



HAL
open science

Signification des variants génétiques à faible ratio allélique détectés par séquençage à haut débit dans le cadre du diagnostic moléculaire des prédispositions aux cancers du sein et de l'ovaire : mosaïque, hématopoïèse clonale ou ADN tumoral circulant

Flavie Boulouard

► To cite this version:

Flavie Boulouard. Signification des variants génétiques à faible ratio allélique détectés par séquençage à haut débit dans le cadre du diagnostic moléculaire des prédispositions aux cancers du sein et de l'ovaire : mosaïque, hématopoïèse clonale ou ADN tumoral circulant. Médecine humaine et pathologie. 2019. dumas-02331669

HAL Id: dumas-02331669

<https://dumas.ccsd.cnrs.fr/dumas-02331669>

Submitted on 24 Oct 2019

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.

UNIVERSITÉ de CAEN NORMANDIE

FACULTÉ de MÉDECINE

Année 2018/2019

THÈSE POUR L'OBTENTION
DU GRADE DE DOCTEUR EN MÉDECINE

Présentée et soutenue publiquement le : mardi 24 septembre 2019

par

Mme BOULOUARD Flavie

Né (e) le 06/04/1990 à Caen (14)

TITRE DE LA THÈSE :

Signification des variants génétiques à faible ratio allélique détectés par séquençage à haut débit dans le cadre du diagnostic moléculaire des prédispositions aux cancers du sein et de l'ovaire : mosaïque, hématopoïèse clonale ou ADN tumoral circulant.

Président : Monsieur le Professeur Stéphane ALLOUCHE

Membres : Monsieur le Professeur Thomas FRERET

Monsieur le Docteur Nicolas GRUCHY

Monsieur le Docteur Laurent CASTERA

Monsieur le Docteur Dominique VAUR

Monsieur le Docteur Matthieu DECAMP

Directeur de thèse : Dr Laurent CASTERA

Année Universitaire 2018/2019**Doyen**

Professeur Emmanuel TOUZÉ

Assesseurs

Professeur Paul MILLIEZ (pédagogie)

Professeur Guy LAUNOY (recherche)

Professeur Sonia DOLLFUS & Professeur Evelyne EMERY (3^{ème} cycle)**Directrice administrative**

Madame Sarah CHEMTOB

PROFESSEURS DES UNIVERSITÉS - PRATICIENS HOSPITALIERS

M.	AGOSTINI Denis	Biophysique et médecine nucléaire
M.	AIDE Nicolas	Biophysique et médecine nucléaire
M.	ALLOUCHE Stéphane	Biochimie et biologie moléculaire
M.	ALVES Arnaud	Chirurgie digestive
M.	AOUBA Achille	Médecine interne
M.	BABIN Emmanuel	Oto-Rhino-Laryngologie
M.	BÉNATEAU Hervé	Chirurgie maxillo-faciale et stomatologie
M.	BENOIST Guillaume	Gynécologie - Obstétrique
M.	BERGER Ludovic	Chirurgie vasculaire
M.	BERGOT Emmanuel	Pneumologie
M.	BIBEAU Frédéric	Anatomie et cytologie pathologique
Mme	BRAZO Perrine	Psychiatrie d'adultes
M.	BROUARD Jacques	Pédiatrie
M.	BUSTANY Pierre	Pharmacologie
Mme	CHAPON Françoise	Histologie, Embryologie
Mme	CLIN-GODARD Bénédicte	Médecine et santé au travail
M.	COQUEREL Antoine	Pharmacologie
M.	DAO Manh Thông	Hépatologie-Gastro-Entérologie
M.	DAMAJ Ghandi Laurent	Hématologie
M.	DEFER Gilles	Neurologie
M.	DELAMILLIEURE Pascal	Psychiatrie d'adultes
M.	DENISE Pierre	Physiologie
M.	DERLON Jean-Michel	Neurochirurgie
Mme	DOLLFUS Sonia	Psychiatrie d'adultes
M.	DREYFUS Michel	Gynécologie - Obstétrique
M.	DU CHEYRON Damien	Réanimation médicale

Mme	ÉMERY Evelyne	Neurochirurgie
M.	ESMAIL-BEYGUI Farzin	Cardiologie
Mme	FAUVET Raffaèle	Gynécologie – Obstétrique
M.	FISCHER Marc-Olivier	Anesthésiologie – réanimation et médecine péri-opératoire
M.	GÉRARD Jean-Louis	Anesthésiologie – réanimation et médecine péri-opératoire
M.	GUILLOIS Bernard	Pédiatrie
Mme	GUITTET-BAUD Lydia	Epidémiologie, économie de la santé et prévention
M.	HABRAND Jean-Louis	Cancérologie option Radiothérapie
M.	HAMON Martial	Cardiologie
Mme	HAMON Michèle	Radiologie et imagerie médicale
M.	HANOUIZ Jean-Luc	Anesthésiologie – réanimation et médecine péri-opératoire
M.	HULET Christophe	Chirurgie orthopédique et traumatologique
M.	HURAUULT de LIGNY Bruno	Néphrologie
	Éméritat jusqu'au 31/01/2020	
M.	ICARD Philippe	Chirurgie thoracique et cardio-vasculaire
M.	JOIN-LAMBERT Olivier	Bactériologie - Virologie
Mme	JOLY-LOBBEDEZ Florence	Cancérologie
M.	JOUBERT Michael	Endocrinologie
Mme	KOTTLER Marie-Laure	Biochimie et biologie moléculaire
M.	LAUNOY Guy	Epidémiologie, économie de la santé et prévention
M.	LE COUTOUR Xavier	Epidémiologie, économie de la santé et prévention
M.	LE HELLO Simon	Bactériologie-Virologie
Mme	LE MAUFF Brigitte	Immunologie
M.	LEPORRIER Michel	Hématologie
	Éméritat jusqu'au 31/08/2020	
M.	LEROY François	Rééducation fonctionnelle
M.	LOBBEDEZ Thierry	Néphrologie
M.	MANRIQUE Alain	Biophysique et médecine nucléaire
M.	MARCÉLLI Christian	Rhumatologie
M.	MARTINAUD Olivier	Neurologie
M.	MAUREL Jean	Chirurgie générale
M.	MILLIEZ Paul	Cardiologie
M.	MOREAU Sylvain	Anatomie/Oto-Rhino-Laryngologie
M.	MOUTEL Grégoire	Médecine légale et droit de la santé
M.	NORMAND Hervé	Physiologie
M.	PARIENTI Jean-Jacques	Biostatistiques, info. médicale et tech. de communication
M.	PELAGE Jean-Pierre	Radiologie et imagerie médicale
Mme	PIQUET Marie-Astrid	Nutrition
M.	QUINTYN Jean-Claude	Ophthalmologie
M.	RAVASSE Philippe	Chirurgie infantile
M.	REZNIK Yves	Endocrinologie

M.	ROD Julien	Chirurgie infantile
M.	ROUPIE Eric	Médecine d'urgence
Mme	THARIAT Juliette	Radiothérapie
M.	TILLOU Xavier	Urologie
M.	TOUZÉ Emmanuel	Neurologie
M.	TROUSSARD Xavier	Hématologie
Mme	VABRET Astrid	Bactériologie - Virologie
M.	VERDON Renaud	Maladies infectieuses
Mme	VERNEUIL Laurence	Dermatologie
M.	VIADER Fausto	Neurologie
M.	VIVIEN Denis	Biologie cellulaire

PROFESSEUR DES UNIVERSITÉS

PROFESSEUR ASSOCIÉ DES UNIVERSITÉS A TEMPS PLEIN

M.	VABRET François	Addictologie
----	-----------------	--------------

PROFESSEURS ASSOCIÉS DES UNIVERSITÉS A MI-TEMPS

M.	de la SAYETTE Vincent	Neurologie
Mme	DOMPMARTIN-BLANCHÈRE Anne	Dermatologie
Mme	LESCURE Pascale	Gériatrie et biologie du vieillissement
M.	SABATIER Rémi	Cardiologie

PRCE

Mme	LELEU Solveig	Anglais
-----	---------------	---------

Année Universitaire 2018 / 2019**Doyen**

Professeur Emmanuel TOUZÉ

Assesseurs

Professeur Paul MILLIEZ (pédagogie)

Professeur Guy LAUNOY (recherche)

Professeur Sonia DOLLFUS & Professeur Evelyne EMERY (3^{ème} cycle)**Directrice administrative**

Madame Sarah CHEMTOB

MAITRES DE CONFERENCES DES UNIVERSITÉS - PRATICIENS HOSPITALIERS

M.	ALEXANDRE Joachim	Pharmacologie clinique
Mme	BENHĀĪM Annie	Biologie cellulaire
M.	BESNARD Stéphane	Physiologie
Mme	BONHOMME Julie	Parasitologie et mycologie
M.	BOUVIER Nicolas	Néphrologie
M.	COULBAULT Laurent	Biochimie et Biologie moléculaire
M.	CREVEUIL Christian	Biostatistiques, info. médicale et tech. de communication
M.	DE BOYSSON Hubert	Médecine interne
Mme	DEBRUYNE Danièle <small>Éméritat jusqu'au 31/08/2019</small>	Pharmacologie fondamentale
Mme	DERLON-BOREL Annie <small>Éméritat jusqu'au 31/08/2020</small>	Hématologie
Mme	DINA Julia	Bactériologie - Virologie
Mme	DUPONT Claire	Pédiatrie
M.	ÉTARD Olivier	Physiologie
M.	GABEREL Thomas	Neurochirurgie
M.	GRUCHY Nicolas	Génétique
M.	GUÉNOLÉ Fabian	Pédopsychiatrie
M.	HITIER Martin	Anatomie - ORL Chirurgie Cervico-faciale
M.	ISNARD Christophe	Bactériologie Virologie
M.	LEGALLOIS Damien	Cardiologie
Mme	LELONG-BOULOUARD Véronique	Pharmacologie fondamentale
Mme	LEPORRIER Nathalie <small>Éméritat jusqu'au 31/10/2020</small>	Génétique
Mme	LEVALLET Guénaëlle	Cytologie et Histologie
M.	LUBRANO Jean	Chirurgie générale
M.	MITTRE Hervé	Biologie cellulaire
M.	REPESSÉ Yohann	Hématologie

M.	SESBOÛÉ Bruno	Physiologie
M.	TOUTIRAIS Olivier	Immunologie
M.	VEYSSIERE Alexis	Chirurgie maxillo-faciale et stomatologie

MAITRES DE CONFERENCES ASSOCIÉS DES UNIVERSITÉS A MI-TEMPS

Mme	ABBATE-LERAY Pascale	Médecine générale
M.	COUETTE Pierre-André (fin 31/08/19)	Médecine générale
M.	LE BAS François (fin 31/08/19)	Médecine générale
M.	SAINMONT Nicolas (fin 31/08/19)	Médecine générale
Mme	NOEL DE JAEGER Sophie (fin 31/08/2021)	Médecine générale

Remerciements

Monsieur le Président du jury, merci de me faire l'honneur d'accepter de juger ma thèse.

Merci à l'ensemble des membres du jury d'avoir accepté aussi de juger ce travail.

Un grand merci à mon directeur de thèse, Laurent Castera. Tu m'as guidée tout au long de ce projet avec un grand investissement et une grande gentillesse. Merci aussi pour tes remarques très constructives lors de la correction du manuscrit.

Je vous remercie Monsieur Dominique Vaur pour votre investissement dans ce projet, les dossiers de mosaïques et les publications intéressantes que, mine de rien, vous me laissiez de côté. Un grand merci aussi pour votre confiance. Je suis très heureuse de faire partie de votre équipe en novembre.

Un grand merci à toi Angelina. Tu m'as beaucoup aidée pour ce projet. Merci pour ta grande réactivité notamment dans mes commandes de kits et d'amorces. Je t'ai souvent sollicitée. C'est un vrai plaisir de travailler avec toi.

Merci beaucoup à toi Valentin Harter pour la partie statistique. Tu m'as beaucoup aidé dans un domaine que je ne maîtrise pas bien. Et tu as même trouvé des $p < 0,05$! Ça, c'est chouette !

Merci à toi Étienne. Tu m'as bien aidé à démarrer dans le projet en me donnant un cours de bio-informatique notamment.

Un grand merci à Alexandre, Anne Laurence, Arnaud, Aurore, Chann, Céline, Florian, Julien, Laurence, Nicolas, Olivia, Robin, Sosthène et Thibaut. Vous avez tous participé à l'élaboration du projet et m'avez toujours aidée avec sympathie et enthousiasme. Je suis ravie de travailler avec vous bientôt.

Merci à vous Sophie Krieger et Agathe Ricou. Vous avez toujours été à l'écoute et c'est également une grande fierté de travailler avec vous bientôt.

Un grand merci à l'ensemble de l'équipe du CHU de Rouen du Professeur Thierry Frebourg pour leur aide et notamment à Stéphanie Baert Desurmont pour avoir partagé un protocole de Snapshot.

Merci à l'ensemble des prescripteurs qui ont tous contribué au projet, en particulier à ceux qui ont été sollicités : Odile Bera, Pascaline Berthet, Louise Crivelli, Stéphanie Chieze Valero, Philippe Denizeau, Dany Galliano, Olivier Ingster et Elodie Lacaze.

Merci à toi Axel pour le soutien que tu m'apportes dans tous mes projets depuis 8 ans bientôt et qui sont parfois des sacrifices. J'ai beaucoup de chance de t'avoir à mes côtés. Merci pour ta gentillesse quotidienne.

Merci, maman, pour m'avoir toujours soutenue depuis 29 ans. Tu as commencé à me pousser à travailler petite et tu me pousses plutôt à me détendre maintenant ! Merci pour ta relecture du manuscrit. Je suis fière de t'avoir comme maman.

Merci à toi papa pour avoir relu mon travail un weekend sous 30°C. Merci pour m'avoir soutenue aussi dans tous mes projets. Je n'ai pas choisi cette discipline par hasard !

Merci à toi Maïlys pour avoir toujours été là. Merci pour m'avoir offert un neveu aussi adorable. Je t'adore.

Merci à Ellen et Thomas. Je vous adore aussi.

Merci Véro pour nous avoir toujours considérées comme ta famille avec Maïlys.

Merci à Denis et Mauricette pour votre soutien et pour m'avoir proposé un resto un jour où Axel m'a abandonnée lâchement pour partir en Australie...:)

Merci Hélène, Anaïs, Éliisa et Coco. Merci à tous mes amis et à mes co-internes de Caen et de Rouen qui sont aussi devenus des amis. Vous avez été importants tout au long de mon parcours et vous continuerez à l'être dans le futur.

Merci aux copines de master 2 et futures collaboratrices pour certaines : Caro, Marine et Zoé.

Merci à ma famille sur laquelle je peux toujours compter.

Merci à toi Ingrid. Tu étais un modèle de tolérance et de respect. Tu continues à m'inspirer chaque jour.

Abréviations

ACMG : American College of Medical Genetics and Genomics
ADN : Acide DésoxyriboNucléique
ADNc : ADN complémentaire
ADNtc : ADN tumoral circulant
ARNm : Acide RiboNucléique messenger
ASXL1 : *ASXL Transcriptional Regulator 1*
ATM : *Ataxia Telangiectasia Mutated*
BAM : Binary Alignment Map
BAP1 : *BRCA1 associated protein 1*
BARD1 : *BRCA1 Associated Ring Domain 1*
BCD : Blépharo-cheilo-odontique
BRCA1 : *Breast Cancer 1*
BRCA2 : *Breast Cancer 2*
BRIP1 : *BRCA1 Interacting Protein c-terminal helicase 1*
CA 125 : Cancer Antigen 125
CA 15-3 : Cancer Antigen 15-3
CanDiD : Cancer Diagnostic Database
CBL : *Calcineurin B-Like*
CDH1 : *Cadhérine 1*
CDKN2A : *Cyclin Dependent Kinase inhibitor 2A*
CEBPA : *CCAAT Enhancer Binding Protein Alpha*
CFB : Centre François Baclesse
CHECK 2 : *Checkpoint Kinase 2*
CHIP : Clonal Hematopoiesis of Indeterminate Potential
CNV : Copy Number Variation
CSL : Cancer du Sein Lobulaire
dbSNP : Single Nucleotide Polymorphism Database
ddNTP : DideoxyriboNucleotide Triphosphate
ddPCR : Droplet Digital PCR
DGOS : Direction Générale de l'Offre de Soins
DNMT3A : *DNA MethylTransferase 3 Alpha*
dNTP : DesoxyriboNucleotide Triphosphate
dPCR : Digital PCR
EDTA : Ethylène-Diamine-Tétra-Acétique
ER : Estrogen Receptor
ETV6 : *ETS variant 6*
ExAC : Exome Aggregation Consortium
EZH2 : *Enhancer of Zest de Polycomb Repressive Complex 2 subunit*
FANCD1 : *Fanconi Anemia Complementation group D1*
FANCN : *Fanconi Anemia Complementation group N*
FANCO : *Fanconi Anemia Complementation group O*
FLT3 : *Fms related Tyrosine kinase 3*

FROG : FRENch OncoGenetic database
GATA2 : *GATA binding protein 2*
GGC : Groupe Génétique et Cancer
GNAS : *GNAS complex locus*
gnomAD : Genome Aggregation Database
HAS : Haute Autorité de Santé
HBOC : Hereditary Breast and Ovarian Cancer
HC : Hématopoïèse clonale
HER 2 : Human Epidermal growth factor Receptor-2
HGVS : Human Genome Variation Society
HNPCC : Hereditary Non-polyposis Colorectal Cancer
HOXB13 : *HOMeoboX B13*
IARC : International Agency for Research on Cancer
IDH1/2 : *isocitrate dehydrogenase (NADP(+))1/2*
INCa : Institut National du Cancer
INHA : *INHibin subunit Alpha*
INBHA : *INHibin subunit Beta A*
InVS : Institut de Veille Sanitaire
IRM : Imagerie par Résonance Magnétique
JAK2 : *Janus Kinase 2*
KRAS : *KRAS Proto-Oncogene, GTPase*
LAM : Leucémie Aigüe Myéloïde
LAM-t : Leucémie Aigüe Myéloïde post-traitement
LFS : Li Fraumeni Syndrom
LOH : Lost Of Heterozygoty
MLH1 : *MutL Homolog 1*
MLH3 : *MutL Homolog 3*
MLPA : Multiplex Ligation-dependent Probe Amplification
MMR : MisMatch Repair
MPL : *MyeloProliferative Leukemia virus*
MRE11A : *MRE11 homolog, double strand break repair nuclease*
MSH2 : *MutS Homolog 2*
MSH6 : *MutS Homolog 6*
MUTYH : *MutY DNA glycosylase*
NBN : *Nibrin*
NCCN : National Comprehensive Cancer Network
NF1 : *Neurofibromin 1*
NF2 : *Neurofibromin 2*
NGS : Next generation sequencing
NPM1 : *Nucleophosmin 1*
NRAS : *NRAS proto-oncogene GTPase*
OMIM : Online Mendelian Inheritance in Man
OR : Odds Ratio
PALB2 : *Partner and Localizer of BRCA2*

PARP : Poly-ADP-Ribose-Polymérase
PB : paires de base
PCR : Polymerase Chain Reaction
PHF6 : *PHD Finger protein 6*
PMS1 : *PMS1 Homolog 1*
PMS2 : *PMS1 Homolog 2*
PPM1D : *Protein Phosphatase, Mg²⁺/Mn²⁺ Dependent 1D*
PTEN : *Phosphatase and TENsin homolog*
QMPSF : Quantitative Multiplex PCR of Short Fluorescent Fragments
RAD50 : *RAD50 double strand break repair protein*
RAD51 : *RAD51 Recombinase*
RAD51B : *RAD51 paralog B*
RAD51C : *RAD51 paralog C*
RAD51D : *RAD51 paralog D*
RB1 : *Retinoblastoma 1*
RCP : Réunion de Concertation Pluridisciplinaire
RH+ : Récepteur Hormonal positif
RIHN : Référentiel des actes Innovants Hors Nomenclature
RINT1 : *RAD50 INTERactor 1*
RR : Risque relatif
RUNX1 : *Runt-related transcription factor 1*
SF3B1 : *Splicing Factor 3b subunit 1*
SRSF2 : *Serine and arginine Rich Splicing Factor 2*
SMD : Syndrome MyéloDysplasique
SNP : Single Nucleotide Polymorphism
SNV : Single Nucleotide Variation
SOLAMEN : Segmental Overgrowth, Lipomatosis, Arteriovenous Malformation and Epidermal Nevus
STAG2 : *Stromal Antigen 2*
STK11 : *Serine/Threonine Kinase 11*
TCGA : The Cancer Genome Atlas
TET2 : *Ten-Eleven-Translocation 2*
TP53 : *Tumor Protein 53*
U2AF1 : *U2 Small Nuclear RNA Auxiliary Factor 1*
UMD : Universal Mutation Database
UV : UltraViolet
VAF : Variant Allele Fraction
VEGF : Vascular Endothelial Growth Factor
VSI : Variant de Signification Inconnue
WT1 : *Wilms Tumor 1*
XRCC2 : *X-Ray repair Cross Complementing 2*
XRCC3 : *X-Ray repair Cross Complementing 3*

Tableaux et Figures

Tableau 1 : Score d'Eisinger	6
Figure 1 : Trois exemples de mécanismes d'inactivation d'un gène dans la tumeur	9
Figure 2 : Représentation du séquençage par la technique Illumina	18
Figure 3 : Représentation de la profondeur et couverture en séquençage haut débit.....	20
Figure 4 : Le principe de l'analyse par Outlyzer.....	21
Figure 5 : Les différentes formes de mosaïques dans les gènes de transmissions autosomiques dominants.....	28
Figure 6 : Représentation de l'ADN tumoral circulant	33
Figure 7: Graphique représentant la quantité d'ADN muté en fonction du type de cancer	35
Figure 8 : La CHIP précurseur des néoplasie hématologiques	37
Tableau 2 : Gènes les plus fréquemment mutés dans la CHIP	38
Tableau 3 : Liste du panel 34 gènes réalisé dans le cadre du panel HBOC au laboratoire.	41
Tableau 4 : Amorces utilisées pour l'amplification en vue du séquençage Sanger	46
Figure 9 : Exemple de représentation d'une réaction d'extension d'amorce (Snapshot).....	47
Tableau 5 : Amorces d'amplification pour la technique Snapshot.....	48
Tableau 6 : Amorces d'extension utilisées pour la technique Snapshot.....	49
Tableau 7 : Différents fluorochromes utilisés pour le marquage des ddNTP pour la technique Snapshot	50
Tableau 8 : Données phénotypiques des patients ayant un ou plusieurs variants de VAF faible avec les résultats des VAF obtenus par en NGS à partir des prélèvements sanguins, des frottis jugaux et des tissus tumoraux et/ sains.	52
Figure 10 : Vérification des variants de ratio allélique supérieur à 10 % par technique Sanger	55
Figure 11 : Résultats des analyses Snapshot et de NGS réalisés chez le patient n°1 à partir du prélèvement sanguin, des cellules jugales et de la tumeur.....	56
Figure 12 : Données obtenues pour le patients n°2 par analyse Snapshot et Tapstation.....	57
Figure 13 : Analyse par Tapestation de la taille des fragments des ADN extraits à partir des prélèvements sanguins	58
Figure 14 : Résultats par analyse Snapshot de l'ensemble des échantillons des patients ayant une VAF supérieure à 5 % et inférieur à 35 %	61
Figure 15 : Cas index sélectionnés pour l'analyse statistique cas/contrôles.....	62
Tableau 9 : Caractéristiques des patients ayant des variants <i>TP53</i> de faible VAF.....	63
Tableau 10 : Evaluation des effets conjoints de l'âge au prélèvement, type de cancer et de la profondeur de couverture moyenne sur <i>TP53</i> dans une analyse multivariée entre le groupe des patients porteurs d'un variant de <i>TP53</i>	64
Figure 16: Diagramme d'aide diagnostique suite à la detection d'un variant potentiellement délétère de faible VAF.....	75
Tableau 11 : Récapitulatif des analyses réalisées chez les différents patients.....	90

Table des matières

INTRODUCTION.....	1
I. Le cancer du sein et de l’ovaire.....	1
1. Le cancer du sein.....	1
a. Données épidémiologiques	1
b. Diagnostic et marqueurs pronostiques	1
c. Prise en charge	3
2. Le cancer de l’ovaire	3
a. Données épidémiologiques	3
b. Diagnostic et histologie.....	3
c. Prise en charge	4
3. Cancer évocateur ou non d’une prédisposition héréditaire	4
II. Indication à la recherche d’une prédisposition aux cancers du sein et de l’ovaire	5
4. Organisation de la consultation d’oncogénétique	5
5. Les critères d’indication de recherche de prédisposition au cancer du sein et de l’ovaire	6
III. Modèle de transmission de la prédisposition aux cancers du sein et de l’ovaire et gènes utiles au diagnostic.....	8
1. Description générale des mécanismes de transmission de la prédisposition et d’apparition du cancer.....	8
2. <i>BRCA1</i> et <i>BRCA2</i>	10
3. <i>PALB2</i>	11
4. <i>RAD51C ET D</i>	11
5. <i>CDH1</i>	11
6. <i>PTEN</i>	11
7. Les gènes MMR	12
8. <i>TP53</i>	13
IV. Détection des individus à risque génétique et prise en charge spécifique.....	15
V. Diagnostic moléculaire chez les cas index : Le séquençage à haut débit	16
1. Historique des techniques antérieures	16
2. Le principe du séquençage à haut débit : exemple de la technique Illumina	16
3. Le traitement des données par analyse bioinformatique	19
4. Les défis de l’interprétation des variants par NGS	22
5. La détection et l’interprétation des variants présentant des VAF faible	23
VI. Les néomutations.....	24

1.	Rappels de principes généraux : mutations héritées / mutations <i>de novo</i> / mosaïques	24
2.	Les néomutations constitutionnelles en pathologie humaine	24
3.	Néomutations dans les gènes impliqués dans le syndrome de prédisposition au cancer du sein et de l’ovaire.....	26
a.	Mutations <i>de novo</i> et néomutations postzygotiques de <i>BRCA1</i> et <i>BRCA2</i>	26
b.	Mutations <i>de novo</i> et mosaïques constitutionnelles du gène <i>TP53</i>	27
c.	Mutations <i>de novo</i> et mosaïques constitutionnelles du gène <i>PTEN</i>	28
d.	Les autres gènes du panel	29
VII. NGS et mutation somatique acquise : sources d’un diagnostic différentiel difficile.....		30
1.	Les mutations tumorales.....	30
a.	Mutations somatiques tumorales testées dans le cancer du sein et de l’ovaire ..	30
b.	Définition de l’ADN tumoral circulant	31
c.	Techniques de détection de l’ADN tumoral circulant.....	33
d.	Applications en cancérologie	34
2.	L’hématopoïèse clonale.....	36
a.	Définition de l’hématopoïèse clonale	36
b.	Recommandations en termes de diagnostic et de suivi	38
VIII. Objectifs de la thèse		39
PATIENTS ET METHODE.....		40
1.	Population étudiée	40
2.	Recherche de variant constitutionnel par séquençage à haut débit	40
3.	Sélection des variants à partir des données de séquençage	42
4.	Analyse par séquençage à haut débit des blocs tumoraux.....	43
5.	Détermination de la taille des ADN extraits et purification des ADN des hauts poids et bas poids moléculaire.....	44
6.	Analyse Sanger.....	45
7.	Technique Snapshot	46
8.	Analyse statistique.....	50
RESULTATS		51
DISCUSSION		65
CONCLUSION.....		74
BIBLIOGRAPHIE		76

INTRODUCTION

I. Le cancer du sein et de l'ovaire

1. Le cancer du sein

a. Données épidémiologiques

Approximativement, une femme sur huit dans la génération actuelle se fera diagnostiquer un cancer du sein dans sa vie (1). C'est la première cause de cancer chez la femme en termes d'incidence, suivie par les cancers colorectaux et le cancer du poumon. Selon les chiffres de l'Institut National du Cancer (INCa), 58459 nouveaux cas de cancers du sein ont été détectés en France métropolitaine en 2018 et 12146 décès ont été recensés faisant de cette pathologie la première cause de décès par cancer chez la femme (14 % des décès féminins par cancer)(2). Les cancers du sein chez l'homme sont rares (1 % des cancers du sein) mais de plus mauvais pronostic. Depuis 2005, on constate toutefois une diminution de la mortalité liée au cancer du sein du fait des progrès thérapeutiques et de la mise en place du dépistage permettant un diagnostic et une prise en charge précoce. Un programme national de dépistage organisé du cancer du sein ayant débuté en 2004 propose ainsi à toutes les femmes âgées de 50 à 74 ans la réalisation d'une mammographie mammaire tous les 2 ans. Ce programme concerne toutes les femmes dont le seul facteur de risque de cancer est l'âge. Ainsi, sur la période 2015-2016, le taux de participation des femmes à ce programme était de 51,1 % ce qui représente environ 5 millions de femmes dépistées (chiffres de l'Institut de Veille Sanitaire (InVS)(3)).

b. Diagnostic et marqueurs pronostiques

Les facteurs pronostiques du cancer du sein sont bien documentés. En effet, en plus des facteurs pronostiques généraux tels que la taille de la tumeur, l'atteinte ganglionnaire, la présence ou non de métastase..., d'autres facteurs pronostiques tels que la positivité ou non au niveau tumoral des récepteurs hormonaux, la surexpression du facteur de croissance des cellules épithéliales humaines « Human Epidermal growth factor Receptor 2 » (HER2), ou l'étude du Ki-67, un marqueur de prolifération cellulaire peuvent être pris en compte (4) (5) (6). Plusieurs sous-types tumoraux ont ainsi été mis en évidence et sont différenciés par l'étude immunohistochimique des différents marqueurs exprimés ou non au niveau de la tumeur :

- Les cancers du sein luminaux A (30 % à 40 % des cancers du sein) : ils expriment de manière importante les récepteurs aux œstrogènes et à la progestérone. Ils sont

négatifs pour le marqueur HER2. Ce sont des cancers hormonodépendants de bon pronostic ne nécessitant pas forcément le recours à la chimiothérapie.

- Les cancers du sein luminaux B (20 % des cancers du sein) : ils sont moins hormonosensibles que les luminaux A et présentent parfois une surexpression légère de HER2. Leur pronostic est moins bon que le sous-type précédant. Ils nécessitent ainsi un traitement par chimiothérapie et hormonothérapie avec parfois du trastuzumab (Herceptin®) en cas de marqueur HER2 surexprimé. Le trastuzumab est un anticorps monoclonal humanisé recombinant de classe IgG1 dirigé contre le récepteur 2 du facteur de croissance épidermique humain (HER2).
- Les cancers HER2 enrichi (20 % des cancers du sein) : ils n'expriment pas de récepteurs aux œstrogènes et à la progestérone mais sur-expriment de manière importante le marqueur HER2. Ce sont des cancers très agressifs qui étaient de mauvais pronostic avant l'arrivée du traitement par trastuzumab (Herceptin®) qui a modifié nettement le pronostic de ces tumeurs.
- Le cancer du sein triple négatif ou « basal like » (20 % des cancers du sein) : il est de mauvais pronostic indifférencié et n'exprime ni les récepteurs hormonaux ni le facteur HER2.

Aussi, des tests pronostiques d'expression génique ont été développés (Prosigna® PAM50 (NanoString), Endopredict® (Myriad)...) au cours de ces dernières années afin de prédire le pronostic de la tumeur et de proposer une prise en charge adaptée. Ces tests de signature génomique sont fondés sur l'identification et la quantification d'acide ribonucléique messenger (ARNm) ou d'acide désoxyribonucléique complémentaire (ADNc) impliqués dans la prolifération tumorale. Par exemple, le test Prosigna® est un test pronostique qui se base sur la signature PAM50 en évaluant l'expression de 50 gènes. Ce test permet de déterminer le sous-type intrinsèque de la tumeur parmi les 4 listés ci-dessus ainsi que le risque de récurrence à 10 ans (7) (8). L'un des intérêts de ces tests est d'évaluer ou non la nécessité d'instaurer un traitement par chimiothérapies pour les cancers luminaux A. En France, en l'absence de preuve d'utilité clinique suffisante à l'heure actuelle, ces tests ne sont pas remboursés par l'assurance maladie actuellement mais uniquement par des financements dérogatoires dans le cadre du référentiel des actes innovants hors nomenclature (RIHN), dans une logique de soutien à l'innovation (Haute Autorité de Santé (HAS) 2019 (9)). Cependant, cela pourrait évoluer dans les années futures.

c. Prise en charge

Le premier traitement du cancer du sein est la chirurgie (tumorectomie ou mastectomie selon le degré d'envahissement). Par ailleurs, la progression des techniques ainsi que la pratique de la chimiothérapie néoadjuvante ont permis de diminuer les indications de mastectomie. De plus, la technique du « ganglion sentinelle » qui consiste à évaluer l'atteinte des ganglions qui drainent directement le sein atteint permet d'évaluer l'atteinte ganglionnaire sans avoir recours au curage axillaire systématique qui occasionne un risque important de lymphœdème. La chirurgie peut être précédée d'une chimiothérapie néoadjuvante et adjuvante. La radiothérapie est souvent indiquée en post chirurgie afin de diminuer le risque de récurrence (6). Enfin, en cas de positivité des récepteurs hormonaux, le tamoxifène, un modulateur sélectif des récepteurs aux œstrogènes ou les inhibiteurs de l'aromatase, enzyme contribuant à la production de l'œstrogène, sont disponibles (10). De plus, le recours à certaines thérapies ciblées telles que le trastuzumab (Herceptine®) pour les tumeurs exprimant le marqueur HER2, le bevacizumab (Avastin®), un « antirécepteur » au facteur de croissance VEGF (« vascular endothelial growth factor ») pour les cancers du sein métastatiques a modifié le pronostic de ces tumeurs (11). De manière globale, 70 % des patients sont indemnes de la maladie 5 ans après le diagnostic (12). Par ailleurs, sur le plan esthétique, les méthodes de reconstruction mammaire ont largement évolué ces dernières années avec différentes techniques proposées comme la reconstruction par prothèse mammaire et la reconstruction par lambeau qui consiste à prélever un lambeau de muscle d'une partie du corps ainsi que de la peau et de la graisse pour reconstruire le sein (13).

