

الجمهورية الجزائرية الديمقراطية الشعبية  
République algérienne démocratique et populaire  
وزارة التعليم العالي والبحث العلمي  
Ministère de l'enseignement supérieur et de la recherche scientifique  
جامعة عين تموشنت بلحاج بوشعيب  
Université–Ain Temouchent- Belhadj Bouchaib  
Faculté des Sciences et de Technologie  
Département des Sciences de la Nature et de la Vie



Projet de Fin d'Etudes  
Pour l'obtention du diplôme de Master en: Biochimie  
Domaine : Sciences de la Nature et de la Vie  
Filière : Sciences Biologiques  
Spécialité : Biochimie  
Thème

**Etude *in silico* des mutations du gène *BRCA1* identifiées chez  
des patients cancéreux Algériens**

**Présenté Par :**

- 1) M. BENABID Abd El Hakim
- 2) M. ZENASNI Youcef

**Devant le jury composé de :**

Mme. BRIXI GORMAT Nassima MCB UAT.B.B (Ain Temouchent) Président  
M. BAKLI Mahfoud MCA UAT.B.B (Ain Temouchent) Examineur  
Mme. MOGHTIT Fatima Zohra MCB UAT.B.B (Ain Temouchent) Encadrant

*Année Universitaire 2020/2021*

# *Remerciements*

*Nous tenons tout d'abord à remercier Dieu le tout puissant et miséricordieux, qui nous a donné la force et la patience d'accomplir ce modeste travail.*

*Nos vifs remerciements vont aux membres du jury pour l'intérêt qu'ils ont porté à ce travail .*

*Nos remerciements les plus respectueux vont à **Mme.BRIXI GORMATNassima** qui a bien voulu nous faire l'honneur de présider le jury. Veuillez trouver ici Madame le témoignage de notre reconnaissance et notre respect le plus profond.*

*Nous remercions également **M.BAKLI Mahfoud** pour l'honneur qu'il nous a fait en acceptant d'examiner ce travail et de siéger parmi cet honorable jury. Veuillez accepter ce travail, en preuve de notre grand respect et de toutes nos profondes reconnaissances.*

*Nous remercions notre encadreur **Mme.MOGHTIT Fatima Zohra** qui nous a permis de réaliser ce travail et nous avoir fait bénéficier de son expérience, de ses précieux conseils, pour son soutien et son encouragement vraiment merci.*

*Nous tenons à exprimer nos sincères remerciements à tous les professeurs qui nous ont enseigné et qui par leurs compétences nous ont soutenu dans la poursuite de nos études.*

*N'oublions pas de remercier nos très chers parents, frères et sœurs, nos camarades et amis pour leur contribution, leur soutien, leurs encouragements et leur patience*

# *Dédicace*

*Je dédie ce mémoire :*

Premièrement, je dédie ce travail à mes parents, qui ont été la

cause de tout ce que j'ai accompli et ma sœur et mon

frère aussi à ma tante.

mes chères amis et

camarades Sarah , Salah , Fayçal ; zaki , Meriem ,

nour , Khalid et tous mes amis sans oublier mon

encadrant Madame Moghtit Fatima Zohra

et tous les enseignants qui m'ont aidé

Merci à tous.

**YOUCEF**

# *Dédicace*

*Je dédie ce mémoire à :*

A mes parents qui m'ont soutenu et encouragé durant ces années d'études. Qui trouvent ici le témoignage de ma profonde reconnaissance.

A Ceux qui ont partagé avec moi tous les moments d'émotion lors de la réalisation de ce travail.

A ma famille, mes proches et à ceux qui me donnent de l'amour et de la vivacité.

A tous mes amis qui m'ont toujours encouragé et à qui je souhaite plus de succès.

A tous ceux que j'aime.

**ABD EL HAKIM**

## Résumé

Le gène *BRCA1* est un gène suppresseur de tumeur et son altération moléculaire pourra conduire à un risque accru de cancer du sein et ovaire. Dans le but de contribuer à l'étude moléculaire, nous nous sommes intéressés à l'étude *in silico* des fonctionnalités de deux mutations du gène *BRCA1* identifiées chez des patients Algériens à savoir la mutation : c.122A>G et la mutation c.135G>C.

Ainsi, les effets délétères de deux mutations ont été prédites en utilisant un protocole d'étude *in silico* constituée de différents logiciels (I-Mutant 2.0 ; SIFT ; Polyphen-2 ; Align GVGD ; Project HOPE).

La mutation c.122A>G est prédite comme étant délétère altérant la protéine BRCA1 tandis que la substitution c.135G>C est prédite comme non délétère. En fait, les changements physicochimiques entre les acides aminés natifs et mutés perturbent la stabilité de la protéine et peuvent engendrer la perte d'interaction avec d'autres molécules. Nos résultats ont montré que ces différents logiciels sont cohérents et que leur combinaison pourrait augmenter les performances de prédiction de l'effet de ces mutations.

Cette étude nous a permis de mieux comprendre l'effet des mutations et a dévoilé l'importance des études de prédiction par des logiciels bio-informatiques

**Mots clés** : *BRCA1* ; *in silico* ; mutation, faux sens .

## *Abstract*

The *BRCA1* gene is a tumor suppressor gene and its molecular alteration may lead to an increased risk of breast and ovarian cancer. In order to contribute to the molecular study, we were interested in the *in silico* study of the functionalities of two mutations of the *BRCA1* gene identified in Algerian patients, namely, c.122A>G mutation and c135G>C mutation.

Thus, the deleterious effects of two mutations were predicted using an *in silico* study protocol consisting of different software (I-Mutant 2.0; SIFT; Polyphen-2; AlignGVGD; Project HOPE).

The c.122AG mutation is predicted to be deleterious altering the *BRCA1* protein while the c.135G>C substitution is predicted to be non-deleterious. Indeed, physicochemical changes between native and mutated amino acids disrupt the stability of the protein and can cause the loss of interaction with other molecules. Our results have shown that these different software are consistent and that their combination could increase the performance of predicting the effect of these mutations.

This study allowed us to better understand the effect of mutations and revealed the importance of prediction studies by bioinformatics software.

**Keywords:** *BRCA1*; *in silico*; mutation ; missense.

# الملخص

إن الجين *BRCA1* هو جين كابح للورم وقد يؤدي تغييره الجزيئي إلى زيادة خطر الإصابة بسرطان الثدي والمبيض. من أجل المساهمة في الدراسة الجزيئية ، قد اهتمنا في هذا البحث بدراسة (*in silico*) لوظائف طفرتين في الجين *BRCA1* المحدد عند مرضى جزائريين تسمى الطفرة *G > c.122A*؛ و الطفرة *C > c.135G*

وبالتالي ، تم التنبؤ بالآثار الضارة للطفرتين باستخدام بروتوكول دراسة *in silico* الذي يتكون من برامج مختلفة (I-Mutant 2.0 ؛ SIFT ؛ Polyphen-2 ؛ محاذاة GVGD ؛ Project HOPE).

من المتوقع أن تكون طفرة *G > c.122A* ضارة في تغيير بروتين *BRCA1* بينما من المتوقع أن يكون استبدال *C > c.135G* غير ضار. في الواقع ، تؤدي التغييرات الفيزيائية والكيميائية بين الأحماض الأمينية الأصلية والمتحورة إلى تعطيل استقرار البروتين ويمكن أن تؤدي إلى فقدان التفاعل مع الجزيئات الأخرى. كما أظهرت نتائجنا أن هذه البرامج المختلفة متناسقة وأن دمجها يمكن أن يزيد من أداء التنبؤ بتأثير هذه الطفرات.

كما سمحت لنا هذه الدراسة بفهم تأثير الطفرات بشكل أفضل وكشفت عن أهمية دراسات التنبؤ بواسطة برامج المعلومات الحيوية .

**الكلمات المفتاحية:** *BRCA1*, *in silico*, طفرة ,

## Table des matières

<b>Liste des figures</b> .....	I
<b>Liste des tableaux</b> .....	III
<b>Liste des abréviations</b> .....	IV
<b>Introduction</b> .....	2
<b>Revue Bibliographique</b> .....	4
I. Notion oncogène et anti anti-oncogène .....	4
II. Description du gène BRCA1 : .....	4
III. Structure de la protéine : .....	5
IV. Rôle et fonctionde la protéine.....	7
V. Implication des mutations du gène dans le développement des cancers : .....	9
VI. La bioinformatique : .....	10
VI-1 definition et historique : .....	10
VI-2 Les banques de données : .....	11
VI-3-Outils bio-informatiques d'analyse de séquences .....	13
<b>Matérielset méthodes</b> .....	16
I. Population d'étude : .....	16
II. Méthode : .....	16
II-1-Prédiction de l'effet de la mutation sur la stabilité de la protéine par le logiciel I-Mutant 2.0 .....	16

II-2- Prédiction de l'effet délétère des mutations exoniques sur la protéine par le logiciel Sorting Intolerant From Tolerant (SIFT) .....	17
II-3- Prédiction des effets des mutatonns par le logiciel Polymorphism Phenotyping-2 (PolyPhen-2).....	18
II-4- Prédiction des effets des mutations par Align Grantham Variation Grantham Deviation (Align-GVGD) .....	19
II-5- Project HOPE ( Have (y)Our Protein Explained): .....	20
<b>Résultats et discussion</b> .....	22
I. Résultats : .....	22
I-1- Analyse de la mutation exonique c.122A>G (p.His41Arg) .....	22
I-1-1-Résultat obtenu par le logiciel I-Mutant 2.0 .....	22
I-1-2- Résultats obtenus par SIFT : .....	22
I-1-3-Résultats obtenus par le logiciel POLYPHEN-2 .....	24
I-1-4- Résultats obtenus par le logiciel Align GVGD :.....	25
I-1-5- Comparaison des propriétés physicochimiques entre les résidus normaux et mutés Project HOPE.....	26
I-2-Analyse de la mutation exonique c.135G>C (p.Lys45Asn) .....	27
I-2-1-Résultat obtenu par le logiciel I-Mutant 2.0 .....	27
I.2.2. Résultats obtenus par SIFT : .....	27
I.2.3. Résultats obtenus par le logiciel POLYPHEN-2 .....	28
I.2.4. Résultats obtenus par le logiciel Align GVGD .....	29
I.2.5. Comparaison des propriétés physicochimiques entre les résidus normaux et mutés Project HOPE.....	30

II. Discussion.....	31
<b>Conclusion et perspective .....</b>	<b>34</b>
<b>Références bibliographiques .....</b>	<b>36</b>

## Liste des figures

Figure 1:Localisation du gène BRCA1(Brody et al., 1995).....	5
Figure 2:Représentation schématique des domaines fonctionnels de la protéine BRCA1.(Savage et al., 2014).....	6
Figure 3:Interface du logiciel I-Mutant2.0.....	17
Figure 4:Interface du logiciel SIFT.....	17
Figure 5:Interface du logiciel Poly Phen-2.....	19
Figure 6:Interface du logiciel Align-GVGD.....	20
Figure 7:Résultat obtenu par le logiciel I-Mutant 2.0 pour la mutation c.122A>G (p.His41Arg). .....	22
Figure 8:Résultat obtenu par le logiciel SIFT pour la mutation c.122A>G (p.His41Arg). .....	23
Figure 9: Résultat obtenu par le logiciel PolyPhen-2 pour la mutation c.122A>G. ....	24
Figure 10:Résultat obtenu par Align GVGD pour la mutation c.122A>G.....	25
Figure 11: Les structures schématiques de l’histidine et l’arginine .....	26
Figure 12: Résultat obtenu par le logiciel I-Mutant 2.0 pour la mutation c.135G>C (p.Lys45Asn).....	27
Figure 13: Résultat obtenu par le logiciel SIFT pour la mutation .....	28
Figure 14:Résultat obtenu par le logiciel PolyPhen-2 pour.....	28
Figure 15: Résultat obtenu par Align GVGD pour la mutation c.135G>C (p.Lys45Asn). ....	29
Figure 16: Les structures schématiques de la lysine et l’asparagine .....	30

