

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE SAAD DAHLAB, BLIDA -1-
FACULTE DES SCIENCES DE LA NATURE ET DE LA VIE
DEPARTEMENT DE BIOLOGIE ET PHYSIOLOGIE CELLULAIRE



Mémoire de fin d'études

En vue de l'obtention du diplôme de Master

En Sciences de la Nature et de la Vie

Filière : SCIENCES BIOLOGIQUES

Option : Génétique

THÈME

**Polymorphisme de l'interleukine1 et statut hormonal
chez les patientes atteintes du cancer du sein.**

Présenté par :

M^{elle} Zane Zineb Intisar

M^{elle} Zeraimi Samah

Soutenu le : 29/09/2020

Devant le jury :

Belkhiter S.

MCA

USDB 1

Présidente.

Benazouz F.

MAA

USDB 1

Examinatrice.

Eddaikra A.

MCB

USDB 1

Promotrice.

Promotion : 2019 – 2020

Remerciements

*Mes remerciements vont tout d'abord à Dieu le tout puissant qui a
Éclairé mon chemin.*

*Je voudrais aussi adresser toute ma gratitude à mon encadreur,
Mme Atika Eddaikra pour
Son précieux aide durant le déroulement de mon projet.*

*Je remercie vivement le président du jury d'avoir
Accepté de présider mon jury de soutenance et aux membres de jury d'avoir accepté de
faire partie et d'évaluer mon
Mémoire de fin d'études. Votre participation à ma soutenance a été un
grand honneur pour moi.*

*Mes remerciements vont aussi à tous mes professeurs, enseignants et
toutes les personnes qui m'ont soutenus jusqu'au bout, et qui n'ont
pas cessé de me donner des conseils très importants au cours de mon
cursus.*

*Mes remerciements vont également à ma famille qui m'a aidé tout au
long de mon chemin avec leur patience, encouragement et confiance et
encouragement tout au long du chemin.*

*Mon derniers remerciements, mais non les moindres s'adressent à tous
mes amies de la promotion 2020, pour avoir beaucoup de sollicitudes et
de chaleur humaine dont nous avions tant besoin.*

*Et en fin, je remercie tous ceux qui ont contribué de loin ou de près à la
Réalisation de ce travail.*

Dédicace

**Je dédie ce travail aux personnes qui m'ont toujours aidé et encouragé,
qui étaient toujours à mes côtés, et qui m'ont accompagné durant mon
chemin d'études supérieures**

**Surtout mes chers parents qui m'ont donné la vie et aimé
inconditionnellement.**

**Aux personnes dont j'ai bien aimé la présence dans ce jour, à toute
mes sœurs, mes aimables amies**

Résumé

Le gène humain IL1B code une cytokine pro-inflammatoire secrété d'une manière régulée dans un organisme sain. Ce gène est situé en position 2q14.1 et se compose de 07 exons.

La modélisation et la clarification de la variation génétique constituent un véritable défi en génétique humaine, aussi bien pour les personnes en bonne santé que pour les personnes malades.

L'étude in silico du Le polymorphisme du gène IL1B en utilisant une méthodologie bioinformatique au moyen de l'analyse de diverses bases de données et de navigateurs Web open source a montré que ce gène est caractérisé par un profil polymorphe qui a un très grand nombre de polymorphismes nucléotidique simple (SNPs) associés à différents types de cancer.

Nous nous sommes intéressés au polymorphisme nucléotidique simple (SNP) dans le gène d'IL1B qui peuvent être en association avec le cancer du sein, nous avons constaté que les trois SNP -511 (C>T ; dbSNP: rs16944), -31 (T>C; dbSNP: rs1143627) localisés dans la région du promoteur et +3954 (C>T; dbSNP: rs1143634) localisé dans l'exon 5, sont liés à la manifestation du cancer du sein

Différents outils de bioinformatique s'offrent à nous pour ce travail telques : NCBI, Ensembl, UCSC, GALAXY et DISGENET.

La bioinformatique est un outil incontournable dans l'identification et la prédiction de la fonction du gène IL1B et leur association avec le cancer du sein.

Mots Clés : IL1B, Cancer du sein, SNP, Polymorphisme, Bioinformatique.

Abstract

The human IL1B gene encodes a pro-inflammatory cytokine secreted in a regulated manner in a healthy organism. This gene is located at position 2q14.1 and consists of 07 exons.

Modeling and clarifying genetic variation is a real challenge in human genetics, for both healthy and sick people.

The in silico study of the polymorphism of the IL1B gene using a bioinformatics methodology through the analysis of various databases and open source web browsers has shown that this gene is characterized by a polymorphic profile which has a very large number of single nucleotide polymorphisms (SNPs) associated with different types of cancer.

We were interested in single nucleotide polymorphism (SNP) in the IL1B gene that may be associated with breast cancer, we found that the three SNPs -511 (C> T; dbSNP: rs16944), -31 (T> C; dbSNP : rs1143627) located in the promoter region and 3954 (C> T; dbSNP: rs1143634) located in exon 5, are linked to the manifestation of breast cancer

Different bioinformatics tools are available to us for this work such as: NCBI, Ensembl, UCSC, GALAXY and DISGENET.

Bioinformatics is an essential tool in the identification and prediction of the function of the IL1B gene and their association with breast cancer.

Keywords : IL1B, Breast cancer, SNP, Polymorphism, Bioinformatic.

المخلص

يقوم جين IL1B البشري بتشفير السيتوكين المؤيد للالتهابات والذي يتم إفرازه بطريقة منظمة في كائن حي صحي. يقع هذا الجين في الموضع q14.1 2 ويتكون من 07 إكسونات.

تعد نمذجة التنوع الجيني وتوضيحه تحديًا حقيقيًا في علم الوراثة البشرية لكل من الأشخاص الأصحاء والمرضى. أظهرت الدراسة في السيليكو لتعدد الأشكال لجين IL1B باستخدام منهجية المعلوماتية الحيوية من خلال تحليل قواعد البيانات المختلفة ومتصفحات الويب مفتوحة المصدر أن هذا الجين يتميز بملف تعريف متعدد الأشكال يحتوي على عدد كبير جدًا من تعدد الأشكال النكليوتيدات الوحيد SNPs المرتبطة بأنواع مختلفة من السرطان.

كنا مهتمين بتعدد الأشكال النكليوتيدات الوحيد SNPs في جين IL1B الذي قد يكون مرتبطًا بسرطان الثدي، وجدنا أن SNPs الثلاثة، (rs16944; C>T) -511 (dbSNP) ، و (rs1143627; T>C) -31 (dbSNP) ، الموجود في منطقة المروج و (rs1143634; C>T) +3954 (dbSNP) الموجود في exon 5 ، مرتبط بمظاهر سرطان الثدي

تتوفر لنا أدوات معلوماتية حيوية مختلفة لهذا العمل مثل NCBI و Ensembl و UCSC و GALAXY و

.DISGENET

المعلوماتية الحيوية هي أداة أساسية في تحديد وظيفة جين IL1B والتنبؤ بها وارتباطها بسرطان الثدي.

الكلمات المفتاحية: IL1B ، سرطان الثدي ، SNP ، تعدد الأشكال ، المعلوماتية الحيوية.

Sommaire

Introduction.....	1
-------------------	---

Chapitre I : rappels bibliographiques

1. Cancer du sein.....	3
1.1. Cancer du sein en Algérie.....	4
2. La formation de la tumeur primaire.....	5
3. Signature moléculaire de cancer du sein.....	6
3.1. Méthodes employées dans la signature moléculaire.....	8
4. Cas de cancer hormono-dépendant.....	9
5. Les facteurs de risques.....	10
6. Polymorphisme IL1 et le cancer du sein.....	11
6.1. Définition de polymorphisme génétique.....	12
6.2. Gène IL1B.....	13
6.2.1. Structure et localisation.....	13
6.2.2. Expression.....	13
6.2.3. Biosynthèse et activation.....	14
6.2.4. Régulation.....	14
6.2.5. Homologie.....	15
6.2.6. Protéine de gène il1b et sa fonction.....	15
6.3. Polymorphisme génétique d'IL1B.....	15
6.3.1. Association du polymorphisme IL1B et le cancer du sein.....	16
6.3.2. Association d'IL1B et le statut hormonal dans le cancer du sein.....	17

Chapitre II : Matériel et méthodes

2.1. Matériel.....	19
2.1.1. Matériel biologique.....	19
2.1.2. Matériels bio-informatiques.....	20
2.2. Méthodes.....	21
2.2.1. Interrogation des bases de données.....	23
2.2.1.1. Outil NCBI.....	23
2.2.2. Recherche des polymorphismes « SNPs ».....	24
2.2.2.1. Outil UCSC.....	24
2.2.2.2. Outil ensembl genome browser.....	26
2.2.2.3. Outil galaxy.....	29
2.2.3. Identification d'association pathologique.....	33
2.2.3.1. Outil ALFRED.....	33
2.2.3.2. Outil DISGENET.....	34
2.2.3.3. Outil GeneMania.....	35
2.2.3.4. Outil STRING.....	36

Chapitre III : Résultats et discussion

Conclusion.....	56
-----------------	----

Références bibliographiques.....	59
----------------------------------	----

Annexes

Liste des abréviations

CS : Cancer du sein
ER : Estrogen Receptor
PR : Progesteron Receptor
HER2 : Human Epidermal Growth Factor Receptor 2
IL1 : Interleukin 1
IL1A : Interleukin 1 alpha
IL1B : Interleukin 1 beta
IL1RN : Interleukin 1 Receptor Antagonist
TN : Triple Négative
qRT-PCR : Real Time Quantitative Reverse Transcription
SNP : Single Nucleotide Polymorphism
Kb : kilo base
RPKM : Unité d'expression
PAMP : Pathogen Associated Molecular Pattern
DAMP : Damage Associated Molecular Pattern
NLRP3 : NOD Like Receptor family pyrin domain containing 3
Myd88 : Myeloid differentiation primary response 88
TRAF6 : TNF Receptor Associated Factor 6
NFKB : Nuclear Factor Kappa B
NLR : NOD Like Receptor
COX2 : Cyclooxygénase 2
iNOS : Inductible Nitric Oxide Synthase
TNF : Tumoral Necrosis Factors
IL6 : Interleukine 6
dbSNP : base de données SNP
rs : Référence snip
LPS : Lipopolysaccharide
OR : Odds ratio
IC : Intervalle de confiance
RefSeq : Reference sequence
UTR : Untranslated Transcribed Region
HWE : Hardy Weinberg Equilibrium
MAP3K3 : Mitogen-Activated Protein Kinase Kinase Kinase 3
IL1R1 : Interleukin 1 Receptor Type 1
A2M : Alpha-2-Macroglobulin
SQSTM1 : Sequestosome 1
CASP1 : Caspase 1
MAPK8IP2 : Mitogen-Activated Protein Kinase 8 Interacting Protein 2
IL1R2 : Interleukin 1 Receptor type 2

IL1RAP : Interleukin 1 Receptor Accessory Protein
TICAM2 : Toll Like Receptor Adaptor Molecule 2
IRAK3 : Interleukin 1 Receptor Associated Kinase 3
CMA1 : Chymase 1
UBE2N : Ubiquitin Conjugating Enzyme E2 N
CASP3 : Caspase 3
TOLLIP : Toll Interacting Protein
IL18 : Interleukin 18
CXCL8 : C-X-C Motif Chemokine Ligand 8
ZNF675 : Zinc Finger Protein 675
IRAK2 : Interleukin 1 Receptor Associated Kinase 2

Tableau des figures

	Titre	Page
Figure1	Estimation du nombre de cas incidents et de décès dans le monde, femmes, tous âges. GLOBOCAN 2018	4
Figure2	Estimation du nombre de cas en Algérie, femmes, tous âges. GLOBOCAN 2018	5
Figure3	Estimation du nombre de décès en Algérie, femmes, tous âges. GLOBOCAN 2018	5
Figure4	Activation illégitime du récepteur des œstrogènes	10
Figure5	Diagramme schématique des facteurs de risque et de la prévention du cancer du sein.	11
Figure6	Gènes de la famille de l'IL-1 sur le chromosome 2 et principaux SNPs fonctionnels.	12
Figure7	Gène d'IL1B	13
Figure8	Expression d'IL1B dans les organes	13
Figure9	Signalisation IL-1B	14
Figure10	Régulations transcriptionnelles d'IL-1 β au niveau de leur promoteur	15
Figure11	Méthode de recherche des SNP du gène IL1B	22
Figure12	les étapes de recherche dans le site NCBI	23
Figure13	Méthode de recherche des SNPs sur le navigateur UCSC	25
Figure14	Méthode d'utilisation d'outils Ensemble genome	28
Figure15	Méthode d'utilisation d'outils galaxy	30
Figure16	L'organigramme de notre recherche sur Galaxy	33
Figure17	Page d'accueil du site ALFRED	34
Figure18	Page d'accueil du site DigestNet	35
Figure19	Page d'accueil du site GeneMania	36
Figure20	Page d'accueil de site STRING	37
Figure21	Résumé graphique de l'alignement	41
Figure22	Localisation chromosomique du gene IL1B	43

Figure23	Liste de tous les SNPs positionnées sur le gène IL1B via UCSC Génome Browser	44
Figure24	Les différents SNPs du lien HUGO	45
Figure25	Distribution des allèles ancestraux et des allèles dérivés du rs11687624	45
Figure26	Structure du gène IL1B selon Ensembl génome browser	46
Figure27	Résultats des SNP du gène IL1B de la base ALFRED	49
Figure28	Variantes IL1B dans le cancer du sein	50
Figure29	Résultats de l'analyse du réseau GeneMania des interactions avec le gène IL1B	51
Figure30	Interaction protéine-protéine du gène IL1B	52
Figure31	Co-expression des gènes avec le gène IL1B	53
Figure32	Modélisation comparative de protéine IL1B par SUISSE MODEL	53

Liste des tableaux

	Titre	Page
Tableau I	Récapitulatif des principales classifications moléculaires pronostiques	9
Tableau II	Les références et identifiants des séquences moléculaires (refseq gene) du gène IL1B	20
Tableau III	Les bases des données	20
Tableau IV	Références Genbank « NCBI » du gène humain IL1B	39
Tableau V	Transcrits du gène IL1B	40
Tableau VI	Les isoformes de la protéine IL1B	40
Tableau VII	Séquences produisant un alignement significatif	41
Tableau VIII	Nombre d'étude cliniques des variantes du gène IL1B chez l'homme	42
Tableau IX	Les variantes de transcription du gène IL1B	46
Tableau X	Les régions qui correspondent les exons le plus de SNP	47
Tableau XI	Nombre de SNP par exon	48
Tableau XII	Classement des exons par nombre de SNP	48
Tableau XIII	Les cinq exons qui ont le plus grand nombre de SNP	48
Tableau XIV	Pathologies associées aux gènes IL1B	50

Introduction

Le cancer du sein est actuellement la première cause de morbidité et de mortalité chez les femmes algériennes (Mehemmai et al., 2019).

Il est considéré comme étant hormonodépendant car Les mutations génétiques peuvent conduire au cancer du sein et ont été liées expérimentalement à certains marqueurs tumoraux, notamment les statuts des récepteurs de la progestérone (PR), des récepteurs des œstrogènes (ER) et du récepteur du facteur de croissance épidermique humain de type 2 (HER2), qui sont utilisés en clinique dans la classification du cancer du sein(Perou et al., 2000).

Cette maladie a une susceptibilité génétique complexe. L'immunité innée et adaptative peut en outre augmenter le risque génétique de développement du cancer du sein(Kaarvatn et al., 2012) , d'autre part les cellules cancéreuses du sein ont la capacité de libérer des cytokines(Bachelot et al., 2003)(Bouma-Ter Steege et al., 2004).

L'interleukine (IL) -1 est une famille de cytokines endogènes impliquées dans l'inflammation et la réponse immunitaire, y compris l'IL-1 alpha (IL-1A), l'IL-1 bêta (IL-1B) et l'antagoniste des récepteurs de l'IL-1 (IL-1Ra) codé par IL-1RN (Bensen et al., 2001). L'interleukine-1 (IL-1) est connue pour jouer un rôle important dans la carcinogenèse du cancer du sein. L'expression de l'IL-1A et de l'IL-1B a été augmentée dans les tissus du cancer du sein humain. Les études actuelles se sont principalement concentrées sur les associations entre les polymorphismes du gène IL1B et le risque de cancer du sein(Zuo et al., 2018).

Dans notre étude nous allons chercher les SNP sur le gène IL1B qui sont en relation avec le cancer du sein en débutant par la récolte de toute les informations liées au gène ensuite la définition des SNP qui peuvent exister sur ce gène enfin la précision des SNP qui sont en relation avec l'apparition du cancer du sein , et à cause de confinement face au corona virus Ce travail a été réorienté vers une étude in silico en utilisant les outils bioinformatique .

Chapitre I : Rappels bibliographiques

1. Cancer du sein

Le cancer du sein est le cancer le plus fréquent chez les femmes et l'un des principales causes de décès parmi elles (Momenimovahed et Salehiniya., 2019). Selon les statistiques GLOBOCAN 2018, près de 2 millions de femmes ont reçu un diagnostic de cancer du sein avec 626679 décès liés (**Fig 01**). C'est une maladie multifactorielle (Zendehtdel *et al.*, 2018) et divers facteurs contribuent à son apparition. Bien que la maladie sévisse partout dans le monde, son incidence, sa mortalité et ses taux de survie varient considérablement d'une région à l'autre du monde, ce qui pourrait être dû à de nombreux facteurs tels que la structure de la population, le mode de vie, les facteurs génétiques et l'environnement (Hortobagyi *et al.*, 2005) Changements du risque facteurs ont conduit à une augmentation de la prévalence du cancer du sein, qui augmente chaque jour.

Les médecins diagnostiquent le cancer du sein et le classent morphologiquement à partir de l'analyse de certains critères : type histologique, taille de la tumeur, éventuel envahissement ganglionnaire, analyse de la présence de récepteurs hormonaux (œstrogènes et/ou progestérone), surexpression de la protéine HER2 (Human Epidermal Growth Factor Receptor-2).

Parmi les différents types de cancers du sein, les adénocarcinomes sont les plus courants (95%). Les cancers du sein se développent à partir des canaux (cancers canaux) et des lobules (cancers lobulaires) de la glande mammaire. Ils sont dits « in situ » lorsque les cellules cancéreuses sont confinées aux canaux et lobules, et « infiltrants » lorsque les cellules cancéreuses sont présentes dans les tissus qui les entourent (Lecarpentier et Andrieu., 2012)

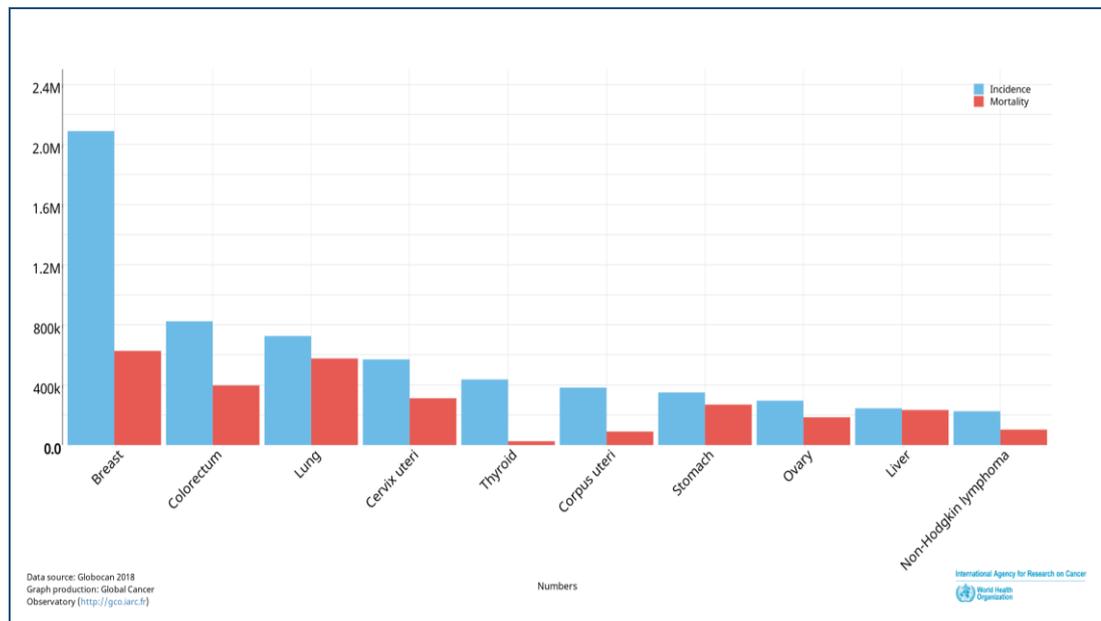


Figure 01 : Estimation du nombre de cas incidents et de décès dans le monde, femmes, tous âges. GLOBOCAN 2018. <https://gco.iarc.fr>

1.1. Cancer du sein en Algérie

Le cancer du sein dont la forte progression prend des proportions épidémiques inquiétantes avec plus de 11 000 nouveaux cas en 2018 et plus de 3000 décès liées, soit 54 nouveaux cas pour 100 000 femmes ; une augmentation de 500 nouveaux cas par an. De plus, le cancer du sein, en Algérie touche autant la femme jeune que la femme ménopausée, ce qui complique son dépistage : l'âge médian est de 47 ans c'est-à-dire que 50% des cancers du sein ont déjà eu lieu avant cet âge(Plan national CANCER 2015 - 2019, 2014).

