

Université de Montréal

**Etude du transcriptome des cellules
non tumorales de l'épithélium de surface de l'ovaire
des femmes porteuses d'une mutation des gènes
BRCA1 et *BRCA2***

par

Diala Abd Rabbo

Département de Biochimie

Faculté de Médecine

Mémoire présenté à la faculté des études supérieures
en vue de l'obtention du grade de M.Sc. en bio-informatique

Avril 2009

© Diala Abd Rabbo, 2009

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé:

Etude du transcriptome des cellules non tumorales
de l'épithélium de surface de l'ovaire des femmes porteuses
d'une mutation des gènes *BRCA1* et *BRCA2*

Présenté par:

Diala Abd Rabbo

a été évalué par un jury composé des personnes suivantes:

Daniel Sinnet Ph.D., président-rapporteur

Christine Maugard M.D. Ph.D., directeur de recherche

Sébastien Lemieux Ph.D., parrain et membre du jury

Résumé

Nous avons étudié le transcriptome de neuf échantillons d'ARN extraits de cultures primaires de cellules non tumorales de l'épithélium de surface de l'ovaire (NOSE) provenant de quatre donneuses non porteuses de mutation, deux mutées sur *BRCA1* et trois sur *BRCA2*, ainsi que de quatre échantillons d'ARN extraits de cultures primaires de cellules tumorales de l'ovaire (TOV) provenant de trois donneuses porteuses de mutation sur *BRCA1* et une sur *BRCA2*. Nous avons identifié, pour la première fois, les signatures moléculaires associées à la présence d'une mutation de *BRCA1* et *BRCA2* dans les cellules NOSEs ainsi que la signature associée à la transformation tumorale des cellules NOSEs en TOVs chez les porteuses de mutation de *BRCA1*. Nous avons également localisé les domaines chromosomiques comportant des gènes corégulés en association avec la présence d'une mutation de *BRCA1* dans les cellules NOSEs. Les allèles sauvage et muté de *BRCA2* étaient exprimés dans les cellules TOVs provenant des porteuses de la mutation 8765delAG sur *BRCA2*. Nous avons observé que le niveau d'expression des transcrits de *BRCA2* était plus élevé dans les cellules provenant des tumeurs ovariennes les plus agressives chez les femmes porteuses de la mutation 8765delAG sur *BRCA2*, les transcrits correspondants à l'allèle muté contribuant avec un pourcentage élevé du niveau d'expression total du gène. Le phénotype tumoral observé chez les Canadiennes Françaises porteuses de cette mutation pourrait résulter d'un effet de dosage de l'allèle muté.

Mots clés : Transcriptome, Domaines chromosomiques de corégulation, Cellules non tumorales de l'épithélium de surface de l'ovaire, *BRCA1*, *BRCA2*, Cellules tumorales ovariennes, *BRCA2* 8765delAG.

Abstract

We analyzed the transcriptome of nine primary cultures of non-tumor ovarian surface epithelium cells (NOSE) from four non-carriers, two *BRCA1* and three *BRCA2* carriers, and four primary cultures of tumor ovarian cells (TOV) from three *BRCA1* and one *BRCA2* carriers. We identified the first molecular signatures associated with the presence of *BRCA1* and *BRCA2* mutations in NOSEs and the first molecular signature associated with the transformation from NOSEs to TOVs in French Canadian women carriers of *BRCA1* mutation. Moreover, we localized some co-regulated chromosomal domains associated with the presence of a *BRCA1* mutation in NOSE cells. Wild-type and mutated *BRCA2* allelic transcripts were expressed in tumor cells from 8765delAG *BRCA2* mutation carriers, with the highest level of *BRCA2* transcript expression and the highest contribution of the mutated allele in cells originating from the most aggressive ovarian tumors. The observed phenotype in *BRCA2*-mutated cells as well as the aggressiveness of the tumor could result from a dosage effect of the *BRCA2* mutated allele.

Keywords: Transcriptome, Co-regulated chromosomal domains, Non-tumor ovarian surface epithelium, *BRCA1*, *BRCA2*, Epithelial ovarian tumor, *BRCA2* 8765delAG.

Table des Matières

| | |
|--|-------------|
| Liste des tableaux | viii |
| Liste des figures | x |
| Liste des sigles et abréviations | xi |
| A. Introduction | 1 |
| A.1. Préambule : situation du sujet | 2 |
| A.2. Appareil reproducteur de la femme | 3 |
| A.3. Histologie des cancers de l’ovaire | 4 |
| A.3.1. Tumeurs épithéliales | 4 |
| A.3.2. Tumeurs germinales | 5 |
| A.3.3. Tumeurs embryonnaires | 5 |
| A.4. Epidémiologie des cancers de l’ovaire au Canada et au Québec | 5 |
| A.5. Facteurs de risque de cancers de l’ovaire | 6 |
| A.5.1. Facteurs associés à la vie reproductive et hormonale | 6 |
| A.5.2. Facteurs associés au mode de vie | 7 |
| A.5.3. Facteurs associés à différents types d’exposition | 7 |
| A.5.4. Facteurs génétiques | 7 |
| A.6. Théories de la carcinogenèse ovarienne | 9 |
| A.6.1. Théorie I - Théorie mécanique de l’ovulation | 9 |
| A.6.2. Théorie II – Théorie du rôle des suppresseurs de tumeurs et oncogènes dans la carcinogenèse ovarienne | 10 |
| A.6.3. Théorie III – Théorie de la carcinogenèse associée aux gènes <i>BRCA1</i> et <i>BRCA2</i> | 10 |
| A.7. Notions d’oncogènes, antioncogènes, suppresseurs de tumeurs et théorie de Knudson | 11 |
| A.7.1. Proto-oncogènes et oncogènes | 11 |
| A.7.2. Anti-oncogènes et gènes suppresseurs de tumeurs | 13 |
| A.7.3. Théorie de Knudson | 14 |
| A.7.4. Tumorigenèse associée à <i>BRCA1</i> et <i>BRCA2</i> et théorie de Knudson | 15 |
| A.7.5. Gènes cerbères, aides-soignants et paysagistes | 18 |
| A.7.5.1. Gènes cerbères ou gardiens du génome | 18 |
| A.7.5.2. Gènes aides-soignants | 18 |
| A.7.5.3. Gènes paysagistes | 20 |
| A.8. <i>BRCA1</i> et <i>BRCA2</i> | 20 |
| A.8.1. Structure et fonctions de <i>BRCA1</i> et <i>BRCA2</i> | 21 |
| A.8.1.1. Structure de <i>BRCA1</i> | 21 |
| A.8.1.2. Domaines fonctionnels de <i>BRCA1</i> | 23 |

| | | |
|--------------|---|-----------|
| A.8.1.2.1. | Domaine à doigt de zinc de type RING _____ | 24 |
| A.8.1.2.2. | Signaux de localisation et d'exportation nucléaire ____ | 25 |
| A.8.1.2.3. | Domaines BRCT _____ | 25 |
| A.8.1.3. | Fonctions de BRCA1 _____ | 26 |
| A.8.1.3.1. | Réparation de l'ADN _____ | 26 |
| A.8.1.3.2. | Régulation transcriptionnelle _____ | 28 |
| A.8.1.3.3. | Contrôle du cycle cellulaire _____ | 28 |
| A.8.1.3.4. | Régulation de l'invasion cellulaire _____ | 29 |
| A.8.1.4. | Structure de BRCA2 _____ | 30 |
| A.8.1.5. | Domaines fonctionnels de BRCA2 _____ | 30 |
| A.8.1.6. | Fonctions de BRCA2 _____ | 31 |
| A.8.1.6.1. | Réparation de l'ADN _____ | 31 |
| A.8.1.6.2. | Régulation transcriptionnelle _____ | 32 |
| A.8.1.6.3. | Contrôle du cycle cellulaire _____ | 33 |
| A.8.2. | Pénétrance des mutations de <i>BRCA1</i> et <i>BRCA2</i> _____ | 34 |
| A.8.3. | Mutations de <i>BRCA1</i> et <i>BRCA2</i> et particularités canadiennes françaises ____ | 34 |
| A.9. | Haplo-insuffisance et gènes <i>BRCA</i> _____ | 39 |
| A.9.1. | <i>BRCA1</i> , <i>BRCA2</i> et cellules tumorales _____ | 40 |
| A.9.2. | Lymphocytes et hétérozygotie aux loci <i>BRCA1</i> et <i>BRCA2</i> _____ | 41 |
| A.10. | <i>BRCA1</i>, <i>BRCA2</i> et cellules NOSEs _____ | 43 |
| A.11. | Du transcriptome à l'ADN _____ | 43 |
| A.12. | Hypothèses _____ | 44 |
| A.13. | Objectifs _____ | 45 |

B. Matériel et Méthodes _____ **46**

| | | |
|-------------|--|-----------|
| B.1. | Cultures primaires et échantillons d'ARN _____ | 48 |
| B.2. | Identification du transcriptome des F-NOSEs et des TOVs _____ | 54 |
| B.2.1. | Biopuce HuGeneFL® _____ | 54 |
| B.2.2. | Traitement des échantillons _____ | 56 |
| B.2.3. | Plateforme Bioconductor _____ | 56 |
| B.2.4. | Procédures de normalisation _____ | 57 |
| B.2.4.1. | Normalisation avec la procédure « <i>Robust Multi-array Average</i> » __ | 57 |
| B.2.4.2. | Normalisation avec la procédure « <i>Guanine Cytosine Robust Multi-array Average</i> » _____ | 60 |
| B.2.5. | Analyse supervisée avec la méthode « <i>Linear Model for Microarrays Data</i> » __ | 62 |
| B.2.5.1. | Analyse différentielle _____ | 62 |
| B.2.5.2. | Correction des valeurs de p avec la procédure « <i>False Discovery Rate</i> » de Benjamini et Hochberg _____ | 66 |
| B.3. | Méthodes de Validation _____ | 67 |
| B.3.1. | Agrégation hiérarchique ascendante _____ | 67 |
| B.3.2. | Quantification de quatre gènes candidats par PCR quantitative en temps réel __ | 68 |
| B.3.3. | Quantification de <i>BRCA1</i> et <i>BRCA2</i> par PCR quantitative en temps réel ____ | 70 |
| B.3.4. | Quantification spécifique d'allèles _____ | 70 |

| | | |
|-------------|---|------------|
| B.4. | Localisation des régions comportant des gènes coréglés chez les porteuses d'une mutation de <i>BRCA1</i> ou de <i>BRCA2</i> | 71 |
| <hr/> | | |
| C. | Résultats | 73 |
| <hr/> | | |
| C.1. | Normalisation des données | 74 |
| C.2. | Profil d'expression différentielle associé à une mutation sur <i>BRCA1</i> ou <i>BRCA2</i> dans les cellules NOSEs et les TOVs | 75 |
| C.3. | Agrégation hiérarchique | 75 |
| C.4. | Quantification par PCR quantitative en temps réel des quatre gènes candidats et de <i>BRCA1</i> et <i>BRCA2</i> dans la série initiale | 77 |
| C.5. | Quantification par PCR quantitative en temps réel des quatre gènes candidats et de <i>BRCA1</i> et <i>BRCA2</i> dans la série étendue | 81 |
| C.6. | Quantification spécifique des allèles G6085T et 8765delAG de <i>BRCA2</i> | 83 |
| C.7. | Localisation des régions chromosomiques coréglées colocalisées chez les porteurs de mutation de <i>BRCA1</i> | 86 |
| <hr/> | | |
| D. | Discussion et conclusions | 92 |
| <hr/> | | |
| D.1. | Originalité et limites de notre étude | 93 |
| D.2. | Avantages et limites des méthodes utilisées | 94 |
| D.3. | Discussion des résultats | 101 |
| D.3.1. | Existe-t'il des signatures moléculaires associées à une mutation sur <i>BRCA1</i> ou <i>BRCA2</i> dans les cellules NOSEs et TOVs des porteuses ? | 101 |
| D.3.2. | L'hétérozygotie pour une mutation sur <i>BRCA2</i> dans les cellules NOSEs et TOVs est-elle suffisante à l'initiation de la carcinogenèse ? | 103 |
| D.3.3. | Existe-t'il des domaines de co-expression associés à une mutation du gène <i>BRCA1</i> dans les cellules NOSEs ? | 105 |
| D.4. | Conclusions et perspectives | 107 |
| <hr/> | | |
| | Bibliographie | 110 |
| <hr/> | | |

Liste des tableaux

| | | |
|-----------------------|--|----|
| Tableau I : | Gènes impliqués dans la prédisposition héréditaire aux cancers de l’ovaire. | 8 |
| Tableau II : | Caractéristiques de <i>BRCA1</i> et <i>BRCA2</i> . | 22 |
| Tableau III : | Risque relatif de survenue de cancers autres que les cancers du sein et de l’ovaire chez les porteurs de mutation sur <i>BRCA1</i> et <i>BRCA2</i> et apparentés. | 34 |
| Tableau IV : | Caractéristiques des donneuses et des cellules NOSEs correspondantes. | 51 |
| Tableau V : | Caractéristiques des donneuses et des TOVs correspondants | 53 |
| Tableau VI : | Résumé des résultats de l’analyse différentielle avec LIMMA pour chacun des quatre gènes candidats sélectionnés. | 79 |
| Tableau VII : | Régions chromosomiques enrichies (RE) en gènes différentiellement exprimés identifiées dans les cellules NOSEs des porteuses d’une mutation sur <i>BRCA1</i> tel que défini avec la fonction <code>geneSetTest</code> ($p < 0,01$). | 86 |
| Tableau VIII : | Régions coréglées colocalisées (RCC) putatives au sein des régions enrichies en ESDEs identifiées dans les cellules NOSEs des porteuses d’une mutation sur <i>BRCA1</i> . | 87 |
| Tableau IX : | Profil d’expression de la région enrichie en ESDEs identifiée sur le chromosome 19 dans les cellules NOSEs des porteuses d’une mutation sur <i>BRCA1</i> tel que défini avec la fonction <code>geneSetTest</code> ($p = 0.0012$) et de la RCC.19 correspondante. | 88 |
| Tableau X : | Profil d’expression de la région enrichie en ESDEs identifiée sur le chromosome 3 dans les cellules NOSEs des porteuses d’une mutation sur <i>BRCA1</i> tel que défini avec la fonction <code>geneSetTest</code> ($p = 0.0042$) et de la RCC.3 correspondante. | 89 |
| Tableau XI : | Profil d’expression de la région enrichie en ESDEs identifiée sur le chromosome 16 dans les cellules NOSEs des porteuses d’une mutation sur <i>BRCA1</i> tel que défini avec la fonction <code>geneSetTest</code> ($p = 0.0070$) et de la RCC.16 correspondante. | 89 |
| Tableau XII : | Profil d’expression de la région enrichie en ESDEs identifiée sur le chromosome 11 dans les cellules NOSEs des porteuses d’une mutation sur <i>BRCA1</i> tel que défini avec la fonction <code>geneSetTest</code> ($p = 0.0073$) et de la RCC.11 correspondante. | 90 |
| Tableau XIII : | Profil d’expression de la région enrichie en ESDEs identifiée sur le chromosome 8 dans les cellules NOSEs des porteuses d’une mutation sur <i>BRCA1</i> tel que défini avec la fonction <code>geneSetTest</code> ($p = 0.0093$) et de la RCC.8 correspondante. | 90 |

Tableau XIV : Profil d'expression de la région enrichie en ESDEs identifiée sur le chromosome 2 dans les cellules NOSEs des porteuses d'une mutation sur *BRCA1* tel que défini avec la fonction `geneSetTest` ($p = 0.0094$) et de la RCC.2 correspondante. _____ 91

Liste des figures

| | | |
|------------------|---|----|
| Fig. 1 : | Appareil reproducteur féminin. _____ | 4 |
| Fig. 2 : | Développement du cancer de l'ovaire chez les porteuses de mutations sur <i>BRCA1</i> ou <i>BRCA2</i> . _____ | 10 |
| Fig. 3 : | Hypothèse des deux étapes de Knudson. _____ | 15 |
| Fig. 4 : | Carcinogenèse multi-étape. _____ | 19 |
| Fig. 5 : | Variants d'épissage de <i>BRCA1</i> qui codent pour une protéine fonctionnelle. ____ | 23 |
| Fig. 6 : | Structure primaire de la protéine <i>BRCA1</i> . _____ | 24 |
| Fig. 7 : | Structure primaire de la protéine <i>BRCA2</i> . _____ | 31 |
| Fig. 8 : | Distribution des mutations de <i>BRCA1</i> et <i>BRCA2</i> dans les familles prédisposées aux cancers du sein et de l'ovaire. _____ | 38 |
| Fig. 9 : | Distribution des mutations fondatrices de <i>BRCA1</i> et <i>BRCA2</i> dans les familles prédisposées aux cancers du sein et/ou de l'ovaire d'origine Canadienne Française. _____ | 38 |
| Fig. 10 : | Représentation de la biopuce HuGeneFL [®] de la plateforme Affymetrix [™] . ____ | 55 |
| Fig. 11 : | Modélisation de l'intensité observée pour les sondes PM (IO_{PM}) en fonction du signal (S) et du bruit de fond (BF). _____ | 58 |
| Fig. 12 : | Effet d'une abase en fonction de sa position dans la séquence de la sonde. ____ | 61 |
| Fig. 13 : | Agrégation hiérarchique ascendante _____ | 68 |
| Fig. 14 : | Boîtes à moustache représentant les niveaux d'expression des gènes mesurés par la HuGeneFL [®] dans chacun des échantillons d'ARN avant (graphique de gauche), après normalisation par RMA (graphique du milieu) et après normalisation par GCRMA (graphique de droite). _____ | 74 |
| Fig. 15 : | Résultats de LIMMA et des agrégations hiérarchiques correspondantes sous forme, respectivement, de « Volcano Plot » (graphiques de gauche) et de « Heat Map » (graphiques de droite). _____ | 77 |
| Fig. 16 : | Comparaison des niveaux d'expression des quatre gènes candidats mesurés par q-RT-PCR et par HuGeneFL [®] dans la série initiale. _____ | 80 |
| Fig. 17 : | Niveaux d'expression de <i>BRCA1</i> et <i>BRCA2</i> mesurés par q-RT-PCR dans la série initiale. _____ | 81 |
| Fig. 18 : | Niveaux d'expression des quatre gènes candidats et de <i>BRCA1</i> et <i>BRCA2</i> mesurés par q-RT-PCR dans la série étendue. _____ | 82 |
| Fig. 19 : | Quantification spécifique par q-RT-PCR des allèles G6085T et 8765delAG de <i>BRCA2</i> . _____ | 85 |

Liste des sigles et abréviations

| | |
|------------------|--|
| AF | Anémie de Fanconi |
| CDB | Cassure Double-Brin de l'ADN |
| ES | Ensemble de Sondes |
| ESDE | Ensemble de Sondes Différentiellement Exprimées |
| F | Familial (mutation identifiée) |
| HF | Histoire Familiale de cancers |
| FRSQ | Fonds de la Recherche en Santé du Québec |
| GST | Gène Suppresseur de Tumeurs |
| HNS | Hybridation Non Spécifique |
| HNPC | Pour « <i>Hereditary Non Polyposis Colorectal Cancer</i> » ou cancer colorectal héréditaire non polyposique |
| IOSE | Pour « <i>Immortalised Ovarian Surface Epithelium cells</i> » ou cellules immortalisées de l'épithélium de surface de l'ovaire |
| LSP | Lymphocytes du Sang Périphérique |
| M1 | Porteuses de Mutation sur <i>BRCA1</i> |
| M2 | Porteuses de Mutation sur <i>BRCA2</i> |
| MM | Milieu Modifié |
| MMR | Pour « <i>MisMatch Repair</i> » ou réparation des mésappariements |
| NF | Non Familial |
| NM | Non Muté |
| NOSE | Pour « <i>Non-tumoral Ovarian Surface Epithelium</i> » ou non tumorale de l'épithélium de surface de l'ovaire |
| OSE | Pour « <i>Ovarian Surface Epithelium</i> » ou épithélium de surface de l'ovaire |
| PDH | Perte D'Hétérozygotie |
| PJ | Polypose Juvénile |
| PMK | Procédure Modifiée de Kruk |
| q-RT-PCR | Pour « <i>quantative Real Time Polymerase Chain Reaction</i> » ou PCR quantitative en temps réel |
| RCC | Région Coréglées Colocalisées |
| RCG | Région de Contrôle Global |
| RE | Région Enrichie |
| RH | Recombinaison Homologue |
| RR-Cancer | Réseau de Recherche sur le Cancer |
| SNP | Pour « <i>Single Nucleotide Polymorphism</i> » |
| SVF | Sérum de Veau Fœtal |
| TM1 | Tumoral et Muté sur <i>BRCA1</i> |
| TM2 | Tumoral et Muté sur <i>BRCA2</i> |
| TOV | Tumeur de l'OVaire |

VNC Variation du Nombre de Copies de l'ADN
VRS Virus du Sarcome de *Rous*

*A ma chère mère, rayon de soleil de
ma vie sans elle mon existence et ma
réussite n'auraient pas été possibles.
A la mémoire de mes grands-parents*

Remerciements

Je tiens tout d'abord à exprimer ma gratitude et ma reconnaissance envers Dieu le Tout-Puissant pour le courage et la force qu'il m'a donnée pour finir ce travail.

Mes sincères remerciements s'adresseront à mon superviseur, Dr. Christine Maugard, pour m'avoir accepté dans son équipe, me donnant la chance de collaborer dans un projet aussi intéressant et motivant, pour sa présence et ses conseils et pour avoir dirigé ce travail avec beaucoup de rigueur, de compétence professionnelle et de patience. Merci également à Dr. Sébastien Lemieux pour ses conseils judicieux. Merci aux membres du jury, en l'occurrence les Drs Daniel Sinnett et Sébastien Lemieux d'avoir accepté d'évaluer ce travail. Merci à Christine Abaji et Caroline Arous. Sans elles, la partie moléculaire du projet n'aurait pu être réalisée.

Je tiens aussi à remercier l'Institut Nazareth et Louis-Braille (INBL) pour sa générosité et l'aide précieuse qu'il m'a offert, me donnant la chance de surmonter un problème de déficience visuelle et d'accomplir mon projet de recherche dans des conditions agréables. Merci pour tout les membres de l'équipe de l'INBL particulièrement pour Madame Bouttet, Madame Malette et Monsieur Quirion.

Finalement, à tous les membres de ma famille, spécialement à ma mère, et à tous mes amis en particulier à Seraya, Dima, Youngbee, Fatti et Malika, vous étiez toujours à mes côtés. Veuillez trouver dans ce travail mes vifs sentiments d'amour et d'affection. Que Dieu vous garde, vous protège et vous offre une vie pleine de bonheur et de succès.

A. Introduction

A.1. Préambule : situation du sujet

BRCA1 et *BRCA2* sont des gènes majeurs de prédisposition aux cancers du sein et de l’ovaire [1-3]. Leur pénétrance incomplète et leur expressivité variable compliquent la prédiction du risque de cancers chez les porteuses d’une mutation de ces gènes [4, 5]. En effet, en l’absence de marqueurs de détection précoce fiables, les chirurgies prophylactiques demeurent l’unique alternative efficace pour la prévention de ces cancers, avec toutes leurs conséquences en terme de qualité de vie et de morbidité pour les patientes qui vont opter pour ce type d’intervention prophylactique. L’annexectomie prophylactique réduit le risque de cancers du sein de 51% (95% IC 35%-63%) et celui de cancers de l’ovaire et des trompes de Fallope de 79% (95% IC 61%–88%) chez les porteuses d’une mutation de *BRCA1* ou *BRCA2* [6].

Pour améliorer la prédiction du risque chez les porteuses de mutation et établir de nouvelles méthodes de diagnostic et de prévention des cancers de l’ovaire, nous voulions identifier les événements associés à l’hétérozygotie dans les cellules qui, du fait de la présence d’une mutation, présentent un risque élevé de transformation néoplasique, telles que les cellules non tumorales de l’épithélium de surface de l’ovaire (NOSE).

Ce mémoire présente l’étude du transcriptome des cellules NOSEs provenant de femmes porteuses d’une mutation germinale des gènes *BRCA1* et *BRCA2*. Nos résultats présentés ici sont uniques parce que, d’une part nous avons utilisé des biopuces afin d’effectuer une étude à l’échelle du génome au lieu de nous limiter à l’étude d’un nombre restreint de gènes comme dans certaines des études précédentes, d’autre part et surtout nous avons utilisé un matériel rare, difficile d’accès, les cellules NOSEs, cellules non immortalisées, donc à un stade plus précoce de la carcinogenèse que les cellules IOSEs (cellules immortalisées de l’épithélium de surface de l’ovaire) utilisées par d’autres équipes [7].

Nous avons effectué une analyse différentielle de données d’expression issues des biopuces HuGeneFL[®]. Nous avons identifié, pour la première fois, les signatures moléculaires associées à une mutation de *BRCA1* et *BRCA2* dans les cellules NOSEs, ainsi que le profil d’expression différentielle correspondant à la transformation tumorale associée à une mutation du gène *BRCA1*. Nous présenterons ici les méthodes d’identification des profils d’expression différentielle associés à une mutation sur *BRCA1* et *BRCA2* dans les cellules NOSEs et TOVs des porteuses de mutation, les signatures moléculaires que nous avons

identifiées ainsi que la validation sur le plan informatique et biologique des résultats d'analyse du transcriptome de ces cellules.

Nous avons également quantifié les transcrits alléliques de *BRCA2* dans les cellules NOSEs afin de vérifier si une perte d'hétérozygotie (PDH) au locus *BRCA2* était nécessaire pour l'initiation de la carcinogenèse chez les patientes porteuses d'une mutation de *BRCA2*. Nous avons ainsi mis en évidence la rétention et l'expression de l'allèle sauvage de *BRCA2* non seulement dans les cellules NOSEs mais aussi dans les adénocarcinomes ovariens des porteuses de mutation de ce gène. Contrairement à ce qui était communément admis jusqu'alors, nous avons montré qu'une PDH n'est pas nécessaire pour le développement de cancers de l'ovaire chez les hétérozygotes pour une mutation sur *BRCA2*. Nous suggérons donc que l'hétérozygotie de *BRCA2* pourrait suffire au développement de cancers de l'ovaire chez les porteuses de mutation et qu'un effet de dosage de l'allèle muté de *BRCA2* interviendrait pour déterminer le phénotype clinique chez ces patientes.

Notre étude ne se limite pas à l'identification des profils d'expression différentielle associés à une mutation des gènes *BRCA*, mais s'intéresse également à la localisation de régions chromosomiques corégulées associées à une mutation de *BRCA1* dans les cellules NOSEs. L'identification de ces régions pourrait nous orienter vers des phénomènes d'épistasie entre gènes contigus ou dans la même voie métabolique. Elle permettrait aussi d'établir de nouvelles méthodes de diagnostic et de prévention de cancers de l'ovaire chez les porteuses de mutation. Nous présenterons et discuterons dans cet exposé des méthodes que nous avons utilisées, des régions que nous avons identifiées et de l'implication putative de ces régions dans la carcinogenèse ovarienne, ainsi que des perspectives ouvertes par ces études.

A.2. Appareil reproducteur de la femme

Les ovaires font partie de l'appareil reproducteur de la femme. Ce sont les glandes génitales féminines qui ont deux fonctions principales : la sécrétion des hormones stéroïdes femelles et la production des ovules. Les deux ovaires sont situés à l'extrémité de chaque trompe de Fallope de part et d'autre de l'utérus (Fig. 1) [8].

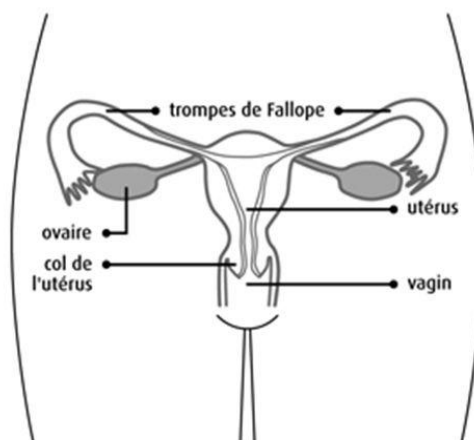


Fig. 1 : Appareil reproducteur féminin.

Figure adaptée d'après <http://www.cancer.ca/>

L'ovaire a une forme ovoïde et allongée de 2 cm de large et 3,5 à 4 cm de long en dehors de la période d'ovulation [9]. Une coupe de l'ovaire permet de distinguer trois zones distinctes de l'extérieur vers l'intérieur : l'épithélium cœlomique qui entoure l'ovaire, le cortex ovarien qui représente la partie périphérique de cet organe où se retrouvent les follicules ovariens et finalement la médulla qui forme la partie centrale de l'ovaire [10].

A.3. Histologie des cancers de l'ovaire

Les tumeurs de l'ovaire peuvent se développer à partir des trois constituants principaux des gonades : l'épithélium cœlomique produisant les tumeurs épithéliales, les cellules germinales produisant les tumeurs germinales et enfin le stroma spécialisé produisant les tumeurs embryonnaires [11]. Nous nous intéresserons ici aux tumeurs épithéliales, de loin les plus fréquentes.

A.3.1. Tumeurs épithéliales

Les tumeurs épithéliales dérivent de l'épithélium cœlomique et représentent 90 % des cancers de l'ovaire [12]. Elles comprennent : les tumeurs séreuses dérivant de l'épithélium tubaire, les tumeurs mucineuses dérivant de l'épithélium endo-cervical, les tumeurs endométrioïdes dérivant de l'endomètre, les tumeurs à cellules claires dérivant de kystes endométriosiques, les tumeurs de Brenner dérivant du mésovarium cœlomique et d'autres

tumeurs plus rares. Les tumeurs séreuses sont les plus fréquentes. Elles représentent 40 à 50 % de l'ensemble des tumeurs épithéliales de l'ovaire [11].

A.3.2. Tumeurs germinales

Les tumeurs germinales dérivent des cellules du même nom. Elles sont surtout fréquentes chez les femmes jeunes. Elles comprennent les tératomes matures kystiques ou kystes dermoïdes de l'ovaire [9].

A.3.3. Tumeurs embryonnaires

Les tumeurs embryonnaires dérivent du mésenchyme. Elles surviennent généralement après la ménopause et comprennent : les tumeurs à cellules thécales ou thécomes, les tumeurs à cellules fusiformes ou fibromes et les tumeurs à cellules de Leydig ou Leydigome [9].

A.4. Epidémiologie des cancers de l'ovaire au Canada et au Québec

Au Canada, une femme sur 72 risque d'être atteinte d'un cancer de l'ovaire. Ce dernier est la 7^{ème} cause de cancer chez la femme en terme d'incidence, mais la 5^{ème} en terme de décès et la 1^{ère} cause de décès par cancer gynécologique au Canada [13]. En 2008, le nombre de nouveaux cas a été estimé à 2 500 femmes au Canada et 640 au Québec [13]. Au Canada, l'espérance de vie d'une femme atteinte d'un cancer de l'ovaire à cinq ans est de 40 % [13]. En 2008, le nombre de décès par cancers de l'ovaire a été estimé à 1 700 femmes au Canada et 370 au Québec [13], ce qui est très élevé comparativement à l'incidence de ce cancer et témoigne de notre inefficacité sur le plan thérapeutique, faisant du cancer de l'ovaire un tueur silencieux.

Le pronostic de cette maladie est effroyable du fait de l'absence de symptômes aux stades précoces, puis plus tard de l'absence de symptômes typiques. De plus les méthodes d'investigations cliniques, radiologiques (échographie) ou biologiques (dosage du marqueur tumoral CA125), sont peu sensibles et spécifiques aux stades précoces de cette maladie. Seuls 10 % des cas de cancers de l'ovaire sont diagnostiqués à ces stades (stade I ou stade II) avec un taux de survie estimé à cinq ans de 80 à 90 % [14]. Soixante-quinze à 85 % de ces cancers sont diagnostiqués à un stade avancé de la carcinogenèse (stade III ou stade IV) [14]. C'est pourquoi, malgré les avancées thérapeutiques récentes, le taux de survie à cinq

ans de ces cancers reste faible : 15 à 20 % [15]. Il est donc indispensable de développer des recherches permettant de mieux comprendre la carcinogenèse ovarienne particulièrement aux stades précoces et d'identifier de nouveaux marqueurs tumoraux, ce qui permettra d'améliorer les méthodes de diagnostic précoce et dans un second temps le pronostic de ces cancers.

A.5. Facteurs de risque de cancers de l'ovaire

A.5.1. Facteurs associés à la vie reproductive et hormonale

L'hystérectomie, la ligature des trompes de Fallope ainsi que les facteurs réduisant le nombre d'ovulations comme la grossesse, l'allaitement et l'utilisation de contraceptifs oraux diminuent le risque de développer un cancer de l'ovaire [16, 17]. Au contraire, les facteurs qui augmentent le nombre d'ovulations comme une puberté précoce [18], une ménopause tardive [18], la nulliparité [19], l'utilisation pendant de longues périodes de médicaments pour le traitement de l'infertilité féminine [20] et certains traitements hormonaux substitutifs de la ménopause [21] augmentent le risque de développer un cancer de l'ovaire.

Ces facteurs de risque ont fait l'objet de plusieurs études afin de déterminer leur effet sur les cancers du sein et de l'ovaire chez les femmes porteuses d'une mutation de *BRCA1* ou *BRCA2* [22-25]. La parité et la nulliparité n'ont pas d'effet sur le risque de survenue d'un cancer du sein chez les porteuses d'une mutation des gènes *BRCA* [25]. L'utilisation de contraceptifs, une grossesse avant l'âge de 40 ans augmentent ce risque [24], alors qu'une grossesse après l'âge de 40 ans le diminue chez les hétérozygotes [25]. L'allaitement pourrait diminuer le risque d'incidence de cancers du sein chez les porteuses de mutation indépendamment de l'âge [24].

Les facteurs réduisant le nombre d'ovulations diminuent le risque de développer un cancer de l'ovaire chez les porteuses d'une mutation de *BRCA1* [22, 23]. L'utilisation de contraceptifs diminue ce risque alors que la parité l'augmente chez les hétérozygotes pour une mutation de *BRCA2* [22, 23]. Les résultats sur l'efficacité de la ligature des trompes de Fallope sur la réduction du risque de survenue de cancer de l'ovaire chez les femmes hétérozygotes sont contradictoires [24]. Enfin, l'effet de la grossesse et l'allaitement sur ce

risque chez les porteuses d'une mutation de *BRCA2* reste inconnu en raison du nombre limité de femmes atteintes de cancers de l'ovaire hétérozygotes pour *BRCA2* [22].

En conclusion, les études doivent se poursuivre pour mieux élucider l'épidémiologie des cancers de l'ovaire en particulier chez les porteuses de mutation des gènes *BRCA1* et *BRCA2*.

A.5.2. Facteurs associés au mode de vie

L'obésité, la nutrition et l'activité physique sont trois facteurs fortement liés. Cependant leur effet sur le risque d'incidence du cancer de l'ovaire n'est pas bien déterminé. Les facteurs alimentaires associés au risque de développer ce cancer restent encore mal définis. Selon une revue de la littérature menée par Bandera et al. [26], les études cas-contrôles ont montré une relation entre l'augmentation du risque de survenue de cancer l'ovaire et l'obésité. Le rôle de l'activité physique reste encore peu clair. Les données épidémiologiques actuelles sur les facteurs diététiques sont généralement encore insuffisantes pour permettre de générer des lignes directrices en ce qui concerne la prévention du cancer de l'ovaire [26]. Selon le rapport de la Société Américaine du Cancer en 2002, la consommation de légumes et de fruits pourrait réduire ce risque [27]. Un régime alimentaire typiquement occidental riche en viande et pauvre en légumes pourrait être associé à une augmentation d'incidence des cancers de l'ovaire. Cependant, l'association entre les facteurs alimentaires et le risque de cancers épithéliaux de l'ovaire demeure encore peu claire et nécessite des études plus approfondies.

A.5.3. Facteurs associés à différents types d'exposition

L'utilisation du talc dans l'aire génitale [28], l'exposition professionnelle à la triazine (un herbicide) [29], aux radiations ionisantes [30] et aux produits chimiques [31] constituent autant de facteurs de risques de cancers de l'ovaire.

A.5.4. Facteurs génétiques

La grande majorité des cas de cancers de l'ovaire seraient sporadiques, c'est-à-dire qu'il s'agirait de cancers résultant de l'accumulation de mutations acquises au cours de la vie. Cependant, l'existence d'une histoire familiale de cancers de l'ovaire représente le facteur de risque prédominant pour une femme indemne de cancers [32]. Les cancers héréditaires de l'ovaire sont observés dans des familles présentant des syndromes prédisposant au

développement de cancers. Nous citerons le syndrome sein-ovaire et le syndrome-HNPCC « *Hereditary Non Polyposis Colorectal Cancer* ». Le syndrome sein-ovaire est défini par la présence de cancers du sein et d'au moins un cancer de l'ovaire dans une même branche d'hérédité chez des sujets apparentés au premier et/ou second degré entre eux [33]. Le syndrome HNPCC associe des cancers digestifs, des cancers gynécologiques (endomètre et plus rarement ovaires) ainsi que parfois des cancers des voies urinaires, chez des sujets apparentés au premier et/ou second degré [33]. Nous citerons également le syndrome de l'ovaire spécifique de site ou syndrome de l'ovaire seul défini par la présence d'au moins deux cas de cancers de l'ovaire chez des sujets apparentés au premier degré [33].

Les syndromes sein-ovaire et ovaire spécifique de site sont associés à des mutations germinales sur les gènes *BRCA1* et *BRCA2*. Ils représentent 90 % des cancers héréditaires de l'ovaire [34]. Le syndrome HNPCC, lui, est associé à une mutation sur un des gènes « *Mismatch Repair* » (MMR) : *MLH1*, *MSH2*, *MSH6*, *PMS2* [35]. Il représente la plus grande partie des 10 % restants [34]. Les gènes MMR codent pour les protéines du système de réparation des mésappariements pendant la réplication de l'ADN. D'autres syndromes prédisposant aux cancers héréditaires de l'ovaire existent mais ils sont plus rares et sont associés à des mutations sur d'autres gènes (Tableau I).

Tableau I : Gènes impliqués dans la prédisposition héréditaire aux cancers de l'ovaire.

| Gène(s) | Syndrome familial | Pourcentage estimé des cancers héréditaires de l'ovaire associés à une mutation de ces gènes | Risque cumulé de cancers de l'ovaire à 70 ans chez les porteurs de mutation |
|-------------------------------|-------------------|--|---|
| <i>BRCA1</i> | Sein-Ovaire | 75 % | 20 - 50% |
| <i>BRCA2</i> | Sein-Ovaire | 10 % | 10 - 30% |
| <i>MLH1, MSH2, MSH6, PMS2</i> | HNPCC | 5 - 10 % | < 10 % |
| <i>STK11</i> | Peutz-Jeghers | < 1 % | < 2 % |
| <i>PTEN</i> | Cowden | < 1 % | < 2 % |
| <i>PTC</i> | Gorlin | < 1 % | < 1 % |
| <i>EXT1, EXT2, EXT3</i> | Ollier | < 1 % | < 1 % |

Tableau tiré de [36].

A.6. Théories de la carcinogenèse ovarienne

Plusieurs théories ont été avancées quant à la carcinogenèse ovarienne. Nous citerons les théories les plus fréquemment admises.

A.6.1. Théorie I - Théorie mécanique de l'ovulation

La théorie la plus communément avancée actuellement est la théorie mécanique de l'ovulation [37]. Cette théorie est basée sur le fait que pendant l'ovulation, l'ovule se libère en produisant une brèche dans l'épithélium de surface de l'ovaire, épithélium que nous appellerons OSE « *Ovarian Surface Epithelium* » dans la suite de notre exposé. Pour réparer l'épithélium de surface, et sous l'effet de facteurs de croissance, les cellules OSEs se multiplient en effectuant plusieurs cycles de mitoses. Ce processus continue jusqu'à ce que la brèche soit réparée et la continuité de l'épithélium de surface soit restaurée.

Dans le cadre d'une étude qui s'intéressait à élucider la pathogenèse des inclusions kystiques de l'ovaire [38], plusieurs milliers de sections provenant de 1 000 ovaires ont été examinées. Ce matériel a été obtenu lors d'ovariectomies chez des femmes ayant présenté une grossesse extra-utérine, une pathologie non cancéreuse de l'ovaire (kystes ovariens) ou de problèmes d'endométriose. Parfois, l'ovariectomie a pu accompagner une hystérectomie chez des femmes présentant une pathologie non cancéreuse de l'utérus. Un petit nombre d'entre elles ont été opérées à la suite du diagnostic de cancers du sein. Cette étude a montré que pendant l'ovulation, les cellules OSEs s'enfouissent dans le cortex de l'ovaire. Elles se mettent ensuite à proliférer produisant des inclusions kystiques. Rien ne prouve que le mécanisme de remodelage de l'épithélium soit le seul facteur intervenant dans son enfouissement en profondeur. Par contre, il est certain que l'ovulation y joue un rôle [39]. Normalement, l'épithélium de surface de l'ovaire est isolé du stroma par une zone fibreuse appelée la tunique albuginée de l'ovaire. A la suite du processus d'internalisation, l'épithélium va entrer en contact avec les hormones et les facteurs de croissance produits par le stroma tels que les œstrogènes, l'acide lysophosphatidique et les gonadotrophines [40], ce qui pourrait favoriser la transformation maligne des cellules OSE.

A.6.2. Théorie II – Théorie du rôle des supresseurs de tumeurs et oncogènes dans la carcinogénèse ovarienne

Plusieurs études ont suggéré que l'apparition d'une tumeur ovarienne pourrait résulter d'une cascade d'altérations touchant des gènes qui jouent un rôle primordial dans la régulation du cycle cellulaire [41, 42]. Elles vont causer l'activation d'oncogènes ou l'inhibition de supresseurs de tumeurs et ainsi favoriser l'initiation de la cancérogénèse.

A.6.3. Théorie III – Théorie de la carcinogénèse associée aux gènes *BRCA1* et *BRCA2*

Afin d'élucider le scénario suivant lequel les cancers héréditaires de l'ovaire apparaissent, le groupe de Sowter et al. a suggéré une troisième théorie qui intègre les deux théories précédentes [43]. L'épithélium de surface de l'ovaire s'internalisant dans le cortex va former des inclusions kystiques. Les cellules non tumorales de l'épithélium de surface de l'ovaire (NOSEs) seraient alors exposées aux hormones et facteurs de croissance produits par les cellules du stroma ovarien, exposition qui pourrait déterminer l'apparition de lésions au niveau de l'ADN. Les femmes porteuses d'une mutation sur *BRCA1* ou *BRCA2* n'étant plus capables de réparer ces lésions, celles-ci vont s'accumuler lors de la multiplication des cellules présentes dans ces inclusions. L'accumulation de lésions de l'ADN favoriserait dès lors l'initiation de la carcinogénèse (Fig. 2) [43].

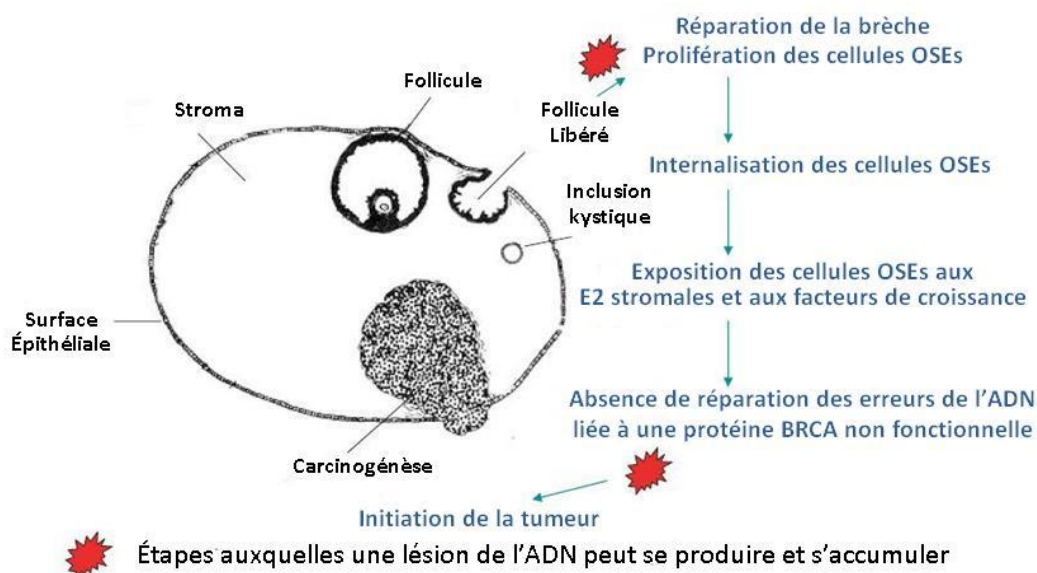


Fig. 2 : Développement du cancer de l'ovaire chez les porteuses de mutations sur *BRCA1* ou *BRCA2*.