## Liste des tableaux

**Tableau I** :Liste des mutations identifiées dans la population Algérienne----- 16

**Tableau II** : Les différentes propriétés physico-chimiques correspondant aux couleurs des aa obtenus par SIFT pour la mutation c.122A>G (p.His41Arg).-----23

**Tableau III** :Les propriétés physicochimiques des deux aa Histidine et Arginine. -----26

**Tableau IV** :Les propriétés physicochimiques des deux aa Lysine et Asparagine-----30

## Liste des abréviations

**Aa** : Acide aminée.

**ADN** : Acide Désoxyribose Nucléique.

**ADNc**: Acide Désoxyribose Nucléique complémentaire

**Align GVGD**: Align Grantham Variation Grantham Deviation.

**Arg**: arginine

**ARNm** : Acide ribonucléique messenger

**Asn**: asparagine

**ATM**: ataxiatelangiectaciamentated

**Brc1**: abbreviation de breast cancer 1

**BRCT**: brca1 carboxyl terminal

**CDS**: Coding DNA Sequence

**CTL**: c-15 terminal building protein interacting protein

**COE**: Centre d'Excellence

**Da**: Dalton..

**DDG**:DeltaDeltaG.

**His**:histidine

**KB**:Kilobase.

**KDa**:Kilodalton.

**Lys** :lysine

**NES**: nuclear export signal

**P** : Bras court du chromosome.

## *Liste des abréviations*

**P** : groupement phosphate.

**Pb** : paire de base.

**PDB**: protéin Data Bank.

**Project HOPE**: Have (y) Our Protein Explained

**PolyPhen-2**: Polymorphism Phenotyping-2

**SIFT**: Sorting Intolerant From Tolerant

**STG**: spécialité traditionnelle garantie

**PALB1**: partner and localizer of brca1

**q**: Bras longs du chromosome.

**RAD51**: radiation sensitivité anormale 51

**SCD**: serine cluster domain

**UV**: ultraviolet

**3D** : tridimensionnelle.

# *Introduction*

Le cancer est une pathologie caractérisée par la multiplication et la propagation anarchiques de cellules malignes initialement normales ayant subi un panel de mutations ou une instabilité génétique au sein d'un tissu normal et pouvant affecter d'autres organes, ce sont les métastases. Si la propagation n'est pas contrôlée, le cancer peut entraîner la mort.(Benyelles, 2018)

L'existence d'une composante multigénique du cancer a été démontrée par différentes études depuis longtemps et aide désormais à mieux comprendre les déterminants génétiques de l'apparition du cancer et leur interaction avec les facteurs de risque environnementaux.(Delort, 2007)

Les gènes considérés comme ayant une implication significative dans la mise en place du processus de cancérisation sont les gènes de réparation de l'ADN, les gènes du métabolisme des xénobiotiques, les oncogènes et les anti-oncogènes. S'agissant de ces derniers, de nombreuses études ont souligné l'existence d'une relation entre les altérations moléculaires de nombreux entre eux et la survenue des cancers. En effet, les altérations constitutionnelles du gène suppresseur du tumeur *BRCA1* sont à l'origine des formes héréditaires de cancers du sein et de l'ovaire.(Taste, 2007)

Bien que plusieurs études sont réalisées sur le gène *BRCA1* dans différentes populations, les données de la littérature sur le spectre des mutations de ce gène en Algérie sont pauvres en raison d'un manque d'études. (Cherbal, 2010)

Au cours de ce travail, en se référant aux données de la littérature, les mutations semblant spécifiques à la population Algérienne sont analysées afin de prédire leurs éventuels effets délétères en utilisant différents logiciels d'étude *in silico*(*Salhi et al., 2006*).

*Revue*

*Bibliographique*

## **I. Notion oncogène et anti anti-oncogène**

La transformation maligne résulte de l'accumulation d'événements mutationnels successifs: des mutations germinales ou acquises altèrent le fonctionnement normal de certains gènes impliqués dans le contrôle de la division et la différenciation cellulaire(Andrieu, Colonna, & Lévy, 1997).Toutefois, l'apparition d'anomalies chromosomiques constitue une étape fondamentale de la cancérogenèse (Boveri, 1929).

L'étude des cancers a permis d'identifier deux catégories de gènes: les anti-oncogènes et les proto-oncogènes. Ces derniers ont une action stimulatrice sur la division cellulaire, mais leur expression est régulée. Ils s'activent en oncogènes lorsqu'ils subissent des altérations sporadiques ou plus rarement constitutionnelles. Leur mode d'action est dominant car il suffit qu'un seul des deux allèles soit muté pour que leur action puisse s'exercer. Les anti-oncogènes ont une action inhibitrice de la division cellulaire. Leur mode de fonctionnement est récessif, puisque la perte de fonction nécessite l'inactivation des deux allèles de ces mêmes gènes. Ils subissent aussi bien des mutations germinales que des mutations somatiques(Zografos, 2002). En effet, la transformation cancéreuse résulte des mutations entraînant l'augmentation de l'activité des proto-oncogènes ou la diminution de l'activité des anti-oncogènes(Coppola, 2010).Le gène *BRCA1* (Breast Cancer susceptibility gene 1) fait partie de la famille des gènes suppresseurs des tumeurs dont les mutations inactivatrices prédisposent au cancer du sein et de l'ovaire.

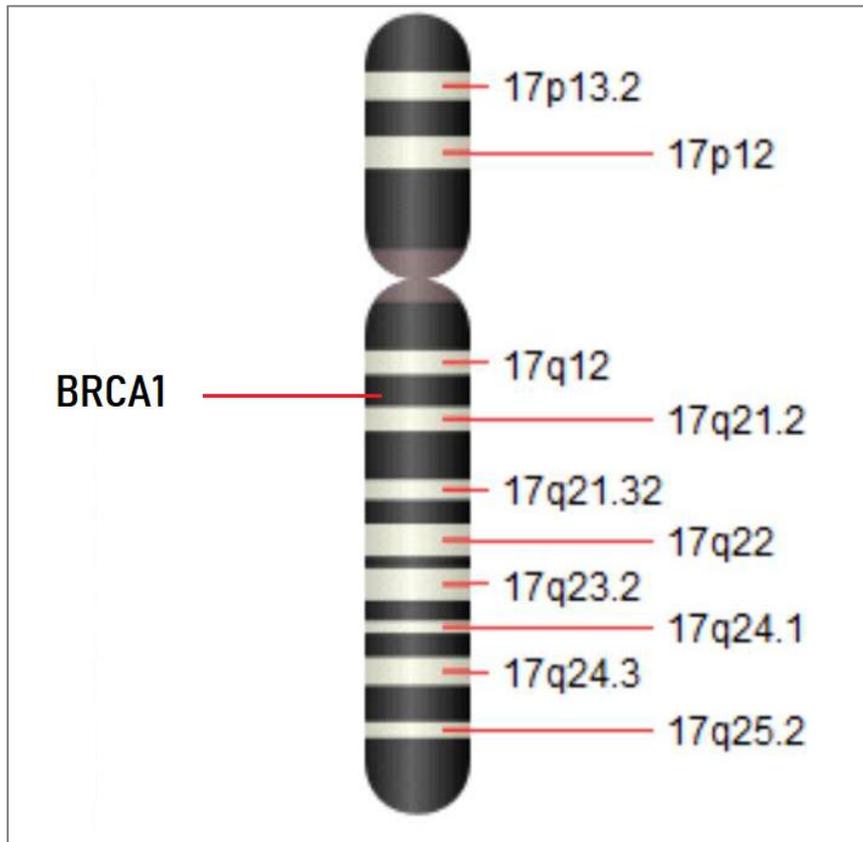
## **II. Description du gène BRCA1**

Le gène *BRCA1* est porté par le chromosome 17 situé dans la région 17q21 (Miki et al., 1994) (Figure 01).Il comporte 22 exons occupant environ 110kilobases (kb) sur l'ADN génomique, dont la taille varie de 40 paires de bases (pb) à 3425 pb. L'exon 11 représente à lui seul 3426 pb (60% de la séquence codante de *BRCA1*). Cette taille importante rend d'ailleurs son étude délicate. Les autres exons quant à eux sont de plus petites tailles. Parmi les exons, 22 sont codants et deux considérés comme non codants :

-**L'exon 1**: en amont du codon d'initiation de traduction ATG (situé dans l'exon 2). Il correspond à la majeure partie de la région 5'UTR du gène

—**L'exon 4** : identifié initialement comme un exon, il s'agit en fait d'une séquence Alu non codante. Il n'est d'ailleurs même pas représenté le plus souvent.

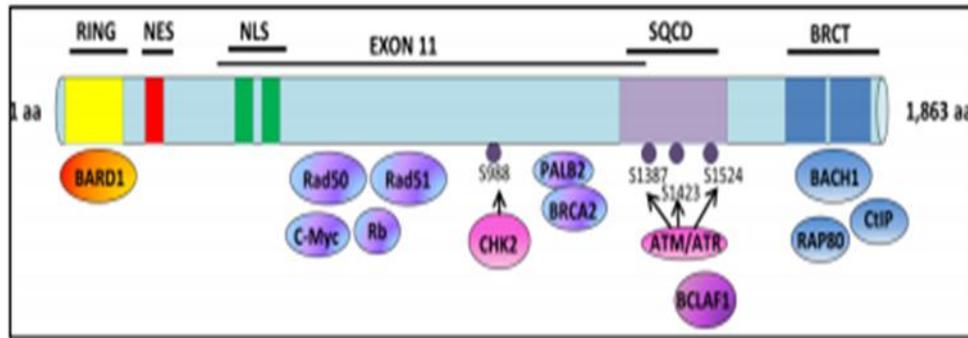
Les séquences introniques (de tailles variant de 403 pb à 9,2 kb) représentent 91% du gène et contiennent une forte quantité de séquences répétées de type Alu.(Ruiu, 2015)



**Figure 1:** Localisation du gène *BRCA1* (Brody et al., 1995)

### III. Structure de la protéine

La région codante du gène «Coding DNA Sequence (CDS)» s'étend sur 5,7kb et code pour une phosphoprotéine nucléaire appelée BRCA1 (UNIPROT :Q3B891) dont la masse moléculaire est de 190KiloDalton (kDa) et constituée de 1863 acides aminés (aa)



**Figure 2:** Représentation schématique des domaines fonctionnels de la protéine BRCA1. (Savage et al., 2014)

La protéine BRCA1 comprend différents domaines structuraux et fonctionnels suivants :

-Au niveau de l'extrémité N-terminale, on retrouve le domaine RING (Really Interesting New gene) : il s'agit d'un domaine en doigt de Zinc, qui grâce à sept résidus cystéine et un résidu histidine, se lie à deux ions  $Zn^{+2}$ , ce qui stabilise sa structure (Brzovic, Rajagopal, Hoyt, King, & Klevit, 2001; Meza, Brzovic, King, & Klevit, 1999). Il est responsable de l'activité E3-ubiquitine ligase de BRCA1, ainsi que l'interaction avec BARD1 (BRCA1 Associated RING Domain protein 1) (Brzovic et al., 2001). Cette extrémité comporte également deux séquences d'exportation nucléaire NES (Nuclear export signal) (Rodríguez & Henderson, 2000).

- La région centrale codée par les exons 11-13 représente 65 % de la séquence peptidique et comporte deux domaines NLS (Nuclear Localization sequences) qui interviennent dans le transport de BRCA1 du cytosol vers le noyau ((C.-F. Chen et al., 1996), ainsi que différents sites de liaison pour de nombreuses protéines (Clark, Rodriguez, Snyder, Hankins, & Boehning, 2012) : la protéine Rb (Retinoblastomaprotéine) qui a un rôle régulateur du cycle cellulaire, un facteur de transcription c-Myc et RAD50 et RAD51 impliqués dans la réparation de l'ADN.

Au sein de cette région, il existe également un domaine en super hélice (coiled-coil domain) qui s'associe avec PALB2 (Partner and localizer of BRCA2) et un site de phosphorylation par CHK2 au niveau de la sérine 988 (S988) (Y. Zhang et al., 2004).

