

République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
جامعة أبو بكر بلقايد- تلمسان
Université ABOUBEKR BELKAID – TLEMCEM
كلية علوم الطبيعة والحياة، وعلوم الأرض والكون
Faculté des Sciences de la Nature et de la Vie, et Sciences de la Terre et de
l'Univers
Département Biologie



MÉMOIRE

Présenté par

BENSMINE Assala

En vue de l'obtention du

Diplôme de MASTER

En Nutrition et diététique

Thème

Analyse *in silico* des nsSNP associés aux allergies alimentaires

Soutenu le 5 / 10/ 2022, devant le jury composé de :

Président	Aribi Mourad	(Professeur)	Abou BakrBelkaid
Encadrant	Hadjidj Zineb	(Professeur)	Abou BakrBelkaid
Examineur	Brahmi Nabila	(MCA)	Abou BakrBelkaid

Année universitaire 2021/2022

Résumé

Les allergies alimentaires (FA) sont des problèmes majeurs qui affectent l'être humain. En effet, les changements environnementaux et nutritionnels au cours des dernières décennies ont conduit à une augmentation de sa prévalence dans le monde, mais les données épidémiologiques précises manquent. Les symptômes des allergies alimentaires peuvent mettre la vie en danger, tandis que les symptômes des intolérances alimentaires sont nettement moins graves. Leur diagnostic repose sur une combinaison d'approches pour éviter les résultats faux positifs ou faux négatifs, et il n'existe actuellement aucun remède à ces effets indésirables. Ils sont gérés par un régime qui exclut l'allergène coupable de l'alimentation quotidienne d'un individu. Il est bien connu que les deux facteurs de risque les plus importants pour le développement d'allergies alimentaires sont les facteurs génétiques et environnementaux, ainsi que les interactions entre eux.

Les études d'association à l'échelle du génome (GWAS), souvent des polymorphismes nucléotidiques simples (SNP), ont identifié et ouvert la voie à la compréhension de l'implication de ces gènes sensibles. Les gènes STAT3, FOXP3 sont les gènes les mieux étudiés liés à l'allergie alimentaire

Les analyse *in silico* des nsSNP aident à déterminer leurs impacts et les effets sur la structure et la fonction des protéines à l'aide de divers outils bioinformatiques disponibles en ligne.

Mots clés : FA, GWAS, SNP, nsSNP, STAT3, FOXP3, Analyse *in silico*.

ملخص

تعتبر الحساسية الغذائية (FA) من المشاكل الرئيسية التي تؤثر على البشر. في الواقع ، أدت التغيرات البيئية والتغذية في العقود الأخيرة إلى زيادة انتشاره في جميع أنحاء العالم ، لكن البيانات الوبائية الدقيقة غير متوفرة. يمكن أن تكون أعراض الحساسية الغذائية مهددة للحياة ، بينما تكون أعراض عدم تحمل الطعام أقل حدة بشكل ملحوظ. يعتمد تشخيصهم على مجموعة من الأساليب لتجنب النتائج الإيجابية أو السلبية الكاذبة ، ولا يوجد حالياً علاج لهذه الآثار الضارة. تتم إدارتها من خلال نظام غذائي يستبعد مسببات الحساسية من النظام الغذائي اليومي للفرد. من المعروف أن أهم عاملين من عوامل الخطر لتطور الحساسية الغذائية هما العوامل الوراثية والبيئية ، وكذلك التفاعلات بينهما.

حددت دراسات الارتباط على مستوى الجينوم (GWAS) ، والتي غالباً ما تكون متعددة الأشكال للنيوكليوتيدات المفردة (SNPs) ، ومهدت الطريق لفهم مشاركة هذه الجينات الحساسة. تعد جينات STAT3 وFOXP3 من أفضل الجينات المدروسة المرتبطة بحساسية الطعام

تساعد تحليلات السليكو لـnsSNPs في تحديد آثارها وتأثيراتها على بنية البروتين ووظيفته باستخدام أدوات المعلوماتية الحيوية المختلفة المتاحة عبر الإنترنت.

الكلمات المفتاحية : FA, GWAS, SNP,nsSNP ,STAT3,FOXP3,insilicoanalysis

Abstract

Food allergies (FA) are major problems that affect human beings. Indeed, environmental and nutritional changes in recent decades have led to an increase in its prevalence worldwide, but precise epidemiological data are lacking. Symptoms of food allergies can be life-threatening, while symptoms of food intolerances are significantly less severe. Their diagnosis relies on a combination of approaches to avoid false positive or false negative results, and there is currently no cure for these adverse effects. They are managed by a diet that excludes the culprit allergen from an individual's daily diet. It is well known that the two most important risk factors for the development of food allergies are genetic and environmental factors, as well as the interactions between them.

Genome-wide association studies (GWAS), often single nucleotide polymorphisms (SNPs), have identified and paved the way for understanding the involvement of these sensitive genes. STAT3, FOXP3 genes are the best studied genes linked to food allergy

In silico analyses of nsSNPs help determine their impacts and effects on protein structure and function using various bioinformatics tools available online.

Keywords :FA, GWAS, SNP, nsSNP , STAT3, FOXP3, *in silico* analysis

Remerciements

Au terme de ce travail, nous devons remercier tout d'abord dieu qui nous a donné la force et le courage de suivre nos études et d'arriver à ce stade et à nos parents qui nous ont beaucoup soutenus pendant tout le long de notre parcours.

Un grand merci à mon encadreur Dr. HADJIDJ Zineb Qu'elle m'a beaucoup aidé, soutenu et me permis d'arriver à ce niveau-là et pour ses excellents conseils et surtout pour son temps passé avec nous et sa patience, sans elle j'aurais pas pu réaliser ce modeste travail et pour sa confiance en moi.

Je tenais aussi à remercier les membres de jury qui m'ont fait honneur d'examiner ce travail.

Enfin, je renouvelle mes remerciements à ceux qui m'ont aidés de près ou de loin pour réaliser ce travail sans oublier les enseignants qui ont contribué à notre formation.

Dédicace

Merci mon dieu de m'avoir donné la capacité d'écrire et de réfléchir, la patience d'aller jusqu'au bout du rêve J'ai l'honneur de dédier ce modeste travail :

A ma mère celle qui m'a donné la vie, le symbole de tendresse qui s'est sacrifiée pour mon bonheur et ma réussite.

A mon père, école de mon enfance, qui a été mon ombre durant toutes les années des études et qui a veillé tout au long de ma vie à m'encourager à me donner l'aide et à me protéger.

A à ma sœurs Ayate et mon frère Achrafà Oussama, je vous souhaite un avenir plein de joie de bonheur de réussite et de sérénités, je vous exprime à travers ce travail mes sentiments de fraternité et d'amour.

A tous les membres de ma famille, petite et grande.

A ma chère amie : Meriem.

A mon encadreur Docteur HADJIDJ Zineb.

A tous ceux qui ont sacrifié leur temps pour la science et à tous ceux qui utilisent la science pour le bien et la prospérité de l'humanité.

BENSMINE Assala

LISTES D'ABREVIATIONS

ARN : Acide Rubonucléique.

dbSNP : base de données Single Nucleotide Polymorphism.

GWAS : Genome Wide Association Studies.

NCBI: National Center for Biotechnology Information.

nsLTP :

NsSNP : Non synonyme polymorphisme nucléotidique simple.

PolyPhen v2 : Polymorphism Phenotyping v2.

PROVEAN: Protein Variation Effect Analyzer.

PP2A: serine/threonine phosphatase 2A.

rs: Rapport SNP.

SIFT: Sorting Intolerant from Tolerant.

SNAP2: Screening of Non-Acceptable Polymorphism 2.

SNP : Polymorphisme nucléotidique simple.

Stat3 : Signal Transducer And Activator Of Transcription 3.

LISTE DES FIGURES

Figure 1 : Classification Des Allergies Alimentaire.

Figure 2 : Single nucleotide polymorphism (SNP).

Figure 2.1 : SNP localisation.

Figure 2.2 : Études d'association à l'échelle du génome (GWAS) (<http://mmg-233-2014-geneticsgenomics.wikia.com/wiki/File:GWAS.jpg>).

LISTE DES TABLEAUX

Tableau 1 : Résultats des nsSNP à haut risques identifiés la méthode in silico SIFT pour le STAT3.

Tableau 2 : Liste des résultats de l'analyse des nsSNPs FOXP3 par SIFT

Table des matières

Résumé

ملخص

Abstract

Remerciement

Dédicace

Listes des abréviations

Liste des figures

Liste des tableaux

PARTIE BIBLIOGRAPHIQUE

Introduction.....	2
Chapitre I : Revue de la littérature.....	4
1.1. Allergies alimentaires.....	5
1.1.1. Historique.....	5
1.1.2. Mécanismes des allergies alimentaires	6
1.1.3. Classification des allergies alimentaires.....	6
1.1.4. Impact des allergies alimentaire sur la qualité de vie.....	8
1.1.5. Méthodes de diagnostique.....	8
1.1.6. Traitements.....	9
1.2. Polymorphisme nucléotidique simple « SNP ».....	10
1.2.1. Généralités	10
1.2.2. L'effet de la localisation des polymorphismes nucléotidique.....	11
1.2.3. Types de polymorphisme nucléotidique	11
1.2.4. Études d'association à l'échelle du génome (GWAS).....	12
1.2.5. Application de polymorphisme nucléotidique.....	13
1.2.5.1. Les SNP comme marqueurs biologiques de maladies humaines.....	13
1.2.5.2. Polymorphisme nucléotidique et le développement de médicaments.....	13
1.2.6. L'avantage de polymorphisme nucléotidique.....	14
1.2.7. Rôle des polymorphismes nucléotidiques dans le traitement des allergies alimentaire.....	14
1.2.8. Analyse in silico des SNP.....	15