2. Le cancer de l'ovaire

a. Données épidémiologiques

Le cancer de l'ovaire est la 7ème cause de cancer chez la femme en termes d'incidence. En 2008, 4430 nouveaux cas avaient été recensés. L'âge médian au moment du diagnostic est de 65 ans. Le cancer de l'ovaire est la 4ème cause de décès par cancer chez la femme. Il a été estimé responsable de 3 000 décès en 2008. Malheureusement, ce cancer est toujours aujourd'hui diagnostiqué tardivement ce qui explique son pronostic souvent sombre. Par ailleurs, il n'existe pas à l'heure actuelle de moyens de dépister efficacement ce cancer. La survie à 5 ans, tous stades confondus, est d'environ 45 % (HAS 2010 (14)).

b. Diagnostic et histologie

La majorité des cancers de l'ovaire sont des tumeurs épithéliales dont les principaux types histologiques sont les carcinomes séreux, endométrioïdes, mucineux, à cellules claires,

mixtes et les tumeurs frontières dites « borderline ». Les patientes présentant une prédisposition génétique au cancer de l’ovaire provoqué par des mutations des gènes *BRCA1* (« *Breast Cancer 1* ») ou *BRCA2* (« *Breast Cancer 2* ») présentent le plus souvent des tumeurs séreuses de haut grade (15) . Les tumeurs de types mixtes (endométriode/mucineux/à cellules claires) sont plus fréquemment retrouvées chez les femmes présentant un syndrome de Lynch, c’est-à-dire une prédisposition héréditaire aux cancers colorectaux héréditaires non polyposiques (HNPCC) (16). Ce cancer se manifeste souvent par des signes tardifs à types de ballonnements intestinaux et de douleurs traduisant souvent une atteinte péritonéale.

c. Prise en charge

Le traitement du cancer de l’ovaire se compose en général d’un traitement chirurgical suivie d’une chimiothérapie par carboplatine et paclitaxel (17). Le cancer de l’ovaire est souvent platino-sensible, particulièrement lorsqu’il s’agit de tumeurs comportant des mutations constitutionnelles ou somatiques des gènes *BRCA1* et *BRCA2* qui sont ainsi de meilleur pronostic. Des thérapies ciblées ont montré une efficacité sur la survie sans progression telles que le bevacizumab (18) (19) et surtout les inhibiteurs de la poly-ADP-ribose-polymérase-1 (PARP) qui ont révolutionné la prise en charge des patientes atteintes d’un cancer de l’ovaire muté *BRCA* (20) (21).

3. Cancer évocateur ou non d’une prédisposition héréditaire

5 à 10 % des cancers du sein (INCa 2016) et 20 % des cancers de l’ovaire (22) sont liés à des altérations génétiques constitutionnelles ; c’est-à-dire à des altérations génétiques présentes, en général, dans l’ensemble des cellules d’un organisme depuis la naissance et potentiellement transmissibles à la descendance. La prédisposition génétique est la présence chez une personne d’un variant génétique appelée communément « mutation » pathogène qui prédispose l’individu à certaines pathologies. Ainsi, les individus porteurs de cette prédisposition auront un risque élevé de développer la pathologie. Dans le syndrome de prédisposition aux cancers du sein et de l’ovaire, les premiers gènes décrits dans ce cadre sont les gènes *BRCA1* identifiés en 1994 (23) et le gène *BRCA2* en 1995 (24). Le syndrome de prédisposition héréditaire aux cancers du sein et de l’ovaire aussi appelé syndrome HBOC (« Hereditary Breast and Ovarian Cancer ») est ainsi provoqué par des altérations génétiques de certains gènes dits « suppresseurs de tumeurs » qui confère un risque augmenté de cancer du sein et de l’ovaire.

L'histoire personnelle et familiale est déterminante pour s'orienter ou non vers une analyse génétique à la recherche de l'événement responsable du trait génétique. En effet, l'observation de plusieurs cas de cancers dans une famille sera suspecte de prédisposition héréditaire. Cependant, un cancer du sein ou de l'ovaire même sporadique, c'est-à-dire sans antécédent familial évocateur, peut aussi être suspect de prédisposition s'il est de survenue précoce ou multiple.

L'identification d'une mutation causale permettra de proposer à l'individu une prise en charge adaptée du patient et de sa famille et un suivi spécifique en fonction du gène muté qui sera détaillé plus loin.

II. Indication à la recherche d'une prédisposition aux cancers du sein et de l'ovaire

1. Organisation de la consultation d'oncogénétique

En France, le diagnostic des prédispositions génétiques et le suivi des personnes à risque élevé ou très élevé de cancer sont mis en œuvre dans le cadre du dispositif national d'oncogénétique. Ce dispositif s'articule autour de 147 sites de consultations répartis dans 104 villes. Par ailleurs, 25 laboratoires répartis sur l'ensemble du territoire permettent la réalisation des tests génétiques. L'objectif de ce dispositif porté par l'INCa, en association avec la Direction Générale de l'Offre de Soins (DGOS, Ministère des Solidarités et de la Santé) est d'identifier toutes les personnes présentant un risque héréditaire de cancer en France et de leur assurer un suivi personnalisé et multidisciplinaire. Lorsqu'une altération génétique est détectée chez un cas index, c'est-à-dire le premier cas testé dans la famille, ses apparentés bénéficient d'un test ciblé à la recherche de la mutation identifiée. Le cas index a l'obligation légale d'informer l'ensemble de ses apparentés s'il est porteur d'une mutation délétère ou il peut déléguer cette obligation à l'onco-généticien. Il s'expose, en effet, à des poursuites en cas de rétention d'information car le médecin est contraint par le secret médical à ne pas divulguer lui-même cette donnée sans consentement. Ainsi, en 2016, 6 417 apparentés ont été testés dans un contexte de syndrome de prédisposition aux cancers du sein et de l'ovaire appelé syndrome HBOC pour 1762 cas index ayant une mutation délétère identifiée parmi les 17821 cas index testés (INCa 2016 (25)).

2. Les critères d'indication de recherche de prédisposition au cancer du sein et de l'ovaire

Actuellement, la recherche de prédisposition aux cancers du sein et de l'ovaire n'est pas réalisée de manière systématique dans la population générale mais chez les personnes dont l'histoire personnelle et/ou familiale est évocatrice d'une prédisposition au cancer. Le premier cas étudié dans la famille est en général le patient ayant eu un cancer et dont l'histoire personnelle est la plus évocatrice dans la famille d'une prédisposition. Cependant, l'analyse chez le cas index optimal n'est pas toujours envisageable pour différentes raisons, par exemple par refus de la personne, décès ou encore absence de contact possible avec cette dernière. Dans ce cas, un autre cas dans la famille est choisi. De même, une analyse peut être réalisée à titre exceptionnel chez un patient non atteint de cancer dit « asymptomatique » en l'absence d'analyse réalisable chez des patients plus évocateurs dans la famille.

Le guide HAS 2014 détaille bien l'organisation du dépistage du cancer du sein et de l'ovaire en France ainsi que la détermination des femmes à haut risque de prédisposition génétique. Tout d'abord, l'indication à adresser ou non un patient en consultation d'oncogénétique est déterminé par le score d'Eisinger (**tableau 1**). Ce score évalue le risque de prédisposition en fonction de l'histoire familiale. Il est destiné à l'ensemble des professionnels de santé et en particulier aux médecins généralistes.

CALCUL DU SCORE D'EISINGER ET CONDUITE À TENIR	
Mutation BRCA1/2 identifiée dans la famille	5
Cancer du sein chez une femme avant 30 ans	4
Cancer du sein chez une femme entre 30 et 39 ans	3
Cancer du sein chez une femme entre 40 et 49 ans	2
Cancer du sein chez une femme entre 50 et 70 ans	1
Cancer du sein chez un homme	4
Cancer de l'ovaire avant 70 ans	4

RÉSULTATS
3 ou + : consultation d'oncogénétique
Inférieur à 3 : dépistage organisé

Tableau 1 : Score d'Eisinger (Source : Quelles modalités de dépistage pour quelles femmes ? INCa, septembre 2015)

Le calcul du score se réalise en additionnant les cotations d'une même branche parentale. Un score supérieur > 3 doit conduire à consulter un onco-généticien.

L'onco-généticien déterminera l'indication à réaliser un test génétique et/ou les modalités de surveillance selon le risque personnel de cancer du sein. Il dispose d'outils de

calculs individuels de risques de cancer du sein tels que le score Bodicea. Ce score permet de différencier les femmes à risque très élevé telles que les femmes mutées *BRCA1* ou les femmes ayant un score supérieur à 20 % et les femmes à risque élevé dont le score est compris entre 12 % et 20 %. De manière générale, pour la prévention du cancer du sein, les femmes à risque élevé bénéficieront de mammographies annuelles à partir de 40 ans tandis que les femmes à risque très élevé passent des imageries par résonance magnétique (IRM) mammaires annuelles à partir de 30 ans ou 5 ans avant le cas familial le plus précoce. Même si les recommandations d'analyse répondent à des critères généraux, chaque laboratoire définit des critères d'indications d'analyse des cas index. Par exemple, au laboratoire du centre François Baclesse (CFB) à Caen, les critères d'indication d'analyse sont les suivants :

- 1 seul cas de cancer du sein :
 - < 36 ans
 - < 51 ans si grade III, triple négatif
 - < 51 ans si cancer bilatéral (dont le 2^{ème} est de survenue avant 71 ans)
 - < 61 ans si cancer du sein médullaire
 - < 71 ans en cas de cancer du sein chez un homme
 - 1 cas de cancer du sein < 51 ans associé à
 - Un cancer de la prostate < 61 ans
 - Ou un cancer du pancréas < 61 ans
 - Ou un cancer de l'ovaire < 71 ans
 - ➔ Chez la même personne ou un apparenté au premier degré
- 1 seul cas de cancer de l'ovaire épithélial
 - < 71 ans
- 2 cas familiaux (unis entre eux au 1^{er} degré ou au 2^{ème} degré en passant par un homme)
 - Au moins un cas < 51 ans et un cas < 71 ans
- 3 cas familiaux (unis entre eux par un lien au 1^{er} degré ou au 2^{ème} degré)
 - Dont au moins un cas < 61 ans

Actuellement, en France, selon les recommandations du Groupe Génétique et Cancer (GGC), dans le cadre du diagnostic moléculaire, les cas index répondant aux critères d'indication de recherche de prédisposition au cancer du sein et de l'ovaire devraient bénéficier d'une analyse d'un panel de 13 gènes. Ceci est rendu possible avec le séquençage à haut débit qui a révolutionné le paysage de la génomique depuis son introduction dans les laboratoires

dans les années 2010. Les apparentés quant à eux bénéficient d'une recherche ciblée de l'altération génétique identifiée chez le cas index par une technique ciblée telle que, par exemple, le séquençage Sanger.

III. Modèle de transmission de la prédisposition aux cancers du sein et de l'ovaire et gènes utiles au diagnostic

Les facteurs de risques génétiques sont variés. Certaines susceptibilités génétiques induisent un faible risque de cancer nettement influencé par les risques environnementaux tandis que d'autres susceptibilités génétiques induisant un risque fort répondent au modèle mendélien de transmission autosomique dominante. Par exemple, des variations pathogènes des gènes *BRCA1* et *BRCA2* induisent un risque fort de prédisposition au cancer du sein tandis que les variants pathogènes des gènes « *Checkpoint kinase 2* » (*CHEK2*) ou « *ataxia telangiectasia mutated* » (*ATM*) induisent un risque plus modéré. De même, le risque génétique peut répondre à un modèle monogénique, c'est-à-dire conféré par une mutation germinale d'un seul gène ou répondre à un modèle multigéniques dû à l'action conjointe de plusieurs gènes provoquant un risque à partir d'un certain seuil. Actuellement, dans le cadre du diagnostic, seuls les facteurs de risque génétique monogéniques élevés sont pris en compte aujourd'hui. Ceci correspond donc plutôt à des modèles mendéliens répondant à une transmission autosomique dominante dont la pénétrance est forte mais néanmoins incomplète. En 2017, le GGC a établi un panel de 13 gènes reconnus d'utilité clinique dans le contexte de prédisposition aux cancers du sein et de l'ovaire.

1. Description générale des mécanismes de transmission de la prédisposition et d'apparition du cancer

Le mode de transmission de cette prédisposition par les différents gènes étudiés actuellement en routine diagnostique est de type autosomique dominant. Ainsi, étant donné que chaque individu transmet la moitié de son patrimoine génétique, une femme ou un homme porteur de la mutation a un risque sur deux de la transmettre à sa descendance.

L'ensemble des gènes testés dont les résultats sont utiles au conseil génétique, c'est-à-dire que le niveau de risque de cancer induit par les variants pathogènes de ces gènes est suffisamment élevé pour les mesures de prise en charge adaptée soit jugée comme bénéfique, sont des gènes dits « supprimeurs de tumeurs ». Les mutations décrites sont généralement rares ayant une fréquence inférieure à 0,1 % dans la population générale et sont des mutations

inactivatrices, de type perte de fonction telles que des mutations « frameshift » ou « non-sens » entraînant l'apparition d'un codon stop prématuré ou des mutations ponctuelles « faux sens » dans des domaines importants pour la fonction de la protéine et suffisantes pour aboutir à une perte de fonction de la protéine. Il existe aussi des évènements de plus grandes tailles dit CNV (« Copy Number Variation ») pouvant entraîner la délétion ou la duplication de tout ou une partie du locus génique.

De manière assez générale, le mécanisme d'apparition du cancer répond au modèle « 2 hits » de Knudson (26). Pour rappel, chaque gène est présent en double au sein de notre patrimoine génétique provenant pour moitié de la mère et pour moitié du père. Dans le syndrome HBOC, la personne prédisposée possède une altération génétique située au niveau d'un gène parmi ceux impliqués dans le syndrome HBOC à l'état hétérozygote. Ainsi, l'ensemble des cellules de l'individu atteint de ce syndrome possède un allèle portant le gène *BRCA1* muté et un allèle porteur du gène *BRCA1* non muté dit sauvage. Si une cellule acquiert par malchance une altération du gène *BRCA1* sauvage, elle ne produira plus de protéine BRCA1 fonctionnelle contribuant à l'apparition du cancer. On parle de modèle « 2 hits ».

Dans le cadre de ces syndromes de prédisposition, l'apparition du cancer est initiée par l'inactivation du second allèle non muté, dit sauvage, le plus souvent par perte d'hétérozygotie soit en anglais le terme « lost of heterozygosity » (« LOH ») lié à un remaniement chromosomique. L'inactivation du second allèle peut être due à l'apparition d'une seconde mutation inactivatrice (« second hit ») entraînant ainsi une perte totale de la fonction de la protéine dans la cellule impliquée, favorisant le processus de cancérogénèse (**Figure 1**) (27).

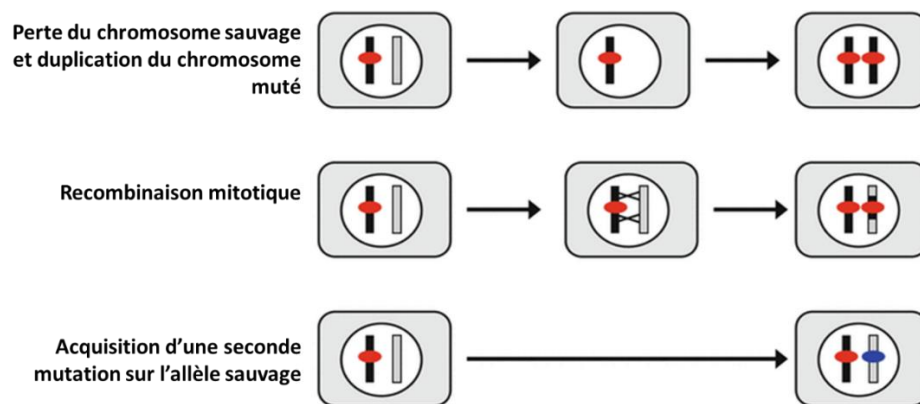


Figure 1 : Trois exemples de mécanismes d'inactivation d'un gène dans la tumeur (image traduite de Bunz, 2016 (27))

2. BRCA1 et BRCA2

Les gènes *BRCA1* et *BRCA2* sont les deux gènes les plus fréquemment impliqués dans le syndrome de prédisposition aux cancers du sein et de l'ovaire. Ils ont été découverts par étude de liaison génétique (« Linkage ») dans des familles très évocatrices de prédisposition au cancer du sein. Ainsi, le locus du gène *BRCA1* en 17q21 a été mis en évidence en 1990 (28). Puis, le locus du gène *BRCA2* en 13q13 a été mis en évidence en 1994 (29). Ces études sont basées sur la ségrégation concomitante des allèles portant des marqueurs polymorphes avec le locus du gène responsable de la maladie, dans plusieurs familles.

Les deux gènes ont fait l'objet à l'origine d'un dépôt de brevet par la société Myriad Genetics (30) engendrant des coûts d'analyses colossaux mais de nombreuses controverses ont conduit à l'invalidation de ces brevets en mars 2010 par la Cour Fédérale des Etats Unis (31). Ce sont aujourd'hui les gènes les plus séquencés dans le monde. Dans la population générale, approximativement une personne sur 400 présentent une mutation du gène *BRCA* mais cette prévalence peut atteindre le ration 1/200 dans certaines populations (32). Les protéines *BRCA1* et *BRCA2* sont connues pour être impliquées dans la réparation des cassures doubles brins de l'ADN par recombinaison homologue, mais ces protéines interagissent avec de nombreuses autres protéines impliquées dans diverses fonctions telles que la progression du cycle cellulaire, la régulation des gènes de transcription et l'ubiquitinylation (33). Cela fait d'eux de véritables gardiens de l'intégrité du génome.

Les mutations pathogènes des gènes *BRCA1* et *BRCA2* prédisposent principalement au risque de cancers du sein et de l'ovaire. Le risque cumulé de cancer du sein à l'âge de 80 ans est de 72 % pour les patientes ayant une mutation du gène *BRCA1* et de 69 % pour les patientes ayant une mutation du gène *BRCA2*. L'incidence des cancers du sein chez les porteuses de mutations *BRCA1* et *BRCA2* sont majeures entre 30 et 60 ans. Concernant le cancer de l'ovaire, le risque cumulé à 80 ans a été estimé à 44 % pour les porteuses de mutations du gène *BRCA1* et 17 % pour les patientes mutées *BRCA2* (34). Dans une moindre mesure, le gène *BRCA2* principalement et le gène *BRCA1* induisent aussi un risque beaucoup plus faible néanmoins de cancer de la prostate (35) et du pancréas (36) (37). Aussi, un porteur homozygote ou hétérozygote composite du gène *BRCA2* nommé aussi *FANCD1* (« *Fanconi anemia complementation group D1* ») présentera un syndrome de Fanconi sévère caractérisé notamment par des signes malformatifs ainsi qu'une prédisposition aux hémopathies et à certaines tumeurs solides (38).

3. PALB2

Le gène *PALB2* (« *partner and localizer of BRCA2* ») ou *FANCN* (« *Fanconi anemia complementation group N* ») code pour une protéine ayant un domaine de liaison à la protéine BRCA2 et permet notamment sa stabilisation intranucléaire et son accumulation en cas d'erreur de dommage à l'ADN (39). L'occurrence des mutations du gène *PALB2* varie de 0,1 % à 1,5 % selon les populations. Plusieurs études ont estimé que le risque de cancer du sein conféré par les mutations du gène *PALB2* à l'état hétérozygote sont voisines de celles dues aux mutations des gènes *BRCA1* et *BRCA2* (40) (41) (42). En revanche, ce gène n'a pas montré de surrisque significatif de prédisposition au cancer de l'ovaire (40) (43).

4. RAD51C ET D

La protéine RAD51 possède un rôle majeur dans la réparation de l'ADN par recombinaison homologue. Les gènes *RAD51C* (ou *FANCO*) et *RAD51D* codent pour des paralogues de RAD51 qui sont des protéines médiatrices de RAD51 qui favorisent son activité. Les mutations constitutionnelles de ces gènes confèrent un risque augmenté de cancer de l'ovaire du même ordre que celui conféré par les mutations du gène *BRCA2* avec un risque relatif (RR) estimé à 5,88 (IC = 2,91-11,88 $p=7,65 \times 10^{-7}$) pour le gène *RAD51C* (44) et un odds ratio (OR) à 6,3 (IC = 2,89-13,85 $p=4,8 \times 10^{-6}$) pour le gène *RAD51D* (45). En revanche, la surveillance mammaire n'est pas recommandée à l'heure actuelle pour ces deux gènes même si le gène *RAD51C* est associé à un surrisque de cancer du sein (46).

5. CDH1

Le gène de la *cadhérine 1* (*CDH1*) (NM_004360.3) code pour la E-cadhérine, une protéine d'adhésion cellulaire qui agit comme suppresseur de tumeur en inhibant l'invasion cellulaire (47). Les mutations germinales de *CDH1* sont à l'origine du cancer gastrique diffus héréditaire et du syndrome du cancer du sein lobulaire (CSL) (OMIM 137215). Pour les porteurs d'altérations pathogènes, le risque estimé au cours de la vie de développer un cancer gastrique diffus atteint 70% pour les hommes et 56% pour les femmes. Les femmes ont également un risque de CSL à vie d'environ 42 % (48).

6. PTEN

Le gène *PTEN* (« *Phosphatase and TENsin homolog* ») agit comme suppresseur de tumeur en régulant négativement la voie PI3K(phosphoinositide-3-kinase)-AKT - impliquée

notamment dans la prolifération, la survie et la croissance cellulaire. Les mutations constitutionnelles du gène *PTEN* (« *Phosphatase and Tensin homolog* ») à l'état hétérozygote sont responsables d'un syndrome appelé maladie de Cowden (49). Les critères majeurs d'indication à la recherche d'une mutation constitutionnelle du gène *PTEN* selon les « National Comprehensive Cancer Network » (NCCN) guidelines version 1.2014 sont les suivants :

- Les lésions cutanéomuqueuses multiples telles que des trichilemmomes par exemple c'est-à-dire des proliférations des follicules pileux
- La macrocéphalie
- Les hamartomes multiples gastro-intestinaux
- La maladie de Lhermitte Duclos provoquée par une lésion hamartomateuse de la fosse postérieure
- La survenue d'un cancer du sein, de l'endomètre et/ou de la thyroïde

L'inclusion de ce gène dans le panel HBOC est justifiée parce qu'il induit un risque très élevé de cancer du sein avec un risque cumulé à 70 ans supérieur à 75 % (50) (51) même si l'inclusion des cas index dans les études réalisées tend à surestimer ce risque. De plus, ce syndrome augmente le risque de cancer de l'endomètre et de la thyroïde. Dans une moindre mesure, une augmentation du risque de cancer du rein et du colon a aussi été rapportée pour les porteurs de ce syndrome (50).

7. Les gènes MMR

Les gènes MMR (« Mismatch repair ») soit les gènes *MLH1*, *MSH2*, *MSH6* et *PMS2* sont impliqués dans la réparation des mésappariements de l'ADN. Les mutations constitutionnelles à l'état hétérozygotes de ces gènes sont responsables du syndrome de Lynch qui est la cause de 2-4 % des cancers colorectaux. Ce syndrome prédispose à différents cancers dont le cancer colorectal sans contexte de polypose, le cancer de l'endomètre, le cancer urothélial et le cancer de l'intestin grêle. Il prédispose aussi à d'autres cancers si l'on considère le spectre tumoral élargi comme le cancer gastrique, le cholangiocarcinome, le cancer de l'ovaire, le glioblastome et le carcinome sébacé. Le syndrome de Lynch est souvent évoqué chez un patient lors du diagnostic de cancer colorectal lorsque le contexte clinique est évocateur et que ce dernier répond notamment aux critères d'Amsterdam et de Bethesda basés en grande partie sur les antécédents personnels et familiaux de cancers du spectre.

Les gènes MMR ont été inclus dans le panel car des études ont montré que les porteurs de mutation hétérozygote d'un de ces gènes avaient un risque cumulé de cancer de l'ovaire à 80 ans de 8 % à 15 % selon le gène responsable (52). Le gène conférant le plus de risque est le gène *MSH2*. L'âge moyen au diagnostic de ces cancers est de 45 ans. Par ailleurs, la prévalence des tumeurs à composante endométrioïde et mucineuse est importante dans ce syndrome (16).

Concernant le cancer du sein, les femmes porteuses de mutations d'un des gènes MMR ont un risque proche de celui de la population générale. Cependant, ces évaluations du risque de cancer ont été réalisées dans le cadre d'une recherche de syndrome de Lynch et non dans les familles de cancers du sein. Des données sur les familles de cancer du sein vont permettre à l'avenir d'affiner l'estimation de ces risques.

8. *TP53*

Le gène *TP53* est un facteur de transcription d'importance majeure impliqué dans la régulation de nombreux gènes ayant un rôle dans le cycle cellulaire, la réparation de l'ADN et l'apoptose. Les mutations constitutionnelles du gène *TP53* à l'état hétérozygote sont responsables du syndrome de Li Fraumeni (LFS) dominé par l'apparition de tumeurs précoces du jeune adulte et de l'enfant. Les tumeurs du spectre sont les cancers du sein pré-ménopausique, sarcome des tissus mous, ostéosarcome, tumeur du système nerveux central, corticosurréalome.

La protéine TP53 agit en formant un tétramère. Dans ce syndrome, les mutations décrites sont en majorité des mutations ponctuelles à type de mutations faux sens (66 %) dont certaines sont dominantes négatives. Ces mutations dominantes négatives sont responsables d'une protéine non fonctionnelle du fait de la formation d'un tétramère non fonctionnel responsable d'un phénotype plus sévère et plus précoce chez les individus porteurs. Au contraire, les mutations perte de fonction telles que les mutations « frame shift » ou non-sens entraînent souvent un phénotype moins sévère (53).

Le spectre tumoral du syndrome de LFS est détaillé dans les critères de Chompret version 2015 (53). Il s'agit des critères d'indication d'analyse du gène *TP53*. Ces critères sont les suivants :

(1) **La présentation familiale**

Probant avec une tumeur du spectre du LFS ci-décrits avant 46 ans et au moins un apparenté au premier ou au deuxième degré ayant développé une tumeur du spectre LFS

(sauf le cancer du sein si le probant a aussi développé un cancer du sein) avant 56 ans ou avec des tumeurs multiples

(2) Des tumeurs multiples primitives

Probant avec des tumeurs multiples du spectre LFS (sauf les cancers du sein multiples) dont l'un est diagnostiqué avant 46 ans.

(3) Des tumeurs rares

Patient avec un corticosurrénaome, une tumeur du plexus choroïde, un rhabdomyosarcome de sous-type embryonnaire anaplasique, indépendamment de l'histoire familiale.

(4) Le cancer du sein précoce

Un cancer du sein avant 31 ans.

Les enfants atteints de ce syndrome développent principalement des ostéosarcomes, des corticosurrénaomes, des tumeurs du système nerveux central et des sarcomes des tissus mous. En revanche, chez l'adulte se développe en grande majorité des cancers du sein et des sarcomes des tissus mous. Concernant l'histologie des cancers du sein développés dans ce syndrome, 55 % sont de type HER2+ et 37 % sont des tumeurs triple positifs soit RH+ (récepteur hormonal +) HER2+ (53). Il est important de connaître le statut *TP53* des patientes porteuses car cela peut modifier la prise en charge de ces dernières. En effet, les patients porteurs d'un variant pathogène de *TP53* développent très souvent des tumeurs secondaires en cas de traitement par radiothérapie, au niveau du champs d'irradiation (53). De même, la plupart des chimiothérapies sauf les poisons du fuseau ont entraîné une génotoxicité sur des modèles de souris LFS (54). La chirurgie est donc à privilégier chez eux en cas de développement d'un cancer.

Il s'agit d'un syndrome à pénétrance élevée mais incomplète et le risque actuel de développer un cancer dans le cadre d'une mutation *TP53* est probablement surestimé du fait d'un biais de sélection dû au fait que les tests génétiques sont réalisés dans des familles au phénotype évocateur. En effet, une étude réalisée chez 63 983 contrôles a montré que la prévalence des mutations pathogènes et probablement pathogènes de *TP53* dans la population générale est de 0,2 % (55). Il est important de tenir compte du contexte familial lorsqu'un

variant pathogène de *TP53* a été mis en évidence dans le cadre d'un test en panel de plusieurs gènes.

IV. Détection des individus à risque génétique et prise en charge spécifique

Pour l'ensemble des 13 gènes inclus dans le panel GGC-unicancer, lorsque le variant causal est identifié chez le cas index, un test prédictif est possible chez les apparentés afin de préconiser une surveillance adaptée selon le gène concerné. Ainsi différents référentiels recommandent un suivi spécifique et parfois une chirurgie prophylactique afin de réduire considérablement le risque d'apparition de cancer.

Par exemple, en ce qui concerne les gènes *BRCA1* et 2, les référentiels de l'HAS 2014 et de l'INCa 2017 recommandent tout d'abord la réalisation d'un suivi gynécologique annuel à partir de 20 ans. Ils préconisent aussi la réalisation d'une IRM et d'une mammographie mammaire annuelle de 30 à 65 ans (à partir de 65 ans, l'IRM est remplacée par la mammographie seule). En outre, une chirurgie mammaire prophylactique de réduction du risque peut être proposée après validation en réunion de concertation pluridisciplinaire (RCP). Concernant le risque de cancer de l'ovaire, aucun suivi n'a montré d'efficacité. Ainsi, seule l'annexectomie prophylactique est fortement encouragée à partir de 40 ans mais à adapter selon l'histoire familiale et le gène concerné, après validation en RCP (56). Concernant le gène *PALB2*, seul le suivi mammaire est préconisé et identique aux gènes *BRCA1* et *BRCA2* tandis que pour les gènes *RAD51C* et *RAD51D* seule l'annexectomie prophylactique est préconisée.

Pour le gène *TP53*, il n'a pas été mis en évidence de surrisque de cancer de l'ovaire. En revanche, les femmes porteuses d'une mutation *TP53* ont un risque élevé de cancer du sein précoce et l'IRM mammaire est préconisée dès 20 ans avec une échographie. La mammographie est en effet, à éviter dans cette prédisposition du fait de la radiosensibilité augmentée. La chirurgie mammaire de réduction du risque peut être proposée. Concernant les autres risques de cancer, le suivi doit se faire selon le référentiel Li Fraumeni (56) qui comporte un suivi lourd avec notamment la réalisation d'une IRM annuelle du corps entier. Cependant, la découverte incidente de mutations *TP53* dans des familles dont l'histoire est non évocatrice amène à des discussions sur la surveillance qui pourrait évoluer dans les prochaines années.

En ce qui concerne les autres gènes du panel, en plus de la surveillance gynécologique recommandée, une surveillance spécifique est préconisée selon le gène atteint par des

référentiels variés (selon le référentiel du Lynch pour les gènes MMR, selon le référentiel de la maladie de Cowden pour *PTEN*...)

V. Diagnostic moléculaire chez les cas index : Le séquençage à haut débit

1. Historique des techniques antérieures

Jusqu'à très récemment, seuls les gènes *BRCA1* et *BRCA2* étaient analysés pour le diagnostic des prédispositions aux cancers du sein et de l'ovaire. La recherche de mutations était principalement effectuée par séquençage Sanger, technique inventée par Frédérick Sanger en 1977 en Angleterre. Cependant, l'inconvénient de cette technique est qu'elle nécessite beaucoup de temps et de réactifs pour séquencer ces deux gènes. En effet, cette technique de séquençage par synthèse de nucléotide permet de séquencer des fragments de 400 à 900 paires de bases. Or, la région codante du gène *BRCA1* et *BRCA2* comporte respectivement 7,2 kb et 10,98 kb, ce qui représente de nombreux fragments à séquencer. Les plus performants des séquenceurs Sanger comme le 3130xL, techniques NGS (« Next Generation Sequencing ») de première génération sont considérés comme techniques de haut débit car ils sont capables de paralléliser 96 électrophorèses en utilisant 96 capillaires. Ils permettent ainsi l'analyse de 96 x 800 nucléotides lus en 1 heure. L'offre diagnostic était néanmoins très limitée en nombre de cas et de gènes séquencés et les délais de rendu des résultats longs d'autant que la détection des CNV était souvent réalisée en parallèle par des techniques dédiées, MPLA (« multiplex ligation-dependent probe amplification ») ou en QMPSF (« quantitative multiplex PCR of short fluorescent fragments »). L'utilisation de méthode de pré-criblage (DHPLC ("Denaturing High Performance Chromatography ») (57), HRM (« High Resolution Melt ») (58), EMMA (« Enhanced Mismatch Mutation Analysis ») (59) par exemple), dont certaines pouvaient détecter simultanément les mutations de petites tailles et de grandes tailles, ont permis d'améliorer les conditions de réalisation de ces analyses mais sans commune mesure avec le séquençage à haut débit de deuxième génération.

2. Le principe du séquençage à haut débit : exemple de la technique Illumina

Le développement des techniques de séquençage à haut débit de deuxième génération depuis à peine une quinzaine d'années a révolutionné la génétique. Le séquençage à haut débit aussi appelé NGS permet le séquençage de plusieurs centaines de milliers de bases par l'utilisation d'approche massivement parallèle. Cette nouvelle technologie a permis d'augmenter considérablement le débit d'analyse tout en réduisant les coûts. Les séquenceurs

NGS de deuxième génération sont capables de lire des fragments de 150 à 300 paires de bases (pb) mais jusqu'à plusieurs milliards de fragments à la fois. Ainsi, ces nouvelles technologies sont capables de générer l'analyse de plusieurs génomes complets de 3,2 milliards de paires de bases en moins de 48h.

Plusieurs techniques de séquençage à haut débit de deuxième génération existent mais celle utilisée pour l'étude est la technologie Illumina qui est une technique de séquençage par synthèse et qui sera donc la seule détaillée ci-après.

La première étape du séquençage consiste à préparer la librairie, c'est-à-dire les fragments d'ADN à séquencer (**Figure 2**). L'ADN de chaque patient analysé est fragmenté soit par sonication ou soit par l'utilisation d'enzymes de restriction puis sélectionné de manière à obtenir des fragments d'une taille d'environ 200 pb. Ces fragments vont subir une étape de réparation appelée *End-repair* d'adénylation afin d'éviter la ligation de ces derniers entre eux. Ensuite, des adaptateurs vont être ajoutés aux extrémités des fragments. Ces adaptateurs sont des fragments d'ADN simple brin qui vont permettre plus tard l'étape de séquençage par hybridation de l'amorce. Dans le cadre d'un re-séquençage ciblé comme proposé pour le diagnostic moléculaire des maladies génétiques aujourd'hui, l'étape d'après consiste à enrichir les régions d'intérêt soit par amplification soit par capture. La technique par capture nécessite l'utilisation de sondes nucléotidiques biotinylées d'une centaine de paires de bases complémentaires des régions d'intérêts. Ces dernières vont se lier aux fragments d'ADN ciblés et vont pouvoir venir se fixer à des billes magnétiques qui seront elles même récupérées par des aimants. Après avoir récupéré et purifié les fragments capturés, des index nucléotides spécifiques (ou barre code moléculaire) à chaque patient vont être ajoutés aux fragments afin de pouvoir réunir (multiplexage) l'ensemble des échantillons à séquencer sur la même lame de séquençage (dites « Flow Cell »), qui seront donc séquencés simultanément (**Figure 2**).