-L'extrémité C-terminale présente, un domaine SCD (Serine Cluster Domain), dont une portion est localisée dans la région codée par l'exon 11-13 (Paul & Paul, 2014), contient de nombreux résidus sérineglutamine (SQ) ou thréonine-glutamine (TQ) qui peuvent être phosphorylés par différentes kinases suite aux lésions de l'ADN tel PIKKs(phosphatidylinositol 3 kinase-related kinases) ; ATM (ataxia-telangiectasia mutated) et ATR (ATM and Rad3-related) et dépendamment du résidu phosphorylé va être responsable de nombreuses fonctions de BRCA1(Traven & Heierhorst, 2005) et deux domaines BRCT (BRCA1 Carboxyl Terminal) en tandem (Williams, Green, & Glover, 2001).

#### **IV. Rôle et fonction de la protéine**

La protéine BRCA1 est exprimée de façon ubiquitaire chez l'être humain. Cette expression corrèle de près avec les degrés de prolifération cellulaire et varie au cours du cycle cellulaire avec des niveaux maximum à la transition G1/S(Sopik, Phelan, Cybulski, & Narod, 2015). C'est une protéine dite « caretaker » dont l'inactivation crée un état permissif dans lequel la cellule accumule des défauts cellulaires qui amènent une instabilité chromosomique.Elle est impliquée dans des fonctions essentielles et selon son partenaire elle joue un rôle de pivot à la fois dans la réparation de l'ADN, la régulation du cycle cellulaire, la transcription et l'ubiquitination(Hadj Bekkouche, 2018).

##### **✓ *BRCA1 et le remodelage de la chromatine***

La protéine BRCA1 joue un rôle dans le remodelage de la chromatine en se liant aux histones acétyle transférases (HAT) p300/CBP (Pao, Janknecht, Ruffner, Hunter, & Verma, 2000) et SWI/SNF (Mating type Switch/sucrose non-fermenting) (Harte et al., 2010).

L'activité de régulation translationnelle de BRCA1 semble s'effectuer à la fois par le biais de co-activateurs et co-répresseurs, mais aussi par le recrutement de protéines impliquées dans le remodelage de la chromatine.

##### **✓ *BRCA1 et la régulation du cycle cellulaire***

L'activité de BRCA1 dans la régulation de plusieurs points de restriction du cycle cellulaire est effectuée en partie par le contrôle transcriptionnel de gènes clés (exemple : *GADD45A*) et par son association à des protéines de régulation(Hadj Bekkouche, 2018).

BRCA1 forme des complexes qui régulent le cycle cellulaire : au sein du complexe BACH1 (BRCA1-associated C-terminal hélices) ou FANCI (Fanconi anemia group J

protéine) ou BRIP1 (BRCA1 interacting protéine C-terminal hélices 1) régule le point de contrôle G1/S (Kumaraswamy & Shiekhattar, 2007) et le point de contrôle G2/M (Yu, Chini, He, Mer, & Chen, 2003). BRCA1 et TOPBP1 (Topo isomérase DNA II Binding Protein 1) colocalisent dans la fourche de réplication durant la phase S après traitement par différents agents génotoxiques (exemple : hydroxyurée, UV, zeocin) (Makiniemi et al., 2001).

✓ *Activité ubiquitine ligase de BRCA1*

BRCA1 possède une activité ubiquitine ligase acquise par son interaction avec BARD1, sur différents substrats: les histones, l'ARN polymérase II, la tubuline gamma, le récepteur alpha de l'œstrogène (Boulton, 2006). De plus, Le complexe BRCA1- BARD1 par son activité ubiquitine ligase (E3 ligase) régule le point de contrôle G2/M (Shabbeer et al., 2013; Yu et al., 2003).

✓ *Régulation de la transcription*

BRCA1 intervient également dans la régulation de la transcription en se liant avec l'ARN polymérase II via l'interaction avec ARN Helicase (Anderson, Schlegel, Nakajima, Wolpin, & Parvin, 1998). BRCA1 a aussi un rôle Co-activateur de p53, en se liant à la région promotrice de nombreux gènes régulés par p53 favorisant ainsi leur transcription (H. Zhang et al., 1998).

✓ *Réparation de l'ADN*

Dans un premier temps elle intervient dans la détection des lésions de l'ADN et la transmission du signal d'activation; en effet la protéine BRCA1 est phosphorylée par les kinases ATM, ATR et CHK2 activées en réponse aux cassures double-brin induites par les radiations ionisantes, UV (Ultraviolets) etc. (Gatei et al., 2001; Traven & Heierhorst, 2005). Elle prend part à la recombinaison homologue à plusieurs niveaux; la protéine BRCA1 au sein du complexe BRCA1-MRN-CtIp, l'associant à MRN (MRE11/RAD50/NBS1) et CtIP (C-15 terminal building proteine inter acting proteine) participe à la résection de l'ADN, étape nécessaire à l'initiation de la recombinaison homologue (L. Chen, Nievera, Lee, & Wu, 2008) et par sa liaison avec RAD51 (Radiation sensibilité anormale 51) qui est une recombinase essentielle à toutes les phases de la recombinaison homologue (Cousineau, Abaji, & Belmaaza, 2005). En interagissant avec RAD50 et l'ADN dans le complexe protéique MRN,

BRCA1 intervient également dans la réparation par le NHEJ et régule négativement le processus alt-NHEJ, très mutagène (Zhuang et al., 2006).

En plus, de ces activités, BRCA1 possède d'autres fonctions cellulaires généralement moins bien connues.

## **V. Implication des mutations du gène dans le développement des cancers**

La protéine BRCA1 est impliquée dans une voie de réponse aux dommages à l'ADN qui est associée à l'activation de la recombinaison homologe et à la réparation des cassures à double brin, et donc au maintien de l'intégrité du génome (Cantor et al., 2001). Un rôle pour BRCA1 dans la transcription a été suggéré puisque BRCA1 est un composant de l'holoenzyme ARN polymérase II et possède des domaines d'activation de la transcription (Scully, Chen, Ochs, et al., 1997). De plus, le BRCA1 s'associe physiquement à la p53 et stimule l'activité transcriptionnelle des gènes impliqués dans l'apoptose et l'arrêt du cycle cellulaire (Derynck, Zhang, & Feng, 1998)

Le gène BRCA I est un gène impliqué dans les formes héréditaires de cancer du sein et de l'ovaire. Le risque dans la population générale de développer un cancer du sein est de 12% et celui d'un cancer de l'ovaire est de 1,4%. Ces risques sont augmentés chez une femme présentant une anomalie du gène *BRCA I*(Senechal & Mouret-Fourme) .

Le fait d'être porteur d'une mutation du *BRCAI* de la lignée germinale augmente considérablement le risque de développer un cancer du sein et/ou de l'ovaire. La mutation du *BRCAI* se produit dans environ 34 à 58 % des familles de cancers du sein et/ou de l'ovaire, et la majorité des mutations du *BRCAI* (environ 80 %) entraînent l'encodage d'une protéine tronquée. Il est intéressant de noter que les porteurs de mutation *BRCAI* atteints d'un cancer de l'ovaire avancé ont de meilleurs taux de survie à 5 ans que les cas non familiaux (77 mois contre 29 mois)(Rubin et al., 1996).

En effet, plus de 600 mutations ou variations de séquence ont été identifiées sur ce gène. une anomalie du gène *BRCA I* est retrouvée chez 3 à 5% des femmes présentant un cancer du sein. chez une femme porteuse, le risque de développer au cours de sa vie un cancer du sein est d'environ 65% et un cancer de l'ovaire d'environ 45%(Wojnarowicz, 2012).

De plus, une mutation de ce gène augmente également le risque de survenue d'un cancer de la prostate chez l'homme et celui du colon.(Wojnarowicz, 2012).

Les mutations les plus fréquentes sont : C4446T, 2953delCT et GlnC. Les mutations intragéniques somatiques du *BRCA1* sont rarement rapportées dans le COU, ou d'autres cancers. Bien que la mutation somatique du *BRCA1* ait rarement été signalée, la méthylation du *BRCA1* par le promoteur a été signalée dans 10 à 20 % des cancers du sein /ovaire (Scully, Chen, Plug, et al., 1997)

## **VI. La bioinformatique**

### **VI.1. Définition et historique**

La bioinformatique est une discipline scientifique récente qui englobe la collecte, l'organisation et l'analyse de données biologiques au moyen de l'outil informatique. Le terme bioinformatique regroupe, sous le même mot, 2 approches. La première approche correspond à la sensibilité plus informatique du domaine et consiste en l'élaboration d'algorithmes et le développement de programmes pour extraire l'information biologique ainsi que la création de banques de données. La seconde approche correspond à la bio-analyse dont le but est centré sur l'analyse de ces données et leur signification dans un contexte biologique. Ces 2 approches, bio-analytique et informatique, sont complémentaires et sont naturellement liées l'une à l'autre dans leur évolution. En effet, l'étude d'un système biologique complet n'est rendue possible que par l'existence d'outils adaptés et performants et la bonne compréhension des systèmes biologiques permet de mettre au point des programmes adaptés et efficaces (Blanc et al., 2007).

L'essor de la bioinformatique est dû à l'émergence et à la diffusion de l'informatique qui a permis de disposer d'ordinateurs puissants capables de stocker et de traiter un nombre croissant de données, mais également, son développement se fait grâce à la compréhension des objets biologiques majeurs (ADN, ARN et protéine) et l'avènement d'outils de prédictions (Blanchet, 1999). L'évolution conjointe de ses 2 approches de la bioinformatique en fait actuellement une discipline indispensable à la compréhension du vivant (Glass et al., 2009).

Initialement, la bioinformatique consistait à créer et à maintenir des bases de données pour le stockage de l'information biologique comme les séquences nucléotidiques et d'acides aminés. Le passage de la recherche en biologie de l'étude d'un seul gène à l'étude de génomes et de protéomes d'organismes complets, a nécessité des moyens de calculs de plus

en plus importants, le rassemblement des banques de données existantes comme GenBank/EMBL/DDBJ, le développement de programmes performants (BLAST, FASTA...) et le développement de nouvelles techniques (les puces à ADN,...). L'analyse de cette grande quantité de données n'est rendue possible que par l'utilisation de la bioinformatique. En effet, la bioinformatique est un élément essentiel pour le passage de la biologie à l'ère post génomique et aux études à haut débit (Aube et al., 2004).

## **VI.2. Les banques de données**

La nécessité d'accès rapide et facile à l'information a toujours été une demande de la part des biologistes et un moteur de développement de la bioinformatique, ce besoin s'est fait encore plus pressant avec l'évolution des techniques du séquençage. Cet accès rapide et organisé se traduit par la création de banques de données (Barbier & Rasmont, 2003).

C'est au début des années 1980s, que les premières banques nucléiques sont apparues. Avant, il existait uniquement des banques de séquences protéiques et de structures tridimensionnelles (Bunnell, González, & Stein, 2003).

Les banques de données sont de deux types : celles qui correspondent à une collecte la plus exhaustive possible, offrant finalement un ensemble hétérogène d'informations (banques généralistes) et celles qui correspondent à des données plus homogènes établies autour d'une thématique, offrant une valeur ajoutée à partir d'une technique particulière ou d'un intérêt suscité par un groupe d'individus (banques spécialisées).

Les caractéristiques des banques de données sont : l'unification, la standardisation et l'interconnexion (Bunnell et al., 2003)

- **Les banques nucléiques :**

Les trois grandes banques généralistes de séquences nucléotidiques, gratuitement accessibles sont :

EMBL (<http://www.ebi.ac.uk/embl>) en Europe,

Gen Bank (<http://www.ncbi.nlm.nih.gov/genbank/>) aux Etats-Unis

la DDBJ (<http://www.ddbj.nig.ac.jp/>) au Japon.

L'EMBL, a été créée en 1980 à Heidelberg (Allemagne) et est maintenue depuis 1994 par l'EBI (*European Bioinformatic Institute*), à Cambridge (Angleterre) (Stoesser, Sterk, Tuli, Stoehr, & Cameron, 1997).

GenBank, a été mise en place en 1979 au LANL (*Los Alamos National Laboratory*), à Los Alamos (Etats-Unis). Depuis 1992, elle est maintenue au NCBI (*National Center for Biotechnology Information*), à Bethesda (Etats-Unis) (Benson et al., 2002).

