



HAL
open science

Les ADN tumoraux circulant : rappel bibliographique et mise en place d'un pipeline d'analyses bioinformatique

Pierre Filhine-Trésarrieu

► To cite this version:

Pierre Filhine-Trésarrieu. Les ADN tumoraux circulant : rappel bibliographique et mise en place d'un pipeline d'analyses bioinformatique. Sciences pharmaceutiques. 2018. hal-03297540

HAL Id: hal-03297540

<https://hal.univ-lorraine.fr/hal-03297540>

Submitted on 23 Jul 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.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-thesesexercice-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

UNIVERSITE DE LORRAINE
2018

FACULTE DE PHARMACIE

MEMOIRE
du DIPLOME D'ETUDES SPECIALISEES
de Biologie Médicale

Soutenu devant le Jury Interrégional

le 25 mai 2018

par Pierre FILHINE-TRESARRIEU
né le 5 février 1989

Conformément aux dispositions de l'arrêté
du 4 octobre 1988 tient lieu de

THESE
pour le DIPLOME D'ETAT
de DOCTEUR en PHARMACIE

**Les ADN tumoraux circulant : rappel bibliographique et mise en place d'un pipeline
d'analyses bioinformatique**

Directeur de thèse
M. Alexandre HARLE

Membres du Jury

Président : M. Jean-Louis MERLIN, Professeur, Faculté de Pharmacie de Nancy

Juges :

M. Jean-Yves JOUZEAU, Professeur, Faculté de Médecine de Nancy
M. Alexandre HARLE, Maître de Conférences, Faculté de Pharmacie de Nancy
M. Kevin DALLEAU, Docteur en Pharmacie

UNIVERSITÉ DE LORRAINE
FACULTÉ DE PHARMACIE
Année universitaire 2016-2017

DOYEN

Francine PAULUS

Vice-Doyen

Béatrice FAIVRE

Directeur des Etudes

Virginie PICHON

Conseil de la Pédagogie

Président, Brigitte LEININGER-MULLER

Collège d'Enseignement Pharmaceutique Hospitalier

Président, Béatrice DEMORE

Commission Prospective Facultaire

Président, Christophe GANTZER

Vice-Président, Jean-Louis MERLIN

Commission de la Recherche

Président, Raphaël DUVAL

Responsable de la filière Officine

Béatrice FAIVRE

Responsables de la filière Industrie

Isabelle LARTAUD,

Jean-Bernard REGNOUF de VAINS

Responsable de la filière Hôpital

Béatrice DEMORE

Responsable Pharma Plus ENSIC

Jean-Bernard REGNOUF de VAINS

Responsable Pharma Plus ENSAIA

Raphaël DUVAL

Responsable Pharma Plus ENSGSI

Igor CLAROT

Responsable de la Communication

Marie-Paule SAUDER

**Responsable de la Cellule de Formation Continue
et individuelle**

Béatrice FAIVRE

**Responsable de la Commission d'agrément
des maîtres de stage**

Béatrice FAIVRE

Responsable ERASMUS

Mihayl VARBANOV

DOYENS HONORAIRES

Chantal FINANCE

Claude VIGNERON

PROFESSEURS EMERITES

Jeffrey ATKINSON

Jean-Claude BLOCK

Max HENRY

Alain MARSURA

Claude VIGNERON

PROFESSEURS HONORAIRES

Pierre DIXNEUF

Marie-Madeleine GALTEAU

Thérèse GIRARD

Michel JACQUE

Pierre LABRUDE

Vincent LOPPINET

Janine SCHWARTZBROD

Louis SCHWARTZBROD

MAITRES DE CONFERENCES HONORAIRES

Monique ALBERT

Mariette BEAUD

Gérald CATAU

Jean-Claude CHEVIN

Jocelyne COLLOMB

Bernard DANGIEN

Marie-Claude FUZELLIER

Françoise HINZELIN

Marie-Hélène LIVERTOUX

Bernard MIGNOT

Jean-Louis MONAL

ASSISTANTS HONORAIRES

Marie-Catherine BERTHE
Annie PAVIS

Blandine MOREAU
Dominique NOTTER
Christine PERDICAKIS
Marie-France POCHON
Anne ROVEL
Gabriel TROCKLE
Maria WELLMAN-ROUSSEAU
Colette ZINUTTI

ENSEIGNANTS

Section
CNU*

Discipline d'enseignement

PROFESSEURS DES UNIVERSITES - PRATICIENS HOSPITALIERS

Danièle BENSOUSSAN-LEJZEROWICZ	82	<i>Thérapie cellulaire</i>
Jean-Louis MERLIN	82	<i>Biologie cellulaire</i>
Alain NICOLAS	80	<i>Chimie analytique et Bromatologie</i>
Jean-Michel SIMON	81	<i>Economie de la santé, Législation pharmaceutique</i>
Nathalie THILLY	81	<i>Santé publique et Epidémiologie</i>

PROFESSEURS DES UNIVERSITES

Christine CAPDEVILLE-ATKINSON	86	<i>Pharmacologie</i>
Igor CLAROT	85	<i>Chimie analytique</i>
Joël DUCOURNEAU	85	<i>Biophysique, Acoustique, Audioprothèse</i>
Raphaël DUVAL	87	<i>Microbiologie clinique</i>
Béatrice FAIVRE	87	<i>Biologie cellulaire, Hématologie</i>
Luc FERRARI	86	<i>Toxicologie</i>
Pascale FRIANT-MICHEL	85	<i>Mathématiques, Physique</i>
Christophe GANTZER	87	<i>Microbiologie</i>
Frédéric JORAND	87	<i>Eau, Santé, Environnement</i>
Isabelle LARTAUD	86	<i>Pharmacologie</i>
Dominique LAURAIN-MATTAR	86	<i>Pharmacognosie</i>
Brigitte LEININGER-MULLER	87	<i>Biochimie</i>
Pierre LEROY	85	<i>Chimie physique</i>
Philippe MAINCENT	85	<i>Pharmacie galénique</i>
Patrick MENU	86	<i>Physiologie</i>
Jean-Bernard REGNOUF de VAINS	86	<i>Chimie thérapeutique</i>
Bertrand RIHN	87	<i>Biochimie, Biologie moléculaire</i>

MAITRES DE CONFÉRENCES DES UNIVERSITÉS - PRATICIENS HOSPITALIERS

Béatrice DEMORE	81	<i>Pharmacie clinique</i>
Alexandre HARLE	82	<i>Biologie cellulaire oncologique</i>
Julien PERRIN	82	<i>Hématologie biologique</i>
Marie SOCHA	81	<i>Pharmacie clinique, thérapeutique et biotechnique</i>

MAITRES DE CONFÉRENCES

Sandrine BANAS	87	<i>Parasitologie</i>
Xavier BELLANGER	87	<i>Parasitologie, Mycologie médicale</i>
Emmanuelle BENOIT	86	<i>Communication et Santé</i>
Isabelle BERTRAND	87	<i>Microbiologie</i>
Michel BOISBRUN	86	<i>Chimie thérapeutique</i>
François BONNEAUX	86	<i>Chimie thérapeutique</i>
Ariane BOUDIER	85	<i>Chimie Physique</i>
Cédric BOURA	86	<i>Physiologie</i>
Joël COULON	87	<i>Biochimie</i>

Sébastien DADE	85	Bio-informatique
Dominique DECOLIN	85	Chimie analytique
Roudayna DIAB	85	Pharmacie galénique
Natacha DREUMONT	87	Biochimie générale, Biochimie clinique
Florence DUMARCAY	86	Chimie thérapeutique
François DUPUIS	86	Pharmacologie
Adil FAIZ	85	Biophysique, Acoustique
Anthony GANDIN	87	Mycologie, Botanique
Caroline GAUCHER	86	Chimie physique, Pharmacologie
Stéphane GIBAUD	86	Pharmacie clinique
Thierry HUMBERT	86	Chimie organique
Olivier JOUBERT	86	Toxicologie, Sécurité sanitaire

ENSEIGNANTS (suite)

Section
CNU*

Discipline d'enseignement

Alexandrine LAMBERT	85	Informatique, Biostatistiques
Julie LEONHARD	86/01	Droit en Santé
Christophe MERLIN	87	Microbiologie environnementale
Maxime MOURER	86	Chimie organique
Coumba NDIAYE	86	Epidémiologie et Santé publique
Marianne PARENT	85	Pharmacie galénique
Francine PAULUS	85	Informatique
Caroline PERRIN-SARRADO	86	Pharmacologie
Virginie PICHON	85	Biophysique
Sophie PINEL	85	Informatique en Santé (e-santé)
Anne SAPIN-MINET	85	Pharmacie galénique
Marie-Paule SAUDER	87	Mycologie, Botanique
Guillaume SAUTREY	85	Chimie analytique
Rosella SPINA	86	Pharmacognosie
Sabrina TOUCHET	86	Pharmacochimie
Mihayl VARBANOV	87	Immuno-Virologie
Marie-Noëlle VAULTIER	87	Mycologie, Botanique
Emilie VELOT	86	Physiologie-Physiopathologie humaines
Mohamed ZAIYOU	87	Biochimie et Biologie moléculaire

PROFESSEUR ASSOCIE

Anne MAHEUT-BOSSER	86	Sémiologie
--------------------	----	------------

PROFESSEUR AGREGE

Christophe COCHAUD	11	Anglais
--------------------	----	---------

✎ En attente de nomination

***Disciplines du Conseil National des Universités :**

80 : Personnels enseignants et hospitaliers de pharmacie en sciences physico-chimiques et ingénierie appliquée à la santé

81 : Personnels enseignants et hospitaliers de pharmacie en sciences du médicament et des autres produits de santé

82 : Personnels enseignants et hospitaliers de pharmacie en sciences biologiques, fondamentales et cliniques

85 ; Personnels enseignants-chercheurs de pharmacie en sciences physico-chimiques et ingénierie appliquée à la santé

86 : Personnels enseignants-chercheurs de pharmacie en sciences du médicament et des autres produits de santé

87 : Personnels enseignants-chercheurs de pharmacie en sciences biologiques, fondamentales et cliniques

11 : Professeur agrégé de lettres et sciences humaines en langues et littératures anglaises et anglo-saxonnes

« LA FACULTE N'ENTEND DONNER AUCUNE APPROBATION,
NI IMPROBATION AUX OPINIONS EMISES DANS LES
THESES, CES OPINIONS DOIVENT ETRE CONSIDEREES
COMME PROPRES A LEUR AUTEUR ».

SERMENT DES APOTHICAIRES



Je jure, en présence des maîtres de la Faculté, des conseillers de l'ordre des pharmaciens et de mes condisciples :

D'honorer ceux qui m'ont instruit dans les préceptes de mon art et de leur témoigner ma reconnaissance en restant fidèle à leur enseignement.

D'exercer, dans l'intérêt de la santé publique, ma profession avec conscience et de respecter non seulement la législation en vigueur, mais aussi les règles de l'honneur, de la probité et du désintéressement.

De ne jamais oublier ma responsabilité et mes devoirs envers le malade et sa dignité humaine ; en aucun cas, je ne consentirai à utiliser mes connaissances et mon état pour corrompre les mœurs et favoriser des actes criminels.

Que les hommes m'accordent leur estime si je suis fidèle à mes promesses.

Que je sois couvert d'opprobre et méprisé de mes confrères si j'y manque.



En préambule à ce manuscrit de thèse, je tiens à remercier sincèrement l'ensemble des personnes ayant permis sa réalisation :

A Monsieur le Docteur Alexandre Harlé

Pour m'avoir fait l'honneur de diriger cette thèse. Pour l'intérêt que vous avez porté à mon travail, pour votre disponibilité, vos conseils, votre relecture attentive et votre aide qui ont permis son aboutissement.

Veillez trouver dans ce travail l'expression de mon profond respect.

Sincères remerciements.

A Monsieur le Professeur Jean-Louis Merlin

Pour m'avoir fait l'honneur de présider cette thèse.

Pour l'intérêt que vous avez porté à mon travail, votre disponibilité et vos encouragements.

Veillez trouver dans ce travail l'expression de mon profond respect.

A Monsieur le Professeur Jean-Yves Jouzeau

Pour l'intérêt porté à mon travail et vos encouragements.

Veillez trouver dans ce travail l'expression de mon profond respect.

A Monsieur Kevin Dalleau

Pour avoir accepté, avec un grand intérêt, de juger ce travail.

Trouve dans ce travail l'expression de ma profonde amitié.

À ma famille, qui êtes toujours présent malgré les tempêtes. Je vous aime.

À tous mes amis.

Aux personnes qui auraient aimé être là.

Table des matières

1	Revue de la littérature	1
1.1	Premiers travaux	1
1.2	Biologie des ADN tumoraux circulant	3
1.2.1	Origines cellulaires des ADN tumoraux circulant	5
1.2.2	Structures et macromolécules impliquées dans la biologie des ADN tumoraux circulant	7
1.2.3	Tailles et fragmentations des ADN tumoraux circulant	8
1.2.4	Dégradation, élimination des ADN libres circulant et demi-vie	10
1.2.5	Implications des ADN tumoraux circulant dans la biologie des tumeurs : messenger intercellulaire, génomemétastase et inflammation	11
1.3	Utilités cliniques des ADN tumoraux circulant	14
1.3.1	<i>Screening</i> et diagnostic	14
1.3.2	Localisation des lésions cancéreuses	22
1.3.3	Pronostic, suivi de la maladie résiduelle et du risque de récurrence	23
1.3.4	ADN tumoraux circulant et médecine personnalisée : classification des cancers, théranostique, hétérogénéité tumorale et évolution	33
1.3.5	Les ADN tumoraux circulant, bientôt en pratique courante? .	44
1.4	Méthodes d'analyses des ADN tumoraux circulant	45
1.4.1	Phase pré-analytique	45
1.4.2	Extraction des ADN tumoraux circulant	46
1.4.3	Méthodes d'analyses	48
1.4.4	Méthodes d'analyses en point unique ou pauci- <i>loci</i>	49
1.4.5	Méthodes d'analyses basées sur le séquençage nouvelle génération	54
1.4.6	Conclusion sur les méthodes d'analyses des ADN tumoraux circulant	60

2	Matériel et Méthodes	63
2.1	Contrôles de qualité, alignement, statistiques, re-calibration	64
2.1.1	Contrôles de qualité	64
2.1.2	Alignement sur le génome et traitement des séquences alignées	66
2.1.3	Appels des variants	67
2.1.4	Annotation des variants et priorisation	70
2.2	Évaluation des performances du <i>pipeline</i>	72
2.2.1	Évaluation des annotations et de la priorisation des variants .	72
2.2.2	Évaluation de la sensibilité et spécificité de l'appel des variants	72
3	Résultats	73
3.1	Annotations et priorisation des variants	73
3.2	Sensibilité et spécificité de l'appel des variants	75
3.2.1	Sans tri des séquences par <i>Trimmomatic</i> et appel par VarDict	78
4	Discussion	79

Table des figures

1	Découvertes majeures pour les ADNlc et de leurs applications cliniques, notamment en oncologie.	2
2	Mécanismes de libération des ADNlc.	3
3	Origines cellulaires des ADNlc dans le cas de lésions tumorales	6
4	Enrichissements en fragments de petite taille	9
5	Transfection de mutation <i>KRAS</i> des cellules NIH/3T3 par du plasma	12
6	Les acides désoxyribonucléiques tumoraux circulant (ADNtc) en tant que biomarqueurs des pathologies tumorales	13
7	Protocole d'analyses du kit Epi proColon Assay®	16
8	Approches pour améliorer la détection précoces des cancers par l'analyse des ADNtc	18
9	Mutations somatiques dans les fibroblastes issus de sujets sains	21
10	Croissance possible d'une tumeur solide : cinétique de Gompertz . . .	25
11	ADNtc monitoring et pronostic : exemple du cancer du sein	28
12	ADNtc monitoring et pronostic : exemple du cancer du poumon . . .	30
13	ADNtc et risque de récurrence : augmentation de sensibilité des analyses, amélioration de l'identification des patients à risque?	33
14	Cibles des principales molécules employées dans les chimiothérapies .	34
15	ADNtc et adaptation des thérapies ciblées : exemple du cancer du poumon	38
16	Résistances acquises après une thérapie ciblée chez des patients porteurs de cancer colorectal métastatique	40
17	Classification des cancers à l'aide des ADNtc : exemple des LBDG . .	43
18	Méthodes pour la détection des ADNtc	48
19	PCR digitale : procédure schématisée	50
20	BEAMing : procédure schématisée	51
21	Intpelx : procédure schématisée	52
22	COLD-PCR : principe de l'enrichissement en allèles mutées	53

23	Principe de la technologie de séquençage Illumina/Solexa	55
24	Méthode TAM-seq	57
25	Protocole de capture de régions d'intérêts	59
26	Protocole de capture de régions d'intérêts	64
27	Structure d'un fichier FASTQ	65
28	Format d'un fichier vcf	69
29	Fichier vcf annoté	71
30	Variant faux positif, intérêt des filtres	75
31	Couvertures des cibles, avec tri des séquences	77
32	Détection de la fusion des gènes <i>ROS1-SLC34A2</i>	78
33	Fractions d'allèles mutées rapportées par le pipeline, avec tri	79
34	Couvertures des cibles, sans tri des séquences	80
35	Adaptations des thérapies ciblées	84

Liste des tableaux

1	Biopsie liquide : évolution des techniques, exemple du cancer colorectal	37
2	Méthodes d'analyses pour l'ADNtc. D'après WAN et collab. (2017)	62
3	Comparaison des annotations entre Strom et collab. et celles du <i>pipeline</i>	74
4	Statistiques de couvertures après tri	76
5	Mutations décrites pour la lignée cellulaire HCC-78 capturées par Capp-Seq	78
6	Précision, rappel et scores F1 des variants appelés par <i>VarDict</i> , sans tri	79

Glossaire

- BRAF*** B-Raf proto-oncogene, serine/threonine kinase. 15, 30, 35, 39, 56, 83, 84
- CDKN2A*** cyclin dependent kinase inhibitor 2A. 17
- EGFR*** epidermal growth factor receptor. 26, 38, 39, 49, 53, 82, 83
- KIT*** KIT proto-oncogene receptor tyrosine kinase. 34
- KRAS*** KRAS proto-oncogene, GTPase. 1, 11, 12, 15, 20, 25, 35, 36, 39, 56, 83, 84
- MET*** MET proto-oncogene, receptor tyrosine kinase. 83
- MGMT*** O-6-methylguanine-DNA methyltransferase. 17
- MYC*** MYC proto-oncogene. 41
- NRAS*** NRAS proto-oncogene, GTPase. 39
- PIK3CA*** phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha. 19, 27, 40, 84
- PTEN*** phosphatase and tensin homolog. 41
- SEPT9*** septin 9. 16
- TP53*** tumor protein p53. 20, 21, 27, 56, 57, 83, 84
- 5-FU** 5-fluorouracile. 33, 34
- ACE** antigène carcino-embryonnaire. 23, 26
- ADN** acide désoxyribonucléique. 1, 3, 4, 6–10, 12, 16, 19, 20, 22, 31, 33–35, 39, 45–47, 49–52, 54–59, 70, 72
- ADNlc** acide désoxyribonucléique libre circulant. 1–12, 14, 15, 18, 20–22, 45–48, 52, 72
- ADNtc** acides désoxyribonucléiques tumoraux circulant. v, 1, 4–11, 13–15, 17–20, 22–33, 35, 38–48, 52, 58, 60, 61, 66, 68, 79, 82–85
- AMM** autorisations de mise sur le marché. 38
- AP1** protéine activatrice 1. 12
- ARMS** *amplification refractory mutation system*. 49, 51, 60
- ARN** acides ribonucléiques. 23, 58
- ARNm** acides ribonucléiques messagers. 8
- ARNmi** micro acides ribonucléiques. 3, 8
- ASCO** *American Society of Clinical Oncology*. 45, 46

AuROCC aire sous une courbe ROC. 15, 19

Bcr-Abl Breakpoint Cluster Region–Abelson. 34

CA 125 antigène tumoral 125. 23

CA 15-3 antigène tumoral 15-3. 27, 28

CA 19-9 antigène carbohydate 19-9. 15, 23

CAPP-Seq *cancer personalized profiling by deep sequencing*. 29, 31, 32, 41–43, 59, 72, 77, 80, 82

CHC carcinome hépato-cellulaire. 9, 12, 19

COLD-PCR *co-amplification at lower denaturation temperature-PCR*. 52, 53, 57

COSHG carcinomes ovarien séreux de haut grade. 20

COSMIC *Catalogue of Somatic Mutations in Cancer*. 67, 71, 72, 82

CPNPC cancer du poumon non à petites cellules. 14, 17–19, 26, 29, 44, 49, 53, 59, 60, 69, 72, 82, 83

CTC cellules tumorales circulantes. 3, 7, 27, 28

ddPCR *droplet digital Polymerase Chain Reaction*. 19, 27, 32

DPNI diagnostic pré-natal non invasif. 20

EMA *european medicines agency*. 49

FAM fractions d’allèles mutées. 47, 51, 52, 56, 57, 66, 68, 72, 73, 77, 79, 81–83

FDA Food and Drugs Administration. 16, 49

HR *Hazard ratio*. 21, 26–28, 43, 83

IC95% Intervalle de confiance à 95%. 21, 25–28

iDES *integrated digital error suppression*. 82

IFN interféron. 12

IL interleukine. 12

INCa Institut national du cancer. 61, 63

kpb kilo paires de bases. 4, 8, 9, 60

LBDGC lymphome B diffus à grandes cellules. 31, 42, 69

LDH lactate déshydrogénase. 24, 30
LLC leucémie lymphoïde chronique. 31
LMC leucémie myéloïde chronique. 34
NF- κ B nuclear factor-kappa B. 12
OR *Odds ratio*. 25
pb paires de bases. 9, 35
PCR réaction en chaîne par polymérase. 16, 49–53, 56–58, 60, 66
PCRd Polymerase Chain Reaction digitale. 28, 47–50
PCRtr réaction en chaîne par polymérase en temps réel. 15, 51
PDGFR *Platelet-derived growth factor receptor*. 34
PSA *Prostate Specific Antigen*. 23
S100B S100 calcium-binding protein B. 24
Safe-SeqS *Safe-Sequencing System*. 57, 58, 82
SMD syndromes myélodysplasiques. 31, 32, 35, 83
SNC système nerveux central. 5, 32, 47
Tam-Seq *tagged-amplicon deep sequencing*. 41, 56, 57
TAMseq *tagged-amplicon deep sequencing*. 27
TEP tomographie par émission de positons. 29
TLR Toll-like receptors. 12
TNF- α Tumor necrosis factor α . 12
VNC variation du nombre de copies. 19, 31, 32, 41, 57, 60, 64, 67–70, 81, 82
VPN valeur prédictive négative. 15
VPP valeur prédictive positive. 31, 42
VS variants structuraux. 26, 27, 32, 68, 69, 82

1 Revue de la littérature

1.1 Premiers travaux

En 1948, Mandel et Métais rapportèrent la présence d'acides nucléiques dans le plasma (MANDEL et METAIS, 1948). Jusqu'aux années 1980, cette découverte ne suscita que peu d'intérêt, hormis en rhumatologie. En 1966, Tan *et al.* montrèrent que de l'acide désoxyribonucléique libre circulant (ADNlc) était présent dans le plasma de patients avec un lupus et que celui-ci conduisait à l'apparition d'auto-anticorps anti-acide désoxyribonucléique (ADN) double brin (TAN et collab., 1966). Aujourd'hui, les ADNlc sont des biomarqueurs prometteurs en oncologie mais aussi dans d'autres domaines de la médecine (DIAZ JR et BARDELLI, 2014). Par exemple, le diagnostic pré-natal des trisomies 13, 18 et 21 par simple prise de sang de la mère est aujourd'hui possible. La greffe, l'infarctus du myocarde ou encore l'insuffisance rénale chronique sont aussi des domaines où les ADNlc peuvent présenter un intérêt en tant que biomarqueurs (DIAZ JR et BARDELLI, 2014; THIERRY et collab., 2016).

Les premières études dans le domaine de l'oncologie furent réalisées par Leon *et al.* (LEON et collab., 1977). Les auteurs mesurèrent la quantité totale d'ADNlc dans le sérum des patients avec une méthode utilisant de l'ADN marqué à l'iode 125 comme antigène et du sérum issu de patients atteints d'un lupus érythémateux disséminé comme anticorps (LEON et collab., 1977). Cinquante-cinq sujets sains furent recrutés pour établir des valeurs de référence et la quantité d'acides nucléiques totale fut aussi évaluée dans le sérum de 173 patients atteints de tout type de cancer. Les concentrations d'acides nucléiques dans le sérum de sujets sains varièrent de 0 à 100 ng/mL et 93% des sujets présentèrent moins de 50 ng/mL (LEON et collab., 1977). Pour les patients atteints d'affections oncologiques, les taux varièrent de 0 à 5000 ng/mL. Si la quantité totale d'ADNlc ne présentait que peu d'intérêt en tant que biomarqueur pour détecter un cancer (50% des patients présentaient des taux < 50 ng/mL), une concentration plus importante d'acides nucléiques était retrouvée chez les patients qui se trouvaient au stade métastatique ($p < 0,02$) (LEON et collab., 1977). Ainsi, les ADNlc peuvent être des biomarqueurs potentiels en oncologie.

Si ces premiers travaux s'intéressaient à la quantité totale d'acides nucléiques, ils ne mettaient pas en évidence les altérations somatiques, soit la présence d'ADNtc. Ce n'est qu'à la fin des années 1980 que Stroun *et al.* démontrèrent que l'ADNlc issu de plasma de patients atteints d'affections oncologiques porte des altérations somatiques (STROUN et collab., 1989, 1987). En 1994, des mutations somatiques touchant le gène KRAS proto-oncogene, GTPase (*KRAS*) furent retrouvées dans le

plasma de patients atteints de cancer pancréatique (SORENSEN et collab., 1994). Les auteurs conclurent que la détection d'altération somatique dans le plasma ou le sérum de patients représenterait un biomarqueur potentiel pour le diagnostic, le suivi ou pour des considérations théragnostiques (SORENSEN et collab., 1994). Ce travail majeur a permis de définir un nouveau concept, celui de la biopsie liquide. Cette dernière, bien moins invasive que la biopsie "classique", serait un biomarqueur pour le diagnostic, la caractérisation, la théragnostique et le suivi des affections oncologiques (DIAZ JR et BARDELLI, 2014). Néanmoins, de nombreux problèmes techniques et de validations des méthodes ainsi que de l'intérêt en tant que biomarqueurs ont dû être résolus avant d'utiliser la biopsie liquide dans la pratique quotidienne (THIERRY et collab., 2016) (**Figure 1**).

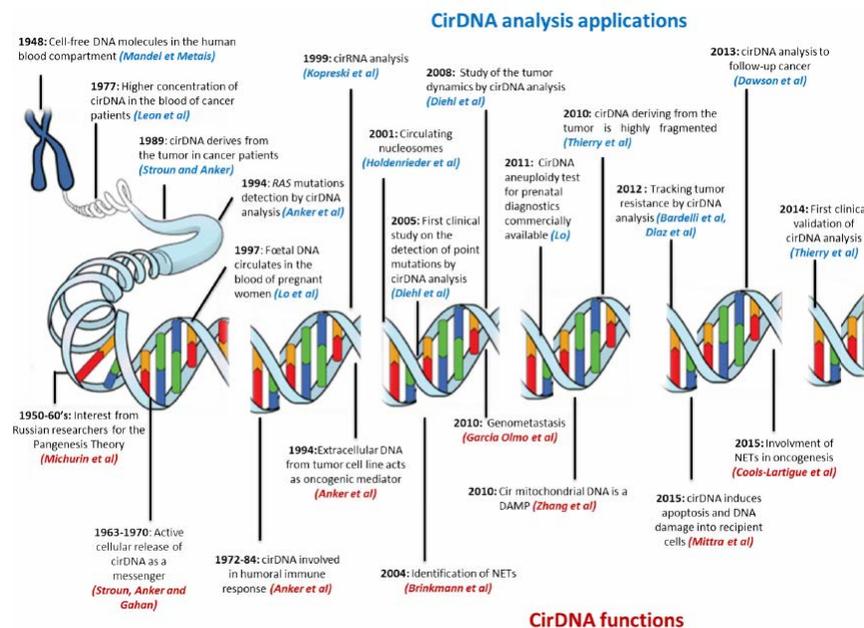


FIGURE 1 – Évolution des ADNlc et de leurs applications cliniques, notamment en oncologie. D'après THIERRY et collab. (2016).

1.2 Biologie des ADN tumoraux circulant

Le terme ADNlc désigne l'ADN nucléaire ou mitochondrial libéré d'une cellule (THIERRY et collab., 2016). Ces ADNlc peuvent être trouvés dans de nombreux fluides biologiques, comme le plasma, l'urine, les selles, le liquide céphalo-rachidien, le lait maternel (THIERRY et collab., 2016).

Les taux d'ADNlc sont généralement plus élevés chez les patients atteints de cancer que chez les sujets sains (DIAZ JR et BARDELLI, 2014). L'augmentation de la masse tumorale entraîne un accroissement du nombre de cellules apoptotiques et nécrotiques, qui relarguent leurs contenus, notamment de l'ADN (JAHR et collab., 2001; DIAZ JR et BARDELLI, 2014) (**Figure 2**). Si ces deux premiers mécanismes représentent les principales sources d'ADNlc, une sécrétion active par des exosomes a aussi été décrite comme mécanisme de libération (**Figure 2**).

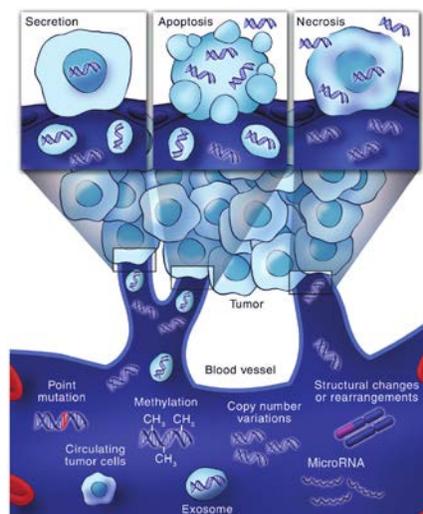


FIGURE 2 – Mécanismes de libération des ADNlc. Les ADNlc proviennent de l'apoptose, de la nécrose et de la sécrétion des cellules tumorales. L'ADN fait parti des contenus relargués par ces mécanismes. Ainsi les altérations génomiques (mutations somatiques, aberrations épigénétiques ou du nombre de copies, etc.) se retrouvent dans la circulation sanguine et peuvent être détectées en analysant l'ADNlc. D'autres altérations, notamment au niveau des ARNmi et des CTC peuvent aussi être retrouvées. D'après DIAZ JR et BARDELLI (2014).

Chez les sujets sains, les premiers travaux pour déterminer l'origine des ADNlc se sont intéressés à leur séquence et n'ont pas été probants (THIERRY et collab., 2016). En analysant les *loci* de méthylation différents entre les tissus, Sun et collab. ont montré que les ADNlc proviennent majoritairement de cellules de la moelle osseuse (SUN et collab., 2015). Le séquençage haut-débit des ADNlc de sujets sains a permis de faire un *footprinting* des nucléosomes et de leurs positions sur le génome (SNYDER et collab., 2016). Ces dernières corrèlent fortement avec les positions épigénétiques retrouvées dans les cellules qui dérivent de la moëlle osseuse (SNYDER et collab., 2016). Ces études mettent en évidence que l'apoptose des cellules provenant de la moëlle osseuse sont à l'origine de la majorité des ADNlc des sujets sains (SUN et collab., 2015; SNYDER et collab., 2016). Une augmentation de la concentration de ces premiers est retrouvée dans les syndromes inflammatoires et les maladies autoimmunes (DIAZ JR et BARDELLI, 2014). En analysant le méthylome des ADNlc, Lehmann-Werman et collab. ont mis en évidence l'origine cellulaire des ADNlc (LEHMANN-WERMAN et collab., 2016). Ainsi, ces auteurs ont montré qu'une partie des ADNlc chez les sujets atteints de diabète de type I proviennent des cellules β des îlots de Langerhans, d'oligodendrocytes pour les patients atteints de sclérose en plaque ou encore du pancréas exocrine en cas de pancréatite ou de cancers du pancréas (LEHMANN-WERMAN et collab., 2016). De même, les auteurs ont mis en évidence que l'importance de la mort de cellules spécifiques peut être évaluée par cette méthode (LEHMANN-WERMAN et collab., 2016). Il existe une grande variété d'ADNlc, en fonction de leur origine cellulaire (THIERRY et collab., 2016). Dans le domaine de l'oncologie, les premières caractérisations par Stroun et collab. montrèrent que les ADNlc avaient une taille comprise entre 0,5 et 21 kilo paires de bases (kpb). Au vu de ces tailles bien inférieures à celle de l'ADN génomique, il a été émis que les ADNlc étaient issus de fragments (STROUN et collab., 1987). Deux ans plus tard, les mêmes auteurs démontrèrent que l'ADNlc présentait une stabilité moins grande que l'ADNlc provenant de cellules non cancéreuses (STROUN et collab., 1989). L'ADNlc est présent dans la circulation sous forme moléculaire ou macromoléculaire et même internalisé dans des vésicules (THIERRY et collab., 2016). Ces structures permettent aux ADNlc d'être protégés de l'action des nucléases et de ne pas être reconnus comme signaux de danger par le système immunitaire (THIERRY et collab., 2016). La structure de l'ADNlc est en lien direct avec son mécanisme à l'origine de sa libération (apoptose, nécrose, phagocytose, etc.) (THIERRY et collab., 2016). Différentes cellules sont à l'origine des ADNlc.

1.2.1 Origines cellulaires des ADN tumoraux circulant

La quantité d'ADNtc dépend de plusieurs facteurs, dont l'évolution de la masse tumorale (THIERRY et collab., 2016). Cette cinétique est la résultante de plusieurs facteurs comme le temps de doublement de la masse tumorale, la perte cellulaire et le taux de prolifération cellulaire (THIERRY et collab., 2016). Ces paramètres varient fortement selon les cancers et le statut de la maladie (THIERRY et collab., 2016). Par exemple, pour un cancer colorectal, la perte cellulaire est très importante (96%), ce qui laisse supposer une grande quantité d'ADNtc libérée (THIERRY et collab., 2016). Dans une étude à grande échelle, Bettgowda et collab. rapportèrent une grande variabilité des taux d'ADNtc selon les cancers et les patients (BETTEGOWDA et collab., 2014). Les cancers thyroïdiens ou les gliomes, ainsi que les métastases du système nerveux central (SNC) provoquent une libération moins importante d'ADNtc que les cancers colorectaux ou mammaires (BETTEGOWDA et collab., 2014). Les travaux de Sorenson et collab. ont montré la présence d'altérations somatiques, ce qui démontre que les tumeurs sont une source d'ADNlc (SORENSEN et collab., 1994). Néanmoins, la part tumorale par rapport à la totalité varie grandement, allant de 0,003% à 95% (MOULIERE et collab., 2013). Cela démontre que les origines des ADNlc ne sont pas uniquement d'origine tumorale dans le cas d'une affection oncologique (THIERRY et collab., 2016). De même, un taux faible est aussi retrouvé chez les sujets sains (LEON et collab., 1977). Dans le cas d'un cancer, trois sources majeures pour les ADNlc sont décrites : les cellules tumorales, les cellules du micro-environnement et les cellules saines (THIERRY et collab., 2016) (**Figure 3**).

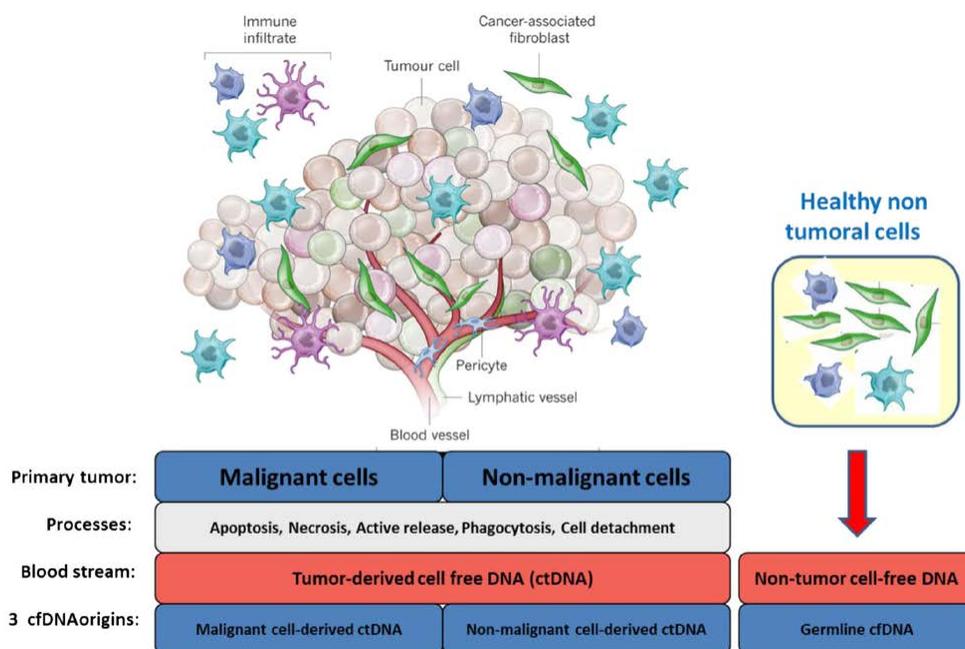


FIGURE 3 – Origines cellulaires des ADNlc dans le cas de lésions tumorales. Trois sources sont possibles : les cellules tumorales, du microenvironnement et saines. Les différents processus pour la libération des ADNlc influent sur la structure de ceux-ci. La cinétique de la masse tumorale, sa localisation ainsi que le statut du patient influent sur la quantité d’ADNlc retrouvée dans les fluides biologiques. D’après THIERRY et collab. (2016).

Même si les taux d’ADNlc sont supérieurs à ceux des sujets sains (5–1500 ng/mL *vs.* 1-5 ng/mL), une grande hétérogénéité de la part tumorale a été mise en évidence (THIERRY et collab., 2016; MOULIERE et collab., 2013). Une des hypothèses de cette hétérogénéité est que les cellules tumorales sécrètent de manière variable les ADNlc (THIERRY et collab., 2016). Par exemple, des cancers où l’autophagie (comme source d’énergie par auto-digestion) ou l’hypoxie (comme les cancers pancréatiques) présentent des taux d’ADNlc supérieurs comparés à des lésions tumorales au sein desquelles ces mécanismes sont peu impliqués (THIERRY et collab., 2016). De plus, durant l’autophagie, les histones sont citrulinisées, ce qui provoque leur détente et l’expulsion de l’ADN (THIERRY et collab., 2016). L’hypoxie influence aussi sur le relargage des ADNlc (CORTESE et collab., 2015). Ainsi, les ADNlc peuvent être des biomarqueurs de suivi notamment dans les cancers où l’hypoxie est un mécanisme

essentiel, comme les cancers pancréatiques (THIERRY et collab., 2016). La nécrose, souvent retrouvée dans les cancers colorectaux, participe grandement au relargage des ADNtc (THIERRY et collab., 2016). À l'inverse, la barrière hémato-encéphalique filtre les ADNtc (MOULIERE et ROSENFELD, 2015). Ainsi, une quantité bien moins importante d'ADNtc est retrouvée chez les patients atteints de gliomes (BETTEGOWDA et collab., 2014). Il est important de noter que les ADNtc ne proviennent pas des CTC (THIERRY et collab., 2016). En effet, une cellule humaine contient environ 6 pg d'ADN et la concentration médiane d'ADNtc chez les patients atteints de cancer est de 17 ng/mL. Si l'ADNtc provenait des CTC, il faudrait retrouver 2000 CTC/mL alors qu'il est habituel de mesurer 10 CTC pour 7,5 mL de sang total (CROWLEY et collab., 2013).

