étaient exclusifs à la tumeur tels que *BRD7* et 579 au LCR tels que *JAK2* ou *MLL2*. Huit des 10 mutations retenues par le panel SAFIR ont été détectées dans le CSF et la tumeur telles qu'un SNVs dans le gène *KRAS* ou *mTOR* (Figure 5.2.5).

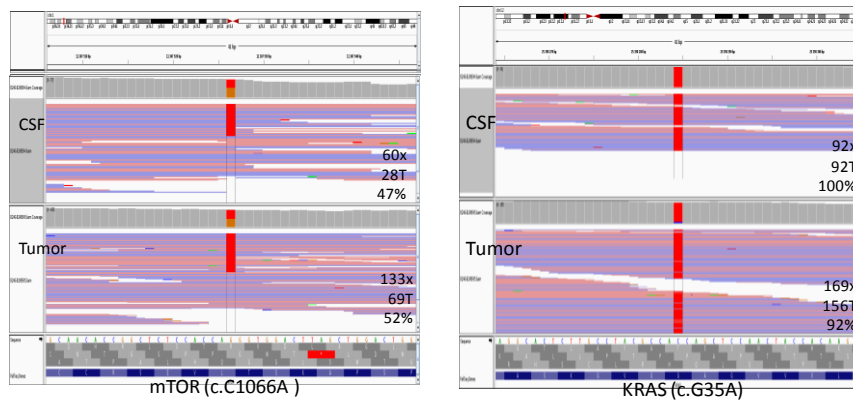


Figure 5.2.5 : SNVs observés entre la tumeur et le LCR dans les gènes *mTOR* et *KRAS*.

Des disparités d'altérations en nombre de copie sont également observées entre la tumeur et le LCR. Une perte du chromosome 4, du chromosome 16 et un gain du 12 sont seulement observés dans la tumeur tandis qu'une perte des chromosomes 9, 10, 17, 18 et X sont uniquement observées dans le LCR et ceci étant validé par deux méthodes indépendantes le WES et l'Oncoscan™ (Figure 5.2.6). Des altérations communes sont tout de même observées telles qu'un gain du chromosome 8 très souvent observé dans les rhabdomyosarcomes mais également une perte du 11p et une perte plus prononcée du 11q. Ces différences de pertes pourraient être dues à différents clones portant ces altérations et donc une hétérogénéité clonale au sein de la tumeur primaire et des métastases.

L'hétérogénéité clonale entre la tumeur primaire et les métastases serait donc très prononcée avec de nombreuses altérations (mutationnelles ou chromosomiques) suggérant une dérive très tôt en clone tumoraux et métastatiques.

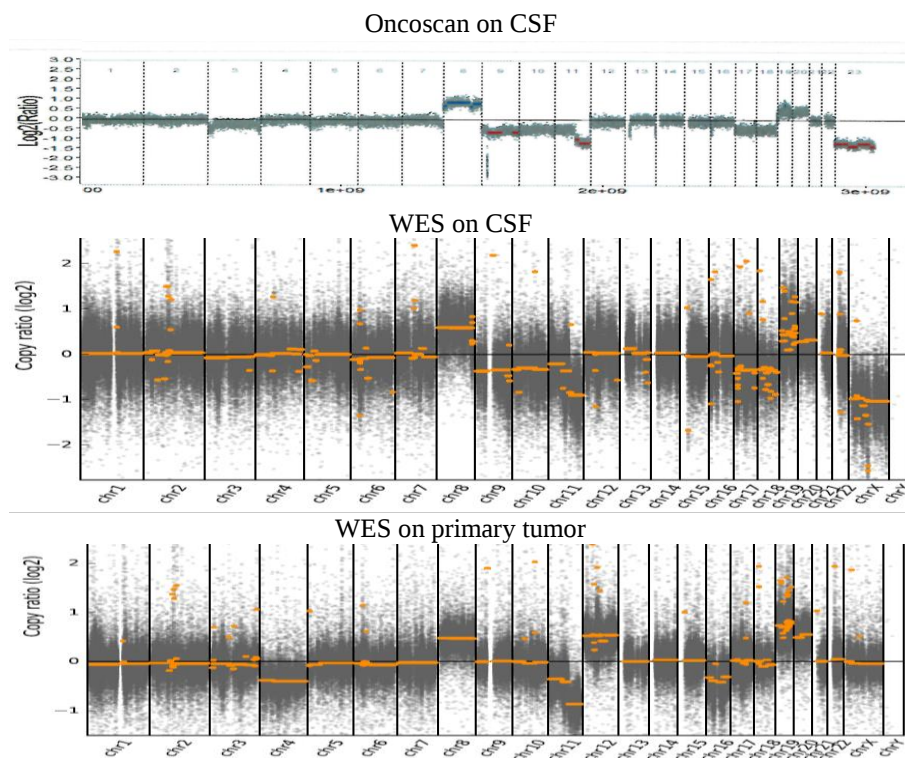


Figure 5.2.6 : Profils d'altérations en nombre de copie sur le LCR obtenus par Oncoscan (en haut) et WES (au milieu) comparés au WES de la tumeur (en bas).

Au total, le WES sur ADN circulant a été possible sur 80% des LCR mais le rendement des librairies a été plus aléatoire et a nécessité l'usage d'un kit différent à celui utilisé des plasmas. La ligation des adaptateurs semble moins efficace peut-être à cause de la forte quantité de sel présente dans le LCR qui n'aurait peut-être pas été totalement éliminé lors de l'extraction. Le LCR contient un faible taux de protéines (0,3%) et environ 5600 ont été isolées mais la majorité n'a pas été caractérisée⁴³⁰⁻⁴³². Les quantités de ces différentes protéines varient d'un individu à un autre et pourraient expliquer des interférences dans la ligation voire lors de la PCR avec des protéines restant fixées à l'ADN semblable à l'héparine responsable de l'inhibition de polymérase⁴³³.

Le WES sur LCR a permis de classifier certains médulloblastomes en groupe tels que les groupes WNT et SHH et de classifier également des ETMR. La séparation entre le groupe 3 et 4 reste plus compliquée mais pourrait être possible par l'étude du méthylome. Même si le

WES n'a pas séparé tous les groupes de tumeurs embryonnaires, des variants d'intérêt ont été détectés et pourraient être des cibles potentielles telles que *DICER1* avec des combinaisons de drogues comme vincristine/dactinomycin/cyclophosphamide ou *ARID1A* avec de l'Olaparib^{434,435}.

Ainsi, un panel contenant des gènes d'intérêt et le TSS dont l'expression des gènes associés à des groupes de tumeurs a été dessiné pour permettre la classification des tumeurs cérébrales embryonnaires comme les médulloblastomes, les ATRT ou les ETMR.

3.6.2.5 Design du panel "Tumeur Cérébrale Embryonnaire"

Pour la prédiction de l'expression, 8 TSS ont été sélectionnés comme contrôle dont 4 gènes ubiquitaires avec une expression forte ou modérée tels que *ACTB*, *B2M*, *GAPDH* ou *SDHA*⁴³⁶. A partir de la base de données Protein atlas, 2 gènes non exprimés dans le système nerveux central (*BAGE4* et *ACPT*) et 2 gènes exprimés (*SLC32A1* et *CNTNAP2*) ont été sélectionnés⁴⁰⁴.

Issues d'un panel Nanostring® à l'Institut Curie, 121 gènes ont été sélectionnés pour classer les tumeurs embryonnaires tels que *MYC* pour les médulloblastomes et ATRT de groupe MYC, *TYR* pour ATRT du groupe TYR ou *JAK3* pour des tumeurs cérébrales proche des sarcomes d'Ewing^{73,152,301,437}. Les TSS de ces gènes ont donc été inclus dans le panel "tumeurs embryonnaires".

En plus de la prédiction d'expression de certains gènes, les séquences codantes de 6 gènes (*CTNNB1*, *PTCH1*, *SMARCB1*, *SMARCA4*, *SMO*, *SUFU*) ont été ajoutées pour la recherche de mutation driver de ce type de tumeur ainsi que 5 hotspots plutôt caractéristiques des gliomes (*H3F3A_K27M*, *HIST1H3B_K27M*, *BRAF_V600E*, *IDH1_R132C* et *IDH2_R172K*)⁴³⁸⁻⁴⁴⁰.

Des SNPs précédemment utilisés ont également été ajoutés comme identifiants.

3.6.2.6 Classification des échantillons

Quinze LCR provenant de patients atteints de médulloblastome (n=10), d'ATRT (n=3), d'ETMR (n=1) et de chordome (n=1) ont été sélectionnés (Tableau Annexe 3). Trois LCR ont été séquencés par WES.

Les LCR provenaient de ponction lombaire et ont été extraits avec le kit ccfDNA de Qiagen®. La quantité médiane de cfDNA était de 7,7 ng/mL de LCR (de 4,96 à 442,75 ng/mL) après dosage Qubit™. Le même protocole que l'exome sur LCR a été utilisé pour les bibliothèques ainsi que la double capture utilisée pour le précédent panel de séquençage ciblé.

Les échantillons ont ensuite été passés sur des flowcell MiSeq™ V3 ou NovaSeq™ SP pour une couverture théorique de 500x.

Les profils de couverture ont ensuite été générés par deepTools dans un premier temps³⁹⁸.

Les premières observations montrent des différences de profils autour du TSS notamment dans la région du TSS pouvant être déplétée d'un nucléosome. Pour un cas d'ATRT (cas 16), une déplétion est observée dans plusieurs TSS dans des gènes impliqués dans le groupe nommé TYR tels que *ENPP2* et *HTRC2* (Figure 5.3.1)^{437,441,442}.

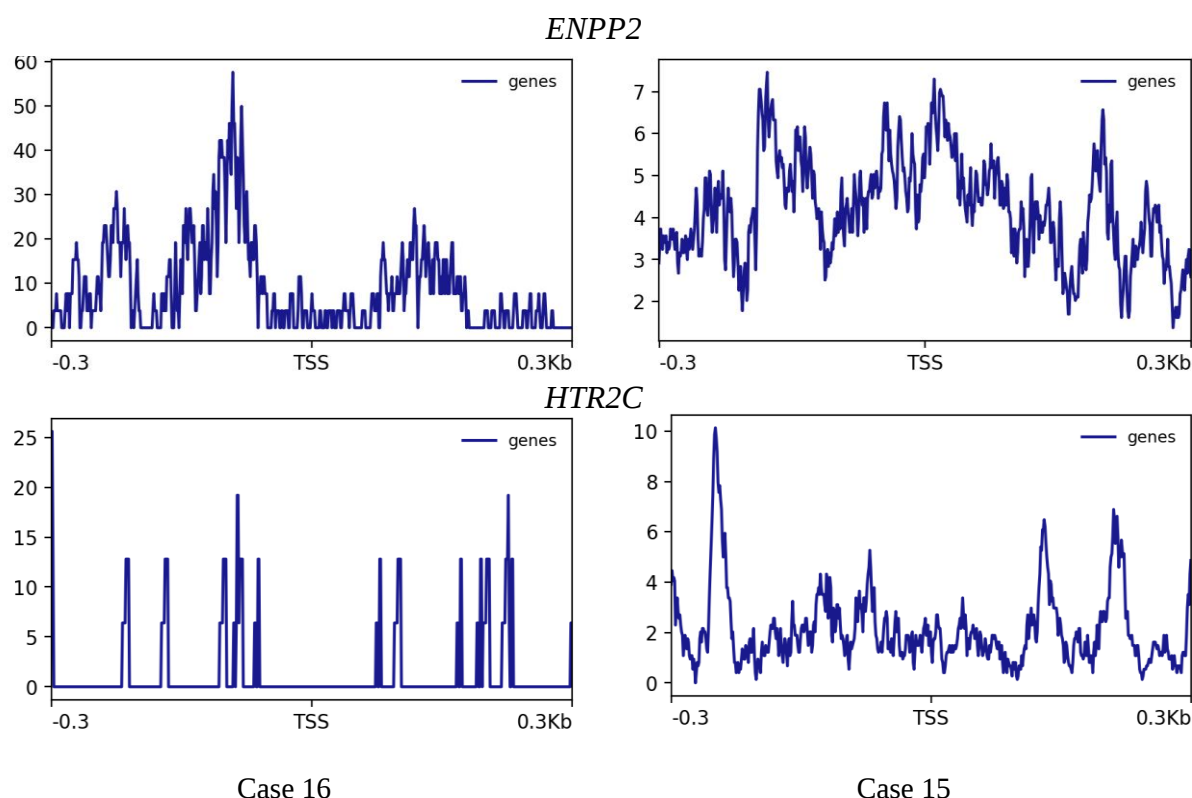


Figure 5.3.1 : Déplétions observées autour des TSS des gènes *ENPP2* (en haut) et *HTR2C* (en bas) pour le cas 16 (à gauche) et non observés pour le cas 15 (à droite). Ces observations pourraient classier l'ATRT dans le groupe TYR.

Pour un autre de cas de tumeurs embryonnaires, les TSS des gènes *CELF3* et *PFDN2* semblent également non couverts suggérant une expression de ces derniers. Ces cas pourraient correspondre au groupe de NB-FOXR2 récemment décrit (Figure 5.3.2)^{301,367}.

Ces premières données sont très préliminaires mais avec un panel couvert de 89 à 97% et l'observation d'empreintes autour de certains TSS, cette méthode pourrait permettre de classier et de caractériser les tumeurs cérébrales de types embryonnaires.

Cependant, une automatisation bio-informatique reste nécessaire pour cette classification. Le développement d'un outil basé sur du Machine Learning est donc en cours. Indépendamment de la recherche d'altération génétique, la classification pourrait être également possible avec l'ADN circulant par d'autres méthodes telles que l'analyse du méthylome.

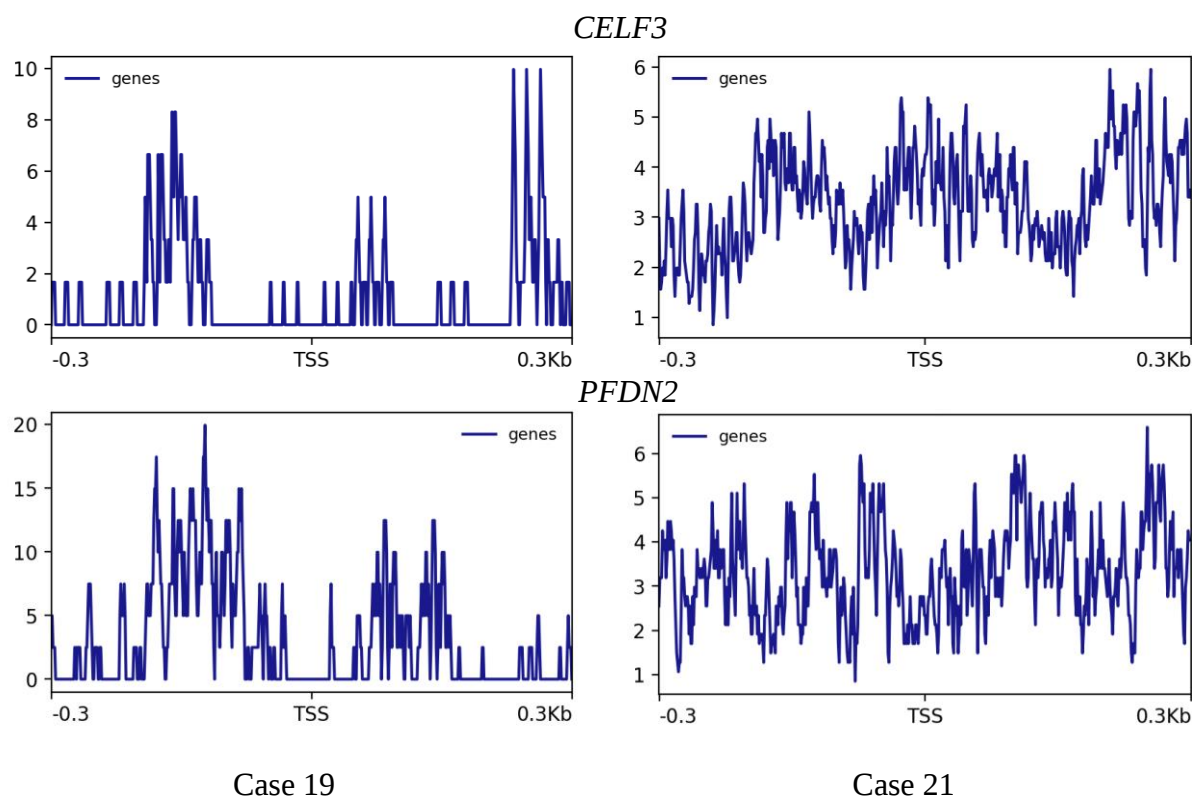


Figure 5.3.2 : Déplétions observés autour des TSS des gènes *CELF3* et *PFDN2* pour les cas 19 et non observés pour le cas 21. Ces observations pourraient classier le cas 19 dans le groupe CNS-NB-FOX2R.

3.6.3 Classification par le méthylome

Le WES sur cfDNA nous a permis de caractériser les tumeurs mais également d'observer l'émergence d'altérations à la rechute ou mettre en évidence des cibles thérapeutiques. Cependant, ce type d'analyse ne permet pas de classer les tumeurs sans informations cliniques.

Dans le cadre d'une étude collaborative sur la classification des tumeurs par la méthylation du cfDNA supervisée par le Cancer Research Institute de Ghent en Belgique, 6 plasmas de la série NGS Kids ont ainsi pu être analysés.

3.6.3.1 Choix du RRBS

La majorité des analyses de méthylome sont réalisées par puces de méthylation telles que les Illumina Infinium Human Methylation 450K™. Cependant, ces protocoles, très standardisés, demandent des quantités importantes d'ADN de départ (500 à 1000 ng) et contiennent une étape de fragmentation enzymatique indispensable et non compatible avec l'ADN circulant. Les WGS bisulfites pourraient également être adaptés mais restent très coûteux.

L'équipe de Ghent a ainsi proposé un protocole adapté de RRBS pour l'ADN circulant⁴⁴³. Ainsi après la digestion du cfDNA par MspI découpant les îlots CpG, des adaptateurs en épingle à cheveux sont fixés aux extrémités des fragments d'ADN. Si la digestion a bien eu lieu de part et d'autre du cfDNA, la circularisation de la structure d'ADN est possible. Tout ADN non circulaire est alors digéré par un mix d'exonucléases. Les adaptateurs sont ensuite ouverts pour permettre la réaction bisulfite. Pour finir, le séquençage a été réalisé par 12 échantillons sur NextSeq™ en Single-End Read de 75 pb⁴⁴⁴.

Une méthode commerciale permettant de la capture les ilots CpG a également été testée, le SeqCap EpiGiant™ de Roche® mais malgré des résultats similaires, le protocole était plus long (5 jours contre 3 jours) et plus couteux.

3.6.3.2 Classification des tumeurs pédiatriques

Soixante cfDNA provenant de 58 patients atteints de tumeurs pédiatriques ont été analysés dans le cadre de ce travail collaboratif. Le plasma a été collecté pour des cas de neuroblastome (n=10), néphroblastome (n=16), sarcome à cellules claires du rein (n=2), tumeur rhabdoïde du rein (n=1), rhabdomyosarcome (alvéolaire n=9 et embryonnaire n=8), sarcome d'Ewing (n=6) et ostéosarcome (n=4) et le LCR pour les tumeurs cérébrales, médulloblastome (n=3) et ATRT (n=1). Dix ng de cfDNA était suffisant pour la construction des banques RRBS.

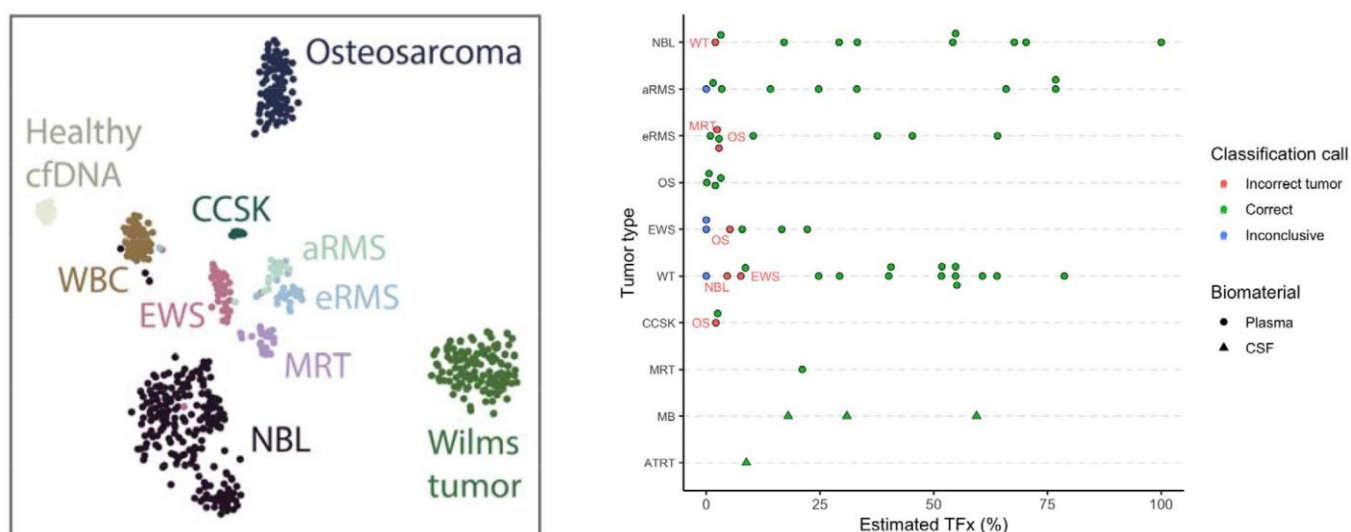


Figure 5.4.1 : Graphique t-SNE à 2 composantes pour les tumeurs extracraniales issues des bases de données (A) : Neuroblastome (NBL), Tumeur Rhabdoïde (MRT), Sarcome d'Ewing (EWS), Rhabdomyosarcome (alvéolaire : aRMS et embryonnaire : eRMS), Carcinome à Cellule Claire du Rein (CCSK), Tumeur de Wilms, Ostéosarcome et les contrôles sains tels que les globules blancs (WBC). Classification des tumeurs (B) des plasmas et LCR (MB : Medulloblastoma ; ATRT : Tumeur Rhabdoïde Teratoïde Atypique).

Une base de données a ensuite été construite à partir de données publiques de méthylation de tumeurs pédiatriques telles que TARGET pour permettre le clustering des cfDNA^{302,376,445-447}. Le méthylome des globules blancs et d'ADN circulant sains fut également

ajouté^{448,449}. La comparaison du RRBS et des puces de méthylation a été possible pour la découpe des îlots CpG en 14 103 régions et a permis ainsi de proposer des clusters selon les types de maladies (Figure 5.4.1). Certaines tumeurs ont dû être retirées suite à une clusterisation avec les globules blancs peut-être due à une contamination de la biopsie ou par une forte infiltration par ces cellules.

Dans 81,66% des cas, la classification était en accord avec le diagnostic clinique (Figure 5.4.1). La classification échoua pour des échantillons à faible contenance tumorale (<10%) ou avec une forte contamination d'ADN de haut poids moléculaire. Les deux types de rhabdomyosarcomes ont pu être également différenciés pour 13 des 14 cas du cluster de ce cancer.

En conclusion, le méthylome est de plus en plus utilisé dans la classification des tumeurs notamment pour les tumeurs cérébrales avec la mise en évidence de certains groupes auparavant inclassifiables tels que les tumeurs neuroectodermiques cérébrales (PNET)^{301,302}. Au total, le séquençage par RRBS mis au point par l'équipe de Ghent permet une classification par type tumoral avec une spécificité de 81,6%. Cette classification est également possible en routine sur des petites quantités de plasma ou du liquide céphalo-rachidien (200 µL avec 10 ng de cfDNA) et pourrait être développée pour appuyer le diagnostic.

Cet article est disponible en libre accès sur BioRxiv <https://doi.org/10.1101/795047>⁴⁵⁰

4 Discussion, perspectives et conclusion

L'ADN tumoral circulant est un formidable outil permettant d'avoir accès à de nombreuses informations de la tumeur juste par un prélèvement de fluide corporel. Cet accès facile et relativement non invasif donne également une temporalité inédite au suivi de la tumeur grâce à des prélèvements successifs possibles sans mise en danger du patient.

La comparaison entre tumeur et plasma par une analyse globale de séquençage telle que le WES a permis d'observer la très grande hétérogénéité des tumeurs pédiatriques telles que le neuroblastome. Certaines altérations étaient ainsi uniquement détectées dans le plasma et pourraient être portées par un groupe de cellules résistantes à certains traitements. L'ADN tumorale circulant semble ainsi provenir de cellules tumorales agressives et/ou des métastases pour les stades avancés. Les métastases possèdent peu des mutations driver privées comparé à la tumeur mais leurs altérations spécifiques peuvent alors donner un avantage, responsable de la résistance au traitement⁴⁵¹.

Le séquençage par WES de plasmas issus d'une étude permettant de proposer des traitements adaptés suite au profil moléculaire d'une tumeur telle que MappyActs permet de proposer un nouvel outil dans les décisions du clinicien. En plus de son utilité en absence de biopsie de la tumeur, l'ADN tumoral circulant propose un autre profil de la maladie permettant ainsi d'identifier de potentielles cibles thérapeutiques. L'ADN circulant est ainsi de plus en plus utilisé dans des protocoles cliniques et internationaux gérés par des associations internationales telles que le SIOPEN, le ctDNA étant détectable dans le plasma ou le LCR pour les tumeurs cérébrales pour une majorité de type de cancers pédiatriques.

La combinaison du WES et du séquençage ciblé sur des plasmas prélevés lors du suivi de patient a permis de mieux comprendre la spatialité clonale de la tumeur mais également sa temporalité. L'évolution clonale a pu ainsi être observée à l'aide de panels ciblant les altérations détectées par WES. En fonction de la dynamique des SNVs dans le temps, des clones ont pu être inférés aux SNVs. L'émergence d'un clone était corrélée à une progression de la maladie telle qu'une rechute ou une résistance à une chimiothérapie.

L'évolution clonale semble jouer un rôle important de l'échappement de la maladie. Le protocole clinique MICCHADO pourrait permettre de proposer ce type d'étude à une échelle inédite (600 patients). L'objectif de cette étude nationale est d'améliorer la compréhension de l'oncogenèse et des mécanismes de progression des tumeurs en identifiant et caractérisant des altérations moléculaires génétiques et les réponses immunitaires chez des patients de moins de 25 ans ayant un cancer à haut risque (neuroblastome, rhabdomyosarcome, sarcome d'Ewing, ostéosarcome, leucémie).

Ainsi, les données issues du WES de la tumeur et du plasma pourraient être utilisées pour chaque diagnostic et pour développer des panels personnalisés pour chaque patient permettant de surveiller l'émergence de clone résistant. Lors d'une rechute, le plasma devra être séquencé une nouvelle fois en WES pour caractériser de nouveaux clones.

Cependant, le nombre d'altérations reste faible dans les tumeurs pédiatriques et les outils bio-informatiques devront donc être optimisés pour permettre un modèle statistique fiable.

L'ADN circulant présente également des caractéristiques particulières permettant d'étudier le transcriptome mais également l'épigénétique.

Des empreintes sont ainsi observées selon son découpage particulier autour des nucléosomes en fonction des tissus émetteurs de l'ADN ou au niveau du TSS de certains gènes. Cette empreinte nous a ainsi permis de prédire l'expression ou non de certains gènes

sur la totalité du génome. Pour des patients atteints de neuroblastome, des profils différents ont été observés entre les cas avec une amplification du gène *MYCN* et sans amplification. Ces différences semblent être dues à certaines cibles du facteur de transcription *MYCN* mais également des gènes impliqués dans la plasticité cellulaire. Cette différence de "plasticité" a déjà été observé au sein de lignées et quelques tumeurs¹⁷⁶. L'analyse des TSS dans l'ADN circulant pourrait ainsi contribuer à l'étude du changement d'identité de la tumeur.

Toutefois, les études précédentes sont restées limitées à quelques TSS et aucune étude n'a proposé une analyse sur l'intégralité des TSS du génome. De nouveaux outils bio-informatique doivent donc être créés et une meilleure prédiction de l'expression de chaque gène doit passer par une optimisation du Machine Learning.

En outre, l'enroulement de l'ADN circulant autour d'un nucléosome a également rendu possible la détection de marque épigénétique des histones. Des marques activatrices et répressives observés chez des patients atteints de neuroblastome ont ainsi été observées dans le plasma en particuliers dans des promoteurs de gènes impliqués dans l'identité cellulaire. Néanmoins, seules les marques de méthylation ont fonctionné lors du ChIP-Seq. La marque d'acétylation H3K27ac n'était pas interprétable. Cette marque, synonyme d'enhancer, pourrait être supprimée lors d'apoptose. Au moment de l'apoptose, une désacétylation massive des histones est observée, sans savoir si ce phénomène est une cause ou une conséquence de la mort cellulaire^{452,453}. L'ADN circulant qui est majoritairement issu de l'apoptose ne pourrait donc pas posséder les marques d'acétylation. De plus, la structure des enhancers en boucle ne peut être compatible avec le découpage de l'ADN circulant et pourrait être détruite lors de cette étape.

