

FRÉDÉRIC GUÉNARD

**ÉVALUATION DE LA CONTRIBUTION DE GÈNES  
DE SUSCEPTIBILITÉ ET DE GÈNES CANDIDATS  
DU CANCER DU SEIN CHEZ DES FAMILLES  
CANADIENNES-FRANÇAISES À RISQUE ÉLEVÉ**

Thèse présentée  
à la Faculté des études supérieures de l'Université Laval  
dans le cadre du programme de doctorat en Physiologie-endocrinologie  
pour l'obtention du grade de Philosophiae doctor (Ph.D.)

DÉPARTEMENT DE MÉDECINE MOLÉCULAIRE  
FACULTÉ DE MÉDECINE  
UNIVERSITÉ LAVAL  
QUÉBEC

2010

## RÉSUMÉ

Le cancer du sein est une maladie comportant à la fois une composante environnementale et une composante génétique. Dans le cadre de mes études doctorales je me suis intéressé à la composante génétique du cancer du sein. Puisque la distribution de l'héritabilité chez les familles présentant plusieurs cas de cancer du sein suggère la prédominance de facteurs génétiques afin d'expliquer l'agrégation familiale de ce type de cancer, mes études ont porté sur des cas de cancer du sein provenant de familles à risque élevé non porteuses de mutations délétères dans les gènes *BRCA1* et *BRCA2*.

Deux volets principaux ont été abordés. Le premier volet, intitulé « évaluation de la contribution de gènes de susceptibilité au cancer du sein », m'a amené à évaluer la contribution des gènes de susceptibilité à forte pénétrance *PTEN* et *STK11*. Ces gènes sont respectivement responsables des syndromes de Cowden et de Peutz-Jeghers avec lesquels un risque accru de cancer du sein a été associé. Trois autres gènes de susceptibilité au cancer du sein, soit *CHEK2*, *BRIP1* et *PALB2*, ont également été analysés. Des mutations dans ces gènes confèrent un risque modéré de cancer du sein.

Lors du deuxième volet de mon projet, « identification de gènes de susceptibilité au cancer du sein », quatre gènes candidats du cancer du sein ont été analysés. Ces gènes ont été sélectionnés sur la base de leur interaction avec *BRCA1*. En effet, puisque les protéines encodées par ces gènes interagissent avec *BRCA1*, elles pourraient réguler ses diverses fonctions. Les gènes *AURKA*, *BAP1*, *BARD1* et *DHX9*, encodant respectivement une kinase, une ubiquitine hydrolase, une ubiquitine ligase et une hélicase à ARN, ont été sélectionnés pour leur rôle possible dans la modulation de l'activité de *BRCA1* dans la réparation couplée à la transcription. L'étude de ces gènes a permis de mettre en évidence une implication possible d'un variant de *BARD1* dans la susceptibilité au cancer du sein.

L'étude de ces gènes a permis d'acquérir une meilleure connaissance de la contribution de ceux-ci à la susceptibilité au cancer du sein chez les familles à risque élevé provenant de la population canadienne-française.

## ABSTRACT

Breast cancer is a disease with both environmental and genetic components. During my Ph.D., my principal interest was focused on the genetic component of breast cancer susceptibility. Studies on families with breast cancer revealed that familial aggregation of breast cancer is mainly explained by genetic factors. Non-*BRCA1/BRCA2* breast cancer individuals from French Canadian breast cancer families were thus analyzed in this project, which is divided in two components.

The first component was the evaluation of the contribution of established breast cancer susceptibility genes in our familial breast cancer cohort. We analyzed the *PTEN* and *STK11* genes, two high penetrance breast cancer susceptibility genes. Mutations in the *PTEN* and *STK11* genes are responsible for Cowden and Li-Faumeni syndromes, respectively, both associated with breast cancer susceptibility. We also analyzed three moderate penetrance breast cancer susceptibility genes, namely *CHEK2*, *BRIP1* and *PALB2*.

The second component of my project consisted in the analysis of four breast cancer candidate genes. These genes were selected on the basis of their interaction with *BRCA1*, one of the major breast cancer susceptibility gene. The proteins encoded by the *AURKA*, *BAP1*, *BARD1* and *DHX9* genes, which are known as kinase, ubiquitin hydrolase, ubiquitin ligase and RNA helicase, respectively, were selected for their possible effect on *BRCA1* function in the transcription coupled repair mechanism. This component led to the identification of a difference in allele frequency for one *BARD1* variant.

In summary, this project led to a better knowledge regarding the contribution of these genes in breast cancer susceptibility among families with high risk of breast cancer from the French Canadian population.

## AVANT-PROPOS

Cette thèse est présentée avec insertion d'articles. Quatre articles dont je suis l'auteur principal constitueront les chapitres II à V et formeront le corps de cette thèse. Ces articles ont été préparés dans le cadre du projet de recherche entrepris pour l'obtention du grade de Philosophiae Doctor. Le premier de ces articles est paru dans la revue *Familial Cancer* en décembre 2007. Le deuxième et le troisième article ont été publiés dans la revue *Journal of Human Genetics* en juillet 2008 et en mars 2009. Une version du manuscrit du quatrième article inclus dans ma thèse a été soumise dans la revue *Genetic Testing and Molecular Biomarkers* en février 2010. Ce manuscrit est présentement en révision par les évaluateurs de la revue. Il est à noter que l'ordre de présentation des publications incluses dans le présent ouvrage ne respecte pas l'ordre chronologique de publication. Afin de faciliter la présentation des deux volets de mon projet, les gènes de susceptibilité au cancer du sein et les gènes candidats du cancer du sein seront présentés séparément. Ainsi, les chapitres II à IV seront composés des articles portant sur les gènes de susceptibilité au cancer du sein alors que l'article concernant les gènes candidats du cancer du sein sera présenté au chapitre V.

Pour chaque article composant le corps de cette thèse, la Dre Durocher fut responsable de la mise au point de l'étude, de la conception de la cohorte et de la révision du manuscrit. La présence de INHERIT BRCA (INterdisciplinary HEalth Research International Team on BREast CAncer susceptibility) dans la liste des auteurs dénote la participation des autres membres de notre groupe de recherche international, dans le recrutement des patients et dans l'évaluation de la présence de mutations de *BRCA1* et *BRCA2* chez les femmes qui composent notre cohorte. Dans certains cas, la présence du Dr Jacques Simard parmi la liste des auteurs indique son implication à titre de Directeur du groupe INHERIT BRCA dans la mise sur pied de ce programme de recherche clinique intégré et dans le recrutement des familles à risque élevé de cancer du sein. Bien qu'il n'apparaisse pas parmi la liste des auteurs des articles composant le corps de cette thèse, je tiens également à souligner l'implication du Dr Damian Labuda qui nous a donné accès à l'ADN génomique d'individus sains nous servant à évaluer la présence des variants identifiés dans une cohorte de femmes non-atteintes.

Le deuxième chapitre, composé de l'article concernant le gène *PTEN*, présente les analyses nous ayant permis de conclure que des mutations de *PTEN* sont rares et qu'il est peu probable que ce gène contribue à une proportion significative des cas de cancer du sein familiaux de la population canadienne-française. Dans le cadre de ce projet, j'ai été responsable de la production et de l'analyse des données ainsi que de la rédaction du manuscrit. Geneviève Ouellette et Charles Joly Beauparlant m'ont appuyé au point de vue technique alors que le Dr Yvan Labrie a agi à titre d'assistant technique. Il a aussi été impliqué dans la rédaction du manuscrit. Les autres auteurs de cet article sont les chercheurs cliniciens qui ont été responsables du recrutement des familles à risque élevé de cancer du sein composant notre cohorte.

Le deuxième article inséré dans cette thèse présente les résultats obtenus lors de l'analyse du gène de susceptibilité *BRIP1* aussi appelé *FANCF* en raison de son implication dans le sous-type J de l'anémie de Fanconi. Dans le cadre de ce projet sous la supervision de la Dre Francine Durocher, j'ai effectué le séquençage de ce gène avec l'aide technique de Geneviève Ouellette et de Charles Joly Beauparlant. Puisque les analyses bioinformatiques menées en collaboration avec le Dr Yvan Labrie laissaient croire à une implication des variants identifiés sur la liaison de facteurs de transcription, j'ai procédé à des essais de gène rapporteur avec le support technique de Geneviève Ouellette, la responsable de la culture cellulaire dans notre équipe. Ceci nous a permis d'évaluer l'effet potentiel des variants situés dans la région promotrice du gène *BRIP1* sur son niveau d'expression.

Le plus récent des articles faisant partie du corps du présent ouvrage a porté sur trois gènes de susceptibilité au cancer du sein. Ce manuscrit, soumis dans la revue *Human Genomics*, présente les analyses nous ayant permis d'exclure une forte implication des gènes *CHEK2*, *STK11* et *PALB2* dans la susceptibilité au cancer du sein chez les familles canadiennes-françaises à risque élevé. Nos analyses suggèrent également une implication potentielle des variants introniques de *STK11* dans la susceptibilité au cancer du sein. Lors de cette étude, j'ai effectué le séquençage des gènes *STK11* et *PALB2* avec la collaboration respective de Christopher St-Laurent Pedneault (étudiant stagiaire sous ma supervision) et de Geneviève Ouellette. Cette dernière a particulièrement été responsable du séquençage de *CHEK2*. De grands réarrangements génomiques ayant été identifiés pour *STK11* et étant suspectés pour

*PALB2*, j'ai effectué avec l'aide de Christopher St-Laurent Pedneault des analyses nous permettant d'exclure la présence de tels réarrangements dans ces deux gènes. Toujours en étroite collaboration avec le Yvan Labrie, j'ai effectué l'analyse des résultats obtenus, j'ai utilisé divers outils bioinformatiques afin d'évaluer l'effet potentiel des variants identifiés et j'ai effectué la rédaction du manuscrit.

Le dernier article inséré dans cette thèse présente l'évaluation de gènes candidats du cancer du sein. Dans le cadre de ce projet, quatre gènes candidats encodant des protéines interagissant avec *BRCA1* ont été à l'étude. De ceux-ci, les gènes *BAP1* et *DHX9* ont été analysés par Geneviève Ouellette alors qu'avec la collaboration de Charles Joly Beauparlant, étudiant à la maîtrise dans le groupe du Dre Durocher, j'ai effectué le séquençage des gènes *AURKA* et *BARD1*. Après avoir effectué différentes analyses bioinformatiques afin d'évaluer l'effet potentiel des variants identifiés dans ces quatre gènes candidats, j'ai effectué la rédaction du manuscrit. Ce projet ayant été initié au début de mes études graduées, le support technique fourni tout au long de ce projet par le Dr Yvan Labrie ainsi que son implication constante dans l'analyse des résultats et dans la rédaction du manuscrit lui ont valu le titre de deuxième auteur de cet article.

La réalisation de mon projet de doctorat a également mené à quatre autres publications qui ne font pas l'objet de cette thèse. La première fut en tant que deuxième auteur d'un chapitre du livre intitulé « Molecular Genetics of Cancer », publié aux éditions Research Signpost et dont l'éditeur est le Dr Daniel Sinnett. Une deuxième publication dans laquelle j'ai été impliqué a été réalisée par un consortium international ayant pour but d'évaluer l'effet de gènes modificateurs sur le risque de cancer du sein chez les porteurs de mutations dans les gènes *BRCA1* et *BRCA2*. Ma contribution à ce projet a consisté au génotypage du gène *AURKA* dans notre cohorte de femmes porteuses de mutations de *BRCA1* et *BRCA2*. Cette publication est fournie en annexe de cette thèse. Sur invitation de l'éditeur, et avec la collaboration de la Dre Durocher, j'ai contribué à deux reprises à titre d'auteur principal à l'élaboration d'une base de données nommée « Atlas of Genetics and Cytogenetics in Oncology and Haematology » et disponible sur Internet. Chaque contribution à cette base de données a consisté en une revue de littérature portant sur un gène candidat étudié faisant partie intégrante de mon projet de doctorat. Ainsi, notre revue de littérature sur le gène

*DHX9* est parue en juillet 2009 alors que notre contribution concernant le gène *BAP1* a été publiée lors de la mise à jour de septembre 2009.

## REMERCIEMENTS

Je tiens à remercier la Dre Francine Durocher pour sa supervision et sa disponibilité, tout comme pour les discussions scientifiques et professionnelles que nous avons eues au cours de ces années. Elle a su me communiquer son intérêt pour la science et élargir mes horizons à plusieurs aspects de la recherche. Par son dévouement et sa grande implication sociale, elle fut pour moi un exemple à suivre.

Merci également au Dr Yvan Labrie, pour son expertise de laboratoire, pour sa supervision en cours de rédaction d'articles, pour son regard critique sur la science et son franc parler. Je voudrais exprimer ma gratitude aux collègues et amis que j'ai côtoyés au cours des dernières années, notamment Geneviève Ouellette pour son travail en culture cellulaire et son aide technique dans plusieurs projets et Sylvie Desjardins pour les discussions que nous avons eues sur nos projets respectifs, qui finalement, s'avèrent étroitement reliés. Je tiens à remercier Charles Joly Beuparlant pour l'aide technique qu'il m'a fournie avant de prendre son envol comme étudiant gradué dans notre équipe. Merci à Christopher St-Laurent Pedneault et aux autres étudiants que j'ai eu l'occasion de superviser. Ils ont tous, à divers niveaux, participé à l'avancement de mes recherches et à mon épanouissement personnel et professionnel. J'ai aussi une pensée spéciale pour Mélanie Houde, qui m'a accompagné au laboratoire pour plus d'une année et qui continue à m'accompagner en pensée.

Je tiens à exprimer ma reconnaissance au Dr Jacques Simard pour l'opportunité qu'il m'a donnée de découvrir le monde de la recherche en génétique et génomique des cancers, un domaine vaste et stimulant qui est à la fois fondamental et près des patientes. Je tiens aussi à exprimer ma reconnaissance aux membres du Laboratoire de génomique des cancers pour la gestion des échantillons, le support technique et informatique, ainsi qu'aux infirmières de recherche impliquées dans les projets de recherche du groupe INHERIT BRCA<sub>s</sub>. Merci aux autres membres du CREMOG pour leur assistance technique et leurs discussions constructives, tout spécialement aux personnes responsables de la plateforme de séquençage avec qui j'ai eu le plaisir de collaborer étroitement.



Au cours de ces années, j'ai eu le support moral et financier de la Fondation René Bussièrès. Savoir que les responsables de cette fondation (messieurs René Bussièrès et Denis Durocher) croyaient en moi et que j'avais leur appui a été pour moi une source de réconfort et de motivation incroyable. C'est pour cela que je tiens à leur exprimer toute ma reconnaissance.

Le support moral de mes proches a contribué à ma réussite. Dans les moments de réjouissance tout comme dans les moments de doute, ils ont su m'accompagner. Leur vision de la vie et leur entrain m'ont permis de relativiser les choses et de prendre la vie du bon côté.

*À Doris, Michel et Julie,  
qui m'ont accompagné et  
encouragé toutes ces années.  
À Josée, qui a laissé plusieurs  
projets de vie en plan pour  
mon accomplissement  
académique et professionnel.*

## TABLE DES MATIÈRES

<b>RÉSUMÉ</b> .....	<b>II</b>
<b>ABSTRACT</b> .....	<b>III</b>
<b>AVANT-PROPOS</b> .....	<b>IV</b>
<b>REMERCIEMENTS</b> .....	<b>VIII</b>
<b>TABLE DES MATIÈRES</b> .....	<b>XI</b>
<b>LISTE DES TABLEAUX</b> .....	<b>XIII</b>
<b>LISTE DES FIGURES</b> .....	<b>XV</b>
<b>LISTE DES ABRÉVIATIONS</b> .....	<b>XVII</b>
<b>CHAPITRE I - INTRODUCTION</b> .....	<b>1</b>
1. Le cancer.....	1
1.1. Statistiques sur le cancer.....	1
1.2. Développement du cancer et formation de tumeurs.....	1
2. Le cancer du sein, une maladie multifactorielle.....	3
2.1. Facteurs environnementaux.....	3
2.1.1. Facteurs démographiques et sanitaires.....	4
2.1.2. Facteurs hormonaux.....	5
2.1.3. Facteurs liés aux habitudes de vie et de nutrition.....	6
2.2. Facteurs génétiques impliqués dans le cancer du sein.....	7
3. Gènes de susceptibilité au cancer du sein.....	9
3.1. Gènes majeurs de susceptibilité au cancer du sein.....	9
3.1.1. <i>BRCA1</i> .....	10
3.1.2. <i>BRCA2</i> .....	14
3.2. Autres gènes de susceptibilité au cancer du sein.....	17
3.2.1. <i>TP53</i> .....	17
3.2.2. <i>PTEN</i> .....	19
3.2.3. <i>STK11</i> .....	24
3.2.4. <i>ATM</i> .....	28
3.2.5. <i>CHEK2</i> .....	30
3.2.6. Gènes de l'anémie de Fanconi.....	32
3.2.6.1. <i>BRIP1/FANCF</i> .....	37
3.2.6.2. <i>PALB2/FANCD1</i> .....	39
3.3. Gènes de susceptibilité au cancer du sein : participation à une voie commune.....	41
3.4. Classification des gènes de susceptibilité.....	44
3.5. Gènes de susceptibilité à faible pénétrance.....	47
3.5.1. L'effet polygénique.....	48
3.5.2. Modèle multiplicatif.....	50
4. Recherche de gènes de susceptibilité au cancer du sein.....	51
4.1. Utilisation de populations fondatrices.....	52
4.2. Utilisation de familles enrichies.....	53
4.3. Utilisation de blocs d'haplotypes.....	54
4.4. Utilisation d'outils bioinformatiques.....	54
4.4.1. Variants situés dans la région codante.....	55
4.4.2. Variants pouvant affecter l'épissage.....	55
4.4.3. Variants pouvant affecter l'expression.....	57

5. Partenaires d'interaction de BRCA1 .....	59
5.1. <i>BARD1</i> .....	59
5.2. <i>BAP1</i> .....	63
5.3. <i>AURKA</i> .....	65
5.4. <i>DHX9</i> .....	67
6. Problématique .....	70
<b>CHAPITRE II - Germline mutations in the breast cancer susceptibility gene <i>PTEN</i> are rare in high-risk non-<i>BRCA1/2</i> French Canadian breast cancer families .....</b>	<b>74</b>
1. Résumé .....	74
2. Manuscrit .....	75
<b>CHAPITRE III - Mutational analysis of the breast cancer susceptibility gene <i>BRIP1/BACH1/FANCI</i> in high-risk non-<i>BRCA1/BRCA2</i> Breast Cancer Families .....</b>	<b>96</b>
1. Résumé .....	96
2. Manuscrit .....	97
<b>CHAPITRE IV - Evaluation of the contribution of the three breast cancer susceptibility genes <i>CHEK2</i>, <i>STK11</i> and <i>PALB2</i> in non-<i>BRCA1/2</i> French Canadian families with high risk of breast cancer .....</b>	<b>130</b>
1. Résumé .....	130
2. Manuscrit .....	131
<b>CHAPITRE V - Genetic sequence variations of BRCA1-interacting genes <i>AURKA</i>, <i>BAP1</i>, <i>BARD1</i> and <i>DHX9</i> in French Canadian Families with high risk of breast cancer .....</b>	<b>172</b>
1. Résumé .....	172
2. Manuscrit .....	173
<b>DISCUSSION .....</b>	<b>205</b>
<b>CONCLUSION .....</b>	<b>212</b>
<b>BIBLIOGRAPHIE .....</b>	<b>216</b>
<b>ANNEXE - <i>AURKA</i> F31I Polymorphism and Breast Cancer Risk in <i>BRCA1</i> and <i>BRCA2</i> Mutation Carriers: A Consortium of Investigators of Modifiers of BRCA1/2 Study .....</b>	<b>244</b>

## LISTE DES TABLEAUX

### CHAPITRE I - INTRODUCTION

<b>Tableau 1 :</b> Facteurs environnementaux modifiant le risque de cancer du sein.....	4
<b>Tableau 2 :</b> Risque associé à des mutations dans les gènes de susceptibilité au cancer du sein à forte pénétrance et à pénétrance modérée .....	46
<b>Tableau 3 :</b> Pourcentage de réduction de la taille de l'échantillon obtenu par l'utilisation de cas avec une histoire familiale de cancer du sein.....	54

### CHAPITRE II

<b>Tableau 1 :</b> Observed intronic sequence variants and genotype frequencies in familial breast cancer cases.....	91
<b>Tableau 2 :</b> Splicing scores observed for the WT and putative splicing sites in the <i>PTEN</i> gene.....	92
<b>Tableau supplémentaire 1 :</b> Oligonucleotide primers used for amplification of the <i>PTEN</i> gene.....	95

### CHAPITRE III

<b>Tableau 1 :</b> Observed coding and intronic sequence variants and genotype frequencies in familial breast cancer cases and controls .....	122
<b>Tableau 2 :</b> Sequence variants detected in human <i>FANCI</i> and residues found in orthologues.....	123
<b>Tableau 3 :</b> Identification of Transcription Factor Binding Sites (TFBS) potentially affected by variants located in the <i>FANCI</i> promoter region.....	124
<b>Tableau 4 :</b> Identification of Exonic Splicing Enhancer (ESE) potentially affected by variants in the <i>FANCI</i> gene.....	125
<b>Tableau supplémentaire 1 :</b> Oligonucleotides used to amplify and sequence promoter, exons and flanking intronic regions of the <i>FANCI</i> gene.....	128
<b>Tableau supplémentaire 2 :</b> Oligonucleotides used for cloning and site-directed mutagenesis of the <i>FANCI</i> promoter constructs .....	129

### CHAPITRE IV

<b>Tableau 1 :</b> Observed sequence variations of <i>PALB2</i> , <i>STK11</i> and <i>CHEK2</i> genes and their genotype in breast cancer and unaffected control individuals .....	161
<b>Tableau 2 :</b> Allele frequencies of variations of <i>PALB2</i> , <i>STK11</i> and <i>PALB2</i> in breast cancer and unaffected control individuals .....	164
<b>Tableau 3 :</b> Non-synonymous sequence variants detected in <i>PALB2</i> and <i>STK11</i> genes and corresponding residues present in orthologues.....	166
<b>Tableau 4 :</b> Location of non-synonymous sequence variants detected in human <i>PALB2</i> and <i>STK11</i> genes and amino acid effect prediction on protein function.....	167

<b>Tableau supplémentaire 1 :</b> Oligonucleotides used to amplify and sequence the <i>STK11</i> and <i>PALB2</i> genes.....	169
<b>Tableau supplémentaire 2 :</b> Description of the regions re-sequenced for the <i>CHEK2</i> , <i>STK11</i> and <i>PALB2</i> genes .....	170
<b>CHAPITRE V</b>	
<b>Tableau 1 :</b> Sequence variations and genotype frequencies of <i>AURKA</i> , <i>BAP1</i> , <i>BARD1</i> and <i>DHX9</i> genes in cases and controls .....	195
<b>Tableau 2 :</b> Non-synonymous sequence variants detected in <i>AURKA</i> , <i>BAP1</i> , <i>BARD1</i> and <i>DHX9</i> genes and corresponding residues present in orthologues.....	198
<b>Tableau 3 :</b> Location of non-synonymous sequence variants detected in <i>AURKA</i> , <i>BAP1</i> , <i>BARD1</i> and <i>DHX9</i> genes and prediction on protein function.....	199
<b>Tableau 4 :</b> Estimated haplotype frequencies in cases and controls using PHASE.....	200
<b>Tableau supplémentaire 1 :</b> Oligonucleotides used to amplify and sequence <i>AURKA</i> , <i>BAP1</i> , <i>BARD1</i> and <i>DHX9</i> genes in familial breast cancer (cDNA) and controls (gDNA).....	204

## LISTE DES FIGURES

### CHAPITRE I - INTRODUCTION

<b>Figure 1</b> : Fonctions des protéines BRCA1 et BRCA2.....	17
<b>Figure 2</b> : Implication de PTEN dans le maintien de la stabilité chromosomique.....	22
<b>Figure 3</b> : Interrelations entre les protéines p53, PTEN et STK11 .....	28
<b>Figure 4</b> : FA/BRCA pathway .....	35
<b>Figure 5</b> : BRIP1 joue des rôles dépendants et indépendants de BRCA1 .....	38
<b>Figure 6</b> : Interrelations de BRCA1 et BRCA2 avec les protéines encodées par les autres gènes de susceptibilité au cancer du sein.....	42
<b>Figure 7</b> : Catégories de gènes de susceptibilité au cancer du sein et risque leur étant associé.....	47
<b>Figure 8</b> : Distribution du risque de cancer du sein chez la population en général et chez les cas de cancer du sein .....	49
<b>Figure 9</b> : Effet fondateur au Québec .....	52
<b>Figure 10</b> : Éléments régulateurs et protéines impliqués dans le mécanisme d'épissage .....	56
<b>Figure 11</b> : Mécanisme d'ubiquitination.....	60
<b>Figure 12</b> : BARD1 fait partie de complexes ayant des rôles différents dans le cycle cellulaire.....	61
<b>Figure 13</b> : AURKA régule l'activité de modulation des centrosomes de BRCA1 .....	66

### CHAPITRE II

<b>Figure 1</b> : Homology analysis of <i>PTEN</i> gene between relevant species.....	93
<b>Figure 2</b> : RPA and Western analyses of identified variants in carriers.....	94

### CHAPITRE III

<b>Figure 1</b> : Promoter activity of long (1068 bp) and short (633 bp) constructs carrying variations present in the promoter.....	126
<b>Figure 2</b> : Haplotype frequencies and combined frequencies of haplotypes reconstructed by the PHASE program and haplotype blocks and tSNPs identified .....	127

### CHAPITRE IV

<b>Figure 1</b> : Pedigrees of <i>CHEK2</i> c.1100delC carriers' families and <i>STK11</i> Phe354Leu carrier's family.....	168
<b>Figure supplémentaire 1</b> : Graphical overview of pairwise linkage disequilibrium (LD) measures of $ D' $ (lower left part) and $r^2$ (higher right part) for each SNPs pair identified inbreast cancer cases and unaffected individuals.....	171

### CHAPITRE V

<b>Figure 1</b> : Pedigree of AURKA:Met373Val carrier's family.....	201
---	-----

<b>Figure 2 :</b> Graphical overview of pairwise linkage disequilibrium (LD) measures of $D'$ and $r^2$ for each SNPs pair identified in breast cancer cases and healthy individuals.....	<b>202</b>
<b>Figure 3 :</b> Haplotype blocks and tSNPs identified for the <i>AURKA</i> , <i>BAP1</i> , <i>BARD1</i> and <i>DHX9</i> genes.....	<b>203</b>



## LISTE DES ABRÉVIATIONS

53BP1	p53 binding protein 1
ADNc	ADN complémentaire
AF	anémie de Fanconi
AML	leucémie myéloïde sévère (acute myeloid leukemia)
ARN pol II	ARN polymérase II
AT	ataxie télangiectasie
ATM	ataxia telangiectasia mutated
ATR	ataxia telangiectasia and Rad3 related
BASC	BRCA1-associated genome surveillance complex
BRCT	BRCA1 C-terminal
BRRS	syndrome de Bannayan-Riley-Ruvalcaba (Bannayan-Riley-Ruvalcaba syndrome)
CASP8	caspase-8
CHEK2	CHK2 checkpoint homolog ( <i>S. pombe</i> )
CS	syndrome de Cowden (Cowden syndrome)
DSB	bris d'ADN double brins (double-strand breaks)
dsRBD	double-stranded RNA binding domain
ER	récepteur des estrogènes (estrogen receptor)
ESE	activateurs d'épissage exoniques (exonic splicing enhancer)
ESS	répresseurs d'épissage exoniques (exonic splicing silencer)
GGR	réparation génomique globale (global genomic repair)
FAAP	FANCA associated polypeptide
FAK	focal adhesion kinase
FAT	FRAP, ATM, TRAPP
FGF	Fibroblast Growth Factors
FHA	fork head-associated domain
FRAP	FKBP-rapamycin associated protein
HCF-1	host cell factor 1
hnRNP	heterogenous nuclear ribonucleoprotein particles
HR	recombinaison homologue (homologous recombination)
ICL	ponts inter-chainés (interstrand crosslink)
IGF	facteurs de croissance insulín-like (insulin-like growth factor)
IR	radiations ionisantes (ionizing radiations)
ISE	activateurs d'épissage introniques (intronic splicing enhancer)
ISS	répresseurs d'épissage introniques (intronic splicing silencer)
LFS	syndrome de Li-Fraumeni (Li-Fraumeni syndrome)
LFS2, LFS-like	Li-Fraumeni like syndrome
LD	déséquilibre de liaison (linkage disequilibrium)
<i>MAP3K1</i>	mitogen activated protein kinase kinase kinase 1
MDC1	mediator of DNA damage checkpoint protein 1
MLPA	multiplex ligation-dependant probe amplification
MO25	mouse protein 25
MRN	MRE11/RAD50/NBS1
NBS1	Nijmegen breakage syndrome 1 (nibrin)
NER	réparation par excision de nucléotides (nucleotide excision repair)

NES	signal d'exportation nucléaire (nuclear export signal)
NHEJ	ressection non-homologue des extrémités d'ADN (non homologous end joining)
OR	rapport de cotes (odd ratio)
PALB2	partner and localizer of BRCA2
PCAF	P300/CBP-associated factor
PHTS	PTEN hamartoma-tumor syndrome
PI3K	phosphatidylinositol 3-kinase
PIP3	phosphoinositol 3,4,5-triphosphate
PJS	syndrome de Peutz-Jeghers (Peutz-Jeghers syndrome)
PP1	protéine phosphatase 1
protéines SR	protéines riches en sérine et arginine (serine/arginine-rich proteins)
RPA	replication protein A
RR	risque relatif (relative risk)
ser/thr kinase	sérine/thréonine kinase
snRNP	petites ribonucléoprotéines nucléaires (small nuclear ribonucleoprotein)
STRAD	ste20-related adaptor protein
TCR	réparation couplée à la transcription (transcription-coupled repair)
TFBS	sites de liaison de facteurs de transcription (transcription factors binding sites)
THBS1	thrombospondin 1
THR	traitements hormonaux de remplacement
TRAPP	transport protein particle
tSNPs	SNPs marqueurs (tagging SNPs)
voie AF/BRCA	voie anémie de Fanconi/BRCA (FA/BRCA pathway)
VUS	variants de significativité inconnue (variants of unknown significance)

# **CHAPITRE I - INTRODUCTION**

## **1. Le cancer**

### **1.1. Statistiques sur le cancer**

Le cancer est une maladie qui atteint un grand nombre de Canadiennes et de Canadiens. Selon les taux d'incidence actuels, 40% des Canadiennes et 45% des Canadiens seront atteints au cours de leur vie. Le quart des Canadiens mourront du cancer, soit 24% des femmes et près de 29% des hommes. Il a été estimé qu'au Canada seulement, lors de l'année 2008, 166 400 nouveaux cas de cancer auraient été diagnostiqués alors que 73 800 décès seraient dus à cette maladie<sup>1</sup>. Chaque année, un nombre grandissant de nouveaux cas sont diagnostiqués. Cette augmentation est principalement attribuable à la croissance démographique et au vieillissement de la population. Toutefois, la mortalité par cancer diminue chez les hommes de tous les âges et chez les femmes de moins de 70 ans.

Au cours de sa vie, 1 femme sur 9 sera atteinte du cancer du sein. Ce cancer se classe au premier rang pour l'incidence chez les Canadiennes, les 22 400 nouveaux cas équivalant au double de celui des nouveaux cas de cancer du poumon qui est néanmoins la principale cause de décès par cancer chez les Canadiennes en 2008. Il a été estimé que 9 200 décès lui seront attribuables, comparativement à 5 300 pour le cancer du sein. Les taux de mortalité dus au cancer du sein chez la femme diminuent depuis le milieu des années 80. Cette tendance est sans doute attribuable à la fois au dépistage par mammographie et à l'efficacité accrue des traitements. Le taux de mortalité par cancer du sein a atteint le niveau le plus bas jamais enregistré depuis 1950. Une tendance similaire à la baisse est observée aux États-Unis, au Royaume-Uni et en Australie<sup>1</sup>.

À l'échelle de la planète, le cancer du sein est le cancer le plus fréquent chez les femmes avec près d'un million de nouveaux cas par année<sup>2</sup>. Pour les États-Unis, il a été estimé que 182 000 nouveaux cas auraient été diagnostiqués en 2008<sup>3</sup>.

### **1.2. Développement du cancer et formation de tumeurs**

Le processus de formation et de développement d'une tumeur est un processus complexe qui requiert une suite d'événements favorisant la tumorigenèse. Ces événements peuvent être retrouvés dans tous les cancers<sup>4</sup>. Le cancer se développe à partir d'une cellule

somatique et de sa progéniture. Les cellules néoplasiques dérivant de cette cellule accumulent alors une série de changements génétiques et épigénétiques qui mènent à des changements dans l'activité de gènes puis à des phénotypes cellulaires sujets à une sélection<sup>5</sup>. Ultimement, une cellule évolue de manière à échapper au contrôle normal de la prolifération, ce qui permet le développement du cancer.

Selon le modèle classique de développement tumoral en plusieurs étapes chez l'homme<sup>6</sup>, des changements importants survenant tant sur une petite échelle (mutations ponctuelles) qu'à grande échelle (grands réarrangements chromosomiques) confèrent un avantage sélectif à une cellule, ce qui lui permettra d'accumuler d'autres altérations menant à la malignité. Ces étapes importantes comprennent sept fonctions que les cellules cancéreuses doivent acquérir : 1) elles doivent acquérir la capacité de devenir résistantes aux facteurs causant l'inhibition de contact, 2) proliférer en absence de facteurs de croissance exogènes, 3) envahir et devenir métastatiques, 4) développer un potentiel illimité de réplifications, 5) s'évader du mécanisme d'apoptose, 6) créer l'angiogenèse de manière à être alimentées en sang et 7) s'évader de la destruction par les cellules du système immunitaire. L'acquisition de ces fonctions est généralement favorisée par l'instabilité génomique trouvée chez les cellules cancéreuses humaines<sup>7</sup>, ce qui peut être causé par des mutations héritées dans des gènes assurant l'intégrité du génome ou des mutations somatiques acquises lors du développement de la tumeur. L'accumulation d'une combinaison d'altérations génétiques dans un sous-groupe de cellules à haut potentiel de progression tumorale pourrait donc être crucial dans la progression d'une tumeur<sup>8</sup>.

## **2. Le cancer, une maladie multifactorielle**

### **2.1. Facteurs environnementaux**

L'expression désordonnée des gènes peut être causée par des mutations génétiques ou des facteurs épigénétiques qui peuvent réduire ou activer l'expression de gènes cibles. La grande variation des taux d'incidence et de mortalité dans plusieurs populations à travers le monde met en évidence la contribution de la susceptibilité génétique dans l'étiologie du cancer du sein. Les études migratoires indiquent par contre que le niveau de risque des immigrants devient similaire à celui des habitants du pays d'accueil, ce qui suggère que l'environnement et l'alimentation influencent aussi le risque de cancer du sein. Le cancer du sein apparaît donc comme une maladie multifactorielle dans laquelle il y a une intervention simultanée de facteurs génétiques et environnementaux<sup>9</sup>.

Certains facteurs environnementaux sont reconnus pour conférer un risque accru de cancer du sein alors que d'autres semblent avoir un effet protecteur (Tableau 1). Parmi les facteurs environnementaux reconnus comme conférant un risque accru de cancer du sein, plusieurs sont en lien de près ou de loin avec l'exposition aux estrogènes. Les composants de la nourriture que nous ingérons peuvent aussi agir comme agents mutagènes ou interférer avec l'action de mutagènes externes. Ces facteurs nutritionnels peuvent également influencer la prolifération et la croissance d'une population de cellules ou modifier l'environnement cellulaire en altérant les niveaux d'hormones. La régulation de ces processus par différents nutriments peut influencer l'expression de gènes centraux en causant la méthylation des régions promotrices ou l'acétylation des histones, ce qui influence la structure de l'ADN et l'accessibilité des gènes pour la transcription<sup>10</sup>.

**Tableau 1**  
**Facteurs environnementaux modifiant le risque de cancer du sein.**

<b>Facteurs de risque</b>	<b>Effet sur le risque</b>
<b>Facteurs démographiques et sanitaires</b>	
Âge <sup>11</sup>	↑
Forte densité mammaire <sup>12</sup>	↑
Exposition aux radiations ionisantes <sup>13</sup>	↑
<b>Hormonaux</b>	
Âge précoce aux premières menstruations <sup>14</sup>	↑
Ménopause tardive <sup>15,16</sup>	↑
Utilisation de contraceptifs oraux <sup>15</sup>	↑
Traitement hormonaux de remplacement <sup>18</sup>	↑
<b>Reproductifs</b>	
Jeune âge à la première grossesse <sup>19</sup>	↓
Plusieurs grossesses à terme <sup>20</sup>	↓
Allaitement maternel <sup>16</sup>	↓
<b>Habitudes de vie et nutrition</b>	
Niveau élevé d'activité physique <sup>24-26</sup>	↓
Obésité et surplus de poids <sup>24</sup>	↑
Consommation d'alcool <sup>31</sup>	↑
Consommation de thé vert <sup>33</sup>	↓
Consommation élevée de lycopène <sup>35,36</sup>	↓

### 2.1.1. Facteurs démographiques et sanitaires.

L'âge est le facteur de risque le plus important vis-à-vis du cancer du sein<sup>11</sup>. La maladie est rare chez les femmes de moins de 30 ans, alors que le risque augmente entre 50 et 75 ans, représentant près de deux tiers de tous les cas de cancer du sein. Le risque de cancer du sein augmente aussi avec le niveau de densité des tissus mammaires, le risque étant multiplié de deux à six fois chez les femmes ayant des seins denses en mammographie<sup>12</sup>. Il a été démontré que l'exposition du tissu mammaire aux radiations ionisantes (ionizing radiations; IR) endommageant l'ADN avant l'âge de 40 ans est susceptible de provoquer un cancer du sein dans les années ultérieures<sup>13</sup>.

### *2.1.2. Facteurs hormonaux*

#### Facteurs hormonaux endogènes

Un âge précoce aux premières menstruations (survenues avant l'âge de 12 ans) augmente le risque de cancer du sein<sup>14</sup>. Le fondement biologique de cette association correspond à l'exposition précoce et prolongée aux hormones durant la période d'activité des ovaires. D'un autre côté, une exposition prolongée aux estrogènes en raison d'une ménopause tardive résulte également en un risque accru de cancer du sein<sup>15,16</sup>, venant ainsi appuyer l'utilisation des anti-estrogènes en médecine préventive<sup>17</sup>.

#### Facteurs hormonaux exogènes

Des facteurs hormonaux exogènes sont également impliqués dans le risque de cancer du sein. Le risque de cancer du sein est augmenté d'environ 25% chez les femmes utilisant couramment des contraceptifs oraux. Cependant, cette augmentation du risque chute dès l'arrêt de la consommation, de sorte que 10 ans après l'arrêt de l'utilisation, aucune augmentation significative n'est constatée<sup>15</sup>. L'utilisation de traitements hormonaux de remplacement (THR), qui retarde les effets de la ménopause, est également à la base d'un risque accru de cancer du sein<sup>18</sup>.

#### Facteurs reliés à la reproduction

Le fait d'avoir mené à terme plusieurs grossesses ou d'avoir eu la première maternité à un jeune âge a un effet protecteur sur le cancer du sein. En effet, les femmes ayant mené au moins une grossesse à terme avant l'âge de 30 ans présentent en moyenne un risque de cancer du sein diminué de 25% par rapport aux femmes n'ayant pas eu d'enfant<sup>19</sup>. De plus, l'effet protecteur de l'accouchement semble augmenter proportionnellement au nombre d'accouchements<sup>20</sup>. La grossesse provoque une différenciation accélérée du tissu mammaire et une prolifération rapide de l'épithélium<sup>21</sup>.

L'effet de l'allaitement maternel sur le risque de cancer du sein est controversé. Certains résultats ont toutefois identifié que les femmes qui ont allaité pendant une durée totale d'au moins 25 mois présentent un risque de cancer du sein réduit de 33% par rapport à celles qui n'ont jamais allaité<sup>16</sup>. Plusieurs mécanismes pourraient expliquer l'association inverse entre l'allaitement et le risque de cancer du sein. La lactation produit des changements

hormonaux endogènes, en particulier une réduction d'estrogènes et une augmentation de la production de prolactine, qui diminuent l'exposition cumulative aux estrogènes chez la femme<sup>22</sup>. L'effet protecteur de l'allaitement pourrait également être attribuable à son rôle dans le décalage du rétablissement de l'ovulation.

### *2.1.3. Facteurs liés aux habitudes de vie et de nutrition*

Le World Cancer Research Fund a publié son deuxième rapport sur le rôle des aliments, de la nutrition et de l'activité physique dans la prévention du cancer<sup>23</sup>. Cet organisme incite à adopter une saine alimentation et un suivi nutritionnel supervisé pour les patients atteints du cancer. Un niveau élevé d'exercice a été lié à un risque diminué de cancer post-ménopause variant en moyenne entre 20 et 40%<sup>24-26</sup>. Les mécanismes biologiques par lesquels l'activité physique serait associée à une diminution de risque impliqueraient une réduction de la production d'estrogènes et le maintien de l'équilibre énergétique<sup>27</sup>. L'activité physique pourrait moduler les niveaux d'insuline et de facteurs de croissance insulin-like (insulin-like growth factor; IGF), augmenter la réponse immunitaire<sup>28</sup> et diminuer la prise de poids, en particulier après la ménopause. Bien que diminuer la prise alimentaire et augmenter la quantité d'activité physique sont efficaces afin de réduire le poids, l'activité physique semble, selon certaines études, diminuer le gras intra-abdominal<sup>29</sup>.

Un poids plus élevé serait associé à un risque accru de cancer du côlon, du rein, de l'œsophage, de l'endomètre, de la thyroïde et de cancer du sein post-ménopause<sup>24</sup>. En étant associé à un profil hormonal soupçonné de favoriser le développement du cancer du sein, l'obésité augmente d'environ 50% le risque de cancer du sein chez les femmes ménopausées, probablement en raison de l'augmentation de concentrations sériques d'estradiol libre<sup>30</sup>. Après la ménopause, l'aromatation des androgènes dans le tissu adipeux est l'une des plus importantes sources d'estrogènes circulants.

L'alcool est le seul facteur naturel établi de risque de cancer du sein. Ce risque augmente d'environ 7% pour une consommation moyenne d'une boisson alcoolisée par jour<sup>31</sup>. Les femmes atteintes d'un cancer du sein et consommant au moins une boisson alcoolisée par jour ont une durée de survie diminuée de 15% à 40%<sup>32</sup>. L'alcool provoque une



augmentation des hormones dans le sérum et une production accrue des facteurs de croissance IGF. Les IGF agissent comme des mitogènes, inhibent l'apoptose et interagissent avec les estrogènes.

L'association entre le risque de cancer du sein et les principales composantes de l'alimentation humaine incluant les fruits et les légumes, les produits laitiers, la viande, les vitamines, les fibres et les phyto-estrogènes a fait l'objet de nombreuses études. D'une manière générale, les résultats restent discordants<sup>14</sup>.

Un rôle possible pour le thé vert dans la prévention du cancer a été suggéré sur la base de l'incidence plus faible du cancer du sein dans certaines régions telles le Japon et la Chine où la consommation de thé vert est élevée. Une réduction significative du cancer du sein associée à la consommation de thé vert, a été identifiée chez des Américaines d'origine asiatique<sup>33</sup>. Le mécanisme potentiel par lequel le thé vert protégerait du cancer passerait par la séquestration des espèces réactives de l'oxygène, l'inhibition de la formation de carcinogènes, une augmentation de la détoxification des carcinogènes, une augmentation de la réparation de l'ADN et par la modulation du métabolisme des estrogènes. De plus, le thé vert a été lié à un arrêt du cycle cellulaire et à l'apoptose<sup>34</sup>.

Une supplémentation en lycopène chez des sujets sains a démontré une fréquence diminuée des bris d'ADN<sup>35</sup>, ce qui a mené à la suggestion qu'une alimentation riche en lycopène pourrait réduire le risque de différentes maladies en diminuant les dommages oxydatifs à l'ADN. Une revue des études épidémiologiques a suggéré une association inverse entre la consommation de tomates, la consommation de lycopène ou le niveau de lycopène sanguin et le risque de cancer du sein<sup>36</sup>.

## **2.2. Facteurs génétiques impliqués dans le cancer du sein**

Des études épidémiologiques populationnelles ont mené à l'élaboration de modèles génétiques plutôt qu'environnementaux afin d'expliquer l'occurrence familiale du cancer<sup>37,38</sup>. Il a été démontré que certains facteurs génétiques pourraient expliquer une partie substantielle des cancers. La démonstration de liaison génétique chez des familles présentant plusieurs cas de cancer du sein<sup>39</sup> a dissipé les doutes à propos de l'implication des facteurs génétiques dans le cancer. D'autres études épidémiologiques ont montré que

les femmes ayant une apparentée de premier degré atteinte du cancer du sein ont un risque de cancer du sein approximativement deux fois plus élevé en comparaison avec la population en général<sup>16</sup>. Les études menées chez des jumeaux indiquent que la majeure partie de l'excès du risque familial est attribuable à une prédisposition héritée<sup>40</sup>. Les jumeaux monozygotes étant identiques et les jumeaux dizygotes partageant la moitié de leur bagage génétique, la susceptibilité accrue en raison de facteurs génétiques devrait se refléter par une meilleure concordance de la maladie chez les jumeaux monozygotes que dizygotes<sup>41</sup>. Ainsi, les études sur le cancer du sein menées chez des jumelles<sup>42,43</sup>, l'incidence de tumeurs contralatérales chez des personnes ayant déjà été atteintes par un cancer du sein<sup>42,43</sup> et la distribution de l'héritabilité chez certaines familles<sup>44</sup> ont suggéré la prédominance des facteurs génétiques afin d'expliquer l'aggrégation familiale du cancer du sein.

### 3. Gènes de susceptibilité au cancer du sein

#### 3.1. Gènes majeurs de susceptibilité au cancer du sein

Les gènes *BRCA1* et *BRCA2* sont les deux gènes majeurs de susceptibilité au cancer du sein. La fréquence des mutations de ces gènes est faible, mais certaines mutations fondatrices relativement fréquentes ont été retrouvées chez différentes populations, comme par exemple chez les juifs Ashkenazes et chez les Canadiens français<sup>45</sup>.

##### *BRCA1 et BRCA2, découverte et risque associé*

Des analyses de liaison menées chez de grandes familles à risque élevé présentant plusieurs cas de cancer du sein ont mené à l'identification de *BRCA1* et *BRCA2* par clonage positionnel<sup>46-48</sup>. Ces gènes de susceptibilité agissent d'une manière autosomique dominante. Au niveau cellulaire, les gènes *BRCA1* et *BRCA2* agissent d'une manière récessive, l'allèle sauvage étant inactivé dans le cancer.

Des mutations de *BRCA1* et *BRCA2* confèrent un risque de cancer du sein accru de 10 à 20 fois<sup>49</sup>. Les femmes porteuses de mutations ont un risque de développer un cancer du sein entre 37 et 85% au cours de leur vie. Le risque de cancer de l'ovaire chez ces femmes est estimé entre 15 et 40%<sup>50,51</sup>. Les estimations faites à partir de cas de cancer du sein choisis sur une base populationnelle et non-sélectionnés sur la base d'une histoire familiale ont indiqué un risque moindre. Le risque associé serait de 65% (51%-75%) chez les porteurs de mutations dans le gène *BRCA1* et de 45% (33%-54%) pour *BRCA2*<sup>52</sup>. Le risque relatif de cancer du sein chez l'homme est élevé chez les porteurs de mutations dans ces gènes, particulièrement chez les porteurs de mutations dans *BRCA2*<sup>53</sup>. De plus un risque élevé de cancer de la prostate a également été rapporté, spécialement chez les porteurs de mutations dans *BRCA2* et avant l'âge de 65 ans<sup>54,55</sup>. Un risque accru de cancer du pancréas, des trompes de Fallope, de l'estomac et du côlon a été associé à des mutations dans *BRCA1* alors que des mutations dans *BRCA2* confèrent un risque accru de cancer de l'estomac, de la vessie, du pancréas et du pharynx<sup>54,56</sup>. Des mutations de l'un de ces gènes seraient retrouvées chez 1 personne sur 1000 au Royaume-Uni. Ceci permettrait d'expliquer 15 à 20% de l'excès de risque familial du cancer du sein<sup>44,57,58-60</sup>.

Dans la présente section, un survol de la structure et des fonctions des gènes *BRCA1* et *BRCA2* sera effectué. Le tout sera accompagné d'une figure récapitulant les fonctions de ces deux gènes.

### 3.1.1. *BRCA1*

#### Structure de BRCA1

Le gène *BRCA1* est composé de 24 exons situés en 17q21. La protéine BRCA1 est une protéine de 1863 résidus possédant plusieurs caractéristiques structurales. Le domaine RING situé en N-terminal de la protéine a été démontré comme étant impliqué dans un grand nombre d'interactions protéine-protéine, notamment avec l'ubiquitine ligase BARD1<sup>61</sup>. La région C-terminale de BRCA1 contient un motif répété d'environ 110 acides aminés (acides aminés 1646-1859 de BRCA1), appelé les répétitions BRCT (BRCT repeats; BRCA1 C-terminal repeats)<sup>62,63</sup>. Ces séquences répétées sont retrouvées dans des protéines impliquées dans la réparation de l'ADN et dans la régulation du cycle cellulaire. La protéine BRCA1 peut agir à différents niveaux dans la réparation de l'ADN. Elle agit à titre de transducteur de signal, à titre d'effecteur de la réparation des dommages à l'ADN ou de point de contrôle du cycle cellulaire. Elle agit aussi comme régulateur des protéines impliquées dans la réparation, tel les protéines du BASC (BRCA1-associated genome surveillance complex)<sup>64</sup>. Dans cette section, les différents rôles de BRCA1 seront passés en revue et une attention particulière sera portée sur certains gènes candidats du cancer du sein.

#### Rôle de BRCA1 dans la réparation des bris d'ADN double brins

Une réparation efficace des bris d'ADN requiert un mécanisme de reconnaissance des dommages et des signaux de transduction activant des effecteurs qui arrêtent le cycle cellulaire aux points de restriction et la réparation de dommages<sup>65,66</sup>. Au moins deux processus distincts contribuent à la réparation des bris d'ADN double brins (double-strand breaks; DSB) chez les mammifères, la recombinaison homologue (homologous recombination; HR) et la ressection non-homologue des extrémités d'ADN (non homologous end joining; NHEJ).

BRCA1 localise aux foci nucléaires lors de la phase S du cycle cellulaire. Des dommages à l'ADN résultent en une hyperphosphorylation de BRCA1 et en une localisation aux foci nucléaires<sup>67,68</sup>. BRCA1 colocalise avec l'histone H2AX, une protéine rapidement phosphorylée aux sites de bris double brins suite aux dommages à l'ADN<sup>69</sup>. De plus, BRCA1 colocalise avec RAD51, un facteur participant à la HR, et avec le complexe MRN (MRE11/RAD50/NBS1) qui est impliqué à la fois dans la HR et dans la NHEJ<sup>70,71</sup>. BRCA1 lie les structures ramifiées d'ADN et la liaison de BRCA1 à l'ADN inhibe l'activité exonucléase du complexe MRN<sup>72</sup>. Cette observation suggère que BRCA1 s'accumule aux sites de bris d'ADN, forme différents complexes protéiques, et est impliquée dans différents aspects de la réparation de l'ADN. Des mutations de *BRCA1* ne résulteraient pas directement en la formation de tumeurs, mais causeraient plutôt de l'instabilité génétique, ce qui favoriserait la transformation cellulaire<sup>73</sup>.

#### Implication de BRCA1 dans la recombinaison homologue

Lors de la HR, une forme très efficace de réparation, des séquences identiques ou similaires sont utilisées pour diriger la réparation de l'ADN endommagé par un mécanisme de conversion ce qui permet la réparation des bris sans perte d'information ou de modification à la séquence d'ADN<sup>74</sup>. Dans ce mécanisme de réparation, les kinases ATM (ataxia telangiectasia mutated) et ATR (ataxia telangiectasia and Rad3 related) jouent un rôle central en contrôlant la phosphorylation de plusieurs protéines impliquées dans les points de contrôles du cycle cellulaire et la réparation des dommages à l'ADN, tel p53, NBS1 et CHEK2 ainsi que BRCA1 et BRCA2<sup>75,76</sup>.

Dans la réparation des bris d'ADN par HR, les protéines RAD51 et RPA (replication protein A) sont essentielles. Le rôle spécifique de BRCA1 dans la réparation des bris d'ADN double brins peut être défini par son interaction avec les protéines BRIP1/FANCD1 et BARD1. Ces deux protéines ont un impact sur l'efficacité de la réparation de l'ADN. Le gène *BRIP1* a récemment été identifié comme gène de susceptibilité au cancer du sein alors que le gène *BARD1* a été étudié à titre de gène candidat dans le cadre de mon projet.

### Jonction non-homologue des extrémités

Dans l'éventualité où la HR est impossible ou inefficace, un mécanisme alternatif de réparation est utilisé, la NHEJ<sup>77</sup>. Dans la voie de réparation par NHEJ, la protéine kinase dépendante de l'ADN, DNA-PK, et les protéines XRCC5 et XRCC6 sont recrutées aux sites de bris d'ADN double brins. Ceci permet l'arrivée subséquente de la ligase XRCC4. Deux voies de NHEJ existeraient, l'une favorisant l'incorporation d'erreur dans l'ADN (error-prone NHEJ) et l'autre n'incorporant pas d'erreur (error-free NHEJ)<sup>78,79</sup>. BRCA1 serait impliquée dans la réparation « error-free » NHEJ ou inhiberait le mécanisme « error-prone » NHEJ<sup>80,81</sup>. L'interaction de BRCA1 avec XRCC5 permet l'accumulation rapide de BRCA1 aux DSBs<sup>82</sup>. De plus, l'interaction de BRCA1 est mutuellement exclusive à XRCC5 ou à BARD1<sup>82</sup>.

### Réparation par excision de nucléotides (réparation couplée à la transcription et réparation génomique globale)

La protéine BRCA1 pourrait également jouer un rôle dans la réparation par excision de nucléotides (nucleotide excision repair; NER)<sup>83</sup>. Les bases endommagées sont enlevées par la nucléase d'excision, une enzyme composée de plusieurs sous-unités ciblant les lésions sur le brin endommagé<sup>83-85</sup>. La NER peut être sous-divisée en deux principaux mécanismes, la réparation couplée à la transcription (transcription-coupled repair; TCR), et la réparation génomique globale (global genomic repair; GGR)<sup>83</sup>. La TCR est le processus par lequel des lésions sur des bases de l'ADN causées par des rayons X ou une oxydation des bases sont enlevées préférentiellement du brin d'ADN transcrit en utilisant le brin non-transcrit comme gabarit. La TCR agirait sur les brins transcrits des gènes actifs<sup>86</sup> et serait activée lorsque la machinerie transcriptionnelle basale est arrêtée par des dommages à l'ADN. BRCA1 serait essentiel pour la TCR<sup>87</sup> et pourrait être connectée à la machinerie transcriptionnelle basale puisque la protéine BRCA1 co-purifie avec la RNA polymérase II holoenzyme par l'interaction qu'elle fait avec RHA, la composante hélicase de l'enzyme<sup>88</sup>. Le gène *RHA*, nommé aussi *DHX9*, est d'ailleurs un gène candidat intéressant pour l'identification de nouveaux gènes de susceptibilité au cancer du sein. L'implication de BRCA1 dans la régulation de la GGR<sup>89</sup>, un mécanisme de réparation ciblant et enlevant les lésions à l'échelle du génome entier<sup>83</sup>, s'effectuerait possiblement en contrôlant la transcription des gènes de la NER, tels *XPC* et *DDB2*<sup>89</sup>.

### Régulation transcriptionnelle par BRCA1

La protéine BRCA1 a été associée à la régulation transcriptionnelle de plusieurs gènes activés en réponse aux dommages à l'ADN. Lorsque BRCA1 est surexprimé, il supprime la transactivation par ER (estrogen receptor)<sup>90</sup>. BRCA1 peut s'associer avec le facteur de transcription ZBRK1 pour former un complexe co-répresseur ciblant un motif spécifique du gène *GADD45*. BRCA1 sert de relais à la détection et au signalement de dommages à l'ADN par ATM à la régulation transcriptionnelle de *GADD45*. De plus, il a été démontré que BRCA1 interagit physiquement et régule certains facteurs de transcription tels p53, c-myc et STAT1<sup>91-93</sup>.

### Remodelage de la chromatine par BRCA1

BRCA1 contribue à la réponse aux dommages à l'ADN par son interaction avec des protéines altérant la structure de l'ADN et de la chromatine<sup>94</sup>. Une association de BRCA1 avec les protéines SWI/SNF et avec des kinases régulatrices de l'acétylation et de la déacétylation a été démontrée<sup>95,96</sup>. BRCA1 interagit également avec BLM, une protéine participant à la réplication de l'ADN<sup>64</sup> et avec l'hélicase BRIP1<sup>97</sup>. La protéine BRIP1, nommée aussi FANCI car impliquée dans l'anémie de Fanconi (AF), est un gène de susceptibilité au cancer du sein qui a été étudié dans le cadre de mes études doctorales. Les rôles de BRCA1 dans la transcription et la réparation de l'ADN pourraient donc être reliés au remodelage des structures d'ADN autour de la machinerie basale de transcription aux fourches de réplication bloquées et aux bris d'ADN double brins.

### Implication de BRCA1 dans le contrôle du cycle cellulaire

L'habileté à contrôler précisément les événements du cycle cellulaire est essentielle pour le maintien de l'intégrité génomique et pour prévenir l'apparition de mutations affectant la croissance normale des cellules. Les cellules déficientes en *BRCA1* démontrent une réparation des dommages à l'ADN diminuée, une amplification des centrosomes et des anomalies des points de restriction du cycle cellulaire<sup>73,98</sup>. L'instabilité génétique causée par une déficience en *BRCA1* déclenche la réponse cellulaire aux dommages à l'ADN ce qui bloque la prolifération cellulaire et induit l'apoptose<sup>99</sup>. Une déficience en *BRCA1* cause des anomalies du point de restriction de la phase S, de celui de la phase G2/M, du point de

contrôle de la formation des fuseaux mitotiques (point de contrôle métaphase/anaphase) et la duplication du centrosome.

La protéine BRCA1 est hyperphosphorylée lors des phases G1 et S, et est déphosphorylée après la phase M<sup>100</sup>. Les protéines ATM, ATR et CHEK2 sont responsables de la phosphorylation de BRCA1<sup>75,76,101,102</sup>. Il a été démontré que p53 est impliquée dans l'induction par BRCA1 du point de restriction G1/S du cycle cellulaire<sup>103</sup>. La phosphorylation de BRCA1 par ATM est requise pour l'activation du point de restriction G2/M<sup>104,105</sup>. BRCA1 interagit aussi avec plusieurs protéines jouant un rôle essentiel dans les points de restriction en phase S, tel l'histone H2AX et le complexe MRN, qui forment des foci nucléaires suite aux IR et arrêtent les cellules en phase S<sup>106</sup>. L'un des gènes que nous avons étudié à titre de gènes candidats du cancer du sein est le gène *AURKA*. La protéine AURKA/Aurora-A, une sérine/thréonine kinase (ser/thr kinase) régulant la progression mitotique<sup>107-110</sup>, lie physiquement et phosphoryle BRCA1. Cet événement corrèle avec une fonction altérée de BRCA1 à réguler la transition G2/M<sup>111</sup>.

Le point de contrôle de la formation des fuseaux mitotiques assure l'efficacité de la ségrégation des chromosomes. Il empêche les cellules contenant des chromosomes mal alignés de procéder à la mitose. Lors de la métaphase, les chromatides soeurs sont attachés aux fuseaux mitotiques par les kinétochores. Le point de contrôle des fuseaux supervise l'attachement des chromatides soeurs. Les anomalies chromosomiques chez les cellules déficientes en *BRCA1* ainsi que la poursuite de l'anaphase et de la télophase suggèrent que le point de contrôle de la formation des fuseaux est déficient<sup>112</sup>. Les défauts en points de contrôle du cycle cellulaire, combinés avec les duplications anormales des centrosomes et les défauts de réparation des dommages à l'ADN pourraient être la cause de l'instabilité génétique présente dans les cellules déficientes en *BRCA1*.

### 3.1.2. *BRCA2*

#### Structure de BRCA2

Le gène *BRCA2*, localisé en 13q12-q13<sup>113</sup>, est composé de 27 exons distribués sur une distance de 70 kb. Le clonage de la protéine *BRCA2* possédant 3418 résidus a permis de déterminer que *BRCA2* ne possède pas de similarité à d'autres protéines connues<sup>47,48</sup>. Huit



motifs de 30-40 résidus encodés par l'exon 11 et appelés les répétitions BRC sont conservés à travers plusieurs espèces de mammifères<sup>114</sup>. Ces répétitions permettent l'interaction de BRCA2 avec RAD51, une protéine essentielle à la réparation de l'ADN et à la recombinaison.

#### Rôle de BRCA2 dans la recombinaison homologue

Le gène *BRCA2* est requis pour les niveaux normaux de réparation de l'ADN par HR<sup>77,115</sup>. Les cellules mutantes *BRCA2* démontrent un haut degré d'aberrations chromosomiques induites et spontanées<sup>116-118</sup>, ainsi qu'une sensibilité accrue aux agents endommageant l'ADN<sup>118-121</sup>, incluant les IR et les agents pontants de l'ADN.

L'implication de BRCA2 dans la réparation des bris d'ADN double brins a principalement été confirmée par son interaction avec la recombinase RAD51<sup>120,122,123</sup>. RAD51 favorise l'échange des brins provenant de deux différentes molécules d'ADN lors de la HR<sup>124</sup>. Il a été démontré que RAD51 et BRCA2 colocalisent aux foci nucléaires suite aux dommages à l'ADN causés par des IR<sup>125</sup>. BRCA2 serait nécessaire au recrutement de RAD51 aux sites de bris de l'ADN et à la fonction de RAD51<sup>120,126</sup>.

#### Régulation et liaison de RAD51 par sa liaison aux motifs BRC de BRCA2

Certains des motifs BRC de BRCA2 interagissent avec le domaine central de RAD51<sup>123,127</sup>. RAD51 s'associe également avec une série de résidus localisés dans la région C-terminale de BRCA2<sup>120</sup>. Selon un modèle suggéré récemment, RAD51 lierait les BRC de BRCA2 dans une forme inactive et transloquerait vers la région C-terminale lorsque nécessaire. Lorsque lié à la région C-terminale de BRCA2, RAD51 serait sous sa forme active et serait libéré en réponse aux dommages à l'ADN. RAD51 pourrait ensuite se lier à l'ADN simple brin au site de bris et former les filaments de nucléoprotéines. Les protéines CHEK1 et CHEK2 peuvent phosphoryler la région C-terminale de BRCA2 et seraient impliquées dans le contrôle de ces interactions entre RAD51 et BRCA2<sup>128</sup>.

#### Rôle de BRCA2 dans les autres mécanismes de réparation

La protéine BRCA2 ne serait pas essentielle à la réparation des bris d'ADN double brins par NHEJ<sup>77,129</sup> ou pour la résection des bris double brins induits par des radiations<sup>64</sup>. Les cellules déficientes en BRCA2 démontrent des effets moins sévères sur la TCR que les

cellules déficientes en BRCA1<sup>87</sup>. Toutefois, les souris déficientes en *BRCA1* ou en *BRCA2* à l'état homozygote démontrent toutes deux un phénotype de mortalité embryonnaire<sup>130</sup>.

#### Régulation transcriptionnelle par BRCA2

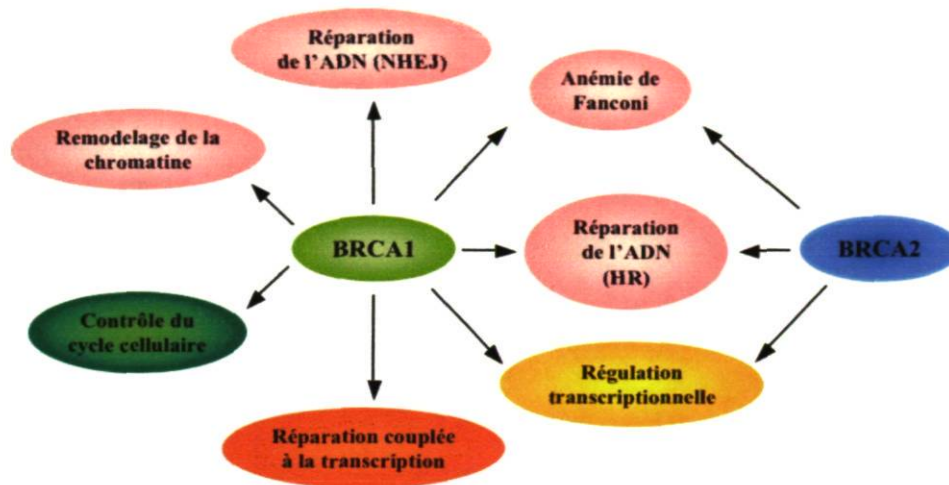
La protéine BRCA2 est impliquée à différents niveaux dans la régulation transcriptionnelle de plusieurs gènes activés en réponse aux dommages à l'ADN. La région N-terminale de BRCA2, une région importante pour son activité de transactivation, interagit avec la protéine EMSY<sup>131</sup>. Un excès d'EMSY atténue cette fonction de BRCA2<sup>131</sup>. De plus, il a été démontré que BRCA2 exprimée de manière endogène peut stimuler la transcription de gènes régulés par le récepteur aux androgènes<sup>132</sup>.

#### Implication de BRCA2 dans l'anémie de Fanconi

L'AF est une maladie récessive rare caractérisée par des malformations congénitales, une défaillance progressive de la moëlle osseuse et une prédisposition au cancer. L'AF est une maladie génétique hétérogène pouvant être causée par des mutations bialléliques dans 13 gènes différents. Les cellules des patients atteints de l'AF démontrent un phénotype cellulaire d'instabilité génomique semblable à celui des cellules déficientes en *BRCA1* ou *BRCA2*. La première évidence démontrant un lien entre les protéines de l'AF et les protéines BRCA1 et BRCA2 a été fournie par l'identification de mutations bialléliques de *BRCA2* comme étant responsables du sous-type D1 de l'AF<sup>133</sup>. Récemment, une nouvelle protéine a été identifiée sur la base de son interaction avec BRCA2. Cette protéine appelée PALB2 (partner and localizer of BRCA2) colocalise avec BRCA2 aux foci nucléaires et cause une localisation et une stabilisation de BRCA2 aux structures nucléaires, rendant ainsi possible la réparation par recombinaison et la fonction de BRCA2 dans les points de contrôles du cycle cellulaire<sup>134</sup>. Des mutations bialléliques tronquantes de PALB2 ont été retrouvées chez des familles atteintes du sous-type N de l'AF<sup>135,136</sup>.

Une implication de plus en plus importante des protéines de l'AF dans la susceptibilité au cancer du sein étant identifiée, *PALB2* est devenu rapidement un candidat intéressant. Ceci a permis d'identifier ce gène comme étant un gène de susceptibilité au cancer du sein à pénétrance modérée<sup>137,138</sup>. Une attention particulière lui sera portée subséquemment.

Il est donc possible de dégager que les protéines BRCA1 et BRCA2 possèdent de nombreuses fonctions et sont étroitement interreliées (Figure 1).



**Figure 1**  
**Fonctions des protéines BRCA1 et BRCA2.**

### 3.2. Autres gènes de susceptibilité au cancer du sein

#### 3.2.1. *TP53* - Syndrome de Li-Fraumeni

##### LFS et mutations germinales de *TP53*

Le syndrome de Li-Fraumeni (Li-Fraumeni syndrome, LFS; OMIM #151623) est un syndrome autosomique dominant associé à une prédisposition à plusieurs cancers. Le LFS et un autre syndrome y étant associé, le Li-Fraumeni like syndrome (LFS-like, LFS2), ont une fréquence estimée de 1/5000 dans la population en général<sup>139</sup>. Des mutations germinales dans le gène *TP53*, situé en 17p13.1, sont responsables du LFS. Les individus atteints du LFS ont un risque accru de développer des sarcomes, des cancers du sein, du cerveau et des tumeurs adénocorticales à un âge précoce<sup>140</sup>. Les patients atteints du LFS développent particulièrement un cancer du sein à un jeune âge, en moyenne vers 36 ans<sup>141,142</sup>. Le risque à vie de développer un cancer chez les patients atteints du LFS approche 100%. Il a été estimé que 50% des patients LFS développeront un cancer avant l'âge de 30 ans<sup>143</sup>. Des mutations dans *TP53* sont peu communes chez les familles atteintes de cancer du sein pour lesquelles aucun cas du LFS n'a été rapporté<sup>144,145</sup>.

### Rôles de p53

Le gène suppresseur de tumeurs *TP53* est un membre central dans la réponse au stress et préserve la stabilité génomique<sup>146</sup>. La protéine p53 agit afin d'engendrer une variété de processus anti-prolifératifs. Suite au stress cellulaire, p53 est stabilisée et se lie à l'ADN sous forme de tétramère engendrant ainsi une régulation transcriptionnelle des gènes cibles impliqués dans la régulation de processus cellulaires clés, tels la réparation de l'ADN, l'arrêt du cycle cellulaire, la sénescence et l'apoptose<sup>147,148</sup>.

Une des fonctions les plus importantes de p53 est sa capacité à activer l'apoptose. P53 favorise l'apoptose par des mécanismes dépendants et indépendants de la transcription qui agissent de concert pour assurer l'efficacité du programme de mort cellulaire. En se liant à des éléments de réponse spécifiques, p53 module la transcription de gènes qui gouvernent les défenses majeures contre la croissance tumorale, incluant l'arrêt du cycle cellulaire, l'apoptose, la maintenance de l'intégrité génétique, l'inhibition de l'angiogenèse et la sénescence cellulaire<sup>149</sup>. P53 interagit aussi avec plusieurs protéines impliquées dans le contrôle de la mort cellulaire programmée et ces interactions pourraient contribuer à son rôle d'inhibition de la tumorigenèse<sup>150-154</sup>.

En plus de son activité de transactivation de gènes favorisant directement l'apoptose, p53 peut aussi induire des gènes qui court-circuitent les voies de signalisation anti-apoptotiques. L'exemple le plus évident de l'action de p53 est la régulation de PTEN, un régulateur négatif de la voie PI3K kinase aussi responsable d'un syndrome prédisposant au cancer du sein. PTEN est une phosphatase qui atténue la voie PI3K en déphosphorylant les phosphatidylinosides phosphorylés. P53 peut transactiver le promoteur de *PTEN* menant à une expression augmentée de PTEN. Ceci peut avoir un effet important puisqu'un défaut de PTEN peut compromettre l'apoptose déclenchée par p53 dans certains types cellulaires<sup>155</sup>.

### Régulation et activation de p53

L'amplitude ou la durée du signal d'activation de p53 peut affecter la réponse à p53. Un faible niveau d'expression de p53 occasionne l'arrêt du cycle cellulaire alors qu'un niveau plus élevé de p53 engendre l'apoptose<sup>156,157</sup>. De nombreux senseurs de stress convergent

vers p53 et résultent en la phosphorylation, l'acétylation, l'ubiquitination et la méthylation de certains de ses résidus<sup>158,159</sup>. La kinase ATM phosphoryle et active p53 en réponse aux dommages à l'ADN<sup>160</sup>. Des mutations d'*ATM* sont responsables de l'ataxie telangiectasie, une maladie associée à un risque accru de cancer du sein<sup>161</sup>. La protéine MDM2 est le régulateur négatif de p53 le plus étudié. Elle inhibe p53 en régulant sa localisation sub-nucléaire, son activité à activer la transcription et sa stabilité. Sous des conditions normales, MDM2 lie le domaine de transactivation de p53 et ubiquitine la protéine p53 la ciblant ainsi pour la dégradation<sup>162</sup>.

### 3.2.2. *PTEN* - Syndrome de Cowden

Parmi les maladies génétiques associées à un risque accru de cancer du sein, deux de ces syndromes sont caractérisés par des polypes hamartomateux. Le syndrome de Cowden (Cowden syndrome, CS) est une maladie autosomique dominante associée un risque accru de développer une multitude de cancers. Le syndrome de Peutz-Jeghers (Peutz-Jeghers syndrome; PJS), dans lequel une prédisposition au cancer du sein a aussi été identifiée, est une maladie autosomique dominante caractérisée par une pigmentation mucocutanée et des polypes hamartomateux gastrointestinaux.

#### Syndromes associés à *PTEN*

Des mutations germinales dans le gène suppresseur de tumeurs *PTEN*, localisé sur le chromosome 10q23<sup>163,164</sup>, ont été identifiées comme étant responsables du CS (OMIM #158350)<sup>165,166</sup>. Le CS est une maladie autosomique dominante caractérisée par de multiples hamartomes (peau, sein, intestin et cerveau) et un risque de cancer du sein, de la thyroïde et de l'endomètre<sup>166,167</sup>. Les caractéristiques plus fréquemment retrouvées chez les patients atteints du CS sont des lésions mucocutanées, une macrocéphalie et des polypes hamartomateux au niveau des intestins<sup>167</sup>. Plus de 90% des individus atteints du CS manifesteront un phénotype caractéristique de cette maladie avant l'âge de 20 ans<sup>165,167</sup>. Des études ont évalué l'incidence du CS à au moins un cas par 200 000<sup>170,171</sup> et 40-65% des cas du CS seraient d'origine familiale<sup>172,173</sup>.

Des mutations germinales de *PTEN* ont également été retrouvées chez des cas du syndrome de Bannayan-Riley-Ruvalcaba (BRRS; OMIM #153480)<sup>167,173</sup>, dans 20% des patients avec

le syndrome de Proteus, dans 50% des syndromes Proteus-like<sup>168</sup> et dans certains individus atteints de polyposes juvéniles<sup>169</sup>. Ces maladies phénotypiquement différentes partagent toutefois quelques caractéristiques cliniques sévères et une étiologie commune, ce qui a mené à leur classification dans un même groupe de maladies appelé PHTS (PTEN hamartoma-tumor syndrome)<sup>172,173</sup>. Parmi ceux-ci, seul le syndrome de Cowden est associé à un risque élevé de cancer du sein, bien que certaines patientes atteintes du BRRS aient développé un cancer du sein<sup>168</sup>. Des mutations de *PTEN* ont aussi été retrouvées chez 10 à 20% des patients présentant de l'autisme et une macrocéphalie<sup>174,175</sup>.

#### Cancer chez les patients atteints du syndrome de Cowden

Les cancers les plus documentés chez les patients atteints du CS sont le cancer du sein, le cancer de la thyroïde et de l'endomètre<sup>169,176</sup>. Le risque de cancer du sein chez les femmes atteintes du CS a été estimé à environ 25 à 50% et l'apparition du cancer du sein chez celles-ci se fait à un âge précoce (40-45 ans)<sup>177</sup>.

#### Mutations retrouvées dans *PTEN*

Des mutations germinales de *PTEN* ont été identifiées chez des individus atteints du cancer colorectal<sup>178</sup> et du cancer du sein<sup>179</sup>. Des mutations du promoteur ou du site de début de la transcription ont été observées chez des cas de CS pour lesquels des mutations n'avaient pas été identifiées dans la région codante de *PTEN*<sup>174,180</sup>. Des sites de liaison pour les activateurs de transcription p53 et EGR-1<sup>174</sup> ont notamment été identifiés dans le promoteur de *PTEN*. De grandes délétions génomiques de *PTEN* et de son promoteur ont aussi été rapportées chez des patients CS et BRRS<sup>174,181</sup>. La possibilité que certains variants d'épissage soient associés au risque de développer un cancer a aussi poussé les chercheurs à évaluer l'impact d'altérations d'épissage chez les patients CS, CS-like et BRRS<sup>182</sup>.

Des mutations somatiques de *PTEN* ont été identifiées chez différents cas de cancer du sein sporadiques<sup>179,183</sup>, chez des cas de cancer colorectal<sup>178,184</sup>, de la prostate<sup>185</sup>, de la vessie et de l'endomètre<sup>186</sup>. La fréquence de mutations de *PTEN* dans les cas de glioblastomes est de près de 30% alors que des mutations de *PTEN* sont identifiées dans 3,5% des tumeurs de cancer du sein<sup>187</sup>. Étant l'un des gènes les plus fréquemment mutés dans le cancer, *PTEN* devient la victime de la réparation déficiente de l'ADN causé par des mutations des gènes

impliqués dans la réparation de l'ADN. La perte de *BRCA1* mène à de larges mutations de *PTEN*<sup>188</sup>.

### Rôles de PTEN

Le gène *PTEN* encode une phosphatase à double spécificité pouvant déphosphoryler des substrats lipidiques et des protéines<sup>189,190</sup>. Des études d'expression de *PTEN* dans des lignées cellulaires cancéreuses ont montré que *PTEN* sauvage affecte l'arrêt du cycle cellulaire en G1 et/ou l'apoptose<sup>191-194</sup>.

#### *Activité phosphatase*

L'activité lipide phosphatase de *PTEN* a comme substrat le phosphoinositol 3,4,5-triphosphate (PIP3), un composant central de la voie de signalisation phosphatidylinositol 3-kinase (PI3K)/AKT<sup>195-197</sup>. *PTEN* exerce alors sa fonction de suppresseur de tumeurs en régulant la prolifération, la survie, la migration et l'invasion des cellules ainsi que l'angiogenèse. L'activité protéine phosphatase de *PTEN* permet quant à elle de déphosphoryler des substrats protéiques sur les résidus sérine, thréonine et tyrosine<sup>189</sup>. En déphosphorylant la protéine tyrosine kinase non-récepteur « Focal adhesion kinase » (FAK), *PTEN* régule la dynamique des interactions à la surface des cellules et inhibe la migration et l'invasion cellulaire<sup>198</sup>.

#### *Activité non-phosphatase*

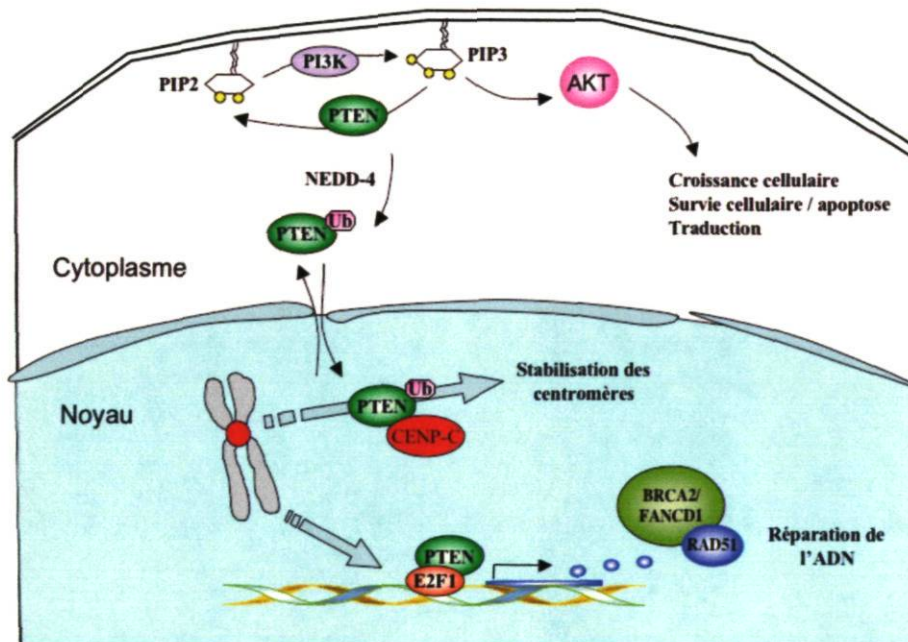
*PTEN* possède aussi des fonctions biologiques indépendantes de son activité phosphatase, telles la régulation de la migration cellulaire, de la transition du cycle cellulaire et de l'intégrité chromosomique<sup>198-201</sup>. *PTEN* pourrait supprimer la croissance tumorale d'une manière indépendante de son activité phosphatase, une version double mutante de *PTEN* pour son activité lipide et protéine phosphatase ayant démontré une inhibition de l'invasion cellulaire semblable à celle observée pour *PTEN* sauvage<sup>202</sup>.

#### *Rôle de PTEN dans la stabilité chromosomique*

Un lien entre la déficience en *PTEN* et l'instabilité a été démontré par le fait que les cellules embryonnaire déficientes en *PTEN* démontrent des défauts des points de contrôle de la réparation de l'ADN en réponse aux IR, ce qui résulte en une accumulation de bris chromosomiques et en des bris d'ADN double brins<sup>203</sup>. Des études récentes ont révélé

l'importance de la région C-terminale de PTEN dans la détermination de sa localisation sub-cellulaire<sup>204,205</sup> ainsi que dans le maintien de la stabilité génomique<sup>201</sup>. PTEN nucléaire peut utiliser deux différents mécanismes pour préserver l'intégrité chromosomique (Figure 2). PTEN affecte la réparation en régulant un joueur important dans la réparation des bris d'ADN, la protéine RAD51. De plus, PTEN s'associe physiquement avec un composant des centromères et la perte de cette association engendre des aberrations chromosomiques<sup>201</sup>.

Des études ont révélé que des changements d'expression globaux diffèrent selon la présence ou l'absence de PTEN<sup>206,207</sup>, ce qui suggère que PTEN pourrait être impliquée dans la régulation de la transcription. PTEN a d'ailleurs été rapportée comme diminuant l'activité promotrice de MDM2, un régulateur négatif de p53<sup>208</sup>. Ces données suggèrent que PTEN agit sur la chromatine et régule la transcription de gènes importants impliqués dans des processus cellulaires tel que le contrôle du cycle cellulaire et la réparation de l'ADN.



**Figure 2**

**Implication de PTEN dans le maintien de la stabilité chromosomique.** PTEN participe au maintien de la stabilité chromosomique en stabilisant les centromères et en favorisant la réparation de l'ADN. Adapté des références <sup>193,194,201,204,209</sup>.



### Expression et régulation de PTEN

La protéine PTEN est constitutivement exprimée, mais la régulation du niveau de cette protéine demeure un élément important dans sa fonction. L'activation inappropriée de la dégradation de PTEN prédispose à la tumorigenèse. Ceci peut survenir suite à une mutation de PTEN ou à une altération de la machinerie de dégradation, tel la surexpression ou suractivation de NEDD4-1, l'ubiquitine ligase ciblant PTEN<sup>205</sup>. À l'opposé, une surexpression de PTEN sauvage dans des cellules cancéreuses induit l'apoptose et arrête la progression cellulaire, la formation de colonies et la migration des cellules<sup>198,199,210,211</sup>.

La protéine PTEN interagit avec STK11 et est un substrat de STK11 *in vitro*. L'interaction de PTEN avec STK11 affecte la localisation cellulaire de STK11, possiblement en le retenant au cytoplasme<sup>212</sup>. Ceci vient mettre en évidence que ces deux suppresseurs de tumeurs sont impliqués dans une voie de signalisation commune<sup>212</sup>. Le suppresseur de tumeurs *TP53*, dont il a été mention précédemment, est relié à PTEN d'une manière complexe. En plus de réguler l'expression de PTEN en agissant à titre de facteur de transcription augmentant son expression<sup>155</sup>, p53 interagit physiquement avec PTEN, ce qui résulte en une stabilisation de p53, mais en une déstabilisation de PTEN<sup>180</sup>.

### *La localisation cellulaire de PTEN régule aussi sa fonction*

Le cytoplasme a été considéré comme le site primaire pour l'action suppresseur de tumeurs de PTEN et sa capacité à bloquer le sentier PI3-kinase comme l'élément central de la fonction de PTEN agissant sur la carcinogenèse<sup>213</sup>. Une proportion significative de PTEN est localisée au noyau dans plusieurs tissus humains et la perte d'expression de PTEN au noyau corrèle avec le développement du cancer<sup>214,215</sup>. La quantité de PTEN nucléaire varie en fonction du cycle cellulaire<sup>216</sup>. La protéine PTEN ne contient pas de NLS, mais la monoubiquitination de sa région C-terminale par l'ubiquitine ligase NEDD4-1 a été démontrée comme favorisant son importation au noyau<sup>204</sup>. Il est à noter que l'induction de la dégradation de PTEN par le protéasome et la translocation nucléaire de PTEN sont tous deux régulés par NEDD4-1. L'importation de PTEN au noyau pourrait aussi impliquer d'autres mécanismes, notamment par le transport passif par diffusion<sup>217</sup> et le transport actif grâce à des signaux ressemblant à des NLS<sup>218,219</sup>.