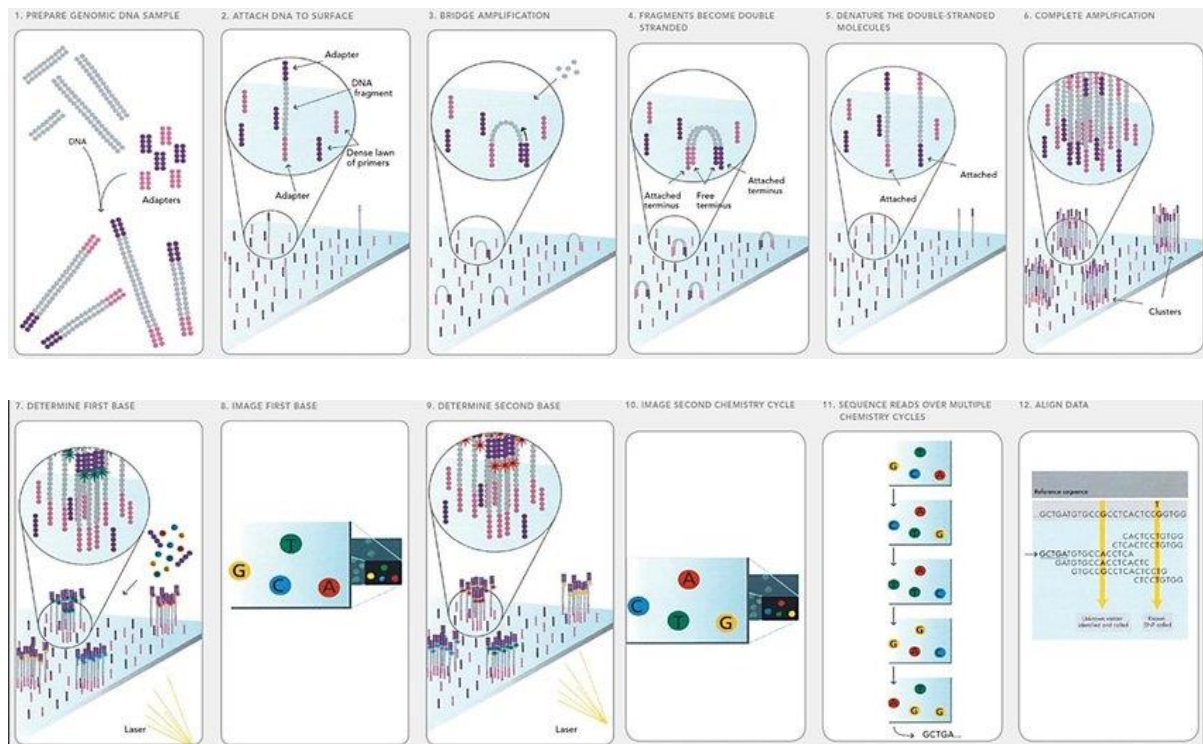


Figure 2 : Représentation du séquençage par la technique Illumina (image issue de www.Illumina.com)

- 1- Préparation des bibliothèques avec ligation des adaptateurs
- 2- Les ADN se lient à la Flow Cell grâce aux adaptateurs
- 3- Amplification en ponts
- 4- Présence de fragments doubles brins
- 5- Dénaturation des ADN double brin
- 6- Amplification complète
- 7- Première base de chacun des clusters séquencés
- 8- Image lue par la caméra haute résolution : chaque point représente théoriquement un même fragment d'ADN amplifié. La couleur de chaque point correspond à la première base séquencée
- 9- Séquençage de la deuxième base de l'ensemble des clusters
- 10- Image lue par la caméra haute résolution : chaque point représente théoriquement un même fragment d'ADN amplifié. La couleur correspond à la deuxième base séquencée
- 11- Le séquençage continue de la même façon sur plusieurs cycles jusqu'au séquençage complet des fragments d'ADN
- 12- Les données brutes obtenues sont converties en données analysables bio informatiquement pour permettre l'alignement des *reads*, soit les fragments séquencés, à partir d'une séquence de référence.

L'ADN va ensuite être dénaturé afin que les brins d'ADN simple brin se fixent à la *flow cell*, une plaque de verre sur laquelle sont fixées de courtes séquences d'ADN complémentaires des adaptateurs. Ensuite, une amplification clonale par « Bridge PCR » (« Polymerase Chain reaction ») va être réalisée à la surface de la *flow cell* permettant la formation de nombreux

clusters, chaque cluster correspondant à l'amplification d'un fragment d'ADN unique au départ. Une fois les clusters réalisés, l'étape de séquençage à proprement parler va pouvoir commencer. Cette étape consiste en l'incorporation des quatre nucléotides terminateurs réversibles marqués par des fluorophores (différent pour les quatre nucléotides). Après excitation par un laser, le signal émis par chaque cluster monoclonal est lu par une caméra haute-résolution correspondant à la première base incorporée. L'identification des bases ou *base calling* est la sélection de la base émettant la plus forte intensité au sein de chaque cluster. Par exemple, si le signal d'un cluster émet principalement un signal lumineux correspondant à la base A, cette dernière est appelée. Puis, un cycle suivant se déroule en incorporant à nouveau les quatre terminateurs réversibles marqués. La deuxième base est lue. Puis, les cycles suivants se déroulent de la même manière jusqu'à la lecture complète de l'ensemble des fragments à séquencer. Une lecture du faisceau lumineux est effectuée simultanément à chaque séquence pour l'ensemble des clusters. L'intérêt des clusters est simplement d'augmenter l'intensité du signal lumineux et de permettre sa détection par la caméra.

3. Le traitement des données par analyse bioinformatique

La bio-informatique joue un rôle central dans l'analyse des données générées par le séquençage haut débit. Le NGS a fait entrer la bio-informatique comme nouvelle compétence indispensable au sein des laboratoires de génétique moléculaire. La capacité à détecter des variations génétiques de différents types est possible le plus souvent en combinant plusieurs programmes, libres ou commerciaux et dont chacun est capable de détecter un type d'événement mutationnel spécifique parmi les suivants :

- les CNV sont des délétions ou des duplications de grands fragments de séquences qui peuvent concerner par exemple un gène complet ou seulement certains exons composant le gène.
- les « single nucleotide variations » (SNV) qui sont des variations qui touchent un seul nucléotide tel que le remplacement d'une base par une autre base, une délétion ou une duplication d'une base. Certains de ces algorithmes permettent de détecter dans des milieux hétérogènes des variants à faible ratio allélique (VAF pour « variant allele fraction »)
- Les indels qui désignent des délétions ou des insertions de plusieurs bases dans une séquence biologique. Cela peut parfois être une délétion avec une insertion d'une nouvelle séquence.

- Les réarrangements complexes structuraux, tels que des translocations, inversions mais aussi la détection d'éléments mobiles si le génome complet est séquencé.

L'analyse bioinformatique se déroule ainsi en plusieurs étapes :

- Le basecalling : il s'agit de l'étape traduisant les signaux émis par le séquenceur tels que les signaux de fluorescence par exemple pour le séquenceur Illumina en séquences (appelées « *Reads* ») de quelques dizaines à centaines de bases. Cette étape fournit aussi une évaluation de la qualité sous la forme de « Phred quality scores » qui sont des scores de vraisemblances permettant de définir la probabilité que la base lue est exacte. On estime que les séquenceurs font 1% d'erreur. Mais grâce à l'intégration de ce score et aux traitements des données en aval, la très grande majorité des erreurs de séquençage sont filtrées.

- Le démultiplexage : les *reads* de chacun des patients qui avaient été identifiés par des codes-barres lors de la préparation des librairies sont séparés. Chaque *read* porte en effet une courte séquence qui a été insérée lors de la préparation des librairies et qui est spécifique d'un patient.

- L'alignement : il s'agit de l'étape d'alignement des *reads* sur une séquence de référence telle que l'hg19 (**Figure 3**). Cet alignement est réalisé grâce à des algorithmes d'alignement qui positionnent chaque « *read* » sur une position génomique selon des analyses probabilistes. Il existe plusieurs aligneurs et certains sont plus adaptés aux SNV, d'autres aux CNV et d'autres aux indels. Le choix de l'aligneur est donc primordial. On peut citer l'algorithme de Burrows Wheeler (60). On obtient alors un fichier BAM (« Binary Alignment Map ») associant à chaque *read* ses coordonnées génomiques.

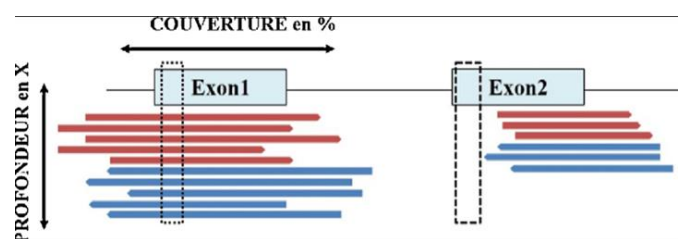


Figure 3 : Représentation de la profondeur et couverture en séquençage haut débit (Image issue de Lacoste et al., 2017(55))

Chaque *read* est représenté en rouge ou en bleu selon le sens de la séquence. Les *reads* sont alignés sur une séquence de référence. La profondeur est le nombre de fois qu'une base est lue en X. Par exemple, l'exon 1 a une couverture de 100% car l'ensemble de l'exon est lu tandis que l'exon 2 à une couverture de 80% environ. La couverture représente une zone lue par au moins 1 *read* en %. La zone encadrée dans l'exon 1 a une profondeur de 10X qui correspond simplement à 10 *reads*.

- L'appel des variants : c'est l'étape de mise en évidence des variants par les « variants caller » qui identifient les divergences observées par rapport à la séquence de référence et les nomment. Les variants caller ont aussi leur spécificité propre dans les événements mutationnels détectés. Pour l'étude ici, nous avons utilisé le logiciel Outlyzer dont la particularité est de détecter des événements de faibles fréquences alléliques avec une bonne spécificité (**Figure 4**). La VAF, ratio du nombre de *reads* portant l'allèle muté sur le nombre de *reads* portant l'allèle sauvage, est proche de 50% pour un variant constitutionnel, présent à l'état hétérozygote. Outlyzer avait été développé pour l'étude des tumeurs qui nécessite de détecter des variants pathogènes, parfois peu représentés due en particulier à l'hétérogénéité tumorale (61), avec une VAF potentiellement très faible tout en évitant les faux positifs.

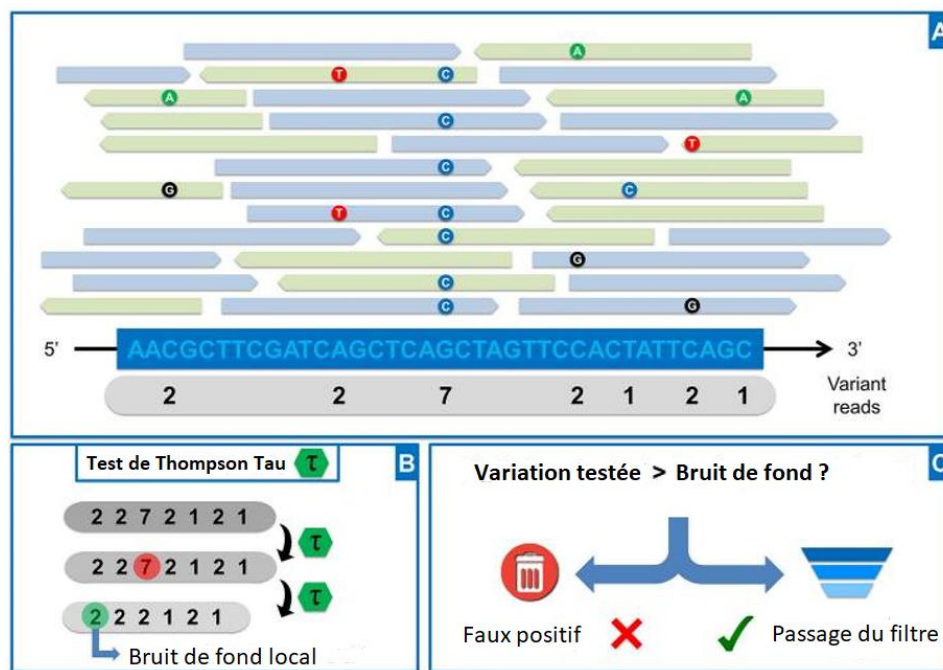


Figure 4 : Le principe de l'analyse par Outlyzer (Image traduite de Muller et al., 2016 (62))

- (A) Les *reads* sont alignés le long d'une séquence de référence. Chaque base qui diffère de la séquence de référence est représenté par un point de couleur. A chaque position génomique, le nombre de variants trouvé est calculé (représenté dans la bannière grise sur le schéma)
- (B) Un test de Thompson Tau est appliqué à la liste des variants trouvés pour permettre le calcul du bruit de fond local
- (C) Le nombre de *reads* portant le variant analysé est comparé au bruit de fond local afin de déterminer si le variant est un faux positif ou non. Si le variant est supérieur au bruit de fond, il sera soumis aux filtres habituels tels que la balance des *reads Forward* et *Reverse*, la moyenne et la déviation standard du score Phred (62).

- L'annotation des variants : elle peut en partie être prise en charge par des annotateurs automatiques de variants comme ANNOVAR(63). Ce logiciel annote chaque variant en donnant sa position génomique, le gène concerné, des scores prédits, la nomenclature des transcrits en format HGVS (« Human Genome Variation Society »)(64). Elle donne aussi d'autres informations utiles à l'interprétation des variants telles que les prédictions des effets du variant sur l'épissage et sur la protéine ainsi que les fréquences du variant dans des grandes bases de données de populations, comme celle maintenue par le Broad Institut, gnomAD (« Genome Aggregation Database »), rendues possible grâce au NGS.

4. Les défis de l'interprétation des variants par NGS

Avec le NGS, des nombreux variants sont détectés chez les patients et la difficulté actuelle est de les interpréter. Un guide de recommandation consensuel publié par l'ACMG (« American College of Medical Genetics and Genomics ») en 2015 est utilisé au niveau international pour l'interprétation et la classification des variants (65) En pratique, le guide de l'ACMG permet une traçabilité des décisions et doit être adapté à chaque pathologie et se décline parfois par des versions adaptées par pays ou syndrome. Ce guide tend à standardiser l'interprétation des variants et se base sur un modèle de catégorisation permettant de stratifier les variants en 5 classes. L'introduction de cette classification en 5 classes a premièrement été proposée par S.E. Plon en 2008 et rapidement utilisée dans une approche de classification basée sur un calcul de probabilité de pathogénicité dans un modèle multifactoriel (66).

- Classe 1 : bénin
- Classe 2 : probablement bénin
- Classe 3 : variant de signification inconnue
- Classe 4 : variant probablement délétère
- Classe 5 : variant délétère

Les principaux outils permettant cette classification sont :

- Les bases de données : Les bases de données en population générale telles que ExAC (« Exome Agregation Consortium ») ou dbSNP (« Single Nucleotide Polymorphism database ») (67) et maintenant gnomAD permettent de connaître la fréquence du variant dans la population générale. Un variant de fréquence > 1% est un polymorphisme. De

manière générale, plus un variant est fréquent, plus il a de chance d'être bénin. Il existe aussi des bases de données de variants recensant les différentes interprétations et pour certaines les phénotypes des patients correspondants. Elles ne sont pas forcément toutes de la même qualité et certaines ne sont pas ouvertes à l'ensemble de la communauté. Par exemple, pour le cancer du sein, il existe UMD (« Universal Mutation Database ») (68), BRCAshare (69), BRCA exchange Global alliance (70) et plus récemment la nouvelle initiative française de base multigénique de partage de connaissances promue pour les laboratoires d'oncogénétique sous l'égide d'UNICANCER et de l'INCa : FrOG (« the French OncoGenetic database »).

- Les logiciels de prédictions bioinformatiques permettant de prédire les effets délétères ou non des variants sur la protéines (Align GVGD (66), SIFT (71), Polyphen, CADD par exemple) en tenant compte de la conservation du nucléotide entre les différentes espèces, l'écart physico-chimique entre l'ancien et le nouvel acide aminé et la localisation de la variation sur un site fonctionnel ou non de la protéine.
- Le logiciel de prédiction de l'effet du variant sur l'épissage tels que SSF (72), MES (73), SpiCE (74)
- Les études de coségrégation du variant avec la pathologie
- Les diverses études fonctionnelles.

Ainsi, ces divers arguments alimentent la classification du variant. Seuls les variants de classe 4 et 5 sont utilisables au titre du conseil génétique aujourd'hui c'est-à-dire qu'ils entraînent la réalisation de tests pré-symptomatiques chez les apparentés. Les apparentés porteurs du variant bénéficient d'un suivi rapproché tandis que pour les non-porteurs du variant, le suivi recommandé est celui de la population générale.

5. La détection et l'interprétation des variants présentant des VAF faible

L'analyse HBOC réalisée dans un cadre diagnostique au laboratoire est validée selon la norme ISO15189. Elle comporte un panel de 12 gènes recommandés par le GGC dans ce cadre. Le panel effectué au laboratoire du CFB est en fait composé de 34 gènes mais seuls les 12 gènes rendus au titre du diagnostic font l'objet d'une interprétation par les biologistes à chaque série de patients. Depuis, la mise en place de ce panel de 12 gènes il y a un peu plus d'un an et notamment l'inclusion du gène *TP53*, les biologistes sont de plus en plus confrontés à la détection de variants de faible fréquence allélique difficiles à interpréter. En effet, ces

variants ne sont pas présents par définition à l'état hétérozygote chez l'individu. Ils peuvent tout de même être constitutionnels mais apparus durant les étapes post zygotiques au tout début de l'embryogénèse ou être le reflet de mutations somatiques tumorales ou pré-cancéreuses détectées dans le sang des patients. Ces diverses causes n'ont pas les mêmes conséquences pour la prise en charge des patients et de leurs familles d'où la nécessité de connaître leur signification. C'est un des challenges d'aujourd'hui auquel notre étude veut contribuer à répondre.

VI. Les néomutations

1. Rappels de principes généraux : mutations héritées / mutations *de novo* / mosaïques

Une mutation constitutionnelle est une mutation présente avant la fécondation ou apparaissant lors des premières divisions du zygote. Une mutation constitutionnelle sera présente dans toutes les cellules somatiques de l'individu et également dans ses cellules germinales si elle est héritée ou apparue avant la première division cellulaire. Si elle apparaît après la première division cellulaire alors l'ensemble des cellules de l'individu ne sera pas atteint. On parle dans ce cas de mosaïque constitutionnelle. Toute mutation nouvellement apparue, en mosaïque ou non, est aussi appelée mutation « *de novo* » ou « néomutation ». Les études d'exomes en trio par séquençage haut débit ont permis d'estimer la proportion de néomutations chez l'homme à 1,58 par exome (75).

2. Les néomutations constitutionnelles en pathologie humaine

En biologie, le terme mosaïque désigne l'état d'un organisme dans lequel cohabitent plusieurs populations de cellules de génotype différent. Ainsi, un être vivant comportant un phénotype en mosaïque possède des différences génotypiques entre certains de ses tissus qui ne répondent pas aux lois de l'hérédité mendélienne.

La détection d'un variant en mosaïque chez un individu implique l'émergence d'une mutation dite « *de novo* », ayant eu lieu précocement en postzygotique lors de l'embryogénèse. Plus la néomutation est précoce dans le développement, plus la proportion de cellules atteinte par la néomutation sera importante. A l'origine, les variants constitutionnels d'origine postzygotique ont montré leur implication dans plusieurs maladies monogéniques d'expressivité variable telles que la neurofibromatose de type 1 et 2 par exemple (76) (77). La néomutation peut toucher les cellules de la lignée germinale, les cellules de la lignée somatique

ou les deux (78) (79). Il est nécessaire de connaître cette distinction car seul un individu porteur d'un variant atteignant les cellules germinales aura un risque de transmettre la néomutation à sa descendance. Cependant, en pratique, il est impossible à l'heure actuelle de s'assurer de manière certaine de l'absence de variant en mosaïque germinale en particulier chez la femme dont les gamètes sont difficilement accessibles pour des tests génétiques (78). Néanmoins, une néomutation postzygotique atteignant les cellules germinales peut être diagnostiquée chez un parent asymptomatique dont plusieurs enfants présentent une même maladie génétique avec un mode de transmission dominante, en particulier s'il s'agit d'une maladie à pénétrance complète.

En cancérologie, on parle aussi de mosaïque tumorale pour désigner l'hétérogénéité génotypique au sein d'une tumeur. En effet, Theodor Boveri en 1929 proposa un modèle selon lequel le cancer est originaire d'une modification génétique acquise au sein d'une cellule. On sait actuellement que le cancer est dû à une accumulation de mutations au cours du temps conférant un avantage sélectif à la cellule par une amélioration de ses capacités de division cellulaire, de résistance à l'apoptose, de migration... Par ailleurs, les tumeurs sont composées de plusieurs sous-populations de cellules qui ont acquis au fil du temps des néomutations indépendamment les unes des autres expliquant ainsi l'hétérogénéité tumorale. Cette hétérogénéité tumorale explique d'ailleurs en partie la réponse incomplète des patients aux divers traitements thérapeutiques, certaines cellules tumorales ayant acquises des mutations leur conférant une résistance au traitement.

Certains gènes impliqués dans les prédispositions aux cancers sont depuis longtemps décrits avec des taux rapportés important de mutation *de novo* et de mosaïques. C'est le cas de la neurofibromatose segmentaire de type 1 décrite pour la première fois en 1936 (80) qui est due à une mutation en mosaïque du gène *NF1* (« *Neurofibromin 1* »). La neurofibromatose, en plus de prédisposer aux neurofibromes, se caractérise par la présence de lésions cutanées à type de taches café au lait. Dans les formes en mosaïque, les signes cutanés ne sont pas répartis sur l'ensemble du corps mais localisés (81). Dans la neurofibromatose de type 2, qui est un syndrome causé par une mutation perte de fonction du gène *NF2* (« *Neurofibromin 2* ») et prédisposant aux schwannomes qui sont des tumeurs nerveuses bénignes, 50 % des mutations apparaissent *de novo* (82) et parmi elles selon une étude récente, 22 % sont à l'état de mosaïque (83). Enfin, concernant le gène *RBI* (« *retinoblastoma 1* »), dont les mutations perte de fonction à l'état hétérozygote prédisposent chez l'enfant jeune au rétinoblastome, une tumeur maligne de la rétine, de nombreux cas de mutations *de novo* ont été rapportés. La fraction d'évènements postzygotiques parmi les cas sporadiques de rétinoblastome a été évalué à 15 %

(84). Le point commun de ces gènes est leur pénétrance importante, sauf pour le gène *RBI* pour lequel quelques rares mutations sont moins pénétrantes. De plus, ces trois syndromes de prédispositions se manifestent par des pathologies spécifiques et facilement diagnosticables sans risque de phénocopie. Ainsi, les mosaïques sont parfois diagnostiquées avant même le séquençage du gène comme le cas de la neurofibromatose segmentaire de type 1.

3. Néomutations dans les gènes impliqués dans le syndrome de prédisposition au cancer du sein et de l'ovaire.

a. Mutations *de novo* et néomutations postzygotiques de *BRCA1* et *BRCA2*

Depuis la découverte des gènes *BRCA1* et *BRCA2* dans le milieu des années 90, des milliers de mutations ont été rapportées pour ces gènes, mais la grande majorité des cas concerne des mutations héritées. Seuls des cas exceptionnels de mutations *de novo* ont été décrits. En effet, 9 cas de mutations *de novo* du gène *BRCA1* ont été rapportés dans la littérature et 6 cas de mutations *de novo* du gène *BRCA2* (85).

Lisa Golmard et al. ont déterminé les taux de mutations *de novo* des gènes *BRCA1* et *BRCA2* sur 12805 patientes non apparentées ayant présenté un cancer du sein et/ou de l'ovaire. Ainsi, sur les 1527 patientes porteuses d'une mutation *BRCA1* ou *BRCA2* (12 %), seules 3 patientes présentaient une mutation *de novo* du gène *BRCA1* dont un cas en mosaïque et une patiente présentait une mutation *de novo* du gène *BRCA2* soit des taux de mutation *de novo* calculés de 0,4 % (0,1 % ; 1,1 %) pour le gène *BRCA1* et de 0,1% (0,02 % ; 0,8 %) pour le gène *BRCA2* (85). Ce taux était probablement légèrement sous-estimé car l'ensemble des parents n'avaient pas pu être séquencés. Cependant, dans la plupart des cas, un apparenté porteur de la mutation permettait de confirmer son caractère hérité. De manière intéressante, un des facteurs majeurs pouvant expliquer un taux élevé de mutation *de novo* pour un gène est dû à l'effet péjoratif que la mutation délétère de ce gène entraîne sur la santé du sujet (86) et en particulier sur la fertilité. Ainsi, plusieurs études ont montré que les mutations *BRCA1* et *BRCA2* n'entraînaient pas de diminution de la fertilité (87) mais au contraire pourraient accroître cette dernière (88). Concernant les mosaïques constitutionnelles, seuls trois cas de mosaïques *BRCA1* ont été rapportés dans la littérature. Il s'agissait dans tous les cas de femmes ayant présenté un cancer du sein (89) (90) (85). Aucun cas de mutation constitutionnelle en mosaïque du gène *BRCA2* n'a été rapporté.

b. Mutations *de novo* et mosaïques constitutionnelles du gène *TP53*

Les mutations constitutionnelles du gène *TP53* à l'état hétérozygote sont responsables du LFS (OMIM # 151623). Chez l'adulte, la distribution tumorale est caractérisée par la prédominance des carcinomes du sein observés chez environ 80% des femmes (53). Indépendamment de l'histoire familiale, le taux de détection de mutations *TP53* chez les femmes ayant présenté un cancer du sein avant 30 ans est de 6% (91) (92) (53). Par ailleurs, les taux de détection de mutation *TP53* chez les enfants atteints de corticosurréalome ou de carcinome du plexus choroïde ont été estimés à 45% et à 42% respectivement (53) ce qui en fait des tumeurs très spécifiques de ce syndrome.

En 2009, le taux de mutations *de novo* du gène *TP53* avait été estimé à 7% (7/75) et en considérant les cas suspects de mutations *de novo* par l'histoire familiale, ce taux atteignait 20% (93). Mariette Renaux-Petel et al. ont aussi évalué la contribution des mutations *de novo* et des mosaïques dans le LFS. Parmi leur grande cohorte de patients atteints de mutations *TP53*, ils ont observé que 14% des mutations au minimum étaient *de novo* (48/336). De manière intéressante, une analyse de l'ADN sanguin par NGS de 108 patients très évocateurs de LFS sans mutation *TP53* préalablement détecté, effectué à une profondeur supérieure à 500X, leur a permis d'identifier 6 cas de mutations *de novo* en mosaïques dont deux enfants atteints de corticosurréalomes anaplasiques, deux enfants atteints de tumeurs du plexus choroïde, une femme ayant présenté un cancer du sein avant 31 ans et un patient qui a présenté un ostéosarcome à 12 ans et une femme qui a présenté un sarcome du sein à 35 ans (94). En tout, sur les 48 mutations *de novo* détectées, 8 étaient des mutations en mosaïque soit un taux de 17% de mutations postzygotiques parmi les mutations *de novo* observées (8/48). Ce taux est potentiellement sous-estimé car il est tout-à-fait possible que certaines mutations en mosaïque ne soient pas détectées dans le sang des patients.

D'autres mutations constitutionnelles postzygotiques de *TP53* avaient déjà été décrites. Le premier cas est celui d'une enfant ayant développé un corticosurréalome à un an et un ostéosarcome à 5 ans. Chez elle, le variant R282W était détecté à un ratio allélique de 20% dans le sang et était détecté à l'état hétérozygote dans le frottis buccal (95). Par ailleurs, Sam Bejahti et al. (96) ont décrit le cas d'un enfant de 2 ans qui a présenté deux tumeurs mésoenchymateuses des tissus mous et un neuroblastome métastatique en 6 mois. Chez ce dernier, ils ont mis en évidence une mutation *TP53* détectée à 3-20% à partir du prélèvement sanguin en NGS et ils ont surtout observé une perte d'hétérozygotie dans chacune des tumeurs analysées confirmant le caractère constitutionnel de la mutation.

c. Mutations *de novo* et mosaïques constitutionnelles du gène *PTEN*

Jessica Mester et al. (97) ont estimé la fréquence des mutations *de novo* du gène *PTEN*. Ils ont ainsi confirmé de manière certaine la présence de 20 mutations *de novo* sur 187 familles ayant une mutation *PTEN* soit une fréquence de 10,7%. Par ailleurs, si l'on ajoute les cas de mutations *de novo* fortement suspectés mais non prouvés par la réalisation de tests génétiques chez les parents de l'individu atteint, cette valeur s'élève à 19,3% (36/187).

La présence d'une mosaïque a été confirmée expérimentalement sur 10 à 40% des patients avec des mutations *de novo* de *PTEN* (98). Ces patients présentaient des hamartomes cutanés et des lésions internes mixtes cliniquement difficiles à distinguer des maladies héréditaires ou *de novo* (99) (100) (101). Par ailleurs, il a également été rapporté des patients présentant des mutations constitutionnelles de *PTEN* et dont certains tissus ne possèdent aucun allèle fonctionnel de *PTEN*. Ces individus ont ainsi acquis durant les étapes post zygotiques une perte d'hétérozygotie de l'allèle sauvage de *PTEN* entraînant dans ce cas, non pas un cancer, mais des malformations importantes dans les tissus concernés comme par exemple des hypertrophies segmentaires et des lipomes (102). On parle de syndrome de « SOLAMEN » (« Segmental overgrowth, lipomatosis, arteriovenous malformation and epidermal nevus »). Il s'agit d'une mosaïque segmentaire de type 2 représenté dans la **figure 5**.

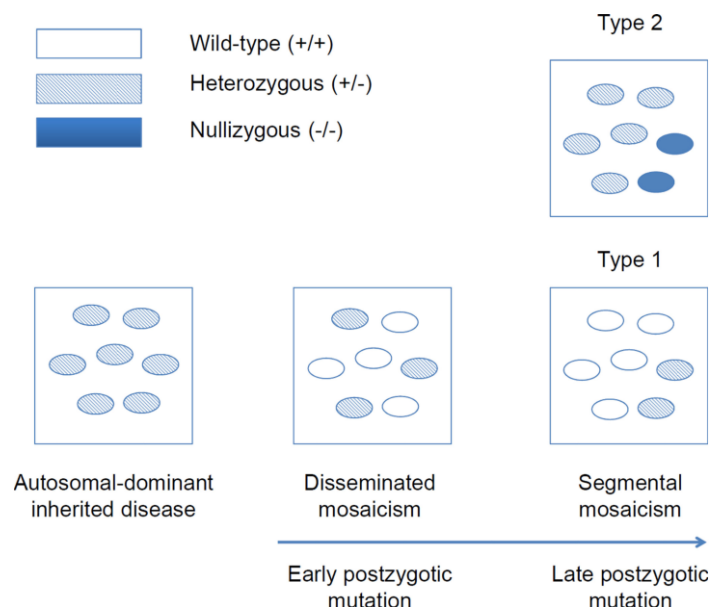


Figure 5 : Les différentes formes de mosaïques dans les gènes de transmissions autosomiques dominants (Figure de Neera Nathan et al.(103))

d. Les autres gènes du panel

Le taux de mutations pathogènes *de novo* dans les gènes MMR a été estimé à environ 2,3% des patients porteurs d'une mutation MMR. Ainsi, la très grande majorité des mutations MMR sont héritées (104). De ce fait, les cas de néomutations constitutionnelles post-zygotiques ont été exceptionnellement décrits. Par exemple, un cas de mosaïque constitutionnelle du gène *MSH2* a été rapporté. Il s'agit d'une femme chez laquelle la mutation du gène *MSH2* : c.2541del était détectée au niveau de sa tumeur colorectale ainsi qu'au sein de son tissu sain colorectal périphérique mais non détectée dans son sang. La preuve de la mosaïque constitutionnelle a été faite lorsque cette dernière a eu un fils présentant un syndrome de Lynch et porteur de la même mutation à l'état hétérozygote confirmant ainsi la mosaïque constitutionnelle somatique et germinale de la patiente (105).

Des cas de mutations *de novo* constitutionnelles du gène *CDHI* ont été rapportés dans la littérature car certaines mutations faux sens à l'état hétérozygote du gène se manifestent par un syndrome rare appelé le syndrome blépharochéilodontique (BCD) (OMIM # 119580). Ce syndrome se caractérise par des anomalies diverses des paupières, des lèvres telles qu'une fente labiopalatine et des dents (106). De manière intéressante, il n'a pas été mis en évidence de surrisque de cancer dans ce syndrome. On parle de relation génotype phénotype car les mutations pathogènes impliquées dans le syndrome BCD sont systématiquement des mutations faux sens situées dans un domaine de liaison au calcium extracellulaire alors que ces mutations ne sont pas identifiées dans les syndromes de prédisposition au cancer de l'estomac (107). Des études d'exome en trio ont permis de révéler la présence d'une néomutation du gène *CDHI* chez l'enfant atteint du syndrome BCD et non présente, par définition, chez les parents. Seuls trois articles rapportent des cas de mutation *de novo* du gène *CDHI* dans un cadre de prédisposition au cancer. Le premier décrit une femme qui a présenté un cancer du sein lobulaire à 54 ans avec mise en évidence de la mutation causale dans le dernier exon du gène *CDHI* (108) et le deuxième cas concerne un homme ayant présenté un cancer gastrique diffus sporadique sur le plan familial à 41ans avec la mise en évidence d'une large délétion hétérozygote du gène *CDHI* (109). Le troisième cas concerne une patiente asymptomatique porteuse du variant délétère de *CDHI* dont la chirurgie prophylactique a mis en évidence quatre petits foyers d'adénocarcinome intra-muqueux en anneaux (110).

VII. NGS et mutation somatique acquise : sources d'un diagnostic différentiel difficile.

Comme on l'a vu précédemment, avec le séquençage haut débit, il devient plus aisé de détecter des variants génétiques à une fréquence allélique inférieure à 50% (nombre de *reads* mutés / nombre de *reads* sauvages). Si ces situations peuvent être évocatrices de la présence d'un variant en mosaïque chez l'individu, ces VAF faibles peuvent aussi être dues à la détection d'ADN tumoral circulant (ADNtc), d'autant plus si la population testée est atteinte de cancer. Enfin, ils peuvent aussi être expliqués par l'hématopoïèse clonale (HC), c'est-à-dire par l'acquisition au niveau des cellules souches hématopoïétiques, précurseurs des cellules sanguines d'une mutation qui sera ainsi détectée dans une partie des cellules sanguines issues de cette cellule mère.

