



HAL
open science

Apport du séquençage haut débit et de la caractérisation fonctionnelle de variants somatiques pour l'accessibilité aux thérapies ciblées en oncologie

Louise-Marie Chiron Chevalier

► To cite this version:

Louise-Marie Chiron Chevalier. Apport du séquençage haut débit et de la caractérisation fonctionnelle de variants somatiques pour l'accessibilité aux thérapies ciblées en oncologie. Médecine humaine et pathologie. Université d'Angers, 2022. Français. NNT : 2022ANGE0017 . tel-03907954

HAL Id: tel-03907954

<https://theses.hal.science/tel-03907954>

Submitted on 20 Dec 2022

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.

THESE DE DOCTORAT DE

L'UNIVERSITE D'ANGERS

ECOLE DOCTORALE N° 605

Biologie Santé

Spécialité : Cancérologie

Par

Louise-Marie CHIRON-CHEVALIER

**Apport du séquençage haut débit et de la caractérisation
fonctionnelle de variants somatiques pour l'accessibilité
aux thérapies ciblées en oncologie**

Thèse présentée et soutenue à Angers, le 19 septembre 2022

Unité de recherche : Unité de Génomique Fonctionnelle - Institut de Cancérologie de l'Ouest

INSERM U1307 CRCI2NA Equipe 4 « Immunité Innée et Cancer »

Thèse N° : 227696

Rapporteurs avant soutenance :

Marie-Dominique Galibert
Lucie Karayan-Tapon

Professeur d'Université, CHU de Rennes
Professeur d'Université, CHU de Poitiers

Composition du Jury :

Président : Mario Campone

Professeur d'Université, Institut de Cancérologie de l'Ouest - Angers

Examineurs : Marie-Christine Copin
Sophie Krieger
Olivier Molinier

Professeur d'Université, CHU d'Angers
Maître de conférence d'Université, Centre François Baclesse - Caen
Praticien Hospitalier, CH Le Mans

Dir. de thèse : Alain Morel

Professeur d'Université, Institut de Cancérologie de l'Ouest - Angers

TABLE DES MATIERES

LISTE DES ABREVIATIONS.....	3
INTRODUCTION.....	7
1. LA MEDECINE PERSONNALISEE EN ONCOLOGIE	8
1.1. Historique de la prise en charge des cancers et développement de la médecine personnalisée en oncologie	8
1.2. Thérapie ciblée et biomarqueurs	10
1.3. Connaissance des processus tumoraux pour identifier les biomarqueurs d'intérêt.	15
1.4. Evolution technologique pour la détection des biomarqueurs en cancérologie : la généralisation du séquençage haut débit.....	17
1.5. Interprétation des variants, bases de données et problématique des variants de signification inconnue.....	24
2. EGFR ET CANCERS BRONCHOPULMONAIRES : ENJEUX ET CHOIX TECHNOLOGIQUES.....	32
2.1. Les cancers bronchopulmonaires : généralités	32
2.2. Hétérogénéité mutationnelle des CBNPC et opportunités thérapeutiques	33
2.3. CBNPC, EGFR et thérapies ciblées.....	36
a) Structure et fonction du récepteur EGFR	36
b) Mutations oncogéniques d'EGFR et thérapies ciblées dans les CBNPC	38
c) ITK-EGFR de première, deuxième et troisième génération	39
d) Résistance aux ITK-EGFR	42
e) Insertions dans l'exon 20 d'EGFR et thérapies ciblées	45
2.4. Détection des mutations <i>EGFR</i> pour la prise en charge clinique des CBNPC...46	
a) Tests moléculaires sur échantillons de tissus.....	46
b) Tests moléculaires dans l'ADNtc	48
c) Les techniques de détection des mutations <i>EGFR</i>	50

3. INTERPRETATION DES VSI POUR LA PRESCRIPTION DES INHIBITEURS DE PARP 53

3.1. Mécanisme d'action des inhibiteurs de PARP et principe de létalité synthétique.....	53
3.2. La voie de recombinaison homologue.....	54
3.3. Tumeurs mutées <i>BRCA1</i> et <i>BRCA2</i>	59
a) <i>BRCA1</i> : généralités.....	59
b) <i>BRCA2</i> : généralités.....	60
c) <i>BRCA1</i> , <i>BRCA2</i> et prédisposition aux cancers	62
d) Pathologie moléculaire des gènes <i>BRCA1</i> et <i>BRCA2</i>	62
e) Analyse des gènes <i>BRCA1</i> et <i>BRCA2</i>	64
3.4. Anomalies moléculaires des tumeurs de profil <i>BRCAness</i> et signatures HRD..	65
a) Mécanismes alternatifs de déficit de la voie de RH	65
b) Les signatures HRD	66
3.5. Indications et biomarqueurs de prescription des iPARP.....	70
a) Cancers de l'ovaire	70
b) Cancers du sein	72
c) Cancers du pancréas	74
d) Cancers de la prostate	76
3.6. Interprétation des VSI des gènes <i>BRCA1</i> et <i>BRCA2</i> pour la prescription des iPARP : place des tests fonctionnels	78
a) VSI de <i>BRCA1</i> et <i>BRCA2</i>	78
b) Caractérisation biologique fonctionnelle des VSI d'épissage des gènes <i>BRCA1</i> et <i>BRCA2</i>	81
c) Caractérisation biologique fonctionnelle des VSI des gènes <i>BRCA1/2</i> par édition génomique CRISPR-Cas9.....	84

OBJECTIFS DE LA THESE..... 87

RESULTATS ET DISCUSSION..... 88

1. DETECTION DES MUTATIONS *EGFR* DANS LES CBNPC..... 89

Article 1 : Caractérisation moléculaire de l'*EGFR* dans les cancers bronchiques non a petites cellules : étude prospective comparative des technologies NGS et automate Idylla.....89

2. INTERPRETATION FONCTIONNELLE DES VSI *BRCA1* et *BRCA2*..... 105

Article 2 : Somatic mRNA analysis of *BRCA1* splice variants provides a direct theranostic impact on PARP inhibitors.....106

Article 3 : Functional pre-therapeutic evaluation by genome editing of variants of uncertain significance of essential tumor suppressor genes.....120

DISCUSSION GENERALE ET PERSPECTIVES 148

BIBLIOGRAPHIE 152

TABLE DES FIGURES 178

TABLE DES TABLEAUX 180

LISTE DES ABREVIATIONS

AA : acide aminé
AAC : autorisation d'accès compassionnel
AAP : autorisation d'accès précoce
Ac : anticorps
ACMG : *american college of medical genetics*
ADNtc : ADN tumoral circulant
AMM : autorisation de mise sur le marché
AMP : *association for molecular pathology*
ANSM : agence nationale de sécurité du médicament
ARNm : ARN messenger
ATU : autorisation temporaire d'utilisation
BAC : *bacterial artificial chromosome*
BER : base excision repair
BRCT : *BRCA1 C-Terminal domain*
CBNPC : cancer bronchique non à petites cellules
CBPC : cancer bronchique à petites cellules
CGI : *cancer genome interpreter*
CIViC : *clinical interpretation of variants in cancer*
COFRAC : comité français d'accréditation
CPC : cadre de prescription compassionnelle
CRISPR-Cas9 : *clustered regulatory interspaced palindromic repeats-CRISPR associated 9*
crRNA : *CRISPR RNA*
CRPCm : cancer de la prostate métastatique résistant à la castration
Del19 : délétions sans perte du cadre de lecture dans l'exon 19 du gène *EGFR*
dNTP : didésoxyribonucléotides triphosphates
DoCM : *database of curated mutations*
DSB : *Double Strand Break*
EGFR : récepteur au facteur de croissance épidermique
EMT : transition épithélio-mésenchymateuse
ESCAT : *ESMO scale for clinical actionability of molecular targets*
ESMO : *european society for medical oncology*
FDA : *food and drug administration*
FFPE : *formalin-fixed and paraffin-embedded*
FLGSA : *full-length gene splicing assay*
GIS : genomic instability status
GIST : tumeurs stromales gastro-intestinales

GSS : *genomic scar score*
HAS : haute autorité de santé
HER2 : récepteur 2 du facteur de croissance épidermique humain
HGMD : *human gene mutation database*
HGSOC : *high grade serous ovarian cancer*
HGVS : *human genome variation society*
HRD : *homologous recombination deficiency*
HRM : high resolution melting
ICGC : international cancer genome consortium
Ins20 : insertions dans l'exon 20 du gène *EGFR*
InDels : insertions et délétions de petite taille
iPARP : inhibiteurs de poly(ADP-ribose) polymérase
ITK : inhibiteur de tyrosine kinase
ITK-EGFR : inhibiteur de tyrosine kinase du facteur de croissance épidermique de type 1
LMC : leucémie myéloïde chronique
LAL : leucémie aiguë lymphoblastique
LAM : leucémie aiguë myéloïde
LOH : *loss of heterozygosity*
LOVD : *leiden open variation database*
LP-WGS : *low-pass whole genome sequencing*
LST : large scale state transition
Mb : megabase
MLPA : *multiplex ligation-dependent probe amplification*
NCCN : *national comprehensive cancer network*
NCI : *national cancer institute*
NGS : *next-generation sequencing*
NHEJ : *non-homologous end-joining*
NIH : *national institute of health*
NLS : *nuclear localization signals*
OMIM : *online mendelian inheritance in man*
PAM : *protospacer adjacent motif*
PARP : poly(ADP-ribose) polymérase
pb : paires de bases
PCa : cancer de la prostate
PCRd : PCR digitale
PGMC : plateformes de génétique moléculaire des cancers
Ph+ : présence du chromosome Philadelphie
RCP : réunion de concertation pluridisciplinaire

RGT : réarrangement génomique de grande taille
RH : recombinaison homologue
RNAseq : séquençage de l'ARN par NGS
ROS : *reactive oxygen stress*
RTK : récepteur tyrosine kinase
SGE : *saturation genome editing*
SMD/SMP : syndromes myélodysplasiques / syndromes myéloprolifératifs
SMRT : *single molecule real time sequencing*
SNP : *single nucleotide polymorphism*
SNUPE : *single nucleotide primer extension*
SNV : *single nucleotide variation*
SRE : *splicing regulatory elements*
SSP : survie sans progression
TAI : *telomeric allelic imbalance*
TCGA : *the cancer genome atlas*
TGS : *third generation sequencing*
TMB : *tumor mutational burden*
TNBC : *triple negative breast cancer*
tracrRNA : *transactivating CRISPR RNA*
UMI : *unique molecular identifiers*
VEGF : *vascular endothelial growth factor*
VSI : variant de signification inconnue
WES : *whole exome sequencing*
WGS : *whole genome sequencing*

INTRODUCTION

1. LA MEDECINE PERSONNALISEE EN ONCOLOGIE

1.1. Historique de la prise en charge des cancers et développement de la médecine personnalisée en oncologie

Le cancer, pathologie connue depuis l'Antiquité, est provoqué par une instabilité génétique entraînant la prolifération anarchique et excessive de cellules anormales dites tumorales.¹ Selon l'OMS, à l'échelle mondiale le cancer est la deuxième cause de décès avec environ 10 millions de morts par an et près d'un décès sur six dû au cancer²⁻⁴.

En 2018 en France, l'incidence des cancers est estimée à 382 000 nouveaux cas, la mortalité à 157 400 par an et le nombre de nouveaux cas de cancers détectés correspond à plus de 1 000 par jour. Les cancers les plus fréquents sont ceux de la prostate, du sein, du côlon-rectum et du poumon⁵. Ces dernières années, le taux d'incidence du cancer baisse ou se stabilise. (Figure 1) Le taux de mortalité diminue grâce au développement de la prévention, du dépistage et à l'émergence de traitements innovants ciblés, plus efficaces avec des effets secondaires mieux contrôlés.

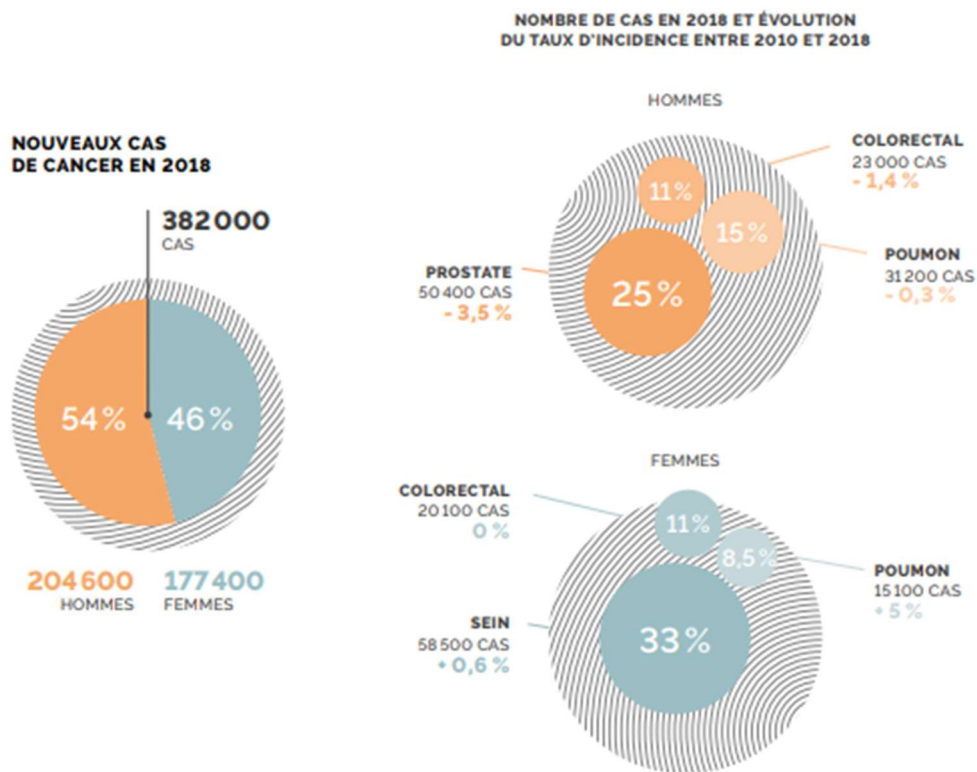


Figure 1 : Les cancers en chiffres en 2018 en France, données de l'INCa⁵

Concernant les traitements, 7,5 millions d'hospitalisations en lien avec le diagnostic, le traitement ou la surveillance d'un cancer, ont été enregistrées en 2019 (+ 1,3 % par

rapport à 2018)⁵. En un siècle, la prise en charge des cancers solides a connu une évolution constante liée au vieillissement de la population, aux évolutions diagnostiques mais aussi aux nombreuses innovations thérapeutiques consécutives à la connaissance intime des processus oncogéniques. La chirurgie a longtemps été la seule option thérapeutique, les techniques chirurgicales ont significativement évolué à la fin du 19^{ème} siècle et au début du 20^{ème} siècle après la mise au point de l'anesthésie générale en 1846. A la même époque, face au constat des récurrences locales et métastatiques après exérèse chirurgicale, apparaît la nécessité de développer des traitements complémentaires. Ainsi débute l'utilisation des traitements systémiques par chimiothérapies et des traitements locorégionaux par radiothérapie.

Au fil du temps est apparue l'évidence qu'il n'existe pas un cancer par organe mais de nombreux sous-types de cancers présentant chacun des anomalies moléculaires spécifiques au sein des cellules cancéreuses et de leur environnement. Au début des années 2000 : la cancérologie « d'organe » a donc évolué progressivement vers une cancérologie adaptée à l'individu et à sa pathologie. Ce changement de paradigme s'inscrit dans ce qui est communément nommé la « médecine personnalisée » ou « médecine de précision ». L'objectif étant de bloquer, contourner ou corriger les anomalies moléculaires impliquées dans le processus tumoral.⁶ Les grandes instances nationales et internationales définissent cette médecine personnalisée de différentes façons :

- Le National Institute of Health (NIH) américain définit une « pratique émergente de la médecine qui utilise le profil génétique des individus pour guider les décisions concernant la prévention, le diagnostic et le traitement des maladies »⁷
- L'objectif selon la Food and Drug Administration (FDA) américaine est d'obtenir « les meilleurs résultats médicaux en choisissant les traitements qui correspondent au profil génomique du patient ou à certaines caractéristiques de ses protéines, circulantes ou localisées à la surface cellulaire »⁸
- L'Institut National du Cancer (INCa) formalise la définition suivante : « La médecine de précision, également appelée médecine personnalisée, a ainsi pour objectif de proposer au patient un traitement adapté aux anomalies de sa tumeur. Elle repose actuellement sur deux types de traitements, les thérapies ciblées et l'immunothérapie spécifique. »⁹

Le but de cette médecine « sur mesure » est donc de classer les patients en sous-populations les plus homogènes possibles des points de vue de leur susceptibilité au développement d'une maladie, du comportement évolutif et/ou de la réponse thérapeutique (efficacité, tolérance). Les tests moléculaires visant à identifier par différentes techniques des anomalies moléculaires spécifiques, sont devenus ces dernières années de toute évidence indispensables pour cette médecine de précision dont les enjeux en oncologie sont :

- Dépister certains cancers
- Préciser le diagnostic et évaluer le pronostic
- Identifier les patients éligibles à des traitements spécifiques
- Caractériser les modes de résistance et les effets toxiques des médicaments
- Suivre l'évolution de la tumeur au cours des traitements et adapter en temps réel la thérapeutique
- Identifier les risques familiaux ou environnementaux de développer une pathologie cancéreuse.

En outre, si depuis les années 2000 la multiplicité des traitements offre de nombreuses opportunités, elle a indéniablement complexifié la décision thérapeutique en cancérologie. Ainsi, le décret du 21 mars 2007 a introduit une modification majeure dans la prise en charge des patients en créant des réunions de concertation pluridisciplinaire (RCP). La prise de décision collégiale permet une évaluation plus objective des situations, elle réduit le risque d'erreurs et permet d'uniformiser les pratiques. Aujourd'hui, s'organisent des RCP dites moléculaires, locales ou régionales voire nationales, avec des experts oncologues, biologistes moléculaires, anatomopathologistes, pharmaciens, pour utiliser à bon escient les données de génomiques au bénéfice des patients.

1.2. Thérapie ciblée et biomarqueurs

L'efficacité des cytotoxiques généraux utilisés en chimiothérapies classiques, encore largement utilisées, exploitent la prolifération rapide des cellules tumorales en bloquant la division cellulaire ou en induisant l'accumulation de dommages de l'ADN, conduisant à leur mort. Ces agents chimiothérapeutiques agissent donc préférentiellement sur les cellules tumorales mais impactent également les cellules saines à capacités proliférantes, expliquant les nombreux effets secondaires observés. Les thérapies dites « ciblées » ont pour objectif de bloquer la croissance ou la propagation tumorale, en interférant avec des anomalies moléculaires propres à la tumeur et absentes des tissus sains. Ces thérapies ciblées promettent d'améliorer l'efficacité thérapeutique en entraînant moins d'effets secondaires sur les tissus sains que les chimiothérapies classiques.¹⁰

Il existe deux approches pharmacologiques principales pour le traitement ciblé du cancer : les anticorps (Ac) monoclonaux (« -mab ») et les petites molécules (« -ib »). Les anticorps sont généralement caractérisés par une forte spécificité pour une cible précise offrant potentiellement un traitement peu toxique, ils ne peuvent atteindre que des cibles présentes à la surface externe des membranes ou extracellulaires. Leur nature protéique et leur poids moléculaire élevé imposent la contrainte d'une administration par voie parentérale (intraveineuse ou sous-cutanée). Depuis 1975, date de la mise en place de

production contrôlée d'anticorps monoclonaux par C.Milstein et G.Köhler¹¹, la preuve de concept de l'intérêt des anticorps thérapeutiques en cancérologie a été acquise dans les années 1990 par l'utilisation d'un anticorps monoclonal chimérique souris/homme dirigé contre le CD20 à la surface des lymphocytes B, le rituximab utilisé en oncohématologie. La première thérapie ciblée autorisée en France en cancérologie solide a été un anticorps monoclonal humanisé : le trastuzumab (Herceptin), dirigé contre le domaine extracellulaire du récepteur 2 du facteur de croissance épidermique humain ou HER2, une protéine de la famille des récepteurs à activité tyrosine kinase. L'action de cet anticorps inhibe la dimérisation de HER2, soit la forme active induite par les ligands, et par conséquent l'activation du récepteur. Le trastuzumab a obtenu l'Autorisation de Mise sur le Marché (AMM) en 2000 dans le traitement du cancer du sein métastatique pour des tumeurs présentant une surexpression de la protéine de HER2 à la surface cellulaire qui était associée jusqu'alors à un mauvais pronostic. Cette découverte majeure effectuée par M.Shepard, D.Slamon et A.Ulrich, honorée par le prix Lasker en 2019, a été le début d'un nombre croissant d'anticorps thérapeutiques en cancérologie mais aussi dans de nombreuses pathologies.¹²

Les petites molécules inhibitrices permettent quant à elles d'atteindre des cibles extracellulaires ou intracellulaires mais avec une sélectivité variable. Leur faible poids moléculaire leur permet des formes galéniques administrables par voie orale. En 2001, la première de ces petites molécules inhibitrices qui a obtenu l'AMM en France est l'imatinib (Glivec), indiqué dans le traitement des leucémies myéloïdes chroniques (LMC) présentant une anomalie moléculaire de type translocation t(9;22)¹³ appelée chromosome Philadelphie (Ph+). Cette translocation produit une protéine chimérique BCR-ABL intracellulaire ayant une activité tyrosine kinase constitutionnelle. L'imatinib est un inhibiteur de tyrosine kinase (ITK) qui entre en compétition avec l'ATP et inhibe l'activité tyrosine kinase constitutive de la protéine chimérique BCR-ABL et la prolifération des cellules hématopoïétiques exprimant BCR-ABL. Les résultats de l'étude initiale comparant imatinib et association interféron alpha - cytosine arabinoside ont montré des résultats spectaculaires avec une réponse cytogénétique complète passant de 14.5% à 76.2% confirmant tous les espoirs reposant sur les thérapies ciblées en oncologie.¹³ L'efficacité remarquable de cette molécule, qui de plus épargne les cellules saines n'exprimant pas BCR-ABL, illustre un siècle plus tard le concept de « magic bullet » postulé par P.Erlich le père fondateur de la chimiothérapie en oncologie. L'imatinib cible également d'autres récepteurs à activité tyrosine kinase dont c-KIT et PDGFR qui élargiront les AMM en particulier dans les tumeurs stromales gastro-intestinales (GIST).^{14,15} En 2005 et 2009, deux ITK, respectivement l'Erlotinib et le Gefitinib, des petites molécules ciblant le domaine tyrosine kinase intracellulaire du récepteur du facteur de croissance épidermique de type 1 (ITK-EGFR) avec mutations oncogéniques, obtiennent l'AMM dans les cancers

bronchiques non à petites cellules (CBNPC). L'usage de molécules thérapeutiques agissant préférentiellement sur des oncogènes activés illustre le concept des thérapies ciblées. Malgré une sélectivité préférentielle des cibles oncogéniques et un effet sur le contrôle ou la régression tumorale, l'administration d'ITK peut engendrer par manque de spécificité des effets secondaires altérant la qualité de vie des patients.

L'essor des thérapies dites ciblées utilisant des petites molécules inhibitrices a entraîné un véritable changement de paradigme dans la prise en charge des cancers dans les années 2000. Fin 2015, un rapport publié par l'INCa faisait état de 43 thérapies ciblées autorisées, dont 28 avaient fait leur entrée dans l'arsenal thérapeutique entre 2012 et 2015¹⁶, et ce nombre ne cesse de croître. Cette dynamique traduit, en termes cliniques, les progrès réalisés dans la compréhension et la caractérisation des altérations moléculaires des cellules cancéreuses. Ainsi, de nombreuses autres cibles oncogéniques ont été identifiées (BRAF, ALK, MET, ROS1, RET, NTRK, PIK3CA...) et les AMM évoluent constamment avec des élargissements multi-localisations. Aujourd'hui plus de cinquante thérapies ciblées sont utilisées en France en AMM ou en ATU (Autorisation Temporaire d'Utilisation), ces molécules peuvent être regroupées en plusieurs familles selon le(s) mécanisme(s) moléculaire(s) qu'elles ciblent (Tableau 1) (Figure 2).

Des altérations oncogéniques ou biomarqueurs moléculaires peuvent être associés (mais pas systématiquement) aux thérapies ciblées (Tableau 1). Ils permettent d'identifier les patients porteurs ou non d'une altération moléculaire qui guidera la décision de traitement (médecine stratifiée). On parle alors d'une approche théranostique, néologisme construit à partir des termes *thérapie* et *diagnostic*, qui correspond à une approche médicale visant à privilégier le développement simultané des aspects diagnostique et thérapeutique.

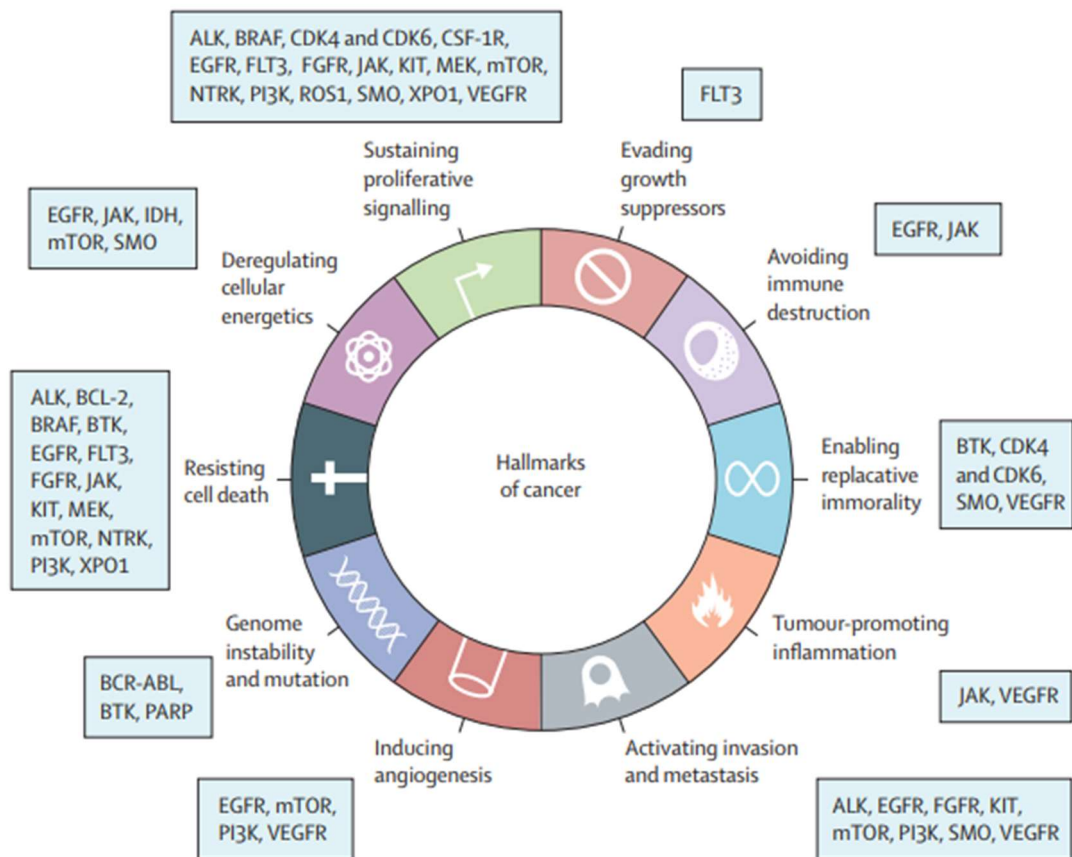


Figure 2 : Caractéristiques des cellules tumorales et cibles des petites molécules inhibitrices, d'après Bedard et al., 2020¹⁷

Famille	Mécanisme d'action	Molécules	BM	Indications	
Récepteurs de la famille HER	Ac monoclonaux	Cetuximab	Oui	CCR métastatique, cancers ORL	
		Panitumumab	Oui	CCR métastatique	
		Trastuzumab	Oui	Cancer du sein HER2+, cancer de l'estomac HER2+	
		Pertuzumab	Oui	Cancer du sein HER2+ en association au trastuzumab	
	Ac bispécifiques	ITK	Amivantamab	Oui	CBNPC métastatique EGFR muté par insertion dans l'exon 20
			Lapatinib	Oui	Cancer du sein métastatique HER2+
			Gefitinib	Oui	CBNPC EGFR muté
			Erlotinib	Oui	CBNPC EGFR muté, cancer du pancréas
			Afatinib	Oui	CBNPC EGFR muté
			Dacomitinib	Oui	CBNPC EGFR muté
			Osimertinib	Oui	CBNPC EGFR muté
			Poziotinib	Oui	CBNPC métastatique EGFR muté par insertion dans l'exon 20
			Mobocertinib	Oui	CBNPC EGFR muté par insertion dans l'exon 20
Voies de signalisation et seconds messagers	Inhibiteur de BRAF	Vémurafenib	Oui	Mélanome BRAF muté, autres tumeurs BRAF mutées	
		Dabrafenib	Oui	Mélanome BRAF muté, CBNPC BRAF muté	
		Encorafenib	Oui	Mélanome BRAF muté, CCR BRAF muté	
	Inhibiteur de KRAS	Sotorasib	Oui	CBNPC métastatique KRAS p.(G12C) muté	
			Oui		
	Inhibiteur de MEK	Trametinib	Oui	En association aux inhibiteurs de BRAF dans les tumeurs BRAF mutées, en monothérapie dans le cancer de l'ovaire séreux de bas grade	
			Oui	Mélanome BRAF muté en association à l'encorafenib	
			Oui	Mélanome BRAF muté en association au vemurafenib	
	Inhibiteurs de la voie PI3K/PDEN/AKT/mTOR	Everolimus	Non	Cancer du sein métastatique, cancer du rein métastatique	
			Non	Cancer du rein avancé, lymphomes à cellules du manteau	
			Oui	Cancer du sein métastatique avec mutation PI3K	
			Non	LLC (2 ^{ème} intention ou 1 ^{ère} intention avec délétion 17p ou mutation TP53), lymphome folliculaire	
	Inhibiteurs ALK, ROS	Crizotinib	Oui	CBNPC avancé ou métastatique	
			Oui		
			Oui		
			Oui		
			Oui		
	ITK BCR-ABL +/- inhibiteur KIT et PDGFR	Imatinib	Oui	LMC Ph+, LAL Ph+, GIST CD117+, SMD / SMP PDGFR+, dermatofibrosarcome protuberans	
			Oui	LMC Ph+	
			Oui	LAL Ph+	
			Oui	LMC Ph+	
			Oui	LAL Ph+ avec mutation T315I, LMC Ph+	
	ITK Bruton (BTK)	Ibrutinib	Non	Lymphomes à cellules du manteau, LLC, macroglobulinémie de Waldenström	
			Non	LLC	
	Inhibiteur BCL2	Venetoclax	Non	LLC	
	Inhibiteur FLT3	Gilteritinib	Oui	LAM avec mutation FLT3	
	ITK FGFR	Pemigatinib	Oui	Cholangiocarcinome avancé ou métastatique avec fusion FGFR2	
Oui			Carcinome urothélial avec altération FGFR2 ou FGFR3		
ITK JAK	Ruxolitinib	Non	Myélofibrose primitive ou secondaire à une polyglobulie de Vaquez ou une thrombocytemie essentielle		
Inhibiteur RET	Pralsetinib	Oui	CBNPC avec fusion RET, carcinome médullaire de la thyroïde avec altération RET		
		Non	Cancer médullaire de la thyroïde		
Inhibiteur TRK	Larotrectinib	Oui	Tumeurs solides avec fusion NTRK		
		Oui			
Inhibiteur voie Hedgehog	Sonidegib	Non	Carcinome basocellulaire localement avancé ou métastatique		
		Non			
Angiogénèse	Inhibiteur VEGF (molécules pour la plupart multikinases)	Selinexor	Non	Lymphome B diffus de haut grade à grandes cellules, myélome multiple	
		Bevacizumab	Non	CBNPC, CCR métastatique, cancer de l'ovaire, cancer du rein métastatique, cancer du sein métastatique, cancer du col de l'utérus	
		Ramucirumab	Non	CCR métastatique, CBNPC, cancer gastrique	
		Axitinib	Non	Cancer du rein métastatique	
		Pazopanib	Non	Cancer du rein métastatique, Sarcomes des tissus mous	
		Sunitinib	Non	Cancer du rein métastatique, GIST, tumeurs neuroendocrines du pancréas	
		Cabozantinib	Non	Cancer du rein métastatique, CHC, cancer médullaire de la thyroïde	
		Lenvatinib	Non	Cancer de la thyroïde métastatique, CHC	
		Sorafenib	Non	CHC, carcinome rénal avancé, carcinome thyroïdien différencié avancé	
		Regorafenib	Non	CCR métastatique, GIST métastatique, CHC	
Activité enzymatique	IDH1	Ivosidenib	Oui	LAM avec mutation IDH1, Cholangiocarcinome avancé ou métastatique avec mutation IDH1, gliome de bas grade avec mutation IDH1	
Cycle cellulaire	Inhibition des CDK4/6	Palbociclib	Non	Cancer du sein avancé ou métastatique RH+ HER2-	
		Abemaciclib	Non		
		Ribociclib	Non		
Réparation de l'ADN	Inhibiteurs de PARP (iPARP)	Olaparib	Oui	Cancer de l'ovaire avec mutation BRCA ou un phénotype HRD, cancer du sein avec mutation BRCA	
		Niraparib	Non	Cancer de l'ovaire	
		Rucaparib	Non	Cancer de l'ovaire	
		Talazoparib	Oui	Cancer du sein avec mutation BRCA	

Tableau 1 : Les thérapies ciblées autorisées en France en oncologie regroupées par mécanismes d'action (liste non exhaustive)^{9,18-20}

BM : thérapies ciblées disposant d'un biomarqueur associé à leur prescription, CCR : cancer colo-rectal, CBNPC : cancer bronchique non à petites cellules, LLC : leucémie lymphoïde chronique, LMC : leucémie myéloïde chronique, LAL : leucémie aiguë lymphoblastique, LAM : leucémie aiguë myéloïde, CHC : carcinome hépatocellulaire

1.3. Connaissance des processus tumoraux pour identifier les biomarqueurs d'intérêt

Dans ce contexte d'individualisation thérapeutique, il est aujourd'hui nécessaire d'identifier des biomarqueurs prédisant la réponse de l'individu à un traitement. C'est la compréhension des caractéristiques moléculaires des cellules tumorales (Figure 2) et des mécanismes de tumorigénèse qui permet d'identifier ces potentiels biomarqueurs.

Il existe une régulation complexe assurant l'intégrité du génome et sa stabilité au cours de l'évolution. L'instabilité génétique remet en cause cette intégrité et va fortement contribuer au processus de tumorigénèse.¹ Cette instabilité génétique peut avoir deux origines (Figure 3) :

- Les agents mutagènes exogènes (rayons UV, tabac, produits chimiques, chimiothérapies, radiothérapies) et endogènes (réactifs oxygénés)^{21,22}
- Les altérations des mécanismes de réparation de l'ADN²².

L'instabilité génétique entraîne une accumulation d'anomalies génétiques (Figure 3) qui peuvent être des mutations ponctuelles au niveau nucléotidique, des perturbations d'ordre chromosomique (aneuploïdies, délétions, inversions, translocations, amplifications) ou des modifications de la régulation épigénétique. Dans les tumeurs solides, 40 à 100 mutations peuvent être présentes dans la séquence codante des gènes mais seulement 5 à 15 vont avoir un réel impact sur la biologie de la cellule tumorale.²³ Les mutations tumorales conférant un avantage prolifératif sélectif (drivers oncogéniques) peuvent ainsi constituer des biomarqueurs pour les thérapies ciblées, au contraire des mutations passagers sélectivement neutres ou provoquant un désavantage tel que la sénescence ou la mort cellulaire.²⁴ Le nombre de ces mutations peut également constituer un biomarqueur, c'est le principe de la charge mutationnelle ou *tumor mutational burden* (TMB) basée sur l'hypothèse que le génome tumoral est un driver de l'immunité antitumorale par l'expression potentielle de néoantigènes qui peuvent être reconnus par les lymphocytes T. Le TMB est défini par le nombre de mutations acquises présentes dans les régions codantes de l'exome somatique (exprimé en nombre de mutations par megabase (Mb)). Ce TMB est un biomarqueur quantitatif associé, selon les tumeurs, à une possibilité de réponse à l'immunothérapie (inhibiteurs de checkpoints immunitaires).^{25,26} On peut cependant noter que peu de néoantigènes efficaces sont produits par les tumeurs, ainsi l'intérêt du TMB reste en discussion (Figure 4).

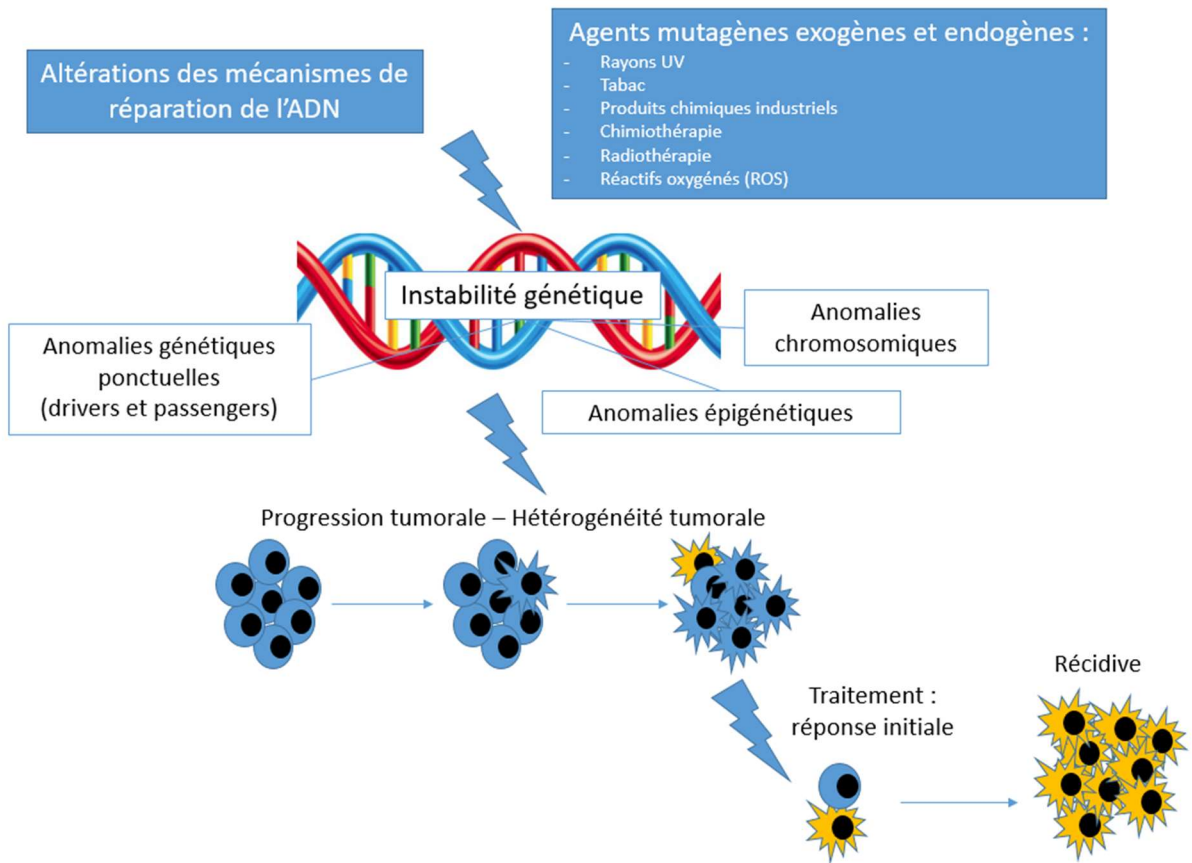


Figure 3 : Causes et conséquences de l'instabilité génétique

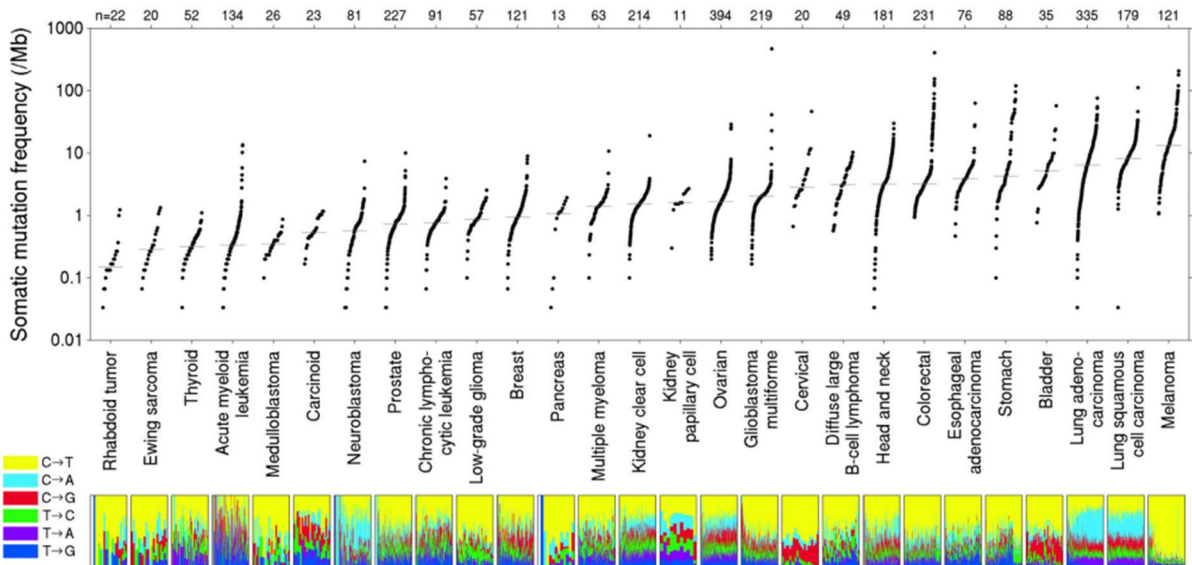


Figure 4 : Charge mutationnelle par type de pathologie, d'après Lawrence et al., 2013²⁷

1.4. Evolution technologique pour la détection des biomarqueurs en cancérologie : la généralisation du séquençage haut débit

Choisir et mettre en place les technologies indispensables pour une détection fiable, sensible et spécifique de ces biomarqueurs est une réflexion quotidienne dans les laboratoires de génétique moléculaire des cancers. En effet, le besoin constant des nouveaux biomarqueurs moléculaires pour une médecine « faite sur mesure », impose une veille technologique permanente liée à l'abondance de connaissances et d'innovations technologiques ces dernières années, particulièrement en génétique et en génomique qui occupent une place prépondérante. Par ailleurs, on peut noter que la quantité croissante de données génétiques nécessaires et générées implique une adaptation continue des besoins informatiques dans les laboratoires de génétique moléculaire. Depuis sa création en 2005, l'INCa soutient la structuration de la génétique moléculaire en oncologie. Ainsi, 28 plateformes de génétique moléculaire des cancers (PGMC) réparties sur l'ensemble du territoire ont été labellisées par l'INCa. Celles-ci permettent d'offrir à tous les patients les analyses de génétique moléculaire indispensables pour toutes les localisations concernées. La répartition territoriale en France est ainsi suffisante pour que les prélèvements tumoraux parvenant dans les laboratoires habituels d'anatomopathologie ou d'hématocytologie puissent être analysés dans des délais compatibles avec la prise en charge thérapeutique des patients. Cette initiative et la répartition géographique des PGMC labellisées INCa en France a été saluée pour son efficacité opérationnelle par l'ensemble de la communauté scientifique internationale.

Les PGMC réalisent les tests moléculaires :

- Déterminant l'accès à une thérapie ciblée ;
- Orientant le processus diagnostique ;
- Participant au diagnostic en complémentarité de paramètres cliniques, morphologiques, biologiques ;
- Orientant la stratégie de traitement du patient ;
- Permettant le suivi de la maladie résiduelle.

En 2005, les choix de thérapies ciblées étant restreints à quelques molécules et donc à quelques biomarqueurs, les PGMC utilisaient une multitude de techniques dites « ciblées ». On distingue les méthodes ciblées indirectes et les méthodes de séquençage ciblées. Concernant les approches ciblées indirectes, on peut citer parmi d'autres la PCR quantitative (ou PCR temps réel), la fusion haute résolution ou *high resolution melting* (HRM), le Snapshot, l'HPLC dénaturante, l'analyse MLPA (*multiplex ligation-dependent probe amplification*). Pour les approches de séquençage, le séquençage historique par arrêt d'élongation dit de Sanger et depuis les années 2000 le pyroséquençage, ont été largement utilisés. Une rupture technologique dans l'analyse des génomes a été matérialisée en 2008 par la caractérisation du génome de J.Watson par séquençage haut débit²⁸. Depuis cette

date, des innovations technologiques successives ont révolutionné l'usage de la caractérisation des génomes au sein des laboratoires au bénéfice des patients. Les demandes en génétique constitutionnelle ont rapidement imposé le séquençage massif ou NGS pour « *next generation sequencing* » comme une approche incontournable de caractérisation des génomes. Parallèlement, les développements et la disponibilité croissante des thérapies ciblées en oncologie depuis les années 2014/2015 ont rendu incontournable l'usage de ces méthodologies.

Suite à ces développements, selon l'INCa, en 2016, 83 000 patients ont pu bénéficier d'un test pour l'accès à une thérapie ciblée et 35 000 tests ont été réalisés par NGS.¹⁶ L'analyse de l'ADN par NGS peut-être plus ou moins étendue, allant d'un panel restreint de gènes (généralement limité aux régions codantes et jonctions intron-exon) à l'ensemble de l'exome « *whole exome sequencing* » (WES) ou même du génome entier « *whole genome sequencing* » (WGS). Le NGS permet également le séquençage plus ou moins étendu de l'ARN, on parle de RNAseq qui a remplacé progressivement les technologies *microarrays*. (Figure 5) Les développements de plus en plus sophistiqués des méthodes de séquençage permettent d'aborder l'épigénétique, la caractérisation globale des séquences promotrices, l'étude de la conformation chromatinienne ou plus récemment d'aborder la génomique sur cellule unique. Pour encadrer l'utilisation du NGS en oncologie, des listes minimales de gènes pour les analyses tumorales (tumeurs solides et cancers hématologiques) par localisation ont été définies initialement par un groupe d'experts coordonné par l'INCa. Ces listes sont en constante évolution en fonction des nouveaux schémas thérapeutiques mis en place.²⁹

Le NGS a transformé radicalement le criblage moléculaire dans le cadre de l'activité diagnostique en oncologie car il présente les avantages majeurs suivants :

- La détection multiple des mutations somatiques, c'est-à-dire présentes uniquement dans la tumeur, à très faible fréquence allélique.
- La capacité de multiplexage permettant l'analyse simultanée :
 - ✓ d'un grand nombre de gènes (exons ou gènes entiers) offrant une recherche exhaustive des cibles donnant accès aux thérapies ciblées
 - ✓ d'un grand nombre de patients en parallèle grâce aux codes-barres ou barcodes qui sont des séquences uniques à chaque échantillon intégrées en début d'analyse NGS.
- L'économie du matériel biologique : quelques nanogrammes d'ADN suffisent pour séquencer une multitude de gènes, avantage capital pour des prélèvements tumoraux précieux et souvent uniques et exigus.
- Le NGS permet en outre l'analyse moléculaire de l'ADN tumoral circulant (ADNtc) dans les fluides biologiques comme le sang, l'urine, le liquide céphalo-rachidien entres autres. L'ADNtc pouvant constituer une alternative lorsque le matériel

tumoral n'est pas disponible (prélèvement non réalisable, tissu tumoral altéré ou non disponible, patient refusant les biopsies), pour suivre la maladie résiduelle ou détecter une progression précoce (on parle de « *monitoring* »).³⁰⁻³²

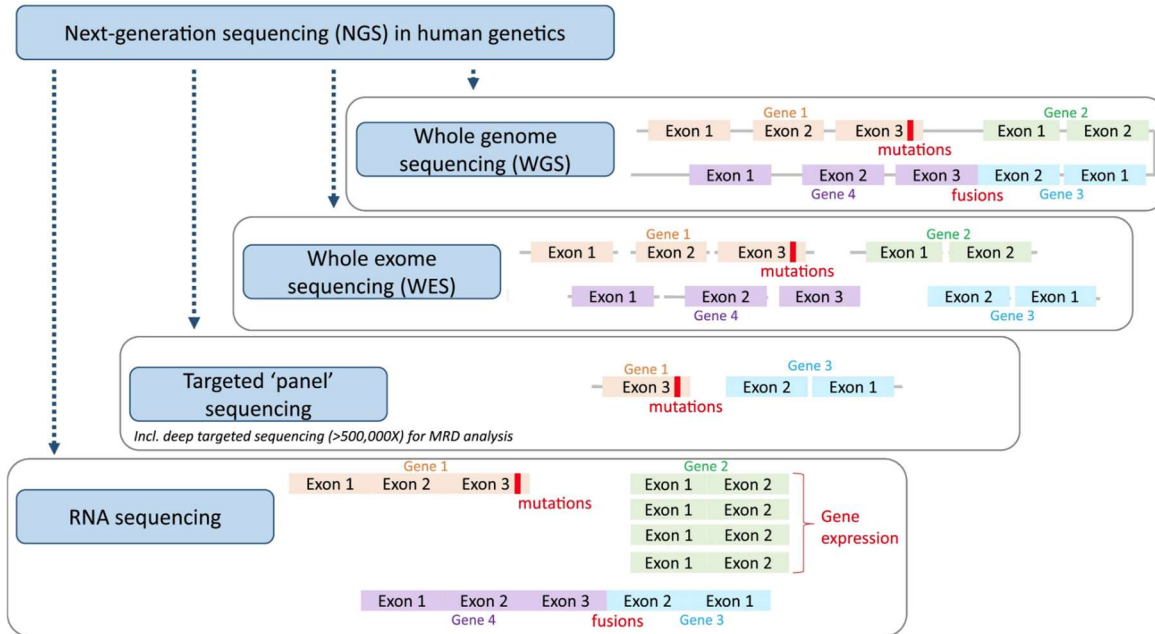


Figure 5 : Les applications du NGS en génétique humaine³³

Deux technologies NGS (Ion Torrent et Illumina) dominent aujourd'hui le marché et sont basées sur le principe de séquençage par synthèse, c'est-à-dire une méthode de séquençage durant laquelle un brin complémentaire est synthétisé par addition successive de désoxyribonucléotides triphosphates à partir d'un brin matrice et d'une amorce complémentaire. Ces deux technologies NGS présentent quatre étapes communes présentées dans la Figure 6 ³³ :

- L'extraction des acides nucléiques à partir d'échantillons de diverses matrices biologiques.
- La préparation de bibliothèques : il s'agit d'une multitude de fragments d'ADN ou d'ARN, générés par coupures aléatoires soit enzymatiques soit mécaniques, et sur lesquels sont fixés des séquences connues appelées « adaptateurs » et des identifiants appelés « barcodes » permettant de séquencer de nombreux échantillons simultanément. Il existe deux stratégies résumées dans le Tableau 2 : les bibliothèques par capture (hybridation de l'ADN à des sondes couvrant les régions cibles) et les bibliothèques par amplicons (amplification multiplexe des régions d'intérêt par PCR grâce à des lots d'amorces). Ces bibliothèques sont ensuite purifiées.
- L'amplification clonale : les bibliothèques sont amplifiées par PCR afin de générer suffisamment de matériel pour l'étape suivante, chaque technologie propose une amplification clonale différente (Tableau 2). Pour la technologie Ion Torrent, cette étape est une PCR en émulsion. La majorité des gouttelettes aqueuses après amplification contient une microbille recouverte d'un fragment unique d'ADN qui sera séquencé. Pour la technologie Illumina, l'amplification est réalisée par PCR en pont ou « *bridge PCR* » sur le support solide (lame de verre appelée « *flow cell* »). Un fragment s'hybride par son extrémité libre sur une amorce adjacente formant un pont, et un brin complémentaire est synthétisé. Les ponts d'ADN double-brin sont ensuite rompus par dénaturation, et un nouveau cycle d'amplification commence. À l'issue de l'amplification, les groupes de fragments d'ADN identiques, ou *clusters*, recouvrent la *flow cell*.
- Le séquençage : un brin complémentaire du fragment d'ADN immobilisé est polymérisé à partir des amorces universelles de séquençage. L'incorporation des nucléotides au niveau du brin d'ADN néo-synthétisé est suivie en temps réel par différents systèmes de mesure (Tableau 2) : l'acquisition d'un signal fluorescent émis par le nucléotide incorporé pour la technologie Illumina (système de détection optique), ou la mesure d'une baisse de pH consécutive à la libération d'un proton par liaison phosphodiester formée pour la technologie Ion Torrent.
- L'analyse bioinformatique des données brutes de séquençage : les données brutes de séquençage doivent être transformées en données de séquence, puis « filtrées » pour éliminer les séquences des adaptateurs, les séquences de mauvaise qualité,

les séquences polyclonales et les séquences de petite taille. Puis chaque séquence, appelée « *read* », est alignée sur le génome de référence pour mettre en évidence les variations. L'ensemble de ce traitement bioinformatique des données constitue le « *pipeline* bioinformatique » que chaque laboratoire de génétique moléculaire a la liberté de mettre en place en interne ou d'externaliser auprès de prestataires.

Les technologies NGS sont puissantes mais présentent quelques faiblesses, notamment la faible longueur des fragments séquencés (< 150 paires de bases (pb)). Les génomes contiennent souvent de nombreuses séquences répétées d'une longueur supérieure à celle des fragments lus par NGS, ce qui entraîne des difficultés dans l'assemblage sur le génome de référence. Par conséquent, un grand nombre de génomes séquencés sont disponibles sous une forme très fragmentée, en centaines ou milliers de contigs individuels (séquence génomique continue obtenue par l'assemblage des fragments chevauchants d'une librairie d'ADN). De même, les variants structuraux (réarrangements génomiques de plus de 50 pb) entre les génomes de différents individus sont difficiles à détecter par NGS, or ces réarrangements peuvent être des altérations génétiques d'importance clinique majeure. De plus les méthodes NGS ne permettent pas de caractériser des isoformes de transcrits générés par épissage alternatif de façon certaine. Enfin, du fait que le NGS nécessite une étape d'amplification par PCR, les régions génomiques avec un pourcentage élevé de GC sont souvent sous-représentées, car elles sont moins bien amplifiées par PCR. Ainsi, rapidement après l'apparition du NGS, des technologies dites de troisième génération (TGS, *third generation sequencing*) ou de lectures longues (*long read*) sont apparues et sont actuellement d'usages limités en pratique clinico-biologique. Elles se caractérisent par le séquençage, en temps réel de molécules uniques et permettent de générer des lectures d'une longueur de plusieurs kilobases voire même des centaines de kilobases.³⁴ Ces technologies ont été utilisées pour effectuer entre autres les séquençages télomère à télomère des chromosomes pour finaliser la caractérisation du génome humain.³⁵ La Figure 7 résume les avantages des technologies *long read* par rapport au NGS. Deux alternatives TGS sont actuellement disponibles :

- ✓ La technologie *single molecule real time sequencing* (SMRT) avec le système PacBio (société Pacific Biosciences)³⁶,
- ✓ La méthode nanopore avec l'équipement MinION (société Oxford Nanopore Technologies).³⁷

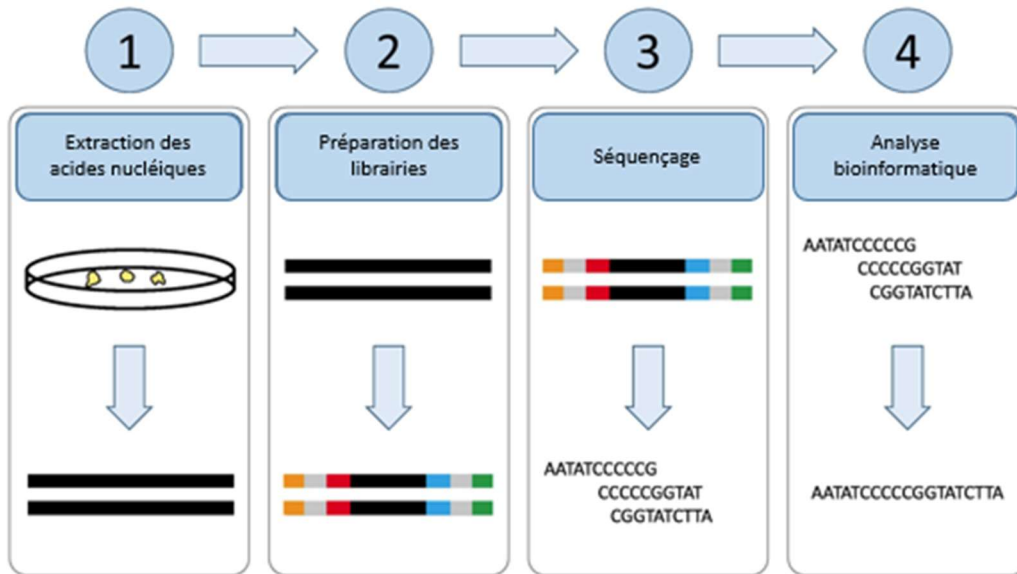


Figure 6 : Principe du NGS, adapté de Hess et al., 2020³³

1. Extraction des acides nucléiques à partir de l'échantillon. 2. Préparation des librairies pour obtenir une population de fragments d'ADN de longueur définie avec des séquences d'oligomères définies aux deux extrémités pour être compatibles avec la technique de séquençage appliquée. 3. Séquençage des librairies. 4. Analyse bioinformatique des données.

	Librairies et amplification clonale	Séquençage
	<p>PCR en émulsion</p>	<p>Séquençage par semi-conduction (ions H⁺)</p>
Ion Torrent	<p>Les différents fragments d'ADN simple brin sont fixés sur les microbilles via les adaptateurs (dans un rapport quantitatif permettant majoritairement la présence d'une seule molécule d'ADN sur chaque bille). Au sein de réacteurs individualisés, les fragments initiaux à la surface des billes vont subir différentes étapes d'amplification.</p>	<p>Détection de la libération des protons à chaque intégration d'un nucléotide au cours de l'élongation. L'incorporation séquentielle de chaque nucléotide permet de détecter ceux qui ont été incorporés au brin complémentaire.</p>
	<p>« Bridge » PCR ou PCR en pont</p>	<p>Séquençage avec détection optique (fluorescence)</p>
Illumina	<p>Hybridation d'un fragment (support solide), synthèse du brin complémentaire, hybridation de l'extrémité libre, création d'un double brin puis dénaturation et nouvelle synthèse aboutissant à la création d'une multitude de clusters de brins d'ADN identiques.</p>	<p>Chaque base émet un signal de fluorescence unique lorsqu'elle est ajoutée au brin en cours de synthèse, permettant de déterminer la séquence du brin complémentaire. Le fluorochrome est ensuite éliminé laissant un 3'OH libre pour l'addition d'un nouveau nucléotide.</p>

Tableau 2 : Principales différences entre les technologies NGS Illumina et Ion Torrent.^{38,39}

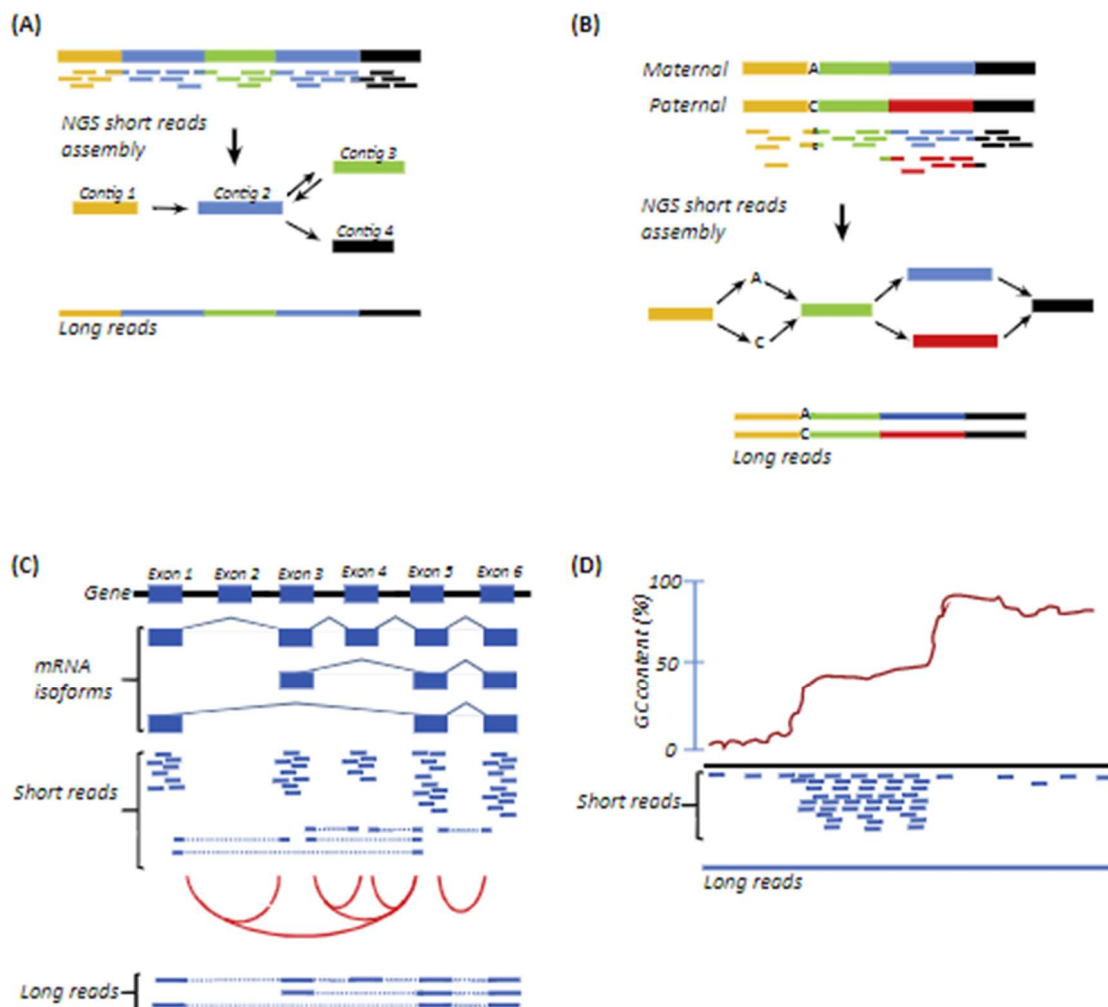


Figure 7 : Comparaison des performances des lectures courtes NGS et des lectures longues TGS, adapté de van Dijk et al., 2018³⁴

(A) Dans un génome, les régions répétées (en bleu) ne peuvent être positionnées à des endroits uniques avec des lectures courtes (NGS). Des lectures qui se situent à l'intérieur de ces régions seront assemblées en un seul contig. La région entre les répétitions (vert) peut être placée en amont ou en aval de ce contig bleu (illustré par les flèches en double-sens), ce qui génère une ambiguïté. Par contre, des lectures longues (long reads, TGS) qui traversent ces régions répétées ne laissent aucune ambiguïté. (B) Mise en phase des haplotypes. Les SNP (singlenucléotide polymorphism : A ou C) ou les variations plus importantes (rouge ou bleu) entre les allèles maternels et paternels situés trop loin les uns des autres pour être couverts par une seule lecture seront difficiles à mettre en phase avec l'allèle parental d'origine. Cela conduira à des trajectoires ambiguës qui aboutiront à des assemblages fragmentés. (C) Plusieurs transcrits (isoformes) peuvent être générés à partir d'un seul gène par épissage alternatif. Des lectures courtes de ces isoformes donneront des lectures à l'intérieur des exons présents dans le mélange (lignes continues) ainsi que des lectures qui chevauchent les jonctions entre les exons (lignes interrompues). Ainsi, les évènements d'épissage alternatif seront détectés ; les jonctions exon-exon détectées dans le mélange sont indiquées par des lignes rouges. Toutefois, les informations concernant les combinaisons des jonctions exon-exon dans les transcrits individuels manquent. Des lectures longues qui couvrent les transcrits entiers fournissent ces informations. (D) Les méthodes NGS dépendent de l'amplification par PCR, ce qui introduit des biais dans des régions pauvres en GC (à gauche de la courbe) ou très riche en GC (à droite). Par conséquent, ces régions seront peu couvertes. Les technologies de TGS n'ont pas besoin de l'amplification PCR et présentent donc beaucoup moins de biais avec ce type de régions.

1.5. Interprétation des variants, bases de données et problématique des variants de signification inconnue

Les variations nucléotidiques par rapport au génome de référence mises en évidence par NGS, appelées « variants » ou plus communément « mutations », doivent être classées grâce à une interprétation basée sur divers arguments. Par définition, un variant pathogène est une altération génétique qui modifie la transcription, la stabilité des ARN et/ou la fonction des protéines, et qui est responsable d'une maladie ou augmente la susceptibilité ou la prédisposition à la développer. À l'opposé, un variant bénin est une variation nucléotidique sans effet pathogène identifié.

L'enjeu en génétique des cancers est triple :

- définir l'impact des variants sur le processus tumoral,
- caractériser les cibles thérapeutiques,
- déterminer les anomalies conférant un potentiel risque héréditaire, périmètre d'une discipline appelée oncogénétique.

Une harmonisation internationale des pratiques est indispensable pour permettre une exploitation et un partage efficace des informations génomiques. Celle-ci est conditionnée en partie par une nomenclature uniforme, fondée sur un ensemble de critères normalisés. Cette nomenclature standard des variants génétiques est maintenue et mise à jour par la *Human Genome Variation Society* (HGVS).⁴⁰ D'autre part, une interprétation consensuelle des variants de génétique moléculaire est nécessaire, elle repose sur un faisceau d'arguments ayant un poids plus ou moins important dans l'interprétation du variant mis en évidence. Cet argumentaire correspond aux recommandations internationales publiées par l'*American College of Medical Genetics* (ACMG) et l'*Association for Molecular Pathology* (AMP) en 2015 et 2016. Cette interprétation est sous la responsabilité exclusive du biologiste et consiste à combiner ces arguments pondérés afin d'assigner une des 5 classes suivantes au variant étudié⁴¹⁻⁴³:

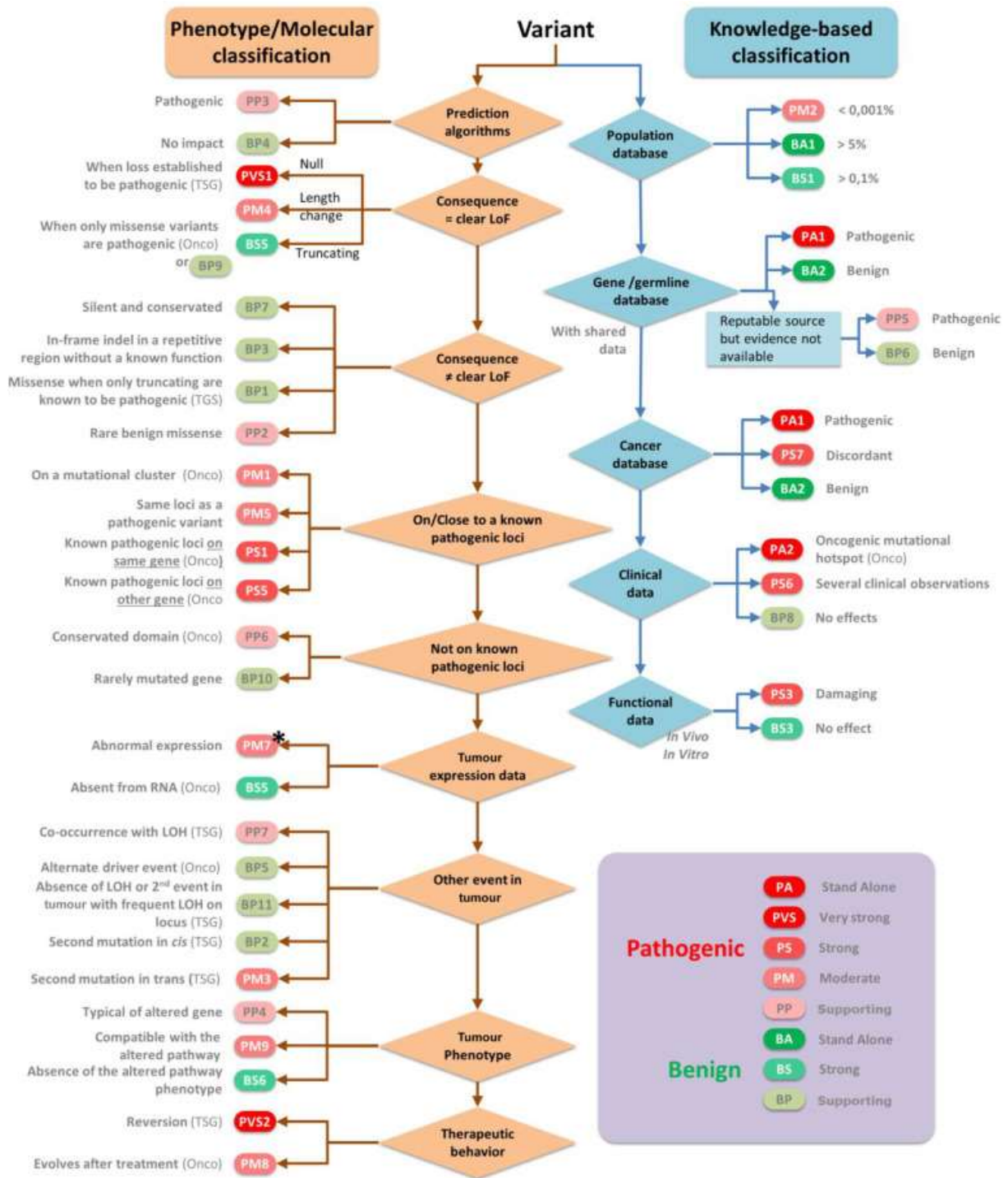
- Classe 1 : Variant bénin
- Classe 2 : Variant probablement bénin
- Classe 3 : Variant de signification inconnue (VSI)
- Classe 4 : Variant probablement pathogène
- Classe 5 : Variant pathogène

Lorsqu'un variant est assigné en classe 3, cela ne permet ni de l'utiliser dans la prise en charge éventuelle du patient et de sa famille, ni d'exclure son caractère causal dans la pathologie du patient. Le variant peut faire l'objet d'études complémentaires visant à apporter des arguments permettant de le re-classer en variant d'une autre classe (études fonctionnelles *in vitro*, étude de l'épissage, étude de ségrégation familiale).⁴⁴

Ces recommandations internationales de classification en 5 classes ont été établies pour la génétique constitutionnelle bénéficiant de réseaux scientifiques et de bases de

données historiques. La génétique tumorale ou génétique somatique est une discipline plus récente et de nombreux variants nouveaux ou exclusivement somatiques sont trouvés. En 2021, des recommandations ont été publiées pour standardiser la classification des altérations somatiques dans les tumeurs solides et tumeurs malignes hématologiques⁴⁵ (Figure 8). De même, les bases de données se mettent en place progressivement. Les bases de données utilisées en génétique somatique sont de deux types :

- Les bases de données de populations contrôles : elles permettent de connaître la fréquence du variant dans la population générale (une fréquence importante est en faveur du caractère bénin du variant). Les bases 1000 Genomes⁴⁶, ESP (*NHLBI Exome Sequencing Project*)⁴⁷, FLOSSIES (*Fabulous Ladies Over Seventy*)⁴⁸ et GnomAD (*Genome Aggregation Database*)⁴⁹ en sont des exemples (Tableau 3).
- Les bases de données de patients : elles se sont mises en place progressivement parallèlement au développement du NGS. En génétique tumorale, les plus consultées sont les suivantes (Tableau 4) : COSMIC⁵⁰, Clinvar⁵¹, *The Cancer Genome Atlas* (TCGA)⁵², OncoKB⁵³, My cancer genome⁵⁴, HGMD (*Human Gene Mutation Database*)⁵⁵, LOVD (*Leiden Open Variation Database*)⁵⁶, OMIM (*Online Mendelian Inheritance in Man*)⁵⁷, DoCM (*Database of Curated Mutations*)⁵⁸, CIViC (*Clinical Interpretation of Variants in Cancer*)⁵⁹, CGI (*Cancer Genome Interpreter*)⁶⁰. Certaines bases sont spécifiques de gènes ou spécifiques de certaines pathologies et ne peuvent pas être citées exhaustivement. Il existe également de nombreuses bases de données locales propres à chaque laboratoire qui restent à exploiter. Certains variants sont rapportés dans les bases de données avec des informations erronées ou incomplètes, il est donc indispensable de vérifier la source des informations et de les croiser.



* **Oncogene:** over-expression of the gene or co-occurrence with amplification / LOH
TSG: Homogeneous abnormal expression in favour of the accumulation of the mutant protein or expression loss in favour of bi-allelic loss has been reported in several cases with alteration

Figure 8 : Arbre décisionnel pour la classification des variants somatiques, d'après Koeppl, F. et al.⁴⁵

Base de données	Contenu – Spécificités de la base de données
1000 Genomes	<ul style="list-style-type: none"> - 3511 génomes de sujets non apparentés issus de 26 populations différentes - principalement adaptée à l'interprétation de variants non exoniques
ESP	6503 exomes de sujets non apparentés, les individus séquencés sont stratifiés en deux populations : EA (european-americans) et AA (african-americans)
FLOSSIES	variants identifiés dans une population contrôle de 10 000 femmes de plus de 70 ans pour 27 gènes impliqués (ou possiblement impliqués) dans les prédispositions au cancer du sein
GnomAD	<ul style="list-style-type: none"> - 123 136 exomes et 15 496 génomes entiers (permettant d'appréhender la fréquence des variants non-exoniques de manière plus précise que 1000 Genomes) de sujets non apparentés issus de diverses études populationnelles ou spécifiques de certaines maladies - les données incluent les données des bases ExAC, 1000 Genomes et partiellement d'ESP - intègre des populations qui ne sont pas représentées dans les autres bases de données, comme la population juive ashkenaze - base de données la plus puissante actuellement en termes de nombre d'individus séquencés - contient possiblement des données génomiques d'individus avec des maladies Mendéliennes ou complexes, comme des cancers, des maladies psychiatriques, des démences, ou des maladies inflammatoires. Le nom des séries incluses est consultable sur le site internet mais l'information sur la présence de malades n'est pas précisée.