1.2.9. Implication de gène STAT3 dans les allergies alimentaires.....	15
1.2.10. Implication de gène FOXP3 dans les allergies alimentaires.....	15
II. Chapitre II : Matériels et Méthodes.....	16
2.1. Recherche des nsSNPs et des séquences protéiques.....	18
2.2. Récupération des données des nsSNP.....	18
2.3. Prédiction des nsSNP délétères par différents outils bioinformatique.....	19
2.4. L'identification de nsSNP sur les domaines des gènes.....	20
2.5. Analyse de l'effet des nsSNP sur la stabilité des protéines.....	20
2.6. Analyse de la conservation évolutive des protéines.....	21
2.7. Modélisation des emplacements nsSNP sur la structure des protéines.....	21
2.8. Base de données sur les polymorphismes nucléotidiques simple (SNP) et gène Ontologie (GO) (SNPs&GO).....	21
III. Chapitre III : Résultats et Interprétation.....	22
3.1. Récupération des nsSNP dans les gènes STAT3 et FOXP3.....	
3.2. Prédiction des nsSNP délétères par SIFT.....	
3.3. Identification des domaines des protéines STAT3 et FOXP3 et catégorisation des nsSNP dans ces domaines.....	
IV. Chapitre IV : Discussion.....	
V. Chapitre V : Conclusion et Perspective	28
VI. Chapitre VI : Bibliographie.....	30

INTRODUCTION

Introduction

L'allergie est une hypersensibilité trouble de l'intégrité du système immunitaire. Les réactions allergiques se produisent normalement lors de contact avec des substances environnementales inoffensives connus sous le nom d'allergènes ; ces réactions sont acquises, prévisible et rapides. (**byron Y Newman, 2011**). Les allergènes sont des substances que l'on retrouve dans notre environnement et qui peuvent être à l'origine de réactions de défense de la part de notre organisme ; on a les pneumallergènes, les allergènes médicamenteux, les substances provoquant des allergies de contact, le venin des insectes et les allergènes alimentaires.

L'allergie alimentaire est une maladie multifactorielle complexe dans laquelle on pense que des facteurs de risque environnementaux et génétiques contribuent à sa pathogénèse. Il déclenche une réponse immunitaire anormale après un contact avec certaines protéines alimentaires, entraînant des réactions cliniques indésirables, notamment l'anaphylaxie, qui peut mettre la vie en danger (**Allen et al, 2006**).

Un polymorphisme mononucléotidique(SNP, *single nucleotide polymorphism*)est une modification unique dans une séquence nucléotidique d'ADN qui peut ou non affecter le phénotypiques d'un organisme.Les SNP non synonymes (nsSNP) se trouve dans les régions non codantes peuvent modifier la synthèse des protéineson affectant l'épissage des gènes, l'expression des gènes et la séquence d'ARN non codant.

Les SNP peuvent jouer un rôle direct ou indirect dans l'expression phénotypique dans la pathogénèse des allergies alimentaires. Cependant, il y a quelques études qui ont exploré l'association entre les SNP et le pronostic des allergies alimentaires. (**Paula Brown et al, 2012**)

Dans notre étude, nous effectuons une analyse in silico des polymorphismes nucléotidiques (SNP) ayant des conséquences fonctionnelles sur le STAT3 et FOXP3, des associées aux allergies alimentaire afin de les identifier in silico et les proposer à des tests de confirmation in vitro et in vivo et explorer de nouvelles perspectives dans le diagnostic et le traitement de ces allergies.

L'objectif de cette étude est de réaliser des analyses *insilico* SNP dans les gènes impliqués dans les allergies alimentaires et les identifier en considérant leurs effets dans ces gènes, déterminer les SNPs fonctionnels et les SNPs qui ont un effet sur la stabilité des protéines, déterminer les SNPs qui ont un effet sur la conservation de la protéine.

Le but de notre étude est d'identifier les SNPs du gène STAT3 et FOXP3 par des outils *insilico* pour la prévention du cancer

Chapitre I : Revue de la littérature

I. REVUE DE LA LITTÉRATURE

1.1. Allergies alimentaires

Les maladies allergiques sont considérées comme un problème de santé important pour l'ensemble de la société. L'allergie ou l'hypersensibilité aux aliments est déclenchée par le système immunitaire, tandis que l'intolérance alimentaire est causée par un manque d'enzymes ou par des substances pharmaco-logiquement actives dans les aliments (**Muraro et al, 2014**). Au cours des 20 dernières années, des efforts ont été consacrés à l'identification des protéines allergènes dans les aliments d'origine végétale et animale, à l'étude de leurs propriétés physicochimiques et de leurs interactions avec les cellules immunitaires. Par la suite, des bases de données sur les allergènes a été créée et maintenue pour fournir des informations détaillées sur les allergènes (par exemple, www.allergen.org). De toute évidence, seules quelques-unes de toutes les familles de protéines connues contiennent des allergènes alimentaires (**Radauer et al, 2008**). La famille de protéines la plus pertinente pour les allergènes alimentaires végétaux est l'albumine 2S, dérivée des graines (ex. sésame, tournesol...) et des fruits à coque (noisettes, noix, noix du Brésil...) et des arachides. Les protéines de stockage des graines identifiées, suivies des non -les protéines de transfert des lipides spécifiques (nsLTP), représentent les allergènes majeurs des fruits tels que les pêches, les pommes et les noix (**Hoffmann-Sommergruber et Mills, 2009**). De plus, le principal allergène de poisson, la parvalbumine, est présent dans la plupart des poissons (**Hoffmann-Sommergruber, 2016**).

1.1.1 Historique

Les sensibilités alimentaires sont des réponses physiologiques anormales à des aliments spécifiques. Pour la grande majorité des consommateurs, le même aliment est sans danger. Les allergies alimentaires peuvent affecter les humains depuis l'Antiquité. Les premiers cas documentés d'allergie alimentaire se sont produits au début du XXe siècle, mais jusqu'à récemment, l'allergie alimentaire a été largement ignorée (**Taylor, 2000**). Les allergies alimentaires les plus courantes sont le lait (CML), les œufs, le blé et les arachides (**Urisu et al, 2011**). Allergies alimentaires courantes autodéclarées (**Eriksson et al, 2004**). Les réactions alimentaires ne sont pas nouvelles et sont décrites depuis plus de deux mille ans. L'ancien

médecin grec Hippocrate a décrit une réaction au lait au 1er siècle. Des réactions allergiques aux œufs et aux poissons ont été décrites dès les XVIe et XVIIe siècles(**Sampson, 1999**).

1.1.2. Mécanismes des allergies alimentaires

La prévalence des allergies alimentaires augmente et le développement d'approches diagnostiques, préventives et thérapeutiques plus précises nécessite une meilleure compréhension des mécanismes sous-jacents. La tolérance orale est la réponse physiologique normale aux antigènes ingérés, et la perturbation de ce processus peut entraîner une sensibilité aux allergènes alimentaires (**Sampson et Sicherer, 2014**).

Des recherches sont en cours pour mieux comprendre les mécanismes conduisant à la sensibilisation et à la maladie et à la désensibilisation et à la tolérance à court et à long terme(**Hussain et al, 2018**). De nombreux modèles animaux ont été développés pour étudier les événements cellulaires et moléculaires qui conduisent à la sensibilisation et à la tolérance aux réponses allergiques(**Galand et al, 2016**). L'une des principales découvertes est que l'administration orale de protéines aux animaux induit généralement une tolérance, mais conduit à une sensibilisation et à des maladies allergiques (**Oyoshiet al, 2014**).Il a été démontré que les réponses des modèles animaux sont influencées par des facteurs qui compromettent la barrière épithéliale (**Barcik et al, 2016**). Ces modèles suggèrent également que la sensibilité aux allergènes alimentaires peut en fait se produire par d'autres sites, tels que les voies respiratoires ou la peau, contrairement à l'intestin, où la tolérance orale est souvent la réponse par défaut. Les résultats de ces modèles de souris soutiennent également l'observation selon laquelle une perturbation précoce de la barrière cutanée causée par une inflammation ou des anomalies génétiques (par exemple, des mutations du gène de la filaggrine) est associée à des taux accrus de sensibilisation alimentaire chez les sujets humains(**Kelleher et al, 2016**). Cependant, des études sur les réponses IgE et la digestibilité des protéines alimentaires ont montré que les voies d'exposition orale sont également des voies importantes de sensibilité aux allergènes alimentaires (**Reboldi et al, 2016**).

1.1.3 Classification des allergies alimentaires

Les allergies alimentaires peuvent être causées par les IgE, non médiées par les IgE ou une combinaison de réactions médiées par les IgE et non IgE. Cliniquement, il peut s'agir de la peau, du tractus gastro-intestinal, des systèmes respiratoire, respiratoire et/ou cardiovasculaire. Les prévalences d'allergies alimentaires varient de 1 % à 10 % chez les

enfants de moins de 5 ans, diminuant chez les adultes en raison de certaines allergies périmées. Les sensibilités alimentaires peuvent être divisées en deux catégories principales : l'allergie, qui est une réponse systémique anormale à une intolérance alimentaire, c'est-à-dire toute forme de sensibilisation alimentaire qui ne fait pas intervenir de mécanisme immunologique. Les différences ont des implications pratiques tant du point de vue clinique que du point de vue réglementaire (**Taylor,2000**).

La réaction alimentaire indésirable est un terme général qui représente une réaction clinique anormale liée à l'ingestion d'aliments et est classée en intolérance alimentaire ou en allergie alimentaire en fonction du mécanisme physiopathologique de la réaction. L'intolérance alimentaire fait référence à une réponse physiologique non désiré à la nourriture et peut être due à des causes inhérentes propriétés alimentaires (par exemple, contaminants toxiques, principes pharmacologiquement actifs) ou des caractéristiques de l'hôte (troubles métaboliques, réactions idiosyncrasiques, troubles psychiatriques, etc.), ils peuvent ne pas être reproductibles et souvent reproductibles. Dépendant de la dose. Considérée comme une intolérance alimentaire L'allergie alimentaire fait référence à une réponse immunitaire anormale à la nourriture qui se produit chez un hôte sensible. Ces réponses sont reproductibles à toutes les prises alimentaires et souvent non dose-dépendantes.

