

Faculté de Médecine

Année 2021

Thèse N°

Thèse pour le diplôme d'État de docteur en Médecine

Présentée et soutenue publiquement

le 29 juin 2021

Par PERANI Alexandre

Né(e) le 28 février 1992 à Mulhouse

Oncogénétique théranostique du cancer du sein : étude des résultats de séquençage NGS de 103 patientes

Thèse dirigée par Madame le Professeur Catherine YARDIN

Examineurs :

Mme le Professeur Catherine YARDIN

Mme le Docteur Sylvie BOURTHOUMIEU

Mme le Docteur Karine DURAND

Mme le Docteur Françoise ESCLAIRE

Mme le Docteur Laurence VENAT-BOUVET

Présidente

Juge

Juge

Juge

Membre invitée



Faculté de Médecine

Année 2021

Thèse N°

Thèse pour le diplôme d'État de docteur en Médecine

Présentée et soutenue publiquement

Le 29 juin 2021

Par PERANI Alexandre

Né(e) le 28 février 1992 à Mulhouse

Oncongénétique théranostique du cancer du sein : étude des résultats de séquençage NGS de 103 patientes

Thèse dirigée par Madame le Professeur Catherine YARDIN

Examineurs :

Mme le Professeur Catherine YARDIN

Mme le Docteur Sylvie BOURTHOUMIEU

Mme le Docteur Karine DURAND

Mme le Docteur Françoise ESCLAIRE

Mme le Docteur Laurence VENAT-BOUVET

Présidente

Juge

Juge

Juge

Membre invitée

Professeurs des Universités - praticiens hospitaliers

Le 7 septembre 2020

ABOYANS Victor	CARDIOLOGIE
ACHARD Jean-Michel	PHYSIOLOGIE
AJZENBERG Daniel	PARASITOLOGIE et MYCOLOGIE
ALAIN Sophie	BACTERIOLOGIE-VIROLOGIE
AUBARD Yves	GYNECOLOGIE-OBSTETRIQUE
AUBRY Karine	O.R.L.
BERTIN Philippe	THERAPEUTIQUE
CAIRE François	NEUROCHIRURGIE
CHARISSOUX Jean-Louis	CHIRURGIE ORTHOPEDIQUE et TRAUMATOLOGIQUE
CLAVERE Pierre	RADIOTHERAPIE
CLEMENT Jean-Pierre	PSYCHIATRIE d'ADULTES
CORNU Elisabeth	CHIRURGIE THORACIQUE et CARDIOVASCULAIRE
COURATIER Philippe	NEUROLOGIE
DARDE Marie-Laure	PARASITOLOGIE et MYCOLOGIE
DAVIET Jean-Christophe	MEDECINE PHYSIQUE et de READAPTATION
DESCAZEAUD Aurélien	UROLOGIE
DES GUETZ Gaëtan	CANCEROLOGIE
DESSPORT Jean-Claude	NUTRITION
DRUET-CABANAC Michel	MEDECINE et SANTE au TRAVAIL
DURAND-FONTANIER Sylvaine	ANATOMIE (CHIRURGIE DIGESTIVE)
FAUCHAIS Anne-Laure	MEDECINE INTERNE
FAUCHER Jean-François	MALADIES INFECTIEUSES
FAVREAU Frédéric	BIOCHIMIE et BIOLOGIE MOLECULAIRE

FEUILLARD Jean	HEMATOLOGIE
FOURCADE Laurent	CHIRURGIE INFANTILE
GAUTHIER Tristan	GYNECOLOGIE-OBSTETRIQUE
GUIGONIS Vincent	PEDIATRIE
HANTZ Sébastien	BACTERIOLOGIE-VIROLOGIE
HOUETO Jean-Luc	NEUROLOGIE
JACCARD Arnaud	HEMATOLOGIE
JAUBERTEAU-MARCHAN M. Odile	IMMUNOLOGIE
JESUS Pierre	NUTRITION
LABROUSSE François	ANATOMIE et CYTOLOGIE PATHOLOGIQUES
LACROIX Philippe	MEDECINE VASCULAIRE
LAROCHE Marie-Laure	PHARMACOLOGIE CLINIQUE
LIENHARDT-ROUSSIE Anne	PEDIATRIE
LOUSTAUD-RATTI Véronique	HEPATOLOGIE
LY Kim	MEDECINE INTERNE
MABIT Christian	ANATOMIE
MAGY Laurent	NEUROLOGIE
MARIN Benoît	EPIDEMIOLOGIE, ECONOMIE de la SANTE et PREVENTION
MARQUET Pierre	PHARMACOLOGIE FONDAMENTALE
MATHONNET Muriel	CHIRURGIE DIGESTIVE
MELLONI Boris	PNEUMOLOGIE
MOHTY Dania	CARDIOLOGIE
MONTEIL Jacques	BIOPHYSIQUE et MEDECINE NUCLEAIRE
MOUNAYER Charbel	RADIOLOGIE et IMAGERIE MEDICALE
NATHAN-DENIZOT Nathalie	ANESTHESIOLOGIE-REANIMATION
NUBUKPO Philippe	ADDICTOLOGIE

OLLIAC Bertrand	PEDOPSYCHIATRIE
PARAF François	MEDECINE LEGALE et DROIT de la SANTE
PLOY Marie-Cécile	BACTERIOLOGIE-VIROLOGIE
PREUX Pierre-Marie	EPIDEMIOLOGIE, ECONOMIE de la SANTE et PREVENTION
ROBERT Pierre-Yves	OPHTALMOLOGIE
SALLE Jean-Yves	MEDECINE PHYSIQUE et de READAPTATION
STURTZ Franck	BIOCHIMIE et BIOLOGIE MOLECULAIRE
TCHALLA Achille	GERIATRIE ET BIOLOGIE DU VIEILLISSEMENT
TEISSIER-CLEMENT Marie-Pierre	ENDOCRINOLOGIE, DIABETE et MALADIES METABOLIQUES
TOURE Fatouma	NEPHROLOGIE
VALLEIX Denis	ANATOMIE
VERGNENEGRE Alain	EPIDEMIOLOGIE, ECONOMIE de la SANTE et PREVENTION
VERGNE-SALLE Pascale	THERAPEUTIQUE
VIGNON Philippe	REANIMATION
VINCENT François	PHYSIOLOGIE
YARDIN Catherine	CYTOLOGIE et HISTOLOGIE

PROFESSEUR ASSOCIE DES UNIVERSITES A MI-TEMPS DES DISCIPLINES MEDICALES

BRIE Joël	CHIRURGIE MAXILLO-FACIALE ET STOMATOLOGIE
KARAM Henri-Hani	MEDECINE D'URGENCE
MOREAU Stéphane	EPIDEMIOLOGIE CLINIQUE

MAITRES DE CONFERENCES DES UNIVERSITES - PRATICIENS HOSPITALIERS

BALLOUHEY Quentin	CHIRURGIE INFANTILE
BARRAUD Olivier	BACTERIOLOGIE-VIROLOGIE
BOURTHOUMIEU Sylvie	CYTOLOGIE et HISTOLOGIE

COUVE-DEACON Elodie	BACTERIOLOGIE-VIROLOGIE
DURAND Karine	BIOLOGIE CELLULAIRE
ESCLAIRE Françoise	BIOLOGIE CELLULAIRE
JACQUES Jérémie	GASTRO-ENTEROLOGIE ; HEPATOLOGIE
LE GUYADER Alexandre	CHIRURGIE THORACIQUE et CARDIOVASCULAIRE
LIA Anne-Sophie	BIOCHIMIE et BIOLOGIE MOLECULAIRE
RIZZO David	HEMATOLOGIE
TERRO Faraj	BIOLOGIE CELLULAIRE
WOILLARD Jean-Baptiste	PHARMACOLOGIE FONDAMENTALE

P.R.A.G.

GAUTIER Sylvie	ANGLAIS
-----------------------	---------

MAITRES DE CONFERENCES DES UNIVERSITES ASSOCIES A MI-TEMPS

SALLE Laurence	ENDOCRINOLOGIE (du 01-09-2020 au 31-08-2021)
-----------------------	---

PROFESSEUR DES UNIVERSITES DE MEDECINE GENERALE

DUMOITIER Nathalie	(Responsable du département de Médecine Générale)
---------------------------	--

MAITRE DE CONFERENCES ASSOCIE A MI-TEMPS DE MEDECINE GENERALE

HOUDARD Gaëtan	(du 01-09-2019 au 31-08-2022)
LAUCHET Nadège	(du 01-09-2020 au 31-08-2023)
PAUTOUT-GUILLAUME Marie-Paule	(du 01-09-2018 au 31-12-2020)
SEVE Léa	(du 01-09-2020 au 31-08-2023)

PROFESSEURS EMERITES

ADENIS Jean-Paul	du 01-09-2017 au 31-08-2021
ALDIGIER Jean-Claude	du 01.09.2018 au 31.08.2020
BESSEDE Jean-Pierre	du 01-09-2018 au 31-08-2020

BUCHON Daniel	du 01-09-2019 au 31-08-2021
MERLE Louis	du 01.09.2017 au 31.08.2020
MOREAU Jean-Jacques	du 01-09-2019 au 31-08-2021
TREVES Richard	du 01-09-2020 au 31-08-2021
TUBIANA-MATHIEU Nicole	du 01-09-2018 au 31-08-2021
VALLAT Jean-Michel	du 01.09.2019 au 31.08.2022
VIROT Patrice	du 01.09.2018 au 31.08.2021

Assistants Hospitaliers Universitaires – Chefs de Clinique

Le 12 juin 2020

ASSISTANTS HOSPITALIERS UNIVERSITAIRES

AUDITEAU Emilie	EPIDEMIOLOGIE (CEBIMER)
DAURIAT Benjamin	HISTOLOGIE, EMBRIOLOGIE ET CYTOGENETIQUE
DERBAL Sophiane	CHIRURGIE ANATOMIE
DOUCHEZ Marie	ANESTHESIOLOGIE-REANIMATION
DUPONT Marine	HEMATOLOGIE BIOLOGIQUE
DURIEUX Marie-Fleur	PARASITOLOGIE
GUYOT Anne	LABORATOIRE ANAPATHOLOGIE
HERMINEAUD Bertrand	LABORATOIRE ANAPATHOLOGIE
HUMMEL Marie	ANESTHESIOLOGIE-REANIMATION
LABRIFFE Marc	PHARMACOLOGIE
LEFEBVRE Cyrielle	ANESTHESIE REANIMATION
LOPEZ Stéphanie	MEDECINE NUCLEAIRE
PASCAL Virginie	IMMUNOLOGIE CLINIQUE
PIHAN Franck	ANESTHESIOLOGIE-REANIMATION
RIVAILLE Thibaud	CHIRURGIE-ANATOMIE
SANSON Amandine	ANESTHESIE REANIMATION
TCHU HOI NGNO Princia	BIOPHYSIQUE ET MEDECINE NUCLEAIRE

CHEFS DE CLINIQUE - ASSISTANTS DES HOPITAUX

ALBOUYS Jérémie	HEPATO GASTRO ENTEROLOGIE
ARMENDARIZ-BARRIGA Matéo	CHIRURGIE ORTHOPEDIQUE ET TRAUMATOLOGIQUE
AUBLANC Mathilde	GYNECOLOGIE-OBSTETRIQUE
BAÏSSE Arthur	REANIMATION POLYVALENTE

BEEHARRY Adil	CARDIOLOGIE
BLOSSIER Jean-David	CHIRURGIE THORACIQUE et CARDIOVASCULAIRE
BRISSET Josselin	MALADIES INFECTIEUSES ET TROPICALES
CHASSANG-BRUZEAU Anne-Hélène	RADIOLOGIE
CHAUVET Romain	CHIRURGIE VASCULAIRE
CISSE Fatou	PSYCHIATRIE
COMPAGNAT Maxence	MEDECINE PHYSIQUE et de READAPTATION
DE POUILLY-LACHATRE Anaïs	RHUMATOLOGIE
DESCHAMPS Nathalie	NEUROLOGIE
DESVAUX Edouard	MEDECINE GERIATRIQUE
DUVAL Marion	NEPHROLOGIE
EL OUAFI Zhour	NEPHROLOGIE
FAURE Bertrand	PSYCHIATRIE d'ADULTES
FAYEMENDY Charlotte	RADIOLOGIE et IMAGERIE MEDICALE
FROGET Rachel	CENTRE D'INVESTIGATION CLINIQUE (pédiatrie)
GEYL Sophie	GASTROENTEROLOGIE
GHANEM Khaled	ORL
GILBERT Guillaume	REANIMATION POLYVALENTE
GUTTIEREZ Blandine	MALADIES INFECTIEUSES
HANGARD Pauline	PEDIATRIE
HARDY Jérémy	CHIRURGIE ORTHOPEDIQUE ET TRAUMATOLOGIQUE
HESSAS-EBELY Miassa	GYNECOLOGIE OBSTETRIQUE
LALOZE Jérôme	CHIRURGIE PLASTIQUE
LEGROS Maxime	GYNECOLOGIE-OBSTETRIQUE
MAURIANGE TURPIN Gladys	RADIOTHERAPIE

MEUNIER Amélie	ORL
MICLE Liviu-Ionut	CHIRURGIE INFANTILE
MOWENDABEKA Audrey	PEDIATRIE
PARREAU Simon	MEDECINE INTERNE ET POLYCLINIQUE
PELETTE Romain	CHIRURGIE UROLOGIE et ANDROLOGIE
PEYRAMAURE Clémentine	ONCOLOGIE MEDICALE
PLAS Camille	MEDECINE INTERNE B
QUILBE Sébastien	OPHTALMOLOGIE
SIMONNEAU Yannick	PNEUMOLOGIE
SURGE Jules	NEUROLOGIE
TRICARD Jérémy	CHIRURGIE THORACIQUE et CARDIOVASCULAIRE MEDECINE VASCULAIRE
VAIDIE Julien	HEMATOLOGIE CLINIQUE
VERLEY Jean-Baptiste	PSYCHIATRIE ENFANT ADOLESCENT
VIDAL Thomas	OPHTALMOLOGIE

CHEF DE CLINIQUE – MEDECINE GENERALE

BERTRAND Adeline

SEVE Léa

PRATICIEN HOSPITALIER UNIVERSITAIRE

LERAT Justine O.R.L. (du 1er mai 2016 au 31 octobre 2020)

MATHIEU Pierre-Alain CHIRURGIE ORTHOPEDIQUE et
TRAUMATOLOGIQUE
(du 1er mai 2016 au 31 octobre 2020)

« I am aware that I am surrounded by people who feel that they could do the job better. Strong people with powerful characters, more natural leaders, perhaps better-suited to leading from the front, making a mark. But, for better or worse, the Crown has landed on my head. »

Claire FOY as *Queen Elizabeth II*, **The Crown**, S1 : E8, « *Pride and joy* », **Netflix**

Remerciements

A Mesdames les membres de ce jury,

A Madame le Professeur Catherine YARDIN

Chef du service de Cytogénétique, Génétique Médicale et Biologie de la Reproduction, CHU de LIMOGES

Au moment de nos discussions sur la composition du jury, je me souviens vous avoir un jour dit que mon parcours avait « particulièrement été influencé par les femmes ». C'est indéniablement vrai vous concernant. De ma première année, jusqu'au jour de la soutenance de ce travail, en passant par mon choix de spécialité à l'issue des ECN, je ne saurais vous être assez reconnaissant d'avoir toujours gardé un œil sur moi. Je vous remercie tout particulièrement pour vos précieux conseils, vos corrections apportées à ce travail, votre disponibilité, et votre bienveillance. Je vous remercie d'avoir accepté de me faire l'honneur de présider ce jury, et de diriger ce travail. J'espère qu'il saura refléter ce que vous m'avez apporté tant sur le plan professionnel que personnel tout au long de ces quatre années d'internat. Puisse-t-il vous faire honneur.

A Madame le Docteur Sylvie BOURTHOMIEU

Je vous remercie d'avoir accepté de juger mon travail. Je garde un excellent souvenir de vos cours d'histologie, qui m'ont passionné tout au long de mon premier cycle. Il s'en sera fallu de peu pour que je choisisse l'anatomopathologie au lieu de la génétique médicale ... mais je ne regrette pas mon choix. Merci infiniment pour vos compétences scientifiques et techniques extrêmement pointues, qui ont grandement influencé ce travail. Votre bienveillance, votre douceur et votre gentillesse m'ont particulièrement touché. J'espère pouvoir continuer à apprendre et collaborer à vos côtés pendant encore longtemps.

A Madame le Docteur Karine DURAND

Merci d'avoir accepté de participer à ce jury. Mon passage dans votre service a été bien trop court ... Je crois y avoir sincèrement passé un des plus difficiles, mais aussi et surtout un des meilleurs semestres de mon internat. J'ai été impressionné par votre PGMC, son expertise technique et biologique, tout autant que pour son personnel si incroyablement attachant. Merci pour ce semestre si particulier, aussi enrichissant sur tous les plans, pour toute cette bonne humeur, et ces connaissances que j'ai pu y acquérir. J'espère encore longtemps avoir l'occasion de discuter avec vous de ces si improbables tumeurs, et je serais déjà honoré d'avoir un jour acquis ne serait-ce que la moitié de vos connaissances en matière de cancérogenèse et de biologie tumorale.

A Madame le Docteur Françoise ESCLAIRE

Je vous remercie de participer à ce jury et de juger ce travail. Merci de m'avoir initié durant mon premier cycle d'étude à ce monde si fascinant de la biologie cellulaire et des techniques de cytogénétique. Je suis infiniment reconnaissant que mon parcours m'ait fait aboutir à ce service, que vous savez si bien habiter de votre bonne humeur et de votre personnalité aussi attachante. Merci d'avoir suivi de loin mon parcours tout au long de ces années, de mon premier cycle à aujourd'hui. J'espère avoir également pouvoir continuer à apprendre auprès de vous, aussi bien dans le domaine de la cytogénétique, que du si exigeant domaine de la Qualité.

A Madame le Docteur Laurence VENAT-BOUVET

Merci infiniment d'avoir accepté de participer à ce jury de thèse. C'est non sans une certaine fierté que ce travail a été réalisé, et il m'est impossible de mesurer et de restituer à quel point vous y avez contribué. Je vous suis reconnaissant de m'avoir transmis le goût de cette discipline si particulière, à l'intersection de l'oncologie et de la génétique. Merci également pour le semestre que j'ai pu effectuer en oncologie. Je garde un souvenir particulier de ces week-ends d'astreinte à vos côtés, et de vos impressionnantes compétences cliniques. J'espère sincèrement un jour maîtriser aussi bien ce domaine que vous ... Votre œil de clinicien m'est indispensable et j'espère devenir un meilleur généticien grâce à vous.

Aux praticiens de l'Institut Bergonié, qui m'ont si gentiment accueilli durant un an. **A Monsieur le Professeur Nicolas SEVENET**, à **Monsieur le Docteur Michel LONGY**, à **Mesdames les Docteurs Virginie BUBIEN et Anika BENSEN**, ainsi qu'aux **Docteurs Françoise BONNET-DORION, Natalie JONES, Pierre MACQUERE et Sabine RAAD**. Merci de m'avoir initié au monde passionnant de la génétique moléculaire et de l'interprétation des variants, cet art nécessitant de solides connaissances scientifiques. Merci aussi de m'avoir permis d'affiner mes compétences cliniques, auprès de vous et de vos adorables conseillères en génétique, **Angela BABIN, Anaïs DUPRE, et Eglantine JOLLY**.

A ma famille, ma mère bien évidemment, mon père, ainsi qu'à mes deux sœurs. J'espère que ce travail vous rendra fier. Merci de votre soutien durant ces longues années. C'est grâce à vous que je suis celui que je suis devenu aujourd'hui.

A Julien, pour ta patience et pour avoir d'avantage cru en moi que je n'y croyais moi-même. Merci de ton indéfectible soutien. Je ne saurais t'être assez reconnaissant d'avoir été là, à mes côtés dans les moments les plus difficiles... Mais tu as avant tout contribué aux plus beaux d'entre eux. Merci pour ces années à tes côtés, et pour ta tendresse. Tu es mon roc, mon pilier au quotidien.

Aux Docteurs Benjamin DAURIAT, Valentine MARQUET, et Clémentine PEYRAMAURE que je suis honoré de fréquenter et pour leurs collaborations si enrichissantes sur des dossiers aussi compliqués que passionnants. A **Séverine DEBORD**, amie toujours présente, et qui sait si bien m'aider à lutter contre ma phobie administrative, à **Marie-Line AUTEF**, dont le caractère et la gentillesse sont si attachants, et partie en retraite bien trop tôt à mon goût, à **Séverine VIGNAUD** modèle de gentillesse et de douceur, et pourtant redoutable d'efficacité, et à **Aurélie CARER**, pour sa douceur, et sa réactivité. **A Marion BARNIQUE et Anaïs GRANGIER**, pour m'avoir épaulé durant mes premiers pas d'interne, parfois difficiles.

A toutes les techniciennes que j'ai rencontrées, **Delphine, Isabelle, Nathalie, Patricia et Sandrine**, merci pour votre expertise technique et ces pauses déjeuner que j'ai prises avec vous durant tout mon internat. **A Maéva, Murat, Sylvain et Alain** qui ont si bien su nous accueillir durant un semestre avec **Maxence RAUTURIER**, le meilleur des co-internes. Merci de votre efficacité, de m'avoir transmis un peu de vos compétences, et pour nos pauses café-cigarettes, si agréables.

A mes amis, ici et ailleurs, pour ce premier cycle passé à vos côtés sur les bancs de la faculté, **aux infirmiers et aides-soignants d'oncologie** pour votre bonne humeur dans ce service si particulier, votre accueil et votre au revoir mémorables. **A mes frères et sœurs de la SIGF**, cette si attachante famille, et **à mes amis** aux quatre coins de la France.

Merci à toutes et à tous.

Droits d'auteurs

Cette création est mise à disposition selon le Contrat :

« **Attribution-Pas d'Utilisation Commerciale-Pas de modification 3.0 France** »

disponible en ligne : <http://creativecommons.org/licenses/by-nc-nd/3.0/fr/>



Table des matières

Liste des abréviations.....	19
Introduction	21
I. Cancer du sein, oncogénétique et génétique moléculaire.....	23
I.1. Epidémiologie	23
I.1.1. Incidence du cancer du sein en France	23
I.1.2. Mortalité par cancer du sein en France.....	24
I.1.3. Facteurs de risques et facteurs protecteurs du cancer du sein	25
I.2. Oncogénétique appliquée aux cancers du sein.....	28
I.2.1. Grands principes de la cancérogenèse.....	28
I.2.2. Principes fondateurs de l'oncogénétique	31
I.2.3. Syndrome HBOC ou syndrome sein-ovaire	32
I.2.3.1. Critères orientant vers le diagnostic de syndrome HBOC.....	34
I.2.3.2. Notions de bases sur la recombinaison homologue	35
I.2.3.3. <i>BRCA1</i> (NM_007294.3)	37
I.2.3.4. <i>BRCA2</i> (NM_000059.3).....	38
I.2.3.5. <i>PALB2</i> (NM_024675.3)	39
I.2.4. Autres syndromes et autres gènes	40
I.2.4.1. Syndromes de Li-Frauméni de type 1 et 2.....	41
I.2.4.2. Syndrome de Cowden et mutations du gène <i>PTEN</i> (NM_000314.4)	42
I.2.4.3. Syndrome de prédisposition lié au gène <i>CDH1</i> (NM_0044360.3).....	44
I.2.4.4. Ataxie télangiectasie et Ataxie télangiectasie-Like	45
I.2.4.5. Syndrome de Peutz-Jeghers lié au gènes <i>STK11</i> (NM_000455.4).....	47
I.2.4.6. Autres gènes étudiés dans le cadre du syndrome HBOC.....	48
I.2.4.6.1. Gènes où l'atteinte ovarienne est au premier plan	48
I.2.4.6.1.1. Gènes impliqués dans le syndrome de Lynch	49
I.2.4.6.1.2. Gènes des recombinaisons : homologues de <i>RAD51</i>	52
I.2.4.6.2. Gènes d'utilité clinique moins certaine	53
I.3. Principales techniques utilisées en oncogénétique moléculaire	55
I.3.1. Techniques qualitatives	57
I.3.1.1. Technique de Sanger	57
I.3.1.2. Le séquençage de nouvelle génération.....	59
I.3.2. Techniques quantitatives	63
I.3.2.1. Multiplex Ligation dependent probe amplification (MLPA)	63
I.3.2.2. Analyse chromosomique par puce à ADN (ACPA ou CGH)	65
I.3.3. Interprétation des variants	66
II. Article	71
Conclusion	97
Références bibliographiques	99
Serment d'Hippocrate.....	107

Table des illustrations

Figure 1 : Taux d'incidence des cancers du sein chez la femme entre 1990 et 2018.....	24
Figure 2 : Taux de mortalité par cancer du sein chez les femmes entre 1990 et 2018.....	25
Figure 3 : Représentation schématique des mécanismes de la cancérogenèse.	29
Figure 4 : Propriétés des cellules tumorales	30
Figure 5 : Théorie du double événement	31
Figure 6 : Représentation schématique des mécanismes de la recombinaison homologe..	37
Figure 7 : Représentation des protéines BRCA1, BRCA2 et PALB2 et de leur interactions..	40
Figure 8 : Représentation simplifiée de la voie PIK3-AKT-mTOR	43
Figure 9 : Positions chromosomique des gènes <i>EPCAM</i> et <i>MSH2</i>	51
Figure 10 : Exemple de séquençage par méthode de Sanger	57
Figure 11 : Electrophorégramme d'un séquençage Sanger par électrophorèse capillaire.....	58
Figure 12 : Grandes étapes du NGS	59
Figure 13 : Exemple de visualisation d'un BAM	62
Figure 14 : Description des amorces et des étapes utilisées en MLPA	63
Figure 15 : Résultats d'une ACPA, exemple du locus de l'EGFR.....	66

Table des tableaux

Tableau 1 : Facteurs de risques et facteurs protecteurs du cancer du sein.	27
Tableau 2 : Risques tumoraux associés au syndrome HBOC.....	33
Tableau 3 : Indication à une consultation d'oncogénétique pour syndrome HBOC	34
Tableau 4 : Critères individuels et familiaux à l'étude constitutionnelle HBOC	35
Tableau 5 : Critères diagnostique du syndrome de Li-Frauméni.....	41
Tableau 6 : Critères diagnostic du syndrome de Cowden.....	42
Tableau 7 : Risques tumoraux associés au syndrome de Lynch	49
Tableau 8 : Types de mutations et conséquences protéiques, exemple de <i>BRCA1</i>	55

Liste des abréviations

ACMG : *American College of Medical Genetics*

ACPA, CGH : Analyse Chromosomique par Puce à ADN, *Comparative Genomic Hybridisation*

ADN : Acide Désoxyribo-Nucléique

ARN : Acide Ribo-Nucléique

ATCD : Antécédents

BAM : *Binary Aligned Map*

BRC : *BRCA2 Repeat Domain*

BRCT : *BRCA1 C-Terminus Domain*

BRRS : *Bannayan-Riley-Ruvalvaca Syndrome*

CLI : Carcinome Lobulaire Infiltrant

CMMRD : *Constitutional Mismatch Repair Deficiency*

COP : Contraception Œstro-Progestative

DBD : *DNA Binding Domain*

ENIGMA : *Evidence-based Network for the Interpretation of Germline Mutant Alleles*

GGC : Groupe Génétique et Cancer

HAS : Haute Autorité de la Santé

HBOC : *Hereditary Breast-Ovarian Cancer*

HGDC : *Hereditary Gastric Diffuse Cancer*

HNPC : *Hereditary Non Polyposis Colon Cancer*

HR : *Homologous Recombination*, Recombinaison homologue

INCa : Institut National du Cancer

InSiGHT : *International Society for Gastrointestinal Hereditary Tumors*

IVG : Interruption volontaire de grossesse

Kb : milliers de paires de bases

LFS : *Li-Frauméni Syndrome*

LHS : *Left Hybridisation Sequence*

LOH : *Loss of heterozygosity*

LPO : *Left Probe Oligo*

MLPA : *Multiplex Lagation dependent Probe Amplification*

MMR : *Mismatch Repair*

MRN complex : *MRE11, RAD50, NBN complex*

MS-MLPA : *Methylation Specific Multiplex Ligatation dependent Probe Amplification*

NES : *Nuclear Exporting Signal*
NGS : *Next Generation Sequencing*
NLS : *Nuclear Localizing Signal*
NSCLC : *Non Small Cells Lung Carcinoma*
PARP : *Poly(ADP-Ribose) Polymerase*
PCR : *Polymerase Chain Reaction*
PHTRS : *PTEN Hamartoma Tumor Syndrome*
PLS : *Proteus-Like Syndrome*
PRTS : *PTEN-Related-Proteus Syndrome*
RE, ER : *Récepteurs des œstrogènes, Estrogen Receptors*
RHS : *Right Hybridisation Sequence*
RING : *Really Interesting New Gene domain*
RP : *Récepteurs de la Progestérone*
RPO : *Right Probe Oligo*
SA : *Semaines d'aménorrhée*
SAM : *Sequence Aligned Map*
SC : *Syndrome de Cowden*
SCST : *malignan Sex Cord Stromal Tumor of ovary*
SIR : *age-ajusted Standard Incidence Ratio*
SNC : *Système Nerveux Central*
THSM : *Traitement Hormonal Substitutif de la Ménopause*
TSM : *Taux d'incidence Standardisé sur la population Mondiale*
TSV : *Tab Separated Values*
UV : *rayonnements Ultra-Violets*
VAM : *Variation Annuelle Moyenne*
vcf : *variant calling file*
VOUS, VSI : *Variant Of Unknown Significance ; Variant de Signification Inconnue*

Introduction

L'étude des caractéristiques génétiques d'un individu atteint de cancer, et/ou de celles de la tumeur qu'il présente, est devenue ces dernières années une étape importante dans la prise en charge de nombreux cancers. Dans le cas du cancer du côlon par exemple, l'étude de certains gènes de la voie RAS/MAP-Kinase et du gène *BRAF* (en particulier la recherche du variant « V600E »), permet de prédire la réponse ou l'absence de réponse aux traitements anti-EGFR.

Longtemps, dans le cancer du sein, les seules caractéristiques tumorales dont le clinicien disposait pour adapter sa stratégie thérapeutique, étaient la présence ou l'absence de certains récepteurs aux hormones (œstrogènes et progestérone) à la surface tumorale, complété par la suite par la présence ou l'absence de surexpression d'un proto-oncogène, le gène *ERBB2* (ou codant pour la protéine Her2). Ces informations permettent, en plus de connaître le pronostic d'une patiente en classant sa tumeur dans un groupe standardisé, de savoir si elle peut tirer un bénéfice de l'administration d'une hormonothérapie (en cas de présence de récepteurs aux hormones à la surface de la tumeur), ou d'un traitement ciblant la protéine Her2 (trastuzumab ou HERCEPTIN ©).

À présent, la donne est en train de changer concernant la prise en charge du cancer du sein. À l'image du cancer de l'ovaire, la prescription d'une nouvelle classe de médicaments, les inhibiteurs de PARP (Olaparib ou LYNPARZA ©), est possible pour certains cancers du sein localement avancés ou en situation métastatique. L'utilisation de ces traitements, qui se sont révélés être une arme thérapeutique précieuse dans les cancers de l'ovaire chez des patientes porteuses d'une altération constitutionnelle ou tumorale des gènes *BRCA1* et *BRCA2*, est maintenant possible dans la prise en charge du cancer le plus fréquent de la femme en France, et dans le monde plus généralement.

Les avancées dans le domaine des biotechnologies de ces dernières années ont permis de s'affranchir de la méthode Sanger, méthode certes éprouvée, mais qualifiée d'étude à bas débit car permettant l'étude dans le même temps technique d'une seule région de 100 à 1200 bp environ. Actuellement, l'étude de ces gènes se pratique selon une méthode de séquençage de nouvelle génération (ou NGS, par amplification des régions d'intérêt et séquençage des bibliothèques d'amplicons après enrichissement), où sont étudiés en panel de nombreux gènes. Il est ainsi plus facile de repérer, parmi l'immense nombre de cancers du sein diagnostiqués en France chaque année, ceux qui sont liés à une prédisposition héréditaire.

Concernant le cancer du sein, les études se font à l'heure actuelle sur échantillon d'ADN constitutionnel. Le panel de gènes utilisé comprend cependant des gènes n'ayant aucun rapport avec ce type de cancer. Ainsi, lorsqu'une étude pour recherche de sensibilité aux inhibiteurs de PARP est prescrite par un oncologue prenant en charge une femme atteinte d'un cancer du sein, de nombreux autres gènes sont étudiés.

À côté des gènes dont l'étude est recommandée lorsqu'une femme présente un cancer du sein (*BRCA1*, *BRCA2*, et *PALB2* entre autres), d'autres informations deviennent accessibles concernant des gènes prédisposant à de tout autres types de cancers comme les cancers digestifs (gènes *MMR* du syndrome de Lynch, ou gènes *MUTYH* et *APC* des polyposes, etc...) ou concernant des gènes dont l'implication médicale en oncologie n'est actuellement pas clairement établie (comme *CHEK2*, *BARD1*, ou encore *BRIP1*). Enfin, l'approche en panel conduit également à avoir des informations concernant des gènes que l'on ne retrouve que

très rarement altérés à l'état constitutionnel (comme le gène *TP53*), ou des gènes retrouvés dans des tumeurs très spécifiques ou des syndromes de prédisposition héréditaire très peu fréquents, comme le gène *PTCH1*, impliqué dans le médulloblastome et le syndrome de Gorlin par exemple.

Toutes ces données non-sollicitées, que l'on qualifie d' « incidentes », ne sont actuellement transmises au prescripteur qu'en cas d'intérêt médical clairement établi. L'identification d'un variant délétère ou probablement délétère (de classe 5 ou de classe 4), au sein d'un gène responsable du syndrome de Lynch par exemple, sera ainsi transmis à l'oncologue même si la recherche initiale concernait une patiente qui ne présentait pas de cancer digestif, mais un cancer du sein.

Une question se pose cependant. Que peut-on faire des autres données que l'on observe ? L'étude nationale TUMOSPEC, actuellement menée à l'initiative du groupe Génétique et Cancer (GGC) tente de répondre à cette question, en ayant pour ambition d'évaluer le risque de développer un cancer du sein ou de l'ovaire, en étudiant des familles au sein desquelles un variant délétère est identifié dans un gène autre que ceux du syndrome HBOC (*BRCA1*, *BRCA2* ou *PALB2*).

Cependant, le cancer du sein est une pathologie fréquente, et il paraît justifié de remettre en cause le système de pensée, selon lequel on considère que seules de rares altérations délétères puissent en être à l'origine au sein de familles où de nombreux cas sont observés.

En ce sens, l'objectif de ce travail est de réaliser une étude observationnelle visant à décrire les altérations moléculaires retrouvées au sein d'un panel de gènes analysés par approche NGS, chez un groupe de patientes présentant un cancer du sein, et ce, à visée thérapeutique.

I. Cancer du sein, oncogénétique et génétique moléculaire

I.1. Epidémiologie

Dans le monde, le cancer du sein concerne chaque année entre **1,4 et 2 millions de femmes**, et le taux de mortalité observé est estimé à environ **450.000 décès par an**. [1;2]

En 2018, la plus forte incidence était retrouvée en Belgique, avec un taux standardisé d'incidence à 113,2 nouveaux cas pour 100.000 personnes-années.

I.1.1. Incidence du cancer du sein en France

En février 2019, l'Institut National du Cancer (INCa) a rendu publique les résultats de son étude portant sur les registres du cancer tenus entre 1990 et 2018. [3] Ces données constituent un outil précieux pour mieux prévenir et prendre en charge le cancer du sein en France. En ce sens, cette démarche s'inscrit pleinement dans le cadre du troisième plan cancer 2014-2019.