Il est important de faire la distinction entre ces trois origines car les implications pour la prise en charge du patient sont très différentes. De plus, la mosaïque constitutionnelle implique un risque de transmission possible à la descendance ce qui n'est pas le cas des deux autres évènements.

1. Les mutations tumorales

a. Mutations somatiques tumorales testées dans le cancer du sein et de l'ovaire

La détection des mutations somatiques acquises au niveau de la tumeur par réalisation d'un séquençage haut débit directement à partir d'un échantillon tumoral est de plus en plus pratiqué dans la prise en charge de cancer (111).

Le profil moléculaire des tumeurs est un élément fondamental de la médecine de précision car il permet d'identifier une dérégulation de certaines voies impliquées dans la cancérogenèse pouvant être ciblées de manière thérapeutique. Pour le cancer de l'ovaire, l'arrivée des inhibiteurs de PARP illustre bien cet avantage. Les inhibiteurs de PARP agissent en empêchant la réparation des cassures simple brin de l'ADN qui, lors de la phase de réplication, deviennent des cassures double brin. Ainsi, dans les cellules tumorales mutées dans les gènes *BRCA1* ou *BRCA2*, le mécanisme de réparation des cassures double brin par recombinaison homologue étant déficitaire, des mécanismes de réparation alternes moins fidèles se mettent en place entraînant une augmentation importante des anomalies génomiques jusqu'à la mort de la cellule. Ce synergisme est appelé létalité synthétique. Il est recommandé de rechercher une mutation somatique de BRCA quand aucune mutation constitutionnelle n'a été identifiée car environ 3 à 10 % des patientes avec un cancer de l'ovaire présentent une

mutation somatique exclusive de *BRCA*, sans altération constitutionnelle (112) (113) (114) (115) (116) Selon les données du Cancer Genome Atlas, 10-20 % des patients atteints d'un cancer du sein triple négatifs sont porteurs d'une mutation *BRCA* délétère et 2-3 % portent une mutation de *BRCA* acquise au niveau tumoral (117) (118).

Le gène le plus fréquemment muté dans les tumeurs est le gène *TP53* (119). La fréquence des mutations ponctuelles de *TP53* est de 36 % dans le cancer du sein et à 91 % dans le cancer de l'ovaire selon les données du TCGA (« The Cancer Genome Atlas »). Ainsi, trouver une mutation pathogène de *TP53* au niveau tumoral est fréquent et ne doit pas faire évoquer en premier lieu une mosaïque de *TP53* sans être confirmé sur un autre tissu.

Concernant les autres mutations somatiques ponctuelles rapportées parmi les 12 gènes étudiés, les plus fréquentes dans le cancer du sein concernent les gènes *CDH1* et *PTEN* avec des fréquences à 15 % et 7 % (TCGA). Dans le cancer de l'ovaire, l'autre gène le plus muté après *TP53* toujours parmi les 12 gènes est *BRCA1* avec une fréquence de mutations somatiques ponctuelles rapportées à 5,5 % (Données du TCGA).

b. Définition de l'ADN tumoral circulant

En cancérologie, différents bio-marqueurs protéiques relargués dans la circulation sanguine par la tumeur sont utilisés pour suivre l'évolution du cancer chez l'individu et en particulier la réponse au traitement. Par exemple, dans le cancer de l'ovaire le biomarqueur utilisé est le CA-125 (« Cancer Antigen-25 ») et dans le cancer du sein il s'agit du CA 15.3 (« Cancer Antigen-15.3 »). Cependant, l'expression de ces marqueurs est parfois peu élevée voire nulle chez certains patients qui présentent pourtant des cancers avancés. Ainsi, ces marqueurs ne peuvent permettre un suivi que des patients présentant un taux de marqueur augmenté avant le traitement (120). Par ailleurs, l'expression de ces marqueurs peut aussi être augmentée dans des pathologies bénignes (121).

Une nouvelle génération de biomarqueurs plus spécifiques a été mise en évidence depuis la découverte des altérations génétiques dans les tumeurs responsables de l'initiation et de la progression tumorale (122). L'étude des mutations somatiques acquises dans la tumeur est désormais permise notamment avec l'arrivée du séquençage à haut débit. Dans les cellules non tumorales, les mutations acquises sont en nombre négligeables. Ainsi, les mutations d'origine tumorale offrent des perspectives intéressantes du fait de leur spécificité comme potentiels biomarqueurs pour le suivi des patients (122) ou comme potentielles cibles thérapeutiques .

En 1869, Ashworth, un universitaire australien, découvrit des cellules tumorales similaires dans le sang périphérique d'un patient atteint d'une maladie métastatique et proposa pour la première fois le concept de cellules tumorales circulantes (123). Ces cellules tumorales circulantes peuvent aussi être détectées dans la circulation sanguine par des marqueurs membranaires spécifiques détectés par immunohistochimie ou immunophénotypage (124).

L'ADN libre circulant a été décrit pour la première fois en 1948. Sa concentration peut être augmentée chez les patients atteints d'un cancer en raison de la présence d'une fraction supplémentaire d'ADNtc, mais également dans d'autres situations physiologiques comme dans les cas de traumatismes, d'infections, d'inflammations mais aussi chez la femme enceinte. L'ADNtc est un fragment d'ADN double brin, généralement sous la forme d'un complexe protéique, dont la taille varie de 70 à 200 pb. Ce dernier est libéré dans la circulation sanguine à partir des cellules tumorales par apoptose, par nécrose et même de manière active (125) (126) (Figure 6). La quantité d'ADN circulant dans le sang des patients souffrant d'un cancer varie considérablement, allant de 10 à 100 ng / ml, c'est-à-dire similaires à ceux des patients non atteints de cancer à des taux atteignant les 5000ng d'ADN / ml de plasma (126) (127). Cependant, des taux d'ADN significativement plus élevés ont été trouvés dans le sérum de patients atteints d'une maladie métastatique par rapport à des patients non métastatiques (127) . Par ailleurs, la clairance de l'ADN libre circulant n'est pas entièrement comprise. L'ADN libre circulant a une demi-vie courte de moins de 2 h.

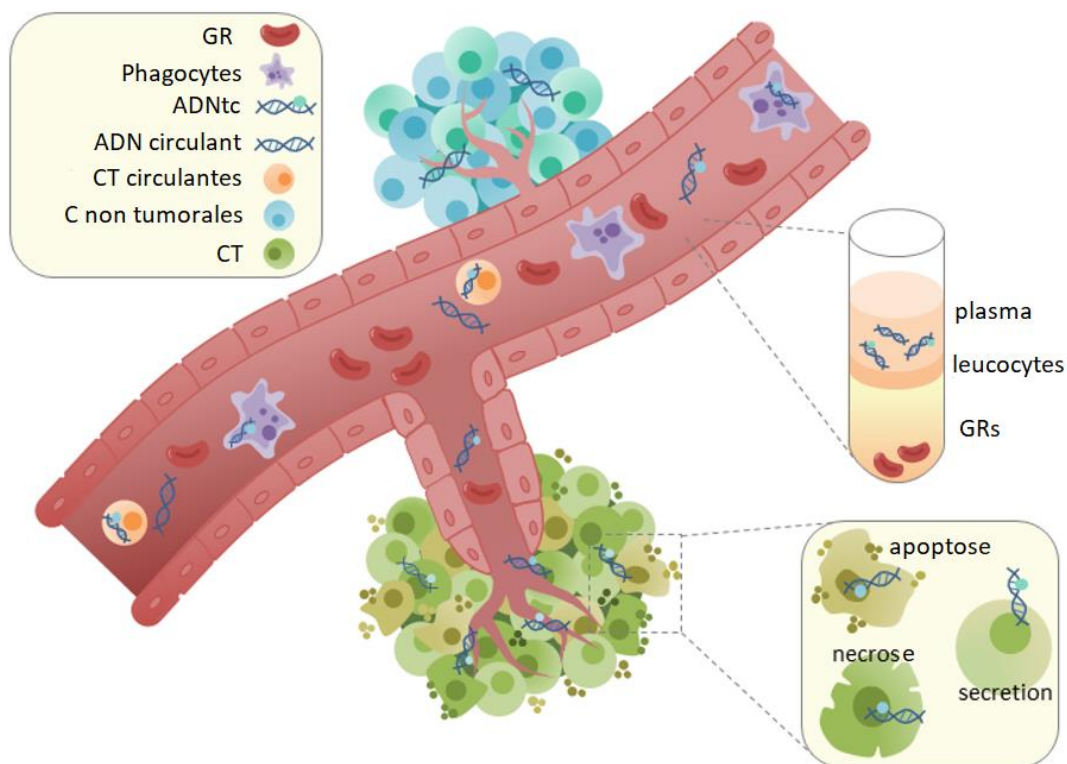


Figure 6 : Représentation de l'ADN tumoral circulant (Image traduite de Rachel June Wong, article issu de EuropeansBiotechnology, 08/08/2019 (128))

L'ADNtc est retrouvée dans les fractions de sérum et de plasma sanguin. Le mécanisme de relargage de l'ADN tumoral circulant n'est pas encore totalement élucidé. Il pourrait provenir de la sécrétion, de la nécrose et/ou de l'apoptose des cellules tumorales.

Abréviations : GR : Globules rouges ; ADNtc : ADN tumoral circulant ; CT : cellules tumorales ; C : cellules

c. Techniques de détection de l'ADN tumoral circulant

Afin de quantifier l'ADNtc d'un patient, un prélèvement de sang total sur tube EDTA (Éthylène-Diamine-Tétra-Acétique) est possible mais dans ce cas, il est nécessaire d'isoler le plasma des cellules circulantes car autrement, l'ADNtc sera contaminé par l'ADN génomique issu de la lyse cellulaire. Ainsi, il existe des tubes spécifiques tels que les tubes STRECK® qui empêchent la lyse cellulaire et permettent ainsi une conservation plus longue du prélèvement avant sa prise en charge. L'isolation du plasma nécessite une centrifugation douce dans un premier temps pour éliminer les cellules sans les lyser puis une centrifugation plus rapide afin de retirer les débris cellulaires. L'extraction se fait avec des kits dédiés aux petits fragments.

A l'origine, les méthodes de détection de l'ADNtc étaient des techniques ciblées. Ainsi, soit il fallait identifier la mutation dans la tumeur primitive pour la rechercher dans le sang, soit seules des mutations hotspot bien définies et conférant par exemples des résistances au traitement étaient recherchées. Parmi ces méthodes ciblées, des techniques telles que la digital PCR (dPCR) permettant une quantification absolue de l'ADNtc peuvent descendre à un niveau de sensibilité très bas (129). La méthode dPCR la plus courante est la « droplet digital PCR » (ddPCR), caractérisée par la discrétisation et l'amplification de chaque matrice d'ADN dans des gouttelettes d'émulsion simples.

Cependant, depuis plusieurs années, une recherche plus globale est non ciblée de l'ADNtc est rendu possible grâce à l'arrivée du séquençage à haut débit. L'avantage principal de cette technique est de permettre la recherche de l'ensemble des mutations somatiques d'un panel de gènes jusqu'à l'étude de l'exome (130).

d. Applications en oncologie

L'une des applications les plus immédiates de l'ADNtc est appelée « biopsie liquide » (131). L'avantage de la biopsie liquide par rapport à la biopsie solide (ou d'organe) est qu'il s'agit d'une méthode non invasive, répétable, permettant un suivi (130) notamment de la réponse au traitement et reflétant l'ensemble de la tumeur. En effet, les biopsies de tumeurs ou de blocs sont parfois difficiles voire impossibles à obtenir en cas de localisation non opérable et ne sont pas toujours de bonne qualité.

Bettegowda et al., ont montré par l'étude en digital PCR de 641 patients atteints de divers types de tumeurs que la quantité d'ADNtc détecté dans le plasma est corrélée à la taille de la tumeur et au caractère métastatique de la maladie. Par ailleurs, ils ont aussi montré que les taux d'ADNtc dépendent de la localisation initiale de la tumeur (Figure 7) (127).

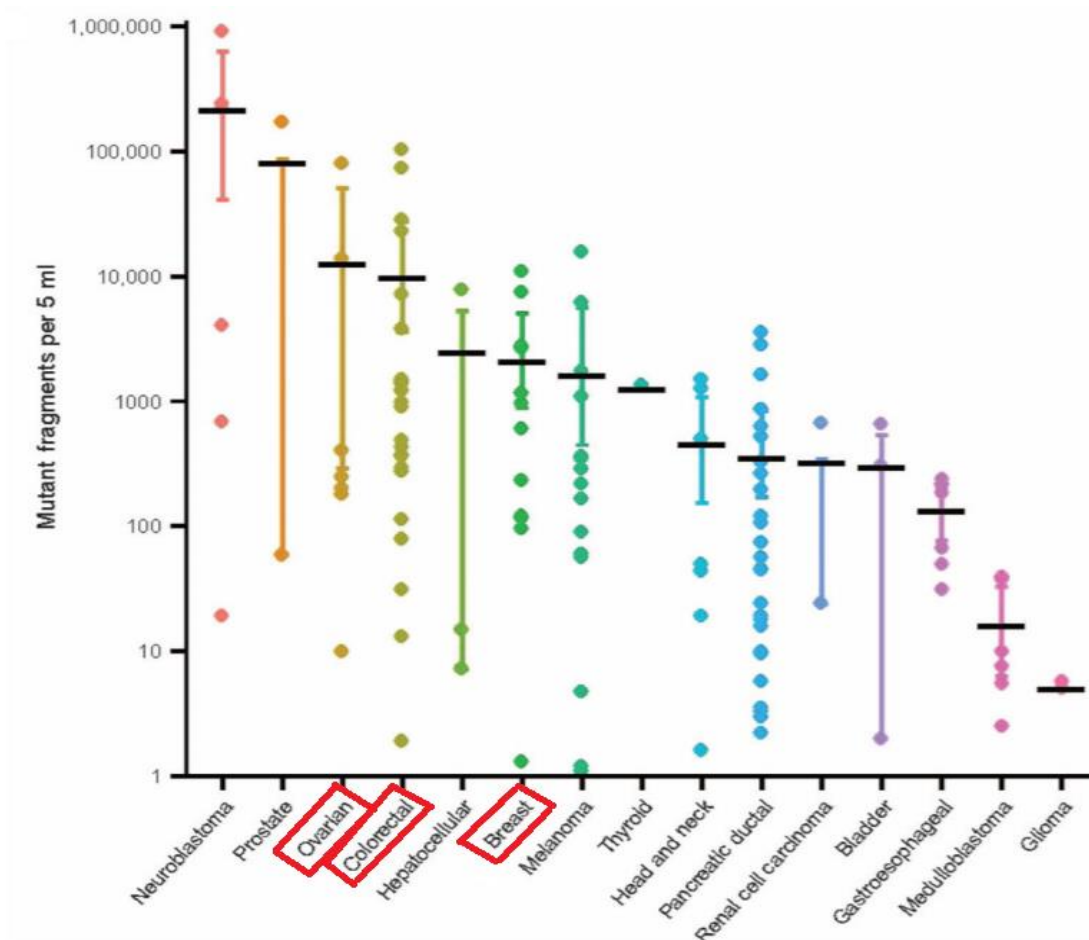


Figure 7: Graphique représentant la quantité d'ADN muté en fonction du type de cancer (Image issue de Bettgowda *et al.*, 2014 (132))

Les barres d'erreur représentent l'intervalle de confiance de 95 % de la moyenne. Les cancers encadrés en rouge sont ceux qui nous intéresseront particulièrement pour notre étude.

Dans le cadre du cancer du sein, les études de l'ADNtc sont assez récentes (133). Par exemple, Dawson *et al.* ont montré un intérêt à la détection de l'ADNtc dans le cancer du sein. Sur les 69 patientes évaluées, l'ADNtc était détecté chez 41 %, 20 % et 5 % des patientes avant chimiothérapie adjuvante, après 2 semaines de traitement et juste avant la chirurgie, respectivement. La détection de l'ADNtc avant chimiothérapie néoadjuvante était significativement associée à un âge avancé et à un statut ER (« Estrogen receptor ») négatif. Enfin, la détection de l'ADNtc avant chimiothérapie néoadjuvante était associée à une probabilité moindre d'atteindre une réponse complète (134). De nombreuses autres études ont montré les nombreux intérêts à la recherche de l'ADNtc dans le cadre du cancer du sein pour le suivi, la classification pronostique et la mise en place d'un traitement ciblé (135).

La quantité d'ADNtc libéré dans le cancer de l'ovaire est plus important en moyenne que dans le cancer du sein (132). L'une des applications évidentes à la recherche d'ADNtc dans le cancer du sein et de l'ovaire est de rechercher la présence de mutations somatiques ou constitutionnelles des gènes *BRCA1* et *BRCA2* orientant dans les deux cas vers l'indication d'une thérapie par inhibiteur de PARP. Ainsi Ratajska et al ont détecté des mutations de *BRCA1* et *BRCA2* chez 30 patientes sur 121 atteintes de cancer de l'ovaire (24,8 %) incluant 22 mutations germinales, 7 mutations exclusivement somatiques et une patiente porteuse à la fois d'une mutation constitutionnelle et d'une mutation somatique (136). Ainsi, lorsque l'accès à la tumeur n'est pas possible ou lorsque l'ADN tumoral n'est pas amplifiable, cette option paraît une stratégie prometteuse. Par ailleurs, elle permettrait une analyse à la fois constitutionnelle et somatique en un seul temps optimisant ainsi la prise en charge thérapeutique.

2. L'hématopoïèse clonale

a. Définition de l'hématopoïèse clonale

L'HC se réfère généralement à toute expansion clonale de cellules hématopoïétiques, détectée par n'importe quelle méthode. Elle n'est pas connue pour avoir une implication clinique ou pronostique. L'HC de potentiel indéterminé (CHIP) est l'HC qui concerne des mutations récurrentes dans les gènes associés aux hémopathies malignes tels que *DNMT3A*, *TET2*, *JAK2*, *SF3B1*, *ASXL1*, *TP53*, *CBL*, *PPM1D* (**Figure 8**). Ces clones doivent être détectés à une VAF d'au moins 0,02 chez l'adulte sans antécédents connus d'hémopathie maligne (137). Ainsi, les conséquences cliniques néfastes de l'HC sont influencées à la fois par le type de mutation génétique présente dans le clone et par le nombre de cellules composant le clone (138).

Population et fréquence	Gènes mutés	Association
Mutations les plus communes dans la population non sélectionnée	<i>DNMT3A, TET2, ASXL1</i>	Mutations driver dans les SMD (Syndrome myélodysplasique) / leucémie ; toutes codent pour des régulateurs épigénétiques ou chromatinien qui peuvent conférer aux CSH un avantage de croissance ou de survie. → Associés à la fois à un risque de progression vers la maladie hématopoïétique et à des événements vasculaires
Mutations communes après traitement cytotoxique	<i>TP53, PPM1D</i>	Facteurs de risque importants de néoplasie myéloïde liée au traitement. Il s'agit de mutations préexistantes sélectionnées par le traitement anticancéreux.
Autres mutations retrouvées un peu moins communes	Facteurs d'épissages (<i>SF3B1, SRSF2, U2AF1...</i>), <i>EZH2, RUNX1, JAK2, MPL</i> , Voie RAS, Complexe des cohésines, <i>CUX1, ATM, GATA2, WT1, ETV6, IDH1, IDH2, PHF6</i>	Risque néoplasique variable. Certaines mutations peuvent conduire à une expansion clonale avec un seul événement, d'autres ne le peuvent probablement pas.

Tableau 2 : Gènes les plus fréquemment mutés dans la CHIP (Tableau adapté de Gibson et Steensma, 2018 (138))

Abréviations : CHS : cellules souches hématopoïétiques

b. Recommandations en termes de diagnostic et de suivi

Dans l'ensemble des études réalisées jusqu'à maintenant, la détection d'une CHIP est associée à une augmentation de la mortalité par rapport au groupe contrôle apparié sur l'âge (138). Le risque de progression de la CHIP vers une hémopathie maligne est de 0,5 % à 1 % par an (137).

Le séquençage de nouvelle génération étant désormais bien intégré, il est devenu relativement facile d'un point de vue technique d'identifier les patients ayant une HC. Cependant, à l'heure actuelle, il n'est pas recommandé de dépister la CHIP chez des personnes asymptomatiques car il n'existe pas de mesure préventive évitant l'apparition éventuelle d'une hémopathie maligne. De même, lorsque cette CHIP est détectée par « accident » lors de la réalisation d'un séquençage à haut débit pour une autre recherche, la détection de ce clone ne modifie pas la prise en charge du patient.

VIII. Objectifs de la thèse

L'objectif de l'étude est de déterminer la fréquence et la signification de la détection de variants pathogènes de VAF faibles détectés par séquençage à haut débit dans le cadre du diagnostic moléculaire des prédisposition aux cancers du sein et de l'ovaire depuis un an au sein du laboratoire du CFB à Caen. En effet, le principal but de l'étude est d'éliminer la présence d'une mosaïque chez les cas index qui impliquerait une surveillance adaptée pour le patient porteur et un risque de transmission du variant à la descendance. L'objectif secondaire est d'établir un protocole simple en cas de détection de ces variants de faible VAF pour identifier les vraies mosaïques constitutionnelles. Enfin, le dernier objectif est d'évaluer si les variants à faible VAF pourraient faire l'objet d'une recommandation pour le diagnostic constitutionnel de prédisposition au cancer du sein et de l'ovaire.

PATIENTS ET METHODE

1. Population étudiée

La population étudiée concerne l'ensemble des patients ayant bénéficié d'une analyse de séquençage à haut débit avec le panel HBOC 12 gènes pour la recherche de prédisposition aux cancers du sein et de l'ovaire au sein du laboratoire du CFB à Caen de mai 2018 à mars 2019. Au total, 2007 femmes ont été incluses dans l'étude. L'ensemble des patientes ont signé un consentement pour la réalisation d'un test génétique pour la recherche de prédisposition au cancer du sein et de l'ovaire.

Pour les patients présentant un variant à VAF pathogène ou probablement pathogène des analyses complémentaires étaient effectuées dans le cadre du diagnostic moléculaire. Un courrier était envoyé aux prescripteurs afin d'accéder au prélèvement tumoral (inclus en paraffine) du patient.

Aussi, cette demande permettait de colliger des informations complémentaires concernant notamment l'état clinique du patient au moment du prélèvement, en particulier s'il était en début de traitement, en rechute ou en rémission.

2. Recherche de variant constitutionnel par séquençage à haut débit

L'ADN des patients a été extrait du sang périphérique et des frottis jugaux à l'aide d'un robot Qiagen QIASymphony™ SP par la méthode « DNA BLOOD 400 V6 DSP ». Les ADN extraits ont ensuite été dosés selon la technique Qubit à l'aide du robot Beckman Biomek FX®. L'ADN a été traité par sonification à l'aide d'un Covaris S2 (Covaris®, MS). Les échantillons ont été préparés avec un SPRIworks HT (Beckman®). Les adaptateurs Illumina ont été remplacés par des adaptateurs indexés (NEB®, Milan, Italie) publiés antérieurement par l'équipe de Huentelman (142). La préparation des bibliothèques a été effectuée avec le kit d'enrichissement Sureselect (Agilent®, Santa Clara, CA). Le panel de gènes utilisé au laboratoire du CFB est représenté ci-dessous. Seuls les gènes de rang A et B sont rendus dans un cadre diagnostique et accrédités selon la norme ISO15189 (**Tableau 3**). Seuls ces 12 gènes étaient analysés pour cette étude dont le principal objectif était diagnostique.

Pour chacun des 12 gènes, l'ensemble des exons et au moins les 20 premières bases et les 20 dernières bases de l'intron étaient couverts par la capture. Le processus d'enrichissement a été réalisé conformément à Kenny et al (138, 139). Le protocole a été robotisé sur un Biomek

FX®. Les bibliothèques ont été séquencées sur un NextSeq 500 (Illumina®, San Diego, Californie) en utilisant le programme paired-end 2 x150 pb. Concernant l'analyse bio-informatique des données générées par le séquençage, le démultiplexage a été effectué avec CASAVA Suite v1.8, l'alignement par BWA 0.7.12(143) et la recalibration des fichiers BAM avec la suite ineq GATK(144) v3.3 pour produire des fichiers BAM, conformément aux recommandations du Broad Institute.

Le variant caller Outlyzer v.2 a été utilisé pour la détection des SNV et des indels car il a une bonne sensibilité pour détecter les variants de faible ratio allélique (62). Les variants ont été annotés avec Alamut-Batch (Interaction Biosoftware). Le génome utilisé est l'Homo sapiens, hg19. Les régions exploitées pour le diagnostic moléculaire des prédispositions au cancer du sein et de l'ovaire sont les régions codantes et les régions autour de celle-ci (-20pb/+10pb dans les introns).

GENE	TRANSCRIT ANALYSE	RANG	GENE	TRANSCRIT ANALYSE	RANG
<i>ATM</i>	NM_000051	C	<i>MUTYH</i>	NM_001048171	C
<i>BAP1</i>	NM_004656	C	<i>NBN</i>	NM_002485	C
<i>BARD1</i>	NM_000465	C	<i>PALB2</i>	NM_024675	A
<i>BRCA1</i>	NM_007294	A	<i>PMS1</i>	NM_000534	C
<i>BRCA2</i>	NM_000059	A	<i>PMS2</i>	NM_000535	B (uniquement exons 6-7-8-10)
<i>BRIP1</i>	NM_032043	C	<i>PTEN</i>	NM_000314	B
<i>CDHI</i>	NM_004360	B	<i>RAD50</i>	NM_005732	C
<i>CDKN2A</i>	NM_000077	C	<i>RAD51</i>	NM_001164269	C
<i>CDKN2A</i>	NM_058195	C	<i>RAD51B</i>	NM_133509	C
<i>CHEK2</i>	NM_001005735	C	<i>RAD51C</i>	NM_058216	A
<i>FAM175A</i>	NM_139076	C	<i>RAD51D</i>	NM_002878	A
<i>HOXB13</i>	NM_006361	C	<i>RINT1</i>	NM_021930	C
<i>INHA</i>	NM_002191	C	<i>STK11</i>	NM_000455	C
<i>INHBA</i>	NM_002192	C	<i>TP53</i>	NM_000546	B
<i>MLH1</i>	NM_000249	B	<i>XRCC2</i>	NM_005431	C
<i>MLH3</i>	NM_001040108	C	<i>XRCC3</i>	NM_001100119	C
<i>MRE11A</i>	NM_005591	C			
<i>MSH2</i>	NM_000251	B			
<i>MSH6</i>	NM_000179	B			

Tableau 3 : Liste du panel 34 gènes réalisé dans le cadre du panel HBOC au laboratoire.

3. Sélection des variants à partir des données de séquençage

La recherche des variants ayant un faible ratio allélique a été réalisée à partir d'extractions de la base de données CanDiD (« Cancer Diagnostic Database ») développée au laboratoire. Il s'agit d'une base de données interfacée qui stocke l'ensemble des variants annotés appelés par les *variant-caller* à chaque série de patient. L'interface permet l'écriture de pipeline de validation basée sur des successions de requêtes standardisées et validées, permettant de filtrer et sélectionner les variants utiles au diagnostic des patients testés. Pour cette étude il a été rédigé une requête interrogeant les variants appelés par Outlyzer pour la détection des variants de faible ratio allélique. Par ailleurs, seuls les variants des 12 gènes du panel HBOC ayant un ratio allélique compris entre 3 % et 35 % ont été sélectionnés. Concernant les données recueillies pour chaque variant, nous avons choisi d'appeler les données suivantes :

- Le nom du variant avec le coding name c. et le protéine p.
- Le nom du gène et la séquence de référence
- La fraction allélique du variant (VAF)
- La profondeur de séquençage soit le nombre de *reads* lus à la position du variant
- Le nombre de *reads* mutés lus en Forward et en Reverse
- Les données gnomAD et ExAC de fréquence dans la population du variant
- Les données de prédictions protéiques : le score Align GVGD (145) (146), le score Phred (147) (148) et le score CADD (149)
- Les données de prédictions sur l'épissage : les scores MES, SSF et le score SPiCE qui combine les scores MES et SSF(150)
- L'interprétation du variant par InterVar (151) qui est un logiciel de bioinformatique qui pré-classe les variants selon les directives de l'ACMG 2015 en « bénin », « probablement bénins », « signification incertaine », « probablement pathogène » et « pathogène » à partir d'un fichier annoté généré par ANNOVAR.
- Les annotations privées du laboratoire lorsque les variants avaient déjà été analysés antérieurement

Les variants extraits de CanDiD avec les critères de filtre détaillés ci-dessus étaient traités sur un tableur Excel par série de patients. Une série correspondait aux patients reçus en une semaine soit une soixantaine de patients en moyenne. Les filtres étaient effectués de la façon suivante :

- (1) Les variants redondants observés chez plus d'un patient dans la série étaient retirés. Cela permettait notamment d'enlever les faux positifs récurrents.
- (2) Un filtre était appliqué pour retirer les variants classés 1 ou 2 par le LBGC score qui correspondait à des SNP (« Single Nucleotide Polymorphism »).
- (3) L'ensemble des séries de variants étaient réunies pour éliminer les nouvelles redondances.
- (4) Un dernier filtre était appliqué pour éliminer les variants présents sur moins de 8 *reads* ou observés dans un seul sens (en *forward* ou *reverse*) grâce à la balance allélique.
- (5) Enfin, les variants du gène *PMS2* hors des exons 6, 7, 8 et 10 étaient aussi retirés de la sélection car ces zones correspondent au pseudogène du gène *PMS2* et le ratio allélique déséquilibré a certainement plus de probabilité d'être expliqué par la capture et le séquençage du pseudogène qu'à une véritable mosaïque.

Enfin, les variants restants étaient explorés un par un sur Alamut Visual, UMD-*BRCA1* et 2, Insight pour les gènes *MMR* et classés selon les critères de l'ACMG (65) afin de ne garder que les variants pathogènes ou probablement pathogènes. La base de données IARC (« International Agency for Research on Cancer ») a été utilisée pour interpréter les variants de *TP53*, en particulier les résultats des analyses fonctionnelles(152). Un variant rare de *PTEN* de classe 3 ne semblant pas délétère par les outils de prédiction a aussi été exploré car les variants exoniques de *PTEN* sont rares dans la population générale contrairement aux autres gènes étudiés. Et, il nous paraissait intéressant d'étudier la présence ou non de ce variant sur la tumeur afin de contribuer à sa classification. Afin d'éliminer la sélection de variants dus à une éventuelle cross-contamination, c'est-à-dire la contamination d'un échantillon entre les séries ou dans une même série, chaque variant a été recherché dans la base de données de CanDiD afin de vérifier s'il n'aurait pas été détecté à l'état hétérozygote chez un autre patient d'une même série ou d'une série effectuée peu de temps avant la série du patient concerné.

4. Analyse par séquençage à haut débit des blocs tumoraux

Les blocs reçus inclus en paraffine étaient extraits sur Qiacube avec le kit QIAamp® DNA FFPE tissu (QIAGEN®) après découpage de copeaux au microtome. Une vérification du pourcentage de cellules tumorales sur lame à partir d'une coloration à l'hématoxyline éosine safran était effectuée. Si possible du tissu sain avoisinant était extrait après détournement du bloc tumoral inclus en paraffine. Onze échantillons de tumeurs inclus en paraffine provenant de 11

patientes dont 8 cancers de l’ovaire, 3 tumeurs du sein et 1 tumeur du colon ont pu être analysées par séquençage à haut débit. La préparation des bibliothèques a été réalisée par Agilent SureDesign (Agilent®, Santa Clara, Californie, États-Unis). Les régions d’intérêt ont été capturées avec le protocole SureSelect XT (Agilent, Santa Clara, Californie, États-Unis) et séquencées sur Illumina Miseq (Illumina, San Diego, Californie, États-Unis) à l’aide du programme pair-end 2 × 150 pb. L’analyse bio-informatique des données générées par le séquençage, le démultiplexage a été effectué avec l’alignement par BWA 0.7.12 (143) et la recalibration des fichiers BAM avec la suite GATK v3.3 (144) pour produire des fichiers BAM, conformément aux recommandations du Broad Institute. Le variant caller utilisé pour l’analyse était Outlyzer v2.

5. Détermination de la taille des ADN extraits et purification des ADN des hauts poids et bas poids moléculaire.

L’ensemble des 21 échantillons d’ADN issus de prélèvements sanguins ayant un variant d’intérêt ont été analysés en TapeStation® à l’aide du kit D1000 ScreenTape® (Agilent®) selon les recommandations du fabricant. Cette technique permet à la fois d’obtenir la taille de fragment d’un échantillon ainsi que leur concentration.

Lorsque l’analyse en TapeStation montrait la présence de fragments de petites tailles, une électrophorèse en gel d’agarose à 1 % à 110 V et 70mA était effectuée afin de séparer les fragments d’ADN en fonction de leur taille. Une photographie du gel était effectuée sur le ChemiDoc® Imaging system à l’aide du logiciel Image Lab4.1 (Biorad®). Ensuite, un découpage des bandes d’intérêt était effectué à l’aide d’un scalpel et d’une lampe à UV (UltraViolet). Les fragments de taille inférieure à 500 pb étaient découpés ensemble lorsqu’il était difficile de discriminer visuellement de manière précise la taille des fragments du fait de leur faible concentration dans le gel. En revanche, pour un échantillon pour lequel les fragments de petite taille étaient facilement distincts, il a été possible de séparer les fragments de taille de 150 pb et ceux à 200 pb. Enfin, la purification de l’ADN contenu dans les bandes de gel d’agarose était effectuée avec le kit NucleoSpin® Gel PCR Cleanup (Macherey-Nagel) selon les recommandations du fabricant en vue de la réalisation ultérieure d’une analyse Snapshot. L’ADN était conservé ensuite à -20°C avant son utilisation.

6. Analyse Sanger

Parmi les 23 variants d'intérêts détectés chez 21 patients, 10 échantillons d'ADN issus de prélèvements sanguins ont été analysés par la technique de séquençage par Sanger car la fréquence allélique à laquelle ils étaient détectés était supérieure à 10 %. Les variants pour lesquels la VAF était inférieure à 10 % n'ont pas été explorés par Sanger à cause de la limite de détection de cette technique. Lorsqu'un frottis jugal était disponible, le séquençage par Sanger a également été réalisé sur ce prélèvement.