L'augmentation de l'incidence du cancer du sein en Algérie coïncide avec "l'occidentalisation" du mode de vie de la femme algérienne (recul de l'âge au mariage, diminution du nombre d'enfants, âge tardif à la première grossesse, diminution de la période d'allaitement maternel, contraception orale fréquente et longue, augmentation de l'indice de masse corporelle)(H. Guendouz, W. Chetibi, A. Abdelouahab, 2011).

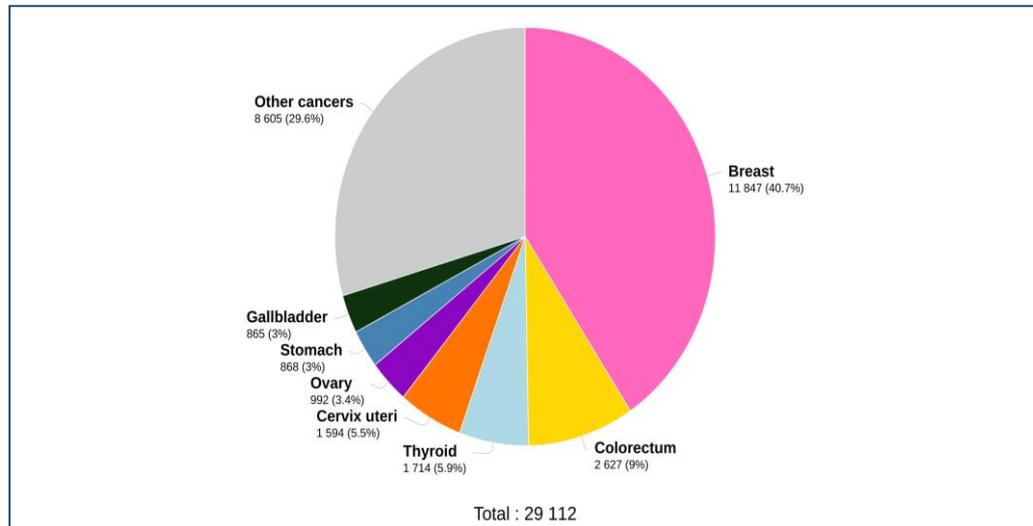


Figure 02 : Estimation du nombre de cas en Algérie, femmes, tous âges.
GLOBOCAN 2018. <https://gco.iarc.fr>

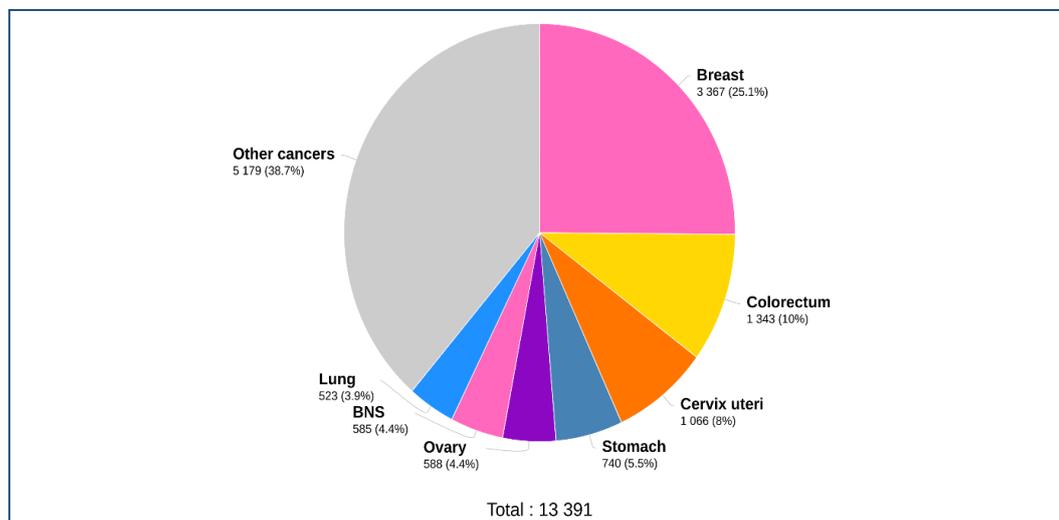


Figure 03 : Estimation du nombre de décès en Algérie, femmes, tous âges.
GLOBOCAN 2018. <https://gco.iarc.fr>

2. La formation de la tumeur primaire

Les cellules épithéliales subissent des altérations génétiques successives qui aboutissent à l'activation d'oncogènes et l'inactivation de gènes suppresseurs de tumeurs qui peuvent provoquer des modifications phénotypiques de l'épithélium mammaire telle que l'hyperplasie et la dysplasie. L'hyperplasie correspond à une prolifération cellulaire incontrôlée. Ces cellules deviennent

Rappels bibliographiques

métaphasiques, elles prolifèrent de façon anarchique et commencent à se transformer. La dysplasie correspond à une lésion précancéreuse avec une anomalie du renouvellement cellulaire associé à des degrés divers d'atypies, de différenciation et de modifications architecturales. Elle est due à un trouble acquis de l'homéostasie cellulaire résultant d'anomalies génétiques qui perturbent la division et la maturation cellulaire. Les cellules dysplasiques peuvent se transformer en cellules cancéreuses dans des délais très variables. L'état dysplasique peut être diagnostiqué par des examens cytologiques, histologiques et anatomopathologiques pour évaluer le pronostic, informer sur le grade d'avancement de la cancérisation et par suite adapter la thérapie.

Le tissu précancéreux présente des lésions histopathologies détectables avant l'apparition d'un cancer. La formation de la tumeur primaire se divise en deux étapes : la dysplasie et le carcinome in situ. Une fois que la taille de la tumeur dépasse les 2 mm de diamètre le système antigénique se met en place. Quand la tumeur atteint le stade final de dysplasie, elle commence le stade de carcinome in situ qui se définit comme étant une prolifération des cellules cancéreuses épithéliales sans parvenir à franchir la membrane basale, elle y reste confinée. Ceci est dû à la présence des éléments de cohésion intercellulaires et structures adhésives tels que l'E-cadhérine (Martin et Jiang., 2009) (Singhai et al., 2011). Il est quasiment impossible de distinguer histologiquement le carcinome in situ et la dysplasie de haut grade. Le dépistage est en effet très important à ce stade pour le pronostic, le traitement peut être local et efficace.

3. Signature moléculaire de cancer du sein

Les cancers sont classés aujourd'hui morphologiquement selon certains critères : taille, type histologique, capacité d'envahir les ganglions, présence des récepteurs hormonaux (œstrogène-ER et progestérone PR) et surexpression de récepteurs HER2.

En fonction de ces caractéristiques, le pronostic sera évalué et le traitement sera adapté. Le progrès des approches transcriptomiques et histopathologiques a permis une classification plus affinée. Les données transcriptomiques ont permis de distinguer six types de cancers du sein : les cancers luminaux A et B, normal-like, basal-like, claudin-low et HER2+ (Prat *et al.*, 2010) (Perou *et al.*, 2000).

Les cancers Luminaux A et B: Ces cancers se distinguent par l'expression des récepteurs aux œstrogènes (ER). Les cancers luminaux A sont de bas grade, ont tendance à se développer assez lentement, et ont un meilleur pronostic alors que les cancers luminaux B ont un niveau de multiplication généralement plus rapide que les cancers luminaux A et leurs pronostics ne sont pas aussi bons.

Rappels bibliographiques

Les cancers du sein dits normal-like : Sont caractérisés par un profil triple négatif (ER- ; PRet ne surexprimant pas HER2) et par l'expression de gènes observés dans le tissu mammaire normal et le tissu adipeux.

Les cancers de type basal: La plupart de ces cancers sont dit triple négatif, ce qui signifie qu'ils manquent d'expression d'œstrogène, de progestérone et de HER2. Ce type est plus fréquent chez les femmes présentant des mutations du gène BRCA1. Pour des raisons qui ne sont pas bien comprises, ces cancers sont plus fréquents chez les jeunes et les femmes afro-américaines.

Les cancers HER2+: Ces cancers ont des copies supplémentaires du gène HER2 et parfois quelques autres. Ces cancers ont tendance à croître rapidement et ont un mauvais pronostic, mais ils peuvent souvent être traités avec succès avec des thérapies ciblées visant à HER2. Une partie des cancers triple négatifs sont des « claudin-low » probablement issues des cellules souches. Les tumeurs de cette catégorie se caractérisent par un déficit d'expression en protéines « Claudin » : des composés importants dans la jonction entre cellules épithéliales. Les cancers associés au plus mauvais pronostic sont les basal-like, claudin-low et Her2+ (Dai *et al.*,2015). De point de vue clinique, les cancers peuvent être classés en trois groupes thérapeutiques

Les tumeurs ER+ : Ce groupe est le plus diversifié. Ces cancers sont caractérisés par l'expression des récepteurs estrogènes avec ou sans l'expression des récepteurs de progestérone et comprennent les cancers luminaux. Les luminaux A et B comptent à peu près 70% des cancers du sein (Schnitt, 2010). Comparées aux tumeurs de sous-type Luminal B, Les tumeurs lumineales A surexpriment les RE et les gènes régulés par les RE comme GATA-3, Fox A1... Les gènes de prolifération sont faiblement exprimés dans les tumeurs lumineales A, contrairement aux cancers luminaux B dans lesquels l'expression des gènes liés à la prolifération est élevée. Respectivement les cancers luminaux A et B présentent 13% et 66% une mutation du suppresseur de tumeurs P53.

Les tumeurs HER2+ : Des études in vitro et in vivo ont démontré le rôle de HER2 à différents niveaux du processus tumoral :

- la prolifération cellulaire ;
- la transformation maligne de lignées cellulaires en culture ;
- la mobilité cellulaire (élément déterminant dans le potentiel métastatique des cellules tumorales).

HER2 se comporterait donc comme un amplificateur de signaux de la croissance cellulaire par l'intermédiaire de son activité tyrosine kinase intrinsèque lors d'une liaison avec un autre

Rappels bibliographiques

récepteur de la famille HER ou lors de sa surexpression, par la formation possible de dimères HER2–HER2 (Ravanal et al., 2005).

15% des cancers du sein surexpriment HER2 (Figuroa-Magalhães *et al.*, 2014). L'expression de HER2 dans ces cancers est un marqueur de mauvais pronostic à cause de la capacité proliférative accélérée (Engel et Kaklamani., 2007) (Slamon *et al.*, 1987). Certains de ces cancers expriment les récepteurs hormonaux, récepteurs aux oestrogènes et la progestérone. Ces tumeurs surexpriment les gènes de l'amplicon HER2 (Vincent et Sigal., 2009) par exemple le GRB7 (Perou *et al.*, 200) (Dai *et al.*, 2014), PGAP3 (Dai *et al.*, 2014), TGFB1-induced antiapoptotic factor 1 et TNF receptor associated factor 4. 40 à 80% des tumeurs HER+ présentent une mutation de P53.

Les triples négatifs : Représentant 10 à 15% des cancers du sein (Ovcaricek *et al.*, 2010). Ils sont caractérisés par l'absence des récepteurs aux œstrogènes, à la progestérone et Her2. Ils englobent les sous types basaux et les claudin-low qui partagent quelques caractéristiques communes comme la sous-expression des marqueurs luminaux comme les cytokératines Cks 8 et 18 et la surexpression des marqueurs basaux comme les Cks 5, 14 et 17. Les claudin-low sont plus riches en marqueurs de l'EMT par exemple, la perte d'expression des protéines de jonctions cellulaire de type E-cadhérine, Claudine 3,4 et 7 et l'acquisition des caractéristique de « souchitude » ou la signature des cellules initiatrices des tumeurs : CD44+/CD24- (Prat *et al.*, 2010). Les tumeurs de type triple négatif sont le plus souvent des carcinomes canauxaires infiltrants, mais le carcinome canalaire in situ peut aussi être triple négatif. Les femmes présentant des mutations du gène BRAC1 sont plus à risque de développer un cancer du sein de type basal. 82% des cancers de cette classe présentent une mutation de P53.

1.3 Méthodes employés dans la signature

Les méthodes employées dans le diagnostic par les signatures moléculaires afin d'aider le pronostic clinique sont : qRT-PCR, microarray ou dérivés type n-counter.

Ces Méthodes ont les mêmes objectifs : calculer un score pronostique, fondé sur les niveaux d'expression de gènes impliqués dans la cancérogenèse, et si possible prédire la réponse au traitement. Ces méthodes sont applicables essentiellement aux tumeurs lumineales RE+, elles ont prouvé leur valeur pronostique dans de vastes essais prospectifs, et les experts souhaitent les intégrer dans la décision thérapeutique, actuellement établie sur les critères clinicopathologiques (Penault-Llorca et Lacroix-Triki, 2017).

Tableau I : Récapitulatif des principales classifications moléculaires pronostiques

	Nom					
	70-gene signature MammaPrint® (Agendia)	21-gene signature Oncotype DX® (Genomic health)	PAM50 Prosigna (Nanostring)	Genomic Grade MapQuant DX™ (Ipsogen/Haliidx)	HOXB13: IL17BR BCI (biotheranostics)	11-gene assay EndoPredict® (Myrid Genetics)
Méthode	Microarray	qRT-PCR	n-counter	Microarray qRT-PCR	qRT-PCR	qRT-PCR
Matériel	Cryo/FFPE	FFPE	FFPE	Cryo/FFPE	FFPE	FFPE
Gènes analysés	70 gènes	ER, PR, BCL2, SCUBE2, Ki67, STK15, BIRC5, CCNB1, MYBL2, HER2, GRB7, MMP11, CTSL2, GSTM1, CD68, BAG1	50 gènes	97 gènes	HOXB13, IL17BR, BUB1, CENPA, NEK2, RACGAP1, RRM2	DHCR7, AZGP1, MGP, STC2, BIRC5, UBE2C, RBBP8, IL6ST
Valeur pronostique	Récidive (5 ans)	Récidive (10 ans)	Récidive (10 ans)	Récidive	Récidive (5 et 10 ans)	Récidive (10 ans)
Indications	RE+/N- ou N+ (1-3) RE-/N- ou N+ (1-3)	RE+/HER2-/N-/HT RE+/HER2- /N+ (1-3)	RH+/HER2- N- ou N+	RE+/N- (grade 2) sous tamoxifène	RE+/N- sous tamoxifène	RE+/ HER2- N- ou N+ (1-3) sous HT
Résultats		RS = 0 à 100	Type moléculaire ROR = 0 à 100		0 à 10	0 à 15
	Haut	Haut > 30	Haut	Haut	Haut	Haut
	Bas	Intermédiaire	Intermédiaire	Équivoque	Intermédiaire	Bas
Essai prospectif	MINDACT	TAILORx RxPONDER		ASTER 70s		

qRT-PCR: transcriptase inverse-réaction en chaîne polymérase quantitative; **Cryo:** cryoconservation; **FFPE:** Formol fixé et inclus en paraffine; **RE (+ ou -):** statut des récepteurs aux œstrogènes; **N (+ ou -):** statut des ganglions lymphatiques; **HER2:** récepteur 2 du facteur de croissance épidermique humain ; **HT:** hormonothérapie; **RH (+ ou -):** statut des hormonaux; **RS:** score de récurrence; **ROR:** score de risque de récurrence; **CT:** chimiothérapie; **MINDACT:** Les puces à ADN dans la maladie des ganglions lymphatiques Node-négatifs et 1-3 positifs peuvent éviter la chimiothérapie; **TAILORx:** Essai d'attribution d'options individualisées pour le traitement Rx; **RxPONDER:** Rx pour Positive Node, cancer du sein endocrinien sensible ; **ASTER 70s:** Traitement systémique adjuvant du carcinome du sein HER2-négatif positif (ER) chez les femmes de plus de 70 ans selon le grade génomique.(Penault-Llorca et Lacroix-Triki, 2017)

4. Cas de cancer hormonodépendant

La glande mammaire normale est un tissu cible pour l'action d'hormones variées, et en particulier les œstrogènes et la progestérone. Ces cellules contiennent des protéines particulières appelées récepteurs hormonaux (**Fig 04**).

Dans le cas des hormones stéroïdes (œstrogène et progestérone), ces récepteurs sont cytoplasmiques. Pour la prolactine, en revanche, ils sont membranaires. Les récepteurs ayant capté l'hormone spécifique, le complexe hormone-récepteur va dans le noyau où il va induire l'action

Rappels bibliographiques

spécifique de l'hormone. Par ailleurs, les œstrogènes (œstradiol) vont stimuler la resynthèse de leur propre récepteur et augmenter la synthèse des récepteurs de progestérone (PgR). À l'opposé, la synthèse du récepteur à œstrogènes est bloquée par la présence de progestérone. Ainsi peut-on définir à l'étage cellulaire l'hormonodépendance d'une tumeur (J. Lansac, 2018).

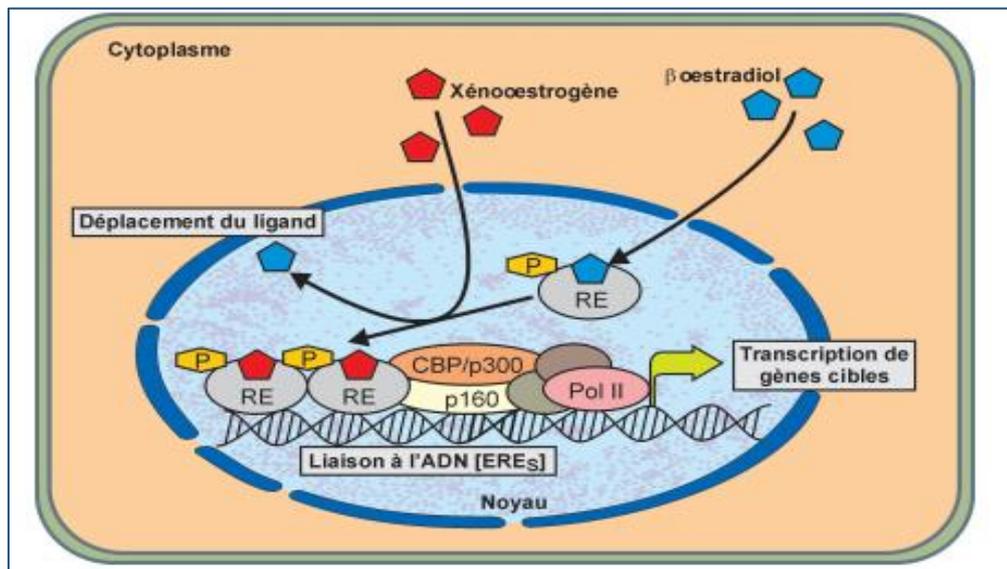


Figure 4 : Activation illégitime du récepteur des œstrogènes (Massaad et Barouki., 1999)

RE : récepteur des œstrogènes ; **ERE** : élément de réponse d'œstrogène ; **Pol II** : Polymérase 2 ; **CBP** : CREB- binding protein

5. Les facteurs de risques

Ils sont souvent à tort considérés comme des facteurs devant jouer un rôle dans le processus carcinogénétique. En réalité, leur seule caractéristique est de présenter une liaison statistique significative avec la maladie. Leur identification a un double intérêt : servir de base à l'élaboration d'hypothèses explicatives à vérifier par des études expérimentales, repérer un sujet qui peut faire l'objet d'une surveillance accrue. Certains ne sont pas modifiables : l'âge, les antécédents familiaux ; d'autres le sont : l'obésité, les traitements hormonaux, l'alcool (J. Lansac, 2018).

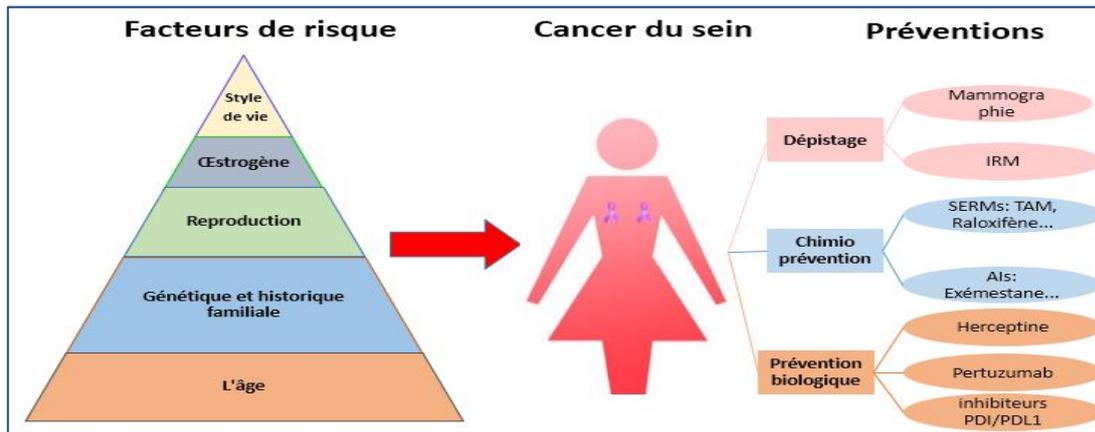


Figure 05 : Diagramme schématique des facteurs de risque et de la prévention du cancer du sein.