Sur la base des mécanismes immunologiques impliqués, les allergies alimentaires peuvent être classées comme :

- a) Médinée par les IgE médinée par les anticorps, il appartient à l'immunoglobuline E (IgE) et est la réaction allergique alimentaire la plus caractéristique.
- b) À médiation cellulaire, lorsqu'un composant cellulaire du système immunitaire est responsable de l'allergie alimentaire et affecte principalement le tractus gastro-intestinal.
- c) Réactions mixtes à médiation cellulaire méditées par les IgE lorsque les IgE et les cellules immunitaires sont impliquées dans la réaction (**Sicherer et Sampson, 2009**).

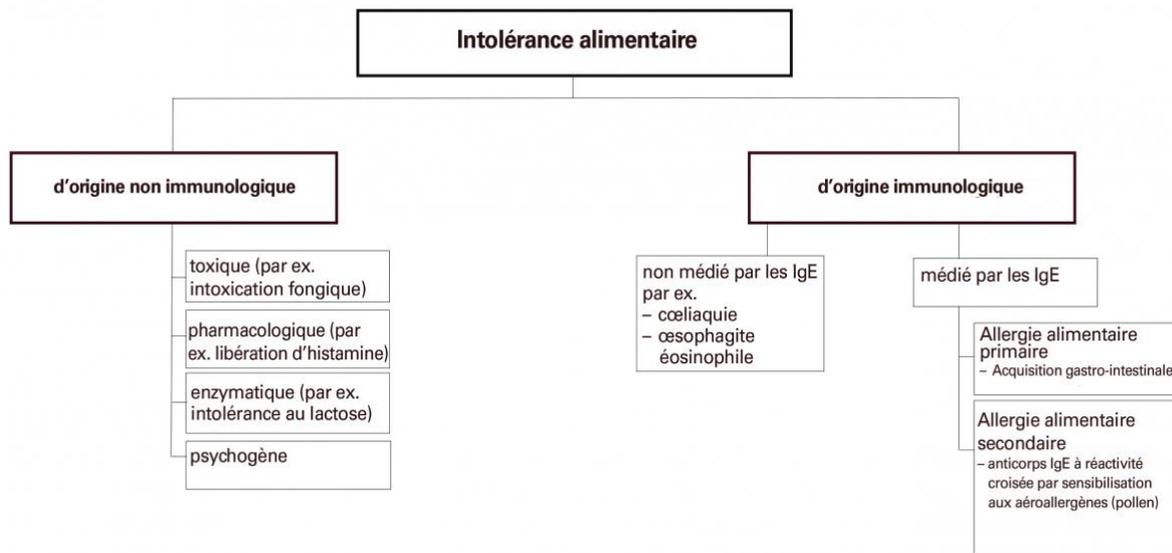


Figure1 : Classification Des Allergies Alimentaire (Cianferoni et Spergel, 2009).

1.1.4. Impact des allergies alimentaire sur la qualité de vie

Il est évident que l'allergie alimentaire a un effet psychosocial important sur les gens notamment les enfants et leurs membres de familles. De coup les personnes allergiques aux aliments doivent rester vigilant pour éviter la nourriture allergène. Cette hyper vigilance affecte négativement leur qualité de vie (Liane et Bacal, 2013). Dans l'ensemble, il y a un super besoin d'éduquer les gens allergiques et de les faire savoir gérer leurs allergies alimentaires. Une étude a montré que les familles des enfants allergiques aux arachides ont beaucoup plus de perturbations dans leur activités quotidienne et dans leur interactions sociales-sociales par rapport à d'autres familles (Primeau et al, 2000).

1.1.5. Méthodes de diagnostique

Les antécédents cliniques et l'examen physique constituent l'approche de première ligne de diagnostic d'une allergie alimentaire, qui doit aborder des facteurs spécifiques tels que la dermite atopique et les antécédents familiaux. Lorsqu'une allergie alimentaire à médiation IgE est suspectée de doser le taux d'IgE sérique ; des niveaux élevés des IgE peuvent suggérer une allergie alimentaire. Les tests cutanés sont recommandés pour identifier l'aliment responsable. Ces tests provoquent une dégranulation des mastocytes médiée par les allergènes alimentaires dans la peau. Une lancette d'allergène est utilisée pour perforer la barrière épithéliale afin de provoquer une réaction cutanée ; si une papule se forme, elle est

mesurée après 15 minutes. Des seuils de taille de papule indiquant une forte probabilité d'allergie alimentaire ont été définis pour certains aliments courants. Cependant, le taux des IgE et les tests cutanés indiquent uniquement une sensibilisation aux allergènes, ce qui nécessaire, mais pas suffisant pour une allergie alimentaire.

Il existe une méthode de diagnostic définitive pour l'allergie alimentaire est l'*Oral Food Challenge* OFC, dans laquelle des doses croissantes d'un allergène alimentaire potentiel sont ingérées sur des intervalles de temps fixes, jusqu' à ce qu'une réaction allergique soit observée. Cependant, cette méthode est utilisée dans des centres spécialisées, équipées et dotées en personnel pour surveiller s'il ya un risque d'une réaction anaphylactique(Wong Yu et al, 2016).

1.1.6. Traitementset prévention

Actuellement, il n'existe pas de traitement définitif pour l'allergie alimentaire ; l'évitement des allergènes alimentaires et le traitement des réactions systémiques restent le schéma thérapeutique le plus utilisé.

Des produits pharmaceutiques tels que l'adrénaline et les antihistaminiques sont utilisées pour traiter les symptômes d'allergie alimentaire. Cependant ces molécules actuellement disponibles ne contrôlent que les symptômes de l'allergie et ne traitent pas le trouble immunitaire sous-jacent.L'immunothérapie pour désensibiliser les individus aux allergènes alimentaires potentiels représente un progrès considérable dans le traitement des allergies alimentaires. En particulier, la compréhension des marqueurs et des mécanismes qui différencient la tolérance, la désensibilisation et les insuffisances durables seront utiles pour la surveillance de l'immunothérapie et fournir la prochaine génération de traitements. (Yu et al, 2016).

La prévention primaire vise à est contrôler l'introduction précoce des aliments a potentiel allergique (tel que les cacahouètes) pour les nourrissons(Sichere et al, 2018). Enrevanche la prévention secondaire vise à prévenir l'expression clinique de l'allergie chez les personnes qui sont soit sensibilisées aux allergènes, soit qui manifestent déjà d'autres troubles allergique tels que la dermite atopique.(Ralf et Heine, 2018)

1.2 Polymorphisme nucléotidique simple

1.2.1. Généralités

Un polymorphisme nucléotidique simple (SNP, *single nucleotide polymorphism*) est défini comme une variation unique dans une séquence nucléotidique d'ADN qui peut ou non affecter les caractéristiques phénotypiques d'un organisme.

Selon **Brookes (1999)**, la définition de travail d'un SNP est qu'il s'agit d'un changement de base unique dans l'ADN génomique où des substitutions de séquence sont présentes chez les individus normaux et la fréquence de l'allèle le moins commun est d'au moins 1 %. Cependant, le terme SNP est souvent utilisé dans un contexte plus général pour inclure toute variante de base unique observée et éventuellement associée à une maladie (**Cavallo et Martin, 2005**).

Dans la population humaine, plus de 99 % de l'ADN est partagé, tandis que les SNP représentent 90 % de la différence de 1 % (**Fairweather-Tait et al., 2007**). Les SNP influencent les réponses individuelles à la maladie (**Wellcome Trust Case Control, 2007**), aux médicaments (**Ma et Lu, 2011**) et aux facteurs environnementaux (**Bresciani et al., 2013**). En moyenne, l'ADN humain est composé d'un SNP de 300 bases (**Nelson et al., 2004**). Cela signifie que pour l'ensemble du génome (3 100 millions de bases), il y aura environ 10 millions de SNP. Plus de 60 000 SNP sont situés dans les régions codantes du gène considéré (**Sachidanandam et al., 2001**).

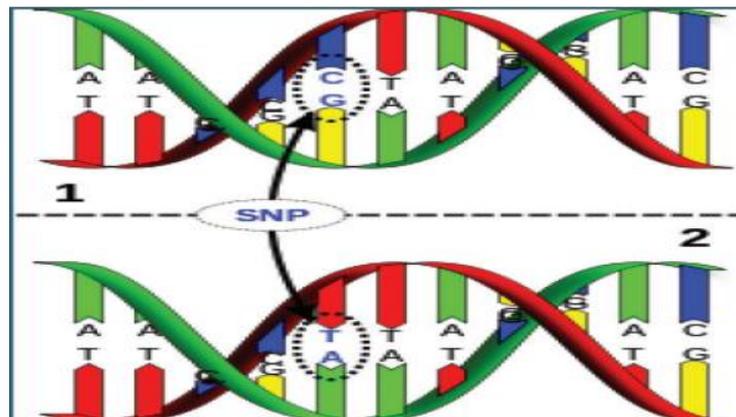


Figure 2 : Single nucleotide polymorphism (SNP). (http://en.wikipedia.org/wiki/Single-nucleotide_polymorphism).

1.2.2. L'effet de la localisation des polymorphismes nucléotidique

L'emplacement des SNP affecte les gènes et d'autres produits. Les SNP dans les gènes peuvent modifier la structure des protéines. Les SNP dans les régions régulatrices extra-géniques affectent quand et comment les gènes sont activés, affectant ainsi la quantité de protéines produites. Ils affectent également l'épissage des gènes, la liaison des facteurs de transcription ou les séquences d'ARN non codantes. Les SNP qui ne sont pas proches des gènes peuvent être utilisés comme marqueurs génétiques pour cartographier les gènes pathogènes (Figure 2.1).