1.2.2 Structures et macromolécules impliquées dans la biologie des ADN tumoraux circulant

Plusieurs structures et macromolécules sont impliquées dans la biologie des ADNtc (THIERRY et collab., 2016). Le nucléosome est une des macromolécules essentielles dans la biologie des ADNtc. Celui-ci est composé d'un octamère d'histones, chargées positivement, autour duquel s'enroule de l'ADN double brin, chargé négativement (THIERRY et collab., 2016). Les histones H1 assurent la stabilité de ce complexe (THIERRY et collab., 2016). Cent quarante-sept bases d'ADN sont enroulées autour d'un histone et l'ADN "linker" entre deux octamères d'histone est long de 20 à 90 bases. Cette structure protège l'ADNtc de l'action d'enzymes (THIERRY et collab., 2016). L'apoptose se définit comme un mécanisme de mort cellulaire caractérisé par le rétrécissement des cellules, l'apparition de blebs, la condensation de la chromatine et la fragmentation du noyau (THIERRY et collab., 2016). Lors de l'apoptose, l'ADN chromosomique est fragmenté par une Dnase activée par une caspase. Ces fragments ont une taille comprise entre 180 et 200 pb (soit un nucléosome et l'ADN "linker") ou plus long (dans le cas d'oligonucléosome) (THIERRY et collab., 2016). L'analyse par méthode immunoenzymatique retrouve des quantités de nucléosomes plus importantes dans le plasma de patients atteints de cancers que de sujets sains (HOLDENRIEDER et STIEBER, 2009). Si une grande partie de l'ADNtc provient de l'apoptose des cellules, d'autres mécanismes et d'autres structures sont aussi impliquées (THIERRY et collab., 2016). Les cellules cancéreuses sécrètent une quantité très importante de micro-vésicules qui circulent dans les fluides biologiques des patients (THIERRY et collab., 2016). Ces dernières sont représentées par les exosomes, les corps apoptotiques et des microvésicules hétérogènes (THIERRY et collab., 2016). Ces vésicules contiennent de l'ADN double brin. Les exosomes sont des vésicules de

30-100 nm, sécrétées par une grande variété de cellules (THIERRY et collab., 2016). Composés de lipides et de protéines, les exosomes contiennent aussi de l'ADN. Ce dernier peut être accroché à la membrane de l'exosome, la taille du fragment est alors importante, plus de 2,5 kpb. À l'intérieur des exosomes, l'ADN est double-brin et sa taille varie d'une centaine de pb à 2,5 kpb (THIERRY et collab., 2016). Il existe un intérêt croissant pour les exosomes, car l'extraction d'acides ribonucléiques messagers (ARNm) et ARNmi à partir de ces structures permet d'obtenir une quantité importante de ces acides nucléiques circulant (THIERRY et collab., 2016). Les blebs formés durant l'apoptose provoquent le relargage de vésicules de tailles comprises entre 1 et 5 μ m. Ces vésicules contiennent de l'ADN fragmenté par l'apoptose, qui se retrouve protégé de la dégradation par les Dnases (THIERRY et collab., 2016). Différentes macromolécules sont aussi impliquées dans la biologie des ADNlc. Le virtosome est un complexe de lipoprotéines et d'acides nucléiques, sécrétés par des cellules qui ne se divisent pas, comme les lymphocytes. Lors de leur mise en cultures, de l'ADN est retrouvé dans le milieu de culture (THIERRY et collab., 2016). Ces acides nucléiques sont protégés de l'action des nucléases par les lipoprotéines (THIERRY et collab., 2016). Le virtosome serait un messenger intercellulaire (THIERRY et collab., 2016). Les "Neutrophils extracellular traps" dérivent de la mort des polynucléaires neutrophiles (THIERRY et collab., 2016). Ces macromolécules sont constituées de granules anti-microbiennes et d'ADN des neutrophiles et constituent ainsi une source non-négligeable d'ADNlc (THIERRY et collab., 2016). L'ADNlc peut aussi être associé à d'autres molécules, comme la fibronectine ou la fraction C1q du complément (THIERRY et collab., 2016). Les fluctuations pathologiques des taux protéiques entraînent des fluctuations des taux d'ADNlc (THIERRY et collab., 2016). Enfin, des fragments d'ADN de grande taille (parfois supérieur à 20 kpb) peuvent adhérer aux membranes des cellules sanguines (THIERRY et collab., 2016). Cet ADN peut être nu ou lié aux structures précédemment décrites (THIERRY et collab., 2016).

L'origine d'ADNlc, et les structures auxquelles il est lié, influent sur sa taille et sa fragmentation, deux métriques importantes dans le cadre des ADNtc.

1.2.3 Tailles et fragmentations des ADN tumoraux circulant

La taille des fragments d'ADNlc renseigne aussi sur leur origine et le mécanisme de leur libération (THIERRY et collab., 2016). La non-standardisation des procédures pré-analytiques a entraîné des controverses sur la taille des ADNtc (THIERRY et collab., 2016). Le nombre de paires de bases des ADNtc est en majorité de la taille d'un nucléosome (ou même d'oligonucléosomes) (THIERRY et collab., 2016). Ces fragments proviendraient de l'apoptose des cellules avec la fragmentation de l'ADN

par une Dnase (THIERRY et collab., 2016). Néanmoins, une variance plus grande de la longueur des fragments de l'ADNtc que l'ADNlc "sain" est rapportée. Certaines études rapportent une taille plus grande que l'ADNlc d'origine somatique (UMETANI et collab., 2006b,a). La nécrose peut libérer des fragments plus grands que l'apoptose. En effet, durant ce mécanisme, la fragmentation de l'ADN nucléaire s'effectue de manière plus anarchique. Des fragments jusqu'à 50 kpb ont été retrouvés (THIERRY et collab., 2016). Des études plus récentes montrent un enrichissement des petits fragments (< 145 paires de bases (pb)) (MOULIERE et ROSENFELD, 2015). L'analyse de plasma de patients avec un carcinome hépato-cellulaire (CHC) montre un enrichissement des petits fragments comparé aux plasmas de sujets sains (JIANG et collab., 2015) (**Figure 4**). L'étude du *fingerprinting* des nucléosomes montre aussi une augmentation des fragments de petites tailles lorsque de l'ADNtc est présent (SNYDER et collab., 2016). Durant cette étude les auteurs ont aussi mis au point une méthode de préparation des bibliothèques pour le séquençage nouvelle génération permettant d'améliorer l'amplification des ADN dégradés (SNYDER et collab., 2016).

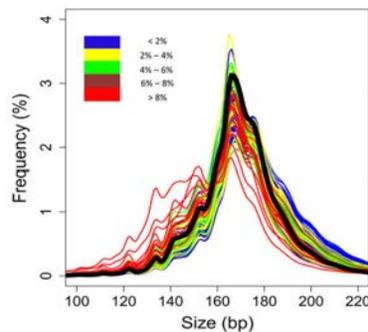


FIGURE 4 – Même si un pic à 166 pb, semblable aux plasmas de sujets sains, est retrouvé, plus la fraction d'ADNtc augmente et plus les fragments d'ADNlc de petites tailles augmentent. La distribution médiane des sujets sains est représentée par la courbe noire épaisse. D'après JIANG et collab. (2015).

L'analyse de la structure des ADNtc par l'équipe d'Alain Thierry a montré une fragmentation importante de ces molécules (THIERRY et collab., 2010; MOULIERE et collab., 2013). La présence de fragments de tailles inférieures à celle d'un nucléosome laisse supposer que l'apoptose n'est pas le seul mécanisme impliqué dans la libération des ADNtc (THIERRY et collab., 2016). Après libération active de l'ADN dans le compartiment circulant la dégradation, par phagocytose ou enzymatique, peut intervenir et conduit à des fragments de tailles inférieures (THIERRY et collab., 2016).

Cette connaissance de la taille et de la fragmentation permettrait *in fine* de mieux caractériser l'origine des ADNlc (cellules tumorales, microenvironnements de la lésion, cellules saines) et de proposer une meilleure détection et quantification des ADNtc (THIERRY et collab., 2016).

1.2.4 Dégradation, élimination des ADN libres circulant et demi-vie

Comme décrit dans les sections précédentes, les ADNlc peuvent être liés ou non à des structures protéiques ou lipoprotéiques. Ces molécules protègent les ADNlc contre une dégradation par des nucléases (THIERRY et collab., 2016). L'élimination de ces acides nucléiques circulant s'effectue au sein du foie ou des reins (THIERRY et collab., 2016). La cinétique de l'élimination des ADNlc a été explorée par l'équipe de Denis Lo en 1999 à partir de plasma de femmes qui accouchaient (LO et collab., 1999). Cette dernière a décrit une élimination biphasique et une demi-vie très courte de 16,3 min des ADNlc d'origine fœtale (LO et collab., 1999). La même équipe a confirmé cette élimination biphasique en 2013, mais avec des demie-vies plus longues (1h pour la phase 1 et 13h pour la phase 2). Enfin, Diehl et collab. ont exploré la demi-vie des ADNtc chez un patient ayant eu une résection complète d'un cancer colorectal et ces auteurs ont décrit une demi-vie de 114 min (DIEHL et collab., 2005). Ainsi, peu de travaux ont exploré la demi-vie des ADNlc et les données obtenues sont conflictuelles. De nouvelles études s'avèrent donc nécessaires afin de déterminer la demi-vie précise des ADNlc et celles-ci doivent notamment s'appuyer sur une meilleure standardisation (THIERRY et collab., 2016).

1.2.5 Implications des ADN tumoraux circulant dans la biologie des tumeurs : messenger intercellulaire, génomemétastase et inflammation

Les ADNtc ne sont pas seulement de simples biomarqueurs potentiels des tumeurs, ils participeraient aussi à l'évolution de ces dernières (THIERRY et collab., 2016). D'abord, les ADNtc seraient des messagers intercellulaires potentiels (THIERRY et collab., 2016). Des travaux ont montré l'existence de virtosomes, complexes lipoprotéiques capables de transporter dans la circulation des ADNlc. Ces particules sont retrouvées dans le surnageant des milieux de cultures (GAHAN et STROUN, 2010). Elles sont aussi capables d'être absorbées par des cellules et les transformer (GAHAN et STROUN, 2010). Une étude a montré une transformation des cellules NIH/3T3 reprenant des virtosomes contenant des mutations *KRAS* de cellules SW480 (ANKER et collab., 1994). Les ADNtc pourraient participer à la génomemétastase en transformant les cellules par transfection (THIERRY et collab., 2016). La culture de cellules NIH/3T3 *KRAS* sauvages avec du plasma d'un patient porteur de mutation *KRAS* montre une transformation de ces premières (GARCÍA-OLMO et collab., 2010) (**Figure 5**). Ces cellules *KRAS* mutantes, lorsqu'elles sont injectées à des souris NOD-SCID, donnent lieu à des tumeurs (GARCÍA-OLMO et collab., 2010).

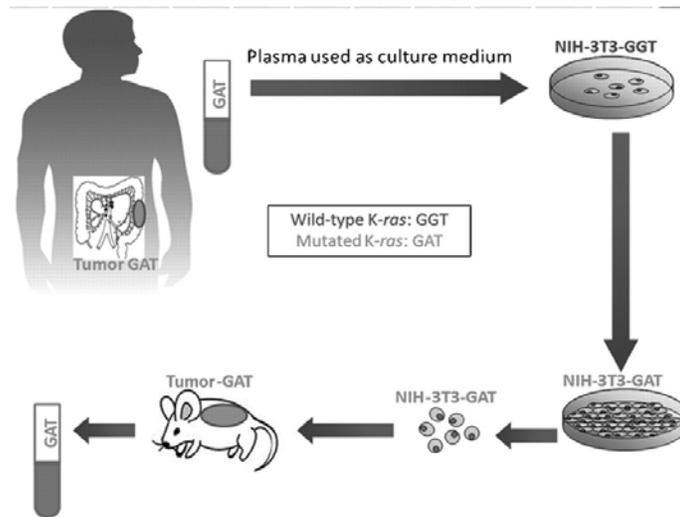


FIGURE 5 – Le plasma d’un patient porteur d’une tumeur avec un mutant *KRAS* a été ajouté au milieu de culture de cellules NIH/3T3. Ces dernières sont ensuite devenues tumorales et la mutation de *KRAS* a été détectée dans les cellules. Enfin, ces dernières ont été injectées a des souris NOD-SCID. Des tumeurs portant la mutations *KRAS* se sont développées et la mutation a aussi été retrouvée dans le plasma des animaux. D’après GARCÍA-OLMO et collab. (2010).

Enfin les ADNlc sont pro-inflammatoires (THIERRY et collab., 2016). L’inflammation est un mécanisme impliqué dans la tumorigenèse et la progression des lésions tumorales. L’exposition aux cellules d’ADN double brin provoque l’activation de gènes qui régulent la synthèse d’interféron (IFN) (THIERRY et collab., 2016). Les structures auxquelles l’ADNlc est complexé peuvent aussi entraîner une réponse inflammatoire (THIERRY et collab., 2016). L’ADNlc peut être reconnu par les Toll-like receptors (TLR) 3-7-8-9. Ces derniers sont des récepteurs impliqués dans l’immunité innée, avec notamment la synthèse de cytokines pro-inflammatoires (Tumor necrosis factor α (TNF- α), interleukine (IL)-1 et IL-6) par l’activation de nuclear factor-kappa B (NF- κ B)/protéine activatrice 1 (AP1) (THIERRY et collab., 2016). Le TLR-9 est impliqué dans la reconnaissance d’oligonucléotides contenant des séquences CpG déméthylées et de l’ADN mitochondrial (THIERRY et collab., 2016). L’ADN tumoral est globalement déméthylé, ainsi sa fraction circulante peut être reconnue par le TLR-9. Ce dernier provoque l’activation de NF- κ B, ce qui entraîne la synthèse de cytokines pro-inflammatoires (THIERRY et collab., 2016). L’ADN mitochondrial se retrouve en grande quantité dans la circulation, notamment dans les CHC (JIANG

et collab., 2015).

Ainsi les ADNtc participent à la physiopathologie tumorale et peuvent représenter de potentielles cibles thérapeutiques (THIERRY et collab., 2016). Mais les ADNtc représentent aussi des biomarqueurs d'intérêt dans toutes les étapes de la prise en charge d'une pathologie cancéreuse (THIERRY et collab., 2016; WAN et collab., 2017) (**Figure 6a**).

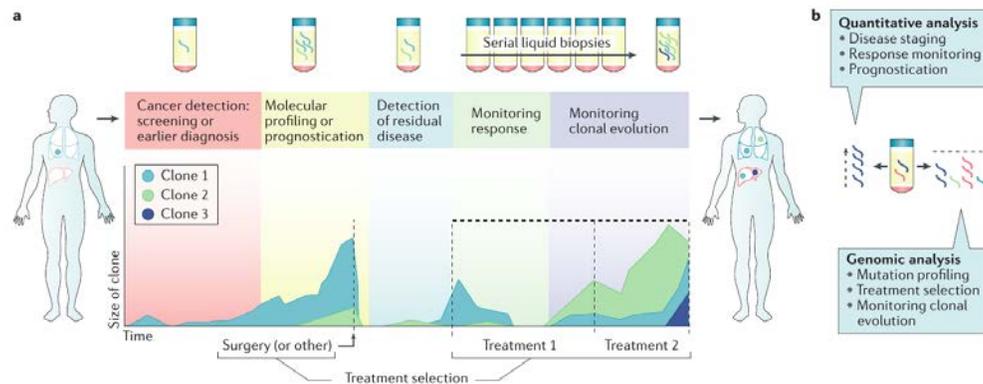


FIGURE 6 – **a**. Les ADNtc peuvent être utilisés comme biomarqueurs dans chaque phase de la maladie. D'abord, ils peuvent être utilisés dans la phase de *screening* et le diagnostic précoce. Ensuite, les ADNtc sont d'intérêt pronostic et grâce à leur demi-vie faible et leur grande spécificité, peuvent servir dans le suivi de la maladie résiduelle. Enfin, les ADNtc sont aussi utiles dans la théranostic, pouvant remplacer les biopsies tissulaires. Les ADNtc permettent aussi de détecter des résistances aux thérapies de manière précoce. **b**. Les ADNtc peuvent donner des informations sur deux plans : quantitatif et qualitatif. La partie qualitative est corrélée à la masse tumorale tandis que la seconde rapporte des informations génomiques. En point unique, le stade, le pronostic ou la théranostique peuvent être évalués. En étude longitudinale, la réponse au traitement et l'évolution clonale sont explorées par l'analyse des ADNtc. D'après WAN et collab. (2017).

1.3 Utilités cliniques des ADN tumoraux circulant

Il existe une forte demande d'amélioration des marqueurs existants dans les cancers (WAN et collab., 2017). Les ADNtc sont des biomarqueurs potentiels pour chaque stade de la maladie (WAN et collab., 2017) (**Figure 6a**). Cela est possible car les ADNtc permettent aussi d'obtenir des informations quantitatives (en rapport avec la masse tumorale) et qualitative (génomique somatique) (WAN et collab., 2017) (**Figure 6b**).

1.3.1 *Screening* et diagnostic

Un diagnostic précoce améliore le pronostic d'une pathologie tumorale. En effet, un patient traité aux premiers stades de la maladie peut être pris en charge avec une chirurgie curative (WAN et collab., 2017). Des marqueurs, comme la recherche de sang dans les selles pour le cancer colorectal, permettent un diagnostic précoce et une diminution de la mortalité de ce cancer (une méta-analyse rapporte une diminution de 16% de la mortalité grâce à l'utilisation de l'Hémocult®) (HEWITSON et collab., 2008). L'imagerie est aussi utilisée dans le diagnostic des cancers mais elle peut entraîner une exposition accrue aux radiations ionisantes (SMITH-BINDMAN et collab., 2009). Néanmoins, même s'il existe de nombreuses avancées dans ces marqueurs afin d'améliorer leur sensibilité et spécificité, la détection de la maladie à son stade le plus précoce demeure difficile. De nombreux patients présentent des métastases lors du diagnostic initial (par exemple 20% des patients avec un cancer colorectal) (LEUFKENS et collab., 2011; BRENNER et collab., 2014). Les ADNtc, marqueurs non-invasifs, peuvent améliorer la détection précoce des cancers (DIAZ JR et BARDELLI, 2014; WAN et collab., 2017). En effet, comme dit précédemment, les ADNtc peuvent se retrouver dans des fluides biologiques différents (plasma, selles, salives, crachats, etc.) (THIERRY et collab., 2016). Plusieurs méthodes peuvent être utilisées pour la détection des cancers au stade précoce : la concentration en ADNtc, l'analyse de la fragmentation, la recherche d'aberrations génomiques somatiques (mutations, nombres de copies, épigénétiques) (WAN et collab., 2017).

Des premiers travaux se sont intéressés à la concentration totale d'ADNtc. Les masses tumorales sont une source plus importante que les cellules saines d'ADNtc, ce qui entraîne une plus forte concentration chez les sujets atteints de cancer (THIERRY et collab., 2016). Des études ont montré que les patients avec un cancer du poumon non à petites cellules (CPNPC) présentent une concentration plasmatique médiane en ADNtc huit fois plus importante que des fumeurs non atteints (SOZZI et collab., 2001, 2003). La sensibilité et la spécificité de ce simple test sont respectivement de 75% et

86% (SOZZI et collab., 2001, 2003). Dans les cancers colorectaux, la concentration moyenne des ADNlc a été retrouvée 50 fois supérieure à celle des contrôles (BONI et collab., 2007). Mais ces études n'étaient pas standardisées, d'un point de vue pré-analytique comme analytique.

Un index d'intégrité des ADNlc a été proposé comme biomarqueur afin de détecter précocement les cancers (UMETANI et collab., 2006b). L'ADNlc provenant aussi de nécrose cellulaire, des fragments plus grands peuvent être retrouvés (JAHR et collab., 2001). La méthode d'analyse mise au point par Umetani et collab. se base sur une réaction en chaîne par polymérase en temps réel (PCRtr) des séquences ALU (UMETANI et collab., 2006a). Ces séquences représentent environ 10% du génome humain et sont dispersées sur l'ensemble des chromosomes (LANDER et collab., 2001). Les aire sous une courbe ROC (AuROCC) de cette méthode pour les cancers colorectaux, les cancers péri-ampullaires et mammaires ont été évaluées respectivement à 0,78, 0,80 et 0,79 (UMETANI et collab., 2006a,b). Mais ces méthodes qui se basent uniquement sur une concentration ou une intégrité ne sont pas idéales. En effet, des pathologies autres que tumorales peuvent provoquer une augmentation des concentrations de ADNlc (THIERRY et collab., 2016). De même, il est admis aujourd'hui que les ADNlc sont plus petits que les ADNlc issus de cellules saines (THIERRY et collab., 2016). Ainsi ces tests peuvent conduire à des faux positifs et des faux négatifs.

Des méthodes basées sur une analyse des mutations somatiques des cancers ont été aussi utilisées pour détecter les cancers dans les premiers stades de la pathologie. La recherche de mutations somatiques *hotspot* comme *KRAS* ou B-Raf proto-oncogene, serine/threonine kinase (*BRAF*) dans les liquides biologiques représente un test hautement spécifique. De plus, des techniques ultra-sensibles, qui seront développées ultérieurement au sein de ce travail, ont été mises au point (WAN et collab., 2017). Mais ces mutations doivent être présentes dès les premiers stades de la maladie afin que le test soit spécifique. Les mutations au sein de l'exon 2 de *KRAS* remplissent ces conditions pour les cancers pancréatiques (ALMOGUERA et collab., 1988). Une étude a montré qu'une absence de mutation de *KRAS* dans le plasma, couplé à l'antigène carbohydre 19-9 (CA 19-9), a une meilleure valeur prédictive négative (VPN) que le CA 19-9 seul (MAIRE et collab., 2002). Mais cette approche de recherche de mutations uniques ou de quelques mutations manque de sensibilité (WAN et collab., 2017). Pour s'affranchir de cette limitation, des méthodes ciblant plusieurs altérations somatiques ont été mises au point (WAN et collab., 2017).

Les altérations épigénétiques peuvent être impliquées dans l'initiation tumorale (JONES et BAYLIN, 2002; CHIK et collab., 2011). Ainsi ces modifications somatiques

sont présentes aux premiers stades de la maladie et peuvent servir de marqueurs de diagnostic précoce (FUJIWARA et collab., 2005; WARTON et SAMIMI, 2015). Une des méthodes permettant d'évaluer la méthylation du promoteur de septin 9 (*SEPT9*), Epi proColon Assay®(Epigenomics AG, Berlin, Allemagne), est disponible en Europe (avec marquage CE) et est approuvée par la Food and Drugs Administration (FDA). Une partie de la population est réticente à effectuer les tests au gâïac ou immunologiques qui recherchent la présence de sang dans les selles (SCHOLEFIELD et collab., 2011; YOUNG et collab., 2012). De plus ces méthodes manquent de spécificité (LAMB et DHILLON, 2017). Un test basé sur une prise de sang permettrait d'augmenter la compliance de la population générale (LAMB et DHILLON, 2017). Cette analyse se déroule en deux étapes principales : la préparation de l'échantillon (extraction, bisulfitation) et la réaction en chaîne par polymérase (PCR) en elle-même (LAMB et DHILLON, 2017) (**Figure 7**).

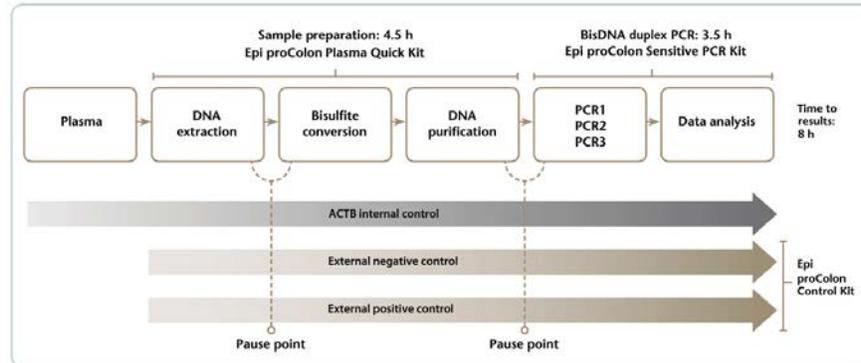


FIGURE 7 – L'analyse du kit Epi proColon Assay se déroule en deux étapes. D'abord une extraction de l'ADN à partir du plasma d'un patient, suivi d'une bisulfitation et une purification. Ensuite, une PCR en triplicat est réalisée sur 45 cycles. L'*ACTB* (β -actine) est utilisée comme témoin d'amplification. D'après POTTER et collab. (2014).

Les premières études cas-contrôles qui ont évalué cette technique montrent une forte sensibilité et spécificité (LAMB et DHILLON, 2017). Une étude prospective à grande échelle, l'essai PRESEPT (NCT00855348), a été menée aux États-Unis et en Allemagne. Celle-ci a recruté plus de 7900 sujets, dont 1516 ont été analysés par le kit : 53 cas de cancer colorectal et 1463 sujets sains (CHURCH et collab., 2014). La sensibilité de ce test a été plus basse que dans les études cas-contrôles (50,9%) mais la spécificité reste élevée (91,5%) (CHURCH et collab., 2014). Une réévaluation des cas en modifiant les seuils décisionnels (2 tests positifs au lieu de 3) a permis d'augmenter la sensibilité à 68% au prix d'une spécificité à 80% (POTTER et collab., 2014). D'autres méthodes d'analyse de la méthylation existent, comme le méthyl-BEAMing, utilisé aussi dans le cancer colorectal (LI et collab., 2009b). Cette analyse du promoteur de la vimentine a montré une sensibilité de 59% (52% pour les premiers stades de la maladie) et une spécificité de 93% (LI et collab., 2009b). Dans les CPNPC, les méthylations anormales des promoteurs de cyclin dependent kinase inhibitor 2A (*CDKN2A*) et O-6-méthylguanine-DNA methyltransferase (*MGMT*) peuvent être détectées dans les crachats jusqu'à 35 mois avant le diagnostic clinique (PALMISANO et collab., 2000). Cette hyperméthylation est très spécifique des lésions tumorales et n'est retrouvée que chez 3% des sujets à haut risque (fumeurs ou mineurs d'uranium) (PALMISANO et collab., 2000). Une autre étude a montré que la recherche de méthylation des promoteurs de quatre gènes (*DCC*, *Kif1a*, *NISCH*, et *Rarb*) permet de différencier les patients avec des résultats d'imagerie douteux (OSTROW et collab., 2010). Les auteurs ont montré que la consommation de cigarettes influence sur la méthylation (OSTROW et collab., 2010). La méthylation d'un seul promoteur chez les patients n'ayant pas fumé ou moins de 20 paquets-année a une spécificité de 100% pour le CPNPC (OSTROW et collab., 2010). Chez les sujets ayant une consommation de plus de 20 paquets-années, 30% des patients du groupe contrôle montrent une méthylation (OSTROW et collab., 2010). C'est pourquoi des sujets ayant une consommation de tabac importante doivent être inclus dans les contrôles afin de déterminer la spécificité de façon plus rigoureuse (OSTROW et collab., 2013).

Avec une méthode analysant un grand nombre de mutations et qui permet de détecter une copie de mutation pour 5 mL de plasma, les ADNtc se retrouvent dans le plasma de 82% des patients aux stades IV et seulement 47% des patients ont des ADNtc au stade I (BETTEGOWDA et collab., 2014). De plus, pour certaines atteintes tumorales, notamment des patients atteints de gliome, ce pourcentage est encore plus bas (BETTEGOWDA et collab., 2014). Plusieurs approches ont été proposées pour améliorer la détection précoce des cancers (WAN et collab., 2017) (**Figure 8**).

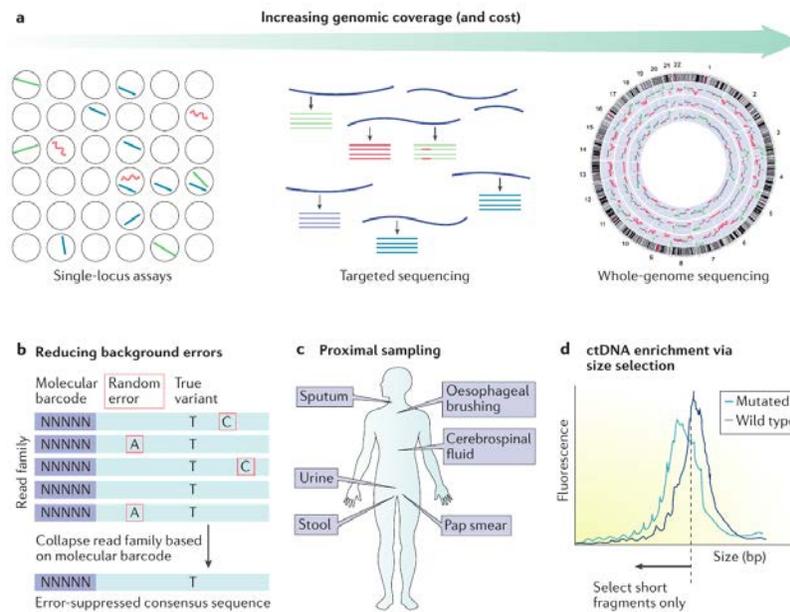


FIGURE 8 – **a.** Les méthodes en point unique ont l’avantage d’être pour la plupart standardisées et possèdent une grande sensibilité. Néanmoins elles n’interrogent qu’une seule altération somatique, qui peut ne pas être retrouvée dans tous les cancers. Les méthodes pouvant analyser de multiples *loci* peuvent détecter un plus grand nombre de cancers différents, mais leur sensibilité est limitée par le coût et le bruit de fond. **b.** Des méthodes comme le *barcoding* permettent de diminuer le bruit de fond des analyses haut débit, permettant une amélioration de leur sensibilité. **c.** D’autres fluides biologiques que le plasma peuvent être considérés pour la recherche d’ADNtc. En effet, dans certains cancers, les fluides plus proches (par exemple l’urine dans le cancer de la vessie) ont des concentrations en ADNtc supérieures aux plasmatiques. **d.** Les fragments d’ADNtc sont plus petits que les ADNtc issus de cellules saines. Une des approches consisterait en la sélection des fragments plus petits, ce qui enrichirait en ADNtc. D’après WAN et collab. (2017).

Les ADNtc sont en effet relativement fragiles et avec une demi-vie courte, la première approche a été l’optimisation et la standardisation des méthodes pré-analytiques et d’extraction des ADNtc (WAN et collab., 2017). Ensuite, pour certaines tumeurs, des fluides biologiques contiennent une fraction d’ADNtc plus élevée que le plasma (WAN et collab., 2017). C’est le cas des selles pour les cancers colorectaux, de l’urine pour les cancers de la vessie, des crachats pour les CPNPC ou des lavages utérins pour les cancers cervicaux (WAN et collab., 2017) **Figure 8c**).

Des méthodes d'analyses se sont basées sur la recherche simultanée d'un nombre élevé de mutations. La technologie Digital SequencingTM est une analyse de 50 gènes oncogéniques en séquençage haut débit avec un *barcoding* qui permet de diminuer par 100 les erreurs lors de l'analyse (LANMAN et collab., 2015). Avec cette méthode, de l'ADNtc était détecté chez 50% des patients au stade I d'un CPNPC (LANMAN et collab., 2015). Une technique utilisant la *droplet digital Polymerase Chain Reaction* (ddPCR) pour analyser les mutations de phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (*PIK3CA*) dans les cancers du sein a montré une sensibilité de 93,3% lorsque les tumeurs portaient d'une mutation de ce gène (BEAVER et collab., 2014).

Mais ces méthodes se basant sur des mutations ne sont pas adaptées à la détection de toutes les pathologies tumorales (MOULIERE et ROSENFELD, 2015). Certains cancers ont très peu de mutations somatiques et aucune *hotspot* (par exemple, les CHC ont environ 4,2 mutations par megabase) (MOULIERE et ROSENFELD, 2015). Néanmoins, ces tumeurs sont caractérisées par d'autres aberrations génomiques, comme des variants du nombre de copies ou des altérations épigénétiques. Les CHC sont caractérisés par des aberrations structurales de grandes tailles, touchant le chromosome 1 et le 8 (KAN et collab., 2013). Les bras courts (p) de ces chromosomes sont fréquemment délétés tandis que les longs (q) sont amplifiés dans les CHC (KAN et collab., 2013). Jiang et collab. ont mis au point un test basé sur un *z-score* qui permet de déterminer si des altérations du nombre de copies diffèrent par rapport à un groupe contrôle (JIANG et collab., 2015). Ils ont recruté pour cette étude 225 sujets, dont 90 patients atteints de CHC, des sujets sains mais aussi des patients infectés chroniquement par le virus de l'hépatite B, avec ou sans cirrhose (JIANG et collab., 2015). Ce test a permis de trouver au moins une variation du nombre de copies (VNC) chez 84,4% des patients avec un CHC et aucune altération chez les sujets sains (JIANG et collab., 2015). De plus, les auteurs ont démontré que les fragments portant les altérations sont plus petits que ceux des sujets sains (JIANG et collab., 2015). Deux patients infectés par le virus de l'hépatite B, un avec et un sans cirrhose, ont montré des altérations chromosomiques dans le plasma et ont été diagnostiqués atteints de CHC quelques mois plus tard (JIANG et collab., 2015). En plus de démontrer l'utilité des altérations chromosomiques dans le diagnostic des CHC, Jiang et collab. ont montré que l'ADN mitochondrial, présentait un intérêt en tant que biomarqueur diagnostique (JIANG et collab., 2015). Les fractions d'ADN mitochondrial pour les patients atteints de CHC et les sujets sains étaient respectivement de 0,0014% et 0,0004% (JIANG et collab., 2015). Avec un seuil de 0,00084%, l'AuROCC est de 0,93 tandis que la sensibilité était de 80% et la spécificité de 94% (JIANG et collab., 2015). Les cancers ovariens ont aussi un nombre élevé de variants

du nombre de copies (THE CANCER GENOME ATLAS RESEARCH NETWORK, 2011). Une méthode dérivant de celle employée en diagnostic pré-natal non invasif (DPNI) a été utilisée pour le diagnostic des carcinomes ovariens sévères de haut grade (COSHG) (COHEN et collab., 2016). Les techniques utilisées en DPNI permettent d'identifier d'éventuelles l'aneuploïdies fœtales à partir du plasma de la mère (COHEN et collab., 2016). Cette technique a été évaluée chez 32 patientes, 16 aux stades FIGO I-II et 16 aux stades FIGO III-IV (COHEN et collab., 2016). Une altération a été détectée chez 13 patientes, dont 6/16 patientes aux premiers stades (COHEN et collab., 2016). Les altérations de l'épigénome peuvent aussi être analysées en points multiples. Les patients avec une pancréatite chronique sont à risque de développer un cancer pancréatique (LIGGETT et collab., 2010). Néanmoins le diagnostic précoce de la tumeur est difficile, les symptômes entre les deux pathologies étant similaires (LIGGETT et collab., 2010). Une puce a été mise au point pour évaluer la méthylation de 56 fragments de promoteurs (COHEN et collab., 2016). Cette méthode a une sensibilité de 76% et une spécificité de 59% (COHEN et collab., 2016). Enfin, pour les cancers viro-induits, comme les cancers naso-pharyngés ou du col de l'utérus, la détection de l'intégration du génome viral dans l'ADN peut représenter un test plus utile que la détection d'autres altérations somatiques (HOLMES et collab., 2016; WAN et collab., 2017). Ces analyses permettraient de détecter précocement les patients porteurs de lésions pré-cancéreuses (CHAN et collab., 2013; GUERRERO-PRESTON et collab., 2016). L'avantage de ces méthodes en multi-points est que la sensibilité de détection est améliorée sans augmenter la prise d'essai (WAN et collab., 2017). En effet le nombre d'altérations somatiques présentes dépend de la quantité de plasma prélevée, ainsi des méthodes comme la plasmaphérèse ou des ligands des ADNtc pourraient représenter des méthodes d'enrichissement (WAN et collab., 2017). Des méthodes analytiques ne sélectionnant que les fragments de petites tailles (plus fréquemment d'origine tumorale que saine) peuvent enrichir en ADNtc (WAN et collab., 2017). La méthode idéale serait d'analyser le génome entier circulant avec une forte couverture (WAN et collab., 2017). Mais le coût de cette approche est trop élevé de même que l'analyse bioinformatique est à fort risque de faux-positifs (WAN et collab., 2017). De plus, pour être certains du diagnostic de cancer, il est nécessaire de mettre en évidence des altérations spécifiques (par exemple des mutations sur *KRAS* ou tumor protein p53 (*TP53*)). Une étude en séquençage haut débit de biopsies cutanées issues de peaux saines exposées aux rayons ultra-violet a montré une accumulation de mutations somatiques (MARTINCORENA et collab., 2015). En effet, quatre à six mutations par megabase ont été retrouvées, avec des signatures caractéristiques des carcinomes épidermoïdes cutanés (MARTINCORENA et collab., 2015). De plus, les oncogènes (comme les gènes impliqués dans la voie NOTCH) ou les gènes sup-

presseurs de tumeurs (comme *TP53*) portent les mêmes mutations que les carcinomes épidermoïdes cutanés (**Figure 9a**). Les mutations s'accumulent même dans plusieurs sous-clones (MARTINCORENA et collab., 2015) (**Figure 9b**). Ce travail montre que les mutations dans les gènes "driver" peuvent être remarquablement élevées dans la peau saine (MARTINCORENA et collab., 2015). Les clones peuvent être de taille importante, ce qui peut entraîner un relargage de ADNlc dans lequel des mutations de gènes "driver" sont présentes (WAN et collab., 2017). Cela entraînerait une augmentation du "bruit biologique" et diminuerait la spécificité de l'analyse (WAN et collab., 2017). En outre, il est retrouvé chez 10% des sujets de plus de 65 ans des clones hématopoïétiques avec des mutations associées aux leucémies (GENOVESE et collab., 2014). Néanmoins, le risque d'évolution vers une pathologie oncohématologique est de l'ordre d'un pourcent (GENOVESE et collab., 2014). Néanmoins, la présence d'une mutation est un facteur de risque de développer une pathologie hématologique (*Hazard ratio* (HR) 12,9, Intervalle de confiance à 95% (IC95%) 5,8–28,7) (GENOVESE et collab., 2014). Enfin, des mutations associées aux lésions cancéreuses ont été retrouvées dans le plasma de sujets sains (GORMALLY et collab., 2006).

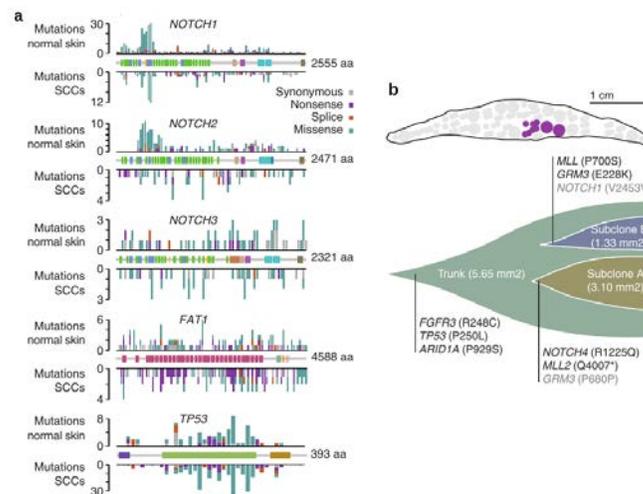


FIGURE 9 – **a**. Les mutations retrouvées dans cinq gènes "driver" montrent une distribution ressemblant à celle retrouvée dans les cancers épidermoïdes cutanés et de la tête et cou. **b**. Représentation d'un clone de taille importante qui est retrouvé dans six biopsies. Plusieurs sous-clones coexistent, avec une accumulation de mutations dans les gènes "driver". D'après MARTINCORENA et collab. (2015).