C'est pourquoi les profils de découpage de l'ADN circulant ainsi que sa méthylation peuvent être une aide dans la classification de la maladie et la prise de décision du clinicien. Néanmoins, ces classifications ne sont pas possibles sans une importante base de données. Ces dernières restent limitées pour la majorité de cancers pédiatriques.

Cependant, une soixantaine de plasmas au diagnostic de patients atteints de cancers pédiatriques ont été séquencés par WGS en faible couverture. Ces résultats pourraient être utilisés dans une étude bio-informatique permettant de décrire un profil de découpage de l'ADN circulant différents entre types de tumeurs pédiatriques et proposer un outil de classification.

En plus de la caractérisation d'altérations en nombre de copie, le WGS low coverage permet d'estimer le contenu tumoral d'un échantillon, tumeur ou plasma. Le WGS low coverage possède un coût faible et peut utiliser à partir de libraires pré-captures et/ou avec une quantité minime d'ADN. Il peut donc être utilisé pour orienter le type d'analyse telle qu'un séquençage ciblé plutôt qu'un WES en cas de faible contenu tumoral dans l'échantillon traité.

Les propriétés de l'ADN circulant sont incroyablement prometteuses. Même s'il peut être utile en absence de biopsie, l'ADN circulant ne peut remplacer totalement cette dernière. Il reste complémentaire et permet une nouvelle vision de la maladie. Toutes ces propriétés permettent également de mieux comprendre et de suivre l'évolution et l'adaptation de la tumeur dans l'espace et le temps. D'autres aspects de l'ADN circulant permettent de suivre la maladie à un niveau résiduel (Minimal Residual Disease, MRD) et de prévenir la rechute. Cependant la sensibilité des méthodes utilisées au cours de cette étude ne permet pas de surveiller la maladie à un niveau résiduel.

L'ADN circulant peut contenir des altérations non détectées dans la tumeur et voir émerger certains clones ne répondant pas à un type de chimiothérapie. Un traitement adapté pourrait être ainsi proposé.

La possibilité d'étudier le profil de transcriptomique ou épigénétique peut être également une aide dans la classification de certains types de cancers tels que les tumeurs embryonnaires cérébrales. Ces analyses pourraient servir à l'observation de changement d'identité avec des transformations de la tumeur vers des types mésenchymateux. La tumeur ne prolifère plus mais devient plus résistante aux traitements tels que certaines chimiothérapies mais également la radiothérapie. Une observation de ce changement d'état chez un patient pourrait faire l'objet d'un suivi particulier avec l'utilisation de l'ADN circulant et des prélèvements rapprochés pour détecter une résurgence de la maladie.

Pour conclure, l'ADN circulant a permis de mieux comprendre l'échappement aux traitements par la possibilité de suivre l'évolution clonale et la plasticité d'une tumeur au cours du temps.

Ainsi l'ADN circulant se révèle être un outil indispensable pour une meilleure prise en charge du patient et pourrait permettre de briser "le plafond de verre" des 20% des patients atteints de tumeurs pédiatriques ne répondant pas aux traitements conventionnels.

Bibliographie

1. Brodeur, G. M. Spontaneous regression of neuroblastoma. *Cell Tissue Res.* **372**, 277–286 (2018).
2. Smith, M. A. *et al.* Outcomes for children and adolescents with cancer: challenges for the twenty-first century. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **28**, 2625–2634 (2010).
3. Ma, X. *et al.* Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. *Nature* **555**, 371–376 (2018).
4. Gröbner, S. N. *et al.* The landscape of genomic alterations across childhood cancers. *Nature* **555**, 321–327 (2018).
5. Alexandrov, L. B. *et al.* The repertoire of mutational signatures in human cancer. *Nature* **578**, 94–101 (2020).
6. Eleveld, T. F. *et al.* Relapsed neuroblastomas show frequent RAS-MAPK pathway mutations. *Nat. Genet.* **47**, 864–871 (2015).
7. Rokita, J. L. *et al.* Genomic Profiling of Childhood Tumor Patient-Derived Xenograft Models to Enable Rational Clinical Trial Design. *Cell Rep.* **29**, 1675-1689.e9 (2019).
8. Lawrence, M. S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
9. Matthay, K. K. *et al.* Neuroblastoma. *Nat. Rev. Dis. Primer* **2**, 16078 (2016).
10. Marshall, G. M. *et al.* The prenatal origins of cancer. *Nat. Rev. Cancer* **14**, 277–289 (2014).
11. Shimada, H. *et al.* The International Neuroblastoma Pathology Classification (the Shimada system). *Cancer* **86**, 364–372 (1999).
12. Brodeur, G. M. *et al.* International criteria for diagnosis, staging, and response to treatment in patients with neuroblastoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **6**, 1874–1881 (1988).
13. Monclair, T. *et al.* The International Neuroblastoma Risk Group (INRG) Staging System: An INRG Task Force Report. *J. Clin. Oncol.* **27**, 298–303 (2009).
14. Parikh, N. S. *et al.* SIOP-PODC adapted risk stratification and treatment guidelines: Recommendations for neuroblastoma in low- and middle-income settings. *Pediatr. Blood Cancer* **62**, 1305–1316 (2015).
15. Molenaar, J. J. *et al.* Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. *Nature* **483**, 589–593 (2012).
16. Pugh, T. J. *et al.* The genetic landscape of high-risk neuroblastoma. *Nat. Genet.* **45**, 279–284 (2013).
17. Uryu, K. *et al.* Identification of the genetic and clinical characteristics of neuroblastomas using genome-wide analysis. *Oncotarget* **8**, 107513–107529 (2017).
18. Berbegall, A. P. *et al.* Heterogeneous MYCN amplification in neuroblastoma: a SIOP Europe Neuroblastoma Study. *Br. J. Cancer* **118**, 1502–1512 (2018).
19. Brodeur, G. M., Seeger, R. C., Schwab, M., Varmus, H. E. & Bishop, J. M. Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage. *Science* **224**, 1121–1124 (1984).
20. Seeger, R. C. *et al.* Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas. *N. Engl. J. Med.* **313**, 1111–1116 (1985).
21. Huang, M. & Weiss, W. A. Neuroblastoma and MYCN. *Cold Spring Harb. Perspect. Med.* **3**, (2013).
22. Slack, A. *et al.* The p53 regulatory gene MDM2 is a direct transcriptional target of MYCN in neuroblastoma. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 731–736 (2005).

23. Chen, L. *et al.* p53 is a Direct Transcriptional Target of MYCN in Neuroblastoma. *Cancer Res.* **70**, 1377–1388 (2010).
24. Knoepfler, P. S. *et al.* Myc influences global chromatin structure. *EMBO J.* **25**, 2723–2734 (2006).
25. Corvetta, D. *et al.* Physical Interaction between MYCN Oncogene and Polycomb Repressive Complex 2 (PRC2) in Neuroblastoma. *J. Biol. Chem.* **288**, 8332–8341 (2013).
26. Marshall, G. M. *et al.* Transcriptional upregulation of histone deacetylase 2 promotes Myc-induced oncogenic effects. *Oncogene* **29**, 5957–5968 (2010).
27. Poole, C. J., Lodh, A., Choi, J.-H. & van Riggelen, J. MYC deregulates TET1 and TET2 expression to control global DNA (hydroxy)methylation and gene expression to maintain a neoplastic phenotype in T-ALL. *Epigenetics Chromatin* **12**, (2019).
28. Brenner, C. *et al.* Myc represses transcription through recruitment of DNA methyltransferase corepressor. *EMBO J.* **24**, 336–346 (2005).
29. Janoueix-Lerosey, I. *et al.* Overall genomic pattern is a predictor of outcome in neuroblastoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **27**, 1026–1033 (2009).
30. Schleiermacher, G. *et al.* Accumulation of segmental alterations determines progression in neuroblastoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **28**, 3122–3130 (2010).
31. Bown, N. *et al.* Gain of chromosome arm 17q and adverse outcome in patients with neuroblastoma. *N. Engl. J. Med.* **340**, 1954–1961 (1999).
32. Attiyeh, E. F. *et al.* Chromosome 1p and 11q Deletions and Outcome in Neuroblastoma. *N. Engl. J. Med.* **353**, 2243–2253 (2005).
33. Henrich, K.-O., Schwab, M. & Westermann, F. 1p36 tumor suppression--a matter of dosage? *Cancer Res.* **72**, 6079–6088 (2012).
34. Janoueix-Lerosey, I. *et al.* Somatic and germline activating mutations of the ALK kinase receptor in neuroblastoma. *Nature* **455**, 967–970 (2008).
35. Chen, Y. *et al.* Oncogenic mutations of ALK kinase in neuroblastoma. *Nature* **455**, 971–974 (2008).
36. Bellini, A. *et al.* Deep Sequencing Reveals Occurrence of Subclonal ALK Mutations in Neuroblastoma at Diagnosis. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **21**, 4913–4921 (2015).
37. Trigg, R. M. & Turner, S. D. ALK in Neuroblastoma: Biological and Therapeutic Implications. *Cancers* **10**, (2018).
38. Osajima-Hakomori, Y. *et al.* Biological Role of Anaplastic Lymphoma Kinase in Neuroblastoma. *Am. J. Pathol.* **167**, 213–222 (2005).
39. Nieborowska-Skorska, M. *et al.* Role of signal transducer and activator of transcription 5 in nucleophosmin/ anaplastic lymphoma kinase-mediated malignant transformation of lymphoid cells. *Cancer Res.* **61**, 6517–6523 (2001).
40. Slupianek, A. *et al.* Role of phosphatidylinositol 3-kinase-Akt pathway in nucleophosmin/anaplastic lymphoma kinase-mediated lymphomagenesis. *Cancer Res.* **61**, 2194–2199 (2001).
41. Mossé, Y. P. *et al.* Identification of ALK as the Major Familial Neuroblastoma Predisposition Gene. *Nature* **455**, 930–935 (2008).
42. Zhu, S. *et al.* Activated ALK collaborates with MYCN in neuroblastoma pathogenesis. *Cancer Cell* **21**, 362–373 (2012).
43. Schönherr, C. *et al.* Anaplastic Lymphoma Kinase (ALK) regulates initiation of transcription of MYCN in neuroblastoma cells. *Oncogene* **31**, 5193–5200 (2012).
44. Cazes, A. *et al.* Activated Alk triggers prolonged neurogenesis and Ret upregulation providing a therapeutic target in ALK-mutated neuroblastoma. *Oncotarget* **5**, 2688–2702 (2014).

45. Lopez-Delisle, L. *et al.* Activated ALK signals through the ERK-ETV5-RET pathway to drive neuroblastoma oncogenesis. *Oncogene* **37**, 1417–1429 (2018).
46. Ceccon, M. *et al.* Excess of NPM-ALK oncogenic signaling promotes cellular apoptosis and drug dependency. *Oncogene* **35**, 3854–3865 (2016).
47. Inoue, H., Hashimura, M., Akiya, M., Chiba, R. & Saegusa, M. Functional role of ALK-related signal cascades on modulation of epithelial-mesenchymal transition and apoptosis in uterine carcinosarcoma. *Mol. Cancer* **16**, 37 (2017).
48. Miyake, I. *et al.* Activation of anaplastic lymphoma kinase is responsible for hyperphosphorylation of ShcC in neuroblastoma cell lines. *Oncogene* **21**, 5823–5834 (2002).
49. Bagci, O., Tumer, S., Olgun, N. & Altungoz, O. Copy number status and mutation analyses of anaplastic lymphoma kinase (ALK) gene in 90 sporadic neuroblastoma tumors. *Cancer Lett.* **317**, 72–77 (2012).
50. Wang, M. *et al.* ALK amplification and protein expression predict inferior prognosis in neuroblastomas. *Exp. Mol. Pathol.* **95**, 124–130 (2013).
51. Schleiermacher, G. *et al.* Emergence of new ALK mutations at relapse of neuroblastoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **32**, 2727–2734 (2014).
52. Cheung, N.-K. V. *et al.* Association of age at diagnosis and genetic mutations in patients with neuroblastoma. *JAMA* **307**, 1062–1071 (2012).
53. Valentijn, L. J. *et al.* TERT rearrangements are frequent in neuroblastoma and identify aggressive tumors. *Nat. Genet.* **47**, 1411–1414 (2015).
54. Peifer, M. *et al.* Telomerase activation by genomic rearrangements in high-risk neuroblastoma. *Nature* **526**, 700–704 (2015).
55. Dagg, R. A. *et al.* Extensive Proliferation of Human Cancer Cells with Ever-Shorter Telomeres. *Cell Rep.* **19**, 2544–2556 (2017).
56. Zeineldin, M. *et al.* MYCN amplification and ATRX mutations are incompatible in neuroblastoma. *Nat. Commun.* **11**, 913 (2020).
57. Duan, X.-F. & Zhao, Q. TERT-mediated and ATRX-mediated Telomere Maintenance and Neuroblastoma. *J. Pediatr. Hematol. Oncol.* **40**, 1–6 (2018).
58. Bollmann, F. M. Targeting ALT: The role of alternative lengthening of telomeres in pathogenesis and prevention of cancer. *Cancer Treat. Rev.* **33**, 704–709 (2007).
59. Dilley, R. L. & Greenberg, R. A. ALTERNATIVE Telomere Maintenance and Cancer. *Trends Cancer* **1**, 145–156 (2015).
60. Bellini, A. *et al.* Study of chromatin remodeling genes implicates SMARCA4 as a putative player in oncogenesis in neuroblastoma. *Int. J. Cancer* **145**, 2781–2791 (2019).
61. Sausen, M. *et al.* Integrated genomic analyses identify ARID1A and ARID1B alterations in the childhood cancer neuroblastoma. *Nat. Genet.* **45**, 12–17 (2013).
62. Masliah-Planchon, J., Bièche, I., Guinebretière, J.-M., Bourdeaut, F. & Delattre, O. SWI/SNF chromatin remodeling and human malignancies. *Annu. Rev. Pathol.* **10**, 145–171 (2015).
63. Schramm, A. *et al.* Mutational dynamics between primary and relapse neuroblastomas. *Nat. Genet.* **47**, 872–877 (2015).
64. Nicolai, S., Pieraccioli, M., Peschiaroli, A., Melino, G. & Raschellà, G. Neuroblastoma: oncogenic mechanisms and therapeutic exploitation of necroptosis. *Cell Death Dis.* **6**, e2010–e2010 (2015).
65. Ackermann, S. *et al.* A mechanistic classification of clinical phenotypes in neuroblastoma. *Science* **362**, 1165–1170 (2018).
66. Ikegaki, N., Shimada, H. & International Neuroblastoma Pathology Committee. Subgrouping of Unfavorable Histology Neuroblastomas With Immunohistochemistry Toward Precision Prognosis and Therapy Stratification. *JCO Precis. Oncol.* **3**, (2019).

67. Mosse, Y. P. *et al.* Germline PHOX2B mutation in hereditary neuroblastoma. *Am. J. Hum. Genet.* **75**, 727–730 (2004).
68. Trochet, D. *et al.* Germline mutations of the paired-like homeobox 2B (PHOX2B) gene in neuroblastoma. *Am. J. Hum. Genet.* **74**, 761–764 (2004).
69. Raabe, E. H. *et al.* Prevalence and functional consequence of PHOX2B mutations in neuroblastoma. *Oncogene* **27**, 469–476 (2008).
70. Rombaut, D. *et al.* Integrative analysis identifies lincRNAs up- and downstream of neuroblastoma driver genes. *Sci. Rep.* **9**, 1–13 (2019).
71. Zeid, R. *et al.* Enhancer invasion shapes MYCN-dependent transcriptional amplification in neuroblastoma. *Nat. Genet.* **50**, 515–523 (2018).
72. Koche, R. P. *et al.* Extrachromosomal circular DNA drives oncogenic genome remodeling in neuroblastoma. *Nat. Genet.* **52**, 29–34 (2020).
73. Hovestadt, V. *et al.* Medulloblastomics revisited: biological and clinical insights from thousands of patients. *Nat. Rev. Cancer* **20**, 42–56 (2020).
74. Ellison, D. W. *et al.* beta-Catenin status predicts a favorable outcome in childhood medulloblastoma: the United Kingdom Children’s Cancer Study Group Brain Tumour Committee. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **23**, 7951–7957 (2005).
75. Northcott, P. A. *et al.* The whole-genome landscape of medulloblastoma subtypes. *Nature* **547**, 311–317 (2017).
76. Raleigh, D. R. *et al.* Hedgehog signaling drives medulloblastoma growth via CDK6. *J. Clin. Invest.* **128**, 120–124 (2018).
77. Remke, M. *et al.* TERT promoter mutations are highly recurrent in SHH subgroup medulloblastoma. *Acta Neuropathol. (Berl.)* **126**, 917–929 (2013).
78. Lindsey, J. C. *et al.* TERT promoter mutation and aberrant hypermethylation are associated with elevated expression in medulloblastoma and characterise the majority of non-infant SHH subgroup tumours. *Acta Neuropathol. (Berl.)* **127**, 307–309 (2014).
79. Northcott, P. A. *et al.* Medulloblastoma Comprises Four Distinct Molecular Variants. *J. Clin. Oncol.* **29**, 1408–1414 (2011).
80. Schwalbe, E. C. *et al.* DNA methylation profiling of medulloblastoma allows robust sub-classification and improved outcome prediction using formalin-fixed biopsies. *Acta Neuropathol. (Berl.)* **125**, 359–371 (2013).
81. Hovestadt, V. *et al.* Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. *Nature* **510**, 537–541 (2014).
82. Roussel, M. F. & Stripay, J. L. Epigenetic Drivers in Pediatric Medulloblastoma. *Cerebellum Lond. Engl.* **17**, 28–36 (2018).
83. Skapek, S. X. *et al.* Rhabdomyosarcoma. *Nat. Rev. Dis. Primer* **5**, 1–19 (2019).
84. Stevens, M. C. G. *et al.* Treatment of Nonmetastatic Rhabdomyosarcoma in Childhood and Adolescence: Third Study of the International Society of Paediatric Oncology—SIOP Malignant Mesenchymal Tumor 89. *J. Clin. Oncol.* **23**, 2618–2628 (2005).
85. Barr, F. G. *et al.* Rearrangement of the PAX3 paired box gene in the paediatric solid tumour alveolar rhabdomyosarcoma. *Nat. Genet.* **3**, 113–117 (1993).
86. Davis, R. J., D’Cruz, C. M., Lovell, M. A., Biegel, J. A. & Barr, F. G. Fusion of PAX7 to FKHR by the variant t(1;13)(p36;q14) translocation in alveolar rhabdomyosarcoma. *Cancer Res.* **54**, 2869–2872 (1994).
87. Shern, J. F. *et al.* Comprehensive genomic analysis of rhabdomyosarcoma reveals a landscape of alterations affecting a common genetic axis in fusion-positive and fusion-negative tumors. *Cancer Discov.* **4**, 216–231 (2014).
88. Breneman, J. C. *et al.* Prognostic factors and clinical outcomes in children and adolescents with metastatic rhabdomyosarcoma--a report from the Intergroup

- Rhabdomyosarcoma Study IV. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **21**, 78–84 (2003).
89. Scrabble, H. *et al.* A model for embryonal rhabdomyosarcoma tumorigenesis that involves genome imprinting. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 7480–7484 (1989).
 90. Taylor, A. C. *et al.* P53 mutation and MDM2 amplification frequency in pediatric rhabdomyosarcoma tumors and cell lines. *Med. Pediatr. Oncol.* **35**, 96–103 (2000).
 91. Stratton, M. R., Fisher, C., Gusterson, B. A. & Cooper, C. S. Detection of point mutations in N-ras and K-ras genes of human embryonal rhabdomyosarcomas using oligonucleotide probes and the polymerase chain reaction. *Cancer Res.* **49**, 6324–6327 (1989).
 92. Yang, P. *et al.* Association of childhood rhabdomyosarcoma with neurofibromatosis type I and birth defects. *Genet. Epidemiol.* **12**, 467–474 (1995).
 93. Shukla, N. *et al.* Oncogene mutation profiling of pediatric solid tumors reveals significant subsets of embryonal rhabdomyosarcoma and neuroblastoma with mutated genes in growth signaling pathways. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **18**, 748–757 (2012).
 94. Taylor, J. G. *et al.* Identification of FGFR4-activating mutations in human rhabdomyosarcomas that promote metastasis in xenotransplanted models. *J. Clin. Invest.* **119**, 3395–3407 (2009).
 95. Doros, L. *et al.* DICER1 mutations in embryonal rhabdomyosarcomas from children with and without familial PPB-tumor predisposition syndrome. *Pediatr. Blood Cancer* **59**, 558–560 (2012).
 96. Grünewald, T. G. P. *et al.* Ewing sarcoma. *Nat. Rev. Dis. Primer* **4**, 1–22 (2018).
 97. Sheffield, N. C. *et al.* DNA methylation heterogeneity defines a disease spectrum in Ewing sarcoma. *Nat. Med.* **23**, 386–395 (2017).
 98. Riggi, N. *et al.* Development of Ewing's Sarcoma from Primary Bone Marrow-Derived Mesenchymal Progenitor Cells. *Cancer Res.* **65**, 11459–11468 (2005).
 99. Aurias, A., Rimbaut, C., Buffe, D., Zucker, J.-M. & Mazabraud, A. Translocation involving chromosome 22 in Ewing's Sarcoma. A cytogenetic study of four fresh tumors. *Cancer Genet. Cytogenet.* **12**, 21–25 (1984).
 100. Delattre, O. *et al.* Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **359**, 162–165 (1992).
 101. Jeon, I. S. *et al.* A variant Ewing's sarcoma translocation (7;22) fuses the EWS gene to the ETS gene ETV1. *Oncogene* **10**, 1229–1234 (1995).
 102. Urano, F. *et al.* Molecular analysis of Ewing's sarcoma: another fusion gene, EWS-E1AF, available for diagnosis. *Jpn. J. Cancer Res. Gann* **89**, 703–711 (1998).
 103. Peter, M. *et al.* A new member of the ETS family fused to EWS in Ewing tumors. *Oncogene* **14**, 1159–1164 (1997).
 104. Ng, T. L. *et al.* Ewing Sarcoma with Novel Translocation t(2;16) Producing an In-Frame Fusion of FUS and FEV. *J. Mol. Diagn. JMD* **9**, 459–463 (2007).
 105. Armengol, G. *et al.* Recurrent gains of 1q, 8 and 12 in the Ewing family of tumours by comparative genomic hybridization. *Br. J. Cancer* **75**, 1403–1409 (1997).
 106. Ozaki, T. *et al.* Genetic imbalances revealed by comparative genomic hybridization in Ewing tumors. *Genes. Chromosomes Cancer* **32**, 164–171 (2001).
 107. Tirode, F. *et al.* Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations. *Cancer Discov.* **4**, 1342–1353 (2014).
 108. Crompton, B. D. *et al.* The genomic landscape of pediatric Ewing sarcoma. *Cancer Discov.* **4**, 1326–1341 (2014).
 109. Rivera, M. N. & Haber, D. A. Wilms' tumour: connecting tumorigenesis and organ development in the kidney. *Nat. Rev. Cancer* **5**, 699–712 (2005).

110. Coorens, T. H. *et al.* Embryonal precursors of Wilms tumor. *Science* **366**, 1247–1251 (2019).
111. Ooms, A. H. A. G. *et al.* Significance of TP53 Mutation in Wilms Tumors with Diffuse Anaplasia: A Report from the Children's Oncology Group. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **22**, 5582–5591 (2016).
112. Gadd, S. *et al.* A Children's Oncology Group and TARGET initiative exploring the genetic landscape of Wilms tumor. *Nat. Genet.* **49**, 1487–1494 (2017).
113. Treger, T. D., Chowdhury, T., Pritchard-Jones, K. & Behjati, S. The genetic changes of Wilms tumour. *Nat. Rev. Nephrol.* **15**, 240–251 (2019).
114. Grundy, P. E. *et al.* Loss of heterozygosity for chromosomes 1p and 16q is an adverse prognostic factor in favorable-histology Wilms tumor: a report from the National Wilms Tumor Study Group. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **23**, 7312–7321 (2005).
115. Wittmann, S. *et al.* Loss of 11q and 16q in Wilms tumors is associated with anaplasia, tumor recurrence, and poor prognosis. *Genes. Chromosomes Cancer* **46**, 163–170 (2007).
116. Williams, R. D. *et al.* Molecular profiling reveals frequent gain of MYCN and anaplasia-specific loss of 4q and 14q in Wilms tumor. *Genes. Chromosomes Cancer* **50**, 982–995 (2011).
117. Pelletier, J. *et al.* WT1 mutations contribute to abnormal genital system development and hereditary Wilms' tumour. *Nature* **353**, 431–434 (1991).
118. Pelletier, J. *et al.* Germline mutations in the Wilms' tumor suppressor gene are associated with abnormal urogenital development in Denys-Drash syndrome. *Cell* **67**, 437–447 (1991).
119. Palculict, T. B. *et al.* Identification of germline DICER1 mutations and loss of heterozygosity in familial Wilms tumour. *J. Med. Genet.* **53**, 385–388 (2016).
120. Dancey, J. E., Bedard, P. L., Onetto, N. & Hudson, T. J. The genetic basis for cancer treatment decisions. *Cell* **148**, 409–420 (2012).
121. Le Tourneau, C. *et al.* Designs and challenges for personalized medicine studies in oncology: focus on the SHIVA trial. *Target. Oncol.* **7**, 253–265 (2012).
122. Servant, N. *et al.* Bioinformatics for precision medicine in oncology: principles and application to the SHIVA clinical trial. *Front. Genet.* **5**, (2014).
123. Massard, C. *et al.* High-Throughput Genomics and Clinical Outcome in Hard-to-Treat Advanced Cancers: Results of the MOSCATO 01 Trial. *Cancer Discov.* (2017) doi:10.1158/2159-8290.CD-16-1396.
124. Darwin, C. On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. *Br. Foreign Medico-Chir. Rev.* **25**, 367–404 (1860).
125. Cahill, D. P., Kinzler, K. W., Vogelstein, B. & Lengauer, C. Genetic instability and darwinian selection in tumours. *Trends Cell Biol.* **9**, M57-60 (1999).
126. Gerlinger, M. & Swanton, C. How Darwinian models inform therapeutic failure initiated by clonal heterogeneity in cancer medicine. *Br. J. Cancer* **103**, 1139–1143 (2010).
127. Greaves, M. & Maley, C. C. CLONAL EVOLUTION IN CANCER. *Nature* **481**, 306–313 (2012).
128. Ding, L., Raphael, B. J., Chen, F. & Wendl, M. C. Advances for studying clonal evolution in cancer. *Cancer Lett.* **340**, 212–219 (2013).
129. Williams, M. J., Sottoriva, A. & Graham, T. A. Measuring Clonal Evolution in Cancer with Genomics. *Annu. Rev. Genomics Hum. Genet.* **20**, 309–329 (2019).
130. Nordling, C. O. A New Theory on the Cancer-inducing Mechanism. *Br. J. Cancer* **7**, 68–72 (1953).
131. Knudson, A. G. Mutation and Cancer: Statistical Study of Retinoblastoma. *Proc. Natl. Acad. Sci. U. S. A.* **68**, 820–823 (1971).