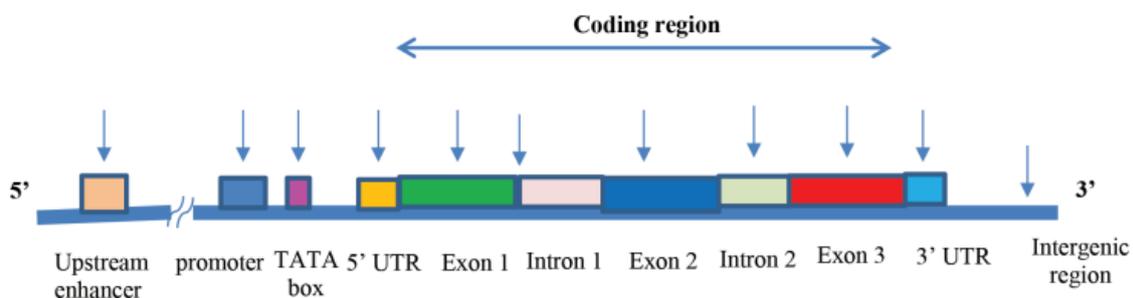


Figure 2.1 : SNP localisation.

1.2.3. Types de polymorphisme nucléotidique

Un SNP peut être classé selon sa localisation dans le génome : dans la région codante du gène, dans la région non codante du gène ou dans la région intergénique. La plupart des SNP sont situés dans des régions non codantes (Syvanen, 2001). Cependant, les SNP présents dans les régions codantes sont importants, car ils peuvent affecter diverses activités biologiques et moléculaires importantes, telles que la stabilité (Casadio et al., 2011), les niveaux d'expression (Heinzen et al, 2012) et la fonction des protéines (Chaseman et Adams, 2001). Cependant, des études récentes (Ward et Kellis, 2012, Visel et al., 2010, Prado-Montes de Oca et al, 2009) ont montré que les SNP dans les régions non codantes peuvent encore affecter d'autres activités telles que l'épissage des gènes, l'expression des gènes et la séquence d'ARN non codant.

Sur les 10 millions de SNP estimés, seuls 3 % à 5 % sont situés dans des régions codantes, la moitié d'entre eux entraînent des substitutions d'acides aminés, appelées SNP non synonymes (nsSNP) (Cargill et al., 1999). Les SNP dans les régions codantes sont classés en fonction de leur rôle. Les SNP synonymes (sSNP) sont des SNP qui ne modifient pas la séquence d'acides

aminés d'une protéine, tandis que les SNP non synonymes (nsSNP) modifient les protéines. Il a été suggéré que 20 % des nsSNP pourraient endommager la protéine (Sunyaev et al., 2001). Les nsSNP peuvent être trompeurs ou absurdes. Les nsSNP faux-sens entraînent différents acides aminés, tandis que les changements de nsSNP non-sens entraînent des codons d'arrêt prématurés. Les nsSNP faux-sens peuvent être des modifications conservatrices ou non conservatrices. Des modifications non conservatrices entraînent des résidus distincts avec des propriétés physicochimiques significativement différentes. Par exemple, si une séquence d'ADN code pour la lysine, une modification conservatrice peut entraîner l'arginine. Cependant, sous des changements non conservateurs, des protéines avec des chaînes latérales neutres en charge peuvent être générées.

1.2.4. Études d'association à l'échelle du génome (GWAS)

GWAS examine des centaines de milliers de SNP communs à l'aide d'un équipement de génotypage à haut débit et compare ces variantes génétiques communs dans un grand nombre de cas affectés (patients) avec ceux de témoins non affectés (non-patients). Des comparaisons sont effectuées pour déterminer la présence de variantes pathogènes associées à la maladie (Figure 3) (Lander, 1996). Dans la plupart des régions chromosomiques, il existe de fortes associations entre les SNP, donc seuls quelques-uns des SNP de chaque région sont sélectionnés pour le séquençage afin de prédire les allèles des SNP restants dans cette région. La sélection des meilleurs SNP d'étiquetage nécessite une cartographie précise des modèles LD entre les SNP de différentes populations ancestrales. Carte LD précise requise pour faciliter les études d'association génétique et stimuler le développement de cartes d'haplotypes humains (Eberle et al, 2007).

GWAS identifie les gènes qui peuvent mener à un risque de créer une maladie. Les données dérivées de GWAS informent sur l'étiologie d'une certaine maladie, les cibles thérapeutiques et le fonctionnement des gènes (Manolio et Collins, 2009).

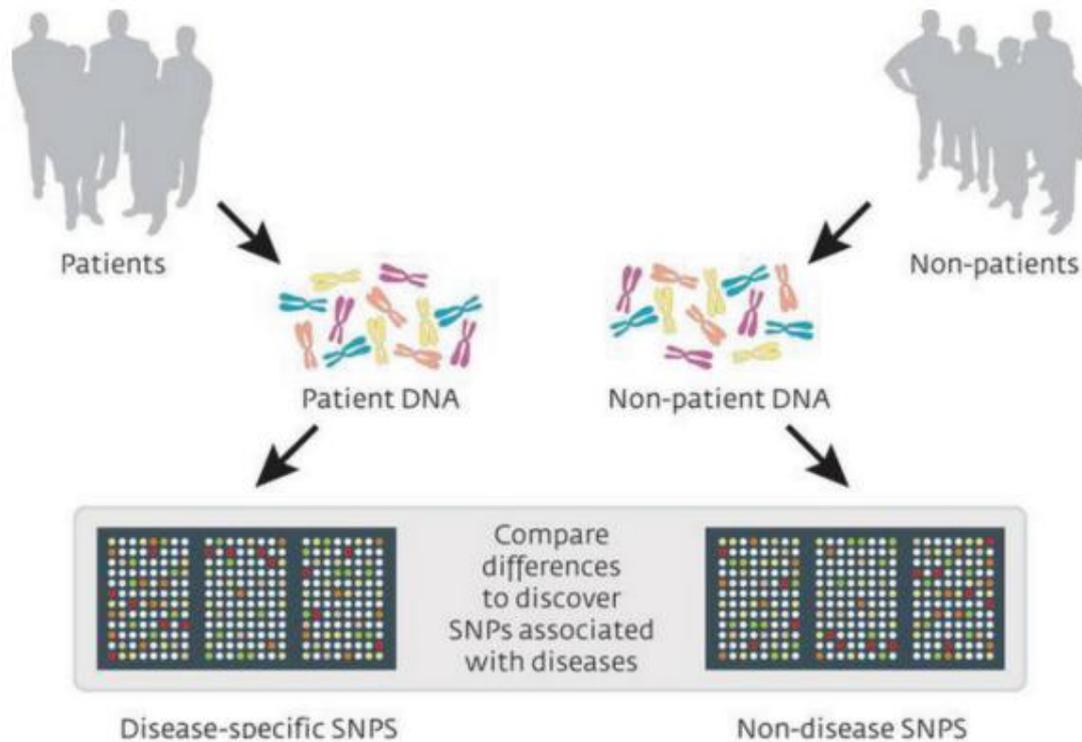


Figure 2.2 : Études d'association à l'échelle du génome (GWAS) (<http://mmg-233-2014-genetics-genomics.wikia.com/wiki/File:GWAS.jpg>).

1.2.5. Application de polymorphisme nucléotidique

1.2.5.1. Les SNP comme marqueurs biologiques de maladies humaines

La plupart des SNP ne sont pas responsables des états pathologiques. Ils sont utilisés comme biomarqueurs pour localiser la maladie sur la carte du génome humain. Les SNP sont présents en moyenne toutes les 200 paires de bases dans le génome humain (Venter *et al*, 2001). Les SNP communs (une gamme de fréquences d'allèles mineurs allant de 5% à > 20%) peuvent expliquer un sous-ensemble de maladies humaines courantes. La plupart des SNP ne se trouvent pas dans les régions codantes des gènes, ni même à l'intérieur des gènes (Weissman D *et al*, 2001). Les SNP non synonymes suspectés de provoquer une maladie chez l'homme ne représentent pas tous les SNP qui provoquent une maladie ou une susceptibilité à la maladie. D'autres SNP fonctionnels associés à la maladie humaine ou à la susceptibilité aux

maladies comprennent ceux situés dans les promoteurs(Lin et al, 2003), introns(Tokuhiro S et al, 2003), sites d'épissage (Betticheret al, 1995), et les régions intragéniques(Gordon D et al, 2003).Même des SNP synonymes ont été impliqués dans des conséquences fonctionnelles passant par un mécanisme inconnu (Duan J et al, 2003).

1.2.5.2. Polymorphisme nucléotidique et le développement de médicaments

Des variantes génétiques codant pour des enzymes métabolisant des médicaments ou des cibles médicamenteuses ont été étudiés en relation avec des réponses médicamenteuses individuelles. Les SNP sont des marqueurs moléculaires populaires dans de telles études pharmaco-génomiques. L'utilisation des SNP pour étudier la génétique de la réponse aux médicaments aidera à la création de médicaments ou de médicaments personnalisés qui conviennent le mieux à un individu et peuvent être déterminés en analysant le profil SNP d'un patient avant le traitement. Les SNP peuvent être impliqués dans l'absorption et la clairance d'agents thérapeutiques. L'association de différents SNP avec une variété de maladies humaines telles que le cancer, les maladies auto-immunes, les maladies neuropsychiatriques et d'autres maladies infectieuses peut être utilisée comme cible pour la pharmacothérapie (Fareed et Afzal, 2013)

1.2.6. L'avantage de polymorphisme nucléotidique

Les SNP sont utilisés dans un large éventail d'études, y compris l'estimation de la prédisposition à une maladie (Voronko et al, 2008), la prédiction de traits génétiques spécifiques (Yip et Lange, 2011), agissant comme un marqueur biologique pour des maladies complexes (Stefansson et al, 2009), prédire l'efficacité des médicaments (Giacomini et al, 2007) et suivre la migration ancestrale (Underhill et Kivisild, 2007).