Ainsi, la détection d'une altérations génomique dans l'ADNlc, même si la mutation touche un gène "driver", ne constitue pas un diagnostic de certitude du cancer. Il est nécessaire de suivre l'évolution de la clinique des sujets sains chez lesquels une mutation somatique a été détectée dans un fluide biologique (WAN et collab., 2017). Enfin, si de nombreuses méthodes ont été mises au point, peu ont été évaluées avec une étude randomisée, contrôlée et de forte puissance.

1.3.2 Localisation des lésions cancéreuses

Une partie non négligeable des patients présente des métastases dès le diagnostic initial de cancer (LEUFKENS et collab., 2011; BRENNER et collab., 2014). Il s'avère parfois délicat de caractériser le tissu primitif, on parle alors de métastase(s) de primitif inconnu . Ces lésions cancéreuses représentent 3% de la totalité des cancers (PAVLIDIS et collab., 2003). La chimiothérapie dépend du tissu d'origine des lésions cancéreuses, les patients avec des métastases de primitif inconnu sont à risque de présenter des résistances au traitement (PAVLIDIS et collab., 2003; ROSENFELD et collab., 2008). La biologie moléculaire présente un intérêt dans la détermination du tissu primitif (ROSENFELD et collab., 2008). Le "footprinting" des nucléosomes ainsi que la méthylation sont spécifiques des tissus et peuvent être retrouvés dans le plasma (SNYDER et collab., 2016; LEHMANN-WERMAN et collab., 2016). Les altérations du méthylome sont plus dispersées au sein du génome, comparé aux mutations somatiques (GUO et collab., 2017). L'analyse par bloc haplotypique mise au point par Guo et collab. permet de mieux couvrir cette dispersion (GUO et collab., 2017). Une valeur de méthylation globale des blocs, appelée *mutation haplotype load* peut ainsi être calculée (GUO et collab., 2017). Cette mesure permet la classification des tissus, ces derniers ayant une méthylation qui leur est spécifique (GUO et collab., 2017). À partir du plasma de 59 patients atteints de cancers du poumon et colorectaux et 75 sujets sains, cette méthode permet de déterminer le tissu d'origine de la méthylation dans l'ADNlc (GUO et collab., 2017). De plus, cette analyse permet d'estimer la part d'ADN tumoral dans l'ADNlc (GUO et collab., 2017).

L'analyse des ADNtc permettrait de déterminer le tissu d'origine des lésions tumorales, en complément des méthodes anatomo-pathologiques ou en remplaçant ces dernières, notamment dans les cas où la biopsie est impossible.

1.3.3 Pronostic, suivi de la maladie résiduelle et du risque de récurrence

Les ADNtc peuvent représenter des biomarqueurs pronostic, pour le suivi de la maladie résiduelle ou le risque de récurrence. Ce dernier est parfois difficile à caractériser mais néanmoins essentiel d'évaluer le bénéfice d'une chimiothérapie adjuvante par exemple. Dans le cancer du sein, les patientes qui ont eu une exérèse doivent être traitées par une chimiothérapie adjuvante si elles ont un fort risque de récurrence. Aujourd'hui des critères tumoraux, ganglionnaires et immunohistochimiques sont employés afin de discuter de l'indication d'une chimiothérapie adjuvante (OAKMAN et collab., 2010; EARLY BREAST CANCER TRIALISTS' COLLABORATIVE GROUP AND OTHERS, 2012). Ce traitement n'apporte aucun bénéfice aux patientes avec un risque de récurrence faible, en outre sa toxicité est importante (OAKMAN et collab., 2010; EARLY BREAST CANCER TRIALISTS' COLLABORATIVE GROUP AND OTHERS, 2012). Une signature de 70 gènes de mauvais pronostic a été isolée par étude du transcriptome et commercialisée sous le nom de Mammaprint® (VAN'T VEER et collab., 2002; OAKMAN et collab., 2010). Un essai clinique dans lequel 6693 femmes ont été incluses a montré que cette signature pouvait aider à sélectionner les patientes (CARDOSO et collab., 2016). En effet, les patientes avec un fort risque clinique mais une signature de faible risque pouvaient ne pas bénéficier de l'effet d'une chimiothérapie adjuvante (CARDOSO et collab., 2016). Quarante-six pour cent des patientes pourraient ne recevoir aucune chimiothérapie adjuvante au prix d'une augmentation de 1,5 point de pourcentage de développer une métastase à cinq ans (CARDOSO et collab., 2016). Néanmoins, cette signature se base sur la biopsie de la tumeur initiale, qui est parfois impossible à effectuer. De plus, cette méthode est basée sur l'expression des gènes donc sur les acides ribonucléiques (ARN). Même s'il existe une amélioration des protocoles d'échantillonnage, d'extraction de conservation et d'analyses des biopsies en vue de l'analyse du transcriptome, les ARN demeurent fragiles et leur dégradation modifie celle-ci (RAVO et collab., 2008; OPITZ et collab., 2010). Le dosage de protéines et de carbohydrates, comme le CA 19-9, l'antigène tumoral 125 (CA 125), le *Prostate Specific Antigen* (PSA) ou l'antigène carcino-embryonnaire (ACE), est aussi utilisé pour détecter une récurrence (OESTERLING, 1991; BALLEHANINNA et CHAMBERLAIN, 2012). Néanmoins, peu de marqueurs fiables ont été trouvés pour le suivi des cancers (DUFFY, 2001). De plus, la spécificité et la sensibilité de ces analytes n'est pas excellente. Par exemple, le CA 19-9 n'est pas exprimé chez les patients Lewis négatif et la cholestase entraîne une augmentation des taux de ce carbohydrate (OHSHIO et collab., 1990; NARIMATSU et collab., 1996). De plus ces analytes ont une demi-vie longue (ITO et collab., 2002; RIEDINGER et collab., 2006). Cette cinétique limite leur utilité, notamment pour déterminer si une exérèse

a été complète ou non. En outre, les marqueurs protéiques possèdent des variations peu importantes. Par exemple, le suivi de la masse tumorale des mélanomes se base actuellement sur plusieurs marqueurs comme la S100 calcium-binding protein B (S100B) ou la lactate déshydrogénase (LDH) (CALAPRE et collab., 2017). Pour cette dernière, les variations s'échelonnent sur 2 log, tandis que les ADNtc fluctuent sur 4 voire 5 log (CALAPRE et collab., 2017). Enfin, le suivi de la pathologie tumorale implique l'imagerie. Cette méthode, bien qu'en constante évolution, est limitée par certains aspects (FRANGIONI, 2008). L'itération des examens entraîne une exposition importante aux rayons ionisants (SMITH-BINDMAN et collab., 2009). De plus, le tissu cicatriciel post-radiothérapie peut être confondu avec une masse tumorale (FRANGIONI, 2008). Par exemple, dans le cancer colorectal, de nombreux faux positifs sont rapportés avec l'utilisation du scanner pour le suivi (CHAO et GIBBS, 2009). Les lésions tumorales peuvent aussi être de taille inférieure au seuil de détection des techniques d'imagerie (FRANGIONI, 2008). Enfin ces dernières ont des limites pour le suivi de certains cancers, comme celui de l'ovaire (FRANGIONI, 2008).

La croissance d'une tumeur solide suit généralement une cinétique de Gompertz (NORTON et collab., 1976; BENZEKRY et collab., 2014). Le point de départ d'une lésion est une cellule unique, qui poursuit son expansion clonale de façon stationnaire, une phase exponentielle avec un *switch* angiogénique et une deuxième phase stationnaire culminant avec la mort du patient (FRANGIONI, 2008) (**Figure 10**). Le seuil de détection d'une lésion tumorale est de l'ordre du gramme, soit 10^9 cellules (FRANGIONI, 2008). Cela implique que le patient est déclaré en rémission mais il peut porter de 1 à 10^9 cellules tumorales et la lésion peut déjà être vascularisée (FRANGIONI, 2008). Ainsi, les seuils de détection d'une récurrence ne sont pas assez sensibles pour assurer une prise en charge optimale (FRANGIONI, 2008). De plus, dans le cancer du sein, un nombre important de patientes présente des micrométastases au stade du diagnostic et du fait de leur taille, ne seront pas réséquées par l'intervention chirurgicale (HOWLADER et collab., 2014; GARCIA-MURILLAS et collab., 2015). Les ADNtc peuvent représenter des biomarqueurs de choix pour le suivi de la maladie tumorale (WAN et collab., 2017). En effet, leur spécificité est grande et l'évolution des techniques augmente leur sensibilité (WAN et collab., 2017). De plus, leur demi-vie courte en font un biomarqueur de choix, notamment dans les situations post-exérèse (THIERRY et collab., 2016).

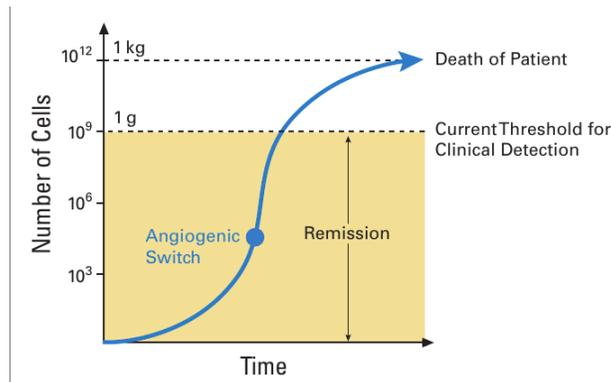


FIGURE 10 – Certaines tumeurs, comme celles d’origine mammaire, suivent une cinétique de Gompertz. Après une phase stationnaire où un seul clone se multiplie, un *switch* angiogénique intervient. Cette néovascularisation permet à la tumeur d’échapper à l’apport des nutriments par diffusion uniquement et entraîne une phase de croissance exponentielle. Enfin, une phase stationnaire a lieu, allant jusqu’à la mort du patient. Les méthodes diagnostiques actuelles permettent la détection d’une récurrence à partir de 10^9 cellules soit environ un gramme de tissu. D’après FRANGIONI (2008).

Les premiers travaux qui se sont intéressés à cet aspect ont néanmoins fait appel à des méthodes peu sensibles. Les plasmas dans lesquels une mutation était détectée provenaient de patients avec une forte masse tumorale. Dans le cas des cancers colorectaux, plusieurs études ont évalué l’intérêt de l’ADNtc et le pronostic. La détection de mutants *KRAS* ou l’hyperméthylation du promoteur de p16 dans le plasma de patients est associée à une survie plus courte (LECOMTE et collab., 2002). En effet, le taux de survie à deux ans a été estimé à 48% dans le groupe avec une altération détectée *versus* 100% si aucune altération n’a été mise en évidence ($p < 0,03$) (LECOMTE et collab., 2002). Une étude prospective sur une cohorte de 94 patients a mis en évidence que la persistance de la mutation *KRAS* dans le sérum est associée à un fort risque de récurrence (RYAN et collab., 2003). La mutation *KRAS* a été mise en évidence dans 60 biopsies de la tumeur (RYAN et collab., 2003). Au sein de ces 60 patients, 16 (27%) avaient une persistance de la mutation *KRAS* dans le sérum en post-exérèse (RYAN et collab., 2003). Dix d’entre eux ont eu une récurrence et un seul dans les 44 patients *KRAS* sérique négatif (*Odds ratio* (OR) 71,7 IC95% 7,7–663,9; $p < 0,0001$) (RYAN et collab., 2003). Dans le cancer du poumon, une étude sur la méthylation du promoteur de 14-3-3 σ dans le sérum a montré une utilité en tant que biomarqueur pronostic (RAMIREZ et collab., 2005). Les patients chez lesquels

une méthylation était détectée avaient une survie médiane plus longue (15,1 mois *vs.* 9,8 mois, $p = 0,004$) (RAMIREZ et collab., 2005). De même, dans le CPNPC, les mutations sur les exons 19 ou 21 d'epidermal growth factor receptor (*EGFR*) ont été recherchées dans le plasma de patients avec des sondes TaqMan (JIAN et collab., 2010). Comparée aux patients chez lesquels aucune altération n'était détectée, la survie globale (11,2 *vs.* 2,7 mois $p = 0,005$) était plus longue (JIAN et collab., 2010).

La première méthode sensible utilisée fut le BEAMing dans le cancer colorectal (DIEHL et collab., 2005, 2008). Une cohorte de 18 patients, qui ont eu une ou plusieurs résections, a été constituée (DIEHL et collab., 2008). Pour les patients avec une exérèse complète, le taux d'ADNtc diminue fortement au jour de sortie (avec une diminution médiane de 99%) (DIEHL et collab., 2008). Cette diminution apparaît même dès le premier jour post-intervention (diminution médiane de 96,7%) et les prélèvements multiples chez un patient ont montré une demi-vie de 114 minutes des ADNtc (DIEHL et collab., 2008). Dans le cas d'une résection incomplète, la diminution des ADNtc est bien moins importante (diminution de 55-56%) (DIEHL et collab., 2008). Trois patients ont présenté une augmentation des ADNtc, certainement liée aux lésions sur le tissu tumoral restant provoqué par l'intervention (DIEHL et collab., 2008). Au niveau du suivi post-opératoire, il a été possible dans 20 cas : quatre avec une disparition complète des ADNtc et 16 avec des taux résiduels (DIEHL et collab., 2008). Il est intéressant de noter qu'aucun patient avec une disparition n'a eu de récurrence, tandis que 15 patients avec des ADNtc encore détectables ont eu une récurrence (test *log-rank* $p = 0,006$) (DIEHL et collab., 2008). Les ADNtc présentent aussi de meilleures performances en tant que biomarqueurs que l'ACE (sensibilité supérieure en situation pré-chirurgicale, diminution plus importante après l'exérèse) (DIEHL et collab., 2008). De plus, le risque de récurrence est mieux apprécié par les ADNtc que l'ACE (test *log-rank* de l'ACE $p = 0,03$). La mise au point d'amorces spécifiques à partir de variants structuraux (VS) spécifiques a montré que la détection des ADNtc était plus précoce de 2 à 15 mois avant la récurrence clinique (REINERT et collab., 2016). Cette approche, bien que complexe, permet une plus grande sensibilité car les VS sont moins sensibles au bruit de fond que les mutations (REINERT et collab., 2016; WAN et collab., 2017). Une étude prospective dans laquelle 230 sujets avec des cancers colorectaux de stade II ont été inclus a montré que la détection d'ADNtc après une chirurgie était associée à un fort risque de récurrence (HR=18, IC95% 7,9-40) (TIE et collab., 2016). Ce dernier est bien plus grand que les scores cliniques validés (HR=3,3, IC95% 1,6-7,0) (TIE et collab., 2016). De plus, la détection d'une récurrence par les ADNtc était plus précoce que par l'ACE (médiane de 167 jours *vs.* 61, $p=0,04$) (TIE et collab., 2016). Une étude longitudinale a montré que l'ADNtc peut représenter un biomarqueur de choix dans le suivi de la

maladie résiduelle et le suivi dynamique de la tumeur (SCHØLER et collab., 2017). Durant le suivi longitudinal de 27 patients, la présence d'ADNtc a été trouvée chez tous les patients avec une récurrence et chez aucun patient sans récurrence (SCHØLER et collab., 2017). Vingt-et-un patients ont été suivis en post-opératoire : après trois mois, de l'ADNtc a été détecté chez six patients (SCHØLER et collab., 2017). Tous ont eu une récurrence en comparaison à seulement quatre dans le groupe ADNtc-négatif (HR=37,7, IC95% 4,2–335,5) (SCHØLER et collab., 2017). De plus, les ADNtc montrent une corrélation avec le volume de la tumeur ($\kappa=0,41$, $\rho=0,4$) (SCHØLER et collab., 2017).

Dans le cancer du sein, Dawson et collab. ont comparé les performances de l'antigène tumoral 15-3 (CA 15-3), des CTC et des ADNtc pour le suivi des cancers du sein (DAWSON et collab., 2013). Cinquante-deux femmes ont été recrutées dans cette étude prospective et mono-centrique et des altérations somatiques (mutations *PIK3CA* et *TP53* ou VS) ont été détectées chez 30 femmes (DAWSON et collab., 2013). Le suivi dans le plasma des altérations somatiques a été fait par ddPCR ou *tagged-amplicon deep sequencing* (TAMseq) et elles ont été détectées chez 29 femmes (97%) et dans 115 des 141 plasmas (82%) (DAWSON et collab., 2013). Une augmentation de la sensibilité est notée avec l'ADNtc comparé au CA 15-3 (85% vs. 59%) et aussi comparé aux CTC (90% vs. 67%) (DAWSON et collab., 2013). Au niveau dynamique, les ADNtc présentent aussi des caractéristiques supérieures au CA 15-3 et CTC : l'étendue des mesures et la corrélation sont meilleures (DAWSON et collab., 2013) (**Figure 11a**). Les ADNtc représentent aussi un biomarqueur pronostic : les patientes avec les taux plus élevés avaient le pronostic le plus mauvais (DAWSON et collab., 2013) (**Figure 11b**).

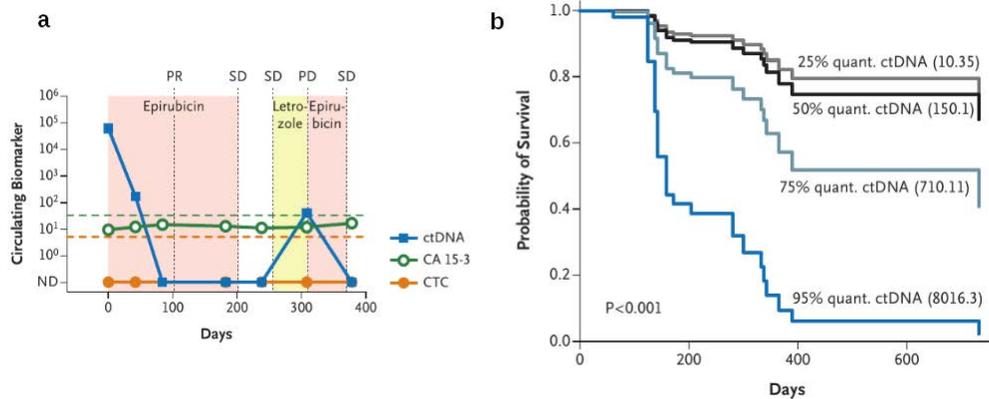


FIGURE 11 – **a**. Le monitoring par les ADNtc présente plusieurs avantages. D’abord l’étendue des mesures est plus grande que celles du CA 15-3 et des CTC. De plus, la corrélation avec la masse tumorale est meilleure. **b**. Cette image représente une régression de Cox qui montre une relation inverse entre les quantiles supérieurs de ADNtc et la survie globale. **PD** : *progressive disease* récurrence, **PR** : *partial response* réponse partielle, **SD** : *stable disease* maladie stable. D’après DAWSON et collab. (2013).

Néanmoins, cette méthode présente plusieurs inconvénients (LIPPMAN et OSBORNE, 2013). En effet, des altérations génomiques somatiques ont été détectées chez 60% des patientes (DAWSON et collab., 2013; LIPPMAN et OSBORNE, 2013). Pour pallier cela, des approches avec séquençage du génome de la tumeur et optimisation des amorces ont été mises au point (GARCIA-MURILLAS et collab., 2015; OLSSON et collab., 2015). Dans une cohorte de 55 patientes, le séquençage initial de la tumeur a permis de détecter au moins une mutation somatique chez 78% des patientes et des amorces spécifiques pour la Polymerase Chain Reaction digitale (PCRd) (GARCIA-MURILLAS et collab., 2015). Ces dernières sont des patients qui ont reçu une chimiothérapie néo-adjuvante ainsi qu’une exérèse complète (GARCIA-MURILLAS et collab., 2015). Les auteurs ont montré que le *monitoring* par les ADNtc en situation post-chirurgicale prédit la récurrence (GARCIA-MURILLAS et collab., 2015). En effet, en point unique, deux à quatre semaines après la chirurgie, la détection de l’ADNtc est associé à la récurrence (HR 25,1 IC95% 4,08–130,5) (GARCIA-MURILLAS et collab., 2015). Le suivi multi-points, ou *mutational tracking* permet d’augmenter la sensibilité de la détection : 50% (6/12) des patientes avec une récurrence ont été détectées en point unique *versus* 80% (12/15) détectées en multi-points (GARCIA-MURILLAS

et collab., 2015). La récurrence de la maladie est détectée 7,9 mois plus tôt par les ADNtc par rapport aux techniques conventionnelles (GARCIA-MURILLAS et collab., 2015). Une cohorte de 20 patientes a été suivie en situation post-chirurgicale par cette méthode : la présence ou l'absence d'une récurrence a été détectée à 93% et 100% (OLSSON et collab., 2015). La détection de cette dernière par les ADNtc a été faite avant la clinique chez 86% des patientes, avec un gain moyen de 11 mois (OLSSON et collab., 2015).

Dans le cancer du poumon, Newman et collab. ont démontré l'intérêt d'un kit de capture dans le suivi de la pathologie (NEWMAN et collab., 2014b). Leur méthode, appelée *cancer personalized profiling by deep sequencing* (CAPP-Seq), permet de détecter au moins une mutation somatique dans 95% des cancers (NEWMAN et collab., 2014b). La diminution de la taille couverte permet de séquencer avec une meilleure couverture que celui d'un exome complet et pour un moindre coût (NEWMAN et collab., 2014b). Dans le cas du suivi des cancers pulmonaires, le niveau des ADNtc dans le plasma avant traitement est corrélé avec le volume tumoral mesuré par tomographie par émission de positons (TEP) ($R^2 = 0,89$, $p = 0,0002$) (NEWMAN et collab., 2014b). Un des intérêts des ADNtc est le suivi en complément de l'imagerie (WAN et collab., 2017). Dans le cas des patients avec CPNPC de stade II ou III qui ont reçu une radiothérapie, le suivi par scanner ou TEP s'avère compliqué : du tissu cicatriciel inflammatoire ou fibrotique peut apparaître et être confondu avec des lésions tumorales (NEWMAN et collab., 2014b). Dans la cohorte suivie par CAPP-Seq, Newman et collab. rapportent le cas d'un patient dont le volume tumoral estimé par l'imagerie demeure important après radiothérapie mais dont l'ADNtc demeure indétectable dans le plasma : 22 mois plus tard ce patient reste toujours en rémission (NEWMAN et collab., 2014b) (**Figure 12a**). Un autre patient présente le cas inverse : au niveau de l'imagerie, une régression du volume tumoral est rapportée mais dans le plasma une augmentation des niveaux d'ADNtc sont décrits (NEWMAN et collab., 2014b). Cette progression suggère des lésions microscopiques et en effet le patient a eu une récurrence sept mois plus tard et décède peu de temps après (NEWMAN et collab., 2014b) (**Figure 12b**).

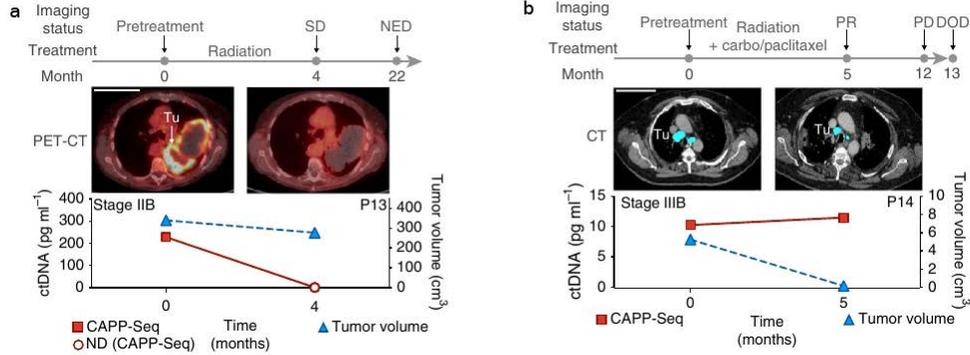


FIGURE 12 – **a.** Suivi d'un patient post-radiothérapie : le volume tumoral est rapporté comme important tandis que les ADNtc demeurent indétectables. Le tissu devait être inflammatoire, car le patient est resté en rémission 22 mois plus tard. **b.** Suivi d'un patient avec une diminution du volume tumoral mais augmentation des taux des ADNtc. Ces derniers devaient refléter des lésions occultes : le patient présente une récurrence sept mois plus tard et décède peu de temps après. **carbo** : carboplatine, **ctDNA** : ADNtc, **DOD** : *dead of disease* jour de décès du patient, **PD** : *progressive disease* récurrence, **PR** : *partial response* réponse partielle, **SD** : *stable disease* maladie stable. D'après NEWMAN et collab. (2014b).

Les mélanomes multiples sont les cancers cutanés les plus agressifs (CALAPRE et collab., 2017). L'inhibition par des molécules anti-*BRAF*, chez les patients *BRAF* mutés, permet une réduction importante de la masse tumorale (CALAPRE et collab., 2017). Une étude sur 48 patients a mis en évidence que les taux élevés d'ADNtc sont associés à une moins bonne réponse aux inhibiteurs de *BRAF* (GRAY et collab., 2015). Ces résultats ont été confirmés par l'analyse d'une cohorte importante (732 sujets) (SANTIAGO-WALKER et collab., 2015). Une survie plus longue a été rapportée chez les patients avec un statut *BRAF* sauvage dans le plasma (SANTIAGO-WALKER et collab., 2015). Pour le suivi de la masse tumorale, les ADNtc présentent aussi plusieurs avantages. Gray et collab. ont montré que l'augmentation des ADNtc en cours de traitement est corrélée à une progression de la pathologie et ce, avant la détection par l'imagerie (GRAY et collab., 2015). Une étude longitudinale de six patients atteints de mélanome au stade IV a montré que le suivi par ADNtc est plus corrélé avec le statut et l'état que la LDH (TSAO et collab., 2015). L'analyse d'une cohorte de 52 patients a montré que les patients avec une diminution précoce du taux de l'ADNtc ont une survie sans progression plus longue (WONG et collab., 2017).

Les ADNtc peuvent aussi être utilisés dans le suivi des maladies oncohéma-

tologiques. Dans le cas des lymphomes, Hohaus et collab. ont montré que les patients avec un taux plus élevé étaient à risque de récurrence, que ce soit pour les lymphomes de Hodgkin ($p = 0,01$) ou les lymphome B diffus à grandes cellules (LBDGC) ($p = 0,03$) (HOHAUS et collab., 2009). Le LBDGC est une maladie curable, mais néanmoins les rechutes sont de mauvais pronostic (ROSCHEWSKI et collab., 2015). L'évaluation des rechutes par les techniques d'imagerie sont limitées par leur sensibilité et l'exposition aux radiations (ROSCHEWSKI et collab., 2015). Durant la lymphoïèse des lymphocytes B, un réarrangement unique des chaînes lourdes *VDJ* intervient et donne lieu à un clonotype (ROSCHEWSKI et collab., 2016). Le suivi dans le plasma de ce réarrangement permet de suivre la masse tumorale et de ce fait, la maladie résiduelle (ROSCHEWSKI et collab., 2016). Dans une étude rétrospective de 126 patients, Roschewski et collab. ont montré que la détection de l'ADNtc après deux cycles de chimiothérapie prédit la récurrence du LBDGC (ROSCHEWSKI et collab., 2015). La valeur prédictive positive (VPP) d'une récurrence est de 88,2% et le gain médian par rapport à la détection avec l'imagerie de 3,5 mois (ROSCHEWSKI et collab., 2015). Le CAPP-Seq a aussi été utilisé dans les LBDGC (SCHERER et collab., 2016). Dans cette étude menée par Alizadeh, les auteurs rapportent que des taux élevés d'ADNtc sont corrélés à une survie sans progression plus faible (SCHERER et collab., 2016). De plus, dans le suivi de la maladie résiduelle, de l'ADNtc a été détecté chez 100% des patients au moment de la rechute radiologique et chez 73% avant, avec un gain moyen de six mois avec une spécificité de 100% (SCHERER et collab., 2016). Certaines pathologies hématologiques possèdent une phase circulante et une phase solide (par exemple ganglionnaire) : c'est le cas de la leucémie lymphoïde chronique (LLC). Dans cette pathologie, le suivi par les ADNtc est corrélée à l'évolution de la lymphoadénopathie mise en évidence par l'imagerie (YEH et collab., 2017b). De plus, la LLC peut conduire au syndrome de Richter, un LBDGC secondaire très agressif avec une médiane de survie de 12 mois (YEH et collab., 2017b). Le diagnostic de ce syndrome s'effectue avec une ponction ganglionnaire, effectuée lorsque l'état du patient se dégrade. Ainsi, le diagnostic de cette complication est tardif et le suivi par ADNtc peut s'avérer intéressant. En effet, des VNC récurrentes sont décrites dans le syndrome de Richter (YEH et collab., 2017b). Le monitoring de ces VNC dans le plasma pourrait représenter un marqueur (YEH et collab., 2017b). Les syndromes myélodysplasiques (SMD) sont des pathologies hématologiques avec un risque d'évolution vers la leucémie aiguë. Le suivi de ces affections s'effectue par des frottis sanguins et par des myélogrammes, prélèvements pour lesquels il existe une grande variabilité inter-opérateurs (YEH et collab., 2017a). Les ADNtc représentent une alternative très peu invasive pour le suivi de ces pathologies avec une dynamique équivalente entre les ADNtc et l'ADN extrait de moelle osseuse (YEH et collab.,

2017a). Les VS sont aussi des marqueurs pronostics dans les SMD (YEH et collab., 2017a). Dans le suivi d'un patient, Yeh et collab. ont montré que les VS détectés dans les cellules de la moelle peuvent de même être mis en évidence dans le plasma (YEH et collab., 2017a). Ce sujet a eu une réponse prolongée au traitement par azacitidine mais a eu une progression du nombre de blastes au sein de la moelle (YEH et collab., 2017a). Le suivi par ADNtc a montré une disparition de sa VNC (del9q) durant la réponse et une réémergence de celle-ci durant la progression, et ce trois mois avant d'être détectée dans la moelle (YEH et collab., 2017a). La méthode CAPP-Seq a aussi été mise au point pour les lymphomes (SCHERER et collab., 2015; KURTZ et collab., 2015). Neuf patients sur 10 (90%) d'une petite cohorte de sujets ayant été déclarés en rémission complète mais ayant eu une récurrence ont été détectés par cette approche (SCHERER et collab., 2015). Il est intéressant de noter que les trois récurrences qui ont touché le SNC ont été détecté dans le plasma (SCHERER et collab., 2015). De plus, le temps médian entre la détection par les ADNtc et la clinique est de 162 jours (SCHERER et collab., 2015). Enfin, une méthode mathématique a été mise au point afin de prédire le volume tumoral et la dynamique (KURTZ et collab., 2015). À partir des mesures des taux d'ADNtc des deux premiers cycles de chimiothérapie, ce modèle a pronostiqué de manière adéquate l'évolution clinique des patients, même ceux déclarés en rémission (KURTZ et collab., 2015).

Les ADNtc peuvent représenter des biomarqueurs performants pour le suivi des pathologies tumorales. Néanmoins, les études publiées sont principalement de faisabilité et avec des petites cohortes. De plus, les recommandations actuelles sont de partir d'une altération somatique trouvée dans les lésions tumorales. Mais cette approche est chronophage, de plus elle est possiblement à risque de faux-négatif si le clone disparaît complètement. D'autres approches, comme le CAPP-Seq permettrait de se passer de l'analyse de la tumeur, mais néanmoins leur coût est supérieur aux méthodes comme la ddPCR. Enfin, l'analyse bioinformatique des méthodes se basant sur les séquençage haut débit est complexe, avec un risque non négligeable de faux-positifs et faux-négatifs. En situation post-traitement (chiomiothérapie ou chirurgie), les études montrent des performances meilleures que les biomarqueurs conventionnels. Néanmoins, certains patients avec des ADNtc négatifs rechutent. Nitzan Rosenfeld a émis l'hypothèse que l'augmentation de la sensibilité des analyses permettrait d'identifier de façon plus certaine les patients en rémission complète (**Figure 13**).

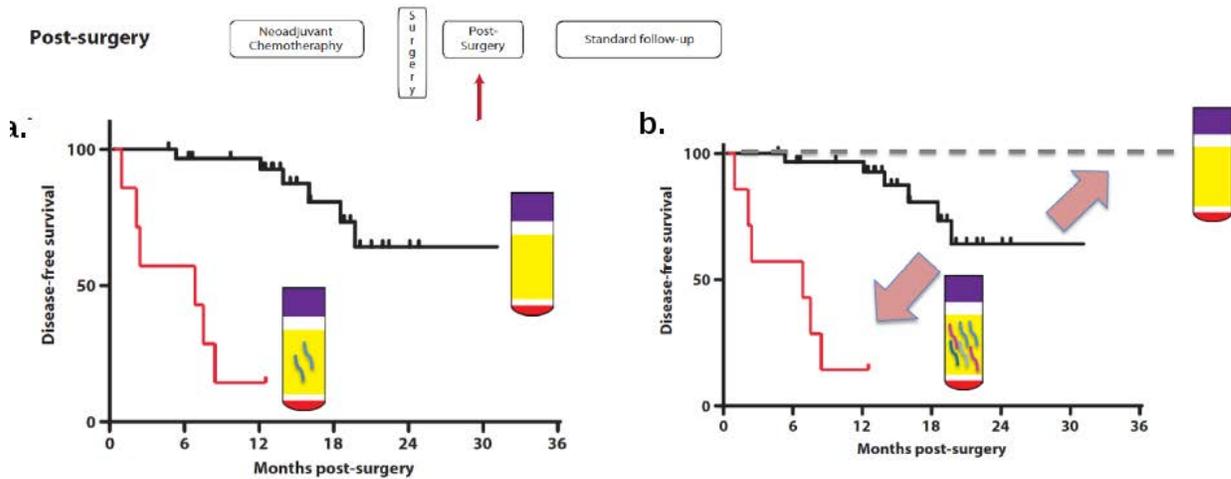


FIGURE 13 – **a.** Les analyses des ADNtc actuelles, même si elles présentent de meilleures performances que les biomarqueurs conventionnels, n’identifient pas correctement tous les patients à risque. **b.** L’augmentation de la sensibilité des analyses permettrait de détecter des lésions de plus petites tailles, facilitant l’identification correcte des patients en rémission complète. D’après une présentation par Nitzan Rosenfeld et image d’après GARCIA-MURILLAS et collab. (2015).

1.3.4 ADN tumoraux circulant et médecine personnalisée : classification des cancers, théranostique, hétérogénéité tumorale et évolution

Les chimiothérapies employées dans les cancers possèdent une forte toxicité (BAUDINO, 2015). Les molécules inhibent la synthèse des bases nucléotidiques à partir des folates (méthotrexate), s’intercalent entre les bases d’ADN lors de la répllication (5-fluorouracile (5-FU), mercaptopurine) ou encore ciblent les topoisomérases (irinotecan, topotecan) (BAUDINO, 2015) (**Figure 14**). Ainsi ces agents interviennent lors de la division cellulaire et agissent contre les cellules qui se divisent rapidement (BAUDINO, 2015). Si les cellules cancéreuses sont à division rapide, des cellules saines (comme celles de l’épithélium intestinal) le sont aussi (BAUDINO, 2015). Ainsi, les chimiothérapies conventionnelles sont très efficaces mais leur toxicité peut s’avérer limitante (BAUDINO, 2015).

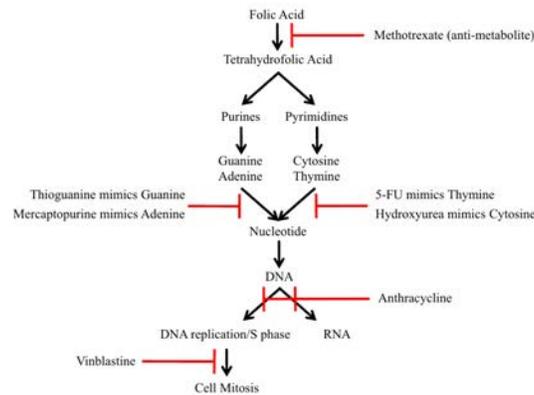


FIGURE 14 – Les chimiothérapies s’appuient sur des agents qui interviennent sur le cycle cellulaire. Les anti-métabolites, comme le méthotrexate, inhibent la synthèse d’acide tétrahydrofolique, ce qui empêche la synthèse de base nucléotidiques. Les analogues de ces dernières bloquent la réplication de l’ADN : par exemple le 5-FU bloque la thymidilate synthase. Les anthracyclines ont un mécanisme d’action triple : ils forment des radicaux libres, ils s’intercalent et peuvent inhiber la topoisomérase de type II. Enfin des agents agissent sur la phase M du cycle cellulaire : la vinblastine empêche la formation des microtubules . D’après BAUDINO (2015).

Des molécules ont été mises au point afin d’augmenter la spécificité anti-tumorale ce qui favorise l’efficacité sans accroissement de la toxicité (SAWYERS, 2004). Ces agents agissent sur une cible spécifique qui est en général une protéine (SAWYERS, 2004). On parle alors de thérapies ciblées, qui n’ont pu être mises au point que par une connaissance précise des voies de signalisation impliquées dans la genèse tumorale (SAWYERS, 2004). La première molécule fut le Gleevec®(Imatinib), employé dans la leucémie myéloïde chronique (LMC) et les cancers gastro-intestinaux (SAWYERS, 2004). Cet agent inhibe la protéine issue de la fusion Breakpoint Cluster Region–Abelson (Bcr-Abl) dans les LMC ou l’activation de c-KIT ou *Platelet-derived growth factor receptor* (PDGFR) dans les cancers gastro-intestinaux (SAWYERS, 2004). La recherche de chromosome Philadelphie (t(9 :22)) ou des mutations qui touchent KIT proto-oncogene receptor tyrosine kinase (*KIT*) permet de prédire l’efficacité du traitement par Imatinib (SAWYERS, 2004). On parle alors de tests compagnons, qui prédisent l’efficacité des thérapies ciblées (OLSEN et JØRGENSEN, 2014). L’analyse d’une biopsie de tissu tumoral est le *gold-standard* actuel, mais des problèmes sont inhérents à cette approche. D’abord, la biopsie est un acte invasif qui requière un plateau technique et du personnel compétent, avec un taux de complica-

tions non-négligeable (RAAIJMAKERS et collab., 2002; OVERMAN et collab., 2012). En outre, cet examen peut ne pas être applicable (notamment pour les tumeurs irradiées) ou la qualité de l'échantillon peut être insuffisante (SCHWARZENBACH et collab., 2011). Habituellement, ces derniers sont fixés à la formaline et inclus en paraffine ce qui dégrade l'ADN (formation de *cross-links*, fragments d'ADN courts de 250 pb...) (TALAULIKAR et collab., 2008). De plus, la biopsie ne produit que l'image à un temps donné d'une seule lésion tumorale. Cela ne permet pas d'apprécier l'hétérogénéité inter-métastases (SHAH et collab., 2009; CAMPBELL et collab., 2010). Enfin, il n'est pas éthique d'effectuer plusieurs biopsies, c'est pourquoi cette approche ne permet pas d'analyser l'évolution tumorale qui est affectée par divers paramètres, dont les traitements (GREAVES et MALEY, 2012). Les ADNtc possèdent plusieurs avantages (DIAZ JR et BARDELLI, 2014). En effet, les prélèvements sont peu invasifs, ce qui permet d'effectuer des séries (DIAZ JR et BARDELLI, 2014). Même si les ADNtc sont composés de fragments de petites tailles, ils représentent une source d'ADN non dégradé (THIERRY et collab., 2016). Finalement, les ADNtc sont issus de toutes les lésions tumorales : ils permettent d'évaluer l'hétérogénéité tumorale et l'évolution des lésions (WAN et collab., 2017).