Enfin, la DDBJ, a débuté son activité en 1987, cette banque a été créée, et est toujours maintenue au NIG (*National Institute of Genetics*), à Mishima (Japon) (Ding et al., 2008).

EMBL, GenBank et la DDBJ sont distribuées par les centres sous la forme d'un ensemble de fichiers plats regroupant les séquences en fonction de critères taxonomiques (procaryotes, Virus, Primates, ...) ou de leur origine (EST, STS).

A l'intérieur de ces fichiers, chaque séquence est contenue dans une structure appelée "entrée", une entrée comprenant une quantité variable d'informations liée à la séquence considérée.

Les informations en question sont introduites au niveau de "champs" bien définis. Le format de stockage utilisé peut être différent d'une banque à l'autre.

- **Les banques protéiques :**

La première banque informatique de séquence de protéines date de la fin des années 1970s avec la création de la NBRF/PIR (*Protein Identification Resource*) grâce au GUMC (*Georgetown University Medical Center*) et de la NBRF (*National Biomedical Research Foundation*), résultant de la version imprimée de l'Atlas of Protein Sequence and Structure de Margaret Dayhoff (USA). En 1986, une version reformatée et corrigée de la NBRF/PIR nommée Swiss-Prot voit le jour à l'EMBL (Bairoch & Apweiler, 2000).

Depuis plus de 30 ans, l'importance de ces banques de séquences n'a fait que croître et elles constituent désormais la pierre angulaire de la compréhension des fonctions des protéines dans la vie cellulaire (Boyer, 2015).

Les trois banques de séquences protéiques majeures (Swiss-Prot, TrEMBL et PIR) sont réunies en une seule et unique banque qui est distribuée sous le nom d'UniProt (*United Protein Databases*) (Mulder et al., 2005). De plus, il existe plusieurs banques protéiques spécialisées, dont la plus célèbre la banque PDB.

**PDB:**

La PDB (Protein Data Bank), est la principale banque internationale de structures tridimensionnelles. Cette banque a été fondée en 1971 par le BNL (*Brookhaven National Laboratory*) et contenait 7 structures. Depuis 1998, elle est sous la tutelle du RSCB (*ResearchCollaboratory for Structural Bioinformatics*) (Harkness, Sussman, Davies-Jones, Greenwood, & Woodroffe, 2003)

Les structures de protéines constituent l'essentiel des entrées de la PDB avec plus de 32500 entrées en 2006, mais il existe également des structures de molécules d'ARN et d'ADN, de complexes protéines-acides nucléiques (Kouranov et al., 2006) Ces structures sont déterminées expérimentalement par cristallographie aux rayons X, par RMN (*Résonance Magnétique Nucléaire*) ou encore microscopie électronique. On peut noter que plus de 90% des structures déposées dans la PDB sont résolues par la technique de cristallographie aux rayons X (Berman et al., 2000).

Les entrées de la banque comprennent des informations sur les structures primaires et secondaires des molécules considérées, les coordonnées atomiques, souvent des détails des expériences (conditions de cristallisation, empilement cristallin, statistiques d'affinement, etc.) ainsi que des références bibliographiques(Sussman et al., 1998).

**VI.3.Outils bio-informatiques d'analyse de séquences :**

Les outils bioinformatiques disponibles offrent plusieurs alternatives: comparer la séquence d'intérêt à celles qui sont connues, rechercher des motifs particuliers ou signatures fonctionnelles permettant de rattacher la protéine à une famille fonctionnelle. Ensuite, il peut être judicieux d'aligner ces protéines par le calcul d'un alignement multiple, offrant une base de travail pour l'étude expérimentale approfondie de sa fonction. D'autres outils bioinformatiques permettent de prédire les régions hydrophobes (trans-membranaires par exemple), des régions accessibles (antigènes potentiel pour la synthèse d'anticorps de détection), des régions d'interaction avec d'autres molécules comme les motifs d'interaction avec l'ADN. Ainsi la bioinformatique offre un panel d'outils dans le domaine de l'analyse de séquence des protéines, qui judicieusement combinés sont des moyens efficaces et rapides pour l'étude d'une protéine dont seule la séquence est connue. Néanmoins, les déductions obtenues *in silico* doivent être vérifiées par des approches expérimentales *in vitro* et/*in vivo* (DuBrin, 2002).

Les variations de séquences au niveau des gènes sont, le plus souvent, responsables de troubles fonctionnels des protéines concernées menant à l'apparition de nombreuses maladies. Ces variations peuvent se produire aussi bien au niveau des régions géniques codantes (exons), qu'au niveau des introns et de ce fait avoir un effet sur le processus d'épissage alternatif, ce qui se traduit par diverses conséquences sur la structure et/ou la fonction de la protéine. De nombreux logiciels ont été développés afin de prédire les différents effets et impacts de ces mutations délétères. La modélisation tridimensionnelle (3D) vient s'ajouter à ces nombreux logiciels de prédiction et permet une meilleure appréciation des défauts moléculaires (Lambert, 2003).

*Population d'étude*

*et méthodes*

## **I. Population d'étude**

L'étude *in silico* a porté sur deux mutations identifiées chez des patients Algériens atteints du cancer du sein (Salhi et al., 2006). Les mutations sont résumées dans le tableau I.

**Tableau I:** Liste des mutations identifiées dans la population Algérienne  
(Hammes et al., 2006)

<b>Exon/Intron</b>	<b>Type de mutation</b>	<b>DNAc</b>	<b>Protéine</b>
Exon 2	Faux –sens	c.122A>G	p.His41Arg
Exon 2	Faux –sens	c.135G>C	p.Lys45Asn

## **II. Méthode**

Les variations de bases au niveau des gènes sont, le plus souvent, responsables de troubles fonctionnels des protéines concernées menant à l'apparition de plusieurs maladies. Ces variations peuvent se produire au niveau d'une région codante (exon) ou au niveau des régions non codantes introns et peuvent donc avoir des effets sur le processus d'épissage alternatif et/ou sur la structure/fonction de la protéine. Plusieurs logiciels destinés à prédire les différents effets délétères des mutations ont été développés.

Les effets délétères des mutations c.122A>G et c.135G>C n'ont jamais été décrits, ni fait l'objet d'études. Ces mutations ont été explorées en utilisant quatre logiciels bio-informatiques: SIFT, PolyPhen-2, Align-GVGD, Mutation Taster et Project HOPE.

### **II.1. Prédiction de l'effet de la mutation sur la stabilité de la protéine par le logiciel I-Mutant 2.0**

L'effet stabilisant ou déstabilisant des mutations a été évalué par le logiciel I-Mutant 2.0 (Figure 15). Ce dernier a été conçu en 2005 par Capriati et al, il est basé sur une approche de type Support Vector Machine (SVM) qui estime la différence d'énergie libre notée  $\Delta\Delta G$  entre la protéine sauvage et mutante. La prédiction se base sur la structure tertiaire ou de la séquence de la protéine (Capriotti, Fariselli, Calabrese, & Casadio, 2005). Une valeur  $\Delta\Delta G > 0$

implique une augmentation de la stabilité de la protéine et à l'inverse une valeur  $DGG < 0$  suggère un effet déstabilisant de la mutation.

**Figure 3:**Interface du logiciel I-Mutant2.0

(<http://folding.biofold.org/cgi-bin/i-mutant2.0.cgi>).

## II.2. Prédiction de l'effet délétère des mutations exoniques sur la protéine par le logiciel SortingIntolerantFromTolerant (SIFT)

SortingIntolerantFromTolerant (SIFT) est un logiciel gratuit accessible en ligne (Figure 12). Ce logiciel s'appuie sur l'homologie des séquences nucléotidiques ou protéiques pour prédire les effets délétères des mutations (P. C. Ng & S. Henikoff, 2003). Il permet de trier la substitution en fonction du degré de conservation d'un acide aminé dit intolérant par rapport au tolérant et prédit l'effet phénotypique de cette substitution sur la protéine (P. C. Ng & S. Henikoff, 2003). SIFT permet de prédire si une substitution est tolérée à chaque position de la séquence (Pauline C Ng & Steven Henikoff, 2003).

Human Genome DB	Tool Description
SIFT/PROVEAN Human SNPs	Get SIFT and PROVEAN predictions for SNPs and indels (Ensembl 66) (Sample format)
SIFT Human SNPs	Get SIFT predictions for nonsynonymous SNPs (Ensembl 63) (Sample format)

**Figure 4:**Interface du logiciel SIFT

(<http://sift.jcvi.org/>)

Une substitution est dite tolérante car elle est considérée comme n'ayant pas d'effet délétère sur la fonction de la protéine. A l'inverse, elle est dite intolérante lorsque celle-ci semble avoir un impact entraînant la perte partielle ou totale de la fonction de la protéine (P. C. Ng & S. Henikoff, 2003). Afin de prédire si le changement d'acide aminé aura un impact sur la protéine SIFT prend en compte la position et la nature de l'acide aminé.

Le logiciel SIFT se base sur le postulat que toute évolution d'une protéine est étroitement corrélée à sa fonction. Cela implique que les parties fonctionnelles soumises à de forte pression de sélection sont conservées. Il suppose que les acides aminés importants seront conservés dans la famille de protéines, et les changements donc à des positions bien conservées ont tendance à être prédits comme délétères. La forte probabilité d'avoir un effet délétère provient d'une substitution d'un acide aminé survenant à une position fortement conservée d'une famille de protéine et inversement, une faible probabilité suggère que la substitution siège en une position peu conservée (Luu, 2012).

La probabilité pondérée en fonction du nombre de séquences alignées est calculée en comparant la matrice de mutation issue de l'alignement multiple avec celles prédites par block substitution Matrix 62 (BLOSUM 62). SIFT calcule des probabilités normalisées pour toutes les substitutions possibles à chaque position. Les probabilités inférieures à un seuil prédéfini (<0.05) seront considérées comme délétères (Awad, 2014).

En réalité, trois différents scores basés sur la similarité entre deux acides aminés sont obtenus :

- Intolérant (Intolérant): [0.00-0.05].
- Potentiallytolérant (potentiellement tolérant): [0.051-0.20].
- Tolérant (tolérant): [0.201-1.00].

### **II.3. Prédiction des effets des mutations par le logiciel Polymorphism Phenotyping-2 (PolyPhen-2)**

Polymorphism Phenotyping-2 (PolyPhen-2) est un logiciel gratuit accessible sur le lien: <http://genetics.bwh.harvard.edu/pph2/> (Figure 6). Il a été publié par Shamil Sunyaev et Vasily Ramensky en 2000 (Sunyaev, Ramensky, & Bork, 2000).

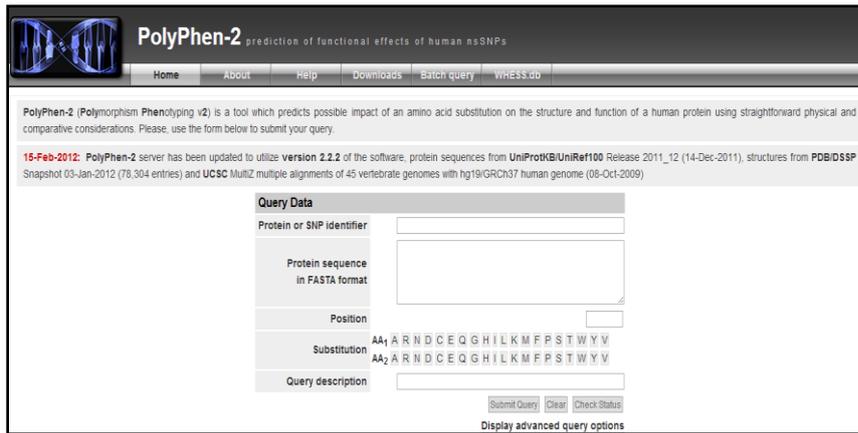


Figure 5: Interface du logiciel Poly Phen-2

<http://genetics.bwh.harvard.edu/pph2/>

Ce logiciel prédit l'impact d'une substitution d'un acide aminé sur la structure et la fonction d'une protéine en se basant sur trois informations:

- L'alignement multiple.
- Les informations structurales provenant des banques de données de structure (PDB).
- Les propriétés physico-chimiques des acides aminés.