132. Dimaras, H. *et al.* Retinoblastoma. *Nat. Rev. Dis. Primer* **1**, 1–23 (2015).
133. Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
134. Snuderl, M. *et al.* Mosaic Amplification of Multiple Receptor Tyrosine Kinase Genes in Glioblastoma. *Cancer Cell* **20**, 810–817 (2011).
135. Navin, N. E. & Hicks, J. Tracing the tumor lineage. *Mol. Oncol.* **4**, 267–283 (2010).
136. Navin, N. *et al.* Tumor Evolution Inferred by Single Cell Sequencing. *Nature* **472**, 90 (2011).
137. Yachida, S. *et al.* Distant Metastasis Occurs Late during the Genetic Evolution of Pancreatic Cancer. *Nature* **467**, 1114–1117 (2010).
138. Nik-Zainal, S. *et al.* The Life History of 21 Breast Cancers. *Cell* **149**, 994–1007 (2012).
139. Welch, J. S. *et al.* The origin and evolution of mutations in Acute Myeloid Leukemia. *Cell* **150**, 264–278 (2012).
140. Ding, L. *et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**, 506–510 (2012).
141. Deveau, P. *et al.* QuantumClone: clonal assessment of functional mutations in cancer based on a genotype-aware method for clonal reconstruction. *Bioinforma. Oxf. Engl.* **34**, 1808–1816 (2018).
142. Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
143. Turajlic, S. *et al.* Deterministic Evolutionary Trajectories Influence Primary Tumor Growth: TRACERx Renal. *Cell* **173**, 595-610.e11 (2018).
144. Mitchell, T. J. *et al.* Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. *Cell* **173**, 611-623.e17 (2018).
145. Turajlic, S. *et al.* Tracking Cancer Evolution Reveals Constrained Routes to Metastases: TRACERx Renal. *Cell* **173**, 581-594.e12 (2018).
146. Jamal-Hanjani, M. *et al.* Tracking the Evolution of Non-Small-Cell Lung Cancer. *N. Engl. J. Med.* **376**, 2109–2121 (2017).
147. Cleary, A. S., Leonard, T. L., Gestl, S. A. & Gunther, E. J. Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature* **508**, 113–117 (2014).
148. Marusyk, A. *et al.* Non-cell autonomous tumor-growth driving supports sub-clonal heterogeneity. *Nature* **514**, 54–58 (2014).
149. Ambros, I. M., Brunner, C., Abbasi, R., Frech, C. & Ambros, P. F. Ultra-High Density SNParray in Neuroblastoma Molecular Diagnostics. *Front. Oncol.* **4**, (2014).
150. Durand, S. *et al.* ALK mutation dynamics and clonal evolution in a neuroblastoma model exhibiting two ALK mutations. *Oncotarget* **10**, 4937–4950 (2019).
151. Carpenter, E. L. & Mossé, Y. P. Targeting ALK in neuroblastoma--preclinical and clinical advancements. *Nat. Rev. Clin. Oncol.* **9**, 391–399 (2012).
152. Ramaswamy, V. *et al.* Recurrence patterns across medulloblastoma subgroups: an integrated clinical and molecular analysis. *Lancet Oncol.* **14**, 1200–1207 (2013).
153. Morrissy, A. S. *et al.* Divergent clonal selection dominates medulloblastoma at recurrence. *Nature* **529**, 351–357 (2016).
154. Cresswell, G. D. *et al.* Intra-Tumor Genetic Heterogeneity in Wilms Tumor: Clonal Evolution and Clinical Implications. *EBioMedicine* **9**, 120–129 (2016).
155. Hoffman, M. *et al.* Intratumoral genetic and functional heterogeneity in pediatric glioblastoma. *Cancer Res.* (2019) doi:10.1158/0008-5472.CAN-18-3441.

156. Vinci, M. *et al.* Functional diversity and co-operativity between subclonal populations of paediatric glioblastoma and diffuse intrinsic pontine glioma cells. *Nat. Med.* **24**, 1204–1215 (2018).
157. Mengelbier, L. H. *et al.* Intratumoral genome diversity parallels progression and predicts outcome in pediatric cancer. *Nat. Commun.* **6**, 1–10 (2015).
158. Karlsson, J. *et al.* Four evolutionary trajectories underlie genetic intratumoral variation in childhood cancer. *Nat. Genet.* **50**, 944–950 (2018).
159. Clevers, H. The cancer stem cell: premises, promises and challenges. *Nat. Med.* **17**, 313–319 (2011).
160. Flavahan, W. A., Gaskell, E. & Bernstein, B. E. Epigenetic plasticity and the hallmarks of cancer. *Science* **357**, (2017).
161. da Silva-Diz, V., Lorenzo-Sanz, L., Bernat-Peguera, A., Lopez-Cerda, M. & Muñoz, P. Cancer cell plasticity: Impact on tumor progression and therapy response. *Semin. Cancer Biol.* **53**, 48–58 (2018).
162. Yuan, S., Norgard, R. J. & Stanger, B. Z. Cellular Plasticity in Cancer. *Cancer Discov.* **9**, 837–851 (2019).
163. Boumahdi, S. & Sauvage, F. J. de. The great escape: tumour cell plasticity in resistance to targeted therapy. *Nat. Rev. Drug Discov.* **19**, 39–56 (2020).
164. Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3983–3988 (2003).
165. Suvà, M. L., Riggi, N. & Bernstein, B. E. Epigenetic reprogramming in cancer. *Science* **339**, 1567–1570 (2013).
166. Halliday, J. *et al.* In vivo radiation response of proneural glioma characterized by protective p53 transcriptional program and proneural-mesenchymal shift. *Proc. Natl. Acad. Sci.* **111**, 5248–5253 (2014).
167. Tschaharganeh, D. F. *et al.* p53-Dependent Nestin Regulation Links Tumor Suppression to Cellular Plasticity in Liver Cancer. *Cell* **158**, 579–592 (2014).
168. Mills, J. C. & Goldenring, J. R. Metaplasia in the Stomach Arises From Gastric Chief Cells. *Cell. Mol. Gastroenterol. Hepatol.* **4**, 85–88 (2017).
169. Zhang, H. *et al.* Lkb1 inactivation drives lung cancer lineage switching governed by Polycomb Repressive Complex 2. *Nat. Commun.* **8**, 14922 (2017).
170. Sun, B., Zhang, D., Zhao, N. & Zhao, X. Epithelial-to-endothelial transition and cancer stem cells: two cornerstones of vasculogenic mimicry in malignant tumors. *Oncotarget* **8**, 30502–30510 (2017).
171. Ottersbach, K. Endothelial-to-haematopoietic transition: an update on the process of making blood. *Biochem. Soc. Trans.* **47**, 591–601 (2019).
172. Saxena, M., Stephens, M. A., Pathak, H. & Rangarajan, A. Transcription factors that mediate epithelial–mesenchymal transition lead to multidrug resistance by upregulating ABC transporters. *Cell Death Dis.* **2**, e179 (2011).
173. Zheng, X. *et al.* EMT Program is Dispensable for Metastasis but Induces Chemoresistance in Pancreatic Cancer. *Nature* **527**, 525–530 (2015).
174. Terry, S. *et al.* New insights into the role of EMT in tumor immune escape. *Mol. Oncol.* **11**, 824–846 (2017).
175. Franzetti, G.-A. *et al.* Cell-to-cell heterogeneity of EWSR1-FLI1 activity determines proliferation/migration choices in Ewing sarcoma cells. *Oncogene* **36**, 3505–3514 (2017).
176. van Groningen, T. *et al.* Neuroblastoma is composed of two super-enhancer-associated differentiation states. *Nat. Genet.* **49**, 1261–1266 (2017).
177. Boeva, V. *et al.* Heterogeneity of neuroblastoma cell identity defined by transcriptional circuitries. *Nat. Genet.* **49**, 1408–1413 (2017).

178. Weiss, T. *et al.* Schwann cell plasticity regulates neuroblastic tumor cell differentiation. *bioRxiv* 2020.04.01.019422 (2020) doi:10.1101/2020.04.01.019422.
179. Barton, N. H. Natural selection and random genetic drift as causes of evolution on islands. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **351**, 785–794; discussion 795 (1996).
180. Lawson, D. A., Kessenbrock, K., Davis, R. T., Pervolarakis, N. & Werb, Z. Tumour heterogeneity and metastasis at single-cell resolution. *Nat. Cell Biol.* **20**, 1349–1360 (2018).
181. Crowley, E., Di Nicolantonio, F., Loupakis, F. & Bardelli, A. Liquid biopsy: monitoring cancer-genetics in the blood. *Nat. Rev. Clin. Oncol.* **10**, 472–484 (2013).
182. Saini, A. *et al.* Liquid Biopsy in Gastrointestinal Cancers. *Diagnostics* **8**, 75 (2018).
183. Alix-Panabières, C., Schwarzenbach, H. & Pantel, K. Circulating tumor cells and circulating tumor DNA. *Annu. Rev. Med.* **63**, 199–215 (2012).
184. Bidard, F.-C., Proudhon, C. & Pierga, J.-Y. Circulating tumor cells in breast cancer. *Mol. Oncol.* **10**, 418–430 (2016).
185. Delaloge, S. *et al.* Dépistage du cancer du sein : en route vers le futur. *Bull. Cancer (Paris)* **103**, 753–763 (2016).
186. Nguyen, D. X., Bos, P. D. & Massagué, J. Metastasis: from dissemination to organ-specific colonization. *Nat. Rev. Cancer* **9**, 274–284 (2009).
187. Cristofanilli, M. *et al.* Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N. Engl. J. Med.* **351**, 781–791 (2004).
188. Bidard, F.-C. *et al.* Circulating Tumor Cells and Circulating Tumor DNA Detection in Potentially Resectable Metastatic Colorectal Cancer: A Prospective Ancillary Study to the Unicancer Prodiges-14 Trial. *Cells* **8**, (2019).
189. Thery, L. *et al.* Circulating Tumor Cells in Early Breast Cancer. *JNCI Cancer Spectr.* **3**, pkz026 (2019).
190. Sharma, S. *et al.* Circulating Tumor Cell Isolation, Culture, and Downstream Molecular Analysis. *Biotechnol. Adv.* **36**, 1063–1078 (2018).
191. Ameri, K. *et al.* Circulating tumour cells demonstrate an altered response to hypoxia and an aggressive phenotype. *Br. J. Cancer* **102**, 561–569 (2010).
192. Alix-Panabières, C. *et al.* Detection and characterization of putative metastatic precursor cells in cancer patients. *Clin. Chem.* **53**, 537–539 (2007).
193. Raposo, G. & Stoorvogel, W. Extracellular vesicles: Exosomes, microvesicles, and friends. *J. Cell Biol.* **200**, 373–383 (2013).
194. Tai, Y., Chen, K., Hsieh, J. & Shen, T. Exosomes in cancer development and clinical applications. *Cancer Sci.* **109**, 2364–2374 (2018).
195. Ge, R., Tan, E., Sharghi-Namini, S. & Asada, H. H. Exosomes in Cancer Microenvironment and Beyond: have we Overlooked these Extracellular Messengers? *Cancer Microenviron.* **5**, 323–332 (2012).
196. Osaki, M. & Okada, F. Exosomes and Their Role in Cancer Progression. *Yonago Acta Med.* **62**, 182–190 (2019).
197. Ha, D., Yang, N. & Nadithe, V. Exosomes as therapeutic drug carriers and delivery vehicles across biological membranes: current perspectives and future challenges. *Acta Pharm. Sin. B* **6**, 287–296 (2016).
198. Peng, Y. & Croce, C. M. The role of MicroRNAs in human cancer. *Signal Transduct. Target. Ther.* **1**, 1–9 (2016).
199. Mitchell, P. S. *et al.* Circulating microRNAs as stable blood-based markers for cancer detection. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 10513–10518 (2008).
200. Pérez, D. de M. *et al.* Extracellular vesicle-miRNAs as liquid biopsy biomarkers for disease identification and prognosis in metastatic colorectal cancer patients. *Sci. Rep.* **10**, 1–13 (2020).

201. Arroyo, J. D. *et al.* Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc. Natl. Acad. Sci.* **108**, 5003–5008 (2011).
202. Koh, W. *et al.* Noninvasive in vivo monitoring of tissue-specific global gene expression in humans. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 7361–7366 (2014).
203. Souza, M. F. de *et al.* Circulating mRNAs and miRNAs as candidate markers for the diagnosis and prognosis of prostate cancer. *PLOS ONE* **12**, e0184094 (2017).
204. Umu, S. U. *et al.* A comprehensive profile of circulating RNAs in human serum. *RNA Biol.* **15**, 242–250 (2017).
205. Xue, V. W. *et al.* Non-invasive Potential Circulating mRNA Markers for Colorectal Adenoma Using Targeted Sequencing. *Sci. Rep.* **9**, 1–10 (2019).
206. Nilsson, R. J. A. *et al.* Blood platelets contain tumor-derived RNA biomarkers. *Blood* **118**, 3680–3683 (2011).
207. Sol, N. & Wurdinger, T. Platelet RNA signatures for the detection of cancer. *Cancer Metastasis Rev.* **36**, 263–272 (2017).
208. Yao, B. *et al.* Delivery of platelet TPM3 mRNA into breast cancer cells via microvesicles enhances metastasis. *FEBS Open Bio* **9**, 2159–2169 (2019).
209. Fan, H. C., Blumenfeld, Y. J., Chitkara, U., Hudgins, L. & Quake, S. R. Analysis of the Size Distributions of Fetal and Maternal Cell-Free DNA by Paired-End Sequencing. *Clin. Chem.* **56**, 1279–1286 (2010).
210. Yu, S. C. Y. *et al.* Size-based molecular diagnostics using plasma DNA for noninvasive prenatal testing. *Proc. Natl. Acad. Sci.* **111**, 8583–8588 (2014).
211. Underhill, H. R. *et al.* Fragment Length of Circulating Tumor DNA. *PLOS Genet.* **12**, e1006162 (2016).
212. Hellwig, S. *et al.* Automated size selection for short cell-free DNA fragments enriches for circulating tumor DNA and improves error correction during next generation sequencing. *PLOS ONE* **13**, e0197333 (2018).
213. Mouliere, F. *et al.* Enhanced detection of circulating tumor DNA by fragment size analysis. *Sci. Transl. Med.* **10**, (2018).
214. Anker, P., Stroun, M. & Maurice, P. A. Spontaneous release of DNA by human blood lymphocytes as shown in an in vitro system. *Cancer Res.* **35**, 2375–2382 (1975).
215. Fournié, G. J. *et al.* Plasma DNA as a marker of cancerous cell death. Investigations in patients suffering from lung cancer and in nude mice bearing human tumours. *Cancer Lett.* **91**, 221–227 (1995).
216. Bronkhorst, A. J., Ungerer, V. & Holdenrieder, S. The emerging role of cell-free DNA as a molecular marker for cancer management. *Biomol. Detect. Quantif.* **17**, 100087 (2019).
217. Wu, T.-L. *et al.* Cell-free DNA: measurement in various carcinomas and establishment of normal reference range. *Clin. Chim. Acta* **321**, 77–87 (2002).
218. Vaart, M. V. D. & Pretorius, P. J. Circulating DNA. *Ann. N. Y. Acad. Sci.* **1137**, 18–26 (2008).
219. Aucamp, J., Bronkhorst, A. J., Badenhorst, C. P. S. & Pretorius, P. J. The diverse origins of circulating cell-free DNA in the human body: a critical re-evaluation of the literature. *Biol. Rev.* **93**, 1649–1683 (2018).
220. Ariga, H. *et al.* Kinetics of fetal cellular and cell-free DNA in the maternal circulation during and after pregnancy: implications for noninvasive prenatal diagnosis. *Transfusion (Paris)* **41**, 1524–1530 (2001).
221. To, E. W. H. *et al.* Rapid Clearance of Plasma Epstein-Barr Virus DNA After Surgical Treatment of Nasopharyngeal Carcinoma. *Clin. Cancer Res.* **9**, 3254–3259 (2003).

222. Diehl, F. *et al.* Circulating mutant DNA to assess tumor dynamics. *Nat. Med.* **14**, 985–990 (2008).
223. Yao, W., Mei, C., Nan, X. & Hui, L. Evaluation and comparison of in vitro degradation kinetics of DNA in serum, urine and saliva: A qualitative study. *Gene* **590**, 142–148 (2016).
224. Tamkovich, S. N. *et al.* Circulating DNA and DNase Activity in Human Blood. *Ann. N. Y. Acad. Sci.* **1075**, 191–196 (2006).
225. Botezatu, I. *et al.* Genetic analysis of DNA excreted in urine: a new approach for detecting specific genomic DNA sequences from cells dying in an organism. *Clin. Chem.* **46**, 1078–1084 (2000).
226. Diehl, F. *et al.* Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 16368–16373 (2005).
227. Khier, S. & Lohan, L. Kinetics of circulating cell-free DNA for biomedical applications: critical appraisal of the literature. *Future Sci. OA* **4**, (2018).
228. Pisetsky, D. S. & Fairhurst, A.-M. The origin of extracellular DNA during the clearance of dead and dying cells. *Autoimmunity* **40**, 281–284 (2007).
229. Ivanov, M., Baranova, A., Butler, T., Spellman, P. & Mileyko, V. Non-random fragmentation patterns in circulating cell-free DNA reflect epigenetic regulation. *BMC Genomics* **16 Suppl 13**, S1 (2015).
230. Snyder, M. W., Kircher, M., Hill, A. J., Daza, R. M. & Shendure, J. Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. *Cell* **164**, 57–68 (2016).
231. Javillier, M. & Fabrykant, M. Recherches experimentales sur le phosphore sanguin et particulièrement sur variations de la phosphatémie. *Bull. Société Chim. Fr.* **13**, (1931).
232. Watson, J. D. & Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).
233. Mandel, P. & Metais, P. Les acides nucléiques du plasma sanguin chez l’Homme. *C. R. Seances Soc. Biol. Fil.* **142**, 241–243 (1948).
234. Schmidt, G. & Thannhauser, S. J. A method for the determination of desoxyribonucleic acid, ribonucleic acid, and phosphoproteins in animal tissues. *J. Biol. Chem.* **161**, 83–89 (1945).
235. Koffler, D., Agnello, V., Winchester, R. & Kunkel, H. G. The Occurrence of Single-Stranded DNA in the Serum of Patients with Systemic Lupus Erythematosus and Other Diseases. *J. Clin. Invest.* **52**, 198–204 (1973).
236. Leon, S. A., Shapiro, B., Sklaroff, D. M. & Yaros, M. J. Free DNA in the serum of cancer patients and the effect of therapy. *Cancer Res.* **37**, 646–650 (1977).
237. Shapiro, B., Chakrabarty, M., Cohn, E. M. & Leon, S. A. Determination of circulating DNA levels in patients with benign or malignant gastrointestinal disease. *Cancer* **51**, 2116–2120 (1983).
238. Stroun, M., Anker, P., Lyautey, J., Lederrey, C. & Maurice, P. A. Isolation and characterization of DNA from the plasma of cancer patients. *Eur. J. Cancer Clin. Oncol.* **23**, 707–712 (1987).
239. Maebo, A. [Plasma DNA level as a tumor marker in primary lung cancer]. *Nihon Kyobu Shikkan Gakkai Zasshi* **28**, 1085–1091 (1990).
240. Vasioukhin, V. *et al.* Point mutations of the N-ras gene in the blood plasma DNA of patients with myelodysplastic syndrome or acute myelogenous leukaemia. *Br. J. Haematol.* **86**, 774–779 (1994).
241. Anker, P. *et al.* K-ras mutations are found in DNA extracted from the plasma of patients with colorectal cancer. *Gastroenterology* **112**, 1114–1120 (1997).

242. Kopreski, M. S. *et al.* Detection of mutant K-ras DNA in plasma or serum of patients with colorectal cancer. *Br. J. Cancer* **76**, 1293–1299 (1997).
243. Nawroz, H., Koch, W., Anker, P., Stroun, M. & Sidransky, D. Microsatellite alterations in serum DNA of head and neck cancer patients. *Nat. Med.* **2**, 1035–1037 (1996).
244. Walknowska, J., Conte, F. A. & Grumbach, M. M. Practical and theoretical implications of fetal-maternal lymphocyte transfer. *Lancet Lond. Engl.* **1**, 1119–1122 (1969).
245. Lo, Y. M. *et al.* Prenatal sex determination by DNA amplification from maternal peripheral blood. *Lancet Lond. Engl.* **2**, 1363–1365 (1989).
246. Lo, Y. M. *et al.* Presence of fetal DNA in maternal plasma and serum. *Lancet Lond. Engl.* **350**, 485–487 (1997).
247. Tjoa, M. L., Cindrova-Davies, T., Spasic-Boskovic, O., Bianchi, D. W. & Burton, G. J. Trophoblastic oxidative stress and the release of cell-free feto-placental DNA. *Am. J. Pathol.* **169**, 400–404 (2006).
248. Alberry, M. *et al.* Free fetal DNA in maternal plasma in anembryonic pregnancies: confirmation that the origin is the trophoblast. *Prenat. Diagn.* **27**, 415–418 (2007).
249. Breveglieri, G., D'Aversa, E., Finotti, A. & Borgatti, M. Non-invasive Prenatal Testing Using Fetal DNA. *Mol. Diagn. Ther.* **23**, 291–299 (2019).
250. Daniels, G., Finning, K., Martin, P. & Massey, E. Noninvasive prenatal diagnosis of fetal blood group phenotypes: current practice and future prospects. *Prenat. Diagn.* **29**, 101–107 (2009).
251. Costa, J.-M., Benachi, A. & Gautier, E. New strategy for prenatal diagnosis of X-linked disorders. *N. Engl. J. Med.* **346**, 1502 (2002).
252. Wright, C. F. & Burton, H. The use of cell-free fetal nucleic acids in maternal blood for non-invasive prenatal diagnosis. *Hum. Reprod. Update* **15**, 139–151 (2009).
253. Tong, Y. K. *et al.* Noninvasive prenatal detection of fetal trisomy 18 by epigenetic allelic ratio analysis in maternal plasma: Theoretical and empirical considerations. *Clin. Chem.* **52**, 2194–2202 (2006).
254. Chim, S. S. C. *et al.* Systematic search for placental DNA-methylation markers on chromosome 21: toward a maternal plasma-based epigenetic test for fetal trisomy 21. *Clin. Chem.* **54**, 500–511 (2008).
255. Lo, Y. M. D. *et al.* Digital PCR for the molecular detection of fetal chromosomal aneuploidy. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 13116–13121 (2007).
256. Fan, H. C., Blumenfeld, Y. J., Chitkara, U., Hudgins, L. & Quake, S. R. Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 16266–16271 (2008).
257. Chiu, R. W. K. *et al.* Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 20458–20463 (2008).
258. Palomaki, G. E. *et al.* DNA sequencing of maternal plasma reliably identifies trisomy 18 and trisomy 13 as well as Down syndrome: an international collaborative study. *Genet. Med. Off. J. Am. Coll. Med. Genet.* **14**, 296–305 (2012).
259. Osborne, C. M. *et al.* Discordant noninvasive prenatal testing results in a patient subsequently diagnosed with metastatic disease. *Prenat. Diagn.* **33**, 609–611 (2013).
260. Dharajiya, N. G. *et al.* Incidental Detection of Maternal Neoplasia in Noninvasive Prenatal Testing. *Clin. Chem.* **64**, 329–335 (2018).
261. Domínguez-Vigil, I. G., Moreno-Martínez, A. K., Wang, J. Y., Roehrl, M. H. A. & Barrera-Saldaña, H. A. The dawn of the liquid biopsy in the fight against cancer. *Oncotarget* **9**, 2912–2922 (2018).

262. Burnham, P., Khush, K. & De Vlaminc, I. Myriad Applications of Circulating Cell-Free DNA in Precision Organ Transplant Monitoring. *Ann. Am. Thorac. Soc.* **14**, S237–S241 (2017).
263. Lo, Y. M. *et al.* Presence of donor-specific DNA in plasma of kidney and liver-transplant recipients. *Lancet Lond. Engl.* **351**, 1329–1330 (1998).
264. Lui, Y. Y. N. *et al.* Origin of plasma cell-free DNA after solid organ transplantation. *Clin. Chem.* **49**, 495–496 (2003).
265. Snyder, T. M., Khush, K. K., Valentine, H. A. & Quake, S. R. Universal noninvasive detection of solid organ transplant rejection. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 6229–6234 (2011).
266. Knight, S. R., Thorne, A. & Lo Faro, M. L. Donor-specific Cell-free DNA as a Biomarker in Solid Organ Transplantation. A Systematic Review. *Transplantation* **103**, 273–283 (2019).
267. De Vlaminc, I. *et al.* Temporal response of the human virome to immunosuppression and antiviral therapy. *Cell* **155**, 1178–1187 (2013).
268. Moustafa, A. *et al.* The blood DNA virome in 8,000 humans. *PLoS Pathog.* **13**, e1006292 (2017).
269. Huang, Y.-F. *et al.* Analysis of microbial sequences in plasma cell-free DNA for early-onset breast cancer patients and healthy females. *BMC Med. Genomics* **11**, (2018).
270. Burd, E. M. Human papillomavirus and cervical cancer. *Clin. Microbiol. Rev.* **16**, 1–17 (2003).
271. Holmes, A. *et al.* Mechanistic signatures of HPV insertions in cervical carcinomas. *Npj Genomic Med.* **1**, 1–16 (2016).
272. Vogelstein, B. & Kinzler, K. W. Digital PCR. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 9236–9241 (1999).
273. Zhu, Z. *et al.* Single-molecule emulsion PCR in microfluidic droplets. *Anal. Bioanal. Chem.* **403**, 2127–2143 (2012).
274. Tsao, S. C.-H. *et al.* Monitoring response to therapy in melanoma by quantifying circulating tumour DNA with droplet digital PCR for BRAF and NRAS mutations. *Sci. Rep.* **5**, 11198 (2015).
275. Garcia-Murillas, I. *et al.* Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Sci. Transl. Med.* **7**, 302ra133 (2015).
276. Seoane, J., De Mattos-Arruda, L., Le Rhun, E., Bardelli, A. & Weller, M. Cerebrospinal fluid cell-free tumour DNA as a liquid biopsy for primary brain tumours and central nervous system metastases. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **30**, 211–218 (2019).
277. Ballester, L. Y. *et al.* Evaluating Circulating Tumor DNA From the Cerebrospinal Fluid of Patients With Melanoma and Leptomeningeal Disease. *J. Neuropathol. Exp. Neurol.* **77**, 628–635 (2018).
278. Martínez-Ricarte, F. *et al.* Molecular Diagnosis of Diffuse Gliomas through Sequencing of Cell-Free Circulating Tumor DNA from Cerebrospinal Fluid. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **24**, 2812–2819 (2018).
279. Avogbe, P. H. *et al.* Urinary TERT promoter mutations as non-invasive biomarkers for the comprehensive detection of urothelial cancer. *EBioMedicine* **44**, 431–438 (2019).
280. Hosen, M. I. *et al.* Urinary TERT promoter mutations are detectable up to 10 years prior to clinical diagnosis of bladder cancer: Evidence from the Golestan Cohort Study. *EBioMedicine* **53**, (2020).
281. Garlan, F. *et al.* Early Evaluation of Circulating Tumor DNA as Marker of Therapeutic Efficacy in Metastatic Colorectal Cancer Patients (PLACOL Study). *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **23**, 5416–5425 (2017).

282. Riva, F. *et al.* Patient-Specific Circulating Tumor DNA Detection during Neoadjuvant Chemotherapy in Triple-Negative Breast Cancer. *Clin. Chem.* **63**, 691–699 (2017).
283. Decraene, C. *et al.* Single Droplet Digital Polymerase Chain Reaction for Comprehensive and Simultaneous Detection of Mutations in Hotspot Regions. *J. Vis. Exp. JoVE* (2018) doi:10.3791/58051.
284. Silveira, A. B. *et al.* High-Accuracy Determination of Microsatellite Instability Compatible with Liquid Biopsies. *Clin. Chem.* (2020) doi:10.1093/clinchem/hvaa013.
285. Kinde, I., Wu, J., Papadopoulos, N., Kinzler, K. W. & Vogelstein, B. Detection and quantification of rare mutations with massively parallel sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 9530–9535 (2011).
286. Forsheew, T. *et al.* Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. *Sci. Transl. Med.* **4**, 136ra68 (2012).
287. Pécuchet, N. *et al.* Analysis of Base-Position Error Rate of Next-Generation Sequencing to Detect Tumor Mutations in Circulating DNA. *Clin. Chem.* **62**, 1492–1503 (2016).
288. Newman, A. M. *et al.* An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat. Med.* **20**, 548–554 (2014).
289. Pentsova, E. I. *et al.* Evaluating Cancer of the Central Nervous System Through Next-Generation Sequencing of Cerebrospinal Fluid. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **34**, 2404–2415 (2016).
290. Bettgowda, C. *et al.* Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci. Transl. Med.* **6**, 224ra24 (2014).
291. Leary, R. J. *et al.* Detection of chromosomal alterations in the circulation of cancer patients with whole-genome sequencing. *Sci. Transl. Med.* **4**, 162ra154 (2012).
292. Klevebring, D. *et al.* Evaluation of exome sequencing to estimate tumor burden in plasma. *PloS One* **9**, e104417 (2014).
293. Dietz, S. *et al.* Low Input Whole-Exome Sequencing to Determine the Representation of the Tumor Exome in Circulating DNA of Non-Small Cell Lung Cancer Patients. *PloS One* **11**, e0161012 (2016).
294. Koepfel, F. *et al.* Whole exome sequencing for determination of tumor mutation load in liquid biopsy from advanced cancer patients. *PloS One* **12**, e0188174 (2017).
295. Butler, T. M. *et al.* Exome Sequencing of Cell-Free DNA from Metastatic Cancer Patients Identifies Clinically Actionable Mutations Distinct from Primary Disease. *PloS One* **10**, e0136407 (2015).
296. Murtaza, M. *et al.* Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature* **497**, 108–112 (2013).
297. Fernandez, A. F. *et al.* A DNA methylation fingerprint of 1628 human samples. *Genome Res.* **22**, 407–419 (2012).
298. Schultz, M. D. *et al.* Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature* **523**, 212–216 (2015).
299. Hoadley, K. A. *et al.* Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* **173**, 291–304.e6 (2018).
300. Koelsche, C. *et al.* Array-based DNA-methylation profiling in sarcomas with small blue round cell histology provides valuable diagnostic information. *Mod. Pathol.* **31**, 1246–1256 (2018).
301. Sturm, D. *et al.* New Brain Tumor Entities Emerge from Molecular Classification of CNS-PNETs. *Cell* **164**, 1060–1072 (2016).
302. Capper, D. *et al.* DNA methylation-based classification of central nervous system tumours. *Nature* **555**, 469–474 (2018).

