



الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la
Recherche Scientifique
جامعة البليدة 1
Université Blida 1



Faculté des Sciences de la Nature et de la Vie
Département de Biotechnologies & Agro-Ecologie

Mémoire

En vue de l'obtention du Diplôme d'un Master Académique

Option

Biotechnologie et Pathologie Moléculaire

Thème

**Etude *in Silico* du polymorphisme du gène Nox4 (NADPH
oxydase) et son association au diabète**

Présenté par Melle. Drif Amina

Melle. Tahar Yasmine

Devant le Jury :

<i>Mr Boukhatem MN.</i>	<i>Pr</i>	<i>SNV, Blida1</i>	<i>Président</i>
<i>Mme Guessaibia N.</i>	<i>MCA</i>	<i>SNV, Blida1</i>	<i>Examinatrice</i>
<i>Mme Eddaikra A</i>	<i>MCB</i>	<i>SNV, Blida1</i>	<i>Promotrice</i>

Remerciements

On remercie Dieu le tout puissant de nous avoir donné la santé et la volonté d'entamer et de terminer ce travail.

*Nous exprimons notre profonde gratitude à notre promotrice **Dr. Eddaikra A**, maître de conférences au département de biologie de l'Université SAAD DAHLA Blida 1 de nous avoir fait confiance et ainsi de nous avoir formé avec rigueur et patience et aussi pour nous avoir accordé sa confiance, ses conseils et son soutien continu tout au long de ce trajet.*

*Nous exprimons notre profonde gratitude à **Dr. Guessaibia N**, Maître de conférences au département de biologie de l'université SAAD DAHLAB Blida 1 pour l'honneur qu'il nous a fait en acceptant d'examiner notre travail*

*Nous exprimons notre profonde gratitude à **Mr. Boukhatem MN**, Professeur au département de Biologie de l'université SAAD DAHLAB Blida 1 d'avoir accepté de présider notre jury de soutenance.*

On souhaite adresser nos remerciements aussi à tout l'ensemble du personnel du département de Biotechnologie de la faculté des sciences, aussi tous ceux qui ont contribué de près ou de loin à la réalisation de ce travail.

Dédicaces

Je dédie ce travail à mes très chers parents, qui m'ont toujours soutenu et encouragé durant toutes mes études, ma mère ZAHRA que j'aime plus que tout qui toujours me soutient dans les moments difficiles. je lui souhaite longue vie et à mon père OMAR défunt Allah yerahmou.

A ma chère sœur Imane qui n'a pas cessée de me conseiller, encourager et soutenir tout au long de mes études, que Dieu la protège et lui offre la chance et le bonheur.

A mon adorable petite sœur Sara qui sait toujours comment procurer la joie et le bonheur pour toute la famille.

A mon chère frère Khaled qui m'a aidé et supporté.

A mes grands-mères, mes tantes Saliha, Sabrina et Zohour et mes cousins, que dieu leur donne une longue et joyeuse vie

A ma chère binôme Yasmine pour son soutien et sa patience et sa compréhension tout au long de ce projet.

A mes chères amies Yasmine, Manel, Sabrina, Lina, Fatima, Kaouther, Nadjwa, Nadhir.

A tout ma grande famille Drif, Nachef

Merci pour leurs amours et leurs encouragements

Amina

Dedicaces

First and foremost, praise and appreciation are due to God almighty for the numerous blessings he has bestowed upon me

For given me mercy, guidance, and good company that encouraged me throughout all the challenging moments of completing my thesis.

To my precious parents Tahar and Ghania, whose words of encouragement and push for tenacity ring in my ears, I wouldn't be here without you, through you I saw and understood faith, patience, honesty and generosity, may God bless and protect you.

To my sisters Lila and Sabrina, my brothers Mohammed and walid, my aunts Assia and khadidja and all my family.

To my dear thesis partner Amina for her support and patience and understanding throughout this project.

To my teachers Ms.Farida, Ms.Bachira, Mr.Ali, Ms.Rahmani, Mr.khadraoui, Mr.Akloule, Ms.Jemiaa , Ms.Ben youcef, Mr.Ismail Mokhtar, Ms.lazregue, Mr.Henni, for being such incredible and inspiring human beings.