Figure adaptée d'après [43].

A.7. Notions d'oncogènes, antioncogènes, suppresseurs de tumeurs et théorie de Knudson

La découverte des oncogènes et des gènes suppresseurs de tumeurs (GST) a permis une avancée dans la connaissance des phénomènes de la carcinogenèse. La prolifération cellulaire au sein des tissus est rigoureusement contrôlée au cours de notre vie. Certaines cellules telles que les neurones ne nécessitent pas un renouvellement constant, d'autres telles que les cellules sanguines ou cutanées sont perpétuellement en multiplication. Le contrôle de cette multiplication cellulaire se fait grâce à un équilibre permanent entre facteurs activateurs et facteurs inhibiteurs de la division cellulaire, équilibre appelé homéostasie cellulaire. Toute altération de cet équilibre peut entraîner une prolifération cellulaire incontrôlée pouvant favoriser l'apparition d'un cancer.

A.7.1. Proto-oncogènes et oncogènes

En 1975, Stehelin et al. ont identifié le premier oncogène *src* dans le rétrovirus VRS « Virus du Sarcome de *Rous* ». Ils ont montré que cet oncogène était capable de transformer les fibroblastes embryonnaires aviaires en culture et d'induire des sarcomes chez le poulet [44]. La même année, ils ont retrouvé dans les cellules normales un gène semblable qu'ils ont appelé *c-src* « *cellular src* », non seulement chez le poulet, mais également chez tous les vertébrés y compris l'homme [45]. Les gènes homologues des oncogènes viraux dans la cellule normale ont alors été appelés proto-oncogènes. La découverte majeure dans le domaine de l'oncogenèse a été la mise en évidence de la capacité du proto-oncogène humain *Ha-ras* « *Harvey-ras* », de transformer les fibroblastes murins NIH-3T3, fibroblastes qui avaient cependant la particularité d'être déjà immortalisés [46].

Les proto-oncogènes ont alors été définis comme des gènes dont la surexpression ou des modifications de structure les transforme en oncogènes, engendrant une prolifération incontrôlée des cellules et contribuant à leur transformation en cellules tumorales [47]. Ainsi, tout gène exerçant une pression de sélection positive sur la prolifération et/ou bloquant les mécanismes apoptiques de la cellule rendant la cellule immortelle, et/ou concourant par ses effets angiogéniques à la croissance tumorale, en facilitant la dissémination des cellules tumorales et l'apparition de métastases est un oncogène.

Les oncogènes sont regroupés en quatre catégories principales selon les propriétés fonctionnelles des proto-oncogènes dont ils sont issus : facteurs de croissance, récepteurs de surface de la membrane plasmique, protéines de signalisation et facteurs de régulation transcriptionnelle.

Un oncogène devient apte à transformer des cellules en acquérant une nouvelle fonction par rapport au proto-oncogène dont il dérive. Ce phénomène est connu sous le nom de gain de fonction (activation). Une fois activé, un oncogène fonctionne d'une façon dominante dans la cellule, ce qui signifie que l'activation d'un seul allèle suffit pour que la cellule échappe au contrôle normal de la prolifération cellulaire.

Il existe plusieurs mécanismes qui, au sein d'une cellule somatique, aboutissent à l'activation des oncogènes. En dehors de mutations ponctuelles ou de micro-délétions chromosomiques causant l'activation constitutive de la protéine codée du fait d'un changement dans sa structure, deux autres événements peuvent être à l'origine de l'amplification d'un proto-oncogène : le premier est l'insertion d'un rétrovirus dans le génome d'une cellule hôte, le deuxième est une anomalie chromosomique le plus souvent translocation d'un proto-oncogène d'un chromosome à un autre amenant celui-ci sous l'influence du promoteur d'un autre gène [47] et enfin la formation d'un produit chimère ayant un potentiel oncogénique [48].

Les oncogènes fréquemment identifiés chez les femmes atteintes de cancers sporadiques de l'ovaire sont *K-RAS*, *c-MYC*, *BRAF*, *HER-2/neu* et *AKT* [49-51]. Les mutations des gènes *K-RAS* et *BRAF* ont été rarement observées dans les carcinomes de l'ovaire à un stade avancé mais elles ont été trouvées dans des adénocarcinomes à un stade précoce, souvent dans les cellules épithéliales adjacentes bénignes et dans 30 à 50 % des cas de tumeurs épithéliales à la limite de la malignité ou « *borderline* » de l'ovaire, tumeurs représentant une forme intermédiaire entre la tumeur bénigne et le cancer [50, 52, 53]. L'amplification de *c-MYC* a été rapportée dans environ 30 % des cancers épithéliaux de l'ovaire [54]. *HER2* et *AKT* sont surexprimés respectivement dans 66 % et dans 12 à 30 % des carcinomes ovariens à un stade avancé mais rarement dans les tumeurs à un stade précoce et à la limite de la malignité [51, 55]. La surexpression de *HLA-G*, gène ayant un rôle dans les mécanismes d'échappement tumoral à l'action du système immunitaire, a été observée dans 61 % des cancers de l'ovaire à un stade avancé mais pas dans les néoplasmes à un stade précoce ou à la limite de la malignité [56]. Enfin, ni de mutation de *K-RAS* ni d'amplification de *c-MYC*,

de *HER-2/neu* ou d'*AKT* n'ont été observées dans les tumeurs ovariennes des porteuses de mutation de *BRCA1* ou *BRCA2* [57].

A.7.2. Anti-oncogènes et gènes suppresseurs de tumeurs

Les premières expériences qui ont conduit par la suite au concept de GSTs sont les expériences d'hybridations cellulaires entre une cellule normale et une cellule tumorale. La cellule hybride formée perdait le plus souvent son phénotype tumoral. Cette réversion vers un phénotype non tumoral impliquait qu'il existe un ou plusieurs gènes dont l'introduction dans ces cellules était capable de corriger la(les) fonction(s) déficiente(s) responsable du phénotype tumoral. En 1971, Knudson et al. ont introduit la notion de GSTs en proposant le modèle à deux étapes qui a été complété par Comings [58, 59]. Cependant, il a fallu attendre jusqu'en 1986, plus de dix ans après la découverte des oncogènes, pour identifier et cloner le premier GST *RBI* [60, 61]. Un an plus tard, Huang et al. ont démontré, pour la première fois, la propriété suppressive de tumeurs de *RBI* en réintroduisant un allèle sauvage dans les cellules cancéreuses [62].

Un gène suppresseur de tumeurs (GST) ou anti-oncogène est défini comme un gène dont l'absence favorise la transformation maligne et dont le remplacement dans les cellules défectives permet la réversion de ce phénotype malin [63]. Ainsi, un gène capable d'induire les mécanismes apoptiques de la cellule, ou de réguler le cycle cellulaire afin d'inhiber la prolifération cellulaire ou d'activer la différenciation cellulaire, ou de permettre la réparation des lésions de l'ADN, maintenant ainsi l'intégrité du génome ou encore de favoriser l'adhésion cellulaire empêchant ainsi la propagation tumorale pourrait intervenir soit comme anti-oncogène soit comme un GST.

Les gènes impliqués dans la réparation de l'ADN tels que *MLH1*, *MSH2* et *MSH6*, les gènes *BRCA1* et *BRCA2* intervenant dans la réparation des cassures double-brin de l'ADN par recombinaison homologue [64, 65] ainsi que le gène *Rbi* [62] représentent de bons exemples de GSTs.

Les mutations du GST *P53* entraînent dans la plupart des cas une prolongation de la demi-vie de la protéine de sorte que son accumulation devient détectable par immunohistochimie et se traduit par une surexpression de la protéine [66]. La surexpression de la protéine *P53* représente l'événement moléculaire le plus observés dans les cancers sporadiques et héréditaires de l'ovaire [49]. Dans une étude menée sur une cohorte de 22 femmes

australiennes et 26 norvégiennes atteintes de cancers sporadiques de l'ovaire, appariées avec autant de femmes atteintes d'un cancer de l'ovaire et porteuses d'une mutation de *BRCA1*, une surexpression de *P53* a été observée dans 73 % des cancers héréditaires et 81% des cas sporadiques chez les femmes norvégiennes et dans 86 % des cancers héréditaires et 64 % des cas sporadiques chez les femmes australiennes [67]. La différence observée n'était pas significative suggérant que l'inactivation de *P53* joue un rôle important dans les tumeurs ovariennes qu'elles soient sporadiques ou associées à une mutation de *BRCA1*.

A.7.3. Théorie de Knudson

Knudson, observant les rétinoblastomes sporadiques et familiaux, a émis l'hypothèse que deux événements et par conséquent la perte de fonction des deux allèles d'un GST serait nécessaire pour permettre l'initiation de la cancérogenèse (Fig. 3). Le premier événement (la première étape) pourrait être hérité ou acquis, permettant de distinguer les rétinoblastomes héréditaires des rétinoblastomes sporadiques. Le deuxième événement serait acquis dans les deux cas [68]. Le plus souvent, c'est l'association de la mutation d'un allèle et de la perte de l'allèle sauvage qui sont observées. Mais d'autres modes d'inactivation des deux allèles sont possibles dans ce modèle. Le premier événement pourrait se produire sous la forme d'une mutation ou d'un grand réarrangement (duplication ou délétion concernant un ou plusieurs exons d'un gène) de l'un des deux allèles du GST et le deuxième événement serait causé soit par une mutation intragénique, soit par une délétion de la bande chromosomique où est situé le gène, soit par une délétion du chromosome entier emportant l'allèle sauvage, soit par une perte chromosomique par non-disjonction lors de la mitose, enfin par recombinaison somatique avec réplication de l'allèle muté [68]. Cette succession d'événements provoquant la perte de l'expression du gène du fait soit de l'absence d'expression de la protéine, soit de l'expression d'une protéine non fonctionnelle va favoriser ainsi l'initiation de la carcinogenèse. A l'inverse des oncogènes qui auraient un effet dominant, les GSTs auraient un effet récessif au niveau cellulaire, l'inactivation des deux allèles étant nécessaire à la transformation tumorale. En revanche, la prédisposition aux cancers, elle, serait transmise sur un mode dominant. Ce modèle a été validé par les méthodes caryotypiques [69, 70] et par les études moléculaires qui ont amené à la localisation et au clonage du gène *RBI* [60, 71]. Comme nous le verrons plus loin, la théorie de Knudson va être remise en cause avec l'introduction de la notion de gènes aides-soignants ou « *caretakers* ».

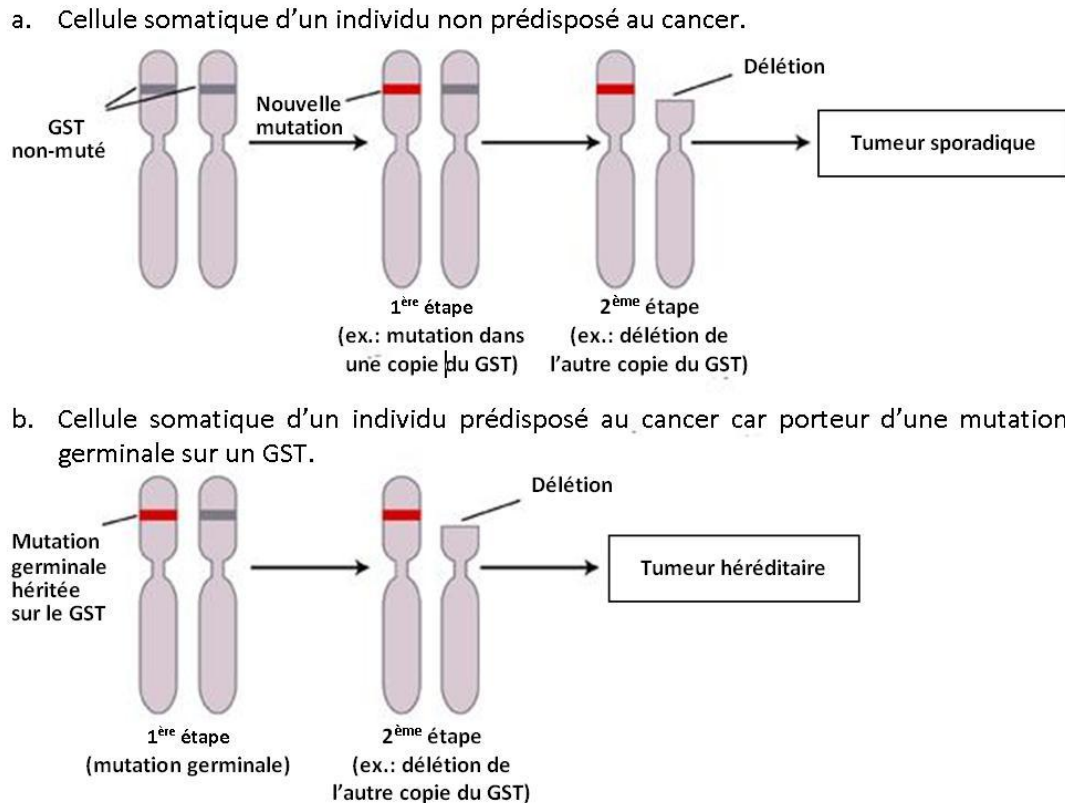


Fig. 3 : Hypothèse des deux étapes de Knudson.

a) Cette figure montre une paire de chromosomes dans une cellule somatique d'un individu non prédisposé. Nous observons les deux allèles d'un GST sous forme sauvage en mauve. L'exposition à un carcinogène par exemple peut entraîner l'apparition sur un des allèles sauvages d'une mutation en rouge (première étape). Les personnes non porteuses de mutation sur le GST possèdent deux copies normales de celui-ci. Il faut donc au moins deux étapes (mutations) indépendantes dans la même cellule, pour que survienne la carcinogénèse.

b) Cette figure présente une cellule chez un individu porteur d'une mutation germinale. Nous observons une paire de chromosomes, avec un allèle d'un GST sous forme sauvage en mauve et l'autre allèle avec une mutation figurée en rouge (première étape). Une délétion, représentant la deuxième étape, apparaît ensuite emportant la copie sauvage du gène. Les personnes avec une mutation germinale sur le GST ont déjà franchi cette première étape dans chacune des cellules de l'organisme de l'individu concerné. Une seule étape serait alors suffisante dans une cellule pour aboutir à l'initiation de la carcinogénèse.

Figure adaptée d'après [58].

Le modèle de Knudson permet de rendre compte de deux caractéristiques cliniques fréquemment observées dans les cancers héréditaires :

- la fréquence plus élevée des cancers se développant à des sites différents dans un même tissu ou bien dans des tissus différents chez un individu porteur d'une mutation germinale. En effet, toutes les cellules de cet individu étant porteuses de cette mutation, la probabilité qu'un second événement survienne à plusieurs sites distincts est alors plus élevée.
- l'âge plus précoce de diagnostic des tumeurs observées dans les formes familiales : l'individu étant déjà porteur d'une mutation, la probabilité qu'un second événement survienne à un jeune âge est alors plus élevée.

A.7.4. Tumorigenèse associée à *BRCA1* et *BRCA2* et théorie de Knudson

Bien que le modèle décrit initialement par Knudson puisse expliquer l'inactivation de certains GSTs comme *Rb*, la théorie des deux étapes rend finalement assez peu compte de ce qui est identifié en particulier avec des gènes tels que *APC*, *BRCA1* ou *BRCA2*. D'autres mécanismes peuvent être responsables de l'inactivation de ces gènes. Ainsi, *BRCA1* et *BRCA2* sont très rarement mutés dans les cancers sporadiques du sein [72, 73] et de l'ovaire [74-77]. Par conséquent, à première vue, il serait raisonnable de conclure que les gènes *BRCA* ne jouent pas un rôle important dans le développement des tumeurs sporadiques. Le premier groupe qui a conclu que *BRCA1* possède un rôle fondamental dans les cancers sporadiques a observé une diminution significative des niveaux d'expression de *BRCA1* lors de la transition du stade de cancers in situ au stade de tumeurs mammaires invasives [78]. Il a montré que l'inhibition de *BRCA1* par des oligonucléotides anti-sens entraînait une croissance accélérée des cellules mammaires normales et des cellules tumorales, montrant ainsi que *BRCA1* se comportait comme un régulateur négatif de la croissance des cellules épithéliales mammaires [78]. Depuis lors, un certain nombre d'études ont confirmé la fréquence d'une sous-expression de l'ARNm de *BRCA1* ou de sa protéine dans les cancers sporadiques du sein [79] et de l'ovaire [80], et certains ont suggéré une corrélation du niveau d'expression avec le degré de progression tumorale [81]. Une perte d'hétérozygotie (PDH) au locus *BRCA1* est fréquemment observée chez les femmes atteintes de cancers sporadiques de l'ovaire (38 % - 44 % [82-84]). Une diminution ou une perte d'expression de la protéine a été observée dans 90 % des cas, alors que seules 44 % des tumeurs présentaient une PDH et 3,5 % une mutation [84]. De ce fait, un pourcentage élevé de tumeurs présentent une diminution d'expression de *BRCA1* sans PDH. De plus, une

perte totale d'expression protéique a été observée dans 18 % des tumeurs ovariennes en l'absence de PDH et de mutation [84], suggérant dès lors qu'un autre mécanisme, tel qu'un mécanisme épigénétique ou une mutation dans la région régulatrice de *BRCA1* contribueraient à l'inactivation de ce gène dans les tumeurs sporadiques. En effet, la perte d'expression d'un des allèles de *BRCA1* par méthylation de son promoteur a pu être observée chez 31 % des femmes atteintes d'un cancer sporadique de l'ovaire [85]. De même, plusieurs études suggèrent un rôle de *BRCA2* dans les cancers sporadiques du sein [86] et de l'ovaire [86, 87]. Chan et al. ont rapporté la surexpression de l'ARNm de *BRCA2* dans les cancers sporadiques de l'ovaire et ont montré que cette surexpression est hautement corrélée avec une hypométhylation de la région promotrice de *BRCA2*, confirmant que l'absence de méthylation de *BRCA2* était un phénomène fréquent dans les cancers sporadiques [87]. Une PDH a été observée au locus *BRCA2* dans 31 % – 37 % des cas de cancers sporadiques de l'ovaire [82, 83, 88] avec une fréquence plus élevée (50 %) dans les cancers séreux [89]. De plus, Hilton et al. ont aussi observé dans 12 % des tumeurs sporadiques ovariennes, l'absence d'expression de l'ARNm de *BRCA2* sans pouvoir détecter dans ces tumeurs ni PDH ni anomalie de la méthylation de ce gène [82]. Cette dernière observation suggère que l'inactivation de *BRCA2* dans ces échantillons pourrait être due à un autre mécanisme, tel que l'inactivation d'un gène requis pour la transcription de *BRCA2* ou au contraire l'activation d'un gène requis pour l'inactivation de sa transcription.

BRCA1 et *BRCA2* présentent beaucoup de similitudes dans les tumeurs sporadiques de l'ovaire, mutations somatiques rares, haute fréquence de PDH, en 17q21-q24 et 13q12-q13, survenant presque exclusivement dans les tumeurs malignes épithéliales et rarement dans les tumeurs à faible malignité, prédominance dans les tumeurs de l'ovaire associée à une différenciation séreuse, finalement coexistence de PDH sur ces deux gènes. Tous ces résultats suggèrent que ces deux gènes pourraient agir en synergie dans le développement d'un sous-groupe de tumeurs sporadiques, particulièrement celles qui présentent une différenciation séreuse [82, 84, 85]. Il est aussi intéressant de noter que 58,2 % des cas avec une expression aberrante de l'ARNm et 56,3 % des cas avec une méthylation anormale ont des profils d'expression et de méthylation opposés de ces gènes [87].

La fréquence élevée d'anomalies d'un de ces gènes, voire des deux gènes présentes dans les tumeurs ovariennes (jusqu'à 83 %), témoignent finalement de leur importance dans la

tumorigénèse ovarienne, qu'elle soit sporadique ou héréditaire [82]. Cette importance est confirmée par les technologies génomiques à haut débit qui ont permis l'identification de profils d'expression associés à des mutations de *BRCA1* ou *BRCA2* dans les tumeurs mammaires [90-93] ou ovariennes [94]. Des profils distincts ont été obtenus en comparant des tumeurs associées à *BRCA1* et à *BRCA2* [92, 95, 96], suggérant que ces deux gènes appartiennent à deux voies distinctes de carcinogénèses. Ces techniques ont permis de différencier le profil d'expression des tumeurs familiales associées à *BRCA* et tumeurs non associées à *BRCA* [93]. De plus, il a été possible de subdiviser les tumeurs sporadiques en deux groupes selon leur profil d'expression, celles dont le profil est proche de celui des tumeurs associées à *BRCA1* et celles dont le profil est proche des tumeurs associées à *BRCA2* [94]. Ce qui vient corroborer le rôle de ces gènes dans les tumeurs sporadiques.

Les limites de la théorie de Knudson, soulevées par la mise en évidence de tumeurs avec PDH isolée, ou de tumeurs sans PDH ni mutations mais avec surexpression de ce gène ont amené à développer de nouveaux concepts aboutissant à la notion de carcinogénèse multi-étapes dont un des modèles le plus représentatif a été la carcinogénèse colique [68].

A.7.5. Gènes cerbères, aides-soignants et paysagistes

L'évolution des concepts de la carcinogénèse a amené à distinguer trois catégories de gènes : les gènes cerbères ou « *gatekeepers* » [97], les gènes aides-soignants ou « *caretakers* » [97], enfin les gènes paysagistes ou « *landscapers* » [98].

A.7.5.1. Gènes cerbères ou gardiens du génome

Les gènes cerbères assurent le maintien d'une prolifération cellulaire normale. Ils gardent constant le nombre de cellules dans un tissu en régulant l'entrée de la cellule dans les différentes phases du cycle cellulaire et en induisant l'apoptose en réponse à une agression de la cellule. Ainsi, les GSTs *P53* et *APC* représentent de bons exemples de gènes cerbères [99].

A.7.5.2. Gènes aides-soignants

Les gènes aides-soignants assurent le maintien de la stabilité du génome en détectant les dommages de l'ADN et en contrôlant la fidélité des processus de réparation. L'inactivation d'un gène aide-soignant va favoriser l'initiation de la carcinogénèse, en augmentant l'instabilité génomique. De ce fait, le taux de mutations sur tous les gènes y compris les

gènes cerbères va s'accroître. Parmi les gènes aides-soignants, nous citerons les gènes intervenant dans la réparation des mésappariements tels que *MSH2*, *MSH6*, *PMS2* et *MLH1* [100], les gènes impliqués dans la reconnaissance et/ou la signalisation des cassures de l'ADN comme *ATM* [99] et les gènes participant à la réparation des cassures double brin de l'ADN tels que *RAD50*, *MRE11*, *NBS1* [101]. Enfin, les gènes *BRCA1* et *BRCA2* de par leur fonction dans le maintien de l'intégrité du génome que nous décrirons plus loin se classent aussi dans cette catégorie [63].

Deux événements seraient en théorie nécessaires pour inactiver les deux copies d'un gène aide-soignant dans une cellule et entraîner l'apparition d'une instabilité génomique. Deux événements supplémentaires survenant dans la même cellule permettraient l'inactivation d'un gène cerbère et l'initiation tumorale. L'instabilité génomique résultant de l'inactivation des gènes aides-soignants augmente la probabilité de survenue de mutations sur un gène cerbère. L'initiation tumorale représenterait donc un processus multi-étapes (Fig. 4) [97].

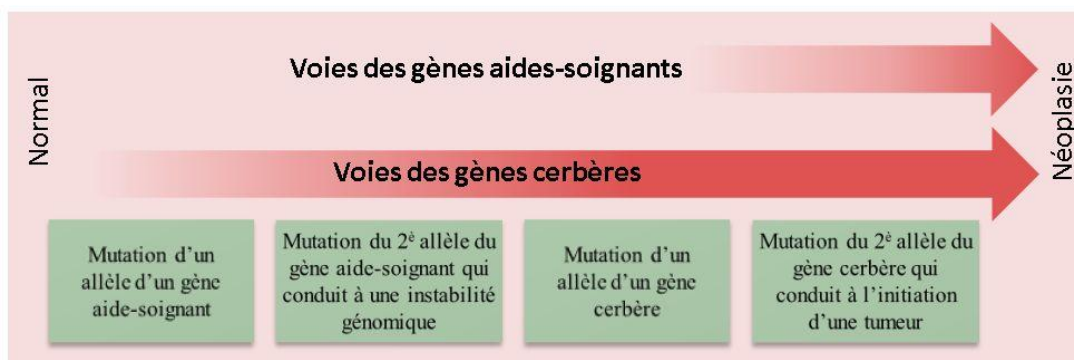


Fig. 4 : Carcinogenèse multi-étape.

Figure adaptée d'après [97].

Cette théorie pourrait rendre compte du processus tumoral pour la majorité des GSTs. Ceci a été également modélisé par Fearon et Vogelstein dans la séquence polype-carcinome [102] conduisant à élaborer une théorie en cinq étapes [68]. Cependant dans certains cas, un seul allèle, serait inactivé comme nous avons pu le montrer avec *BRCA1* [84] ou même est surexprimé comme il a été observé avec *BRCA2* [87]. La génétique du cancer est en voie de se complexifier et les récentes découvertes de la génomique, comme les variations du nombre de copies de gène ne font que la complexifier créant de nouveaux défis.

A.7.5.3. Gènes paysagistes

La notion de gènes paysagistes a été introduite après celle des gènes cerbères et aides-soignants. Ils modulent le micro-environnement des cellules de façon à inhiber leur transformation tumorale. Cette catégorie de gènes a été identifiée, pour la première fois, chez les personnes atteintes d'une polypose juvénile (PJ). Ce syndrome héréditaire est caractérisé par le développement, le long du tube digestif, de proliférations cellulaires bénignes de types hamartomateuses, proliférations dérivant du stroma. Les cellules épithéliales normales associées à ces proliférations tendent à subir une transformation néoplasique du fait de leur exposition à un micro-environnement anormal [103].

PTEN représente un exemple de gènes paysagistes favorisant la survenue de proliférations hamartomateuses [104]. La protéine PTEN est une phosphatase qui déphosphoryle principalement le phosphatidylinositol 3,4,5-triphosphate (PIP3), un phospholipide membranaire qui active un ensemble de voies de signalisation. La surexpression du PTEN dans les cultures cellulaires de mammifères favorise l'apoptose en abaissant le taux de PIP3 diminuant ainsi l'activité de la protéine kinase B (PKB) anti-apoptique. Les cellules dépourvues de PTEN contiennent des taux élevés de PIP3 et une PKB active. Puisque PKB exerce un effet anti-apoptique, la perte de PTEN empêche indirectement la mort cellulaire programmée et contribue ainsi à une prolifération cellulaire incontrôlée [105].

La mise en évidence de cette catégorie de GSTs souligne la nécessité de ne pas se limiter à l'étude d'un gène isolé, ni d'un type cellulaire isolé pour comprendre les mécanismes intervenant dans le développement tumoral, mais plutôt de tenir compte des interactions gènes-gènes, des interactions gènes-environnement mais aussi du rôle joué par le stroma et le microenvironnement tumoral dans les tumeurs.

A.8. *BRCA1* et *BRCA2*

L'identification des gènes *BRCA1* [2] et *BRCA2* [106] a représenté une avancée majeure dans la compréhension de la carcinogenèse héréditaire mammaire et ovarienne. Les gènes *BRCA1* et *BRCA2* sont des gènes majeurs de prédisposition aux cancers du sein et de l'ovaire. La prédisposition aux cancers est héritée sous forme de trait dominant. En d'autres termes, il suffirait d'hériter d'une copie mutée de *BRCA1* ou *BRCA2* pour être prédisposé à un risque élevé de développer de cancers du sein ou de l'ovaire. En revanche, jusqu'à un

passé récent, il était admis que l'initiation des tumeurs ne serait possible qu'après qu'un second événement inactive la copie sauvage de ce gène (mutation somatique) [43, 107], comme si l'allèle muté se comportait comme un allèle récessif en terme de transformation tumorale.

A.8.1. Structure et fonctions de BRCA1 et BRCA2

Bien qu'il existe des différences importantes entre les structures primaires des protéines BRCA1 et BRCA2, l'étude de la structure génomique des gènes correspondants révèle certains parallèles frappants. Ces deux gènes comportent de très longues séquences nucléotidiques, un exon central (exon 11) codant pour environ 50 % de la protéine, soit ~ 45 % pour BRCA1 et ~ 60 % pour BRCA2 [108]. Dans les deux cas, leur séquence nucléotidique est plus riche en A/T qu'en G/C et leur codon d'initiation se trouve dans le deuxième exon [109].

A.8.1.1. Structure de BRCA1

Le gène *BRCA1* « *breast cancer 1, early onset* » selon la nomenclature HUGO [110] a été localisé par Hall et al. en 1990 [111] et cloné par Miki et al. en 1994 [2]. Il est constitué de 81 155 pb et est situé sur la bande 17q21-q24 [110] du brin négatif du chromosome 17. Il s'étend de la base 38 449 840 à la base 38 530 994 (NM_007294). Ce gène est composé de 24 exons dont 22 codants épissés en un ARNm majoritaire de 7,8 kb (Tableau II). Ce dernier code pour une protéine nucléaire ubiquitaire de 220 kDa composée de 1 863 aa (acides aminés) [2].

L'épissage alternatif de l'ARN prémessager de *BRCA1* engendre au moins 12 transcrits alternatifs (NM_007294, <http://www.ncbi.nlm.nih.gov/sites/entrez>) (Fig. 5) [112]. Une grande hétérogénéité dans l'utilisation des sites de jonction d'épissage en 5' de l'ARN prémessager a été observée. Xu et al. ont mis en évidence l'existence de deux exons 1 appelés 1a et 1b à l'extrémité 5' de l'ARNm de *BRCA1*, mais ceux-ci ne sont pas retrouvés au niveau protéique puisque le site d'initiation de la traduction se situe sur l'exon 2. Ces deux exons sont retrouvés dans le préARNm codant pour la protéine pleine longueur. Ils sont mutuellement exclusifs et ne sont donc jamais retrouvés simultanément. Ils correspondent à deux promoteurs différents en 5' du gène qui ont plusieurs sites de liaison pour fixer les facteurs de transcription. C'est pourquoi la distribution de certains facteurs de régulation pourrait expliquer l'existence d'expression spécifique de tissus de certains

transcrits. Il existe au moins 12 variants d'épissage. Plusieurs d'entre eux seraient non fonctionnels et représenteraient le produit d'un phénomène d'épissage aberrant suite à une instabilité génomique dans les tumeurs où ils ont été identifiés (Fig. 5) [112-114]. Parmi ces variants, ceux ayant une terminaison prématurée, résulteraient d'une régulation cellulaire subtile survenant devant un changement du mode d'épissage alternatif engendrant des variants qui seraient dégradés par la voie de dégradation des ARNm portant un codon stop précoce. Les ARNm pleine longueur de *BRCA1*, les variants $\Delta(9,10)$, $\Delta(11q)$ et $\Delta(9,10,11q)$ représentent les variants d'épissage de *BRCA1* les plus connus. Ils sont appelés variants d'épissage prédominants parce qu'ils ont été rapportés par plusieurs études dans une variété de tissus, dans des conditions différentes (Fig. 5) [112]. Finalement, Fortin et al. ont mis en évidence un nouveau transcrit contenant un exon additionnel (exon 13a) n'entraînant pas un décalage du cadre de lecture [115]. Celui-ci est généré par l'insertion de 66 nucléotides entre les exons 13 et 14 suite à un épissage alternatif dans l'intron 13 (IVS13-2786-2720). De plus, il a été détecté dans tous les échantillons d'ARN provenant de 12 tissus normaux et de plusieurs lignées cellulaires cancéreuses y compris une lignée de cancer du sein.

Tableau II : Caractéristiques de *BRCA1* et *BRCA2*.

| Gène (HUGO) | Nomenclature (HUGO) | Bande (HUGO) | Localisation en bases (NCBI) | Longueur en paires de bases (pb) | Brin | Nombre d'exons | Séquence de Référence (Genbank) | Taille de la protéine (aa) | Nombre de transcrits publiés |
|----------------|---------------------------------|-----------------|------------------------------------|---|------|-------------------|--|-------------------------------------|---------------------------------------|
| <i>BRCA1</i> | breast cancer 1, early onset | 17q21- q24 | de 38 449 840 à 38 530 994 | 81 155 | - | 24 | NM_007294 | 1 863 | 12 |
| <i>BRCA2</i> | breast cancer 2, early onset | 13q12- q13 | de 31 787 617 à 31 871 809 | 84 193 | + | 27 | NM_000059 | 3 418 | 2 |

Tableau tiré de [2, 106].

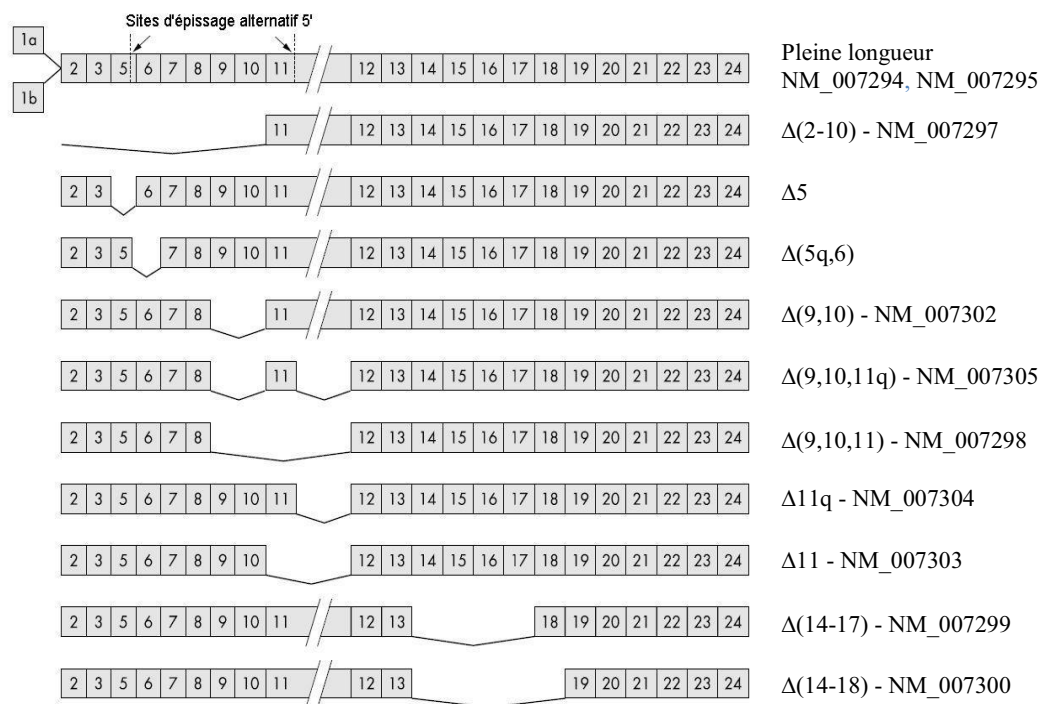


Fig. 5 : Variants d'épissage de *BRCA1* qui codent pour une protéine fonctionnelle.

Les segments numérotés représentent les exons constituant l'ARNm, les exons excisés correspondent aux régions figurées par des lignes qui connectent entre eux les exons présents. Les exons ne sont pas représentés à l'échelle. Ainsi, les deux lignes obliques correspondent à une interruption sur le 11^{ème} exon qui ne peut être représenté en totalité du fait de sa grande taille. Figure adaptée d'après [112].

A.8.1.2. Domaines fonctionnels de BRCA1

Depuis l'identification de *BRCA1* et *BRCA2*, une grande attention a été accordée à l'étude de leurs fonctions. Bien que ces dernières aient fait l'objet de multiples revues de la littérature [65, 116-124], elles sont encore loin d'être totalement élucidées. Cependant, les domaines fonctionnels présents dans un gène fournissent les premiers indices quant à leurs fonctions biologiques.

La protéine BRCA1 contient six domaines fonctionnels : une région N-terminale en anneau (Cys-3-His-Cys4), domaine à doigt de zinc de type RING, un signal d'exportation nucléaire (SEN) localisé à l'extrémité N-terminale, deux signaux de localisation nucléaire (SLN) situés dans l'exon 11 et une région C-terminale répétée en tandem constituée de deux domaines BRCT « *BRCA1 C-Terminal* » (Fig. 6) [125].

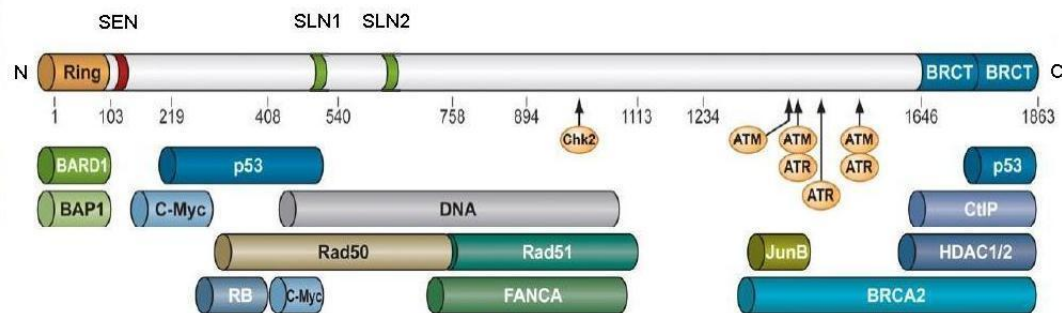


Fig. 6 : Structure primaire de la protéine BRCA1.

BRCA1 contient un domaine N-terminal à doigt de zinc de type motif RING, un signal d'exportation nucléaire (SEN), deux signaux de localisation nucléaire (SLN) et deux domaines BRCT C-terminaux. Les flèches noires désignent les sites de phosphorylation de BRCA1. Les protéines dont les interactions avec BRCA1 ont été validées expérimentalement sont représentées par des cylindres au-dessous de leurs sites de liaison. Figure adaptée d'après [126].

A.8.1.2.1. Domaine à doigt de zinc de type RING

Le domaine à doigt de zinc de type RING est un motif de fixation du zinc, composé d'une séquence conservée riche en cystéines permettant des interactions protéine-protéine ou protéine-ADN [127]. Le motif RING joue le rôle d'une ligase de type E3 permettant à BRCA1 de se dimériser avec la protéine BARD1 « *BRCA1-associated RING* », formant un complexe qui stabilise les deux protéines, permettant de diriger BRCA1 vers les sites de cassures de l'ADN et de jouer un rôle important dans la réparation de l'ADN par Recombinaison Homologue (RH) ciblée [128, 129]. De plus, il peut jouer le rôle d'une ligase ubiquitine-protéine dirigeant les protéines à dégrader, y compris BRCA1, vers le protéasome [130]. Il permet aussi à BAP1 « *BRCA1-associated protein 1* », une ubiquitine hydrolase C-terminale, de se fixer sur BRCA1 pour former un complexe qui augmente de quatre fois l'activité suppressive de tumeurs de BRCA1 [131]. Plusieurs hypothèses ont été formulées pour expliquer cette observation, entre autres que BAP1 stabiliserait la protéine BRCA1 en la protégeant contre la dégradation par l'ubiquitination. Ceci représenterait un mécanisme important pour empêcher la dégradation des protéines suppressives de tumeurs. Cependant, il a été observé qu'un des mutants inactifs sur le plan catalytique de BAP1 était toujours capable d'inhiber l'ubiquitination et donc de protéger BRCA1 de la dégradation, impliquant un second mécanisme d'action. Il faut souligner que l'inhibition de BAP1 entraînait une hypersensibilité des cellules aux radiations ionisantes et un retard dans la progression en phase S, ce retard pouvant être dû au fait que la cellule devenait incapable de

répondre aux dommages de l'ADN se produisant pendant cette phase. Ceci a amené à émettre une autre hypothèse sur le rôle de BAP1 suggérant que le complexe BRCA1/BARD1 et BAP1 réguleraient l'ubiquitination en réponse aux dommages de l'ADN [132]. Une mutation dans ce domaine pourrait donc non seulement rendre la désubiquitination par BAP1 inefficace mais aussi affecter la réparation de l'ADN contribuant ainsi au développement de cancers.

A.8.1.2.2. Signaux de localisation et d'exportation nucléaire

Les signaux de localisation et d'exportation nucléaire déterminent la localisation cellulaire de BRCA1. Ils représentent un moyen par lequel la cellule régule la réponse de BRCA1 aux lésions de l'ADN [133]. En effet, les signaux de localisation nucléaire SLN1 et SLN2 assurent l'importation de BRCA1 du cytoplasme vers le noyau par la voie des récepteurs importines afin de lui permettre de réparer les lésions de l'ADN. Une mutation dans SLN1 abolit l'interaction entre BRCA1 et son substrat, alors qu'une mutation dans SLN2 entraîne une abolition partielle de cette interaction [127]. La protéine BRCA1 mutée dans le domaine SLN1 ou SLN2 s'accumulerait donc dans le cytoplasme et serait incapable, en cas de dommages de l'ADN, de migrer au noyau afin de les réparer. Le signal d'exportation nucléaire (SEN), quant à lui, permet l'exportation de BRCA1 du noyau au cytoplasme par la voie exportine/CRM1 « *Chromosome Region Maintenance 1* ». Il semble que cette exportation ne puisse survenir, en réponse aux dommages de l'ADN, que dans les cellules ayant une copie sauvage de la protéine p53. Bien que BRCA1 ne possède pas de fonction spécifique dans le cytoplasme, son élimination du noyau, en réponse aux lésions de l'ADN, serait requise lors de l'arrêt du cycle cellulaire afin de prévenir un conflit entre la réponse aux lésions de l'ADN dépendante de BRCA1 et celle dépendante de p53. Un échec de l'exportation de BRCA1 dans le cytoplasme durant l'arrêt du cycle cellulaire pourrait créer des signaux contradictoires générant ainsi une cellule instable [133].

A.8.1.2.3. Domaines BRCT

Les domaines BRCT sont des motifs conservés communs aux protéines qui participent à la réparation de l'ADN et au contrôle du cycle cellulaire [125]. En plus de leur identification dans BRCA1, ils ont été retrouvés à l'extrémité C-terminale de plusieurs protéines impliquées dans la suppression de tumeurs [134], telles que p53BP1 interagissant avec Tp53 [135] et Rad9p jouant un rôle essentiel dans le contrôle du point de l'arrêt du cycle cellulaire lors de dommages à l'ADN dans la levure [135].

Les domaines BRCT de BRCA1 reconnaissent et lient les résidus pSer et pThr des protéines phosphorylées par les kinases ATM et ATR en réponse aux dommages à l'ADN, permettant ainsi à BRCA1 de former des complexes protéiques impliqués dans la réparation de l'ADN [136]. Ils permettent ainsi à BRCA1 de jouer le rôle d'un adaptateur moléculaire ou d'un médiateur assurant le recrutement des protéines ATM, ART et de leurs cibles (protéines phosphorylées par ATM et ART) aux sites des lésions de l'ADN [137]. En outre, ils lient plusieurs protéines, entre autres l'holoenzyme ARN polymérase II, p53, CtIP et RB [138]. Ils interagissent également avec des protéines altérant la chromatine et la structure de l'ADN telles que le remodeleur de la chromatine SWI/SNF [139], l'hélicase BACH1 [140], l'histone acétylase p300/CBP [141] et l'histone déacétylase HDAC1/2 [142], facilitant l'accès de la machinerie de réparation à l'ADN endommagé [118]. Magnard et al. ont mis en évidence que BRCA1 interagit au niveau du domaine BRCT avec l'Acétyl-Coenzyme A Carboxylase A, une enzyme intervenant dans le métabolisme lipidique [143]. Celle-ci catalyse la carboxylation de l'Acétyl-CoA en Malonyl-CoA, étape limitante de la synthèse des acides gras. Or les facteurs nutritionnels font partie des facteurs de risque de cancers du sein et de l'ovaire.

A.8.1.3. Fonctions de BRCA1

La protéine BRCA1 est exprimée de manière ubiquitaire et joue un rôle dans plusieurs processus fondamentaux pour le fonctionnement normal d'une cellule : réparation de l'ADN, régulation de la transcription et contrôle du cycle cellulaire [125].