303. Kang, S. *et al.* CancerLocator: non-invasive cancer diagnosis and tissue-of-origin prediction using methylation profiles of cell-free DNA. *Genome Biol.* **18**, 53 (2017).
304. Moss, J. *et al.* Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. *Nat. Commun.* **9**, 1–12 (2018).
305. Guo, S. *et al.* Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA. *Nat. Genet.* **49**, 635–642 (2017).
306. Shen, S. Y. *et al.* Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature* **563**, 579–583 (2018).
307. Phallen, J. *et al.* Direct detection of early-stage cancers using circulating tumor DNA. *Sci. Transl. Med.* **9**, (2017).
308. Cristiano, S. *et al.* Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature* **570**, 385–389 (2019).
309. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213 (2013).
310. Adalsteinsson, V. A. *et al.* Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. *Nat. Commun.* **8**, 1324 (2017).
311. Murtaza, M. *et al.* Multifocal clonal evolution characterized using circulating tumour DNA in a case of metastatic breast cancer. *Nat. Commun.* **6**, 8760 (2015).
312. De Mattos-Arruda, L. *et al.* Cerebrospinal fluid-derived circulating tumour DNA better represents the genomic alterations of brain tumours than plasma. *Nat. Commun.* **6**, 8839 (2015).
313. Genovese, G. *et al.* Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *N. Engl. J. Med.* **371**, 2477–2487 (2014).
314. Razavi, P. *et al.* High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. *Nat. Med.* **25**, 1928–1937 (2019).
315. Chabon, J. J. *et al.* Integrating genomic features for non-invasive early lung cancer detection. *Nature* **580**, 245–251 (2020).
316. Ulz, P. *et al.* Inferring expressed genes by whole-genome sequencing of plasma DNA. *Nat. Genet.* **48**, 1273–1278 (2016).
317. Song, C.-X. *et al.* 5-Hydroxymethylcytosine signatures in cell-free DNA provide information about tumor types and stages. *Cell Res.* **27**, 1231–1242 (2017).
318. Zeng, H., He, B., Yi, C. & Peng, J. Liquid biopsies: DNA methylation analyses in circulating cell-free DNA. *J. Genet. Genomics Yi Chuan Xue Bao* **45**, 185–192 (2018).
319. Li, W. *et al.* CancerDetector: ultrasensitive and non-invasive cancer detection at the resolution of individual reads using cell-free DNA methylation sequencing data. *Nucleic Acids Res.* **46**, e89–e89 (2018).
320. Sadeh, R. *et al.* ChIP-seq of plasma cell-free nucleosomes identifies cell-of-origin gene expression programs. *bioRxiv* 638643 (2019) doi:10.1101/638643.
321. Combaret, V. *et al.* Circulating MYCN DNA as a Tumor-specific Marker in Neuroblastoma Patients. *Cancer Res.* **62**, 3646–3648 (2002).
322. Combaret, V. *et al.* Determination of 17q gain in patients with neuroblastoma by analysis of circulating DNA. *Pediatr. Blood Cancer* **56**, 757–761 (2011).
323. Combaret, V. *et al.* Detection of tumor ALK status in neuroblastoma patients using peripheral blood. *Cancer Med.* **4**, 540–550 (2015).
324. Chicard, M. *et al.* Genomic Copy Number Profiling Using Circulating Free Tumor DNA Highlights Heterogeneity in Neuroblastoma. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **22**, 5564–5573 (2016).

325. Van Roy, N. *et al.* Shallow Whole Genome Sequencing on Circulating Cell-Free DNA Allows Reliable Noninvasive Copy-Number Profiling in Neuroblastoma Patients. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **23**, 6305–6314 (2017).
326. Klega, K. *et al.* Detection of Somatic Structural Variants Enables Quantification and Characterization of Circulating Tumor DNA in Children With Solid Tumors. *JCO Precis. Oncol.* **2018**, (2018).
327. Morini, M. *et al.* Exosomal microRNAs from Longitudinal Liquid Biopsies for the Prediction of Response to Induction Chemotherapy in High-Risk Neuroblastoma Patients: A Proof of Concept SIOPEX Study II. *Cancers* **11**, 1476 (2019).
328. Misawa, A. *et al.* RASSF1A hypermethylation in pretreatment serum DNA of neuroblastoma patients: a prognostic marker. *Br. J. Cancer* **100**, 399–404 (2009).
329. Berger, M. *et al.* Genomic EWS-FLI1 fusion sequences in Ewing sarcoma resemble breakpoint characteristics of immature lymphoid malignancies. *PloS One* **8**, e56408 (2013).
330. Krumbholz, M. *et al.* Genomic EWSR1 Fusion Sequence as Highly Sensitive and Dynamic Plasma Tumor Marker in Ewing Sarcoma. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **22**, 4356–4365 (2016).
331. Eguchi- Ishimae, M. *et al.* Early detection of the PAX3-FOXO1 fusion gene in circulating tumor-derived DNA in a case of alveolar rhabdomyosarcoma. *Genes. Chromosomes Cancer* **58**, 521–529 (2019).
332. Barris, D. M. *et al.* Detection of circulating tumor DNA in patients with osteosarcoma. *Oncotarget* **9**, 12695–12704 (2018).
333. Shulman, D. S. *et al.* Detection of circulating tumour DNA is associated with inferior outcomes in Ewing sarcoma and osteosarcoma: a report from the Children’s Oncology Group. *Br. J. Cancer* **119**, 615–621 (2018).
334. Jiménez, I. *et al.* Circulating tumor DNA analysis enables molecular characterization of pediatric renal tumors at diagnosis. *Int. J. Cancer* **144**, 68–79 (2019).
335. Kurihara, S. *et al.* Circulating free DNA as non-invasive diagnostic biomarker for childhood solid tumors. *J. Pediatr. Surg.* **50**, 2094–2097 (2015).
336. Huang, T. Y. *et al.* Detection of Histone H3 mutations in cerebrospinal fluid-derived tumor DNA from children with diffuse midline glioma. *Acta Neuropathol. Commun.* **5**, (2017).
337. Xie, M. *et al.* Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* **20**, 1472–1478 (2014).
338. Risberg, B. *et al.* Effects of Collection and Processing Procedures on Plasma Circulating Cell-Free DNA from Cancer Patients. *J. Mol. Diagn. JMD* **20**, 883–892 (2018).
339. Markus, H. *et al.* Evaluation of pre-analytical factors affecting plasma DNA analysis. *Sci. Rep.* **8**, (2018).
340. Trigg, R. M., Martinson, L. J., Parpart-Li, S. & Shaw, J. A. Factors that influence quality and yield of circulating-free DNA: A systematic review of the methodology literature. *Heliyon* **4**, (2018).
341. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
342. Van der Auwera, G. A. *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinforma.* **43**, 11.10.1–11.10.33 (2013).
343. McKenna, A. *et al.* The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
344. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).

345. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
346. Chun, S. & Fay, J. C. Identification of deleterious mutations within three human genomes. *Genome Res.* **19**, 1553–1561 (2009).
347. Schwarz, J. M., Rödelberger, C., Schuelke, M. & Seelow, D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods* **7**, 575–576 (2010).
348. Ng, P. C. & Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **31**, 3812–3814 (2003).
349. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
350. Koboldt, D. C. *et al.* VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012).
351. Venkatraman, E. S. & Olshen, A. B. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinforma. Oxf. Engl.* **23**, 657–663 (2007).
352. Favero, F. *et al.* Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **26**, 64–70 (2015).
353. Chicard, M. *et al.* Whole-Exome Sequencing of Cell-Free DNA Reveals Temporo-spatial Heterogeneity and Identifies Treatment-Resistant Clones in Neuroblastoma. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **24**, 939–949 (2018).
354. Lee, H. *et al.* Improving the efficiency of genomic loci capture using oligonucleotide arrays for high throughput resequencing. *BMC Genomics* **10**, 646 (2009).
355. Jiao, W., Vembu, S., Deshwar, A. G., Stein, L. & Morris, Q. Inferring clonal evolution of tumors from single nucleotide somatic mutations. *BMC Bioinformatics* **15**, 35 (2014).
356. Miller, C. A. *et al.* Visualizing tumor evolution with the fishplot package for R. *BMC Genomics* **17**, 880 (2016).
357. Shen, R. & Seshan, V. E. FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. *Nucleic Acids Res.* **44**, e131 (2016).
358. Gharibi, A. *et al.* ITGA1 is a pre-malignant biomarker that promotes therapy resistance and metastatic potential in pancreatic cancer. *Sci. Rep.* **7**, (2017).
359. O'Day, E. M., Idos, G. E., Hill, C., Chen, J. W. & Wagner, G. Cytidine monophosphate N-acetylneuraminic acid synthetase enhances invasion of human triple-negative breast cancer cells. *OncoTargets Ther.* **11**, 6827–6838 (2018).
360. Chen, Z. X. *et al.* RNA helicase A is a downstream mediator of KIF1B β tumor-suppressor function in neuroblastoma. *Cancer Discov.* **4**, 434–451 (2014).
361. Wang, H.-W. *et al.* Pediatric primary central nervous system germ cell tumors of different prognosis groups show characteristic miRNome traits and chromosome copy number variations. *BMC Genomics* **11**, 132 (2010).
362. Zheng, W., Jiang, C. & Li, R. Integrin and gene network analysis reveals that ITGA5 and ITGB1 are prognostic in non-small-cell lung cancer. *OncoTargets Ther.* **9**, 2317–2327 (2016).
363. Xu, M.-D. *et al.* Genomic characteristics of pancreatic squamous cell carcinoma, an investigation by using high throughput sequencing after in-solution hybrid capture. *Oncotarget* **8**, 14620–14635 (2017).
364. Westerman, B. A. *et al.* Basic helix-loop-helix transcription factor profiling of lung tumors shows aberrant expression of the proneural gene atonal homolog 1 (ATOH1, HATH1, MATH1) in neuroendocrine tumors. *Int. J. Biol. Markers* **22**, 114–123 (2007).

365. Nguyen, P. N. N., Huang, C.-J., Sugii, S., Cheong, S. K. & Choo, K. B. Selective activation of miRNAs of the primate-specific chromosome 19 miRNA cluster (C19MC) in cancer and stem cells and possible contribution to regulation of apoptosis. *J. Biomed. Sci.* **24**, (2017).
366. Mong, E. F. *et al.* Chromosome 19 microRNA cluster enhances cell reprogramming by inhibiting epithelial-to-mesenchymal transition. *Sci. Rep.* **10**, 1–13 (2020).
367. Lambo, S. *et al.* The molecular landscape of ETMR at diagnosis and relapse. *Nature* **576**, 274–280 (2019).
368. Huang, C., Jia, P., Chastain, M., Shiva, O. & Chai, W. The human CTC1/STN1/TEN1 complex regulates telomere maintenance in ALT cancer cells. *Exp. Cell Res.* **355**, 95–104 (2017).
369. Weber-Hall, S. *et al.* Gains, losses, and amplification of genomic material in rhabdomyosarcoma analyzed by comparative genomic hybridization. *Cancer Res.* **56**, 3220–3224 (1996).
370. Parham, D. M. & Barr, F. G. Classification of Rhabdomyosarcoma and Its Molecular Basis. *Adv. Anat. Pathol.* **20**, 387–397 (2013).
371. Nishimura, R. *et al.* Characterization of genetic lesions in rhabdomyosarcoma using a high-density single nucleotide polymorphism array. *Cancer Sci.* **104**, 856–864 (2013).
372. Shukla, N. N. *et al.* Plasma DNA-based molecular diagnosis, prognostication, and monitoring of patients with EWSR1 fusion-positive sarcomas. *JCO Precis. Oncol.* **2017**, (2017).
373. Hayashi, M. *et al.* Highly Personalized Detection of Minimal Ewing Sarcoma Disease Burden from Plasma Tumor DNA. *Cancer* **122**, 3015–3023 (2016).
374. SU, Y.-H., KIM, A. K. & JAIN, S. Liquid biopsies for hepatocellular carcinoma. *Transl. Res. J. Lab. Clin. Med.* **201**, 84–97 (2018).
375. Brohl, A. S., Kahen, E., Yoder, S. J., Teer, J. K. & Reed, D. R. The genomic landscape of malignant peripheral nerve sheath tumors: diverse drivers of Ras pathway activation. *Sci. Rep.* **7**, (2017).
376. Pinto, E. M. *et al.* Malignant rhabdoid tumors originating within and outside the central nervous system are clinically and molecularly heterogeneous. *Acta Neuropathol. (Berl.)* **136**, 315–326 (2018).
377. Croft, D. *et al.* Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* **39**, D691–D697 (2011).
378. Lamoliatte, F., McManus, F. P., Maarifi, G., Chelbi-Alix, M. K. & Thibault, P. Uncovering the SUMOylation and ubiquitylation crosstalk in human cells using sequential peptide immunopurification. *Nat. Commun.* **8**, 14109 (2017).
379. Jürgen Dohmen, R. SUMO protein modification. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1695**, 113–131 (2004).
380. Han, Z.-J., Feng, Y.-H., Gu, B.-H., Li, Y.-M. & Chen, H. The post-translational modification, SUMOylation, and cancer (Review). *Int. J. Oncol.* **52**, 1081–1094 (2018).
381. Zhu, S., Sachdeva, M., Wu, F., Lu, Z. & Mo, Y.-Y. Ubc9 promotes breast cell invasion and metastasis in a sumoylation-independent manner. *Oncogene* **29**, 1763–1772 (2010).
382. Zhao, X.-Y. *et al.* Hypoxia inducible factor-1 mediates expression of galectin-1: the potential role in migration/invasion of colorectal cancer cells. *Carcinogenesis* **31**, 1367–1375 (2010).
383. Rytinki, M. M. & Palvimo, J. J. SUMO wrestling in cell movement. *Cell Res.* **21**, 3–5 (2011).
384. Wang, Q. *et al.* SUMO-specific protease 1 promotes prostate cancer progression and metastasis. *Oncogene* **32**, 2493–2498 (2013).

385. Wang, Z., Jin, J., Zhang, J., Wang, L. & Cao, J. Depletion of SENP1 suppresses the proliferation and invasion of triple-negative breast cancer cells. *Oncol. Rep.* **36**, 2071–2078 (2016).
386. Zhang, W. *et al.* SENP1 regulates hepatocyte growth factor-induced migration and epithelial-mesenchymal transition of hepatocellular carcinoma. *Tumour Biol. J. Int. Soc. Oncodevelopmental Biol. Med.* **37**, 7741–7748 (2016).
387. Li, C. *et al.* Quantitative SUMO proteomics identifies PIAS1 substrates involved in cell migration and motility. *Nat. Commun.* **11**, 1–14 (2020).
388. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
389. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338 (2019).
390. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).
391. Mootha, V. K. *et al.* PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
392. Butler, T. M. *et al.* Circulating tumor DNA dynamics using patient-customized assays are associated with outcome in neoadjuvantly treated breast cancer. *Cold Spring Harb. Mol. Case Stud.* **5**, (2019).
393. Magbanua, M. J. M. *et al.* Circulating tumor DNA in neoadjuvant treated breast cancer reflects response and survival. *medRxiv* 2020.02.03.20019760 (2020) doi:10.1101/2020.02.03.20019760.
394. Dang, H. X. *et al.* ClonEvol: clonal ordering and visualization in cancer sequencing. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **28**, 3076–3082 (2017).
395. Abécassis, J. *et al.* Assessing reliability of intra-tumor heterogeneity estimates from single sample whole exome sequencing data. *PloS One* **14**, e0224143 (2019).
396. Kent, W. J. *et al.* The Human Genome Browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
397. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
398. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
399. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
400. Lizio, M. *et al.* Gateways to the FANTOM5 promoter level mammalian expression atlas. *Genome Biol.* **16**, 22 (2015).
401. Valent, A. *et al.* MYCN gene overrepresentation detected in primary neuroblastoma tumour cells without amplification. *J. Pathol.* **198**, 495–501 (2002).
402. Peirce, S. K. & Findley, H. W. High level MycN expression in non-MYCN amplified neuroblastoma is induced by the combination treatment nutlin-3 and doxorubicin and enhances chemosensitivity. *Oncol. Rep.* **22**, 1443–1449 (2009).
403. Hsu, C.-L. *et al.* Unveiling MYCN regulatory networks in neuroblastoma via integrative analysis of heterogeneous genomics data. *Oncotarget* **7**, 36293–36310 (2016).
404. Uhlén, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
405. GEZER, U. *et al.* Correlation of histone methyl marks with circulating nucleosomes in blood plasma of cancer patients. *Oncol. Lett.* **3**, 1095–1098 (2012).
406. Molenaar, J. J. *et al.* Cyclin D1 and CDK4 activity contribute to the undifferentiated phenotype in neuroblastoma. *Cancer Res.* **68**, 2599–2609 (2008).

407. Selmi, A. *et al.* TWIST1 is a direct transcriptional target of MYCN and MYC in neuroblastoma. *Cancer Lett.* **357**, 412–418 (2015).
408. Malouf, G. G. *et al.* Architecture of epigenetic reprogramming following Twist1-mediated epithelial-mesenchymal transition. *Genome Biol.* **14**, R144 (2013).
409. Malek, R. *et al.* TWIST1-WDR5-Hottip regulates Hoxa9 chromatin to facilitate prostate cancer metastasis. *Cancer Res.* **77**, 3181–3193 (2017).
410. Harikumar, A. & Meshorer, E. Chromatin remodeling and bivalent histone modifications in embryonic stem cells. *EMBO Rep.* **16**, 1609–1619 (2015).
411. Shlyueva, D., Stampfel, G. & Stark, A. Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.* **15**, 272–286 (2014).
412. Hao, N., Shearwin, K. E. & Dodd, I. B. Positive and Negative Control of Enhancer-Promoter Interactions by Other DNA Loops Generates Specificity and Tunability. *Cell Rep.* **26**, 2419–2433.e3 (2019).
413. Voigt, P., Tee, W.-W. & Reinberg, D. A double take on bivalent promoters. *Genes Dev.* **27**, 1318–1338 (2013).
414. Rogers, M. F. *et al.* FATHMM-XF: accurate prediction of pathogenic point mutations via extended features. *Bioinformatics* **34**, 511–513 (2018).
415. McLaren, W. *et al.* Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **26**, 2069–2070 (2010).
416. Carter, H., Douville, C., Stenson, P. D., Cooper, D. N. & Karchin, R. Identifying Mendelian disease genes with the Variant Effect Scoring Tool. *BMC Genomics* **14**, S3 (2013).
417. Tate, J. G. *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **47**, D941–D947 (2019).
418. Dommange-Romero, F., Collardeau-Frachon, S. & Hameury, F. Pleuropneumoblastome de l'enfant. *Bull. Cancer (Paris)* **97**, 1047–1052 (2010).
419. Shi, W. *et al.* Reliability of Whole-Exome Sequencing for Assessing Intratumor Genetic Heterogeneity. *Cell Rep.* **25**, 1446–1457 (2018).
420. Talevich, E., Shain, A. H., Botton, T. & Bastian, B. C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLOS Comput. Biol.* **12**, e1004873 (2016).
421. Huether, R. *et al.* The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. *Nat. Commun.* **5**, 3630 (2014).
422. Teng, H. *et al.* Prevalence and architecture of posttranscriptionally impaired synonymous mutations in 8,320 genomes across 22 cancer types. *Nucleic Acids Res.* **48**, 1192–1205 (2020).
423. Robertson, J. C., Jorcyk, C. L. & Oxford, J. T. DICER1 Syndrome: DICER1 Mutations in Rare Cancers. *Cancers* **10**, (2018).
424. de Bont, J. M. *et al.* Differential expression and prognostic significance of SOX genes in pediatric medulloblastoma and ependymoma identified by microarray analysis. *Neuro-Oncol.* **10**, 648–660 (2008).
425. de la Rocha, A. M. A., Sampron, N., Alonso, M. M. & Matheu, A. Role of SOX family of transcription factors in central nervous system tumors. *Am. J. Cancer Res.* **4**, 312–324 (2014).
426. Sin-Chan, P. *et al.* A C19MC-LIN28A-MYCN Oncogenic Circuit Driven by Hijacked Super-enhancers Is a Distinct Therapeutic Vulnerability in ETMRs: A Lethal Brain Tumor. *Cancer Cell* **36**, 51–67.e7 (2019).
427. Zhang, J. *et al.* Germline Mutations in Predisposition Genes in Pediatric Cancer. *N. Engl. J. Med.* **373**, 2336–2346 (2015).

428. Waszak, S. M. *et al.* Spectrum and prevalence of genetic predisposition in medulloblastoma: a retrospective genetic study and prospective validation in a clinical trial cohort. *Lancet Oncol.* **19**, 785–798 (2018).
429. Liang, X. *et al.* Targeted next-generation sequencing identifies clinically relevant somatic mutations in a large cohort of inflammatory breast cancer. *Breast Cancer Res. BCR* **20**, (2018).
430. Felgenhauer, K. Protein size and cerebrospinal fluid composition. *Klin. Wochenschr.* **52**, 1158–1164 (1974).
431. Wiederkehr, F., Ogilvie, A. & Vonderschmitt, D. J. Cerebrospinal fluid proteins studied by two-dimensional gel electrophoresis and immunoblotting technique. *J. Neurochem.* **49**, 363–372 (1987).
432. Schilde, L. M. *et al.* Protein variability in cerebrospinal fluid and its possible implications for neurological protein biomarker research. *PLoS ONE* **13**, (2018).
433. Yokota, M., Tatsumi, N., Nathalang, O., Yamada, T. & Tsuda, I. Effects of heparin on polymerase chain reaction for blood white cells. *J. Clin. Lab. Anal.* **13**, 133–140 (1999).
434. Schultz, K. A. *et al.* DICER1-pleuropulmonary blastoma familial tumor predisposition syndrome: a unique constellation of neoplastic conditions. *Pathol. Case Rev.* **19**, 90–100 (2014).
435. Williamson, C. T. *et al.* ATR inhibitors as a synthetic lethal therapy for tumours deficient in ARID1A. *Nat. Commun.* **7**, 13837 (2016).
436. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, research0034.1-research0034.11 (2002).
437. Johann, P. D. *et al.* Atypical Teratoid/Rhabdoid Tumors Are Comprised of Three Epigenetic Subgroups with Distinct Enhancer Landscapes. *Cancer Cell* **29**, 379–393 (2016).
438. Wu, G. *et al.* The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma. *Nat. Genet.* **46**, 444–450 (2014).
439. Miller, A. M. *et al.* Tracking tumour evolution in glioma through liquid biopsies of cerebrospinal fluid. *Nature* **565**, 654–658 (2019).
440. Behling, F. & Schittenhelm, J. Oncogenic BRAF Alterations and Their Role in Brain Tumors. *Cancers* **11**, (2019).
441. Torchia, J. *et al.* Molecular subgroups of atypical teratoid rhabdoid tumours in children: an integrated genomic and clinicopathological analysis. *Lancet Oncol.* **16**, 569–582 (2015).
442. Japp, A. S., Klein-Hitpass, L., Denkhaus, D. & Pietsch, T. OTX2 Defines a Subgroup of Atypical Teratoid Rhabdoid Tumors With Close Relationship to Choroid Plexus Tumors. *J. Neuropathol. Exp. Neurol.* **76**, 32–38 (2017).
443. Meissner, A. *et al.* Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33**, 5868–5877 (2005).
444. Koker, A. D., Paemel, R. V., Wilde, B. D., Preter, K. D. & Callewaert, N. A versatile method for circulating cell-free DNA methylome profiling by reduced representation bisulfite sequencing. *bioRxiv* 663195 (2019) doi:10.1101/663195.
445. Huertas-Martínez, J. *et al.* DNA methylation profiling identifies PTRF/Cavin-1 as a novel tumor suppressor in Ewing sarcoma when co-expressed with caveolin-1. *Cancer Lett.* **386**, 196–207 (2017).
446. Wu, S. P. *et al.* DNA Methylation–Based Classifier for Accurate Molecular Diagnosis of Bone Sarcomas. *JCO Precis. Oncol.* 1–11 (2017) doi:10.1200/PO.17.00031.
447. Seki, M. *et al.* Integrated genetic and epigenetic analysis defines novel molecular subgroups in rhabdomyosarcoma. *Nat. Commun.* **6**, 7557 (2015).

448. Almstrup, K. *et al.* Pubertal development in healthy children is mirrored by DNA methylation patterns in peripheral blood. *Sci. Rep.* **6**, 1–12 (2016).
449. Chan, K. C. A. *et al.* Noninvasive detection of cancer-associated genome-wide hypomethylation and copy number aberrations by plasma DNA bisulfite sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 18761–18768 (2013).
450. Paemel, R. V. *et al.* Minimally invasive classification of pediatric solid tumors using reduced representation bisulfite sequencing of cell-free DNA: a proof-of-principle study. *bioRxiv* 795047 (2020) doi:10.1101/795047.
451. Hu, Z., Li, Z., Ma, Z. & Curtis, C. Multi-cancer analysis of clonality and the timing of systemic spread in paired primary tumors and metastases. *Nat. Genet.* 1–8 (2020) doi:10.1038/s41588-020-0628-z.
452. Füllgrabe, J., Hajji, N. & Joseph, B. Cracking the death code: apoptosis-related histone modifications. *Cell Death Differ.* **17**, 1238–1243 (2010).
453. Th'ng, J. P. Histone modifications and apoptosis: cause or consequence? *Biochem. Cell Biol. Biochim. Biol. Cell.* **79**, 305–311 (2001).