Cette étude montre qu'en 2018, près de **soixante mille** (58.459) nouveaux cas de cancer du sein ont été diagnostiqués chez des femmes de plus de 15 ans. Comparativement à 1990, ce nombre a pratiquement doublé (+95%), puisqu'en 1990 près de trente mille (28.489) nouveaux cas de cancer du sein étaient recensés. Le cancer du sein suit la tendance globale de l'augmentation de l'incidence des cancers de la femme en France, toute localisation confondue. Il reste le cancer le plus fréquent, puisqu'il représente près d'un nouveau cas sur trois (32,9% des 177.433 nouveaux cas de cancer). Le Taux Standardisé d'incidence sur la population Mondiale (TSM) du cancer du sein est de **99,0 nouveaux cas pour 100.000 personnes-années en 2018**. Ce nombre est de loin supérieur au TSM du second cancer de la femme, le cancer colorectal (TSM à 23,9), et du troisième cancer, le cancer du poumon (TSM à 23,3). La France se classe donc au quatrième rang mondial, après la Belgique, première avec 113,2 nouveaux cas pour 100.000 personnes-années, le Luxembourg, deuxième avec un TSM à 109,3, et les Pays-Bas, troisième avec un TSM à 105,9. [1]

Sur près de trente ans, la progression du nombre de nouveaux cas de cancer du sein n'est cependant pas linéaire. Les taux d'incidence standardisés de plusieurs années (1990, 1995, 2000, 2005, 2010, 2015 et 2018), ont été calculés, et il apparaît **deux périodes** distinctes. [Figure 1]

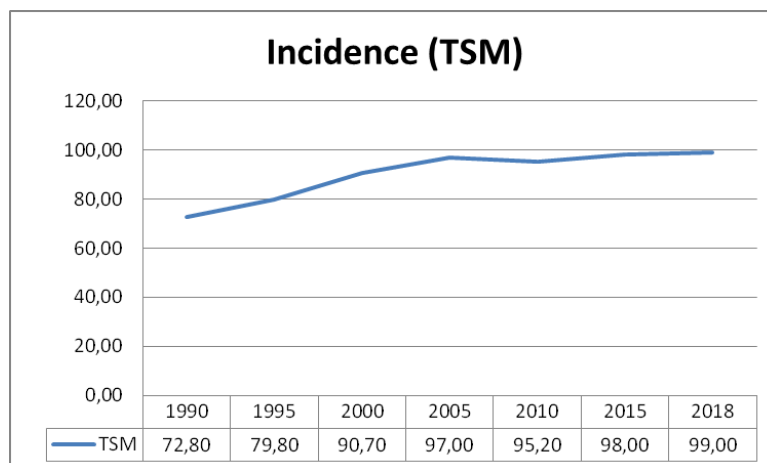


Figure 1 : Taux d'incidence des cancers du sein chez la femme entre 1990 et 2018. Exprimé en taux standardisé sur la population mondiale (TSM). D'après [2]

- Entre 1990 et 2005, on observe une forte augmentation du nombre de nouveaux cas de cancers du sein. Sur cette première période, son passe de 72,8 nouveaux cas pour 100.000 personnes-années à 97,0, soit une augmentation brute de 33,24% en 15 ans pour une variation annuelle moyenne (VAM) de **+2,02% par an**.
- Entre 2010 et 2018, on observe une stabilisation du nombre de nouveaux cas. L'augmentation de l'incidence est moins importante, passant de 95,2 nouveaux cas de cancers du sein pour 100.000 personnes-années en 2005, contre 99,0 en 2018. Cela représente une augmentation brute de 4,94% sur cette période, soit une VAM de **+0,6 % par an**.

La plus forte augmentation d'incidence concerne les années 1995 à 2000 (+13,66% de nouveaux cas pour 100.000 personnes-années), soit une VAM de 2,73% par an. C'est plus du double de la VAM sur la période entière du rapport, calculée à +1,1% entre 1990 et 2018. Mais le nombre de nouveaux cas sur une année n'est cependant pas une information suffisante pour décrire complètement l'épidémiologie du cancer du sein en France. En effet, en étudiant le nombre de nouveaux cas de cancer du sein recensés en 2018 selon des tranches d'âge de 10 ans, il apparaît une période de la vie d'une femme, particulièrement à risque. Située entre les âges de 50 et 69 ans, elle survient donc chez les femmes après leur **ménopause**, et regroupe à elle seule presque la moitié (45,5%) de tous les nouveaux cas de cancer du sein diagnostiqués en 2018.

I.1.2. Mortalité par cancer du sein en France

Si le taux d'incidence du cancer du sein entre 1990 et 2018 est en augmentation, celui de la mortalité est **en diminution**. [3]

De manière similaire à celle du taux d'incidence, les auteurs rapportent le taux de mortalité par cancer du sein entre 1990 et 2018, avec les mêmes années repères [Figure 2].

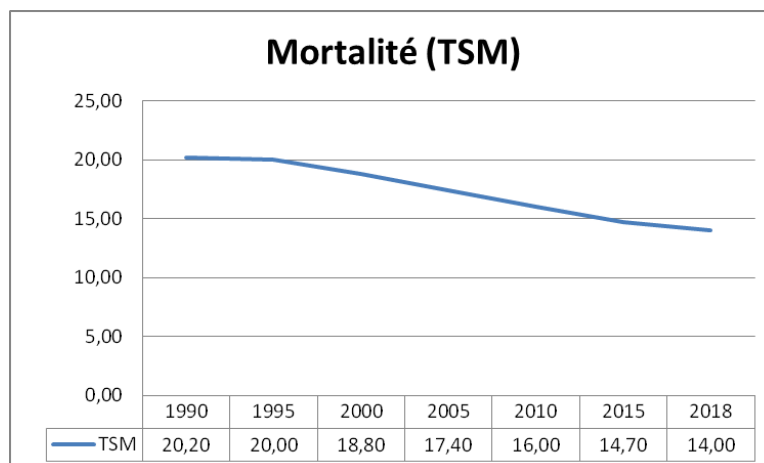


Figure 2 : Taux de mortalité par cancer du sein chez les femmes entre 1990 et 2018. Exprimé en taux standardisé sur la population mondiale (TSM). D'après [2]

Ainsi, on observe un net recul de la mortalité par cancer du sein sur la période 1990-2018, passant d'un **taux de 20,2 décès** par cancer du sein pour 100.000 personnes-années en 1990, à un **taux de 14,0** en 2018. En termes de VAM du taux de mortalité, cela correspond à une VAM de **-1,3% par an** sur la période 1990-2018.

Entre 1990 et 1995, le TSM de la mortalité par cancer du sein est plutôt stable, mais la diminution est quasi-linéaire sur la période suivante. Cette diminution est encore plus accentuée sur la période plus récente allant de 2010-2018. La VAM sur cette période est estimée à -1,6% par an. Exprimé en nombre de décès pour l'année 2018, cela représente 12.146 décès par cancer du sein sur les 67.817 décès par cancer tout type confondu, soit 17,9%.

Le cancer du sein est donc le cancer le plus fréquent chez la femme, mais également le cancer entraînant le plus de décès, avec un TSM à 14,0 décès pour 100.000 personnes-année, à égalité avec le cancer du poumon. Et ce, bien que le cancer du poumon soit en terme d'incidence le troisième cancer chez la femme (TSM à 23,3), derrière le cancer colorectal (TSM à 23,9), qui lui entraîne une mortalité plus faible (TSM à 6,9).

I.1.3. Facteurs de risques et facteurs protecteurs du cancer du sein

Compte-tenu de la forte incidence du cancer du sein à travers le monde, de nombreuses études épidémiologiques ont été menées pour tenter d'identifier les facteurs de risques et les facteurs protégeant du risque de développer une tumeur maligne mammaire. Plusieurs catégories de facteurs peuvent être individualisées, avec deux grands groupes : celui des facteurs de risques non-modifiables (endogènes), et celui des facteurs de risques modifiables (exogènes). L'ensemble de ces facteurs sont repris dans le tableau 1.

- Concernant les facteurs de risques non-modifiables, la première catégorie est celle des facteurs de risques endogènes généraux. Ils regroupent **l'âge** (le risque augmente en fonction de l'âge, et il est maximal autour de la ménopause. Il diminue ou reste constant ensuite), le **sexe féminin**, les **antécédents** personnels et familiaux de cancer du sein, les antécédents personnels de pathologie du sein, les **facteurs génétiques** (que le gène incriminé soit connu ou non), et le fait d'avoir une **densité mammaire importante**. [4]
- La seconde catégorie est celle des facteurs de risques endogènes hormonaux, concernant les principaux paramètres liés aux hormones sexuelles chez la femme, le cancer du sein en population générale étant majoritairement hormono-dépendant (présence de récepteurs aux hormones à la surface tumorale). Dans cette catégorie, se retrouve surtout **l'âge tardif de la ménopause** (après 50 ans), ainsi que **l'âge précoce de la puberté** (inférieur à 12 ans). [3 ;5;6] Certains facteurs liés à la grossesse entrent également cette catégorie, car ils ne sont pas entièrement modifiables. Parmi eux, les **grossesses multiples** (gémellaires *versus* mono-fœtales), **celles non-menées à terme** (avec un accouchement avant 33 SA), ainsi que **l'âge tardif au premier accouchement** (plus de 30 ans). [6]
- Les facteurs de risque exogènes peuvent être également déclinés selon ces mêmes catégories. [7] Dans les facteurs de risque généraux, qui concernent également plusieurs autres types de cancers, on peut lister **l'obésité** et le **surpoids** [8], qui peuvent fréquemment s'associer au **diabète de type II**, ou encore la **consommation d'alcool et de tabac**.
- Concernant les facteurs de risque hormonaux, le choix d'une **contraception œstroprogestative** est un facteur de risque connu, de même que la **prise d'un traitement hormonal substitutif de la ménopause**. A noter que ce sur-risque décroît, voire disparaît, 5 à 10 ans après l'arrêt de la contraception ou du traitement substitutif.
- La catégorie « environnement et le mode de vie », dans sa composante modifiable, comprend essentiellement **le fait de travailler de nuit**, ou une **alimentation pauvre en fruits et légumes** et riche en produits ultra-transformés. Peuvent également y être inclus le **statut socio-économique**, conditionnant souvent ces facteurs, **l'exposition aux radiations** (dans le cadre de certains métiers, ou de certaines pathologies comme le lymphome de Hodgkin), et **l'exposition à la pollution atmosphérique**, pour les femmes vivant en zone très densément peuplées. [5;7]

Les facteurs protecteurs vis-à-vis du cancer du sein sont malheureusement moins nombreux. En revanche, ils sont majoritairement modifiables (exogènes), comme la **pratique d'une activité physique régulière**. Cela concerne aussi bien les femmes durant la période pré-ménopausique que post-ménopausique de leur vie, comme montré par une récente méta-analyse. [8] Des facteurs environnementaux peuvent également protéger du cancer du sein. Au même titre qu'une alimentation suffisante en fruit et légumes, le fait d'avoir un taux de vitamine D suffisant est un facteur protecteur. [7] Enfin, certains facteurs hormonaux semblent protéger du cancer du sein. Ils sont aussi bien exogènes, comme le choix **d'allaiter son enfant** (allaiter pendant 13 mois diminuerait le risque de développer un cancer du sein de 50%), qu'endogènes, comme le fait d'avoir des **cycles menstruels réguliers**. [7]

A côté de ces facteurs de risques et de ces facteurs protecteurs, d'autres facteurs sont plus controversés, et ne font pas l'objet de consensus clairs, certains étant discutables. Parmi eux, se retrouvent entre autre le groupe sanguin, ou encore une durée insuffisante de sommeil. Pour certains, il est licite de se demander s'il ne s'agit pas de facteurs de confusion, comme par exemple le groupe sanguin, sa répartition n'étant pas homogène au sein de la population générale.

Tableau 1 : Facteurs de risques et facteurs protecteurs du cancer du sein.

	Facteurs de risque	Facteurs protecteurs	Facteurs controversés
Facteurs endogènes			
Généraux	Sexe féminin âge > 50 ans (6,5) ATCD de cancer (5,4-11) ATCD sénologiques (1,5-4,4) Facteurs génétiques Densité des seins (1,8-6)		Groupe sanguin
	Hormonaux	Puberté précoce (1,2-1,5) Ménopause tardive (1,5-2,0)	Cycles menstruels réguliers
Liés à la grossesse	Multi-fœtales terme < 33 SA		
Facteurs exogènes			
Généraux	Obésité/surpoids (1,9 à 5,7) Diabète de type II Alcool (1,3-2,5) et tabac		
Hormonaux	COP THSM (1,5)	Durée de lactation	Recours à l'IVG Médicaments stimulant l'ovulation
Environnement et mode de vie	Pollution et irradiation* (1-20) Travail de nuit Alimentation Statut socio-économique	Activité physique Taux de vitamine D	Consommation de café Durée de sommeil

ATCD = Antécédents, SA = Semaine d'aménorrhée, COP = Contraception Œstroprogestative, THSM = Traitement hormonal substitutif de la ménopause, IVG = Interruption volontaire de grossesse. Les risques relatifs sont présentés entre parenthèse lorsqu'ils sont clairement identifiés.
* en cas d'irradiation thoracique pour maladie de Hodgkin. D'après [5-8]

I.2. Oncogénétique appliquée aux cancers du sein

Chaque année donc, ce sont près de 60.000 cas de cancers du sein qui sont diagnostiqués en France. [1] Il n'est donc pas rare d'observer des familles dans lesquelles il existe **plusieurs femmes atteintes, et ce, sur plusieurs générations**.

Pour l'immense majorité, il ne s'agit que de la simple **association fortuite** de plusieurs cas de cancers sporadiques, en lien avec la forte incidence de cette pathologie dans la population générale. Seule une faible proportion est génétiquement favorisée. **Environ 10% sont en lien avec la transmission intrafamiliale d'une altération moléculaire dans des gènes de prédisposition**, que cette altération soit connue ou non. [3] Pour les femmes concernées, le risque de développer un cancer du sein est alors très augmenté tout au long de la vie par rapport aux femmes exposées uniquement aux autres facteurs de risque. [9]

Parmi les gènes impliqués, *BRCA1* et *BRCA2* ont été les premiers décrits, en 1994 et 1995 respectivement. [10; 11] **Une altération dans l'un de ces deux gènes est retrouvée dans près d'une famille suspecte de forme héréditaire sur six**. [12] Depuis, bien d'autres gènes ont été associés à une augmentation du risque de tumeurs mammaires, qu'elles soient au premier ou au second plan de la symptomatologie. Les gènes *TP53*, *CDH1*, ou encore *PTEN* n'en sont que des exemples. [13 ; 4] Les altérations responsables ségrègent selon le mode **autosomique dominant**, mais certains gènes ont la caractéristique d'induire des phénotypes extrêmes en cas d'inactivation bi-allélique constitutionnelle, à l'image du CMMRD ou de l'anémie de Fanconi. Pour la part restante des cancers héréditaires, en l'absence de variant délétère identifié, l'association de plusieurs cas semble être le témoin du partage de gènes de susceptibilité et de facteurs environnementaux communs entre les différents membres d'une même famille.

L'identification d'une altération moléculaire au sein des gènes de prédisposition aux cancers du sein est aujourd'hui d'une grande utilité. Pour une patiente atteinte de cancer, ces informations peuvent conduire le clinicien à modifier sa prise en charge, avec des impacts potentiels sur le traitement chirurgical, ou le traitement des patientes en phase métastatique par exemple. Mais elle a également une conséquence sur le plan familial, puisqu'elle permet d'identifier au sein d'une famille les personnes à risques de développer un cancer avant qu'ils ne l'aient présenté. Il leur sera alors proposé un suivi médical différent de celui réalisé en population générale, guidé par les risques tumoraux associés. En ce sens, l'oncogénétique s'intègre donc pleinement dans une **démarche de diagnostic de précision et de médecine personnalisée**.

I.2.1. Grands principes de la cancérogenèse

Les connaissances sur la cancérogenèse ont largement progressé ces dernières années. Initialement décrite à l'aide de modèles assez simples essentiellement centrés sur le génome des cellules atteintes, il apparaît clair aujourd'hui que les **cellules saines avoisinantes** (stromales et immunitaires) ont également un rôle fondamental. [14] Un parallèle en ce sens

peut être fait avec les mécanismes de l'embryogenèse. En effet, la différenciation des cellules souches migrantes embryonnaires se fait en fonction des cellules stromales qu'elles colonisent, par le biais de facteurs de croissances, par exemple.

Schématiquement, le développement d'un cancer résulte de l'échec de la réparation des **lésions qui surviennent au sein du génome d'une cellule**. Ces lésions, peuvent avoir diverses origines, notamment virale dans le cas des cancers hépatiques liés aux virus des hépatites ; ou encore environnementale comme les radiations ionisantes (UV) qui peuvent conduire au développement d'un cancer de la peau (mélanome). De l'accumulation de ces lésions, il résulte une instabilité génomique entretenue notamment à **l'activation d'oncogènes** (impliqués dans la synthèse de facteurs de croissance par exemple), et à **l'inactivation simultanée de gènes suppresseurs de tumeurs**, *gate keepers* et *care takers*, ayant respectivement pour rôle le contrôle du cycle cellulaire et le maintien de l'intégrité du génome. [Figure 3]

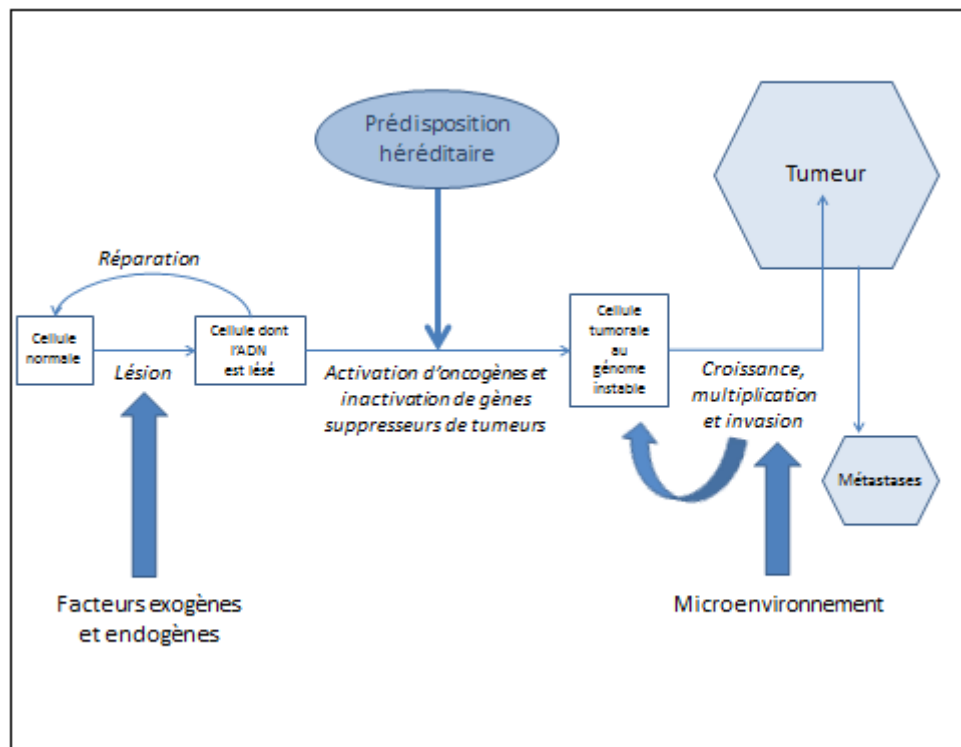


Figure 3 : Représentation schématique des mécanismes de la cancérogenèse. (D'après [15-16])

Cet **échappement aux mécanismes d'homéostasie tissulaire** permet à la cellule tumorale de se multiplier. [15]. En plus d'avoir acquis la capacité à échapper à la réponse immunitaire anti-tumorale, la cellule est capable de produire de manière non-régulée des facteurs de croissance. Les facteurs stimulant la néo-angiogenèse qu'elle synthétise par ailleurs, permettent de recruter de nombreuses cellules immunitaires qui génèrent un climat pro-inflammatoire au niveau du microenvironnement tumoral riche en facteurs de croissance,

utilisés par les cellules tumorales pour croître. Il en résulte finalement une capacité d'invasion non seulement au niveau local, mais également à distance, sous forme de métastases. [16] Les cellules cancéreuses au sein des tumeurs acquièrent donc un certain nombre de propriétés, pour la plupart en détournant à leur profit les voies de signalisations qui à l'état physiologique permettent l'homéostasie cellulaire. [Figure 4]

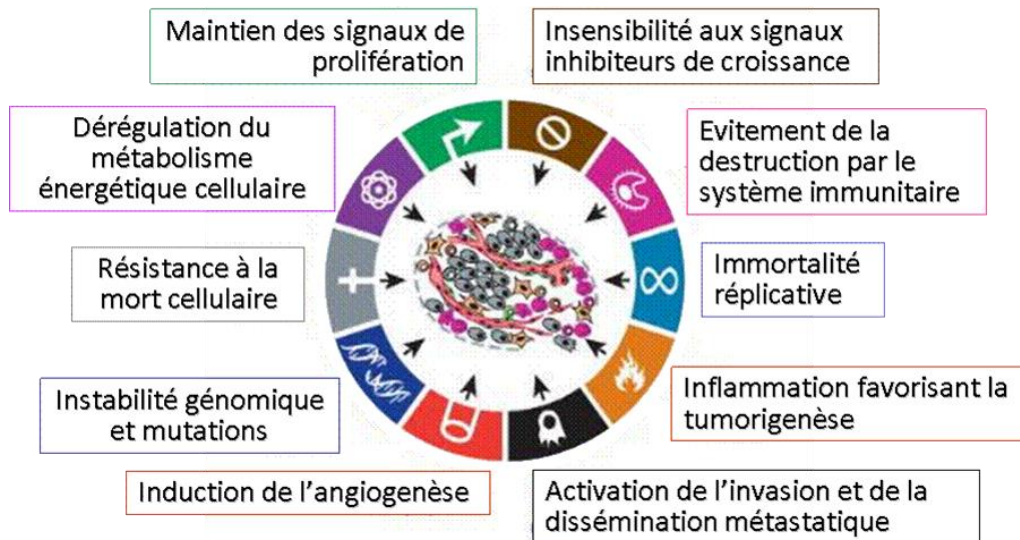


Figure 4 : Propriétés des cellules tumorales (D'après [15])

Pour exemple, on peut citer les voies PI3K/AKT/mTOR et RAS/MAP-Kinase, activées physiologiquement par la liaison à l'EGFR de son ligand. Dans certains cas de cancers du poumon, *EGFR* est lui-même muté, entraînant une activation constitutive indépendante du ligand, permettant aux cellules tumorales de stimuler l'angiogenèse et de proliférer tout en perdant leur sensibilité à l'apoptose.

Enfin, toutes les cellules cancéreuses au sein des tumeurs **n'ont pas le même patrimoine génétique**, car elles ne sont pas issues de la prolifération d'un seul et unique clone cellulaire. C'est de cette manière que l'on peut expliquer **l'apparition de résistances aux thérapies ciblées**. En ce sens un autre parallèle peut être fait, cette fois avec les bactéries. Lors d'une infection, l'utilisation d'antibiotiques peut conduire à l'apparition de résistances, par pression de sélection. Une antibiothérapie conduite de manière non-optimale conduira ainsi à sélectionner les bactéries qui y sont résistantes, et qui pourront proliférer à l'arrêt du traitement. En cancérologie, c'est l'exemple de la mutation p.(Thr790Met) du gène de l'EGFR, conduisant à l'apparition d'une résistance aux inhibiteurs de tyrosine kinase (Erlotinib, Gefitinib) utilisés dans le cadre du traitement du cancer du poumon non à petites cellules (NSCLC) [17].

I.2.2. Principes fondateurs de l'oncogénétique

L'oncogénétique est une discipline assez récente. Elle considère les cas, où **il existe, à l'état constitutionnel, une altération au sein des proto-oncogènes ou des gènes suppresseurs de tumeurs.**

L'exemple classique dans le domaine est celui du rétinoblastome, décrit par Knudson à partir d'études statistiques en 1971 [18], et dont le raisonnement est repris et expliqué par Bonaiti-Pellier et Feingold en 1990. [19] Partant des travaux de ses prédécesseurs, qui avaient remarqué que les formes bilatérales ou multifocales étaient plus rares que les formes unilatérales, et que 50% de la descendance de ces sujets étaient eux-mêmes atteints de rétinoblastome, ils en conclurent que ces formes étaient génétiquement favorisées. Afin d'expliquer ces différents constats, ils formulèrent l'hypothèse que deux événements génétiques distincts étaient nécessaires au développement de ce type de tumeur :

- une première mutation survenant soit de manière **constitutionnelle** (et donc présente dans toutes les cellules des deux rétines), soit de **manière acquise** (atteignant seulement quelques cellules);
- une seconde mutation, **acquise**, aboutissant à la perte de fonction homozygote d'un anti-oncogène ou à l'activation homozygote d'un proto-oncogène. [Figure 5]

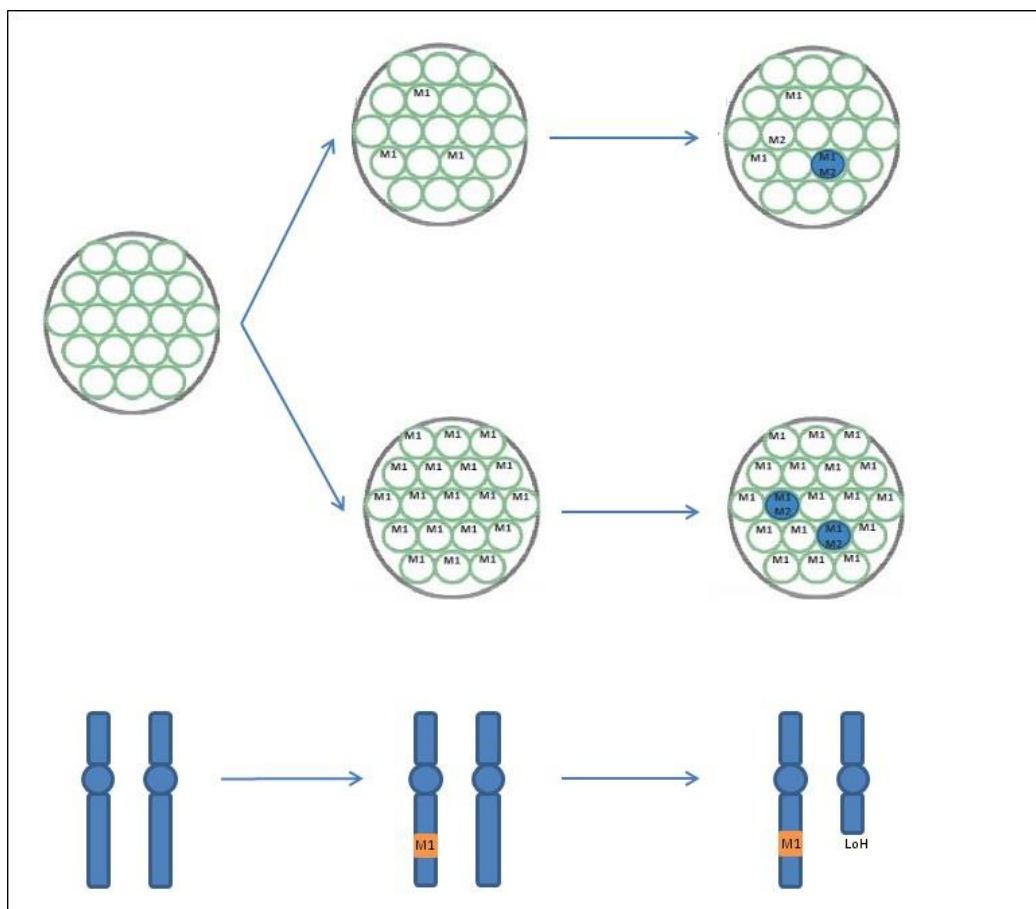


Figure 5 : Théorie du double événement (D'après [18-19])

En génétique tumorale, si le premier évènement est majoritairement une **altération de la séquence du gène**, le second évènement peut soit être de même nature soit être une **perte d'hétérozygotie** (ou LOH pour *Loss Of Heterozygosity*). Ce mécanisme de LOH permet donc l'inactivation complète des gènes suppresseurs de tumeurs, par perte de l'allèle sauvage, et conservation de l'allèle altéré. [20] Dans le cas des proto-oncogènes, ce n'est pas à proprement parler une perte d'hétérozygotie qui permet le développement du cancer, mais souvent une **amplification** parfois très importante du gène, comme c'est le cas dans le cancer du sein où il peut exister une amplification du gène *ERBB2* (Her2).

Ainsi, on explique pourquoi les formes héréditaires sont volontiers plus bilatérales ou multifocales que les formes sporadiques, puisque la survenue de deux événements indépendants au sein de la même cellule dans chacune des deux rétines (ou dans plusieurs cellules de la même rétine) est extrêmement peu probable.

Ce n'est que plus tard, en 1987 que l'observation de rétinoblastomes en lien avec une micro-délétion au niveau du locus 13q14 ont permis l'identification et le clonage du gène *RB1*, responsable de cette prédisposition héréditaire. [21] Depuis, de nombreux autres syndromes de prédisposition aux cancers ont été décrits, impliquant différents organes – seins, ovaires, côlon, etc. – ouvrant le champ d'une médecine vaste, dans lequel les connaissances scientifiques en matière de génétique sont un précieux atout.

I.2.3. Syndrome HBOC ou syndrome sein-ovaire

Le syndrome « HBOC » (pour *Hereditary Breast and Ovarian Cancer*) est responsable d'une importante augmentation du risque de développer un **cancer du sein** chez la femme comme chez l'homme (incluant également le risque de développer une seconde atteinte primitive homo ou controlatérale), ainsi que du risque de développer un **cancer de l'ovaire** (également de la trompe et du péritoine), chez les individus porteurs d'une altération constitutionnelle au sein des gènes *BRCA1*, *BRCA2*, ou *PALB2*. [22 ;23]

D'autres risques tumoraux sont décrits, **variables en fonction du gène**, notamment au niveau prostatique chez l'homme, ainsi que pancréatiques ou dermatologiques quel que soit le sexe de l'individu. Plusieurs auteurs ont estimé le risque de développer un cancer chez les individus porteurs d'une altération constitutionnelle des gènes *BRCA1*, et *BRCA2*. [24 – 31]. De description plus récente, les risques associés aux altérations du gène *PALB2* sont moins connus, mais une récente étude portant sur 524 familles a permis d'apporter des premières données, méritant d'être approfondies à l'avenir. [32] Ces risques sont repris dans le tableau suivant [tableau 2].

Tableau 2 : Risques tumoraux associés au syndrome HBOC (D'après [24-31])

Type de cancer	BRCA1	BRCA2	PALB2
Sein (femme)	46-87%	38-84%	25-66%
Sein (seconde atteinte)	21,1% (à 10 ans)	10,8% (à 10 ans)	-
Sein (homme)	1,2%	8,9%	0,9%
Ovaire	39-63%	16,5-27%	0,6-4,2%
Prostate	8,6% (à 65 ans)	15% (à 65 ans)	-
Pancréas	1-3%	2-7%	2,2-2,8%
Mélanome	-	5%	-

Certaines caractéristiques histologiques ou cliniques sont un peu plus fréquemment rencontrées dans ce syndrome.

- Concernant les cancers de l'ovaire et de la trompe, il s'agit **d'adénocarcinomes séreux** [33], volontiers de haut grade.
- Concernant le cancer du sein c'est classiquement le phénotype tumoral dit « **triple négatif** », c'est-à-dire n'exprimant ni récepteurs hormonaux (œstrogènes RE, progestérone RP), ni surexpression de l'oncogène *ERBB2* (protéine Her2), qui est associé aux mutations des gènes HBOC. Plusieurs études ont tenté de préciser cette association. Ainsi, une importante étude portant sur 1824 patientes atteintes d'un cancer du sein de ce phénotype, triple négatif donc, non sélectionnées sur leurs antécédents familiaux, a permis de retrouver chez **10% d'entre elles** une mutation au sein des gènes *BRCA1* (8.5%), *BRCA2* (2.7%) ou *PALB2* (1.2%). [34]

De manière intéressante, lorsqu'est étudié le phénotype des tumeurs présentées par les patientes portant une mutation au sein de ces gènes, il apparaît que 80% des tumeurs ont un phénotype triple négatif chez les patientes *BRCA1* mutées, alors que cette surreprésentation phénotypique n'est pas retrouvée chez les patientes *BRCA2* mutées. [35].

Dernier élément remarquable, et permettant souvent de suspecter ce syndrome, **l'âge au diagnostic de la première tumeur** est plus précoce chez les femmes *BRCA1* mutées et *BRCA2* mutées, comparé aux femmes issues de la population générale. Si on s'appuie sur les données issues de l'étude de Eerola et al. portant sur 262 tumeurs (51 issues de patientes appartenant à une famille où *BRCA1* est muté, 59 issues de familles *BRCA2* muté, et 152 issues de familles sans mutation *BRCA1* ou *BRCA2*), il apparaît que l'âge moyen au diagnostic est de **44 ans** pour *BRCA1* et de **47 ans** pour *BRCA2*, alors que pour les patientes issues de la population générale, cet âge moyen était de 56 ans. [36]

I.2.3.1. Critères orientant vers le diagnostic de syndrome HBOC

La HAS a proposé des critères permettant l'orientation des femmes vers une consultation d'oncogénétique. L'indication à l'étude des gènes *BRCA1*, *BRCA2* et *PALB2* est ensuite étudiée par le clinicien, en fonction de l'histoire personnelle et familiale, et des caractéristiques phénotypiques de la tumeur présentée par la patiente. [37] Une patiente est donc orientée vers une consultation d'oncogénétique sur son histoire personnelle si elle a présenté :

- un adénocarcinome du sein avant l'âge de 30 ans
- Un cancer du sein associé à un cancer de l'ovaire, de la trompe, ou du pancréas
- Des cancers multiples, dont un cancer du sein et deux autres (hors cancers ORL, du poumon ou du col utérin)

Pour ce qui est de son histoire familiale, la HAS recommande le **calcul du score d'Eisinger** [38] détaillé ci-dessous [tableau 3]. Les femmes sont orientées vers une consultation d'oncogénétique si elles ont **un score supérieur ou égal à trois**, après addition des points calculés séparément dans chacune des deux branches parentales.

Tableau 3 : Indication à une consultation d'oncogénétique pour syndrome HBOC (ou score d'Eisinger ; D'après HAS 2011)

Critère	Score
Mutation <i>BRCA1</i> ou <i>BRCA2</i> identifiée dans la famille	5
Cancer du sein chez une femme avant 30 ans	4
Cancer du sein chez une femme entre 30 et 40 ans	3
Cancer du sein chez une femme entre 40 et 50 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	3

Après évaluation de l'histoire personnelle et familiale par le médecin oncogénéticien, l'indication d'une étude constitutionnelle des gènes est retenue ou non, sur la base de critères développés dans le tableau 4, repris depuis le site de l'institut du cancer dans l'article d'Odile Cohen-Haguenaer. [39]

Tableau 4 : Critères individuels et familiaux à l'étude constitutionnelle HBOC

Critères familiaux

Au moins trois cas de cancers du sein chez des apparentés au premier ou second degré, au sein de la même branche familiale

Deux cas de cancers du sein chez les apparentés au premier ou second degré *via* un homme, si :

Cancer du sein avant 40 ans ou cancer bilatéral

Cancer du sein chez un homme

Cancer du sein et un cancer de l'ovaire

Deux cancers de l'ovaire

Au moins deux cas de cancers du sein ou de l'ovaire et au moins un cancer du pancréas chez les apparentés au premier et second degré au sein de la même branche familiale

Critères individuels

Cancer du sein avant 36 ans

Cancer du sein et un cancer de l'ovaire chez la même patiente

Cancer du sein de phénotype triple négatif avant 51 ans

Cancer du sein chez un homme

Cancer de l'ovaire ou tubaire épithélial avant 70 ans, sans limite d'âge si séreux de haut grade

I.2.3.2. Notions de bases sur la recombinaison homologue

Il paraît difficile d'étudier les gènes de prédisposition aux cancers du sein, sans évoquer, même brièvement, les différents mécanismes de réparation des lésions de l'ADN. Schématiquement, plusieurs patterns de lésions sont décrits :

- la modification chimique des bases (méthylation, oxydation), insertion d'agents alkylants
- la formation de pontages intra-brin ou inter-brin
- les mésappariements de bases
- les cassures simple et double brin

Pour maintenir l'intégrité de son patrimoine génétique, la cellule possède **plusieurs mécanismes de réparation** faisant intervenir de nombreuses protéines. [40] Les mutations des gènes codant pour ces protéines sont fréquemment rencontrées en pathologie cancéreuse humaine. Si les mutations des gènes intervenant dans la correction des mésappariements (ou *Mismatch Repair, MMR*) prédisposent plutôt aux cancers digestifs dans le cadre du syndrome de Lynch (*cf infra*), la majorité des gènes impliqués dans la prédisposition au cancer du sein interviennent eux, dans la **réparation des cassures double brin** par le mécanisme de **recombinaison homologue** (*Homologous Recombination, HR*). Ce mécanisme a particulièrement bien été décrit par Li et Heyer en 2008 [41], bien que RAD51, protéine centrale du mécanisme et une des premières décrites, ait été identifiée bien des années avant, en 1993 (*cf infra*).