Tableau 3 : Bases de données de populations contrôles et leurs spécificités

Base de données	Contenu – Spécificités de la base de données
COSMIC	Base internationale la plus complète pour explorer l'impact des mutations somatiques dans le cancer humain (>10 000 000 de mutations cataloguées). Les mutations sont répertoriées avec une prédiction <i>in silico</i> d'impact fonctionnel, une distribution par tissu, les voies de signalisation impactées et des références bibliographiques. Cette base ne propose pas directement d'impact clinique, elle est utile par son exhaustivité. Il existe deux types de données dans COSMIC : les données évaluées par un groupe défini d'experts et les données issus d'autres bases (TCGA ⁵² , ICGC Data Portal ⁶¹ par exemple).
ClinVar	Base rapportant des variants constitutionnels et des variants somatiques avec leur annotation selon les 5 classes recommandées. ^{41-43,51} Les publications prises en compte dans l'annotation des variants sont exhaustivement rapportées et un niveau de confiance dans l'interprétation est fourni pour chaque variant par un groupe d'experts issus de laboratoires cliniques ou de laboratoires de recherche. ⁵¹
TCGA	TCGA est un programme lancé en 2005 par le National Cancer Institute (NCI) et le National Human Genome Research Institute. Ce projet a permis la caractérisation moléculaire de plus de 20 000 échantillons primaires de cancer et d'échantillons normaux appariés couvrant 33 types de cancer. Plus de 2,5 pétaoctets de données génomiques, épigénomiques, transcriptomiques et protéomiques ont été générées. Ces données qui ont déjà permis d'améliorer le diagnostic, le traitement et la prévention du cancer, sont accessibles au public et utilisées par tous les membres de la communauté des chercheurs. Les variants sont rapportés mais non annotés. ⁵²
OncoKB	Données théranostiques en oncologie : pour chaque altération spécifique de 682 gènes du cancer, sont répertoriés l'effet biologique, la prévalence, les informations pronostiques et les implications en termes de traitement. La base est développée et maintenue par le groupe Knowledge Systems du Marie Josée and Henry R. Kravis Center for Molecular Oncology du Memorial Sloan Kettering Cancer Center. Les informations proviennent de diverses sources, (FDA, NCCN, ASCO, ClinicalTrials.gov et littérature scientifique). ⁵³
My Cancer Genome	Données théranostiques en oncologie : contient des données sur l'impact clinique de 16 871 biomarqueurs moléculaires et sur l'utilisation de 2 861 médicaments dans 955 types de cancer. Ces informations sont issues de la FDA, des directives du NCCN et d'autres sociétés professionnelles, de 9 809 essais cliniques, de publications. ⁵⁴
HGMD	Collection de données sur les mutations germinales dans les gènes associés aux maladies héréditaires humaines. Actuellement, cette base répertorie 210341 variants dans 8780 gènes. ⁵⁵
LOVD	Base de variants germinaux rapportant une classification clinique, le phénotype et les références (bases de données et publications). LOVD contient plus de 370000 variants uniques dans 23 030 gènes. ^{56,62}
OMIM	Base mise à jour quotidiennement contenant des informations sur toutes les maladies mendéliennes connues et sur plus de 16 000 gènes. OMIM se concentre sur la relation entre le phénotype et le génotype. ⁵⁷
DoCM	DoCM est un dépôt de données qui facilite l'agrégation des informations sur les gènes et variants ayant un rôle pronostique, diagnostique, prédictif ou fonctionnel en cancérologie. Elle fournit des listes de variants facilement interprétables grâce à des liens directs vers les sources et publications. Elle contient actuellement 1364 variants affectant 132 gènes. ⁵⁸
CIVIC	Base alimentée par des experts pour l'interprétation clinique des variants en oncologie, décrivant la pertinence thérapeutique, pronostique, diagnostique des variants héréditaires et somatiques de tous types. Elle contient actuellement 1678 interprétations de pertinence clinique pour 713 variants affectant 283 gènes. ⁵⁹
CGI	CGI s'appuie sur des méthodes <i>in silico</i> (BoostDM et OncodriveMut) ainsi que sur les connaissances recueillies dans le domaine public pour annoter les altérations d'une tumeur selon plusieurs niveaux de preuve et prédire la réponse aux médicaments. La base contient 1574 biomarqueurs génomiques, 221 médicaments anti-cancéreux et 79 types de cancer. ⁶⁰

Tableau 4 : Bases de données patients et leurs spécificités

Lorsque les variants sont VSI ou absents des bases de données, l'argumentaire de la causalité prend en compte les éléments suivants⁴⁴ :

- Des données structurales liées à la nature du variant nucléotidique : variants non-sens, faux-sens ou synonymes mais aussi les changements de cadre de lecture (frameshift), les modifications des sites d'épissage, des codons d'initiation de la traduction, et les grands réarrangements (hors phase ou en phase).
- Des prédictions bio-informatiques : lorsque qu'un variant est mis en évidence, il est nécessaire d'interroger plusieurs logiciels de prédiction, tels que SIFT⁶³, MutationTaster^{64,65}, Polyphen-2⁶⁶ ou UMD-Predictor⁶⁷. Ces algorithmes attribuent un score et une classification en faveur de la probabilité du caractère délétère ou bénin d'un variant faux-sens en fonction de différents critères : conservation de l'acide aminé (AA) au cours de l'évolution, distance physico-chimique entre le résidu aminoacide sauvage et muté, position dans un domaine fonctionnel connu de la protéine. Certains outils comme Mutation Taster^{64,65} ou UMD-Predictor⁶⁷ intègrent à la fois la prédiction d'un effet de l'AA et d'un effet sur l'épissage. Cependant, pour évaluer un éventuel effet sur l'épissage, il est préférable d'utiliser des outils de prédiction spécifiques comme MaxEntScan⁶⁸, NNSPLICE⁶⁹, GeneSplicer⁷⁰, Human Splicing Finder⁷¹, spliceAI⁷², SPiP⁷³ ou encore SpliceLauncher⁷⁴. Il est essentiel de combiner les approches, c'est-à-dire d'interroger plusieurs outils et algorithmes, afin d'obtenir la meilleure sensibilité et la meilleure spécificité possible. La complémentarité des outils tient au fait qu'ils utilisent des algorithmes distincts et qu'ils ont été entraînés et validés sur des sets de variants « vrais positifs » et « vrais négatifs » distincts.
- Des données bibliographiques : une revue de la littérature concernant le variant à classer est nécessaire.
- Des critères cliniques : les informations cliniques et anatomocytopathologiques sont indispensables. En effet, une confrontation clinico-biologique dans le cadre d'une RCP avec le clinicien prenant en charge le patient et en présence d'un biologiste et d'un pharmacien est nécessaire afin de préciser la corrélation phénotype-génotype et d'étayer le choix thérapeutique le plus adapté, on parle de « RCP moléculaires ». Afin d'harmoniser l'interprétation des analyses moléculaires cliniquement pertinentes, la classification « *ESMO Scale for Clinical Actionability of molecular Targets* » (ESCAT)⁷⁵ a été proposée en 2018 par l'*European Society for Medical Oncology* (ESMO). L'ESCAT classe les altérations génomiques selon les niveaux d'évidence soutenant leur valeur en tant que cibles cliniques. Cette classification définit 6 classes dont l'ESCAT I où la thérapie est associée dans les essais à une amélioration de la survie et l'ESCAT II où la thérapie est associée à une réponse anti-tumorale dans le type de cancer étudié sans connaître l'amplitude du bénéfice. Au-delà (ESCAT III à VI), aucune preuve de bénéfice clinique n'est apportée. (Tableau 5)

- Les variants associés : au-delà de la cible théranostique, certaines mutations de résistance peuvent apparaître entraînant un switch thérapeutique (vers une nouvelle thérapie ciblée ou vers une chimiothérapie), les sites métastatiques peuvent présenter un profil moléculaire différent de la tumeur primitive expliquant une progression tumorale hétérogène. Parfois plusieurs tumeurs concomitantes ou certaines associations de mutations peuvent être prédictives de la réponse au traitement⁷⁶.
- Les données fonctionnelles disponibles : des études fonctionnelles *in vivo* ou *in vitro* montrant l'impact du variant sur le gène ou son produit peuvent constituer un argument de poids dans l'interprétation. Les types de tests fonctionnels sont nombreux. Il est à noter qu'il n'existe que très peu de tests fonctionnels concernant les gènes impliqués dans la cancérogénèse, qu'il s'agisse des oncogènes tels qu'*EGFR* dans les cancers bronchopulmonaires ou des suppresseurs de tumeurs tels que *BRCA1* et *BRCA2* dans les tumeurs du spectre du syndrome seins-ovaires, étudiés dans ce travail.

	ESCAT evidence tier	Required level of evidence	Clinical value class	Clinical implication
Ready for routine use	I: Alteration-drug match is associated with improved outcome in clinical trials	I-A: prospective, randomised clinical trials show the alteration-drug match in a specific tumour type results in a clinically meaningful improvement of a survival end point I-B: prospective, non-randomised clinical trials show that the alteration-drug match in a specific tumour type, results in clinically meaningful benefit as defined by ESMO MCBS 1.1 I-C: clinical trials across tumour types or basket clinical trials show clinical benefit associated with the alteration-drug match, with similar benefit observed across tumour types	Drug administered to patients with the specific molecular alteration has led to improved clinical outcome in prospective clinical trial(s)	Access to the treatment should be considered standard of care
Investigational	II: alteration-drug match is associated with antitumour activity, but magnitude of benefit is unknown	II-A: retrospective studies show patients with the specific alteration in a specific tumour type experience clinically meaningful benefit with matched drug compared with alteration-negative patients II-B: prospective clinical trial(s) show the alteration-drug match in a specific tumour type results in increased responsiveness when treated with a matched drug, however, no data currently available on survival end points	Drug administered to a molecularly defined patient population is likely to result in clinical benefit in a given tumour type, but additional data are needed	Treatment to be considered 'preferable' in the context of evidence collection either as a prospective registry or as a prospective clinical trial
Hypothetical target	III: alteration-drug match suspected to improve outcome based on clinical trial data in other tumour type(s) or with similar molecular alteration	III-A: clinical benefit demonstrated in patients with the specific alteration (as tiers I and II above) but in a different tumour type. Limited/absence of clinical evidence available for the patient-specific cancer type or broadly across cancer types III-B: an alteration that has a similar predicted functional impact as an already studied tier I abnormality in the same gene or pathway, but does not have associated supportive clinical data	Drug previously shown to benefit the molecularly defined subset in another tumour type (or with a different mutation in the same gene), efficacy therefore is anticipated for but not proved	Clinical trials to be discussed with patients
	IV: pre-clinical evidence of actionability	IV-A: evidence that the alteration or a functionally similar alteration influences drug sensitivity in preclinical <i>in vitro</i> or <i>in vivo</i> models IV-B: actionability predicted <i>in silico</i>	Actionability is predicted based on preclinical studies, no conclusive clinical data available	Treatment should 'only be considered' in the context of early clinical trials. Lack of clinical data should be stressed to patients
Combination development	V: alteration-drug match is associated with objective response, but without clinically meaningful benefit	Prospective studies show that targeted therapy is associated with objective responses, but this does not lead to improved outcome	Drug is active but does not prolong PFS or OS, probably in part due to mechanisms of adaptation	Clinical trials assessing drug combination strategies could be considered
	X: lack of evidence for actionability	No evidence that the genomic alteration is therapeutically actionable	There is no evidence, clinical or preclinical, that a genomic alteration is a potential therapeutic target	The finding should not be taken into account for clinical decision

Tableau 5 : Classification ESCAT. D'après Mateo et al., 2018⁷⁵

2. EGFR ET CANCERS BRONCHOPULMONAIRES : ENJEUX ET CHOIX TECHNOLOGIQUES

2.1. Les cancers bronchopulmonaires : généralités

En 2018, le cancer bronchopulmonaire est le 3^{ème} cancer le plus fréquent en France avec 46 300 nouveaux cas, dont 67% d'hommes et 33% de femmes, avec une forte progression chez les femmes ces dernières années. C'est la première cause de décès par cancer en France. En 2017, la prévalence de ce cancer est estimée à 169 718 personnes.⁵

Le premier facteur de risque de cancer bronchopulmonaire est le tabagisme actif, responsable de 75% d'entre eux. Le risque relatif de cancer du poumon chez le fumeur est multiplié par 10 à 15 selon la quantité fumée, et le risque est d'autant plus important que le tabagisme est débuté précocement. Le tabagisme passif ou environnemental correspondant à l'exposition à la fumée émise par les fumeurs est également incriminé. Un excès de risque de 20 à 26% est estimé pour un non-fumeur vivant avec un fumeur.⁷⁷⁻⁷⁹ D'autres facteurs de risque exogènes environnementaux ou professionnels sont rapportés : hydrocarbures aromatiques induisant un stress oxydatif, radioactivité, amiante, poussière de silice...⁷⁷

Depuis plus de 25 ans, le taux de survie nette à 5 ans augmente mais reste mauvais : 13% pour les personnes diagnostiquées entre 1989 et 1993 puis 17% entre 2005 et 2010 et enfin 20% entre 2010 et 2015.⁸⁰ L'amélioration peut être attribuée à l'augmentation de la détection des phases précoces, ainsi qu'aux progrès en termes de traitements tels que la chirurgie thoroscopique, la radiothérapie, les thérapies ciblées et, depuis 2015, l'immunothérapie.

Histologiquement, le cancer bronchopulmonaire constitue un groupe hétérogène de cancers qui est divisé en deux grandes catégories suivant le type cellulaire à partir duquel il se développe :

- Les cancers bronchiques à petites cellules (CBPC) représentant 15% des cancers du poumon.
- Les CBNPC, sur lesquels se concentrera ce travail, représentant plus de 80 % des cancers du poumon et regroupant les adénocarcinomes (60% des cas), les carcinomes épidermoïdes (30% des cas) et les carcinomes à grandes cellules (plus rares).

Depuis les années 2000, la prise en charge du CBNPC a connu d'importantes avancées thérapeutiques. La prise en charge des stades localisés (< IIIb selon la classification de l'*International Association for the Study of Lung Cancer*⁸¹) repose sur la chirurgie complétée par une chimiothérapie adjuvante en fonction des critères pronostiques. Pour les stades localement avancés non résécables, le traitement de référence est la radio-chimiothérapie

qui peut être suivie par un traitement d'entretien par immunothérapie anti-PD-L1 pendant un an (Durvalumab) pour les carcinomes avec expression de PD-L1 $\geq 1\%$ ^{82,83}. Pour les stades métastatiques (IV), la prise en charge repose sur les traitements systémiques seuls ou en association : thérapies ciblées, immunothérapie, chimiothérapie.⁸⁴

Malgré les avancées thérapeutiques récentes, le pronostic reste médiocre en raison notamment de la présence de tumeurs localement avancées ou métastatiques chez la majorité des patients au moment du diagnostic et par l'apparition de multiples résistances. Le pronostic est à la fois lié au stade et à la précocité du diagnostic, le cancer du poumon est généralement évoqué devant la présence ou la persistance de troubles respiratoires. Les pratiques de dépistage précoces restent malheureusement peu développées dans ce type de cancer et sont sans doute une perspective à développer dans les années à venir même si le conseil majeur en ce qui concerne la prévention du cancer du poumon reste l'arrêt du tabagisme actif et passif.

2.2. Hétérogénéité mutationnelle des CBNPC et opportunités thérapeutiques

Le CBNPC est un cancer associé à de nombreuses altérations génomiques. Mieux définir l'hétérogénéité génétique intratumorale et préciser la diversité des oncogènes activés dans les CBNPC est un préalable incontournable au succès des thérapies ciblées. En 2004, seules les mutations de sensibilité de l'*EGFR* aux ITK et les variants oncogéniques de *KRAS* étaient recherchés par des techniques ciblées. Les progrès techniques et le NGS permettent maintenant de mettre en évidence de nombreuses altérations génomiques en sus de *KRAS* et *EGFR* donnant potentiellement accès à des thérapies ciblées (Figure 9). Par ordre de fréquence, on peut citer :

- ***KRAS*** : 25% des CBNPC. Le plus souvent associées au tabagisme, les mutations de l'oncogène *KRAS* affectent principalement les codons 12 et 13 de l'exon 2 et sont décrites comme étant activatrices des voies de transduction intracellulaire RAS/RAF/MAPkinases^{85,86}. Il y a actuellement deux principales implications des mutations *KRAS* dans la prise en charge des CBNPC. D'une part, le sotorasib, petite molécule ciblant sélectivement la mutation *KRAS* p.Gly12Cys (ou G12C), a obtenu en 2021 une ATU en monothérapie dans le traitement des patients adultes atteints d'un CBNPC métastatique muté *KRAS* G12C et prétraités.⁸⁷ D'autre part, les altérations inactivatrices du gène suppresseur de tumeur *STK11/LKB1* associées à une mutation activatrice de *KRAS* sont

liées à une moindre efficacité des inhibiteurs de checkpoint immunitaires (immunothérapie).⁷⁶

- EGFR : 17% des CBNPC, cette proportion varie avec l'origine ethnique et le sous-type histologique, cette fréquence est par exemple de 40% chez les femmes asiatiques non fumeuses d'une quarantaine d'années.^{88,89} La première partie de ce travail est dédiée à l'activité théranostique et à la détection des altérations de l'EGFR.
- ALK : le gène ALK, codant pour un récepteur tyrosine kinase (RTK), est impliqué dans des réarrangements aboutissant à des protéines de fusion oncogéniques. Ces altérations sont rapportées dans 7% des CBNPC, et le plus souvent chez les patients jeunes, non fumeurs, de sexe féminin et présentant une histologie d'adénocarcinome.⁹⁰⁻⁹³ Les réarrangements avec le gène EML4 sont les plus fréquemment décrits dans les CBNPC.⁹⁴ Le premier inhibiteur d'ALK à avoir obtenu des autorisations est le crizotinib, petite molécule ciblant les tyrosines kinases ALK, ROS1 et MET. D'autres molécules sont aujourd'hui autorisées ou en cours d'étude.^{95,96}
- MET : l'oncogène MET code pour un RTK. Les altérations de MET représentent 6% des CBNPC et sont principalement des mutations entraînant un saut d'exon 14 (3%), des amplifications (3%) ou de très rares fusions avec des gènes partenaires. Des études précliniques et cliniques ont suggéré que l'activation de MET peut être un driver oncogénique ou un facteur de résistance acquise au(x) traitement(s). Les agents ciblant MET (ITK, petites molécules ou Ac plus ou moins sélectifs de MET) sont une stratégie de traitement prometteuse pour le CBNPC : certaines molécules possèdent déjà des autorisations et d'autres sont en développement.^{88,97,98}
- NF1 : Les mutations inactivatrices de NF1 sont présentes dans environ 4% des CBNPC. Le gène NF1 est un suppresseur de tumeur qui régule négativement la voie de signalisation RAS/MAPkinases ouvrant la voie aux thérapies ciblées inhibant cette voie.⁹⁹⁻¹⁰¹
- HER2, BRAF, RET : à une fréquence de 2%
- ROS1, NTRK1, PIK3CA, AKT1, MEK1 : à une fréquence de 1%

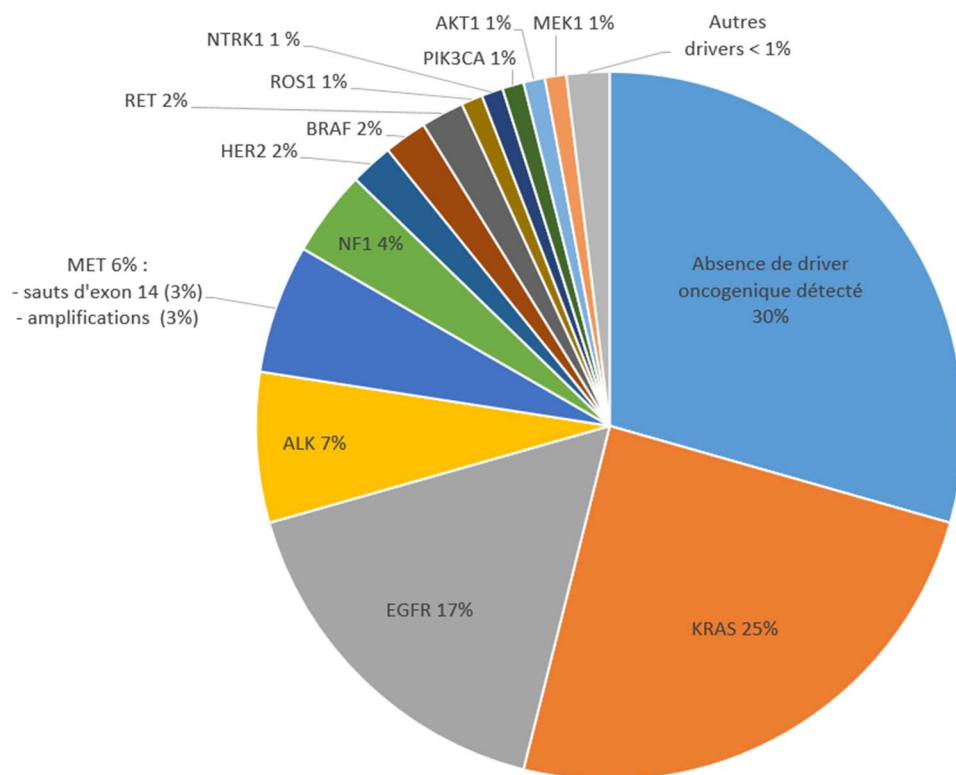


Figure 9 : Fréquence des altérations génomiques retrouvées dans les CBNPC

Les altérations au niveau des principaux drivers oncogéniques sont clonales et généralement mutuellement exclusives (*KRAS*, *EGFR*, *MET*, *BRAF*, *ALK*, *HER2*, ...). Les autres mutations ont tendance à être sous-clonales et interviennent plus tardivement au cours de l'évolution tumorale.^{102,103} Du point de vue de l'évolution clonale, la co-occurrence d'altérations implique une coopération fonctionnelle favorisant la progression tumorale, tandis que l'exclusivité traduit plutôt une redondance ou un antagonisme.¹⁰⁴

Le développement de traitements ciblés sur les voies oncogéniques activées, en particulier l'EGFR, a permis l'accès à un nouvel arsenal thérapeutique efficace dans la prise en charge du CBNPC (Tableau 6), globalement mieux toléré et offrant aux patients une meilleure qualité de vie.

Altérations	Molécule autorisée (AMM, AAP, AAC, CPC)
<i>KRAS G12C</i>	Sotorasib (2L)
<i>EGFR : mutations activatrices (sauf insertions dans l'exon 20)</i>	Gefitinib Erlotinib Afatinib Osimertinib
<i>EGFR : insertions dans l'exon 20</i>	Amivantamab (2L) Mobocertinib (2L) Poziotinib (2L)
<i>MET : mutations entraînant un saut d'exon 14</i>	Capmatinib (2L) Tepotinib (2L) Crizotinib
<i>BRAF V600E et MEK</i>	Dabrafenib + trametinib (2L)
<i>Réarrangement NTRK</i>	Entrectinib Larotrectinib Selitrectinib
<i>Réarrangement ALK</i>	Alectinib Crizotinib Brigatinib Ceritinib Lorlatinib (2L) Entrectinib
<i>Réarrangement ROS1</i>	Crizotinib Entrectinib Lorlatinib (2L)
<i>Réarrangement RET</i>	Pralsetinib (2L) Selpercatinib (2L)

Tableau 6. Thérapies ciblées disponibles en France dans les CBNPC

AMM : Autorisation de mise sur le marché délivrée par l'Agence Nationale de Sécurité du Médicament (ANSM)
 AAP : Autorisation d'accès précoce délivrée par la Haute Autorité de Santé (HAS), sur demande d'un laboratoire, prescrit à un groupe de patients
 AAC : Autorisation d'Accès Compassionnel délivrée par l'ANSM sur demande d'un médecin ou d'une institution sanitaire pour un patient nommé désigné
 CPC : Cadre de prescription compassionnelle délivrée par l'ANSM
 2L : à partir de la 2^{ème} ligne de traitement

2.3. CBNPC, EGFR et thérapies ciblées

a) Structure et fonction du récepteur EGFR

Egalement appelé HER1, l'EGFR est un récepteur transmembranaire à activité tyrosine kinase codé par le gène *EGFR* comprenant 28 exons et situé sur le chromosome 7 (7p12). La glycoprotéine EGFR (poids moléculaire de 170kDa) appartient à la famille des facteurs de croissance épidermiques ErbB/HER constituée de 4 récepteurs transmembranaires : EGFR/ErbB1, ErbB2/HER2/Neu, ErbB3/HER3 et ErbB4/HER4. La protéine EGFR est exprimée dans la plupart des types cellulaires à un taux de 20 000 à 200 000 récepteurs par cellule. Au niveau de l'appareil respiratoire, son expression est forte dans les cellules épithéliales bronchiques et modérée dans les cellules épithéliales

alvéolaires. Physiologiquement, la protéine EGFR est impliquée dans les mécanismes de signalisation intracellulaire contrôlant la croissance, la survie, l'adhésion, la migration ainsi que la différenciation cellulaire. L'EGFR correspond au prototype des récepteurs membranaires à activité tyrosine kinase et présente trois domaines fonctionnels¹⁰⁵⁻¹⁰⁷ (Figure 10) :

- un domaine extracellulaire NH₂-terminal situé dans le milieu extérieur qui contient le site de liaison capable de fixer ses ligands spécifiques, le TGF- α et l'EGF sont les principaux ligands spécifiques de l'EGFR qui jouent un rôle clé dans la modulation de la biologie cellulaire ;
- un domaine transmembranaire hélicoïdal amphiphatique qui permet l'ancrage du récepteur dans la membrane plasmique ;
- un domaine intracellulaire COOH-terminal, présentant entre autres un domaine à activité tyrosine kinase. Ce domaine comporte un petit lobe qui contient un site de liaison de l'ATP et un grand lobe qui contient le site catalytique de l'enzyme ainsi qu'une extrémité cible des phosphorylations.

Le récepteur est inactif sous sa forme de monomère. La fixation du ligand au domaine extracellulaire entraîne un changement de conformation qui favorise le rapprochement de deux monomères. Le récepteur activé est formé d'un homo- ou d'un hétéro-dimère, entre l'EGFR et un autre récepteur de la famille Erb. Le rapprochement des deux domaines tyrosine-kinases aboutit à l'activation de la fonction tyrosine kinase et permet l'auto- ou la trans-phosphorylation des différents résidus tyrosine situés à l'extrémité COOH-terminale intracellulaire du récepteur. Le récepteur activé va pouvoir s'associer à différentes protéines adaptatrices intracellulaires permettant la transduction du signal de la surface cellulaire à son noyau. La transduction du signal fait intervenir trois types de molécules : les protéines à domaines SH₂/SH₃, les protéines G monomériques et les protéines kinases. (Figure 10)¹⁰⁸

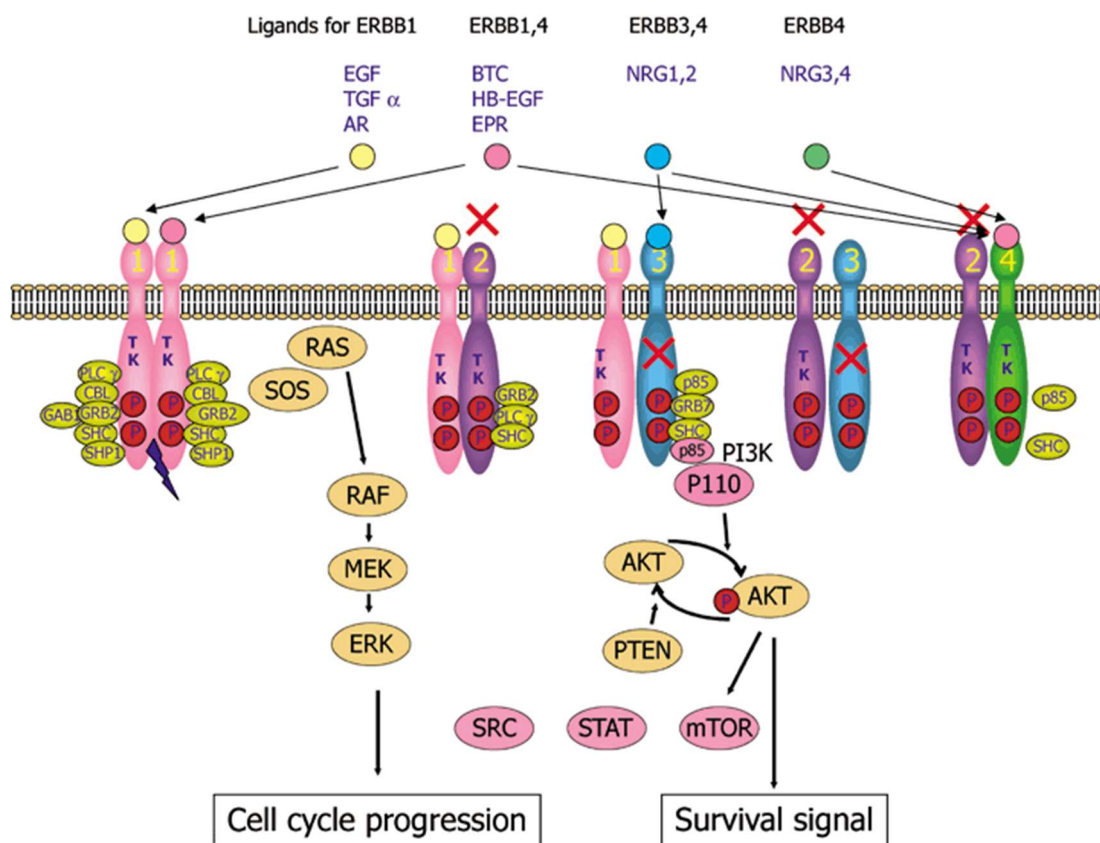


Figure 10 : La famille Epidermal Growth Factor Receptor et ses voies de signalisation. D'après Mitsudomi, T. & Yatabe, 2007¹⁰⁹

b) Mutations oncogéniques d'EGFR et thérapies ciblées dans les CBNPC

Dans les CBNPC, les mutations activatrices du gène EGFR, essentiellement situées dans les exons 18 à 21 du gène codant pour le domaine tyrosine-kinase (Figure 11), provoquent une addiction oncogénique, c'est-à-dire une dépendance des cellules cancéreuses à l'activité d'un oncogène qui favorise la survie, la croissance et la progression des cellules tumorales.¹¹⁰ Le récepteur sous sa forme oncogénique présente une augmentation de l'affinité à l'ATP présent dans le cytoplasme, il est ainsi activé de façon constitutive en l'absence de ligand. Les mutations activatrices les plus fréquentes sont la mutation p.Leu858Arg dans l'exon 21 (40%) et les délétions sans perte du cadre de lecture dans l'exon 19 (Del19) (45%) impactant principalement la séquence LeuArgGluAla (Figure 11). Certaines mutations activatrices du gène *EGFR* sont associées à une sensibilité accrue et d'autres à une résistance primaire ou secondaire aux inhibiteurs de l'activité tyrosine kinase de l'EGFR (ITK-EGFR).

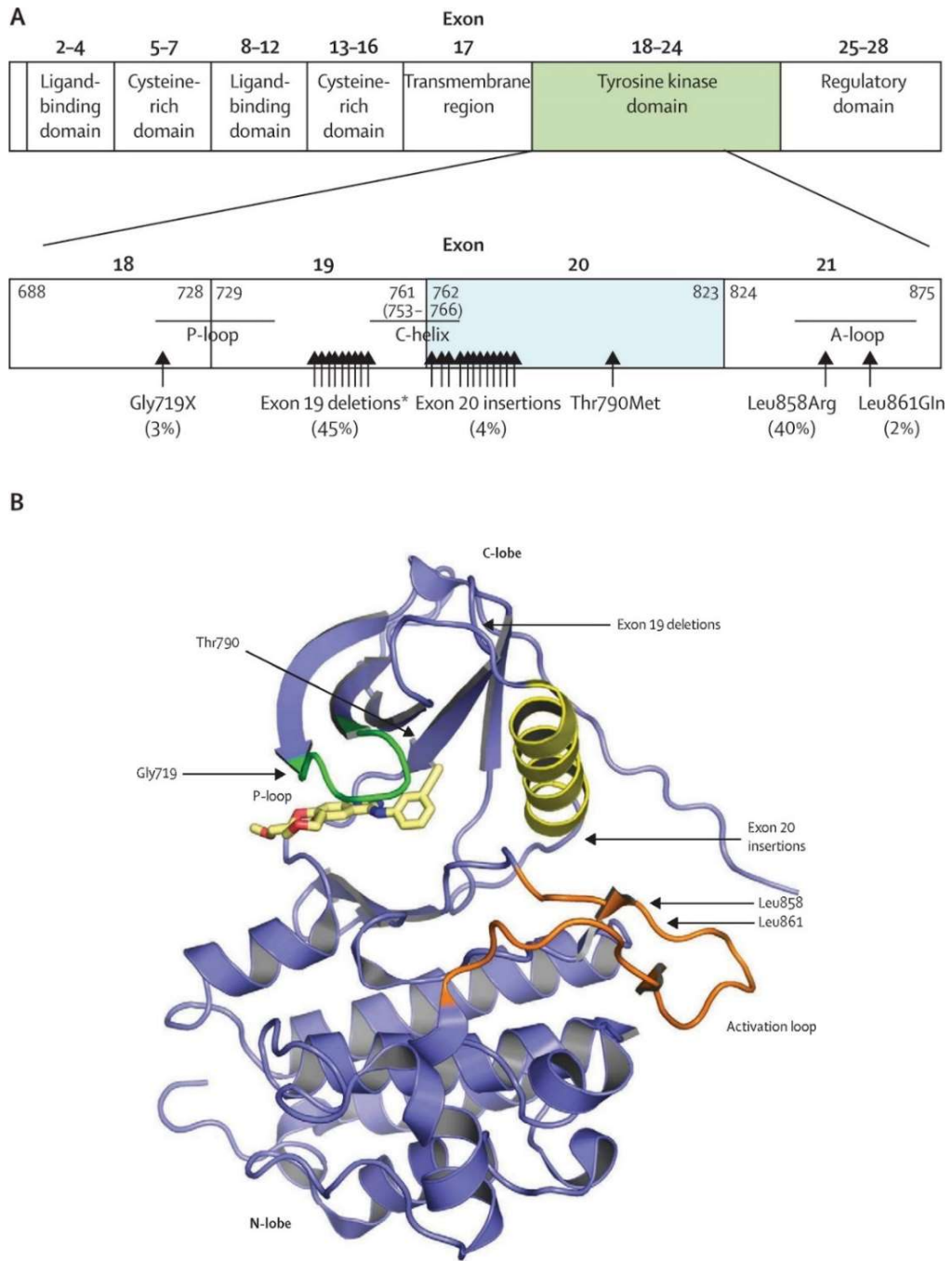


Figure 11. Mutations dans le domaine tyrosine-kinase du gène *EGFR*, d'après Yasuda et al., 2012¹¹¹

A. Structure et exons du gène *EGFR* et fréquence des mutations tumorales *EGFR* dans les CBNPC
 B. Représentation tridimensionnelle du domaine tyrosine-kinase de l'*EGFR* et position des acides-aminés concernés par les mutations tumorales *EGFR* les plus fréquentes

c) ITK-EGFR de première, deuxième et troisième génération

Les ITK-EGFR agissent par compétition avec l'ATP au niveau de la poche de fixation à l'ATP (Figure 12), bloquant les signaux de survie et de prolifération (« choc

oncogénique ») et induisant un arrêt du cycle cellulaire pouvant conduire à l'apoptose de la cellule tumorale.^{112,113} Quatre molécules ITK-EGFR possèdent actuellement des autorisations en France : gefitinib et erlotinib (1^{ère} génération), afatinib (2^{ème} génération) et osimertinib (3^{ème} génération).

- ITK-EGFR de première génération

Les premiers essais cliniques sur le gefitinib et l'erlotinib, inhibiteurs sélectifs et réversibles, ont permis d'identifier les mutations activatrices p.Leu858Arg et Del19, puis de les caractériser comme facteurs de sensibilité aux ITK-EGFR dans les CBNPC de type adénocarcinomes.¹¹⁴⁻¹¹⁶ Les ITK-EGFR de première génération ont d'abord obtenu leurs autorisations en deuxième ou en troisième ligne après un traitement à base de sels de platine suite à une étude réalisée chez des patients métastatiques ayant un profil de patients répondeurs (non-fumeurs, de sexe féminin, d'origine asiatique et atteints d'adénocarcinome).¹¹⁷ Les études ont ensuite porté sur des populations de patients mutés ou sur des populations enrichies en patients mutés : en plus d'une toxicité moindre, un bénéfice sur la survie sans progression (SSP) sous ITK-EGFR a été observé comparativement aux traitements conventionnels (SSP comprises entre 9,2 et 13,1 mois, versus 4,6 à 6,3 mois pour les différents doublets de chimiothérapie testés) (Tableau 7). La seule étude européenne comparant la bichimiothérapie à un ITK-EGFR chez des patients mutés *EGFR* (étude EURTAC) retrouve des SSP conformes aux précédents résultats (9,4 versus 5,2 mois; HR = 0,42) et une médiane de survie globale (SG) qui n'est pas atteinte après un suivi médian de 19 mois.¹¹⁸ Les ITK-EGFR sont alors devenus un standard dans la prise en charge des patients porteurs d'une mutation activatrice de l'*EGFR*, avec des AMM obtenues en 2009 pour le gefitinib et en 2011 pour l'erlotinib. Ces essais ont ouvert la voie aux traitements pilotés par une caractérisation génétique ou encore à la notion de théranostic.

Etude/population	Effectif	SSP en mois ITK-EGFR (IC95) vs chimio	HR (IC95)	Tx réponse (%)	Tx contrôle (%)
IPASS/asiatique (Géfitinib vs carboplatine-paclitaxel)	261	ND (pour la population mutée)	0,48 (0,36-0,64)	71,2	ND
WJTOG/asiatique (Géfitinib vs cisplatine-docétaxel)	177	9,2 (8-13,9) vs 6,3	0,49 (0,34-0,71)	ND	93
NEJ002/asiatique (Géfitinib vs carboplatine-paclitaxel)	228	10,8 (ND) vs 5,4	0,3 (0,22-0,41)	73,7	89,5
OPTIMAL/asiatique (Erlotinib vs carboplatine-gemcitabine)	165	13,1 (10,6-16,5) vs 4,6	0,16 (0,1-0,26)	82	96
EURTAC/caucasienne (Erlotinib vs doublet avec cisplatine)	174	9,7 (8,4-12,3) vs 5,2	0,37 (0,25-0,54)	58	79

Tableau 7. Essais de phase III comparant les ITK-EGFR de première génération à un doublet de chimiothérapie chez des patients porteurs d'une mutation de l'EGFR¹¹⁸⁻¹²²

SSP : survie sans progression ; Tx contrôle : taux de contrôle de la maladie (réponses complètes + réponses partielles + maladies stabilisées); Tx réponse : taux de réponse (réponses complètes + réponses partielles).

- ITK-EGFR de deuxième génération

Suite aux études LUX-LUNG¹²³⁻¹²⁵ réalisées à l'échelle internationale, l'afatinib, inhibiteur irréversible des récepteurs de la famille ErBb, dit ITK-EGFR de « deuxième génération », obtient en 2014 l'AMM dans les CBNPC, localement avancés ou métastatiques qui présentent une (des) mutation(s) activatrice(s) de l'EGFR. Compte tenu des nombreux effets indésirables observés, l'afatinib se positionne progressivement dans les contextes de mutations activatrices rares de l'EGFR comme p.Gly719X ou p.Ser768Ile.^{126,127} Toutefois, des données françaises récentes incitent à envisager une chimiothérapie plutôt qu'un ITK en cas de mutations rares (hors p.Thr790Met et insertions de l'exon 20, bénéficiant de thérapies spécifiques) puisque la SG s'établit à 27,7 mois vs 16,9 mois sous ITK-EGFR.¹²⁸ Actuellement, en cas de mutation *EGFR* rare ou d'association de mutations *EGFR*, il est recommandé de discuter la stratégie thérapeutique en RCP.

- ITK-EGFR de troisième génération

Un ITK-EGFR de troisième génération inhibiteur irréversible de l'EGFR, l'osimertinib, a ensuite été développé. Cette molécule de la classe des mono-anilino-pyrimidines se fixe dans le site de liaison à l'ATP en se liant de manière covalente sur la Cys797 de l'EGFR. Il a, dans un premier temps, ciblé avec succès la mutation de résistance p.Thr790M de l'*EGFR* et les autres mutations de sensibilité en deuxième ligne de traitement après erlotinib et/ou gefitinib (étude AURA2).¹²⁹ Désormais, en cas de mutation activatrice *EGFR*, il est recommandé de proposer un traitement de première ligne par osimertinib. L'essai FLAURA, comparant l'osimertinib à un traitement par ITK-EGFR de première génération (erlotinib ou gefitinib) en 1ère ligne chez les patients présentant une Del19 ou une mutation p.Leu858Arg dans l'exon 21 d'*EGFR* a montré une amélioration significative de la SSP (18,9 mois versus 10,2 mois ; HR=0,46 [0,37-0,57] ; p< 0.001)¹³⁰ et de la SG (38,6 mois versus 31,8 mois ; HR=0,799 (0,641-0,957) ; p=0,0462) en faveur du bras osimertinib.¹³¹ D'autant plus qu'environ 30% des patients dans chaque bras n'ont pas bénéficié d'un traitement de seconde ligne. Le profil de tolérance (tous grades confondus et grades 3/4) est en faveur de l'osimertinib. Par ailleurs, l'osimertinib, ITK de troisième génération, traverse mieux la barrière hémato-encéphalique et montre une efficacité chez les patients présentant des méningites carcinomateuses par comparaison au bras contrôle.¹³²

- Perspectives d'association des ITK-EGFR

Des essais sont actuellement en cours pour évaluer l'association des ITK-EGFR à des anti-VEGF¹³³⁻¹³⁵ ou à des chimiothérapies par doublet sels de platine-pemetrexed¹³⁶⁻¹³⁸, montrant des bénéfices en terme de SSP mais au prix d'une majoration de la toxicité.

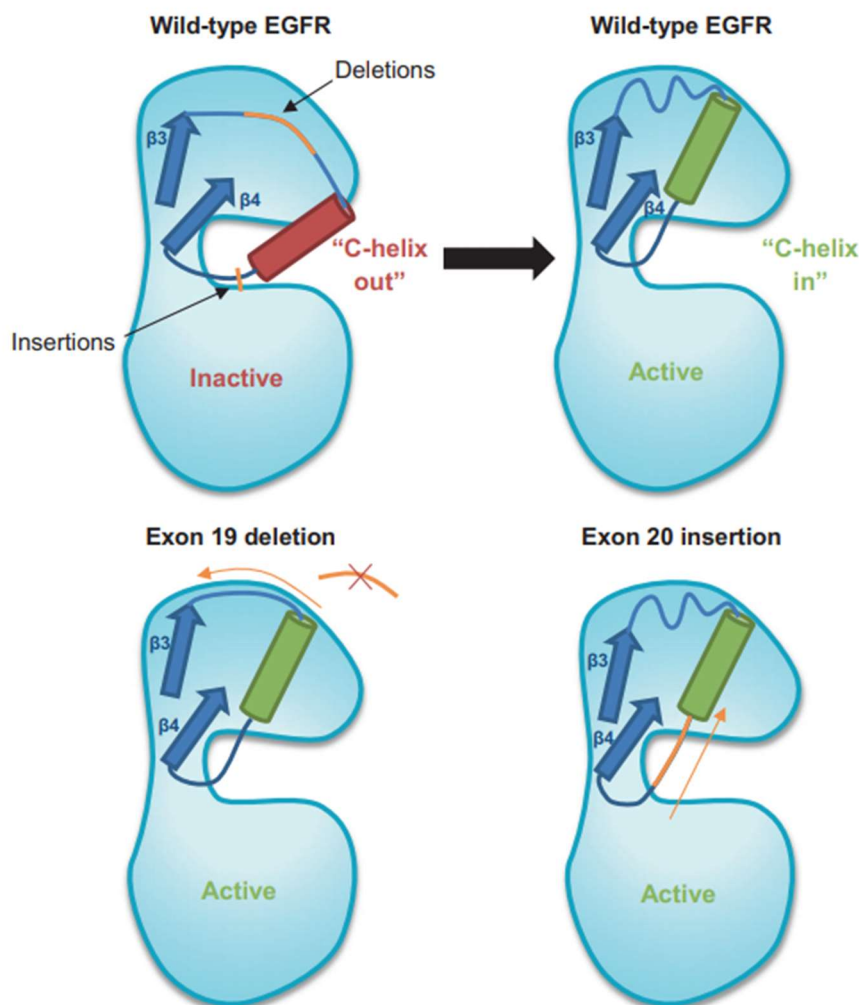


Figure 12 : Impact des délétions et insertions sur l'activation de l'EGFR, d'après Vyse & Huang, 2019¹³⁹

En absence de mutation activatrice, lors de la liaison avec le ligand, l'hélice C régulatrice pivote d'une conformation inactive vers l'extérieur à une conformation active vers l'intérieur pour former des interactions clés avec la boucle p du site actif. Les mutations oncogènes telles que les délétions de l'exon 19 peuvent "tirer" l'hélice C du côté N-terminal, tandis que les insertions de l'exon 20 "poussent" du côté C-terminal pour stabiliser l'état actif de l'EGFR même en l'absence de ligand.

d) Résistance aux ITK-EGFR

On différencie la résistance primaire ou innée caractérisée par l'absence de réponse initiale au traitement de la résistance acquise pour laquelle un bénéfice clinique a été suivi d'une progression de la maladie au cours du temps de traitement.¹⁴⁰

- Résistance primaire

Observée chez 4 à 10% des patients présentant des CBNPC mutés *EGFR*, la résistance primaire aux ITK-EGFR peut avoir deux origines : l'hôte ou la tumeur.¹⁴¹ Les différents facteurs de résistance primaire aux ITK-EGFR sont résumés dans le Tableau 8.

	Facteurs de résistance primaire aux ITK-EGFR et mécanisme(s) envisagé(s) ou démontré(s)	Référence(s)
Facteurs liés à l'hôte	Tabagisme par : - Induction du cytochrome CYP1A1 impliqué dans le métabolisme des ITK - Inhibition de l'ubiquitinylation du récepteur entraînant sa surexpression	Alfieri <i>et al</i> , 2011 ¹⁴² Cortot et Jänne, 2014 ¹⁴¹
	Polymorphismes de la protéine BIM entraînant une dérégulation de la fonction proapoptotique de BIM	Ng <i>et al.</i> , 2012 ¹⁴³
	Variants constitutionnels <i>EGFR</i> activant le récepteur EGFR et prédisposant au CBNPC : - p.Thr790Met* - p.Val843Ile*	Bell <i>et al</i> , 2005 ¹⁴⁴ Matsushima <i>et al</i> , 2014 ¹⁴⁵
	Activation de la voie NFκB : via la surexpression de c-FLIP ou de IKK ou via l'inhibition de IκB.	Bivona <i>et al</i> , 2011 ¹⁴⁶
Facteurs liés à la tumeur	Variants <i>EGFR</i> tumoraux <i>de novo</i> diminuant d'affinité du récepteur pour les ITK-EGFR : - insertions dans l'exon 20 - variant p.Thr790Met* - variant p.Asp761Tyr* - variants p. Ser768Ile* et p.Val769Leu* - variant p.Thr854Ala*	Vyse et Huang, 2019 ¹³⁹ Fujita <i>et al.</i> , 2012 ¹⁴⁷ Toyooka <i>et al.</i> , 2007 ¹⁴⁸ Asahina <i>et al</i> , 2006 ¹⁴⁹ et Niogret <i>et al</i> , 2018 ¹⁵⁰ Zhang <i>et al</i> , 2021 ¹⁵¹
	Activation des récepteurs ErbB2 et ErbB3	Erjala <i>et al.</i> , 2006 ¹⁵²
	Pertes de PTEN	Yamamoto <i>et al</i> , 2010 ¹⁵³

Tableau 8 : Facteurs de résistance primaire aux ITK-EGFR

*Résistance démontrée pour les ITK-EGFR de première et deuxième générations, sensibilité à l'osimertinib en première ligne non étudiée ou en cours d'étude.

- Résistance acquise

Une rechute est inexorablement constatée quel que soit l'ITK-EGFR administré, illustrant les effets de l'instabilité génomique et donc les capacités d'adaptation des cellules tumorales. Les mécanismes de résistance acquise sont hétérogènes et peuvent résulter de deux processus : la sélection clonale (sélection d'un clone préexistant) et la pression sélective (apparition d'une altération secondaire au traitement conférant un avantage prolifératif). Dans les CBNPC mutés *EGFR*, la connaissance des mécanismes de résistance acquise aux ITK-EGFR est primordiale pour adapter au mieux le traitement : on différencie la résistance *on-target* impliquant l'*EGFR*, de la résistance *off-target* mettant en jeu un changement d'addiction oncogénique et des changements phénotypiques¹⁵⁴ (Tableau 9). Comme l'illustrent les Figures 13 et 14, la résistance acquise aux ITK-EGFR n'est pas superposable d'une génération d'ITK à l'autre, aussi bien en terme de type de mécanisme qu'en terme de fréquence de survenue. En effet, suite à l'utilisation de l'osimertinib, dans un premier temps en deuxième ligne puis en première ligne de traitement, de nouvelles

mutations *on-target* ont été identifiées mais aussi des mécanismes *off-target* non décrits jusqu'alors.

Les résistances *on-target* sont de deux types : les mutations secondaires de l'*EGFR* et les surexpressions du récepteur (Tableau 9). La mutation p.Thr790Met de l'exon 20 du gène *EGFR* est responsable de 50% des résistances acquises aux ITK-EGFR de première et deuxième générations, elle apparaît le plus souvent en cis de la mutation driver primaire. Cette mutation entraîne une affinité supérieure du récepteur pour l'ATP et ainsi une inhibition compétitive de la fixation des ITK-EGFR.¹⁵⁵⁻¹⁵⁸ D'autres mutations *EGFR* moins fréquentes causent également une résistance acquise et certaines mutations sont spécifiques à l'osimertinib (en deuxième ligne et en première ligne de traitement), la plus fréquente étant la p.Cys797Ser, site de fixation de la molécule.^{159,160} (Figure 13)

Mutations EGFR secondaires :

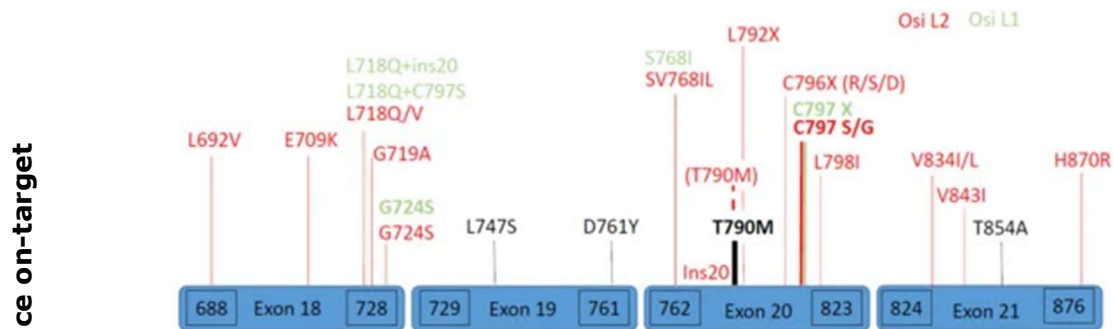


Figure 13. Mutations secondaires EGFR. Adapté de Reita et al, 2021¹⁶⁰

Résistance on-target

Noir : à progression après ITK-EGFR de 1^{ère} et 2^{ème} générations
 Rouge : à progression après ITK-EGFR de 3^{ème} génération en 2^{ème} ligne de traitement
 Vert : à progression après ITK-EGFR de 3^{ème} génération en 1^{ère} ligne de traitement

Surexpression du récepteur : l'amplification de l'allèle muté de l'*EGFR* contribue à la résistance acquise avec un impact mal connu^{161,162}

Amplification / surexpression :

- MET (dans la moitié des cas associées à l'apparition du variant p.Thr790Met)¹⁶³
- HER2¹⁶⁴
- FGFR1¹⁶⁵
- AXL¹⁶⁶

Résistance off-target

Autres altérations génomiques¹⁶⁰ :

- Mutations ponctuelles : *PIK3CA*, *BRAF*, *KRAS*, *NRAS*, *MET*, *CTNNB1*...
- Pertes de PTEN
- Gènes de fusion impliquant : *NTRK*, *RET*, *ALK*, *ROS1*, *FGFR3*, *BRAF*...

Transformations histologiques / changements phénotypiques des cellules tumorales¹⁶¹ :

- Transformation en CBPC
- Transition épithélio-mésenchymateuse (EMT)

Tableau 9 : Principaux mécanismes de résistance acquise aux ITK-EGFR

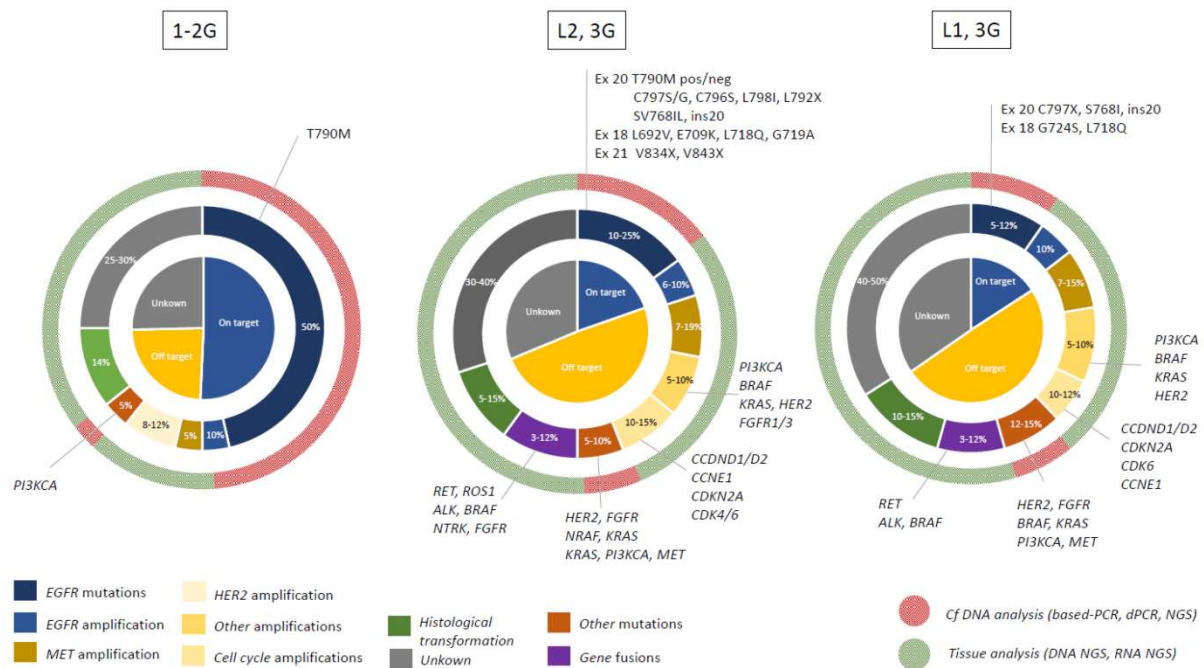


Figure 14 : Mécanismes de résistance aux ITK-EGFR dépendants de la génération d'ITK-EGFR et de la ligne de thérapie ciblée. D'après Reita et al, 2021¹⁶⁰

1-2G : ITK-EGFR de première et seconde générations ; 3G : ITK-EGFR de troisième génération ; L1 : première ligne de traitement ; L2 : seconde ligne de traitement

e) Insertions dans l'exon 20 d'EGFR et thérapies ciblées

Les insertions dans l'exon 20 du gène *EGFR* (Ins20) sont généralement situées après l'hélice C du domaine tyrosine kinase (Figures 11 et 12), et sont caractérisées par des insertions sans modification du cadre de lecture de 3 à 21 bases entre les AA 762 et 774 de la protéine *EGFR*.¹⁶⁷ Ces mutations sont généralement exclusives des autres drivers oncogéniques et représentent 4 à 10% de l'ensemble des mutations *EGFR* dans les CBNPC (Figure 11).^{111,139,167,168} Les Ins20 entraînent une modification conformationnelle et un mécanisme d'activation de la kinase différents des autres mutations activatrices *EGFR* : les Ins20 sont responsables d'une activation constitutive de l'*EGFR* en "poussant" l'hélice C du côté C-terminal du récepteur, stabilisant ainsi l'état actif (Figure 12). Ce mécanisme unique d'activation entraînerait une résistance *de novo* aux ITK-EGFR. Les modèles *in vitro* ont en effet montré que les protéines mutées par Ins20 les plus répandues sont résistantes aux doses cliniquement utilisables des ITK-EGFR réversibles (gefitinib, erlotinib) et irréversibles (neratinib, afatinib, osimertinib).^{167,169,170} L'expérience clinique croissante des patients dont les tumeurs portent des Ins20 correspond aux données précliniques ; très peu de patients ont répondu aux ITK-EGFR.^{127,171,172}

Le développement récent de plusieurs nouvelles thérapies ciblées inhibant sélectivement les Ins20 est prometteur pour les futures options thérapeutiques efficaces des patients atteints de CBNPC :

- le poziotinib et le mobocertinib : ITK inhibant les Ins20 d'*EGFR*¹⁷³⁻¹⁷⁷
- l'amivantamab : anticorps bispécifique de type IgG1 humain dirigé contre les récepteurs EGFR et MET, produit par une lignée cellulaire de mammifère (lignée ovarienne de Hamster Chinois) par les techniques de l'ADN recombinant¹⁷⁸⁻¹⁸⁰.

Ces trois options thérapeutiques sont actuellement autorisées en France (ACC) en deuxième ligne de traitement des CBNPC avancés ou métastatiques mutés *EGFR* Ins20.

Il apparaît donc qu'une caractérisation moléculaire de l'*EGFR* est un préalable essentiel à la prise en charge des patients atteints de CBNPC.

2.4. Détection des mutations *EGFR* pour la prise en charge clinique des CBNPC

a) Tests moléculaires sur échantillons de tissus

La caractérisation des variants génétiques du gène *EGFR* dans un but diagnostique des CBNPC doit privilégier les prélèvements histologiques. Les cyto blocs (inclusion en paraffine du culot de centrifugation) permettent également la réalisation d'études immunohistochimiques et moléculaires. Depuis 2016, l'*European Expert Group* recommande au moins 5 biopsies bronchiques pour le diagnostic et 5 biopsies supplémentaires pour le phénotypage et le génotypage.¹⁸¹

Les tests de détection des variants génétiques activateurs de l'*EGFR* et de BRAF, NRAS et KRAS ont reçu un avis favorable de la HAS pour leur remboursement en 2014 et 2015. Les conditions de réalisation de ces tests sur tissus tumoraux ont fait l'objet de recommandations de la HAS dans un rapport en septembre 2021. Les populations cibles des tests correspondent aux patients concernés par les indications définies dans les AMM des thérapies ciblées associées à leurs tests compagnons ; elles sont estimées à partir des données d'activité de génétique somatique de l'INCa.¹⁸²

La phase pré-analytique est sous la responsabilité d'un médecin anatomo-pathologiste. Elle comprend des étapes de prélèvement, de préparation, de qualification et de sélection de l'échantillon :

- La préparation de l'échantillon s'effectue soit par la cryoconservation soit par la méthode de fixation par le formol suivie d'une inclusion en paraffine (FFPE pour *formalin-fixed and paraffin-embedded*). La fixation des prélèvements histologiques doit se faire immédiatement après le prélèvement. Les fixateurs à base d'acide

picrique et d'alcool – formol – acide acétique sont à proscrire. Il a été démontré que les échantillons de tissus FFPE permettaient à la fois l'analyse morphologique et la garantie d'une qualité acceptable des acides nucléiques à analyser.¹⁸³⁻¹⁸⁵

- La qualification de l'échantillon correspond à un contrôle macroscopique et microscopique de l'échantillon. Elle permet notamment d'identifier la présence d'éventuels facteurs confondants (tissu nécrosé, cellule stromale), de préciser qu'un minimum de 100 cellules tumorales étaient analysables (sinon des réserves doivent être émises) et surtout de déterminer la cellularité tumorale de l'échantillon.¹⁸² L'échantillon destiné à l'analyse génétique doit être le plus riche possible en cellules tumorales, car sa contamination par du tissu non tumoral entraîne une dilution du pourcentage d'allèles mutés diminuant de ce fait la sensibilité du test. La cellularité tumorale de l'échantillon doit être supérieure à la limite de détection du test ou/ de la technique d'analyse utilisée, le seuil de 20 à 30% de cellules tumorales est généralement décrit comme seuil minimal pour interpréter les analyses de biologie moléculaire. Cependant chaque laboratoire de génétique moléculaire est libre de pratiquer les analyses sur des échantillons de cellularité tumorale très faible (< 20%) en fonction du contexte clinique et du matériel disponible. Dans le cas des CBNPC, les prélèvements tumoraux sont souvent pauvres en terme de quantité et de qualité de matériel.¹⁸⁶

La phase analytique du test comprend deux étapes :

- Une première étape d'extraction des acides nucléiques à partir du tissu fixé (FFPE) ou cryoconservé : lyse cellulaire, élimination des protéines et purification / concentration des acides nucléiques.
- Puis la recherche d'altérations par technique moléculaire qui repose essentiellement sur la PCR ciblée sous différents formats et sur le séquençage, il n'existe pas de technique de référence.

La phase post-analytique comprend l'interprétation, la validation et la communication des résultats via l'envoi du compte-rendu de génétique moléculaire aux médecins prescripteurs. Les résultats de l'analyse sont présentés de manière qualitative (présence ou absence de variants) et les variants identifiés sont notés selon la nomenclature internationale HGVS. L'interprétation et la conclusion des résultats doivent établir l'association variant/sensibilité à la thérapie ciblée sans proposition de traitement. Le résultat du test génétique va permettre au(x) médecin(s) de statuer sur la stratégie thérapeutique à adopter. La communication des résultats au patient se fait ensuite par le médecin clinicien au cours d'une consultation durant laquelle il communique les résultats du test génétique et le (ou les) choix thérapeutique(s) qui en découle(nt). Le délai de rendu

des résultats doit donc être le plus court possible pour pouvoir éclairer la prise de décision thérapeutique puis initier sans retard le traitement sans perte de chance.

Conformément à la loi en vigueur, tous les laboratoires de biologie moléculaire doivent être accrédités par le Comité français d'accréditation (COFRAC) selon les critères de la norme ISO 15189 (norme de qualité internationale de référence pour les laboratoires de biologie médicale). Chaque technique d'analyse utilisée par le laboratoire doit faire l'objet d'un dossier de validation de la méthode reposant essentiellement sur l'évaluation des performances de la technique (sensibilité analytique, répétabilité, reproductibilité, robustesse) et sur la maîtrise des risques à chaque étape du processus analytique du patient, du prélèvement de l'échantillon jusqu'à la communication des résultats d'analyse.

Dans le contexte particulier des CBNPC et dans un délai maximum de 3 semaines calendaires suivant la date de prélèvement, une recherche d'altérations moléculaires doit systématiquement être demandée lorsque suffisamment de tissu a pu être obtenu pour le diagnostic : dans tous les CBNPC non épidermoïdes et dans les carcinomes épidermoïdes des non-fumeurs de stades avancés. Les cibles à rechercher sont les suivantes^{182,187} :

- Dès le diagnostic initial : l'expression de PDL1 (immunohistochimie), les variants génétiques de l'*EGFR* et de *KRAS* et les fusions des gènes *ALK* et *ROS1*.
- En option au diagnostic initial et systématiquement avant la seconde ligne : les variants de *BRAF*, les altérations d'épissage de l'exon 14 de *MET* et les fusions des gènes *RET* et *NTRK*.
- Systématiquement avant la seconde ligne : les autres altérations de *MET*, les variants de *HER2* et les autres fusions (notamment *NRG1*).

En cas d'insuffisance ou d'absence de tissu tumoral contributif pour l'analyse moléculaire, ce qui représente en moyenne 10% des prélèvements tissulaires adressés aux laboratoires de biologie moléculaire³¹, une alternative s'est progressivement développée : l'utilisation de biopsies liquides pour étudier l'ADN tumoral circulant (ADNtc).

b) Tests moléculaires dans l'ADNtc

L'identification de cibles moléculaires par l'analyse de l'ADNtc est une évolution importante dans les pratiques diagnostiques en oncologie moléculaire somatique appliquées à la médecine personnalisée.^{30,188,189}

Les cellules saines ou tumorales libèrent leur contenu et donc leurs acides nucléiques dans la circulation sanguine, essentiellement lors de l'apoptose ou de la nécrose.³⁰ On retrouve par conséquent de l'ADN (mais également des ARN de différentes natures) dans la circulation sanguine. Chez les patients atteints de cancer, cet ADN

circulant peut être d'origine tumorale et ainsi porter les anomalies génétiques des cellules tumorales dont il provient. L'ADNtc est en proportions très variables et peut représenter une très faible fraction de l'ADN libre circulant, souvent moins de 0,01%.¹⁹⁰ Grâce au développement de techniques de biologie moléculaire sensibles et spécifiques, il est souvent possible de détecter ces anomalies génétiques somatiques dans l'ADNtc. Du fait des difficultés et de la paucité tumorale des prélèvements, le cancer broncho-pulmonaire est le type tumoral pour lequel l'intérêt clinique de l'analyse de l'ADNtc est le plus documenté et le plus pertinent. Cette analyse est recommandée au moment du diagnostic d'un CBNPC métastatique, dans les cas où l'échantillon tissulaire tumoral utilisé est de quantité ou de qualité insuffisante pour conduire les analyses moléculaires somatiques : tissu tumoral épuisé après analyse histologique, problème de fixation, traitement délétère du prélèvement tissulaire (décalcification), mauvaise qualité de l'ADN tissulaire extrait et/ou résultat non contributif pour diverses raisons.^{188,191,192} De même, lors de la progression tumorale des patients présentant un CBNPC avec mutation activatrice *EGFR* et progressant sous traitement par ITK-*EGFR*, les mutations de résistance *on-target* et *off-target* peuvent être recherchées à partir d'ADNtc.¹⁹³⁻¹⁹⁷ Si le mécanisme de résistance n'est pas identifié dans l'ADNtc, une nouvelle biopsie (de la tumeur primitive ou d'un site ganglionnaire ou secondaire) doit être systématiquement envisagée pour une analyse moléculaire sur tissu. Si la biopsie n'est pas réalisable, un second prélèvement sanguin pour analyse sur ADNtc est une option dans les situations de progression lente si cela ne génère pas de retard de prise en charge. Il est important de préciser que l'absence de mutation détectée dans l'ADNtc n'exclut pas la présence de mutation(s) dans le tissu tumoral.³¹ De nombreux travaux suggèrent que les cinétiques de disparition et d'apparition des mutations d'*EGFR* seraient associées à la réponse des patients aux traitements par ITK^{194,195,198}, mais l'utilité clinique de cette observation en pratique courante reste à démontrer. Aussi, il n'est à ce jour pas recommandé de suivre de manière régulière la disparition de la mutation princeps d'*EGFR* et/ou la survenue d'une mutation de résistance au cours du traitement, celle-ci doit être effectuée en cas de progression documentée de la maladie.

De l'ADNtc peut également être retrouvé dans d'autres fluides biologiques et être analysé dans les mêmes indications que celles du plasma : liquide broncho-alvéolaire, liquide céphalorachidien, liquide pleural. L'élargissement des types de liquides biologiques utilisés représente une chance supplémentaire de mettre en évidence des cibles de manière non-invasive, notamment en cas de métastases cérébrales.¹⁹⁹⁻²⁰²

La phase pré-analytique est une étape clé de l'analyse de l'ADNtc plasmatique car la quantité d'ADN normal libéré par la lyse des cellules sanguines vient progressivement diluer l'ADNtc. Deux possibilités : une centrifugation rapide des prélèvements sur tubes

EDTA (dans les 4h suivant le prélèvement) ou l'utilisation de tubes de conservation spécifiques. La technique analytique utilisée pour étudier l'ADNtc doit être suffisamment sensible, avec une limite de détection inférieure ou égale à 2% de fréquence allélique.³¹

c) Les techniques de détection des mutations *EGFR*

De nombreuses méthodes d'analyse ont été appliquées à la détection d'altérations génétiques dans les tumeurs. Le choix de la technique employée repose sur un compromis entre divers critères : la sensibilité, l'adéquation avec le type d'échantillon analysé, le temps technique, le coût et la disponibilité.

L'ADN extrait de tissus FFPE et des fluides biologiques est fragmenté²⁰³, ainsi diverses techniques analytiques peuvent être appliquées à ces ADN. Deux types de techniques peuvent être utilisées pour détecter les mutations *EGFR* : les techniques dites ciblées permettant une détection sélective de mutations connues *a priori* et les techniques de NGS permettant le séquençage de régions choisies pour un set de gènes d'intérêt. Les techniques les plus couramment utilisées pour détecter les mutations somatiques *EGFR* sont décrites dans le Tableau 10.

Concernant les techniques ciblées, l'historique séquençage de Sanger a progressivement laissé la place à de nouvelles technologies ; pour en citer quelques-unes (Tableau 10) :

- Le pyroséquençage qui est une technique de séquençage de l'ADN permettant d'effectuer un séquençage rapide ciblé basé sur la détection chimioluminescente de pyrophosphate libéré lors de la liaison phosphodiester lors de l'addition d'un désoxyribonucléotide complémentaire. Sa sensibilité analytique est d'environ 5% d'allèles mutés.²⁰⁴
- La technique *high resolution melting* (HRM) ou analyse des courbes de fusion à haute résolution. La région de l'ADN potentiellement porteuse de la mutation est d'abord amplifiée par PCR puis les amplicons sont chauffés très progressivement et précisément jusqu'au point où les deux brins d'ADN se séparent. Le principe de l'HRM est d'observer la température de fusion qui témoigne de la complémentarité des séquences.²⁰⁵
- Des solutions commerciales de PCR quantitative temps réel spécifique d'allèles : développées pour un set de mutations *EGFR*. Elles permettent une mise en œuvre technique facilitée. Leur sensibilité est variable d'une cible à l'autre (fréquence allélique de 1 à 10%).
- La PCR digitale (PCRd) consiste à répartir et partitionner le volume réactionnel total d'une PCR contenant les cibles et les réactants dans un grand nombre de compartiments (de quelques centaines à plusieurs millions) de très faible volume

(de l'ordre du nanolitre, voire moins). Ces compartiments peuvent se présenter comme des logettes sur un support solide (micro-chambres) ou d'une émulsion huile/eau sous la forme de « nano-gouttelettes ». Ainsi, en principe, la partition des cibles permet d'isoler les copies sauvages des copies mutées. Le signal fluorescent des sondes d'hydrolyses (Taqman) spécifiques des copies mutées peut plus facilement se démarquer de celui des copies sauvages, permettant d'améliorer significativement la sensibilité analytique (de l'ordre de 0,01%). Pour les PCRd, chaque laboratoire choisit les sondes d'hydrolyse spécifiques des mutations qu'il souhaite détecter.

- La méthode BEAMing (pour *Beads, Emulsion, Amplification, Magnetics*) combine PCRd et cytométrie en flux. Après une étape de pré-amplification par PCR des séquences d'intérêt, les séquences amplifiées sont liées à des billes magnétiques puis compartimentées en microgouttelettes dans une émulsion eau-huile. Dans chaque microgouttelette se produit une réaction PCR. Les billes sont ensuite purifiées à l'aide d'un aimant puis incubées avec des marqueurs fluorescents. Les séquences d'ADN ainsi marquées sont ensuite analysées par cytométrie en flux.²⁰⁶

Principe	Technique / description	Nombre de mutations <i>EGFR</i> détectées	Recherche sur tissu	Recherche sur ADNtc	Limite de détection (FA en %)
Pyroséquençage	Technique de séquençage de l'ADN basée sur la détection de l'activité de l'ADN polymérase avec une autre enzyme chimioluminescente ²⁰⁴	Choix du laboratoire	oui	non	5%
HRM	Technique d'analyse des courbes de fusion ²⁰⁵	Choix du laboratoire	oui	non	5%
PCR quantitative en temps réel spécifique d'allèles	kit cobas [®] EGFR Mutation Test v2 (Roche) ²⁰⁷	18	oui	oui	Selon les mutations de 1,3% à 13,4%
	kit Therascreen EGFR RGQ PCR Kit v2 (Qiagen) utilisant les technologies ARMS (Amplification Refractory Mutation System) et Scorpions ²⁰⁸	29	oui	Non validée	Selon les mutations de 0,66% à 16,87%
	Cartouches Idylla <i>EGFR</i> Mutation Test (Biocartis) utilisant les technologies PlexZyme et PlexPrime ^{209,210}	51	oui	non	Selon les mutations de 1 à 19,1%
PCR digitale	PCR Digitale en microgouttelettes à haut débit : - BioRad - Raindance - Stilla	Choix du laboratoire	oui	oui	Jusqu'à 0,01% selon la technique utilisée
	PCR Digitale en compartiments solides ThermoFisher QuantStudio	Choix du laboratoire	oui	oui	
BEAMing PCR	Technique combinant PCRd et cytométrie en flux ²⁰⁶	Choix du laboratoire	oui	oui	0,01%
NGS	Panels de séquençage haut débit	Choix du laboratoire	oui	oui	2%

Tableau 10 : Techniques actuellement utilisées pour la détection des mutations *EGFR*

Le nombre de thérapies ciblant des altérations moléculaires spécifiques n'a cessé d'augmenter en cancérologie ces vingt dernières années et de nombreuses molécules sont en phase de développement clinique. Le choix du traitement des patients est de plus en plus orienté par le résultat de la détermination d'un panel de biomarqueurs spécifique à chaque localisation tumorale. En conséquence, les laboratoires de génétique tumorale ont pour la plupart amorcé un virage technologique vers le NGS au détriment des méthodes ciblées. Cette implémentation du NGS en routine clinique génère un grand nombre de VSI pour lesquels des tests complémentaires de caractérisation fonctionnelle sont nécessaires et participent à la décision thérapeutique. Par exemple, la prescription de certaines molécules de la récente classe thérapeutique des inhibiteurs de PARP (iPARP) est liée à la présence de variants pathogènes des gènes *BRCA1* et *BRCA2* pour lesquels des dizaines de milliers de VSI sont aujourd'hui rapportés.