Annexes

Tableau Annexe 1 : Descriptif de cas de l'étude évolution clonale et plasticité	1
Tableau Annexe 2 : Descriptif de cas de l'étude MappyAct	2
Tableau Annexe 3 : Descriptif de cas de l'étude LCR.....	5
Figure annexe 1 : Suivi des mutations en séquençage ciblé des cas NGS Kids	6

Article : Whole-Exome Sequencing of Cell-Free DNA Reveals Temporo-spatial Heterogeneity and Identifies Treatment-Resistant Clones in Neuroblastoma

Tableau Annexe 1 : Descriptif de cas de l'étude évolution clonale et plasticité

ID Patients	Center	Sex	Disease Group	Histology group level 1	Histology group level 2, VP	Genomic Analysis	(cDNA) par mL de plasma (ng/mL)										aSNV common	aSNV Tumor	aSNV Plasma
							Copy Number	Tumor contributit y	cDNA contributit y	Contribution %ctDNA	Number of Tumor SNVs	Number of Plasma SNVs							
MAP002	Gustave Roussy	Female	Sarcoma	other	other SC	PANEL	26.4	Yes	Both	0,26	4	8	12						
MAP003	Gustave Roussy	Male	Carcinoma	Carcinoma	CA	eRMS	26.4	no	None	0									
MAP004	Gustave Roussy	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	74.2	no	Plasma	0,28	0	0	18			HRAS			
MAP005	Gustave Roussy	Female	Neuroblastoma	Neuroblastoma	NBL		WES + RNAseq	52.0	Both	Both	0,56	12	74	26	ATR	CHD7			
MAP007	Gustave Roussy	Male	Osteosarcoma	Osteosarcoma	OS		WES + RNAseq	7.9	Yes	Both	0	70	11	0					
MAP009	Gustave Roussy	Male	Other	NRSTS	NRSTS	other NRSTS	13.7	Yes	No	Tumor	0,41	3	1	1					
MAP010	CHU Nantes	Male	Ewing Sarcoma	Ewing	EWS		WES + RNAseq	17.9	Yes	No	Tumor	0	26	0	0				
MAP011	La Timone Marseille	Male	Neuroblastoma	Neuroblastoma	NBL		WES + RNAseq	41.6	Yes	No	Tumor	0	31	2	0				
MAP012	Institut Curie Paris	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	38.5	Yes	Both	0,23	0	29	0		GLI1, MYCN, TACC1, TIA1				
MAP014	CHU Nantes	Female	Ewing Sarcoma	Ewing	EWS		WES + RNAseq	6.1	No	None	0								
MAP015	CHU Strasbourg	Female	Rhabdoid Tumor	Rhabdoid tumor	RT		WES + RNAseq	26.1	Yes	No	Tumor	0	11	2	0				
MAP016	Gustave Roussy	Male	Carcinoma	Carcinoma	CA		WES + RNAseq	11.5	Yes	Both	0	32	29	3	TP53, BRAF				
MAP017	CHU Nantes	Female	Ewing Sarcoma	Ewing	EWS		WES + RNAseq	19.1	No	Plasma	0,21	0	0	29					
MAP018	La Timone Marseille	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	106.4	Both	Both	0,4	3	58	3	FBXW7, MAP2K4, NRAS					
MAP019	CHU Nantes	Male	Other	NRSTS	NRSTS	other NRSTS	14.9	No	Yes	Plasma	0	0	0	5		MLL3			
MAP020	CHU Nantes	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	6.9	Both	Both	0,23	0	36	0	CDKN2A, PIK3CA, TP53					
MAP023	Institut Curie Paris	Female	Neuroblastoma	Neuroblastoma	NBL		WES + RNAseq	84.0	Both	Both	0,45	19	31	1	ATRX	ARI10B			
MAP025	CHU Nancy	Male	Neuroblastoma	Neuroblastoma	NBL		WES + RNAseq	9.1	no	None	0,3								
MAP027	Institut Curie Paris	Male	MPNST	NRSTS	NRSTS	MPNST	57.4	Yes	Both	Both	0,29	1	18	0	FGFR4				
MAP029	Institut Curie Paris	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	75.6	Yes	Both	Both	0,57	17	8	18		TCF12			
MAP030	La Timone Marseille	Male	Neuroblastoma	Neuroblastoma	NBL		WES	45.1	Plasma	Plasma	0,6	1	0	473		BRCA2, GFI1R, KANK1, MSH6, NUP214, PDGFRB, POLD1, RICTOR, SMARCB1, TP73			
MAP032	Gustave Roussy	Male	Neuroblastoma	Neuroblastoma	NBL		WES	16.2	no	None	0								
MAP033	Gustave Roussy	Male	Osteosarcoma	Osteosarcoma	OS		WES + RNAseq	23.0	Tumor	Yes	Both	0,81	6	26	1	0			
MAP035	CHU Nantes	Male	Neuroblastoma	Neuroblastoma	NBL		WES	1338.0	Both	Yes	Both	0	14	0	0				
MAP036	CHU Strasbourg	Male	Ewing Sarcoma	Ewing	EWS		WES + RNAseq	22.3	Yes	No	Tumor	0	14	0	1				
MAP037	La Timone Marseille	Female	Ewing Sarcoma	Ewing	EWS		WES + RNAseq	119.7	Both	Both	0,31	1	15	3					
MAP038	Centre Oscar Lamb	Female	Osteosarcoma	Osteosarcoma	OS		WES + RNAseq	33.0	Both	Both	0,32	20	27	3					
MAP040	Centre Léon Bérard	Female	Neuroblastoma	Neuroblastoma	NBL		WES + RNAseq	222.3	Both	Both	0,76	0	63	0	CDK5				
MAP042	CHU Nantes	Male	Nephroblastoma	Nephroblastoma	WT		WES + RNAseq	7.7	Tumor	Yes	Both	0	18	0	0		SMARCB1		
MAP044	CHU Strasbourg	Male	Other	NRSTS	NRSTS	DSRCT	20.5	Tumor	Yes	Both	0,42	29	6	5					
MAP048	Centre Léon Bérard	Male	Ewing Sarcoma	Ewing	EWS		WES + RNAseq	24.0	Tumor	Yes	Both	0	2	0	1				
MAP049	Institut Curie Paris	Female	Neuroblastoma	Neuroblastoma	NBL		WES + RNAseq	1172.5	Both	Yes	Both	0,88	0	84	0	BCOR, PTPRK			
MAP052	Institut Curie Paris	Female	Osteosarcoma	Osteosarcoma	OS		WES + RNAseq	24.0	Tumor	Yes	Both	0	34	10	5	BRAF			
MAP054	Centre Oscar Lamb	Female	Ewing Sarcoma	Ewing	EWS		WES + RNAseq	33.0	Both	Both	0,86	2	34	5	TP53	POLD1			
MAP055	CHU Nantes	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	316.8	Both	Yes	Both	0,28	0	7	0					
MAP056	Gustave Roussy	Female	Osteosarcoma	Osteosarcoma	OS		WES + RNAseq	20.2	no	None	0								
MAP058	CHU Nantes	Male	Neuroblastoma	Neuroblastoma	NBL		WES	21.5	Tumor	no mutat	No								
MAP060	Gustave Roussy	Male	Osteosarcoma	NRSTS	NRSTS	US	28.3	Tumor	Yes	Both	0,22	5	15	1	NF1				
MAP061	Centre Oscar Lamb	Male	Sarcoma	NRSTS	NRSTS	SS	21.9	Tumor	Yes	Both	0	13	4	3					
MAP062	La Timone Marseille	Male	Other	NRSTS	NRSTS	DSRCT	77.9	Tumor	Yes	Both	0	7	4	0					
MAP064	La Timone Marseille	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	40.4	Tumor	Yes	Both	0,41	3	8	0	TP53				
MAP065	Gustave Roussy	Female	Neuroblastoma	Neuroblastoma	NBL		no	27.5	no	None	0								
MAP066	CHU Nantes	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	22.9	Both	Yes	Both	0,32	20	30	1	ATM, FGFR4	POLD1			
MAP070	Centre Oscar Lamb	Female	Osteosarcoma	Osteosarcoma	OS		no	13.0	No	None	0,29								
MAP076	La Timone Marseille	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	WES + RNAseq	35.6	Tumor	Yes	Both	0	5	26	0	TP53			
MAP077	La Timone Marseille	Male	Other	other ST	other ST		WES + RNAseq	20.6	Tumor	Yes	Both	0,23	4	15	0	FBXW7, NRAS, TP53			
MAP078	Gustave Roussy	Male	Osteosarcoma	Osteosarcoma	OS		WES + RNAseq	47.5	Both	Yes	Both	0,39	1	51	0	CREBBP			
MAP086	Gustave Roussy	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	18.6	Tumor	Yes	No	Tumor	0	10	1	0				
MAP090	Centre Oscar Lamb	Female	Osteosarcoma	Osteosarcoma	OS		WES + RNAseq	26.6	Tumor	Yes	No	0	106	2	0				
MAP092	Institut Curie Paris	Female	Nephroblastoma	Nephroblastoma	WT		WES	18.6	Tumor	Yes	Tumor	0,31	25	0	1				
MAP094	Gustave Roussy	Male	Sarcoma	NRSTS	NRSTS	US	15.4	No	Yes	No	Tumor	0	8	0	0				
MAP095	Gustave Roussy	Female	Osteosarcoma	Osteosarcoma	OS		WES + RNAseq	28.1	Tumor	no mutat	No								
MAP097	Gustave Roussy	Male	Nephroblastoma	Nephroblastoma	WT		WES + RNAseq	30.5	Tumor	Yes	Both	0	10	0	0				
MAP099	Gustave Roussy	Female	Nephroblastoma	Nephroblastoma	WT		WES + RNAseq	23.3	Both	Yes	Both	0,33	0	23	1				
MAP101	Centre Oscar Lamb	Female	Neuroblastoma	Neuroblastoma	NBL		WES + RNAseq	924.0	Both	Yes	Both	0,73	20	57	63	ALK	ATRX, DICER1		
MAP104	Institut Curie Paris	Male	Sarcoma	NRSTS	NRSTS	SS	16.9	Tumor	Yes	No	Tumor	0,17	15	0	0				
MAP106	Centre Oscar Lamb	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	WES + RNAseq	64.8	Tumor	Yes	Both	0	7	3	0	BRAF			
MAP108	Gustave Roussy	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	WES + RNAseq	37.1	Tumor	Yes	Tumor	0	19	0	0				
MAP110	Institut Curie Paris	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	WES + RNAseq	176.1	Tumor	Yes	Tumor	0,2	30	0	0				
MAP111	Institut Curie Paris	Female	Neuroblastoma	Neuroblastoma	NBL		no	70.6	No	None	0,54								
MAP115	Hopital Trouseau	Female	Ewing Sarcoma	Ewing	EWS		WES + RNAseq	51.5	Both	Both	0,49	1	18	1					
MAP116	CHU Bordeaux	Female	MPNST	NRSTS	NRSTS	MPNST	25.3	Tumor	Yes	No	Tumor	0	39	0	1				
MAP117	La Timone Marseille	Male	Carcinoma	Carcinoma	CA		WES	267.4	Both	Yes	Both	0,63	54	32	129	SMARCB1	JAK2, KDM5A		
MAP121	Centre Léon Bérard	Male	Ewing Sarcoma	Ewing	EWS		WES + RNAseq	24.2	Tumor	Yes	Both	0	14	12	1	INFATC4, TSC2			
MAP122	Centre Léon Bérard	Female	Nephroblastoma	Nephroblastoma	WT		WES + RNAseq	31.5	Both	Yes	Both	0,47	0	10	0				
MAP123	CHU Nantes	Female	Nephroblastoma	Nephroblastoma	WT		WES + RNAseq	13.8	Tumor	Yes	Tumor	0	30	1	0				
MAP124	Institut Curie Paris	Female	Neuroblastoma	Neuroblastoma	NBL		WES + RNAseq	14.3	No	Tumor	0,3	72	0	0					
MAP125	CHU Nantes	Male	Neuroblastoma	Neuroblastoma	NBL		WES + RNAseq	27.7	Both	Yes	Both	0	2	24	0				
MAP133	La Timone Marseille	Male	Sarcoma	NRSTS	NRSTS	US	WES	25.6	Tumor	Yes	Both	0	4	3	2	TP53			
MAP134	Centre Léon Bérard	Female	Other	other	other ST		WES + RNAseq	15.5	Tumor	Yes	Both	0	7	1	0				
MAP138	La Timone Marseille	Male	Osteosarcoma	Osteosarcoma	OS		WES + RNAseq	19.1	Tumor	Yes	Both	0,31	5	40	0	NBS, TCF3, TP53			
MAP143	CHU Strasbourg	Male	Neuroblastoma	Neuroblastoma	NBL		WES + RNAseq	120.2	Both	Yes	Both	0,37	24	94	0	NOTCH1			
MAP144	CHU Strasbourg	Female	Neuroblastoma	Neuroblastoma	NBL		WES + RNAseq	26.4	No	None	0,48								
MAP151	CHU Toulouse	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	WES + RNAseq	58.3	Both	Yes	Both	0,76	3	15	4	RB1, TP53	TP53		
MAP153	CHU Nantes	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	eRMS	WES + RNAseq	9.4	No	None	0								

Tableau Annexe 2 : Descriptif de cas de l'étude MappyAct

ID Patients	Center	Sex	Disease Group	Histology group level 1	Histology group level 2 TP	Genomic Analysis	(cDNA) per mL de plasma (ng/mL)													aSNV common	aSNV Tumor	aSNV Plasma
							Copy Number	Tumor contributivity	cDNA contributivity	% cDNA	Number of Tumor SNVs	Number of Plasma SNVs										
MAP160	CHU Nantes	Female	Neuroblastoma	Neuroblastoma	NBL	no	29.2	none	No	No	None	0	None	0	103	1	0					
MAP162	La Timone Marseille	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	ERMS	22.8	Tumor	Yes	No	Tumor	0	None	0.75	0	0						
MAP173	Centre Léon Bérard	Male	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	63.0	None	No	No	None	0	None	0.75	0	0						
MAP177	CHU Toulouse	Male	Osteosarcoma	Osteosarcoma	OS	WES	10.0	Tumor	Yes	No	Tumor	0	None	0	34	1	0					
MAP178	La Timone Marseille	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	41.0	Tumor	Yes	Yes	Both	0.38	Both	0.38	6	78	0					
MAP180	La Timone Marseille	Male	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	26.4	Tumor	Yes	No	Tumor	0.45	None	0.45	25	1	0					
MAP183	La Timone Marseille	Male	Nephroblastoma	Nephroblastoma	WT	WES + RNAseq	14.0	Tumor	Yes	Yes	Both	0.31	Both	0.31	0	2	0					
MAP189	La Timone Marseille	Male	Carcinoma	Carcinoma	CA	WES + RNAseq	20.2	Tumor	Yes	No	Tumor	0	None	0	10	0	0					
MAP194	Gustave Roussy	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	11.0	Both	Yes	Yes	Both	0.39	Both	0.39	11	35	3	RB1, TP53				
MAP195	Centre Léon Bérard	Female	Neuroblastoma	Neuroblastoma	NBL	WES	23.7	None	No	No	None	0.68	None	0.68	0	0	0					
MAP199	CHU Toulouse	Male	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	136.1	Both	Yes	Yes	Both	0.38	Both	0.38	2	74	0					
MAP200	CHU Bordeaux	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	arMS	11.2	Both	Yes	Yes	Both	0.25	Both	0.25	2	6	0					
MAP203	Centre Léon Bérard	Female	Sarcoma	NRSTS	other NRSTS	WES + RNAseq	19.1	Tumor	Yes	No	Tumor	0.26	8	1	0	0						
MAP204	CHU Toulouse	Female	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	20.1	Tumor	Yes	Yes	Both	0	Both	0	42	11	0	GLI1				
MAP206	CHU Bordeaux	Male	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	57.4	Plasma	No	Yes	Plasma	0.77	0	1	54	0						
MAP207	Gustave Roussy	Male	Ewing Sarcoma	Ewing	Ewing	WES + RNAseq	19.4	None	Yes	No	Tumor	0	None	0	15	1	0					
MAP208	La Timone Marseille	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	arMS	39.0	Tumor	Yes	Yes	Both	0	Both	0	15	0	0					
MAP215	Gustave Roussy	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	17.4	Tumor	Yes	Yes	Both	0	Both	0	14	6	1	TP53				
MAP217	Gustave Roussy	Female	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	42.0	Tumor	Yes	Yes	Both	0.51	Both	0.51	41	30	3					
MAP221	Centre Léon Bérard	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	15.2	Tumor	Yes	Yes	Both	0	Both	0	22	1	0	PRKACA				
MAP222	Gustave Roussy	Male	Osteosarcoma	Osteosarcoma	OS	no	0.0	None	No	No	None	0.3	0	0	0	3	0					
MAP231	Centre Léon Bérard	Male	Carcinoma	other	other ST	WES + RNAseq	15.7	Tumor	Yes	No	Tumor	0	None	0	57	5	0					
MAP232	Institut Curie Paris	Female	Ewing Sarcoma	Ewing	Ewing	WES + RNAseq	6.1	Tumor	Yes	Yes	Both	0	Both	0	16	8	0					
MAP234	Gustave Roussy	Male	Nephroblastoma	Nephroblastoma	WT	WES	12.3	Tumor	Yes	Yes	Both	0.27	Both	0.27	7	17	0	TP53				
MAP235	CHU Nantes	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	8.8	Tumor	Yes	No	Tumor	0	None	0	21	1	0					
MAP241	Institut Curie Paris	Female	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	18.4	Tumor	Yes	No	Tumor	0	None	0	20	15	4	XPO1				
MAP242	Centre Léon Bérard	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	arMS	19.2	Tumor	Yes	Yes	Both	0	Both	0	21	15	4					
MAP244	Institut Curie Paris	Male	Rhabdoid Tumor	Rhabdoid tumor	RT	WES + RNAseq	30.2	None	Yes	Yes	Both	0	Both	0	4	43	0	CHD7, TP53				
MAP248	Hopital Trousseau	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	ERMS	475.0	Tumor	Yes	Yes	Both	0	Both	0	64	4	0					
MAP249	Centre Oscar Lamb	Female	Ewing Sarcoma	Ewing	Ewing	WES + RNAseq	55.4	Both	Yes	Yes	Both	0.3	Both	0.3	4	19	4	CCND1				
MAP250	Centre Oscar Lamb	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	arMS	31.8	Tumor	Yes	Yes	Both	0	Both	0	17	2	2	CDKN2A, MYCN, TP53				
MAP252	Gustave Roussy	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	18.7	Tumor	Yes	No	Tumor	0	None	0	67	0	0					
MAP253	La Timone Marseille	Female	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	26.8	Tumor	Yes	Yes	Both	0	Both	0	12	17	1					
MAP255	Gustave Roussy	Male	Ewing Sarcoma	Ewing	EWS	WES	50.0	Tumor	Yes	No	Tumor	0.42	Both	0.42	4	1	0					
MAP258	Centre Oscar Lamb	Male	Nephroblastoma	Nephroblastoma	WT	WES + RNAseq	16.9	Tumor	Yes	Yes	Both	0.18	Both	0.18	22	4	0					
MAP261	CHU Toulouse	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	9.5	Both	Yes	Yes	Both	0.31	Both	0.31	2	63	0	TP53				
MAP262	La Timone Marseille	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	48.2	Both	Yes	Yes	Both	0.56	Both	0.56	0	12	1					
MAP264	CHU Nantes	Female	Other	NRSTS	other NRSTS	WES + RNAseq	5.5	None	no	No	None	0	None	0	1	1	1					
MAP266	Gustave Roussy	Female	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	0.0	None	Yes	No	Tumor	0.67	12	0	0	3						
MAP268	Institut Curie Paris	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	17.0	Tumor	Yes	Yes	Both	0	Both	0	12	6	3					
MAP271	CHU Nancy	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	ERMS	8.0	None	Yes	No	Tumor	0.59	28	1	0	0		TP53				
MAP273	Institut Curie Paris	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	7.6	Tumor	Yes	No	Tumor	0	Both	0	7	0	2					
MAP274	Centre Léon Bérard	Female	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	47.2	Both	Yes	Yes	Both	0.34	Both	0.34	8	3	9	PAPPA				
MAP279	La Timone Marseille	Male	MPNST	NRSTS	MPNST	WES + RNAseq	31.8	Tumor	Yes	No	Tumor	0	None	0	14	0	0					
MAP280	Gustave Roussy	Male	Ewing Sarcoma	Ewing	EWS	no	0.0	None	No	No	None	0	None	0	0	0	0					
MAP281	CHU Nantes	Female	Neuroblastoma	Neuroblastoma	NBL	WES	8.3	Tumor	Yes	No	Tumor	0.69	2	1	0	0						
MAP282	Centre Léon Bérard	Female	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	22.9	Tumor	Yes	Yes	Both	0.2	Both	0.2	7	6	3	TP53				
MAP283	CHU Nantes	Female	Sarcoma	NRSTS	other NRSTS	WES + RNAseq	15.9	Both	Yes	Yes	Both	0.23	Both	0.23	0	34	1	MIL3				
MAP285	Centre Léon Bérard	Female	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	12.1	Tumor	Yes	No	Tumor	0	None	0	101	2	0					
MAP288	Centre Léon Bérard	Male	Ewing Sarcoma	Ewing	EWS	Sarcoma BOR/C	11.6	None	No	Tumor	0	None	0	32	0	1						
MAP289	Gustave Roussy	Male	Sarcoma	NRSTS	US	WES + RNAseq	25.5	Tumor	no mutation	No	Tumor	0.29	Both	0.29	4	46	2	0				
MAP299	CHU Nancy	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	17.0	Tumor	Yes	No	Tumor	0.3	Both	0.3	46	2	0					
MAP300	Centre Oscar Lamb	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	arMS	22.6	Tumor	Yes	No	Tumor	0	Both	0	27	0	0					
MAP301	Gustave Roussy	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	arMS	29.7	Tumor	Yes	Yes	Both	0.25	Both	0.25	15	96	0	CTNNB1, FGFR4				
MAP303	CHU Nantes	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	other	10.1	Tumor	Yes	No	Tumor	0	Both	0	13	2	0					
MAP304	CHU Angers	Female	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	66.0	Both	Yes	Yes	Both	0.31	Both	0.31	1	19	0					
MAP305	Centre Oscar Lamb	Male	MPNST	NRSTS	MPNST	WES + RNAseq	21.9	Both	Yes	Yes	Both	0.23	Both	0.23	6	18	0	PTCH1, PTPN11				
MAP307	Gustave Roussy	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	ERMS	16.5	Tumor	Yes	Yes	Both	0.18	Both	0.18	19	2	1					
MAP308	CHU Strasbourg	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	21.4	Both	Yes	Yes	Both	0.24	Both	0.24	0	15	0					
MAP309	Institut Curie Paris	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	26.2	Tumor	Yes	Yes	Both	0.15	Both	0.15	6	18	1					
MAP315	Centre Léon Bérard	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	37.0	Tumor	Yes	Yes	Both	0	Both	0	49	15	3	TET2				
MAP316	CHU Bordeaux	Male	Other	other	other ST	WES + RNAseq	36.5	Tumor	Yes	Yes	Both	0.3	Both	0.3	4	22	0	DICER1, HRAS				
MAP319	CHU Nancy	Male	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	13.6	Tumor	Yes	Yes	Both	0.22	Both	0.22	5	59	0	ALK				
MAP320	Centre Léon Bérard	Male	Ewing Sarcoma	Ewing	EWS	WES	37.4	Tumor	Yes	Yes	Both	0.16	Both	0.16	10	23	3					
MAP322	CHU Toulouse	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	arMS	8.9	Tumor	Yes	Yes	Both	0.16	Both	0.16	0	3	1	TP53				
MAP324	Gustave Roussy	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	11.4	Tumor	Yes	Yes	Both	0.2	Both	0.2	2	40	1	ATM, RAD54B				
MAP332	IN Milan	Female	Carcinoma	Carcinoma	CA	WES + RNAseq	25.1	Tumor	Yes	Yes	Both	0	Both	0	1	1	0					
MAP333	CHU Angers	Male	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	400.0	Plasma	Yes	Yes	Both	0.73	Both	0.73	9	2	66	ORST1, PIK3CA, RICTOR				
MAP336	La Timone Marseille	Male	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	13.3	Tumor	Yes	Yes	Both	0.18	Both	0.18	9	80	2	ALK, IDH1				
MAP340	Institut Curie Paris	Female	Carcinoma	Carcinoma	CA	WES + RNAseq	23.8	Tumor	Yes	Yes	Both	0.26	Both	0.26	3	5	4					
MAP342	Gustave Roussy	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	21.9	Both	Yes	Yes	Both	0.29	Both	0.29	6	71	2					
MAP344	Centre Léon Bérard	Female	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	12.0	Tumor	Yes	No	Tumor	0	Both	0	41	1	2					
MAP345	CHU Nantes	Male	NRSTS	NRSTS	other NRSTS	WES + RNAseq	9.7	None	no	No	None	0.17	None	0.17	0	0	0					
MAP347	La Timone Marseille	Female	Carcinoma	Carcinoma	CA	WES + RNAseq	137.9	None	Yes	No	Tumor	0	Both	0	3	0	0					
MAP348	Centre Léon Bérard	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	ERMS	3.9	Tumor	Yes	Yes	Both	0.25	Both	0.25	74	2	1					

ID Patients	Center	Sex	Disease Group	Histology group level 1	Histology group level 2 TP	Genomic Analysis		cDNA per mL de plasma (ng/mL)	Copy. Number	Tumor contributivity	cDNA contributivity	Contriutivity % cDNA	Number of Plasma SNVs	aSNV common	aSNV Tumor	aSNV Plasma
						WES + RNAseq	WES + RNAseq									
MAP349	Centre Oscar Lambi	Female	Carcinoma	Carcinoma	CA	WES + RNAseq	29.4	Tumor	Yes	Yes	Both	0.22	3	44	6	SMARCA4
MAP354	Centre Léon Bérard	Female	Sarcoma	NRSTS	SS	WES + RNAseq	11.6	Tumor	Yes	No	Tumor	0.2	31	1	2	
MAP355	CHU Bordeaux	Female	Carcinoma	Carcinoma	CA	WES + RNAseq	5.1	None	No	No	None	0				
MAP357	CHU Strasbourg	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	WES + RNAseq	20.3	Tumor	Yes	Yes	Both	0.19	14	10	1	
MAP359	La Timone Marseille	Male	Rhabdoid Tumor	Rhabdoid Tumor	RT	WES + RNAseq	22.0	None	No	No	Plasma	0	0	4	14	
MAP361	Institut Curie Paris	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	WES + RNAseq	120.9	None	No	Yes	Plasma	0.3	0	1	9	CDKN2A
MAP362	Centre Léon Bérard	Male	Carcinoma	Carcinoma	CA	WES + RNAseq	110.8	Tumor	Yes	Yes	Both	0	10	22	1	
MAP364	Centre Léon Bérard	Male	Sarcoma	NRSTS	OS	WES + RNAseq	24.7	Tumor	Yes	Yes	Both	0.18	4	5	2	
MAP370	CHU Nantes	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	18.2	Tumor	Yes	Yes	Both	0	4	28	1	TP53
MAP371	CHU Nancy	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	WES + RNAseq	18.8	Tumor	Yes	Yes	Both	0.22	14	14	0	
MAP372	CHU Angers	Male	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	9.5	Tumor	Yes	No	Tumor	0.17	15	0	0	
MAP376	Gustave Roussy	Female	Carcinoma	Carcinoma	CA	WES	14.2	Tumor	Yes	No	Tumor	0	57	1	1	
MAP379	Institut Curie Paris	Male	Neuroblastoma	Neuroblastoma	NBL	no	26.4	None	No	Yes	Plasma	0	0	2	7	
MAP382	Institut Curie Paris	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	9.4	None	No	Yes	Plasma	0	0	2	3	
MAP384	Gustave Roussy	Female	Sarcoma	NRSTS	US	WES + RNAseq	43.0	Tumor	Yes	Yes	Both	0.29	7	34	0	TP53
MAP386	La Timone Marseille	Female	Other	other ST	other ST	WES + RNAseq	10.6	None	No	No	None	0				
MAP387	CHU Nancy	Male	Other	NRSTS	DSRCT	WES + RNAseq	46.6	None	Yes	No	Tumor	0.21	3	0	0	
MAP388	Institut Curie Paris	Female	Other	other ST	other ST	WES + RNAseq	12.8	Tumor	Yes	Yes	Both	0	25	10	3	DICER1, KRAS, TP53
MAP389	Centre Léon Bérard	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	23.2	Tumor	Yes	Yes	Both	0.27	6	27	0	CDKN2A
MAP397	CHU Nantes	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	WES + RNAseq	12.1	Tumor	Yes	No	Tumor	0	8	0	0	
MAP399	Institut Curie Paris	Male	Other	NRSTS	other NRSTS	WES + RNAseq	5.6	Tumor	Yes	No	Both	0	4	0	0	
MAP403	Institut Curie Paris	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	9.9	Tumor	Yes	No	Tumor	0	12	0	1	
MAP404	CHU Toulouse	Female	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	7.3	Both	Yes	Yes	Both	0	7	3	1	ARI01A
MAP405	La Timone Marseille	Male	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	16.6	Both	Yes	Yes	Both	0.35	0	38	0	DDX3X, MCM7
MAP406	IN Milan	Male	Sarcoma	NRSTS	other NRSTS	WES + RNAseq	11.6	Tumor	Yes	Yes	Both	0.31	8	10	0	
MAP408	CHU Bordeaux	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	WES + RNAseq	16.2	None	Yes	No	Tumor	0	1	0	0	
MAP409	Institut Curie Paris	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	83.5	Tumor	Yes	No	Tumor	0	49	1	0	
MAP417	Institut Curie Paris	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	49.2	Plasma	No	Yes	Plasma	0.37	0	0	18	PIK3R1
MAP418	CHU Nantes	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	10.5	Tumor	Yes	No	Tumor	0	56	0	0	
MAP419	Gustave Roussy	Male	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	35.0	Tumor	Yes	Yes	Both	0.21	20	33	1	
MAP420	CHU Nantes	Male	Rhabdoid Tumor	Rhabdoid Tumor	RT	WES + RNAseq	46.2	Tumor	Yes	Yes	Both	0.29	3	2	0	
MAP422	Hopital Trousseau	Male	Neuroblastoma	Neuroblastoma	NBL	WES	28.0	None	Yes	No	Tumor	0.15				
MAP423	Gustave Roussy	Female	Neuroblastoma	Neuroblastoma	NBL	WES	23.3	Tumor	Yes	Yes	Both	0.3	20	11	0	TP53
MAP425	CHU Toulouse	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	WES + RNAseq	11.4	Tumor	Yes	No	Tumor	0.28	19	3	0	
MAP426	La Timone Marseille	Female	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	26.9	Tumor	Yes	Yes	Both	0.21	5	38	2	
MAP429	Gustave Roussy	Male	Ewing Sarcoma	Ewing	EWS	WES	8.1	Tumor	Yes	No	Tumor	0	20	0	0	
MAP430	Centre Léon Bérard	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	54.0	Tumor	Yes	Yes	Both	0.39	10	21	6	OR151, TP53
MAP432	Hopital Trousseau	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	WES + RNAseq	19.4	Tumor	Yes	Yes	Both	0.4	1	29	0	ATM
MAP434	CHU Strasbourg	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	6.8	Tumor	Yes	Yes	Both	0	25	3	0	
MAP437	Institut Curie Paris	Female	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	17.8	Tumor	Yes	Yes	Both	0.22	16	59	1	ALK
MAP438	Centre Léon Bérard	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	WES + RNAseq	7.2	Tumor	Yes	No	Tumor	0.18	11	1	0	CREBBP, KDM5A, SCMH1
MAP442	CHU Bordeaux	Female	Carcinoma	Carcinoma	CA	WES + RNAseq	11.4	None	No	No	None	0				
MAP443	CHU Strasbourg	Male	Hepatoblastoma	Hepatoblastoma	HB	WES + RNAseq	13.0	Tumor	Yes	Yes	Both	0.28	2	9	0	TOP2A
MAP446	Gustave Roussy	Male	Neuroblastoma	Neuroblastoma	NBL	WES	17.2	Tumor	Yes	Yes	Both	0.3	127	9	0	ALK
MAP447	La Timone Marseille	Female	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	57.0	Both	Yes	Yes	Both	0.53	1	46	4	KANK1
MAP448	CHU Nantes	Female	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	80.0	Both	Yes	Yes	Both	0.35	39	19	41	ERC2
MAP449	Centre Léon Bérard	Male	Other	NRSTS	other NRSTS	WES + RNAseq	10.2	None	Yes	Yes	Both	0	1	2	4	TP53
MAP450	IN Milan	Female	Neuroblastoma	Neuroblastoma	NBL	WES	462.5	Both	Yes	Yes	Both	0.57	47	103	22	BARD1
MAP454	CHU Nantes	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	WES + RNAseq	9.3	Tumor	Yes	Yes	Both	0	8	7	0	NRAS
MAP455	CHU Nantes	Female	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	11.4	Tumor	Yes	Yes	Both	0	53	3	1	PIK3C2B
MAP456	La Timone Marseille	Female	Neuroblastoma	Neuroblastoma	NBL	no	37.6	Plasma	No	Yes	Plasma	0.61	0	0	60	ML2, MYCN
MAP457	Centre Léon Bérard	Male	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	13.8	Tumor	Yes	Yes	Both	0	2	7	0	ALK
MAP458	Institut Curie Paris	Female	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	149.0	Tumor	Yes	Yes	Both	0.95	2	23	9	
MAP459	Institut Curie Paris	Male	Ewing Sarcoma	Ewing	EWS	WES	11.1	Tumor	Yes	No	Tumor	0	20	0	0	
MAP460	Institut Curie Paris	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	WES + RNAseq	9.0	Tumor	Yes	Yes	Both	0	43	14	5	BCOR
MAP463	Institut Curie Paris	Female	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	23.7	Tumor	Yes	Yes	Both	0	19	1	5	IGF1R, SMARCA2
MAP464	Centre Léon Bérard	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	33.9	Both	Yes	Yes	Both	0.26	2	25	4	PTPRD
MAP466	Institut Curie Paris	Female	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	WES + RNAseq	5.3	Tumor	Yes	No	Tumor	0	20	1	2	
MAP467	CHU Strasbourg	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	26.0	Tumor	Yes	Yes	Both	0.22	1	39	5	
MAP469	Gustave Roussy	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	16.5	Tumor	Yes	Yes	Both	0.21	18	51	0	
MAP470	Gustave Roussy	Female	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	11.8	Tumor	Yes	No	Tumor	0	47	0	0	
MAP471	Gustave Roussy	Female	Hepatoblastoma	Hepatoblastoma	HB	WES + RNAseq	16.6	Tumor	Yes	No	Tumor	0.25	34	1	0	
MAP474	Centre Léon Bérard	Female	Other	Sarcoma BCOR/C	Other	WES + RNAseq	14.5	Tumor	Yes	Yes	Both	0	8	6	3	
MAP477	Centre Léon Bérard	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	WES + RNAseq	12.7	Tumor	Yes	Yes	Both	0	12	21	2	KRAS, TP53
MAP479	CHU Strasbourg	Male	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	12.4	Tumor	Yes	Yes	Both	0	54	3	1	PAPPA
MAP484	CHU Nantes	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	45.3	Both	Yes	Yes	Both	0.28	19	8	9	
MAP486	Centre Léon Bérard	Male	Neuroblastoma	Neuroblastoma	NBL	WES	79.0	Tumor	Yes	Yes	Both	0	9	14	6	
MAP488	CHU Angers	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	36.8	Both	Yes	Yes	Both	0.38	0	25	0	
MAP490	La Timone Marseille	Male	NRSTS	NRSTS	US	WES + RNAseq	17.6	Tumor	Yes	Yes	Both	0.17	1	23	3	TP53
MAP492	CHU Bordeaux	Male	Ewing Sarcoma	Ewing	EWS	WES + RNAseq	47.9	Both	Yes	Yes	Both	0.37	0	5	0	
MAP493	La Timone Marseille	Male	Rhabdomyosarcoma	Rhabdomyosarcoma	RMS	WES + RNAseq	21.7	Tumor	Yes	No	Tumor	0	28	1	2	
MAP494	Institut Curie Paris	Female	Sarcoma	NRSTS	other NRSTS	WES + RNAseq	11.5	Tumor	Yes	Yes	Both	0	9	4	3	
MAP495	Our's Lady's-Dublin	Female	Osteosarcoma	Osteosarcoma	OS	WES + RNAseq	5.5	Tumor	Yes	No	Tumor	0.25	54	0	0	
MAP498	La Timone Marseille	Male	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	17.0	None	No	No	None	0				
MAP500	Gustave Roussy	Male	Neuroblastoma	Neuroblastoma	NBL	WES + RNAseq	357.0	Both	Yes	Yes	Both	0.43	1	61	24	