1.2.7. Rôle des polymorphismes nucléotidiques dans le traitement des allergies alimentaire

Les polymorphismes mononucléotidiques (SNP) peuvent jouer un rôle direct ou indirect dans l'expression phénotypique qui peut jouer un rôle important dans la pathogenèse des allergies alimentaires (Brown et al, 2012).Les études d'association et d'association de l'atopie et de l'allergie ont identifié un certain nombre des gènes qui sont importants dans la spécificité et la reconnaissance des allergènes et représentent donc de bons candidats pour les facteurs de susceptibilité. Les marqueurs génétiques SNP. Plusieurs études ont fourni des preuves solides

que la susceptibilité génétique de l'hôte et les facteurs environnementaux déterminent la régulation complexe de l'allergie alimentaire médiée par les IgE(Cameron et al, 2009). Cependant, seules quelques études ont exploré la relation entre les SNP et le développement et le pronostic de l'allergie alimentaire(Smith et al, 2008)

1.2.8. Analyse *insilico*

3.1.Analyse *in silico* des polymorphisme nucléotidique

C'est Défini comme une discipline scientifique couvrant tous les aspects de l'acquisition, du traitement, du stockage, de la distribution, de l'analyse et de l'interprétation de l'information biologique(Benton, 1996). Ce domaine interdisciplinaire tente de relier les théories et les applications des mathématiques, de la biologie et de l'informatique pour comprendre la signification et la relation entre les données biologiques quantitatives et qualitatives aux niveaux cellulaire, biochimique et moléculaire.

1.2.9. Implication de gène STAT3 dans les allergies alimentaires :

Des études cliniques ont montré que l'incidence des IgE spécifiques aux allergènes alimentaires est 10 fois supérieure à l'incidence des allergies alimentaires , ce qui suggère un niveau supplémentaire de régulation de la tolérance au-delà de la simple prévention de l'amorçage immunologique vers Th2 et IgE. De plus, chez les patients porteurs de mutations Stat3 conduisant au syndrome d'hyper-IgE, la réactivité anaphylactique aux allergènes alimentaires est en fait diminuée.

1.2.9. Implication de gène FOXP3 dans les allergies alimentaires :

L'allergie alimentaire médiée par les IgE fait face à un lourd fardeau économique et à des événements potentiellement mortels, en particulier pour les patients souffrant d'anaphylaxie d'origine alimentaire. L'anaphylaxie d'origine alimentaire est une réaction allergique

potentiellement mortelle et la prévalence n'a cessé d'augmenter. L'anaphylaxie induite par l'exercice dépendant du blé est un type rare d'anaphylaxie, causé par la combinaison du blé et de facteurs d'augmentation supplémentaires, comme l'exercice, l'alcool ou l'acide acétylsalicylique. CD4 + CD25 + Foxp3 , un facteur suppresseur critique dans le maintien de l'homéostasie immunitaire, tandis que les cellules T auxiliaires productrices d'IL-17 ont l'effet inverse. Nous supposons que les facteurs de transcription (FOXP3, ROR γ t, PU.1) ainsi que les cytokines correspondantes (IL-6, IL-9, IL-10, IL-17) sont essentiels dans les cas d'anaphylaxie.

Chapitre II : Matériel et méthodes

Matériel et méthodes

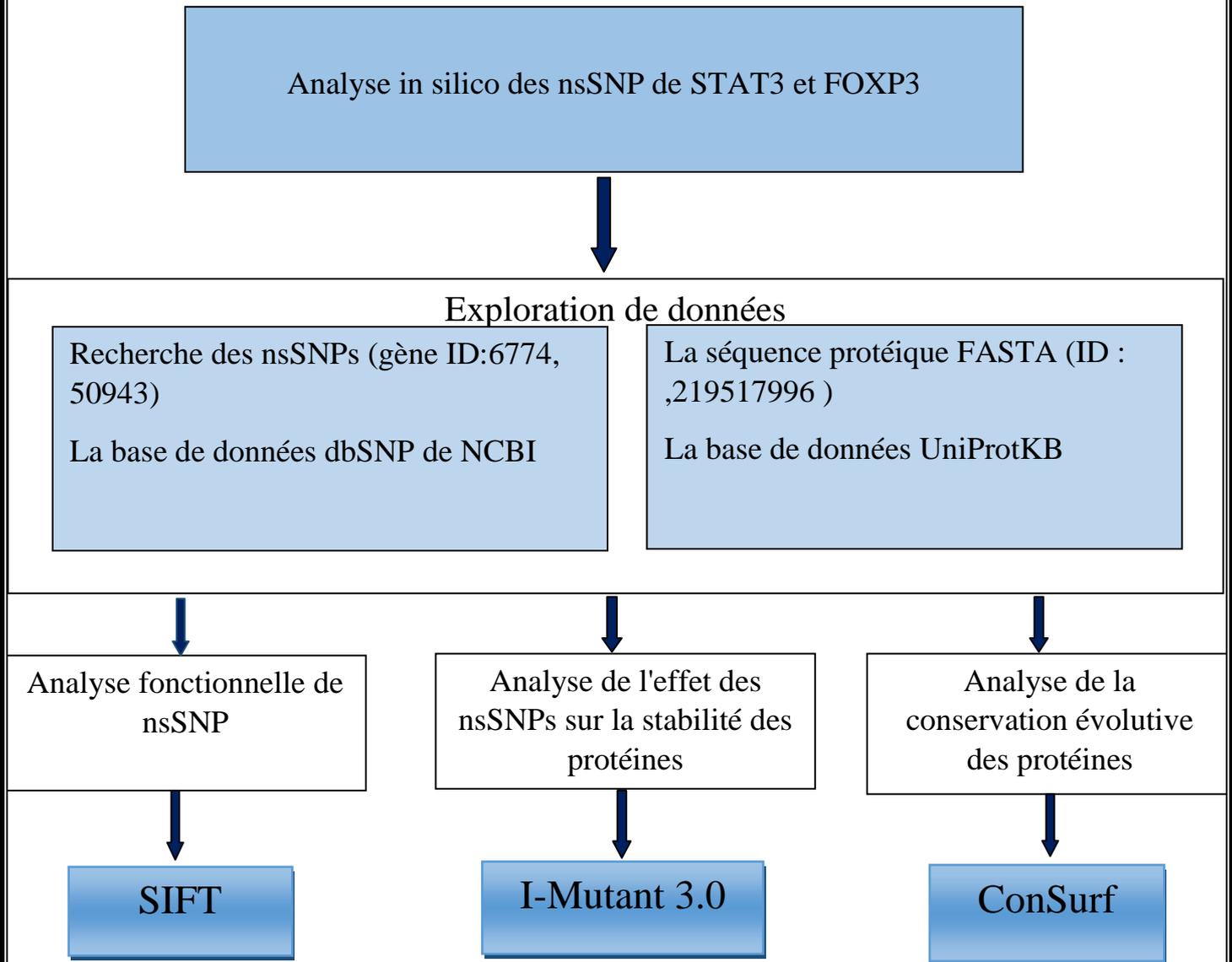


Figure :Study design.

2.1. Recherche des nsSNPs et des séquences protéiques

Les données génétiques humaines sont collectées dans Entrez Gene sur le National Centre de Biologie Site Web d'information (NCBI).

2.2. Récupération des données des nsSNP

La distribution nsSNP du gène FOXP3 et STAT3 a été recueillie de la base de données dbSNP (<http://www.ncbi.nlm.nih.gov/projets/SNP/>) (Sherry et al, 2001), la base de données ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar>) (Landrum et al, 2018) du National Center for Biotechnology Information (NCBI), le Gène Humain Base de Données sur les Mutations (HGMD, <http://www.hgmd.cf.ac.uk/ac/index.php>) (Stenson et al, 2017), et la base de données DisGeNET (<http://www.disgenet.org>) (Pinero et al, 2017) en utilisant les limites de "Homo sapiens" et "codage non synonyme". La littérature antérieure a également été revue. Les séquences d'acides aminés et d'ADN, ID SNP, acides aminés de type sauvage, positions des acides aminés, acides aminés faux-sens, fréquence des allèles mineurs (MAF), et d'autres informations ont été recueillies

The screenshot displays the dbSNP search results for the variant rs1053004. The interface is from the National Library of Medicine (NIH) National Center for Biotechnology Information. The search results show 1 item out of 26861. The variant details are as follows:

Variant type:	SNV
Alleles:	G>A [Show Flanks]
Chromosome:	17:42314074 (GRCh38) 17:40466092 (GRCh37)
Canonical SPDI:	NC_000017.11:42314073:G:A
Gene:	STAT3 (VarView)
Functional Consequence:	genic_downstream_transcript_variant,3_prime_UTR_variant
Clinical significance:	benign
Validated:	by frequency, by alfa, by cluster
MAF:	G=0.38237/100746 (ALFA) G=0.240113/85 (SGDP_PRJ)

Figure: Base de données dbSNP.

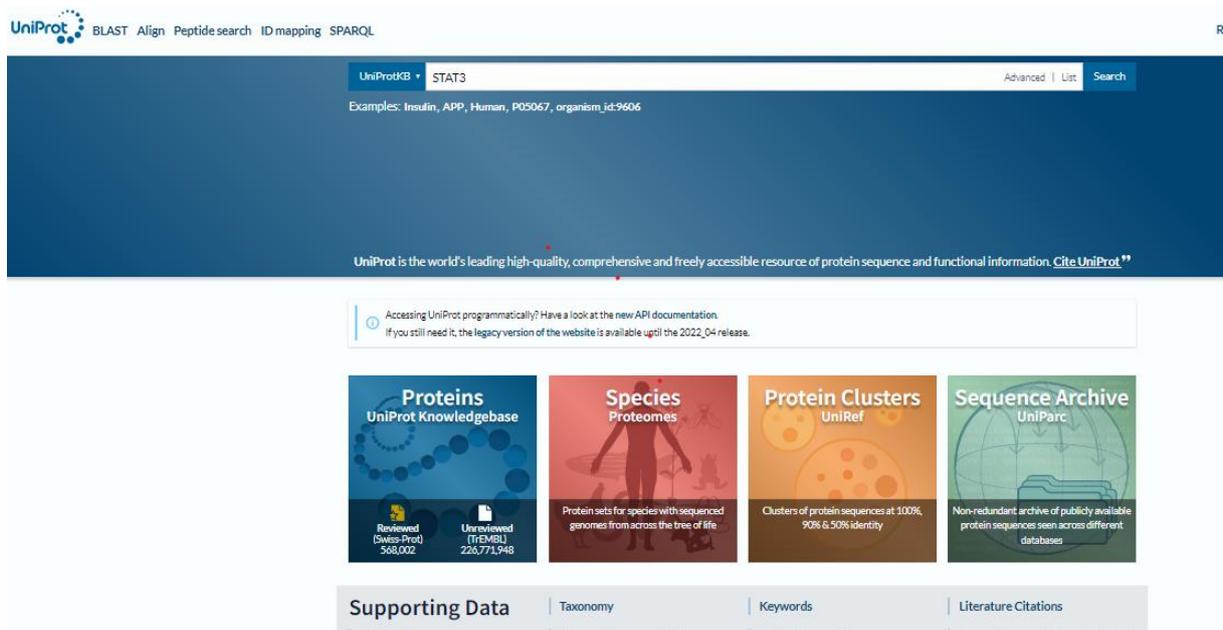


Figure 2.3 : Base de données UniProt.