A.8.1.3.1. Réparation de l'ADN

En réponse aux lésions de l'ADN, la protéine BRCA1 est hyperphosphorylée par les kinases ATM et ATR. Elle est localisée au niveau des fourches de réplication de l'ADN [144, 145], suggérant qu'elle jouerait un rôle dans l'initiation de la réparation des lésions de l'ADN. BRCA1 est impliqué dans des complexes protéiques qui activent la réparation des cassures double-brin (CDB) de l'ADN et qui initient la RH. BRCA1 et BRCA2 colocalisent pour former un complexe avec RAD51 [146], protéine nécessaire à la recombinaison durant la mitose et la méiose ainsi qu'à la réparation des CDB par RH. La mise en évidence de la colocalisation des protéines BRCA avec RAD51 au niveau des sites de recombinaison et des lésions de l'ADN a suggéré que les protéines BRCA1 et BRCA2 étaient impliquées à la fois dans la détection et la réparation de ces CDBs [117]. Certaines études ont proposé des interactions directes entre Rad51 et BRCA1 [144], mais d'autres ont suggéré le contraire

étant donné l'existence d'interactions négligeables sur le plan stoechiométrique entre ces deux protéines [147]. Il a été observé que l'accumulation de RAD51 dans les cellules souches embryonnaires de souris déficientes en *Brcal* était réduite après traitement par des agents endommageant l'ADN [148]. De plus, ces cellules présentent des anomalies de la réparation des CDBs par RH.

Les embryons murins *Brcal* $-/-$ outre qu'ils présentent un retard de développement, sont hypersensibles aux radiations γ . Les études du caryotype spectral (SKY) retrouvent un nombre élevé d'anomalies de nombre et de structure des chromosomes [149], suggérant que BRCA1 serait responsable du maintien de l'intégrité du génome. Les cellules souches embryonnaires *Brcal* $-/-$ sont hypersensibles à la mitomycine-C, agent induisant des ponts intercaténaires de l'ADN [150]. L'introduction d'un transgène *Brcal* dans les cellules *Brcal* $-/-$ ne restaure que partiellement le phénotype réparateur alors que la correction d'une mutation de *Brcal*, telle que la mutation *Brcal* $\Delta 11b$ restaure complètement la résistance à la mitomycine-C, et corrige l'instabilité chromosomique, suggérant que l'incapacité de la réparation des CDBs par RH ciblée entraînerait une instabilité du génome pouvant permettre l'initiation de la cancérogenèse [150].

En plus de son rôle dans la réparation par RH, BRCA1 est également impliqué dans les mécanismes de réparation par excision de nucléotide qui comportent deux voies : une voie globale qui répare les dommages de l'ADN indépendamment de leur localisation et une voie couplée à la transcription qui répare plus rapidement les lésions localisées dans le brin transcrit des gènes actifs [151, 152]. BRCA1 interviendrait donc à la fois dans la réparation couplée à la transcription et dans la réparation globale du génome [138].

L'ensemble de ces résultats supporte le rôle de gène aide-soignant pour *BRCA1* participant au maintien de l'intégrité du génome en favorisant la RH et en limitant les processus de réparation non homologue, qui auraient des effets mutagènes.

Finalement, des données récentes suggèrent que des complexes de BRCA1 seraient recrutés par la protéine RAP80 aux sites de dommages de l'ADN [153-155]. D'autres études suggèrent que le complexe BRCA1-CtIP interagirait avec le complexe MRN (RAD50-MRE11-NBS1) pour faciliter la résection des extrémités des CDBs pour permettre la réparation par RH [156, 157]. Tous ces résultats élucident les mécanismes de réparation de l'ADN dans lesquels BRCA1 est impliqué afin de préserver l'intégrité du génome.

A.8.1.3.2. Régulation transcriptionnelle

L'activité transcriptionnelle de BRCA1 dépend de sa capacité d'interagir directement ou indirectement avec différents facteurs de transcription. BRCA1 s'associe à plus de 15 protéines impliquées dans la transcription et se retrouve paradoxalement impliqué à la fois dans l'activation et la répression de la transcription [158]. Cable et al. [159] ont mis en évidence l'interaction des complexes protéiques ainsi formés avec des séquences spécifiques de l'ADN. Ces complexes sont sous-exprimés dans les lignées cellulaires de cancers sporadiques du sein qui ne présentent pas de mutation germinale de *BRCA1*. En revanche, les complexes contenant certaines mutations faux-sens de *BRCA1* causeraient une anomalie de l'activation du gène impliqué, montrant qu'une autre des fonctions de *BRCA1* serait de réguler l'expression d'autres gènes impliqués dans différents processus cellulaires.

Lane et al. ont proposé un modèle reliant les rôles de BRCA1 dans la transcription et la réparation [160]. Ils suggèrent que BRCA1 se lierait à la polymérase RNAPII de transcription processive dans les cellules non endommagées pour participer à la surveillance du génome. BRCA1 serait ainsi localisée à proximité des régions de la chromatine associées à la transcription des gènes actifs. En réponse aux dommages à l'ADN, BRCA1 serait alors phosphorylé et transporté aux sites actifs de la réparation. Des niveaux faibles de BRCA1 dans les tissus mammaires, ovariens et prostatiques, cibles potentielles pour la carcinogenèse associée à ce gène chez les porteurs de mutations de *BRCA1*, pourraient prédisposer ces tissus à des phénotypes d'instabilité génomique. En revanche, d'autres tissus en prolifération, notamment ceux du thymus et la rate, expriment des niveaux plus élevés de BRCA1 ce qui les protégerait contre ce phénotype.

BRCA1 pourrait auto-ajuster son expression grâce à la boucle que son extrémité terminale fait pour se lier à son promoteur [161]. Enfin, COBRA1, cofacteur de BRCA1, semblerait être impliquée dans la suppression de tumeurs par des voies de régulation de transcription communes avec BRCA1 [162, 163]. Pour conclure, ces observations mettent en évidence l'importance du rôle transcriptionnel de BRCA1 dans la suppression de tumeurs.

A.8.1.3.3. Contrôle du cycle cellulaire

Les points de contrôle du cycle cellulaire jouent un rôle essentiel dans la survie d'une cellule en empêchant la propagation des lésions de l'ADN au cours du cycle cellulaire avant leur réparation. L'utilisation de cellules diversement déficientes en protéines impliquées

dans la réponse aux dommages de l'ADN a permis de montrer que la présence simultanée des protéines ATM et BRCA1 était nécessaire pour induire l'arrêt du cycle cellulaire en phase S (phase de synthèse de l'ADN) et G₂ (gap 2 ou lacune 2, avant la mitose, préparation pour division cellulaire) en cas de lésions de l'ADN, permettant ainsi la réparation de l'ADN avant l'entrée en mitose [164]. En réponse à ces lésions, la protéine BRCA1 phosphorylée par ATM pourrait induire l'arrêt du cycle cellulaire en phase G2/M déterminant l'activation transcriptionnelle de GADD45 [165], de 14-3-3-sigma [166] et de la kinase Chk1 [167]. De plus, en réprimant celle de la cycline B1, BRCA1 produirait l'inactivation de la kinase mitotique CDC2 [167]. BRCA1 serait aussi impliquée dans la décaténation de l'ADN, suggérant son intervention dans la régulation de la ségrégation des chromosomes pendant la mitose et la dynamique du fuseau mitotique [168, 169].

Récemment, Yun et al. [170] ont suggéré qu'en réponse aux dommages à l'ADN, la phosphorylation du résidu Ser 327 de CtIP lors de l'entrée des cellules en phase S et le recrutement de BRCA1 favoriseraient la réparation des CDBs par RH. Ceci élucide le rôle majeur que BRCA1 et CtPI jouent dans la promotion d'un maintien efficace de la stabilité génomique.

Finalement, le complexe BRCA1-BARD1 est nécessaire pour l'accumulation de TPX2, un facteur principal dans la stabilisation des microtubules et l'attachement amphitélique des microtubules aux chromatides sœurs [171]. Ces résultats suggèrent que le rôle de BRCA1 dans la régulation du cycle cellulaire représente un facteur important dans la réparation des lésions de l'ADN et ainsi dans le maintien de l'intégrité génomique.

A.8.1.3.4. Régulation de l'invasion cellulaire

BRCA1 joue un rôle important dans l'inhibition de la progression tumorale. Wang et al. suggèrent que BRCA1 pourrait transactiver la région promotrice de CAV1, « *Caveolin-1* », composant principal de la membrane plasmique des cavéoles, générant un niveau élevé d'expression de ce suppresseur de tumeurs qui intervient dans le contrôle de l'invasion [172]. BRCA1 pourrait ainsi inhiber la capacité d'invasion et le pouvoir métastasant des cellules tumorales mammaires en induisant la migration de CAV1 du cytoplasme à la membrane cellulaire.

En conclusion, BRCA1 est une protéine multifonctionnelle qui joue un rôle important dans le maintien de l'intégrité génomique. Ces fonctions font de *BRCA1* un gène aide-soignant. Nous reviendrons sur les éléments qui confirment ce rôle dans notre discussion.

A.8.1.4. Structure de BRCA2

En 1995, Wooster et al. identifiaient les deux tiers de la séquence codante du gène *BRCA2* « *breast cancer 2, early onset* » [3]. Un an plus tard, Tavtigian et al. publiait la séquence complète de *BRCA2* [1]. Composé de 84 193 pb, ce gène est situé sur la bande 13q12.q13 [110] du brin positif du chromosome 13. Il commence à la base 31 787 617 et se termine à la base 31 871 809. Ce gène comporte 27 exons dont 26 exons codants épissés en un ARNm majoritaire de 10 à 12 kb (NM_000059). Ce dernier code pour une protéine nucléaire ubiquitaire de 384 kDa composée de 3 418 aa (Tableau II) [1]. Contrairement à *BRCA1*, seul un petit nombre de transcrits alternatifs de *BRCA2* ont été identifiés à l'heure actuelle : un transcrit alternatif majoritaire avec délétion de l'exon 12 ($\Delta 12$) [173], plusieurs isoformes contenant une insertion de 64 paires de bases ne décalant pas le cadre de lecture, appelés collectivement isoformes *BRCA2 v20*, ont été décrits dans les leucocytes. Cette insertion a été appelée exon 20a. Elle dérive d'un site cryptique d'épissage situé dans l'intron 20 [174]. Deux autres isoformes ont été identifiées. Celles-ci éliminent l'exon 22 utilisant un site cryptique d'épissage dans l'exon 23. Toutes ces isoformes codent pour une même protéine tronquée qui a perdu les acides aminés codés par les sept derniers exons du gène [174]. Enfin, Yang et al. [177] ont identifié deux polymorphismes synonymes dans *BRCA2*, Ser2414Ser (7242A>G) et Ser455Ser (1365A>G) entraînant un biais dans l'utilisation préférentielle de codons, sans association avec le risque de survenue de cancers du sein [175].

A.8.1.5. Domaines fonctionnels de BRCA2

La protéine BRCA2 n'a aucune homologie avec la protéine BRCA1 en ce qui concerne sa structure primaire. Elle comporte un domaine N-terminal ayant une activité transcriptionnelle putative, huit domaines BRC répétés distribués sur différents sites de la région centrale codée par l'exon 11 et deux signaux de localisation nucléaire situés à l'extrémité C-terminale (Fig. 7).

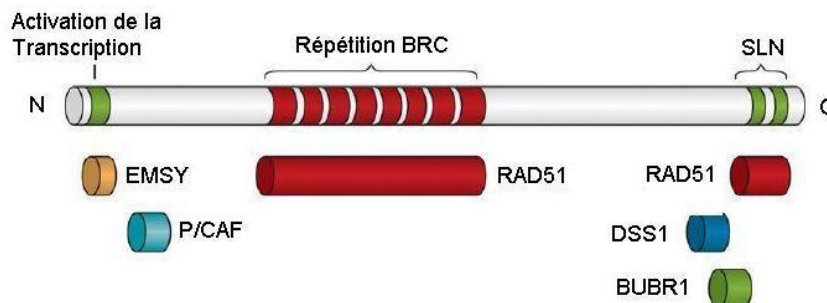


Fig. 7 : Structure primaire de la protéine BRCA2.

BRCA2 contient un domaine d'activation de la transcription, huit domaines répétés BRC et deux signaux de localisation nucléaire (SLN). Les protéines qui interagissent avec BRCA2 sont représentées par des cylindres au-dessous de leurs sites de liaison. Figure adaptée d'après [176].

A.8.1.6. Fonctions de BRCA2

La protéine BRCA2 est exprimée de façon ubiquitaire. Elle joue un rôle majeur dans la réparation de l'ADN, cependant son rôle dans la régulation de la transcription et du point de contrôle du cycle cellulaire n'est pas encore bien défini.

A.8.1.6.1. Réparation de l'ADN

BRCA2 jouerait un rôle essentiel dans la réparation des CDBs. La migration de RAD51 vers le noyau est déficiente dans les cellules privées de BRCA2, suggérant que BRCA2 transporterait RAD51 au site des lésions de l'ADN [177]. Le complexe BRCA2-RAD51 existerait sous deux formes : une forme inactive empêchant la fixation de RAD51 sur un brin monocaténaire de l'ADN et une forme active, RAD51 formant alors des filaments nucléoprotéiques après son transport par BRCA2 aux sites des dommages de l'ADN. La transition de la forme inactive à la forme active serait rendue possible par la phosphorylation de ces protéines en réponse aux dommages de l'ADN.

Les cellules déficientes en BRCA2 présentent une sensibilité accrue aux radiations ionisantes, induisant un défaut dans la réparation des CDBs [178, 179]. Elles accumulent en culture des cassures chromosomiques. Des échanges aberrants se produisent pendant la mitose, induisant l'apparition de translocations chromosomiques non homologues, fragments acentriques et de larges délétions [178, 179]. Des phénotypes similaires sont observés dans les cellules déficientes en RAD51 [180]. Ces deux observations montrent que

l'interaction de BRCA2 avec Rad51 joue un rôle fondamental dans le maintien de la structure chromosomique.

Une partie des fonctions de BRCA2 a pu être élucidé grâce à l'étude de l'anémie de Fanconi (AF). L'AF est un désordre autosomal récessif caractérisé par une anémie, une insuffisance médullaire, des fausses couches et l'existence d'une myélodysplasie qui va se transformer en leucémie myéloïde aigue [181]. Il s'agit d'un désordre héréditaire causée par une homozygotie pour au moins huit gènes différents (de FA-A à FA-G).

Les cellules lymphoblastoïdes des patients atteints d'AF ont une hypersensibilité caractéristique aux agents capable de créer des ponts intercaténaire tel que la Mytomycine C (MMC) et le diepoxybutane, à la base de l'identification en clinique des sujets atteints, un test fonctionnel permettant de classer les sujets atteints par rapport à leur groupe de complémentation [181]. Des études s'intéressant à la génétique de l'AF ont identifié l'importance de BRCA2 dans la réparation des ponts intercaténaire de l'ADN [182, 183] chez ces patients. Il s'est avéré que les patients atteints d'AF et appartenant au groupe de complémentation AF-D1 présentaient en fait une mutation biallélique de *BRCA2*, lequel s'est avéré par la suite être le gène *FANCD1* [183]. Il faut noter que les cellules du groupe de complémentation de *FANCD1* de même que les cellules qui ont perdu la fonction de *BRCA2* présentaient un défaut de formation des foci nucléaires de *RAD51* après irradiation par les rayons X ou le traitement par les MMC [184]. *BRCA1* jouerait aussi un rôle dans la voie FA en interagissant avec *FANCD2* [184]. Finalement, *BRCA2* supprimerait les glissements des chromatides sœurs lors de la réplication, suggérant ainsi le rôle de *BRCA2* dans la cohésion et/ou l'alignement des chromatides sœurs [65].

Ainsi, *BRCA2* jouerait un rôle direct dans la réparation des lésions de l'ADN. Son absence ou la perte de sa fonction pourrait donc jouer un rôle fondamental dans l'apparition d'une instabilité génomique et par conséquent dans l'initiation de la carcinogenèse. Ceci l'apparente à un gène aide-soignant.

A.8.1.6.2. Régulation transcriptionnelle

Le rôle de *BRCA2* dans le contrôle de la transcription n'est pas bien défini. Une surexpression de *BRCA2* semblerait être associée à une diminution de l'activité transcriptionnelle basale de p53 [125]. *Brca2* pourrait aussi activer p53 d'une façon indirecte. En effet, la perte de *Brca2* entraîne un défaut de réparation des dommages de

l'ADN, causant leur accumulation, ce qui entrainerait l'activation de p53 entraînant l'arrêt du cycle cellulaire [185]. En revanche, BRCA2 pourrait activer la transcription via P/CAF, protéine co-activatrice de transcription ayant une activité d'acétylation des histones [186]. De plus, la région codée par l'exon 3 de BRCA2 pourrait avoir une activité transcriptionnelle [187]. Ainsi, Nordling et al. ont rapporté la délétion de cette région dans une famille suédoise à risque élevé de cancers du sein et de l'ovaire, suggérant que cette délétion pourrait prédisposer aux cancers observés dans cette famille [188]. Il a ensuite été montré que cette activité transcriptionnelle serait réprimée par la fixation de la protéine EMSY sur l'exon 3 de BRCA2 [86], Ainsi la protéine EMSY, se lie à *BRCA2* sur une région (exon 3) qui peut être délétée dans certains cancers. EMSY est capable de supprimer le potentiel activateur de l'exon 3 de *BRCA2*, de s'associer au régulateur de la chromatine HP1 β et BS69, et de se localiser aux sites de réparation après une lésion de l'ADN [86]. La surexpression du gène EMSY qui intervient via son implication génique dans les tumeurs du sein et de l'ovaire rend compte d'un mécanisme de la carcinogenèse associée au gène *BRCA2* distinct de ce qui avait été observé jusqu'ici et dont l'exploration doit se poursuivre [189].

A.8.1.6.3. Contrôle du cycle cellulaire

Comme nous venons de le mentionner, BRCA2 inhibe la transcription de p53 [190]. Cette inhibition permettrait de réguler la durée de l'arrêt du cycle cellulaire induit par p53 pour réparer les lésions de l'ADN. De plus, BRCA2 contrôlerait le point de contrôle G2/M en interagissant avec la protéine BRAF35 « *BRCA2-associated factor 35* » qui se lie à la structure de l'ADN [191]. Le complexe BRAF35/BRCA2 s'associe avec la chromatine condensée pendant la mitose, suggérant l'implication de ce complexe dans le remodelage de la chromatine et dans la progression du cycle cellulaire [191]. De plus, en réponse aux dommages du fuseau mitotique, BRCA2 est phosphorylée par hBUBR1, une protéine impliquée dans le point de contrôle de la phase M (la mitose) lors de ce type de dommages [192]. La coexpression et la colocalisation de BRCA2 et hBUBR1 ont été observées au noyau des cellules ayant un fuseau mitotique endommagé, mais pas dans les cellules ayant des fuseaux intacts, proposant un rôle de BRCA2 au niveau du point de contrôle mitotique. Enfin, BRCA2 interagit avec DSS1, une protéine requise pour une croissance cellulaire normale [193].

En conclusion, une des fonctions majeures de *BRCA2* est son implication dans la réparation de l'ADN et le maintien de la stabilité du génome. Si cette fonction est déficiente, l'accumulation de nombreuses anomalies chromosomiques est susceptible de permettre l'apparition d'une anomalie sur un gène cerbère responsable de l'initiation de la carcinogenèse. *BRCA2* serait donc un gène aide-soignant.

A.8.2. Pénétrance des mutations de *BRCA1* et *BRCA2*

Une méta-analyse de 22 études, effectuée par Antoniou et al. [4], a estimé le risque cumulé de cancers du sein chez les porteuses d'une mutation sur *BRCA1* et *BRCA2* respectivement à 65% (95% IC : 51% – 75%) et 45% (95 % IC : 33% – 54%), à l'âge de 70 ans. Les estimations correspondantes pour le cancer de l'ovaire sont de 39% (95% IC :22% - 51%) pour *BRCA1* et de 11% (95% IC : 4,1% - 18%) pour *BRCA2*.

Une autre méta-analyse a été menée par Friedenson [5], portant sur 30 études s'adressant soit à des cohortes de porteurs de mutation de *BRCA1* ou *BRCA2*, soit à des populations d'individus à haut risque d'être porteurs d'une mutation de *BRCA1* ou *BRCA2* (apparentés de sujets atteints et porteurs). L'incidence des cancers autres que les cancers du sein et de l'ovaire a été calculée indépendamment des catégories d'âge dans l'ensemble de ces cohortes montrant une augmentation du risque relatif cumulé sur la vie d'autres cancers : cancers de l'estomac, de la prostate, du pancréas et du côlon (Tableau III).

Tableau III : Risque relatif de survenue de cancers autres que les cancers du sein et de l'ovaire chez les porteurs de mutation sur *BRCA1* et *BRCA2* et apparentés.

| Cancers | Risque relatif de cancers chez les porteurs |
|---------------|---|
| Estomac | 1,69 [1,21 - 2,38] |
| Prostate | 1,62 [1,31 - 2,00] |
| Pancréas | 1,32 [1,13 - 1,56] |
| Colo-rectoral | 1,19 [1,02 - 1,38] |

Tableau tiré de [5].

A.8.3. Mutations de *BRCA1* et *BRCA2* et particularités canadiennes françaises

Jusqu'à ce jour, près de 12 000 mutations germinales de *BRCA1* et 11 300 de *BRCA2* identifiées dans les familles à risque élevé de cancers du sein et/ou de l'ovaire ont été répertoriées dans la base de données BIC « *Breast Cancer Information Core* »

(<http://research.nhgri.nih.gov/bic/>) [194]. La majorité des mutations qui sont responsables d'une prédisposition héréditaire aux cancers du sein et/ou de l'ovaire sont des insertions ou délétions entraînant un décalage de phase du cadre de lecture ou des mutations non-sens, mutations qui vont entraîner l'apparition d'un codon stop et la production d'une protéine tronquée. Des mutations d'épissage peuvent également créer un site d'épissage alternatif et déterminer la synthèse d'une protéine non fonctionnelle.

Il existe également des mutations faux-sens substituant un acide aminé à un autre acide aminé entraînant pour un certain nombre d'entre elles un changement de la fonction de la protéine. Mis à part quelques mutations faux-sens affectant les résidus cystéines des domaines BRCT et à doigt de zinc de type RING de *BRCA1*, la conséquence de ces faux-sens sur la fonction de la protéine est longtemps restée inconnue. Un grand nombre de ces substitutions nucléotidiques n'étaient pas enregistrées dans la base de données du BIC. Actuellement, elles sont soit enregistrées comme des variants de signification inconnue, soit comme des polymorphismes. Un petit nombre d'entre elles a été reconnu comme mutations délétères. Les substitutions nucléotidiques ou SNP « *Single Nucleotide Polymorphisms* » font actuellement l'objet de recherche pour identifier leur impact comme modificateur du risque de cancers, par différentes méthodes : utilisation de modèle *in silico* [195, 196], utilisation de modèles cliniques [197], études d'association [198] ou techniques de biologie moléculaire basées sur des essais fonctionnels [199]. L'avènement du projet HapMap [200, 201] rend compte de l'intérêt croissant pour le rôle modificateur du risque de cancers des SNPs associés à une mutation dans l'un des gènes *BRCA* [202, 203].

Trente à 50 % des personnes prédisposées aux cancers du sein et/ou de l'ovaire sont porteurs d'une mutation germinale de l'un des gènes *BRCA* [204, 205]. La fréquence de ces mutations dans les familles à risque varie d'une population à une autre. De plus, certaines populations présentent une grande diversité de mutations, alors que certains groupes ethniques présentent un nombre restreint de mutations avec une fréquence élevée en raison d'un effet fondateur [206], qui correspond à la transmission de mutations germinales dans une population issue de la migration d'un nombre restreint d'individus appartenant à une même population mère [207]. Un effet fondateur a ainsi pu être identifié dans différents groupes ethniques, tout particulièrement les populations Juives Ashkénazes, Islandaises, Hollandaises, et Canadiennes Françaises [206, 208, 209].

Entre 1608 et 1795, 8 500 immigrants venant de France ont migré dans ce qui a été appelé la Nouvelle-France [207]. Certaines régions de France ont été sur-représentées parmi les immigrants, notamment les régions de la côte atlantique et des environs de Paris [210], créant ainsi un effet fondateur dans la population Canadienne Française [210]. De ce fait, un groupe de mutations prédisposant aux cancers du sein et de l’ovaire est sur-représenté chez les Canadiens Français sans que la proportion de ces cancers soit plus élevée au Québec qu’ailleurs. Les mutations fondatrices les plus fréquentes chez les Canadiennes Françaises sont *BRCA1*-C4446T, *BRCA2*-8765delAG et *BRCA2*-3398delAAAAG (1,7 % des femmes atteintes de cancers du sein incidents diagnostiqués avant l’âge de 41 ans et 1,3 % des femmes atteintes de cancers de l’ovaire) [211-213].

Les mutations de *BRCA1* et *BRCA2* identifiées chez les membres des familles prédisposées aux cancers du sein et/ou de l’ovaire sont dispersées sur l’ensemble de la séquence des deux gènes dans la majorité des populations (Fig. 8). Dans la population Canadienne française, une mutation de *BRCA1*, la mutation C4446T (R1443X) située sur l’exon 13, représente 60 % des mutations de *BRCA1* observées dans les familles porteuses d’une mutation sur ce gène. Elle concerne 18 familles / 28 (Simard et al. [209]) et 46 familles / 78 (Maugard et al., données personnelles non publiées), (Fig. 9). La mutation la plus fréquemment observée sur *BRCA2* est la mutation 8765delAG localisée sur l’exon 20. Elle représente 59 % des mutations de *BRCA2* observées dans les familles porteuses d’une mutation sur ce gène (28 familles / 33 pour Simard et al. [209], 42 familles / 86 pour Maugard et al., données personnelles non publiées). Il est remarquable de noter que la mutation G6085T n’a été relevée dans aucune famille de la cohorte recrutée par Simard et al., alors qu’elle a été observée 16 fois dans notre cohorte à la clinique des cancers familiaux du CHUM, suggérant ainsi que cette mutation serait liée à des flux migratoires distincts entre la population recrutée au CHUM et celle recrutée à Québec, ce qui justifierait de préciser l’origine des ascendants des porteurs de cette mutation.

Il est intéressant de noter que 60 % des mutations fondatrices de *BRCA1* sont localisées sur l’exon 13 (46 familles / 78 pour Maugard et al. et 18 familles / 28 pour Simard et al.) et que 59 % de celles de *BRCA2* sont localisées sur l’exon 20 (42 familles / 86 pour Maugard et al. et 28 / 33 pour Simard et al.) (Fig. 9) qui sont tous deux des exons de petite taille. Alors que l’exon 11 représente plus de 50% de la séquence codante pour chacun de ces deux gènes, seules 35% des mutations fondatrices de *BRCA1* (9 familles / 28 pour Simard et al., 28

familles / 78 pour Maugard et al.) et 38 % de celles de *BRCA2* (3 familles / 33 Simard et al. et familles 42 / 86 pour Maugard et al.) sont localisées sur cet exon.

L'identification des mutations fondatrices dans la population Canadienne Française permet au généticien de proposer une stratégie diagnostique spécifique adaptée à cette population, stratégie en deux temps, évitant ainsi de faire un séquençage extensif de ces gènes à tous les patients [208, 209].

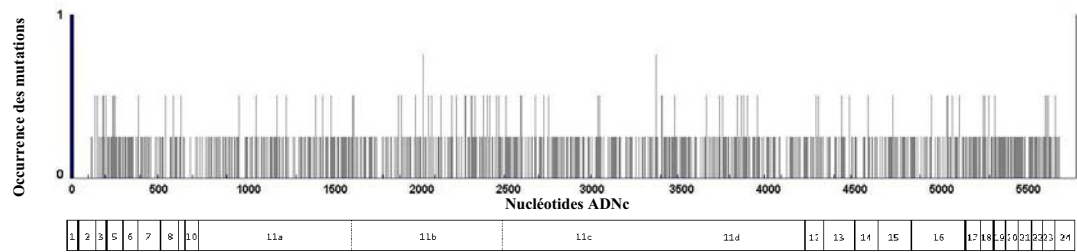
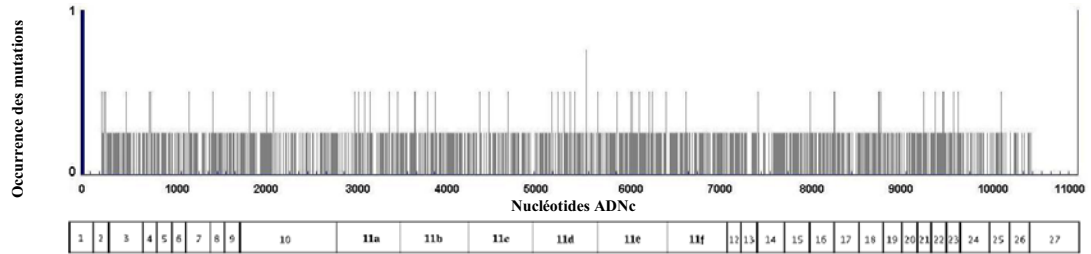
BRCA1*BRCA2*

Fig. 8 : Distribution des mutations de *BRCA1* et *BRCA2* dans les familles prédisposées aux cancers du sein et de l’ovaire.

Figure adaptée d’après [194].

BRCA1

1

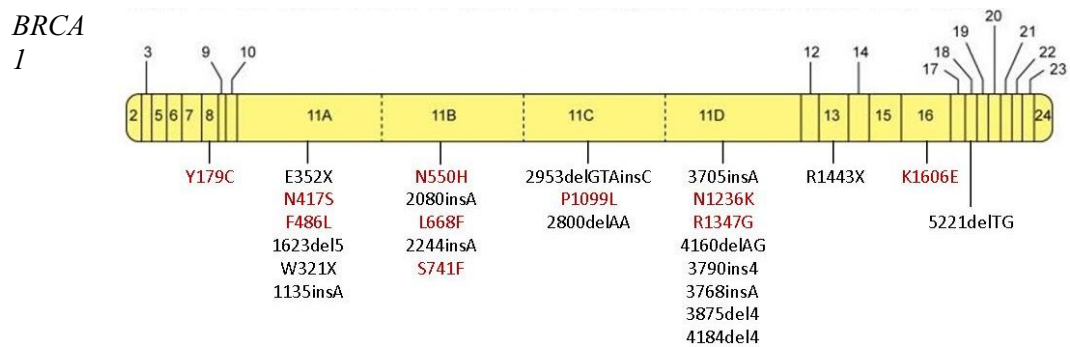
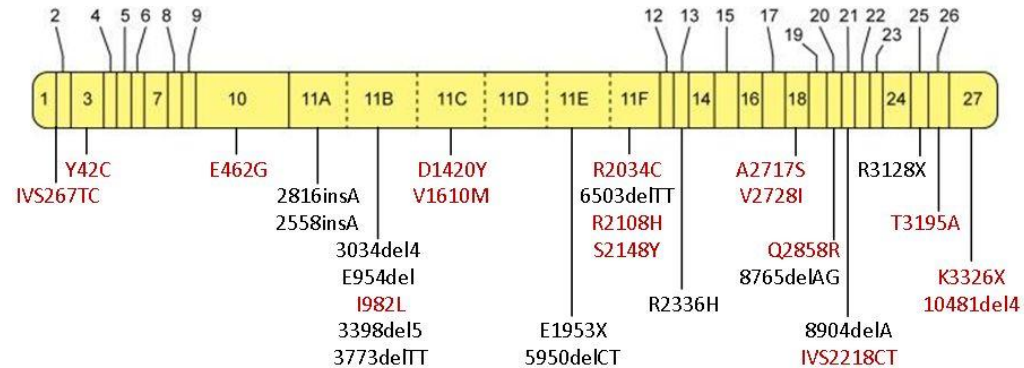
*BRCA2*

Fig. 9 : Distribution des mutations fondatrices de *BRCA1* et *BRCA2* dans les familles prédisposées aux cancers du sein et/ou de l’ovaire d’origine Canadienne Française. D’après Simard et al. [209] et données personnelles de Maugard et al.

A.9. Haplo-insuffisance et gènes *BRCA*

Bien que la théorie des deux étapes expliquant l'inactivation d'un GST soit admise depuis plusieurs années, les données récentes de la littérature montrent que cette théorie n'est pas universelle. En particulier, nous l'avons montré plus haut, cette théorie ne rend pas compte de ce qui est observé dans la carcinogenèse associée aux gènes *BRCA1* et *BRCA2*.

Plusieurs phénomènes peuvent expliquer qu'un phénotype tumoral apparaisse chez un individu porteur d'une mutation sur un allèle d'un gène. L'exemple le plus trivial est celui du chromosome X et du Y chez l'homme. Dans ce cas, il s'agit de gènes hémizygotés présents seulement sur le X ou sur le Y chez le mâle [184]. Nous retiendrons également l'inactivation d'un des deux X chez la femme ou l'inactivation d'un allèle d'un autosome par un phénomène d'empreinte (méthylation d'un des promoteurs) [184].

L'existence d'une mutation dominante négative qui bloque la fonction de l'allèle sauvage [214, 215], la présence d'une mutation gain-de-fonction responsable d'une activité aberrante du gène muté et finalement une mutation entraînant la perte de fonction d'un allèle suivie d'une perte de l'allèle sauvage résultant en un phénotype nul représentent trois de ces phénomènes. Une autre possibilité est l'existence d'une mutation perte de fonction sans perte de l'allèle sauvage, mutation responsable d'une haplo-insuffisance avec effet de dosage du gène.

La perte d'un seul allèle d'un GST a longtemps été considérée comme n'ayant pas ou peu de signification fonctionnelle. La notion d'haplo-insuffisance et de dosage génique mérite d'être considérée ici bien qu'elle ait longtemps été contestée pour les gènes *BRCA* en particulier. Cependant, des études récentes suggèrent que la perte d'un seul allèle pourrait être impliquée dans le processus de carcinogenèse associé à ces gènes, nous y reviendrons.

Pour certains gènes, plus d'un mécanisme peut être en cause. Ainsi, certaines mutations de *P53* peuvent présenter un effet dominant négatif, alors que d'autres peuvent entraîner une perte de fonction de *P53* [184]. Dans la leucémie aiguë promyélocytaire (LAP), la translocation impliquant les gènes *PML* et *RAR α* conduit à une réduction du dosage de *PML*, mais aussi à la création de la protéine chimérique *PML-RAR α* qui peut avoir un effet dominant négatif [184]. De plus, l'haplo-insuffisance de plusieurs gènes peut se traduire par une pénétrance incomplète d'un GST selon la définition classique du modèle de Knudson.

L'haplo-insuffisance lorsqu'elle concerne des gènes tels que *BRCA1* et *BRCA2* impliqués dans la stabilité du génome pourrait favoriser la transformation tumorale en rendant la cellule génétiquement instable favorisant la survenue et l'accumulation de mutations. Ainsi, nous nous attendons à ce que la perte ou du moins la réduction de la capacité de la cellule de réparer l'ADN joue un rôle dans l'initiation de la carcinogenèse.

Nous allons voir quelques exemples qui suggèrent que les gènes *BRCA1* et *BRCA2* répondent au mécanisme d'haplo-insuffisance et que cette haplo-insuffisance peut se traduire par un phénotype particulier, y compris dans les cellules dites normales.

A.9.1. *BRCA1*, *BRCA2* et cellules tumorales

Jusqu'alors, la perte de l'allèle sauvage au locus *BRCA1* ou *BRCA2* était considérée comme un événement nécessaire pour la transformation tumorale dans les tumeurs héréditaires du sein et/ou de l'ovaire. S'il paraissait clair qu'être hétérozygote pour une mutation sur l'un de ces gènes prédisposait aux cancers du sein et de l'ovaire, la discussion portait essentiellement sur le fait de savoir si la PDH survenait comme un événement stochastique ou bien si l'hétérozygotie en elle-même favorisait la perte de l'allèle sauvage, comme cela a pu être observé pour *Rb* [184]. Quoi qu'il en soit elle reposait sur le prémisses qu'une PDH au locus correspondant à l'allèle sauvage était nécessaire pour permettre la progression tumorale.

Cependant, récemment un niveau élevé d'hétérogénéité de PDH a été retrouvé dans les lésions invasives et entre lésions pré-invasives et invasives chez les patientes atteintes de cancers du sein et porteuses d'une mutation de *BRCA1* ou *BRCA2*, suggérant que la perte de l'allèle sauvage de *BRCA1* ou *BRCA2* n'était pas nécessaire à la carcinogenèse mammaire associée aux gènes *BRCA* [216]. En revanche, cette étude avait mis en évidence une PDH dans tous les cancers de l'ovaire analysés. La rétention de l'allèle sauvage a été rapportée par Hamel et al. pour une tumeur ovarienne de type germinale chez une femme porteuse d'une mutation sur *BRCA2* [217]. Cet histotype particulier n'étant pas jusqu'alors associé au spectre des tumeurs observées chez les porteuses de mutation, la question de l'haplo-insuffisance n'avait pas été posée.

Pour notre part, nous nous sommes intéressés à la traduction de l'hétérozygotie des gènes *BRCA1* et *BRCA2* sur le plan phénotypique, et plus précisément à sa traduction en terme de phénotype d'expression (transcriptome) dans les cellules non-tumorales.

A.9.2. Lymphocytes et hétérozygotie aux loci *BRCA1* et *BRCA2*

L'hétérozygotie pour *BRCA2* détermine dans la lignée de poulet DT40, lignée cellulaire B, un phénotype particulier caractérisé par un taux de prolifération cellulaire réduit, une augmentation de la mort cellulaire, une sensibilité élevée aux agents capables de provoquer des lésions de l'ADN et une diminution de la formation de foci avec *RAD51* après irradiation [218]. Le Dr Maugard l'a également observé dans la lignée cellulaire TOV81D, lignée tumorale ovarienne, hétérozygote pour *BRCA2* établie dans le laboratoire de nos collaborateurs au CHUM [211, 219]. Cette lignée provient d'une tumeur ovarienne bien différenciée et très peu agressive qui n'a pas entraîné la mort du donneur, suggérant que l'hétérozygotie pour ce gène est capable de déterminer dans une cellule normale l'apparition d'un phénotype spécifique proche d'un phénotype tumoral.

De nombreuses études rapportent également un phénotype particulier chez les sujets porteurs d'une mutation de *BRCA1* ou *BRCA2*, dans les lymphocytes du sang périphérique (LSPs) ou dans des cellules lymphoblastoïdes (lymphoblastes immortalisés par EBV). Ce phénotype se traduit par une sensibilité accrue au rayonnement γ [220, 221], au peroxyde d'hydrogène [222-224] et aux mutagènes [225], mais pas aux rayons X [225]. Les LSPs des porteurs de mutation se distinguent par leur capacité réduite à la formation de colonies, une augmentation de la formation de micronoyaux, des défauts dans la réparation de l'ADN, une augmentation des cassures de la chromatine. Une fréquence élevée de déviation du ratio d'inactivation du chromosome X a été aussi rapportée dans les LSPs de patientes porteuses d'une mutation de *BRCA1* atteintes d'un cancer du sein [226] ou d'un cancer invasif de l'ovaire [227, 228]. Un groupe n'a pas retrouvé de différence entre les cellules lymphoblastoïdes de 41 femmes non-porteuses et 38 porteuses d'une mutation de *BRCA1*, infirmant la possibilité d'utiliser les analyses du biais d'inactivation de l'X pour prédire la présence d'une mutation germinale de *BRCA1* [229]. Cependant, l'utilisation de lignées lymphoblastoïdes au lieu de LSPs et le petit nombre d'échantillons pourraient être responsables de la discordance observée avec une seconde étude menée sur un plus grand nombre de patientes (735 non-porteuses, 120 porteuses d'une mutation de *BRCA1* et 24 de *BRCA2*). Cette dernière a finalement conclu que le ratio d'inactivation de l'X était plus élevé chez les porteuses d'une mutation de *BRCA1* et possiblement aussi de *BRCA2* que chez les non-porteuses et était de plus associée à une augmentation statistiquement significative de l'âge au diagnostic de cancer du sein et de l'ovaire [230].

Les données de la littérature suggèrent que les LSPs pourraient représenter un outil efficace pour l'étude des maladies complexes. En examinant l'ADN des porteurs de mutation germinale de *TP53* dans des familles présentant un syndrome de Li-Fraumeni, caractérisé par une augmentation du risque de cancers du sein précoce, de sarcomes, de tumeurs cérébrales et d'autres cancers, Malkin et al. ont montré qu'un nombre significatif de variations du nombre de copies de l'ADN (VNCs) étaient transmis entre les porteurs de mutations et ainsi identifié un grand nombre de régions génomiques incluant des gènes associés au cancer [231]. Cette approche pourrait être appliquée à une meilleure compréhension des mécanismes expliquant la variabilité de pénétrance des mutations de *BRCA1* ou *BRCA2*, survenue ou non de cancers, âge de diagnostic et surtout l'apparition de cancers affectant des sites distincts chez les porteurs de mutation. L'étude du transcriptome des LSPs permettrait également de mieux comprendre la survenue de maladies humaines. Des changements de profils d'expression ont ainsi été détectés dans les cellules lymphoblastoïdes d'hétérozygotes porteurs d'une mutation d'*ATM* [232]. Des découvertes récentes confirment l'intérêt de l'étude des LSPs des sujets porteurs de mutation. Non seulement les LSPs des porteurs de mutation de *BRCA1* ont un phénotype particulier comme nous l'avons montré auparavant : sensibilité accrue aux rayons γ et aux mutagènes, mais aussi une fréquence plus élevée qu'attendue de perte d'allèles au locus *GPA* dans les érythrocytes et au locus *HPRT* chez des femmes atteintes de cancer du sein avec mutation de *BRCA1*. Ce phénotype unique avec mutagenèse accrue est susceptible de contribuer au développement de ce cancer [233]. Avec une approche par PCR quantitative en temps réel spécifique d'allèle, Chen et al. ont d'ailleurs montré qu'un déséquilibre d'expression des transcrits alléliques de *BRCA1* et à un moindre degré de *BRCA2* dans les LSPs pouvait contribuer au développement de cancers du sein qu'ils soient ou non familiaux [234]. Toutes ces données confirment l'importance de l'étude du transcriptome chez les porteuses de mutation. De plus, elles suggèrent que les LSPs représenteraient un modèle intéressant pour disséquer les mécanismes contribuant au phénotype individuel des porteurs de mutations de *BRCA1* ou *BRCA2*. Finalement, nous pensons que les LSPs pourraient s'avérer utiles pour prédire le phénotype clinique dans l'avenir.

A.10. *BRCA1*, *BRCA2* et cellules NOSEs

Un phénotype particulier a été associé à l'hétérozygotie pour *BRCA1* et *BRCA2* dans les cellules NOSEs provenant de femmes ayant une histoire familiale de cancers (FH) du sein et/ou de l'ovaire [235]. Après plusieurs passages en cultures primaires, les cellules NOSEs provenant de ces femmes, acquièrent des caractéristiques similaires aux cellules tumorales ovariennes (TOV) : persistance d'une morphologie épithéliale [236], perte de leur capacité de réponse aux signaux permettant la transition épithélio-mésenchymateuse qui caractérise les cellules NOSEs non familiales (NF-NOSE) en culture. Un niveau d'expression élevé des deux marqueurs de différenciation des cellules épithéliales : l'E-cadherine et le CA125, une instabilité télomérique, ainsi que la coexpression des récepteurs MET et HGF ont aussi été décrits dans les cultures de cellules FH-NOSEs [235].

Des données récentes issues d'études protéomiques ont permis d'identifier un petit groupe de protéines différentiellement exprimées dans trois lignées de cellules épithéliales ovariennes obtenues après immortalisation de cellules F-NOSE provenant de donneuses hétérozygotes pour *BRCA1* [7]. Bien que l'existence d'une signature moléculaire associée à une mutation de *BRCA1* ait été mise en évidence par ce groupe, il faut noter que le matériel utilisé dans cette étude est constitué de cellules immortalisées, donc à un stade plus avancé de la carcinogenèse que les cellules non immortalisées que nous avons utilisées.

Toutes ces observations suggèrent qu'il existerait des événements moléculaires précoces associés à une mutation de *BRCA1* ou de *BRCA2* dans les cellules NOSEs. L'identification de ces événements permettrait d'améliorer nos connaissances sur la carcinogenèse ovarienne précoce, en particulier en ce qui concerne les partenaires des gènes *BRCA1* et/ou *BRCA2*.

A.11. Du transcriptome à l'ADN

Jusqu'à il y a peu, il était admis que le niveau d'expression d'un gène était déterminé d'une manière autonome par un ensemble d'éléments régulateurs (promoteurs, « *enhancers* » et « *silencers* ») et que le génome était composé d'une série de gènes contigus séparés les uns des autres par des éléments isolants agissant comme des frontières permettant à chaque gène d'échapper au contrôle des éléments régulateurs de ses voisins. Récemment, ce modèle génomique a été remis en cause par plusieurs études comparant d'une manière globale

l'expression des gènes en fonction de leur localisation génomique [237, 238]. Ces études ont montré que le génome est constitué de domaines de coexpression.