Case	Age (M)	Disease	Group	Metastasis	NGS	Remarks
1	64	Medulloblastoma	ND	M0	WES	
2	65	Medulloblastoma	4	M2	WES	
3	109	Medulloblastoma	ND		WES	
4	205	Medulloblastoma	ND		WES	
5	36	Medulloblastoma	ND	M2	WES	
6	141	Medulloblastoma	4		WES	
7	59	Medulloblastoma	ND		WES	MYCN amplification
8	348	Medulloblastoma	SHH		WES	
9	85	Medulloblastoma	WNT		WES	
10	123	Medulloblastoma	NoWnt/NoSHH		WES	
11	105	Medulloblastoma	WNT	M0	WES+Target	
12	410	Medulloblastoma	ND		WES+Target	
13	40	ETMR		M1	WES+Target	19q13 amplification
14	MD	Rhabdomyosarcoma	ND		WES+Target	
15	MD	ATRT	ND		Target	SMARCB1 Mutation
16	MD	ATRT	ND		Target	
17	MD	ATRT	ND		Target	
18	MD	Medulloblastoma	3		Target	
19	MD	Medulloblastoma	ND		Target	
20	MD	Medulloblastoma	ND		Target	
21	MD	Medulloblastoma	NoWnt/NoSHH		Target	
22	MD	Medulloblastoma	NoWnt/NoSHH		Target	
23	MD	Medulloblastoma	WNT		Target	CTNNB1 mutation
24	MD	Medulloblastoma	4		Target	
25	MD	Medulloblastoma	3		Target	
26	MD	Chordoma			Target	

Tableau Annexe 3 : Descriptif de cas de l'étude LCR

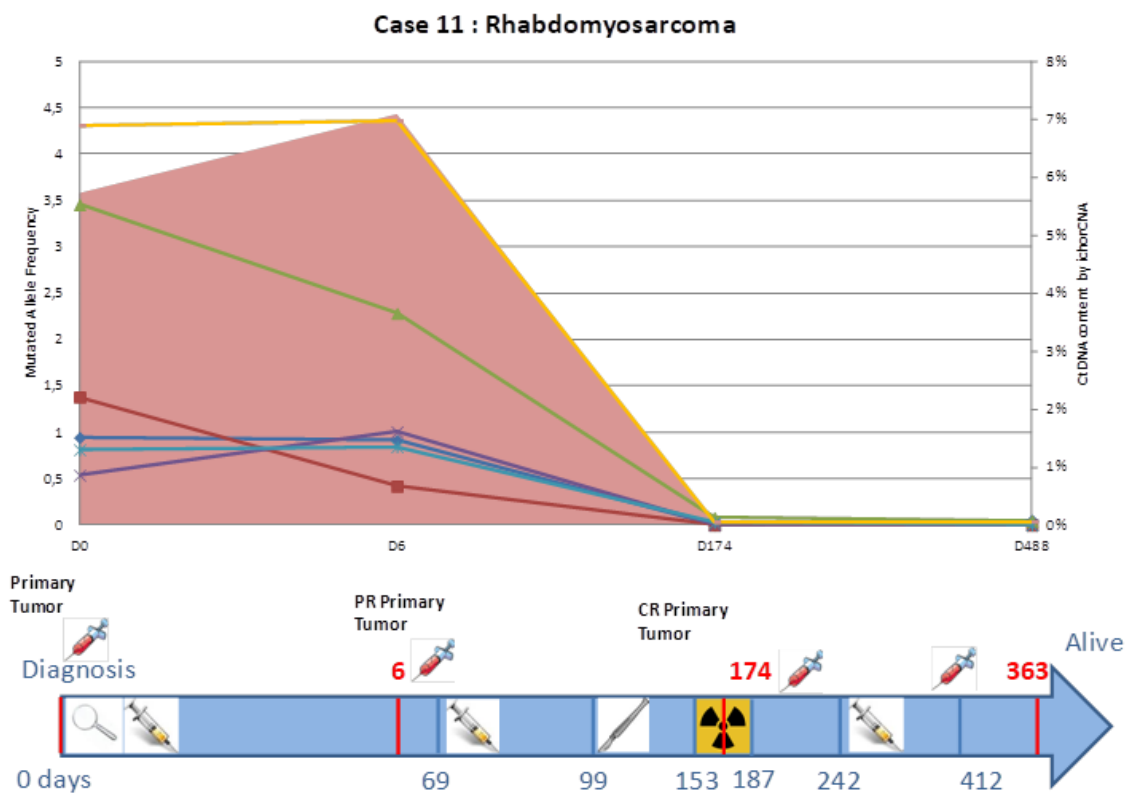
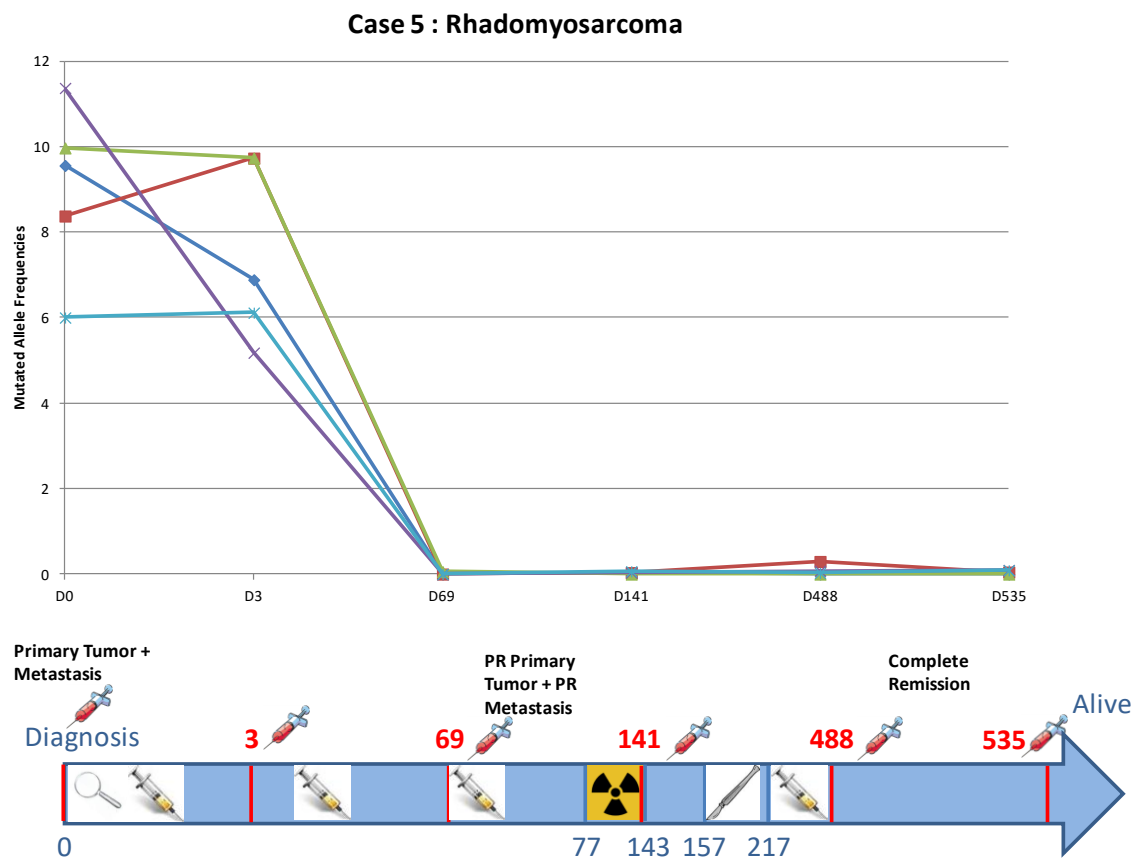
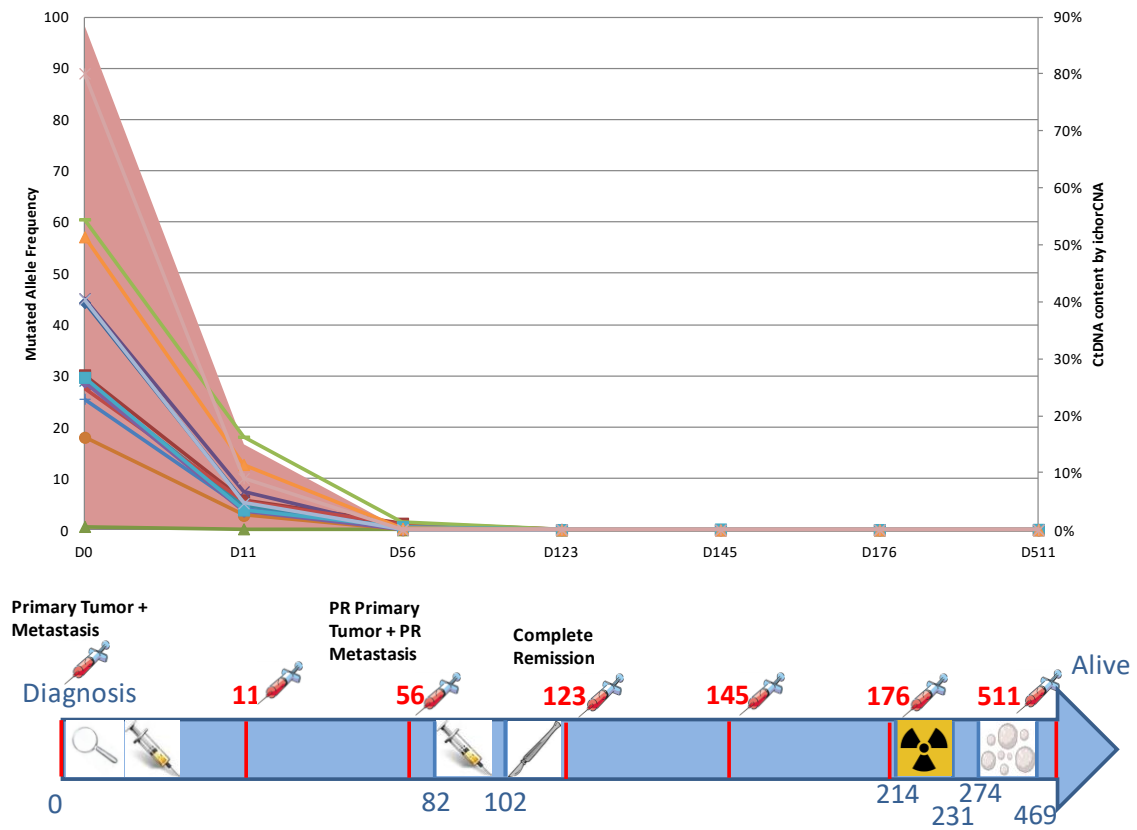
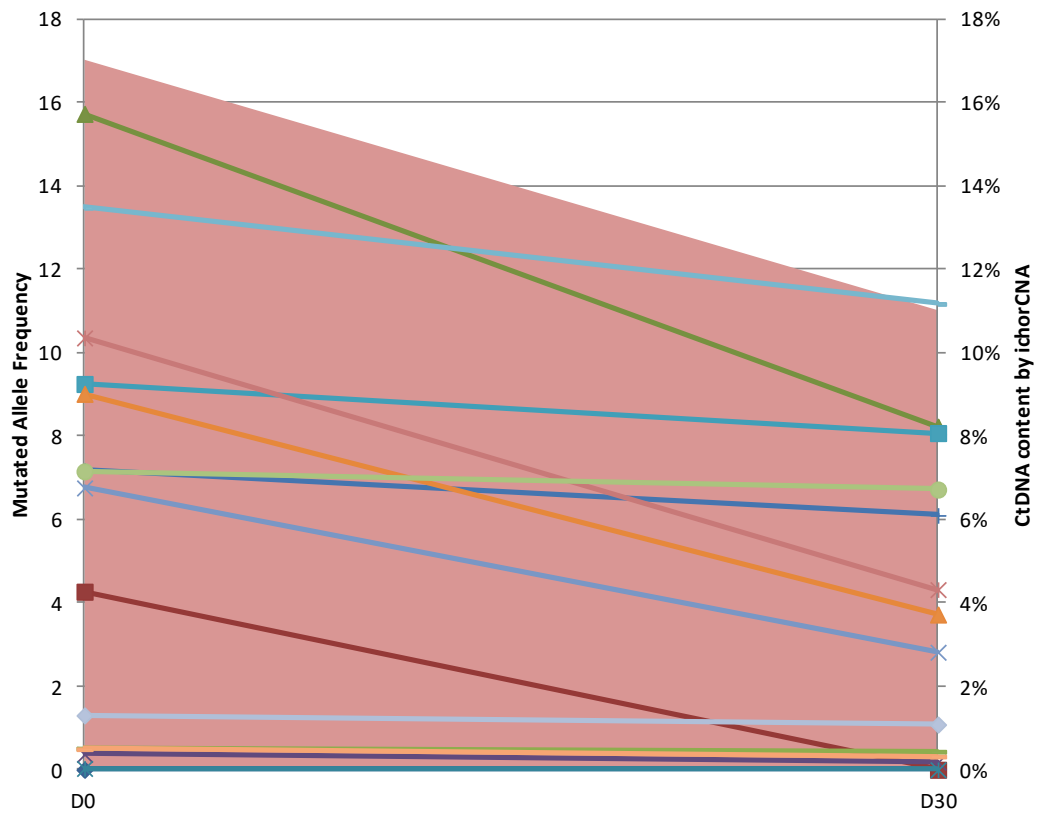


Figure annexe 1 : Suivi des mutations en séquençage ciblé des cas NGS Kids

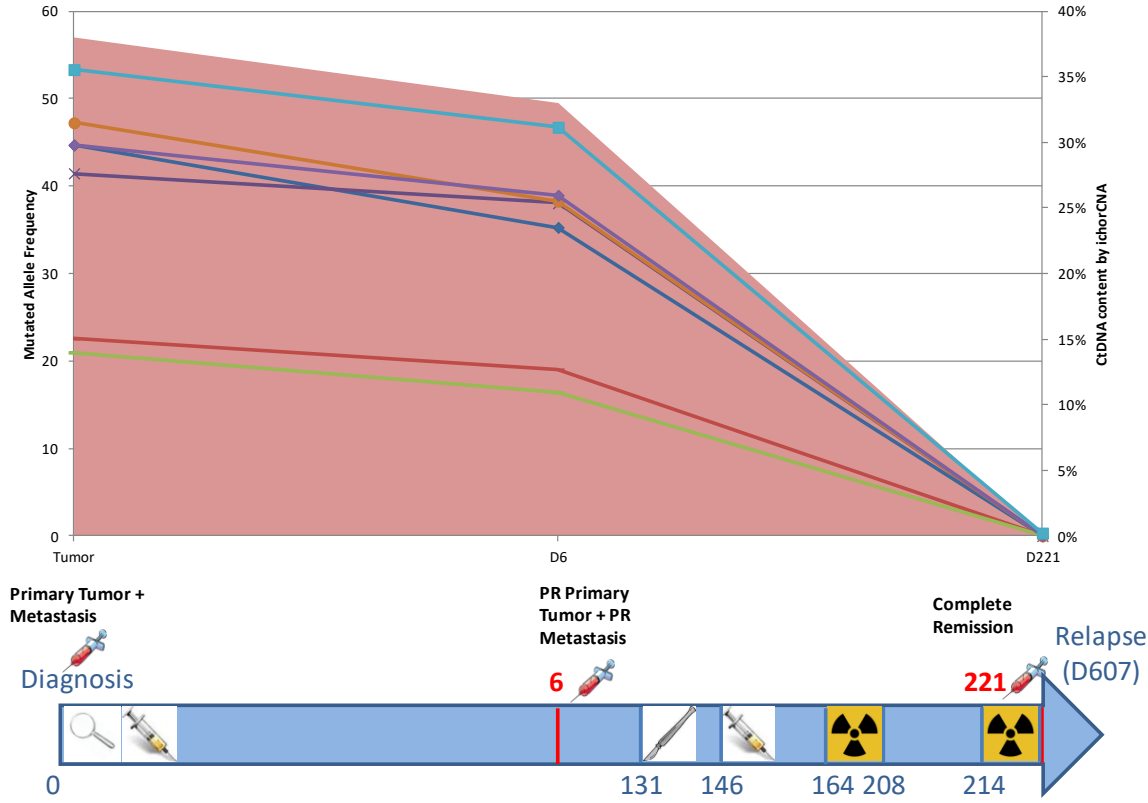
Case 14 : Neuroblastoma



Case 22 : MPNST



Case 25 : Rhabdomyosarcoma



Whole-Exome Sequencing of Cell-Free DNA Reveals Temporo-spatial Heterogeneity and Identifies Treatment-Resistant Clones in Neuroblastoma



Mathieu Chicard^{1,2}, Leo Colmet-Daage^{1,2}, Nathalie Clement², Adrien Danzon^{1,2}, Mylène Bohec³, Virginie Bernard³, Sylvain Baulande³, Angela Bellini^{1,2}, Paul Deveau^{1,2}, Gaëlle Pierron⁴, Eve Lapouble⁴, Isabelle Janoueix-Lerosey¹, Michel Peuchmaur⁵, Nadège Corradini⁶, Anne Sophie Defachelles⁷, Dominique Valteau-Couanet⁸, Jean Michon⁹, Valérie Combaret^{10,11}, Olivier Delattre^{1,3,4}, and Gudrun Schleiermacher^{1,2,9}

Abstract

Purpose: Neuroblastoma displays important clinical and genetic heterogeneity, with emergence of new mutations at tumor progression.

Experimental Design: To study clonal evolution during treatment and follow-up, an innovative method based on circulating cell-free DNA (cfDNA) analysis by whole-exome sequencing (WES) paired with target sequencing was realized in sequential liquid biopsy samples of 19 neuroblastoma patients.

Results: WES of the primary tumor and cfDNA at diagnosis showed overlap of single-nucleotide variants (SNV) and copy number alterations, with 41% and 93% of all detected alterations common to the primary neuroblastoma and cfDNA. CfDNA WES at a second time point indicated a mean of 22 new SNVs for patients with progressive disease. Relapse-specific alterations included genes of the MAPK pathway and targeted the protein kinase A signaling pathway. Deep coverage target

sequencing of intermediate time points during treatment and follow-up identified distinct subclones. For 17 seemingly relapse-specific SNVs detected by cfDNA WES at relapse but not tumor or cfDNA WES at diagnosis, deep coverage target sequencing detected these alterations in minor subclones, with relapse-emerging SNVs targeting genes of neuritogenesis and cell cycle. Furthermore a persisting, resistant clone with concomitant disappearance of other clones was identified by a mutation in the ubiquitin protein ligase *HERC2*.

Conclusions: Modelization of mutated allele fractions in cfDNA indicated distinct patterns of clonal evolution, with either a minor, treatment-resistant clone expanding to a major clone at relapse, or minor clones collaborating toward tumor progression. Identification of treatment-resistant clones will enable development of more efficient treatment strategies. *Clin Cancer Res*; 1–11. ©2017 AACR.

¹INSERM U830, Laboratoire de Génétique et Biologie des Cancers, Research Center, PSL Research University, Institut Curie, Paris, France. ²SIRIC RTOP « Recherche Translationnelle en Oncologie Pédiatrique », Translational Research Department, Research Center, PSL Research University, Institut Curie, Paris, France. ³Institut Curie Genomics of Excellence (ICGex) Platform, Research Center, Institut Curie, Paris, France. ⁴Unité de Génétique Somatique, Service de Génétique, Hospital Group, Institut Curie, Paris, France. ⁵Service de Pathologie, Hôpital Robert Debré, APHP and Université Diderot P7/SPC, Paris, France. ⁶Institut d'Hématologie et d'Oncologie Pédiatrique, Centre Léon Bérard, Lyon, France. ⁷Service d'Oncologie Pédiatrique, Centre Oscar Lambret, Lille, France. ⁸Department of Pediatric Oncology, Institut Gustave Roussy, Villejuif, France. ⁹Department of Pediatric Oncology, Hospital Group, Institut Curie, Paris, France. ¹⁰Plateforme de Génomique des Cancers, Centre Léon Bérard, Lyon, France. ¹¹Laboratoire de Recherche Translationnelle, Centre Léon-Bérard, Lyon, France.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (<http://clincancerres.aacrjournals.org/>).

Prior presentation: This article has been presented in part at the AACR 2017 conference (Abstract number 4952).

Corresponding Author: Gudrun Schleiermacher, Institut Curie, 26 rue d'Ulm, Paris 75248, France. Phone: 331-4432-4550; Fax: 331-5310-4005; E-mail: gudrun.schleiermacher@curie.fr

doi: 10.1158/1078-0432.CCR-17-1586

©2017 American Association for Cancer Research.

Introduction

Analysis of circulating tumor DNA (ctDNA), a fraction of cell-free DNA (cfDNA), is a revolutionary tool for the study of tumor-specific genetic alterations in patients with cancer. Extracted from blood, it can be used as a surrogate marker for molecular diagnosis, and for estimation of tumor burden (1–3). These techniques enable an access of tumor molecular information when a biopsy is impossible and render sequential molecular analyses possible. Shift of mutational patterns over time has been demonstrated in particular in leukemia, but in most solid tumors, sequential analysis of tumor tissue is not possible due to the limited number of accessible tumor tissue time points (4–8).

Neuroblastoma, the most frequent extracranial solid tumor of early childhood, is characterized by both clinical and genetic heterogeneity with few recurrent biomarkers. Amplification of the *MYCN* oncogene is observed in 20% of cases and is associated with poor prognosis. Other copy number alterations (CNA) occur over more extensive chromosome regions, with the segmental chromosome alterations being associated with a poor outcome (9). Few recurrent mutations have been described, the most frequent targeting *ALK* (10%–12%), genes involved in chromatin remodeling (*ATRX*, *ARID1A*), or *TERT* rearrangements.

Translational Relevance

In cancer with genetic heterogeneity, the identification of clones resistant to upfront treatment will lead to new therapeutic strategies targeting these resistant clones. We present a new approach based on whole-exome sequencing (WES) techniques of cell-free DNA (cfDNA) samples in neuroblastoma patients, enabling for the first time sequential analyses of mutational profiles in these patients. We now show how the study of cfDNA by WES can contribute to the understanding of clonal evolution in this malignancy, indicating the emergence of subclonal events present at diagnosis to clonal events at disease progression. The modelization of clonal evolution highlights mechanisms of treatment resistance and of escape. The identification of treatment-resistant clones will enable adaptation of treatment strategies.

Both spatial and temporal genetic heterogeneity play an important role in neuroblastoma. Spatial heterogeneity has been described within a tumor, and between a tumor and its metastasis, for genetic features such as *MYCN*, or other CNAs (10). Temporal heterogeneity has also been observed with an accumulation of new CNAs, and mutations including *ALK* and *RAS-MAPK* mutations at the time of progression (11–13). However, more extensive studies of these phenomena are hampered by the scarcity of paired diagnosis–relapse samples, highlighting the importance of surrogate markers.

In neuroblastoma, several studies have indicated the presence of ctDNA in blood, enabling the detection of *MYCN* or *ALK* alterations (14–16). The overall copy number profile can also be determined in cfDNA using different technologies such as

commercially available platforms or low coverage whole-genome sequencing (WGS; refs. 10, 17, 18).

The characterization of both CNA and a full mutational spectrum, of importance when analyzing mechanisms of tumor progression or treatment resistance, would be possible by whole-exome sequencing (WES) of cfDNA. Although a few limited studies have demonstrated its feasibility, the proposed protocols have not been applied to larger scale studies (19–21).

We now demonstrate for the first time the feasibility of WES cfDNA based on a modification of standard WES protocols enabling sequencing of extremely low amounts of cfDNA, enabling both CNA and SNV analysis in sequential neuroblastoma cfDNA samples. Deep coverage target sequencing was performed for validation and for analysis of intermediate time points (Fig. 1).

Materials and Methods

Patients and samples

Patients with neuroblastoma were included in this study if samples for WES of the primary tumor and plasma obtained at diagnosis were both available ($n = 19$ patients; Table 1; Supplementary Table S1). For 2 patients, plasma samples obtained at diagnosis were available, 9 had plasma collected at diagnosis and relapse, and 9 had 2 to 6 plasma collected during treatment and follow-up (Supplementary Fig. S1).