To my friends Amina, Sabrina, Nadjoua, lina, yasmine, kaouthar and Nadhir who were there for me all along the University journey, you are the best.

To Lilly, Mohamed and Marwa, I love you.

In memory of

Karima Dahdouh, the purest soul, it is a great blessing to have known her.

Tahar Ahmed, my sweet grandfather for he was an honorable man.

Yasmine

Sommaire

Introduction	1
Chapitre1 : Données bibliographique	3
1.1 Gène NOX4	3
1.1.1 Structure et localisation	3
1.1.2 La protéine de Nox4	3
1.1.3 Interaction entre Nox4 et P22phox	4
1.1.4 Le rôle physiologique de Nox4	4
1.1.5 La famille de Nox	5
1.1.6 Isoformes de Nox4	6
1.2 Polymorphismes.....	7
1.2.1 Polymorphisme d'un seul nucléotide.....	7
1.2.2 Polymorphisme du gène NOX4.....	10
1.3 Diabète	10
1.3.1 Diabète de type 1	11
1.3.2 Diabète de type 2	11
1.3.3 Complications du diabète et NOX4.....	12
Chapitre 2 : Matériel et Méthodes	15
2. Matériel et Méthodes.....	15
2.1 Matériel.....	15
2.1.1 Matériel biologique	15
2.1.2 Matériel non biologique	15
2.2 Méthodes.....	17
2.2.1 Etude de gène et recherche des SNP	18
2.2.2 Identification des rs SNPs du gène de NOX4	24
2.2.3 Docking Moléculaire	38
2.2.4 Méthode du Docking moléculaire	45
Chapitre 3 : Résultats	46
3. Résultats	46
3.1 Résultats de l'OUTIL NCBI.....	46
3.2 Cas cliniques de l'association des « rs » du gène NOX4.....	46
3.3 Résultats de l'outil Ensembl	47
3.4 Résultats de l'OUTIL UCSC	47
3.5 Résultats de l'OUTIL LitVar	48
3.6 Résultat de la sélection des variants (SNPs).....	49
3.6.1 Identification fonctionnelle des mutations délétères	50

3.6.2	Prédiction de l'effet de substitution d'acides aminés sur la stabilité de la protéine	50
3.6.3	Prédiction des effets sur les domaines protéiques par MutPred.	51
3.6.4	Résultat Hope Project	52
3.7	Prédiction in silico de la structure 3D du gène NOX4.....	54
3.7.1	Résultat de l'outil SUISSE MODEL	54
3.8	Résultat du Docking Moléculaire	56
3.8.1	Préparation de la cible (enzyme)	56
3.8.2	Préparation du ligand (FAD)	57
3.8.3	Docking moléculaire.....	58
Chapitre 4 : Discussion.		61
Conclusion		66
Références bibliographiques.....		68
Annexes.....	