Les premières méthodes manquaient de sensibilité, ainsi une faible concordance entre les résultats de la biopsie et du plasma était rapportée (le **Tableau 1** en rapporte quelques exemples dans le cancer du colon). L'évolution des techniques et l'apparition de nouvelles méthodes a permis d'augmenter la concordance entre le tissu biopsié et l'analyse en ADNtc. Une des méthodes, nommée IntPlex, a été évaluée par une étude multicentrique, prospective et en double aveugle (THIERRY et collab., 2014). Cent-six patients avec un cancer colorectal métastatique ont été recrutés au sein de ce travail, sept mutations *KRAS* et une touchant *BRAF* ont été recherchées dans la biopsie et dans le plasma (THIERRY et collab., 2014). Pour la mutation *BRAF* V600E, une spécificité et sensibilité de 100% sont rapportées, tandis que pour les mutations *KRAS*, la spécificité est de 98% et la sensibilité de 92% (THIERRY et collab., 2014). La concordance entre la biopsie et le plasma est de 96% et de l'ADNtc a été détecté chez tous les patients (THIERRY et collab., 2014). La biopsie liquide peut aussi être employée dans les atteintes oncohématologiques (YEH et collab., 2017a). En effet, les aspirations de moelles osseuses peuvent être diluées par du sang et une leucopénie peut accompagner les SMD (YEH et collab., 2017a). Une excellente concordance entre les résultats de la moelle osseuse et en ADNtc a été retrouvée (YEH et collab., 2017a).

Des résistances aux thérapies ciblées apparaissent en cours de traitement (GILLIES et collab., 2012). Dans le cas des cancers colorectaux, les sous-clones résistants sont

sélectionnés par les thérapies anti-EGFR (DIAZ JR et collab., 2012; MISALE et collab., 2012). Des études longitudinales ont montré que les mutants *KRAS* émergeaient avec les thérapies ciblées et diminuaient avec l'arrêt de celle-ci (SIRAVEGNA et collab., 2015; MORELLI et collab., 2015).

TABLE 1 – Biopsie liquide : évolution des techniques, exemple du cancer colorectal

Premier auteur Journal Année	Design	Ma- trice	Méthode d'analyse pour les ADNtc	Patients Contrôles	Gene(s)	Patients avec mu- tation(s) détec- tée(s) dans la biopsie	Patients avec mu- tations(s) détectée en ADNtc	Concor- dance
Ryan, Gut, 2003	Prospect- tive	Sérum	PCR semi- nichée et séquençage Sanger	78 pré- chirurgie/94 en suivi 20	<i>KRAS</i>	32 patients en pré- chirurgie (41%)	41 (53%) préchirurgie 60 (64%) suivi	76%
Diehl, PNAS, 2005	Non décrit	Plasma	BEAMing	56 Pas de contrôle	<i>APC</i>	33/56 (56%)	17/33 (52%)	NR
Trevisiol, Int J Biol Markers, 2006	Prospect- tive et consécu- tive	Sérum	Me-PCR	86 Pas de contrôle	<i>KRAS</i>	28 (3%)	11 (13%)	67%
Taly, Clin Chem, 2013	En cours de chimio- thérapie	Plasma	RainDance	50 10	<i>KRAS</i> <i>BRAF</i>	24/50 (48%)	14/50 (28%) 5/50 (10%)	Multiplex 74%, Duplex 84% Duplex 82%
Bettgowda, Sci Transl Med, 2014	Pré- thérapie	Plasma	dPCR	206 Pas de contrôle	<i>KRAS</i>	79 (38%)	69 (33%)	95%
Thierry, Nat Med, 2014	Prospect- tive, multicen- trique, en aveugle	Plasma	Intplex	95 29	<i>KRAS</i> <i>BRAF</i>	39 (38%)		96% 100%

Dans les cancers du poumon, les autorisations de mise sur le marché (AMM) des inhibiteurs d'EGFR comme l'Erlotinib disposent que les ADNtc peuvent être utilisés si un échantillon de la tumeur n'est pas analysable. Ces molécules provoquent la sélection de clones résistants, porteurs par exemple de la mutation *EGFR* T790M (SORENSEN et collab., 2014). Ces mutations peuvent être suivies dans le plasma des patients, ce qui permettrait de mettre en place des inhibiteurs de nouvelle génération comme l'Osimertinib de façon plus précoce (SORENSEN et collab., 2014; WAN et collab., 2017) (**Figure 15**).

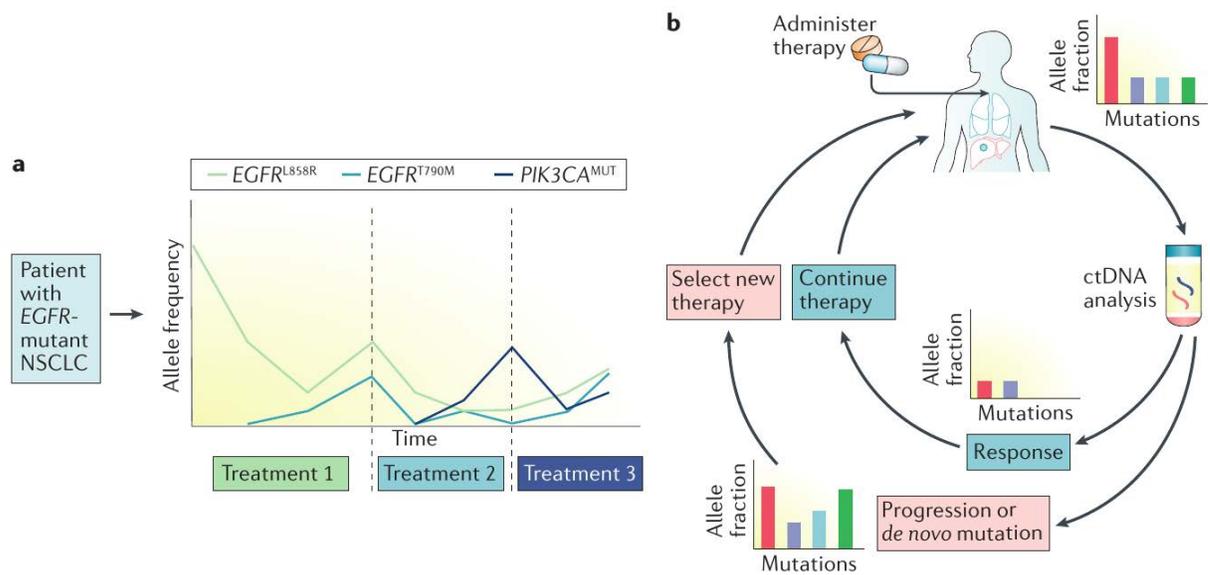


FIGURE 15 – **a.** Les ADNtc peuvent s'avérer utiles dans de nombreuses situations pour la théranostique. D'abord, étant le reflet de toutes les tumeurs, ils peuvent être utilisés comme test compagnon. Ensuite les ADNtc sont issus de prélèvements peu invasifs, ce qui permet de suivre l'évolution de la résistance avec l'émergence de clones porteurs de mutations. **b.** Cette figure schématise l'utilisation des ADNtc. Les analyses se doivent d'être rapides, afin que les cliniciens adaptent les thérapies en fonction des réponses obtenues. Une réponse conduit au maintien de la thérapie et à son suivi comme décrit dans la figure **a**. Dans le cas contraire, une nouvelle thérapie est initiée. D'après WAN et collab. (2017).

Les mélanomes sont aussi traités par des inhibiteurs de BRAF, dont le test compagnon est la présence de mutations sur l'exon 15 qui touchent le codon 600 (GRAY et collab., 2015; SCHADENDORF et collab., 2015). L'évolution de cette pathologie peut être très rapide, notamment dans les formes métastatiques (SCHADENDORF et collab., 2015). Les ADNtc, pouvant être prélevés par un grand nombre de personnes, peuvent augmenter la rapidité de l'analyse, notamment pour mettre en place les traitements. Les ADNtc possèdent aussi l'avantage d'être une source d'ADN frais. Enfin, une résistance aux inhibiteurs de BRAF peut apparaître durant le traitement (GRAY et collab., 2015). Les mutations NRAS proto-oncogène, GTPase (*NRAS*) peuvent entraîner une résistance aux inhibiteurs de BRAF. Dans une petite cohorte, ces mutations ont été retrouvées dans le plasma de 3 patients sur 7 (GRAY et collab., 2015). Ces altérations ont été retrouvées avant l'imagerie, ainsi les ADNtc peuvent représenter des marqueurs de choix pour le suivi des thérapies dans le mélanome et l'adaptation de ces dernières (GRAY et collab., 2015).

Ces études ont analysé un seul gène pour évaluer l'évolution de la résistance aux thérapies mais plusieurs gènes peuvent être impliqués. Bettgowda et collab. ont analysé le plasma de patients avec un cancer colorectal métastatique traités par panitumumab ou cetuximab avec un panel de gènes dans la voie de signalisation de l'EGFR (BETTEGOWDA et collab., 2014). Des mutations non présentes avant traitement ont été mises en évidence par cette méthode chez 23 des 24 patients initialement répondeurs (BETTEGOWDA et collab., 2014). Il a été retrouvé des mutations de l'exon 2 de *KRAS* mais aussi sur le codon 600 de *BRAF* ou encore l'*EGFR* (BETTEGOWDA et collab., 2014). En moyenne 2,9 mutations ont été retrouvées en post-traitement (de 0 à 12 altérations) (**Figure 16**) : les auteurs avancent l'hétérogénéité inter-métastases et la réponse des différentes lésions tumorales aux molécules (BETTEGOWDA et collab., 2014). Ainsi, le suivi par une seule mutation et non un panel de gènes peut s'avérer limitant.

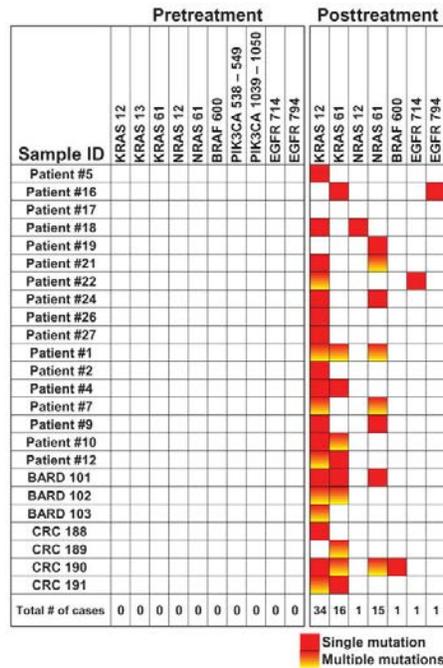


FIGURE 16 – Le plasma de 24 patients traités par anti-EGFR a été analysé par un panel de gènes. Avant la thérapie, aucune mutation n’a été mise en évidence par l’analyse des ADNtc. En post-traitement, des mutations ont été mises en évidence chez 23 patients (96%). En moyenne 2,9 mutations ont été détectées : cela reflète l’hétérogénéité inter-métastase. D’après BETTEGOWDA et collab. (2014).

L’exome a été utilisé pour le suivi des mécanismes de résistance aux différents traitements (MURTAZA et collab., 2013; WAN et collab., 2017). Six patients avec des cancers avancés (ovaires, sein et poumon) ont été suivis avec des séries de prélèvements durant un à deux ans (MURTAZA et collab., 2013). L’analyse simultanée d’une biopsie et des ADNtc a mis en évidence que les altérations retrouvées dans la première se reflètent dans les seconds (MURTAZA et collab., 2013). De plus, les fractions alléliques des mutations augmentent dans le plasma lors d’une résistance à un traitement (MURTAZA et collab., 2013) Par exemple, chez une patiente avec un cancer du sein traitée par paclitaxel, les auteurs rapportent une augmentation de la fraction allélique de plusieurs mutations (*PIK3CA*, *BMI1* et *SMC4*) (MURTAZA et collab., 2013). Les gènes touchés appartiennent à plusieurs voies de signalisation ce qui complique la mise au point d’un panel de gènes réduits (MURTAZA et collab., 2013). Cette étude montre aussi que les ADNtc pourraient aussi être des biomarqueurs pour le

monitoring des chimiothérapies conventionnelles (MURTAZA et collab., 2013). Mais l'exome est une approche coûteuse, qui ne peut être utilisée que pour les fortes fractions alléliques ($> 5\%$) (WAN et collab., 2017). Des méthodes basées sur des panels de gènes ont été mises au point afin de diminuer les coûts et augmenter la sensibilité (WAN et collab., 2017). Elles peuvent faire appel à soit des amplicons (*tagged-amplicon deep sequencing* (Tam-Seq) soit à de la capture CAPP-Seq (WAN et collab., 2017). Néanmoins, ces approches étant basées sur des altérations déjà décrites, aucune *de novo* n'est détectée (WAN et collab., 2017). Le séquençage à faible couverture du génome entier en ADNtc permet de mettre en évidence les VNC (WAN et collab., 2017). Ces dernières sont plus spécifiques des lésions tumorales, ainsi elles sont moins sensibles au bruit de fond que les mutations en point unique (WAN et collab., 2017). Dans les cancers de la prostate métastatiques, des amplifications focales d'oncogènes (MYC proto-oncogene (*MYC*)) ou des délétions (phosphatase and tensin homolog (*PTEN*)) ont été retrouvées (ULZ et collab., 2016). Un délai moyen de 26,4 semaines entre deux amplifications a été retrouvé lors de l'analyse de séries de plasma, ce qui suggère une adaptation rapide suite aux pressions de sélection (ULZ et collab., 2016).

L'hétérogénéité tumorale découle de l'évolution clonale des cancers. Celle-ci existe non seulement au sein d'une même lésion mais aussi entre les métastases (GERLINGER et collab., 2012; YANCOVITZ et collab., 2012). Une analyse pan-cancer a montré que l'hétérogénéité intra-tumorale est un marqueur pronostic pour la survie (MORRIS et collab., 2016). Cet aspect est difficile à mettre en évidence avec les biopsies : il faudrait effectuer l'analyse de plusieurs sites au sein de la même lésion ou prélever plusieurs métastases, ce qui n'est pas éthique (DIAZ JR et BARDELLI, 2014). Toutes les tumeurs peuvent donner des ADNtc, ainsi l'hypothèse que les ADNtc reflètent tous les tissus a été émise. Une étude menée par Rosenfeld a suivi une patiente avec un cancer du sein métastatique sur 1193 jours (MURTAZA et collab., 2015). Les auteurs ont évalué l'architecture génomique au sein de huit biopsies et neuf plasmas ont été aussi évalués (MURTAZA et collab., 2015). Les mutations partagées par toutes les métastases (*stem mutations*), les mutations des clades entre les métastases et même les mutations privées (qui n'appartiennent qu'à une lésion) ont été détectées dans les ADNtc (MURTAZA et collab., 2015). Même si ces résultats doivent être validés avec une cohorte plus large, Murtaza et collab. montrent une concordance entre les tumeurs et les ADNtc dans le contexte de la phylogénie tumorale (MURTAZA et collab., 2015).

L'analyse du transcriptome a permis de mettre en évidence plusieurs signatures de gènes ayant un intérêt clinique. C'est le cas de MammaPrint[®], signature de 70 gènes liés au risque de récurrence dans le cancer du sein (VAN'T VEER et collab.,

2002). Les LBDGC se caractérisent par une grande hétérogénéité dans la réponse aux thérapies (SCHERER et collab., 2016). Des classifications pour cette pathologie ont été décrites, qu’elles soient cliniques, radiologiques ou basées sur la transcriptomique. L’analyse de cette dernière a permis d’isoler deux signatures qui dépendent de la cellule d’origine du lymphome (*germinal center B cell-like* [GCB] ou *activated B cell-like* [ABC]) (ALIZADEH et collab., 2000). Ces signatures sont d’intérêt pronostic mais aussi liées à la réponse aux thérapies (THIEBLEMONT et collab., 2011). Mais l’analyse de l’expression des gènes est difficile à mettre en place en pratique, car elle nécessite des tissus congelés mais aussi du personnel formé. Une classification par l’histologie, l’algorithme de Hans, est utilisée dans la pratique courante, mais elle est peu reproductible (variabilité inter-opérateurs, dépendance de la qualité des tissus) (SCOTT et collab., 2014). Scherer et collab. proposent une classification à l’aide de l’approche CAPP-Seq pour les LBDGC (SCHERER et collab., 2016). Un premier groupe de patients, dont la tumeur a été analysée par transcriptomique, a permis de mettre au point une classification par CAPP-Seq (SCHERER et collab., 2016). La concordance effectuée sur 76 patients avec la classification CAPP-Seq et celle de Hans est de 80 % (SCHERER et collab., 2016) **Figure 17 a.**). De plus, cette dernière échoue à stratifier de façon adéquate les patients, au contraire de l’approche par ADNtc (SCHERER et collab., 2016) **Figure 17 b-d.**). Enfin, cette dernière présente une concordance élevée que le CAPP-Seq soit effectué sur le tissu tumoral ou le plasma (SCHERER et collab., 2016) **Figure 17 e.**). Les auteurs montrent aussi l’intérêt des ADNtc dans l’évaluation de l’hétérogénéité tumorale (SCHERER et collab., 2016). Les lymphomes folliculaires sont généralement indolents mais peuvent évoluer en LBDGC, qui sont très agressifs (SCHERER et collab., 2016). En utilisant la méthode CAPP-Seq, Scherer et collab. ont pu différencier les types histologiques des lymphomes ainsi que leur évolution (SCHERER et collab., 2016). Une VPP de 91% pour les lymphomes folliculaires transformés *versus* non-transformés a été rapportée (sensibilité 83%, spécificité 89%), et ce en moyenne 66 jours avant le diagnostic clinique (SCHERER et collab., 2016). Chez un patient, les résultats sont discordants entre l’anatomopathologie et la classification CAPP-Seq. La biopsie inguinale rapporte une histologie de lymphome folliculaire tandis que l’analyse des ADNtc classe le patient en lymphome folliculaire transformé (SCHERER et collab., 2016). Les mutations détectées provenaient d’un sous-clone de lymphome folliculaire transformé, qui a été détecté et prélevé neuf mois plus tard après un échec du rituximab (SCHERER et collab., 2016) (**Figure 17 f.**). Les ADNtc, au moment du diagnostic, montrent une similitude avec les lymphomes folliculaires transformés, ce que confirme l’algorithme de classification proposé par les auteurs (SCHERER et collab., 2016). Ces lymphomes sont très peu sensibles au rituximab, ainsi les ADNtc permettraient de différencier

les sous-types histologiques d'intérêt clinique, de guider les thérapies et de suivre l'évolution clonale (SCHERER et collab., 2016).

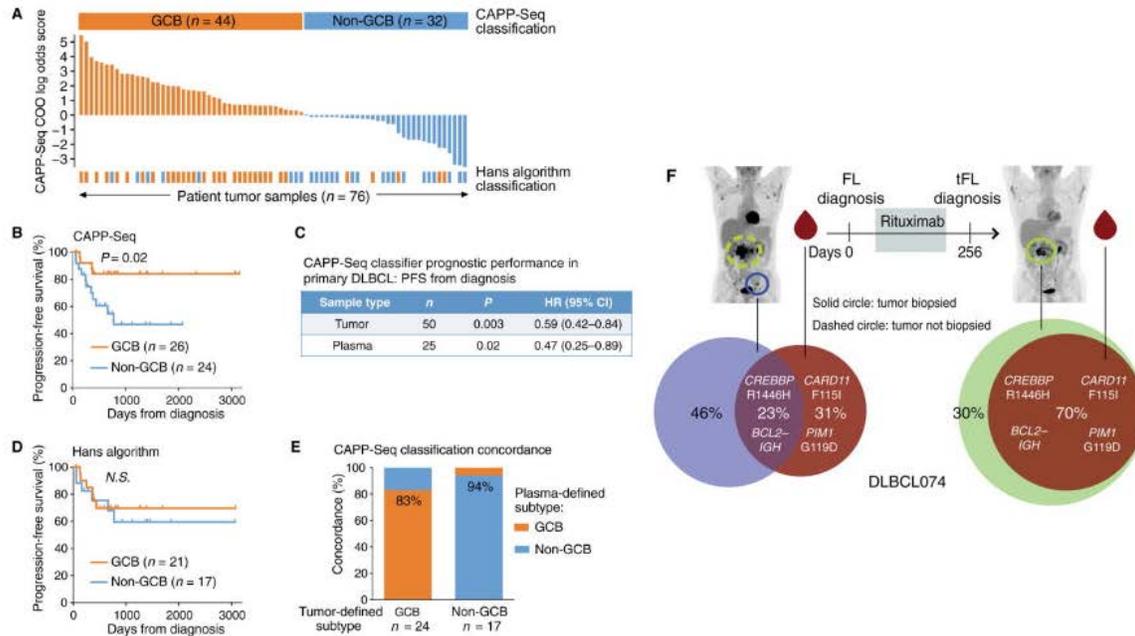


FIGURE 17 – **a.** Soixante-seize patients ont été classifiés par l'approche CAPP-Seq, 59 d'entre eux ont eu aussi une classification par immunohistochimie. La concordance entre les ADNtc et l'algorithme de Hans est de 80%. **b.** Survie sans progression des patients classés par l'approche CAPP-Seq (n=50) (test *log-rank*, $p=0,02$). **c.** Analyse de Cox univariée du HR pour la classification par CAPP-Seq. **d.** Survie sans progression des patients (n=38) classés par la méthode de Hans (test *log-rank* non significatif). **e.** Concordance de la classification CAPP-Seq réalisée sur la tumeur ou le plasma (n=41). **f.** Suivi d'un patient porteur d'un clone de lymphome folliculaire non-transformé (cercle bleu, localisation inguinale) et transformé (cercle vert, localisation rétropéritonéale). Au diagnostic, l'analyse de la lésion inguinale montre une histologie qui se rapproche d'un lymphome indolent. Une hétérogénéité tumorale a été détectée au moment du diagnostic par les ADNtc avec présence de mutations qui appartiennent au lymphome agressif. Après un échec du rituximab, un diagnostic de lymphome folliculaire transformé est posé. L'analyse par CAPP-Seq de la biopsie rétropéritonéale et les ADNtc partagent un grand nombre de mutations. D'après SCHERER et collab. (2016).

1.3.5 Les ADN tumoraux circulant, bientôt en pratique courante ?

Ces différents travaux montrent un intérêt croissant pour les ADNtc en tant que "biopsie liquide" (WAN et collab., 2017). Issus de plusieurs lésions différentes ainsi que provenant de prélèvements peu invasifs, les ADNtc permettent d'évaluer l'hétérogénéité inter-métastases et d'analyser l'évolution clonale (WAN et collab., 2017). Ainsi les ADNtc sont d'un grand intérêt en théranostic (WAN et collab., 2017). Néanmoins, des études de puissance doivent être réalisées avant que la biopsie liquide ne puisse être utilisée en complément ou remplacement des biopsies tissulaires : la plupart des travaux sont soit de faisabilité, soit basés sur des cohortes de petites tailles. Un panel d'experts a rendu une revue systématique de la littérature en 2017 (WONG et collab., 2017). Les auteurs mettent en évidence que peu de méthodes pour l'analyse des ADNtc sont valides et utiles en clinique pour les cancers avancés (MERKER et collab., 2018). Des discordances subsistent entre les ADNtc et les biopsies, ainsi les experts recommandent d'effectuer les tests sur les tissus quand rien n'est détecté en ADNtc (MERKER et collab., 2018). De plus, les auteurs ne rapportent une utilité clinique de la biopsie liquide que pour les cancers colorectaux et les CPNPC (MERKER et collab., 2018). De plus, ils mettent en garde par rapport à la capacité des ADNtc de détecter des sous-clones (MERKER et collab., 2018). En effet, une mutation qui provient d'un sous-clone, peut être détectée par les ADNtc et ne prédira pas de la même façon la réponse au traitement qu'une détection dans le tissu (MERKER et collab., 2018). Enfin, aucune évidence d'utilité n'a été relevée pour les ADNtc dans les premiers stades des cancers, le monitoring des thérapies et le suivi de la maladie résiduelle (MERKER et collab., 2018). Si les experts approuvent la corrélation entre les taux de ADNtc et la masse tumorale dans de nombreux cancers, le manque d'études prospectives et de puissance nuancent l'utilité clinique de ces analyses (MERKER et collab., 2018). De plus, dans le monitoring des thérapies, les résultats des ADNtc ne sont pas binaires : la quantification des changements s'avère complexe et peu d'études décrivent de façon rigoureuse la méthode de suivi (MERKER et collab., 2018). Ce qui n'est pas le cas du suivi de la maladie résiduelle, où les résultats sont binaires (présence *vs.* absence d'ADNtc). Néanmoins, les experts rapportent que les faux-négatifs et les faux-positifs n'ont pas été évalués de manière adéquate à ce jour (MERKER et collab., 2018). De plus, aucune étude ne s'est encore intéressée à la thérapie guidée par les résultats d'ADNtc (MERKER et collab., 2018). Enfin, pour le screening, les auteurs rappellent que la faisabilité des ADNtc pour détecter un cancer chez les patients asymptomatiques n'a pas été évaluée (MERKER et collab., 2018). En outre, il existe un risque de sur-diagnostic par cette approche, tout comme avec la mammographie (MERKER et collab., 2018). Il serait intéressant

d'évaluer les capacités des ADNtc dans des populations à risque plutôt que générale. La requête "liquid biopsy" sur clinicaltrials.gov retrouve 66 essais cliniques en cours et les experts de l'*American Society of Clinical Oncology* (ASCO) concluent sur une évolution rapide de la recherche (MERKER et collab., 2018).

1.4 Méthodes d'analyses des ADN tumoraux circulant

L'analyse des ADNtc est complexe, du processus pré-analytique au post-analytique. La courte demi-vie et la biologie complexe des ADNtc compliquent la standardisation de la phase pré-analytique ainsi que l'extraction (EL MESSAOUDI et collab., 2013). Les ADNtc se retrouvent en très faible quantité dans les différents liquides biologiques, que ce soit en terme pondéral (5-1500 ng/mL) ou en fraction par rapport à l'ADN constitutionnel (0,0003% à 95%) (MOULIERE et collab., 2011, 2014). De plus, même si une forte hétérogénéité est rapportée au niveau des proportions d'ADNtc, seule une fraction des patients présente une proportion supérieure à 25% d'allèles mutées (MOULIERE et collab., 2014; THIERRY et collab., 2014). Les méthodes d'analyse classiques, comme la méthode de Sanger, sont caractérisées par une sensibilité faible, ne détectant les ADNtc que chez les patients avec une forte proportion d'allèles mutées (**Tableau 1**) (DIAZ JR et BARDELLI, 2014). C'est pourquoi de nouvelles méthodes ont été mises au point, afin d'augmenter la sensibilité de détection et de pouvoir suivre un plus grand nombre de patients (**Tableau 1**) (DIAZ JR et BARDELLI, 2014). Enfin, au niveau post-analytique, des analyses bioinformatiques rigoureuses s'avèrent nécessaires, notamment pour les méthodes d'analyse basées sur le séquençage nouvelle génération. Ces dernières sont aussi peu standardisées et en constante évolution.

1.4.1 Phase pré-analytique

De nombreux facteurs influencent la qualité, la fragmentation et la quantité des ADNtc, comme la matrice prélevée, les conditions de conservation ou encore les conditions de prélèvements (EL MESSAOUDI et collab., 2013) Dans les premiers travaux qui portent sur les ADNtc, la phase pré-analytique n'était pas standardisée (EL MESSAOUDI et collab., 2013). Une revue de la littérature sous la direction d'Alain Thierry a permis en 2013 de proposer une standardisation de cette phase pour les ADNtc issus de prélèvements sanguins (EL MESSAOUDI et collab., 2013). Le panel d'experts de l'ASCO a aussi rendu des recommandations en 2018 (MERKER et collab., 2018). Le choix de la matrice (plasma *vs.* sérum) est déterminant : une

plus forte quantité d'ADNlc est présente dans les sérums (EL MESSAOUDI et collab., 2013). Cela résulte de la lyse de certains leucocytes durant la formation du caillot, ce qui provoque le relargage d'ADN (EL MESSAOUDI et collab., 2013). Dans le cadre de l'oncologie, cela s'accompagne par une augmentation du "bruit de fond", avec une diminution des fractions de mutants (EL MESSAOUDI et collab., 2013). Il n'existe aucune donnée concernant les conditions de prélèvement (tube de purge, veines à prélever, etc.), les experts de l'ASCO recommandent de suivre les instructions des fabricants (MERKER et collab., 2018). Les ADNlc sont habituellement prélevés sur des tubes K₃EDTA, mais il existe des tubes spécifiques aux prélèvements d'ADNlc, les tubes Strecktm. Des études ont montré peu de différences en terme de quantité entre les tubes K₃EDTA et Strecktm si les prélèvements sont pris en charge immédiatement (EL MESSAOUDI et collab., 2013). En revanche, si cette dernière est retardée, les taux d'ADNlc diminuent rapidement dans les tubes K₃EDTA, tandis qu'ils restent stables jusqu'à 14 jours avec les tubes Strecktm (EL MESSAOUDI et collab., 2013). Au vu de la demi-vie des ADNlc et de la lyse des leucocytes dans les tubes K₃EDTA, El Messaoudi et collab. recommandent d'isoler le plasma dans les quatre heures qui suivent le prélèvement tandis que Menker et collab. donnent un délai de six heures (EL MESSAOUDI et collab., 2013; MERKER et collab., 2018). L'hémolyse doit être évitée si possible, car elle provoque une augmentation des taux de nucléosomes (EL MESSAOUDI et collab., 2013). Le transport des tubes de sang total doit être effectué sans variations importantes de la température et sans chocs importants (MERKER et collab., 2018). Aucune donnée sur l'usage des pneumatiques n'a été décrite. Si un transport long doit être effectué, il est recommandé d'acheminer du plasma congelé et d'éviter les cycles de congélation-décongélation (MERKER et collab., 2018). L'isolation du plasma a été optimisée par El Messaoudi et collab., avec une double centrifugation : la première à 1200 g pendant 10 minutes suivie d'une centrifugation du plasma à 16000 g pendant 10 minutes (EL MESSAOUDI et collab., 2013). Cette approche permet de garantir l'absence de cellules dans le plasma et de diminuer la contamination de l'ADN issu de cellules saines (EL MESSAOUDI et collab., 2013). Le plasma peut être stocké jusqu'à neuf mois à -20 ° C ou -80 ° C (EL MESSAOUDI et collab., 2013). Une altération de l'intégrité des ADNlc est notée après trois cycles de congélation-décongélation (EL MESSAOUDI et collab., 2013).

1.4.2 Extraction des ADN tumoraux circulant

De nombreux kits commerciaux pour l'extraction des ADNlc existent. Néanmoins peu d'études d'évaluation de ces méthodes ont été conduites (PÉREZ-BARRIOS et collab., 2016). Une nécessité d'harmonisation et de standardisation s'avère aussi

nécessaire. Plusieurs paramètres guident le choix d'un kit d'extraction. Le premier est la quantité d'ADNlc circulant extrait. La concentration d'ADNlc dépend du stade de la maladie (très peu d'ADNlc dans les cancers localisés ou atteignant le SNC) ou si le patient est traité ou non. La quantité d'ADN n'est pas la même selon l'analyse : une PCRd nécessite 10 ng tandis que 50 à 200 ng sont requis pour une analyse en séquençage nouvelle génération, ce qui représente environ 10 mL de plasma. Des pistes comme la plasmaphérèse ou des méthodes avec des anticorps sont proposées afin d'augmenter la quantité d'ADNlc en contre-partie ces approches sont lourdes pour le patient et/ou coûteuses (WAN et collab., 2017). Pérez-Barríos et collab. ont évalué trois kits commerciaux : Le MagNa Pure Compact Nucleic Acid Isolation Kit I, le Maxwell RSC ccfDNA Plasma Kit et le QIAamp Circulating Nucleic Acid Kit (PÉREZ-BARRIOS et collab., 2016). Les concentrations d'ADN extraits ont été mesurées par le Qubit dsDNA HS Assay kit et le profil électrophorétique par le système Agilent 2100. Enfin les auteurs ont évalué les fractions d'allèles mutées (FAM) par PCRd (PÉREZ-BARRIOS et collab., 2016). Au niveau des quantités extraites, des différences statistiquement significatives sont rapportées entre le kit MagNa Pure et Maxwell ($p < 0,00001$) mais pas entre ce dernier et le kit QIAamp (PÉREZ-BARRIOS et collab., 2016). Si un profil oligonucléosomal est retrouvé chez 88% des patients, les profils s'avèrent différents entre les kits MagNa Pure et Maxwell (PÉREZ-BARRIOS et collab., 2016). L'extraction des petits fragments par le kit Maxwell n'est pas aussi performante que celle par le kit MagNa Pure (PÉREZ-BARRIOS et collab., 2016). Hors ces fragments représentent de l'ADN issus de cellules apoptotiques et sont enrichis en ADN qui provient de cellules tumorales (THIERRY et collab., 2016). Enfin, les FAM ne sont pas différentes selon les kits (PÉREZ-BARRIOS et collab., 2016). Néanmoins, les auteurs avancent l'hypothèse que si un kit n'a pas les mêmes capacités d'extractions des fragments d'ADNlc, cela peut affecter la quantification de la masse tumorale par l'analyse des ADNtc (PÉREZ-BARRIOS et collab., 2016). De plus, les auteurs rapportent le cas d'un patient chez qui une mutation somatique n'a pas été détectée après extraction par le kit Maxwell mais bien mise en évidence après extraction avec le kit MagNa Pure (PÉREZ-BARRIOS et collab., 2016).

C'est pourquoi le choix d'un kit d'extraction d'ADNlc n'est pas une étape à négliger. La quantité, mais aussi la capacité à extraire différents fragments, doivent être évaluées (PÉREZ-BARRIOS et collab., 2016). De plus, cette variable doit être prise en compte dans les études multi-centriques ainsi qu'au niveau du suivi longitudinal d'un patient (PÉREZ-BARRIOS et collab., 2016).

1.4.3 Méthodes d'analyses

La distinction entre les ADNtc de l'ADNlc sain peut se faire en se basant sur les mutations somatiques (DIAZ JR et BARDELLI, 2014). Mais les ADNtc représentent une fraction très faible (généralement $< 1\%$) des ADNlc, ainsi les méthodes d'analyses des mutations de première génération (séquençage de Sanger ou pyroséquençage) ne peuvent détecter des altérations que chez des patients avec une masse tumorale importante et des taux de ADNtc importants (DIAZ JR et BARDELLI, 2014). De nouvelles technologies ont permis d'augmenter la sensibilité des analyses et de rechercher la présence d'ADNtc chez un plus grand nombre de patients (**Figure 18**).

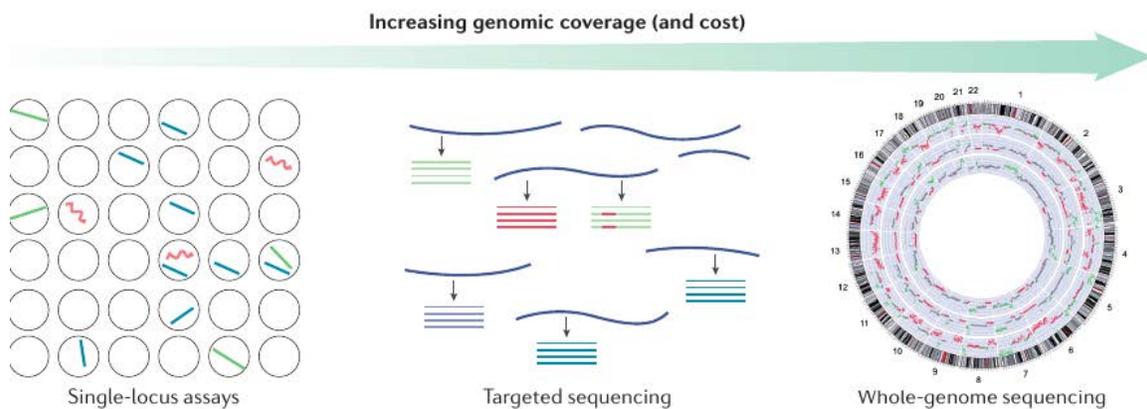


FIGURE 18 – Les méthodes disponibles pour analyser l'ADNtc peuvent se subdiviser en trois catégories. D'abord les analyses point unique, comme la PCRd : elles permettent d'analyser que peu d'altérations, mais avec une grande sensibilité. Ensuite, les analyses ciblées avec l'aide du séquençage haut débit. Ces méthodes peuvent analyser de quelques exons à l'ensemble de ces derniers. Enfin les méthodes génome entier : elles permettent de mettre en évidence des altérations du nombre de copies principalement, mais avec une sensibilité moindre (de l'ordre de 5 à 10%. D'après WAN et collab. (2017).

Il est possible de diviser ces nouvelles méthodes en deux grandes catégories : celles qui analysent des mutations *hotspot* ou les approches avec une couverture plus grande.

1.4.4 Méthodes d'analyses en point unique ou pauci-*loci*

La sensibilité de ces approches est grande avec une procédure de réalisation et d'analyse simple (WAN et collab., 2017). Certaines de ces méthodes ont été mises au point il y a des décennies, comme l'*amplification refractory mutation system* (ARMS) ou la PCRd, ce qui permet d'avoir un recul important (NEWTON et collab., 1989; VOGELSTEIN et KINZLER, 1999). Des altérations de l'épigénome peuvent être analysées par ces approches, comme avec le méthyl-BEAMing (LI et collab., 2009b).

1.4.4.1 Méthodes microfluidiques ou PCR allèle-spécifique

amplification refractory mutation system Cette PCR est allèle-spécifique, des amorces dont la dernière base en 3' est spécifique de l'allèle à amplifier (NEWTON et collab., 1989). Dans le cas d'une non-correspondance de la base 3' avec la molécule d'ADN, l'amplification est semblable au bruit de fond (NEWTON et collab., 1989). Des kits commerciaux basés sur cette approche, approuvés par la FDA ou l'*european medicines agency* (EMA), sont aussi disponibles (WAN et collab., 2017). C'est le cas du kit Cobas ou du kit Therascreen pour la recherche de mutations *EGFR* dans le CPNPC (WAN et collab., 2017). Le premier permet la recherche de 41 mutations et utilise l'exon 28 d'*EGFR* comme contrôle interne (sensibilité de 25–100 copies/mL) tandis que le second couvre 29 variants d'*EGFR* (sensibilité variant de 0,05%–12,47% selon les mutations) (WAN et collab., 2017). L'utilisation de ces kits sur prélèvements plasmatiques est autorisée si l'analyse sur la biopsie est impossible.