Patients were treated in French centers of the SFCE (Société Française de lutte contre les Cancers et les leucémies de l'Enfant et de l'adolescent) according to the relevant protocols (Supplementary Information S1). Written informed consent was obtained from parents/guardians according to national law. Samples were collected following inclusion of patients in the national PHRC IC2007-09 study, and this study was authorized by the ethics

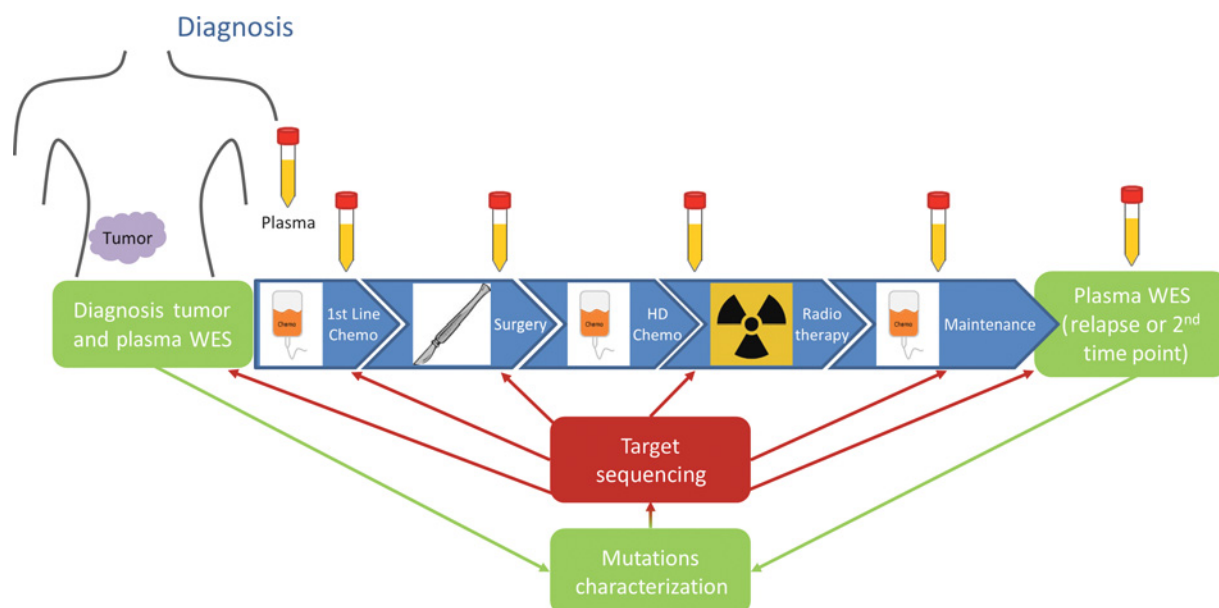


Figure 1.

Overview of study design. A WES was realized on primary tumor and cfDNA obtained at diagnosis as well as on cfDNA from a second time point (relapse or after treatment). Sequencing results allowed SNVs and CNA characterization. SNVs were studied at intermediate time points by deep coverage targeted sequencing. Deep coverage targeted sequencing was also performed on primary tumor samples to search for mutations at a subclonal level.

Table 1. Summary of patients' characteristics and genetic analyses at diagnosis

Case	Clinical information					Plasma: cfDNA Quantity (ng/mL)	Genetic alterations detected at the time of diagnosis (0- none; 1-primary tumor only; 2-cfDNA only; 3-both primary tumor and cfDNA)										Genes known to be recurrently altered in neuroblastoma or in other tumors			
	INSS Stage at diagnosis	Age at diagnosis (M)	Delay diagnosis-2nd time point	Status at 2nd time point	Delay diagnosis-last follow-up (M)		Status at last follow-up	Genomic		MYCN	Amplification	1p-	1q+	2p+	3p-	4p-		11q-	14q-	17q+
								copy number	profile											
1	4	20	7	SD	14	Dead	13,533	71	Numerical	3	0	0	0	0	0	0	0	0	0	MAP3K1: 3; MLLT3: 3
2 ^a	3	4	53	CR	47	Alive	406	104	Segmental	3	0	0	0	0	0	0	0	0	3	MLLT4: 3; RUNX3: 2; ATXNL3: 2; GATA6: 2
3	4	25	7	CR	50	Alive	2,170	193	Segmental	3	3	0	0	2	0	0	0	0	0	ADCY4: 3; KRAS: 1; STAG2: 1
4	3	19	6	CR	51	Alive	76	262	Segmental	3	3	0	0	3	0	0	0	0	3	ALK: 3; MAP2K2: 3
5	4	47	6	CR	40	Alive	1,295	1267	Segmental	0	0	0	0	3	3	0	3	0	3	ARHGAP29: 3; ARHGAP33: 3
6	4	101	1	PR	7	Dead	794	371	Segmental	3	3	3	0	0	0	0	0	3	0	ARHGAP18: 3; ITIH2: 3
7	4s	3	7	PD	16	Dead	1,248	3617	Segmental	3	3	3	0	0	0	0	0	0	3	
8	4	19	12	CR	27	Alive	1,113	232	Segmental	3	3	0	0	0	0	0	0	0	3	MLL4: 2; NBPFI0: 3
9	3	20	4	CR	36	Alive	2,328	73	Segmental	3	2	0	2	0	0	0	0	0	2	ARHGAP6: 2; ERG: 2
10	4	32	27	Relapse	31	Dead	387	368	Segmental	0	0	0	3	3	0	3	0	3	3	NKTR: 2; RASAI: 2; FLG:3
11	4	11	23	Relapse	31	Dead	639	105	Segmental	0	3	3	0	0	0	3	0	3	0	ALK: 3; AXINI: 3; MAP2K7: 3; NOTCH1: 3
12	1	58	8	Relapse	69	Alive	210	137	Segmental	0	NC									ITPR2: 3
13	4	31	22	Relapse	24	Dead	1,278	364	Segmental	3	3	3	0	0	0	3	0	3	3	ADCY4: 3; ADCY7: 3; ARHGAP30: 3; CHD5: 3; XRCC5: 3; RAD23A: 3
14	4	14	9	Relapse	11	Dead	930	/	Segmental	0	3	3	3	0	0	0	0	0	3	ALK: 3; MAPK10: 1; BRCA2: 1; CDK19: 3
15	4	30	17	Relapse	18	Dead	1,530	376	Segmental	0	3	0	0	0	0	0	0	0	3	ARHGAP18: 2; DICER1: 2; NOTCH2: 2; SHH: 2
16	4	66	26	Relapse	49	Alive	39	65	Segmental	0	3	3	2	3	3	3	0	3	3	NOTCH4: 3
17	2	30	14	Relapse	21	Dead	26	30	Segmental	0	0	0	2	0	0	0	0	0	0	RNASEL: 3
18	2	9	/	/	12	Dead	350	/	Segmental	0	3	3	0	0	0	0	0	0	0	CDK3: 3
19	4	18	14	Relapse	21	Dead	42	1,210	Segmental	3	3	0	0	0	0	3	0	3	0	MLL5: 1

NOTE: cfDNA WES at diagnosis only permitted SNV calling but not CNA analysis.

Abbreviations: SD, stable disease; M, months; NC, noncontributive.

^aCase 2: stage IVs at diagnosis with serum available at diagnosis. The study of this case is based on a local relapse with WES on tumor and plasma at relapse and target sequencing on plasma follow-up.

committees "Comité de Protection des Personnes Sud-Est IV," references L07-95/L12-171, and "Comité de Protection des Personnes Ile de France," reference 0811728.

Sample collection and processing

For each case, DNA was extracted from a single tumor sample obtained at diagnosis according to standard procedures. Plasma samples were obtained at diagnosis or during treatment and follow-up (1–6 samples/patient; Supplementary Table S2) by blood sampling directly on standard EDTA tubes and prepared by centrifugation at 2,000 rpm for 10 minutes followed by careful aliquoting and freezing at -80°C within 1 to 24 hours after collection.

cfDNA purification and quantification

cfDNA was extracted from 100 μL to 3.3 mL of plasma using QIAamp Circulating Nucleic Acid Kit (Qiagen) with the Qia-vac24s system, according to the manufacturers' recommendation. cfDNA concentration was measured by Qubit fluorometric assay (Invitrogen) with dsDNA HS (High Sensitivity) Assay Kit. The total cfDNA concentration per mL of plasma was calculated. Its quality was defined by bioanalyzer agilent 2100 (Agilent) analysis using the High Sensitivity DNA chip with cfDNA quality expressed as the 200-bp fragment fraction. For 11 plasma samples obtained at diagnosis, cfDNA copy number analysis using Oncoscan has been reported previously (10).

WGS/WES of primary tumors

Primary tumors and paired germline genomic DNA were whole-genome/exome sequenced. For 4 patients, WGS of germline, tumor at diagnosis, and tumor at relapse DNA using an Illumina HiSeq 2500, with 100-bp paired-end reads has been reported previously (11). For the other patients, WES was performed following either an Agilent SureSelect Human All Exon v5 or a Roche Nimblegen SeqCap EZ Exome V3.

Library construction and exome capture of cfDNA

cfDNA libraries were constructed without fragmentation (to account for the mean cfDNA fragment size of 160 bp) using Kapa Library Preparation Kit Illumina platforms (Kapa Biosystems) with Indexed Adapters included in SeqCap EZ Human Exome Kit v3.0 (Nimblegen Roche Sequencing). The manufacturer's protocol was modified with a ligation of 16 hours at 20°C using an adapter:insert molar ratio of 10:1; 10 μL of Pre-LM-PCR oligo 1&2 (5 $\mu\text{mol/L}$) on 9 cycles of precaptured LM-PCR and 99 μL of SeqCap EZ Purification Beads ($1.8\times$) were used for clean-up of the amplified sample library. Library quantification and quality were determined as above.

For exome capture, SeqCap EZ Exome Enrichment Kit v3.0 (Nimblegen Roche Sequencing) was used according to the manufacturer's protocol. WES was performed on germline, tumor, and plasma samples from 19 patients using Illumina Hi-seq2500 leading to paired-ends 100×100 bp. Eight samples were multiplexed for the exome capture and put on 3 lines of high-output flow cell (expected coverage: $100\times$).

For custom deep coverage targeted capture sequencing, a panel was designed encompassing 1,211 SNVs, corresponding to all SNVs observed in WES sequencing. Libraries of cfDNA were constructed using a double capture procedure. Target sequencing was performed on all plasma samples and tumor samples whenever available (PE 150×150 bp). Twelve samples were

multiplexed for the capture and sequenced with MiSeq V3 reagents (expected coverage: $1,000\times$).

Bioinformatics pipeline

Following alignment with Bowtie2 (22) allowing up to 4% of mismatches, bam files were cleaned according to the Genome Analysis Toolkit recommendations (23). Targeted bam files were cleaned up without removing duplicates.

WES variant calling was performed using 3 variant callers: GenomeAnalysisTK-3.5 UnifiedGenotyper, HaplotypeCaller, and Samtools-0.1.18 (24). Annovar-v2013-07-29 with cosmic-v64, dbsnp-v137, and RefSeq were used for annotations, and functional prediction was performed using Polyphen2, LRT, MutationTaster, and Sift (25–28). SNVs with a quality <30 , a depth of coverage <20 in tumor or plasma samples, or <2 reads supporting the variant were filtered out. Only SNVs within exons of coding genes or splice sites were kept. Then, variants reported in more than 1% of the population in the 1000 genomes (1000gApr1_2012) or Exome Sequencing Project (ESP6500) were discarded to filter out polymorphisms. Finally, synonymous variants were filtered out except those with a COSMIC ID. Variant calling comparison between germline and somatic DNAs allowed us to focus only on tumor-specific SNVs. Only nongermine SNVs supported by >2 reads and a position coverage $>20\times$ were taken into account as tumor-specific alterations.

CNAs were analyzed using VarScan.v2.3.5 (29) and DNACopy-1.42.0 (30) or HMMCopy (31). Estimation of tumor cellularity in analyzed samples (primary tumor or cfDNA) was performed using Sequenza (32).

For the definition of clones, mutated allele fractions (MAF) with similar values were taken into account, while correcting for the copy number status, all similar MAFs ($\pm 5\%$) with values corrected for copy number status defining a clone, as described previously (33). Clonal evolution graphs were generated with Fishplot-0.3 (34). Serial dilutions copy number profiles were generated using CNVkit, and the visualization of log ratio segments was done using matplotlib plotting library (35, 36).

Variants were considered confirmed if observed by two independent sequencing techniques of WES and deep coverage targeted sequencing.

The sequencing data have been deposited at the European Genome-phenome Archive (EGA) under study accession number: EGAS00001002705.

Results

Optimization of library construction for WES cfDNA analysis and bioinformatics pipeline

We have developed an experimental and bioinformatics pipeline to enable WES of cfDNA. Library construction was performed with low cfDNA input (mean, 63 ng; range, 7–100 ng). An additional step of amplification was omitted to reduce false positive rates and further improved quality of sequencing data (37, 38). The step of dual size selection only retained the cfDNA and excluded genomic DNA in case of lymphocyte degradation (Supplementary Fig. S2). These libraries could then be subjected directly to whole-exome or target capture experiments. This optimization enabled a good rate of on-target reads (mean, 82%; range, 69%–88%) and a low percentage of duplicates (mean, 32%; range, 13%–63%) for a WES at $100\times$ coverage (Supplementary Fig. S3), thus enabling a detection of MAF at

5%. The bioinformatics analytic process was established to filter out all germline events based on comparison of primary tumor and cfDNA sequencing results with the germline sequencing results, retaining only somatic alterations for further analysis.

To estimate the sensitivity of cfDNA WES and deep coverage target sequencing, patient cfDNA from one case (Case 14) harboring both significant CNAs and SNVs, and containing a high amount of ctDNA (>90%) was serially diluted with cfDNA from a healthy donor, and dilutions were submitted to both WES and deep coverage capture sequencing. No significant difference of MAFs detected by the two independent techniques for a given SNV was observed, confirming these SNVs and ruling out sequencing errors. Application of the bioinformatics pipeline enabled to establish a threshold for SNV detection of 5% MAFs by WES and 1% by deep coverage target sequencing of cfDNA, below which SNVs were not distinguishable from the background (Supplementary Fig. S4). Furthermore, apart from MYCN amplification, a threshold for detection of CNAs was determined around a proportion of 20% of patient cfDNA in the healthy donor cfDNA (Supplementary Fig. S4).

Comparison between primary tumor and cfDNA at diagnosis

For 19 neuroblastoma patients, WES or WGS data of the primary tumor obtained at diagnosis were available. These data were compared with WES cfDNA analysis at the time of diagnosis.

Analysis based on the Sequenza tool (32) revealed a mean of 73% of tumor cells in the primary tumor sample (range,

15%–98%), and a mean of 60% of ctDNA in the cfDNA fraction (range, 3%–99%; Supplementary Fig. S5). Calling of SNVs could be performed on all diagnostic cfDNA samples, whereas CNAs could not be determined in 1 case (Case 12) due to a ctDNA fraction below the threshold enabling detection of copy number changes.

Following filtering on germline DNA to filter out any constitutional alterations or germline polymorphisms, a total of 861 somatic SNVs were detected by WES analysis in cfDNA, with 353 common to the primary tumor and cfDNA, 97 specific to the primary tumor, and 411 specific to cfDNA (Supplementary Table S3). Overall, a mean of 19 SNVs (range, 9–69) per case common to primary neuroblastoma and cfDNA was observed (Fig. 2A). SNVs observed recurrently at diagnosis in both primary tumor and cfDNA by WES concerned among others the genes *ALK* (2 cases; Fig. 3D), genes of the MAPK pathway (3 cases), *ARHGAP* genes (4 cases), *NOTCH* (1 case), *MLL* genes (4 cases). SNVs of genes described to play a role in neuroblastoma oncogenesis or other tumors are highlighted in Table 1, and all detected SNVs are listed in Supplementary Table S3. No correlation of MAFs was observed between the solid tumor and cfDNA ($R^2 = 0.017$; Supplementary Fig. S6A). A mean of 6 SNVs per case were specific to primary neuroblastoma and 22 specific to the cfDNA, respectively, including among others genes of the MAPK pathway.

Furthermore, of a total of 162 CNAs, 151 CNAs were common to analyses of primary neuroblastoma and cfDNA (Table 1; Figs. 2B and 3A), with only 3 and 8 specific to either. Concerning the

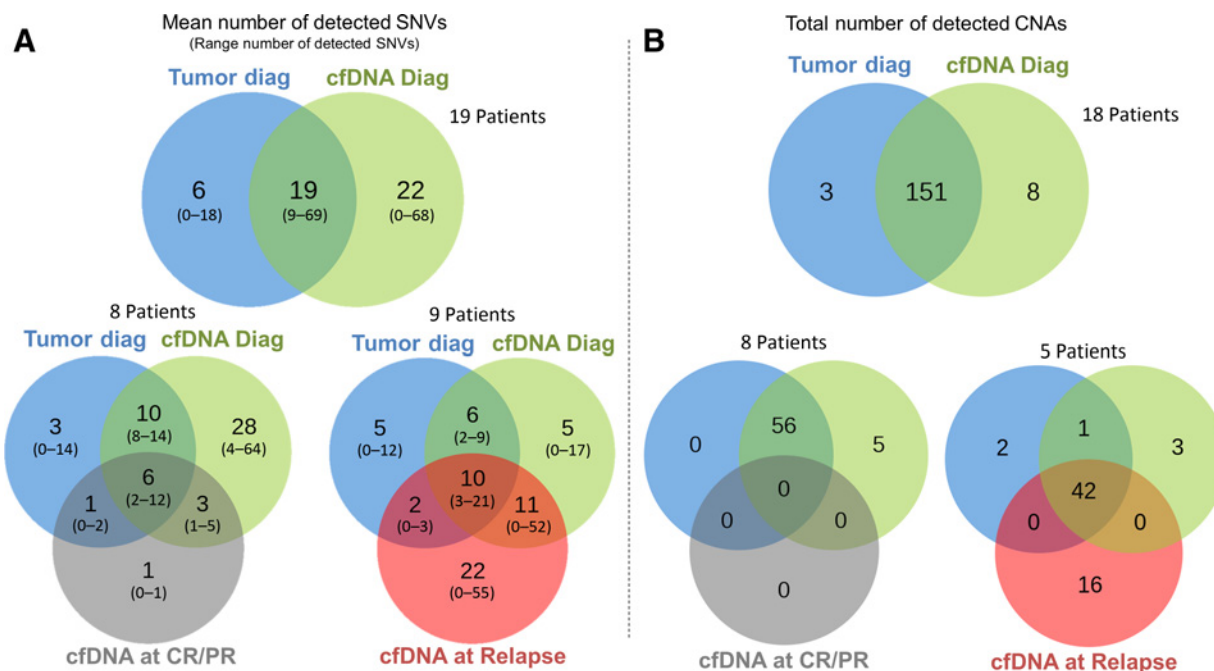
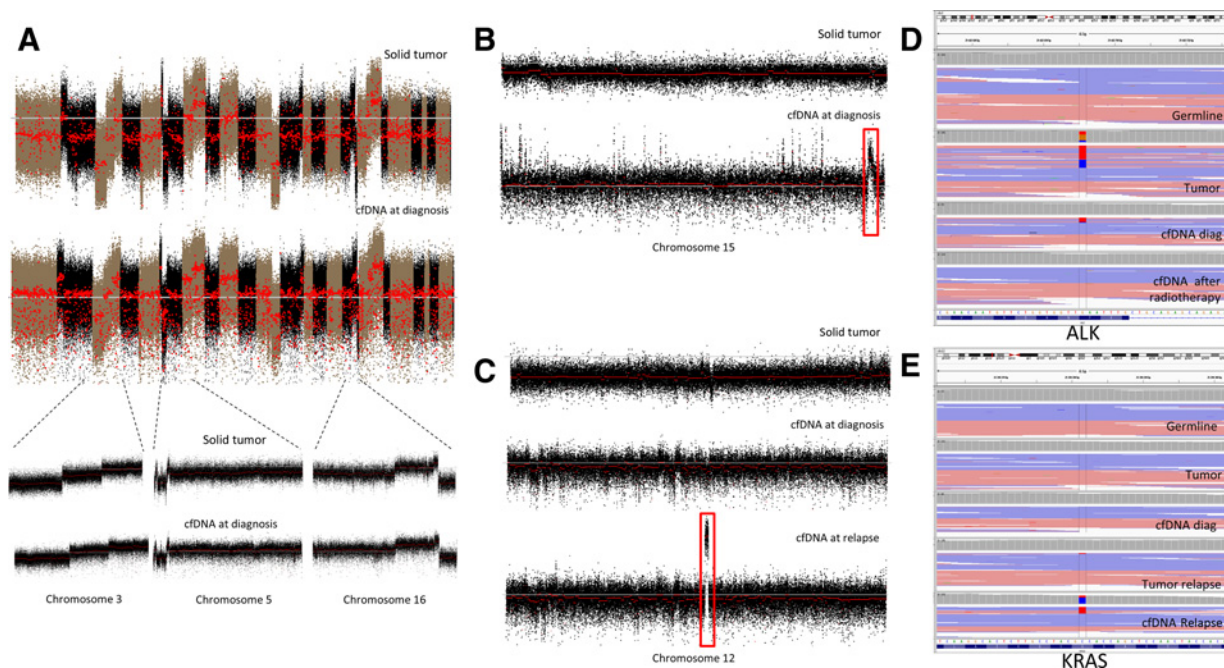


Figure 2.

Venn diagram of alterations detected in primary tumors and cfDNA (SNV and CNA). Alterations detected in primary tumor WES are in blue, in cfDNA WES at diagnosis in green, in cfDNA WES after treatment in gray, and in cfDNA WES at relapse in red. **A**, Mean number of SNVs detected by WES/WGS (range: number in square brackets). At diagnosis, the majority of tumor SNVs are detected in cfDNA. When cfDNA WES was realized at a second time point without evidence of disease, a decrease of SNVs was observed. When cfDNA WES was realized at a second time point corresponding to relapse, an increase of number of mutations was observed. **B**, Total number of CNAs detected by WES/WGS. At diagnosis, solid tumor and cfDNA share the majority of CNA. At a second time point without evidence of disease, CNAs are not detected. At the time of relapse, the majority of alterations detected at diagnosis are conserved, and few new alterations were observed (only cases with interpretable CNA profiles were counted).

**Figure 3.**

Genetic alterations observed in primary neuroblastoma and cfDNA by WES. **A**, Comparison of copy number profiles of primary neuroblastoma compared with cfDNA (Case 10). Identical CNAs are observed between the solid tumor and the cfDNA WES at diagnosis. Each chromosome (3, 5, and 16 in figure) presents the same breakpoints in tumor and cfDNA. **B**, A gain in chromosome 15 was detected only in cfDNA, containing the gene *IGF1R* (Case 1; chr15: 99200000-99600000). **C**, An amplification of *CDK4* was detected in the cfDNA at relapse only (Case 15; chr12: 58068384-58638722). Relapse tumor was not available. **D**, IGV representation of double mutation of *ALK* presents at diagnosis on tumor (30%-17%) and cfDNA (17%-2%), disappearing during treatment (Case 4; G>T-G>C, p.F1174L, chr2: 29443695). **E**, IGV representation of a *KRAS* mutation, which emerged at relapse at a subclonal level in tumor (3%) and at clonal level in cfDNA (28%; Case 11; C>T p.G13D, chr12: 25398281).

known recurrent CNA of neuroblastoma, *MYCN* amplification, 1p deletion, 11q deletion, and 17q gain were observed in both primary neuroblastoma and cfDNA in 10, 12, 6, and 12 of 19 cases, respectively, with concordant breakpoints. Alterations seen only in cfDNA but not in the primary were 1p deletion and 17q gain in 1 case each. No discrepancies were observed between CNA determined in cfDNA by WES as compared with an analysis using OncoScan for 11 cases reported previously (10).

Amplification of *MDM2* and *PTPRB* are observed for two cases (Case 1 and 17) in primary neuroblastoma and cfDNA. One case (Case 1) demonstrated a focal gain of chromosome 15q harboring the *IGF1R* gene detected only in cfDNA (Fig. 3B), as confirmed by two independent techniques (OncoScan and qPCR; ref. 10).

Analysis of a second time point cfDNA

cfDNA WES was then performed at a second time point (Fig. 1; Supplementary Table S2). For 8 patients, this corresponded to complete or partial remission (CR/PR), whereas for 9 other patients, this corresponded to disease progression or relapse (PD) as documented by clinical and/or radiological evaluations. At the second time point, a significantly higher fraction of ctDNA in the overall cfDNA was observed in patients with PD: for CR/PR patients (8 cases), the mean fraction of ctDNA in the overall cfDNA fraction was 33% (range, 0%-40%), versus a mean of 67% (range, 43%-89%) for PD patients (Supplementary Fig. S5; $P = 0.002898$). Calling of SNVs could be performed on all second

time point cfDNA samples, whereas CNAs could only be determined in 5 cases.

For patients in CR/PR (8 cases), a total of 64 SNVs were detected at the second time point. Among these, a total of 45 SNVs, with a mean of 6 SNVs (range, 2-12) per case, were common to primary neuroblastoma and the diagnostic ctDNA time point (Fig. 2A). For patients in CR/PR at the second time point, the majority of SNVs disappeared, including an *ALK* mutation (Case 4; Fig. 3D): Of a total of 323 SNVs seen in cfDNA at diagnosis, only 61 SNVs were seen at the second time point. Only in 3 cases was one new SNV observed at this second ctDNA time point (*ZNF814*, *ILDR2*, *RREB1*).

For patients in PD (9 cases), a total of 392 SNVs were detected at the second time point. Among these, 182 were common to SNVs seen in cfDNA at diagnosis, including a *ALK* mutation (Case 11). A total of 90 SNVs common to both the primary neuroblastoma and the diagnostic ctDNA time point were detected, with a mean of 10 SNVs per case (range, 3-21; Fig. 2A). A total of 198 new SNVs, with a mean of 22 new SNVs per case (range, 0-55) were observed at this second cfDNA time point. Among the new, seemingly relapse/progression-specific SNVs, several SNVs shown previously to play a role in cancer were detected targeting genes such as *KRAS* (Case 11 relapse; c.G38A; p.G13D; Fig. 3E), *ADCY7* (Case 13; c.T2834C;p.V945A), or *NCOR1* (Case 15; c.C4367T;p.A1456V). Pathways targeted by these new seemingly relapse-specific SNVs concerned genes of the MAPK pathways and targeted protein kinase A signaling [Ingenuity Pathways Analysis

(IPA)]. In addition, in 5 interpretable cases, a total of 16 new CNAs were observed (Fig. 2B), including *CDK4* amplification (Case 15, Fig. 3C) or *CDK6* amplification associated with *BCL11A* deletion (Case 16).

Deep coverage capture sequencing of cfDNA at intermediate time points

A panel was then designed encompassing all SNVs observed in any of the primary tumors or cfDNA samples studied by WES, and also taking into account SNVs observed in a relapse tumor for the 4 cases where sequencing data from the relapse tumor were available ($n = 1,211$; ref. 11), to enable deep coverage capture sequencing of all intermediate time points. In cases with confirmed SNVs, for a given SNV, MAFs observed either by cfDNA WES or by deep coverage capture sequencing did not show a statistically significant difference ($P = 0.897$, t test) and good positive correlation ($R^2 = 0.899$; Supplementary Fig. S6B), indicating the concordance of the two techniques.

Capture sequencing enabled to analyze cfDNA samples obtained at diagnosis and at the second time point for all patients, as well as intermediate time points in 9 patients (Fig. 1) aiming at a high coverage to search for subclonal events. Among the SNVs seen at relapse but not at diagnosis, targeted high coverage sequencing was employed to determine whether they were present in a subclone at diagnosis or not. Interestingly for 17 seemingly relapse-specific SNVs detected by cfDNA WES at relapse but not by tumor or cfDNA WES at diagnosis, higher coverage target analysis of primary tumor or cfDNA from diagnosis showed the presence of these alterations in a minor subclone (MAF mean, 0.84%; range, 0.33%–2.91%), indicating that these minor subclonal populations at diagnosis emerged to major clones during tumor progression and suggesting clonal evolution in 5 of 9 cases with PD (Supplementary Fig. S7). Genes targeted by the alterations in the emerging clones included 10 genes involved in neuritogenesis and 5 in cell cycle and connective tissue development (IPA network analysis; Supplementary Table S1).

Modelization of clonal evolution based on MAFs

Deep coverage capture sequencing also enabled to analyze intermediate time points in 9 patients (Cases 1–9; Supplementary Table S2), aiming at a high coverage to search for subclonal events during treatment or follow-up (Fig. 1). For these 9 cases, the MAFs of SNVs detected in cfDNA were followed in a series of 2 to 6 intermediate plasma samples. For each case, the MAFs corresponded to the clinical disease status. For one case, a decrease of MAFs concurred with disease remission (Case 9; Fig. 4A). For another case (Case 6; Fig. 4B), a decrease of the MAFs was observed during first-line chemotherapy, but an increase of the MAFs was observed corresponding to disease progression (Supplementary Fig. S8A).

Among these 9 cases, the MAFs of all SNVs observed at all time points were used to develop more detailed models of clonal evolution in neuroblastoma in patients with PD (2 patients). Clones were defined on the basis of MAFs with values corrected for copy number status, and the evolution of clones was inferred from MAFs evolving in a similar fashion. This enabled to emit hypotheses with regards to the clonal composition of the tumor at different time points.

In one case with intermediate metastatic progression, then stable disease, a persisting, resistant clone with concomitant disappearance of other clones, was identified by a mutation in

HERC2, a gene coding for a ubiquitin protein ligase (Case 1; Fig. 5A). In another case, two minor clones previously detected at the time of diagnosis emerged with an increase of MAFs from 10% to 20% and 1% to 5%, involving the genes *ADRM1*, *BMPR2*, *CELSR1* (Case 6, Figs. 4B and 5B; Supplementary Fig. S8B).