Brièvement, lors de la formation d'une cassure double brin de l'ADN, les extrémités de la cassure sont digérées, et la réparation se fait **à partir de la séquence intacte portée par la chromatide sœur**. Trois grandes étapes sont décrites. [Figure 6]

- La première est la phase dite « **présynaptique** ». Durant cette étape, la cassure est détectée, les différentes protéines nécessaires à la réparation recrutées et les complexes protéiques assemblés. C'est également durant cette phase que le site de cassure est partiellement dégradé de part et d'autres du site de la coupure sur plusieurs kilobases. Les différentes protéines impliquées sont MRE11, RAD50 et NBN au sein du complexe MRN, à la fois détecteur des lésions, et effecteur de la résection. D'autres protéines interviennent également dans cette dernière fonction : il s'agit des protéines BRCA1 et BLM entre autres.
- La seconde phase est celle dite « **synaptique** », où la séquence homologue est recherchée, et où une boucle de réplication se forme. Cette étape nécessite la formation d'un filament pour « capter » la chromatide sœur et stabiliser la boucle de réplication. Ce filament est formé par un homopolymère de protéine RAD51, après recrutement des monomères par un complexe formé par les protéines PALB2, BRCA2 et RAD51C. Le filament ainsi formé est lui-même stabilisé par la protéine BRCA1 associée à la protéine BARD1. [42]
- Enfin, la dernière phase est la phase « **post-synaptique** » où la synthèse d'ADN est effectuée, permettant de reconstituer l'intégrité des brins lésés, à partir de la séquence intacte de la chromatide sœur.

A noter que lors de la phase présynaptique, en parallèle de la résection des extrémités rompues, le complexe MRN active les protéines ATM et ATR, qui vont elles-mêmes recruter CHEK2 et p53 pour stopper le cycle cellulaire et favoriser l'apoptose en cas d'échec de réparation. [43]

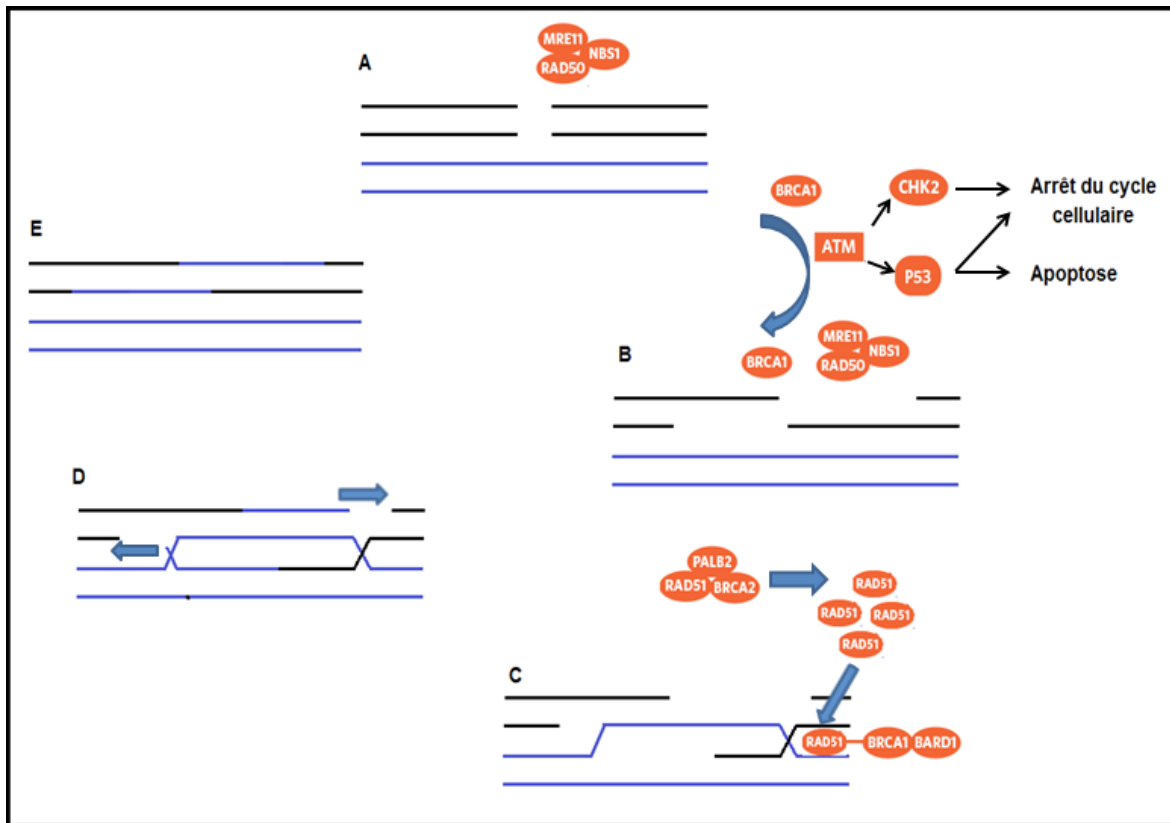


Figure 6 : Représentation schématique des mécanismes de la recombinaison homologue (D'après [42-43])

I.2.3.3. *BRCA1* (NM_007294.3)

Le gène *BRCA1* est situé sur le bras long du chromosome 17, au locus 17q21.31. Orienté en anti-sens, il s'étend des positions génomiques g.41279500 à g.41194312, pour une taille totale d'environ 85,2 Kb. Il comprend 24 exons, d'une taille totale d'environ 7,8 Kb, dont 23 sont codants. [44]

A l'heure actuelle, un très grand nombre de variants délétères a été identifié au sein de ce gène, **sans réelle mutation fréquente, ni hotspot**. Par ailleurs, en 2012, Eggington et son équipe ont montré que près de 3% de l'ensemble des variants reportés au sein des gènes *BRCA1* et *BRCA2* étaient de signification inconnue (de classe 3).

Comme développé précédemment, la protéine *BRCA1* joue un rôle central dans le processus de **recombinaison homologue** [45], par interaction directe avec d'autres protéines, comme *BARD1* et *BRIP1* ; mais également dans la **régulation du cycle cellulaire** en recrutant la protéine *ATM*. Elle possède une dernière fonction d'**inhibition de la synthèse d'ARN messagers** issus de la transcription de gènes lésés. [46]

Elle comporte 1863 résidus [47], avec trois grands domaines fonctionnels décrits.

- Le premier est nommé **RING**, et se situe en position N-terminale. Il est constitué par les acides aminés 24 à 64. C'est ce premier domaine qui est impliqué dans la régulation de la synthèse des ARN messagers, grâce à sa fonction ubiquitine-ligase. En effet, le recrutement de BRCA1 aux sites de lésions de l'ADN entraîne la dégradation de l'ARN polymérase II, empêchant ainsi la synthèse de protéines aberrantes issus de gènes lésés. Cette fonction est effectuée à l'aide de la protéine BARD1, avec laquelle BRCA1 se lie par l'intermédiaire de leur domaines RING respectifs. [46]
- Le deuxième domaine est plus structural que fonctionnel. Il s'agit d'une conformation protéique particulière, appelée **Coiled-coil**, ou super-hélice, constituée par plusieurs hélices alpha enroulées les unes autour des autres. Ce domaine est constitué par les résidus 1397 et 1424, et permet l'interaction de BRCA1 avec la protéine codée par PALB2, par l'intermédiaire de leur domaine *Coiled-coil* respectifs.
- Le troisième domaine, est constitué de deux structures nommées **BRCT1** et **BRCT2** (pour *BRCA1 C-terminal Domain*), et permet le recrutement et la localisation de la protéine vers les zones lésées de l'ADN à réparer. Ils sont situés en C-terminal de la protéine, et sont constitués par les résidus 1642 à 1736 pour le premier, et 1756 à 1855 pour le second. C'est grâce à ces domaines que ce forment les hétéro-dimères formés par la protéine BRCA1, lorsqu'elle est associée à la protéine BRIP1, par association de leurs domaines BRCT respectifs.

Enfin, la protéine BRCA1 porte **trois courtes séquences**, permettant son importation et son exportation vers le compartiment nucléaire. Ces séquences signal sont appelées NLS pour *Nuclear localizing signals*, et NES pour *Nuclear exporting signal*.

La transcription de *BRCA1* s'effectue majoritairement durant les phase G1 et S du cycle cellulaire [48]. La protéine est exprimée au niveau du noyau, en association avec les protéines BRCA2 et RAD51 au niveau des sites de lésion à l'ADN [49].

De manière expérimentale, l'inactivation bi-allélique de la protéine chez les souris conduit à une mortalité embryonnaire, due en partie à une incapacité à réparer les lésions à l'ADN, mais également à une instabilité génomique de manière plus globale, qui se traduit par des dommages chromosomiques similaires à ceux observés dans l'**anémie de Fanconi**. [50 ;51] Ce phénotype extrême, en cas d'inactivation bi-allélique du gène *BRCA1*, peut également être observé en pathologie humaine.

1.2.3.4. **BRCA2 (NM_000059.3)**

Contrairement à *BRCA1*, *BRCA2* est orienté en sens, sur le chromosome 13, au locus 13q13.1. Il est de plus grande taille que le gène *BRCA1*, puisqu'il s'étend entre les positions génomiques g.32887617 et g.32975809 (soit une taille totale de 88,2 Kb). Il comprend 27 exons pour une taille totale d'environ 10,4 Kb.

Comme pour *BRCA1*, on dénombre environ 1800 variants délétères reportés dans les différentes bases de données, et parmi tous les variants décrits, 3% environ sont de classe 3.

La protéine BRCA2 est constituée de 3418 résidus. Associée aux protéines PALB2 et RAD51, elle joue un rôle majeur dans le processus de **recombinaison homologue**, puisqu'elle permet le recrutement et la polymérisation sous forme de filament de la protéine RAD51, nécessaire à la formation de la boucle de réplication, filament que BRCA1 stabilise. Il faut cependant noter que BRCA2 entre également en interaction avec de nombreuses autres, dont p53, et participe ainsi au contrôle du cycle cellulaire et de l'apoptose. [52].

Elle possède trois grands domaines fonctionnels :

- le premier étant pour la **liaison avec la protéine PALB2**, constituées par les résidus 1 à 40.
- Le second domaine est en réalité constitué par **un motif de 35 acides aminés répétés huit fois**. Ils sont nommés BRC1 à BRC8. Les quatre premiers, constitués par les résidus 1002 à 1551 permettent l'interaction avec la protéine RAD51, et les 4 suivants, constitués par les résidus 1564 à 2085, permettent la régulation de l'activité de la protéine RAD51. [53; 54]
- Le dernier domaine fonctionnel individualisable est le domaine **DBD** (pour *DNA Binding Domain*), constitué par les résidus 2479 à 3192. [55]

Tout comme BRCA1, BRCA2 est exprimée durant les phases G1 et S du cycle cellulaire [56]; et les souris KO *Brca2* montrent les mêmes anomalies que celles constatées chez les souris *Brca1-KO* [50; 51]).

I.2.3.5. PALB2 (NM_024675.3)

Le dernier gène responsable du syndrome HBOC est le gène *PALB2*, de description plus récente. Il s'agit d'un petit gène, situé au locus 16p12.2 et s'étendant de la position g.23654678 à la position g.23612483, soit une taille totale d'environ 42,1 Kb. Il comporte 13 exons, ayant une taille totale de 4 Kb environ.

La première description de l'association entre cancer du sein le gène *PALB2* est issue de l'étude de Rahman en 2007. [57] En étudiant 923 femmes atteintes de cancer pour lesquelles aucune mutation *BRCA1* ou *BRCA2* n'était retrouvée, ils ont mis en évidence 10 variants tronquants au sein du gène *PALB2*. Cette étude, s'appuie sur une étude antérieure, qui étudiait les interactions de la protéine BRCA2 au niveau de la chromatine avec d'autres protéines. C'est ainsi qu'a été découverte la protéine PALB2, qui co-localisait avec la protéine BRCA2, d'où son nom de partenaire et localiseur de BRCA2 (*Partner and Localizer of BRCA2*). [58]

Enfin, comme pour BRCA1 et BRCA2, des modèles animaux ont été créés, afin d'étudier le phénotype des souris *Palb2-KO*. [59] Expérimentalement, il a été montré qu'une inactivation complète de PALB2 chez la souris conduisait aux mêmes conséquences que celles observées chez les souris *Brca1* et *Brca2 KO*, à savoir une létalité embryonnaire précoce.

La protéine PALB2 que ce gène code, comprend 1186 résidus, répartis en trois grands domaines fonctionnels individualisables. [45]

- Le premier domaine que l'on peut décrire se situe au niveau de résidus 1 à 579, il permet l'**association de la protéine à l'ADN**.
- Au sein du premier domaine, on peut en décrire un second, constitué par les résidus 1 à 319, permettant l'**association de la protéine PALB2 aux protéines BRCA1, et RAD51 lorsqu'elle est active**, et PALB2 elle-même pour inactivation de la protéine par dimérisation. Ce domaine est constitué d'un motif structural de type *Coiled-coil*.
- Le dernier domaine, C-terminal, nommé WD40 est constitué par les résidus 853 à 1186 permet l'**interaction avec les protéines BRCA2 et RAD51**. [60]

L'ensemble de ces informations, sont résumées dans la figure suivante [figure 4].

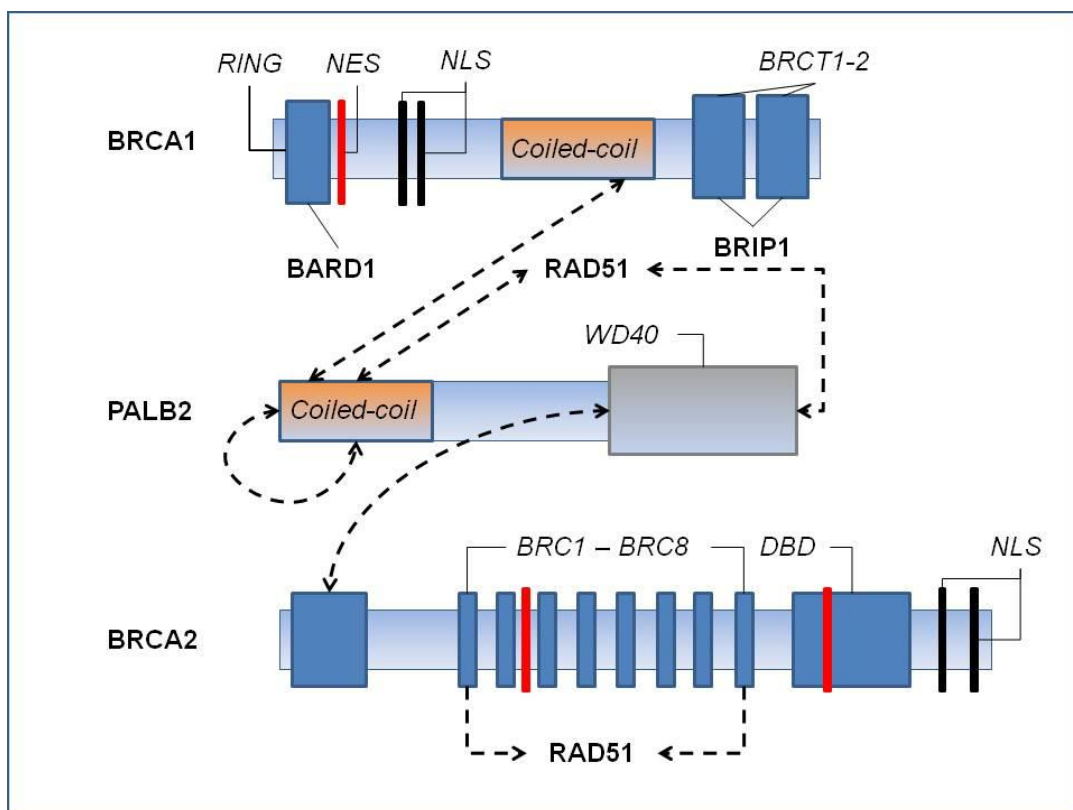


Figure 7 : Représentation des protéines BRCA1, BRCA2 et PALB2 et de leur interactions. D'après [45]
 Les interactions entre protéines sont représentées par les flèches discontinues. NES = Nuclear Exporting Signal, NLS = Nuclear Localizing Signal, DBD = DNA Binding Domain

I.2.4. Autres syndromes et autres gènes

D'autres gènes augmentant le risque de développer une tumeur maligne du sein sont décrits. Pour la plupart, ils appartiennent à des **syndromes distincts**, associés à des spectres tumoraux différents, avec pour certains des expressions phénotypiques dépassant le cadre de l'oncologie. Par ailleurs, lorsqu'ils sont responsables de cancers du sein, les caractéristiques histopathologiques sont le plus souvent différentes, avec une représentation des tumeurs triple négatives moindre.

I.2.4.1. Syndromes de Li-Frauméni de type 1 et 2

Le syndrome de Li-Frauméni est un syndrome rare, mais particulier en oncogénétique, caractérisé par 1) **l'extrême précocité des atteintes** et 2) **la grande hétérogénéité des organes pouvant être touchés**. [61] Deux formes ont été décrites, différant par le gène en cause.

- Gène TP53 (NM_000546.5)

Le gène TP53 est situé au locus 17p13.1. C'est un petit gène, qui s'étend des positions génomiques g.7592868 à g.7569720 (environ 23,1 Kb), et qui comprend 17 exons. Il code pour la protéine p53, qui joue un rôle central dans plusieurs grandes fonctions cellulaires. Elle est impliquée dans **l'arrêt du cycle cellulaire et dans l'apoptose**, ainsi que la **réplication de l'ADN et la réparation des lésions qu'il subit** [62].

A l'état basal, le taux de protéine p53 libre est maintenu à un faible niveau, par liaison avec la protéine MDM2. Et cas de stress cellulaire, cette liaison est rompue permettant de rendre p53 disponible, et elle va s'accumuler au niveau du noyau pour effectuer ses fonctions. [63] En oncogénétique, une altération du gène TP53 est suspectée lorsqu'un même individu présente **plusieurs cancers** (métachrones ou synchrones), ou lorsqu'un **grand nombre d'individus** au sein d'une même famille sont atteints. En réalité, la mise en évidence d'une altération constitutionnelle est rare, seulement **400 familles** ayant été rapportées dans la littérature, et des critères diagnostics, dits de Chompret [Tableau 5], ont été proposés pour faciliter l'identification des sujets atteints [64].

Tableau 5 : Critères diagnostique du syndrome de Li-Frauméni

- 1 Survenue d'une tumeur appartenant au spectre LFS (cancer du sein avant la ménopause, sarcome des tissus mous, ostéosarcome, corticosurréalome, tumeur cérébrale, cancer du poumon, leucémie) avant l'âge de 46 ans

- 2 Au moins un parent au premier ou second degré avec un tumeur du spectre LFS avant l'âge de 56 ans ou des cancers multiples (sauf cancer du sein avant la ménopause si présenté par le cas-index)

- 3 Individu avec tumeurs multiples, sauf si deux des tumeurs appartiennent au spectre LFS avec la première survenue après 56 ans

- 4 Individu ayant présenté soit un corticosurréalome, soit une tumeur des plexus choroïdes, quelle que soit l'histoire familiale.

Le diagnostic de syndrome de Li-Frauméni est ainsi retenu, sur l'association des critères 1 et 2, ou sur l'unique présence des critères 3 ou 4. Cependant, il a été décrit des patients suspects d'être atteints de ce syndrome, sans altération *TP53* retrouvée, ayant permis l'identification d'une seconde forme de ce syndrome, liée cette fois au gène *CHEK2*.

- Gène CHEK2 (NM_007194.1)

Le gène CHEK2 est situé au locus 22q12.1 et s'étend des positions génomiques g.29139822 à g.29081731, et comprend 22 exons. Bien que découvert depuis 1998, le rôle de cette protéine n'a été décrit que deux ans plus tard, en 2000, grâce aux études réalisées par l'équipe menée par Chebab [65]. En effet, cette équipe a montré que CHEK2 permettait l'**arrêt du cycle cellulaire** en phase G1, en réponse à la présence de lésions à l'ADN. Cette activité, inhibée en cas de mutation du gène *TP53*, a permis de mettre en évidence la position de CHEK2 dans la cascade conduisant à l'arrêt du cycle cellulaire, en amont de TP53 donc. La même année, l'équipe de Lee a décrit l'interaction de cette protéine avec BRCA1, également impliquée dans la réparation des lésions à l'ADN. [66]

I.2.4.2. Syndrome de Cowden et mutations du gène PTEN (NM_000314.4)

Bien que rare, le syndrome de Cowden (SC) est un syndrome classique en oncogénétique tant il est reconnaissable. Il est dû aux altérations constitutionnelles du gène *PTEN* [67] situé en 10q22.23. Ses critères diagnostics ont initialement été décrits par le Consortium international du syndrome de Cowden en 1995, et ont été révisés par Pilarsky et Eng en 2004 [68] ; [Tableau 6].

Tableau 6 : Critères diagnostic du syndrome de Cowden (D'après [68])

Critères pathognomoniques	Gangliocytome dysplasique du cervelet (Lhermitte Duclos) Atteintes dermatologiques spécifiques (papillomatose orofaciale, trichilemmomes, kératose acrale)
Critères majeurs	Atteinte tumorale : cancer du sein, de l'endomètre, cancer papillaire de la thyroïde Macrocéphalie
Critères mineurs	Pathologies non cancéreuses : nodules thyroïdiens, lipomes ; fibromes cutanés et utérins ; mastopathie fibrokystique Tumeurs génito-urinaires, Polypes hamartomateux digestifs Déficience intellectuelle, autisme

Son diagnostic est retenu en cas de signes **pathognomoniques**, ou sur l'association à la fois de **signes majeurs** et de **signes mineurs**. Un élément important à souligner concernant les critères diagnostics du SC, est la présence de **polypes hamartomateux**, également appelés polypes « juvéniles », terme cependant à éviter car potentiellement confusogène. *PTEN* est

ainsi également un gène phare en oncogénétique digestive, au sein du sous-ensemble des polyposes hamartomateuses.

D'un point de vue moléculaire, le potentiel oncogénique des mutations inactivatrices de *PTEN* est expliqué par son **implication dans la voie de cancérogenèse PI3K-AKT-mTOR**, qu'il régule négativement. [69] Bien que l'implication de *PTEN* dans cette voie ait particulièrement été décrite au niveau cérébral [70], d'autres équipes ont permis de préciser son rôle dans le développement d'autres types de cancers, en particuliers hormonaux dépendants. [71 ;72]. Globalement, la voie PI3K-AKT-mTOR est impliquée dans la prolifération cellulaire, grâce à la **régulation transcriptionnelle** de certains gènes comme celui l'IGF-1, via le complexe mTORC1 [Figure 8], comme le font d'ailleurs certaines hormones sexuelles via leur récepteur nucléaire (ER). Cette voie, permet en outre, de diminuer le taux de protéine p53 en stimulant MDM2, **favorisant ainsi la survie cellulaire**. De manière intéressante, cette voie est interconnectée à une autre voie importante dans les processus de cancérogenèse, la voie RAS-MAPK-ERK stimulant le complexe mTORC1.

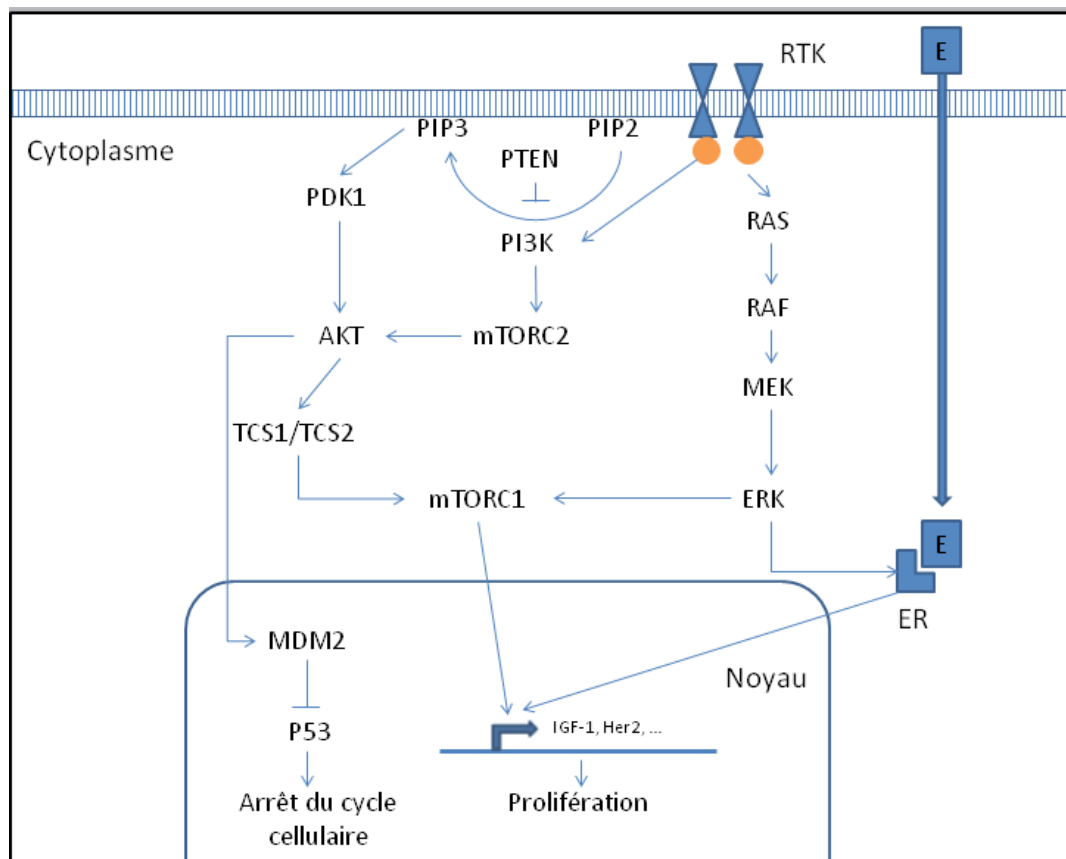


Figure 8 : Représentation simplifiée de la voie PI3K-AKT-mTOR (D'après [69-72]) Les interactions stimulatrices des différentes protéines sont représentées par les flèches continues. E = *Estrogen*, ER = *Estrogen Receptor*

Concernant le cancer du sein, Hobert et Eng [73] ont estimé un risque de développer une tumeur maligne du sein situé entre 25 et 50% chez les patientes atteinte d'un SC. En 2012, ce risque a été précisé par Tan et son équipe [74], avec un ratio d'incidence standardisé sur l'âge

(ou age-adjusted standard incidence ratio, SIR) à **25,4** (IC95% : 19,8-32,0) et un risque de **82,5%** sur la vie entière. Les autres risques de développer une tumeur du spectre ont également été rapportés, correspondant à 35,2% pour le cancer de la thyroïde, 28,2% pour le cancer de l'endomètre, 9,0 % pour le cancer colorectal, 33,6% pour le cancer du rein, et enfin 6,0% pour le mélanome. Deux éléments enfin sont importants à préciser.

- Le premier, c'est que de par l'implication de la voie PI3K-AKT-mTOR dans le développement du syndrome de Cowden, les altérations d'autres gènes de cette voie peuvent être responsables d'un phénotype similaire. C'est notamment le cas des gènes *PIK3CA* et *AKT1*. [75]
- Le second, est que bien qu'étant de taille modeste, puisque s'étendant sur environ 120kb, et étant constitué de 9 exons, les altérations de *PTEN* sont responsables de **plusieurs phénotypes individualisables**. Les syndromes alléliques au syndrome de Cowden correspondent au Bannayan-Riley-Ruvalbaca syndrome (BRRS), au PTEN-Related-Proteus-syndrome (PRTS), et au Proteus-Like syndrome (PLS), tant est si bien qu'il est préférable de retenir le terme de « PTEN Hamartoma Tumor Syndrome » (PHTS), en cas de mutation de ce gène.

I.2.4.3. Syndrome de prédisposition lié au gène *CDH1* (NM_0044360.3)

Le gène *CDH1*, d'une taille d'environ 100 kb et situé en 16q22.1, code pour la **cadhérine E**. [76]. Cette protéine transmembranaire est une molécule d'adhésion cellulaire, essentielle pour la morphogenèse, la polarisation et la migration cellulaire, ainsi que pour le maintien de l'architecture tissulaire [77], tant durant la période embryonnaire que durant la vie adulte. Elle est impliquée dans la formation des **jonctions intermédiaires**, ou *zonula adherens*. Leur domaine extracellulaire permet la liaison homo-dimérique avec les molécules de cadhérine d'une cellule voisine, tandis que leur domaine intracellulaire est intimement lié au cytosquelette, en particulier aux filaments d'actine, par l'intermédiaire de protéines appelées caténines [78].

Au niveau tumoral, la perte de fonction du gène de la cadhérine E est responsable de la **transformation maligne** par la perte de la capacité d'adhésion des cellules, et en favorisant leurs capacités de migration, donc la diffusion métastatique. Les mutations de ce gène sont observées au sein de certaines tumeurs d'origine épithéliale (carcinomes), en particulier du **sein et de l'estomac**, mais ont également été décrites dans les cancers de la prostate, de l'endomètre, et de l'ovaire. [79]

Les mutations constitutionnelles du gène *CDH1* sont quant à elles responsables d'un syndrome de prédisposition au cancer particulièrement reconnaissable, en raison du **phénotype histologique** des tumeurs qui y sont associées. Elles favorisent ainsi :

- en premier lieu le développement de l'**adénocarcinome de l'estomac à cellules indépendantes** (ou *Hereditary Gastric Diffuse Cancer*, HGDC) caractérisé en anatomopathologie par l'absence d'une formation tumorale solide au profit de cellules diffuses de morphologie particulière, dites « en bague à chaton », infiltrant la paroi gastrique [80].

- En second lieu, le développement chez les femmes atteintes de l'**adénocarcinome lobulaire du sein** (ou Carcinome Lobulaire Infiltrant, CLI). [81] Cette tumeur est classiquement caractérisée en anatomopathologie par la présence de cellules très infiltrantes, soit isolées, soit disposées en cordons de cellules uniques en file indienne, envahissant le tissu stromal [82].

En France, l'analyse du gène *CDH1* est donc réalisée dans deux situations. Dans le cadre de la présentation gastrique du syndrome, il est séquencé lorsque les critères diagnostiques de l'*International Gastric Cancer Linkage Consortium* sont satisfaits. Ceux-ci correspondent soit à 1/ l'**association de deux cas de cancers gastriques diffus**, dont au moins un diagnostiqué avant 50 ans ; ou 2/ l'**association familiale de trois cas de cancers gastriques diffus**, quel que soit l'âge du diagnostic. [83]

Dans la présentation mammaire de la maladie, *CDH1* est séquencé en panel de gènes, aux côtés des gènes *BRCA1*, *BRCA2*, *PALB2* et des autres gènes dont l'étude est recommandée en cas de suspicion de syndrome HBOC.

Enfin, en 2015, une étude portée par Hansford et publiée dans le *Journal of the American Medical Association (JAMA) Oncology* [84] a permis de préciser les risques de développer un adénocarcinome de l'estomac ou du sein chez les individus porteurs d'une mutation du gène *CDH1*. Cette étude, basée sur 183 cas-index satisfaisant aux critères cliniques du syndrome HGDC (cf. *supra*) fait actuellement référence. Ainsi, les risques cumulés à 80 ans de développer un cancer gastrique diffus sont de 70% pour les hommes (IC95% : 59-80) et de 56% pour les femmes (IC95% : 44-69). Ces dernières ont par ailleurs un risque cumulé à 80 ans de **42%** de développer un CLI (IC95% : 23-68).

I.2.4.4. Ataxie télangiectasie et Ataxie télangiectasie-Like

Le gène *ATM* (NM_000051.3) est un grand gène d'environ 150 kb, situé sur le chromosome 11 au locus 11q22.3, et s'étendant de la position génomique g.108.222.483 à la position g.108.369.101. [85] Il comporte 63 exons, dont 62 codants pour une protéine de 3056 acides aminés, comportant trois grands domaines fonctionnels, dont un domaine à **activité kinase**, et un domaine impliqué dans le maintien de l'**intégrité des télomères**.

Les altérations moléculaires bi-alléliques de ce gène sont responsables d'un syndrome classique en neurologie, de transmission autosomique récessive, appelé **Ataxie Télangiectasie** (OMIM #208900). Il est caractérisé entre autre, par la présence d'une ataxie cérébelleuse associée à des télangiectasies conjonctivales, et à une apraxie oculomotrice. Les patients atteints ont également un sur-risque de développer des cancers, estimé à environ **25%** durant la vie-entière. Ce sur-risque concerne différentes localisations, comme le sein, le foie, l'estomac, et l'œsophage, mais de manière plus minoritaire comparée aux tumeurs les plus fréquentes dans ce syndrome, qui sont essentiellement de nature hématologique [86]. Ce risque concerne essentiellement les leucémies (lymphoïdes aiguës) et les lymphomes

(habituellement de type B), observés de manière préférentielle chez les jeunes patients, en règle avant 20 ans.

Les altérations **mono-alléliques du gène ATM** ne sont responsables d'aucune manifestation neurologique, même mineure, chez les patients qui ont donc un statut hétérozygote. En revanche, l'étude de plusieurs familles comprenant des sujets atteints d'ataxie télangiectasie a permis de mettre en évidence un **sur-risque tumoral**, non seulement chez les sujets porteurs de mutations bi-allélique (atteints du syndrome), mais également chez leurs apparentés **porteurs de mutations mono-alléliques** (hétérozygotes sains). En ce sens, l'étude de Thompson et son équipe, datant de 2005 est particulièrement informative. [87] Basée sur l'observation de 1160 apparentés, dans 132 familles d'individus présentant une ataxie télangiectasie, ils ont pu montrer que le risque chez les hétérozygotes concerne essentiellement des localisations **digestives** (côlon, et vésicule biliaire), mais également le **sein**. D'autres atteintes ont été observées, concernant notamment l'estomac (avec 10 cas observés, pour 4,74 attendus) et le **pancréas** (4 observés pour 2,63 attendus), mais ils n'ont pas pu montrer de différence statistiquement significative (respectivement RR = 3,39 ; IC95% : 0,86-13,4 ; et RR = 2,41 ; IC95% : 0,34 à 17,1).

Ces observations ont par ailleurs été étayées en 2016 par une méta-analyse menée par Van Os et son équipe [88] qui a notamment permis d'affiner et de préciser ces sur-risques, notamment pour le sein avec **un risque relatif de 3,0**, pour un intervalle de confiance à 95% (IC95%) compris entre 2,1 et 4,5. Ainsi, les altérations mono-alléliques du gène ATM sont aujourd'hui considérée comme des **facteurs de risque modéré du cancer du sein**, justifiant ainsi son étude au sein des panels réalisés pour diagnostiquer un cancer du sein héréditaire. De ce fait, les altérations du gène ATM sont fréquemment retrouvées en tissus tumoral, comme le montrent les données issues de la base de donnée COSMIC (*Catalogue Of Somatic Mutation In Cancer* ; <http://cancer.sanger.ac.uk>) la plus utilisée dans le domaine de la génétique tumorale. Les tumeurs concernées sont les tumeurs hématologiques bien sûr, mais également celles du sein, de la prostate, du pancréas ; mais sont également retrouvées dans les cancers colorectaux, et pulmonaires.