L'âge, les antécédents familiaux, les facteurs reproductifs, les œstrogènes et le style de vie sont cinq facteurs de risque importants de cancer du sein, représentés dans le diagramme pyramidal. Le dépistage (mammographie et IRM), la chimioprévention (avec SERM et IA) et la prévention biologique (utilisant Herceptin et pertuzumab) sont actuellement utilisés pour prévenir le cancer du sein. Les inhibiteurs de PD1 / PDL1 sont des médicaments d'immunothérapie et pourraient être des stratégies prometteuses dans le traitement du TNBC. (Sun *et al.*, 2017)

6. Polymorphisme IL1 et le cancer du sein

La famille IL-1 de cytokines est codée par deux gènes séparés, IL1A et IL1B sont situés dans une région de 430 kb sur le chromosome 2q14.2 dans un cluster (**Fig 06**) et contient également les gènes des récepteurs IL1 de type 1 et 2 (IL1R1 et IL1R2), qui sont connus pour être la régulation clé de l'inflammation et de la réponse immunitaire aux cytokines multifonctionnelles et qui affectent presque tous les types de cellules (Bensen *et al.*, 2001).

Dans de nombreux types de tumeurs, on pense que l'IL-1 est régulée à la hausse et est exprimée en tant que facteur de progression tumorale par métastase et expression de gènes angiogéniques et de facteurs de croissance. Il a récemment été suggéré que l'IL-1 joue un rôle dans le développement du cancer du sein (Zuo *et al.*, 2018).

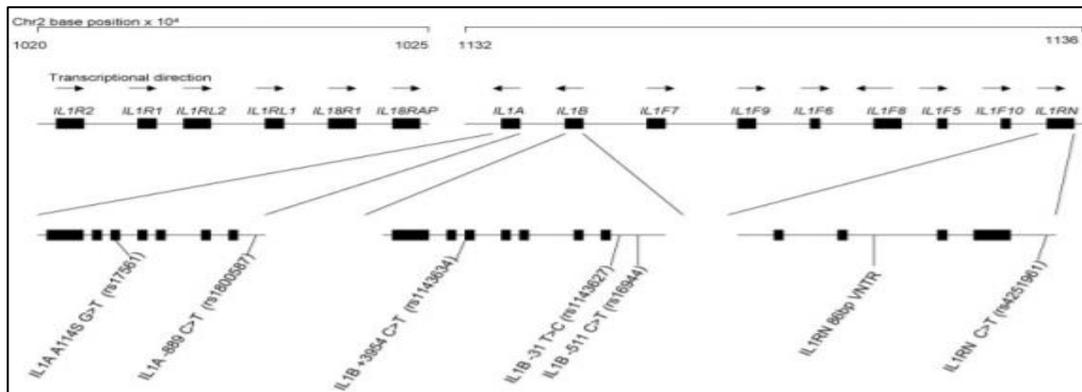


Figure 06 : Gènes de la famille de l'IL-1 sur le chromosome 2 et principaux SNPs fonctionnels. (Smith et Humphries., 2009).

6.1. Définition du polymorphisme génétique

Par convention, une variation rare (affectant par définition moins de 1 % des chromosomes d'un individu) est appelée une mutation et une variation fréquente (à partir de 1 %), un polymorphisme ou variant génétique. Sur les trois milliards de bases du génome humain, plus de 99 % sont identiques d'un individu à l'autre. Les autres représentent les variantes génétiques ou polymorphismes. Ce pourcentage peut paraître négligeable, mais représente pourtant environ trois millions de nucléotides qui sont à la base des différences entre les individus.

Le polymorphisme le plus fréquent est le SNP. Il s'agit d'une variation simple et unique dans la séquence nucléotidique d'un gène, retrouvée chez plusieurs individus d'une même espèce et qui se transmet de façon mendélienne. Ces SNP représentent 90 % de l'ensemble des variantes génétiques. La plupart sont bi-alléliques et se caractérisent par la fréquence d'un allèle minoritaire ou variant. On classe ces SNP en fonction de la modification nucléotidique (Courivaud, Saas et Ducloux., 2012).

Des différentes maladies associées aux SNP, qui sont exprimées dans toutes les cellules d'un organisme, certaines maladies, comme le cancer, sont causées par des variations génétiques typiques d'un petit sous-ensemble de cellules somatiques (Karki et al., 2015).

6.2. Gène d'IL1B

Le gène couvre une région de 7,5 kb et la partie codante est divisée en sept exons (Giraldo et al., 2008).

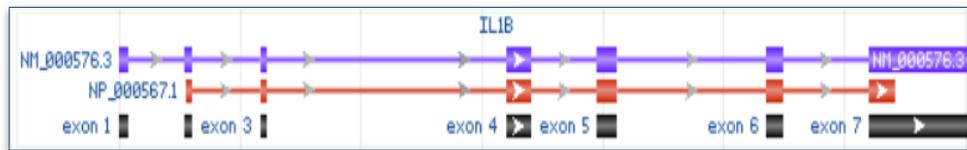


Figure 07 : Gène d'IL1B (<https://www.ncbi.nlm.nih.gov/gene/3553>)

6.2.1. Structure et localisation

L'IL-1bêta est située sur le chromosome 2 à l'emplacement 113 303 808-113 310 827. Il est situé au milieu de l'IL-1A et de l'IL1-RN entre 40 et 110 kb de l'IL-1A. Le gène pro IL-1 β est composé de sept exons avec une longueur de produit de transcription primaire de 7 008 nucléotides (Giraldo et al., 2008).

6.2.2. Expression

Expression biaisée dans la moelle osseuse (RPKM 72.5), l'appendice (RPKM 32.9) et 12 autres tissus.

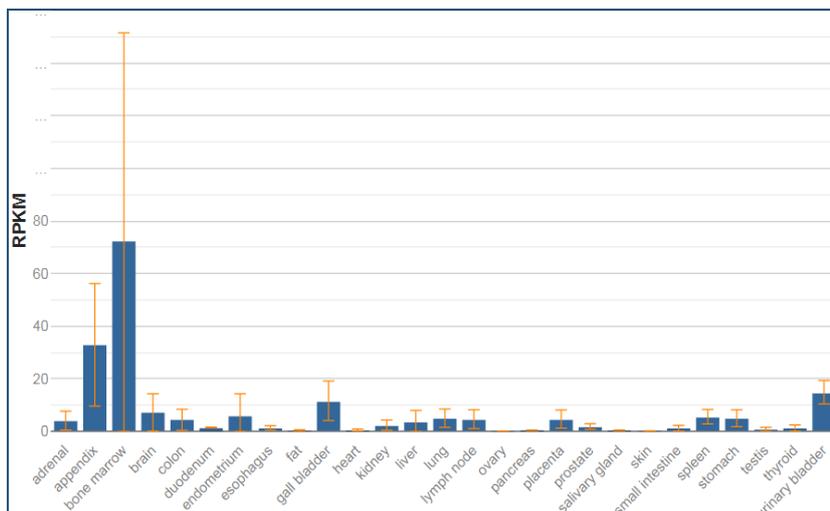


Figure 08 : Expression d'IL1B dans les organes humaines (<https://www.ncbi.nlm.nih.gov/gene/3553>)

6.2.3. Biosynthèse et activation

IL-1 β est produite principalement par des monocytes sanguins, les macrophages tissulaires, les cellules dendritiques et la peau microglie du cerveau (Garlanda, Dinarello et Mantovani., 2013). La synthèse d'IL-1 β se produit lors de la stimulation par des modèles moléculaires associés à des motifs (PAMP), tandis que son activation est liée à des stimuli secondaires, tels que des modèles moléculaires associés aux dommages (DAMP), en présence d'agents pathogènes envahissants et sur des signaux de danger stériles (Brough, Tyrrell et Allan., 2011) (Baroja-Mazo et al., 2014). L'IL-1 β est traitée par l'inflammasome, un complexe multiprotéique activé via la famille des récepteurs de type diabétique non obèse (NOD) (NLR), les protéines contenant un domaine pyrine (NLRP) et nécessaire à l'activation de la caspase 1, nécessaire pour le clivage de l'IL-1 β et activation (Miller et al., 1993) (Li et al., 1995) (Fig 09). La pyroptose médiée par la caspase-1 est une forme lytique de mort cellulaire et un mécanisme effecteur immunitaire inné. Au cours de l'inflammation, après le clivage et l'activation médiés par la caspase-1, l'IL-1 β est libérée d'une

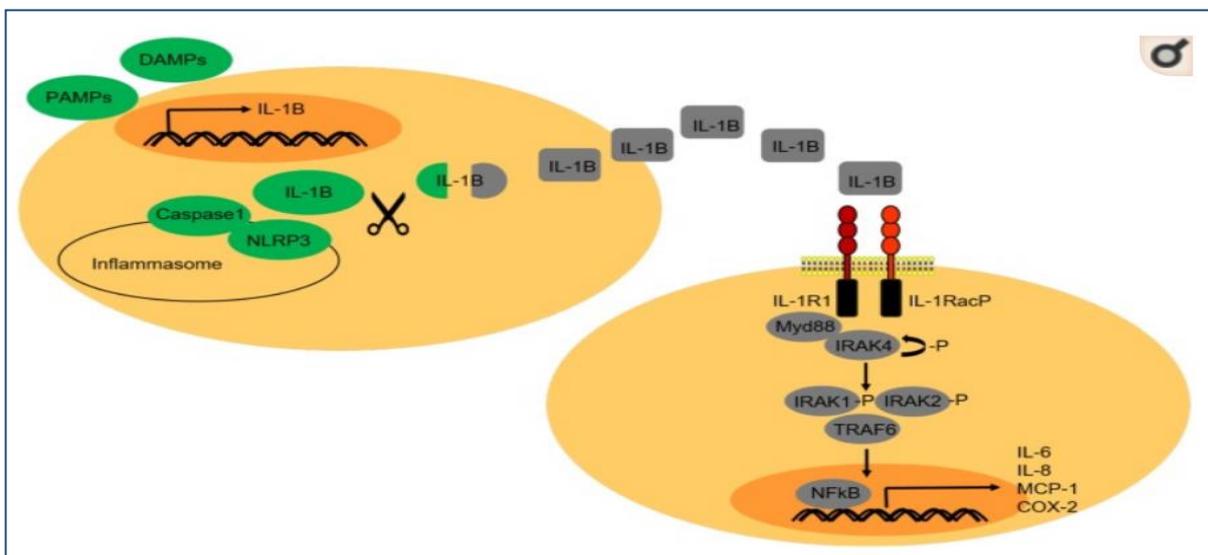


Figure 09 : Signalisation IL-1 β . (Tulotta et Ottewell., 2018)

manière sécrétoire " non conventionnelle " encore mal comprise (Vince et Silke., 2016). Une hypothèse simple suggère que l'IL-1 β activée est libérée passivement après la pyroptose induite par la caspase 1, soulignant le lien entre la mort cellulaire et l'inflammation.

6.2.4. Régulation

L'expression d'IL-1 β est induite et strictement associée à l'inflammation. Pour cette raison IL-1 β est majoritairement exprimée in vivo par les monocytes et macrophages activés. La présence de séquences remarquables en cis reconnues par des facteurs de régulations trans en amont des

Rappels bibliographiques

promoteurs de ces interleukines présage de l'existence de nombreuses régulations de leur expression (**Fig 10**) (Roussel-Vavasseur., 2008).

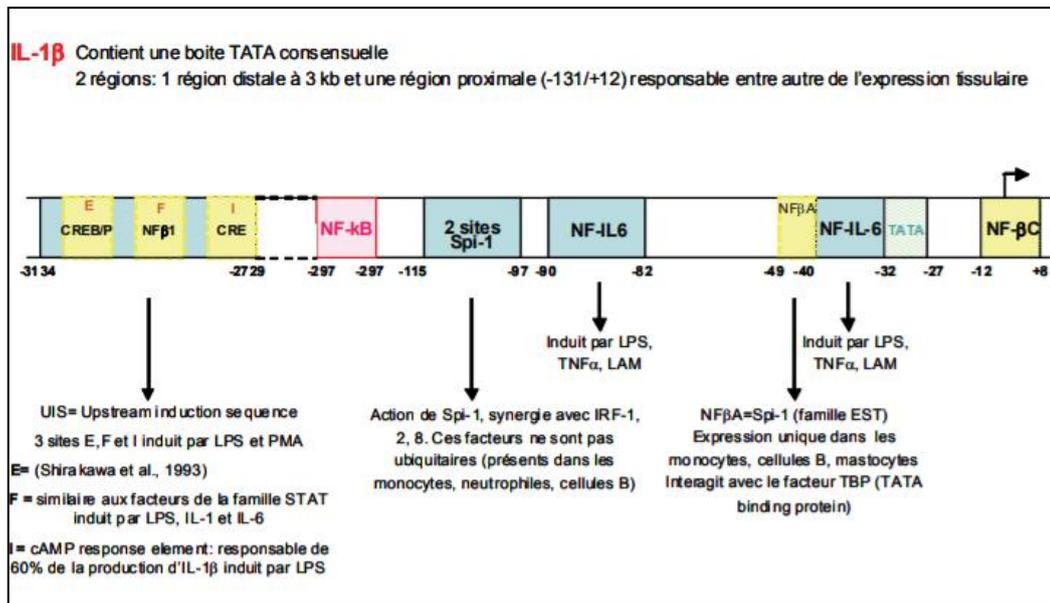


Figure 10 : Régulations transcriptionnelles d'IL-1β au niveau de leur promoteur. (Roussel-Vavasseur., 2008)

6.2.5. Homologie

IL-1beta partage une homologie de séquence avec IL-1A et IL1-RN. IL-1B et IL-1A ont une homologie d'acides aminés entre eux d'environ 25% (Giraldo et al., 2008).

6.2.6. Protéines du gène de l'IL1B et fonction

La protéine codée par ce gène est un membre de la famille des cytokines de l'interleukine 1 qui contient 269 acides aminés.

L'IL-1B induit l'expression de la cyclo-oxygénase (COX) -2 et de l'oxyde nitrique synthase inductible (iNOS). L'IL-1B augmente également l'expression de presque toutes les autres cytokines, telles que le facteur de nécrose tumorale (TNF) et l'IL-6, et les chimiokines ainsi que les molécules d'adhésion. De plus, l'IL-1B augmente l'expression de diverses protéases tissulaires et métalloprotéases matricielles et inhibe la synthèse des protéoglycanes (Dinarello., 2005).

6.3. Polymorphisme génétique d'IL1B

Le gène IL-1B est hautement polymorphe et transitions de bases entre C et T aux positions -511 (C>T; dbSNP: rs16944), -31 (T>C; dbSNP: rs1143627) et +3954 (C>T; dbSNP: rs1143634)

Rappels bibliographiques

paires de bases du site transcriptionnel ont été largement rapportés. En particulier, les deux premiers polymorphismes sont localisés dans la région du promoteur et présentent un fort déséquilibre de liaison (El-Omar et al., 2000)(Machado et al., 2001).

La substitution IL-1B -31T / C provoque la perturbation d'un motif de boîte TATA et a été trouvée pour affecter de manière marquée l'affinité de liaison de plusieurs facteurs de transcription (El-Omar et al., 2000) , (Chen et al., 2006) , (Lind, Haugen et Zienolddiny., 2007) et affecter ainsi l'activité de transcription de l'IL -1B (Chen et al., 2006) . L'IL-1B +3954 C / T dans l'exon 5 est un polymorphisme nucléotidique unique (SNP). Des études in vitro ont montré que -511T et + 3954T sont associés à une augmentation de la sécrétion d'IL-1 β à partir de la sécrétion de protéine IL-1 β induite par le lipopolysaccharide (LPS) (Hall et al., 2004) (Hernandez-Guerrero et al., 2003).

6.3.1. Association polymorphisme d'IL1B et cancer du sein

L'IL-1 β agit comme une cytokine pro-inflammatoire et est principalement produite par des monocytes et des macrophages activés (Xu et al., 2013). Dans le développement du cancer du sein, l'IL-1 β régule à la hausse la cyclooxygénase-2, augmente l'expression du facteur de croissance endothélial vasculaire dans les cellules tumorales et participe à l'angiogenèse causée par une inflammation chronique (Esquivel-Velázquez et al., 2015). De plus, l'IL-1 β induit la prolifération des cellules cancéreuses du sein en augmentant l'activité aromatasase qui convertit les androgènes en œstrogènes et en augmentant l'activité stéroïde sulfatase, qui convertit les œstrogènes conjugués en œstrogènes à action libre. De plus, la leptine et l'IL-1 β , qui sont des adipocytokines pro-inflammatoires, ont également un effet direct sur les cellules épithéliales du sein pour augmenter la production d'aromatasase et favoriser la croissance tumorale, la migration et l'invasion par des facteurs de croissance (Perrier, Caldefie-Chézet et Vasson., 2009).

Plusieurs études ont été menées sur la base de cette association étroite entre les niveaux d'IL-1 β et l'invasivité, l'agressivité et le grade tumoral élevé du cancer du sein (Escobar et al., 2015)(Kurtzman et al., 1999). Une étude a montré que l'IL-1 β était présente dans 90% des carcinomes mammaires invasifs, et il a été suggéré que des niveaux élevés d'IL-1 β étaient associés à une invasion tumorale et à une agressivité tumorale (Jin et al., 1997). Une autre étude a indiqué que les tumeurs de grade III avaient des niveaux d'expression d'IL-1 β plus élevés que les tumeurs de grade I et II, et que la protéine-1 β inflammatoire des macrophages était associée à l'inflammation (Chavey et al., 2007).

Rappels bibliographiques

Sur la base du rôle important de l'IL-1 β dans le développement du cancer du sein, un certain nombre d'études ont rapporté l'association des polymorphismes de l'IL-1 β avec le risque de cancer du sein. En raison de la petite taille relative d'une seule étude, les résultats sont incohérents (Ito et al., 2002) premier rapport, rs1143627 est significativement associé au risque de cancer du sein (CC vs TT: OR = 1,82, IC à 95% = 1,03–3,23), en raison de la petite taille de l'échantillon (231 cas et 186 témoins), une validation supplémentaire diverses populations devraient être faites. Une autre étude cas-témoins de (Liu et al., 2006b) ont soutenu l'association dans une population chinoise (CC vs TT: OR ajusté = 1,72, IC à 95% = 1,16–2,54).

En revanche, une analyse cas-témoins sur une population chinoise a suggéré que le génotype -511 TT était un facteur de risque de cancer du sein (Liu et al., 2006b). De même, le rapport précédent sur le risque associé au polymorphisme de l'IL-1b (+3954 C> G) est incohérent. La présence d'allèles mutants et de génotypes au site +3954 C> T est associée à un risque accru de cancer du sein dans une population indienne (Pooja et al., 2012). Seules deux autres études ont analysé ce polymorphisme, dont l'une a rapporté une association hautement significative entre l'allèle mutant et le phénotype agressif du carcinome du sein défini par le grade histologique élevé, les métastases des ganglions axillaires et la taille de la tumeur (Snoussi et al., 2005). La deuxième étude a analysé sa corrélation avec la survie des patientes atteintes d'un cancer du sein (Grimm et al., 2009). Il a été constaté que la présence d'un génotype mutant homozygote était associée à la survie globale en analyse univariée mais pas en analyse multivariée. L'étude de (Pooja et al., 2012) est le premier rapport suggérant l'association de l'allèle mutant et des génotypes de l'IL-1b (+3954 C> G) avec le risque de cancer du sein.

6.3.2. Association IL1B et le statut hormonal dans le cancer du sein

Il a été démontré que l'œstrogène influence la production de cytokines pro-inflammatoires; Des études sur l'IL-1B et l'IL1RN et les œstrogènes ont montré une réduction de la production d'IL-1B en réponse à une exposition aux œstrogènes (Rogers et al., 2007). D'autres ont suggéré que la famille IL-1, en particulier IL1RN, est un médium important du microenvironnement du cancer du sein en raison de sa relation avec le statut ER et PR (Miller et al., 2000); Dans l'étude de (Slattery et al., 2014), IL1RN était associé à des tumeurs ER + / PR +.

Il a été découvert que l'IL-1 β se combine avec le récepteur des œstrogènes (ER) α , qui est un facteur pronostique du cancer du sein, entraînant une activation transcriptionnelle dans les cellules cancéreuses du sein (Speirs et al., 1999). (Honma et al., 2002) ont montré que l'IL-1 β pouvait réguler la prolifération des cellules cancéreuses du sein grâce à la production d'œstrogènes par des enzymes catalysant les stéroïdes dans le tissu.

Chapitre II : Matériel et méthodes

2. Matériels et méthodes

Ce présent travail est une thématique proposé par Mme Eddaikra qui s'est déroulé pendant la période de confinement liée à la pandémie covid19 du mois d'Avril jusqu'au mois de juillet 2020.

Rappelons que notre objectif principal était l'étude du polymorphisme du gène interleukine1B(IL1B) et ses associations pathologiques au cancer du sein en utilisant les bases de données internationales et les outils bio-informatique open source.

Initialement notre travail « étude cas témoins » devait se réaliser à l'institut pasteur pour le génotypage en collaboration avec Dr Zaabat N service d'immunologie et le CHU frantz fanon service d'oncologie avec la collaboration du Pr CAID.