### Lien entre p53, BRCA1 et PTEN

Suite aux dommages à l'ADN, BRCA1 devient hyperphosphorylé par ATM<sup>75</sup> et CHEK2<sup>220</sup>. BRCA1 interagit physiquement et régule certains facteurs de transcription dont p53<sup>91</sup>. Une rétroaction existe entre BRCA1 et p53. BRCA1 a été démontrée comme augmentant la transcription par p53, alors que p53 activée réprime la transcription de *BRCA1*. Lorsque surexprimée dans des fibroblastes de souris, la protéine BRCA1 agit de manière synergique avec p53 pour causer un arrêt de la croissance irréversible<sup>221</sup>.

Une relation similaire existe entre PTEN et p53. PTEN a été rapportée comme étant une cible transcriptionnelle en aval de p53 dans l'apoptose<sup>222</sup>. D'autres études ont cependant suggéré que PTEN agirait en amont de p53 pour réguler son expression et son activité<sup>209,223</sup>. PTEN peut accroître la stabilité de p53 et son activité de liaison à l'ADN en s'y associant physiquement<sup>209</sup>.

Puisque PTEN est hautement exprimée, PTEN pourrait agir en temps normal pour le maintien général de la cellule. À l'opposé, p53 est principalement responsable de la réponse aux urgences tels les dommages à l'ADN. Conséquemment, p53 favorise les points de contrôle du cycle cellulaire et la mort cellulaire, dépendamment du type de dommage survenu. PTEN agirait principalement dans la mitose, un processus activé lors de la croissance normale d'une cellule<sup>187</sup>.

### 3.2.3. *STK11* - Syndrome de Peutz-Jeghers

Des mutations du gène *STK11/LKB1* sont responsables du syndrome de Peutz-Jeghers (Peutz-Jeghers syndrome, PJS; OMIM #175200) une maladie autosomique dominante<sup>224,225</sup> qui se caractérise par des taches de mélanine sur les lèvres, les muqueuses buccales et les doigts. Cette maladie rare, ayant une prévalence estimée de 1/25 000 à 1/280 000<sup>226</sup>, se manifeste par une pigmentation mucocutanée, des polypes intestinaux hamartomateux et une incidence accrue de plusieurs cancers. Le cancer gastrointestinal est le type de cancer le plus souvent diagnostiqué chez les patients PJS. Le risque de développer d'autres cancers, notamment le cancer du sein, de l'intestin, de l'estomac, de la thyroïde, des poumons, de l'endomètre, des ovaires, des testicules, du col de l'utérus et du pancréas est également élevé<sup>227,228</sup>. Les patients PJS ont un risque de développer un cancer environ 18

fois plus élevé que la population en général<sup>229</sup> et 37-93% des patients atteints du PJS auront développé un cancer à l'âge de 65 ans<sup>230</sup>. Une augmentation de cinq fois du risque de cancer du sein à développement précoce a été associée au PJS, ce qui pourrait représenter un risque de 29-50% de développer un cancer du sein avant l'âge de 65 ans<sup>230</sup>.

### Structure du gène *STK11*

Le locus responsable de la majorité des cas de PJS est retrouvé en 19p13.3<sup>224,225,231</sup>. Ce gène, nommé *STK11*, s'étend sur 23 kb et est composé de neuf exons codants et d'un exon non codant situé à l'extrémité 3' du gène<sup>224</sup>. La protéine de 433 acides aminés encodée par ce gène possède une activité ser/thr kinase et contient un signal de localisation nucléaire dans sa région N-terminale non catalytique (résidus 38-43). La partie centrale de la protéine (résidus 49-309) forme le domaine kinase alors que la région C-terminale de la protéine ne semble pas contenir d'autres domaines<sup>224,225</sup>, mais contient toutefois le site majeur d'auto-phosphorylation (Thr336)<sup>232</sup>.

### Mutations de *STK11*

Les mutations identifiées dans le PJS mènent habituellement à la perte de son activité kinase<sup>233</sup> et la perte de l'allèle sauvage n'est pas absolument nécessaire au développement d'hamartomes ou de la maladie<sup>234</sup>. Des mutations situées dans la partie C-terminale, une région importante pour le contrôle de la voie AMPK et de la polarité cellulaire, mais en dehors du domaine kinase, tel Phe354Leu<sup>235,236</sup>, peuvent aussi avoir un effet sur la fonction de cette protéine<sup>237,238</sup>.

Les premières études ont révélé que des mutations de *STK11* étaient présentes dans seulement 69% des cas de PJS. Avec le développement de nouvelles méthodes d'analyses, tel le MLPA (Multiplex Ligation-dependant Probe Amplification), et l'analyse du promoteur, un pourcentage de plus en plus grand des cas de PJS est expliqué par des mutations de *STK11*, variant notamment entre 66% et 94% selon les études<sup>239-241</sup>. Des mutations germinales de *STK11* ont également été identifiées dans certains cancers, notamment le cancer du côlon et du poumon<sup>235,242</sup>. La méthylation du promoteur et une perte d'expression de *STK11* ont été identifiées chez certains cas, mais ce mécanisme serait responsable d'une faible proportion des cas<sup>243,244</sup>.

### Rôles de STK11

STK11 est une kinase multi-tâches avec un grand potentiel d'orchestration de l'activité cellulaire. La fonction suppresseuse de tumeurs de STK11 joue un rôle dans la polarité des cellules, dans le métabolisme de l'énergie, l'apoptose, l'arrêt du cycle cellulaire et dans la prolifération cellulaire. Son activité kinase est nécessaire à cette fonction de suppression de tumeurs<sup>245-247</sup>.

La formation de polypes dans le PJS serait due à une prolifération accrue des cellules, à une apoptose diminuée ou à une combinaison de ces deux facteurs<sup>248</sup>. STK11 aide à établir la polarité cellulaire<sup>249,250</sup> et l'activité catalytique de STK11 est nécessaire afin de bloquer la division cellulaire. Une perte d'expression de *STK11* semble conditionner à la formation de polypes<sup>251</sup>.

Des études sur l'absence d'apoptose dans les polypes ont révélé que STK11 est impliqué dans l'induction de l'apoptose<sup>252-254</sup>. La protéine STK11 interagit avec p53 et possède la capacité de la phosphoryler<sup>253,255</sup>. STK11 peut donc potentiellement agir sur l'activité de ce membre central de la préservation de la stabilité génomique via l'apoptose. Il a été suggéré que STK11 pourrait agir à titre d'inducteur ou de suppresseur de l'apoptose dépendamment de la sévérité du stress énergétique et du type cellulaire<sup>256</sup>. La voie de signalisation STK11-AMPK joue un rôle dans la protection des cellules de l'apoptose en réponse aux agents augmentant le ratio AMP/ATP<sup>257</sup>. Toutefois, si les cellules ne sont pas capables de renverser ce ratio, elles iraient en apoptose<sup>256,257</sup>. STK11 ferait partie d'une voie de signalisation inhibant la kinase mTOR, un médiateur clé dans les signaux de survie occasionnés par PI3K/Akt.

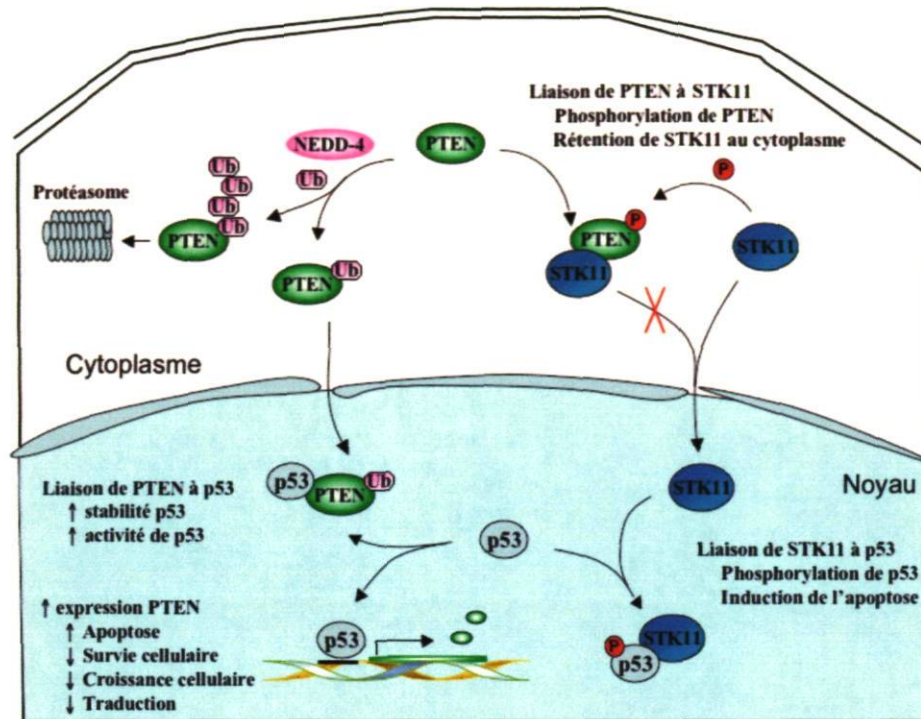
### Régulation et partenaires d'interaction de STK11

STK11 est exprimée de manière ubiquitaire dans les tissus adultes et fœtaux<sup>258-261</sup>. La signalisation par STK11 est régulée par sa phosphorylation et sa localisation cellulaire. STK11 est phosphorylée sur plusieurs résidus et sa phosphorylation joue un rôle important dans l'arrêt du cycle cellulaire, la suppression de tumeurs et la polarité cellulaire<sup>247</sup>.

*In vivo*, STK11 forme un complexe hétérotrimère avec les protéines STRAD (Ste20-related adaptor protein) et MO25 (mouse protein 25), qui jouent un rôle crucial dans la régulation

de la stabilité protéique, de l'activité kinase et de la localisation cellulaire de STK11<sup>255,262,263</sup>. STRAD lierait STK11 au cytoplasme et cette liaison serait essentielle à la phosphorylation d'AMPK par STK11<sup>262</sup>. La liaison de la protéine MO25 au complexe STK11/STRAD vient le stabiliser<sup>264</sup>, stimulant ainsi l'activité catalytique de STK11 d'environ 10 fois<sup>255</sup>. STK11 interagit aussi avec la protéine suppresseuse de tumeurs PTEN. Une surexpression de PTEN résulte en une rétention de STK11 au cytoplasme<sup>212</sup>. De plus la protéine PTEN est un substrat de STK11 *in vitro*. L'activité kinase de STK11 n'est pas nécessaire à sa liaison à PTEN, mais la pathogénicité de certaines mutations de STK11 pourrait être due non seulement à une perte d'activité kinase, mais elle serait due à une plus faible affinité pour PTEN<sup>212</sup>.

Il est possible de dégager jusqu'à maintenant un lien étroit entre les gènes *TP53*, *PTEN* et *STK11*, responsables de syndromes associés à un risque accru de cancer du sein (Figure 3). Alors que l'interaction de PTEN avec p53 vient accroître la stabilité et l'activité de liaison de l'ADN de p53, l'interaction de PTEN avec STK11 retient STK11 au cytoplasme. D'un autre côté, STK11 agit à la fois sur PTEN et p53 en les phosphorylant. L'interaction entre ces trois joueurs centraux de la régulation des mécanismes cellulaires et les syndromes causés par des mutations dans ces gènes vient mettre en évidence leur importance dans le maintien du bon fonctionnement de la machinerie cellulaire et dans le développement tumoral.



**Figure 3**

**Interrelations entre les protéines p53, PTEN et STK11.** L'interaction de PTEN avec p53 vient accroître la stabilité et l'activité de liaison de l'ADN de p53. L'interaction de PTEN avec STK11 retient STK11 au cytoplasme. STK11 agit à la fois sur PTEN et p53 en les phosphorylant. Adapté des références <sup>155,204,212,253,255</sup>.

### 3.2.4. ATM - Ataxie télangiectasie

L'ataxie télangiectasie (AT; OMIM # 208900) est une maladie héréditaire autosomique récessive causée par des mutations dans le gène *ATM* (ataxia telangiectasia mutated)<sup>265</sup>. L'AT est une maladie neurodégénérative progressive à pénétrance complète qui démontre une fréquence de 1/40 000 à 1/300 000<sup>266,267</sup> et pour laquelle la fréquence de porteurs hétérozygotes a été estimée à 1/200<sup>267</sup>. Au point de vue cellulaire, l'AT se caractérise par de l'instabilité génomique, des bris chromosomiques et une sensibilité aux IR. Les patients atteints de cette maladie démontrent fréquemment de l'ataxie cérébelleuse, des télangiectasies, une immunodéficience et une prédisposition à différents cancers. Environ un tiers des patients AT développent des cancers<sup>267</sup>. Des lymphomes chez les enfants, des leucémies et des cas de cancer du sein sont parmi les cancers les plus fréquemment rencontrés chez les patients atteints d'AT<sup>268</sup>. D'autres cancers pourraient toutefois être associés à l'AT, tels le cancer de l'estomac, du côlon, du pancréas, de la vessie et de l'ovaire<sup>269,270</sup>.

### Mutations d'ATM dans le cancer

Un risque accru de cancer a aussi été identifié chez les apparentés de patients atteints d'AT (porteurs hétérozygotes de mutations). Un excès de cancer du sein a été identifié chez les mères de patients atteints de l'AT<sup>271</sup> et l'augmentation du risque de cancer du sein chez les porteuses hétérozygotes de mutations dans *ATM* a été estimée à 5,1 fois<sup>269</sup>. D'autres études menées chez les familles de patients AT ont démontré un risque relatif moins élevé<sup>270,272-274</sup> alors qu'une analyse de ségrégation a révélé qu'un risque relatif de 2,37 serait associé à des mutations hétérozygotes d'*ATM*<sup>161</sup>. Il a aussi été proposé que certaines mutations d'*ATM* pourraient conférer un risque de cancer du sein du même ordre que des mutations des gènes *BRCA1* et *BRCA2*<sup>275</sup>. Dans un contexte de cancer du sein familial, certaines études ont identifié une plus grande proportion de porteurs de mutations d'*ATM* chez des cas familiaux<sup>276-278</sup> et d'autres ont révélé une association des mutations d'*ATM* avec le cancer du sein<sup>275,279</sup>.

### Structure d'ATM

Le gène *ATM* est un gène composé de 66 exons localisé en 11q22-23<sup>265</sup>. Deux transcrits différents ont été identifiés pour ce gène. L'isoforme la plus commune de ce gène encode une protéine de 3056 acides aminés<sup>280</sup>. La protéine ATM fait partie de la famille des kinases ressemblant aux PI3K<sup>281</sup>, contient un domaine FAT (FRAP, ATM, TRAPP). Une région C-terminale de 35 résidus hautement conservés, appelée le domaine FATC<sup>282</sup>, serait importante pour la régulation de l'activité kinase d'ATM et pour sa liaison avec des protéines régulatrices<sup>283</sup>. La région N-terminale d'ATM contient un domaine HEAT influençant ses interactions avec d'autres protéines<sup>284</sup> et une région essentielle pour sa liaison aux substrats<sup>285</sup>.

### Rôle de la protéine ATM

Le gène *ATM* encode une PI3K exprimée dans un large spectre de tissus. Cette protéine est localisée majoritairement au noyau sous forme de dimères ou de multimères inactifs qui sont activés suite aux dommages à l'ADN<sup>286</sup>. En réponse aux bris d'ADN double brins, ATM agit à titre de transmetteur de signal des dommages à l'ADN en se dissociant sous forme de monomères hautement actifs. Lors de ce processus, ATM s'auto-phosphoryle et est recrutée aux sites de dommage<sup>287</sup>. ATM phosphoryle alors une panoplie de substrats

impliqués dans la réparation de l'ADN, la réplication, dans les points de contrôle du cycle cellulaire, l'apoptose et la recombinaison<sup>286</sup>. On peut dénoter les protéines encodées par les gènes de susceptibilité au cancer du sein *BRCA1*, *TP53* et *CHEK2*<sup>288</sup> parmi les cibles d'ATM.

### 3.2.5. *CHEK2* - Syndrome s'apparentant au syndrome de Li-Fraumeni

La protéine *CHEK2* (CHK2 checkpoint homolog (S. pombe)) est une ser/thr kinase dépendante d'ATM qui module la voie des points de restriction des dommages à l'ADN<sup>289</sup>. *CHEK2* engendre la phosphorylation de plusieurs cibles impliquées dans la progression du cycle cellulaire et dans la réparation des dommages à l'ADN<sup>160,290</sup>. Des mutations germinales de *CHEK2* ont été initialement identifiées chez des familles atteintes d'un syndrome s'apparentant au syndrome de Li-Fraumeni (Li-Fraumeni like, LFS2; OMIM #609265), mais ne répondant pas aux critères classiques de classification du LFS et pour lesquelles des mutations de *TP53* n'ont pas été identifiées<sup>291</sup>.

#### Mutations de *CHEK2* dans le cancer

Une étude populationnelle a démontré que la mutation 1100delC de *CHEK2* est présente dans approximativement 1% de la population et dans 5% des cas de cancer du sein familiaux pour lesquels des mutations dans les gènes *BRCA1* et *BRCA2* n'ont pas été identifiées<sup>292</sup>. Des analyses de pénétrance menées chez ces familles ont par la suite permis d'estimer le risque relatif de cancer du sein chez les porteuses de cette mutation à 1,70 (95% IC = 1,32-2,20) alors qu'une autre étude menée par un consortium international a par la suite démontré que le rapport de cotes (odd ratio; OR) associé à la présence de cette mutation est de 2,34 (95% CI = 1,72-3,20)<sup>293</sup>. Une méta-analyse récemment parue a démontré que le variant 1100delC de *CHEK2* est un élément important de prédisposition au cancer du sein, augmentant le risque de cancer du sein de trois à cinq fois. Il a été estimé que la mutation 1100delC chez les femmes avec une histoire familiale de cancer du sein confère un risque cumulatif de 37%<sup>294</sup>. Cette mutation a été associée à un risque accru de cancer du sein chez l'homme et pourrait être responsable de 9% de ceux-ci<sup>292</sup>. Toutefois, cette association n'a pas été clairement répliquée depuis<sup>295-297</sup>. D'autres variants de *CHEK2* ont également été associés avec un risque accru de cancer du sein, notamment Ile57Thr<sup>298-300</sup>, et d'autres mutations spécifiques à certaines populations<sup>300-303</sup>.



Un défaut de fonction de CHEK2 pourrait aussi prédisposer les cellules de plusieurs organes au développement de tumeurs. La mutation 1100delC pourrait conférer un risque de cancer de la prostate<sup>292,298,301,302,304,305</sup> alors qu'une association du variant Ile157Thr avec un risque élevé de cancer colorectal, de la thyroïde et du rein a été démontrée<sup>298,306</sup>.

### Structure de CHEK2

Le gène *CHEK2* est localisé en 22q12.1 et encode pour un membre conservé du réseau cellulaire de réponse aux dommages à l'ADN qui protège l'intégrité du génome<sup>290</sup>. La protéine CHEK2 contient trois domaines distincts. Un domaine N-terminal SQ/TQ, un fork head-associated domain (FHA) et un domaine ser/thr kinase. Le domaine SQ/TQ est un domaine de régulation contenant les sites potentiels de phosphorylation par ATM<sup>290,307</sup>. Le domaine FHA est reconnu pour être impliqué dans la liaison d'autres protéines par la reconnaissance de résidus phosphothréonines. Ce domaine participe activement à la liaison protéine-phosphoprotéine de CHEK2 lors de son activation et du signalement des dommages à l'ADN<sup>308</sup>. La région C-terminale de la protéine CHEK2 contient son domaine catalytique.

Au niveau moléculaire, des défauts de CHEK2 mènent à sa perte d'expression ou de fonction. Des mutations dans la région régulatrice SQ/TQ préviennent son activation par ATM et des mutations dans son domaine FHA préviennent les interactions protéine-protéine avec ses partenaires tels BRCA1 et p53<sup>308,309</sup>. Des mutations de CHEK2 dans son domaine catalytique altèrent son activité kinase. D'autres défauts de CHEK2 mènent à une mauvaise localisation de CHEK2 au cytoplasme<sup>310</sup> ou à un épissage aberrant de l'ARN messager.

### Rôles de CHEK2

La protéine CHEK2 est une ser/thr kinase dépendante d'ATM qui module les points de restriction contrôlant la réponse aux dommages à l'ADN<sup>289</sup>. L'activation de CHEK2 en réponse aux dommages est un processus comportant plusieurs étapes qui est initié par une phosphorylation rapide de plusieurs sites SQ/TQ par ATM<sup>289,311</sup>. Suite à son homodimérisation et à sa phosphorylation intermoléculaire, CHEK2 est déplacée au noyau pour propager les signaux de dommages à l'ADN et cibler ses substrats<sup>312</sup>. Elle y bloque la

prolifération cellulaire et initie la réparation de l'ADN en réponse aux bris d'ADN double brins<sup>311,313</sup>. La kinase CHEK2 aide à maintenir l'intégrité génomique et prévient l'incorporation de mutations délétères<sup>307</sup> en engendrant la phosphorylation de plusieurs cibles impliquées dans la progression du cycle cellulaire et dans la réparation des dommages à l'ADN<sup>160,290</sup>.

La modulation de BRCA1 et BRCA2 par CHEK2 révèle son importance dans la voie de réparation des dommages à l'ADN. CHEK2 affecte la localisation nucléaire de BRCA1<sup>220</sup> et régule l'interaction de BRCA2 avec la recombinaison RAD51, ce qui joue un rôle important dans le recrutement subséquent de RAD51 aux sites de dommages à l'ADN<sup>128</sup>. D'autres cibles de CHEK2, tel p53 et son partenaire d'interaction MDM2, l'impliquent dans le contrôle de la mort cellulaire<sup>289,307,311</sup>. La kinase CHEK2 est exprimée dans une vaste majorité de tissus normaux<sup>314,315</sup>. La fonction de médiateur central de CHEK2 dans la réponse cellulaire aux dommages à l'ADN et son rôle dans la prévention de l'incorporation de mutations délétères font de CHEK2 un joueur central dans le maintien de l'intégrité génomique.

### 3.2.6. Les gènes de l'anémie de Fanconi

L'anémie de Fanconi (OMIM #227650) est une maladie génétique récessive caractérisée par une défaillance de la moëlle osseuse et entraînant la mort à un jeune âge. Les principales malformations congénitales associées à l'AF sont une aplasie radiale, une hyperpigmentation de la peau, un retard de croissance, de la microphthalmie et des malformations aux reins. Cette maladie se caractérise aussi par une prédisposition à la leucémie myéloïde aigüe (acute myeloid leukemia; AML) et l'apparition de certaines tumeurs solides<sup>316,317</sup>. L'incidence de cette maladie est de 1 cas par 200 000 à 400 000 naissances. La fréquence des porteurs hétérozygotes est estimée à 1 sur 250<sup>318</sup>.

Au point de vue génétique, l'AF se présente principalement sous la forme d'une maladie héréditaire autosomique récessive, bien qu'un des gènes responsables de cette maladie soit situé sur le chromosome X<sup>319</sup>. Treize groupes de complémentation ont été décrits (A, B, C, D1, D2, E, F, G, I, J, L, M et N), chacun d'eux étant causé par des mutations bialléliques du gène associé et nommés selon le groupe de complémentation auquel il est relié

(*FANCA*, *FANCB*, *FANCC*, etc). Le sous groupe le plus fréquent d'AF est le sous groupe A<sup>320</sup>. La sévérité du phénotype clinique pourrait être différente dépendamment du gène en cause. En effet, les patients du groupe de complémentation G sont plus à risque de développer l'AML que les patients des deux autres groupes de complémentation les plus fréquents, notamment les groupes A et C<sup>321</sup>. Les patients du groupe de complémentation *FANCD1/BRCA2* sont plus à risque de développer un phénotype de cancer que les patients des autres groupes de complémentation<sup>322</sup>.

#### Phénotype cellulaire de l'anémie de Fanconi

Les cellules des patients atteints de l'AF sont sensibles aux agents pontants l'ADN tels la mytomycine C, le diepoxybutane et le cisplatine<sup>318</sup>. Le nombre accru de bris chromosomiques causé par ces agents est à la base des tests diagnostiques pour cette maladie. L'AF démontre donc un phénotype cellulaire d'instabilité génomique semblable à celui des cellules déficientes en *BRCA1* ou *BRCA2*. Ceci a permis de suggérer que les protéines de l'AF seraient impliquées dans la reconnaissance ou la réparation des dommages à l'ADN. La première évidence fournissant un lien clair entre les protéines de l'AF et la réparation des dommages à l'ADN fut l'identification de *BRCA2* comme étant responsable du sous groupe D1<sup>133</sup>. D'autres évidences sont par la suite venues appuyer cette observation, notamment l'identification des hélicases *BRIP1/FANCI*<sup>323</sup> et *FANCM*<sup>324</sup> parmi les protéines de l'AF.

#### *Cancer chez les cas d'anémie de Fanconi*

Le risque de cancer en général chez les patients atteints d'AF est 44 fois plus élevé que ce qui serait attendu, estimé à 26 fois plus élevé qu'attendu pour les tumeurs solides et 868 fois plus élevé pour l'AML. De plus, un risque significativement élevé de cancer de l'oesophage, de la tête et du cou, tout comme celui de cancer du sein, de l'utérus, du foie et du cerveau a été rapporté<sup>325</sup>.

#### Cancer du sein et anémie de Fanconi

Outre *BRCA2/FANCD1*, une association entre un risque accru de cancer du sein et les gènes *BRIP1/FANCI* et *PALB2/FANCN* a récemment été identifiée<sup>137,138,326</sup>. La fonction de

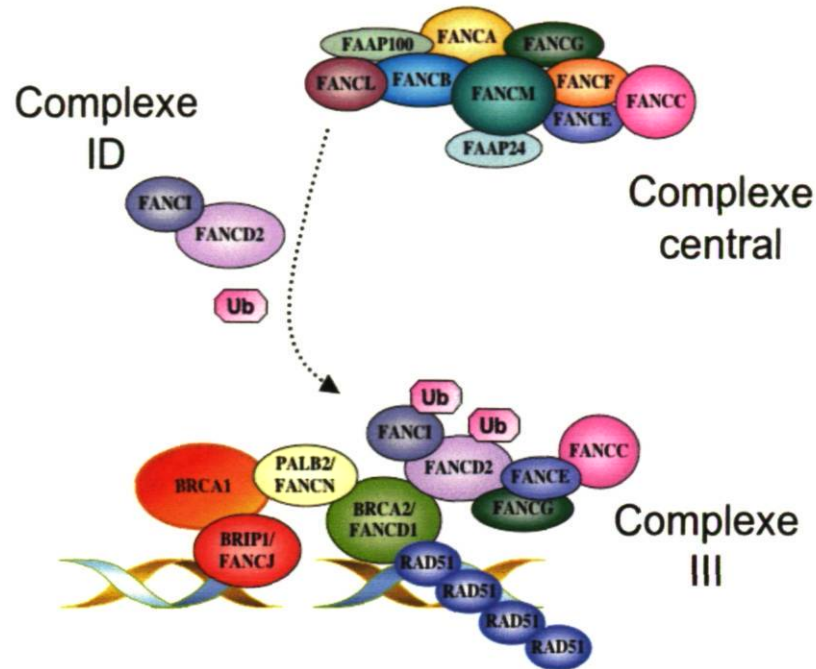
ces gènes, leur identification à titre de gènes de l'AF et l'association de ceux-ci avec un risque accru de cancer du sein seront discutés ultérieurement.

#### Implication des gènes de l'anémie de Fanconi dans les cancers sporadiques

Plusieurs études ont examiné la contribution des protéines de l'AF dans des cancers sporadiques. Une perte d'expression ou un niveau réduit des protéines FANC ont été observés dans des lignées cellulaires d'AML<sup>327</sup>. Des variants de FANCD2 ont été associés au développement de cas sporadiques de cancer du sein<sup>328</sup>. De plus, des mutations des gènes de l'AF ainsi qu'une diminution de leur expression ont été démontrées dans différents cancers sporadiques et chez des individus non-atteints de l'AF<sup>329,330</sup>.

#### Voie AF/BRCA (FA/BRCA pathway)

L'implication des gènes de l'AF dans une même voie métabolique de réponse aux dommages à l'ADN et la participation des protéines BRCA1 et BRCA2 ont menés à l'émergence de la voie AF/BRCA (FA/BRCA pathway). La découverte d'une interaction entre BRCA1 et FANCD2<sup>331</sup> et l'identification de *BRCA2* comme étant un gène de l'AF<sup>133</sup> en sont deux événements majeurs. Des interactions entre BRCA1-FANCA, BRCA1-BRIP1/FANCJ, BRCA2-FANCD2 et BRCA2-FANCG ont également été rapportées<sup>97,332-334</sup>. De plus, la protéine PALB2/FANCN lie à la fois BRCA1 et BRCA2<sup>134-136,335,336</sup>. Le FA/BRCA pathway peut se résumer en une séquence de trois étapes consécutives pendant lesquelles certaines des protéines FANC s'assemblent pour former différents complexes menant ultimement à la réparation des bris d'ADN en collaboration avec BRCA1 et BRCA2 (Figure 4).



**Figure 4**

**FA/BRCA pathway.** Représentation du mécanisme général et des différents complexes formés dans le FA/BRCA pathway. Adapté des références <sup>134,335-340</sup>.

Les protéines FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL et FANCM forment un complexe avec les protéines FAAP100 et FAAP24 (FANCA associated polypeptide). Ce complexe protéique forme un premier groupe de protéines de l'AF et est appelé le **complexe central**. Chacune des protéines du complexe central est nécessaire à la monoubiquitination de FANCD2 et de FANCI<sup>337-339</sup>. La protéine FANCL est la sous-unité ubiquitine ligase du complexe central.

Les protéines de l'AF faisant partie du deuxième groupe, les protéines FANCD2 et FANCI, forment un complexe appelé le **complexe ID**<sup>339</sup>. Ces deux protéines sont monoubiquitinées par le complexe central et leur monoubiquitination est interdépendante<sup>338,339</sup>. Suite à leur monoubiquitination, les protéines du complexe ID sont ciblées aux sites de dommages de la chromatine. À cet endroit, la protéine FANCD2 s'associe à un complexe comprenant FANCD1/BRCA2 et FANCE<sup>341</sup>. Les protéines FANCD2 et FANCI monoubiquitinées se retrouvent aux foci nucléaires reconnus pour être des sites de réparation de l'ADN. Cette relocalisation dépend à la fois du complexe central, de ATR, de BRCA1 et de l'histone H2AX<sup>331,342</sup>.

Le troisième groupe des protéines de l'AF (**groupe III**) est constitué des protéines FANCD1/BRCA2, FANCI/BRIP1 et FANCN/PALB2. Ces trois protéines agissent en aval de la monoubiquitination du complexe ID ou via une voie différente puisque les cellules déficientes en ces protéines possèdent un niveau de FANCD2 monoubiquitiné normal en réponse aux dommages à l'ADN<sup>343</sup>. Les gènes encodant les protéines du groupe III de l'AF sont étroitement liés au développement du cancer du sein, des mutations monoalléliques prédisposant les femmes porteuses au cancer du sein.

L'interaction des protéines de l'AF avec d'autres protéines impliquées dans la réparation des bris d'ADN démontre leur rôle important dans ce processus. Plusieurs protéines, dont les kinases ATR et CHEK1, activent les protéines de la voie AF/BRCA en facilitant la phosphorylation de certains membres du complexe central, du complexe ID, de BRCA1, BRCA2 et de PALB2<sup>324,344-347</sup>. Les protéines du complexe central agissent en collaboration avec l'hélicase BLM, la Topoisomérase III et RPA, qui ont comme rôle de stabiliser la fourche de réplication bloquée et de permettre la réinitiation de la réplication d'une manière dépendante de la HR<sup>348</sup>.

Plusieurs protéines impliquées dans la réparation des bris d'ADN interagissent avec les protéines du groupe III. Les protéines BRCA2 et PALB2 font partie d'un complexe contenant la protéine RAD51. La protéine PALB2 a été identifiée comme une composante essentielle de ce complexe et l'inactivation de PALB2 résulte en un défaut dans la réparation de l'ADN par HR<sup>134</sup>. La protéine BRIP1/FANCI, quant à elle, fait partie d'un complexe contenant les protéines BRCA1 et TOPBP1, une protéine liant la topoisomérase II, et d'un autre complexe composé des protéines MLH1 et PMS2, impliqué dans la ségrégation post-méiotique. L'interaction de FANCI avec les protéines MLH1 et PMS2 est d'ailleurs essentielle pour sa fonction dans la voie de l'AF<sup>349</sup>.

En plus de sa participation à la voie AF, il a été suggéré que la protéine FANCD2 pourrait être impliquée dans une autre voie de signalisation activée par ATM en réponse aux IR causant des bris d'ADN double brins<sup>350</sup>. En effet, ATM phosphoryle FANCD2 ce qui active le point de contrôle en phase S. D'autres fonctions ont également été associées aux protéines de l'AF, tel la prévention des dommages oxydatifs par des espèces réactives de

l'oxygène<sup>351</sup>, la régulation de l'apoptose, la réponse aux cytokines et aux facteurs de croissance<sup>352</sup>.

### 3.2.6.1. *BRIP1/FANCI*

Le gène *BRIP1/FANCI*, situé en 17q22-24, encode une protéine de 1249 aa possédant un domaine hélicase dans sa partie N-terminale. La région C-terminale de cette protéine contient son domaine d'interaction avec les domaines BRCT de *BRCA1*<sup>97</sup>.

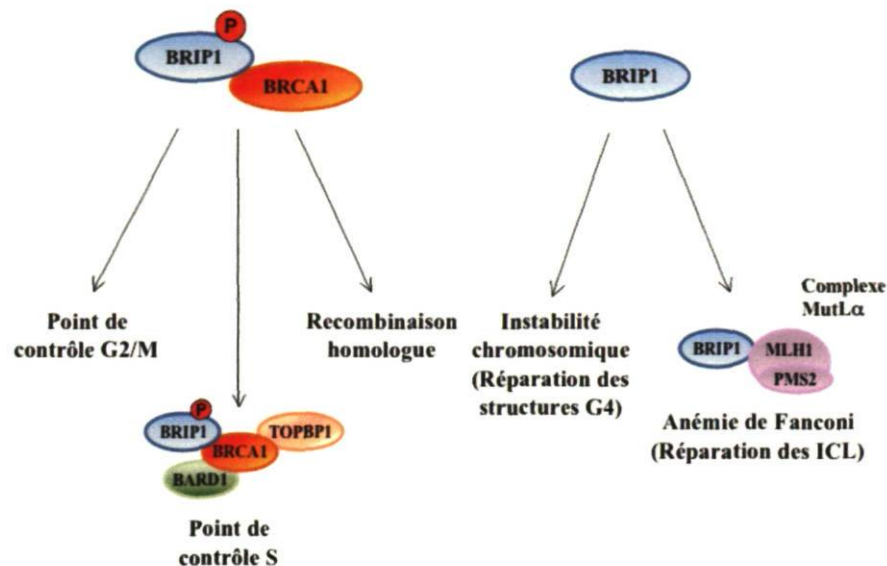
#### Mutations de *BRIP1* dans le cancer du sein

Lors de l'identification de l'interaction de la protéine *BRIP1* avec *BRCA1*, deux mutations faux-sens causant une perte d'activité hélicase de la protéine<sup>353</sup> ont été identifiées chez des cas familiaux<sup>97</sup>. L'analyse ce gène dans des cas de cancer du sein sporadiques a suggéré une association du variant Ser919Pro avec le développement du cancer du sein avant 50 ans<sup>354</sup>. Différentes études ont également été menées chez des familles exemptes de mutations dans les gènes *BRCA1* et *BRCA2*<sup>355-359</sup>. Le point culminant dans l'association de *BRIP1* avec la susceptibilité au cancer du sein est venu suite à l'identification de son implication dans l'AF de sous-type J<sup>323</sup>. Des mutations tronquantes de *BRIP1* ont été identifiées chez des cas de cancer du sein provenant de familles à risque élevé dépourvues de mutations dans les gènes *BRCA1* et *BRCA2*, permettant ainsi d'associer clairement des mutations de *BRIP1* avec un risque accru de cancer du sein, alors que le risque relatif conféré par des mutations de *BRIP1* a été estimé à 2,0 (95% CI = 1,2-3,2, p=0,012)<sup>326</sup>.

#### Rôles attribués à *BRIP1*

La protéine *BRIP1* est impliquée dans les points de contrôle du cycle cellulaire et dans la réparation des bris d'ADN où elle possède des fonctions dépendantes et indépendantes de *BRCA1*<sup>97,360</sup> (Figure 5). La protéine *BRIP1* interagit avec la protéine *BRCA1*<sup>97</sup> et est une hélicase qui déroule les duplex d'ADN d'une manière dépendante à l'ATP<sup>353,361</sup>. L'interaction de *BRIP1* avec *BRCA1* est dépendante de l'état de phosphorylation de *BRIP1* et cette interaction est requise pour le point de contrôle G2/M du cycle cellulaire<sup>353</sup>. L'assemblage d'un complexe contenant *BRCA1*, *BRIP1* et *BARD1* permet l'interaction de ce complexe avec *TOPBP1*, un facteur jouant un rôle important dans le point de contrôle en

phase S<sup>362</sup>. De plus, l'activité hélicase de BRIP1 est nécessaire pour la progression à travers la phase S du cycle cellulaire<sup>363</sup>.



**Figure 5**

**BRIP1 joue des rôles dépendants et indépendants de BRCA1.** La protéine BRIP1, lorsque liée à BRCA1, serait impliquée dans les points de contrôle du cycle cellulaire et dans la recombinaison homologue. Elle posséderait aussi des rôles indépendants de BRCA1 dans la réparation des structures G4 et des ICL. Adapté des références 97,349,353,363,364.

L'interaction de BRIP1 avec BRCA1<sup>97</sup> a permis d'impliquer cette protéine dans la réparation des bris d'ADN double brins par HR<sup>97</sup>. En réponse aux dommages à l'ADN ou au stress causé par la réplication, BRIP1 colocalise aux foci nucléaires avec la protéine RPA qui lie l'ADN simple brin et est impliquée dans le signalement des fourches de réplication arrêtées. La colocalisation de ces deux protéines est dépendante de BRCA1<sup>365</sup>. RPA régulerait l'activité hélicase de BRIP1, qui est importante pour le maintien de la stabilité génomique<sup>365</sup>. Par son activité dépendante de l'ATP, BRIP1 peut déstabiliser les duplex protéine-ADN. Ainsi, il inhibe l'échange des brins effectué par RAD51, ce qui serait important pour le contrôle de la réparation de l'ADN par la HR<sup>366</sup>.

BRIP1 fonctionne dans la voie de l'AF indépendamment de BRCA1 et son action se situe en aval de l'activation de FANCD2<sup>367</sup>. Les cellules déficientes en BRIP1 sont hypersensibles aux agents causant des ICL<sup>367,368</sup>. La correction des ICL dans les cellules requiert l'interaction de BRIP1 avec le complexe de réparation des mésappariements MutLα<sup>349</sup> composé des protéines MLH1 et PMS2.



Les défauts causés par des mutations inactivatrices de BRIP1 implique aussi d'autres types de dommages structuraux de l'ADN nuisant à la réplication. En effet, BRIP1 aurait un rôle important dans le déroulement des structures G4 de l'ADN qui pourraient affecter la réplication et causer de l'instabilité chromosomique<sup>364</sup>. Ces sites G4 ont été prouvés comme étant mutagéniques puisqu'ils peuvent adopter une conformation particulière en quadruplex lors de la réplication pouvant ainsi bloquer la réplication et induire des réarrangements au niveau des sites G4 du génome<sup>369</sup>.

### 3.2.6.2. *PALB2/FANCN*

Le gène *PALB2* (partner and localizer of BRCA2), localisé en 16p12.1, est composé de 13 exons et encode une protéine interagissant avec BRCA2. La protéine encodée par ce gène est composée de 1186 aa et contient un segment homologue à la Prefoldin, une protéine interagissant avec l'actine et la tubuline. La région C-terminale de *PALB2* contient les domaines WD-40 qui sont requis pour son interaction avec BRCA2. Le domaine de liaison entre *PALB2* et BRCA2 couvre la région C-terminale du domaine de transactivation de BRCA2. Il a été récemment identifié que la région N-terminale de *PALB2* (résidus 9-42) lie BRCA1<sup>335,336</sup>.

Des mutations bi-alléliques tronquantes de *PALB2* ont été retrouvées chez des familles atteintes du sous-type N de l'AF<sup>135,136</sup>. Les cellules déficientes en *PALB2/FANCN* démontrent une monoubiquitination normale de FANCD2 et les fibroblastes déficients en *PALB2* ne présentent pas de formation de foci nucléaires de RAD51 suite aux IR. Ceci est similaire aux cellules déficientes en *BRCA2* et diffère des autres sous-types de cellules de l'AF<sup>343</sup>. Les phénotypes associés à des mutations de *BRCA2* et *PALB2* sont semblables entre eux, mais différent du phénotype associé aux autres gènes de l'AF. En particulier, le types D1 et N de l'AF sont associés avec un risque élevé de tumeurs solides chez l'enfant tels les tumeurs de Wilms et des médulloblastomes, ce qui survient très rarement dans les autres sous-types de l'AF<sup>135,370,371</sup>.

### Mutations de *PALB2* dans le cancer

Le gène *PALB2* est devenu rapidement un candidat important en tant que gène de susceptibilité au cancer du sein. Une étude menée chez des cas de cancer du sein provenant

de familles dépourvues de mutations dans les gènes *BRCA1* et *BRCA2* a permis de mettre en évidence l'implication de ce gène dans la susceptibilité au cancer du sein. Des analyses de ségrégation menées chez ces familles ont permis d'estimer que le risque relatif associé à des mutations de *PALB2* est de 2,3<sup>138</sup>. Une étude finlandaise a ensuite estimé l'augmentation du risque de cancer du sein à 4 fois<sup>137</sup> alors qu'une étude subséquente a proposé que certaines mutations seraient associées à un risque plus élevé<sup>372</sup>. L'analyse de *PALB2* chez 50 Canadiennes françaises provenant de familles à risque élevé ou atteintes du cancer du sein à un âge précoce a révélé la présence d'une mutation tronquante (c.2323C>T, Gln775X) qui s'est avérée être une mutation fondatrice dans la population canadienne-française<sup>373</sup>. Des mutations fondatrices de *PALB2* sont retrouvées chez les Finlandais et représenteraient 1% de tous les cas de cancer du sein, alors que dans la population canadienne-française, la mutation fondatrice de *PALB2* serait responsable de près de 0,5% des cas de cancer du sein à développement précoce<sup>137,373</sup>. Dans la population du Royaume-Uni, des mutations de *PALB2* seraient responsables de 1% des cas de cancer du sein familiaux. Des mutations de *PALB2* pourraient également conférer un risque élevé de cancer du sein chez l'homme comme le font des mutations de *BRCA2*<sup>138</sup>. L'analyse de la région promotrice de *PALB2* chez des individus porteurs de mutations de *PALB2* a révélé une hyperméthylation et une diminution d'expression d'environ 28 fois dans les tumeurs provenant de ces individus en comparaison aux tissus normaux<sup>374</sup>.

Les familles ayant initialement permis d'identifier *PALB2* à titre de gène de susceptibilité au cancer du sein présentaient également d'autres types de cancers, tels des cancers colorectaux, des cancers de l'estomac, de l'endomètre, du pancréas et des leucémies<sup>137</sup>. L'analyse de *PALB2* dans des cas familiaux de cancer de la prostate a été effectuée sans toutefois révéler d'association entre une mutation fondatrice finlandaise et le cancer de la prostate<sup>375</sup>. L'analyse d'une cohorte de cas familiaux de cancers pancréatiques a toutefois permis d'établir *PALB2* à titre de gène de susceptibilité au cancer du pancréas<sup>376</sup>.

### Rôle de PALB2

*PALB2*, en interagissant avec *BRCA2* et *BRCA1*, vient se positionner en aval de la monoubiquitination du complexe ID dans la voie de l'AF. Tout comme *BRCA2*, *PALB2* participe à la réponse aux dommages à l'ADN en étant recrutée aux bris d'ADN double

brins et en favorisant la bonne localisation et la stabilité de BRCA2<sup>134</sup>. L'association de BRCA2 avec PALB2 semble être essentielle pour sa fonction dans la HR et la réparation des bris d'ADN double brins et pour l'attachement de BRCA2 aux structures nucléaires<sup>134</sup>. Comme BRCA2, PALB2 colocalise avec l'histone H2AX phosphorylée suite aux radiations ionisantes<sup>134</sup>. De multiples mutations faux-sens de BRCA2 identifiées chez des patients atteints du cancer du sein viennent altérer l'interaction entre PALB2 et BRCA2 affectant ainsi la fonction de BRCA2 dans la HR et la réparation des bris d'ADN.

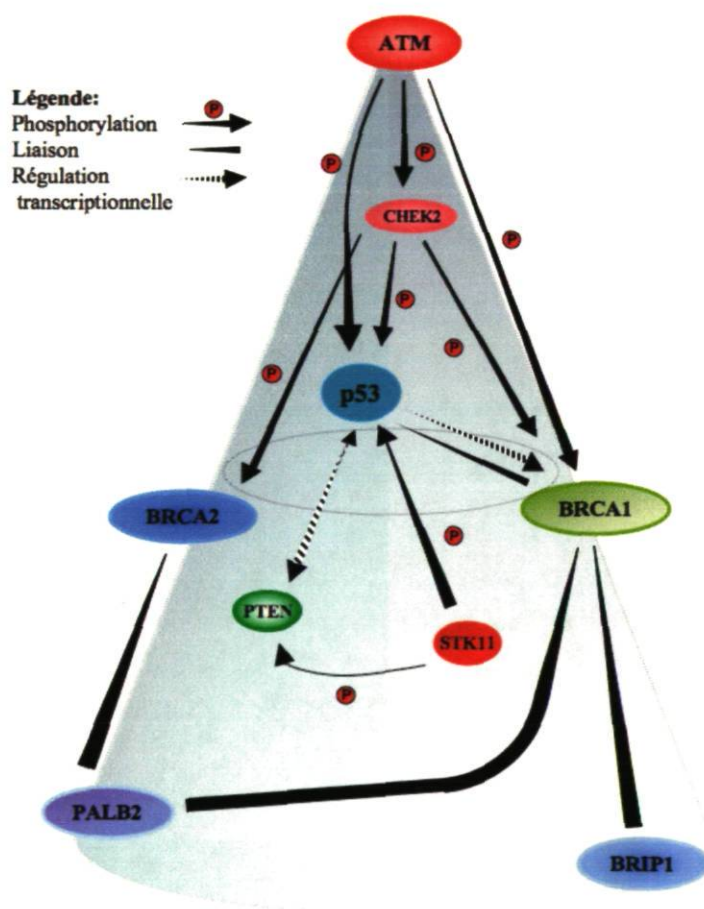
La protéine PALB2 lie également BRCA1 et servirait de protéine d'échaffaudage entre BRCA1 et BRCA2 dans la formation du complexe BRCA nécessaire à la réparation par HR<sup>335,336</sup>. Tout comme pour l'interaction PALB2-BRCA2, un défaut d'interaction entre PALB2 et BRCA1 résulterait en des défauts dans la réparation par HR. L'interaction directe de BRCA1 et PALB2 pourrait réguler la réparation par recombinaison en modulant le recrutement de BRCA2 et RAD51 aux sites de bris<sup>335</sup>. En accord avec ce modèle, il a été identifié que PALB2 existe sous la forme d'homo-oligomères et que l'oligomérisation de PALB2 aux sites de bris est essentielle à la formation de foci. Ainsi, l'association de PALB2 à la chromatine et son oligomérisation initieraient la réparation par recombinaison en coordonnant l'association de BRCA2 et RAD51 avec la chromatine et l'arrivée du complexe de réparation aux sites de dommages à l'ADN<sup>377</sup>.

### **3.3. Gènes de susceptibilité au cancer du sein : participation à une voie commune**

#### *BRCA1 et BRCA2 sont activés par phosphorylation*

Les gènes de susceptibilité au cancer du sein présentés jusqu'ici sont interreliés (Figure 6) et sont impliqués dans le contrôle du cycle cellulaire, de l'apoptose et de la réparation des bris d'ADN. En effet, ATM est majoritairement localisée au noyau sous forme de dimères inactifs qui sont activés suite aux dommages à l'ADN<sup>286</sup>. En réponse aux bris d'ADN double brins, ATM agit à titre de transmetteur de signal des dommages à l'ADN en étant recrutée aux sites de dommage<sup>287</sup>. ATM phosphoryle une panoplie de substrats impliqués dans la réparation de l'ADN, la réplication, les points de contrôle du cycle cellulaire, l'apoptose et la recombinaison<sup>286</sup>. On peut dénoter les protéines BRCA1, p53 et CHEK2<sup>288</sup> parmi les cibles d'ATM.

La protéine CHEK2 est une ser/thr kinase dépendante d'ATM qui module les points de restriction contrôlant la réponse aux dommages à l'ADN<sup>289</sup>. CHEK2 phosphoryle BRCA1, affectant ainsi la localisation nucléaire de BRCA1<sup>220</sup>. CHEK2 phosphoryle également le domaine C-terminal de BRCA2, ce qui régule l'interaction de BRCA2 avec la recombinaise RAD51. Cette phosphorylation de BRCA2 par CHEK2 joue un rôle important dans le recrutement subséquent de RAD51 aux sites de dommages à l'ADN<sup>128</sup>. CHEK2 pourrait également être impliquée dans la régulation de la localisation nucléaire de BRCA2. D'autres cibles de CHEK2, tel p53 et son partenaire d'interaction MDM2, l'impliquent dans le contrôle de la mort cellulaire<sup>289,307,311</sup>.



**Figure 6**  
**Interrelations de BRCA1 et BRCA2 avec les protéines encodées par les autres gènes de susceptibilité au cancer du sein.** Les kinases apparaissent en rouge alors que les protéines de l'anémie de Fanconi portent des teintes de bleu.

### *Rôle central de BRCA1 et BRCA2*

Suite aux dommages à l'ADN, BRCA1 devient hyperphosphorylé par ATM<sup>75</sup> et CHEK2<sup>220</sup>. L'interaction de BRCA1 avec p53 a pour effet d'augmenter la transcription par p53, alors que p53 réprime la transcription de *BRCA1*<sup>91</sup>. Suite à une surexpression de *BRCA1* endogène, BRCA1 agit de manière synergique avec p53 pour causer un arrêt de la croissance irréversible<sup>221</sup>. Il a été démontré que p53 est impliquée dans l'induction par BRCA1 du point de restriction G1/S du cycle cellulaire<sup>103</sup>.

BRCA1 contribue également à la réponse aux dommages à l'ADN par son interaction avec des protéines altérant la structure de l'ADN et de la chromatine<sup>94</sup>, en particulier avec les hélicases BLM, une protéine participant à la réplication de l'ADN<sup>64</sup> et avec l'hélicase BRIP1<sup>97</sup>. Le rôle spécifique de BRCA1 dans la réparation des bris d'ADN double brins peut être défini par son interaction avec les protéines BRIP1/FANCD1 et BARD1. Ces deux protéines ont un impact sur l'efficacité de la réparation de l'ADN.

Les protéines BRCA2 et RAD51 interagissent entre elles et sont impliquées dans la HR<sup>104,120,122-124</sup>. BRCA2 est nécessaire au recrutement et à la fonction de RAD51 aux sites de bris de l'ADN<sup>120,126</sup>. La liaison de BRCA2 à RAD51 pourrait faciliter l'arrivée de RAD51 ou l'organisation des filaments de RAD51<sup>378,379</sup>. Récemment, une nouvelle protéine a été identifiée sur la base de son interaction avec BRCA2. Cette protéine, appelée PALB2, colocalise avec BRCA2 aux foci nucléaires et cause une localisation et une stabilisation de BRCA2 aux structures nucléaires, rendant ainsi possible la réparation par recombinaison et la fonction de BRCA2 dans les points de contrôles du cycle cellulaire<sup>134</sup>.

### *Interrelations entre les autres gènes de susceptibilité*

Tel que mentionné précédemment, la kinase ATM phosphoryle et active p53 en réponse aux dommages à l'ADN<sup>160</sup>. Lorsqu'activée, p53 peut induire des gènes qui court-circuitent les voies de signalisation anti-apoptotiques. L'exemple le plus évident de l'action de p53 est la régulation de PTEN, un régulateur négatif de la voie PI3K kinase aussi responsable d'un syndrome prédisposant au cancer du sein (voir section 3.2.2). En effet, p53 peut transactiver le promoteur de *PTEN* menant à une expression augmentée de PTEN. Ceci peut avoir un effet important puisqu'un défaut de PTEN peut compromettre l'apoptose

déclenchée par p53 dans certains types cellulaires<sup>155</sup>. *PTEN* a été rapporté comme étant une cible transcriptionnelle en aval de p53 dans l'apoptose<sup>222</sup>. D'autres études ont cependant suggéré que *PTEN* agirait en amont de p53 pour réguler son expression et son activité<sup>209,223</sup>. *PTEN* peut accroître la stabilité de p53 et son activité de liaison à l'ADN en s'y associant physiquement<sup>209</sup>.

*PTEN* affecte également la réparation de l'ADN en régulant un joueur important dans la réparation des bris d'ADN, la protéine RAD51. De plus, *PTEN* interagit avec STK11 et affecte la localisation cellulaire de STK11, possiblement en la retenant au cytoplasme<sup>212</sup>. STK11 lie et phosphoryle p53, ce qui résulte en une induction de l'apoptose<sup>253,255</sup>.

Le troisième groupe des protéines de l'AF (groupe III) est constitué des protéines FANCD1/BRCA2, FANCI/BRIP1 et FANCD3/PALB2. Plusieurs protéines impliquées dans la réparation des bris d'ADN interagissent avec les protéines de ce groupe. La protéine BRIP1 interagit avec la protéine BRCA1<sup>97</sup>. L'interaction de BRIP1 avec BRCA1 est dépendante de l'état de phosphorylation de BRIP1. L'interaction de BRIP1 avec BRCA1<sup>97</sup> a permis d'impliquer cette protéine dans la réparation des bris d'ADN double brins par HR<sup>97</sup>. Cette interaction est requise pour le point de contrôle G2/M du cycle cellulaire induit par les dommages à l'ADN<sup>353</sup>. Les protéines BRCA2 et PALB2 font partie d'un complexe contenant la protéine RAD51. La protéine PALB2 a été identifiée comme une composante essentielle de ce complexe et l'inactivation de PALB2 résulte en un défaut dans la réparation de l'ADN par HR<sup>134</sup>. La protéine PALB2 lie également BRCA1 et servirait de protéine d'échaffaudage entre BRCA1 et BRCA2 dans la formation du complexe BRCA nécessaire à la réparation par HR<sup>335,336</sup>. Tout comme pour l'interaction PALB2-BRCA2, un défaut d'interaction entre PALB2 et BRCA1 résulterait en des défauts dans la réparation par HR.

### **3.4. Classification des gènes de susceptibilité**

Notre connaissance des gènes de susceptibilité au cancer du sein nous a permis de les classer en trois catégories selon leur risque associé. Les deux premières catégories renferment les gènes de forte pénétrance et de pénétrance modérée, pour lesquels des mutations sont rares (Tableau 2, Figure 7). Jusqu'à maintenant, les gènes présentés dans

cette thèse font partie de ces deux catégories. Les deux gènes majeurs de susceptibilité au cancer du sein, *BRCA1* et *BRCA2*, ont été classés en tant que gènes de forte pénétrance. Cette catégorie renferme aussi les gènes *TP53*, *PTEN* et *STK11* qui sont responsables de syndromes associés à une forte susceptibilité au cancer du sein. Les autres gènes de susceptibilités présentés, notamment *ATM*, *CHEK2*, *BRIP1* et *PALB2*, confèrent un risque modéré de développer un cancer du sein qui est d'environ 2 à 4 fois. La troisième catégorie est constituée de gènes et d'allèles à faible pénétrance, pour lesquels les variations associées au cancer du sein sont des variations communes retrouvées à une fréquence élevée. Jusqu'à maintenant, les gènes de faible pénétrance au cancer du sein ont été identifiés grâce à des études d'association dont la plupart ont été menées à l'échelle du génome entier. Le risque associé à des variations dans ces gènes est de l'ordre de 1,5. Puisque l'identification de gènes de faible pénétrance nécessite des études d'association comportant des milliers de sujets et que de telles études ne peuvent être réalisées que dans le cadre de grands consortiums internationaux, les gènes de susceptibilité à faible pénétrance n'ont pas été la cible de mon projet de recherche. Ainsi, une attention moindre sera portée aux gènes de cette catégorie lors du présent ouvrage et une présentation et description succincte de ceux-ci sera faite.

Tableau 2

Risque associé à des mutations dans les gènes de susceptibilité au cancer du sein à forte pénétrance et à pénétrance modérée.

	Gènes	Position	Syndromes associés	État (dominant /récessif)	Fréquence des porteurs hétérozygotes	Risque de cancer du sein	Références
Forte pénétrance	<i>BRCA1</i>	17q	Cancer du sein et de l'ovaire	Dominant	1/1000	65%	44,52,57, 58,59,60
	<i>BRCA2</i>	13q	Cancer du sein et de la prostate	Dominant	1/1000	45%	44,52,57, 58,59,60
	<i>TP53</i>	17p	Li-Fraumeni	Dominant	1/5000	50-60%	139,380*
	<i>PTEN</i>	10q	Cowden	Dominant	1/200 000	25-50%	170,171,177
	<i>STK11</i>	19p	Peutz-Jeghers	Dominant	1/25 000 à 1/280 000	29-50%	226,230
Pénétrance modérée	<i>ATM</i>	11q	Ataxie télangiectasie	Récessif	1/200	RR = 2,37	161,267
	<i>CHEK2</i>	22q	Li-Fraumeni « like »	Dominant	1/100	OR = 2,34	292,293
	<i>BRIP1/ FANCI</i>	17q	Anémie de Fanconi (sous-type J)	Récessif	ND	RR = 2,0	326
	<i>PALB2/ FANCN</i>	16p	Anémie de Fanconi (sous-type N)	Récessif	ND	RR = 2,3	138

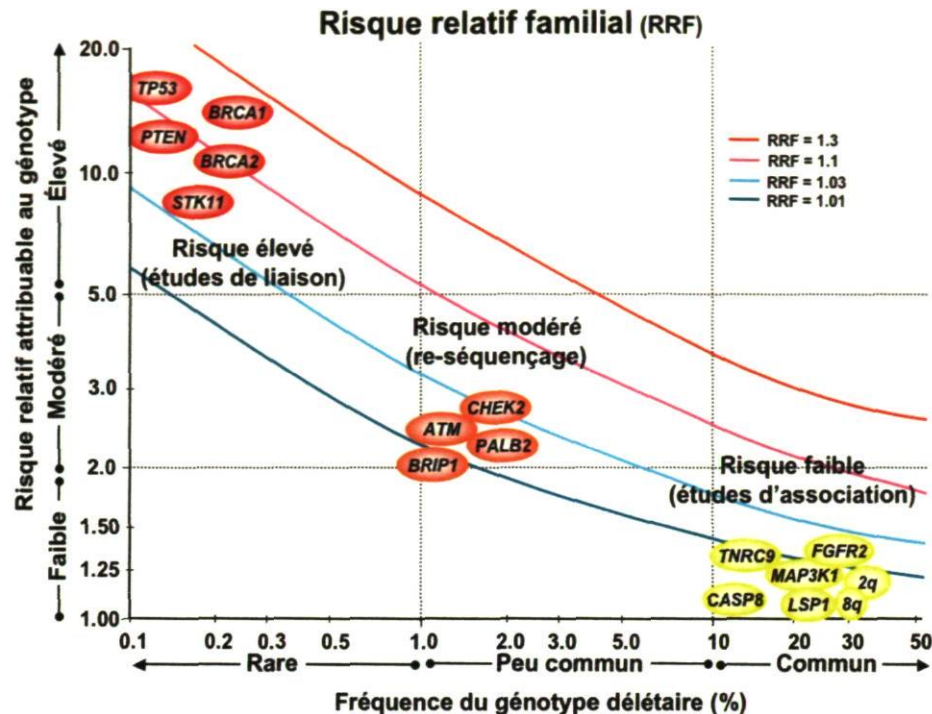
\* Risque estimé de cancer du sein à 45 ans

ND : Non disponible

RR : Risque relatif

OR : Odd ratio, rapport de cotes





**Figure 7**

**Catégories de gènes de susceptibilité au cancer du sein et risque leur étant associé.** RRF : risque relatif familial. Adapté des références <sup>177,230,381</sup>.

### 3.5. Gènes de susceptibilité à faible pénétrance

La plus forte évidence pour une association avec une susceptibilité de faible pénétrance est localisée dans l'intron 2 du gène *FGFR2* (« fibroblast growth factor receptor 2 »)<sup>382-384</sup>. Ce gène encode un récepteur à haute affinité pour plusieurs FGF (Fibroblast Growth Factors) agissant sur la croissance des cellules épithéliales et mésenchymales. Dans ce gène, un groupe de 6 variants fortement liés entre eux forment un allèle de risque de cancer du sein pour lequel le risque de cancer du sein à l'âge de 70 ans a été estimé à 10,5% chez les porteurs homozygotes, 6,7% chez les hétérozygotes et 5,5% chez les non-porteurs de l'allèle de risque.

Un variant localisé dans une région de fort déséquilibre de liaison sur le chromosome 16, et situé à proximité de l'extrémité 5' du gène *TNRC9/TOX3* a permis d'identifier *TNRC9* comme un gène de susceptibilité à faible pénétrance<sup>382</sup>. Ce gène de fonction inconnue contient un motif potentiel du groupe de haute mobilité, ce qui suggère qu'il pourrait exercer la fonction de facteur de transcription.

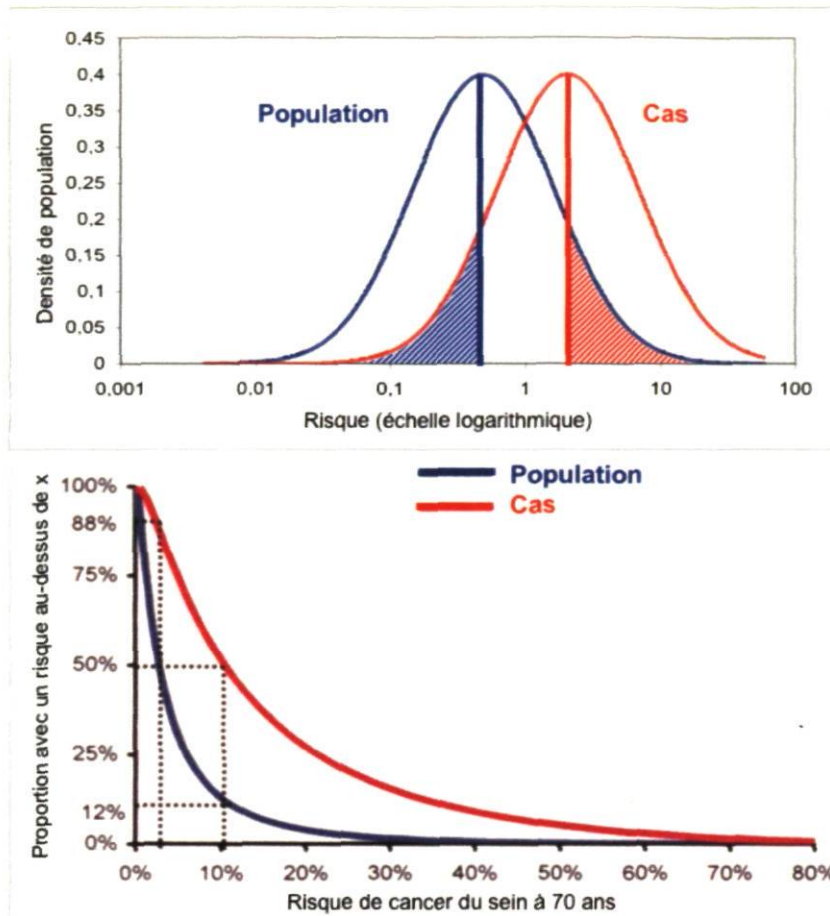
Des variants dans de grandes régions de déséquilibre de liaison ont permis d'identifier les gènes *MRPS30* (PDCD9, programmed cell death protein 9), *MAP3KI* (Mitogen Activated Protein Kinase Kinase Kinase 1) et *LSP1*, comme étant des gènes de faible pénétrance<sup>384</sup>.

Un variant entraînant un changement d'acide aminé (D302H; Asp302His) dans le gène Caspase-8 (*CASP8*), impliqué dans l'apoptose, a été associé à une faible diminution du risque de cancer du sein<sup>385</sup>. D'autres régions de fort déséquilibre de liaison ne contenant toutefois pas de région codante connue ou potentielle ont aussi été associées à une faible pénétrance<sup>381,382</sup>.

Jusqu'à maintenant, quelques allèles de susceptibilité au cancer du sein à faible pénétrance ont été identifiés. Il est peu probable que les autres gènes de susceptibilité à faible pénétrance qui restent à identifier aient un effet aussi substantiel que pour l'allèle de *FGFR2*. Les gènes avec un effet plus important sont plus faciles à identifier que ceux avec un effet plus faible en raison de la taille des échantillons à utiliser pour identifier ces derniers. Toutefois, certaines évidences appuient l'existence de plusieurs autres allèles communs de susceptibilité avec des effets plus faibles<sup>382</sup>.

### 3.5.1. L'effet polygénique

Les résultats des études menées sur des jumeaux<sup>40,386</sup>, ceux des études de liaison effectuées chez des familles présentant plusieurs cas de cancer du sein<sup>387</sup> et les résultats d'un modèle basé sur la distribution de cas familiaux de cancer du sein suggèrent l'existence de plusieurs gènes de susceptibilité conférant un large éventail de risque<sup>388</sup>. Ce modèle polygénique permettrait de représenter les effets d'un grand nombre de gènes, conférant chacun un faible risque de cancer du sein, mais agissant entre eux de manière multiplicative<sup>44</sup>. Ainsi, selon un modèle polygénique, il a été estimé que la moitié de tous les cas de cancer du sein surviendrait dans 12% de la population avec un risque élevé de cancer du sein<sup>388</sup> (Figure 8).



**Figure 8**

**Distribution du risque de cancer du sein dans la population en général et chez les cas de cancer du sein. Partie supérieure.** La distribution du risque chez les individus qui développeront le cancer (cas) serait plus élevée que celle chez les individus qui ne le développeront pas (population). Ceci se reflète par un risque moyen plus élevé chez les cas. **Partie inférieure.** Proportion de la population et des cas avec un risque supérieur à une valeur de x%. Adapté de Pharoah et al., Nat Genet 2002<sup>388</sup>.

Deux modèles généraux ont été proposés afin d'illustrer la prédisposition polygénique. Le premier modèle, appelé « common variant-common disease », est basé sur la présence d'un variant ayant apparu précocement lors de l'histoire d'une population. Un tel variant aurait potentiellement une fréquence relativement élevée mais un effet faible. Le deuxième modèle stipule que des variations isolées et survenues récemment pourraient expliquer le modèle polygénique, ce qui ferait en sorte qu'une multitude de variants rares seraient responsables d'un même phénotype et qu'ils resteraient à identifier, demandant donc une étude détaillée de chacun de ceux-ci.

### 3.5.2. Modèle multiplicatif - Interaction entre les différents gènes de prédisposition au cancer du sein

Le modèle le plus employé pour la prédiction de l'effet de plusieurs facteurs de prédisposition est le modèle multiplicatif. Selon ce modèle, les différents facteurs agiraient d'une manière indépendante et multiplicative. Le risque combiné associé à la présence de deux facteurs serait toutefois dépendant de la nature de l'interaction de ceux-ci. Les dernières études semblent démontrer que les variants communs à faible pénétrance interagissent généralement entre eux d'une manière multiplicative<sup>382</sup>. De plus, certains variants communs à faible pénétrance semblent agir d'une manière multiplicative sur le risque de cancer du sein chez les porteurs de mutations à forte pénétrance dans les gènes *BRCA1* ou *BRCA2*<sup>389</sup>. Toutefois, ce type d'interaction ne semble pas être présent dans tous les cas. Par exemple, le variant 1100delC de *CHEK2* n'influence pas le risque de cancer du sein chez des porteurs de mutations dans les gènes *BRCA1* et *BRCA2*<sup>292</sup>. De la même manière, certains facteurs environnementaux peuvent moduler le risque chez les porteurs de mutations dans *BRCA1* et *BRCA2*<sup>390</sup>.

#### **4. Recherche de gènes de susceptibilité au cancer du sein**

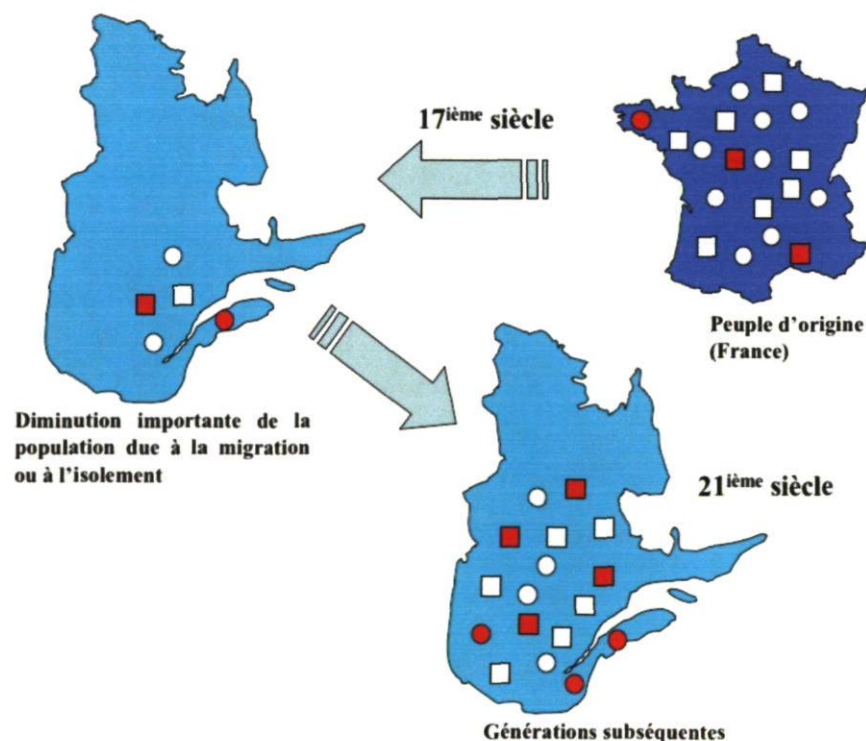
*Les types d'études pouvant mener à l'identification de gènes de susceptibilité au cancer du sein*

Trois principaux types d'approches peuvent servir à identifier et localiser les gènes contribuant au développement de maladies communes. Jusqu'à maintenant, des études de liaison, d'association et des études de gènes candidats ont mené à l'identification de gènes de susceptibilité au cancer du sein. Les études de liaison et les d'études d'association peuvent être toutes deux menées sur l'ensemble du génome. Les études de liaison et d'association menées sur l'ensemble du génome sont basées sur le fait que les variants ou les marqueurs génétiques situés à proximité de gènes causant des maladies ségrègent avec la maladie. Ce principe de ségrégation simultanée de marqueurs génétiques entre eux, appelé déséquilibre de liaison (linkage disequilibrium; LD) est à la base des études de liaison. Puisque les variants qui causent des maladies monogéniques sont habituellement des variants rares, les allèles ségrégant avec la maladie peuvent être identifiés chez les individus porteurs car ils partagent le même allèle. Toutefois, dans le cadre de l'identification et la localisation de facteurs génétiques associés à des mutations de faible pénétrance ou à des maladies complexes et génétiquement hétérogènes, les études de liaison remportent moins de succès.

Une approche alternative aux études menées sur l'ensemble du génome est d'effectuer des études de gènes candidats afin d'examiner directement l'implication de variants ou de gènes choisis sur le risque de cancer. Cette approche, pouvant être menée sous forme d'étude cas-contrôles ou sur une base familiale, possède une puissance plus grande que les études de liaison pour identifier des gènes de faible pénétrance. Cette méthode peut être effectuée de deux manières différentes. L'une implique la sélection de variants précis à analyser lors de grandes études d'association. L'autre consiste au re-séquençage de gènes candidats chez un nombre limité d'individus ou de familles dans le but d'identifier des variants potentiellement associés à la maladie. Cette étape de re-séquençage peut ensuite être suivie d'études d'association permettant de confirmer l'association identifiée précédemment. Le re-séquençage de gènes candidats suivie d'études d'association chez des femmes ou des familles atteintes permet à la fois l'identification de variants rares ayant une forte pénétrance et l'identification de gènes de faible pénétrance.

#### 4.1. Utilisation de populations fondatrices pour la recherche de gènes de susceptibilité

Il est possible de remarquer une différence dans la prévalence de mutations d'une population à l'autre. Pour certaines populations, une sur-représentation d'une mutation pourrait être attribuable à un effet fondateur (Figure 9). Le phénomène de l'effet fondateur est lié à l'établissement d'une population sur un nouveau territoire. Il se caractérise par un changement du profil des fréquences alléliques du groupe migrant par rapport à sa population d'origine. Il y a apparition d'un effet fondateur lorsqu'un petit nombre d'individus sont à la base de l'établissement ou de la fondation d'une population, ou lorsqu'un goulot d'étranglement survient dans l'histoire d'une population, réduisant cette population à un faible nombre d'individus. C'est d'ailleurs le cas pour la population du Québec. La population actuelle du Québec est formée de 7,3 millions d'habitants, desquels approximativement 6 millions dérivent de près de 8500 colons français venus coloniser la Nouvelle-France entre 1608 et 1759<sup>391</sup>.



**Figure 9**

**Effet fondateur au Québec.** Près de 8500 colons français sont venus coloniser la Nouvelle-France entre 1608 et 1759. De nos jours, approximativement 6 millions des 7,3 millions d'habitants de la population du Québec descendent de ces colons français. Ceci pourrait avoir causé une sur-représentation de certaines mutations dans notre population. Adapté de De Braekeleer et Dao, Hum Biol 1994<sup>391</sup>.

L'analyse d'une population fondatrice s'avère utile lors d'études génétiques pour l'identification et la localisation de marqueurs ou de gènes liés à des traits complexes. Dans ce type de population, l'hétérogénéité génétique est réduite, ce qui fait en sorte qu'un plus grand nombre d'individus avec une maladie seront porteurs d'une même mutation, facilitant ainsi son identification. Il y aura également un plus fort déséquilibre de liaison entre la mutation et les marqueurs environnants, donc la région porteuse de la mutation sera plus facilement identifiée en raison de la coségrégation de cette région avec la maladie<sup>392</sup>.

#### **4.2. Utilisation de familles enrichies**

Les gènes de susceptibilité au cancer du sein à forte pénétrance contribuent approximativement à 20% du risque familial de cancer du sein. Le reste du risque familial de cancer du sein doit donc être expliqué par des gènes conférant un risque plus modeste. Le re-séquençage de gènes candidats chez des patients atteints et chez des contrôles sains semble être la méthode la plus efficace afin d'identifier de nouveaux gènes de susceptibilité au cancer du sein à pénétrance modérée. Puisque le risque associé à des mutations dans de tels gènes est seulement de 2 à 4 fois, et que la fréquence de ces mutations est faible, le nombre d'individus requis dans de telles études doit être relativement élevé, ce qui complique significativement les analyses. L'utilisation d'individus provenant de familles à risque élevé de cancer du sein, en comparaison à des cas sélectionnés sur la base d'études populationnelles, pourrait permettre de pallier aux problèmes associés au re-séquençage d'un grand nombre d'individus puisque des allèles de plus faible susceptibilité au cancer du sein sont généralement enrichis chez des cas familiaux en comparaison avec des cas non-familiaux<sup>137</sup>. L'utilisation d'individus ayant une histoire familiale de cancer du sein permet de diminuer considérablement la taille de l'échantillon dans des études d'association. En effet, en comparaison avec une étude d'association standard dans laquelle les individus n'ont pas été sélectionnés sur la base de la présence d'une histoire familiale de cancer du sein, la taille de l'échantillon nécessaire pour détecter un allèle de susceptibilité avec une fréquence commune est réduite de plus de deux fois si les cas de cancer du sein ont été sélectionnés sur la base de la présence d'un apparenté de premier degré atteint et de plus de 4 fois si les cas ont été sélectionnés sur la base de la présence de deux apparentés de premier degré atteints<sup>49</sup> (Tableau 3).

**Tableau 3**

**Pourcentage de réduction de la taille de l'échantillon obtenu par l'utilisation de cas avec une histoire familiale de cancer du sein en comparaison à des cas non sélectionnés sur la base d'une histoire familiale selon un modèle multiplicatif et avec une puissance statistique de 90% à un niveau de significativité de  $10^{-4}$  \*. Adapté d'Antoniou et Easton, Genet Epidemiol 2003<sup>49</sup>.**

Type d'histoire familiale	Risque relatif de 1,5 Fréquence allélique					Risque relatif de 2,0 Fréquence allélique				
	0,01	0,05	0,10	0,20	0,30	0,01	0,05	0,10	0,20	0,30
Mère atteinte	58,7	56,8	56,2	54,5	53,0	60,3	59,3	57,7	54,5	51,4
Mère et soeur atteintes	77,6	76,4	75,7	73,9	72,3	79,8	78,4	76,6	73,2	70,0
Cas bilatéraux	78,4	77,4	76,9	75,6	74,4	79,8	78,8	77,5	74,6	72,1
Fortte histoire familiale	83,7	82,1	80,7	77,7	75,1	87,5	85,2	82,1	76,3	71,0

\* Âge du cas référence au diagnostic : 45 ans

#### 4.3. Utilisation de blocs d'haplotypes

La recherche d'allèles de faible pénétrance pour le cancer du sein a permis d'identifier plusieurs associations se retrouvant dans les régions introniques ou intergéniques pour lesquelles aucun gène n'est connu. Ceci tend à démontrer l'importance du LD dans des études de ce type ainsi que le fait qu'un variant causal d'une maladie ne doit pas obligatoirement se situer dans la région codante d'un gène. En accord avec le principe de LD, la majorité des variants communs sont séparés en groupes qui sont presque parfaitement corrélés entre eux. Ainsi, le génotype de l'un permet de prédire le génotype des variants environnants avec lesquels il est parfaitement corrélé. Nous pouvons donc décrire le génome comme une série de régions à haut LD, appelées des blocs de LD. Ces blocs de LD sont séparés par de courtes régions avec un faible LD communément appelés "hotspots" de recombinaison<sup>393</sup>. Lorsque la distribution du LD pour une région est connue, quelques SNPs marqueurs (tagging SNPs; tSNPs) peuvent être sélectionnés de manière à ce que seuls ou en combinaison (haplotypes), ils puissent capturer la majorité des variations dans une région. Ainsi, l'identification d'haplotypes communs dans chacun des blocs permet également de sélectionner un ensemble minimal de SNPs qui permettent de distinguer les haplotypes communs entre eux, donc d'inférer le génotype des autres variants faisant partie du même bloc.

#### 4.4. Utilisation d'outils bioinformatiques

La recherche de mutations ne résulte pas seulement en l'identification de mutations tronquantes ou délétères, mais peut résulter en l'identification de mutations faux-sens, de



variants affectant potentiellement l'épissage et d'insertions ou de délétions préservant le cadre de lecture. Ces divers types de variants sont regroupés sous le terme variant de significativité inconnue (variants of unknown significance; VUS). Différentes données peuvent être combinées afin de déterminer la probabilité de pathogénicité d'un variant. Certaines données directes, telles la ségrégation d'un variant avec la maladie, la sévérité de l'histoire familiale de cancer et la co-occurrence du variant d'intérêt avec un variant pathogène reconnu dans le même gène peuvent être analysées pour déterminer la pathogénicité d'un variant<sup>394</sup>. Des mesures indirectes pour déterminer l'effet d'un variant peuvent découler d'analyses bioinformatiques de la séquence variante, notamment la conservation évolutive (conservation entre les espèces), la sévérité prédite des substitutions d'acides aminés sur la séquence et la structure, et l'effet prédit des variants sur les séquences impliquées dans l'épissage<sup>395,396</sup>. Cette section s'attardera aux différents types de VUS pouvant être retrouvés lors de la recherche de mutations. Quelques méthodes d'analyses *in silico* et expérimentales seront également présentées.

#### 4.4.1. *Variants situés dans la région codante*

Des analyses *in silico* peuvent faciliter la classification des substitutions faux-sens. Les caractéristiques physicochimiques des variants et les modifications structurales engendrées par ceux-ci peuvent être analysées, tout comme le niveau de conservation des résidus variants dans l'évolution à cette position de la protéine. Ceci s'effectue par l'alignement multiple de séquences de la protéine d'intérêt parmi différentes espèces. En général, une surabondance de mutations faux-sens associées à des maladies est retrouvée à des positions hautement conservées<sup>397</sup>. Des programmes intégrant plusieurs caractéristiques simultanément, tels les programmes SIFT, Polyphen, AGVGD et PMUT ont été développés afin d'estimer l'effet potentiel de VUS situés dans la région codante.

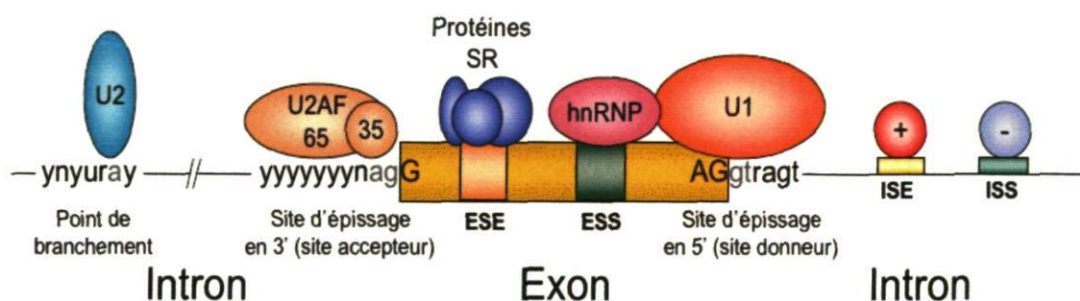
#### 4.4.2. *Variants pouvant potentiellement affecter l'épissage*

Bien que les propriétés physiques et chimiques des protéines, tout comme leur stabilité, leur activité et leur localisation puissent être régulées par des modifications post-traductionnelles<sup>398</sup>, les événements d'épissage alternatifs sont la première source de diversité protéomique humaine<sup>399,400</sup>. Plusieurs altérations causales de maladies affectant l'épissage sont situées dans les régions introniques<sup>401</sup>. Ces altérations d'épissage pouvant

altérer le niveau d'expression, la taille ou la fonction d'une protéine, il est important d'investiguer l'effet potentiel des variants identifiés afin de déterminer s'ils pourraient être responsables d'événements d'épissage alternatif.

### La machinerie d'épissage

L'excision des introns est effectuée par un assemblage de facteurs d'épissage qui sont collectivement recrutés aux ARN pré-messagers pour former le spliceosome (Figure 10), une machine nucléaire sophistiquée comprenant différents types de petites ribonucléoprotéines nucléaires (small nuclear ribonucleoprotein; snRNP) et de protéines auxiliaires. L'assemblage du spliceosome requiert l'identification du site d'épissage en 5' et du site de branchement reconnus respectivement par les protéines U1 et U2. La séquence de pyrimidines et le dinucléotide AG situé dans la région 3' de l'intron sont respectivement reconnus par les sous-unités 65 kDa et 35 kDa d'U2AF. Plusieurs séquences auxiliaires modulent aussi l'efficacité de l'épissage, incluant les activateurs d'épissage exoniques et introniques (exonic splicing enhancer et intronic splicing enhancer; ESE et ISE) et les répresseurs d'épissage exoniques et introniques (exonic splicing silencer et intronic splicing silencer; ESS et ISS). Alors que les ESE sont reconnus par les protéines SR (serine/arginine-rich proteins; SR proteins), les ESS interagissent avec les protéines hnRNP (heterogenous nuclear ribonucleoprotein particles). Les ISE et les ISS peuvent interagir avec une variété de protéines régulatrices qui agissent en activant ou en réprimant l'épissage<sup>402</sup>.



**Figure 10**

**Éléments régulateurs et protéines impliqués dans le mécanisme d'épissage.** Le mécanisme d'épissage nécessite la reconnaissance des sites d'épissage en 5', en 3' et du site de branchement. Des modulateurs d'épissage exoniques (ESE et ESS) et introniques (ISE et ISS) peuvent aussi être reconnus par des protéines régulatrices pour affecter l'épissage. Adapté de Graveley, Trends Genet 2008<sup>402</sup>.