D'un point de vue moléculaire, la protéine ATM intervient dans plusieurs grandes fonctions cellulaires, telles que la **régulation du cycle cellulaire**, la **croissance cellulaire**, et la **réparation des lésions de l'ADN**. En 2016, Choi, Kipps et Kurzrock ont publié dans *Molecular Cancer Therapeutics* [89] une revue de la littérature très complète portant sur l'implication du gène dans ces différentes voies, et les conséquences de ses altérations dans le domaine de l'oncologie.

Brièvement, en cas de lésion à l'ADN, ATM est recruté au niveau de la cassure par le complexe NRN lors de la phase présynaptique de la recombinaison homologue. La réparation se fait via l'intervention d'autres gènes, comme *BRCA1*, *RAD51*, etc. En cas d'échec de réparation, *ATM* joue encore un rôle, en **favorisant l'apoptose** par l'interaction avec la protéine p53, **en stoppant le cycle cellulaire** par l'interaction avec la protéine CHEK2, ou encore en **inhibant la croissance et la survie cellulaire** en régulant négativement la voie PI3K-AKT-mTOR.

Ainsi, plusieurs gènes avec lesquels interagit le gène ATM peuvent également être séquencés dans le cadre des panels destinés au diagnostic du cancer du sein héréditaire, de manière superposable aux partenaires des gènes *BRCA1*, *BRCA2* et *PALB2*. C'est notamment le cas

du gène *MRE11A*, appartenant au complexe NRN (*cf supra*), dont les mutations bi-alléliques sont responsables d'un tableau clinique d'ataxie télangiectasie, appelé ataxie télangiectasie-like.

I.2.4.5. Syndrome de Peutz-Jeghers lié au gènes *STK11* (NM_000455.4)

Les altérations du gène *STK11* sont responsables du syndrome de Peutz-Jeghers (OMIM #175200). Il s'agit d'un gène situé sur le chromosome 19, au locus 19p13.3. Il s'étend au total sur environ 23kb, et comprend 10 exons dont 9 codants pour une protéine kinase de 443 acides aminés. [90].

Tout comme le gène *PTEN*, *STK11* est impliqué dans les **cancers digestifs** au sein du sous-ensemble des polyposes hamartomateuses, aux côtés d'autres gènes comme *SMAD4*, et *BMPR1A*. Ce n'est d'ailleurs pas la seule caractéristique qu'il partage avec *PTEN*, puisque les altérations de *STK11* rendent compte d'un ensemble phénotypique assez large, comprenant des **atteintes cutanées**, et un sur-risque de **cancer du sein** entre autre.

D'un point de vue moléculaire, le fait que *PTEN* et *STK11* possèdent un phénotype relativement superposable s'explique par leur implication respective dans la voie de signalisation cellulaire **PIK3-AKT-mTOR**. Ainsi, si *PTEN* est un inhibiteur de cette voie par son action répressive sur *AKT* (*cf supra*), *STK11* lui exerce son activité inhibitrice directement sur un effecteur plus terminal, qu'est le gène *TSC2*, comme l'ont rapporté Corradetti et Guan en 2006 dans *Oncogene*. [91]

Cliniquement un syndrome de Peutz-Jeghers est suspecté devant l'association de plusieurs signes, définis par une conférence de consensus européenne en 2010, et repris dans un article de la revue *Gut* [92]. Ces critères comprennent évidemment la présence de polypes hamartomateux (deux au moins), et la présence d'antécédents familiaux. S'y associe dans le la présence de signes cutanés caractéristiques, à type de **lentignes au niveau des muqueuses et de la peau** de couleur bleutée à brune, préférentiellement en zone péri-orificielle (bouche, yeux, narines, périnée).

Le spectre tumoral associé aux altérations hétérozygotes du gène *STK11* est aujourd'hui assez bien délimité, notamment grâce à l'*American College of Gastroenterology* [93]. Ainsi, les risques majeurs concernent en particulier le **cancer colorectal** (risque cumulé de 39%), le cancer de l'**estomac** (29%), celui de l'**intestin grêle** (13%), celui du **pancréas** (11 à 36%) et celui du **sein** (32 à 54%). Des risques plus modérés sont également reportés, concernant le cancer du col de l'utérus, de l'utérus en lui-même, du testicule (environ 10% pour chacun d'entre eux), et celui du poumon (7 à 17%). Enfin, une tumeur ovarienne en particulier semble être caractéristique du syndrome de Peutz-Jeghers, non seulement en raison de sa fréquence, mais également en raison de son âge moyen de survenue. Il s'agit de la **tumeur maligne des cordons sexuels et du stroma de l'ovaire** (SCST), touchant 21% des femmes porteuses d'une mutation *STK11* (contre 1,4% en population générale), avec un âge moyen de survenue estimé à 28 ans.

Pour finir, deux éléments semblent intéressants à préciser concernant les tumeurs liées à ce gène.

- Le premier concerne les mécanismes moléculaires du développement du cancer du côlon, qui ont été assez bien décrits chez les patients atteints de ce syndrome, en particulier par l'équipe de Miyaki en 2000. [94] En se basant sur l'étude en tissus tumoral de 27 polypes hamartomateux, ils ont montré, de manière assez attendue, que ces lésions étaient dues à une inactivation tumorales **bi-allélique** de *STK11*. Par ailleurs, en se focalisant sur les lésions hamartomateuses ayant une composante adénomateuse, ils ont montré qu'une évolution de ces lésions vers une histologie adénomateuse voire adénocarcinomateuse était possible, par l'acquisition de **mutations supplémentaires**. Ainsi, en plus de l'inactivation bi-allélique de *STK11*, ces tumeurs présentaient des mutations au niveau d'autres gènes, comme *TP53*, ou du gène de la bêta-caténine (*CTNNB1*, par ailleurs impliquée dans les tumeurs desmoïdes liées au gène *APC*).
- Le second point concerne la fréquence des mutations du gène *STK11* en tissus tumoral dans les cancers sporadiques, en particulier le **cancer du poumon** non à petites cellules (NSCLC). Dans cette localisation précise, les mutation *STK11* sont retrouvées dans près d'un tiers des cas, et constitue le troisième gène le plus fréquemment muté, après *TP53* et *KRAS*. [95] Cette importante représentation du gène en tissus tumoral est également retrouvée dans une seconde localisation, le cancer du **col de l'utérus**, où il est retrouvé altéré dans près de 20% des cas. [96]

I.2.4.6. Autres gènes étudiés dans le cadre du syndrome HBOC

La plupart des laboratoires de génétique moléculaire en France utilisent le même panel de gènes pour faire le diagnostic de syndrome HBOC, issu des **recommandations du Groupe Génétique et Cancer**. Ainsi, à côté des gènes dont l'implication est majeure dans le développement du cancer du sein, ceux impliqués préférentiellement dans le cancer de l'ovaire sont également étudiés. Le sur-risque de cancer du sein pour ces derniers est moins connu ou moins documenté pour la plupart, mais nettement moins important en tout cas.

Enfin, d'autres gènes sont fréquemment étudiés au sein des laboratoires, soit en fonction de leur domaine d'expertise, soit à visée de recherche.

I.2.4.6.1. Gènes où l'atteinte ovarienne est au premier plan

Dans cette partie, deux grands cadres peuvent être distingués. Le premier, concerne le **syndrome de Lynch**, bien connu en oncogénétique digestive, et les mutations affectant le gène *RAD51* et ses paralogues (*RAD51B*, *RAD51C* et *RAD51D*).

I.2.4.6.1.1. Gènes impliqués dans le syndrome de Lynch

Le syndrome de Lynch, ou HNPCC pour *Hereditary Non Polyposis Colon Cancer*, est un syndrome de **prédisposition aux cancers digestifs** au sens large, mais d'autres organes extra-digestifs, peuvent également être touchés. Il s'agit d'un des syndromes de prédisposition aux cancers les plus fréquents. L'étude de De la Chapelle, publiée en 2005 dans *Familial Cancer* [97] a permis de préciser son incidence, en se basant sur l'étude de deux types de tumeurs évocatrices du syndrome non seulement en raison de leur **localisation** (côlon et endomètre), mais avant tout en raison de leurs **caractéristiques moléculaires**, car présentant des mutations dans les gènes responsables du syndrome HNPCC. Ainsi, il a estimé l'incidence du syndrome de Lynch entre **1/2000 à 1/660** ; taux confirmé par d'autres auteurs [98].

Dans le syndrome de Lynch, deux spectres tumoraux sont définis, en fonction des risques associés [99]. Le **spectre étroit**, comprend les localisations où le risque relatif (RR) est le plus élevé. Il s'agit du côlon et du rectum, de l'intestin grêle, de l'endomètre, et du cancer urothélial des voies urinaires supérieures. Le **spectre large**, concerne les localisations où le RR est plus faible, et comprend l'estomac, le glioblastome, les adénomes sébacés, le cholangiocarcinome, et l'ovaire. C'est en raison de cette dernière localisation qu'il est présent dans le panel HBOC recommandé par le Groupe Génétique et Cancer. [Tableau 7]. Le risque relatif pour chaque localisation peut être apprécié par le risque cumulé à 70 ans de présenter une de ces tumeurs. [Tableau 7]

Tableau 7 : Risques tumoraux associés au syndrome de Lynch (D'après [99])

Spectre	Localisation	Risque cumulé à 70 ans
Étroit	Colon et rectum	10-74%
	Intestin grêle	4,0-22%
	Endomètre	14-54%
	Voies urinaires supérieures	0,2-25%
Étendu	Ovaire	4,0-22%
	Estomac	0,2-22%
	SNC (glioblastome)	1,0-22%
	Voies biliaires (cholangiocarcinome)	0,2-22%
	Cutanée (adénomes sébacées)	1-9%

Ici doit être mentionné que la grande amplitude du risque cumulé observé à 70% est lié à **l'hétérogénéité du syndrome de Lynch**, certaines tumeurs étant plus fréquemment observées en fonction du gène atteint, ou du sexe du patient. C'est par exemple le cas du

cholangiocarcinome, où le risque est estimé entre 0,2 et 4% en cas d'atteinte des gènes *MLH1* ou *MSH2*, alors qu'en cas d'atteinte de *MSH6* il est estimé à 6% chez les hommes et 22% chez les femmes, et en cas de mutation de *PMS2*, il est estimé à 6%. Enfin, d'autres risques ont été associés au syndrome de Lynch, mais font l'objet de débats. C'est notamment le cas du cancer du pancréas, du cancer de la prostate ou du cancer du sein.

Concernant ce dernier, une revue de la littérature menée par Win et publiée en 2012 dans *Breast Cancer Research* [100] et reportant 21 études a montré que 13 études ne rapportaient pas de sur-risque, contre 8 ayant des conclusions opposées. Fait intéressant, parmi les 21 études revues, une était de type prospectif [101], toujours à l'initiative de Win et publiée un an plus tôt dans le *Journal of Clinical Oncology*, et a permis de mettre en évidence un **sur-risque modéré**, via un SIR de 3,95 (IC95% 1,59-8,13). Des études plus récentes semblent renforcer l'idée que le risque de développer un cancer du sein varie en fonction du gène impliqué, comme pour les autres localisations tumorales. Ainsi, il a récemment été mis en évidence à partir d'une étude prospective allemande portant sur 530 femmes atteintes d'un syndrome de Lynch, une **association entre les mutations *MSH6* et le cancer du sein, avec un SIR à 4,4** (IC95% 1,2-11,2) [102]. Une autre étude, rétrospective cette fois, a elle montré que les mutations du gène *PMS2* était de manière statistiquement significative ($p = 0,0043$) plus fréquemment rencontrées chez les femmes **atteinte d'un syndrome de Lynch et ayant présenté un cancer du sein**, comparées à celles porteuses d'une mutation *MLH1*, *MSH2* ou *MSH6*. [103] Il n'en demeure pas moins que la question de l'association entre syndrome HNPCC et cancer du sein restera sans réponse certaine encore pendant longtemps, compte-tenu de l'importante incidence du cancer du sein sporadique, compliquant ce type d'analyses.

D'un point de vue moléculaire, le syndrome de Lynch est extrêmement intéressant, pour plusieurs raisons.

- La première, est qu'il est dû aux altérations de plusieurs gènes, définissant la notion de **variabilité « inter-locus »**. Plusieurs gènes ont été incriminés, tous appartenant au système de réparation des mésappariements de l'ADN (ou *Mismatch Repair*, *MMR*) : il s'agit des gènes *MLH1*, *MSH2*, *MSH6*, *PMS2* et *EPCAM*.
- La deuxième raison est qu'en raison de la fréquence du cancer colorectal en population générale, il est nécessaire de faire appel à plusieurs niveaux de compétence en matière de techniques de biologie afin de l'identifier, afin de le distinguer d'un cancer d'origine sporadique.

Ainsi, un algorithme diagnostique a été proposé en 2010 par Bouzourene et son équipe [104], repris par l'Association Américaine de Gastroentérologie en 2015 [105]. Il est donc recommandé de procéder en premier lieu soit à une étude **immunohistochimique** des 4 protéines (*MLH1*, *MSH2*, *MSH6* et *PMS2*) sur matériel tumoral, la perte d'expression d'une protéine pouvant être le reflet d'un variant constitutionnel perte de fonction du gène correspondant (associé à une perte d'hétérozygotie en tissus tumoral) ; soit à une **recherche d'instabilité de microsatellites** normalement monomorphes en population générale, leur variation de longueur en tissus tumoral pouvant être le reflet d'un défaut du système *MMR*. En cas de **perte d'expression de la protéine *MLH1***, il est recommandé de procéder à la fois à la recherche ciblée de la mutation **p.(Val600Glu) du gène *BRAF***, la mutation

« V600E » ; et à la recherche d'une **hyperméthylation du promoteur du gène de MLH1**, toutes deux étant classiquement absentes chez les patients atteints d'un syndrome de Lynch. Une fois cette possibilité éliminée, ou en cas de perte d'expression d'une protéine autre que MLH1 associée à une instabilité des microsatellites, l'analyse constitutionnelle des gènes MMR est réalisée.

- La troisième raison est représentée par le cas particulier du syndrome de Lynch lié aux anomalies du gène *EPCAM*. En 2009, Ligtenberg et son équipe [106] ont décrit un nouveau mécanisme moléculaire responsable du syndrome de Lynch, impliquant le gène *EPCAM*, n'appartenant pas au système MMR. En se basant sur l'étude de tumeurs MSH2-déficientes, ils ont montré que **les délétions des derniers exons du gène EPCAM**, gène situé en 5', conduisaient à la **méthylation du promoteur du gène MSH2 situé en Cis**, et donc à son inactivation. [Figure 9]

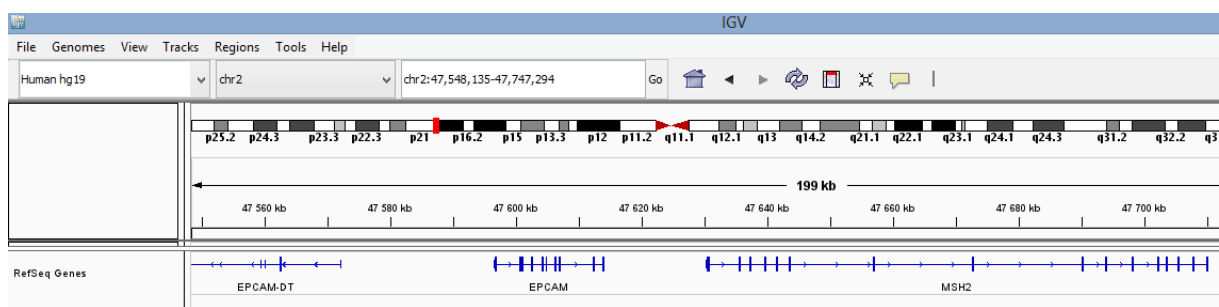


Figure 9 : Positions chromosomique des gènes *EPCAM* et *MSH2*
(Visualisation à l'aide du logiciel IGV (Integrative Genomics Viewer))

- Le dernier élément à souligner concernant le syndrome de Lynch, et le nombre de syndromes qui lui sont allélique, ou qui sont considérés comme étant des variants de ce syndrome. Peuvent être mentionnés ici le syndrome de **Muir-Torre** (OMIM #158320), caractérisé par la prédominance des atteintes cutanées [107], et le syndrome de **Turcot**, caractérisé lui par une prédominance des tumeurs cérébrales [108]. Enfin, un phénotype pédiatrique particulier peut être caractérisé en cas d'inactivation bi-allélique d'un des gènes *MMR*, appelé d'après son acronyme anglais *CMMRD* pour *Constitutional Mismatch Repair Deficiency*. Des études rétrospectives portant sur ce phénotype rare [109-110] ont permis de préciser les atteintes les plus fréquentes chez les enfants concernés. Ainsi, il a été relevé une prédominance des atteintes cérébrales, gastro-intestinales précoces et hématologiques. Par ailleurs, la majorité des patients présentent des tâches-café au lait, signe dermatologique important pour orienter vers ce syndrome, et faire le distinguo avec un syndrome de Li-Frauméni qu'il peut mimer par exemple, par la précocité et la nature des atteintes observées.

I.2.4.6.1.2. Gènes des recombinaisons : paralogues de RAD51

Il est important ici de rappeler la notion importante, et assez vite oubliée que le génome en général, et le génome humain en particulier ne sont pas figés, et que des changements de structure ou de nombre ont permis à l'espèce humaine d'émerger à partir d'un ancêtre commun à d'autres espèces de mammifères. Ce rappel étant fait, on peut donc préciser ce que sont deux gènes dits « **paralogues** ». En effet, au cours de l'évolution, un gène peut se dupliquer, et avec le temps, chacune des deux copies accumuler des variations, tout en gardant un certain degré d'homologie. Ces deux gènes différents, issus d'un même « ancêtre génique commun » sont alors appelés paralogues. Cela est notamment le cas du gène *RAD51*, impliqué dans la réparation des cassures double brin, qui comporte plusieurs paralogues, dont les gènes *RAD51B*, *RAD51C* et *RAD51D*.

Le gène *RAD51* humain a en premier lieu été identifié en 1993 (par analogie avec le gène d'espèces bactériennes telles que *E.coli* ou *Saccharomyces cerevisiae*) [111]. Par la suite, ont successivement été caractérisés les gènes *RAD51B* en 1997 [112]), puis *RAD51C* [113] et enfin *RAD51D* [114], tous deux en 1998. Par la suite, les connaissances se sont rapidement accumulées concernant les fonctions biologiques de ces différents gènes, et en 2006 est parue une étude dans le journal européen de biologie moléculaire (*European Molecular Biology Organization, EMBO*) reprenant les rôles de ces différents gènes, dans le processus de recombinaison homologue. [115]

On retiendra simplement que si *RAD51* est un des **effecteurs centraux de la recombinaison homologue**, la présence de *RAD51C* est nécessaire pour sa stabilisation, ainsi que pour la **régulation du cycle cellulaire**, comme *RAD51B*, puisque leur inhibition conduit à son arrêt. Cependant, *RAD51C* semble être une protéine centrale, puisque présente dans les deux complexes protéiques permettant le recrutement de *RAD51* aux sites de lésion de l'ADN, le premier formé de l'hétéro-dimère *RAD51C-XRCC3*, le second formé de l'hétéro-tétramère *RAD51B-RAD51C-RAD51D-XRCC2*.

L'implication des paralogues de *RAD51* dans les cancers du sein et de l'ovaire est de découverte assez récente. En 2017, une équipe française a publié dans le Journal Européen de Génétique Humaine, une étude permettant de préciser la fréquence des mutations de ces gènes dans les cancers du sein et de l'ovaire. [116] Basée sur 2649 individus présentant un cancer du sein ou de l'ovaire avec des antécédents familiaux de ces mêmes cancers, les auteurs ont pu montrer que ces variants étaient retrouvés à une fréquence de 0,73% chez les patients présentant une histoire **personnelle et familiale** de cancer du sein, et de 2,63% chez les patientes présentant une **histoire personnelle et familiale de cancer du sein et de l'ovaire associée**, quelle que soit leur atteinte personnelle.

Enfin, récemment, une équipe a précisé les risques tumoraux associés aux altérations des gènes *RAD51C* et *RAD51D*, à partir de l'étude de 125 familles où ségrège le premier gène, et 60 le second. [117] Ils ont ainsi défini un risque relatif de 7,55 (IC95% : 5,6-10,19) pour le cancer tubo-ovarien lié à *RAD51C*, et de 1,99 (IC95% : 1,39-2,85) pour le cancer du sein. Concernant *RAD51D*, ces risques ont été estimés respectivement à 7,60 (IC95% : 5,61-10,3) et à 1,83 (IC95% : 1,24-2,72), confirmant ainsi le risque prédominant de cancer tubo-ovarien.

Concernant le gène RAD51B, son rôle en pathologie humaine n'est **pas encore clairement établi**. L'étude de la base de données ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar>) ne reporte à l'heure actuelle qu'un seul **variant ponctuel** le concernant, un variant synonyme classé probablement pathogène (Classe 4 selon les critères de l'ACMG). Ce variant (rs764896402), nomenclaturé c.84G>A soit p.(Gln28=) sur le transcrit de référence NM_133509.4, est considéré comme probablement pathogène en raison d'un probable effet sur l'épissage (car étant le dernier nucléotide de l'exon 2), et de sa détection chez plusieurs individus présentant un cancer du sein ou de l'ovaire. Des données supplémentaires sont cependant nécessaires pour confirmer ou infirmer son rôle dans le cadre du syndrome HBOC.

I.2.4.6.2. Gènes d'utilité clinique moins certaine

Dans cette partie, les gènes évoqués ne seront que très brièvement décrits, le sur-risque tumoral associé à leurs altérations étant mal connu.

- *BARD1* (NM_000465.2)

Le gène *BARD1* est un petit gène situé sur le chromosome 2. Il code pour une protéine de 777 acides aminés, ayant une relative similarité de structure avec la protéine BRCA1. [118] Plusieurs domaines protéiques sont décrits, également retrouvés au sein de la protéine BRCA1, comme le domaine **RING** en N-terminal, et les domaines **BRCT**, en C-terminal. Comme évoqué dans la partie plus haut, l'interaction des protéines BRCA1 et BARD1, *via* leurs domaines RING, est fondamentale pour la stabilisation du filament de RAD51, durant la phase synaptique de la réparation des cassures double brin par recombinaison homologue.

- *BRIP1* (NM_032043.2)

De découverte plus récente que *BARD1*, les fonctions du gène *BRIP1* sont encore mal caractérisées. [119] Tout comme la protéine BARD1, celle codée par ce gène interagit avec BRCA1, *via* ses domaines **BRCT** cette fois. L'association de ces deux protéines permet la réparation des cassures doubles brin de l'ADN, par mécanisme de recombinaison homologue.

- *NBN* (NM_002485)

Le gène *NBN*, situé en 8q21.13 – q21.3 code pour une protéine de 754 AA, appelée nibrine. Les altérations moléculaires associées à ce gène ont été reportées comme étant responsables du syndrome de **Nijmegen**, appartenant au groupe de syndromes dits « d'instabilité chromosomique », aux côtés des syndromes d'Ataxie-Télangiectasie et du syndrome de Bloom. [120]

La nibrine est impliquée dans les phases **précoces** de la recombinaison homologue, au sein du complexe MRN [41], associé aux protéines RAD50 et MRE11, comme dit plus haut ; et dont le rôle majeur est la détection des cassures double brin, et le recrutement des protéines nécessaires à sa réparation.

I.3. Principales techniques utilisées en oncogénétique moléculaire

Les techniques de biologie moléculaire utilisées dans le cadre de l'oncogénétique constitutionnelle peuvent être regroupées en **deux grandes catégories**, en fonction du type d'altérations moléculaires qu'elles permettent de détecter.

- Les techniques qualitatives, permettent de mettre en évidence des variants **ponctuels** de séquence, à type de substitutions, petites délétions ou petites insertions. Ce type d'altérations a des conséquences variées sur la séquence protéique, et elles sont qualifiées de mutations synonymes, faux-sens, non-sens, décalage du cadre de lecture (ou *frameshift*), et peuvent également avoir un impact délétère sur l'épissage. [Tableau 8] Elles peuvent être recherchées de manière ciblée sur des régions d'intérêt (technique de Sanger), ou plus globale (Séquençage de nouvelle génération).

Tableau 8 : Types de mutations et conséquences protéiques, exemple de *BRCA1*

(D'après Clinvar : <https://www.ncbi.nlm.nih.gov/clinvar>)

Type de mutation	Conséquences protéiques	Exemple
Synonyme	Peu importantes en règle. Rares altérations de l'épissage si touchent les régions exoniques des jonctions intron-exon	<i>BRCA1</i> NM_007294.4 c.4812A>G ; p.(Gln1604=)
Faux-sens	Changement d'un acide aminé. Difficiles à qualifier : parfois gain ou perte de fonction, voire modification de structure tertiaire	<i>BRCA1</i> NM_007294.4 c.5117G>C ; p.(Gly1706Ala)
Non-sens	Création d'un codon stop (arrêt de la traduction) prématuré. Protéine tronquée, perte de fonction	<i>BRCA1</i> NM_007294.4 c.5449G>T ; p.(Glu1817*)
<i>Frameshift</i>	Décalage du cadre de lecture avec substitution de plusieurs acides aminés, jusqu'à la création d'un codon stop prématuré. Protéine tronquée, perte de fonction	<i>BRCA1</i> NM_007294.4 c.5551delG ; p.(Asp1851Thrfs*4)
Epissage	Destruction ou création de sites d'épissages. Délétion d'exon, inclusion de séquences introniques dans le cadre de lecture. Protéine tronquée, intégration de séquences non-codantes.	<i>BRCA1</i> NM_007294.4 c.5075-2A>G ; p.(Asp1692_Trp1718delinsGly) Δex18insGly

- Les techniques quantitatives, permettent de révéler la présence de certains types de **réarrangements** de plus ou moins grande taille selon la technique employée. Peuvent ainsi être mis en évidence :

1/ des altérations limités à **un ou plusieurs exons** tout au plus, en règle ne dépassant pas l'échelle du gène. Il s'agit le plus souvent de délétions ou de duplications de régions géniques, détectées par des techniques ciblées (PCR quantitative en temps réel, MLPA, etc.);

2/ des altérations de plus grande taille, **dépassant l'échelle du gène**, et pouvant aller jusqu'à l'échelle chromosomique, détectées par étude pan-génomique grâce à l'analyse chromosomique par puce à ADN (ACPA). Cette technique offre par ailleurs l'avantage, grâce à sa grande résolution et à son *design* modulable, de détecter les points de cassures lorsqu'une délétion est détectée, en enrichissant en oligonucléotides les régions d'intérêt.

D'autres types de mutation peuvent **exceptionnellement** être détectés, comme les translocations, les insertions de séquences Alu, ou encore les variants ponctuels présents à faible fréquence allélique (appelé mosaïcismes), mais les laboratoires ne peuvent garantir leur mise en évidence systématique lorsqu'ils sont présents.

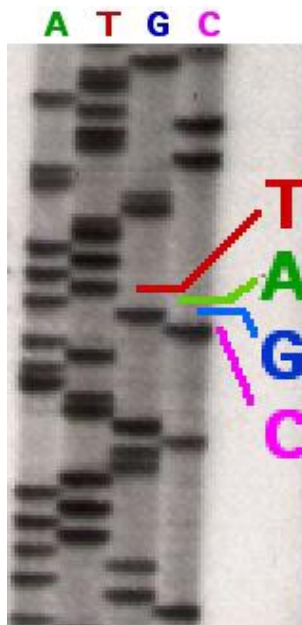
- Les **translocations** peuvent être détectées en NGS, lorsqu'un des points de cassure survient au sein d'une région exonique d'un gène étudié dans le panel. Il est alors schématiquement visualisé une délétion hétérozygote, associée à une insertion hétérozygote d'une même séquence, qu'il est alors possible de caractériser à l'aide d'outil d'alignement de séquence sur le génome.
- La présence d'une **insertion de séquence Alu**, peuvent être suspectées par analyse des profondeurs de couvertures de gènes en NGS. Il est alors classiquement observé une répartition non-gaussienne des *reads* (interruption brutale, ou aspect de double bosse). Une adaptation du *pipeline* bio-informatique avec des outils de détection des séquences Alu permettent alors de les caractériser.
- Les **mosaïcismes** sont quant à eux détectés en NGS, lorsqu'ils sont visibles à une fréquence allélique suffisante. Ils apparaissent alors comme des variants non-hétérozygotes et non-homozygotes. La confirmation de leur présence nécessite la réalisation d'analyse ciblée de type Sanger sur la région portant l'anomalie, idéalement sur un prélèvement d'ADN issu d'un autre type tissulaires (biopsie cutanée, etc.)

Enfin, certains mécanismes oncogéniques, plus fréquents dans le cadre de l'oncogénétique tumorale peuvent parfois être recherchés à l'état constitutionnel, comme les modifications épigénétiques de promoteurs de gènes. C'est notamment le cas de l'étude de la méthylation du promoteur du gène *MLH1*, décrite depuis peu à l'état constitutionnel de manière extrêmement rare. [121] Ce dernier type d'étude ne sera pas abordé dans cette partie.

I.3.1. Techniques qualitatives

I.3.1.1. Technique de Sanger

Initialement décrite par Sanger et Nicklen en 1977, la technique de séquençage par réaction de polymérisation en chaîne terminale est toujours utilisée dans la pratique courante actuelle, tant il s'agit d'une méthode **simple, fiable, et rapide** d'étude de séquences nucléotidiques. [122] Cette méthode repose sur l'utilisation d'une ADN polymérase, de désoxyribonucléotides utilisés en excès, auxquels on associe des **didésoxyribonucléotides** (sans groupement hydroxyle, ou –OH, en position 3') marqués et en concentration limitante, afin de stopper l'élongation du brin néoformé à partir d'une courte séquence d'ADN complémentaire à la région d'intérêt, appelée amorce. Jusqu'à nos jours, la technique a considérablement évolué, permettant de gagner en rapidité.



A l'origine, les didésoxyribonucléotides terminateurs d'élongation utilisés étaient marqués à l'aide d'un radio-isotope, le Phosphore 32 (^{32}P) et la réaction de séquençage réalisée en parallèle dans 4 tubes différents. Chaque tube comprenait, en plus des 4 bases nécessaires à la synthèse du nouveau brin, un seul type de base modifiée et marqué (didésoxyriboadénine, didésoxyriboguanine, didésoxyribothymine, ou didésoxyribocytosine). Au cours de l'élongation, dans chacun des tubes, l'élongation de la molécule néoformée se stoppe ainsi de manière aléatoire par l'incorporation d'une base marquée, générant ainsi des fragments de longueur différente. Enfin, les produits générés sont mis à migrer au sein d'un gel d'électrophorèse de polyacrylamide, et la révélation des radio-isotopes permet de reconstituer la séquence, en fonction de la taille des fragments. [Figure 10]

Figure 10 : Exemple de séquençage par méthode de Sanger (Par John Schmidt). Ici la séquence étudiée est donc : TACGAGATATATGGCGTTAATACGATATATTGGAACCTTCTATTGC

Par la suite, la technique a bénéficié de plusieurs améliorations, lui permettant de gagner :

- en **sécurité** en abandonnant l'utilisation de marqueurs radio-isotopiques au profit de marqueurs fluorescents en 1986 [123]
- en **rapidité** et en **sensibilité**, en se reposant sur l'utilisation de la réaction de polymérisation en chaîne (ou PCR) [124]

- en **coût**, aussi bien économique qu'humain, en automatisant la technique, grâce à l'abandon de l'électrophorèse classique en gel, au profit de l'électrophorèse capillaire en 1991 [125].

Les résultats obtenus sous forme d'électrophorégramme, sont ainsi plus simples à interpréter, et peuvent par ailleurs être automatiquement alignés sur la séquence de référence afin de visualiser rapidement et simplement les variants de séquence d'intérêt. [Figure 11]. Cette technique est particulièrement efficace en génétique constitutionnelle, puisque l'électrophorégramme d'un variant hétérozygote est représenté par un double pic de hauteurs similaires pour chacune des bases détectées à une position donnée (G et C par exemple, dans le cas d'une transversion à une position connue). Il est cependant important de noter deux limites :

- la première concernant les variants homozygotes, qui ne seront détectables que si la base de référence est connue ; les logiciels d'alignement ne détectant que les pics supplémentaires à une position donnée
- la seconde concernant les mosaïcismes, qui est également applicable en génétique tumorale ; puisque leur détection nécessite des fréquences alléliques relativement importantes, afin de pouvoir être distinguées du bruit de fond (en règle, supérieur à 10%).

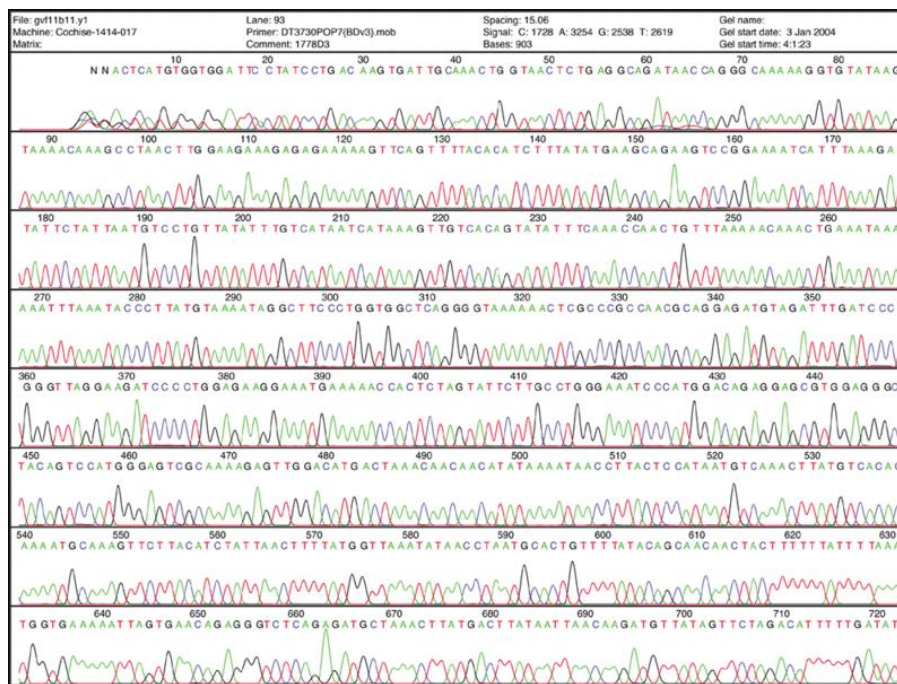


Figure 11 : Electrophorégramme d'un séquençage Sanger par électrophorèse capillaire (Par National Human Genome Research Institute)

Cette technique reste cependant qualifiée d'analyse à « **bas-débit** », car ne permettant l'étude que d'une région d'intérêt relativement courte, limité à environ 1Kb correspondant, en règle générale, à tout ou partie d'exon, plus ou moins étendue aux jonctions exons-introns sur quelques dizaines de bases.

I.3.1.2. Le séquençage de nouvelle génération

Si la technique de Sanger garde donc encore aujourd'hui toute sa pertinence dans **l'étude ciblée de gènes ou de régions spécifiques de gènes**, l'avancée majeure de ces 15 dernières années a été l'émergence des technologies de séquençage haut-débit, ou NGS (pour *Next Generation Sequencing*).