2.3. Prédiction des nsSNP délétères par différents outils bioinformatique

Dans cette étude, huit outils Web ont été utilisés pour prédire l'impact fonctionnel et la pathogénicité des nsSNP. Tous les outils sont utilisés conformément à leurs paramètres par défaut, sauf indication contraire.

SIFT

SIFT (Trier les intolérants des tolérants, <https://sift.bii.a-star.edu.sg/>) prédit si une substitution d'acides aminés cause des effets délétères en se basent sur l'homologie de séquence et des propriétés physiques des acides aminés. Une variante faux-sens est prédite comme délétère lorsque le score SIFT < 0,05, tandis qu'un score $\geq 0,05$ indique qu'une variante est bénigne(Sim et al ,2012).

PROVEAN

PROVEAN (Protein Variation Effects Analyzer, <http://provean.jcvi.org/index.php>) est un outil de prédiction Effets fonctionnels des substitutions, des insertions et des suppressions d'acides aminés, qui introduisent des scores d'alignement delta pour les séquences d'interrogation des protéines afin de mesurer l'impact des variantes. Les scores delta élevés étaient considérés comme des variantes à effets neutres, tandis que les scores delta faibles étaient considérés comme des variantes d'acides aminés ayant des effets négatifs sur la fonction des protéines. Pour fournir des prédictions binaires, le seuil du score PROVEAN est fixé à 2,5 pour obtenir une précision équilibrée élevée(Choi et al, 2010).

PolyPhen2

PolyPhen2 (polymorphismPhenotyping v2, [http:// genetics.bwh.harvard.edu/pph2/](http://genetics.bwh.harvard.edu/pph2/)), qui utilise les ensembles de données HumVar et HumDiv, basés sur un classificateur Bayes naïf formé avec l'apprentissage automatique supervisé. Un algorithme glouton itératif a été utilisé pour sélectionner des caractéristiques prédictives, dont huit caractéristiques basées sur la séquence et trois caractéristiques basées sur la structure, par lesquelles différentes mutations ont été classées comme « éventuellement délétères », « éventuellement délétères » ou « bénignes » (Adzhubei et al,2010).

PANTHERE-PSEP

PANTHER-PSEP (PANTHER-position-specific evolution preservation, <http://pantherdb.org/tools/csnpscoreform.jsp>) utilise une métrique pertinente mais différente de la « préservation évolutive » : les séquences possibles des protéines ancestrales aux nœuds d'un arbre phylogénétique sont reconstruites à partir de protéines homologues. D'état actuel de chaque acide aminé, son histoire peut être retracé pour calculer la durée que l'acide aminé a été conservé dans ses ancêtres. Le score PSEP était classé comme « probablement dommageable » (la préservation temps > 450 ma), "possiblement dommageable" (200 ma < temps de conservation < 450 ma) et "probablement bénin" (temps de conservation < 200 ma).(Tang et Thomas, 2016).

SNAP2

Dépistage des polymorphismes non acceptables (SNAP2) (disponible sur <https://roslab.org/services/snap>) est un neural outil basé sur un réseau qui classe les substitutions d'acides aminés en efficace et neutre sur la fonction des protéines en prenant un diversité des séquences et des caractéristiques différentes dans considération. SNAP2 fournit une liste de toutes les substitutions possibles dans la séquence protéique avec un score, un effet fonctionnel (neutre ou effet) et la précision attendue pour tout remplacement. La précision attendue indique le niveau de confiance pour chaque prédiction. Les résultats sont également affichés sous forme de carte thermique(Emadi et al, 2020).

2.4. L'identification de nsSNP sur les domaines des gènes :

La base de données **InterPro** (<http://www.ebi.ac.uk/interpro/>) regroupe des modèles prédictifs représentant des domaines, des familles et des sites fonctionnels de protéines provenant de plusieurs bases de données de sources diverses (Figure)(Hunter et al, 2009).

2.5. Analyse de l'effet des nsSNP sur la stabilité des protéines :

PMut

PMut (<http://mmb.irbbarcelona.org/PMut/>) a été formé et testé par la base de données créée manuellement SwissVar (version d'octobre 2016), qui comprend 27 203 mutations nocives et 38 078 mutations bénignes pour 12 141 protéines. Les scores de prédiction de PMut vont de 0 à 1 et la valeur seuil est fixée à 0,5 (neutre, 0 à 0,5 ; pathologique, 0,5 à 1) (**Lopez-Ferrando et al, 2017**).

I-mutant 3.0

I-Mutant 3.0 (<http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi>) est entraîné et testé sur l'ensemble de données $\Delta\Delta G$ Mut obtenu auprès de ProTherm. Le prédicteur peut estimer le changement stable au cours d'une mutation ponctuelle, mesuré par la valeur $\Delta\Delta G$ (kcal/mol), basée sur la structure protéique ou la séquence protéique. Une valeur $\Delta\Delta G$ inférieure à "0" indique que le variant réduit la stabilité de la protéine. A l'inverse, des valeurs de $\Delta\Delta G$ supérieures à 0 indiquent que le variant améliore la stabilité des protéines (**Capriotti et al, 2005**).

2.6. Analyse de la conservation évolutive des protéines

ConSurf

(consurf.tau.ac.il/) est un serveur web que nous avons utilisé pour l'identification et le classement des mutations situé dans la région conservée de les protéines STAT3 et FOXP3 (Figure) (**Halder et al, 2022**). Cet outil utilise la séquence protéique FASTA comme entrée et classé les acides aminés en différents domaines de protéines ayant une importance fonctionnelle.

2.7. Modélisation des emplacements nsSNP sur la structure des protéines

I-Mutant (version 2.0) et FOLD-X I-Mutant (<http://folding.uib.es/i-mutant/i-mutant2.0.html>) est un outil basé sur un réseau de neurones pour l'analyse de routine de la stabilité et des altérations des protéines en prenant en compte les mutations à site unique. La séquence FASTA de la protéine extraite d'UniProt est utilisée comme entrée pour prédire l'effet mutationnel sur stabilité des protéines. I-Mutant fournit également les scores d'altérations d'énergie libre, calculés avec le FOLD-X serveur web basé sur l'énergie (**Dabhi, Bet al, 2014**).

2.8. Base de données sur les polymorphismes nucléotidiques simple (SNP) et gène Ontologie (GO) (SNPs&GO)

(Disponible sur snps.biofold.org/snps-and-go/snps-and-go.html) est un prédicteur unique basé sur SVM et intégré à GO. Il prédit si un acide aminé la substitution est associé à la maladie ou n'utilise pas de GO fonctionnel termes, structure protéique 3D et informations évolutives sur la séquence protéique. La substitution d'acides aminés est associée à la maladie si le score de probabilité est supérieur à 0,5 (**Emadi et al , 2020**).

Résultats et interprétation :

3. Résultats et interprétation :

CONFIDENTIEL

IV. Discussions

CONFIDENTIEL

V. Conclusion et perspectives

L'étude actuelle a examiné les effets des SNPs fonctionnels associés aux gènes STAT3 et FOXP3 humain par des méthodes computationnelles en raison de la relation de ces gènes avec l'allergie alimentaire. La présente étude suggère que la fonction des protéines STAT3 et FOXP3 pourraient être perturbées par divers SNP non synonyme. Ces nsSNPs peuvent être considérés comme des cibles principales dans la cause de maladies telles que l'allergie alimentaire. Nous supposons que les résultats de la présente étude constituent une base pour de futures études expérimentales et in silico. Nos résultats serviront comme guide important pour d'autres analyses structurelle permettant de prédire les interactions protéine-protéine ou protéine-peptide ainsi que les modifications post-traductionnelles

VI. Bibliographie

- Allen KJ, Hill DJ, Heine RG. Food allergy in childhood. *Med J Aust.* 2006;**185**(7):394-400.
- Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res.* 2005;**33**(Web Server issue):W306–10.
- Dabhi, B.; Mistry, K. N. In Silico Analysis of Single Nucleotide Polymorphism (SNP) in Human TNF- α Gene. *Meta Gene***2014**, 2, 586–595.
- Liu X, Zhang S, Tsai H-J, et al. Genetic and environmental contributions to allergen sensitization in a Chinese twin study. *Clin ExpAllergy.* 2009;**39**(7):991-998.
- Tsai H-J, Kumar R, Pongracic J, et al. Familial aggregation of food allergy and sensitization to food allergens: a family-based study. *ClinExp Allergy.* 2009;**39**(1):101-109.
- Taylor, S.L., Emerging problems with food allergens. *Food Nutrition and Agriculture*, (26), 2000, pp.14-23.
- Urisu, A., Ebisawa, M., Mukoyama, T., Morikawa, A. and Kondo, N., Japanese guideline for food allergy. *Allergology International*, 60(2), 2011, pp.221-236.
- Eriksson, N.E., Moller, C., Werner, S., Magnusson, J., Bengtsson, U. and Zolubas, M., Self-reported food hypersensitivity in Sweden, Denmark, Estonia, Lithuania and Russia. *Journal of investigational allergology and clinical immunology*, 14(1), 2004, pp.70-79.
- Eberle MA, Ng PC, Kuhn K, et al. Power to detect risk alleles using genome-wide tag SNP panels. *PLoS Genetics.* 2007;**3**:1827-1837.
- Sampson, H.A. Food allergy. Part 1: immuno-pathogenesis and clinical disorders. *Journal of Allergy and Clinical Immunology*, 103(5), 1999, pp.717-728.
- Manolio TA, Collins FS. The HapMap and genome-wide association studies in diagnosis and therapy. *Annual Review of Medicine.* 2009;**60**:443-456.
- Muraro, A., Werfel, T., Hoffmann-Sommergruber, K., Roberts, G., Beyer, K., Bindslev-Jensen, C., Cardona, V., Dubois, A., Dutoit, G., Eigenmann, P. and Fernandez Rivas, M. EAACI food allergy and anaphylaxis guidelines: diagnosis and management of food allergy. *Allergy*, 69(8), 2014, pp.1008-1025.
- Radauer, C., Bublin, M., Wagner, S., Mari, A. and Breiteneder, H. Allergens are distributed into few protein families and possess a restricted number of biochemical functions. *Journal of Allergy and Clinical Immunology*, 121(4),2008, pp.847-852.