Liste des figures

Figure 1: la localisation chromosomique de gène Nox4 (chromosome 11 q14.2).....	3
Figure 2:structure de la protéine du gène Nox4.	4.
Figure 3:structure de la famille des NADPH Oxydase.	6
Figure 4:Types de polymorphisme	7
Figure 5:Polymorphisme de nucléotide unique (SNP) dans le génome individuel et dans la population.	8
Figure 6:Modifications génomique d'un seul nucléotide.....	9
Figure 7:Conséquences possibles qui peuvent résulter des SNP.....	9
Figure 8:Physiopathologie du diabète sucre type 1 et 2.	11
Figure 9:Interaction HuR/NOX4 dans la néphropathie diabétique.	13
Figure 10:Rôle des isoformes NOX dans la pathogenèse de la néphropathie diabétique.	14
Figure 11:méthode d'étude et d'analyse du polymorphisme du gène NOX4.	17
Figure 12:Les étapes de recherche sur NCBI.	18
Figure 13: Les étapes de recherche sur PubMed.	19
Figure 14:Les étapes de recherche sur GeneCards.....	20
Figure 15: la base des données LitVar.....	21
Figure 16:Les étapes de recherche sur SNPinfo.....	22
Figure 17: Méthode de recherche et d'étude des rs du gène NOX4, exploités et non exploités dans des cas cliniques « études cas-témoins ».....	23
Figure 18:Les étapes de recherche sur UCSC.	24
Figure 19: Les étapes de recherche sur la base des données Ensembl.	25
Figure 20: tableaux des transcrits de gène NOX4.....	26
Figure 21: la séquence nucléotidique de gène NOX4.....	26
Figure 22:Les étapes de recherche sur SIFT.	27
Figure 23:Les étapes de recherche sur PolyPhen-2.	28
Figure 24:Les étapes de recherche sur SNAP2.	29
Figure 25:Les étapes de recherche sur PhD-SNP.....	29
Figure 26:les étapes de recherche sur I-Mutant2.0.....	30
Figure 27:Les étapes de recherche sur MutPred2.....	31
Figure 28:les étapes de recherche avec HOPE project.	32
Figure 29:Les étapes de recherche sur UniProt.....	33
Figure 30:Les différentes informations fournis par UniProt sur le gène NOX4.	33
Figure 31:Les étapes de recherche sur SMART.....	34
Figure 32:Les étapes de recherche des ORF (NCBI ORF-finder).....	36
Figure 33: les nombre des ORF (31) obtenues et l'ORF le plus long dans le gène NOX4.....	36
Figure 34:la modélisation dans la base de données SWISS-MODEL.....	37
Figure 35: Les étapes de la modélisation.	37
Figure 36:Méthode du Docking moléculaire (Amarrage moléculaire).	39
Figure 37:La base de données PDB.....	40
Figure 38:Récupération du format PDB de la protéine.	41
Figure 39:Recherche des ligands dans SWISS-MODEL.	41.
Figure 40:la recherche du ligand Flavine adénine dinucléotide (FAD) de la protéine NOX4..	42.
Figure 41:La récupération du ligand FAD du site PubChem.	43
Figure 42:l'identification du site de liaison avec PrankWeb.....	44
Figure 43:Nombre des variants polymorphiques récupérés à partir des différentes plateformes.	46

Figure 44:Diagramme circulaire représentant les différents types de variations de NOX4 selon le navigateur Ensembl.	47
Figure 45:Affichage des donnes dans le navigateur UCSC.	48
Figure 46:Diagramme circulaire représentant les variations nucléotidiques du gène NOX4 détectées par SIFT.	49
Figure 47:Close-up sur la mutation. La protéine est colorée en gris, les chaînes latérales du résidu de type sauvage et du résidu mutant sont représentées et colorées respectivement en vert et en rouge.	53
Figure 48:Modélisation comparative de la protéine NOX4 à l'aide de SWISS-MODEL.	54
Figure 49:Structure 3D de la protéine 7D3E qui code pour le gène NOX4 présenté par SWISS-MODEL.	55
Figure 50: Acides aminés formant le site actif selon PyMOL.	57
Figure 51:L'interaction entre ligand FAD et les résidus de site actif de NOX4 selon Discovery studio visualize.	60
Figure 52:L'interaction entre ligand FAD et les résidus de site actif de NOX4 selon Discovery studio visualize.	60
Annexes:	
Figure 53: séquence nucléotidique (FASTA) du gène NOX4 "NM_016931.5"	
Figure 54:les déférentes variants de gène NOX4 (NCBI dbvar)	
Figure 55:les transcrits de gène NOX4 (Ensembl).	
Figure 56: les déférentes SNP de gène NOX4 (Ensembl).....	
Figure 57:les Exons de gène NOX4 (Ensemble).....	
Figure 58: la séquence protéique du NOX4 (UniProt).	
Figure 59: recherche des ORF (ORF Finder)	
Figure 60: L'ORF le plus longue (ORF-Finder NCBI).	
Figure 61: cartographie génétique du gène NOX4.	
Figure 62:Préparation de la protéine de NOX4 (7D3E) (Autodock tools).....	
Figure 63:Préparation du site actif de la protéine NOX4 (7D3E) pour le Docking moléculaire (Autodock tools).....	
Figure 64:Préparation du ligand Flavin adenine dinucleotide (FAD) (Autodock tools)	
Figure 65:La configuration des coordonnées du Grid box pour le Docking moléculaire (Autodock tools).	

Liste des tableaux.