Les prédictions sont classées en 3 niveaux: «*probablydamaging*» (probablement délétère), «*possiblydamaging*» (potentiellement délétère) et «*benign*» bénigne.

La prédiction est dite bénigne lorsque la fonction de la protéine ne semble pas affectée par cette variation.

A l'inverse, une prédiction est dite délétère lorsque la protéine a une forte probabilité d'être altérée par la variation (Sunyaev et al., 2000).

#### II.4. Prédiction des effets des mutations par Align Grantham Variation Grantham Deviation (Align-GVGD)

Align Grantham Variation Grantham Deviation (Align-GVGD) est un logiciel gratuit accessible en ligne (<http://agvgd.iarc.fr/>), publié par Sean V. Tavtigia en 2008 (Tavtigian, Byrnes, Goldgar, & Thomas, 2008).

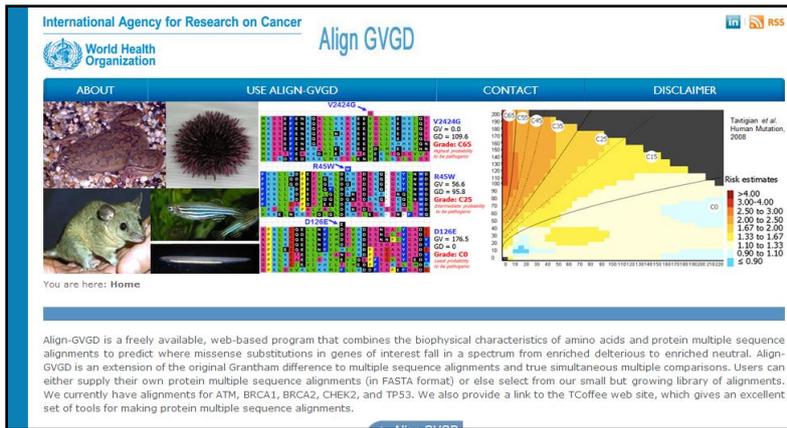


Figure 6: Interface du logiciel Align-GVGD

[\(http://agvgd.iarc.fr/\)](http://agvgd.iarc.fr/).

Ce logiciel combine alignement de séquences multiples et les caractéristiques biophysiques des acides aminés qui s'appuient sur la distance de Grantham (Figure 14). En effet, la distance de Grantham calcule l'écart physico-chimique entre deux acides aminés. Plus cette distance est élevée, plus les acides aminés en cause sont différents. Ce logiciel prédit l'impact des mutations faux-sens sur la fonction de la protéine.

Les résultats sont établis sous forme de classes [C0 à C65]. Les classes [C45 à C65] signifient que la mutation affecte la fonction, elle est dite «*more likely to affect the function*». A l'inverse, les classes [C0 à C35] indiquent que la substitution n'affecte pas la fonction, elle est dite «*less likely to affect the function*» (Mathe et al., 2006)

## II.5. Project HOPE ( Have (y)Our ProteinExplained)

Project HOPE est un programme automatique gratuit accessible sur le lien suivant (<http://www.cmbi.ru.nl/hope/input?3>). Il analyse les effets structuraux et fonctionnels des mutations ponctuelles. Le système HOPE utilise toutes les informations collectées combinées avec des propriétés connues du type sauvage et muté d'un aa, comme la taille, la charge et l'hydrophobicité pour prédire l'effet de la mutation sur la structure et la fonction de la protéine (Elshaikh, Ismaiel, Osman, & Shokri, 2016).

*Résultats*

*et*

*discussion*

## I. Résultats

### I.1. Analyse de la mutation exonique c.122A>G (p.His41Arg)

#### I.1.1. Résultat obtenu par le logiciel I-Mutant 2.0

L'analyse des effets de la mutation faux-sens c.122A>G sur la stabilité de la protéine BRCA1 a été prédite par le logiciel I-Mutant 2.0. Le résultat obtenu par ce logiciel est représenté sur la figure 7 .

```
*****
**                                     **
**                               I-Mutant v2.0                               **
**       Predictor of Protein Stability Changes upon Mutations               **
**                                     **
*****

SEQ File: fileseq.txt

Position  WT  NEW  Stability  RI   pH   T
         41  H   R   Decrease  5   7.0  25

WT: Aminoacid in Wild-Type Protein
NEW: New Aminoacid after Mutation
RI:  Reliability Index
T:   Temperature in Celsius degrees
pH:  -log[H+]

*****
```

**Figure 7:** Résultat obtenu par le logiciel I-Mutant 2.0 pour la mutation c.122A>G (p.His41Arg).

Le logiciel I-Mutant 2.0 a permis de calculer la différence d'énergie (DDG) entre la protéine normale et mutée. La valeur DDG calculée pour cette mutation est négative cela indique que la mutation testée est considérée comme déstabilisante pour la structure de la protéine BRCA1.

#### I.1.2. Résultats obtenus par SIFT

Le résultat obtenu par le logiciel SIFT est représenté sur la figure 8. Ce résultat est présenté sous forme de deux colonnes. La colonne de gauche regroupe les aa qui sont prédits intolérants pour une position donnée.

A l'inverse, celle de droite regroupe les aa qui sont prédits tolérants. Les aa sont colorés en fonction de leurs propriétés physico-chimiques (Tableau II).

w h y f m q r d n c e k l v l g S T A	34P	0.94	P
h d w n q g e p c s k y t a R F M L I	35V	0.94	V
w y h f c g e d k p n Q a V R M I L T	36S	0.94	S
g w y h d n r f q e k c m s P A L I V	37T	0.94	T
w c f y m i v L n t G H R A P S E D	38K	0.94	Q K
y w v t s r q p n m l k i h g f e d a	39C	0.94	C
w m i f c v y R h p t k L Q S A E N	40D	0.94	G D
y w v t s r q p n m l k i g f e d c a	41H	0.94	H
w h g m e q d f k p C Y a L R N V T S	42I	0.94	I
h n k r q d g e p c t s a m v i w L Y	43F	0.94	F
y w v t s r q p n m l k i h g f e d a	44C	0.94	C
c d f m i W v g p n e A Y T H S Q L R	45K	0.94	K
c w Q t M P N h I V D G E y L K S A R	46F	0.94	F
y w v t s r q p n m l k i h g f e d a	47C	0.94	C

**Figure 8:**Résultat obtenu par le logiciel SIFT pour la mutation c.122A>G (p.His41Arg).

**TableauII:**Les différentes propriétés physico-chimiques correspondant aux couleurs des aa obtenus par SIFT pour la mutation c.122A>G (p.His41Arg).

Couleurs des aa	Propriété physico-chimique
Rouge	Basique
Vert	Polaire non chargé
Noire	Non polaire
Bleu	Acide

L'analyse de la mutation montre qu'aucun acide aminé ne peut remplacer l'Histidine à la position 41. Le remplacement de l'Histidine par l'Arginine est intolérant ce qui signifie que la mutation pourrait avoir un impact sur la protéine.

1.1.3. Résultats obtenus par le logiciel POLYPHEN-2

La prédiction de l'impact de la mutation c.122A>G(p.His41Arg) sur la protéine pRb par le logiciel PolyPhen-2 nous a permis d'avoir le résultat présenté sur la figure 9.

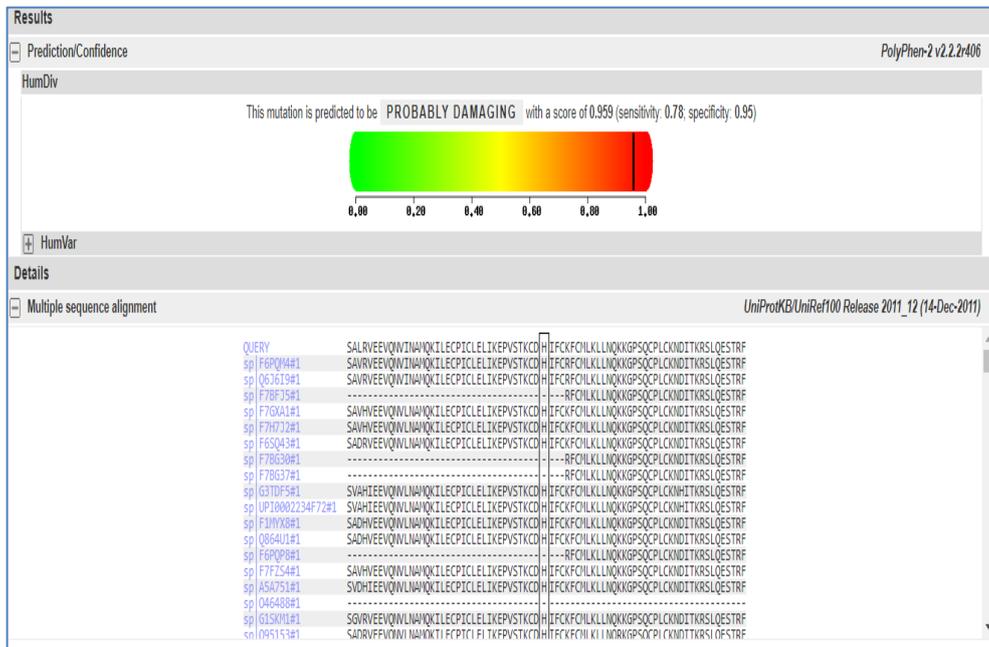


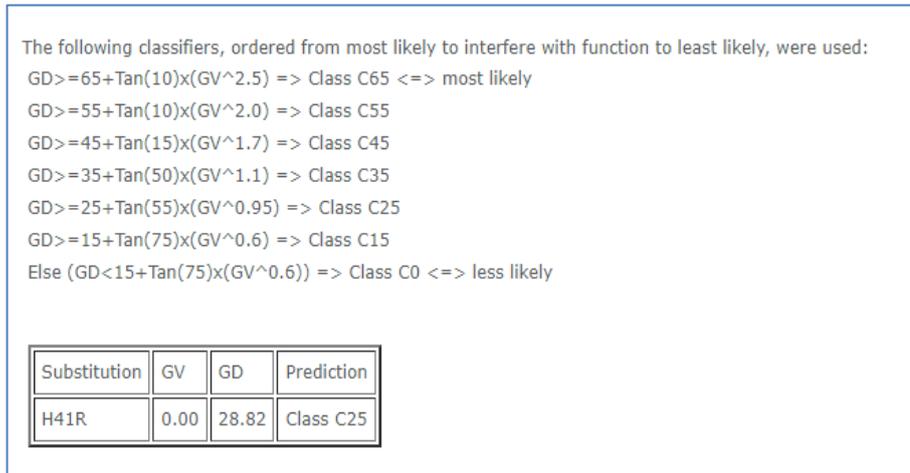
Figure 9: Résultat obtenu par le logiciel PolyPhen-2 pour la mutation c.122A>G.

La mutation c.122A>G a été prédite par le logiciel PolyPhen-2 comme étant probablement délétères à la structure protéique (Probablydamaging). Cela signifie que cette dernière a un effet sur la structure et la fonction protéique.

Ce logiciel a permis également d'obtenir le résultat d'alignement multiple par comparaison de la protéine BRCA1 avec toutes les espèces enregistrées sur UniProKB. A partir de l'alignement multiple, il ressort que cette mutation siège dans une région conservée entre plusieurs espèces. Les régions conservées sont considérées comme les régions les plus fonctionnelles des protéines.

**I.1.4. Résultats obtenus par le logiciel Align GVGD**

Le logiciel Align GVGD a également été utilisé afin de prédire les effets de la mutation testée sur la protéine BRCA1. Le résultat obtenu par ce logiciel est présenté sur la figure 10.



**Figure 10:** Résultat obtenu par Align GVGD pour la mutation c.122A>G.

L'analyse de la mutation c.122A>G par Align GVGD a permis de calculer les valeurs GV et GD. La valeur de la variation de Grantham GV est égale à 0.00, ce qui signifie que cette substitution siège au niveau d'une région fortement conservée entre les espèces analysées.