Là où la technique de Sanger ne permet que l'étude de courtes séquences nucléotidiques pour un seul patient dans un même temps opératoire, les techniques de séquençage massivement parallèles permettent elles, l'étude de **plusieurs gènes entiers, et ce pour plusieurs patients**. Grâce aux progrès technologiques réalisés, l'étude de l'ensemble des exons de tous les gènes (exome), voire l'étude du génome entier est à présent possible, dans un temps relativement court et pour un coût maîtrisé. Ceci quand, rappelons-le, le *Human Genome Project* (l'étude du génome entier par méthode de Sanger) a nécessité 13 ans (entre 1990 et 2003), et la collaboration de 20 laboratoires issus de 7 nations différentes, pour un coût d'environ 3 Milliards de dollars. [126] Plusieurs technologies sont actuellement disponibles, proposées par différentes sociétés, et de nombreux auteurs en ont fait la synthèse comparative. [127-128] Dans cette partie, seront simplement et brièvement décrites les principales étapes de cette technique. [Figure 12]

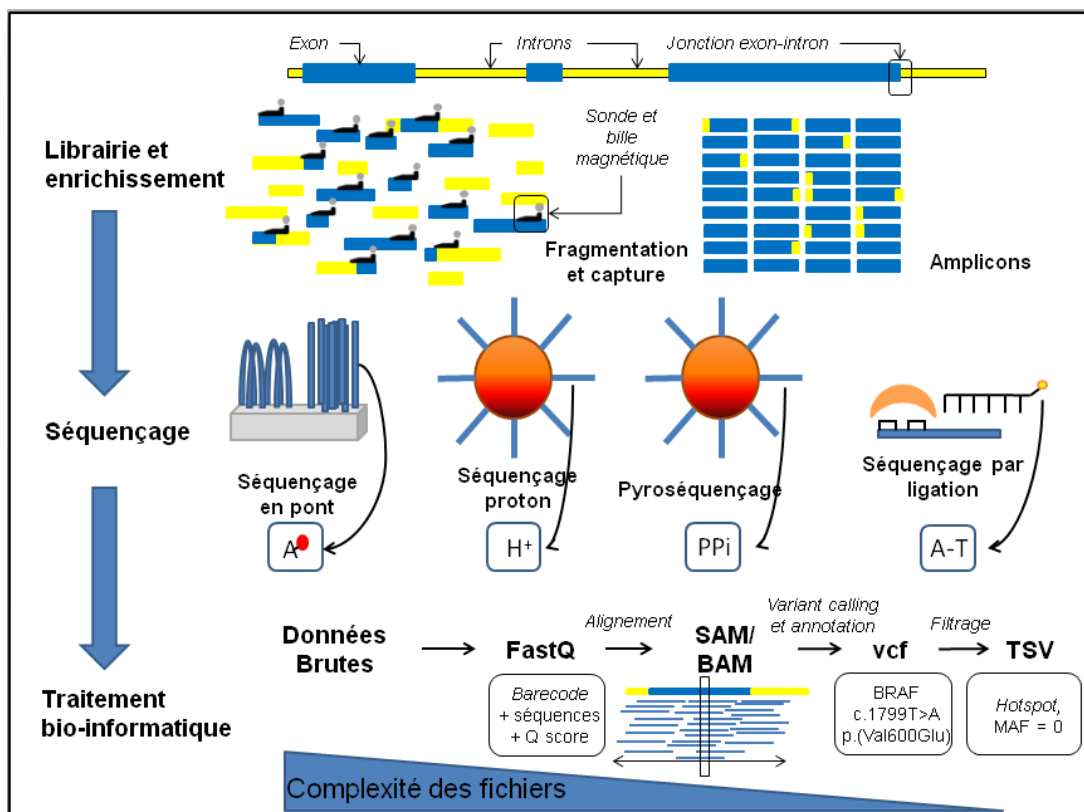


Figure 12 : Grandes étapes du NGS (D'après [127-128]). Pour chaque technique, les éléments détectés pour reconnaître la base insérée sont représentés (nucléotide marqué par fluorescence, proton (H⁺), pyrophosphate (PPi), doublet de bases).

La première étape, après l'extraction de l'ADN génomique à étudier, consiste à sélectionner et amplifier les régions d'intérêt, afin de générer de courtes séquences nucléotidiques qui pourront être facilement « lues » par le séquenceur. C'est ce que l'on appelle « **constituer et enrichir une librairie** ». Deux méthodes sont alors possibles :

- La **fragmentation** de l'ADN (soit par technique enzymatique à l'aide d'une enzyme de restriction, soit par technique mécanique par ultrasons ou « sonication ») suivie par une phase de « **capture** » de régions d'intérêt à l'aide de sondes d'ADN complémentaire portant une microbille magnétique (« stratégie d'enrichissement par capture ») ;
- La constitution d'une librairie directement enrichie, car dépourvue des séquences hors-région d'intérêt, par **PCR multiplex** ciblant spécifiquement la zone à séquencer (« stratégie d'enrichissement par **amplicons** »).

L'étape suivante consiste à préparer les librairies pour le séquençage. Lors de cette phase, des « **adaptateurs** » sont incorporés de part et d'autre des séquences d'intérêt générées. Ces courtes séquences nucléotidiques ont plusieurs fonctions :

- elles permettent la **fixation de l'amplicon au support de séquençage** (bille, ou lame de verre) ;
- elles contiennent une **amorce universelle de PCR** pour le séquençage en lui-même ; et
- elles permettent l'**identification des patients** à l'aide d'une petite séquence de nucléotides d'une longueur et d'un motif unique (code barre ou « **barcode** »). A l'issue de cette étape, l'ensemble des échantillons des patients différents peuvent être réunis (ou « poolés »), quand, durant l'étape précédente, les manipulations étaient réalisées parallèlement et séparément pour chaque patient.

Après plusieurs purifications, la troisième étape est l'étape d'**amplification et de séquençage** à proprement parler. Plusieurs technologies sont disponibles, spécifiques à chaque société. Quatre grandes méthodes sont disponibles.

- l'amplification **en pont** sur lame de verre (ou puce) et séquençage par synthèse, où la détection de la base insérée pour le séquençage se fait par mesure de fluorescence ;
- l'amplification par **émulsion à l'aide de microbille en puits** sur plaque, où la détection de la base insérée se fait par mesure de la variation du pH (détection des protons, ions H⁺) ;
- la méthode de **pyroséquençage**, utilisant également une amplification par émulsion, où la détection de la base insérée est effectuée par mesure d'un signal lumineux, la libération d'un groupement pyrophosphate (PPi) aboutissant à l'activation d'une luciférase

- la méthode par **ligation**, qui diffère des trois précédentes. Avec cette méthode, l'amplification se fait également par émulsion, mais le séquençage repose sur l'utilisation d'oligonucléotides de 8 bases, portant un fluorochrome. Les deux premières bases de l'oligonucléotide sont celles détectées pour le séquençage, après ligature et clivage des 6 autres, permettant la détection du signal fluorescent.

Enfin, la dernière étape correspond à **l'intégration et au traitement des données de séquençage**, à l'aide d'un algorithme de programmes bio-informatiques (« *pipeline* bio-informatique » ou « *workflow* bio-informatique »). L'objectif final est la création d'un rapport regroupant l'ensemble des variants d'intérêt, pour interprétation par le biologiste. Pour cela, plusieurs étapes sont nécessaires, de même que la création de différents formats de fichiers informatiques.

- les **données brutes** issues du séquenceur, à savoir les séquences lues appelées « *reads* » sont enregistrées dans un fichier « Fastq ». Il associe, à chaque *read* par ailleurs identifiée par son *barcode*, un **score de qualité** du séquençage pour chaque nucléotide de la séquence, évaluant la probabilité d'erreur de lecture. Ce score, appelé phred (symbolisé par la lettre Q) est logarithmique, selon la formule :

$$Q = - 10 \log_{10}(P)$$

avec P étant la probabilité d'erreur. Ainsi, un score Q à 30, correspond à une probabilité que le variant observé soit une erreur de séquençage de 1/1000, soit une précision de détection des mutations de 99,99%. [129]

- Ensuite, par l'étape d'**alignement** des *reads* sur le génome de référence, un nouveau format de fichier est généré, appelé « SAM » (pour *Sequence Alignment Map* »). Ce format de fichier est ensuite converti en version binaire, sous la forme d'un fichier « BAM » (pour *Binary Alignment Map* »). Les séquences lues sont donc précisément **positionnées sur le génome de référence**, et associées à leurs coordonnées génomiques. Deux informations supplémentaires deviennent alors accessibles. [Figure 13]

1/ La première est la **profondeur de séquençage**, correspondant, pour une position génomique précise, le nombre de fois qu'elle a été séquencée (correspondant au nombre de *reads* la comprennent) ; il s'agit donc d'une information « verticale ». Ainsi, une base donnée peut avoir une profondeur de séquençage estimée à 200X, signifiant qu'elle a été lue 200 fois, car 200 *reads* la comprennent.

2/ La seconde est la **couverture de séquençage**, correspondant au pourcentage des zones d'intérêt totales séquencées ; il s'agit donc d'une information concernant l'étalement des *reads* sur la zone que l'on veut étudier, en somme, une information « horizontale ». Ainsi, si la zone d'intérêt est l'ensemble des gènes impliqués dans le cancer du sein (panel ciblé sur une pathologie), la couverture peut correspondre à 99%,

signifiant que 99% des séquences codantes et jonctions introns-exons ont été correctement séquencées.

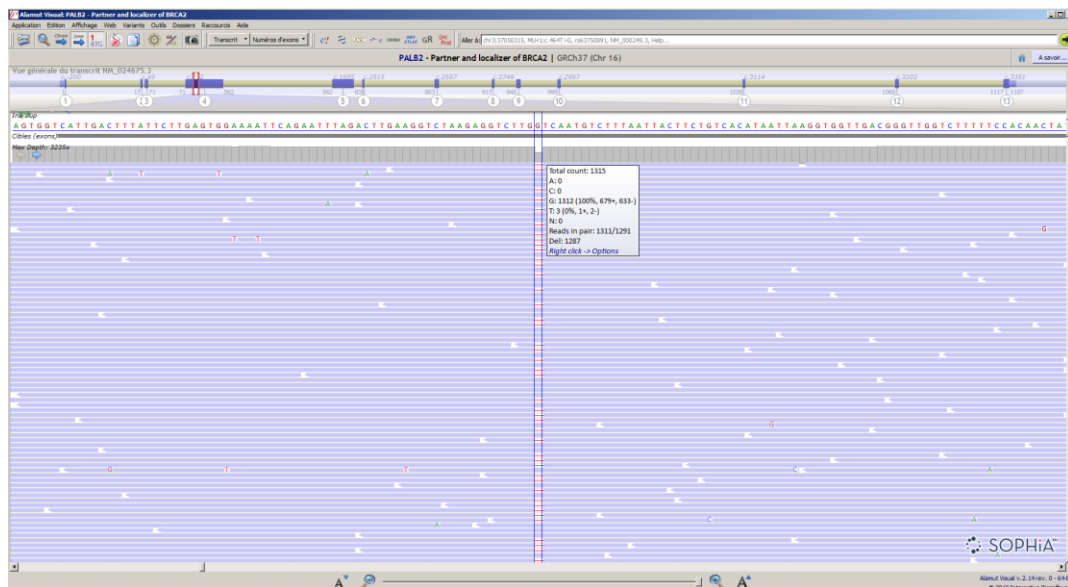


Figure 13 : Exemple de visualisation d'un BAM (Visualisation à l'aide du logiciel Alamut®, Sophia Genetics). Pour cette séquence au sein de l'exon 4 de *PALB2*, la couverture est de 100%, et la profondeur est de 1315X pour la base étudiée. Visualisation d'une délétion hétérozygote.

- Après l'étape d'alignement, survient celle de l'appel de variant (ou *variant calling*), aboutissant à la création d'un fichier « vcf » (pour « *Variant Calling File* »), comprenant l'**ensemble de variants détectés chez un même patient**, identifié via le *barcode* qui lui est propre.

Pour chaque variant détecté, un algorithme informatique associe différentes informations utiles pour l'interprétation de sa pathogénicité. C'est l'étape d'**annotation des variants**. Ces informations sont de différente nature, comme la fréquence en population générale, l'impact protéique prévisible, etc.

- Enfin, un dernier fichier est généré pour chaque patient, appelé « TSV » pour *Tab Separated Values*. C'est ce fichier qui est réellement interprété par le biologiste, après une étape informatique de **filtration des variants**, afin de ne retenir que ceux potentiellement pathogènes. Ainsi, les variants de trop forte fréquence en population générale, les variants avec de trop faibles Q scores, les variants situés hors *hotspots*, ou encore les possibles artéfacts sont filtrés. Cette dernière étape de filtration réalisée par le *pipeline* bio-informatique se fait selon des critères définis par le laboratoire. Le fichier TSV est visualisable sous la forme d'un tableur, reprenant chaque variant et chaque annotation qui lui est associée.

A noter qu'en cas de doute sur la nature artéfactuelle d'un variant, le biologiste peut accéder et visualiser les données de séquençage du patient (BAM et vcf).

I.3.2. Techniques quantitatives

Si l'approche NGS permet une estimation du nombre de copies à un locus donné par **analyse relative des profondeurs de séquençage**, la présence de réarrangements de grande taille à l'échelle de l'exon ou du gène sont tout de même fréquemment vérifiés par une seconde technique. Ces techniques sont spécifiquement quantitatives, et deux grandes techniques sont fréquemment utilisées.

I.3.2.1. Multiplex Ligation dependent probe amplification (MLPA)

Décrite en 2002 [130], la technique d'amplification multiplexe dépendante de la ligation (ou MLPA pour *Multiplex Ligation dependent Probe Amplification*) est une technique simple, rapide, et peu coûteuse pour **étudier simultanément une quarantaine de loci**, dans le même temps opératoire.

Comme la grande majorité des techniques utilisées en biologie moléculaire, la MLPA se base sur le principe de la PCR. La particularité de cette technique réside dans trois éléments fondamentaux. [Figure 14]

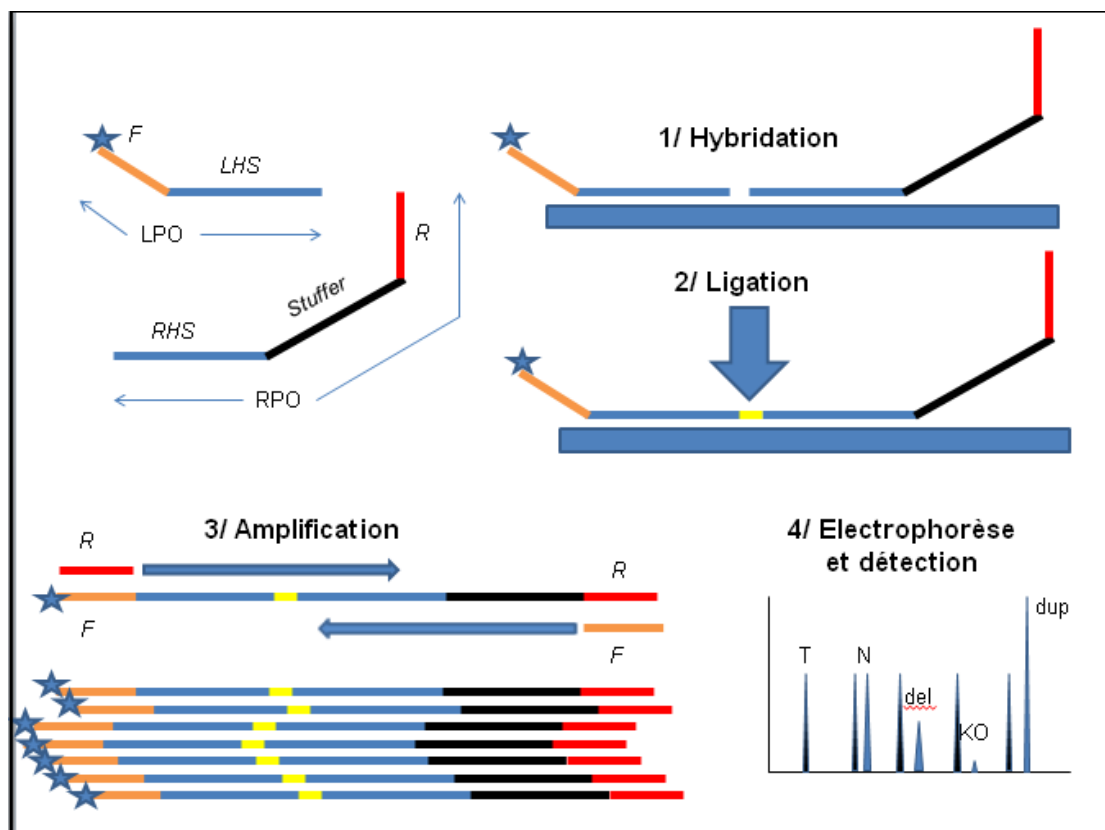


Figure 14 : Description des amorces et des étapes utilisées en MLPA (D'après [130])
F = Amorce Forward ou sens, LHS = Left Hybridization Sequence, LPO = Left Probe Oligo, RHS = Right Hybridization Sequence, R = Amorce Reverse ou anti-sens, RPO = Right Probe Oligo, T = Témoin, N = Normal, Del = délétion hétérozygote, KO = délétion homozygote, Dup = duplication

- le premier est la **construction particulière des amorces utilisées**. En effet, dans cette technique, deux oligonucléotides sont utilisés. 1/ Le **LPO** (pour *Left Probe Oligo*), constitué de deux parties : une amorce sens (*forward*, ou F) et une demie-sonde, constitué d'une courte séquence d'ADN complémentaire à la région d'intérêt. Le LPO est, de plus, marqué par un fluorochrome, afin de distinguer les différents patients étudiés dans le même temps opératoire. 2/ Le **RPO** (pour *Right Probe Oligo*) quant à lui constitué de trois parties distinctes : une demie sonde d'ADN complémentaire à la zone d'intérêt et adjacente à la séquence du LPO ; un *stuffer*, correspondant à une séquence oligonucléotidique d'une longueur connue et choisie spécifiquement pour distinguer les différents *loci* étudiés, et une amorce anti-sens (*reverse*, ou R). [Figure 14]
- le deuxième est le recours à une **ADN ligase**. Ainsi, après une première étape d'incubation des amorces en présence des ADN à étudier permettant leur hybridation, une étape de ligation est effectuée. Cette étape permet ainsi la constitution d'**un seul fragment d'ADN complémentaire** à celui de la région d'intérêt, portant, de chaque extrémité, les amorces universelles de PCR sens et anti-sens (*Forward* et *Reverse*). De cette manière, chez les patients porteurs d'une délétion au locus étudié, la ligation ne pourra s'effectuer, les séquences *LHS* et *RHS* ne pouvant pas trouver leur partenaire. Les oligos non ligués sont enfin éliminés par lavage, afin de ne pas créer d'interférence lors des dernières étapes d'amplification et de détection.
- le troisième est la **nature du produit amplifié et détecté**. Contrairement à la majorité des techniques utilisées en biologie moléculaire, en MLPA ce n'est pas directement la séquence d'ADN d'intérêt que l'on amplifie et détecte, mais bien les fragments d'ADN complémentaires issus de la ligation des deux oligos utilisés (LPO et RPO).

Tout l'intérêt de la technique réside dans le fait qu'il est possible, *via* l'utilisation de *stuffer* de différentes tailles, de distinguer **différentes régions d'une même séquence d'ADN**, correspondant en règle aux différents exons des gènes d'intérêt. L'utilisation de sondes de fluorescences distinctes, permettent en outre de passer **plusieurs ADN issus de patients différents dans le même temps opératoire**. L'utilisation d'une électrophorèse pour distinguer les zones d'intérêt, la détection des fluorescences, ainsi que leurs intensités relatives par rapport à celle mesurée sur un témoin, permettent d'estimer le nombre de copies à un locus défini.

A noter qu'une adaptation de cette technique est disponible, la MS-MLPA (pour *Methylation Specific Multiplex Ligation dependent Probe Amplification*) pour détecter les anomalies (pertes ou gain) de méthylation à certains *loci*. Cette technique permet ainsi le diagnostic de syndromes liés à des anomalies de l'empreinte comme celui des syndromes de Prader-Willi et Angelman, ou encore ceux de Silver-Russell et de Beckwith-Wiedemann.

I.3.2.2. Analyse chromosomique par puce à ADN (ACPA ou CGH)

La technique d'hybridation génomique comparative (ou *CGH* pour *Comparative Genomic Hybridization*), également appelée Analyse Chromosomique par Puce à ADN (ACPA) est plus ancienne que la MLPA, car ayant initialement été décrite en 1992. [131] Cette étude de cytogénétique moléculaire, est **pan-génomique**, car elle étudie l'ensemble du génome dans le même temps opératoire, contrairement à la MLPA qui est une technique ciblée.

Son principe est simple. Des **quantités égales d'ADN génomique** d'un patient et d'un témoin sont marquées par un fluorochrome différent, mélangées, puis mises en compétition pour l'hybridation à un ADN de référence.

Initialement décrite par Kallioniemi sur des chromosomes en métaphase étalés sur lame de verre, la technique a connu une évolution à la fin des années 1990, avec une hybridation non plus réalisées sur des chromosomes entiers, mais sur des oligonucléotides d'environ 1Mb fixés sur une lame de verre, appelée « puce ». [132-133]. Aujourd'hui, la résolution des puces est considérablement augmentée, **de l'ordre d'1Kb**, grâce à la fixation de millions d'oligonucléotides de 60 à 80 bases sur la lame de verre. Malgré cette évolution majeure, l'interprétation est demeurée la même.

Après hybridation, les signaux de fluorescence sont **numérisés par un scanner**, et un logiciel calcule les rapports de fluorescence du patient par rapport à celle du témoin (P/T). Les résultats sont visualisés au niveau de chaque oligonucléotide de la puce, sous la forme d'un point aligné ensuite sur un idéogramme du chromosome portant la séquence de l'oligonucléotide en question. Ces rapports sont exprimés sous la forme d'un **logarithme binaire**, particulièrement utile en génétique tumorale en cas d'amplifications de *loci*, car permettant de centrer les rapports autour de faibles valeurs (3 ou 4, correspondant à des rapports patient/témoin compris entre 8 et 16). L'estimation du nombre de copies à un locus donné est calculé selon la formule suivante :

$$\text{Log}_2(x) = y \leftrightarrow x = 2^y$$

avec x correspondant au rapport patient/témoin. Le nombre de copies au locus étudié est donc la moitié de ce rapport, le génome du témoin étant diploïde. [Figure 15]

Dans le cadre de la génétique **constitutionnelle**, trois valeurs de logarithme binaire sont en règle observés : **0**, correspondant à une absence d'anomalie du nombre de copies ; **-1** correspondant à une délétion hétérozygote du patient au locus donné ; et **+0,58** correspondant à une duplication hétérozygote du patient au locus étudié.

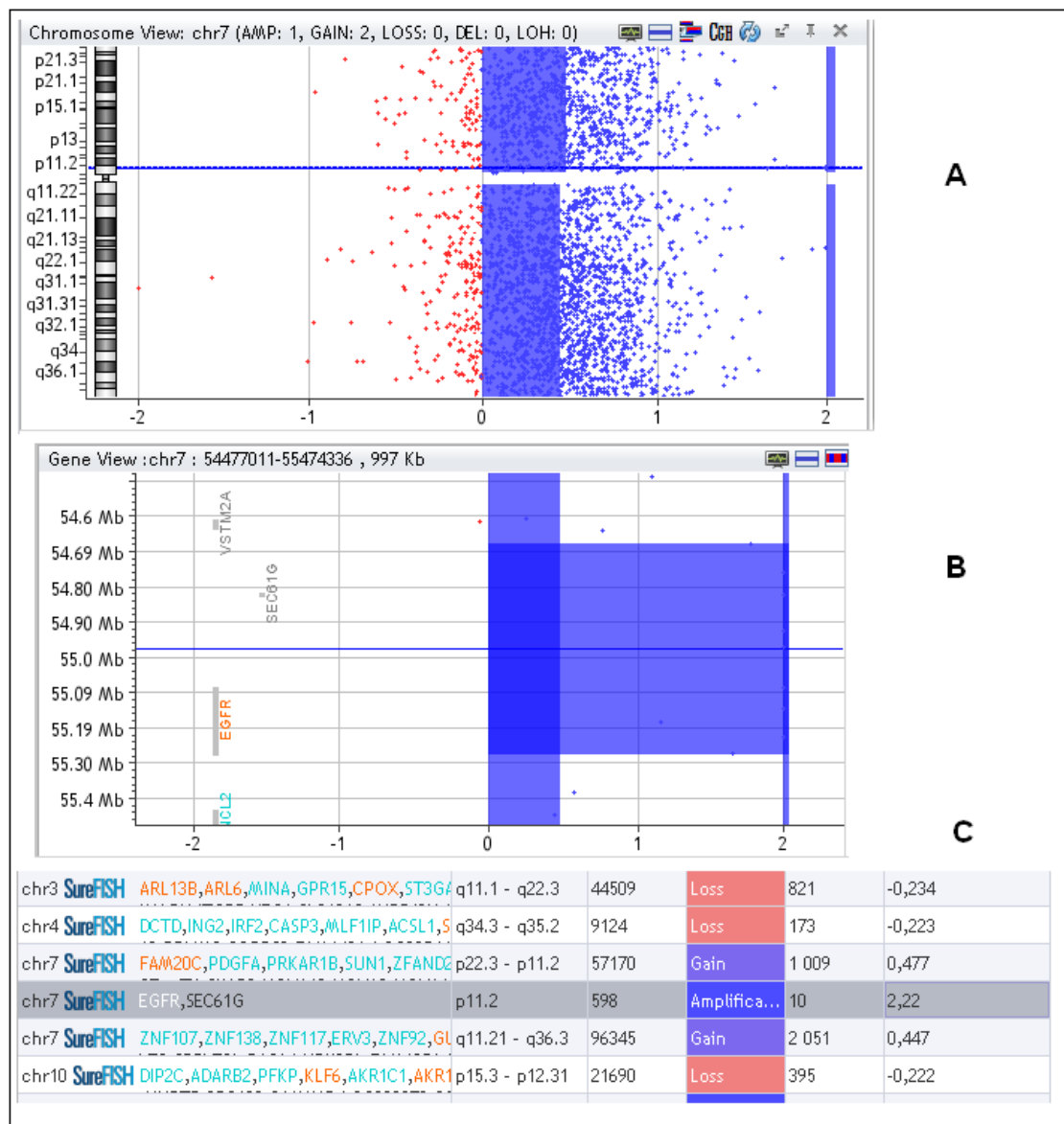


Figure 15 : Résultats d'une ACPA, exemple du locus de l'EGFR (Cytogenomics, Agilent Inc. ®). A : vue globale sur le chromosome 7, B : vue au niveau du locus d'EGFR ; C : logarithmes binaires des ratios P/T pour les loci où une anomalie du nombre de copies a été détectée

I.3.3. Interprétation des variants

L'interprétation des variants nécessite, au-delà d'une solide connaissance des techniques de biologie moléculaires utilisées et d'une certaine expertise des gènes étudiés, de l'expérience.

En 2015, un groupe dirigé par Richards [134] pour le compte du Collège Américain de Génétique Médicale (*American College of Medical Genetics and Genomics*, ou ACMG) a proposé un outil relativement simple pour aider à l'interprétation des variants détectés. L'objectif de ce travail était d'**uniformiser les pratiques**, au sein des différents laboratoires de génétique moléculaire.

Plusieurs critères pour l'interprétation sont pris en compte, sachant que dans chaque grand groupe de données, plusieurs arguments sont définis.

- Les données **épidémiologiques** : ces informations font partie des plus prépondérantes, puisque dans la majorité des laboratoires, elles constituent le **premier degré de filtration des variants**. Ainsi, partant du postulat que les pathologies d'ordre génétique sont rares en population générale, puisque correspondant à une prévalence inférieure à **5 pour 10.000** [135], les altérations moléculaires qui en sont responsables ne peuvent être retrouvées dans les bases de données des caractéristiques génétiques des individus issus de la population générale (indemnes), **à une fréquence supérieure à 0,05%**. Il faut cependant noter que ce postulat est théoriquement valable pour les pathologies dont le mode de transmission est autosomique dominant, comme la majorité des syndromes de prédisposition aux cancers.

Hors cadre de l'oncogénétique du cancer du sein, pour les syndromes dont les altérations moléculaires ségrègent sur le mode de transmission autosomique récessif, ce seuil de 0,05% peut être **adapté à la prévalence de la pathologie selon la loi de Hardy et de Weinberg**, permettant d'estimer approximativement la proportion de sujets hétérozygotes pour une pathologie autosomique récessive en se basant sur sa prévalence, selon la formule :

$$2pq = 2\sqrt{q^2}$$

avec $2pq$ correspondant à la fréquence des hétérozygotes en population générale et q^2 correspondant à la prévalence de la maladie.

- Les **données prédictives et *in silico*** : elles correspondent essentiellement aux prédictions de pathogénicité issues des différents algorithmes bio-informatiques disponibles. Ces derniers se basent eux-mêmes sur plusieurs critères, comme le **degré de conservation** du résidu en question à travers les espèces, ou encore l'**écart physico-chimique** entre le résidu de référence et celui par lequel il est substitué en cas de mutation faux-sens. De nombreux logiciels sont disponibles, parmi les plus fréquemment utilisés, **SIFT** (<https://sift.bii.a-star.edu.sg>), **Mutation Taster** (<http://www.mutationtaster.org/>) et **Polyphen2** (<http://genetics.bwh.harvard.edu/pph2/>) pour les algorithmes de prédiction de pathogénicité, et **GERP** (<http://mendel.stanford.edu/SidowLab/downloads/gerp>) pour les données de conservation du résidu. Il est important de préciser que d'autres éléments sont également pris en compte. Ils se basent notamment sur la **comparaison avec des variants délétères connus**, comme un nouveau variant nucléotidique avec la même conséquence protéique qu'un variant déjà reporté ; ou encore sur l'**effet prédit sur la protéine ou sur sa fonction**, comme un variant tronquant ou entraînant une modification de sa fonction (perte ou gain).

- Les **données fonctionnelles** : ce sont des données extrêmement précieuses pour l'interprétation des variants détectés. Cependant, elles sont relativement peu disponibles, car un nombre relativement limité de variants ont été testés en conditions expérimentales pour faire la preuve de leur caractère délétère.

La recherche de ces informations est probablement celle nécessitant le plus de temps pour le biologiste, et également celle faisant le plus appel à son expertise et à sa connaissance des gènes étudiés. Dans le cas de mutations **d'épissage** notamment, ces données sont particulièrement importantes, car les études de transcrits lorsqu'elles sont disponibles, permettent de caractériser précisément l'impact protéique (nature du transcrit majoritaire, celle du transcrit minoritaire et le rapport des deux s'il y a lieu, saut d'exon, rétention intronique induisant la création d'un codon stop prématuré, utilisation d'un site d'épissage alternatif, etc...).

Enfin, pour le gène *TP53*, une base de données, IARC TP53 (<https://p53.iarc.fr/TP53GeneVariations.aspx>), conduite à l'initiative de l'*International Agency for Research on Cancer* est disponible, et permet de préciser le caractère fonctionnel ou non fonctionnel de la protéine en fonction du variant étudié. Les données compilées se basent sur l'étude menée *in vitro* sur la levure de l'activité transcriptionnelle résultant de différents variants au sein de ce gène, de Kato et son équipe publiée en 2003 [136].

- Les **données bibliographiques** : tout comme les données fonctionnelles, elles requièrent un certain temps ainsi qu'une certaine expertise de la méthodologie de réalisation des études publiées dans la littérature. Ces données se réfèrent au poids que l'on peut accorder aux travaux affirmant le caractère délétère ou bénin d'un variant. Ici aussi, des bases de données peuvent être consultées, notamment la base de données ClinVar du *National Center for Biotechnology Information* (NCBI) (<https://www.ncbi.nlm.nih.gov/clinvar>), avec un poids prépondérant associés aux données d'interprétation des variants ayant fait l'objet d'une **relecture par des groupes internationaux experts**, comme les groupes ENIGMA (*Evidence-based Network for the Interpretation of Germline Mutant Alleles*) pour le syndrome HBOC (<https://enigmaconsortium.org/>), ou encore le groupe InSiGHT (*International Society for Gastrointestinal Hereditary Tumors*) pour les syndromes de prédisposition aux cancers digestifs (<https://www.insight-group.org/>). A côté de la base de données ClinVar, de très nombreuses autres sont disponibles, certaines **spécialisées dans un gène**, comme pour le gène *VHL* (<http://www.umd.be/VHL/>); d'autres **plus généralistes**, comme la base de données LOVD (<https://www.lovd.nl/>).
- Les données alléliques : celles-ci donnent un argument en faveur ou au contraire en défaveur de la pathogénicité d'un variant, **lorsqu'il est associé à un autre variant connu dans le même gène**. Ces données sont cependant difficilement exploitables, car elles nécessitent de savoir si les deux variants sont présents en *Cis* ou en *Trans*, et donc d'avoir accès aux génotypes des apparentés du cas-index.
- Les données de **ségrégation**, et le **caractère de novo** du variant : comme les précédentes, elles sont difficilement accessibles, et leur utilisation dans le cadre de l'oncogénétique est relativement peu fréquente. En effet, elles permettent d'attribuer

un argument en faveur de la pathogénicité du variant, lorsqu'il est **apparu de novo chez le cas-index en l'absence d'antécédents familiaux**, ou lorsqu'il **ségrége avec la pathologie dans les familles suspectes d'une forme héréditaire de prédisposition**. Or, dans le cadre de l'oncogénétique diagnostique, les études de ségrégation ne sont pas réalisées pour les variants non délétères ou de signification inconnue. Ces critères servent donc essentiellement en génétique médicale, et dans le domaine de la recherche.

- Les données **phénotypiques** : Elles permettent d'apporter un poids supplémentaire en faveur du caractère délétère d'un variant, lorsque le phénotype observé est **très évocateur d'une altération du gène étudié**. Cependant, ces données sont à manier avec précaution, en particulier dans le domaine de l'oncogénétique. La stratégie d'étude elle-même, en panel de gènes impliqués dans le même phénotype, ne permet que très rarement d'avoir recours à cet argument, tant **l'hétérogénéité inter-locus est grande**. Bien-sûr, certaines exceptions existent, lorsqu'un signe est considéré comme pathognomonique ou très spécifique à une atteinte d'un gène en particulier, comme c'est par exemple le cas de la macrocéphalie dans le syndrome de Cowden, ou le type histologique de cancer gastrique à cellules indépendantes dans le cas des mutations du gène de l'E-Cadhérine. Quoi qu'il en soit, cet argument **reste de faible poids dans l'interprétation finale** faite par le biologiste.

Cette notion de poids associé aux différents arguments développés ci-dessus permet de préciser la manière dont est, à la fin du travail de recueil de données, réalisée leur synthèse afin de procéder à l'interprétation et à la classification des variants à proprement parler. En effet, pour **chaque groupe de données, un poids plus ou moins important est associé à chaque argument** (mineur, modéré, fort, ou très fort). La combinaison des différents arguments associés à leur poids, permet au biologiste de classer chaque variant dans une des cinq classes de variants selon un algorithme proposé par Richards, et utilisé de manière universelle au sein des laboratoires de génétique moléculaire à présent.

- Les variants de classe 1 correspondent aux variants **bénins**
- Les variants de classe 2 correspondent aux variants **probablement bénins**
- Les variants de classe 4 correspondent aux variants **probablement délétères**
- Les variants de classe 5 correspondent aux variants clairement **délétères**

Une dernière catégorie est décrite, celles de variants de classe 3, correspondant aux **variants de signification inconnue** (VSI, ou *VOUS* pour *Variant of Unknown Significance*, encore appelés *UV* pour *Unclassified Variants*), impossibles à classer dans une des quatre autres classes à partir des arguments relevés, parfois discordants.

Plusieurs remarques doivent ici être faites.