Les variants pouvant potentiellement résulter en des altérations d'épissage forment une classe de variants pour laquelle les conséquences peuvent être étudiées par la combinaison d'outils bioinformatiques et d'essais *in vitro*. Tout d'abord, il est possible d'utiliser une variété de programmes de prédiction utilisés pour prédire la probabilité qu'un variant affecte le patron d'épissage. Certains programmes ont été développés de manière à prédire des activateurs d'épissage (ESE et ISE) et répresseurs d'épissage (ESS et ISS) possibles. Ils sont basés sur l'identification de séquences potentiellement reconnues par les facteurs d'épissage protéiques. La séquence de ces sites étant relativement mal définie, l'existence d'autres éléments importants pour l'épissage tel la distance du site d'épissage et la force du site d'épissage<sup>403</sup> font en sorte qu'il est relativement difficile de déterminer si ces sites sont utilisés sous des conditions physiologiques. Toutefois, les variants situés dans les dinucléotides introniques flanquant les exons résultent pratiquement dans tous les cas en des altérations d'épissage. Parmi les différents programmes pouvant être utilisés, trois ont été utilisés dans le cadre de nos études, *Splice Site Finder*, *GeneSplicer* et *Splice Site Prediction by Neural Network*. Puisque l'échelle et la méthode d'attribution des valeurs pour la prédiction des sites d'épissage varie d'un programme à l'autre, il est difficile de comparer entre eux les résultats obtenus avec ces différents modèles de prédiction. Il faut bien sûr noter que les analyses *in silico* effectuées à l'aide de ces différents programmes fournissent des prédictions de l'impact potentiel des variants à l'étude, et que les prédictions sont tributaires des connaissances que nous possédons sur les différents mécanismes d'épissage. L'approche générale pour la prédiction et la vérification d'aberrations d'épissage associées à des variants situés dans ou à proximité de sites consensus est relativement directe. Lorsque du matériel provenant de l'individu porteur est disponible, les prédictions peuvent être testées *in vitro* par l'analyse de l'ARN messager. Le variant pourra être considéré comme pathogénique s'il résulte en un ou des transcrits portant un codon stop prématuré, ou qu'une délétion respectant le cadre de lecture mène à la perte d'un domaine fonctionnel.

#### 4.4.3. *Variants pouvant affecter l'expression*

La régulation transcriptionnelle est accomplie par l'activité coordonnée de protéines régulatrices spécifiques qui reconnaissent et lient des éléments de régulation. Les éléments de régulation tels les sites de liaison de facteurs de transcription (transcription factors

binding sites; TFBS), les activateurs et les répresseurs, sont couramment localisés dans la région promotrice des gènes. Parmi les nombreuses approches bioinformatiques développées dans le but de comprendre et de prédire le rôle des éléments de contrôle de la transcription, plusieurs voies ont été explorées. La stratégie la plus commune pour l'identification de potentiels TFBS à partir des séquences d'ADN fait appel à des bases de données d'éléments de régulation expérimentalement identifiés ou dérivés<sup>404,405</sup>. Des algorithmes utilisant la conservation de séquences chez des gènes orthologues pour l'identification des éléments de régulation ont aussi été développés<sup>406,407</sup>. Les programmes de prédiction de sites de liaison de facteurs de transcription possèdent des caractéristiques propres rendant difficile la comparaison des résultats de prédiction obtenus entre eux. Ces prédictions demandent également des confirmations expérimentales.

## 5. Partenaires d'interaction de BRCA1 (gènes candidats analysés)

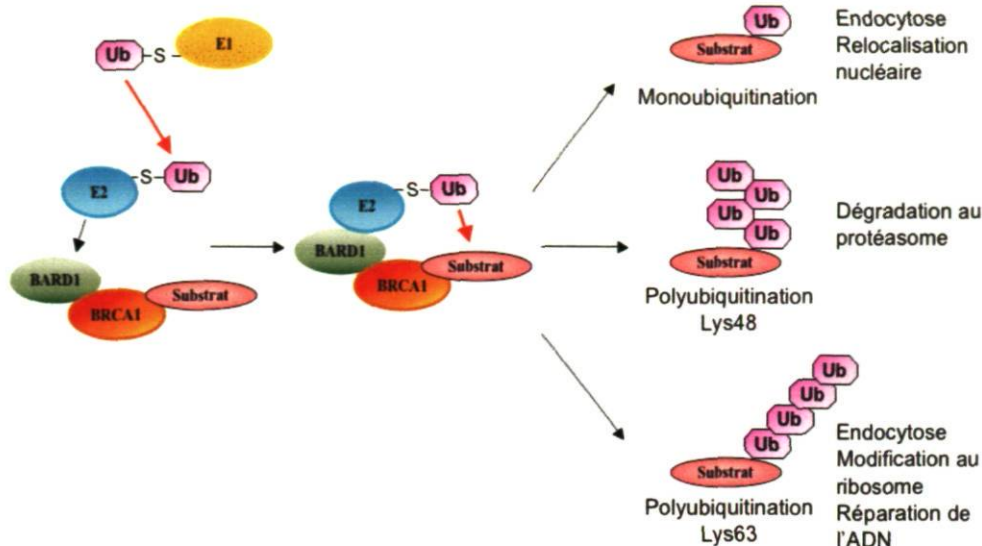
Tel que mentionné à la section 3, BRCA1 joue un rôle dans plusieurs sentiers métaboliques et interagit avec divers partenaires. Notre attention s'est arrêtée sur les gènes *BARD1*, *BAP1*, *AURKA* et *DHX9* en raison de leur interaction avec BRCA1 et leur implication biologique plausible.

### 5.1. *BARD1*

Le gène *BARD1*, localisé en 2q34-35 est composé de 11 exons. La protéine encodée par ce gène interagit avec BRCA1<sup>61</sup>. Ces deux protéines possèdent une organisation structurale semblable. Elles sont toutes deux constituées d'un domaine N-terminal RING de type doigt de zinc et deux motifs BRCT en C-terminal. Le domaine RING de BARD1 est nécessaire pour son interaction avec BRCA1<sup>61</sup>. La protéine BARD1 contient également trois répétitions ankyrines dans sa partie centrale, reconnues pour être impliquées dans des interactions protéine-protéine. Le complexe BRCA1/BARD1 possède une activité E3 ubiquitine ligase et peut catalyser son auto-ubiquitination de type polyubiquitination sur de multiples résidus lysines, ce qui résulte en une augmentation de l'activité ubiquitine ligase du complexe<sup>408</sup>.

### Ubiquitination

Le mécanisme d'ubiquitination implique une action séquentielle d'une enzyme activant l'ubiquitine (E1), une enzyme conjuguant l'ubiquitine (E2) et d'une ubiquitine ligase (E3) (Figure 11). Le type d'ubiquitination du substrat va déterminer le devenir de la protéine cible. Alors que le type de polyubiquitination le plus reconnu (liaison Lys-48) sert au ciblage du substrat au protéasome pour dégradation, la polyubiquitination par des chaînes liées Lys-63 dirige le substrat pour l'endocytose, pour des modifications au ribosome ou pour la réparation de l'ADN. La monoubiquitination aurait quant à elle pour effet de cibler le substrat pour l'endocytose ou pour la localisation nucléaire<sup>409</sup>.



**Figure 11**

**Mécanisme d'ubiquitination.** L'ubiquitination d'un substrat requiert l'activation de l'ubiquitine par une enzyme E1, son transfert à une protéine E2 liée à une ubiquitine ligase E3 permettant ainsi la spécificité pour un substrat. Le type d'ubiquitination du substrat influence alors le devenir de celui-ci. Adapté d'Ohta et Fukuda, *Oncogene* 2004<sup>410</sup>.

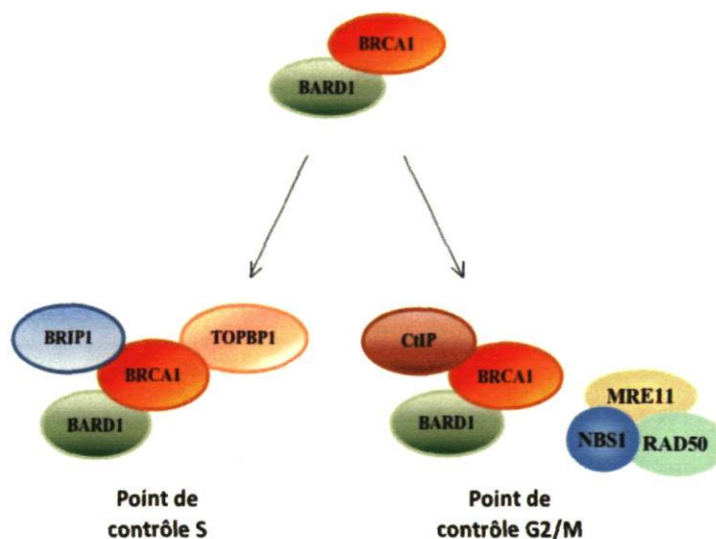
### Rôles de BARD1

Un rôle dans la transcription et la réparation couplée à la transcription (TCR) a été attribué à BARD1. Suite aux dommages à l'ADN, l'hétérodimère BRCA1/BARD1 ubiquitine Rbp1, la plus grande des sous-unités de l'ARN polymérase. Ainsi, la réponse aux dommages à l'ADN par l'ubiquitination de l'ARN polymérase II (ARN pol II) phosphorylée viendrait lier BARD1 à la transcription et à la TCR<sup>411</sup>.

L'hétérodimère BRCA1/BARD1 augmente la stabilité chromosomique en favorisant la réparation des bris d'ADN double brins<sup>412</sup>. Le complexe BRCA1/BARD1 est responsable de la monoubiquitination de l'histone H2AX<sup>408</sup> et de l'ubiquitination de la  $\gamma$ -tubuline impliquée dans la duplication des centrosomes et la ségrégation des chromosomes, ce qui est en accord avec un rôle de BARD1 dans l'intégrité génomique et dans la mitose<sup>413</sup>.

Le complexe BRCA1/BARD1 formerait deux complexes différents possédant des fonctions distinctes dans les points de contrôles du cycle cellulaire<sup>362</sup>. L'un contiendrait la protéine BRIP1 alors que l'autre contiendrait CtIP, un régulateur de la prolifération cellulaire. Le complexe que forme BARD1 avec BRCA1, BRIP1 et TOPBP1 serait impliqué dans le point de contrôle en phase S et contribuerait à retarder la réplication suite à la formation de

bris d'ADN double brins (Figure 12). Le complexe BRCA1/BARD1/CtIP en association avec le complexe MRN serait quant à lui impliqué dans le point de contrôle en G2/M. À ce titre, l'activité enzymatique du complexe BRCA1/BARD1 a été démontrée comme régulant le point de contrôle G2/M en catalysant la polyubiquitination de CtIP induite par les dommages à l'ADN<sup>414</sup>. De plus, BARD1 peut engendrer un arrêt du cycle cellulaire en phase G1 lorsque surexprimée. Cette activité est dépendante de sa localisation nucléaire et de ses NLS<sup>415</sup>.



**Figure 12**

**BARD1 fait partie de complexes ayant des rôles différents dans le cycle cellulaire.** Représentation des partenaires et des rôles attribués au complexe BRCA1/BARD1. Adapté des références <sup>362,414</sup>.

La localisation cellulaire de BRCA1 est régulée par son transport entre le noyau et le cytoplasme. BRCA1 est principalement nucléaire<sup>416</sup> en raison du fait que la protéine BARD1 lie BRCA1 et importe BRCA1 au noyau en cachant son signal d'exportation nucléaire (nuclear export signal; NES)<sup>417</sup>. L'exportation nucléaire de BRCA1 a pour effet de stimuler l'apoptose indépendante de p53. En séquestrant BRCA1 au noyau, BARD1 régule l'apoptose dépendante de BRCA1. BARD1 inhiberait donc l'apoptose, ce qui suggère que la rétention nucléaire du complexe BRCA1/BARD1 contribuerait à la réparation de l'ADN et à la survie cellulaire<sup>418</sup>. D'un autre côté, BARD1 lie p53 et la sous-unité XRCC6 de la DNA-PK<sup>419</sup>, venant l'impliquer dans l'apoptose dépendante de p53. BARD1 joue un rôle de médiateur de l'apoptose en liant et stabilisant p53. De plus, les cellules dans lesquelles BARD1 est réprimé sont résistantes à l'apoptose<sup>419</sup>.

La protéine BARD1 possède une autre activité proapoptotique indépendante de BRCA1. La localisation de BARD1 et son activité apoptotique sont régulées par son transport nucléocytoplasmique. Un NES a d'ailleurs été identifié à proximité de son domaine N-terminal. Une diminution du niveau d'expression de BRCA1 ou une déstabilisation de l'interaction entre BRCA1 et BARD1 résulte en la localisation cytoplasmique de BARD1. Un niveau accru de BARD1 cytoplasmique corrèle avec une augmentation de l'apoptose<sup>420</sup>.

Le complexe BRCA1/BARD1 est requis lors de la mitose pour l'assemblage adéquat des fuseaux mitotiques, pour la transition entre la mitose et l'interphase, pour la ségrégation des chromosomes et l'assemblage nucléaire. Ce complexe, avec son activité E3 ubiquitine ligase, participe à l'organisation des microtubules et à l'assemblage des pôles des fuseaux<sup>421</sup>. La protéine BARD1, lorsqu'elle interagit avec BRCA1, est aussi responsable de l'ubiquitination et de la dégradation de la protéine Aurora B impliquée dans la formation de la plaque équatoriale et dans la progression de la mitose. L'expression aberrante de certaines isoformes de BARD1 a été associée à une fonction proproliférative indépendante de BRCA1 et à un mauvais pronostic de cancer du sein et de l'ovaire<sup>422,423</sup>. L'isoforme  $\beta$  de BARD1, incapable d'interagir avec BRCA1, forme un complexe avec Aurora B et BRCA2 et stabilise Aurora B. La forme sauvage de BARD1 ne peut pas lier BRCA2 et Aurora B<sup>424-426</sup>. Ceci pourrait expliquer le rôle proprolifératif de l'isoforme  $\beta$  de BARD1.

#### Implication de *BARD1* dans la susceptibilité au cancer du sein

Des mutations de *BARD1* sont des événements rares dans les tumeurs. Différentes mutations ont toutefois été identifiées dans des cancers du sein, de l'ovaire et de l'endomètre<sup>427</sup>. Un variant faux-sens (Cys557Ser) de *BARD1* a été lié à une perte de stabilité et d'activité apoptotique du suppresseur de tumeurs p53<sup>428</sup>. Ceci a mené à l'analyse de ce variant à plusieurs reprises dans un contexte de cancer du sein, suggérant une association de ce variant avec le cancer du sein selon certaines études<sup>429-431</sup> alors que d'autres études ne sont pas venues confirmer ces résultats<sup>432-435</sup>. Il a aussi été proposé que certains haplotypes de BARD1 pourraient être associés à un risque accru ou diminué de cancer du sein<sup>436</sup> et des isoformes aberrantes de BARD1 ont été associées avec un mauvais pronostic de cancer du sein et de l'ovaire<sup>422,423</sup>.



La liaison de BARD1 avec BRCA1 et la régulation de l'ubiquitination, son implication dans la transcription, la TCR, la réparation de l'ADN et l'apoptose, tout comme le nombre croissant d'évidences liant le gène *BARD1* à la susceptibilité au cancer du sein nous ont mené à évaluer ce gène dans notre cohorte de familles à risque élevé de cancer du sein.

## **5.2. BAP1**

Le gène encodant cette protéine s'étend sur une distance de 9 kb, est composé de 17 exons et est localisé à 3p21.3. La protéine BAP1, étant composée de 729 acides aminés, est une protéase à thiol appartenant à la famille des UCH et localisée au noyau. Le domaine UCH de BAP1 est localisé dans sa partie N-terminale (résidus 5-214). La région C-terminale de BAP1 (résidus 598-729) interagit avec le domaine RING de BRCA1 et contient également deux NLS potentiels. L'interaction de BAP1 avec BRCA1 a mené à la suggestion qu'elle pourrait influencer la fonction de BRCA1 par son action de déubiquitination. La voie ubiquitine-protéasome pourrait donc être régulée par BRCA1 ou serait un effecteur de BRCA1. La présence de BAP1 et sa liaison à BRCA1 augmente l'activité de répression de la croissance des cellules cancéreuses du sein. Cette observation a d'ailleurs mené à la suggestion que *BAP1* serait un gène suppresseur de tumeurs qui agirait dans la voie de contrôle de la croissance cellulaire par BRCA1<sup>437</sup>.

### Rôles de BAP1

Puisque BAP1 lie BRCA1, il a été proposé que BAP1 pourrait jouer un rôle dans les processus régulationnels dépendants de l'ubiquitine se déroulant au noyau, tels la transcription, le remodelage de la chromatine, le contrôle du cycle cellulaire, la recombinaison et la réparation de l'ADN<sup>437</sup>. Deux rôles principaux ont été initialement proposés pour BAP1. L'association entre BAP1 et BRCA1 pourrait permettre à BAP1 de cibler les protéines interagissant avec BRCA1. Il se pourrait également que BAP1 soit impliquée dans la régulation de la localisation cellulaire de BRCA1 ou des protéines interagissant avec BRCA1<sup>438</sup>.

La protéine BAP1 a été suggérée comme étant impliquée dans la TCR<sup>439</sup> par son interaction avec BRCA1 qui est reconnu pour participer à ce processus cellulaire<sup>87,440</sup>, par son implication dans la réparation de l'ADN<sup>441</sup> et son interaction avec l'ARN pol II<sup>442</sup>.

L'activité ubiquitine hydrolase de BAP1 et sa localisation nucléaire ont été démontrées comme étant nécessaires à sa fonction suppresseur de tumeurs<sup>443</sup>. L'activité de suppression de tumeurs assignée à BAP1 s'exécute via le contrôle de la prolifération cellulaire et par son association avec des gènes impliqués dans la transition G1/S<sup>444</sup>. Tout comme BRCA1, BAP1 vient lier le domaine RING de BARD1, ce qui a comme effet d'inhiber l'activité E3 ubiquitine ligase du complexe. L'inhibition de l'expression de BAP1 cause une hypersensibilité des cellules aux IR et retarde la progression en phase S, suggérant ainsi que BAP1 et le complexe BRCA1/BARD1 agissent afin de réguler l'ubiquitination lors de la réponse aux dommages à l'ADN et du cycle cellulaire<sup>445</sup>.

#### BAP1 et la susceptibilité au cancer

Le gène *BAP1* fait partie d'une région démontrant fréquemment une perte d'hétérozygotie dans le cancer du sein. En effet, une étude a démontré une perte d'hétérozygotie de la région 3p21-22 dans 32% des tumeurs du sein<sup>446</sup>. Une étude subséquente possédant une meilleure résolution a démontré une perte d'hétérozygotie dans 69% des tumeurs primaires de cancer du sein pour la région chromosomique 3p21.3 contenant le gène *BAP1*<sup>447</sup>.

Les premières analyses menées sur le gène *BAP1* ont permis d'identifier des réarrangements génomiques, des délétions et des mutations faux-sens dans des lignées cellulaires de cancer du poumon et dans des tumeurs du sein<sup>437,438</sup> suggérant son implication dans ces deux types de cancers. Toutefois, une seule autre étude s'est attardée à la contribution de *BAP1* au cancer du sein<sup>448</sup>. Selon cette étude menée chez 47 Françaises atteintes du cancer du sein et ayant été testées négatives pour des mutations *BRCA1* et *BRCA2*, le gène *BAP1* ne serait pas être relié à un risque élevé de cancer du sein. Ce groupe n'avait toutefois pas été en mesure d'éliminer la possibilité que des variants de ce gène soient associés à une faible susceptibilité au cancer du sein.

Les rôles potentiels multiples de BAP1 et le faible nombre d'études ayant porté sur ce gène candidat du cancer du sein nous ont mené à nous attarder à ce partenaire d'interaction de BRCA1 afin d'évaluer sa contribution au cancer du sein dans les familles à risque élevé provenant de la population canadienne-française.

### 5.3. *AURKA*

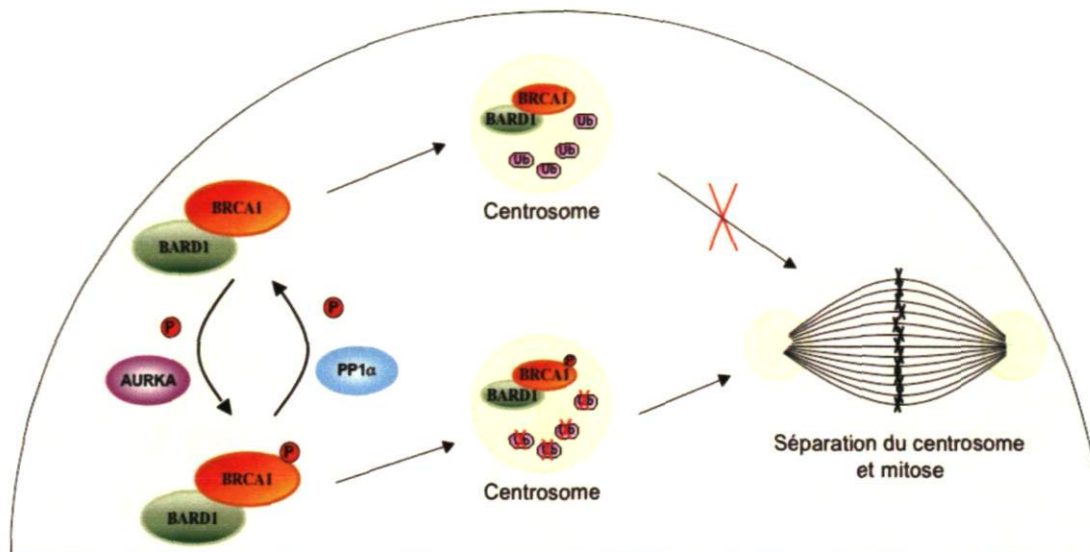
Le gène *AURKA*, aussi connu sous le nom d'*Aurora A*, de *STK15* et de *BTAK*, est localisé en 20q13.2-q13.3. Ce gène encode une ser/thr kinase impliquée dans le point de contrôle G2/M et dans la mitose. La protéine *AURKA* serait impliquée dans la formation des microtubules et dans la stabilisation des pôles des fuseaux lors de la ségrégation des chromosomes. Elle est retrouvée au centrosome lors de l'interphase et aux pôles des fuseaux lors de la mitose. Son niveau d'expression est régulé lors du cycle cellulaire<sup>449</sup>. La partie N-terminale d'*AURKA* est constituée de deux régions conservées chez les protéines de la famille Aurora alors que la région centrale d'*AURKA* contient le domaine kinase<sup>450</sup>.

#### Rôles d'*AURKA*

À la fin de la phase G1 du cycle cellulaire, le centrosome de la cellule se réplique de manière que lors de la transition G2/M, la cellule contienne deux centrosomes qui formeront les fuseaux mitotiques. Le suppresseur de tumeurs BRCA1 localise aux centrosomes<sup>451,452</sup>, est important pour la régulation du nombre de centrosomes<sup>453</sup> et régule directement l'activité de nucléation des microtubules du centrosome<sup>452,454</sup>. L'activité ubiquitine ligase du complexe BRCA1/BARD1 est critique pour la fonction du centrosome<sup>454</sup>. La protéine *AURKA* serait impliquée dans la séparation du centrosome<sup>455</sup> ainsi que dans la maturation, l'assemblage et la stabilité des fuseaux<sup>456</sup>. La protéine *AURKA* lie BRCA1 et possède la capacité de la phosphoryler, ce qui a mené à la suggestion que la phosphorylation de BRCA1 par *AURKA* jouerait un rôle dans la transition G2/M<sup>111</sup>. L'activité kinase d'*AURKA* rend les cellules mitotiques réfractaires à l'inhibition par BRCA1. Que ce soit en ciblant directement BRCA1 ou en ciblant les protéines présentes aux centrosomes à l'interphase ou à la mitose, *AURKA* affecterait la fonction inhibitrice de BRCA1 sur les centrosomes<sup>457</sup>.

Selon un modèle intégrant ces données (Figure 13), la protéine *AURKA* modulerait l'inhibition de la fonction des centrosomes de BRCA1 en inhibant son activité E3 ubiquitine ligase. À l'opposé, la déphosphorylation de BRCA1 par la protéine phosphatase 1 (PP1) augmenterait l'activité ubiquitine ligase de BRCA1. L'inhibition du potentiel de nucléation des microtubules du centrosome par l'activité E3 ubiquitine ligase de BRCA1

serait donc contrôlée par sa régulation effectuée par la kinase AURKA et la phosphatase PP1 lors du cycle cellulaire<sup>457</sup>.



**Figure 13**

**AURKA régule l'activité de modulation des centrosomes de BRCA1.** La protéine AURKA, en phosphorylant BRCA1, diminue l'activité ubiquitine ligase du complexe BRCA1/BARD1 qui régule la séparation du centrosome. Ainsi, AURKA favorise la formation des fuseaux mitotiques et la progression de la mitose. Adapté de Sankaran et al., *Cancer Res* 2007<sup>457</sup>.

Une autre cible d'AURKA vient montrer son implication dans la mitose. La phosphatase CDC25B est un initiateur des événements précoces de la mitose<sup>458</sup>. La phosphorylation de CDC25B aux centrosomes est dépendante d'AURKA, qui pourrait participer à la régulation temporelle et spatiale d'événements biochimiques nécessaires pour l'entrée en mitose<sup>108,459</sup>.

#### Implication d'AURKA dans la susceptibilité au cancer

AURKA a été impliquée initialement dans la formation de tumeurs. Une expression élevée d'AURKA a été démontrée dans plusieurs types de tumeurs, incluant des tumeurs du côlon, de la vessie, de l'ovaire et du sein<sup>450,460,461</sup>. La surexpression d'AURKA est corrélée avec les tumeurs du foie et de la prostate à haut grade<sup>462,463</sup>.

L'identification d'AURKA à titre de gène de susceptibilité à faible pénétrance pour le cancer chez l'humain et la souris<sup>449</sup> a mené à l'analyse de ce gène dans différents types de cancer. Des associations entre des variants d'AURKA et un risque accru de carcinomes des cellules squameuses de l'oesophage<sup>464</sup>, de cancer du côlon<sup>465</sup> et du cancer de l'ovaire<sup>466,467</sup>

ont été rapportées. Des variants d'*AURKA* ont également été évalués dans des cas de cancer du sein<sup>468-475</sup> sans toutefois en venir à des conclusions définitives. Dans un même ordre d'idée, *AURKA* a également été évalué à titre de gène modificateur de risque de cancer du sein chez des porteurs de mutations dans *BRCA1* et *BRCA2*. Ceci a permis d'éliminer un effet modificateur de risque du variant Phe31Ile<sup>476</sup>.

*AURKA* régule la ségrégation des chromosomes lors de la mitose, un élément important pour le maintien de la stabilité génomique. Son interaction avec *BRCA1* et la régulation de son activité E3 ubiquitine ligase, nous ont poussé à évaluer ce gène dans notre cohorte de familles à risque élevé de cancer du sein non-porteuses de mutations dans les gènes *BRCA1* et *BRCA2*, de même que chez des porteuses de mutations de *BRCA1* et *BRCA2*.

#### 5.4. *DHX9*

Le gène *DHX9* est situé en 1q25, une région rapportée comme étant un locus de susceptibilité au cancer de la prostate<sup>477</sup>. Ce gène composé de 27 exons est exprimé de manière ubiquitaire<sup>477</sup> et encode une hélicase de la famille des DEAH<sup>478-480</sup>. Cette hélicase possède la capacité de dérouler l'ADN et l'ARN double brins<sup>478,481,482</sup>. Son activité est dépendante de l'ATP et nécessite sa région centrale hélicase ainsi que l'intégrité de ses domaines N-terminaux (dsRBD I et II; double-stranded RNA binding domain) et de son domaine RGG (Arg-Gly-Gly)<sup>483</sup>. Les domaines dsRBD sont reconnus pour lier préférentiellement l'ARN. Il a été démontré que le domaine RGG de *DHX9* lie préférentiellement l'ADN simple brin<sup>483</sup>. La partie C-terminale de la protéine *DHX9* contient une séquence signal facilitant son transport à travers les pores nucléaires<sup>484</sup>.

#### Rôles de *DHX9*

La protéine *DHX9* joue différents rôles au niveau du noyau et du cytoplasme de la cellule. En effet, *DHX9* a été impliquée dans la transcription, le transport et la traduction des ARN. Il a été démontré que *DHX9* lie CBP/p300 et l'ARN pol II et qu'elle est requise pour la formation du complexe contenant ces deux protéines<sup>485</sup>. Des mutations tronquantes de *DHX9* affectent l'interaction entre *DHX9* et l'ARN pol II<sup>88</sup>. De plus, *DHX9* interagit avec la sous-unité du facteur nucléaire NF- $\kappa$ B et agit à titre de coactivateur transcriptionnel<sup>486</sup> permettant à la sous-unité p65 de recruter le co-activateur CBP/p300. Ainsi, par son

interaction avec l'ARN pol II, DHX9 semble faciliter le recrutement et/ou l'assemblage du complexe de transactivation<sup>486</sup>.

Un fait important permettant de faire un lien entre *DHX9* et le cancer du sein est l'identification d'une interaction entre DHX9 et BRCA1<sup>88</sup>. Il a été démontré que DHX9 est la protéine permettant de relier BRCA1 et les autres composantes de l'holoenzyme ARN pol II. En effet, les résidus 230-325 de RHA lient les résidus 1560-1863 de BRCA1. Cette région de BRCA1 semble être celle requise pour la liaison de BRCA1 à l'ARN pol II. DHX9, liant à la fois BRCA1 et le coactivateur transcriptionnel CBP, serait donc la protéine faisant le pont entre BRCA1 et l'holoenzyme ARN pol II<sup>88</sup>. Des mutations tronquantes de *DHX9* affectent l'interaction entre ces deux protéines, ce qui a pour effet de diminuer l'activité transcriptionnelle de BRCA1<sup>88</sup>. Ceci pourrait possiblement affecter la fonction de BRCA1 dans la transcription et dans la TCR, liant ainsi DHX9 à la TCR. Plusieurs autres protéines nucléaires impliquées dans l'épissage, le traitement et la maturation de l'ARN interagissent avec DHX9, notamment les ribonucléoprotéines hétérogènes nucléaires et les petites ribonucléoprotéines nucléaires<sup>487</sup>.

DHX9 est aussi impliquée dans l'exportation des transcrits d'ARN messager du noyau vers le cytoplasme par son interaction avec les protéines de transport nucléaire TAP et HAP95<sup>488,489</sup>. Au cytoplasme, DHX9 est préférentiellement associé aux polyribosomes en cours de traduction et est impliqué à titre d'effecteur de l'expression génique de la transcription à la traduction par reconnaissance des ARN possédant une structure complexe à leur extrémité 5'<sup>490</sup>.

La protéine DHX9 a également été impliquée dans la réparation des dommages à l'ADN et dans la tumorigenèse. DHX9 interagit avec l'hélicase WRN, la protéine déficiente dans le syndrome de Werner. Ce syndrome est associé avec un vieillissement prématuré et une prédisposition au développement de tumeurs. L'interaction entre ces deux protéines a comme effet d'inhiber les activités NTPase et hélicase de DHX9 alors que cette même interaction a comme effet d'augmenter l'activité exonucléase 3'→5' de WRN<sup>491</sup>. Ceci vient impliquer DHX9 dans le maintien de la stabilité génomique en favorisant la processivité de WRN<sup>491</sup>.

### Implication de *DHX9* dans le cancer

Bien que *DHX9* n'ait pas été évalué dans un contexte de cancer du sein, des altérations de ce gène ont été rapportées dans le cancer du poumon et dans le cancer de la prostate. Une étude a démontré une surexpression de *DHX9* dans des tumeurs du poumon<sup>492</sup> et la région chromosomique contenant *DHX9* est amplifiée dans des tumeurs de la prostate<sup>493</sup>.

L'effet de mutations tronquantes de *DHX9* sur la fonction de BRCA1 a mené à l'évaluation de ce gène à titre de gène candidat dans notre cohorte de cas de cancer du sein familiaux. Le rôle central de *DHX9* dans la transcription et son implication dans une multitude de processus, incluant la réparation de l'ADN, démontrent également la logique du choix de ce gène à titre de gène candidat du cancer du sein. De plus, la contribution de ce gène n'a jamais été évaluée au niveau du cancer du sein.

## 6. Problématique

### *Mise en contexte*

Des analyses menées chez des familles à risque élevé de cancer du sein ont mené à l'identification de *BRCA1* et *BRCA2*<sup>46-48</sup>. Des mutations dans ces gènes confèrent un risque de cancer du sein accru de 10 à 20 fois<sup>49</sup>. Au niveau populationnel, des mutations de ces gènes permettraient d'expliquer 15 à 20% de l'excès du risque familial du cancer du sein<sup>44,57-60</sup>.

Les découvertes récentes ont permis de dresser une image plus claire de l'architecture de la susceptibilité au cancer du sein. Les gènes *BRCA1* et *BRCA2* sont les gènes majeurs de susceptibilité au cancer du sein et avec les autres gènes de forte pénétrance, *TP53*, *PTEN* et *STK11*, ils contribuent approximativement à 20% du risque familial de cancer du sein. Le reste du risque familial de cancer du sein doit donc être expliqué par des gènes conférant un risque plus modeste. Les gènes *ATM*, *CHEK2*, *BRIP1* et *PALB2* sont des gènes de susceptibilité possédant des similarités biologiques à *BRCA1* ou *BRCA2*. Ils confèrent un risque de cancer du sein de l'ordre de 2 à 4 fois. Une dernière catégorie de gènes de susceptibilité au cancer du sein est apparue récemment. Cette catégorie renferme les gènes de faible pénétrance qui confèrent un risque de l'ordre de 1,5. Les variations dans ces gènes sont communes, mais n'expliqueraient que 3,6% de l'excès du risque familial de cancer du sein<sup>382</sup>.

En 2001, une équipe de recherche internationale nommée INHERIT BRCA (INterdisciplinary HEalth Research International Team on BREast CANcer susceptibility) a été formée. L'un des buts de cette équipe était de caractériser le profil des mutations de *BRCA1* et *BRCA2* dans la population canadienne-française. L'étude de cette population fondatrice a permis de déterminer que ces deux gènes de susceptibilité au cancer du sein sont responsables d'environ 24% des familles à risque élevé<sup>494</sup>. Ainsi, une grande proportion du risque familial de cancer du sein demeure à expliquer dans les familles à risque élevé provenant de la population canadienne-française.



### *Hypothèse*

Puisque des mutations de *BRCA1* et *BRCA2* ne sont responsables que d'une faible proportion des familles à risque élevé de cancer du sein provenant de la population canadienne française, il est possible que les autres gènes de susceptibilité identifiés jusqu'à maintenant puissent expliquer une proportion des cas. En raison de l'effet fondateur présent chez les Canadiens français, ces gènes de susceptibilité pourraient possiblement avoir une plus grande contribution au cancer du sein dans cette population que dans d'autres populations. Il est également possible qu'une proportion des cas de cancer du sein familiaux présents dans la population canadienne-française soit expliquée par des gènes toujours inconnus.

### *Objectifs de l'étude*

Dans le cadre du premier volet de mon projet de doctorat, « évaluation de la contribution de gènes de susceptibilité au cancer du sein », les gènes à forte pénétrance *PTEN* et *STK11*, ainsi que les gènes à pénétrance modérée *CHEK2*, *BRIP1/FANCI* et *PALB2/FANCD1* ont été évalués dans une cohorte d'individus à risque élevé de cancer du sein provenant de la population canadienne-française. Cette cohorte composée de 96 femmes atteintes du cancer du sein a été sélectionnée à partir d'une cohorte plus grande mise sur pied dans le cadre du projet INHERIT BRCA. Ces patientes ont été sélectionnées sur la base d'une forte histoire familiale de cancer du sein et sur leur jeune âge au diagnostic. Les ADN des patientes ayant fait l'objet de mes études doctorales ont tous été analysés et testés négatifs pour des mutations et des réarrangements génomiques dans les gènes *BRCA1* et *BRCA2*. En se basant sur la présence d'un grand nombre de cancers du sein chez leurs apparentés de premier, deuxième et troisième degré, nous avons évalué l'implication d'autres gènes de susceptibilité au cancer du sein chez ces participantes.

Dans le cadre du deuxième volet de ce projet de doctorat, « identification de gènes de susceptibilité au cancer du sein », des analyses de gènes candidats ont été priorisées afin d'identifier de nouveaux gènes de susceptibilité au cancer du sein. Comme il en a été question à la section 3.3, les gènes de susceptibilité au cancer du sein présentement connus agissent dans des voies cellulaires interreliées. Plusieurs des protéines encodées par ces gènes interagissent entre elles et influent sur la fonction des autres gènes de susceptibilité.

C'est pourquoi les gènes candidats étudiés dans le cadre du deuxième volet de mes études doctorales encodent des protéines interagissant avec BRCA1, étant ainsi susceptibles de moduler certaines de ses fonctions.

Afin d'évaluer les gènes de susceptibilité et les gènes candidats à l'étude, une approche de re-séquençage a été utilisée. Pour identifier de nouveaux gènes de susceptibilité au cancer du sein, le re-séquençage de gènes candidats a comme avantage de posséder une puissance plus grande que les études de liaison qui sont fortement affectées par l'hétérogénéité génétique et la pénétrance incomplète<sup>495</sup>. De plus, le re-séquençage de gènes candidats permet d'identifier des gènes à pénétrance faible, à pénétrance modérée ou des gènes de susceptibilité à forte pénétrance.

Notre étude a porté sur notre cohorte de 96 femmes atteintes provenant de la population canadienne-française, une population fondatrice pour laquelle un plus grand nombre d'individus avec une maladie pourrait être porteur d'une même mutation, facilitant ainsi son identification. L'utilisation d'individus provenant de familles à risque élevé de cancer du sein a été priorisée. Ceci permet de diminuer considérablement la taille de l'échantillon à utiliser en comparaison avec une étude dans laquelle les individus n'ont pas été sélectionnés. Pour une même puissance statistique, la taille de l'échantillon nécessaire pour détecter un allèle de susceptibilité avec une fréquence commune est réduite de plus de deux fois si les cas de cancer du sein ont été sélectionnés sur la base de la présence d'un apparenté de premier degré atteint et de plus de quatre fois si les cas ont été sélectionnés selon la présence de deux apparentés de premier degré atteints<sup>49</sup>.

L'effet potentiel des variants identifiés a été analysé à l'aide d'une combinaison d'outils bioinformatiques et d'analyses expérimentales. Alors que l'effet potentiel de variations codantes a été évalué par l'alignement multiple de séquences ou à l'aide de programmes intégrant cette information, d'autres programmes ont été utilisés pour analyser l'effet potentiel des variants pouvant affecter l'épissage ou l'expression d'un gène. Dans le cadre de mon projet, lorsque possible, le séquençage de l'ADNc provenant d'individus atteints a été effectué afin d'identifier des aberrations d'épissage dans les gènes étudiés. Des essais de protection à la ribonucléase ont été effectués afin d'identifier la présence d'ARN messager différenciellement épissés, alors que des analyses de type Western ont été menées pour

vérifier le niveau d'expression de la protéine et la présence de protéines altérées. L'effet des variants affectant potentiellement des sites de liaison de facteurs de transcription a été analysé par des essais de gène rapporteur, une méthode permettant de déterminer *in vitro* l'effet d'un variant sur l'activité transcriptionnelle d'un promoteur.

Dans le but de fournir des outils nécessaires à l'évaluation des gènes d'intérêt à grande échelle chez les Canadiennes françaises, nous avons effectué l'identification d'haplotypes, de blocs d'haplotypes et la détermination de tSNPs. La densité de marqueurs modifie généralement la taille des blocs identifiés lors de l'analyse d'une région et la distribution du LD est différente d'une population à l'autre<sup>393</sup>. Pour ces raisons, nous avons généré ces informations pour les gènes *AURKA*, *BAP1*, *BARD1*, *DHX9* et *BRIP1/FANCI*, qui ont été étudiés dans notre cohorte provenant de la population canadienne-française. Ceci facilitera ainsi l'analyse subséquente de ces gènes dans notre population.

## CHAPITRE II

### **Germline mutations in the breast cancer susceptibility gene *PTEN* are rare in high-risk non-*BRCA1/2* French Canadian breast cancer families**

#### **1. Résumé**

Le syndrome de Cowden, causé par des mutations dans le gène *PTEN*, est associé à une susceptibilité accrue au cancer du sein. Afin d'évaluer la proportion des familles canadiennes-françaises à risque élevé de cancer du sein potentiellement porteuses de mutations de *PTEN*, la région codante et les régions introniques flanquantes de ce gène ont été analysées dans notre cohorte de cas de cancer du sein. Bien qu'aucune mutation germinale n'ait été identifiée dans la région codante, notre étude a mené à l'identification de quatre variants introniques. Des analyses plus poussées ont été effectuées afin de déterminer l'effet de ces variants sur l'épissage et la production de la protéine. La présence de ces variants ne semble pas altérer l'épissage de l'ARN ou le niveau de la protéine produite. Nos analyses suggèrent que les mutations germinales de *PTEN* sont rares et qu'il est peu probable qu'elles contribuent à une proportion significative des cas de cancer du sein familiaux dans la population canadienne-française.

Guénard F, Labrie Y, Ouellette G et al. Germline mutations in the breast cancer susceptibility gene *PTEN* are rare in high-risk non-*BRCA1/2* French Canadian breast cancer families. *Fam Cancer*. 2007;6(4):483-90.

#### **2. Manuscrit**

**Running title: *PTEN* mutations in non-*BRCA1/2* French Canadian breast cancer families**

**Germline mutations in the breast cancer susceptibility gene *PTEN* are rare in high-risk non-*BRCA1/2* French Canadian breast cancer families**

Frédéric Guénard<sup>a</sup>, Yvan Labrie<sup>a</sup>, Geneviève Ouellette<sup>a</sup>, Charles Joly Beauparlant<sup>a</sup>, Paul Bessette<sup>b</sup>, Jocelyne Chiquette<sup>c</sup>, Rachel Laframboise<sup>d</sup>, Jean Lépine<sup>e</sup>, Bernard Lespérance<sup>f</sup>, Roxane Pichette<sup>f</sup>, Marie Plante<sup>g</sup>, INHERIT BRCA<sup>h</sup>, Francine Durocher<sup>a, \*</sup>

<sup>a</sup>Cancer Genomics Laboratory, Oncology and Molecular Endocrinology Research Centre, CHUQ, Pavillon CHUL and Laval University, Québec, G1V 4G2, Canada;

<sup>b</sup>Service de gynécologie, Centre Hospitalier Universitaire de Sherbrooke, Fleurimont, J1H 5N4, Canada;

<sup>c</sup>Clinique des maladies du sein Deschênes-Fabia, Hôpital du Saint-Sacrement, Québec c, G1S 4L8, Canada;

<sup>d</sup>Service de médecine génétique, CHUQ, Pavillon CHUL, Québec, G1V 4G2, Canada;

<sup>e</sup>Centre hospitalier régional de Rimouski, Rimouski, G5L 5T1, Canada;

<sup>f</sup>Service d'héματο-oncologie, Hôpital du Sacré-Cœur, Montréal, H4J 1C5, Canada;

<sup>g</sup>Service de gynécologie, CHUQ, L'Hôtel-Dieu de Québec, Québec, G1R 2J6, Canada;

<sup>h</sup>Other members of INHERIT BRCA<sup>s</sup> involved in this study are listed in Appendix 1

\*Corresponding author:

Dr Francine Durocher

Cancer Genomics Laboratory

CHUL Research Centre, CHUQ

2705 Laurier Boulevard, T2-53

Québec city, Québec,

Canada, G1V 4G2

Tel: (418)-654-2296

Fax: (418)-654-2278

**e.mail: [francine.durocher@crchul.ulaval.ca](mailto:francine.durocher@crchul.ulaval.ca)**

**Abstract**

Cowden syndrome is a disease associated with an increase in breast cancer susceptibility. Alleles in *PTEN* and other breast cancer susceptibility genes would be responsible for ~25% of the familial component of breast cancer risk, *BRCA1* and *BRCA2* being the two major genes responsible for this inherited risk. In order to evaluate the proportion of high-risk French Canadian non-*BRCA1/BRCA2* breast/ovarian cancer families potentially harboring a *PTEN* germline mutation, the whole coding and flanking intronic sequences were analyzed in a series of 98 breast cancer cases. Although no germline mutation has been identified in the coding region, our study led to the identification of four intronic variants. Further investigations were performed to analyze the effect of these variants, alone and/or in combination, on splicing and PTEN protein levels. Despite suggestive evidence emerging from *in silico* analyses, the presence of these intronic variants do not seem to alter RNA splicing or PTEN protein levels. In addition, as loss of PTEN or part of it has been reported, Western blot analysis has also been performed. No major deletion could be identified in our cohort. Therefore, assuming a Poisson distribution for the frequency of deleterious mutation in our cohort, if the frequency of such deleterious mutation was 2%, we would have had a 90% or greater chance of observing at least one such mutation. These results suggest that *PTEN* germline mutations are rare and are unlikely to account for a significant proportion of familial breast cancer cases in the French Canadian population.

**Keywords:** *PTEN/MMAC1/TEP1*; Breast cancer susceptibility; Germline mutations; Non-*BRCA1/2* families; Splicing

## Introduction

In developed countries, breast cancer incidence has been estimated to 7.8% at 80 years old [1]. Mutations in the two most important breast cancer susceptibility genes in high-risk families, *BRCA1* and *BRCA2*, confer an estimated risk of breast cancer by age 70 of 65% and 45%, respectively [2, 3]. A lower proportion of hereditary breast cancers is explained by mutations in *TP53*, *ATM*, *LKB1*, *CHEK2* [4], *BRIP1* [5] and *PTEN*. Indeed, the risk of breast cancer in Cowden disease (CD) associated with mutations in the *PTEN* gene has been estimated to 30-50% by age 70 [6], and both somatic and germline *PTEN* mutations have been identified in breast cancers [7-10]. In non-*BRCA1/BRCA2* familial breast cancer, *PTEN* has been excluded on the basis of linkage analysis [11, 12] and was also examined in few studies [9, 13], which suggested that *PTEN* does not play a major role in breast cancer susceptibility when signs of CD are not present.

Breast cancer is a genetically heterogeneous disease [reviewed in 14]. Therefore, considering that deleterious mutations in *BRCA1/BRCA2* were identified in only 24% of 256 high-risk French Canadian breast/ovarian cancer families [15], evaluating or excluding other breast cancer susceptibility genes such as *PTEN* is very important, especially in the context of finding other breast cancer susceptibility genes. Moreover, although all of the proven breast cancer susceptibility alleles identified to date are rare, specific genetic alterations in a given gene can be more frequent in a relatively recent, small founder population, such as the French Canadian population [15, 16].

Based on the risk of breast cancer in CD associated with mutations in the *PTEN* gene, on the identification of germline *PTEN* mutations in breast cancers and on the high proportion of high-risk French Canadian breast/ovarian families not due to *BRCA1/2* mutations, evaluation of the implication of *PTEN* in these families is imperative. Therefore, in order to evaluate the excess of familial risk resulting from germline mutations in the *PTEN* gene, the complete sequence of the exonic and flanking intronic regions were analyzed in DNA samples from breast cancer cases drawn from non-related *BRCA1/BRCA2*-negative high-risk French Canadian breast/ovarian families.

## Materials and methods

### *Patient samples*

Families participating in the present study were first recruited as part of a research project started in 1996, which thereafter evolved in a large ongoing project designated INHERIT BRCA [15]. More details regarding ascertainment criteria, experimental and clinical procedures, assessment of the *BRCA1/2* status of the subset of 98 high-risk French Canadian breast/ovarian cancer families used in this study as well as the INHERIT BRCA research program have been described previously [15, 17].

### *Experimental procedures and expression analyses*

In order to avoid non-specific amplification of the *PTEN* pseudogene and to avoid reporting mutations reflecting sequence differences between *PTEN* and the pseudogene [18, 19], intronic primers were designed using the Primer Express 2.0 software (Applied Biosystems, Foster, CA, USA). Primers used for DNA amplification and sequencing are listed in Supplemental Table 1. Sequencing procedures and analyses were performed as described previously [17]. Total RNA was extracted from EBV-transformed  $\beta$ -lymphoblastoid cell lines using TRI REAGENT<sup>®</sup> (Molecular Research Center, inc., Cincinnati, OH, USA) according to the manufacturer's instructions.

Reverse transcription of 5  $\mu$ g of total RNA was performed as previously described [17]. A 224 bp fragment covering the junction of exons 1 and 2 of *PTEN* mRNA was amplified from cDNA using the following set of primers: 5'-CCATTTCCATCCTGCAGAAGAA-3', and 5'-GTATACGCCTTCAAGTCTTTCTGCA-3'. Following subcloning, a 356 bp high-specific activity single-stranded RNA probe was synthesized *in vitro*. Ribonuclease protection assay (RPA III<sup>™</sup> Ribonuclease Protection Assay Kit, Ambion®, Austin, TX, USA) was conducted following supplier's protocol. This specific cRNA probe was used to evaluate the effect of c.79+35C>T and c.80-96A>G variants and was hybridized with 20  $\mu$ g of total RNA from a wild-type (WT) individual, one heterozygous individual for the c.79+35C>T variant as well as from one heterozygous and one homozygous individuals for the c.80-96A>G variant. Following digestion and precipitation, samples were run on a 6% polyacrylamide gel. Using the same procedure, a 239 bp fragment covering the junction of exons 8 and 9 was amplified using the following set of primers: 5'-CATACCAGGACCAGAGGAAACCT-3', and 5'-CCTCTGGATTGACGGCTCCT-3',



leading to a 370 bp cRNA probe, and hybridized with 20 µg of total RNA from a WT, one heterozygous and one homozygous individuals for the c.1026+32T>G variant.

### ***Western blot analysis***

Total proteins were extracted by resuspending cells in phosphate buffered saline and cell lysis was performed using standard procedures. Following protein extraction from immortalized cell lines, total protein extracts were quantified using Bio-Rad Protein Assay (Bio-Rad, Hercules, CA, USA). Twenty five µg of total protein extracts from a WT individual, as well as 6 individuals, carriers of the different genotype combinations for c.79+35C>T, c.80-96A>G, c.253+109insTCTTA and c.1026+32T>G, were separated on a 10% denaturing SDS gel and transferred by electroblotting (Bio-Rad, Hercules, CA, USA) onto a nitrocellulose membrane (Hybond™-ECL™ nitrocellulose membrane, Amersham Pharmacia Biotech, Germany) and thereafter hybridized to the anti-PTEN-specific monoclonal mouse IgG antibody 6H2.1 (diluted 1:1000) raised against the last 100 C-terminal amino acids of the PTEN protein (Cascade Bioscience, Winchester, MA, USA). Hybridization procedures were performed as previously described [20]. The PTEN protein band was revealed using enhanced chemiluminescence (Perkin-Elmer, Boston, MA) while the β-actin protein (clone AC-74) was used as an internal control (Sigma, Saint Louis, MO, USA). The same procedure as the one described above was followed for the evaluation of major deletions among our breast cancer cases.

An N-terminal affinity purified goat polyclonal antibody (N-19) raised against the N-terminal part of human PTEN protein (Santa Cruz Biotechnology inc., Santa Cruz, CA, USA) was also used to evaluate the level of PTEN protein. This antibody was diluted 1:200 in membrane blocking reagent and membranes were incubated and washed as described above, and the PTEN protein band was revealed using the same procedure.

### ***Computational analysis***

Alignment of the *PTEN* gene orthologues [*Homo sapiens* (*Hs*), *Pan troglodytes* (*Pt*), *Mus musculus* (*Mm*), *Monodelphis domestica* (*Md*), *Gallus gallus* (*Gg*), *Takifugus rubipes* (*Tr*) and *Dario rerio* (*Dr*)] was performed using the default settings of the ClustalW program [21]. *In silico* analyses were conducted using three different splicing prediction programs to

evaluate the effect of variants identified on the splicing scores, namely GeneSplicer web interface [22], SpliceSiteFinder [23, 24] and NNSPLICE 0.9 [25] programs.

## Results & Discussion

### *Variants identified in the PTEN gene*

The complete analysis of the nine exons and intronic flanking sequences of *PTEN* did not identify any deleterious or missense mutations within the *PTEN* coding region. However, four intronic variants have been detected among our high-risk French Canadian breast cancer cases (Table 1). Of these, three were reported previously (c.80-96A>G, c.253+109insTCTTA and c.1026+32T>G) and showed similar frequencies to those found in the present study (Table 1). The rare c.79+35C>T variant [minor allele frequency (MAF) of 0.005] was observed only once in our series of breast cancer cases, and has also been found in an unaffected first-degree relative of 65 years. Out of the two variants that are reported in HapMap database (c.80-96A>G and c.1026+32T>G), only the latter differs in frequency, this variant being observed at a much higher MAF in the CEPH cohort than in our French Canadian breast cancer cases. This could be attributed to population-specific differences, the CEPH cohort being composed of Northern and Western Europeans. Indeed, consistent with this observation, large population-specific MAF differences are reported in HapMap database for this specific variant. No significant deviation from Hardy-Weinberg equilibrium was observed for any of the four intronic variants identified in our cohort (data not shown).

### *Homology analysis and conservation among PTEN orthologues*

The functionality and possible importance of sequences in which variants were identified were evaluated by aligning the *PTEN* nucleotide sequences of relevant species. As shown in Figure 1, alignment of the respective intronic sequences showed that the c.79+35C>T variant is located in a highly conserved region of 13 nucleotides in *Hs*, *Pt* and *Mm*. However, the other three variants identified in this study (c.80-96A>G, c.253+109insTCTTA, c.1026+32T>G) are located in non-conserved regions. These changes do not represent isolated events in human, since c.80-96A>G, c.253+109insTCTTA and c.1026+32T>G rare alleles are also observed in *Pt*. Most known elements responsible for splicing are usually located in conserved regions near exon boundaries (within 50 bp). Therefore, the non-conservation of variants found in this study may be related to their deeper location in introns or their lower importance on selective pressure. However, intronic enhancers and silencers are needed for proper splicing of

exonic sequences, and these regulatory elements can be near (50-100 bp) or far (hundreds to thousands bp) from a given splice site [26].

### ***Splice site prediction scores***

The function of a protein can be altered by different mechanisms, namely post-translational modifications, alternative splicing or gene expression level modified by epigenetic changes [26, 27]. Therefore, we then assessed the putative effect of the variants identified in this study on mRNA splicing, using three different splicing prediction programs (GeneSplicer, SpliceSiteFinder and NNSPLICE). According to Table 2, *in silico* analyses demonstrate that c.79+35C>T, c.80-96A>G and c.253+109insTCTTA sequence variations could have a weak effect on the strength of different pseudo splice sites, while c.1026+32T>G could create an additional splice site 31bp further from the existing junction of exon 8, which could result in a putative truncated protein lacking two important domains (PEST and PDZ), both being closely involved in protein stability and protein-protein interactions. Since the three splicing prediction programs do not use the same splicing score scale, it is not possible to directly compare between the different splicing scores for a given polymorphism, nor for a difference between splice sites predictions. Interestingly, the GeneSplicer program failed to identify WT splice sites in all four instances. These preliminary analyses (splicing scores and sequence homology) prompted us to investigate the potential effect of the two relevant variants (c.79+35C>T and c.1026+32T>G) on splicing at the RNA and protein levels.

### ***Ribonuclease Protection Assay (RPA) and Western Analysis***

Although some studies have identified polymorphisms or loss of PTEN expression by immunohistochemistry in certain malignancies [7, 18 28-33], the effect of sequence variations potentially leading to an alteration in PTEN expression remains to be established for the variants identified in the present study. Therefore a cRNA probe of 224 bp, covering exons 1 and 2 of PTEN mRNA, was synthesized (Figure 2, panel a). This cRNA probe, hybridized to the only c.79+35C>T carrier identified in our cohort (lane 4) and to carriers of the other variant identified in this region (c.80-96A>G), under both the heterozygous (lane 5) and homozygous (lane 6) states revealed a band of 224 bp, which represents the expected full protected fragment (lane 3). This cRNA fragment is observed in all tested breast cancer high-risk individuals, including the WT subject, confirming the non-effect of

these variants on mRNA splicing. This observation is in agreement with a recent study performed on *PTEN* mutation-positive patients and patients with no mutation but presenting with classic features associated with CD and BRRS [34]. A much weaker additional band of 165 bp, likely corresponding to the protected fragment of exon 1 alone, was also observed for all tested individuals, including the WT sample (Figure 2, panel A). This could be explained either by inefficient splicing, or by the weak expression of ESTs comprising the sequence corresponding to the beginning of intron 1, as reported in UCSC database (AI378794.1).

Similarly, in order to evaluate the effect of c.1026+32T>G on splicing, a cRNA probe of 239 bp covering the junction of exons 8 and 9 was used (Figure 2, panel b). Alternative splicing of exon 8 of *PTEN* mRNA, as predicted by the SpliceSiteFinder program, would theoretically lead to protection of either mRNA sequence of exon eight or nine, resulting in the detection of 190 bp and 49 bp fragments, respectively. No aberrant pattern of digestion for either a WT (lane 3), heterozygous (lane 4) or homozygous (lane 5) individual for this genetic variant was detected, thus ruling out c.1026+32T>G as being responsible for alternative splicing, despite suggestive evidence from *in silico* analysis.

Western blot analysis, using both C-terminal and N-terminal antibodies, was conducted to determine the level of WT *PTEN* protein (using the C-terminal antibody) or to detect the presence of putative aberrant spliced proteins (using the N-terminal antibody). The effect of a given variant, alone or in combination with other variant(s), was evaluated in six selected individuals bearing different genotype combinations. As shown in Figure 2 (panel c), none of the genotype combinations examined seem neither to affect the level of *PTEN* protein, nor to produce any alternative protein product, as indicated when using the N-terminal antibody (data not shown). Moreover, in comparison to the WT genotype (lane 1), the level of proteins expressed by any of the variant-containing individuals seems unaffected, further confirming the absence of effect of the identified variants on protein expression. Using the same C-terminal antibody, Western blot analysis was also performed on all but one -due to material availability- breast cancer individuals in order to evaluate loss or altered level of *PTEN* expression that might occur following major deletion in the *PTEN* gene. No such deletion seems to be present in our cohort (data not shown).

Among the series of 98 breast cancer cases selected from high-risk non-BRCA1/2 French Canadian families, no deleterious mutation or amino acid change was identified in the coding region of the *PTEN* gene. However, we found four intronic variants, two of which potentially affecting splice sites. The data presented here clearly demonstrate that none of the intronic variants identified in this study leads to any alternative splicing or significant modification of PTEN protein expression. Therefore, we can conclude that germline mutations in the *PTEN* gene, if present, are rare and do not contribute substantially to breast cancer susceptibility in high-risk French Canadian breast cancer families. Indeed, assuming a Poisson distribution of the probability of finding a deleterious mutation in our cohort, if the frequency of such deleterious mutation was 2%, we would have had 90% or greater chance of observing at least one such mutation. Therefore, these high-risk families are yet to be explained by other gene(s).

**Abbreviations**

BRRS	Bannayan-Riley-Ruvalcaba Syndrome
CD	Cowden disease
<i>Dr</i>	<i>Dario rerio</i>
EBV	Epstein-Barr virus
EST	Expressed sequence tag
<i>Gg</i>	<i>Gallus gallus</i>
<i>Hs</i>	<i>Homo sapiens</i>
MAF	Minor allele frequency
<i>Md</i>	<i>Monodelphis domestica</i>
<i>Mm</i>	<i>Mus musculus</i>
PDZ	PSD-95/Dlg/ZO-1 domain
PEST	proline, glutamic acid, serine, and threonine domain
<i>Pt</i>	<i>Pan troglodytes</i>
RPA	Ribonuclease protection assay
<i>Tr</i>	<i>Takifugus rubipes</i>
WT	Wild type

**Conflict of interest statement**

The authors declare that they have no competing interests.

## **Acknowledgements**

The authors are indebted to the participants and their families for their generosity and providing DNA samples. We would like to thank Dr Martine Dumont, Gilles Leblanc, Carolle Samson and Martine Tranchant for sample management, mutation screening, and skillful technical assistance as well as Claire Brousseau, Marie-Andrée Lajoie, Pascale Léger, Hélène Malouin and Josée Rhéaume, for genetic counselling and clinical data management at the Cancer Genomics Laboratory. We thank Anne-Marie Moisan and Lucie Larouche for MLPA analyses. A special thank to Mélanie Houde for skillful laboratory assistance. We also appreciate advice received from ethics committees. This work was supported by the Canadian Institutes of Health Research (CIHR) and Institute of Cancer and Institute of Gender and Health for the INHERIT BRCA<sub>s</sub> research program, the Fonds de la Recherche en Santé du Québec (FRSQ)/Réseau de Médecine Génétique Appliquée (RMGA), the Canadian Breast Cancer Research Alliance and the CURE Foundation. F.G. is recipient of a studentship from Fondation René Bussières, and F.D. is a recipient of a Research Career Award in the Health Sciences from CIHR/Rx&D Health Research Foundation.

## **Appendix 1**

Peter Bridge, Molecular Diagnostic Laboratory, Alberta Children's Hospital, Calgary, Canada.

Yann Joly, Centre de recherche en droit public, Montreal University, Montreal, Canada.

Bartha Maria Knopper, Chaire de recherche en droit et médecine, Centre de recherche en droit public, Montreal University, Montreal, Canada.

Jacques Simard, Canada Research Chair in Oncogenetics, Cancer Genomics Laboratory, Oncology and Molecular Endocrinology Research Centre, CHUQ, Pavillon CHUL and Laval University, Québec, Canada.



## References

1. Collaborative Group on Hormonal Factors in Breast Cancer (2001) Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet* 358(9291):1389-99
2. Ford D, Easton DF, Stratton M, et al. (1998) Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet* 62(3):676-89
3. Antoniou A, Pharoah PD, Narod S, et al (2003) 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* 72(5):1117-30
4. CHEK2 Breast Cancer Case-Control Consortium. (2004) CHEK2\*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. *Am J Hum Genet.* 74(6):1175-82
5. Seal S, Thompson D, Renwick A, et al (2006) Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nat Genet* 38(11):1239-41
6. Ball S, Arolker M, Purushotham AD (2001) Breast cancer, Cowden disease and PTEN-MATCHS syndrome. *Eur J Surg Oncol* 27(6):604-6
7. Lynch ED, Ostermeyer EA, Lee MK, et al (1997) Inherited mutations in PTEN that are associated with breast cancer, cowden disease, and juvenile polyposis. *Am J Hum Genet* 61(6):1254-60
8. Mercapide J, Zhang SY, Fan X, et al (2002) CCND1- and ERBB2-gene deregulation and PTEN mutation analyses in invasive lobular carcinoma of the breast. *Mol Carcinog* 35(1):6-12
9. Figer A, Kaplan A, Frydman M, et al (2002) Germline mutations in the PTEN gene in Israeli patients with Bannayan-Riley-Ruvalcaba syndrome and women with familial breast cancer. *Clin Genet* 62(4):298-302
10. Kurose K, Gilley K, Matsumoto S, et al (2002) Frequent somatic mutations in PTEN and TP53 are mutually exclusive in the stroma of breast carcinomas. *Nat Genet* 32(3):355-7.
11. Shugart YY, Cour C, Renard H, et al (1999) Linkage analysis of 56 multiplex families excludes the Cowden disease gene PTEN as a major contributor to familial breast cancer. *J Med Genet* 36(9):720-1
12. Haiman CA, Stram DO, Cheng I, et al (2006) Common genetic variation at PTEN and risk of sporadic breast and prostate cancer. *Cancer Epidemiol Biomarkers Prev* 15(5):1021-5
13. Carroll BT, Couch FJ, Rebbeck TR et al (1999) Polymorphisms in PTEN in breast cancer families. *J Med Genet* 36(2):94-6. Erratum in: *J Med Genet* 2000, 37(7):559
14. Durocher F, Guénard F, Desjardins S, et al (2005) Inherited susceptibility to breast cancer: accomplishments and challenges. In Sinnett D (eds) *Molecular Genetics of Cancer*. Research Signpost, Kerala
15. Simard J, Dumont M, Moisan AM, et al (2006) Evaluation of BRCA1 and BRCA2 mutation prevalence, risk prediction models and a multi-step testing approach in French-Canadian high-risk breast and ovarian cancer families. *J Med Genet* Aug 11; [Epub ahead of print]

16. Vezina H, Durocher F, Dumont M, et al (2005) Molecular and genealogical characterization of the R1443X BRCA1 mutation in high-risk French-Canadian breast/ovarian cancer families. *Hum Genet* 117(2-3):119-32
17. Durocher F, Labrie Y, Soucy P et al (2006) Mutation Analysis and Characterization of ATR Sequence Variants in Breast Cancer Cases from High-Risk French Canadian Breast/Ovarian Cancer Families. *BMC Cancer* 6:230
18. Rhei E, Kang L, Bogomolny F et al (1997) Mutation analysis of the putative tumor suppressor gene PTEN/MMAC1 in primary breast carcinomas. *Cancer Res* 57(17):3657-9
19. Teng DH, Hu R, Lin H, et al (1997) MMAC1/PTEN mutations in primary tumor specimens and tumor cell lines. *Cancer Res* 57(23):5221-5
20. Weng LP, Smith WM, Dahia PL, et al (1999) PTEN suppresses breast cancer cell growth by phosphatase activity-dependent G1 arrest followed by cell death. *Cancer Res* 59(22):5808-14
21. Chenna R, Sugawara H, Koike T, et al. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* 31(13):3497-500
22. Pertea M, Lin X, Salzberg SL. (2001) GeneSplicer: a new computational method for splice site prediction. *Nucleic Acids Res* 29(5):1185-90
23. Shapiro MB, Senapathy P (1987) RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res* 15(17):7155-74
24. Senapathy P, Shapiro MB, Harris NL (1990) Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome project. *Methods Enzymol* 183:252-78
25. Reese MG, Eeckman FH, Kulp D, et al (1997) Improved splice site detection in Genie. *J Comput Biol* 4(3):311-23
26. Pagani F, Baralle FE. (2004) Genomic variants in exons and introns: identifying the splicing spoilers. *Nat Rev Genet* 5(5):389-96.
27. Lim LP, Burge CB (2001) A computational analysis of sequence features involved in recognition of short introns. *Proc Natl Acad Sci U S A* 98(20):11193-8
28. Perren A, Weng LP, Boag AH et al (1999) Immunohistochemical evidence of loss of PTEN expression in primary ductal adenocarcinomas of the breast. *Am J Pathol* 155(4):1253-60
29. Depowski PL, Rosenthal SI, Ross JS (2001) Loss of expression of the PTEN gene protein product is associated with poor outcome in breast cancer. *Mod Pathol* 14(7):672-6
30. Bose S, Wang SI, Terry MB et al (1998) Allelic loss of chromosome 10q23 is associated with tumor progression in breast carcinomas. *Oncogene* 17(1):123-7
31. Chen ST, Yu SY, Tsai M, et al. (1999) Mutation analysis of the putative tumor suppression gene PTEN/MMAC1 in sporadic breast cancer. *Breast Cancer Res Treat* 55(1):85-9
32. Olschwang S, Serova-Sinilnikova OM, Lenoir GM et al (1998) PTEN germ-line mutations in juvenile polyposis coli. *Nat Genet* 18(1):12-4
33. Nassif NT, Lobo GP, Wu X, et al (2004) PTEN mutations are common in sporadic microsatellite stable colorectal cancer. *Oncogene* 23(2):617-28

34. Agrawal S, Pilarski R, Eng C (2005) Different splicing defects lead to differential effects downstream of the lipid and protein phosphatase activities of PTEN. *Hum Mol Genet* 14(16):2459-68
35. Hamilton JA, Stewart LM, Ajayi L, et al (2000) The expression profile for the tumour suppressor gene PTEN and associated polymorphic markers. *Br J Cancer* 82(10):1671-6
36. Kong D, Suzuki A, Zou TT, et al (1997) PTEN1 is frequently mutated in primary endometrial carcinomas. *Nat Genet* 17(2):143-4
37. Sakurada A, Suzuki A, Sato M, et al (1997) Infrequent genetic alterations of the PTEN/MMAC1 gene in Japanese patients with primary cancers of the breast, lung, pancreas, kidney, and ovary. *Jpn J Cancer Res* 88(11):1025-8

### Legends to Figures

**Figure 1. Homology analysis of *PTEN* gene between relevant species.** The location of identified variants is indicated by an arrow above the nucleotide sequences while the high homology region for the c.79+35C>T variant is shaded. The location of the c.253+109insTCTTA variant is indicated by an arrow and a parenthesis above the nucleotide sequence. *Hs*: *Homo sapiens*; *Pt*: *Pan troglodytes*; *Ms*: *Mus musculus*; *Md*: *Monodelphis domestica*; *Gg*: *Gallus gallus*; *Tr*: *Takifugus rubripes*; *Dr*: *Dario rerio*.

**Figure 2. RPA and Western analyses of identified variants in carriers. Panel a.** Schematic representation and localization of exons 1-2 cRNA probe for c.79+35C>T and c.80-96A>G polymorphisms carriers. Lane 1, undigested cRNA probe; lane 2, tRNA (negative control); lane 3, WT individual; lane 4, heterozygous c.79+35C>T carrier; lane 5, heterozygous c.80-96A>G carrier; lane 6, homozygous c.80-96A>G carrier. **Panel b.** Schematic representation and localization of exons 8-9 cRNA probe for the c.1026+32T>G polymorphism carriers. Lane 1, undigested cRNA probe; lane 2, tRNA (negative control); lane 3, WT individual; lane 4, heterozygous c.1026+32T>G carrier; lane 5, homozygous c.1026+32T>G carrier. **Panel c.** Western blot analysis of samples representing different alleles variant combinations using a C-terminal PTEN antibody (6H2.1). Order of genotypes is c.79+35C>T; c.80-96A>G; c.253+109insTCTTA; c.1026+32T>G. Lane 1 : wt;wt;wt;wt. Lane 2 : het;wt;wt;wt. Lane 3 : wt;hom;wt;hom. Lane 4 : wt;het;hom;het. Lane 5 : wt;wt;het;wt. Lane 6 : wt;het;wt;het. Lane 7 : wt;hom;wt;wt. Abbreviations: wt : wild type alleles, het : heterozygous alleles, hom : homozygote alleles, MW : molecular weight, kDa : kiloDaltons.

Table 1

Observed intronic sequence variants and genotype frequencies in familial breast cancer cases

SNP Designation		Location	Genotype Frequency			Minor Allele Frequency (MAF)		
SNP ID <sup>a</sup>	dbSNP ID		Common homozygote (expected) <sup>b</sup>	Heterozygote (expected) <sup>b</sup>	Rare homozygote (expected) <sup>b</sup>	French Canadian	HapMap CEU	Other studies
c.79+35C>T	N/A	Intron 1	0.99 (0.99)	0.01 (0.01)	0.00 (0.00)	0.005	N/A	N/A
c.80-96A>G	rs1903858	Intron 1	0.52 (0.48)	0.35 (0.42)	0.13 (0.09)	0.306	0.350	0.360 [32] 0.400 <sup>c</sup> [35]
c.253+109 insTCTTA	rs3830675	Intron 4	0.51 (0.48)	0.37 (0.42)	0.12 (0.09)	0.306	N/A	0.370 [13] 0.400 [32] 0.640 <sup>c</sup> [35] 0.540 [36] 0.530 [37]
c.1026+32T>G	rs555895	Intron 8	0.51 (0.49)	0.38 (0.42)	0.11 (0.09)	0.301	0.650	0.370 <sup>c</sup> [35]

<sup>a</sup> According to the nomenclature of the Human Genome Variation Society<sup>b</sup> As expected under Hardy-Weinberg equilibrium<sup>c</sup> Heterozygosity value

N/A: Information not available for this SNP (not reported in dbSNP)

Table 2

Splicing scores observed for the WT and the putative splicing sites in the PTEN gene

Variants	Program	WT splice site	Score	Alternative Splice site affected by variant	Distance from WT junction (bp)	Score change WT / variant
c.79+35C>T	GeneSplicer	DS1	ND	DS1	0	NA
				pDS1	+64	4.45/4.74
	SpliceSiteFinder	DS1	47.6	DS1	0	47.6/47.6
				pDS1	+36	54.8/55.1
	NNSPLICE	DS1	ND	DS1	0	NA
c.80-96A>G	GeneSplicer	AS2	ND	AS2	0	NA
				pAS2	+29	1.75/1.73
	SpliceSiteFinder	AS2	ND	AS2	0	NA
				NNSPLICE	AS2	0
				pAS2	-111	0.89/0.85
c.253+109 insTCTTA	GeneSplicer	DS4	ND	DS4	0	NA
				pAS5	+367	3.73/4.08
	SpliceSiteFinder	DS4	89.2	DS4	0	89.2/89.2
				NNSPLICE	DS4	0.98
c.1026+32T>G	GeneSplicer	DS8	ND	DS8	0	NA
				pDS8	+229	1.83/2.18
	SpliceSiteFinder	DS8	90.0	DS8	0	90.0/90.0
				pDS8a	+27	56.9/71.4
	NNSPLICE	DS8	0.90	pDS8b	+31	---/74.6
				DS8	0	0.90/0.90
				pDS8	+38	0.32/0.17

AS: Acceptor Site

DS: Donor Site

pAS: putative Acceptor Site

pDS: putative Donor Site

NA: Information not Available

ND: Splicing site not Detected

Figure 1

**c.79+35C>T**

↓

<i>Hs</i>	TTT-CTGCG-GCTGCTCCTCTTTACCTTTCTGTCACTCTC- TTAGAACGTGGGAGTAGACG
<i>Pt</i>	TTT-C-GCG-GCTGCTC---TTTACCTTTCTGTCACTCTC- TTAGAACGTGGGAGTAGACG
<i>Mm</i>	TTT-CTGCG-GCTGTTCTCCTCTTTGCTTTTCTGTCACTCTG-AT--AACGTGGGAGTAGACG
<i>Md</i>	TCC-TGGTC-GCCACCCCTCCCC-CCGGGCCGGGGTCACG-GCCGCCCCCGCCCGCTGCC
<i>Gg</i>	CTTCCTTTGTAAAAAAGGAAACTCTTATTTTATTCTGCTT-TCAGCTTTAAGGCAAAGATC
<i>Tr</i>	TGTGTAGCA-GAGTCCCCCCCCCCCCCTCGCTCGTCCTC-TCCGCTCACACTTGGCTGTC
<i>Dr</i>	CCAGTGCCCCGACTGACACATGGTGGCCTTCCAGCAGCATGAGGAGAGTTTCACAGTAGGTA

**c.80-96A>G**

↓

<i>Hs</i>	TTATTA--CTCCA-GCTATAGTGGGGAAAACCTTCTTTTCATAACTAGCTAATGTTTTTAAA
<i>Pt</i>	TTATTA--CTCCA-GCTATAGTGGGGAAAAGCTTCTTTTCATAACTAGCTAATGTTTTTAAA
<i>Mm</i>	TCCCTAGGCTCCATGCCCATCTAGCTACAGGCTTTTGGCCTCATTTTGACAATGTT---AG
<i>Md</i>	GCCTTCAATGTTGGTCATTTCTTAATAGAACAGTAATATTCCATCCTATTTGTGTCCATT
<i>Gg</i>	TGCCATTTCTACAAGTCAGGCCGCACAGCCTTTTCTAAGCAGGACATAAAAAGGTGCAGGA
<i>Tr</i>	ACTTTTCACCTAACTTCAAGGTAAAAAACAACAAAAAATCATGAGCTGGGGGAAAA
<i>Dr</i>	AGATCGGTGCTCAGAGCCCCGCTCGCACTCAGCTCGCGCTCCGTCTAGCCTCAACGGC-TA

**c.253+109insTCTTA**

↓

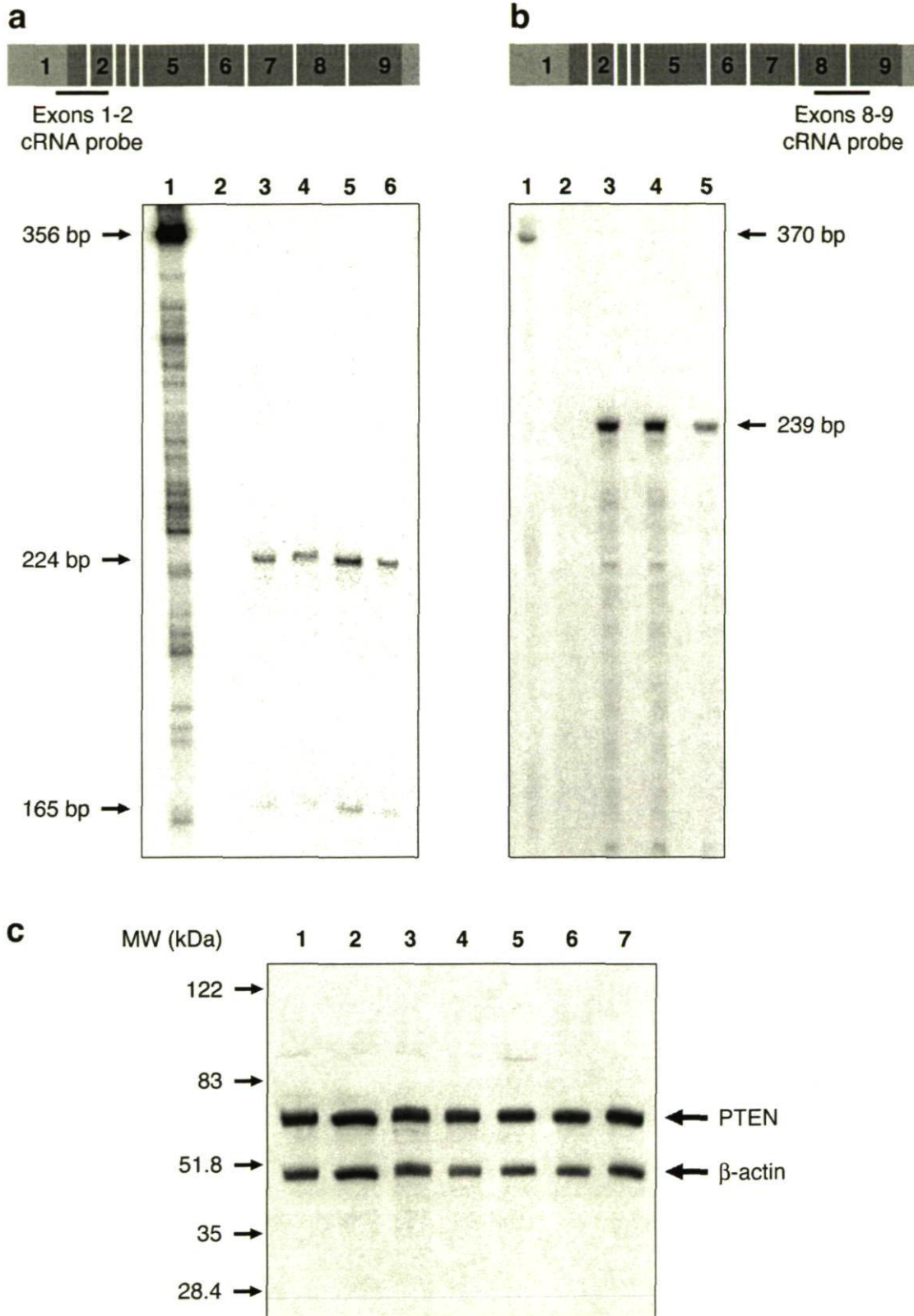
<i>Hs</i>	TATCGAGTGA-----GATACATATTTA-----AGAATTA--TCTTTAAAAATTTCAAAAAT
<i>Pt</i>	TATCGAGTGA-----GATACATATTTATCTTAAGAATTA--TCTTTAAAAATTTCAAAAAT
<i>Mm</i>	TCACAAAAGGTTTTGGAAACAATTTTATGAAGAAAGTCGG-TCACTGGAGAACACATGCAG
<i>Md</i>	TCCCAGCCTGTGC--CTTTTGCCTCGACTGCCCTCCTTACTTCCCTGACCACGGCTTCCTC
<i>Gg</i>	TAGTGTCTCCTTACCACCTCAGCTGCTTAAAAATTACAATTAAGCAAAAGACTTCTCCCAG
<i>Tr</i>	TAAAGAGTGG-----GATGCGTTTTAAGGGCGATCGGCTG--TCTCTCTTCACTGTCTTTCCG
<i>Dr</i>	TAACAAACAAAATAAAACCTAAATTA AAAAAGATTTTTTG-TGTTAATACATGATAAAAAT

**c.1026+32T>G**

↓

<i>Hs</i>	GTAAATTAACATTTTGTGG-GGGTT-GTTGACTTGTATGTATGTGATGTGTGTTAATT
<i>Pt</i>	GTAAATTAACATTTTGTGG-GGGTT-GGTGACTTGTATGTATGTGATGTGTGTTAATT
<i>Mm</i>	GTAAA---ACCATTTGGGGTGGGAG-GGTGTGTTTTATTCTGATTGTGAA-GCTAAGAG
<i>Md</i>	AGAGATACTAATAGAATTCCTTGGCTC-CTCCCTTAGCTTAGTAAAAGTTGATGAAAAGTG
<i>Gg</i>	GAAAACCAGGACCTGTTCTTCAACATCCGCTGACATTACAAAATCACTTTTACAAAATAAAA
<i>Tr</i>	AAAAAA--AAACAAAAA--ACATG-AGCTGGGGGAAAAAAGAGACTGATCGCTAATTA
<i>Dr</i>	AAAGAAAACCTGCACTGAATGAAGCAACCATTGTACCCTACACTAGCCACA-ACAGATATGG

Figure 2





**Supplemental Table 1****Oligonucleotide primers used for amplification of the *PTEN* gene.**

Exon	Forward primer	Reverse primer	Annealing temperature (°C)	PCR product length (bp)
1	5'-CCAGTCGCTGCAACCATCCA-3' <sup>a</sup>	5'-CTAGCCAAACACACCCTAGGTTGA-3'	60	451
2	5'-CTTGTCTTTTCAGGCAGGTGTC-3' <sup>a</sup>	5'-CCCTGAAGTCCATTAGGTACGGT-3'	60	574
3	5'-CCCATAGAAGGGGTATTTGTTGGA-3'	5'-CAAGTAACATGAATCTGTGCCAACA-3' <sup>a</sup>	60	428
4	5'-GTGGCATCACAAAGTTTTTAAGCAA-3' <sup>a</sup>	5'-CCTCCTCATCCATCAGACTCCT-3'	60	723
5	5'-GTCCGTATAGCGTAAATCCCAGA-3'	5'-GGGAGGAACACAAGATGAACAAGA-3' <sup>a</sup>	60	871
6	5'-GGTCAGAGCGCTGTTGTGACCT-3' <sup>a</sup>	5'-GGCTTCTTAGCCCAATGAGTTGA-3'	60	588
7	5'-TCTTTGCCGCATGCTAGAGACT-3' <sup>a</sup>	5'-CAATGCCATAAGGCCTTTTCCT-3'	60	864
8	5'-TCATGTGAATGAAAATGCAACAGA-3'	5'-CCAGATTCAGCAAACAACGAA-3' <sup>a</sup>	60	673
9.1 <sup>b</sup>	5'-GGCCTCTTAAAGATCATGTTTGTACA-3' <sup>a</sup>	5'-TCGGAAACCTCTCTTAGCCAACT-3' <sup>a</sup>	60	798
9.2 <sup>b</sup>	5'-GGCACTTCCCGTTTTATTC-3' <sup>a</sup>	5'-TGGAACGGCTGACAGCTATTGA-3' <sup>a</sup>	60	908

<sup>a</sup> Used as sequencing primers<sup>b</sup> Exon 9 has been amplified in two PCR fragments.

## CHAPITRE III

### Mutational analysis of the breast cancer susceptibility gene *BRIP1/BACH1/FANCF* in high-risk non-*BRCA1/BRCA2* Breast Cancer Families

#### 1. Résumé

Le gène *BRIP1* encode une hélicase interagissant avec *BRCA1* et contribuant à sa fonction de réparation de l'ADN. Des mutations germinales de *BRIP1/FANCF* ont été identifiées chez des patientes atteintes du cancer de sein et chez certains patients atteints de l'anémie de Fanconi (AF). Nous avons cherché à évaluer la contribution des altérations du gène *FANCF* dans la susceptibilité au cancer du sein dans notre cohorte d'individus provenant de familles canadiennes-françaises. Aucune mutation délétère, de délétion d'exon ou de rétention de séquences introniques n'a pu être identifiée. Notre analyse a révélé 42 variants, dont 24 variants introniques, 13 exoniques et cinq localisés dans la région promotrice. Une analyse des facteurs de transcription a identifié un effet potentiel des variants du promoteur sur l'expression, ce qui a été vérifié à l'aide d'essais de gène rapporteur. Bien que la présente étude n'implique pas *FANCF* à titre de gène de susceptibilité au cancer du sein à forte pénétrance chez les familles canadiennes-françaises, il n'est pas exclu que les variants faux-sens de *FANCF* puissent être associés avec des phénotypes de l'AF, tel les cancers chez l'enfant.