PCR digitale Cette méthode se base sur une isolation des allèles par dilution limite (VOGELSTEIN et KINZLER, 1999). Cela permet d'amplifier par PCR individuellement des molécules d'ADN unique. Des sondes fluorescentes sont ensuite ajoutées, ce qui permet une analyse séparée de chacun des produits de PCR (VOGELSTEIN et KINZLER, 1999) (**Figure 19**). Des résultats "digitaux" sont obtenus : dans chaque puits il existe soit des produits de PCR issus d'allèle sauvage ou mutée, mais jamais de mélange des deux (VOGELSTEIN et KINZLER, 1999). La sensibilité de cette méthode a été évaluée à 0,01% (soit 1 allèle mutée pour 10000 sauvages) (DIAZ JR et BARDELLI, 2014).

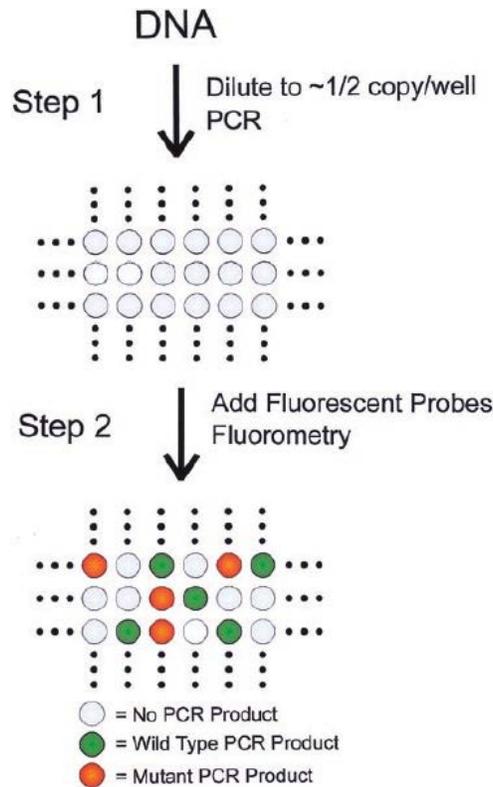


FIGURE 19 – Une PCR est effectuée sur des échantillons d'ADN dilués. Ensuite, des sondes fluorescentes spécifiques des allèles sauvages et mutées sont ajoutées à chaque puit et la fluorescence est lue. D'après VOGELSTEIN et KINZLER (1999).

Plusieurs améliorations ont été portées à cette méthode, dont l'utilisation de pico-gouttes (*picodroplets*) pour "compartmentaliser" l'ADN avec la plateforme RainDance[®] (TALY et collab., 2013). Cette approche permet d'analyser plusieurs mutations (TALY et collab., 2013).

BEAMing BEAMing est un acronyme pour beads, emulsion, amplification, magnetics (DRESSMAN et collab., 2003). Cette méthode dérive des PCRd et combine une PCR dans une émulsion eau-dans-huile et la cytométrie en flux (DIEHL et collab., 2006) (**Figure 20**). D'après Sysmex, qui commercialise ce kit sous le nom d'OncoBEAMtm, l'analyse peut être effectuée en deux jours. La sensibilité est de 0,02 à 0,04%. De plus, cette méthode a été évaluée dans les cancers colorectaux, pulmonaires, mammaires et le mélanome. L'analyse peut être faite sur plusieurs mu-

tations et a même été évaluée pour quantifier la méthylation d'un promoteur (Li et collab., 2009b).

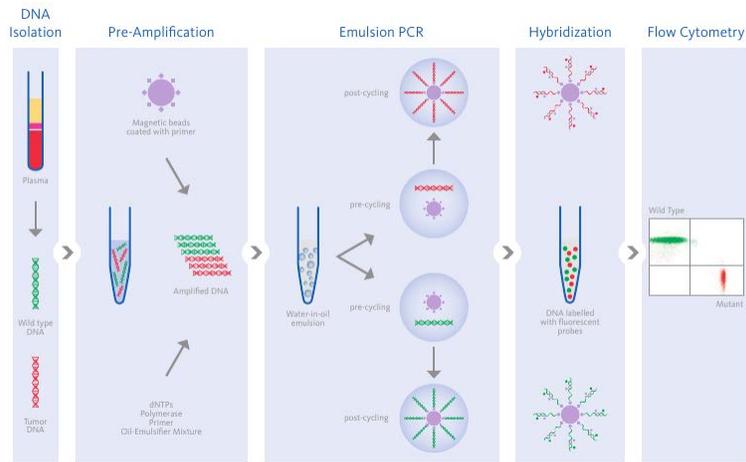


FIGURE 20 – Après extraction d'ADN issu de fluides biologiques, une pré-amplification par PCR est effectuée. Des billes magnétiques couvertes d'amorces ainsi que les réactifs pour la PCR en émulsion (amorces, nucléotides, mixture) sont ajoutés. L'émulsion eau-dans-huile permet de séparer les molécules d'ADN amplifiées, avec une seule allèle par goutte. Une PCR a lieu dans cette dernière, ce qui permet de recouvrir les billes magnétiques avec des molécules d'ADN. Après une étape de purification des billes, les sondes fluorescentes sont hybridées aux allèles amplifiés. Enfin, la lecture s'effectue par cytométrie de flux, ce qui permet la détection et quantification des allèles mutés. D'après une documentation Sysmex.

Intplex Cette méthode allèle-spécifique est présentée comme simple, peu coûteuse et ne requière que 2 mL de plasma (THIERRY et collab., 2014). De plus, aucun matériel "spécifique" n'est nécessaire pour effectuer cette analyse. Par rapport à une méthode ARMS conventionnelle, l'amorce pour l'amplification des allèles mutés possède une température de fusion basse et les amplicons sont de petites tailles (environ 60 pb) (THIERRY et collab., 2014). Un oligonucléotide bloqueur inhibe l'amplification des allèles sauvages à la localisation des mutations (THIERRY et collab., 2014). L'amplification des allèles sauvages en dehors de la région *hotspot* permet de déterminer la FAM avec précision. Une PCRtr suivie de l'analyse des courbes de fusion permet de déterminer le statut mutationnel. Des contrôles négatifs et positifs (issus de lignées cellulaires) sont employés (THIERRY et collab., 2014) (**Figure 21**). La sensibilité de cette méthode est évaluée entre 0.004% et 0.014% (THIERRY et collab., 2014).

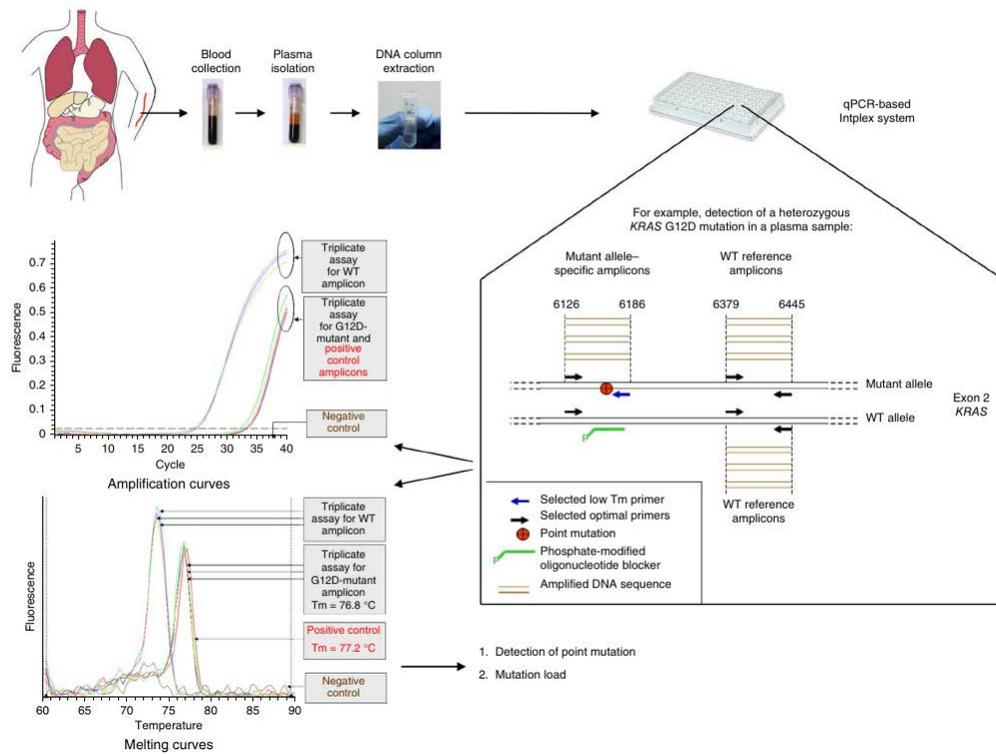


FIGURE 21 – Après prélèvement sanguin et extraction des ADNlc, ces derniers sont amplifiés par une PCR allèle-spécifique propres à la recherche d’ADNtc. Des amorces spécifiques de chaque mutation somatique recherchée, avec une température de fusion basse, ainsi que des amorces spécifiques des allèles sauvages, sont utilisées. Un oligonucléotide bloqueur permet d’empêcher l’amplification des séquences non somatiques pour les régions *hotspot*. Cette PCR est suivie en temps réel et une analyse des courbes de fusion permet de déterminer avec précision la présence ou non de mutations somatiques. De l’ADN issu de lignée cellulaire est utilisé comme contrôle positif. Les amplicons issus de séquence non mutée permettent de quantifier la FAM. D’après THIERRY et collab. (2014).

1.4.4.2 Méthodes avec enrichissement des allèles mutées

Ces approches permettent d’amplifier sélectivement les allèles supportant un variant même avec un grand nombre d’allèles sauvages présentes (LI et collab., 2008). La première méthode fut la *co-amplification at lower denaturation temperature-PCR* (COLD-PCR) (LI et collab., 2008). La **Figure 22** résume les étapes d’un cycle de

COLD-PCR.

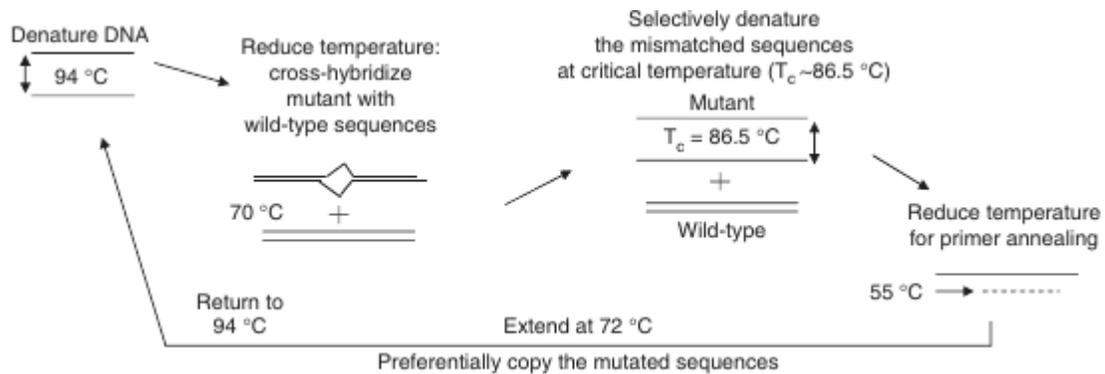


FIGURE 22 – La COLD-PCR comporte deux étapes principales. D’abord, une PCR classique sur peu de cycles, qui amplifie sans distinction les allèles sauvages et mutées. Après une dénaturation à 94 °C et une réduction à 70 °C, les amplicons issus d’allèles mutées, largement minoritaires, forment des hétéroduplexes avec les amplicons issus d’allèles sauvages. En choisissant une température de dénaturation plus basse, les hétéroduplexes sont séparés contrairement aux homoduplexes. Cela permet d’amplifier en priorité les allèles mutées. D’après LI et collab. (2008).

Par cette approche, Li et collab. rapportent une augmentation de la sensibilité par un facteur 100 comparativement à une PCR "classique" (LI et collab., 2009b). De plus, de 10 à 100 loci peuvent être amplifiés de manière simultanée avec cette approche (WAN et collab., 2017). Par la suite, l’analyse des produits de COLD-PCR peut être faite par des techniques conventionnelles de biologie moléculaire (comme le séquençage de Sanger ou le pyroséquençage) (LI et collab., 2008).

1.4.4.3 Avantages et limites des méthodes en pauci-loci

Ces approches sont généralement peu coûteuses et permettent de rendre des résultats en peu de temps. Elles permettent de suivre les mutations de résistance récurrentes, comme les mutations $EGFR^{t790m}$ (WAN et collab., 2017). Le recul par rapport à ces méthodes est grand pour certaines d’entre elles. Deux kits commerciaux disposent même d’autorisations des agences américaines et européennes comme tests compagnons pour les anti-EGFR dans les CPNPC (WAN et collab., 2017). De plus, avec la mise au point d’amorces spécifiques, ces approches peuvent être utilisées pour le monitoring de tous les patients (WAN et collab., 2017). Néanmoins, ces

méthodes ne permettent pas d'analyser l'évolution clonale des lésions tumorales et les mutations de résistance *de novo* ne sont pas détectées. Les approches basées sur le séquençage nouvelle génération peuvent représenter une alternative afin de remédier à ces désavantages (WAN et collab., 2017).

1.4.5 Méthodes d'analyses basées sur le séquençage nouvelle génération

Le séquençage nouvelle génération est apparu au milieu des années 2000, avec le développement des technologies de pyroséquençage 454 et Solexa/Illumina (METZKER, 2010). Cette dernière étant la plus utilisée de nos jours, ces principes seront les seuls décrits.

1.4.5.1 Principe de la technologie Illumina/Solexa

Cette méthode est une amplification sur phase solide (METZKER, 2010). Après amplification de fragments d'ADN (fragmentés par sonication ou lyse enzymatique) et liaison d'adaptateurs spécifiques, ces échantillons sont placés sur une lame de verre (METZKER, 2010). Sur cette dernière, deux séquences complémentaires aux adaptateurs sont présents, ainsi la première molécule se lie de manière non-covalente à la lame de verre (METZKER, 2010) (**Figure 23 a.**). La *bridge amplification* de ces molécules permet de former des clusters de molécules d'ADN (**Figure 23 a.**). La *bridge amplification* débute par l'adaptateur opposé, qui est lié à la lame. Cela engendre des clusters composés de molécules d'ADN semblables et rapprochées sur la surface de la lame de verre (**Figure 23 a.**). Enfin, une lyse chimique permet de ne conserver les molécules d'ADN que dans un seul sens de lecture (5'-3') (METZKER, 2010).

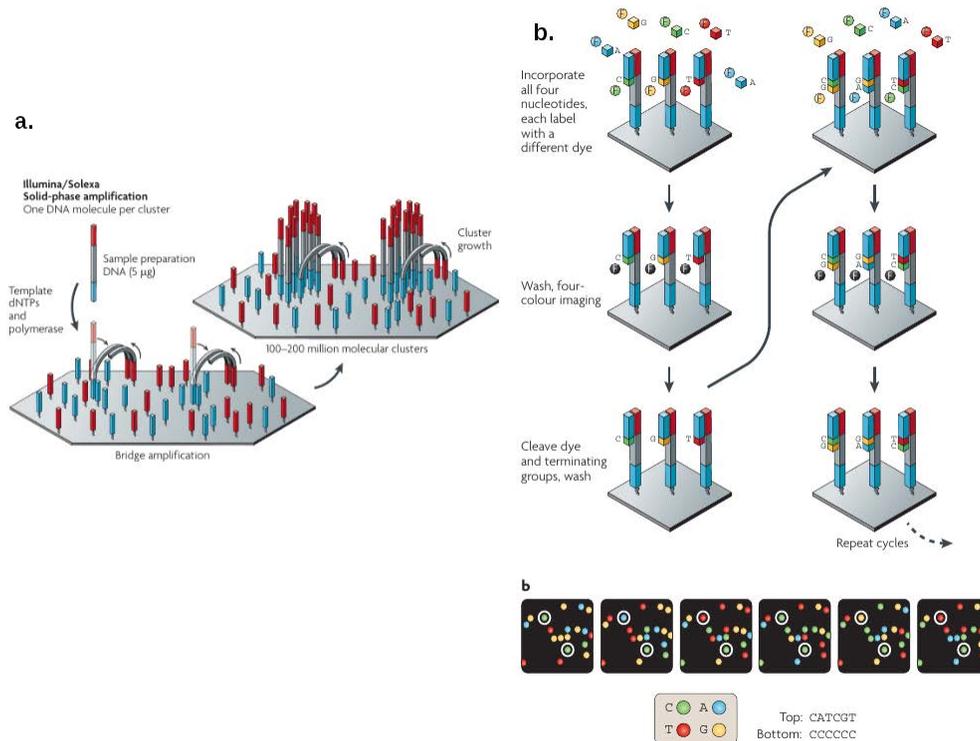


FIGURE 23 – **a.** Une seule molécule d’ADN préparée par des étapes préalables (avec notamment la liaison d’adaptateurs spécifiques) est présente par cluster. Les adaptateurs spécifiques reconnaissent ceux présents sur la lame de verre de manière non-covalente. La première étape est une *bridge amplification*, qui permet de former un cluster de molécules d’ADN semblables, liées de façon covalente à la lame de verre et proches. **b.** Chaque molécule des clusters possède une amorce universelle, à partir de laquelle une élongation débute. Les nucléotides incorporés portent tous un fluorophore bloqueur de l’élongation. Après lecture de la fluorescence par un capteur, les fluorophores sont lysés et l’élongation reprend. D’après METZKER (2010).

Ensuite, des cycles répétitifs sont effectués afin de déterminer la séquence des clusters (METZKER, 2010) (**Figure 23 b.**). D’abord les quatre nucléotides de l’ADN portant un fluorophore différent sont incorporés et ce dernier bloque l’élongation (METZKER, 2010) (**Figure 23 b.**). La fluorescence est lue et la couleur correspond à un nucléotide (METZKER, 2010) (**Figure 23 b., bas**). Enfin, le fluorophore est lysé, ce qui permet l’élongation de la molécule d’ADN synthétisée (METZKER, 2010).

Après un nombre de cycles déterminé par l'utilisateur, soit le séquençage s'arrête, on parle alors de lecture *single-end*. Sinon, les clusters sont régénérés par une nouvelle *bridge-amplification* et la lyse chimique permet la présence de clusters lus dans le sens inverse, la lecture est cette fois-ci appelée *paired-end* (METZKER, 2010). Après travail des images obtenues, les séquenceurs Illumina fournissent des fichiers FASTQ, où chaque cluster donne un *read*, qui contient la séquence nucléotidique ainsi que des informations sur la qualité du séquençage (METZKER, 2010).

Les fragments d'ADN de départ proviennent du génome entier. La technologie Illumina/Solexa permet de séquencer le génome entier d'un individu humain mais en contrepartie le coût est élevé. De plus, la couverture ne sera pas très importante, ce qui s'avère problématique dans la recherche de mutations dont la FAM est faible. Pour diminuer les montants ainsi qu'augmenter la profondeur de l'analyse, des approches par captures (soit ciblées, soit de l'exome entier) ou des amplicons (comme la méthode Tam-Seq) ont été mises au point.

1.4.5.2 Méthode par amplicons

Cette technique a été développée par l'équipe de Nitzan Rosenfeld (FORSHEW et collab., 2012). Le premier panel couvrait 5995 pb et incluait des gènes comme *KRAS*, *TP53* ou encore *BRAF* (FORSHEW et collab., 2012). Cette méthode est basée sur des amplicons de petites tailles dont les amorces se recouvrent parfois (FORSHEW et collab., 2012) (**Figure 24 a.**). L'amplification des régions se fait en parallèle, avec au départ 15 cycles pour toutes les allèles d'intérêt. Ensuite une PCR *single-plex* est utilisée pour amplifier les régions d'intérêt, afin d'éliminer les produits non-spécifiques (FORSHEW et collab., 2012). Enfin, le barcoding moléculaire pour la distinction des échantillons et la liaison d'adaptateurs de séquençage est faite (FORSHEW et collab., 2012). La lecture des amplicons obtenus est ensuite faite sur une plateforme de séquençage haut débit (FORSHEW et collab., 2012).

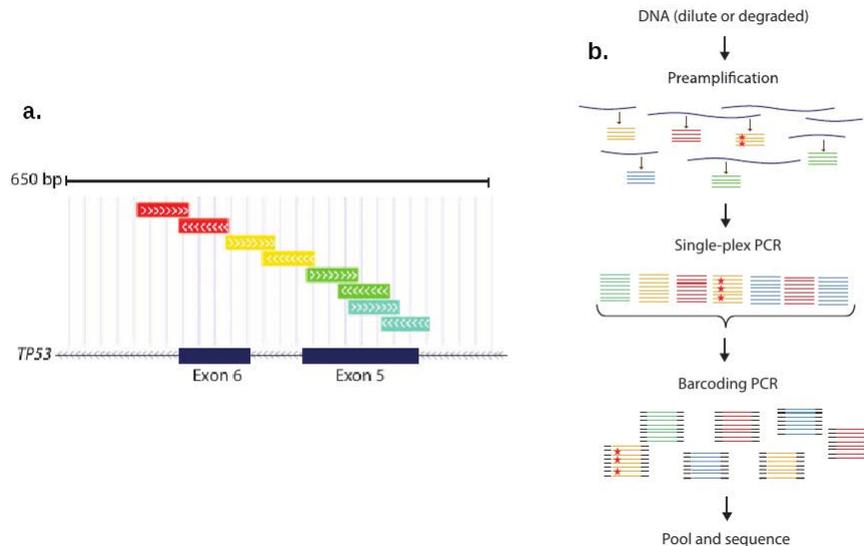


FIGURE 24 – **a.** Schéma des différents amplicons pour les exons 5 et 6 de *TP53*. Du fait de la taille réduite (< 200 pb) des produits de PCR, les amorces se recouvrent parfois. **b.** Procédure Tam-Seq schématisée. Après une préamplification de toutes les allèles présentes, une PCR *single-plex* est utilisée pour ne conserver que les régions d'intérêt. Cela permet d'exclure les produits non-spécifiques. Enfin, un barcoding moléculaire et des adaptateurs de séquençage sont liés aux produits de PCR. D'après FORSHEW et collab. (2012).

La première évaluation de cette approche rapporte que les mutations présentes à une FAM > 2% sont retrouvées avec une spécificité et une sensibilité supérieures à 97% (FORSHEW et collab., 2012). Ainsi, une limite de cette approche est qu'elle ne peut être employée que chez les patients avec un cancer avancé (FORSHEW et collab., 2012). Les auteurs avancent que la COLD-PCR pourrait améliorer la sensibilité de la méthode (FORSHEW et collab., 2012). Les gènes de fusion sont difficilement détectables par cette approche, il faudrait connaître de façon exacte les points de cassure. Des améliorations portées à la méthode Tam-Seq ont permis d'augmenter la sensibilité à une FAM de 0,5% et de même détecter les VNC (GALE et collab., 2016).

Au sein de la technologie *Safe-Sequencing System* (Safe-SeqS), chaque fragment d'ADN est identifié par un barcoding moléculaire unique (KINDE et collab., 2011). Seules les molécules "tagguées" sont ensuite amplifiées, ce qui donne des "familles"

d'amplicons qui ne dérivent que d'une seule molécule. Ces "familles" sont ensuite séquencées : si plus de 95% des molécules qui la composent portent la même mutation, les fragments de PCR sont considérés comme mutés (KINDE et collab., 2011). Cette approche permet de diminuer les faux-positifs induits par des erreurs techniques et de corriger les biais induits par la PCR (KINDE et collab., 2011). La sensibilité est aussi considérablement augmentée (KINDE et collab., 2011). Néanmoins, le protocole pour l'analyse Safe-SeqS est complexe, avec une purification sur gel à effectuer (KINDE et collab., 2011).

1.4.5.3 Méthode par capture

Ces techniques utilisent des sondes complémentaires des régions d'intérêts. Ces sondes sont soit fixées sur une lame (capture par phase solide) soit par capture en solution. Dans cette dernière, les sondes complémentaires sont biotinylées et les billes couvertes de streptavidines (c'est alors une capture en solution). Différents kits sont proposés mais le principe général reste similaire (la **Figure 25** décrit le protocole SeqCap de Roche). Les sondes sont nucléotidiques (ADN ou ARN) et d'une longueur de l'ordre de la centaine de bases. Le design de la capture fait souvent appel au *tailing*, c'est-à-dire au chevauchement des sondes, ce qui permet qu'une base d'intérêt soit couverte par plusieurs sondes. L'intérêt de la capture est sa grande optimisation, favorisée par les fabricants. En effet, cette technologie permet de séquencer quelques exons jusqu'à leur ensemble. Dans le cas de l'ADNtc, plusieurs méthodes sont proposées.

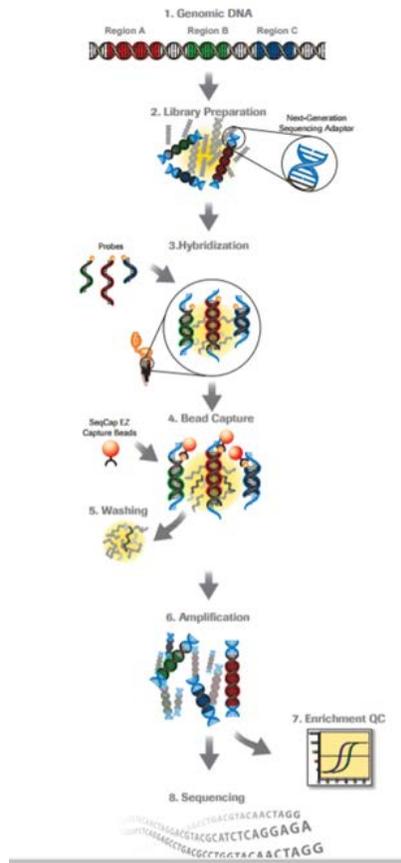


FIGURE 25 – Le protocole SeqCap est divisé en huit étapes. 1) le choix des régions d'intérêt. 2) La fragmentation de l'ADN extrait se fait par sonication. Ensuite des adaptateurs sont liés aux fragments 3) Hybridation : les séquences complémentaires biotinylés s'hybrident aux régions génomiques d'intérêt. 4) Les billes magnétiques couvertes de streptavidine permettent une capture des complexes. 5) Un lavage élimine les fragments non liés. 6) Les fragments capturés sont amplifiés. 7) Un contrôle de qualité de la capture et de l'amplification est effectué. 8) Les produits obtenus peuvent être séquencés. D'après une documentation Roche.

CAPP-Seq Cette technologie a été mise au point par l'équipe de Diehn (NEWMAN et collab., 2014b). Les auteurs ont d'abord analysé les données de séquençage disponibles pour les CPNPC et ont construit un algorithme de récurrence qui permet d'isoler les exons qui sont les plus porteurs de mutations (NEWMAN et collab., 2014b). Les gènes de fusion ont aussi été rajoutés dans ce panel, notamment les fusions *ALK-*

MET (NEWMAN et collab., 2014b). À partir d'une capture de 125 kpb, au moins une mutation somatique peut être détectée chez 95% des patients atteints de CPNPC (NEWMAN et collab., 2014b). La sensibilité de cette méthode est élevée, de 0,01% (NEWMAN et collab., 2014b). La détection des ADNtc a été améliorée en combinant le barcoding moléculaire et la suppression *in silico* d'erreurs de séquençage (NEWMAN et collab., 2016). Cette approche permet de séquencer un grand nombre de régions d'intérêts, avec une grande profondeur et pour un coût bien moins élevé que de l'exome ou du génome (NEWMAN et collab., 2014b).

Exome Le séquençage de l'exome entier a été utilisé avec succès pour détecter des ADNtc et suivre l'évolution clonale des tumeurs (MURTAZA et collab., 2013). Néanmoins, le coût de cette approche est élevé et la sensibilité est faible (5%) (WAN et collab., 2017). C'est pourquoi cette analyse ne peut pas être utilisée en pratique courante et n'est utilisée qu'en recherche pour des patients avec une forte masse tumorale (WAN et collab., 2017).

1.4.5.4 Génome entier avec couverture faible

Les VNC sont représentés par des amplifications et délétions de régions du génome et peuvent être détectées même avec une faible couverture (WAN et collab., 2017). Le séquençage du génome avec une faible couverture (*shallow whole genome sequencing*) permet de détecter les VNC dans le plasma en comparant des échantillons *versus* des contrôles (WAN et collab., 2017). Cette méthode a été utilisée pour détecter les aneuploïdies fœtales mais aussi les VNC spécifiques des cancers (WAN et collab., 2017). Ces altérations somatiques présentent l'intérêt d'être moins sensibles au bruit de fond, comparées aux mutations ne touchant qu'une seule base (WAN et collab., 2017). Néanmoins, la sensibilité de cette analyse reste faible (entre 5 et 10%) et son coût élevé (WAN et collab., 2017).

1.4.6 Conclusion sur les méthodes d'analyses des ADN tumoraux circulant

Depuis la mise au point de l'ARMS-PCR en 1989, un grand nombre de méthodes n'analysant qu'un locus ou pauci-loci ont été mises au point (WAN et collab., 2017) (**Tableau 2**). La révolution du séquençage haut débit au milieu des années 2000 a permis de diminuer grandement les coûts de séquençage (METZKER, 2010). Dans le cadre des ADNtc, cette nouvelle technologie a permis de séquencer un grand nombre de loci voire même le génome entier circulant (WAN et collab., 2017). Cette

approche permet de compenser la sensibilité plus faible des techniques qui utilisent le séquençage haut débit par rapport aux méthodes locus unique, en augmentant la probabilité de détecter une autre altération (WAN et collab., 2017) (**Tableau 2**). La sélection de fragments de petites tailles, enrichis en ADNtc, pourrait augmenter la sensibilité des analyses (WAN et collab., 2017) Néanmoins, cet accroissement du nombre de régions analysées induit une nette augmentation du risque de faux-positifs (WAN et collab., 2017). De plus, les données générées par cette approche sont importantes et nécessitent l'utilisation de la bioinformatique afin d'extraire les altérations d'intérêt (WAN et collab., 2017). Le panel d'experts de l'Institut national du cancer (INCa) exige de valider les *pipelines* d'analyses (INCA, 2017).

TABLE 2 – Méthodes d’analyses pour l’ADNtc. D’après WAN et collab. (2017)

Type d’analyses	Exemples	Nombre de <i>loci</i>	Limite de détection	Utilité clinique
Singleplex ou multiplex	Microfluidique ou PCR allèle-spécifique	1-10	Selon la méthode, 0,001%–0,01% voire une mutation somatique par millilitre	Détection et quantification des mutations hotspot Monitoring des mutations de résistance récurrentes Analyse rapide
	PCR digitale BEAming Intplex	ADNtc et ADNlc pour Intplex		
	Enrichissement en allèles mutées COLD-PCR SCODA NaME-PrO	10-100 <i>loci</i>		
	ARMS-PCR, kits commerciaux Cobas EGFR Therascreen EGFR	Cobas EGFR : 7 analyses couvrant 41 mutations Therascreen EGFR : 3 analyses couvrant 29 mutations	A une sensibilité > 95% Cobas EGFR 25-100 copies par millilitre Therascreen EGFR : médiane 1,42% (0,05%–12,47% selon le variant)	Approuvé par : FDA (COBAS EGFR) CE-IVD (Therascreen EGFR)
Séquençage haut débit ciblé	Basée sur des amplicons : TAm-Seq et enhanced Tam-Seq Safe-SeqS	10 kb à 50 Mb	<0,01%-0,5% pour les panels spécifiques 1% pour les panels multiplex 5% pour l’exome	Profil d’un panel de gènes Détection de mutation de novo pour les résistances Évaluation de l’évolution clonale Augmentation de sensibilité par test d’un grand nombre de loci
	Basée sur de la capture : Exome CAPP-seq			
Génome entier	Séquençage du génome entier à base couverture	3,2 Gb	5%–10%	Identification des variants structuraux Stratification des patients selon la charge tumorale Détection des aberrations chromosomiques

2 Matériel et Méthodes

Un *pipeline* est l'ensemble des outils nécessaires pour aligner les séquences, appeler les variants et annoter ces derniers. Comme décrit ci-dessus, l'augmentation des régions séquencées implique une augmentation des faux-positifs (erreurs techniques, erreurs de la polymérase, de lectures, etc.) (WAN et collab., 2017). Des filtres peuvent être appliqués afin de diminuer ces faux-positifs, mais le risque peut s'inverser. C'est pourquoi les *pipelines* "participe[nt] à l'évaluation de la sensibilité et de la spécificité" (texte issu du guide pour la validation de méthode publié par l'INCa (INCA, 2017)). La **Figure 26** schématise le *pipeline* mis en place et évalué dans ce travail. Une partie de ce pipeline dérive des *best-practices* proposées par le Broad Institute (VAN DER AUWERA et collab., 2013).

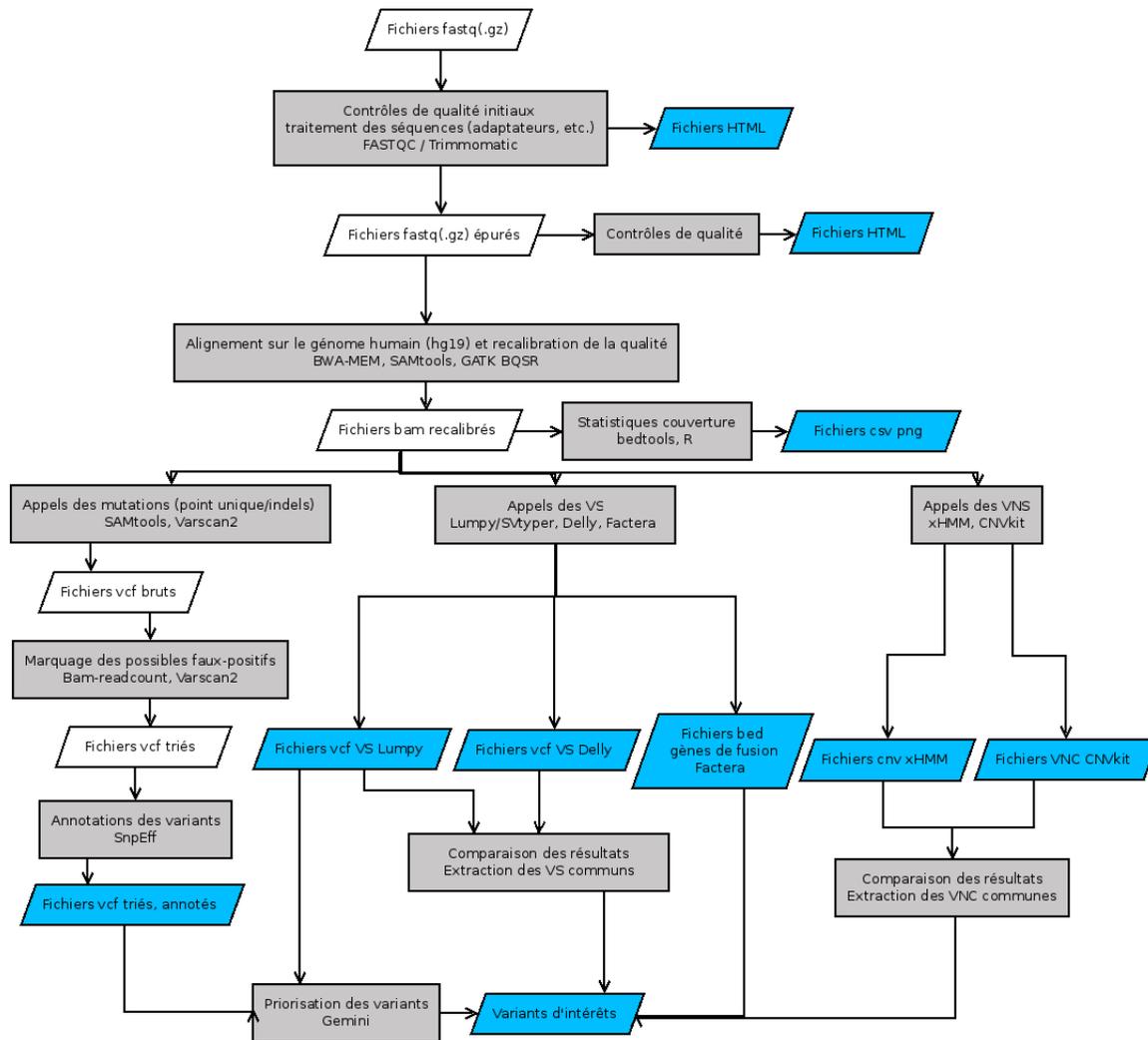


FIGURE 26 – Ce pipeline comporte plusieurs étapes majeures. D’abord l’alignement des séquences produites par l’appareil ainsi que les statistiques (qualités, couvertures, etc.) sont produites. Ensuite, les différents variants (des mutations en point unique aux VNC) sont appelés et filtrés. En outre, ces altérations sont annotées afin d’évaluer leur impact. Enfin, une priorisation permet d’extraire les variants d’intérêt. Le fond gris désigne les tâches effectuées par le *pipeline* avec sur la seconde ligne les logiciels employés. Les fichiers devant être contrôlés par le biologiste sont en fond bleu ciel

2.1 Contrôles de qualité, alignement, statistiques, re-calibration

2.1.1 Contrôles de qualité

La première étape fait appel à des fichiers *fastq*. Ces derniers sont des fichiers textes qui comportent les séquences nucléotidiques et la qualité des bases (**Figure 27**).

2.1.2 Alignement sur le génome et traitement des séquences alignées

Cette partie permet de déterminer la localisation génomique des *reads*. Néanmoins, des discordances avec les nucléotides du génome (qui résultent d'erreur technique ou des mutations, polymorphismes d'une seule base, insertion/délétions) sont présentes dans les séquences, ce qui représente un challenge dans l'alignement. En effet, les aligneurs doivent retrouver la localisation la plus probable, avec un nombre acceptable d'erreurs. BLAST (pour *Basic local alignment search tool*) est l'aligneur le plus employé et utilise un algorithme heuristique (ALTSCHUL et collab., 1990). Cet outil très puissant et précis nécessite néanmoins une puissance et un temps de calcul trop important dans le cadre de l'alignement pour le séquençage haut débit. C'est pourquoi de nouveaux algorithmes ont été mis au point, qui permettent un gain de temps et de puissance considérable. Dans ce travail, l'aligneur utilisé est *bwa* (*Burrows-Wheeler Alignment*) avec son algorithme *mem* et les options -M (permettant de marquer les secondes localisations des *reads* comme secondaire) et -B 3 (diminuant l'erreur d'une discordance à 3 au lieu de 4) (LI, 2013). Le génome humain de référence utilisé est l'hg19, avec des séquences leurres. Cela permet un alignement de qualité tout en excluant les *reads* qui sont localisés sur les leurres (ALIOTO et collab., 2015). Les fichiers obtenus après cette étape sont appelés des fichiers SAM (ou BAM lorsqu'ils sont compressés). Ces derniers sont triés par *samtools* afin que les positions génomiques des alignements se succèdent (LI et collab., 2009a). Les duplicats de PCR sont des erreurs techniques qui apparaissent lors de l'amplification. Une erreur de la polymérase par exemple peut se retrouver dans un grand nombre de *read* de ce fait et être considérée à tort comme un variant. Les duplicats de PCR, qui sont représentés par des *reads* similaires en tout point, sont marqués par l'option *MarkDuplicates* du logiciel *picard*, ce qui permet de les exclure lors de l'appel des variants (<http://broadinstitute.github.io/picard/>). Néanmoins, l'ADNtc comporte des petits fragments, il existe un risque de marquage de duplicats à tort (NEWMAN et collab., 2014b). De plus, ce marquage diminue la profondeur, ce qui s'avère limitant lors de l'appel des variants avec une faible FAM (NEWMAN et collab., 2014b). Dans ce travail, l'approche par marquage des duplicats et non-marquage a été évalué. Les statistiques effectuées sur les données alignées permettent d'explorer de nouveau la qualité de séquençage. L'utilisation de *bedtools* permet de déterminer la couverture des régions d'intérêt, mettant en évidence des erreurs techniques, par exemple une capture insuffisante (QUINLAN et HALL, 2010).