Sachant que nous ne pouvions ne pas réaliser notre stage pratique « extraction d'ADN, étude phénotypique et génotypique » suite aux conditions de pandémie COVID19. Notre travail a été réorienté vers une étude in silico bioinformatique.

Notre étude in silico est basée sur l'utilisation des bases de données protéomique et génomique des bases de données open source pour la recherche du polymorphisme du gène IL1B associés au cancer du sein et ses interaction hormonal.

Nous tenons à préciser que le suivi de notre formation par notre promotrice s'est réalisé par visioconférence pendant la durée de confinement.

2.1. Matériel

2.1.1. Matériel biologique

Pour répondre à notre problématique, nous avons adopté une méthodologie in silico basée sur l'interrogation de bases de données open source : NCBI, UCSC genome browser, le navigateur du génome Ensembl et Galaxy.

Les références et identifiants des séquences moléculaires (refseq gene) du gène IL1B (Interleukine 1 Beta) que nous avons utilisées dans notre étude in silico sont présentés dans le tableau II.

Nous avons utilisé des séquences nucléotidiques transcrites et protéiques (format Fasta), ce format nous permis de retrouver par la suite les SNPs avec leurs rs (références snip) associés aux au cancer du sein.

Tableau II : Références et identifiants des séquences moléculaires (refseq gene) du gène IL1B

NCBI Reference Sequence	NA	Format	Source	nbp	Exon	Début et fin de la séquence
IL 1B (refseq gene)	NM_000576.3	Fasta	NCBI (GenBank)	7,025 bp	7	Chr 2:113,587,332-113,594,356
IL 1B (refseq proteine)	NP_000567.1	Fasta	NCBI (Genpep)	269 aa	-	-

2.1.2. Matériel bio-informatique

Notre étude a nécessité l'utilisation du matériel informatique (ordinateur portable) et d'une connexion internet à haut débit.

Pour l'analyse de nos échantillons moléculaires nous avons utilisé des interfaces web, des logiciels, des softwares qui se base sur des algorithmes. Ces outils bioinformatiques sont répertoriés dans le tableau III.

Tableau III : Bases des données

Bases de données	But d'utilisation	url
NCBI	Recherche de séquences et études cliniques.	https://www.ncbi.nlm.nih.gov/
UCSCgenome browser	Visualisation du gène filtrer les données des SNP codants.	https://genome.ucsc.edu

Matériel et Méthodes

Ensembl	Recherche de gène, de variantes et de phénotypes.	https://www.ensembl.org/index.html
Usegalaxy	Recherche biomédicale	https://usegalaxy.org/
ALFRED	Recherche de pathologies associées	https://alfred.med.yale.edu/alfred/index.asp
Disgenet	Recherche de pathologies associées	http://www.disgenet.org
Genemania	Interactions moléculaires	https://genemania.org
STRING	Interaction protéine-protéine	https://string-db.org

2.2. Méthodes

Le protocole de notre étude est schématisé par l'organigramme (**Fig 11**).

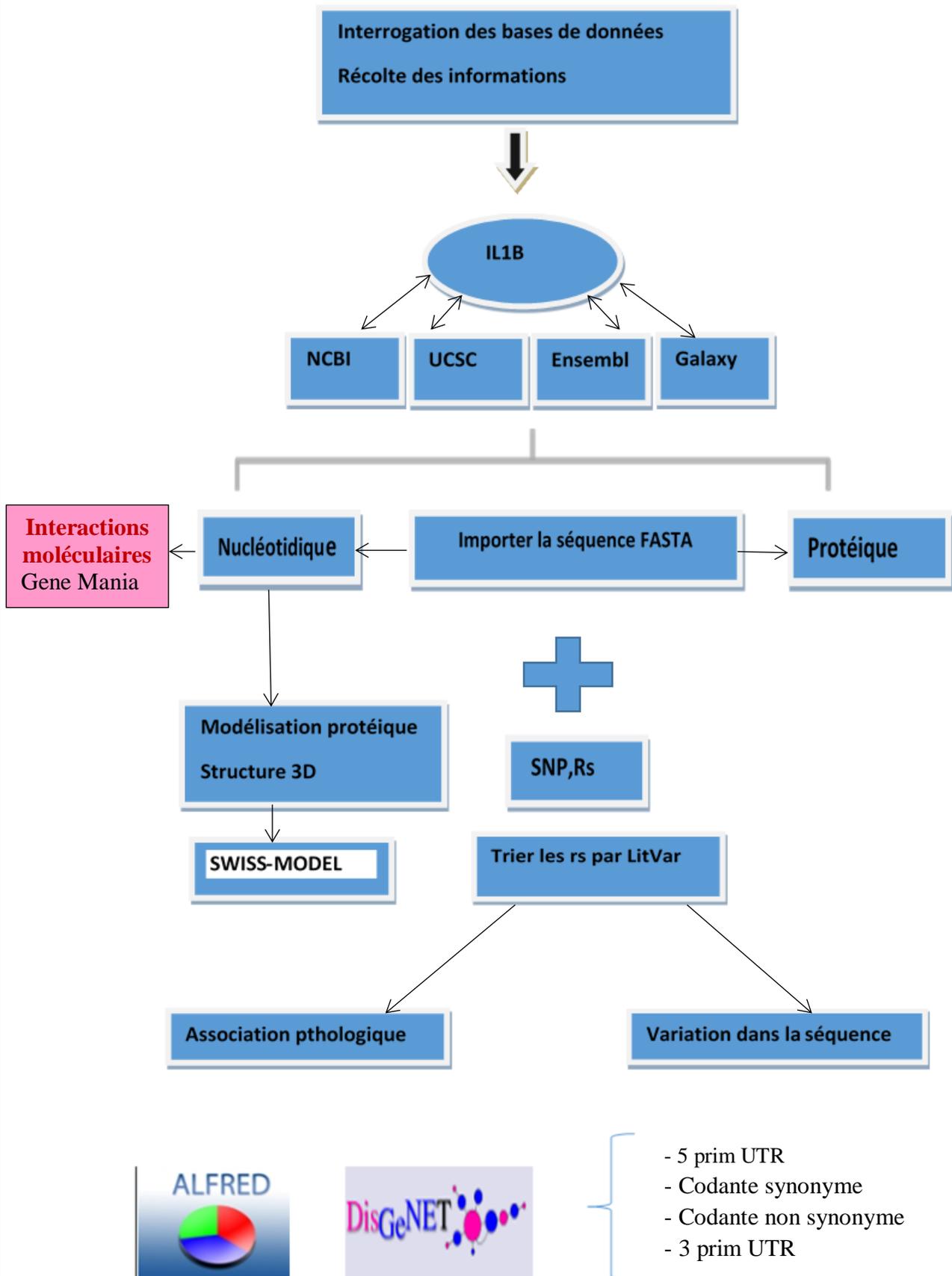


Figure 11 : Méthode de recherche des SNP du gène IL1B

2.2.1. Interrogation des bases de données

2.2.1.1. Outil NCBI

Principe

L’initiation de notre travail a sollicité l’Interrogation de la base de données NCBI (*National Center for Biotechnology Information*). C’est une base de données d’Information de littérature biomédicale et de biomolécule. (<https://www.ncbi.nlm.nih.gov/>).

Méthode

- Accéder à <https://www.ncbi.nlm.nih.gov/>
- Lancer la recherche en introduisant notre mot clé caractérisé par le symbole de notre gène « IL1B ». Plusieurs possibilités vont nous permettre d’accéder au polymorphisme du gène d’intérêt
- Choisir l’onglet « gène ». Puis choisir et cliquer sur « IL1B Homo sapiens ». Une autre page va s’ouvrir indiquant toutes les informations sur le gène d’intérêt (**Fig 12**).

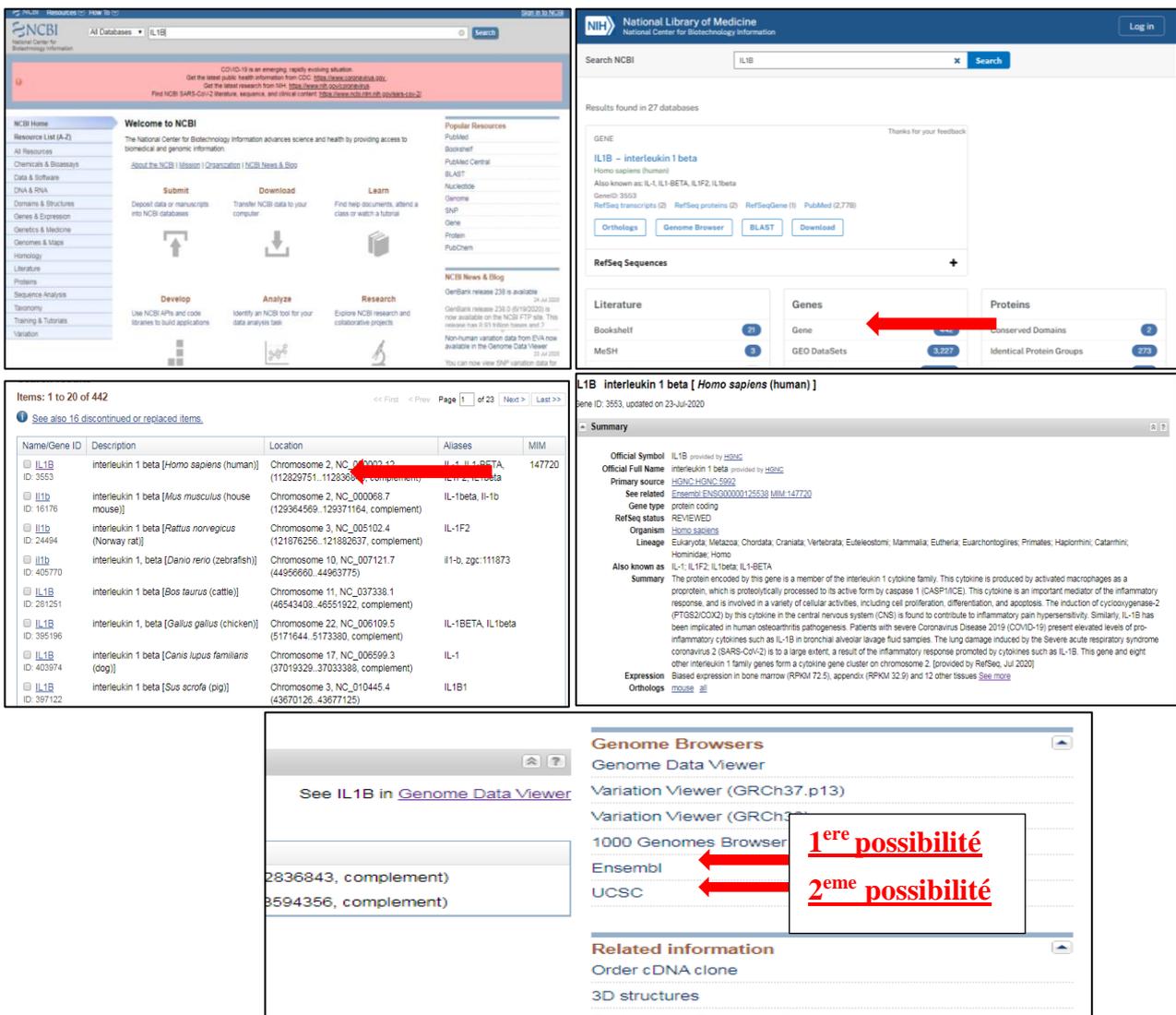


Figure 12 : les étapes de recherche dans le site NCBI

2.2.2. Recherche des polymorphimes « SNPs »

Il existe plusieurs possibilités pour accéder au polymorphisme du gène d'intérêt « IL1B » plus particulièrement les SNPs.

Afin de répondre à notre problématique nous avons choisi trois possibilités : UCSC, Ensembl, et galaxy.

2.2.2.1. Outil UCSC

Principe

Le navigateur de génome d'UCSC est un navigateur de génome en ligne et téléchargeable hébergé par l'Université de Californie à Santa Cruz.

Le site Web sur la bio-informatique du génome de l'Université de Californie à Santa Cruz (UCSC) à l'adresse <http://genome.ucsc.edu> fournit des liens vers divers outils d'analyse du génome, notamment le Navigateur de génome d'UCSC, un outil graphique permettant de visualiser une région spécifique d'un génome et une collection de « pistes » d'annotations alignées. Un autre outil du site Web, le navigateur de tables UCSC, fournit un accès pratique aux tables de la base de données MySQL sous-jacent aux annotations du navigateur de génome. Les deux navigateurs prennent en charge une fonctionnalité de suivi des annotations personnalisée qui permet aux utilisateurs de télécharger leurs propres données à des fins d'affichage et de comparaison (Kent *et al.*, 2010).

Méthode

- Accéder à <https://genome.ucsc.edu>
- sélectionner : le taxon, l'espèce
- lancer la recherche en introduisant le symbole du gène « IL1B »
- simplifier l'affichage en masquant toutes les pistes de donnée en appuyant sur « hide all », puis on active les pistes suivantes en cliquant sur pack : common SNPs \ ALL SNPs \ Flagged SNPs pour variation, et Old UCSC genes \ UCSC all events pour Genes. Puis on clique sur actualiser (refresh).

Matériel et Méthodes

On obtient un très grand nombre de SNP avec leurs numéros à l'écran, dont : les rouges sont des SNPs codants non synonymes, les verts sont des SNPs codants synonymes, les bleus sont les SNPs des régions d'épissage ou des régions non traduites et les noirs représentent les SNPs des régions ironiques.

The figure illustrates the workflow for searching SNPs on the UCSC Genome Browser. It shows the main interface, a genomic track view, the 'Regulation' and 'Variation' tracks, and the 'Table Browser' configuration for selecting fields from a specific SNP set.

Figure 13 : Méthode de recherche des SNPs sur le navigateur UCSC. <https://genome.ucsc.edu>

- Sélectionner ;
- ✓ Assembly : GRCh38/hg38
- ✓ Group : Variation
- ✓ Track : Common SNPs(151)
- ✓ Table : snp151Common
- ✓ Region : position (ces coordonnées seront la position du gène (IL1B))
- ✓ Pour limiter à quelques colonnes, choisissez : Output format : “selected fields from primary and related tables”
- ✓ Appuyer sur “get output”
 - Une deuxième table apparaît
- Sélectionner uniquement ces champs : bin, chrom, chromStart, chromEnd, name, score, observed, func, allele Freq Count.
- Appuyer sur “get output”
- On obtient une liste des SNPs du gène IL1B

2.2.2.2. Outil Ensembl genome browser

Principe

Le projet Ensembl <https://www.ensembl.org/index.html> a été initialement lancé en 1999 dans le but de développer des méthodologies pour l'annotation automatique de la séquence génomique (humaine) avec les gènes et leurs transcrits constitutifs.

Il fournit une multitude de données génomiques librement disponibles et accessibles pour les chercheurs en génétique, en génomique et en biologie moléculaire. Ce navigateur nous permet de filtrer les données des SNP codants : promoteur \ exon du gène IL1B (Newman *et al.*, 2018).

Méthode

- Accéder à la page d'accueil du navigateur Ensembl genome browser via le lien <http://www.ensembl.org> (Fig 14)

Matériel et Méthodes

- Choisir l'espèce « human » et introduit le symbole de notre gène d'intérêt « IL1B » dans la zone de recherche.

- Par la suite une page s'affiche avec toutes les informations concernant le gène IL1B : description, synonyme du gène, localisation du gène, et les transcrit du gène.

- Cliquer sur « séquences » nous obtenons la séquence FASTA nucléotidique ou protéique des exons ou des séquences codante synonyme et non synonyme, des introns, 5prim UTR, et 3prim UTR. Dans ce cas une configuration est nécessaire dans l'onglet « configure this page ».

- Cliquer sur "Variant table" dans la sequence « genetic variation » pour obtenir un tableau décrivant toutes les variantes dans les régions du gène IL1B nous obtenant un tableau des rs (références snip) des SNPs associés ou non à des pathologies et stockés dans des bases de données.

- Cliquer sur variant image une image de différentes descriptives de notre requête va s'afficher.

2.2.2.3. Outil Galaxy

Principe

Galaxy est un flux de travail scientifique ouvert, de l'intégration des données, de la persistance d'analyse et une plateforme de publication, Galaxy est utilisé pour l'expression des gènes, l'assemblage du génome protéomique, épigénomique, transcriptomique, et une foule d'autres disciplines dans les sciences de la vie (Zhanjiang J Liu., 2017).

Méthode

Pour télécharger les données sur la plateforme galaxy. On se rend sur le site <https://usegalaxy.org/>, on clique sur « Get Data » puis sur « UCSC Main table browser », une page s'ouvrira, on remplit les informations nécessaire pour Exon avec la position de notre gène « IL1B » et on clique sur « get output » ceci nous conduira vers une autre page ou on fera d'autre réglage et on clique sur « send query to Galaxy », on attend quelque second pour visualiser la liste des Exon. On refait la même chose pour les snps (**Fig 15**).

Matériel et Méthodes

Galaxy interface showing the workflow for generating BED files from knownGene and snp151Common tracks.

Tools interface showing the search for tools, listing various servers like modENCODE fly server, InterMine server, etc.

Table Browser configuration for knownGene track:

- clade: Mammal, genome: Human, assembly: Dec. 2013 (GRCh38/hg38)
- group: Genes and Gene Predictions, track: GENCODE v32
- region: genome, position: chr2:112,829,751-112,836,843
- identifiers (names/accessions): paste list, upload list
- filter: create
- intersection: create
- correlation: create
- output format: BED - browser extensible data
- output file: (leave blank to keep output in browser)
- file type returned: plain text
- Send output to: Galaxy
- get output | summary/statistics

Table Browser configuration for snp151Common track:

- clade: Mammal, genome: Human, assembly: Dec. 2013 (GRCh38/hg38)
- group: Variation, track: Common SNPs(151)
- region: genome, position: chr2:112,829,751-112,836,843
- identifiers (names/accessions): paste list, upload list
- filter: create
- intersection: create
- correlation: create
- output format: BED - browser extensible data
- output file: (leave blank to keep output in browser)
- file type returned: plain text
- Send output to: Galaxy
- get output | summary/statistics

Output knownGene as BED configuration:

- Include custom track header: name= tb_knownGene, description= table browser query on knownGene, visibility= pack
- Create one BED record per: Coding Exons
- Note: if a feature is close to the beginning or end of a chromosome and upstream/downstream edge of the chromosome.
- Send query to Galaxy

Output snp151Common as BED configuration:

- Include custom track header: name= tb_snp151Common, description= table browser query on snp151Common, visibility= pack
- Create one BED record per: Whole Gene
- Note: if a feature is close to the beginning or end of a chromosome and upstream/downstream edge of the chromosome.
- Send query to Galaxy

1: UCSC Main on Human: knownGene (chr2:112,829,751-112,836,843)

18 regions
format: bed, génome de référence: hg38

display in IGB View
display with IGV local Human hg38
display at UCSC main

1.Chrom	2.Start	3.End	4.Name	5	6.Strand
chr2	11283060	11283073	ENST00000263341.7.cdx.0.0	chr2:11283061	r -
chr2	112831291	112831422	ENST00000263341.7.cdx.1.0	chr2:112831292	r -
chr2	112832661	112832826	ENST00000263341.7.cdx.2.0	chr2:112832662	r -
chr2	112833373	112833575	ENST00000263341.7.cdx.3.0	chr2:112833374	r -
chr2	112835565	112835617	ENST00000263341.7.cdx.4.0	chr2:112835566	r -

1: UCSC Main on Human: snp151Common (chr2:112,829,751-112,836,843)

28 regions
format: bed, génome de référence: hg38

display in IGB View
display with IGV local Human hg38
display at UCSC main

1.Chrom	2.Start	3.End	4.Name	5	6.Strand
chr2	112829855	112829856	rs1071676		-
chr2	112830259	112830260	rs10176689		+
chr2	112830724	112830725	rs1143643		-
chr2	112830814	112830815	rs3917380		-
chr2	112830935	112830936	rs3917379		-

Figure 15 : Méthode d'utilisation d'outils Galaxy. (<https://usegalaxy.org/>)

✓ Trouver les exons avec le plus de SNP

Dans le menu des outils de Galaxy, on clique sur « Join tool the intervals of two datasets side-by-side », une page s'ouvrira sur Galaxy, on remplit les informations nécessaires suivantes :

- “Join” : sélectionner « Exons ».
- “With” : sélectionner « SNPs ».
- “with min overlap” : sélectionnez « 1 ».
- “Return” : sélectionnez « Only records that are joined (INNER JOIN) », ce qui signifie que seules les correspondances sont incluses dans la sortie (c'est-à-dire uniquement les exons contenant des SNP et uniquement les SNP qui tombent en exons).
- Cliquer sur « Execute ».

✓ Compter le nombre de SNP par exon

On va utiliser l'outil « Group tool data by a column and perform aggregate operation on other columns » et on remplit les informations:

- “Select data” : sélectionner le “output dataset” de “Join tool”
- “Group by column” : Column: 4 (la colonne avec les ID d'exon)
- “Insert Operation” :
- 1 : Operation
- “Type” : Count
- “On column” : Column: 4
- Cliquer sur « Execute » pour effectuer le regroupement.