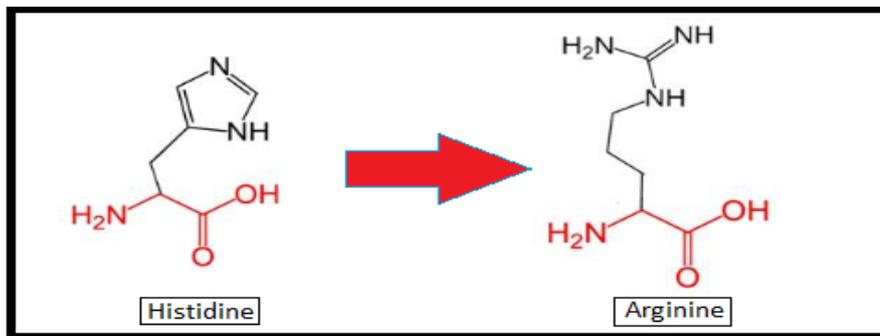
D'autre part, ce même logiciel nous a permis de calculer l'écart de Grantham GD entre les aa sauvages et mutés. La valeur GD pour la mutation c.122A>G est égale à 28.82 ce qui signifie que les propriétés physico-chimiques entre l'histidine et l'arginine sont modérément conservées.

Ce logiciel a permis de prédire que cette dernière correspond à la classe C25 indiquant que la substitution n'affecte pas la fonction de la protéine.

**I.1.5. Comparaison des propriétés physicochimiques entre les résidus normaux et mutés (Project HOPE)**

Pour rappel, la mutation c.122A>G est responsable de la substitution de l'aa histidine par l'aa arginine. Les structures de ces deux aa sont représentées dans la figure 11. Tandis que les propriétés physicochimiques de chaque résidu sont résumées dans le tableau. La figure montre les structures schématiques de l'aa normal (à gauche) et le muté (à droite). La partie commune (squelette) entre les deux aa est de couleur rouge. Tandis que la chaîne latérale qui est unique pour chaque résidu est de couleur noire.

Ces deux aa présentent des propriétés chimiques particulières en raison de la nature de leurs chaînes latérales.



**Figure 11:** Les structures schématiques de l'histidine et l'arginine

**Tableau III:** Les propriétés physicochimiques des deux aa histidine et arginine.

Résidu	Taille	Charge
histidine	155Da	Positive
arginine	174Da	Positive

Les différences entre ces deux aa résident dans leurs tailles (tableau III). Cette différence peut entraîner un changement de conformation de la protéine ainsi qu'une perte d'interaction avec d'autres molécules ou résidus au sein du BRCA1.

**I.2. Analyse de la mutation exonique c.135G>C (p.Lys45Asn) :**

***I.2.1. Résultat obtenu par le logiciel I-Mutant 2.0***

L'analyse des effets de la mutation faux-sens c.135G>C sur la stabilité de la protéine BRCA1 a été prédite par le logiciel I-Mutant 2.0. Le résultat obtenu par ce logiciel est représenté sur la figure.

```
*****
**                                     **
**                               I-Mutant v2.0                               **
**       Predictor of Protein Stability Changes upon Mutations               **
**                                     **
*****

SEQ File: fileseq.txt

Position  WT  NEW  DDG  pH  T
         45  K   N   0.29  7.0  25

WT:  Aminoacid in Wild-Type Protein
NEW:  New Aminoacid after Mutation
DDG:  DG(NewProtein)-DG(WildType) in Kcal/mol
      DDG<0: Decrease Stability
      DDG>0: Increase Stability
T:    Temperature in Celsius degrees
pH:   -log[H+]
```

**Figure 12:** Résultat obtenu par le logiciel I-Mutant 2.0 pour la mutation c.135G>C (p.Lys45Asn)

Le logiciel I-Mutant 2.0 a permis de calculer la différence d'énergie (DDG) entre la protéine normale et mutée. La valeur DDG calculée pour cette mutation est égale à 0,29. Étant donné que cette valeur est positive, donc la mutation testée n'affecte pas la stabilité de la structure de la protéine BRCA1.

***I.2.2. Résultats obtenus par SIFT***

Les résultats de l'analyse de la mutation c.135G>C (p.Lys45Asn) montrent que la substitution de la lysine en asparagine en position 45 est intolérante (Intolerated). Ce résultat signifie que cette dernière pourrait avoir un impact sur la protéine (Figure 13).

y w v t s r q p n m l k i h g f e d a	39C	0.94	C
w m i f c v y R h p t k L Q S A E N	40D	0.94	G D
y w v t s r q p n m l k i g f e d c a	41H	0.94	H
w h g m e q d f k p C Y a L R N V T S	42I	0.94	I
h n k r q d g e p c t s a m v i w L Y	43F	0.94	F
y w v t s r q p n m l k i h g f e d a	44C	0.94	C
c d f m i W v g p n e A Y T H S Q L R	45K	0.94	K
c w Q t M P N h I V D G E y L K S A R	46F	0.94	F
y w v t s r q p n m l k i h g f e d a	47C	0.94	C
d g h n e c s w r k y p q a T v F L	48M	0.92	I M
w p m q n G K C a D H F E y V I R T S	49L	0.93	L
w c f y m i h v l p G N D A T Q S R E	50K	0.93	K
p d e q r M K N H C S T I W G A V Y F	51L	0.82	L

Figure 13: Résultat obtenu par le logiciel SIFT pour la mutation c.135G>C (p.Lys45Asn)

1.2.3. Résultats obtenus par le logiciel POLYPHEN-2

La prédiction de l'impact de la mutation c.135G>C (p.Lys45Asn) sur la protéine BRCA1 par le logiciel PolyPhen-2 nous a permis d'avoir le résultat présenté sur la figure 14.

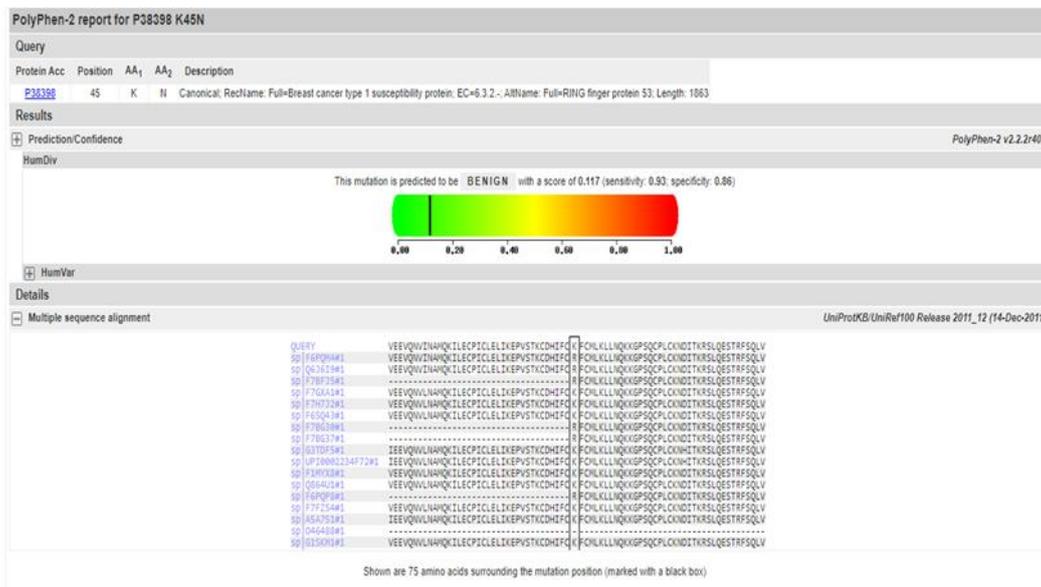
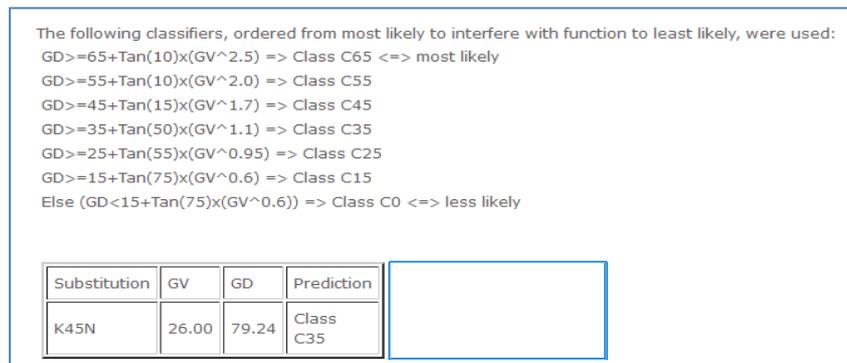


Figure 14: Résultat obtenu par le logiciel PolyPhen-2 pour la mutation c.135G>C (p.Lys45Asn)

La mutation c.135G>C (p.Lys45Asn) a été prédite par le logiciel PolyPhen-2 comme bénigne (*benign*). Cela signifie que cette dernière n'a pas d'effet sur la structure et la fonction protéique. Ce logiciel a permis également d'obtenir le résultat d'alignement multiple par comparaison de la protéine pR avec toutes les espèces enregistrées sur UniProKB, ce qui nous a permis de confirmer que la mutation ne siège pas dans une région conservée entre plusieurs espèces.

#### ***1.2.4. Résultats obtenus par le logiciel Align GVGD***

Le logiciel Align GVGD a également été utilisé afin de prédire les effets de la mutation testée sur la protéine BRCA1. Le résultat obtenu par ce logiciel est présenté sur la figure 15.



**Figure 15:** Résultat obtenu par Align GVGD pour la mutation c.135G>C (p.Lys45Asn).

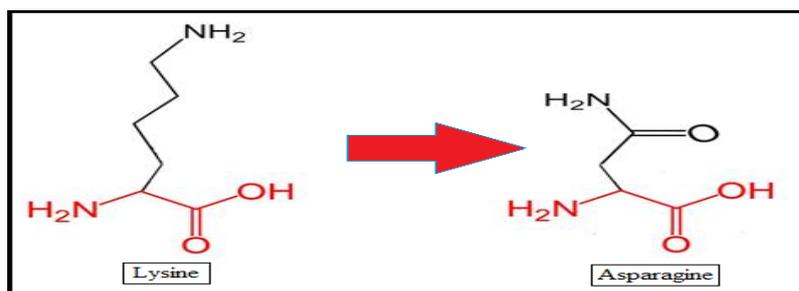
L'analyse de la mutation c.135G>C (p.Lys45Asn) par Align GVGD a permis de calculer les valeurs GV et GD. La valeur de la variation de Grantham GV est égale à 26, ce qui signifie que cette substitution ne siège pas au niveau d'une région fortement conservée entre les espèces analysées. D'autre part, le calcul de GD montre une valeur égale à 79,24 ceci montre que les propriétés physicochimique entre asparagine et l'histidine sont modérément conservées. Ce logiciel a permis de prédire que cette dernière correspond à la classe C35. La mutation est prédite comme affectant la fonction du BRCA1.

***I.2.5. Comparaison des propriétés physicochimiques entre les résidus normaux et mutés (Project HOPE)***

Pour rappel, la mutation c.135G>C est responsable de la substitution de l'aa lysine par l'aa asparagine. Les structures de ces deux aa sont représentées dans la figure 16., tandis que les propriétés physicochimiques de chaque résidu sont résumées dans le tableau IV.

La figure montre les structures schématiques de l'aa normal (à gauche) et le muté (à droite). La partie commune (squelette) entre les deux aa est de couleur rouge.

Tandis que la chaîne latérale qui est unique pour chaque résidu est de couleur noire. Ces deux aa présentent des propriétés chimiques particulières en raison de la nature de leurs chaînes latérales.



**Figure 16:** Les structures schématiques de la lysine et l'asparagine

**Tableau IV :** Les propriétés physicochimiques des deux aa Lysine et Asparagine.

Résidu	Taille	Charge
lysine	146 Da	Positive
asparagine	132 Da	Neutre

Les différences entre ces deux acides aminés résident dans leur taille ainsi que leur charge (tableau IV). Ces différences sont souvent à l'origine de changement de conformations mais aussi de fonction lorsqu'ils touchent des sites importants.