Tableau I: Distribution Tissulaire et Fonctions de la Nox4.....	5
Tableau II: Description du gène Nox4.	15
Tableau III: Les Baes des données utilisé dans la recherche.....	15
Tableau IV: Les bases des données et les logiciels utilisés dans le dorking de la molécule.	16
Tableau V: Références GenBank NCBI du gène NOX4.....	18
Tableau VI: les domaines du la proteine NOX4 (SMART).	35
Tableau VII: Résultats des « rs » du gène de la NOX4 associés dans les cas cliniques.	46
Tableau VIII: Polymorphismes de NOX4 associes au diabète selon LitVar.	48
Tableau IX: Prédiction de SIFT, PolyPhen-2, SNAP2 pour les rs de NOX4.	50
Tableau X: Prédiction des SNPs délétères et analyse de leurs effets sur la stabilité de la protéine.	51
Tableau XI: Prédiction du serveur MutPred pour l'effet des substitutions d'acides aminés de NOX4.	51
Tableau XII: Résidus de site actif de NOX4	56
Tableau XIII: Résultats énergétiques de ligand de référence FAD.	57
Tableau XIV: les coordonnées de Grid Box pour l'amarrage du NOX4-FAD	58
Tableau XV: Résultats de SCORE et RMSD de docking moléculaire.	58
Tableau XVI: Rapport des interactions entre les résidus du site actif avec le ligand FAD.....	59

Liste des abréviations

3D : *Three-Dimensional*.
AGE : *Advanced Glycation End-product*.
ATP5F1P6 : *ATP synthase F(0) complex subunit B1, mitochondrial*.
CASP9 : *Caspase-9*.
castp : *Computed Atlas of Surface Topography of proteins*.
CELA2B : *Chymotrypsin-like Elastase family member 2B*.
DDG : *Delta Delta G*.
DUOX1 : *Dual Oxidases 1*.
FRO : *Radicaux Libres Oxygénés*.
GCID : *Gene Card Identification*.
GMQE : *Global Model Quality Estimation*.
GoDARTS : *Genetics of Diabetes Audit and Research, Tayside and Scotland*.
GRCh38 : *Genome Reference Consortium Human Build 38*.
HG : *Hyperglycémie*.
HIS : *Histidine*.
HuR : *Human antigen R*.
IC : *Interval of Confidence*.
ID : *Identification*.
INPP5A : *Inositol Polyphosphate-5-Phosphatase A*.
NCBI : *National Center for Biotechnology Information*.
NLM : *National Library of Medicine*.
NODAT : *New onset diabetes after transplantation*.
OR : *Odd Ratio*.
ORF : *Open Reading Frame*.
PhD-SNP : *Predictor of human Deleterious Single Nucleotide Polymorphisms*
PHE : *Phenylalanine*.
PolyPhen-2 : *Polymorphism Phenotyping v2*.
PTP1B : *Tyrosine-Protein Phosphatase 1B*.
RAGE : *Receptor for Advanced Glycation End-products*.
RCSB PDB : *Research Collaboratory for Structural Bioinformatics PDB*.
Refseq : *Reference sequence*
RNAseq : *séquençage de l'ARN*
rs : *Reference SNP*.
RSMD : *Root-Mean-Square Deviation (déviation de la racine de la moyenne des carrés)*.
SAV : *Single Amino acid Variants*.
SIFT : *Sorting Intolerant From Tolerant*
SMART : *Simple Modular Architecture Research Tool*
SNAP2 : *Screening for Nonacceptable Polymorphisms*.
TRP : *Tryptophan*.
UCSC : *University of California Santa Cruz*.
UTR *Untranslated Region*.
vdW : *Rayons de van der Waals des atomes*.

Résumé

Pour sa prévention et pour comprendre la physiopathologie du diabète plusieurs voies d'études ont été sollicitées. En effet, au cours des dernières décennies, l'étude du polymorphisme SNP (polymorphisme à un seul nucléotide) porté par un gène permet d'estimer son association au diabète ainsi qu'à d'autres pathologies.

Le gène NOX4 code pour un membre de la famille d'enzymes NOX qui fonctionne comme la sous-unité catalytique du complexe NADPH oxydase. Ce gène situé sur le chromosome 11 en position 14 du bras long (11q14) est connue par sa fonction principale qui est la génération des espèces réactives en oxygène (ROS), mais qui a notamment d'autres rôles physiologiques tels que la régulation de la cascade de signalisation de l'insuline.