Les réactions de PCR en vue du séquençage Sanger ont été réalisées avec les amorces décrites dans le **tableau 4**. Le mix de PCR de 21 μL comportait 1,9 mM de MgCl_2 (Applied Biosystems®), 760 μM de dNTPs (« desoxyribonucleotide triphosphate ») (Applied Biosystems®), 0,8 X de tampon (Applied Biosystems®), 0,5 μM d'amorce et 0,85 U de Taq Gold (Applied Biosystems®) et de l'eau stérile auquel était ajoutée 4 μL d'ADN extrait dont la concentration est comprise entre 1 et 30 $\text{ng}/\mu\text{L}$. Le programme de PCR comportait une activation de l'enzyme à 95°C pendant 10 minutes suivie de 26 cycles de gradient de température comportant 20 secondes à 94°C, 20 s 63°C (-0,5°C par cycle), et 45 s à 72°C enfin 20 nouveaux cycles étaient réalisés, composés de 20 s à 94°C, 20 s à 50°C et 45 s à 72 °C suivie par une étape finale d'extension à 72°C pendant 7 minutes. Le produit de PCR était purifié avec le kit Illustra™ ExoProStar™ selon les instructions du fabricant. Le séquençage était réalisé en utilisant des amorces universelles avec le kit de séquençage Big Dye Terminator v1.1 (Applied Biosystem®) dans les conditions suivantes : 3 μL de mix, 6 μL d'amorce à 0,5 μM et 1 μL de produit de PCR purifié. Le cycle de séquençage comportait une étape de dénaturation initiale à 96°C pendant 1 minute suivie de 25 cycles comportant les étapes suivantes : 10s à 96°C, 5s à 50°C et 30s à 60°C dans un thermocycleur Eppendorf Mastercycler pro S (Applied Biosystem®). Le produit de séquençage était purifié en bille Séphadex G-50 Superfine Part (GE Healthcare) et la séparation réalisée par électrophorèse capillaire sur un séquenceur ABI 3130xl Genetic Analyzer (Applied Biosystems®) avec un capillaire de 36 cm et un polymère POP-7™. L'analyse des séquences a été effectuée avec le logiciel Sequencing Analysis Software v.5.4 (Applied Biosystems®/Life Technologies, Foster City, CA).

Gène	Position des variants explorés	Amorce Forward	Amorce Reverse	Taille de l'amplicon
<i>BRCA1</i>	898	GGCATCCAGAAAAGTATCAGG	TTTTTTTCTGTGCTGGGAGTC	250 pb
<i>TP53</i>	427, 470, 526, 589, 659	GAGGTGTAGACGCCAACTCT	GGTCAAATAAGCAGCAGGAG	627 pb
<i>TP53</i>	742	TTGCCACAGGTCTCCCCAA	GTGATGAGAGGTGGATGGGTA GTAGTATG	290 pb
<i>TP53</i>	817, 839	AAGGGTGGTTGGGAGTAGAT	TGTCTTTGAGGCATCACTGC	532 pb
<i>PTEN</i>	698	AAATCGTTTTTGACAGTTTGACA	GCTTTTAATCTGTCCTTATTTTG	295 pb

Tableau 4 : Amorces utilisées pour l'amplification en vue du séquençage Sanger

Une queue M13 a été ajoutée à l'ensemble des amorces afin de pouvoir réaliser la réaction de séquençage avec des amorces universelles utilisées au laboratoire et s'hybridant avec ces queues M13.

7. Technique Snapshot

Le Snapshot a été mis au point pour l'étude au laboratoire pour diminuer le seuil de détection des variants à VAF faible par rapport au Sanger et permettre l'analyse des frottis jugaux non séquençables par la technique NGS. Les tests ont été effectués sur l'ensemble des échantillons disponibles des 21 patients soit si possible le sang, le frottis buccal et l'ADN issu du bloc tumoral afin de confirmer la présence du variant et de quantifier de manière relative l'allèle muté par rapport à l'allèle sauvage entre les différentes origines de prélèvement chez un même patient.

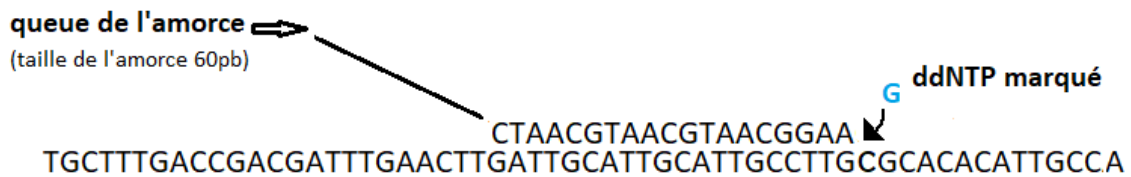
La technique Snapshot consiste tout d'abord à effectuer une PCR afin d'amplifier la séquence comportant la variation génétique. Ensuite, une amorce complémentaire dont l'extrémité 3' est située juste en amont du variant est ajoutée afin de réaliser une polymérisation d'un didésoxynucléotide triphosphate (ddNTP) marqué complémentaire du brin matrice à partir de l'extrémité 3' de l'amorce (extension d'amorce). Le ddNTP forme, en effet, une liaison ester entre le groupe 3'OH de l'amorce et le groupe 5'P du ddNTP. Cependant, la réaction s'arrête après la liaison du ddNTP car ce dernier ne possède pas de complexe 3'OH libre permettant la liaison d'un nouveau ddNTP. L'ajout du ddNTP provoque ainsi l'arrêt de l'élongation. Pour détecter la liaison du ddNTP, les différents types de ddNTP sont marqués par des fluorochromes différents.

Les produits de la réaction d'extension d'amorce subissent une électrophorèse sur un séquenceur à capillaire automatique. La taille de ces produits d'extension correspond à la taille de l'amorce d'extension plus le ddNTP ajouté. Un laser excite le fluorochrome et une caméra

capture l'émission de fluorescence pour obtenir un électrophorégramme après analyse des données brutes. Enfin l'électrophorégramme obtenu est comparé à celui d'un sujet contrôle de manière à identifier toute modification du profil. Lorsqu'il y a une variation chez un individu, on observe sur l'électrophorégramme deux pics de couleurs différentes migrant à la même taille, l'un correspondant au brin muté et l'autre au brin sauvage (**Figure 9**).

a)

Allèle sauvage :



Allèle muté :



b)

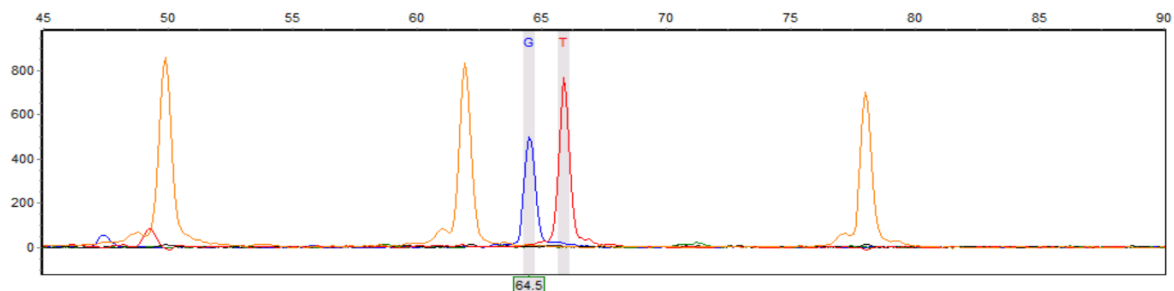


Figure 9 : Exemple de représentation d'une réaction d'extension d'amorce (Snapshot)

- a) Les sondes d'extension se fixent sur le brin matrice et un ddNTP marqué est incorporé. Le ddNTP G marqué par le dR110 visible en bleu se fixe sur l'allèle sauvage tandis que le ddNTP T marqué par le dROX™ visible en rouge se fixe sur l'allèle muté
- b) Résultat de l'électrophorégramme obtenu avec un pic bleu représentant l'allèle sauvage et un pic rouge représentant l'allèle muté. Le marqueur de taille marqué au LIZ est visible en orange. Les pics sont bien à la taille attendue soit aux environs de 60 pb qui est la taille de l'amorce d'extension.

Le protocole de PCR initial ainsi que la purification du produit de PCR étaient identiques à celui détaillé plus haut pour la réalisation du séquençage Sanger. Seules les amorces utilisées étaient différentes. En effet, afin de pouvoir réaliser une amplification des ADN issus des blocs tumoraux, des amorces ont été dessinées dans le but d'obtenir des produits d'amplifications de 150 pb environs compatibles avec les petits fragments d'ADN extraits ou purifiés. Ces amorces sont représentées dans le **tableau 5**. En effet, les ADN issus des blocs tumoraux sont dégradés et plus courts que les ADN issus des prélèvements sanguins.

Localisation de l'amorce	Amorce Forward	Amorce Reverse	Taille de l'amplicon
BRCA1 PCR 898	AGCTGAGAGGCATCCAGAAA	AGCCAGGCTGTTTGCTTTTA	179
TP53 PCR 814-844	GCCTCTTGCTTCTCTTTTCC	GTGAGGCTCCCCTTCTTG	166
TP53 PCR 711-742	TTGGGCCTGTGTTATCTCCT	GGCTCCTGACCTGGAGTCTT	147
TP53 PCR 659	GTGGAAGGAAATTTGCGTGT	CTTAACCCCTCCTCCCAGAG	120
TP53 PCR 577-589	GTCCCCAGGCCTCTGATT	TAGGGCACCACCACACTATG	137
TP53 PCR 517-526	CGCCATGGCCATCTACAA	AACCAGCCCTGTCGTCTCT	123
TP53 PCR 455-470	AGTACTCCCCTGCCCTCAAC	CTCACAACCTCCGTCATGTG	148
TP53 PCR 406-427	TGCCCTGACTTTCAACTCTGT	GCTGTGACTGCTTGATAGTGG	167
TP53 PCR 321	CTGGCCCCTGTCATCTTCT	AACTGACCGTGCAAGTCACA	113
PTEN PCR 601	GCTACCTGTAAAGAATCATCTGGA	GGAAGGATGAGAATTTCAAGCA	126
PTEN PCR 698	TGCAGATCCTCAGTTTGTGG	CACACACAGGTAACGGCTGA	122

Tableau 5 : Amorces d'amplification pour la technique Snapshot.

Une queue M13 a été ajoutée à l'ensemble des amorces afin de pouvoir réaliser des Sanger avec les amorces universelles utilisées au laboratoire et s'hybridant sur ces queues M13.

L'analyse Snapshot a été réalisée en utilisant le kit Snapshot Multiplex (Applied Biosystem®). Le mix de séquençage comportait 1,5 µl d'eau stérile, 5 µl de Snapshot Multiplex reaction Mix, 0,5 µl de Primer d'extension à 2 µM et 3 µl d'amplicon. Par ailleurs, les amorces d'extension spécifiques des variants trouvés sont représentées dans le **tableau 6**. Parfois, l'amorce choisie était complémentaire à la séquence située en aval du variant sur le brin reverse, lorsque la zone d'amont comportait une stretch par exemple. Les couleurs des pics correspondent dans ces cas à la séquence reverse et non à la séquence en forward. Le cycle de séquençage comportait une étape de dénaturation initiale à 96°C pendant 1 minute suivie de 25 cycles avec les étapes suivantes : 10 s à 96°C, 5s à 50°C et 30s à 60°C dans un thermocycleur Eppendorf Mastercycler pro S (Applied Biosystem®).

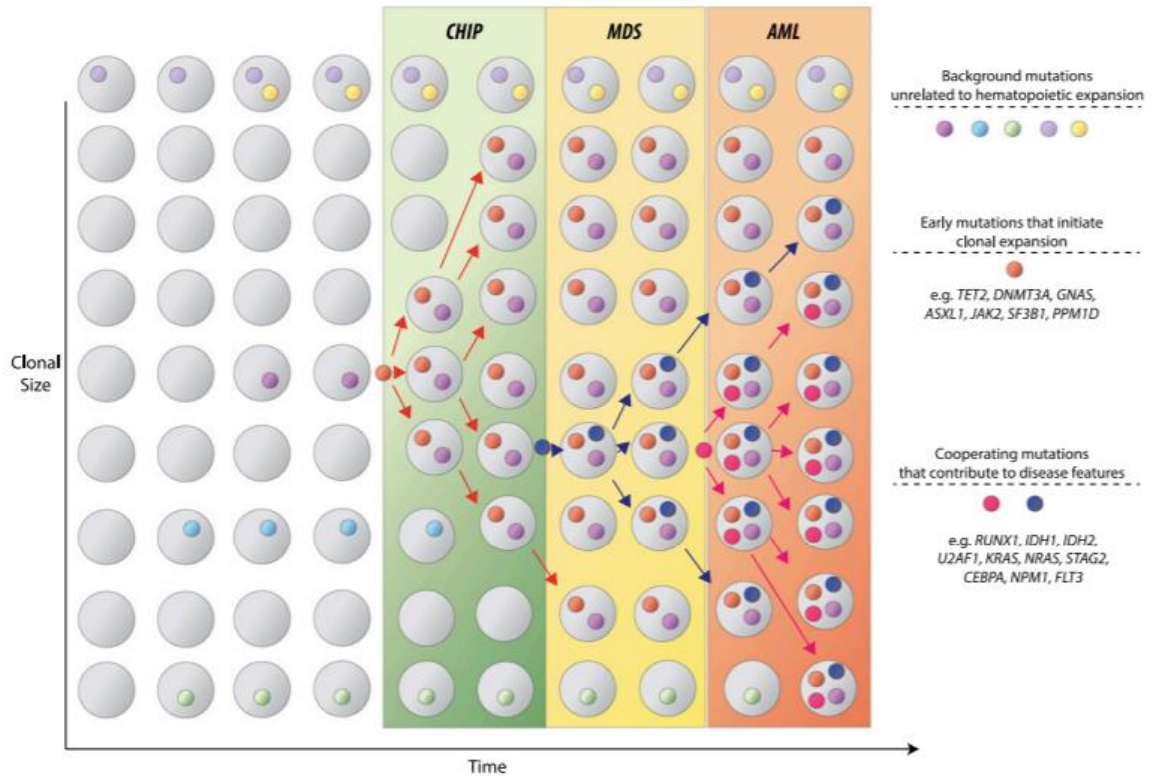


Figure 8 : La CHIP précurseur des néoplasie hématologiques (Image de Steensma et al., 2015 (137))

Représentation de l'évolution dans le temps d'un clone hématopoïétique vers une néoplasie myéloïde avec acquisition de mutations initiatrices ou « driver » de couleur saumon initiant l'expansion du clone, des mutations acquises plus tardivement et favorisant la survie et l'expansion du clone en rose et bleu tandis que les autres mutations sont des mutations passagères ne favorisant pas le clone.

En 2014, trois études importantes ont permis une compréhension de l'HC en utilisant des données de séquençage de nouvelle génération issues d'études visant à répondre à d'autres questions (138). Deux de ces études portaient sur le séquençage d'exome visant à identifier les variants constitutionnels augmentant le risque de diabète (139) et de schizophrénie (140), tandis qu'une troisième incluait des échantillons de sang de contrôle associés à un séquençage d'exome de tumeurs solides réalisé pour le TCGA (141). Chaque étude comprenait ainsi des milliers de sujets. Les gènes principalement concernés par l'HC de potentiel indéterminé sont ainsi représentés dans le **tableau 2**.

Nom des amorces	Séquence (+ ajout d'une queue)	Taille de l'amorce
BRCA1_Snap_898	GCTTTTATTACAGAATTCAGCCTTTT	71 pb
TP_53_Snap_c.321	ACCTACCAGGGCAGCTA	57 pb
TP_53_Snap_c.406	CCTCAACAAGATGTTTTGC	59 pb
TP_53_Snap_c.421	CCAACTGGCCAAGACC	56 pb
TP_53_Snap_c.427	ATCAACCCACAGCTGCA	57 pb
TP_53_Snap_c.455	GGGTTGATTCCACACCCC	58 pb
TP_53_Snap_c.470_471	GTAGATGGCCATGGCGCG	58 pb
TP_53_Snap_c.517	CAGCACATGACGGAGGTT	58 pb
TP_53_Snap_c.526	GGAGGTTGTGAGGCGC	56 pb
TP_53_Snap_c.577	TCTGGCCCCCTCCTCAG	56 pb
TP_53_Snap_c.578	CTGGCCCCCTCCTCAGC	56 pb
TP_53_Snap_c.589	ACACGCAAATTCCTTCCA	59 pb
TP_53_Snap_c.659	ATAGTGTGGTGGTGCCCT	58 pb
TP_53_Snap_c.711	CACCATCCACTACAACACTACAT	61 pb
TP_53_Snap_c.715	ATCCACTACAACACTACATGTGT	61 pb
TP_53_Snap_c.742	GGTGAGGATGGGCCTCC	57 pb
TP_53_Snap_c.814	GACGGAACAGCTTTGAG	57 pb
TP_53_Snap_c.817	CAGGACAGGCACAAACAC	58 pb
TP_53_Snap_c.818	CCAGGACAGGCACAAACA	58 pb
TP_53_Snap_c.839	CTCTGTGCGCCGGTCT	56 pb
TP_53_Snap_c.844	GCCTGTCCTGGGAGAGAC	58 pb
PTEN_Snap_c.601	GTTGTTTCACAAGATGATGTTT	62 pb
PTEN_Snap_c.698	CCAATTCAGGACCCACAC	58 pb

Tableau 6 : Amorces d'extension utilisées pour la technique Snapshot

Une queue comportant la répétition de la séquence cgtagatag qui est une séquence de virus a été ajoutée pour les amorces des variants *PTEN* et *TP53* dans le but d'augmenter la taille du produit d'extension afin qu'il ne migre pas en début de capillaire et soit de moins bonne qualité. Elle a été utilisée par l'équipe de Renaux Petel et al., 2017 (94). Pour le variant *BRCA1*, il s'agissait d'une queue polyA (x45).

Les produits d'extension ont été traités avec 1 U de SAP (« Shrimp Alkaline Phosphatase ») à 37°C pendant 60 minutes et 75°C pendant 15 min. 0,5µl du produit d'extension ont ensuite été dilués dans 10 µl d'Hi-Di (« highly de-ionized ») formamide et 0,3 µl de GeneScan-120LIZ size standard (Applied Biosystems®), le marqueur de taille. L'étape de dénaturation s'est faite à 95°C pendant 5 min avant analyse sur le séquenceur ABI 3130 Genetic Analyzer (Applied Biosystems®) sur un capillaire de 36 cm et avec un polymère POP-7. L'analyse des résultats a été effectuée avec le logiciel GeneMarker V.2.2.0. Les différents fluorochromes utilisés pour le marquage des ddNTP sont décrits dans le **tableau 7**.

ddNTP	Fluorochrome	Couleur du pic obtenu sur l'électrophorégramme
A	dR6G	Vert
C	dTAMRA™	Noir
G	dR110	Bleu
T	dROX™	Rouge

Tableau 7 : Différents fluorochromes utilisés pour le marquage des ddNTP pour la technique Snapshot

8. Analyse statistique

Dans un échantillon de cas index diagnostiqués avec un cancer du sein ou de l'ovaire, des associations entre la présence de variants délétères à faible fréquence allélique (hors mosaïque et hors ADNtc documentés) et différents facteurs ont été recherchées. Ces associations ont été mesurées à l'aide du test non paramétrique de Wilcoxon-Mann-Whitney pour les facteurs quantitatifs et du test exact de Fisher pour les facteurs qualitatifs. Le seuil de significativité retenu était de 5 %.

Etant donné les corrélations entre ces différents facteurs et afin d'estimer leurs effets propres, une analyse en régression logistique avec sélection backward des variables a été réalisée. Cette méthode permet de mettre en évidence des effets propres aux facteurs. Le modèle final présente donc les effets de chacun des facteurs de l'âge présentés sous forme d'OR avec intervalles de confiance à 95 %.

RESULTATS

Au total 2007 patients ont bénéficié d'un panel HBOC par séquençage à haut débit à partir d'un prélèvement sanguin au CFB (**Tableau 8**). Vingt patients étaient porteurs d'un variant potentiellement délétère avec une VAF comprise entre 3 % et 35 %. Parmi les 12 gènes du panel, 20 variants de *TP53*, 1 variant de *BRCA1* et 2 variants de *PTEN* (dont 1 variant de signification inconnue (VSI)) ont été identifiés (**Tableau n°8**). Deux patients étaient porteurs de 2 variants ; La **patiente n°2** possédait à la fois un variant de *TP53* faux sens de classe 5, le variant c.742C>T, p.Arg248Trp et, un variant non-sens de *PTEN*, le variant c.601G>T, p.Glu201*. La **patiente n°6** portait 2 variants faux sens de *TP53*, le variant c.578A>G, p.His193Arg et le variant c.589G>T, p.Val197Leu. Ainsi, en tout, 23 variants ont été analysés. Seul le **patient n°18** qui a présenté un cancer du sein à 82 ans était un homme, les autres patientes étaient des femmes. Un résumé des analyses effectuées pour chaque patient est détaillé dans le **tableau en annexe 1**.

N°	Age	P + âge	Gène	c., p. selon HGVS	VAF (%)		% cellules tumorales tissu	Hypothèse	Conclusion
					(Nbre reads Mut+/Tot)	(Nbre read Mut+/Total)			
1	50	Ov 49	BRCA1	c.898G>T, p.Glu300*	17 (66/388)	Tt : 86 (12/14) Tnt (foie) : 28 (5/18)	Tt : 90 % Tnt(foie) : 0%	ADNtc Mosaïque	Mosaïque
				c.601G>T, p.Glu201*	32 (102/319)	23 (180/796)	40%		
2	66	Colon 66	PTEN TP53	c.742C>T, p.Arg248Trp	27 (69/255)	58 (515/911)		HC ADNtc	ADNtc
				c.455C>T, p.Pro152Leu	8 (13/154)	0 (4/815)	60%	HC ADNtc	
3	71	S 61/70 Ov 70	TP53	c.470_471delinsAT, p.Val157Asp	23 (84/266)	3 (62/1901)	2%	HC ADNtc	HC
4	59	Ov 59	TP53	c.517G>A, p.Val173Met	4 (9/202)	Tt : 0 (3/1028) Tnt : 0 (1417)	Tt: 45 % Tnt: 0 %	HC ADNtc mosaïque	HC
				c.578A>G, p.His193Arg	3 (8/244)	0 (0/1118)	10%		
5	66	S 40-65	TP53	c.589G>T, p.Val197Leu	14 (34/251)	0 (0/1118)	10%	HC ADNtc	HC
				c.659A>G, p.Tyr220Cys	18 (80/435)	Tt : 0 (7/501) Tnt : 0 (8/744)	Tt : 50 % Tnt: 0%	HC ADNtc	
6	78	Ov 77	TP53	c.711G>A, p.Met237Ile	6 (13/233)	0 (1/697)	40%	HC ADNtc	HC
7	54	S 48 TN	TP53	c.817C>A, p.Arg273Ser	28 (88/318)	0 (16/825)	75%	HC ADNtc	HC
8	49	Ov 48	TP53	c.818G>A, p.Arg273His	7 (16/222)	0 (4/159)	80%	HC ADNtc	HC
9	63	Ov 64	TP53	c.844C>T, p.Arg282Trp	4 (10/276)	0 (0/1509)	60%	HC ADNtc	HC
10	73	Ov 65	TP53	c.526T>A, p.Cys176Ser	24 (81/334)	NA	NA	HC ADNtc mosaïque	HC Mosaïque
11	60	Ov 60	TP53	c.839G>T, p.Arg280Ile	19 (36/193)	NA	NA	HC ADNtc	HC
12	81	S 64 RH+	TP53	c.321C>A, p.Tyr107*	5 (19/415)	NA	NA	HC Mosaïque	HC ?
13	59	Ov 35 - 54 (en rechute)	TP53	c.715A>G, p.Asn239Asp	5 (12/251)	NA	NA	HC	HC
14	69	S 37/37 RI	TP53	c.406C>G, p.Gln136Glu	8 (22/283)	NA	NA	HC Mosaïque	HC Mosaïque
15	70	Ov 69	TP53	c.427G>A, p.Val143Met	12 (50/414)	NA	NA	HC ADNtc	HC
16	84	Ov 84	PTEN	c.698G>T, p.Arg233Leu	3 (13/382)	0 (3/447)	80%	HC Mosaïque	HC ?
17	82	Ov 82	TP53	c.421T>G, p.Cys141Gly	3 (14/440)	NA	NA	HC ADNtc	HC
18	83	S 82	TP53	c.577C>T, p.His193Tyr	4 (18/431)	NA	NA	HC ADNtc	HC
19	60	Ov 53	TP53	c.814G>A, p.Val272Met	4 (12/324)	NA	NA	HC ADNtc	HC

Tableau 8 : Données phénotypiques des patients ayant un ou plusieurs variants de VAF faible avec les résultats des VAF obtenus par en NGS à partir des prélèvements sanguins, des frottis juxtaux et des tissus tumoraux et/ sains.

Abréviations : ADNtc : ADN tumoral circulant ; HC : Hématopoïèse clonale ; Mut + : allèle muté ; NA : non adapté ; NR : non reçu ; Ov : Cancer de l’ovaire ; Nbre : nombre ; P : Phénotype ; S : Cancer du sein ; VAF : « Variant allele Fraction » ; WT : « Wild type » = allèle sauvage

Un seul variant pathogène non-sens de *BRCA1* a été trouvé chez la **patiente n°1 (Tableau 8)** atteinte d'un cancer de l'ovaire à 49 ans sporadique sur le plan familial donc non évocateur d'une prédisposition héréditaire. Cette patiente avait déjà bénéficié d'une chirurgie et d'une chimiothérapie et était en cours de traitement d'entretien par Avastin® au moment du prélèvement. La maladie était stabilisée au moment du prélèvement. Le variant dont elle était porteuse, *BRCA1* : c.898G>T, p.Glu300*, avait une VAF à 17 % dans l'ADN extrait du prélèvement sanguin, une VAF à 86% dans l'ADN issu du bloc tumoral ovarien composé de 90 % de cellules tumorales et une VAF à 28 % dans l'ADN issue d'une biopsie de foie exempte de cellules tumorales. Les résultats de séquençage du prélèvement tumoral ovarien montrent une perte d'hétérozygotie dans la tumeur avec un enrichissement du variant par rapport au prélèvement sanguin. Ce résultat est en faveur d'une mosaïque constitutionnelle ou de la détection d'ADNtc avec un clone dominant dans la tumeur qui serait détecté au niveau du prélèvement sanguin.

La patiente n°2 (Tableau 8) porteuse d'un variant de *PTEN* et d'un variant de *TP53* était en cours de traitement pour un cancer du côlon. La maladie était évolutive au moment du prélèvement. Le variant *PTEN* était détecté à 32 % et le variant *TP53* à 27 % à partir du prélèvement sanguin. Le prélèvement à partir du bloc inclus en paraffine a montré un enrichissement du variant *TP53* avec une fréquence allélique à 58 % mais le variant *PTEN* avait une VAF de seulement 23 %. Le variant de *TP53* montrait ainsi une potentielle LOH au niveau tumoral car sur 40 % de cellules tumorales, on retrouvait 58 % de variant *TP53* (**Tableau 8**). Au stade de l'analyse par NGS uniquement, seule l'hypothèse de l'HC pouvait être éliminée car, même s'il est possible de détecter un clone hématopoïétique au niveau tumoral du fait de l'infiltration de la tumeur par des cellules lymphocytaires, les taux détectés devraient être théoriquement inférieurs à ceux détectés dans le sang. Ainsi, les deux hypothèses retenues pour le deuxième patient à ce stade étaient la détection d'ADNtc ou une mosaïque constitutionnelle.

Les 18 autres patients porteurs d'un variant probablement délétère ou potentiellement délétère de *TP53* avaient présenté un cancer de l'ovaire pour 11 d'entre eux et d'un cancer du sein pour les 6 autres (**Tableau 8**). Une patiente avait présenté à la fois un cancer du sein et un cancer de l'ovaire. Il s'agissait pour la très grande majorité de variants faux sens prédits non fonctionnels dans la base de données IARC. Seul le variant c.589G>T, p.Val197Leu avait montré un effet partiellement fonctionnel par étude *in vitro*. Les VAF des variants lus étaient compris entre 3% et 28%. Concernant les 9 patients (**patients n° 3, 4, 5, 6, 7, 8, 9, 10 et 11**)

pour lesquelles la tumeur a pu être analysée, la VAF était plus faible au niveau tumoral qu'au niveau du prélèvement sanguin. Parfois, le variant n'était pas vu du tout au niveau tumoral comme pour le **patiente n°6** et la **patiente n°11 (Tableau 8)**. Ces patientes avaient été diagnostiquées de leur cancer dans l'année. L'hypothèse la plus probable chez ces dernières est que ces variants signent la présence d'un clone hématopoïétique étant donné que le variant n'est pas enrichi au niveau tumoral, ce que l'on pourrait attendre dans le cas d'une mosaïque ou d'un clone d'ADNtc détecté au niveau sanguin. Concernant les 9 autres patients (**patiente n° 12, 13, 14, 15, 16, 17, 19, 20 et 21**), la tumeur n'a pas pu être analysée et il n'est pas possible de conclure avec certitude entre les trois hypothèses (**Tableau 8**).

La **patiente n°18**, qui a présenté un cancer de l'ovaire à 84 ans est porteuse d'un variant de *PTEN*, c.698G>T, p.Arg233Leu de signification indéterminé (classe 3). Le variant était détecté sur 3/447 *reads* au niveau tumoral ce qui prouve que le variant n'est pas initiateur dans le développement tumoral de la patiente.

Une analyse Sanger a été réalisée pour l'ensemble des 9 patientes dont les variants avaient un ratio allélique supérieur à 10 % soit au-dessus du seuil de détection de la technique (**Figure 10**). Pour 8 patients, une analyse à la fois du sang et des cellules jugales a été réalisée tandis que pour la **patiente n°13** seul le sang était disponible au laboratoire. Les variants des **patients 1, 4, 6, 7,9, 12, 17** analysés à partir des prélèvements sanguins ont été confirmés sur le Sanger et avaient une intensité de pic relativement comparables aux valeurs des VAF trouvées en NGS (**Tableau 8** et **Figure 10**). Concernant les frottis jugaux, les intensités des pics étaient globalement moins fortes que ceux du sang pour l'ensemble des échantillons sauf pour le **patient 1** chez lequel on observe un pic de taille importante dans le frottis (**Figure 10**). Il s'agit de la patiente chez laquelle nous suspicions une mosaïque constitutionnelle ou la détection d'ADNtc. Ce résultat conforte l'hypothèse de la mosaïque constitutionnelle qui toucherait aussi les cellules jugales. Pour les autres patientes, ces analyses ont confirmé la présence du variant détecté en NGS et le frottis buccal assure du caractère non hétérozygote de certains variants et élimine un biais technique de séquençage qui induirait une diminution de la VAF.

Seul les variants des gènes *TP53* et *PTEN* du **patient n°2** n'étaient pas confirmés sur le Sanger réalisé à partir du prélèvement sanguin et du frottis buccal alors que les valeurs des VAF détectés en NGS à partir des prélèvements sanguins étaient respectivement de 32 % et 27 % (**Figure 10**).

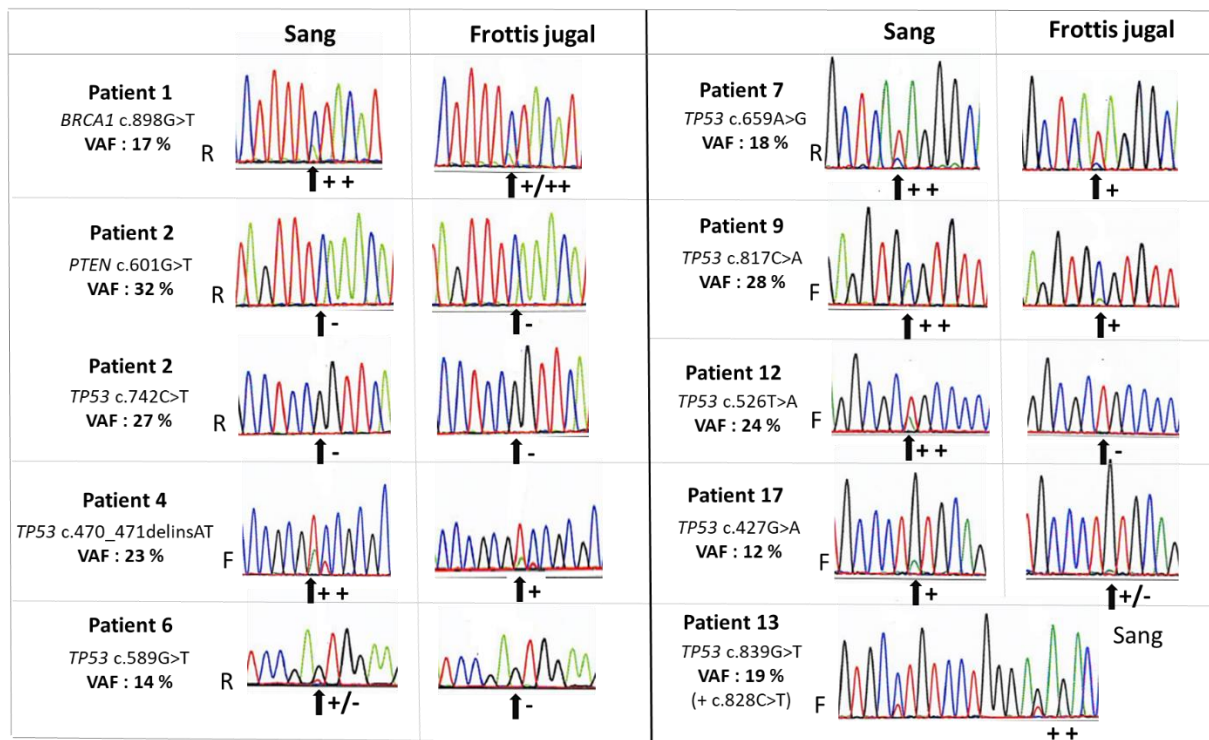


Figure 10 : Vérification des variants de ratio allélique supérieur à 10 % par technique Sanger

Abréviation : VAF : « Variant Allele Fraction »

Les variants ayant une valeur de VAF supérieure à 10 % ont été confirmés par technique Sanger à partir du prélèvement sanguin et d'un frottis jugal sauf pour l'échantillon du **patient n°2** pour lequel les variants de *PTEN* et *TP53* n'ont pas été visualisés sur le Sanger alors que les données de VAF étaient respectivement de 32 % et 27 %. Pour le **patient n°13**, seul l'échantillon sanguin était disponible au laboratoire.