En effet, en utilisant des techniques génomiques à haut débit, Roy et al. [237] et Lercher et al. [238] ont retrouvé non seulement dans des microorganismes mais aussi chez l'homme des régions chromosomiques ordonnées le long du génome, contenant des gènes colocalisés et coexprimés correspondant à des domaines de coexpression. Ces régions sont sous le contrôle de régions appelées régions de contrôle global (RCG), régions très conservées contenant des agrégats d'éléments régulateurs « *enhancers* » capables de contrôler plusieurs gènes distincts à la fois dans leur structure mais aussi dans leur fonction. Les RCG sont parfois éloignées de centaines de kilo bases du groupe de gènes qu'elles contrôlent. Ainsi, Spitz et al. [239] ont réussi à identifier une RCG au locus murin complexe *HoxD* qui régule l'expression des gènes responsables du développement des membres. L'introduction de nouveaux gènes par des mécanismes de translocation dans une RCG entraînerait la diminution du niveau d'expression des gènes contrôlés par cette RCG [240]. Une mutation ponctuelle pourrait via les interactions entre gènes voisins contenus dans une RCG, déterminer des modifications dans le profil d'expression de la région coréglée qu'elle contrôle.

Ainsi, dans le cadre de ce projet, l'étude du transcriptome des NOSEs peut nous orienter vers des phénomènes d'épistasie entre gènes contigus ou dans la même voie métabolique. L'identification des régions chromosomiques comportant des gènes colocalisés et coréglés chez les porteurs de mutation sur l'un des deux gènes *BRCA1* et *BRCA2* nous ouvrira la voie dans un second temps vers l'identification d'haplotypes en déséquilibre de liaison avec ces régions, et plus avant vers l'identification d'haplotypes régulateurs. Une fois ces haplotypes définis, nous pourrions développer des outils diagnostics utilisant l'ADN à la place de l'ARN. Cette substitution nous permettrait ainsi de nous affranchir des contraintes liées à l'utilisation de l'ARN, en particulier des problèmes de stabilité.

A.12. Hypothèses

Les données de la littérature que nous venons d'exposer nous ont amenées à émettre les hypothèses suivantes :

1. les cellules NOSEs, cibles potentielles de la carcinogenèse associées aux gènes *BRCA1* et *BRCA2*, auraient un profil d'expression signature de la présence d'une mutation de ces gènes,
2. il existerait des régions chromosomiques comportant des gènes corégulés (domaines de corégulation) en association avec une mutation de *BRCA1* et *BRCA2*.

A.13. Objectifs

Afin de tester ces hypothèses, nous nous sommes fixées les objectifs suivants :

1. identifier les profils d'expression différentielle signatures de la présence d'une mutation des gènes *BRCA1* ou *BRCA2* dans les cellules NOSEs,
2. identifier les régions chromosomiques comportant des gènes corégulés (domaines de corégulation) en association avec une mutation sur *BRCA1*.

B. Matériel et Méthodes

B.1. Cultures primaires et échantillons d'ARN

Les études s'intéressant aux événements moléculaires précoces associés aux mutations de *BRCA1* et *BRCA2* précédant la transformation tumorale dans les cellules NOSEs n'étaient pas nombreuses jusqu'alors en raison de l'accessibilité réduite à ce type d'échantillon [241, 242]. Elles se sont limitées le plus souvent à l'étude d'un petit nombre de gènes du fait de la faible quantité de matériel disponible [235]. De plus, l'architecture monocouche de l'épithélium de surface de l'ovaire (OSE) limite le nombre de cellules disponibles pour l'extraction d'ARN. Une des manières de résoudre ce problème est de cultiver ces cellules. Notre groupe au sein du Réseau de Recherche sur le Cancer (RR-Cancer) du Fonds de la Recherche en Santé du Québec (FRSQ) a été confronté à la difficulté d'obtenir un nombre de passages en culture permettant d'extraire une quantité suffisante d'ARN.

Dans notre série initiale, les cellules NOSEs ont été recueillies par grattage de la surface du fragment d'ovaire normal fourni par le pathologiste immédiatement après l'exérèse chirurgicale réalisée dans le département de gynécologie-oncologique du CHUM, puis cultivées dans du milieu OSE supplémenté avec 15 % de sérum de veau foetal (SVF) [219, 243]. L'ARNm total était extrait avec du TRIzol™ (Gibco/IBRL, Life Technologies Inc., Burlington, ON, Canada) lorsque les cellules atteignaient 80 % de confluence, puis entreposé dans la banque de tissus du RR-Cancer du FRSQ. Tous les échantillons ont été collectés chez des donneuses qui avaient signé au préalable un consentement éclairé en accord avec les règles éthiques en vigueur dans notre établissement. Initialement, notre rendement de mise en culture n'était que de 50 %. Par la suite, notre groupe a établi les cultures primaires des cellules NOSEs avec un milieu OSE modifié supplémenté avec de l'EGF, de l'hydrocortisone, de l'insuline et de l'extrait pituitaire bovin [244], ce qui a augmenté considérablement notre rendement de mise en culture et nous a permis d'étendre notre série initiale. Un sous-groupe des cellules NOSEs a été établi en utilisant la procédure modifiée de Kruk (PMK) [245]. Les cellules TOVs, elles, ont été cultivées dans du milieu OSE avec 10 % de SVF [219, 243]. Ces cultures ont été réalisées grâce à une collaboration entre le Dr. C. Maugard et L. Portelance, technicienne de la banque du RR-Cancer du CHUM.

Notre série initiale comportait :

1. quatre échantillons d'ARN extraits de cellules NOSEs provenant de donneuses Canadiennes françaises soit n'ayant pas d'histoire familiale de cancers et non-porteuses d'une des huit mutations suivantes de *BRCA1* ou de *BRCA2* du panel dit « panel canadien français », *BRCA1* (2953delGTA1C, 3875delGTCT et 4446C>T) et *BRCA2* (2816insA, 3398delAAAAG, 6085G>T, 6503delTT et 8765delAG), soit ayant une histoire familiale de cancers mais non porteuses de la mutation identifiée dans leur famille [208]. Ces femmes ont été opérées pour des pathologies gynécologiques non néoplasiques (Tableau IV). Ces échantillons définissent la classe des NM.
2. deux échantillons d'ARN extraits de cellules NOSEs provenant de femmes ayant choisi de pratiquer une annexectomie bilatérale prophylactique, parce qu'elles avaient un risque élevé de développer un cancer de l'ovaire étant porteuses d'une mutation sur *BRCA1* (M1) (Tableau IV). Ces échantillons définissent la classe des M1.
3. trois échantillons d'ARN extraits de cellules NOSEs obtenues chez des donneuses ayant bénéficié d'une annexectomie bilatérale prophylactique, car elles avaient un risque élevé de développer un cancer de l'ovaire étant porteuses d'une mutation sur *BRCA2* (M2) (Tableau IV). Ces échantillons définissent la classe des M2.
4. trois échantillons d'ARN extraits de cellules TOVs de donneuses ayant une mutation sur *BRCA1* (TM1) (Tableau V). Ces échantillons définissent la classe des TM1.
5. un échantillon d'ARN extrait de cellules TOVs provenant d'une donneuse avec une mutation sur *BRCA2* (TM2) (Tableau V). Cet échantillon définit la classe des TM2.

Ce sont les échantillons de cette série initiale qui ont été utilisés pour générer les profils d'expression avec la biopuce HuGeneFL[®].

L'amélioration de nos méthodes de culture nous a permis d'étendre notre série initiale en incluant de nouveaux échantillons :

1. quatre échantillons provenant de donneuses non porteuses de mutation de *BRCA1* ni de *BRCA2*, classe des NM (Tableau IV),
2. un échantillon muté sur *BRCA1*, classe des M1 (Tableau IV),
3. quatre échantillons mutés sur *BRCA2*, classe des M2 (Tableau IV),

4. un échantillon provenant d'une femme atteinte d'un cancer de l'ovaire et porteuse d'une mutation sur *BRCA2*, classe des TM2 (Tableau V).

La biopuce HuGeneFL[®] n'étant plus disponible au Centre du Génome du fait de l'évolution des technologies bio-informatiques, nous n'avons pu réaliser l'analyse avec cette biopuce sur les échantillons supplémentaires inclus dans la série étendue, série qui nous a cependant permis d'effectuer la validation biologique de nos résultats.

Tableau IV : Caractéristiques des donneuses et des cellules NOSEs correspondantes.

| N° | Age | Donneuse | | Statut génétique | | | Caractéristique des cultures primaires | | | | Analyses effectuées |
|----|-----|--|--|---|-------------------------------|----------------|--|--|---------|----------------------|-------------------------------|
| | | Histoire personnelle et/ou familiale de cancer | Motif d'intervention / Diagnostic chirurgical | Porteuse de mutation sur <i>BRCA1</i> ou <i>BRCA2</i> | Mutation familiale identifiée | N° Echantillon | Milieu de culture | Cellules d'origine et latéralité de l'ovaire | Passage | Méthodes | |
| 1 | 54 | Non | Léiomyome (Muqueuse utérine) Kyste séreux (Ovaire D) | Non | **NA | 31 | OSE | NOSE G | p11 | | q-RT-PCR |
| 2 | 26 | Non | Endosalpingiose Kyste paratubaire | Non | NA | 61 | OSE | NOSE D | p7 | | q-RT-PCR |
| 3 | 51 | Non | Léiomyome (Muqueuse utérine) Corps lutéal (Ovaire G) | Non | NA | 116 | OSE | NOSE G | p13 | §§ BP (HuGeneFL®) | q-RT-PCR |
| 4 | 46 | *HF | Léiomyome (Muqueuse utérine) Kyste folliculaire (Ovaire G) | Non | <i>BRCA1</i> C4446T | 390 | OSE | NOSE §I | p8 | | q-RT-PCR |
| 5 | 37 | HF | Dysplasie cervicale | Non | <i>BRCA1</i> 2953del3+C | 504 | OSE | NOSE D | p8 | | BP (HuGeneFL®) q-RT-PCR |
| 6 | 43 | HF | Léiomyome (Muqueuse utérine) Endométriose (Trompe G) | Non | <i>BRCA2</i> G6085T | 900 | OSE | NOSE G | p8 | | BP (HuGeneFL®) |
| 7 | 49 | Non | Léiomyome (Muqueuse utérine) Kyste bénin fonctionnel (Ovaire D) | Non | NA | 910 | OSE | NOSE G | p5 | | BP (HuGeneFL®) q-RT-PCR |
| 8 | 50 | Non | Léiomyome (Muqueuse utérine) | Non | NA | 2457 | †PMK/ ‡MM | NOSE D | p3 IXT | | q-RT-PCR |
| 9 | 44 | HF | Annexectomie prophylactique Kyste bénin séreux (Ovaire G) | Oui | <i>BRCA1</i> 2953del3+C | 319 | OSE | I NOSE | p9 | | BP (HuGeneFL®) q-RT-PCR |
| 10 | 39 | HF | Annexectomie prophylactique Dysplasie cervicale | Oui | <i>BRCA1</i> C4446T | 653 | OSE | NOSE G | p8 | | BP (HuGeneFL®) q-RT-PCR |

Tableau IV (suite) :

| | | | | | | | | | | |
|----|----|-------------------------------|--|-----|---------------------------|------|------------|--------|--------|-------------------------------|
| 11 | 42 | HF | Annexectomie prophylactique Kyste folliculaire (Ovaire D) Kyste du corps lutéal (Ovaire G) | Oui | <i>BRCA1</i> 3790ins4 | 2039 | PMK/ MM | NOSE D | p2 2XT | q-RT-PCR |
| 12 | 50 | Côlon (47) Sein (49) HF | Annexectomie prophylactique Léiomyome (Muqueuse utérine) | Oui | <i>BRCA2</i> G6085T | 848 | OSE | NOSE D | p4 | BP (HuGeneFL®) q-RT-PCR |
| 13 | 49 | HF | Annexectomie prophylactique Léiomyome (Muqueuse utérine) | Oui | <i>BRCA2</i> 8765de1AG | 1181 | OSE | NOSE I | p6 | BP (HuGeneFL®) q-RT-PCR |
| 14 | 43 | HF | Annexectomie prophylactique Léiomyome (Muqueuse utérine) Kyste du corps lutéal (Ovaire D) | Oui | <i>BRCA2</i> G6085T | 1275 | OSE | NOSE G | p4 | BP (HuGeneFL®) q-RT-PCR |
| 15 | 45 | HF | Annexectomie prophylactique | Oui | <i>BRCA2</i> 8765de1AG | 2016 | PMK/ MM | NOSE D | p4 1XT | q-RT-PCR |
| 16 | 45 | HF | Annexectomie prophylactique | Oui | <i>BRCA2</i> 8765de1AG | 2016 | OSE | NOSE G | p4 | q-RT-PCR |
| 17 | 52 | HF | Annexectomie prophylactique Léiomyome (Muqueuse utérine) | Oui | <i>BRCA2</i> 8765de1AG | 2064 | MM | NOSE D | p2 | q-RT-PCR |
| 18 | 50 | HF | Annexectomie prophylactique Endométriose (Ovaire D) | Oui | <i>BRCA2</i> G6085T | 2075 | PMK/ MM | NOSE G | p6 8XT | q-RT-PCR |

HF** = Histoire familiale de cancer, *NA** = Non applicable, †**PMK** = procédure modifiée de Kruk, ‡**MM** = milieu modifié,
§**I** = NOSE dont le site ovarien d'origine est inconnu et §§**BP** = biopuce.

Tableau V : Caractéristiques des donneuses et des TOV's correspondants

| N° | Histoire médicale et diagnostic chirurgical du sujet | | | Statut génétique | | Caractéristiques des cultures | | | Analyses | | | |
|----|--|-------------------------------|--------------------------------------|------------------|-------|-------------------------------|------------------------|--------|----------------------------|-------------------|---------|--------------------------------|
| | Age à la chirurgie | Histoire familiale de cancers | Diagnostic chirurgical | Stade | Grade | Porteuse | Mutation familiale | Nom | Type de culture et origine | Milieu de culture | Passage | Méthodes |
| 19 | 41 | Oui | Cystadénocarcinome séreux | IIIC | 3 | Oui | <i>BRCA1</i> C4446T | EP730 | *CP ‡Ep (‡TOV) | OSE | p11 | §BP (HuGeneFL®) q-RT-PCR |
| 20 | 60 | Oui | Cystadénocarcinome séreux | IIIC | 3 | Oui | <i>BRCA1</i> C4446T | OV747 | CP Ascitis (TOV) | OSE | p7 | BP (HuGeneFL®) q-RT-PCR |
| 21 | 57 | Oui | Cystadénocarcinome Endométrioïde | IV | 3 | Oui | <i>BRCA1</i> C4446T | OV893 | CP Ascite (TOV) | OSE | p13 | BP (HuGeneFL®) q-RT-PCR |
| 22 | 56 | Oui | Cystadénocarcinome papillaire séreux | IV | 3 | Oui | <i>BRCA2</i> 8765delAG | OV552 | CP Ascite (TOV) | OSE | P8 | BP (HuGeneFL®) q-RT-PCR |
| 23 | 66 | Oui | Cystadénocarcinome | IIIC | 1 | Oui | <i>BRCA2</i> 8765delAG | T0V81D | Lignée cellulaire (TOV) | OSE | p14 | q-RT-PCR |

*CP = Culture Primaire, ‡Ep = Epiploon, ‡TOV = Tumeur ovarienne, §BP = biopuce.

B.2. Identification du transcriptome des F-NOSEs et des TOVs

Les biopuces ont représenté une véritable révolution technologique dans le domaine de la transcriptomique, en permettant de mesurer l'expression des gènes à l'échelle du génome entier, alors qu'auparavant les méthodes existantes se limitaient à la mesure de l'expression d'un nombre restreint de gènes. Les profils d'expression correspondant à nos échantillons d'ARN de la série initiale ont été générés avec la biopuce HuGeneFL[®]. Le traitement de l'ARN, l'hybridation des biopuces et la détection des signaux ont été réalisés au Centre d'Innovation de Génome Québec (Montréal, QC, Canada) en suivant la méthodologie décrite dans [246]. Les données d'expression ont été analysées sur la plateforme bio-informatique Bioconductor, plateforme ouverte permettant le développement de logiciels pour l'analyse et la compréhension de données génomiques, et utilisant le langage de programmation R [247]. Les signaux bruts ont été recalculés (soustraction du bruit de fond, normalisation et transformation en \log_2) avec le module RMA « *Robust Multiarray Average expression measure* » [248, 249]. Les profils d'expression différentielle ont été identifiés avec le module LIMMA « *Linear Model for Microarrays Data* » [250] après correction des valeurs de p avec la correction de Benjamini et Hochberg [251].

B.2.1. Biopuce HuGeneFL[®]

Il s'agit d'une puce à ADN. Des séquences de 25 nucléotides monocaténares appelées sondes sont synthétisées in situ directement sur la puce utilisant des techniques alliant la photolithographie et la chimie combinatoire. Au départ, tous les sites de fixation des sondes sont protégés par un groupement photolabile. En utilisant un masque photosensible, la stimulation lumineuse ciblée permet d'enlever ce groupement et de fixer un nouveau nucléotide protégé par un groupement photolabile. Le procédé est successivement renouvelé avec les différentes bases jusqu'à la synthèse complète de la sonde requise [252]. Des millions de copies de la même sonde se trouvent dans un espace de $24 \mu\text{m}^2$ sur la puce, formant une cellule de sondes.

Deux types de sondes appariées ont été utilisés pour fabriquer cette biopuce :

- des sondes s'hybridant avec les séquences de référence appelées sondes PM « *Perfect Match* » ou appariement parfait, car elles s'hybrident parfaitement avec la séquence cible à quantifier (Fig. 10).

- des sondes associées aux sondes PM, mais qui en diffèrent par une seule base située au milieu de la séquence nucléotidique (Fig. 10). Ces sondes sont appelées sondes MM « *MissMatch* » ou mésappariement).

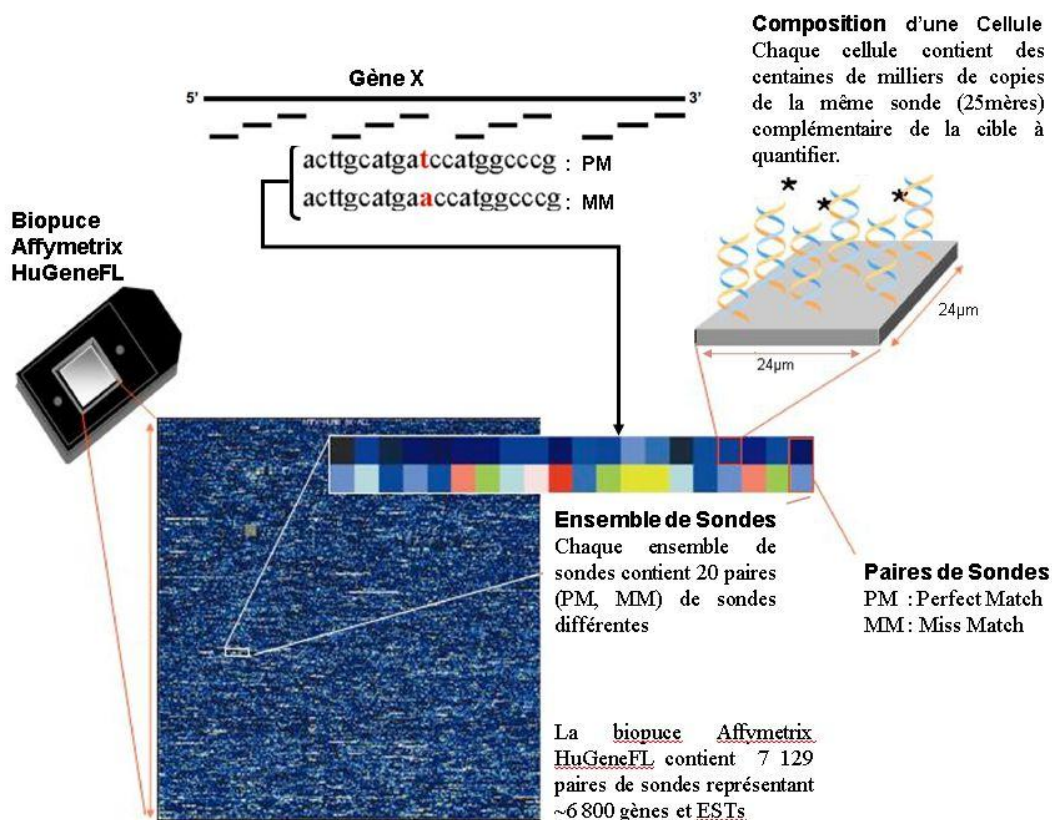


Fig. 10 : Représentation de la biopuce HuGeneFL[®] de la plateforme Affymetrix[™].
Figure adaptée d'après www.affymetrix.com).

Une paire de sondes est formée d'une cellule PM et d'une cellule MM correspondante. Vingt de ces paires de sondes forment un ensemble de sondes (ES) qui permet de quantifier l'expression d'un gène. Dans le cas de la biopuce HuGeneFL[®], les différentes paires de sondes constituant un ES sont distribuées sur l'ensemble de la séquence du gène. La biopuce HuGeneFL[®] comporte 7 129 ensembles de 20 paires de sondes correspondant à ~ 6 800 gènes bien identifiés [253]. Dans la suite de ce document, nous désignerons par valeur PM et valeur MM les niveaux d'expression respectifs de la sonde PM et de la sonde MM mesurés par la biopuce HuGeneFL[®].

B.2.2. Traitement des échantillons

L'ADN complémentaire (ADNc) a été synthétisé à partir de 20 µg d'ARN extrait de chaque échantillon de la série initiale en utilisant le promoteur oligo-dT contenant le site de dimérisation de l'ARN polymérase T7. L'ARN complémentaire (ARNc) produit par la transcription in vitro de l'ADNc a été biotinylé par incorporation de dUTPs et dCTPs marqués avec de la Biotine, puis hydrolysé en fragments par incubation à 95 °C dans un tampon composé de 40 mM de Tris-Acétate, 100 mM de potassium acétate et 30 mM de MgCl₂ ajusté à un pH de 8.1 afin de réduire sa structure secondaire. Enfin l'hybridation avec la biopuce HuGeneFL[®] d'Affymetrix[™] a été réalisée avec 15 µg d'ARNc. Après lavage et marquage par le conjugué streptavidine-phycoérythrine selon le protocole proposé par Affymetrix[™], la lecture des biopuces a été effectuée avec un scanner Hewlett-Packard commercialisé par Affymetrix[™] qui génère une image où chaque cellule est représentée par des pixels dont l'intensité est fonction de la quantité d'ARN hybridée avec chaque sonde oligonucléotidique.

B.2.3. Plateforme Bioconductor

La plateforme Bioconductor a été mise en place et associée au langage de programmation statistique R en 2001. Il s'agit d'une initiative collaborative entre statisticiens, mathématiciens, biologistes et informaticiens visant à créer des solutions informatiques (algorithmes, logiciels) pour l'analyse et la compréhension des données génomiques [254]. Ce projet a permis le développement et l'utilisation de logiciels innovants à large échelle, la réduction des barrières limitant la recherche scientifique interdisciplinaire et la génération de résultats de recherches reproductibles.

Cette source ouverte fournit l'accès à un éventail de méthodes statistiques et d'outils de représentation graphique puissants permettant l'analyse des données génomiques. De plus, elle facilite l'intégration des métadonnées biologiques dans l'analyse des données expérimentales comme les données de la littérature provenant de PubMed et les données d'annotation provenant de LocusLink. Elle représente aussi une plateforme commune permettant le développement et l'application de logiciels évolutifs et inter-opérationnels. Les outils de Bioconductor sont distribués sous forme de modules développés principalement avec le langage R. Au départ, la plupart de ces modules étaient des outils bio-informatiques consacrés uniquement à l'analyse des données issues des biopuces à ADN. Par la suite, la plateforme a étendu son application pour inclure l'analyse de données

génomiques de toute nature (SAGE « *Serial Analysis of Gene Expression* », séquences, données SNPs) générées avec différents types de biopuces. De ce fait, la plateforme Bioconductor est devenue un outil d'analyse incontournable dans le domaine de la bio-informatique.

B.2.4. Procédures de normalisation

La normalisation est la première étape à effectuer dans une analyse de données de biopuces. Elle vise à rendre comparables les mesures des niveaux d'expression des gènes quantifiés dans chacun des échantillons par une biopuce différente. Elle applique une série de transformations soit aux valeurs PM et MM, soit aux seules valeurs PM obtenues avec chaque biopuce, en fonction de la procédure utilisée. Trois étapes successives vont être appliquées : l'ajustement du bruit de fond, la normalisation des données et, finalement, le calcul de l'expression des gènes.

Plusieurs méthodes de normalisation sont disponibles. Nous allons nous limiter à la description des méthodes « *Robust Multi-array Average* » (RMA) [248, 255] et « *Guanine Cytosine Robust Multi-Array Average* » (GCRMA) [256] que nous avons toutes deux appliquées afin de choisir la méthode fournissant la meilleure normalisation de nos données d'expression générées avec la biopuce HuGeneFL[®].

B.2.4.1. Normalisation avec la procédure « *Robust Multi-array Average* »

Correction du bruit de fond [255] : Le bruit de fond est une mesure de l'intensité du signal causé par les hybridations non spécifiques et par l'autofluorescence de la surface de la biopuce elle-même [257]. La correction du bruit de fond est une étape essentielle dans l'opération de normalisation. Elle vise à estimer le bruit de fond de chaque biopuce puis à le soustraire des valeurs PM. L'intensité observée des sondes PM (IO_{PM}) est modélisée sous forme de la somme d'un signal (S) et d'un bruit de fond (BF) (Fig. 11).

$$IO_{PM} = S + BF$$

S et BF représentent deux variables aléatoires et indépendantes : S suit une distribution exponentielle $Exp(\lambda)$ de moyenne $1/\lambda$ et de variance égale à $1/\lambda^2$. BF suit une distribution normale $N^+(\mu, \sigma^2)$ de moyenne μ et de variance σ^2 . N^+ indique que la distribution est tronquée à gauche du zéro. Les paramètres μ , λ et σ de chaque biopuce sont estimés en ajustant le modèle illustré ci-dessous pour les valeurs IO_{PM} .

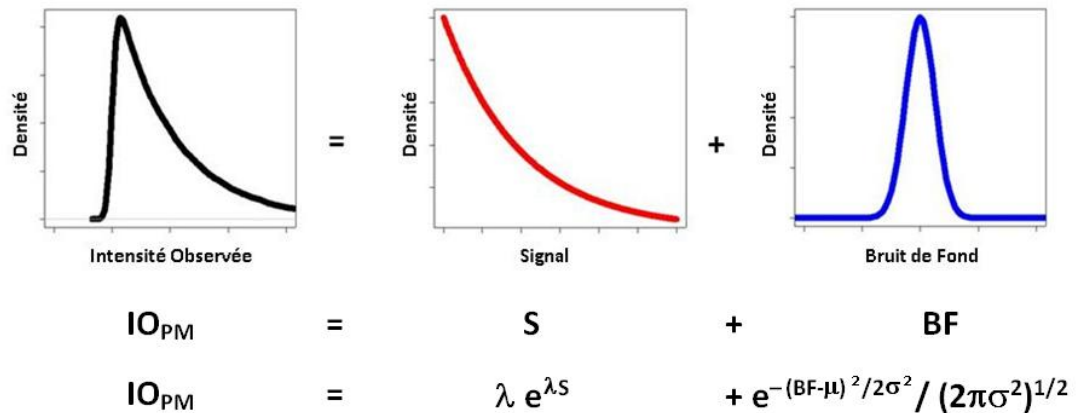


Fig. 11 : Modélisation de l'intensité observée pour les sondes PM (IO_{PM}) en fonction du signal (S) et du bruit de fond (BF).

Nous définissons par $\varphi(x)$ la fonction de la distribution normale standard (de moyenne 0 et variance 1) et par $\phi(x)$ sa fonction de densité. Nous désignons par PM^* les valeurs PM obtenues après ajustement du BF. Les valeurs PM^* sont calculées par l'équation suivante :

$$PM^* = E(S|IO_{PM}) = IO_{PM} - \mu - \lambda\sigma^2 + \sigma \cdot \frac{\varphi\left(\frac{IO_{PM} - \mu - \lambda\sigma^2}{\sigma}\right) - \varphi\left(\frac{\mu + \lambda\sigma^2}{\sigma}\right)}{\phi\left(\frac{IO_{PM} - \mu - \lambda\sigma^2}{\sigma}\right) + \phi\left(\frac{\mu + \lambda\sigma^2}{\sigma}\right) - 1}$$

↘
BF

Normalisation des données proprement dite par la méthode des quantiles [255] : Cette opération vise à minimiser les variabilités techniques accumulées pendant les différentes étapes de l'expérience afin de rendre les données des biopuces comparables entre elles. Elle consiste à transformer les valeurs PM^* de façon à amener chaque biopuce à avoir la même distribution de valeurs PM^* . Cette transformation est justifiée sur la base du postulat suivant : **Les distributions des niveaux d'expression des gènes sont proches les unes des autres dans les différents échantillons étudiés.** Sur le plan biologique, ceci revient à supposer que peu de gènes sont différentiellement exprimés dans les différentes expériences et lorsque leur expression est dérégulée, elle l'est aussi bien positivement que négativement [249].

Définissons par X la matrice des valeurs PM^* quantifiées par les biopuces de sorte que les colonnes correspondent aux biopuces et les lignes correspondent aux ES. La méthode des quantiles consiste à :

1. trier chaque colonne de la matrice X pour obtenir X_1 ,
2. calculer les moyennes à travers toutes les lignes de X_1 et remplacer chaque élément de la ligne par la moyenne calculée pour avoir X_2 .

réarranger chaque colonne de X_2 de sorte à avoir l'ordre original de X .

La matrice résultante contient les valeurs PM^* normalisées qui seront désignées par $N(PM^*)$ dans les sections qui suivent.

Calcul de l'expression des gènes [248] : Cette opération vise à attribuer une valeur représentative de l'expression de chaque gène sur la biopuce à l'ES mesurant le transcrite de ce gène. En théorie, les 20 valeurs PM du même ES devraient avoir des intensités identiques puisqu'elles examinent l'expression du même gène. Or en pratique, ce n'est pas le cas. En effet, les sondes appartenant au même ES s'apparient avec différentes régions de l'ARNc qui correspondent à différents exons. A cause du phénomène d'épissage alternatif, ces sondes peuvent ainsi s'apparier avec différents transcrits ayant des niveaux d'expression distincts. De plus, la longueur des sondes (25 bases), la faible affinité de certaines sondes, en particulier celles riches en GC, causeraient des problèmes d'hybridation croisée [258]. Ceci se traduit par l'appariement partiel d'une sonde à l'ARN cible, l'ARN correspondant à d'autres gènes étant capable également de s'apparier à la même sonde ou à des sondes appartenant au même ES.

La procédure RMA transforme les valeurs $N(PM^*)$ issues de l'étape précédente en \log_2 afin de rendre leur distribution proche d'une distribution normale. Les valeurs en \log_2 ainsi calculées seront désignées par Y . Pour calculer l'expression des gènes, cette procédure suppose que les valeurs Y suivent un modèle linéaire additif représenté par l'équation suivante :

$$Y_{ijn} = \mu_{in} + \alpha_{jn} + \varepsilon_{ijn} \quad , i = 1, \dots, I \quad , j = 1, \dots, J \text{ et } n = 1, \dots, N$$

avec, pour chaque ES n , Y représentant les valeurs en \log_2 de $N(PM^*)$, μ_{in} le niveau d'expression de l'ES n que nous cherchons à calculer en \log_2 sur la biopuce i , (μ ici représente le niveau d'expression exprimé en \log_2 de l'ES que nous cherchons à calculer à partir de ses 20 sondes). α_{jn} l'effet de l'affinité d'une sonde j donnée, ε_{ijn} un terme d'erreur indépendant distribué autour d'une moyenne de zéro. La procédure RMA suppose que la

somme des α_j est égale à zéro ($\sum_j \alpha_j = 0$) pour toutes les sondes. Ceci revient à supposer qu’Affymetrix™ a choisi des sondes qui sont en moyenne représentatives de l’expression des gènes à mesurer.

La procédure RMA utilise la méthode robuste « *Median Polish* » [259] pour estimer les paramètres du modèle linéaire additif, en suivant les étapes suivantes :

1. pour chaque ES, définir par Y la matrice constituée des valeurs Y_{ij} de la sonde j sur la biopuce i .
2. chercher la médiane de chaque ligne de la matrice Y et la soustraire des éléments de la ligne à laquelle elle correspond,
3. chercher la médiane de chaque colonne de la matrice résultant de l’étape précédente et la soustraire des éléments de la colonne à laquelle elle correspond,
4. itérer les étapes 2 et 3 jusqu’à ce que les médianes des lignes et colonnes deviennent nulles,
5. soustraire la dernière matrice obtenue de la matrice du départ Y pour obtenir les valeurs corrigées,

la moyenne des valeurs corrigées des sondes appartenant à l’ES en question va fournir finalement la mesure de l’expression correspondant à cet ES.

B.2.4.2. Normalisation avec la procédure « Guanine Cytosine Robust Multi-array Average »

La procédure GCRMA est une version améliorée de RMA utilisant la même méthode de normalisation et de calcul de l’expression des gènes. Elle en diffère par la méthode de correction du bruit de fond [256]. Au contraire de la procédure RMA qui utilise seulement les valeurs PM pour ajuster le bruit de fond, la procédure GCRMA, elle, utilise également les valeurs MM ainsi que des informations sur la séquence des sondes pour estimer l’affinité de celles-ci. Ceci permet d’estimer le bruit causé par l’hybridation non spécifique (HNS). La procédure GCRMA modélise les valeurs PM et MM comme la somme d’un bruit optique (O) introduit par le scanneur lors de la lecture de la biopuce, d’un bruit (N) causé par l’HNS et d’un signal (S) selon les équations suivantes :

$$\begin{aligned} PM &= O_{PM} + N_{PM} + S \\ MM &= O_{MM} + N_{MM} + \varphi S \quad ; \quad \varphi \in [0, 1] \end{aligned}$$

Le paramètre φ est employé pour tenir compte des sondes MM qui détectent un vrai signal à la place d’un bruit de fond. La procédure GCRMA suppose que : O suit une distribution

log-normale, $\log(S)$ suit une fonction exponentielle et $\log(N_{PM})$ et $\log(N_{MM})$ suivent une distribution bivariable normale de moyennes respectives μ_{PM} et μ_{MM} , de variances égales à σ^2 donnant une corrélation de 0,7. Les moyennes μ_{PM} et μ_{MM} suivent une fonction de lissage (h) de l'affinité des sondes PM et MM. Cette affinité, désignée par α , tient compte des bases b (A, T, G ou C) de la sonde et de leur position le long de la séquence de la sonde. En effet, l'affinité d'une sonde est modélisée comme une somme d'effets (μ) dépendant de la position des bases dans cette séquence :

$$\alpha = \sum_{k=1}^{25} \sum_{j \in \{A, T, G, C\}} \mu_{j,k} 1_{b_k=j} \quad ; \quad 1 \leq k \leq 25$$

avec k qui indique la position de la base dans la séquence de la sonde, b_k qui représente la base se trouvant à la position k et j l'une des quatre bases A, G, C ou T. $1_{b_k=j}$ est une variable booléenne égale à 1 si la $k^{\text{ème}}$ base est j et 0 dans le cas contraire et $\mu_{j,k}$ représente l'effet de la base j à la position k . L'effet d'une base est estimé en utilisant une fonction de lissage à 5 degrés de liberté (Fig. 12).

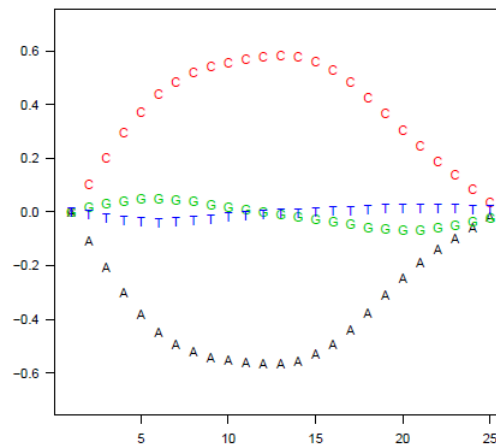


Fig. 12 : Effet d'une abase en fonction de sa position dans la séquence de la sonde.

L'abscisse correspond aux 25 positions de la base dans la séquence de la sonde et l'ordonnée représente l'effet de cette base. Figure tirée de [256].

La procédure GCRMA considère le bruit optique (O) comme une constante et elle l'estime à l'intensité minimale observée dans chaque biopuce parce que la variance du bruit optique est négligeable comparée à celle du bruit causé par l'HNS. Pour estimer les paramètres μ_{PM} , μ_{MM} et σ^2 , la procédure GCRMA ajuste une courbe loess reliant les valeurs MM

après correction pour le bruit optique ($\log(\text{MM}-\hat{O}$; \hat{O} = estimation du bruit de fond), à l'affinité des sondes MM (αMM). Enfin, la procédure GCRMA calcule l'intensité du signal recherchée comme la moyenne de l'appariement spécifique calculée *a posteriori* en utilisant les valeurs PM et MM observées ainsi que les paramètres estimés.

B.2.5. Analyse supervisée avec la méthode « Linear Model for Microarrays Data »

B.2.5.1. Analyse différentielle

Afin d'identifier les profils d'expression différentielle associés à une mutation de l'un des gènes *BRCA1* ou *BRCA2* dans les cellules NOSEs et TOVs, nous avons effectué une analyse supervisée en utilisant la méthode « *Linear Model for Microarrays Data* » (LIMMA) [250]. Une analyse est décrite comme étant supervisée si elle utilise des données connues *a priori* comme des paramètres de classification. Dans cette analyse, nous connaissons le phénotype tumoral ou non tumoral des cellules dont proviennent nos échantillons ainsi que le statut génétique des donneuses. Ces différents paramètres nous ont permis de répartir nos échantillons entre cinq classes : la classe des échantillons non tumoraux provenant de donneuses non mutées appelée NM, les classes des échantillons non tumoraux provenant de donneuses avec une mutation soit sur *BRCA1*, soit sur *BRCA2* appelées respectivement M1 et M2, enfin les classes correspondant aux échantillons tumoraux ayant une mutation sur *BRCA1* ou sur *BRCA2* appelées respectivement TM1 et TM2.

Notre analyse statistique vise à identifier les ES différentiellement exprimés (ESDE) entre les différentes classes préalablement définies. Pour ce faire, nous effectuons un test d'hypothèse nous permettant d'inférer c'est-à-dire d'accepter ou de rejeter l'hypothèse qu'un ES est différentiellement exprimé. Cette hypothèse appelée hypothèse nulle est représentée sous forme de contraste c'est-à-dire d'une combinaison linéaire de paramètres correspondant aux échantillons des classes comparées. Nous avons identifié les ESDEs associés à une mutation de *BRCA1* (contraste NM vs M1) ou de *BRCA2* (contraste NM vs M2) dans les cellules NOSEs, les ES ayant une expression différentielle entre les classes M1 et M2 (contraste M1 vs M2), les ES différentiellement exprimés au cours de la transformation tumorale associée à une mutation de *BRCA1* (contraste M1 vs TM1), enfin les ESDEs entre la classe M des échantillons non tumoraux mutés sur *BRCA1* ou sur

BRCA2 ($M = M1 \cup M2$) et la classe TM des échantillons tumoraux mutés sur *BRCA1* ou *BRCA2* ($TM = TM1 \cup TM2$) (contraste M vs TM).

LIMMA est une approche qui utilise un modèle linéaire afin d'identifier les ESDEs après quantification par la méthode des biopuces. Elle permet l'analyse d'expériences complexes impliquant simultanément différents ARN cibles aussi facilement qu'une expérience simple impliquant des répliquats du même ARN. Cette approche requiert une matrice de design représentant les différents ARN cibles hybridés sur les différentes biopuces et optionnellement une matrice de contrastes. La matrice de design définit les coefficients de régression du modèle linéaire étudié. Elle peut être définie soit suivant l'approche de paramétrage moyennes-groupes « *groups-mean parametrization* » ou bien avec l'approche de paramétrage contrastes-traitements « *treatment-contrasts parametrization* ». Selon la première définition, chaque coefficient correspond à une classe d'échantillons, alors que selon la deuxième, la matrice de design contient deux types de coefficients : un correspondant à la classe des échantillons contrôles et les autres coefficients décrivant les contrastes d'intérêts.

Dans notre étude, nous avons choisi d'utiliser l'approche de paramétrage groupes-moyennes afin de définir notre matrice de design. Nous avons ainsi cinq coefficients, chacun pour une des classes NM, M1, M2, TM1 et TM2 définies préalablement. La matrice de contrastes permet de combiner ces coefficients en contrastes d'intérêt. Chaque contraste correspond à une comparaison des coefficients des ARN cibles. Si la matrice de contrastes n'est pas définie, le module LIMMA calculera par défaut toutes les comparaisons possibles entre les différentes classes d'échantillons. Si elle est définie, ceci permet soit d'exclure certaines comparaisons, soit de rajouter d'autres comparaisons complexes tel que, dans notre étude, le contraste M vs TM.

La première étape consiste à ajuster un modèle linéaire en utilisant la fonction `lmFit()` qui modélise entièrement la partie systématique des données d'expression. Chaque ligne de la matrice de design correspond à une biopuce de l'expérience et chaque colonne correspond à un coefficient. Avec des données générées avec des puces à ADN, le nombre de coefficients est égal au nombre de classes d'échantillons. Le modèle linéaire appliqué à chaque ES est illustré par l'équation $\mathbf{E}[Y_j] = \mathbf{X} \alpha_j$, avec Y_j qui correspond au vecteur contenant les niveaux d'expression d'un ES j pour chaque biopuce i calculés après normalisation par RMA, et α_j qui représente le vecteur des coefficients. Finalement \mathbf{X} correspond à la matrice

de design qui définit le design de l'expérience, en précisant quel échantillon d'ARN a été hybridé sur quelle biopuce. Un des objectifs de cette étape est d'estimer la variabilité des données.

Ainsi, pour cette analyse, nous avons :

$$Y_j = \begin{bmatrix} Y_{NM-3} \\ Y_{NM-5} \\ Y_{NM-6} \\ Y_{NM-7} \\ Y_{M1-9} \\ Y_{M1-10} \\ Y_{M2-12} \\ Y_{M2-13} \\ Y_{M2-14} \\ Y_{TM1-19} \\ Y_{TM1-20} \\ Y_{TM1-21} \\ Y_{TM2-22} \end{bmatrix} \quad X = \begin{array}{c} \begin{array}{ccccc} \text{NM} & \text{M1} & \text{M2} & \text{TM1} & \text{TM2} \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \quad \alpha_j = \begin{bmatrix} \alpha_{NM} \\ \alpha_{M1} \\ \alpha_{M2} \\ \alpha_{TM1} \\ \alpha_{TM2} \end{bmatrix}$$

L'étape suivante correspond à établir la matrice des contrastes. Celle-ci vise à comparer les coefficients correspondant aux classes impliquées dans les contrastes étudiés et ceci en utilisant l'équation suivante $\beta_j = C^T \alpha_j$, avec β_j qui représente l'estimateur des contrastes étudiés, C^T représente la transposée de la matrice des contrastes C et α_j le vecteur des coefficients. Si la matrice de design a été définie en utilisant l'approche de paramétrage moyennes-groupes, α_j représente la moyenne du niveau d'expression en \log_2 de la classe i . Si la matrice de design a été définie suivant l'approche de paramétrage contrastes-traitements, le premier coefficient (α_1) représente la moyenne du niveau d'expression en \log_2 de la première classe (classe contenant les échantillons contrôles) et les autres α_j correspondent aux différences entre la moyenne des niveaux d'expression des échantillons contrôles (α_1) et celle des échantillons traitements. α_j est le vecteur des coefficients représentant les valeurs estimées des moyennes ou des «*fold change*» (variabilité biologique) en \log_2 du niveau d'expression des gènes des échantillons appartenant aux classes étudiées. Ces coefficients sont obtenues par ajustement du modèle linéaire.

Ceci nous donne pour cette analyse,

$$C = \begin{array}{c} \begin{array}{ccccc} \text{M1-NM} & \text{M2-NM} & \text{M2-M1} & \text{TM1-M1} & \text{TM-M} \end{array} \\ \begin{bmatrix} -1 & -1 & 0 & 0 & 0 \\ 1 & 0 & -1 & -1 & -0.5 \\ 0 & 1 & 1 & 0 & -0.5 \\ 0 & 0 & 0 & 1 & 0.5 \\ 0 & 0 & 0 & 0 & 0.5 \end{bmatrix} \begin{array}{l} \text{NM} \\ \text{M1} \\ \text{M2} \\ \text{TM1} \\ \text{TM2} \end{array} \end{array}$$

$$\Rightarrow \beta_j = \begin{bmatrix} \alpha_{M1} - \alpha_{NM} \\ \alpha_{M2} - \alpha_{NM} \\ \alpha_{M2} - \alpha_{M1} \\ \alpha_{TM1} - \alpha_{M1} \\ \alpha_{TM1} + \alpha_{TM2} - \alpha_{M1} - \alpha_{M2} \end{bmatrix}$$

La matrice β_j détermine les contrastes que nous désirons effectuer entre les classes. Or chacun de ces contrastes n'est autre que notre hypothèse nulle dans LIMMA. En effet, nous voulons tester si la matrice β_j est égale à la matrice nulle, en d'autres termes tester l'égalité des coefficients entre eux ($\alpha_{M1} = \alpha_{NM}$, $\alpha_{M2} = \alpha_{NM}$, $\alpha_{M1} = \alpha_{M2}$, $\alpha_{TM1} = \alpha_{M1}$ et $\alpha_{M1} + \alpha_{M2} = \alpha_{TM1} + \alpha_{TM2}$) et ceci pour chaque ES.