Guénard F, Labrie Y, Ouellette G et al. Mutational analysis of the breast cancer susceptibility gene *BRIP1/BACH1/FANCF* in high-risk non-*BRCA1/BRCA2* breast cancer families. *J Hum Genet.* 2008;53(7):579-91.

#### 2. Manuscrit

**Mutational analysis of the breast cancer susceptibility gene  
*BRIP1/BACH1/FANCI* in high-risk non-*BRCA1/BRCA2* Breast Cancer  
Families**

Frédéric Guénard<sup>1</sup>, Yvan Labrie<sup>1</sup>, Geneviève Ouellette<sup>1</sup>, Charles Joly Beauparlant<sup>1</sup>, Jacques Simard<sup>1,2</sup>, INHERIT BRCA<sup>3</sup>, Francine Durocher<sup>1\*</sup>

<sup>1</sup>Cancer Genomics Laboratory, Oncology and Molecular Endocrinology Research Centre, Centre Hospitalier Universitaire de Québec and Laval University, Québec, G1V 4G2, Canada;

<sup>2</sup>Canada Research Chair in Oncogenetics

<sup>3</sup>Other members of INHERIT BRCA<sup>3</sup> involved clinical aspects of this study are listed in Appendix

\*Corresponding author:

Dr Francine Durocher

Cancer Genomics Laboratory, Oncology and Molecular Endocrinology Research Centre,  
Centre Hospitalier Universitaire de Québec and Laval University

2705 Laurier Boulevard

Québec city, Québec,

Canada, G1V 4G2

Tel: (418)-654-2296

Fax: (418)-654-2278

e.mail: [francine.durocher@crchul.ulaval.ca](mailto:francine.durocher@crchul.ulaval.ca)

**Abstract**

The *BRIP1* gene encodes a helicase interacting with BRCA1, which contributes to BRCA1-associated DNA repair function. Germline *BRIP1* mutations affecting the helicase domain activity have been identified in early onset breast cancer patients. In addition, *BRIP1* was recently identified as deficient in Fanconi Anemia (FA) complementation group J. Given the growing number of evidence now linking BRCA1, BRCA2, and the FA pathway, as well as the involvement of FA proteins (BRCA2/FANCD1 and PALB2/FANCN) in breast cancer susceptibility, we sought to evaluate the contribution of *FANCF* gene alterations regarding breast cancer susceptibility among our cohort of 96 breast cancer individuals from high-risk non-*BRCA1/2* French Canadian families. No deleterious mutation, exon deletion or retention of intronic portions could be identified. However, extensive analysis of the promoter and whole exonic and flanking intronic regions of *FANCF* led to the identification of 42 variants, including 22 novel variants not previously reported, four of which located in the promoter region. Transcription factors analysis revealed a potential involvement of *FANCF* promoter variants in regulation of *FANCF* expression, and reporter gene assays were performed. The allelic frequency was assessed in a cohort of 73 unaffected French Canadian individuals, and haplotype analysis and tagging SNPs identification were also performed. Although our current study unlikely involves *FANCF* as a high-risk predisposition gene in non-*BRCA1/2* high-risk French Canadian families, the possible association of *FANCF* missense variants with phenotypes associated with FA, such as childhood cancer, cannot be excluded.

*Keywords:* *BRIP1/FANCF* gene, Breast cancer, Sequence analysis, Haplotype, Tagging SNP

## Introduction

In developed countries, breast cancer incidence has been estimated to 7.8% at 80 years of age (Collaborative Group on Hormonal Factors in Breast Cancer 2001). Mutations in the two most important breast cancer susceptibility genes, *BRCA1* and *BRCA2*, account for only 25% of families with hereditary breast cancer (Easton 1999; Ponder 2001). A lower proportion of hereditary breast cancers is explained by mutations in *TP53* (Lalloo et al. 2003; Gasco et al. 2003), *PTEN* (Ueda et al. 1998; Ball et al. 2001; Guenard et al. 2007), *STK11* (Giardiello et al. 2000), *ATM* (Stankovic et al. 1998; Chenevix-Trench et al. 2002; Szabo et al. 2004; Renwick et al. 2006), *CHEK2* (CHEK2 Breast Cancer Case-Control Consortium 2004) and more recently, specific *FANC* genes have also been associated with breast cancer susceptibility, namely *FANCD1* subsequently identified as *BRCA2*, *FANCN/PALB2* and *FANCI/BRIP1* (Durocher et al. 2005; Seal et al. 2006; Rahman et al. 2007; Erkkö et al. 2007; Tischkowitz et al. 2007).

*FANCI*, also called *BRIP1/BACH1*, has been identified as deficient in Fanconi Anemia (FA) complementation group J (Levitus et al. 2005). Many of the proteins encoded by FA genes, including the more recently identified *FANCI* (Smogorzewska et al. 2007), interact directly with each other to form a multi-subunit nuclear complex and are involved in DNA repair processes (Medhurst et al. 2001). An increased risk to develop cancer has been reported for homozygous FA patients (Alter et al. 2003).

The *BRIP1/BACH1/FANCI* gene is located at 17q22-q24 and encodes a helicase in which the C-terminal part is known to interact with *BRCA1* via its *BRCA1* C-terminal (BRCT) repeats (Cantor et al. 2001). The corresponding protein contributes to *BRCA1*-associated DNA repair function, and mutation of an essential catalytic residue (Pro47) found in most members of the DEAH family interferes with normal double-strand break (DSB) repair in a manner that was dependent on *FANCI* binding to *BRCA1* (Cantor et al. 2001). In addition, phosphorylation of the Ser990 residue triggers *FANCI* interaction with BRCT domains of *BRCA1*, which is required for establishment of the G2 cell cycle checkpoint response to DNA damage (Yu et al. 2003). Proper localization of *FANCI* to the nucleus is impaired by the c.517C>T variant (p.Arg173Cys) (Lei et al. 2003) which has been suggested to have an effect on breast cancer susceptibility. Furthermore, enzymatic activity of *FANCI* was found defective in two patients with germline *FANCI* coding sequence mutations who

experienced early-onset breast cancer (Cantor et al. 2004). Germline mutations affecting domain activity or mRNA expression were also identified in early-onset breast cancer patients thus suggesting an implication for *FANCI* in breast cancer susceptibility (Cantor et al. 2001; Cantor et al. 2004). Given the close relationship between *BRCA1* and *FANCI* and the observation that other helicases belonging to the RecQ family are implicated in cancer-predisposing syndromes, namely the Bloom (*BLM* gene), the Werner (*WRN* gene), and the Rothmund-Thomson syndromes (*RecQL4* gene) (Ellis et al. 1995; Yu et al. 1996; Kitao et al. 1998; Kitao et al. 1999), this reinforces the rational basis for a potential implication of *FANCI* in breast cancer predisposition.

In recent years, a number of studies have been conducted to evaluate the possible implication of *FANCI* gene with regard to breast cancer susceptibility (Rutter et al. 2003; Sigurdson et al. 2004; Lewis et al. 2005; Garcia-Closas et al. 2006; Vahteristo et al. 2006; Seal et al. 2006; Song et al. 2007; Frank et al. 2007). However, the great majority was performed either on specific variants (such as P919S and -64G>A) (Sigurdson et al. 2004; Frank et al. 2007), on breast cancer cases not selected on the basis of a family history (Garcia-Closas et al. 2006; Song et al. 2007), or on a limited number of cases from *BRCA1/2*-negative breast cancer families (Rutter et al. 2003; Vahteristo et al. 2006). Recently published results have clearly shown that inactivating truncating mutations of *FANCI* confer susceptibility to breast cancer in high-risk non-*BRCA1/BRCA2* families (Seal et al. 2006).

In this regard, given that deleterious mutations in *BRCA1/BRCA2* were identified in only 24% of 256 high-risk French Canadian breast/ovarian cancer families (Simard et al. 2007), and that germline mutations in other genes are rare and do not seem to contribute substantially to breast cancer susceptibility in high-risk French Canadian breast cancer families (Durocher et al. 2006, 2007; Guenard et al. 2007; Desjardins et al. 2008), evaluation of the *FANCI* contribution to breast cancer in these families is imperative. Therefore, the complete coding, promoter and flanking intronic regions of *FANCI* were analyzed in 96 breast cancer cases drawn from nonrelated *BRCA1/BRCA2*-negative French Canadian families with high risk of breast cancer.

## **Material and Methods**

**Ascertainment of families:** Recruitment of high-risk French Canadian breast and/or ovarian cancer families from Canada was part of a large ongoing interdisciplinary research program INterdisciplinary HEalth Research International Team on BREast CANcer (INHERIT BRCA). More details regarding ascertainment criteria, experimental procedures as well as the INHERIT BRCA research program have been described elsewhere (Vezina et al. 2005; Antoniou et al. 2006; Simard et al. 2007). Briefly, all patients were referred by physicians from the Québec province, and they were also responsible for *BRCA1/2* test results disclosure to participants. Ethics committees approved the study and patients signed an informed consent. Patients were screened for mutations and large genomic rearrangements of the *BRCA1* and *BRCA2* genes (Moisan et al. 2006). Subsequently, another component was designed for the “localization and identification of new breast cancer susceptibility loci/genes.” Ethics approval for this latter study was also obtained from the different institutions participating in this research project and each participant knowing their inconclusive *BRCA1/2* test results status had to sign a specific informed consent for their participation in this component. All participants had to be at least 18 years of age and mentally capable. A subset of 96 high-risk French Canadian breast/ovarian cancer families were recruited in the present study according to the ascertainment criteria described previously (Durocher et al. 2006, 2007; Guenard et al. 2007; Desjardins et al. 2007). The mean age at diagnosis of these 96 individuals affected with breast cancer was 48.5 years old (32-74). Genomic DNA extraction of the 96 French Canadian breast cancer cases as well as 73 healthy unrelated French Canadian individuals has been performed as previously described (Durocher et al. 2006, 2007; Guenard et al. 2007; Desjardins et al. 2007). These control individual samples were recruited on a non-nominative basis, as part of a long-term study aiming at the characterization of the genetic variability in human populations approved by Institutional Ethics Review Board.

**DNA/RNA isolation from immortalized cell lines and cDNA synthesis:** Lymphocytes were isolated and immortalized from blood samples using the Epstein-Barr Virus (EBV) in 15% RPMI media as previously described (Boukamp et al. 1988, 1990). Total RNA was extracted from EBV-transformed  $\beta$ -lymphoblastoid cell lines using TRI REAGENT® (Molecular Research Center inc., Cincinnati, OH, USA) according to the manufacturer’s

instructions. Following RNA extraction, reverse transcription of 5 µg of RNA was performed as previously described (Durocher et al. 2006). Genomic DNA from 96 high-risk French Canadian breast cancer individuals and 73 controls was extracted from peripheral blood using standard methods either phenol-chloroform, Gentra kits (Minneapolis, MN, USA) or the QIAamp DNA Blood MAXI Kit (QIAGEN inc., Mississauga, ON, Canada) following supplier instructions while genomic DNA from immortalized cell lines was extracted using the CTAB method (Sambrook et al. 2001).

**PCR amplification, sequence analysis and variant characterization:** Amplification and direct sequencing of complementary DNA (cDNA) and genomic DNA of the *FANCF* gene were conducted using primers selected with the Primer Express 2.0 software (Applied Biosystems, Foster City, CA, USA). These primers, listed in Supplemental Table 1, were chosen in order to amplify exonic, promoter and flanking intronic regions of the *FANCF* gene. A fluorescence-based direct sequencing method using BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) was performed to analyze both cDNA and genomic sequences. Sequence reactions were run on an Applied Biosystems 3730xl DNA Analyzer and sequence data were analyzed using the Staden Pregap4 and Gap4 programs (Bonfield et al. 1998).

Each SNP was tested for departure from Hardy-Weinberg equilibrium (HWE) by means of a Chi<sup>2</sup> test. All p-values were two-sided with 1 degree of freedom. Allelic distribution in both series was tested using a Chi<sup>2</sup> test. P-values less than 0.05 were considered as significant.

**Computational analysis:** In order to identify potential effects of variants located in the promoter region, transcription factor search was performed using the program MatInspector from the GenomatixSuite (Quandt et al. 1995; Cartharius et al. 2005).

The ESE finder web-based program (Cartegni et al. 2003) was used to determine the potential effect of variants found in the coding region of the *FANCF* gene on pre-mRNA splicing. All tests were run under default threshold values. The NNSPLICE 0.9 program (Reese et al. 1997) was used to evaluate the potential effect of intronic variants identified in this study on the strength of potential splice sites leading to the splicing of pre-messenger RNA, while the SIFT and PolyPhen web-based softwares were used to predict the effect of



amino acid substitution on protein structure.

**Reporter gene assay:** *FANCI* promoter constructs were amplified from genomic DNA using either a 1068 bp (long) or a 633 bp (short) fragment with primers introducing 5' *KpnI* and 3' *XhoI* sites (Supplemental Table 2), and these amplicons were thereafter sub-cloned into pGL3 basic vector (Promega, Madison, WI). Variant constructs were produced by site-directed mutagenesis (QuikChange Site-Directed Mutagenesis Kit; Stratagene, La Jolla, CA) or obtained by PCR amplification and sub-cloning. Transient transfection assays of MCF-7 cells and dual-luciferase reporter assays were conducted as previously described (Duguay et al. 2004). Cells were harvested 48 hours after transfection, which were performed in triplicate. Reporter gene activity was corrected both by subtracting background signal and by normalizing firefly activity against renilla activity obtained from the same sample. Results are expressed as means  $\pm$  SD for each triplicate and luciferase activity was compared between the variant and its respective wild-type (WT) construct, with a 2-tailed unpaired Student *t* test. P-values less than 0.05 were considered significant.

**Linkage Disequilibrium and haplotype analyses:** The LDA program (Ding et al. 2003) was used to calculate pairwise linkage disequilibrium (LD; Lewontin's  $|D'|$  and  $r^2$  values) for each SNP pair identified in our breast cancer series (Lewontin 1964; Devlin and Risch 1995). The PHASE 2.1.1 software (Stephens et al. 2001; Stephens and Donnelly 2003) was used to estimate haplotype frequencies. SNPs having a minor allele frequency (MAF)  $\geq 5\%$  identified in both sample sets were used to perform this haplotype analysis. The algorithm was run five times, with a minimum of 100 permutations, under default conditions. Genotyping data from both series, and from the CEPH cohort (HapMap database) were used to identify LD blocks within the *FANCI* gene using the Haploview software. Identification of tagging SNPs (tSNPs) within each LD block was conducted using the same software.

#### **Electronic databases**

**ESE finder:** [http://rulai.cshl.edu/cgi-bin/tools/ESE3/ese\\_finder.cgi?process=home](http://rulai.cshl.edu/cgi-bin/tools/ESE3/ese_finder.cgi?process=home)

**Haploview:** <http://www.broad.mit.edu/mpg/haploview>

**HapMap:** <http://www.hapmap.org>

**Matinspector :** <http://www.genomatix.de/matinspector.html>

**NCBI:** <http://www.ncbi.nlm.nih.gov>

**PHASE:** <http://www.stat.washington.edu/stephens/software.html>

**PolyPhen:** <http://genetics.bwh.harvard.edu/pph/>

**SIFT:** <http://blocks.fhrc.org/sift/SIFT.html>

**Splice Site Prediction Program using Neural Networks (NNSPLICE; SSPNN):**  
[http://www.fruitfly.org/seq\\_tools/splice.html](http://www.fruitfly.org/seq_tools/splice.html)

**TFSEARCH:** <http://www.cbrc.jp/research/db/TFSEARCH.html>

**Transcription Element Search System (TESS):** <http://www.cbil.upenn.edu/tess>

**UCSC Genome Bioinformatics:** <http://genome.ucsc.edu/>

## Results

***FANCI sequence analysis:*** Analysis of the promoter, whole exonic and intronic flanking sequences of the *FANCI* gene in our 96 breast cancer individuals led to the identification of 42 variants, of which 17 display a MAF  $\geq 5\%$  (Table 1). These sequence variations are composed of 37 nucleotide substitutions, three deletions and two insertions. Among these 42 variants, 13 are exonic, five and 24 are located in promoter and intronic regions, respectively, while 22 are novel, the remaining 20 being reported in the SNP database (dbSNP build 127). Analysis of the promoter region, consisting of approximately 1400 bp upstream of exon 1, led to the identification of five additional variants, including four not reported in databases.

Sequence variations displaying a MAF  $\geq 5\%$  in breast cancer cases were further assessed in a series of 73 unaffected individuals, also of French Canadian origin. Rare changes (MAF  $< 5\%$ ) located in the proximal region of these common variants (MAF  $\geq 5\%$ ) were also analyzed, yielding a total of 38 nucleotide variants characterized in both sample sets. Of these, 14 were not observed in control individuals. Corresponding frequencies are denoted in Table 1. No deviation from HWE is observed for any of the nucleotide changes identified, nor a significant difference in terms of allelic frequency was observed between both series based on single-marker analysis. The segregation of the variants identified was examined in families for which DNA material was available from multiple individuals (34 of the 42 variants), but no clear segregation between the identified variants and breast cancer was observed (data not shown).

***Conservation of human FANCI residues:*** Among the five amino acid substitutions found in our breast cancer cases, the common c.2755T>C variant (p.Pro919Ser) is located within the BRCA1-binding domain (Cantor et al. 2001) while the 4 remaining variants (c.415T>G, c.517C>T, c.577G>A, c.823A>G) are situated within the Rad3-related DNA helicase domain of *FANCI* (residues 1 to 888) (Cantor et al. 2001). As illustrated in Table 2, comparison of these amino acid changes was performed across relevant species using data extracted from the UCSC database. Alignment of *FANCI* orthologue sequences revealed that p.Ser139, p.Arg173, and p.Ile275 are strongly conserved in higher species, with p.Ile275 also found in more distant species such as *Xenopus tropicalis* and *Tetraodon nigroviridis*. The p.Val193 residue is poorly conserved while the p.Pro919 residue is highly

variable and not conserved in other species. Taken together, the comparison of these orthologues indicate that 3 variants, namely c.415T>G (p.Ser139Ala), c.517C>T (p.Arg173Cys) and c.823A>G (p.Ile275Val), are under strong functional constraint. Of these five amino acid substitutions, only p.Arg173Cys is predicted as damaging or not tolerated for protein folding. The remaining four, which include 2 unreported amino acid changes, are considered as tolerated or benign.

***In silico analysis:*** Variants located in the promoter region were evaluated for their potential effect on transcription factor binding sites (TFBS). As illustrated in Table 3, while two variants (c.-141-331T>C and c.-141-280G>T) seem to have a limited effect, the c.-141-785delCTTTT, c.-141-64G>A and c.-141-55A>G changes demonstrate a stronger potential impact on several promoter binding sites (BS), involving particularly Androgen Receptor (AR), Myogenic Factor 4 (MF4), KLF4, ras-responsive element binding 1 (RREB1) and aryl hydrocarbon receptor (AHR) BS.

Regarding the possible effect of exonic variants on exonic splicing enhancer (ESE), analysis of the scores obtained reveals that certain exonic variants markedly alter the binding capacity of putative ESE elements, either by reducing their values or by creating new binding sites for SR proteins. As illustrated in Table 4, binding site scores affecting binding of SC35, SRp55 and SRp40 proteins are mainly disturbed by 4 variants, namely c.517C>T (SRp55), c.2637G>A (SC35), c.2755T>C (SC35, SRp40 and SRp55) and the 3'-UTR c.3750+396delCT variation, which seems to reduce the binding of the SC35 protein. All these variations could result in promotion or repression of splicing, or more particularly on alternative splicing.

The potential effect of the 24 intronic variants on the strength of splice sites leading to the splicing of pre-mRNA was also evaluated by *in silico* analyses. The results obtained indicate that few sequence variations could affect pre-mRNA splicing, the score variations unlikely causing any marked changes in alternative splicing (data not shown). To further assess the possible effect of exonic and intronic variants on splicing at the mRNA level, cDNA material from our 96 high-risk breast cancer individuals was also analyzed. This analysis allowed to confirm the presence of all exonic variants previously identified on genomic DNA, with the exception of c.-138T>C which could not be detected due to 5'-end

localization of this variant on mRNA. In agreement with *in silico* studies, analysis of cDNA material revealed no aberrant splicing, even in carriers of variants that were suggestive to alter splicing. In addition, analysis of cDNA material from immortalized lymphocytes of a number of these breast cancer individuals treated with puromycin, an inhibitor of nonsense-mediated decay, confirmed these results (data not shown).

**Luciferase assay:** As illustrated in Figure 1, transient transfection of *FANCF* promoter constructs revealed a slight significant reduction for the c.-141-280G>T variant in reporter activity relative to the WT *FANCF* promoter construct, showing a 15% reduced ability to drive the firefly luciferase reporter gene. As for the four other promoter variants examined, no significant effect was observed on transcriptional activation.

**LD and haplotype analyses:** Pairwise LD between all 42 SNPs identified in the case series was calculated using LDA program (data not shown). Complete LD was found between SNPs 1 (c.-141-785delCTTTT) and 42 (c.3750+483T>C), which are located ~181 kb apart ( $|D'|=1.0$ ), indicating that LD at the *FANCF* locus does not decline significantly with distance. However, regarding specific association involving SNPs 16 (c.508-281A>G) and 17 (c.508-31C>G), lower LD values are observed, mainly with variants located within the 3'UTR region of the gene. As expected,  $r^2$  coefficient calculated for *FANCF* gene region displays lower values, as this measure is dependent on allelic frequency, which is well represented by the large spectrum of  $r^2$  values ranging from 0 to 1.0. Indeed the majority of lowest  $r^2$  values are observed for 19 of the 25 SNPs displaying a MAF below 5%.

PHASE program estimation of *FANCF* haplotypes reconstructed from the 17 SNPs having a MAF  $\geq 5\%$  in both sample sets identified 17 haplotypes exhibiting a frequency  $\geq 1\%$  (Figure 2a). These 17 haplotypes represent >80% of all haplotypes estimated in both series. The permutation test of these 17 SNP-haplotypes indicates no significant difference in the estimated haplotype frequency distributions between both groups ( $p=0.327$ , data not shown).

Given that no significant difference was observed between cases and controls following PHASE analyses, that no deviation from HWE for all SNPs in both series was observed, and that no significant difference in haplotype frequencies was identified, these observations support the use of individuals from both series as one population in order to

identify LD blocks and tSNPs of *FANCF* gene in the French Canadian population.

LD block assignment was conducted using the Haploview software, and led to the identification of 3 LD blocks in the *FANCF* gene (Figure 2b). The 5'-block identified (promoter region to intron 5) encompasses approximately 16 kb while LD blocks two (intron 8 to intron 16) and three (exon 19 to exon 20) cover 59 kb and 3 kb, respectively. Following LD-block partition, tSNPs have been identified within each block. When considering haplotypes having a frequency  $\geq 5\%$ , eight tSNPs were identified in the French Canadian population, which represent  $>95\%$  of all haplotypes. SNPs 4 (rs2048718), 7 (rs4988340) and 12 were identified as tSNPs within the first block, while SNPs 22 and 25 (rs11390869, rs2191248) represent the middle block, the third block being defined by SNPs 34, 35 and 36 (rs4986765, rs4986764 and rs4988357).

Using HapMap data from the CEPH/CEU cohort, 3 LD blocks were also identified (Figure 2b). Out of 17 SNPs used in our haplotype blocks analysis in French Canadian individuals, 11 were part of the CEPH/CEU sample set, with 9 present at a MAF  $\geq 5\%$  in the CEPH/CEU sample set. Despite using a different panel of SNPs, the composition of the three blocks is relatively the same between French Canadian and CEPH/CEU data sets. Indeed similar regions of LD breakage are observed following both analyses.

## Discussion

To increase the power of our study aiming at finding genetic variants involved in breast cancer susceptibility, we selected individuals from our cohort of French Canadian non-*BRCA1/2* high-risk breast cancer families (one individual per family). Given that this population is considered as a founder population, this allowed us to increase the likelihood of potentially identifying genetic variants associated with breast cancer (Antoniou and Easton 2003).

The previous identification of *BRCA2/FANCD1* and *PALB2/FANCN* (Rahman et al. 2007; Erkko et al. 2007) as breast cancer susceptibility genes further strengthens the implication of FA genes in breast cancer susceptibility. Inactivating truncating mutations of *FANCF* have been demonstrated to confer susceptibility to breast cancer in high-risk non-*BRCA1/BRCA2* families from the UK (Seal et al. 2006). Although no deleterious germline mutation leading to a premature termination of the protein was identified in our study, 42 nucleotide variations were found, including 22 novel variants not previously reported in nucleotide databases (dbSNP and UCSC databases).

A number of case-control studies investigated the potential association of *FANCF* germline mutations with breast cancer susceptibility. Among the few polymorphisms/mutations not identified in our study, and for which a significant association has been demonstrated, P47A and M299I were originally described by Cantor and al. (2001). However these breast cancer associations were not subsequently confirmed in large breast cancer cohorts (Rutter et al. 2003; Lewis et al. 2005; Seal et al. 2006). Two silent variants observed in our French Canadian cohort (c.2637G>A: p.Glu879Glu and c.3411C>T: p.Tyr1137Tyr) have been investigated in several breast cancer case-control studies regarding their implication with breast cancer susceptibility but none of these studies reported a significant association with breast cancer for both nucleotide alterations (Cantor et al. 2001; Luo et al. 2002; Karppinen et al. 2003; Rutter et al. 2003; Lewis et al. 2005; Vahteristo et al. 2006).

Whereas the c.517C>T (p.Arg173Cys) variant has been reported to impair nuclear translocation (Yu et al. 2003), this nucleotide change, together with another rare missense variant (c.577G>A: p.Val193Ile), do not seem to be associated with breast cancer in other studies (Luo et al. 2002; Rutter et al. 2003; Lewis et al. 2005; Vahteristo et al. 2006; Seal et

al. 2006). As for the common c.2755T>C variant (p.Pro919Ser) located in the BRCA1-binding domain spanning residues 888-1063 (Cantor et al. 2001) and which has been previously reported to be associated with an increased risk of breast cancer to age 50 (Sigurdson et al. 2004), we and others did not highlight any significant association with breast cancer risk (Cantor et al. 2001; Luo et al. 2002; Karppinen et al. 2003; Rutter et al. 2003; Vahteristo et al. 2006; Garcia-Closas et al. 2006; Seal et al. 2006; Frank et al. 2007).

Of the five variations identified in the promoter region, the sole variant assessed in previous studies (c.-141-64G>A) demonstrated no association with either breast cancer (Rutter et al. 2003; Sigurdson et al. 2004; Song et al. 2007; Pharoah et al. 2007; Frank et al. 2007) or with bladder cancer risk (Figueroa et al. 2007). In concordance with these findings, we found this sequence variation to be more frequent in the control series, further supporting its nonassociation with breast cancer risk. With the exception of c.205+162T>C and c.1281+91insT, all reported intronic variants were evaluated in the context of breast cancer (Karppinen et al. 2003; Rutter et al. 2003; Lewis et al. 2005; Vahteristo et al. 2006; Garcia-Closas et al. 2006; Song et al. 2007; Pharoah et al. 2007), and none were shown to be associated with an increased risk.

In agreement with our results, two recent studies, performed on large cohorts of breast cancer cases, included four variants identified in the present study (c.1340+109G>A, c.-141-64G>A, c.508-31C>G, c.508-281A>G) and came to the conclusion that none were associated significantly with breast cancer risk (Pharoah et al., 2007; Song et al., 2007), although rs4988344 (c.508-31C>G) seemed to be associated with an increased risk of ovarian cancer (Song et al., 2007). Although eight variants could not be evaluated further due to material availability from other family members, none of the variations examined clearly displayed a segregation with the disease. However, partial or incomplete segregation patterns were recently identified for susceptibility alleles that confer modest increase in breast cancer risk (Rahman et al. 2007), thus reinforcing the need for extensive analysis of the variants identified in our study.

Three of the five promoter variants reported here are predicted to impact TFBS affinity, which is consistent with other results obtained in 16 cell-cycle checkpoint genes (Belanger et al. 2005). The identification of RREB1 and KLF4 BS in the *FANCF* promoter, involved



in the regulation of the cell-cycle checkpoint regulator p16, and in G1/S and G2/M transitions of the cell-cycle, respectively, points toward a possible cell-cycle-dependent regulation of *FANCI* which is concordant with the known cell-cycle-dependent phosphorylation and interaction of *FANCI* with *BRCA1* (Yu et al. 2003). The promoter variants identified could also be involved in hormone activation or hormone response. Indeed, the c.-141-785delCTTTT variant seems to abolish the sole AR BS present in the promoter region analyzed, while the c.-141-64G>A variant creates a potential AHR BS, the latter known to be implicated in repression of E2F interaction with promoter of cell-cycle genes (Marlowe et al. 2004), and of estrogen-receptor (ER) binding to estrogen-responsive elements (Kharat and Saatcioglu 1996), resulting in inhibition of proliferation. Moreover, a BS for the Sp2 transcription factor, disrupted by both c.-141-64G>A and c.-141-55A>G variants, is known to interact with the G1/S phase transition-associated E2F1 transcription factor, and to regulate luteinizing hormone (LH) receptor at the transcription level.

With the exception of the c.-141-280G>T variant for which a slight significant difference in the ability to drive the firefly luciferase reporter gene was identified, no other variant identified in this study seems to cause a change in the basal transcriptional activity of the *FANCI* gene. Further investigations would be required to determine the potential impact of this promoter variant on binding of transcription factors to the *FANCI* promoter region.

Regarding alternative splicing of the *FANCI* gene, our findings are consistent with results from Cantor et al. (2001), where no aberrant spliced mRNA species was identified in individuals from breast cancer families. Altered splicing has, however, been reported in individuals from FA complementation group J (Levrán et al. 2005; Levitus et al. 2005), and has also been observed in an individual affected with breast cancer from a *BRCA1/2*-negative family (Seal et al. 2006).

Haplotypes reconstruction from the 17 SNPs having a MAF  $\geq 5\%$  in both sample sets led to the identification of 11 haplotypes (frequency  $\geq 2\%$ ) (Figure 2a). The only other haplotype analysis identified 9 haplotypes with a frequency  $\geq 2\%$  (Rutter et al, 2003). Although similar, the greater diversity obtained here is a reflection of the number of SNPs used for haplotype analysis: 17 in the present study, as opposed to 9 (all included in our set) in Rutter et al. (2003).

Analysis conducted with the Haploview software identified three distinct blocks (Figure 2b). A previous study identified two haplotype blocks located at the 5' and 3' ends with a large region of LD breakage comprising exons 1 to 12 (Rutter et al. 2003), while another investigation performed by Song et al. (2007) also identified two haplotype blocks, but with LD breakage between intron 6 and intron 9. We identified three haplotype blocks, which is concordant with results obtained from HapMap data, where three haplotype blocks were also identified. Another study including 9kb upstream and 9kb downstream of the *FANCF* gene identified 12 tSNPs (Song et al. 2007) whereas eight tSNPs were identified among our French Canadian individuals. This number of tSNPs is consistent with the number of tSNPs normally required at other gene loci and in other populations (Johnson et al. 2001; Weale et al. 2003), including the French Canadian population (Durocher et al. 2006, 2007). Although only two tSNPs (rs2191248 and rs4986765) have been tagged in both sample sets, this could be explained by the fact that several equivalent SNPs may represent a given LD block.

Analysis of promoter, exonic and flanking intronic regions of *FANCF* in 96 high-risk individuals from non-*BRCA1/BRCA2* French Canadian families did not lead to the identification of deleterious truncating mutation, exon deletion or retention of an intronic portion. However, this analysis led to the identification of 42 variants, 22 of which are novel, including two common variants displaying a frequency of 43% and 38%. Moreover, this study reports four novel variations in the promoter region of *FANCF* potentially affecting TFBS, one of these significantly altering the promoter activity in a reporter gene assay. Further studies are therefore warranted to fully establish the functional contribution of these sequence changes on *FANCF* gene expression.

**Acknowledgements**

The authors are indebted to the participants and their families for their generosity and providing DNA samples. We thank Dr Damian Labuda and Claudia Moreau at the Centre de recherche de l'Hôpital Ste-Justine for providing control DNA samples. We would like to thank Dr Martine Dumont, Dr Penny Soucy, Gilles Leblanc, Carolle Samson and Martine Tranchant for sample management, mutation screening, and skillful technical assistance as well as Claire Brousseau, Marie-Andrée Lajoie, Pascale Léger, Hélène Malouin and Joséé Rhéaume, for genetic counselling and clinical data management at the Cancer Genomics Laboratory. A special thank to Mélanie Houde for skillful laboratory assistance. We would like to thank Professor Bartha Maria Knoppers and her colleagues from the Centre de recherche en droit public de l'Université de Montréal for their precious help with ELSI issues related to our research program. We also appreciate advice received from ethics committees. This work was supported by the Canadian Institutes of Health Research (CIHR) and Institute of Cancer and Institute of Gender and Health for the INHERIT BRCA's research program, the Fonds de la Recherche en Santé du Québec (FRSQ)/Réseau de Médecine Génétique Appliquée (RMGA), the Canadian Breast Cancer Research Alliance, the CURE Foundation and the Canadian Fanconi Anemia Research Fund. F.G. is recipient of a studentship from Fondation René Bussières, J.S. is Chairholder of the Canada Research Chair in Oncogenetics, and F.D. is a recipient of a chercheur boursier from the Fonds de la Recherche en Santé du Québec (FRSQ) and a Research Career Award in the Health Sciences from CIHR/Rx&D Health Research Foundation.

**Appendix**

Other members of INHERIT BRCA involved in clinical aspects of this study

Paul Bessette: Service de gynécologie, Centre Hospitalier Universitaire de Sherbrooke, Fleurimont, QC, J1H 5N4, Canada;

Jocelyne Chiquette: Clinique des maladies du sein Deschênes-Fabia, Hôpital du Saint-Sacrement, Québec, QC, G1S 4L8, Canada;

Rachel Laframboise: Service de médecine génétique, CHUQ, Pavillon CHUL, Québec, QC, G1V 4G2, Canada;

Jean Lépine: Centre Hospitalier régional de Rimouski, Rimouski, QC, G5L 5T1, Canada;

Bernard Lespérance, Roxane Pichette: Service d'hémato-oncologie, Hôpital du Sacré-Cœur, Montréal, QC, H4J 1C5, Canada;

Marie Plante: Service de gynécologie, CHUQ, L'Hôtel-Dieu de Québec, Québec, QC, G1R 2J6, Canada;

## References

- Alter BP (2003) Cancer in Fanconi anemia, 1927-2001. *Cancer* 97(2):425-40.
- Antoniou AC, Easton DF (2003) Polygenic inheritance of breast cancer: Implications for design of association studies. *Genet Epidemiol* 25:190-202.
- Antoniou AC, Durocher F, Smith P, Simard J, Easton DF; INHERIT BRCA program members (2006) BRCA1 and BRCA2 mutation predictions using the BOADICEA and BRCAPRO models and penetrance estimation in high-risk French-Canadian families. *Breast Cancer Res* 8(1):R3.
- Ball S, Arolker M, Purushotham AD (2001) Breast cancer, Cowden disease and PTEN-MATCHS syndrome. *Eur J Surg Oncol* 27(6):604-6.
- Belanger H, Beaulieu P, Moreau C, Labuda D, Hudson TJ, Sinnett D (2005) Functional promoter SNPs in cell cycle checkpoint genes. *Hum Mol Genet* 14(18):2641-8.
- Bonfield JK, Rada C, Staden R (1998) Automated detection of point mutations using fluorescent sequence trace subtraction. *Nucleic Acids Res* 26(14):3404-9.
- Boukamp P, Petrussevska RT, Breitkreutz D, Hornung J, Markham A, Fusenig NE (1988) Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line. *J Cell Biol* 106(3):761-71.
- Boukamp P, Stanbridge EJ, Foo DY, Cerutti PA, Fusenig NE (1990) c-Ha-ras oncogene expression in immortalized human keratinocytes (HaCaT) alters growth potential in vivo but lacks correlation with malignancy. *Cancer Res* 50(9):2840-7.
- Cantor SB, Bell DW, Ganesan S, Kass EM, Drapkin R, Grossman S, Wahrer DC, Sgroi DC, Lane WS, Haber DA, Livingston DM (2001) BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell* 105(1):149-60.
- Cantor S, Drapkin R, Zhang F, Lin Y, Han J, Pamidi S, Livingston DM (2004) The BRCA1-associated protein BACH1 is a DNA helicase targeted by clinically relevant inactivating mutations. *Proc Natl Acad Sci U S A* 101(8):2357-62.
- Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR (2003) ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res* 31(13):3568-71.
- Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M, Bayerlein M, Werner T (2005) MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* 21(13):2933-42.
- CHEK2 Breast Cancer Case-Control Consortium (2004) CHEK2\*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. *Am J Hum Genet* 74(6):1175-82.
- Chenevix-Trench G, Spurdle AB, Gatei M, Kelly H, Marsh A, Chen X, Donn K, Cummings M, Nyholt D, Jenkins MA, Scott C, Pupo GM, Dork T, Bendix R, Kirk J, Tucker K, McCredie MR, Hopper JL, Sambrook J, Mann GJ, Khanna KK (2002) Dominant negative ATM mutations in breast cancer families. *J Natl Cancer Inst* 94(3):205-15.
- Collaborative Group on Hormonal Factors in Breast Cancer (2001) Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet* 358(9291):1389-99.
- Desjardins S, Belleau P, Labrie Y, Ouellette G, Bessette P, Chiquette J, Laframboise R, Lepine J, Lesperance B, Pichette R, Plante M, INHERIT BRCA, Durocher F (2007) Genetic variants and haplotype analyses of the ZBRK1/ZNF350 gene in high-risk non-

- BRCA1/2 French Canadian breast and ovarian cancer families. *Int J Cancer* [Epub ahead of print].
- Devlin B, Risch N (1995) A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* 29(2):311-22.
- Ding K, Zhou K, He F, Shen Y (2003) LDA--a java-based linkage disequilibrium analyzer. *Bioinformatics* 19(16):2147-8.
- Duguay Y, Báár C, Skorpen F, Guillemette C (2004) A novel functional polymorphism in the uridine diphosphate-glucuronosyltransferase 2B7 promoter with significant impact on promoter activity. *Clin Pharmacol Ther* 75(3):223-33.
- Durocher F, Guénard F, Desjardins S, Ouellette G, Labrie Y (2005) Inherited susceptibility to breast cancer: accomplishments and challenges. In Sinnett D, eds. *Molecular Genetics of Cancer*. Research Signpost, Kerala, India, pp 19-93.
- Durocher F, Labrie Y, Soucy P, Sinilnikova O, Labuda D, Bessette P, Chiquette J, Laframboise R, Lepine J, Lesperance B, Ouellette G, Pichette R, Plante M, Tavtigian SV, Simard J (2006) Mutation analysis and characterization of ATR sequence variants in breast cancer cases from high-risk French Canadian breast/ovarian cancer families. *BMC Cancer* 6:230.
- Durocher F, Labrie Y, Ouellette G, INHERIT BRCAs, Simard J (2007) Genetic sequence variations and ADPRT haplotype analysis in French Canadian families with high risk of breast cancer. *J Hum Genet* 52(12):963-77.
- Easton DF (1999) How many more breast cancer predisposition genes are there? *Breast Cancer Res* 1(1):14-7.
- Ellis NA, Groden J, Ye TZ, Straughen J, Lennon DJ, Ciocci S, Proytcheva M, German J (1995) The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell* 83(4):655-66.
- Erkko H, Xia B, Nikkila J, Schleutker J, Syrjakoski K, Mannermaa A, Kallioniemi A, Pylkas K, Karppinen SM, Rapakko K, Miron A, Sheng Q, Li G, Mattila H, Bell DW, Haber DA, Grip M, Reiman M, Jukkola-Vuorinen A, Mustonen A, Kere J, Aaltonen LA, Kosma VM, Kataja V, Soini Y, Drapkin RI, Livingston DM, Winqvist R (2007) A recurrent mutation in PALB2 in Finnish cancer families. *Nature* 446(7133):316-9.
- Figuroa JD, Malats N, Rothman N, Real FX, Silverman D, Kogevinas M, Chanock S, Yeager M, Welch R, Dosemeci M, Tardon A, Serra C, Carrato A, Garcia-Closas R, Castano-Vinyals G, Garcia-Closas M (2007) Evaluation of genetic variation in the double-strand break repair pathway and bladder cancer risk. *Carcinogenesis* 28(8):1788-93.
- Frank B, Hemminki K, Meindl A, Wappenschmidt B, Sutter C, Kiechle M, Bugert P, Schmutzler RK, Bartram CR, Burwinkel B (2007) BRIP1 (BACH1) variants and familial breast cancer risk: a case-control study. *BMC Cancer* 7:83.
- Garcia-Closas M, Egan KM, Newcomb PA, Brinton LA, Titus-Ernstoff L, Chanock S, Welch R, Lissowska J, Peplonska B, Szeszenia-Dabrowska N, Zatonski W, Bardin-Mikolajczak A, Struewing JP (2006) Polymorphisms in DNA double-strand break repair genes and risk of breast cancer: two population-based studies in USA and Poland, and meta-analyses. *Hum Genet* 119(4):376-88.
- Gasco M, Yulug IG, Crook T (2003) TP53 mutations in familial breast cancer: functional aspects. *Hum Mutat* 21(3):301-6.
- Giardiello FM, Brensinger JD, Tersmette AC, Goodman SN, Petersen GM, Booker SV, Cruz-Correa M, Offerhaus JA (2000) Very high risk of cancer in familial Peutz-

- Jeghers syndrome. *Gastroenterology* 119(6):1447-53.
- Guenard F, Labrie Y, Ouellette G, Beauparlant CJ, Bessette P, Chiquette J, Laframboise R, Lepine J, Lesperance B, Pichette R, Plante M; INHERIT BRCAs, Durocher F (2007) Germline mutations in the breast cancer susceptibility gene PTEN are rare in high-risk non-BRCA1/2 French Canadian breast cancer families. *Fam Cancer* [Epub ahead of print].
- Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, Ueda H, Cordell HJ, Eaves IA, Dudbridge F, Twells RC, Payne F, Hughes W, Nutland S, Stevens H, Carr P, Tuomilehto-Wolf E, Tuomilehto J, Gough SC, Clayton DG, Todd JA (2001) Haplotype tagging for the identification of common disease genes. *Nat Genet* 29(2):233-7.
- Karppinen SM, Vuosku J, Heikkinen K, Allinen M, Winqvist R (2003) No evidence of involvement of germline BACH1 mutations in Finnish breast and ovarian cancer families. *Eur J Cancer* 39(3):366-71.
- Kharat I, Saatcioglu F (1996) Antiestrogenic effects of 2,3,7,8-tetrachlorodibenzo-p-dioxin are mediated by direct transcriptional interference with the liganded estrogen receptor. Cross-talk between aryl hydrocarbon- and estrogen-mediated signaling. *J Biol Chem* 271(18):10533-7.
- Kitao S, Ohsugi I, Ichikawa K, Goto M, Furuichi Y, Shimamoto A (1998) Cloning of two new human helicase genes of the RecQ family: biological significance of multiple species in higher eukaryotes. *Genomics* 54(3):443-52.
- Kitao S, Shimamoto A, Goto M, Miller RW, Smithson WA, Lindor NM, Furuichi Y. (1999) Mutations in RECQL4 cause a subset of cases of Rothmund-Thomson syndrome. *Nat Genet* 22(1):82-4.
- Laloo F, Varley J, Ellis D, Moran A, O'Dair L, Pharoah P, Evans DG; Early Onset Breast Cancer Study Group (2003) Prediction of pathogenic mutations in patients with early-onset breast cancer by family history. *Lancet* 361(9363):1101-2.
- Lei H, Vorechovsky I (2003) BACH1 517C-->T transition impairs protein translocation to nucleus: a role in breast cancer susceptibility? *Int J Cancer* 104(3):389-91.
- Levitus M, Waisfisz Q, Godthelp BC, de Vries Y, Hussain S, Wiegant WW, Elghalbzouri-Maghrani E, Steltenpool J, Rooimans MA, Pals G, Arwert F, Mathew CG, Zdzienicka MZ, Hiom K, De Winter JP, Joenje H (2005) The DNA helicase BRIP1 is defective in Fanconi anemia complementation group J. *Nat Genet* 37(9):934-5.
- Levrán O, Attwooll C, Henry RT, Milton KL, Neveling K, Rio P, Batish SD, Kalb R, Velleuer E, Barral S, Ott J, Petrini J, Schindler D, Hanenberg H, Auerbach AD (2005) The BRCA1-interacting helicase BRIP1 is deficient in Fanconi anemia. *Nat Genet* 37(9):931-3.
- Lewis AG, Flanagan J, Marsh A, Pupo GM, Mann G, Spurdle AB, Lindeman GJ, Visvader JE, Brown MA, Chenevix-Trench G; Kathleen Cuninghame Foundation Consortium for Research into Familial Breast Cancer (2005) Mutation analysis of FANCD2, BRIP1/BACH1, LMO4 and SFN in familial breast cancer. *Breast Cancer Res* 7(6):R1005-16.
- Lewontin RC (1964) The interaction of selection and linkage. II. Optimum models. *Genetics* 50:757-82.
- Luo L, Lei H, Du Q, von Wachenfeldt A, Kockum I, Luthman H, Vorechovsky I, Lindblom A (2002) No mutations in the BACH1 gene in BRCA1 and BRCA2 negative breast-cancer families linked to 17q22. *Int J Cancer* 98(4):638-9.

- Marlowe JL, Knudsen ES, Schwemberger S, Puga A (2004) The aryl hydrocarbon receptor displaces p300 from E2F-dependent promoters and represses S phase-specific gene expression. *J Biol Chem* 279(28):29013-22.
- Medhurst AL, Huber PA, Waisfisz Q, de Winter JP, Mathew CG (2001) Direct interactions of the five known Fanconi anaemia proteins suggest a common functional pathway. *Hum Mol Genet* 10(4):423-9.
- Moisan AM, Fortin J, Moisan MD, Samson C, Bessette P, Chiquette J, Laframboise R, Lepine J, Lesperance B, Pichette R, Plante M, Provencher L, Voyer P, Goldgar D, Bridge P, Simard J (2006) No Evidence of BRCA1/2 Genomic Rearrangements in High-Risk French-Canadian Breast/Ovarian Cancer Families. *Genet Test* 10(2):104-15.
- Pharoah PD, Tyrer J, Dunning AM, Easton DF, Ponder BA; SEARCH Investigators (2007) Association between common variation in 120 candidate genes and breast cancer risk. *PLoS Genet* 3(3):e42.
- Ponder BA (2001) Cancer genetics. *Nature* 411(6835):336-41.
- Quandt K, Frech K, Karas H, Wingender E, Werner T (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res* 23(23):4878-84.
- Rahman N, Seal S, Thompson D, Kelly P, Renwick A, Elliott A, Reid S, Spanova K, Barfoot R, Chagtai T, Jayatilake H, McGuffog L, Hanks S, Evans DG, Eccles D, Easton DF, Stratton MR (2007) PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat Genet* 39(2):165-7.
- Reese MG, Eeckman FH, Kulp D, Haussler D (1997) Improved splice site detection in Genie. *J Comput Biol* 4(3):311-23.
- Renwick A, Thompson D, Seal S, Kelly P, Chagtai T, Ahmed M, North B, Jayatilake H, Barfoot R, Spanova K, McGuffog L, Evans DG, Eccles D; Breast Cancer Susceptibility Collaboration (UK), Easton DF, Stratton MR, Rahman N (2006) ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nat Genet* 38(8):873-5.
- Rutter JL, Smith AM, Davila MR, Sigurdson AJ, Giusti RM, Pineda MA, Doody MM, Tucker MA, Greene MH, Zhang J, Struwing JP (2003) Mutational analysis of the BRCA1-interacting genes ZNF350/ZBRK1 and BRIP1/BACH1 among BRCA1 and BRCA2-negative probands from breast-ovarian cancer families and among early-onset breast cancer cases and reference individuals. *Hum Mutat* 22(2):121-8.
- Sambrook J, Russell DW (2001) Irwin N, Janssen KA, Argentine J, Curtis S, Zierler M, Dickerson M, Sialiano I, McInerney N, Brown D, Schaefer S, deBruin D, Atkeson E, Weiss D, Curtis MD (eds) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Seal S, Thompson D, Renwick A, Elliott A, Kelly P, Barfoot R, Chagtai T, Jayatilake H, Ahmed M, Spanova K, North B, McGuffog L, Evans DG, Eccles D, Easton DF, Stratton MR, Rahman N (2006) Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nat Genet* 38(11):1239-41.
- Sigurdson AJ, Hauptmann M, Chatterjee N, Alexander BH, Doody MM, Rutter JL, Struwing JP (2004) Kin-cohort estimates for familial breast cancer risk in relation to variants in DNA base excision repair, BRCA1 interacting and growth factor genes. *BMC Cancer* 4:9.
- Simard J, Dumont M, Moisan AM, Gaborieau V, Malouin H, Durocher F, Chiquette J,



- Plante M, Avard D, Bessette P, Brousseau C, Dorval M, Godard B, Houde L; INHERIT BRCA, Joly Y, Lajoie MA, Leblanc G, Lepine J, Lesperance B, Vezina H, Parboosingh J, Pichette R, Provencher L, Rheaume J, Sinnett D, Samson C, Simard JC, Tranchant M, Voyer P, Easton D, Tavtigian SV, Knoppers BM, Laframboise R, Bridge P, Goldgar D (2007) 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* 44(2):107-21.
- Smogorzewska A, Matsuoka S, Vinciguerra P, McDonald ER 3rd, Hurov KE, Luo J, Ballif BA, Gygi SP, Hofmann K, D'Andrea AD, Elledge SJ (2007) Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. *Cell* 129(2):289-301.
- Song H, Ramus SJ, Kjaer SK, Hogdall E, Dicioccio RA, Whittemore AS, McGuire V, Hogdall C, Jacobs IJ, Easton DF, Ponder BA, Dunning AM, Gayther SA, D P Pharoah P (2007) Tagging Single Nucleotide Polymorphisms in the BRIP1 Gene and Susceptibility to Breast and Ovarian Cancer. *PLoS ONE* 2:e268.
- Stankovic T, Kidd AM, Sutcliffe A, McGuire GM, Robinson P, Weber P, Bedenham T, Bradwell AR, Easton DF, Lennox GG, Haites N, Byrd PJ, Taylor AM (1998) ATM mutations and phenotypes in ataxia-telangiectasia families in the British Isles: expression of mutant ATM and the risk of leukemia, lymphoma, and breast cancer. *Am J Hum Genet* 62(2):334-45.
- Stephens M, Donnelly P (2003) A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet* 73(5):1162-9.
- Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68(4):978-89.
- Szabo CI, Schutte M, Broeks A, Houwing-Duistermaat JJ, Thorstenson YR, Durocher F, Oldenburg RA, Wasielewski M, Odefrey F, Thompson D, Floore AN, Kraan J, Klijn JG, van den Ouweland AM, Wagner TM, Devilee P, Simard J, van 't Veer LJ, Goldgar DE, Meijers-Heijboer H (2004) Are ATM mutations 7271T-->G and IVS10-6T-->G really high-risk breast cancer-susceptibility alleles? *Cancer Res* 64(3):840-3.
- Tischkowitz M, Xia B, Sabbaghian N, Reis-Filho JS, Hamel N, Li G, van Beers EH, Li L, Khalil T, Quenneville LA, Omeroglu A, Poll A, Lepage P, Wong N, Nederlof PM, Ashworth A, Tonin PN, Narod SA, Livingston DM, Foulkes WD (2007) Analysis of PALB2/FANCN-associated breast cancer families. *Proc Natl Acad Sci U S A* 104(16):6788-93.
- Ueda K, Nishijima M, Inui H, Watatani M, Yayoi E, Okamura J, Yasutomi M, Nakamura Y, Miyoshi Y (1998) Infrequent mutations in the PTEN/MMAC1 gene among primary breast cancers. *Jpn J Cancer Res* 89(1):17-21.
- Vahteristo P, Yliannala K, Tamminen A, Eerola H, Blomqvist C, Nevanlinna H (2006) BACH1 Ser919Pro variant and breast cancer risk. *BMC Cancer* 6:19.
- Vezina H, Durocher F, Dumont M, Houde L, Szabo C, Tranchant M, Chiquette J, Plante M, Laframboise R, Lepine J, Nevanlinna H, Stoppa-Lyonnet D, Goldgar D, Bridge P, Simard J (2005) Molecular and genealogical characterization of the R1443X BRCA1 mutation in high-risk French-Canadian breast/ovarian cancer families. *Hum Genet* 117(2-3):119-32.
- Weale ME, Depondt C, Macdonald SJ, Smith A, Lai PS, Shorvon SD, Wood NW, Goldstein DB (2003) Selection and evaluation of tagging SNPs in the neuronal-sodium-channel gene SCN1A: implications for linkage-disequilibrium gene mapping.

- Am J Hum Genet 73(3):551-65.
- Yu CE, Oshima J, Fu YH, Wijsman EM, Hisama F, Alisch R, Matthews S, Nakura J, Miki T, Ouais S, Martin GM, Mulligan J, Schellenberg GD (1996) Positional cloning of the Werner's syndrome gene. *Science* 272(5259):258-62.
- Yu X, Chini CC, He M, Mer G, Chen J (2003) The BRCT domain is a phospho-protein binding domain. *Science* 302(5645):639-42.

## Legends to Figures

### Figure 1

Promoter activity of long (1068 bp) and short (633 bp) constructs carrying variations present in the promoter. An asterisk (\*) indicates  $P < 0.05$  (Student t test) relative to its respective WT construct.

### Figure 2

**Panel a.** Haplotype frequencies and combined frequencies of haplotypes exhibiting a MAF  $\geq 1\%$  in cases and controls, as reconstructed by the PHASE program. SNPs denoted with an asterisk (\*) are insertions and are represented by the I letter in haplotypes.

**Panel b.** Haplotype blocks and tSNPs identified using SNPs showing a MAF higher than 5% (17 SNPs). Tagging SNPs identified on a block-by-block basis using haplotypes showing a frequency higher than 5% are denoted with an asterisk (\*). Haplotype frequencies are displayed on the right of each haplotype combination while the level of recombination is displayed above the connections between two blocks. Thin connections represent haplotypes with frequencies between 1 and 10% while haplotypes with frequencies higher than 10% are represented by thick lines.

Table 1

Observed coding and intronic sequence variants and genotype frequencies in familial breast cancer cases and controls

SNP	SNP ID <sup>1</sup>	dbSNP ID	Amino acid change	Location	Cases (N=96)				Controls			
					Common Homoz.	Heteroz.	Rare Homoz.	MAF	Common Homoz.	Heteroz.	Rare Homoz.	MAF (N=)
1	c.-141-785delCTTTT	N/A	-	Promoter	94	2	0	0.010	62	0	0	0.000 (62)
2	c.-141-331T>C	N/A	-	Promoter	93	3	0	0.016	59	2	0	0.016 (61)
3	c.-141-280G>T	N/A	-	Promoter	95	1	0	0.005	61	0	0	0.000 (61)
4	c.-141-64G>A	rs2048718	-	Promoter	38	41	17	0.391	23	32	15	0.443 (70)
5	c.-141-55A>G	N/A	-	Promoter	95	1	0	0.005	70	0	0	0.000 (70)
6	c.-138T>C	N/A	-	5' UTR	95	1	0	0.005	70	0	0	0.000 (70)
7	c.-31+12G>A	rs4988340	-	Intron 1	52	40	4	0.250	36	23	10	0.312 (69)
8	c.-31+22G>C	rs4988341	-	Intron 1	95	1	0	0.005	68	1	0	0.007 (69)
9	c.205+162T>C	N/A	-	Intron 3	94	2	0	0.010	N/A	N/A	N/A	N/A
10	c.249A>G	N/A	p.Gln83Gln	Exon 4	95	1	0	0.005	72	0	0	0.000 (72)
11	c.379+76A>G	N/A	-	Intron 4	95	1	0	0.005	72	0	0	0.000 (72)
12	c.380-226insA	N/A	-	Intron 4	31	48	17	0.427	33	28	11	0.347 (72)
13	c.380-223A>T	N/A	-	Intron 4	95	1	0	0.005	72	0	0	0.000 (72)
14	c.380-28G>A	rs4988343	-	Intron 4	91	5	0	0.026	70	2	0	0.014 (72)
15	c.415T>G	N/A	p.Ser139Ala	Exon 5	95	1	0	0.005	71	0	0	0.000 (71)
16	c.508-281A>G	rs9908659	-	Intron 5	31	48	17	0.427	29	28	11	0.368 (68)
17	c.508-31C>G	rs4988344	-	Intron 5	75	18	3	0.125	53	15	0	0.110 (68)
18	c.517C>T	rs4988345	p.Arg173Cys	Exon 6	95	1	0	0.005	68	0	0	0.000 (68)
19	c.577G>A	rs4988346	p.Val193Ile	Exon 6	94	2	0	0.010	67	1	0	0.007 (68)
20	c.823A>G	N/A	p.Ile275Val	Exon 7	95	1	0	0.005	66	0	0	0.000 (66)
21	c.1140+11G>A	N/A	-	Intron 8	95	1	0	0.005	69	0	0	0.000 (69)
22	c.1140+91insT	rs11390869	-	Intron 8	3	33	60	0.797	2	26	41	0.783 (69)
23	c.1140+198C>G	N/A	-	Intron 8	35	49	12	0.380	27	31	9	0.366 (67)
24	c.1340+53G>A	N/A	-	Intron 9	95	1	0	0.005	72	0	0	0.000 (72)
25	c.1340+109G>A	rs2191248	-	Intron 9	36	47	13	0.380	32	31	9	0.340 (72)
26	c.1474-196delCCCT	N/A	-	Intron 10	88	8	0	0.042	N/A	N/A	N/A	N/A
27	c.1628+99G>A	N/A	-	Intron 11	93	3	0	0.016	N/A	N/A	N/A	N/A
28	c.1628+163G>A	N/A	-	Intron 11	95	1	0	0.005	N/A	N/A	N/A	N/A
29	c.1795-47C>G	rs4988351	-	Intron 12	58	30	8	0.240	44	22	3	0.203 (69)
30	c.2097+7G>A	rs4988352	-	Intron 14	94	2	0	0.010	70	0	0	0.000 (70)
31	c.2379+204T>G	rs12937232	-	Intron 16	62	31	3	0.193	33	15	2	0.190 (50)
32	c.2379+218G>A	N/A	-	Intron 16	92	4	0	0.021	50	1	0	0.010 (51)
33	c.2379+306T>C	rs9904294	-	Intron 16	62	31	3	0.193	32	16	3	0.216 (51)
34	c.2637G>A	rs4986765	p.Glu879Glu	Exon 19	36	49	11	0.370	35	29	7	0.303 (71)
35	c.2755T>C	rs4986764	p.Pro919Ser	Exon 19	18	48	30	0.563	12	32	26	0.600 (70)
36	c.2905+83T>A	rs4988357	-	Intron 19	45	40	11	0.323	27	29	11	0.375 (67)
37	c.2906-86G>A	N/A	-	Intron 19	95	1	0	0.005	61	0	0	0.000 (61)
38	c.3411C>T	rs4986763	p.Tyr1137Tyr	Exon 20	27	49	20	0.464	25	34	12	0.408 (71)
39	c.3750+128A>G	N/A	-	3' UTR	93	3	0	0.016	58	2	0	0.017 (60)
40	c.3750+158C>T	rs1978111	-	3' UTR	27	49	20	0.464	27	32	12	0.394 (71)
41	c.3750+396delCT	N/A	-	3' UTR	95	1	0	0.005	61	2	0	0.016 (63)
42	c.3750+483T>C	rs7213430	-	3' UTR	28	49	19	0.453	23	29	10	0.395 (62)

<sup>1</sup> According to the nomenclature of the Human Genome Variation Society  
N/A : Information not available for this SNP or not reported in dbSNP  
MAF : Minor Allele Frequency

Table 2

SNP	SNP ID*	Amino acid change	<i>Canis familiaris</i>	<i>Mus musculus</i>	<i>Rattus norvegicus</i>	<i>Gallus gallus</i>	<i>Monodelphis domestica</i>	<i>Xenopus tropicalis</i>	<i>Tetraodon nigroviridis</i>
15	c.415T>G	p.Ser139Ala	Ser	Ser	Ser	Ser	Ser	-	-
18	c.517C>T	p.Arg173Cys	Arg	Arg	Arg	Arg	Arg	-	-
19	c.577G>A	p.Val193Ile	Thr	Val	Val	Lys	Met	-	-
20	c.823A>G	p.Ile275Val	Ile	Ile	Ile	Ile	Ile	Ile	Ile
35	c.2755T>C	p.Pro919Ser	Leu	Phe	Leu	Ser	Leu	-	-

\* According to the nomenclature of the Human Genome Variation Society

- No corresponding residue in this species

Table 3

Identification of Transcription Factor Binding Sites (TFBS) potentially affected by variants located in the *FANCL* promoter region

Variant	TFBS potentially affected	Strand	Sequence affected (wt sequence shown)*	Matrix similarity score change (wt/var)
c.-141-785delCTTTT	Androgen Receptor	plus	gatgaacctc <b>TTTT</b> cttt	0.836/ND
	Krüppel-like Factor 4	minus	ctc <b>AAA</b> G <b>aaa</b> aggaggtt	ND/0.990
	Myogenic Factor 4	minus	cttta <b>ATT</b> Tactcaaa <b>gaaa</b> agg	0.786/ND
c.-141-331T>C	Forkhead box I1	plus	tttaaac <b>AAA</b> Cactagg	0.990/0.997
c.-141-280G>T	Krüppel-like Factor 4	minus	aaaaaaaa <b>ca</b> AAGG	0.955/0.972
c.-141-64G>A	RREB1	plus	CCCC <b>Ag</b> tcctgcaca	ND/0.823
	Sp2	minus	ttgtgtgca <b>GGACT</b> g	0.800/ND
c.-141-55A>G	Aryl hydrocarbon Receptor	minus	ccggcctt <b>gTGT</b> Gcaggactgg	ND/0.974
	Sp2	minus	ttgtgtgca <b>GGACT</b> g	0.800/ND

\* Positions in bold are variant positions while capital letters are positions representing the core sequence of the binding site

\*\* wt: wild type allele; var: variant allele

ND: Not Detected

Table 4

Identification of Exonic Splicing Enhancer (ESE) potentially affected by exonic variants in the *FANCF* gene

Variant	Location	Binding site potentially affected	Motif in wild-type sequence*	Score	Motif in variant sequence*	Score
c.415T>G	Exon 5	SC35	ATCTGCTA	2.473	AGCTGCTA	3.121
c.517C>T	Exon 6	SRp55	AACGTC	3.640	AATGTC	2.089
c.577G>A	Exon 7	SRp40	TGACTAT	2.308	TGACTGT	2.875
c.2637G>A	Exon19	SC35	GAGTCCTT	1.292	<b>G</b> AATCCTT	2.589
		SC35	AGTCCTTG	2.401	<b>A</b> ATCCTTG	2.041
c.2755T>C	Exon19	SF2/ASF	CTCACCT	2.536	CCCACCT	1.932
		SC35	GTACCTCA	1.970	<b>G</b> TACCCCA	3.700
		SRp40	CCTCACC	3.772	CCCCACC	2.424
		SRp55	TACCTC	3.223	TACCCC	0.618
c.3411C>T	Exon 20	SC35	CGATCCTG	3.181	<b>T</b> GATCCT	3.160
c.3750+396delCT	Exon 20	SC35	CTTCTCTG	2.447	CTTCTGCC	1.400

\* Variant positions are in bold

Figure 1

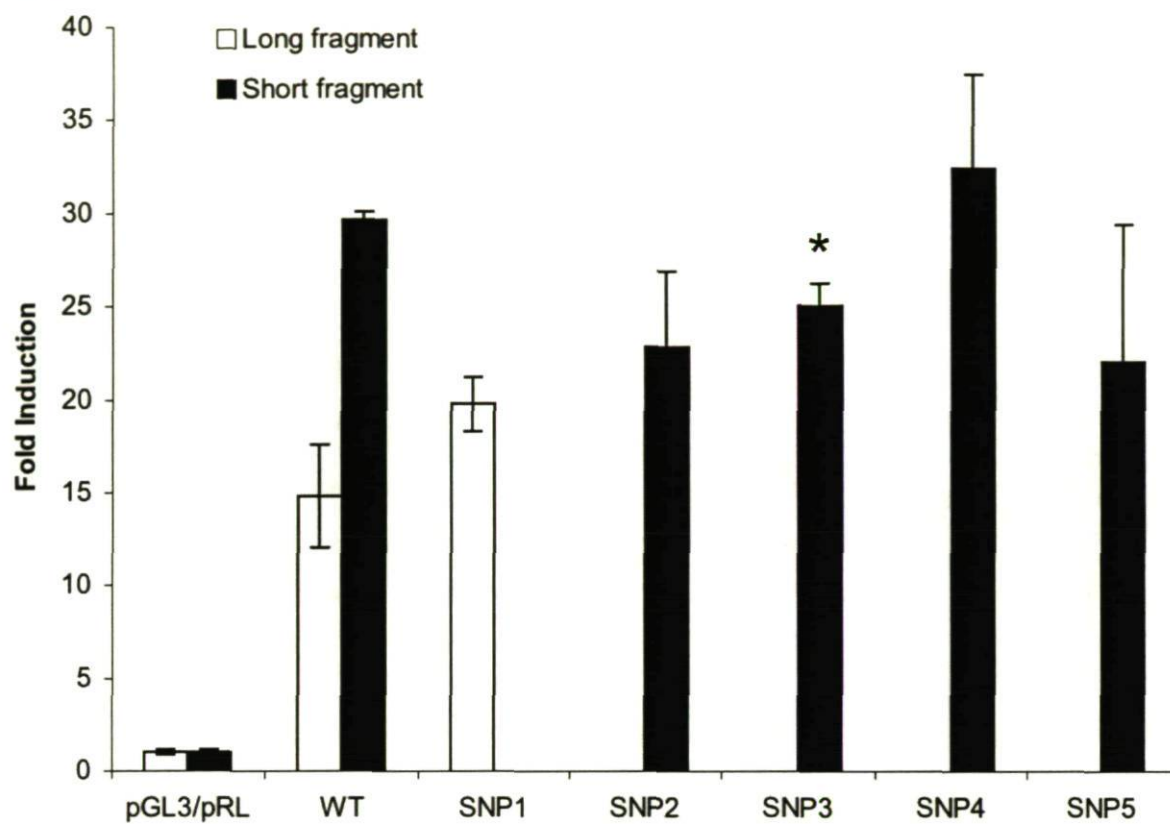
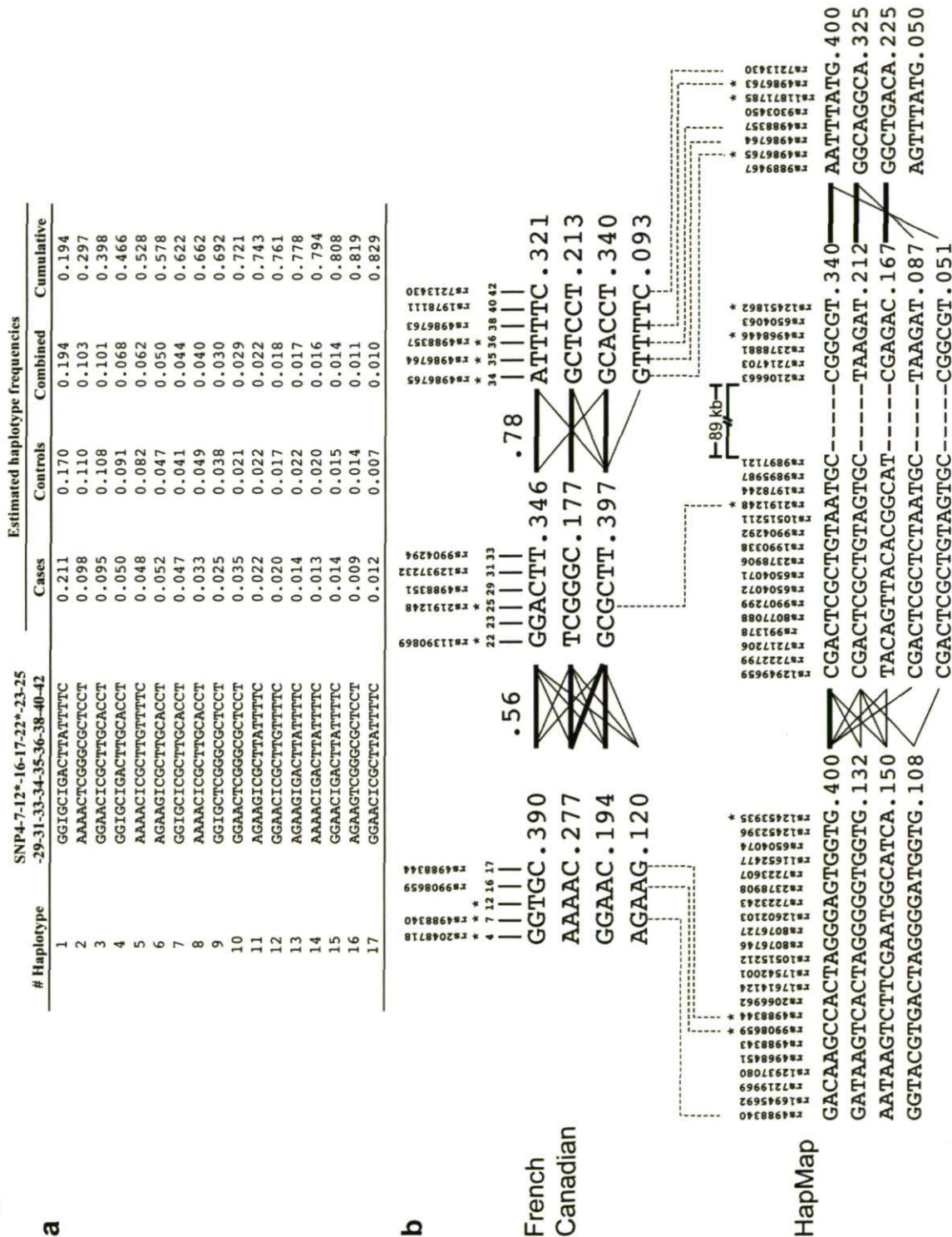




Figure 2



Supplemental table 1  
Oligonucleotides used to amplify and sequence promoter, exons and flanking intronic regions of the *FANCI* gene.

Fragment	Material	Forward primer	Reverse primer	Sequence primer	Annealing temperature (°C)	Fragment length (bp)
PCR 1	cDNA	GGGCTCCGCTTTATTGCTCT	GGCTGTGTAATATGGACAGGCCT	CGGTTTCCCTCCAGACAGTTAGGA, GGCTGTGTAATATGGACAGGCCT and CTAGCAGATAACTTTGCAGCCAGA	64	897
PCR 2	cDNA	GCACACAAGCAGATTGCTCA	CTCCTCCCTGTGGTTCTACAATGA	GCACACAAGCAGATTGCTCA, CTCCTCCCTGTGGTTCTACAATGA and TGATATGATTAGCTCCAGCTGG	64	1447
PCR 3	cDNA	CAAGGGTCGGAATCTCTGTGCT	GGCACTATTCTCTGATGACCCGA	CAAGGGTCGGAATCTCTGTGCT and GGCACTATTCTCTGATGACCCGA	63	1126
PCR 4	cDNA	CAGTATTCTGGAAGAAGCAGGGA	CGTGAAACTAGAAGTGCCTTAGGA	CAGTATTCTGGAAGAAGCAGGGA, CGTGAAACTAGAAGTGCCTTAGGA and GGATAAATGAACTGGAAGTGGGAA	64	1458
Promoter	gDNA	CCCGCCAAGAGTTGATGTT	TACCTTTCCTCGACCCACCAGA	CCCGCCAAGAGTTGATGTT, TACCTTTCCTCGACCCACCAGA and GATTTGGCAAGATCATGATGAACCT	64	1388
Exon 1	gDNA	GACTTCCATTGGATGCCGAAGT	TCCCGCCAAGAATTAACGCA	GACTTCCATTGGATGCCGAAGT	62	583
Exon 2	gDNA	CATGCTCTTGACCGCTGTGA	CTGCAATCAGTAGTTTCCAGAGGT	CTGCAATCAGTAGTTTCCAGAGGT	64	506
Exon 3	gDNA	GATGCTTGATAAATTCTGGCACA	TGCCAAGTCTTTATGTACTTTCCGA	GATGCTTGATAAATTCTGGCACA	60	615
Exon 4	gDNA	GATGCTGCAGATAAAGTGACCAGCT	GCCAGGCAAATATCAAGAATGA	GATGCTGCAGATAAAGTGACCAGCT	62	745
Exon 5	gDNA	GGTTTTGCTCTAAAGAAGCTGGAGA	ATGCAGTTTCACTGAAACGACATGT	ATGCAGTTTCACTGAAACGACATGT and GGTTTTGCTCTAAAGAAGCTGGAGA	60	732
Exon 6	gDNA	GCATGTCATCAGGTAGGGCTCA	CTCCTTGAGAACTGTTGCCCTAAGA	CTCCTTGAGAACTGTTGCCCTAAGA	61	1151
Exon 7	gDNA	GATTCCATGTGAGGTTGATAACGT	CCAAGCTTTAAGACCAATCCCAT	CCAAGCTTTAAGACCAATCCCAT	62	545
Exon 8	gDNA	GGGTATTGGAGCTCTCCGAT	CCTGCTGTAATCCCCTCAGTCA	GGGTATTGGAGCTCTCCGAT and CACTTCTCAAAGCTCACAAGGCA	60	701
Exon 9	gDNA	GTTGCTGTGATAGGTTTGCAAAGT	GGGGCTATTATGTTATCCTGCCA	GGGGCTATTATGTTATCCTGCCA	60	705
Exon 10	gDNA	GACCTGGCTCTCAGTTGAAGAGACA	CCCACAATTAACCATGCCAT	GACCTGGCTCTCAGTTGAAGAGACA	60	617
Exon 11	gDNA	ACAGAGCACTAACTCACCCCTCT	CAGGTGATGAGGGAAAGACTCTA	ACAGAGCACTAACTCACCCCTCT	64	1321
Exon 12	gDNA	GAGCATGCAGAGACTCTGGACT	TCTTTAGAAGAGGCTGGCAAAGT	GAGCATGCAGAGACTCTGGACT	62	918
Exon 13	gDNA	CTAGACTCAAGCGATCTCTCGTCT	GCTTAAACACGAGTGAGCCAAGA	ACTTTGCCAGCCTCTTCTAAAGA, GCTTAAACACGAGTGAGCCAAGA and TGATATGATTAGCTCCAGCTGG	No annealing	649
Exon 14	gDNA	CTATGGCAAATATACTGGCCTTCA	ACGTCCAGCCTTGCTCTTTCA	ACGTCCAGCCTTGCTCTTTCA	60	788
Exon 15	gDNA	CAGCAGCTAACTGTGCCAGCA	AGACTGTGAGATGATGGCATCGT	AGACTGTGAGATGATGGCATCGT	64	619
Exon 16	gDNA	TGCAGAGTTTGATAAATTGGTTTGCT	GGGAAGGGTAGTCAGTGGAGCA	GGGAAGGGTAGTCAGTGGAGCA	60	716
Exon 17	gDNA	AACTGCTTGAGATCACACAGCTGAT	CTGAGCTTTGATGTTACAAGACAA	AACTGCTTGAGATCACACAGCTGAT	60	669
Exon 18	gDNA	AATGCAAGGATGATTCAGTGTGTTGA	GGGCATGTGTTTCACTGTTGGT	GGGCATGTGTTTCACTGTTGGT	60	787
Exon 19	gDNA	CATCTCTGCCCTACTTCTACTGCCT	GACAAATGACTTGCTCATGATCACA	GACAAATGACTTGCTCATGATCACA	62	843
Exon 20*	gDNA	TTGGGCTAGGAAACATTAGAGACCT	CTTCAGAACAGAGCGGATGTTCA	CTTCAGAACAGAGCGGATGTTCA	62	589
Exon 20*	gDNA	CAGTATTCTGGAAGAAGCAGGGA	CGTGAAACTAGAAGTGCCTTAGGA	GGTAAAATACCGAAGGCAACACCT, GCAGAGAAGGGCAGATCACGA and GGATAAATGAACTGGAAGTGGGAA	62	1398
Exon 20*	gDNA	ACTGCAACCTTCACCTCCAGGTT	CAGAGGGGATAAATTGTTTCCACCT	ACTGCAACCTTCACCTCCAGGTT	62	837

\* Exon 20 has been amplified in three fragments from 5' position to 3' position.

Supplemental table 2  
 Oligonucleotides used for cloning and site-directed mutagenesis of the *FANCI* promoter constructs.

Fragment	Material	Forward primer*	Reverse primer*
Long promoter construct	gDNA	<b>CTTGTCGGTACCCAAATTGACAATGAGCAAAACAATC</b>	<b>CTTGTCCTCGAGTCTGAGCTCCGATTCACTGAAAGA</b>
Short promoter construct	gDNA	<b>CTTGTCGGTACCCCTGAGTCTTCGCTATCCTGGCTACT</b>	<b>CTTGTCCTCGAGTCTGAGCTCCGATTCACTGAAAGA</b>
Promoter site-directed mutagenesis SNP1	Promoter construct	CATGATGAACCTCCTTTGAGTAAAT	ATTTACTCAAAGGAGGTTTCATCATG
Promoter site-directed mutagenesis SNP2	Promoter construct	TTAAACAAACACCCAGGGATTTGCTG	CAGCAAAATCCCCTGGTGTGTTGTTTAA
Promoter site-directed mutagenesis SNP3	Promoter construct	CCTTGAGCCCTTTTTTTTTTTTTTTT	AAAAAAAAAAAAAAAAAGGGCTCAAGG

\* Bold nucleotides are added to amplification primers to introduce artificially engineered *KpnI* (forward) and *XhoI* (reverse) restriction sites.

## CHAPITRE IV

### **Evaluation of the contribution of the three breast cancer susceptibility genes *CHEK2*, *STK11* and *PALB2* in non-*BRCA1/2* French Canadian families with high risk of breast cancer**

#### **1. Résumé**

Des mutations inactivatrices de *CHEK2* et *STK11* sont responsables des syndromes de Li-Fraumeni et de Peutz-Jeghers, deux syndromes autosomiques dominants associés à un risque accru de cancer du sein. Des mutations tronquantes monoalléliques de *PALB2* ont aussi été associées à un risque accru de cancer du sein. Nous avons analysé la région codante et les régions introniques flanquantes de ces gènes dans notre cohorte de cas de cancer du sein. Cette étude a révélé une mutation de *CHEK2* (c.1100delC) et une mutation de *STK11* rapportée chez une famille atteinte du syndrome de Peutz-Jeghers. Un total de 17, 25 et 27 variants dans les gènes *CHEK2*, *STK11* et *PALB2* respectivement ont été identifiés. L'analyse de la fréquence de ces variants chez 96 femmes non-atteintes a suggéré une différence de fréquence entre les individus atteints et non-atteints pour deux variants introniques de *STK11*. Des mutations de ces gènes seraient donc rares et ne contribueraient pas à une proportion substantielle de la susceptibilité au cancer du sein chez les familles canadiennes-françaises à risque élevé.

Guénard F, St-Laurent Pedneault C, Ouellette G et al. Evaluation of the contribution of the three breast cancer susceptibility genes *CHEK2*, *STK11* and *PALB2* in non-*BRCA1/2* French Canadian families with high risk of breast cancer. *Genet Test Mol Biomarkers* 2010  
*Soumis*

#### **2. Manuscrit**

**Evaluation of the contribution of the three breast cancer susceptibility genes *CHEK2*, *STK11* and *PALB2* in non-*BRCA1/2* French Canadian families with high risk of breast cancer**

Frédéric Guénard<sup>a</sup>, Christopher St-Laurent Pedneault<sup>a</sup>, Geneviève Ouellette<sup>a</sup>, Yvan Labrie<sup>a</sup>, Jacques Simard<sup>a,b</sup>, INHERIT<sup>c</sup>, Francine Durocher<sup>a,\*</sup>

<sup>a</sup> Cancer Genomics Laboratory, Division of Endocrinology and Genomics of CHUQ Research Center and Laval University, Québec, G1V 4G2, Canada;

<sup>b</sup> Canada Research Chair in Oncogenetics;

<sup>c</sup> Clinicians involved in this study are listed in Appendix

\*Corresponding author:

Dr Francine Durocher

Cancer Genomics Laboratory

CHUL Research Center, CHUQ

2705 Laurier Boulevard, T2-53

Québec city, Québec,

Canada, G1V 4G2

Tel: 418-654-2296

Fax: 418-654-2278

**e.mail: [francine.durocher@crchul.ulaval.ca](mailto:francine.durocher@crchul.ulaval.ca)**

**Keywords:** *CHEK2*; *STK11*; *PALB2/FANCN*; breast cancer; non-*BRCA1/2* families;

MLPA

**Abstract**

Inactivating mutations of the *CHEK2* and *STK11* genes are responsible for Li-Fraumeni and Peutz-Jeghers syndrome, respectively, both autosomal dominant syndromes associated with an increased risk of breast cancer. The *PALB2/FANCN* gene encodes a nuclear partner of BRCA2 and acts as a linker between BRCA1 and BRCA2. Monoallelic *PALB2* truncating mutations were shown to confer higher risk of breast cancer. In order to evaluate the proportion of French Canadian non-*BRCA1/BRCA2* families with high risk of breast cancer potentially harboring alterations in these three breast cancer susceptibility genes, the whole coding and flanking intronic sequences were analyzed in a series of 96 high risk breast cancer individuals. Despite no *PALB2* deleterious truncating mutations being identified, the c.1100delC breast cancer-associated *CHEK2* mutation and a *STK11* mutation reported to be the causative mutation in a Peutz-Jeghers family were identified. This extensive analysis also led to the identification of a total of 17, 25 and 27 variants in the *CHEK2*, *STK11* and *PALB2* genes, respectively. Ascertainment of allele frequency of these variants in a cohort of 96 healthy unrelated women suggests a difference in allele frequency for two *STK11* intronic variants. In addition, large genomic rearrangements in both *STK11* and *PALB2* were also examined. Our analysis led to the conclusion that *CHEK2*, *STK11* and *PALB2* mutations or large genomic rearrangements of either *STK11* or *PALB2* are rare, and do not contribute to a substantial fraction of breast cancer susceptibility in high-risk French Canadian breast cancer families.

## Introduction

In developed countries, breast cancer incidence has been estimated to 7.8% at 80 years old. Mutations in the two most important breast cancer susceptibility genes in high-risk families, *BRCA1* and *BRCA2*, confer an estimated risk of breast cancer by age 70 of 65% and 45%, respectively (Antoniou et al., 2003). A lower proportion of hereditary breast cancers is explained by mutations in *TP53*, *ATM*, *PTEN*, *BRIP1*, *CHEK2*, *STK11* and *PALB2* (Stratton and Rahman, 2008).

Maintenance of genome integrity depends on the coordination of cell cycle checkpoints, DNA repair systems, and apoptosis. DNA damage checkpoints are associated with delay or arrest of cell cycle progression. The checkpoint kinase 2 (*CHEK2/CHK2*) gene encodes a serine/threonine (ser/thr) kinase involved in cell cycle regulation following DNA damage (Ahn et al., 2004). *CHEK2* is a major downstream effector of *ATM* (ataxia telangiectasia-mutated) in response to double-strand breaks (DSBs). *BRCA1*, *BRCA2* and *p53* are known substrates of *CHEK2* and are implicated in several intracellular processes to maintain genomic stability (Ahn et al., 2004; Bahassi et al., 2008).

Heterozygous germline mutations in *CHEK2* occur in Li-Fraumeni syndrome (LFS) (Bell et al., 1999), a rare familial cancer phenotype usually associated with inherited mutations in the *TP53* gene. LFS is characterized by the occurrence of sarcomas, breast cancer, brain tumors, leukemia, and adrenal cortical tumors in multiple relatives (Li and Fraumeni, 1969). Germline mutations or variants in the *CHEK2* gene have been reported in non-*BRCA1/2* breast cancer families (Bell et al., 1999; Vahteristo et al., 2002). Most notably, the c.1100delC mutation was originally found in cases of the cancer-predisposing LFS (Bell et al., 1999) and this protein truncating mutation located within the kinase domain was reported to abolish the kinase activity of *CHEK2* (Wu et al., 2001). In 2002, a population-based cohort revealed an association of the *CHEK2* c.1100delC mutation among patients with a positive family history (Vahteristo et al., 2002). In the same year, the *CHEK2*-Breast Cancer Consortium observed a significant difference between breast cancer individuals from non-*BRCA1/BRCA2* families and healthy controls, which led to the estimation that the *CHEK2* c.1100delC mutation shared an approximately two-fold increase of breast cancer risk in women while there is a ten-fold increased risk in men (Meijers-Heijboer et al., 2002). Later study confirmed that *CHEK2* c.1100delC confers an increased

risk of breast cancer and that risk was apparent in women unselected for family history (CHEK2 Breast Cancer Case-Control Consortium, 2004). It was thereafter suggested that *CHEK2* c.1100delC is an important breast cancer-predisposing gene, which increases the risk three- to five-fold (Weischer et al., 2008).