Les flèches noires montrent la localisation du variant recherché. Les sigles ++, ++/+, +, +/- et - représentent de manière décroissante et relative, l'intensité du pic muté détecté par rapport au pic sauvage. Le sigle ++ signifie que l'allèle muté est bien visible sur l'électrophorégramme tandis que le sigle - l'absence de pic correspondant à l'allèle muté.

Dans le but de d'améliorer limite de détection et d'analyser les prélèvements de cellules jugales, l'ensemble des échantillons des patients ont été analysés par Snapshot (sang +/- cellules jugales +/- tissu en paraffine). Cependant, seuls les variants avec un ratio allélique supérieur à 5 % ont été confirmé par cette technique car les variants de VAF inférieure étaient trop proches du bruit du fond pour s'assurer de la spécificité des pics détectés (**Figures 11, 12 et 13**). Concernant le **patient n°17**, le variant n'a pas pu être confirmé à cause d'un problème technique et non par manque de sensibilité.

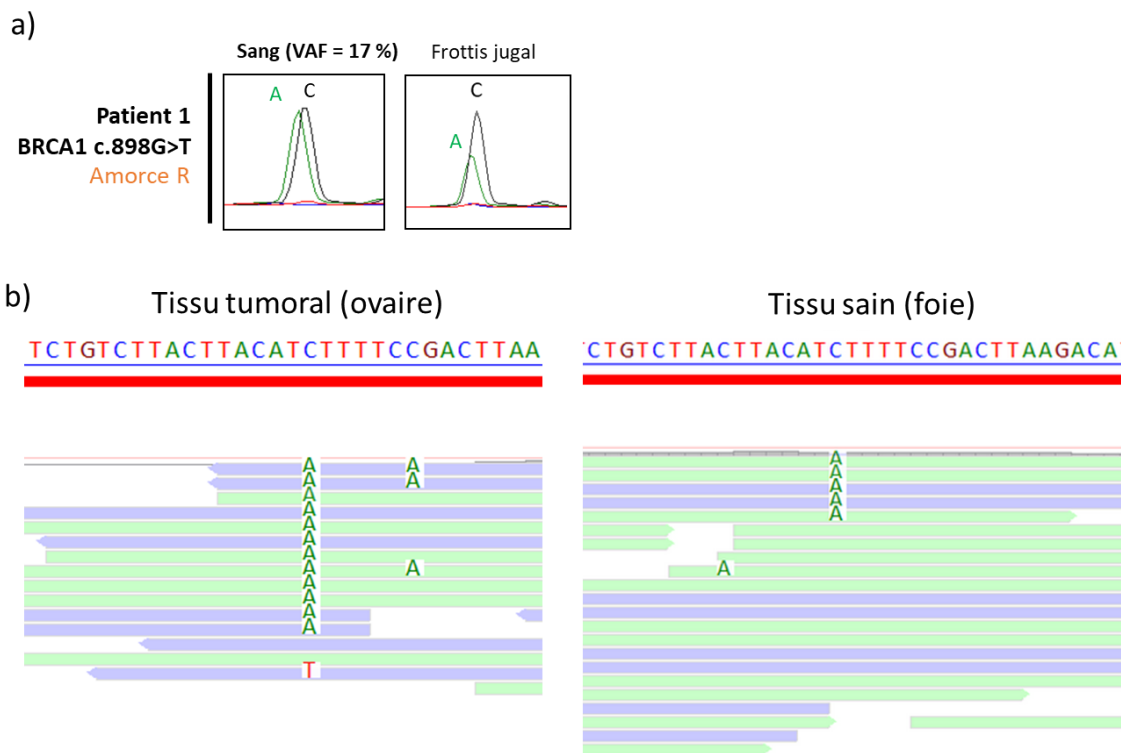


Figure 11 : Résultats des analyses Snapshot et de NGS réalisés chez le patient n°1 à partir du prélèvement sanguin, des cellules jugales et de la tumeur.

- a) Le variant de *BRCA1* lu en reverse a été confirmé par Snapshot à l'aide du Snapshot® Multiplex kit (Applied Biosystem®) à partir du prélèvement sanguin et des cellules jugales.
- b) Visualisation des BAM sur Alamut Visual issus de l'analyse NGS de l'ADN extrait des prélèvements de tissu tumoral et de tissu sain.

La technique Snapshot a confirmé la présence du variant sur le prélèvement du **patient n°1 (Figure 11)**. L'analyse a montré que le pic correspondant au variant du frottis jugal était un peu moins intense que le variant détecté dans le prélèvement sanguin. Cependant, le pic représentant l'allèle sauvage était encore important dans les cellules jugales confortant l'hypothèse de la mosaïque atteignant les cellules épithéliales jugales. L'ADN tumoral a aussi été analysé par Snapshot mais ce dernier étant faiblement amplifiable, aucun signal n'a pu être détecté. Nous n'avons malheureusement pas pu vérifier la perte d'hétérozygotie identifiée en NGS par la méthode Snapshot.

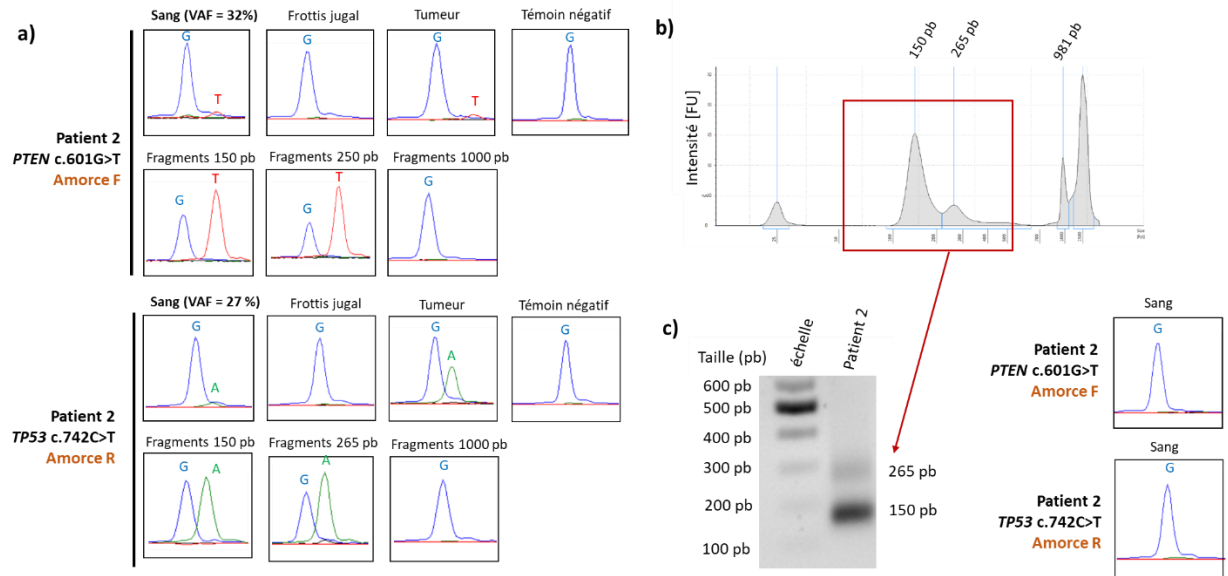


Figure 12 : Données obtenues pour le patients n°2 par analyse Snapshot et Tapstation

- Images des électrophorégrammes obtenus après analyse Snapshot à l'aide du Snapshot® Multiplex kit (Applied Biosystem®) des différents prélèvements disponibles au laboratoire du **patient n°2** soit l'ADN tumoral, l'ADN issu du prélèvement sanguin et l'ADN extrait à partir du bloc tumoral. Le témoin négatif est de l'ADN issu d'un patient non porteur du variant. Les fragments d'ADN de différentes tailles (150 pb, 265 pb et > 1000 pb) ont pu être analysés séparément par migration de l'ADN issu du prélèvement sanguin sur un gel d'agarose puis découpage des bandes et purifications de ces dernière à l'aide du kit NucleoSpin® Gel PCR Cleanup (Macherey-Nagel). Enfin, l'analyse a été réalisée avec l'utilisation des amorces en Sanger permettant de séquencer uniquement les fragments de taille supérieure à 500 pb.
- Image issue de l'analyse par TapeStation de l'ADN extrait du sang du **patient n°2**
- Photographie partielle du gel d'agarose à 1 % obtenu après migration de l'ADN extrait du sang du **patient n°2** juste avant le découpage et la purification des bandes de tailles différentes.
- Résultats de l'analyse Snapshot dont l'amplification a été réalisée avec des amorces produisant des amplicons de taille supérieure à 250 pb (les mêmes amorces que celles utilisées pour l'étape d'amplification avant le séquençage Sanger). Le pic correspondant au variant muté est absent.

Pour le **patient n°2** chez lequel nous suspicions la détection d'ADNtc, une analyse en Snapshot a aussi été réalisée. Dans le sang, les pics correspondant aux allèles mutés étaient visibles contrairement après technique Sanger (**Figure 10**) mais faibles par rapport aux VAF trouvée en NGS. Sur les ADN issus des cellules jugales, les pics correspondant aux allèles mutés n'étaient plus détectables (**Figure 12**). Par ailleurs, les ADN tumoraux ont montré un pic semblable à celui du sang pour le variant *PTEN* et un pic enrichi pour le variant de *TP53*. Ces données confirment les données obtenues par séquençage à haut débit. Concernant les échantillons d'ADN issus de la purification des fragments de taille de 150 pb et 265 pb après migration sur gel d'agarose, le pic correspondant à l'allèle muté obtenu est supérieur à celui du pic de l'allèle sauvage pour les 2 variants (**Figure 12a, b et c**). De plus, les fragments de tailles

environ égale à 1000 pb ont aussi été analysés et ont montré une absence de pic correspondant à l'allèle muté confirmant la présence des deux variants uniquement sur les fragments de petite taille (**Figure 12 a, b et c**). Enfin, le Snapshot réalisé avec une autre PCR utilisant des amorces plus éloignées l'une de l'autre, et permettant l'amplification de fragments d'ADN de taille supérieure à 250 pb (amorces utilisées pour le séquençage en Sanger) a confirmé, de manière indirecte, l'absence du variant muté dans les grands fragments et expliqué la négativité de l'analyse Sanger réalisée en amont (**Figure 12d**). Ces résultats combinés démontrent que les 2 variants sont uniquement dans les fragments de petites tailles. Les variants de *PTEN* et *TP53* du **patient n°2** détectés par séquençage à haut débit sont donc très probablement liés à la détection d'ADNtc.

L'ensemble des échantillons ont été analysés en TapeStation après extraction de l'ADN issu des prélèvements sanguins afin de vérifier la taille des fragments obtenus pour voir si d'autres échantillons présentaient des fragments de petite taille en faveur d'ADNtc, comme le patient n°2 (**Figure 13**) afin de s'assurer de l'absence de bande de petite taille pouvant correspondre à de l'ADNtc. Contrairement au **patient n°2**, qui présente deux bandes d'intensité importante à 150 pb principalement et 265 pb, les autres échantillons ne présentent pas le même profil. Les échantillons **des patients 5, 6, 8, 14, 15, 16 et 17** montrent cependant plusieurs bandes de tailles différentes en échelle (**Figure 13**). Ces échantillons correspondent aux échantillons les plus concentrés en ADN. Pour ces derniers, le fragment de plus faible taille se situe autour de 180 pb et les autres fragments ont des tailles correspondant à des multiples de 2 de 180 pb. Il pourrait s'agir d'ADN dégradé.

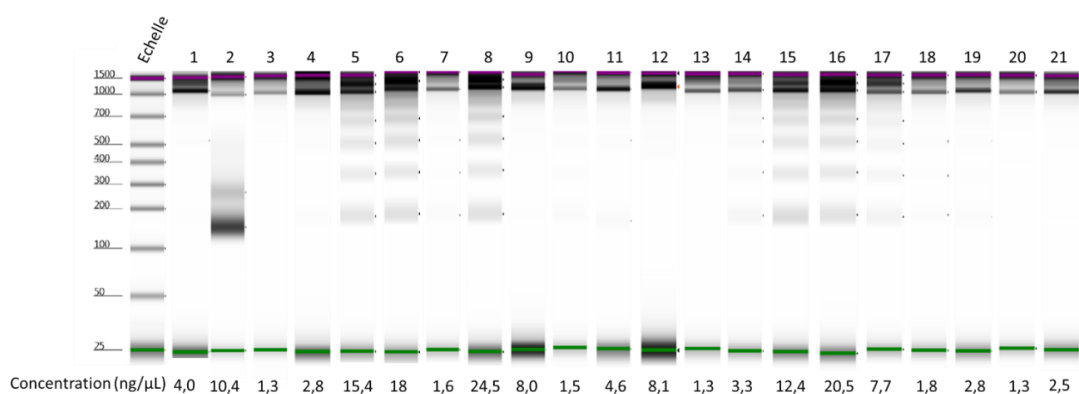
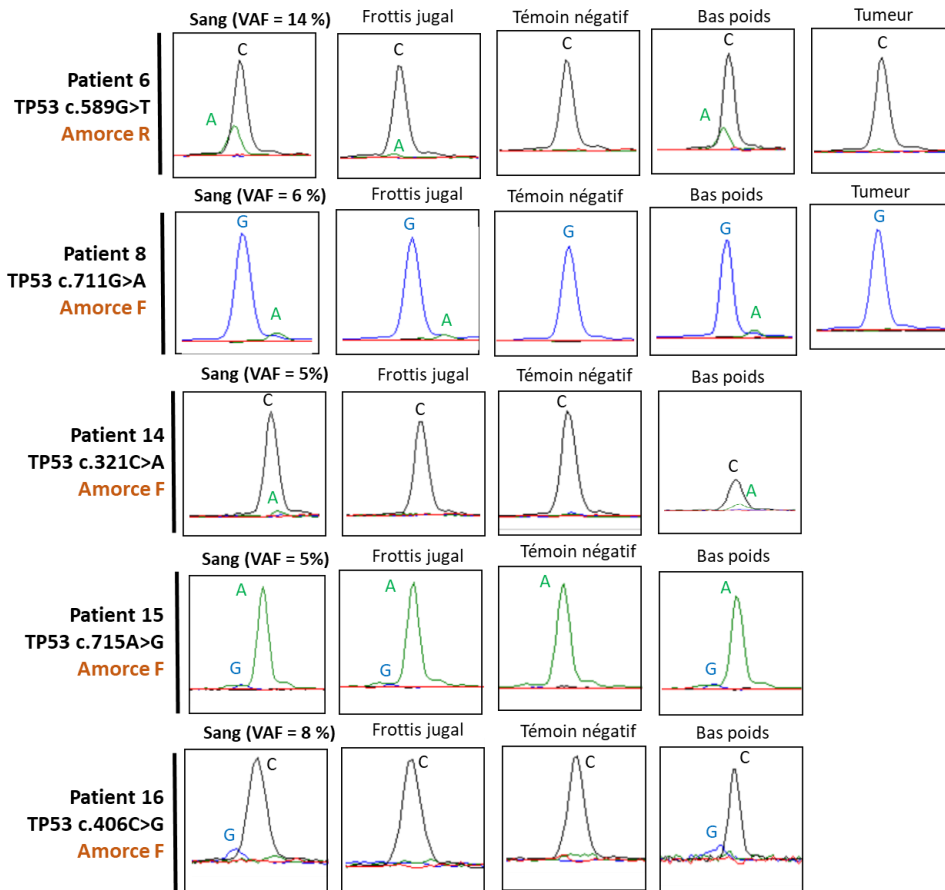


Figure 13 : Analyse par TapeStation de la taille des fragments des ADN extraits à partir des prélèvements sanguins

Image du gel obtenu par analyse TapeStation des ADN extraits à partir des prélèvements sanguins des 21 patients pour lesquels une détection de variant de faible ratio allélique a été détecté par séquençage à haut débit. En bas est représenté la concentration en ng/μL mesurée en TapeStation.

La technique Snapshot a été réalisée sur l'ensemble des échantillons disponibles des patients. Ils ont permis de confirmer la présence des variants de VAF supérieure à 5 % détectés par NGS. Les variants des **patients n°5, 11, 18, 19, 20 et 21** n'ont pas pu être confirmés par cette technique non plus. Pour **les patients 6, 8, 14, 15 et 16** pour lesquels des fragments de petites tailles, en escalier, étaient observés en TapeStation (**Figure 14**), les fragments de bas poids moléculaires (taille inférieure à 500 pb) ont pu être analysés après migration sur un gel d'agarose et découpage de bande. Pour l'ensemble des 5 échantillons analysés (**Figure 14a**), les fragments de petite taille avaient une différence d'intensité de pic allèle muté / allèle sauvage superposable à celui de l'ADN total issu du prélèvement sanguin. Cela confirme que le variant détecté dans le sang n'est pas enrichi dans les fragments de petites tailles contrairement à l'échantillon de **patient n°2 (Figure 12)**. Au niveau des cellules jugales, les pics correspondant à l'allèle muté étaient en majorité inférieurs à ceux détectés dans le sang (**Figure 14a et b**) sauf pour **les patients n°4 et n°10 (Figure 14b)**. Cependant, pour ces deux patients, l'ADN tumoral a pu être analysé et a montré l'absence de pic correspondant à l'allèle muté ce qui n'est pas en faveur d'une mosaïque. Concernant les 7 ADN tumoraux analysés par Snapshot, la technique confirme l'absence d'enrichissement du variant au niveau tumoral. Pour les patients n°3, 4, 5, 6, 7, 8, 9 et 10, pour lesquels on a pu analyser la tumeur, on peut attribuer sans ambiguïté le variant de faible VAF à la détection d'un clone hématopoïétique (**Tableau 8 et Figure 14**).

a)



b)

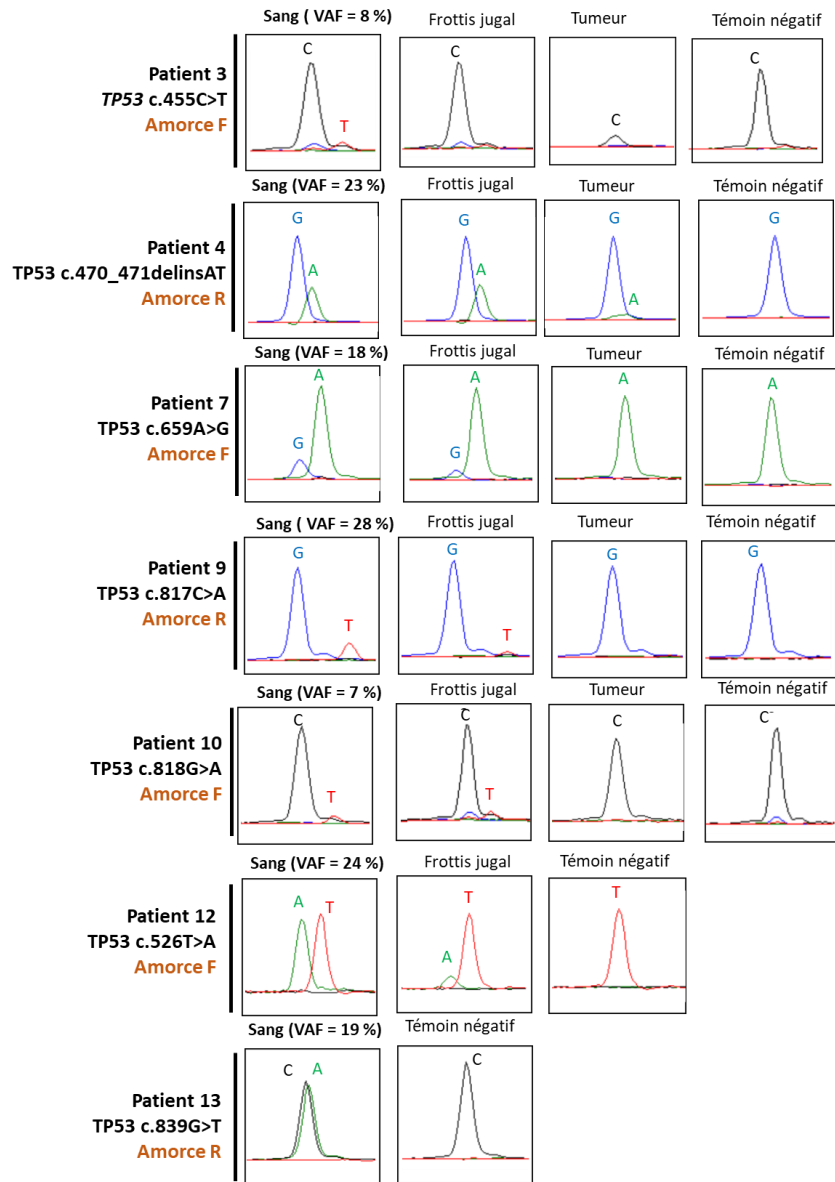


Figure 14 : Résultats par analyse Snapshot de l'ensemble des échantillons des patients ayant une VAF supérieure à 5 % et inférieure à 35 %

Les patients n°1 et n°2 ne sont pas présentés ici (cf Figure 11 et 12)

L'ensemble des échantillons disponibles au laboratoire des patients ont été analysés par Snapshot grâce Snapshot® Multiplex kit (Applied Biosystem®). Au mieux nous avons : le prélèvement sanguin, le frottis buccal et le bloc tumoral. Le témoin négatif correspond à l'ADN d'un patient non porteur du variant. L'amorce F correspond à l'amorce *Forward* et l'amorce R à l'amorce *Reverse*. Lorsque l'amorce *Reverse* est utilisée, les bases attendues pour les variants sauvages et mutés sont les complémentaires des bases notées à gauche. Par exemple, pour le patient n°6, le variant 589G>T devient C>A.

- Les patients 6, 8, 14, 15 et 16 ont pu bénéficier d'une analyse des fragments de petites tailles détectés (< 500 pb) issus des prélèvements sanguins par migration de l'ADN issu du prélèvement sanguin sur un gel d'agarose à 1 % puis découpage des bandes et purifications de ces dernières à l'aide du kit NucleoSpin® Gel PCR Cleanup (Macherey-Nagel).
- Les patients 3, 4, 7, 9, 10, 12 et 13 n'ont pas bénéficié de l'analyse des petits fragments car ils n'étaient pas détectés en TapeStation (Figure 13)

Afin d'évaluer les caractéristiques cliniques des patients ayant un variant pathogène de *TP53* avec une VAF faible, nous avons réalisé une analyse statistique des patients porteurs du variant par rapport aux non porteurs (**Tableau 9**). Pour l'analyse, nous avons retiré la **patiente n°1** porteuse de variant *BRCA1* à l'état constitutionnel, la **patiente n°2** porteuse de deux variants attribués à la détection d'ADNtc et la **patiente n°18** porteuse d'un variant de *PTEN* de classe 3.

Parmi les 2007 patients analysés, les cas indemnes de maladie ainsi que les cas index avec une tumeur différente du sein et de l'ovaire ont été exclues (**Figure 15**).

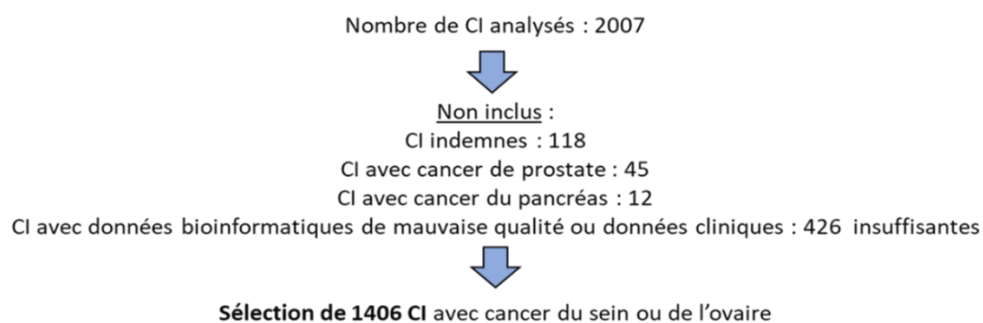


Figure 15 : Cas index sélectionnés pour l'analyse statistique cas/contrôles

Abréviation : CI : cas index

Nous avons montré que les patients porteurs d'un variant de *TP53* à faible VAF sont en moyenne significativement plus âgés au moment du prélèvement (**Tableau 9**). La moyenne est, en effet, de 55,7 ans pour le groupe contrôle contre 67,6 ans pour le groupe des porteurs ($p < 0,001$). Les patients porteurs d'un variant *TP53* présentent un diagnostic de cancer plus tardif que les patients non porteurs (50,2 vs 59,9 ; $p < 0,01$). Enfin, de manière très significative, la population porteuse d'un variant à faible VAF de *TP53* est statistiquement enrichie en cas de cancer de l'ovaire en comparaison de la population contrôle (310/1388) ($p < 0,001$) (**Tableau 9**).

Nous n'avons pas montré de différence entre les deux groupes concernant le délai entre le diagnostic du cancer et le prélèvement (**Tableau 9**). Enfin, nous avons observé que dans le groupe des patients pour lesquels nous avons détecté un variant de *TP53* délétère, la couverture moyenne du gène était plus importante (**Tableau 9**).

	Contrôle n = 1388	Patients avec variants de <i>TP53</i> VAF < 35 % n = 18	p-value*
Age au prélèvement	55.7 (13.4)	67.6 (9.97)	p < 0,001
Age au 1er diagnostic de cancer	50.2 (12.8)	59.9 (14.0)	p < 0.01
Délai entre le premier diagnostic et le prélèvement	5.4 (8.1)	7.5 (9.6)	NS
≤ 1 an	745	9	
2 à 3 ans	155	0	
≥ 4 ans	470	8	
Non renseigné	18	1	
Type de cancer			
Sein	1078 (78 %)	7 (39 %)	p < 0,001
Ovaire	310 (22 %)	11 (61 %)	
Profondeur de couverture moyenne sur <i>TP53</i>	213 (76)	262 (70)	p < 0.01

* Test non paramétrique de Wilcoxon pour les variables quantitatives, test exact de Fisher pour les variables qualitatives

Tableau 9 : Caractéristiques des patients ayant des variants *TP53* de faible VAF

Ayant noté une forte corrélation entre l'âge au diagnostic et le type de cancer avec un âge moyen de 46,9 ans ($\pm 11,1$) ans pour le cancer du sein et de 61,9 ans ($\pm 11,4$) pour le cancer de l'ovaire, nous avons souhaité voir si les deux variables étaient différentes dans les deux groupes indépendamment l'une de l'autre. Ainsi, l'analyse multivariée prenant en compte trois facteurs confondants importants soit le type de cancer et la profondeur de couverture moyenne du gène *TP53* et l'âge au prélèvement montre que l'âge au prélèvement est toujours plus élevé dans le groupe des patients porteurs d'un variant *TP53* de manière significative (OR=1,79 ; $p < 0,01$) (**Tableau 10**). De même, le groupe des porteurs présente plus de cancers de l'ovaire que le groupe contrôle indépendamment de l'âge et de la moyenne de profondeur de couverture du gène (OR = 3,28 ; $p = 0,02$) (**Tableau 10**).

	Odd ratio	Intervalle de confiance	p-value
Age au prélèvement (OR par tranche de 10 années)	1.79	[1.15 ; 2.78]	p = 0.01
Type de cancer Sein	--		
Ovaire	3.28	[1.20 ; 8.96]	p = 0.02
Profondeur de couverture moyenne sur TP53 (OR par tranche de 100 reads)	2.10	[1.19 ; 3.71]	p = 0.01

Modèle de régression logistique avec sélection backward des variables (les variables d'âge au diagnostic et de délai sont sorties du modèle)

Tableau 10 : Evaluation des effets conjoints de l'âge au prélèvement, type de cancer et de la profondeur de couverture moyenne sur TP53 dans une analyse multivariée entre le groupe des patients porteurs d'un variant de TP53

DISCUSSION

Parmi les 2007 cas index ayant bénéficié du panel de 12 gènes (*BRCA1*, *BRCA2*, *PALB2*, *RAD51C*, *RAD51D*, *MLH1*, *MSH2*, *MSH6*, *PMS2*, *CDH1*, *PTEN* et *TP53*), ont été identifiés 20 patients porteurs d'un variant de ratio allélique faible compris entre 3 et 32 % pathogène ou probablement pathogène et 1 patient porteur d'un variant de signification inconnu de *PTEN* (**Tableau 8**). La très grande majorité des variants détectés étaient localisés sur le gène *TP53* (20/23). Les autres variants se situaient sur le gène *PTEN* pour deux patientes et le gène *BRCA1* pour une patiente.

Cette série d'étude a mis en évidence une exceptionnelle mosaïque constitutionnelle de gène *BRCA1* détecté à 17 % dans le sang d'une patiente (**patiente n°1**) (**Figure 11**). Il s'agit du quatrième cas rapporté dans la littérature et le premier cas chez une patiente atteinte du cancer de l'ovaire. Le séquençage de la tumeur de la patiente a mis en évidence une perte d'hétérozygotie au locus correspondant confirmant que la tumeur de la patiente a bien été initiée dans un contexte de prédisposition au cancer du sein et de l'ovaire. Aussi, l'hypothèse d'une détection d'ADNtc a pu être affirmée par le fait que seuls des fragments de grande taille dans l'échantillon d'ADN ont été extraits du prélèvement sanguin. En effet, le variant était visible en Sanger amplifiant des fragments de 250 pb, ce qui prouve de manière indirecte que le variant était présent sur les grands fragments supérieurs à 250 pb. Par ailleurs, le variant a aussi été mis en évidence à partir du prélèvement de cellules jugales avec une fraction allélique semblant assez importante par la technique Snapshot et Sanger, confortant l'hypothèse de la présence de ce variant en mosaïque en testant un autre tissu normalement moins contaminé par le sang. D'autre part, la patiente était stabilisée au moment du prélèvement en cours de traitement d'entretien par Avastin®. Ce deuxième argument n'élimine pas l'hypothèse d'une détection d'ADNtc mais n'est pas en faveur d'une détection à un taux aussi élevé de 17 %. Enfin, la dernière hypothèse de l'HC n'était pas cohérente avec le fait que ce variant présente une fraction allélique élevée au niveau tumoral et par le fait que ce gène ne fait pas partie des gènes les plus concernés par l'HC (138).

L'étude a mis en évidence la détection d'ADNtc chez la **patiente n°2** atteinte d'un cancer du côlon. En effet, cette patiente était en cours de traitement de son cancer et n'était pas stabilisée au moment de son prélèvement sanguin. La patiente avait bénéficié d'un panel HBOC du fait de ses antécédents familiaux et non pour sa pathologie. Un variant pathogène de *PTEN* (c.601G>T, p.Glu201*) avec une VAF à 32 % et un variant pathogène de *TP53* (c.742C>T,

p.Arg248Trp) à 27 % (**Tableau 8 et Figure 12**). Puisque ces deux variants pathogènes étaient détectés uniquement dans les petits fragments de taille à 150 pb environs et 265 pb, l'hypothèse de l'origine tumorale de ce variant était forte. En effet, les fragments étaient bien visibles en TapeStation et l'analyse par Sanger et Snapshot ont permis de confirmer la détection des variants uniquement dans les petits fragments. L'analyse sur la tumeur était en faveur de la présence de l'enrichissement du variant du variant *TP53* dans la tumeur probablement présente à l'état hémizygote tandis que le variant de *PTEN* était plus faible mais probablement à l'état hétérozygote dans la tumeur puisque que le prélèvement extrait contenait seulement 40 % de cellules tumorales. L'analyse sur la tumeur n'était pas suffisante pour étayer l'une ou l'autre des hypothèses entre la mosaïque ou l'ADNtc. Seul le séquençage des fragments nous a permis clairement de conclure. Ces variants atteignaient des valeurs proches de 50 % ce qui peut amener à penser que des variants dus à la détection d'ADNtc pourrait faussement être pris pour des variants constitutionnels d'autant plus que pour certains patients, la confirmation sur un deuxième prélèvement se fait par la réalisation d'un deuxième prélèvement sanguin réalisé quasiment simultanément. Bettegoya et al. avait montré que la détection d'ADNtc était importante dans le cadre du cancer colorectal mais moins que pour l'ovaire (132). La détection des variants est donc probablement plus expliquée par le fait que le prélèvement a été effectué à un stade de relargage important au niveau sanguin d'ADNtc. Par ailleurs, les gènes *TP53* et *PTEN* sont très fréquemment mutés au niveau somatique dans le cancer du côlon. La manière de ne pas risquer de détecter l'ADNtc pour une analyse constitutionnelle serait de rajouter une étape après l'extraction qui consisterait à réaliser une sélection des fragments de taille supérieure à 500 pb de l'ADN par exemple.

A l'inverse, l'HC peut potentiellement conduire à détecter faussement au niveau de la tumeur des variants d'origine hématopoïétique pris pour des variants tumoraux. Pour le **patient n°4**, par exemple, le variant présent à 23 % dans le sang est détecté à 3 % dans la tumeur (**Tableau 8**). Si la tumeur avait été séquencée seule, on aurait pu penser à un sous clone tumoral ayant acquis la mutation. En ce qui concerne les mutations *TP53*, cela n'entraîne pas de conséquence thérapeutique directe chez les patientes mais cela pourrait en avoir pour d'autres gènes. Coombs et al avaient déjà mis en évidence cette observation en étudiant sur des tumeurs solides plusieurs gènes fréquemment impliqués dans l'HC tels que *DNMT3A*, *TET2*, *ASXL1*, *TP53*, *ATM*, *CHEK2*, *SF3B1*, *CBL*, *JAK2*. Ainsi, en utilisant des échantillons de sang appariés à la tumeur, ils ont montré que 8 % (18/226) des mutations observées dans ces gènes dans les tumeurs étaient de véritables événements d'HC dont 2 mutations *TP53*(153). Cela pourrait avoir

des conséquences en terme thérapeutique si un traitement ciblé est donné devant la positivité d'un variant pensé d'origine tumorale alors qu'il s'agit d'un clone hématopoïétique. A l'inverse, l'ADNtc peut être contaminé par l'HC car l'ADN circulant est issu en grande partie des cellules du sang périphérique (154). Dans tous les cas, il est difficile de conclure sur la signification des variants de VAF faible sans tester à la fois l'ADN issu d'un prélèvement sanguin et l'ADN issu du bloc tumoral, principalement pour les variants fréquemment rencontrés dans le contexte d'HC.