Pour réaliser l'analyse différentielle, LIMMA fournit les fonctions `topTable` et `decideTable` qui résument les résultats du modèle linéaire, effectuent les tests, et ajustent les p en tenant compte des comparaisons multiples. Les résultats sont fournis en terme de « *fold change* », erreur standard et p.

Le test statistique utilisé est un t modéré calculé pour chaque ES et chaque contraste. Il représente le ratio d'une valeur M ou valeur du contraste qui représente le \log_2 de la variation biologique entre deux conditions expérimentales ou bien le \log_2 des niveaux d'expression des ES sur l'écart type de ces niveaux d'expression [250]. Un t modéré possède la même interprétation qu'un t de Student à l'exception près que les écarts types sont modérés sur l'ensemble des ES, empruntant effectivement l'information à partir de l'ensemble des ES et les réduisant à une valeur commune, permettant d'établir un modèle bayésien simple pour inférer l'hypothèse nulle pour chaque ES. Cette méthode permet d'éviter d'estimer *a priori* le nombre des ESDEs.

Le t de Student n'est pas adéquat ici parce qu'il peut générer un t élevé en valeur absolue en raison d'un faible écart type et non pas en raison d'une différence significative d'expression.

Le test B représente le « *log-odds* » c'est-à-dire le log du risque relatif approché qu'un ES est différentiellement exprimé. Ainsi, si $B = 1,5$, le risque relatif approché d'avoir une expression différentielle est $\exp(1,5) = 4,48$. Ainsi, la probabilité qu'un ES soit différentiellement exprimé est de 82 % ($4,48 / (4,48 + 1) = 0,82$). Le test B est automatiquement ajusté pour le problème des tests multiples en supposant que la proportion attendue des ES différentiellement exprimés est de 1 % (ou un autre pourcentage décidé *a priori* par l'expérimentateur). Un test B égal à zéro signifie qu'il y a 50 % de probabilité que l'ES soit différentiellement exprimé et autant de chances qu'il ne le soit pas. Plus le B est élevé, plus le résultat est significatif.

Le test B, le test t et le p (probabilité) sont calculés en utilisant le module LIMMA. Après calcul du t, LIMMA va déterminer la significativité statistique de chaque expression différentielle, en lui attribuant une probabilité p. Le p, appelé aussi erreur α , représente la probabilité de rejeter à tort l'hypothèse nulle alors que nous aurions dû l'accepter. Il représente dans notre cas la probabilité de déclarer un ES comme étant significativement exprimé alors qu'il ne l'est pas. Après calcul des p, LIMMA doit les ajuster pour tenir compte du fait que nous effectuons simultanément de multiples tests. Nous y reviendrons dans la section suivante. Les p ajustés ainsi calculés avec LIMMA seront notés q.

B.2.5.2. Correction des valeurs de p avec la procédure « *False Discovery Rate* » de Benjamini et Hochberg

Avec les biopuces, les analyses statistiques présentent une particularité supplémentaire. En effet, nous testons l'hypothèse nulle sur un grand nombre d'ES simultanément (7 129). Cette situation rejoint le problème des comparaisons multiples. Le calcul du p doit être ajusté, sinon la significativité serait grandement sur-estimé, car il faut tenir compte du fait que nos 7 129 tests s'appliquent aux mêmes échantillons et ne sont donc pas indépendants. Le risque d'erreur lorsque nous nous adressons à des comparaisons multiples est donc très supérieur à celui qui est considéré lorsque nous considérons une comparaison unique. La question est de déterminer le nombre de faux positifs résultant de comparaisons effectuées indépendamment les unes des autres tout en tenant compte de la multiplicité des comparaisons effectuées simultanément. Il s'agit alors de définir le nombre de faux positifs correspondant à une valeur donnée du p ajusté et ainsi de définir le seuil acceptable qui nous donnera une significativité suffisante avec le nombre de faux positifs le plus bas.

Des procédures doivent être appliquées pour corriger les valeurs de p en fonction du nombre de tests effectués. Nous avons utilisé la procédure de Benjamini et Hochberg (B&H) [251] qui contrôle le taux de fausses découvertes ou taux de faux positifs, « *False Discovery Rate* » (FDR), soit en fait le taux d'hypothèses nulles rejetées à tort. Elle consiste à ordonner les gènes en fonction de leur p. La valeur de p la plus grande reste telle qu'elle. Les suivantes sont ajustées successivement par ordre décroissant en multipliant chaque valeur p_i par n/i , où n désigne le nombre total d'ES et i le rang de l'ES dans la liste d'ES étudiés. La procédure va déclarer différentiellement exprimés les ES qui donnent une valeur de p ajustée inférieure au seuil de significativité que nous avons fixé. La correction devient plus stricte au fur et à mesure que la valeur de p diminue. Nous avons calculé le FDR, taux de faux positifs parmi les ES déclarés par LIMMA comme étant différentiellement

exprimés. Le FDR retenu est celui qui donne le meilleur compromis entre un p qui permet de sélectionner un nombre suffisant de candidats avec un nombre restreint de faux positifs. Nous avons finalement retenu un FDR de 0,05 pour l'ensemble des contrastes à l'exception du contraste entre échantillons mutés et tumoraux, pour lequel le FDR retenu était de 0,2. Il est à noter que ce contraste correspond à la comparaison de la moyenne des niveaux d'expression des gènes exprimés dans les échantillons non-tumoraux mutés sur *BRCA1* ou *BRCA2* à celle des échantillons tumoraux mutés sur *BRCA1* ou *BRCA2*, chaque classe contenant des échantillons mutés sur l'un ou l'autre de ces gènes. Par conséquent, la variation des niveaux d'expression des gènes exprimés dans les échantillons de chacune des deux classes en question est élevée. De ce fait, la différence entre leur moyenne devient faible. La détection d'une expression différentielle entre ces deux classes nécessite donc le choix d'un seuil élevé. Quoiqu'il en soit, ceci témoigne du fait que ces deux gènes appartiennent à des voies de carcinogenèse différentes sur le plan biologique, rendant très hétérogène chacune des classes comparées.

B.3. Méthodes de Validation

Pour valider la pertinence des gènes candidats sélectionnés avec LIMMA à partir des données d'expression générées avec la biopuce HuGeneFL[®], nous avons utilisé deux types de méthodes : une méthode bio-informatique et une méthode biologique.

B.3.1. Agrégation hiérarchique ascendante

Nous voulions vérifier si le niveau d'expression des ESDEs identifiés avec LIMMA était capable de reclasser chacun de nos échantillons dans la classe appropriée. En utilisant la fonction *hClust* du module *Stats*, nous avons effectué une analyse d'agrégation hiérarchique ascendante sur les ESDEs pour chacun des contrastes examinés.

Une agrégation hiérarchique ascendante procède à l'agglomération progressive des groupes en agrégeant deux à deux les groupes les plus proches. Nous désignerons ici par groupe à la fois les échantillons à classer et les regroupements d'échantillons générés par l'algorithme d'agglomération progressive utilisé. Cet algorithme est défini par deux facteurs, soit le critère de regroupement des groupes et la métrique de calcul de distances entre deux groupes. Dans cette analyse, nous avons utilisé le critère de regroupement du saut maximal sur des distances euclidiennes (Fig. 13A) [260]. Selon ce critère, la distance entre deux

groupes est égale à la plus grande distance entre les échantillons du premier et ceux du deuxième groupe (Fig. 13A).

L'algorithme d'agglomération progressive considère, au départ, chaque échantillon comme un groupe. Par la suite, l'algorithme calcule les distances entre les différents groupes selon le critère de regroupement du saut maximal. Il agrège les deux groupes les plus proches c'est-à-dire ayant entre eux la plus petite distance et répète les deux étapes précédentes jusqu'à ce qu'il n'y ait plus de groupe à agréger. A la fin de cette procédure, nous obtenons une partition hiérarchique représentée sous forme d'une arborescence appelée dendrogramme (Fig. 13B), indiquant la structure de similitude de nos échantillons.

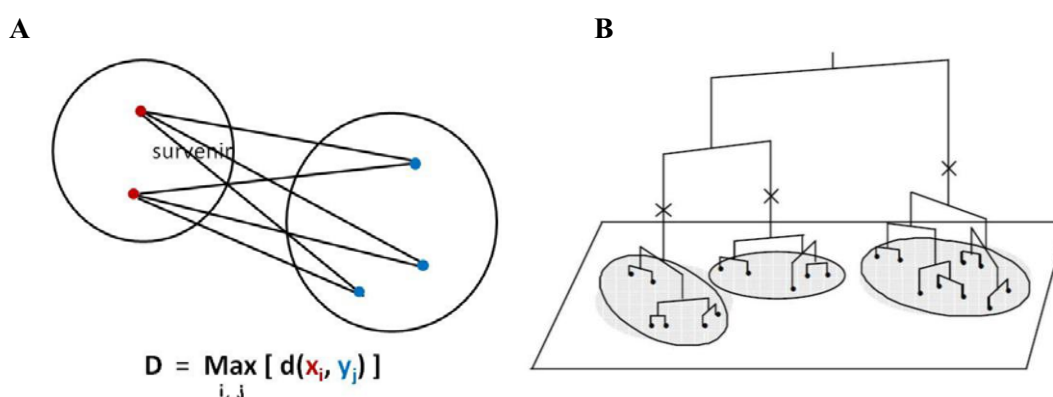


Fig. 13 : Agrégation hiérarchique ascendante

A) Calcul de la distance entre deux groupes en utilisant le critère de regroupement du saut maximal. $d(x_i, y_j)$ représente la distance entre le $i^{\text{ème}}$ échantillon (x_i) du premier groupe et le $j^{\text{ème}}$ échantillon (y_j) du deuxième groupe.

B) Dendrogramme résultant d'un algorithme d'agrégation hiérarchique d'après [261].

B.3.2. Quantification de quatre gènes candidats par PCR quantitative en temps réel

Afin de vérifier que les ESDEs identifiés par LIMMA correspondent à de vrais positifs, c'est-à-dire qu'ils présentent réellement une expression différentielle, nous avons quantifié le niveau d'expression de quatre gènes candidats par PCR quantitative en temps réel (q-RT-PCR). Nous avons sélectionné ces gènes pour leur fonction biologique significative dans la carcinogenèse ainsi que pour la significativité de la différence d'expression calculée par LIMMA ($q \leq 0.05$) pour au moins l'un des contrastes étudiés. Cette partie du travail a été effectuée par C. Abaji dans le laboratoire du Dr Maugard et pour le gène *MPI* par C. Arous. La quantification d'autres gènes candidats pour lesquels nous avons obtenu une différence

significative se poursuit dans le laboratoire ainsi que les quantifications spécifiques d'allèles correspondant aux mutations sur *BRCA1*.

Le profil d'expression obtenu avec les biopuces pour les transcrits de ces quatre gènes candidats, a été validé par q-RT-PCR, avec le Rotor-Gene™ 3000 (*Corbett Research, Sydney, NSW, Australia*). Nous avons effectué un cycle de décontamination (50°C, 2') puis un cycle de dénaturation (95°C, 10'), enfin 45 cycles d'amplification de l'ADNc (95°C, 30'', 60°C, 40''). Les amorces ainsi que les sondes TaqMan® utilisées pour quantifier les niveaux d'expression de chacun de nos quatre gènes candidats ont été dessinées sur le site internet *Primer3* [262], de telle sorte que l'amplicon cible chevauche deux exons et corresponde à une région du(des) transcrit(s) du gène reconnu par l'ES de la biopuce HuGeneFL®. La spécificité de reconnaissance pour cette région a été vérifiée avec Blast (NCBI) et Blat (<http://genome.ucsc.edu/>). Les sondes ont été synthétisées par IDT™ (*Integrated DNA Technologies, Coralville, IA*) en utilisant la chimie TaqMan®. Les sondes reconnaissant les gènes d'intérêt (gènes cibles) ont été couplées avec le fluorochrome émetteur (reporter) 6-FAM, celles reconnaissant le gène normaliseur (gène de référence) ont été couplées avec l'émetteur HEX, ce qui a permis de les co-amplifier en duplex. Le BHQ-1 « *black hole quencher-1* » a été utilisé comme « *quencher* » pour neutraliser la fluorescence avant amplification. Toutes les réactions ont été effectuées avec le kit « *Platinum Quantitative PCR SuperMix-UDG* (*Invitrogen Canada Inc., Burlington, ON*). Nous avons choisi le *TBP* « *Tata-binding protein* » comme gène de référence car son niveau d'expression n'a pas montré de variation significative dans les TOVs en q-RT-PCR [87], ni dans les NOSEs et TOVs dans nos propres données de biopuces. Nos données de q-RT-PCR ont confirmé qu'il s'agissait d'un gène de référence approprié. Les réactions ont été optimisées pour obtenir la meilleure efficacité et une bonne reproductibilité des résultats. Une courbe standard a été réalisée avec des dilutions sériées d'un pool d'ADNc obtenu à partir de l'ARN extrait soit de la lignée cellulaire ovarienne TOV112D, soit de la lignée OV90, pour chacun des gènes d'intérêt coamplifié en duplex avec le gène normaliseur *TBP* excepté pour *MPI*. Ce dernier a été amplifié séparément de *TBP* mais dans la même expérience car il n'a pas été possible d'optimiser la réaction de PCR en duplex. Un échantillon calibreur a été introduit dans chaque expérience nous permettant de comparer les échantillons quantifiés dans différentes expériences. Nous avons utilisé la dilution initiale de la courbe standard qui sert à étalonner chaque expérience. Le ratio gène cible / gène de référence a été calculé avec l'algorithme de Pfaffl [263] :

$$R = \frac{(E_{\text{gène cible}})^{\Delta CT_{\text{gène cible}} (\text{échantillon d'intérêt} - \text{échantillon contrôle})}}{(E_{\text{gène de référence}})^{\Delta CT_{\text{gène de référence}} (\text{échantillon d'intérêt} - \text{échantillon contrôle})}}$$

avec R qui représente le ratio relatif de l'expression des gènes cible et de référence, E l'efficacité de la PCR pour un gène donné et ΔCT la différence entre le CT (cycle à partir duquel l'intensité de la fluorescence émise sort du bruit de fond) d'un échantillon cible et d'un échantillon contrôle qui sert de calibre dans les expériences successives réalisées. Les mesures ont été effectuées sur deux échantillons d'ADNc indépendants et chaque point a été mesuré au minimum en duplicata, au mieux quand cela était possible en triplicata voire dans le cas de *MPI* en quadriplicata.

B.3.3. Quantification de *BRCA1* et *BRCA2* par PCR quantitative en temps réel

Les sondes TaqMan[®] utilisées pour quantifier les niveaux d'expression de *BRCA1* et de *BRCA2* ont été conçues indépendamment de la région codante correspondant aux ES qui mesurent le niveau d'expression de *BRCA1* et *BRCA2* sur la HuGeneFL[®], puisqu'aucun des ES de la HuGeneFL[®] ne donnait d'expression différentielle dans aucun des contrastes étudiés. Nous avons donc choisi pour chaque gène une zone cible qui était présente à la fois dans la séquence de référence mais aussi dans l'ensemble des transcrits correspondants aux différents variants d'épissage connus de ce gène, ceci nous permettant a priori de quantifier l'ensemble des transcrits du gène. Par ailleurs, nous avons également réalisé une quantification spécifique des allèles sauvages et mutés de *BRCA2* avec des sondes reconnaissant spécifiquement le transcrit correspondant à l'allèle sauvage et à l'allèle muté et ce, pour chacune des deux mutations de *BRCA2* présentes dans notre échantillon.

B.3.4. Quantification spécifique d'allèles

Pour vérifier la présence et l'expression des allèles muté et sauvage, nous avons quantifié chacun des transcrits alléliques de *BRCA2* (sauvage et muté) par q-RT-PCR dans les échantillons d'ARN non tumoraux non mutés et mutés des porteuses de mutation G6085T et 8765delAG de *BRCA2* ainsi que dans les échantillons tumoraux des porteuses de la mutation 8765delAG de *BRCA2*. Nous ne disposons malheureusement d'aucun échantillon de tumeurs ovariennes de patientes porteuses de la mutation G6085T. Comme précédemment, chaque expérience a été réalisée en duplicata sur deux ADNc indépendants. Les courbes standards ont été réalisées à partir de dilutions sériées de la bande purifiée du produit de PCR d'un échantillon hétérozygote pour la mutation de *BRCA2* étudiée. Pour la

mutation *BRCA2*-8765delAG, nous avons utilisé un échantillon provenant de la lignée cellulaire TOV81D comme calibre. Pour la mutation *BRCA2*-G6085T, nous ne disposons pas d'une tumeur ovarienne hétérozygote pour cette mutation pour servir de calibre, c'est pourquoi nous avons utilisé une tumeur mammaire provenant d'une femme hétérozygote pour la mutation et nous avons démontré par séquençage l'hétérozygotie de ces cellules tumorales. Cette partie du travail a été effectuée par C. Abaji dans le laboratoire du Dr Maugard.

Pour la mutation 8765delAG de *BRCA2*, la spécificité de discrimination allélique a été obtenue en ajoutant une petite quantité de la sonde sauvage au mélange de réaction utilisé pour quantifier l'allèle muté. La quantification de l'allèle sauvage n'étant pas possible en compétition, elle a été réalisée séparément de celle de l'allèle muté. Pour la mutation G6085T de *BRCA2*, la quantification des allèles sauvage et muté a été effectuée en compétition. Dans les deux cas, nous avons utilisé le gène *TBP* comme gène de référence. Il a été pour sa part quantifié séparément.

B.4. Localisation des régions comportant des gènes corégulés chez les porteuses d'une mutation de *BRCA1* ou de *BRCA2*

Afin d'identifier les régions enrichies en gènes différenciellement exprimés en présence d'une mutation du gène *BRCA1*, nous avons effectué une étude exploratoire de différentes méthodes de localisation de régions comportant des gènes colocalisés et corégulés. Nous avons finalement choisi d'utiliser la fonction `geneSetTest` du module LIMMA [250]. Cette approche repose sur le test U de Mann-Whitney-Wilcoxon [264]. Il s'agit de vérifier si un ensemble de gènes sélectionnés selon des critères établis *a priori* est enrichi en gènes différenciellement exprimés. Elle repose sur l'hypothèse nulle suivante : la probabilité que les gènes de l'ensemble testé soient différenciellement exprimés n'est pas plus élevée que celle d'un groupe de gènes sélectionnés au hasard. Ainsi, lorsqu'une région est enrichie en gènes différenciellement exprimés, soient surexprimés soient sous-exprimés, l'hypothèse nulle sera rejetée avec un risque d'erreur fourni par le p calculé par `geneSetTest`. L'utilisation de ce test suppose que les gènes soient indépendants les uns des autres, ce postulat n'est réalisable si le nombre des ES testés est faible ce qui est le cas ici [265].

Nous avons appliqué `geneSetTest` sur les résultats générés par LIMMA pour le contraste entre échantillons non tumoraux non mutés et échantillons non tumoraux mutés sur *BRCA1*. Dans le but d'éliminer les données qui peuvent représenter une source de bruit supplémentaire ou bien nuire à l'identification des régions coréglées, une étape de prétraitement de données a été nécessaire. Nous avons exclu les ES dont les données d'annotation sont incomplètes (EntrezID, chromosome, localisation chromosomique), ainsi que les ES ayant plus d'une localisation chromosomique connue. Lorsque plusieurs ES quantifiaient l'expression du même gène, nous avons conservé l'ES présentant la valeur d'expression différentielle la plus élevée dans le contraste étudié. Les ES obtenus après prétraitement des données ont été classés en fonction du chromosome sur lequel ils sont localisés. Nous les avons ensuite ordonnés sur chacun des brins (sens ou anti-sens) en fonction de leurs coordonnées en bases. Par la suite, nous avons balayé chaque brin chromosomique avec des fenêtres incluant 10 ES contigus. Nous avons retenu les régions qui donnaient un p calculé par le `geneSetTest` inférieur à 0,01.

C. Résultats

C.1. Normalisation des données

Nous avons appliqué les procédures RMA et GCRMA sur les niveaux d'expression générés avec la biopuce HuGeneFL[®] sur nos échantillons afin de choisir la procédure qui fournit une meilleure normalisation. Nous avons représenté ci-dessous les résultats des différentes méthodes de normalisation sous forme de boîtes à moustache (Fig. 14).

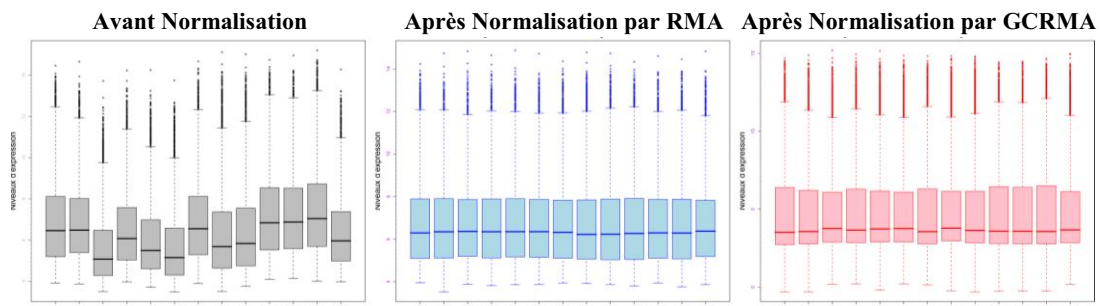


Fig. 14 : Boîtes à moustache représentant les niveaux d'expression des gènes mesurés par la HuGeneFL[®] dans chacun des échantillons d'ARN avant (graphique de gauche), après normalisation par RMA (graphique du milieu) et après normalisation par GCRMA (graphique de droite).

Une boîte à moustache, pour une biopuce donnée, représente la variance des expressions de l'ensemble des gènes analysés par cette biopuce; la petite ligne horizontale au centre de chaque boîte à moustache correspond à la médiane de l'expression des gènes (percentile 50) et le haut et le bas de chaque boîte, respectivement les percentiles 75 et 25. La hauteur de la boîte représente l'écart inter-quantile. Les extrémités inférieure et supérieure des segments représentent respectivement les percentiles 5 et 95. Les points au-dessus du percentile 95 correspondent aux gènes ayant des niveaux d'expression aberrants (fréquemment appelés dans la littérature *outliers*). L'écart inter-quantile ainsi que la médiane de l'expression des gènes pour chaque biopuce sont plus variables après normalisation avec la procédure GCRMA qu'avec le RMA. De plus, la procédure GCRMA génère davantage de valeurs aberrantes que la procédure RMA. C'est pourquoi, nous avons choisi d'utiliser RMA pour normaliser nos données d'expression.

C.2. Profil d'expression différentielle associé à une mutation sur *BRCA1* ou *BRCA2* dans les cellules NOSEs et les TOVs

Nous avons comparé les profils moléculaires entre NOSEs mutées sur *BRCA1* (F-NOSE1) et NOSEs non mutées (NF-NOSE) (Fig. 15A), NOSEs mutées sur *BRCA2* (F-NOSE2) et NOSEs non mutées (NF-NOSE) (Fig. 15B), entre NOSEs mutées sur *BRCA2* (F-NOSE2) et NOSEs mutées sur *BRCA1* (F-NOSE1) (Fig. 15C), entre TOVs mutées sur *BRCA1* (F-TOV1) et NOSEs mutées sur *BRCA1* (F-NOSE1) (Fig. 15D) et finalement entre TOVs mutées sur *BRCA1* ou *BRCA2* (F-TOV1-2) et NOSEs mutées sur *BRCA1* ou sur *BRCA2* (F-NOSE1-2) (données non présentées).

Les résultats de l'analyse différentielle effectuée avec LIMMA pour ces contrastes sont représentés sous forme de « volcano plot » (Fig. 15, graphiques de gauche). Un volcano plot représente la relation entre le Log Odds (B), risque relatif approché des niveaux d'expression différentielle exprimé en \log_2 (axe des ordonnées) et le \log_2 du « *fold change* » (axe des abscisses). B correspond à la probabilité que chaque ES soit différentiellement exprimé. Plus B est élevé, plus la probabilité que cet ensemble de sondes différentiellement exprimé (ESDE) soit un vrai positif est grande. Chaque ES est représenté par un point. Les ESDE ayant un $q \leq 0.05$ sont représentés par des points rouges. Ceux avec un « *fold change* » positif correspondent aux ES surexprimés, ceux ayant un « *fold change* » négatif correspondent aux ES sous-exprimés dans cette classe. Cinq ESDEs (correspondant à trois gènes distincts) ont été identifiés dans le contraste F-NOSE1 vs NF-NOSE (Fig. 15A), 35 ESDEs dans le contraste F-NOSE2 vs NF-NOSE (Fig. 15B). Cent trente-trois ESDEs ont été identifiés dans le contraste F-NOSE1 vs F-NOSE2 (Fig. 15C), confirmant le fait que *BRCA1* et *BRCA2* appartiennent à des voies de carcinogenèse différentes. Enfin, 791 ES ont été associés à la transformation tumorale des NOSEs en TOVs chez les porteuses de mutation sur *BRCA1* (Fig. 15D).

C.3. Agrégation hiérarchique

Nous avons effectué une agrégation hiérarchique sur les ESDEs identifiés avec LIMMA pour chacun des contrastes étudiés. Les résultats de cette agrégation hiérarchique effectuée avec la fonction hClust sont représentés d'une part sous forme de dendrogramme illustrant la hiérarchie selon laquelle les échantillons sont regroupés (Fig. 15, graphiques de droite) et

d'autre part de « Heat Map » illustrant le niveau d'expression des ESDEs (Fig. 15, graphiques de gauche). Chaque colonne du « Heat Map » correspond à un échantillon et chaque ligne correspond à un ES. La couleur de chaque cellule représente le niveau d'expression de l'ES de l'échantillon correspondant. Le rouge indique une surexpression et le vert indique une sous-expression. Plus la couleur est foncée, plus le transcrit correspondant est exprimé.

Les ES identifiés avec LIMMA sont capables de reclasser les échantillons dans la classe appropriée, et ce pour chacun des contrastes examinés à l'exception des ESDEs du contraste F-TOV1-2 vs F-NOSE1-2 (données non présentées). Ceci n'est pas surprenant puisque l'identification des ESDEs devrait garder les ES qui séparent le mieux ces classes. Bien entendu, les résultats de l'agrégation hiérarchique des ESDEs de ce contraste exclue cette analyse des ensembles des validations. Il est à noter que les dendrogrammes correspondant aux contrastes NM vs M2 (Fig. 15B, graphique de droite) et M1 vs M2 (Fig. 15C, graphique de droite) montrent que les échantillons M2-13 et M2-14 s'agrègent ensemble bien que les donneuses de ces échantillons soient porteuses de mutations différentes sur *BRCA2* (Tableau IV). L'échantillon M2-12, lui est classé dans une branche différente de celle des échantillons M2-13 et M2-14, bien que la donneuse de M2-12 présente la même mutation que celle de M2-14 (Tableau IV), (Fig. 15B et 15C, graphiques de droite).

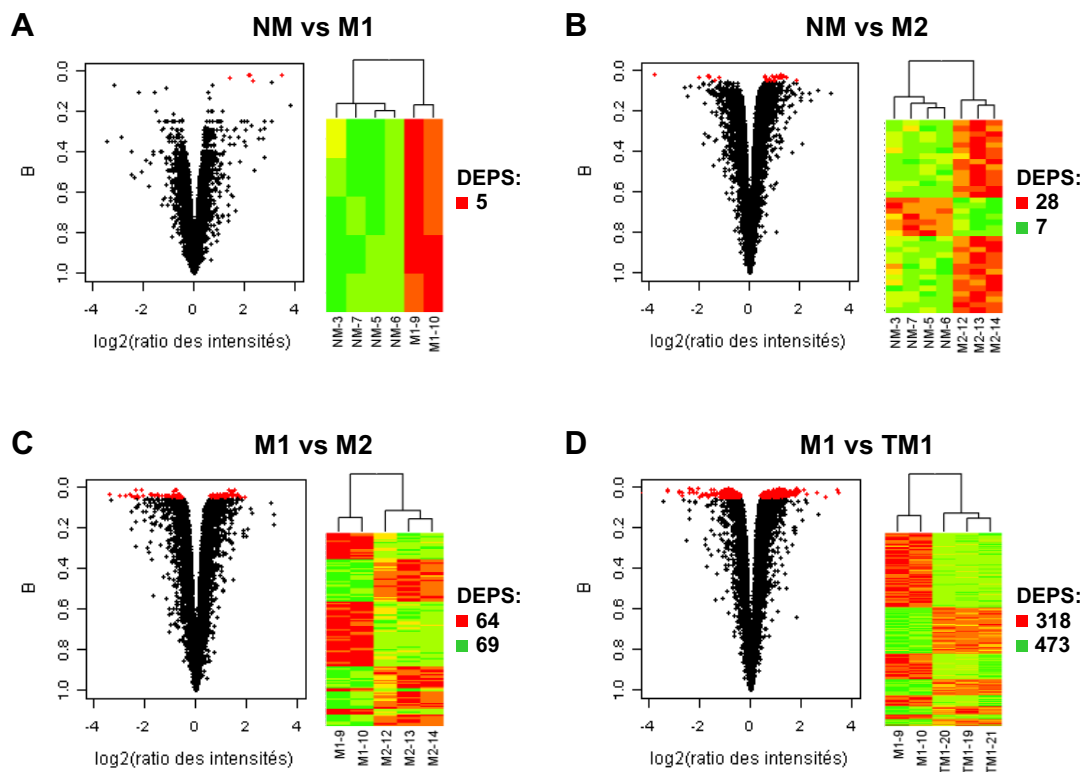


Fig. 15 : Résultats de LIMMA et des agrégations hiérarchiques correspondantes sous forme, respectivement, de « Volcano Plot » (graphiques de gauche) et de « Heat Map » (graphiques de droite).

C.4. Quantification par PCR quantitative en temps réel des quatre gènes candidats et de *BRCA1* et *BRCA2* dans la série initiale

De façon à valider d'un point de vue biologique les résultats obtenus avec les méthodes bio-informatiques, nous avons confirmé par q-RT-PCR les profils d'expression différentielle des quatre gènes candidats que nous avons sélectionné tel que précisé plus haut. La validation initiale a été limitée aux échantillons pour lesquels la quantification avait été faite avec la biopuce HuGeneFL[®] permettant ainsi de comparer les profils obtenus avec les deux méthodes de quantification.

Nous avons observé en général une bonne concordance entre les profils d'expression observés avec la biopuce HuGeneFL[®] et ceux obtenus avec la quantification q-RT-PCR pour les quatre gènes candidats (Fig. 16).

Les transcrits du gène « *Serum amyloid A2* » (*SAA2*), correspondant aux ES X51441_at et X51441_s_at, sont significativement surexprimés dans les F-NOSE1 comparés aux NF-NOSEs (Tableau VI et Fig. 16A). Les transcrits de *SAA2*, correspondant aux ES J03474_at et X51441_at, sont significativement sous-exprimés dans les F-NOSE2 comparés aux F-NOSE1 (Tableau VI et Fig. 16A). De plus, les transcrits correspondant aux trois ES J03474_at, X51441_at et X51441_s_at (données non illustrées) sont significativement sous-exprimés dans les F-TOV1 comparés aux F-NOSE1 (Tableau VI et Fig. 16A). Le transcrit du « *Filamin A interacting protein 1-like* » (*FILIP1L*) correspondant à l'ES U53445_at est significativement sous-exprimé dans les F-NOSE2 comparés aux NF-NOSEs (Tableau VI et Fig. 16B). Le transcrit de l'« *Annexin A8-like 2* » (*ANXA8L2*) correspondant à l'ES X16662_at est significativement surexprimé dans les F-TOV1 vs. F-NOSE1 (Tableau VI et Fig. 16C). Finalement, le transcrit du « *Mannose Phosphate Isomerase* » (*MPI*) mesuré par l'ES X76057_at est significativement surexprimé dans les F-NOSE2 comparé aux NF-NOSEs et aux F-NOSE1 (Tableau VI et Fig. 16D). Son expression est significativement plus élevée dans les F-TOV1 comparée aux F-NOSE1 (Tableau VI), mais cette surexpression dans les F-TOV1 n'est pas confirmée par q-RT-PCR (Fig. 16D).

La comparaison des niveaux d'expression de *BRCA1*, de *BRCA2* et des quatre gènes candidats suggère que *SAA2* pourrait être coréglulé avec *BRCA2* dans les échantillons F-NOSE2 de la série initiale (échantillons M2-12, M2-13 et M2-14) (Fig. 16 et 17). En effet, les niveaux d'expression de *BRCA2* quantifiés par q-RT-PCR sont faibles dans les échantillons F-NOSE2 à l'exception de l'échantillon M2-12 (Fig. 17B). Le niveau élevé d'expression de *BRCA2* dans cet échantillon comparé aux autres F-NOSE2 (M2-13 et M2-14) (Fig. 17B) s'accompagne d'un niveau élevé d'expression de *SAA2* (Fig. 16A). Nous avons également observé une bonne concordance entre les niveaux d'expression de *BRCA2* (Fig. 17B), *BRCA1* (Fig. 17A) et *FILIP1L* (Fig. 16B) dans les échantillons F-NOSE2 de la série initiale (Fig. 16B et 17). Cependant, cette concordance n'a pas été observée pour les niveaux d'expression d'*ANXA8L2* et *MPI*.

Tableau VI : Résumé des résultats de l'analyse différentielle avec LIMMA pour chacun des quatre gènes candidats sélectionnés.

| Contraste | ES (Affymetrix) | Symbole du gène (HUGO) | Description du gène (HUGO) | t | q (p ajusté) |
|-----------|--------------------|------------------------------|---|-------|-----------------|
| NM vs M1 | X51441_at | <i>SAA2</i> | <i>Serum Amyloid A2</i> | 8,25 | 0,017 |
| | X51441_s_at | <i>SAA2</i> | <i>Serum Amyloid A2</i> | 6,76 | 0,045 |
| NM vs M2 | U53445_at | <i>FILIP1L</i> | <i>Filamin A Interacting Protein 1-like</i> | -5,84 | 0,034 |
| | X76057_at | <i>MPI</i> | <i>Mannose Phosphate Isomerase</i> | 7,11 | 0,026 |
| M1 vs M2 | J03474_at | <i>SAA2</i> | <i>Serum Amyloid A2</i> | -4,64 | 0,044 |
| | X51441_at | <i>SAA2</i> | <i>Serum Amyloid A2</i> | -5,88 | 0,031 |
| | X76057_at | <i>MPI</i> | <i>Mannose Phosphate Isomerase</i> | 8,48 | 0,011 |
| M1 vs TM1 | J03474_at | <i>SAA2</i> | <i>Serum Amyloid A2</i> | -5,77 | 0,017 |
| | X51441_at | <i>SAA2</i> | <i>Serum Amyloid A2</i> | -7,44 | 0,010 |
| | X51441_s_at | <i>SAA2</i> | <i>Serum Amyloid A2</i> | -6,51 | 0,013 |
| | X16662_at | <i>ANXA8L2</i> | <i>Annexin A8-Like 2</i> | 4,84 | 0,024 |
| | X76057_at | <i>MPI</i> | <i>Mannose Phosphate Isomerase</i> | 6,56 | 0,013 |

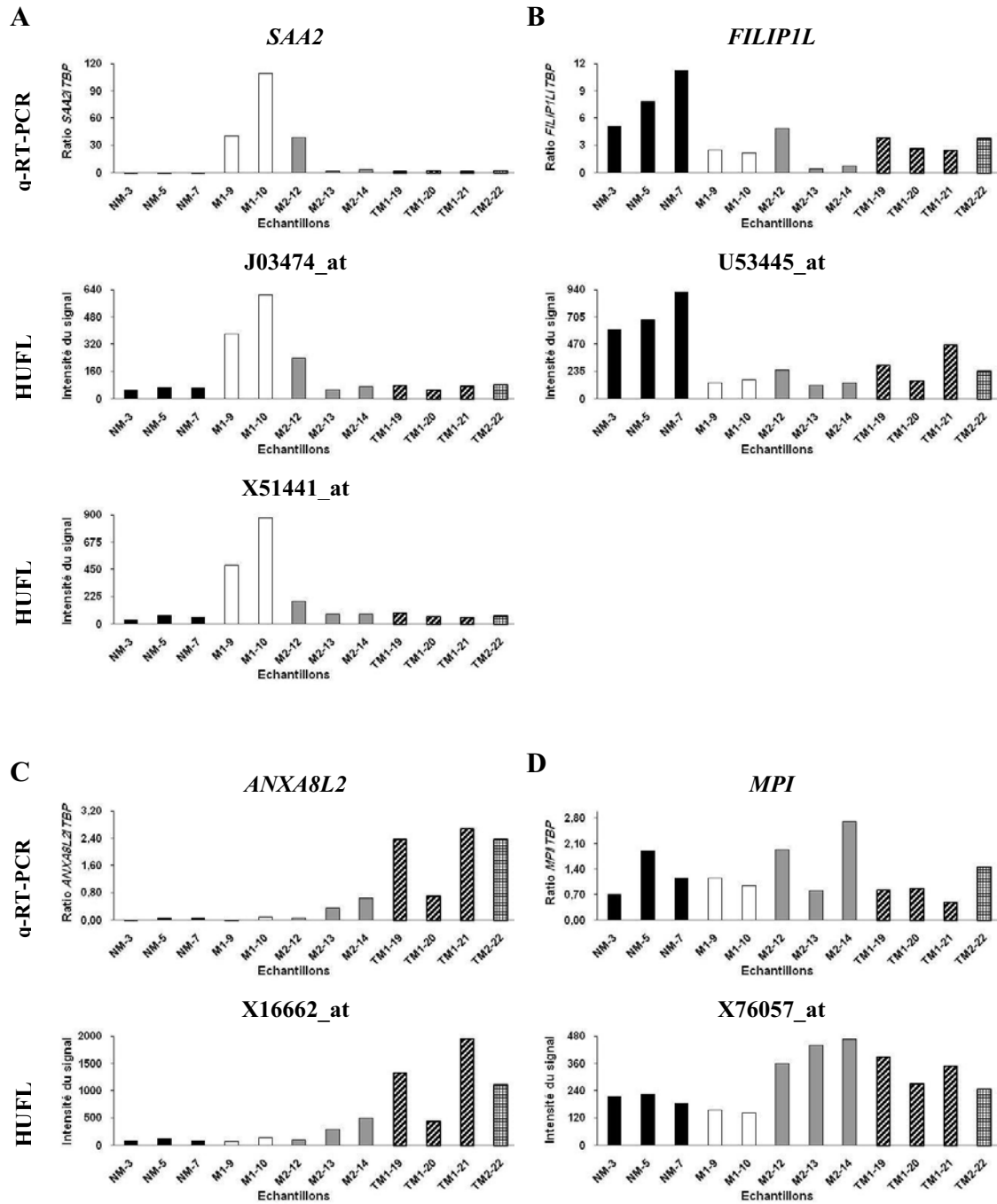


Fig. 16 : Comparaison des niveaux d'expression des quatre gènes candidats mesurés par q-RT-PCR et par HuGeneFL[®] dans la série initiale.

Les graphiques supérieurs représentent la quantification par q-RT-PCR. Les graphiques inférieurs représentent les niveaux d'expression mesurés par la biopuce HuGeneFL[®].

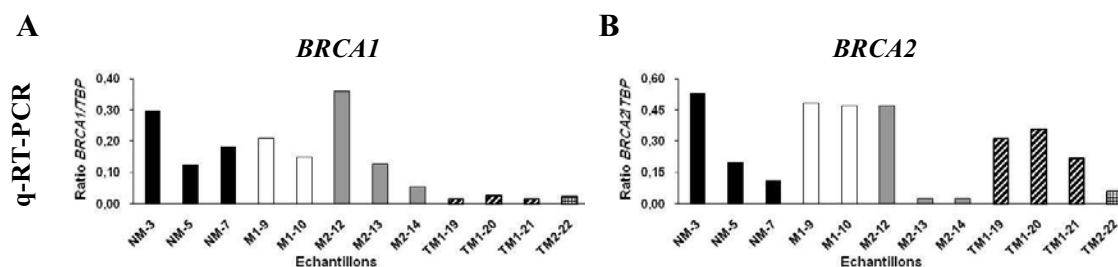


Fig. 17 : Niveaux d'expression de *BRCA1* et *BRCA2* mesurés par q-RT-PCR dans la série initiale.

C.5. Quantification par PCR quantitative en temps réel des quatre gènes candidats et de *BRCA1* et *BRCA2* dans la série étendue

L'amélioration des conditions de culture nous a permis d'étendre notre série initiale et de vérifier si les profils obtenus avec les échantillons initiaux étaient confirmés sur les échantillons supplémentaires provenant de donneuses indépendantes que nous avons inclus dans notre série étendue (Tableaux IV et V).

Nous avons également quantifié par q-RT-PCR les niveaux d'expression des quatre gènes candidats (Fig. 18A, 18B, 18C et 18D), ainsi que ceux de *BRCA1* et *BRCA2* (Fig. 18E et 18F) sur la série étendue. En général, les profils d'expression de la série étendue se sont avérés concordants avec les résultats que nous avons observés dans la série initiale avec la HuGeneFL[®] (Tableau VI et Fig. 18). Cependant, nous n'avons pas pu confirmer la surexpression du *MPI* dans les F-TOV1 comparés aux F-NOSE1 (Tableau VI et Fig. 18D).

Nous n'avons trouvé aucune différence significative entre les niveaux d'expression de *BRCA2* dans les échantillons M2-16 (cultivé dans du milieu OSE standard) et M2-17 (cultivé dans du milieu MM), ni entre les échantillons M2-15 et M2-18 cultivés dans du milieu MM selon la procédure PMK (Fig. 18F). En revanche, nous avons observé des différences importantes quant au niveau du transcrite de *BRCA2* dans les échantillons M2-12 vs (M2-13, M2-14 et M2-16), qui sont tous cultivés dans le milieu OSE standard (Fig. 18F). Ces résultats suggèrent que le milieu de culture utilisé n'est pas le facteur responsable des variations du niveau d'expression de *BRCA2* dans les échantillons des porteuses de mutation sur ce gène.

Nous confirmons ci-dessous que la variation des niveaux du transcrit de *SAA2* dans les F-NOSE2 de la série étendue est concordante avec celle de l'expression de *BRCA2*, ce qui plaide en faveur d'une corégulation de ces deux gènes chez les porteuses de mutation de *BRCA2* (Fig. 18A et 18F).