3. INTERPRETATION DES VSI POUR LA PRESCRIPTION DES INHIBITEURS DE PARP

3.1. Mécanisme d'action des inhibiteurs de PARP et principe de létalité synthétique

Les poly(ADP-ribose) polymérases (PARP) sont des enzymes nucléaires impliquées dans des modifications post-traductionnelles de protéines et dans la synthèse de poly(ADP-ribose). Parmi la famille des PARP composée de 17 membres, les PARP 1 et 2, les plus connues, jouent un rôle important dans la réponse cellulaire aux altérations de l'ADN. Elles sont en particulier impliquées dans la réparation des coupures simple-brin par le mécanisme de réparation par excision de base (BER pour *Base Excision Repair*) en permettant le recrutement sur le site de la coupure du complexe protéique de réparation.²¹¹ Si le mécanisme BER est bloqué par l'inhibition des PARP, la coupure simple-brin va devenir double après passage de la fourche de réplication. La coupure double-brin pouvant être réparée par la voie de réparation de l'ADN par recombinaison homologue (RH), il s'ensuit que l'inhibition pharmacologique des PARP dans une cellule normale n'a pas de conséquence significative. Le terme de létalité synthétique a été utilisé pour décrire l'effet de l'inhibition des PARP dans des cellules déficientes pour la recombinaison homologue du fait, en particulier, d'une mutation de *BRCA1* ou *BRCA2* qui participent à cette voie. L'inefficacité conjointe des deux systèmes de réparation entraîne des cassures chromosomiques, un arrêt en phase G2 du cycle et finalement l'apoptose (Figure 15).^{212,213} Les autres protéines PARP ont d'autres fonctions biologiques diverses, comme par exemple la formation du fuseau mitotique, la spermatogénèse, l'immunité innée ou le trafic intracellulaire.²¹³

Les iPARP sont des inhibiteurs compétitifs analogues du NAD⁺ essentiel à l'activité de parylation. Ils forment une nouvelle classe thérapeutique en cancérologie, dont le chef de file est l'olaparib. Parmi les molécules actuellement disponibles en France, certaines ont une prescription liée à un biomarqueur de déficience de la voie de recombinaison homologue (HRD pour *homologous recombination deficiency*) (Tableau 11).

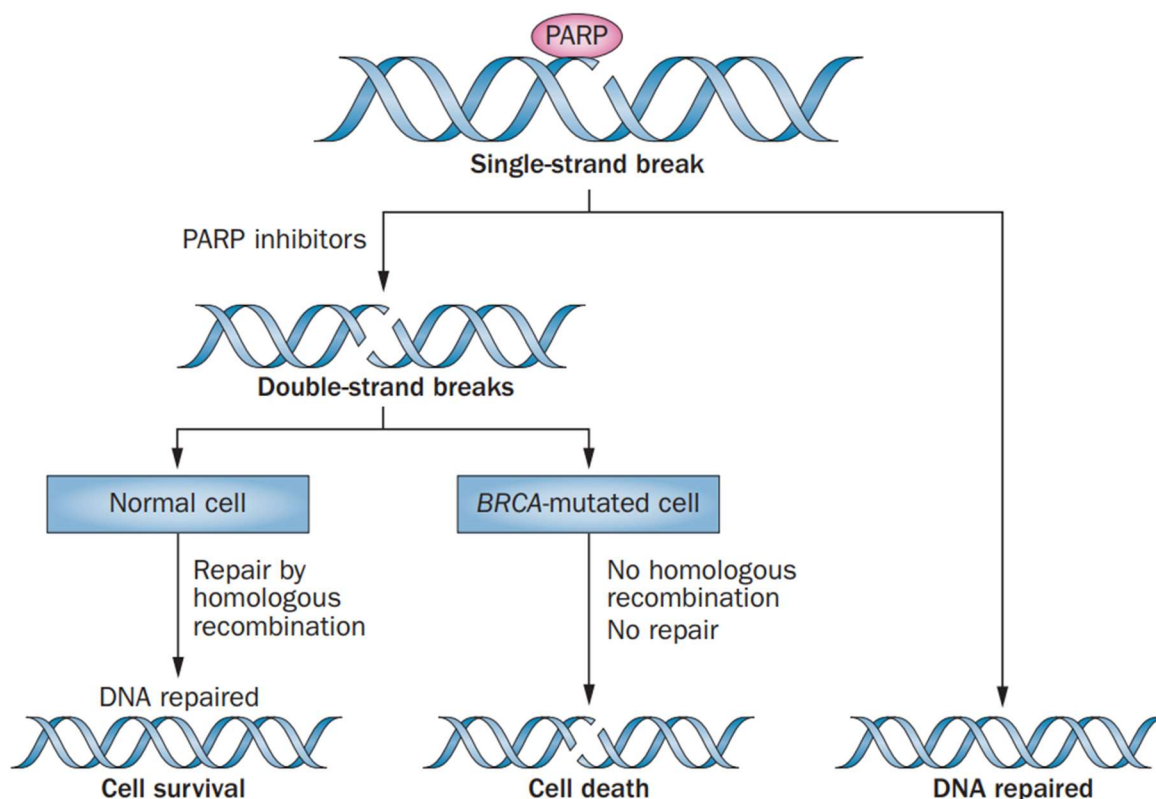


Figure 15. Rôle des iPARP dans la létalité synthétique d'après Sonnenblick, Amir et al.²¹³

Molécule iPARP	Prescription liée à un biomarqueur	Biomarqueur concerné	Localisations tumorales concernées par les autorisations
Olaparib	Oui	Mutation <i>BRCA1/2</i> Statut HRD	Ovaire/sein/pancréas/prostate Ovaire
Niraparib	Non	-	Ovaire
Rucaparib	Non en 2 ^{ème} ligne Oui à partir de la 3 ^{ème} ligne	- Mutation <i>BRCA1/2</i>	Ovaire
Talazoparib	Oui	Récepteur HER2 négatif <u>et</u> mutation <i>BRCA1/2</i>	Sein

Tableau 11. iPARP autorisés en France en oncologie et biomarqueur de HRD associé¹⁸⁻²⁰

3.2. La voie de recombinaison homologue

Parmi les milliers de lésions de l'ADN que subissent quotidiennement les cellules, cancéreuses ou non, les plus redoutables sont les cassures double-brin (DSB pour *Double Strand Break*), car une seule peut conduire à la mort cellulaire si elle n'est pas réparée correctement. Ces DSB de l'ADN peuvent être dues à des mécanismes physiologiques lors de la méiose et de la maturation des cellules lymphoïdes T et B. Cependant, la plupart sont pathologiques, souvent liées à des accidents de la réplication, ou à l'action d'agents exogènes. Deux grandes voies de réparation de l'ADN peuvent réparer ces cassures : la

voie de jonction d'extrémités non homologues (NHEJ pour *Non-Homologous End-Joining*) et la voie de RH.²¹⁴ L'intervention d'une voie par rapport à une autre est déterminée par plusieurs facteurs dont principalement la phase du cycle cellulaire dans laquelle la lésion est détectée. Lorsque la cellule présente une lésion de l'ADN, les points de contrôle du cycle cellulaire ou *checkpoints* induisent le ralentissement ou l'arrêt de la progression du cycle cellulaire et activent en parallèle les mécanismes de réparation appropriés, afin d'empêcher le transfert des erreurs aux cellules filles.^{215,216} Les mécanismes de réparation par RH et NHEJ sont en compétition. Lors d'un défaut des protéines impliquées dans la RH, la NHEJ prend le relais. De même, la RH est augmentée lorsqu'un des composants NHEJ est muté.²¹⁷

La NHEJ est non conservatrice et va fusionner les extrémités des DSB sans tenir compte de la séquence homologue. Cette voie est la seule active tout au long du cycle cellulaire, notamment en G1, mais présente l'inconvénient d'être infidèle et de créer des insertions ou de délétions de matériel génétique, voire des translocations, lors de la réparation.

La voie de RH est la seule pouvant parfaitement réparer la lésion. Elle utilise en effet une copie intacte du génome, le plus souvent la chromatide sœur, comme matrice (ou le chromosome homologue lors de la méiose). Elle n'est active que lors de la réplication en phases S et G2. Le mécanisme de réparation par RH des DSB fait appel à 3 types d'acteurs protéiques (Figure 16) : des « détecteurs » (protéines qui vont identifier les extrémités d'ADN cassées), des « effecteurs » (protéines qui vont réparer l'ADN) et des « médiateurs » (protéines qui vont permettre le lien entre les premiers et les deuxièmes acteurs protéiques).²¹⁸ Tout d'abord, lorsqu'une DSB est formée, le complexe MRN la détecte et recrute la kinase ATM (*AtaxiaTealangiectasia Mutated*), menant à la phosphorylation du variant d'histone H2AX. MDC1 reconnaît cette forme phosphorylée de H2AX, s'y associe et recrute de nouveau le complexe MRN (formé des protéines Mre11, Rad50 et NSb1) et des E3 ubiquitine ligases (RNF8 et RNF168).^{219,220} Le complexe MRN et la protéine CtIP interviennent pour initier la résection des extrémités de la cassure.²²¹ L'activation de ce complexe induit l'activation d'ATM qui va phosphoryler diverses protéines impliquées dans le remodelage de la chromatine, dans l'arrêt du cycle cellulaire et dans la reconnaissance des extrémités de la cassure. Les extrémités simples brins obtenues à la suite de la résection sont stabilisées par la protéine RPA (*Replication Protein A*) qui prévient également la formation de structures secondaires. Les protéines BRCA1, BRCA2, PALB2, 53BP1, RAD51 et ses homologues (RAD51B, RAD51C, RAD51D) sont ensuite recrutées afin d'amener la protéine RAD51 à former un filament nucléoprotéique à partir de l'ADN simple brin. Une fois celui-ci formé et l'homologie avec la chromatide sœur trouvée, l'ADN polymérase allonge l'ADN simple brin par son extrémité 3' en utilisant l'ADN complémentaire comme matrice pour restaurer les nucléotides perdus lors de la cassure.

Les extrémités sont ensuite refermées grâce à une ligase et les brins entrecroisés des deux chromatides sœurs (« jonctions de Holliday ») sont enfin coupés par la protéine MuS81 résultant en deux brins d'ADN intacts. (Figure 17)²²²

Les gènes *BRCA1* et *BRCA2* sont des gènes dits « supprimeurs de tumeur », ils codent respectivement pour les protéines BRCA1 et BRCA2 qui jouent des rôles essentiels dans ce mécanisme de RH (Figure 17) et dans le maintien de l'intégrité du génome. Dans un premier temps, la protéine BRCA1 interagit directement avec la protéine CtIP phosphorylée couplée au complexe MRN formant ainsi le complexe BRCA1-C. Cette interaction permet de catalyser la résection des extrémités de la DSB afin de générer des extrémités simple brin. La protéine BRCA1 agit également comme antagoniste de 53BP1, protéine ayant pour fonction de bloquer la résection des DSB et ainsi, par inhibition du système de RH, de conduire à la mort cellulaire. Par la suite, BRCA1 interagit avec la protéine PALB2 permettant ainsi le recrutement de la protéine BRCA2. La protéine BRCA1 est un acteur indispensable du processus de RH, les cellules déficientes pour BRCA1 étant incapables de fixer PALB2, BRCA2 et RAD51 au niveau de la cassure.^{223,224} La protéine BRCA2 est également un acteur essentiel de la RH qui participe au recrutement de la protéine RAD51 sur le site de la DSB, et assure ainsi son action d'invasion des extrémités simple brin. De même les cellules déficientes pour BRCA2 sont incapables de lier RAD51, perturbant le processus de RH.^{223,224}

De manière synthétique, les tumeurs présentant une HRD sont susceptibles de répondre aux mêmes thérapeutiques et peuvent être classées en deux grandes catégories selon les anomalies moléculaires :

- les tumeurs présentant une anomalie moléculaire délétère des gènes *BRCA1* ou *BRCA2*,
- les tumeurs présentant un profil *BRCAness*, terme apparu en 2004 pour définir un phénotype similaire à celui des cancers mutés *BRCA1* ou *BRCA2*.²²⁵

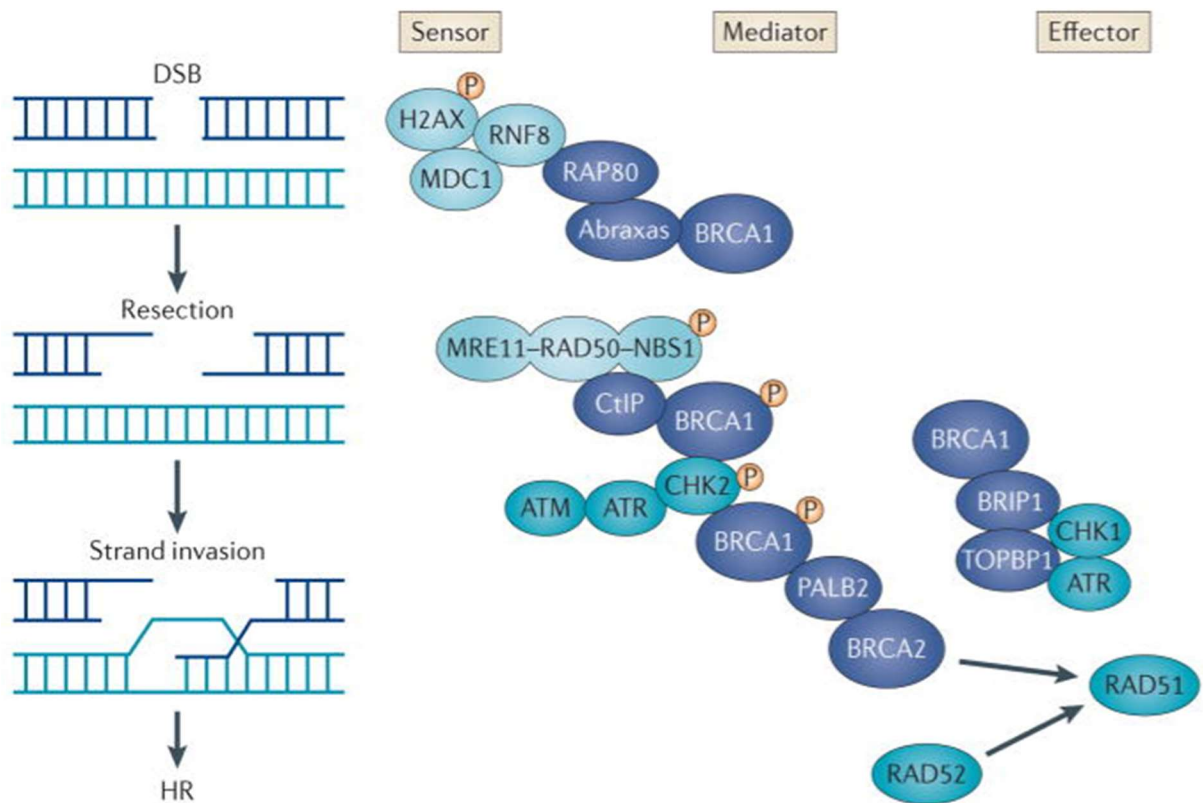


Figure 16 : Les principaux acteurs protéiques de la réparation de l'ADN par recombinaison homologue, d'après Roy et al.²¹⁸

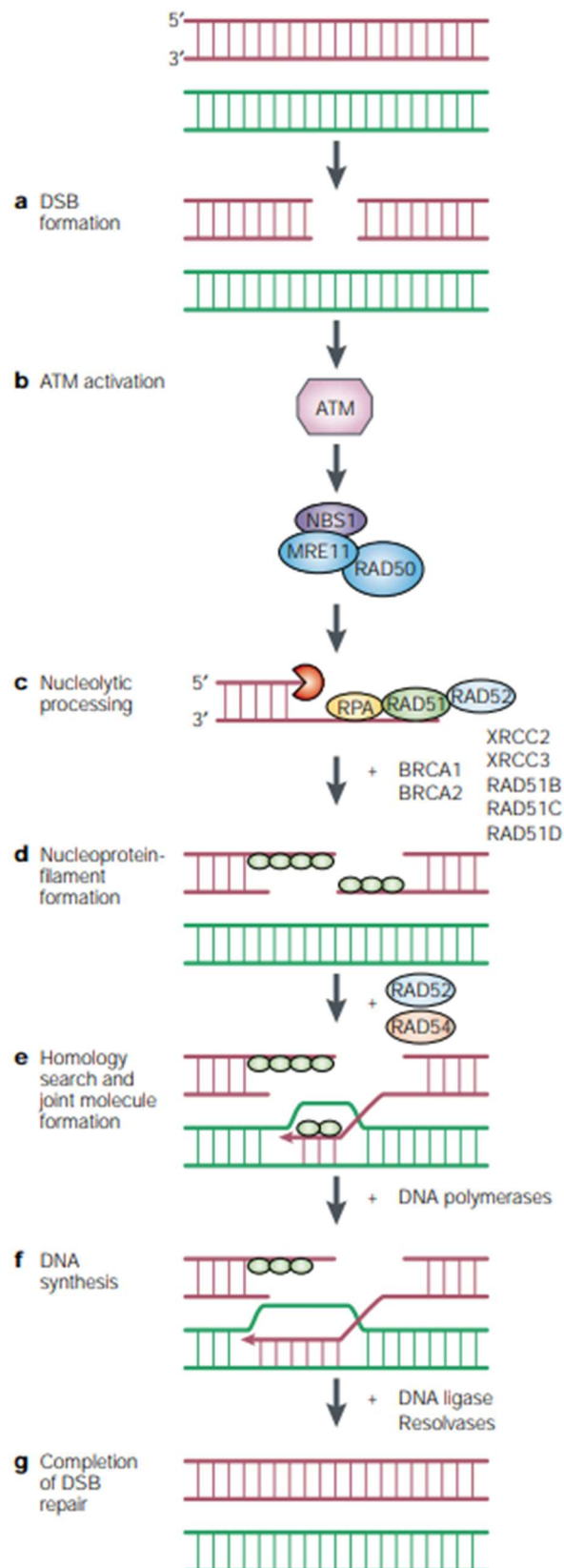


Figure 17 : Les étapes de la réparation des cassures double brin de l'ADN par recombinaison homologue, d'après van Gent et al.²²²

3.3. Tumeurs mutées *BRCA1* et *BRCA2*

a) *BRCA1* : généralités

- Structure du gène *BRCA1*

A partir de l'hypothèse d'un modèle de transmission autosomique dominant préalablement établi par étude de ségrégation, Mary-Claire King est parvenue à localiser *BRCA1* en 17q21 en 1990.²²⁶ Sa séquence est composée de 5592 nucléotides et code une protéine de 1863 AA. L'ADN génomique s'étend sur 81 kb et comporte 24 exons dont deux considérés comme non codants : l'exon 1 (majeure partie de la région 5'UTR, en amont du codon d'initiation de traduction ATG situé dans l'exon 2) et l'exon 4 (séquence Alu non codante initialement identifiée comme un exon). L'exon 11 représente 3426 pb (60% de la séquence codante de *BRCA1*), les autres exons sont de plus petite taille. Les séquences introniques (de tailles variant de 403 pb à 9,2 kb) représentent 91% du gène et contiennent une forte quantité de séquences répétées de type Alu.

- Protéine *BRCA1*

Il s'agit d'une protéine de 1863 AA dont seules les régions N- et C- terminales sont fortement conservées chez les mammifères.²²⁷ Les principaux domaines fonctionnels de *BRCA1* sont (Figure 17)²¹⁸:

- Un domaine catalytique RING en N-terminal (AA 24 à 64) riche en résidus cystéines. Ces AA organisés en 4 hélices alpha antiparallèles groupées formant des structures en doigt de zinc indispensables aux interactions protéiques dans cette région, en particulier dans la formation d'hétérodimères *BRCA1* avec la protéine *BARD1*.
- Deux domaines BRCT (*BRCA1 C-Terminal domain*) retrouvés à l'extrémité C-terminal. Ils sont constitués d'environ 100 AA. Ils sont repliés en 4 feuillets Beta parallèles entourés de deux hélices alpha d'un côté et d'une hélice alpha unique de l'autre. Ils permettent l'interaction de *BRCA1* avec de nombreux partenaires dont *Abraxas*, *BRIP1* et *CtIP*.
- Deux signaux de localisation nucléaire (NLS pour *Nuclear Localization Signals*)
- Un domaine superhélice (*coiled coil*) permettant à *BRCA1* d'interagir avec la protéine *PALB2* lors du processus de réparation par RH.

- Fonctions de *BRCA1* en dehors de la RH

Outre son rôle dans la RH, *BRCA1* est souvent qualifié de « gardien de l'intégrité du génome » car *BRCA1* intervient dans de très nombreux processus cellulaires, nucléaires ou cytoplasmiques, la plupart du temps par l'intermédiaire d'interactions avec d'autres protéines :

- La régulation du cycle cellulaire, en assurant le passage d'une phase à l'autre ou l'arrêt de la progression du cycle^{228,229}

- La régulation de la transcription, en modulant l'activité de différents facteurs de transcription^{230,231}
- Le remodelage de la chromatine²³²
- L'activité ubiquitine ligase : BRCA1 s'associe de manière constitutive à BARD1 (*BRCA1 Associated Ring Domain 1*), ce complexe hétérodimérique permet d'induire l'activité ubiquitine ligase E3 de BRCA1, à l'origine de l'ubiquitination des histones mais également de nombreuses protéines interagissant avec BRCA1^{233,234}
- La régulation du stress oxydatif : le mécanisme est encore mal compris, mais il a été rapporté qu'en absence de BRCA1, le niveau des réactifs oxygénés appelés ROS (*Reactive Oxygen Stress*) augmente tandis qu'une surexpression de BRCA1 fait diminuer la concentration des ROS^{235,236}
- La mitophagie (voie spécifique d'autophagie visant à dégrader les mitochondries défectueuses) : l'absence de protéine BRCA1 fonctionnelle entraîne l'inhibition de la voie de mitophagie compromettant ainsi tout le réseau mitochondrial et donc la survie cellulaire.²³⁷

b) BRCA2 : généralités

- Structure du gène *BRCA2*

Wooster *et al.* ont identifié en 1995 le gène *BRCA2* dans la région chromosomique 13q12-q13 impliquée dans des histoires familiales de cancer du sein dans des familles islandaises.²³⁸ Par la suite, Tavtigian *et al.* (1996) ont déterminé la séquence codante complète du gène et sa structure exonique ainsi que son mode d'expression.²³⁹ La séquence codante du gène *BRCA2* est composée de 10254 nucléotides et code une protéine de 3418 AA. L'ADN génomique s'étend sur 84 kb et comporte 27 exons avec trois grands exons : l'exon 10 (1,1 kb), l'exon 11 (4,9 kb) et l'exon 27 (1,5 kb). Les autres exons ont une taille variant de 41 pb à 428 pb. Le codon initiateur ATG est situé dans l'exon 2 et l'extrémité 5'UTR est représentée par l'exon 1.

- Protéine BRCA2

Le gène *BRCA2* encode une protéine de 3418 AA. Celle-ci présente très peu d'homologie avec d'autres protéines et aucune avec BRCA1. Les principaux domaines fonctionnels sont les suivants (Figure 17)²¹⁸ :

- Huit motifs répétés BRC d'environ 35 AA chacun dans la partie centrale de la protéine, ce domaine est impliqué dans l'interaction de BRCA2 avec RAD51.²⁴⁰
- Un motif comprenant 190 AA principalement en hélices alpha (domaine dit «hélical»)

- Trois domaines d'environ 110 AA (OB1, OB2 et OB3) présentant une homologie avec le repliement OB présent chez plusieurs protéines liant l'ADN simple brin
- Deux NLS
- Un site de phosphorylation par des kinases cyclines dépendantes (CDK) qui semble réguler l'association de RAD51 à un deuxième motif de liaison à cet endroit.²⁴¹

- Fonctions de BRCA2 en dehors de la RH

Tout comme BRCA1, la protéine BRCA2 est impliquée dans de nombreux processus cellulaires et participe à la stabilité du génome :

- La régulation du cycle cellulaire, en contrôlant le passage de la phase G2 à la phase M²⁴²
- La régulation de la transcription, en médiant l'activité de facteurs de transcription²⁴³
- La mitophagie²³⁷
- Le maintien de l'intégrité des télomères : grâce au recrutement de RAD51 sur l'ADN des télomères, BRCA2 prévient la dégradation et la fusion des chromosomes et facilite la réplication^{244,245}
- La stabilisation des fourches de réplication de l'ADN.²⁴⁶

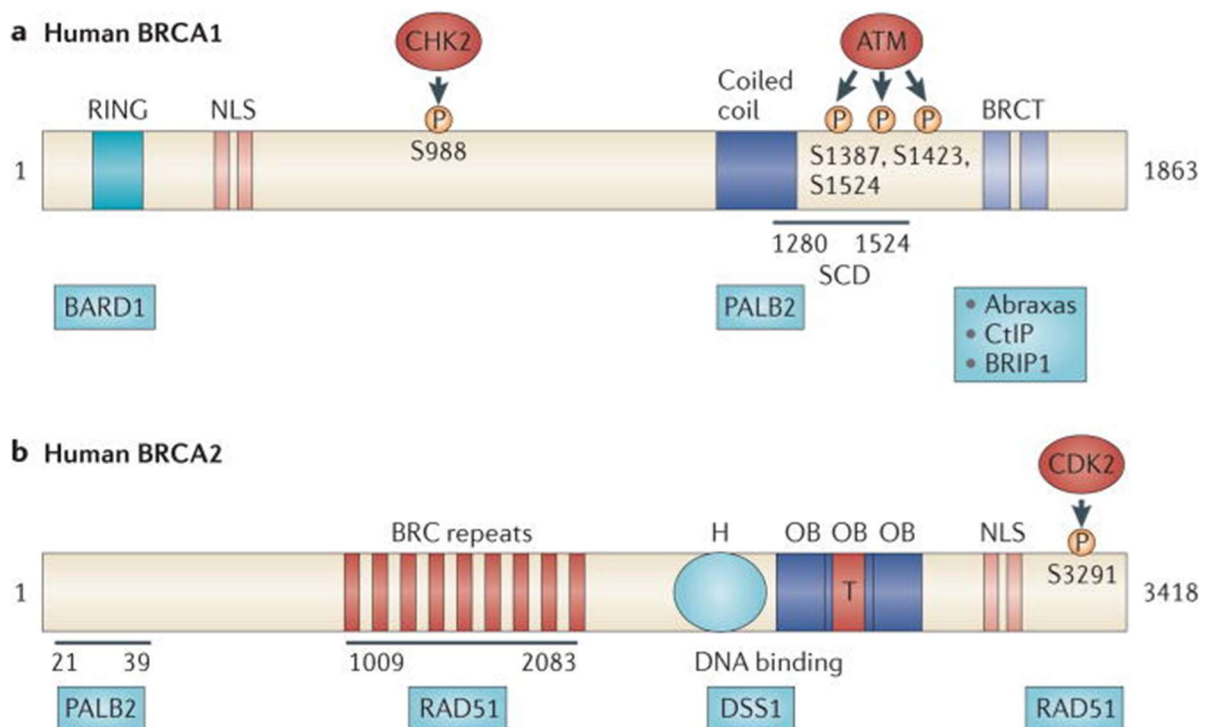


Figure 18. Structure de domaines fonctionnels de BRCA1 et BRCA2, d'après Roy et al.²¹⁸

c) *BRCA1*, *BRCA2* et prédisposition aux cancers

Les altérations constitutionnelles pathogènes des gènes suppresseurs de tumeur *BRCA1* et *BRCA2* augmentent fortement le risque de cancer, principalement au niveau des tissus mammaires et ovariens. L'expression de ces deux gènes étant ubiquitaire, plusieurs hypothèses sont envisagées pour expliquer cette spécificité tissulaire : la présence de mécanismes de réparation redondants dans les autres tissus²⁴⁷, la perte d'hétérozygotie très fréquente dans les tissus mammaires et ovariens²⁴⁸, l'influence des hormones de type œstrogène²⁴⁹ ou encore l'impact de facteurs cancérigènes non génétiques^{250,251}.

Les altérations constitutionnelles hétérozygotes pathogènes des gènes *BRCA* sont caractérisées par une forte pénétrance toutefois incomplète et variable selon le gène altéré, l'âge et le tissu. Ainsi, une anomalie responsable de la perte de fonction d'un de ces gènes sera à l'origine d'un risque plus au moins important de développer une tumeur du spectre, sans induire un risque total. Concernant les cancers mammaires, le risque cumulé d'atteinte d'un cancer du sein à l'âge de 80 ans est de 72% et 69% chez les femmes porteuses d'une altération pathogène de *BRCA1* ou *BRCA2*, respectivement. Ce risque augmente très rapidement entre l'âge de 30 et 40 ans pour les porteuses d'une variation de *BRCA1*, et entre 40 et 50 ans pour les porteuses d'une variation de *BRCA2*. De plus, une femme ayant été atteinte d'un premier cancer du sein présentera 40% et 26% de probabilité de développer un cancer du sein controlatéral 20 ans après le premier diagnostic, si celle-ci est porteuse d'une variation pathogène de *BRCA1* ou *BRCA2*, respectivement.²⁵² Chez les hommes, une anomalie de *BRCA2* conduira à un sur-risque de cancer du sein.^{253,254} Concernant le tissu ovarien, le risque cumulé d'atteinte d'un cancer de l'ovaire à l'âge de 80 ans est de 40% et 17% chez les femmes porteuses d'une anomalie de *BRCA1* et *BRCA2*, respectivement²⁵². Concernant les autres tissus, une variation pathogène de *BRCA2* détectée chez un individu, homme ou femme, conduira à un risque cumulé estimé à 5 à 10% de développer un cancer du pancréas²⁵⁵ et à un sur-risque de développer un mélanome ou un cancer de l'estomac.²⁵⁶⁻²⁵⁸ Chez les hommes, une anomalie de *BRCA2* conduira à un risque supplémentaire de 3 à 8,6% de développer un cancer de la prostate.²⁵⁵

d) Pathologie moléculaire des gènes *BRCA1* et *BRCA2*

Les gènes *BRCA* sont considérés comme des gènes suppresseurs de tumeur et répondent donc à l'hypothèse « *two-hit* » que Knudson avait initialement décrite en 1971 sur des cas de rétinoblastomes.²⁵⁹ Selon ce modèle, les deux allèles du gène doivent être inactivés pour rendre la cellule vulnérable à l'instabilité génomique et induire la tumorigenèse. Le premier « *hit* » correspond à l'inactivation du premier allèle,

conséquence d'une altération constitutionnelle pathogène entraînant une « perte de fonction » de l'un des gènes *BRCA* tandis que le second « *hit* » survient sur le deuxième allèle dans les cellules à l'origine du tissu tumoral, sous la forme d'une altération somatique acquise au cours de la vie. Ce deuxième « *hit* » peut correspondre à une perte d'hétérozygotie (LOH pour *Loss Of Heterozygosity*), une variation nucléotidique ponctuelle pathogène, ou encore une altération épigénétique comme une hyperméthylation du promoteur.²⁶⁰ Du fait de l'inactivation des deux allèles, la voie de réparation par RH n'est plus fonctionnelle et des lésions s'accumulent dans le génome de la cellule modifiant ainsi son phénotype basal vers un éventuel phénotype tumoral.

En pratique, on reconnaît deux types d'altération de séquence des gènes *BRCA* :

- **les mutations ponctuelles** dont l'effet dépend de la position au niveau :
 - Des séquences régulatrices en amont du gène ou dans la région promotrice : les altérations peuvent entraîner une modification quantitative de l'expression du gène (par altération de séquences consensus transcriptionnelles)
 - Des régions codantes : les substitutions de type variations faux-sens ou neutres peuvent affecter le codon d'initiation de la traduction, les différents domaines fonctionnels ou les sites d'interaction avec les partenaires. Les insertions ou délétions de petite taille non multiples de trois bases peuvent décaler le cadre de lecture de l'ARN messager (ARNm) (*frameshift*) aboutissant le plus souvent à l'apparition d'un codon stop prématuré entraînant la synthèse d'une protéine tronquée ou l'absence de protéine. Les insertions ou délétions de petite taille multiples de trois bases peuvent aussi ne pas décaler le cadre de lecture (*inframe*), dans ce cas la traduction de l'ARNm altéré aboutit à une protéine présentant un ou plusieurs acides aminés ajoutés ou supprimés. Certaines variations, dites splicéogéniques, sont capables d'affecter l'épissage. Il peut s'agir d'une part de variations au niveau des jonctions exon/intron reconnues par le splicéosome sur la base de séquences spécifiques consensus, les sites 5' et 3' d'épissage (5'/3'ss, *splice site*), également appelés sites donneurs et accepteurs d'épissage, respectivement. La séquence des 5'ss inclut les trois dernières bases de l'exon et les six premières bases de l'intron tandis que la séquence des 3'ss comprend les vingt dernières bases de l'intron et les trois premières bases de l'exon.²⁶¹⁻²⁶³ D'autre part, certaines variations au niveau du site de branchement et la région riche en pyrimidine essentiels dans l'interaction du splicéosome et des facteurs de régulation avec la séquence du préARNm peuvent affecter l'épissage.²⁶⁴ Enfin toute variation à distance des sites d'épissage constitutifs peut impacter l'épissage en modifiant des éléments de régulation (SRE, *Splicing Regulatory Elements*), la force d'un site d'épissage alternatif ou cryptique ou en créant un site d'épissage de novo.^{265,266}

- Des introns : les mutations introniques peuvent perturber l'épissage par altération des séquences consensus régulatrices. A noter que l'introduction d'un nouveau site d'épissage dans une région intronique est susceptible de créer un nouvel exon appelé pseudo-exon si un site complémentaire préexiste à proximité. A titre d'exemple, il a été montré que le variant c.6937+594T>G de *BRCA2* induit une augmentation de la force d'un 5'ss cryptique et conduit ainsi à l'inclusion d'un pseudo-exon de 95 nucléotides entre l'exon 12 et 13 du gène *BRCA2*, responsable d'un décalage du cadre de lecture.²⁶⁷ De même, la variation ponctuelle intronique profonde c.4185+4105C>T dans l'intron 12 de *BRCA1* est à l'origine d'un pseudo-exon générant un transcrit aberrant et une protéine non fonctionnelle.²⁶⁸
- Des régions en aval du gène : elles peuvent perturber l'arrêt de la transcription ou la stabilité de l'ARNm.
 - **les réarrangements génomiques de grande taille (RGT)** (translocations, inversions, délétions ou duplications d'un exon au gène entier, voir plus) : ceux intéressant les régions promotrices peuvent altérer l'expression d'un gène ; les autres, en fonction de leur taille et de leur position, aboutissent à une protéine tronquée ou à l'absence de protéine.

e) Analyse des gènes *BRCA1* et *BRCA2*

Les gènes *BRCA* sont de grande taille avec des mutations pouvant être réparties sur toute la longueur des gènes, ce qui nécessite un séquençage complet des parties codantes et introniques avoisinantes. L'analyse des gènes *BRCA* est principalement réalisée par NGS avec la technologie de capture ou par amplicons. La capture semble être plus appropriée et plus performante.²⁶⁹ Le point faible du NGS est le défaut de détection de grands réarrangements génomiques (notamment les délétions et duplications), nécessitant une analyse complémentaire du nombre de copies par technique MLPA dans le cadre de l'analyse germinale.²⁷⁰ Cette technique n'est malheureusement pas applicable dans le cadre de l'analyse somatique à cause de la qualité de l'ADN tumoral extrait très fragmenté. Par ailleurs, les techniques de NGS, et plus particulièrement la méthode par amplicons, s'avèrent perfectibles pour la mise en évidence de mutations dans les régions homopolymériques. Comme conclu dans l'article de Wallace en 2016, compte tenu des difficultés que présente l'analyse tumorale des gènes *BRCA* (qualité de l'ADN tumoral, non détection des RGT), celle-ci ne peut se substituer à l'analyse constitutionnelle.²⁷¹

Il est cependant crucial de noter que certaines anomalies pathogènes des gènes *BRCA* peuvent n'être que somatiques, c'est-à-dire présentes uniquement dans les cellules

tumorales, c'est le cas de 5 à 10% des carcinomes ovariens^{272,273}, environ 3% des cancers du sein²⁷⁴ et plus de 12% des cancers avancés de la prostate²⁷⁵.

3.4. Anomalies moléculaires des tumeurs de profil *BRCAness* et signatures HRD

a) Mécanismes alternatifs de déficit de la voie de RH

De nombreux travaux dont ceux du TCGA ont montré qu'il y avait d'autres mécanismes de HRD que l'inactivation directe des gènes *BRCA*. Les principaux mécanismes décrits sont l'hyperméthylation du promoteur de *BRCA1*, les amplifications de *EMSY* (co-facteur de *BRCA2* et répresseur transcriptionnel) et l'inactivation d'autres gènes individuels altérant la réparation de l'ADN. Progressivement le terme statut HRD vient remplacer le terme *BRCAness* dans la littérature.

- L'hyperméthylation du promoteur de *BRCA1* correspond à une méthylation aberrante des cytosines des îlots CpG qui sont sous-représentés dans le génome et généralement localisés au niveau des régions promotrices des gènes. Ce mécanisme épigénétique est responsable d'une perte d'expression plus ou moins marquée de *BRCA1*. Différentes études ont démontré que l'hyperméthylation du promoteur de *BRCA1* est spécifique aux cellules tumorales, généralement absente des cellules saines, et fréquemment associée à une LOH. Ce mécanisme est principalement rapporté dans les cancers du sein et de l'ovaire à une même fréquence de 10 à 15%.²⁷⁶⁻²⁷⁸ Les caractéristiques tumorales de l'hyperméthylation du promoteur de *BRCA1* sont similaires à celles observées en présence de variations pathogènes des gènes *BRCA*, cependant elle n'a pas la même valeur pronostique.²⁷²
- La surexpression de *EMSY*, qui interagit avec l'exon 3 de *BRCA2*, pourrait constituer un mécanisme d'inactivation de *BRCA2* dans les cancers du sein et de l'ovaire non héréditaires.^{279,280}
- Altérations d'autres gènes individuels : le concept *BRCAness* comprend initialement les gènes acteurs de la RH ou régulant la fourche de réplication (Figure 16) et les gènes de la famille FANC. Des altérations de ces gènes sont retrouvées dans de nombreux cancers dont ceux de l'ovaire, du sein, de la prostate et du pancréas.²⁸¹ Afin de mieux caractériser les altérations de ces gènes dans les différentes localisations tumorales, différentes études utilisant de larges catalogues de gènes voire l'analyse de l'exome ou du génome entier ont été menées (Tableau 12). Cependant les fréquences de

HRD publiées sont dépendantes de nombreux facteurs liés aux études : technique utilisée, gènes de RH pris en compte, types histologiques de cancers étudiés (par exemple cancers du sein de type triple négatif versus tous sous-types confondus), stade du cancer (primitif ou à progression, localisé ou métastatique...).

Compte tenu de la diversité et de la complexité des mécanismes en cause, l'identification de signatures de HRD est progressivement devenue indispensable pour évaluer l'efficacité de la RH plutôt que des déficiences sur des mécanismes individuels.

Localisation tumorale	Nombre de patients testés	Fréquence de HRD	Référence
Pancréas	100	24%	Waddell et al., 2015 ²⁸²
	262	19%	Park et al., 2020 ²⁸³
Prostate	150	19%	Robinson et al., 2015 ²⁸⁴
	333	19%	Cancer Genome Atlas Network., 2015 ²⁸⁵
Sein	507	30%	Cancer Genome Atlas Network., 2012 ²⁸⁶
Ovaire	390	31%	Pennington et al., 2014 ²⁷³
	489	49%	Cancer Genome Atlas Research Network, 2011 ²⁷²

Tableau 12 : Quelques fréquences de HRD publiées sur de nombreux échantillons (n>100)

b) Les signatures HRD

Plusieurs équipes se sont attelées à l'identification des tumeurs HRD via des signatures transcriptionnelles et mutationnelles. Les premières signatures sont basées sur la mise en évidence de grands réarrangements génomiques témoins d'une instabilité génomique due à un système de réparation inefficace de l'ADN. Il s'agit plutôt de « cicatrices », témoins d'une instabilité génomique, que de tests fonctionnels. Ces signatures seraient alors des biomarqueurs permettant de prédire la réponse au iPARP.²⁸⁷⁻²⁸⁹

Les premières études transcriptionnelles s'intéressent à l'expression de tumeurs mutées *BRCA* pour définir un profil *BRCAness*. De manière générale, ce profil montre une meilleure SG et prédit une réponse aux iPARP et aux sels de platine.^{290,291} Cependant les difficultés rencontrées lors des analyses transcriptomiques et les progrès technologiques fulgurants des analyses génomiques ont progressivement favorisé l'élaboration de nouvelles signatures mutationnelles.

Le développement d'outils bio-informatiques a considérablement amélioré l'exploitation des analyses génomiques. Au-delà des SNV et InDels, la bioinformatique donne aujourd'hui accès à des données concernant la ploïdie, les nombres de copies et le

statut allélique de chaque région chromosomique aboutissant à un paysage de plus en plus précis du génome tumoral. En 2013 sur cette base, trois équipes indépendantes ont proposé des signatures performantes pour la HRD :

- La signature développée à l'Institut Curie repose sur le nombre de changements (ou cassures) de statut chromosomique (variation du nombre de copies et/ou du nombre d'allèles), témoignant d'un remaniement chromosomique. De façon surprenante, le nombre total de cassures n'est pas corrélé au statut de RH. Mais en limitant la mesure aux cassures séparant de grandes régions chromosomiques de plus de 10 Mb (> 107 pb), la mesure du nombre de changements d'état de grande taille appelées LST (*Large Scale state Transition*) (Figure 19), est parfaitement corrélée au statut de RH. L'analyse combinée des LST et des données de séquençage pangénomique d'un nombre restreint de tumeurs montre une bonne corrélation entre LST et translocation interchromosomique.²⁹² La mise au point de cette signature a été faite dans des séries de cancers du sein de sous-type triple négatif (TNBC). La signature a été ensuite validée dans l'ensemble des histotypes de cancers du sein et dans le cancer de l'ovaire.^{293,294}
- La signature développée à Boston par l'équipe d'Andrea Richardson repose sur la mesure du déséquilibre allélique dans les régions subtélomériques appelé TAI (*Telomeric Allelic Imbalance*) (Figure 19). Ce biomarqueur a été mis au point dans des TNBC et dans des cancers séreux ovariens dans lesquels sa positivité a été associée à une meilleure réponse initiale aux sels de platine.²⁹⁵
- Celle initialement développée par la société Myriad Genetics repose sur la mesure globale de l'équilibre allélique c'est-à-dire de LOH. La signature HRD-LOH repose sur la détection de grandes régions (> 15 Mb) de perte d'hétérozygotie (Figure 19).²⁹⁶

Il est probable qu'il s'agisse de trois approches bio-informatiques mesurant le même paramètre d'instabilité génomique. L'exploitation de ces trois signatures a été obtenue par Myriad Genetics qui propose une solution commerciale (myChoice® HRD) basée sur la recherche de mutations et RGT des gènes *BRCA* et un score d'instabilité génomique appelé GIS (*Genomic Instability Status*) associant LST, TAI et LOH (Figure 19). Myriad Genetics a présenté un ensemble de données montrant la plus grande performance de ce score combiné dont le *cut-off* est 42, notamment pour prédire la sensibilité aux sels de platine et aux iPARP.²⁹⁷ L'utilisation de ce GIS dans un essai clinique a permis de doubler le nombre de femmes qui, atteintes d'un carcinome de l'ovaire, pouvaient bénéficier d'iPARP.²⁹⁸ Le GIS et le test myChoice® HRD sont par la suite progressivement apparus dans les recommandations cliniques internationales.²⁹⁹⁻³⁰¹ Ce test n'est actuellement disponible que de manière délocalisée via un envoi de matériel FFPE au laboratoire Myriad de Salt

Lake City, la mise en place de ce test en Europe est en cours dans quelques laboratoires satellites.

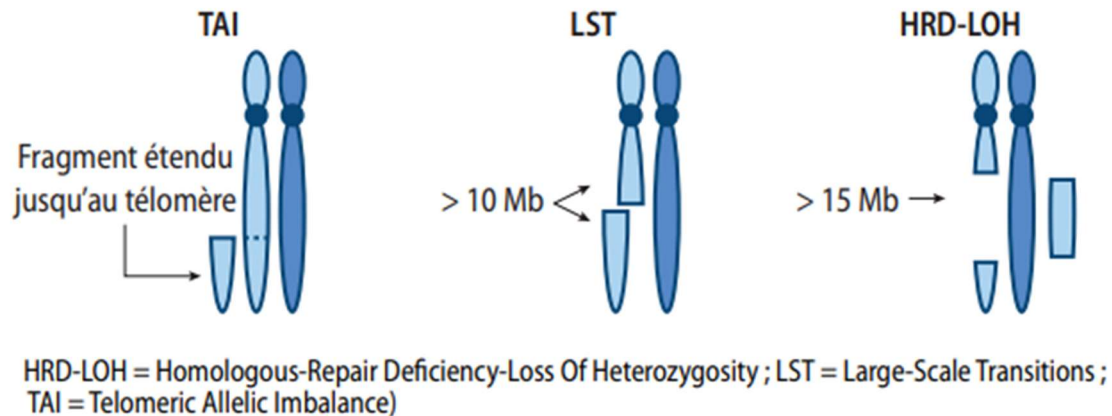


Figure 19 : Principe des trois biomarqueurs constituant le GIS de la solution commerciale myChoice® HRD, d'après Rodrigues et al., 2016³⁰²

D'autres signatures HRD sont peu à peu développées :

- FoundationFocus™ CDX_{BRCA} LOH commercialisé par la société Roche (également disponible dans le test FoundationOne®CDx) : ce test, à ce jour utilisé uniquement dans les cancers ovariens, repose sur la recherche tumorale des mutations pathogènes des gènes *BRCA* et sur la détection d'une LOH rendue « *high* » ou « *low* » en fonction du pourcentage de perte d'hétérozygotie (avec un cut-off à 16). Les régions de LOH sont déduites sur les 22 chromosomes autosomes en utilisant le profil du nombre de copies à l'échelle du génome et les fréquences alléliques sur les SNP. Cette signature HRD a été validée lors de l'essai ARIEL3 dans lequel le statut HRD montrant un bénéfice de l'iPARP (rucaparib) dans la population présentant un statut HRD FoundationFocus CDX_{BRCA} LOH est associé à une amélioration de la survie sans progression avec traitement d'entretien par rucaparib.^{303,304} Ce test moléculaire tumoral sur échantillon FFPE n'est aujourd'hui disponible que de manière externalisée.
- AmoyDx® HRD Focus Panel (marqué CE-IVD, société AmoyDx) : il s'agit d'un test réalisé par la technologie HANDLE (*Halo-shape ANnealing and Defer-Ligation Enrichment*) évaluant le statut HRD par détection des mutations pathogènes des gènes *BRCA* et par un score de cicatrice génomique appelé GSS (*Genomic Scar Score*) avec cut-off à 50. Une seule étude utilisant ce test est rapportée à ce jour : il s'agit d'une étude chinoise évaluant l'efficacité et la tolérance de l'iPARP (niraparib) dans des cancers ovariens.³⁰⁵
- Illumina HRD GIS via le test TruSight™ Oncology 500 à usage de recherche pour le moment (société Illumina) : ce test réalisé par NGS permet d'identifier des variants

dans 523 gènes et évalue l'instabilité microsatellitaire, le TMB et le statut HRD de l'échantillon tumoral testé.

- Sophia DDM HRD Solution (société Sophia Genetics, test commercialisé en février 2022) : ce test est basé sur l'analyse NGS d'un panel de 28 gènes combinée à la technique WGS « *low-pass* » (LP-WGS). Le LP-WGS est une technologie de NGS dont le principe est de séquencer le génome entier à faible profondeur et donc à faible coût. Les informations recueillies sur les SNP, les CNV et les déséquilibres chromosomiques sont exploitées par une analyse bioinformatique aboutissant à un score d'instabilité.
- Développement de scores à partir de techniques SNParray : OncoScanDx (société ThermoFisher) et Infinium CytoCNP-850K v1.2 BeadChip (société Illumina). Le principe de la technique SNParray est d'identifier des déséquilibres chromosomiques en étudiant les fréquences alléliques d'un grand nombre de SNP. Les algorithmes bioinformatiques sont en cours d'étude pour définir des scores à partir des données produites par ces tests moléculaires.
- Des solutions académiques sont en cours de développement et utilisent diverses techniques (*SNParray*, NGS, *deepHRD*, *CGHarray*...).

Ainsi, la recherche d'HRD tumorale, a minima dans les cancers de l'ovaire, est devenue un impératif. Il semble primordial, pour des raisons de coût et de délai de prise en charge clinique, de réaliser la recherche d'altérations *BRCA* et de statut HRD lors de la même analyse. L'utilisation de signatures validées d'HRD dans la caractérisation génétique des tumeurs représente un enjeu majeur, la comparaison des différentes solutions est une étape importante dans le souci d'une égalité de prise en charge des patients. Les tests AmoyDx® HRD Focus Panel et FoundationOne®CDx ont montré une concordance, mais pas une équivalence totale, avec myChoice® HRD dans une étude publiée en 2021.³⁰⁶ La même équipe a comparé le score Illumina HRD GIS avec myChoice® HRD : les GIS étaient fortement corrélés (0,98), les prévalences des statuts HRD et des altérations *BRCA* étaient similaires et les discordances concernaient des cas avec des valeurs proches du seuil de positivité.³⁰⁷ Le test développé par Sophia Genetics n'est aujourd'hui comparé avec myChoice® HRD que sur de faibles nombres d'échantillons, sans données publiées à l'heure actuelle.

3.5. Indications et biomarqueurs de prescription des iPARP

a) Cancers de l'ovaire

On estime à environ 5190 le nombre de nouveaux cas de cancers de l'ovaire en 2018 en France. Il s'agit du 8^{ème} cancer le plus fréquent chez les femmes, responsable de 3500 décès par an.⁵ Le cancer de l'ovaire est fréquemment découvert à un stade avancé, principalement sous la forme d'une carcinose péritonéale, en raison de la continuité entre la surface des ovaires et des trompes avec le péritoine. Le tableau clinique initial conduit le plus souvent à la réalisation d'une coelioscopie exploratrice à visée diagnostique lors de laquelle des biopsies sont effectuées. Ces prélèvements sont souvent les seuls disponibles au cours de la maladie pour réaliser les tests moléculaires.

Les cancers ovariens d'origine épithéliale représentent 90% des cas et regroupent des sous-types histologiques distincts qui présentent des caractéristiques génomiques différentes en fonction des types de dommages de l'ADN.³⁰⁸ Les cancers ovariens séreux de haut grade ou HGSOC (*high grade serous ovarian cancer*) représentent 70% des cancers épithéliaux de l'ovaire et sont ceux concernés par les indications des iPARP. Les autres sous-types épithéliaux sont nettement moins fréquents : endométrioides (environ 10%), à cellules claires (environ 12%), mucineux (3%), tumeurs de Brenner (1%) et histologies mixtes (environ 5%).³⁰⁹ Dans 18 à 23% des cancers ovariens, des mutations d'origine héréditaire augmentent le risque de développement tumoral.^{310,311} Deux types d'altérations en sont principalement responsables : les variations pathogènes des gènes *BRCA1* et *BRCA2* présentes dans 20% des HGSOC et les anomalies du système *mismatch repair* (MMR) préférentiellement retrouvées dans les sous-types endométrioides.^{311,312}

Le profil moléculaire des HGSOC (Figure 20) a été largement étudié au cours de ces deux dernières décennies. Les altérations les plus fréquentes sont : les mutations de *TP53* retrouvées dans 96% des HGSOC, les mutations des gènes *BRCA* (20-25%) et d'autres gènes de la voie de RH, les hyperméthylations de *BRCA1*, les altérations du nombre de copies. Globalement les HRD représentent la moitié des tumeurs.^{272,313}

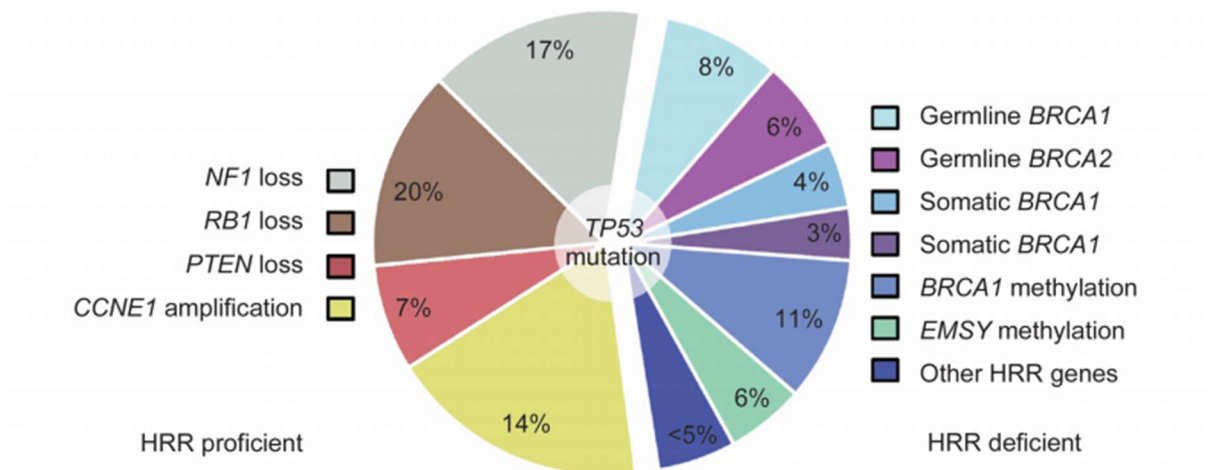


Figure 20 : Principales altérations moléculaires identifiées dans les HGSOV, d'après Hollis et Gourley, 2016³¹³

La prise en charge thérapeutique des patientes atteintes de HGSOV est décidée en RCP et comprend une chirurgie et/ou une chimiothérapie en fonction notamment de l'histologie, du stade de la maladie, de l'état général et des comorbidités éventuelles de la patiente. La chimiothérapie standard en première ligne associe généralement carboplatine et paclitaxel pour une durée de 6 cycles. Depuis les années 2010, on a assisté à une formidable avancée dans le cancer de l'ovaire avec l'entrée de cette pathologie dans l'ère de la médecine personnalisée grâce aux iPARP. Des essais cliniques ont dans un premier temps montré une efficacité très significative des iPARP en termes de survie sans récurrence chez les femmes atteintes de HGSOV porteuses d'une altération *BRCA1* ou *BRCA2* et présentant une maladie platine-sensible. L'étude SOLO1 a prouvé la spectaculaire efficacité de 2 ans d'olaparib en traitement de maintenance après une chimiothérapie de première ligne pour les patientes de stade FIGO III/IV ayant une mutation *BRCA1* ou *BRCA2*. A 5 ans, 48 % des patientes traitées par olaparib ne présentaient pas de progression tumorale versus 21% avec le placebo.³¹⁴ L'étude SOLO-2, a étudié l'efficacité de l'olaparib en maintenance après une rechute platine-sensible chez les patientes ayant une mutation *BRCA1* ou *BRCA2* : la survie sans progression était de 19,1 mois sous olaparib versus 5,5 mois sous placebo.³¹⁵ L'étude PAOLA-1 a ensuite rapporté le bénéfice de l'adjonction de 2 ans d'olaparib en maintenance au traitement standard par bevacizumab. La survie sans progression médiane était allongée de 15 mois en cas de mutation *BRCA* et de 11 mois en cas de déficit de la RH non lié à *BRCA*, introduisant l'HRD comme nouveau biomarqueur.²⁹⁸ Un autre inhibiteur de PARP, a été évalué en entretien après chimiothérapie de première ligne dans l'étude PRIMA. Dans cette étude, le niraparib améliorait la survie sans progression en cas de mutation *BRCA* ou de déficit de la RH non

lié à *BRCA* mais également, de façon plus modeste, en l'absence de déficit de la RH.³¹⁶ En fin l'étude VELIA a montré une PFS significativement plus longue avec le veliparib (versus placebo) en association avec la chimiothérapie puis en traitement de maintenance.³¹⁷ Ces différentes études, qui ne sont malheureusement pas comparables entre elles, ont abouti aux indications des iPARP dans le cancer de l'ovaire (Tableau 13), l'olaparib étant le premier iPARP à avoir obtenu son AMM en 2014.

Traitement d'entretien des cancer épithéliaux avancés de haut grade de l'ovaire, en réponse à une première ligne de chimiothérapie à base de platine	
Olaparib	En monothérapie en cas de mutation <i>BRCA</i> En association avec le bevacizumab en cas de mutation <i>BRCA</i> ou de statut HRD
Niraparib	En l'absence de mutation <i>BRCA</i> ou en cas de contre-indication au bevacizumab
Traitement d'entretien des cancers épithéliaux de haut grade de l'ovaire récidivants et sensibles au platine et qui sont en réponse à une chimiothérapie à base de platine.	
Olaparib	En cas de mutation <i>BRCA</i>
Niraparib	Quel que soit le statut <i>BRCA</i> , seulement dans les cancers séreux
Rucaparib	Quels que soient le statut <i>BRCA</i> et l'histologie

Tableau 13. Disponibilité des iPARP en France dans le cancer de l'ovaire (mars 2022)^{18,19}

b) Cancers du sein

En 2018, on recensait 58500 nouveaux cas de cancer du sein en France, ce qui le place au 2^{ème} rang des cancers les plus fréquents dans la population générale française et au premier rang chez les femmes. La mortalité s'élève à 12100 décès par an en 2018. Au contraire du cancer de l'ovaire, 60 % des cancers du sein sont détectés à un stade précoce. De nombreuses actions de dépistage ont été progressivement mises en place pour diminuer l'incidence de cancer : autopalpation et consultation médicale en cas de modification des seins, examen clinique tous les ans à partir de 25 ans, mammographie de dépistage tous les deux ans entre 50 et 74 ans sans symptôme ni facteur de risque autre que l'âge. Des modalités de suivi spécifiques sont recommandées pour les femmes présentant des antécédents médicaux personnels ou familiaux, ou certaines prédispositions génétiques.⁵

Le cancer du sein est une maladie hétérogène dont les classifications histologiques et cliniques ont été progressivement complétées par une classification moléculaire basée sur l'expression de trois marqueurs : les récepteurs aux estrogènes (RE), les récepteurs à la progestérone (RP), récepteurs HER2 (HER2). Actuellement, quatre sous-types sont bien individualisés : luminal (comprenant les sous-types A et B), basal, HER2+ et normal. Ces catégories moléculaires ont des pronostics différents apportant un intérêt majeur dans la

prise en charge thérapeutique.³¹⁸ Les données du TCGA publiées en 2012 montrent que les profils génomiques et protéomiques présentent une hétérogénéité entre les sous-types mais aussi dans les sous-types de cancers mammaires (Tableau 14). A noter que, tout sous-type confondu, des mutations sur 3 gènes (*TP53*, *PIK3CA* et *GATA3*) ont une incidence de plus de 10%. De manière intéressante, la comparaison des tumeurs mammaires de sous-type basal comprenant une grande majorité de TNBC avec les tumeurs ovariennes a révélé de nombreuses similitudes moléculaires, laissant supposer une étiologie et des possibilités thérapeutiques communes. (Tableau 14)²⁸⁶ Enfin, on estime que 5 à 10% des cancers du sein féminins sont liés à une prédisposition héréditaire par mutations des gènes *BRCA*, cette fréquence augmente dans les TNBC (14 à 20% des cas).^{319,320}

Subtype	Luminal A	Luminal B	Basal-like	HER2E
ER+/HER2- (%)	87	82	10	20
HER2+ (%)	7	15	2	68
TNBCs (%)	2	1	80	9
TP53 pathway	<i>TP53</i> mut (12%); gain of <i>MDM2</i> (14%)	<i>TP53</i> mut (32%); gain of <i>MDM2</i> (31%)	<i>TP53</i> mut (84%); gain of <i>MDM2</i> (14%)	<i>TP53</i> mut (75%); gain of <i>MDM2</i> (30%)
PIK3CA/PTEN pathway	<i>PIK3CA</i> mut (49%); <i>PTEN</i> mut/loss (13%); <i>INPP4B</i> loss (9%)	<i>PIK3CA</i> mut (32%); <i>PTEN</i> mut/loss (24%); <i>INPP4B</i> loss (16%)	<i>PIK3CA</i> mut (7%); <i>PTEN</i> mut/loss (35%); <i>INPP4B</i> loss (30%)	<i>PIK3CA</i> mut (42%); <i>PTEN</i> mut/loss (19%); <i>INPP4B</i> loss (30%)
RB1 pathway	Cyclin D1 amp (29%); <i>CDK4</i> gain (14%); low expression of <i>CDKN2C</i> ; high expression of RB1	Cyclin D1 amp (58%); <i>CDK4</i> gain (25%)	<i>RB1</i> mut/loss (20%); cyclin E1 amp (9%); high expression of <i>CDKN2A</i> ; low expression of RB1	Cyclin D1 amp (38%); <i>CDK4</i> gain (24%)
mRNA expression	High ER cluster; low proliferation	Lower ER cluster; high proliferation	Basal signature; high proliferation	HER2 amplicon signature; high proliferation
Copy number	Most diploid; many with quiet genomes; 1q, 8q, 8p11 gain; 8p, 16q loss; 11q13.3 amp (24%)	Most aneuploid; many with focal amp; 1q, 8q, 8p11 gain; 8p, 16q loss; 11q13.3 amp (51%); 8p11.23 amp (28%)	Most aneuploid; high genomic instability; 1q, 10p gain; 8p, 5q loss; <i>MYC</i> focal gain (40%)	Most aneuploid; high genomic instability; 1q, 8q gain; 8p loss; 17q12 focal <i>ERBB2</i> amp (71%)
DNA mutations	<i>PIK3CA</i> (49%); <i>TP53</i> (12%); <i>GATA3</i> (14%); <i>MAP3K1</i> (14%)	<i>TP53</i> (32%); <i>PIK3CA</i> (32%); <i>MAP3K1</i> (5%)	<i>TP53</i> (84%); <i>PIK3CA</i> (7%)	<i>TP53</i> (75%); <i>PIK3CA</i> (42%); <i>PIK3R1</i> (8%)
DNA methylation	—	Hypermethylated phenotype for subset	Hypomethylated	—
Protein expression	High oestrogen signalling; high MYB; RPPA reactive subtypes	Less oestrogen signalling; high FOXM1 and MYC; RPPA reactive subtypes	High expression of DNA repair proteins, <i>PTEN</i> and <i>INPP4B</i> loss signature (pAKT)	High protein and phosphoprotein expression of EGFR and HER2

Percentages are based on 466 tumour overlap list. Amp, amplification; mut, mutation.

Tableau 14 : Caractéristiques moléculaires des sous-types de cancer du sein (données TCGA 2012)²⁸⁶

Il existe aujourd’hui de nombreux traitements utilisés dans les cancers du sein : la chirurgie, la radiothérapie, l'hormonothérapie, la chimiothérapie et les thérapies ciblées. Le choix du traitement, validé en RCP, repose sur des critères clinico-pathologiques comme la taille de la tumeur, l’atteinte ganglionnaire, le grade histo-pronostique, les embolies vasculaires, les récepteurs hormonaux, la surexpression d’HER2, le Ki67 et l’âge, mais aussi sur des critères génomiques. Il existe aujourd’hui 4 signatures génomiques basées sur l’analyse de 12 à 70 gènes (Oncotype DX, MammaPrint, Prosigna et EndoPredict) qui proposent des scores prédictifs de réponse à la chimiothérapie.³²¹ Concernant les thérapies ciblées par iPARP, les études de phase III OlympiAD et Embraca ont permis l’obtention des AMM respectivement de l’olaparib et du talazoparib dans le cancer du sein localement avancé ou métastatique HER2 négatif et avec une mutation germinale des gènes *BRCA1/2*. Dans chacune de ces 2 études, la supériorité de l’iPARP versus une chimiothérapie a été démontrée sur la PFS (la médiane de PFS a été de 7,03 mois dans le groupe olaparib versus

4,17 mois dans le groupe chimiothérapie dans OlympiAD et 8,6 mois dans le groupe talazoparib versus 5,6 mois dans le groupe chimiothérapie dans Embraca). En revanche, la supériorité des iPARP n'a pas été démontrée sur la SG. En pratique, le choix entre olaparib et talazoparib est fait au cas par cas en prenant notamment en compte le profil de tolérance de chacun de ces traitements.^{19,322,323} Compte tenu de la forte incidence du cancer sein, ces autorisations d'iPARP en présence d'une mutation *BRCA* lèvent le problème de la population à tester. Les recommandations de l'ESMO en 2021 proposent de rechercher ces altérations au minimum pour les patients avec un cancer du sein récidivant métastatique de sous type triple négatif ou RE+ HER2- ayant progressé sous hormonothérapie.^{324,325} Des études sont en cours pour évaluer l'efficacité des iPARP en fonction du statut HRD, en présence de mutations somatiques ou encore dans les cancers du sein localisés.

c) Cancers du pancréas

En 2018, 14100 nouveaux cas de cancer du pancréas ont été recensés en France. Malgré des progrès diagnostiques et thérapeutiques, le cancer du pancréas reste un cancer de mauvais pronostic dont l'incidence progresse, en majorité chez les personnes de plus de 50 ans. Le taux de survie nette standardisée à 5 ans n'est que de 11%. Du fait d'une évolution longtemps asymptomatique, seuls 10 à 20 % des patients sont diagnostiqués à un stade où la tumeur est résécable. Plus 90% des cancers du pancréas sont des adénocarcinomes.⁵

Les analyses moléculaires tumorales révèlent qu'une large majorité des cancers du pancréas présentent une mutation activatrice de l'oncogène *KRAS* et des altérations du suppresseur de tumeur *TP53* (respectivement 93% et 72% selon TCGA). Les mutations germinales et somatiques des gènes de réparation de l'ADN *BRCA2*, *PALB2* et *ATM* sont observées dans 8% des échantillons, représentant une classe de patients bons répondeurs à la chimiothérapie à base de platine et/ou aux iPARP. (Figure 21)³²⁶ Environ 10% des cancers du pancréas sont liés à des syndromes génétiques de prédisposition héréditaire dont le syndrome sein-ovaire associé à une mutation des gènes *BRCA* fait partie. Les altérations des gènes *BRCA1/2* sont présentes chez 4 à 7% des patients atteints de cancer du pancréas, les mutations *BRCA2* étant largement majoritaires.^{327,328}

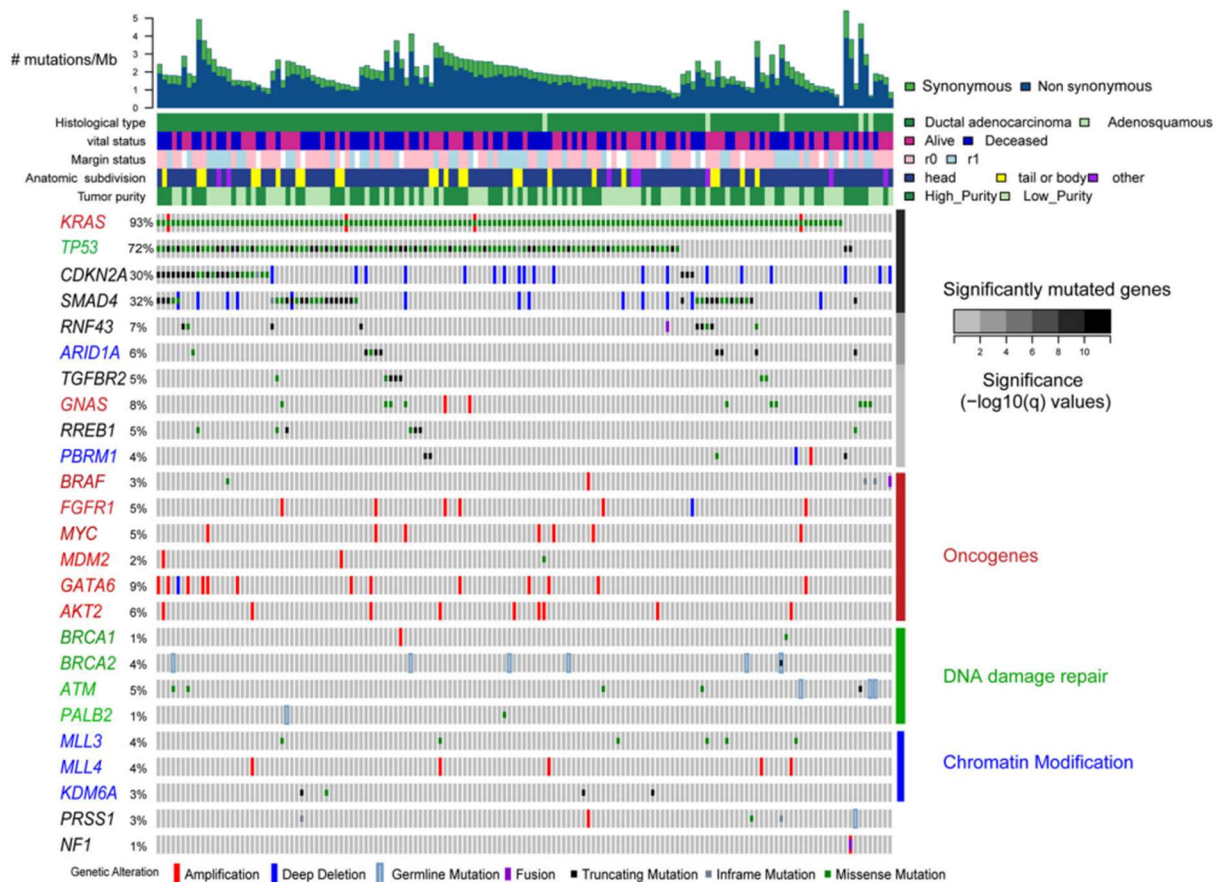


Figure 21 : Profil moléculaire de 149 cancers du pancréas (données TCGA)³²⁶

La chirurgie et la chimiothérapie conventionnelle, parfois associées à une radiothérapie, sont les principaux traitements des adénocarcinomes du pancréas (seuls ou associés). Jusqu'en 2020, il n'existait pas de médicament disposant d'une AMM dans le traitement d'entretien dans cette pathologie. Les patients en première ligne (protocole FOLFIRINOX ou association nab-paclitaxel/gemcitabine) étaient traités jusqu'à progression de la maladie ou la survenue d'une toxicité importante ; ensuite le relais était pris par une chimiothérapie de seconde ligne. Fin 2020, les conclusions de l'étude POLO (étude de phase III, randomisée, en double aveugle versus placebo) ont permis l'obtention de l'AMM de l'olaparib dans le traitement d'entretien des patients adultes atteints d'un adénocarcinome du pancréas métastatique avec mutation germinale des gènes *BRCA1/2* sans signe de progression après au minimum 16 semaines de chimiothérapie de première ligne à base de platine et non éligibles à la poursuite de celle-ci. En effet, un gain de PFS de +3,6 mois a été observé sous olaparib versus placebo dans cette population cible, sans démonstration de gain en SG. Cette option de traitement d'entretien concerne donc actuellement un nombre très limité de patients.^{19,329}

d) Cancers de la prostate

Le cancer de la prostate (PCa) représente 25 % des cancers masculins, c'est le cancer le plus fréquent chez l'homme. Rare avant 50 ans, son incidence augmente progressivement avec l'âge. C'est un cancer de bon voire très bon pronostic, avec un taux de survie à 5 ans élevé (93%). Le facteur pronostique majeur de ce cancer est le stade au diagnostic : 80 % des cancers sont diagnostiqués alors qu'ils sont encore localisés à la prostate. Environ 50 000 nouveaux cas de PCa sont estimés en France chaque année. La principale caractéristique de ce cancer est son évolution généralement lente, sur plusieurs années. Dans plus de 90% des cas, l'histologie du PCa est un adénocarcinome, mais on rencontre également des sarcomes et des carcinomes de la prostate.⁵

Plusieurs classifications moléculaires de PCa ont été établies, elles peuvent être liées ou non à l'évolution de la maladie. Les analyses moléculaires ont permis d'identifier de multiples altérations génétiques récurrentes, qui conduisent à la dérégulation de différents processus biologiques, comme ceux impliqués dans le développement de la prostate, la régulation du cycle cellulaire, la voie de signalisation des androgènes, l'organisation de la chromatine ou la réparation de l'ADN. Parmi les altérations impliquées, les plus fréquentes concernent les fusions de gènes impliquant la famille de facteurs de transcription *ETS*, les mutations des gènes *SPOP*, *FOXA1*, *IDH1*, des altérations chromosomiques au niveau des gènes *c-MYC*, *EGFR*, et *NKX3.1*, pour les formes localisées de PCa ; auxquelles s'ajoutent des altérations des gènes *PTEN*, *TP53*, *AR* (récepteur aux androgènes), *CDK12*, *RB1*, *CDKN1B* et des gènes de réparation de l'ADN pour les formes plus avancées. (Figure 22)³³⁰

On considère qu'environ 5% des PCa sont liés à une prédisposition héréditaire. Les principaux gènes identifiés dans les formes héréditaires avec un âge précoce de survenue sont *HOXB13* impliqué dans le développement de la prostate (notamment la mutation G84E)³³¹ et certains gènes de réparation de l'ADN, le plus fréquemment *BRCA2*, mais aussi *ATM*, *CHEK2*, *BRCA1*, *PALB2*, *RAD51D*, *ATR*, *NBN*, *PMS2* (Figure 23).³³²

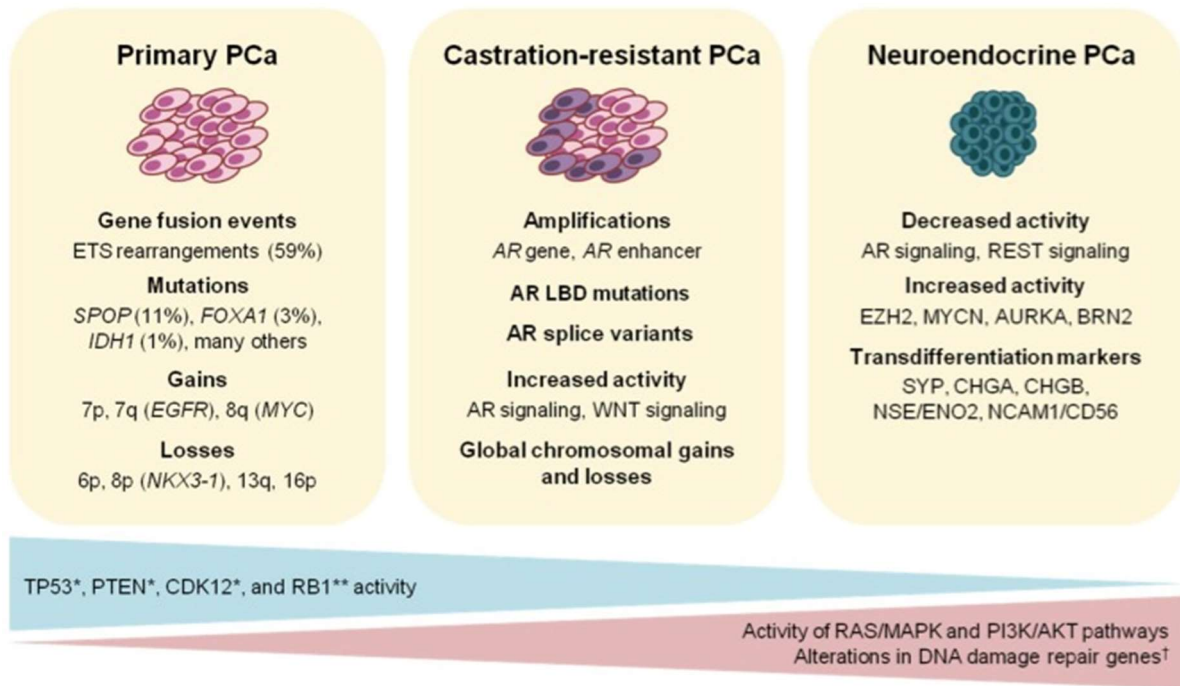


Figure 22 : Caractéristiques moléculaires des différents stades de cancer de la prostate³³⁰

L'inactivation est davantage observée dans les PCa résistant à la castration (CR)* et dans les tumeurs neuroendocrines (NE)**. †Les altérations des voies de réparation de l'ADN sont les plus marquées dans les PCa CR et quasiment absent des tumeurs NE. AR : androgen receptor, LBD : ligand binding domain

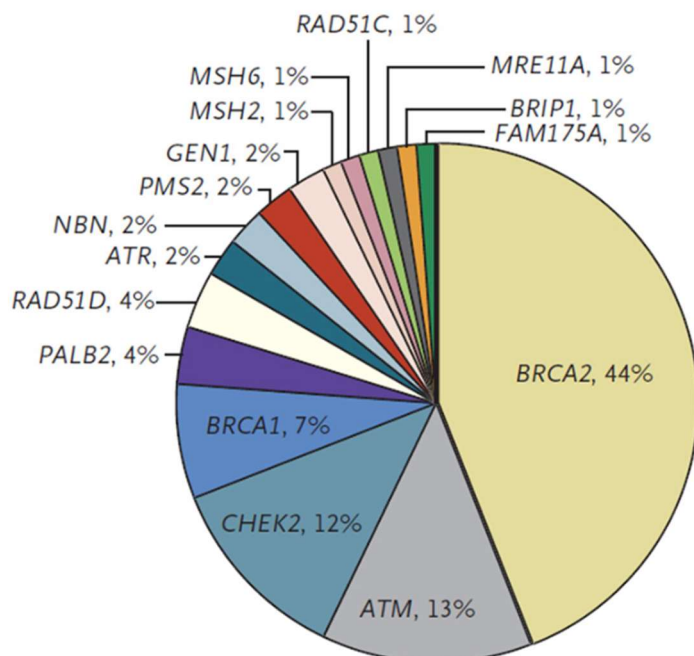


Figure 23 : Distribution des mutations pathogènes de gènes de réparation de l'ADN impliquées dans les PCa métastatiques.³³²

Plusieurs traitements peuvent être proposés dans les PCa : chirurgie, radiothérapie externe, curiethérapie, hormonothérapie, surveillance active. Cette stratégie dépend des caractéristiques du cancer déterminées lors du diagnostic : localisation, type histologique, stade, grade, âge, état général et comorbidités du patient. Dans les PCa, les iPARP ont trouvé leur place dans les stades métastatiques résistants à la castration (CPRCm). Selon les recommandations du Comité de Cancérologie de l'Association Française d'Urologie, la stratégie thérapeutique du CPRCm reposait jusqu'en 2021 sur les hormonothérapies de nouvelle génération (acétate d'abiratérone et enzalutamide) ou sur une chimiothérapie par taxane (docétaxel ou cabazitaxel). L'étude de phase III PROfound a été la première étude randomisée évaluant un iPARP dans le CPRCm. Cette étude a fait la démonstration d'une supériorité de l'olaparib par rapport à l'enzalutamide ou à l'acétate d'abiratérone dans la cohorte incluant les patients avec mutations *BRCA1/2* ou *ATM* avec un gain de +3,84 mois en médiane de PFS radiologique (7,39 mois versus 3,55 mois) et de +4,40 mois en SG (19,09 mois versus 14,69 mois).^{333,334} L'olaparib a donc obtenu en mai 2021 une AMM en monothérapie pour le traitement des patients atteints d'un CPRCm, avec mutation des gènes *BRCA1/2* (germinale et/ou somatique) et qui ont progressé après un traitement antérieur incluant une hormonothérapie de nouvelle génération (enzalutamide ou acétate d'abiratérone).¹⁹ Des études sont en cours pour évaluer l'efficacité des iPARP en première ligne, en association à l'abiratérone ou encore en fonction du statut HRD.