- En premier lieu, les termes de « **mutations** » et de « **polymorphismes** » doivent être abandonnés. Les recommandations de l'ACMG recommandent l'utilisation préférentielle du terme de « variants » associés à leur classe ; bien que le terme « polymorphisme » se retrouve encore fréquemment en pratique courante (comprendre par-là « variant de classe 1, fréquent en population générale, classiquement à plus de 1%).
- Les critères proposés par Richards **peuvent être adaptés** aux spécificités qu'on peut rencontrer dans certaines pathologies. Ainsi, il est possible par exemple dans le cadre de l'étude des gènes MMR pour le diagnostic d'un syndrome de Lynch, de prendre en compte certaines caractéristiques tumorales évocatrices de leurs altérations. Une perte d'expression d'une ou deux des quatre protéines et/ou la présence d'une instabilité microsatellitaire au sein de la tumeur constituera, de cette façon, un argument supplémentaire en faveur du caractère délétère d'un variant identifié. Par ailleurs, les critères de fréquence en population générale peuvent également être adaptés à la fréquence de cette pathologie.
- Enfin, bien que codifiée pour assurer le maximum de reproductibilité entre laboratoires, l'interprétation des variants reste à l'appréciation du biologiste, avec la possibilité de pondérer certains arguments de son choix, en fonction du contexte clinique du patient dont il a la charge.

II. Article

Pour cette partie, l'étude réalisée ainsi que ses résultats sont présentés sous la forme du manuscrit d'un article scientifique original, soumis pour lecture et éventuelle publication dans une revue internationale de cancérologie.

Néanmoins, il convient de rappeler le but de ce travail qui était de réaliser une **étude descriptive** des variants retrouvés lors de **l'analyse constitutionnelle d'un panel de gènes** impliqués dans les **prédispositions héréditaires aux cancers** selon une stratégie par approche **NGS**, chez des femmes présentant un **cancer du sein dans le cadre d'une indication théranostique**.

Cette étude a été réalisée à partir des données collectées entre **mai 2016 et septembre 2020**, au sein du laboratoire de biologie moléculaire des cancers de **l'Institut Bergonié** à Bordeaux, où a été réalisé l'ensemble des séquençages des patientes sur cette période.

Germline Molecular Alterations in Breast Cancer Theranostic Multigenes

Panel Analysis : a Study of 103 Patients

Alexandre Perani (1), Laurence Venat-Bouvet (2), Clémentine Peyramaure (2), Françoise Bonnet-Dorion (3), Natalie Jones (3), Pierre Macquère (3), Catherine Yardin (1,4) and Nicolas Sevenet (3)

(1) Genetics unit, University Hospital of Limoges, France ; (2) Oncogenetic unit, University Hospital of Limoges, France ; (3) Molecular Genetics Laboratory, Bergonié Institute, Bordeaux, France ; (4) Univ. Limoges, CNRS, XLIM, UMR 7252, F-87000 Limoges, France.

Abstract

Detection of BRCA genes mutations in various cancers is currently predictive to PARP inhibitor sensibility. Concerning breast cancer, administration of PARP inhibitors is possible in case of germline alteration in these two genes for patients with metastatic or locally advanced disease, and their detection is to date widely possible through Next-Generation Sequencing. Rapid analysis circuits have been established, leading to definition of a recent term in oncology : « theranostic indication ». We report here a study, conducted on 103 patients with breast cancer. Multigenes panel analysis was performed, and 10 pathogenic alterations in HBOC related genes was identified (2 in *BRCA1*, 5 in *BRCA2*, 1 in *PALB2* and 2 in *RAD51C*) leading to a diagnosis performance of 9.71%. Moreover, 68 unclassified variations were detected, and 54 among them were discussed in order to provide additional arguments in favor or in disfavor of their deleteriousity. Our findings suggested the necessity to enhance links between germline and somatic genetics in order to classify detected variants of unknown significance (VUS). Genetic study of patients with breast cancer remains thus a complex challenge for oncogeneticists and molecular biologists, but new technologies and constant increasing of informations in datasets should tend to improve diagnosis rate and number of patients eligible to PARP inhibitors.

Key words : *Breast Cancer, PARP Inhibitor, BRCA1, BRCA2, Next Generation Sequencing, Multigenes panel analysis, Variants of uncertain significance, Theranostic indication*

Abbreviations : **ACMG** = American College of Medical Genetics ; **CGH** = Comparative Genomic Hybridization ; **CNV** = Copy Number Variation ; **FGCC** = French Genetic and Cancer Consortium ; **HBOC** = Hereditary Breast and Ovarian Cancer ; **HER2** = Human Epidermal growth factor Receptor 2 ; **HRD** = Homologous Recombination Deficient ; **MAF** = Minor Allele Frequency ; **MMR** = Mismatch Repair ; **MSS** = Manchester Scoring System ; **NGS** = Next Generation Sequencing ; **PAP-1** = Poly(ADP-Ribose) Polymerase-1 ; **PCR** = Polymerase Chain Reaction ; **QMPSF** = Quantitative Multiplex PCR of Short fluorescent Fragments ; **SNV** = Single Nucleotide Variant ; **TNBC** = Triple Negative Breast Cancer ; **VUS** = Variant of Unknown Significance ; **WGS** = Whole Genome Sequencing

Introduction

Based on dramatic efficiency of poly(ADP-ribose) polymerase-1 (PARP-1) inhibitors in treatment of ovarian cancer associated with deleterious germline or somatic mutations of *BRCA1* or *BRCA2* genes [1,2], a first clinical trial was conducted to assess benefits of this molecule in patients with breast cancer [3]. This randomized open-label phase III trial, named OlympiAD, compared administration of Olaparib to standard treatment in patients with human epidermal growth factor receptor 2 (Her2) negative advanced or metastatic breast cancer, and harboring germline mutation in *BRCA1* and *BRCA2* genes. Results showed a higher median progression-free survival, and a lower risk of disease progression or death for patient treated with PARP inhibitor. Following this trial, many other ones were performed [4,5] with similar conclusions. Other molecules in this new pharmacological class were moreover developed and tested, both in treatment of ovarian cancer with Niraparib [6] and Rucaparib [7], and breast cancer with Talazoparib [8]. As a consequence European Medicine Agency recently approved the use of Olaparib and Talazoparib in treatment of breast cancer, indicated after anthracyclines and taxanes treatment in patients with local advanced or metastatic stage, presenting *BRCA1* or *BRCA2* germline mutation [9,10].

Germline *BRCA* genes study is performed in routine with a genes panel approach strategy. In 2018, the French Genetic and Cancer Consortium edited guidelines, specifying 13 genes to be screened in patients with breast cancer, included in the Hereditary Breast and Ovarian Cancer (HBOC) genes panel [11]. Thus molecular data not only about *BRCA1* and *BRCA2*, but also about *PALB2*, *TP53*, *CDH1*, *PTEN*, *RAD51C*, *RAD51D*, *MLH1*, *MSH2*, *MSH6*, *PMS2*, and *EPCAM* genes are available for each tested patient. Moreover this genes list in accordance with own expertise of each laboratory. *RAD51B*, *CHEK2*, *STK11*, *ATM*, *BARD1*, *BRIP1*, and other genes are therefore often studied in the same technical time. The management of potential « incidental » or « additional » data is consequently becoming a novel issue for geneticists, even clinicians or biologists, in the same way that those found in Whole Genome Sequencing (WGS) [12]. Currently, only incidental « actionable » deleterious alterations (class 5) are reported to oncogeneticist, such as those in Mismatch Repair (MMR) genes, *MLH1*, *PMS2*, *MSH2* and *MSH6*, predisposing in Lynch syndrome. Others alterations (class 3, and not actionable mutations) do not appear on genetic report. However, such informations could be precious for future tumor medical management when they are known in early stage of cancer, or for patient's genetic counselling.

The aim of this study is to describe i) the germline genetic characteristics found in theranostic multigene analysis for breast cancer in order to determine the diagnosis performance in this indication, and ii) the possible actionable alterations found.

Materials and methods

Patients included in study

Blood samples were collected from the oncogenetics service of Limoges University Hospital (France) between May 2016 and September 2020. Patients with theranostic indication were selected for this study (n = 209).. [Figure 1]

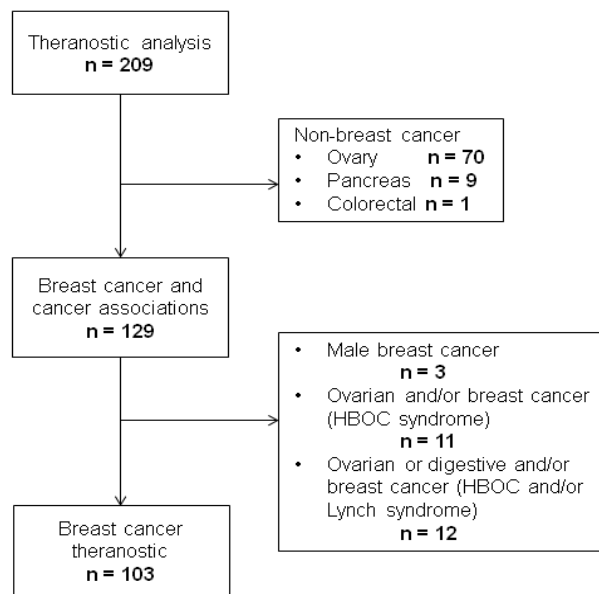


Figure 1: inclusion criteria and number of patients recruited in study : Flowchart of study. A total of 209 patients were included in study. 106 were excluded, in order to select only female with breast cancer (n=103). 26 patients were excluded because tumor was not specified (breast alone, or associated with another tumor primary site). Males (n=3) were excluded of study because they could lead to a selection bias in the estimation of the detection rate of BRCA1/2 genes alteration.

After exclusion of patients without breast cancer (n = 80), 129 prescription were assessed, and 26 were excluded (breast cancer in male patients n = 3; not discriminable tumor : breast and/or ovarian cancer n = 11 ; association of breast cancer with ovarian and/or digestive tumor, n = 12). A total of 103 patients were included, corresponding to those with metastatic breast disease, recurrent breast disease, and patients in « neoadjuvant stage » *i.e.* non-recurrent disease and non-metastatic disease, but surgical approach depending on *BRCA* genes status (breast conserving surgery vs. risk-reducing double mastectomy with/without oophorectomy).

Genetic analysis was performed in the molecular genetics laboratory of Bergonié Institute, Bordeaux (France). All patients signed a consent for germline genetic analysis. Clinical data were collected, and age, tumoral hormonal status (*i.e.* Triple Negative Breast Cancer or TNBC), histopathological data (type, grade according to Elston and Ellis, and Ki67 group as defined by Nielsen and al [13]), stage

(neoadjuvant, recurrent or metastatic disease), and nature of cancers in family were reported. In order to assess the phenotype of recruited patients, familial breast, ovarian, prostate and pancreas cancers history was assessed using Manchester Scoring System, as described by Evans in 2004 [14] and adjusted in 2008 [15]. Characteristics of included patients for the study are presented in following table. [Table 1]

Table 1 : Characteristics of recruited patients and tumors for multi-genes panel analysis

		N	Percentage	age (m ± sd)	(min - max)
Stage	neoadjuvant	79	76.70	44.89 ± 12.63	(27 – 83)
	recurrent	7	6.80	57.86 ± 14.3	(41 – 78)
	metastatic	17	16.50	59.53 ± 13.65	(40 – 87)
Tumor Hormonal status	TNBC	59	57.3	49.63 ± 15.05	(27 – 87)
	non TNBC	38	36.9	54.5 ± 8.7	(45-71)
	no data	6	5.8	44.94 ± 12.8	(59-78)
Histology	IDC	69	67.00	49.98 ± 17.52	(27 – 87)
	is IDC	2	1.94	55.5 ± 0.7	(55 – 56)
	aIDC	4	3.88	52.75 ± 14.52	(37 – 72)
	mIDC	2	1.94	36 ± 8.49	(30 – 42)
	IDC NST	10	9.7	44.1 ± 9.7	(29 – 58)
	LBC	6	5.84	57.17 ± 14.9	(38 – 72)
	no data	10	9.7	54.3 ± 13.6	(37-82)
Grade	I	2	1.94	50.5 ± 12.02	(42 – 59)
	II	27	26.2	52.3 ± 15.62	(30 – 81)
	III	59	57.3	44.46 ± 12.54	(27 – 87)
	no data	15	14.56	55.13 ± 14.0	(33-82)
Ki67	≤5	3	2.91	56.67 ± 18.17	(42 – 77)
	5 – 30	27	26.2	48.36 ± 15.83	(30 – 82)
	> 30	59	57.3	46.77 ± 13.9	(27 – 87)
	no data	14	13.59	52.2 ± 10.8	(35-72)

TNBC : Triple Negative Breast Cancer ; IDC : Intra-Ductal carcinoma ; is IDC : in situ IDC ; a IDC : aprocin IDC ; mIDC : mucinous IDC ; IDC NST : IDC of No Special Type ; LBC : Lobular Breast Carinoma. Note that lacking data for histopathological characteristics corresponded to recruited patients with a metastatic disease. When a biopsy of the metastase was performed, Ki67 and grade were assessed. Mean age under 50 and majority of TNBC validated the indication for rapid screening procedure.

Mean of age at screening was 48 years old (standard deviation: ± 14.12), but repartition of age was dispersed. [Figure 1 in appendix] Majority of patients were recruited in « neoadjuvant » stage (76.7%). Tumors were essentially grade III (67.05%) intra-ductal (74.19%), triple negative (60.82%), and with high Ki67 index (66.29% over 30%). HBOC tumor familial history was found in 44.79% of patients (data not shown), with a low Manchester score (mean : 4.33). Patients with recurrent breast cancer were either in neoadjuvant stage or with locally advanced/metastatic disease. Data were missing for one of them. As expected, metastatic patients were older than patients in neoadjuvant or recurrent stages. Patients with grade I tumor were younger than patients with grade II tumors, but youngest patients were those with grade II tumors. Same age distribution was observed for low (≤5), intermediate (5-30) and high (>30) ki67 tumors.

Libraries preparation and Next Generation Sequencing

Lymphocyte genomic DNA was isolated using Autopure automaton (QIAGEN, Hilden, Germany) according to manufacturer's instructions. In case of extraction failure, DNA was isolated manually with same reagents. DNA was quantified by fluorometry (Qubit, Life Technologies, Grand Island, NY, USA), and quality was assessed using TapeStation 2200 (Agilent Technologies, Santa Clara, CA, USA). NGS was first performed using Miseq, then NextSeq500 platform capture technology (Illumina, San Diego, CA). The studied genes in panel were designed in the unit according to The French Genetic and Cancer Consortium guidelines for diagnosis of HBOC syndrome. Genes involved in digestive cancers predisposition, Gorlin and Cowden syndromes were also studied. The 25 genes included in panel were *APC*, *ATM*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CDH1*, *CHEK2*, *EPCAM*, *MLH1*, *MRE11A*, *MSH2*, *MSH6*, *MUTYH*, *PALB2*, *PIK3CA*, *PMS2*, *PTCH1*, *PTCH2*, *PTEN*, *RAD50*, *RAD51C*, *STK11*, *SUFU*, and *TP53* (*accession number and detail of panel in Appendix Table 1*). Sensibility of SNV and indel detection was over 99% in exonic and intron-exon junctions (located in -12 and +8bp) regions. For each patient, 200 ng of DNA were used. Fragmentation was performed by sonication (Covaris, Woburn, MA, USA) in order to generate fragments of 200 bp. Amplified libraries (tagged with sequencing adapters and specific barcodes for each patient) were purified with AMPure PCR purification kit (Life Sciences, Branford, CT, USA). For capture step, SureSelect XT enrichment kit (Agilent Technologies) was used. Then re-amplified pooled libraries were loaded onto array sequencer.

Bioinformatics analysis

FASTQ generation and secondary data analysis after sequencing run was performed using Miseq Reporter (Illumina). PCR replicate, short and low quality reads were not considered. Remaining reads generated were aligned against the whole genome build hg19. Mutations were detected with BWA for single nucleotide variants and GATK for indels. Mutation were annotated with data from Variant Effect Predictor (Ensembl). BAM were visualized with Alamut Software (Interactive Biosoftware). Presence of *BRCA1*, *BRCA2* and *PALB2* copy number variation was assessed by relative sequencing depth analysis, using « CNVSeq » a tool developed in the laboratory. Each gene was divided into fragments of about 200 bp, and number of reads generated were compared to internal controls. Deletions were considered present when subject over control ratio was 0.7 or under ; duplications were considered present when ratio was 1.5 or over.

Variants interpretation and mutations confirmation

After exclusion of substitution variants with strand bias over 50% and those with supporting reads under 2%, all the variants with minor allele frequency (MAF) in ESP and/or GnomAD [16-17] databases under 0.005 were assessed. Presence of *CHEK2* gene variant c.1100delC ; p.(Thr367Metfs*15) was also checked out, because present in GnomAD database with the highest reported MAF greater than 0.005 (MAF = 0.008). Known benign and likely benign variants located in genes in panel, with or without MAF over 0.005, in sequenced genes were excluded from study. Remaining sequence variants found were interpreted using American College of Medical Genetics (ACMG) and French Genetic and Cancer Consortium (FGCC) guidelines [18,11].

In order to define a cut-off for « BS1 » evidence, corresponding to « MAF too high for disorder », frequencies repartition of known pathogenic variants in each was assessed with Varsome, an *in silico* prediction tool [19]. Based on frequency data present in dbSNP [20] Uniprot and Pubmed databases, and the highest frequency of pathogenic known variant, a cut-off was set under which 99% of known pathogenic variants are present [Appendix Table 2]. Variants with a superior MAF in general population databases (GnomAD or ESP) received a « BS1 » evidence of non-pathogenicity. « Reputable sources » supporting « BP6 » and « PP5 » argues were international consortia of each specific syndrome or gene when existing ; *i.e* « Evidence-based Network for the Interpretation of Germline Mutant Alleles » (or ENIGMA) for *BRCA1* and *BRCA2* gene mutations [21] ; and the « International Society for Gastrointestinal Hereditary Tumours » (on InSiGHT) variant classifications for MMR genes [22]. *In silico* data, *i.e* supporting « BP4 », « PP3 » and « BP7 » evidences were obtained from SIFT [23], Mutation taster [24], Polyphen2 [25], Align GVGD [26] for SNV and short indels ; NNSplice [27] and MaxEntScan [28] for splicing variants. Integrated scores were also consulted *i.e* CADD [29] and FATHMM [30], with cut-off supporting variants pathogenicity as suggested by authors (respectively 20 and 0.7).

Only pathogenic variants (ACMG class 5), likely pathogenic variants (ACMG class 4) and variants of unknown significance (ACMG class 3) in *BRCA1*, *BRCA2* and *PALB2* genes were confirmed with Sanger method using BigDye Terminator 3.1 System with ABI instrument (Applied Biosystem, UK). In case of splicing variants in these genes, effect on transcript (or mRNA) was assed using RT-PCR analysis, according standard procedures. Copy number variations were also assessed, using CGH-array or QMPSF, depending on considered gene.

Statistical analysis

Box plot and scatter plot constructions were made using GeoGebra free mathematics software available online [31]. Statistical analysis was performed using the same software. Analysis of difference among sub-groups was performed with Fisher F test (or one-way ANOVA) and Pearson correlation coefficient (noted « r ») was calculated using linear regression models. Kruskal-Wallis one-way analysis of variance was used to assess differences among age medians, with a free online tool [32]. Radar charts were constructed using Excel software [33].

Results

Variants description and diagnostic performance

Description of detected variants

We identified 108 rare or private variants satisfying study criteria in 70 patients among 103 included (67.96%), and 33 of them (32.04%) had only known ACMG class 1 and 2 variants (with or without MAF over 0.005) in sequenced genes. A large gene deletion or loss was identified using relative depth sequencing analysis, and confirmed by CGH array. Among 107 point mutations or short indels detected, 7 found in intronic sequences (*i.e.* out of canonical donor or acceptor splice sites). We found 76 non-synonymous variants, most of them (82.9%, n = 63) were missense changes. In addition, 10 truncating variants (2 nonsense and 8 frameshift, thus 13.2%), 1 in frame deletion and 2 splice site alterations were detected. No relevant variation satisfying study criteria was detected in three genes of the panel : *EPCAM*, *PIK3CA* and *PMS2*.

ACMG and FGCC guidelines were used for variants interpretation. About two thirds of variants were classified as variant of uncertain significance (68 ACMG class and 3 gene alterations among 108, 63%). We found 26 benign or likely benign rare or private genic changes (24.07% ACMG class 1 and 2 variants) and 14 likely pathogenic or pathogenic variants (4 ACMG class 4 and 10 ACMG class 5 alterations, *i.e.* 12.93%). Results are presented in **Figure 2A**.

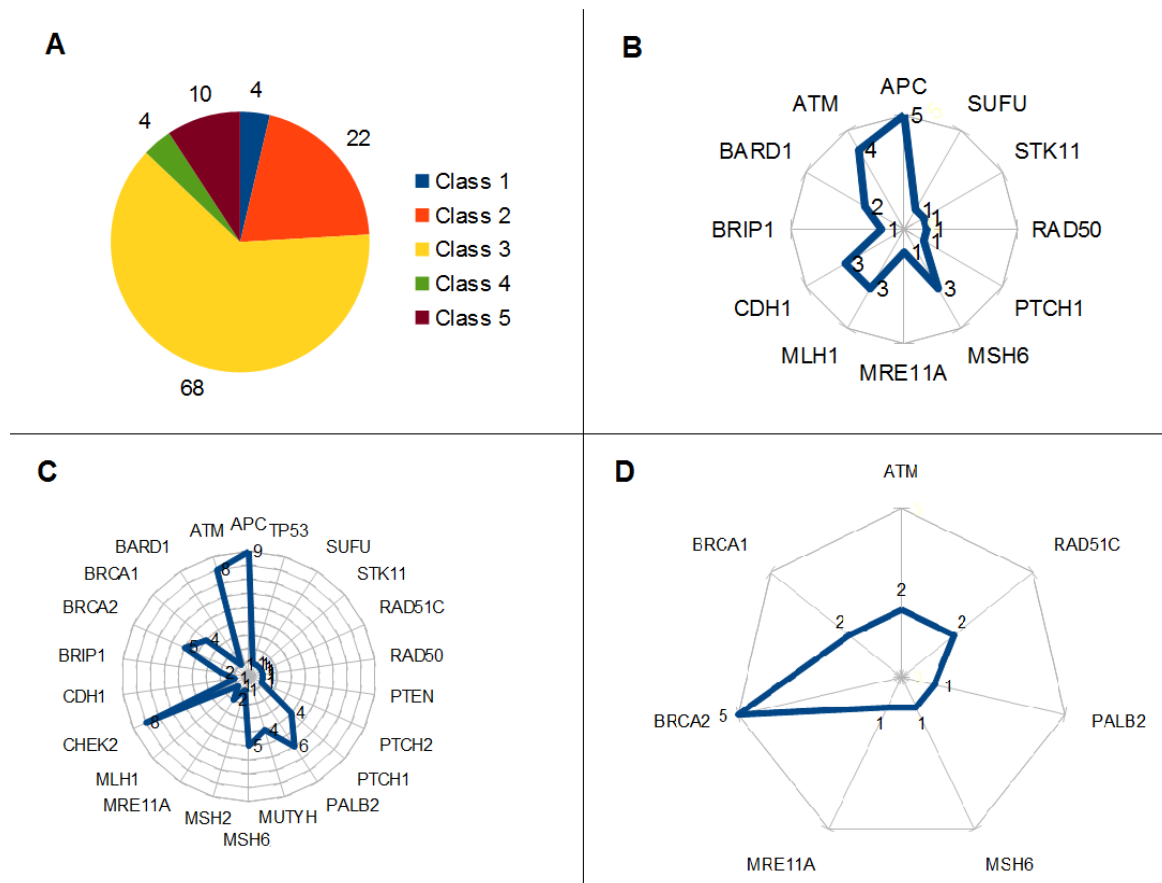


Figure 2: Results of multi-genes panel analysis. A: Number and ACMG class of variants identified. Majority of variants detected were rare or private variants of uncertain significance. Pathogenic or likely pathogenic variants corresponded to 12.93% of 108 variants detected, and benign or likely benign corresponded to 24.07%. **B: Benign and likely benign variants identified.** A third of these variants affected APC and ATM. **C: Variants of uncertain significance detected.** APC, ATM, CHEK2 and MSH6 represented about the half genes. Note that 15 VUS are located in BRCA1, BRCA2 and PALB2 genes. **D: Pathogenic and likely pathogenic variants detected.** We found 5 ATM, MRE11A and RAD51C ACMG class 4 and 5 variants, in addition of the 8 detected in BRCA1, BRCA2 and PALB2 genes. Note that 1 likely pathogenic variant was identified in MSH6 gene.

Benign and likely benign variants [Figure 2B]

We could determine the absence of pathogenicity for about a quarter of total rare or private variants detected. Once the synonymous variants distant from the exon-intron junctions were excluded ($n = 15$), 4 non-exonic variants could be classified as benign or likely benign, because of a MAF in ESP or GnomAD datasets greater than the BS1 cut-off defined as described above. A rare *MLH1* variant (rs35502531, c.1852_1853delinsGC resulting in a lysine to alanine change at codon 618) was classified as benign according to the verdict of the InSiGHT expert panel, based on multifactorial likelihood model ($p < 0.001$). In addition, a frequent in frame deletion of 21 bases between positions 1075 and 1095 affecting *BARD1* exon 4 (absent of ESP datasets) and resulting in p.(Leu359_Pro365del) was classified as benign, because present in GnomAD datasets with a MAF of 0,028 and carried by 180 reported homozygotes.

In order to classify the 5 missense variants detected, three data were collected. We assessed in a first time their relative MAF in GnomAD compared to the BS1 cut-off, the presence of homozygotes carriers reported in a second one, and eventually their CADD score. Homozygosis was reported for two of them (rs1801673, and rs498834). The three remaining missense variants (rs35187787, rs41557217 and rs41557217) had both a MAF greater than BS1 cut-off and a low CADD score, allowing them to be classified as benign. In addition, a third of ACMG class 1 and class 2 variants identified were found in *ATM* and *APC* genes, which are among the biggest sequenced (respectively 63 and 16 exons, with total exonic sequences covering 13147 and 10730 bp).

Pathogenic and likely pathogenic variants [Figure 2D]

We detected 14 ACMG class 4 and class 5 variants, concerning 7 genes: *BRCA1* (n=2), *BRCA2* (n=5), *PALB2* (n=1), *RAD51C* (n=2), *ATM* (n=2), *MRE11A* (n=1) and *MSH6* (n=1). Thus, diagnosis of Hereditary Breast and Ovarian Cancer susceptibility syndrome (HBOC) was set for 10 patients among the 103 included in the study, corresponding to a positivity rate of 9.71%. Remaining variants, affecting *ATM*, *MRE11A* and *MSH6* were not reported to the clinician, because present in the heterozygous state for the two former genes, and because classified as likely pathogenic for the latter.

Majority of variants detected (11 among 14, corresponding to 78.5%) were predicted to be truncating, due to frameshift (n=7), stop gained (n=2) or splice alteration (n=2). [Table 2]

Table 2 : Characteristics of mutated patients and with ACMG class 5 and class 4 variants detected

Syndrome	Gene	HGVS Nomenclature		Patient ID	age	TNBC ?	Histology	Grade	ki67	Neoadj.	Rec.	Meta.	MSS	other cancers in family
HBOC	BRCA1	c.1386delG	p.(Thr464Profs*11)	31	55	yes	IDC	3	80	yes	no	no	0	CRC, kidney, skin
	BRCA1	c.2043dupT	p.(Asn682*)	4	46	yes	IDC	3	40	yes	no	no	11	breast, prostate
	BRCA2	c.-227-?_67+?del	p.(?)	63	72	yes	aIDC	2	60	yes	no	no	NA	none
	BRCA2	c.2808_2811delACAA	p.(Ala938Profs*21)	10	48	no	IDC	3	50	yes	no	no	8	pancreas breast
	BRCA2	c.5208delA	p.(Asp1737Ilefs*4)	78	34	yes	IDC	3	60	yes	no	no	2	myeloma
	BRCA2	c.6209_6212delAAAAG	p.(Glu2070Valfs*10)	11	53	no	IDC	3	80	yes	no	no	9	skin
	BRCA2	c.8755-2A>C	p.(Gly2919Leufs*3)	25	42	yes	IDC NST	3	80	yes	no	no	6	CRC
	PALB2	c.635delC	p.(Pro212Glnfs*11)	102	48	no	IDC	3	40	no	yes	no	6	CRC, gastric
	RAD51C	c.732delT	p.(Ile224Metfs*9)	7	60	yes	IDC	3	70	yes	no	no	0	none
RAD51C	c.837+1G>T	p.(?)	52	55	no	IS-IDC	NA	NA	no	yes	no	8	breast, CRC	
AT	ATM	c.1464G>T	p.(Trp488Cys)	23	44	yes	IDC	3	80	yes	no	no	1	lung, hodgkin
	ATM	c.6776_6777delCT	p.(Ser2259Tyrfs*13)	41	38	no	LBC	2	NA	yes	no	no	1	pancreas breast
	MRE11A	c.1714C>T	p.(Arg572*)	92	54	NA	NA	NA	NA	no	no	yes	0	none
HNPCC	MSH6	c.3801G>A	p.(Met1267Ile)	8	40	yes	IDC	2	NA	yes	no	no	0	breast, brain

ID = identifiant ; TNBC = Triple Negative Breast Cancer ; IDC : Intra-Ductal carcinoma ; is IDC : *in situ* IDC ; a IDC : aprocin IDC ; mIDC : mucinous IDC ; IDC NST : IDC of No Special Type ; LBC : Lobular Breast Carinoma ; H = High ; Neoadj = Neoadjuvant ; Rec = Recurrent ; Meta = Metastatic ; NA = Not Applicable ; CRC = Colorectal Cancer. Note that Ki67 was not assessed for patient ID41 and ID8. No biopsy was performed for patient ID52 screened with recurrent breast disease. Patient ID63 had no information concerning her family. Breast and pancreas cancer familial history were mentioned, when affecting brothers and sisters of index-case, because not considered in Manchester Scoring (MS).

Effect of *BRCA2* c.8755-2A>C acceptor splice site alteration on transcript was assessed using RT-PCR and Sanger Sequencing, showing a frameshift due to a glycine to leucine change at codon 2919 and creation of a premature stop codon 3 positions further on protein level (data not shown). *RAD51C* c.837+1G>T donor splice site alteration (corresponding to rs760235677) was classified as likely pathogenic, according to ClinVar reports, even if this prediction has not been confirmed by published transcriptional studies. Expected effect of this variant disrupting intron 5 donor splice site on protein level results in a loss of function.

Interestingly, we detected a Copy Number Variation (CNV) affecting *BRCA2* exon 1 and 2, using relative sequencing depth analysis. CGH-array was performed to assess size and relative positions of deletion breakpoints corresponding to g.(32880759_32881119)_(32890871_32891687)del (9753bp, 28 deviated probes). Eventually, we used specific PCR and Sanger Sequencing with primers we designed, and exact breakpoints (g.32880956 and g.32891087) genomics locations could be defined (data not shown). CNV corresponded to a large deletion, with a total size of about 10.2 kb leading to a deletion of *BRCA2* exon 1 and 2 and affecting exon 1 and 2 *ZARIL* nearby gene. **[Figure 2 in Appendix]**. Based on a previous report (variation VCV000873424.1, corresponding to g.32889493_32890429del, *i.e.* exon 1 and part of *BRCA2* promoter gene deletion), and transcriptional study supporting haploinsufficiency due to deleted allele [34], CNV we detected was classified as pathogenic (ACMG class 5). However exact effect on *BRCA2* protein could not be determined.

Eventually, two missense variants affecting *ATM* and *MSH6*, corresponding respectively to W488C and M1267I were classified as likely pathogenic, based on Renault AL, et al. study [35] for the former, and predicted splice alteration for the latter (located in the first base of *MSH6* exon 8). Moreover, both variant had MAF lower than BS1 cut-off. Majority of patients with likely pathogenic or pathogenic variants were screened in neo-adjuvant stage (11/14, corresponding to 78.6%). Two were screened with a recurrent disease, and one in metastatic stage. Even if designed to estimate the probability of germline alteration in *BRCA1* and *BRCA2* genes, Manchester Score was calculated before the screening for all patients. Almost all patients with ACMG class 4 and 5 variants in these two genes had familial history of breast and/or ovarian cancer (5/6). Only patient ID31 had null Manchester score. Mean age at screening was 49 (sd \pm 9.09). Interestingly, patient ID8 with *MSH6* M1267I variant had no familial history of colorectal cancer, but a brain tumor was reported in her family. *MSH6* germline alterations are responsible in HNPCC syndrome, and patients affected are more susceptible to develop a wide spectrum of tumors, including brain tumors [36]. Eventually, one patient included in study with an ACMG class 5 variant in *ATM* gene (patient ID92) benefited of multigenes panel analysis presented a lobular breast cancer. She had a low familial history of breast and ovarian cancer (MSS=1), but a



pancreatic cancer was reported in her family, which is consistent with the recent review of Nanda and Roberts [37] supporting role of germline *ATM* alterations in pancreatic cancer.

Variants of uncertain significance [Figure 2C]

ACMG class 3 variants represented about two thirds of total variants detected (n=68 among 108), and this proportion was respected for *BRCA1*, *BRCA2* and *PALB2* genes (n=15 among 23, corresponding to 65.2%). [Figure 3]

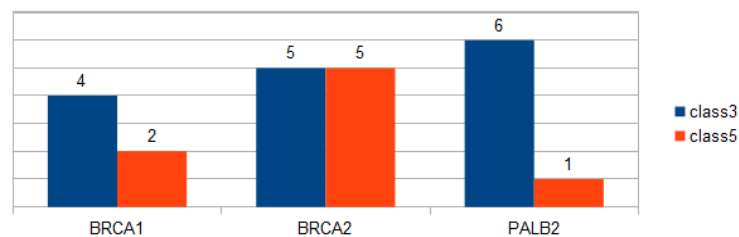


Figure 3: Respective count of ACMG class 3 and class 5 variants detected in *BRCA1*, *BRCA2* and *PALB2*. Blue bars represent UV and red ones, pathogenic variants detected. Note that none ACMG class 4 variant was detected. Proportion of UV in panel and in *BRCA/PALB2* genes is comparable (respectively 63% and 65.2%). *BRCA2* was the most impacted gene, with 10 variants identified, whatever their classes. *BRCA1* and *PALB2* counted each more UV than deleterious variants. Moreover, among UV detected in *BRCA1* and *BRCA2* genes two were synonymous variants (respectively D1337= and V1969=)

Within -12;+8bp exon-intron junctions sequenced, 3 intronic variants were detected that we could not classify using ACMG of FGGC guidelines. Two of them were reported in GnomAD or ESP datasets with a low MAF, lower than BS1 cut-off used : *MUTYH* c.989-9C>T (MAF = 0.0003, BS1 cut-off : 0.00039) and *CHEK2* c.320-5T>A (MAF = 0.0004, BS1 cut-off : 0.00053). The last one, found in *MRE11A* gene was absent of datasets consulted (BS1 cut-off : 0.00027). We found no published transcriptional studies concerning these variants, and used *in silico* splice alteration prediction tools generated conflicting conclusions.

Moreover, we detected 9 synonymous variants that we could not classify as benign or likely benign, because it satisfied the contradictory PM2 (corresponding to variant absent from the datasets of unaffected individuals) and BP7 (*i.e.* a synonymous variant with no predicted splicing alteration) evidences of ACMG. These variants corresponded to *ATM* : H269= and K971= ; *BRCA1* : D1337= ; *BRCA2* : V1969= ; *MSH6* : P1082= ; *PTCH1* : H189= ; *PTEN* : Y240= ; *STK11* : G15= ; and *SUFU* : F197=). Only *BRCA1* and *BRCA2* synonymous variants were reported to clinicians. Interestingly, a frame shift variant affecting *PTCH2* exon 12 gene was detected. This variation (rs1329842135) corresponding to c.1526dupA and p.(Asn509Lysfs*55) on protein level, was absent from ESP datasets but reported in GnomAD (MAF =0.000011, lower than BS1 cut-off used : 0.00012). This variant was

detected in patient ID72, aged 50 at screening, and presenting recurrent grade 3 IDC (Ki67 : 60%) breast cancer. Neither skin cancer nor keratocystic odontogenic tumors were reported in personal or familial history, consistent with the absence of Gorlin-Like syndrome affecting her. Consequently, we could not classify this variation, and its significance remained uncertain. Remaining 55 variations that we were not able to classify were missense variants, affecting *APC* (n=9), *ATM* (n=6), *BARD1* (n=1), *BRCA1* (n=3), *BRCA2* (n=4), *BRIP1* (n=2), *CDH1*(n=1), *CHEK2* (n=7), *MLH1* (n=1), *MRE11A* (n=1), *MSH2* (n=1), *MSH6* (n=4), *MUTYH* (n=3), *PALB2* (n=6), *PTCH1* (n=3), *RAD50* (n=1), *RAD51C* (n=1) and *TP53* (n=1) genes. Their characteristics are developed in « Prioritization of missense Unclassified Variants » following part.