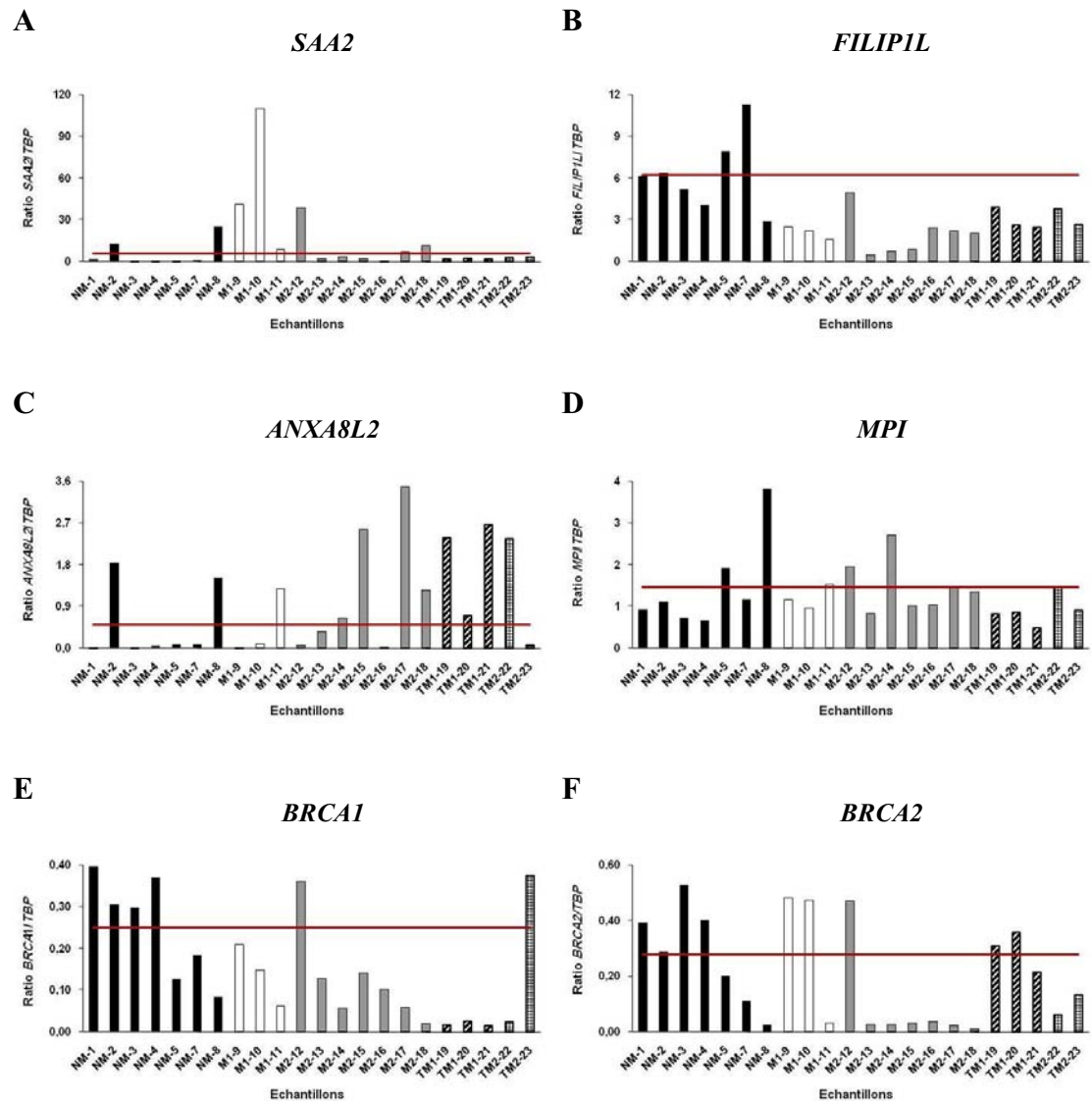


Fig. 18 : Niveaux d'expression des quatre gènes candidats et de *BRCA1* et *BRCA2* mesurés par q-RT-PCR dans la série étendue.

Dans chaque figure, la ligne horizontale représente la moyenne des ratios des échantillons NF-NOSE du transcrit correspondant.

C.6. Quantification spécifique des allèles G6085T et 8765delAG de *BRCA2*

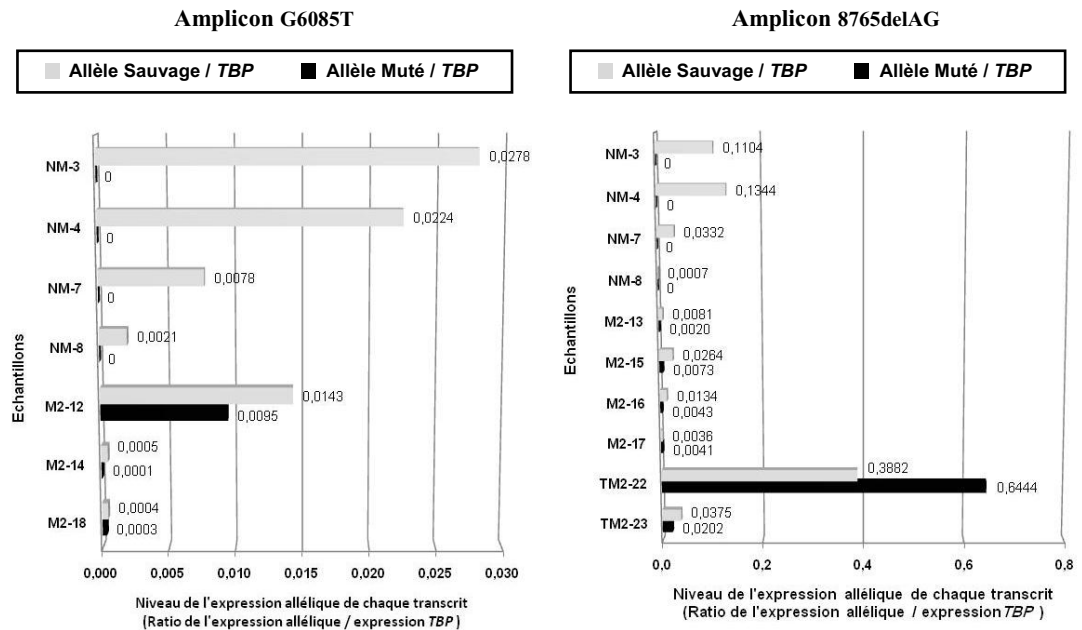
Afin de déterminer la contribution relative des transcrits de chacun des allèles mutés et sauvage au niveau d'expression total de *BRCA2*, nous avons mesuré le niveau d'expression de chacun des deux allèles dans les échantillons des porteuses des mutations *BRCA2*-G6085T et *BRCA2*-8765delAG, quantification exercée au site de la mutation mais aussi dans un sous-groupe d'échantillons non tumoraux provenant de donneuses non porteuses (Fig. 19). Ce sous-groupe a servi de contrôle pour confirmer la spécificité de la sonde mutée. La rétention et l'expression de l'allèle sauvage dans chacun des deux échantillons hétérozygotes soit pour *BRCA2*-G6085T ou *BRCA2*-8765delAG utilisés comme calibreur a été vérifiée par séquençage de chacun des amplicons purifié après amplification par q-RT-PCR (données non représentées).

Nous avons quantifié l'allèle G6085T sur des échantillons non tumoraux provenant de donneuses porteuses de la mutation G6085T sur *BRCA2* et de donneuses non porteuses de mutations. Notre groupe a montré que l'allèle sauvage était retenu et exprimé dans les trois échantillons non tumoraux mutés sur *BRCA2* que nous avons étudié (Fig. 19A, graphique de gauche). De plus, le niveau d'expression total de *BRCA2* était plus élevé dans l'échantillon M2-12 provenant d'une donneuse ayant développé un cancer du sein et du côlon que dans les deux autres échantillons (Fig. 19A, graphique de gauche), M2-14 et M2-18 qui ne présentent aucun phénotype tumoral (Fig. 19A, graphique de gauche). Il est important de noter que le niveau d'expression que nous avons quantifié pour l'allèle muté de *BRCA2* représente 40 % du niveau d'expression total de *BRCA2* dans cet échantillon, alors qu'il représente seulement 12 % du total dans l'échantillon M2-14 (Fig. 19B, graphique de gauche). Pour M2-18, il représente quand même 42% du total, mais comme le niveau d'expression de *BRCA2* est très bas dans ce dernier, le niveau d'expression absolu de l'allèle muté reste faible. Comme nous ne disposions pas d'échantillon tumoral ovarien de femme porteuse de cette mutation, nous n'avons pu confirmer la rétention de l'allèle sauvage dans les tumeurs ovariennes. Mais, nous avons confirmé la rétention et l'expression de l'allèle sauvage sur notre échantillon calibreur de cancer du sein provenant d'une femme porteuse de la mutation G6085T (non représenté dans notre tableau).

Pour la mutation 8765delAG, nous disposions non seulement d'échantillons mutés sur *BRCA2*, mais aussi de deux échantillons tumoraux : l'un provenant de la tumeur OV552

(TM2-22) et l'autre provenant de la lignée cellulaire tumorale TOV81D (TM2-23). Il est intéressant de noter que nous avons mis en évidence la rétention et l'expression de l'allèle sauvage de *BRCA2* non seulement dans les échantillons mutés non tumoraux mais aussi, et ce à notre grande surprise, dans les échantillons tumoraux des femmes porteuses de la mutation 8765delAG. (Fig. 19A, graphique de droite). Le niveau d'expression total des allèles sauvage et muté était plus élevé dans les échantillons tumoraux que dans les échantillons non tumoraux des porteuses. De plus, le niveau d'expression total de *BRCA2* était plus élevé dans TM2-22 (obtenu à partir d'une tumeur agressive, OV552) que dans TM2-23 (dérivant d'une lignée cellulaire de cancer ovarien qui se multiplie à temps de doublement lent, la lignée TOV81D) (Fig. 19A, graphique de droite). Or, il se trouve que les cellules ovariennes tumorales obtenues à partir de l'échantillon TM2-23 bien que provenant d'un adénocarcinome ovarien invasif (stade IIIC), étaient bien différenciées (grade 1). Elles présentaient un comportement en culture très voisin de celui des cellules non tumorales. La donneuse n'est d'ailleurs pas décédée de son cancer ovarien mais d'un cancer du sein diagnostiqué douze ans plus tard. Au contraire, l'échantillon TM2-22, lui, provenait de cellules présentes dans l'ascite accompagnant un cystadénocarcinome papillaire séreux agressif de stade IV avec des cellules très indifférenciées (grade 3). Alors que le niveau d'expression total de *BRCA2* que nous avons quantifié dans l'échantillon TM2-23 était faible (Fig. 19A, graphique de droite), le niveau d'expression total de *BRCA2* dans TM2-22, lui, était élevé non seulement en comparaison à celui de TM2-23 mais aussi à celui mesuré dans les F-NOSE2 et NF-NOSEs (Fig. 19A, graphique de droite). Le niveau d'expression de l'allèle sauvage de *BRCA2* dans TM2-22 représentait seulement 38 % du niveau d'expression total (Fig. 19B, graphique de droite) alors que l'allèle muté représentait lui, 62 % du total. Dans TM2-23, il ne représentait que 35 % d'un niveau total d'expression faible (Fig. 19B, graphique de droite). Cette observation suggère que le phénotype tumoral pourrait résulter de la contribution relative du transcrite de l'allèle muté dans le niveau d'expression total, mais aussi de la valeur absolue du niveau d'expression de cet allèle, suggérant un effet de dosage du transcrite de l'allèle muté *BRCA2*-8765delAG.

A Quantification spécifique d'allèle pour chaque amplicon au site de la mutation



B Distribution relative de l'expression des allèles sauvage et muté par rapport à l'expression globale de l'amplicon

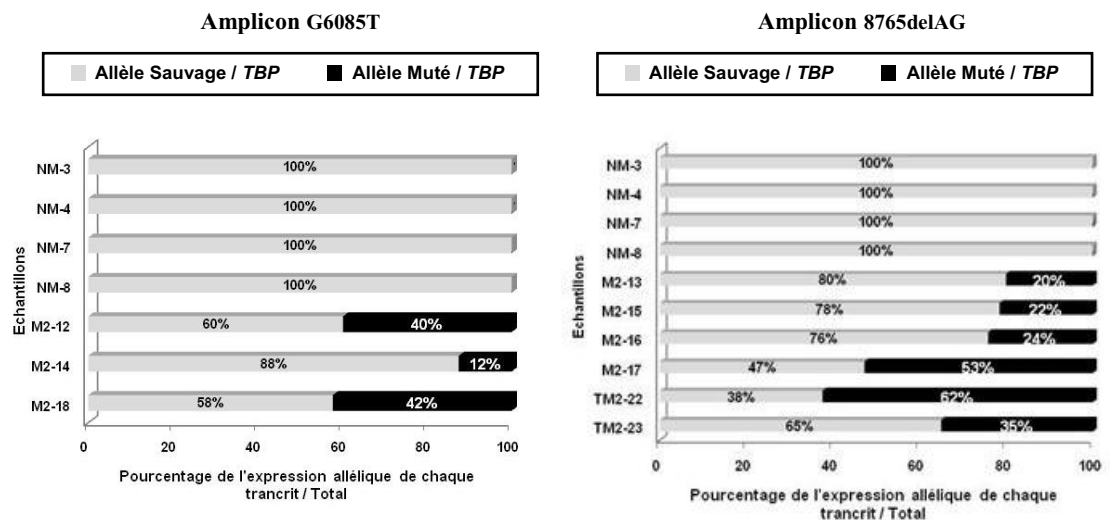


Fig. 19 : Quantification spécifique par q-RT-PCR des allèles G6085T et 8765delAG de *BRCA2*.

L'allèle sauvage est représenté en gris. L'allèle muté est représenté en noir.

A) Quantification par q-RT-PCR du niveau d'expression des allèles G6085T et 8765delAG de *BRCA2* et des allèles sauvages correspondants dans les échantillons porteurs de ces mutations et dans un sous-groupe de NF-NOSEs utilisées comme témoins pour confirmer la spécificité de la sonde sauvage.

B) Contributions relatives des allèles muté et sauvage dans l'expression totale de *BRCA2* dans les échantillons des femmes porteuses de la mutation *BRCA2*-G6085T ou de la *BRCA2*-8765delAG.

C.7. Localisation des régions chromosomiques coréglées colocalisées chez les porteurs de mutation de *BRCA1*

En utilisant la fonction `geneSetTest` du module LIMMA, nous avons localisé six régions chromosomiques enrichies (RE) en gènes différentiellement exprimés dans les cellules NOSEs provenant de donneuses porteuses d'une mutation sur *BRCA1* (Tableau VII).

Tableau VII : Régions chromosomiques enrichies (RE) en gènes différentiellement exprimés identifiées dans les cellules NOSEs des porteuses d'une mutation sur *BRCA1* tel que défini avec la fonction `geneSetTest` ($p < 0,01$).

| Région | Chr. (HUGO) | Brin (HUGO) | Bornes (b*) (Affymetrix) | Longueur (Mb**) (Mb**) | n | N (Ensembl) | P |
|--------|-------------|-------------|---------------------------|------------------------|----|-------------|--------|
| RE.19 | 19 | + | 43 557 029 – 45 973 139 | 2,8 | 10 | 37 | 0.0012 |
| RE.3 | 3 | + | 134 947 924 – 150 028 277 | 15,2 | 10 | 36 | 0.0042 |
| RE.16 | 16 | - | 65 513 089 – 54 952 865 | 10,6 | 10 | 30 | 0.0070 |
| RE.11 | 11 | - | 65 178 392 – 62 316 173 | 2,8 | 10 | 50 | 0.0073 |
| RE.8 | 8 | + | 19 840 861 – 26 491 337 | 6,6 | 10 | 31 | 0.0093 |
| RE.2 | 2 | - | 97 739 232 – 85 475 381 | 12,2 | 10 | 45 | 0.0094 |

*b = base, **Mb = Mégabase, n = Nombre de gènes quantifiés par la biopuce HuGeneFL[®], N = Nombre total des gènes présents dans la région chromosomique sur le brin correspondant.

En raison de la faible densité de la biopuce utilisée, les régions RE identifiées avec `geneSetTest` couvrent de larges régions chromosomiques. Elles peuvent ainsi correspondre à plusieurs régions coréglées colocalisées (RCC) et contenir des gènes ayant des profils d'expression différents qui pourraient être contrôlés par différentes régions de contrôle global. Nous avons identifié des RCC putatives (Tableau VIII et IX) existant au sein des régions enrichies en ESDEs identifiées par `geneSetTest`. Ces RCC peuvent être délimitées par des gènes présentant un profil d'expression différent de ceux appartenant à la RCC si la RCC est incluse en totalité dans la région étudiée ou bien par des gènes ayant le même niveau d'expression suggérant alors que la RCC pourrait chevaucher la région suivante ou précédente.

Tableau VIII : Régions coréglées colocalisées (RCC) putatives au sein des régions enrichies en ESDEs identifiées dans les cellules NOSEs des porteuses d'une mutation sur *BRCA1*.

| Région | Chr. (HUGO) | Brin (HUGO) | Bornes (b*) (Affymetrix) | Bande Chromosomique (HUGO) | Longueur (Mb**) | n | N (Ensembl) |
|--------|-------------|-------------|---------------------------|----------------------------|-----------------|----|-------------|
| RCC.19 | 19 | + | 43 557 029 – 45 973 139 | 19q13 | 2,8 | 10 | 37 |
| RCC.3 | 3 | + | 142 688 615 – 150 028 277 | 3q21-q22-q23-q24 | 7,3 | 6 | 12 |
| RCC.16 | 16 | - | 65 513 089 – 56 749 316 | 16q23-q22-q21 | 8,8 | 7 | 17 |
| RCC.11 | 11 | - | 65 121 802 – 62 330 945 | 11q13-q12 | 2,8 | 8 | 48 |
| RCC.8 | 8 | + | 22 158 563 – 22 280 736 | 8p21 | 0,1 | 2 | 3 |
| RCC.2 | 2 | - | 97 739 232 – 86 224 565 | 2p12-p11-q11 | 11,5 | 8 | 37 |

*b = base, **Mb = Mégabase, n = Nombre de gènes quantifiés par la biopuce HuGeneFL[®], N = Nombre total des gènes présents dans la sous-région chromosomique.

La région RCC.19 recouvre 2,8 Mb (Méga bases) sur le brin positif du chromosome 19. Elle est composée de 37 gènes dont 10 quantifiés avec la HuGeneFL[®] qui sont surexprimés dans les échantillons F-NOSE1 comparés aux NF-NOSEs (Tableau IX). RCC.19 est superposable à la région enrichie RE.19 identifiée par geneSetTest. Selon Spitz, les régions de co-expression contrôlées par une RCG contiennent plusieurs gènes distincts (jusqu'à 20) et ne correspondent pas aux classiques complexes multigéniques produits par la duplication d'un gène ancestral qui partagent souvent les séquences régulatrices héritées de leur ancêtre commun [266]. RCC.19 pourrait elle-même inclure une ou plusieurs RCC, car elle s'étend sur une large fenêtre chromosomique comportant plus de 20 gènes. La région RCC.3 s'étend sur 7,3 Mb sur le brin positif du chromosome 3. Elle est formée de 12 gènes dont six sont quantifiés avec la HuGeneFL[®], et sont sous-exprimés dans les échantillons F-NOSE1 vs NF-NOSEs (Tableau X). Cette région est mal délimitée avec un gène adjacent qui tend à être sous-exprimé dans les F-NOSE1.

La région RCC.16 s'étend sur 8,8 Mb sur le chromosome 16. Elle renferme 17 gènes et comporte sept gènes dont l'expression est quantifiée par la HuGeneFL[®], (encadrés en rouge, Tableau XI). Ces sept gènes sont sous-exprimés dans les F-NOSE1 comparés aux NF-NOSEs. Ils sont précédés par un gène ayant un profil d'expression distinct (Tableau XI). De même, les régions RCC.11, RCC.2 contiennent respectivement 48 et 37 gènes (revérifie le nombre de gènes dans l'encadré). Huit gènes contigus sont quantifiés avec la HuGeneFL[®] dans RCC.11 (encadrés en rouge, Tableau XII) et huit dans RCC.2 (encadrés en rouge, Tableau XIV). Ils présentent une sous-expression dans les échantillons

F-NOSE1 comparés aux NF-NOSEs. Les gènes de part et d'autre s'en distinguent par leur profil d'expression délimitant ainsi les domaines de co-expression.

Enfin, La région RCC.8 comporterait deux gènes sous-exprimés dans les F-NOSE1 comparés aux NF-NOSEs. Cette région inclus trois gènes dont deux sont quantifiés avec la HuGeneFL[®]. Les gènes *POLR3D* et *SLC39A14* pourraient faire partie d'une RCC dans les cellules F-NOSE1 (Tableau XIII). Cependant, il est difficile en réalité de déterminer ici d'emblée les bornes du ou des domaines de co-expression dans cette région du fait du nombre limité de gènes adjacents ayant un profil d'expression similaire (Tableau XIII). De plus, il existerait 13 gènes entre *PPP3CC* et *NEFM*. Il est donc possible que des gènes contigus non quantifiés par la biopuce utilisée définissent d'autres domaines de corégulation et que celui que nous avons sélectionné ici ne corresponde pas réellement à un tel domaine mais à un faux positif lié à la méthode.

Tableau IX : Profil d'expression de la région enrichie en ESDEs identifiée sur le chromosome 19 dans les cellules NOSEs des porteuses d'une mutation sur *BRCA1* tel que défini avec la fonction *geneSetTest* ($p = 0.0012$) et de la RCC.19 correspondante.

| Symbole (HUGO) | Description (HUGO) | Position (bp) (Affymetrix) | Test t LIMMA | Position | | | | | |
|-------------------|--|----------------------------------|-----------------|----------|--------|--------|--------|--------|--------|
| | | | | NM.3 | NM.5 | NM.6 | NM.7 | M1.9 | M1.10 |
| <i>PSMD8</i> | proteasome (prosome, macropain) 26S subunit, non-ATPase, 8 | 43 557 029 | 1.17 | 10.792 | 10.805 | 10.924 | 11.267 | 11.080 | 11.223 |
| <i>RYR1</i> | ryanodine receptor 1 (skeletal) | 43 616 179 | 0.96 | 5.207 | 5.071 | 5.510 | 5.052 | 5.222 | 5.551 |
| <i>LGALS7</i> | lectin, galactoside-binding, soluble, 7 | 43 971 693 | 0.59 | 6.385 | 7.000 | 6.707 | 7.103 | 7.002 | 6.847 |
| <i>NFKB1B</i> | nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, beta | 44 082 454 | 0.64 | 6.819 | 4.250 | 6.493 | 6.649 | 6.357 | 6.572 |
| <i>ZFP36</i> | zinc finger protein 36, C3H type, homolog (mouse) | 44 589 326 | 1.94 | 7.929 | 7.889 | 7.739 | 8.214 | 8.254 | 8.269 |
| <i>SUPT5H</i> | suppressor of Ty 5 homolog (S. cerevisiae) | 44 628 163 | 1.12 | 8.698 | 7.931 | 8.723 | 9.047 | 9.016 | 8.813 |
| <i>MAP3K10</i> | mitogen-activated protein kinase kinase kinase 10 | 45 389 490 | 1.83 | 8.797 | 9.250 | 9.438 | 9.215 | 9.499 | 9.548 |
| <i>ITPKC</i> | inositol 1,4,5-trisphosphate 3-kinase C | 45 914 847 | 1.21 | 5.112 | 5.314 | 5.670 | 5.207 | 5.638 | 5.462 |
| <i>SNRPA</i> | small nuclear ribonucleoprotein polypeptide A | 45 948 618 | 3.22 | 8.673 | 8.543 | 8.386 | 8.299 | 9.010 | 8.938 |
| <i>MIA</i> | melanoma inhibitory activity | 45 973 139 | 0.93 | 6.439 | 6.750 | 6.800 | 6.371 | 6.741 | 6.756 |

Tableau X : Profil d'expression de la région enrichie en ESDEs identifiée sur le chromosome 3 dans les cellules NOSEs des porteuses d'une mutation sur *BRCA1* tel que défini avec la fonction *geneSetTest* ($p = 0.0042$) et de la RCC.3 correspondante.

| Symbole (HUGO) | Description (HUGO) | Position (bp) (Affymetrix) | Test t LIMMA | Position | | | | | |
|-------------------|--|----------------------------------|-----------------|----------|--------|--------|--------|--------|--------|
| | | | | NM.3 | NM.5 | NM.6 | NM.7 | M1.9 | M1.10 |
| <i>TF</i> | transferrin | 134 947 924 | -1.30 | 4.623 | 4.976 | 4.672 | 4.797 | 4.583 | 4.555 |
| <i>EPHB1</i> | EPH receptor B1 | 135 996 949 | -0.60 | 5.992 | 6.296 | 6.112 | 5.983 | 5.811 | 6.178 |
| <i>PCCB</i> | propionyl Coenzyme A carboxylase, beta polypeptide | 137 451 856 | -0.78 | 6.010 | 6.164 | 5.928 | 6.014 | 5.926 | 5.919 |
| <i>NCK1</i> | NCK adaptor protein 1 | 138 063 762 | -0.16 | 5.061 | 5.034 | 4.711 | 4.796 | 4.950 | 4.799 |
| <i>RASA2</i> | RAS p21 protein activator 2 | 142 688 615 | -1.41 | 4.467 | 4.954 | 4.810 | 5.056 | 4.534 | 4.584 |
| <i>ATP1B3</i> | ATPase, Na ⁺ /K ⁺ transporting, beta 3 polypeptide | 143 078 159 | -1.53 | 11.106 | 10.791 | 11.230 | 10.941 | 10.819 | 10.713 |
| <i>PLS1</i> | plastin 1 (1 isoform) | 143 825 033 | -3.40 | 5.219 | 5.093 | 5.178 | 5.063 | 4.520 | 4.777 |
| <i>TRPC1</i> | transient receptor potential cation channel, subfamily C, member 1 | 143 925 955 | -0.83 | 7.503 | 6.915 | 7.002 | 7.635 | 7.138 | 7.011 |
| <i>ZIC1</i> | Zic family member 1 (odd-paired homolog, | 148 609 870 | -0.48 | 4.511 | 4.509 | 4.555 | 4.459 | 4.427 | 4.466 |
| <i>CPB1</i> | carboxypeptidase B1 (tissue) | 150 028 277 | -1.12 | 6.232 | 6.369 | 6.306 | 6.254 | 6.185 | 6.099 |

Tableau XI : Profil d'expression de la région enrichie en ESDEs identifiée sur le chromosome 16 dans les cellules NOSEs des porteuses d'une mutation sur *BRCA1* tel que défini avec la fonction *geneSetTest* ($p = 0.0070$) et de la RCC.16 correspondante.

| Symbole (HUGO) | Description (HUGO) | Position (bp) (Affymetrix) | Test t LIMMA | Position | | | | | |
|-------------------|---|----------------------------------|-----------------|----------|-------|-------|-------|-------|-------|
| | | | | NM.3 | NM.5 | NM.6 | NM.7 | M1.9 | M1.10 |
| <i>RRAD</i> | Ras-related associated with diabetes | -65 513 089 | -3.11 | 7.222 | 7.091 | 7.070 | 8.114 | 6.644 | 6.212 |
| <i>NAE1</i> | NEDD8 activating enzyme E1 subunit 1 | -65 394 282 | -1.32 | 7.875 | 6.718 | 7.065 | 7.125 | 6.681 | 6.942 |
| <i>TK2</i> | thymidine kinase 2, mitochondrial | -65 100 852 | -0.97 | 5.968 | 5.774 | 5.530 | 5.596 | 5.528 | 5.588 |
| <i>CDH11</i> | cadherin 11, type 2, OB-cadherin (osteoblast) | -63 538 185 | -0.64 | 7.772 | 8.977 | 9.002 | 9.303 | 8.504 | 8.536 |
| <i>CDH8</i> | cadherin 8, type 2 | -60 244 735 | -0.96 | 4.721 | 4.842 | 4.616 | 4.932 | 4.645 | 4.630 |
| <i>GOT2</i> | glutamic-oxaloacetic transaminase 2, mitochondrial (aspartate aminotransferase 2) | -57 298 535 | -0.58 | 9.541 | 7.926 | 8.846 | 9.153 | 8.609 | 8.677 |
| <i>CSNK2A2</i> | casein kinase 2, alpha prime polypeptide | -56 749 316 | -3.29 | 6.165 | 6.652 | 6.112 | 6.688 | 5.530 | 5.850 |
| <i>CNGB1</i> | cyclic nucleotide gated channel beta 1 | -56 475 205 | 0.86 | 6.970 | 7.443 | 7.131 | 7.577 | 7.534 | 7.367 |
| <i>MT1G</i> | metallothionein 1G | -55 258 153 | -0.51 | 7.565 | 8.071 | 7.648 | 8.764 | 7.986 | 7.712 |
| <i>AMFR</i> | autocrine motility factor receptor | -54 952 865 | -1.61 | 6.584 | 6.267 | 6.619 | 6.092 | 6.210 | 5.948 |

Tableau XII : Profil d'expression de la région enrichie en ESDEs identifiée sur le chromosome 11 dans les cellules NOSEs des porteuses d'une mutation sur *BRCA1* tel que défini avec la fonction *geneSetTest* ($p = 0.0073$) et de la RCC.11 correspondante.

| Symbole (HUGO) | Description (HUGO) | Position (bp) (Affymetrix) | Test t LIMMA | Position | | | | | |
|-------------------|--|----------------------------------|-----------------|----------|-------|-------|-------|-------|-------|
| | | | | NM.3 | NM.5 | NM.6 | NM.7 | M1.9 | M1.10 |
| <i>RELA</i> | v-rel reticuloendotheliosis viral oncogene homolog A (avian) | -65 178 392 | 1.68 | 5.753 | 6.455 | 6.481 | 6.443 | 6.593 | 6.734 |
| <i>MAP3K11</i> | mitogen-activated protein kinase kinase kinase 11 | -65 121 802 | -2.15 | 6.751 | 6.984 | 6.699 | 7.025 | 6.443 | 6.610 |
| <i>MEN1</i> | multiple endocrine neoplasia I | -64 327 571 | -1.65 | 6.809 | 6.726 | 6.876 | 6.546 | 6.517 | 6.473 |
| <i>MAP4K2</i> | mitogen-activated protein kinase kinase kinase 2 | -64 313 185 | -2.06 | 8.069 | 8.500 | 8.127 | 8.240 | 7.823 | 7.964 |
| <i>PYGM</i> | phosphorylase, glycogen; muscle | -64 270 605 | -1.90 | 5.801 | 6.196 | 5.818 | 5.964 | 5.492 | 5.756 |
| <i>BAD</i> | BCL2-associated agonist of cell death | -63 793 877 | -1.07 | 8.232 | 8.147 | 8.120 | 8.078 | 8.067 | 7.933 |
| <i>PLA2G16</i> | phospholipase A2, group XVI | -63 098 824 | -1.99 | 6.646 | 7.093 | 7.453 | 7.802 | 6.560 | 6.774 |
| <i>CHRM1</i> | cholinergic receptor, muscarinic 1 | -62 432 727 | -0.96 | 7.848 | 8.256 | 8.026 | 7.979 | 7.834 | 7.920 |
| <i>STX5</i> | syntaxin 5 | -62 330 945 | -0.73 | 7.357 | 7.368 | 7.227 | 7.080 | 7.146 | 7.158 |
| <i>NXF1</i> | nuclear RNA export factor 1 | -62 316 173 | 0.08 | 6.316 | 5.927 | 5.394 | 5.698 | 6.046 | 5.663 |

Tableau XIII : Profil d'expression de la région enrichie en ESDEs identifiée sur le chromosome 8 dans les cellules NOSEs des porteuses d'une mutation sur *BRCA1* tel que défini avec la fonction *geneSetTest* ($p = 0.0093$) et de la RCC.8 correspondante.

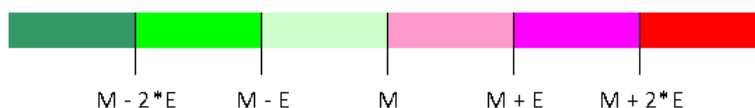
| Symbole (HUGO) | Description (HUGO) | Position (bp) (Affymetrix) | Test t LIMMA | Position | | | | | |
|-------------------|--|----------------------------------|-----------------|----------|--------|--------|--------|-------|--------|
| | | | | NM.3 | NM.5 | NM.6 | NM.7 | M1.9 | M1.10 |
| <i>LPL</i> | lipoprotein lipase | 19 840 861 | -1.38 | 4.640 | 4.879 | 4.492 | 4.955 | 4.351 | 4.635 |
| <i>ATP6V1B2</i> | ATPase, H ⁺ transporting, lysosomal 56/58kDa, V1 subunit B2 | 20 098 983 | -1.04 | 7.931 | 6.921 | 7.932 | 7.860 | 7.443 | 7.271 |
| <i>EPB49</i> | erythrocyte membrane protein band 4.9 (dematin) | 21 972 679 | 0.02 | 7.713 | 7.314 | 7.802 | 7.420 | 7.718 | 7.417 |
| <i>SFTPC</i> | surfactant protein C | 22 075 128 | -1.29 | 8.195 | 8.466 | 8.271 | 8.203 | 7.847 | 8.280 |
| <i>POLR3D</i> | polymerase (RNA) III (DNA directed) polypeptide D, 44kDa | 22 158 563 | -1.93 | 7.279 | 7.353 | 6.975 | 7.383 | 6.715 | 7.088 |
| <i>SLC39A14</i> | solute carrier family 39 (zinc transporter), member 14 | 22 280 736 | -2.73 | 10.616 | 10.612 | 10.636 | 10.245 | 9.967 | 10.176 |
| <i>PPP3CC</i> | protein phosphatase 3 (formerly 2B), catalytic subunit, gamma isoform | 22 354 540 | 0.24 | 5.220 | 4.971 | 5.257 | 5.010 | 5.467 | 4.857 |
| <i>NEFM</i> | neurofilament, medium polypeptide | 24 827 211 | 0.69 | 5.900 | 4.684 | 5.336 | 5.482 | 5.450 | 5.662 |
| <i>PPP2R2A</i> | protein phosphatase 2 (formerly 2A), regulatory | 26 204 950 | -2.34 | 8.936 | 9.031 | 8.735 | 9.192 | 8.602 | 8.584 |
| <i>DPYSL2</i> | dihydropyrimidinase-like 2 | 26 491 337 | -1.24 | 7.810 | 7.882 | 7.437 | 7.158 | 7.312 | 7.289 |

Tableau XIV : Profil d'expression de la région enrichie en ESDEs identifiée sur le chromosome 2 dans les cellules NOSEs des porteuses d'une mutation sur *BRCA1* tel que défini avec la fonction *geneSetTest* ($p = 0.0094$) et de la RCC.2 correspondante.

| Symbole (HUGO) | Description (HUGO) | Position (bp) (Affymetrix) | Test t LIMMA | Position | | | | | |
|-------------------|---|----------------------------------|-----------------|----------|-------|-------|-------|-------|-------|
| | | | | NM.3 | NM.5 | NM.6 | NM.7 | M1.9 | M1.10 |
| <i>TMEM131</i> | transmembrane protein 131 | -97 739 232 | -1.32 | 6.745 | 6.993 | 6.449 | 7.211 | 6.566 | 6.567 |
| <i>ACTR1B</i> | ARP1 actin-related protein 1 homolog B, centractin beta (yeast) | -97 638 862 | -0.90 | 6.488 | 6.709 | 6.199 | 6.671 | 6.410 | 6.304 |
| <i>DUSP2</i> | dual specificity phosphatase 2 | -96 172 643 | -1.91 | 7.479 | 7.566 | 7.490 | 7.628 | 7.244 | 7.326 |
| <i>ADRA2B</i> | adrenergic, alpha-2B-, receptor | -96 142 352 | -1.05 | 8.976 | 9.035 | 9.223 | 8.921 | 8.811 | 8.956 |
| <i>FABP1</i> | fatty acid binding protein 1, liver | -88 203 624 | -1.06 | 4.385 | 4.531 | 4.623 | 4.439 | 4.288 | 4.403 |
| <i>CD8A</i> | CD8a molecule | -86 865 239 | -1.59 | 6.517 | 6.680 | 6.526 | 6.565 | 6.338 | 6.382 |
| <i>RNF103</i> | ring finger protein 103 | -86 684 026 | -4.36 | 8.044 | 7.463 | 7.652 | 7.598 | 6.930 | 6.849 |
| <i>IMMT</i> | inner membrane protein, mitochondrial (mitofilin) | -86 224 565 | -2.52 | 6.924 | 6.164 | 6.470 | 6.424 | 6.156 | 5.651 |
| <i>GGCK</i> | gamma-glutamyl carboxylase | -85 629 715 | 1.11 | 5.631 | 5.680 | 5.795 | 5.602 | 5.783 | 5.875 |
| <i>CAPG</i> | capping protein (actin filament), gelsolin-like | -85 475 381 | 0.96 | 7.818 | 8.157 | 8.102 | 8.109 | 7.946 | 8.520 |

Les gènes encadrés en rouge correspondraient à une ou plusieurs régions coréglées localisées (RCC) chez les porteuses de mutation de *BRCA1*

L'échelle des couleurs est définie par rapport à la moyenne (M) de l'ensemble des valeurs d'expression du gène sur la totalité des échantillons et par l'écart type (E) à la moyenne des valeurs de chacun des échantillons.



D. Discussion et conclusions

D.1. Originalité et limites de notre étude

L'arrivée des biopuces à ADN a permis l'identification des signatures moléculaires liées aux mutations de *BRCA1* et *BRCA2* dans les tissus non tumoraux. Cependant, en raison de l'accessibilité réduite à de tels échantillons et de la quantité limitée de matériel disponible, les études permettant l'identification de ces signatures moléculaires sont encore peu nombreuses. De plus, la plupart de ces études se sont limitées à un nombre restreint de gènes [235]. Certaines d'entre elles ont été réalisées dans des cellules NOSEs obtenues après une annexectomie prophylactique réalisée chez des femmes sur la notion de risque élevé *a priori* de développer un cancer du sein. Ce risque a été calculé à partir de l'histoire familiale de cancers sans que la présence ou l'absence d'une mutation germinale ait été vérifiée. Mais, puisque nous sommes en transmission autosomique dominante, le risque de ces dernières d'avoir hérité de la mutation transmise par un des parents est de 50 %. Il est donc possible qu'une partie de ces NOSEs correspondent à des femmes non porteuses et qu'elles auraient dû être intégrées dans la classe que nous avons appelée NM, c'est-à-dire non tumorale et non mutée.

Notre étude se distingue dès lors des précédentes, d'une part parce qu'elle est réalisée à l'échelle génomique, d'autre part parce que nous avons recruté uniquement des patientes d'origine Canadienne Française chez lesquelles soit les huit mutations les plus fréquentes chez les Canadiennes Françaises ont été recherchées, soit la mutation familiale si la patiente appartenait à une famille pour laquelle une mutation était déjà connue.

Le groupe de Jazaeri a comparé le profil d'expression de tumeurs primaires ovariennes (ARN extrait de tissus congelés) provenant de donneuses non porteuses de mutation et de donneuses porteuses de mutation sur *BRCA1* et *BRCA2* à celui de lignées cellulaires IOSEs (cellules NOSEs immortalisées par SV-40) afin d'identifier les gènes surexprimés dans les tumeurs ovariennes indépendamment de leur statut génétique [95]. Par conséquent, une partie des gènes identifiés par cette étude pourrait refléter les différences entre modèles utilisés (tissu *vs* culture) et non pas la transformation tumorale, biais expérimental déjà identifié par ma directrice, le Dr Maugard avec ses collaborateurs [246]. Une étude protéomique a également permis d'identifier des protéines différenciellement exprimées dans les cellules IOSEs provenant de porteuses de mutation sur *BRCA1* [7, 267]. Ces données protéomiques, bien qu'elles viennent confirmer l'existence d'une signature moléculaire associée à une mutation sur *BRCA1*, ne peuvent être comparées aux nôtres, car

elles ont été effectuées sur des cellules immortalisées, à un stade plus avancé de la carcinogenèse. En effet, les cellules NOSEs transformées avec SV-40 acquièrent une autonomie accrue avec des caractéristiques voisines de celles des TOVs même en l'absence de mutations [235]. Une augmentation importante du nombre d'aberrations chromosomiques a également été observée après immortalisation [268]. Enfin, les profils d'expression générés avec l'ARN extrait des cellules IOSEs sont distincts de ceux obtenus avec des cellules isolées par broyage de la surface ovarienne ou avec des cellules NOSEs en cultures primaires [269]. Ainsi, nos données sont uniques, puisqu'elles interrogent les premières étapes de la carcinogenèse héréditaire ovarienne.

D.2. Avantages et limites des méthodes utilisées

La recherche post-génomique pendant ces 20 dernières années a connu un essor remarquable avec l'arrivée des biopuces. Cet outil permet aux chercheurs de mesurer simultanément, dans un tissu donné, l'expression de dizaines de milliers de transcrits et de réaliser des études à l'échelle du génome. En ce sens, la HuGeneFL[®] représentait, lorsqu'elle a été proposée pour la première fois, une avancée majeure par rapport aux techniques classiques de la biologie moléculaire, puisqu'elle permettait de mesurer le niveau d'expression d'environ 6 800 gènes simultanément.

L'un des avantages de la HuGeneFL[®] réside dans le fait que ses ES, composés de 20 paires de sondes, sont distribués sur l'ensemble de la séquence du gène étudié et qu'elles ne sont pas biaisées en 3' comme dans la biopuce HG-U133 Plus 2[®]. Ceci permet une quantification plus robuste de l'ensemble des messagers. En effet, dans le cas où il existe des phénomènes d'épissage alternatif, certains des transcrits pourraient ne pas être reconnus dans le cas de sonde biaisées en 3', problème qui n'est pas observé avec la HuGeneFL[®]. Son inconvénient majeur est sa faible densité de couverture du génome (7 129 ES correspondant à ~6 800 gènes) qui nous a conduit à envisager pour nos études ultérieures l'utilisation des biopuces ayant une plus forte densité.

Parmi les méthodes de normalisation existantes, nous avons choisi d'utiliser la procédure RMA car elle est la mieux adaptée aux données d'expression générées par les biopuces de la plateforme Affymetrix[™]. Notre choix est aussi justifié par les résultats des deux études comparatives effectuées par Bolstad et al. [249] et Irizarry et al. [255] pour évaluer la performance de différentes méthodes de normalisation. En se basant sur des données

empiriques, Bolstad et al. [249] ont comparé la performance de trois méthodes de normalisation qu'ils ont proposées : méthode cyclique de loess [270], méthodes basées sur les contrastes [271] et les quantiles [249], trois méthodes dites de données complètes qui combinent l'information de toutes les puces pour établir la relation de normalisation, à deux autres méthodes utilisant une biopuce de référence : une méthode appelée par les auteurs méthode de standardisation similaire à l'approche d'Affymetrix [257] et une seconde appelée méthode non linéaire similaire à l'approche de Schadt et al. [272]. Bolstad et al. ont montré que la méthode basée sur les quantiles, sur laquelle repose la procédure RMA, est la plus rapide à utiliser et la plus efficace pour réduire la variance sans augmenter le biais de quantification des niveaux d'expression des gènes. Irizarry et al. pour leur part, ont comparé la procédure RMA qu'ils avaient développée à trois procédures de normalisation : AvDiff, MAS 5.0 et MBEI [255]. Les auteurs ont montré que la procédure RMA est celle qui donne la meilleure spécificité et sensibilité de détection des expressions différentielles. Enfin, l'un des avantages les plus importants de cette méthode réside dans le fait qu'elle se limite aux valeurs PM sans utiliser les valeurs MM pour normaliser les données d'expression des gènes. Théoriquement, les sondes MM servent au contrôle des hybridations non spécifiques, mais pratiquement l'intensité de ces sondes peut correspondre aussi à un vrai signal, si un polymorphisme présent sur un des transcrits était reconnu par une des sondes MM [255]. Aucune des méthodes de normalisation actuellement disponibles, y compris la procédure GCRMA, n'a réussi jusqu'à maintenant à intégrer les sondes MM dans ce calcul, sans ajouter de variabilité additionnelle. Or, nous cherchons à éliminer toute variabilité non biologique. Les résultats de normalisation obtenus avec les procédures RMA et GCRMA sur nos propres données d'expression confirment que l'utilisation de la GCRMA ne permet pas d'améliorer la normalisation de ces données.

La méthode LIMMA que nous avons retenue pour l'analyse différentielle de nos données d'expression possède plusieurs avantages. D'une part, elle ne se limite pas à un seul type de design d'expérience ce qui lui confère une grande souplesse d'utilisation, qu'il s'agisse d'une application à des plateformes distinctes, ou bien d'effectuer simultanément des analyses différentielles sur plusieurs contrastes, tel que nous l'avons réalisé, ce qui l'apparente à une méthode d'analyse de variances de type ANOVA, d'autre part, l'utilisation des statistiques bayésiennes lui permet de générer des résultats robustes même dans le cas d'expériences conduites sur un petit nombre d'échantillons, ce qui est tout à fait adapté à notre propre série de données. En effet, LIMMA emploie le t modéré à la place du t de Student pour mieux estimer la variance globale des ES quantifiés par toutes les

biopuces impliquées dans l'analyse différentielle. Dans le cas d'un nombre d'échantillons inférieur à 30, le test t de Student est applicable à condition que les échantillons suivent, dans chaque classe, une distribution normale et que l'hypothèse d'égalité des variances entre les deux classes soit acceptable. Cependant lorsque le nombre d'échantillons est limité, tel que dans le cas de nos données (2 à 5 échantillons par classe), l'hypothèse d'égalité des variances entre les classes ne peut plus être retenue. Le t modéré proposé par Smyth [250] vient résoudre ce problème en modérant l'ensemble des écarts type de chaque classe pour tous les ES et en les réduisant à une valeur commune grâce à un modèle bayésien simple. De ce fait, le t modéré produit des valeurs de p de la même manière que le t de Student mais avec un nombre de degrés de liberté plus élevé conférant ainsi à LIMMA une puissance accrue pour les analyses appliquées sur un petit nombre d'échantillons.