L'attribution des scores *PHRED-like* par les séquenceurs n'est pas correctement décrite par les fabricants (VAN DER AUWERA et collab., 2013). Une surestimation de la qualité par les séquenceurs Illumina a été rapportée (QU et collab., 2009). La qual-

ité des nucléotides est une donnée essentielle lors de l'appel des variants. En effet, un score *PHRED-like* bas est plus en faveur d'une erreur technique et il est rapporté une baisse de celui-ci si la base est un variant. En s'appuyant sur des bases de variants connus, il est possible de recalibrer la qualité des bases (VAN DER AUWERA et collab., 2013). Globalement, un nucléotide correspondant à celui du génome de référence ou à un variant préalablement décrit verra son score *PHRED-like* augmenter tandis qu'un nucléotide portant un variant totalement inconnu aura une qualité abaissée (VAN DER AUWERA et collab., 2013). Cette recalibration est effectuée par l'algorithme *BaseRecalibrator* du logiciel *GATK* (*Genome Analysis Toolkit*) (MCKENNA et collab., 2010). Les bases de données utilisées sont les suivantes : 1000genomes pour les polymorphismes d'une base, Mills et 1000genomes pour les insertions/délétions polymorphiques et le *Catalogue of Somatic Mutations in Cancer* (COSMIC) pour les mutations somatiques.

Cette approche, recommandée par le *Broad Institute*, permet d'obtenir des alignements avec une qualité optimale pour l'appel des variants (VAN DER AUWERA et collab., 2013). De plus, les nombreux contrôles qualités effectués permettent la mise en évidence des possibles problèmes techniques comme la capture des régions d'intérêt. L'analyse de ces derniers permet une amélioration de la partie séquençage.

2.1.3 Appels des variants

Après l'obtention d'alignements sur un génome de référence, les variants par rapport à ce dernier sont appelés. Ceux-ci sont de longueur variable : ils peuvent toucher une seule base jusqu'à des régions de tailles importante, de l'ordre de la mégabase (dans le cas des VNC). Ainsi, différents outils ont été mis au point afin de les retrouver à partir des alignements.

D'abord l'utilisation de *varscan2* permet de retrouver les mutations en point unique tout aussi bien que les insertions/délétions de quelques dizaines de paires de bases (KOBOLDT et collab., 2012). L'évaluation de ce logiciel a montré une excellente sensibilité et spécificité (ALIOTO et collab., 2015). Ce logiciel, dont l'algorithme est basé sur une méthode statistique/heuristique, utilise des fichiers *mpileup* issus de *samtools* (LI et collab., 2009a). Ces derniers représentent un empilement des nucléotides séquencés à chacune des positions génomiques. Le *pipeline* présenté ici fait appel à deux fichiers *mpileup* différents. Le premier est destiné à l'appel des mutations en point unique, la qualité des nucléotides et de l'alignement des *reads* sont conservées. Le deuxième est moins stringent, ces indicateurs ne sont pas retenus : ce fichier est utilisé pour la découverte des insertions/délétions, plus difficiles à mettre

en évidence. Une approche par échantillon plutôt qu'un empilement de toute la cohorte est préférée. En effet, du fait de la couverture importante lors de l'analyse des ADNtc, il existe un risque de faux-négatifs par un *mpileup* global, la couverture cumulée dépassant la capacité d'analyse de *samtools*. Ensuite, *varscan2* analyse chaque position afin d'extraire les variants à l'aide d'une méthode heuristique et statistique (KOBOLDT et collab., 2012). Pour les mutations en point unique, la couverture minimum est de 8x, le variant doit être présent sur 2 *reads*, le score *PHRED* moyen des nucléotides doit être de 20, la *p-value* < 0,05 et la FAM est au minimum à 0,0001. Les critères sont moins stringents pour la recherche des insertions/délétions mais leur FAM minimale est de 0,001.

Avec ces paramètres, de nombreux faux-positifs sont présents. Des paramètres de *varscan2* plus restrictifs permettent de diminuer leur taux (WARDEN et collab., 2014). Néanmoins, cette approche est à risque de faux-négatifs. Il a été préféré d'effectuer un filtre sur une analyse sensible afin d'augmenter la spécificité, comme recommandé par Koboldt et collab., Alioto et collab. ainsi que le *Broad Institute* (KOBOLDT et collab., 2012; VAN DER AUWERA et collab., 2013; ALIOTO et collab., 2015). Dans ce travail l'utilisation de *bam-readcount* et de la fonction *fpfilter* de *varscan2* permet de marquer les variants comme potentiellement faux-positifs, afin de faciliter leur priorisation (KOBOLDT et collab., 2012). Par exemple, les critères comme la différence du score *PHRED-like* entre les nucléotides variants et ceux de références ou la position des variants aux seins des *reads* (plus la position est en 3' sur le *read* et plus la qualité diminue) sont utilisés. Cette analyse renseigne dans la colonne "INFO" du fichier *vcf* les critères "PASS" pour les variants de haute qualité ou le critère de filtration (**Figure 28**). Les filtres utilisés pour *VarDict* sont ceux proposés par les auteurs : biais de sens, différence de qualité entre les nucléotides variants et de référence, positions des variants au sein des *reads*, si le variant fait partie d'une délétion et le rapport signal sur bruit (LAI et collab., 2016).

Un deuxième logiciel pour appeler les variants, *Vardict*, est aussi utilisé (LAI et collab., 2016). *VarDict* utilise directement les fichiers BAM contrairement à *VarScan2* et grâce à un réalignement local, est plus précis pour l'appel des indels (LAI et collab., 2016).

Les VS ainsi que les VNC, touchant un grand nombre de gènes, sont impliqués dans l'oncogénèse. Les outils pour les rechercher font appels à d'autres approches que ceux pour les variants n'impactant que quelques nucléotides. Les VS sont mis en évidence dans ce travail par trois logiciels. Le premier est *lumpy* qui recherche les points de cassure sur des critères comme la couverture (pour détecter les duplications et délétions), les discordances d'alignement sur le génome de paires de *reads* ou les *reads*

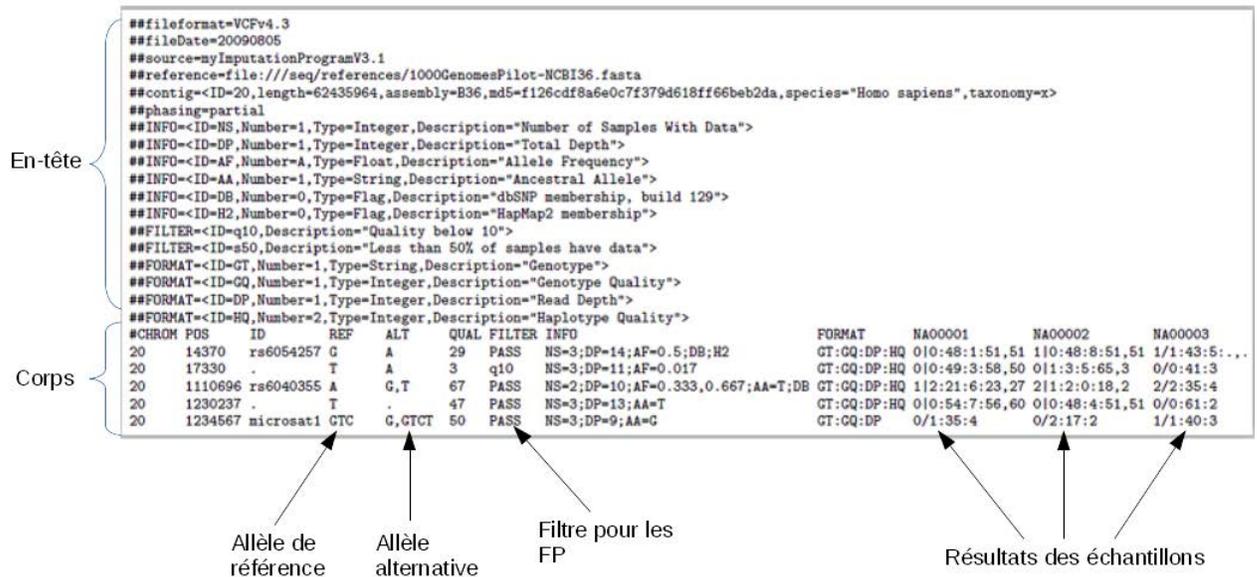


FIGURE 28 – Ces fichiers textes contiennent les variants détectés. Ces premiers sont divisés en plusieurs parties. D’abord l’entête, qui comprend la description des différents champs. Ensuite le corps de fichier comporte des colonnes dans lesquelles les variants sont décrits. D’abord les trois premières indiquent leurs positions génomiques, ainsi que l’allèle de référence et alternative. Ensuite, des colonnes d’informations sont présentes. Celle nommée "FILTER" comporte les filtres présents ou non pour trier les faux des vrais-positifs. Ensuite, les résultats des échantillons sont présents : 0/0 indique que le patient est homozygote sauvage, 0/1 hétérozygote et 1/1 homozygote muté

dont la paire n’est pas alignée (LAYER et collab., 2014). Le deuxième logiciel, *delly*, fait appel à d’autres méthodes pour appeler les VS que *lumpy*. La combinaison de deux logiciels permet d’augmenter la spécificité et la sensibilité pour l’appel des VS. Les gènes de fusion, même s’ils peuvent être mis en évidence par *lumpy*, sont détectés par *factera*, une approche spécialisée pour la détection de ceux-ci (NEWMAN et collab., 2014a). Ce logiciel a été utilisé avec succès pour mettre en évidence les gènes de fusion dans les CPNPC et les LBDGC (NEWMAN et collab., 2014b; KURTZ et collab., 2015). Enfin la détection des VNC paraît simple : les régions amplifiées présentent une couverture supérieure à la médiane tandis qu’une diminution est corrélée à une délétion. Néanmoins, cette approche peut être utilisée pour la détection des VNC dans le séquençage du génome entier mais cette mise en évidence s’avère plus com-

pliquée dans les méthodes par capture. En effet cette dernière est influencée par de nombreux facteurs, qu'ils soient génomiques (pourcentage en GC, régions répétitives ou complexes) ou techniques (effet *batch* sous l'influence de l'opérateur, des conditions environnementales, etc.). Il s'avère donc nécessaire de normaliser la couverture avant d'appeler les VNC ou de comparer les séquençages des échantillons tumoraux *versus* des paires d'échantillons issus d'ADN germinale (méthode qui se rapproche des puces CGH). Afin d'augmenter la sensibilité et la spécificité de la détection des VNC, une approche par deux logiciels présentant une normalisation est utilisée. Le premier algorithme utilisé est *xHMM* qui se base sur une normalisation par analyse en composante principale afin de transformer les couvertures en z-score (FROMER et collab., 2012). Le second logiciel est *cnvkit*, qui normalise les couvertures en utilisant les *reads* présents sur les cibles et les *reads* en dehors des cibles (ces derniers représentant une capture aspécifique) (TALEVICH et collab., 2016). Cette approche permet de calculer une couverture moyenne sur chaque intervalle, qui est ensuite combinée (TALEVICH et collab., 2016). En comparant cette dernière à une table issue d'échantillons contrôles, une table des copies en \log_2 est extraite (TALEVICH et collab., 2016). Cette dernière sert pour la visualisation et l'appel des VNC (TALEVICH et collab., 2016). Ces logiciels ont été mis en place seulement mais non évalués, car nécessitant plus d'échantillons (*xHMM*) ou des contrôles sains (*cnvkit*).

2.1.4 Annotation des variants et priorisation

Les *vcf* ne comportent initialement qu'une entête, les positions génomiques des variants ainsi que des paramètres comme la profondeur, la qualité des bases, etc. (**Figure 28**). Il s'avère donc difficile d'évaluer l'impact des variants sans une annotation de ceux-ci. Différents outils et bases de données existent afin de réaliser ce processus. Dans le *pipeline* mis au point ici, le logiciel utilisé est *SnpEff* avec la base *GRCh37.75* d'*Ensembl* (CINGOLANI et collab., 2012). Cette étape permet de déterminer quelle partie (exons, introns, site d'épissage, etc.) est impactée, si le variant modifie la structure protéique ou non, est-il délétère ou non. De plus, l'utilisation de *SnpEff* permet de connaître l'effet du variant sur les différents transcrits (CINGOLANI et collab., 2012). En effet, ces derniers ne sont pas impactés de façon similaire par le même variant (**Figure 29**). il existe une influence importante des outils d'annotation utilisés (MCCARTHY et collab., 2014). Si une grande concordance existe entre les annotations pour les variants situés sur les exons, elle est beaucoup plus faible en ce qui concerne les altérations pouvant entraîner une perte de fonction (*stop-gain* ou décalage du cadre de lecture) (MCCARTHY et collab., 2014).

L'annotation est une des premières étapes de la priorisation des variants. Plusieurs

```

##$npEffVersion="4.3t (build 2017-11-24 10:18) by Pablo Cingolani"
##$npEffFile="SnEff_6Rch37_75_..76_filtered_vcf/SRR1200742_filtered_variants_normalized.html"
##$INFO~ID=ANI,Number=,Type=String,Description="Functional annotations: 'Allele | Annotation | Gene_Name | Gene_ID | Feature_Type | Feature_ID | Transcript_BioType | Rank | HGVS.p | HGVS.c | cDNA_pos / cDNA_length | CDS_pos / CDS_length | AA_pos / AA_length | Distance | ERRORS | WARNINGS | INFO'">
##$INFO~ID=LOF,Number=,Type=String,Description="Predicted loss of function effects for this variant. Format: 'Gene_Name | Gene_ID | Number_of_transcripts_in_gene | Percent_of_transcripts_affected'">
##$INFO~ID=MD,Number=,Type=String,Description="Predicted nonsense mediated decay effects for this variant. Format: 'Gene_Name | Gene_ID | Number_of_transcripts_in_gene | Percent_of_transcripts_affected'">
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT Sample1
chr1 77752549 A T PASS ADP=589;MT=0;HET=1;HOM=0;NC=0;ANN=C|intron_variant|MODIFIER|AKS|ENSG00000154027|transcript|ENST00000354567|protein_coding|1/13|c.61-77A>T|||||T|intron_variant|MODIFIER|AKS|ENSG00000154027|transcript|ENST00000317704|processed_transcript|1/5|n.317-77A>T|||||T|intron_variant|MODIFIER|AKS|ENSG00000154027|transcript|ENST00000344720|protein_coding|1/13|c.-16-79A>T|||||T|intron_variant|MODIFIER|AKS|ENSG00000154027|transcript|ENST00000478407|protein_coding|1/4|c.-18-77A>T|||||WARNING_TRANSCRIPT_INCOMPLETE GT:GQ:SDP:DP:RD:AD_0:FREQ:PVAL:RBQ:ABQ:RDF:RDR:ADF:ADR:AD 0/1:255:589:589:293:296:50,25%:0,7402E-112:40:40:271:22:204:32:293:296
chr1 77752996 T C PASS ADP=256;MT=0;HET=0;HOM=1;NC=0;ANN=C|intron_variant|MODIFIER|AKS|ENSG00000154027|transcript|ENST00000354567|protein_coding|2/13|c.247+184T>C|||||C|intron_variant|MODIFIER|AKS|ENSG00000154027|transcript|ENST00000317704|processed_transcript|2/5|n.503+184T>C|||||C|intron_variant|MODIFIER|AKS|ENSG00000154027|transcript|ENST00000344720|protein_coding|2/13|c.169+184T>C|||||C|intron_variant|MODIFIER|AKS|ENSG00000154027|transcript|ENST00000478407|protein_coding|2/4|c.169+184T>C|||||WARNING_TRANSCRIPT_INCOMPLETE GT:GQ:SDP:DP:RD:AD_0:FREQ:PVAL:RBQ:ABQ:RDF:RDR:ADF:ADR:AD 1/1:255:256:256:0:256:100%:2,1162E-153:0:37:0:0.5:251:0,256

```

Colonne info annotée

FIGURE 29 – Ces fichiers vcf annotés comprennent les mêmes parties que décrites dans le **Figure 28**. L'annotation, présente en colonne "INFO", permet de donner un sens biologique à un variant, en le rattachant au gène (notés ENSGXXX) potentiellement impacté, si la mutation est synonyme ou non-sens, etc. L'annotation par des logiciels comme *SnEff* permet de voir l'impact de l'altération sur les différents transcrits (notés ENSTXXX).

critères peuvent être utilisés afin de mettre en évidence les potentiels variants d'intérêt. D'abord des filtres précédemment décrits pour éliminer les variants appelés mais qui sont potentiellement des erreurs techniques. Ensuite, les annotations fonctionnelles sont utilisées. Par exemple, les mutations exoniques n'entraînant pas de changement d'acide aminé (mutation silencieuse) sont moins délétères que les mutations non-sens. Il existe des scores qui permettent d'évaluer l'impact des variants non décrits. *SIFT* (Sorting Intolerant From Tolerant) et *POLYPHEN* (Polymorphism Phenotyping) sont deux outils de prédiction couramment utilisés (COOPER et SHENDURE, 2011). Le premier fait appel à la génomique comparative et indirectement à la composition de la protéine et aux propriétés physico-chimiques des acides aminés : plus un amino-acide est conservé au sein des l'évolution et plus son changement est délétère (NG et HENIKOFF, 2003). Le score *POLYPHEN* s'appuie sur des données de génomique comparative, biochimique et structurale (ADZHUBEI et collab., 2010). Enfin, il est possible d'utiliser des bases de données diverses. Certaines permettent de connaître la fréquence allélique de l'altération. Si ce dernier entraîne une maladie grave, sa fréquence doit être faible au sein de l'ensemble des populations, ce qui permet d'exclure les variants avec une fréquence allélique élevée. Dans ce travail, plusieurs bases sont utilisées, comme ExAC 1000genomes ou gnomAD, avec un seuil pour la fréquence allélique de 5%. D'autres archives publiques, telles que ClinVar ou COSMIC, sont exploitées. ClinVar rapporte les relations entre les variations de séquence et les phénotypes (LANDRUM et collab., 2013). De plus, l'interprétation concernant le degré de pathogénicité est aussi décrit (LANDRUM et collab., 2013).

COSMIC représente un catalogue de mutations somatiques rapportées par des experts.

Dans ce *pipeline*, le logiciel utilisé pour le tri et la priorisation des variants est *GEMINI* (PAILA et collab., 2013). Cet outil permet d'extraire les variants d'intérêts dans un fichier *csv*, lisible depuis n'importe quelle plateforme.

Après une sélection et mise en place d'outils, les performances du *pipeline* (dont le code se trouve en annexe) doivent être évaluées.

2.2 Évaluation des performances du *pipeline*

Les différents séquençages utilisés pour évaluer le *pipeline* décrit ont été extraits de la base *Sequence Read Archive* (<https://www.ncbi.nlm.nih.gov/sra>).

2.2.1 Évaluation des annotations et de la priorisation des variants

Cette partie, essentielle au gain de temps pour le biologiste en charge de la validation des résultats, est évaluée en analysant les séquençages rendus publics par Quest Diagnostic (STROM et collab., 2015). Vingt-sept échantillons ont été analysés par le *pipeline* et les variants priorisés ont été comparés à ceux rapportés par Strom et collab.

2.2.2 Évaluation de la sensibilité et spécificité de l'appel des variants

Des séquençages d'ADN issu de cellule HCC-78 provenant d'un adénocarcinome pulmonaire type CPNPC. Cette lignée porte un nombre important de mutations et un réarrangement de *ROS1*. Lors de la mise au point de la méthode CAPP-Seq, Newman et collab. ont réalisé le séquençage à très haute couverture d'ADNlc issu de plasma d'un sujet sain, des cellules HCC-78 ainsi que de dilutions de ces dernières dans le fluide biologique (1%, 0,5%, 0,1%, 0,05% et 0,025%) (NEWMAN et collab., 2014b). Cette analyse permet de tester la sensibilité et la spécificité du *pipeline* à détecter des variants somatiques avec une très faible FAM. Les impacts de l'utilisation de *Trimmomatic* et de la recalibration de la qualité des bases sont aussi évalués. Les variants des cellules HCC-78 sont appelés par *VarDict* (options par défauts, sauf la FAM qui est abaissée à 0,0001). Ces mutations sont ensuite extraites par *GEMINI* et ensuite triées sur *R* avec les paramètres suivants : profondeur > 50, FAM > 7,5% et

aucun filtre. À partir de cette liste de variants, le rappel est calculé selon la formule suivante :

$$Rappel = \frac{\text{Nombre de mutations correctement détectées}}{\text{Nombre de mutation HCC} - 78} \quad (1)$$

Pour la précision, les variants présents dans le plasma du sujet sain sont exclus, ainsi que ceux filtrés par *VarDict* (à l'exception du filtre de la FAM). Cette mesure est calculée selon l'équation :

$$Précision = \frac{\text{Nombre de mutations correctement détectées}}{\text{Nombre de mutations détectées totales}} \quad (2)$$

La moyenne harmonique des ces deux mesures donne le score F1 (lorsqu'il est égal à 1 est parfait) :

$$Score F1 = 2 \cdot \frac{Précision \cdot Rappel}{Précision + Rappel} \quad (3)$$

3 Résultats

3.1 Annotations et priorisation des variants

La priorisation à l'aide de *gemi* permet d'isoler un nombre restreint de variants : en médiane, seul deux altérations sont retenues. Le **Tableau 3** résume la comparaison entre les annotations des Strom et collab. et celles obtenues par le *pipeline*. Des discordances peuvent être apparentes entre les deux annotations, notamment au niveau des indels. Par exemple, pour l'échantillon GM13709, Strom et collab. rapportent une délétion en c.2068, tandis que le *pipeline* indique une position en c.2071. Néanmoins, le rsid reste le même (rs80357688) et l'annotation peut être rapportée des deux manières. C'est pourquoi il est conclu que la concordance entre les deux annotations est de 100%.

TABLE 3 – Comparaison des annotations entre Strom et collab. et celles du *pipeline*

Noms		Résultats Strom et collab.			Résultats pipeline		rsid	Filtre
Échantillon	Archive	Gène	Codon	Protéine	Codon	Protéine		
GM13705	SRR2925409	BRCA1	c.3756_3759delGTCT	p.Leu1252_Ser1253delinsLeufs	c.3756_3759delGTCT	p.Ser1253fs	rs80357868	Non
GM13708	SRR2925410	BRCA1	c.4752C>G	p.Tyr1584Ter	c.4752C>G	p.Tyr1584*	rs80357433	Non
GM13709	SRR2925411	BRCA1	c.2068delA	p.Lys690 = fs	c.2071delA	p.Arg691fs	rs80357688	Non
GM13710	SRR2925412	BRCA1	c.4327C>G	p.Arg1443Gly	c.4327C>G	p.Arg1443Gly	rs41293455	Non
GM13711	SRR2925413	BRCA1	c.3119G>A	p.Ser1040Asn	c.3119G>A	p.Ser1040Asn	rs4986852	Non
GM13712	SRR2925414	BRCA1	c.2155_2156insA	p.Lys719delinsLysArgfs	c.2157dupA	p.Glu720fs	rs397508942	MinMMQSDiff
GM13713	SRR2925415	BRCA1	c.3748G>T	p.Glu1250Ter	c.3748G>T	p.Glu1250*	rs28897686	Non
GM13714	SRR2925416	BRCA1	c.5382_5383insC	p.Asn1795Glnfs	c.5382dupC	p.Asn1795fs	rs80357823	MinMMQSDiff
GM13715	SRR2925417	BRCA1	c.5326_5327insC	p.Ser1776delinsSerProfs	c.5329dupC	p.Gln1777fs	rs397507247	Non
GM14090	SRR2925418	BRCA1	c.66_67delAG	p.Leu22_Glu23delinsLeuValfs	c.68_69delAG	p.Glu23fs	rs386833395	Non
GM14092	SRR2925419	BRCA1	c.5201T>C	p.Val1734Ala	c.5201T>C	p.Val1734Ala	rs80357132	Non
GM14093	SRR2925420	BRCA1	c.1204delG	p.Glu402Serfs	c.1204delG	p.Glu402fs	rs80357859	Non
GM14094	SRR2925421	BRCA1	c.1175_1214del40	p.Leu392_Ser405delinsGlnfs	c.1175_1214del40	p.Leu392fs	rs80359874	Non
GM14095	SRR2925422	BRCA1	c.5200delG	p.Val1734Terfs	c.5200delG	p.Val1734fs	rs80357997	Non
GM14096	SRR2925423	BRCA1	c.3481_3491delGAAGATACTAG	p.Glu1161_Ser1164delinsPhefs	c.3481_3491delGAAGATACTAG	p.Glu1161fs	rs80357877	Non
GM14097	SRR2925424	BRCA1	c.181T>G	p.Cys61Gly	c.181T>G	p.Cys61Gly	rs28897672	Non
GM14170	SRR2925425	BRCA2	c.5946delT	p.Ser1982Argfs	c.5946delT	p.Ser1982fs	rs80359550	Non
GM14622	SRR2925426	BRCA2	c.6275_6276delTT	p.Leu2092Profs	c.6275_6276delTT	p.Leu2092fs	rs11571658	Non
GM14623	SRR2925427	BRCA2	c.125A>G	p.Tyr42Cys	c.125A>G	p.Tyr42Cys	rs4987046	Non
GM14624	SRR2925428	BRCA2	c.5718_5719delCT	p.Asn1906_Ser1907 = fs	c.5722_5723delCT	p.Leu1908fs	rs80359530	Non
GM14626	SRR2925429	BRCA2	c.9976A>T	p.Lys3326Ter	c.9976A>T	p.Lys3326*	rs11571833	Non
GM14634	SRR2925430	BRCA1	c.4065_4068delTCAA	p.Asn1355_Gln1356delinsLysfs	c.4065_4068delTCAA	p.Asn1355fs	rs80357508	Non
GM14636	SRR2925431	BRCA1	c.5621_5622insA	p.Tyr1874delinsTerProfs	c.5621dupA	p.Tyr1874fs	rs80357629	Non
GM14637	SRR2925432	BRCA1	c.4327C>T	p.Arg1443Ter	c.4327C>T	p.Arg1443*	rs41293455	Non
GM14638	SRR2925433	BRCA1	c.213-11T>G	-	c.213-11T>G	-	rs80358061	Non
GM14639	SRR2925434	BRCA2	c.6198_6199delTT	p.Val2066_Ser2067delinsValHisfs	c.6198_6199delTT	p.Ser2067fs	rs80359564	Non
GM14684	SRR2925435	BRCA1	c.797_798delTT	p.Val266 = fs	c.798_799delTT	p.Ser267fs	rs80357724	Non

Note : Les variants proviennent de 27 lignées cellulaires du *Coriell Institute*. Strom et collab. ont utilisés les transcrits protéines suivantes, respectivement : *BRCA1* NM_007300.3 et NP_009231.2 ; *BRCA2* NM_000059.3 et NP_000050. Pour le *pipeline*, les transcrits et protéines sont respectivement : *BRCA1* ENST00000471181 et NP_009231.2 ; *BRCA2* ENST00000380152 et NP_000050.

Dans le cas de l'échantillon GM14093 (SRR2925420), la région chr6 : 25,261,169-25,261,974 présente une couverture importante et un variant classé avec un fort impact a été appelé (**Figure 30**). Ce variant n'appartient pas aux régions d'intérêts, ce qui en fait un faux-positif très probable. L'utilisation de filtres limitant aux exons de *BRCA1/2* permet de ne pas retenir cette altération, ce qui diminue le nombre de faux-positifs.



FIGURE 30 – La région chr6 :25,261,169-25,261,974 présente une couverture importante et une délétion a été appelée. Le kit de capture mis au point par Strom et collab. ne capture que les exons de gènes *BRCA1/2*, ce variant est sûrement un faux-positif

3.2 Sensibilité et spécificité de l'appel des variants

La description des échantillons analysés ainsi que les statistiques d'alignement, comparées à celles de Newman et collab., se trouvent dans le **Tableau 4**. Le tri des séquences initiales à l'aide de *Trimmomatic* diminue la couverture des échantillons, sauf pour le SRR1200742.

TABLE 4 – Statistiques de couvertures après tri

Nom échantillon	Description	Newman <i>et al.</i>			<i>Pipeline</i>			Différence profondeur (%)
		Séquences alignées (%)	Séquences alignées avec paire (%)	Profondeur médiane (n)	Séquences alignées (%)	Séquences alignées avec paire (%)	Profondeur médiane (n)	
SRR1197524	Plasma sujet sain	97,5	86,7	8280	99,9	87,9	6578	-20,6
SRR1197554	HCC78 0,025%	98,2	87,0	3913	99,9	87,4	3204	-18,1
SRR1197555	HCC78 0,05%	98,1	86,1	6549	99,9	86,2	5508	-15,9
SRR1197556	HCC78 0,1%	98,4	88,1	6897	99,9	88,4	5694	-17,4
SRR1197557	HCC78 0,5%	98,8	89,8	8096	99,9	90,4	6594	-18,6
SRR1197558	HCC78 1%	98,5	89,8	7779	99,9	90,4	6234	-19,9
SRR1200742	HCC78 100 %	99,0	96,9	7602	100	99,2	8346	9,8

Les données statistiques de couverture (**Figure 35**) montre que l'échantillon SRR1197554 présente une profondeur beaucoup plus faible que les autres échantillons, ce qui suggèrent des problèmes techniques.

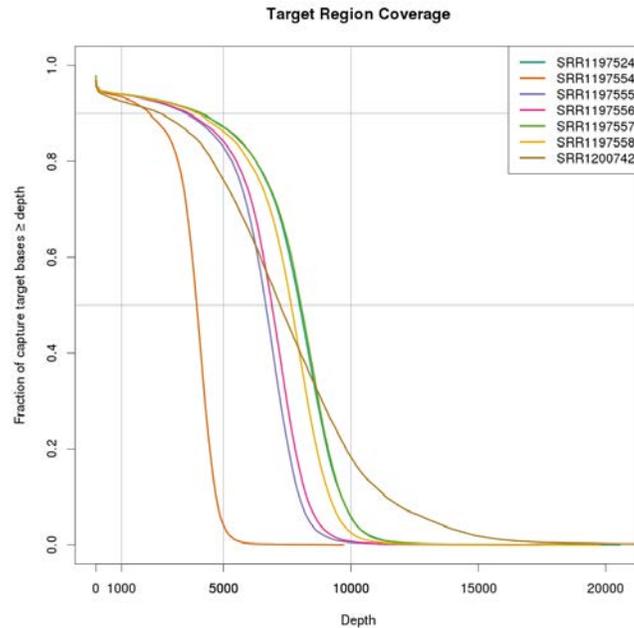


FIGURE 31 – La couverture représente le nombre de séquences alignées sur les cibles. La profondeur représente le nombre de lecture totale de chaque base. Plus celle-ci est élevée et plus les variants avec une faible FAM peuvent être détectés facilement. L'analyse de la profondeur permet de visualiser si la partie analytique s'est déroulée correctement. Ici, des problèmes analytiques de l'échantillon SRR1197554 peuvent être suspectés, sa profondeur médiane étant plus faible que les autres. Figure générée par un *script* R maison

Après utilisation de *Trimmomatic*, les séquençages des dilutions de cellules HCC-78 à 0,1%, 0,05% et 0,025% dans le plasma présentent des profondeurs médianes trop faibles pour un appel de variants avec certitudes par ce *pipeline*. Elles ont été exclues de l'analyse.

Concernant le séquençage par CAPP-Seq de la lignée HCC-78, trois mutations sont en commun avec celles rapportées par le *Broad Institute*. Ces mutations ont toutes été retrouvées (**Tableau 5**). La fusion *ROS1-SLC34A2* a aussi été retrouvée par *Factera* (**Figure 32**) pour la lignée cellulaire pure.

TABLE 5 – Mutations décrites pour la lignée cellulaire HCC-78 capturées par Capp-Seq

Mutation rapportée			HCC-78 100 %			HCC-78 1 %		
Gene	Codon	Protéine	Profondeur (n)	Variants (n)	FAM (%)	Profondeur (n)	Variants (n)	FAM
TP53	c.722C>T	p.Ser241Phe	3576	3575	99,999	5155	14	0,27
CDKN2A	c.172C>T	p.Arg58*	77	77	100	2073	9	0,675
SI	c.1582C>T	p.Pro528Ser	4536	42	0,926	6580	6	0*
U2AF1	c.101C>T	p.Ser34Phe	4106	1210	29,469	5146	8	0*

Note : * Variants non appelés par *varscan2*

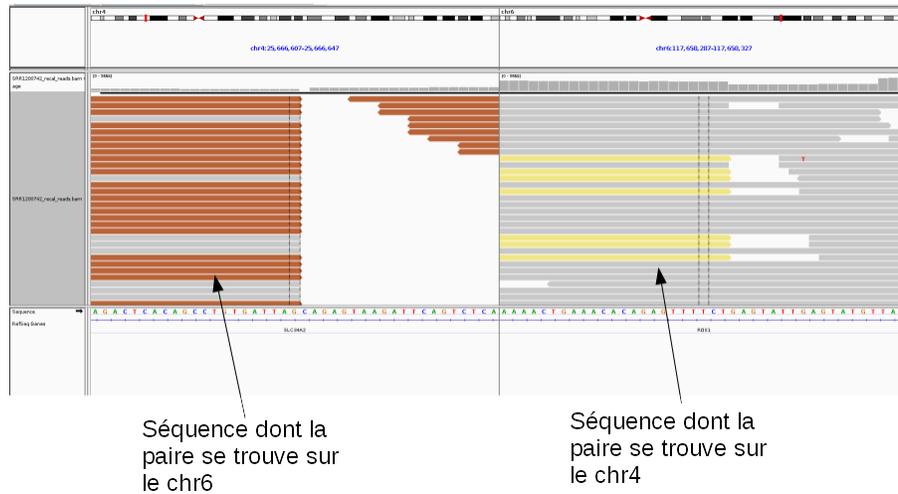


FIGURE 32 – La fusion entre les gènes *ROS1-SLC34A2* a été mise en évidence par *Factera* et visualisée à l’aide du logiciel *Igv*.

La détermination de la fraction allélique retrouvée par le pipeline *versus* celles obtenues par les dilutions se trouve dans la **Figure 33**. Avec l’usage de *Trimmomatic*, la sensibilité de l’analyse est estimée à 0,5%.

3.2.1 Sans tri des séquences par *Trimmomatic* et appel par *VarDict*

Ce tri, qui permet d’enlever les séquences de mauvaise qualité, diminue la couverture (**Tableau 4**) mais aussi les variations techniques. La non-utilisation de *Trimmomatic* permet d’augmenter les profondeurs (**Figure 34**, les couvertures avec tri se trouvent en **Figure 35**).

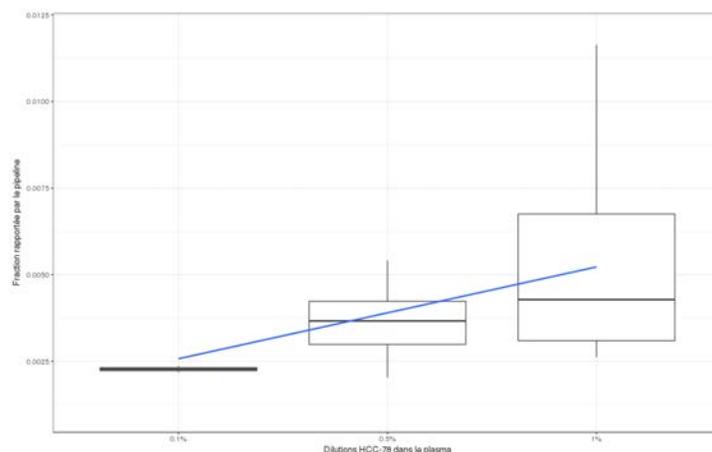


FIGURE 33 – Ce graphique représente les FAM rapportées par le *pipeline* versus les dilutions de HCC-78 dans le plasma. Figure générée par un *script* R maison

TABLE 6 – Précision, rappel et scores F1 des variants appelés par *VarDict*, sans tri

Dilution lignée HCC-78	Sans limitation de sensibilité			Sensibilité limitée à 0,25 %		
	Rappel	Précision	Score F1	Rappel	Précision	Score F1
1%	0,26 (113/442)	0,01 (113/8539)	0,025	0,22 (99/442)	0,03 (99/3254)	0,05
0,5%	0,16 (71/442)	0,02 (71/3580)	0,04	0,11 (48/442)	0,06 (48/765)	0,08
0,1%	0,07 (33/442)	0,008 (33/3918)	0,02			
0,05%	0,05 (22/442)	0,005 (22/4515)	0,009			
0,025%	0,02 (11/442)	0,002 (11/4809)	0,004			

Les scores F1 obtenus sont bas, à cause d’un bruit de fond élevé (notamment pour l’échantillon correspondant à une dilution à 1% de lignée HCC-78) (**Tableau 6**). La limitation à une sensibilité de 0,25% augmente les scores F1 (**Tableau 6**).

4 Discussion

L’analyse de l’ADNtc suscite de grandes attentes pour la médecine personnalisée des cancers. L’augmentation de la sensibilité des méthodes permet de sélectionner de façon adéquate les patients qui nécessitent une chimiothérapie adjuvante, de détecter les mutations prédictives de réponse ou de résistances ainsi que d’évaluer la charge tumorale. Les méthodes en pauci-*loci* présentent de grande sensibilité (FAM de 0,01%) et sont peu sujettes aux faux-positifs. En revanche, elles ne peuvent pas détecter des mutations *de novo*. Pouvant interroger un grand nombre de points, les analyses se basant sur le séquençage haut-débit peuvent détecter plus

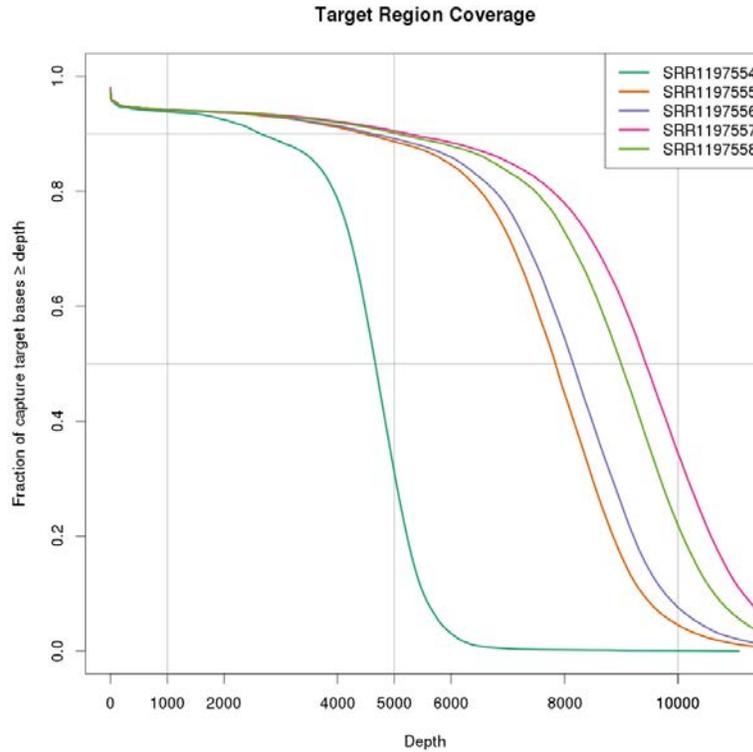


FIGURE 34 – Sans tri par *Trimmomatic* des *reads* initiaux, les profondeurs sont plus grandes, ce qui permet d’augmenter la sensibilité de l’appel des variants. Figure générée par un *script* R maison

de variants (par exemple, quatre altérations somatiques sont détectées en médiane par CAPP-Seq par échantillon), ce qui augmente la sensibilité de détection d’une maladie résiduelle (NEWMAN et collab., 2014b). De même, les mutations de résistance non récurrentes peuvent être mise en évidence. Mais ces méthodes nécessitent une plate-forme technique conséquente, du personnel compétent et un traitement bioinformatique des données (WAN et collab., 2017).