Prioritization of missense Unclassified Variants (VUS)

Then, we assessed unclassified missense variants. Among 55 variations identified, one was rejected (*CHEK2* c.1116_1117inv, K373E), because with 2% supporting reads. In order to provide additional arguments in favor or in disfavor of their deleteriousity, and to identify variations possibly actionable, we collected genetic tumoral data (total report count in COSMIC [17], OncoKB verdict [18]), and integrated scores were also consulted *i.e* CADD [29] and FATHMM (when available in COSMIC) [30], with cut-off supporting variants pathogenicity as suggested by authors (respectively 20 and 0.7) [Figure 4]. Moreover, location of variant (in a structural or functional known domain, noted SFD, of the protein or not) was reported.

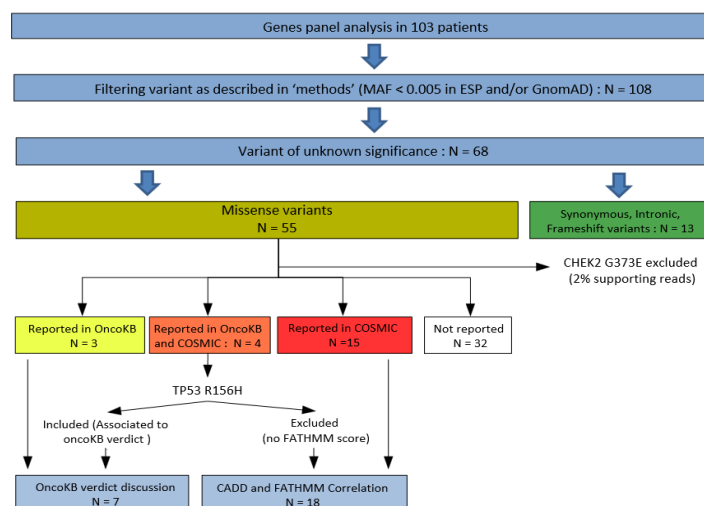


Figure 4 : Variants filtering for VUS interpretation Among 108 variations identified, 68 were classified as VUS. 55 among them were missense variations. *CHEK2* G373E was excluded because detected with 2% supporting reads (possibly corresponding to a sequencing artifact). Variants reported in OncoKB and/or COSMIC were assessed to discuss OncoKB verdict, or correlation between CADD and FATHMM scores.

Among 54 variants, we identified 24 with a CADD score below 20, and 30 with a CADD score over this cut-off. About a third (9/24) of the former, and half (n = 15) of the latter impacted SFD. Regardless their CADD score, 19 were reported in COSMIC database: 6 had a FATHMM prediction score under 0.7 and 12 had a FATHMM over this cut off (FATHMM score of TP53 R156H was not available on COSMIC). A positive correlation was found between these two scores ($r = 0.8453$). [Figure 5]

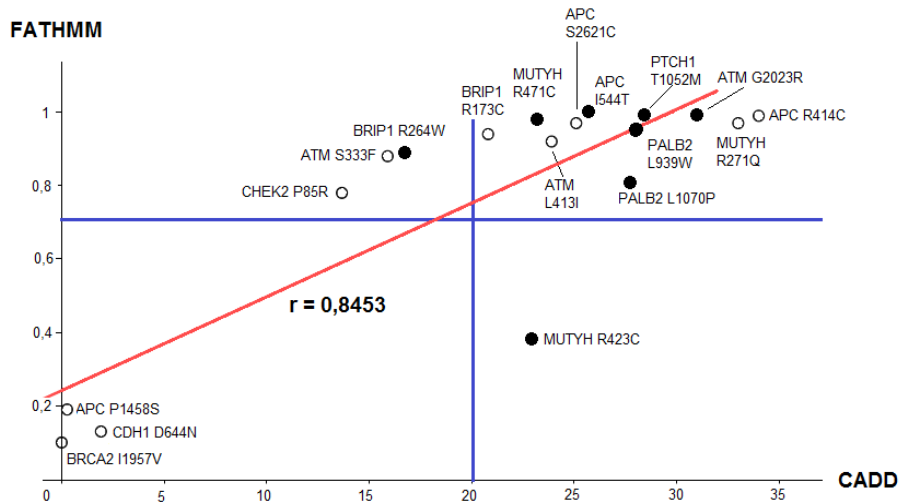


Figure 5: Correlation between CADD and FATHMM scores of UV detected and reported in COSMIC datasets. Blue bars figure cut-offs recommended for CADD and FATHMM score interpretation (respectively 20 and 0.7). Each variant we detected corresponds to a circle. White circles denote variations not affecting structural or functional domains, and black circle those located in SFD. A positive linear correlation was found (red line, $r = 0.8453$), despite the presence of scattered values.

Majority of variants that we detected and reported in COSMIC (14/18 *i.e.* 77.8%) had concordant CADD and FATHMM scores. Interestingly, none variations with concordant CADD score below 20 and FATHMM score below 0.7 cut-off (corresponding to rs143796828: *APC* P1458S, rs587781696 : *CDH1* D644N and rs80358817 : *BRCA2* I1957V) are located in SFD. We detected 11 variations with concordant high CADD and FATHMM score: 3 in *APC* (R414C, I544T, and S2621C), 2 in *ATM* (L413I and G2023R), 1 in *BRIP1* (R173C), 2 in *MUTYH* (R271Q, and R471C), 2 in *PALB2* (L939W and L1070P) and 1 in *PTCH1* (T1052M). About a half are located in SFD (6/11), and 5 impacted 3 genes involved in homologous recombination (*ATM*, *PALB2* and *BRIP1*). However we observed 4 variations with discordant predictions, corresponding to *MUTYH* R423C, *BRIP1* R264W, *ATM* S333F and *CHEK2* P85R. Half of these variants impacted SFD (rs150792276 : *MUTYH* R423C, and rs28997569 : *BRIP1* R264W). When available, OncoKB verdict was consulted. [Table 3]

Table 3 : OncoKB verdict and characteristics of missense unclassified variations detected

OncoKB classification	Variant	Functional or structural domain ?	GnomAD	CADD	FATHMM	Cosmic Count
likely neutral	BRCA1 R170Q	no	0.000034	1.252	NF	NF
Inconclusive	ATM S333F	no	0.0016	15.9	0.88	9
	BRCA2 L3011P	no	0	28.1	NF	NF
Likely oncogenic	CHEK2 E64K	no	0.00016	23	NF	NF
	PALB2 L939W	yes	0.00092	28	0.95	3
	TP53 R156H	yes	0.000013	23.7	NA	15
oncogenic	BRIP1 R173C	yes	0.0026	20.8	0.94	4

NF = not found in datasets, NA = not assessed. No OncoKB benign unclassified variation was detected. We detected 2 variants with inconclusive verdict in *ATM* and *BRCA2*. 4 variations we classified as ACMG class 3 are classified likely oncogenic and 1 as oncogenic, and 3 of them are potentially actionable (*CHEK2*, *PALB2* and *BRIP1*).

Among 54 unclassified missense variations that we detected, 7 were reported in OncoKB datasets. Interestingly, 3 of them were not reported in COSMIC. No variation is classified as neutral, but BRCA1 R170Q (with CADD score = 1.252 and located outside SFD) is mentioned as likely neutral. 2 variations that we detected have inconclusive verdict: ATM S333F, with discordant CADD and FATHMM score and located outside SFD, and BRCA2 L3011P, with high CADD score (28.1) but absent of COSMIC datasets. Interestingly, 4 variations are classified as Likely oncogenic or oncogenic in OncoKB database, and 3 of them are involved in homologous recombination, and 2 (*PALB2* L939W and *BRIP1* R173C, impacting SFD) have concordant high CADD and FATHMM score. OncoKB verdict provided additional arguments in favor of the deleterious nature of the two variants *CHEK2* E64K and *TP53* R156H, respectively absent of COSMIC and present in these datasets but without FATHMM score associated. Moreover, the former variation (rs141568342) impacted a gene involved in homologous recombination.

Patients Sub-groups analysis

Patients were divided into three groups, to assess their characteristics in term of age, tumoral stage and phenotype, or cancer familial breast and ovarian cancer history. Sub-group A (n=33) corresponded to patients with no relevant variation (*i.e.* only presence of variants not satisfying study criteria, with MAF in datasets greater than 0,005, or known ACMG class 1 and 2 in genes of panel). Sub-group B (n=22) consisted of patients with ACMG class 3, class 4 and class 5 variants in *BRCA1*, *BRCA2* and *PALB2* genes. It corresponded to patients whose variations were reported to prescriber. Sub-group C (n = 48) regrouped patients with rare or private variations affecting others genes of the panel, regardless their ACMG classification after we interpreted them. Patients with 2 variants or more were included in sub-group B if one of them were located in *BRCA1*, *BRCA2* or *PALB2* gene; and in sub-group C if variants were located in other gene in panel.

Mean age in sub-group A (mean 50, standard deviation 15.54) was greater than in sub-groups B (mean 45, standard deviation 12.06) and C (mean 48, standard deviation 14), but no statistically significant difference was found ($p=0.53$). However, a statistically significant difference was observed between mean age in sub-group A versus overall patients mean age (48 sd 14.12 ; $p=0.022$). Characteristics of patients in the three sub-groups are presented in [Table 4].

Table 4: Characteristics of patients, tumoral stage, phenotype, and familial history of breast and ovarian cancer in each sub-group

		Sub-goup A (n=33)		Sub-group B (n=22)		Sub-group C (n=48)		p value
		N	%	N	%	N	%	
Age	<30	1	3.03	1	4.55	1	2.08	0.711
	30-39	9	27.27	5	22.73	15	31.25	0.355
	40-49*	9	27.27	9	40.91	12	25.0	0.028
	50-59	7	21.21	5	22.73	12	25.0	0.799
	>60*	7	21.21	2	9.09	8	16.67	0.041
Stage	neoadjuvant*	26	78.79	19	86.36	34	70.83	0.022
	recurrent	1	3.03	1	4.55	5	2.08	0.487
	metastatic*	6	18.18	2	9.09	9	18.75	0.041
THS	TNBC	21	67.74	12	57.14	26	57.78	0.212
	no data	2	-	1	-	3	-	-
Grade	I	0	0	1	5.0	1	2.63	0.087
	II	10	33.33	4	20.0	13	34.21	0.052
	III	20	66.67	15	75.0	24	63.16	0.179
	no data	3	-	2	-	10	-	-
Ki67	<5	0	0	2	9.52	1	2.5	0.366
	5-30*	6	21.43	4	19.05	15	37.5	0.006
	>30*	22	78.57	15	71.43	24	60.0	0.013
	no data	5	-	1	-	8	-	-
HBOC history	yes	13	41.93	11	55	19	42.2	0.086
	no data	2	-	2	-	3	-	-

THS = Tumor Hormonal Status; HBOC = Hereditary Breast and Ovarian Cancer ; TNBC = Triple Negative Breast Cancer. Note that lacking data for histopathological characteristics corresponded to patient recruited with a metastatic disease. When a biopsy of the metastasis was performed, Ki67 (3 cases) and grade (2 cases) was assessed. Asterisks denote statistically significant difference among ratios.

Age repartitions were comparable among sub-groups, once younger patients (age <30) were excluded. Sub-group A was the most homogenous, with about a quarter of patients in each remaining categories. In sub-group B, the majority of patients was aged between 40 and 49 years (40.91%, *versus* 27.27% in A sub-group and 25.0% in C sub-group, $p=0.0268$). In sub-group C majority of patients were younger (between 30 and 39), but no significant difference was found. Majority of patients with a detected germline alteration in *BRCA1*, *BRCA2* or *PALB2* genes was recruited in neoadjuvant stage (86.36%). A statistically significant difference was observed when compared to rate of patients recruited in neoadjuvant stage in other sub-groups ($p=0.022$). Repartition of patients stages was comparable

between sub-groups A and C, but the lower rate of patients screened in metastatic stage was observed in sub-group B (9.09% *versus* 18.18% and 18.75% in sub-groups A and C respectively, $p=0.041$).

Surprisingly, detection of germline alteration in *BRCA1*, *BRCA2* or *PALB2* genes did not influence the rate of patients with triple negative breast cancer. When comparing sub-group B rate (57.14%) to sub-groups A and C (respectively 67.74% and 57.78%), no difference was shown ($p=0.212$), denoting the relative homogeneity of tumor hormonal status in overall patients included. No statistically significant difference among sub-groups of tumoral rate was shown, supporting that repartition was comparable in each subgroups, with a majority of high grade tumors and a minority of low grade tumors. However, difference among Ki67 rate was found statistically significant. Contrary to what could have been expected, and even if repartition of Ki67 is respected among sub-groups, sub-group C showed a higher rate of intermediate Ki67 tumors (37.5% versus 21.43% and 19.05% in sub-groups A and B respectively, $p=0.006$) and a lower rate of high Ki67 tumors (60% versus 78.57% and 71.43% in sub-groups A and B respectively, $p=0.013$).

Moreover, rates of patients with familial breast and ovarian cancer history was compared among sub-groups, but even if this rate is higher in sub-group B (55%) than those observed in other sub-groups (41.93% and 42.2% in A and C one respectively), statistical analysis failed in showing a significant difference ($p=0.086$). Furthermore, count and ages of patients was reported for each category. [Appendix Table 4]. No statistically significant difference was found among age median in sub-groups using Kruskal-Wallis non-parametric test (data not shown).

Discussion

Multigenes panel analysis in breast cancer has two main objectives in theranostic indication : the detection of patients which could benefit of PARP-1 inhibitors administration or surgical procedure adaptation (conserving vs risk-reducing surgery), and diagnosis of hereditary predisposition to breast cancer. Part of breast cancer due to *BRCA1* and *BRCA2* gene pathogenic alteration represents about 5 to 10%, when a genetic factor is suspected in about 20% of breast cancers [40]. Next Generation Sequencing is an effective method to answer these questions, with sensitivity close to 100% for point mutation identification [41,42], and large rearrangements detection, precised with complementary methods, such as CGH-array or MLPA.

Various predictive models have been developed in order to increase diagnosis performance, by estimation of the probability for a patient to have a germline alteration in these two genes [43-44]. Manchester scoring system designed by Evans and colleagues [14] is one of the most practical in clinical routine. Among our 103 patients included for multigenes panel analysis, we identified 7 patients with

deleterious germline alteration in *BRCA1* and *BRCA2* genes, corresponding to a rate of 6.71%, which is consistent with the low Manchester score of included patients with HBOC tumor familial history. However, Manchester scoring system as used in our study has limitations, such as the fact that it does not consider HBOC history affecting siblings or histopathological tumor characteristics of proband. Adjustments have been made and a second (MSS2) [15] then a third version of MSS [45] were presented integrating weightings according to tumor characteristics. However, the addition of supplementary parameters makes the estimation sensitive to missing data, and elaboration of MSS3 it-self was based on 71% of complete pathological reports. As a consequence we accepted that Manchester scores in our cohort were probably undervalued, especially since the majority of patients recruited had no HBOC familial history.

In our cohort, *BRCA1* and *BRCA2* gene alterations represented 60% of total pathogenic variants detected in HBOC genes, and larger studies, like this conducted in 2018 by Momozawa and colleagues in 7051 unselected patients [46] showed a comparable rate. In addition, we detected one large rearrangement in *BRCA2*, consisting in deletion of exon 1 and 2. Compared to point mutations (single nucleotide variants and short indels), this molecular mechanism represents about 10% of alterations affecting *BRCA1* and *BRCA2* genes [47-48], and expertise of molecular biologist is necessary to define their consequences. In our small cohort, one patient out of 7 with altered *BRCA1* and *BRCA2* genes has a large deletion, representing 14% of the total number close to the 10% reported by Jenkins et al, and Ewald and al.

Next-generation sequencing is helpful in the diagnosis of hereditary predisposition to breast and ovarian cancers, but one of its limitations is the rate of detected variants of unknown significance [49]. In our study, these variations represented about two thirds of all identified variations (68 among 108). ACMG criteria are useful to standardize the interpretation of variations found in NGS, but in theranostic indications it is not possible to assess segregation, *de novo* or allelic data, which contributes to this high rate of unclassified variations. In the particular field of oncology, databases of somatic mutations in cancer such as COSMIC and OncoKB are precious to provide additional arguments in interpretation of germline variations identified. In our study, it was thus possible to identify 4 “suspicious” variations : such as CHEK2 E64K (rs141568342) classified as « likely oncogenic » in OncoKB datasets based on yeast assay assessing CHEK2-mediated response to DNA damages [50] ; or BRIP1 R173C (rs4988345) classified as « oncogenic » based on *in vitro* study of Lei et Vorechovsky [51]. Nevertheless, these conclusions should be considered with precautions. Indeed, PALB2 L939W (rs45478192) classified as « likely oncogenic », was showed to have a mild impact on homologous recombination [52], and its benign character was confirmed by a study based on 42671 breast cancer cases and 42164 controls [53] showing absence of statistical difference of prevalence of this variant in cases and controls (OR = 1.05, 95 % IC = 0.83–1.32, $p = 0.70$). To date, determination of homologous recombination deficiency is

possible in ovarian tumor, by detection of BRCA gene mutations and genomic instability resulting in loss of heterozygosity, telomeric allelic imbalance and large-scale state transitions for *BRCAness*-HRD tumors. Such tests are not yet validated in breast cancer but their use could be a precious help for interpretation of germline unclassified variations detected in genes involved in homologous recombination, as functional studies on yeast used for TP53 variations interpretation. Indeed, last « suspicious » variation detected, TP53 R156H (rs371524413) and classified as « likely oncogenic » in OncoKB datasets, but was recently reviewed by ClinGen TP53 Variant Curation Expert Panel and considered as an unclassified variant. Quesnel et colleagues [54] first reported this variation in probands with Li-Frauméni syndrome due to compound heterozygosity R156H and R267Q on maternal allele associated to R290H on paternal one. Mother of index-case presented a metachronous bilateral breast cancer at ages of 35 and 43, and father was unaffected, suggesting deleterious character of TP53 R156H moreover detected by whole-exome sequencing at somatic state in adrenocortical carcinoma [55]. However, functional studies of Kato et al [56] and Giacomeli et al [57] did not show p53 loss of function or dominant effect of this variant, leading to inconclusive pathogenicity verdict.

Eventually, multigenes panel analysis strategy may generate unsolicited informations. Through the analysis of various genes, diagnosis rate could be increase in our study to 9.71%, with identification of two deleterious variant in RAD51C, but two pathogenic variants in ATM, and one in MRE11A were also detected. To date, there is no recommendation to mention heterozygous pathogenic variants in these two genes, but considering the very close links existing between germline and somatic mutations in cancer, detection of non-BRCA gene alterations should incite to determine molecular characteristics of breast tumors. Such a strategy could be helpful to identify patients who could potentially be included in ongoing clinical trials, assessing efficacy of PARP inhibitors, or other drugs in treatment of non-BRCA mutated HRD tumors.

Conclusion : The theranostic Multigenes Panel Analysis used in our study for 103 patients presenting breast cancer leads to near 10% of diagnosis rate for hereditary predisposition to breast cancer. Such strategy may help in one hand to guide the treatment, for example for the use of PARP inhibitors, and in the other hand is essential in the familial follow-up.

Acknowledgements : We would like to thank Molecular Genetics Laboratory's technicians of Bergonié Institute and Limoges CHU, for their technical expertise.

Disclosure of Interest : Authors declare no conflict of interests.

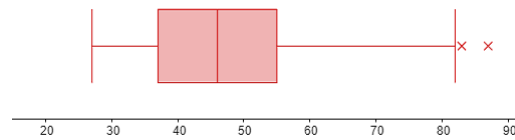
References

- [1] Audeh MW, et al., Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and recurrent ovarian cancer: a proof-of-concept trial. *Lancet*. 2010 Jul 24;376(9737):245-51.
- [2] Ledermann J, et al., Olaparib maintenance therapy in platinum-sensitive relapsed ovarian cancer. *N Engl J Med*. 2012 Apr 12;366(15):1382-92. doi: 10.1056/NEJMoa1105535. Epub 2012 Mar 27.
- [3] Robson M, et al., Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. *N Engl J Med*. 2017 Aug 10;377(6):523-533.
- [4] Griguolo G, et al., Olaparib for the treatment of breast cancer. *Expert Rev Anticancer Ther*. 2018 Jun;18(6):519-530.
- [5] Le D, et al., Olaparib tablets for the treatment of germ line BRCA-mutated metastatic breast
- [6] Mirza MR, et al., Niraparib Maintenance Therapy in Platinum-Sensitive, Recurrent Ovarian Cancer. *N Engl J Med*. 2016 Dec 1;375(22):2154-2164.
- [7] Coleman RL, et al., Rucaparib maintenance treatment for recurrent ovarian carcinoma after response to platinum therapy (ARIEL3): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet*. 2017 Oct 28;390(10106):1949-1961.
- [8] Litton JK, et al., Talazoparib in Patients with Advanced Breast Cancer and a Germline BRCA Mutation. *N Engl J Med*. 2018 Aug 23;379(8):753-763.
- [9] <https://www.ema.europa.eu/en/medicines/human/EPAR/lynparza>
- [10] <https://www.ema.europa.eu/en/medicines/human/EPAR/talzenna>
- [11] Moretta J, 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 [The French Genetic and Cancer Consortium guidelines for multigene panel analysis in hereditary breast and ovarian cancer predisposition]. *Bull Cancer*. 2018 Oct;105(10):907-917.(French)
- [12] Pujol, P., et al., Guidelines for reporting secondary findings of genome sequencing in cancer genes: the SFMPP recommendations. *Eur J Hum Genet* 26, 1732–1742 (2018).
- [13] Nielsen TO, et al., Assessment of Ki67 in Breast Cancer: Updated Recommendations from the International Ki67 in Breast Cancer Working Group. *J Natl Cancer Inst*. 2020 Dec 28
- [14] Evans DG et al., A new scoring system for the chances of identifying a *BRCA1/2* mutation outperforms existing models including BRCAPRO, *Journal of Medical Genetics* 2004;41:474-480
- [15] Evans DG, and al., Probability of *BRCA1/2* mutation varies with ovarian histology: results from screening 442 ovarian cancer families. *Clin Genet*. 2008 Apr;73(4):338-45
- [16] Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (URL: <http://evs.gs.washington.edu/EVS/>) [Accessed September 2020]
- [17] Karczewski K.J, et al., The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020(581) :434–443
- [18] Richards S, et al., Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015 May;17(5):405-24
- [19] Christos K, et al., VarSome: the human genomic variant search engine, *Bioinformatics*. 2019 June;35(11):1978-1980
- [20] Sherry,S.T., Ward,M. and Sirotkin,K. (1999) **dbSNP**—Database for Single Nucleotide Polymorphisms and Other Classes of Minor Genetic Variation. *Genome Res.*, 9, 677–679

- [21] Spurdle AB, et al., ENIGMA--evidence-based network for the interpretation of germline mutant alleles: an international initiative to evaluate risk and clinical significance associated with sequence variation in BRCA1 and BRCA2 genes. *Hum Mutat.* 2012 Jan;33(1):2-7
- [22] Plazzer JP, et al., The InSiGHT database: utilizing 100 years of insights into Lynch syndrome. *Fam Cancer.* 2013 Jun;12(2):175-80
- [23] Vaser R., et al. SIFT missense predictions for genomes. *Nat Protoc* 11, 1–9 (2016)
- [24] Schwarz JM, et al., MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods.* 2014 Apr;11(4):361-2
- [25] Adzhubei IA, Schmidt S, et al., A method and server for predicting damaging missense mutations. *Nat Methods.* 2010 Apr;7(4):248-9
- [26] Tavtigian SV, et al., Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J Med Genet.* 2006 Apr;43(4):295-305
- [27] Reese MG, et al., Improved Splice Site Detection in Genie. *J Comp Biol.* 1997 4(3) : 311-23
- [28] Yeo G, et al., Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol.* 2004;11(2-3):377-94
- [29] Rentzsch P, et al., CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* 2019 Jan 8;47(D1):D886-D894
- [30] Shihab HA, et al., Predicting the functional consequences of cancer-associated amino acid substitutions. *Bioinformatics.* 2013 Jun 15;29(12):1504-10
- [31] GeoGebra, 2020. *Classic - GeoGebra.* [online] Geogebra.org. Available at: <<https://www.geogebra.org/classic#spreadsheet>> [Accessed September 2020]
- [32] Kruskal-Wallis Test for k=3. Richard Lowry 2001-2017. [online] Vassarstats.net. Available at : <<https://web.archive.org/web/20170525004804/http://vassarstats.net/kw3.html>>
- [33] Microsoft Corporation. Microsoft Excel [Internet]. 2007. Available from: <https://office.microsoft.com/excel>
- [34] Casadei S, et al., Characterization of splice-altering mutations in inherited predisposition to cancer. *Proc Natl Acad Sci U S A.* 2019 Dec 16;116(52):26798–807
- [35] Renault AL, et al., Morphology and genomic hallmarks of breast tumours developed by ATM deleterious variant carriers. *Breast Cancer Res.* 2018 Apr 17;20(1):28
- [36] Zhao YS, Hu FL, Wang F, Han B, Li DD, Li XW, Zhu S. Meta-analysis of MSH6 gene mutation frequency in colorectal and endometrial cancers. *J Toxicol Environ Health A.* 2009;72(11-12):690-7
- [37] Nanda N, Roberts NJ. *ATM Serine/Threonine Kinase and its Role in Pancreatic Risk.* *Genes (Basel).* 2020 Jan 17;11(1):108
- [38] Bamford S, et al., The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br J Cancer.* 2004 Jul 19;91(2):355-8
- [39] Chakravarty D, et al., OncoKB: A Precision Oncology Knowledge Base. *JCO Precis Oncol.* 2017 Jul;2017:PO.17.00011
- [40] Claus EB, Schildkraut JM, Thompson WD, Risch NJ. The genetic attributable risk of breast and ovarian cancer. *Cancer.* 1996 Jun 1;77(11):2318-24

- [41] Palma MD, Domchek SM, Stopfer J, Erlichman J, Siegfried JD, Tigges-Cardwell J, Mason BA, Rebbeck TR, Nathanson KL. The relative contribution of point mutations and genomic rearrangements in BRCA1 and BRCA2 in high-risk breast cancer families. *Cancer Res.* 2008 Sep 1;68(17):7006-14
- [42] Feliubadaló, L., Lopez-Doriga, A., Castellsagué, E. et al. Next-generation sequencing meets genetic diagnostics: development of a comprehensive workflow for the analysis of BRCA1 and BRCA2 genes. *Eur J Hum Genet* 21, 864–870 (2013)
- [43] Antoniou AC, Pharoah PP, Smith P, Easton DF. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. *Br J Cancer.* 2004 Oct 18;91(8):1580-90
- [44] Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med.* 2004 Apr 15;23(7):1111-30
- [45] Evans DG, Harkness EF, Plaskocinska I, Wallace AJ, Clancy T, Woodward ER, Howell TA, Tischkowitz M, Laloo F. Pathology update to the Manchester Scoring System based on testing in over 4000 families. *J Med Genet.* 2017 Oct;54(10):674-681
- [46] Momozawa, Y., Iwasaki, Y., Parsons, M.T. *et al.* Germline pathogenic variants of 11 breast cancer genes in 7,051 Japanese patients and 11,241 controls. *Nat Commun* 9, 4083 (2018)
- [47] Judkins T, Rosenthal E, Arnell C, Burbidge LA, Geary W, Barrus T, Schoenberger J, Trost J, Wenstrup RJ, Roa BB. Clinical significance of large rearrangements in BRCA1 and BRCA2. *Cancer.* 2012 Nov 1;118(21):5210-6
- [48] Ewald IP, Ribeiro PL, Palmero EI, Cossio SL, Giugliani R, Ashton-Prolla P. Genomic rearrangements in BRCA1 and BRCA2: A literature review. *Genet Mol Biol.* 2009 Jul;32(3):437-46
- [49] van Marcke C, Collard A, Vikkula M, Duhoux FP. Prevalence of pathogenic variants and variants of unknown significance in patients at high risk of breast cancer: A systematic review and meta-analysis of gene-panel data. *Crit Rev Oncol Hematol.* 2018 Dec;132:138-144
- [50] Roeb W, Higgins J, King MC. Response to DNA damage of CHEK2 missense mutations in familial breast cancer. *Hum Mol Genet.* 2012 Jun 15;21(12):2738-44
- [51] Lei H, Vorechovsky I. BACH1 517C-->T transition impairs protein translocation to nucleus: a role in breast cancer susceptibility? *Int J Cancer.* 2003 Apr 10;104(3):389-91
- [52] Boonen et al., Functional analysis of genetic variants in the high-risk breast cancer susceptibility gene PALB2. *Nat Commun* 10, 5296 (2019) ; low risk OR 1,05 (Southey MC, et al., PALB2, CHEK2 and ATM rare variants and cancer risk: data from COGS. *J Med Genet.* 2016 Dec;53(12):800-811
- [53] Catucci I, et al., The PALB2 p.Leu939Trp mutation is not associated with breast cancer risk. *Breast Cancer Res.* 2016 Nov 9;18(1):111
- [54] Quesnel S, et al., p53 compound heterozygosity in a severely affected child with Li-Fraumeni syndrome. *Oncogene.* 1999 Jul 8;18(27):3970-8
- [55] Juhlin CC, et al., Whole-exome sequencing characterizes the landscape of somatic mutations and copy number alterations in adrenocortical carcinoma. *J Clin Endocrinol Metab.* 2015 Mar;100(3):E493-502
- [56] Kato S, et al., Understanding the function-structure and function-mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. *Proc Natl Acad Sci U S A.* 2003 Jul 8;100(14):8424-9
- [57] Giacomelli AO, et al., Mutational processes shape the landscape of TP53 mutations in human cancer. *Nat Genet.* 2018 Oct;50(10):1381-1387

Appendix



App. Figure 1 : Age of patients included in study. Scatter plot showing age repartition of patients included in the study. The whiskers indicate the lowest (27 yo) and highest (82 yo) values. Crosses (83 and 87 yo) stand for outliers. The edges of the box are the 25th (37 yo) and 75th (55 yo) percentiles. The vertical line in the box denotes the median (46 yo).

App. Table 1 : Multi-gene panel characteristics

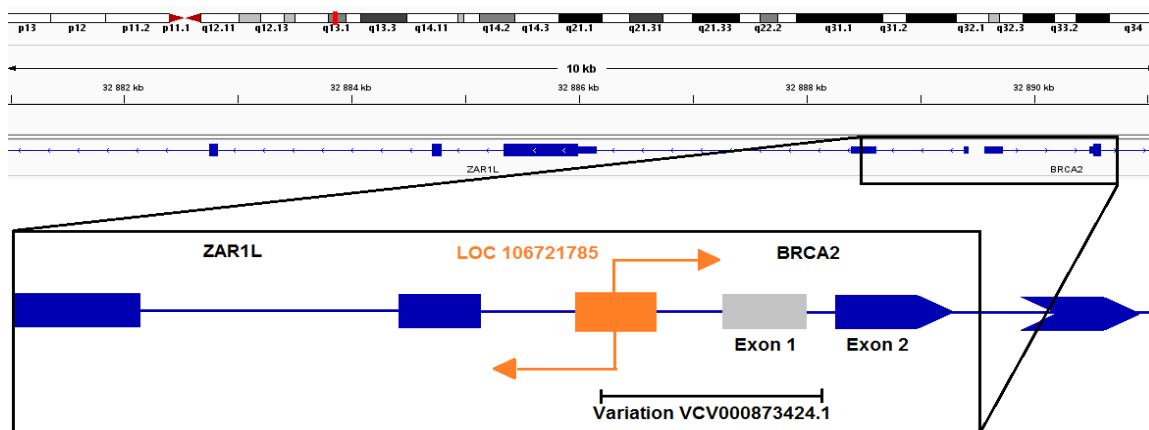
Gene	Chr	Genomic start	Genomic end	Accession number	Exon count	Total exon size (bp)
APC	5	112005202	112219936	NM_000038.5	16	10730
ATM	11	108091559	108241826	NM_000051.3	63	13147
BARD1	2	215676428	215591262	NM_000465.2	11	2607
BRCA1	17	41279500	41194312	NM_007294.3	23	7207
BRCA2	13	32887617	32975809	NM_000059.3	27	11386
BRIP1	17	59942920	59754547	NM_032043.2	20	8166
CDH1	16	68769195	68871444	NM_004360.3	16	4815
CHEK2	22	29139822	2081731	NM_007194.3	15	1858
EPCAM	2	47594287	47616167	NM_002354.2	9	1718
MLH1	9	37032841	37094337	NM_000249.3	19	2662
MRE11A	11	94229040	94148466	NM_005591.3	20	5164
MSH2	2	47628263	47712360	NM_000251.3	16	3145
MSH6	2	48008221	48036092	NM_000179.2	10	4328
MUTYH	1	45808142	457922914	NM_012222.2	16	1921
PALB2	16	23654678	23612483	NM_024675.3	13	4058
PIK3CA	3	178864311	178954500	NM_006218.2	21	3712
PMS2	7	6050737	6010870	NM_000535.5	15	2836
PTCH1	9	98281247	98203264	NM_000264.3	24	7943
PTCH2	1	45310616	45283516	NM_003738.4	22	4298
PTEN	10	89621195	89730532	NM_000314.4	9	5547
RAD50	5	131890616	131982313	NM_005732.3	25	6597
RAD51C	17	56767963	56813692	NM_058216.1	9	1282
STK11	19	1203798	1230434	NM_000455.3	10	3276
SUFU	10	104261719	104359214	NM_016169.3	13	4948
TP53	17	7592868	7569720	NM_000546.5	11	2586
Total					453	125937

Table presenting gene characteristics sequenced in our panel. Genomic start and genomic end are reported. A total of 453 exons are sequenced for each patient, representing about 126Kb. APC, ATM, and BRCA2 are the biggest genes studied, with exon size over 10Kb. Note that concerning EPCAM gene only terminal exonic deletions are assessed, and concerning PMS2 gene only exons 6, 7, 8 and 10 are considered, due to homology with pseudogenes for other exons.

App. Table 2 : Cut-off used for « BS1 »

Gene	Pathogenic variants (N)	Highest frequency known for pathogenic variant	Variant ID	Cut-off used	Se	Sp
APC	1072	0.0002295	rs587779353	0.00023	99.9	11.9
ATM	1387	0.002383	rs11212587	0.00042	97.9	12.5
BARD1	232	0.000186	rs746325928	0.00019	98.7	19.5
BRCA1	3196	0.004068	rs80357914	0.0024	100	11.3
BRCA2	3971	0.005857	rs80359550	0.0012	99.8	11.6
BRIP1	386	0.0013	rs1483527885	0.001282	99.7	7.3
CDH1	248	0.00006202	rs771085839	0.0001	100	22.5
CHEK2	445	0.02471	rs17879961	0.0011	93.3	5.5
MLH1	1057	0.001848	rs637550650	0.0019	99.9	9.7
MRE11A	84	0.0002721	rs759130031	0.00027	97.6	13.2
MSH2	1042	0.000435	rs63750790	0.00044	99.7	11.5
MSH6	905	0.00174	rs63749973	0.00086	99.7	6.8
MUTYH	243	0.004917	rs36053993	0.0049	89.3	5.9
PALB2	735	0.002033	rs180177102	0.0012	99.6	8.1
PTCH1	318	0.0003086	rs752765582	0.00031	99.7	46.8
PTCH2	2	0.00012	rs121434397	0.00006176	100	80.7
PTEN	649	0.0001955	rs587782455	0.00019	99.5	17.9
RAD50	269	0.00462	rs760146707	0.00065	98.5	7.4
RAD51C	158	0.0005227	rs587782818	0.00053	99.4	8.3
STK11	222	0.0001394	rs764449808	0.00014	99.5	19.5
SUFU	38	0.0001148	rs1060501109	0.00012	97.4	25.8
TP53	1055	0.000231	rs587781371	0.00019	99.6	22.1
Total	17714				99.9	17.5

Table presenting cut-off used for « BS1 » ACMG criteria. Numbers of pathogenic known variants considered for each gene are presented. Sensibility (« Se ») and Specificity (« Sp ») base on 17714 pathogenic and 14988 benign variants was calculated. Note that no BS1 cut-off was defined for EPCAM, PIK3CA and PMS2 genes, because without variants satisfying study criteria.