The *STK11* gene encodes a ser/thr kinase that inhibits cellular proliferation, controls cell polarity and interacts with the mTOR pathway (Alessi et al., 2006). Inactivating mutations of this gene are responsible for Peutz-Jeghers syndrome (PJS), an autosomal dominant syndrome characterized by multiple hamartomatous polyps in the gastrointestinal tract which is associated with an increased risk of gastrointestinal and extraintestinal cancers, including breast cancer. Indeed, the cumulative risk of breast cancer in patients with PJS, which is associated with mutations in the *STK11* gene, has been estimated to be 45-54% (Giardiello et al., 2000; Hearle et al., 2006). Despite the high cumulative risk of breast cancer reported in PJS patients, few studies evaluated the *STK11* gene in the context of breast cancer. No somatic mutations were identified in sporadic breast cancers (Bignell et al., 1998; Forster et al., 2000), nor in 14 breast cancer families with at least one tumor showing loss of heterozygosity at 19p, the location of the *STK11* gene (Chen and Lindblom, 2000).

The PALB2 protein was identified as a nuclear partner of BRCA2 (Xia et al., 2006), by promoting its localization and stability in key nuclear structures, which in turn facilitates BRCA2 functions in DNA repair. Like BRCA2, PALB2 appears to participate in DNA damage response and is recruited to a subset of DNA DSBs (Xia et al., 2006). It was recently demonstrated that PALB2 also binds BRCA1 (Sy et al., 2009a; Zhang et al., 2009). Moreover, PALB2 is an integral component of the BRCA complex and acts as a linker between BRCA1 and BRCA2 (Sy et al., 2009a). While biallelic *PALB2* mutations were associated with a new subtype of Fanconi anemia, FA-N (Reid et al., 2007; Xia et al., 2007), monoallelic truncating *PALB2* mutations were identified in both familial and unselected breast cancer individuals and were reported to confer a 2.3- to 6-fold higher risk of breast cancer (Rahman et al., 2007; Erkkö et al., 2008). It has been shown that disease-associated variants in *PALB2* are rare in the UK population, and only one pathogenic sequence variant accounting for a frequency of 1% was identified in Spanish *BRCA1/2*-negative breast cancer families (García et al., 2009). However, a single *PALB2* mutation



accounts for 1% of all breast cancers in Finland (Erkko et al., 2007), and a founder truncating mutation (Gln775X) has been identified in the French Canadian population (Foulkes et al., 2007). These two founder populations were also analyzed for *PALB2* large genomic rearrangements but such alterations were not found (Foulkes et al., 2007; Pylkäs et al., 2008).

In our cohort of high-risk French Canadian breast/ovarian cancer families, deleterious mutations in *BRCA1/BRCA2* were identified in only 24% of 256 families (Simard et al., 2007) without any evidence of large genomic rearrangements (Moisan et al., 2006). The high proportion of high-risk French Canadian breast/ovarian families not due to *BRCA1/2* mutations prompted us to evaluate the contribution of the breast cancer susceptibility genes *CHEK2*, *STK11* and *PALB2* in these families. Therefore, in order to evaluate the extent of germline mutations in these genes, all coding exons and their flanking intronic regions were analyzed in distinct *BRCA1/BRCA2*-negative high-risk French Canadian breast/ovarian families. In addition to evaluating deleterious truncating mutations, the potential effect of sequence variants identified was investigated both by assessing their frequency in healthy unrelated French Canadian individuals and by evaluating their potential effect on protein function. The use of individuals from families with high risk of breast cancer is recognized to increase the power of a study in comparison with unselected breast cancer cases (Antoniou and Easton, 2003). However, the identification of an association between a variant and breast cancer risk requires thousands of individuals.

## Materials and methods

### *Ascertainment of high-risk families*

High-risk French Canadian breast and/or ovarian families were recruited through a large ongoing interdisciplinary research program designated INHERIT BRCA. Clinicians from seven participating hospitals were directly involved in this translational research program and approval was obtained from ethics committees corresponding to the different institutions participating in the INHERIT BRCA program. More details regarding ascertainment criteria, experimental and clinical procedures, as well as the INHERIT BRCA research program have been described elsewhere (Durocher et al., 2006; Simard et al., 2007). Patients were screened for mutations of the *BRCA1* and *BRCA2* genes and *BRCA1/BRCA2* deletions or large genomic rearrangements were excluded in these individuals (Moisan et al., 2006; Simard et al., 2007). Subsequently, another component was designed for the “Localization and identification of new breast cancer susceptibility loci/genes”. Ethics approval for this latter study was also obtained from the different institutions participating in this research project and each participant knowing their inconclusive *BRCA1/2* test results status had to sign a specific informed consent for their participation in this component. All patients were referred by a physician from the Québec province and had to be at least 18 years of age and mentally capable. From these *BRCA1/2*-negative families, a subset of 96 high-risk French Canadian breast cancer individuals were selected from 96 families according to the following ascertainment criteria 1) three or more breast cancer cases diagnosed before the age of 65, 2) two or more breast cancer cases (<65) if one breast cancer was diagnosed before 45 years, 3) or when there was a strong family history of breast/ovarian cancer (e.g. daughter-mother-grand-mother) (Guénard et al., 2007). When two or more subjects were available within a family, the youngest subject was systematically chosen for this study. The mean age at diagnosis of these 96 subjects affected with breast cancer was 48.5 years old (31-74 years). Of these, 59 have been diagnosed before the age of 50.

The recruitment of 96 healthy unrelated French Canadian individuals was also conducted on a non-nominative basis as part of a long term study aiming at the characterization of the genetic variability in human population and approved by the Institutional Ethics Review Board of the Hôpital Sainte-Justine. Few information is available for these anonymized

individuals. Age was available for nearly 70 of them, the mean age calculated from these individuals was 48.5 years old (30-72 years).

#### ***DNA extraction, RNA isolation from immortalized cell lines and cDNA synthesis***

Lymphocytes from breast cancer individuals were isolated and immortalized from 7 to 9 mL of blood samples using the Epstein-Barr Virus in 15% RPMI media as previously described (Durocher et al., 2006). Subsequent extraction, storage and reverse transcription of total RNA, as well as genomic DNA extraction were performed as previously described (Durocher et al., 2006). Genomic DNA from 96 controls was extracted from peripheral blood using standard methods, either phenol–chloroform, Gentra kits (Minneapolis, MN, USA), or the QIAamp DNA Blood MAXI Kit (QIAGEN Inc., Mississauga, ON, Canada) following supplier's instructions. These DNA samples were thereafter subjected to whole-genome amplification using Illustra GenomiPhi V2 DNA amplification kit (GE Healthcare, NJ, USA) according to the manufacturer's instructions.

#### ***PCR amplification and mutation analysis***

Amplification and direct sequencing of genomic coding sequence of the *PALB2* and *STK11* genes were conducted using primers chosen in order to amplify and sequence exonic and flanking intronic regions of these genes. cDNA amplicons covering the entire mRNA-encoding sequences of the *STK11* gene (GenBank mRNA record NM\_000455.4) were also obtained by PCR amplification from cDNA of breast cancer cases. Primers used to amplify and sequence the *PALB2* and *STK11* genes are listed in Supplemental Table 1. All exons of *PALB2* were amplified and sequenced. An average of 176 bp of intronic sequence was obtained on each side of each exon while 200 bp and 300 bp were sequenced in the 5'UTR and 3'UTR regions, respectively (Supplemental Table 2). As for the *STK11* gene, all coding exons were amplified and sequenced (exons 1 to 9).

Regarding *CHEK2*, a long range PCR using primers described previously (Sodha et al., 2002) was used to amplify exons 11 to 15. The amplification was performed with the Expand Long Template PCR system (Roche Diagnostics, IN, USA) following the manufacturer's instructions. The PCR product was then used as a template to amplify each individual exons (11-15). Only 95 breast cancer individuals were analysed for the *CHEK2* gene due to material availability. Intronic primers used for amplification of exons 2 to 10 were listed in Dufault and coll. (Dufault et al., 2004), while amplification of the non-coding

exon 1 was carried using these intronic primers : F5'-CTCAACGTTGCTTCCCTCAG-3' and F5'-GACACGCAGAACATCCTCCT-3'. More information on the regions analyzed here is available in Supplemental Table 2. Amplification of variant-containing amplicons of *CHEK2*, *STK11* and *PALB2* from healthy unrelated French Canadian individuals was also conducted to assess the frequency of variants identified in breast cancer individuals.

A fluorescence-based direct sequencing method using BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) was used to analyze cDNA and genomic DNA from breast cancer individuals as well as genomic DNA from healthy individuals. Sequences were run on an Applied Biosystems 3730xl DNA Analyzer and sequence data were analyzed using the Staden Pregap4 and Gap4 programs. The segregation of deleterious *CHEK2* and *STK11* variants identified in breast cancer individuals was also analyzed in carrier's relatives for which DNA material was available.

Hardy-Weinberg equilibrium (HWE) was used as genotyping quality-control check. Each single nucleotide polymorphism (SNP) was tested for HWE deviation by means of a Chi<sup>2</sup> test. All p-values were two-sided with 1 degree of freedom. In order to estimate the genetic association of each SNP with breast cancer, differences in allele frequencies between both sample sets were tested using a Chi<sup>2</sup> test. P-values less than 0.05 were initially considered as significant. However, to address the issue of multiple comparison, Bonferroni correction was applied for each estimate of genetic association (corrected P-value threshold:  $P < 7.25 \times 10^{-4}$ ).

#### ***Exon deletion analysis***

Multiplex ligation-dependent probe amplification (MLPA) was performed with genomic DNA samples from breast cancer individuals using the SALSA MLPA kit P101 STK11 and SALSA MLPA kit P057 FANCD2-PALB2 kits (MRC-Holland b.v., Amsterdam, The Netherlands) following manufacturer's recommendations. These MLPA kits cover all *STK11* and *PALB2* exons. Samples were analyzed using a capillary electrophoresis (ABI 3100 genetic analyzer; Applied Biosystems) and fragment analysis tools.

#### ***Conservation in other species and computational analyses***

Protein alignment was performed using data extracted from the UCSC database, and both the SIFT and PolyPhen web-based softwares were used to predict the effect of amino acid substitutions. All tests were run under default threshold values.

## Results

The entire coding sequence of the breast cancer susceptibility genes *CHEK2*, *STK11* and *PALB2* was analyzed in almost 100 breast cancer individuals, each coming from a distinct family with high risk of breast cancer. This analysis led to the identification of a total of 17, 25 and 27 variants detected in the *CHEK2*, *STK11* and *PALB2* genes, respectively (Table 1). *CHEK2* analysis led to the identification of the deleterious c.1100delC mutation in two high-risk individuals. The Phe354Leu *STK11* mutation reported to be the causative mutation in a PJS family was also identified. MLPA analysis was also successfully conducted for 90 and 95 breast cancer individuals for the *STK11* and *PALB2* genes, respectively, and did not reveal any genomic rearrangements in these two genes.

### *CHEK2* sequence analysis

A total of 17 variants were identified in the *CHEK2* gene (Table 1). Among these changes, four variants are located in the coding region and 5 in the introns while 6 and 2 are located in the 3' and 5'UTR regions, respectively. Nine unreported variations were identified in this study, five of them located in the 3'UTR region, three in the introns and the other being a conservative substitution located in exon 15. With the exception of the c.1100delC variant, exonic variants located in the coding region (c.252A>G, c.1556C>T and c.1608A>G) do not result in amino acid change. The c.1100delC mutation, located in the kinase domain, creates a frameshift in exon 11 that leads to a premature stop at codon 381. We found the c.1100delC mutation in 2/95 (2%) cases from non-*BRCA1/2* breast cancer families while this mutation was not observed among healthy control individuals. From these two c.1100delC carriers, one was diagnosed with breast cancer at the early age of 39. Analysis of other family members for which material was available revealed that another c.1100delC carrier was affected with breast cancer while the two non-carriers were unaffected (Figure 1 panel A). The second individual from our cohort for which c.1100delC mutation was identified was diagnosed with a bilateral breast cancer, both before 50 years. Another individual from this family was available and found to be an unaffected carrier of the c.1100delC mutation (Figure 1 panel B). In terms of allele frequency, none of the 17 variants identified in the *CHEK2* gene shows a significant difference between breast cancer and unaffected individuals. Of these, six display a significant deviation from HWE either in the cases and/or unaffected groups, with an overrepresentation of rare homozygote

genotypes, which can be expected with relatively small sample sets and rare variants.

### ***STK11* sequence analysis**

Analysis of the *STK11* gene in 96 unrelated breast cancer individuals led to the identification of 25 variants (Table 1). Twelve variants are novel and not reported in dbSNP database (Build 130). Despite the fact that this analysis did not lead to the identification of germline deleterious truncating mutations, 21 intronic variants and 4 rare exonic variants could be identified (Table 1). Among the 21 intronic variants, five are located within 50 bp from the nearest exon/intron junction, one being 7 bp from the junction (c.2035+7G>C). While the two first exonic variants are located in the 5'UTR region (c.-311C>T and c.-127T>C), the other two (c.1062C>T and c.1185A>G) are located in the coding region, including one resulting in an amino acid change (c.1062C>T, Phe354Leu). The segregation of the c.1062C>T variant with breast cancer was thus analyzed in the sole carrier's family (Figure 1 panel C). An affected sister diagnosed at 51 years could be genotyped and was found non-carrier of the variant allele. Analysis of variants identified in *STK11* in a series of unaffected individuals suggested a difference in allele frequency for the c.375-49G>A (rs34928889; P = 0.032) and the c.464+32insGGGGCC (rs58579265; P = 0.018) intronic variants located within introns 2 and 3, respectively. However, these associations do not reach statistical significance following Bonferroni correction. Regarding the distribution of *STK11* genotypes, deviation from HWE is observed in both the breast cancer cases and unaffected individuals for the common variants 11, 12 and 13 showing overrepresentation of the homozygote variant genotypes. Significant deviation from HWE is also observed in unaffected individuals for variant 15, with a lower number of homozygote variant genotypes than expected while genotype distribution for this variant follows HWE in the breast cancer sample set.

Many alternatively spliced mRNA species have been reported in the *STK11* gene (Aretz et al., 2005; Hearle et al., 2006; Mehenni et al., 2006). As several intronic variants have been identified in this study, cDNA from breast cancer individuals was sequenced in order to verify for the presence of aberrantly spliced mRNA isoforms. Both sequencing chromatograms and gel electrophoresis did not reveal any major aberrantly spliced mRNA, indicating that none of the *STK11* variants predicted to potentially alter SR proteins binding

sites or strength of potential splice sites seem to have an impact on mRNA splicing (data not shown).

The functionality and potential importance of the Phe354Leu sequence variant identified was evaluated by aligning the STK11 amino acid sequences of relevant species. As shown in Table 2, alignment of the respective sequences showed that the common allele (Phe) of this variant seems to be highly conserved, with the exception of *Ornithorhynchus anatinus* and *Gallus gallus*. Further evaluation of the potential effect of this amino acid change, located outside the catalytic domain of the STK11 protein, reveals that it is predicted to be tolerated by the SIFT program while the PolyPhen program predicts that the Phe354Leu variant could be possibly damaging (Table 3).

### ***PALB2* sequence analysis**

Out of the 27 variants identified in the *PALB2* gene, twelve are coding substitutions while one variant is located in the 5'UTR region, the remaining 14 nucleotide changes being located within introns (Table 1). Among the 14 intronic variants, one rare novel variant is located 8 bp from exon 12 (c.3202-8G>T). Twelve of the 27 *PALB2* variants are not reported in dbSNP database (Build 130). No significant difference was observed between both sample sets based on allele frequencies. In terms of HWE, 5 of the 27 variations show a deviation in the healthy individuals, with overrepresentation of the rare homozygous genotype. Similarly, the deep intronic c.2587-190A>C variant and the exonic c.2590C>T variant (rs45568339) show overrepresentation of the rare homozygote genotype in breast cancer individuals.

Within the *PALB2* gene, 8 variants result in an amino acid change. Two residues (Leu939 and Gly998) are completely conserved in distant species suggesting that these residues are under strong functional constraint or may affect the protein conformation (Table 2). For the Glu672 and Pro864 residues, modest level of conservation is observed while either a Val or Ile residue is found for the Val932Met variant. Interestingly, the human Val425 residue seems to be relatively highly conserved with a Leu residue found at the corresponding position for all other orthologues (Table 2). Evaluation of the potential effect of *PALB2* amino acid changes on protein function reveals that Leu939Trp and Glu998Gly are predicted to be deleterious by both programs (SIFT and PolyPhen), while the Asp219Gly, Glu672Gln and Pro864Ser amino acid changes are predicted to affect the protein function

by at least one of the two programs (Table 3). Albeit none of the variants identified in this study is located within a domain, the Gly998Glu variant is located near the second WD40 repeat (residues 1008-1050) (Erkko et al., 2007) known to be a common protein-protein interaction domain.



## Discussion

Residual familial risk of breast cancer not caused by *BRCA1* or *BRCA2* genes is recognized to be explained by a polygenic or high-risk heterogeneity model. Recent studies on new breast cancer susceptibility genes tend to demonstrate that multiple rare alleles are responsible for a proportion of breast cancer susceptibility. Evaluation of the contribution of known breast cancer susceptibility genes is imperative, especially in a founder population. Specific genetic alterations in a given gene can be more frequent in an isolated and small founder populations, such as the French Canadian population (Vézina et al., 2005; Fackenthal et al., 2007; Simard et al., 2007). Although no deleterious truncating mutation was identified in the *PALB2* gene, this analysis led to the identification of the previously reported breast cancer-associated *CHEK2* c.1100delC mutation in two high-risk breast cancer individuals, and the Phe354Leu *STK11* mutation reported to be the causative mutation in a PJS family (Amos et al., 2004) was also identified in another family.

As alternative spliced mRNA species were reported in *STK11* (Aretz et al., 2005; Hearle et al., 2006; Mehenni et al., 2006), cDNA analysis was also conducted for this gene. Both sequencing chromatograms and gel electrophoresis were thoroughly analyzed to detect any aberrant spliced forms, leading us to conclude that no major spliced *STK11* mRNAs are highly expressed in the cDNA material analyzed. As demonstrated previously by our group, alternative transcripts expressed at low levels (~10% compared to the WT form) can be detected by sequencing chromatogram analysis (Durocher et al., 2006). However, we cannot exclude the presence of minor, weakly expressed forms which could have not been detected here. In addition, both the *PALB2* and *STK11* genes were also analyzed for large genomic deletions/rearrangements and no such alteration was identified. We can state that such large deletions must be rare in the French Canadian population, our study having nearly 90% chance of observing at least one such deletion or rearrangement if the frequency of these alterations was 2% in high-risk French Canadian breast cancer families. This is concordant with other studies in which no *PALB2* deletions in French Canadian (Foulkes et al., 2007), Finnish (Pylkäs et al., 2008) and Dutch (Ameziane et al., 2009) breast cancer individuals were identified. As for the *STK11* gene, large genomic deletions were found exclusively in PJS patients (Aretz et al., 2005; Hearle et al., 2006; Volikos et al., 2006).

Germline *CHEK2* mutations are responsible for LFS (Bell et al., 1999). High penetrance of cancer was reported in LFS females and particularly in the 16-45 age-class, for which breast cancer represents 80% of all cancer cases in this group (Chompret et al., 2000). The c.252A>G (Glu84Glu) variant identified in our study was previously reported as a neutral variant in LFS patient (Bell et al., 1999), in unselected breast cancer (Kilpivaara et al., 2004), in breast cancer families (Schutte et al., 2003) as well as in French Canadian breast cancer families (Novak et al., 2008). Our study led to the identification of the c.1100delC mutation in two high-risk breast cancer individuals. This mutation initially associated with a two-fold increase of breast cancer risk (Vahteristo et al., 2002; CHEK2 Breast Cancer Case-Control Consortium, 2004) was thereafter suggested to increase the risk of breast cancer three- to five-fold (Weischer et al., 2008). Results obtained from segregation in our breast cancer individuals are in concordance with the reported increased risk of breast cancer and the incomplete penetrance of the c.1100delC mutation, an unaffected mutation carrier being identified. This protein truncating mutation was shown to abolish the kinase activity of CHEK2 (Wu et al., 2001), and a lower level of CHEK2 expression was reported in tumors of c.1100delC carriers (Vahteristo et al., 2002). Our results are in agreement with those reported from the New York Cancer Project where a low prevalence of the c.1100delC has been reported in breast cancer cases (1.0%) and in controls (0.3%) (Offit et al., 2003), as well as with results from other studies conducted in early onset breast cancer (Ghadirian et al., 2009) and in non-*BRCA1/2* breast cancer families (Novak et al., 2008) from the French Canadian population.

Mutations in the *STK11* gene are responsible for PJS, a cancer predisposing syndrome associated with an increased risk of breast cancer. Cumulative risk of breast cancer in PJS patients associated with mutations in the *STK11* gene has been estimated to be 45-54% (Giardiello et al., 2000; Hearle et al., 2006). Our analysis of the *STK11* breast cancer susceptibility gene in individuals from French Canadian families with high risk of breast cancer led to the identification of one missense variation (Phe354Leu) and a novel silent variant located in exon 9 (c.1185A>G; Thr395Thr). Missense mutations are believed to be responsible for approximately one fifth of the *STK11* mutations found in PJS (Launonen, 2005). Conflicting results were published on the functional implication of the Phe354Leu variant. This variation was first identified in a colorectal adenocarcinoma tumor (Dong et

al., 1998). It was thereafter suggested to be a rare polymorphism on the basis that this variation has been found in Korean left-sided colorectal cancers (6.3%) as well as in cancer-free controls (5.6%) from the same population (Launonen et al., 2000). Nonetheless, Phe354Leu was reported to cosegregate with the disease in one PJS family (Amos et al., 2004).

Our results suggest a possible association of two *STK11* intronic variants with breast cancer. However, following Bonferroni correction, these associations do not reach statistical significance. Furthermore, as association of these variants could be indirect, we should keep in mind that these variants may be in LD with the causal variant, and the organization of variations in *STK11* as well as in the *CHEK2* and *PALB2* genes might be useful to detect such indirect association (Supplemental Figure 1). Intronic variants located within 50 bp from exon-intron junctions, as well as other variants identified in this study, could affect mRNA splicing. Analysis of cDNA from breast cancer individuals did not identify such alteration, thus indicating that the *STK11* variants identified in this study do not have a significant impact on mRNA splicing. Our results are in agreement with those obtained in other studies conducted on the *STK11* gene where mRNA splicing defects were reported in PJS cases (Launonen, 2005) but not in breast cancer individuals (Forster et al., 2000).

The PALB2 protein is an integral component of the BRCA complex which acts as a linker between BRCA1 and BRCA2 (Sy et al., 2009a). Biallelic *PALB2* mutations were associated with the FA-N subtype (Reid et al., 2007; Xia et al., 2007) and monoallelic truncating *PALB2* mutations were reported to confer higher risk of breast cancer (Rahman et al., 2007; Erkkö et al., 2008). From the twelve coding variants identified in our study, eleven have been evaluated in a breast cancer context, and none was reported to be associated with breast cancer (Erkkö et al., 2007; Rahman et al., 2007; Tischkowitz et al., 2007; Cao et al., 2009; García et al., 2009; Sluiter et al., 2009). Among the missense variants identified in this study, Pro864Ser, Val932Met, Leu939Trp and Gly998Glu are located within the region required for BRCA2 interaction (residues 850-1186) but outside the WD40 repeats (Erkkö et al., 2007; Sy et al., 2009b). It is noteworthy that we did not identify the Gln775X French Canadian founder mutation previously reported and found in approximately 0.5% of unselected early-onset breast cancers from this population (Foulkes

et al., 2007; Ghadirian et al., 2009).

A recent analysis conducted on randomly selected breast cancer cases from China suggested that the c.2586+58C>T variant (rs249954) was associated with a 36% increase of breast cancer risk (Chen et al., 2008). However, data from our study do not support the increased risk associated with this intronic variant. Intronic variants located in the vicinity of exon boundaries (4 and 5 nt from exons), resulting in splicing defect, frame shift and premature termination of the protein, were identified in *FANCN* patients (Reid et al., 2007). However, few intronic variants were identified in individuals from Spanish breast cancer families, and are predicted to be most unlikely pathogenic (García et al., 2009). In agreement with this observation, the *in silico* analysis performed in our study favors the classification of *PALB2* intronic variants as neutral (Splice Site Prediction by Neural Network; data not shown). Difference in allele frequency and associations found between populations for *PALB2* mutations, namely for the Finnish, Chinese and French Canadian populations could be attributable to the founder effect found in some of these populations, to the type of breast cancer evaluated in different studies (sporadic cases vs breast cancer cases from high-risk families) or to experimental variations.

We report here the identification of the *CHEK2* c.1100delC mutation in 2 of our breast cancer individuals (MAF = 0.011). Moreover, the Phe354Leu variant, reported to cosegregate with the disease in a PJS family, was also found once in our cohort of high-risk individuals drawn from non-*BRCA1/BRCA2* French Canadian families. Our analysis did not reveal any deleterious truncating mutation in the *PALB2* gene, nor a significant difference in allele frequency between breast cancer and unaffected individuals. Splicing defects being reported in PJS patients, and our results suggesting a possible association of two intronic variants with breast cancer, further analysis was conducted on cDNA for the *STK11* gene. However, no aberrant mRNA species in breast cancer individuals was identified. In addition, MLPA analysis performed on both *STK11* and *PALB2* genes did not demonstrate any large genomic deletions or rearrangements. Therefore, we can conclude that deleterious germline mutations in these three genes are rare and do not contribute to a substantial proportion of the families with high risk of breast cancer from the French Canadian population. Indeed, assuming a Poisson distribution of the probability of finding a deleterious mutation in our cohort, if the frequency of such deleterious mutation was 2%,

we would have had 90% or greater chance of observing at least one such mutation. Much larger cohorts will be required to define a possible association of *STK11* intronic variants and *PALB2* missense variants with a modest or weak impact on breast cancer susceptibility.

## Discussion

Residual familial risk of breast cancer not caused by *BRCA1* or *BRCA2* genes is recognized to be explained by a polygenic or high-risk heterogeneity model. Recent studies on new breast cancer susceptibility genes tend to demonstrate that multiple rare alleles are responsible for a proportion of breast cancer susceptibility. Evaluation of the contribution of known breast cancer susceptibility genes is imperative, especially in a founder population. Specific genetic alterations in a given gene can be more frequent in an isolated and small founder populations, such as the French Canadian population.<sup>29,38,39</sup> Although no deleterious truncating mutation was identified in the *PALB2* gene, this analysis led to the identification of the previously reported breast cancer-associated *CHEK2* c.1100delC mutation in two high-risk breast cancer individuals, and the Phe354Leu *STK11* mutation reported to be the causative mutation in a PJS family<sup>40</sup> was also identified in another family. HWE was used mainly as genotyping quality-control check and to detect potential deleterious effect of homozygote variant genotypes. However, a deviation from HWE in either breast cancer or unaffected individuals could also be due to the recent emergence of a variant in a population, and this does not necessarily reflect a potential effect of a variant, especially with relatively small sample sets and rare variants. The occurrence of a homozygote variant genotype might explain the deviation from HWE for most deviations seen in this study, especially for rare variants. Otherwise, in some instances the occurrence of heterozygote genotypes in amplicons bearing variants showing deviation from HWE allow to exclude the amplification of a single allele as a cause for these deviations.

As alternative spliced mRNA species were reported in *STK11*,<sup>14,36,37</sup> cDNA analysis was also conducted for this gene. Both sequencing chromatograms and gel electrophoresis were thoroughly analyzed to detect any aberrant spliced forms, leading us to conclude that no major spliced *STK11* mRNAs are highly expressed in the cDNA material analyzed. As demonstrated previously by our group, alternative transcripts expressed at low levels (~10% compared to the WT form) can be detected by sequencing chromatogram analysis.<sup>32</sup> However, we cannot exclude the presence of minor, weakly expressed forms which could have not been detected here. In addition, both the *PALB2* and *STK11* genes were also analyzed for large genomic deletions/rearrangements and no such alteration was identified. We can state that such large deletions must be rare in the French Canadian population, our

study having nearly 90% chance of observing at least one such deletion or rearrangement if the frequency of these alterations was 2% in high-risk French Canadian breast cancer families. This is concordant with two other studies in which no *PALB2* deletions in French Canadian<sup>27</sup> and Finnish<sup>28</sup> breast cancer individuals were identified. As for the *STK11* gene, large genomic deletions were found exclusively in PJS patients.<sup>14,36,41</sup>

Germline *CHEK2* mutations are responsible for LFS.<sup>5</sup> High penetrance of cancer was reported in LFS females and particularly in the 16-45 age-class, for which breast cancer represents 80% of all cancer cases in this group.<sup>42</sup> The c.252A>G (Glu84Glu) variant identified in our study was previously reported as a neutral variant in LFS patient,<sup>5</sup> in unselected breast cancer,<sup>43</sup> in breast cancer families<sup>44</sup> as well as in French Canadian breast cancer families.<sup>45</sup> Our study led to the identification of the c.1100delC mutation in two high-risk breast cancer individuals. This mutation initially associated with a two-fold increase of breast cancer risk<sup>7,10</sup> was thereafter suggested to increase the risk of breast cancer three- to five-fold.<sup>11</sup> Results obtained from segregation in our breast cancer individuals are in concordance with the reported increased risk of breast cancer and the incomplete penetrance of the c.1100delC mutation, an unaffected mutation carrier being identified. This protein truncating mutation was shown to abolish the kinase activity of CHEK2,<sup>8</sup> and a lower level of CHEK2 expression was reported in tumors of c.1100delC carriers.<sup>7</sup> Our results are in agreement with those reported from the New York Cancer Project where a low prevalence of the c.1100delC has been reported in breast cancer cases (1.0%) and in controls (0.3%),<sup>46</sup> as well as with results from another study conducted in non-*BRCAl/2* French Canadian breast cancer families.<sup>45</sup>

Mutations in the *STK11* gene are responsible for PJS, a cancer predisposing syndrome associated with an increased risk of breast cancer. Cumulative risk of breast cancer in PJS patients associated with mutations in the *STK11* gene has been estimated to be 45-54%.<sup>13,14</sup> Our analysis of the *STK11* breast cancer susceptibility gene in individuals from French Canadian families with high risk of breast cancer led to the identification of one missense variation (Phe354Leu) and a novel silent variant located in exon 9 (c.1185A>G; Thr395Thr). Missense mutations are believed to be responsible for approximately one fifth of the *STK11* mutations found in PJS.<sup>47</sup> Conflicting results were published on the functional implication of the Phe354Leu variant. This variation was first identified in a colorectal

adenocarcinoma tumor.<sup>48</sup> It was thereafter suggested to be a rare polymorphism on the basis that this variation has been found in Korean left-sided colorectal cancers (6.3%) as well as in cancer-free controls (5.6%) from the same population.<sup>49</sup> Nonetheless, Phe354Leu was reported to cosegregate with the disease in one PJS family.<sup>40</sup>

Our results suggest a possible association of two *STK11* intronic variants with breast cancer. However, following Bonferroni correction, these associations do not reach statistical significance. Furthermore, as association of these variants could be indirect, we should keep in mind that these variants may be in LD with the causal variant, and the organization of variations in *STK11* as well as in the *CHEK2* and *PALB2* genes might be useful to detect such indirect association (Supplemental Figure 1). Intronic variants located within 50 bp from exon-intron junctions, as well as other variants identified in this study, could affect mRNA splicing. Analysis of cDNA from breast cancer individuals did not identify such alteration, thus indicating that the *STK11* variants identified in this study do not have a significant impact on mRNA splicing. Our results are in agreement with those obtained in other studies conducted on the *STK11* gene where mRNA splicing defects were reported in PJS cases<sup>47</sup> but not in breast cancer individuals.<sup>16</sup>

The *PALB2* protein is an integral component of the BRCA complex which acts as a linker between BRCA1 and BRCA2.<sup>19</sup> Biallelic *PALB2* mutations were associated with the FA-N subtype<sup>21,22</sup> and monoallelic truncating *PALB2* mutations were reported to confer higher risk of breast cancer.<sup>23,24</sup> From the twelve coding variants identified in our study, eleven have been evaluated in a breast cancer context, and none was reported to be associated with breast cancer.<sup>23,25,26,50-52</sup> Among the missense variants identified in this study, Pro864Ser, Val932Met, Leu939Trp and Gly998Glu are located within the region required for BRCA2 interaction (residues 850-1186) but outside the WD40 repeats.<sup>26,53</sup> It is noteworthy that we did not identify the Gln775X French Canadian founder mutation previously reported and found in approximately 0.5% of unselected early-onset breast cancers from this population.<sup>27</sup>

A recent analysis conducted on randomly selected breast cancer cases from China suggested that the c.2586+58C>T variant (rs249954) was associated with a 36% increase of breast cancer risk.<sup>54</sup> However, data from our study do not support the increased risk



associated with this intronic variant. Intronic variants located in the vicinity of exon boundaries (4 and 5 nt from exons), resulting in splicing defect, frame shift and premature termination of the protein, were identified in *FANCN* patients.<sup>22</sup> However, few intronic variants were identified in individuals from Spanish breast cancer families, and are predicted to be most unlikely pathogenic.<sup>25</sup> In agreement with this observation, the *in silico* analysis performed in our study favors the classification of *PALB2* intronic variants as neutral (Splice Site Prediction by Neural Network; data not shown). Difference in allele frequency and associations found between populations for *PALB2* mutations, namely for the Finnish, Chinese and French Canadian populations could be attributable to the founder effect found in some of these populations, to the type of breast cancer evaluated in different studies (sporadic cases vs breast cancer cases from high-risk families) or to experimental variations.

We report here the identification of the *CHEK2* c.1100delC mutation in 2 of our breast cancer individuals (MAF = 0.011). Moreover, the Phe354Leu variant, reported to cosegregate with the disease in a PJS family, was also found once in our cohort of high-risk individuals drawn from non-*BRCA1/BRCA2* French Canadian families. Our analysis did not reveal any deleterious truncating mutation in the *PALB2* gene, nor a significant difference in allele frequency between breast cancer and unaffected individuals. Splicing defects being reported in PJS patients, and our results suggesting a possible association of two intronic variants with breast cancer, further analysis was conducted on cDNA for the *STK11* gene. However, no aberrant mRNA species in breast cancer individuals was identified. In addition, MLPA analysis performed on both *STK11* and *PALB2* genes did not demonstrate any large genomic deletions or rearrangements. Therefore, we can conclude that deleterious germline mutations in these three genes are rare and do not contribute to a substantial proportion of the families with high risk of breast cancer from the French Canadian population. Indeed, assuming a Poisson distribution of the probability of finding a deleterious mutation in our cohort, if the frequency of such deleterious mutation was 2%, we would have had 90% or greater chance of observing at least one such mutation. However, we cannot discard the fact that a small proportion of hereditary, family-positive breast cancers could be assigned to *CHEK2*. Much larger cohorts will be required to determine the prevalence of the *CHEK2* c.1100delC mutation in the French Canadian

population as well as to define a possible association of *STK11* intronic variants and *PALB2* missense variants with a modest or weak impact on breast cancer susceptibility.

**Acknowledgements**

The authors are indebted to the participants and their families for their generosity and providing DNA samples. We would like to thank Dr Martine Dumont, Gilles Leblanc and Martine Tranchant for sample management, mutation screening, and skillful technical assistance as well as Josée Rhéaume for genetic counseling and clinical data management at the Cancer Genomics Laboratory. We thank Damian Labuda and Claudia Moreau at the Centre de Recherche de l'Hôpital Ste-Justine for providing control DNA samples. We would also like to thank Professor Bartha Maria Knoppers and her colleagues from the Centre de recherche en droit public de l'Université de Montréal for their precious help with ELSI issues related to our research program. We also appreciate advice received from ethics committees. This work was supported by the Canadian Institutes of Health Research (CIHR) through the INHERIT BRCA's research program, the Fonds de la Recherche en Santé du Québec (FRSQ)/Réseau de Médecine Génétique Appliquée (RMGA), the CURE Foundation, the Canadian Breast Cancer Research Alliance (CBCRA) and Fanconi Canada. F.G. holds a studentship from Fondation René Bussièrès, J.S. is Chairholder of the Canada Research Chair in Oncogenetics, and F.D. is a recipient of a chercheur-boursier from the Fonds de la Recherche en Santé du Québec (FRSQ) and a Research Career Award in the Health Sciences from CIHR/Rx&D Health Research Foundation.

**Authors Disclosure Statement**

No competing financial interests exist.

**Appendix**

Clinicians of INHERIT involved in this study

Paul Bessette: Service de gynécologie, Centre Hospitalier Universitaire de Sherbrooke, Fleurimont, QC, J1H 5N4, Canada;

Jocelyne Chiquette: Clinique des maladies du sein Deschênes-Fabia, Hôpital du Saint-Sacrement, Québec, QC, G1S 4L8, Canada;

Rachel Laframboise: Service de médecine génétique, CHUQ, Pavillon CHUL, Québec, QC, G1V 4G2, Canada;

Jean Lépine: Centre Hospitalier régional de Rimouski, Rimouski, QC, G5L 5T1, Canada;

Bernard Lespérance, Roxane Pichette: Service d'hémato-oncologie, Hôpital du Sacré-Cœur, Montréal, QC, H4J 1C5, Canada;

Marie Plante: Service de gynécologie, CHUQ, L'Hôtel-Dieu de Québec, Québec, QC, G1R 2J6, Canada.

## References

- Ahn J, Urist M, Prives C (2004) The Chk2 protein kinase. *DNA Repair (Amst)* 3(8-9):1039-1047.
- Alessi DR, Sakamoto K, Bayascas JR (2006) LKB1-dependent signaling pathways. *Annu Rev Biochem* 75:137-163.
- Ameziane N, van den Ouweland AM, Adank MA *et al.* (2009) Lack of large genomic deletions in BRIP1, PALB2, and FANCD2 genes in BRCA1/2 negative familial breast cancer. *Breast Cancer Res Treat* 118(3):651-653.
- Amos CI, Keitheri-Cheteri MB, Sabripour M, *et al.* (2004) Genotype-phenotype correlations in Peutz-Jeghers syndrome. *J Med Genet* 41(5):327-333.
- Antoniou AC, Easton DF (2003) Polygenic inheritance of breast cancer: Implications for design of association studies. *Genet Epidemiol* 25(3):190-202.
- Antoniou A, Pharoah PD, Narod S *et al.* (2003) 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* 72(5):1117-1130.
- Aretz S, Stienen D, Uhlhaas S *et al.* (2005) High proportion of large genomic STK11 deletions in Peutz-Jeghers syndrome. *Hum Mutat* 26(6):513-519.
- Bahassi EM, Ovesen JL, Riesenberger AL *et al.* (2008) The checkpoint kinases Chk1 and Chk2 regulate the functional associations between hBRCA2 and Rad51 in response to DNA damage. *Oncogene* 27(28):3977-3985.
- Bell DW, Varley JM, Szydlo TE *et al.* (1999) Heterozygous germ line hCHK2 mutations in Li-Fraumeni syndrome. *Science* 286(5449):2528-2531.
- Bignell GR, Barfoot R, Seal S *et al.* (1998) Low frequency of somatic mutations in the LKB1/Peutz-Jeghers syndrome gene in sporadic breast cancer. *Cancer Res* 58(7):1384-1386.
- Cao AY, Huang J, Hu Z *et al.* (2009) The prevalence of PALB2 germline mutations in BRCA1/BRCA2 negative Chinese women with early onset breast cancer or affected relatives. *Breast Cancer Res Treat* 114(3):457-462.
- CHEK2 Breast Cancer Case-Control Consortium (2004) CHEK2\*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. *Am J Hum Genet* 74(6):1175-1182.

- Chen J, Lindblom A (2000) Germline mutation screening of the STK11/LKB1 gene in familial breast cancer with LOH on 19p. *Clin Genet* 57(5):394-397.
- Chen P, Liang J, Wang Z *et al.* (2008) Association of common PALB2 polymorphisms with breast cancer risk: a case-control study. *Clin Cancer Res* 14(18):5931-5937.
- Chompret A, Brugières L, Ronsin M *et al.* (2000) P53 germline mutations in childhood cancers and cancer risk for carrier individuals. *Br J Cancer* 82(12):1932-1937.
- Dong SM, Kim KM, Kim SY *et al.* (1998) Frequent somatic mutations in serine/threonine kinase 11/Peutz-Jeghers syndrome gene in left-sided colon cancer. *Cancer Res* 58(17):3787-3790.
- Dufault MR, Betz B, Wappenschmidt B *et al.* (2004) Limited relevance of the CHEK2 gene in hereditary breast cancer. *Int J Cancer* 110(3):320-325.
- Durocher F, Labrie Y, Soucy P *et al.* (2006) Mutation analysis and characterization of ATR sequence variants in breast cancer cases from high-risk French Canadian breast/ovarian cancer families. *BMC Cancer* 6:230.
- Erkko H, Xia B, Nikkilä J *et al.* (2007) A recurrent mutation in PALB2 in Finnish cancer families. *Nature* 446(7133):316-319.
- Erkko H, Dowty JG, Nikkilä J *et al.* (2008) Penetrance analysis of the PALB2 c.1592delT founder mutation. *Clin Cancer Res* 14(14):4667-4671.
- Fackenthal JD, Olopade OI (2007) Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. *Nat Rev Cancer* 7(12):937-948.
- Forster LF, Defres S, Goudie DR *et al.* (2000) An investigation of the Peutz-Jeghers gene (LKB1) in sporadic breast and colon cancers. *J Clin Pathol* 53(10):791-793.
- Foulkes WD, Ghadirian P, Akbari MR *et al.* (2007) Identification of a novel truncating PALB2 mutation and analysis of its contribution to early-onset breast cancer in French-Canadian women. *Breast Cancer Res* 9(6):R83.
- García MJ, Fernández V, Osorio A *et al.* (2009) Analysis of FANCB and FANCN/PALB2 fanconi anemia genes in BRCA1/2-negative Spanish breast cancer families. *Breast Cancer Res Treat* 113(3):545-551.
- Ghadirian P, Robidoux A, Zhang P *et al.* (2009) The contribution of founder mutations to early-onset breast cancer in French-Canadian women. *Clin Genet* 76(5):421-426.

- Giardiello FM, Brensinger JD, Tersmette AC *et al.* (2000) Very high risk of cancer in familial Peutz-Jeghers syndrome. *Gastroenterology* 119(6):1447-1453.
- Guénard F, Labrie Y, Ouellette G *et al.* (2007) Germline mutations in the breast cancer susceptibility gene PTEN are rare in high-risk non-BRCA1/2 French Canadian breast cancer families. *Fam Cancer* 6(4):483-490.
- Hearle N, Schumacher V, Menko FH *et al.* (2006) Frequency and spectrum of cancers in the Peutz-Jeghers syndrome. *Clin Cancer Res* 12(10):3209-3215.
- Kilpivaara O, Vahteristo P, Falck J *et al.* (2004) CHEK2 variant I157T may be associated with increased breast cancer risk. *Int J Cancer* 111(4):543-547.
- Launonen V (2005) Mutations in the human LKB1/STK11 gene. *Hum Mutat* 26(4):291-297.
- Launonen V, Avizienyte E, Loukola A *et al.* (2000) No evidence of Peutz-Jeghers syndrome gene LKB1 involvement in left-sided colorectal carcinomas. *Cancer Res* 60(3):546-548.
- Li FP, Fraumeni JF Jr (1969) Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome? *Ann Intern Med* 71(4):747-752.
- Meijers-Heijboer H, van den Ouweland A, Klijn J *et al.* (2002) Low-penetrance susceptibility to breast cancer due to CHEK2(\*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet* 31(1):55-59.
- Moisan AM, Fortin J, Dumont M *et al.* (2006) No Evidence of BRCA1/2 genomic rearrangements in high-risk French-Canadian breast/ovarian cancer families. *Genet Test* 10(2):104-115.
- Novak DJ, Chen LQ, Ghadirian P *et al.* (2008) Identification of a novel CHEK2 variant and assessment of its contribution to the risk of breast cancer in French Canadian women. *BMC Cancer* 8:239.
- Offit K, Pierce H, Kirchoff T *et al.* (2003) Frequency of CHEK2\*1100delC in New York breast cancer cases and controls. *BMC Med Genet* 4:1.
- Pylkäs K, Erkkö H, Nikkilä J *et al.* (2008) Analysis of large deletions in BRCA1, BRCA2 and PALB2 genes in Finnish breast and ovarian cancer families. *BMC Cancer* 8:146.
- Rahman N, Seal S, Thompson D *et al.* (2007) PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat Genet* 39(2):165-167.

- Reid S, Schindler D, Hanenberg H *et al.* (2007) Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nat Genet* 39(2):162-164.
- Schutte M, Seal S, Barfoot R *et al.* (2003) Variants in CHEK2 other than 1100delC do not make a major contribution to breast cancer susceptibility. *Am J Hum Genet* 72(4):1023-1028.
- Simard J, Dumont M, Moisan AM *et al.* (2007) 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* 44(2):107-121.
- Sluiter M, Mew S, van Rensburg EJ (2009) PALB2 sequence variants in young South African breast cancer patients. *Fam Cancer* 8(4):347-353.
- Sodha N, Houlston RS, Bullock S *et al.* (2002) Increasing evidence that germline mutations in CHEK2 do not cause Li-Fraumeni syndrome. *Hum Mutat* 20(6):460-462.
- Stratton MR, Rahman N (2008) The emerging landscape of breast cancer susceptibility. *Nat Genet* 40(1):17-22.
- Sy SM, Huen MS, Chen J (2009a) PALB2 is an integral component of the BRCA complex required for homologous recombination repair. *Proc Natl Acad Sci U S A* 106(17):7155-7160.
- Sy SM, Huen MS, Zhu Y *et al.* (2009b) PALB2 regulates recombinational repair through chromatin association and oligomerization. *J Biol Chem* 284(27):18302-18310.
- Tischkowitz M, Xia B, Sabbaghian N *et al.* (2007) Analysis of PALB2/FANCN-associated breast cancer families. *Proc Natl Acad Sci U S A* 104(16):6788-6793.
- Vahteristo P, Bartkova J, Eerola H *et al.* (2002) A CHEK2 genetic variant contributing to a substantial fraction of familial breast cancer. *Am J Hum Genet* 71(2):432-438.
- Vézina H, Durocher F, Dumont M *et al.* (2005) Molecular and genealogical characterization of the R1443X BRCA1 mutation in high-risk French-Canadian breast/ovarian cancer families. *Hum Genet* 117(2-3):119-132.
- Volikos E, Robinson J, Aittomäki K *et al.* (2006) LKB1 exonic and whole gene deletions are a common cause of Peutz-Jeghers syndrome. *J Med Genet* 43(5):e18.



Weischer M, Bojesen SE, Ellervik C *et al.* (2008) CHEK2\*1100delC genotyping for clinical assessment of breast cancer risk: meta-analyses of 26,000 patient cases and 27,000 controls. *J Clin Oncol* 26(4):542-548.

Wu X, Webster SR, Chen J (2001) Characterization of tumor-associated Chk2 mutations. *J Biol Chem* 276(4):2971-2974.

Xia B, Sheng Q, Nakanishi K *et al.* (2006) Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Mol Cell* 22(6):719-729.

Xia B, Dorsman JC, Ameziane N *et al.* (2007) Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. *Nat Genet* 39(2):159-161.

Zhang F, Ma J, Wu J *et al.* (2009) PALB2 links BRCA1 and BRCA2 in the DNA-damage response. *Curr Biol* 19(6):524-529.

**Legend to Figure**

**Figure 1.** Pedigrees of *CHEK2* c.1100delC carriers' families (panels A and B) and *STK11* Phe354Leu carrier's family (panel C). The age at onset of cancer, age of death or current age are indicated below family members. The affected case that was sequenced first is designated by an arrow.

**Legend to Supplemental figure**

**Supplemental figure 1.** Graphical overview of pairwise linkage disequilibrium (LD) measures of  $|D'|$  (lower left part) and  $r^2$  (higher right part) for each SNPs pair identified in breast cancer cases and unaffected individuals. All SNPs are denoted numerically with reference to Table 1. Distance between first and last SNPs for each gene is indicated on each LD representation. The LDA program was used to calculate pairwise LD ( $D'$  and  $r^2$ ) values.

Table 1  
Observed sequence variations of *PALB2*, *STK11* and *CHEK2* genes and their genotype in breast cancer and unaffected control individuals.

Gene	SNP	SNP ID <sup>1</sup>	dbSNP ID	Amino acid change	Location	Series	Number of individuals	Common homozygote (expected) <sup>2</sup>	Heterozygote (expected) <sup>2</sup>	Rare homozygote (expected) <sup>2</sup>	Dev from HWE p-value
<b>CHEK2</b>	1	c.-72-136G>C	rs9613669	-	5'UTR	Cases	95	38 (39.8)	47 (43.4)	10 (11.8)	0.415
	2	c.-72-48G>A	rs2236141	-	5'UTR	Controls	89	34 (35.9)	45 (41.3)	10 (11.9)	0.393
	3	c.-7+59delT	N/A	-	Intron 1	Cases	90	66 (67.4)	28 (25.3)	1 (2.4)	0.291
	4	c.252A>G	rs1805129	Glu84Glu	Exon 2	Controls	95	58 (58.4)	29 (28.2)	3 (3.4)	0.786
	5	c.444+24C>T	N/A	-	Intron 3	Cases	91	93 (93.0)	2 (2.0)	0 (0.0)	0.917
	6	c.684-74G>C	N/A	-	Intron 5	Controls	95	89 (89.0)	2 (2.0)	0 (0.0)	0.916
	7	c.846+384A>G	rs5752774	-	Intron 7	Cases	95	90 (90.1)	5 (4.9)	0 (0.1)	0.792
	8	c.846+436T>C	rs2073327	-	Intron 7	Controls	95	91 (91.0)	4 (3.9)	0 (0.0)	0.834
	9	c.1100delC	N/A	Thr367MetfsX14	Exon 11	Cases	95	94 (94.0)	1 (1.0)	0 (0.0)	0.959
	10	c.1566C>T	N/A	Pro522Pro	Exon 15	Controls	95	95 (95.0)	0 (0.0)	0 (0.0)	-
	11	c.1608A>G	rs17886242	Pro536Pro	Exon 15	Cases	94	94 (93.0)	0 (0.0)	0 (0.0)	0.959
	12	c.*6G>A	N/A	-	3'UTR	Controls	94	93 (92.0)	1 (1.0)	0 (0.0)	<0.001
	13	c.*18C>T	rs17884403	-	3'UTR	Cases	95	89 (87.1)	3 (6.7)	2 (0.1)	<0.001
	14	c.*76delT	N/A	-	3'UTR	Controls	94	94 (94.0)	0 (0.0)	0 (0.0)	0.959
	15	c.*85T>C	N/A	-	3'UTR	Cases	95	93 (92.0)	1 (1.0)	0 (0.0)	<0.001
	16	c.*88T>A	N/A	-	3'UTR	Controls	94	93 (93.0)	1 (1.0)	0 (0.0)	0.959
	17	c.*99G>A	N/A	-	3'UTR	Cases	94	93 (93.0)	0 (0.0)	0 (0.0)	<0.001
<b>STK11</b>	1	c.-311C>T	N/A	-	5' UTR	Cases	96	92 (92.0)	1 (1.0)	0 (0.0)	0.958
	2	c.-127T>C	N/A	-	5' UTR	Controls	88	95 (95.0)	1 (1.0)	0 (0.0)	0.959
	3	c.290+36G>T	rs3764640	-	Intron 1	Cases	96	86 (86.0)	2 (2.0)	0 (0.0)	0.914
	4	c.290+78C>T	rs3764641	-	Intron 1	Controls	93	92 (92.0)	4 (3.9)	0 (0.0)	0.835

5	c.290+84C>T	N/A	-	Intron 1	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.959
6	c.290+177G>C	N/A	-	Intron 1	Controls	93	93 (93.0)	0 (0.0)	0 (0.0)	-
7	c.374+24G>T	rs2075604	-	Intron 2	Cases	96	94 (94.0)	2 (2.0)	0 (0.0)	0.918
8	c.374+141G>A	N/A	-	Intron 2	Controls	93	93 (93.0)	0 (0.0)	0 (0.0)	-
9	c.374+200delC	N/A	-	Intron 2	Cases	96	68 (65.8)	23 (27.3)	5 (2.8)	0.121
10	c.374+272G>A	N/A	-	Intron 2	Controls	94	75 (74.2)	17 (18.7)	2 (1.2)	0.390
11	c.375-132A>G	rs2075605	-	Intron 2	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.959
12	c.375-49G>A	rs34928889	-	Intron 2	Controls	94	94 (94.0)	0 (0.0)	0 (0.0)	-
13	c.464+32 insGGGGGCC	rs58579265	-	Intron 3	Cases	96	94 (94.0)	2 (2.0)	0 (0.0)	0.918
14	c.465-194G>A	N/A	-	Intron 3	Controls	94	92 (92.0)	2 (2.0)	0 (0.0)	0.917
15	c.465-51T>C	rs2075606	-	Intron 3	Cases	96	87 (85.3)	7 (10.4)	2 (0.3)	0.001
16	c.735-51C>T	rs9282860	-	Intron 5	Controls	93	80 (77.7)	10 (14.6)	3 (0.7)	0.002
17	c.863-224C>T	N/A	-	Intron 6	Cases	96	48 (40.0)	28 (43.9)	20 (12.0)	<0.001
18	c.920+7G>C	rs2075607	-	Intron 7	Controls	93	32 (26.9)	36 (46.2)	25 (19.9)	0.032
19	c.920+32G>A	N/A	-	Intron 7	Cases	96	83 (78.8)	8 (16.3)	5 (0.8)	<0.001
20	c.920+263A>G	rs60977562	-	Intron 7	Controls	90	85 (84.1)	4 (5.8)	1 (0.1)	0.003
21	c.920+311C>T	rs2075608	-	Intron 7	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.959
22	c.1062C>G	rs59912467	Phe354Leu	Exon 8	Controls	94	94 (94.0)	0 (0.0)	0 (0.0)	-
23	c.1185A>G	N/A	Thr395Thr	Exon 9	Cases	96	51 (51.0)	38 (37.9)	7 (7.0)	0.983
24	c.*16+110C>T	rs2277746	-	Intron 9	Controls	94	32 (39.6)	58 (42.8)	4 (11.59)	<0.001
25	c.*16+113G>A	N/A	-	Intron 9	Cases	96	86 (86.3)	10 (9.5)	0 (0.3)	0.590
1	c.-47G>A	rs8053188	-	Exon 1	Controls	96	81 (81.1)	4 (3.9)	0 (0.1)	0.824
2	c.49-352delGA	N/A	-	Intron 1	Cases	85	94 (94.0)	2 (2.0)	0 (0.0)	0.918
3	c.656A>G	rs45594034	Asp219Gly	Exon 4	Controls	88	85 (85.0)	3 (3.0)	0 (0.0)	0.871
4	c.1194G>A	rs61755173	Val398Val	Exon 4	Cases	96	61 (59.4)	29 (32.2)	6 (4.4)	0.324
					Controls	90	52 (50.6)	31 (33.8)	7 (5.6)	0.440
					Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.959
					Controls	89	87 (87.0)	2 (2.0)	0 (0.0)	0.915
					Cases	96	25 (22.5)	43 (48.0)	28 (25.5)	0.311
					Controls	93	22 (20.4)	43 (46.3)	28 (26.4)	0.491
					Cases	96	61 (61.0)	31 (31.1)	4 (4.0)	0.980
					Controls	94	54 (52.9)	33 (35.3)	7 (5.9)	0.536
					Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.959
					Controls	89	88 (88.0)	1 (1.0)	0 (0.0)	0.957
					Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.959
					Controls	86	86 (86.0)	0 (0.0)	0 (0.0)	-
					Cases	96	56 (54.0)	32 (36.0)	8 (6.0)	0.276
					Controls	88	53 (52.6)	30 (30.9)	5 (4.6)	0.783
					Cases	96	94 (94.0)	2 (2.0)	0 (0.0)	0.918
					Controls	90	87 (87.0)	3 (3.0)	0 (0.0)	0.872
					Cases	96	92 (92.0)	4 (3.9)	0 (0.0)	0.835
					Controls	92	87 (87.1)	5 (4.9)	0 (0.1)	0.789
					Cases	96	88 (88.2)	8 (7.7)	0 (0.2)	0.670
					Controls	95	88 (87.2)	6 (7.7)	1 (0.2)	0.034
					Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.959
					Controls	88	88 (88.0)	0 (0.0)	0 (0.0)	-
					Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.959
					Controls	95	95 (95.0)	0 (0.0)	0 (0.0)	-

5	c.1273G>A	N/A	Val425Met	Exon 4	Cases	96	93 (93.0)	3 (3.0)	0 (0.0)	0.876
6	c.1572A>G	rs45472400	Ser524Ser	Exon 4	Controls	93	92 (92.0)	1 (1.0)	0 (0.0)	0.958
7	c.1676A>G	rs152451	Gln559Arg	Exon 4	Cases	96	94 (94.0)	2 (2.0)	0 (0.0)	0.918
8	c.1684+41dupTGA	N/A	-	Intron 4	Controls	95	91 (91.0)	4 (3.9)	0 (0.0)	0.834
9	c.2014G>C	rs45532440	Glu672Gln	Exon 5	Cases	96	80 (78.8)	14 (16.3)	2 (0.8)	0.165
10	c.2442G>A	N/A	Glu814Glu	Exon 5	Controls	94	86 (85.2)	7 (8.6)	1 (0.2)	0.076
11	c.2514+1361T>C	N/A	-	Intron 5	Cases	96	94 (94.0)	2 (2.0)	0 (0.0)	0.918
12	c.2586+58C>T	rs249954	-	Intron 6	Controls	93	92 (92.0)	1 (1.0)	0 (0.0)	0.958
13	c.2587-404A>G	rs13339215	-	Intron 6	Cases	96	88 (88.2)	8 (7.7)	0 (0.2)	0.670
14	c.2587-190A>C	N/A	-	Intron 6	Controls	94	91 (90.0)	2 (3.9)	1 (0.0)	<0.001
15	c.2590C>T	rs45568339	Pro864Ser	Exon 7	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.959
16	c.2748+15C>T	N/A	-	Intron 7	Controls	84	94 (94.0)	0 (0.0)	0 (0.0)	-
17	c.2794G>A	rs45324036	Val932Met	Exon 8	Cases	86	84 (84.0)	36 (33.9)	4 (5.0)	0.547
18	c.2816T>G	rs45478192	Leu939Trp	Exon 8	Controls	94	92 (91.1)	2 (2.7)	0 (0.0)	0.418
19	c.2834+328C>T	N/A	-	Intron 8	Cases	96	95 (95.0)	23 (21.7)	1 (1.6)	0.572
20	c.2835-27C>T	N/A	-	Intron 8	Controls	92	72 (72.6)	28 (23.7)	0 (2.1)	0.085
21	c.2993G>A	rs45551636	Gly998Glu	Exon 9	Cases	96	64 (66.1)	3 (4.9)	1 (0.1)	<0.001
22	c.2997-108insT	N/A	-	Intron 9	Controls	94	92 (91.1)	0 (0.0)	0 (0.0)	-
23	c.3113+196A>G	N/A	-	Intron 10	Cases	96	94 (94.0)	3 (4.9)	0 (0.0)	<0.001
24	c.3113+204G>A	rs8060124	-	Intron 10	Controls	91	91 (91.0)	0 (0.0)	0 (0.0)	-
25	c.3113+325A>G	rs8061190	-	Intron 10	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.959
26	c.3202-8G>T	N/A	-	Intron 11	Controls	84	84 (84.0)	0 (0.0)	0 (0.0)	-
27	c.3300T>G	rs45516100	Thr1100Thr	Exon 12	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.876
					Controls	95	92 (91.0)	2 (3.9)	1 (0.0)	0.959

<sup>1</sup> According to the nomenclature of the Human Genome Variation Society

<sup>2</sup> As expected under Hardy-Weinberg equilibrium

Table 2  
Allele frequencies of variations of *PALB2*, *STK11* and *PALB2* in breast cancer and unaffected control individuals.

Gene	SNP	SNP ID	MAF cases (# of alleles)	MAF controls (# of alleles)	P-value <sup>1</sup>
<b>CHEK2</b>	1	c.-72-136G>C	0.353 (190)	0.365 (178)	0.802
	2	c.-72-48G>A	0.158 (190)	0.194 (180)	0.356
	3	c.-7+59delT	0.011 (190)	0.011 (182)	0.966
	4	c.252A>G	0.026 (190)	0.021 (190)	0.736
	5	c.444+24C>T	0.005 (190)	0.000 (190)	0.317
	6	c.684-74G>C	0.005 (190)	0.000 (180)	0.330
	7	c.846+384A>G	0.353 (190)	0.271 (188)	0.088
	8	c.846+436T>C	0.205(190)	0.197 (188)	0.838
	9	c.1100delC	0.011 (190)	0.000 (190)	0.156
	10	c.1566C>T	0.011 (188)	0.027 (184)	0.241
	11	c.1608A>G	0.011 (190)	0.037 (188)	0.089
	12	c.*6G>A	0.005 (190)	0.011 (188)	0.556
	13	c.*18C>T	0.011 (190)	0.037 (188)	0.089
	14	c.*76delT	0.005 (190)	0.022 (186)	0.169
	15	c.*85T>C	0.011 (188)	0.000 (186)	0.158
	16	c.*88T>A	0.005 (188)	0.000 (186)	0.320
	17	c.*99G>A	0.021 (188)	0.005 (186)	0.181
<b>STK11</b>	1	c.-311C>T	0.005 (192)	0.011 (176)	0.544
	2	c.-127T>C	0.021 (192)	0.005 (182)	0.197
	3	c.290+36G>T	0.240 (192)	0.261 (186)	0.678
	4	c.290+78C>T	0.161 (192)	0.113 (186)	0.171
	5	c.290+84C>T	0.005 (192)	0.000 (186)	0.324
	6	c.290+177G>C	0.010 (192)	0.000 (186)	0.163
	7	c.374+24G>T	0.172 (192)	0.112 (188)	0.093
	8	c.374+141G>A	0.005 (192)	0.000 (188)	0.322
	9	c.374+200delC	0.010 (192)	0.011 (188)	0.983
	10	c.374+272G>A	0.010 (192)	0.011 (188)	0.983
	11	c.375-132A>G	0.057 (192)	0.086 (186)	0.278
	12	c.375-49G>A	0.354 (192)	0.462 (186)	0.032
	13	c.464+32insGGGGGCC	0.094 (192)	0.033 (180)	0.018
	14	c.465-194G>A	0.005 (192)	0.000 (188)	0.322
	15	c.465-51T>C	0.271 (192)	0.351 (188)	0.091
	16	c.735-51C>T	0.052 (192)	0.024 (170)	0.160
	17	c.863-224C>T	0.010 (192)	0.017 (178)	0.583
	18	c.920+7G>C	0.214 (192)	0.250 (180)	0.405
	19	c.920+32G>A	0.005 (192)	0.011 (178)	0.518
	20	c.920+263A>G	0.516 (192)	0.532 (186)	0.746
	21	c.920+311C>T	0.203(192)	0.250 (188)	0.275
	22	c.1062C>G	0.005 (192)	0.006 (178)	0.957
	23	c.1185A>G	0.005 (192)	0.000 (172)	0.343

	24	c.*16+110C>T	0.250 (192)	0.227 (176)	0.610
	25	c.*16+113G>A	0.010 (192)	0.017 (180)	0.601
<b>PALB2</b>	1	c.-47G>A	0.021 (192)	0.027 (184)	0.163
	2	c.49-352delGA	0.042 (192)	0.042 (190)	0.600
	3	c.656A>G	0.005 (192)	0.000 (176)	0.338
	4	c.1194G>A	0.005 (192)	0.000 (190)	0.319
	5	c.1273G>A	0.016 (192)	0.005 (186)	0.330
	6	c.1572A>G	0.010 (192)	0.021 (190)	0.403
	7	c.1676A>G	0.094 (192)	0.048 (188)	0.082
	8	c.1684+41dupTGA	0.010 (192)	0.005 (186)	0.581
	9	c.2014G>C	0.042 (192)	0.021 (188)	0.256
	10	c.2442G>A	0.005 (192)	0.000 (188)	0.332
	11	c.2514+136T>C	0.005 (192)	0.000 (168)	0.349
	12	c.2586+58C>T	0.229 (192)	0.192 (172)	0.384
	13	c.2587-404A>G	0.130 (192)	0.152 (184)	0.541
	14	c.2587-190A>C	0.026 (192)	0.000 (188)	0.061
	15	c.2590C>T	0.026 (192)	0.000 (182)	0.061
	16	c.2748+15C>T	0.005 (192)	0.005 (188)	0.982
	17	c.2794G>A	0.005 (192)	0.000 (168)	0.349
	18	c.2816T>G	0.005 (192)	0.000 (168)	0.349
	19	c.2834+328C>T	0.016 (192)	0.000 (176)	0.100
	20	c.2835-27C>T	0.005 (192)	0.000 (160)	0.360
	21	c.2993G>A	0.026 (192)	0.017 (178)	0.639
	22	c.2997-108insT	0.042 (192)	0.018 (166)	0.197
	23	c.3113+196A>G	0.031 (192)	0.022 (184)	0.544
	24	c.3113+204G>A	0.141 (192)	0.147 (184)	0.866
	25	c.3113+325A>G	0.188 (192)	0.190 (184)	0.866
	26	c.3202-8G>T	0.005 (192)	0.000 (192)	0.317
	27	c.3300T>G	0.042 (192)	0.021 (190)	0.248

<sup>1</sup> Bonferroni correction,  $P < 7.25 \times 10^{-4}$  (0.05/69)

Table 3

Non-synonymous sequence variants detected in *PALB2* and *STK11* genes and corresponding residues present in orthologues<sup>1</sup>

Gene	SNP <sup>2</sup>	SNP ID <sup>3</sup>	Amino acid change	<i>Macaca mulatta</i>	<i>Mus musculus</i>	<i>Canis familiaris</i>	<i>Equus caballus</i>	<i>Monodelphis domestica</i>	<i>Ornithorhynchus anatinus</i>	<i>Gallus gallus</i>
<b>STK11</b>	22	c.1062C>G	Phe354Leu	Phe	Phe	Phe	Phe	Phe	Pro	Tyr
<b>PALB2</b>	3	c.656A>G	Asp219Gly	Asp	Glu	Arg	Asp	Cys	Gly	Ser
	5	c.1273G>A	Val425Met	Leu	Leu	Leu	Leu	Leu	Leu	Leu
	7	c.1676A>G	Gln559Arg	Arg	-	Gln	Arg	Gln	-	-
	9	c.2014G>C	Glu672Gln	Glu	Glu	Glu	Glu	Val	Arg	Thr
	15	c.2590C>T	Pro864Ser	Pro	Pro	Pro	Thr	Pro	Pro	Ser
	17	c.2794G>A	Val932Met	Val	Ile	Val	Ile	Val	Val	Val
18	c.2816T>G	Leu939Trp	Leu	Leu	Leu	Leu	Leu	Leu	Leu	Leu
21	c.2993G>A	Gly998Glu	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly

<sup>1</sup> No amino acid change was identified in the *CHEK2* gene<sup>2</sup> According to Table 1<sup>3</sup> According to the nomenclature of the Human Genome Variation Society

- No corresponding residue



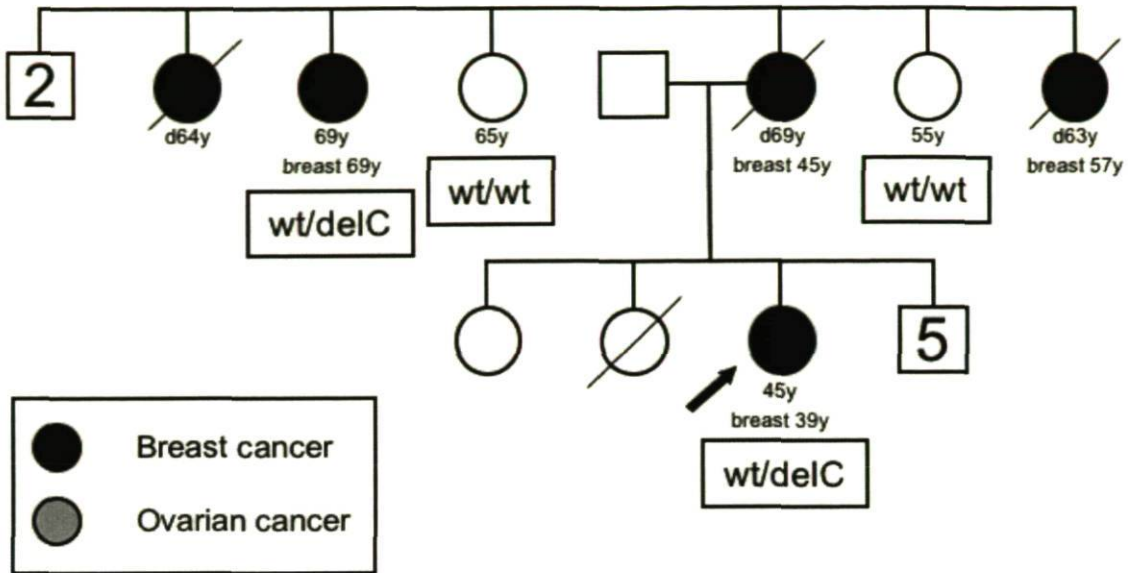
Table 4  
 Location of non-synonymous sequence variants detected in human *PALB2* and *STK11* genes and amino acid effect prediction on protein function<sup>1</sup>

Gene	SNP <sup>2</sup>	Amino acid change	Domain	SIFT	PolyPhen (score difference)
<b>STK11</b>	22	Phe354Leu	Near kinase domain (50-337)	Tolerated	Possibly damaging
<b>PALB2</b>	3	Asp219Gly	-	Affect protein function	Benign
	5	Val425Met	-	Tolerated	Benign
	7	Gln559Arg	-	Tolerated	Benign
	9	Glu672Gln	-	Affect protein function	Benign
	15	Pro864Ser	Near WD40 repeat (867-908)	Tolerated	Probably damaging
	17	Val932Met	-	Tolerated	Benign
	18	Leu939Trp	-	Affect protein function	Possibly damaging
	21	Gly998Glu	Near WD40 repeat (1008-1050)	Affect protein function	Probably damaging

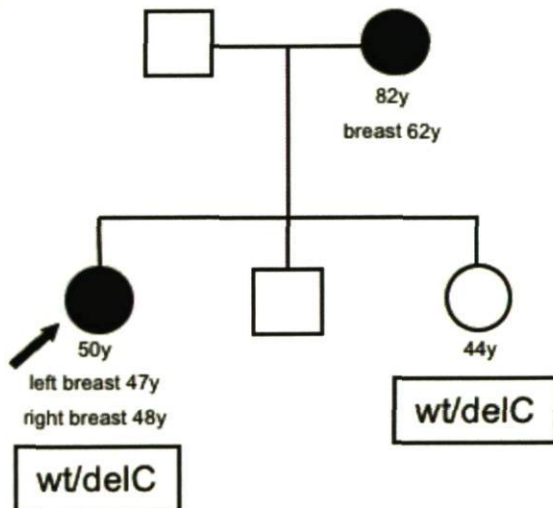
<sup>1</sup> No amino acid change was identified in the *CHEK2* gene

<sup>2</sup> According to Table 1

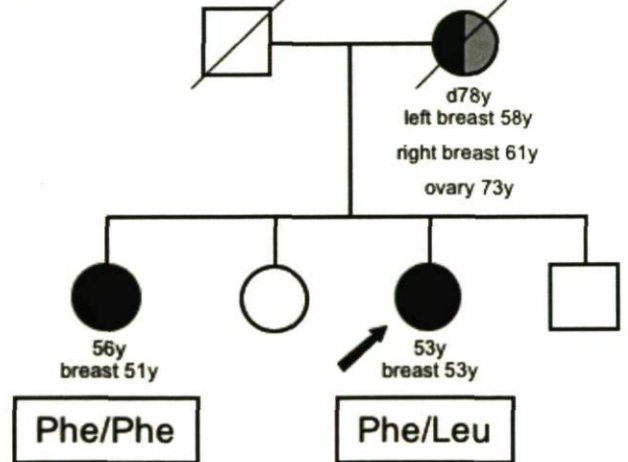
A



B



C



Supplemental table 1  
Oligonucleotides used to amplify and sequence the *STK11* and *PALB2* genes<sup>1</sup>.

Gene	Material	Fragment	Forward primer	Reverse primer	Sequence primer	Annealing (°C)	Fragment length (bp)
STK11	gDNA	STK11-Ex1	GGCCGGTTCATACTTGCCGT	TCTCAGGAGGTTTCGCTACCA	GGCCGGTTCATACTTGCCGT and TCTCAGGAGGTTTCGCTACCA	60	980
		STK11-Ex2	CAGCACTGTGAACACAGCTTCT	GCTGTGCCCGTAAGACCCAT	CAGCACTGTGAACACAGCTTCT and GCTGTGCCCGTAAGACCCAT	66	616
		STK11-Ex3	CTTACGGGCACAGCCTCCCTA	GGGCCTTCATGTCAATGAATATCA	CTTCCCGTAGGCTCCTCCT and GGGCCTTCATGTCAATGAATATCA	62	650
		STK11-Ex4-5	CAGGAAAATGAGACCTGTGGACA	CCACCATCTGCCGTATGAGTTACA	CCACCATCTGCCGTATGAGTTACA and GGTACTCCAGGCCGTCATCA	60	766
		STK11-Ex6	GTAACATACGGCAGATGGTGT	GGTCGGCCTCTCCACTCAGT	GTAACATACGGCAGATGGTGT	61	611
		STK11-Ex7	CTGGCTTGCAAGGCTGTCA	GCTCTGTCTTCCCTGTCATGGA	GCTCTGTCTTCCCTGTCATGGA	58	789
		STK11-Ex8	GGCTGAGCTTCTGTGGTACA	GAGAGGCTTCCCTCCTGAAAAGGT	GGACCCCTCTCTCATGCCAGCT	65	896
		STK11-Ex9	GTCAGCTTGCCCTCCTACTCGTGA	CCACACGTAGGCCGCCATGA	CTACTCGTAGGTTCCCTGCAGTCA and CCACACGTAGGCCCTCCATGA	54	752
		STK11	cDNA	STK11-PCR1	AGGGAAGTCGGAACACAAGGA	TTGTAGATGTTGTCCCTTCGAA	AGGGAAGTCGGAACACAAGGA and TTGTAGATGTTGTCCCTTCGAA
STK11	cDNA	STK11-PCR2	GCACCCTCAAAAATCTCCGACCT	GTCACAAGCCCGGTGCACCT	GCACCCTCAAAAATCTCCGACCT and GTCACAAGCCCGGTGCACCT	58	976
PALB2	gDNA	PALB2-Ex1	GGGGCATCGCTAGATAATCC	AGCAAGGGTTCGATACAGC	GGGGCATCGCTAGATAATCC	58	885
		PALB2-Ex2-3	CCTGGACCATAACGGATAAA	CACACTGTGGAAAAAGAACAA	CCTGGACCATAACGGATAAA	60	815
		PALB2-Ex4.1 <sup>2</sup>	TGCACTGATCTTCTTAAATAAATGTTT	TTGTGCCCTCCAAACTTACAGG	TTGTGCCCTCCAAACTTACAGG	60	787
		PALB2-Ex4.2 <sup>2</sup>	GCAAAAATCCTGCTAGATCACC	AAGGAAGTGCCAGGCCAAATA	AGGGCGACTACAGTTCCCTT and AAGGAAGTGCCAGGCCAAATA	57	1209
		PALB2-Ex5	ATGACCCTGTTTGAACCCCTT	TAAAACGTGGAAGGCCCAAT	GGGTATTACATTTAAGAATGGT and TAAAACGTGGAAGGCCCAAT	62	1232
		PALB2-Ex6	CATCCGTCGAACAGGTACA	TTGCCAATAGGTTGGCATAG	CATCCGTCGAACAGGTACA	60	671
		PALB2-Ex7	ACTCCCTCCGAGGAATAACA	ATTTGTATAAAGCGAACAAAGC	CTTTCATGGCCGGCACCGGT and ATTTGTATAAAGCGAACAAAGC	63	1079
		PALB2-Ex8-9	TACCTTGCAATAGGTAATGCTCA	TGTTGATGCGGTACATGCTT	TACCTTGCAATAGGTAATGCTCA and TGTTGATGCGGTACATGCTT	58	1478
		PALB2-Ex10	AATGAAATTCGAGGTGCTTGCA	TCGAAGGTAGACAGTTGGTCTG	AATGAAATTCGAGGTGCTTGCA, TCGAAGGTAGACAGTTGGTCTG and GTGATAGGAGGAGACTGTGCTCA	64	1096
		PALB2-Ex11	TGATGGAACCTTCCCTCATGG	CGGGGAAGGTTTGTTCATTA	TGATGGAACCTTCCCTCATGG	60	450
		PALB2-Ex12	CAGAGCCATCGGTCATTGC	TGTGTTTGACACAGTGCCTTT	CAGAGCCATCGGTCATTGC	60	324
		PALB2-Ex13	CAGATCCCGTATGTGTGTC	GCAGCACTGGCATGAAAGAT	CACCTTACTAGGCCAAAGACTATTGAT	60	1235

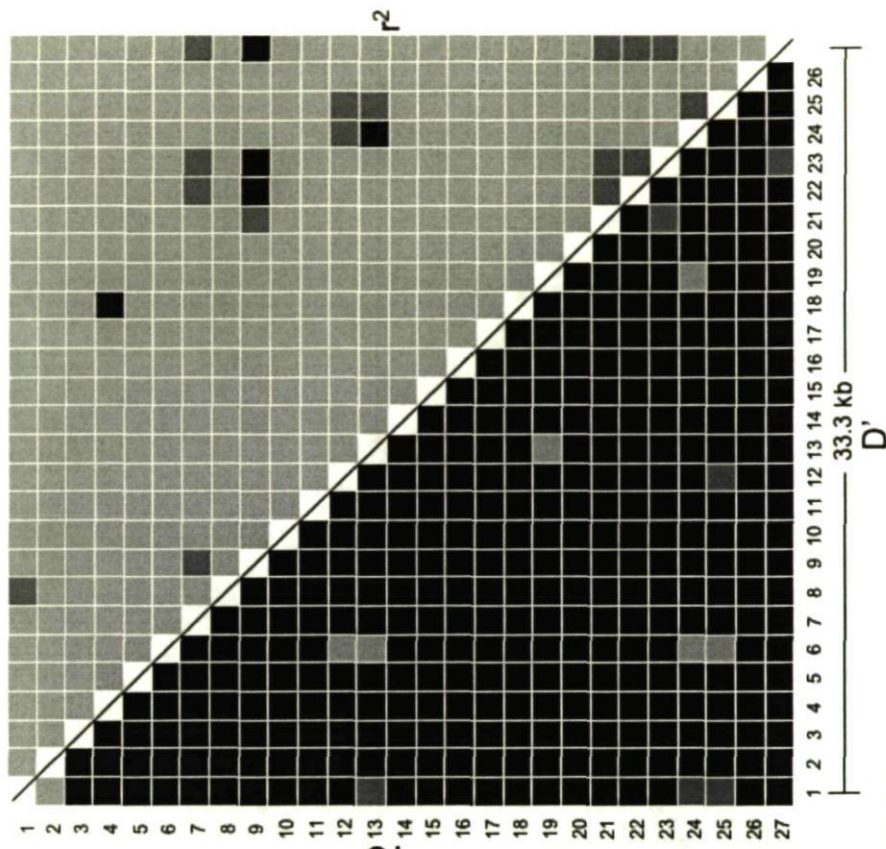
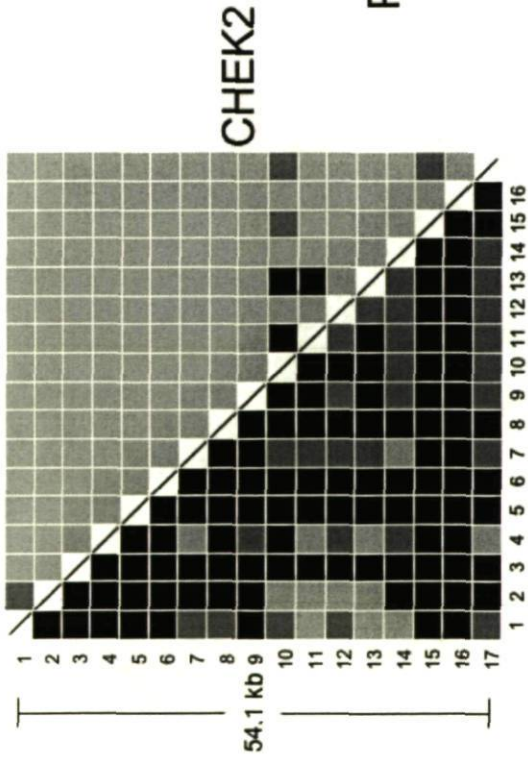
<sup>1</sup> These primers were selected with the Primer Express 2.0 software (Applied Biosystems, Foster City, CA, USA).  
<sup>2</sup> PALB2 exon 4 has been amplified in two fragments from 5' position to 3' position.

## Supplemental table 2

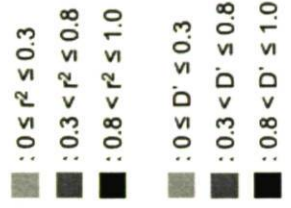
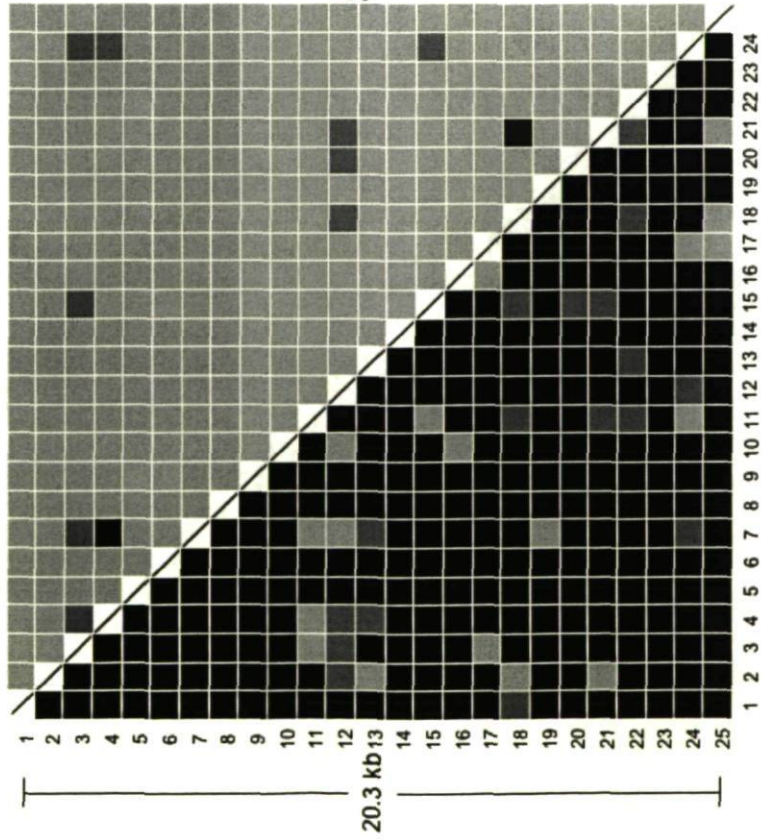
Description of the regions re-sequenced for the *CHEK2*, *STK11* and *PALB2* genes.

<b>Gene</b>	<b>Material</b>	<b>Number of exons analyzed</b>	<b>Total length sequenced (kb)</b>	<b>Average intronic bp analyzed (range)</b>	<b>5'UTR analyzed (bp)</b>	<b>3'UTR analyzed (bp)</b>
<b>CHEK2</b>	gDNA	15	5,4	102 (34-470)	72	125
<b>STK11</b> <sup>1</sup>	gDNA	9	6,0	207 (38-330)	360	16 <sup>2</sup>
	cDNA	---	1,6	---	80	227
<b>PALB2</b>	gDNA	13	11,2	176 (60-450)	200	300

<sup>1</sup> Exon 10 of the *STK11* gene was not analyzed (composed only of 3'UTR sequence)<sup>2</sup> 3'UTR region sequenced is located in exon 9



**STK11**



## CHAPITRE V

### Genetic sequence variations of BRCA1-interacting genes *AURKA*, *BAP1*, *BARD1* and *DHX9* in French Canadian Families with high risk of breast cancer

#### 1. Résumé

Les gènes de susceptibilité au cancer du sein représentent environ 20-25% de l'agrégation familiale. Une proportion significative demeure donc à être expliquée par d'autres gènes. En raison de la susceptibilité conférée par *BRCA1* et de son implication dans plusieurs processus cellulaires, il est justifié d'évaluer les gènes encodant des protéines interagissant avec *BRCA1* à titre de gènes candidats du cancer du sein. Quatre gènes encodant des protéines interagissant avec *BRCA1* ont ainsi été analysés. Un total de 10, 4, 11 et 6 variants ont été identifiés respectivement dans les gènes *AURKA*, *BAP1*, *BARD1* et *DHX9*. L'analyse de ces variants dans une cohorte de 98 Canadiennes françaises non-atteintes du cancer du sein a suggéré une différence de fréquence allélique pour un variant de *BARD1*. L'estimation d'haplotypes, l'identification de blocs d'haplotypes et de SNPs marqueurs ont été effectuées pour ces gènes, fournissant ainsi des outils efficaces et utiles pour l'identification d'allèles de susceptibilité au cancer du sein à pénétrance faible ou modérée dans ces gènes, et ce, dans de plus grandes cohortes.