Les annotations et la priorisation des variants par le *pipeline* mis en place dans ce travail présentent une grande concordance avec les résultats de l’étude de Strom et collab. Cela permet un gain de temps considérable pour les personnes en charge de la validation. La sensibilité a été évaluée à 0,5% pour l’analyse par capture. Des mutations d’intérêt jusqu’à la dilution de 0,025% ont pu être détectées mais sans évaluation précise du bruit de fond, il est difficile de les différencier des faux-

positifs. La capture par hybridation présente de nombreux avantages par rapport aux méthodes par amplicons, par exemple une plus grande flexibilité ou une couverture uniforme (NEWMAN et collab., 2016). La grande souplesse quant au design des panels permet d'ajouter de créer des panels adaptés aux attentes cliniques. De plus, la capture permet de détecter les gènes de fusion de manière plus aisée que les amplicons. Enfin, la couverture relativement uniforme des cibles favorise la détection des VNC.

Néanmoins, des faux-positifs peuvent être induits par ces méthodes. Les transversions G>T représentent la majorité des erreurs (NEWMAN et collab., 2016). Des transitions C>T ou G>A sont aussi notés. L'hypothèse émise pour expliquer ces erreurs techniques sont des dommages d'oxydation qui conduisent à la formation de 8-oxoguanine et à la déamination des cytosines (COSTELLO et collab., 2013; CHEN et collab., 2014). Ces artefacts proviendraient de la capture. En faisant varier les temps de capture jusqu'à trois jours, Newman et collab. ont montré une augmentation des erreurs G>T et C>A (NEWMAN et collab., 2016). Les polymérase sont aussi une source d'erreur, avec un taux à 0,01% (DIAZ JR et BARDELLI, 2014). Des erreurs peuvent aussi être présentes lors de séquençage des nucléotides. Dans le cas de la technologie Illumina[®], une forte corrélation entre les intensités de fluorescence pour les adénosines et cytosines tout comme entre guanosines et thymines sont décrites (SCHIRMER et collab., 2015). De même, les spectres de fluorescence entre ces nucléotides sont proches, les filtres utilisés peuvent ne pas séparer efficacement les signaux (SCHIRMER et collab., 2015). Tous ces artefacts entraînent un bruit de fond, qui entraînent des faux-positifs surtout lors de la recherche des mutations avec une faible FAM.

Les résultats de ce travail montrent qu'un *pipeline* d'analyse de données est critique et des procédures de validation commencent à être émises (ROY et collab., 2017). En effet, les choix effectués impactent la sensibilité et la spécificité des résultats. Dans ce travail, il a été préféré d'analyser les performances du *pipeline* sur des mutations réelles, comme recommandé par l'étude de l'équipe de Gut (ALIOTO et collab., 2015). Néanmoins, il est préférable d'évaluer la sensibilité par rapport à des échantillons connus comme étant avec des altérations d'une faible FAM plutôt que par dilution d'une lignée cellulaire (ROY et collab., 2017). Le tri des séquences par *Trimomatic*, s'il provoque une diminution de la couverture et *in fine* de la sensibilité, permet de diminuer le bruit de fond. Les différents logiciels pour appeler les variants influent aussi sur le résultat, d'un point de vue sensibilité et spécificité. Les résultats de ce travail montrent de meilleures performances de *VarDict* par rapport à *VarScan2* pour les variants avec une faible FAM. Cette observation concorde avec des résultats de la littérature : Sandmann et collab décrivent de meilleures performances pour

VarDict (SANDMANN et collab., 2017). Néanmoins, les mutations avec une très faible FAM s'avèrent difficile à différencier de la variance technique par l'approche présentée. Les variants complexes, tels que les VNC et VS, n'ont pas été évalués au sein de ce travail. En effet, pour les VNC, il est fortement recommandé d'avoir des échantillons témoins afin d'obtenir de meilleurs résultats. Pour les VS, aucun n'a été décrit dans la base COSMIC (<https://cancer.sanger.ac.uk/cosmic/search?q=HCC78>, accédé en avril 2018).

Plusieurs alternatives ont été mises au point afin de diminuer le nombre de faux-positifs en séquençage haut débit. D'abord, des approches par de type *barcoding* comme Safe-SeqS, décrites précédemment dans ce travail (KINDE et collab., 2011). Ces kits rajoutent néanmoins un coût à l'analyse. Ensuite, le bruit de fond a été calculé par approche statistique (NEWMAN et collab., 2014b; PÉCUCHE et collab., 2016b). Ces méthodes se basent sur un panel de plasmas de sujets sains. Ces derniers sont analysés par le même kit utilisé et avec une couverture similaire des plasmas de patients. À l'exclusion des polymorphismes, le nombre d'allèles alternatives à chaque position est décompté (NEWMAN et collab., 2014b; PÉCUCHE et collab., 2016b). Cela permet de construire une base de donnée qui contient le bruit de fond spécifique pour l'ensemble des localisations génomiques du kit (NEWMAN et collab., 2014b; PÉCUCHE et collab., 2016a). Enfin, un test binomial est utilisé pour différencier les mutations des erreurs techniques pour les patients (NEWMAN et collab., 2014b; PÉCUCHE et collab., 2016a). Cette approche permet d'augmenter la sensibilité tout en diminuant le nombre de faux-positifs. Newman et collab. rapportent une sensibilité à 0,02%, tandis qu'elle a été évaluée à 0,003% pour les mutations en point unique et 0,001% pour les indels par l'équipe de Laurent-Puig (NEWMAN et collab., 2014b; PÉCUCHE et collab., 2016a). Des approches combinées *barcoding* et correction du bruit de fond. C'est le cas de la méthode *integrated digital error suppression* (iDES), qui améliore les résultats de l'approche CAPP-Seq par un facteur 15 (NEWMAN et collab., 2016). L'analyse par ADNtc des mutations *EGFR* d'une petite cohorte de patients atteints de CPNPC a montré une sensibilité de 92% et une spécificité de 96% et une détection d'une allèle mutée pour 25000 allèles sauvages (NEWMAN et collab., 2016).

Les évolutions techniques pourraient permettre d'utiliser l'ADNtc comme test compagnon (WAN et collab., 2017). Cela favoriserait la détermination du statut mutationnel du patient au travers d'un prélèvement non-invasif et le suivi de la réponse (WAN et collab., 2017). Cet aspect est essentiel dans les thérapies ciblées : de nombreuses mutations de résistances proviennent de sous-clones (VENKATESAN et SWANTON, 2016; JAMAL-HANJANI et collab., 2017). Ces derniers, moins adaptés à

l'environnement que les cellules sensibles, deviennent majoritaires avec la pression de sélection induit par les traitements (VENKATESAN et SWANTON, 2016). Deux stratégies ont été mises au point afin de réduire l'effet de cette évolution sur la clinique du patient. D'abord la thérapie adaptative qui se base sur l'évolution de la masse tumorale (VENKATESAN et SWANTON, 2016). Dans cette approche, les plus hautes doses sont administrées durant les phases de croissance et les plus faibles lors de la régression (VENKATESAN et SWANTON, 2016). Cela permet de stabiliser la charge tumorale et d'empêcher les sous-clones résistants d'atteindre une proportion significative (VENKATESAN et SWANTON, 2016). Cette méthode a été utilisée dans les mélanomes traités par anti-*BRAF* (SUN et collab., 2014). L'ADNtc permet de connaître avec précision la charge tumorale ainsi que la maladie résiduelle, constituant ainsi un biomarqueur de choix pour l'adaptation d'un traitement par cette stratégie (WAN et collab., 2017). Ensuite, une approche par thérapie séquentielle peut être utilisée. Au sein de cette méthode, les molécules sont utilisées en fonction des résistances détectées (VENKATESAN et SWANTON, 2016). Par exemple, dans les CPNPC, les anti-*EGFR* de première génération (erlotinib ou gefitinib) provoquent l'émergence de mutation *EGFR*^{t790m} (VENKATESAN et SWANTON, 2016). Ces dernières sont sensibles aux anti-*EGFR* de troisième génération (rociletinib, osimertinib), un changement pour ces molécules peut être proposé. Le monitoring par l'ADNtc permettrait de suivre l'acquisition des mutations de résistance et de proposer un changement de thérapie de manière rapide (WAN et collab., 2017).

Néanmoins, aucun seuil de FAM n'a été défini, ce qui peut induire un risque de sur-diagnostic si l'ADNtc provient d'un sous-clone, pour les thérapies ciblées ou le pronostic (MERKER et collab., 2018). Plusieurs études sur biopsies tumorales ont évalués l'impact de la présence de sous-clones sur le bénéfice d'une thérapie ou le pronostic. Dans le cas des cancers colorectaux, une étude rétrospective a montré que les patients mutés *KRAS* avec une FAM $\leq 1\%$ avaient les mêmes bénéfices d'un traitement par anti-*EGFR* comme les patients *KRAS* sauvage (LAURENT-PUIG et collab., 2014). Dans le cas des SMD, les patients porteur d'une mutation de *TP53* avec une FAM $> 6\%$ avait une survie plus faible que ceux avec une FAM $< 6\%$ (43,5 mois *vs.* 138 mois, HR = 12,2, $p = 0,003$) (BELICKOVA et collab., 2016). Ces travaux montrent l'intérêt de définir des seuils, afin d'améliorer la prise en charge des patients. Enfin, il est nécessaire de connaître de manière plus précise les cellules à l'origine des ADNtc. En effet, Charles Swaton et collab. estiment que pour être efficace, une thérapie doit cibler des altérations présentes dès l'origine des lésions (JAMAL-HANJANI et collab., 2017). Dans les CPNPC, les premiers résultats du projet TracerX ont montré que les mutations touchant des oncogènes n'apparaissent pas au même stade d'évolution (JAMAL-HANJANI et collab., 2017). *EGFR*, *MET* proto-oncogene, receptor tyrosine

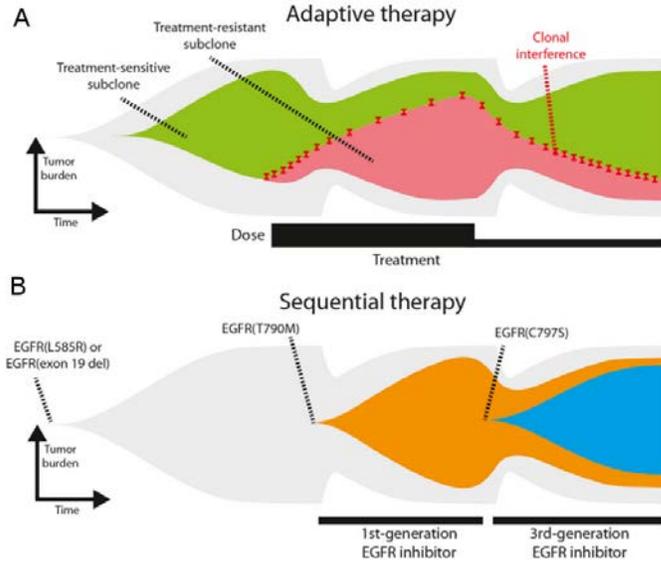


FIGURE 35 – **a.** À l’initiation de la thérapie, de fortes doses peuvent être utilisées ce qui conduit à l’émergence d’un sous-clone de résistance. Afin d’empêcher ce dernier d’atteindre une taille importante, les posologies sont diminuées ce qui décroît l’interférence clonale. **b.** L’adaptation séquentielle fait appel au suivi des mutations de résistance. Les thérapies ciblées sont choisies en fonction des clones émergents. Le principal risque de cette stratégie est un échappement thérapeutique, à cause de la sélection d’un clone porteur d’une mutation de résistance *de novo* et résistant aux thérapies ciblées. D’après VENKATESAN et SWANTON (2016).

kinase (*MET*) ou encore *BRAF* sont des gènes dont les mutations apparaissent tôt et sont retrouvés dans chacun des clones (JAMAL-HANJANI et collab., 2017). Au contraire, *PIK3CA*, *NF1*, *KRAS*, *TP53* ou *NOTCH* sont des altérations présentes dans des sous-clones (JAMAL-HANJANI et collab., 2017). Si l’ADNtc permet d’apprécier l’hétérogénéité intra et inter-tumorale, il s’avère difficile de déterminer si une mutation provient du tronc ou de la clade. Il s’avère aussi nécessaire de déterminer si le suivi du patient doit s’effectuer au travers d’une recherche de quelques mutations ou plus largement. En effet, les méthodes pauci-*loci* s’avèrent plus rapides que les recherches plus étendues. Cela permettrait une adaptation des traitements plus précoce et surtout pour un coût moindre. Néanmoins ces méthode ne permettent de suivre qu’un nombre limité de mutations, ce qui est à risque de faux-négatifs notamment pour les mutations *de novo* (WAN et collab., 2017). De plus, la corrélation avec

la charge tumorale s'avère meilleure avec plusieurs mutations au lieu qu'une seule (NEWMAN et collab., 2014b). Tous ces impacts doivent être évalués par des études prospectives à grande échelle avant que l'ADNtc soit utilisé en pratique courante.

Références

- ADZHUBEI, I. A., S. SCHMIDT, L. PESHKIN, V. E. RAMENSKY, A. GERASIMOVA, P. BORK, A. S. KONDRASHOV et S. R. SUNYAEV. 2010, «A method and server for predicting damaging missense mutations», *Nature methods*, vol. 7, n° 4, p. 248.
- ALIOTO, T. S., I. BUCHHALTER, S. DERDAK, B. HUTTER, M. D. ELDRIDGE, E. HOVIG, L. E. HEISLER, T. A. BECK, J. T. SIMPSON, L. TONON et collab.. 2015, «A comprehensive assessment of somatic mutation detection in cancer using whole-genome sequencing», *Nature communications*, vol. 6, p. 10 001.
- ALIZADEH, A. A., M. B. EISEN, R. E. DAVIS, C. MA, I. S. LOSSOS, A. ROSENWALD, J. C. BOLDRICK, H. SABET, T. TRAN, X. YU et collab.. 2000, «Distinct types of diffuse large b-cell lymphoma identified by gene expression profiling», *Nature*, vol. 403, n° 6769, p. 503.
- ALMOGUERA, C., D. SHIBATA, K. FORRESTER, J. MARTIN, N. ARNHEIM et M. PERUCHO. 1988, «Most human carcinomas of the exocrine pancreas contain mutant *ck-ras* genes», *Cell*, vol. 53, n° 4, p. 549–554.
- ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS et D. J. LIPMAN. 1990, «Basic local alignment search tool», *Journal of molecular biology*, vol. 215, n° 3, p. 403–410.
- ANDREWS, S. et collab.. 2010, «Fastqc : a quality control tool for high throughput sequence data», .
- ANKER, P., J. LYAUTEY, F. LEFORT, C. LEDERREY et M. STROUN. 1994, «Transformation de cellules nih», *Comptes rendus de l'Académie des sciences. Série 3, Sciences de la vie*, vol. 317, n° 10, p. 869–874.
- VAN DER AUWERA, G. A., M. O. CARNEIRO, C. HARTL, R. POPLIN, G. DEL ANGEL, A. LEVY-MOONSHINE, T. JORDAN, K. SHAKIR, D. ROAZEN, J. THIBAUTL et collab.. 2013, «From fastq data to high-confidence variant calls : the genome analysis toolkit best practices pipeline», *Current protocols in bioinformatics*, p. 11–10.
- BALLEHANINNA, U. K. et R. S. CHAMBERLAIN. 2012, «The clinical utility of serum ca 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma : An evidence based appraisal», *Journal of gastrointestinal oncology*, vol. 3, n° 2, p. 105.

- BAUDINO, T. A. 2015, «Targeted cancer therapy : the next generation of cancer treatment», *Current drug discovery technologies*, vol. 12, n° 1, p. 3–20.
- BEAVER, J. A., D. JELOVAC, S. BALUKRISHNA, R. L. COCHRAN, S. CROESSMANN, D. J. ZABRANSKY, H. Y. WONG, P. V. TORO, J. CIDADO, B. G. BLAIR et collab.. 2014, «Detection of cancer dna in plasma of patients with early-stage breast cancer», *Clinical cancer research*, vol. 20, n° 10, p. 2643–2650.
- BELICKOVA, M., J. VESELA, A. JONASOVA, B. PEJSOVA, H. VOTAVOVA, M. D. MERKEROVA, Z. ZEMANOVA, J. BREZINOVA, D. MIKULENKOVA, M. LAUERMANNOVA et collab.. 2016, «Tp53 mutation variant allele frequency is a potential predictor for clinical outcome of patients with lower-risk myelodysplastic syndromes», *Oncotarget*, vol. 7, n° 24, p. 36 266.
- BENZEKRY, S., C. LAMONT, A. BEHESHTI, A. TRACZ, J. M. EBOS, L. HLATKY et P. HAHNFELDT. 2014, «Classical mathematical models for description and prediction of experimental tumor growth», *PLoS computational biology*, vol. 10, n° 8, p. e1003 800.
- BETTEGOWDA, C., M. SAUSEN, R. J. LEARY, I. KINDE, Y. WANG, N. AGRAWAL, B. R. BARTLETT, H. WANG, B. LUBER, R. M. ALANI et collab.. 2014, «Detection of circulating tumor dna in early-and late-stage human malignancies», *Science translational medicine*, vol. 6, n° 224, p. 224ra24–224ra24.
- BOLGER, A. M., M. LOHSE et B. USADEL. 2014, «Trimmomatic : a flexible trimmer for illumina sequence data», *Bioinformatics*, vol. 30, n° 15, p. 2114–2120.
- BONI, L., E. CASSINOTTI, M. CANZIANI, G. DIONIGI, F. ROVERA et R. DIONIGI. 2007, «Free circulating dna as possible tumour marker in colorectal cancer», *Surgical oncology*, vol. 16, p. 29–31.
- BRENNER, H., M. KLOOR et C. P. POX. 2014, «Colorectal cancer», *Lancet*, vol. 383, n° 9927, p. 1490–1502.
- CALAPRE, L., L. WARBURTON, M. MILLWARD, M. ZIMAN et E. S. GRAY. 2017, «Circulating tumour dna (ctdna) as a liquid biopsy for melanoma», *Cancer letters*, vol. 404, p. 62–69.
- CAMPBELL, P. J., S. YACHIDA, L. J. MUDIE, P. J. STEPHENS, E. D. PLEASANCE, L. A. STEBBINGS, L. A. MORSBERGER, C. LATIMER, S. MCLAREN, M.-L. LIN et collab.. 2010, «The patterns and dynamics of genomic instability in metastatic pancreatic cancer», *Nature*, vol. 467, n° 7319, p. 1109.

- CARDOSO, F., L. J. VAN'T VEER, J. BOGAERTS, L. SLAETS, G. VIALE, S. DELALOGUE, J.-Y. PIERGA, E. BRAIN, S. CAUSERET, M. DELORENZI et collab.. 2016, «70-gene signature as an aid to treatment decisions in early-stage breast cancer», *New England Journal of Medicine*, vol. 375, n° 8, p. 717–729.
- CHAN, K., E. C. HUNG, J. K. WOO, P. K. CHAN, S.-F. LEUNG, F. LAI, A. S. CHENG, S. W. YEUNG, Y. W. CHAN, T. K. TSUI et collab.. 2013, «Early detection of nasopharyngeal carcinoma by plasma epstein-barr virus dna analysis in a surveillance program», *Cancer*, vol. 119, n° 10, p. 1838–1844.
- CHAO, M. et P. GIBBS. 2009, «Caution is required before recommending routine carcinoembryonic antigen and imaging follow-up for patients with early-stage colon cancer», *Journal of Clinical Oncology*, vol. 27, n° 36, p. e279–e280.
- CHEN, G., S. MOSIER, C. D. GOCKE, M.-T. LIN et J. R. ESHLEMAN. 2014, «Cytosine deamination is a major cause of baseline noise in next-generation sequencing», *Molecular diagnosis & therapy*, vol. 18, n° 5, p. 587–593.
- CHIK, F., M. SZYF et S. A. RABBANI. 2011, «Role of epigenetics in cancer initiation and progression», dans *Human Cell Transformation*, Springer, p. 91–104.
- CHURCH, T. R., M. WANDELL, C. LOFTON-DAY, S. J. MONGIN, M. BURGER, S. R. PAYNE, E. CASTAÑOS-VÉLEZ, B. A. BLUMENSTEIN, T. RÖSCH, N. OSBORN et collab.. 2014, «Prospective evaluation of methylated sept9 in plasma for detection of asymptomatic colorectal cancer», *Gut*, vol. 63, n° 2, p. 317–325.
- CINGOLANI, P., A. PLATTS, L. L. WANG, M. COON, T. NGUYEN, L. WANG, S. J. LAND, X. LU et D. M. RUDEN. 2012, «A program for annotating and predicting the effects of single nucleotide polymorphisms, snpeff : Snps in the genome of drosophila melanogaster strain w1118 ; iso-2 ; iso-3», *Fly*, vol. 6, n° 2, p. 80–92.
- COHEN, P. A., N. FLOWERS, S. TONG, N. HANNAN, M. D. PERTILE et L. HUI. 2016, «Abnormal plasma dna profiles in early ovarian cancer using a non-invasive prenatal testing platform : implications for cancer screening», *BMC medicine*, vol. 14, n° 1, p. 126.
- COOPER, G. M. et J. SHENDURE. 2011, «Needles in stacks of needles : finding disease-causal variants in a wealth of genomic data», *Nature Reviews Genetics*, vol. 12, n° 9, p. 628.

- CORTESE, R., I. ALMENDROS, Y. WANG et D. GOZAL. 2015, «Tumor circulating dna profiling in xenografted mice exposed to intermittent hypoxia», *Oncotarget*, vol. 6, n° 1, p. 556.
- COSTELLO, M., T. J. PUGH, T. J. FENNEL, C. STEWART, L. LICHTENSTEIN, J. C. MELDRIM, J. L. FOSTEL, D. C. FRIEDRICH, D. PERRIN, D. DIONNE et collab.. 2013, «Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative dna damage during sample preparation», *Nucleic acids research*, vol. 41, n° 6, p. e67–e67.
- CROWLEY, E., F. DI NICOLANTONIO, F. LOUPAKIS et A. BARDELLI. 2013, «Liquid biopsy : monitoring cancer-genetics in the blood», *Nature reviews Clinical oncology*, vol. 10, n° 8, p. 472–484.
- DAWSON, S.-J., D. W. TSUI, M. MURTAZA, H. BIGGS, O. M. RUEDA, S.-F. CHIN, M. J. DUNNING, D. GALE, T. FORSHEW, B. MAHLER-ARAUJO et collab.. 2013, «Analysis of circulating tumor dna to monitor metastatic breast cancer», *New England Journal of Medicine*, vol. 368, n° 13, p. 1199–1209.
- DIAZ JR, L. A. et A. BARDELLI. 2014, «Liquid biopsies : genotyping circulating tumor dna», *Journal of clinical oncology*, vol. 32, n° 6, p. 579–586.
- DIAZ JR, L. A., R. T. WILLIAMS, J. WU, I. KINDE, J. R. HECHT, J. BERLIN, B. ALLEN, I. BOZIC, J. G. REITER, M. A. NOWAK et collab.. 2012, «The molecular evolution of acquired resistance to targeted egfr blockade in colorectal cancers», *Nature*, vol. 486, n° 7404, p. 537.
- DIEHL, F., M. LI, D. DRESSMAN, Y. HE, D. SHEN, S. SZABO, L. A. DIAZ, S. N. GOODMAN, K. A. DAVID, H. JUHL et collab.. 2005, «Detection and quantification of mutations in the plasma of patients with colorectal tumors», *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, n° 45, p. 16 368–16 373.
- DIEHL, F., M. LI, Y. HE, K. W. KINZLER, B. VOGELSTEIN et D. DRESSMAN. 2006, «Beaming : single-molecule pcr on microparticles in water-in-oil emulsions», *Nature methods*, vol. 3, n° 7, p. 551.
- DIEHL, F., K. SCHMIDT, M. A. CHOTI, K. ROMANS, S. GOODMAN, M. LI, K. THORNTON, N. AGRAWAL, L. SOKOLL, S. A. SZABO et collab.. 2008, «Circulating mutant dna to assess tumor dynamics», *Nature medicine*, vol. 14, n° 9, p. 985.

- DRESSMAN, D., H. YAN, G. TRAVERSO, K. W. KINZLER et B. VOGELSTEIN. 2003, «Transforming single dna molecules into fluorescent magnetic particles for detection and enumeration of genetic variations», *Proceedings of the National Academy of Sciences*, vol. 100, n° 15, p. 8817–8822.
- DUFFY, M. 2001, «Clinical uses of tumor markers : a critical review», *Critical reviews in clinical laboratory sciences*, vol. 38, n° 3, p. 225–262.
- EARLY BREAST CANCER TRIALISTS' COLLABORATIVE GROUP AND OTHERS. 2012, «Comparisons between different polychemotherapy regimens for early breast cancer : meta-analyses of long-term outcome among 100 000 women in 123 randomised trials», *The Lancet*, vol. 379, n° 9814, p. 432–444.
- EL MESSAOUDI, S., F. ROLET, F. MOULIERE et A. R. THIERRY. 2013, «Circulating cell free dna : preanalytical considerations», *Clinica Chimica Acta*, vol. 424, p. 222–230.
- FORSHEW, T., M. MURTAZA, C. PARKINSON, D. GALE, D. W. TSUI, F. KAPER, S.-J. DAWSON, A. M. PISKORZ, M. JIMENEZ-LINAN, D. BENTLEY et collab.. 2012, «Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma dna», *Science translational medicine*, vol. 4, n° 136, p. 136ra68–136ra68.
- FRANGIONI, J. V. 2008, «New technologies for human cancer imaging», *Journal of clinical oncology*, vol. 26, n° 24, p. 4012–4021.
- FROMER, M., J. L. MORAN, K. CHAMBERT, E. BANKS, S. E. BERGEN, D. M. RUDERFER, R. E. HANDSAKER, S. A. MCCARROLL, M. C. O'DONOVAN, M. J. OWEN et collab.. 2012, «Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth», *The American Journal of Human Genetics*, vol. 91, n° 4, p. 597–607.
- FUJIWARA, K., N. FUJIMOTO, M. TABATA, K. NISHII, K. MATSUO, K. HOTTA, T. KOZUKI, M. AOE, K. KIURA, H. UEOKA et collab.. 2005, «Identification of epigenetic aberrant promoter methylation in serum dna is useful for early detection of lung cancer», *Clinical Cancer Research*, vol. 11, n° 3, p. 1219–1225.
- GAHAN, P. B. et M. STROUN. 2010, «The virtosome—a novel cytosolic informative entity and intercellular messenger», *Cell biochemistry and function*, vol. 28, n° 7, p. 529–538.

- GALE, D., V. PLAGNOL, A. LAWSON, M. PUGH, S. SMALLEY, K. HOWARTH, M. MADI, B. DURHAM, V. KUMANDURI, K. LO et collab.. 2016, «Analytical performance and validation of an enhanced tam-seq circulating tumor dna sequencing assay», .
- GARCIA-MURILLAS, I., G. SCHIAVON, B. WEIGELT, C. NG, S. HREBIEN, R. J. CUTTS, M. CHEANG, P. OSIN, A. NERURKAR, I. KOZAREWA et collab.. 2015, «Mutation tracking in circulating tumor dna predicts relapse in early breast cancer», *Science translational medicine*, vol. 7, n° 302, p. 302ra133–302ra133.
- GARCÍA-OLMO, D. C., C. DOMÍNGUEZ, M. GARCÍA-ARRANZ, P. ANKER, M. STROUN, J. M. GARCÍA-VERDUGO et D. GARCÍA-OLMO. 2010, «Cell-free nucleic acids circulating in the plasma of colorectal cancer patients induce the oncogenic transformation of susceptible cultured cells», *Cancer research*, vol. 70, n° 2, p. 560–567.
- GENOVESE, G., A. K. KÄHLER, R. E. HANDSAKER, J. LINDBERG, S. A. ROSE, S. F. BAKHOUM, K. CHAMBERT, E. MICK, B. M. NEALE, M. FROMER et collab.. 2014, «Clonal hematopoiesis and blood-cancer risk inferred from blood dna sequence», *New England Journal of Medicine*, vol. 371, n° 26, p. 2477–2487.
- GERLINGER, M., A. J. ROWAN, S. HORSWELL, J. LARKIN, D. ENDESFELDER, E. GRONROOS, P. MARTINEZ, N. MATTHEWS, A. STEWART, P. TARPEY et collab.. 2012, «Intratumor heterogeneity and branched evolution revealed by multi-region sequencing», *New England journal of medicine*, vol. 366, n° 10, p. 883–892.
- GILLIES, R. J., D. VERDUZCO et R. A. GATENBY. 2012, «Evolutionary dynamics of carcinogenesis and why targeted therapy does not work», *Nature Reviews Cancer*, vol. 12, n° 7, p. 487.
- GORMALLY, E., P. VINEIS, G. MATULLO, F. VEGLIA, E. CABOUX, E. LE ROUX, M. PELUSO, S. GARTE, S. GUARRERA, A. MUNNIA et collab.. 2006, «Tp53 and kras2 mutations in plasma dna of healthy subjects and subsequent cancer occurrence : a prospective study», *Cancer research*, vol. 66, n° 13, p. 6871–6876.
- GRAY, E. S., H. RIZOS, A. L. REID, S. C. BOYD, M. R. PEREIRA, J. LO, V. TEMBE, J. FREEMAN, J. H. LEE, R. A. SCOLYER et collab.. 2015, «Circulating tumor dna to monitor treatment response and detect acquired resistance in patients with metastatic melanoma», *Oncotarget*, vol. 6, n° 39, p. 42008.

- GREAVES, M. et C. C. MALEY. 2012, «Clonal evolution in cancer», *Nature*, vol. 481, n° 7381, p. 306.
- GUERRERO-PRESTON, R., B. L. VALLE, A. JEDLICKA, N. TURAGA, O. FOLAWIYO, F. PIRINI, F. LAWSON, A. VERGURA, M. NOORDHUIS, A. DZIEDZIC et collab.. 2016, «Molecular triage of premalignant lesions in liquid-based cervical cytology and circulating cell-free dna from urine, using a panel of methylated human papilloma virus and host genes», *Cancer Prevention Research*, vol. 9, n° 12, p. 915–924.
- GUO, S., D. DIEP, N. PLONGTHONGKUM, H.-L. FUNG, K. ZHANG et K. ZHANG. 2017, «Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma dna», *Nature genetics*, vol. 49, n° 4, p. 635.
- HEWITSON, P., P. GLASZIOU, E. WATSON, B. TOWLER et L. IRWIG. 2008, «Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult) : an update», *The American journal of gastroenterology*, vol. 103, n° 6, p. 1541–1549.
- HOHAUS, S., M. GIACHELIA, G. MASSINI, G. MANSUETO, B. VANNATA, V. BOZZOLI, M. CRISCUOLO, F. D'ALÒ, M. MARTINI, L. M. LAROCCA et collab.. 2009, «Cell-free circulating dna in hodgkin's and non-hodgkin's lymphomas», *Annals of oncology*, vol. 20, n° 8, p. 1408–1413.
- HOLDENRIEDER, S. et P. STIEBER. 2009, «Clinical use of circulating nucleosomes», *Critical reviews in clinical laboratory sciences*, vol. 46, n° 1, p. 1–24.
- HOLMES, A., S. LAMEIRAS, E. JEANNOT, Y. MARIE, L. CASTERA, X. SASTRE-GARAU et A. NICOLAS. 2016, «Mechanistic signatures of hpv insertions in cervical carcinomas», *NPJ Genomic Medicine*, vol. 1, p. 16 004.
- HOWLADER, N., S. F. ALTEKRUSE, C. I. LI, V. W. CHEN, C. A. CLARKE, L. A. RIES et K. A. CRONIN. 2014, «Us incidence of breast cancer subtypes defined by joint hormone receptor and her2 status», *JNCI : Journal of the National Cancer Institute*, vol. 106, n° 5.
- INCA. 2017, «Séquençage de nouvelle génération d'un panel de gènes pour l'analyse en génétique somatique, guide pour la validation de méthode», .

- ITO, K., K. HIBI, H. ANDO, K. HIDEMURA, T. YAMAZAKI, S. AKIYAMA et A. NAKAO. 2002, «Usefulness of analytical cea doubling time and half-life time for overlooked synchronous metastases in colorectal carcinoma», *Japanese journal of clinical oncology*, vol. 32, n° 2, p. 54–58.
- JAHR, S., H. HENTZE, S. ENGLISCH, D. HARDT, F. O. FACKELMAYER, R.-D. HESCH et R. KNIPPERS. 2001, «Dna fragments in the blood plasma of cancer patients : quantitations and evidence for their origin from apoptotic and necrotic cells», *Cancer research*, vol. 61, n° 4, p. 1659–1665.
- JAMAL-HANJANI, M., G. A. WILSON, N. MCGRANAHAN, N. J. BIRKBAK, T. B. WATKINS, S. VEERIAH, S. SHAFI, D. H. JOHNSON, R. MITTER, R. ROSENTHAL et collab.. 2017, «Tracking the evolution of non–small-cell lung cancer», *New England Journal of Medicine*, vol. 376, n° 22, p. 2109–2121.
- JIAN, G., Z. SONGWEN, Z. LING, D. QINFANG, Z. JIE, T. LIANG et Z. CAICUN. 2010, «Prediction of epidermal growth factor receptor mutations in the plasma/pleural effusion to efficacy of gefitinib treatment in advanced non-small cell lung cancer», *Journal of cancer research and clinical oncology*, vol. 136, n° 9, p. 1341–1347.
- JIANG, P., C. W. CHAN, K. A. CHAN, S. H. CHENG, J. WONG, V. W.-S. WONG, G. L. WONG, S. L. CHAN, T. S. MOK, H. L. CHAN et collab.. 2015, «Lengthening and shortening of plasma dna in hepatocellular carcinoma patients», *Proceedings of the National Academy of Sciences*, vol. 112, n° 11, p. E1317–E1325.
- JONES, P. A. et S. B. BAYLIN. 2002, «The fundamental role of epigenetic events in cancer», *Nature reviews genetics*, vol. 3, n° 6, p. 415–428.
- KAN, Z., H. ZHENG, X. LIU, S. LI, T. D. BARBER, Z. GONG, H. GAO, K. HAO, M. D. WILLARD, J. XU et collab.. 2013, «Whole-genome sequencing identifies recurrent mutations in hepatocellular carcinoma», *Genome research*, vol. 23, n° 9, p. 1422–1433.
- KINDE, I., J. WU, N. PAPADOPOULOS, K. W. KINZLER et B. VOGELSTEIN. 2011, «Detection and quantification of rare mutations with massively parallel sequencing», *Proceedings of the National Academy of Sciences*, vol. 108, n° 23, p. 9530–9535.
- KOBOLDT, D. C., Q. ZHANG, D. E. LARSON, D. SHEN, M. D. MCLELLAN, L. LIN, C. A. MILLER, E. R. MARDIS, L. DING et R. K. WILSON. 2012, «Varscan

- 2 : somatic mutation and copy number alteration discovery in cancer by exome sequencing», *Genome research*, vol. 22, n° 3, p. 568–576.
- KURTZ, D. M., F. SCHERER, A. M. NEWMAN, A. F. LOVEJOY, D. M. KLASS, J. J. CHABON, S. GAMBHIR, M. DIEHN et A. A. ALIZADEH. 2015, «Dynamic noninvasive genomic monitoring for outcome prediction in diffuse large b-cell lymphoma», .
- LAI, Z., A. MARKOVETS, M. AHDESMAKI, B. CHAPMAN, O. HOFMANN, R. MCEWEN, J. JOHNSON, B. DOUGHERTY, J. C. BARRETT et J. R. DRY. 2016, «Vardict : a novel and versatile variant caller for next-generation sequencing in cancer research», *Nucleic acids research*, vol. 44, n° 11, p. e108–e108.
- LAMB, Y. N. et S. DHILLON. 2017, «Epi procolon® 2.0 ce : A blood-based screening test for colorectal cancer», *Molecular diagnosis & therapy*, vol. 21, n° 2, p. 225–232.
- LANDER, E. S., L. M. LINTON, B. BIRREN, C. NUSBAUM, M. C. ZODY, J. BALDWIN, K. DEVON, K. DEWAR, M. DOYLE, W. FITZHUGH et collab.. 2001, «Initial sequencing and analysis of the human genome», *Nature*, vol. 409, n° 6822, p. 860–921.
- LANDRUM, M. J., J. M. LEE, G. R. RILEY, W. JANG, W. S. RUBINSTEIN, D. M. CHURCH et D. R. MAGLOTT. 2013, «Clinvar : public archive of relationships among sequence variation and human phenotype», *Nucleic acids research*, vol. 42, n° D1, p. D980–D985.
- LANMAN, R. B., S. A. MORTIMER, O. A. ZILL, D. SEBISANOVIC, R. LOPEZ, S. BLAU, E. A. COLLISSON, S. G. DIVERS, D. S. HOON, E. S. KOPETZ et collab.. 2015, «Analytical and clinical validation of a digital sequencing panel for quantitative, highly accurate evaluation of cell-free circulating tumor dna», *PLoS one*, vol. 10, n° 10, p. e0140712.
- LAURENT-PUIG, P., D. PEKIN, C. NORMAND, S. K. KOTSOPOULOS, P. NIZARD, K. PEREZ-TORALLA, R. ROWELL, J. OLSON, P. SRINIVASAN, D. LE CORRE et collab.. 2014, «Clinical relevance of kras-mutated subclones detected with picodroplet digital pcr in advanced colorectal cancer treated with anti-egfr therapy», *Clinical cancer research*.
- LAYER, R. M., C. CHIANG, A. R. QUINLAN et I. M. HALL. 2014, «Lumpy : a probabilistic framework for structural variant discovery», *Genome biology*, vol. 15, n° 6, p. R84.