## **II. Discussion**

La nature de la mutation est fondamentale pour évaluer son caractère pathogène. Les mutations localisées au niveau des sites donneurs et accepteurs d'épissage, les délétions et les insertions entraînent un décalage du cadre de lecture responsable de l'apparition d'un codon stop prématuré ou la synthèse d'une protéine allongée. Le caractère délétère de ces mutations est donc admis. En revanche, les effets délétères des mutations faux-sens sont sujets à discussion essentiellement si elles sont responsables de forme sévère d'une pathologie.

C'est la raison pour laquelle les effets délétères de deux nouvelles mutations rapportées dans cette étude ont été prédits par différents logiciels de prédiction *in silico*.

Les effets délétères des mutations faux-sens sur la stabilité ont été évalués par I mutant 2.0. Tandis que les effets de ces dernières sur la fonction protéique ont été prédits par trois logiciels différents (SIFT, Polyphen et Align GVGD).

La stabilité d'une protéine est évaluée par le calcul d'énergie ; plus le champ de force de l'aa est grand, moins la protéine est stable. D'autre part, les logiciels de prédiction des effets de mutations sur la fonction protéique reposent sur trois approches de prédiction (le calcul du score matriciel de Grantham, les différences des propriétés physico-chimiques et l'analyse de l'état de conservation inter-espèces du résidu étudié). Toutefois, ces différents résultats obtenus sont généraux et ne spécifient pas les effets exacts des deux mutations.

Le gène *BRCA1* appartient à une classe de gènes suppresseurs de tumeur et code pour une protéine qui joue un rôle dans le contrôle négatif du cycle cellulaire afin de prévenir la prolifération incontrôlée des cellules.

La mutation c.122A>G agit sur la région codante de ce gène et consiste en une substitution d'une Histidine par une Arginine au niveau du codon 41 affectant ainsi la protéine synthétisée. Les résultats obtenus par le logiciel I Mutant 2.0 sont en faveur d'un effet déstabilisant. De même, les logiciels (SIFT, Polyphen2 et) ont prédit des effets délétères sur la protéine. Toutefois, le résultat obtenu par AGAVD montre que cette variation n'a pas d'impact sur la protéine.

En fait, l'effet délétère réside essentiellement aux changements des propriétés physico-chimiques des deux acides aminés. L'acide aminé histidine est plus petit (155Da) que l'arginine (174Da). De plus, l'arginine sert de donneur de liaisons hydrogènes aux autres

acides aminés. Ces différences pourraient entraîner la perte ou un gain d'interaction avec d'autres molécules et donc un changement au niveau structural de la protéine.

De même, la mutation c.135G>C a également été analysée et les résultats montrent qu'elle n'est pas déstabilisante et n'a pas d'impact sur la fonction de la protéine. Cependant, les différences des propriétés chimiques des acides aminés pourraient expliquer l'effet des mutations. Concernant la taille, le résidu muté est petit (132 Da) que le résidu sauvage (146Da). La charge diffère également, le résidu muté est neutre et le sauvage est positif. Ces différences pourraient perturber la protéine.

En effet, les deux substitutions sont situées dans la région du domaine en doigt de zinc. Ce motif fait partie d'un domaine plus étendu contenant des hélices alpha antiparallèles indispensables aux interactions protéiques. Ce domaine est impliqué dans le processus de l'ubiquitination (Shabbeer et al., 2013; Yu et al., 2003). Il lie également la protéine BAP1 (BRCA-associated protein 1) enzyme impliquée dans la prédisposition au cancer. Ces mutations peuvent entraîner un changement de conformation de la protéine et ce qui pourrait provoquer l'inhibition de l'interaction entre cette protéine et les autres protéines et facteurs. Ce qui aura comme effet une prolifération incontrôlable des cellules et par conséquent le développement du cancer (Kumaraswamy & Shiekhattar, 2007).

En fait, la combinaison de ces logiciels est très informative pour la prédiction *in silico* des caractères délétères des nouvelles mutations. Les études structurales *in silico* permettent d'évaluer le retentissement des mutations sur la structure ainsi que la fonction de la protéine; il est intéressant d'en confronter les résultats d'étude *in silico* avec ceux des études fonctionnelles afin d'évaluer les fonctionnalités des nouvelles mutations. Toutefois, cette dernière approche est rarement utilisée en pratique de diagnostic génétique car elle est très lente, coûteuse et nécessite une équipe spécialisée.

*Conclusion*

*et*

*perspectives*

Dans cette étude, nous nous sommes contribué à la compréhension des mécanismes moléculaires à l'origine du défaut de la protéine BRCA1. Les effets délétères de deux mutations spécifiques à la population Algérienne atteinte de cancer de sein étés étudiés en utilisant une combinaison de logiciels d'étude *in silico*. Ces mutations correspondent à deux mutations faux sens au niveau de l'exon 2 (c.122A>G) et (c.135G>C).

La mutation c.122A>G (p.His41Arg) a été prédite comme délétère au niveau structural de la protéine BRCA1. Les changements physicochimiques entre les acides aminés native et muté sont responsables de perte des interactions moléculaires ayant des effets déstabilisants au niveau de la structure de la protéine. D'autre part, la variation c.135G>C (p.Lys45Asn) a été prédite comme non délétère au niveau fonctionnel de la protéine. Mais, même si la mutation est sans effet délétère sur la protéine, la conséquence de cette mutation faux sens dépend principalement de la position de l'acide aminé muté et de ses propriétés chimiques. Une étude fonctionnelle de la protéine reste nécessaire afin de déterminer l'impact de ces mutations.

Notre travail a permis de montrer l'efficacité de l'utilisation associée de plusieurs logiciels prédictifs afin d'évaluer l'impact des mutations au niveau protéique,; il est intéressant d'en confronter les résultats d'étude *in silico* avec ceux des études fonctionnelles afin d'évaluer les fonctionnalités des nouvelles mutations. Par ailleurs, les logiciels bioinformatiques de prédiction de l'impact des mutations sur la protéine d'intérêt restent un outil de choix non couteux et confèrent les moyens de fournir rapidement une valeur pronostique aux résultats du génotypage, ce qui les rend potentiellement précieux pour les soins aux patients ainsi que pour la recherche scientifiques. Néanmoins, ils doivent être validés par des études *in vitro* et/ou *in vivo*.

*References*

*bibliographiques*

- Anderson, S. F., Schlegel, B. P., Nakajima, T., Wolpin, E. S., & Parvin, J. D. (1998). BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A. *Nature genetics*, 19(3), 254-256.
- Andrieu, J. M., Colonna, P., & Lévy, R. (1997). Cancers: guide pratique d'évaluation de traitement et de surveillance. *Boeck Secundair*, 19, 1170-1189.
- Aube, A., Cabarrocas, J., Neunlist, M., Doulay, F., Liblau, R., & Galmiche, J. (2004). *Altered intestinal motility and permeability in a transgenic mice model of glia disruption*. Paper presented at the Gastroenterology.
- Awad, D. (2014). *Vers un système perceptuel de reconnaissance d'objets*. Université de La Rochelle.
- Bairoch, A., & Apweiler, R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Research*, 28(1), 45-48.
- Barbier, Y., & Rasmont, P. (2003). Des banques de données biogéographiques vers les banques de données biotopographiques. De nouvelles techniques pour de nouveaux enjeux. *Phytoprotection*, 84(2), 51-57.
- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Rapp, B. A., & Wheeler, D. L. (2002). GenBank. *Nucleic Acids Research*, 30(1), 17-20.
- BENYELLES, M. K. (2018). *Identification des cellules souches cancéreuses dans les cancers du sein et évaluation du niveau d'expression des marqueurs CD44 et ALDH1*.
- Berman, H. M., Bhat, T. N., Bourne, P. E., Feng, Z., Gilliland, G., Weissig, H., & Westbrook, J. (2000). The Protein Data Bank and the challenge of structural genomics. *Nature structural biology*, 7(11), 957-959.
- Blanc, G., Ogata, H., Robert, C., Audic, S., Suhre, K., Vestris, G., . . . Raoult, D. (2007). Reductive genome evolution from the mother of Rickettsia. *PLoS genetics*, 3(1), e14.
- Blanchet, C. (1999). Logiciel MPSA et ressources bioinformatiques client-serveur Web dédiés à l'analyse de séquences de protéine. *Lyon, France: Université Claude Bernard, Lyon, 1*.
- Boulton, S. J. (2006). BRCA1-mediated ubiquitylation. *Cell Cycle*, 5(14), 1481-1486.
- Boveri, T. (1929). The origin of malignant tumors. *Williams & Wilkins, Baltimore*, 6, 27.
- Boyer, S. b. (2015). *Analyse statistique de la sélection dans des banques minimalistes de protéines*. Université Grenoble Alpes.

## Références bibliographiques

- Brody, L. C., Abel, K. J., Castilla, L. H., Couch, F. J., McKinley, D. R., Yin, G., . . . Xu, J. (1995). Construction of a transcription map surrounding the *BRCA1* locus of human chromosome 17. *Genomics*, 25(1), 238-247.
- Brzovic, P. S., Rajagopal, P., Hoyt, D. W., King, M.-C., & Kleit, R. E. (2001). Structure of a BRCA1–BARD1 heterodimeric RING–RING complex. *Nature structural biology*, 8(10), 833-837.
- Bunnell, D. B., González, M. J., & Stein, R. A. (2003). Zooplankton biomass enhances growth, but not survival, of first-feeding *Pomoxis* spp. larvae. *Canadian Journal of Fisheries and Aquatic Sciences*, 60(11), 1314-1323.
- Cantor, S. B., Bell, D. W., Ganesan, S., Kass, E. M., Drapkin, R., Grossman, S., . . . Haber, D. A. (2001). BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell*, 105(1), 149-160.
- Capriotti, E., Fariselli, P., Calabrese, R., & Casadio, R. (2005). Predicting protein stability changes from sequences using support vector machines. [Research Support, Non-U.S. Gov't]. *Bioinformatics*, 21 Suppl 2, ii54-58. doi: 10.1093/bioinformatics/bti1109
- Chen, C.-F., Li, S., Chen, Y., Chen, P.-L., Sharp, Z. D., & Lee, W.-H. (1996). The nuclear localization sequences of the BRCA1 protein interact with the importin- $\alpha$  subunit of the nuclear transport signal receptor. *Journal of Biological Chemistry*, 271(51), 32863-32868.
- Chen, L., Nievera, C. J., Lee, A. Y.-L., & Wu, X. (2008). Cell cycle-dependent complex formation of BRCA1· CtIP· MRN is important for DNA double-strand break repair. *Journal of Biological Chemistry*, 283(12), 7713-7720.
- Cherbal, F. (2010). *Contribution des mutations germinales des gènes BRCA1 et BRCA2 dans le syndrome de prédisposition héréditaires aux cancers du sein et de l'ovaire chez des familles algériennes*. Alger.
- Clark, S. L., Rodriguez, A. M., Snyder, R. R., Hankins, G. D., & Boehning, D. (2012). Structure-function of the tumor suppressor *BRCA1*. *Computational and structural biotechnology journal*, 1(1), e201204005.
- Coppola, D. (2010). Mechanisms of Oncogenesis: An Update on Tumorigenesis. *Pringer*, 19, 306-309.
- Cousineau, I., Abaji, C., & Belmaaza, A. (2005). BRCA1 regulates RAD51 function in response to DNA damage and suppresses spontaneous sister chromatid replication slippage: implications for sister chromatid cohesion, genome stability, and carcinogenesis. *Cancer research*, 65(24), 11384-11391.