In conclusion, this novel technology enabled to follow MAFs in sequential cfDNA samples from neuroblastoma patients and indicated different mutational evolutionary patterns. In patients with response to treatment, as expected, evidence of tumor-specific genetic alterations disappeared in cfDNA. On the other hand, in case of persistent disease or relapse, tumor cell-specific alterations could be detected and intermediate analyses revealed either resistance of preexisting or emergence of new clones.

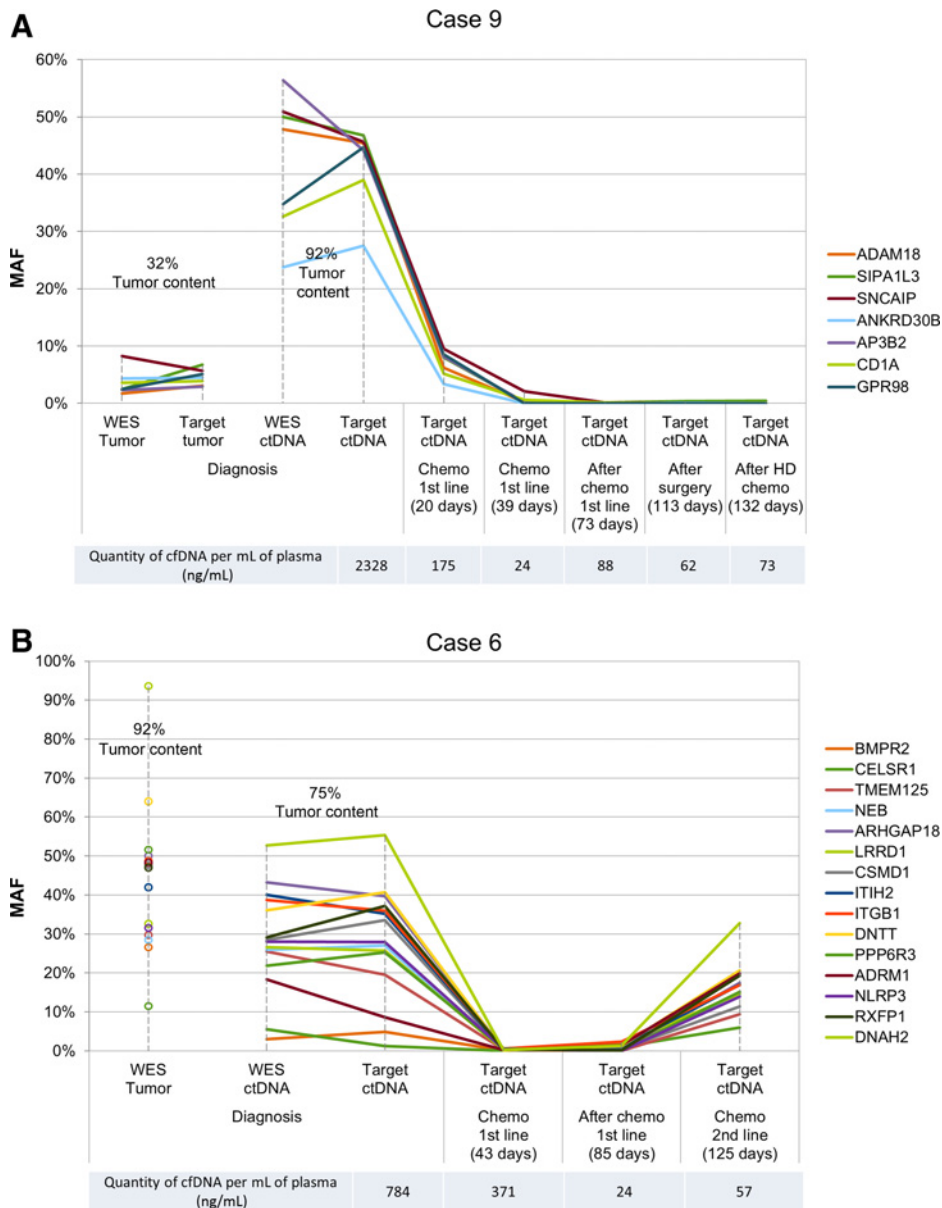
Discussion

Recently, cfDNA has emerged for the analysis of ctDNA, as this surrogate marker can detect tumor cell-specific alterations in a noninvasive setting. The majority of ctDNA studies are based on tools such as digital droplet PCR or targeted sequencing, requiring prior characterization of the biomarkers to be studied (2, 3, 15). However, clonal evolution has been shown to play an important role in many solid cancers, and new clones will not be detected when basing analysis on a comparison with the primary tumors (12, 39, 40), suggesting that WES will be more adapted in particular for cancers with few recurrent biomarkers such as neuroblastoma, despite its lower sensitivity.

Neuroblastoma has been shown to shed important amounts of ctDNA in particular in metastatic cases and high-risk disease and thus represents a relevant model for further ctDNA analysis. In this study, the ctDNA fraction of the overall cfDNA could be calculated by Sequenza (32), confirming a high proportion of ctDNA in the cfDNA fraction (mean, 60%; range, 3%–99%; Supplementary Table S1), whereas for healthy adults, in general, cfDNA quantities of approximately 10 ng/mL of plasma are observed; in high-risk neuroblastoma patients at diagnosis, high amounts of cfDNA (mean, 1,034 ng/mL of plasma; range, 26–13,533) were documented, further confirming the observation of an important shredding of ctDNA into the bloodstream by high-risk neuroblastoma at diagnosis (10, 16). Further studies on larger series will enable to establish whether the amount of ctDNA is linked to the overall disease burden, the degree of necrosis, or both. Importantly, an increase of the ctDNA fraction in the overall cfDNA was observed at the time of disease progression, versus a decrease with tumor response, suggesting that the ctDNA fraction of cfDNA in general is predictive of the underlying disease status.

WES analysis could be applied successfully to cfDNA samples, enabling calling of SNVs in 36 of 36 and calling of CNAs in 23 of 36 samples. For a given SNV, there was no correlation between the observed MAF in the primary neuroblastoma versus the plasma obtained at the same time point. This might be explained by differences of tumor content between the tumor sample and the ctDNA fraction in the total cfDNA or by spatial heterogeneity with specific SNVs detected in each sample on the other hand. A wide range of MAFs was observed in a given sample, which can be explained by subclonal events on the one hand, or the possibility of differential shredding of ctDNA from different cells.

The WES analysis of primary tumor and cfDNA identified an overlap of only 41% of SNVs at diagnosis, whereas an overlap of 93% was observed for CNAs. This suggests a greater intratumor variability for SNVs and highlights the potential of cfDNA analysis

**Figure 4.**

Evolution of MAFs in sequential cfDNA samples. Using both WES and target sequencing, evolution of MAFs correlates with disease status. MAFs detected by WES in the primary tumor and in cfDNA at diagnosis and second time points, and MAFs detected by deep coverage target sequencing in intermediate time points, are indicated. The estimated cfDNA quantity per mL of plasma is indicated below the graphs.

A, Example of decrease of MAFs during treatment (Case 9). All MAFs decrease during the first chemotherapy and all mutations disappeared after the chemotherapy. **B**, Example of resurgence of MAFs (Case 6), which will be followed by a relapse. At diagnosis, 15 SNVs were detected with MAFs identified by cfDNA WES and confirmed by deep coverage target sequencing. All SNVs become undetectable during the first chemotherapy but increase during follow-up, coinciding with metastatic progression, but with different proportions of increase. A resurgence of SNVs coinciding with disease progression is observed. For a majority of SNVs, the MAFs at relapse are lower than at diagnosis, decreasing by a factor of 2 (*ARGHAP18*: 39.75%; 0.51%; 0.00%; 17.43%; *CSMD1*: 33.56%; 0.19%; 0.50%; 11.43%; *ITGB1*: 35.16%; 0.34%; 1.04%; 16.99%; *ITIH2*: 35.93%; 0.62%; 2.31%; 16.87%). However, for other genes, an increase of MAFs can be observed, by a factor of 2 or 3: *ADRM1*: 8.66; 0.30%; 1.47%; 19.89%; *BMPR2*: 4.88%; 0.11%; 0.31%; 14.83%; *CELSR1*: 1.29%; 0.00%; 1.01%; 5.94%). Genes presenting a similar dynamic can be grouped together, enabling to infer different cell clones (Supplementary Fig. S8).

when exploration of the whole spectrum of tumor-specific alterations is of importance, for instance, in the setting of precision medicine. Multiple samples from distinct tumor sites, or tumor and metastatic sites, might give further insight into such heterogeneity. In case of primary tumor-specific alterations, these alterations might correspond to clones that might release less ctDNA, which according to one hypothesis could correspond to less aggressive cells. On the other hand, ctDNA-specific alterations might correspond to more aggressive clones or to those originating from metastatic sites (20, 41, 42). The overall ctDNA MAFs corresponded to the clinical disease status, indicating that ctDNA SNVs might serve as a surrogate marker for disease burden and/or minimal residual disease (43, 44). Furthermore, the emergence of new SNVs coincided with disease progression. As plasma samples in this study were not obtained within a specific systematic surveillance program, we cannot conclude whether cfDNA WES

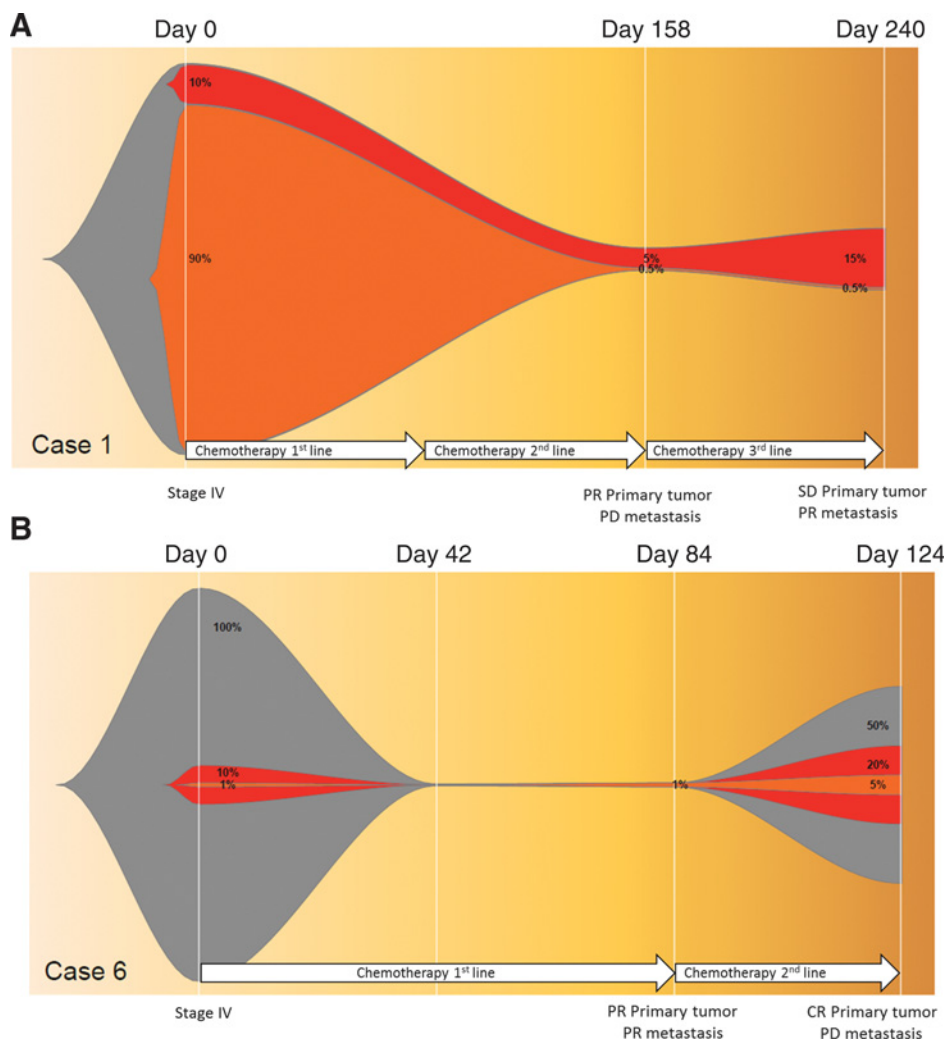
might have served as an early marker for progression. This important question will be addressed in prospective studies.

The study of SNVs and comparison of SNVs between diagnosis and relapse allows the identification of pathways involved in neuroblastoma progression, including MAPK genes, as described previously, and with one pathway specific to relapse involving protein kinase A signaling (45). Furthermore, SNVs that emerged from a subclonal level at diagnosis to a clonal level at relapse preferentially affected pathways of neuritogenesis (Supplementary Table S1; ref. 46).

To study whether seemingly relapse-specific alterations were already present at the time of diagnosis, but in a minor subclone not detected by standard coverage WES, the depth of coverage was increased. This enabled the detection of 17 SNVs, which were initially deemed relapse-specific but which are already present in a minor subclone at diagnosis, underlining

Figure 5.

Hypothesis of clonal evolution based on MAFs of SNVs detected in sequential plasma samples. The sum of MAFs represents the ctDNA proportion in the cfDNA. **A**, Modelization of clonal evolution based on MAFs. MAFs defined two clones derived from the primary clone. A subclone at diagnosis (in red), which resists to chemotherapy, takes precedence of the major clone, which is present at diagnosis (orange). The red clone then becomes a major clone at the relapse. **B**, At diagnosis, two subclones are derived from the primary clone, which remains major. After chemotherapy, these subclones appear resistant and collaborate with the primary clone for a relapse.



the importance of high depth of coverage when searching for subclones. Aiming for a depth of coverage of 1,000 $\times$, MAFs >1% were retained as significant. However, increasing the depth of coverage in the current experimental procedure was also associated with PCR duplicates. These duplicates were filtered out during the bioinformatics analysis pipeline, but it cannot be excluded that some identical reads not corresponding to PCR duplicates are also falsely removed. Thus, this protocol could be further improved using a UMI (unique molecular identifiers) prior to library construction to separate real PCR duplicates from alignment duplicates (47).

Altogether, the study of SNVs and analysis of MAFs of a given SNV over time enables the study of clonal evolution and to develop different models of tumor evolution (Figs. 4 and 5; ref. 33). In one model, among different clones derived from the primary clone, one subclone at diagnosis resists to chemotherapy and takes precedence, becoming a major clone at relapse. Indeed, a clone resistant to treatment was identified by the presence of a mutation in a ubiquitin ligase gene. In another model, subclones derived from the primary clone, which remains major, become resistant and collaborate with the primary clone for a relapse. Indeed, it can be hypothesized that these emerging clones might

contribute to tumor progression by collaboration with other nonemerging clones (48).

In conclusion, the study of clonal evolution based on cfDNA WES highlights mechanisms of treatment resistance and of escape. These techniques are of high clinical potential in the context of precision medicine and sequential analysis of molecular targets, enabling adaptation of treatment strategies depending on clonal composition. Indeed, biomarker-based clinical trials take into account a molecular profile obtained at a given time point. The evolution of predictive biomarkers over time, which might occur under selective pressure, such as targeted therapies, might now be further evaluated on the basis of noninvasive ctDNA analysis and thus be useful for identification of drug-resistant clones and will in the future lead to further sequential adaptation of biomarker-based treatment approaches.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: M. Chicard, O. Delattre, G. Schleiermacher
Development of methodology: M. Chicard, G. Schleiermacher

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): N. Clement, S. Baulande, G. Pierron, E. Lapouble, M. Peuchmaur, N. Corradini, J. Michon, V. Combaret, O. Delattre, G. Schleiermacher

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): M. Chicard, N. Clement, S. Baulande, P. Deveau, G. Pierron, E. Lapouble, M. Peuchmaur, N. Corradini, J. Michon, V. Combaret, O. Delattre, G. Schleiermacher

Writing, review, and/or revision of the manuscript: M. Chicard, L. Colmet-Daage, A. Danzon, V. Bernard, S. Baulande, G. Pierron, I. Janoueix-Lerosey, M. Peuchmaur, N. Corradini, A.S. Defachelles, D. Valteau-Couanet, V. Combaret, O. Delattre, G. Schleiermacher

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): M. Chicard, N. Clement, M. Bohec, A. Bellini, G. Pierron, I. Janoueix-Lerosey, M. Peuchmaur, A.S. Defachelles, G. Schleiermacher

Study supervision: M. Chicard, G. Schleiermacher

Acknowledgments

This study was supported by the Annenberg Foundation, the Nelia and Amadeo Barletta Foundation (FNAB), and the Association Hubert Gouin

Enfance et Cancer. This study was also funded by the Associations Enfants et Santé, Les Bagouzes à Manon, Les amis de Claire. Funding was also obtained from SiRIC/INCa (grant INCa-DGOS-4654), from the CEST of Institut Curie, and PHRC IC2007-09 grant. High-throughput sequencing has been performed by the ICGex NGS platform of the Institut Curie supported by the grants ANR-10-EQPX-03 (Equipex) and ANR-10-INBS-09-08 (France Génomique Consortium) from the Agence Nationale de la Recherche ("Investissements d'Avenir" program), by the Canceropole Ile-de-France and by the SiRIC-Curie program - SiRIC Grant "INCa-DGOS- 4654". We would like to thank Roche Diagnostic France and Magdalena Nawara for their help in protocol modification and the Circulating Biomarkers Team of Institut Curie (Aurore Rampanou, Caroline Garnier-Hego, Jordan Madic, Charles Decraene, and Charlotte Proudhon) for use of their laboratory and their advice.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received June 2, 2017; revised October 6, 2017; accepted November 21, 2017; published OnlineFirst November 30, 2017.

References

- Schwarzenbach H, Hoon DSB, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer* 2011;11:426–37.
- Newman AM, Bratman SV, To J, Wynne JF, Eclow NCW, Modlin LA, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat Med* 2014;20:548–54.
- Riva F, Bidard F-C, Houy A, Saliou A, Madic J, Rampanou A, et al. Patient-specific circulating tumor DNA detection during neoadjuvant chemotherapy in triple-negative breast cancer. *Clin Chem* 2017;63:691–9.
- Maley CC, Galipeau PC, Finley JC, Wongsurawat VJ, Li X, Sanchez CA, et al. Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nat Genet* 2006;38:468–73.
- Wang Y, Waters J, Leung ML, Unruh A, Roh W, Shi X, et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* 2014;512:155–60.
- Yachida S, Jones S, Bozic I, Antal T, Leary R, Fu B, et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* 2010;467:1114–7.
- Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 2012;481:506–10.
- Landau DA, Carter SL, Getz G, Wu CJ. Clonal evolution in hematological malignancies and therapeutic implications. *Leukemia* 2014;28:34–43.
- Janoueix-Lerosey I, Schleiermacher G, Michels E, Mosseri V, Ribeiro A, Lequin D, et al. Overall genomic pattern is a predictor of outcome in neuroblastoma. *J Clin Oncol* 2009;27:1026–33.
- Chicard M, Boyault S, Colmet Daage L, Richer W, Gentien D, Pierron G, et al. Genomic copy number profiling using circulating free tumor DNA highlights heterogeneity in neuroblastoma. *Clin Cancer Res* 2016;22:5564–73.
- Elefeld TF, Oldridge DA, Bernard V, Koster J, Daage LC, Diskin SJ, et al. Relapsed neuroblastomas show frequent RAS-MAPK pathway mutations. *Nat Genet* 2015;47:864–71.
- Schramm A, Köster J, Assenov Y, Althoff K, Peifer M, Mahlow E, et al. Mutational dynamics between primary and relapse neuroblastomas. *Nat Genet* 2015;47:872–7.
- Schleiermacher G, Javanmardi N, Bernard V, Leroy Q, Cappel R, Frio T, et al. Emergence of new ALK mutations at relapse of neuroblastoma. *J Clin Oncol* 2014;32:2727–34.
- Combaret V, Bergeron C, Noguera R, Iacono I, Puisieux A. Circulating MYCN DNA predicts MYCN-amplification in neuroblastoma. *J Clin Oncol* 2005;23:8919–20.
- Combaret V, Iacono I, Bellini A, Bréjon S, Bernard V, Marabelle A, et al. Detection of tumor ALK status in neuroblastoma patients using peripheral blood. *Cancer Med* 2015;4:540–50.
- Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med* 2014;6:224ra24.
- 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;50:2094–7.
- Leary RJ, Sausen M, Kinde I, Papadopoulos N, Carpten JD, Craig D, et al. Detection of chromosomal alterations in the circulation of cancer patients with whole-genome sequencing. *Sci Transl Med* 2012;4:162ra154.
- Murtaza M, Dawson S-J, Tsui DWY, Gale D, Forshew T, Piskorz AM, et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature* 2013;497:108–12.
- Butler TM, Johnson-Camacho K, Peto M, Wang NJ, Macey TA, Korkola JE, et al. Exome sequencing of cell-free DNA from metastatic cancer patients identifies clinically actionable mutations distinct from primary disease. *PLoS One* 2015;10:e0136407.
- Dietz S, Schirmer U, Mercé C, von Bubnoff N, Dahl E, Meister M, et al. Low input whole-exome sequencing to determine the representation of the tumor exome in circulating DNA of non-small cell lung cancer patients. *PLoS One* 2016;11:e0161012.
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–9.
- Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 2013;43:11.10.1–33.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297–303.
- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–9.
- Chun S, Fay JC. Identification of deleterious mutations within three human genomes. *Genome Res* 2009;19:1553–61.
- Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods* 2010;7:575–6.
- Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res* 2003;31:3812–4.
- Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012;22:568–76.
- Seshan VE, Olshen AB. DNACopy: DNA copy number data analysis. R package version 1.48.0 [Internet]. Bioconductor 2016[cited 2017 Jan 25]. Available from: <http://bioconductor.org/packages/DNACopy/>.

31. Ha G, Roth A, Lai D, Bashashati A, Ding J, Goya R, et al. Integrative analysis of genome-wide loss of heterozygosity and mono-allelic expression at nucleotide resolution reveals disrupted pathways in triple negative breast cancer. *Genome Res* 2012;22:1995–2007.
32. Favero F, Joshi T, Marquard AM, Birkbak NJ, Krzystanek M, Li Q, et al. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann Oncol* 2015;26:64–70.
33. Jiao W, Vembu S, Deshwar AG, Stein L, Morris Q. Inferring clonal evolution of tumors from single nucleotide somatic mutations. *BMC Bioinformatics* 2014;15:35.
34. Miller CA, McMichael J, Dang HX, Maher CA, Ding L, Ley TJ, et al. Visualizing tumor evolution with the fishplot package for R. *BMC Genomics* 2016;17:880.
35. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Comput Biol* 2016;12:e1004873.
36. Hunter JD. Matplotlib: a 2D graphics environment. *Comput Sci Eng* 2007;9:90–5.
37. Zanini F, Brodin J, Albert J, Neher RA. Error rates, PCR recombination, and sampling depth in HIV-1 whole genome deep sequencing. *Virus Res* 2017;239:106–114.
38. de Bourcy CFA, De Vlaminc I, Kanbar JN, Wang J, Gawad C, Quake SR. A quantitative comparison of single-cell whole genome amplification methods. *PLoS One* 2014;9:e105585.
39. Morrissy AS, Garzia L, Shih DJH, Zuyderduyn S, Huang X, Skowron P, et al. Divergent clonal selection dominates medulloblastoma at recurrence. *Nature* 2016;529:351–7.
40. Crompton BD, Stewart C, Taylor-Weiner A, Alexe G, Kurek KC, Calicchio ML, et al. The genomic landscape of pediatric Ewing sarcoma. *Cancer Discov* 2014;4:1326–41.
41. De Mattos-Arruda L, Mayor R, Ng CKY, Weigelt B, Martínez-Ricarte F, Torrejon D, et al. Cerebrospinal fluid-derived circulating tumour DNA better represents the genomic alterations of brain tumours than plasma. *Nat Commun* 2015;6:8839.
42. Zha J, Lackner MR. Targeting the insulin-like growth factor receptor-1R pathway for cancer therapy. *Clin Cancer Res* 2010;16:2512–7.
43. Kohlmann A, Nadarajah N, Alpermann T, Grossmann V, Schindela S, Dicker F, et al. Monitoring of residual disease by next-generation deep-sequencing of RUNX1 mutations can identify acute myeloid leukemia patients with resistant disease. *Leukemia* 2014;28:129–37.
44. van Wezel EM, Zwiijnenburg D, Zappeij-Kannegieter L, Bus E, van Noesel MM, Molenaar JJ, et al. Whole-genome sequencing identifies patient-specific DNA minimal residual disease markers in neuroblastoma. *J Mol Diagn* 2015;17:43–52.
45. Beuschlein F, Fassnacht M, Assié G, Calebiro D, Stratakis CA, Osswald A, et al. Constitutive activation of PKA catalytic subunit in adrenal cushing's syndrome. *N Engl J Med* 2014;370:1019–28.
46. Pugh TJ, Morozova O, Attiyeh EF, Asgharzadeh S, Wei JS, Auclair D, et al. The genetic landscape of high-risk neuroblastoma. *Nat Genet* 2013;45:279–84.
47. Cai X, Janku F, Zhan Q, Fan J-B. Accessing genetic information with liquid biopsies. *Trends Genet* 2015;31:564–75.
48. Cleary AS, Leonard TL, Gestl SA, Gunther EJ. Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature* 2014;508:113–7.

Clinical Cancer Research

Whole-Exome Sequencing of Cell-Free DNA Reveals Temporo-spatial Heterogeneity and Identifies Treatment-Resistant Clones in Neuroblastoma

Mathieu Chicard, Leo Colmet-Daage, Nathalie Clement, et al.

Clin Cancer Res Published OnlineFirst November 30, 2017.

Updated version	Access the most recent version of this article at: doi: 10.1158/1078-0432.CCR-17-1586
Supplementary Material	Access the most recent supplemental material at: http://clincancerres.aacrjournals.org/content/suppl/2017/11/30/1078-0432.CCR-17-1586.DC1

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org .
Permissions	To request permission to re-use all or part of this article, use this link http://clincancerres.aacrjournals.org/content/early/2018/01/23/1078-0432.CCR-17-1586 . Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.

Titre : Utilisation de l'ADN circulant pour l'étude de l'hétérogénéité tumorale et l'évolution clonale en oncologie pédiatrique

Mots clés : ADN circulant, Biomarqueurs, Hétérogénéité, Plasticité, Cancers Pédiatriques

Résumé : Les cancers pédiatriques représentent un défi thérapeutique, et une compréhension des mécanismes d'échappement aux traitements est nécessaire pour pouvoir développer de nouvelles approches thérapeutiques.

L'ADN circulant peut être libéré par une tumeur dans les fluides corporels et permet de détecter et suivre des altérations génétiques tumorales par des prélèvements successifs peu invasifs tels qu'une prise de sang.

Dans ce travail, une technique de séquençage de type « whole exome » (séquençage de l'ensemble des exons) de l'ADN circulant a été mis au point pour permettre l'étude des altérations génétiques tumorales dans le plasma chez des enfants atteints de cancer.

Ces analyses soulignent la grande hétérogénéité génétique spatiale et temporelle des cancers pédiatriques. De plus, un rôle important de l'évolution clonale dans la progression de la maladie a ainsi pu être mis en évidence.

Des approches utilisant les caractéristiques particuliers de l'ADN circulant ont également permis d'inférer le profil d'expression, basées sur l'empreinte des sites de début de transcription, ou le profil épigénétique de la tumeur. En plus d'une aide à une classification des tumeurs, ces particularités pourront aider à l'observation d'un changement d'identité cellulaire sous traitement.

L'ADN circulant est donc un formidable outil pour mieux comprendre l'échappement aux traitements d'une tumeur par son hétérogénéité spatiale et temporelle et sa plasticité.

Title : Use of circulating DNA for the study of tumor heterogeneity and clonal evolution in pediatric oncology

Keywords : cfDNA, Biomarkers, Heterogeneity, Plasticity, Pediatrics Cancers

Abstract : Pediatric cancers represent a therapeutic challenge, and an understanding of the mechanisms of escape from treatment is necessary in order to be able to develop new therapeutic approaches. Circulating DNA can be released by a tumor in bodily fluids and enable to detect and follow genetic alterations in tumors by successive minimally invasive samples such as a blood test.

In this work, a "whole exome" sequencing technique for circulating DNA was developed to allow the study of genetic tumor alterations in plasma in children with cancer. These analyzes highlight the frequent spatial and temporal genetic heterogeneity of pediatric cancers. In addition, an important role of clonal evolution in the progression of the disease was thus highlighted.

Approaches using the particular characteristics of circulating DNA have also made it possible to infer the expression profile, based on the imprint of the transcription start sites, or the epigenetic profile of the tumor. In addition to an aid in the classification of tumors, these features may help to observe a change in cellular identity under treatment.

Circulating DNA is therefore an important tool for better understanding mechanisms of escape from treatment of a tumor based on spatial and temporal heterogeneity and cellular plasticity.