Guénard F, Labrie Y, Ouellette G et al. Genetic sequence variations of BRCA1-interacting genes *AURKA*, *BAP1*, *BARD1* and *DHX9* in French Canadian families with high risk of breast cancer. *J Hum Genet.* 2009 Mar;54(3):152-61.

#### 2. Manuscrit

**Genetic sequence variations of BRCA1-interacting genes *AURKA*, *BAP1*, *BARD1* and *DHX9* in French Canadian Families with high risk of breast cancer**

Frédéric Guénard<sup>1</sup>, Yvan Labrie<sup>1</sup>, Geneviève Ouellette<sup>1</sup>, Charles Joly Beuparlant<sup>1</sup>, INHERIT BRCA<sup>2</sup> and Francine Durocher<sup>1\*</sup>

<sup>1</sup>Cancer Genomics Laboratory, Oncology and Molecular Endocrinology Research Centre, Centre Hospitalier Universitaire de Québec and Laval University, Québec, G1V 4G2, Canada;

<sup>2</sup>Other members of INHERIT BRCA<sup>2</sup> involved this study are listed in Appendix

**Short title:** BRCA1-interacting gene variants in BOC families

\*Corresponding author:

Dr Francine Durocher

Cancer Genomics Laboratory, Oncology and Molecular Endocrinology Research Centre, Centre Hospitalier Universitaire de Québec and Laval University

2705 Laurier Boulevard

Québec city, Québec,

Canada, G1V 4G2

Tel: (418)-654-2296

Fax: (418)-654-2278

**e.mail: [Francine.durocher@crchul.ulaval.ca](mailto:Francine.durocher@crchul.ulaval.ca)**

**Abstract**

Breast cancer is a heterogeneous disease displaying some degree of familial clustering. Highly penetrant breast cancer susceptibility genes represent approximately 20-25% of the familial aggregation of breast cancer. A significant proportion of this familial aggregation of breast cancer is thus yet to be explained by other breast cancer susceptibility genes. Given the high susceptibility conferred by the two major breast cancer predisposition genes, *BRCA1* and *BRCA2*, and the implication of these genes in many key cellular processes, assessment of genes encoding BRCA1-interacting proteins as plausible breast cancer candidate genes is thus attractive. In this study, four genes encoding BRCA1-interacting proteins were analyzed in a cohort of 96 breast cancer individuals from high-risk non-*BRCA1/BRCA2* French Canadian families. Although no deleterious truncating germline mutations or aberrant spliced mRNA species were identified, a total of 10, 4, 11 and 6 variants were found in the *AURKA*, *BAP1*, *BARD1* and *DHX9* genes, respectively. The allele frequency of each variant was further ascertained in a cohort of 98 healthy French Canadian unrelated women and a difference in allele frequency was observed for one *BARD1* variant based on single-marker analysis. Haplotype estimation, haplotype blocks and tagging SNPs identification were then performed for each gene, providing a valuable tool for further searches of common disease-associated variants in these genes and therefore further analyses on these genes in larger cohorts is warranted in the search of low-to-moderate penetrance breast cancer susceptibility alleles.

**Keywords:** BRCA1-interacting genes; breast cancer; haplotypes; tagging SNPs; variants



## Introduction

Breast cancer is the most common malignant form of cancer among occidental women, with hereditary cases representing approximately 5-10%.<sup>1</sup> Mutations in high penetrance *BRCA1* and *BRCA2* genes account for only ~25% of families with hereditary breast cancer<sup>2</sup> while other lower penetrance genes such as *ATM*, *CHEK2*, *TP53*, *PTEN*, *STK11*, *CDH1*, *PALB2/FANCN* and *BRIP1/FANCF*<sup>3</sup> explain a much smaller proportion of families with hereditary breast cancer. Deleterious mutations in *BRCA1/BRCA2* were identified in 24% of our 256 high-risk French Canadian breast/ovarian cancer families,<sup>4</sup> supporting that a significant proportion of high-risk breast cancer families still remains unexplained by either germline mutations in high penetrance *BRCA1/2* alleles or other lower penetrance genes.<sup>5-10</sup>

Several structural motifs allow the BRCA1 protein to interact with cellular proteins to regulate diverse biological functions such as control of transcription<sup>11</sup> and DNA damage repair.<sup>12</sup> Therefore, BRCA1-interacting proteins implicated in these cellular pathways represent attractive candidates with regard to breast cancer susceptibility. Based on a candidate gene strategy, four genes encoding BRCA1-interacting proteins, namely *AURKA*, *BAP1*, *BARD1* and *DHX9*, were thoroughly investigated in an attempt to identify germline deleterious mutations and/or aberrant splicing events which could potentially be associated with an increase in breast cancer risk. In the present study, familial breast cancer cases were purposely selected. Indeed, such a study design has been demonstrated to substantially decrease the number of cases and controls to achieve the same power, as compared with studies where cases are unselected for family history.<sup>13</sup>

The *AURKA* protein is a centrosome-associated Ser/Thr kinase which localizes to the centrosome and could be involved in centrosome duplication,<sup>14</sup> maturation and separation, as well as in the control of bipolar spindle assembly.<sup>15</sup> *AURKA* binds and phosphorylates BRCA1, and it was suggested that this BRCA1 phosphorylation plays a role in G2/M transition.<sup>16</sup> The *AURKA* gene has been recently identified as a low-penetrance tumor susceptibility gene,<sup>17</sup> and was subsequently analyzed for its implication with breast cancer albeit leading conflicting results.<sup>18-27</sup>

*BAP1* is an ubiquitin carboxy-terminal hydrolase which binds to the BRCA1 RING finger domain and enhances BRCA1 growth suppression properties.<sup>28</sup> *BAP1* was suggested to

regulate the BRCA1/BRCA2-RAD51 DNA repair complex.<sup>29</sup> Given the known implication of BRCA1 in DNA repair<sup>12</sup> and its interaction with RNA pol II,<sup>11</sup> which is linked to transcription-coupled repair (TCR) processes, a role for BAP1 in TCR has been suggested.<sup>30</sup> It has been reported that mutation of an amino acid of BAP1 in its BRCA1-binding domain (L691P) abolishes the BAP1-BRCA1 interaction.<sup>28</sup> On the other hand, a study on 47 non-*BRCA1/BRCA2* familial breast cancer cases from the French population did not reveal any association with breast cancer predisposition.<sup>31</sup>

The BARD1 protein interacts with BRCA1 and both proteins form a complex which possesses a dual E3 ubiquitin ligase activity.<sup>32</sup> The heterodimer BRCA1/BARD1 promotes its own ubiquitination<sup>33</sup> and ubiquitinates RNA pol II subunit A.<sup>34</sup> The interaction of BRCA1 with RNA pol II<sup>11</sup> and DHX9<sup>35</sup> suggested a role for BARD1 in TCR. The *BARD1* gene has been reported to be targeted by somatic<sup>36</sup> and germline<sup>37-39</sup> mutations in breast and in ovarian<sup>36,40</sup> cancer cohorts. *BARD1* mutations have been identified in hereditary breast/ovarian cancers from *BRCA1/2*-negative patients<sup>37,41</sup> and it was suggested that the Cys557Ser variant may represent a low penetrance breast cancer susceptibility allele.<sup>39,42,43</sup> However, a recent study did not identify any increased risk of breast cancer associated with specific *BARD1* variations,<sup>44</sup> while another study reported an association of *BARD1* haplotypes with either an increased or decreased risk of breast cancer.<sup>45</sup>

As for the DHX9 helicase, a member of the DEAH helicase family,<sup>46</sup> it binds directly to both CBP and pol II and is required for complex formation between both proteins.<sup>47</sup> The DHX9 protein (aa 230-325) binds the full-length BRCA1 protein through a sub-region of the BRCT domain of BRCA1 (Anderson et al. 1998). *DHX9* truncating mutations were reported to affect the interaction of both proteins (BRCA1 and RNA pol II) and to result in decreased transcriptional activity of BRCA1,<sup>35</sup> thus possibly affecting its function in transcription and TCR. Despite an interaction with BRCA1 and a role in transcriptional activity of BRCA1, *DHX9* mutations have not been assessed thus far in non-*BRCA1/BRCA2* breast cancer families.

Based on the multiple functions of these BRCA1-interacting genes, they obviously represent attractive candidates that might explain a proportion of the remaining inherited breast cancer cases. The goal of this study being the identification of deleterious mutation,

and given that such deleterious disease-causing mutations producing a premature truncation of the protein occur mostly within the coding region, complementary DNA (cDNA) mutation screening was therefore performed. In addition, since alternative splicing is now recognized as the primary source of human proteomic diversity and gene regulation and that an increasing number of disease-causing mutations affecting the correct splicing of genes are identified,<sup>48</sup> cDNA material definitively represents a resource of choice as alternative splice forms could not be detected when extensive re-sequencing is performed on genomic DNA.

Thus the goal of this study was to evaluate the possible involvement of *AURKA*, *BAP1*, *BARD1* and *DHX9* germline deleterious mutations or on breast cancer susceptibility by analyzing all mRNA-encoding sequences as well as alternative splicing in cDNA samples from 96 breast cancer cases, in which no *BRCA1/2* mutations were detected.

## **Material and methods**

### *Ascertainment of families and DNA extraction*

High-risk French Canadian breast and/or ovarian families were recruited through a large ongoing interdisciplinary research program designated INHERIT BRCA. Clinicians (listed in Appendix) were directly involved in this translational research program and were responsible for *BRCA1/2* test result disclosure to participants. Approval was obtained from ethics committees corresponding to the different participating institutions. More details regarding ascertainment criteria, experimental and clinical procedures, as well as the INHERIT BRCA research program have been described elsewhere.<sup>4,5</sup> Subsequently, another component was designed for the “Localization and identification of novel breast cancer susceptibility loci/genes”. Ethics approval for this latter study was also obtained and each participant knowing their inconclusive *BRCA1/2* test results status had to sign a specific informed consent for their participation in this component. A subset of 96 high-risk French Canadian breast/ovarian cancer families were recruited in the present study according to the ascertainment criteria described previously.<sup>5,7,9</sup> Recruitment of 98 healthy unrelated French Canadian individuals was conducted on a non-nominative basis, as part of a long term study aiming at the characterization of the genetic variability in human population, and approved by the Institutional Ethics Review Board of the Hôpital Sainte-Justine.

### *DNA extraction, RNA isolation from immortalized cell lines and cDNA synthesis*

Lymphocytes from breast cancer individuals were isolated and immortalized as previously described.<sup>5</sup> Total RNA extraction, storage and reverse transcription of RNA, as well as genomic DNA extraction from healthy unrelated French Canadian individuals were also performed as previously described.<sup>5,6,8-10</sup>

### *PCR amplification and mutation analysis*

cDNA amplicons covering the entire mRNA-encoding sequences of the GenBank mRNA records (*AURKA*: NM\_198436.1; *BAP1*: NM\_004656.2; *BARD1*: NM\_000465.1; *DHX9*: NM\_001357.3) were obtained by PCR amplification from cDNA of breast cancer cases using the primer pairs listed in Supplemental Table 1. Direct sequencing and sequence data analysis were performed as previously described.<sup>5</sup> Genomic DNA sequences from healthy unrelated French Canadian individuals were obtained using the same direct sequencing

method, but with amplification of relevant variant-containing exons from amplicons obtained using intronic primers (Supplemental Table 1). The segregation of rare missense and deletion variants (minor allele frequency; MAF <5%), and variants demonstrating a suggestive difference in genotype frequency between breast cancer cases and unaffected individuals were also analyzed in families for which DNA material was available from multiple individuals.

Each variant was tested for deviation from HWE by means of a  $\chi^2$  test. All p-values were two-sided with 1 degree of freedom. In order to estimate the genetic association of each variation or haplotype with breast cancer, differences in genotype distributions between the case and control groups were tested using a  $\chi^2$  test. P-values less than 0.05 are considered as significant. However, to address the issue of multiple comparison, Bonferroni correction was applied for each estimate of genetic association.

#### *Conservation in other species and computational analyses*

Protein alignment was performed using data extracted from the UCSC database, and both the SIFT and PolyPhen web-based softwares were used to predict the effect of amino acid substitutions. All tests were run under default threshold values.

#### *LD analysis and haplotype estimation*

To estimate the pattern of LD, the LDA program was used to calculate pairwise LD ( $D'$  and  $r^2$ ) values. Haplotype analysis was performed with the PHASE 2.1.1 software. Haplotype frequencies were estimated using all variants genotyped in both sample series (cases and controls) and a permutation test was conducted in order to determine the significance of differences in the distribution of inferred haplotypes between both groups. All association tests were run under default conditions, with 1000 permutations. Haplotype blocks were identified using genotyping data from both series using the Haploview software. tSNPs from each LD block were thereafter identified using the same software, and haplotype frequencies were also estimated using only the tSNPs identified.

## Results

The entire coding sequence of *AURKA*, *BAP1*, *BARD1* and *DHX9* genes were analyzed using cDNA from 96 high-risk breast cancer individuals, each coming from a distinct high-risk family. No deleterious truncating germline mutations were found. In addition, based on gel electrophoresis and sequence chromatogram analysis, no major alternative splice forms have been detected for any of the genes analyzed, unlikely involving this kind of alterations with breast cancer risk. However, a number of sequence variations were identified within each of these four genes, as described below.

### *AURKA* gene analysis

Five of the 10 variants identified in the *AURKA* gene are coding substitutions, whereas the remaining five are located within the 3'UTR region (Table 1). Four of the *AURKA* variants are novel, being not reported in dbSNP database (Build 127). Albeit no significant difference was observed between both sample sets based on single-marker analysis, the novel c.1130T>A variation is present exclusively among three breast cancer cases at the heterozygous state. As for the c.1117A>G (Met373Val) variant present in four breast cancer individuals at the heterozygous state, segregation analysis suggests a possible association with breast cancer. Indeed as illustrated in the pedigree from one of the four carrier's families (Figure 1), all heterozygous carriers are affected with breast cancer while among the two non-carriers, one is affected (diagnosed at 45 years) and the other is unaffected (55 years). A DNA sample from an additional family member was also available for two other families. An unaffected first-degree relative (59 years) could be genotyped and was found non-carrier of the Met373Val variant in one family, whereas an affected third-degree relative diagnosed with breast cancer at 74 years was found heterozygote carrier for this variant in the other family (data not shown). No additional family member was available for the remaining family. With the exception of the c.1212+423G>C variant which displays a slight excess of homozygotes (HWE  $P=0.03$ ) among the breast cancer cases, no variants display a significant deviation from HWE.

As shown in Table 2, all variants located in the coding sequence result in an amino acid change. Of these, three residues (Val57, Ala213 and Val377) are completely conserved in distant species (apart from Leu377 in *Md*), suggesting that they are under strong functional constraint or may have a specific role on protein conformation. The three variants

(Ala213Val, Met373Val and Val377Glu) located within the Ser/Thr protein kinase domain are predicted to be not tolerated by SIFT, with Val377Glu also evaluated as probably damaging by the PolyPhen program (Table 3).

A graphical representation of LD with both  $r^2$  and Lewontin's  $|D'|$  and using data from both series is shown in Figure 2. Most single nucleotide polymorphism (SNP) pairs display strong LD ( $0.8 < D' \leq 1$ ), while lower LD is observed for combination involving SNPs 1, 2, 4 and 6. The  $r^2$  coefficient calculated displays a large spectrum of values ranging from 0 to 0.6, given the high dependence of  $r^2$  on allele frequency. Analysis of haplotypes estimated with the PHASE program led to 6 haplotypes exhibiting a frequency  $\geq 1\%$  (Table 4). Based on solid block algorithm, two LD blocks were identified at the *AURKA* locus, with LD breakage being located in the region of exon 9 (Figure 3). However, when performing a similar analysis using CEPH/CEU data, three LD blocks are obtained, including two relatively small blocks (0.5 and 1.2 kb) in the 5' region and a larger block covering intron 1 to intron 7 (17 kb) (data not shown). As indicated in Figure 3, three tagging SNPs (tSNPs) (rs2273535, rs1047972 and rs34189236) have been identified, representing 95% of our estimated haplotypes.

### ***BAP1 gene analysis***

Analysis of *BAP1* sequence in our breast cancer individuals led to the identification of only three variants (Table 1). Two are novel synonymous variants, c.294C>T being detected only once among our 96 breast cancer individuals, whereas c.1026C>T was present exclusively in two unaffected individuals at the heterozygous state. No significant difference in allele frequency was obtained between both series for any of these nucleotide substitutions based on single-marker analysis.

All *BAP1* variations are comprised within one block of strong LD (Figure 3), which is confirmed by the perfect LD ( $D'=1$ ) observed between the two most distant intragenic SNPs (SNP1 and SNP4) (Figure 2). As displayed in Table 4, PHASE analysis estimated two major haplotypes which are represented by only one tSNP (rs123598; SNP4) (Figure 3), and both haplotypes display a similar frequency in both sample sets. Given that only three SNPs, all located in intron 3, are available from HapMap data, it was not relevant to perform haplotype and tSNP identification using the CEU population.

### ***BARD1 gene analysis***

Among the *BARD1* sequence variations identified in this study, six are missense variants, while another consists of an in frame 21 bp deletion (seven amino acids) in exon 4 (c.1075del21). Three (c.820A>G, c.1203T>C and c.2212A>G) are considered as novel, being not reported in dbSNP database (Table 1). All variants follow HWE, with the exception of the c.1670G>C variant, where a significant deviation is observed due to the occurrence of one rare homozygote. A significant difference is also observed for the c.1518C>T substitution, which is slightly more frequent within cases ( $P=0.013$ , data not shown). However, this association does not reach statistical significance following Bonferroni correction, and segregation of this variant in a large informative family does not reveal any obvious association with breast cancer (data not shown). Of the seven amino acid changes identified, only the Ile738Val residue displays a certain level of conservation in higher species, with the alternative Val residue being observed in rodents and *Tetraodon nigroviridis* (Table 2). Apart from the deletion of 7 amino acids, *in silico* evaluation of the potential effect of amino acid changes on the protein function (Table 3) revealed that four amino acid changes are predicted to affect the protein function by at least one of the two programs used, including the Arg378Ser change which is predicted to potentially affect the protein function by both SIFT and PolyPhen softwares.

The graphical representation of the pairwise LD between the 11 variants identified in both series combined is shown in Figure 2. The low pairwise  $D'$  values observed involve mainly SNP8, while moderate pairwise values are observed for SNP1. A large spectrum of  $r^2$  values ranging from 0 to 1.0 is observed for *BARD1*. The Haploview software (Figure 3) determined two LD blocks, with breakage of LD located between SNP3 and SNP4 in the region of exon 4. Within these two LD blocks, 2 and 6 tSNPs were identified, block 1 being defined by SNPs 1 and 2, while SNPs 4, 6, 8, 9, 10 and 11 represent the second block. Similar analyses using CEPH/CEU data set led to the identification of 7 LD blocks (data not shown). The discordance obtained is present mainly in the 5'-region of the *BARD1* gene, in which the four haplotype blocks identified from HapMap data are represented by a single block in the French Canadian population. Regarding haplotypes estimation, the analysis led to the identification of 11 haplotypes, all exhibiting a frequency  $\geq 1\%$  (Table 4). No significant difference in single or global haplotype distribution was identified between



both groups.

### ***DHX9 gene analysis***

The analysis of *DHX9* in our breast cancer individuals led to the identification of five variants, with an additional nucleotide substitution present exclusively among unaffected individuals (Table 1). Out of these variations, two involve an amino acid change, among which c.1873C>G is located in the proximity of the helicase C-terminal domain (aa 631-776). The c.1745C>T (rs2275177) sequence variation demonstrates a significant deviation from HWE in breast cancer cases, where an excess of rare homozygous genotypes is observed. Single-marker analysis did not reveal a significant difference in MAF for any of the nucleotide changes identified. Regarding the amino acid conservation of the Pro89Ala variant in other orthologues, the Pro allele is conserved in the three species in which this residue is available (*Canis* and rodents, Table 2). The two *DHX9* amino acid changes identified are predicted to be possibly damaging by the PolyPhen program. In addition, the Ser625Cys located within the RNA pol II binding domain (aa 230-650) is predicted to be not tolerated according to SIFT (Table 3).

Most pairs of adjacent SNPs display strong LD ( $0.8 < D' \leq 1$ ) following LDA analysis. As displayed in Figure 2,  $r^2$  values under 0.3 are obtained for all pairwise comparisons. The perfect  $D'$  value associated with the two most distant variations identified is confirmed by Haploview analysis (Figure 3). Three tSNPs were highlighted for *DHX9*, namely c.915T>G, c.1745C>T and c.1976C>T (Figure 3). Four haplotypes exhibiting a frequency  $\geq 1\%$  have been estimated with the PHASE program (Table 4).

## Discussion

Given that deleterious mutations in *BRCA1/BRCA2* were identified in only 24% of the 256 high-risk French Canadian breast/ovarian cancer families,<sup>4</sup> the analysis of the coding region of four genes encoding BRCA1-interacting proteins was undertaken in 96 high-risk individuals drawn from non-*BRCA1/BRCA2* French Canadian families. To our knowledge, this is the first study to investigate the potential association of *DHX9* germline mutations with breast cancer susceptibility in high-risk non-*BRCA1/BRCA2* breast cancer families. Although *AURKA* and *BARD1* genes have been investigated either in tumor samples, population-based studies or high-risk families, only one investigation assessed the involvement of *BAP1* in breast cancer susceptibility.<sup>31</sup>

The Phe31Ile variant was reported to confer an increased risk of breast cancer, either alone or in combination with Val57Ile.<sup>18-21,23,27</sup> However our study and others did not replicate this association.<sup>22,24-26</sup> With the exception of c.1212+423G>C (rs8173), evaluated in both breast<sup>44</sup> and ovarian<sup>49</sup> cancer, none of the 3'UTR *AURKA* variants identified in our study were assessed in other studies. As for the Met373Val variant, we report here a possible implication of this variant with breast cancer susceptibility, segregation analysis indicating that all heterozygous carriers are affected with breast cancer. However, this rare variant is observed at a similar frequency in affected and unaffected individuals from our study, thus requiring further analyses.

Many *BARD1* variants identified in this study have been evaluated in the context of breast cancer. The c.1670G>C variant was associated with an increased risk of breast cancer either in high-risk families or in population-based studies,<sup>38,39</sup> whereas our results and others<sup>43,50</sup> do not support these findings. None of the other variants identified here (i.e. c.609A>C, c.1053G>C, c.1075del21 and c.1518C>T) was previously associated with an increased risk of breast cancer,<sup>38,41</sup> despite our results suggesting an association of the c.1518C>T (His506His) with an increased risk of breast cancer (P=0.013), which is however not significant following Bonferroni correction. This finding would nonetheless have to be confirmed in much bigger cohorts.

The *BAP1* analysis led to the identification of the c.1026C>T (Ser342Ser) variant, which was reported once in a cohort of 47 French familial breast cancer cases negative for *BRCA1/BRCA2* mutations.<sup>31</sup> Interestingly the common c.2190+444C>T sequence variation

localized in the 3'UTR region was identified in our population with a MAF>5% in both series but was not analyzed in the French cohort.<sup>31</sup>

As for *DHX9* sequence variations, we report here the first assessment of its implication in breast cancer. Of the six sequence variants identified, four are rare, novel variants, two of which being missense variants. The Ser625Cys variation is located within the RNA pol II binding domain (aa 230-650) close to the C-terminal helicase domain (aa 631-776), while the Pro89Ala is situated in the N-terminal basic region capable of binding efficiently the C/H3 region of CBP transcription coactivator (aa 1-250).<sup>47</sup>

In addition to establishing the allele frequency of each variant in a series of unaffected controls, the effect of the variants identified were also assessed using conservation analysis and web-based programs. Both sequencing chromatograms and gel electrophoresis were thoroughly analyzed to detect any aberrant spliced forms, and we were able to conclude that no major spliced mRNAs are highly expressed in the cDNA material from our high-risk breast cancer individuals. Despite the fact that alternative transcripts expressed at low levels (~10% compared to the wild type form) can be detected by sequencing chromatogram analysis,<sup>5,51</sup> we cannot exclude the presence of minor, weakly expressed forms which could not have been detected here. The identification of alternative spliced forms of *BARD1*<sup>40,52,53</sup> and *AURKA*<sup>54</sup> mRNA have been reported, which is in accordance with UCSC database data and other alternative transcript databases such as ASAP II and Fast DB. As for *BAP1*, a splice variant was reported in a small cell lung cancer cell line.<sup>28</sup> However discordant results are presented in different databases, UCSC identifying many alternative spliced mRNA species for both *BAP1* and *DHX9* genes, while the database Fast DB identified only few alternative spliced mRNA species for *BAP1*, and none for *DHX9*.

In terms of linkage disequilibrium, relatively strong LD ( $0.8 < D' \leq 1$ ) is observed for the four genes when using the  $D'$  measure (Figure 2), and the  $D'$  values obtained here are concordant with data from the sole study reporting LD values ( $D'$ ) for the *AURKA* gene.<sup>22</sup> Using the  $r^2$  correlation coefficient, much lower LD values were obtained. Indeed differences between  $D'$  and  $r^2$  values are likely to be explained by the fact that  $r^2$  is more stringent and sensitive to frequency than  $D'$ .<sup>55</sup>

To our knowledge, the current study is the first to report haplotype estimation for the *BAP1*

and *DHX9* genes. The sole other study assessing *BARD1* haplotypes<sup>45</sup> reported a significant difference in haplotype distribution between cases and controls for two different haplotypes. Although both haplotypes have indeed been estimated in our analysis (*BARD1*-h1 and *BARD1*-h6), this association was not observed. The current study is also the most extensive investigation conducted on *AURKA* haplotypes, which includes 11 variations encompassing exons 3 to 9. Our results do not confirm any association of *AURKA* haplotypes with an increased risk of breast cancer, as previously demonstrated in other studies.<sup>19,22,23</sup>

The identification of haplotype blocks was thereafter conducted for the *AURKA*, *BAP1*, *BARD1* and *DHX9* genes. With the exception of the *AURKA* gene, for which the 5'-end can be roughly separated into two clusters,<sup>22</sup> this analysis was not previously performed or reported for the *BAP1*, *BARD1* and *DHX9* genes. The haplotype block identification for *BAP1* is concordant with results obtained from HapMap data, whereas similar analyses using French Canadian data for the *AURKA*, *BARD1* and *DHX9* genes revealed slight differences when compared with HapMap data (data not shown). However, we have to keep in mind that the greater extent of LD generally found in the French Canadian population can be responsible for a lower number of haplotype blocks.<sup>56</sup> In addition such a difference between HapMap and this study may be due to the higher coverage of the gene regions in the HapMap data set which include both exonic and intronic SNPs, the latter not assessed in our study.

The analysis of the coding region of the *AURKA*, *BAP1*, *BARD1* and *DHX9* genes did not identify any deleterious truncating mutation or aberrant splicing mRNAs. We report here a possible association of the *BARD1* c.1518C>T variant with an increased risk of breast cancer. Further studies on much bigger cohorts are however needed to fully evaluate the association of this variant with breast cancer risk and to assess the functional contribution of this variant on mRNA stability and protein expression. Haplotype analysis, LD block and tSNPs identification were also conducted for all four genes, and permitted to state that haplotypes displaying a frequency  $\geq 1\%$  represent more than 94% of all haplotypes estimated for each gene in our population. This first haplotype-based analysis to establish tSNPs for both the *BAP1* and *DHX9* genes will provide a valuable tool for further large association studies.

### **Acknowledgements**

We are indebted to the participants and their families for their generosity and providing DNA samples. We would like to thank Dr Martine Dumont, Gilles Leblanc, Carolle Samson and Martine Tranchant for sample management, mutation screening, and skillful technical assistance as well as Claire Brousseau, Marie-Andrée Lajoie, Pascale Léger, Hélène Malouin and Josée Rhéaume, for genetic counseling and clinical data management. We also thank Anne-Marie Moisan and Lucie Larouche for MLPA analyses. We thank Damian Labuda and Claudia Moreau at the Centre de Recherche de l'Hôpital Ste-Justine for providing control DNA samples. We thank Professor Bartha Maria Knoppers and her colleagues from the Centre de recherche en droit public de l'Université de Montréal for their precious help with ELSI issues related to our research program. We also appreciate advice received from ethics committees. This work was supported by the Canadian Institutes of Health Research (CIHR) through the INHERIT BRCA's research program, the Fonds de la Recherche en Santé du Québec (FRSQ)/Réseau de Médecine Génétique Appliquée (RMGA), the CURE Foundation and the Canadian Breast Cancer Research Alliance (CBCRA). FG holds a studentship from Fondation René Bussièrès, CJB holds a Frederick Banting and Charles Best Canada Graduate Scholarships - Master's Award studentship from CIHR, JS is Chairholder of the Canada Research Chair in Oncogenetics, and FD is a recipient of a chercheur-boursier from the Fonds de la Recherche en Santé du Québec (FRSQ) and a Research Career Award in the Health Sciences from CIHR/Rx&D Health Research Foundation.

**Appendix**

Clinicians and other members (\*) of INHERIT BRCAs involved in this study

Paul Bessette: Service de gynécologie, Centre Hospitalier Universitaire de Sherbrooke, Fleurimont, QC, J1H 5N4, Canada;

Peter Bridge: Molecular Diagnostic Laboratory, Alberta Children's Hospital, Calgary, AB, T2T 5C7, Canada;

Jocelyne Chiquette: Clinique des maladies du sein Deschênes-Fabia, Hôpital du Saint-Sacrement, Québec, QC, G1S 4L8, Canada;

Rachel Laframboise: Service de médecine génétique, CHUQ, Pavillon CHUL, Québec, QC, G1V 4G2, Canada;

Jean Lépine: Centre Hospitalier régional de Rimouski, Rimouski, QC, G5L 5T1, Canada;

Bernard Lespérance, Roxane Pichette: Service d'hémo-oncologie, Hôpital du Sacré-Cœur, Montréal, QC, H4J 1C5, Canada;

Marie Plante: Service de gynécologie, CHUQ, L'Hôtel-Dieu de Québec, Québec, QC, G1R 2J6, Canada;

\* Jacques Simard, Canada Research Chair in Oncogenetics: Cancer Genomics Laboratory, Oncology and Molecular Endocrinology Research Centre, CHUQ, Pavillon CHUL, Québec, QC G1V 4G2, Canada

## References

- 1 Bradbury AR, Olopade OI (2007) Genetic susceptibility to breast cancer. *Rev Endocr Metab Disord* 8(3):255-267
- 2 Easton DF (1999) How many more breast cancer predisposition genes are there? *Breast Cancer Res* 1(1):14-17
- 3 Stratton MR, Rahman N (2008) The emerging landscape of breast cancer susceptibility. *Nat Genet* 40(1):17-22
- 4 Simard J, Dumont M, Moisan AM, Gaborieau V, Malouin H, Durocher F, Chiquette J, Plante M, Avaré D, Bessette P, Brousseau C, Dorval M, Godard B, Houde L; INHERIT BRCA, Joly Y, Lajoie MA, Leblanc G, Lepine J, Lesperance B, Vezina H, Parboosingh J, Pichette R, Provencher L, Rheaume J, Sinnott D, Samson C, Simard JC, Tranchant M, Voyer P, Easton D, Tavtigian SV, Knoppers BM, Laframboise R, Bridge P, Goldgar D (2007) 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* 44(2):107-121
- 5 Durocher F, Labrie Y, Soucy P, Sinilnikova O, Labuda D, Bessette P, Chiquette J, Laframboise R, Lépine J, Lesperance B, Ouellette G, Pichette R, Plante M, Tavtigian SV, Simard J (2006) Mutation analysis and characterization of ATR sequence variants in breast cancer cases from high-risk French Canadian breast/ovarian cancer families. *BMC Cancer* 6:230
- 6 Durocher F, Labrie Y, Ouellette G; INHERIT BRCA, Simard J (2007) Genetic sequence variations and ADPRT haplotype analysis in French Canadian families with high risk of breast cancer. *J Hum Genet* 52(12):963-977
- 7 Guénard F, Labrie Y, Ouellette G, Beauparlant CJ, Bessette P, Chiquette J, Laframboise R, Lepine J, Lesperance B, Pichette R, Plante M; INHERIT BRCA, Durocher F (2007) Germline mutations in the breast cancer susceptibility gene PTEN are rare in high-risk non-BRCA1/2 French Canadian breast cancer families. *Fam Cancer* 6(4):483-490
- 8 Guénard F, Labrie Y, Ouellette G, Joly Beauparlant C, Simard J; INHERIT BRCA, Durocher F (2008) Mutational analysis of the breast cancer susceptibility gene BRIP1/BACH1/FANCD1 in high-risk non-BRCA1/BRCA2 breast cancer families. *J Hum Genet* 53(7):579-591
- 9 Desjardins S, Belleau P, Labrie Y, Ouellette G, Bessette P, Chiquette J, Laframboise R, Lepine J, Lesperance B, Pichette R, Plante M, Durocher F; INHERIT BRCA Other members of INHERIT BRCA involved in this study are listed in Appendix 1 (2008a) Genetic variants and haplotype analyses of the ZBRK1/ZNF350 gene in high-risk non BRCA1/2 French Canadian breast and ovarian cancer families. *Int J Cancer* 122(1):108-116
- 10 Desjardins S, Ouellette G, Labrie Y, Simard J; INHERIT BRCA, Durocher F (2008b) Analysis of GADD45A sequence variations in French Canadian families with high risk of breast cancer. *J Hum Genet* 53(6):490-498
- 11 Scully R, Anderson SF, Chao DM, Wei W, Ye L, Young RA, Livingston DM, Parvin JD (1997) BRCA1 is a component of the RNA polymerase II holoenzyme. *Proc Natl Acad Sci U S A* 94(11):5605-5610
- 12 Durocher F, Guénard F, Desjardins S, Ouellette G, Labrie Y (2005) Inherited susceptibility to breast cancer: accomplishments and challenges. In: Sinnott D (ed)

- Molecular genetics of cancer. Research Signpost, Kerala, pp 19–93
- 13 Antoniou AC, Easton D (2003) Polygenic inheritance of breast cancer: Implications for design of association studies. *Genet Epidemiol* 25(3):190-202
  - 14 Zhou H, Kuang J, Zhong L, Kuo WL, Gray JW, Sahin A, Brinkley BR, Sen S (1998) Tumour amplified kinase STK15/BTAK induces centrosome amplification, aneuploidy and transformation. *Nat Genet* 20(2):189-193
  - 15 Dutertre S, Descamps S, Prigent C (2002) On the role of aurora-A in centrosome function. *Oncogene* 21(40):6175-6183
  - 16 Ouchi M, Fujiuchi N, Sasai K, Katayama H, Minamishima YA, Ongusaha PP, Deng C, Sen S, Lee SW, Ouchi T (2004) BRCA1 phosphorylation by Aurora-A in the regulation of G2 to M transition. *J Biol Chem* 279(19):19643-19648
  - 17 Ewart-Toland A, Briassouli P, de Koning JP, Mao JH, Yuan J, Chan F, MacCarthy-Morrogh L, Ponder BA, Nagase H, Burn J, Ball S, Almeida M, Linardopoulos S, Balmain A (2003) Identification of Stk6/STK15 as a candidate low-penetrance tumor-susceptibility gene in mouse and human. *Nat Genet* 34(4):403-412
  - 18 Dai Q, Cai QY, Shu XO, Ewart-Toland A, Wen WQ, Balmain A, Gao YT, Zheng W (2004) Synergistic effects of STK15 gene polymorphisms and endogenous estrogen exposure in the risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 13(12):2065-2070
  - 19 Egan KM, Newcomb PA, Ambrosone CB, Trentham-Dietz A, Titus-Ernstoff L, Hampton JM, Kimura MT, Nagase H (2004) STK15 polymorphism and breast cancer risk in a population-based study. *Carcinogenesis* 25(11):2149-2153. Erratum in: *Carcinogenesis* 26(3):701
  - 20 Sun T, Miao X, Wang J, Tan W, Zhou Y, Yu C, Lin D (2004) Functional Phe31Ile polymorphism in Aurora A and risk of breast carcinoma. *Carcinogenesis* 25(11):2225-2230
  - 21 Ewart-Toland A, Dai Q, Gao YT, Nagase H, Dunlop MG, Farrington SM, Barnetson RA, Anton-Culver H, Peel D, Ziogas A, Lin D, Miao X, Sun T, Ostrander EA, Stanford JL, Langlois M, Chan JM, Yuan J, Harris CC, Bowman ED, Clayman GL, Lippman SM, Lee JJ, Zheng W, Balmain A (2005) Aurora-A/STK15 T+91A is a general low penetrance cancer susceptibility gene: a meta-analysis of multiple cancer types. *Carcinogenesis* 26(8):1368-1373
  - 22 Lo YL, Yu JC, Chen ST, Yang HC, Fann CS, Mau YC, Shen CY (2005) Breast cancer risk associated with genotypic polymorphism of the mitosis-regulating gene Aurora-A/STK15/BTAK. *Int J Cancer* 115(2):276-283
  - 23 Cox DG, Hankinson SE, Hunter DJ (2006) Polymorphisms of the AURKA (STK15/Aurora Kinase) Gene and Breast Cancer Risk (United States). *Cancer Causes Control* 17(1):81-83
  - 24 Fletcher O, Johnson N, Palles C, dos Santos Silva I, McCormack V, Whittaker J, Ashworth A, Peto J (2006) Inconsistent association between the STK15 F31I genetic polymorphism and breast cancer risk. *J Natl Cancer Inst* 98(14):1014-1018
  - 25 Couch FJ, Sinilnikova O, Vierkant RA, Pankratz VS, Fredericksen ZS, Stoppa-Lyonnet D, Coupier I, Hughes D, Hardouin A, Berthet P, Peock S, Cook M, Baynes C, Hodgson S, Morrison PJ, Porteous ME, Jakubowska A, Lubinski J, Gronwald J, Spurdle AB; kConFab, Schmutzler R, Versmold B, Engel C, Meindl A, Sutter C, Horst J, Schaefer D, Offit K, Kirchhoff T, Andrulis IL, Ilyushik E, Glendon G, Devilee P, Vreeswijk MP, Vasen HF, Borg A, Backenhorst K, Struwing JP, Greene



- MH, Neuhausen SL, Rebbeck TR, Nathanson K, Domchek S, Wagner T, Garber JE, Szabo C, Zikan M, Foretova L, Olson JE, Sellers TA, Lindor N, Nevanlinna H, Tommiska J, Aittomaki K, Hamann U, Rashid MU, Torres D, Simard J, Durocher F, Guenard F, Lynch HT, Isaacs C, Weitzel J, Olopade OI, Narod S, Daly MB, Godwin AK, Tomlinson G, Easton DF, Chenevix-Trench G, Antoniou AC; Consortium of Investigators of Modifiers of BRCA1/2 (2007) AURKA F31I polymorphism and breast cancer risk in BRCA1 and BRCA2 mutation carriers: a consortium of investigators of modifiers of BRCA1/2 study. *Cancer Epidemiol Biomarkers Prev* 16(7):1416-1421
- 26 Tchatchou S, Wirtenberger M, Hemminki K, Sutter C, Meindl A, Wappenschmidt B, Kiechle M, Bugert P, Schmutzler RK, Bartram CR, Burwinkel B (2007) Aurora kinases A and B and familial breast cancer risk. *Cancer Lett* 247(2):266-272
- 27 Vidarsdottir L, Bodvarsdottir SK, Hilmarsdottir H, Tryggvadottir L, Eyfjord JE (2007) Breast cancer risk associated with AURKA 91T -->A polymorphism in relation to BRCA mutations. *Cancer Lett* 250(2):206-212
- 28 Jensen DE, Proctor M, Marquis ST, Gardner HP, Ha SI, Chodosh LA, Ishov AM, Tommerup N, Vissing H, Sekido Y, Minna J, Borodovsky A, Schultz DC, Wilkinson KD, Maul GG, Barlev N, Berger SL, Prendergast GC, Rauscher FJ 3<sup>rd</sup> (1998) BAP1: a novel ubiquitin hydrolase which binds to the BRCA1 RING finger and enhances BRCA1-mediated cell growth suppression. *Oncogene* 16(9):1097-1112
- 29 Jensen DE, Rauscher FJ 3<sup>rd</sup> (1999) BAP1, a candidate tumor suppressor protein that interacts with BRCA1. *Ann N Y Acad Sci* 886:191-194
- 30 Hu JJ, Mohrenweiser HW, Bell DA, Leadon SA, Miller MS (2002) Symposium overview: genetic polymorphisms in DNA repair and cancer risk. *Toxicol Appl Pharmacol* 185(1):64-73
- 31 Coupier I, Cousin PY, Hughes D, Legoix-Ne P, Trehin A, Sinilnikova OM, Stoppa-Lyonnet D (2005) BAP1 and breast cancer risk. *Fam Cancer* 4(4):273-277
- 32 Hashizume R, Fukuda M, Maeda I, Nishikawa H, Oyake D, Yabuki Y, Ogata H, Ohta T (2001) The RING heterodimer BRCA1-BARD1 is a ubiquitin ligase inactivated by a breast cancer-derived mutation. *J Biol Chem* 276(18):14537-14540
- 33 Chen A, Kleiman FE, Manley JL, Ouchi T, Pan ZQ (2002) Autoubiquitination of the BRCA1\*BARD1 RING ubiquitin ligase. *J Biol Chem* 277(24):22085-22092
- 34 Starita LM, Horwitz AA, Keogh MC, Ishioka C, Parvin JD, Chiba N (2005) BRCA1/BARD1 ubiquitinate phosphorylated RNA polymerase II. *J Biol Chem* 280(26):24498-24505
- 35 Anderson SF, Schlegel BP, Nakajima T, Wolpin ES, Parvin JD (1998) BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A. *Nat Genet* 19(3):254-256
- 36 Thai TH, Du F, Tsan JT, Jin Y, Phung A, Spillman MA, Massa HF, Muller CY, Ashfaq R, Mathis JM, Miller DS, Trask BJ, Baer R, Bowcock AM (1998) Mutations in the BRCA1-associated RING domain (BARD1) gene in primary breast, ovarian and uterine cancers. *Hum Mol Genet* 7(2):195-202
- 37 Ghimenti C, Sensi E, Presciuttini S, Brunetti IM, Conte P, Bevilacqua G, Caligo MA (2002) Germline mutations of the BRCA1-associated ring domain (BARD1) gene in breast and breast/ovarian families negative for BRCA1 and BRCA2 alterations. *Genes Chromosomes Cancer* 33(3):235-242
- 38 Karppinen SM, Heikkinen K, Rapakko K, Winqvist R (2004) Mutation screening of

- the BARD1 gene: evidence for involvement of the Cys557Ser allele in hereditary susceptibility to breast cancer. *J Med Genet* 41(9):e114
- 39 Stacey SN, Sulem P, Johannsson OT, Helgason A, Gudmundsson J, Kostic JP, Kristjansson K, Jonsdottir T, Sigurdsson H, Hrafnkelsson J, Johannsson J, Sveinsson T, Myrdal G, Grimsson HN, Bergthorsson JT, Amundadottir LT, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K (2006) The BARD1 Cys557Ser variant and breast cancer risk in Iceland. *PLoS Med* 3(7):e217
- 40 Wu JY, Vlastos AT, Pelte MF, Caligo MA, Bianco A, Krause KH, Laurent GJ, Irminger-Finger I (2006) Aberrant expression of BARD1 in breast and ovarian cancers with poor prognosis. *Int J Cancer* 118(5):1215-1226
- 41 Ishitobi M, Miyoshi Y, Hasegawa S, Egawa C, Tamaki Y, Monden M, Noguchi S (2003) Mutational analysis of BARD1 in familial breast cancer patients in Japan. *Cancer Lett* 200(1):1-7
- 42 Karppinen SM, Barkardottir RB, Backenhorn K, Sydenham T, Syrjakoski K, Schleutker J, Ikonen T, Pylkas K, Rapakko K, Erkko H, Johannesdottir G, Gerdes AM, Thomassen M, Agnarsson BA, Grip M, Kallioniemi A, Kere J, Aaltonen LA, Arason A, Moller P, Kruse TA, Borg A, Winqvist R (2006) Nordic collaborative study of the BARD1 Cys557Ser allele in 3956 patients with cancer: enrichment in familial BRCA1/BRCA2 mutation-negative breast cancer but not in other malignancies. *J Med Genet* 43(11):856-862
- 43 Jakubowska A, Cybulski C, Szymanska A, Huzarski T, Byrski T, Gronwald J, Debniak T, Gorski B, Kowalska E, Narod SA, Lubinski J (2008) BARD1 and breast cancer in Poland. *Breast Cancer Res Treat* 107(1):119-122
- 44 Pharoah PD, Tyrer J, Dunning AM, Easton DF, Ponder BA; SEARCH Investigators (2007) Association between common variation in 120 candidate genes and breast cancer risk. *PLoS Genet* 3(3):e42
- 45 Huo X, Hu Z, Zhai X, Wang Y, Wang S, Wang X, Qin J, Chen W, Jin G, Liu J, Gao J, Wei Q, Wang X, Shen H (2007) Common non-synonymous polymorphisms in the BRCA1 associated RING Domain (BARD1) gene are associated with breast cancer susceptibility: a case-control analysis. *Breast Cancer Res Treat* 102(3):329-337
- 46 Lee CG, Hurwitz J (1992) A new RNA helicase isolated from HeLa cells that catalytically translocates in the 3' to 5' direction. *J Biol Chem* 267(7):4398-4407
- 47 Nakajima T, Uchida C, Anderson SF, Lee CG, Hurwitz J, Parvin JD, Montminy M (1997) RNA helicase A mediates association of CBP with RNA polymerase II. *Cell* 90(6):1107-1112
- 48 Srebrow A, Kornblihtt AR (2006) The connection between splicing and cancer. *J Cell Sci* 119(Pt 13):2635-41
- 49 Dicioccio RA, Song H, Waterfall C, Kimura MT, Nagase H, McGuire V, Hogdall E, Shah MN, Luben RN, Easton DF, Jacobs IJ, Ponder BA, Whittemore AS, Gayther SA, Pharoah PD, Kruger-Kjaer S (2004) STK15 polymorphisms and association with risk of invasive ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 13(10):1589-1594
- 50 Goringe KL, Choong DY, Visvader JE, Lindeman GJ, Campbell IG (2008) BARD1 variants are not associated with breast cancer risk in Australian familial breast cancer. *Breast Cancer Res Treat*. 2008 Oct;111(3):505-9
- 51 Fortin J, Moisan AM, Dumont M, Leblanc G, Labrie Y, Durocher F, Bessette P, Bridge P, Chiquette J, Laframboise R, Lépine J, Lespérance B, Pichette R, Plante M,

- Provencher L, Voyer P, Simard J (2005) A new alternative splice variant of BRCA1 containing an additional in-frame exon. *Biochim Biophys Acta* 1731(1):57-65
- 52 Lombardi G, Falaschi E, Di Cristofano C, Naccarato AG, Sensi E, Aretini P, Roncella M, Bevilacqua G, Caligo MA (2007) Identification of novel alternatively spliced BRCA1-associated RING domain (BARD1) messenger RNAs in human peripheral blood lymphocytes and in sporadic breast cancer tissues. *Genes Chromosomes Cancer* 46(9):791-795
- 53 Tsuzuki M, Wu W, Nishikawa H, Hayami R, Oyake D, Yabuki Y, Fukuda M, Ohta T (2006) A truncated splice variant of human BARD1 that lacks the RING finger and ankyrin repeats. *Cancer Lett* 233(1):108-116
- 54 Shin SO, Lee KH, Kim JH, Baek SH, Park JW, Gabrielson EW, Kwon TK (2000) Alternative splicing in 5'-untranslational region of STK-15 gene, encoding centrosome associated kinase, in breast cancer cell lines. *Exp Mol Med* 32(4):193-196
- 55 Devlin B, Risch N (1995) A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* 29(2):311-322
- 56 Vézina H, Durocher F, Dumont M, Houde L, Szabo C, Tranchant M, Chiquette J, Plante M, Laframboise R, Lépine J, Nevanlinna H, Stoppa-Lyonnet D, Goldgar D, Bridge P, Simard J (2005) Molecular and genealogical characterization of the R1443X BRCA1 mutation in high-risk French-Canadian breast/ovarian cancer families. *Hum Genet* 117(2-3):119-132

## **Titles and legends to figures**

### **Figure 1**

Pedigree of *AURKA*:Met373Val carrier's family. The age of onset of cancer, age of death or current age are indicated below family members. The affected case that was sequenced first is designated by an *arrow*. The genotype call is indicated below the respective individual analyzed. *d*, deceased.

### **Figure 2**

Graphical overview of pairwise linkage disequilibrium (LD) measures of  $|D'|$  (lower left part) and  $r^2$  (higher right part) for each SNPs pair identified in breast cancer cases and healthy individuals. All SNPs are denoted numerically with reference to Table 1. Distance between first and last SNPs for each gene is indicated on the left end on each LD representation.

### **Figure 3**

Haplotype blocks and tSNPs identified for the *AURKA* (A), *BAP1* (B), *BARD1* (C) and *DHX9* (D) genes. Tagging SNPs identified on a block-by-block basis using haplotypes showing a frequency higher than 1% are denoted with an asterisk (\*). Haplotype frequencies are displayed on the right of each haplotype combination while the level of recombination is displayed above the connections between two blocks. Thin connections represent haplotypes with frequencies between 1 and 10% while haplotypes with frequencies higher than 10% are represented by thick lines.

Table 1  
Sequence variations and genotype frequencies of *AURKA*, *BAP1*, *BARD1* and *DHX9* genes in cases and controls

Gene	SNP	SNP ID <sup>1</sup>	dbSNP ID	Amino acid change	Location	Series	Number of individuals	Common homozygote (expected) <sup>2</sup>	Heterozygote (expected) <sup>2</sup>	Rare homozygote (expected) <sup>2</sup>	Minor allele frequency (MAF)	Dev from HWE pvalue
<b>AURKA</b>	1	c.91T>A	rs2273535	Phe31Ile	Exon 4	Cases	96	64 (61.0)	25 (31.1)	7 (4.0)	0.203	0.06
						Controls	96	63 (61.8)	28 (30.5)	5 (3.8)	0.198	0.43
	2	c.169G>A	rs1047972	Val57Ile	Exon 4	Cases	96	62 (61.8)	30 (30.5)	4 (3.8)	0.198	0.88
						Controls	96	69 (67.5)	23 (26.0)	4 (2.5)	0.161	0.26
	3	c.638C>T	N/A	Ala213Val	Exon 7	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.005	0.96
						Controls	97	97 (0.0)	0 (0.0)	0 (0.0)	0.000	---
	4	c.1117A>G	rs33923703	Met373Val	Exon 10	Cases	96	92 (92.0)	4 (3.9)	0 (0.0)	0.021	0.83
						Controls	95	89 (89.1)	6 (5.8)	0 (0.1)	0.032	0.75
	5	c.1130T>A	N/A	Val377Glu	Exon 10	Cases	96	93 (93.0)	3 (3.0)	0 (0.0)	0.016	0.88
						Controls	96	96 (0.0)	0 (0.0)	0 (0.0)	0.000	---
<b>BAP1</b>	1	c.1212+3A>C	rs34361943	-	3'-UTR	Cases	96	92 (92.0)	4 (3.9)	0 (0.0)	0.021	0.83
						Controls	95	89 (89.1)	6 (5.8)	0 (0.1)	0.032	0.75
	2	c.1212+69G>A	rs34189236	-	3'-UTR	Cases	96	94 (94.0)	2 (2.0)	0 (0.0)	0.010	0.92
						Controls	95	94 (94.0)	1 (1.0)	0 (0.0)	0.005	0.96
	3	c.1212+370C>T	N/A	-	3'-UTR	Cases	96	94 (94.0)	2 (2.0)	0 (0.0)	0.010	0.92
						Controls	95	95 (0.0)	0 (0.0)	0 (0.0)	0.000	---
	4	c.1212+422C>G	N/A	-	3'-UTR	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.005	0.96
						Controls	95	95 (0.0)	0 (0.0)	0 (0.0)	0.000	---
	5	c.1212+423G>C	rs8173	-	3'-UTR	Cases	96	60 (56.3)	27 (34.5)	9 (5.3)	0.234	0.03
						Controls	95	54 (54.6)	36 (34.9)	5 (5.6)	0.242	0.75
<b>BAP1</b>	1	c.294C>T	N/A	Ser98Ser	Exon 5	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.005	0.96
						Controls	94	94 (0.0)	0 (0.0)	0 (0.0)	0.000	---
	2	c.1002A>G	rs28997577	Leu334Leu	Exon 11	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.005	0.96
					Controls	92	91 (91.0)	1 (1.0)	0 (0.0)	0.005	0.96	
3	c.1026C>T	N/A	Ser342Ser	Exon 11	Cases	96	96 (0.0)	0 (0.0)	0 (0.0)	0.000	---	
					Controls	93	91 (91.0)	2 (2.0)	0 (0.0)	0.011	0.92	

<b>BARDI</b>	4	c.2190+444C>T	rs123598	-	3'-UTR	Cases	96	81 (80.7)	14 (14.7)	1 (0.7)	0.083	0.66
						Controls	87	74 (73.6)	12 (12.9)	1 (0.6)	0.080	0.53
	1	c.70C>T	rs1048108	Pro24Ser	Exon 1	Cases	96	52 (48.9)	33 (39.2)	11 (7.9)	0.286	0.12
						Controls	95	46 (45.2)	39 (40.7)	10 (9.2)	0.311	0.69
	2	c.609A>C	rs28997574	Gly203Gly	Exon 4	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.005	0.96
						Controls	86	81 (81.1)	5 (4.8)	0 (0.1)	0.029	0.78
	3	c.820A>G	N/A	Ser274Gly	Exon 4	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.005	0.96
						Controls	90	90 (0.0)	0 (0.0)	0 (0.0)	0.000	---
	4	c.1053G>C	rs2070096	Thr351Thr	Exon 4	Cases	96	56 (59.4)	39 (32.2)	1 (4.4)	0.214	0.04
	5	c.1075del21	rs28997575	delLPECSSP	Exon 4	Cases	96	92 (92.0)	4 (3.9)	0 (0.0)	0.021	0.83
				aa 359-365		Controls	87	85 (85.0)	2 (2.0)	0 (0.0)	0.000	0.91
6	c.1134G>C	rs2229571	Arg378Ser	Exon 4	Cases	96	12 (13.9)	49 (45.2)	35 (36.9)	0.620	0.42	
7	c.1203T>C	N/A	Ser401Ser	Exon 4	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.005	0.96	
					Controls	87	87 (0.0)	0 (0.0)	0 (0.0)	0.000	---	
8	c.1518C>T	rs2070093	His506His	Exon 6	Cases	96	52 (52.5)	38 (37.0)	6 (6.5)	0.260	0.79	
9	c.1519G>A	rs2070094	Val507Met	Exon 6	Controls	98	68 (69.5)	29 (26.1)	1 (2.5)	0.158	0.27	
10	c.1670G>C	rs28997576	Cys557Ser	Exon 7	Cases	96	93 (93.0)	3 (3.0)	0 (0.0)	0.016	0.88	
					Controls	98	95 (94.0)	2 (3.9)	1 (0.0)	0.020	<0.01	
11	c.2212A>G	N/A	Ile738Val	Exon 11	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.005	0.96	
					Controls	97	92 (92.1)	5 (4.9)	0 (0.1)	0.026	0.79	
<b>DHX9</b>	1	c.265C>G	N/A	Pro89Ala	Exon 4	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.005	0.96
						Controls	94	94 (0.0)	0 (0.0)	0 (0.0)	0.000	---
	2	c.915T>G	N/A	Ser305Ser	Exon 10	Cases	96	93 (93.0)	3 (3.0)	0 (0.0)	0.016	0.88
						Controls	96	92 (92.0)	4 (3.9)	0 (0.0)	0.021	0.83
	3	c.1745C>T	rs2275177	Thr582Thr	Exon 16	Cases	96	81 (78.8)	12 (16.3)	3 (0.8)	0.094	0.01
						Controls	91	80 (79.4)	10 (11.2)	1 (0.4)	0.066	0.30

4	c.1873C>G	N/A	Ser625Cys	Exon 17	Cases	96	96 (0.0)	0 (0.0)	0 (0.0)	0.000	---
					Controls	91	90 (90.0)	1 (1.0)	0 (0.0)	0.005	0.96
5	c.1976C>T	rs35533583	Gly659Gly	Exon 17	Cases	96	91 (91.1)	5 (4.9)	0 (0.1)	0.026	0.79
					Controls	91	84 (84.1)	7 (6.7)	0 (0.1)	0.038	0.70
6	c.2708C>T	N/A	p.Ile903Ile	Exon 23	Cases	96	95 (95.0)	1 (1.0)	0 (0.0)	0.005	0.96
					Controls	94	93 (93.0)	1 (1.0)	0 (0.0)	0.005	0.96

<sup>1</sup> According to the nomenclature of the Human Genome Variation Society

<sup>2</sup> As expected under Hardy-Weinberg equilibrium

Table 2

Non-synonymous sequence variants detected in *AURKA*, *BAP1*, *BARD1* and *DHX9* genes and corresponding residues present in orthologues

Gene	SNP <sup>1</sup>	SNP ID <sup>2</sup>	Amino acid change	<i>Canis familiaris</i>	<i>Mus musculus</i>	<i>Rattus norvegicus</i>	<i>Monodelphis domestica</i>	<i>Gallus gallus</i>	<i>Xenopus tropicalis</i>	<i>Tetraodon nigroviridis</i>
<b>AURKA</b>	1	c.91T>A	Phe31Ile	Val	Ile	Ile	Ile	-	-	-
	2	c.169G>A	Val57Ile	Val	Val	Val	Val	Val	-	-
	3	c.638C>T	Ala213Val	Ala	Ala	Ala	Ala	Ala	Ala	Ala
	4	c.1117A>G	Met373Val	Thr	Thr	Thr	Thr	Pro	Ser	Pro
	5	c.1130T>A	Val377Glu	Val	Val	Val	Leu	Val	Val	Val
<b>BARD1</b>	1	c.70C>T	Pro24Ser	Pro	Pro	Pro	Ser	Ala	-	-
	3	c.820A>G	Ser274Gly	Ser	Ser	Ser	Val	Asn	-	-
	5	c.1075del21	delILPECSSP	LPGCSSP	LVDCSSL	LVDCSSL	LTESAPS	DCPIP	QR-----	-
	6	c.1134G>C	Arg378Ser	Arg	Arg	Arg	Val	Lys	Arg	-
	9	c.1519G>A	Val507Met	Leu	Ile	Ile	Val	Val	Ile	Leu
<b>DHX9</b>	10	c.1670G>C	Cys557Ser	Tyr	Cys	Cys	Cys	Pro	Cys	-
	11	c.2212A>G	Ile738Val	Ile	Val	Val	Ile	Ile	Val	Ile
	1	c.265C>G	Pro89Ala	Pro	Pro	Pro	-	-	-	-
	4	c.1873C>G	Ser625Cys	Ala	Ser	Ser	Ala	-	Ser	Ala

<sup>1</sup> According to Table 1

<sup>2</sup> According to the nomenclature of the Human Genome Variation Society

- No corresponding residue



Table 3  
Location of non-synonymous sequence variants detected in *AURKA*, *BAP1*, *BARD1* and *DHX9* genes and prediction on protein function

Gene	SNP <sup>1</sup>	Amino acid change	Domain	SIFT	PolyPhen (score difference)
<b>AURKA</b>	1	Phe31Ile	None	Tolerated	Benign
	2	Val57Ile	None	Tolerated	Benign
	3	Ala213Val	Ser/Thr protein kinase	Not Tolerated	Benign
	4	Met373Val	Ser/Thr protein kinase	Not Tolerated	Benign
	5	Val377Glu	Ser/Thr protein kinase	Not Tolerated	Probably damaging (2.39)
<b>BARD1</b>	1	Pro24Ser	None	Not Tolerated	Benign
	3	Ser274Gly	None	Tolerated	Benign
	5	delLPECSSP	None	N/A	N/A
	6	Arg378Ser	None	Not Tolerated	Possibly damaging (1.79)
	9	Val507Met	Ankyrin repeats	Not Tolerated	Benign
	10	Cys557Ser	close from BRCT (aa 568-767)	Tolerated	Probably damaging (2.52)
<b>DHX9</b>	11	Ile738Val	BRCT	Tolerated	Benign
	1	Pro89Ala	CBP-binding domain (aa 1-250)	Tolerated	Possibly damaging (1.56)
	4	Ser625Cys	RNA pol II binding domain (aa 230-650)	Not Tolerated	Possibly damaging (1.53)

<sup>1</sup> According to Table 1

Table 4  
Estimated haplotype frequencies in cases and controls using PHASE

Gene	Haplotype number <sup>1</sup>	Haplotypes (all SNPs used) <sup>2</sup>	Haplotypes (tSNPs only)	Estimated haplotype frequencies (tSNPs only)		
				Cases	Controls	Cumulative
<b>AURKA</b>	h1	TGCATAGCCG	TGA <sup>3</sup>	0.600	0.635	0.618
	h2	AGCATAGCCC	AGA <sup>3</sup>	0.196	0.186	0.809
	h3	TACATAGCCG	TAA <sup>3</sup>	0.182	0.145	0.973
	h4	AGCATAGCCG	AGA <sup>3</sup>	---	---	---
	h5	TGCATAGCCC	TGA <sup>3</sup>	---	---	---
	h6	TACATAGCCC	TAA <sup>3</sup>	---	---	---
<b>BAPI</b>	h1	CACC	C	0.917	0.919	0.918
	h2	CACT	T	0.083	0.081	1.000
<b>BARDI</b>	h1	TAAGACTCAGA	TAGCCAGA	0.214	0.227	0.220
	h2	CAACAGTCGGA	CACGCGGA	0.187	0.233	0.430
	h3	CAAGACTTGA	CAGCTGGA	0.176	0.099	0.567
	h4	CAAGACTCGGA	CAGCCGGA	0.097	0.088	0.660
	h5	CAAGAGTCGGA	CAGGCGGA	0.083	0.092	0.748
	h6	CAAGACTCAGA	CAGCCAGA	0.053	0.080	0.814
	h7	CAAGAGTTGA	CAGGTGGA	0.061	0.045	0.867
	h8	TAAGACTCGGA	TAGCCGGA	0.039	0.044	0.909
	h9	CAAGACTCACA	CAGCCACA	0.021	0.020	0.930
	h10	CCAGACTCAGG	CCGCCAGG	0.005	0.025	0.945
	h11	TAAGAGTCGGA	TAGGCGGA	0.014	0.013	0.959
<b>DHX9</b>	h1	CTCCC	TCC	0.905	0.887	0.856
	h2	CTTCC	TTC	0.036	0.072	0.910
	h3	CTTCTC	TTT	0.031	0.020	0.935
	h4	CGCCC	GCC	0.021	0.015	0.953

<sup>1</sup> Only haplotypes exhibiting a frequency  $\geq 1\%$  are displayed.

<sup>2</sup> Wild type allele for the *BARDI* SNP 5 (c.1075del21) is represented by the letter T in haplotypes while variant allele for this SNP is not present in haplotypes exhibiting a frequency  $\geq 1\%$ .

<sup>3</sup> These tSNPs haplotypes both tag simultaneously a common and a rare *AURKA* haplotypes estimated using all SNPs.

Figure 1

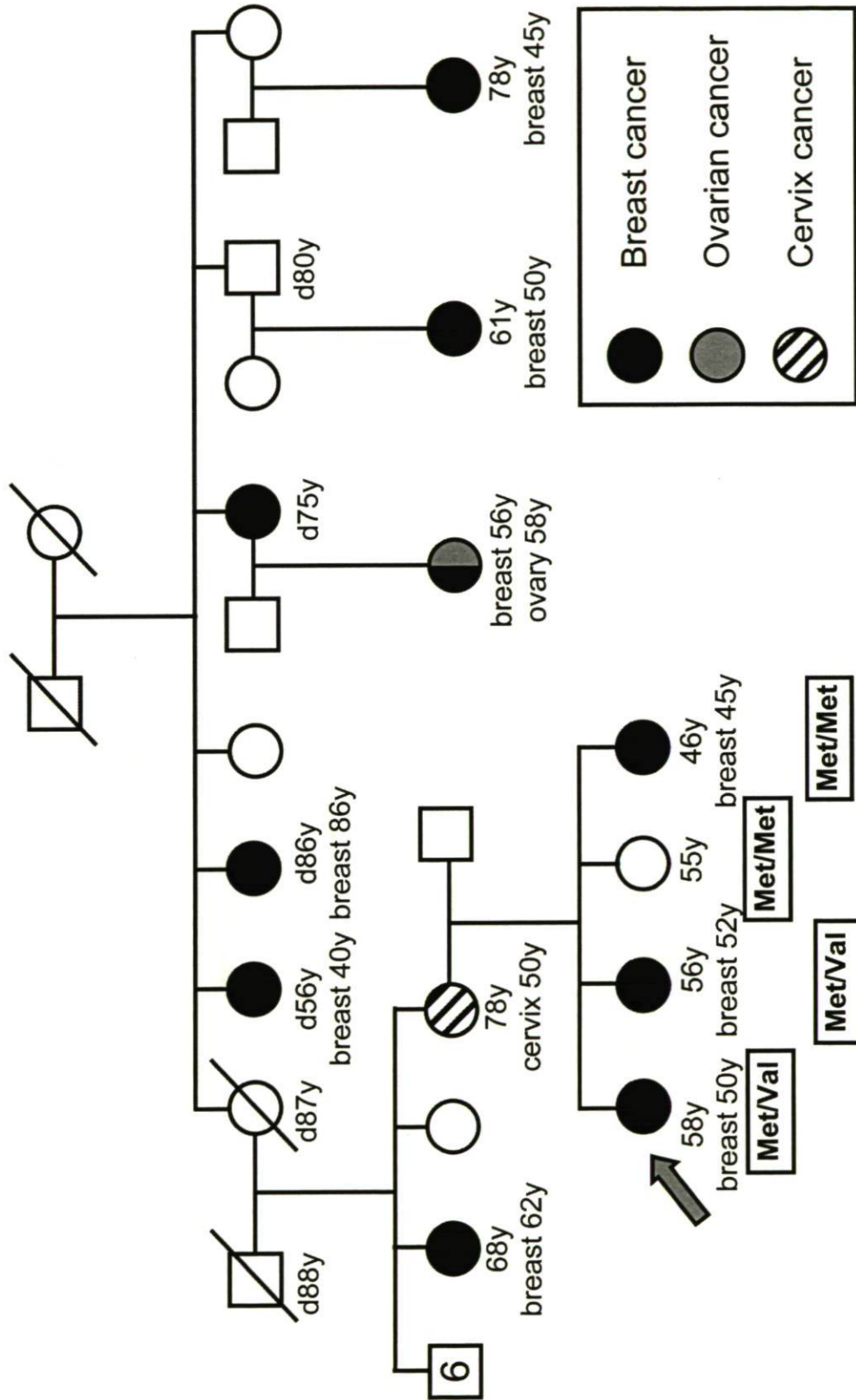


Figure 2

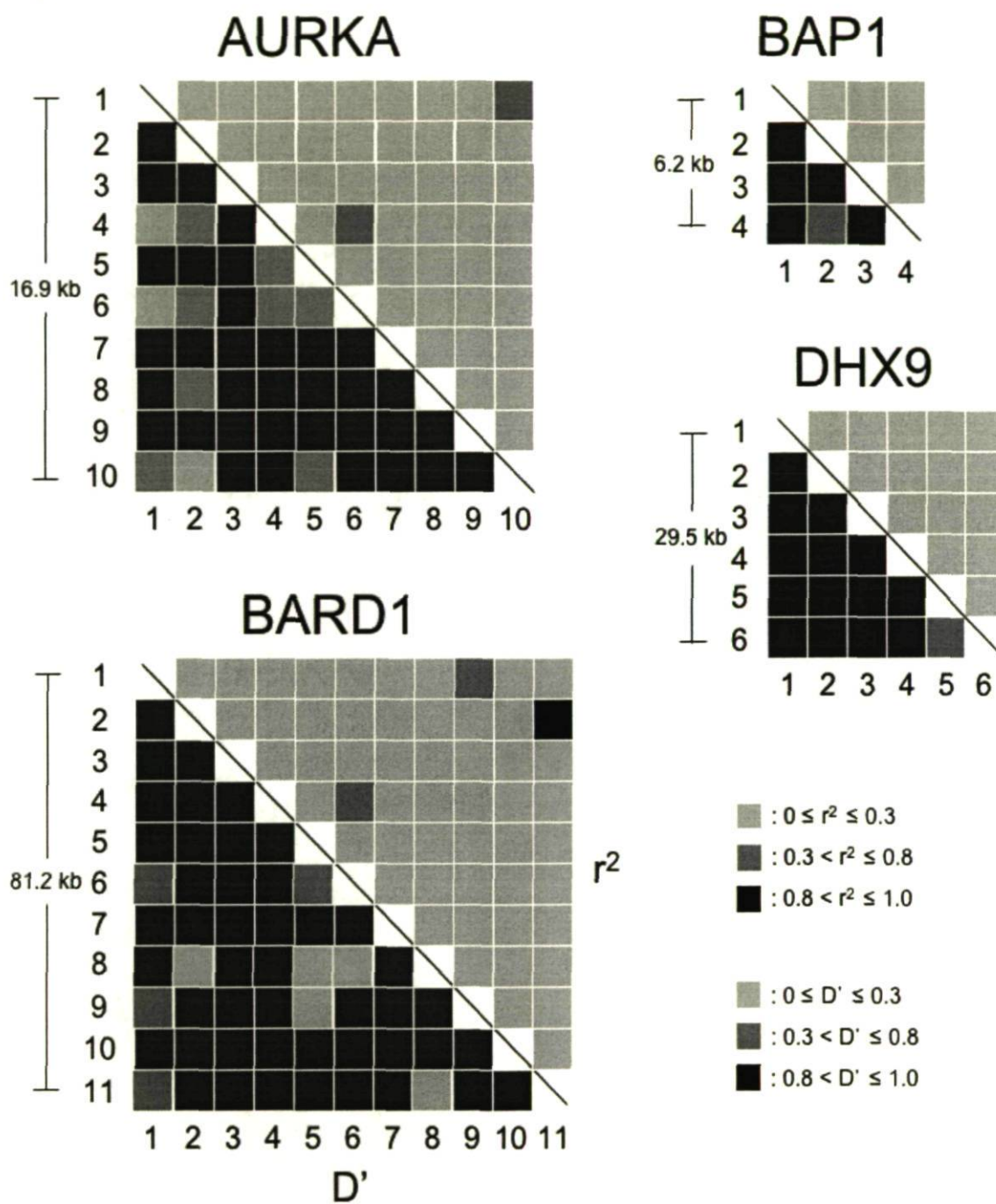


Figure 3

## A: AURKA

rs2273535 \*  
rs1047972 \*

1 2 3  
| | |

TGC .619  
AGC .196  
TAC .183



rs34361943  
rs34189236 \*

5 6 7 8 9  
| | | | |

TAGCC .955  
TCGCC .021

## B: BAP1

rs28997577  
rs123598 \*

1 2 3 4  
| | | |

CACC .910  
CACT .077

## C: BARD1

rs1048108 \*  
rs28997574 \*

1 2 3  
| | |

CAA .676  
TAA .305  
CAA .015



rs2070096 \*  
rs28997575  
rs2229571 \*  
rs2070093 \*  
rs2070094 \*  
rs28997576 \* \*

4 5 6 7 8 9 10 11  
| | | | | | | | |

CAGTCGGA .211  
GACTCAGA .296  
GACTTGGA .136  
GACTCGGA .127  
GAGTCGGA .094  
GAGTTGGA .059  
GACTCACA .022  
GACTCAGG .016

## D: DHX9

rs2275177 \*  
rs35533583 \*

1 2 3 4 5 6  
| | | | | |

CTCCCC .888  
CTTCCC .054  
CTTCTC .026  
CGCCCC .019

Supplemental table 1  
Oligonucleotides used to amplify and sequence *AURKA*, *BAP1*, *BARD1* and *DXH9* genes in familial breast cancer (cDNA) and controls (gDNA).

Gene	Material	Fragment	Forward primer	Reverse primer	Sequence primer	Annealing (°C)	Fragment length (bp)
AURKA	cDNA	AURKA-c1	CCAGAGTGGGGGATATTGA	GCTCTGAGACACATGGCCCTCT	GATCCGAAGTGGGACTGT, GAGATCCACCCTTCTCATCATGCA, GCACCACCTTGAACACAGTTTATAGAGA	66	1991
	gDNA	AURKA-ex4	GGGTGAGGAATTGGAGGGGAT	GGACACCAATTTATGCTGTGTCCT	GGGTGAGGAATTGGAGGGGAT	64	749
	gDNA	AURKA-ex7	GGGTTGGAGTGTGGTTTCATG	CAGCTGTGCAAACTAGGCACAA	GGGTTGGAGTGTGGTTTCATG	60	490
	gDNA	AURKA-ex10	CAGGCCAAGATCCGAACATCT	GCTCTGAGACACATGGCCCTCT	CAGGCCAAGATCCGAACATCT and GGGAGATAAGTGGTTAAGGAGGACT	64	948
BAP1	cDNA	BAP1-c1	GCCTGGACATGGCGCTGA	CTCCTCTGCACCATCTGTGTGGT	GCCTGGACATGGCGCTGA and ACTCTTATCAGTGTCTGCAGAGCCT	58	982
	cDNA	BAP1-c2	CGGGTCATCATGGAGCGTAT	GGACAGCACGGTTGTAGCGT	CGGGTCATCATGGAGCGTAT and GGACAGCACGGTTGTAGCGT	58	1030
	cDNA	BAP1-c3	TGAGATCGGCAGTGCTTTCAA	CCCCAGCTAGGACCCTGTAGTT	TGAGATCGGCAGTGCTTTCAA and CCCCAGCTAGGACCCTGTAGTT	58	1167
	cDNA	BAP1-c4	GGCACAGCGAGGTA CTGCA	CAGAGGCAGCTGCCACACTT	GGCTGTTCAAGACTGCTCTCCA and CAGAGGCAGCTGCCACACTT	58	975
	gDNA	BAP1-ex5	CAGATAGACTGACCTGCTCTGGA	CCTTTGCTCTACGGGGAAGAAAATA	CAGATAGACTGACCTGCTCTGGA	60	727
	gDNA	BAP1-ex11	TTGCCCTGTCTCATTGCGGGAT	AATCCACAGCCACGTA ACTAGAT	AATCCACAGCCACGTA ACTAGAT	60	586
	gDNA	BAP1-ex17-1'	GGGATGCGCTGCTGCTTAACT	TAGGGCCACAACACTGAAGGGCGAA	GGGATGCGCTGCTGCTTAACT and TAGGGCCACAACACTGAAGGGCGAA	60	1096
	gDNA	BAP1-ex17-2'	AGACTGCTCTCCATAGCAAGGTT	GCCTAGGATTC CAAGGATCTCT	GCCTAGGATTC CAAGGATCTCT	60	697
BARD1	cDNA	BARD1-c1	CTTCCCTGTGGTTTCCCGA	CAGCCATATGGGCAACAGCT	CTTCCCTGTGGTTTCCCGA and CAGCCATATGGGCAACAGCT	62	1335
	cDNA	BARD1-c2	GACTGGGCTCTCACAAACCGT	TGTTCTTACCACCTTCATGCA	GACTGGGCTCTCACAAACCGT and TGTTCTTACCACCTTCATGCA	60	1310
	gDNA	BARD1-ex1	CTTCCCTGTGGTTTCCCGA	GGTTCCTCAGCCCGCTGAAGT	GGTTCCTCAGCCCGCTGAAGT	58	490
	gDNA	BARD1-ex4	GATATGGCACAAAGTCGGATTCAGT	GCCACTGCTCCA ACTGAAGTAGAGA	GATATGGCACAAAGTCGGATTCAGT and GGCCTTCTCTGTTTCAGAGGAAGT	62	1335

gDNA	BAR1-ex6	CACGGATACTTACTTTTGGTTCATGA	GCTGAATCCTCAAGTGTAGAACCCA	CCAAGGCTCAAAATGAGAATAGCCT	62	808
gDNA	BAR1-ex7	CATTTCTCCTTTGAGTTCAGCAGCT	CAACGACACTCTACAACATCTTCGA	CATTTCTCCTTTGAGTTCAGCAGCT	62	853
gDNA	BAR1-ex11	GAATCTGAAATGCTTCTCTCTCCAA	GGCTCTCACAAAACCGTGCAA	GGCTCTCACAAAACCGTGCAA	61	522
cDNA	DHX9-c1	GGTAGAAGGCCAGAGTCACA	AGTTGGATTGTGGAGGTGACCA	GAATTCGGCCAGAGTCACA and AGTTGGATTGTGGAGGTGACCA	60	1095
cDNA	DHX9-c2	GGCAGAAAGGATTTTGCACGA	TGGGGCAATTGAAGAAAATTTACAAA	GGCAGAAAGGATTTTGCACGA and TGGGGCAATTGAAGAAAATTTACAAA	60	1012
cDNA	DHX9-c3	GTTTTGTAAGTGTGCTCCTG	GATTCGTCCTCCAAAGGAGTCAAC	GTTTTGTAAGTGTGCTCCTG and GATTCGTCCTCCAAAGGAGTCAAC	60	1102
cDNA	DHX9-c4	CACATGACACCAGAGATGTTCCG	GATGATTTGTGGATAAGTGCAATTACG	CACATGACACCAGAGATGTTCCG and GATGATTTGTGGATAAGTGCAAG	60	716
cDNA	DHX9-c5	AATGTATGCTATCATAAGGAAAAGAGG	CAATGAACCACAGAATGGGTG	CAATGAACCACAGAATGGGTG and CTGCCGGAGACCAGTGATA	60	941
gDNA	DHX9-ex4	ATTTACTAGTCCCGTGAAGGCAA	AGCAAAGAAGTATGCTGTACCA	AGCAAAGAAGTATGCTGTACCA	60	373
gDNA	DHX9-ex10	TTCCATAGGTGGAGCCTTACAAA	TGCTCTGGAGTAGCCTGTAAAA	TGCTCTGGAGTAGCCTGTAAAA	60	512
gDNA	DHX9-ex16	AAACAAGTGGCCTACACTACCT	CCAGATGCTGCAGCCTACTG	AAACAAGTGGCCTACACTACCT and CCAGATGCTGCAGCCTACTG	62	1150
gDNA	DHX9-ex17	GTAGTGCTTGAGAGTTGGACGAC	CTCCAAGTGGTACTGGATCAAAAC	GTAGTGCTTGAGAGTTGGACGAC and CTCCAAGTGGTACTGGATCAAAAC	62	1021
gDNA	DHX9-ex23	TATGTGGCCATTAGAGCATCAGA	GGAAGTACAGCAGTGTTCACGA	TATGTGGCCATTAGAGCATCAGA	60	745

<sup>1</sup> BAP1 exon 17 has been amplified in two fragments from 5' position to 3' position.

## DISCUSSION

Le cancer est une maladie qui atteint une proportion importante des Canadiennes et des Canadiens. Chaque année, un nombre grandissant de nouveaux cas sont diagnostiqués. Le cancer du sein est celui qui a la plus forte incidence chez les Canadiennes<sup>1</sup>. Il en est de même aux États-Unis<sup>3</sup> et à l'échelle mondiale<sup>2</sup>.

Des facteurs génétiques et environnementaux sont impliqués dans l'étiologie du cancer du sein, conférant à cette maladie le titre de maladie multifactorielle<sup>9</sup>. Tel que présenté dans la section 2.1.2, plusieurs facteurs environnementaux conférant un risque accru de cancer du sein sont liés de près ou de loin avec l'exposition aux estrogènes. D'autres facteurs environnementaux, notamment de nature démographique ou liés aux habitudes de vie ou de nutrition, sont aussi impliqués dans le cancer du sein. La démonstration de liaison génétique chez des familles présentant plusieurs cas de cancer du sein<sup>39</sup>, le risque accru de cancer chez les femmes ayant une apparentée de premier degré atteinte du cancer du sein<sup>16</sup>, les études menées chez des jumeaux<sup>40</sup> et la distribution de l'hérédité chez certaines familles<sup>44</sup> ont démontré que les facteurs génétiques jouent un rôle prépondérant dans l'étiologie de cette maladie. Il a été proposé que plusieurs gènes seraient responsables de la susceptibilité génétique au cancer du sein, que ces gènes conféreraient un large éventail de risque<sup>388</sup>, et qu'ils pourraient agir entre eux d'une manière multiplicative<sup>44</sup>.

Les gènes majeurs de susceptibilité au cancer du sein *BRCA1* et *BRCA2* sont responsables d'environ 24% des familles à risque élevé de cancer du sein provenant de la population canadienne-française<sup>494</sup>. Des mutations fondatrices de *BRCA1* et *BRCA2* ont d'ailleurs été identifiées dans cette population<sup>494,496</sup>. Dans le cadre de mes études doctorales, j'ai évalué la contribution des gènes de susceptibilité au cancer du sein à forte pénétrance *PTEN* et *STK11* dans notre cohorte de cas familiaux de cancer du sein provenant de la population canadienne-française. Alors qu'aucune mutation de *PTEN* n'a été identifiée, une mutation de *STK11* rapportée comme étant une mutation responsable du syndrome de Peutz-Jeghers<sup>236</sup> a été identifiée. Cette mutation pourrait potentiellement expliquer l'agrégation de cancer du sein chez une famille de notre cohorte. Néanmoins, l'analyse des gènes de susceptibilité au cancer du sein *PTEN* et *STK11* nous permet d'affirmer que ces gènes de forte pénétrance ne semblent pas responsables d'une proportion significative des cas de



cancer familiaux chez les femmes exemptes de mutations dans les gènes *BRCA1* et *BRCA2*. Toutefois, certains des résultats obtenus pour le gène *STK11* laissent présager une implication possible des variations introniques communes de ce gène dans la susceptibilité au cancer du sein. Cela demande par contre confirmation dans de plus grandes cohortes. Par exemple, pour confirmer l'association d'un variant retrouvé à une fréquence de 0,30 et conférant un risque relatif de 2,0, il faudrait mettre sur pied une étude d'association contenant 400 cas et 400 contrôles (puissance statistique de 90%; seuil statistique de  $P < 10^{-4}$ )<sup>495</sup>.

L'analyse des gènes de susceptibilité à pénétrance modérée *CHEK2*, *BRIP1* et *PALB2* a aussi été effectuée dans le cadre de mon projet. La mutation c.1100delC de *CHEK2* a été identifiée chez deux familles. La fréquence à laquelle cette mutation a été retrouvée dans notre cohorte de cas de cancers du sein semble écarter la présence d'un effet fondateur de ce gène dans la population canadienne-française, bien que cette mutation puisse contribuer à une faible proportion des cas de cancer du sein familiaux dans cette population<sup>497,498</sup>. L'analyse des gènes *BRIP1* et *PALB2* n'a pas permis d'identifier de mutation tronquante de ces gènes. Des analyses de gène rapporteur ont toutefois suggéré qu'un variant situé dans le promoteur de *BRIP1* pourrait être associé à une activité de transactivation diminuée<sup>499</sup>. Bien que notre cohorte n'ait pas permis de détecter la présence de la mutation fondatrice canadienne-française de *PALB2* (Q775X) pour laquelle la fréquence des porteurs a été estimée à 0,56% chez des cas de cancer du sein à développement précoce<sup>373</sup>, nos analyses viennent confirmer l'absence de grands réarrangements génomiques du gène *PALB2* chez des cas de cancer du sein provenant de cette population<sup>373</sup>. Notre analyse n'a toutefois pas permis de révéler la présence de la mutation fondatrice de *PALB2*. Ce volet de mon projet de doctorat a donc permis d'éliminer la possibilité d'une contribution substantielle des gènes *PTEN*, *STK11*, *CHEK2*, *BRIP1* et *PALB2* dans la susceptibilité au cancer du sein chez les familles à risque élevé provenant de notre population.