Compte tenu de la lente évolution des PCa, l'indication des iPARP au stade métastatique implique l'analyse de tissus FFPE très anciens pour détecter des variations pathogènes des gènes *BRCA*. Les contraintes techniques liées au séquençage d'ADN tumoral très altéré entraîne le développement progressif de ces analyses sur l'ADNtc. Mais la détection des altérations sur l'ensemble de ces deux gènes suppresseurs de tumeurs dans l'ADNtc est encore un défi.

3.6. Interprétation des VSI des gènes *BRCA1* et *BRCA2* pour la prescription des iPARP : place des tests fonctionnels

a) VSI de *BRCA1* et *BRCA2*

Le développement des consultations d'oncogénétique, des iPARP et du NGS ont permis de mettre en évidence des milliers de variants *BRCA1* et *BRCA2*. Les différentes bases de données classifient de nombreux polymorphismes et variants pathogènes mais aussi un grand nombre de VSI qui restent aujourd'hui à annoter que ce soit dans un but de conseil génétique ou dans un but théranostique. En avril 2021, environ 40 000 et 25 000 variants *BRCA1/BRCA2* ont été signalés dans les bases de données internationales BRCAexchange et ClinVar, respectivement, dont près de 80% et 50%, respectivement, ne

sont pas classés en termes d'effet pathogène.³³⁵ La plupart des variants présents dans ces bases sont des variants constitutionnels, quelques-uns sont somatiques. Comme le montrent les données de la base française FrOG (*French OncoGenetics*) qui remplace la base UMD-BRCAsShare (Figure 24), le nombre de VSI est en constante augmentation compte tenu des extensions d'indications des iPARP dans différentes localisations tumorales.

Il existe trois abords pour évaluer ces VSI *BRCA1/2* :

- L'abord clinique (pour les variants constitutionnels) via l'étude de co-ségrégation familiale du VSI avec la maladie, l'histoire familiale, les caractéristiques tumorales. En France, le Groupe génétique et cancer (GGC) d'Unicancer, créé en 1991, évalue les risques familiaux de cancer, élabore les recommandations de prise en charge des familles et contribue à l'amélioration des connaissances sur les prédispositions génétiques aux cancers. Selon les données du GGC, en 2010, une mutation causale utilisable pour le conseil génétique est mise en évidence dans environ 13% des cas index testés. Des VSI concernent 11000 familles pour *BRCA1* et 13500 familles pour *BRCA2*, ils sont mis en évidence chez plus de 10% des patients testés. Pour ces familles, des tests génétiques ne peuvent pas être proposés aux apparentés et le conseil génétique est orienté par l'histoire familiale. Depuis 2012 et jusqu'en décembre 2022, le GGC et l'Institut Curie proposent partout en France un protocole de recherche appelé COVAR qui organise des études de co-ségrégation pour les VSI de *BRCA1/2* et *PALB2*, afin de définir le caractère causal ou non de ces variations. A ce jour, parmi les 653 variants sélectionnés dans l'étude COVAR, 101 variants distincts partagés par 1 624 familles ont été classés comme étant pathogènes/probablement pathogènes ou bénins/probablement bénins grâce à cette étude. Soixante-six des 101 variants classés par COVAR seraient restés des VSI sans les données de co-ségrégation qui s'avèrent donc être un outil puissant de classification.^{336,337}
- L'abord bioinformatique avec les signatures notamment HRD (développées précédemment) et les outils de prédiction *in silico* (qui ne seront pas détaillés dans ce manuscrit).
- L'abord biologique via l'étude fonctionnelle du VSI qu'elle soit *in vitro* ou *in vivo* (paragraphe ci-après).

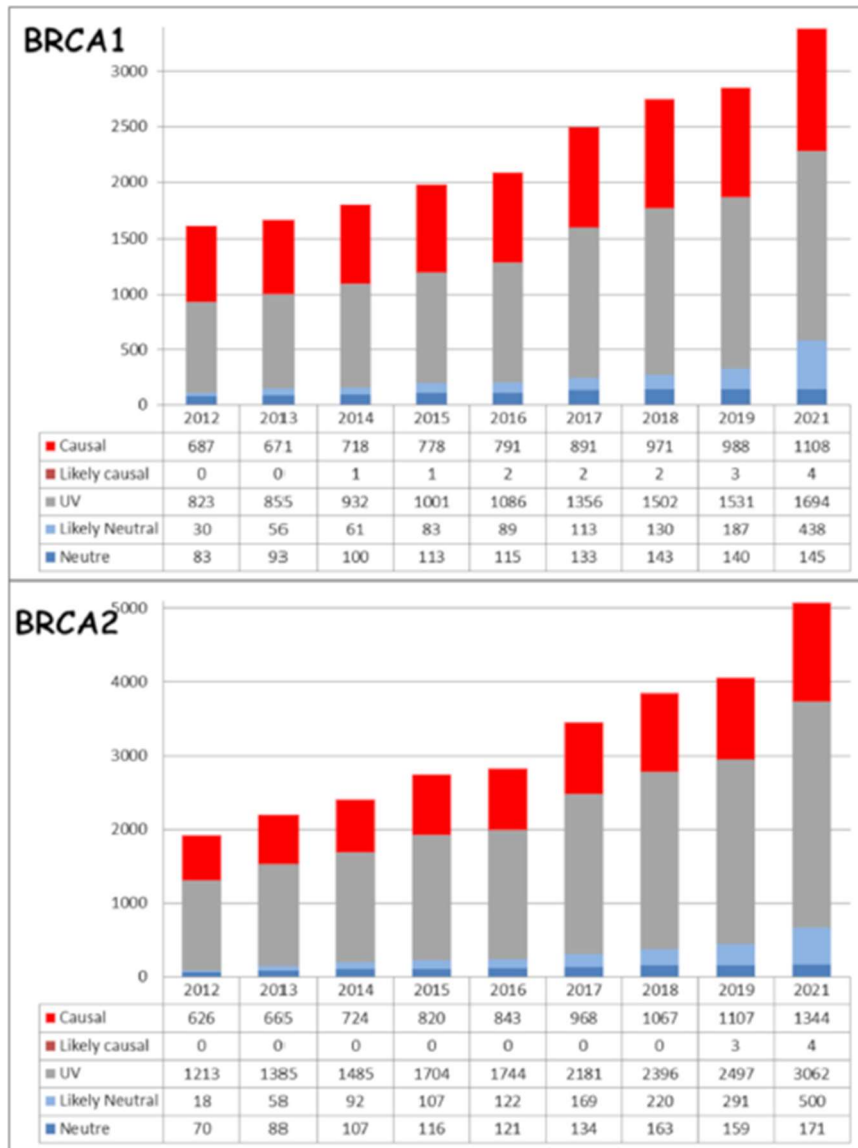


Figure 24 : Augmentation du nombre de variants dans la base de données FrOG depuis 2012, d'après Caputo et al., 2021³³⁷

La caractérisation fonctionnelle des variants est un enjeu majeur en oncogénétique et en thérapeutique. Certaines approches ont été proposées mais sont difficilement applicables en pratique clinique quotidienne, par exemple l'étude de l'inactivation de *BRCA1/2* consistant à mettre en évidence l'arrêt de la relocalisation de RAD51 sur les sites de dommages et donc la disparition des foci nucléaires. A l'heure actuelle ces tests ne sont pas validés comme marqueurs de prédiction de réponse au iPARP.^{292,338} Les approches fonctionnelles les plus intéressantes à ce jour sont celles utilisées pour mettre en évidence les anomalies d'épissage et celles utilisant très récemment les technologies d'édition génomique CRISPR-Cas9.

b) Caractérisation biologique fonctionnelle des VSI d'épissage des gènes *BRCA1* et *BRCA2*

L'épissage est un processus extrêmement dynamique qui met en jeu de multiples signaux et des facteurs de régulation permettant une régulation complexe très fine de ce mécanisme moléculaire. La complexité de celle-ci est illustrée notamment par la spécificité tissulaire de certains épissages alternatifs³³⁹ mais également par les modifications de l'épissage observées lors du processus de tumorigenèse³⁴⁰. Concernant les gènes *BRCA*, des travaux menés à partir de matériels biologiques dérivés de sang d'individus témoins ont permis de détecter expérimentalement par RT-PCR, en plus du transcrite de référence, la présence de 68 et 27 transcrits alternatifs générés à partir des gènes *BRCA1* et *BRCA2*, respectivement.³⁴¹⁻³⁴³ A noter que toutes les isoformes ne conduisent pas systématiquement à une protéine ou une protéine fonctionnelle.³⁴⁴ L'épissage peut être altéré par des variations génétiques, si celles-ci touchent un des nombreux acteurs intervenant dans ce processus avec pour conséquence éventuelle une modification du profil d'épissage des transcrits. Théoriquement toute variation nucléotidique, qu'elle soit exonique ou intronique, est susceptible d'altérer l'épissage.³⁴⁵ Ces anomalies d'épissage, sont très fréquentes puisque 9% des variations pathogènes de *BRCA1* seraient des variants d'épissage.³⁴⁶ De nombreux outils bioinformatiques de prédiction mais également un certain nombre d'approches expérimentales ont été implémentés dans le but de contribuer à la mise en évidence et à l'interprétation précise des variations à l'origine d'un défaut d'épissage.

Les approches expérimentales de détection des anomalies d'épissage induites par des variations génétiques reposent sur une analyse comparative des profils d'épissage des transcrits exprimés dans le contexte sauvage avec ceux produits dans le contexte muté. Ces données de nature qualitative et semi-quantitative sont essentielles pour appréhender les conséquences fonctionnelles et donc le caractère pathogène potentiel des mutations splicéogéniques.

- Analyse des profils d'épissage des transcrits par RT-PCR à partir d'ARN de patients : après migration des produits de RT-PCR par électrophorèse sur gel d'agarose, les profils d'épissage obtenus pour les échantillons sont comparés et les différents amplicons observés sont isolés et séquencés afin de déterminer la nature exacte des isoformes produites. La simple migration des produits de RT-PCR sur gel d'agarose ne permet a priori ni de discriminer les transcrits présentant une faible différence de taille ni d'évaluer avec précision les taux relatifs des différents transcrits potentiellement produits.³⁴⁷⁻³⁴⁹ Plus récemment, des approches quantitatives ont été intégrées à la RT-PCR, ainsi l'analyse de la cinétique complète de la réaction d'amplification permet d'obtenir une quantification précise des transcrits : PCR temps réel³⁵⁰, PCR digitale³⁵¹.

- Analyse de l'expression allélique à partir de l'ARN de patients : Le but est d'appréhender la contribution précise de l'allèle muté dans la production respective des transcrits normaux et aberrants. L'une des méthodes les plus utilisées est celle basée sur l'extension d'amorces, appelée SNUPE (*Single Nucleotide Primer Extension*) ou SNaPshot. Cette technique repose sur l'extension d'une amorce hybridée sur l'ADNc en 5' ou 3' de la variation exonique d'intérêt. L'extension de cette amorce via l'incorporation de didésoxyriboNucléotides TriPhosphates (dNTP) fluorescents complémentaires à la séquence d'intérêt permet de discriminer les transcrits produits à partir de l'allèle sauvage ou muté. Cette approche permet de détecter avec précision un éventuel déséquilibre allélique résultant de la dégradation par le système de régulation *nonsense-mediated decay* (NMD) de l'allèle porteur d'un codon stop prématuré mais également d'évaluer la sévérité du défaut d'épissage induit par la variation. Ces deux points ont pu être mis en évidence dans le cas de variations localisées dans les gènes *BRCA*.³⁵² L'analyse de l'expression allélique peut également être appréhendée par la méthode quantitative de pyroséquençage qui peut de la même manière directement cibler la variation d'intérêt dans l'ADNc du patient. Cette technique permet, de détecter la perte d'expression d'un allèle et de quantifier l'impact de variations exoniques sur les profils d'épissage des transcrits comme démontré dans le cas de variations exoniques du gène *BRCA2*.³⁴⁸
- Analyses des défauts d'épissage par la technique de RNA-seq : L'étude du transcriptome par RNA-seq présente le double avantage de déterminer la séquence de l'ensemble des transcrits exprimés dans un échantillon et d'estimer leurs taux relatifs. Le RNA-seq ciblé est une approche visant à concentrer l'analyse sur les transcrits de certains gènes permettant la détection d'un plus grand nombre d'isoformes d'épissage. Les épissages alternatifs naturels des gènes *BRCA* ont ainsi pu être caractérisés et sont une aide à l'interprétation des variations splicéogéniques à l'origine d'évènements aberrants d'épissage.^{353,354} Le RNAseq ciblé permet donc à la fois d'identifier d'éventuelles anomalies d'épissage résultant de VSI mais également d'analyser l'expression allélique. Cette approche est cependant dépendante de la disponibilité de l'ARN extrait à partir d'un matériel biologique du patient et ne tient pas compte de la dégradation éventuelle de certains transcrits par le système NMD.

- Autres approches expérimentales : minigènes et transgènes

Des approches expérimentales évaluant l'effet de l'introduction de variations sur l'épissage dans des systèmes plus ou moins artificiels ont été développées.

Les méthodes de clonage moléculaire basées sur l'utilisation de minigènes permettent de caractériser finement l'effet des variations sur l'épissage. Le principe repose sur la transfection transitoire dans des cellules de vecteurs d'expression comportant un promoteur capable d'induire la transcription d'une séquence d'intérêt appelée minigène (séquences comprenant au moins deux exons séparés par un intron, représentant ainsi une version simplifiée d'un gène multi-exon). Après la transfection en parallèle des minigènes sauvage et muté, il est alors possible de réaliser une analyse comparative de leurs profils d'épissage par RT-PCR et séquençage. Cette approche est utilisée pour la caractérisation fonctionnelle de variation d'épissage dans différentes régions des gènes *BRCA1* et *BRCA2*.^{355,356}

Plus récemment, des tests basés sur l'utilisation de transgènes ont été développés pour étudier l'effet des variations sur l'épissage dans un contexte génomique le plus naturel possible. Ainsi, l'approche FLGSA, (*Full-Length Gene Splicing Assay*) utilise la transfection d'un vecteur d'expression renfermant certains gènes, du codon d'initiation de la traduction au codon de terminaison de la traduction.³⁵⁷ Des vecteurs spécifiques appelés BAC (*Bacterial Artificial Chromosome*), dans lesquels il est possible de cloner de larges fragments génomiques (jusqu'à 300 Kb) peuvent renfermer l'intégralité des gènes humains *BRCA1* ou *BRCA2* (avec leurs promoteurs respectifs) pour étudier l'impact des variations sur l'épissage mais également sur la fonction de la protéine (survie cellulaire, activité de RH et sensibilité à certains agents génotoxiques). Après introduction de ces vecteurs sauvages ou mutés dans des cellules souches de souris, l'impact des variations introduites par mutagenèse est examiné par l'analyse RT-PCR des profils d'épissage des transcrits produits à partir des transgènes BAC.^{358,359}

Ces stratégies permettent de se rapprocher du contexte génomique naturel mais elles requièrent l'expression d'un transgène. Afin de dépasser cette limite, il est aujourd'hui possible d'introduire les variations génétiques d'intérêt, non limitée aux variations d'épissage, directement dans le locus chromosomique endogène grâce notamment à un changement de paradigme d'ingénierie ciblée du génome par les méthodes dérivées initialement de l'édition génomique par le système CRISPR-Cas9 (*Clustered Regulatory Interspaced Palindromic Repeats - CRISPR associated 9*).

c) Caractérisation biologique fonctionnelle des VSI des gènes *BRCA1/2* par édition génomique CRISPR-Cas9

Cette rupture technologique CRISPR-Cas9 développée par le Dr Charpentier et le Dr Doudna (Prix Nobel de Chimie 2020) s'impose aujourd'hui comme la stratégie privilégiée d'édition du génome étant donné sa versatilité, sa précision satisfaisante, sa simplicité relative et son faible coût.

L'approche CRISPR-Cas9 utilise l'endonucléase Cas9 qui réalise une coupure double brin de l'ADN au niveau d'une séquence précise, reconnue par l'intermédiaire d'un ARN guide (ARNg). Cet outil d'origine procaryote permet une défense contre les infections virales (appelées bactériophages) en clivant l'ADN viral.³⁶⁰ Cette réaction immunitaire peut être divisée en trois phases. Tout d'abord la phase d'acquisition lors de laquelle un fragment viral est capturé et intégré comme « *spacer* » entre les loci palindromiques : il servira alors de mémoire génétique en cas de nouvelle infection par ce même virus. Ensuite vient la phase d'expression du pre-crRNA (*CRISPR RNA*) puis crRNA, homologue au génome du bactériophage, lors de la nouvelle infection. Enfin, l'étape d'interférence correspond à l'hydrolyse endonucléasique de la cible par le complexe ribonucléoprotéique crRNA-Cas9 dégradant l'ADN viral. La librairie de fragments viraux ainsi constituée joue alors le rôle de mémoire immunitaire pouvant être transmise à la descendance, CRISPR-Cas9 fait donc partie de l'immunité innée bactérienne. (Figure 25)³⁶¹

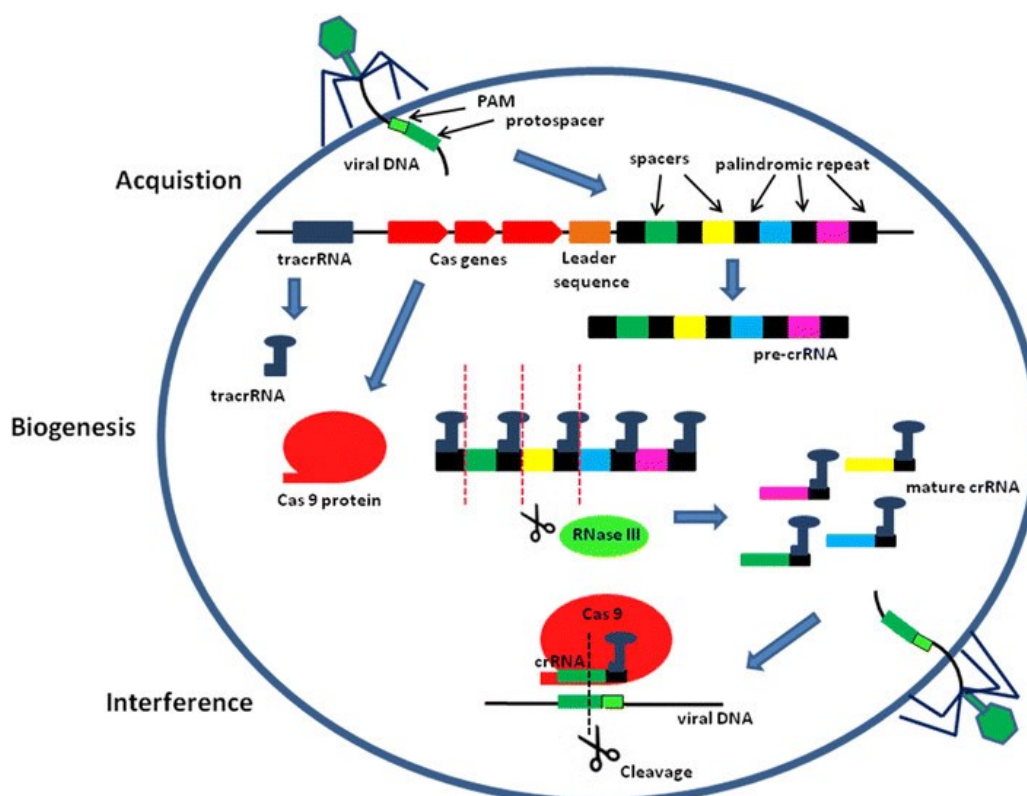


Figure 25 : Les trois étapes du système immunitaire adaptatif bactérien CRISPR-Cas : acquisition, biogénèse du crRNA et interférence de l'ADN viral³⁶¹

Après sa découverte, ce mécanisme a été adapté afin de servir d'outil d'édition génomique dans divers types cellulaires. Pour cela, deux éléments sont indispensables : l'endonucléase Cas9 et l'ARNg. Cet ARNg est constitué de deux parties : le crRNA complémentaire de la cible et le tracrRNA (*transactivating CRISPR RNA*) qui associe l'ARNg à la Cas9. Autre élément important, le motif PAM (*protospacer adjacent motif*) est situé directement en aval de la séquence ciblée par l'ARNg, la coupure double est réalisée 3 nucléotides en amont de ce PAM. Le motif NGG de ce PAM est caractéristique de la Cas9 mais peut varier en fonction des espèces procaryotes dont proviennent les Cas. Les Cas existantes sont nombreuses, elles sont liées à l'évolution des virus et dépendent également des espèces bactériennes dont elles proviennent.^{362,363} Actuellement si la Cas9 est la plus utilisée, d'autres Cas comme Cpf1³⁶⁴ ou Cas12b³⁶⁵, présentant des caractéristiques et spécificités différentes, seraient également utilisables. L'édition génomique CRISPR-Cas9 offre des applications multiples dont l'interprétation des VSI sur la réalisation de « *knock-in* » avec insertion de mutations précises dans un modèle cellulaire.

Dès 2014, l'équipe du Dr Shendure a mis à profit cette technologie pour étudier l'impact, au niveau des transcrits, de toutes les variations possibles localisées dans l'exon 18 du gène *BRCA1* de cellules humaines haploïdes. Couplée à l'utilisation d'une librairie complexe de matrices porteuses des mutations, cette approche de saturation par édition du génome (*SGE, Saturation Genome Editing*) a permis d'analyser simultanément l'effet d'un grand nombre de variations introduites dans le génome des cellules transfectées. Ces travaux ont permis d'identifier de très nombreuses variations splicéogéniques responsables de la modification de régions de régulation de l'épissage.³⁶⁶ De façon tout à fait remarquable, après cette preuve de concept, cette approche a été utilisée pour l'analyse de près de 4 000 substitutions ponctuelles localisées dans 13 exons fonctionnellement essentiels (exons 2 à 5 et 15 à 23) et aux régions introniques flanquantes de ces exons dans ce même gène *BRCA1*. Ainsi, cette approche permet de relier chacune des multiples variations testées dans une seule et unique expérience à un phénotype moléculaire (abondance des transcrits) ou cellulaire (survie) sous la forme d'un score de fonctionnalité individuel. En effet, l'impact de chacune des milliers de variations testées a été évalué sur la base de l'analyse par RNAseq de l'abondance des transcrits mais également sur la capacité à produire une protéine *BRCA1* fonctionnelle, indispensable à la survie des cellules utilisées dans cet essai. Les données générées dans le cadre de cette étude ont été en partie validées sur la base de la concordance observée ($\geq 96\%$) avec la classification établie pour une série de variations pathogènes et neutres. (Figure 26)³⁶⁷

La mise en place de l'édition génomique comme outil de caractérisation fonctionnelle des VSI de *BRCA1* et *BRCA2* est un des enjeux de cette thèse, et ce avec la contrainte de respecter les délais de prise en charge du patient.

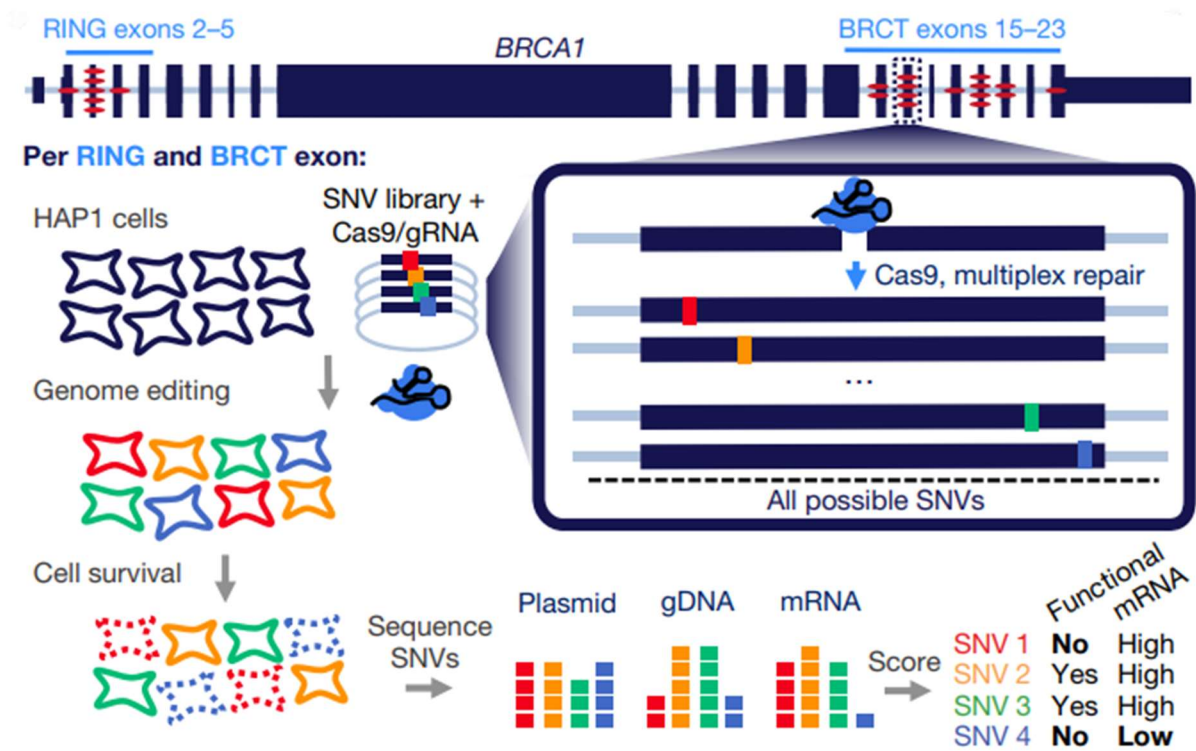


Figure 26 : Schéma expérimental de SGE introduisant tous les SNV dans les domaines RING et BRCT de BRCA1. D'après Findlay et al., 2018³⁶⁷

OBJECTIFS DE LA THESE

Lors des deux dernières décennies, le développement et la profusion des thérapies ciblées en cancérologie a transformé le concept de médecine de précision ou personnalisée en une réalité. Désormais, les plateformes de génétique moléculaire des cancers réparties en France effectuent des dizaines de milliers de tests chaque année pour caractériser morphologiquement et moléculairement le plus finement possible chaque tumeur afin de les traiter de la façon la plus spécifique possible. La quantité de données nécessaires ne fait que croître de manière à préciser de plus en plus le paysage de la tumeur et de son environnement. L'obtention, le délai de production de données et leur interprétation sont ainsi un enjeu capital pour la prise en charge optimale des patients par les innovations pharmacologiques disponibles et en forte croissance.

Le premier volet de ce travail repose sur les enjeux technologiques de la détection des mutations du gène *EGFR* pour permettre une prise en charge optimisée des patients éligibles aux thérapies ciblées dans les CBNPC au diagnostic et lors des récidives.

Le deuxième volet de cette thèse repose sur l'interprétation des altérations génétiques de la tumeur et la validation biologique fonctionnelle, pour différencier les variants à impact fonctionnel fort des variants sans effet biologique et ainsi reclassifier les VSI. Cet aspect est fondamental en ce sens qu'il peut déterminer le choix thérapeutique du clinicien envers son patient et ne se limite pas aux différents prédicteurs. De ce fait, une attention particulière a été apportée sur la réalisation de ces analyses dans un temps compatible avec la prise en charge du patient concerné par la présence d'un VSI. Ce travail a ainsi été réalisé sur les variants des gènes suppresseurs de tumeur *BRCA1* et *BRCA2* impliqués dans la réparation de l'ADN par RH, les altérations pathogènes de ces gènes donnant accès aux thérapies ciblées par iPARP. Deux approches ont été utilisées : d'une part la caractérisation des défauts d'épissage à partir d'ARN tumoral du patient par RT-PCR et clonage rapide des produits de RT-PCR, et d'autre part l'évaluation rapide *in vitro* des VSI par édition du génome en utilisant la technologie CRISPR-Cas.

RESULTATS ET DISCUSSION

1. DETECTION DES MUTATIONS *EGFR* DANS LES CBNPC

Dès la première ligne de traitement, la détection des mutations activatrices du domaine tyrosine kinase de l'*EGFR* est un enjeu majeur dans la prise en charge des CBNPC métastatiques. En effet, cette analyse conditionne l'accès aux ITK-*EGFR*, thérapie ciblée orale ayant montré un important bénéfice pour les patients en terme de survie sans progression et de survie globale. Lors de la progression, des altérations de résistance *on target* mais aussi *off target* doivent être recherchées sur tissu tumoral ou dans l'ADNtc, dans le but d'adapter la thérapeutique le plus rapidement possible. Dans les laboratoires de génétique tumorale, cette caractérisation moléculaire des CBNPC peut être restreinte aux mutations bien documentées de l'*EGFR*, ou plus large par NGS pour séquencer exhaustivement les exons concernés de l'*EGFR* mais aussi d'autres cibles d'intérêt théranostique dont le panel s'élargit continuellement au cours de ces dernières années.

Nous avons réalisé une étude prospective comparant la caractérisation de l'*EGFR* sur une technique ciblée et rapide (automate Idylla - Biocartis) ou plus étendue par NGS (Ion Torrent). Cent échantillons tumoraux FFPE ont été testés parallèlement par les deux techniques et la comparaison a porté sur les aspects techniques et pratiques au laboratoire.

ARTICLE 1 : CARACTERISATION MOLECULAIRE DE L'EGFR DANS LES CANCERS BRONCHIQUES NON A PETITES CELLULES : ETUDE PROSPECTIVE COMPARATIVE DES TECHNOLOGIES NGS ET AUTOMATE IDYLLA

Louise-Marie Chevalier, Amandine Billaud, Christophe Passot, Adélaïde Renoult, Frédéric Bigot, Véronique Verrière, Alain Morel

Publié le 18 février 2020 – Annales de Pathologie



Disponible en ligne sur
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com



ARTICLE ORIGINAL

Caractérisation moléculaire de l'EGFR dans les cancers bronchiques non à petites cellules : étude prospective comparative des technologies NGS et automate Idylla



EGFR molecular characterization in non-small cell bronchic cancer: comparative prospective study by NGS and Idylla platform technologies

Louise-Marie Chevalier^{a,*,b}, Amandine Billaud^{a,b},
Christophe Passot^a, Adélaïde Renoult^a,
Frédéric Bigot^a, Véronique Verrièle^a, Alain Morel^{a,b}

^a Département de Biopathologie, Institut de Cancérologie de l'Ouest, site Paul Papin, 15, rue André Boquel, 49 055 Angers cedex 02, France

^b CRCINA U1232, Inserm, Université de Nantes, Université d'Angers, 4, rue Larrey, Angers, France

Accepté pour publication le 23 janvier 2020
Disponible sur Internet le 17 février 2020

MOTS CLÉS

EGFR ;
Cancer bronchique non à petites cellules ;
Séquençage de nouvelle génération ;
PCR temps réel ;
Comparaison technique

Résumé

Introduction. – La détection des altérations génétiques du domaine tyrosine kinase de l'EGFR est un enjeu majeur dans la prise en charge des cancers bronchiques non à petites cellules, car elle conditionne l'accès aux inhibiteurs de tyrosine kinase. En pratique, il est possible de faire une caractérisation restreinte aux mutations bien documentées ou de séquencer tous les exons concernés de l'EGFR et aussi d'autres cibles d'intérêt théranostique. Cette étude prospective compare la caractérisation ciblée de l'EGFR sur automate Idylla (Biocartis) ou plus étendue par séquençage nouvelle génération par la technologie Ion Torrent.

Matériel et méthodes. – Cent échantillons tumoraux fixés au formol et inclus en paraffine ont été testés simultanément par les deux techniques dans les conditions recommandées par les fournisseurs. La comparaison a porté sur l'ensemble des aspects techniques et pratiques au laboratoire.

Résultats. – Au moins une mutation EGFR d'intérêt pour les inhibiteurs de tyrosine kinase a été détectée par séquençage pour 9 échantillons et 7 par le système Idylla. Pour trois échantillons, des mutations EGFR de sensibilité aux inhibiteurs de tyrosines kinases ont été détectées uniquement par séquençage nouvelle génération. De plus, pour 37 échantillons, des mutations d'intérêt hors de l'EGFR ont été rendues par séquençage au prescripteur.

* Auteur correspondant.

Adresses e-mail : louisemarie.chevalier@ico.unicancer.fr (L.-M. Chevalier), amandine.billaud@univ-angers.fr (A. Billaud), christophe.passot@ico.unicancer.fr (C. Passot), adelaide.renoult@ico.unicancer.fr (A. Renoult), frederic.bigot@ico.unicancer.fr (F. Bigot), veronique.verriele@ico.unicancer.fr (V. Verrièle), alain.morel@ico.unicancer.fr (A. Morel).

<https://doi.org/10.1016/j.annpat.2020.01.006>
0242-6498/© 2020 Elsevier Masson SAS. Tous droits réservés.

Conclusion. – La technologie Idylla permet la caractérisation rapide d'une majorité de variants de l'*EGFR* et peut être optimisée en exploitant les courbes d'amplification avec l'outil Idylla Explore ou en augmentant la quantité de matériel initial. Une analyse complémentaire par séquençage nouvelle génération pour les résultats non contributifs par Idylla doit être préconisée.

© 2020 Elsevier Masson SAS. Tous droits réservés.

KEYWORDS

EGFR;
Non-small cell lung cancer;
Next-generation sequencing;
Real-time PCR;
Technical comparison

Summary

Introduction. – Detection of genetic alterations in the *EGFR* tyrosine kinase domain is a major concern in the management of non-small cell lung cancer because it conditions access to tyrosine kinase inhibitors. In practice, it is possible to characterize only well-documented mutations or to sequence all relevant *EGFR* exons and also other targets of therapeutic interest. This prospective study compares the targeted *EGFR* characterization on Idylla platform (Biocartis) and a more extensive one by next generation sequencing using Ion Torrent technology.

Material and methods. – A total of 100 formalin-fixed paraffin-embedded tumour samples were tested simultaneously by both techniques under the conditions recommended by the suppliers. The comparison covered all technical and practical aspects of the laboratory.

Results. – At least one *EGFR* mutation of interest for tyrosine kinase inhibitors for 9 and 7 samples was detected respectively by sequencing and by the Idylla system. For three samples, *EGFR* sensitive mutations to tyrosine kinase inhibitors were detected only by next-generation sequencing. In addition, for 37 samples, mutations of clinical interest outside *EGFR* were characterized by sequencing and communicated to the prescriber.

Conclusion. – Idylla technology allows the rapid characterization of a majority of *EGFR* variants. The result can be optimized by careful analysis of the amplification curves with the Idylla Explore tool or by increasing the amount of initial material. A complementary new generation sequencing analysis for non-contributory results by Idylla should also be recommended.

© 2020 Elsevier Masson SAS. All rights reserved.

Introduction

Avec 49 109 nouveaux cas estimés en France en 2017, le cancer du poumon se situe au 4^e rang des cancers incidents, tous sexes confondus, et au premier rang des décès par cancer [1].

Depuis les années 2000, la prise en charge du cancer bronchique non à petites cellules (CBNPC) a connu d'importantes avancées thérapeutiques. La prise en charge des stades localisés (<IIIb) repose sur la chirurgie complétée par une chimiothérapie adjuvante en fonction des critères pronostiques. Pour les stades localement avancés non résecables, le traitement de référence est la radiochimiothérapie qui peut être suivie par un entretien par immunothérapie pendant un an (Durvalumab) pour les carcinomes avec expression de PD-L1 ≥ 1 %. Pour les stades métastatiques (IV), la prise en charge repose sur les traitements systémiques seul ou en association (thérapies ciblées, immunothérapie, chimiothérapie). Le développement de nouveaux traitements ciblés sur les voies oncogéniques activées a permis l'accès à un nouvel arsenal thérapeutique efficace, globalement mieux tolérés et offrant aux patients une meilleure qualité de vie. Mieux définir l'hétérogénéité génétique intratumorale et préciser la diversité des oncogènes activés dans les CBNPC est, depuis, un préalable incontournable au succès des thérapies ciblées. Le principal oncogène ciblé à ce jour dans les CBNPC est le récepteur au facteur de croissance épidermique (*EGFR*), muté sur son domaine tyrosine kinase, qui est à l'origine de l'usage des inhibiteurs de tyrosine kinase de l'*EGFR* (ITK-*EGFR*).

L'*EGFR* qui appartient à la famille des récepteurs tyrosine kinase est codé par le gène *EGFR*, situé au niveau du chromosome 7. L'*EGFR* joue un rôle important dans la physiologie normale, mais aussi en carcinogenèse en favorisant la prolifération, la survie et l'invasion [2]. Dans la population caucasienne, la fréquence des mutations ciblées par les ITK-*EGFR* est rapportée entre 10 et 15 % des CBNPC. Dans la littérature, il est bien connu que cette proportion varie en fonction du sous-type de CBNPC, du genre, du statut fumeur ou non fumeur des patients ainsi que de l'origine ethnique de la population [3,4]. L'utilisation des ITK-*EGFR* en tant que traitement de première ligne dans les CBNPC présentant des altérations génétiques oncogéniques de l'*EGFR* a permis une augmentation significative du taux de survie sans progression, par comparaison avec les traitements de chimiothérapies conventionnelles [5].

La détection spécifique, sensible et rapide des mutations de sensibilité ou de résistance de l'*EGFR* aux ITK-*EGFR* est un défi quotidien des laboratoires de génétique moléculaire des cancers. Pour répondre au mieux aux prescriptions, deux approches techniques existent : les techniques dites « ciblées » qui caractérisent généralement les mutations bien documentées de l'*EGFR* et le séquençage haut débit ou séquençage de nouvelle génération (NGS) qui peut explorer l'ensemble des exons d'intérêt (18 à 21) de l'*EGFR*. De plus, le séquençage par NGS permet de caractériser d'autres oncogènes ou gènes d'intérêt tels que *KRAS*, *HER2*, *BRAF*, *MET* par le séquençage d'un panel d'exons de gènes recommandé par l'INCa pour les CBNPC.

Le but de l'étude a été d'évaluer, de manière prospective, la caractérisation des variants de l'EGFR réalisée par deux technologies : la PCR en temps réel sur une plateforme Idylla (Biocartis) et le séquençage par NGS par la technologie Ion Torrent sur un séquenceur Ion S5 (Thermo Fisher Scientific). La comparaison a porté aussi bien sur des aspects proprement techniques que sur des aspects pratiques au laboratoire (délais de rendu, mise en œuvre).

Ion 530 Kit Chef (ThermoFisher Scientific). Les données ont été analysées par Torrent Suite 5.4 et pipeline bioinformatique développé et utilisé au laboratoire (PICO-NGS v1.0). Les attentes analytiques du séquençage par NGS étaient les suivantes : profondeur de séquençage minimum 300X, couverture 100 %, sensibilité analytique 2,5 %, génome de référence hg19.

Matériels et méthodes

Type d'étude

L'étude est une évaluation prospective, diagnostique et monocentrique comparant deux technologies dans le cadre de prescriptions de statut mutationnel de l'EGFR faites au laboratoire de génétique moléculaire de l'Institut de Cancérologie de l'Ouest (Tableau 1).

Échantillons

De manière systématique et prospective, d'octobre 2018 à février 2019, 100 échantillons tumoraux de CBNPC fixés au formol et inclus en paraffine (FFPE pour Formalin-Fixed Paraffin-Embedded) validés par un pathologiste ont été transmis au laboratoire de biologie moléculaire sous forme de sections de 10 µm (lames ou copeaux) dont une section était utilisée pour la technique Biocartis. La quantité de matériel devait permettre simultanément le séquençage par NGS (technique de référence au laboratoire) et une analyse sur automate Idylla selon les spécifications du fournisseur (Tableau 1).

Tout type histologique de CBNPC était éligible à l'étude. Les tumeurs avec une cellularité tumorale < 10 % et/ou fixées au Bouin et/ou dont la quantité de matériel était insuffisante pour réaliser les deux techniques ont été exclues (Tableau 1). Les cas exclus ne font pas partie des 100 échantillons étudiés.

Technologie NGS

L'ADN génomique (ADNg) somatique a été extrait avec le kit NucleoSpin DNA FFPE XS (Macherey Nagel) puis repris dans 100 µL de tampon d'élution. La quantité d'ADNg somatique recommandée pour chaque échantillon était de 20 ng ou un volume maximum de 7,5 µL. La recherche de mutations des gènes par NGS a été réalisée avec un panel de 21 gènes comprenant les exons séquencés suivants : *AKT1* (NM_005163) exon 3, *ALK* (NM_004304) exons 23 à 25, *BRAF* (NM_004333) exons 11 et 15, *CTNNB1* (NM_001904) exon 3, *EGFR* (NM_005228) exons 18 à 21, *ERBB2* (NM_004448) exons 19 et 20, *ERBB4* (NM_005235) exons 10 et 12, *ESR1* (NM_000125) exon 8, *FGFR2* (NM_000141) exons 7, 12 et 14, *FGFR3* (NM_000142) exons 7, 9 et 14, *HRAS* (NM_005343) exons 2 à 4, *KIT* (NM_000222) exons 8, 9, 11, 13, 17 et 18, *KRAS* (NM_033360) exons 2 à 4, *MAP2K1* (NM_002755) exon 2, *MET* (NM_001127500) exons 2, 14 et 16 à 20, *NRAS* (NM_002524) exons 2 à 4, *PDGFRA* (NM_006206) exons 12, 14 et 18, *PIK3CA* (NM_006218) exons 10 et 21, *POLE* (NM_006231) exons 9 à 14, *STK11* (NM_000455) exons 1 à 9, *TP53* (NM_000546) exons 2 à 11. Les librairies ont été préparées par PCR multiplex (Kit Ion AmpliSeq Library 2.0 – Custom Panel, ThermoFisher Scientific), amplifiées par PCR en émulsion sur IonChef System et séquencées par NGS sur Ion S5 System avec le kit Ion 510 & Ion 520 &

Technologie Idylla (Biocartis)

Le système Idylla est une solution de diagnostic moléculaire basée sur une PCR temps réel spécifique d'allèles entièrement automatisée. Elle associe dans une cartouche à usage unique l'extraction des échantillons et la PCR afin de détecter les cibles moléculaires à partir d'une variété d'échantillons solides ou liquides et notamment des sections FFPE. Dans cette étude, pour chaque échantillon, une section FFPE de 10 µm (lame ou copeau) a été placée directement dans la cartouche Biocartis du système Idylla, sans prétraitement, selon les recommandations du fournisseur. La cartouche a ensuite été insérée dans le système Idylla, où la technologie des ultrasons focalisés de haute intensité, les tampons, les réactifs et la chaleur induisent un déparaffinage, une rupture des tissus, une lyse cellulaire et une libération ultérieure d'ADN. Des amorces spécifiques d'allèles permettent une amplification et une détection en temps réel des mutations par des sondes d'hybridations fluorescentes selon la technologie PlexZyme et PlexPrime [6]. La cartouche testée était l'Idylla EGFR Mutation Test, contenant tous les consommables nécessaires. La console Idylla et les cartouches portent le marquage CE-IVD. Le test de mutation Idylla EGFR permet d'amplifier et de détecter 51 mutations dans les exons 18 à 21 de l'EGFR (Tableau 2). Un ensemble de paramètres en rapport avec les courbes d'amplification générées est déterminé par le logiciel Idylla, notamment la valeur ΔCq (différence entre la valeur de cycle de quantification (Cq) du signal de contrôle interne du gène EGFR (Cq contrôle EGFR) et le Cq du signal du variant). Un échantillon est classé comme muté si les paramètres de la courbe d'amplification générée se situent dans la plage validée soit avec un Cq contrôle EGFR < 26. Le fournisseur indique des limites de sensibilité par mutation ou groupe de mutations en fonction du Cq contrôle EGFR. En cas de Cq contrôle EGFR > 26, le fournisseur recommande d'estimer l'échantillon comme non interprétable. Certains échantillons sont rendus « invalides » par le logiciel Idylla, le statut EGFR de ces échantillons est alors non déterminé. Les résultats obtenus lors du test Idylla EGFR Mutation Test au cours de cette étude n'ont pas été utilisés à des fins cliniques ou diagnostiques et n'ont pas été mis à la disposition des patients ou de leurs cliniciens. Le fournisseur Biocartis donne accès pour tous les échantillons aux courbes d'amplification générées dans chaque chambre de PCR avec l'outil informatique Idylla Explore (non marqué CE-IVD). Dans chaque chambre est amplifié un groupe de mutations EGFR spécifiques et un contrôle de traitement d'échantillon correspondant à la quantification de l'EGFR total. Les courbes générées correspondent à la quantité d'ADN amplifié en unité de fluorescence arbitraire en fonction du nombre de cycles d'amplification.

Étude des mutations EGFR discordantes

Pour les échantillons présentant une ou plusieurs mutations de l'EGFR par NGS et non retrouvées par le système Idylla dans les conditions recommandées par le fournisseur,

Tableau 1 Caractéristiques de l'étude.
Characteristics of the study.

Nom de l'étude	Recherche de mutations du gène <i>EGFR</i> chez les patients atteints d'un CBNPC : comparaison des technologies Idylla et NGS Ion Torrent
Type d'étude	Évaluation prospective, diagnostique, monocentrique
Population	Patients atteints d'un CBNPC pour lesquels une demande de statut mutationnel <i>EGFR</i> est faite à la plateforme de génétique moléculaire de l'Institut de Cancérologie de l'Ouest
Nombre de patients	100 patients
Objectifs	Tester la concordance des résultats de la recherche des mutations du gène <i>EGFR</i> entre l'automate Idylla et l'analyse par NGS d'un panel personnalisé de gènes Comparer la sensibilité des deux techniques Comparer le délai de réponse entre les 2 techniques
Matériel	Tissu FFPE : lames ou copeaux
Critère d'inclusion	Patient atteint d'un cancer bronchique non à petites cellules (tous types histologiques)
Critères de non-inclusion	Cellularité tumorale < 10 % Fixation au Bouin Quantité de matériel insuffisante pour réaliser simultanément le séquençage par NGS et l'analyse sur automate Idylla selon les spécifications du fournisseur

une deuxième cartouche était réalisée. Afin d'augmenter la quantité d'ADN, cette deuxième analyse a été réalisée en augmentant le matériel FFPE ou à partir d'ADNg somatique initialement extrait pour l'analyse NGS.

Résultats

Données anatomopathologiques

Parmi les 100 échantillons de CBNPC étudiés, les types histologiques précisés par le laboratoire d'anatomopathologie étaient les suivants : 78 adénocarcinomes, 7 carcinomes/adénocarcinomes non classés, 7 carcinomes épidermoïdes, 7 carcinomes indifférenciés, et un carcinome neuroendocrine à grandes cellules.

Concernant la nature des prélèvements, 87 échantillons étaient des biopsies, 4 étaient des pièces de résection et 9 étaient des cytoblocs.

Les cellularités tumorales variaient entre 10 % et 80 % avec une moyenne de 38,95 %, un écart-type de 22,21 % et une médiane à 40 %. Quarante-six échantillons présentaient une cellularité tumorale inférieure ou égale à 30 % (Fig. 1). Six échantillons FFPE ont été obtenus sous la forme de sections de 10 µm étalées sur lame et 94 sous forme de copeaux de même épaisseur.

Résultats des caractérisations moléculaires de l'*EGFR*

Les résultats de l'analyse moléculaire des échantillons FFPE par les deux techniques sont synthétisés dans le Tableau 3. Neuf échantillons ont été détectés avec au moins une mutation de l'*EGFR* par NGS et 7 par le système Idylla. L'ensemble des variants *EGFR* retrouvés avec les deux techniques sont rapportés dans le Tableau 4. Un variant de signification inconnue p.(Gly729Glu) dans l'exon 18 du gène *EGFR* a été détecté par NGS. Ce variant d'intérêt théranostique non établi ne fait pas partie des mutations de l'*EGFR* recherchées par le système Idylla. Quatre-vingt-trois échantillons ont été rendus « non mutés *EGFR* » par l'analyse NGS. L'automate Idylla indique 83 échantillons avec aucune mutation détectée sur l'ensemble des cibles recherchées de l'*EGFR*. Sept échantillons sont non interprétables ou non contributifs par NGS soit en raison de la dégradation importante de l'ADN,

soit liés à la faible quantité d'ADN initiale. Dix échantillons sont non contributifs avec le système Idylla : 7 résultats sont « invalides » et 3 échantillons présentent une valeur de Cq contrôle *EGFR* supérieure à 26 considérée par le fournisseur comme la limite de sensibilité. Trois échantillons sont non contributifs avec l'une et l'autre des techniques (Fig. 2).

Analyse des résultats discordants entre les deux techniques d'analyse de l'*EGFR*

Trois échantillons se sont révélés discordants (Tableau 4), pour chaque cas les conditions matérielles requises par le fournisseur pour la bonne exécution du test ont été respectées.

Patient 1

Le prélèvement était une biopsie d'adénocarcinome pulmonaire TTF1+ d'une femme de 66 ans, avec une cellularité tumorale évaluée à 20 % (plus de 100 cellules analysables). L'analyse par NGS met en évidence deux mutations *EGFR* : un variant c.2154_2155GG > TT, p.(Gly719Cys) dans l'exon 18 et un variant c.2303G > T, p.(Ser768Ile) dans l'exon 20, avec des fréquences alléliques respectives de 18,4 % et 14,7 %. L'analyse sur Idylla ne détecte aucune mutation (Cq contrôle *EGFR*=24). Cependant, l'outil Idylla Explore montre des courbes d'amplification avec un nombre de cycles supérieur au seuil de 26 pour les cibles G719C et S768I (Fig. 3E). Une deuxième analyse Idylla réalisée avec 65 ng d'ADNg utilisé pour l'analyse par NGS permet au système Idylla de détecter ces deux mutations.

Patient 2

Les copeaux d'échantillon FFPE reçus au laboratoire étaient issus d'une biopsie d'adénocarcinome pulmonaire TTF1+ chez une femme de 51 ans. La cellularité tumorale était de 10 % (avec moins de 100 cellules analysables). L'analyse par NGS a détecté une délétion complexe c.2237_2255delinsT, p.(Glu646_Ser752delinsVal) en phase dans l'exon 19 du gène *EGFR* avec une fréquence allélique de 1,6 %. La présence de cette délétion a été ensuite confirmée par analyse de fragment.

L'analyse sur Idylla avec un copeau puis, dans un second temps, avec 115 ng dans 50 µL d'ADNg somatique extrait

Tableau 2 Mutations détectées par le test Idylla EGFR Mutation Test.
Mutations detected by the Idylla EGFR Mutation Test.

Oncogène	Exon	Résultat rendu par Idylla	Mutation
EGFR	18	G719A/C/S	c.2156G>C
			c.2155G>T
	19	Exon 19 deletions	c.2155G>A
			c.2154_2155delinsTT
			c.2238_2248delinsGC
			c.2239_2248delinsC
			c.2240_2248del
			c.2239_2247del
			c.2239_2251delinsC
			c.2240_2251del
			c.2235_2249del
			c.2236_2250del
20	T790M S768I Exon 20 insertions	c.2239_2253del	
		c.2240_2254del	
		c.2238_2252del	
		c.2237_2251del	
		c.2235_2252delinsAAT	
		c.2237_2252delinsT	
		c.2234_2248del	
		c.2236_2253delinsCTA	
		c.2237_2253delinsTA	
		c.2235_2251delinsAG	
		c.2236_2253delinsCAA	
		c.2230_2249delinsGTCAA	
21	L858R L861Q	c.2240_2257del	
		c.2237_2255delinsT	
		c.2239_2256del	
		c.2236_2253del	
		c.2239_2258delinsCA	
		c.2237_2254del	
		c.2238_2255del	
		c.2237_2257delinsTCT	
		c.2236_2255delinsAT	
		c.2236_2256delinsATC	
		c.2237_2256delinsTT	
		c.2237_2256delinsTC	
c.2235_2255delinsGGT			
20	T790M S768I Exon 20 insertions	c.2238_2258del	
		c.2236_2256del	
		c.2253_2276del	
		c.2369C > T	
		c.2303G > T	
		c.2310_2311insGGT	
		c.2307_2308insGCCAGCGTG	
		c.2309_2310delinsCCAGCGTGGAT	
		c.2311_2312insGCGTGGACA	
		c.2319_2320insCAC	
		c.2573T > G	
		c.2573_2574delinsGT	
21	L858R L861Q	c.2573_2574delinsGA	
		c.2582T > A	

pour l'analyse NGS ne détecte aucune mutation. Dans ces deux analyses sur le système Idylla, les valeurs de Cq contrôles EGFR sont respectivement de 24,6 et 23,6, respectant la limite de 26 du fabricant. L'apparition de produits d'amplification mutation spécifique avant le contrôle EGFR total semble indiquer une amplification non spécifique importante dans certaines chambres (Fig. 4A–E). Ce bruit

de fond important ne permet pas à l'algorithme du logiciel Idylla d'interpréter les résultats.

Patient 3

Le prélèvement concerne une biopsie d'adénocarcinome pulmonaire chez un homme de 85 ans. La cellularité

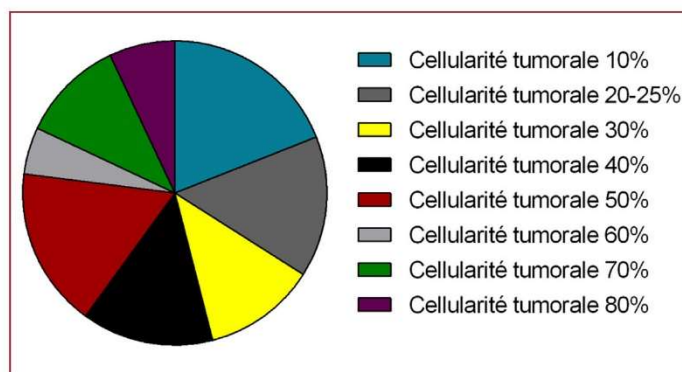


Figure 1. Répartition des cellularités tumorales des échantillons.
Distribution of samples tumor cellularities.

Résultats	NGS Ion Torrent	Idylla
Mutés <i>EGFR</i>	9	7
Non mutés <i>EGFR</i>	83	83
Variante de signification inconnue <i>EGFR</i>	1	0
Non interprétables/non contributifs	7	10 dont : 7 échantillons Cq contrôle <i>EGFR</i> > 26 3 « invalides »

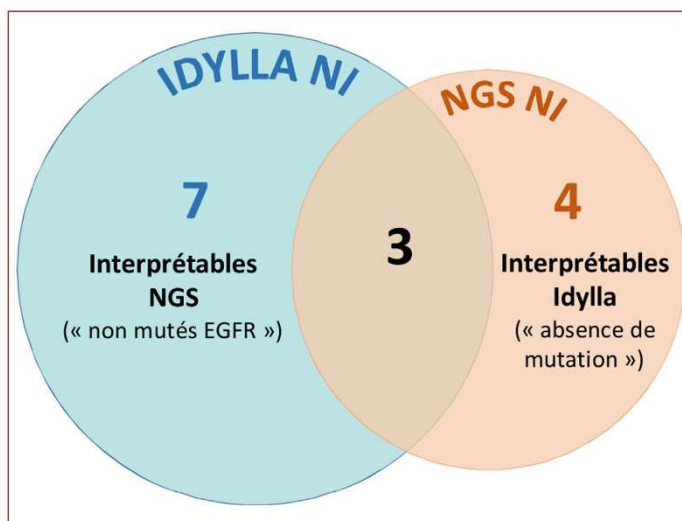


Figure 2. Échantillons non interprétables (NI).
Uninterpretable (NI) samples.

tumorale des copeaux était estimée à 20 % (avec plus de 100 cellules analysables). L'analyse par NGS a mis en évidence deux mutations *EGFR* : un variant c.2573T>G, p.(Leu858Arg) dans l'exon 21 à une fréquence allélique de 34,8 % et un variant c.2369C>T, p.(Thr790Met) dans l'exon

20 à une fréquence allélique 9,9 %, proche d'un polymorphisme c.2361G>A, p.(Gln787Gln) (Fig. 5). L'analyse sur l'automate Idylla à partir d'une section de 10 µm ou à partir de 50 ng dans 50 µL d'ADNg somatique extrait pour l'analyse NGS détecte uniquement la mutation *EGFR* p.(Leu858Arg)

Tableau 4 Mutations EGFR caractérisées par chaque technique (discordances : patients 1, 2, 3).
EGFR mutations characterized by each technique (discordances : patients 1, 2, 3).

Patient	% Cellules tumorales	NGS		Exon	Fréquence allélique %	Idylla	
		Mutation EGFR (nomenclature HGVS)	Mutation EGFR présente sur le compte rendu Idylla			Mutation EGFR présente sur le compte rendu Idylla	Contrôle EGFR total
1	20	c.2154_2155GG>TT,	18	18,4	Aucune mutation détectée	CQ 24	
		p.(Gly719Cys) c.2303G>T,					20
2	10	p.(Ser768Ile) Déletion c.2237_2255delinsT,	19	1,6	Aucune mutation détectée	CQ 24.6	
		p.(Glu746_Ser752delinsVal) c.2573T>G, p.(Leu858Arg) c.2369C>T, p.(Thr790Met) Insertion c.2317_2319dup,					21
3	20		20	9,9	CQ 22.7		
						20	39,3
4	70		20	39,3	CQ 22.8		
						21	6,8
5	50		21	6,8	CQ 20.2		
						18	12,7
6	70		18	10,5	CQ 16.5		
						21	16,8
7	70		21	16,8	CQ 23.3		
						21	59,9
8	40		21	59,9	CQ 22.4		
						19	6,1
9	40		19	6,1	CQ 20.1		
						19	14,2
10	80		19	14,2	CQ 21.2		



Figure 3. Patient 1 : courbes d'amplification générées par le système Idylla dans l'outil Idylla Explore. Les chambres A, B, C et D ne montrent pas d'amplification spécifique. Dans la chambre E, malgré un résultat négatif rendu par l'algorithme du logiciel Idylla, les courbes d'amplification visualisées laissent suspecter la présence des cibles G719C et S768I.
Patient 1: amplification curves generated by the Idylla system in the Idylla Explore tool.

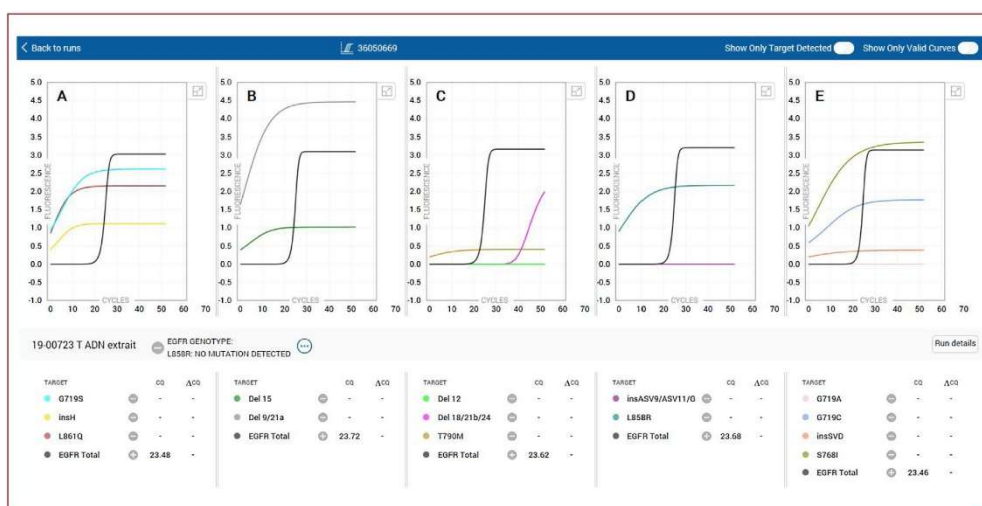


Figure 4. Patient 2 : courbes d'amplification générées par le système Idylla dans l'outil Idylla Explore. La présence d'amplification non spécifique dans les chambres A, B, D et E ne permet pas à l'algorithme du logiciel Idylla d'interpréter les résultats et d'identifier la mutation mise en évidence par NGS (délétion complexe c.2237_2255delInsT, p.(Glu646_Ser752delInsVal) en phase dans l'exon 19 du gène EGFR). On note une absence d'amplification dans la chambre C.
Patient 2: amplification curves generated in each room by the Idylla system in the Idylla Explore tool.

(Cq contrôles EGFR respectivement à 22,7 et 22,6). Les courbes de PCR dans Idylla Explore ne montrent pas de bruit de fond dans la chambre de détection de la mutation p.(Thr790Met) (Fig. 6 C), le système Idylla valide donc l'absence de cette mutation.

Autres cibles moléculaires mises en évidence par l'analyse NGS

Pour 37 échantillons sur les 100 analysés, les mutations non EGFR d'intérêt pour la prise en charge des patients

(Tableau 5) ont été rendues au prescripteur après analyse NGS avec le panel de 21 gènes, et quarante-huit échantillons ont été rendus non mutés.

Aspects pratiques : délais, ressources humaines, procédure qualité

Séquençage par NGS

En termes de délai, l'analyse par séquençage NGS d'une série d'échantillons de tumeurs thoraciques au laboratoire de biologie moléculaire de l'Institut de Cancérologie de

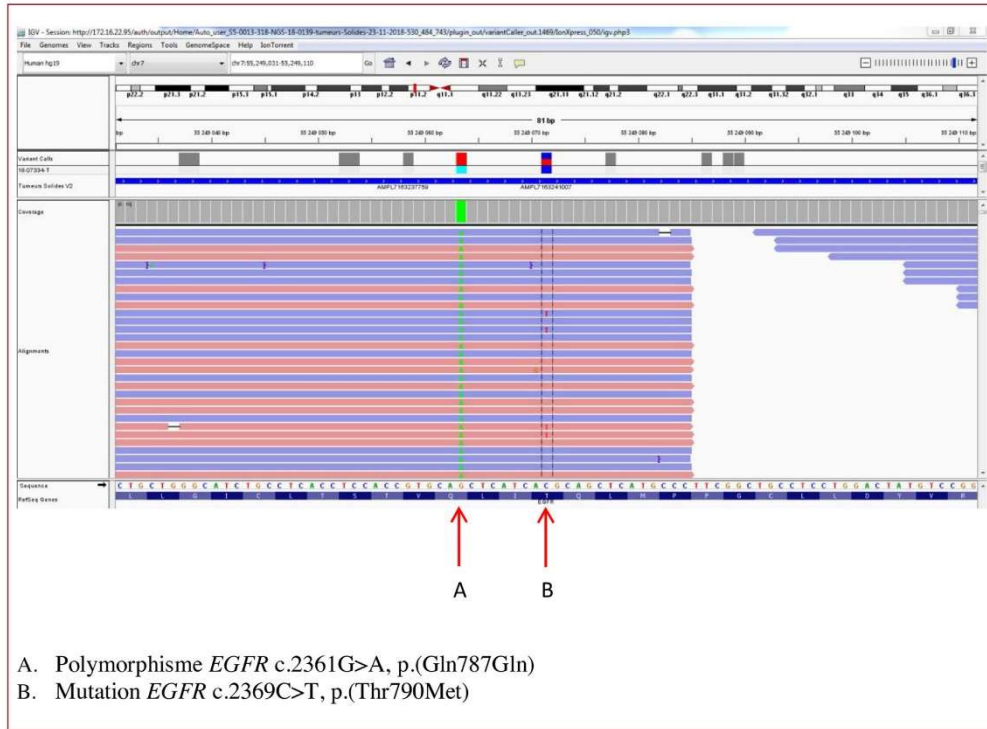


Figure 5. Visualisation avec l’outil Integrative Genomic Viewer du variant c.2369C>T, p.(Thr790Met) détecté par NGS (patient 3). Polymorphisme *EGFR* c.2361G>A, p.(Gln787Gln) ; mutation *EGFR* c.2369C>T, p.(Thr790Met). *Integrative Genomic Viewer* visualization of c.2369C>T, p.(Thr790Met) mutation detected by NGS (patient 3).

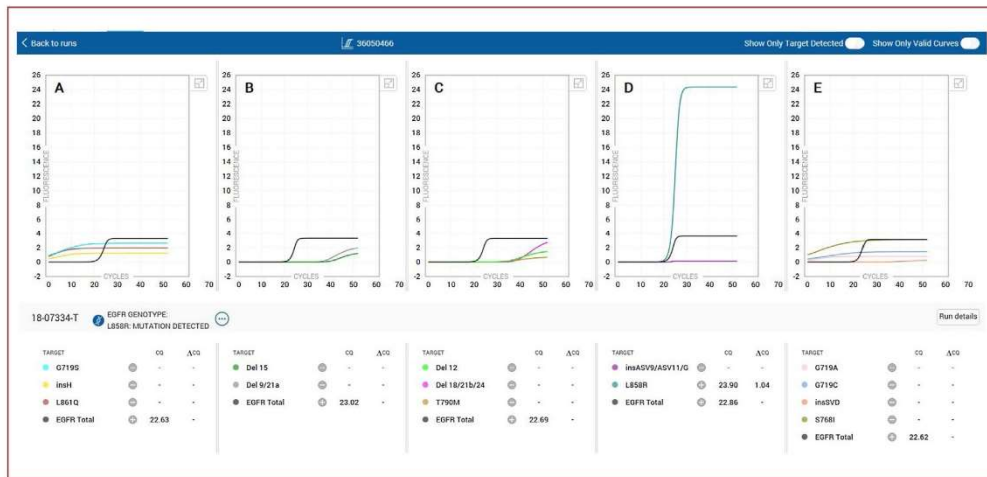


Figure 6. Patient 3 : courbes d’amplification générées par le système Idylla dans l’outil Idylla Explore. Dans la chambre D, la mutation L858R est clairement détectée. En revanche, dans la chambre C, les courbes d’amplification ne permettent pas de visualiser la mutation T790M mise en évidence par NGS. Les chambres A, B et E ne montrent pas d’amplification spécifique. *Patient 3: amplification curves generated by the Idylla system in the Idylla Explore tool.*

Tableau 5 Mutations indiquées sur le compte rendu de biologie moléculaire après analyse NGS.
Mutations on molecular biology report from NGS analysis.

Gène(s)	Nombre d'échantillons
KRAS	22
KRAS et STK11	9
MET	2
BRAF V600E	1
BRAF non V600E	1
NRAS	1
HER2	2

L'Ouest nécessite : la partie technique semi-automatisée de 3 jours réalisée par des techniciens qualifiés et habilités puis le séquençage et l'analyse bioinformatique effectués en 2 jours. L'interprétation et la validation biologique par des biologistes moléculaires ajoutent 2 jours selon le nombre de variants identifiés et le nombre d'échantillons de la série. Le délai de rendu des résultats annoncé au prescripteur de 10 jours ouvrés est en général respecté. L'analyse par NGS requiert une infrastructure adaptée (locaux pré-PCR et post-PCR, thermocycleur, séquenceur et consommables et un personnel qualifié). La démarche qualité comprenant l'habilitation-qualification du personnel, les procédures et l'analyse bioinformatique selon la norme ISO 15189 est complexe.