✓ Trier les exons par nombre de SNP

On va sur « Sort tool data in ascending or descending order » puis on remplit les informations suivantes :

- “Sort Query”: Output de Group tool
- “Column Selections” :
- 1 : Column Selections
- “on column” : Column : 2
- “in” : Descending order
- “Flavor”: Fast numeric sort (-n)
- Cliquer sur « Execute ».

Matériel et Méthodes

✓ Sélectionnez les cinq premiers exons

Pour sélectionner les cinq premiers exons on va sur « Select first Tool lines from a dataset » ensuite on remplit ces informations :

- “*Select first*” : « 5 ».
- “*from*” : «le output de Sort tool ».
- Cliquer sur « Execute ».

✓ Récupération des informations d'exon

Après toutes les étapes passées, on prend notre résultat et compare avec note Exon par l'outil « Compare two Datasets tool to find common or distinct row » qui est sur Galaxy dans la barre des outils et une page s'ouvrira sur Galaxy pour remplir quelque info :

- “*Compare*” : « Exons ».
- “*Using column*” : « Column: 4 ».
- “*against*” : «the output from Select first tool».
- “*Using column*” : «Column : 1».
- “*to find*” : «Matching rows of 1st dataset».
- Cliquer sur « Execute ».

Enfin, on s'affiche nos données dans le navigateur génomique UCSC par un clic sur « display at UCSC main » qui on le trouve sur notre fichier qui est sur historique.

Comme vous avez pu le constater, enchaîner toutes ces étapes est fastidieux. C'est pourquoi il est intéressant d'automatiser leur enchaînement. Pour cela nous allons créer un workflow. On rend sur « History Options » puis on choisit sur le menu « Extract Workflow ». En final on sauvegarde notre « workflow » (**Fig 16**).

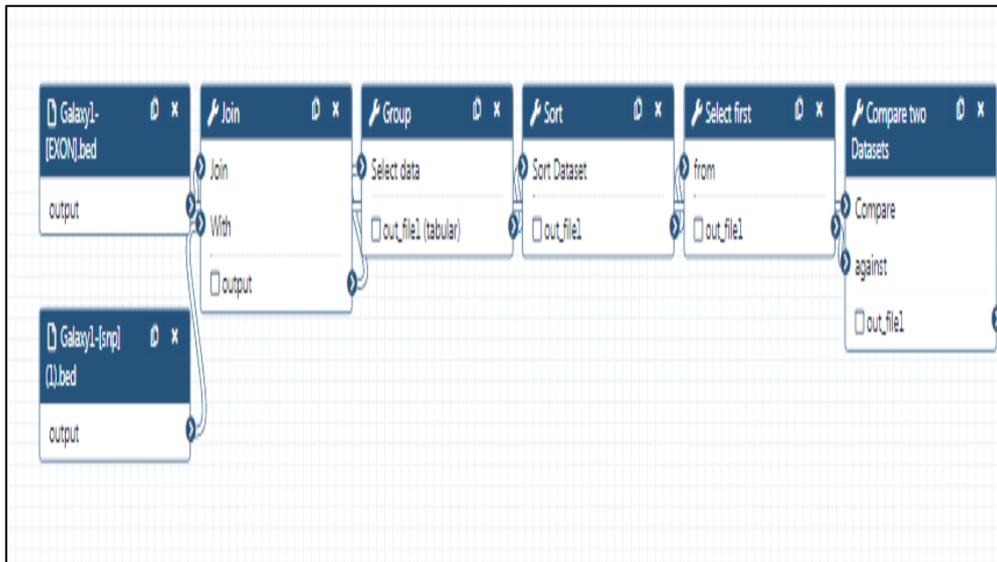


Figure 16 : L'organigramme de notre travail sur Galaxy

2.2.3. Identification d'association pathologique

2.2.3.1. Outil ALFRED

Principe

ALFRED (The ALlele FREquency Database) est conçu pour stocker et diffuser les fréquences d'allèles sur des sites polymorphes humains pour plusieurs populations, principalement pour les communautés de la génétique des populations et de l'anthropologie moléculaire. Actuellement, ALFRED dispose d'informations sur plus de 180 sites polymorphes pour plus de 70 populations. ALFRED est conçu pour que les données sur la fréquence des allèles sur des échantillons de population humaine soient facilement accessibles aux communautés scientifiques et éducatives (Osier *et al.*, 2001).

Méthode

- Accéder à la page d'accueil via le lien (<https://alfred.med.yale.edu/alfred/index.asp>) (Fig 17)
- Remplir les champs de recherches par le symbole du gène IL1B pour générer les résultats de la recherche sous forme de tableau avec les « rs » lié à des pathologies

- En cliquant sur chaque « rs » deux pages sont générées : l'une vers une description du site ALFRED et l'autre vers une description du site dvbar.

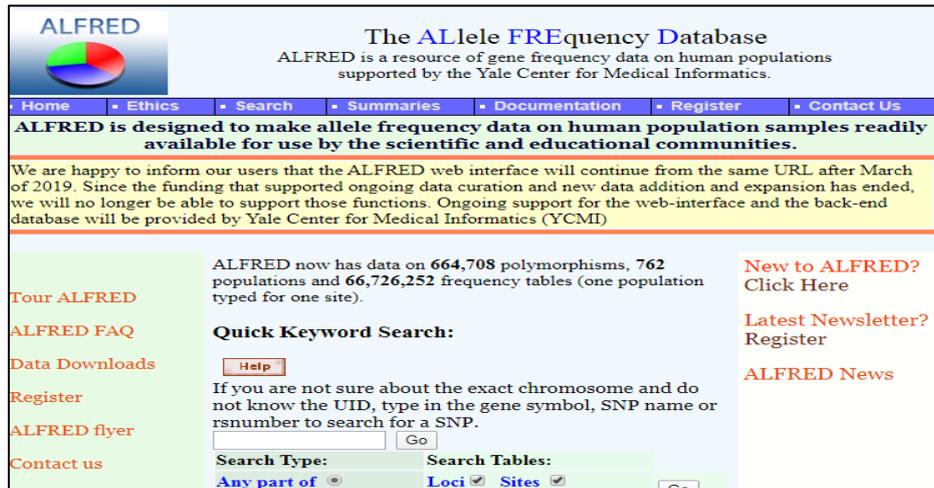


Figure 17 : Page d'accueil du site ALFRED
<https://alfred.med.yale.edu/alfred/index.asp>

2.2.3.2. Outil DISGENET

Principe

DisGeNET est une plate-forme de découverte complète conçue pour répondre à diverses questions concernant le fondement génétique des maladies humaines. DisGeNET contient plus de 380 000 associations entre plus de 16 000 gènes et 13 000 maladies, ce qui en fait l'un des plus grands référentiels du genre. DisGeNET intègre des bases de données sélectionnées par des experts avec des données extraites au format texte, couvre des informations sur les maladies mendéliennes et complexes et inclut des données provenant de modèles de maladies animales. Il présente un score basé sur les preuves à l'appui pour prioriser les associations gène-maladie. Ainsi, DisGeNET offre l'une des collections les plus complètes d'associations gène-maladie humaines et un ensemble précieux d'outils pour étudier les mécanismes moléculaires sous-jacents aux maladies d'origine génétique, conçus pour répondre aux besoins de différents profils d'utilisateur, notamment les bioinformaticiens, les biologistes et les professionnels de la santé (Piñero *et al.*, 2015).

Méthode

- Accéder à la page d'accueil plate-forme de découverte DigestNet via le lien <http://www.disgenet.org> (Fig 18).

- Il suffit de mettre le symbole du gène IL1B dans la zone de recherche et tableau contenant un ensemble d'informations concernant les associations pathologiques est généré.

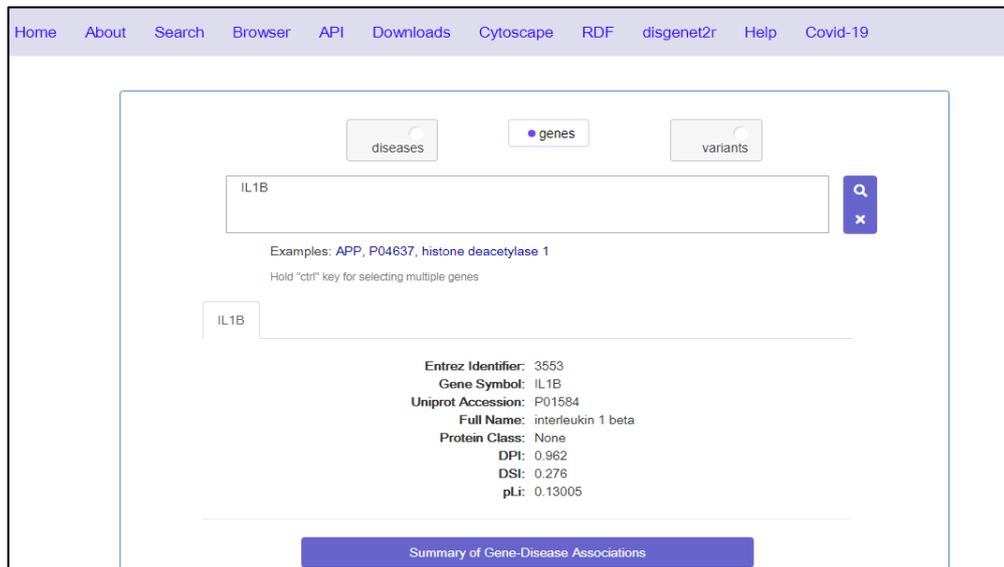


Figure 18 : Page d'accueil du site DigesNet. (<http://www.disgenet.org/search>)

2.2.3.3. Outil GenMania

Principe

GeneMANIA est un site Web flexible et convivial destiné à générer des hypothèses sur la fonction des gènes, à analyser des listes de gènes et à hiérarchiser les gènes pour des tests fonctionnels. À partir d'une liste de gènes interrogés, GeneMANIA trouve des gènes fonctionnellement similaires à l'aide de données génomiques et Protéomique. GeneMANIA est une autre utilisation de la prédiction de la fonction des gènes. Avec un seul gène d'interrogation, GeneMANIA trouve des gènes susceptibles de partager une fonction avec gène IL1B en fonction de leurs interactions (Franz *et al.*, 2018).

Méthode

Le lien « <http://genemania.org> » nous permet d'avoir un accès sur la page recherche de GenMania. Nous introduisons dans la zone recherche le symbole du gène IL1B. Nous pouvons par la suite grâce à une sélection dans un menu déroulant choisir les paramètres des interactions

que nous voulons pour notre travail puis lancer la recherche. En cliquant sur enregistrer nous obtenant la figure sous forme d'image.

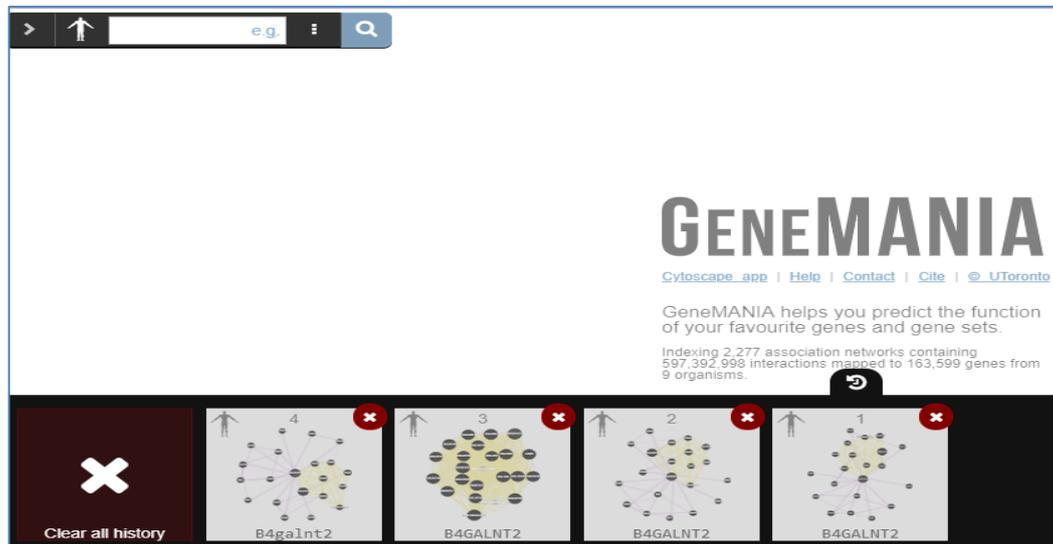


Figure 19 : Page d'accueil du site GeneMania.
(<http://genemania.org/data>)

2.2.3.4. STRING

Principe

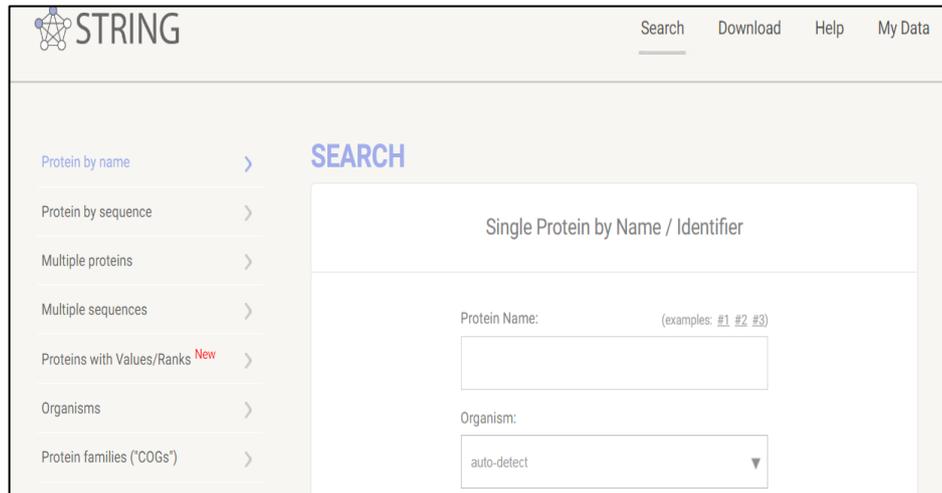
La base STRING est l'une des ressources les plus complètes. Elle possède une très bonne interface de navigation permettant l'exploration et la visualisation des associations protéine-protéine connues et prédites à partir de différents critères : voisinage physique des gènes sur le chromosome, existence d'un événement de fusion entre deux gènes, co-occurrence de deux gènes dans différentes espèces, coexpression des gènes, interaction protéique connue obtenue expérimentalement, cocitation des gènes ou protéines dans une référence bibliographique (Tagu et Risler., 2010).

Méthode

- Accéder à la page d'accueil plate-forme de STRING via le lien <https://string-db.org/> (Fig 20)

Matériel et Méthodes

- Remplir dans la zone de recherche le nom de la protéine «IL1B » et de l'organisme « homo sapiens » puis cliquer sur chercher.
- Plusieurs possibilités vont nous permettre d'accéder au « IL1B » sélectionner un qui correspond et appuyez sur Continuer pour obtenir la figure.



The screenshot shows the STRING database search interface. At the top left is the STRING logo. To the right are links for 'Search', 'Download', 'Help', and 'My Data'. On the left side, there is a vertical menu with the following items: 'Protein by name', 'Protein by sequence', 'Multiple proteins', 'Multiple sequences', 'Proteins with Values/Ranks' (with a 'New' tag), 'Organisms', and 'Protein families ("COGs")'. The main content area is titled 'SEARCH' and 'Single Protein by Name / Identifier'. It contains a 'Protein Name' input field with a placeholder '(examples: #1 #2 #3)' and an 'Organism' dropdown menu currently set to 'auto-detect'.

Figure 20 : Page d'accueil de STRING. (<https://string-db.org/>)

Chapitre III : Résultats et Discussion

3. Résultats et Discussion

3.1. Résultats

3.1.1. NCBI

L'interrogation de la base de données « NCBI » nous a permis d'obtenir les informations associées au gène IL1B « humain » (**tableau IV**). Les résultats de la recherche montre que ce gène IL1B humain est caractérisé par un profil polymorphique caractérisé par un très grand nombre de SNP.

Tableau IV : Références Genbank « NCBI » du gène humain IL1B

	NCBI
Nature	Nucléotide
Source	Homo sapiens (humain)
Organisme	Homo sapiens
Locus	NG_008851
Bp	14020bp
RNA	linéaire
Accession	NG_008851
Version (Séquence Reference de NCBI)	NG_008851.1
Cytogénétique localisation ou map	2q14.1
Synonyme	IL-1B ; IL1-BETA ; IL1beta ; IL1F2
Nom	interleukin 1 beta
Alias symbole (HGHC : HUGO Gene Nomenclature committee)	IL1B
Gene groups	Interleukins
Nombre d'exons	7

Résultats et Discussion

Nb des SNPs

2440

3.1.1.1. Transcrits du gène IL1B

Les résultats de l'interrogation de la base de données NCBI montrent que le gène IL1B a été transcrit en deux variantes (**Tableau V**).

Tableau V : Transcrits du gène IL1B

Variante de transcription	Transcript_ID	Longueur de variantes	Nombre d'exons	Exons codants	Emplacement
Variante 1	NM_000576.3	1507 bp	7	7	Chr2:112,829,751-112,836,779
Variante 2	XM_017003988.2	1502 bp	6	–	Chr2:112,829,758-112,836,843

3.1.1.2. Isoformes du gène IL1B

Les résultats de l'interrogation de la base de données UniProtKB montrent que le gène IL1B présente trois isoformes

Tableau VI : Isoformes de la protéine IL1B

Isoformes	poids de résidu	Charge	Poids moléculaire	Nombre de résidus
C9JWV2	114,633 g / mol	-10,5	16392,54 g / mol	143 aa
C9JSC2	114,649 g / mol	-11,5	14331,16 g / mol	125 aa
C9JVK0	115,346 g / mol	-12,0	12803,44 g / mol	111

3.1.1.3. Résultats d'alignement Blastn

Les résultats de l'alignement de la séquence FASTA du gène IL1B sont représentés par la (Fig 21) et le tableau VII. (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>)

Tableau VII : Séquences produisant un alignement significatif

Descriptions		Graphic Summary	Alignments	Taxonomy		
Sequences producing significant alignments						
		Download	Manage Columns	Show 100		
<input type="checkbox"/> select all 1 sequences selected		GenBank	Graphics	Distance tree of results		
Description	Max Score	Total Score	Query Cover	E value	Per Ident	Accession
<input checked="" type="checkbox"/> Homo sapiens chromosome 2, GRCh38.p13 Primary Assembly	1493	3125	100%	0.0	99.88%	NC_000002.12
<input type="checkbox"/> Homo sapiens chromosome 19, GRCh38.p13 Primary Assembly	50.0	177	9%	0.004	82.98%	NC_000019.10
<input type="checkbox"/> Homo sapiens chromosome 6, GRCh38.p13 Primary Assembly	45.5	123	5%	0.18	90.62%	NC_000006.12
<input type="checkbox"/> Homo sapiens chromosome 9, GRCh38.p13 Primary Assembly	44.6	83.7	4%	0.18	85.37%	NC_000009.12
<input type="checkbox"/> Homo sapiens chromosome 3, GRCh38.p13 Primary Assembly	43.7	244	13%	0.65	84.44%	NC_000003.12
<input type="checkbox"/> Homo sapiens chromosome 7, GRCh38.p13 Primary Assembly	43.7	124	4%	0.65	88.57%	NC_000007.14
<input type="checkbox"/> Homo sapiens chromosome 13, GRCh38.p13 Primary Assembly	43.7	283	12%	0.65	86.11%	NC_000013.11

L'identifiant de notre requête Alignement est ID : NM_000576.3 chez l'homme pour l'interleukine 1 beta (IL1B), transcrit variant. L'alignement de la séquence nucléotidique a généré une longueur de 1507 avec 27 hits de Blast contre 293 séquences d'espèces différentes. Nous avons retenus que ceux qui ont un score de similitude de 99% faisant référence à 1 hit de Blastn (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

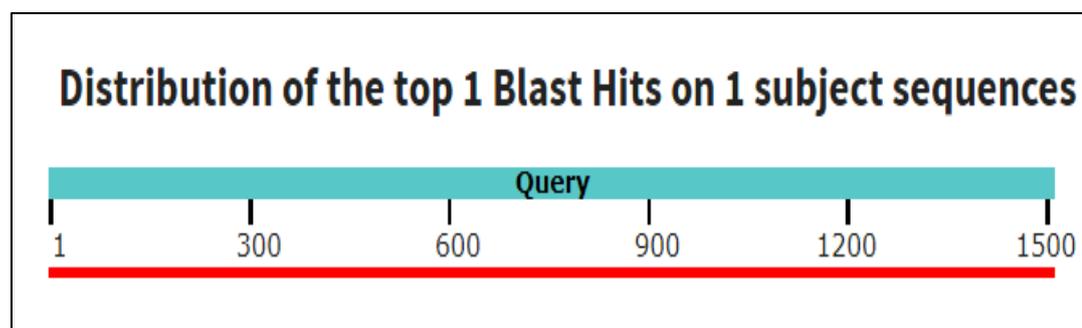


Figure 21 : Résumé graphique de l'alignement (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>)

3.1.1.4. Variantes NCBI

La recherche du polymorphisme du gène humain « IL1B » par dbvar NCBI a généré 102 variantes. En revanche seulement 18 études sont associées au gène « IL1B » chez l'homme jusqu'à aujourd'hui (**Tableau VIII**) (<https://www.ncbi.nlm.nih.gov/dbvar>)

Tableau VIII : Nombre d'étude cliniques des variantes du gène IL1B chez l'homme

Number of Studies: 18			
Study ID	Organism	Study type	Number of variant regions
nstd102	human	Case-Set	62493
nstd37	human	Case-Set	20391
estd210	human	Case-Set	4
nstd11	human	Tumor vs. Matched-Normal	181
nstd166	human	Control Set	304733
nstd189	human	Case-Control	2181
nstd158	human	Collection	1621
nstd151	human	Case-Set	49339
nstd122	human	Case-Control	94923
nstd113	human	Case-Set	9010
nstd100	human	Case-Control	70319
estd211	human	Case-Set	100
nstd85	human	Case-Control	1416
nstd54	human	Case-Control	81345
estd197	human	Control Set	232775
estd192	human	Collection	61187
nstd16	human	Control Set	1290
nstd128	human	Collection	1328521

3.1.2. UCSC génome browser

Le navigateur UCSC nous a permis d'obtenir la localisation chromosomique du gène IL1B (**Fig 22**) et une piste des rs représentées par les numéros d'accessions des SNPs communs et des SNPs Flagged et de tous les SNPs (**Fig 23**).