L'acquisition de connaissances plus approfondies de la pathogenèse du cancer du sein et de la susceptibilité génétique au cancer du sein a mené à l'identification de cibles plausibles comme étant des gènes candidats du cancer du sein. Tel que présenté dans la section 3.3, les gènes de susceptibilités au cancer du sein sont interreliés et sont impliqués dans le

contrôle du cycle cellulaire, l'apoptose et la réparation de l'ADN. En réponse aux dommages à l'ADN, ATM phosphoryle une panoplie de substrats. Les protéines BRCA1, p53 et CHEK2<sup>288</sup> sont parmi les cibles d'ATM. CHEK2, une autre kinase, phosphoryle aussi BRCA1<sup>220</sup> et p53. De plus, CHEK2 phosphoryle BRCA2, régulant ainsi son interaction avec la recombinaise RAD51. Il existe une rétroaction entre plusieurs gènes de susceptibilité au cancer du sein. BRCA1 augmente la transcription par p53, alors que p53 réprime la transcription de *BRCA1*. PTEN accroît la stabilité et l'activité de liaison à l'ADN de p53 en s'y associant physiquement, alors que p53 active l'expression de PTEN<sup>209</sup>. STK11 lie aussi p53 et le phosphoryle<sup>253,255</sup>. La liaison de STK11 à PTEN affecte sa localisation cellulaire, possiblement en le retenant au cytoplasme<sup>212</sup>. De plus, trois gènes de susceptibilité au cancer du sein encodent des protéines faisant partie de la voie de l'anémie de Fanconi. BRIP1/FANCI interagit avec la protéine BRCA1<sup>97</sup>. La protéine PALB2/FANCD1, quant à elle, sert de protéine d'échafaudage entre BRCA1 et BRCA2/FANCD1 dans la formation du complexe BRCA nécessaire à la réparation par recombinaison homologue<sup>335,336</sup>.

Les gènes de susceptibilité faisant partie d'un réseau de protéines interconnectées, il est donc raisonnable de penser que des gènes encodant des protéines faisant partie de ce réseau pourraient aussi être des gènes de susceptibilité au cancer du sein. Ainsi, afin d'identifier des mutations délétères pouvant potentiellement être associées à une susceptibilité au cancer du sein, l'analyse des gènes candidats *AURKA*, *BAP1*, *BARD1* et *DHX9* a été effectuée. En dépit du fait que le choix des gènes candidats analysés soit dépendant des connaissances que nous avons de ceux-ci<sup>501</sup>, les évidences présentées précédemment nous ont poussé à étudier ces quatre gènes encodant des protéines interagissant avec BRCA1. Nous avons priorisé l'approche du re-séquençage de gènes candidats puisque cette méthode permet l'identification de gènes de susceptibilité à forte pénétrance et à pénétrance modérée pour lesquels les mutations sont rares. Sachant que les gènes étudiés font partie de voies biologique reliées et qu'ils pourraient agir d'une manière synergique, nous n'avons délibérément décidé de ne pas procéder à des analyses d'interaction gène-gène. De telles analyses sont fortement dépendantes de la fréquence des variants identifiés et nécessitent de grandes cohortes. Nos analyses se sont limitées principalement aux séquences codantes de ces gènes candidats puisque le but premier de notre étude était d'identifier des mutations

clairement délétères et des événements d'épissage alternatif. Nous sommes toutefois conscients que des variations situées dans la région promotrice de ces gènes pourraient affecter leur niveau d'expression et que de tels effets ne peuvent pas être détectés dans le cadre de notre étude. Bien que les analyses effectuées n'aient pas mené à l'identification claire d'une association de l'un de ces gènes avec une susceptibilité au cancer du sein, elles ont permis de dégager une association possible d'un variant de *BARD1* avec un risque accru de cancer du sein. L'identification d'haplotypes, la détermination de blocs d'haplotypes et l'identification de tSNPs dans ces gènes s'avèrent des informations très utiles afin de permettre leur analyse dans de plus grandes cohortes. Ainsi, ce volet de mes études doctorales a permis de dégager une association potentielle d'un variant de *BARD1* avec une susceptibilité accrue au cancer du sein et de fournir les outils nécessaires à l'étude de ces quatre gènes candidats à plus grande échelle.

### **Gènes de susceptibilité à identifier**

Jusqu'à présent, une stratification des gènes de susceptibilité a pu être effectuée en fonction du risque associé. Alors que les gènes de forte pénétrance confèrent un risque de l'ordre de plus de 10 fois, les gènes de pénétrance modérée augmentent le risque de 2 à 3 fois et les gènes de faible pénétrance d'environ 1,5 fois. Puisque plus de 70% de la prédisposition génétique au cancer du sein est inexplicée, il demeure assurément certains gènes de prédisposition à découvrir. De plus, au fur et à mesure que d'autres gènes de prédisposition au cancer du sein seront identifiés, il est probable que de nouvelles catégories de gènes de susceptibilité soient identifiées ou que les limites des différentes classes, tel que nous les connaissons présentement, deviennent de moins en moins définies.

Les gènes de susceptibilité au cancer du sein à forte pénétrance jusqu'ici identifiés sont des gènes pour lesquels les mutations sont rares et dominantes. Ils ont été identifiés par des analyses de liaison ou par des analyses de ségrégation. L'existence d'autres gènes de susceptibilité à forte pénétrance et l'identification de ceux-ci par des analyses de même type que celles qui ont permis l'identification des premiers gènes de susceptibilité au cancer du sein n'est pas à exclure. Toutefois, les gènes de forte pénétrance restant à découvrir seront responsables d'un nombre limité de cas de cancer du sein et n'expliqueront qu'une très faible partie des cas.

Il est fort probable qu'il existe d'autres gènes de susceptibilité au cancer du sein à pénétrance modérée. Il est toutefois difficile de prédire le nombre exact de gènes qui seront de ce type. Les gènes de susceptibilité à pénétrance modérée identifiés à ce jour sont impliqués dans la réparation des bris d'ADN et sont liés de près ou de loin à BRCA1 et BRCA2. Il est donc justifié d'étudier les partenaires d'interactions de BRCA1 et BRCA2 à titre de gènes candidats dans le but d'identifier de nouveaux gènes de susceptibilité au cancer du sein. L'implication de gènes faisant partie d'autres mécanismes cellulaires et n'interagissant pas avec BRCA1 ou BRCA2 n'est toutefois pas à exclure.

Les mutations associées au cancer du sein dans ce type de gène ne résultent généralement pas en de grands « pedigrees » présentant de multiples cas de cancer du sein, ce qui fait en sorte que ce type de gènes est plutôt difficile à identifier à l'aide d'études de liaison. De plus, comme ces allèles sont peu communs, il est peu probable de les identifier à l'aide d'études d'association. Le re-séquençage de gènes candidats chez des patientes atteintes et chez des individus sains provenant de la même population semble donc être la méthode la plus efficace afin d'identifier de nouveaux gènes de susceptibilité au cancer du sein à pénétrance modérée. En raison du fait que le risque associé à des mutations dans de tels gènes n'est seulement que de 2 à 4 fois, et que la fréquence de ces mutations est faible, le nombre d'individus requis dans de telles études doit être relativement élevé, ce qui complique significativement les analyses. L'utilisation d'individus provenant de familles à risque élevé de cancer du sein, en comparaison à des cas sélectionnés sur la base d'études populationnelles, pourrait permettre de pallier aux problèmes associés au re-séquençage d'un grand nombre d'individus. Des allèles de plus faible susceptibilité au cancer du sein sont généralement enrichis chez des cas familiaux de cancer du sein en comparaison avec des cas non-familiaux<sup>137</sup>.

Des analyses d'association menées sur l'ensemble du génome ont permis l'identification de variants communs à faible pénétrance conférant un risque accru de cancer du sein. Ces tours du génome ont principalement été effectués chez des individus de descendance européenne. Il est probable que dans d'autres populations, de telles analyses mènent à la découverte d'autres allèles de susceptibilité au cancer du sein à faible pénétrance ou bien que les allèles présentement identifiés soient associés à d'autres maladies ou d'autres

phénotypes. De plus, les variants ayant été identifiés comme tSNPs pourraient maintenant être utilisés afin d'estimer le risque associé à la présence de ces allèles et dans le cadre d'analyses épidémiologiques.

### **Étude de familles à risque élevé provenant de la population canadienne-française : recherche de gènes de susceptibilité**

Dans une population fondatrice dérivée d'un petit nombre d'individus, l'hétérogénéité génétique est réduite. L'utilisation d'une telle population est utile pour l'identification et la localisation de marqueurs ou de gènes liés à des traits complexes. Une plus grande proportion des individus atteints d'une maladie sera potentiellement porteuse d'une même mutation et il y aura un plus fort LD entre la mutation et les marqueurs environnants, facilitant ainsi l'identification de la région porteuse de la mutation<sup>391</sup>.

L'utilisation de cas présentant une prédisposition génétique permet d'augmenter la puissance d'une étude sans toutefois avoir à augmenter de façon significative le nombre d'individus participant à cette étude. Par le fait même, l'utilisation de cas présentant une prédisposition génétique permet d'augmenter la puissance d'une étude sans en augmenter les coûts. Parmi les facteurs permettant un enrichissement pour une prédisposition génétique, on peut nommer une histoire familiale, des cas de cancer du sein bilatéraux et un jeune âge au diagnostic. L'utilisation d'une cohorte composée de cas de cancer du sein avec un apparenté du premier degré ayant aussi été diagnostiqué avec un cancer du sein permet, pour une même puissance statistique, de diminuer considérablement la taille de la cohorte en comparaison à une cohorte composée d'individus n'ayant aucun apparenté de premier degré atteint d'un cancer du sein. Par le fait même, une cohorte composée d'individus possédant deux apparentés de premier degré ayant été atteints d'un cancer du sein permet d'obtenir la même puissance qu'une cohorte quatre fois plus grande composée d'individus atteints d'un cancer du sein, mais n'ayant aucun apparenté de premier degré atteint d'un cancer du sein. L'utilisation de cas de cancer du sein ayant un cancer du sein bilatéral permet une réduction de la taille de la cohorte du même ordre que l'utilisation d'individus ayant deux apparentés de premier degrés atteints<sup>49</sup>. La réduction de la taille de l'échantillon obtenue de par l'utilisation d'individus diagnostiqués à un âge précoce (35-40 ans) n'est pas aussi avantageuse. La conception d'une cohorte composée de cas atteints à un âge

précoce permet, pour une même puissance, une réduction de la taille de moins de 40% en comparaison avec une cohorte de cas diagnostiqués à 65 ans<sup>49</sup>. Nous avons donc étudié une cohorte d'individus provenant de familles présentant plusieurs cas de cancer du sein, nous permettant ainsi d'obtenir une meilleure puissance pour notre étude.

## CONCLUSION

### Implications cliniques

Le risque accru de cancer du sein associé aux mutations dans les gènes de pénétrance faible ou modérée n'est aucunement à remettre en doute. L'importance clinique de ces mutations chez les femmes porteuses est toutefois plus difficilement interprétable. Selon le modèle polygénique multiplicatif, le risque associé à un variant agirait d'une manière multiplicative afin de modifier le risque déjà présent chez une femme et causé par d'autres facteurs. Il en serait de même pour les gènes à faible pénétrance. Le risque conféré par un locus de ce type est faible individuellement. Par contre, en présence de plusieurs loci, ce risque semble agir de manière multiplicative. Ainsi, le risque de cancer du sein est approximativement six fois plus élevé chez les femmes porteuses de 14 allèles de risque connus, en comparaison avec les femmes non-porteuses d'un allèle de risque pour ces loci<sup>501</sup>. D'une manière semblable, une analyse des variants associés à une faible pénétrance a permis d'identifier que certains d'entre eux pourraient modifier le risque de cancer du sein chez des porteurs de mutations dans les gènes *BRCA1* ou *BRCA2*. Certains de ces loci semblent moduler le risque de cancer du sein chez les porteurs de mutations à forte pénétrance dans les gènes *BRCA1* ou *BRCA2*<sup>389</sup>.

Bien que l'utilisation clinique d'un gène de susceptibilité à faible pénétrance soit limitée, certains allèles de susceptibilité, lorsque pris ensemble, pourraient nous permettre de discerner les femmes à risque élevé de cancer du sein de celles à faible risque, particulièrement dans le cadre de programme de dépistage au niveau de la population. Stratifier les femmes en fonction de leur risque génétique pourrait améliorer l'efficacité des différents programmes de dépistage. Ainsi, l'efficacité des programmes populationnels de prévention, tel la mammographie, pourrait être améliorée en ciblant les femmes les plus à risque en fonction de leur génotype.

Un intérêt important est présentement porté à l'amélioration de l'estimation du risque de cancer. L'utilisation des découvertes et des méthodes génétiques améliore à la fois l'estimation du risque de cancer chez les individus et facilite le choix d'interventions préventives. Au niveau clinique, les principales applications reliées à la connaissance des facteurs génétiques impliqués dans la susceptibilité au cancer du sein sont l'estimation du

risque d'un individu en particulier et la prévention à plus grande échelle. La connaissance du risque de cancer du sein chez un individu permettra une prise en charge et un suivi adéquats.

### **Implications cliniques de l'étude d'une population fondatrice**

Pour une mutation fondatrice dans laquelle un grand nombre d'individus porteurs d'une même mutation est disponible, il est possible d'évaluer la pénétrance et l'expression de l'allèle muté. L'analyse de modificateurs de risque de nature environnementale ou génétique peut aussi être effectuée. Dans le cadre de tests génétiques, il est avantageux de connaître si des mutations fondatrices sont présentes dans une population. Ceci permet une meilleure estimation de la probabilité d'être porteur d'une mutation chez les individus atteints d'une maladie. Les tests génétiques peuvent cibler les mutations fondatrices, ce qui permet l'obtention de tests plus rapides et une diminution des coûts associés à ceux-ci. Certaines mutations sont difficilement détectables par les méthodes de détection usuelles. Connaître la présence d'une telle mutation fondatrice dans une population permet l'utilisation de techniques appropriées pour leur détection. De plus, dans l'éventualité où plusieurs mutations fondatrices sont présentes dans une population, elles peuvent être analysées simultanément.

### **Tests de prédiction pour tests génétiques et modèles de prédiction du risque**

Certains modèles ont été développés afin de calculer la probabilité d'un individu ou d'une famille d'être porteurs d'une mutation dans *BRCA1* ou *BRCA2*. Ces modèles sont principalement dérivés de modèles statistiques prenant en compte la pénétrance des gènes et l'histoire familiale de cancer du sein. Ces modèles peuvent être spécialisés pour un gène en particulier, tel le modèle de Couch<sup>502</sup> qui prédit le risque d'être porteur d'une mutation dans *BRCA1*, ou évaluer à la fois le risque d'être porteur d'une mutation dans *BRCA1* ou *BRCA2*, tel le modèle de Frank<sup>503</sup>, le modèle BRCAPRO<sup>504</sup> et le système de pointage de Manchester<sup>505</sup>. Ces modèles peuvent être utilisés, entre autres, par les médecins et les conseillers génétiques afin de déterminer la pertinence de référer l'individu ou la famille en question à une clinique de dépistage génétique.

Ces modèles ont été développés afin d'estimer la probabilité d'être porteur d'une mutation dans les gènes *BRCA1* ou *BRCA2*. Ils n'estiment pas le risque de cancer du sein chez un



individu. Toutefois, d'autres modèles ont été développés afin d'estimer le risque d'être atteint du cancer du sein chez les femmes ne répondant pas nécessairement aux critères afin d'être testées pour des mutations dans les gènes *BRCA1* et *BRCA2*. Deux des modèles les plus utilisés sont les modèles de Gail et de Claus. Le modèle de Gail incorpore les informations au sujet de plusieurs facteurs de risque de cancer du sein, incluant le nombre d'apparentés du premier degré atteints, l'âge à la ménopause, l'âge à la première grossesse menée à terme et le nombre de biopsies du sein déjà effectuées<sup>506</sup>. Le modèle de Claus incorpore l'histoire familiale de cancer du sein du côté maternel et paternel (degré d'apparentement et âge au diagnostic), mais ne prend pas en compte les autres facteurs de risque généraux mentionnés plus haut<sup>507</sup>. Plus récemment, un modèle permettant d'estimer le risque de cancer et la probabilité d'être porteur d'une mutation de *BRCA1* ou *BRCA2* a également été développé<sup>508</sup>. Comme aucun des modèles développés ne peut estimer le risque adéquatement et que chacun de ceux-ci comporte certaines limitations, il a été recommandé d'utiliser plusieurs modèles de prédiction de risque en combinaison<sup>509</sup>. Au fil du temps, l'acquisition d'une meilleure connaissance des gènes de susceptibilité au cancer du sein et du risque associé à chacun de ceux-ci permettra de développer des modèles statistiques calculant la probabilité d'être porteur d'une mutation dans l'un de ces gènes. Ultiment, une connaissance approfondie des facteurs génétiques impliqués dans l'étiologie du cancer du sein permettra l'incorporation de ceux-ci dans des modèles de prédiction du risque tel qu'il en a été pour les gènes *BRCA1* et *BRCA2*.

### **La médecine personnalisée**

L'utilisation de tests génétiques afin d'estimer le risque de cancer du sein mène à des situations où les résultats sont relativement difficiles à interpréter. L'identification de variants de significativité inconnue (variants of unknown significance; VUS) cause certains problèmes cliniques pour les individus désirant une estimation personnalisée de leur risque de cancer et pour les cliniciens qui doivent recommander des stratégies de réduction du risque<sup>510,511</sup>. Les méthodes bioinformatiques permettent une classification rapide de l'effet potentiel des VUS et ne requièrent pas la création de « pédigrees », la collecte d'échantillons de tumeurs ou la mise au point d'essais fonctionnels. Bien que les méthodes bioinformatiques ne puissent pas être considérées comme des alternatives aux analyses médicales génétiques, elle peuvent cependant venir en complément à celles-ci. Dans le

futur, une contribution accrue et une forte progression de la recherche dans le domaine des méthodes de prédiction et de classification du risque associé à la présence d'un VUS est souhaitable. Ces modèles prédictifs pourront incorporer l'information génétique médicale et l'information provenant d'essais fonctionnels, ce qui fera en sorte qu'ils deviendront de plus en plus efficaces. Ultimement, une classification pourra être attribuée à chaque variant, réduisant l'incertitude de la prédiction et les inquiétudes associées à la présence d'un VUS chez un patient. Ainsi, l'estimation du risque chez un individu et sa prise en charge seront améliorés.

Jusqu'à maintenant, plusieurs gènes de susceptibilité au cancer du sein ont été identifiés. Bien que le présent projet de recherche ait éliminé plusieurs de ceux-ci en tant que joueurs majeurs dans la susceptibilité au cancer du sein chez des familles à risque élevé provenant de la population canadienne-française, la contribution exacte des facteurs génétiques dans le développement du cancer du sein chez les familles provenant de cette population reste à déterminer. La présente étude a également permis d'identifier ou d'exclure des gènes de susceptibilité potentiels et a procuré les outils nécessaires à l'évaluation à grande échelle de ces gènes candidats chez d'autres populations. Au fil des années, une meilleure connaissance des facteurs génétiques impliqués dans l'étiologie de cette maladie, jumelée à l'amélioration des modèles de prédiction de risque et à l'essor de la médecine personnalisée permettront une meilleure utilisation des méthodes de réduction du risque et une personnalisation du traitement.

## BIBLIOGRAPHIE

- 1) Société canadienne du cancer et Institut national du cancer du Canada : Statistiques canadiennes sur le cancer 2008, Toronto, Canada, 2008.
- 2) Parkin DM, Bray F, Ferlay J, et al. Global cancer statistics, 2002. *CA Cancer J Clin.* 2005 Mar-Apr;55(2):74-108.
- 3) Ries LAG, Melber D, Karpcho M et al., SEER Web Site 2008.
- 4) Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000 Jan 7;100(1):57-70.
- 5) Nowell PC. The clonal evolution of tumor cell populations. *Science.* 1976 Oct 1;194(4260):23-8.
- 6) Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell.* 1990 Jun 1;61(5):759-67.
- 7) Elliott RL, Blobe GC. Role of transforming growth factor Beta in human cancer. *J Clin Oncol.* 2005 Mar 20;23(9):2078-93.
- 8) Brown K, Strathdee D, Bryson S, et al. The malignant capacity of skin tumours induced by expression of a mutant H-ras transgene depends on the cell type targeted. *Curr Biol.* 1998 Apr 23;8(9):516-24.
- 9) Nkondjock A, Ghadirian P. Facteurs de risque du cancer du sein. *Med Sci (Paris).* 2005 Feb;21(2):175-80.
- 10) Wiseman M. The second World Cancer Research Fund/American Institute for Cancer Research expert report. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. *Proc Nutr Soc.* 2008 Aug;67(3):253-6.
- 11) Kelsey JL, Bernstein L. Epidemiology and prevention of breast cancer. *Annu Rev Public Health.* 1996;17:47-67.
- 12) Boyd NF, Lockwood GA, Byng JW, et al. Mammographic densities and breast cancer risk. *Cancer Epidemiol Biomarkers Prev.* 1998 Dec;7(12):1133-44.
- 13) Boice JD Jr. Radiation and breast carcinogenesis. *Med Pediatr Oncol.* 2001 May;36(5):508-13.
- 14) World Cancer Research Fund/American Institute for Cancer Research (1997) Food, Nutrition and the Prevention of Cancer: a Global Perspective. Washington, DC: AICR.
- 15) Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. *Lancet.* 1996 Jun 22;347(9017):1713-27.
- 16) Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. *Lancet.* 2002 Jul 20;360(9328):187-95.
- 17) Cummings SR, Tice JA, Bauer S, et al. Prevention of breast cancer in postmenopausal women: approaches to estimating and reducing risk. *J Natl Cancer Inst.* 2009 Mar 18;101(6):384-98.
- 18) Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. *Lancet.* 1997 Oct 11;350(9084):1047-59.
- 19) Layde PM, Webster LA, Baughman AL, et al. The independent associations of parity, age at first full term pregnancy, and duration of breastfeeding with the risk of breast

- cancer. Cancer and Steroid Hormone Study Group. *J Clin Epidemiol*. 1989;42(10):963-73.
- 20) Hinkula M, Pukkala E, Kyyrönen P, et al. Grand multiparity and the risk of breast cancer: population-based study in Finland. *Cancer Causes Control*. 2001 Aug;12(6):491-500.
  - 21) Russo J, Hu YF, Yang X, et al. Developmental, cellular, and molecular basis of human breast cancer. *J Natl Cancer Inst Monogr*. 2000;(27):17-37.
  - 22) Key TJ, Pike MC. The role of oestrogens and progestagens in the epidemiology and prevention of breast cancer. *Eur J Cancer Clin Oncol*. 1988 Jan;24(1):29-43.
  - 23) World Cancer Research Fund/American Institute for Cancer Research (2007). Food, nutrition, physical activity and the Prevention of cancer: Global Perspective. Washington DC: AICR; <http://dietandcancerreport.org>.
  - 24) IARC. IARC Handbook of Cancer Prevention, Volume 6, Weight Control and Physical Activity. IARC Press: Lyon, 2002.
  - 25) Monninkhof EM, Elias SG, Vlems FA, et al. Physical activity and breast cancer: a systematic review. *Epidemiology*. 2007 Jan;18(1):137-57.
  - 26) Friedenreich CM, Cust AE. Physical activity and breast cancer risk: impact of timing, type and dose of activity and population subgroup effects. *Br J Sports Med*. 2008 Aug;42(8):636-47.
  - 27) McTiernan A. Mechanisms linking physical activity with cancer. *Nat Rev Cancer*. 2008 Mar;8(3):205-11.
  - 28) Yu H, Rohan T. Role of the insulin-like growth factor family in cancer development and progression. *J Natl Cancer Inst*. 2000 Sep 20;92(18):1472-89.
  - 29) McTiernan A, Ulrich C, Slate S, et al. Physical activity and cancer etiology: associations and mechanisms. *Cancer Causes Control*. 1998 Oct;9(5):487-509.
  - 30) Key TJ, Verkasalo PK, Banks E. Epidemiology of breast cancer. *Lancet Oncol*. 2001 Mar;2(3):133-40.
  - 31) Collaborative Group on Hormonal Factors in Breast Cancer. Alcohol, tobacco and breast cancer--collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. *Br J Cancer*. 2002 Nov 18;87(11):1234-45.
  - 32) Feigelson HS, Calle EE, Robertson AS, et al. Alcohol consumption increases the risk of fatal breast cancer (United States). *Cancer Causes Control*. 2001 Dec;12(10):895-902.
  - 33) Wu AH, Yu MC, Tseng CC, et al. Green tea and risk of breast cancer in Asian Americans. *Int J Cancer*. 2003 Sep 10;106(4):574-9.
  - 34) Kavanagh KT, Hafer LJ, Kim DW, et al. Green tea extracts decrease carcinogen-induced mammary tumor burden in rats and rate of breast cancer cell proliferation in culture. *J Cell Biochem*. 2001;82(3):387-98.
  - 35) Pool-Zobel BL, Bub A, Müller H, et al. Consumption of vegetables reduces genetic damage in humans: first results of a human intervention trial with carotenoid-rich foods. *Carcinogenesis*. 1997 Sep;18(9):1847-50.
  - 36) Giovannucci E. Tomatoes, tomato-based products, lycopene, and cancer: review of the epidemiologic literature. *J Natl Cancer Inst*. 1999 Feb 17;91(4):317-31.
  - 37) Newman B, Austin MA, Lee M, et al. Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. *Proc Natl Acad Sci U S A*. 1988 May;85(9):3044-8.

- 38) Claus EB, Risch N, Thompson WD. Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am J Hum Genet.* 1991 Feb;48(2):232-42.
- 39) Hall JM, Lee MK, Newman B, et al. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science.* 1990 Dec 21;250(4988):1684-9.
- 40) Peto J, Mack TM. High constant incidence in twins and other relatives of women with breast cancer. *Nat Genet.* 2000 Dec;26(4):411-4.
- 41) Risch N. The genetic epidemiology of cancer: interpreting family and twin studies and their implications for molecular genetic approaches. *Cancer Epidemiol Biomarkers Prev.* 2001 Jul;10(7):733-41.
- 42) Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med.* 2000 Jul 13;343(2):78-85.
- 43) Peto J. Cancer epidemiology in the last century and the next decade. *Nature.* 2001 May 17;411(6835):390-5.
- 44) Antoniou AC, Pharoah PD, McMullan G, et al. A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *Br J Cancer.* 2002 Jan 7;86(1):76-83.
- 45) Fackenthal JD, Olopade OI. Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. *Nat Rev Cancer.* 2007 Dec;7(12):937-48.
- 46) Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science.* 1994 Oct 7;266(5182):66-71.
- 47) Wooster R, Bignell G, Lancaster J, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature.* 1995 Dec 21-28;378(6559):789-92.
- 48) Tavtigian SV, Simard J, Rommens J, et al. The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. *Nat Genet.* 1996 Mar;12(3):333-7.
- 49) Antoniou AC, Easton DF. Polygenic inheritance of breast cancer: Implications for design of association studies. *Genet Epidemiol.* 2003 Nov;25(3):190-202.
- 50) Ford D, Easton DF, Bishop DT, et al. Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. *Lancet.* 1994 Mar 19;343(8899):692-5.
- 51) Struwing JP, Hartge P, Wacholder S, et al. The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. *N Engl J Med.* 1997 May 15;336(20):1401-8.
- 52) Antoniou A, Pharoah PD, Narod S, 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 May;72(5):1117-30.
- 53) Ford D, Easton DF, Stratton M, et al. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet.* 1998 Mar;62(3):676-89.
- 54) The Breast Cancer Linkage Consortium. Cancer risks in BRCA2 mutation carriers. *J Natl Cancer Inst.* 1999 Aug 4;91(15):1310-6.
- 55) Thompson D, Easton DF; Breast Cancer Linkage Consortium. Cancer Incidence in BRCA1 mutation carriers. *J Natl Cancer Inst.* 2002 Sep 18;94(18):1358-65.
- 56) van Asperen CJ, Brohet RM, Meijers-Heijboer EJ, et al. Cancer risks in BRCA2 families: estimates for sites other than breast and ovary. *J Med Genet.* 2005 Sep;42(9):711-9.

- 57) Anglian Breast Cancer Study Group. Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. *Br J Cancer*. 2000 Nov;83(10):1301-8.
- 58) Antoniou AC, Gayther SA, Stratton JF, et al. Risk models for familial ovarian and breast cancer. *Genet Epidemiol*. 2000 Feb;18(2):173-90.
- 59) Easton DF. How many more breast cancer predisposition genes are there? *Breast Cancer Res*. 1999;1(1):14-7.
- 60) Peto J, Collins N, Barfoot R, et al. Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer. *J Natl Cancer Inst*. 1999 Jun 2;91(11):943-9.
- 61) Wu LC, Wang ZW, Tsan JT, et al. Identification of a RING protein that can interact in vivo with the BRCA1 gene product. *Nat Genet*. 1996 Dec;14(4):430-40.
- 62) Koonin EV, Altschul SF, Bork P. BRCA1 protein products ... Functional motifs... *Nat Genet*. 1996 Jul;13(3):266-8.
- 63) Callebaut I, Mornon JP. From BRCA1 to RAP1: a widespread BRCT module closely associated with DNA repair. *FEBS Lett*. 1997 Jan 2;400(1):25-30.
- 64) Wang Y, Cortez D, Yazdi P, et al. BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes Dev*. 2000 Apr 15;14(8):927-39.
- 65) Khanna KK, Jackson SP. DNA double-strand breaks: signaling, repair and the cancer connection. *Nat Genet*. 2001 Mar;27(3):247-54.
- 66) Rich T, Allen RL, Wyllie AH. Defying death after DNA damage. *Nature*. 2000 Oct 12;407(6805):777-83.
- 67) Jin Y, Xu XL, Yang MC, Wei F, Ayi TC, Bowcock AM, Baer R. Cell cycle-dependent colocalization of BARD1 and BRCA1 proteins in discrete nuclear domains. *Proc Natl Acad Sci U S A*. 1997 Oct 28;94(22):12075-80.
- 68) Scully R, Chen J, Ochs RL, et al. Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. *Cell*. 1997 Aug 8;90(3):425-35.
- 69) Paull TT, Rogakou EP, Yamazaki V, et al. A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Curr Biol*. 2000 Jul 27-Aug 10;10(15):886-95.
- 70) Scully R, Chen J, Plug A, et al. Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell*. 1997 Jan 24;88(2):265-75.
- 71) Zhong Q, Chen CF, Li S, et al. Association of BRCA1 with the hRad50-hMre11-p95 complex and the DNA damage response. *Science*. 1999 Jul 30;285(5428):747-50.
- 72) Paull TT, Cortez D, Bowers B, et al. Direct DNA binding by Brca1. *Proc Natl Acad Sci U S A*. 2001 May 22;98(11):6086-91.
- 73) Deng CX. Tumorigenesis as a consequence of genetic instability in Brca1 mutant mice. *Mutat Res*. 2001 Jun 2;477(1-2):183-9.
- 74) Venkitaraman AR. Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell*. 2002 Jan 25;108(2):171-82.
- 75) Cortez D, Wang Y, Qin J, et al. Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks. *Science*. 1999 Nov 5;286(5442):1162-6.

- 76) Tibbetts RS, Cortez D, Brumbaugh KM, et al. Functional interactions between BRCA1 and the checkpoint kinase ATR during genotoxic stress. *Genes Dev.* 2000 Dec 1;14(23):2989-3002.
- 77) Yu VP, Koehler M, Steinlein C, et al. Gross chromosomal rearrangements and genetic exchange between nonhomologous chromosomes following BRCA2 inactivation. *Genes Dev.* 2000 Jun 1;14(11):1400-6.
- 78) Zhang J, Powell SN. The role of the BRCA1 tumor suppressor in DNA double-strand break repair. *Mol Cancer Res.* 2005 Oct;3(10):531-9.
- 79) Bau DT, Mau YC, Shen CY. The role of BRCA1 in non-homologous end-joining. *Cancer Lett.* 2006 Aug 18;240(1):1-8.
- 80) Wang HC, Chou WC, Shieh SY, et al. Ataxia telangiectasia mutated and checkpoint kinase 2 regulate BRCA1 to promote the fidelity of DNA end-joining. *Cancer Res.* 2006 Feb 1;66(3):1391-400.
- 81) Zhuang J, Zhang J, Willers H, et al. Checkpoint kinase 2-mediated phosphorylation of BRCA1 regulates the fidelity of nonhomologous end-joining. *Cancer Res.* 2006 Feb 1;66(3):1401-8.
- 82) Wei L, Lan L, Hong Z, et al. Rapid Recruitment of BRCA1 to DNA Double-Strand Breaks Is Dependent on Its Association with Ku80. *Mol Cell Biol.* 2008 Dec;28(24):7380-93.
- 83) Wood RD. Nucleotide excision repair in mammalian cells. *J Biol Chem.* 1997 Sep 19;272(38):23465-8.
- 84) Sancar A. DNA excision repair. *Annu Rev Biochem.* 1996;65:43-81.
- 85) Petit C, Sancar A. Nucleotide excision repair: from E. coli to man. *Biochimie.* 1999 Jan-Feb;81(1-2):15-25.
- 86) Hanawalt PC. Subpathways of nucleotide excision repair and their regulation. *Oncogene.* 2002 Dec 16;21(58):8949-56.
- 87) Le Page F, Randrianarison V, Marot D, et al. BRCA1 and BRCA2 are necessary for the transcription-coupled repair of the oxidative 8-oxoguanine lesion in human cells. *Cancer Res.* 2000 Oct 1;60(19):5548-52.
- 88) Anderson SF, Schlegel BP, Nakajima T, et al. BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A. *Nat Genet.* 1998 Jul;19(3):254-6.
- 89) Hartman AR, Ford JM. BRCA1 induces DNA damage recognition factors and enhances nucleotide excision repair. *Nat Genet.* 2002 Sep;32(1):180-4.
- 90) Welch PL, Owens KN, King MC. Insights into the functions of BRCA1 and BRCA2. *Trends Genet.* 2000 Feb;16(2):69-74.
- 91) Ouchi T, Monteiro AN, August A, et al. BRCA1 regulates p53-dependent gene expression. *Proc Natl Acad Sci U S A.* 1998 Mar 3;95(5):2302-6.
- 92) Ouchi T, Lee SW, Ouchi M, et al. Collaboration of signal transducer and activator of transcription 1 (STAT1) and BRCA1 in differential regulation of IFN-gamma target genes. *Proc Natl Acad Sci U S A.* 2000 May 9;97(10):5208-13.
- 93) Wang Q, Zhang H, Kajino K, et al. BRCA1 binds c-Myc and inhibits its transcriptional and transforming activity in cells. *Oncogene.* 1998 Oct 15;17(15):1939-48.
- 94) Irminger-Finger I, Siegel BD, Leung WC. The functions of breast cancer susceptibility gene 1 (BRCA1) product and its associated proteins. *Biol Chem.* 1999 Feb;380(2):117-28.

- 95) Yarden RI, Brody LC. BRCA1 interacts with components of the histone deacetylase complex. *Proc Natl Acad Sci U S A*. 1999 Apr 27;96(9):4983-8.
- 96) Bochar DA, Wang L, Beniya H, et al. BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer. *Cell*. 2000 Jul 21;102(2):257-65.
- 97) Cantor SB, Bell DW, Ganesan S, et al. BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell*. 2001 Apr 6;105(1):149-60.
- 98) Deng CX. Roles of BRCA1 in centrosome duplication. *Oncogene*. 2002 Sep 9;21(40):6222-7.
- 99) Deng CX. BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. *Nucleic Acids Res*. 2006 Mar 6;34(5):1416-26.
- 100) Ruffner H, Verma IM. BRCA1 is a cell cycle-regulated nuclear phosphoprotein. *Proc Natl Acad Sci U S A*. 1997 Jul 8;94(14):7138-43.
- 101) Okada S, Ouchi T. Cell cycle differences in DNA damage-induced BRCA1 phosphorylation affect its subcellular localization. *J Biol Chem*. 2003 Jan 17;278(3):2015-20.
- 102) Turner JM, Aprelikova O, Xu X, et al. BRCA1, histone H2AX phosphorylation, and male meiotic sex chromosome inactivation. *Curr Biol*. 2004 Dec 14;14(23):2135-42.
- 103) Fabbro M, Savage K, Hobson K, et al. BRCA1-BARD1 complexes are required for p53Ser-15 phosphorylation and a G1/S arrest following ionizing radiation-induced DNA damage. *J Biol Chem*. 2004 Jul 23;279(30):31251-8.
- 104) Xu B, Kim St, Kastan MB. Involvement of Brcal in S-phase and G(2)-phase checkpoints after ionizing irradiation. *Mol Cell Biol*. 2001 May;21(10):3445-50.
- 105) Xu B, O'Donnell AH, Kim ST, et al. Phosphorylation of serine 1387 in Brcal is specifically required for the Atm-mediated S-phase checkpoint after ionizing irradiation. *Cancer Res*. 2002 Aug 15;62(16):4588-91.
- 106) Stewart GS, Wang B, Bignell CR, et al. MDC1 is a mediator of the mammalian DNA damage checkpoint. *Nature*. 2003 Feb 27;421(6926):961-6.
- 107) Meraldi P, Honda R, Nigg EA. Aurora kinases link chromosome segregation and cell division to cancer susceptibility. *Curr Opin Genet Dev*. 2004 Feb;14(1):29-36.
- 108) Katayama H, Brinkley WR, Sen S. The Aurora kinases: role in cell transformation and tumorigenesis. *Cancer Metastasis Rev*. 2003 Dec;22(4):451-64.
- 109) Ducat D, Zheng Y. Aurora kinases in spindle assembly and chromosome segregation. *Exp Cell Res*. 2004 Nov 15;301(1):60-7.
- 110) Crane R, Gadea B, Littlepage L, et al. Aurora A, meiosis and mitosis. *Biol Cell*. 2004 Apr;96(3):215-29.
- 111) Ouchi M, Fujiuchi N, Sasai K, et al. BRCA1 phosphorylation by Aurora-A in the regulation of G2 to M transition. *J Biol Chem*. 2004 May 7;279(19):19643-8.
- 112) Wang RH, Yu H, Deng CX. A requirement for breast-cancer-associated gene 1 (BRCA1) in the spindle checkpoint. *Proc Natl Acad Sci U S A*. 2004 Dec 7;101(49):17108-13.
- 113) Wooster R, Neuhausen SL, Mangion J, et al. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science*. 1994 Sep 30;265(5181):2088-90.
- 114) Bignell G, Micklem G, Stratton MR, et al. The BRC repeats are conserved in mammalian BRCA2 proteins. *Hum Mol Genet*. 1997 Jan;6(1):53-8.



- 115) Moynahan ME, Pierce AJ, Jasin M. BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol Cell*. 2001 Feb;7(2):263-72.
- 116) Tirkkonen M, Johannsson O, Agnarsson BA, et al. Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. *Cancer Res*. 1997 Apr 1;57(7):1222-7.
- 117) Patel KJ, Yu VP, Lee H, et al. Involvement of Brca2 in DNA repair. *Mol Cell*. 1998 Feb;1(3):347-57.
- 118) Kraakman-van der Zwet M, Overkamp WJ, van Lange RE, et al. Brca2 (XRCC11) deficiency results in radioresistant DNA synthesis and a higher frequency of spontaneous deletions. *Mol Cell Biol*. 2002 Jan;22(2):669-79.
- 119) Connor F, Bertwistle D, Mee PJ, et al. Tumorigenesis and a DNA repair defect in mice with a truncating Brca2 mutation. *Nat Genet*. 1997 Dec;17(4):423-30.
- 120) Sharan SK, Morimatsu M, Albrecht U, et al. Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking Brca2. *Nature*. 1997 Apr 24;386(6627):804-10.
- 121) Morimatsu M, Donoho G, Hasty P. Cells deleted for Brca2 COOH terminus exhibit hypersensitivity to gamma-radiation and premature senescence. *Cancer Res*. 1998 Aug 1;58(15):3441-7.
- 122) Mizuta R, LaSalle JM, Cheng HL, et al. RAB22 and RAB163/mouse BRCA2: proteins that specifically interact with the RAD51 protein. *Proc Natl Acad Sci U S A*. 1997 Jun 24;94(13):6927-32.
- 123) Wong AK, Pero R, Ormonde PA, et al. RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene brca2. *J Biol Chem*. 1997 Dec 19;272(51):31941-4.
- 124) Baumann P, Benson FE, West SC. Human Rad51 protein promotes ATP-dependent homologous pairing and strand transfer reactions in vitro. *Cell*. 1996 Nov 15;87(4):757-66.
- 125) Tarsounas M, Davies AA, West SC. RAD51 localization and activation following DNA damage. *Philos Trans R Soc Lond B Biol Sci*. 2004 Jan 29;359(1441):87-93.
- 126) Yuan SS, Lee SY, Chen G, et al. BRCA2 is required for ionizing radiation-induced assembly of Rad51 complex in vivo. *Cancer Res*. 1999 Aug 1;59(15):3547-51.
- 127) Esashi F, Galkin VE, Yu X, et al. Stabilization of RAD51 nucleoprotein filaments by the C-terminal region of BRCA2. *Nat Struct Mol Biol*. 2007 Jun;14(6):468-74.
- 128) Bahassi EM, Ovesen JL, Riesenber AL, et al. The checkpoint kinases Chk1 and Chk2 regulate the functional associations between hBRCA2 and Rad51 in response to DNA damage. *Oncogene*. 2008 Jun 26;27(28):3977-85.
- 129) Mérel P, Prieur A, Pfeiffer P, et al. Absence of major defects in non-homologous DNA end joining in human breast cancer cell lines. *Oncogene*. 2002 Aug 15;21(36):5654-9.
- 130) Evers B, Jonkers J. Mouse models of BRCA1 and BRCA2 deficiency: past lessons, current understanding and future prospects. *Oncogene*. 2006 Sep 25;25(43):5885-97.
- 131) Hughes-Davies L, Huntsman D, Ruas M, et al. EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer. *Cell*. 2003 Nov 26;115(5):523-35.
- 132) Shin S, Verma IM. BRCA2 cooperates with histone acetyltransferases in androgen receptor-mediated transcription. *Proc Natl Acad Sci U S A*. 2003 Jun 10;100(12):7201-6.

- 133) Howlett NG, Taniguchi T, Olson S, et al. Biallelic inactivation of BRCA2 in Fanconi anemia. *Science*. 2002 Jul 26;297(5581):606-9.
- 134) Xia B, Sheng Q, Nakanishi K, et al. Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Mol Cell*. 2006 Jun 23;22(6):719-29.
- 135) Reid S, Schindler D, Hanenberg H, et al. Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nat Genet*. 2007 Feb;39(2):162-4.
- 136) Xia B, Dorsman JC, Ameziane N, et al. Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. *Nat Genet*. 2007 Feb;39(2):159-61.
- 137) Erkkö H, Xia B, Nikkilä J, et al. A recurrent mutation in PALB2 in Finnish cancer families. *Nature*. 2007 Mar 15;446(7133):316-9.
- 138) Rahman N, Seal S, Thompson D, et al. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat Genet*. 2007 Feb;39(2):165-7.
- 139) Laloo F, Varley J, Ellis D, et al. Prediction of pathogenic mutations in patients with early-onset breast cancer by family history. *Lancet*. 2003 Mar 29;361(9363):1101-2.
- 140) Malkin D. p53 and the Li-Fraumeni syndrome. *Cancer Genet Cytogenet*. 1993 Apr;66(2):83-92.
- 141) Li FP, Fraumeni JF Jr. Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome? *Ann Intern Med*. 1969 Oct;71(4):747-52.
- 142) Malkin D, Li FP, Strong LC, et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science*. 1990 Nov 30;250(4985):1233-8.
- 143) Lustbader ED, Williams WR, Bondy ML, et al. Segregation analysis of cancer in families of childhood soft-tissue-sarcoma patients. *Am J Hum Genet*. 1992 Aug;51(2):344-56.
- 144) Evans DG, Birch JM, Thorneycroft M, et al. Low rate of TP53 germline mutations in breast cancer/sarcoma families not fulfilling classical criteria for Li-Fraumeni syndrome. *J Med Genet*. 2002 Dec;39(12):941-4.
- 145) Laloo F, Varley J, Moran A, et al. BRCA1, BRCA2 and TP53 mutations in very early-onset breast cancer with associated risks to relatives. *Eur J Cancer*. 2006 May;42(8):1143-50.
- 146) Vousden KH, Lane DP. p53 in health and disease. *Nat Rev Mol Cell Biol*. 2007 Apr;8(4):275-83.
- 147) Levine AJ. p53, the cellular gatekeeper for growth and division. *Cell*. 1997 Feb 7;88(3):323-31.
- 148) Riley T, Sontag E, Chen P, et al. Transcriptional control of human p53-regulated genes. *Nat Rev Mol Cell Biol*. 2008 May;9(5):402-12.
- 149) Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med*. 2004 Aug;10(8):789-99.
- 150) Prives C, Hall PA. The p53 pathway. *J Pathol*. 1999 Jan;187(1):112-26.
- 151) Dumont P, Leu JI, Della Pietra AC 3rd, et al. The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. *Nat Genet*. 2003 Mar;33(3):357-65.
- 152) Mihara M, Erster S, Zaika A, et al. p53 has a direct apoptogenic role at the mitochondria. *Mol Cell*. 2003 Mar;11(3):577-90.
- 153) Erster S, Mihara M, Kim RH, et al. In vivo mitochondrial p53 translocation triggers a rapid first wave of cell death in response to DNA damage that can precede p53 target gene activation. *Mol Cell Biol*. 2004 Aug;24(15):6728-41.

- 154) Chipuk JE, Bouchier-Hayes L, Kuwana T, et al. PUMA couples the nuclear and cytoplasmic proapoptotic function of p53. *Science*. 2005 Sep 9;309(5741):1732-5.
- 155) Stambolic V, MacPherson D, Sas D, et al. Regulation of PTEN transcription by p53. *Mol Cell*. 2001 Aug;8(2):317-25.
- 156) Chen X, Ko LJ, Jayaraman L, et al. p53 levels, functional domains, and DNA damage determine the extent of the apoptotic response of tumor cells. *Genes Dev*. 1996 Oct 1;10(19):2438-51.
- 157) Zhao R, Gish K, Murphy M, et al. Analysis of p53-regulated gene expression patterns using oligonucleotide arrays. *Genes Dev*. 2000 Apr 15;14(8):981-93.
- 158) Kruse JP, Gu W. SnapShot: p53 posttranslational modifications. *Cell*. 2008 May 30;133(5):930-30.e1.
- 159) Tang Y, Zhao W, Chen Y, et al. Acetylation is indispensable for p53 activation. *Cell*. 2008 May 16;133(4):612-26.
- 160) Zhou BB, Elledge SJ. The DNA damage response: putting checkpoints in perspective. *Nature*. 2000 Nov 23;408(6811):433-9.
- 161) Renwick A, Thompson D, Seal S, et al. ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nat Genet*. 2006 Aug;38(8):873-5.
- 162) Bond GL, Hu W, Levine AJ. MDM2 is a central node in the p53 pathway: 12 years and counting. *Curr Cancer Drug Targets*. 2005 Feb;5(1):3-8.
- 163) Li DM, Sun H. TEP1, encoded by a candidate tumor suppressor locus, is a novel protein tyrosine phosphatase regulated by transforming growth factor beta. *Cancer Res*. 1997 Jun 1;57(11):2124-9.
- 164) Li J, Yen C, Liaw D, et al. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science*. 1997 Mar 28;275(5308):1943-7.
- 165) Nelen MR, Padberg GW, Peeters EA, et al. Localization of the gene for Cowden disease to chromosome 10q22-23. *Nat Genet*. 1996 May;13(1):114-6.
- 166) Liaw D, Marsh DJ, Li J, et al. Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. *Nat Genet*. 1997 May;16(1):64-7.
- 167) Eng C. Will the real Cowden syndrome please stand up: revised diagnostic criteria. *J Med Genet*. 2000 Nov;37(11):828-30.
- 168) Eng C. PTEN: one gene, many syndromes. *Hum Mutat*. 2003 Sep;22(3):183-98.
- 169) Eng C, Peacocke M. PTEN and inherited hamartoma-cancer syndromes. *Nat Genet*. 1998 Jul;19(3):223.
- 170) Nelen MR, van Staveren WC, Peeters EA, et al. Germline mutations in the PTEN/MMAC1 gene in patients with Cowden disease. *Hum Mol Genet*. 1997 Aug;6(8):1383-7.
- 171) Nelen MR, Kremer H, Konings IB, et al. Novel PTEN mutations in patients with Cowden disease: absence of clear genotype-phenotype correlations. *Eur J Hum Genet*. 1999 Apr;7(3):267-73.
- 172) Marsh DJ, Coulon V, Lunetta KL, et al. Mutation spectrum and genotype-phenotype analyses in Cowden disease and Bannayan-Zonana syndrome, two hamartoma syndromes with germline PTEN mutation. *Hum Mol Genet*. 1998 Mar;7(3):507-15.

- 173) Marsh DJ, Kum JB, Lunetta KL, et al. PTEN mutation spectrum and genotype-phenotype correlations in Bannayan-Riley-Ruvalcaba syndrome suggest a single entity with Cowden syndrome. *Hum Mol Genet.* 1999 Aug;8(8):1461-72.
- 174) Zhou XP, Waite KA, Pilarski R, et al. Germline PTEN promoter mutations and deletions in Cowden/Bannayan-Riley-Ruvalcaba syndrome result in aberrant PTEN protein and dysregulation of the phosphoinositol-3-kinase/Akt pathway. *Am J Hum Genet.* 2003 Aug;73(2):404-11.
- 175) Butler MG, Dasouki MJ, Zhou XP, et al. Subset of individuals with autism spectrum disorders and extreme macrocephaly associated with germline PTEN tumour suppressor gene mutations. *J Med Genet.* 2005 Apr;42(4):318-21.
- 176) Marsh DJ, Dahia PL, Coulon V, et al. Allelic imbalance, including deletion of PTEN/MMAC1, at the Cowden disease locus on 10q22-23, in hamartomas from patients with Cowden syndrome and germline PTEN mutation. *Genes Chromosomes Cancer.* 1998 Jan;21(1):61-9.
- 177) Ball S, Arolker M, Purushotham AD. Breast cancer, Cowden disease and PTEN-MATCHS syndrome. *Eur J Surg Oncol.* 2001 Sep;27(6):604-6.
- 178) Nassif NT, Lobo GP, Wu X, et al. PTEN mutations are common in sporadic microsatellite stable colorectal cancer. *Oncogene.* 2004 Jan 15;23(2):617-28.
- 179) Rhei E, Kang L, Bogomolny F, et al. Mutation analysis of the putative tumor suppressor gene PTEN/MMAC1 in primary breast carcinomas. *Cancer Res.* 1997 Sep 1;57(17):3657-9.
- 180) Tang Y, Eng C. PTEN autoregulates its expression by stabilization of p53 in a phosphatase-independent manner. *Cancer Res.* 2006 Jan 15;66(2):736-42.
- 181) Zhou XP, Marsh DJ, Morrison CD, et al. Germline inactivation of PTEN and dysregulation of the phosphoinositol-3-kinase/Akt pathway cause human Lhermitte-Duclos disease in adults. *Am J Hum Genet.* 2003 Nov;73(5):1191-8.
- 182) Agrawal S, Eng C. Differential expression of novel naturally occurring splice variants of PTEN and their functional consequences in Cowden syndrome and sporadic breast cancer. *Hum Mol Genet.* 2006 Mar 1;15(5):777-87.
- 183) Chen ST, Yu SY, Tsai M, et al. Mutation analysis of the putative tumor suppression gene PTEN/MMAC1 in sporadic breast cancer. *Breast Cancer Res Treat.* 1999 May;55(1):85-9.
- 184) Zhou XP, Loukola A, Salovaara R, et al. PTEN mutational spectra, expression levels, and subcellular localization in microsatellite stable and unstable colorectal cancers. *Am J Pathol.* 2002 Aug;161(2):439-47.
- 185) Cairns P, Evron E, Okami K, et al. Point mutation and homozygous deletion of PTEN/MMAC1 in primary bladder cancers. *Oncogene.* 1998 Jun 18;16(24):3215-8.
- 186) Mutter GL, Lin MC, Fitzgerald JT, et al. Altered PTEN expression as a diagnostic marker for the earliest endometrial precancers. *J Natl Cancer Inst.* 2000 Jun 7;92(11):924-30.
- 187) Yin Y, Shen WH. PTEN: a new guardian of the genome. *Oncogene.* 2008 Sep 18;27(41):5443-53.
- 188) Saal LH, Gruvberger-Saal SK, Persson C, et al. Recurrent gross mutations of the PTEN tumor suppressor gene in breast cancers with deficient DSB repair. *Nat Genet.* 2008 Jan;40(1):102-7.

- 189) Myers MP, Stolarov JP, Eng C, et al. P-TEN, the tumor suppressor from human chromosome 10q23, is a dual-specificity phosphatase. *Proc Natl Acad Sci U S A*. 1997 Aug 19;94(17):9052-7.
- 190) Maehama T, Dixon JE. The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. *J Biol Chem*. 1998 May 29;273(22):13375-8.
- 191) Furnari FB, Lin H, Huang HS, et al. Growth suppression of glioma cells by PTEN requires a functional phosphatase catalytic domain. *Proc Natl Acad Sci U S A*. 1997 Nov 11;94(23):12479-84.
- 192) Li DM, Sun H. PTEN/MMAC1/TEP1 suppresses the tumorigenicity and induces G1 cell cycle arrest in human glioblastoma cells. *Proc Natl Acad Sci U S A*. 1998 Dec 22;95(26):15406-11.
- 193) Stambolic V, Suzuki A, de la Pompa JL, et al. Negative regulation of PKB/Akt-dependent cell survival by the tumor suppressor PTEN. *Cell*. 1998 Oct 2;95(1):29-39.
- 194) Sun H, Lesche R, Li DM, et al. PTEN modulates cell cycle progression and cell survival by regulating phosphatidylinositol 3,4,5,-trisphosphate and Akt/protein kinase B signaling pathway. *Proc Natl Acad Sci U S A*. 1999 May 25;96(11):6199-204.
- 195) Sulis ML, Parsons R. PTEN: from pathology to biology. *Trends Cell Biol*. 2003 Sep;13(9):478-83.
- 196) Leslie NR, Downes CP. PTEN function: how normal cells control it and tumour cells lose it. *Biochem J*. 2004 Aug 15;382(Pt 1):1-11.
- 197) Sansal I, Sellers WR. The biology and clinical relevance of the PTEN tumor suppressor pathway. *J Clin Oncol*. 2004 Jul 15;22(14):2954-63.
- 198) Tamura M, Gu J, Matsumoto K, et al. Inhibition of cell migration, spreading, and focal adhesions by tumor suppressor PTEN. *Science*. 1998 Jun 5;280(5369):1614-7.
- 199) Weng L, Brown J, Eng C. PTEN induces apoptosis and cell cycle arrest through phosphoinositol-3-kinase/Akt-dependent and -independent pathways. *Hum Mol Genet*. 2001 Feb 1;10(3):237-42.
- 200) Raftopoulou M, Etienne-Manneville S, Self A, et al. Regulation of cell migration by the C2 domain of the tumor suppressor PTEN. *Science*. 2004 Feb 20;303(5661):1179-81.
- 201) Shen WH, Balajee AS, Wang J, et al. Essential role for nuclear PTEN in maintaining chromosomal integrity. *Cell*. 2007 Jan 12;128(1):157-70.
- 202) Gildea JJ, Herlevsen M, Harding MA, et al. PTEN can inhibit in vitro organotypic and in vivo orthotopic invasion of human bladder cancer cells even in the absence of its lipid phosphatase activity. *Oncogene*. 2004 Sep 2;23(40):6788-97.
- 203) Puc J, Keniry M, Li HS, et al. Lack of PTEN sequesters CHK1 and initiates genetic instability. *Cancer Cell*. 2005 Feb;7(2):193-204.
- 204) Trotman LC, Wang X, Alimonti A, et al. Ubiquitination regulates PTEN nuclear import and tumor suppression. *Cell*. 2007 Jan 12;128(1):141-56.
- 205) Wang X, Trotman LC, Koppie T, et al. NEDD4-1 is a proto-oncogenic ubiquitin ligase for PTEN. *Cell*. 2007 Jan 12;128(1):129-39.
- 206) Hong TM, Yang PC, Peck K, et al. Profiling the downstream genes of tumor suppressor PTEN in lung cancer cells by complementary DNA microarray. *Am J Respir Cell Mol Biol*. 2000 Sep;23(3):355-63.

- 207) Matsushima-Nishiu M, Unoki M, Ono K, et al. Growth and gene expression profile analyses of endometrial cancer cells expressing exogenous PTEN. *Cancer Res.* 2001 May 1;61(9):3741-9.
- 208) Chang CJ, Freeman DJ, Wu H. PTEN regulates Mdm2 expression through the P1 promoter. *J Biol Chem.* 2004 Jul 9;279(28):29841-8.
- 209) Freeman DJ, Li AG, Wei G, et al. PTEN tumor suppressor regulates p53 protein levels and activity through phosphatase-dependent and -independent mechanisms. *Cancer Cell.* 2003 Feb;3(2):117-30.
- 210) Gu J, Tamura M, Yamada KM. Tumor suppressor PTEN inhibits integrin- and growth factor-mediated mitogen-activated protein (MAP) kinase signaling pathways. *J Cell Biol.* 1998 Nov 30;143(5):1375-83.
- 211) Li J, Simpson L, Takahashi M, et al. The PTEN/MMAC1 tumor suppressor induces cell death that is rescued by the AKT/protein kinase B oncogene. *Cancer Res.* 1998 Dec 15;58(24):5667-72.
- 212) Mehenni H, Lin-Marq N, Buchet-Poyau K, et al. LKB1 interacts with and phosphorylates PTEN: a functional link between two proteins involved in cancer predisposing syndromes. *Hum Mol Genet.* 2005 Aug 1;14(15):2209-19.
- 213) Cully M, You H, Levine AJ, et al. Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis. *Nat Rev Cancer.* 2006 Mar;6(3):184-92.
- 214) Gimm O, Perren A, Weng LP, et al. Differential nuclear and cytoplasmic expression of PTEN in normal thyroid tissue, and benign and malignant epithelial thyroid tumors. *Am J Pathol.* 2000 May;156(5):1693-700.
- 215) Whiteman DC, Zhou XP, Cummings MC, et al. Nuclear PTEN expression and clinicopathologic features in a population-based series of primary cutaneous melanoma. *Int J Cancer.* 2002 May 1;99(1):63-7.
- 216) Ginn-Pease ME, Eng C. Increased nuclear phosphatase and tensin homologue deleted on chromosome 10 is associated with G0-G1 in MCF-7 cells. *Cancer Res.* 2003 Jan 15;63(2):282-6.
- 217) Liu F, Wagner S, Campbell RB, et al. PTEN enters the nucleus by diffusion. *J Cell Biochem.* 2005 Oct 1;96(2):221-34.
- 218) Chung JH, Ginn-Pease ME, Eng C. Phosphatase and tensin homologue deleted on chromosome 10 (PTEN) has nuclear localization signal-like sequences for nuclear import mediated by major vault protein. *Cancer Res.* 2005 May 15;65(10):4108-16.
- 219) Planchon SM, Waite KA, Eng C. The nuclear affairs of PTEN. *J Cell Sci.* 2008 Feb 1;121(Pt 3):249-53.
- 220) Lee JS, Collins KM, Brown AL, et al. hCds1-mediated phosphorylation of BRCA1 regulates the DNA damage response. *Nature.* 2000 Mar 9;404(6774):201-4.
- 221) Ongusaha PP, Ouchi T, Kim KT, et al. BRCA1 shifts p53-mediated cellular outcomes towards irreversible growth arrest. *Oncogene.* 2003 Jun 12;22(24):3749-58.
- 222) Backman SA, Stambolic V, Suzuki A, et al. Deletion of Pten in mouse brain causes seizures, ataxia and defects in soma size resembling Lhermitte-Duclos disease. *Nat Genet.* 2001 Dec;29(4):396-403.
- 223) Mayo LD, Donner DB. The PTEN, Mdm2, p53 tumor suppressor-oncoprotein network. *Trends Biochem Sci.* 2002 Sep;27(9):462-7.
- 224) Hemminki A, Markie D, Tomlinson I, et al. A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. *Nature.* 1998 Jan 8;391(6663):184-7.

- 225) Jenne DE, Reimann H, Nezu J, et al. Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. *Nat Genet.* 1998 Jan;18(1):38-43.
- 226) Wojciechowska M, Dryjański T. A 21-year-old female patient with Peutz-Jeghers syndrome. *Pol Arch Med Wewn.* 2008 Oct;118(10):585-9.
- 227) Hearle N, Schumacher V, Menko FH, et al. Frequency and spectrum of cancers in the Peutz-Jeghers syndrome. *Clin Cancer Res.* 2006 May 15;12(10):3209-15.
- 228) Ji H, Ramsey MR, Hayes DN, et al. LKB1 modulates lung cancer differentiation and metastasis. *Nature.* 2007 Aug 16;448(7155):807-10.
- 229) Giardiello FM, Welsh SB, Hamilton SR, et al. Increased risk of cancer in the Peutz-Jeghers syndrome. *N Engl J Med.* 1987 Jun 11;316(24):1511-4.
- 230) Lim W, Hearle N, Shah B, et al. Further observations on LKB1/STK11 status and cancer risk in Peutz-Jeghers syndrome. *Br J Cancer.* 2003 Jul 21;89(2):308-13.
- 231) Hemminki A, Tomlinson I, Markie D, et al. Localization of a susceptibility locus for Peutz-Jeghers syndrome to 19p using comparative genomic hybridization and targeted linkage analysis. *Nat Genet.* 1997 Jan;15(1):87-90.
- 232) Sapkota GP, Boudeau J, Deak M, et al. Identification and characterization of four novel phosphorylation sites (Ser31, Ser325, Thr336 and Thr366) on LKB1/STK11, the protein kinase mutated in Peutz-Jeghers cancer syndrome. *Biochem J.* 2002 Mar 1;362(Pt 2):481-90.
- 233) Mehenni H, Gehrig C, Nezu J, et al. Loss of LKB1 kinase activity in Peutz-Jeghers syndrome, and evidence for allelic and locus heterogeneity. *Am J Hum Genet.* 1998 Dec;63(6):1641-50.
- 234) Hernan I, Roig I, Martin B, et al. De novo germline mutation in the serine-threonine kinase STK11/LKB1 gene associated with Peutz-Jeghers syndrome. *Clin Genet.* 2004 Jul;66(1):58-62.
- 235) Dong SM, Kim KM, Kim SY, et al. Frequent somatic mutations in serine/threonine kinase 11/Peutz-Jeghers syndrome gene in left-sided colon cancer. *Cancer Res.* 1998 Sep 1;58(17):3787-90.
- 236) Amos CI, Keitheri-Cheteri MB, Sabripour M, et al. Genotype-phenotype correlations in Peutz-Jeghers syndrome. *J Med Genet.* 2004 May;41(5):327-33.
- 237) Sapkota GP, Kieloch A, Lizcano JM, et al. Phosphorylation of the protein kinase mutated in Peutz-Jeghers cancer syndrome, LKB1/STK11, at Ser431 by p90(RSK) and cAMP-dependent protein kinase, but not its farnesylation at Cys(433), is essential for LKB1 to suppress cell growth. *J Biol Chem.* 2001 Jun 1;276(22):19469-82.
- 238) Forcet C, Etienne-Manneville S, Gaudé H, et al. Functional analysis of Peutz-Jeghers mutations reveals that the LKB1 C-terminal region exerts a crucial role in regulating both the AMPK pathway and the cell polarity. *Hum Mol Genet.* 2005 May 15;14(10):1283-92.
- 239) Aretz S, Stienen D, Uhlhaas S, et al. High proportion of large genomic STK11 deletions in Peutz-Jeghers syndrome. *Hum Mutat.* 2005 Dec;26(6):513-9.
- 240) Hearle NC, Rudd MF, Lim W, et al. Exonic STK11 deletions are not a rare cause of Peutz-Jeghers syndrome. *J Med Genet.* 2006 Apr;43(4):e15.
- 241) Volikos E, Robinson J, Aittomäki K, et al. LKB1 exonic and whole gene deletions are a common cause of Peutz-Jeghers syndrome. *J Med Genet.* 2006 May;43(5):e18.
- 242) Sanchez-Cespedes M, Parrella P, Esteller M, et al. Inactivation of LKB1/STK11 is a common event in adenocarcinomas of the lung. *Cancer Res.* 2002 Jul 1;62(13):3659-62.

- 243) Esteller M, Avizienyte E, Corn PG, et al. Epigenetic inactivation of LKB1 in primary tumors associated with the Peutz-Jeghers syndrome. *Oncogene*. 2000 Jan 6;19(1):164-8.
- 244) Trojan J, Brieger A, Raedle J, et al. 5'-CpG island methylation of the LKB1/STK11 promoter and allelic loss at chromosome 19p13.3 in sporadic colorectal cancer. *Gut*. 2000 Aug;47(2):272-6.
- 245) Baas AF, Kuipers J, van der Wel NN, et al. Complete polarization of single intestinal epithelial cells upon activation of LKB1 by STRAD. *Cell*. 2004 Feb 6;116(3):457-66.
- 246) Marignani PA, Kanai F, Carpenter CL. LKB1 associates with Brg1 and is necessary for Brg1-induced growth arrest. *J Biol Chem*. 2001 Aug 31;276(35):32415-8.
- 247) Alessi DR, Sakamoto K, Bayascas JR. LKB1-dependent signaling pathways. *Annu Rev Biochem*. 2006;75:137-63.
- 248) Katajisto P, Vallenius T, Vaahtomeri K, et al. The LKB1 tumor suppressor kinase in human disease. *Biochim Biophys Acta*. 2007 Jan;1775(1):63-75.
- 249) Barnes AP, Lilley BN, Pan YA, et al. LKB1 and SAD kinases define a pathway required for the polarization of cortical neurons. *Cell*. 2007 May 4;129(3):549-63.
- 250) Shelly M, Cancedda L, Heilshorn S, et al. LKB1/STRAD promotes axon initiation during neuronal polarization. *Cell*. 2007 May 4;129(3):565-77.
- 251) Bardeesy N, Sinha M, Hezel AF, et al. Loss of the Lkb1 tumour suppressor provokes intestinal polyposis but resistance to transformation. *Nature*. 2002 Sep 12;419(6903):162-7.
- 252) Adachi-Yamada T, Fujimura-Kamada K, Nishida Y, et al. Distortion of proximodistal information causes JNK-dependent apoptosis in *Drosophila* wing. *Nature*. 1999 Jul 8;400(6740):166-9.
- 253) Karuman P, Gozani O, Odze RD, et al. The Peutz-Jegher gene product LKB1 is a mediator of p53-dependent cell death. *Mol Cell*. 2001 Jun;7(6):1307-19.
- 254) Lee JH, Koh H, Kim M, et al. JNK pathway mediates apoptotic cell death induced by tumor suppressor LKB1 in *Drosophila*. *Cell Death Differ*. 2006 Jul;13(7):1110-22.
- 255) Boudeau J, Baas AF, Deak M, et al. MO25alpha/beta interact with STRADalpha/beta enhancing their ability to bind, activate and localize LKB1 in the cytoplasm. *EMBO J*. 2003 Oct 1;22(19):5102-14.
- 256) Mukherjee P, Mulrooney TJ, Marsh J, et al. Differential effects of energy stress on AMPK phosphorylation and apoptosis in experimental brain tumor and normal brain. *Mol Cancer*. 2008 May 12;7:37.
- 257) Shaw RJ, Kosmatka M, Bardeesy N, et al. The tumor suppressor LKB1 kinase directly activates AMP-activated kinase and regulates apoptosis in response to energy stress. *Proc Natl Acad Sci U S A*. 2004 Mar 9;101(10):3329-35.
- 258) Rowan A, Churchman M, Jefferey R, et al. In situ analysis of LKB1/STK11 mRNA expression in human normal tissues and tumours. *J Pathol*. 2000 Oct;192(2):203-6.
- 259) Jishage K, Nezu J, Kawase Y, et al. Role of Lkb1, the causative gene of Peutz-Jegher's syndrome, in embryogenesis and polyposis. *Proc Natl Acad Sci U S A*. 2002 Jun 25;99(13):8903-8.
- 260) Yoo LI, Chung DC, Yuan J. LKB1--a master tumour suppressor of the small intestine and beyond. *Nat Rev Cancer*. 2002 Jul;2(7):529-35.
- 261) Conde E, Suarez-Gauthier A, García-García E, Lopez-Rios F, Lopez-Encuentra A, García-Lujan R, Morente M, Sanchez-Verde L, Sanchez-Cespedes M. Specific



- pattern of LKB1 and phospho-acetyl-CoA carboxylase protein immunostaining in human normal tissues and lung carcinomas. *Hum Pathol.* 2007 Sep;38(9):1351-60.
- 262) Baas AF, Boudeau J, Sapkota GP, et al. Activation of the tumour suppressor kinase LKB1 by the STE20-like pseudokinase STRAD. *EMBO J.* 2003 Jun 16;22(12):3062-72.
- 263) Boudeau J, Scott JW, Resta N, et al. Analysis of the LKB1-STRAD-MO25 complex. *J Cell Sci.* 2004 Dec 15;117(Pt 26):6365-75.
- 264) Milburn CC, Boudeau J, Deak M, et al. Crystal structure of MO25 alpha in complex with the C terminus of the pseudo kinase STE20-related adaptor. *Nat Struct Mol Biol.* 2004 Feb;11(2):193-200.
- 265) Savitsky K, Bar-Shira A, Gilad S, et al. A single ataxia telangiectasia gene with a product similar to PI-3 kinase. *Science.* 1995 Jun 23;268(5218):1749-53.
- 266) Swift M, Morrell D, Cromartie E, et al. The incidence and gene frequency of ataxia-telangiectasia in the United States. *Am J Hum Genet.* 1986 Nov;39(5):573-83.
- 267) Chun HH, Gatti RA. Ataxia-telangiectasia, an evolving phenotype. *DNA Repair (Amst).* 2004 Aug-Sep;3(8-9):1187-96.
- 268) Khanna KK. Cancer risk and the ATM gene: a continuing debate. *J Natl Cancer Inst.* 2000 May 17;92(10):795-802.
- 269) Swift M, Morrell D, Massey RB, et al. Incidence of cancer in 161 families affected by ataxia-telangiectasia. *N Engl J Med.* 1991 Dec 26;325(26):1831-6.
- 270) Thompson D, Duedal S, Kirner J, et al. Cancer risks and mortality in heterozygous ATM mutation carriers. *J Natl Cancer Inst.* 2005 Jun 1;97(11):813-22.
- 271) Pippard EC, Hall AJ, Barker DJ, et al. Cancer in homozygotes and heterozygotes of ataxia-telangiectasia and xeroderma pigmentosum in Britain. *Cancer Res.* 1988 May 15;48(10):2929-32.
- 272) Easton DF. Cancer risks in A-T heterozygotes. *Int J Radiat Biol.* 1994 Dec;66(6 Suppl):S177-82.
- 273) Janin N, Andrieu N, Ossian K, et al. Breast cancer risk in ataxia telangiectasia (AT) heterozygotes: haplotype study in French AT families. *Br J Cancer.* 1999 Jun;80(7):1042-5.
- 274) Olsen JH, Hahnemann JM, Børresen-Dale AL, et al. Cancer in patients with ataxia-telangiectasia and in their relatives in the nordic countries. *J Natl Cancer Inst.* 2001 Jan 17;93(2):121-7.
- 275) Chenevix-Trench G, Spurdle AB, Gatei M, et al. Dominant negative ATM mutations in breast cancer families. *J Natl Cancer Inst.* 2002 Feb 6;94(3):205-15.
- 276) Vorechovský I, Luo L, Lindblom A, et al. ATM mutations in cancer families. *Cancer Res.* 1996 Sep 15;56(18):4130-3.
- 277) Bay JO, Grancho M, Pernin D, et al. No evidence for constitutional ATM mutation in breast/gastric cancer families. *Int J Oncol.* 1998 Jun;12(6):1385-90.
- 278) Thorstenson YR, Roxas A, Kroiss R, et al. Contributions of ATM mutations to familial breast and ovarian cancer. *Cancer Res.* 2003 Jun 15;63(12):3325-33.
- 279) Allinen M, Launonen V, Laake K, et al. ATM mutations in Finnish breast cancer patients. *J Med Genet.* 2002 Mar;39(3):192-6.
- 280) Uziel T, Savitsky K, Platzer M, et al. Genomic Organization of the ATM gene. *Genomics.* 1996 Apr 15;33(2):317-20.
- 281) Abraham RT. PI 3-kinase related kinases: 'big' players in stress-induced signaling pathways. *DNA Repair (Amst).* 2004 Aug-Sep;3(8-9):883-7.

- 282) Bosotti R, Isacchi A, Sonnhammer EL. FAT: a novel domain in PIK-related kinases. *Trends Biochem Sci.* 2000 May;25(5):225-7.
- 283) Jiang X, Sun Y, Chen S, et al. The FATC domains of PIKK proteins are functionally equivalent and participate in the Tip60-dependent activation of DNA-PKcs and ATM. *J Biol Chem.* 2006 Jun 9;281(23):15741-6.
- 284) Perry J, Kleckner N. The ATRs, ATMs, and TORs are giant HEAT repeat proteins. *Cell.* 2003 Jan 24;112(2):151-5.
- 285) Fernandes N, Sun Y, Chen S, et al. DNA damage-induced association of ATM with its target proteins requires a protein interaction domain in the N terminus of ATM. *J Biol Chem.* 2005 Apr 15;280(15):15158-64.
- 286) Kurz EU, Lees-Miller SP. DNA damage-induced activation of ATM and ATM-dependent signaling pathways. *DNA Repair (Amst).* 2004 Aug-Sep;3(8-9):889-900.
- 287) Bakkenist CJ, Kastan MB. DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature.* 2003 Jan 30;421(6922):499-506.
- 288) Shiloh Y. The ATM-mediated DNA-damage response: taking shape. *Trends Biochem Sci.* 2006 Jul;31(7):402-10.
- 289) Ahn J, Urist M, Prives C. The Chk2 protein kinase. *DNA Repair (Amst).* 2004 Aug-Sep;3(8-9):1039-47.
- 290) Bartek J, Falck J, Lukas J. CHK2 kinase--a busy messenger. *Nat Rev Mol Cell Biol.* 2001 Dec;2(12):877-86.
- 291) Bell DW, Varley JM, Szydlo TE, et al. Heterozygous germ line hCHK2 mutations in Li-Fraumeni syndrome. *Science.* 1999 Dec 24;286(5449):2528-31.
- 292) Meijers-Heijboer H, van den Ouweland A, Klijn J, et al. Low-penetrance susceptibility to breast cancer due to CHEK2(\*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet.* 2002 May;31(1):55-9.
- 293) CHEK2 Breast Cancer Case-Control Consortium. CHEK2\*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. *Am J Hum Genet.* 2004 Jun;74(6):1175-82.
- 294) Weischer M, Bojesen SE, Ellervik C, et al. CHEK2\*1100delC genotyping for clinical assessment of breast cancer risk: meta-analyses of 26,000 patient cases and 27,000 controls. *J Clin Oncol.* 2008 Feb 1;26(4):542-8.
- 295) Neuhausen S, Dunning A, Steele L, et al. Role of CHEK2\*1100delC in unselected series of non-BRCA1/2 male breast cancers. *Int J Cancer.* 2004 Jan 20;108(3):477-8.
- 296) Sodha N, Wilson C, Bullock SL, et al. Analysis of familial male breast cancer for germline mutations in CHEK2. *Cancer Lett.* 2004 Nov 25;215(2):187-9.
- 297) Syrjäkoski K, Kuukasjärvi T, Auvinen A, et al. CHEK2 1100delC is not a risk factor for male breast cancer population. *Int J Cancer.* 2004 Jan 20;108(3):475-6.
- 298) Cybulski C, Huzarski T, Górski B, et al. A novel founder CHEK2 mutation is associated with increased prostate cancer risk. *Cancer Res.* 2004 Apr 15;64(8):2677-9.
- 299) Kilpivaara O, Vahteristo P, Falck J, et al. CHEK2 variant I157T may be associated with increased breast cancer risk. *Int J Cancer.* 2004 Sep 10;111(4):543-7.
- 300) Bogdanova N, Enssen-Dubrowinskaja N, Feshchenko S, et al. Association of two mutations in the CHEK2 gene with breast cancer. *Int J Cancer.* 2005 Aug 20;116(2):263-6.
- 301) Dong X, Wang L, Taniguchi K, et al. Mutations in CHEK2 associated with prostate cancer risk. *Am J Hum Genet.* 2003 Feb;72(2):270-80.

- 302) Cybulski C, Górski B, Huzarski T, et al. CHEK2 is a multiorgan cancer susceptibility gene. *Am J Hum Genet.* 2004 Dec;75(6):1131-5.
- 303) Dufault MR, Betz B, Wappenschmidt B, et al. Limited relevance of the CHEK2 gene in hereditary breast cancer. *Int J Cancer.* 2004 Jun 20;110(3):320-5.
- 304) Seppälä EH, Ikonen T, Mononen N, et al. CHEK2 variants associate with hereditary prostate cancer. *Br J Cancer.* 2003 Nov 17;89(10):1966-70.
- 305) Thompson D, Seal S, Schutte M, et al. A multicenter study of cancer incidence in CHEK2 1100delC mutation carriers. *Cancer Epidemiol Biomarkers Prev.* 2006 Dec;15(12):2542-5.
- 306) Kilpivaara O, Alhopuro P, Vahteristo P, et al. CHEK2 I157T associates with familial and sporadic colorectal cancer. *J Med Genet.* 2006 Jul;43(7):e34.
- 307) Kastan MB, Bartek J. Cell-cycle checkpoints and cancer. *Nature.* 2004 Nov 18;432(7015):316-23.
- 308) Li J, Williams BL, Haire LF, et al. Structural and functional versatility of the FHA domain in DNA-damage signaling by the tumor suppressor kinase Chk2. *Mol Cell.* 2002 May;9(5):1045-54.
- 309) Falck J, Lukas C, Protopopova M, et al. Functional impact of concomitant versus alternative defects in the Chk2-p53 tumour suppressor pathway. *Oncogene.* 2001 Sep 6;20(39):5503-10.
- 310) Staalesen V, Falck J, Geisler S, et al. Alternative splicing and mutation status of CHEK2 in stage III breast cancer. *Oncogene.* 2004 Nov 4;23(52):8535-44.
- 311) Bartek J, Lukas J. Chk1 and Chk2 kinases in checkpoint control and cancer. *Cancer Cell.* 2003 May;3(5):421-9.
- 312) Lukas C, Falck J, Bartkova J, et al. Distinct spatiotemporal dynamics of mammalian checkpoint regulators induced by DNA damage. *Nat Cell Biol.* 2003 Mar;5(3):255-60.
- 313) Nevanlinna H, Bartek J. The CHEK2 gene and inherited breast cancer susceptibility. *Oncogene.* 2006 Sep 25;25(43):5912-9.
- 314) Lukas C, Bartkova J, Latella L, et al. DNA damage-activated kinase Chk2 is independent of proliferation or differentiation yet correlates with tissue biology. *Cancer Res.* 2001 Jul 1;61(13):4990-3.
- 315) Latella L, Lukas J, Simone C, et al. Differentiation-induced radioresistance in muscle cells. *Mol Cell Biol.* 2004 Jul;24(14):6350-61.
- 316) Yamashita T, Nakahata T. Current knowledge on the pathophysiology of Fanconi anemia: from genes to phenotypes. *Int J Hematol.* 2001 Jul;74(1):33-41.
- 317) Tischkowitz M, Dokal I. Fanconi anaemia and leukaemia - clinical and molecular aspects. *Br J Haematol.* 2004 Jul;126(2):176-91.
- 318) Joenje H, Patel KJ. The emerging genetic and molecular basis of Fanconi anaemia. *Nat Rev Genet.* 2001 Jun;2(6):446-57.
- 319) Meetei AR, Levitus M, Xue Y, et al. X-linked inheritance of Fanconi anemia complementation group B. *Nat Genet.* 2004 Nov;36(11):1219-24.
- 320) D'Andrea AD, Grompe M. The Fanconi anaemia/BRCA pathway. *Nat Rev Cancer.* 2003 Jan;3(1):23-34.
- 321) Faivre L, Guardiola P, Lewis C, et al. Association of complementation group and mutation type with clinical outcome in fanconi anemia. European Fanconi Anemia Research Group. *Blood.* 2000 Dec 15;96(13):4064-70.

- 322) Adachi D, Oda T, Yagasaki H, et al. Heterogeneous activation of the Fanconi anemia pathway by patient-derived FANCA mutants. *Hum Mol Genet.* 2002 Dec 1;11(25):3125-34.
- 323) Levitus M, Waisfisz Q, Godthelp BC, et al. The DNA helicase BRIP1 is defective in Fanconi anemia complementation group J. *Nat Genet.* 2005 Sep;37(9):934-5.
- 324) Meetei AR, Medhurst AL, Ling C, et al. A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nat Genet.* 2005 Sep;37(9):958-63.
- 325) Rosenberg PS, Alter BP, Ebell W. Cancer risks in Fanconi anemia: findings from the German Fanconi Anemia Registry. *Haematologica.* 2008 Apr;93(4):511-7.
- 326) Seal S, Thompson D, Renwick A, et al. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nat Genet.* 2006 Nov;38(11):1239-41.
- 327) Xie Y, de Winter JP, Waisfisz Q, et al. Aberrant Fanconi anaemia protein profiles in acute myeloid leukaemia cells. *Br J Haematol.* 2000 Dec;111(4):1057-64.
- 328) Barroso E, Milne RL, Fernández LP, et al. FANCD2 associated with sporadic breast cancer risk. *Carcinogenesis.* 2006 Sep;27(9):1930-7.
- 329) Taniguchi T, Tischkowitz M, Ameziane N, et al. Disruption of the Fanconi anemia-BRCA pathway in cisplatin-sensitive ovarian tumors. *Nat Med.* 2003 May;9(5):568-74.
- 330) van der Heijden MS, Yeo CJ, Hruban RH, et al. Fanconi anemia gene mutations in young-onset pancreatic cancer. *Cancer Res.* 2003 May 15;63(10):2585-8.
- 331) Garcia-Higuera I, Taniguchi T, Ganesan S, et al. Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. *Mol Cell.* 2001 Feb;7(2):249-62.
- 332) Folias A, Matkovic M, Bruun D, et al. BRCA1 interacts directly with the Fanconi anemia protein FANCA. *Hum Mol Genet.* 2002 Oct 1;11(21):2591-7.
- 333) Hussain S, Witt E, Huber PA, et al. Direct interaction of the Fanconi anaemia protein FANCG with BRCA2/FANCD1. *Hum Mol Genet.* 2003 Oct 1;12(19):2503-10.
- 334) Hussain S, Wilson JB, Medhurst AL, et al. Direct interaction of FANCD2 with BRCA2 in DNA damage response pathways. *Hum Mol Genet.* 2004 Jun 15;13(12):1241-8.
- 335) Sy SM, Huen MS, Chen J. PALB2 is an integral component of the BRCA complex required for homologous recombination repair. *Proc Natl Acad Sci U S A.* 2009 Apr 28;106(17):7155-60.
- 336) Zhang F, Ma J, Wu J, et al. PALB2 links BRCA1 and BRCA2 in the DNA-damage response. *Curr Biol.* 2009 Mar 24;19(6):524-9.
- 337) Jacquemont C, Taniguchi T. The Fanconi anemia pathway and ubiquitin. *BMC Biochem.* 2007 Nov 22;8 Suppl 1:S10.
- 338) Sims AE, Spiteri E, Sims RJ 3rd, et al. FANCI is a second monoubiquitinated member of the Fanconi anemia pathway. *Nat Struct Mol Biol.* 2007 Jun;14(6):564-7.
- 339) Smogorzewska A, Matsuoka S, Vinciguerra P, et al. Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. *Cell.* 2007 Apr 20;129(2):289-301.
- 340) Wang W. Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins. *Nat Rev Genet.* 2007 Oct;8(10):735-48.