Plateforme Idylla

Pour l'analyse d'un échantillon sur le système Idylla : la partie technique consiste à déposer le matériel FFPE dans la cartouche à usage unique et à charger celle-ci dans le module Idylla (temps inférieur à 2 min). L'analyse par l'automate dure 150 min pour un échantillon, temps auquel il faut rajouter l'interprétation biologique qui est très variable en fonction de la complexité du cas. Cette interprétation biologique est réalisée avec la valeur de Cq contrôle *EGFR* grâce au rapport Idylla automatiquement généré. Pour plusieurs échantillons, le délai de rendu des résultats dépend du nombre de modules disponibles au laboratoire. Aucune infrastructure ou matériel particulier n'est nécessaire pour la réalisation de la technique Idylla. La démarche qualité selon la norme ISO 15189 est simple (marquage CE-IVD) et ne requiert pas de compétence particulière.

Discussion/conclusion

La preuve de l'intérêt de la caractérisation des mutations de l'*EGFR* dans les CBNPC n'est plus à faire compte tenu de l'efficacité des ITK-*EGFR* dans cette pathologie. L'enjeu pour les laboratoires réalisant cette analyse est aujourd'hui de répondre aux prescriptions de manière sensible et spécifique dans des délais compatibles avec la prise en charge des patients. Cette étude prospective inédite avait pour objectif de comparer deux technologies : le système de PCR temps réel entièrement automatisé Idylla ciblant 51 mutations de l'*EGFR* et le séquençage par NGS d'un panel de 21 gènes comprenant les exons 18 à 21 de l'*EGFR*. Les études publiées évaluant l'Idylla *EGFR* Mutation Test sont des études rétrospectives comme celle d'Ilie et al. comparant système Idylla et pyroséquençage [7]. Les spécifications des fournisseurs ont été respectées pour les 100 échantillons testés avec les deux techniques. L'intérêt particulier de ce

travail était d'avoir une évaluation concrète à la fois scientifique et pratique des techniques, dans les conditions réelles de l'activité d'un laboratoire de biologie moléculaire des cancers. L'étude a ainsi été effectuée sans biais de choix d'échantillons et avec les contraintes quotidiennes rencontrées dans cette localisation tumorale [8]. Dans quasiment la moitié des cas, les prélèvements avaient une cellularité inférieure ou égale à 30 % dont 19 avaient une cellularité tumorale estimée à 10 % par le pathologiste.

Le taux de mutés *EGFR* retrouvé dans cette étude (9 %) est en deçà du taux attendu dans la littérature [3,4]. Deux arguments expliquent ce résultat. D'une part, la fréquence des mutations de l'*EGFR* est 4,4 fois plus élevée pour les types adénocarcinomes que pour les types non adénocarcinomes [9], or parmi les 100 échantillons de cette étude, 15 échantillons étaient de type histologique non adénocarcinome dont 7 carcinomes épidermoïdes. D'autre part, le taux de mutation est dépendant de la cellularité tumorale, or 46 % des échantillons présentaient une cellularité tumorale < 30 % et donc un risque accru de résultats faussement négatifs.

Les résultats montrent une concordance partielle entre les deux techniques concernant les résultats non contributifs. De même, parmi les mutations *EGFR* retrouvées, 7 échantillons étaient concordants et 3 échantillons étaient discordants avec une mutation *EGFR* caractérisée par NGS et non retrouvées sur Idylla.

La première discordance concernait une co-mutation *EGFR* présente à des fréquences alléliques > 10 % par NGS, une seconde cartouche Idylla avec une quantité plus élevée de matériel a permis la mise en évidence des mutations, indiquant que cette discordance peut être attribuée à une quantité insuffisante de matériel insérée dans la cassette. En pratique, lorsqu'on met dans la cassette Idylla un copeau de 10 µm, selon les recommandations du fournisseur Biocartis, on ignore la surface réelle de l'échantillon. Dans ce premier cas discordant, la cellularité tumorale était de 20 %. L'augmentation de la quantité de matériel avait donc permis de passer d'un Cq contrôle *EGFR* de 24 (proche de 26, seuil à partir duquel le résultat *EGFR* ne peut être rendu) à 22,8. Dans ce cas l'analyse des courbes d'amplification par Idylla Explore a permis de suspecter la présence des deux mutations caractérisées par NGS et d'envisager une approche complémentaire (Fig. 3E).

La seconde discordance concernait un échantillon avec une cellularité tumorale estimée d'environ 10 %. L'analyse par NGS caractérisait une délétion dans l'exon 19 de l'*EGFR* avec une fréquence allélique de 1,6 %, inférieure à notre sensibilité analytique évaluée à 2,5 %. Après confirmation de sa présence par analyse de fragment, cette altération a cependant été rendue auprès du prescripteur permettant ainsi au patient d'accéder à une thérapie ciblée par un ITK-*EGFR*. Dans ce cas, l'augmentation de la quantité de matériel introduite dans la cassette Idylla n'a pas permis la mise en évidence de la mutation malgré des valeurs de Cq contrôle *EGFR* à 24,6 avec 1 copeau et 23,6 avec 50 µL d'ADNg extrait. L'analyse des courbes d'amplification par Idylla Explore montre dans les deux analyses un bruit de fond dans plusieurs chambres de PCR (Fig. 4), le résultat est donc éliminé par l'algorithme. Il ne semble donc pas ici que la faible quantité de matériel soit à l'origine du résultat négatif. Cette discordance souligne à la fois la grande sensibilité de l'analyse NGS et l'intérêt de tester tous les échantillons indépendamment de la cellularité pour donner le maximum d'opportunités thérapeutiques au patient.

L'échantillon impliqué dans la troisième et dernière discordance présentait une cellularité tumorale de 20 %, les deux techniques ont détecté la présence d'une mutation de sensibilité p.(Leu858Arg), mais seul le NGS a retrouvé une mutation p.(Thr790Met) à une fréquence allélique de 9,9 %. Cette mutation présente une grande importance en pratique clinique puisqu'elle est très couramment associée à une résistance acquise aux ITK-EGFR de première et deuxième générations [10–13] et ouvre accès à un ITK-EGFR de troisième génération, l'osimertinib [14]. Le système Idylla n'a détecté la mutation ni dans les conditions CE-IVD avec un copeau ni en présence d'une quantité d'ADNg somatique plus importante bien que dans les deux cas les Cq contrôles *EGFR* étaient très satisfaisants (respectivement 22,7 et 22,6). La sensibilité analytique annoncée par le fournisseur pour la mutation p.(Thr790Met) avec un Cq contrôle *EGFR* de 22,4 est de 3,9 % donc bien inférieure à la fréquence allélique retrouvée au NGS. Cependant, la présence du polymorphisme p.(Gln787Gln) en amont de la mutation p.(Thr790Met) pourrait expliquer l'absence de détection de la mutation p.(Thr790Met). En effet, pour une valeur de Cq contrôle *EGFR* de 22,7, la sensibilité du variant p.(Thr790Met) en présence du polymorphisme p.(Gln787Gln) a été estimée par le service Recherche et Développement Biocartis à 15 % donc très supérieure à la fréquence allélique retrouvée de 9,9 % par NGS. « L'Exome Aggregation Consortium » (ExAC) indique que ce polymorphisme p.(Gln787Gln) est présent chez 52 % de la population. Dans le cadre d'une prescription de recherche de mutation de résistance de l'*EGFR* aux ITK-EGFR de première ou deuxième génération, en cas d'absence de détection du variant p.(Thr790Met) et du fait de l'importance de cette mutation dans la prise en charge des patients, il serait conseillé d'effectuer une analyse complémentaire par une méthode insensible à la co-mutation p.(Gln787Gln).

En outre, l'analyse par NGS d'un panel de 21 gènes comprenant ceux recommandés par l'INCa permet de détecter de manière exhaustive les mutations de l'*EGFR*, dont la mutation de résistance aux ITK-EGFR de troisième génération p.(Cys797Ser) [15] qui est absente des mutations recherchées par le système Idylla. L'analyse par NGS permet également la détection sur d'autres gènes de variants d'intérêt thérapeutique. Par exemple, parmi les mutations retrouvées (Tableau 5), les mutations activatrices des oncogènes *KRAS* ou *BRAF* dans les CBNPC sont généralement mutuellement exclusives des mutations *EGFR* [16] et donnent ainsi une information précieuse sur les traitements à mettre en place. Par ailleurs, la présence d'une co-mutation des gènes *KRAS* et *STK11* a été associée à une moindre réponse aux inhibiteurs de *checkpoints* immunitaires (anti-PD-1, anti-PD-L1) [17]. Ou encore, les mutations susceptibles de créer un saut d'exon 14 du gène *MET* ou les insertions dans l'exon 20 du gène *HER2* sont des mutations d'intérêt dans les CBNPC et peuvent constituer des cibles thérapeutiques [18]. Les fréquences des mutations non *EGFR* retrouvées au NGS dans les 100 échantillons testés (Tableau 5) sont en accord avec les données de la littérature concernant les fréquences des oncogènes activés dans les CBNPC [3,4].

D'un point de vue pratique, le système Idylla est d'utilisation extrêmement simple aussi bien au niveau des infrastructures, du matériel et de la technique en comparaison avec le NGS. La démarche qualité selon la norme ISO15189, les ressources humaines nécessaires et l'habilitation du personnel sont également en faveur d'Idylla. La technique Idylla permet un délai de réponse

de quelques heures pour un échantillon, temps auquel il faut rajouter l'interprétation biologique du résultat et possiblement de visualisation et d'analyse des courbes d'amplification obtenues par Idylla Explore. Selon notre expertise, une analyse des courbes d'amplification devrait être systématique pour des valeurs de Cq comprises entre 23 et 26.

En conclusion, pour les 100 échantillons testés de manière prospective simultanément par séquençage NGS et par l'automate Idylla, 3 faux négatifs ont été rapportés avec le système Idylla, mais aucun faux positif sur un total de 9 patients éligibles à une thérapie ciblée par les ITK. Par ailleurs, il est à noter que sur les 12 variants de l'*EGFR* identifiés par NGS parmi les 9 patients, 4 variants n'ont pas été rendus par la procédure de compte rendu automatique de l'automate Idylla. Ces résultats sont en faveur d'une bonne valeur prédictive positive du système Idylla mais d'une valeur prédictive négative moins satisfaisante. Il faut cependant rappeler le contexte particulier des CBNPC où les échantillons sont très souvent pauvres en quantité de matériel et en cellules tumorales et approchent alors les limites de sensibilité des techniques analytiques.

Pour diminuer le nombre de faux négatifs avec le système Idylla, notre expérience nous incite à recommander : d'une part, augmenter le nombre de copeaux insérés dans la cassette en cas de faible cellularité tumorale du prélèvement, et d'autre part, regarder systématiquement les courbes d'amplification avec Idylla Explore pour les échantillons avec une valeur de Cq contrôle *EGFR* entre 23 et 26, ce qui est le cas pour un tiers des échantillons. L'analyse des courbes d'amplification peut permettre une seconde analyse Idylla avec plus de matériel FFPE ou avec de l'ADNg somatique extrait, ce qui est actuellement une utilisation en dehors des recommandations du fournisseur. La possibilité de récupérer le lysat dans la cassette Idylla préalablement analysée pourrait également être envisagée pour optimiser l'analyse et détecter d'autres cibles d'intérêt. Pour les valeurs de Cq contrôle *EGFR* inférieures à environ 23, l'automate Idylla semble une méthode rapide de réponse pour une caractérisation ciblée d'une majorité de variants de l'*EGFR*. Cependant, une analyse par séquençage NGS est une alternative plus informative pour les échantillons faiblement cellulaires, avec un Cq contrôle *EGFR* limite et pour la recherche de la mutation de résistance p.(Thr790Met) aux ITK-EGFR de première et deuxième générations.

Remerciements

AstraZeneca pour leur soutien financier pour l'achat des réactifs Idylla *EGFR* Mutation Test (Biocartis).

Biocartis pour leur collaboration à l'interprétation des cas discordants.

Déclaration de liens d'intérêts

C.L.-M., B.A., R.A., V.V., M.A. : frais de déplacements et d'hébergement pris en charge par AstraZeneca pour des réunions scientifiques.

Les autres auteurs déclarent ne pas avoir de liens d'intérêts.

Références

[1] Rapport interactif Les cancers en France - édition 2017[Internet]. Disponible sur : <https://www.e-cancer>.

- fr/Professionnels-de-sante/Les-chiffres-du-cancer-en-France/Epidemiologie-des-cancers/Les-cancers-les-plus-frequents/Cancer-du-poumon [consulté 27 mars 2019].
- [2] Hynes NE, MacDonald G. ErbB receptors and signaling pathways in cancer. *Curr Opin Cell Biol* 2009;21:177–84.
- [3] Pao W, Girard N. New driver mutations in non-small-cell lung cancer. *Lancet Oncol* 2011;12:175–80.
- [4] Meador CB, Micheel CM, Levy MA, Lovly CM, Horn L, Warner JL, et al. Beyond histology: translating tumor genotypes into clinically effective targeted therapies. *Clin Cancer Res* 2014;20:2264–75.
- [5] Mok TS, Wu Y-L, Thongprasert S, Yang C-H, Chu D-T, Saijo N, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 2009;361:947–57.
- [6] Tan LY, Walker SM, Lonergan T, Lima NE, Todd AV, Mokany E. Superior Multiplexing Capacity of PlexPrimers Enables Sensitive and Specific Detection of SNPs and Clustered Mutations in qPCR. *PLOS ONE* 2017;12:e0170087.
- [7] Ilie M, Butori C, Lassalle S, Heeke S, Piton N, Sabourin J-C, et al. Optimization of EGFR mutation detection by the fully-automated qPCR-based Idylla system on tumor tissue from patients with non-small cell lung cancer. *Oncotarget* 2017;8:103055–62.
- [8] Breton Q, Valo I, Benallègue F, Morel A, Verrière V. Influence de la cellularité tumorale dans la recherche des mutations des cancers du poumon : expérience du centre Paul Papin (Angers). [Internet]; 2015 <https://www.em-consulte.com/en/article/1008091> [consulté 28 mai 2019].
- [9] INCa. Mutations de l'EGFR dans le cancer du poumon : mise en évidence d'une cible moléculaire permettant un accès spécifique aux thérapies ciblées. Boulogne Billancourt: INCa; 2010.
- [10] Forde PM, Ettinger DS. Managing acquired resistance in EGFR-mutated non-small cell lung cancer. *Clin Adv Hematol Oncol HO* 2015;13:528–32.
- [11] Jotte RM, Spigel DR. Advances in molecular-based personalized non-small-cell lung cancer therapy: targeting epidermal growth factor receptor and mechanisms of resistance. *Cancer Med* 2015;4:1621–32.
- [12] Nguyen KSH, Kobayashi S, Costa DB. Acquired resistance to epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancers dependent on the epidermal growth factor receptor pathway. *Clin Lung Cancer* 2009;10:281–9.
- [13] Stewart EL, Tan SZ, Liu G. Known and putative mechanisms of resistance to EGFR targeted therapies in NSCLC patients with EGFR mutations – a review. *Transl Lung Cancer Res* 2015;4:67–81.
- [14] Jiang T, Zhou C. Clinical activity of the mutant-selective EGFR inhibitor AZD9291 in patients with EGFR inhibitor-resistant non-small cell lung cancer. *Transl Lung Cancer Res* 2014;3:370–2.
- [15] Thress KS, Paweletz CP, Felip E, Cho BC, Stetson D, Dougherty B, et al. Acquired EGFR C797S mediates resistance to AZD9291 in advanced non-small cell lung cancer harboring EGFR T790M. *Nat Med* 2015;21:560–2.
- [16] Sequist LV, Heist RS, Shaw AT, Fidias P, Rosovsky R, Temel JS, et al. Implementing multiplexed genotyping of non-small-cell lung cancers into routine clinical practice. *Ann Oncol* 2011;22:2616–24.
- [17] Skoulidis F, Goldberg ME, Greenawalt DM, Hellmann MD, Awad MM, Gainor JF, et al. STK11/LKB1 Mutations and PD-1 Inhibitor Resistance in KRAS-Mutant Lung Adenocarcinoma. *Cancer Discov* [Internet]; 2018 <http://cancerdiscovery.aacrjournals.org/content/early/2018/06/20/2159-8290.CD-18-0099> [consulté 19 avr 2019].
- [18] Genetically Informed Cancer Medicine – My Cancer Genome [Internet]. Disponible sur : [consulté 19 avr 2019].

DISCUSSION ARTICLE 1

Cette étude prospective avait pour objectif de comparer deux technologies : le système de PCR temps réel entièrement automatisé Idylla ciblant 51 mutations de l'*EGFR* et le séquençage par NGS d'un panel de 21 gènes comprenant les exons 18 à 21 de l'*EGFR*.

Du point de vue du matériel tumoral, les 100 échantillons ont parfaitement mis en lumière les contraintes auxquelles sont confrontés les laboratoires de biologie moléculaire en termes de qualité et quantité de matériel pour les CBNPC. En effet, quasiment la moitié des prélèvements avaient une cellularité tumorale inférieure ou égale à 30% dont 19 estimée à 10% par le pathologiste. Il est donc évident que la première qualité attendue pour les techniques de détection des mutations de l'*EGFR* est la meilleure sensibilité possible, celle-ci étant prioritaire sur la facilité d'utilisation de la technique, ce qui permet de maximiser les chances d'accès aux thérapies ciblées et en particulier pour cette étude aux ITK-EGFR.

Concernant la recherche de mutations activatrices *EGFR*, la technique NGS s'est avérée supérieure à la technique automatisée Idylla en termes de sensibilité et en termes d'efficacité de détection de la mutation de résistance p.Thr790Met aux ITK-EGFR de première et deuxième générations. Depuis cette publication en 2020, le débat concernant la recherche de mutations *EGFR* a changé pour deux raisons. D'une part, l'osimertinib (ITK-EGFR de troisième génération) est recommandé et autorisé dès la première ligne de traitement des CBNPC quelle que soit la mutation activatrice de l'*EGFR* (en dehors des insertions de l'exon 20). Cependant la détection des mutations de résistance *on target* secondaires à l'osimertinib est devenue incontournable pour orienter la prise en charge thérapeutique à progression. La plupart de ces mutations de résistance, dont la plus fréquente p.Cys797Ser³⁶⁸, ne sont pas détectées par les kits commerciaux de PCR temps réel les plus utilisés (Cobas *EGFR* Mutation Test v2, Therascreen *EGFR* RGQ PCR Kit v2, Idylla *EGFR* Mutation). D'autre part, de nouvelles opportunités thérapeutiques sont apparues récemment, et plus particulièrement les médicaments ciblant les insertions de l'exon 20 de l'*EGFR* (poziotinib, mobocertinib, amivantamab). Les kits commerciaux sus-cités ne détectent que les insertions de l'exon 20 les plus fréquentes et avec une sensibilité variable, entraînant une perte d'information importante sur l'addiction oncogénique de la tumeur et une perte de chance pour le patient à partir de la deuxième ligne de traitement.

Malgré les efforts des industriels pour la mise en place de techniques simples et rapides de détection des mutations activatrices de l'*EGFR*, le NGS reste à ce jour la technique de choix dans ce cadre.

Par ailleurs, le NGS présente l'avantage de permettre la détection d'altérations autres que les mutations de l'*EGFR*, ces altérations permettent de mieux comprendre le

mécanisme de tumorigénèse de chaque tumeur et d'élargir le champ de la médecine personnalisée dans le CBNPC. Notre travail a mis en évidence cet avantage du NGS (avec un panel de 21 gènes comprenant les cibles recommandées par l'INCa) avec la détection, parmi les 100 échantillons testés, de mutations activatrices de *KRAS* (dont 10 mutations p.Gly12Cys donnant accès au sotorasib) parfois associées à des mutations du gène suppresseur de tumeurs *STK11*, de mutations altérant l'épissage de l'exon 14 de *MET*, d'insertions dans l'exon 20 d'*HER2*, de mutations activatrices de *BRAF* ou encore de *NRAS*.

De plus, le séquençage de l'ARN par NGS (RNAseq) s'est largement développé ces dernières années, permettant de détecter de manière sensible et spécifique de nouvelles altérations de type réarrangements génomiques aboutissant à des transcrits de fusions pour lesquels des thérapies ciblées sont aujourd'hui autorisées (fusions impliquant *NTRK*, *ALK*, *ROS1*, *RET* ou *MET* avec différents partenaires). Ainsi, en amont du NGS, une extraction des acides nucléiques totaux à partir d'échantillons tumoraux (FFPE ou congelés) permet une stratégie parallèle (analyse simultanée par NGS ciblée sur ADN et ARN) ou une stratégie séquentielle (analyse par NGS ciblée sur ADN dans un premier temps puis, en absence de driver retrouvé au niveau des gènes *EGFR*, *BRAF*, *KRAS*, *ERBB2* et *MET*, une analyse ciblée sur ARN réalisée dans un deuxième temps). Il faut noter que moins de 47% des tumeurs de CBNPC sont pan-négatives sur ces drivers. L'équipe de Cohen *et al.* propose en 2020 une évaluation comparative de ces deux stratégies d'analyse moléculaire de CBNPC (Figure 27) et recommande une stratégie flexible en fonction du statut tabagique : les non-fumeurs devraient bénéficier d'une stratégie simultanée et les fumeurs d'une stratégie séquentielle. En effet, parmi les tumeurs de non-fumeurs testées, 22% présentaient une fusion ou une anomalie d'épissage versus 2% pour les fumeurs.³⁶⁹ En pratique et en attente de recommandations de l'INCa ou des sociétés savantes, les paramètres à prendre en compte sont l'équipement du laboratoire, l'accès au statut tabagique du patient, le coût des tests qui doit rester cohérent avec leur remboursement et le délai de rendu des résultats pour une prise en charge compatible avec la mise en place rapide d'un traitement.

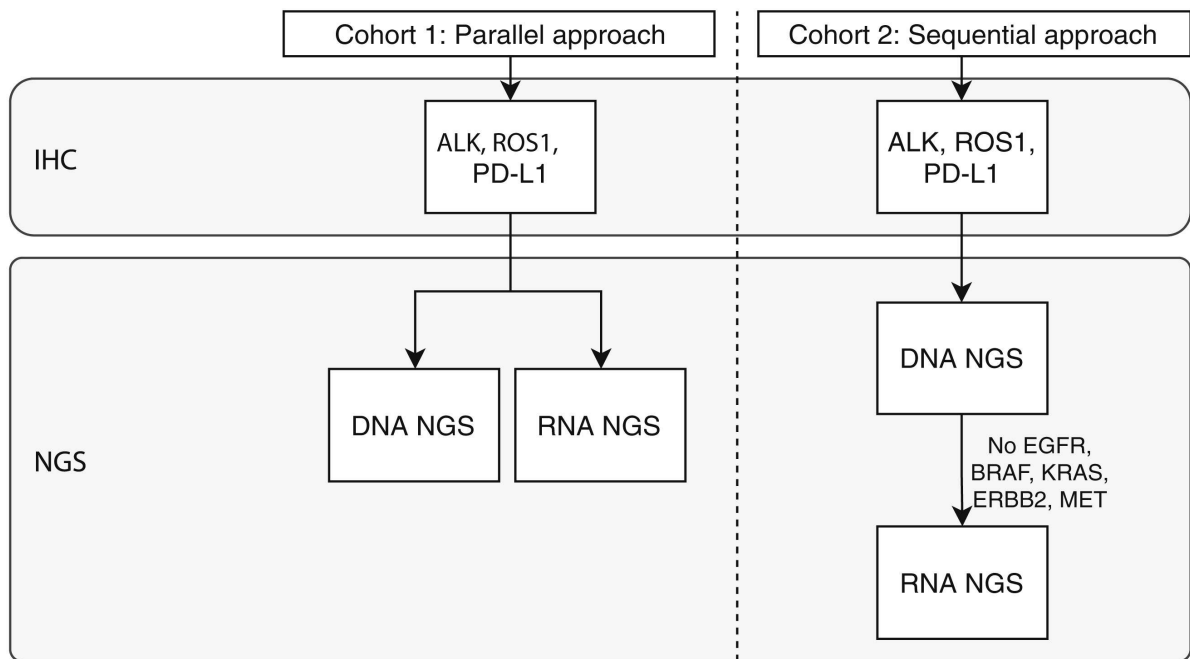


Figure 27 : Approches parallèle et séquentielle pour l'analyse moléculaire des CBNPC. D'après Cohen et al.³⁶⁹

En outre, l'analyse moléculaire de l'ADNtc s'est largement développée dans le contexte de CBNPC. Compte tenu de la faible quantité d'ADNtc, cette recherche doit être réalisée avec des techniques d'une grande sensibilité analytique. Certaines techniques ciblées commerciales comme le kit Cobas *EGFR* Mutation Test v2 marqué CE-IVD annoncent une très bonne sensibilité dans l'ADNtc pour les mutations les plus fréquentes d'*EGFR*, d'autres ne sont qu'au stade du développement (Société Biocartis). La sensibilité du NGS, généralement aux alentours de 2% de fréquence allélique, a été nettement améliorée ces dernières années grâce à une nouvelle approche visant à étiqueter les fragments d'ADN par des barcodes moléculaires uniques appelés UMI (*Unique Molecular Identifiers*) avant toute étape d'amplification par PCR. Cette approche a permis de détecter et même corriger les erreurs dans les lectures (*reads*) issus du NGS ainsi que de quantifier avec précision le nombre des molécules initiales d'acides nucléiques. L'utilisation d'UMI dans les panels de gènes de NGS est particulièrement utile dans le cadre de l'analyse de l'ADNtc mais elle reste cependant coûteuse.

En conclusion concernant ce premier article, malgré un temps un peu plus long, la supériorité du NGS pour l'analyse moléculaire des CBNPC s'impose aujourd'hui pour l'exhaustivité des cibles recherchées, pour sa grande sensibilité notamment dans l'ADNtc grâce aux UMI, et pour le séquençage d'ARN en deuxième voire en première intention. Cependant, les techniques ciblées peuvent trouver leur place en cas d'urgences extrêmes.

2. INTERPRETATION FONCTIONNELLE DES VSI *BRCA1* et *BRCA2*

Les gènes *BRCA1* et *BRCA2* sont des gènes suppresseurs de tumeur principalement impliqués dans les voies de réparation de l'ADN. La mise en évidence d'altérations pathogènes de ces gènes présente un double enjeu en cancérologie : d'une part l'implication dans une prédisposition héréditaire à certains cancers et d'autre part en tant que biomarqueur d'une sensibilité accrue aux thérapies ciblées orales par iPARP.

En génétique tumorale, la recherche de variations pathogènes des gènes *BRCA* requiert un séquençage complet des parties codantes et introniques avoisinantes par NGS. Les variants qui en résultent sont répartis sur toute la longueur des gènes et le rôle des laboratoires est d'évaluer leur pathogénicité en se basant sur des algorithmes décisionnels utilisant des bases de données, des critères cliniques et de co-ségrégation, des outils de prédiction informatiques ou encore des données fonctionnelles. Le développement du NGS, des consultations d'oncogénétique et des iPARP ont progressivement permis de mettre en évidence des milliers de variants *BRCA1* et *BRCA2* incluant un grand nombre de VSI qui restent aujourd'hui à annoter que ce soit dans un but de conseil génétique ou dans un but thérapeutique. Pour progresser dans la caractérisation de ces VSI *BRCA1* et *BRCA2*, le développement de tests fonctionnels s'impose. Par ailleurs, en génétique somatique, seuls des échantillons FFPE sont disponibles et la qualité des acides nucléiques de ces échantillons est médiocre.

Ainsi nous proposons deux méthodologies d'évaluation fonctionnelle de VSI *BRCA* adaptées à la contrainte des tissus FFPE et réalisables dans un délai compatible avec les impératifs de prise en charge des patients par les thérapies ciblées. L'article 2 expose une démarche expérimentale limitée aux variants d'épissage du gène *BRCA1*, applicable à d'autres gènes. A partir de l'ARNm extrait de tissus tumoraux FFPE, le principe repose sur une analyse par électrophorèse capillaire après transcription inverse puis amplification par PCR pour visualiser les potentiels sauts d'exons ou rétentions d'introns. La méthodologie présentée dans l'article 3 permet quant à elle la caractérisation fonctionnelle de VSI non limitée aux altérations d'épissage et basée sur la technologie d'édition génomique CRISPR-Cas9. Chaque mutation est évaluée par un score fonctionnel. Cette méthodologie est applicable à tous les gènes suppresseurs de tumeurs essentiels dans un modèle unique.

ARTICLE 2 : SOMATIC mRNA ANALYSIS OF *BRCA1* SPLICE VARIANTS PROVIDES A DIRECT THERANOSTIC IMPACT ON PARP INHIBITORS

Louise-Marie Chevalier, Amandine Billaud, Sabrina Fronteau, Jonathan Dauvé, Anne Patsouris, Véronique Verrièle and Alain Morel

Publié le 2 mars 2020 – *Molecular Diagnosis & Therapy*



Somatic mRNA Analysis of *BRCA1* Splice Variants Provides a Direct Theranostic Impact on PARP Inhibitors

Louise-Marie Chevalier^{1,2} · Amandine Billaud^{1,2} · Sabrina Fronteau² · Jonathan Dauvé² · Anne Patsouris² · Véronique Verrièle² · Alain Morel^{1,2}

© Springer Nature Switzerland AG 2020

Abstract

Background The identification of pretherapeutic somatic *BRCA* variants can have considerable clinical impact given that they affect response to the new poly (ADP-ribose) polymerase (PARP)-targeted therapy. One major issue with this type of testing is the identification of splicing variants of uncertain significance (VUS) on degraded somatic messenger RNA. It is therefore important to be able to quickly characterize these splice variants.

Objective As part of PARP inhibitor targeted therapy, we have investigated a method for the direct confirmation of potential pathogenic somatic splice variants of *BRCA1* found in fixed tumor samples. Previously these VUS have commonly only been tested by in silico analysis.

Methods Five *BRCA1* variants affecting splicing were characterized from formalin-fixed, paraffin-embedded (FFPE) ovarian carcinoma tissues by next-generation sequencing (NGS). Three patient samples had already been functionally characterized and were used as controls. Total somatic RNA from samples was extracted, reverse-transcribed, and amplified with several primer pairs encompassing the target exon. The polymerase chain reaction (PCR) products were analyzed by capillary gel electrophoresis to assess possible changes in size due to splicing alterations. Finally, we confirmed our results by cloning, followed by Sanger sequencing, and analyzed the expression of the aberrant forms.

Results Our molecular approach made it possible to visualize the splicing outcomes of three variants (c.5194-2A>G, c.5434C>G, and c.547+1G>A) already identified and present in databases and/or identified with prediction tools (ClinVar, UMD, ARUP Utah database, and Human Splice Finder splices sites prediction) and to confirm their exon skipping consequences, their expression in tumors, and thus their pathogenicity. The c.4484+5G>A variant was not found in databases and was predicted to have no impact on splicing, but was found to display altered processing in tumor tissue. This variant also had a major detrimental impact on transcriptional expression.

Conclusion In a break from purely in silico approaches, we propose a simple and rapid pretherapeutic functional analysis of somatic *BRCA1* variants potentially involved in splicing alterations. This approach will allow more ovarian cancer patients to benefit from new therapies targeting PARP.

Louise-Marie Chevalier and Amandine Billaud contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s40291-020-00452-z>) contains supplementary material, which is available to authorized users.

✉ Alain Morel
Alain.morel@ico.unicancer.fr

¹ Université d'Angers, Inserm, CRCINA, 49000 Angers, France

² Institut de Cancérologie de l'Ouest - Paul Papin, 49000 Angers, France

Key Points

Many variants of uncertain significance on *BRCA1* splicing are detected by next-generation sequencing in ovarian tumors.

In silico prediction of the variant function does not always provide functional biologic proof and is a clinical issue.

Somatic messenger RNA analysis from formalin-fixed, paraffin-embedded samples will allow more patients to benefit from targeted therapy.

Published online: 02 March 2020

Adis

1 Introduction

Ovarian cancer (OC) is the seventh most common cancer and the eighth leading cause of death in women worldwide [1]. Epithelial ovarian cancer (EOC) accounts for more than 90% of all OCs, and the most common histological subtype is high-grade serous [2]. Patients are often diagnosed with advanced disease and are typically managed with cytoreductive surgery followed by platinum-based chemotherapy. However, most patients relapse and are further treated with additional lines of salvage chemotherapy that may increase survival but are not curative [3]. Consequently, there is a rising interest in the use of targeted molecular and biologic therapies in EOC treatment, including antiangiogenic agents, signaling pathway inhibitors, immunotherapies [4, 5], and poly (ADP-ribose) polymerase inhibitors (PARPi). PARPi is a pharmacological inhibitor of the poly (ADP-ribose) polymerase (PARP) enzymes that play a key role in regulating several DNA repair pathways and, notably, single-stranded breaks (SSB). Direct inhibition of PARP leads to accumulation of SSBs, causing the collapse of replication forks and accumulation of DNA double-stranded breaks (DSBs) [6]. DSBs are commonly repaired by non-homologous end-joining (NHEJ) or by a homologous recombination repair system. BRCA1, BRCA2, and other homologous recombination proteins are recruited to repair DNA DSBs. Tumor cells with altered BRCA1 or BRCA2 function are markedly sensitive to PARPi due to homologous recombination deficiency (HRD). Cell death due to the simultaneous lack of function of two or more genes is the cornerstone of synthetic lethal therapies [7]. Furthermore, previous treatments with platinum-based chemotherapies induce DNA damage, which is not repaired due to BRCA1/2 lack of function. Increasing genomic instability sensitizes cells to PARPi, resulting in enhanced cytotoxicity [8].

In the European Union, maintenance treatment with PARPi is approved for relapsed platinum-sensitive EOC patients presenting *BRCA1/2* deficient variants. Acquired or inherited variants in different DNA repair pathways may be found in EOC. The introduction of PARPi into clinical practice and the availability of data from the trials SOLO1 (D0816C00001) and SOLO2 (D0816C00002) suggest the clinical utility of *BRCA1/2* testing within a time period compatible with the beginning of targeted therapy (within a maximum of 2 months). Apart from *BRCA1/2* inherited variants, somatic *BRCA1/2* variants occur in approximately 7% of EOC [9], and these women could also benefit from PARPi in terms of improved survival following recurrence. Therefore, genetic testing by next-generation sequencing (NGS) of somatic *BRCA1/BRCA2* variants in EOC is a general practice. Unlike germline variants, many somatic variants are reported but not annotated or not even reported in scientific publications or *BRCA1/2* databases, and the assessment of their clinical

impact is an important issue in theranostic expertise. Among these variants of uncertain significance (VUS), the functional implication of some putative *BRCA1/2* splicing variants is very difficult to interpret. Whereas alteration of the canonical splice site (directed by the GU-AG rules) are well understood, in some reported cases, the rules do not seem to apply, highlighting the importance of functional testing [10]. Moreover, classification of exonic or intronic variants beyond these 2 bps remains difficult. These variants could affect splicing by various means: exon skipping, intron retention, activation of cryptic splice sites, or introduction of new splice sites within introns and exons. Even silent variants could have unexpectedly deleterious impacts on messenger RNA (mRNA) splicing [11]. It is well known that splicing variants are largely underestimated [12], and as a consequence, many patients cannot benefit from PARPi therapy. To assess variant pathogenicity on the splicing machinery, *in silico* prediction algorithms are available, but these tools are not 100% reliable [12].

The aim of this study was to apply, within a period consistent with the implementation of PARPi, a method to determine the pathogenic potential of some *BRCA1* somatic potential splicing VUS commonly tested using *in silico* prediction analysis. For this purpose, we used an experimental direct analysis of somatic mRNA extracted from formalin-fixed, paraffin-embedded (FFPE) tissue. Reverse transcription polymerase chain reaction (PCR) analysis revealed several aberrant transcripts resulting from splicing alterations, which were confirmed by sequencing.

2 Materials and Methods

2.1 Study Design

Potential splice variants characterized in somatic genomic DNA (gDNA) by NGS were confirmed by Sanger sequencing and tested for pathogenicity by comparing databases and predictive bioinformatics tools. Experimental analysis of somatic mRNA was then performed to directly assess the impact of these somatic splice variants on splicing and, consequently, on protein function (Fig. 1a).

2.2 Tissue Samples and Cell Lines

Tumor specimens were obtained from five patients with high-grade serous EOC assessed by an experienced histopathologist. The study was conducted on FFPE tissue samples. Hematoxylin and eosin (H&E)-stained slides were used to determine tumor cell content by area. For each sample, three 10- μ m sections from FFPE blocks were used for DNA or RNA analysis. The NCI-H1975 cell line (ATCC[®] CRL5908[™]) was used for quality control at each step.

 Adis

2.3 DNA and RNA Extraction Protocol

DNA isolation from FFPE samples was performed with the NucleoSpin® DNA FFPE XS Kit (Macherey–Nagel) according to the manufacturer's instructions. RNA isolation from FFPE samples was performed with the RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE (Thermo Fisher) using the manufacturer's instructions. Both concentrations were determined using the Qubit (Thermo Fisher) associated with the Quantifluor dsDNA System Kit (Promega) for DNA and the Qubit RNA HS Assay Kit (Thermo Fisher) for RNA.

2.4 NGS and Data Analysis

The first stages of NGS library preparation were carried out with the Oncomine BRCA Assay Manual Kit (Thermo Fisher Scientific). Briefly, barcoded libraries were generated from 20 ng of DNA per sample. Two premixed pools of 265 primer pairs for entire *BRCA1/2* coding regions and noncoding putative splice boundaries were used to generate the sequencing libraries. Clonal amplification of the libraries was carried out by emulsion PCR using an Ion Chef System (Thermo Fisher Scientific) according to the manufacturer's instructions. The prepared libraries were then sequenced on Ion Torrent S5 Sequencer using Ion 520&530 Kit Chef (Thermo Fisher Scientific). Variants of interest were visualized in Integrative Genomics Viewer (IGV), a high-performance visualization tool for interactive exploration of large integrated genomic datasets on standard desktop computers [13]. Human Genome Variation Society (HGVS)-approved guidelines (<http://www.hgvs.org/mutnomen/>) were used for *BRCA1* nomenclature. The variants found by NGS were researched in the UMD-BRCA1 database (<http://www.umd.be/BRCA1/>) [14], the ClinVar database (<https://www.ncbi.nlm.nih.gov/clinvar/>) [15], and the BRCA Mutation Database (http://www.arup.utah.edu/database/BRCA/Home/BRCA1_landing.php).

2.5 Sanger DNA Sequencing

Sanger sequencing was used to confirm the variants characterized by NGS and to sequence complementary DNA (cDNA) cloned in pGEM®-T Easy Vector Systems (Promega). Primers specific to the PCR template were designed using Primer3 and BLAST (<https://primer3plus.com/> and <https://www.ncbi.nlm.nih.gov/tools/primer-blast/>). Custom DNA oligonucleotides were purchased from Eurofins Genomics (<https://www.eurofinsgenomics.eu/>). PCR using Go Taq G2 Hot Start Colorless Master Mix (Promega) or GoTaq® Flexi DNA Polymerase (Promega) was followed by purification with Nucleospin Gel and PCR Clean-up Kit (Macherey–Nagel). Sequencing was performed with the

dye terminator chemistry BigDye Terminator v1.1 Cycle Sequencing Kit (Thermo Fisher) on the 3130 ABI Genetic Analyzer (Thermo Fisher). The sequencing data were viewed and analyzed using Sequencing Analysis 5.3.1 software.

2.6 RT-PCR and *BRCA1* cDNA Characterization

Reverse transcription PCR (RT-PCR) was carried out with the SuperScript VILO cDNA Synthesis Kit (Invitrogen) using 200 ng of total RNA. Specific primers were designed encompassing the exon of interest for the evaluation of exon skipping (Fig. 1b). These primers were localized on exon junctions. A subsequent PCR was performed using GoTaq® Flexi DNA Polymerase (Promega) on an Eppendorf thermocycler to obtain fragments of interest, whose sizes were evaluated by capillary electrophoresis using the QIAxcel Advanced (Qiagen) system with the QIAxcel DNA High-Resolution Kit (1200). PCR products were separated on 3% agarose gels and visualized with a fluorescent nucleic acid dye [GelRed™ Nucleic Acid Gel Stain (Biotium)], and bands of interest were extracted by the Nucleospin Gel and PCR Clean-up Kit (Macherey–Nagel). DNA fragments were ligated into pGEM®-T Easy Vector Systems (Promega) before bacterial transformation in JM109 High-Efficiency Competent Cells (Promega). The plasmid was extracted by the NucleoSpin® Plasmid Kit (Macherey–Nagel), and sequencing was performed with the BigDye Terminator V1.1 Kit from Thermo Fisher, as described by the manufacturer.

2.7 In Silico Analysis

The Human Splice Finder (HSF) algorithm was employed for splice site prediction [16]. The HSF system combines 12 different algorithms, including the Maximum Entropy Model (MaxEntScan) [17], to identify and predict the effect of variants on splicing motifs, including the acceptor and donor splice sites, the branch point, and auxiliary sequences known to either enhance or repress splicing [exonic splicing enhancers (ESE) and exonic splicing silencers (ESS)].

2.8 RT-qPCR Analysis of *BRCA1* Expression

We evaluated the impact of aberrant transcripts on *BRCA1* expression by performing reverse transcription quantitative polymerase chain reaction (RT-qPCR) on cDNA generated by reverse transcription, as described above. We used primers designed for specific amplification of the expected full-length transcript or the exon-skipping transcript. We thus designed two sets of primers for each variant. Another set of primers was designed to amplify the junction between exons 2 and 3, as an internal control. The primer sequences

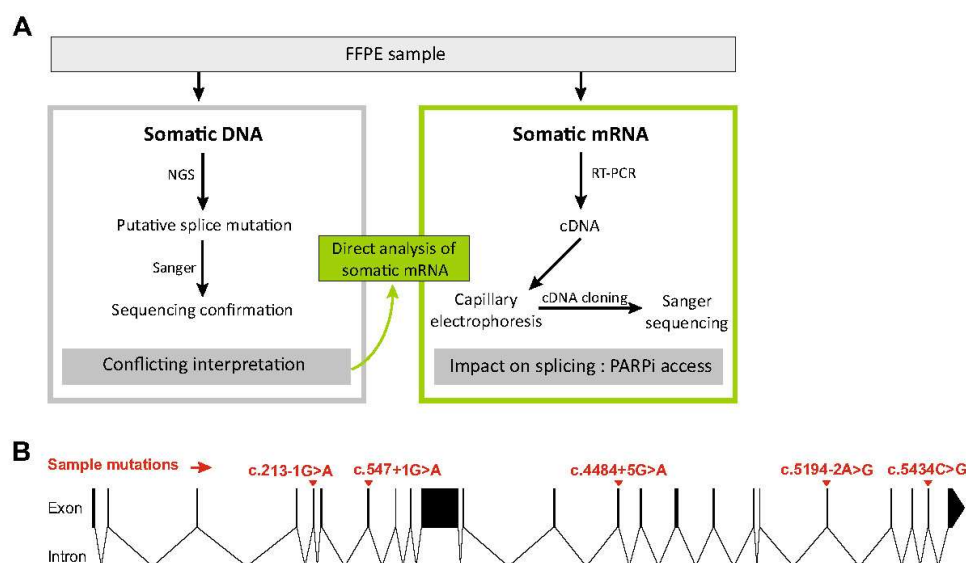


Fig. 1 Workflow for splice variant assessment in this study. **a** Evaluation of the functional impact of variants on splicing by combined somatic DNA sequencing followed by directed cDNA cloning-sequencing from mRNA isolated from FFPE samples. **b** Exon–intron structure of the *BRCA1* gene, according to <http://wormweb.org/exonintron>. The

positions of the sequence variants found in each sample and evaluated here are indicated. *cDNA* complementary DNA, *FFPE* formalin-fixed, paraffin-embedded, *mRNA* messenger RNA, *NGS* next-generation sequencing, *PARPi* poly (ADP-ribose) polymerase inhibitor, *RT-PCR* reverse transcription polymerase chain reaction

are provided in Table 1. Analysis was performed using Sybr Green (Thermo Fisher Scientific) and the comparative *Ct* method [$2^{-(\Delta\Delta Ct)}$].

3 Results

3.1 Patients and Tumors Characteristics

In this study, the mean age at diagnosis was 63.2 years (range 53–72 years). The five patients were Caucasian and had OC. The search for somatic variants of the *BRCA1* and *BRCA2* genes in our laboratory was essential for their therapeutic management, determining whether or not they were given PARPi. Tissue samples came from primary tumors or from metastatic tissue. All tumor samples received in the laboratory had sufficient tumor cellularity (> 10% tumor cells) to search for variants. The characteristics of these tumors are summarized in Table 2. Two patients selected for this study had a putative splicing VUS, and functional characterization was therefore required. Moreover, three patients, with variants already classified as pathogenic, were selected to assess the validity of our protocol.

3.2 DNA Analysis: *BRCA1* Variants, Databases Analysis, and In Silico Prediction

A *BRCA1* variant with a high allelic frequency (mean 63.3%, range 48.5–75.9) was demonstrated by NGS for each patient (Table 3). Variants were visualized on IGV and confirmed by Sanger sequencing. The results presented here are those obtained from patient 4 (Fig. 2a, b). The results obtained for the other patients are shown in the supplementary figure (see the electronic supplementary material). Primer sequences for putative pathogenic variants confirmed by Sanger sequencing are shown in Table 1. With collaboration with the genetics department of the *Institut de Cancérologie de l'Ouest*, we found that two out of the five variants were germline and two were somatic. We were unable to search for a germline variant in one of the five patients (the patient died in the weeks following the search for *BRCA1* and *BRCA2* somatic variants).

For interpretation of the pathogenicity of the sequenced variants (Table 3), one variant [c.4484+5G>A (p.?)] was listed only in the ClinVar database, with the interpretation “uncertain significance,” and was absent from the other databases consulted. The other variants were classified as deleterious (“pathogenic” or “likely pathogenic”) in at least one of the three databases consulted.

△ Adis

Table 1 PCR primer pairs used to amplify and sequence mutated *BRCA1* fragments and to analyze its expression

Variants	Forward primer (5'-3')	Reverse primer (5'-3')	T_m (°C)	
c.5194-2A>G	AGGGGAGTGAATACAGAGT	GCCTGAATGCCTTAAATATG	55	<i>BRCA1</i> variant confirmation by Sanger sequencing
c.4484+5G>A	AAACTGGAGAAAAGTATGGTGA AAAA	GGCCTTCTGCTGACAAGTT	55	
c.5434C>G	CAAGCACAGGTAATGAGTG	AGTGACAGTTCAGTAGTCCT	55	
c.547+1G>A	AAAGCTGCCTACCACAAATA	GGATACAACCTCAAAGACG	55	
c.213-1G>A	ACCTGTGTCAAGCTGAAAAGC	CACTGCCATCACACGGTTTA	55	
c.5194-2A>G	GGTGACCCAGTCTATTAAAGAAA	GGGCCCATAGCAACAGAT	53	<i>BRCA1</i> mutated fragment amplification
c.4484+5G>A	ATCCAGAACAAAGCACATCA	CAACTGTGCATGTACCACCT	55	
c.5434C>G	TACAGCTGTGTGGTGCTTCT	GTGCTACACTGTCCAACACC	55	
c.547+1G>A	CCGTGCCAAAAGACTTCTAC	TGCCTTATTAACGGTATCTTCA	53	
c.213-1G>A	GCCTTCACAGTGTCTTTIATG	CGTTTCTGTAGCCATACT	53	
Exon 2-3	TGCAGAAAATCTTAGAGTGTCCC	AAATATGTGGTCACTTTGTGG	58	<i>BRCA1</i> expression by qPCR analysis
Exon 19-20	GGGTGACCCAGTCTATTAAAGAAA	CCATTGACCACATCTCCTCTGA	59	
Exon 19-21	AAAATGCTGAATGAGATCTTCA	CATTCCAGTTGATCTGTGGGC	56.5	
Exon 13-14	TCCAGAACAAAGCACATCAGA	GCCTTCTGGATTCTGGCTTAT	57.6	
Exon 13-15	GCACATCAGAAAAAGGTCATCC	TCCCAGAGCAACTGTGCAT	57.5	
Exon 22-23	GGAGCTTTCATCATTACCCCTTGG	AGCCATTGTCTCTGTCCAGG	61.8	
Exon 22-24	ATTACCCCTTGGCACACAAT	CACTGTCCAACACCCACTCT	58.5	
Ex7-8	AGACTTCTACAGAGTGAACCCG	ACAGTTCCAAGGTTAGAGAGTTGG	59.5	
Ex7-9	CCGAAAATCCTTCTCTGGAT	CTTGATCTCCACACTGCAA	56.6	

PCR polymerase chain reaction, qPCR quantitative polymerase chain reaction, T_m melting temperature

Table 2 Baseline tumor characteristics

	Sample				
	1	2	3	4	5
Age at diagnosis (years)	53	66	68	72	57
Tumor histology	High-grade Serous ovarian carcinoma	High-grade Serous ovarian carcinoma	High-grade Serous ovarian carcinoma	High-grade Undifferentiated ovarian carcinoma	Ovarian clear cell adenocarcinoma
Tumor type	Primary	Metastatic	Metastatic	Metastatic	Primary
Tumor stage (FIGO)	IIIc	IV	Unknown	IIIc	-
Tumor contents	80%	60%	30%	60%	70%
Tumor fragment type	Surgical sample	Biopsy	Biopsy	Biopsy	Surgical sample

FIGO Federation International of Gynecology and Obstetrics

In silico prediction tools (HSF; Table 3) identified four of the five variants studied as having potentially altered splicing. The c.4484+5G>A (p.?) variant found in sample 2 was predicted by HSF to likely have no impact on splicing.

3.3 RNA Analysis

The fragment amplified from the *BRCA1* wild-type (WT) H1975 cell line produced the expected full-length transcript (Table 1) and was used as a positive control, together with a *BRCA1* WT FFPE sample. Samples 1 (c.5194-2A>G) and

4 (c.547+1G>A) generated two fragments, corresponding to the mutated fragment and the WT fragment, and samples 2 (c.4484+5G>A) and 3 (c.5434C>G) yielded only the mutated amplicon. All PCR products were further characterized by Sanger sequencing to confirm the putative exon skipping (Fig. 2c-e). The results shown are those for sample 4. The results for the other patients are presented in the supplementary figure (see the electronic supplementary material). Sample 5 (c.213-1G>A) did not generate any useful PCR products for further analysis. Sanger sequencing confirmed exon skipping and the presence of aberrant transcripts in the four samples analyzed (Fig. 2e and Table 3).

Adis

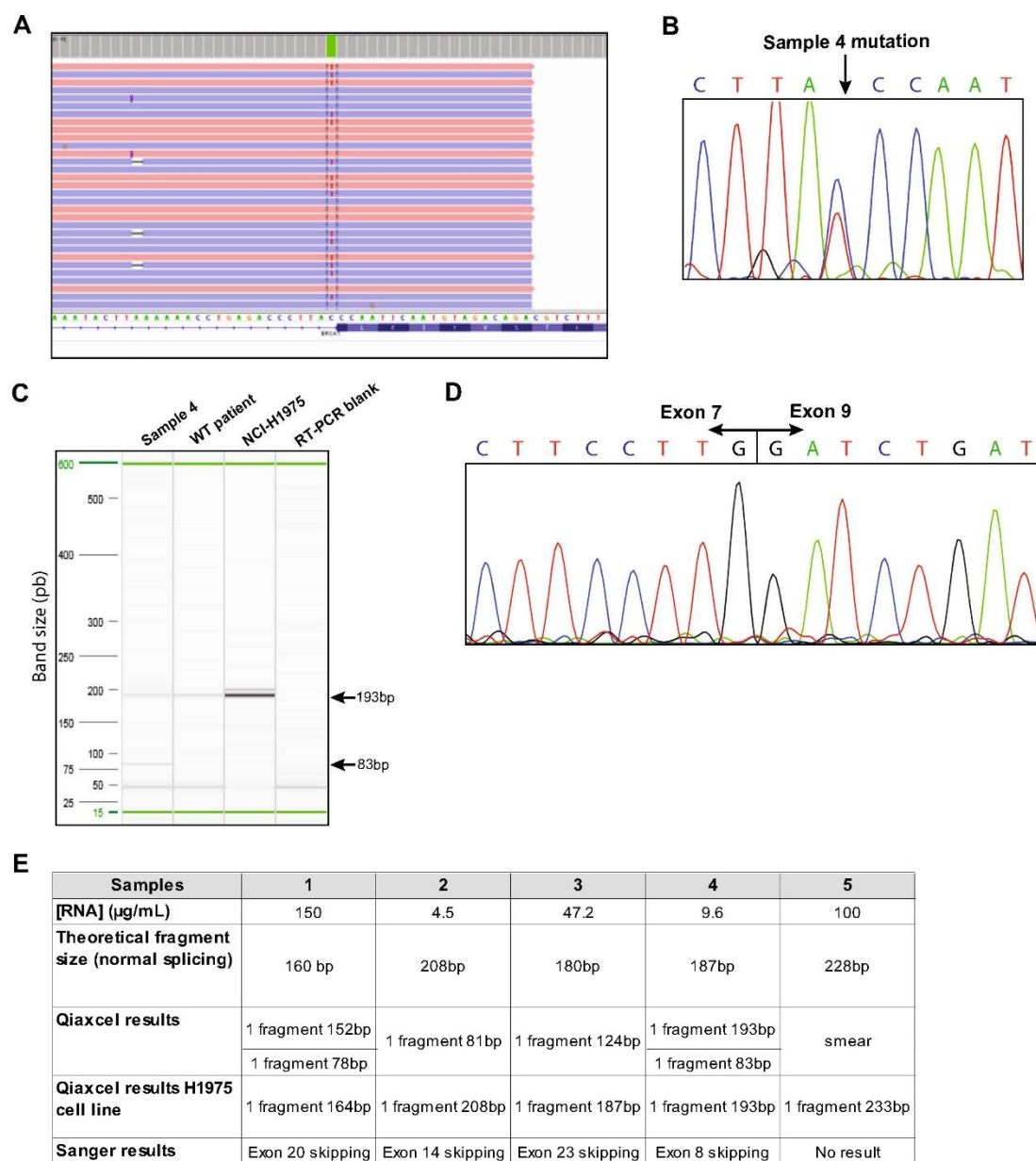


Fig. 2 Analysis of the c.547+1G>A variant from sample 4. **a** The Ion Torrent data displayed in IGV from the *BRCA1* variant characterization. Forward sequences reads are in red, and reverse reads are in blue. **b** Sanger sequencing of the mutated *BRCA1* amplicon for confirmation of the variant. **c** Analysis by capillary electrophoresis of the reverse transcription amplification product encompassing the putative skipped exon. **d** Sanger sequencing of the 83-bp amplicon after clon-

ing indicates the absence of exon 8. The 193-bp product corresponds to the normal splice variant of *BRCA1*. **e** The table shows the results of *BRCA1* mRNA characterization for the five FFPE ovarian tumors analyzed here. *FFPE* formalin-fixed, paraffin-embedded, *IGV* integrative genomics viewer, *mRNA* messenger RNA, *RT-PCR* reverse transcription polymerase chain reaction, *WT* wild-type

△ Adis

Table 3 Results of *BRCA1* analysis

	Sample				
	1	2	3	4	5
DNA (µg/mL)	38.8	2.3	7.4	8.2	9.9
Variants	c.5194-2A>G	c.4484+5G>A	c.5434C>G	c.547+1G>A	c.213-1G>A
AA change	p.?	p.?	p.Pro1812Ala	p.?	p.?
Exon	20	14	23	8	6
Allele frequency (%)	75.7	64.5	75.9	48.5	51.9
Somatic or germline	Somatic	Unknown	Germline	Somatic	Germline
ClinVar database interpretation	Pathogenic/likely pathogenic	Uncertain significance	Pathogenic/likely pathogenic	Likely pathogenic	Pathogenic
UMD database interpretation	Not recorded	Not recorded	Class 5: Causal	Not recorded	Class 5: Causal
ARUP Utah database interpretation	Class 5: Definitely pathogenic	Not recorded	Not recorded	Not recorded	Class 5: Definitely pathogenic
HSF splice sites prediction	Alteration of the WT acceptor site, most probably affecting splicing	Alteration of an intronic ESS site Probably no impact on splicing	Creation of an exonic ESS site Potential alteration of splicing Alteration of an exonic ESE site Potential alteration of splicing	Alteration of the WT donor site, most probably affecting splicing	Alteration of the WT acceptor site, most probably affecting splicing Activation of an intronic cryptic acceptor site Potential alteration of splicing
LOVD database	Pathogenic	Not recorded	Conflicting interpretations of pathogenicity	Not recorded	Pathogenic
ENIGMA database	Not recorded	Not recorded	Not recorded	Not recorded	Not recorded
COSMIC identifiers	Not recorded	Not recorded	Not recorded	COSM3387921	COSM6922026
RNA analysis results	Exon 20 skipping r.5194_5277del	Exon 14 skipping r.4358_4484del	Exon 23 skipping r.5415_5477del	Exon 8 skipping r.442_547del	No result
Predicted protein impact	Alteration of the BRCT2 domain	Alteration of the coiled-coil domain	Alteration of the BRCT2 domain	Alteration of the MB2 domain (interaction with c-Myc)	No result

AA amino acid, *ESE* exonic splicing enhancer, *ESS* exonic splicing silencer, *HSF* Human Splice Finder, *WT* wild-type

3.4 Evaluation of Splicing Alteration Expression by qPCR Analysis

We were able to demonstrate the occurrence of alternative splicing events for these *BRCA1* variants. However, we also performed RT-qPCR and compared full-length and aberrant transcripts to assess the real impact of the variants on transcript levels (Fig. 3). For each patient, the results were compared to those for an FFPE sample with WT *BRCA1* expression, to check the specificity of our primers. Exons 2 and 3 (RING domain) were also amplified, as an internal control. Given the large differences in *BRCA1* expression between samples, results were normalized with this internal control. All primer (Table 1) efficiencies were assessed before the experiment, and ranged from 103.5 to 110.1%. As in the previous analysis, no amplicon was obtained from sample 5. For patients 1 and

4, the production of aberrant transcripts was demonstrated, but the full-length forms were more abundant (Fig. 3a, d). By contrast, patients 2 and 3 mostly produced the aberrant transcript (Fig. 3b, c). These results were consistent with the results of capillary electrophoresis (Fig. 2c and supplementary figure C).

3.5 Theranostic Impact and Clinical Utility

Following the molecular proof of *BRCA1* splice alteration, patients 1 (c.5194-2A>G) and 4 (c.547+1G>A) were treated with a PARPi (olaparib) according to the recommendations (recurrent platinum-sensitive OC with mutation). Both patients presented a relapse, after 3 and 9 months, respectively, of maintenance treatment. Unfortunately, patient 2 (c.4484+5G>A) died a few weeks after the oncologist's request for analysis. Patient 5 (c.213-1G>A)

△ Adis

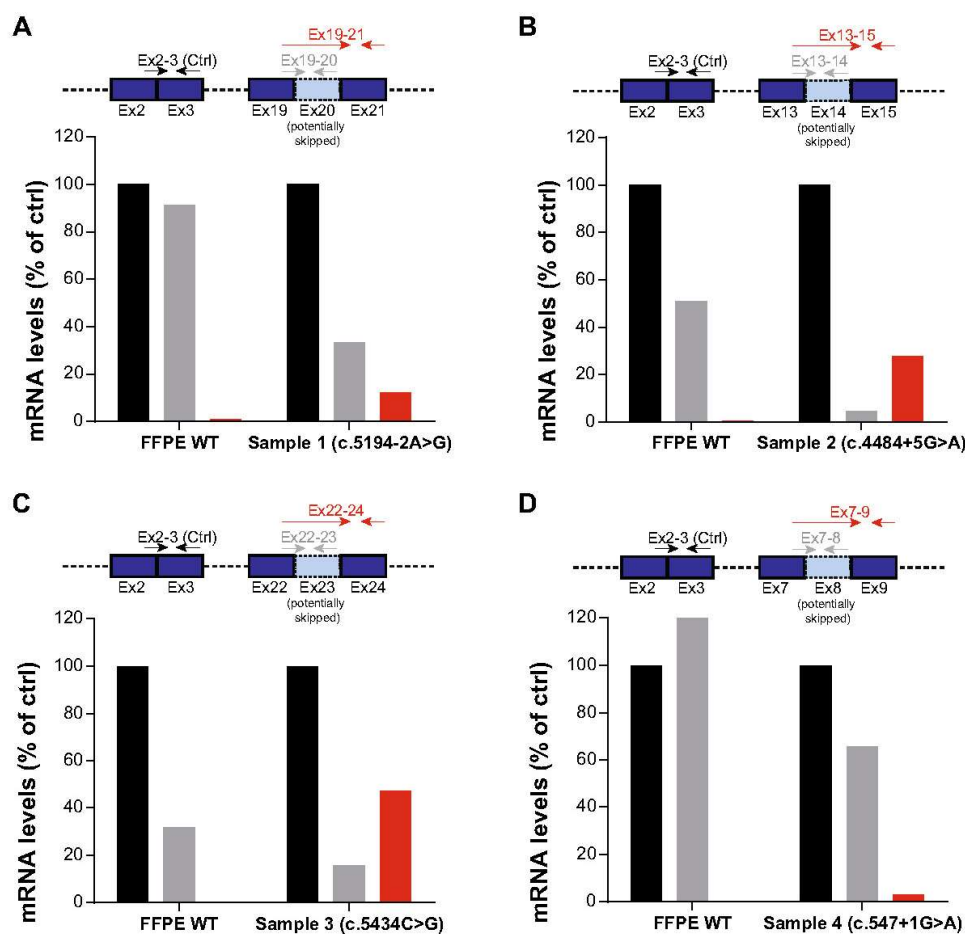


Fig. 3 Analysis of aberrant transcript levels. Aberrant and full-length transcript levels were compared by RT-qPCR analysis. Another FFPE sample with WT *BRCA1* expression was used to check the efficiency of the aberrant form primers. An internal control (exons 2 and 3 of

BRCA1) was used to normalize all results. a Sample 1, b sample 2, c sample 3, and d sample 4. *Ctrl* control, *FFPE* formalin-fixed, paraffin-embedded, *mRNA* messenger RNA, *RT-qPCR* reverse transcription quantitative polymerase chain reaction, *WT* wild-type

is in remission after the first line of platinum-based chemotherapy, and is eligible for treatment should a platinum-sensitive relapse occur. Patient 3 (c.5434C>G) presented a recurrent platinum-sensitive OC, and PARPi (olaparib) maintenance therapy was therefore initiated. The therapeutic decisions taken after screening for *BRCA1* variants and patient outcomes are summarized in Table 4.

4 Discussion

NGS analysis of tumors currently allows the determination of somatic DNA alterations indicative of a potential response to targeted therapy and is now part of clinical treatment

decisions. For high-grade serous ovarian tumors, olaparib was the first clinically recommended PARPi used with efficacy on platinum-sensitive tumors at relapse with inactive *BRCA1* or *BRCA2*. Germline variants of these two genes are widely documented through available databases. However, we and others observed a significant number of exclusive somatic variants frequently without annotation. In these later cases, bioinformatics analysis alone with the usual tools does not always identify a functional consequence, particularly for VUS involved in splice alteration. In the absence of molecular proof of *BRCA1/2* activity loss linked to a putative somatic splice variant, the patient cannot benefit from olaparib treatment.

To answer the functional consequence question, we used poor-quality RNA from FFPE tumors to evaluate their

△ Adis

Table 4 Theranostic impact of *BRCA1* variants

Samples	Variants	Therapeutic decisions	Patients outcomes
1	c.5194-2A>G	Eligible for PARPi	Platinum-sensitive recurrent ovarian cancer Progression after 3 months on PARPi (olaparib) maintenance treatment Chemotherapy combining gemcitabin and bevacizumab Deceased
2	c.4484+5G>A	Eligible for PARPi	Deceased weeks after oncologist request
3	c.5434C>G	Eligible for PARPi	Platinum-sensitive recurrent ovarian cancer Maintenance treatment by PARPi (olaparib) since 6 months
4	c.547+1G>A	Eligible for PARPi	Platinum-sensitive recurrent ovarian cancer Progression after 9 months on PARPi (olaparib) maintenance treatment Chemotherapy combining doxorubicin and carboplatin since 2 months
5	c.213-1G>A	Eligible for PARPi	Remission after first-line chemotherapy

PARPi poly (ADP-ribose) polymerase inhibitor

functional impact (Fig. 1a, b). The principle is to reverse transcribe RNA from fixed tissue extract. After cDNA synthesis, *BRCA1*-specific PCR products were size selected to find evidence of splicing alteration. Lower fragment sizes compared to cDNA from WT transcripts were cloned and sequenced to provide molecular evidence of splicing aberrations. From the *BRCA1* ovarian tumor somatic DNA sequences, five variants of the *BRCA1* gene were identified for further study: four variants located at intron–exon boundaries and one exonic variant were experimentally characterized for their potential splicing regulatory element affection. Three of the patient samples have already been classified and were used as positive controls to assess the validity of our protocol. We then analyzed the expression of the forms with aberrant splicing, to determine whether they really had a detrimental effect on transcript levels and thus protein production. However, our FFPE samples also contained WT cells due to tumor heterogeneity, so the levels of aberrant transcripts may have been underestimated.

Variant c.5194–2A>G was reported to be pathogenic in the databases used and in a previous study [19], and the in silico prediction tool favored a disturbance to splicing. Our experimental analysis of RNA extracted from FFPE tissue showed the skipping of exon 20 for this variant, confirming the expected deleterious character, and validated the method for poor-quality RNA (see the supplementary figure). The production of the transcript displaying exon 20 skipping was confirmed (Fig. 3a). Despite the high frequency of the variant (75.4%), the abundance of this aberrant form was one-third that of the full-length transcript.

Variant c.547+1G>A was only reported in the ClinVar database as “likely pathogenic”, and the in silico tool was in agreement for an alteration of splicing that we confirmed with exon 8 skipping (Fig. 2a–d). However, this aberrant form seems to be minor in terms of transcriptional expression (Fig. 3d).

Variant c.4484+5G>A was only reported in the ClinVar database as “uncertain significance” and the prediction tool

that we used predicted no impact on splicing. We experimentally showed exon skipping in tumor RNA, and the bioinformatic interpretation was in conflict with our results. However, this functional characterization of splicing alteration had a major theranostic impact because we were able to give access to the targeted therapy using PARPi. Moreover, this variant had a major transcriptional impact, because it was overexpressed with respect to the full-length form (Fig. 3b).

The intron–exon junction c.213–1G>A found in patient 5 was pathogenic in databases, and HSF evaluation predicted a splicing aberration already characterized in previous studies [19, 20]. However, this variant could not be confirmed by our functional testing as a consequence of poor RNA quality.

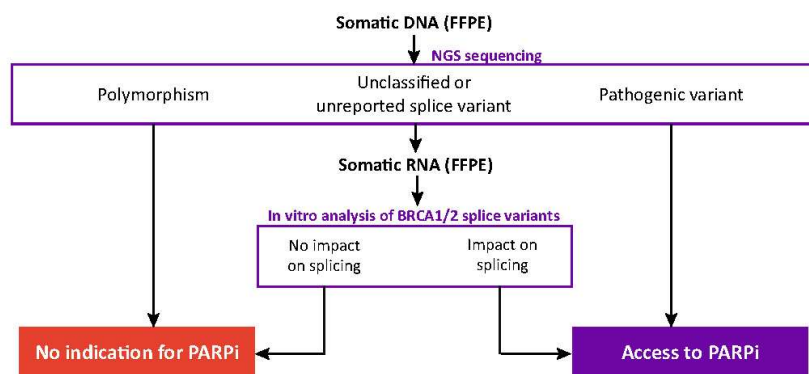
Concerning the exonic variant c.5434C>G from patient 3, our experimental analysis of mRNA from the FFPE sample confirmed that it induces a major splicing alteration, with exon 23 skipping, resulting in predicted protein termination within the second BRCT domain of *BRCA1*. The aberrant form was also a major event in this sample (Fig. 3c). This result confirms published findings [18, 19].

An analysis of patient outcomes (Table 4) revealed a correlation with the levels of the mutated transcripts. Patients 1 (c.5194–2A>G) and 4 (c.547+1G>A) both suffered relapses, after 3 and 9 months, respectively, of PARPi maintenance treatment. Moreover, their aberrant forms seemed to have a minor impact on *BRCA1* expression (Fig. 3a, d). By contrast, patient 3 (c.5434C>G), who presented a variant with a major impact on transcriptional expression (Fig. 3c), is still on PARPi maintenance treatment. Despite the small number of patients studied, this correlation is potentially interesting and merits a more extensive analysis.

First, we were able to demonstrate that mRNA from FFPE tumor samples could be used directly to produce biological evidence on the splicing alteration of *BRCA1* and *BRCA2* transcripts. As a positive control, we verified the deleterious effect of well-documented splice variants (c.5194–2A>G,

△ Adis

Fig. 4 Decision tree. *FFPE* formalin-fixed, paraffin-embedded, *NGS* next-generation sequencing, *PARPi* poly (ADP-ribose) polymerase inhibitor



c.547+1G>A, and c.213–1G>A) in mRNA from ovarian tumor tissue FFPE slices. Furthermore, the clinical utility has also been shown for VUS or unreported variants (c.4484+5G>A and c.5434C>G). For this purpose, our results clearly indicate that a functional analysis of non-pathogenic predicted splice variants by in silico tools is essential to give access to PARPi. Finally, and perhaps most importantly, our protocol is compatible with the delays in the management and implementation of the targeted therapy expected by clinicians in less than 2 months.

However, we must emphasize some limiting biological parameters: variation in the quantity and quality of transcript material extracted from the FFPE tumor sample, low allelic frequency variants that could be more difficult to isolate by the cloning technique, and long exons (mainly more than 250 bp) for which the FFPE RNA fragments will not allow visualization of the WT strand. In that event, if the splicing process is altered, we will only be able to see the shorter mutated strand. Consequently, in the case of normal splicing, we will not see any band for the patient after amplification. In this latter case, a positive control with gDNA from the cell line confirms that the absence of the amplicon is not due to PCR artifacts.

In a break with purely in silico prediction, we propose a rapid pretherapeutic functional analysis of *BRCA1/2* splice site variants of unknown significance at diagnosis with a delay of far less than 2 months. Moreover, our reported data emphasized the possible use of degraded mRNA from FFPE in clinical practice to confirm a potential splice event. The analysis of splice variants directly from FFPE samples has already been reported for variants of *BRCA1* and other tumor suppressors, demonstrating the feasibility of extending our method [21]. However, we chose to check the initial evaluation by cloning and Sanger sequencing. This last step was not done in the previous study, potentially preventing the characterization of somatic VUS with low frequencies due to tumor heterogeneity, as Sanger sequencing is not very sensitive. Finally, investigation of *BRCA1/2* putative splice

variants in our laboratory follows the decision tree shown in Fig. 4. This approach will allow more OC patients to benefit from this new targeted therapy on PARP and, by extension, could be applicable to any theranostic interpretation of putative splice variants, such as the *c-MET* gene in non-small cell lung carcinoma.

Compliance with Ethical Standards

Ethical Approval and Informed Consent The study procedures were approved by the ethics committee of the Angers medical university under ref. 2020-05.

Conflict of Interest LMC, AB, SF, JD, AP, VV and AM have no conflicts of interest to report.

Funding This work was supported in part by a grant from *Cancéropôle Grand Ouest*. Amandine Billaud is supported by a grant from *Institut de Cancérologie de l'Ouest*.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424.
2. Seidman JD, Horkayne-Szakaly I, Haiba M, Boice CR, Kurman RJ, Ronnett BM. The histologic type and stage distribution of ovarian carcinomas of surface epithelial origin. *Int J Gynecol Pathol*. 2004;23(1):41–4.
3. Ledermann JA. Front-line therapy of advanced ovarian cancer: new approaches. *Ann Oncol*. 2017;28(8):46–50.
4. Vetter MH, Hays JL. Use of targeted therapeutics in epithelial ovarian cancer: a review of current literature and future directions. *Clin Ther*. 2018;40(3):361–71.
5. Pujade-Lauraine E. New treatments in ovarian cancer. *Ann Oncol*. 2017;28(8):57–60.
6. Reinbolt RE, Hays JL. The role of PARP inhibitors in the treatment of gynecologic malignancies. *Front Oncol*. 2013;3:237.
7. Farmer H, McCabe N, Lord CJ, Tutt ANJ, Johnson DA, Richardson TB, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature*. 2005;434(7035):917–21.