- Common SNPs(151) qui ont une fréquence allélique mineure d'au moins 1% et sont mappés à un seul emplacement dans l'assemblage du génome de référence.
- All SNP (151) qui ont été mappés à plusieurs emplacements dans l'assemblage du génome de référence.
- Flagged SNPs(151) marqués comme cliniquement associés par dbSNP, mappés à un seul emplacement dans l'assemblage du génome de référence, et *non* connus pour avoir une fréquence allélique mineure d'au moins 1%.

Résultats et Discussion

Dans la section Options de coloration de la page des commandes de piste, les termes de fonction sont regroupés en plusieurs catégories, affichées ici avec des couleurs par défaut :

Locus: en amont/en aval du gène, intron variant

Coding - Synonymous: variant synonyme

Coding non synonymous : gain de fonction, faux sens variant, perte de fonction, changement du cadre de lecture, site d'épissage donneur et accepteur

Untranslated: 5' UTR variant, 3' UTR variant, non codant variant.

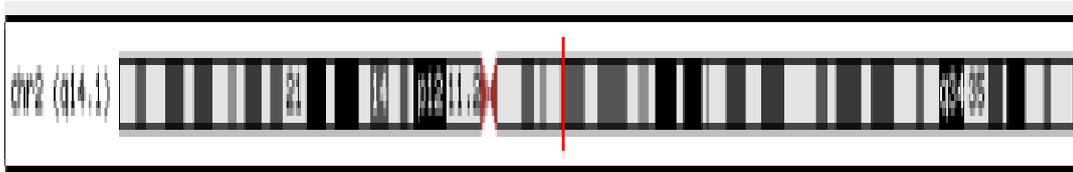


Figure 22 : Localisation chromosomique du gène IL1B (ch2 q14.1).

Résultats et Discussion



Figure 23 : Liste de tous les SNPs positionnées sur le gène IL1B via UCSC Génome Browser. <https://genome.ucsc.edu/cgi-bin/hgTracks>

3.1.3. HUGO (Gene Nomenclature committe)

La configuration des pistes de SNP de UCSC nous a conduit dans vers un autre lien (<http://hgdp.uchicago.edu/cgi-bin/gbrowse/HGDP/>). Ce dernier nous permet de visualiser les références snip « rs » liées au gène IL1B (Fig 24). Aussi la figure 25 visualise la distribution des allèles ancestraux et des allèles dérivés pour chaque rs.

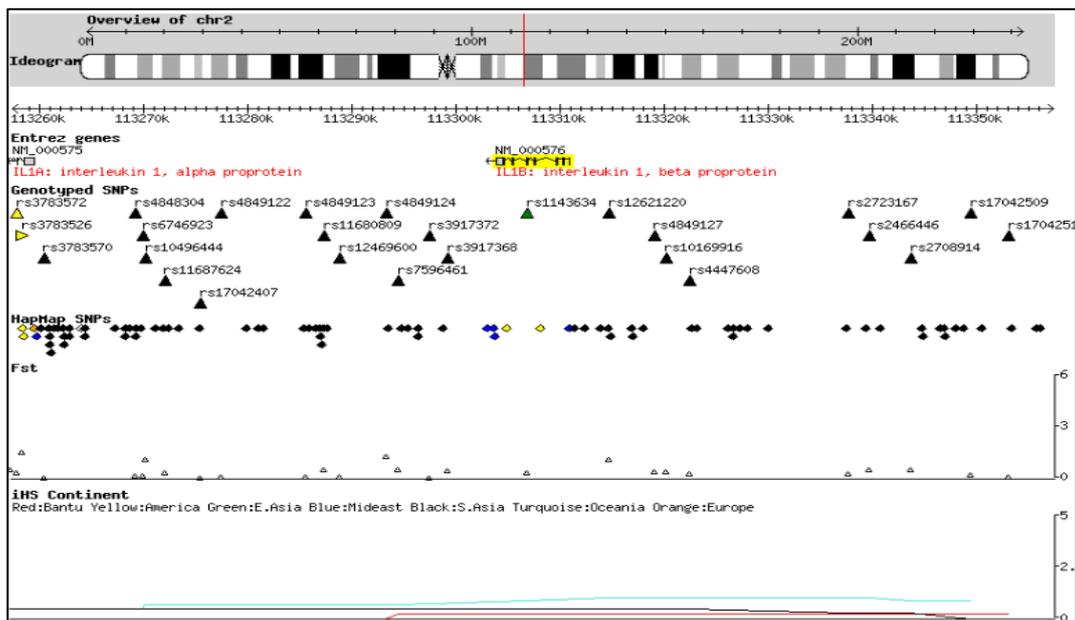


Figure 24 : Les différents SNPs du lien HUGO. <http://hgdp.uchicago.edu/cgi-bin/gbrowse/HGDP/>

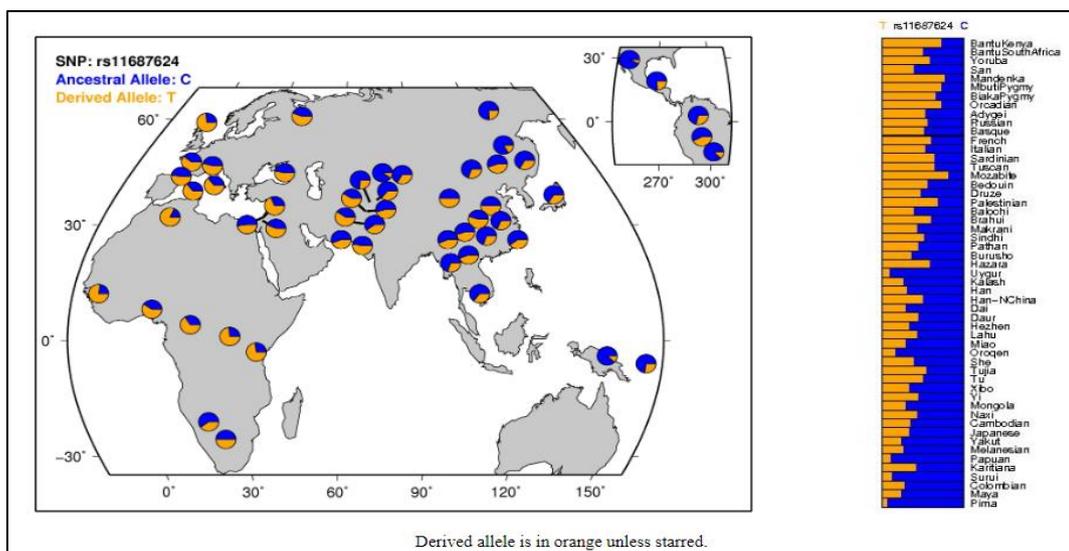


Figure 25 : Distribution des allèles ancestraux et des allèles dérivés du rs11687624 <http://hgdp.uchicago.edu/cgi-bin/alfreqs.cgi?pos=113272063&chr=chr2&rs=rs11687624&imp=false>

3.1.4. Ensemble Genome Browser

Les résultats obtenus par le navigateur Esembl genome browser Human Assembly nous a permis de rassembler les informations représentées par le tableau IX et (Fig 26).

Tableau IX : Variantes de transcription du gène IL1B

Name	Transcript ID	bp	Protein	Biotype	CCDS
IL1B-201	ENST00000263341.7	1507	269aa	Protein coding	CCDS2102
IL1B-203	ENST00000418817.5	581	143aa	Protein coding	-
IL1B-204	ENST00000432018.5	581	125aa	Protein coding	-
IL1B-202	ENST00000416750.1	580	111aa	Protein coding	-
IL1B-207	ENST00000491056.5	1617	No protein	Processed transcript	-
IL1B-208	ENST00000496280.5	849	No protein	Retained intron	-
IL1B-206	ENST00000487639.1	654	No protein	Retained intron	-
IL1B-205	ENST00000477398.1	571	No protein	Retained intron	-

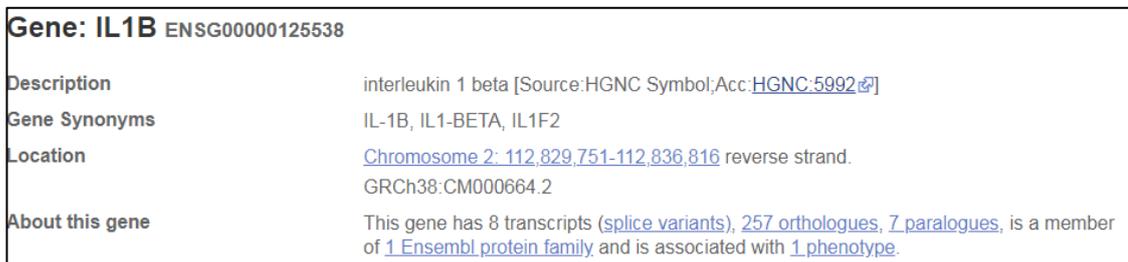
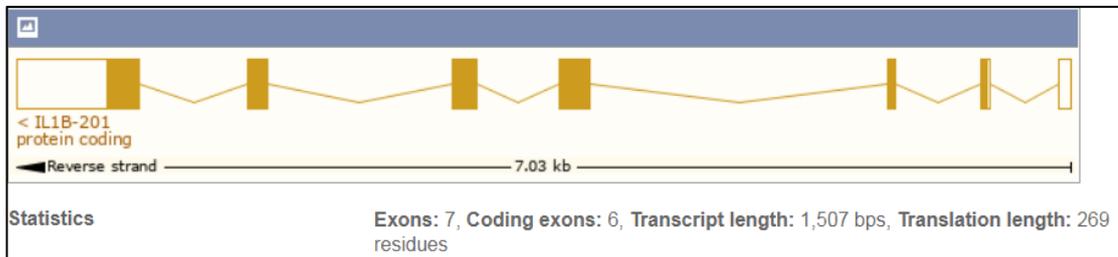


Figure 26 : Structure du gène IL1B selon Ensembl génome browser

Résultats et Discussion

3.1.5. Galaxy

Notre analyse sur la plateforme Galaxy « Usegalaxy » nous a permis d'avoir les informations associées au gène IL1B « humain ». Où elle nous a permis d'obtenir 05 exons qui ont le plus grand nombre de polymorphismes nucléotidique simples (SNP).

Le téléchargement des données des exons et SNPs sur la plateforme permis d'affiché une liste de 18 régions pour les exons et 28 pour le SNPs sont présentent dans les tableaux (Annex01 ; Annex02), donc chaque ligne représente un exon (ou SNP) ; les trois premiers colonnes sont l'emplacement génomique et la quatrième colonne contient le nom « ID » de l'exon (ou SNP).

✓ Trouver les exons avec le plus de SNP

Le résultat de notre fusion de deux fichier « Exon&snp » a permis d'obtenir 10 régions qui correspond les exons les plus de SNP qui sont présentés dans le tableau X ; les six premiers colonnes correspondent aux exons, et les six derniers colonnes correspondent aux SNP, tandis que la colonne 4 contient les ID des exons et la colonne 10 contient les ID des SNPs.

Tableau X : les régions qui correspond les exons le plus de SNP.

Chrom	Start	End	Name	Score	Strand	Thick Start	Thick End	ItemR GB	BlockC ount	Block Sizes	BlockS tarts	
chr2	1283	12832	ENST00000263341.7_cds_2_0_chr2_112832662_r			chr2	11283	11283	rs1143	634	0	-
chr2	1283	12832	ENST00000418817.5_cds_0_0_chr2_112832699_r			chr2	12832	12832	rs1143	634	0	-
chr2	1283	12832	ENST00000432018.5_cds_0_0_chr2_112832752_r			chr2	12832	12832	rs1143	634	0	-
chr2	1283	12832	ENST00000416750.1_cds_0_0_chr2_112832795_r			chr2	12832	12832	rs1143	634	0	-

✓ Compter le nombre de SNP par exon

Il contient deux colonnes. Le premier contient les ID d'exon, et le second le nombre de fois ou cet ID d'exon est apparu dans le fichier (combien de SNP étaient présents dans cet exon).

Tableau XI : nombre de SNP par exon

Résultats et Discussion

1	2
ENST00000263341.7_cds_2_0_chr2_112832662_r	1
ENST00000416750.1_cds_0_0_chr2_112832795_r	1
ENST00000418817.5_cds_0_0_chr2_112832699_r	1
ENST00000432018.5_cds_0_0_chr2_112832752_r	1

✓ Trier les exons par nombre de SNP

Tableau XII : classement des exons par nombre de SNP

1	2
ENST00000263341.7_cds_2_0_chr2_112832662_r	1
ENST00000416750.1_cds_0_0_chr2_112832795_r	1
ENST00000418817.5_cds_0_0_chr2_112832699_r	1
ENST00000432018.5_cds_0_0_chr2_112832752_r	1

✓ Récupération des informations d'exon

Tableau XIII : les cinq exons qui ont le plus grand nombre de SNP.

hr om	start	end	Name	core	transcript	hick Star t	hic kEn d	tem RGB	lock Cou nt	lock Sizes	lockS tarts
			ENST00000263								
hr2	1283 2661	1283 2826	341.7_cds_2_0_chr2_1 12832662_r								
			ENST00000416								
hr2	1283 2794	1283 2826	750.1_cds_0_0_chr2_1 12832795_r								
			ENST00000418								
hr2	1283 2698	1283 2826	817.5_cds_0_0_chr2_1 12832699_r								
			ENST00000432								
hr2	1283 2751	1283 2826	018.5_cds_0_0_chr2_1 12832752_r								

3.1.6. ALFRED (The ALlele FREquency Database)

Les résultats de la recherche des SNP du gène IL1B par le navigateur ALFRED sont présentés par la figure 27.

Name	ALFRED UID	Locus Symbol	Chromosome	Band Position
interleukin 1, beta	LO062034P	IL1B	2	2q14

Synonyms:
Sites List: [See Sites List](#)
External Resources: [Entrez Gene Locus Information](#)
References: [See References](#)
Locus Description:
Sites within this locus ordered by their chromosomal position in the 37.3 NCBI build:

Site Name (Navigates to ALFRED description)	dbSNP rs# (Navigates to dbSNP reference page)	Chr-Position	Status	# Populations typed
rs3917368	rs3917368	113582782		116
rs1143634	rs1143634	113590390		76
rs16944	rs16944	113594867		2
rs12621220	rs12621220	113598255		75

Figure 27 : Résultats des SNP du gène IL1B de la base ALFRED

Nous n'avons constaté que l'interrogation de la base de données ALFRED à générer que 4 rs typés chez des populations. Ces rs ont deux lien : le premier vers le navigateur ALFRED et le second vers le navigateur dbSNP. Nous constatons aussi que les rs les plus étudiés dans les populations sont : rs3917368, rs1143634.

3.1.7. DISGENET

L'interrogation de la base de données DISGENET nous a permis d'obtenir le tableau XIV.

Résultats et Discussion

Tableau XIV : Pathologies associées aux gènes IL1B.

Disease	Type	Disease Class	Semantic Type	N. genes	N. SNPs
Malignant Neoplas...	group	Neoplasms	Neoplastic Process	8621	1641
Asthma	disease	Respiratory Tract Dis...	Disease or Syndr...	2096	1536
Breast Carcinoma	disease	Neoplasms; Skin and...	Neoplastic Process	6776	2793
Malignant neoplas...	disease	Neoplasms; Skin and...	Neoplastic Process	6941	3417
Depressive disorder	disease	Mental Disorders	Mental or Behavi...	1719	297
Multiple Sclerosis	disease	Immune System Dis...	Disease or Syndr...	1600	1022
Schizophrenia	disease	Mental Disorders	Mental or Behavi...	2872	2897
Psoriasis	disease	Skin and Connective...	Disease or Syndr...	1308	705
Pneumonitis	disease	Infections; Respirator...	Disease or Syndr...	697	13

L'examen des résultats de la base de données DISGENET nous a permis d'identifier que le gène IL1B est associé au cancer du sein. 4 SNPs sont mentionnés dans cette association (Fig 28).

IL1B, interleukin 1 beta, 3553

N. diseases: 1801; N. variants: 22

Disease: Breast Carcinoma x

Source: ALL

Results per page: 25

Filter within current results:

Variant	DSI	DPI	Chr	Position	Consequence	Alleles	Class	AF EXOME	AF GENOME	Disease	Score	N. EI	N. PMIDs	First Ref.	Last Ref.
rs1143634	0.597	0.680	2	112832813	synonymous variant	G/A	snv	0.19	0.19	Breast C...	0.030	1.000	3	2012	2019
rs1143623	0.677	0.440	2	112838252	upstream gene variant	C/G	snv		0.24	Breast C...	0.010	1.000	1	2019	2019
rs1143627	0.605	0.760	2	112836810	5 prime UTR variant	G/A	snv		0.56	Breast C...	0.010	1.000	1	2010	2010
rs16944	0.531	0.920	2	112837290	upstream gene variant	A/G	snv		0.57	Breast C...	0.010	1.000	1	2019	2019

Figure 28 : Variantes IL1B dans le cancer du sein

3.1.8. GenMania

La fonction principal de GeneMania est d'effectuer des recherches sur un ensemble de données biologiques accessibles aux publiques (open source) pour trouver des gènes apparentés. Les résultats obtenus par GeneMania sont représentés par la figure 29.

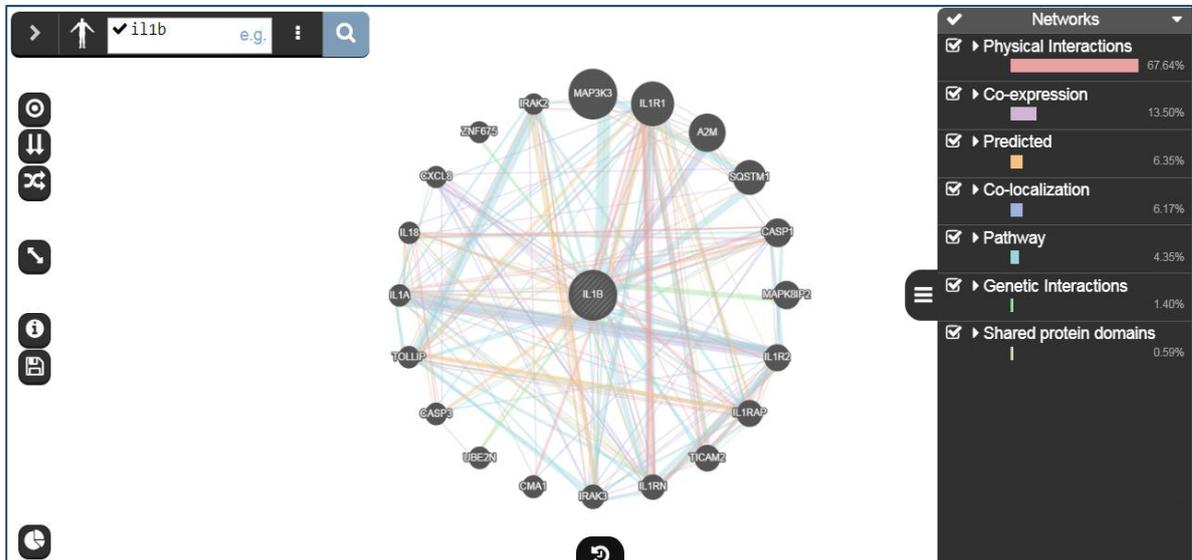


Figure 29 : Résultats de l'analyse du réseau GeneMania des interactions avec le gène IL1B

(<https://genemania.org/search/homo-sapiens//IL1B>)

Résultats et Discussion

Ces résultats regroupent 20 gènes différents qui interagissent avec le gène IL1B avec sept niveaux d'interactions avec 22 liens au total.