Pour l'ensemble les gènes *BRCA2*, *PALB2*, *RAD51C* et *RAD51D*, *CDHI* et les gènes MMR, l'analyse n'a pas mis en évidence variant potentiellement pathogène de VAF faible. Pour certains gènes, aucun cas de mosaïques n'a été rapporté dans la littérature (ex : *BRCA2*, *PALB2*, *RAD51C* et *RAD51D*). Pour les gènes autres que *BRCA2*, cela peut s'expliquer par le fait que l'incidence mutationnelle de ces gènes dans le cadre des panels HBOC est plus faible que celle de *BRCA1* et *BRCA2*), diminuant ainsi la probabilité de mettre en évidence une néomutation postzygotique ». Par exemple, la fréquence des mutations retrouvées dans le cadre de femmes ayant présenté un cancer du sein a été estimé à 0,80 % pour le gène *PALB2* et de 0,09 % et 0,07 % respectivement pour les gènes *RAD51C* et *RAD51D* (155). Pour le gène *CDHI*, il en est de même puisque la fréquence des variants délétères détectés à l'état hétérozygote dans le cadre de l'analyse par panel HBOC a été évaluée à 0,048 % selon les données du GGC rapportés par les laboratoires français. Il en était de même pour les gènes MMR (*MLH1*, *MSH2*, *MSH6* et *PMS2*) inclus dans le panel depuis peu du fait du surrisque de cancer de l'ovaire induit par ces gènes. En effet, peu d'hommes ou femmes étudiés dans ce cadre remplissent les critères de test pour le panel colon. La probabilité de trouver des mutations délétères de ces gènes dans la population étudiée est donc faible, même à l'état hétérozygote. Or, les mosaïques constitutionnelles représentent des événements encore plus rares. De plus, comme les mutations hétérozygotes de ces gènes n'ont pas d'impact sur la fécondité à l'état hétérozygote, l'apparition d'une néo-mutation n'est pas le mécanisme privilégié d'apparition du syndrome.

L'importance de détecter les variants de faible ratio allélique est avant tout de rejeter ou de confirmer la détection éventuelle d'une mosaïque. Sur cette grande série de patientes analysées, seule une mosaïque documentée et validée expérimentalement de *BRCA1* chez une patiente qui a présenté un cancer de l'ovaire à 49 ans a été mis en évidence. Cette tumeur d'apparition précoce est très évocatrice d'une prédisposition au cancer du sein et de l'ovaire due à une mutation constitutionnelle du *BRCA1*. En revanche, pour les autres patients pour

lesquels nous avons détecté des variants délétères de *TP53* ou *PTEN*, l'ensemble des tumeurs présentées ne rentrait pas dans le spectre du LFS ou de la maladie de Cowden. En effet, les patients ont présenté des cancers de l'ovaire ou des cancers du sein à des âges tardifs alors que les cancers de sein présents dans le LFS sont des cancers précoces. En revanche, la question de la présence d'un variant en mosaïque se pose pour la **patiente n°16** qui a présenté un cancer du sein bilatéral à 37 ans RH+ car les cancers du sein du LFS sont préférentiellement des cancers triples positifs c'est-à-dire RH+ HER2+ (156). Le statut HER2 n'était pas connu à l'époque où la patiente a présenté son cancer. Cette patiente est aujourd'hui âgée de 69 ans donc il est aussi très probable que cela puisse être de l'HC. Cependant, lorsque les blocs tumoraux ne sont plus accessibles, il serait intéressant de discuter en réunion de concertation pluridisciplinaire de l'intérêt d'effectuer une recherche ciblée chez les enfants de la patiente dans l'hypothèse d'une mosaïque. En résumé, il est important de connaître le contexte phénotypique du patient pour orienter la prise en charge en cas d'impossibilité de définir avec certitude l'origine du variant déséquilibré. Concernant les 10 autres patientes porteuses d'un variant de *TP53* à faible fréquence allélique et pour lesquelles la tumeur a été analysée, les séquençages ont montré une diminution franche de la VAF au niveau tumoral comparée aux VAF retrouvées par analyse du prélèvement sanguin. Cette observation est plutôt en faveur de l'hypothèse de l'HC. Dans le cas d'une mosaïque, le variant aurait dû être mieux représenté au niveau du tissu concerné par la tumeur. En effet, même si les prélèvements de blocs reçus n'étaient pas composés à 100 % de cellules tumorales, dans le cas d'une mosaïque, il est probable qu'il existe un continuum et que le tissu à proximité de la tumeur présente le variant avec un ratio allélique facilement détectable voire proche de 50 %. Cette « appauvrissement en cellules porteuses de la variation dans la tumeur n'est pas non plus en faveur de la présence d'ADNtc. Cependant, il a déjà été décrit que les variants détectés dans l'ADNtc pourraient ne pas correspondre à la tumeur primitive. Cette divergence pourrait être attribuée au moins en partie aux cellules tumorales circulantes ou à l'ADN libérée par des métastases (157). De plus, les analyses de la taille des variants en TapeStation n'ont pas montré d'enrichissement en variants de petites tailles contrairement au **patient n°2**. Finalement, l'HC est l'hypothèse retenue.

Enfin, concernant les patientes porteuses d'un variant pathogène ou probablement pathogène de *TP53* pour lesquelles le bloc tumoral n'était pas accessible, soit parce qu'il avait été détruit lorsqu'il datait de plus de 15 ans, par exemple, ou que nous n'avions pas eu de retour du prescripteur, il était difficile de conclure. Cependant, les analyses par TapeStation n'ont pas montré d'enrichissement en variant de petite taille peu en faveur d'ADNtc et les phénotypes

des patients en termes de contexte tumoral et d'âge étaient similaires aux patients pour lesquels précédemment analysés avec leur matériel tumoral Ainsi, il est très probable que la plupart des variants détectés soit aussi liés à la présence d'une l'HC sauf pour la patiente n°16 qui a présenté un cancer du sein bilatéral à 37 ans et pour laquelle on ne peut pas exclure l'hypothèse d'une mosaïque.

Une des limites de cette étude a été un manque de sensibilité des techniques de confirmation des variants pour ceux ayant un ratio allélique inférieur à 5 %. En effet, le séquençage par Sanger nous a permis de confirmer des variants de fréquence allélique supérieurs à 10 % et la technique Snapshot, nous a permis de confirmer la présence des variants de fréquence allélique supérieure à 5 %. Pour les variants inférieurs à 5 %, le pic correspondant à l'allèle muté était visible dans la grande majorité des cas mais trop proche du bruit de fond pour s'assurer de la spécificité de ce dernier. D'autres techniques existent pour descendre en limite de détection pour les variants de petite taille telle que la digitale PCR qui permet une quantification absolue des événements rares contrairement au Snapshot et au Sanger permettant une quantification relative. Cependant, la digital PCR nous aurait permis de confirmer la détection des variants de fréquence allélique faible dans le sang mais ne nous aurait pas aidés dans la démarche de compréhension du mécanisme physiopathologique. En effet, pour confirmer une mosaïque, l'objectif est de retrouver un variant à un état proche de l'hétérozygotie dans un tissu autre que le sang. Or, l'analyse Snapshot ou Sanger suffit dans ce cas.

Cette étude montre, par ailleurs, l'importance de réaliser un séquençage à forte profondeur, dans l'idéal supérieur à 300 X. En effet, l'analyse statistique a montré que le groupe des porteurs de variants de faible VAF avaient une profondeur de couverture moyenne plus importante que dans le groupe contrôle (**Tableau 9**). Ainsi, plus un gène est couvert à forte profondeur, plus on a de chance de détecter une mosaïque à faible VAF dans le sang. Jusqu'à maintenant, étaient garanties de profondeurs minimales à 30X pour une analyse constitutionnelle ce qui est insuffisant pour détecter des mosaïques faibles. Cependant, l'amélioration des techniques de préparation des bibliothèques et notamment l'utilisation de kits tels que SureSelect^{XT HS} a permis permettre d'améliorer la profondeur de couverture à des niveaux élevés en diminution notamment le nombre de *reads* dupliqués. L'autre possibilité est d'inclure moins de patients en série pour améliorer la profondeur de couverture si nécessaire. Une autre manière d'augmenter la sensibilité de la technique serait l'utilisation d'UMI (« Unique Molecular Identifiers ») dans l'étape de préparation des bibliothèques (158). Il s'agit d'index soit de

de très courtes séquences d'ADN qui sont indexées aléatoirement aux différents fragments d'ADN. Chaque index est spécifique d'un fragment d'ADN. Ainsi, cela permet après l'étape d'amplification et de séquençage de s'assurer que les séquences lues ne sont pas issues du même fragment d'ADN initial.

D'un point de vue technique, les résultats produits à partir des frottis jugaux effectués systématiquement au moment du prélèvement ont été difficiles à interpréter confortant peu l'une ou l'autre des hypothèses. En effet, même en cas de suspicion importante d'HC, on observait en Snapshot quasi systématiquement la présence du variant sur les cellules jugales (sauf quand le pic correspondant à l'allèle muté était déjà faible dans le sang), probable reflet d'une contamination du frottis jugal par des cellules leucocytaires. En effet, ces prélèvements censés être composés de cellules épithéliales ne devraient pas contenir de cellules sanguines mais le prélèvement est souvent contaminé par des cellules présentes dans la salive. La qualité de réalisation d'un frottis jugal est essentielle pour aider au diagnostic. Il est important que le préleveur soit formé à la réalisation du frottis. En cas de réalisation d'un frottis de qualité, la détection du variant dans le frottis conforterait ainsi l'hypothèse de la mosaïque. Chez la **patiente n°1**, le variant était bien vu dans le frottis jugal, plus que pour les autres patients (**Figures 11 et 14**). Cependant, dans le cas d'une mosaïque, le variant peut être retrouvé à très faible intensité voire pas du tout au niveau du frottis jugal si les cellules jugales ne portent pas toutes le variant. En s'assurant que le frottis est de qualité, même un pic de faible intensité détecté serait en faveur d'une mosaïque. Un frottis négatif n'exclut néanmoins pas une mosaïque ne touchant pas les cellules jugales. D'autres prélèvements que des cellules jugales peuvent aider dans le cas d'une suspicion de mosaïque importante. Par exemple, l'extraction d'ADN pourrait être réalisée à partir d'une culture de fibroblastes issue d'une biopsie de peau ou à partir des bulbes de sourcils (159). Enfin, au-delà de permettre de détecter une mosaïque, la réalisation d'un frottis jugal de qualité permettrait de s'assurer du caractère constitutionnel des variants détectés à des taux se rapprochant des 50 %. En effet, en routine, les patients bénéficient de deux prélèvements indépendants dans le but d'éviter des erreurs d'identitovigilance avant tout. Cependant, il paraît nécessaire dorénavant, principalement depuis l'inclusion du gène *TP53*, de ne pas confirmer les variants avec un prélèvement sanguin qui pourrait induire à rendre faussement une mutation constitutionnelle due à la détection d'un clone hématopoïétique ou d'ADNtc.

En reprenant les données phénotypiques de la population d'études, les patientes porteuses d'un variant *TP53* étaient, de manière significative, plus âgées (**Tableau 9 et 10**). Les

1388 patientes non porteuses dans le modèle multivarié tenant compte de deux facteurs confondants importants qui sont le type de cancer et la profondeur d'analyse du gène. Ces résultats sont concordants avec les données de la littérature dans l'hypothèse où ces variants seraient dus à de l'HC car l'incidence de l'HC augmente avec l'âge. En effet, l'HC est détectable chez moins de 1 % des moins de 40 ans. Inversement, la CHIP est présente chez 10 % à 15 % des personnes âgées de 70 ans et plus, la fréquence dépendant de l'étude et de la méthode d'analyse (138). De plus, 55 % à 65 % des mutations ponctuelles observées dans l'HC sont des transitions C>T qui constituent le type dominant de mutation ponctuelle dans la plupart des tissus « vieillissants » (139) (141). Cependant, dans les mutations que nous avons observées, cela ne semblait pas être le type de mutation majoritaire.

Dans cette population d'étude, un nombre significativement plus important de femmes ayant présenté un cancer de l'ovaire chez les porteuses d'un variant *TP53* potentiellement délétère a été mis en évidence (**Tableau 9 et 10**). Weitzel et al n'avaient pas retrouvé cette observation dans leur étude (160). Cependant, la proportion de cancer de l'ovaire était à 21 % et celle des cancers du sein à 53 %. Les cas de cancer de l'ovaire étaient tout de même largement augmentés. De plus, cela pourrait s'expliquer en partie par la différence de méthodologie dans la sélection des variants déséquilibrés : ici des variants avec une fréquence allélique très faible (3 %) ont été inclus alors que précédemment il n'avait pas été précisé le seuil de détection. Par ailleurs, la population de patientes ayant eu un cancer de l'ovaire recrutée dans notre étude est large et comporte des cancers chez des individus âgés depuis l'arrivée des inhibiteurs de PARP comme option thérapeutique dans leur prise en charge. Ainsi, le facteur dû à l'âge des patientes recrutées pourrait expliquer aussi la surreprésentation des cancers de l'ovaire par rapport à l'étude de Weitzel. Cependant, l'augmentation du nombre des cancers de l'ovaire dans la population des patientes porteuses de VAF faibles résiste au modèle multivarié tenant compte de l'âge. Le résultat semble robuste. Le fait de trouver plus de cancer de l'ovaire dans le groupe porteur de variant *TP53* potentiellement expliqué par un clone hématopoïétique pourrait s'expliquer par la chimiothérapie à base de taxol et de platine particulièrement cytotoxique et reçue systématiquement par les femmes atteintes d'un cancer de l'ovaire. En effet, des études récentes de CHIP dans des populations de patients précédemment exposés à une thérapie cytotoxique ont montré un enrichissement des mutations en *TP53* par rapport à des populations non sélectionnées (161) (162) (163). Une autre étude réalisée sur des patientes atteintes d'un cancer de l'ovaire n'a pas réussi à montrer de manière significative cet enrichissement des mutations *TP53* chez les patientes traitées par chimiothérapie avec 4/326 femmes parmi celles

traitées (1,2 %) contre 7/274 femmes non traitées (2,6 %) ($P = 0.24$), mais la tendance est en faveur de cet enrichissement et l'échantillonnage pourrait expliquer la non-significativité du résultat (164) (165). Le mécanisme d'apparition de ce clone hématopoïétique post traitement peut dorénavant s'expliquer. En effet, dans une étude, des variants de *TP53* identifiés chez un petit échantillon de patients atteints de leucémie aigüe myéloïde post traitement (LAM-t) ont été retrouvés grâce à une analyse de forte profondeur sur les échantillons disponibles avant leur tout premier traitement par chimiothérapie pour leur cancer initial (166). De plus, dans ces cohortes, le clone hématopoïétique détecté présente les mêmes caractéristiques que dans des populations non sélectionnées, notamment une association forte avec l'âge et une prédominance de substitutions C> T (161), ce qui suggère que les processus sous-jacents aux mutations sont similaires quel que soit le contexte. Globalement, ces données appuient un modèle selon lequel le clone hématopoïétique précède habituellement, sinon toujours, la thérapie cytotoxique. Le traitement cytotoxique favoriserait seulement la sélection du clone hématopoïétique préexistant. L'HC post traitement cytotoxique concerne aussi un autre gène de manière plus fréquente. Il s'agit des mutations récurrentes à type gain de fonction du gène *PPM1D*, une phosphatase qui régule négativement de multiples composants de la voie de réponse aux dommages de l'ADN (167). Les données de la littérature suggèrent que les mutations de *PPM1D* sont associées à l'exposition à des agents chimiothérapeutiques spécifiques, tels que le platine, l'étoposide, la cytarabine et la doxorubicine. Il serait intéressant de voir s'il en est de même pour le gène *TP53* pour lequel l'étude a montré un enrichissement de femmes traitées par platine et taxol dans le groupe des porteurs de variants de *TP53* hétérozygotes.

Une question de la nécessité de rapporter ou non la détection de ces clones hématopoïétiques de manière incidente est importante. En effet, les clones hématopoïétiques sont associés à un sur-risque de néoplasie hématologique (hazard ratio, 11.1; 95 % intervalle de confiance [IC], 3.9 to 32.6) et d'accident vasculaire cérébral ischémique (hazard ratio, 2.6; 95 % IC, 1.4 to 4.8) (139). De plus, une étude a montré que des mutations, affectant en particulier les gènes *TP53* et *U21AF1*, impliquent un risque relativement élevé de LAM ultérieure (168). Cette même étude cas/témoin a montré que les cas de pré-LAM présentaient un nombre plus important de mutations détectées par échantillon, des variants de VAF plus élevées et un enrichissement en mutations de gènes spécifiques. Une autre étude cas témoins de 221 femmes ayant développé une LAM et 221 contrôles appariés a montré que des mutations dans les gènes *IDH1*, *IDH2*, *TP53*, *DNMT3A*, *TET2* et les spliceosomes augmentaient significativement les chances de développer une LAM. Le surrisque conféré par les mutations *TP53* (OR : 47,2 ;

intervalle de confiance à 95 % : 2,5–879,1) était d'ailleurs le plus important. Dans l'étude, les 21 patients porteurs de mutations *TP53* ont ainsi développé une LAM (21/21) (169). En analyse univariée des gènes mutés, ils ont observé que la durée de latence entre l'apparition du clone et le diagnostic de LAM variait en fonction du gène muté et que cette durée était très diminuée dans le groupe des mutés *TP53* (4,9 contre 10,1 ans). Desai et Roboz proposent des recommandations pour les clones hématopoïétiques les plus à risque de LAM-t qu'ils décrivent comme ceux porteurs d'une mutation associée à un risque élevé de LAM-t (*TP53*, *PPM1D*, *DNMT3A* R882H), ceux avec des mutations driver de l'HC avec une fraction allélique > 10 % ou les clones ayant plus d'une mutation impliquée dans l'HC (170). Dans ces cas, ils recommandent d'effectuer une surveillance de la formule sanguine tous les 3 à 6 mois pendant 10 ans car l'incidence des LAM-t est la plus élevée dans les 10 ans post traitements. Cependant, mis à part la surveillance, aucun traitement préventif n'est préconisé à l'heure actuelle.

CONCLUSION

Finalement, mon travail de thèse lève le voile sur la complexité d'interpréter les variants de faible VAF au cours du diagnostic moléculaire des prédispositions au cancer du sein et de l'ovaire. En effet, au cours de cette étude, nous avons observé les trois cas possibles parmi la présence d'une néomutation en mosaïque, la détection d'ADNtc et d'HC. De plus, la mise en place du panel HBOC 13 gènes incluant *TP53* sur les recommandations du GGC augmente de manière importante l'incidence de ces détections. Il est alors primordial de tenir compte du phénotype des patientes ainsi que du gène concerné dans la démarche diagnostique. En effet, le gène *TP53* qui a été retrouvé en très grande majorité dans l'étude est, parmi les gènes du panel HBOC, le plus impliqué dans l'HC, en particulier chez les patientes âgées traitées par chimiothérapie. Lorsque le phénotype est évocateur, il est important de s'attarder sur le variant pour ne pas passer à côté d'une véridique mosaïque mais lorsqu'il ne l'est pas, il peut paraître déraisonnable voire délétère d'inquiéter le patient et ses descendants en particulier. Grâce à ce travail nous avons pu conclure à l'origine du variant chez les patientes pour lesquelles nous avons reçu un prélèvement tumoral et proposer un conseil génétique pour une patiente porteuse de mosaïque du gène *BRCA1*. Dans le but de simplifier la prise en charge de ces variants de faible VAF, nous proposons une démarche diagnostique afin de résoudre le diagnostic différentiel (**Figure 16**) qui pourra sûrement évoluer les prochaines années avec l'évolution des techniques et une meilleure compréhension des mécanismes physiopathologiques des trois événements. Concernant les variants de *TP53* dus à l'HC, il conviendrait de préciser le risque hématologique afin d'envisager de rapporter ces découvertes finalement fortuites.

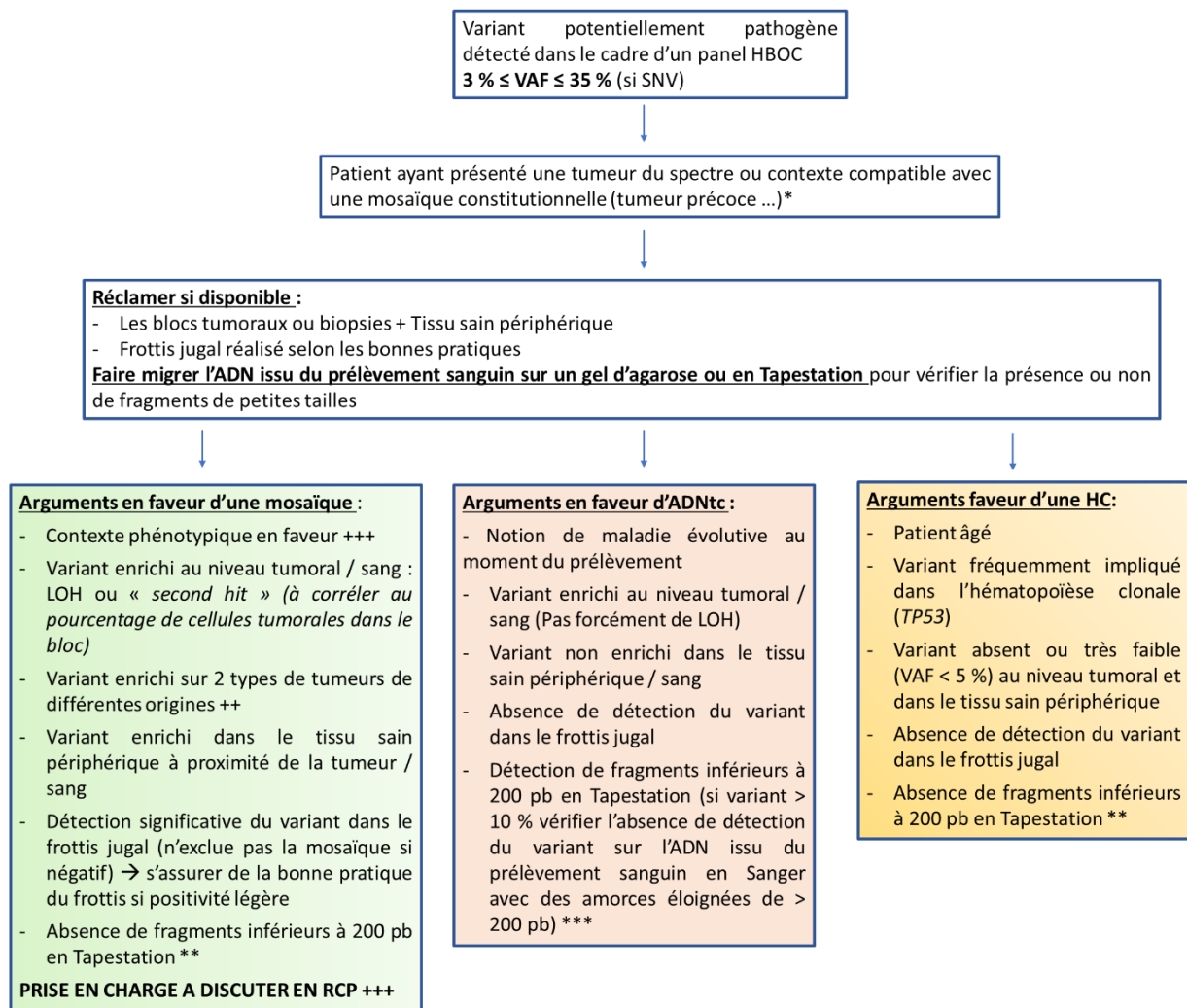


Figure 16: Diagramme d'aide diagnostique suite à la détection d'un variant potentiellement délétère de faible VAF

Abréviations: HBOC : « Hereditary Breast and Ovarian Cancer » ; VAF : « Variant Allele Fraction » ; LOH : « Lost of heterozygotie » ; SNV : « Single Nucleotide Variation »

*Ne pas hésiter à en discuter avec un centre de référence du gène concerné

**Si la concentration en ADN et que la VAF du variant sont faibles, il est possible de ne pas apercevoir les petits fragments qui seraient présents.

***En cas de détection de plusieurs bandes en échelles, multiples de 180 pb environs, il ne s'agit pas d'un argument en faveur de la détection d'ADN tumoral circulant mais plutôt l'observation d'ADN dégradé. En cas de nécessité de prouver la présence du variant dans les fragments de bas poids, il est possible de faire migrer l'ADN sur gel d'agarose, de le purifier et réaliser un séquençage Sanger ou un Snapshot avec des amorces permettant l'amplification d'un fragment < 150 pb.

Toutes les mosaïques suspectées doivent faire l'objet d'une discussion en RCP pour la prise en charge du patient et de ses descendants.

BIBLIOGRAPHIE

1. Ferlay J, Parkin DM, Steliarova-Foucher E. Estimates of cancer incidence and mortality in Europe in 2008. *Eur J Cancer*. mars 2010;46(4):765-81.
2. Le cancer du sein - Les cancers les plus fréquents [Internet]. [cité 29 août 2019]. 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-sein>
3. Données de participation au programme de dépistage organisé du cancer du sein, 2015-2016 [Internet]. [cité 29 août 2019]. Disponible sur: </liste-des-actualites/donnees-de-participation-au-programme-de-depistage-organise-du-cancer-du-sein-2015-2016>
4. Prat A, Fan C, Fernández A, Hoadley KA, Martinello R, Vidal M, et al. Response and survival of breast cancer intrinsic subtypes following multi-agent neoadjuvant chemotherapy. *BMC Med*. 18 déc 2015;13(1):303.
5. Dai X, Li T, Bai Z, Yang Y, Liu X, Zhan J, et al. Breast cancer intrinsic subtype classification, clinical use and future trends. *Am J Cancer Res*. 2015;5(10):2929-43.
6. Espié M. Prise en charge du cancer du sein. *J Radiol Diagn Interv*. juill 2014;95(7-8):740-4.
7. Filipits M, Nielsen TO, Rudas M, Greil R, Stöger H, Jakesz R, et al. The PAM50 Risk-of-Recurrence Score Predicts Risk for Late Distant Recurrence after Endocrine Therapy in Postmenopausal Women with Endocrine-Responsive Early Breast Cancer. *Clin Cancer Res*. 1 mars 2014;20(5):1298-305.
8. Wallden B, Storhoff J, Nielsen T, Dowidar N, Schaper C, Ferree S, et al. Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. *BMC Med Genomics*. déc 2015;8(1):54.
9. Poursuivre la recherche clinique pour positionner utilement les signatures génomiques dans la prise en charge des cancers du sein [Internet]. Haute Autorité de Santé. [cité 29 août 2019]. Disponible sur: https://www.has-sante.fr/jcms/c_2903408/fr/poursuivre-la-recherche-clinique-pour-positionner-utilement-les-signatures-genomiques-dans-la-prise-en-charge-des-cancers-du-sein
10. Peddi PF. Hormone receptor positive breast cancer: state of the art. *Curr Opin Obstet Gynecol*. déc 2017;1.
11. Pathak M, Dwivedi SN, Deo SVS, Thakur B, Sreenivas V, Rath GK. Effectiveness of Added Targeted Therapies to Neoadjuvant Chemotherapy for Breast Cancer: A Systematic Review and Meta-analysis. *Clin Breast Cancer*. juin 2019;S1526820919303234.
12. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100 000 women in 123 randomised trials. *The Lancet*. févr 2012;379(9814):432-44.

13. Strach MC, Prasanna T, Kirova YM, Alran S, O'Toole S, Beith JM, et al. Optimise not compromise: The importance of a multidisciplinary breast cancer patient pathway in the era of oncoplastic and reconstructive surgery. *Crit Rev Oncol Hematol.* févr 2019;134:10-21.
14. Haute Autorité de santé définissant les actes et prestations pour l'ALD n° 30 « Tumeur maligne, affection maligne du tissu lymphatique ou hématopoïétique - cancer de l'ovaire ».
15. Lakhani SR. Pathology of Ovarian Cancers in BRCA1 and BRCA2 Carriers. *Clin Cancer Res.* 1 avr 2004;10(7):2473-81.
16. Helder-Woolderink JM, Blok EA, Vasen HFA, Hollema H, Mourits MJ, De Bock GH. Ovarian cancer in Lynch syndrome; a systematic review. *Eur J Cancer.* mars 2016;55:65-73.
17. Gadducci A, Guarneri V, Peccatori FA, Ronzino G, Scandurra G, Zamagni C, et al. Current strategies for the targeted treatment of high-grade serous epithelial ovarian cancer and relevance of BRCA mutational status. *J Ovarian Res.* déc 2019;12(1):9.
18. Aghajanian C, Blank SV, Goff BA, Judson PL, Teneriello MG, Husain A, et al. OCEANS: A Randomized, Double-Blind, Placebo-Controlled Phase III Trial of Chemotherapy With or Without Bevacizumab in Patients With Platinum-Sensitive Recurrent Epithelial Ovarian, Primary Peritoneal, or Fallopian Tube Cancer. *J Clin Oncol.* 10 juin 2012;30(17):2039-45.
19. Aghajanian C, Goff B, Nycum LR, Wang YV, Husain A, Blank SV. Final overall survival and safety analysis of OCEANS, a phase 3 trial of chemotherapy with or without bevacizumab in patients with platinum-sensitive recurrent ovarian cancer. *Gynecol Oncol.* oct 2015;139(1):10-6.
20. Ledermann J, Harter P, Gourley C, Friedlander M, Vergote I, Rustin G, et al. Olaparib maintenance therapy in patients with platinum-sensitive relapsed serous ovarian cancer: a preplanned retrospective analysis of outcomes by BRCA status in a randomised phase 2 trial. *Lancet Oncol.* juill 2014;15(8):852-61.
21. Ledermann JA, Harter P, Gourley C, Friedlander M, Vergote I, Rustin G, et al. Overall survival in patients with platinum-sensitive recurrent serous ovarian cancer receiving olaparib maintenance monotherapy: an updated analysis from a randomised, placebo-controlled, double-blind, phase 2 trial. *Lancet Oncol.* nov 2016;17(11):1579-89.
22. Weissman SM, Weiss SM, Newlin AC. Genetic Testing by Cancer Site: Ovary. *Cancer J.* 2012;18(4):320-7.
23. Miki Y, Swensen J, Shattuck-Eidens D, Futreal P, Harshman K, Tavtigian S, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science.* 7 oct 1994;266(5182):66-71.
24. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature.* déc 1995;378(6559):789-92.

25. Oncogénétique en 2016 - Consultations, laboratoires et suivi - Ref: ADONCOG17 [Internet]. [cité 30 août 2019]. Disponible sur: <https://www.e-cancer.fr/Expertises-et-publications/Catalogue-des-publications/Oncogenetique-en-2016-Consultations-laboratoires-et-suivi>
26. Knudson AG. Mutation and Cancer: Statistical Study of Retinoblastoma. *Proc Natl Acad Sci.* 1 avr 1971;68(4):820-3.
27. Bunz F. Tumor Suppressor Genes. In: Bunz F, éditeur. *Principles of Cancer Genetics* [Internet]. Dordrecht: Springer Netherlands; 2016 [cité 29 août 2019]. p. 75-134. Disponible sur: https://doi.org/10.1007/978-94-017-7484-0_3
28. Hall J, Lee M, Newman B, Morrow J, Anderson L, Huey B, et al. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science.* 21 déc 1990;250(4988):1684-9.
29. Wooster R, Neuhausen S, Mangion J, Quirk Y, Ford D, Collins N, et al. Localization of a breast cancer susceptibility gene, *BRCA2*, to chromosome 13q12-13. *Science.* 30 sept 1994;265(5181):2088-90.
30. Gold ER, Carbone J. Myriad Genetics: In the eye of the policy storm. *Genet Med Off J Am Coll Med Genet.* avr 2010;12(4 Suppl):S39-70.
31. United States District Court. Association for Molecular Pathology et al. against US Patent and Trademark Office et al., Document 255, Filed 03/29/2010.
32. Cropper C, Woodson A, Arun B, Barcenas C, Litton J, Noblin S, et al. Evaluating the NCCN Clinical Criteria for Recommending *BRCA1* and *BRCA2* Genetic Testing in Patients With Breast Cancer. *J Natl Compr Canc Netw.* juin 2017;15(6):797-803.
33. Petrucelli N, Daly MB, Feldman GL. Hereditary breast and ovarian cancer due to mutations in *BRCA1* and *BRCA2*. *Genet Med.* mai 2010;12(5):245-59.
34. Kuchenbaecker KB, Hopper JL, Barnes DR, Phillips K-A, Mooij TM, Roos-Blom M-J, et al. Risks of Breast, Ovarian, and Contralateral Breast Cancer for *BRCA1* and *BRCA2* Mutation Carriers. *JAMA.* 20 juin 2017;317(23):2402.
35. Castro E, Eeles R. The role of *BRCA1* and *BRCA2* in prostate cancer. *Asian J Androl.* mai 2012;14(3):409-14.
36. Iqbal J, Ragone A, Lubinski J, Lynch HT, Moller P, Ghadirian P, et al. The incidence of pancreatic cancer in *BRCA1* and *BRCA2* mutation carriers. *Br J Cancer.* 25 oct 2012;107:2005.
37. Lucas AL, Frado LE, Hwang C, Kumar S, Khanna LG, Levinson EJ, et al. *BRCA1* and *BRCA2* germline mutations are frequently demonstrated in both high-risk pancreatic cancer screening and pancreatic cancer cohorts: *BRCA1/2* Germline Mutations in Pancreatic Cancer. *Cancer.* 1 juill 2014;120(13):1960-7.
38. Keefe P, Bokhari SRA. Fanconi Syndrome. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2019 [cité 29 août 2019]. Disponible sur: <http://www.ncbi.nlm.nih.gov/books/NBK534872/>

39. Xia B, Sheng Q, Nakanishi K, Ohashi A, Wu J, Christ N, et al. Control of BRCA2 Cellular and Clinical Functions by a Nuclear Partner, PALB2. *Mol Cell*. juin 2006;22(6):719-29.
40. Antoniou AC, Casadei S, Heikkinen T, Barrowdale D, Pylkäs K, Roberts J, et al. Breast-Cancer Risk in Families with Mutations in *PALB2*. *N Engl J Med*. 7 août 2014;371(6):497-506.
41. Thompson ER, Rowley SM, Li N, McInerny S, Devereux L, Wong-Brown MW, et al. Panel Testing for Familial Breast Cancer: Calibrating the Tension Between Research and Clinical Care. *J Clin Oncol*. mai 2016;34(13):1455-9.
42. Castéra L, Harter V, Muller E, Krieger S, Goardon N, Ricou A, et al. Landscape of pathogenic variations in a panel of 34 genes and cancer risk estimation from 5131 HBOC families. *Genet Med Off J Am Coll Med Genet*. 2018;20(12):1677-86.
43. Southey MC, Goldgar DE, Winqvist R, Pylkäs K, Couch F, Tischkowitz M, et al. *PALB2*, *CHEK2* and *ATM* rare variants and cancer risk: data from COGS. *J Med Genet*. déc 2016;53(12):800-11.
44. Loveday C, Turnbull C, Ruark E, Xicola RMM, Ramsay E, Hughes D, et al. Germline RAD51C mutations confer susceptibility to ovarian cancer. *Nat Genet*. 26 avr 2012;44:475.
45. Loveday C, Turnbull C, Ramsay E, Hughes D, Ruark E, Frankum JR, et al. Germline mutations in RAD51D confer susceptibility to ovarian cancer. *Nat Genet*. 7 août 2011;43:879.
46. Li N, McInerny S, Zethoven M, Cheasley D, Lim BWX, Rowley SM, et al. Combined Tumor Sequencing and Case-Control Analyses of RAD51C in Breast Cancer. *JNCI J Natl Cancer Inst*. 5 avr 2019;djz045.
47. Kourtidis A, Lu R, Pence LJ, Anastasiadis PZ. A central role for cadherin signaling in cancer. *Exp Cell Res*. sept 2017;358(1):78-85.
48. Hansford S, Kaurah P, Li-Chang H, Woo M, Senz J, Pinheiro H, et al. Hereditary Diffuse Gastric Cancer Syndrome: *CDHI* Mutations and Beyond. *JAMA Oncol*. 1 avr 2015;1(1):23.
49. Pilarski R, Burt R, Kohlman W, Pho L, Shannon KM, Swisher E. Cowden syndrome and the PTEN hamartoma tumor syndrome: systematic review and revised diagnostic criteria. *J Natl Cancer Inst*. 6 nov 2013;105(21):1607-16.
50. Riegert-Johnson DL, Gleeson FC, Roberts M, Tholen K, Youngborg L, Bullock M, et al. Cancer and Lhermitte-Duclos disease are common in Cowden syndrome patients. *Hered Cancer Clin Pract*. 17 juin 2010;8(1):6.
51. Bubien V, Bonnet F, Brouste V, Hoppe S, Barouk-Simonet E, David A, et al. High cumulative risks of cancer in patients with *PTEN* hamartoma tumour syndrome. *J Med Genet*. avr 2013;50(4):255-63.