△ Adis

8. Calabrese CR, Almassy R, Barton S, Batey MA, Calvert AH, Canan-Koch S, et al. Anticancer chemosensitization and radiosensitization by the novel poly(ADP-ribose) polymerase-1 inhibitor AG14361. *J Natl Cancer Inst.* 2004;96(1):56–67.
9. Chandrasekaran D, Manchanda R. Germline and somatic genetic testing in ovarian cancer patients. *BJOG.* 2018;125:1460.
10. de la Hoya M, Soukarieh O, López-Perolio I, Vega A, Walker LC, van Ierland Y, et al. Combined genetic and splicing analysis of BRCA1 c.[594-2A>C; 641A>G] highlights the relevance of naturally occurring in-frame transcripts for developing disease gene variant classification algorithms. *Hum Mol Genet.* 2016;25(11):2256–68.
11. Baralle D, Baralle M. Splicing in action: assessing disease causing sequence changes. *J Med Genet.* 2005;42(10):737–48.
12. Wappenschmidt B, Becker AA, Hauke J, Weber U, Engert S, Köhler J, et al. Analysis of 30 putative BRCA1 splicing mutations in hereditary breast and ovarian cancer families identifies exonic splice site mutations that escape in silico prediction. *PLoS One.* 2012;7(12). Disponible sur: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3519833/>.
13. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol.* 2011;29(1):24–6.
14. Bérout C, Letovsky SI, Braastad CD, Caputo SM, Beaudoux O, Bignon YJ, et al. BRCA Share: a collection of clinical BRCA gene variants. *Hum Mutat.* 2016;37(12):1318–28.
15. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* 2014;42:980–5.
16. Desmet F-O, Hamroun D, Lalande M, Colod-Bérout G, Claustres M, Bérout C. Human splicing finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 2009;37(9):e67.
17. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol.* 2004;11(2–3):377–94.
18. Gaildrat P, Krieger S, Théry J-C, Killian A, Rousselin A, Berthet P, et al. The BRCA1 c.5434C→G (p.Pro1812Ala) variant induces a deleterious exon 23 skipping by affecting exonic splicing regulatory elements. *J Med Genet.* 2010;47(6):398–403.
19. Findlay GM, Daza RM, Martin B, Zhang MD, Leith AP, Gasperini M, et al. Accurate classification of BRCA1 variants with saturation genome editing. *Nature.* 2018;562(7726):217–22.
20. Steffensen AY, Dandanell M, Jønson L, Ejlertsen B, Gerdes A-M, Nielsen FC, et al. Functional characterization of BRCA1 gene variants by mini-gene splicing assay. *Eur J Hum Genet.* 2014;22(12):1362–8.
21. Jansen AM, van der Klift HM, Roos MA, van Eendenburg JD, Tops CM, Wijnen JT, et al. RNA analysis of cancer predisposing genes in formalin-fixed paraffin-embedded tissue determines aberrant splicing. *Eur J Hum Genet.* 2018;26(8):1143–50.

DISCUSSION ARTICLE 2

La détection dans la tumeur de variations pathogènes des gènes *BRCA1* et *BRCA2*, qu'elles soient somatiques ou constitutionnelles, conditionne la prise en charge du patient et l'accès aux thérapies ciblées par iPARP. Les localisations concernées par ces indications se multiplient (ovaire, sein, pancréas, prostate) générant un nombre croissant d'analyses NGS entraînant inexorablement la mise en évidence d'une myriade de VSI au niveau de ces deux gènes. L'interprétation de ces VSI est donc devenu un enjeu majeur pour les laboratoires de génétique tumorale. Théoriquement, tous les VSI sont susceptibles d'impacter l'épissage, le type d'évènement d'épissage alternatif le plus fréquemment détecté dans les gènes *BRCA* étant le saut d'un ou de plusieurs exons.

A partir de l'ARNm de tissus tumoraux FFPE, 5 variants du gène *BRCA1* ont été étudiés. L'analyse par électrophorèse capillaire après transcription inverse puis amplification par PCR permet de visualiser les potentiels sauts d'exons ou rétentions d'introns. Un clonage suivi d'un séquençage de Sanger a permis de vérifier la présence de sauts d'exons chez 4 des 5 patients (c.5194-2A>G, c.5434C>G, c.547+1G>A et c.4484+5G>A). Pour le 5^{ème}, le matériel extrait était de qualité insuffisante pour être analysé par cette technique. A noter que, depuis la publication de cet article, l'interprétation dans ClinVar du variant c.547+1G>A est passé de « *Likely pathogenic* » à « *Pathogenic / Likely pathogenic* » et celle du variant c.4484+5G>A de « *Uncertain significance* » à « *Conflicting interpretations of pathogenicity* » suite à une soumission de ce variant l'évaluant en classe IV.³⁷⁰ Ces modifications vont dans le sens de nos résultats. Le principal avantage de cette approche d'évaluation de VSI est qu'elle peut permettre de mettre en évidence une anomalie d'épissage dans un temps compatible avec la mise en place d'un traitement par iPARP.

Pour analyser pleinement l'impact d'une variation splicéogénique, il est essentiel d'évaluer la contribution de l'allèle muté dans la production respective des transcrits aberrants, d'une part, et des transcrits normalement épissés, d'autre part. En effet, les modifications apportées aux pré-ARNm telles que des sauts d'exon peuvent être à l'origine d'un décalage du cadre de lecture responsable de l'introduction d'un codon stop prématuré dans les molécules d'ARNm mature. Ces transcrits porteurs d'un codon stop prématuré peuvent être dégradés par le système de surveillance NMD évitant ainsi la production de certaines protéines tronquées susceptibles d'inhiber la fonction de la protéine pleine longueur.³⁷¹ Dans ce travail, l'analyse du niveau d'expression des transcrits mutés a été réalisée par RT-qPCR. De manière intéressante, cette analyse s'est montrée corrélée au devenir des patients traités par iPARP (Figure 3 et Table 4) : le patient dont l'expression du transcrit muté était supérieure à celle du transcrit non muté a montré la réponse la plus longue aux iPARP. Cette corrélation entre impact transcriptionnel et réponse au traitement

serait à étudier dans une cohorte plus importante pour comprendre la plus ou moins grande sensibilité des patients présentant pourtant tous des variants pathogéniques.

Cependant, cette approche présente certaines limites :

- Si aucune anomalie n'est détectée sur l'ARN tumoral, cette technique ne permet en aucun cas d'exclure une anomalie d'épissage.
- Il faut connaître l'exon susceptible d'être concerné par l'anomalie d'épissage, car le design des amorces dépend de cette information. S'il y a un saut d'un autre exon, il ne sera pas mis en évidence.
- Un saut d'exon ou la perte d'un fragment d'exon n'est pas systématiquement associé à une perte de fonction de la protéine. Si la majorité des transcrits alternatifs des gènes *BRCA* induisent un décalage du cadre de lecture conduisant à l'introduction d'un codon stop prématuré potentiellement cible du NMD ou sont non-codants à la suite de l'élimination du codon d'initiation de la traduction, certains transcrits pourraient générer des isoformes fonctionnelles. C'est notamment le cas des transcrits présentant des délétions en phase, des modifications des régions terminales non codantes ou encore des codons stops prématurés non cibles du NMD car localisés au niveau du dernier exon. Certains travaux ont d'ores et déjà démontré le caractère fonctionnel, au moins en partie, de quelques isoformes dont notamment celles résultant du saut de l'exon 9 et 10 de *BRCA1*, du saut des exons 4 à 7 de *BRCA2* ou encore, du saut de l'exon 12 de *BRCA2*.^{350,372-374} L'interprétation de notre approche comprend donc deux parties : dans un premier temps la mise en évidence d'une anomalie d'épissage dans l'ARNm tumoral du patient, et dans un 2^{ème} temps la recherche dans la littérature d'anomalies avec des conséquences protéiques similaires (par exemple variation entraînant le même saut d'exon).

La méthode de caractérisation fonctionnelle des variants affectant l'épissage du gène *BRCA1* présentée dans cet article peut être étendue. Elle est déjà mise en application dans notre laboratoire pour identifier les sauts d'exon 14 du gène *MET* dans l'ARNm tumoral de patient souffrant de CBNPC puisque ce type de variant donne accès aux thérapies ciblées par inhibiteurs de MET. A ce jour, des sauts d'exon 14 de *MET* ont été identifiés pour plus d'une trentaine de VSI ou variants non répertoriés de *MET* détectés par NGS au laboratoire. Pour plus d'efficacité et de rapidité dans la prise en charge, les acides nucléiques totaux sont extraits en première intention dans les CBNPC pour permettre les analyses secondaires sur ARN.

Quels que soient les gènes, les potentiels variants d'épissage sont nombreux, la fonctionnalité et la pertinence biologiques de ces anomalies sont encore très peu appréhendées et nécessitent de mettre en place des analyses fonctionnelles plus robustes parmi lesquelles l'édition génomique est aujourd'hui la plus prometteuse.

**ARTICLE 3 : FUNCTIONAL PRE-THERAPEUTIC EVALUATION BY GENOME EDITING
OF VARIANTS OF UNCERTAIN SIGNIFICANCE OF ESSENTIAL TUMOR
SUPPRESSOR GENES**

Amandine Billaud, Louise-Marie Chevalier, Paule Augereau, Jean-Sébastien Frenel,
Christophe Passot, Mario Campone and Alain Morel

Publié le 9 novembre 2021 – *Genome Medicine*

RESEARCH

Open Access

Functional pre-therapeutic evaluation by genome editing of variants of uncertain significance of essential tumor suppressor genes



Amandine Billaud^{1,2}, Louise-Marie Chevalier^{1,2}, Paule Augereau², Jean-Sebastien Frenel^{2,3}, Christophe Passot², Mario Campone^{2,3} and Alain Morel^{1,2*} 

Abstract

Background: Targeted therapies in oncology are promising but variants of uncertain significance (VUS) limit their use for clinical management and necessitate functional testing in vitro. Using *BRCA1* and *BRCA2* variants, which have consequences on PARP inhibitor sensitivity, and *POLE* variants, potential biomarkers of immunotherapy response, we developed a rapid functional assay based on CRISPR-Cas9 genome editing to determine the functional consequences of these variants having potentially direct implications on patients' access to targeted therapies.

Methods: We first evaluated the functional impact of 26 *BRCA1* and 7 *BRCA2* variants by editing and comparing NGS results between the variant of interest and a silent control variant. Ten of these variants had already been classified as benign or pathogenic and were used as controls. Finally, we extended this method to the characterization of *POLE* VUS.

Results: For the 23 variants that were unclassified or for which conflicting interpretations had been reported, 15 were classified as functionally normal and 6 as functionally abnormal. Another two variants were found to have intermediate consequences, both with potential impacts on splicing. We then compared these scores to the patients' responses to PARP inhibitors when possible. Finally, to prove the application of our method to the classification of variants from other tumor suppressor genes, we exemplified with three *POLE* VUS. Among them, two were classified with an intermediate functional impact and one was functionally abnormal. Eventually, four *POLE* variants previously classified in databases were also evaluated. However, we found evidence of a discordance with the classification, variant p.Leu424Val being found here functionally normal.

* Correspondence: alain.morel@ico.unicancer.fr

¹Université d'Angers, Inserm, CRCINA, SFR ICAT, F-49000 Angers, France

²Institut de Cancérologie de l'Ouest Nantes-Angers, F-49000 Angers, France

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Conclusions: Our new rapid functional assay can be used to characterize the functional implication of *BRCA1* and *BRCA2* variants, giving patients whose variants were evaluated as functionally abnormal access to PARP inhibitor treatment. Retrospective analysis of patients' responses to PARP inhibitors, where accessible, was consistent with our functional score evaluation and confirmed the accuracy of our protocol. This method could potentially be extended to the classification of VUS from all essential tumor suppressor genes and can be performed within a timeframe compatible with clinical applications, thereby having a direct theranostic impact.

Keywords: Variants of uncertain significance, Genome editing, Functional testing, *BRCA1*, *BRCA2*, *POLE*, Cancer, Theranostic

Background

Over the last two decades, the therapeutic options available in oncology have evolved towards therapies targeted based on tumoral genetic information [1]. These new treatments improve patient outcomes and have fewer adverse effects. However, the functional significance of many variants of targetable genes remains unknown. In the ClinVar [2] database, 237 934 variants of uncertain significance (VUS) are registered for the total set of genes considered (32 951 genes represented). Most are missense variants and potential splice site variants. With recent technical improvements and the development of whole-exome and whole-genome sequencing, this number of VUS is likely to rise still further in the coming years [3], with one in three variants classified as VUS overall, and 80% located on tumor suppressor genes [4]. As VUS affect treatment options and patient management, this trend highlights the need for a method of functional testing.

Variants of the *BRCA1* and *BRCA2* genes can be used to illustrate this crucial problem. These two tumor suppressor genes have many roles, mostly in genome protection via the homologous recombination (HR) pathway [5–7]. Inheritable mutations of *BRCA1/2* increase the risk of breast cancer (50–80%) and ovarian cancer (40–60%) [8, 9], and have also been implicated in prostate and pancreatic cancers. With the development of PARP inhibitors, pathogenic variants of these genes are now biomarkers of response to these treatments [10–12]. Their multiple interaction domains and protein partners account for the tremendous diversity of variants found in tumors. More than 2660 *BRCA1* and 4840 *BRCA2* VUS are registered in ClinVar, which means that respectively 35.7% and 45.7% of the total currently known *BRCA1* and *BRCA2* variants are VUS. However, databases, such as ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>), BRCA exchange (<https://brcaexchange.org/>), and UMD [13] (Universal Mutation Database, <http://www.umd.be/BRCA1/>), mostly contain germline variants. Somatic variants, which may be detected in only one or two individuals, also exist and many such variants remain unreported. Functional testing in vitro is currently based on transcriptional activation [14, 15], HR activity [16, 17], and splicing [18–20]. However, such

tests are not necessarily compatible with clinical management in terms of the time taken to obtain results to guide treatment.

Thus, CRISPR-Cas9 genome editing is a promising tool for meeting the challenge of VUS classification [21–23]. We first tested genome editing on diploid cells, but the efficiency of editing was low and clonal expansion, which is time-consuming, was required. Moreover, most of the clones were heterozygous, with highly variable sensitivity to PARP inhibitors depending on proliferation rates. We therefore used haploid HAP1 cells. Here, we adapted the concept developed in *BRCA1* variant characterization by saturation genome editing [24, 25] to the study of specific variants detected in our clinical testing laboratory. By comparing editing frequencies, by next-generation sequencing (NGS), between a variant of interest (testing variant) and a silent variant classified as benign, and therefore functionally normal (control variant), we were able to calculate a functional score and to evaluate the functional consequences of 23 *BRCA1/2* VUS and 10 *BRCA1/2* variants already classified as benign or pathogenic. Among those, 23 were characterized in our laboratory during somatic mutation testing in patients. In this retrospective analysis, where patients had received PARP inhibitors, their responses were compared to the functional score calculated for their variant. We finally extended our method to the evaluation of seven variants of *POLE*, another tumor suppressor gene biomarker for immunotherapy administration, demonstrating the potential utility of this approach for characterization of variants of others essential tumor suppressor genes. The essentiality of the *BRCA1*, *BRCA2*, and *POLE* genes in our haploid model [25, 26] was a key feature, making it possible to evaluate function rapidly, within less than 3 weeks, compatible with a direct benefit of the patient carrying the variant.

Methods

Population

Tumor suppressor gene (*BRCA1*, *BRCA2*, and *POLE*) somatic variants included in this study were selected solely based on their status of variants of uncertain significance or unclassified variants; well-known variants were also analyzed as controls. All of them were

characterized by next-generation sequencing at the *Institut de Cancérologie de l'Ouest* (ICO, Angers, France). Consequently, 25 patients suffering from high-grade serous ovarian cancer were enrolled in this study, originating from the Institut de Cancérologie de l'Ouest and the Centre Hospitalier du Mans (Le Mans, France). The functional consequences of these selected variants were evaluated by genome editing in haploid cells and the results were compared to the patients' response to PARP inhibitors where available. These patients were treated with PARP inhibitors at the ICO, allowing access to their treatment results. This retrospective study was approved by the local ethics committee of Angers medical university (France) under the reference 2020-05.

HAP1 cell culture

Wild-type haploid HAP1 cells were purchased from Horizon Discovery and cultured in Isocove's modified Dulbecco's media (IMDM) containing L-glutamine and 25 mM HEPES (Corning), supplemented with 10% fetal calf serum (Eurobio). Cells were grown at 37 °C, under an atmosphere containing 5% CO₂ and were passaged before confluence, to prevent reversion to the diploid state. Haploidy of HAP1 cells was confirmed by measure of DNA content via coloration with propidium iodide (PI) dye following Vindelov method [27] and cytometry analysis before their use.

Genetically engineered HAP1 cells

Polyclonal LIG4 knock-out cells were generated with CRISPR-Cas9 technology [28]. Briefly, a guide RNA (gRNA) was first designed to target the second exon with an *AflIII* restriction site three nucleotides upstream from the PAM sequence. An Alt-R CRISPR-Cas9 crRNA (IDT DNA) (5'-CAATTACACAGTACGTGTCT-3') and an Alt-R CRISPR-Cas9 tracrRNA with an ATTO550 fluorescent dye (IDT DNA) were complexed at a final concentration of 1 μM with 6 pmol of Alt-R S.p. Hifi Cas9 Nuclease V3 (IDT DNA) in presence of Lipofectamine CRISPR-MAX Cas9 Transfection Reagent (Thermo Fisher Scientific). The mixture was incubated for 20 min, and reverse transfection was then performed by adding RNA-Cas9 ribonucleoprotein complexes to 1.6 × 10⁵ cells. Four hours after transfection, cells were sorted by FACS based on ATTO550 fluorescence. Only 20% of cells with the highest level of fluorescence were retained and used to seed with IMDM supplemented with 1% penicillin-streptomycin (Gibco). The cells were incubated for 5 days and then subjected to limiting dilution. About 20 clones were amplified for DNA extraction with Chelex 100 Resin (Biorad). We used 10 μL of these DNA extract for PCR amplification (forward primer: 5'-CTGGAGAACAGAATTGCAGA-3'; reverse primer: 5'-TAGCAATCATATTCACGGGC-3') followed by digestion with the *AflIII* restriction enzyme

(New England Biolabs) for 1 h at 37 °C. The mixture was then incubated for 20 min at 80 °C for enzyme inactivation. The clones were screened by following their migration in a 2% agarose gel on electrophoresis. Clones that had undergone genomic editing and had lost the restriction site were identified by Sanger sequencing on an ABI 3130 Genetic Analyzer (Thermo Fisher Scientific) with the BigDye Terminator v1.1 Cycle Sequencing Kit (Thermo Fisher Scientific). Results were visualized and analyzed with Sequencing Analysis 5.3.1 software (Thermo Fisher Scientific). The exact same protocol was followed to generate the polyclonal XRCC4 KO HAP1 cell line (RNAg and primer sequences are listed in Additional file 1: Table S1).

VUS selection and gRNA design

First, ten variants of *BRCA1*, three of *BRCA2*, and two of *POLE* were selected for study based on their status as variants of uncertain significance or unclassified variants; 10 well-known variants of *BRCA1* (p.Tyr179Cys, p.Cys197=, p.Tyr422X, p.Gln1604X, p.Gln1604=, and p.Pro1812Ala) and *BRCA2* (p.Asp935Asn, p.Ser1882X, p.Val2728Ile and p.Gln2829Arg) were also analyzed as controls. All these 25 variants were characterized by NGS in our laboratory during somatic mutation testing in patients. To increase the number of characterized VUS, we also selected 10 variants of *BRCA1* (p.Ile31Asn, p.Glu149Ala, p.Val191Asp, p.Gln210=, p.Gly462Arg, p.Arg979Cys, p.Gly1201Ser, p.Thr1394Ile, p.Ala1752Pro, p.Gly1770Val) and a *POLE* variant (p.Arg1826Trp) for which conflicting interpretations had been reported in the databases (ClinVar [2], OncoKB [29]). Four variants of *POLE* (p.Ala31Ser, p.Pro286Ser, p.Leu424Val, p.Phe695Ile) were also selected in databases and used as controls (ClinVar, OncoKB). Alt-R CRISPR-Cas9 crRNA (IDT DNA) were designed with the Alt-R Custom Cas9 crRNA design tool (IDT DNA) (https://eu.idtdna.com/site/order/designtool/index/CRISPR_CUSTOM). The PAM sequence had to be adjacent to the variant to facilitate the editing of KO LIG4 HAP1 cells [30]. We also selected the gRNA based on the possibility of inserting a silent variant into the PAM sequence or the 3 to 5 nucleotides immediately upstream [30]. This increases editing efficiency and will be used as a control in subsequent experiments (reference variant). For each variant, we designed two Ultramer DNA Oligos (IDT DNA) of about 84 nt. The first contained the patient's variant (testing variant) and the second contained a silent variant (control variant), already reported to be benign if possible, and therefore functionally normal. Both contained the silent reference variant mentioned above. All the gRNA and

DNA oligonucleotides designed are reported in Additional file 1: Table S1.

Transfection of LIG4 KO HAP1 cells

For each variant, two transfections were performed simultaneously, both with the same gRNA but with different DNA oligomers (the VUS to be classified in one transfection or the control variant in the other). We used the protocol described earlier for knocking-out *LIG4* in HAP1 cells, but with 2 nmol of DNA oligonucleotides added before the Lipofectamine CRISPRMAX Cas9 Transfection Reagent. A cell suspension containing 400,000 cells/mL in IMDM supplemented with 10% FBS was then prepared, and Alt-R HDR Enhancer (IDT DNA) was added to a final concentration of 2 nM. Reverse transfection was then performed. On day 1 post-transfection, the medium was replaced with fresh IMDM supplemented with 10% FBS. On day 4 to 5, depending on the degree of confluence, the cells were released by trypsin treatment and used to seed 6-cm-diameter plates. Two days after plating, a second transfection was performed with the same protocol for both types of transfection, to enrich the cell preparation in edited cells. The cells were then incubated for a further 4 to 5 days before DNA extraction.

DNA extraction and NGS sequencing

All gDNA were extracted from edited cells with the Maxwell 16 Blood DNA Purification Kit (Promega) and quantified using a Qubit (Thermo Fisher Scientific) and the Quantifluor dsDNA System Kit (Promega). We then used 20 ng of the extracted DNA to generate the NGS library. The libraries were prepared with the Oncomine BRCA Assay Manual Kit (Thermo Fisher Scientific) or a custom assay Ampliseq for *POLE* (Thermo Fisher Scientific), allowing amplification of the entire *BRCA1* and *BRCA2* or *POLE* coding regions and noncoding putative splice boundaries. Then, 20 ng of DNA per sample was needed and barcoded, the libraries were subjected to clonal amplification by PCR emulsion with an Ion Chef System (Thermo Fisher Scientific) according to the manufacturer's instructions. The prepared libraries were then sequenced on an Ion Torrent S5 Sequencer with the Ion 520 and 530 Chef Kit (Thermo Fisher Scientific).

Variants of interest were visualized with Integrative Genomics Viewer (IGV) [31]. Human Genome Variation Society (HGVS)-approved guidelines (<http://www.hgvs.org/mutnomen/>) were used. The variants found by NGS were researched in the UMD-BRCA1 or UMD-BRCA2 databases [13], the ClinVar database, the BRCA exchange database, and the BRCA Mutation Database (http://www.arup.utah.edu/database/BRCA/Home/BRCA1_landing.php) to assess their classification and functional impact where available.

Sequencing analysis and VUS functional score evaluation

Following NGS sequencing, insertions or deletions located around the expected cleavage site, in the eight nucleotides centered on the PAM sequence or the seven nucleotides centered on the VUS, were also counted. Indel frequencies were then calculated by dividing the total amount of indels by the total number of reads. For the evaluation of SNV coverage, the ratio of the total numbers of reads for the VUS evaluated and the control SNV was calculated. Finally, functional scores for all the variants studied were calculated by comparing the sequence frequencies of all the inserted variants (VUS of interest, silent control SNV, and silent reference SNV) and the results contained in the available databases (UMD database, ClinVar, BRCA exchange).

The following formula was used:

$$\text{Functional score} = \frac{1}{2} * ((\log_2((f_{\text{mut}} * f_{\text{PAM}_{\text{mut}}}) / (f_{\text{sil}} * f_{\text{PAM}_{\text{sil}}})) + (\log_2((f_{\text{mut}} * f_{\text{PAM}_{\text{sil}}}) / (f_{\text{sil}} * f_{\text{PAM}_{\text{mut}}})))))$$

With f_{mut} for variant frequency of the VUS of interest, $f_{\text{PAM}_{\text{mut}}}$ for PAM variant frequency (or silent reference SNV) corresponding to the cells edited with the testing variant, f_{sil} for variant frequency of the silent control SNV, and $f_{\text{PAM}_{\text{sil}}}$ for variant frequency of the silent reference SNV corresponding to the control condition. Read covers and indel frequencies must be similar in the control condition and the tested condition. All the variants measured for a given variant must be localized on a same read.

Statistics

Linear regression analyses were performed to compare insertions and deletions frequencies between the control condition (silent control variant) and the studied variant (testing variant). Mann-Whitney tests were conducted to assess the potential statistical significance of the differences between the functional scores of the different classifications of variants (benign, intermediate and deleterious). Thresholds delineating the intermediate classification were calculated following the evaluation of the standard deviation of the functional scores among the benign and deleterious variants. All statistical analyses were performed with GraphPad Prism analysis software.

Results

A comparison of editing frequencies between *BRCA1/2* variants and silent control SNV can be used for functional classification

In oncology, the functional testing of variants of uncertain significance has become a major issue in the context of access to targeted therapies. We used the example of *BRCA1/2* variants to develop a method based on the CRISPR-Cas9 method. First, we promoted DNA repair by homologous recombination after the endonuclease action of Cas9 by inhibiting the non-homologous end-

joining pathway (NHEJ) [32] by knock-out of the *LIG4* gene. The gRNA targeting this gene was selected according to its proximity to the *AflIII* restriction site, which is located at the Cas9 double-stranded cleavage site (Additional file 1: Fig S1A and S1B). The clones that had lost the restriction site were assumed to have undergone editing and were subjected to Sanger sequencing (Additional file 1: Fig S1C-E). All six of the clones sequenced had been edited, but only three also presented a frame-shift (clones 5, 8, and 17). These three clones were pooled together to constitute the polyclonal *LIG4* knock-out HAP1 cell line, to prevent side effects and putative off-target effects associated with clonal selection.

In HAP1 cells, *BRCA1* and *BRCA2* are essential genes [25, 26]. Genomic editing to create a functionally abnormal variant of these genes thus leads to cell death, facilitating the screening of edited cells. Moreover, edited cells with insertions or deletions instead of the variant of interest also die, due to the essential nature of the gene concerned. We checked that the absence of a variant following NGS sequencing was due to the functionality of the variant rather than a problem linked to genomic editing, by simultaneously performing a second transfection, with the same gRNA, but the insertion of a silent variant already classified as benign in databases, and therefore functionally normal, where possible (Fig. 1A). Thus, for each variant, we designed one gRNA and two oligonucleotides (Fig. 1B, C), one carrying the variant of interest, or testing variant, and the other carrying a silent control variant. We also added a second silent reference variant, also classified as benign where possible, located in the corresponding PAM motif or 3 nt upstream from the PAM motif. This second modification had two purposes. First, its insertion prevented from re-editing blocking gRNA association [30]. This second variant was present in both sets of conditions and was therefore also used as reference for comparisons. Following NGS sequencing, variant frequencies were compared and functional score was calculated (Fig. 1B, C). In our examples, the p.Pro1812Ala variant was classified as pathogenic in UMD and BRCA exchange databases and the p.Tyr179Cys variant as benign in UMD, ClinVar, and BRCA exchange databases. When the p.Tyr179Cys variant was analyzed, the frequency of the testing variants was similar to that of the control variant (Fig. 1C). On the contrary, for the p.Pro1812Ala variant, the frequency of the testing variant was considerably lower compared to the control variant (Fig. 1B) due to the essentiality of *BRCA1* gene.

We then validated our method testing 10 variants of *BRCA1* and *BRCA2* already classified as benign or pathogenic in databases. We evaluated the indel frequency to estimate the editing efficiency in the “testing variant” and “control variant” conditions compared for

each variant. Indeed, this indel frequency was identical for both conditions when analyzed for the 8 nt surrounding the PAM sequence (Fig. 1D, $R^2 = 0.8853$) but not for the 7 nt surrounding the edited variant (patient or silent control variant) (Additional file 1: Fig S2, $R^2 = 0.0258$). These results confirmed published findings; the Cas9 protein cleaves the DNA 3 nt upstream from the PAM sequence [33]. Moreover, the observed linear regression made it possible to evaluate genome editing efficiency and to compare the two conditions with the same gRNA. Following NGS sequencing, the coverages of the variant of interest (patient or silent control) and the reference control variant were also checked and shown to be similar in the two conditions (Fig. 1E). Finally, functional scores were calculated from sequencing frequencies (Fig. 2C and Table 1). The significant differences observed between benign and pathogenic variants validated the use of this method to value the function of variants of uncertain significance for the *BRCA1* and *BRCA2* genes. Functional score thresholds dividing variants into a functionally abnormal / intermediate / functionally normal classification were established from the analysis of these already classified variants by using the calculated standard deviations.

Assessment of validity of functional data according to guidelines

To assess the strength of evidence provided by our assay, we followed the four steps of the guidelines of Brnich et al. [34]. First, the disease mechanism seems reasonably well understood, since a loss of function of *BRCA1* and *BRCA2* genes is causal for homolog recombination deficiency [35]. Second, the used model seems appropriate since (i) the physiologic context seems reasonably respected in an haploid genome where the introduction of one variant is equivalent to the presence of the variant in a homozygous state in a diploid genome and (ii) the molecular consequence is very likely to be imputable to the variant itself since it is introduced by CRISPR-Cas9 genome editing and therefore no modification of regulatory sequences is expected. Third, the validity of the assay was evaluated with 10 controls (5 functionally normal and 5 functionally abnormal) and the OddsPath estimated with the table provided by Brnich et al. (Additional file 1: Table S2).

Functional characterization of *BRCA1/2* variants of unknown significance

The fourth step of the guidelines, which is the application of evidence to individual variants, could then be performed. The strength of evidence was moderate for both PS3 (pathogenic) and BS3 (benign) categories with OddsPath > 4.3 and < 0.23 respectively. Thirteen variants of *BRCA1* and *BRCA2* were initially identified in our

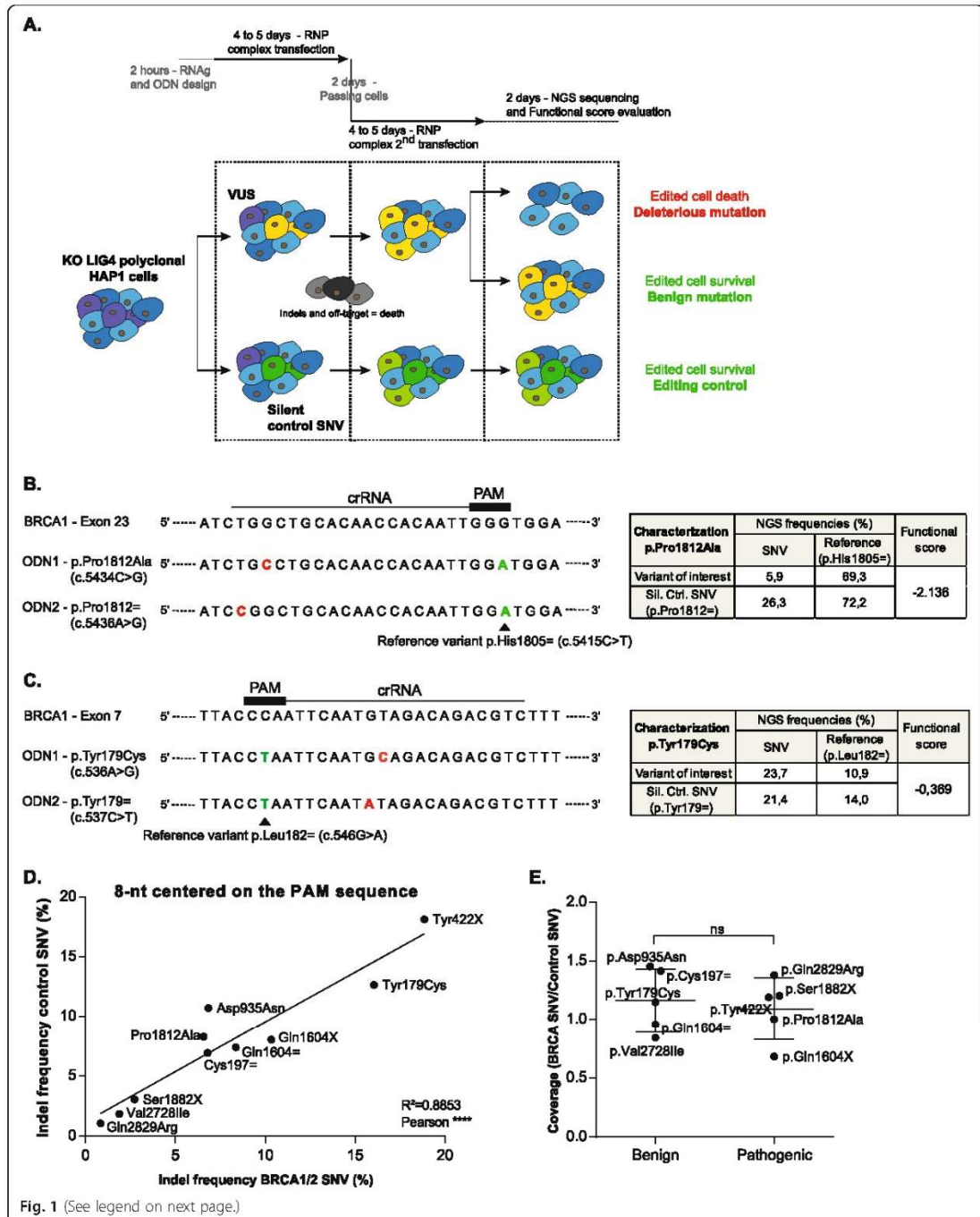


Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 Comparison of editing frequencies between *BRCA1/2* variants and silent control SNVs for functional classification. **A** Experimental process used in this analysis for evaluation of the functional impact of *BRCA1/2* and *POLE* SNVs. **B** Examples of the gRNA and the two corresponding oligonucleotides designed for the p.Pro1812Ala variant editing, pathogenic mutation of the *BRCA1* gene. On this schematic representation, silent control and patient mutations of interest are colored in red, whereas silent reference variants are in green. NGS sequencing results and functional scores for this mutation are associated. **C** Same thing with the p.Tyr179Cys variant, benign variant of the *BRCA1* gene. **D** Insertion and deletion frequencies determined following the NGS sequencing of *BRCA1/2* variants of interest, compared with those for the corresponding silent control. These means include the 8 nt surrounding the PAM sequence. **E** Analysis of NGS sequencing coverage with the following formula: Mutation of interest (Mutation+Reference control)/Silent control (Mutation+Reference control). The results for benign *BRCA1/2* variants are compared with those obtained for pathogenic mutations (Mann-Whitney tests, $p = 0.8016$)

laboratory and considered as VUS based on an absence of annotation concerning their function. We then also studied other 10 variants of *BRCA1* previously classified as VUS or not previously reported in the databases (ClinVar, BRCA exchange, UMD) (Table 1). References mentioning these variants and their potential functional impact are listed in Table 1 where available [20, 25, 36–52]. These 23 VUS and the 10 other variants we used for validating our system affected different domains of the proteins and were distributed along the entire length of these genes (Fig. 2A, B). All reference silent variants, silent control variants, and their classifications in databases are listed in Additional file 1: Table S3. For each variant, we checked the coverage and indel frequencies (Additional file 1: Fig. S2B and S2C, $R^2 = 0.9328$) for the testing variant and the control variant. All functional scores were then calculated from NGS frequencies (Fig. 2C and Table 1), and the variants were classified according to their impact on cell survival and, consequently, their functional impact in our model. Most of the 23 VUS analyzed were functionally normal (about 65.2%), but six were found to have an impact on edited cell survival (p.Gln210=, p.Leu291X, p.Glu1586X, p.Ala1752Thr, p.Ala1752Pro, p.Gly1770Val). Another two variants (p.Leu1080= and c.5194-2A>G) had intermediate functional classification.

Patients' responses to PARP inhibitor administration and comparison with the calculated functional scores

The next step was to confront our functional score to clinical data and the responses to PARP inhibitors of patients presenting alterations of *BRCA1/2*. The study we conducted was retrospective; hence, only 6 of the patients concerned received a PARP inhibitor treatment thus far. The mean age of these patients was 60.5 years and all of them suffered from high-grade serous ovarian cancer with different variants of *BRCA1* and *BRCA2*. The characteristics of their tumors are summarized in Table 2. Their responses to PARP inhibitors ranged from 3 months before resistance to 9 months and continuing. We then compared these data to the functional scores we had calculated for their variants. Five had, indeed, been classified as functionally abnormal and one with intermediate

functional score (Tables 1 and 2). The lowest response to PARP inhibitor, from patient 4, was associated with this intermediate functional score.

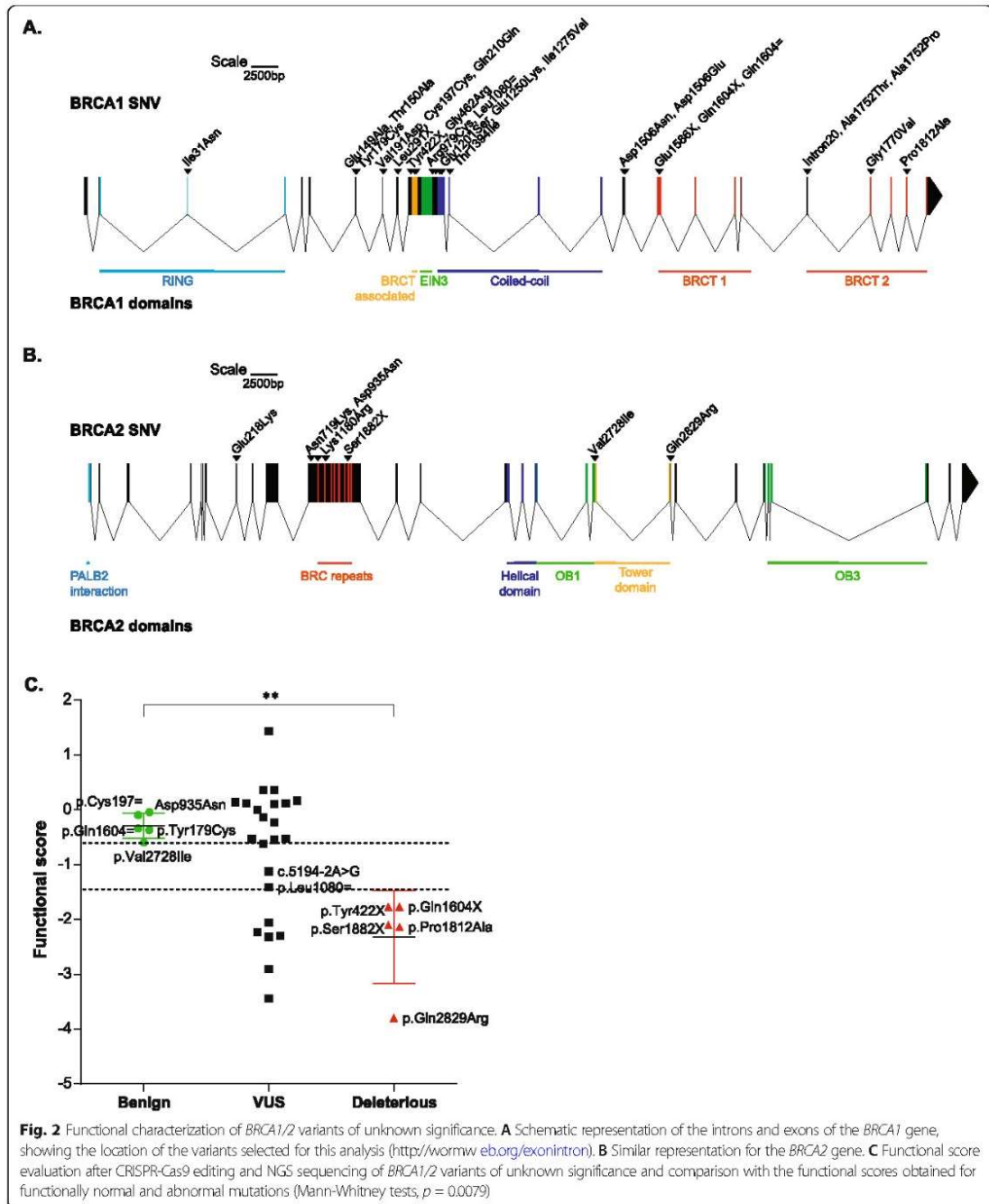
Extension of the experimental process to the functional evaluation of *POLE* variants

We then extended our protocol to the characterization of VUS from other tumor suppressor genes that were also essential in our model. We chose to study variants of the *POLE* gene because of potential interest of their functional impact for determining access to immunotherapy. We therefore selected seven *POLE* variants from databases, which included two classified as benign and two classified as pathogenic (Fig. 3B). The same protocol was followed to evaluate their functional scores, and the results obtained were compared with published findings (Fig. 3A, B). The p.Leu424Val variant, which has been classified as pathogenic in databases, was surprisingly found to have a functional score of -0.031 rather associated with a functionally normal variant according to our experimental protocol and the cut-off we used.

Discussion

Our evaluation of the functional scores of 33 variants of *BRCA1/2*, including 23 unreported or of uncertain significance, allows the classification of 20 of them as functionally normal, six as functionally abnormal (p.Gln210=, p.Leu291X, p.Glu1586X, p.Ala1752Thr, p.Ala1752Pro, p.Gly1770Val) and two as intermediate (p.Leu1080= and c.5194-2A>G) (Table 1). These results were consistent with databases (ClinVar, BRCA exchange, UMD), with previous functional studies (Table 1), where available, and with the published saturation genome editing study of the RING and BRCT domains of *BRCA1* (p.Ile31Asn, p.Ala1752Thr, p.Ala1752Pro, p.Gly1770Val and p.Pro1812Ala) [25]. Indeed, we evaluated, as expected following their publication, variant p.Ile31Asn as functionally normal whereas the four others were functionally abnormal.

Four of the six variants we classified as functionally abnormal were located in the BRCT domain of *BRCA1* (Fig. 2A), responsible of its interaction with various phosphorylated proteins such as Abraxas, BRIP1, or CtIP [6, 53]; the other two were previously unreported nonsense variants. The results were more surprising for the



p.Gln210 = variant of *BRCA1*, which was also found functionally abnormal. However, the p.Gln210 = was described in databases as a variant possibly creating or

strengthen a splice site of *BRCA1*, which is compatible with our results suggesting it is pathogenic. One of the two intermediate variants, c.5194-2A>G, has already

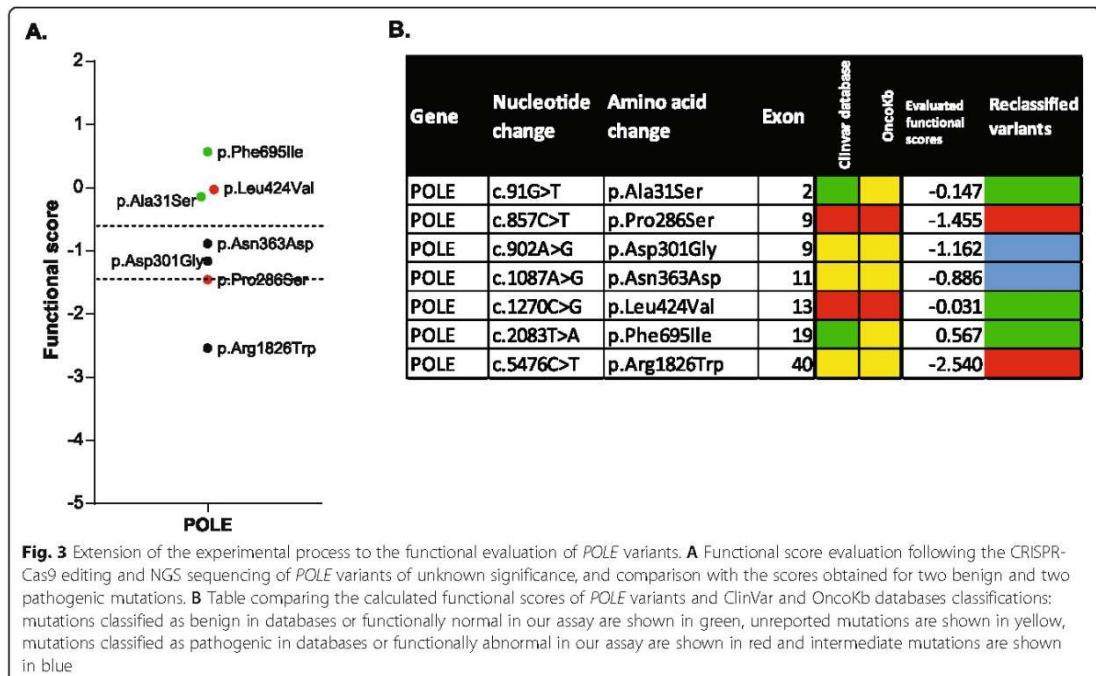
Table 1 Functional scores for the evaluated *BRCA1/2* variants and comparison with database annotations

Gene	Nucleotide change	Amino acid change	Exon	UMD database	ClinVar database	BRCA exchange	Findlay et al. ²⁵ Function scores	Found in our lab	Evaluated functional scores	Functional consequence	Ref.
<i>BRCA1</i>	c.92T>A	p.Ile31Asn	3	+/-	?	?	0.82822 3479		-0.624	Normal	25
<i>BRCA1</i>	c.446A>C	p.Glu149Ala	3	+/-	+/-	+/-			0.361	Normal	
<i>BRCA1</i>	c.448A>G	p.Thr150Ala	8	?	?	?		X	-0.138	Normal	
<i>BRCA1</i>	c.536A>G	p.Tyr179Cys	8	+	+	+		X	-0.369	Normal	54
<i>BRCA1</i>	c.572T>A	p.Val191Asp	8	+/-	+/-	?			-0.229	Normal	55
<i>BRCA1</i>	c.591C>T	p.Cys197=	8	+	+	+		X	-0.096	Normal	56
<i>BRCA1</i>	c.630A>G	p.Gln210=	10	+/-	+/-	+/-			-2.055	Abnormal	
<i>BRCA1</i>	c.872T>G	p.Leu291X	11	?	?	?		X	-2.298	Abnormal	
<i>BRCA1</i>	c.1266T>G	p.Tyr422X	11	-	-	-		X	-1.771	Abnormal	57
<i>BRCA1</i>	c.1384G>A	p.Gly462Arg	11	+/-	+/-	+/-			-0.527	Normal	58
<i>BRCA1</i>	c.2935C>T	p.Arg979Cys	11	+/-	+/-	+			0.122	Normal	
<i>BRCA1</i>	c.3238T>C	p.Leu1080=	11	+/-	+	+		X	-1.415	Intermediate	
<i>BRCA1</i>	c.3601G>A	p.Gly1201Ser	11	+/-	+	+			0.138	Normal	
<i>BRCA1</i>	c.3748G>A	p.Glu1250Lys	11	+	+/-	+		X	0.117	Normal	59,60
<i>BRCA1</i>	c.3823A>G	p.Ile1275Val	11	+	+/-	+/-		X	-0.531	Normal	
<i>BRCA1</i>	c.4181C>T	p.Thr1394Ile	12	+/-	+/-	+/-			1.430	Normal	61
<i>BRCA1</i>	c.4516G>A	p.Asp1506Asn	15	?	?	?		X	-0.004	Normal	
<i>BRCA1</i>	c.4518T>G	p.Asp1506Glu	15	?	?	?		X	0.165	Normal	
<i>BRCA1</i>	c.4756G>T	p.Glu1586X	15	?	?	?		X	-2.319	Abnormal	
<i>BRCA1</i>	c.4810C>T	p.Gln1604X	16	-	-	-		X	-1.769	Abnormal	62
<i>BRCA1</i>	c.4812A>G	p.Gln1604=	16	+	+	+		X	-0.340	Normal	
<i>BRCA1</i>	c.5194-2A>G		20	?	-	-		X	-1.123	Intermediate	20
<i>BRCA1</i>	c.5254G>A	p.Ala1752Thr	20	+/-	+/-	+/-	2.58578 8759	X	-2.909	Abnormal	25,63
<i>BRCA1</i>	c.5254G>C	p.Ala1752Pro	20	+/-	-	+/-	3.12678 1127		-2.235	Abnormal	25,64, 65
<i>BRCA1</i>	c.5309G>T	p.Gly1770Val	21	+/-	-	-	2.59802 6614		-3.439	Abnormal	25,66
<i>BRCA1</i>	c.5434C>G	p.Pro1812Ala	23	-	+/-	-	2.75730 524	X	-2.136	Abnormal	25,67, 68
<i>BRCA2</i>	c.652G>A	p.Glu218Lys	8	?	?	?		X	-0.546	Normal	
<i>BRCA2</i>	c.2157T>G	p.Asn719Lys	11	?	?	?		X	0.359	Normal	
<i>BRCA2</i>	c.2803G>A	p.Asp935Asn	11	+	+	+		X	-0.042	Normal	69
<i>BRCA2</i>	c.3539A>G	p.Lys1180Arg	11	+/-	+/-	+/-		X	0.104	Normal	
<i>BRCA2</i>	c.5645C>A	p.Ser1882X	11	-	-	-		X	-2.104	Abnormal	
<i>BRCA2</i>	c.8182G>A	p.Val2728Ile	18	+	+	+		X	-0.598	Normal	
<i>BRCA2</i>	c.8486A>G	p.Gln2829Arg	19	-	+/-	+/-		X	-3.797	Abnormal	70

Variants classified as benign in databases (+) or functionally normal in our assay (Normal) are shown in green. Variants classified as pathogenic in databases (-) or functionally abnormal in our assay (Abnormal) are shown in red. Variants with conflicting interpretation are shown in orange (+/-), unreported mutations are shown in yellow (?). Variants with an intermediate score are shown in blue. Functional scores of variants already characterized by Findlay et al. [25] are included where available. Variants which were characterized in our laboratory are marked (X), the others were selected in the databases

Table 2 Baseline tumor characteristics of patients who were treated with PARP inhibitors

Samples	1	2	3	4	5	6
Tumor histology	High-grade serous ovarian carcinoma	High-grade serous ovarian carcinoma	High-grade serous ovarian carcinoma	High-grade serous ovarian carcinoma	High-grade serous ovarian carcinoma	High-grade serous ovarian carcinoma
Tumor type	Metastatic	Metastatic	Metastatic	Primary	Metastatic	Metastatic
Tumor content (%)	40	50	30	80	30	70
Tumor fragment type	Surgical sample	Biopsy	Biopsy	Surgical sample	Biopsy	Biopsy
Gene	BRCA1	BRCA1	BRCA1	BRCA1	BRCA1	BRCA2
Exon	11	15	16	20	23	11
Allele frequency (%)	8	23.5	19.9	75.7	75.9	61.7
Nucleotide change	c.872T>G	c.4756G>T	c.4810C>T	c.5194-2A>G	c.5434C>G	c.5645C>A
AA change	p.Leu291X	p.Glu1586X	p.Gln1604X	p.?	p.Pro1812Ala	p.Ser1882X
Somatic or germline	Somatic	Somatic	Somatic	Somatic	Germline	Germline
Patients' outcome following PARPi administration	Maintenance treatment by PARPi (olaparib) since 7 months	Progression after 6 months on PARPi (niraparib)	Maintenance treatment by PARPi (olaparib) since 10 months	Progression after 3 months on PARPi (olaparib)	Maintenance treatment by PARPi (olaparib) since 7 months	Maintenance treatment by PARPi (olaparib) since 8 months
Evaluated functional score	- 2.298	- 2.319	- 1.769	- 1.123	- 2.136	- 2.104
Functional consequence	Functionally abnormal	Functionally abnormal	Functionally abnormal	Intermediate	Functionally abnormal	Functionally abnormal



been reported to affect splicing and may also be functionally abnormal. Its classification as functionally intermediate might reflect the existence of a large number of *BRCA1* splicing variants with a not totally deficient splicing site also called leaky splicing effect [54–56]. The sensitivity of the patient 4, who relapsed after only 3 months on PARP inhibitor, was consistent with these hypotheses (Fig. 3) and might also reflect the low level of expression of the splice variant compared to the full-length transcript despite the high variant frequency in the sample (analysis of the patient's RNA by RT-qPCR and previously published) [20]. The second intermediate variant, p.Leu1080=, is located in the middle of exon 11 of *BRCA1*. This synonymous variant has not been reported in databases having a likelihood of resulting in a splicing alteration according to bioinformatic analyses. However, our intermediate functional score is consistent with the finding of the ESEfinder tool (<http://kramer01.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home>) [57, 58], a bioinformatic tool used to identify exonic splicing enhancers, which predicted that this variant might create a SRp40 ESE site. We cannot rule out the possibility that the cut-off used to classify the variants might be too low, but this value could be refined further after the analysis of more *BRCA1/2* variants. Moreover, all variants were compared to silent variants classified as non-pathogenic or as assumed to be benign. Following NGS, all these variants were indeed found to be functionally normal (Additional file 1: Table S3).

A 3-week period to determine the functional impact of a variant is compatible with clinical management and is one of the main advantages of our protocol (Fig. 1A). No clone selection is required, thus avoiding potential off-target effects that could occur in one cell. This genome editing-based protocol can be used to evaluate exonic, intronic, and even splicing variants. Moreover, the comparison between the calculated functional scores and the response of patients carrying these variants to PARP inhibitors were concordant, included for the intermediate score (Fig. 3). The main difficulty is finding a Cas9 PAM sequence (NGG) close enough to the variant to edit, to ensure binding of the gRNA. However, increasing numbers of Cas proteins from different species are becoming available, increasing PAM diversity [59]. We have already tested our protocol with the Cpf1 endonuclease. Similar results were obtained for the variant studied, implying a stable functional score, but editing efficiency in our hands was lower with this method. The need to improve editing efficiency may cease to be relevant with the appearance of base editing [60, 61] and, now, prime editing [62], which increase efficiency without the need for a double-strand break limiting off-target insertion and deletion frequencies. Moreover, prime editing will make functional characterization

possible not only for point variants, but also for insertions or deletions. We have also generalized in another polyclonal cell line, knock-out for the *XRCC4* gene, also implicated in the NHEJ pathway [63] (Additional file 1: Fig. S3B and S3C). The same functional scores were obtained for the variants analyzed in this line, including p.Gln210= variant of *BRCA1*. However, more already classified variants should be analyzed to evaluate the thresholds in this other polyclonal cell line.

In this era of personalized medicine based on molecular genomics results provided by clinical laboratories, there is a need for harmonization of functional assay validity. To assess the validity of our assay, we followed the guidelines proposed by Brnich et al [34]. It allowed us to determine that our assay provides a "moderate" strength of evidence of the functional consequence for both pathogenic and benign variants.

We then extended our analysis to the characterization of another essential tumor suppressor gene: *POLE*, which is considered important for determining access to immunotherapy and four variants were used as control: two classified as benign and two as pathogenic. However, one of the pathogenic variants (p.Leu424Val) had a functional score typical of a functionally normal variant. We also analyzed three *POLE* VUS (Fig. 3A, B). Prediction algorithms in databases were conflicted concerning the consequences of p.Arg1826Trp variant for protein function: SIFT (deleterious) and PolyPhen-2 (probably damaging). The functional score of this variant was consistent with it being a functionally abnormal variant. Moreover, two VUS were classified as intermediate (p.Asn363Asp and p.Asp301Gly), but the cut-off used here should be refined by the evaluation of more well-known variants. Both these intermediate variants are located in the exonuclease domain and may therefore affect the proofreading activity of the protein, leading to a hypermutation phenotype [64]. The amino acids Asp301 and Asn363 are highly conserved between species. Both are predicted to affect function, but no functional tests have been performed. Moreover, the replacement of Asn with Asp in position 210 greatly decreases the endonuclease activity of *POLE* [65] and p.Asn363Asp variant has not been reported, but p.Asn363Lys is classified as pathogenic. The p.Asp301Gly and p.Asn363Asp variants, which had intermediate functional scores, may, therefore, actually be functionally abnormal. Moreover, the patient carrying the p.Asn363Asp variant had received immunotherapy treatment. However, he relapsed after only 6 months of continuous treatment which is consistent with our findings.

The principal difficulty encountered during this analysis concerned p.Leu424Val *POLE* variant, observed in multiple individuals with either an attenuated polyposis

phenotype or a history of colorectal cancer [66–68]. This variant classified as pathogenic is also located in the exonuclease domain. However, the functional score obtained here suggested that it was more a functionally normal variant (Fig. 3). Leucine and valine have similar properties, so this substitution is generally considered conservative. According to an article on hypermutation, the mutation of leucine in position 424 has consequences very different to those of other well-known pathogenic variants of *POLE* (p.Pro436Ser or p.Pro286Arg) [69]. A strong mutation phenotype was observed when this residue was replaced with a proline or an isoleucine residue, but not when it was replaced with a valine. The mutation load associated with the loss of *POLE* proofreading activity therefore depending on the type of amino acid change involved, potentially accounting for our functional score. Alternatively, there may have been too little time before DNA extraction in our model for missense variants to accumulate, due to the loss of exonuclease activity. The loss of some genes may therefore require more time, despite their essentiality for highly deleterious effects on cell survival.

Conclusions

The method presented here was proved effective for the characterization of the functional impact of *BRCA1* and *BRCA2* VUS and concordant with clinical data where available. More importantly, it can be used to obtain the necessary biological evidence of VUS function required for the prescription of targeted treatment within less than 3 weeks, which is compatible with use in clinical application. The patient carrying the genomic abnormality therefore benefits from an analysis of his or her variant, with potential consequences for relatives. This is particularly important for extremely rare somatic variants, which is compared to orphan diseases. The extension of its application to the study of *POLE* variants is already underway, and this method could be extended to the characterization of all essential tumor suppressor genes in our model [26], not only within the field of oncology. At a time at which purely in silico approaches are being used to guide therapeutic decisions, a method evaluating the functional implications of VUS is essential. Genomic editing is, thus, a promising tool for personalizing medicine and providing access to targeted therapy.

Abbreviations

VUS: Variants of uncertain significance; HR: Homologous recombination; NGS: Next-generation sequencing; IMDM: Isocove's modified Dulbecco's media; PI: Propidium iodide; gRNA: Guide RNA; PAM: Palindromic adjacent motif; KO: Knock-out; IGV: Integrative genomics viewer; HGVS: Human genome variation society; Indels: Insertions or deletions; SNV: Single-nucleotide variant; NHEJ: Non-homologous end-joining; HDR: Homologous-directed repair; RING: Really interested new gene; BRCT: BRCA1 C terminus

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13073-021-00976-x>.

Additional file 1: Figure S1. Generation of polyclonal LIG4 knock-out HAP1 cells. A. Schematic diagram of the LIG4 DNA region targeted by the gRNA, and localization of the AflIII restriction site used to knock out this gene in HAP1 cells. B. Schematic diagram of the protocol used to generate the polyclonal LIG4 KO cell line. C. Electrophoresis results following digestion with the AflIII restriction enzyme for screening of the clones. D. Sanger sequence alignments of the clones selected on basis of AflIII restriction site loss relative to HAP1 parental cells. E. Sanger electrophoregrams of LIG4 KO clones with insertions or deletions causing frameshifts relative to the HAP1 parental cell line. **Figure S2.** Analysis of coverage and frequencies of insertions and deletions obtained following genome editing and NGS sequencing of *BRCA1* and *BRCA2* variants. A. Insertion and deletion frequencies determined following the NGS sequencing of *BRCA1/2* variants of interest, compared with those for the corresponding silent control. These means include the 7 nt surrounding the mutations of interest. B. Analysis of NGS sequencing coverage with the following formula: Mutation of interest (Mutation+Reference control)/ Silent control (Mutation+Reference control). Results for all *BRCA1/2* variants characterized by comparison with the classified benign and pathogenic mutations (Mann-Whitney tests, $p = 0.1934$ and $p = 0.2902$). C. Frequencies of insertions and deletions according to the NGS sequencing of *BRCA1/2* variants and corresponding silent editing controls and linear regression analysis. These means include the 8 nt surrounding the PAM sequence. Variants classified in databases are shown in colours: benign mutations in green and pathogenic mutations in red. **Figure S3.** Evaluation of the reproducibility of our functional assay. A. Replicates of functions scores from four different variants obtained by following the described protocol in the LIG4 KO HAP1 cells ($n = 3$). Among these, the variants p.Pro1812Ala and p.Tyr422X were already classified as pathogenic in databases, the other two were VUS or had a conflicting interpretation. B. Sanger electrophoregrams of XRCC4 KO clones with insertions or deletions causing frameshifts relative to the HAP1 parental cell line. Clones 5 and 17 were pooled to generate the polyclonal XRCC4 KO HAP1 cell line. C. Comparison of the functional scores evaluated after CRISPR-Cas9 editing and NGS sequencing of four *BRCA1* variants in polyclonal LIG4 KO cells and polyclonal XRCC4 KO cells. **Table S1.** gRNA and oligonucleotides sequences designed to edit *BRCA1*, *BRCA2* and *POLE* variants in the study. **Table S2.** Odds of Pathogenicity (OddsPath) estimated by performance of classified variant controls. In this assay, five benign and five pathogenic variants of *BRCA1* and *BRCA2* were used as controls allowing the estimation of the evidence strength of our classification according to Brnich et al [34] guidelines. **Table S3.** Variants used as silent controls or references with their classification according to different databases. Benign mutations are shown in green (+), unreported mutations are shown in yellow (?), mutations with conflicting interpretation are shown in orange (+/-), pathogenic mutations are shown in red (-) and intermediate mutations are shown in blue (0).

Acknowledgements

We would like to thank the clinicians of the Institut de Cancérologie de l'Ouest (Angers, France) and the Centre Hospitalier du Mans (Le Mans, France) for their collaboration in this project. We also appreciated the help of J. Dauve and A. Renoult with NGS sequencing analysis and genome editing.

Authors' contributions

AB and A.M. conceived the project. AB and A.M. designed the experiments. AB performed the experiments and AB, L-M.C., C.P., and A.M. performed analysis. P.A, J-S.F., and M.C. performed the clinical evaluation. AB, C.P., and A.M. wrote the manuscript. All author(s) read and approved the final manuscript.

Funding

This work was supported in part by a grant from Cancéropôle Grand Ouest. Amandine Billaud was supported by a grant from Institut de Cancérologie de l'Ouest.

Availability of data and materials

The data that support the findings of this study are available within the manuscript or publicly deposited. Sequence data have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB47293 (<https://www.ebi.ac.uk/ena/browser/view/PRJEB47293>) [70].

Declarations**Ethics approval and consent to participate**

The retrospective study procedures have been approved by the ethics committee of the Angers medical university (France) under ref 2020-15. All patients provided written informed consent to participate in the study. This research conformed to the principles of the Helsinki Declaration.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interest.

Author details

¹Université d'Angers, Inserm, CRCINA, SFR ICAT, F-49000 Angers, France.

²Institut de Cancérologie de l'Ouest Nantes-Angers, F-49000 Angers, France.

³Université de Nantes, Inserm, CRCINA, F-44000 Nantes, France.

Received: 20 May 2020 Accepted: 23 September 2021

Published online: 09 November 2021

References

- Van't Veer LJ, Bernards R. Enabling personalized cancer medicine through analysis of gene-expression patterns. *Nature*. 2008;452(7187):564–70. <https://doi.org/10.1038/nature06915>.
- Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitpiralla S, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. 2018;46(D1):D1062–7. <https://doi.org/10.1093/nar/glx1153>.
- Robson ME, Bradbury AR, Arun B, Dornchek SM, Ford JM, Hampel HL, et al. American Society of Clinical Oncology Policy Statement Update: genetic and genomic testing for cancer susceptibility. *J. Clin. Oncol*. 2015;33(31):3660–7. <https://doi.org/10.1200/JCO.2015.63.0996>.
- Liu Y, Hu X, Han C, Wang L, Zhang X, He X, et al. Targeting tumor suppressor genes for cancer therapy. *BioEssays*. 2015;37(12):1277–86. <https://doi.org/10.1002/bies.201500093>.
- Roy R, Chun J, Powell SN. BRCA1 and BRCA2: different roles in a common pathway of genome protection. *Nat. Rev. Cancer*. 2011;12(1):68–78. <https://doi.org/10.1038/nrc3181>.
- Huen MSY, Sy SMH, Chen J. BRCA1 and its toolbox for the maintenance of genome integrity. *Nat. Rev. Mol. Cell Biol*. 2010;11(2):138–48. <https://doi.org/10.1038/nrm2831>.
- Shahid T, Soroka J, Kong EH, Malivert L, McIlwraith MJ, Pape T, et al. Structure and mechanism of action of the BRCA2 breast cancer tumor suppressor. *Nat. Struct. Mol. Biol*. 2014;21(11):962–8. <https://doi.org/10.1038/nsmb.2899>.
- Norquist BM, Brady MF, Harrell MI, Walsh T, Lee MK, Gulsuner S, et al. Mutations in homologous recombination genes and outcomes in ovarian carcinoma patients in GOG 218: An NRG oncology/Gynecologic oncology group study. *Clin. Cancer Res*. 2018;24(4):777–83. <https://doi.org/10.1158/1078-0432.CCR-17-1327>.
- Tutt A, Ashworth A. The relationship between the roles of BRCA genes in DNA repair and cancer predisposition. *Trends Mol. Med*. 2002;8(12):571–6. [https://doi.org/10.1016/S1471-4914\(02\)02434-6](https://doi.org/10.1016/S1471-4914(02)02434-6).
- Farmer H, McCabe N, Lord CJ, Tutt ANJ, Johnson DA, Richardson TB, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature*. 2005;434(7035):917–21. <https://doi.org/10.1038/nature03445>.
- Bryant HE, Schultz N, Thomas HD, Parker KM, Flower D, Lopez E, et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature*. 2005;434(7035):913–7. <https://doi.org/10.1038/nature03443>.
- Ledermann J, Harter P, Gourley C, Friedlander M, Vergote I, Rustin G, et al. Olaparib maintenance therapy in platinum-sensitive relapsed ovarian cancer. *N. Engl. J. Med*. 2012;366(15):1382–92. <https://doi.org/10.1056/NEJMoa1205535>.
- Bérout C, Collod-Bérout G, Boileau C, Soussi T, Junien C. UMD (Universal mutation database): a generic software to build and analyze locus-specific databases. *Hum. Mutat*. 2000;15(1):86–94. [https://doi.org/10.1002/\(SICI\)1098-1004\(200001\)15:1<86::AID-HUMU16>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1098-1004(200001)15:1<86::AID-HUMU16>3.0.CO;2-4).
- Konstantinopoulos PA, Spentzos D, Karlan BY, Taniguchi T, Fountzilas E, Francoeur N, et al. Gene expression profile of BRCAness that correlates with responsiveness to chemotherapy and with outcome in patients with epithelial ovarian cancer. *J. Clin. Oncol*. 2010;28(22):3555–61. <https://doi.org/10.1200/JCO.2009.27.5719>.
- Jazaeri AA, Yee CJ, Sotiřou C, Brantley KR, Boyd J, Liu ET. Gene expression profiles of BRCA1-linked, BRCA2-linked, and sporadic ovarian cancers. *J. Natl. Cancer Inst*. 2002;94(13):990–1000. <https://doi.org/10.1093/jnci/94.13.990>.
- Popova T, Manié E, Rieunier G, Caux-Moncoutier V, Tirapo C, Dubois T, et al. Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. *Cancer Res*. 2012;72(21):5454–62. <https://doi.org/10.1158/0008-5472.CAN-12-1470>.
- Graeser M, McCarthy A, Lord CJ, Savage K, Hills M, Salter J, et al. A marker of homologous recombination predicts pathologic complete response to neoadjuvant chemotherapy in primary breast cancer. *Clin. Cancer Res*. 2010;16(24):6159–68. <https://doi.org/10.1158/1078-0432.CCR-10-1027>.
- Fraille-Bethencourt E, et al. Functional classification of DNA variants by hybrid minigenes: Identification of 30 spliceogenic variants of BRCA2 exons 17 and 18. *PLoS Genet*. 2017;13:1–21.
- Spearman AD, Sweet K, Zhou XP, McLennan J, Couch FJ, Toland AE. Clinically applicable models to characterize BRCA1 and BRCA2 variants of uncertain significance. *J. Clin. Oncol*. 2008;26(33):5393–400. <https://doi.org/10.1200/JCO.2008.17.8228>.
- Chevalier L-M, Billaud A, Fronteau S, Dauvé J, Patsouris A, Verrielle V, et al. Somatic mRNA analysis of BRCA1 splice variants provides a direct therapeutic impact on PARP inhibitors. *Mol. diagnosis Ther*. 2020;24(2):233–43. <https://doi.org/10.1007/s40291-020-00452-z>.
- Jinek M, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337:816–22.
- Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc*. 2013;8(11):2281–308. <https://doi.org/10.1038/nprot.2013.143>.
- Doench J, Arn G. I ready for CRISPR? A user's guide to genetic screens. *Nat. Rev. Genet*. 2017;19(2):67–80. <https://doi.org/10.1038/nrg.2017.97>.
- Findlay GM, Boyle EA, Hause RJ, Klein JC, Shendure J. Saturation editing of genomic regions by multiplex homology-directed repair. *Nature*. 2014;513(7516):120–3. <https://doi.org/10.1038/nature13695>.
- Findlay G, M, et al. Accurate functional classification of thousands of BRCA1 variants with saturation genome editing. *Nature*. 2018;559(7601):125–9. <https://doi.org/10.1038/nature17664>.
- Blomen V, et al. Gene essentiality and synthetic lethality in haploid human cells. *Science*. 2015;350:1092, 2015.
- Vindeløv LL, Christensen U, Nissen NI. A Detergent-trypsin method for the preparation of nuclei for flow cytometric DNA analysis. *Cytometry*. 1983;3(5):323–7. <https://doi.org/10.1002/cyto.990030503>.
- Yu X, Liang X, Xie H, Kumar S, Ravinder N, Potter J, et al. Improved delivery of Cas9 protein/gRNA complexes using lipofectamine CRISPRMAX. *Biotechnol. Lett*. 2016;38(6):919–29. <https://doi.org/10.1007/s10529-016-2064-9>.
- Chakravarty, D, et al. OncoKB: a precision oncology knowledge base. *JCO Precis. Oncol*. 1–16 (2017) doi:<https://doi.org/10.1200/PO.17.00011>.
- Paquet D, Kwart D, Chen A, Sproul A, Jacob S, Teo S, et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature*. 2016;533(7601):125–9. <https://doi.org/10.1038/nature17664>.
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat. Biotechnol*. 2011;29(1):24–6. <https://doi.org/10.1038/nbt.1754>.
- Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol*. 2015;33(5):543–8. <https://doi.org/10.1038/nbt.3198>.
- Doudna JA, Charpentier E. The new frontier of genome engineering with CRISPR-Cas9. *Science*. 2014;346:1258096.
- Brnich SE, et al. Recommendations for application of the functional evidence P53/B53 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Med*. 2019;12:3.