1. Interactions physiques en rose entre IL1B et CASP3, IL1RAP, IL1R1, CASP1, IL1R2, IL1RN, IRAK3, A2M, CMA1
2. Interaction de co-expression en mauve avec gène cible IL1B et d'autres gènes : CASP1, IL1R2, IL1RN, IRAK3, IL1A, IL18 et CXCL8
3. Prédite en orange entre IL1B et IL1R1, CASP1, UBE2N, CASP3, IL18.
4. Co-localisation en bleu entre IL1B et d'autres gènes : A2M, CXCL8 et IRAK3.
5. Sentir en bleu vert entre IL1B et les gènes : MAP3K3, SQSTM1, CASP1, IL1R2, IL1RAP, TICAM2, IL1RN, IRAK3, TOLLIP et IRAK2.
6. Interactions génétiques en vert entre IL1B et MAPK8IP2, UBE2N, ZNF675.
7. Domaines protéiques partagés en jaune entre le gène IL1B et autres gènes : MAP3K3, IRAK3, IL1A, IL18.

3.1.9. STRING

Les résultats de la visualisation des interactions protéine-protéine par l'outil STRING (fonctional protein association networks) et qui signifie une chaîne de molécules est représenté par la figure 30.

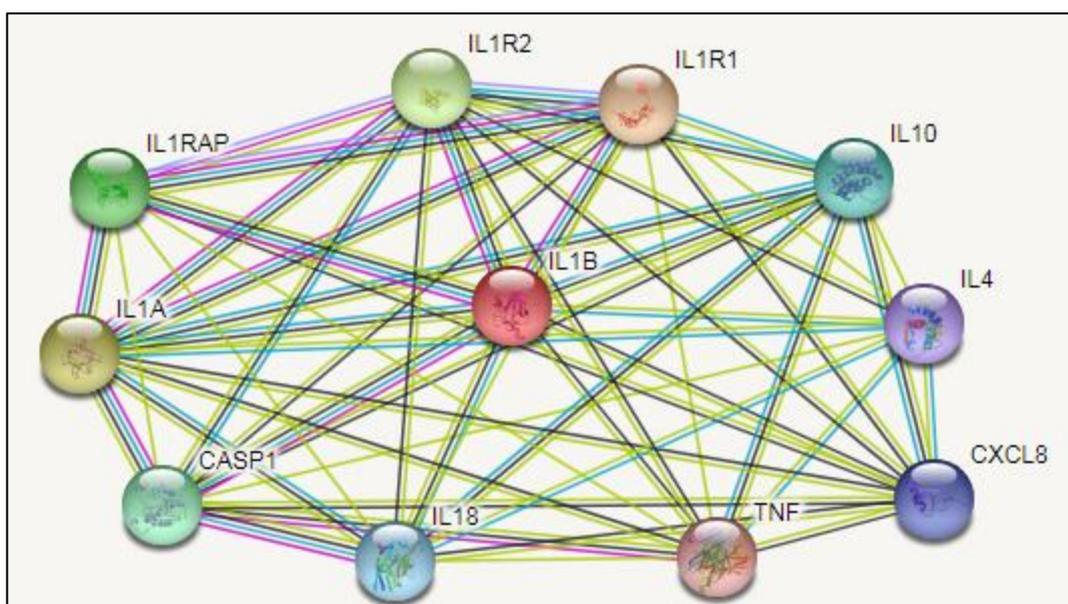


Figure 30 : Interaction protéine-protéine du gène IL1B

(<https://string-db.org/cgi/network.pl?taskId=CIPSwABBwZiM>)

Résultats et Discussion

Le Score d'expression basé sur les modèles d'expression de l'ARN et la co-régulation protéique fournis par ProteomeHD est présenté par la figure 31. Nous constatons une Co-expression en rouge foncé entre le gène IL1B et le gène CXCL8 avec un score d'expression d'ARN de 0,662.

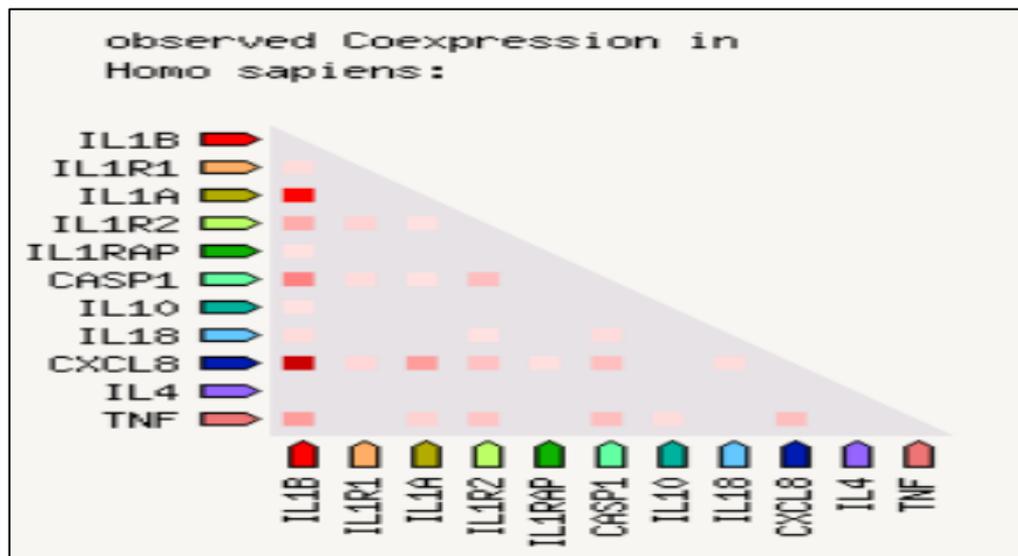


Figure 31 : Co-expression des gènes avec le gène IL1B

(https://string-db.org/cgi/coexpression_evidence.pl?taskId=KmFrteFSazMV&allnodes=1)

Prédiction in silico de la structure 3D d'IL1B

La prédiction de la structure 3D in silico est une étape importante pour mieux comprendre l'architecture globale de celle-ci. Pour ce faire nous avons utilisé le navigateur SUISSE MODEL basé sur la modélisation comparative de protéine IL1B (Fig 32).

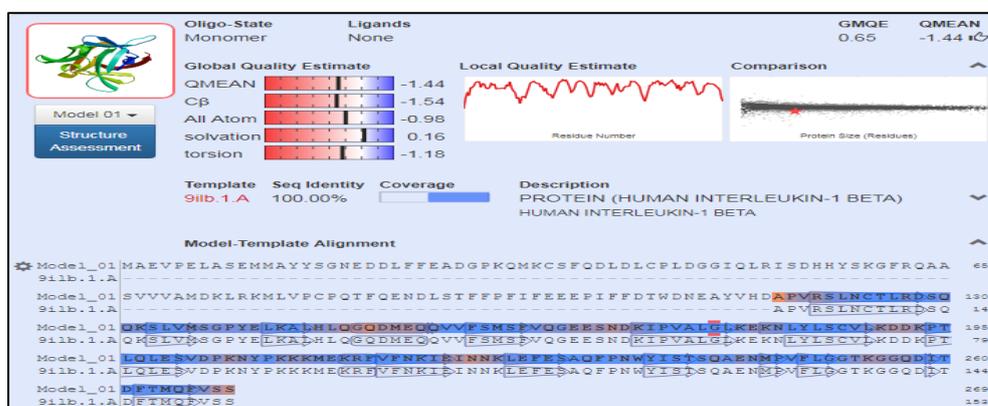


Figure 32 : Modélisation comparative de protéine IL1B par SUISE MODEL
<https://swissmodel.expasy.org/interactive/KrWZ9K/models/>

3.2. Discussion

Au sein de la tumeur, l'IL-1 β est produite et sécrétée par divers types de cellules, telles que les cellules immunitaires, les fibroblastes ou les cellules cancéreuses. Cependant, les mécanismes de production d'IL-1 β ont été les plus largement étudiés dans les cellules immunitaires, en particulier dans les cellules myéloïdes, telles que les macrophages.

Il a été démontré que l'IL-1 β joue un rôle dans de nombreux événements physiologiques. Il peut moduler l'expression génique et la production de cytokines, réguler l'adhésion et la migration cellulaires, l'angiogenèse ou la réponse immunitaire (Rébé et Ghiringhelli., 2020).

Les niveaux d'IL1 β sont élevés dans les tissus du cancer du sein et sont en corrélation avec le caractère invasif et un phénotype agressif (Jin *et al.*, 1997). Ils semblent réguler la prolifération des cellules cancéreuses grâce à la production d'œstrogène par des enzymes catalysant les stéroïdes dans les tissus (Honma *et al.*, 2002). Le gène IL1B est cartographié en 2q14 (Nicklin *et al.*, 2002) et les variants génétiques couramment décrits incluent le - 511C> T et le -31C> T dans le 5'UTR et le polymorphisme + 3954C> T dans l'exon 5 du gène.

Notre étude a été conçue pour examiner l'association génétique entre les polymorphismes du gène candidat d'IL1B et la susceptibilité au cancer du sein chez les femmes.

À ce jour, de nombreuses études épidémiologiques ont évalué le polymorphisme du gène *IL-1 β* avec le risque de cancer comme le cancer du poumon cervical (Lee et Song., 2014), gastrique (Chen *et al.*, 2016) et cancer du poumon non à petites cellules (Pérez-Ramírez *et al.*, 2017). Plusieurs études se sont concentrées sur l'association entre le polymorphisme de l'*IL-1 β -C31T* et le risque de cancer du sein. Cependant, les conclusions des études antérieures sont contradictoires. (Ito *et al.*, 2002) ont montré le rôle protecteur du génotype *TT* (*IL-1 β -C31T* rs1143627) contre l'apparition du cancer du sein dans la population japonaise (OR ajusté en fonction de l'âge, 0,58; IC à 95%, 0,32–1,02). Une méta-analyse menée en 2010, qui comprenait 8 études cas-témoins, a conclu que *CC* génotype était associé à une augmentation de 1,37 fois du risque de cancer du sein (Liu *et al.*, 2010). De plus, (Liu *et al.*, 2006) ont observé une relation positive entre le génotype *IL-1 β -31CC* et le risque de cancer du sein (OR ajusté, 1,72; IC à 95%, 1,16–2,54).

Une étude de (Gong *et al.*, 2013) Chez les femmes ménopausées EA, les SNP dans *IL1B* (rs1143627 et rs16944) étaient associés à un risque accru (tendance *P* pour les allèles C et T = 0,003, 0,002, respectivement). (Liu *et al.*, 2006) ont étudié les associations de

Résultats et Discussion

2 SNP promoteurs du gène IL-1b avec un risque de cancer du sein dans une population chinoise. Ils ont constaté que les génotypes variantes des deux loci IL-1b T-31C et C-511T étaient associés à un risque significativement accru de cancer du sein sous forme d'allèle dose-réponse, et les associations étaient plus évidentes chez les femmes âgées, les femmes ménopausées et les personnes avec un âge avancé de la ménarche. Ces résultats suggèrent que ces 2 variantes du promoteur IL-1b peuvent contribuer au risque de développer un cancer du sein dans les populations chinoises.

Jusqu'à présent, des études émergentes ont été menées pour étudier les relations entre le polymorphisme rs1143634 et le risque de cancer du sein. Par exemple, (Snoussi *et al.*, 2005) ont seulement trouvé que + 3954C> T rs1143634 d'IL-1 β était fortement lié aux formes agressives de cancer du sein alors qu'aucune différence significative entre le polymorphisme de + 3954C> T rs1143634 d'IL-1 β et le risque de cancer du sein a été observée. Les résultats rapportés par (Balasubramanian *et al.*, 2006) ont également indiqué que le polymorphisme + 3954C> T rs1143634 n'était pas une prédisposition à la susceptibilité au cancer du sein. Les données de (Hefler *et al.*, 2005) ont dévié de HWE et aucun lien évident n'a été trouvé entre le polymorphisme + 3954C> T rs1143634 et la présence d'un cancer du sein.

Étant donné que les niveaux d'IL-1 β dans le tissu mammaire peuvent être contrôlés *in vivo* par l'estradiol et sont corrélés à la graisse sous-cutanée abdominale (Abrahamsson *et al.*, 2012), les polymorphismes génétiques de ce gène pourraient être plus pertinents chez les femmes ménopausées, comme le suggèrent les résultats de (Gong *et al.*, 2013), car la graisse périphérique est une principale source de production d'œstrogènes après la ménopause.

Dans notre travail, l'utilisation des outils bioinformatiques permettront d'avoir une meilleure caractérisation du gène IL1B. La recherche de l'association d'IL1B et le cancer du sein est basée sur l'analyse d'un grand nombre de SNPs, cela permis d'approfondir les connaissances sur physiopathologie. La base de données DISGENET a identifié les variantes génétiques associées au cancer du sein. Ils peuvent ensuite être utilisés comme marqueurs génétiques qui permettront d'identifier le risque d'exposition au cancer du sein.

Conclusion

IL1B est une cytokine proinflammatoire qui a une fonction importante dans le maintien d'homéostasie dans l'organisme normal.

L'IL1B a pour rôle d'éteindre la prolifération des cellules souches cancéreuse dans le cas d'un organisme sain, mais selon notre étude il peut aussi allumer leurs prolifération ce qui peut conduire l'organisme à devenir atteint du cancer du sein.

Notre étude a été conçue pour examiner spécifiquement les variantes génétiques putativement fonctionnelles dans le gène d'IL1B liées à un risque de cancer du sein chez les femmes.

D'après nos résultats nous avons conclure que les rs1143627, rs16944, rs1143634 sont corrélés avec le changement de fonction de l'IL1B qui devient inducteur de la proliférations des cellules cancéreuse dans le sein et puis l'apparition du cancer du sein , ces résultats permettre de mieux diagnostiquer et comprendre les différentes voies qui conduisent au développement des cellules cancéreuses dans le sein en outre combattre cette maladie maligne.

La bioinformatique est un moyen essentiel et efficace dans la recherche et l'identification des SNPs et définitions de ceux qui sont en association avec le cancer du sein.

Discussion

Within the tumor, IL-1 β is produced and secreted by various types of cells, such as immune cells, fibroblasts or cancer cells. However, the mechanisms of IL-1 β production have been most widely studied in immune cells, particularly in myeloid cells, such as macrophages.

IL-1 β has been shown to play a role in many physiological events. It can modulate gene expression and cytokine production; regulate cell adhesion and migration, angiogenesis or immune response (Rébé and Ghiringhelli, 2020).

Levels of IL1 β are elevated in breast cancer tissue and correlate with invasiveness and an aggressive phenotype (Jin et al., 1997). They appear to regulate the proliferation of cancer cells through the production of estrogen by steroid-catalyzing enzymes in the tissues (Honma et al., 2002). The IL1B gene is mapped to 2q14 (Nicklin et al., 2002) and commonly described genetic variants include -511C> T and -31C> T in the 5'UTR and the + 3954C> T polymorphism in exon 5 of the gene.

Our study was designed to examine the genetic association between polymorphisms in the candidate IL1B gene and susceptibility to breast cancer in women.

To date, numerous epidemiological studies have evaluated the polymorphism of the IL-1 β gene with the risk of cancer such as cervical (Lee and Song, 2014), gastric (Chen et al., 2016) and non-lung cancer, small cell (Pérez-Ramírez et al., 2017). Several studies have focused on the association between the polymorphism of IL-1 β -C31T and the risk of breast cancer. However, the conclusions of previous studies are contradictory. (Ito et al., 2002) showed the protective role of the TT genotype (IL-1 β -C31T rs1143627) against the appearance of breast cancer in the Japanese population (OR adjusted according to age, 0.58; 95% CI, 0.32–1.02). A meta-analysis conducted in 2010, which included 8 case-control studies, concluded that CC genotype was associated with a 1.37-fold increase in breast cancer risk (Liu et al., 2010). In addition, (Liu et al., 2006) observed a positive relationship between the IL-1 β -31CC genotype and breast cancer risk (adjusted OR, 1.72; 95% CI, 1.16–2, 54).

A study by (Gong et al., 2013) in postmenopausal EA women, SNPs in IL1B (rs1143627 and rs16944) were associated with increased risk (P trend for C and T alleles = 0.003, 0.002, respectively). (Liu et al., 2006) studied the associations of 2 SNP promoters of the IL-1b gene with a risk of breast cancer in a Chinese population. They found that variant genotypes of both

IL-1b T-31C and C-511T loci were associated with a significantly increased risk of breast cancer as a dose-response allele, and the associations were more evident in older women, postmenopausal women and people with advanced menarche age. These results suggest that these 2 variants of the IL-1b promoter may contribute to the risk of developing breast cancer in Chinese populations.

So far, emerging studies have been conducted to investigate the relationships between the rs1143634 polymorphism and breast cancer risk. For example, (Snoussi et al., 2005) only found that + 3954C> T rs1143634 of IL-1 β was strongly linked to aggressive forms of breast cancer while no significant difference between the polymorphism of + 3954C> T rs1143634 of IL-1 β and the risk of breast cancer was observed. The results reported by (Balasubramanian et al., 2006) also indicated that the polymorphism + 3954C> T rs1143634 was not a predisposition to susceptibility to breast cancer. The data from (Hefler et al., 2005) deviated from HWE and no obvious link was found between the polymorphism + 3954C> T rs1143634 and the presence of breast cancer.

Since IL-1 β levels in breast tissue can be controlled in vivo by estradiol and correlates with abdominal subcutaneous fat (Abrahamsson et al., 2012), genetic polymorphisms of this gene could be more relevant in postmenopausal women, as the results of (Gong et al., 2013) suggest, because peripheral fat is a primary source of estrogen production after menopause.

In our work, the use of bioinformatic tools will allow us to have a better characterization of the IL1B gene. The research of the association of IL1B and breast cancer is based on the analysis of a large number of SNPs, which has led to further knowledge on pathophysiology. The DISGENET database has identified genetic variants associated with breast cancer. They can then be used as genetic markers.

Références

- Abrahamsson, A. *et al.* (2012) 'Estradiol, tamoxifen, and flaxseed alter IL-1 β and IL-1Ra levels in normal human breast tissue in vivo', *Journal of Clinical Endocrinology and Metabolism*, 97(11), pp. E2044–E2054. doi: 10.1210/jc.2012-2288.
- Bachelot, T. *et al.* (2003) 'Prognostic value of serum levels of interleukin 6 and of serum and plasma levels of vascular endothelial growth factor in hormone-refractory metastatic breast cancer patients', *British Journal of Cancer*. Nature Publishing Group, 88(11), pp. 1721–1726. doi: 10.1038/sj.bjc.6600956.
- Balasubramanian, S. P. *et al.* (2006) 'Interleukin gene polymorphisms and breast cancer: A case control study and systematic literature review', *BMC Cancer*, 6, pp. 1–21. doi: 10.1186/1471-2407-6-188.
- Baroja-Mazo, A. *et al.* (2014) 'The NLRP3 inflammasome is released as a particulate danger signal that amplifies the inflammatory response', *Nature Immunology*. Nature Publishing Group, 15(8), pp. 738–748. doi: 10.1038/ni.2919.
- Bensen, J. T. *et al.* (2001) 'Identification of a novel human cytokine gene in the interleukin gene cluster on chromosome 2q12-14', *Journal of Interferon and Cytokine Research*. J Interferon Cytokine Res, 21(11), pp. 899–904. doi: 10.1089/107999001753289505.
- Bouma-Ter Steege, J. C. A. *et al.* (2004) 'Angiogenic profile of breast carcinoma determines leukocyte infiltration', *Clinical Cancer Research*, 10(21), pp. 7171–7178. doi: 10.1158/1078-0432.CCR-04-0742.
- Brough, D., Tyrrell, P. J. et Allan, S. M. (2011) 'Regulation of interleukin-1 in acute brain injury', *Trends in Pharmacological Sciences*. Trends Pharmacol Sci, pp. 617–622. doi: 10.1016/j.tips.2011.06.002.
- Chavey, C. *et al.* (2007) 'Oestrogen receptor negative breast cancers exhibit high cytokine content', *Breast Cancer Research*. Breast Cancer Res, 9(1). doi: 10.1186/bcr1648.
- Chen, B. *et al.* (2016) 'Correlation Between Interleukin-1 β -511 C/T Polymorphism and Gastric Cancer in Chinese Populations: A Meta-Analysis', *Medical Science Monitor : International Medical Journal of Experimental and Clinical Research*. International Scientific Information, Inc., 22, p. 1742. doi: 10.12659/MSM.895771.
- Chen, H. *et al.* (2006) 'Single nucleotide polymorphisms in the human interleukin-1B gene affect transcription according to haplotype context', *Human Molecular Genetics*, 15(4), pp. 519–529. doi: 10.1093/hmg/ddi469.
- Courivaud, C., Saas, P. et Ducloux, D. (2012) 'Genetic polymorphisms: How to interpret studies?', *Nephrologie et Thérapeutique*, 8(3), pp. 141–145. doi: 10.1016/j.nephro.2011.07.411.
- Dai, X., Chen, A. et Bai, Z (2014). 'Integrative investigation on breast cancer in ER, PR and HER2- defined subgroups using mRNA and miRNA expression profiling', *Représentant Sci*, 4:6566. doi: 10.1038 / srep06566.