## Références bibliographiques

- Delort, L. (2007). *Facteurs de risque et de protection des cancers dans l'étude épidémiologique COSA (Cancers de l'Ovaire et du Sein en Auvergne): étude des polymorphismes génétiques et des interactions gènes-environnement*. Clermont-Ferrand 1.
- Derynck, R., Zhang, Y., & Feng, X.-H. (1998). Transcriptional activators of TGF- $\beta$  responses: Smads. *Cell*, 95(6), 737-740.
- Ding, H., Richard, P., Nakayama, K., Sugawara, K., Arakane, T., Sekiba, Y., . . . Takahashi, T. (2008). Observation of Fermi-surface-dependent nodeless superconducting gaps in Ba<sub>0.6</sub>K<sub>0.4</sub>Fe<sub>2</sub>As<sub>2</sub>. *EPL (Europhysics Letters)*, 83(4), 47001.
- DuBrin, A. J. (2002). *The winning edge: How to motivate, influence, and manage your company's human resources*: International Thomson Education Group.
- Elshaikh, A., Ismaiel, M. M., Osman, M. M., & Shokri, S. A. I. (2016). Computational analysis of single nucleotide polymorphism (SNPs) in human GRM4 gene. *American Journal of Biomedical Research*, 4(3), 61-73.
- Gatei, M., Zhou, B.-B., Hobson, K., Scott, S., Young, D., & Khanna, K. K. (2001). Ataxia telangiectasia mutated (ATM) kinase and ATM and Rad3 related kinase mediate phosphorylation of Brca1 at distinct and overlapping sites: in vivo assessment using phospho-specific antibodies. *Journal of Biological Chemistry*, 276(20), 17276-17280.
- Glass, H. C., Glidden, D., Jeremy, R. J., Barkovich, A. J., Ferriero, D. M., & Miller, S. P. (2009). Clinical neonatal seizures are independently associated with outcome in infants at risk for hypoxic-ischemic brain injury. *The Journal of pediatrics*, 155(3), 318-323.
- Hadj Bekkouche, N. (2018). Population masculine au sein des familles avec mutation germinale dans les gènes BRCA1 et BRCA2 aspects cliniques et pathologiques des cancers.
- Hammes, F., Salhi, E., Köster, O., Kaiser, H.-P., Egli, T., & Von Gunten, U. (2006). Mechanistic and kinetic evaluation of organic disinfection by-product and assimilable organic carbon (AOC) formation during the ozonation of drinking water. *Water research*, 40(12), 2275-2286.
- Harkness, K., Sussman, J., Davies-Jones, G., Greenwood, J., & Woodroffe, M. (2003). Cytokine regulation of MCP-1 expression in brain and retinal microvascular endothelial cells. *Journal of neuroimmunology*, 142(1-2), 1-9.
- Harte, M. T., O'Brien, G. J., Ryan, N. M., Gorski, J. J., Savage, K. I., Crawford, N. T., . . . Harkin, D. P. (2010). BRD7, a subunit of SWI/SNF complexes, binds directly to

## Références bibliographiques

- BRCA1* and regulates BRCA1-dependent transcription. *Cancer research*, 70(6), 2538-2547.
- Kouranov, A., Xie, L., de la Cruz, J., Chen, L., Westbrook, J., Bourne, P. E., & Berman, H. M. (2006). The RCSB PDB information portal for structural genomics. *Nucleic Acids Research*, 34(suppl\_1), D302-D305.
- Kumaraswamy, E., & Shiekhattar, R. (2007). Activation of *BRCA1/BRCA2*-associated helicase BACH1 is required for timely progression through S phase. *Molecular and cellular biology*, 27(19), 6733-6741.
- Lambert, C. (2003). Développement d'une méthode automatique fiable de modélisation de la structure tridimensionnelle des protéines par homologie et application au protéome de *Brucella melitensis*. département de biologie. *FUNDP, Namur*.
- Luu, T. D. (2012). *Développement d'une infrastructure d'analyse multi-niveaux pour la découverte des relations entre génotype et phénotype dans les maladies génétiques humaines*. Strasbourg.
- Makiniemi, M., Hillukkala, T., Tuusa, J., Reini, K., Vaara, M., Huang, D., . . . Mäkelä, T. P. (2001). BRCT domain-containing protein TopBP1 functions in DNA replication and damage response. *Journal of Biological Chemistry*, 276(32), 30399-30406.
- Mathe, E., Olivier, M., Kato, S., Ishioka, C., Hainaut, P., & Tavtigian, S. V. (2006). Computational approaches for predicting the biological effect of p53 missense mutations: a comparison of three sequence analysis based methods. *Nucleic Acids Research*, 34(5), 1317-1325.
- Meza, J. E., Brzovic, P. S., King, M.-C., & Klevit, R. E. (1999). Mapping the functional domains of BRCA1: interaction of the ring finger domains of BRCA1 and BARD1. *Journal of Biological Chemistry*, 274(9), 5659-5665.
- Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., . . . Ding, W. (1994). A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*, 266(5182), 66-71.
- Mulder, N. J., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., . . . Cerutti, L. (2005). InterPro, progress and status in 2005. *Nucleic Acids Research*, 33(suppl\_1), D201-D205.
- Ng, P. C., & Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Research*, 31(13), 3812-3814.

## Références bibliographiques

- Ng, P. C., & Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. [Research Support, U.S. Gov't, P.H.S.]. *Nucleic Acids Res*, 31(13), 3812-3814.
- Pao, G. M., Janknecht, R., Ruffner, H., Hunter, T., & Verma, I. M. (2000). CBP/p300 interact with and function as transcriptional coactivators of BRCA1. *Proceedings of the National Academy of Sciences*, 97(3), 1020-1025.
- Paul, A., & Paul, S. (2014). The breast cancer susceptibility genes (BRCA) in breast and ovarian cancers. *Frontiers in bioscience (Landmark edition)*, 19, 605.
- Rodríguez, J. A., & Henderson, B. R. (2000). Identification of a functional nuclear export sequence in BRCA1. *Journal of Biological Chemistry*, 275(49), 38589-38596.
- Rubin, S. C., Benjamin, I., Behbakht, K., Takahashi, H., Morgan, M. A., LiVolsi, V. A., . . . Weber, B. L. (1996). Clinical and pathological features of ovarian cancer in women with germ-line mutations of BRCA1. *New England Journal of Medicine*, 335(19), 1413-1416.
- Ruij, J. (2015). *Modélisation d'objets sédimentaires par des surfaces paramétriques et application à l'analyse d'image*. Université de Lorraine.
- Salhi, N., Gargouri, R., Bakour, R., Nacer, B., Boulgua, K., Guargouri, A., & Farid, C. (2006). *BRCA1* mutations in breast cancer: identification of new mutations in the BRCA1 gene in Algerian population: AACR.
- Savage, K. I., Matchett, K. B., Barros, E. M., Cooper, K. M., Irwin, G. W., Gorski, J. J., . . . Madden, A. F. (2014). *BRCA1* deficiency exacerbates estrogen-induced DNA damage and genomic instability. *Cancer research*, 74(10), 2773-2784.
- Scully, R., Chen, J., Ochs, R. L., Keegan, K., Hoekstra, M., Feunteun, J., & Livingston, D. M. (1997). Dynamic changes of *BRCA1* subnuclear location and phosphorylation state are initiated by DNA damage. *Cell*, 90(3), 425-435.
- Scully, R., Chen, J., Plug, A., Xiao, Y., Weaver, D., Feunteun, J., . . . Livingston, D. M. (1997). Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell*, 88(2), 265-275.
- Senechal, C., & Mouret-Fourme, E. Prédiposition génétique aux cancers du sein et de l'ovaire par BRCA1 et BRCA2.
- Shabbeer, S., Omer, D., Berneman, D., Weitzman, O., Alpaugh, A., Pietraszkiewicz, A., . . . Yarden, R. I. (2013). *BRCA1* targets G2/M cell cycle proteins for ubiquitination and proteasomal degradation. *Oncogene*, 32(42), 5005-5016.

## Références bibliographiques

- Sopik, V., Phelan, C., Cybulski, C., & Narod, S. (2015). *BRCA1* and *BRCA2* mutations and the risk for colorectal cancer. *Clinical genetics*, 87(5), 411-418.
- Stoesser, G., Sterk, P., Tuli, M. A., Stoehr, P. J., & Cameron, G. N. (1997). The EMBL nucleotide sequence database. *Nucleic Acids Research*, 25(1), 7-13.
- Sunyaev, S., Ramensky, V., & Bork, P. (2000). Towards a structural basis of human non-synonymous single nucleotide polymorphisms. [Comparative Study Review]. *Trends Genet*, 16(5), 198-200.
- Sussman, J. L., Lin, D., Jiang, J., Manning, N. O., Prilusky, J., Ritter, O., & Abola, E. E. (1998). Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules. *Acta Crystallographica Section D: Biological Crystallography*, 54(6), 1078-1084.
- Taste, C. (2007). *Fonctions et régulation de la protéine suppresseur de tumeurs BRCA1 dans la réponse cellulaire aux poisons du fuseau mitotique*. Toulouse 3.
- Tavtigian, S. V., Byrnes, G. B., Goldgar, D. E., & Thomas, A. (2008). Classification of rare missense substitutions, using risk surfaces, with genetic- and molecular- epidemiology applications. *Human mutation*, 29(11), 1342-1354.
- Traven, A., & Heierhorst, J. (2005). SQ/TQ cluster domains: concentrated ATM/ATR kinase phosphorylation site regions in DNA- damage- response proteins. *Bioessays*, 27(4), 397-407.
- Williams, R. S., Green, R., & Glover, J. M. (2001). Crystal structure of the BRCT repeat region from the breast cancer-associated protein *BRCA1*. *Nature structural biology*, 8(10), 838-842.
- Wojnarowicz, P. M. (2012). *Chromosome 17 and epithelial ovarian cancer: Evidence of genomic loss, altered gene expression, and tumour suppressor gene candidates*. McGill University Libraries.
- Yu, X., Chini, C. C. S., He, M., Mer, G., & Chen, J. (2003). The BRCT domain is a phospho-protein binding domain. *Science*, 302(5645), 639-642.
- Zhang, H., Somasundaram, K., Peng, Y., Tian, H., Zhang, H., Bi, D., . . . El-Deiry, W. S. (1998). *BRCA1* physically associates with p53 and stimulates its transcriptional activity. *Oncogene*, 16(13), 1713-1721.
- Zhang, Y., Zeleznik- Le, N., Emmanuel, N., Jayathilaka, N., Chen, J., Strissel, P., . . . Taki, T. (2004). Characterization of genomic breakpoints in *MLL* and *CBP* in leukemia patients with t (11; 16). *Genes, Chromosomes and Cancer*, 41(3), 257-265.

## *Références bibliographiques*

- Zhuang, J., Zhang, J., Willers, H., Wang, H., Chung, J. H., Van Gent, D. C., . . . Xia, F. (2006). Checkpoint kinase 2-mediated phosphorylation of BRCA1 regulates the fidelity of nonhomologous end-joining. *Cancer research*, 66(3), 1401-1408.
- Zografos, L. ( 2002). Chapitre rétinoblastome Rapport de la société française d'ophtalmologie.

## Résumé

Le gène *BRCA1* est un gène suppresseur de tumeur et son altération moléculaire pourra conduire à un risque accru de cancer du sein et ovaire. Dans le but de contribuer à l'étude moléculaire, nous nous sommes intéressées à l'étude *in silico* des fonctionnalités de deux mutations du gène *BRCA1* identifiées dans une étude algérienne : c.122A>G et c.135G>C.

Ainsi, les effets délétères de deux mutations ont été prédites en utilisant un protocole d'étude *in silico* constituée de différents logiciels (I-Mutant 2.0 ; SIFT ; Polyphen-2 ; Align GVGD ; Project HOPE).

La mutation c.122A>G est prédite comme étant délétère altérant la protéine *BRCA1* tandis que la substitution c.135G>C est prédite comme non délétère. En fait, les changements physicochimiques entre les acides aminés natifs et mutés perturbent la stabilité de la protéine et peuvent engendrer la perte d'interaction avec d'autres molécules. Nos résultats ont montré que ces différents logiciels sont cohérents et que leur combinaison pourrait augmenter les performances de prédiction de l'effet des mutations.

Cette étude nous a permis de mieux comprendre l'effet des mutations et a dévoilé l'importance des études de prédiction par des logiciels bio-informatiques par leur puissance prédictives.

**Mots clés** : *BRCA1* ; *in silico* ; mutation, faux sens .