Nous nous sommes heurtés par la suite au problème des comparaisons multiples. Nous discuterons ici des paramètres que prennent en compte les méthodes de correction et du choix de la correction de Benjamini et Hochberg.

Le paramètre de loin le plus utilisé est le « *False Discovery Rate* » (FDR) qui représente le taux de fausses découvertes ou proportion attendue d'hypothèses nulles rejetées à tort (faux positifs) parmi toutes les hypothèses nulles rejetées [251]. Contrôler le FDR au seuil de 0,05 signifie que le taux attendu de faux positifs parmi les ES déclarés comme différentiellement exprimés est inférieur à 5 %. Un autre paramètre fréquemment utilisé est le « *positive False Discovery Rate* » (pFDR) qui est une extension du FDR. Il représente le taux de faux positifs conditionnel au rejet d'au moins une hypothèse nulle [273]. Ce critère est incontrôlable dans le cas où aucun gène ne serait différentiellement exprimé. Lorsque le nombre de tests est élevé, les procédures de corrections contrôlant le FDR fournissent des résultats comparables à celles contrôlant le pFDR. Enfin, un autre paramètre très utilisé est le « *Family Wise Error Rate* » (FWER) qui représente la probabilité de rejeter à tort au moins l'une des hypothèses nulles testées, en d'autres termes d'avoir au moins un faux positif et de déclarer un gène différentiellement exprimé alors qu'il ne l'est pas en réalité [274]. Fixer le FWER au seuil de 0,05 permet donc d'être confiant à 95 % de n'avoir aucun faux positif parmi les ES déclarés comme étant différentiellement exprimés.

La correction de Benjamini et Hochberg qui contrôle le FDR est moins conservative que les procédures contrôlant le « *Family Wise Error* » (FWER). Ces dernières ont l'inconvénient d'être moins bien adaptées aux multiples hypothèses générées par les biopuces, car la probabilité de découvrir une vraie association avec cette procédure, en cas de comparaisons

multiples peut devenir proche de zéro. La procédure de Benjamini et Hochberg, elle, tolère plus de faux positifs, mais génère moins de faux négatifs. Le contrôle du FDR représente un compromis entre le refus absolu de contrôler la multiplicité, ce qui maximise l'erreur α et l'adhésion stricte au contrôle du FWER qui s'assure de ne rejeter à tort aucune hypothèse nulle, mais qui risque de minimiser la puissance du test utilisé.

Nous avons appliqué une agrégation hiérarchique sur les ESDEs identifiés avec LIMMA en employant le critère de saut maximal sur des distances euclidiennes. Selon ce critère de regroupement, la distance entre deux groupes d'échantillons est égale au maximum des distances entre les échantillons du premier et ceux du deuxième groupe. Par la suite, deux objets G et H sont considérés proches si tous les échantillons des groupes G et H ont des profils d'expression relativement similaires. Ce critère de regroupement produit donc des objets compacts à petit diamètre en d'autres termes des groupements d'échantillons très homogènes. De plus, la distance euclidienne est largement employée dans les agrégations hiérarchiques parce qu'elle est facile à calculer, mais l'utilisation de cette métrique présente certaines limites pour certains types de données. En effet, deux échantillons ayant le même profil d'expression mais avec des amplitudes différentes quant aux niveaux d'expression peuvent être classés loin l'un de l'autre, alors que deux échantillons ayant des profils d'expression différents mais avec de faibles variations d'amplitudes dans leur niveaux d'expression peuvent être classés comme étant proches l'un de l'autre. Une grande différence d'amplitude entre les niveaux d'expression des gènes dans des échantillons appartenant à la même classe pourrait survenir suite à des problèmes techniques tels qu'un problème pendant la lecture des biopuces ou une différence dans la quantité de l'ARN hybridé sur chaque biopuce. Cependant, la normalisation des données d'expression vient résoudre ce problème quand la distance euclidienne est employée pour la classification des échantillons, mais ceci n'est pas vrai si elle est utilisée pour la classification des gènes. Dans ce dernier cas, il faut utiliser une distance de corrélation au lieu de la distance euclidienne pour obtenir une meilleure agrégation.

Le choix de la PCR en temps réel (q-RT-PCR) et surtout celui de la chimie TaqMan[®], nous a permis de dessiner des sondes capables de reconnaître une région du(des) transcrit(s) de chacun de nos gènes candidats correspondant à ceux qui s'hybridaient avec l'ESDE de la HuGeneFL[®]. Ceci nous a permis d'une part de mettre en évidence la bonne concordance entre les profils d'expression des gènes candidats quantifiés par la biopuce et par la q-RT-PCR et d'autre part de confirmer que ces quatre gènes candidats présentent une expression

différentielle tel qu'identifié avec LIMMA. La q-RT-PCR nous a également permis d'effectuer une quantification spécifique des transcrits des allèles muté et sauvage de *BRCA2* en sélectionnant nos amorces et sondes TaqMan[®] de façon à ce qu'elles s'hybrident au site soit de la mutation G6085T, soit de la mutation 8765delAG de *BRCA2*. Nous avons ainsi réussi à montrer que l'allèle sauvage de *BRCA2* était exprimé dans les cellules NOSEs provenant de nos échantillons porteurs de la mutation G6085T de *BRCA2* ainsi que dans les cellules NOSEs et TOVs provenant de nos échantillons porteurs de la mutation 8765delAG de *BRCA2*, et à déterminer la contribution relative de l'allèle muté au total. Ces résultats montrent pour la première fois qu'une perte d'hétérozygotie au locus *BRCA2* n'est pas nécessaire pour la transformation tumorale ovarienne et qu'un effet de dosage de l'allèle muté de ce gène pourrait déterminer le phénotype des cellules tumorales chez les porteuses d'une mutation *BRCA2*-8765delAG.

Nous avons également effectué une étude exploratoire de différentes méthodes de localisation de régions coréglées. Celle-ci nous a permis d'évaluer les avantages et les limites de ces méthodes à partir des données générées avec la HuGeneFL[®] sur notre série initiale. Nous avons exclu l'utilisation de PGE « *Positional gene enrichment* » [275], GSEA « *Gene set enrichment analysis* » [276], GSA « *Gene set analysis* » [277], CHARM « *Chromosomal aberration region miner* » [278] et DIGMAP « *Differential gene locus mapping* » [279] et nous avons retenu la fonction geneSetTest du module LIMMA.

En effet, la **PGE** s'applique sur un ensemble de gènes identifiés par une analyse différentielle. Elle va donc dépendre en grande partie du seuil choisi pour la détermination de l'ensemble de gènes à tester. Elle pourrait ainsi engendrer des faux négatifs, tel qu'il a été rapporté dans [276]. En outre et surtout, elle ne prend pas en charge les données générées par la biopuce HuGeneFL[®]. La **GSEA** ainsi que sa version améliorée la **GSA** consomment du temps machine et exigent un minimum de six échantillons par classe. Or les effectifs de toutes nos classes sont au-dessous de six. La **CHARM** identifie les régions amplifiées ou délétées dans chaque échantillon séparément. Elle ne permet donc pas de localiser les régions coréglées associées à une mutation d'un gène donné et de ce fait elle ne répond pas à notre objectif. Afin de pouvoir réaliser une représentation graphique des régions enrichies en gènes différentiellement exprimés, les auteurs de **DIGMAP** (Yi et al.) ont mis en place la base de données GAP « *Gene Annotation Project* » [279]. Celle-ci contient les annotations des gènes provenant des bases de données du NCBI (<http://www.ncbi.nlm.nih.gov/>) et de GoldenPath (<http://genome.ucsc.edu/>). Nous avons

rencontré plusieurs problèmes pendant l'application de DIGMAP sur une autre série de données générées à partir de la biopuce HG-U133 Plus 2[®]. La majorité correspondait à des problèmes d'annotation provenant de GAP. Nous avons remarqué que plusieurs localisations chromosomiques étaient attribuées à un seul ES, plusieurs « *GeneBank Accession Number* » correspondaient à un seul « *Unigene* » et vice versa. De plus et surtout, cette base de données n'est pas mise à jour sur une base régulière.

L'avantage de la fonction *geneSetTest*, que nous avons finalement utilisé pour identifier les régions enrichies en gènes différentiellement exprimés chez les porteuses d'une mutation de *BRCA1* dans les cellules NOSEs, réside dans sa capacité à détecter une expression différentielle avec un petit nombre d'échantillons ou de gènes. En plus de sa rapidité d'exécution, elle nous permet d'obtenir une précision accrue pour identifier les régions enrichies en gènes différentiellement exprimés en définissant les fenêtres balayées par leurs coordonnées en bases. De plus, elle est purement compétitive. En d'autres termes, elle effectue une comparaison entre le niveau d'expression de l'ensemble des ES à tester (dans notre analyse 10 ES) et celui des ES n'appartenant pas à cet ensemble (le reste des ES).

Une des limites de *geneSetTest* est qu'elle semblerait plus adaptée à des classes d'échantillons ayant entre elles un nombre réduit de variations génotypiques, telles que des échantillons provenant de souris congéniques. Cependant, nous pouvons considérer que les variations de nos échantillons sont limitées du fait de l'origine ethnique commune de nos participants qui sont tous des Canadiens Français. Une autre limite réside dans la contrainte imposée a priori pour pouvoir appliquer cette méthode : ceci suppose que les gènes de l'ensemble à tester sont indépendants pour pouvoir ensuite tester si l'hypothèse nulle s'applique. Ce postulat d'indépendance est raisonnable si l'ensemble de gènes en question est constitué d'un petit nombre de gènes. Nous avons ainsi initialement limité la taille de la fenêtre balayée à 10 ES contigus sur la HuGeneFL[®] pour répondre à cette contrainte.

Nous avons rencontré deux limites principales à cette méthode : d'une part la faible densité de la biopuce générant nos données d'expression, d'autre part la façon dont nous avons défini initialement la résolution des fenêtres balayées. En effet, la HuGeneFL[®] quantifie seulement ~6 800 gènes. Or ~25 000 gènes ont été séquencés par le Projet Génome Humain (PGH) [280]. De ce fait, une région pourrait être jugée comme étant non enrichie en gènes différentiellement exprimés alors qu'en réalité elle l'est, du fait du nombre limité de gènes quantifiés avec la HuGeneFL[®], problème pouvant constituer une source importante de faux

négatifs. Par contre, une région identifiée comme étant enrichie en gènes différentiellement exprimés en se basant sur des résultats générés avec la HuGeneFL[®] pourrait correspondre à un faux positif en terme de RCC. Cet artefact peut être dû au fait que le nombre de gènes dans la région testée est significativement plus petit que le nombre de gènes existant réellement dans cette région. Ainsi, il est possible que trois gènes différentiellement exprimés appartiennent apparemment à une RCC alors qu'en réalité ils ne le sont pas car ils sont séparés par d'autres gènes.

D'autre part, dix ES contigus mesurés par une biopuce à faible densité telle que la HuGeneFL[®] peuvent correspondre à des distances très variables si nous considérons les coordonnées des régions chromosomiques définies par les ES situés à leurs bornes. Une fenêtre couvrant une large distance pourrait être jugée comme étant non enrichie en gènes différentiellement exprimés, alors qu'une partie des gènes localisés dans cette fenêtre font effectivement partie d'une région coréglée non identifiée, du fait de la faible densité des ES dans la fenêtre ainsi définie. Nous envisageons d'améliorer la façon dont nous définissons les fenêtres à tester. Une des solutions pourrait être de déterminer une largeur fixe pour chacune des fenêtres, largeur définie en bases. Cependant, cette largeur devrait être modifiée en fonction du chromosome balayé car la densité en gènes varie d'un chromosome à l'autre.

Les données de la littérature peuvent nous aider à redéfinir la largeur de ces fenêtres, en tenant compte du chromosome balayé. La fenêtre ainsi définie aurait une largeur moyenne de 3,66 Mb et des extrêmes variants entre 0,72 Mb pour le chromosome 19 et 5,69 Mb pour le chromosome 13 avec une moyenne de 6 gènes par fenêtre [281]. En outre, il serait nécessaire d'utiliser des intervalles de balayage se chevauchant pour être capable d'identifier une RCC dans le cas où elle serait à cheval sur deux fenêtres prédéfinies. Nous pensons que ces améliorations pourraient réduire le nombre des faux négatifs et améliorer nos chances de trouver d'éventuelles RCC. Bien entendu, l'utilisation d'une puce à plus haute densité constitue un élément important de solution que nous envisageons donc pour améliorer la résolution de la méthode. Nous avons déjà débuté une autre étude utilisant la biopuce HG-U133 Plus 2[®].

D.3. Discussion des résultats

D.3.1. Existe-t'il des signatures moléculaires associées à une mutation sur *BRCA1* ou *BRCA2* dans les cellules NOSEs et TOVs des porteuses ?

Nous avons identifié pour la première fois un profil d'expression différentielle signature de la présence d'une mutation de *BRCA1* dans les cellules NOSEs et TOVs ainsi qu'un profil d'expression différentielle signature de la présence d'une mutation de *BRCA2* dans les cellules NOSEs [282]. Les transcrits dont l'expression était dérégulée dans les échantillons des porteuses de mutation de *BRCA1* comparés aux échantillons non mutés l'étaient aussi dans les cellules tumorales des porteuses de mutation de *BRCA1* comparés aux mutées sur *BRCA1*. Ceci suggère que les gènes correspondant à ces transcrits sont impliqués dans les étapes précoces du processus de carcinogenèse associée à *BRCA1* [282].

Parmi les gènes identifiés dans notre analyse différentielle, nous avons choisi quatre gènes candidats susceptibles d'avoir un rôle important dans la carcinogenèse ovarienne en raison de leur fonction biologique, des données antérieures de la littérature témoignant de leur implication dans la carcinogenèse et bien entendu parce qu'ils étaient différentiellement exprimés avec un q statistiquement significatif.

Nous avons montré que *SAA2* « *Serum Amyloid A 2* » est surexprimé dans les échantillons F-NOSE1 comparés aux NF-NOSEs (X51441_at : $q = 0,017$ et $\log(\text{Fold Change}) = 3.456$, X51441_s_at : $q = 0,045$ et $\log(\text{Fold Change}) = 2.308$) et est significativement sous-exprimé dans les F-TOV1 comparé aux F-NOSE1 (J03474_at : $q = 0,017$ et $\log(\text{Fold Change}) = -2.828$, X51441_at : $q = 0,010$ et $\log(\text{Fold Change}) = -3.286$, X51441_s_at : $q = 0,013$ et $\log(\text{Fold Change}) = -2.341$). La famille *SAA* contient trois gènes et un pseudogène localisés en 11p15.1 [283]. *SAA1* et *SAA2* partagent une homologie de 96 %. Ils codent pour des apolipoprotéines *SAA* de phase aiguë (A-SAA). De plus, ils sont exprimés dans l'épithélium de plusieurs tissus normaux [284]. Les protéines A-SAA ont également été identifiées comme marqueurs tumoraux sériques potentiels dans les cancers du poumon [285], des ovaires [286] et de la prostate [287]. Il est important de noter que l'expression du gène *SAA1* est modifiée de façon significative dans la muqueuse rectosigmoïdienne macroscopiquement normale des individus ayant une histoire familiale de cancer du côlon [288]. Nos données sont en plein accord avec les données de la

littérature citées plus haut, suggérant que ces gènes pourraient également être impliqués à des stades précoces de la carcinogenèse héréditaire ovarienne.

FILIPIL « *Filamin A Interacting Protein 1-Like* » était clairement sous-exprimé dans les échantillons F-NOSE2 comparés aux NF-NOSEs (U53445_at : $q = 0,034$ et $\log(\text{Fold Change}) = -2.047$). Ce gène serait impliqué dans l'inhibition de l'angiogenèse [289]. *FILIPIL* a été appelé « *Decreased in Ovarian Cancer* » après son identification grâce à une approche soustractive comme l'un des transcrits sous-exprimés dans les lignées cellulaires de cancer de l'ovaire comparées aux cellules OSEs [290]. Il est également sous-exprimé dans les cellules de prostate humaine immortalisées par comparaison aux cellules sénescents. De plus, il a été observé que l'expression de *FILIPIL* serait perdue dans les lignées cellulaires du cancer de la prostate et dans les xénogreffes humaines [291]. Les données d'une étude récente utilisant le transfert chromosomique ont suggéré que la région chromosomique où est localisé *FILIPIL* pourrait jouer un rôle important dans la suppression de cancers sporadiques de l'ovaire [292]. Nos données corroborent ces résultats et indiquent un rôle potentiel du gène *FILIPIL* dans la carcinogenèse précoce associée à *BRCA2*.

ANXA8L2 « *Annexin A8-like 2* » appartient à une famille multigénique codant pour des protéines de liaison calcium/phospholipide. Ces peptides partagent un domaine structural commun appelé le noyau d'annexine composé de quatre copies ou plus d'annexine répétées. Celui-ci est responsable de l'activité de la liaison de l'ANXA8L2 à la membrane cellulaire, activité régulée par la concentration de l'ion Ca^{+2} [293]. Il a été démontré que l'ANXA8 possède des propriétés de liaison aux phospholipides et à la F-actine. Ceci suggère son implication dans l'organisation structurale des domaines membranaires associés à l'actine [294]. Le rôle de ce gène dans l'involution de la glande mammaire a été démontré chez des souris. Chez l'homme, la protéine ANXA8 est coexprimée avec CK5 et CK17 dans un sous-groupe de carcinomes mammaires humains, les carcinomes de type basal avec un faible taux de survie. Il est à noter que 85,7 % des tumeurs du sein associées à *BRCA1*, coexpriment *ANXA8* et *CK5* [41]. Nos résultats indiquent une association possible du gène *ANXA8* avec la transformation tumorale ovarienne chez les porteuses de mutation sur *BRCA1*.

Finalement, **MPI** « *Mannose Phosphate Isomerase* », purifié du placenta en 1994, code pour une protéine de 423 acides aminés relativement bien conservée [43]. À travers sa fonction de catalyseur pour la réaction d'interconversion entre le fructose-6-phosphate et le

mannose-6-phosphate, *MPI* joue un rôle majeur dans la maturation post-traductionnelle de plusieurs protéines. Des mutations de ce gène sont impliquées dans certaines maladies héréditaires, telles que les anomalies congénitales de la glycosylation de type Ib [295]. En 1920, Warburg a identifié un phénotype biochimique particulier (ou distinct) dans les cellules tumorales qui développent un métabolisme énergétique aberrant augmentant l'activation de la voie de la glycolyse pour accroître la production d'adénosine triphosphate (ATP). La surexpression de *MPI* dans les cellules pourrait ainsi augmenter la production d'énergie par la cellule et conférer ainsi un avantage sélectif aux cellules qui en font l'objet, tel que les NOSEs ici. Ces résultats mettent en évidence la nécessité de mieux élucider les événements moléculaires clés impliqués dans l'« effet Warburg » [296].

Les différences retrouvées entre les résultats de quantification de *MPI* obtenus avec la HuGeneFL[®] (Tableau VI) et la q-RT-PCR pour nos échantillons F-TOV1 (Fig. 16D) montrent la nécessité d'étudier et d'identifier de façon plus précise les profils d'épissage alternatif de ce gène dans les tumeurs ovariennes, lesquels profils pourraient montrer de grandes variations dans les tumeurs, expliquant les différences observées entre les deux méthodes de quantification.

Ces données confirment l'existence de profils moléculaires associés à la présence d'une mutation de *BRCA1* ou de *BRCA2* dans les cellules NOSEs. De plus, elles suggèrent que plusieurs des gènes identifiés par notre analyse différentielle pourraient représenter des partenaires putatifs de *BRCA1* et *BRCA2*, intervenant dans les premières étapes de la carcinogenèse chez les femmes porteuses d'une mutation.

D.3.2. L'hétérozygotie pour une mutation sur *BRCA2* dans les cellules NOSEs et TOVs est-elle suffisante à l'initiation de la carcinogenèse ?

Les profils d'expression que nous avons identifiés dans nos échantillons ont été non seulement caractérisés pour la présence d'une mutation sur *BRCA2* mais aussi pour leur niveau d'expression de *BRCA2*. Nous avons observé que les allèles sauvage et muté sont exprimés dans tous nos échantillons mutés sur *BRCA2*. L'allèle sauvage est également retenu et exprimé dans les échantillons tumoraux provenant des porteuses de la mutation *BRCA2-8765delAG*.

Jusqu'à récemment, il était admis qu'une perte d'hétérozygotie aux loci *BRCA1* et *BRCA2* était nécessaire pour la transformation tumorale mammaire et ovarienne. Afin de préciser si

une perte d'hétérozygotie était présente dans ces tumeurs, la plupart des études réalisées précédemment étaient basées sur des examens autoradiographiques ou des méthodes semi-quantitatives de fluorescence. Or les examens autoradiographiques présentent des limites dans l'interprétation des données du fait de leur manque de sensibilité et les méthodes semi-quantitatives de fluorescence reposent sur une définition arbitraire du ratio d'intensité en deçà duquel on parle de PDH pour l'allèle sauvage ce qui pourrait nuire à l'interprétation des résultats obtenus avec ces méthodes.

Récemment, King et al. ont démontré que la perte de l'allèle sauvage de *BRCA1* et *BRCA2* n'était pas nécessaire à l'initiation de la carcinogenèse mammaire associée aux gènes *BRCA*. Cependant, ils ont retrouvé une perte d'hétérozygotie chez leurs patientes atteintes de cancers de l'ovaire [216]. Avec une approche utilisant la quantification spécifique d'allèle basée sur la mesure par PCR en temps réel, Chen et al. ont montré qu'un déséquilibre d'expression des transcrits alléliques de *BRCA1* et à un moindre degré de *BRCA2* pourrait contribuer au développement de cancers du sein qu'ils soient ou non familiaux [234]. L'absence de perte d'hétérozygotie a été rapportée précédemment dans une tumeur ovarienne germinale survenue chez une porteuse de la mutation *BRCA2-6174delT*. Cependant, cet histotype inhabituel n'a pas été considéré comme faisant partie du phénotype clinique, ce type de tumeur étant rapporté pour la première fois chez une porteuse de mutation de *BRCA2* [217]. Dans le cadre de notre étude, la quantification spécifique d'allèle nous a permis de montrer, pour la première fois, que l'allèle sauvage était transcrit dans les cellules adénocarcinomeuses ovariennes de Canadiennes Françaises porteuses de la mutation *BRCA2-8765delAG* [282]. Après avoir démontré l'absence de perte d'hétérozygotie au locus *BRCA2* dans nos échantillons tumoraux provenant de donneurs porteurs de mutation sur ce gène, la question qui se pose alors est de savoir si nos observations sont dues à un phénomène d'haplo-insuffisance de *BRCA2* ou à un effet de dosage de l'allèle muté de ce gène.

Patel et al. [297] ont montré qu'à la différence des lignées cellulaires de fibroblastes embryonnaires de souris homozygotes pour *Brca2*, celles de souris hétérozygotes pour une mutation sur *Brca2* ne présentent pas de déficience dans la réparation de l'ADN in vitro, suggérant ainsi l'absence d'un phénomène d'haplo-insuffisance pour le gène *BRCA2*. Une surexpression de *BRCA2* avait été observée auparavant dans les cancers du sein sporadiques de grade élevé [298] ainsi que dans les adénocarcinomes ovariens [87]. Nous avons pour notre part observé que l'agressivité des tumeurs ovariennes chez les porteuses d'une

mutation germinale de *BRCA2* était associée à une surexpression de ce gène (Tableau V et Fig. 17B). Nous avons de plus mis en évidence un niveau élevé du transcrit allélique muté dans les cancers héréditaires. La surexpression de *BRCA2* dans nos échantillons tumoraux mutés sur *BRCA2* ne serait pas en faveur d'un phénomène d'haplo-insuffisance pour ce gène, du moins en ce qui concerne les mutations étudiées ici. De plus, la surexpression de l'allèle muté de *BRCA2* dans ces mêmes échantillons soutiendrait l'hypothèse qu'un effet de dosage de l'allèle muté pourrait avoir un effet négatif dominant se traduisant par un gain de fonction au lieu d'une perte de fonction. Cette hypothèse permettrait d'interpréter nos observations quant à l'association entre la surexpression de l'allèle muté de *BRCA2* et l'agressivité des tumeurs ovariennes chez nos porteuses d'une mutation germinale de *BRCA2*.

Nos résultats suggèrent que l'hétérozygotie pour une mutation de *BRCA2* pourrait être suffisante pour l'initiation de la carcinogenèse ovarienne. De plus, un effet de dosage, lié non seulement au niveau d'expression total de *BRCA2* mais aussi et surtout à la contribution relative du transcrit muté, pourrait déterminer la pénétrance d'une mutation de ce gène et l'agressivité des tumeurs qui y sont associées.

D.3.3. Existe-t'il des domaines de co-expression associés à une mutation du gène *BRCA1* dans les cellules NOSEs ?

Il était admis jusqu'alors qu'un gène était sous le contrôle d'éléments « *enhancers* » et répresseurs localisés relativement près de l'unité de transcription (régulation en cis) responsable de l'expression spatio-temporelle de ces gènes pendant le développement ou le fonctionnement d'un tissu ou un organe donné. Cependant, les données de la génomique sont venues remettre en cause ces concepts.

Récemment, plusieurs études s'intéressant à la relation entre l'expression des gènes et l'organisation du génome ont montré que les gènes adjacents avaient tendance à avoir le même profil d'expression [237-240, 266]. En effet, des régions chromosomiques ordonnées le long du génome contenant des gènes coordonnés, correspondant à la co-expression de domaines, ont été retrouvées dans des microorganismes [299] et chez l'homme [300]. Ces régions représentent des régions de contrôle global (RCG) très conservées qui sont capables de contrôler plusieurs gènes distincts dans leur structure et leur fonction sur une large distance. En effet, des éléments régulateurs en cis ont été trouvés situés à plusieurs

centaines de Kb, parfois à plus de 800 kb, de leur gène cible principal, régulant différents gènes ne présentant pas la même fonction [301]. Ces régions régulatrices diffèrent des complexes (clusters) de gènes précédemment décrits comme ceux de la globine, car les gènes qui partagent ces éléments régulateurs à longue distance ne sont pas liés entre eux fonctionnellement ni issus de phénomènes de duplication génique. Il semble y avoir une relation entre la taille physique du cluster et la complexité de l'organisme, avec une taille des clusters qui varie de 10 kb chez la levure à plusieurs mégabases chez les mammifères [302].

L'identification de RCG dans les cellules NOSEs provenant des porteuses d'une mutation sur le gène *BRCA1* pourrait améliorer notre compréhension des événements précoces de la carcinogenèse ovarienne. Cependant, ceci nécessite en premier lieu de localiser les domaines de co-expression, ou les régions coréglées colocalisées, contrôlés par ces RCG. Nous avons identifié six régions enrichies en gènes coexprimés contenant une ou plusieurs régions coréglées colocalisées dont une région comportant des gènes surexprimés et cinq régions comportant des gènes sous-exprimés dans les cellules F-NOSE1.

Notre étude est la première à s'intéresser aux régions enrichies en gènes différentiellement exprimés dans le but d'identifier des RCC dans les cellules ovariennes des porteuses d'une mutation de *BRCA1*. Nous avons montré que les régions chromosomiques correspondant aux bandes 8p21 et 16q22 étaient enrichies en gènes sous-exprimés chez les porteuses d'une mutation de *BRCA1*.

Parmi les régions que nous avons identifiées comme coréglées, certaines ont été identifiées dans la littérature comme étant aussi le siège de délétion. En utilisant l'hybridation génomique comparative, Zweemer et al. [303] ont mis en évidence des variations du nombre de copies de l'ADN chez 36 patientes atteintes de cancers de l'ovaire et ayant une histoire familiale : 13 porteuses d'une mutation sur *BRCA1*, 2 sur *BRCA2* et 21 non-porteuses de mutation. Ainsi, les bandes chromosomiques 8p21-pter et 16q22-pter sont parmi les plus fréquemment délétées dans les tumeurs ovariennes héréditaires : 8p21-pter et 16q22-pter sont délétées respectivement dans 23 et 19 sur 36 tumeurs ovariennes [303]. Une délétion de la bande 8p21-pter a été également rapportée avec une fréquence élevée par Tapper et al. [304]. Cette région contient le suppresseur de tumeurs *SFRP1* qui est impliqué dans la pathogenèse de plusieurs types de cancers, entre autres dans les cancers de la prostate [305] et du sein [306, 307].

Nous attirons l'attention du lecteur sur le fait que la région du chromosome 16 que nous avons identifiée comme étant enrichie en gènes sous-exprimés dans les échantillons mutés sur *BRCA1* contient deux gènes *CDH* appartenant à la famille des cadhérines. Dans une étude portant cette fois sur les pertes d'hétérozygotie dans des tumeurs ovariennes incidentes, Sarrió et al. [308] ont observé une fréquence élevée (78%) de perte d'hétérozygotie en 16q22, bande chromosomique contenant plusieurs gènes de la famille des cadhérines jouant un rôle important dans la transition épithélio-mésenchymateuse. De plus, elle contient le gène *RRAD* interagissant avec l'oncogène *RAS*. Une méthylation de *RRAD* a fréquemment été observée dans les cancers du sein et des poumons [309].

Nous avons mis en évidence que la bande 19q13 était enrichie en gènes surexprimés dans les cellules NOSEs des porteuses de mutation de *BRCA1*. En utilisant la technique d'hybridation fluorescente (FISH), Thompson et al. ont montré que la région chromosomique 19q13.1-q13.2, contenant l'oncogène putatif *AKT2*, était amplifiée dans des lignées cellulaires de tumeurs ovariennes humaines [310]. Ils suggèrent que l'oncogène *AKT2* ainsi que d'autres oncogènes non identifiés appartenant à cette bande chromosomique pourraient contribuer à l'initiation ou la progression de cancers de l'ovaire. Nous soulignons qu'*AKT2* est localisé sur le brin négatif du chromosome 19 (45 428 064 bases), ainsi il est situé dans la région que nous avons identifiée, mais il n'est pas quantifié par la HuGeneFL[®].

En conclusion, trois des régions que nous avons identifiées comme étant enrichies en gènes différemment exprimés dans les cellules NOSEs de femmes porteuses d'une mutation de *BRCA1* ont été aussi identifiées comme des régions présentant des altérations moléculaires dans des cellules tumorales ovariennes dans d'autres études. Puisque nos échantillons sont des échantillons non tumoraux provenant de patientes prédisposées au cancer de l'ovaire, nous suggérons que ces trois régions pourraient jouer un rôle dans la carcinogenèse ovarienne précoce chez les porteuses d'une mutation de *BRCA1*.

D.4. Conclusions et perspectives

Les résultats présentés dans cette étude sont uniques car ils sont les premiers à confirmer l'existence des signatures moléculaires associées à la présence d'une mutation de *BRCA1* et *BRCA2* dans les cellules NOSEs. L'identification de ces profils est d'une grande importance clinique, notamment pour distinguer entre les patientes porteuses d'une

mutation sur *BRCA1* et celles porteuses d'une mutation sur *BRCA2* quand la mutation familiale n'a pas pu être déterminée même à la suite d'investigations extensives. La détermination des profils correspondant aux femmes porteuses de mutation sur *BRCA1* ou *BRCA2* dans les familles à risque, d'une part facilitera l'identification des patientes à risque, d'autre part, empêchera le biais introduit dans les études de liaison ou d'association visant à identifier de nouveaux gènes de prédisposition (*BRCAx*). Le phénotype des cellules tumorales dont proviennent nos échantillons dépend non seulement du niveau d'expression de *BRCA2* mais également du niveau d'expression du transcrit allélique muté. Ceci suggère un effet de dosage des transcrits des allèles de *BRCA2* sur le phénotype tumoral. Nous avons observé une concordance entre les profils d'expression de certains de nos gènes candidats et celui de *BRCA2* dans les échantillons mutés sur *BRCA2*. Ceci suggère que ces gènes seraient corégulés dans ces échantillons, nous orientant vers un partenaire putatif du gène *BRCA2*.

De plus, nous avons mis en évidence des RCCs qui correspondent à des régions chromosomiques où ont été identifiés antérieurement des délétions ou des amplifications dans les tumeurs.

En conclusion, non seulement des gènes ayant une activité régulatrice et/ou les modifications épigénétiques peuvent venir modifier le transcriptome des gènes *BRCA1* ou *BRCA2*, mais l'allèle muté de *BRCA2* lui-même par un effet de dosage pourrait rendre compte des différences interindividuelles observées dans les phénotypes cliniques chez les porteurs d'une même mutation.

Nous envisageons de confirmer l'existence des profils d'expression différentielle associés dans les cellules NOSEs à une mutation sur les gènes *BRCA1* et *BRCA2*, sur un nombre plus élevé d'échantillons et en utilisant la biopuce HG-U133 Plus 2[®] d'Affymetrix[™] qui a une densité plus élevée que la HuGeneFL[®].

De plus, nous voulons poursuivre l'identification des profils moléculaires et des RCC dans les lymphocytes du sang périphérique des femmes porteuses de mutation sur l'un de ces gènes.

Une approche intégrée visant à identifier à partir de données d'expression les régions comportant des gènes corégulés colocalisés en association à une mutation nous permettrait ultérieurement d'orienter la découverte d'haplotypes en déséquilibre avec ces régions

chromosomiques chez les porteuses de mutation et de poursuivre par l'identification d'haplotypes régulateurs. L'intégration des données d'expression, données des haplotypes et données cliniques, plus particulièrement le phénotype clinique des patientes devrait ouvrir la voie au développement de nouvelles méthodes de diagnostic précoce et de nouvelles stratégies de prévention.

Bibliographie

1. Tavtigian, S.V., et al., *The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds*. Nat Genet, 1996. **12**(3): p. 333-7.
2. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1*. Science, 1994. **266**(5182): p. 66-71.
3. Wooster, R., et al., *Identification of the breast cancer susceptibility gene BRCA2*. Nature, 1995. **378**(6559): p. 789-92.
4. Antoniou, A., et al., *Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies*. Am J Hum Genet, 2003. **72**(5): p. 1117-30.
5. Friedenson, B., *BRCA1 and BRCA2 pathways and the risk of cancers other than breast or ovarian*. MedGenMed, 2005. **7**(2): p. 60.
6. Rebbeck, T.R., N.D. Kauff, and S.M. Domchek, *Meta-analysis of risk reduction estimates associated with risk-reducing salpingo-oophorectomy in BRCA1 or BRCA2 mutation carriers*. J Natl Cancer Inst, 2009. **101**(2): p. 80-7.
7. Smith-Beckerman, D.M., et al., *Proteome changes in ovarian epithelial cells derived from women with BRCA1 mutations and family histories of cancer*. Mol Cell Proteomics, 2005. **4**(2): p. 156-68.
8. Shier, D., J. Butter, and R. Lewis, *Hole's essentials of human anatomy and physiology*. 6th Edition ed. 1998, Boston: McGraw-Hill.
9. Dauplat, J. and J.-P. Gastalla, *Cancer de l'ovaire*. 1996, Paris: Amette BlackWell.
10. Hould, R., *Histologie descriptive*, ed. Décarie. 1982.
11. Lecuru, F. *Cancer de l'ovaire*. Cancerologie 2003-2004 [cited; Cours - Hôpital Européen Georges Pampidou].
12. Weiss, N.S., T. Homonchuk, and J.L. Young, Jr., *Incidence of the histologic types of ovarian cancer: the U.S. Third National Cancer Survey, 1969-1971*. Gynecol Oncol, 1977. **5**(2): p. 161-7.
13. *Société canadienne du cancer et Institut national du cancer du Canada : Statistiques canadiennes sur le cancer 2008*. 2008: Toronto.
14. Ozols, R., et al., *Epithelial ovarian cancer*, in *Principles and practice of gynecologic oncology*, **W.J. Hoskins**, Editor. 2005, Lippincott Williams & Wilkins: Philadelphia. p. 895-988.
15. Teneriello, M.G. and R.C. Park, *Early detection of ovarian cancer*. CA Cancer J Clin, 1995. **45**(2): p. 71-87.
16. Purdie, D.M., et al., *Ovulation and risk of epithelial ovarian cancer*. Int J Cancer, 2003. **104**(2): p. 228-32.
17. Salehi, F., et al., *Risk factors for ovarian cancer: an overview with emphasis on hormonal factors*. J Toxicol Environ Health B Crit Rev, 2008. **11**(3-4): p. 301-21.
18. Hall, K.L., M.A. Dewar, and J. Perchalski, *Screening for gynecologic cancer. Vulvar, vaginal, endometrial, and ovarian neoplasms*. Prim Care, 1992. **19**(3): p. 607-20.
19. Bristow, R.E. and B.Y. Karlan, *Ovulation induction, infertility, and ovarian cancer risk*. Fertil Steril, 1996. **66**(4): p. 499-507.
20. Rossing, M.A., et al., *Ovarian tumors in a cohort of infertile women*. N Engl J Med, 1994. **331**(12): p. 771-6.
21. Riman, T., et al., *Hormone replacement therapy and the risk of invasive epithelial ovarian cancer in Swedish women*. J Natl Cancer Inst, 2002. **94**(7): p. 497-504.

22. Antoniou, A.C., et al., *Reproductive and hormonal factors, and ovarian cancer risk for BRCA1 and BRCA2 mutation carriers: results from the International BRCA1/2 Carrier Cohort Study*. *Cancer Epidemiol Biomarkers Prev*, 2009. **18**(2): p. 601-10.
23. McLaughlin, J.R., et al., *Reproductive risk factors for ovarian cancer in carriers of BRCA1 or BRCA2 mutations: a case-control study*. *Lancet Oncol*, 2007. **8**(1): p. 26-34.
24. Friedman, L.C. and R.M. Kramer, *Reproductive issues for women with BRCA mutations*. *J Natl Cancer Inst Monogr*, 2005(34): p. 83-6.
25. Andrieu, N., et al., *Pregnancies, breast-feeding, and breast cancer risk in the International BRCA1/2 Carrier Cohort Study (IBCCS)*. *J Natl Cancer Inst*, 2006. **98**(8): p. 535-44.
26. Bandera, E.V., *Nutritional factors in ovarian cancer prevention: what have we learned in the past 5 years?* *Nutr Cancer*, 2007. **59**(2): p. 142-51.
27. Byers, T., et al., *American Cancer Society guidelines on nutrition and physical activity for cancer prevention: Reducing the risk of cancer with healthy food choices and physical activity*. *CA Cancer J Clin*, 2002. **52**(2): p. 92-119.
28. Booth, M., V. Beral, and P. Smith, *Risk factors for ovarian cancer: a case-control study*. *Br J Cancer*, 1989. **60**(4): p. 592-8.
29. Donna, A., et al., *Triazine herbicides and ovarian epithelial neoplasms*. *Scand J Work Environ Health*, 1989. **15**(1): p. 47-53.
30. Hartge, P. and P. Stewart, *Occupation and ovarian cancer: a case-control study in the Washington, DC, metropolitan area, 1978-1981*. *J Occup Med*, 1994. **36**(8): p. 924-7.
31. Walrath, J., et al., *Causes of death among female chemists*. *Am J Public Health*, 1985. **75**(8): p. 883-5.
32. Cook, J., *Family history of ovarian cancer*. *Current Obstetrics & Gynaecology*, 2002. **12**: p. 47-51.
33. Lortholary, A., et al., *Le point sur les prédispositions génétiques aux cancers du sein et de l'ovaire*, in *Dépistage des cancers gynécologiques et mammaires*, Masson, Editor. 2004, Elsevier Masson: Paris. p. 160.
34. Prat, J., A. Ribé, and A. Gallardo, *Hereditary ovarian cancer*. *Human Pathology*, 2005. **36**(8): p. 861-870.
35. Lynch, H.T., T. Smyrk, and J. Lynch, *An update of HNPCC (Lynch syndrome)*. *Cancer Genet Cytogenet*, 1997. **93**(1): p. 84-99.
36. Kasprzak, L., W.D. Foulkes, and A.N. Shelling, *Fortnightly review: Hereditary ovarian carcinoma*. *BMJ*, 1999. **318**(7186): p. 786-789.
37. Padykulha, H., *The female reproductive system in Cell and Tissue Biology*, L. Weiss, Editor. 1988, Urban and Schwarzenberg: Baltimore. p. 851-878.
38. Radisavljevic, S.V., *The pathogenesis of ovarian inclusion cysts and cystomas*. *Obstet Gynecol*, 1977. **49**(4): p. 424-9.
39. Mittal, K.R., et al., *Contralateral ovary in unilateral ovarian carcinoma: a search for preneoplastic lesions*. *Int J Gynecol Pathol*, 1993. **12**(1): p. 59-63.
40. Cramer, D.W. and W.R. Welch, *Determinants of ovarian cancer risk. II. Inferences regarding pathogenesis*. *J Natl Cancer Inst*, 1983. **71**(4): p. 717-21.
41. Berchuck, A., M.F. Kohler, and R.C. Bast, Jr., *Oncogenes in ovarian cancer*. *Hematol Oncol Clin North Am*, 1992. **6**(4): p. 813-27.
42. Okamoto, A., S. Yokoyama, and Y. Terashima, *[Genetic alterations in the genesis and development of ovarian cancer]*. *Gan To Kagaku Ryoho*, 1992. **19**(12): p. 1971-6.
43. Sowter, H.M. and A. Ashworth, *BRCA1 and BRCA2 as ovarian cancer susceptibility genes*. *Carcinogenesis*, 2005. **26**(10): p. 1651-6.

44. Stehelin, D., H.E. Varmus, and J.M. Bishop, *Detection of nucleotide sequences associated with transformation by avian sarcoma viruses*. *Bibl Haematol*, 1975(43): p. 539-41.
45. Stehelin, D., et al., *DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA*. *Nature*, 1976. **260**(5547): p. 170-3.
46. Shih, C., et al., *Transforming genes of carcinomas and neuroblastomas introduced into mouse fibroblasts*. *Nature*, 1981. **290**(5803): p. 261-4.
47. Theoleyre, O. and F. Baklouti, [*Oncogenic transcription factors as splicing regulators*]. *Med Sci (Paris)*, 2004. **20**(11): p. 1014-9.
48. Guasch, G., et al., *FOP-FGFR1 tyrosine kinase, the product of a t(6;8) translocation, induces a fatal myeloproliferative disease in mice*. *Blood*, 2004. **103**(1): p. 309-12.
49. Aunoble, B., et al., *Major oncogenes and tumor suppressor genes involved in epithelial ovarian cancer (review)*. *Int J Oncol*, 2000. **16**(3): p. 567-76.
50. Singer, G., et al., *Mutations in BRAF and KRAS characterize the development of low-grade ovarian serous carcinoma*. *J Natl Cancer Inst*, 2003. **95**(6): p. 484-6.
51. Cheng, J.Q., et al., *AKT2, a putative oncogene encoding a member of a subfamily of protein-serine/threonine kinases, is amplified in human ovarian carcinomas*. *Proc Natl Acad Sci U S A*, 1992. **89**(19): p. 9267-71.
52. Shih Ie, M. and R.J. Kurman, *Ovarian tumorigenesis: a proposed model based on morphological and molecular genetic analysis*. *Am J Pathol*, 2004. **164**(5): p. 1511-8.
53. Mok, S.C., et al., *Mutation of K-ras protooncogene in human ovarian epithelial tumors of borderline malignancy*. *Cancer Res*, 1993. **53**(7): p. 1489-92.
54. Tashiro, H., et al., *c-myc over-expression in human primary ovarian tumours: its relevance to tumour progression*. *Int J Cancer*, 1992. **50**(5): p. 828-33.
55. Ross, J.S., et al., *HER-2/neu oncogene amplification by fluorescence in situ hybridization in epithelial tumors of the ovary*. *Am J Clin Pathol*, 1999. **111**(3): p. 311-6.
56. Singer, G., et al., *HLA-G Is a Potential Tumor Marker in Malignant Ascites*. *Clinical Cancer Research*, 2003. **9**: p. 4460-4464.
57. Rhei, E., et al., *Molecular genetic characterization of BRCA1- and BRCA2-linked hereditary ovarian cancers*. *Cancer Res*, 1998. **58**(15): p. 3193-6.
58. Knudson, A.G., Jr., *Mutation and cancer: statistical study of retinoblastoma*. *Proc Natl Acad Sci U S A*, 1971. **68**(4): p. 820-3.
59. Comings, D.E., *A general theory of carcinogenesis*. *Proc Natl Acad Sci U S A*, 1973. **70**(12): p. 3324-8.
60. Friend, S.H., et al., *A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma*. *Nature*, 1986. **323**(6089): p. 643-6.
61. Lee, W.H., et al., *Human retinoblastoma susceptibility gene: cloning, identification, and sequence*. *Science*, 1987. **235**(4794): p. 1394-9.
62. Huang, H.J., et al., *Suppression of the neoplastic phenotype by replacement of the RB gene in human cancer cells*. *Science*, 1988. **242**(4885): p. 1563-6.
63. Lacave, R., C.-J. Larsen, and J. Robert, *Cancérologie fondamentale*. Collection Société française du cancer ;. 2005, Montrouge: J. Libbey-Eurotext. xiv, 437 p.
64. Scully, R., et al., *Genetic analysis of BRCA1 function in a defined tumor cell line*. *Mol Cell*, 1999. **4**(6): p. 1093-9.
65. Abaji, C., I. Cousineau, and A. Belmaaza, *BRCA2 regulates homologous recombination in response to DNA damage: implications for genome stability and carcinogenesis*. *Cancer Res*, 2005. **65**(10): p. 4117-25.