-
- Dai, X. et al. (2015) ‘Breast cancer intrinsic subtype classification, clinical use and future trends’, *Am. J. Cancer Res.* 5(10), pp. 2929–2943. PMID: 26693050
 - Dinarello, C. A. (2005) ‘Interleukin-1 β ’, *Critical Care Medicine*, 33(12 SUPPL.), pp. 460–462. doi: 10.1097/01.CCM.0000185500.11080.91.
 - El-Omar, E. M. et al. (2000) ‘Interleukin-1 polymorphisms associated with increased risk of gastric cancer’, *Nature*. Nature, 404(6776), pp. 398–402. doi: 10.1038/35006081.
 - Engel, R. H. et Kaklamani, V. G. (2007) ‘HER2-positive breast cancer: current and future treatment strategies’, *Drugs*, 67 (9), pp. 1329–1341. doi: 10.2165 / 00003495-200767090-00006.
 - Escobar, P. et al. (2015) ‘IL-1 β produced by aggressive breast cancer cells is one of the factors that dictate their interactions with mesenchymal stem cells through chemokine production’, *Oncotarget*. Impact Journals LLC, 6(30), pp. 29034–29047. doi: 10.18632/oncotarget.4732.
 - Esquivel-Velázquez, M. et al. (2015) ‘The role of cytokines in breast cancer development and progression’, *Journal of Interferon and Cytokine Research*. Mary Ann Liebert Inc., pp. 1–16. doi: 10.1089/jir.2014.0026.
 - Figueroa-Magalhães, M. C., Jelovac, D., Connolly, R. M. et Wolff, A. C. (2014) ‘Treatment of HER2- positive breast cancer’. *Breast Edinb. Scotl.* 23(2), pp. 128–136. doi: 10.1016 / j.breast.2013.11.011.
 - Franz, M. et al. (2018) ‘GeneMANIA update 2018’, *Nucleic Acids Research*. Oxford University Press, 46(W1), pp. W60–W64. doi: 10.1093/nar/gky311.
 - Garlanda, C., Dinarello, C. A. et Mantovani, A. (2013) ‘The Interleukin-1 Family: Back to the Future’, *Immunity*. Immunity, pp. 1003–1018. doi: 10.1016/j.immuni.2013.11.010.
 - Giordano, S. B. et Gradishar, W. (2017) ‘Breast cancer: Updates and advances in 2016’, *Current Opinion in Obstetrics and Gynecology*. Lippincott Williams and Wilkins, pp. 12–17. doi: 10.1097/GCO.0000000000000343.
 - Giraldo, S. et al. (2008) ‘IL1B (interleukin 1, beta)’, *Atlas of Genetics and Cytogenetics in Oncology and Haematology*, 13(4), pp. 273–275. doi: 10.4267/2042/44449.
 - Gong, Z. et al. (2013) ‘Innate Immunity Pathways and Breast Cancer Risk in African American and European-American Women in the Women’s Circle of Health Study (WCHS)’, *PLoS ONE*. Public Library of Science, 8(8). doi: 10.1371/journal.pone.0072619.
 - Grimm, C. et al. (2009) ‘The prognostic value of four interleukin-1 gene polymorphisms in caucasian women with breast cancer - A multicenter study’, *BMC Cancer*. BioMed Central, 9(1), p. 78. doi: 10.1186/1471-2407-9-78.
 - H. Guendouz, W. Chetibi, A. Abdelouahab, A. B. (2011) ‘Cancer du sein de la femme de moins de 35 ans : étude rétrospective à propos de 612 cas’, *Cancer/Radiothérapie*, 14(6–7), p. 662. doi: 10.1016/j.canrad.2010.07.625.

-
- Hall, S. K. *et al.* (2004) ‘Correlation of polymorphic variation in the promoter region of the interleukin-1 β gene with secretion of interleukin-1 β protein’, *Arthritis and Rheumatism*. *Arthritis Rheum*, 50(6), pp. 1976–1983. doi: 10.1002/art.20310.
 - Hefler, L. A. *et al.* (2005) ‘Interleukin-1 and interleukin-6 gene polymorphisms and the risk of breast cancer in Caucasian women’, *Clinical Cancer Research*. *Clin Cancer Res*, 11(16), pp. 5718–5721. doi: 10.1158/1078-0432.CCR-05-0001.
 - Hernandez-Guerrero, C. *et al.* (2003) ‘In-vitro secretion of proinflammatory cytokines by human amniochorion carrying hyper-responsive gene polymorphisms of tumour necrosis factor- α and interleukin-1 β ’, *Molecular Human Reproduction*. Oxford Academic, 9(10), pp. 625–629. doi: 10.1093/molehr/gag076.
 - Honma, S. *et al.* (2002) ‘The influence of inflammatory cytokines on estrogen production and cell proliferation in human breast cancer cells’, *Endocrine Journal*. Japan Endocrine Society, 49(3), pp. 371–377. doi: 10.1507/endocrj.49.371.
 - Hortobagyi, G. N. *et al.* (2005) ‘The global breast cancer burden: Variations in epidemiology and survival’, *Clinical Breast Cancer*. Elsevier Inc., pp. 391–401. doi: 10.3816/CBC.2005.n.043.
 - Ito, L. S. *et al.* (2002) ‘Significant reduction in breast cancer risk for Japanese women with Interleukin 1B -31 CT/TT relative to CC genotype’, *Japanese Journal of Clinical Oncology*, 32(10), pp. 398–402. doi: 10.1093/jjco/hyf081.
 - J. Lansac, L. O. (2018) *cancer du sein*. Ninth Edit, *Gynecologie pour le Praticien*. Ninth Edit. Elsevier Masson SAS. doi: <https://doi.org/10.1016/B978-2-294-75637-5.00019-3>.
 - Jin, L. *et al.* (1997) ‘Expression of interleukin-1 β in human breast carcinoma’, *Cancer*. *Cancer*, 80(3), pp. 421–434. doi: 10.1002/(SICI)1097-0142(19970801)80:3<421::AID-CNCR10>3.0.CO;2-Z.
 - Joyon, N., Penault-Llorca, F. and Lacroix-Triki, M. (2017) ‘Classification et signatures moléculaires des cancers du sein en 2017’, *Oncologie*, 19(3–4), pp. 64–70. doi: 10.1007/s10269-017-2700-6.
 - Kaarvatn, M. H. *et al.* (2012) ‘Interleukin-1 gene locus polymorphisms are associated with risk to breast cancer in Croatian population’, *Periodicum Biologorum*, 114(4), pp. 497–503.
 - Karki, R. *et al.* (2015) ‘Defining “mutation” and “polymorphism” in the era of personal genomics’, *BMC Medical Genomics*. BioMed Central Ltd. doi: 10.1186/s12920-015-0115-z.
 - Kent, W. J. *et al.* (2010) ‘BigWig and BigBed: Enabling browsing of large distributed datasets’, *Bioinformatics*. Oxford Academic, 26(17), pp. 2204–2207. doi: 10.1093/bioinformatics/btq351.
 - Kurtzman, S. H. *et al.* (1999) ‘Cytokines in human breast cancer: IL-1 α and IL-1 β expression.’, *Oncology Reports*. Spandidos Publications, 6(1). doi: 10.3892/or.6.1.65.
 - Lecarpentier, J. et Andrieu, N. (2012) ‘Étude Des Facteurs Modificateurs Du Risque De

- Lee, Y. H. et Song, G. G. (2014) ‘A meta-analysis of the association between CTLA-4 +49 A/G, -318 C/T, and IL-1 polymorphisms and susceptibility to cervical cancer’, *Neoplasma*. SAP - Slovak Academic Press, spol. s.r.o., 61(4), pp. 481–490. doi: 10.4149/neo_2014_060.
- Li, P. *et al.* (1995) ‘Mice deficient in IL-1 β -converting enzyme are defective in production of mature IL-1 β and resistant to endotoxic shock’, *Cell*. Cell, 80(3), pp. 401–411. doi: 10.1016/0092-8674(95)90490-5.
- Lind, H., Haugen, A. et Zienolddiny, S. (2007) ‘Differential binding of proteins to the IL1B -31 T/C polymorphism in lung epithelial cells’, *Cytokine*. Academic Press, 38(1), pp. 43–48. doi: 10.1016/j.cyto.2007.05.001.
- Liu, J. *et al.* (2006) ‘Functional variants in the promoter of interleukin-1 β are associated with an increased risk of breast cancer: A case-control analysis in a Chinese population’, *International Journal of Cancer*. Int J Cancer, 118(10), pp. 2554–2558. doi: 10.1002/ijc.21652.
- Liu, X. *et al.* (2010) ‘Three polymorphisms in interleukin-1 β gene and risk for breast cancer: A meta-analysis’, *Breast Cancer Research and Treatment*, 124(3), pp. 821–825. doi: 10.1007/s10549-010-0910-3.
- Massaad, C. et Barouki, R. (1999) ‘An assay for the detection of xenoestrogens based on a promoter containing overlapping EREs’, *Environ Health Perspect*, 107(7) , pp. 563-536. doi: 10.1289/ehp.99107563.
- Machado, J. C. *et al.* (2001) ‘Interleukin 1B and interleukin 1RN polymorphisms are associated with increased risk of gastric carcinoma’, *Gastroenterology*. Gastroenterology, 121(4), pp. 823–829. doi: 10.1053/gast.2001.28000.
- Martin, T. A. et Jiang, W. G. (2009) ‘Loss of tight junction barrier function and its role in cancer metastasis’, *Biochim. Biophys. Acta BBA – Biomembr*, 1788(4), pp. 872–891. doi:10.1016/j.bbamem.2008.11.005.
- Mehemmai, C. *et al.* (2019). ‘BRCA1 and BRCA2 Germline Mutation Analysis in Hereditary Breast/Ovarian Cancer Families from the Aures Region (Eastern Algeria): First Report’, *Pathol Oncol Res*, 26(2), pp. 715-726. doi: 10.1007/s12253-019-00586-4.
- Miller, D. K. *et al.* (1993) ‘Purification and characterization of active human interleukin-1 β -converting enzyme from THP.1 monocytic cells’, *Journal of Biological Chemistry*, 268(24), pp. 18062–18069.
- Miller, L. J. *et al.* (2000) ‘Interleukin-1 family expression in human breast cancer: Interleukin-1 receptor antagonist’, *Cancer Investigation*, 18(4), pp. 293–302. doi: 10.3109/07357900009012171.
- Momenimovahed, Z. et Salehiniya, H. (2019) ‘Epidemiological characteristics of and risk factors for breast cancer in the world’, *Breast Cancer: Targets and Therapy*. Dove Medical Press Ltd., pp. 151–164. doi: 10.2147/BCTT.S176070.

-
- Newman, V. *et al.* (2018) ‘The ensembl genome browser: Strategies for accessing eukaryotic genome data’, in *Methods in Molecular Biology*. Humana Press Inc., pp. 115–139. doi: 10.1007/978-1-4939-7737-6_6.
 - Nicklin, M. J. H. *et al.* (2002) ‘A sequence-based map of the nine genes of the human interleukin-1 cluster’, *Genomics*. Academic Press Inc., 79(5), pp. 718–725. doi: 10.1006/geno.2002.6751.
 - Pérez-Ramírez, C. *et al.* (2017) ‘Interleukins as new prognostic genetic biomarkers in non-small cell lung cancer’, *Surgical Oncology*. Elsevier Ltd, 26(3), pp. 278–285. doi: 10.1016/j.suronc.2017.05.004.
 - Perou, C. M. *et al.* (2000) ‘Molecular portraits of human breast tumours’, *Nature*. Nature, 406(6797), pp. 747–752. doi: 10.1038/35021093.
 - Perrier, S., Caldefie-Chézet, F. et Vasson, M. P. (2009) ‘IL-1 family in breast cancer: Potential interplay with leptin and other adipocytokines’, *FEBS Letters*. FEBS Lett, pp. 259–265. doi: 10.1016/j.febslet.2008.12.030.
 - Piñero, J. *et al.* (2015) ‘DisGeNET: A discovery platform for the dynamical exploration of human diseases and their genes’, *Database*. Oxford University Press, 2015. doi: 10.1093/database/bav028.
 - *Plan national CANCER 2015 - 2019* (2014).
 - Pooja, S. *et al.* (2012) ‘Polymorphic variations in IL-1 β , IL-6 and IL-10 genes, their circulating serum levels and breast cancer risk in Indian women’, *Cytokine*. Elsevier Ltd, 60(1), pp. 122–128. doi: 10.1016/j.cyto.2012.06.241.
 - Prat, A. *et al.* (2010) ‘Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer’, *Breast Cancer Res*. BCR 12, R68. doi: 10.1186/bcr2635
 - Rébé, C. et Ghiringhelli, F. (2020) ‘Interleukin-1 β and cancer’, *Cancers*. MDPI AG, pp. 1–31. doi: 10.3390/cancers12071791.
 - Rogers, A. *et al.* (2007) ‘Different effects of raloxifene and estrogen on interleukin-1 β and interleukin-1 receptor antagonist production using in vitro and ex vivo studies’, *Bone*, 40(1), pp. 105–110. doi: 10.1016/j.bone.2006.07.011.
 - Roussel-Vavasseur, L. (2008) *Caractérisation de l’interleukine-33, une cytokine nucléaire associée à la chromatine*.
 - Schnitt SJ. (2010) ‘Classification and prognosis of invasive breast cancer: from morphology to molecular taxonomy’, *Mod Pathol*, 23 Suppl 2: S60-4. doi: 10.1038/modpathol.33. PMID: 20436504.
 - Singhai, R. *et al.* (2011) ‘E-Cadherin as a diagnostic biomarker in breast cancer’, *North Am. J. Med. Sci.* 3, pp. 227–23.
 - Slamon, D. J. *et al.* (1987) ‘Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene’, *Science*, 235(4785), pp. 177–182. doi: 10.1126

/ science.3798106.

- Slattery, M. L. *et al.* (2014) 'Genetic variants in interleukin genes are associated with breast cancer risk and survival in a genetically admixed population: The breast cancer health disparities study', *Carcinogenesis*, 35(8), pp. 1750–1759. doi: 10.1093/carcin/bgu078.
- Smith, A. J. P. et Humphries, S. E. (2009) 'Cytokine and cytokine receptor gene polymorphisms and their functionality', *Cytokine and Growth Factor Reviews*. Cytokine Growth Factor Rev, pp. 43–59. doi: 10.1016/j.cytogfr.2008.11.006.
- Snoussi, K. *et al.* (2005) 'Genetic variation in pro-inflammatory cytokines (interleukin-1 β , interleukin-1 α and interleukin-6) associated with the aggressive forms, survival, and relapse prediction of breast carcinoma', *European Cytokine Network*, 16(4), pp. 253–260.
- Speirs, V. *et al.* (1999) 'Evidence for transcriptional activation of ER α by IL-1 β in breast cancer cells', *International journal of oncology*. Int J Oncol, 15(6), pp. 1251–1254. doi: 10.3892/ijo.15.6.1251.
- Sun, Y. S. *et al.* (2017) 'Risk factors and preventions of breast cancer', *International Journal of Biological Sciences*. Ivyspring International Publisher, pp. 1387–1397. doi: 10.7150/ijbs.21635.
- Tagu, D. et Risler, J.-L. (2010) '*Bio-informatique _ Principes d'utilisation des outils*'. Quae, 270 pages. Available at:
https://books.google.dz/books?hl=fr&lr=&id=WHL0bN9Mh5wC&oi=fnd&pg=PA1&dq=info:30LkvQJHrmQJ:scholar.google.com/&ots=TuYimFHOn5&sig=Xk27ljk4BjIOWQ9gBdPoE45yYTY&redir_esc=y#v=onepage&q&f=false.
- Tulotta, C. et Ottewell, P. (2018) 'The role of IL-1B in breast cancer bone metastasis', *Endocrine-Related Cancer*. BioScientifica Ltd, pp. R421–R434. doi: 10.1530/ERC-17-0309.
- Vince, J. E. et Silke, J. (2016) 'The intersection of cell death and inflammasome activation', *Cellular and Molecular Life Sciences*. Birkhauser Verlag AG, pp. 2349–2367. doi: 10.1007/s00018-016-2205-2.
- Vincent, S et Sigal, Z. (2009) 'Quel est l'apport de la biologie pour les patientes? Vers une nouvelle classification des cancers du sein : apport de la biologie moléculaire', Toward a new breast carcinoma classification based on molecular biology.
- Wirapati, P. *et al.* (2008) 'Meta-analysis of gene expression profiles in breast cancer: Toward a unified understanding of breast cancer subtyping and prognosis signatures', *Breast Cancer Research*. BioMed Central Ltd., 10(4). doi: 10.1186/bcr2124.
- Xu, J. *et al.* (2013) 'Systematic Review and Meta-Analysis on the Association between IL-1B Polymorphisms and Cancer Risk', *PLoS ONE*. Edited by V. De Re. Public Library of Science, 8(5), p. e63654. doi: 10.1371/journal.pone.0063654.
- Zendehdel, M. *et al.* (2018) 'Subtypes of Benign Breast Disease as a Risk Factor for Breast Cancer: A Systematic Review and Meta-Analysis Protocol.', *Iranian journal of medical sciences*. Shiraz University of Medical Sciences, 43(1), pp. 1–8. PMID: 29398746 ; PMCID:

PMC5775989.

- Zhanjiang (John) Liu (2017) *Bioinformatics in aquaculture: principles and methods*. illustrée. Edited by John Wiley & Sons. Available at: <https://books.google.dz/books?id=Up7zDQAAQBAJ&pg=PA255&dq=Galaxy+Genome+bioinformatics&hl=fr&sa=X&ved=2ahUKEwiR8OiJo7vqAhVfRhUIHSLkD38Q6AEwAXoECAIQAg#v=onepage&q=Galaxy+Genome+bioinformatics&f=false>.
- Zuo, X. *et al.* (2018) 'Interleukin gene polymorphisms in Chinese Han population with breast cancer, a case-control study', *Oncotarget*. Impact Journals LLC, 9(26), pp. 17994–18001. doi: 10.18632/oncotarget.23157.

Annexes

Annexe 01 : Tableau de localisation des exons d'IL1B selon Galaxy.

Chrom	Start	End	Name	Score	Strand
chr2	112830360	112830573	ENST000002633417_cds_0_0_chr2_112830361_r	0	-
chr2	112831291	112831422	ENST000002633417_cds_1_0_chr2_112831292_r	0	-
chr2	112832661	112832826	ENST000002633417_cds_2_0_chr2_112832662_r	0	-
chr2	112833373	112833575	ENST000002633417_cds_3_0_chr2_112833374_r	0	-
chr2	112835565	112835617	ENST000002633417_cds_4_0_chr2_112835566_r	0	-
chr2	112836182	112836229	ENST000002633417_cds_5_0_chr2_112836183_r	0	-
chr2	112832698	112832826	ENST000004188175_cds_0_0_chr2_112832699_r	0	-
chr2	112833373	112833575	ENST000004188175_cds_1_0_chr2_112833374_r	0	-
chr2	112835565	112835617	ENST000004188175_cds_2_0_chr2_112835566_r	0	-
chr2	112836182	112836229	ENST000004188175_cds_3_0_chr2_112836183_r	0	-
chr2	112832751	112832826	ENST000004320185_cds_0_0_chr2_112832752_r	0	-
chr2	112833373	112833575	ENST000004320185_cds_1_0_chr2_112833374_r	0	-
chr2	112835565	112835617	ENST000004320185_cds_2_0_chr2_112835566_r	0	-
chr2	112836182	112836229	ENST000004320185_cds_3_0_chr2_112836183_r	0	-
chr2	112832794	112832826	ENST000004167501_cds_0_0_chr2_112832795_r	0	-
chr2	112833373	112833575	ENST000004167501_cds_1_0_chr2_112833374_r	0	-
chr2	112835565	112835617	ENST000004167501_cds_2_0_chr2_112835566_r	0	-
chr2	112836182	112836229	ENST000004167501_cds_3_0_chr2_112836183_r	0	-

Annexe 02 : Tableau d'information SNP d'IL1B selon Galaxy.

Chrom	Start	End	Name	Score	Stra
chr2	112829855	112829856	rs1071676	0	-
chr2	112830259	112830260	rs10176689	0	+
chr2	112830724	112830725	rs1143643	0	-
chr2	112830814	112830815	rs3917380	0	-
chr2	112830935	112830936	rs3917379	0	-
chr2	112830975	112830976	rs1143642	0	-
chr2	112831088	112831089	rs1143641	0	-
chr2	112831178	112831180	rs1143640	0	-
chr2	112831215	112831216	rs1143639	0	-
chr2	112831491	112831492	rs2708929	0	-
chr2	112831755	112831756	rs1143637	0	-
chr2	112831993	112831994	rs10183049	0	+
chr2	112832812	112832813	rs1143634	0	-
chr2	112832889	112832890	rs1143633	0	-
chr2	112833697	112833698	rs3136558	0	-
chr2	112833764	112833765	rs3136557	0	-
chr2	112834014	112834014	rs3917357	0	-
chr2	112834015	112834016	rs1143631	0	-
chr2	112834077	112834078	rs1143630	0	-
chr2	112834785	112834786	rs3917356	0	-
chr2	112834828	112834829	rs2853547	0	-
chr2	112835315	112835316	rs3917354	0	-
chr2	112835317	112835318	rs397841817	0	-
chr2	112835371	112835372	rs3917352	0	-
chr2	112835382	112835383	rs3917351	0	-
chr2	112835940	112835941	rs1143629	0	-
chr2	112836627	112836629	rs3917347	0	-
chr2	112836809	112836810	rs1143627	0	-