App. Figure 2 : Characteristics of 10,2Kb deletion found, affecting BRCA2 exon 1 and 2. View of the large deletion affecting BRCA2 exon 1 and 2, using IGV genome viewer. This large CNV, deleting about 10.2 kb affects exon 1-2 of BRCA2 gene, and part of ZAR1L gene (blue line and boxes in upper part of figure). In the focus region in lower part, bidirectional LOC106721785 promoter is represented in orange box. Grey box correspond to non-coding BRCA1 exon 1. Blue boxes represented exons of BRCA2 and ZAR1L gene. Variation VCV000873424.1 is represented using horizontal black line. This deletion is reported as pathogenic in Clinvar, based on Casadei S, et al., transcript analysis [33], supporting no transcription from deleted allele. (Lower part of figure not at scale).

App. Table 3 : Age means of sub-group patients for each characteristics

		Sub-groupe	N	age (m \pm sd)	(min – max)
Stage	neoadjuvant	A	26	47.81 \pm 14.66	(29 – 83)
		B	19	43.95 \pm 11.17	(27 – 72)
		C	34	43.18 \pm 8.64	(28 – 77)
	recurrent	A	1	50.0	50
		B	1	48.0	48
		C	5	61.4 \pm 15.85	(41 – 78)
	metastatic	A	6	63.67 \pm 14.5	(45 – 87)
		B	2	51.8 \pm 12.05	(53 – 72)
		C	9	56.11 \pm 13.49	(40 – 82)
THS	TNBC	A	21	52.33 \pm 16.78	(34 – 87)
		B	12	45.42 \pm 12.11	(27 – 72)
		C	26	49.38 \pm 14.86	(28 – 82)
	no data	-	6	54.5 \pm 8.7	(45-71)
Grade	I	A	0	-	-
		B	1	42.0	42
		C	1	59.0	59
	II	A	10	52.5 \pm 13.18	(37 – 81)
		B	4	52.5 \pm 17.41	(30 – 72)
		C	13	52.1 \pm 17.97	(30 – 78)
	III	A	20	49.0 \pm 17.18	(29 – 87)
		B	15	42.01 \pm 8.52	(27 – 55)
		C	24	42.17 \pm 9.11	(28 – 60)
	no data	-	15	55.13 \pm 14.0	(33-82)
Ki67	≤ 5	A	0	-	-
		B	2	46.5 \pm 6.36	(42 – 51)
		C	1	77.0	77
	5 – 30	A	6	48.17 \pm 10.16	(39 – 67)
		B	4	45.5 \pm 18.14	(32 – 72)
		C	15	47.87 \pm 19.3	(20 – 82)
	> 30	A	22	50.36 \pm 17.76	(29 – 87)
		B	15	45.33 \pm 11.86	(27 – 72)
		C	24	44.35 \pm 10.49	(28 – 68)
	no data	-	14	52.21 \pm 10.8	(35-72)
HBOC History	yes	A	13	52.69 \pm 14.12	(37 – 83)
		B	11	41.55 \pm 8.70	(27 – 53)
		C	19	47.53 \pm 15.65	(30 – 78)
	no	A	18	47.44 \pm 15.68	(29 – 87)
		B	9	46.88 \pm 12.62	(30 – 72)
		C	26	49.03 \pm 12.9	(28 – 82)
no data	-	7	49.14 \pm 15.4	(31-72)	

Table presenting mean age of patients for each tumoral or personal characteristics considered in the study. No statistically significant difference was shown among age median in sub-groups, using Kruskal-Wallis non-parametric test (data not shown).

Conclusion

Notre étude réalisée sur 103 patientes présentant un cancer du sein souligne parfaitement les enjeux et limites de la stratégie diagnostique utilisée actuellement dans le cadre de l'oncogénétique théranostique de cette pathologie. Parmi les patientes incluses ayant bénéficié d'une analyse constitutionnelle selon une approche NGS du panel de gènes recommandé par le GGC dans l'exploration de ce type de cancer, il a été possible d'identifier 10 altérations délétères, dont 7 au sein des gènes *BRCA1* et *BRCA2*. Ainsi, nous avons pu retenir le diagnostic de prédisposition héréditaire aux cancers du sein et de l'ovaire pour ces patientes, avec un rendement diagnostique de 9,71% sur l'ensemble de la cohorte.

La majorité de ces altérations détectées étaient de type SNV ou *short indels* (7/10), avec pour conséquence la synthèse d'une protéique tronquée, soit par décalage du cadre de lecture (6/7), soit par la création d'un codon de terminaison de traduction prématuré (1/7). Deux altérations de l'épissage ont également été détectées, ainsi qu'un réarrangement de grande taille entraînant une large lésion du gène *BRCA2*, tous confirmés par d'autres techniques et soulignant l'importante nécessité de combiner plusieurs approches d'étude du génome pour les caractériser.

De par la particularité de l'oncogénétique théranostique du cancer du sein, sous-tendue par des études validées chez des sujets présentant une altération des gènes *BRCA1* et *BRCA2* à l'état constitutionnel, ces résultats répondent de manière indirecte à un double objectif. Ils permettront ainsi non seulement une adaptation de la prise en charge thérapeutique des cas-index (soit par la réalisation d'une chirurgie non conservatrice, soit par l'administration de thérapies ciblant les PARP), mais également une prise en charge de leur famille, en détectant et en suivant les individus asymptomatiques porteurs de l'altération familiale.

En raison de ses mécanismes d'oncogenèse multiples, le cancer du sein demeure cependant une maladie complexe. A côté des altérations des gènes de prédisposition majeure comme *BRCA1* et *BRCA2*, d'autres sont décrites, plus fréquentes en population générale et associées à une plus faible pénétrance de la maladie, posant la question de la stratification du risque, enjeu majeur dans la prise en charge de cette pathologie. Ainsi, dans notre étude, trois altérations délétères supplémentaires ont été détectées, deux affectant le gène *ATM*, la dernière touchant le gène *MRE11A*. S'agissant de variants détectés à l'état hétérozygote impliqués dans des pathologies se transmettant sur le mode autosomique récessif, et en l'absence de recommandations consensuelles de suivi spécifiques concernant la gestion du sur-risque de cancer du sein chez ces individus porteurs sains, leur présence n'a pas été mentionnée au prescripteur. Rappelons ici que la récente classe thérapeutique des inhibiteurs de PARP, initialement développée pour les sujets portant une altération délétère des gènes *BRCA1* et *BRCA2*, est actuellement à l'étude pour déterminer les possibilités d'extension de son indication aux altérations d'autres gènes impliqués dans le mécanisme de recombinaison homologue, que leur contribution à un phénotype tumoral HRD soit prouvée ou supposée.

Les technologies de séquençage haut-débit ont donc progressivement permis l'exercice d'une médecine de précision dans le cadre de la prise en charge du cancer du sein. Mais elles sont également à l'origine de nouveaux défis pour les généticiens moléculaires. L'augmentation du nombre d'analyse chaque année, associée à l'élargissement des indications conduit à l'augmentation du nombre de nouveaux variants identifiés, très rares ou privés, dépassant la vitesse d'incrémentation des bases de données ou celle de conception de modèles fonctionnels nécessaire à leur interprétation. Comme le montre l'analyse des 54 variants faux-

sens de signification inconnue détectés dans le cadre de l'étude, l'avenir semble aller, afin d'offrir le meilleur diagnostic moléculaire possible, à la combinaison des approches, l'analyse tumorale pouvant aider à l'interprétation des variants de signification inconnue détectés au niveau constitutionnel. Il nous a ainsi été possible de repérer quatre variations (affectant *CHEK2*, *PALB2*, *TP53* et *BRIP1*) pour lesquelles des données supplémentaires seraient nécessaires pour argumenter leur rôle dans la contribution du phénotype, ou la sensibilité aux inhibiteurs de PARP. De manière plus large, prouver un phénotype *ATM déficient*, détecter une perte d'hétérozygotie, ou encore mettre en évidence les conséquences génomiques d'une déficience en recombinaison homologue, peut constituer une aide précieuse pour classer des variants d'interprétation difficile, car trop rares pour avoir été correctement phénotypés à l'état constitutionnel.

Enfin, l'approche d'étude en panel de gènes de plus en plus large, conduit également à de nouvelles problématiques éthiques lorsque des variants potentiellement actionnables sont détectés. Ces données non désirées conduisent le généticien moléculaire à questionner sa place dans la prise en charge globale du patient, et à le faire sortir de son rôle de diagnosticien d'une situation de prédisposition héréditaire, en suggérant ou non la réalisation d'analyses de génétique complémentaires en tissu tumoral afin d'améliorer sa prise en charge.

En raison de l'augmentation prévisible du nombre de biomarqueurs identifiés dans cette pathologie, fort est à parier que dans un avenir proche l'écart entre génétique constitutionnelle et tumorale tendra à se réduire, avec la réalisation de ces deux niveaux d'analyse de manière combinée, voire associées à d'autres techniques innovantes telles que l'étude de transcrit, l'étude de marqueurs épigénétiques, ou encore l'étude de cicatrices génomiques (déséquilibre allélique des régions télomériques, LOH, ou grandes cassures chromosomiques) ouvrant le champ d'une nouvelle ère multi-omique en oncogénétique.

Références bibliographiques

1. World Cancer Research Fund, American Institute for Cancer Research. « Breast cancer statistics » [En Ligne]. Disponible sur : <https://www.wcrf.org/dietandcancer/cancer-trends/breast-cancer-statistics>
2. Laudisio et al.(2018) Obesity and breast cancer in premenopausal women : Current evidence and future perspectives. *European Journal of Obstetrics and Gynecology and Reproductive Biology* 230 :217-221
3. INCa (2019) Estimations nationales de l'incidence et de la mortalité par cancer en France métropolitaine entre 1990 et 2018
4. Institut National du Cancer. Cancer du sein : facteurs de risques – « Prédipositions génétiques » [En Ligne]. Disponible sur : <https://www.e-cancer.fr/Patients-et-proches/Les-cancers/Cancer-du-sein/Facteurs-de-risque/Predispositions-genetiques>
5. Azevedo Queiroz et al.(2018) Nutritional and environmental risk factors for breast cancer : a case-control study. *Scientia Medica* 2018 ;28(2) :ID28723. <http://doi.org/10.15448/1980-6108.2018.228723>.
6. Friedson et al., Assessing and managing breast cancer risk : clinical tools for advising patients. *Med GenMed*.2004.6-8
7. Momenimovahed et al.(2019) Epidemiological characteristics of and risk factors of breast cancer in the world. *Breast Cancer – Targets and Therapy* 11 :151-164
8. Chen et al.(2019) Physical activity and risk of breast cancer :a meta-analysis of 38 cohort studies in 45 study reports. *Value in Health* 22 :104-128
9. Kenemans et al.(2004) Oncogenic pathways in hereditary and sporadic breast cancer. *Maturitas* 49 : 34-43
10. Kuchenbaecker et al.(2017) Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. *Journal of American Medical Association* 317(23) :2402-2416
11. Miki et al.(1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 266 :66-71
12. Wooster et al.(1995) Identification of the breast cancer susceptibility gene BRCA2. *Nature* 378 : 789-792
13. Caputo et al.(2012) Description and analysis of genetic variants in French hereditary breast and ovarian cancer families recorded in the UMD-BRCA1/BRCA2 databases. *Nucleic Acids Research* 40 : D992-D1002
14. M Tubiana et al., *Biologies* 331 2008
15. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000 Jan 7;100(1):57-70.
16. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011 Mar 4;144(5):646-74
17. Sherma et al., Epidermal growth factor receptor mutations in lung cancer. *Nat Rev cancer* 2007, march 7(3) :169-181

18. Knudson et al., Mutation and cancer : statistical study of retinoblastoma. Proc Natl Acad Sci, 1971 68 820-3
19. Bonaïti-Pellier et al., Rétinoblastome : de l'épidémiologie génétique à la biologie moléculaire, Med/Sci, 6 : 972-9, 1990
20. Nichols et al., Loss of heterozygosity of essential genes represents a widespread class of potential cancer vulnerabilities. Nat Commun, 11 (2517) 2020
21. Lee et al., Human Retinoblastoma susceptibility gene : cloning, identification, and sequence. Science 235 : 1394-1399, 1987
22. <https://www.omim.org/>
23. Antoniou et al. Am J Hum 2003 ; 72 117-30 Genet
24. Moran et al., Risk of cancer other than breast or ovarian in individuals with BRCA1 and BRCA2 mutations. Fam Cancer. 2012 (2)235-242
25. Leongamornlert et al., Germline BRCA1 mutations increase prostate cancer risk. Br J Cancer. 2012(10) :1697-1701
26. Iqbal et al., The incidence of pancreatic cancer in BRCA1 and BRCA2 mutation carriers. Br J Cancer. 2012(12)2005-2009
27. Mavaddat et al., Cancer risks for BRCA1 and BRCA2 mutation carriers : results from prospective analysis of EMBRACE. J Natl Cancer Inst. 2013 (11) :812-822
28. Easton et al., Gene-Panel Sequencing and the Prediction of Breast-Cancer risk. NEJM. 2015(23)2243-2257
29. Kobayashi et al., Hereditary breast and ovarian cancer susceptibility genes (review). Oncol Rep. 2013(30) :1010-1029
30. Antoniou et al., Breast-cancer Risk in Families with Mutations in PALB2. N Engl J Med. 2017(371) :497-506
31. Ginsburg et al., BRCA1 and BRCA2 families and the risk of skin cancer. Fam Cancer 2010 (9): 489–493
32. Yang et al., Cancer Risks associated with Germline PALB2 Pathogenic variants : An international study of 524 families. J Clin Oncol. 2020(7)674-684
33. McLaughlin et al., Long-term ovarian cancer survival associated with mutation in BRCA1 or BRCA2. J Natl Cancer Inst. 2013(105):141–8
34. Fergus et al., Inherited mutation in 17 Breast Cancer susceptibility genes among a large triple-negative Breast Cancer Cohort unselected for family history of breast cancer. J Clin Oncol 2015 (4) :304-311
35. Chapuis et al., Clinico-pathological characteristics of BRCA1- and BRCA2-related breast cancer. Semin Surg Oncol. 2000(18) :287-295
36. Eerola et al., Relationship of patients' age to histopathological features of breast tumours in BRCA1 and BRCA2 and mutation-negative breast cancer families. Breast Cancer Res. 2005(7) :465-469
37. HAS, Dépistage du cancer du sein en France : identification des femmes à haut risque et modalités de dépistage, Avril 2020, consultable sur : <https://www.has->

sante.fr/upload/docs/application/pdf/2011-05/note_de_cadrage_depistage_du_cancer_du_sein_-_identification_des_femmes_a_haut_risque_et_modalites_de_depistag.pdf

38. Eisinger et al., Identification and management of hereditary predisposition to cancer of the breast and the ovary. *Bull Cancer*. 2004(91) :219-237
39. Cohen-Haguenaer, Prédiposition héréditaire au cancer du sein (1). *Med/Sci*. 2019(35) :138-151
40. Pearl LH et al., Therapeutic opportunities within the DNA damage response. *Nat Rev Cancer*. 2015 Mar ;15(3) :166-80
41. Li X, Heyder WD. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res*. 2008 Jan ;18(11) :99-113
42. Yanzinski SA, Functions, Regulation, and Therapeutic Implication of the ATR Checkpoint Pathway. *Annu Rev Genet*. 2016 Nov 23 ;50 :155-173
43. Maréchal A, Zou L. DNA damage sensing by the ATM and ATR kinases. *Cold Spring Harb Perspect Biol*. 2013 Sep 1 ;5(9) :a012716
44. Deng et al., BRCA1 : cell cycle chekpoint, genetic instability, DNA damage response and cancer evolution. *Nucl. Ac. Reas*. 2006 ; 34 :1416-1426
45. Buisson et al., Fonctions des supprimeurs de tumeurs PALB2 et BRCA2 dans la réparation des cassures double brin de l'ADN, *Med/Sci* 2013 ; 29 : 301-7
46. Wu W, Koike A, Takeshita T, Ohta T. The ubiquitin E3 ligase activity of BRCA1 and its biological functions. *Cell Div*. 2008 Jan 7;3:1
47. Chen et al., BRCA1 is a 220-kDa nuclear phosphoprotein that is expressed and phosphorylated in a cell cycle-dependent manner. *Cancer Res*. 1996(56) :3168-3172
48. Gudas et al., Cell cycle regulation of BRCA1 messenger RNA in human breast epithelial cells. *Cell Growth Differ*, 1996 (7) :717-723
49. Cousineau et al., BRCA1 regulates RAD51 function in response to DNA damage and supresses spontaneous sister chromatid replication slippage : implications for sister chromatid cohesion, genome stability, and carcinogenesis. *Cancer Res*, 2005(65)11384-11391
50. Ludwig et al., Targeted mutations of breast cancer susceptibility gene homologs in mice : lethal phenotypes of Brca1, Brca2, Brca1/Brca2, Brca1/p53, and Brca2/p53 nulligous embryos. *Genes Dev*, 1997(11)1226-1241 ;
51. Brodie et al., BRCA1-associated tumorigenesis : what have we learned from knockout mice ? *Tends Genet*, 2001(17) :S18-S22
52. Venkitaraman et al., Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell*, 2002(108)171-182
53. Chen et al., The BRC repeats in BRCA2 are critical for RAD51 binding and resistance to methyl methanesulfonate treatment *Proc Natl Acad Sci USA* 1998 95(9) : 5287-5292
54. Careira et al., The BRC repeats of BRCA2 modulate the DNA-Binding Selectivity of RAD51 ; 2009 *Cell* 136 1032-1043

55. Lee et al., Cycling with BRCA2 from DNA Repair to Mitosis, *Exp Cell Resea* 2014 329 78-84
56. Vaughn et al., Cell Cycle control of BRCA2, *Cancer Res*, 1996(56):4590-4594
57. Rahman et al., PALB2, which encode a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nature Genet.* 2007(39) :165-167
58. Xia et al., Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Molc. Cell.* 2006(22) :719-729
59. Rantakari et al., Inactivation of Palb2 gene leads to mesoderm differentiation defect and early embryonic lethality in mice. *Hum. Molec. Genet.* 2010(19) :3021-3029
60. Buisson et al., Breast cancer proteins PALB2 and BRCA2 stimulate polymerase eta in recombination-associated DNA synthesis at bloked replication forks. *Cell Rep* 2014, 6(3) 553-564
61. Aedma SK, Kasi A. Li-Fraumeni Syndrome. [Updated 2020 May 22]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 Jan-. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK532286>
62. Zerdouni et al., Germline TP53 mutations result into a constitutive defect of p53 DNA binding and transcriptional response to DNA damage. *Hum Mol Genet.* 2017(26)2591-2602
63. Valdez et al., Li-Frauméni syndrome : a paradigm for the understanding of hereditary cancer predisposition. *Br J. Haematol.* 2017(176) :539-552
64. Weitzel et al., Somatic TP53 variants frequently confound germ-line testing results. *Genet Med.* 2018 Aug;20(8):809-816
65. Chebab et al., Chk2/hCds1 functions as a DNA damage chekpoint in G-1 by stabilizing p53, *Genes Dev*, 14 :278-288 2000
66. Lee et al., hCds1-mediated phosphorylation of BRCA1 regulates the DNA damage response. *Nature* 404: 201-204, 2000
67. Nelen MR et al., Localization of the gene for Cowden disease to chromosome 10q22-23. *Nat Genet.* 1996 May ;13(1) :114-6
68. Pilarski R, Eng C. Will the real Cowden syndrome please stand uo (angain) ? Expanding mutational and clinical spectra of the PTEN harmartoma tumour syndrome. *J Med Genet.* 2004 May ;41(5) :323-6
69. Freudsperger C et al., EGFR-PI3K-AKT-mTOR signaling in head and neck squamous cell carcinomas : attractive targets for molecular-oriented therapy. *Expert Opin Ther Targets*, 2011 Jan ;15(1)63-74
70. Zhou XP et al., Germline inactivation of PTEN and dysregulation of the phosphoinositol-3-kinase/Akt pathway cause human Lhermitte-Duclos disease in adults. *Am J Hum Genet.* 2003 Nov ;73(5) :1191-8
71. Pavlidou A, Vlahos NF. Molecular alterations of PI3K/Akt/mTOR pathway : a therapeutic target in endometrial cancer. *ScientificWorldJournal.* 2014 Jan 12 ;2014 :709-736
72. Malbuchi S, et al., The PI3K/AKT/mTOR pathway as a therapeutic target in ovarian cancer. *Gynecol Oncol.* 2015 Apr ;137(1) :173-9

73. Hobert JA, Eng C. PTEN hamartoma tumor syndrome : an overview. *Genet Med.* 2009 Oct ;11(10) :687-94
74. Tan MH, et al., Lifetime cancer risks in individuals with germline PTEN mutations. *Clin Can Res.* 2012 Jan 15 ;18(2) :400-7
75. Orloff MS, et al., Germline PIK3CA and AKT1 mutations in Cowden and Cowden-like syndromes. *Am J Hum Genet.* 2013 Jan 10 ;92(1) :76-80
76. Berx G, et al., Cloning and characterization of the human invasion suppressor gene E-cadherin (CDH1). *Genomics.* 1995 Mar 20 ;26(2) :281-9
77. Riethmacher D, et al., A targeted mutation in the mouse E-cadherin gene results in defective preimplantation development. *Proc Natl Acad Sci USA.* 1995 Jan 31 ;92(3) :855-9
78. Jou TS, et al., Genetic and biochemical dissection of protein linkages in the cadherin-catenin complex. *Proc Natl Acad Sci USA.* 1995 May 23 ;91(11) :5067-71
79. Wijnhoven BP, et al., E-cadherin-catenin cell-cell adhesion complex and human cancer. *Br J Surg.* 2000 Aug ;87(8) :992-1005
80. Carneiro F, et al., Model of the early development of diffuse gastric cancer in E-cadherin mutation carriers and its implications for patient screening. *J Pathol.* 2004 Jun ;203(2) :681-7
81. Schrader KA, et al., Hereditary diffuse gastric cancer :association with lobular breast cancer. *Fam Cancer.* 2008 ;7(1) :73-82
82. Moll R, et al., Differential loss of E-cadherin expression in infiltrating ductal and lobular breast carcinomas. *Am J Pathol.* 1993 Dec ;143(6) :1731-42
83. Fitzgerald RC, et al., Hereditary diffuse gastric cancer : updated consensus guidelines for clinical management and directions for future research. *J Med Genet.* 2010 Jul ;47(7) :436-44
84. Hansford S, et al., Hereditary Diffuse Gastric Cancer Syndrome :CDH1 Mutations and Beyond. *JAMA Oncol.* 2015 Apr ;1(1) :23-32
85. Savitsky K et al., The complete sequence of the coding region of the ATM gene reveals similarity to cell cycle regulators in different species. *Hum Mol Genet.* 1995 Nov ;4(11) :2025-32
86. Swift M, et al., Incidence of cancer in 161 families affected by ataxia-telangiectasia. *N Engl J Med.* 1991 Dec 26 ;325(26) :1831-6
87. Thompson D, et al., Cancer risks and mortality in heterozygous ATM mutation carriers. *J Natl Cancer Inst.* 2005 Jun1 ;97(11) :813-22
88. Van os NJ, et al., Health risks for ataxia-telangiectasia mutated heterozygotes :a systematic review, meta-analysis and evidence-based guideline. *Clin Genet.* 2016 Aug ;90(2) :105-17
89. Choi M, Kipps T, Kurzrock R. ATM Mutations in Cancer :Therapeutic Implications. *Mol Cancer Ther.* 2016 Aug ;15(8) :1781-91
90. Jenne DE, et al., Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. *Nat Genet.* 1998 Jan ;18(1) :38-43

91. Corradetti MN, Guan KL. Upstream of the mammalian target of rapamycin :do all roads pass through mTOR ? *Oncogene*. 2006 Oct 16 ;25(48) :6347-60
92. Beggs AD, et al., Peutz-Jeghers syndrome : a systematic review and recommendations for management. *Gut*. 2010 Jul ;59(7) :975-86
93. Syngal S, et al., ACG clinical guideline : Genetic testing and management of hereditary gastrointestinal cancer syndromes. *Am J Gastroenterol*. 2015 Feb ;110(2) :223-62
94. Miyaki M, et al., Somatic mutations of LKB1 and beta-catenin genes in gastrointestinal polyps from patients with Peutz-Jeghers syndrome. *Cancer Res*. 2000 Nov 15 ;60(22) :6311-3
95. Ji H, et al., LKB1 modulates lung cancer differentiation and metastasis. *Nature*. 2007 Aug 16 ;448(7155) :807-10
96. Wingo SN, et al., Somatic LKB1 mutations promote cervical cancer progression. *PLoS One*. 2009 ;4(4) :e5137
97. de la Chapelle A. The incidence of Lynch syndrome. *Fam Cancer*. 2005 ;4(3) :233-7
98. Haraldsdottir S, et al., Comprehensive population-wide analysis of Lynch syndrome in Iceland reveals founder mutations in MSH6 and PMS2. *Nat Commun*. 2017 May 3 ;8 :14755
99. Bonadona V, et al., Cancer risks associated with germline mutations in MLH1, MSH2, and MSH6 genes in Lynch syndrome. *JAMA*. 2011 Jun 8 ;305(22) :2304-10
100. Win AK, et al., Risk of breast cancer in Lynch syndrome :a systematic review. *Breast Cancer Res*. 2013 Mar 19 ;15(2) :R27
101. Win AK, et al., Colorectal and other cancer risks for carriers and noncarriers from families with a DNA mismatch repair gene mutation : a prospective cohort study. *J Clin Oncol*. 2012 Mar 20 ;30(9) :958-64
102. Bucksch K, et al., Cancer risks in Lynch syndrome, Lynch-like syndrome, and familial colorectal cancer type X :a prospective cohort study. *BMC Cancer*. 2020 May 24 ;20(1) :460
103. Sheehan M, et al., Investigating the Link between Lynch Syndrome and Breast Cancer. *Eur J Breast Health*. 2020 Apr 1 ;16(2) :106-109
104. Bouzourene H, et al., Selection of patients with germline MLH1 mutated Lynch syndrome by determination of MLH1 methylation and BRAF mutation. *Fam Cancer*. 2010 Jun ;9(2) :167-72
105. Rubenstein JH, et al., American Gastroenterological Association Institute Guideline on the Diagnosis and Management of Lynch Syndrome. *Gastroenterology*. 2015 Sep ;149(3) :777-82
106. Ligtenberg MJ, Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. *Nat Genet*. 2009 Jan ;41(1) :112-7
107. Misage N, et al., Sebaceous neoplasms in Muir-Torre syndrome. *Am J Dermatopathol*. 2000 Apr ;22(2) :155-61

108. Hamilton SR, et al., The molecular basis of Turcot's syndrome. *N Engl J Med.* 1995 Mar 30 ;332(13) :839-47
109. Wimmer K, et al., Constitutional mismatch repair-deficiency syndrome :have we so far seen only the tip on an iceberg ? *Hum Genet.* 2008 Sep ;124(2) :105-22
110. Bakry D, et al., Genetic and clinical determinants of constitutional mismatch repair deficiency syndrome : report from the constitutional mismatch repair deficiency consortium. *Eur J Cancer.* 2014 Mar ;50(5) :987-96
111. Shinohara A, et al., Cloning of human, mouse and fission yeast recombination genes homologous to RAD51 and recA. *Nat Genet.* 1993 Jul ;4(3) :239-43
112. Albala JS, et al., Identification of a novel human RAD51 homolog, RAD51B. *Genomics.* 1997 Dec 15 ;46(3) :476-9
113. Dosanjh MK, et al., Isolation and characterization of RAD51C, a new human member of the RAD51 family of related genes. *Nucleic Acids Res.* 1998 Mar 1 ;26(5) :1179-84
114. Pittman DL, et al., Identification, characterization, and genetic mapping of Rad51d, a new mouse and human RAD51/RecA-related gene. *Genomics.* 1998 Apr 1 ;49(1) :103-11
115. Rodrigue A, et al., Interplay between human DNA repair proteins at a unique double-strand break in vivo. *EMBO J.* 2006 Jan 11 ;25(1) :222-31
116. Golmard L, et al., Contribution of germline deleterious variants in the RAD51 paralogs to breast and ovarian cancers. *Eur J Hum Genet.* 2017 Dec ;25(12) :1345-1353
117. Yang X, et al., Ovarian and Breast Cancer Risks Associated With Pathogenic Variants in RAD51C and RAD51D. *J Natl Cancer Inst.* 2020 Dec 14 ;112(12) :1242-1250
118. Watters AK, et al., The Effects of Genetic and Epigenetic Alterations of BARD1 on the Development of Non-Breast and Non-Gynecological Cancers. *Genes (Basel).* 2020 Jul 21 ;11(7) :829
119. Cantor SB, et al., BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell.* 2001 Apr 6;105(1):149-60
120. Varon R, et al., Nibrin, a novel DNA double-strand break repair protein, is mutated in Nijmegen breakage syndrome. *Cell.* 1998 May 1;93(3):467-76
121. Gazzoli I, et al., A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor. *Cancer Res.* 2002 Jul 15 ;62(14) :3925-8
122. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA.* 1977 Dec ;74(12) :5463-7
123. Smith LM, et al., Fluorescence detection in automated DNA sequence analysis. *Nature.* 1986 Jun 12-18 ;321(6071) :674-9
124. Mullis K, et al., Specific enzymatic amplification of DNA in vitro : the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol.* 1986 ;51 Pt 1 :263-73
125. Swerdlow H, et al., Three DNA sequencing methods using capillary gel electrophoresis and laser-induced fluorescence. *Anal Chem.* 1991 Dec 15 ;63(24) :2835-41

126. Chial H., DNA sequencing technologies key to the Human Genome Project. *Nature Education* 2008 1(1) :219
127. Goodwin S, et al., Coming of age : ten years of next-generation sequencing technologies. *Nat Rev Genet.* 2016 May 14 ;17(6) :333-51
128. Slatko BE et al., Overview of Next-Generation Sequencing Technologies. *Curr Protoc Mol Biol.* 2018 Apr ;122(1) :e59
129. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* 1998 Mar ;8(3) :186-94
130. Schouten JP, et al., Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res.* 2002 Jun 15;30(12):e57
131. Kallioniemi A, et al., Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science.* 1992 Oct 30;258(5083):818-21
132. Solinas-Toldo S, et al., Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer.* 1997 Dec;20(4):399-407
133. Pinkel D,. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet.* 1998 Oct;20(2):207-11
134. Richards S, et al., Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015 May;17(5):405-24
135. Moliner AM, Waligora J. The European Union Policy in the Field of Rare Diseases. *Adv Exp Med Biol.* 2017;1031:561-587
136. Kato S, et al., Understanding the function-structure and function-mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. *Proc Natl Acad Sci U S A.* 2003 Jul 8;100(14):8424-9

Serment d'Hippocrate

En présence des maîtres de cette école, de mes condisciples, je promets et je jure d'être fidèle aux lois de l'honneur et de la probité dans l'exercice de la médecine.

Je dispenserai mes soins sans distinction de race, de religion, d'idéologie ou de situation sociale.

Admis à l'intérieur des maisons, mes yeux ne verront pas ce qui s'y passe, ma langue taira les secrets qui me seront confiés et mon état ne servira pas à corrompre les mœurs ni à favoriser les crimes.

Je serai reconnaissant envers mes maîtres, et solidaire moralement de mes confrères. Conscient de mes responsabilités envers les patients, je continuerai à perfectionner mon savoir.

Si je remplis ce serment sans l'enfreindre, qu'il me soit donné de jouir de l'estime des hommes et de mes condisciples, si je le viole et que je me parjure, puissé-je avoir un sort contraire.

Oncogénétique théranostique du cancer du sein : étude des résultats de séquençage NGS de 103 patientes

Les altérations des gènes *BRCA1* et *BRCA2* conférant une sensibilité aux inhibiteurs de PARP-1 (PARPi), les techniques de séquençage haut-débit sont devenues incontournables dans l'évaluation et le traitement des patientes atteintes d'un cancer du sein, définissant l'indication théranostique en oncogénétique. L'objectif de l'étude réalisée était donc de décrire les altérations constitutionnelles délétères retrouvées par approche NGS chez 103 patientes sélectionnées dans cette indication, et de tenter d'identifier parmi les variants de signification inconnue (VSI) détectés, ceux pouvant être potentiellement actionnables. Nous avons mis en évidence 10 variants responsables du syndrome HBOC (performance diagnostique : 9,71%). Parmi eux, 60% touchaient *BRCA1* (n=2) et *BRCA2* (n=5) autorisant l'administration de PARPi chez les patientes concernées ; les autres touchant *PALB2* (n=1), et *RAD51C* (n = 2). Trois autres variants délétères ont été détectés, localisés dans *ATM* (n=2) et *MRE11A* (n=1) ; gènes impliqués dans la recombinaison homologue. Enfin, parmi 54 VSI faux-sens étudiés, il a été possible, en analysant les données issues de la génétique tumorale, d'identifier 3 variants supplémentaires susceptibles d'induire un phénotype HRD (touchant *CHEK2*, *PALB2* et *BRIP1*). Notre étude montre donc que le NGS est une approche efficace pour identifier les patientes pouvant bénéficier d'un traitement par PARPi, que les altérations moléculaires en cause soient ponctuelles ou structurales grâce à la combinaison des techniques d'étude du génome. Avec l'augmentation prévisible du nombre de biomarqueurs disponibles dans cette indication, il est probable que dans un avenir proche, analyses tumorale et constitutionnelle soient réalisées de manière combinée, voire associées à d'autres techniques innovantes définissant une nouvelle approche « multi-omique » de l'oncogénétique.

Mots-clés : NGS, Cancer du sein, Oncogénétique, Théranostique, Recombinaison Homologue, Variant de Signification Inconnue

Germline molecular alterations in breast cancer theranostic multigenes panel analysis : a study of 103 patients

BRCA1 and *BRCA2* genes alterations conferring sensitivity to PARP inhibitors (PARPi), high-throughput sequencing techniques have become essential in the evaluation and treatment of patients with advanced breast cancer, defining theranostic indication in oncogenetics. The objective of the study was thus to describe germline deleterious alterations found by NGS approach in 103 patients selected in this indication, and to attempt to identify among variants of unknown significance (VOUS) detected, those that could potentially be actionable. We identified 10 variants responsible for HBOC syndrome (diagnostic performance : 9.71%). Among them 60% were located in *BRCA1* (n = 2) and in *BRCA2* (n = 5), authorizing the administration of PARPi in the patients concerned; the other altering *PALB2* (n = 1), and *RAD51C* (n=2) genes. Three other deleterious variants were detected, located in *ATM* (n = 2) and *MRE11A* (n = 1), all involved in homologous recombination. Finally, among 54 missense VOUS studied, it was possible, by analyzing data from tumor genetics, to identify 3 additional variants likely to induce an HRD phenotype (affecting *CHEK2*, *PALB2* and *BRIP1*). Our study shows whether that NGS is an effective approach to identify patients who may benefit from treatment with PARPi, whether the molecular alterations involved are punctual or structural thanks to the combination of genome study's techniques. Predictable increasing of biomarkers available in this indication will in a close future, lead to combine germline and somatic approaches, likely associated to innovating techniques, defining a new « multi-omic » approach in oncogenetics.

Keywords : NGS, Breast cancer, Oncogenetics, Theranostic, Homologous Recombination, Variant of Unknow Significance