- 341) Wang X, Andreassen PR, D'Andrea AD. Functional interaction of monoubiquitinated FANCD2 and BRCA2/FANCD1 in chromatin. *Mol Cell Biol.* 2004 Jul;24(13):5850-62.
- 342) Bogliolo M, Lyakhovich A, Callén E, et al. Histone H2AX and Fanconi anemia FANCD2 function in the same pathway to maintain chromosome stability. *EMBO J.* 2007 Mar 7;26(5):1340-51.
- 343) Godthelp BC, van Buul PP, Jaspers NG, et al. Cellular characterization of cells from the Fanconi anemia complementation group, FA-D1/BRCA2. *Mutat Res.* 2006 Oct 10;601(1-2):191-201.
- 344) Yamashita T, Kupfer GM, Naf D, et al. The fanconi anemia pathway requires FAA phosphorylation and FAA/FAC nuclear accumulation. *Proc Natl Acad Sci U S A.* 1998 Oct 27;95(22):13085-90.
- 345) Qiao F, Mi J, Wilson JB, et al. Phosphorylation of fanconi anemia (FA) complementation group G protein, FANCG, at serine 7 is important for function of the FA pathway. *J Biol Chem.* 2004 Oct 29;279(44):46035-45.
- 346) Ho GP, Margossian S, Taniguchi T, et al. Phosphorylation of FANCD2 on two novel sites is required for mitomycin C resistance. *Mol Cell Biol.* 2006 Sep;26(18):7005-15.
- 347) Matsuoka S, Ballif BA, Smogorzewska A, et al. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science.* 2007 May 25;316(5828):1160-6.
- 348) Cobb JA, Bjergbaek L, Shimada K, et al. DNA polymerase stabilization at stalled replication forks requires Mec1 and the RecQ helicase Sgs1. *EMBO J.* 2003 Aug 15;22(16):4325-36.
- 349) Peng M, Litman R, Xie J, et al. The FANCI/MutLalpha interaction is required for correction of the cross-link response in FA-J cells. *EMBO J.* 2007 Jul 11;26(13):3238-49.
- 350) Taniguchi T, Garcia-Higuera I, Xu B, et al. Convergence of the fanconi anemia and ataxia telangiectasia signaling pathways. *Cell.* 2002 May 17;109(4):459-72.
- 351) Cumming RC, Lightfoot J, Beard K, et al. Fanconi anemia group C protein prevents apoptosis in hematopoietic cells through redox regulation of GSTP1. *Nat Med.* 2001 Jul;7(7):814-20.
- 352) Fagerlie S, Lensch MW, Pang Q, et al. The Fanconi anemia group C gene product: signaling functions in hematopoietic cells. *Exp Hematol.* 2001 Dec;29(12):1371-81.
- 353) Cantor S, Drapkin R, Zhang F, et al. The BRCA1-associated protein BACH1 is a DNA helicase targeted by clinically relevant inactivating mutations. *Proc Natl Acad Sci U S A.* 2004 Feb 24;101(8):2357-62.
- 354) Sigurdson AJ, Hauptmann M, Chatterjee N, et al. Kin-cohort estimates for familial breast cancer risk in relation to variants in DNA base excision repair, BRCA1 interacting and growth factor genes. *BMC Cancer.* 2004 Mar 12;4:9.
- 355) Luo L, Lei H, Du Q, et al. No mutations in the BACH1 gene in BRCA1 and BRCA2 negative breast-cancer families linked to 17q22. *Int J Cancer.* 2002 Apr 1;98(4):638-9.
- 356) Karppinen SM, Vuosku J, Heikkinen K, et al. No evidence of involvement of germline BACH1 mutations in Finnish breast and ovarian cancer families. *Eur J Cancer.* 2003 Feb;39(3):366-71.

- 357) Rutter JL, Smith AM, Dávila MR, et al. Mutational analysis of the BRCA1-interacting genes ZNF350/ZBRK1 and BRIP1/BACH1 among BRCA1 and BRCA2-negative probands from breast-ovarian cancer families and among early-onset breast cancer cases and reference individuals. *Hum Mutat.* 2003 Aug;22(2):121-8.
- 358) Lewis AG, Flanagan J, Marsh A, et al. Mutation analysis of FANCD2, BRIP1/BACH1, LMO4 and SFN in familial breast cancer. *Breast Cancer Res.* 2005;7(6):R1005-16.
- 359) Vahteristo P, Yliannala K, Tamminen A, Eerola H, Blomqvist C, Nevanlinna H. BACH1 Ser919Pro variant and breast cancer risk. *BMC Cancer.* 2006 Jan 24;6:19.
- 360) Peng M, Litman R, Jin Z, et al. BACH1 is a DNA repair protein supporting BRCA1 damage response. *Oncogene.* 2006 Apr 6;25(15):2245-53.
- 361) Gupta R, Sharma S, Sommers JA, et al. Analysis of the DNA substrate specificity of the human BACH1 helicase associated with breast cancer. *J Biol Chem.* 2005 Jul 8;280(27):25450-60.
- 362) Greenberg RA, Sobhian B, Pathania S, et al. Multifactorial contributions to an acute DNA damage response by BRCA1/BARD1-containing complexes. *Genes Dev.* 2006 Jan 1;20(1):34-46.
- 363) Kumaraswamy E, Shiekhattar R. Activation of BRCA1/BRCA2-associated helicase BACH1 is required for timely progression through S phase. *Mol Cell Biol.* 2007 Oct;27(19):6733-41.
- 364) Wu Y, Shin-ya K, Brosh RM Jr. FANCD1 helicase defective in Fanconi anemia and breast cancer unwinds G-quadruplex DNA to defend genomic stability. *Mol Cell Biol.* 2008 Jun;28(12):4116-28.
- 365) Gupta R, Sharma S, Sommers JA, et al. FANCD1 (BACH1) helicase forms DNA damage inducible foci with replication protein A and interacts physically and functionally with the single-stranded DNA-binding protein. *Blood.* 2007 Oct 1;110(7):2390-8.
- 366) Sommers JA, Rawtani N, Gupta R, et al. FANCD1 uses its motor ATPase to destabilize protein-DNA complexes, unwind triplexes, and inhibit RAD51 strand exchange. *J Biol Chem.* 2009 Mar 20;284(12):7505-17.
- 367) Bridge WL, Vandenberg CJ, Franklin RJ, et al. The BRIP1 helicase functions independently of BRCA1 in the Fanconi anemia pathway for DNA crosslink repair. *Nat Genet.* 2005 Sep;37(9):953-7.
- 368) Litman R, Peng M, Jin Z, et al. BACH1 is critical for homologous recombination and appears to be the Fanconi anemia gene product FANCD1. *Cancer Cell.* 2005 Sep;8(3):255-65.
- 369) Kruisselbrink E, Guryev V, Brouwer K, et al. Mutagenic capacity of endogenous G4 DNA underlies genome instability in FANCD1-defective *C. elegans*. *Curr Biol.* 2008 Jun 24;18(12):900-5.
- 370) Reid S, Renwick A, Seal S, et al. Biallelic BRCA2 mutations are associated with multiple malignancies in childhood including familial Wilms tumour. *J Med Genet.* 2005 Feb;42(2):147-51.
- 371) Taniguchi T, D'Andrea AD. Molecular pathogenesis of Fanconi anemia: recent progress. *Blood.* 2006 Jun 1;107(11):4223-33.
- 372) Erkkö H, Dowty JG, Nikkilä J, et al. Penetrance analysis of the PALB2 c.1592delT founder mutation. *Clin Cancer Res.* 2008 Jul 15;14(14):4667-71.

- 373) Foulkes WD, Ghadirian P, Akbari MR, et al. Identification of a novel truncating PALB2 mutation and analysis of its contribution to early-onset breast cancer in French-Canadian women. *Breast Cancer Res.* 2007;9(6):R83.
- 374) Potapova A, Hoffman AM, Godwin AK, et al. Promoter hypermethylation of the PALB2 susceptibility gene in inherited and sporadic breast and ovarian cancer. *Cancer Res.* 2008 Feb 15;68(4):998-1002.
- 375) Tischkowitz M, Sabbaghian N, Ray AM, et al. Analysis of the gene coding for the BRCA2-interacting protein PALB2 in hereditary prostate cancer. *Prostate.* 2008 May 1;68(6):675-8.
- 376) Jones S, Hruban RH, Kamiyama M, et al. Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. *Science.* 2009 Apr 10;324(5924):217.
- 377) Sy SM, Huen MS, Zhu Y, et al. PALB2 Regulates Recombinational Repair through Chromatin Association and Oligomerization. *J Biol Chem.* 2009 Jul 3;284(27):18302-10.
- 378) Yang H, Jeffrey PD, Miller J, et al. BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. *Science.* 2002 Sep 13;297(5588):1837-48.
- 379) Wilson JH, Elledge SJ. Cancer. BRCA2 enters the fray. *Science.* 2002 Sep 13;297(5588):1822-3.
- 380) Chompret A, Brugières L, Ronsin M, Gardes M, Dessarps-Freichay F, Abel A, Hua D, Ligot L, Dondon MG, Bressac-de Paillerets B, Frébourg T, Lemerle J, Bonaïti-Pellié C, Feunteun J. P53 germline mutations in childhood cancers and cancer risk for carrier individuals. *Br J Cancer.* 2000 Jun;82(12):1932-7.
- 381) Stratton MR, Rahman N. The emerging landscape of breast cancer susceptibility. *Nat Genet.* 2008 Jan;40(1):17-22.
- 382) Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature.* 2007 Jun 28;447(7148):1087-93.
- 383) Hunter DJ, Kraft P, Jacobs KB, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet.* 2007 Jul;39(7):870-4.
- 384) Stacey SN, Manolescu A, Sulem P, et al. Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet.* 2008 Jun;40(6):703-6.
- 385) Cox A, Dunning AM, Garcia-Closas M, et al. A common coding variant in CASP8 is associated with breast cancer risk. *Nat Genet.* 2007 Mar;39(3):352-8.
- 386) Mack TM, Hamilton AS, Press MF, et al. Heritable breast cancer in twins. *Br J Cancer.* 2002 Jul 29;87(3):294-300.
- 387) Smith P, McGuffog L, Easton DF, et al. A genome wide linkage search for breast cancer susceptibility genes. *Genes Chromosomes Cancer.* 2006 Jul;45(7):646-55.
- 388) Pharoah PD, Antoniou A, Bobrow M, et al. Polygenic susceptibility to breast cancer and implications for prevention. *Nat Genet.* 2002 May;31(1):33-6.
- 389) Antoniou AC, Spurdle AB, Sinilnikova OM, et al. Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers. *Am J Hum Genet.* 2008 Apr;82(4):937-48.
- 390) Narod SA. Modifiers of risk of hereditary breast cancer. *Oncogene.* 2006 Sep 25;25(43):5832-6.

- 391) De Braekeleer M, Dao TN. Hereditary disorders in the French Canadian population of Quebec. I. In search of founders. *Hum Biol.* 1994 Apr;66(2):205-23.
- 392) Neuhausen SL. Founder populations and their uses for breast cancer genetics. *Breast Cancer Res.* 2000;2(2):77-81.
- 393) Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. *Science.* 2002 Jun 21;296(5576):2225-9.
- 394) Goldgar DE, Easton DF, Byrnes GB, et al. Genetic evidence and integration of various data sources for classifying uncertain variants into a single model. *Hum Mutat.* 2008 Nov;29(11):1265-72.
- 395) Tavtigian SV, Greenblatt MS, Goldgar DE, et al. Assessing pathogenicity: overview of results from the IARC Unclassified Genetic Variants Working Group. *Hum Mutat.* 2008 Nov;29(11):1261-4.
- 396) Spurdle AB, Couch FJ, Hogervorst FB, et al. Prediction and assessment of splicing alterations: implications for clinical testing. *Hum Mutat.* 2008 Nov;29(11):1304-13.
- 397) Walker DR, Bond JP, Tarone RE, et al. Evolutionary conservation and somatic mutation hotspot maps of p53: correlation with p53 protein structural and functional features. *Oncogene.* 1999 Jan 7;18(1):211-8.
- 398) Farley AR, Link AJ. Identification and quantification of protein posttranslational modifications. *Methods Enzymol.* 2009;463:725-63.
- 399) Teraoka SN, Telatar M, Becker-Catania S, et al. Splicing defects in the ataxia-telangiectasia gene, ATM: underlying mutations and consequences. *Am J Hum Genet* 1999;64:1617-31.
- 400) Graveley BR. Alternative splicing: increasing diversity in the proteomic world. *Trends Genet.* 2001 Feb;17(2):100-7.
- 401) Pagani F, Baralle FE. Genomic variants in exons and introns: identifying the splicing spoilers. *Nat Rev Genet.* 2004 May;5(5):389-96.
- 402) Graveley BR. The haplo-spliceo-transcriptome: common variations in alternative splicing in the human population. *Trends Genet.* 2008 Jan;24(1):5-7.
- 403) Fairbrother WG, Yeo GW, Yeh R, et al. RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. *Nucleic Acids Res.* 2004 Jul 1;32(Web Server issue):W187-90.
- 404) Steffens NO, Galuschka C, Schindler M, et al. AthaMap: an online resource for in silico transcription factor binding sites in the Arabidopsis thaliana genome. *Nucleic Acids Res.* 2004 Jan 1;32(Database issue):D368-72.
- 405) Matys V, Kel-Margoulis OV, Fricke E, et al. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.* 2006 Jan 1;34(Database issue):D108-10.
- 406) Sandelin A, Wasserman WW, Lenhard B. ConSite: web-based prediction of regulatory elements using cross-species comparison. *Nucleic Acids Res.* 2004 Jul 1;32(Web Server issue):W249-52.
- 407) Liu Y, Liu XS, Wei L, et al. Eukaryotic regulatory element conservation analysis and identification using comparative genomics. *Genome Res.* 2004 Mar;14(3):451-8.
- 408) Mallery DL, Vandenberg CJ, Hiom K. Activation of the E3 ligase function of the BRCA1/BARD1 complex by polyubiquitin chains. *EMBO J.* 2002 Dec 16;21(24):6755-62.
- 409) Pickart CM. Ubiquitin enters the new millennium. *Mol Cell.* 2001 Sep;8(3):499-504.



- 410) Ohta T, Fukuda M. Ubiquitin and breast cancer. *Oncogene*. 2004 Mar 15;23(11):2079-88.
- 411) Starita LM, Horwitz AA, Keogh MC, et al. BRCA1/BARD1 ubiquitinate phosphorylated RNA polymerase II. *J Biol Chem*. 2005 Jul 1;280(26):24498-505.
- 412) Laufer M, Nandula SV, Modi AP, et al. Structural requirements for the BARD1 tumor suppressor in chromosomal stability and homology-directed DNA repair. *J Biol Chem*. 2007 Nov 23;282(47):34325-33.
- 413) Starita LM, Machida Y, Sankaran S, et al. BRCA1-dependent ubiquitination of gamma-tubulin regulates centrosome number. *Mol Cell Biol*. 2004 Oct;24(19):8457-66.
- 414) Yu X, Fu S, Lai M, et al. BRCA1 ubiquitinates its phosphorylation-dependent binding partner CtIP. *Genes Dev*. 2006 Jul 1;20(13):1721-6.
- 415) Schüchner S, Tembe V, Rodriguez JA, et al. Nuclear targeting and cell cycle regulatory function of human BARD1. *J Biol Chem*. 2005 Mar 11;280(10):8855-61.
- 416) Wilson CA, Ramos L, Villaseñor MR, et al. Localization of human BRCA1 and its loss in high-grade, non-inherited breast carcinomas. *Nat Genet*. 1999 Feb;21(2):236-40.
- 417) Fabbro M, Rodriguez JA, Baer R, et al. BARD1 induces BRCA1 intranuclear foci formation by increasing RING-dependent BRCA1 nuclear import and inhibiting BRCA1 nuclear export. *J Biol Chem*. 2002 Jun 14;277(24):21315-24.
- 418) Fabbro M, Schuechner S, Au WW, et al. BARD1 regulates BRCA1 apoptotic function by a mechanism involving nuclear retention. *Exp Cell Res*. 2004 Aug 15;298(2):661-73.
- 419) Feki A, Jefford CE, Berardi P, et al. BARD1 induces apoptosis by catalysing phosphorylation of p53 by DNA-damage response kinase. *Oncogene*. 2005 May 26;24(23):3726-36.
- 420) Rodriguez JA, Schüchner S, Au WW, et al. Nuclear-cytoplasmic shuttling of BARD1 contributes to its proapoptotic activity and is regulated by dimerization with BRCA1. *Oncogene*. 2004 Mar 11;23(10):1809-20.
- 421) Joukov V, Groen AC, Prokhorova T, et al. The BRCA1/BARD1 heterodimer modulates ran-dependent mitotic spindle assembly. *Cell*. 2006 Nov 3;127(3):539-52.
- 422) Wu JY, Vlastos AT, Pelte MF, et al. Aberrant expression of BARD1 in breast and ovarian cancers with poor prognosis. *Int J Cancer*. 2006 Mar 1;118(5):1215-26.
- 423) Li L, Ryser S, Dizin E, et al. Oncogenic BARD1 isoforms expressed in gynecological cancers. *Cancer Res*. 2007 Dec 15;67(24):11876-85.
- 424) Daniels MJ, Wang Y, Lee M, et al. Abnormal cytokinesis in cells deficient in the breast cancer susceptibility protein BRCA2. *Science*. 2004 Oct 29;306(5697):876-9.
- 425) Delaval B, Ferrand A, Conte N, et al. Aurora B -TACC1 protein complex in cytokinesis. *Oncogene*. 2004 Jun 3;23(26):4516-22.
- 426) Ryser S, Dizin E, Jefford CE, et al. Distinct roles of BARD1 isoforms in mitosis: full-length BARD1 mediates Aurora B degradation, cancer-associated BARD1beta scaffolds Aurora B and BRCA2. *Cancer Res*. 2009 Feb 1;69(3):1125-34.
- 427) Ghimenti C, Sensi E, Presciuttini S, et al. Germline mutations of the BRCA1-associated ring domain (BARD1) gene in breast and breast/ovarian families negative for BRCA1 and BRCA2 alterations. *Genes Chromosomes Cancer*. 2002 Mar;33(3):235-42.

- 428) Sauer MK, Andrulis IL. Identification and characterization of missense alterations in the BRCA1 associated RING domain (BARD1) gene in breast and ovarian cancer. *J Med Genet.* 2005 Aug;42(8):633-8.
- 429) Karppinen SM, Heikkinen K, Rapakko K, et al. Mutation screening of the BARD1 gene: evidence for involvement of the Cys557Ser allele in hereditary susceptibility to breast cancer. *J Med Genet.* 2004 Sep;41(9):e114.
- 430) Karppinen SM, Barkardottir RB, Backenhorn K, et al. Nordic collaborative study of the BARD1 Cys557Ser allele in 3956 patients with cancer: enrichment in familial BRCA1/BRCA2 mutation-negative breast cancer but not in other malignancies. *J Med Genet.* 2006 Nov;43(11):856-62.
- 431) Stacey SN, Sulem P, Johannsson OT, et al. The BARD1 Cys557Ser variant and breast cancer risk in Iceland. *PLoS Med.* 2006 Jul;3(7):e217.
- 432) Vahteristo P, Syrjäkoski K, Heikkinen T, et al. BARD1 variants Cys557Ser and Val507Met in breast cancer predisposition. *Eur J Hum Genet.* 2006 Feb;14(2):167-72.
- 433) Goringe KL, Choong DY, Visvader JE, et al. BARD1 variants are not associated with breast cancer risk in Australian familial breast cancer. *Breast Cancer Res Treat.* 2008 Oct;111(3):505-9.
- 434) Jakubowska A, Cybulski C, Szymanska A, et al. BARD1 and breast cancer in Poland. *Breast Cancer Res Treat.* 2008 Jan;107(1):119-22.
- 435) Johnatty SE, Beesley J, Chen X, et al. The BARD1 Cys557Ser polymorphism and breast cancer risk: an Australian case-control and family analysis. *Breast Cancer Res Treat.* 2009 May;115(1):145-50.
- 436) Huo X, Hu Z, Zhai X, et al. Common non-synonymous polymorphisms in the BRCA1 Associated RING Domain (BARD1) gene are associated with breast cancer susceptibility: a case-control analysis. *Breast Cancer Res Treat.* 2007 May;102(3):329-37.
- 437) Jensen DE, Proctor M, Marquis ST, et al. BAP1: a novel ubiquitin hydrolase which binds to the BRCA1 RING finger and enhances BRCA1-mediated cell growth suppression. *Oncogene.* 1998 Mar 5;16(9):1097-112.
- 438) Jensen DE, Rauscher FJ 3rd. Defining biochemical functions for the BRCA1 tumor suppressor protein: analysis of the BRCA1 binding protein BAP1. *Cancer Lett.* 1999 Sep;143 Suppl 1:S13-7.
- 439) Hu JJ, Mohrenweiser HW, Bell DA, et al. Symposium overview: genetic polymorphisms in DNA repair and cancer risk. *Toxicol Appl Pharmacol.* 2002 Nov 15;185(1):64-73.
- 440) Abbott DW, Thompson ME, Robinson-Benion C, et al. BRCA1 expression restores radiation resistance in BRCA1-defective cancer cells through enhancement of transcription-coupled DNA repair. *J Biol Chem.* 1999 Jun 25;274(26):18808-12.
- 441) Durocher F, Guénard F, Desjardins S, et al. (2005) Inherited susceptibility to breast cancer: accomplishments and challenges. In: Sinnett D (ed) *Molecular genetics of cancer.* Research Signpost, Kerala, pp 19–93.
- 442) Scully R, Anderson SF, Chao DM, et al. BRCA1 is a component of the RNA polymerase II holoenzyme. *Proc Natl Acad Sci U S A.* 1997 May 27;94(11):5605-10.
- 443) Ventii KH, Devi NS, Friedrich KL, et al. BRCA1-associated protein-1 is a tumor suppressor that requires deubiquitinating activity and nuclear localization. *Cancer Res.* 2008 Sep 1;68(17):6953-62.

- 444) Misaghi S, Ottosen S, Izrael-Tomasevic A, et al. Association of C-terminal ubiquitin hydrolase BRCA1-associated protein 1 with cell cycle regulator host cell factor 1. *Mol Cell Biol.* 2009 Apr;29(8):2181-92.
- 445) Nishikawa H, Wu W, Koike A, et al. BRCA1-associated protein 1 interferes with BRCA1/BARD1 RING heterodimer activity. *Cancer Res.* 2009 Jan 1;69(1):111-9.
- 446) Chen LC, Matsumura K, Deng G, et al. Deletion of two separate regions on chromosome 3p in breast cancers. *Cancer Res.* 1994 Jun 1;54(11):3021-4.
- 447) Maitra A, Wistuba II, Washington C, et al. High-resolution chromosome 3p allelotyping of breast carcinomas and precursor lesions demonstrates frequent loss of heterozygosity and a discontinuous pattern of allele loss. *Am J Pathol.* 2001 Jul;159(1):119-30.
- 448) Coupier I, Cousin PY, Hughes D, et al. BAP1 and breast cancer risk. *Fam Cancer.* 2005;4(4):273-7.
- 449) Ewart-Toland A, Briassouli P, de Koning JP, et al. Identification of Stk6/STK15 as a candidate low-penetrance tumor-susceptibility gene in mouse and human. *Nat Genet.* 2003 Aug;34(4):403-12.
- 450) Bischoff JR, Anderson L, Zhu Y, et al. A homologue of *Drosophila* aurora kinase is oncogenic and amplified in human colorectal cancers. *EMBO J.* 1998 Jun 1;17(11):3052-65.
- 451) Hsu LC, White RL. BRCA1 is associated with the centrosome during mitosis. *Proc Natl Acad Sci U S A.* 1998 Oct 27;95(22):12983-8.
- 452) Sankaran S, Starita LM, Groen AC, et al. Centrosomal microtubule nucleation activity is inhibited by BRCA1-dependent ubiquitination. *Mol Cell Biol.* 2005 Oct;25(19):8656-68.
- 453) Xu X, Weaver Z, Linke SP, et al. Centrosome amplification and a defective G2-M cell cycle checkpoint induce genetic instability in BRCA1 exon 11 isoform-deficient cells. *Mol Cell.* 1999 Mar;3(3):389-95.
- 454) Sankaran S, Starita LM, Simons AM, et al. Identification of domains of BRCA1 critical for the ubiquitin-dependent inhibition of centrosome function. *Cancer Res.* 2006 Apr 15;66(8):4100-7.
- 455) Giet R, McLean D, Descamps S, et al. *Drosophila* Aurora A kinase is required to localize D-TACC to centrosomes and to regulate astral microtubules. *J Cell Biol.* 2002 Feb 4;156(3):437-51.
- 456) Giet R, Prigent C. The *Xenopus laevis* aurora/Ip11p-related kinase pEg2 participates in the stability of the bipolar mitotic spindle. *Exp Cell Res.* 2000 Jul 10;258(1):145-51.
- 457) Sankaran S, Crone DE, Palazzo RE, et al. Aurora-A kinase regulates breast cancer associated gene 1 inhibition of centrosome-dependent microtubule nucleation. *Cancer Res.* 2007 Dec 1;67(23):11186-94.
- 458) Nilsson I, Hoffmann I. Cell cycle regulation by the Cdc25 phosphatase family. *Prog Cell Cycle Res.* 2000;4:107-14.
- 459) Duttre S, Cazales M, Quaranta M, et al. Phosphorylation of CDC25B by Aurora-A at the centrosome contributes to the G2-M transition. *J Cell Sci.* 2004 May 15;117(Pt 12):2523-31.
- 460) Sen S, Zhou H, White RA. A putative serine/threonine kinase encoding gene BTAK on chromosome 20q13 is amplified and overexpressed in human breast cancer cell lines. *Oncogene.* 1997 May 8;14(18):2195-200.

- 461) Zhou H, Kuang J, Zhong L, et al. Tumour amplified kinase STK15/BTAK induces centrosome amplification, aneuploidy and transformation. *Nat Genet.* 1998 Oct;20(2):189-93.
- 462) Jeng YM, Peng SY, Lin CY, et al. Overexpression and amplification of Aurora-A in hepatocellular carcinoma. *Clin Cancer Res.* 2004 Mar 15;10(6):2065-71.
- 463) Buschhorn HM, Klein RR, Chambers SM, et al. Aurora-A over-expression in high-grade PIN lesions and prostate cancer. *Prostate.* 2005 Sep 1;64(4):341-6.
- 464) Miao X, Sun T, Wang Y, et al. Functional STK15 Phe31Ile polymorphism is associated with the occurrence and advanced disease status of esophageal squamous cell carcinoma. *Cancer Res.* 2004 Apr 15;64(8):2680-3.
- 465) Ewart-Toland A, Dai Q, Gao YT, et al. Aurora-A/STK15 T+91A is a general low penetrance cancer susceptibility gene: a meta-analysis of multiple cancer types. *Carcinogenesis.* 2005 Aug;26(8):1368-73.
- 466) Dicioccio RA, Song H, Waterfall C, et al. STK15 polymorphisms and association with risk of invasive ovarian cancer. *Cancer Epidemiol Biomarkers Prev.* 2004 Oct;13(10):1589-94.
- 467) Ramus SJ, Vierkant RA, Johnatty SE, et al. Consortium analysis of 7 candidate SNPs for ovarian cancer. *Int J Cancer.* 2008 Jul 15;123(2):380-8.
- 468) Dai Q, Cai QY, Shu XO, et al. Synergistic effects of STK15 gene polymorphisms and endogenous estrogen exposure in the risk of breast cancer. *Cancer Epidemiol Biomarkers Prev.* 2004 Dec;13(12):2065-70.
- 469) Egan KM, Newcomb PA, Ambrosone CB, et al. STK15 polymorphism and breast cancer risk in a population-based study. *Carcinogenesis.* 2004 Nov;25(11):2149-53.
- 470) Sun T, Miao X, Wang J, et al. Functional Phe31Ile polymorphism in Aurora A and risk of breast carcinoma. *Carcinogenesis.* 2004 Nov;25(11):2225-30.
- 471) Lo YL, Yu JC, Chen ST, et al. Breast cancer risk associated with genotypic polymorphism of the mitosis-regulating gene Aurora-A/STK15/BTAK. *Int J Cancer.* 2005 Jun 10;115(2):276-83.
- 472) Cox DG, Hankinson SE, Hunter DJ. Polymorphisms of the AURKA (STK15/Aurora Kinase) Gene and Breast Cancer Risk (United States). *Cancer Causes Control.* 2006 Feb;17(1):81-3.
- 473) Fletcher O, Johnson N, Palles C, et al. Inconsistent association between the STK15 F31I genetic polymorphism and breast cancer risk. *J Natl Cancer Inst.* 2006 Jul 19;98(14):1014-8.
- 474) Tchatchou S, Wirtenberger M, Hemminki K, et al. Aurora kinases A and B and familial breast cancer risk. *Cancer Lett.* 2007 Mar 18;247(2):266-72.
- 475) Vidarsdottir L, Bodvarsdottir SK, Hilmarsdottir H, et al. Breast cancer risk associated with AURKA 91T -->A polymorphism in relation to BRCA mutations. *Cancer Lett.* 2007 Jun 8;250(2):206-12.
- 476) Couch FJ, Sinilnikova O, Vierkant RA, et al. AURKA F31I polymorphism and breast cancer risk in BRCA1 and BRCA2 mutation carriers: a consortium of investigators of modifiers of BRCA1/2 study. *Cancer Epidemiol Biomarkers Prev.* 2007 Jul;16(7):1416-21.
- 477) Lee CG, Eki T, Okumura K, et al. The human RNA helicase A (DDX9) gene maps to the prostate cancer susceptibility locus at chromosome band 1q25 and its pseudogene (DDX9P) to 13q22, respectively. *Somat Cell Mol Genet.* 1999 Jan;25(1):33-9.

- 478) Lee CG, Hurwitz J. A new RNA helicase isolated from HeLa cells that catalytically translocates in the 3' to 5' direction. *J Biol Chem*. 1992 Mar 5;267(7):4398-407.
- 479) Lee CG, Hurwitz J. Human RNA helicase A is homologous to the maleless protein of *Drosophila*. *J Biol Chem*. 1993 Aug 5;268(22):16822-30.
- 480) Zhang S, Maacke H, Grosse F. Molecular cloning of the gene encoding nuclear DNA helicase II. A bovine homologue of human RNA helicase A and *Drosophila* Mle protein. *J Biol Chem*. 1995 Jul 7;270(27):16422-7.
- 481) Zhang SS, Grosse F. Purification and characterization of two DNA helicases from calf thymus nuclei. *J Biol Chem*. 1991 Oct 25;266(30):20483-90.
- 482) Zhang S, Grosse F. Nuclear DNA helicase II unwinds both DNA and RNA. *Biochemistry*. 1994 Apr 5;33(13):3906-12.
- 483) Zhang S, Grosse F. Domain structure of human nuclear DNA helicase II (RNA helicase A). *J Biol Chem*. 1997 Apr 25;272(17):11487-94.
- 484) Tang H, McDonald D, Middlesworth T, et al. The carboxyl terminus of RNA helicase A contains a bidirectional nuclear transport domain. *Mol Cell Biol*. 1999 May;19(5):3540-50.
- 485) Nakajima T, Uchida C, Anderson SF, et al. RNA helicase A mediates association of CBP with RNA polymerase II. *Cell*. 1997 Sep 19;90(6):1107-12.
- 486) Tetsuka T, Uranishi H, Sanda T, et al. RNA helicase A interacts with nuclear factor kappaB p65 and functions as a transcriptional coactivator. *Eur J Biochem*. 2004 Sep;271(18):3741-51.
- 487) Zhang S, Buder K, Burkhardt C, et al. Nuclear DNA helicase II/RNA helicase A binds to filamentous actin. *J Biol Chem*. 2002 Jan 4;277(1):843-53.
- 488) Kang Y, Cullen BR. The human Tap protein is a nuclear mRNA export factor that contains novel RNA-binding and nucleocytoplasmic transport sequences. *Genes Dev*. 1999 May 1;13(9):1126-39.
- 489) Westberg C, Yang JP, Tang H, et al. A novel shuttle protein binds to RNA helicase A and activates the retroviral constitutive transport element. *J Biol Chem*. 2000 Jul 14;275(28):21396-401.
- 490) Hartman TR, Qian S, Bolinger C, et al. RNA helicase A is necessary for translation of selected messenger RNAs. *Nat Struct Mol Biol*. 2006 Jun;13(6):509-16.
- 491) Friedemann J, Grosse F, Zhang S. Nuclear DNA helicase II (RNA helicase A) interacts with Werner syndrome helicase and stimulates its exonuclease activity. *J Biol Chem*. 2005 Sep 2;280(35):31303-13.
- 492) Wei X, Pacyna-Gengelbach M, Schlüns K, et al. Analysis of the RNA helicase A gene in human lung cancer. *Oncol Rep*. 2004 Jan;11(1):253-8.
- 493) El Gedaily A, Bubendorf L, Willi N, et al. Discovery of new DNA amplification loci in prostate cancer by comparative genomic hybridization. *Prostate*. 2001 Feb 15;46(3):184-90.
- 494) Simard J, Dumont M, Moisan AM, 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 Feb;44(2):107-21.
- 495) Pharoah PD, Dunning AM, Ponder BA, et al. Association studies for finding cancer-susceptibility genetic variants. *Nat Rev Cancer*. 2004 Nov;4(11):850-60.

- 496) Oros KK, Ghadirian P, Greenwood CM, 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 Nov 10;112(3):411-9.
- 497) Novak DJ, Chen LQ, Ghadirian P, et al. Identification of a novel CHEK2 variant and assessment of its contribution to the risk of breast cancer in French Canadian women. *BMC Cancer*. 2008 Aug 15;8:239.
- 498) Guénard F, St-Laurent Pedneault C, Ouellette G, et al. Evaluation of the contribution of the three breast cancer susceptibility genes CHEK2, STK11 and PALB2 in non-BRCA1/2 French Canadian families with high risk of breast cancer. *Genet Test Mol Biomarkers*. *Submitted 2009*.
- 499) Guénard F, Labrie Y, Ouellette G, et al. Mutational analysis of the breast cancer susceptibility gene BRIP1/BACH1/FANCI in high-risk non-BRCA1/BRCA2 breast cancer families. *J Hum Genet*. 2008;53(7):579-91.
- 500) Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet*. 2002 May;3(5):391-7.
- 501) Pharoah PD, Antoniou AC, Easton DF, et al. Polygenes, risk prediction, and targeted prevention of breast cancer. *N Engl J Med*. 2008 Jun 26;358(26):2796-803.
- 502) Couch FJ, DeShano ML, Blackwood MA, et al. BRCA1 mutations in women attending clinics that evaluate the risk of breast cancer. *N Engl J Med*. 1997 May 15;336(20):1409-15.
- 503) Frank TS, Deffenbaugh AM, Reid JE, et al. Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. *J Clin Oncol*. 2002 Mar 15;20(6):1480-90.
- 504) Parmigiani G, Berry D, Aguilar O. Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2. *Am J Hum Genet*. 1998 Jan;62(1):145-58.
- 505) Evans DG, Eccles DM, Rahman N, et al. A new scoring system for the chances of identifying a BRCA1/2 mutation outperforms existing models including BRCAPRO. *J Med Genet*. 2004 Jun;41(6):474-80.
- 506) Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst*. 1989 Dec 20;81(24):1879-86.
- 507) Claus EB, Risch N, Thompson WD. Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am J Hum Genet*. 1991 Feb;48(2):232-42.
- 508) Antoniou AC, Pharoah PP, Smith P et al. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. *Br J Cancer*. 2004 Oct 18;91(8):1580-90.
- 509) Bradbury AR, Olopade OI. Genetic susceptibility to breast cancer. *Rev Endocr Metab Disord*. 2007 Sep;8(3):255-67.
- 510) Petrucelli N, Lazebnik N, Huelsman KM, et al. Clinical interpretation and recommendations for patients with a variant of uncertain significance in BRCA1 or BRCA2: a survey of genetic counseling practice. *Genet Test*. 2002 Summer;6(2):107-13.
- 511) Chenevix-Trench G, Healey S, Lakhani S, et al. Genetic and histopathologic evaluation of BRCA1 and BRCA2 DNA sequence variants of unknown clinical significance. *Cancer Res*. 2006 Feb 15;66(4):2019-27.

## **ANNEXE**

### ***AURKA* F31I Polymorphism and Breast Cancer Risk in *BRCA1* and *BRCA2* Mutation Carriers: A Consortium of Investigators of Modifiers of *BRCA1/2* Study**

#### **1. Résumé**

L'oncogène *AURKA* est associé à une ségrégation anormale des chromosomes et à une prédisposition au cancer. *AURKA* est plus fréquemment amplifié dans les tumeurs de porteurs de mutations dans les gènes *BRCA1* et *BRCA2* que dans des tumeurs sporadiques, suggérant ainsi qu'une surexpression d'*AURKA* et une inactivation de *BRCA1* et *BRCA2* coopèrent lors du développement et de la progression des tumeurs. Le polymorphisme F31I d'*AURKA* à l'état homozygote a été associé à un risque accru de cancer du sein. Afin de déterminer s'il modifie le risque de cancer du sein chez les porteurs de mutations dans les gènes *BRCA1* et *BRCA2*, un total de 4935 et 2241 porteurs de mutations dans les gènes *BRCA1* et *BRCA2* respectivement et 11 porteurs de mutations dans *BRCA1* et *BRCA2* ont été génotypés pour F31I. Ce polymorphisme d'*AURKA* ne serait pas associé à un risque modifié de cancer du sein chez les porteurs de mutations dans les gènes *BRCA1* et *BRCA2*.

Couch FJ, Sinilnikova O, Vierkant RA, et al. *AURKA* F31I polymorphism and breast cancer risk in *BRCA1* and *BRCA2* mutation carriers: a consortium of investigators of modifiers of *BRCA1/2* study. *Cancer Epidemiol Biomarkers Prev.* 2007 Jul;16(7):1416-21.

#### **2. Manuscrit**

# AURKA F31I Polymorphism and Breast Cancer Risk in BRCA1 and BRCA2 Mutation Carriers: A Consortium of Investigators of Modifiers of BRCA1/2 Study

Fergus J. Couch,<sup>1</sup> Olga Sinilnikova,<sup>4</sup> Robert A. Vierkant,<sup>3</sup> V. Shane Pankratz,<sup>3</sup> Zachary S. Fredericksen,<sup>3</sup> Dominique Stoppa-Lyonnet,<sup>6</sup> Isabelle Coupier,<sup>6</sup> David Hughes,<sup>5</sup> Agnès Hardouin,<sup>7</sup> Pascaline Berthet,<sup>7</sup> Susan Peock,<sup>8</sup> Margaret Cook,<sup>8</sup> Caroline Baynes,<sup>9</sup> Shirley Hodgson,<sup>10</sup> Patrick J. Morrison,<sup>11</sup> Mary E. Porteous,<sup>12</sup> Anna Jakubowska,<sup>13</sup> Jan Lubinski,<sup>13</sup> Jacek Gronwald,<sup>13</sup> Amanda B. Spurdle,<sup>14</sup> kConFab,<sup>15</sup> Rita Schmutzler,<sup>16</sup> Beatrix Versmold,<sup>16</sup> Christoph Engel,<sup>17</sup> Alfons Meindl,<sup>18</sup> Christian Sutter,<sup>19</sup> Jurgen Horst,<sup>21</sup> Dieter Schaefer,<sup>22</sup> Kenneth Offit,<sup>23</sup> Tomas Kirchhoff,<sup>23</sup> Irene L. Andrulis,<sup>24,25</sup> Eduard Ilyushik,<sup>24</sup> Gordon Glendon,<sup>25</sup> Peter Devilee,<sup>27</sup> Maaïke P.G. Vreeswijk,<sup>27</sup> Hans F.A. Vasen,<sup>28</sup> Ake Borg,<sup>29</sup> Katja Backenhorst,<sup>29</sup> Jeffery P. Struwing,<sup>30</sup> Mark H. Greene,<sup>31</sup> Susan L. Neuhausen,<sup>32</sup> Timothy R. Rebbeck,<sup>33</sup> Katherine Nathanson,<sup>33</sup> Susan Domchek,<sup>33</sup> Theresa Wagner,<sup>35</sup> Judy E. Garber,<sup>36</sup> Csilla Szabo,<sup>1</sup> Michal Zikan,<sup>37</sup> Lenka Foretova,<sup>38</sup> Janet E. Olson,<sup>3</sup> Thomas A. Sellers,<sup>39</sup> Noralane Lindor,<sup>2</sup> Heli Nevanlinna,<sup>40</sup> Johanna Tammiska,<sup>40</sup> Kristiina Aittomäki,<sup>41</sup> Ute Hamann,<sup>20</sup> Muhammad U. Rashid,<sup>20</sup> Diana Torres,<sup>20</sup> Jacques Simard,<sup>20</sup> Francine Durocher,<sup>42</sup> Frederic Guenard,<sup>42</sup> Henry T. Lynch,<sup>43</sup> Claudine Isaacs,<sup>44</sup> Jeffrey Weitzel,<sup>45</sup> Olufunmilayo I. Olopade,<sup>46</sup> Steven Narod,<sup>26</sup> Mary B. Daly,<sup>34</sup> Andrew K. Godwin,<sup>34</sup> Gail Tomlinson,<sup>47</sup> Douglas F. Easton,<sup>8</sup> Georgia Chenevix-Trench,<sup>14</sup> and Antonis C. Antoniou<sup>8</sup> on behalf of the Consortium of Investigators of Modifiers of BRCA1/2

<sup>1</sup>Departments of Laboratory Medicine and Pathology, <sup>2</sup>Medical Genetics, and <sup>3</sup>Health Sciences Research, Mayo Clinic College of Medicine, Rochester, Minnesota; <sup>4</sup>Unité Mixte de Génétique Constitutionnelle des Cancers Fréquents, Hospices Civils de Lyon-Centre Léon Bérard; <sup>5</sup>IARC, Lyon, France; <sup>6</sup>Institut National de la Santé et de la Recherche Médicale U509, Service de Génétique Oncologique, Institut Curie, Paris, France; <sup>7</sup>Centre François Baclesse, Caen, France; <sup>8</sup>Cancer Research UK, Genetic Epidemiology Unit, Department of Public Health and Primary Care and <sup>9</sup>Cancer Research UK, Human Cancer Genetics Group, Department of Oncology, University of Cambridge, Cambridge, United Kingdom; <sup>10</sup>South West Thames Regional Genetics Service, St. Georges Hospital, London, United Kingdom; <sup>11</sup>Northern Ireland Regional Genetics Centre, Cancer Genetics Service, Belfast City Hospital, Belfast, United Kingdom; <sup>12</sup>South East of Scotland Clinical Genetics Service, Western General Hospital, Edinburgh, United Kingdom; <sup>13</sup>Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland; <sup>14</sup>Queensland Institute of Medical Research, Brisbane, Queensland, Australia; <sup>15</sup>The Kathleen Cuninghame Foundation Consortium for Research into Familial Breast Cancer, Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia; <sup>16</sup>Division of Molecular Gynecology-Oncology, University of Cologne, Cologne, Germany; <sup>17</sup>Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany; <sup>18</sup>Department of Gynaecology and Obstetrics, Technical University, Munich, Germany; <sup>19</sup>Institute of Human Genetics, University of Heidelberg; <sup>20</sup>Deutsches Krebsforschungszentrum Heidelberg, Heidelberg, Germany; <sup>21</sup>Institute of Human Genetics, University of Muenster, Muenster, Germany; <sup>22</sup>Institute of Human Genetics, University of Frankfurt, Frankfurt, Germany; <sup>23</sup>Clinical Genetics Service, Memorial Sloan-Kettering Cancer Center, New York, New York; <sup>24</sup>Fred A. Litwin Center for Cancer Genetics, Samuel Lunenfeld Research Institute, Mount Sinai Hospital; <sup>25</sup>Ontario Cancer Genetics Network, Cancer Care Ontario; <sup>26</sup>Centre for Research in Women's Health, Toronto, Ontario, Canada; <sup>27</sup>Department of Human Genetics, Leiden University Medical Center; <sup>28</sup>The Foundation for the Detection of Hereditary Tumors, Leiden, the Netherlands; <sup>29</sup>Department of Oncology, Lund University Hospital, Lund, Sweden; <sup>30</sup>Clinical Genetics Branch and <sup>31</sup>Laboratory of Population Genetics, National Cancer Institute, NIH, Bethesda, Maryland; <sup>32</sup>Division of Epidemiology, Department of Medicine, University of California Irvine, Irvine, California; <sup>33</sup>Abramson Cancer Center, University of Pennsylvania School of Medicine; <sup>34</sup>Fox Chase Cancer Center, Philadelphia, Pennsylvania; <sup>35</sup>Division of Senology, Medical University of Vienna, Vienna, Austria; <sup>36</sup>Dana-Farber Cancer Institute, Boston, Massachusetts; <sup>37</sup>Department of Biochemistry and Experimental Oncology, First Faculty of Medicine, Charles University, Prague, Czech Republic; <sup>38</sup>Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute, Brno, Czech Republic; <sup>39</sup>Division of Cancer Prevention and Control, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida; <sup>40</sup>Departments of <sup>41</sup>Obstetrics and Gynecology and <sup>42</sup>Clinical Genetics, Helsinki University Central Hospital, Helsinki, Finland; <sup>43</sup>Cancer Genomics Laboratory, Oncology and Molecular Endocrinology Research Centre, Centre Hospitalier Universitaire de Québec and Laval University, Québec City, Québec, Canada; <sup>44</sup>Department of Preventive Medicine and Public Health, Creighton University, Omaha, Nebraska; <sup>45</sup>Lombardi Cancer Center, Georgetown University, Washington, District of Columbia; <sup>46</sup>City of Hope Cancer Center, Duarte, California; <sup>47</sup>University of Chicago, Chicago, Illinois; and <sup>8</sup>University of Texas Southwestern Medical Center at Dallas, Dallas, Texas

## Abstract

The AURKA oncogene is associated with abnormal chromosome segregation and aneuploidy and predisposition to cancer. Amplification of AURKA has been detected at higher frequency in tumors from BRCA1 and BRCA2 mutation

carriers than in sporadic breast tumors, suggesting that overexpression of AURKA and inactivation of BRCA1 and BRCA2 cooperate during tumor development and progression. The F31I polymorphism in AURKA has been associated

Received 2/19/07; revised 4/10/07; accepted 4/23/07.

**Grant support:** Breast Cancer Research Foundation, U.S. Army Medical Research and Materiel Command grant W81XWH-04-1-0588, and Mayo Clinic Breast Cancer Specialized Program of Research Excellence grant P50-CA116201 (F.J. Couch); Cancer Care Ontario and RFA CA-95-003 as part of the National Cancer Institute Breast Cancer Family Registries (I.L. Andrulis); Canadian Institute for Health Research through the INHERIT BRCA program (J. Simard); Dutch Cancer Society grant UL2001-2471 (P. Devilee); Academy of Finland (110663), Finnish Cancer Society, Helsinki University Central Hospital Research Fund, and Sigrid Juselius Fund (H. Nevanlinna); Deutsches Krebsforschungszentrum Heidelberg (U. Hamann); NIH grants R01-CA74415 (S.L. Neuhausen) and R01-CA083855 and R01-CA102776 (T.R. Rebbeck); The Programme Hospitalier de Recherche Clinique AOR01082 (D. Stoppa-Lyonnet and O. Sinilnikova); Komen Foundation BCTR 0601361 and Weissenbach-Southworth-Niehaus Research Fund (K. Offit); The German Cancer Aid grant 107054 and Center for Molecular Medicine Cologne grant TV 93 (R. Schmutzler). A General Clinic Research Center grant from the NIH (M01 RR00043) awarded to the City of Hope National

Center supports in part the collection and management of the Hereditary Cancer Registry (J. Weitzel). kConFab is supported by grants from the National Breast Cancer Foundation, the National Health and Medical Research Council, and by the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania, and South Australia, and the Cancer Foundation of Western Australia. D.F. Easton is a principal research fellow of Cancer Research UK. EMBRACE and A. Antoniou are funded by an award from Cancer Research UK.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

**Requests for reprints:** Fergus J. Couch, Mayo Clinic College of Medicine, 200 First Street Southwest, Rochester, MN 55905. Phone: 507-284-3623; Fax: 507-266-0824. E-mail: Couch.fergus@mayo.edu

Copyright © 2007 American Association for Cancer Research.

doi:10.1158/1055-9965.EPI-07-0129



with breast cancer risk in the homozygous state in prior studies. We evaluated whether the *AURKA* F31I polymorphism modifies breast cancer risk in *BRCA1* and *BRCA2* mutation carriers from the Consortium of Investigators of Modifiers of *BRCA1/2*. Consortium of Investigators of Modifiers of *BRCA1/2* was established to provide sufficient statistical power through increased numbers of mutation carriers to identify polymorphisms that act as modifiers of cancer risk and can refine breast cancer risk estimates in *BRCA1* and *BRCA2* mutation carriers. A total of 4,935 *BRCA1* and 2,241 *BRCA2* mutation carriers and 11 individuals carrying both *BRCA1* and *BRCA2* mutations was genotyped

for F31I. Overall, homozygosity for the 31I allele was not significantly associated with breast cancer risk in *BRCA1* and *BRCA2* carriers combined [hazard ratio (HR), 0.91; 95% confidence interval (95% CI), 0.77-1.06]. Similarly, no significant association was seen in *BRCA1* (HR, 0.90; 95% CI, 0.75-1.08) or *BRCA2* carriers (HR, 0.93; 95% CI, 0.67-1.29) or when assessing the modifying effects of either bilateral prophylactic oophorectomy or menopausal status of *BRCA1* and *BRCA2* carriers. In summary, the F31I polymorphism in *AURKA* is not associated with a modified risk of breast cancer in *BRCA1* and *BRCA2* carriers. (Cancer Epidemiol Biomarkers Prev 2007;16(7):1416-21)

## Introduction

The *AURORA-A/AURKA/BTAK/STK15* gene encodes a serine/threonine kinase that regulates mitotic chromosome segregation. *AURKA* is amplified and overexpressed in breast and other tumors and is associated with centrosome amplification, failure of cytokinesis, and aneuploidy. Genetic mapping studies in mouse models suggest that *AURKA* is a genetic modifier of cancer risk (1). In addition, mouse models of *AURKA* exhibit infrequent mammary gland tumor formation but display synergy in tumor formation when combined with overexpressed oncogenes or disrupted tumor suppressors, suggesting that *AURKA* is a low-risk cancer susceptibility gene (2).

Further evidence for a role of *AURKA* in breast cancer comes from observations that homozygosity for a F31I polymorphism in *AURKA* is associated with an increased risk for breast cancer. In a study of incident breast cancer cases ( $n = 941$ ) and age-matched population controls ( $n = 830$ ), Egan et al. (3) found that the breast cancer risk for Ile/Ile homozygotes were at increased risk for breast cancer [odds ratio (OR), 1.54; 95% confidence interval (95% CI), 0.96-2.47], although this finding was not significant. Sun et al. (4) observed that the Ile-encoding allele is the common allele in the Chinese population, whereas the Phe-encoding allele is more common in Caucasian populations (4). In addition, an association between Ile/Ile homozygotes and estrogen receptor-negative breast carcinomas (OR, 2.56; 95% CI, 1.24-5.26) was detected. Lo et al. (5) reported a significant association between *AURKA* haplotypes and breast cancer risk. Ewart-Toland et al. (6) also found an increase in cancer risk for the Ile/Ile homozygotes (OR, 1.35; 95% CI, 1.12-1.64;  $P = 0.002$ ) in a meta-analysis of data from four case-control breast cancer populations. Furthermore, postmenopausal women homozygous for the F31I and I57V alleles of *AURKA* in a case-control study nested within the Nurses' Health Study prospective cohort had an increased risk of invasive breast cancer (OR, 1.63; 95% CI, 1.08-2.45; ref. 7). In contrast, Dai et al. (8) did not observe a significant association with breast cancer risk for Ile/Ile homozygotes (OR, 1.2; 95% CI, 0.9-1.6) in a population-based case-control series of Han Chinese, and Fletcher et al. (9) found no association between Ile/Ile homozygotes and risk of bilateral breast cancer (OR, 0.63; 95% CI, 0.34-1.13). Importantly, the F31I variant has been shown to alter the activity of the Aurora box-1 of the *AURKA* protein, resulting in disruption of p53 binding and a decreased rate of degradation of *AURKA*. The stabilized *AURKA* may lead to centrosome amplification and failure of cytokinesis, increased chromosomal instability and aneuploidy, and promotion of tumor formation (1).

Mutations in *BRCA1* and *BRCA2* are correlated with aberrant duplication of the centrosome leading to centrosome amplification, chromosome missegregation, and aneuploidy (10-12). Amplification of *AURKA* has also been detected at much higher frequency in tumors from *BRCA1* and *BRCA2* mutation carriers than in sporadic breast tumors, suggesting that overexpression of *AURKA* and inactivation of *BRCA1* and *BRCA2* cooperate during tumor development and/or progres-

sion. Based on these data, we hypothesized that the F31I polymorphism modifies the risk of breast cancer in *BRCA1* and *BRCA2* mutation carriers. To address this hypothesis, *AURKA* F31I was genotyped on *BRCA1* and *BRCA2* deleterious mutation carriers from 16 clinic and population-based research studies and multicenter consortia participating in the Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA) and the association of F31I with breast cancer risk was assessed.

## Materials and Methods

**Subjects.** *BRCA1* and *BRCA2* mutation carriers were identified through 16 clinic and population-based research studies and multicenter consortia participating in the CIMBA. This international consortium was established in 2005 by a group of investigators interested in identifying modifiers of cancer risk in *BRCA1* and *BRCA2* mutation carriers that could be used to refine cancer risk estimates. Recruitment of mutation carriers for this and other CIMBA studies was approved by institutional review boards or ethics committees at all sites. *BRCA1* and *BRCA2* mutation carriers were defined as carriers of frameshifting small deletions and insertions, nonsense mutations, splice site mutations verified *in vitro*, and large genomic rearrangements that result in a premature stop codon in either *BRCA1* or *BRCA2*. These mutations were identified by a variety of screening techniques and sequence verified. As the K3326X variant in exon 27 is not associated with high risk of breast cancer, this and other mutations causing stop codons in exon 27 were excluded. Missense mutations that have been classified as pathogenic by multifactorial likelihood approaches were included in the deleterious category (12-14), whereas carriers of all other missense and intronic mutations in *BRCA1* and *BRCA2* were excluded from the study. Phenotypic data for mutation carriers were provided by each contributing center. Data were collected on year of birth, mutation description, ethnicity, country of residence, age at last follow-up, ages at breast and ovarian cancer diagnosis, age at bilateral prophylactic mastectomy, age at bilateral prophylactic oophorectomy, and status and age at menopause. These and other available epidemiologic data obtained from risk factor questionnaires and/or medical records were uniformly coded and stored in a centralized CIMBA database.

**Genotyping.** The F31I polymorphism (rs2273535) of *AURKA* was genotyped by 13 groups by the 5' nuclease assay (Taqman) on an ABI 7900HT Sequence Detection System (Applied Biosystems). PCR primers were 5'-CTGGCCAC-TATTTACAGGTAATGGA-3' (forward) and 5'-TGGAGGTC-CAAAACGTTGTTCTC-3' (reverse). Probes were VIC-ACCTA-GCAATTTCTT and FAM-CTCAGCAAATTCCTT. The annealing temperature was 60°C. Lund investigators used an alternative reverse primer (CATCTTTTGCITTCATGA-ATGCCAG) and did the 5' nuclease assay on a RotorGene

Table 1. Characteristics of study subjects by site

Source	Ascertainment	BRCA1 cases	BRCA1 unaff.*	Total BRCA1	BRCA2 cases	BRCA2 unaff.	Total BRCA2	B1/2 <sup>†</sup> cases	B1/2 unaff.	Total B1/2	Total carriers
MAGIC	Clinic	303	428	731	137	160	297	3	0	3	1,031
GEMO	Clinic	413	276	689	223	84	307	0	0	0	996
EMBRACE	Clinic	235	219	454	156	148	304	1	2	3	761
Poland	Clinic	307	427	734	0	0	0	0	0	0	734
kConFab	Clinic	203	201	404	169	143	312	0	0	0	716
GCHBOC	Clinic	286	113	399	173	52	225	3	0	3	627
MSKCC	Clinic	174	117	291	102	70	172	1	0	1	464
Ontario	Clinic and population	125	52	177	100	41	141	0	0	0	318
LUMC	Clinic	99	120	219	12	20	32	0	0	0	251
Lund	Clinic	73	88	161	38	32	70	0	0	0	231
MOD-SQUAD	Clinic	82	67	149	28	15	43	0	0	0	192
HEBCS	Clinic	56	39	95	54	40	94	0	0	0	189
DKFZ	Clinic	82	41	123	30	21	51	0	0	0	174
MAYO	Clinic	53	23	76	26	20	46	0	0	0	122
INHERIT	Clinic	33	37	70	40	41	81	0	0	0	151
NCI	Clinic	47	116	163	17	50	67	0	0	0	230
Total		2,571	2,364	4,935	1,305	937	2,242	8	2	10	7,187

Abbreviations: MAGIC, Modifiers and Genetics in Cancer; GEMO, Genetic Modifiers of cancer risk in BRCA1/2 mutation carriers study; GCHBOC, German Consortium for Hereditary Breast and Ovarian Cancer; EMBRACE, Epidemiological Study of BRCA1 and BRCA2 Mutation Carriers; kConFab, Kathleen Cunningham Consortium for Research into Familial Breast Cancer; INHERIT BRCA, Interdisciplinary Health Research International Team on Breast Cancer susceptibility; MSKCC, Memorial Sloan-Kettering Cancer Center; MAYO, Mayo Clinic; LUMC, Leiden University Medical Center; MOD-SQUAD, Modifier Study of Quantitative Effects on Disease; HEBCS, Helsinki Breast Cancer Study; DKFZ, Deutsches Krebsforschungszentrum Heidelberg; NCI, National Cancer Institute.

\*The term unaff. refers to individuals not affected with breast cancer.

<sup>†</sup>B1/2 refers to individuals with both BRCA1 and BRCA2 deleterious mutations.

(Corbett Research). INHERIT investigators directly sequenced the polymorphism using the following primers: 5'-GGGTG-AGGAATTGGAGGGGAT-3' (forward) and 5'-GGACACCA-ATTTATGCTGTGTCCT-3' (reverse). Genotyping for the HEBCS was done by Amplifluor fluorescent genotyping (KBioscience).<sup>48</sup> Genotyping for the DKFZ and Polish studies was done by fragment analysis. DNA fragments containing the polymorphism were amplified using forward primer 5'-AGTTGGAGGTCCAAAACGTG-3' and Cy5-labeled reverse primer 5'-CGCTGGGAAGTATTTGAAGG-3', digested with 2.5 units *Xap*I (Fermentas), separated on 3% agarose gel (Polish samples) or by capillary gel electrophoresis (German samples) on a CEQ 8000 DNA Analysis System (Beckmann), and sized relative to CEQ DNA Size Standard-400 in each well. Allele sizes were 114 bp for the T allele and 78 bp for the A allele.

**Statistical Methods.** Hazard ratios (HR) were modeled using Cox proportional hazards regression analysis, with breast cancer as the outcome and age as the time variable (15). We corrected for possible ascertainment bias using a weighted cohort approach (16). Briefly, this involves assigning weights to the mutation-carrying subjects such that the reweighted incidence rates observed in the study sample are consistent with the age-dependent penetrances for breast cancer onset established in carriers of inactivating mutations in BRCA1 and BRCA2. Subjects were followed from birth until the earliest occurrence of breast cancer (3,884), bilateral prophylactic mastectomy (232), ovarian cancer (643), age 80 (97), or age at last contact (2,331). Subjects were censored at age 80 because population-based incidence rates for older mutation carriers are unreliable, and accurate sampling weights cannot be assigned. Carriers with both BRCA1 and BRCA2 mutations were included once in overall analyses and were also included in each of the BRCA1 and BRCA2 gene-specific analyses. The number of subjects in each family varied from 1 to 33, with 75% of families represented by a single individual. Because the exact relationships among the family members were not available, we accounted for the nonindependence of

observations within families using a robust variance estimate (17). Primary analyses modeled AURKA as a recessive effect, comparing those with two copies of the minor allele with those with less than two copies. Secondary analyses examined associations using a two degree-of-freedom general model, simultaneously comparing subjects with one copy or with two copies of the minor allele with the subjects with zero copies.

Overall analyses were carried out for all subjects regardless of whether they carried a mutation in BRCA1 or BRCA2 or both. All analyses accounted for birth cohort and country of residence by including them as stratification variables in the Cox regression. The overall analysis also accounted for study site and mutation status. Additional analyses were conducted to obtain risk estimates for individuals with different characteristics, as defined by gene status, menopausal status, oophorectomy status, and study site. Gene-specific results accounted for study site along with birth cohort and country of residence by use of stratification variables. Site-specific results accounted for mutation status, birth cohort, and country of residence. Menopausal status and oophorectomy status were modeled as time-dependent covariates and results accounted for group status and mutation status. In secondary analyses, the influence of benign prophylactic oophorectomy and menopausal status on associations between the Ile/Ile genotype and breast cancer risk was also evaluated. As these covariates did not confound the observed associations, the associations reported in Table 2 are not adjusted for these variables.

Among those who provided ethnicity information, 97% were Caucasian, 2% were Ashkenazi Jewish, and the remaining 1% were "other." Those who did not provide ethnicity information were grouped in a separate "missing" category for analysis purposes. Ethnicity was initially included as an additional stratification variable but was subsequently excluded because of the absence of any effect on the results. We assessed the possible heterogeneity of risk ratios across study site using standard tests of interaction. A sensitivity analysis assessing the effect of possible survival bias was conducted by excluding cases ascertained more than 3 years after diagnosis. All statistical tests were two sided, and all analyses were carried out using the Statistical Analysis System (SAS Institute, Inc.) and S-Plus (Insightful) software systems.

<sup>48</sup><http://www.kbioscience.co.uk>

## Results

A total of 4,935 female *BRCA1*, 2,241 female *BRCA2* deleterious mutation carriers, and 11 individuals carrying both *BRCA1* and *BRCA2* mutations was included in this study. Of these 7,187 mutation carriers, 3,884 had a diagnosis of breast cancer at the end of follow-up and 3,303 were censored as unaffected at a mean age of 43.4 years. The distribution of *BRCA1* and *BRCA2* carriers by study site, gene, and cancer status is shown in Table 1. To avoid overlap between studies, we compared carriers by country of origin, year of birth, mutation, and reported ages. Duplication of samples between MAYO and MAGIC and between GEMO and MAGIC was detected. In both instances, the duplicated samples were excluded from the MAGIC data set.

The distribution of the *AURKA* F31I genotypes is shown in Table 2. Of the 363 (5%) carriers homozygous for the Ile-encoding allele, 188 were affected with breast cancer. The frequency of the recessive Ile/Ile-encoding genotype in the 16 groups varied between 3% and 8%, which is similar to estimates from other populations (6). There was no difference in the frequency of the Ile/Ile recessive genotype across genotyping platforms ( $P = 0.33$ ). Similarly, the study sites with the highest Ile/Ile frequencies did not have ethnic mixtures significantly different to the other study sites. The F31I polymorphism did not deviate significantly from Hardy-Weinberg equilibrium ( $P = 0.07$ ) among all 7,187 affected and unaffected carriers.

The estimated risk of breast cancer associated with the recessive genotype for F31I in *BRCA1* and *BRCA2* carriers using a weighted Cox proportional hazards model is shown in Table 2. Although there was a suggestion of a protective effect (HR, 0.91; 95% CI, 0.77-1.06), overall, the result was not statistically significant. Similarly, no association with risk was observed for individual participating centers other than for two centers (Ontario and HEBCS) that contributed small

numbers of carriers to the study (Table 2). A test for heterogeneity across study site was not significant ( $P = 0.06$ ). In an effort to account for the trend toward heterogeneity, we investigated the influence of the three sites that were significantly different from the other sites [MOD-SQUAD ( $P = 0.02$ ), GEMO ( $P = 0.01$ ), and DKFZ ( $P = 0.03$ )] on the overall effect. Exclusion of each site in turn did not substantially alter the overall HR or the significance of the association.

Because *BRCA1* is phosphorylated by *AURKA* (18), we evaluated whether the Ile/Ile genotype was associated with risk of breast cancer in *BRCA1* or *BRCA2* carriers. No significant association with risk was detected for either *BRCA1* (HR, 0.90; 95% CI, 0.75-1.08) or *BRCA2* carriers (HR, 0.93; 95% CI, 0.67-1.29; Table 2). As other studies have reported an association between the recessive Ile/Ile-encoding genotype and postmenopausal status in noncarriers (3, 7), we considered the influence of menopausal status of carriers on breast cancer risk. At the end of follow-up, 4,201 carriers were premenopausal and 2,986 were postmenopausal. No significant association with risk was detected (Table 2). Because prophylactic oophorectomy substantially reduces the risk of breast cancer in *BRCA1* and *BRCA2* mutation carriers (19), we also evaluated the influence of prophylactic oophorectomy status. A total of 707 individuals reported undergoing prophylactic oophorectomy, 4,298 reported no history of oophorectomy, whereas 2,182 (30%) provided no data at last follow-up. Associations with breast cancer risk by category of prophylactic oophorectomy did not differ markedly from the overall results. Secondary analyses using a two degree-of-freedom general model also failed to detect a significant association for either a single copy ( $P = 0.97$ ) or two copies ( $P = 0.24$ ) of the F31I polymorphism compared with no copies.

In an effort to account for possible survival bias and the inclusion of prevalent cases in the collection of *BRCA1* and

**Table 2. Association of *AURKA* F31I with breast cancer risk**

Group	0 or 1 copy Ile allele			2 copies Ile allele			HR (95% CI), all cases	HR (95% CI)* incident cases
	Unaffected	Affected	Person-years	Unaffected	Affected	Person-years		
Overall	3,128	3,696	296,122	175	188	15,793	0.91 (0.77-1.06)	0.84 (0.65-1.08)
By mutation status								
<i>BRCA1</i>	2,237	2,460	200,406	129	120	10,754	0.90 (0.75-1.08)	0.90 (0.66-1.22)
<i>BRCA2</i>	893	1,245	96,110	46	68	5,039	0.93 (0.67-1.29)	0.67 (0.44-1.03)
By menopausal status								
Premenopausal	1,935	2,049	242,208	111	106	12,834	0.84 (0.69-1.03)	0.83 (0.60-1.15)
Postmenopausal	1,193	1,647	53,914	64	82	2,959	0.96 (0.75-1.23)	0.77 (0.51-1.16)
By oophorectomy status								
No	1,772	2,318	201,303	101	107	10,474	0.85 (0.69-1.05)	0.82 (0.58-1.15)
Yes	510	160	3,793	28	9	213	1.10 (0.56-2.18)	1.03 (0.39-2.78)
Missing	846	1,218	91,026	46	72	5,106	0.97 (0.75-1.26)	0.86 (0.55-1.34)
By study site								
MAGIC	559	423	41,554	29	20	2,002	1.02 (0.63-1.67)	
GEMO	347	597	40,913	13	39	2,266	1.33 (0.97-1.82)	
EMBRACE	353	378	30,757	16	14	1,318	0.70 (0.37-1.32)	
Poland	399	285	30,360	28	22	2,197	0.98 (0.65-1.47)	
kConFab	322	362	29,568	22	10	1,251	0.64 (0.34-1.22)	
GCHBOC	157	432	24,819	8	30	1,698	0.94 (0.65-1.37)	
MSKCC	182	268	19,371	5	9	591	0.79 (0.38-1.66)	
Ontario	79	217	13,069	14	8	1,012	0.33 (0.13-0.82)	
LUMC	129	106	10,350	11	5	715	0.68 (0.32-1.44)	
Lund	113	102	11,401	7	9	803	1.05 (0.55-1.99)	
MOD-SQUAD	78	104	7,760	4	6	388	1.56 (1.04-2.36)	
HEBCS	75	108	8,451	4	2	344	0.27 (0.05-1.96)	
DKFZ	61	110	6,714	1	2	109	7.05 (0.66-75.2)	
MAYO	41	71	4,998	2	8	442	1.41 (0.65-3.07)	
INHERIT	76	70	6,668	2	3	225	1.29 (0.45-3.67)	
NCI	157	63	9,371	9	1	433	0.28 (0.05-1.77)	

NOTE: Weighted Cox proportional hazards regression analysis, modeling *AURKA* F31I as a recessive genotypic effect. Results overall by menopausal status and by oophorectomy status account for birth cohort, group status, country, and mutation status. Mutation-specific results account for birth cohort, group status, and country. Group-specific results account for birth cohort, mutation status, and country. Robust variance estimates were used to correct for possible nonindependence of study subjects.

\*Cox proportional hazards regression analysis restricted to cases for whom genetic diagnosis is less than 3 y after breast cancer diagnosis.

BRCA2 carriers, we repeated our analysis after excluding cases diagnosed more than 3 years before the date of ascertainment. For this analysis, we excluded records where an age at interview was not provided. Overall, the mean difference between age of diagnosis and age at interview for the 3,422 cases with available data was 8.7 years. Of these, 1,322 (38.6%) cases had been diagnosed less than 3 years before the date of ascertainment. When excluding prevalent cases, no association between the Ile/Ile genotype and breast cancer risk was observed, and the risk estimates were similar to those obtained when using both prevalent and incident cases (Table 2).

## Discussion

Overall, no evidence of a significant association between homozygosity for the F31I AURKA polymorphism and breast cancer risk in BRCA1 and BRCA2 mutation carriers in combination or alone was observed. These results were somewhat unexpected given the known functional relationship between AURKA and BRCA1 (18), the known influence of F31I on AURKA protein stability (1), and the significant associations with cancer risk reported in several studies of unselected breast cancer cases and controls. Although the variant does not seem to modify predisposition to cancer in this combined group of mutation carriers, the possibility remains that the Ile/Ile genotype influences tumor progression or clinical outcome or modifies cancer risk in conjunction with other risk factors. The suggestion of a modestly protective effect of the Ile/Ile genotype in this study particularly when restricting the study to incident cases supports this possibility. Interestingly, a study of bilateral breast cancer cases also identified a nonsignificant protective effect for the Ile/Ile genotype (9). This common protective effect among individuals at higher risk of breast cancer in the Caucasian population suggests that homozygosity for the F31I polymorphism may reduce cancer risk in high-risk groups while possibly increasing risk in the general population. Additional studies of other high-risk populations and the combined effects of other risk factors are needed to further evaluate these possibilities.

In this study, we accounted for the effects of both bilateral prophylactic oophorectomy and menopausal status effects by treating these factors as time-dependent variables in the analysis. As bilateral prophylactic oophorectomy is known to reduce breast cancer risk by ~50% in BRCA1 and BRCA2 mutation carriers (19), we chose to account for the remaining risk of cancer in women undergoing prophylactic oophorectomy by assessing it as an additional time-varying covariate rather than by censoring the follow-up of the women at the time they underwent this procedure. In addition, we did a sensitivity analysis to assess the potential for survival bias in our analyses by restricting the study to women more likely to have incident cases of breast cancer. Although no change in the significance of the results was observed following this approach, it is important to evaluate this possibility in any study, whether single site or multicenter, of individuals at significantly elevated risk of cancer.

This report represents the largest association study conducted to date in BRCA1 and BRCA2 carriers. It also is the first report from CIMBA, an international consortium established to provide sufficient statistical power to test candidate single nucleotide polymorphisms as modifiers of cancer risk in BRCA1 and BRCA2 mutation carriers and to refine breast cancer risk prediction in this population. The operating principles of CIMBA are as follows. (a) CIMBA is open to any group that can contribute genotype and phenotype information on at least 92 BRCA1 and/or BRCA2 mutation carriers. Groups with smaller collections of carriers are encouraged to participate through partnership with a larger group. (b) Phenotypic data obtained from risk factor ques-

tionnaires and/or medical records are uniformly coded and stored in a centralized CIMBA database. These data include year of birth, mutation description, ethnicity, country of residence, age at last follow-up, ages at breast and ovarian cancer diagnosis, age at bilateral prophylactic mastectomy, age at bilateral prophylactic oophorectomy, and status and age at menopause. (c) Panels of single nucleotide polymorphisms for genotyping are selected every 6 months at a CIMBA group meeting. (d) Only single nucleotide polymorphisms that show significant associations, either in the published literature or in data available to a member group, at  $P < 0.01$ , are considered. (e) Each investigator/group is free to participate or not in any round of genotyping. (f) Genotyping quality control standards must be followed (2% duplicates, call rates >95%, randomized arrangement of affected and unaffected carriers for genotyping). (g) Genotyping data from participating centers are pooled and analyzed as outlined in the CIMBA analysis plan. This study represents the first genetic modifier study conducted by CIMBA using these guidelines.

This study of 7,187 BRCA1 and BRCA2 carriers had 80% power to detect significant ( $P < 0.05$ ) protective recessive effects with HRs of  $\leq 0.82$  for the F31I allele. We therefore conclude that the present study has a sufficient sample size to assess with reasonable confidence the involvement of the F31I allele in the modification of breast cancer risk among BRCA1 and BRCA2 mutation carriers. It also shows the importance of large consortia, such as CIMBA, in evaluating the associations between genetic markers and cancer risk.

## Appendix 1. Study Collaborators

MAGIC collaborators: Susan Neuhausen, University of California Irvine, Irvine, CA; Timothy Rebbeck, Susan Domchek, Katherine Nathanson, University of Pennsylvania School of Medicine, Philadelphia, PA; Theresa Wagner, Medical University of Vienna, Vienna, Austria; Judy Garber, Dana-Farber Cancer Institute, Boston, MA; Henry Lynch, Creighton University, Omaha, NE; Claudine Isaacs, Lombardi Cancer Center, Georgetown University, Washington, DC; Jeffrey Weitzel, City of Hope Cancer Center, Duarte, CA; Olufunmilayo Olopade, University of Chicago, Chicago, IL; Steven Narod, Centre for Research in Women's Health, Toronto, Ontario, Canada; Mary Daly and Andrew Godwin, Fox Chase Cancer Center, Philadelphia, PA; Gail Tomlinson, University of Texas Southwestern Medical Center at Dallas, Dallas, TX; Fergus Couch, Mayo Clinic, Rochester, MN.

GEMO study collaborators: Agnès Chompret, Brigitte Bressac-de-Paillerets, Véronique Byrde, Corinne Capoulade, Gilbert Lenoir, Institut Gustave Roussy, Villejuif, France; Yves-Jean Bignon, Nancy Uhrhammer, Centre Jean Perrin, Clermont-Ferrand, France; Marion Gauthier-Villars, Muriel Belotti, Antoine de Pauw, Dominique Stoppa-Lyonnet, Institut Curie, Paris, France; Laure Barjhoux, Mélanie Léone, Sophie Giraud, Olga Sinilnikova, Hospices Civils de Lyon/Centre Léon Bérard, Lyon, France; Christine Lasset, Valérie Bonadona, Centre Léon Bérard, Lyon, France; Agnès Hardouin, Pascaline Berthet, Centre François Baclesse, Caen, France; Hagay Sobol, Institut Paoli Calmettes, Marseille, France; Florence Coulet, Chrystelle Colas, Florent Soubrier, Hôpital Pitié-Salpêtrière, Paris, France; Isabelle Coupier, CHU de Arnaud-de-Villeneuve, Montpellier, France; Jean-Philippe Peyrat, Joëlle Fournier, Philippe Vennin, Claude Adenis, Centre Oscar Lambret, Lille, France; Catherine Nogues, Centre René Huguénin, St. Cloud, France; Rosette Lidereau, Institut National de la Santé et de la Recherche Médicale U735, Centre René Huguénin, St. Cloud, France; Danièle Muller, Jean-Pierre Fricker, Centre Paul Strauss, Strasbourg, France; Michel Longy, Institut Bergonié, Bordeaux, France; Christine Toulas, Rosine Guimbaud, Laurence Gladiéff, Viviane Feillel, Institut Claudius

Regaud, Toulouse, France; Sylvie Mazoyer, Centre National de la Recherche Scientifique UMR5201, Lyon, France; Henry T. Lynch, Creighton University, Omaha, NE; Drakoulis Yannoukakos, National Center for Scientific Research Demokritos, Athens, Greece.

EMBRACE collaborators: Coordinating Centre, Cambridge: Douglas Easton, Antonis Antoniou, Susan Peock, Margaret Cook; North of Scotland Regional Genetics Service, Aberdeen: Neva Haites, Helen Gregory; Northern Ireland Regional Genetics Service, Belfast: Patrick J. Morrison; West Midlands Regional Clinical Genetics Service, Birmingham: Trevor Cole, Carole McKeown; South West Regional Genetics Service, Bristol: Alan Donaldson; East Anglian Regional Genetics Service, Cambridge: Joan Paterson; Medical Genetics Services for Wales, Cardiff: Jonathon Gray; St. James's Hospital and National Centre for Medical Genetics, Dublin: Peter Daly, David Barton; South East of Scotland Regional Genetics Service, Edinburgh: Mary Porteus, Michael Steel; Peninsula Clinical Genetics Service, Exeter: Carole Brewer, Julia Rankin; West of Scotland Regional Genetics Service, Glasgow: Rosemarie Davidson, Victoria Murday; South East Thames Regional Genetics Service, London: Louise Izatt, Gabriella Pichert; North West Thames Regional Genetics Service, Harrow: Huw Dorkins; Leicestershire Clinical Genetics Service, Leicester: Richard Trembath; Yorkshire Regional Genetics Service, Leeds: Tim Bishop, Carol Chu; Merseyside and Cheshire Clinical Genetics Service, Liverpool: Ian Ellis; Manchester Regional Genetics Service, Manchester: Gareth Evans, Fiona Lalloo, Andrew Shenton; North East Thames Regional Genetics Service, London: James Mackay, Anne Robinson; Nottingham Centre for Medical Genetics, Nottingham: Susan Ritchie, Sandy Raeburn; Northern Clinical Genetics Service, Newcastle: Fiona Douglas, John Burn; Oxford Regional Genetics Service, Oxford: Sarah Durell; Department of Cancer Genetics, Royal Marsden Hospital: Ros Eeles; North Trent Clinical Genetics Service, Sheffield: Jackie Cook, Oliver Quarrell; South West Thames Regional Genetics Service, London: Shirley Hodgson; and Wessex Clinical Genetics Service, Southampton: Diana Eccles, Anneke Lucassen.

GCHBOC study collaborators: Beatrix Versmold and Rita Schmutzler, Division of Molecular Gynecology-Oncology, University of Cologne, Cologne, Germany; Christoph Engel, Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany; Alfons Meindl, Department of Gynaecology and Obstetrics, Technical University, Munich, Germany; Christian Sutter, Institute of Human Genetics, University of Heidelberg, Heidelberg, Germany; Jurgen Horst, Institute of Human Genetics, University of Muenster, Muenster, Germany; Dieter Schaefer, Institute of Human Genetics, University of Frankfurt, Frankfurt, Germany; Norbert Arnold, University of Schleswig-Holstein, Campus Kiel, Kiel, Germany; Wera Hofmann, Max Delbrück Center for Molecular Medicine, Berlin, Germany; Dieter Niederacher, University of Düsseldorf, Düsseldorf, Germany; Helmut Deissler, University of Ulm, Ulm, Germany; Karin Kast, University of Dresden, Dresden, Germany.

kConFab collaborators: Georgia Chenevix-Trench, Amanda Spurdle, <http://www.kconfab.org/Organisation/Members.shtml>

MOD-SQUAD collaborators: Michal Zikan, Petr Pohlreich, Zdenek Kleibl, Charles University, Prague, Czech Republic; Lenka Foretova, Machackova Eva, Lukesova Miroslava, Masaryk Memorial Cancer Institute, Brno, Czech Republic.

The LUMC collaborators: Peter Devilee, Maaike P.G. Vreeswijk, Hans F.A. Vasen, H. Meijers-Heijboer, and D. Halley.

HEBCS: Heli Nevanlinna, Johanna Tammiska, Kristiina Aittomäki, Carl Blomqvist, Kirsimari Aaltonen, Anmitta Tamminen, Helsinki University Central Hospital, Helsinki, Finland.

INHERIT BRCA5 collaborators: Paul Bessette, Service de Gynécologie, Centre Hospitalier Universitaire de Sherbrooke,

Fleurimont, Quebec, Canada; Peter Bridge, Molecular Diagnostic Laboratory, Alberta Children's Hospital, Calgary, Alberta, Canada; Jocelyne Chiquette and Louise Provencher, Clinique des Maladies du sein Deschênes-Fabia, Hôpital du saint-Sacrement, Quebec City, Quebec, Canada; Rachel Laframboise, Service de Médecine Génétique, CHUQ, Pavillon CHUL, Quebec City, Quebec, Canada; Jean Lépine, Centre Hospitalier Régional de Rimouski, Rimouski, Quebec, Canada; Bernard Lespérance and Roxane Pichette, Service d'hématologie, Hôpital du Sacré-Coeur, Montréal, Quebec, Canada; Marie Plante, Service de Gynécologie, CHUQ, L'Hôtel-Dieu de Québec, Quebec City, Quebec, Canada; and Patricia Voyer, Clinique des maladies du sein, Carrefour de Santé de Jonquière, Jonquière, Quebec, Canada.

MAYO collaborators: Fergus J. Couch, Noralane Lindor, Linda Wadum, Kiley Johnson, Jennifer Mentlick, Janet Olson, Mayo Clinic College of Medicine, Rochester, MN.

### Acknowledgments

We thank Jennifer Scott for assistance with preparation of the manuscript; Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the Clinical Follow-up Study for their contributions to this resource; and the many families who contribute to kConFab. The LUMC team thanks K. Kroeze-Jansema for technical assistance. The HEBCS team thanks Drs. Carl Blomqvist and Kirsimari Aaltonen, as well as Anitta Tamminen, M.Sc., for their kind help.

### References

- Ewart-Toland A, Briassouli P, de Koning JP, et al. Identification of Stk6/STK15 as a candidate low-penetrance tumor-susceptibility gene in mouse and human. *Nat Genet* 2003;34:403-12.
- Zhang D, Hirota T, Marumoto T, et al. Cre-loxP-controlled periodic Aurora-A overexpression induces mitotic abnormalities and hyperplasia in mammary glands of mouse models. *Oncogene* 2004;23:8720-30.
- Egan KM, Newcomb PA, Ambrosone CB, et al. STK15 polymorphism and breast cancer risk in a population-based study. *Carcinogenesis* 2004;25:2149-53.
- Sun T, Miao X, Wang J, et al. Functional Phe311le polymorphism in Aurora A and risk of breast carcinoma. *Carcinogenesis* 2004;25:2225-30.
- Lo YL, Yu JC, Chen ST, et al. Breast cancer risk associated with genotypic polymorphism of the mitosis-regulating gene Aurora-A/STK15/BTAK. *Int J Cancer* 2005;115:276-83.
- Ewart-Toland A, Dai Q, Gao YT, et al. Aurora-A/STK15 T+91A is a general low penetrance cancer susceptibility gene: a meta-analysis of multiple cancer types. *Carcinogenesis* 2005;26:1368-73.
- Cox DG, Hankinson SE, Hunter DJ. Polymorphisms of the AURKA (STK15/Aurora kinase) gene and breast cancer risk (United States). *Cancer Causes Control* 2006;17:81-3.
- Dai Q, Cai QY, Shu XO, et al. Synergistic effects of STK15 gene polymorphisms and endogenous estrogen exposure in the risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 2004;13:2065-70.
- Fletcher O, Johnson N, Palles C, et al. Inconsistent association between the STK15 F311 genetic polymorphism and breast cancer risk. *J Natl Cancer Inst* 2006;98:1014-8.
- Xu X, Weaver Z, Linke SP, et al. Centrosome amplification and a defective G<sub>2</sub>-M cell cycle checkpoint induce genetic instability in BRCA1 exon 11 isoform-deficient cells. *Mol Cell* 1999;3:389-95.
- Deng CX. Roles of BRCA1 in centrosome duplication. *Oncogene* 2002;21:6222-7.
- Wu K, Hinson SR, Ohashi A, et al. Functional evaluation and cancer risk assessment of BRCA2 unclassified variants. *Cancer Res* 2005;65:417-26.
- Goldgar DE, Easton DF, Deffenbaugh AM, Monteiro ANA, Tavtigian SV, Couch FJ. Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. *Am J Hum Genet* 2004;75:535-44.
- Chenevix-Trench G, Healey S, Lakhani S, et al. Genetic and histopathologic evaluation of BRCA1 and BRCA2 DNA sequence variants of unknown clinical significance. *Cancer Res* 2006;66:2019-27.
- Korn EL, Graubard BI, Midthune D. Time-to-event analysis of longitudinal follow-up of a survey: chose of the time-scale. *Am J Epidemiol* 1997;145:72-80.
- Antoniu AC, Goldgar DE, Andrieu N, et al. A weighted cohort approach for analysing factors modifying disease risks in carriers of high-risk susceptibility genes. *Genet Epidemiol* 2005;29:1-11.
- Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. *J Am Stat Assoc* 1989;84:1074-8.
- Ouchi M, Fujiuchi N, Sasai K, et al. BRCA1 phosphorylation by Aurora-A in the regulation of G<sub>2</sub> to M transition. *J Biol Chem* 2004;279:19643-8.
- Rebbeck TR, Lynch HT, Neuhausen SL, et al.; Prevention and Observation of Surgical End Points Study Group. Prophylactic oophorectomy in carriers of BRCA1 or BRCA2 mutations. *N Engl J Med* 2002;346:1616-22.