Hoffmann-Sommergruber, K. and Mills, E.C. Food allergen protein families and their structural characteristics and application in component-resolved diagnosis: new data from the EuroPrevall project. *Analytical and bioanalytical chemistry*, 395(1), 2009, pp.25-35.

Hoffmann-Sommergruber, K. Proteomics and its impact on food allergy diagnosis. *EuPA Open Proteomics*, (12), 2016, pp.10 -12.

Sicherer SH, Sampson HA. Food allergy: epidemiology, pathogenesis, diagnosis, and treatment. *J Allergy Clin Immunol* 2014;133:291-308.

Hussain M, Borcard L, Walsh KP, Pena Rodriguez M, Mueller C, Kim BS, et al. Basophil-derived IL-4 promotes epicutaneous antigen sensitization concomitant with the development of food allergy. *J Allergy Clin Immunol* 2018;141:223-34.

Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* 2018;46(D1):D1062–d7

Galand C, Leyva-Castillo JM, Yoon J, Han A, Lee MS, McKenzie AN, et al. IL33 promotes food anaphylaxis in epicutaneously sensitized mice by targeting mast cells. *J Allergy Clin Immunol* 2016;138:1356-66.

Oyoshi MK, Oettgen HC, Chatila TA, Geha RS, Bryce PJ. Food allergy: insights into etiology, prevention, and treatment provided by murine models. *J Allergy Clin Immunol* 2014;133:309-17.

Barcik W, Untersmayr E, O'Pali-Schöoll I, O'Mahony L, Frei R. Influence of microbiome and diet on immune responses in food allergy models. *Drug Discovery Today: Disease Models* 2016;17:71-80.

Kelleher MM, Dunn-Galvin A, Gray C, Murray DM, Kiely M, Kenny L, et al. Skin barrier impairment at birth predicts food allergy at 2 years of age. *J Allergy Clin Immunol* 2016;137:1111-6, e1-8.

Reboldi A, Arnon TI, Rodda LB, Atakilit A, Sheppard D, Cyster JG. IgA production requires B cell interaction with subepithelial dendritic cells in Peyer's patches. *Science* 2016;352:aaf4822.

Sicherer, S.H. and Sampson, H.A. Food allergy: recent advances in pathophysiology and treatment. *Annual review of medicine*, (60), 2009, pp.261-277.

Stenson PD, Mort M, Ball EV, Evans K, Hayden M, Heywood S, et al. The human gene mutation database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next generation sequencing studies. *Hum Genet.* 2017;136(6):665–77.

- Pinero J, Bravo A, Queralt-Rosinach N, Gutierrez-Sacristan A, Deu-Pons J, Centeno E, et al. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res.* 2017;45(D1):D833–d9.
- Cianferoni, A. and Spergel, J.M. Food allergy: review, classification and diagnosis. *Allergology International*, 58(4), 2009, pp.457-466.
- Bacal, Liane R. 2013. « The Impact of Food Allergies on Quality of Life ». *Pediatric Annals* 42 (7): 141- 45. <https://doi.org/10.3928/00904481-20130619-12>.
- Brown, Paula, Bindukumar Nair, Supriya D. Mahajan, Donald E. Sykes, Gary Rich, Jessica L. Reynolds, Ravikumar Aalinkeel, John Wheeler, et Stanley A. Schwartz. 2012. « Single Nucleotide Polymorphisms (SNPs) in Key Cytokines May Modulate Food Allergy Phenotypes ». *European Food Research and Technology = Zeitschrift Fur Lebensmittel-Untersuchung Und -Forschung. A* 235 (5): 971- 80.
- Heine, Ralf G. 2018. « Food Allergy Prevention and Treatment by Targeted Nutrition ». *Annals of Nutrition & Metabolism* 72 Suppl 3: 33- 45. <https://doi.org/10.1159/000487380>.
- Newman, Byron Y. 2011. « A Bit about Allergies ». *Optometry (St. Louis, Mo.)* 82 (8): 453- 54. <https://doi.org/10.1016/j.optm.2011.07.001>.
- Primeau, M. N., R. Kagan, L. Joseph, H. Lim, C. Dufresne, C. Duffy, D. Prhcal, et A. Clarke. 2000. « The Psychological Burden of Peanut Allergy as Perceived by Adults with Peanut Allergy and the Parents of Peanut-Allergic Children ». *Clinical and Experimental Allergy: Journal of the British Society for Allergy and Clinical Immunology* 30 (8): 1135- 43. <https://doi.org/10.1046/j.1365-2222.2000.00889.x>.
- Sampson, Hugh A., Liam O'Mahony, A. Wesley Burks, Marshall Plaut, Gideon Lack, et Cezmi A. Akdis. 2018. « Mechanisms of Food Allergy ». *The Journal of Allergy and Clinical Immunology* 141 (1): 11- 19. <https://doi.org/10.1016/j.jaci.2017.11.005>.
- Sicherer, Scott H., et Hugh A. Sampson. 2018. « Food Allergy: A Review and Update on Epidemiology, Pathogenesis, Diagnosis, Prevention, and Management ». *The Journal of Allergy and Clinical Immunology* 141 (1): 41- 58. <https://doi.org/10.1016/j.jaci.2017.11.003>.
- Yu, Wong, Deborah M. Hussey Freeland, et Kari C. Nadeau. 2016. « Food Allergy: Immune Mechanisms, Diagnosis and Immunotherapy ». *Nature Reviews. Immunology* 16 (12): 751- 65. <https://doi.org/10.1038/nri.2016.111>.
- Fairweather-Tait, S., Harvey, L., Heath, A.-L. M. & Roe, M. 2007. Effect of SNPs on iron metabolism. *Genes & Nutrition*, 2, 15-19.
- Wellcome Trust Case Control, C. 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, 447, 661-78.
- Ma, Q. & Lu, A. Y. 2011. Pharmacogenetics, pharmacogenomics, and individualized medicine. *PharmacologyReview*, 63, 437-59.

Bresciani, G., Cruz, I. B., De Paz, J. A., Cuevas, M. J. & Gonzalez-Gallego, J. 2013. The MnSOD Ala16Val SNP: relevance to human diseases and interaction with environmental factors. *Free Radical Research*, 47, 781-92.

Nelson, M. R., Marnellos, G., Kammerer, S., Hoyal, C. R., Shi, M. M., Cantor, C. R. & Braun, A. 2004. Large-scale validation of single nucleotide polymorphisms in gene regions. *GenomeResearch*, 14, 1664-8.

Sachidanandam, R., Weissman, D., Schmidt, S. C., Kakol, J. M., Stein, L. D., Marth, G., Sherry, S., Mullikin, J. C., Mortimore, B. J. & Willey, D. L. 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, 409, 928-933.

Syvanen, A. C. 2001. Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nature ReviewGenetics*, 2, 930-42.

Casadio, R., Vassura, M., Tiwari, S., Fariselli, P. & Luigi Martelli, P. 2011. Correlating disease-related mutations to their effect on protein stability: A large-scale analysis of the human proteome. *Human Mutation*, 32, 1161- 1170.

Chasman, D. & Adams, R. M. 2001. Predicting the functional consequences of non-synonymous single nucleotide polymorphisms: structure-based assessment of amino acid variation1. *Journal of MolecularBiology*, 307, 683-706.

Ward, L. D. & Kellis, M. 2012. Interpreting noncoding genetic variation in complex traits and human disease. *Nature Biotechnology*, 30, 1095-1106.

Visel, A., Zhu, Y., May, D., Afzal, V., Gong, E., Attanasio, C., Blow, M. J., Cohen, J. C., Rubin, E. M. & Pennacchio, L. A. 2010. Targeted deletion of 166 the 9p21 non-coding coronary artery disease risk interval in mice. *Nature*, 464, 409-412.

Prado-Montes De Oca, E., Velarde-Félix, J. S., Ríos-Tostado, J. J., Picos-Cárdenas, V. J. & Figuera, L. E. 2009. SNP 668C (- 44) alters a NF-κB1 putative binding site in non-coding strand of human β-defensin 1 (<i> DEFB1) and is associated with lepromatous leprosy. *Infection, Genetics and Evolution*, 9, 617-625.

Cargill, M., Altshuler, D., Ireland, J., Sklar, P., Ardlie, K., Patil, N., Lane, C. R., Lim, E. P., Kalyanaraman, N. & Nemesh, J. 1999. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nature Genetics*, 22, 231-238.

Sunyaev, S., Ramensky, V., Koch, I., Lathe III, W., Kondrashov, A. S. & Bork, P. 2001. Prediction of deleterious human alleles. *Human Molecular Genetics*, 10, 591-597.

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*. 2001; 29(1):308–11.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science*. 2001;291:1304-1351.

Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*. 2001;409:928-933.

Lander ES. The new genomics: Global views of biology. *Science*. 1996;274:536-539.

Lin MT, Storer B, Martin PJ, Tseng LH, Gooley T, Chen PJ, et al. Relation of an interleukin-10 promoter polymorphism to graft-versus-host disease and survival after hematopoietic-cell transplantation. *The New England Journal of Medicine*. 2003;349:2201-2210

Tokuhiro S, Yamada R, Chang X, Suzuki A, Kochi Y, Sawada T, et al. An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. *Nature Genetics*. 2003;35:341-348.