35. Powell SN, Kachnic LA. Roles of BRCA1 and BRCA2 in homologous recombination, DNA replication fidelity and the cellular response to ionizing radiation. *Oncogene*. 2003;22(37):5784–91. <https://doi.org/10.1038/sj.onc.1206678>.
36. Lindor NM, Guidugli L, Wang X, Vallée MP, Monteiro ANA, Tavtigian S, et al. A review of a multifactorial probability-based model for classification of BRCA1 and BRCA2 variants of uncertain significance (VUS). *Hum. Mutat*. 2012;33(1):8–21. <https://doi.org/10.1002/humu.21627>.
37. Towler WI, Zhang J, Ransburgh DJR, Toland AE, Ishioka C, Chiba N, et al. Analysis of BRCA1 variants in double-strand break repair by homologous recombination and single-strand annealing. *Hum. Mutat*. 2013;34(3):439–45. <https://doi.org/10.1002/humu.22251>.
38. Dosi V, Tosar A, Cañadas C, Pérez-Segura P, Díaz-Rubio E, Caldés T, et al. Alternative splicing and molecular characterization of splice site variants: BRCA1 c.591C > T as a case study. *Clin. Chem*. 2010;56(1):53–61. <https://doi.org/10.1373/clinchem.2009.132274>.
39. Coulet F, Pires F, Rouleau E, Lefol C, Martin S, Colas C, et al. A one-step prescreening for point mutations and large rearrangement in BRCA1 and BRCA2 genes using quantitative polymerase chain reaction and high-resolution melting curve analysis. *Genet. Test. Mol. Biomarkers*. 2010;14(5): 677–90. <https://doi.org/10.1089/gtmb.2009.0183>.
40. Lu C, Xie M, Wendt MC, Wang J, McLellan MD, Leiserson MDM, et al. Patterns and functional implications of rare germline variants across 12 cancer types. *Nat. Commun*. 2015;6(1). <https://doi.org/10.1038/ncomms10086>.
41. Bouwman P, van der Gulden H, van der Heijden I, Drost R, Kljin CN, Praseyanti P, et al. A high-throughput functional complementation assay for classification of BRCA1 missense variants. *Cancer Discov*. 2013;3(10): 1142–55. <https://doi.org/10.1158/2159-8290.CD-13-0094>.
42. Tavtigian SV, Deffenbaugh AM, Yin L, Judkins T, Scholl T, Samollow PB, et al. Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J. Med. Genet*. 2006; 43(4):295–305. <https://doi.org/10.1136/jmg.2005.033878>.
43. Woods NT, et al. Functional assays provide a robust tool for the clinical annotation of genetic variants of uncertain significance. *NPJ Genomic Med*. 2016;1:16001.
44. Esteban-Cardeñosa E, Duran M, Infante M, Velasco E, Miner C. High-throughput mutation detection method to scan BRCA1 and BRCA2 based on heteroduplex analysis by capillary array electrophoresis. *Clin. Chem*. 2004;50(2):313–20. <https://doi.org/10.1373/clinchem.2003.023614>.
45. Carvalho RS, Abreu RBV, Velkova A, Marsillac S, Rodarte RS, Suarez-Kurtz G, et al. Probing structure-function relationships in missense variants in the carboxy-terminal region of BRCA1. *PLoS One*. 2014;9(5):1–10. <https://doi.org/10.1371/journal.pone.0097766>.
46. Lee MS, Green R, Marsillac SM, Coquelle N, Williams RS, Yeung T, et al. Comprehensive analysis of missense variations in the BRCT domain of BRCA1 by structural and functional assays. *Cancer Res*. 2010;70(12):4880–90. <https://doi.org/10.1158/0008-5472.CAN-09-4563>.
47. Phelan CM, Dapic V, Tice B, Favis R, Kwan E, Barany F, et al. Classification of BRCA1 missense variants of unknown clinical significance. *J. Med. Genet*. 2005;42(2):138–46. <https://doi.org/10.1136/jmg.2004.024711>.
48. Tudini E, Moghadasi S, Parsons MT, van der Kolk L, van den Ouweland AMW, Niederacher D, et al. Substantial evidence for the clinical significance of missense variant BRCA1 c.5309G>T p.(Gly1770Val). *Breast Cancer Res. Treat*. 2018;172(2):497–503. <https://doi.org/10.1007/s10549-018-4903-y>.
49. Gaildrat P, et al. The BRCA1 c.5434C→G (p.Pro1812Ala) variant induces a deleterious exon 23 skipping by affecting exonic splicing regulatory elements. *J. Med. Genet*. 47, 398–403 (2010).
50. Jarhelle E, Riise Stensland HMF, Mæhle L, Van Ghelue M. Characterization of BRCA1 and BRCA2 variants found in a Norwegian breast or ovarian cancer cohort. *Fam. Cancer*. 2017;16(1):1–16. <https://doi.org/10.1007/s10689-016-9916-2>.
51. Wagner TMU, et al. Global sequence diversity of BRCA2: Analysis of 71 breast cancer families and 95 control individuals of worldwide populations. *Hum. Mol. Genet*. 1999;8(3):413–23. <https://doi.org/10.1093/hmg/8.3.413>.
52. Kraus C, Hoyer J, Vasileiou G, Wunderle M, Lux MP, Fasching PA, et al. Gene panel sequencing in familial breast/ovarian cancer patients identifies multiple novel mutations also in genes others than BRCA1/2. *Int. J. Cancer*. 2017;140(1):95–102. <https://doi.org/10.1002/ijc.30428>.
53. Wang B, et al. Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science*. 2007;316:1194–8.
54. Lovelock PK, et al. Identification of BRCA1 missense substitutions that confer partial functional activity: potential moderate risk variants? *Breast Cancer Res*. 2007;9:1–13.
55. Montalbán G, Bonache S, Moles-Femández A, Gadea N, Tenés A, Torres-Esquius S, et al. Incorporation of semi-quantitative analysis of splicing alterations for the clinical interpretation of variants in BRCA1 and BRCA2 genes. *Hum. Mutat*. 2019;40(12): 2296–317. <https://doi.org/10.1002/humu.23882>.
56. Tubeuf H, Caputo SM, Sullivan T, Rondeaux J, Krieger S, Caux-Moncoutier V, et al. Calibration of pathogenicity due to variant-induced leaky splicing defects by using BRCA2 exon 3 as a model system. *Cancer Res*. 2020;80(17): 3593–605. <https://doi.org/10.1158/0008-5472.CAN-20-0895>.
57. Smith PJ, Zhang C, Wang J, Chew SL, Zhang MQ, Krainer AR. An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Hum. Mol. Genet*. 2006;15(16):2490–508. <https://doi.org/10.1093/hmg/ddl1171>.
58. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res*. 2003; 31(13):3568–71. <https://doi.org/10.1093/nar/gkg616>.
59. Leenay RT, Maksimchuk KR, Slotkowski RA, Agrawal RN, Gomia AA, Briner AE, et al. Identifying and visualizing functional PAM diversity across CRISPR-Cas systems. *Mol. Cell*. 2016;62(1):137–47. <https://doi.org/10.1016/j.molcel.2016.02.031>.
60. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*. 2016;533(7603):420–4. <https://doi.org/10.1038/nature17946>.
61. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, et al. Programmable base editing of A-T to G-C in genomic DNA without DNA cleavage. *Nature*. 2017;551(7681):464–71. <https://doi.org/10.1038/nature24644>.
62. Anzalone AV, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*. 2019;576:149.
63. Pierce AJ, Hu P, Han M, Ellis N, Jasin M, Ku DNA end-binding protein modulates homologous repair of double-strand breaks in mammalian cells. *Genes Dev*. 2001;15(24):3237–42. <https://doi.org/10.1101/gad.946401>.
64. Palles C, et al. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat. Genet*. 2013;45(2):136–43. <https://doi.org/10.1038/ng.2503>.
65. Nishioka M, Mizuguchi H, Fujiwara S, Komatsubara S, Kitabayashi M, Uemura H, et al. Long and accurate PCR with a mixture of KOD DNA polymerase and its exonuclease deficient mutant enzyme. *J. Biotechnol*. 2001;88(2):141–9. [https://doi.org/10.1016/S0168-1656\(01\)00275-9](https://doi.org/10.1016/S0168-1656(01)00275-9).
66. Valle L, Hernández-Illán E, Bellido F, Aiza G, Castillejo A, Castillejo MI, et al. New insights into POLE and POLD1 germline mutations in familial colorectal cancer and polyposis. *Hum. Mol. Genet*. 2014;23(13):3506–12. <https://doi.org/10.1093/hmg/ddu058>.
67. Elsayed FA, Kets CM, Ruano D, van den Akker B, Mensenkamp AR, Schrupp M, et al. Germline variants in POLE are associated with early onset mismatch repair deficient colorectal cancer. *Eur. J. Hum. Genet*. 2015;23(8):1080–4. <https://doi.org/10.1038/ejhg.2014.242>.
68. Spier I, Holzapfel S, Altmüller J, Zhao B, Horpaan S, Vogt S, et al. Frequency and phenotypic spectrum of germline mutations in POLE and seven other polymerase genes in 266 patients with colorectal adenomas and carcinomas. *Int. J. Cancer*. 2015;137(2):320–31. <https://doi.org/10.1002/ijc.29396>.
69. Campbell BB, et al. Comprehensive analysis of hypermutation in human cancer. *Cell*. 2017;1–15. <https://doi.org/10.1016/j.cell.2017.09.048>.
70. Billaud A, Chevalier L-M, Augereau P, Frenel J-S, Passot C, Campone M, et al. Functional pre-therapeutic evaluation by genome editing of variants of uncertain significance of essential tumor suppressor genes. *PRJEB47293*. European Nucleotide Archive (ENA) at EMBL-EBI. 2021; <https://www.ebi.ac.uk/ena/browser/view/PRJEB47293>. Accessed 16 Sept.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information

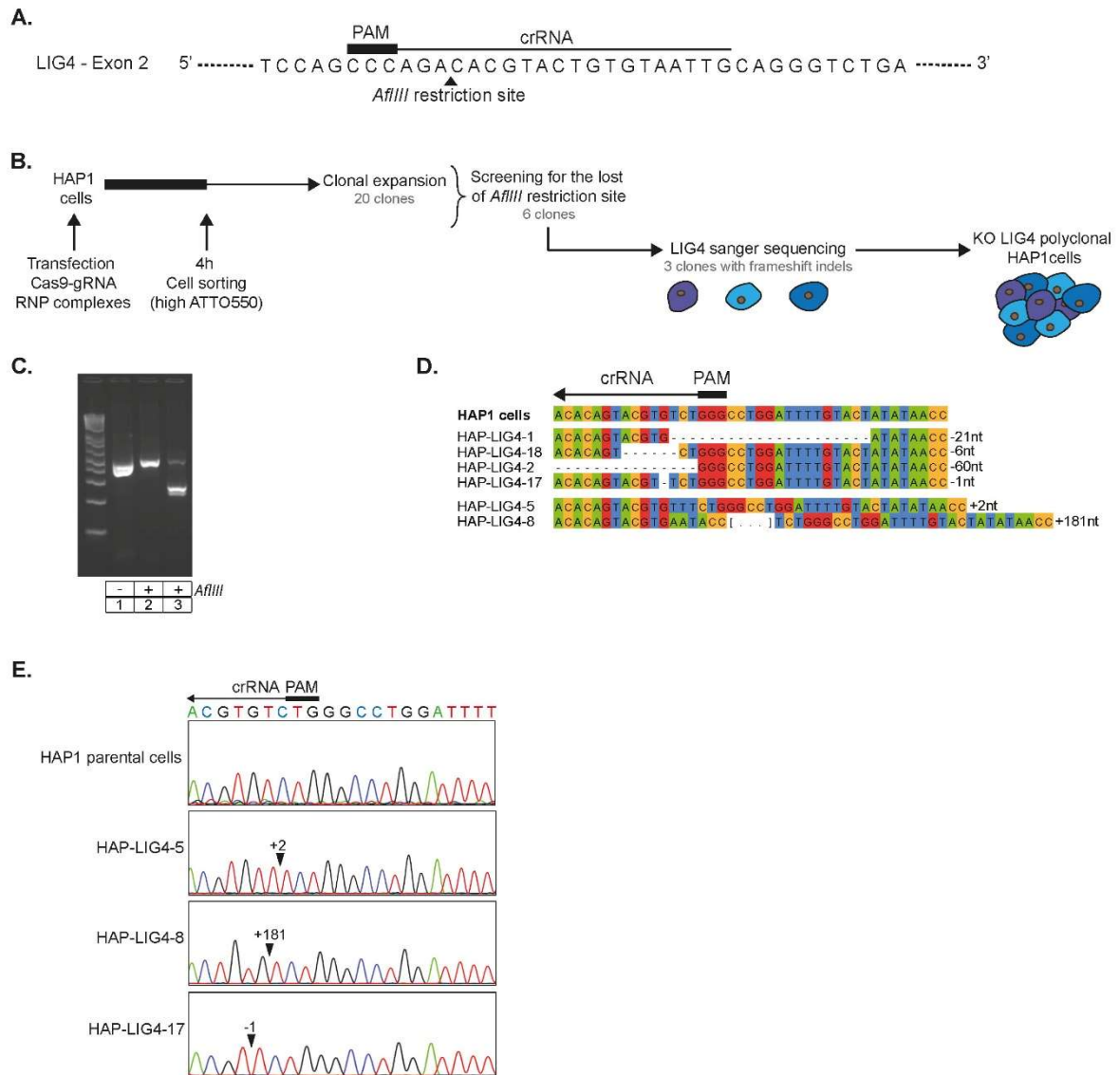


Figure S1: Generation of polyclonal LIG4 knock-out HAP1 cells. **A.** Schematic diagram of the LIG4 DNA region targeted by the gRNA, and localization of the AflIII restriction site used to knock out this gene in HAP1 cells. **B.** Schematic diagram of the protocol used to generate the polyclonal LIG4 KO cell line. **C.** Electrophoresis results following digestion with the AflIII restriction enzyme for screening of the clones. **D.** Sanger sequence alignments of the clones selected on basis of AflIII restriction site loss relative to HAP1 parental cells. **E.** Sanger electrophoregrams of LIG4 KO clones with insertions or deletions causing frameshifts relative to the HAP1 parental cell line.

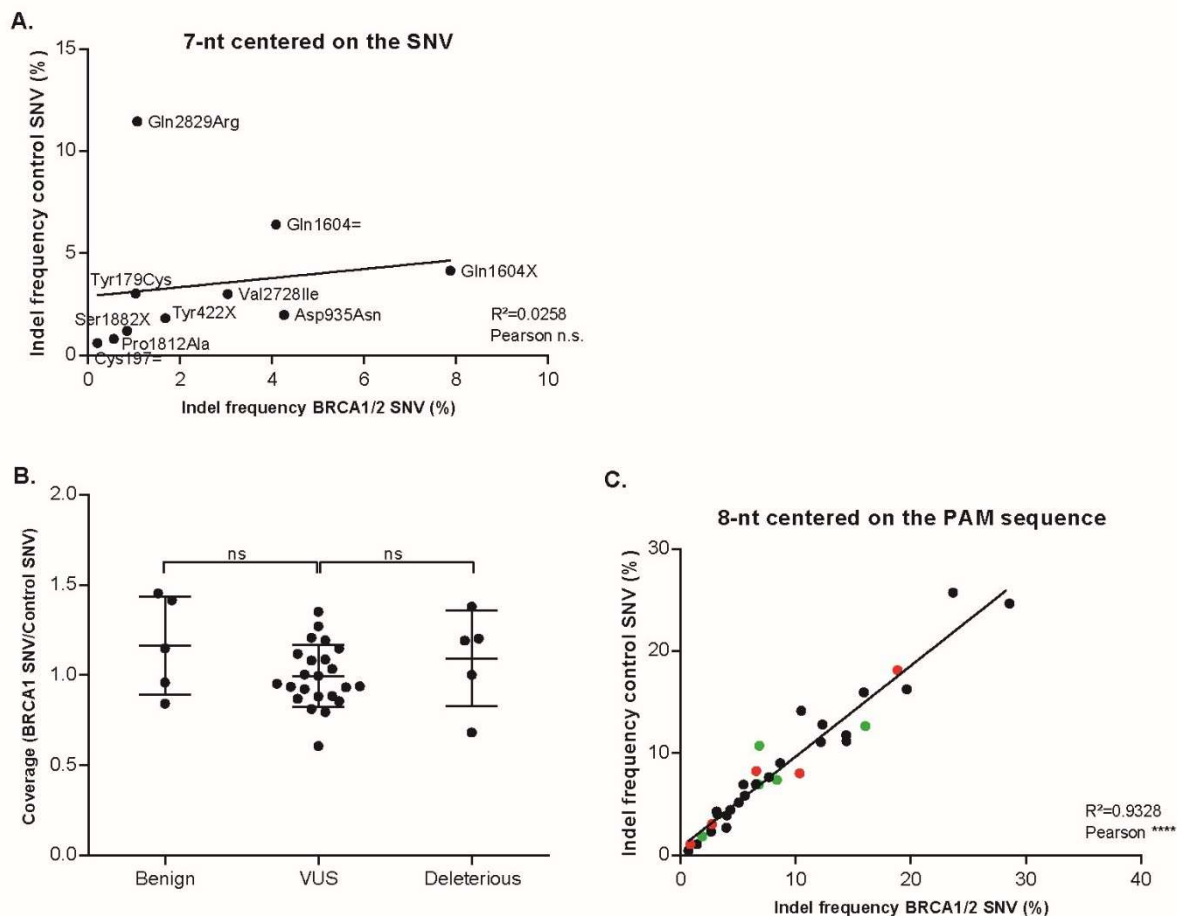


Figure S2: Analysis of coverage and frequencies of insertions and deletions obtained following genome editing and NGS sequencing of *BRCA1* and *BRCA2* variants. **A.** Insertion and deletion frequencies determined following the NGS sequencing of *BRCA1/2* variants of interest, compared with those for the corresponding silent control. These means include the 7 nt surrounding the mutations of interest. **B.** Analysis of NGS sequencing coverage with the following formula: Mutation of interest (Mutation+Reference control)/Silent control (Mutation+Reference control). Results for all *BRCA1/2* variants characterized by comparison with the classified benign and pathogenic mutations (Mann-Whitney tests, $p=0.1934$ and $p=0.2902$). **C.** Frequencies of insertions and deletions according to the NGS sequencing of *BRCA1/2* variants and corresponding silent editing controls and linear regression analysis. These means include the 8 nt surrounding the PAM sequence. Variants classified in databases are shown in colours: benign mutations in green and pathogenic mutations in

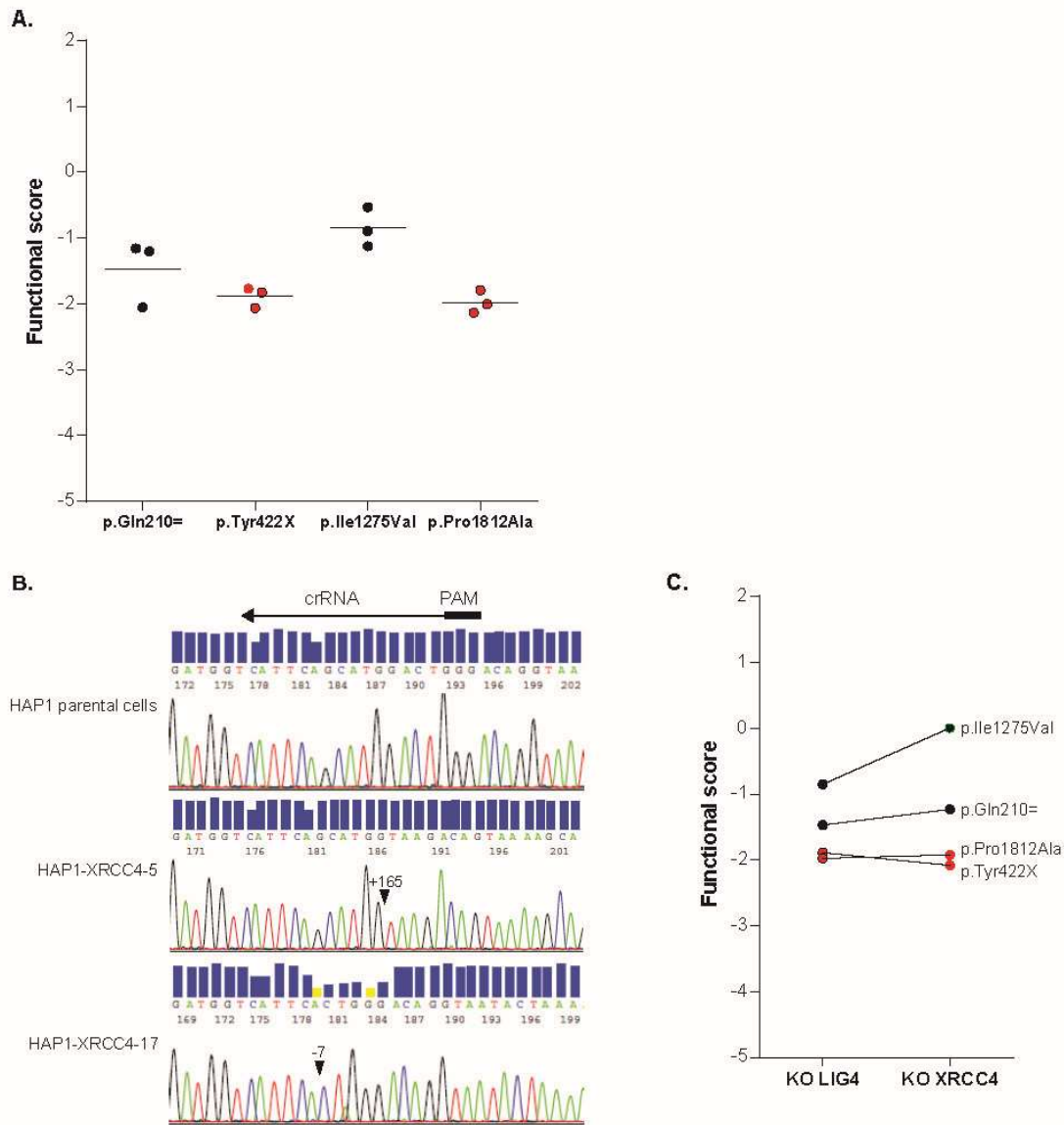


Figure S3: Evaluation of the reproducibility of our functional assay. **A.** Replicates of functions scores from four different variants obtained by following the described protocol in the LIG4 KO HAP1 cells (n=3). Among these, the variants p.Pro1812Ala and p.Tyr422X were already classified as pathogenic in databases, the other two were VUS or had a conflicting interpretation. **B.** Sanger electrophoregrams of XRCC4 KO clones with insertions or deletions causing frame-shifts relative to the HAP1 parental cell line. Clones 5 and 17 were pooled to generate the polyclonal XRCC4 KO HAP1 cell line. **C.** Comparison of the functional scores evaluated after CRISPR-Cas9 editing and NGS sequencing of four BRCA1 variants in polyclonal LIG4 KO cells and polyclonal XRCC4 KO cells.

Table S1: crRNA and oligonucleotides sequences designed to edit *BRCA1*, *BRCA2* and *POLE* variants in the study.

Gene	Names	crRNA sequences
<i>BRCA1</i>	BRCA1-Ile31Asn	TGCTAGTCTGGAGTTGATCA
<i>BRCA1</i>	BRCA1-Glu149Ala	TGGCTTCTGCTAAACAGTA
<i>BRCA1</i>	BRCA1-Thr150Ala	TGGCTTCTGCTAAACAGTA
<i>BRCA1</i>	BRCA1-Tyr179Cys	GACGTCTGTCTACATTGAAT
<i>BRCA1</i>	BRCA1-Val191Asp	TTCTGAAGATACCGTTAATA
<i>BRCA1</i>	BRCA1-Cys197=	TTCTGAAGATACCGTTAATA
<i>BRCA1</i>	BRCA1-Gln210=	ATCCAAACTGATTTTCATCCC
<i>BRCA1</i>	BRCA1-Leu291X	GTTCTCATGCTGTAATGAGC
<i>BRCA1</i>	BRCA1-Tyr422X	TGTATTGGACGTTCTAAATG
<i>BRCA1</i>	BRCA1-Gly462Arg	TGGGAAAACCTATCGGAAGA
<i>BRCA1</i>	BRCA1-Arg979Cys	AAAGTGGTGGTATACGATAT
<i>BRCA1</i>	BRCA1-Leu1080=	GAATGCTATGCTTAGATTAG
<i>BRCA1</i>	BRCA1-Gly1201Ser	CCCTTTCACCCATACACATT
<i>BRCA1</i>	BRCA1-Glu1250Lys	AGACAGACACTCGGTAGCAA
<i>BRCA1</i>	BRCA1-Ile1275Val	TAACCAGGTAATATTGGCAA
<i>BRCA1</i>	BRCA1-Thr1394Ile	ACACACGCTTTTTACCTGAG
<i>BRCA1</i>	BRCA1-Asp1506Asn	ACCACCTATCATCTAATGAT
<i>BRCA1</i>	BRCA1-Asp1506Glu	ACCACCTATCATCTAATGAT
<i>BRCA1</i>	BRCA1-Glu1586X	GCCAACACGAGCTGACTCTG
<i>BRCA1</i>	BRCA1-Gln1604X	TTGGGGAACCTTCAATGCAG
<i>BRCA1</i>	BRCA1-Gln1604=	TTGGGGAACCTTCAATGCAG
<i>BRCA1</i>	BRCA1-Intron20	GAAACCAAACACAACCCATC
<i>BRCA1</i>	BRCA1-Ala1752Thr	AAAGCGAGCAAGAGAATCCC
<i>BRCA1</i>	BRCA1-Ala1752Pro	AAAGCGAGCAAGAGAATCCC
<i>BRCA1</i>	BRCA1-Gly1770Val	ACCTGTGGCATGTTGGTGA
<i>BRCA1</i>	BRCA1-Pro1812Ala	CTGGCTGCACAACCACAATT
<i>BRCA2</i>	BRCA2-Glu218Lys	CTTACAGCAGTAGTATCATG
<i>BRCA2</i>	BRCA2-Asn719Lys	TGATTCTGTGCATGCCTGC
<i>BRCA2</i>	BRCA2-Asp935Asn	ATGGTTTTATATGGAGACAC
<i>BRCA2</i>	BRCA2-Lys1180Arg	TGTCTACCTGACCAATCGAT
<i>BRCA2</i>	BRCA2-Ser1882X	AACCTGCCATAATTTTCGTT
<i>BRCA2</i>	BRCA2-Val2728Ile	AGATGGGTGGTATGCTGTTA
<i>BRCA2</i>	BRCA2-Gln2829Arg	TCAAAGAGCATACCCTATAC
<i>POLE</i>	POLE-Ala31Ser	AGTTTCGGCACTCAAGCGCC
<i>POLE</i>	POLE-Pro286Ser	TGTAGGAAATCATCATAATC
<i>POLE</i>	POLE-Asp301Gly	CACCTGGCCATCGATCATGT
<i>POLE</i>	POLE-Asn363Asp	CATCATGGTCACCTACAACG
<i>POLE</i>	POLE-Leu424Val	GGGCAGTCATAATCTCAAGG
<i>POLE</i>	POLE-Ala456Pro	GAATACGTGGCCAGAGTCTG
<i>POLE</i>	POLE-Phe695Ile	CCATCGGATCCAGCACCAGC

<i>POLE</i>	POLE-Arg1826Trp	AGCGGTAGAAGTGCATCACC
<i>LIG4</i>	LIG4-KO	CAATTACACAGTACGTGTCT
<i>XRCC4</i>	XRCC4-KO	ATGGTCATTACAGCATGGACT
Genes	Names	Donor sequences
<i>BRCA1</i>	BRCA1-Ile31Asn-mut	GCAAAATATGTGGTCACACTTTGTGGAGACAGGTTCTTTGTTCAACTCCAGACTAGCAGGGTAGG GGGGGAGAAAAAGAAAAATAATGAGGC
<i>BRCA1</i>	BRCA1-Ile31Asn-sil	GCAAAATATGTGGTCACACTTTGTGGAGACAGGTTCTTTAATCAACTCCAGACTAGCAGGGTAGG GGGGGAGAAAAAGAAAAATAATGAGGC
<i>BRCA1</i>	BRCA1-Glu149Ala-mut	CCAAGGTTAGAGAGTTGGCACTGAGACTGGTTGCCTGCTAAACAGTATGATAAAGAACAGTCA AGCAATTGTTGGCCAGTTCTGTGC
<i>BRCA1</i>	BRCA1-Glu149Ala-sil	CCAAGGTTAGAGAGTTGGCACTGAGACTGGTTGCCTGCTAAACAGTATGATAAAGAACAGTCA AGCAATTGTTGGCCAGTTCTGTGC
<i>BRCA1</i>	BRCA1-Thr150Ala-mut	CCAAGGTTAGAGAGTTGGCACTGAGACTGGTTCTCTGCTAAACAGTATGATAAAGAACAGTCA AGCAATTGTTGGCCAGTTCTGTGC
<i>BRCA1</i>	BRCA1-Thr150Ala-sil	CCAAGGTTAGAGAGTTGGCACTGAGACTGGTTCTCTGCTAAACAGTATGATAAAGAACAGTCA AGCAATTGTTGGCCAGTTCTGTGC
<i>BRCA1</i>	BRCA1-Tyr179Cys-mut	GCAATTATTATTAATACTTAAAAACCTGAGACCCTTACCTAATCAATGACAGACAGCTCTTT TGAGGTTGTATCCGCTGC
<i>BRCA1</i>	BRCA1-Tyr179Cys-sil	GCAATTATTATTAATACTTAAAAACCTGAGACCCTTACCTAATCAATGACAGACAGCTCTTT TGAGGTTGTATCCGCTGC
<i>BRCA1</i>	BRCA1-Val191Asp-mut	GGTTCTCTTGACTCACCTGCAATAAGTTGCTTTATTATCGGTATCTTCAGAAGAATCAGATCCTA AAAAATTTCCCCC
<i>BRCA1</i>	BRCA1-Val191Asp-sil	GGTTCTCTTGACTCACCTGCAATAAGTTGCTTTATTATCGGTATCTTCAGAAGAATCAGATCCTA AAAAATTTCCCCC
<i>BRCA1</i>	BRCA1-Cys197=-mut	CCAGCTTCATAGACAAAGGTTCTCTTGACTCACCTACAATAAGTTGCTTTATTAACGGTATCTTC GAAGAAATCAGATCC
<i>BRCA1</i>	BRCA1-Cys197=-sil	CCAGCTTCATAGACAAAGGTTCTCTTGACTCACCTGCAGTAAGTTGCTTTATTAACGGTATCTTC AGAAAGAAATCAGATCC
<i>BRCA1</i>	BRCA1-Gln210=-mut	GCCATTACCTTTTTTGCAGAATCCAACTGATTTATCCCTCGTCCCTGAGGGGTGATTTGTAA CAATCTTGATCTCCC
<i>BRCA1</i>	BRCA1-Gln210=-sil	GCCATTACCTTTTTTGCAGAATCCAACTGATTTATCCCTCGTCCCTGAGGGGTGATTTGTAA CAATCTTGATCTCCC
<i>BRCA1</i>	BRCA1-Leu291X-mut	CTACATTCATTCTGCTTTAGTGAGTCATAAACTGCTGTTCTCATGCTGTATGAGCTTGCATGAG TATTTGTGCCACATGGCTCC
<i>BRCA1</i>	BRCA1-Leu291X-sil	CTACATTCATTCTGCTTTAGTGAGTCATAAACTGCTGTTCTCATGCTGTATGAGCTTGCATGAG TATTTGTGCCACATGGCTCC
<i>BRCA1</i>	BRCA1-Tyr422X-mut	CAGTGGCCAGTAAGTCTATTTTCTGAAGAACCAGACTATTCTACTCTATTAGAACGTCCA ATACATCAGCTACTTTGGC
<i>BRCA1</i>	BRCA1-Tyr422X-sil	CAGTGGCCAGTAAGTCTATTTTCTGAAGAACCAGACTATTCTACTCTATTAGAACGTCCA ATACATCAGCTACTTTGGC
<i>BRCA1</i>	BRCA1-Gly462Arg-mut	CAGTTACATGGCTTAAGTTGGGGAGGCTTCTTTCTCCGATAGGTTTTCCAAATATTTGTCTC AATATTACTCTACTG
<i>BRCA1</i>	BRCA1-Gly462Arg-sil	CAGTTACATGGCTTAAGTTGGGGAGGCTTCTTTCTCCGATAGGTTTTCCAAATATTTGTCTC AATATTACTCTACTG
<i>BRCA1</i>	BRCA1-Arg979Cys-mut	CTTACATTTAGTTTTAACAAATGACTTGATGGGAAAAAGTGGTGTATACAATATGATTTTGT AAAGTCCATGTTATTTGG
<i>BRCA1</i>	BRCA1-Arg979Cys-sil	CTTACATTTAGTTTTAACAAATGACTTGATGGGAAAAAGTGGTGTATGCGATATGATTTTGT AAAGTCCATGTTATTTGG
<i>BRCA1</i>	BRCA1-Leu1080=-mut	GGAAGACTTTGTTTATAGACCTCAGGTTGCAAACTCCTAATCTAAGCATAGCATTAGTTTTGGC CCTCTGTTTCTACCTAGTTCTGC
<i>BRCA1</i>	BRCA1-Leu1080=-sil	GGAAGACTTTGTTTATAGACCTCAGGTTGCAAACTCCTAATCTAAGCATAGCATTAGTTTTGGC CCTCTGTTTCTACCTAGTTCTGC
<i>BRCA1</i>	BRCA1-Gly1201Ser-mut	GGACTCTAATTTCTGGCCCTCTCGGTATCCCTGAGCGAAATGTGTATGGGTGAAAGGGCTAG GACTCCTGCTAAGCTCTCC
<i>BRCA1</i>	BRCA1-Gly1201Ser-sil	GGACTCTAATTTCTGGCCCTCTCGGTATCCCTGAGCGAAATGTGTATGGGTGAAAGGGCTAG GACTCCTGCTAAGCTCTCC
<i>BRCA1</i>	BRCA1-Glu1250Lys-mut	GCTATTCTCAATGATAATAAATCTCCTCTGTGTTCTTAGACAGACACTGGTAGCAACAGTGCT ATGCCTAGTAGACTGAGAAGG
<i>BRCA1</i>	BRCA1-Glu1250Lys-sil	GCTATTCTCAATGATAATAAATCTCCTCTGTGTTCTTAGACAGACACTGGTAGCAACAGTGCT ATGCCTAGTAGACTGAGAAGG
<i>BRCA1</i>	BRCA1-Ile1275Val-mut	CCTCACTAAGGTGATGTTCTGAGATGCTTTGGCAATACTACCTGGTACTGCAGTCATTTAAGC TATTTCTCAATGATAATAAATCTCC
<i>BRCA1</i>	BRCA1-Ile1275Val-sil	CCTCACTAAGGTGATGTTCTGAGATGCTTTGGCAATACTACCTGGTACTGCAGTCATTTAAGC TATTTCTCAATGATAATAAATCTCC
<i>BRCA1</i>	BRCA1-Thr1394Ile-mut	GCAAAGGACACACACACGCGATGTGCACACACACACGCTTTTACCTGAATTGTTAAAAATG TCACTCTGAGAGGATAGCCC

BRCA1	BRCA1-Thr1394Ile-sil	GCAAAGGACACCACACACGCGATGTGCACACACACACGCTTTTACCTGAGTTGTTAAGATG TCACTCTGAGAGGATAGCCC
BRCA1	BRCA1-Asp1506Asn-mut	CTCCAGAGCAACTGTGCATGTACCACCTATTATCTAATGATGGACATTTAGAAGGGGATGACCT AGAAAGATAAATGGAAGG
BRCA1	BRCA1-Asp1506Asn-sil	CTCCAGAGCAACTGTGCATGTACCACCTGTCTAATGATGGACATTTAGAAGGGGATGACCT AGAAAGATAAATGGAAGG
BRCA1	BRCA1-Asp1506Glu-mut	CTCCAGAGCAACTGTGCATGTACCACCTCTCATCTAATGATGGACATTTAGAAGGGGATGACCT AGAAAGATAAATGGAAGG
BRCA1	BRCA1-Asp1506Glu-sil	CTCCAGAGCAACTGTGCATGTACCACCTGTCTAATGATGGACATTTAGAAGGGGATGACCT AGAAAGATAAATGGAAGG
BRCA1	BRCA1-Glu1586X-mut	GCAGAGTTGAAGATGGTATGTTGCCAACACGAGCTGACTATGGAGCTGTCTTCAGAAGGAT CAGATTCAGGGTCATCAGAG
BRCA1	BRCA1-Glu1586X-sil	GCAGAGTTGAAGATGGTATGTTGCCAACACGAGCTGATTCTGGAGCTGTCTTCAGAAGGAT CAGATTCAGGGTCATCAGAG
BRCA1	BRCA1-Gln1604X-mut	GCAGCAGCTGGACTCTGGGCAGATTCTGCAACTTCAATTAGGGAACCTTCAATGCAGACGTTGA AGATGGTATGTTGCCAACACGAGCTGACTCTGGGC
BRCA1	BRCA1-Gln1604X-sil	GCAGCAGCTGGACTCTGGGCAGATTCTGCAACTTCAACTGGGGAACCTTCAATGCAGACGTTGA AGATGGTATGTTGCCAACACGAGCTGACTCTGGGC
BRCA1	BRCA1-Gln1604=-mut	GCAGCAGCTGGACTCTGGGCAGATTCTGCAACTTCAACTGGGGAACCTTCAATGCAGACGTTGA AGATGGTATGTTGCCAACACGAGCTGACTCTGGGC
BRCA1	BRCA1-Gln1604=-sil	GCAGCAGCTGGACTCTGGGCAGATTCTGCAACTTCAATTAGGGAACCTTCAATGCAGACGTTGA AGATGGTATGTTGCCAACACGAGCTGACTCTGGGC
BRCA1	BRCA1-Intron20-mut	CCATTGACCACATCTCCTCTGACTTCAAATCATGCCGAAAGAAACCAACACAACCCATCAGCAT AAGAGAAAGAGAAGCTTCCCTCAATGG
BRCA1	BRCA1-Intron20-sil	CCATTGACCACATCTCCTCTGACTTCAAATCATGCTGAAAGAACCAACACAACCCATCAGCAT AAGAGAAAGAGAAGCTTCCCTCAATGG
BRCA1	BRCA1-Ala1752Thr-mut	GAGGGAGGGAGCTTACCTTTCTGCTTGGGATTCTCTTGGCTGCTTGGACCTTGGTGGTTCTT CCATTGACCACATCTCC
BRCA1	BRCA1-Ala1752Thr-sil	GAGGGAGGGAGCTTACCTTTCTGCTTGGGATTCTCTAGCTGCTTGGACCTTGGTGGTTCTT CCATTGACCACATCTCC
BRCA1	BRCA1-Ala1752Pro-mut	GAGGGAGGGAGCTTACCTTTCTGCTTGGGATTCTCTTGGCTGCTTGGACCTTGGTGGTTCTT CCATTGACCACATCTCC
BRCA1	BRCA1-Ala1752Pro-sil	GAGGGAGGGAGCTTACCTTTCTGCTTGGGATTCTCTAGCTGCTTGGACCTTGGTGGTTCTT CCATTGACCACATCTCC
BRCA1	BRCA1-Gly1770Val-mut	GGAACCTGGGGTTCTCCAGGCTTACCTGTGGGCATGTTGGTGAACGGCAGATAGCAACAG ATTTCTAGCCCCCTGAAGATCTGG
BRCA1	BRCA1-Gly1770Val-sil	GGAACCTGGGGTTCTCCAGGCTTACCTGTGGGCATGTTGGTGAACGGCAGATAGCAACAG ATTTCTAGCCCCCTGAAGATCTGG
BRCA1	BRCA1-Pro1812Ala-mut	GGAAGCCATTGTCTCTGTCAGGCATCCGCTGCACAACCAAAATGGATGGACACCCCTGGATC CCAGGAAGGAAAGAGCATT
BRCA1	BRCA1-Pro1812Ala-sil	GGAAGCCATTGTCTCTGTCAGGCATCCGCTGCACAACCAAAATGGATGGACACCCCTGGATC CCAGGAAGGAAAGAGCATT
BRCA2	BRCA2-Glu218Lys-mut	CAGTCAGAAATGAAGAAGCATCTAAAACGTATTTCTCCTCACGATACTACTGCTGTAAGTAAATAT GACATTGATTAGACTGTTG
BRCA2	BRCA2-Glu218Lys-sil	CAGTCAGAAATGAAGAAGCATCTGAGACTGATTTCTCCTCACGATACTACTGCTGTAAGTAAATAT GACATTGATTAGACTGTTG
BRCA2	BRCA2-Asn719Lys-mut	GTTATTTATTACCCAGAAGCTGATTCTGTGCATGCCTGCAAGAAGGACAGTGTGAAAAGGATC CAAAAAGCAAAAAGTTTCAG
BRCA2	BRCA2-Asn719Lys-sil	GTTATTTATTACCCAGAAGCTGATTCTGTGCATGCCTGCAAGAAGGACAGTGTGAAAAGGATC CAAAAAGCAAAAAGTTTCAG
BRCA2	BRCA2-Asp935Asn-mut	CGAACCCATTTCAAGAAGCTTACCATGGTTTTATATGGAGATACAGGTGACAAACAGCAACCC AAGTGTCAATTAAGAAAGATTGG
BRCA2	BRCA2-Asp935Asn-sil	CGAACCCATTTCAAGAAGCTTACCATGGTTTTATATGGAGATACAGGTGACAAACAGCAACCC AAGTGTCAATTAAGAAAGATTGG
BRCA2	BRCA2-Lys1180Arg-mut	GCTGATCTTCATGCATAATGAATGCTCCATCGATTGGTCAGGTAGACAGCAGCAGCAATTGA AGGTACAGTTGAAATTAACGG
BRCA2	BRCA2-Lys1180Arg-sil	GCTGATCTTCATGCATAATGAATGCTCCATCGATTGGTCAGGTAGACAGCAGCAAAATTTGA AGGTACAGTTGAAATTAACGG
BRCA2	BRCA2-Ser1882X-mut	CAGTAAAGTAATTAAGGAAAACAACGAGAATAAATAAAAAATTTGTCAAACGAAAATTATGGCA GGTTGTTACGAGGCATTGG
BRCA2	BRCA2-Ser1882X-sil	CAGTAAAGTAATTAAGGAAAACAACGAGAATAAATAAAAAATTTGTCAAACGAAAATTATGGCA GGTTGTTACGAGGCATTGG
BRCA2	BRCA2-Val2728Ile-mut	GATACCCAAAAAGTGGCCATTATTGAACTTACAGATGGGTGGTATGCTATTAAAGCCAGTTAGA TCCTCCCCTCTAGCTGTC
BRCA2	BRCA2-Val2728Ile-sil	GATACCCAAAAAGTGGCCATTATTGAACTTACAGATGGGTGGTATGCTGTAAGCCAGTTAG ATCCTCCCCTCTAGCTGTC
BRCA2	BRCA2-Gln2829Arg-mut	GGAAATGTTGGTTGTTGATGTAATTATCAAAGAGCATACCCCTATCCGGGTATGATGATTCTT GAACTTACCATATATTC

<i>BRCA2</i>	BRCA2-Gln2829Arg-sil	GGAAATGTTGGTTGTGTTGATGTAATTATTCAAAGAGCATATCCTATCCAGGTATGATGATTCTT GAAACTTACCATATATTTTC
<i>POLE</i>	POLE-Ala31Ser-mut	CCCTTCTTTCACTCAGGGATGATGGCGCCACTTCTCAGTTTCGTCACTCAAGCGCCTCGAACGGA GTCAGTGGACGGATAAGATGG
<i>POLE</i>	POLE-Ala31Ser-sil	CCCTTCTTTCACTCAGGGATGATGGCGCCACTTCTCAGTTTCGGCCCTCAAGCGCCTCGAACGGA GTCAGTGGACGGATAAGATGG
<i>POLE</i>	POLE-Pro286Ser-mut	GACATTGAGACGACAAACTGCCCTCAAGTTTTCTGATGCTGAGACAGATCAGATTATGATGAT TTCCTACATGATCGATGGCCAGG
<i>POLE</i>	POLE-Pro286Ser-sil	GACATTGAGACGACAAACTGCCCTCAAGTTTTCTGATGCTGAGACAGATCAGATTATGATGAT TTCCTACATGATCGATGGCCAGG
<i>POLE</i>	POLE-Asp301Gly-mut	GCTGAGACAGACCAGATTATGATGATTTCTGATGATCGGTGGCCAGGTGAGCAGGTGGCTTC TGGGAAGTAAGCTCCTGGG
<i>POLE</i>	POLE-Asp301Gly-sil	GCTGAGACAGACCAGATTATGATGATTTCTGATGATCGGTGGCCAGGTGAGCAGGTGGCTTC TGGGAAGTAAGCTCCTGGG
<i>POLE</i>	POLE-Asn363Asp-mut	GGTTTGAACACGTCCAGGAGACAAACCACCATCATGGTCACCTACGACGGAGACATTTTTGAC TG GTGAGTCTGTCTTC
<i>POLE</i>	POLE-Asn363Asp-sil	GGTTTGAACACGTCCAGGAGACAAACCACCATCATGGTCACCTACAATGGAGACTTTTTGAC TG GTGAGTCTGTCTTC
<i>POLE</i>	POLE-Leu424Val-mut	GGTGGGTGAAGAGGGACAGTTACCTTCTGTGGGCAGTCATAATGTCAAGGCCGCCAAGGC CAAGCTAGGCTATGATCCCG
<i>POLE</i>	POLE-Leu424Val-sil	GGTGGGTGAAGAGGGACAGTTACCTTCTGTGGGCAGTCATAATGTCAAGGCCGCCAAGGC CAAGCTAGGCTATGATCCCG
<i>POLE</i>	POLE-Ala456Pro-mut	GATGGCCCTGCTCTCTGGCGTTCTTCTCAGACTCTGCCACGTATTCTGTGTCAGATGCTGTG CCACTTACTACCTGTAC
<i>POLE</i>	POLE-Ala456Pro-sil	GATGGCCCTGCTCTCTGGCGTTCTTCTCAGACTCTGCCACATATTCTGTGTCAGATGCTGTG CCACTTACTACCTGTAC
<i>POLE</i>	POLE-Phe695Ile-mut	CAGCCAGTCGACGGAATACCATCGGATCCAGCACCAGCTCGAGTCAGAGAAGATCCCCCTT GTTCCAGAGGGGCCAGCTCGG
<i>POLE</i>	POLE-Phe695Ile-sil	CAGCCAGTCGACGGAATACCATCGGATCCAGCACCAGCTCGAGTCAGAGAAGATTTCCCCCTTG TTCCAGAGGGGCCAGCTCGG
<i>POLE</i>	POLE-Arg1826Trp-mut	CCCAGTACCACAACATCTATGCAGACAATCAGGTGATGCACTTCTACCGCTGGCTTGTGCGCCAT CCTCTCTGCTTCATGACCC
<i>POLE</i>	POLE-Arg1826Trp-sil	CCCAGTACCACAACATCTATGCAGACAATCAGGTGATGCACTTCTACCGCTGGCTTGTGCGCCAT CCTCTCTGCTTCATGACCC
Gene s	Names	Primers sequences
<i>LIG4</i>	LIG4-Fwd	CTGGAGAACAGAATTGCAGA
<i>LIG4</i>	LIG4-Rev	TAGCAATCATATTCACGGGC
<i>XRCC4</i>	XRCC4-Fwd	GAGGCCAGTACAGAAAACAT
<i>XRCC4</i>	XRCC4-Rev	TGGAAAAGTATCCCTGAGGA

Table S2: Odds of Pathogenicity (OddsPath) estimated by performance of classified variant controls.
 In this assay, five benign and five pathogenic variants of *BRCA1* and *BRCA2* were used as controls allowing the estimation of the evidence strength of our classification according to Brnich *et al*³⁰ guidelines.

# Classified			Proportion pathogenic (Prior, P ₁)	Assay result		"+1 proportions" (Posterior, P ₂)		OddsPath=[P ₂ x (1- P ₁)] / [(1- P ₂) x P ₁]	
Total	Pathogenic	Benign		Functionally abnormal	Functionally normal	Pathogenic	Benign	Pathogenic	Benign
10	5	5	0.50	5	5	0.83	0.17	5.0	0.200

Note: this estimate "adds 1.0" to the numerator and denominator in order to do a calculation

Odds of Pathogenicity (OddsPath)	Evidence strength equivalent
<0.0029	BS3_very_strong [^]
<0.053	BS3
<0.23	BS3_moderate*
<0.48	BS3_supporting
0.48-2.1	Indeterminate
>2.1	PS3_supporting
>4.3	PS3_moderate
>18.7	PS3
>350	PS3_very_strong

[^]There are no very strong benign evidence codes included in the Richards, et al. guidance. This category is included here for modeling purposes only.

*Since there are no moderate strength benign evidence codes included in the Richards, et al. guidance, moderate level evidence is equivalent to two instances of supporting-level benign evidence

Table S3: Variants used as silent controls or references with their classification according to different databases.

Benign mutations are shown in green (+), unreported mutations are shown in yellow (?), mutations with conflicting interpretation are shown in orange (+/-), pathogenic mutations are shown in red (-) and intermediate mutations are shown in blue (!).

Gene	Nucleotide change	Amino acid change	Exon	Silent control	Silent reference	UMD database	Clinvar database	BRCA exchange	Characterized variant
BRCA1	c.93C>T	Ile31=	3			?	?	?	Ile31Asn
BRCA1	c.96G>A	Lys32=	3			?	?	?	Ile31Asn
BRCA1	c.442-12C>T		3			?	?	?	Glu149Ala, Thr150Ala
BRCA1	c.450C>G	Thr150=	3			?	?	?	Glu149Ala, Thr150Ala
BRCA1	c.537C>T	Tyr179=	8			?	?	?	Tyr179Cys
BRCA1	c.546G>A	Leu182=	8			?	?	?	Tyr179Cys
BRCA1	c.573T>A	Val191=	8			?	?	?	Val191Asp
BRCA1	c.579G>A	Lys193=	8			?	+	+	Val191Asp, Cys197Cys
BRCA1	c.591C>T	Tyr196=	8			+	+	+	Cys197Cys
BRCA1	c.633A>T	Gly211=	10			?	?	?	Gln210Gln
BRCA1	c.636C>G	Thr212=	10			?	?	?	Gln210Gln
BRCA1	c.840C>A	Ala280=	11			?	?	?	Leu291X
BRCA1	c.873A>G	Leu291=	11			?	?	?	Leu291X
BRCA1	c.1254G>A	Glu418=	11			?	+	+	Tyr422X
BRCA1	c.1266T>C	Tyr422=	11			?	?	?	Tyr422X
BRCA1	c.1386G>A	Gly462=	11			?	+	+	Gly462Arg
BRCA1	c.1404G>A	Lys468=	11			?	+	?	Gly462Arg
BRCA1	c.2928C>T	Asn976=	11			?	?	?	Arg979Cys
BRCA1	c.2937T>C	Arg979=	11			?	?	?	Arg979Cys
BRCA1	c.3243T>C	Asn1081=	11			?	+	+	Leu1080=
BRCA1	c.3261G>A	Gly1087=	11			?	?	?	Leu1080=
BRCA1	c.3594C>G	Leu1198=	11			?	?	?	Gly1201Ser
BRCA1	c.3603T>A	Gly1201=	11			?	?	?	Gly1201Ser
BRCA1	c.3738C>T	Thr1246=	11			?	+	+	Glu1250Lys
BRCA1	c.3750G>A	Glu1250=	11			?	+	+	Glu1250Lys
BRCA1	c.3825A>C	Ile1275=	11			?	?	?	Ile1275Val
BRCA1	c.3834G>A	Lys1278=	11			?	+	+	Ile1275Val
BRCA1	c.4173G>A	Ile1391=	12			?	?	?	Thr1394Ile
BRCA1	c.4179C>A	Thr1393=	12			?	?	?	Thr1394Ile
BRCA1	c.4503C>T	Cys1501=	15			+/-	+	+	Asp1506Asn, Asp1506Glu
BRCA1	c.4518T>C	Asp1506=	15			?	+	+	Asp1506Asn, Asp1506Glu
BRCA1	c.4752C>T	Ala1584=	15			?	+	+	Glu1586X

BRCA1	c.4758G>A	Glu1586=	15			?	?	?	Glu1586X
BRCA1	c.4791C>G	Thr1597=	16			?	?	?	Gln1604X, Gln1604=
BRCA1	c.4809C>A	Pro1603=	16			?	?	?	Gln1604=
BRCA1	c.4812A>G	Gln1604=	16			+	+	+	Gln1604X
BRCA1	c.5194-28C>G		20			?	?	?	c.5194-2A>G
BRCA1	c.5194-10T>G		20			?	?	?	c.5194-2A>G
BRCA1	c.5256A>T	Ala1752=	20			?	?	?	Ala1752Thr, Ala1752Pro
BRCA1	c.5268G>A	Gln1756=	20			?	+	+	Ala1752Thr, Ala1752Pro
BRCA1	c.5310G>T	Gly1770=	21			?	+	+	Gly1770Val
BRCA1	c.5313C>G	Pro1771=	21			?	?	?	Gly1770Val
BRCA1	c.5415C>T	His1805=	23			?	+	?	Pro1812Ala
BRCA1	c.5436A>G	Pro1812=	23			+/-	?	?	Pro1812Ala
BRCA2	c.654A>G	Glu218=	8			?	?	?	Glu218Lys
BRCA2	c.669T>C	His223=	8			?	?	?	Glu218Lys
BRCA2	c.2139G>A	Gln713=	11			?	?	?	Asn719Lys
BRCA2	c.2157T>C	Asn719=	11			?	?	?	Asn719Lys
BRCA2	c.2796C>T	Asp932=	11			?	?	?	Asp935Asn
BRCA2	c.2805T>C	Asp935=	11			?	+	+	Asp935Asn
BRCA2	c.3510C>T	Ala1170=	11			+/-	?	?	Lys1180Arg
BRCA2	c.3540G>A	Lys1180=	11			?	?	?	Lys1180Arg
BRCA2	c.5646A>C	Ser1882=	11			?	?	?	Ser1882X
BRCA2	c.5655C>T	Cys1885=	11			?	?	?	Ser1882X
BRCA2	c.8184T>A	Val2728=	18			?	?	?	Val2728Ile
BRCA2	c.8187G>A	Lys2729=	18			?	+	+	Val2728Ile
BRCA2	c.8478C>T	Tyr2826=	19			?	?	?	Gln2829Arg
BRCA2	c.8484A>C	Ile2828=	19			?	?	?	Gln2829Arg
Gene	Nucleotide change	Amino acid change	Exon	Silent control	Silent reference	Exac	Clinvar database	Oncokb	Characterized variant
POLE	c.93A>C	Ala31=	2			+	+	?	Ala31Ser
POLE	c.105G>C	Leu35=	2			+	?	?	Ala31Ser
POLE	c.864T>C	Ala288=	9			?	+	?	Pro286Ser
POLE	c.874C>T	Asp291=	9			?	?	?	Pro286Ser
POLE	c.891C>G	Ser297=	9			?	?	?	Asp301Gly
POLE	c.900C>T	Ile300=	9			?	+	?	Asp301Gly
POLE	c.1089C>T	Asn363=	11			+	+	?	Asn363Arg
POLE	c.1092G>A	Gly364=	11			+	+	?	Asn363Arg
POLE	c.1272C>G	Leu424=	13			?	+	?	Leu424Val
POLE	c.1278G>C	Ala426=	13			?	+	?	Leu424Val
POLE	c.1360-6C>T		14			+	+	?	Ala456Pro

POLE	c.1371G>A	Thr457=	14			+	+	?	Ala456Pro
POLE	c.2070G>C	Leu690=	19			?	?	?	Phe695Ile
POLE	c.2085C>T	Phe695=	19			?	+/-	?	Phe695Ile
POLE	c.5448C>T	Asn1815=	40			?	?	?	Arg1826Trp
POLE	c.5478G>T	Arg1826=	40			+	+	?	Arg1826Trp

DISCUSSION ARTICLE 3

Suite aux premiers résultats que nous avons obtenus dans les lignées diploïdes et aux résultats de SGE publiés par Findlay et son équipe^{366,367}, la lignée cellulaire HAP1 a été sélectionnée comme modèle d'étude. Dans un premier temps, nous avons réalisé une lignée haploïde polyclonale KO du gène *LIG4* essentiel au fonctionnement de la voie de réparation des jonctions non homologues alternative à la RH lors de cassures double brin de l'ADN. Cette modification permet de favoriser la RH, augmentant l'efficacité d'édition et limitant les effets hors cibles.³⁷⁵

Dans notre modèle, les gènes *BRCA1* et *BRCA2* sont essentiels, l'insertion de mutations pathogènes entraîne donc la mort des cellules éditées, facilitant le criblage et donnant une indication quant à sa conséquence fonctionnelle. Afin d'avoir un contrôle d'édition, nous réalisons en parallèle l'édition d'une mutation silencieuse, évaluée non pathogène dans les bases de données, en présence du même ARNg que la mutation d'intérêt. Enfin, permettant à la fois la comparaison de l'efficacité dans les deux conditions et une amélioration de l'efficacité d'édition par limitation des phénomènes de réédition, une seconde mutation silencieuse est insérée au niveau du PAM ou 3 nucléotides en amont.

Cette approche a permis l'étude de 33 variants *BRCA* : 10 variants d'interprétation connue (bénins ou pathogènes) utilisés comme contrôles et 23 VSI (variants non classifiés ou d'impact fonctionnel indéterminé). A l'issue de l'édition et du séquençage de ces 23 VSI, le score fonctionnel que nous proposons a rapporté 15 variants fonctionnellement normaux, 6 variants fonctionnellement anormaux et 2 variants de conséquences « intermédiaires » (p.Leu1080= et c.5194-2A>G). Cependant, ce score nécessite d'être consolidé sur un échantillonnage plus important pour être plus discriminant, cela pourrait permettre de classer le variant p.Leu1080= qui est à la limite de la pathogénie. Le variant c.5194-2A>G avait été testé dans l'article 2 montrant un saut d'exon 20 de *BRCA1* mais un impact transcriptionnel mineur corrélé à la rechute précoce sous iPARP. Cette observation suggère à nouveau l'importance de l'étude des niveaux d'expression des transcrits aberrants.

L'essentialité des gènes suppresseurs de tumeurs est une caractéristique centrale de notre modèle qui peut être appliqué à d'autres gènes essentiels. Pour illustrer cela, nous avons étendu notre approche à l'étude de variants du gène *POLE* codant la polymérase epsilon. En effet, les altérations pathogènes du domaine exonucléase du gène *POLE* ont aujourd'hui un double enjeu en génétique tumorale. D'une part ces altérations *POLE*, souvent accompagnées d'un phénotype tumoral hypermuté³⁷⁶, pourraient constituer un biomarqueur d'une sensibilité accrue aux inhibiteurs de *checkpoints* immunitaires³⁷⁷. D'autre part, ces mutations *POLE* participent à la nouvelle classification moléculaire des

cancers de l'endomètre basées sur les données TCGA, les tumeurs *POLE* mutées constituant une classe à part entière associée à un bon pronostic et à des recommandations de désescalade thérapeutique.^{378,379} En génétique constitutionnelle, ces variations pathogènes de *POLE* ont été associées à une prédisposition héréditaire au cancer colorectal.³⁸⁰ Sept variants de *POLE* ont été testés : 4 variants contrôles déjà classifiés et 3 VSI. Bien que notre modèle semble tout à fait applicable à la caractérisation des variants de *POLE*, une discordance a été relevée pour la mutation p.Leu424Val : classée pathogène dans les bases de données ClinVar et OncoKB, notre approche a rapporté un score plutôt associé à un impact fonctionnel neutre pour ce variant. Il a été observé que le changement d'AA au niveau de la leucine 424 n'était pas toujours associé à un phénotype hypermuté.³⁸¹ L'absence d'hypermutation en cas de changement leucine en valine pourrait expliquer pourquoi l'édition n'est pas une conséquence létale dans notre modèle. Cette discordance ne remet donc pas en question l'impact de cette mutation sur la prédisposition au cancer colorectal. Une autre hypothèse pouvant expliquer cette discordance serait un temps d'incubation insuffisant suite aux transfections successives, ne permettant pas l'accumulation de mutations conduisant à la mort cellulaire.

La mise au point de cette méthodologie permet donc la caractérisation fonctionnelle de VSI de tous les gènes suppresseurs de tumeur essentiels dans un modèle unique et ce dans un délai de 3 semaines compatible avec les impératifs cliniques.

L'une des principales contraintes actuelles de notre test fonctionnel est l'obligation de disposer d'un motif PAM suffisamment proche de la mutation d'intérêt pour chaque cible. Les technologies CRISPR et leurs nucléases programmables guidées par l'ARNg ne cessent de se développer et offrent une flexibilité de plus en plus importante pour le ciblage de l'ADN. Le répertoire de nucléases disponibles s'étoffe progressivement avec de nombreuses séquences PAM possibles, permettant d'augmenter le champ des variants que nous pourrions évaluer à l'avenir. Récemment des nucléases fonctionnant sans reconnaissance PAM ont été développées³⁸² de même que des fusions de Cas avec des retrotranscriptases permettant d'éditer fidèlement une séquence sans coupure double brin de l'ADN génomique ou « *Prime editing* » pour ne citer que quelques innovations autour de cette rupture technologique. Ainsi, il est maintenant clair que la versatilité de l'édition génomique va prendre une place incontournable dans l'analyse fonctionnelle de variants en cancérologie au sens large.

DISCUSSION GENERALE ET PERSPECTIVES

Depuis une quinzaine d'années, la médecine personnalisée est devenue réalité et s'est considérablement développée en oncologie. Les thérapies ciblées font désormais partie de l'arsenal thérapeutique au même titre que la chirurgie, la radiothérapie et les chimiothérapies.

Les enjeux actuels de la médecine de précision sont multiples :

- Prédire la réponse au traitement : au-delà de l'anomalie moléculaire ciblée, il est essentiel d'identifier d'autres facteurs prédictifs de la réponse au traitement afin d'administrer le médicament uniquement aux patients qui en tireront un bénéfice pour éviter toute perte de temps, pour réduire le risque d'effets indésirables et éviter des coûts inutiles. Ces facteurs peuvent être cliniques ou encore biologiques (caractéristiques moléculaires, statut immunologique, signatures...). Afin de les révéler, certaines études à grande échelle ont été lancées, internationales (TCGA, International cancer genome consortium ou ICGC) ou nationales (Plan France Médecine Génomique).
- Lutter contre les résistances : les thérapies ciblées ont souvent une durée d'action limitée. Ces résistances sont très certainement le résultat de mécanismes adaptatifs rendus possibles par l'hétérogénéité des cellules cancéreuses qui coexistent au sein d'une même tumeur. L'identification de ces résistances et la prédiction évolutive des tumeurs sous pression des traitements devraient permettre de prévoir la stratégie thérapeutique et d'offrir de nouvelles cibles thérapeutiques.
- Augmenter l'efficacité et la tolérance : des équipes travaillent à l'amélioration des thérapies ciblées déjà disponibles pour les rendre encore plus spécifiques de leur cible, accroître leur efficacité et réduire leurs effets secondaires.
- Mieux cibler les tumeurs : la connaissance de l'environnement tumoral et les avancées pharmacologiques ont permis l'émergence des anticorps conjugués (ADC, Antibody Drug Conjugates) dont des dizaines sont actuellement en développement. Le principe repose sur des anticorps spécifiques d'antigènes tumoraux utilisés comme vecteurs afin de diriger une chimiothérapie au sein de la tumeur.
- Améliorer la prévention : la médecine de précision permet d'identifier une éventuelle prédisposition à une maladie et ainsi d'offrir une prévention adaptée et personnalisée.

Cette thèse a été focalisée sur l'identification de marqueurs moléculaires donnant accès à des thérapies ciblées reposant sur deux principes fondamentaux : l'addiction oncogénique et la létalité synthétique. Deux principaux volets ont été abordés : les méthodologies de détection des acides nucléiques biomarqueurs, le choix technique étant directement en lien avec la qualité et le délai de rendu du résultat, et la caractérisation fonctionnelle de variations génétiques, conditionnant l'efficacité des thérapies ciblées.

La première partie de ce travail, avec l'exemple des variants *EGFR* et du CBNPC, a illustré à quel point les technologies NGS maximisent les chances du patient de bénéficier

d'une thérapie ciblée, plus efficace et moins toxique que les chimiothérapies classiques. En effet, Le NGS offre une grande sensibilité analytique, permettant la détection de cibles dans des échantillons de tissus pauvres en cellules tumorales et dans l'ADNtc. Par ailleurs, le NGS donne l'information la plus complète par l'exhaustivité des cibles recherchées simultanément et l'opportunité de travailler sur ADN et ARN. Après l'essor des thérapies ciblées, cette généralisation du NGS et la production phénoménale de données génomiques et transcriptomiques a donné naissance à une deuxième révolution en cancérologie avec l'arrivée du *big data* moléculaire et l'émergence de bases de données de grande dimension, telles que TCGA. Ces bases de données issues du « profilage moléculaire » centralisent des données « omics » (génomique, transcriptomique, épigénomique, protéomique, métabolomique) dans toutes les localisations tumorales. Ces progrès permettent déjà d'identifier des altérations d'intérêt diagnostique, pronostique ou thérapeutique.

La deuxième partie de cette thèse s'est intéressé à la caractérisation fonctionnelle des VSI, en prenant l'exemple des gènes *BRCA* et de la prescription des thérapies ciblées par iPARP. Nous avons constaté une implémentation progressive d'outils prédictifs *in silico*. Des outils d'intelligence artificielle ont récemment été proposés pour caractériser classer des variants *BRCA*. Ils intègrent des données de la littérature, la structure du gène et de la protéine, des données de conservation entre les espèces. En utilisant le « *machine-learning* » sur la base de données ClinVar, Favalli et *al.* ont ainsi développé un algorithme de classification de variants.³⁸³ Cependant, pour veiller à l'homogénéité de l'interprétation des variants dans une même thématique, il serait intéressant de s'accorder sur les outils de prédiction à interroger et des scores de confiance qu'on peut leur attribuer. Associé à la prédiction, des tests biologiques fonctionnels sont donc indispensables pour évaluer ces VSI et affiner les algorithmes d'évaluation des scores de pathogénicité. Des avancées technologiques majeures, telles que l'ingénierie ciblée du génome par l'utilisation des technologies CRISPR-Cas9 et ses déclinaisons multiples, dont les limites d'utilisation diminuent à une vitesse vertigineuse, sont extrêmement prometteuses en particulier en génomique fonctionnelle. Enfin, face à un nombre croissant de VSI mis en évidence dans les tumeurs, des bases de données de variants somatiques annotés et validés ouvertes à la communauté scientifique se mettent en place. Par exemple la base française SoVad, initiée par le GFCO, a pour objectif de permettre l'homogénéisation des interprétations des variants somatiques détectés sur les gènes ayant une application dans la prise en charge des tumeurs solides.

L'optimisation de cette médecine personnalisée est un des objectifs de la stratégie décennale de lutte contre les cancers 2021-2030 insistant pour que l'innovation bénéficie à tous les patients (AXE 4 du Plan Cancer). En effet, malgré la caractérisation de plus en plus performante des cibles pharmacologiques moléculaires des tumeurs comme critère de

prescription, il persiste une proportion importante de patients qui restent insensibles d'emblée ou dont la sensibilité évolue négativement au cours du traitement. Cette remarque indique clairement que l'aspect prédictif de réponse à une thérapie ciblée sur la base de caractéristiques génétiques ciblées est largement insuffisante. Cette limitation est ainsi responsable de perte de chance et d'une dégradation de qualité de vie pour nombre de patients. Elle induit par ailleurs des coûts inutiles de traitement. On peut en outre noter que durant ces dernières années, malgré l'extrême puissance de la production de données génomiques, celle-ci se limite en pratique clinique au maximum à l'exome soit 2 à 3% du génome et en général aux altérations génomiques altérant les séquences codantes des protéines, les sites d'épissage des ARN voire quelques marques épigénétiques. Même si ces étapes ont prouvé leur utilité pour la prise en charge des patients, il est illusoire de penser que la recherche des altérations ciblées et correctement annotées avec quelques données clinico-biologiques seront suffisantes pour prédire la réponse ou la résistance tumorale à une thérapie. A l'avenir, l'analyse génomique moléculaire des tumeurs et du patient devra prendre en compte les mutations silencieuses ainsi que les altérations non codantes présentes dans l'ensemble du génome. Par exemple concernant les mutations silencieuses présentes dans les séquences codantes, la redondance du code génétique et la préférence d'utilisation de ce dernier par les cellules génèrent un autre niveau d'information génétique qui peut être relié à des altérations cellulaires directes ou indirectes de plus ou moins grande importance sur le comportement de la tumeur. De même, l'ensemble des altérations génomiques tumorales et/ou constitutionnelles non codantes identifiées comme le « *dark genome* » ou la « *genomic dark matter* » est clairement un niveau d'information largement sous exploité. Ainsi pour l'avenir on peut imaginer que l'analyse comparative somatique et constitutionnelle de génome entier associée à celle du transcriptome pourra permettre d'améliorer significativement la personnalisation thérapeutique en cancérologie.

L'exploitation conjointe par les méthodes d'intelligence artificielle de ces données multiomiques en masse qui sont en constante progression n'est qu'à son début, et il ne fait aucun doute que dans un premier temps elle fera progresser la connaissance biologique intime des tumeurs au sein du patient. Dans un second temps, de concert avec l'ensemble des multiples caractéristiques clinico-biologiques de chaque patient, ces nouvelles méthodologies feront progresser la prise en charge individualisée des malades dans le but ultime de guérir les patients et sans doute d'améliorer la prévention de la cancérogénèse.

BIBLIOGRAPHIE

1. Burrell, R. A., McGranahan, N., Bartek, J. & Swanton, C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* 501, 338–345 (2013).
2. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* 68, 394–424 (2018).
3. Bray, F., Laversanne, M., Weiderpass, E. & Soerjomataram, I. The ever-increasing importance of cancer as a leading cause of premature death worldwide. *Cancer* 127, 3029–3030 (2021).
4. Ferlay, J. *et al.* Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* 136, E359–E386 (2015).
5. INCa. Panorama des cancers en France (Edition 2021). <https://www.e-cancer.fr/Comprendre-prevenir-depister/Qu-est-ce-qu-un-cancer/Chiffres-cles> (2021).
6. van't Veer, L. J. & Bernards, R. Enabling personalized cancer medicine through analysis of gene-expression patterns. *Nature* 452, 564–570 (2008).
7. NIH. National Institute of Health. Site internet. <https://www.nih.gov>.
8. FDA. US Food and Drug Administration. Site internet. <https://www.fda.gov>.
9. INCa. Institut National du Cancer. Site internet. <https://www.e-cancer.fr/Patients-et-proches/Se-faire-soigner/Traitements/Therapies-ciblees-et-immunotherapie-specifique/Medecine-de-precision>.
10. Sawyers, C. Targeted cancer therapy. *Nature* 432, 294–297 (2004).
11. Köhler, G. & Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256, 495–497 (1975).
12. Pegram, M. D., Pauletti, G. & Slamon, D. J. HER-2/neu as a predictive marker of response to breast cancer therapy. *Breast Cancer Res. Treat.* 52, 65–77 (1998).
13. O'Brien, S. G. *et al.* Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. *N. Engl. J. Med.* 348, 994–1004 (2003).
14. Verweij, J. *et al.* Progression-free survival in gastrointestinal stromal tumours with high-dose imatinib: randomised trial. *Lancet Lond. Engl.* 364, 1127–1134 (2004).
15. Blanke, C. D. *et al.* Long-term results from a randomized phase II trial of standard- versus higher-dose imatinib mesylate for patients with unresectable or metastatic gastrointestinal stromal tumors expressing KIT. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 26, 620–625 (2008).
16. INCa. Les thérapies ciblées dans le traitement du cancer en 2015, état des lieux et enjeux. at (2015).

17. Bedard, P. L., Hyman, D. M., Davids, M. S. & Siu, L. L. Small molecules, big impact: 20 years of targeted therapy in oncology. *The Lancet* 395, 1078–1088 (2020).
18. Accueil - ANSM. <https://ansm.sante.fr/>.
19. Portail HAS Professionnels. *Haute Autorité de Santé* https://www.has-sante.fr/jcms/fc_2873790/fr/professionnels.
20. VIDAL, L'intelligence médicale au service du soin. VIDAL <https://www.vidal.fr/>.
21. Pan, S.-T., Li, Z.-L., He, Z.-X., Qiu, J.-X. & Zhou, S.-F. Molecular mechanisms for tumour resistance to chemotherapy. *Clin. Exp. Pharmacol. Physiol.* 43, 723–737 (2016).
22. Lord, C. J. & Ashworth, A. The DNA damage response and cancer therapy. *Nature* 481, 287–294 (2012).
23. Bozic, I. *et al.* Accumulation of driver and passenger mutations during tumor progression. *Proc. Natl. Acad. Sci. U. S. A.* 107, 18545–18550 (2010).
24. Martincorena, I. *et al.* Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* 171, 1029–1041.e21 (2017).
25. Chalmers, Z. R. *et al.* Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* 9, 34 (2017).
26. Yarchoan, M., Hopkins, A. & Jaffee, E. M. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. *N. Engl. J. Med.* 377, 2500–2501 (2017).
27. Lawrence, M. S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499, 214–218 (2013).
28. Wadman, M. James Watson's genome sequenced at high speed. *Nature* 452, 788 (2008).
29. INCa. Listes de gènes minimales à analyser dans le cadre d'un usage à visée diagnostique du NGS. at (2016).
30. Schwarzenbach, H., Hoon, D. S. B. & Pantel, K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat. Rev. Cancer* 11, 426–437 (2011).
31. Lemoine, A. *et al.* Recommandations du GFCO pour l'utilisation diagnostique des analyses génétiques somatiques sur l'ADN tumoral circulant. *Innov. Thérapeutiques En Oncol.* 2, 225–232 (2016).
32. Shu, Y. *et al.* Circulating Tumor DNA Mutation Profiling by Targeted Next Generation Sequencing Provides Guidance for Personalized Treatments in Multiple Cancer Types. *Sci. Rep.* 7, 583 (2017).
33. Hess, J. F. *et al.* Library preparation for next generation sequencing: A review of automation strategies. *Biotechnol. Adv.* 41, 107537 (2020).
34. van Dijk, E. L., Jaszczyszyn, Y., Naquin, D. & Thermes, C. The Third Revolution in Sequencing Technology. *Trends Genet. TIG* 34, 666–681 (2018).

35. Nurk, S. *et al.* The complete sequence of a human genome. *Science* 376, 44–53 (2022).
36. Eid, J. *et al.* Real-time DNA sequencing from single polymerase molecules. *Science* 323, 133–138 (2009).
37. Lu, H., Giordano, F. & Ning, Z. Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics Proteomics Bioinformatics* 14, 265–279 (2016).
38. Leong, I. U. S., Skinner, J. R. & Love, D. R. Application of Massively Parallel Sequencing in the Clinical Diagnostic Testing of Inherited Cardiac Conditions. *Med. Sci.* 2, 98–126 (2014).
39. Khodakov, D., Wang, C. & Zhang, D. Y. Diagnostics based on nucleic acid sequence variant profiling: PCR, hybridization, and NGS approaches. *Adv. Drug Deliv. Rev.* 105, 3–19 (2016).
40. Nomenclature for the description of sequence variants: Human Genome Variation Society. at <https://www.hgvs.org/mutnomen> (2013).
41. Plon, S. E. *et al.* Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum. Mutat.* 29, 1282–1291 (2008).
42. Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med. Off. J. Am. Coll. Med. Genet.* 17, 405–424 (2015).
43. Amendola, L. M. *et al.* Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium. *Am. J. Hum. Genet.* 98, 1067–1076 (2016).
44. Réseau NGS-Diag. Homogénéisation de l'interprétation de variants de séquence. (2018).
45. Koepfel, F. *et al.* Standardisation of pathogenicity classification for somatic alterations in solid tumours and haematologic malignancies. *Eur. J. Cancer Oxf. Engl.* 1990 159, 1–15 (2021).
46. 1000 Genomes Project - IGSR: The International Genome Sample Resource. at <https://www.internationalgenome.org/>.
47. ESP (NHLBI Exome Sequencing Project). at <https://esp.gs.washington.edu>.
48. FLOSSIES (Fabulous Ladies Over Seventy). at <https://whi.color.com/about>.
49. GnomAD (Genome Aggregation Database). at <https://gnomad.broadinstitute.org/>.
50. Tate, J. G. *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* 47, D941–D947 (2019).
51. Landrum, M. J. *et al.* ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* 46, D1062–D1067 (2018).