- LECOMTE, T., A. BERGER, F. ZINZINDOHOUE, S. MICARD, B. LANDI, H. BLONS, P. BEAUNE, P.-H. CUGNENC et P. LAURENT-PUIG. 2002, «Detection of free-circulating tumor-associated dna in plasma of colorectal cancer patients and its association with prognosis», *International journal of cancer*, vol. 100, n° 5, p. 542–548.
- LEHMANN-WERMAN, R., D. NEIMAN, H. ZEMMOUR, J. MOSS, J. MAGENHEIM, A. VAKNIN-DEMBINSKY, S. RUBERTSSON, B. NELLGÅRD, K. BLENNOW, H. ZETTERBERG et collab.. 2016, «Identification of tissue-specific cell death using methylation patterns of circulating dna», *Proceedings of the National Academy of Sciences*, vol. 113, n° 13, p. E1826–E1834.
- LEON, S., B. SHAPIRO, D. SKLAROFF et M. YAROS. 1977, «Free dna in the serum of cancer patients and the effect of therapy», *Cancer research*, vol. 37, n° 3, p. 646–650.
- LEUFKENS, A. M., M. A. VAN DEN BOSCH, M. S. VAN LEEUWEN et P. D. SIERSEMA. 2011, «Diagnostic accuracy of computed tomography for colon cancer staging : a systematic review», *Scandinavian journal of gastroenterology*, vol. 46, n° 7-8, p. 887–894.
- LI, H. 2013, «Aligning sequence reads, clone sequences and assembly contigs with bwa-mem», *arXiv preprint arXiv :1303.3997*.
- LI, H., B. HANDSAKER, A. WYSOKER, T. FENNEL, J. RUAN, N. HOMER, G. MARTH, G. ABECASIS et R. DURBIN. 2009a, «The sequence alignment/map format and samtools», *Bioinformatics*, vol. 25, n° 16, p. 2078–2079.
- LI, J., L. WANG, H. MAMON, M. H. KULKE, R. BERBECO et G. M. MAKRI-GIORGOS. 2008, «Replacing pcr with cold-pcr enriches variant dna sequences and redefines the sensitivity of genetic testing», *Nature medicine*, vol. 14, n° 5, p. 579.
- LI, M., W.-D. CHEN, N. PAPADOPOULOS, S. N. GOODMAN, N. C. BJERREGAARD, S. LAURBERG, B. LEVIN, H. JUHL, N. ARBER, H. MOINOVA et collab.. 2009b, «Sensitive digital quantification of dna methylation in clinical samples», *Nature biotechnology*, vol. 27, n° 9, p. 858–863.
- LIGGETT, T., A. MELNIKOV, Q.-L. YI, C. REPLOGLE, R. BRAND, K. KAUL, M. TALAMONTI, R. A. ABRAMS et V. LEVENSON. 2010, «Differential methylation of cell-free circulating dna among patients with pancreatic cancer versus chronic pancreatitis», *Cancer*, vol. 116, n° 7, p. 1674–1680.

- LIPPMAN, M. et C. K. OSBORNE. 2013, «Circulating tumor dna—ready for prime time», *N Engl J Med*, vol. 368, p. 1249–50.
- LO, Y. D., J. ZHANG, T. N. LEUNG, T. K. LAU, A. M. CHANG et N. M. HJELM. 1999, «Rapid clearance of fetal dna from maternal plasma», *The American Journal of Human Genetics*, vol. 64, n° 1, p. 218–224.
- MAIRE, F., S. MICARD, P. HAMMEL, H. VOITOT, P. LÉVY, P. CUGNENC, P. RUSZNIEWSKI et P. L. PUIG. 2002, «Differential diagnosis between chronic pancreatitis and pancreatic cancer : value of the detection of kras2 mutations in circulating dna», *British journal of cancer*, vol. 87, n° 5, p. 551–554.
- MANDEL, P. et P. METAIS. 1948, «Comptes rendus séances société biol», *Ses Fil*, vol. 142, p. 241–243.
- MARTINCORENA, I., A. ROSHAN, M. GERSTUNG, P. ELLIS, P. VAN LOO, S. MCLAREN, D. C. WEDGE, A. FULLAM, L. B. ALEXANDROV, J. M. TUBIO et collab.. 2015, «High burden and pervasive positive selection of somatic mutations in normal human skin», *Science*, vol. 348, n° 6237, p. 880–886.
- MCCARTHY, D. J., P. HUMBURG, A. KANAPIN, M. A. RIVAS, K. GAULTON, J.-B. CAZIER et P. DONNELLY. 2014, «Choice of transcripts and software has a large effect on variant annotation», *Genome medicine*, vol. 6, n° 3, p. 26.
- MCKENNA, A., M. HANNA, E. BANKS, A. SIVACHENKO, K. CIBULSKIS, A. KERNYTSKY, K. GARIMELLA, D. ALTSHULER, S. GABRIEL, M. DALY et collab.. 2010, «The genome analysis toolkit : a mapreduce framework for analyzing next-generation dna sequencing data», *Genome research*, vol. 20, n° 9, p. 1297–1303.
- MERKER, J. D., G. R. OXNARD, C. COMPTON, M. DIEHN, P. HURLEY, A. J. LAZAR, N. LINDEMAN, C. M. LOCKWOOD, A. J. RAI, R. L. SCHILSKY, A. M. TSIMBERIDOU, V. PATRICIA, B. L. BILLMAN, T. K. OLIVER, S. S. BRUINOOG, D. F. HAYES et T. N. C. 2018, «Circulating tumor dna analysis in patients with cancer : American society of clinical oncology and college of american pathologists joint review», *Journal of clinical oncology*, vol. 36, p. 1–11.
- METZKER, M. L. 2010, «Sequencing technologies—the next generation», *Nature reviews genetics*, vol. 11, n° 1, p. 31.

- MISALE, S., R. YAEGER, S. HOBOR, E. SCALA, M. JANAKIRAMAN, D. LISKA, E. VALTORTA, R. SCHIAVO, M. BUSCARINO, G. SIRAVEGNA et collab.. 2012, «Emergence of kras mutations and acquired resistance to anti-egfr therapy in colorectal cancer», *Nature*, vol. 486, n° 7404, p. 532.
- MORELLI, M., M. OVERMAN, A. DASARI, S. KAZMI, T. MAZARD, E. VILAR, V. MORRIS, M. LEE, D. HERRON, C. ENG et collab.. 2015, «Characterizing the patterns of clonal selection in circulating tumor dna from patients with colorectal cancer refractory to anti-egfr treatment», *Annals of Oncology*, vol. 26, n° 4, p. 731–736.
- MORRIS, L. G., N. RIAZ, A. DESRICHARD, Y. ŞENBABAĞLU, A. A. HAKIMI, V. MAKAROV, J. S. REIS-FILHO et T. A. CHAN. 2016, «Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival», *Oncotarget*, vol. 7, n° 9, p. 10051.
- MOULIERE, F., S. EL MESSAOUDI, C. GONGORA, A.-S. GUEDJ, B. ROBERT, M. DEL RIO, F. MOLINA, P.-J. LAMY, E. LOPEZ-CRAPEZ, M. MATHONNET et collab.. 2013, «Circulating cell-free dna from colorectal cancer patients may reveal high kras or braf mutation load», *Translational oncology*, vol. 6, n° 3, p. 319IN6–328IN8.
- MOULIERE, F., S. EL MESSAOUDI, D. PANG, A. DRITSCHILO et A. R. THIERRY. 2014, «Multi-marker analysis of circulating cell-free dna toward personalized medicine for colorectal cancer», *Molecular oncology*, vol. 8, n° 5, p. 927–941.
- MOULIERE, F., B. ROBERT, E. A. PEYROTTE, M. DEL RIO, M. YCHOU, F. MOLINA, C. GONGORA et A. R. THIERRY. 2011, «High fragmentation characterizes tumour-derived circulating dna», *PloS one*, vol. 6, n° 9, p. e23418.
- MOULIERE, F. et N. ROSENFELD. 2015, «Circulating tumor-derived dna is shorter than somatic dna in plasma», *Proceedings of the National Academy of Sciences*, vol. 112, n° 11, p. 3178–3179.
- MURTAZA, M., S.-J. DAWSON, K. POGREBNIAK, O. M. RUEDA, E. PROVENZANO, J. GRANT, S.-F. CHIN, D. W. TSUI, F. MARASS, D. GALE et collab.. 2015, «Multifocal clonal evolution characterized using circulating tumour dna in a case of metastatic breast cancer», *Nature communications*, vol. 6, p. 8760.
- MURTAZA, M., S.-J. DAWSON, D. W. TSUI, D. GALE, T. FORSHEW, A. M. PISKORZ, C. PARKINSON, S.-F. CHIN, Z. KINGSBURY, A. S. WONG et collab..

- 2013, «Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma dna», *Nature*, vol. 497, n° 7447, p. 108.
- NARIMATSU, H., H. IWASAKI, S. NISHIHARA, H. KIMURA, T. KUDO, Y. YAMAUCHI et S. HIROHASHI. 1996, «Genetic evidence for the lewis enzyme, which synthesizes type-1 lewis antigens in colon tissue, and intracellular localization of the enzyme», *Cancer research*, vol. 56, n° 2, p. 330–338.
- NEWMAN, A. M., S. V. BRATMAN, H. STEHR, L. J. LEE, C. L. LIU, M. DIEHN et A. A. ALIZADEH. 2014a, «Factera : a practical method for the discovery of genomic rearrangements at breakpoint resolution», *Bioinformatics*, vol. 30, n° 23, p. 3390–3393.
- NEWMAN, A. M., S. V. BRATMAN, J. TO, J. F. WYNNE, N. C. ECLOV, L. A. MODLIN, C. L. LIU, J. W. NEAL, H. A. WAKELEE, R. E. MERRITT et collab.. 2014b, «An ultrasensitive method for quantitating circulating tumor dna with broad patient coverage», *Nature medicine*, vol. 20, n° 5, p. 548.
- NEWMAN, A. M., A. F. LOVEJOY, D. M. KLASS, D. M. KURTZ, J. J. CHABON, F. SCHERER, H. STEHR, C. L. LIU, S. V. BRATMAN, C. SAY et collab.. 2016, «Integrated digital error suppression for improved detection of circulating tumor dna», *Nature biotechnology*, vol. 34, n° 5, p. 547.
- NEWTON, C., A. GRAHAM, L. HEPTINSTALL, S. POWELL, C. SUMMERS, N. KALSHEKER, J. SMITH et A. MARKHAM. 1989, «Analysis of any point mutation in dna. the amplification refractory mutation system (arms)», *Nucleic acids research*, vol. 17, n° 7, p. 2503–2516.
- NG, P. C. et S. HENIKOFF. 2003, «Sift : Predicting amino acid changes that affect protein function», *Nucleic acids research*, vol. 31, n° 13, p. 3812–3814.
- NORTON, L., R. SIMON, H. D. BRERETON et A. E. BOGDEN. 1976, «Predicting the course of gompertzian growth», *Nature*, vol. 264, n° 5586, p. 542.
- OAKMAN, C., L. SANTARPIA et A. DI LEO. 2010, «Breast cancer assessment tools and optimizing adjuvant therapy», *Nature reviews Clinical oncology*, vol. 7, n° 12, p. 725.
- OESTERLING, J. E. 1991, «Prostate specific antigen : a critical assessment of the most useful tumor marker for adenocarcinoma of the prostate», *The Journal of urology*, vol. 145, n° 5, p. 907–923.

- OHSHIO, G., T. MANABE, Y. WATANABE, K. ENDO, H. KUDO, T. SUZUKI et T. TOBE. 1990, «Comparative studies of du-pan-2, carcinoembryonic antigen, and ca19-9 in the serum and bile of patients with pancreatic and biliary tract diseases : evaluation of the influence of obstructive jaundice.», *American Journal of Gastroenterology*, vol. 85, n° 10.
- OLSEN, D. et J. T. JØRGENSEN. 2014, «Companion diagnostics for targeted cancer drugs—clinical and regulatory aspects», *Frontiers in oncology*, vol. 4, p. 105.
- OLSSON, E., C. WINTER, A. GEORGE, Y. CHEN, J. HOWLIN, M.-H. E. TANG, M. DAHLGREN, R. SCHULZ, D. GRABAU, D. VAN WESTEN et collab.. 2015, «Serial monitoring of circulating tumor dna in patients with primary breast cancer for detection of occult metastatic disease», *EMBO molecular medicine*, vol. 7, n° 8, p. 1034–1047.
- OPITZ, L., G. SALINAS-RIESTER, M. GRADE, K. JUNG, P. JO, G. EMONS, B. M. GHADIMI, T. BEISSBARTH et J. GAEDCKE. 2010, «Impact of rna degradation on gene expression profiling», *BMC medical genomics*, vol. 3, n° 1, p. 36.
- OSTROW, K., C. MICHAILIDI, R. GUERRERO-PRESTON, M. HOQUE, A. GREENBERG, W. ROM et D. SIDRANSKY. 2013, «Cigarette smoke induces methylation of the tumor suppressor gene nisch», *Epigenetics*, vol. 8, n° 4, p. 383–388.
- OSTROW, K. L., M. O. HOQUE, M. LOYO, M. BRAIT, A. GREENBERG, J. M. SIEGFRIED, J. R. GRANDIS, A. G. DAVIS, W. L. BIGBEE, W. ROM et collab.. 2010, «Molecular analysis of plasma dna for the early detection of lung cancer by quantitative methylation-specific pcr», *Clinical Cancer Research*, vol. 16, n° 13, p. 3463–3472.
- OVERMAN, M. J., J. MODAK, S. KOPETZ, R. MURTHY, J. C. YAO, M. E. HICKS, J. L. ABBRUZZESE et A. L. TAM. 2012, «Use of research biopsies in clinical trials : are risks and benefits adequately discussed?», *Journal of Clinical Oncology*, vol. 31, n° 1, p. 17–22.
- PAILA, U., B. A. CHAPMAN, R. KIRCHNER et A. R. QUINLAN. 2013, «Gemini : integrative exploration of genetic variation and genome annotations», *PLoS computational biology*, vol. 9, n° 7, p. e1003153.
- PALMISANO, W. A., K. K. DIVINE, G. SACCOMANNO, F. D. GILLILAND, S. B. BAYLIN, J. G. HERMAN et S. A. BELINSKY. 2000, «Predicting lung cancer by

- detecting aberrant promoter methylation in sputum», *Cancer research*, vol. 60, n° 21, p. 5954–5958.
- PAVLIDIS, N., E. BRIASOULIS, J. HAINSWORTH et F. GRECO. 2003, «Diagnostic and therapeutic management of cancer of an unknown primary», *European journal of cancer*, vol. 39, n° 14, p. 1990–2005.
- PÉCUCHET, N., Y. ROZENHOLC, E. ZONTA, D. PIETRAZ, A. DIDELOT, P. COMBE, L. GIBAUT, J.-B. BACHET, V. TALY, E. FABRE et collab.. 2016a, «Analysis of base-position error rate of next-generation sequencing to detect tumor mutations in circulating dna», *Clinical chemistry*, vol. 62, n° 11, p. 1492–1503.
- PÉCUCHET, N., E. ZONTA, A. DIDELOT, P. COMBE, C. THIBAUT, L. GIBAUT, C. LOURS, Y. ROZENHOLC, V. TALY, P. LAURENT-PUIG et collab.. 2016b, «Base-position error rate analysis of next-generation sequencing applied to circulating tumor dna in non-small cell lung cancer : a prospective study», *PLoS medicine*, vol. 13, n° 12, p. e1002199.
- PÉREZ-BARRIOS, C., I. NIETO-ALCOLADO, M. TORRENTE, C. JIMÉNEZ-SÁNCHEZ, V. CALVO, L. GUTIERREZ-SANZ, M. PALKA, E. DONOSO-NAVARRO, M. PROVENCIO et A. ROMERO. 2016, «Comparison of methods for circulating cell-free dna isolation using blood from cancer patients : impact on biomarker testing», *Translational lung cancer research*, vol. 5, n° 6, p. 665.
- POTTER, N. T., P. HURBAN, M. N. WHITE, K. D. WHITLOCK, C. E. LOFTON-DAY, R. TETZNER, T. KOENIG, N. B. QUIGLEY et G. WEISS. 2014, «Validation of a real-time pcr-based qualitative assay for the detection of methylated sept9 dna in human plasma», *Clinical chemistry*, vol. 60, n° 9, p. 1183–1191.
- QU, W., S.-I. HASHIMOTO et S. MORISHITA. 2009, «Efficient frequency-based de novo short-read clustering for error trimming in next-generation sequencing», *Genome research*, vol. 19, n° 7, p. 1309–1315.
- QUINLAN, A. R. et I. M. HALL. 2010, «Bedtools : a flexible suite of utilities for comparing genomic features», *Bioinformatics*, vol. 26, n° 6, p. 841–842.
- RAAIJMAKERS, R., W. J. KIRKELS, M. J. ROOBOL, M. F. WILDHAGEN et F. H. SCHRDER. 2002, «Complication rates and risk factors of 5802 transrectal ultrasound-guided sextant biopsies of the prostate within a population-based screening program», *Urology*, vol. 60, n° 5, p. 826–830.

- RAMIREZ, J. L., R. ROSELL, M. TARON, M. SANCHEZ-RONCO, V. ALBEROLA, R. DE LAS PEÑAS, J. M. SANCHEZ, T. MORAN, C. CAMPS, B. MASSUTI et collab.. 2005, «14-3-3 σ methylation in pretreatment serum circulating dna of cisplatin-plus-gemcitabine-treated advanced non-small-cell lung cancer patients predicts survival : the spanish lung cancer group», *Journal of Clinical Oncology*, vol. 23, n $^{\circ}$ 36, p. 9105–9112.
- RAVO, M., M. MUTARELLI, L. FERRARO, O. M. V. GROBER, O. PARIS, R. TARALLO, A. VIGILANTE, D. CIMINO, M. DE BORTOLI, E. NOLA et collab.. 2008, «Quantitative expression profiling of highly degraded rna from formalin-fixed, paraffin-embedded breast tumor biopsies by oligonucleotide microarrays», *Laboratory Investigation*, vol. 88, n $^{\circ}$ 4, p. 430.
- REINERT, T., L. V. SCHØLER, R. THOMSEN, H. TOBIASEN, S. VANG, I. NORDENTOFT, P. LAMY, A.-S. KANNERUP, F. V. MORTENSEN, K. STRIBOLT et collab.. 2016, «Analysis of circulating tumour dna to monitor disease burden following colorectal cancer surgery», *Gut*, vol. 65, n $^{\circ}$ 4, p. 625–634.
- RIEDINGER, J., J. WAFFLART, G. RICOLLEAU, N. ECHE, H. LARBRE, J. BASUYAU, I. DALIFARD, K. HACENE et M. PICHON. 2006, «Ca 125 half-life and ca 125 nadir during induction chemotherapy are independent predictors of epithelial ovarian cancer outcome : results of a french multicentric study», *Annals of oncology*, vol. 17, n $^{\circ}$ 8, p. 1234–1238.
- ROSCHEWSKI, M., K. DUNLEAVY, S. PITTALUGA, M. MOORHEAD, F. PEPIN, K. KONG, M. SHOVLIN, E. S. JAFFE, L. M. STAUDT, C. LAI et collab.. 2015, «Circulating tumour dna and ct monitoring in patients with untreated diffuse large b-cell lymphoma : a correlative biomarker study», *The Lancet Oncology*, vol. 16, n $^{\circ}$ 5, p. 541–549.
- ROSCHEWSKI, M., L. M. STAUDT et W. H. WILSON. 2016, «Dynamic monitoring of circulating tumor dna in non-hodgkin lymphoma», *Blood*, vol. 127, n $^{\circ}$ 25, p. 3127–3132.
- ROSENFELD, N., R. AHARONOV, E. MEIRI, S. ROSENWALD, Y. SPECTOR, M. ZEPENIUK, H. BENJAMIN, N. SHABES, S. TABAK, A. LEVY et collab.. 2008, «MicroRNAs accurately identify cancer tissue origin», *Nature biotechnology*, vol. 26, n $^{\circ}$ 4, p. 462.
- ROY, S., C. COLDREN, A. KARUNAMURTHY, N. S. KIP, E. W. KLEE, S. E. LINCOLN, A. LEON, M. PULLAMBHATLA, R. L. TEMPLE-SMOLKIN, K. V. VOELK-

- ERDING et collab.. 2017, «Standards and guidelines for validating next-generation sequencing bioinformatics pipelines : A joint recommendation of the association for molecular pathology and the college of american pathologists», *The Journal of Molecular Diagnostics*.
- RYAN, B., F. LEFORT, R. MCMANUS, J. DALY, P. KEELING, D. WEIR et D. KELLEHER. 2003, «A prospective study of circulating mutant kras2 in the serum of patients with colorectal neoplasia : strong prognostic indicator in post-operative follow up», *Gut*, vol. 52, n° 1, p. 101–108.
- SANDMANN, S., A. O. DE GRAAF, M. KARIMI, B. A. VAN DER REIJDEN, E. HELLSTRÖM-LINDBERG, J. H. JANSEN et M. DUGAS. 2017, «Evaluating variant calling tools for non-matched next-generation sequencing data», *Scientific reports*, vol. 7, p. 43169.
- SANTIAGO-WALKER, A., R. GAGNON, J. MAZUMDAR, M. CASEY, G. V. LONG, D. SCHADENDORF, K. FLAHERTY, R. KEFFORD, A. HAUSCHILD, P. HWU et collab.. 2015, «Correlation of braf mutation status in circulating-free dna and tumor and association with clinical outcome across four braf and meki clinical trials», *Clinical Cancer Research*.
- SAWYERS, C. 2004, «Targeted cancer therapy», *Nature*, vol. 432, n° 7015, p. 294.
- SCHADENDORF, D., D. E. FISHER, C. GARBE, J. E. GERSHENWALD, J.-J. GROB, A. HALPERN, M. HERLYN, M. A. MARCHETTI, G. MCARTHUR, A. RIBAS et collab.. 2015, «Melanoma», *Nature reviews Disease primers*, vol. 1, p. 15003.
- SCHERER, F., D. M. KURTZ, A. M. NEWMAN, H. STEHR, A. F. CRAIG, M. S. ESFAHANI, A. F. LOVEJOY, J. J. CHABON, D. M. KLASS, C. L. LIU et collab.. 2016, «Distinct biological subtypes and patterns of genome evolution in lymphoma revealed by circulating tumor dna», *Science translational medicine*, vol. 8, n° 364, p. 364ra155–364ra155.
- SCHERER, F., D. M. KURTZ, A. M. NEWMAN, H. STEHR, C. L. LIU, L. ZHOU, A. F. CRAIG, J. J. CHABON, A. F. LOVEJOY, D. M. KLASS et collab.. 2015, «Noninvasive genotyping and assessment of treatment response in diffuse large b cell lymphoma», .
- SCHIRMER, M., U. Z. IJAZ, R. D'AMORE, N. HALL, W. T. SLOAN et C. QUINCE. 2015, «Insight into biases and sequencing errors for amplicon sequencing with the illumina miseq platform», *Nucleic acids research*, vol. 43, n° 6, p. e37–e37.

- SCHOLEFIELD, J., S. MOSS, C. MANGHAM, D. WHYNES et J. HARDCASTLE. 2011, «Nottingham trial of faecal occult blood testing for colorectal cancer : a 20-year follow-up», *Gut*, p. gutjnl-2011.
- SCHØLER, L. V., T. REINERT, M.-B. W. ØRNTOFT, C. G. KASSENTOFT, S. S. ÁRNADÓTTIR, S. VANG, I. NORDENTOFT, M. KNUDSEN, P. LAMY, D. ANDREASEN et collab.. 2017, «Clinical implications of monitoring circulating tumor dna in patients with colorectal cancer», *Clinical Cancer Research*, vol. 23, n° 18, p. 5437-5445.
- SCHWARZENBACH, H., D. S. HOON et K. PANTEL. 2011, «Cell-free nucleic acids as biomarkers in cancer patients», *Nature Reviews Cancer*, vol. 11, n° 6, p. 426.
- SCOTT, D. W., G. W. WRIGHT, P. M. WILLIAMS, C.-J. LIH, W. WALSH, E. S. JAFFE, A. ROSENWALD, E. CAMPO, W. C. CHAN, J. M. CONNORS et collab.. 2014, «Determining cell-of-origin subtypes of diffuse large b-cell lymphoma using gene expression in formalin-fixed paraffin-embedded tissue», *Blood*, vol. 123, n° 8, p. 1214-1217.
- SHAH, S. P., R. D. MORIN, J. KHATTRA, L. PRENTICE, T. PUGH, A. BURLEIGH, A. DELANEY, K. GELMON, R. GULIANY, J. SENZ et collab.. 2009, «Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution», *Nature*, vol. 461, n° 7265, p. 809.
- SIRAVEGNA, G., B. MUSSOLIN, M. BUSCARINO, G. CORTI, A. CASSINGENA, G. CRISAFULLI, A. PONZETTI, C. CREMOLINI, A. AMATU, C. LAURICELLA et collab.. 2015, «Clonal evolution and resistance to egfr blockade in the blood of colorectal cancer patients», *Nature medicine*, vol. 21, n° 7, p. 795.
- SMITH-BINDMAN, R., J. LIPSON, R. MARCUS, K.-P. KIM, M. MAHESH, R. GOULD, A. B. DE GONZÁLEZ et D. L. MIGLIORETTI. 2009, «Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer», *Archives of internal medicine*, vol. 169, n° 22, p. 2078-2086.
- SNYDER, M. W., M. KIRCHER, A. J. HILL, R. M. DAZA et J. SHENDURE. 2016, «Cell-free dna comprises an in vivo nucleosome footprint that informs its tissues-of-origin», *Cell*, vol. 164, n° 1, p. 57-68.
- SORENSEN, B. S., L. WU, W. WEI, J. TSAI, B. WEBER, E. NEXO et P. MELDGAARD. 2014, «Monitoring of epidermal growth factor receptor tyrosine

- kinase inhibitor-sensitizing and resistance mutations in the plasma dna of patients with advanced non-small cell lung cancer during treatment with erlotinib», *Cancer*, vol. 120, n° 24, p. 3896–3901.
- SORENSEN, G. D., D. M. PRIBISH, F. H. VALONE, V. A. MEMOLI, D. J. BZIK et S.-L. YAO. 1994, «Soluble normal and mutated dna sequences from single-copy genes in human blood.», *Cancer Epidemiology and Prevention Biomarkers*, vol. 3, n° 1, p. 67–71.
- SOZZI, G., D. CONTE, M. LEON, R. CIRINCIONE, L. ROZ, C. RATCLIFFE, E. ROZ, N. CIRENEI, M. BELLOMI, G. PELOSI et collab.. 2003, «Quantification of free circulating dna as a diagnostic marker in lung cancer», *Journal of Clinical Oncology*, vol. 21, n° 21, p. 3902–3908.
- SOZZI, G., D. CONTE, L. MARIANI, S. L. VULLO, L. ROZ, C. LOMBARDO, M. A. PIEROTTI et L. TAVECCHIO. 2001, «Analysis of circulating tumor dna in plasma at diagnosis and during follow-up of lung cancer patients», *Cancer research*, vol. 61, n° 12, p. 4675–4678.
- STROM, C. M., S. RIVERA, C. ELZINGA, T. ANGELONI, S. H. ROSENTHAL, D. GOOS-ROOT, M. SIAW, J. PLATT, C. BRAASTADT, L. CHENG et collab.. 2015, «Development and validation of a next-generation sequencing assay for brca1 and brca2 variants for the clinical laboratory», *PLoS One*, vol. 10, n° 8, p. e0136419.
- STROUN, M., P. ANKER, J. LYAUTEY, C. LEDERREY et P. A. MAURICE. 1987, «Isolation and characterization of dna from the plasma of cancer patients», *European Journal of Cancer and Clinical Oncology*, vol. 23, n° 6, p. 707–712.
- STROUN, M., P. ANKER, P. MAURICE, J. LYAUTEY, C. LEDERREY et M. BELJANSKI. 1989, «Neoplastic characteristics of the dna found in the plasma of cancer patients», *Oncology*, vol. 46, n° 5, p. 318–322.
- SUN, C., L. WANG, S. HUANG, G. J. HEYNEN, A. PRAHALLAD, C. ROBERT, J. HAANEN, C. BLANK, J. WESSELING, S. M. WILLEMS et collab.. 2014, «Reversible and adaptive resistance to braf (v600e) inhibition in melanoma», *Nature*, vol. 508, n° 7494, p. 118.
- SUN, K., P. JIANG, K. A. CHAN, J. WONG, Y. K. CHENG, R. H. LIANG, W.-K. CHAN, E. S. MA, S. L. CHAN, S. H. CHENG et collab.. 2015, «Plasma dna tissue mapping by genome-wide methylation sequencing for noninvasive prenatal,

- cancer, and transplantation assessments», *Proceedings of the National Academy of Sciences*, vol. 112, n° 40, p. E5503–E5512.
- TALAULIKAR, D., J. X. GRAY, B. SHADBOLT, M. MCNIVEN et J. E. DAHLSTROM. 2008, «A comparative study of the quality of dna obtained from fresh frozen and formalin-fixed decalcified paraffin-embedded bone marrow trephine biopsy specimens using two different methods», *Journal of clinical pathology*, vol. 61, n° 1, p. 119–123.
- TALEVICH, E., A. H. SHAIN, T. BOTTON et B. C. BASTIAN. 2016, «Cnvkit : genome-wide copy number detection and visualization from targeted dna sequencing», *PLoS computational biology*, vol. 12, n° 4, p. e1004873.
- TALY, V., D. PEKIN, L. BENHAIM, S. K. KOTSOPOULOS, D. LE CORRE, X. LI, I. ATOCHIN, D. R. LINK, A. D. GRIFFITHS, K. PALLIER et collab.. 2013, «Multiplex picodroplet digital pcr to detect kras mutations in circulating dna from the plasma of colorectal cancer patients», *Clinical chemistry*, vol. 59, n° 12, p. 1722–1731.
- TAN, E., P. SCHUR, R. CARR et H. KUNKEL. 1966, «Deoxybonucleic acid (dna) and antibodies to dna in the serum of patients with systemic lupus erythematosus.», *Journal of Clinical Investigation*, vol. 45, n° 11, p. 1732.
- THE CANCER GENOME ATLAS RESEARCH NETWORK. 2011, «Integrated genomic analyses of ovarian carcinoma», *Nature*, vol. 474, n° 7353, p. 609–615.
- THIEBLEMONT, C., J. BRIERE, N. MOUNIER, H.-U. VOELKER, W. CUCCUINI, E. HIRCHAUD, A. ROSENWALD, A. JACK, C. SUNDSTROM, S. COGLIATTI et collab.. 2011, «The germinal center/activated b-cell subclassification has a prognostic impact for response to salvage therapy in relapsed/refractory diffuse large b-cell lymphoma : a bio-coral study», *Journal of clinical oncology*, vol. 29, n° 31, p. 4079–4087.
- THIERRY, A., S. EL MESSAOUDI, P. GAHAN, P. ANKER et M. STROUN. 2016, «Origins, structures, and functions of circulating dna in oncology», *Cancer and Metastasis Reviews*, vol. 35, n° 3, p. 347–376.
- THIERRY, A. R., F. MOULIERE, S. EL MESSAOUDI, C. MOLLEVI, E. LOPEZ-CRAPEZ, F. ROLET, B. GILLET, C. GONGORA, P. DECHELOTTE, B. ROBERT et collab.. 2014, «Clinical validation of the detection of kras and braf mutations from circulating tumor dna», *Nature medicine*, vol. 20, n° 4, p. 430–435.

- THIERRY, A. R., F. MOULIERE, C. GONGORA, J. OLLIER, B. ROBERT, M. YCHOU, M. DEL RIO et F. MOLINA. 2010, «Origin and quantification of circulating dna in mice with human colorectal cancer xenografts», *Nucleic acids research*, vol. 38, n° 18, p. 6159–6175.
- TIE, J., Y. WANG, C. TOMASETTI, L. LI, S. SPRINGER, I. KINDE, N. SILLIMAN, M. TACEY, H.-L. WONG, M. CHRISTIE et collab.. 2016, «Circulating tumor dna analysis detects minimal residual disease and predicts recurrence in patients with stage ii colon cancer», *Science translational medicine*, vol. 8, n° 346, p. 346ra92–346ra92.
- TSAO, S. C.-H., J. WEISS, C. HUDSON, C. CHRISTOPHI, J. CEBON, A. BEHREN et A. DOBROVIC. 2015, «Monitoring response to therapy in melanoma by quantifying circulating tumour dna with droplet digital pcr for braf and nras mutations», *Scientific reports*, vol. 5, p. 11 198.
- ULZ, P., J. BELIC, R. GRAF, M. AUER, I. LAFER, K. FISCHEREDER, G. WEBERSINKE, K. PUMMER, H. AUGUSTIN, M. PICHLER et collab.. 2016, «Whole-genome plasma sequencing reveals focal amplifications as a driving force in metastatic prostate cancer», *Nature communications*, vol. 7, p. 12 008.
- UMETANI, N., A. E. GIULIANO, S. H. HIRAMATSU, F. AMERSI, T. NAKAGAWA, S. MARTINO et D. S. HOON. 2006a, «Prediction of breast tumor progression by integrity of free circulating dna in serum», *Journal of clinical oncology*, vol. 24, n° 26, p. 4270–4276.
- UMETANI, N., J. KIM, S. HIRAMATSU, H. A. REBER, O. J. HINES, A. J. BILCHIK et D. S. HOON. 2006b, «Increased integrity of free circulating dna in sera of patients with colorectal or periampullary cancer : direct quantitative pcr for alu repeats», *Clinical chemistry*, vol. 52, n° 6, p. 1062–1069.
- VAN'T VEER, L. J., H. DAI, M. J. VAN DE VIJVER, Y. D. HE, A. A. HART, M. MAO, H. L. PETERSE, K. VAN DER KOOY, M. J. MARTON, A. T. WITTEVEEN et collab.. 2002, «Gene expression profiling predicts clinical outcome of breast cancer», *nature*, vol. 415, n° 6871, p. 530.
- VENKATESAN, S. et C. SWANTON. 2016, «Tumor evolutionary principles : How intratumor heterogeneity influences cancer treatment and outcome.», dans *American Society of Clinical Oncology educational book. American Society of Clinical Oncology. Meeting*, vol. 35, p. e141–9.

- VOGELSTEIN, B. et K. W. KINZLER. 1999, «Digital pcr», *Proceedings of the National Academy of Sciences*, vol. 96, n° 16, p. 9236–9241.
- WAN, J. C., C. MASSIE, J. GARCIA-CORBACHO, F. MOULIERE, J. D. BRENTON, C. CALDAS, S. PACEY, R. BAIRD et N. ROSENFELD. 2017, «Liquid biopsies come of age : towards implementation of circulating tumour dna», *Nature Reviews Cancer*, vol. 17, n° 4, p. 223–238.
- WARDEN, C. D., A. W. ADAMSON, S. L. NEUHAUSEN et X. WU. 2014, «Detailed comparison of two popular variant calling packages for exome and targeted exon studies», *PeerJ*, vol. 2, p. e600.
- WARTON, K. et G. SAMIMI. 2015, «Methylation of cell-free circulating dna in the diagnosis of cancer», *Frontiers in molecular biosciences*, vol. 2.
- WONG, S. Q., J. M. RALEIGH, J. CALLAHAN, I. A. VERGARA, S. FTOUNI, A. HATZIMIHALIS, A. J. COLEBATCH, J. LI, T. SEMPLE, K. DOIG et collab.. 2017, «Circulating tumor dna analysis and functional imaging provide complementary approaches for comprehensive disease monitoring in metastatic melanoma», *Precision Oncology*, vol. 1, p. 1–14.
- YANCOVITZ, M., A. LITTERMAN, J. YOON, E. NG, R. L. SHAPIRO, R. S. BERMAN, A. C. PAVLICK, F. DARVISHIAN, P. CHRISTOS, M. MAZUMDAR et collab.. 2012, «Intra-and inter-tumor heterogeneity of brafv600emutations in primary and metastatic melanoma», *PloS one*, vol. 7, n° 1, p. e29 336.
- YEH, P., M. DICKINSON, S. FTOUNI, T. HUNTER, D. SINHA, S. Q. WONG, R. AGARWAL, R. VEDURURU, K. DOIG, C. Y. FONG et collab.. 2017a, «Molecular disease monitoring using circulating tumor dna in myelodysplastic syndromes», *Blood*, vol. 129, n° 12, p. 1685–1690.
- YEH, P., T. HUNTER, D. SINHA, S. FTOUNI, E. WALLACH, D. JIANG, Y.-C. CHAN, S. Q. WONG, M. J. SILVA, R. VEDURURU et collab.. 2017b, «Circulating tumour dna reflects treatment response and clonal evolution in chronic lymphocytic leukaemia», *Nature Communications*, vol. 8, p. 14 756.
- YOUNG, G. P., C. G. FRASER, S. P. HALLORAN et S. COLE. 2012, «Guaiac based faecal occult blood testing for colorectal cancer screening : an obsolete strategy?», *Gut*, p. gutjnl–2011.

DEMANDE D'IMPRIMATUR

Date de soutenance : 25/05/2018

DIPLOME D'ETAT DE DOCTEUR
EN PHARMACIE

présenté par : FILHINE-TRESARRIEU Pierre

Sujet : Les ADN tumoraux circulant : rappel
bibliographique et mise en place d'un pipeline
d'analyses bioinformatiqueJury :Président : M. Jean-Louis MERLIN, Professeur
Directeur : M. Alexandre HARLE, Maître de Conférences
Juges : M Jean-Yves JOUZEAU, Professeur
M Kevin DALLEAU, Pharmacien

Vu,

Nancy, le 23/4/2018

Le Président du Jury

Directeur de Thèse

M. Jean-Louis MERLIN

M. Alexandre HARLE

Vu et approuvé,

Nancy, le 4.05.2018

Doyen de la Faculté de Pharmacie
de l'Université de Lorraine,

Francine PAULUS



Vu,

Nancy, le 25 MAI 2018

Le Président de l'Université de Lorraine,

Pierre MUTZENHARDT



N° d'enregistrement : 10293

N° d'identification :

TITRE

Les ADN tumoraux circulant : rappel bibliographique et mise en place d'un pipeline d'analyses bioinformatique

Thèse soutenue le 25 mai 2018

Par Pierre FILHINE-TRESARRIEU

RESUME :

De nombreux biomarqueurs ont été mis au point afin de suivre efficacement les patients atteints de cancer. Parmi ceux-ci, l'ADN tumoral circulant (ADNtc) est un des plus prometteurs. Composé de diverses altérations somatiques provenant des cellules tumorales, l'ADNtc se retrouve dans divers fluides biologiques comme le sang, ce qui permet d'effectuer une « biopsie liquide », de manière non invasive. L'ADNtc a été évalué comme biomarqueur diagnostique, pronostique, du risque de récurrence mais aussi en théranostic comme potentiellement remplaçant la biopsie des lésions. Si les premières détections d'altérations ont été faites dans le milieu des années 1990, des progrès techniques ont été nécessaires afin de détecter l'ADNtc, celui-ci se trouvant en très faible quantité dans les liquides biologiques. Des méthodes en point unique, comme le BEAMing ou la réaction en chaîne par polymérase digitale ont été développés en parallèle des approches qui s'appuient sur le séquençage haut débit. Ces dernières nécessitent un traitement bioinformatique des données, afin d'extraire les altérations d'intérêt. Dans ce travail, un *pipeline* d'analyses a été mis au point et évalué. Ce *pipeline* permet de détecter des variants d'une seule base aux structuraux. L'évaluation de ce *pipeline* a été faite à partir de dilutions de lignée cellulaire HCC-78, ce qui a permis de déterminer la sensibilité à 1 variant somatique pour 500 sauvages. Les variants structuraux, comme les gènes de fusion, ont aussi été mis en évidence. Néanmoins, la spécificité de ce *pipeline* reste faible, à cause d'un bruit de fond important. Il s'avère nécessaire d'évaluer ce dernier à partir d'échantillons sains afin d'augmenter la sensibilité et la spécificité des analyses.

MOTS CLES :

Cancer, séquençage haut débit, biomarqueur, bioinformatique

Directeur de thèse	Intitulé du laboratoire	Nature	
Dr Alexandre HARLE	CRAN UMR 7039, Université de Lorraine, CNRS	Expérimentale	X
		Bibliographique	<input type="checkbox"/>
		Thème	5

Thèmes

1 – Sciences fondamentales
3 - Médicaments
⑤ - Biologie

2 – Hygiène/Environnements
4 – Alimentation – Nutrition
6 – Pratique professionnelle