66. Matias-Guiu, X. and J. Prat, *Molecular pathology of ovarian carcinomas*. Virchows Arch, 1998. **433**(2): p. 103-11.
67. Aghmesheh, M., et al., *No differences in p53 mutation frequencies between BRCA1-associated and sporadic ovarian cancers*. Gynecol Oncol, 2004. **95**(3): p. 430-6.
68. Knudson, A.G., *Two genetic hits (more or less) to cancer*. Nat Rev Cancer, 2001. **1**(2): p. 157-62.
69. Cavenee, W.K., et al., *Expression of recessive alleles by chromosomal mechanisms in retinoblastoma*. Nature, 1983. **305**(5937): p. 779-84.
70. Knudson, A.G., Jr., et al., *Chromosomal deletion and retinoblastoma*. N Engl J Med, 1976. **295**(20): p. 1120-3.
71. Lee, W.H., et al., *The retinoblastoma susceptibility gene encodes a nuclear phosphoprotein associated with DNA binding activity*. Nature, 1987. **329**(6140): p. 642-5.
72. Janatova, M., et al., *Novel somatic mutations in the BRCA1 gene in sporadic breast tumors*. Hum Mutat, 2005. **25**(3): p. 319.
73. Kenemans, P., R.A. Verstraeten, and R.H. Verheijen, *Oncogenic pathways in hereditary and sporadic breast cancer*. Maturitas, 2004. **49**(1): p. 34-43.
74. Merajver, S.D., et al., *Somatic mutations in the BRCA1 gene in sporadic ovarian tumours*. Nat Genet, 1995. **9**(4): p. 439-43.
75. Hosking, L., et al., *A somatic BRCA1 mutation in an ovarian tumour*. Nat Genet, 1995. **9**(4): p. 343-4.
76. Takahashi, H., et al., *Mutations of the BRCA2 gene in ovarian carcinomas*. Cancer Res, 1996. **56**(12): p. 2738-41.
77. Foster, K.A., et al., *Somatic and germline mutations of the BRCA2 gene in sporadic ovarian cancer*. Cancer Res, 1996. **56**(16): p. 3622-5.
78. Thompson, M.E., et al., *Decreased expression of BRCA1 accelerates growth and is often present during sporadic breast cancer progression*. Nat Genet, 1995. **9**(4): p. 444-50.
79. Dinesh, K.P., et al., *Analysis of loss of heterozygosity and immunohistochemistry in BRCA1 gene in sporadic breast cancers*. Mol Cell Biochem, 2006. **287**(1-2): p. 177-83.
80. Zheng, W., et al., *Reduction of BRCA1 expression in sporadic ovarian cancer*. Gynecol Oncol, 2000. **76**(3): p. 294-300.
81. Mueller, C.R. and C.D. Roskelley, *Regulation of BRCA1 expression and its relationship to sporadic breast cancer*. Breast Cancer Res, 2003. **5**(1): p. 45-52.
82. Hilton, J.L., et al., *Inactivation of BRCA1 and BRCA2 in ovarian cancer*. J Natl Cancer Inst, 2002. **94**(18): p. 1396-406.
83. Zikan, M., et al., *High frequency of BRCA1/2 and p53 somatic inactivation in sporadic ovarian cancer*. J Genet, 2007. **86**(2): p. 169-71.
84. Russell, P.A., et al., *Frequent loss of BRCA1 mRNA and protein expression in sporadic ovarian cancers*. Int J Cancer, 2000. **87**(3): p. 317-21.
85. Wang, C., et al., *Expression of BRCA1 protein in benign, borderline, and malignant epithelial ovarian neoplasms and its relationship to methylation and allelic loss of the BRCA1 gene*. J Pathol, 2004. **202**(2): p. 215-23.
86. Hughes-Davies, L., et al., *EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer*. Cell, 2003. **115**(5): p. 523-35.
87. Chan, K.Y., et al., *Epigenetic factors controlling the BRCA1 and BRCA2 genes in sporadic ovarian cancer*. Cancer Res, 2002. **62**(14): p. 4151-6.

88. Gras, E., et al., *Loss of heterozygosity on chromosome 13q12-q14, BRCA-2 mutations and lack of BRCA-2 promoter hypermethylation in sporadic epithelial ovarian tumors*. *Cancer*, 2001. **92**(4): p. 787-95.
89. Bieche, I., C. Nogues, and R. Lidereau, *Overexpression of BRCA2 gene in sporadic breast tumours*. *Oncogene*, 1999. **18**(37): p. 5232-8.
90. Wessels, L.F., et al., *Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors*. *Cancer Res*, 2002. **62**(23): p. 7110-7.
91. van 't Veer, L.J., et al., *Gene expression profiling predicts clinical outcome of breast cancer*. *Nature*, 2002. **415**(6871): p. 530-6.
92. Hedenfalk, I., et al., *Gene-expression profiles in hereditary breast cancer*. *N Engl J Med*, 2001. **344**(8): p. 539-48.
93. Hedenfalk, I., et al., *Molecular classification of familial non-BRCA1/BRCA2 breast cancer*. *Proc Natl Acad Sci U S A*, 2003. **100**(5): p. 2532-7.
94. Hedenfalk, I.A., *Gene expression profiling of hereditary and sporadic ovarian cancers reveals unique BRCA1 and BRCA2 signatures*. *J Natl Cancer Inst*, 2002. **94**(13): p. 960-1.
95. Jazaeri, A.A., et al., *Gene expression profiles of BRCA1-linked, BRCA2-linked, and sporadic ovarian cancers*. *J Natl Cancer Inst*, 2002. **94**(13): p. 990-1000.
96. Jazaeri, A.A., et al., *BRCA1-mediated repression of select X chromosome genes*. *J Transl Med*, 2004. **2**(1): p. 32.
97. Kinzler, K.W. and B. Vogelstein, *Cancer-susceptibility genes. Gatekeepers and caretakers*. *Nature*, 1997. **386**(6627): p. 761, 763.
98. Michor, F., Y. Iwasa, and M.A. Nowak, *Dynamics of cancer progression*. *Nat Rev Cancer*, 2004. **4**(3): p. 197-205.
99. Macleod, K., *Tumor suppressor genes*. *Curr Opin Genet Dev*, 2000. **10**(1): p. 81-93.
100. Hooper, M.L., *Tumour suppressor gene mutations in humans and mice: parallels and contrasts*. *EMBO J*, 1998. **17**(23): p. 6783-9.
101. Wang, Y., et al., *BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures*. *Genes Dev*, 2000. **14**(8): p. 927-39.
102. Fearon, E.R. and B. Vogelstein, *A genetic model for colorectal tumorigenesis*. *Cell*, 1990. **61**(5): p. 759-67.
103. Kinzler, K.W. and B. Vogelstein, *Landscaping the cancer terrain*. *Science*, 1998. **280**(5366): p. 1036-7.
104. Lynch, E.D., et al., *Inherited mutations in PTEN that are associated with breast cancer, cowden disease, and juvenile polyposis*. *Am J Hum Genet*, 1997. **61**(6): p. 1254-60.
105. Lodish, H.F., *Biologie moléculaire de la cellule*. 3e éd. ed. 2005, Bruxelles: De Boeck. xxxii, 973, [78] p.
106. Wooster, R., et al., *Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13*. *Science*, 1994. **265**(5181): p. 2088-90.
107. Reedy, M.B., et al., *Antisense inhibition of BRCA1 expression and molecular analysis of hereditary tumors indicate that functional inactivation of the p53 DNA damage response pathway is required for BRCA-associated tumorigenesis*. *Gynecol Oncol*, 2001. **81**(3): p. 441-6.
108. Vega, A., et al., *Analysis of BRCA1 and BRCA2 in breast and breast/ovarian cancer families shows population substructure in the Iberian peninsula*. *Ann Hum Genet*, 2002. **66**(Pt 1): p. 29-36.

109. Godwin, A., T. Hamilton, and A. Knudson, *Oncogenes and anti-oncogenes in gynecologic malignancies*, in *Principles and practice of gynecologic oncology*, **W.J. Hoskins**, Editor. 2005, Lippincott Williams & Wilkins: Philadelphia. p. 87-116.
110. Shows, T.B., et al., *International System for Human Gene Nomenclature (1979) ISGN (1979)*. Birth Defects Orig Artic Ser, 1980. **15**(11): p. 96-116.
111. Hall, J.M., et al., *Linkage of early-onset familial breast cancer to chromosome 17q21*. Science, 1990. **250**(4988): p. 1684-9.
112. Orban, T.I. and E. Olah, *Emerging roles of BRCA1 alternative splicing*. Mol Pathol, 2003. **56**(4): p. 191-7.
113. Lu, M., et al., *Characterization of functional messenger RNA splice variants of BRCA1 expressed in nonmalignant and tumor-derived breast cells*. Cancer Res, 1996. **56**(20): p. 4578-81.
114. Xu, C.F., et al., *Mutations and alternative splicing of the BRCA1 gene in UK breast/ovarian cancer families*. Genes Chromosomes Cancer, 1997. **18**(2): p. 102-10.
115. Fortin, J., et al., *A new alternative splice variant of BRCA1 containing an additional in-frame exon*. Biochim Biophys Acta, 2005. **1731**(1): p. 57-65.
116. Venkitaraman, A.R., *Cancer susceptibility and the functions of BRCA1 and BRCA2*. Cell, 2002. **108**(2): p. 171-82.
117. Scully, R., et al., *Association of BRCA1 with Rad51 in mitotic and meiotic cells*. Cell, 1997. **88**(2): p. 265-75.
118. Paull, T.T., et al., *A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage*. Curr Biol, 2000. **10**(15): p. 886-95.
119. Zhong, Q., et al., *Association of BRCA1 with the hRad50-hMre11-p95 complex and the DNA damage response*. Science, 1999. **285**(5428): p. 747-50.
120. Deng, C.X. and S.G. Brodie, *Roles of BRCA1 and its interacting proteins*. Bioessays, 2000. **22**(8): p. 728-37.
121. Scully, R. and N. Puget, *BRCA1 and BRCA2 in hereditary breast cancer*. Biochimie, 2002. **84**(1): p. 95-102.
122. Yuan, S.S., et al., *BRCA2 is required for ionizing radiation-induced assembly of Rad51 complex in vivo*. Cancer Res, 1999. **59**(15): p. 3547-51.
123. Lomonosov, M., et al., *Stabilization of stalled DNA replication forks by the BRCA2 breast cancer susceptibility protein*. Genes Dev, 2003. **17**(24): p. 3017-22.
124. Tutt, A., et al., *Absence of Brca2 causes genome instability by chromosome breakage and loss associated with centrosome amplification*. Curr Biol, 1999. **9**(19): p. 1107-10.
125. Yoshida, K. and Y. Miki, *Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage*. Cancer Sci, 2004. **95**(11): p. 866-71.
126. Mark, W.Y., et al., *Characterization of segments from the central region of BRCA1: an intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions?* J Mol Biol, 2005. **345**(2): p. 275-87.
127. Baer, R. and W.H. Lee, *Functional domains of the BRCA1 and BRCA2 proteins*. J Mammary Gland Biol Neoplasia, 1998. **3**(4): p. 403-12.
128. Westermarck, U.K., et al., *BARD1 participates with BRCA1 in homology-directed repair of chromosome breaks*. Mol Cell Biol, 2003. **23**(21): p. 7926-36.
129. Baer, R. and T. Ludwig, *The BRCA1/BARD1 heterodimer, a tumor suppressor complex with ubiquitin E3 ligase activity*. Curr Opin Genet Dev, 2002. **12**(1): p. 86-91.

130. Lorick, K.L., et al., *RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination*. Proc Natl Acad Sci U S A, 1999. **96**(20): p. 11364-9.
131. Jensen, D.E., et al., *BAP1: a novel ubiquitin hydrolase which binds to the BRCA1 RING finger and enhances BRCA1-mediated cell growth suppression*. Oncogene, 1998. **16**(9): p. 1097-112.
132. Nishikawa, H., et al., *BRCA1-associated protein 1 interferes with BRCA1/BARD1 RING heterodimer activity*. Cancer Res, 2009. **69**(1): p. 111-9.
133. Feng, Z., et al., *DNA damage induces p53-dependent BRCA1 nuclear export*. J Biol Chem, 2004. **279**(27): p. 28574-84.
134. Callebaut, I. and J.P. Mornon, *From BRCA1 to RPA1: a widespread BRCT module closely associated with DNA repair*. FEBS Lett, 1997. **400**(1): p. 25-30.
135. Koonin, E.V., S.F. Altschul, and P. Bork, *BRCA1 protein products ... Functional motifs*. Nat Genet, 1996. **13**(3): p. 266-8.
136. Manke, I.A., et al., *BRCT repeats as phosphopeptide-binding modules involved in protein targeting*. Science, 2003. **302**(5645): p. 636-9.
137. Motoyama, N. and K. Naka, *DNA damage tumor suppressor genes and genomic instability*. Curr Opin Genet Dev, 2004. **14**(1): p. 11-6.
138. Narod, S.A. and W.D. Foulkes, *BRCA1 and BRCA2: 1994 and beyond*. Nat Rev Cancer, 2004. **4**(9): p. 665-76.
139. Bochar, D.A., et al., *BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer*. Cell, 2000. **102**(2): p. 257-65.
140. Cantor, S.B., et al., *BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function*. Cell, 2001. **105**(1): p. 149-60.
141. Pao, G.M., et al., *CBP/p300 interact with and function as transcriptional coactivators of BRCA1*. Proc Natl Acad Sci U S A, 2000. **97**(3): p. 1020-5.
142. Yarden, R.I. and L.C. Brody, *BRCA1 interacts with components of the histone deacetylase complex*. Proc Natl Acad Sci U S A, 1999. **96**(9): p. 4983-8.
143. Magnard, C., et al., *BRCA1 interacts with acetyl-CoA carboxylase through its tandem of BRCT domains*. Oncogene, 2002. **21**(44): p. 6729-39.
144. Scully, R., et al., *Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage*. Cell, 1997. **90**(3): p. 425-35.
145. Thomas, J.E., et al., *Induction of phosphorylation on BRCA1 during the cell cycle and after DNA damage*. Cell Growth Differ, 1997. **8**(7): p. 801-9.
146. Chen, J., et al., *Stable interaction between the products of the BRCA1 and BRCA2 tumor suppressor genes in mitotic and meiotic cells*. Mol Cell, 1998. **2**(3): p. 317-28.
147. Venkitaraman, A.R., *Functions of BRCA1 and BRCA2 in the biological response to DNA damage*. J Cell Sci, 2001. **114**(Pt 20): p. 3591-8.
148. Moynahan, M.E., et al., *Brca1 controls homology-directed DNA repair*. Mol Cell, 1999. **4**(4): p. 511-8.
149. Shen, S.X., et al., *A targeted disruption of the murine Brca1 gene causes gamma-irradiation hypersensitivity and genetic instability*. Oncogene, 1998. **17**(24): p. 3115-24.
150. Moynahan, M.E., T.Y. Cui, and M. Jasin, *Homology-directed dna repair, mitomycin-c resistance, and chromosome stability is restored with correction of a Brca1 mutation*. Cancer Res, 2001. **61**(12): p. 4842-50.
151. Le Page, F., et al., *BRCA1 and BRCA2 are necessary for the transcription-coupled repair of the oxidative 8-oxoguanine lesion in human cells*. Cancer Res, 2000. **60**(19): p. 5548-52.
152. Hartman, A.R. and J.M. Ford, *BRCA1 induces DNA damage recognition factors and enhances nucleotide excision repair*. Nat Genet, 2002. **32**(1): p. 180-4.

153. Kim, H., J. Chen, and X. Yu, *Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response*. *Science*, 2007. **316**(5828): p. 1202-5.
154. Sobhian, B., et al., *RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites*. *Science*, 2007. **316**(5828): p. 1198-202.
155. Wang, B., et al., *Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response*. *Science*, 2007. **316**(5828): p. 1194-8.
156. Chen, L., et al., *Cell cycle-dependent complex formation of BRCA1.CtIP.MRN is important for DNA double-strand break repair*. *J Biol Chem*, 2008. **283**(12): p. 7713-20.
157. Sartori, A.A., et al., *Human CtIP promotes DNA end resection*. *Nature*, 2007. **450**(7169): p. 509-14.
158. Starita, L.M. and J.D. Parvin, *The multiple nuclear functions of BRCA1: transcription, ubiquitination and DNA repair*. *Curr Opin Cell Biol*, 2003. **15**(3): p. 345-50.
159. Cable, P.L., et al., *Novel consensus DNA-binding sequence for BRCA1 protein complexes*. *Mol Carcinog*, 2003. **38**(2): p. 85-96.
160. Lane, T.F., *BRCA1 and transcription*. *Cancer Biol Ther*, 2004. **3**(6): p. 528-33.
161. Tan-Wong, S.M., et al., *Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene*. *Proc Natl Acad Sci U S A*, 2008. **105**(13): p. 5160-5.
162. Aiyar, S.E., et al., *Concerted transcriptional regulation by BRCA1 and COBRA1 in breast cancer cells*. *Int J Biol Sci*, 2007. **3**(7): p. 486-92.
163. Sun, J., et al., *Deregulation of cofactor of BRCA1 expression in breast cancer cells*. *J Cell Biochem*, 2008. **103**(6): p. 1798-807.
164. Xu, B., S. Kim, and M.B. Kastan, *Involvement of Brca1 in S-phase and G(2)-phase checkpoints after ionizing irradiation*. *Mol Cell Biol*, 2001. **21**(10): p. 3445-50.
165. Harkin, D.P., et al., *Induction of GADD45 and JNK/SAPK-dependent apoptosis following inducible expression of BRCA1*. *Cell*, 1999. **97**(5): p. 575-86.
166. Aprelikova, O., et al., *BRCA1 is a selective co-activator of 14-3-3 sigma gene transcription in mouse embryonic stem cells*. *J Biol Chem*, 2001. **276**(28): p. 25647-50.
167. Yarden, R.I., et al., *BRCA1 regulates the G2/M checkpoint by activating Chk1 kinase upon DNA damage*. *Nat Genet*, 2002. **30**(3): p. 285-9.
168. Lotti, L.V., et al., *Subcellular localization of the BRCA1 gene product in mitotic cells*. *Genes Chromosomes Cancer*, 2002. **35**(3): p. 193-203.
169. Lou, Z., K. Minter-Dykhouse, and J. Chen, *BRCA1 participates in DNA decatenation*. *Nat Struct Mol Biol*, 2005. **12**(7): p. 589-93.
170. Yun, M.H. and K. Hiom, *CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle*. *Nature*, 2009.
171. Joukov, V., et al., *The BRCA1/BARD1 heterodimer modulates ran-dependent mitotic spindle assembly*. *Cell*, 2006. **127**(3): p. 539-52.
172. Wang, Y., J. Yu, and Q. Zhan, *BRCA1 regulates caveolin-1 expression and inhibits cell invasiveness*. *Biochem Biophys Res Commun*, 2008. **370**(2): p. 201-6.
173. Bieche, I. and R. Lidereau, *Increased level of exon 12 alternatively spliced BRCA2 transcripts in tumor breast tissue compared with normal tissue*. *Cancer Res*, 1999. **59**(11): p. 2546-50.
174. Speevak, M.D., et al., *Alternatively spliced, truncated human BRCA2 isoforms contain a novel coding exon*. *Eur J Hum Genet*, 2003. **11**(12): p. 951-4.
175. Yang, R., et al., *Polymorphisms in BRCA2 resulting in aberrant codon-usage and their analysis on familial breast cancer risk*. *Breast Cancer Res Treat*, 2009.

176. Fortin, J., *Analyses génomique et transcriptionnelle des gènes de susceptibilité aux cancers de sein et de l'ovaire BRCA1 et BRCA2 chez les Canadiennes Françaises*, in *Département de l'anatomie physiologie*. 2005, Université Laval: Québec. p. 146.
177. Davies, A.A., et al., *Role of BRCA2 in control of the RAD51 recombination and DNA repair protein*. *Mol Cell*, 2001. **7**(2): p. 273-82.
178. Moynahan, M.E., A.J. Pierce, and M. Jasin, *BRCA2 is required for homology-directed repair of chromosomal breaks*. *Mol Cell*, 2001. **7**(2): p. 263-72.
179. Yu, V.P., et al., *Gross chromosomal rearrangements and genetic exchange between nonhomologous chromosomes following BRCA2 inactivation*. *Genes Dev*, 2000. **14**(11): p. 1400-6.
180. Sonoda, E., et al., *Rad51-deficient vertebrate cells accumulate chromosomal breaks prior to cell death*. *EMBO J*, 1998. **17**(2): p. 598-608.
181. Joenje, H. and K.J. Patel, *The emerging genetic and molecular basis of Fanconi anaemia*. *Nat Rev Genet*, 2001. **2**(6): p. 446-57.
182. Witt, E. and A. Ashworth, *Biomedicine. D-Day for BRCA2*. *Science*, 2002. **297**(5581): p. 534.
183. Howlett, N.G., et al., *Biallelic inactivation of BRCA2 in Fanconi anemia*. *Science*, 2002. **297**(5581): p. 606-9.
184. Santarosa, M. and A. Ashworth, *Haploinsufficiency for tumour suppressor genes: when you don't need to go all the way*. *Biochim Biophys Acta*, 2004. **1654**(2): p. 105-22.
185. Connor, F., et al., *Tumorigenesis and a DNA repair defect in mice with a truncating Brca2 mutation*. *Nat Genet*, 1997. **17**(4): p. 423-30.
186. Fuks, F., J. Milner, and T. Kouzarides, *BRCA2 associates with acetyltransferase activity when bound to P/CAF*. *Oncogene*, 1998. **17**(19): p. 2531-4.
187. Milner, J., et al., *Transcriptional activation functions in BRCA2*. *Nature*, 1997. **386**(6627): p. 772-3.
188. Nordling, M., et al., *A large deletion disrupts the exon 3 transcription activation domain of the BRCA2 gene in a breast/ovarian cancer family*. *Cancer Res*, 1998. **58**(7): p. 1372-5.
189. Haber, D.A., *The BRCA2-EMSY connection: implications for breast and ovarian tumorigenesis*. *Cell*, 2003. **115**(5): p. 507-8.
190. Marmorstein, L.Y., T. Ouchi, and S.A. Aaronson, *The BRCA2 gene product functionally interacts with p53 and RAD51*. *Proc Natl Acad Sci U S A*, 1998. **95**(23): p. 13869-74.
191. Marmorstein, L.Y., et al., *A human BRCA2 complex containing a structural DNA binding component influences cell cycle progression*. *Cell*, 2001. **104**(2): p. 247-57.
192. Futamura, M., et al., *Potential role of BRCA2 in a mitotic checkpoint after phosphorylation by hBUBR1*. *Cancer Res*, 2000. **60**(6): p. 1531-5.
193. Li, J., et al., *DSS1 is required for the stability of BRCA2*. *Oncogene*, 2006. **25**(8): p. 1186-94.
194. *Breast Cancer Mutation Data Base*. 2007, National Human Genome Research Institute.
195. Shi, M.M., *Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies*. *Clin Chem*, 2001. **47**(2): p. 164-72.
196. Mooney, S., *Bioinformatics approaches and resources for single nucleotide polymorphism functional analysis*. *Brief Bioinform*, 2005. **6**(1): p. 44-56.
197. Spearman, A.D., et al., *Clinically applicable models to characterize BRCA1 and BRCA2 variants of uncertain significance*. *J Clin Oncol*, 2008. **26**(33): p. 5393-400.

198. Schwab, S.G., et al., *Association of tumor necrosis factor alpha gene -G308A polymorphism with schizophrenia*. Schizophr Res, 2003. **65**(1): p. 19-25.
199. Bonnet, C., et al., *Screening BRCA1 and BRCA2 unclassified variants for splicing mutations using reverse transcription PCR on patient RNA and an ex vivo assay based on a splicing reporter minigene*. J Med Genet, 2008. **45**(7): p. 438-46.
200. *The International HapMap Project*. Nature, 2003. **426**(6968): p. 789-96.
201. *A haplotype map of the human genome*. Nature, 2005. **437**(7063): p. 1299-320.
202. Rajasekaran, R., et al., *Identification and in silico analysis of functional SNPs of the BRCA1 gene*. Genomics, 2007. **90**(4): p. 447-52.
203. Tommasi, S., et al., *Molecular and in silico analysis of BRCA1 and BRCA2 variants*. Mutat Res, 2008. **644**(1-2): p. 64-70.
204. Tonin, P.N., et al., *Founder BRCA1 and BRCA2 mutations in French Canadian breast and ovarian cancer families*. Am J Hum Genet, 1998. **63**(5): p. 1341-51.
205. Oros, K.K., et al., *Significant proportion of breast and/or ovarian cancer families of French Canadian descent harbor 1 of 5 BRCA1 and BRCA2 mutations*. Int J Cancer, 2004. **112**(3): p. 411-9.
206. Ferla, R., et al., *Founder mutations in BRCA1 and BRCA2 genes*. Ann Oncol, 2007. **18 Suppl 6**: p. vi93-8.
207. Laberge, A.M., *[Prevalence and distribution of genetic diseases in Quebec: impact of the past on the present]*. Med Sci (Paris), 2007. **23**(11): p. 997-1001.
208. Oros, K.K., et al., *Application of BRCA1 and BRCA2 mutation carrier prediction models in breast and/or ovarian cancer families of French Canadian descent*. Clin Genet, 2006. **70**(4): p. 320-9.
209. Simard, J., et al., *Evaluation of BRCA1 and BRCA2 mutation prevalence, risk prediction models and a multistep testing approach in French-Canadian families with high risk of breast and ovarian cancer*. J Med Genet, 2007. **44**(2): p. 107-21.
210. De Braekeleer, M. and T.N. Dao, *Hereditary disorders in the French Canadian population of Quebec. I. In search of founders*. Hum Biol, 1994. **66**(2): p. 205-23.
211. Tonin, P.N., et al., *Founder BRCA1 and BRCA2 mutations in early-onset French Canadian breast cancer cases unselected for family history*. Int J Cancer, 2001. **95**(3): p. 189-93.
212. Tonin, P.N., et al., *Founder BRCA1 and BRCA2 mutations in French Canadian ovarian cancer cases unselected for family history*. Clin Genet, 1999. **55**(5): p. 318-24.
213. Oros, K.K., et al., *Haplotype analysis suggest common founders in carriers of the recurrent BRCA2 mutation, 3398delAAAAG, in French Canadian hereditary breast and/or ovarian cancer families*. BMC Med Genet, 2006. **7**: p. 23.
214. Jenkins, J.R., K. Rudge, and G.A. Currie, *Cellular immortalization by a cDNA clone encoding the transformation-associated phosphoprotein p53*. Nature, 1984. **312**(5995): p. 651-4.
215. Hastie, N.D., *The genetics of Wilms' tumor--a case of disrupted development*. Annu Rev Genet, 1994. **28**: p. 523-58.
216. King, T.A., et al., *Heterogenic loss of the wild-type BRCA allele in human breast tumorigenesis*. Ann Surg Oncol, 2007. **14**(9): p. 2510-8.
217. Hamel, N., et al., *Mixed ovarian germ cell tumor in a BRCA2 mutation carrier*. Int J Gynecol Pathol, 2007. **26**(2): p. 160-4.
218. Warren, M., et al., *Phenotypic effects of heterozygosity for a BRCA2 mutation*. Hum Mol Genet, 2003. **12**(20): p. 2645-56.
219. Provencher, D.M., et al., *Characterization of four novel epithelial ovarian cancer cell lines*. In Vitro Cell Dev Biol Anim, 2000. **36**(6): p. 357-61.

220. Buchholz, T.A., et al., *Evidence of haplotype insufficiency in human cells containing a germline mutation in BRCA1 or BRCA2*. Int J Cancer, 2002. **97**(5): p. 557-61.
221. Barwell, J., et al., *Lymphocyte radiosensitivity in BRCA1 and BRCA2 mutation carriers and implications for breast cancer susceptibility*. Int J Cancer, 2007. **121**(7): p. 1631-6.
222. Foray, N., et al., *Gamma-rays-induced death of human cells carrying mutations of BRCA1 or BRCA2*. Oncogene, 1999. **18**(51): p. 7334-42.
223. Rothfuss, A., et al., *Induced micronucleus frequencies in peripheral lymphocytes as a screening test for carriers of a BRCA1 mutation in breast cancer families*. Cancer Res, 2000. **60**(2): p. 390-4.
224. Baldeyron, C., et al., *A single mutated BRCA1 allele leads to impaired fidelity of double strand break end-joining*. Oncogene, 2002. **21**(9): p. 1401-10.
225. Trenz, K., et al., *Enhanced sensitivity of peripheral blood lymphocytes from women carrying a BRCA1 mutation towards the mutagenic effects of various cytostatics*. Mutat Res, 2003. **544**(2-3): p. 279-88.
226. Ganesan, S., et al., *BRCA1 supports XIST RNA concentration on the inactive X chromosome*. Cell, 2002. **111**(3): p. 393-405.
227. Buller, R.E., et al., *Association between nonrandom X-chromosome inactivation and BRCA1 mutation in germline DNA of patients with ovarian cancer*. J Natl Cancer Inst, 1999. **91**(4): p. 339-46.
228. Kristiansen, M., et al., *High frequency of skewed X inactivation in young breast cancer patients*. J Med Genet, 2002. **39**(1): p. 30-3.
229. Helbling-Leclere, A., G.M. Lenoir, and J. Feunteun, *Heterozygote BRCA1 status and skewed chromosome X inactivation*. Fam Cancer, 2007. **6**(1): p. 153-7.
230. Lose, F., et al., *Skewed X chromosome inactivation and breast and ovarian cancer status: evidence for X-linked modifiers of BRCA1*. J Natl Cancer Inst, 2008. **100**(21): p. 1519-29.
231. Shlien, A., et al., *Excessive genomic DNA copy number variation in the Li-Fraumeni cancer predisposition syndrome*. Proc Natl Acad Sci U S A, 2008. **105**(32): p. 11264-9.
232. Watts, J.A., et al., *Gene expression phenotype in heterozygous carriers of ataxia telangiectasia*. Am J Hum Genet, 2002. **71**(4): p. 791-800.
233. Grant, S.G., et al., *Elevated Levels of Somatic Mutation in a Manifesting BRCA1 Mutation Carrier*. Pathol Oncol Res, 2007. **13**(4): p. 276-83.
234. Chen, X., et al., *Allelic imbalance in BRCA1 and BRCA2 gene expression is associated with an increased breast cancer risk*. Hum Mol Genet, 2008. **17**(9): p. 1336-48.
235. Wong, A.S. and N. Auersperg, *Ovarian surface epithelium: family history and early events in ovarian cancer*. Reprod Biol Endocrinol, 2003. **1**: p. 70.
236. Dyck, H.G., et al., *Autonomy of the epithelial phenotype in human ovarian surface epithelium: changes with neoplastic progression and with a family history of ovarian cancer*. Int J Cancer, 1996. **69**(6): p. 429-36.
237. Roy, P.J., et al., *Chromosomal clustering of muscle-expressed genes in Caenorhabditis elegans*. Nature, 2002. **418**(6901): p. 975-9.
238. Lercher, M.J., A.O. Urrutia, and L.D. Hurst, *Clustering of housekeeping genes provides a unified model of gene order in the human genome*. Nat Genet, 2002. **31**(2): p. 180-3.
239. Spitz, F., F. Gonzalez, and D. Duboule, *A global control region defines a chromosomal regulatory landscape containing the HoxD cluster*. Cell, 2003. **113**(3): p. 405-17.

240. Monge, I., T. Kondo, and D. Duboule, *An enhancer-titration effect induces digit-specific regulatory alleles of the HoxD cluster*. Dev Biol, 2003. **256**(2): p. 212-20.
241. Piek, J.M., et al., *Expression of differentiation and proliferation related proteins in epithelium of prophylactically removed ovaries from women with a hereditary female adnexal cancer predisposition*. Histopathology, 2003. **43**(1): p. 26-32.
242. Barakat, R.R., et al., *Absence of premalignant histologic, molecular, or cell biologic alterations in prophylactic oophorectomy specimens from BRCA1 heterozygotes*. Cancer, 2000. **89**(2): p. 383-90.
243. Lounis, H., et al., *Primary cultures of normal and tumoral human ovarian epithelium: a powerful tool for basic molecular studies*. Exp Cell Res, 1994. **215**(2): p. 303-9.
244. Li, N.F., et al., *A modified medium that significantly improves the growth of human normal ovarian surface epithelial (OSE) cells in vitro*. Lab Invest, 2004. **84**(7): p. 923-31.
245. Kruk, P.A., S.L. Maines-Bandiera, and N. Auersperg, *A simplified method to culture human ovarian surface epithelium*. Lab Invest, 1990. **63**(1): p. 132-6.
246. Ouellet, V., et al., *Discrimination between serous low malignant potential and invasive epithelial ovarian tumors using molecular profiling*. Oncogene, 2005. **24**(29): p. 4672-87.
247. Dudoit, S., R.C. Gentleman, and J. Quackenbush, *Open source software for the analysis of microarray data*. Biotechniques, 2003. **Suppl**: p. 45-51.
248. Irizarry, R.A., et al., *Summaries of Affymetrix GeneChip probe level data*. Nucleic Acids Res, 2003. **31**(4): p. e15.
249. Bolstad, B.M., et al., *A comparison of normalization methods for high density oligonucleotide array data based on variance and bias*. Bioinformatics, 2003. **19**(2): p. 185-93.
250. Smyth, G.K., *Linear models and empirical bayes methods for assessing differential expression in microarray experiments*. Stat Appl Genet Mol Biol, 2004. **3**: p. Article3.
251. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing*. J. R. Statist. Soc., 1995. **B.57**: p. 289-300.
252. Kreiner, T., *Rapid genetic sequence analysis using a DNA probe array system*. American laboratory, 1996. **28**(5): p. 39-43.
253. Affymetrix, *Data Sheet - GeneChip® Made-to-Order Array Program*. 2002.
254. Gentleman, R.C., et al., *Bioconductor: open software development for computational biology and bioinformatics*. Genome Biol, 2004. **5**(10): p. R80.
255. Irizarry, R.A., et al., *Exploration, normalization, and summaries of high density oligonucleotide array probe level data*. Biostatistics, 2003. **4**(2): p. 249-64.
256. Wu, Z., et al., *A Model Based Background Adjustment for Oligonucleotide Expression Arrays*. Journal of the American Statistical Association, 2004. **99**(468): p. 909-917.
257. Affymetrix *GeneChip Expression Analysis - Data Analysis Fundamentals. Volume*,
258. Alberts, R., et al., *A verification protocol for the probe sequences of Affymetrix genome arrays reveals high probe accuracy for studies in mouse, human and rat*. BMC Bioinformatics, 2007. **8**: p. 132.
259. Mosteller, F. and J.W. Tukey, *Data analysis and regression : a second course in statistics*. 1977, Reading, Mass. ; Don Mills, Ont.: Addison-Wesley Pub. Co. xvii, 588 p.

260. Hartigan, J.A., *Clustering algorithms*. 1975, New York ; Toronto: Wiley. xiii, 351 p.
261. Lebart, L., *Quelques méthodes de classification*, in *Département des Sciences Economiques et Sociales*. 2003-2004, École Nationale Supérieure des Télécommunications - ENST: Paris. p. 23.
262. Lincoln, S., et al., *Primer3*. http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi.
263. Pfaffl, M.W., G.W. Horgan, and L. Dempfle, *Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR*. *Nucleic Acids Res*, 2002. **30**(9): p. e36.
264. Wilcoxon, F., *Individual Comparisons by Ranking Methods*. *Biometrics Bulletin* 1945. **1**: p. 80-83.
265. Smyth, G., et al. (2004) *The limma package*. **Volume**, 123
266. Spitz, F., [*Genes without borders: global control of transcription activity at the level of several genes*]. *Med Sci (Paris)*, 2004. **20**(1): p. 9-10.
267. He, Q.Y., et al., *Proteomic analysis of a preneoplastic phenotype in ovarian surface epithelial cells derived from prophylactic oophorectomies*. *Gynecol Oncol*, 2005. **98**(1): p. 68-76.
268. Tsao, S.W., et al., *Nonrandom chromosomal imbalances in human ovarian surface epithelial cells immortalized by HPV16-E6E7 viral oncogenes*. *Cancer Genet Cytogenet*, 2001. **130**(2): p. 141-9.
269. Zorn, K.K., et al., *Choice of normal ovarian control influences determination of differentially expressed genes in ovarian cancer expression profiling studies*. *Clin Cancer Res*, 2003. **9**(13): p. 4811-8.
270. Dudoit, S., et al., *Statistical methods for identifying genes with differential expression in replicated cDNA microarray experiments*. *Statistica Sinica*, 2002. **12**(1): p. 111-139.
271. Astrand, M., *Normalizing oligonucleotide arrays*. 2001.
272. Schadt, E.E., et al., *Feature extraction and normalization algorithms for high-density oligonucleotide gene expression array data*. *J Cell Biochem Suppl*, 2001. **Suppl 37**: p. 120-5.
273. Storey, J., *A direct approach to false discovery rates*. *Royal Statistical Society*, 2002. **64**(3): p. 479-498.
274. Hochberg, Y. and A.C. Tamhane, *Multiple comparison procedures*. *Wiley series in probability and mathematical statistics. Applied probability and statistics*. 1987, New York; Toronto: Wiley. xxii, 450 p.
275. De Preter, K., et al., *Positional gene enrichment analysis of gene sets for high-resolution identification of overrepresented chromosomal regions*. *Nucleic Acids Res*, 2008. **36**(7): p. e43.
276. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles*. *Proc Natl Acad Sci U S A*, 2005. **102**(43): p. 15545-50.
277. Efron, B., Tibshirani, R., *On testing the significance of sets of genes*. *Annals of applied statistics*, 2007. **1**(1): p. 107-129.
278. Myers, C.L., et al., *Accurate detection of aneuploidies in array CGH and gene expression microarray data*. *Bioinformatics*, 2004. **20**(18): p. 3533-43.
279. Yi, Y., et al., *Coupled analysis of gene expression and chromosomal location*. *Genomics*, 2005. **85**(3): p. 401-12.
280. *Finishing the euchromatic sequence of the human genome*. *Nature*, 2004. **431**(7011): p. 931-45.

281. Goeman, J.J. and P. Buhlmann, *Analyzing gene expression data in terms of gene sets: methodological issues*. Bioinformatics, 2007. **23**(8): p. 980-7.
282. Abaji, C., et al., *Mutations in Non-Tumor Ovarian Surface Epithelium Cells. Evidence of a Dosage Effect of 8765delAG and G6085T BRCA2 Mutations*. Genome Research 2008.
283. Sellar, G.C. and A.S. Whitehead, *Localization of four human serum amyloid A (SAA) protein superfamily genes to chromosome 11p: characterization of a fifth SAA-related gene sequence*. Genomics, 1993. **16**(3): p. 774-6.
284. Urieli-Shoval, S., et al., *Widespread expression of serum amyloid A in histologically normal human tissues. Predominant localization to the epithelium*. J Histochem Cytochem, 1998. **46**(12): p. 1377-84.
285. Gao, W.M., et al., *Distinctive serum protein profiles involving abundant proteins in lung cancer patients based upon antibody microarray analysis*. BMC Cancer, 2005. **5**: p. 110.
286. Moshkovskii, S.A., et al., *Ovarian cancer marker of 11.7 kDa detected by proteomics is a serum amyloid A1*. Proteomics, 2005. **5**(14): p. 3790-7.
287. Le, L., et al., *Identification of serum amyloid A as a biomarker to distinguish prostate cancer patients with bone lesions*. Clin Chem, 2005. **51**(4): p. 695-707.
288. Hao, C.Y., et al., *Alteration of gene expression in macroscopically normal colonic mucosa from individuals with a family history of sporadic colon cancer*. Clin Cancer Res, 2005. **11**(4): p. 1400-7.
289. Mazzanti, C.M., et al., *Early genetic mechanisms underlying the inhibitory effects of endostatin and fumagillin on human endothelial cells*. Genome Res, 2004. **14**(8): p. 1585-93.
290. Mok, S.C., et al., *Molecular cloning of differentially expressed genes in human epithelial ovarian cancer*. Gynecol Oncol, 1994. **52**(2): p. 247-52.
291. Schwarze, S.R., et al., *Novel pathways associated with bypassing cellular senescence in human prostate epithelial cells*. J Biol Chem, 2002. **277**(17): p. 14877-83.
292. Cody, N.A., et al., *Transfer of chromosome 3 fragments suppresses tumorigenicity of an ovarian cancer cell line monoallelic for chromosome 3p*. Oncogene, 2007. **26**(4): p. 618-32.
293. Stein, T., et al., *Annexin A8 is up-regulated during mouse mammary gland involution and predicts poor survival in breast cancer*. Clin Cancer Res, 2005. **11**(19 Pt 1): p. 6872-9.
294. Goebeler, V., et al., *Annexin A8 displays unique phospholipid and F-actin binding properties*. FEBS Lett, 2006. **580**(10): p. 2430-4.
295. DeRossi, C., et al., *Ablation of mouse phosphomannose isomerase (Mpi) causes mannose 6-phosphate accumulation, toxicity, and embryonic lethality*. J Biol Chem, 2006. **281**(9): p. 5916-27.
296. Pedersen, P.L., *Warburg, me and Hexokinase 2: Multiple discoveries of key molecular events underlying one of cancers' most common phenotypes, the "Warburg Effect", i.e., elevated glycolysis in the presence of oxygen*. J Bioenerg Biomembr, 2007. **39**(3): p. 211-22.
297. Patel, K.J., et al., *Involvement of Brca2 in DNA repair*. Mol Cell, 1998. **1**(3): p. 347-57.
298. Egawa, C., et al., *Quantitative analysis of BRCA1 and BRCA2 mRNA expression in sporadic breast carcinomas and its relationship with clinicopathological characteristics*. Jpn J Cancer Res, 2001. **92**(6): p. 624-30.
299. Roy, P., J. Stuart, and J. Lund, *Chromosomal clustering of muscle-expressed genes in Caenorhabditis elegans*. et al, Nature, 2002. **418**(6901): p. 975-9.

300. Lercher, M. and L. Hurst, *Co-expressed yeast genes cluster over a long range but are not regularly spaced*. J Mol Biol, 2006. **359**(3): p. 825-31.
301. Zuniga, A., *Globalisation reaches gene regulation: the case for vertebrate limb development*. Curr Opin Genet Dev, 2005. **15**(4): p. 403-9.
302. Hurst, L.D., C. Pal, and M.J. Lercher, *The evolutionary dynamics of eukaryotic gene order*. Nat Rev Genet, 2004. **5**(4): p. 299-310.
303. Zweemer, R.P., et al., *Comparative genomic hybridization of microdissected familial ovarian carcinoma: two deleted regions on chromosome 15q not previously identified in sporadic ovarian carcinoma*. Lab Invest, 2001. **81**(10): p. 1363-70.
304. Tapper, J., et al., *Genetic changes in inherited and sporadic ovarian carcinomas by comparative genomic hybridization: extensive similarity except for a difference at chromosome 2q24-q32*. Cancer Res, 1998. **58**(13): p. 2715-9.
305. Watanabe, M., et al., *Allelic loss and microsatellite instability in prostate cancers in Japan*. Oncology, 1998. **55**(6): p. 569-74.
306. Hermsen, M.A., et al., *Genetic analysis of 53 lymph node-negative breast carcinomas by CGH and relation to clinical, pathological, morphometric, and DNA cytometric prognostic factors*. J Pathol, 1998. **186**(4): p. 356-62.
307. Seitz, S., et al., *Deletion mapping and linkage analysis provide strong indication for the involvement of the human chromosome region 8p12-p22 in breast carcinogenesis*. Br J Cancer, 1997. **76**(8): p. 983-91.
308. Sarrío, D., et al., *Expression of cadherins and catenins correlates with distinct histologic types of ovarian carcinomas*. Hum Pathol, 2006. **37**(8): p. 1042-9.
309. Suzuki, M., et al., *Methylation and gene silencing of the Ras-related GTPase gene in lung and breast cancers*. Ann Surg Oncol, 2007. **14**(4): p. 1397-404.
310. Thompson, F.H., et al., *Amplification of 19q13.1-q13.2 sequences in ovarian cancer. G-band, FISH, and molecular studies*. Cancer Genet Cytogenet, 1996. **87**(1): p. 55-62.