Betticher DC, Thatcher N, Altermatt HJ, Hoban P, Ryder WD, Heighway J. Alternate splicing produces a novel cyclin D1 transcript. *Oncogene*. 1995;11:1005-1011.

Helms C, Cao L, Krueger JG, Wijsman EM, Chamian F, Gordon D, et al. A putative RUNX1 binding site variant between SLC9A3R1 and NAT9 is associated with susceptibility to psoriasis. *Nature Genetics*. 2003;35:349-356.

Duan J, Wainwright MS, Comeron JM, Saitou N, Sanders AR, Gelernter J, et al. Synonymous mutations in the human dopamine receptor D2 (DRD2) affect mRNA stability and synthesis of the receptor. *Human Molecular Genetics*. 2003;12:205-216.

Fareed M, Afzal M. Single-nucleotide polymorphism in genome-wide association of human population: A tool for broad spectrum service. *Egyptian Journal of Medical Human Genetics*. 2013;14(2):123-134.

Voronko, O., Bodoev, N. & Archakov, A. 2008. The use of SNP markers for estimation of individual genetic predisposition to diabetes mellitus type 1 and 2. *Biochemistry (Moscow) Supplemental Series B: Biomedical Chemistry*, 2, 126-132.

Yip, W. & Lange, C. 2011. Quantitative trait prediction based on genetic marker array data, a simulation study. *Bioinformatics*, 27, 745-748.

Stefansson, H., Ophoff, R. A., Steinberg, S., Andreassen, O. A., Cichon, S., Rujescu, D., Werge, T., Pietiläinen, O. P. H., Mors, O. & Mortensen, P. B. 2009. Common variants conferring risk of schizophrenia. *Nature*, 460, 744-747.

Giacomini, K., Brett, C., Altman, R., Benowitz, N., Dolan, M., Flockhart, D., Johnson, J., Hayes, D., Klein, T. & Krauss, R. 2007. The pharmacogenetics research network: from SNP discovery to clinical drug response. *ClinicalPharmacology&Therapeutics*, 81, 328-345.

Underhill, P. A. & Kivisild, T. 2007. Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. *AnnualReview. Genetics*, 41, 539-56.

Cameron L, Depner M, Kormann M, Klopp N, Illig T, von Mutius E, Kabesch M (2009) Genetic variation in CRTh2 influences development of allergic phenotypes. *Allergy* 64(10):1478– 1485.

Smith M, Tourigny MR, Noakes P, Thornton CA, Tulic MK, Prescott SL (2008) Children with egg allergy have evidence of reduced neonatal CD4(+) CD25(+)CD127(lo/-) regulatory T cell function. *J Allergy ClinImmunol* 121(6):1460–1466, 1466 e1461–1467.

Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* 2012;40:W452–7.

Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. *PLoS One.* 2012;7(10):e46688.

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010;7(4):248–9.

Tang H, Thomas PD. PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation. *Bioinformatics.* 2016;32(14):2230–2.

Emadi, E.; Akhoundi, F.; Kalantar, S. M.; Emadi-Baygi, M. Predicting the Most Deleterious Missense NsSNPs of the Protein Isoforms of the Human HLA-G Gene and in Silico Evaluation of Their Structural and Functional Consequences. *BMC Genet.* 2020, 21 (1), 94.

Lopez-Ferrando V, Gazzo A, de la Cruz X, Orozco M, Gelpi JL. PMut: a webbased tool for the annotation of pathological variants on proteins, 2017 update. *NucleicAcidsRes.* 2017;45(W1):W222–w8.

ANNEXE

Tableau 1 : Résultats des nsSNP à haut risques identifiés la méthode in silico SIFT pour le STAT3.

SNP ID	Posistion d'AA	SIFT
rs1386600261	N5Y	Deleterious
rs1041081083	N5S	Deleterious
rs780720013	R13Q	Deleterious
rs1462777520	L21V	Deleterious
rs751281347	D24N	Deleterious
rs1803125	Q32K	Deleterious
rs1340218209	F33V	Deleterious
rs748204289	A35S	Deleterious
rs1292306960	P36S	Deleterious
rs868859792	A46V	Deleterious
rs766875947	S48I	Deleterious
rs766875947	S48N	Deleterious
rs775122881	D65G	Deleterious
rs375760579	Q66E	Deleterious
rs1187104418	Y68H	Deleterious
rs1281834571	Y79C	Deleterious
rs776494537	N82S	Deleterious
rs1423582073	R85K	Deleterious
rs1344374375	Q91R	Deleterious
rs1452139691	M99V	Deleterious
rs1408283351	R103W	Deleterious
rs1417279405	R103Q	Deleterious
rs749626783	R107Q	Deleterious
rs780604324	C108Y	Deleterious
rs1215714766	S113L	Deleterious
rs964892419	R114C	Deleterious
rs1344978308	R114H	Deleterious
rs1360347337	L115F	Deleterious
rs757305223	T118S	Deleterious
rs1234764558	T118I	Deleterious
rs751614036	T121S	Deleterious
rs1440082733	A122P	Deleterious
rs774724351	A122V	Deleterious
rs1410411796	A123T	Deleterious
rs1475754644	A123V	Deleterious
rs574370336	Q125E	Deleterious
rs777883261	G126E	Deleterious
rs1222451818	A129T	Deleterious

rs1195908534	N130S	Deleterious
rs1442943787	T133P	Deleterious
rs1008624238	V136L	Deleterious
rs1008624238	V136L	Deleterious
rs1008624238	V136M	Deleterious
rs17878478	M143I	Deleterious
rs1477283446	H147Y	Deleterious
rs1029086530	L158I	Deleterious
rs967009897	V165L	Deleterious
rs758408552	N175H	Deleterious
rs995902901	N175K	Deleterious
rs748237742	L179F	Deleterious
rs866556426	S181I	Deleterious
rs748184550	D184G	Deleterious
rs749529243	L188M	Deleterious
rs781034418	N189K	Deleterious
rs1423110870	S194T	Deleterious
rs895801743	S194L	Deleterious
rs751411956	T196N	Deleterious
rs1048207886	R197K	Deleterious
rs993574412	Q198K	Deleterious
rs374063766	Q198H	Deleterious
rs1356637796	A209V	Deleterious
rs753370239	R215G	Deleterious
rs1175109539	S216G	Deleterious
rs1325088493	V218M	Deleterious
rs762301765	S219N	Deleterious
rs764674735	A222V	Deleterious
rs1004741102	A227V	Deleterious
rs1348311948	E229K	Deleterious
rs993916477	E229G	Deleterious
rs1284111065	V231M	Deleterious
rs1205738213	T234S	Deleterious
rs773138024	T236S	Deleterious
rs371541785	T236R	Deleterious
rs376677265	D237N	Deleterious
rs755746762	D242E	Deleterious
rs1279915769	R245G	Deleterious
rs1221396883	R246Q	Deleterious
rs757347742	Q248P	Deleterious
rs767159289	I249T	Deleterious
rs1269919034	I249M	Deleterious
rs1451984094	R262Q	Deleterious
rs745393806	S273P	Deleterious
rs1161466672	R302Q	Deleterious
rs765856200	V310M	Deleterious
rs759527569	P327A	Deleterious

rs1249457196	G342S	Deleterious
rs769072802	L358W	Deleterious
rs781724933	V375F	Deleterious
rs781724933	V375I	Deleterious
rs1455469288	N400Y	Deleterious
rs749583802	D427G	Deleterious
rs146620441	A498G	Deleterious
rs1281575473	W510R	Deleterious
rs372641163	G536V	Deleterious
rs1471271759	A555T	Deleterious
rs865836493	P603Q	Deleterious
rs1340488106	V667L	Deleterious
rs1244109795	A697T	Deleterious
rs774279920	T721S	Deleterious
rs11547455	S727F	Deleterious
rs757779747	N740H	Deleterious
rs151033214	G743V	Deleterious
rs759106857	T762I	Deleterious
rs759106857	T762S	Deleterious
rs1191676333	S768F	Deleterious
rs1372796820	M770V	Deleterious

Tableau 3.2 : Liste des résultats de l'analyse des nsSNPs FOXP3 par SIFT

SNP ID	Changement d'AA	SIFT
rs1389154157	P4S	Deleterious
rs782104014	S10L	Deleterious
rs868982237	G17V	Deleterious
rs782799706	P20S	Deleterious
rs782138321	G21R	Deleterious
rs782426843	A22V	Deleterious
rs782304327	S23L	Deleterious
rs1194116356	G37V	Deleterious
rs1432262933	G47V	Deleterious
rs1170772176	R51Q	Deleterious
rs1382549860	S67L	Deleterious
rs138970962	V79L	Deleterious
rs376715533	R85W	Deleterious
rs1304549886	H91P	Deleterious
rs782303740	R114Q	Deleterious
rs782640594	H121Y	Deleterious
rs782665522	T135P	Deleterious
rs1274949988	T138I	Deleterious
rs1243920194	V161G	Deleterious
rs140222626	S184L	Deleterious
rs368066149	C204Y	Deleterious
rs1045776635	H217R	Deleterious
rs2232369	A220V	Deleterious
rs781965746	M238R	Deleterious
rs782071730	S254N	Deleterious
rs781960959	M256I	Deleterious
rs781844954	A273T	Deleterious

rs782789484	D276N	Deleterious
rs781902347	S279F	Deleterious
rs74162067	V292I	Deleterious
rs1224110402	E300V	Deleterious
rs782351468	R309W	Deleterious
rs782731811	R309Q	Deleterious
rs782591493	K332R	Deleterious
rs782649847	A343T	Deleterious
rs1428528910	R358Q	Deleterious
rs886821043	F367L	Deleterious
rs782443295	R369H	Deleterious
rs781854593	R414C	Deleterious
rs782786757	R414H	Deleterious
rs145368924	R417W	Deleterious
rs925367928	R417Q	Deleterious
rs1317916668	S418T	Deleterious
rs781796588	S418R	Deleterious
rs369332983	N426K	Deleterious
rs782088695	P429S	Deleterious