52. TCGA (The Cancer Genome Atlas). at <https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>.
53. Chakravarty, D. *et al.* OncoKB: A Precision Oncology Knowledge Base. *JCO Precis. Oncol.* 2017, (2017).
54. Kusnoor, S. V. *et al.* My Cancer Genome: Evaluating an Educational Model to Introduce Patients and Caregivers to Precision Medicine Information. *AMIA Jt. Summits Transl. Sci. Proc. AMIA Jt. Summits Transl. Sci.* 2016, 112–121 (2016).
55. Pd, S. *et al.* Human Gene Mutation Database (HGMD): 2003 update. *Hum. Mutat.* 21, (2003).
56. If, F. *et al.* LOVD v.2.0: the next generation in gene variant databases. *Hum. Mutat.* 32, (2011).
57. OMIM - Online Mendelian Inheritance in Man. <https://www.omim.org/>.
58. Ainscough, B. J. *et al.* DoCM: a database of curated mutations in cancer. *Nat. Methods* 13, 806–807 (2016).
59. Griffith, M. *et al.* CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nat. Genet.* 49, 170–174 (2017).
60. Tamborero, D. *et al.* Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med.* 10, 25 (2018).
61. ICGC Data Portal. at <https://dcc.icgc.org/>.
62. Ifac, F. *et al.* The LOVD3 platform: efficient genome-wide sharing of genetic variants. *Eur. J. Hum. Genet. EJHG* 29, (2021).
63. Vaser, R., Adusumalli, S., Leng, S. N., Sikic, M. & Ng, P. C. SIFT missense predictions for genomes. *Nat. Protoc.* 11, 1–9 (2016).
64. Schwarz, J. M., Rödelsperger, C., Schuelke, M. & Seelow, D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods* 7, 575–576 (2010).
65. Schwarz, J. M., Cooper, D. N., Schuelke, M. & Seelow, D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods* 11, 361–362 (2014).
66. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* 7, 248–249 (2010).
67. Salgado, D. *et al.* UMD-Predictor: A High-Throughput Sequencing Compliant System for Pathogenicity Prediction of any Human cDNA Substitution. *Hum. Mutat.* 37, 439–446 (2016).
68. Yeo, G. & Burge, C. B. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol. J. Comput. Mol. Cell Biol.* 11, 377–394 (2004).
69. BDGP: Splice Site Prediction by Neural Network (NNSplice). at <https://www.fruitfly.org/>.

70. Pertea, M., Lin, X. & Salzberg, S. L. GeneSplicer: a new computational method for splice site prediction. *Nucleic Acids Res.* 29, 1185–1190 (2001).
71. Desmet, F.-O. *et al.* Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 37, e67 (2009).
72. Jaganathan, K. *et al.* Predicting Splicing from Primary Sequence with Deep Learning. *Cell* 176, 535-548.e24 (2019).
73. Leman, R. *et al.* SPiP: Splicing Prediction Pipeline, a machine learning tool for massive detection of exonic and intronic variant effect on mRNA splicing. (2021). *in preparation* doi:10.22541/au.164544915.57104749/v1.
74. Leman, R. *et al.* SpliceLauncher: a tool for detection, annotation and relative quantification of alternative junctions from RNAseq data. *Bioinforma. Oxf. Engl.* 36, 1634–1636 (2020).
75. Mateo, J. *et al.* A framework to rank genomic alterations as targets for cancer precision medicine: the ESMO Scale for Clinical Actionability of molecular Targets (ESCAT). *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* 29, 1895–1902 (2018).
76. Skoulidis, F. *et al.* STK11/LKB1 Mutations and PD-1 Inhibitor Resistance in KRAS-Mutant Lung Adenocarcinoma. *Cancer Discov.* 8, 822–835 (2018).
77. Herbst, R. S., Heymach, J. V. & Lippman, S. M. Lung cancer. *N. Engl. J. Med.* 359, 1367–1380 (2008).
78. Gandini, S. *et al.* Tobacco smoking and cancer: a meta-analysis. *Int. J. Cancer* 122, 155–164 (2008).
79. Wirth, N., Abou-Hamdan, K., Spinosa, A., Bohadana, A. & Martinet, Y. Le tabagisme passif. *Rev. Pneumol. Clin.* 61, 7–15 (2005).
80. Defossez, G. *et al.* Estimations nationales de l'incidence et de la mortalité par cancer en France métropolitaine entre 1990 et 2018. Synthèse. at (2019).
81. Goldstraw, P. *et al.* The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* 11, 39–51 (2016).
82. Antonia, S. J. *et al.* Durvalumab after Chemoradiotherapy in Stage III Non-Small-Cell Lung Cancer. *N. Engl. J. Med.* 377, 1919–1929 (2017).
83. Antonia, S. J. *et al.* Overall Survival with Durvalumab after Chemoradiotherapy in Stage III NSCLC. *N. Engl. J. Med.* 379, 2342–2350 (2018).
84. Brunelli, A. *et al.* ERS/ESTS clinical guidelines on fitness for radical therapy in lung cancer patients (surgery and chemo-radiotherapy). *Eur. Respir. J.* 34, 17–41 (2009).
85. Román, M. *et al.* KRAS oncogene in non-small cell lung cancer: clinical perspectives on the treatment of an old target. *Mol. Cancer* 17, 33 (2018).

86. Dogan, S. *et al.* Molecular epidemiology of EGFR and KRAS mutations in 3,026 lung adenocarcinomas: higher susceptibility of women to smoking-related KRAS-mutant cancers. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 18, 6169–6177 (2012).
87. Hong, D. S. *et al.* KRASG12C Inhibition with Sotorasib in Advanced Solid Tumors. *N. Engl. J. Med.* 383, 1207–1217 (2020).
88. Tsao, A. S. *et al.* Scientific Advances in Lung Cancer 2015. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* 11, 613–638 (2016).
89. Zhang, X.-T. *et al.* The EGFR mutation and its correlation with response of gefitinib in previously treated Chinese patients with advanced non-small-cell lung cancer. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* 16, 1334–1342 (2005).
90. Koivunen, J. P. *et al.* EML4-ALK fusion gene and efficacy of an ALK kinase inhibitor in lung cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 14, 4275–4283 (2008).
91. Wong, D. W.-S. *et al.* The EML4-ALK fusion gene is involved in various histologic types of lung cancers from nonsmokers with wild-type EGFR and KRAS. *Cancer* 115, 1723–1733 (2009).
92. Shaw, A. T. *et al.* Clinical features and outcome of patients with non-small-cell lung cancer who harbor EML4-ALK. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 27, 4247–4253 (2009).
93. Takahashi, T. *et al.* Clinicopathologic features of non-small-cell lung cancer with EML4-ALK fusion gene. *Ann. Surg. Oncol.* 17, 889–897 (2010).
94. Soda, M. *et al.* Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature* 448, 561–566 (2007).
95. Gerber, D. E. & Minna, J. D. ALK inhibition for non-small cell lung cancer: from discovery to therapy in record time. *Cancer Cell* 18, 548–551 (2010).
96. Sullivan, I. & Planchard, D. ALK inhibitors in non-small cell lung cancer: the latest evidence and developments. *Ther. Adv. Med. Oncol.* 8, 32–47 (2016).
97. Kawakami, H. *et al.* Targeting MET Amplification as a New Oncogenic Driver. *Cancers* 6, 1540–1552 (2014).
98. Fujino, T., Suda, K. & Mitsudomi, T. Lung Cancer with MET exon 14 Skipping Mutation: Genetic Feature, Current Treatments, and Future Challenges. *Lung Cancer Auckl. NZ* 12, 35–50 (2021).
99. Redig, A. J. *et al.* Clinical and Molecular Characteristics of NF1-Mutant Lung Cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 22, 3148–3156 (2016).
100. Tlemsani, C. *et al.* NF1 mutations identify molecular and clinical subtypes of lung adenocarcinomas. *Cancer Med.* 8, 4330–4337 (2019).
101. Tao, J., Sun, D., Dong, L., Zhu, H. & Hou, H. Advancement in research and therapy of NF1 mutant malignant tumors. *Cancer Cell Int.* 20, 492 (2020).

102. Sequist, L. V. *et al.* Implementing multiplexed genotyping of non-small-cell lung cancers into routine clinical practice. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* 22, 2616–2624 (2011).
103. Jamal-Hanjani, M. *et al.* Tracking the Evolution of Non-Small-Cell Lung Cancer. *N. Engl. J. Med.* 376, 2109–2121 (2017).
104. Skoulidis, F. & Heymach, J. V. Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. *Nat. Rev. Cancer* 19, 495–509 (2019).
105. Reiter, J. L. *et al.* Comparative genomic sequence analysis and isolation of human and mouse alternative EGFR transcripts encoding truncated receptor isoforms. *Genomics* 71, 1–20 (2001).
106. Hynes, N. E. & MacDonald, G. ErbB receptors and signaling pathways in cancer. *Curr. Opin. Cell Biol.* 21, 177–184 (2009).
107. Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* 103, 211–225 (2000).
108. Zhang, X., Gureasko, J., Shen, K., Cole, P. A. & Kuriyan, J. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell* 125, 1137–1149 (2006).
109. Mitsudomi, T. & Yatabe, Y. Mutations of the epidermal growth factor receptor gene and related genes as determinants of epidermal growth factor receptor tyrosine kinase inhibitors sensitivity in lung cancer. *Cancer Sci.* 98, 1817–1824 (2007).
110. Weinstein, I. B. & Joe, A. Oncogene addiction. *Cancer Res.* 68, 3077–3080; discussion 3080 (2008).
111. Yasuda, H., Kobayashi, S. & Costa, D. B. EGFR exon 20 insertion mutations in non-small-cell lung cancer: preclinical data and clinical implications. *Lancet Oncol.* 13, e23–31 (2012).
112. Tracy, S. *et al.* Gefitinib induces apoptosis in the EGFR^{L858R} non-small-cell lung cancer cell line H3255. *Cancer Res.* 64, 7241–7244 (2004).
113. Weinstein, I. B. Cancer. Addiction to oncogenes--the Achilles heel of cancer. *Science* 297, 63–64 (2002).
114. Lynch, T. J. *et al.* Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* 350, 2129–2139 (2004).
115. Paez, J. G. *et al.* EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 304, 1497–1500 (2004).
116. Pao, W. *et al.* EGF receptor gene mutations are common in lung cancers from 'never smokers' and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc. Natl. Acad. Sci. U. S. A.* 101, 13306–13311 (2004).
117. Shepherd, F. A. *et al.* Erlotinib in previously treated non-small-cell lung cancer. *N. Engl. J. Med.* 353, 123–132 (2005).

118. Rosell, R. *et al.* Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. *Lancet Oncol.* 13, 239–246 (2012).
119. Mok, T. S. *et al.* Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N. Engl. J. Med.* 361, 947–957 (2009).
120. Mitsudomi, T. *et al.* Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): an open label, randomised phase 3 trial. *Lancet Oncol.* 11, 121–128 (2010).
121. Maemondo, M. *et al.* Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. *N. Engl. J. Med.* 362, 2380–2388 (2010).
122. Zhou, C. *et al.* Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomised, phase 3 study. *Lancet Oncol.* 12, 735–742 (2011).
123. Yang, J. C.-H. *et al.* Afatinib versus cisplatin-based chemotherapy for EGFR mutation-positive lung adenocarcinoma (LUX-Lung 3 and LUX-Lung 6): analysis of overall survival data from two randomised, phase 3 trials. *Lancet Oncol.* 16, 141–151 (2015).
124. Yang, J. C.-H. *et al.* Afatinib for patients with lung adenocarcinoma and epidermal growth factor receptor mutations (LUX-Lung 2): a phase 2 trial. *Lancet Oncol.* 13, 539–548 (2012).
125. Yang, J. C.-H. *et al.* Symptom control and quality of life in LUX-Lung 3: a phase III study of afatinib or cisplatin/pemetrexed in patients with advanced lung adenocarcinoma with EGFR mutations. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 31, 3342–3350 (2013).
126. Yang, J. C.-H. *et al.* Afatinib for the Treatment of NSCLC Harboring Uncommon EGFR Mutations: A Database of 693 Cases. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* 15, 803–815 (2020).
127. Yang, J. C.-H. *et al.* Clinical activity of afatinib in patients with advanced non-small-cell lung cancer harbouring uncommon EGFR mutations: a combined post-hoc analysis of LUX-Lung 2, LUX-Lung 3, and LUX-Lung 6. *Lancet Oncol.* 16, 830–838 (2015).
128. Brindel, A. *et al.* Uncommon EGFR mutations in lung adenocarcinoma: features and response to tyrosine kinase inhibitors. *J. Thorac. Dis.* 12, 4643–4650 (2020).
129. Goss, G. *et al.* Osimertinib for pretreated EGFR Thr790Met-positive advanced non-small-cell lung cancer (AURA2): a multicentre, open-label, single-arm, phase 2 study. *Lancet Oncol.* 17, 1643–1652 (2016).

130. Soria, J.-C. *et al.* Osimertinib in Untreated EGFR-Mutated Advanced Non-Small-Cell Lung Cancer. *N. Engl. J. Med.* 378, 113–125 (2018).
131. Ramalingam, S. S. *et al.* Overall Survival with Osimertinib in Untreated, EGFR-Mutated Advanced NSCLC. *N. Engl. J. Med.* 382, 41–50 (2020).
132. Reungwetwattana, T. *et al.* CNS Response to Osimertinib Versus Standard Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors in Patients With Untreated EGFR-Mutated Advanced Non-Small-Cell Lung Cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* JCO2018783118 (2018) doi:10.1200/JCO.2018.78.3118.
133. Seto, T. *et al.* Erlotinib alone or with bevacizumab as first-line therapy in patients with advanced non-squamous non-small-cell lung cancer harbouring EGFR mutations (JO25567): an open-label, randomised, multicentre, phase 2 study. *Lancet Oncol.* 15, 1236–1244 (2014).
134. Saito, H. *et al.* Erlotinib plus bevacizumab versus erlotinib alone in patients with EGFR-positive advanced non-squamous non-small-cell lung cancer (NEJ026): interim analysis of an open-label, randomised, multicentre, phase 3 trial. *Lancet Oncol.* 20, 625–635 (2019).
135. Chen, F., Chen, N., Yu, Y. & Cui, J. Efficacy and Safety of Epidermal Growth Factor Receptor (EGFR) Inhibitors Plus Antiangiogenic Agents as First-Line Treatments for Patients With Advanced EGFR-Mutated Non-small Cell Lung Cancer: A Meta-Analysis. *Front. Oncol.* 10, 904 (2020).
136. Hosomi, Y. *et al.* Gefitinib Alone Versus Gefitinib Plus Chemotherapy for Non-Small-Cell Lung Cancer With Mutated Epidermal Growth Factor Receptor: NEJ009 Study. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 38, 115–123 (2020).
137. Noronha, V. *et al.* Gefitinib Versus Gefitinib Plus Pemetrexed and Carboplatin Chemotherapy in EGFR-Mutated Lung Cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 38, 124–136 (2020).
138. Cheng, Y. *et al.* Randomized Phase II Trial of Gefitinib With and Without Pemetrexed as First-Line Therapy in Patients With Advanced Nonsquamous Non-Small-Cell Lung Cancer With Activating Epidermal Growth Factor Receptor Mutations. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 34, 3258–3266 (2016).
139. Vyse, S. & Huang, P. H. Targeting EGFR exon 20 insertion mutations in non-small cell lung cancer. *Signal Transduct. Target. Ther.* 4, 5 (2019).
140. Jackman, D. *et al.* Clinical definition of acquired resistance to epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 28, 357–360 (2010).
141. Cortot, A. B. & Jänne, P. A. Molecular mechanisms of resistance in epidermal growth factor receptor-mutant lung adenocarcinomas. *Eur. Respir. Rev. Off. J. Eur. Respir. Soc.* 23, 356–366 (2014).

142. Alfieri, R. R. *et al.* Metabolism of the EGFR tyrosin kinase inhibitor gefitinib by cytochrome P450 1A1 enzyme in EGFR-wild type non small cell lung cancer cell lines. *Mol. Cancer* 10, 143 (2011).
143. Ng, K. P. *et al.* A common BIM deletion polymorphism mediates intrinsic resistance and inferior responses to tyrosine kinase inhibitors in cancer. *Nat. Med.* 18, 521–528 (2012).
144. Bell, D. W. *et al.* Inherited susceptibility to lung cancer may be associated with the T790M drug resistance mutation in EGFR. *Nat. Genet.* 37, 1315–1316 (2005).
145. Matsushima, S. *et al.* V843I, a lung cancer predisposing EGFR mutation, is responsible for resistance to EGFR tyrosine kinase inhibitors. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* 9, 1377–1384 (2014).
146. Bivona, T. G. *et al.* FAS and NF- κ B signalling modulate dependence of lung cancers on mutant EGFR. *Nature* 471, 523–526 (2011).
147. Fujita, Y. *et al.* Highly sensitive detection of EGFR T790M mutation using colony hybridization predicts favorable prognosis of patients with lung cancer harboring activating EGFR mutation. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* 7, 1640–1644 (2012).
148. Toyooka, S., Date, H., Uchida, A., Kiura, K. & Takata, M. The epidermal growth factor receptor D761Y mutation and effect of tyrosine kinase inhibitor. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 13, 3431; author reply 3431–3432 (2007).
149. Asahina, H. *et al.* Non-responsiveness to gefitinib in a patient with lung adenocarcinoma having rare EGFR mutations S768I and V769L. *Lung Cancer Amst. Neth.* 54, 419–422 (2006).
150. Niogret, J., Coudert, B. & Boidot, R. Primary Resistance to Afatinib in a Patient with Lung Adenocarcinoma Harboring Uncommon EGFR Mutations: S768I and V769L. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* 13, e113 (2018).
151. Zhang, L. *et al.* Molecular Characteristics of the Uncommon EGFR Exon 21 T854A Mutation and Response to Osimertinib in Patients With Non-Small Cell Lung Cancer. *Clin. Lung Cancer* S1525-7304(21)00304-1 (2021) doi:10.1016/j.clcc.2021.12.008.
152. Erjala, K. *et al.* Signaling via ErbB2 and ErbB3 associates with resistance and epidermal growth factor receptor (EGFR) amplification with sensitivity to EGFR inhibitor gefitinib in head and neck squamous cell carcinoma cells. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 12, 4103–4111 (2006).
153. Yamamoto, C. *et al.* Loss of PTEN expression by blocking nuclear translocation of EGR1 in gefitinib-resistant lung cancer cells harboring epidermal growth factor receptor-activating mutations. *Cancer Res.* 70, 8715–8725 (2010).
154. Rotow, J. & Bivona, T. G. Understanding and targeting resistance mechanisms in NSCLC. *Nat. Rev. Cancer* 17, 637–658 (2017).

155. Nguyen, K.-S. H., Kobayashi, S. & Costa, D. B. Acquired resistance to epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancers dependent on the epidermal growth factor receptor pathway. *Clin. Lung Cancer* 10, 281–289 (2009).
156. Jotte, R. M. & Spigel, D. R. Advances in molecular-based personalized non-small-cell lung cancer therapy: targeting epidermal growth factor receptor and mechanisms of resistance. *Cancer Med.* 4, 1621–1632 (2015).
157. Engelman, J. A. *et al.* Allelic dilution obscures detection of a biologically significant resistance mutation in EGFR-amplified lung cancer. *J. Clin. Invest.* 116, 2695–2706 (2006).
158. Yun, C.-H. *et al.* The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc. Natl. Acad. Sci. U. S. A.* 105, 2070–2075 (2008).
159. Mok, T. S. *et al.* Osimertinib or Platinum-Pemetrexed in EGFR T790M-Positive Lung Cancer. *N. Engl. J. Med.* 376, 629–640 (2017).
160. Reita, D. *et al.* Molecular Mechanism of EGFR-TKI Resistance in EGFR-Mutated Non-Small Cell Lung Cancer: Application to Biological Diagnostic and Monitoring. *Cancers* 13, 4926 (2021).
161. Sequist, L. V. *et al.* Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. *Sci. Transl. Med.* 3, 75ra26 (2011).
162. Sharma, S. V., Bell, D. W., Settleman, J. & Haber, D. A. Epidermal growth factor receptor mutations in lung cancer. *Nat. Rev. Cancer* 7, 169–181 (2007).
163. Engelman, J. A. *et al.* MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* 316, 1039–1043 (2007).
164. Takezawa, K. *et al.* HER2 amplification: a potential mechanism of acquired resistance to EGFR inhibition in EGFR-mutant lung cancers that lack the second-site EGFR T790M mutation. *Cancer Discov.* 2, 922–933 (2012).
165. Terai, H. *et al.* Activation of the FGF2-FGFR1 autocrine pathway: a novel mechanism of acquired resistance to gefitinib in NSCLC. *Mol. Cancer Res. MCR* 11, 759–767 (2013).
166. Zhang, Z. *et al.* Activation of the AXL kinase causes resistance to EGFR-targeted therapy in lung cancer. *Nat. Genet.* 44, 852–860 (2012).
167. Yasuda, H. *et al.* Structural, biochemical, and clinical characterization of epidermal growth factor receptor (EGFR) exon 20 insertion mutations in lung cancer. *Sci. Transl. Med.* 5, 216ra177 (2013).
168. Arcila, M. E. *et al.* EGFR exon 20 insertion mutations in lung adenocarcinomas: prevalence, molecular heterogeneity, and clinicopathologic characteristics. *Mol. Cancer Ther.* 12, 220–229 (2013).

169. Yang, M. *et al.* NSCLC harboring EGFR exon-20 insertions after the regulatory C-helix of kinase domain responds poorly to known EGFR inhibitors. *Int. J. Cancer* 139, 171–176 (2016).
170. Floc'h, N. *et al.* Antitumor Activity of Osimertinib, an Irreversible Mutant-Selective EGFR Tyrosine Kinase Inhibitor, in NSCLC Harboring EGFR Exon 20 Insertions. *Mol. Cancer Ther.* 17, 885–896 (2018).
171. Naidoo, J. *et al.* Epidermal growth factor receptor exon 20 insertions in advanced lung adenocarcinomas: Clinical outcomes and response to erlotinib. *Cancer* 121, 3212–3220 (2015).
172. Beau-Faller, M. *et al.* Rare EGFR exon 18 and exon 20 mutations in non-small-cell lung cancer on 10 117 patients: a multicentre observational study by the French HERMETIC-IFCT network. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* 25, 126–131 (2014).
173. Robichaux, J. P. *et al.* Mechanisms and clinical activity of an EGFR and HER2 exon 20-selective kinase inhibitor in non-small cell lung cancer. *Nat. Med.* 24, 638–646 (2018).
174. Prelaj, A. *et al.* Poziotinib for EGFR and HER2 exon 20 insertion mutation in advanced NSCLC: Results from the expanded access program. *Eur. J. Cancer Oxf. Engl.* 1990 149, 235–248 (2021).
175. Riely, G. J. *et al.* Activity and Safety of Mobocertinib (TAK-788) in Previously Treated Non-Small Cell Lung Cancer with EGFR Exon 20 Insertion Mutations from a Phase I/II Trial. *Cancer Discov.* 11, 1688–1699 (2021).
176. Gonzalez, F. *et al.* Mobocertinib (TAK-788): A Targeted Inhibitor of EGFR Exon 20 Insertion Mutants in Non-Small Cell Lung Cancer. *Cancer Discov.* 11, 1672–1687 (2021).
177. Zhou, C. *et al.* Treatment Outcomes and Safety of Mobocertinib in Platinum-Pretreated Patients With EGFR Exon 20 Insertion-Positive Metastatic Non-Small Cell Lung Cancer: A Phase 1/2 Open-label Nonrandomized Clinical Trial. *JAMA Oncol.* 7, e214761 (2021).
178. Köhler, J. & Jänne, P. A. Amivantamab: Treating EGFR Exon 20-Mutant Cancers With Bispecific Antibody-Mediated Receptor Degradation. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 39, 3403–3406 (2021).
179. Park, K. *et al.* Amivantamab in EGFR Exon 20 Insertion-Mutated Non-Small-Cell Lung Cancer Progressing on Platinum Chemotherapy: Initial Results From the CHRYSALIS Phase I Study. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 39, 3391–3402 (2021).
180. Yun, J. *et al.* Antitumor Activity of Amivantamab (JNJ-61186372), an EGFR-MET Bispecific Antibody, in Diverse Models of EGFR Exon 20 Insertion-Driven NSCLC. *Cancer Discov.* 10, 1194–1209 (2020).
181. Dietel, M. *et al.* Diagnostic procedures for non-small-cell lung cancer (NSCLC): recommendations of the European Expert Group. *Thorax* 71, 177–184 (2016).

182. HAS. Définition des conditions de réalisation des tests de détection des mutations activatrices de l'EGFR et des mutations BRAF, NRAS, et KRAS. 48 (2021).
183. Hedegaard, J. *et al.* Next-generation sequencing of RNA and DNA isolated from paired fresh-frozen and formalin-fixed paraffin-embedded samples of human cancer and normal tissue. *PLoS One* 9, e98187 (2014).
184. McDonough, S. J. *et al.* Use of FFPE-derived DNA in next generation sequencing: DNA extraction methods. *PLoS One* 14, e0211400 (2019).
185. Bass, B. P., Engel, K. B., Greytak, S. R. & Moore, H. M. A review of preanalytical factors affecting molecular, protein, and morphological analysis of formalin-fixed, paraffin-embedded (FFPE) tissue: how well do you know your FFPE specimen? *Arch. Pathol. Lab. Med.* 138, 1520–1530 (2014).
186. Breton, Q., Valo, I., Benallègue, F., Morel, A. & Verrière, V. Influence de la cellularité tumorale dans la recherche des mutations des cancers du poumon : expérience du centre Paul Papin (Angers). </data/revues/02426498/v35i5/S0242649815001510/> (2015).
187. S Couraud, Westeel, V. & Ranchon. Référentiel sur le cancer bronchique non à petites-cellules: actualisation 2021. ARISTOT 2021. <http://referentiels-aristot.com/>.
188. Tissot, C., Villar, S., Olivier, M. & Couraud, S. [Free circulating DNA as a tool for lung cancer patients management]. *Rev. Pneumol. Clin.* 72, 61–71 (2016).
189. Jiang, T., Ren, S. & Zhou, C. Role of circulating-tumor DNA analysis in non-small cell lung cancer. *Lung Cancer Amst. Neth.* 90, 128–134 (2015).
190. Diehl, F. *et al.* Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proc. Natl. Acad. Sci. U. S. A.* 102, 16368–16373 (2005).
191. Qiu, M. *et al.* Circulating tumor DNA is effective for the detection of EGFR mutation in non-small cell lung cancer: a meta-analysis. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 24, 206–212 (2015).
192. Sacher, A. G. *et al.* Prospective Validation of Rapid Plasma Genotyping for the Detection of EGFR and KRAS Mutations in Advanced Lung Cancer. *JAMA Oncol.* 2, 1014–1022 (2016).
193. Zheng, D. *et al.* Plasma EGFR T790M ctDNA status is associated with clinical outcome in advanced NSCLC patients with acquired EGFR-TKI resistance. *Sci. Rep.* 6, 20913 (2016).
194. Marcq, M., Vallée, A., Bizieux, A. & Denis, M. G. Detection of EGFR mutations in the plasma of patients with lung adenocarcinoma for real-time monitoring of therapeutic response to tyrosine kinase inhibitors? *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* 9, e49–e50 (2014).

195. Oxnard, G. R. *et al.* Noninvasive detection of response and resistance in EGFR-mutant lung cancer using quantitative next-generation genotyping of cell-free plasma DNA. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 20, 1698–1705 (2014).
196. Ni, J. *et al.* Dynamic monitoring of EGFR mutations in circulating cell-free DNA for EGFR-mutant metastatic patients with lung cancer: Early detection of drug resistance and prognostic significance. *Oncol. Lett.* 13, 4549–4557 (2017).
197. Helman, E. *et al.* Cell-Free DNA Next-Generation Sequencing Prediction of Response and Resistance to Third-Generation EGFR Inhibitor. *Clin. Lung Cancer* 19, 518-530.e7 (2018).
198. Marchetti, A. *et al.* Early Prediction of Response to Tyrosine Kinase Inhibitors by Quantification of EGFR Mutations in Plasma of NSCLC Patients. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* 10, 1437–1443 (2015).
199. Zeng, D. *et al.* Cell-free DNA from bronchoalveolar lavage fluid (BALF): a new liquid biopsy medium for identifying lung cancer. *Ann. Transl. Med.* 9, 1080 (2021).
200. Liu, Y. *et al.* Cell-free DNA from cerebrospinal fluid can be used to detect the EGFR mutation status of lung adenocarcinoma patients with central nervous system metastasis. *Transl. Lung Cancer Res.* 10, 914–925 (2021).
201. Mokánszki, A. *et al.* Cell-free DNA From Pleural Effusion Samples: Is It Right for Molecular Testing in Lung Adenocarcinoma? *Pathol. Oncol. Res. POR* 27, 613071 (2021).
202. Kim, Y. J., Ji, W. J., Lee, J.-C., Chun, S.-M. & Choi, C.-M. Assessment of Anti-tumor Efficacy of Osimertinib in Non-Small Cell Lung Cancer Patients by Liquid Biopsy Using Bronchoalveolar Lavage Fluid, Plasma, or Pleural Effusion. *Cancer Res. Treat.* (2022) doi:10.4143/crt.2021.857.
203. Underhill, H. R. *et al.* Fragment Length of Circulating Tumor DNA. *PLoS Genet.* 12, e1006162 (2016).
204. Ronaghi, M., Uhlén, M. & Nyrén, P. A sequencing method based on real-time pyrophosphate. *Science* 281, 363, 365 (1998).
205. Heideman, D. a. M. *et al.* A panel of high resolution melting (HRM) technology-based assays with direct sequencing possibility for effective mutation screening of EGFR and K-ras genes. *Cell. Oncol. Off. J. Int. Soc. Cell. Oncol.* 31, 329–333 (2009).
206. Dressman, D., Yan, H., Traverso, G., Kinzler, K. W. & Vogelstein, B. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proc. Natl. Acad. Sci. U. S. A.* 100, 8817–8822 (2003).
207. Malapelle, U. *et al.* Profile of the Roche cobas® EGFR mutation test v2 for non-small cell lung cancer. *Expert Rev. Mol. Diagn.* 17, 209–215 (2017).

208. Vallée, A., Le Loupp, A.-G. & Denis, M. G. Efficiency of the Therascreen® RGQ PCR kit for the detection of EGFR mutations in non-small cell lung carcinomas. *Clin. Chim. Acta Int. J. Clin. Chem.* 429, 8–11 (2014).
209. Tan, L. Y. *et al.* Superior Multiplexing Capacity of PlexPrimers Enables Sensitive and Specific Detection of SNPs and Clustered Mutations in qPCR. *PLOS ONE* 12, e0170087 (2017).
210. Ilie, M. *et al.* Optimization of EGFR mutation detection by the fully-automated qPCR-based Idylla system on tumor tissue from patients with non-small cell lung cancer. *Oncotarget* 8, 103055–103062 (2017).
211. Do, K. & Chen, A. P. Molecular pathways: targeting PARP in cancer treatment. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 19, 977–984 (2013).
212. Farmer, H. *et al.* Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* 434, 917–921 (2005).
213. Sonnenblick, A., de Azambuja, E., Azim, H. A. & Piccart, M. An update on PARP inhibitors--moving to the adjuvant setting. *Nat. Rev. Clin. Oncol.* 12, 27–41 (2015).
214. Konstantinopoulos, P. A., Ceccaldi, R., Shapiro, G. I. & D'Andrea, A. D. Homologous Recombination Deficiency: Exploiting the Fundamental Vulnerability of Ovarian Cancer. *Cancer Discov.* 5, 1137–1154 (2015).
215. Ceccaldi, R., Rondinelli, B. & D'Andrea, A. D. Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends Cell Biol.* 26, 52–64 (2016).
216. Hustedt, N. & Durocher, D. The control of DNA repair by the cell cycle. *Nat. Cell Biol.* 19, 1–9 (2016).
217. Weinstock, D. M. & Jasin, M. Alternative pathways for the repair of RAG-induced DNA breaks. *Mol. Cell. Biol.* 26, 131–139 (2006).
218. Roy, R., Chun, J. & Powell, S. N. BRCA1 and BRCA2: different roles in a common pathway of genome protection. *Nat. Rev. Cancer* 12, 68–78 (2012).
219. Cook, P. J. *et al.* Tyrosine dephosphorylation of H2AX modulates apoptosis and survival decisions. *Nature* 458, 591–596 (2009).
220. Yan, J. & Jetten, A. M. RAP80 and RNF8, key players in the recruitment of repair proteins to DNA damage sites. *Cancer Lett.* 271, 179–190 (2008).
221. Reilly, N. M., Yard, B. D. & Pittman, D. L. Homologous Recombination-Mediated DNA Repair and Implications for Clinical Treatment of Repair Defective Cancers. *Methods Mol. Biol. Clifton NJ* 1999, 3–29 (2019).
222. van Gent, D. C., Hoeijmakers, J. H. & Kanaar, R. Chromosomal stability and the DNA double-stranded break connection. *Nat. Rev. Genet.* 2, 196–206 (2001).
223. Prakash, R., Zhang, Y., Feng, W. & Jasin, M. Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins. *Cold Spring Harb. Perspect. Biol.* 7, a016600 (2015).

224. Chen, C.-C., Feng, W., Lim, P. X., Kass, E. M. & Jasin, M. Homology-Directed Repair and the Role of BRCA1, BRCA2, and Related Proteins in Genome Integrity and Cancer. *Annu. Rev. Cancer Biol.* 2, 313–336 (2018).
225. Turner, N., Tutt, A. & Ashworth, A. Hallmarks of 'BRCAness' in sporadic cancers. *Nat. Rev. Cancer* 4, 814–819 (2004).
226. Hall, J. M. *et al.* Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 250, 1684–1689 (1990).
227. Szabo, C. I. *et al.* Human, canine and murine BRCA1 genes: sequence comparison among species. *Hum. Mol. Genet.* 5, 1289–1298 (1996).
228. Ruffner, H. & Verma, I. M. BRCA1 is a cell cycle-regulated nuclear phosphoprotein. *Proc. Natl. Acad. Sci. U. S. A.* 94, 7138–7143 (1997).
229. Wang, Q., Zhang, H., Fishel, R. & Greene, M. I. BRCA1 and cell signaling. *Oncogene* 19, 6152–6158 (2000).
230. Monteiro, A. N., August, A. & Hanafusa, H. Evidence for a transcriptional activation function of BRCA1 C-terminal region. *Proc. Natl. Acad. Sci. U. S. A.* 93, 13595–13599 (1996).
231. Nadeau, G. *et al.* BRCA1 can stimulate gene transcription by a unique mechanism. *EMBO Rep.* 1, 260–265 (2000).
232. Zhu, Q. *et al.* BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* 477, 179–184 (2011).
233. Mallery, D. L., Vandenberg, C. J. & Hiom, K. Activation of the E3 ligase function of the BRCA1/BARD1 complex by polyubiquitin chains. *EMBO J.* 21, 6755–6762 (2002).
234. Xia, Y., Pao, G. M., Chen, H.-W., Verma, I. M. & Hunter, T. Enhancement of BRCA1 E3 ubiquitin ligase activity through direct interaction with the BARD1 protein. *J. Biol. Chem.* 278, 5255–5263 (2003).
235. Martinez-Outschoorn, U. E. *et al.* BRCA1 mutations drive oxidative stress and glycolysis in the tumor microenvironment: implications for breast cancer prevention with antioxidant therapies. *Cell Cycle Georget. Tex* 11, 4402–4413 (2012).
236. Saha, T., Rih, J. K. & Rosen, E. M. BRCA1 down-regulates cellular levels of reactive oxygen species. *FEBS Lett.* 583, 1535–1543 (2009).
237. Arun, B., Akar, U., Gutierrez-Barrera, A. M., Hortobagyi, G. N. & Ozpolat, B. The PARP inhibitor AZD2281 (Olaparib) induces autophagy/mitophagy in BRCA1 and BRCA2 mutant breast cancer cells. *Int. J. Oncol.* 47, 262–268 (2015).
238. Wooster, R. *et al.* Identification of the breast cancer susceptibility gene BRCA2. *Nature* 378, 789–792 (1995).
239. Tavtigian, S. V. *et al.* The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. *Nat. Genet.* 12, 333–337 (1996).

240. Wong, A. K., Pero, R., Ormonde, P. A., Tavtigian, S. V. & Bartel, P. L. RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene *brca2*. *J. Biol. Chem.* 272, 31941–31944 (1997).
241. Esashi, F. *et al.* CDK-dependent phosphorylation of BRCA2 as a regulatory mechanism for recombinational repair. *Nature* 434, 598–604 (2005).
242. Menzel, T. *et al.* A genetic screen identifies BRCA2 and PALB2 as key regulators of G2 checkpoint maintenance. *EMBO Rep.* 12, 705–712 (2011).
243. Milner, J., Ponder, B., Hughes-Davies, L., Seltmann, M. & Kouzarides, T. Transcriptional activation functions in BRCA2. *Nature* 386, 772–773 (1997).
244. Min, J. *et al.* The breast cancer susceptibility gene BRCA2 is required for the maintenance of telomere homeostasis. *J. Biol. Chem.* 287, 5091–5101 (2012).
245. Badie, S. *et al.* BRCA2 acts as a RAD51 loader to facilitate telomere replication and capping. *Nat. Struct. Mol. Biol.* 17, 1461–1469 (2010).
246. Lomonosov, M., Anand, S., Sangrithi, M., Davies, R. & Venkitaraman, A. R. Stabilization of stalled DNA replication forks by the BRCA2 breast cancer susceptibility protein. *Genes Dev.* 17, 3017–3022 (2003).
247. Sedic, M. *et al.* Haploinsufficiency for BRCA1 leads to cell-type-specific genomic instability and premature senescence. *Nat. Commun.* 6, 7505 (2015).
248. Clarke, C. L. *et al.* Mapping loss of heterozygosity in normal human breast cells from BRCA1/2 carriers. *Br. J. Cancer* 95, 515–519 (2006).
249. Chen, J.-Q., Brown, T. R. & Yager, J. D. Mechanisms of hormone carcinogenesis: evolution of views, role of mitochondria. *Adv. Exp. Med. Biol.* 630, 1–18 (2008).
250. Néné, N. R. *et al.* Association between the cervicovaginal microbiome, BRCA1 mutation status, and risk of ovarian cancer: a case-control study. *Lancet Oncol.* 20, 1171–1182 (2019).
251. Singh, A. K. & Yu, X. Tissue-Specific Carcinogens as Soil to Seed BRCA1/2-Mutant Hereditary Cancers. *Trends Cancer* 6, 559–568 (2020).
252. Kuchenbaecker, K. B. *et al.* Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. *JAMA* 317, 2402–2416 (2017).
253. Thorlacius, S. *et al.* Linkage to BRCA2 region in hereditary male breast cancer. *Lancet Lond. Engl.* 346, 544–545 (1995).
254. Ford, D. *et al.* Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am. J. Hum. Genet.* 62, 676–689 (1998).
255. Pilarski, R. The Role of BRCA Testing in Hereditary Pancreatic and Prostate Cancer Families. *Am. Soc. Clin. Oncol. Educ. Book Am. Soc. Clin. Oncol. Annu. Meet.* 39, 79–86 (2019).

256. Breast Cancer Linkage Consortium. Cancer risks in BRCA2 mutation carriers. *J. Natl. Cancer Inst.* 91, 1310–1316 (1999).
257. Gumaste, P. V. *et al.* Skin cancer risk in BRCA1/2 mutation carriers. *Br. J. Dermatol.* 172, 1498–1506 (2015).
258. Li, S. *et al.* Cancer Risks Associated With BRCA1 and BRCA2 Pathogenic Variants. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* JCO2102112 (2022) doi:10.1200/JCO.21.02112.
259. Knudson, A. G. Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl. Acad. Sci. U. S. A.* 68, 820–823 (1971).
260. Van Heetvelde, M. *et al.* Accurate detection and quantification of epigenetic and genetic second hits in BRCA1 and BRCA2-associated hereditary breast and ovarian cancer reveals multiple co-acting second hits. *Cancer Lett.* 425, 125–133 (2018).
261. Mount, S. M. A catalogue of splice junction sequences. *Nucleic Acids Res.* 10, 459–472 (1982).
262. Senapathy, P., Shapiro, M. B. & Harris, N. L. Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome project. *Methods Enzymol.* 183, 252–278 (1990).
263. Fredericks, A. M., Cygan, K. J., Brown, B. A. & Fairbrother, W. G. RNA-Binding Proteins: Splicing Factors and Disease. *Biomolecules* 5, 893–909 (2015).
264. Mercer, T. R. *et al.* Genome-wide discovery of human splicing branchpoints. *Genome Res.* 25, 290–303 (2015).
265. Liu, H. X., Cartegni, L., Zhang, M. Q. & Krainer, A. R. A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. *Nat. Genet.* 27, 55–58 (2001).
266. Sterne-Weiler, T., Howard, J., Mort, M., Cooper, D. N. & Sanford, J. R. Loss of exon identity is a common mechanism of human inherited disease. *Genome Res.* 21, 1563–1571 (2011).
267. Anczuków, O. *et al.* BRCA2 deep intronic mutation causing activation of a cryptic exon: opening toward a new preventive therapeutic strategy. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 18, 4903–4909 (2012).
268. Montalban, G. *et al.* Screening of BRCA1/2 deep intronic regions by targeted gene sequencing identifies the first germline BRCA1 variant causing pseudoexon activation in a patient with breast/ovarian cancer. *J. Med. Genet.* 56, 63–74 (2019).
269. Samorodnitsky, E. *et al.* Evaluation of Hybridization Capture Versus Amplicon-Based Methods for Whole-Exome Sequencing. *Hum. Mutat.* 36, 903–914 (2015).
270. Judkins, T. *et al.* Clinical significance of large rearrangements in BRCA1 and BRCA2. *Cancer* 118, 5210–5216 (2012).

271. Wallace, A. J. New challenges for BRCA testing: a view from the diagnostic laboratory. *Eur. J. Hum. Genet. EJHG* 24 Suppl 1, S10-18 (2016).
272. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* 474, 609–615 (2011).
273. Pennington, K. P. *et al.* Germline and somatic mutations in homologous recombination genes predict platinum response and survival in ovarian, fallopian tube, and peritoneal carcinomas. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 20, 764–775 (2014).
274. Winter, C. *et al.* Targeted sequencing of BRCA1 and BRCA2 across a large unselected breast cancer cohort suggests that one-third of mutations are somatic. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* 27, 1532–1538 (2016).
275. Decker, B. *et al.* Biallelic BRCA2 Mutations Shape the Somatic Mutational Landscape of Aggressive Prostate Tumors. *Am. J. Hum. Genet.* 98, 818–829 (2016).
276. Esteller, M. *et al.* Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. *J. Natl. Cancer Inst.* 92, 564–569 (2000).
277. Mancini, D. N. *et al.* CpG methylation within the 5' regulatory region of the BRCA1 gene is tumor specific and includes a putative CREB binding site. *Oncogene* 16, 1161–1169 (1998).
278. Baldwin, R. L. *et al.* BRCA1 promoter region hypermethylation in ovarian carcinoma: a population-based study. *Cancer Res.* 60, 5329–5333 (2000).
279. Hughes-Davies, L. *et al.* EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer. *Cell* 115, 523–535 (2003).
280. Wilkerson, P. M. *et al.* Functional characterization of EMSY gene amplification in human cancers. *J. Pathol.* 225, 29–42 (2011).
281. Lord, C. J. & Ashworth, A. BRCAness revisited. *Nat. Rev. Cancer* 16, 110–120 (2016).
282. Waddell, N. *et al.* Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature* 518, 495–501 (2015).
283. Park, W. *et al.* Genomic Methods Identify Homologous Recombination Deficiency in Pancreas Adenocarcinoma and Optimize Treatment Selection. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 26, 3239–3247 (2020).
284. Robinson, D. *et al.* Integrative clinical genomics of advanced prostate cancer. *Cell* 161, 1215–1228 (2015).
285. Cancer Genome Atlas Research Network. The Molecular Taxonomy of Primary Prostate Cancer. *Cell* 163, 1011–1025 (2015).
286. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* 490, 61–70 (2012).
287. Lord, C. J., Tutt, A. N. J. & Ashworth, A. Synthetic lethality and cancer therapy: lessons learned from the development of PARP inhibitors. *Annu. Rev. Med.* 66, 455–470 (2015).

288. Wagener-Ryczek, S., Merkelbach-Bruse, S. & Siemanowski, J. Biomarkers for Homologous Recombination Deficiency in Cancer. *J. Pers. Med.* 11, 612 (2021).
289. Takamatsu, S. *et al.* Utility of Homologous Recombination Deficiency Biomarkers Across Cancer Types. *JCO Precis. Oncol.* 5, PO.21.00141 (2021).
290. Jazaeri, A. A. *et al.* Gene expression profiles of BRCA1-linked, BRCA2-linked, and sporadic ovarian cancers. *J. Natl. Cancer Inst.* 94, 990–1000 (2002).
291. Konstantinopoulos, P. A. *et al.* Gene expression profile of BRCAness that correlates with responsiveness to chemotherapy and with outcome in patients with epithelial ovarian cancer. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 28, 3555–3561 (2010).
292. Popova, T. *et al.* Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. *Cancer Res.* 72, 5454–5462 (2012).
293. Manié, E. *et al.* Genomic hallmarks of homologous recombination deficiency in invasive breast carcinomas. *Int. J. Cancer* 138, 891–900 (2016).
294. Goundiam, O. *et al.* Histo-genomic stratification reveals the frequent amplification/overexpression of CCNE1 and BRD4 genes in non-BRCAness high grade ovarian carcinoma. *Int. J. Cancer* 137, 1890–1900 (2015).
295. Birkbak, N. J. *et al.* Telomeric allelic imbalance indicates defective DNA repair and sensitivity to DNA-damaging agents. *Cancer Discov.* 2, 366–375 (2012).
296. Abkevich, V. *et al.* Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. *Br. J. Cancer* 107, 1776–1782 (2012).
297. Mills, G. B. *et al.* Homologous recombination deficiency score shows superior association with outcome compared with its individual score components in platinum-treated serous ovarian cancer. *Gynecol. Oncol.* 141, 2–3 (2016).
298. Ray-Coquard, I. *et al.* Olaparib plus Bevacizumab as First-Line Maintenance in Ovarian Cancer. *N. Engl. J. Med.* 381, 2416–2428 (2019).
299. Tew, W. P. *et al.* PARP Inhibitors in the Management of Ovarian Cancer: ASCO Guideline. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 38, 3468–3493 (2020).
300. Miller, R. E. *et al.* ESMO recommendations on predictive biomarker testing for homologous recombination deficiency and PARP inhibitor benefit in ovarian cancer. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* 31, 1606–1622 (2020).
301. Armstrong, D. K. *et al.* Ovarian Cancer, Version 2.2020, NCCN Clinical Practice Guidelines in Oncology. *J. Natl. Compr. Cancer Netw. JNCCN* 19, 191–226 (2021).
302. Rodrigues, M., Manié, É., Popova, T. & Stern, M. H. BRCAness/défauts de la recombinaison homologue dans les cancers : mécanismes, diagnostic et conséquences thérapeutiques. *Corresp. En Onco-Théranostic* 6 (2016).

303. Coleman, R. L. *et al.* Rucaparib maintenance treatment for recurrent ovarian carcinoma after response to platinum therapy (ARIEL3): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Lond. Engl.* 390, 1949–1961 (2017).
304. Oza, A. M. *et al.* Patient-Centered Outcomes in ARIEL3, a Phase III, Randomized, Placebo-Controlled Trial of Rucaparib Maintenance Treatment in Patients With Recurrent Ovarian Carcinoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 38, 3494–3505 (2020).
305. Ni, J. *et al.* The efficacy and safety of niraparib for ovarian cancer: a single-center observational study from China. *J. Ovarian Res.* 14, 68 (2021).
306. Weichert, W. *et al.* 216 An evaluation of the performance of molecular assays to identify homologous recombination deficiency-positive tumours in ovarian cancer. *Int. J. Gynecol. Cancer* 31, (2021).
307. Weichert, W. *et al.* 978 Concordance of the FDA-approved companion diagnostic and a next-generation sequencing assay kit for assessing homologous recombination deficiency in ovarian cancer. *Int. J. Gynecol. Cancer* 31, (2021).
308. Wang, Y. K. *et al.* Genomic consequences of aberrant DNA repair mechanisms stratify ovarian cancer histotypes. *Nat. Genet.* 49, 856–865 (2017).
309. Rojas, V., Hirshfield, K. M., Ganesan, S. & Rodriguez-Rodriguez, L. Molecular Characterization of Epithelial Ovarian Cancer: Implications for Diagnosis and Treatment. *Int. J. Mol. Sci.* 17, 2113 (2016).
310. Norquist, B. M. *et al.* Inherited Mutations in Women With Ovarian Carcinoma. *JAMA Oncol.* 2, 482–490 (2016).
311. Pietragalla, A., Arcieri, M., Marchetti, C., Scambia, G. & Fagotti, A. Ovarian cancer predisposition beyond BRCA1 and BRCA2 genes. *Int. J. Gynecol. Cancer Off. J. Int. Gynecol. Cancer Soc.* 30, 1803–1810 (2020).
312. Hodan, R. *et al.* Prevalence of Lynch syndrome in women with mismatch repair-deficient ovarian cancer. *Cancer Med.* 10, 1012–1017 (2021).
313. Hollis, R. L. & Gourley, C. Genetic and molecular changes in ovarian cancer. *Cancer Biol. Med.* 13, 236–247 (2016).
314. Moore, K. *et al.* Maintenance Olaparib in Patients with Newly Diagnosed Advanced Ovarian Cancer. *N. Engl. J. Med.* 379, 2495–2505 (2018).
315. Pujade-Lauraine, E. *et al.* Olaparib tablets as maintenance therapy in patients with platinum-sensitive, relapsed ovarian cancer and a BRCA1/2 mutation (SOLO2/ENGOT-Ov21): a double-blind, randomised, placebo-controlled, phase 3 trial. *Lancet Oncol.* 18, 1274–1284 (2017).
316. González-Martín, A. *et al.* Niraparib in Patients with Newly Diagnosed Advanced Ovarian Cancer. *N. Engl. J. Med.* 381, 2391–2402 (2019).

317. Coleman, R. L. *et al.* Veliparib with First-Line Chemotherapy and as Maintenance Therapy in Ovarian Cancer. *N. Engl. J. Med.* 381, 2403–2415 (2019).
318. van 't Veer, L. J. *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415, 530–536 (2002).
319. Gonzalez-Angulo, A. M. *et al.* Incidence and outcome of BRCA mutations in unselected patients with triple receptor-negative breast cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 17, 1082–1089 (2011).
320. Liang, Y. *et al.* Prevalence and Spectrum of BRCA1/2 Germline Mutations in Women with Breast Cancer in China Based on Next-Generation Sequencing. *Med. Sci. Monit. Int. Med. J. Exp. Clin. Res.* 24, 2465–2475 (2018).
321. Sestak, I. *et al.* Comparison of the Performance of 6 Prognostic Signatures for Estrogen Receptor-Positive Breast Cancer: A Secondary Analysis of a Randomized Clinical Trial. *JAMA Oncol.* 4, 545–553 (2018).
322. Robson, M. *et al.* Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. *N. Engl. J. Med.* 377, 523–533 (2017).
323. Litton, J. K. *et al.* Talazoparib in Patients with Advanced Breast Cancer and a Germline BRCA Mutation. *N. Engl. J. Med.* 379, 753–763 (2018).
324. Gennari, A. *et al.* ESMO Clinical Practice Guideline for the diagnosis, staging and treatment of patients with metastatic breast cancer. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* 32, 1475–1495 (2021).
325. Condorelli, R. *et al.* Genomic alterations in breast cancer: level of evidence for actionability according to ESMO Scale for Clinical Actionability of molecular Targets (ESCAT). *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* 30, 365–373 (2019).
326. Cancer Genome Atlas Research Network. Electronic address: andrew_aguirre@dfci.harvard.edu & Cancer Genome Atlas Research Network. Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell* 32, 185–203.e13 (2017).
327. Lucas, A. L. *et al.* BRCA1 and BRCA2 germline mutations are frequently demonstrated in both high-risk pancreatic cancer screening and pancreatic cancer cohorts. *Cancer* 120, 1960–1967 (2014).
328. Holter, S. *et al.* Germline BRCA Mutations in a Large Clinic-Based Cohort of Patients With Pancreatic Adenocarcinoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 33, 3124–3129 (2015).
329. Golan, T. *et al.* Maintenance Olaparib for Germline BRCA-Mutated Metastatic Pancreatic Cancer. *N. Engl. J. Med.* 381, 317–327 (2019).
330. Angeles, A. K., Bauer, S., Ratz, L., Klauk, S. M. & Sültmann, H. Genome-Based Classification and Therapy of Prostate Cancer. *Diagn. Basel Switz.* 8, E62 (2018).

331. Ewing, C. M. *et al.* Germline mutations in HOXB13 and prostate-cancer risk. *N. Engl. J. Med.* 366, 141–149 (2012).
332. Pritchard, C. C. *et al.* Inherited DNA-Repair Gene Mutations in Men with Metastatic Prostate Cancer. *N. Engl. J. Med.* 375, 443–453 (2016).
333. de Bono, J. *et al.* Olaparib for Metastatic Castration-Resistant Prostate Cancer. *N. Engl. J. Med.* 382, 2091–2102 (2020).
334. Hussain, M. *et al.* Survival with Olaparib in Metastatic Castration-Resistant Prostate Cancer. *N. Engl. J. Med.* 383, 2345–2357 (2020).
335. Cline, M. S. *et al.* BRCA Challenge: BRCA Exchange as a global resource for variants in BRCA1 and BRCA2. *PLoS Genet.* 14, e1007752 (2018).
336. Institut Curie. *Study of Family COsegregation of Nucleotide VARiants in the Panel of Genes to Validate Their Use in Genetic Counseling.* <https://clinicaltrials.gov/ct2/show/NCT01689584> (2021).
337. Caputo, S. M. *et al.* Classification of 101 BRCA1 and BRCA2 variants of uncertain significance by cosegregation study: A powerful approach. *Am. J. Hum. Genet.* 108, 1907–1923 (2021).
338. Graeser, M. *et al.* A marker of homologous recombination predicts pathologic complete response to neoadjuvant chemotherapy in primary breast cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 16, 6159–6168 (2010).
339. Dvinge, H. Regulation of alternative mRNA splicing: old players and new perspectives. *FEBS Lett.* 592, 2987–3006 (2018).
340. Escobar-Hoyos, L., Knorr, K. & Abdel-Wahab, O. Aberrant RNA Splicing in Cancer. *Annu. Rev. Cancer Biol.* 3, 167–185 (2019).
341. Colombo, M. *et al.* Comprehensive annotation of splice junctions supports pervasive alternative splicing at the BRCA1 locus: a report from the ENIGMA consortium. *Hum. Mol. Genet.* 23, 3666–3680 (2014).
342. Romero, A. *et al.* BRCA1 Alternative splicing landscape in breast tissue samples. *BMC Cancer* 15, 219 (2015).
343. Fackenthal, J. D. *et al.* Naturally occurring BRCA2 alternative mRNA splicing events in clinically relevant samples. *J. Med. Genet.* 53, 548–558 (2016).
344. Tress, M. L., Abascal, F. & Valencia, A. Alternative Splicing May Not Be the Key to Proteome Complexity. *Trends Biochem. Sci.* 42, 98–110 (2017).
345. Scotti, M. M. & Swanson, M. S. RNA mis-splicing in disease. *Nat. Rev. Genet.* 17, 19–32 (2016).
346. Baralle, D., Lucassen, A. & Buratti, E. Missed threads. The impact of pre-mRNA splicing defects on clinical practice. *EMBO Rep.* 10, 810–816 (2009).

347. Théry, J. C. *et al.* Contribution of bioinformatics predictions and functional splicing assays to the interpretation of unclassified variants of the BRCA genes. *Eur. J. Hum. Genet. EJHG* 19, 1052–1058 (2011).
348. Gaildrat, P. *et al.* Multiple sequence variants of BRCA2 exon 7 alter splicing regulation. *J. Med. Genet.* 49, 609–617 (2012).
349. Di Giacomo, D. *et al.* Functional analysis of a large set of BRCA2 exon 7 variants highlights the predictive value of hexamer scores in detecting alterations of exonic splicing regulatory elements. *Hum. Mutat.* 34, 1547–1557 (2013).
350. de la Hoya, M. *et al.* Combined genetic and splicing analysis of BRCA1 c.[594-2A>C; 641A>G] highlights the relevance of naturally occurring in-frame transcripts for developing disease gene variant classification algorithms. *Hum. Mol. Genet.* 25, 2256–2268 (2016).
351. Colombo, M. *et al.* The BRCA2 c.68-7T > A variant is not pathogenic: A model for clinical calibration of spliceogenicity. *Hum. Mutat.* 39, 729–741 (2018).
352. Caux-Moncoutier, V. *et al.* Impact of BRCA1 and BRCA2 variants on splicing: clues from an allelic imbalance study. *Eur. J. Hum. Genet. EJHG* 17, 1471–1480 (2009).
353. Davy, G. *et al.* Detecting splicing patterns in genes involved in hereditary breast and ovarian cancer. *Eur. J. Hum. Genet. EJHG* 25, 1147–1154 (2017).
354. Brandão, R. D. *et al.* Targeted RNA-seq successfully identifies normal and pathogenic splicing events in breast/ovarian cancer susceptibility and Lynch syndrome genes. *Int. J. Cancer* 145, 401–414 (2019).
355. Raponi, M., Douglas, A. G. L., Tammamo, C., Wilson, D. I. & Baralle, D. Evolutionary constraint helps unmask a splicing regulatory region in BRCA1 exon 11. *PLoS One* 7, e37255 (2012).
356. Acedo, A., Hernández-Moro, C., Curiel-García, Á., Díez-Gómez, B. & Velasco, E. A. Functional classification of BRCA2 DNA variants by splicing assays in a large minigene with 9 exons. *Hum. Mutat.* 36, 210–221 (2015).
357. Lin, J.-H. *et al.* First estimate of the scale of canonical 5' splice site GT>GC variants capable of generating wild-type transcripts. *Hum. Mutat.* 40, 1856–1873 (2019).
358. Yang, Y. & Sharan, S. K. A simple two-step, 'hit and fix' method to generate subtle mutations in BACs using short denatured PCR fragments. *Nucleic Acids Res.* 31, e80 (2003).
359. Kuznetsov, S. G., Liu, P. & Sharan, S. K. Mouse embryonic stem cell-based functional assay to evaluate mutations in BRCA2. *Nat. Med.* 14, 875–881 (2008).
360. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821 (2012).

361. Hryhorowicz, M., Lipiński, D., Zeyland, J. & Słomski, R. CRISPR/Cas9 Immune System as a Tool for Genome Engineering. *Arch. Immunol. Ther. Exp. (Warsz.)* 65, 233–240 (2017).
362. Makarova, K. S. *et al.* Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9, 467–477 (2011).
363. Makarova, K. S. *et al.* An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13, 722–736 (2015).
364. Zetsche, B. *et al.* Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163, 759–771 (2015).
365. Ming, M. *et al.* CRISPR-Cas12b enables efficient plant genome engineering. *Nat. Plants* 6, 202–208 (2020).
366. Findlay, G. M., Boyle, E. A., Hause, R. J., Klein, J. C. & Shendure, J. Saturation editing of genomic regions by multiplex homology-directed repair. *Nature* 513, 120–123 (2014).
367. Findlay, G. M. *et al.* Accurate classification of BRCA1 variants with saturation genome editing. *Nature* 1 (2018) doi:10.1038/s41586-018-0461-z.
368. Schoenfeld, A. J. *et al.* Tumor Analyses Reveal Squamous Transformation and Off-Target Alterations As Early Resistance Mechanisms to First-line Osimertinib in EGFR-Mutant Lung Cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 26, 2654–2663 (2020).
369. Cohen, D. *et al.* Optimizing Mutation and Fusion Detection in NSCLC by Sequential DNA and RNA Sequencing. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* 15, 1000–1014 (2020).
370. ClinVar. at <https://www.ncbi.nlm.nih.gov/clinvar/>.
371. Chang, Y.-F., Imam, J. S. & Wilkinson, M. F. The nonsense-mediated decay RNA surveillance pathway. *Annu. Rev. Biochem.* 76, 51–74 (2007).
372. Li, L. *et al.* Functional redundancy of exon 12 of BRCA2 revealed by a comprehensive analysis of the c.6853A>G (p.I2285V) variant. *Hum. Mutat.* 30, 1543–1550 (2009).
373. Thirthagiri, E. *et al.* BRCA2 minor transcript lacking exons 4-7 supports viability in mice and may account for survival of humans with a pathogenic biallelic mutation. *Hum. Mol. Genet.* 25, 1934–1945 (2016).
374. Mesman, R. L. S. *et al.* Alternative mRNA splicing can attenuate the pathogenicity of presumed loss-of-function variants in BRCA2. *Genet. Med. Off. J. Am. Coll. Med. Genet.* 22, 1355–1365 (2020).
375. Chu, V. T. *et al.* Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* 33, 543–548 (2015).

376. Bell, D. W. & Ellenson, L. H. Molecular Genetics of Endometrial Carcinoma. *Annu. Rev. Pathol.* 14, 339–367 (2019).
377. Wang, F. *et al.* Evaluation of POLE and POLD1 Mutations as Biomarkers for Immunotherapy Outcomes Across Multiple Cancer Types. *JAMA Oncol.* 5, 1504–1506 (2019).
378. Cancer Genome Atlas Research Network *et al.* Integrated genomic characterization of endometrial carcinoma. *Nature* 497, 67–73 (2013).
379. Talhouk, A. & McAlpine, J. N. New classification of endometrial cancers: the development and potential applications of genomic-based classification in research and clinical care. *Gynecol. Oncol. Res. Pract.* 3, 14 (2016).
380. Palles, C. *et al.* Germline mutations in the proof-reading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat. Genet.* 45, 136–144 (2013).
381. Campbell, B. B. *et al.* Comprehensive Analysis of Hypermutation in Human Cancer. *Cell* 171, 1042–1056.e10 (2017).
382. Collias, D. & Beisel, C. L. CRISPR technologies and the search for the PAM-free nuclease. *Nat. Commun.* 12, 555 (2021).
383. Favalli, V. *et al.* Machine learning-based reclassification of germline variants of unknown significance: The RENOVO algorithm. *Am. J. Hum. Genet.* 108, 682–695 (2021).

TABLE DES FIGURES

Figure 1 : Les cancers en chiffres en 2018 en France, données de l'INCa ⁵	8
Figure 2 : Caractéristiques des cellules tumorales et cibles des petites molécules inhibitrices, d'après Bedard et al., 2020 ¹⁷	13
Figure 3 : Causes et conséquences de l'instabilité génétique	16
Figure 4 : Charge mutationnelle par type de pathologie, d'après Lawrence et al., 2013 ²⁷	16
Figure 5 : Les applications du NGS en génétique humaine ³³	19
Figure 6 : Principe du NGS, adapté de Hess et al., 2020 ³³	22
Figure 7 : Comparaison des performances des lectures courtes NGS et des lectures longues TGS, adapté de van Dijk et al., 2018 ³⁴	23
Figure 8 : Arbre décisionnel pour la classification des variants somatiques, d'après Koepfel, F. et al. ⁴⁵	26
Figure 9 : Fréquence des altérations génomiques retrouvées dans les CBNPC	35
Figure 10 : La famille Epidermal Growth Factor Receptor et ses voies de signalisation. D'après Mitsudomi, T. & Yatabe, 2007 ¹⁰⁹	38
Figure 11. Mutations dans le domaine tyrosine-kinase du gène EGFR, d'après Yasuda et al., 2012 ¹¹¹	39
Figure 12 : Impact des délétions et insertions sur l'activation de l'EGFR, d'après Vyse & Huang, 2019 ¹³⁹	42
Figure 13. Mutations secondaires EGFR. Adapté de Reita et al, 2021 ¹⁶⁰	44
Figure 14 : Mécanismes de résistance aux ITK-EGFR dépendants de la génération d'ITK-EGFR et de la ligne de thérapie ciblée. D'après Reita et al, 2021 ¹⁶⁰	45
Figure 15. Rôle des iPARP dans la létalité synthétique d'après Sonnenblick, Amir et al. ²¹³	54
Figure 16 : Les principaux acteurs protéiques de la réparation de l'ADN par recombinaison homologue, d'après Roy et al. ²¹⁸	57
Figure 17 : Les étapes de la réparation des cassures double brin de l'ADN par recombinaison homologue, d'après van Gent et al. ²²²	58
Figure 18. Structure de domaines fonctionnels de BRCA1 et BRCA2, d'après Roy et al. ²¹⁸	61
Figure 19 : Principe des trois biomarqueurs constituant le GIS de la solution commerciale myChoice® HRD, d'après Rodrigues et al., 2016 ³⁰²	68
Figure 20 : Principales altérations moléculaires identifiées dans les HGSOC, d'après Hollis et Gourley, 2016 ³¹³	71
Figure 21 : Profil moléculaire de 149 cancers du pancréas (données TCGA) ³²⁶	75

Figure 22 : Caractéristiques moléculaires des différents stades de cancer de la prostate ³³⁰	77
Figure 23 : Distribution des mutations pathogènes de gènes de réparation de l'ADN impliquées dans les PCa métastatiques. ³³²	77
Figure 24 : Augmentation du nombre de variants dans la base de données FrOG depuis 2012, d'après Caputo et al., 2021 ³³⁷	80
Figure 25 : Les trois étapes du système immunitaire adaptatif bactérien CRISPR-Cas : acquisition, biogénèse du crRNA et interférence de l'ADN viral ³⁶¹	84
Figure 26 : Schéma expérimental de SGE introduisant tous les SNV dans les domaines RING et BRCT de BRCA1. D'après Findlay et al., 2018 ³⁶⁷	86
Figure 27 : Approches parallèle et séquentielle pour l'analyse moléculaire des CBNPC.	104

TABLE DES TABLEAUX

Tableau 1 : Les thérapies ciblées autorisées en France en oncologie regroupées par mécanismes d'action (liste non exhaustive) ^{9,18-20}	14
Tableau 2 : Principales différences entre les technologies NGS Illumina et Ion Torrent. ^{38,39}	22
Tableau 3 : Bases de données de populations contrôles et leurs spécificités	27
Tableau 4 : Bases de données patients et leurs spécificités	28
Tableau 5 : Classification ESCAT. D'après Mateo et al., 2018 ⁷⁵	31
Tableau 6. Thérapies ciblées disponibles en France dans les CBNPC	36
Tableau 7. Essais de phase III comparant les ITK-EGFR de première génération à un doublet de chimiothérapie chez des patients porteurs d'une mutation de l'EGFR ¹¹⁸⁻¹²² ..	40
Tableau 8 : Facteurs de résistance primaire aux ITK-EGFR	43
Tableau 9 : Principaux mécanismes de résistance acquise aux ITK-EGFR	44
Tableau 10 : Techniques actuellement utilisées pour la détection des mutations EGFR ..	51
Tableau 11. <i>iPARP autorisés en France en oncologie et biomarqueur de HRD associé</i> ¹⁸⁻²⁰	54
Tableau 12 : Quelques fréquences de HRD publiées sur de nombreux échantillons (n>100)	66
Tableau 13. Disponibilité des iPARP en France dans le cancer de l'ovaire (mars 2022) ^{18,19}	72
Tableau 14 : Caractéristiques moléculaires des sous-types de cancer du sein (données TCGA 2012) ²⁸⁶	73

REMERCIEMENTS

A M. le Professeur Alain Morel qui a dirigé ce travail avec autant de disponibilité. Je vous exprime toute ma gratitude pour votre enseignement et pour les échanges que nous avons eus au quotidien au laboratoire. C'est en travaillant à vos côtés que j'ai orienté mon activité vers la génétique tumorale, discipline passionnante et stimulante, veuillez trouver ici toute ma reconnaissance et mon respect.

A M. le Professeur Mario Campone, directeur de l'Institut de Cancérologie de l'Ouest, qui m'a fait l'honneur de présider ce jury de thèse. Je vous suis reconnaissante pour toute la confiance que vous m'accordez dans cette belle institution qu'est l'ICO.

A Mme le Professeur Marie-Dominique Galibert et à Mme le Professeur Lucie Karayan-Toker d'avoir accepté, en tant que rapporteurs, d'évaluer ce travail et d'y apporter leur expertise.

A Mme le Professeur Marie-Christine Copin, Mme le Docteur Sophie Krieger et M. le Docteur Olivier Molinier d'avoir accepté de faire partie de mon jury de thèse et d'examiner mon travail.

A l'équipe du laboratoire de Biologie Médicale de l'ICO : Clémence, Christophe et Guillaume, mes collègues biologistes, Amandine, Jonathan et toute l'équipe du laboratoire. J'ai beaucoup apprécié l'opportunité de travailler avec vous tous dans une atmosphère stimulante, amicale et bienveillante.

A mes collègues anatomopathologistes, Véronique, Isabelle et Farida, ainsi qu'à mes collègues du service d'oncogénétique, pour tous nos échanges et votre constante disponibilité.

Aux membres de l'équipe 4 de l'unité CRC12NA INSERM U1307, et à M. le Docteur Yves Lacroix, responsable de l'équipe.

A Mme le Docteur Isabelle Tournier et M. le Professeur Vincent Procaccio, qui ont accepté de faire partie de mon comité de suivi de thèse, pour vos conseils.

A mes parents, ma sœur et toute ma famille, HQOSA.

A Cédric, mon mari, pour tes encouragements, ton soutien et pour tout ce que nous faisons ensemble au quotidien.

A Louis, Alexandre et Maxence, mes trois garçons chéris.

L'auteur du présent document vous autorise à le partager, reproduire, distribuer et communiquer selon les conditions suivantes :



- Vous devez le citer en l'attribuant de la manière indiquée par l'auteur (mais pas d'une manière qui suggérerait qu'il approuve votre utilisation de l'œuvre).
- Vous n'avez pas le droit d'utiliser ce document à des fins commerciales.
- Vous n'avez pas le droit de le modifier, de le transformer ou de l'adapter.

**Consulter la licence creative commons complète en français :
<http://creativecommons.org/licences/by-nc-nd/2.0/fr/>**

Ces conditions d'utilisation (attribution, pas d'utilisation commerciale, pas de modification) sont symbolisées par les icônes positionnées en pied de page.



ENGAGEMENT DE NON PLAGIAT

Je, soussignée CHIRON-CHEVALIER Louise-Marie
déclare être pleinement conscient(e) que le plagiat de documents ou d'une
partie d'un document publiée sur toutes formes de support, y compris l'internet,
constitue une violation des droits d'auteur ainsi qu'une fraude caractérisée.
En conséquence, je m'engage à citer toutes les sources que j'ai utilisées
pour écrire ce rapport ou mémoire.

signé par l'étudiant(e) le **11 / 07 / 2022**

**Cet engagement de non plagiat doit être signé et joint
à tous les rapports, dossiers, mémoires.**

Présidence de l'université
40 rue de rennes – BP 73532
49035 Angers cedex
Tél. 02 41 96 23 23 | Fax 02 41 96 23 00



Titre : Apport du séquençage haut débit et de la caractérisation fonctionnelle de variants somatiques pour l'accessibilité aux thérapies ciblées en oncologie

Mots clés : thérapie ciblée, *EGFR*, *BRCA1/2*, VSI, test fonctionnel, CRISPR-Cas9

Résumé : Le récent changement de paradigme lié au développement de la médecine personnalisée en oncologie nécessite une évolution des pratiques impliquant les laboratoires de génétique moléculaire des cancers.

Le premier volet de ce travail repose sur les enjeux technologiques de la détection des mutations du gène *EGFR* pour permettre une prise en charge optimisée des patients éligibles aux thérapies ciblées dans les cancers bronchiques non à petites cellules au diagnostic et lors des récidives. Le séquençage haut débit s'est ici présenté comme la meilleure alternative pour identifier les mutations de sensibilité et de résistance acquise de l'*EGFR*.

Le deuxième volet de cette thèse concerne

l'interprétation biologique des altérations génétiques de la tumeur, en particulier les variants des gènes suppresseurs de tumeur *BRCA1* et *BRCA2* conditionnant la sensibilité aux thérapies ciblées par inhibiteurs de PARP. Deux approches de caractérisation fonctionnelle de variants de signification incertaine (VSI) ont été utilisées, toutes deux réalisables dans un temps compatible avec les impératifs cliniques. La première consiste à caractériser des défauts d'épissage à partir d'ARN tumoral du patient par RT-PCR et clonage rapide des produits de RT-PCR. La deuxième est une évaluation rapide *in vitro* des VSI par édition du génome en utilisant la technologie CRISPR-Cas9 et un score fonctionnel.

Title : Contribution of new generation sequencing and functional characterization of somatic variants for the accessibility of targeted therapies in oncology

Keywords : targeted therapies, *EGFR*, *BRCA1/2*, VUS, functional impact, CRISPR-Cas9

Abstract : The recent paradigm shift related to the development of personalized medicine in oncology requires an evolution of practices involving cancer molecular genetic laboratories.

The first part of this work is based on the technological challenges of *EGFR* gene mutation detection to enable optimized management of patients eligible for targeted therapies in non-small cell lung cancer at diagnosis and recurrence. New generation sequencing was the best alternative to identify *EGFR* sensitivity and acquired resistance mutations.

The second part of this work concerns the biological interpretation of genetic alterations in

the tumor, in particular variants in the tumor suppressor genes *BRCA1* and *BRCA2* conditioning the sensitivity to PARP inhibitor targeted therapies. Two approaches for functional characterization of variants of uncertain significance (VUS) were used, both of which can be achieved in a time frame compatible with clinical imperatives. The first one consists in characterizing splicing defects from patient tumor RNA by RT-PCR and rapid cloning of RT-PCR products. The second is a rapid *in vitro* evaluation of VUS by genome editing using CRISPR-Cas9 technology and a functional score.