52. Bonadona V. Cancer Risks Associated With Germline Mutations in MLH1, MSH2, and MSH6 Genes in Lynch Syndrome. *JAMA*. 8 juin 2011;305(22):2304.
53. Bougeard G, Renaux-Petel M, Flaman J-M, Charbonnier C, Fermeijer P, Belotti M, et al. Revisiting Li-Fraumeni Syndrome From *TP53* Mutation Carriers. *J Clin Oncol*. 20 juill 2015;33(21):2345-52.
54. Kasper E, Angot E, Colasse E, Nicol L, Sabourin J-C, Adriouch S, et al. Contribution of genotoxic anticancer treatments to the development of multiple primary tumours in the context of germline *TP53* mutations. *Eur J Cancer*. sept 2018;101:254-62.
55. de Andrade KC, Mirabello L, Stewart DR, Karlins E, Koster R, Wang M, et al. Higher-than-expected population prevalence of potentially pathogenic germline *TP53* variants in individuals unselected for cancer history. *Hum Mutat*. déc 2017;38(12):1723-30.
56. Moretta J, Berthet P, Bonadona V, Caron O, Cohen-Haguenaer O, Colas C, et al. Recommandations françaises du Groupe Génétique et Cancer pour l'analyse en panel de gènes dans les prédispositions héréditaires au cancer du sein ou de l'ovaire. *Bull Cancer (Paris)*. oct 2018;105(10):907-17.
57. Liu W, Smith DI, Reetz KJ, Thibodeau SN, James CD. Denaturing high performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations. *Nucleic Acids Res*. 1 mars 1998;26(6):1396-400.
58. Riahi A, Kharrat M, Lariani I, Chaabouni-Bouhamed H. High-resolution melting (HRM) assay for the detection of recurrent *BRCA1/BRCA2* germline mutations in Tunisian breast/ovarian cancer families. *Fam Cancer*. déc 2014;13(4):603-9.
59. Caux-Moncoutier V, Castéra L, Tirapo C, Michaux D, Rémon M-A, Laugé A, et al. EMMA, a cost- and time-effective diagnostic method for simultaneous detection of point mutations and large-scale genomic rearrangements: application to *BRCA1* and *BRCA2* in 1,525 patients. *Hum Mutat*. mars 2011;32(3):325-34.
60. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):R25.
61. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *N Engl J Med*. 8 mars 2012;366(10):883-92.
62. Muller E, Goardon N, Brault B, Rousselin A, Paimparay G, Legros A, et al. OutLyzer: software for extracting low-allele-frequency tumor mutations from sequencing background noise in clinical practice. *Oncotarget [Internet]*. 29 nov 2016;7(48).
63. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 1 sept 2010;38(16):e164-e164.

64. den Dunnen JT, Dalgleish R, Maglott DR, Hart RK, Greenblatt MS, McGowan-Jordan J, et al. HGVS Recommendations for the Description of Sequence Variants: 2016 Update. *Hum Mutat.* juin 2016;37(6):564-9.
65. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, 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.* mai 2015;17(5):405-24.
66. Plon SE, Eccles DM, Easton D, Foulkes WD, Genuardi M, Greenblatt MS, et al. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat.* nov 2008;29(11):1282-91.
67. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 1 janv 2001;29(1):308-11.
68. 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.
69. Bérout C, Letovsky SI, Braastad CD, Caputo SM, Beaudoux O, Bignon YJ, et al. BRCA Share: A Collection of Clinical BRCA Gene Variants: HUMAN MUTATION. *Hum Mutat.* déc 2016;37(12):1318-28.
70. Cline MS, Liao RG, Parsons MT, Paten B, Alquaddoomi F, Antoniou A, et al. BRCA Challenge: BRCA Exchange as a global resource for variants in BRCA1 and BRCA2. *Eng C, éditeur. PLOS Genet.* 26 déc 2018;14(12):e1007752.
71. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* avr 2010;7(4):248-9.
72. Desmet F-O, Hamroun D, Lalande M, Collod-Bérout G, Claustres M, Bérout C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* mai 2009;37(9):e67-e67.
73. Yeo G, Burge CB. Maximum Entropy Modeling of Short Sequence Motifs with Applications to RNA Splicing Signals. *J Comput Biol.* mars 2004;11(2-3):377-94.
74. Leman R, Gaildrat P, Gac GL, Ka C, Fichou Y, Audrezet M-P, et al. Novel diagnostic tool for prediction of variant spliceogenicity derived from a set of 395 combined in silico/in vitro studies: an international collaborative effort. *Nucleic Acids Res.* 6 sept 2018;46(15):7913-23.
75. Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BWM, Willemsen MH, et al. Genome sequencing identifies major causes of severe intellectual disability. *Nature.* juill 2014;511(7509):344-7.

76. Ainsworth PJ, Chakraborty PK, Weksberg R. Example of somatic mosaicism in a series of de novo neurofibromatosis type 1 cases due to a maternally derived deletion. *Hum Mutat.* 1997;9(5):452-7.
77. Evans DGR, Wallace AJ, Wu CL, Trueman L, Ramsden RT, Strachan T. Somatic Mosaicism: A Common Cause of Classic Disease in Tumor-Prone Syndromes? Lessons from Type 2 Neurofibromatosis. *Am J Hum Genet.* sept 1998;63(3):727-36.
78. Youssoufian H, Pyeritz RE. Mechanisms and consequences of somatic mosaicism in humans. *Nat Rev Genet.* oct 2002;3(10):748-58.
79. Hall JG. Review and hypotheses: somatic mosaicism: observations related to clinical genetics. *Am J Hum Genet.* oct 1988;43(4):355-63.
80. Parkhurst HJ. CLEVELAND DERMATOLOGICAL SOCIETY. *Arch Dermatol.* 1 oct 1931;24(4):712.
81. García-Romero MT, Parkin P, Lara-Corrales I. Mosaic Neurofibromatosis Type 1: A Systematic Review. *Pediatr Dermatol.* janv 2016;33(1):9-17.
82. Evans DG. Neurofibromatosis 2. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJ, Stephens K, et al., éditeurs. *GeneReviews®* [Internet]. Seattle (WA): University of Washington, Seattle; 1993 [cité 20 juill 2019].
83. Evans DG, Hartley CL, Smith PT, King AT, Bowers NL, Tobi S, et al. Incidence of mosaicism in 1055 de novo NF2 cases: much higher than previous estimates with high utility of next-generation sequencing. *Genet Med* [Internet]. 5 juill 2019;
84. Amitrano S, Marozza A, Somma S, Imperatore V, Hadjistilianou T, De Francesco S, et al. Next generation sequencing in sporadic retinoblastoma patients reveals somatic mosaicism. *Eur J Hum Genet.* oct 2015;23(11):1523-30.
85. Golmard L, Delnatte C, Laugé A, Moncoutier V, Lefol C, Abidallah K, et al. Breast and ovarian cancer predisposition due to de novo BRCA1 and BRCA2 mutations. *Oncogene.* mars 2016;35(10):1324-7.
86. Veltman JA, Brunner HG. De novo mutations in human genetic disease. *Nat Rev Genet.* août 2012;13(8):565-75.
87. Moslehi R, Singh R, Lessner L, Friedman JM. Impact of *BRCA* mutations on female fertility and offspring sex ratio. *Am J Hum Biol.* 2009;NA-NA.
88. Kwiatkowski F, Arbre M, Bidet Y, Laquet C, Uhrhammer N, Bignon Y-J. BRCA Mutations Increase Fertility in Families at Hereditary Breast/Ovarian Cancer Risk. *Fei P*, éditeur. *PLOS ONE.* 5 juin 2015;10(6):e0127363.
89. Friedman E, Efrat N, Soussan-Gutman L, Dvir A, Kaplan Y, Ekstein T, et al. Low-level constitutional mosaicism of a de novo BRCA1 gene mutation. *Br J Cancer.* févr 2015;112(4):765-8.

90. Delon I, Taylor A, Molenda A, Drummond J, Oakhill K, Girling A, et al. A germline mosaic *BRCA1* exon deletion in a woman with bilateral basal-like breast cancer: Letter to the Editor. *Clin Genet.* sept 2013;84(3):297-9.
91. Mouchawar J, Korch C, Byers T, Pitts TM, Li E, McCredie MRE, et al. Population-Based Estimate of the Contribution of *TP53* Mutations to Subgroups of Early-Onset Breast Cancer: Australian Breast Cancer Family Study: Table 1. *Cancer Res.* 15 juin 2010;70(12):4795-800.
92. Evans DGR, Moran A, Hartley R, Dawson J, Bulman B, Knox F, et al. Long-term outcomes of breast cancer in women aged 30 years or younger, based on family history, pathology and *BRCA1/BRCA2/TP53* status. *Br J Cancer.* mars 2010;102(7):1091-8.
93. Gonzalez KD, Buzin CH, Noltner KA, Gu D, Li W, Malkin D, et al. High frequency of de novo mutations in Li-Fraumeni syndrome. *J Med Genet.* 1 oct 2009;46(10):689-93.
94. Renaux-Petel M, Charbonnier F, Théry J-C, Fermey P, Lienard G, Bou J, et al. Contribution of de novo and mosaic *TP53* mutations to Li-Fraumeni syndrome. *J Med Genet.* 25 oct 2017;jmedgenet-2017-104976.
95. Prochazkova K, Pavlikova K, Minarik M, Sumerauer D, Kodet R, Sedlacek Z. Somatic *TP53* mutation mosaicism in a patient with Li-Fraumeni syndrome. *Am J Med Genet A.* févr 2009;149A(2):206-11.
96. Behjati S, Maschietto M, Williams RD, Side L, Hubank M, West R, et al. A Pathogenic Mosaic *TP53* Mutation in Two Germ Layers Detected by Next Generation Sequencing. Mazoyer S, éditeur. *PLoS ONE.* 8 mai 2014;9(5):e96531.
97. Mester J, Eng C. Estimate of de novo mutation frequency in probands with *PTEN* hamartoma tumor syndrome. *Genet Med.* sept 2012;14(9):819-22.
98. Mester J, Eng C. When Overgrowth Bumps Into Cancer: The *PTEN*-Opathies: AMERICAN JOURNAL OF MEDICAL GENETICS PART C (SEMINARS IN MEDICAL GENETICS). *Am J Med Genet C Semin Med Genet.* mai 2013;163(2):114-21.
99. Pritchard CC, Smith C, Marushchak T, Koehler K, Holmes H, Raskind W, et al. A mosaic *PTEN* mutation causing Cowden syndrome identified by deep sequencing. *Genet Med.* déc 2013;15(12):1004-7.
100. Salo-Mullen EE, Shia J, Brownell I, Allen P, Girotra M, Robson ME, et al. Mosaic partial deletion of the *PTEN* gene in a patient with Cowden syndrome. *Fam Cancer.* sept 2014;13(3):459-67.
101. Gammon A, Jasperson K, Pilarski R, Prior T, Kuwada S. *PTEN* mosaicism with features of Cowden syndrome: *PTEN* mosaicism with features of CS. *Clin Genet.* déc 2013;84(6):593-5.
102. Caux F, Plauchu H, Chibon F, Faivre L, Fain O, Vabres P, et al. Segmental overgrowth, lipomatosis, arteriovenous malformation and epidermal nevus (SOLAMEN) syndrome is related to mosaic *PTEN* nullizygosity. *Eur J Hum Genet.* juill 2007;15(7):767-73.

103. Nathan N, Keppler-Noreuil KM, Biesecker LG, Moss J, Darling TN. Mosaic Disorders of the PI3K/PTEN/AKT/TSC/mTORC1 Signaling Pathway. *Dermatol Clin.* janv 2017;35(1):51-60.
104. Win AK, Jenkins MA, Buchanan DD, Clendenning M, Young JP, Giles GG, et al. Determining the frequency of de novo germline mutations in DNA mismatch repair genes. *J Med Genet.* 1 août 2011;48(8):530-4.
105. Sourrouille I, Coulet F, Lefevre JH, Colas C, Eyries M, Svrcek M, et al. Somatic mosaicism and double somatic hits can lead to MSI colorectal tumors. *Fam Cancer.* mars 2013;12(1):27-33.
106. Ababneh FK, Al-Swaid A, Elhag A, Youssef T, Alsaif S. Blepharo-cheilo-dontic (BCD) syndrome: Expanding the phenotype, case report and review of literature. *Am J Med Genet A.* juin 2014;164(6):1525-9.
107. Kievit A, Tessadori F, Douben H, Jordens I, Maurice M, Hoogeboom J, et al. Variants in members of the cadherin–catenin complex, CDH1 and CTNND1, cause blepharocheilodontic syndrome. *Eur J Hum Genet.* févr 2018;26(2):210-9.
108. Krempely K, Karam R. A novel de novo CDH1 germline variant aids in the classification of carboxy-terminal E-cadherin alterations predicted to escape nonsense-mediated mRNA decay. *Cold Spring Harb Mol Case Stud.* 2018;4(4).
109. Sugimoto S, Yamada H, Takahashi M, Morohoshi Y, Yamaguchi N, Tsunoda Y, et al. Early-onset diffuse gastric cancer associated with a de novo large genomic deletion of CDH1 gene. *Gastric Cancer Off J Int Gastric Cancer Assoc Jpn Gastric Cancer Assoc.* oct 2014;17(4):745-9.
110. Shah MA, Salo-Mullen E, Stadler Z, Ruggeri JM, Mirander M, Pristyazhnyuk Y, et al. De novo CDH1 mutation in a family presenting with early-onset diffuse gastric cancer. *Clin Genet.* sept 2012;82(3):283-7.
111. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med.* juin 2017;23(6):703-13.
112. Merajver SD, Pham TM, Caduff RF, Chen M, Poy EL, Cooney KA, et al. Somatic mutations in the BRCA1 gene in sporadic ovarian tumours. *Nat Genet.* avr 1995;9(4):439-43.
113. Foster KA, Harrington P, Kerr J, Russell P, DiCioccio RA, Scott IV, et al. Somatic and germline mutations of the BRCA2 gene in sporadic ovarian cancer. *Cancer Res.* 15 août 1996;56(16):3622-5.
114. Berchuck A, Heron KA, Carney ME, Lancaster JM, Fraser EG, Vinson VL, et al. Frequency of germline and somatic BRCA1 mutations in ovarian cancer. *Clin Cancer Res Off J Am Assoc Cancer Res.* oct 1998;4(10):2433-7.
115. Hennessy BTJ, Timms KM, Carey MS, Gutin A, Meyer LA, Flake DD, et al. Somatic Mutations in *BRCA1* and *BRCA2* Could Expand the Number of Patients That Benefit

- From Poly (ADP Ribose) Polymerase Inhibitors in Ovarian Cancer. *J Clin Oncol.* août 2010;28(22):3570-6.
116. Dougherty BA, Lai Z, Hodgson DR, Orr MCM, Hawryluk M, Sun J, et al. Biological and clinical evidence for somatic mutations in *BRCA1* and *BRCA2* as predictive markers for olaparib response in high-grade serous ovarian cancers in the maintenance setting. *Oncotarget* [Internet]. 4 juill 2017;8(27).
 117. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature.* oct 2012;490(7418):61-70.
 118. Sharma P, Klemp JR, Kimler BF, Mahnken JD, Geier LJ, Khan QJ, et al. Germline BRCA mutation evaluation in a prospective triple-negative breast cancer registry: implications for hereditary breast and/or ovarian cancer syndrome testing. *Breast Cancer Res Treat.* juin 2014;145(3):707-14.
 119. Kandath C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. *Nature.* oct 2013;502(7471):333-9.
 120. Mansha M, Gill A, Thomson PC. Potential risk factors of ovarian cancer and analysis of CA125, a biomarker used for its monitoring and diagnosis. *Mol Biol Rep.* juin 2019;46(3):3325-32.
 121. Sikaris KA. CA125—A Test with a Change of Heart. *Heart Lung Circ.* oct 2011;20(10):634-40.
 122. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer Genome Landscapes. *Science.* 29 mars 2013;339(6127):1546-58.
 123. Ashworth, T.R. (1869) A Case of Cancer in Which Cells Similar to Those in the Tumours Were Seen in the Blood after Death. *The Medical Journal of Australia*, 14, 146-147. In.
 124. Racila E, Euhus D, Weiss AJ, Rao C, McConnell J, Terstappen LWMM, et al. Detection and characterization of carcinoma cells in the blood. *Proc Natl Acad Sci.* 14 avr 1998;95(8):4589-94.
 125. Stroun M, Lyautey J, Lederrey C, Olson-Sand A, Anker P. About the possible origin and mechanism of circulating DNA apoptosis and active DNA release. *Clin Chim Acta Int J Clin Chem.* nov 2001;313(1-2):139-42.
 126. Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, et al. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res.* 15 févr 2001;61(4):1659-65.
 127. Free DNA in the Serum of Cancer Patients and the Effect of Therapy S. A. Leon, B. Shapiro, D. M. Sklaroff and M. J. Yaros DOI: Published March 1977.
 128. AG B. Liquid biopsy identifies residual breast cancer [Internet]. [cité 29 août 2019]. Disponible sur: <https://european-biotechnology.com/up-to-date/latest-news/news/liquid-biopsy-identifies-residual-breast-cancer.html>

129. Tong Y, Shen S, Jiang H, Chen Z. Application of Digital PCR in Detecting Human Diseases Associated Gene Mutation. *Cell Physiol Biochem*. 2017;43(4):1718-30.
130. Dawson S-J, Tsui DWY, Murtaza M, Biggs H, Rueda OM, Chin S-F, et al. Analysis of Circulating Tumor DNA to Monitor Metastatic Breast Cancer. *N Engl J Med*. 28 mars 2013;368(13):1199-209.
131. Hench IB, Hench J, Tolnay M. Liquid Biopsy in Clinical Management of Breast, Lung, and Colorectal Cancer. *Front Med*. 2018;5:9.
132. Bettgowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al. Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies. *Sci Transl Med*. 19 févr 2014;6(224):224ra24-224ra24.
133. Zhang X, Ju S, Wang X, Cong H. Advances in liquid biopsy using circulating tumor cells and circulating cell-free tumor DNA for detection and monitoring of breast cancer. *Clin Exp Med*. 1 août 2019;19(3):271-9.
134. Rothé F, Silva MJ, Venet D, Campbell C, Bradburry I, Rouas G, et al. Circulating Tumor DNA in HER2-Amplified Breast Cancer: A Translational Research Substudy of the NeoALTTO Phase III Trial. *Clin Cancer Res*. 15 juin 2019;25(12):3581-8.
135. Buono G, Gerratana L, Bulfoni M, Provinciali N, Basile D, Giuliano M, et al. Circulating tumor DNA analysis in breast cancer: Is it ready for prime-time? *Cancer Treat Rev*. févr 2019;73:73-83.
136. Ratajska M, Koczkowska M, Żuk M, Gorczyński A, Kuźniacka A, Stukan M, et al. Detection of *BRCA1/2* mutations in circulating tumor DNA from patients with ovarian cancer. *Oncotarget* [Internet]. 24 nov 2017;8(60).
137. Steensma DP, Bejar R, Jaiswal S, Lindsley RC, Sekeres MA, Hasserjian RP, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood*. 2 juill 2015;126(1):9-16.
138. Gibson CJ, Steensma DP. New Insights from Studies of Clonal Hematopoiesis. *Clin Cancer Res*. 1 oct 2018;24(19):4633-42.
139. Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, et al. Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *N Engl J Med*. 25 déc 2014;371(26):2488-98.
140. Genovese G, Kähler AK, Handsaker RE, Lindberg J, Rose SA, Bakhoum SF, et al. Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *N Engl J Med*. 25 déc 2014;371(26):2477-87.
141. Xie M, Lu C, Wang J, McLellan MD, Johnson KJ, Wendl MC, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med*. déc 2014;20(12):1472-8.
142. Craig DW, Pearson JV, Szelinger S, Sekar A, Redman M, Corneveaux JJ, et al. Identification of genetic variants using bar-coded multiplexed sequencing. *Nat Methods*. oct 2008;5(10):887-93.

143. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 15 juill 2009;25(14):1754-60.
144. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 1 sept 2010;20(9):1297-303.
145. Mathe E, Olivier M, Kato S, Ishioka C, Hainaut P, Tavtigian SV. Computational approaches for predicting the biological effect of p53 missense mutations: a comparison of three sequence analysis based methods. *Nucleic Acids Res*. 2006;34(5):1317-25.
146. Tavtigian SV. Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J Med Genet*. 9 sept 2005;43(4):295-305.
147. Ewing B, Green P. Base-Calling of Automated Sequencer Traces Using *Phred*. II. Error Probabilities. *Genome Res*. 1 mars 1998;8(3):186-94.
148. Ewing B, Hillier L, Wendl MC, Green P. Base-Calling of Automated Sequencer Traces Using *Phred*. I. Accuracy Assessment. *Genome Res*. 1 mars 1998;8(3):175-85.
149. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*. 8 janv 2019;47(D1):D886-94.
150. Raphaël L, + S, Krieger, + P, Gaildrat, Le Gac G, et al. NOUVELLES RECOMMANDATIONS POUR LA PRÉDICTION DES VARIANTES D'ÉPISSAGE À PARTIR D'ÉTUDES COMBINÉES IN SILICO/IN VITRO DE 395 VARIANTES Claude Houdayer 5+# des groupes UGG + et ENIGMA # Introduction. 2017.
151. Li Q, Wang K. InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *Am J Hum Genet*. févr 2017;100(2):267-80.
152. Bouaoun L, Sonkin D, Ardin M, Hollstein M, Byrnes G, Zavadil J, et al. *TP53* Variations in Human Cancers: New Lessons from the IARC *TP53* Database and Genomics Data: Human Mutation. *Hum Mutat*. sept 2016;37(9):865-76.
153. Coombs CC, Gillis NK, Tan X, Berg JS, Ball M, Balasis ME, et al. Identification of Clonal Hematopoiesis Mutations in Solid Tumor Patients Undergoing Unpaired Next-Generation Sequencing Assays. *Clin Cancer Res*. 2018;24(23):5918-24.
154. Hu Y, Ulrich BC, Supplee J, Kuang Y, Lizotte PH, Feeney NB, et al. False-Positive Plasma Genotyping Due to Clonal Hematopoiesis. *Clin Cancer Res*. 15 sept 2018;24(18):4437-43.
155. Couch FJ, Shimelis H, Hu C, Hart SN, Polley EC, Na J, et al. Associations Between Cancer Predisposition Testing Panel Genes and Breast Cancer. *JAMA Oncol*. 1 sept 2017;3(9):1190.
156. Packwood K, Martland G, Sommerlad M, Shaw E, Moutasim K, Thomas G, et al. Breast cancer in patients with germline *TP53* pathogenic variants have typical tumour

- characteristics: the Cohort study of TP53 carrier early onset breast cancer (COPE study). *J Pathol Clin Res.* 1 juill 2019;5(3):189-98.
157. Schwarzenbach H, Hoon DSB, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer.* juin 2011;11(6):426-37.
 158. Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* mars 2017;27(3):491-9.
 159. Castillo D, Herzog J, Sand S, O'Connor T, Clark C, Weitzel J Well-groomed participants: eyebrow plucks as surrogates for biomarker samples and a viable source of constitutional DNA. Annual Meeting of the American Society of Human Genetics, Vancouver, Canada, 18–22 October 2016. In.
 160. Weitzel JN, Chao EC, Nehoray B, Van Tongeren LR, LaDuca H, Blazer KR, et al. Somatic TP53 variants frequently confound germ-line testing results. *Genet Med.* août 2018;20(8):809-16.
 161. Coombs CC, Zehir A, Devlin SM, Kishtagari A, Syed A, Jonsson P, et al. Therapy-Related Clonal Hematopoiesis in Patients with Non-hematologic Cancers Is Common and Associated with Adverse Clinical Outcomes. *Cell Stem Cell.* sept 2017;21(3):374-382.e4.
 162. Takahashi K, Wang F, Kantarjian H, Doss D, Khanna K, Thompson E, et al. Preleukaemic clonal haemopoiesis and risk of therapy-related myeloid neoplasms: a case-control study. *Lancet Oncol.* janv 2017;18(1):100-11.
 163. Gillis NK, Ball M, Zhang Q, Ma Z, Zhao Y, Yoder SJ, et al. Clonal haemopoiesis and therapy-related myeloid malignancies in elderly patients: a proof-of-concept, case-control study. *Lancet Oncol.* janv 2017;18(1):112-21.
 164. Swisher EM, Harrell MI, Norquist BM, Walsh T, Brady M, Lee M, et al. Somatic Mosaic Mutations in *PPM1D* and *TP53* in the Blood of Women With Ovarian Carcinoma. *JAMA Oncol.* 1 mars 2016;2(3):370.
 165. Hsu JI, Dayaram T, Tovy A, De Braekeleer E, Jeong M, Wang F, et al. *PPM1D* Mutations Drive Clonal Hematopoiesis in Response to Cytotoxic Chemotherapy. *Cell Stem Cell.* nov 2018;23(5):700-713.e6.
 166. Wong TN, Ramsingh G, Young AL, Miller CA, Touma W, Welch JS, et al. Role of TP53 mutations in the origin and evolution of therapy-related acute myeloid leukaemia. *Nature.* févr 2015;518(7540):552-5.
 167. Ruark E, Snape K, Humburg P, Loveday C, Bajrami I, Brough R, et al. Mosaic *PPM1D* mutations are associated with predisposition to breast and ovarian cancer. *Nature.* 16 déc 2012;493(7432):406-10.
 168. Abelson S, Collord G, Ng SWK, Weissbrod O, Mendelson Cohen N, Niemeyer E, et al. Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature.* juill 2018;559(7714):400-4.

169. Desai P, Mencia-Trinchant N, Savenkov O, Simon MS, Cheang G, Lee S, et al. Somatic mutations precede acute myeloid leukemia years before diagnosis. *Nat Med.* 1 juill 2018;24(7):1015-23.
170. Desai P, Roboz GJ. Clonal Hematopoiesis and therapy related MDS/AML. *Best Pract Res Clin Haematol.* mars 2019;32(1):13-23.

ANNEXE

N° patient	Prélèvement sanguin					Frottis jugal		Tissu tumoral	
	NGS	Tapestation	Sanger (fait si VAF (sang) > 10 %)	Snapshot	Snapshot petits fragments (fait si Tapestation +)	Sanger (fait si VAF (sang) > 10 %)	Snapshot	NGS	Snapshot
1	A	A	A	A		A	A	A	NA
2	A	A	A	A	A	A	A	A	A
2	A	A	A	A	A	A	A	A	A
3	A	A		A			A	A	A
4	A	A	A	A		A	A	A	A
5	A	A		A			A	A	A
6	A	A	A	A	A	A	A	A	A
7	A	A	A	A		A	A	A	A
8	A	A		A	A		A	A	A
9	A	A	A	A		A	A	A	A
10	A	A		A			A	A	A
11	A	A		A			NR	A	A
12	A	A		A			A	NR	NR
13	A	A	A	A		NR	NR	NR	NR
14	A	A		A	A		A	NR	NR
15	A	A		A	A		A	NR	NR
16	A	A	A	A	A	A	A	NR	NR
17	A	A	A	NA		A	NA	NR	NR
18	A	A		A			A	A	A
19	A	A		A			NR	NR	NR
20	A	A		A	A		NR	NR	NR
20	A	A		A			NR	NR	NR
21	A	A		A			NR	NR	NR

Tableau 11 : Récapitulatif des analyses réalisées chez les différents patients

Abréviations : A : analysés ; NA : non analysable ; NGS : « Next Generation Sequencing » NR : Non reçu ; VAF : « Variant Allele Fraction »

Les analyses Snapshot à partir des prélèvements sanguins, des frottis buccaux et des tissus tumoraux (si disponibles) ont été réalisés pour l'ensemble des patients à l'aide du kit Snapshot Multiplex (Applied Biosystem ®). Cependant, les variants dont les VAF avaient des valeurs inférieures à 5 % dans le sang en NGS n'ont pas été présentés dans les résultats de Snapshot car le pic représentant l'allèle muté était trop proche du bruit de fond. Pour ces patients, le pic correspondant à l'allèle muté sur le Snapshot issu du frottis buccal était encore plus faible.

« Par délibération de son Conseil en date du 10 Novembre 1972, l'Université n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ou mémoires. Ces opinions doivent être considérées comme propres à leurs auteurs ».

VU, le Président de Thèse

VU, le Doyen de la Faculté

VU et permis d'imprimer
en référence à la délibération
du Conseil d'Université
en date du 14 Décembre 1973

Pour le Président
de l'Université de CAEN et P.O

Le Doyen

ANNEE DE SOUTENANCE : 2019

NOM ET PRENOM DE L'AUTEUR : BOULOUARD Flavie

TITRE DE LA THESE :

Signification des variants génétiques à faible ratio allélique détectés par séquençage à haut débit dans le cadre du diagnostic moléculaire des prédispositions aux cancers du sein et de l'ovaire : mosaïque, hématopoïèse clonale ou ADN tumoral circulant.

RESUME DE LA THESE EN FRANÇAIS :

Depuis que le Groupe Génétique et Cancer (GGC) recommande d'effectuer le diagnostic des prédispositions aux cancers du sein et de l'ovaire par l'étude de 13 gènes par séquençage à haut débit (notamment BRCA1, BRCA2 et TP53), des variants ayant un ratio allélique (ou VAF pour "variant allele fraction") faible sont détectés régulièrement posant la question du diagnostic différentiel de néomutation en mosaïque, d'hématopoïèse clonale (HC) et d'ADN tumoral circulant (ADNtc). Pour y répondre, une recherche des patients porteurs de variants de VAF faibles a été effectuée sur 2007 cas index. La présence de ces variants a été vérifiée par Sanger et/ou SNaPshot sur différents tissus disponibles tels que le sang, les cellules jugales et la tumeur. Parmi 20 patients porteurs d'un variant potentiellement pathogène présentant une VAF faible, ont été prouvés un cas de mosaïque de BRCA1, la présence d'ADNtc chez une patiente porteuse d'un variant de TP53 et de PTEN avec une VAF proche de 30 %, et 9 cas d'HC. Pour les autres patients, l'hypothèse était l'HC sous réserve de l'accessibilité à des tissus différents. Plus de femmes étaient atteintes d'un cancer de l'ovaire chez les porteuses de variants de TP53 de faible VAF (11/18) par rapport au groupe contrôle (310/1388) ($p < 0,001$) et elles étaient plus âgées (67,6 vs 55,7 ans ; $p < 0,001$). Cette étude met en évidence la complexité d'interpréter ces variants. Les états de mosaïque restent des événements rares qu'il faut démontrer par des analyses complémentaires avant de tester les descendants pour éviter les erreurs diagnostiques traumatisantes qu'induirait la présence d'une HC ou d'ADNtc notamment en cas de variant de TP53.

MOTS CLES : Mosaïque, hématopoïèse clonale, ADN tumoral circulant, BRCA1, TP53, cancer du sein, cancer de l'ovaire

TITRE DE LA THESE EN ANGLAIS :

Significance of genetic variants with low allelic fraction detected by high-throughput sequencing in the molecular diagnosis of predispositions to breast and ovarian cancers: mosaic, clonal hematopoiesis or circulating tumor DNA.

RESUME DE LA THESE EN ANGLAIS :

Since the new recommendations of the « Groupe Génétique et Cancer » (GGC) to perform a panel of 13 genes (including BRCA1, BRCA2 and TP53), for the diagnosis of breast and ovarian cancers predispositions, deleterious variants of low variant allele fraction (VAF) are regularly detected. It raises the question of the differential diagnoses of mosaic neomutation, clonal hematopoiesis (CH), and circulating tumor DNA (ctDNA). To answer this question, a search of patients with low VAF variants was performed on 2007 index cases analyzed by high throughput sequencing. The presence of these variants was verified by Sanger sequencing and/or SNaPshot on different available tissues such as blood, oral smear and tumor. Among the 20 patients with potentially pathogenic low VAF variants, it was possible to document a case of BRCA1 mosaic, to detect ctDNA in a patient carrying a TP53 and a PTEN variant with a VAF close to 30%, and 9 cases of CH. For the other patients, the preferred hypothesis was CH subject to accessibility to different tissues. More women were diagnosed with ovarian cancer in carriers of low VAF TP53 variants (11/18) relative to the control group (310/1388) ($p < 0.001$) and they were older (67.6 vs 55.7 years, $p < 0.001$). This study highlights the complexity of interpreting these variants. Mosaic remains a rare event that should be explored before testing the offspring in order to avoid traumatic diagnostic errors that would result from the presence of HC or ctDNA, particularly in the cases of TP53 variants.

KEY WORDS: Mosaic, clonal hematopoiesis, circulating tumor DNA, BRCA1, TP53, breast cancer, ovarian cancer