L'objectif de notre étude est d'étudier l'association du polymorphisme du gène NOX4 avec le diabète par une méthode *in silico*. Cette dernière vise à identifier les références snip « rs » des polymorphismes a nucléotide unique (SNPs) associés au diabète. Pour ce faire, nous avons utilisé une méthode bioinformatique faisant recourt à plusieurs bases de données open sources dont : Pub Med, NCBI, Ensembl, UCSC, SIFT, PolyPhen-2. Les données ont été récoltées, triées puis analysées.

Nos résultats montrent bien une association entre polymorphisme du gène NOX4 et le diabète. En effet, nous avons identifié quatre « rs » : rs614128 (Intron 2) ; rs490934 (Intron 2) ; rs553635 (Promoteur) et rs3017887 (5'UTR) du gène NOX4 significativement associés à la néphropathie diabétique (DN). En outre, nous avons aussi identifier trois « rs » rs10765219 (Intron 13) et rs11018670 (Intron 12), rs3913535 (Intron 16) du gène NOX4 significativement associés à la rétinopathie diabétique (RD). Cependant, le rs1836882 (Promoteur) a été identifier significativement associer aux NODAT (Nouvelle apparition du diabète après une transplantation rénale).

L'analyse *in silico* de plus de quarante-sept mille SNPs de NOX4 a permis d'identifies quatre autres rs de NOX4 à effet délétère : rs13931533, rs56061986, rs373220761 et rs148405952 non-explorés dans les études cas-témoins associées au diabète, tandis que l'étude du Docking moléculaire a révélé les trois acides aminés (TRP1305, HIS1319, PHE1321) du site de liaison de la protéine NOX4 dans l'état native qui interagissent directement avec son ligand flavin adenine dinucleotide (FAD).

Nos résultats suggèrent, que ces polymorphismes peuvent altérer la fonction du gène NOX4 en modifiant sa structure. Cette altération pourrait affecter considérablement l'expression du gène NOX4 et la fonction de la protéine qui code pour ce gène.

Les mots clés : NOX4, polymorphisme, SNP, diabète, *in silico*, bioinformatique.

Abstract

For its prevention and to understand the pathophysiology of diabetes several avenues of study have been solicited. Indeed, in the last decades, the study of SNP polymorphism (single nucleotide polymorphism) carried by a gene allows to estimate its association with diabetes as well as with other pathologies.

The NOX4 gene codes for a member of the NOX family of enzymes that functions as the catalytic subunit of the NADPH oxidase complex. This gene located on chromosome 11 at position 14 of the long arm (11q14) is known for its primary function which is the generation of reactive oxygen species (ROS), but has other physiological roles such as the regulation of the insulin signaling cascade.

The objective of our study is to investigate the association of the NOX4 gene polymorphism with diabetes by an *in silico* method. The latter aims at identifying the snip "rs" references of the single nucleotide polymorphisms (SNPs) associated with diabetes. To do this, we used a bioinformatics method that makes use of several open source databases including: Pub Med, NCBI, Ensembl, UCSC, SIFT, PolyPhen-2. The data were collected, sorted and analyzed.

Our results show an association between NOX4 gene polymorphism and diabetes. Indeed, we identified four "rs": rs614128 (Intron 2); rs490934 (Intron 2); rs553635 (Promoter) and rs3017887 (5'UTR) of the NOX4 gene significantly associated with diabetic nephropathy (DN). In addition, we also identified three "rs" rs10765219 (Intron 13) and rs11018670 (Intron 12), rs3913535 (Intron 16) of NOX4 gene significantly associated with diabetic retinopathy (DR). However, rs1836882 (Promoter) was identified significantly associated with NODAT (New onset diabetes after kidney transplantation).

The *in silico* analysis of more than forty seven thousand NOX4 SNPs identified four additional deleterious NOX4 rs: rs13931533, rs56061986, rs373220761, and rs148405952 not explored in diabetes-associated case-control studies, while molecular docking study revealed the three amino acids (TRP1305, HIS1319, PHE1321) of the NOX4 protein binding site in the native state that interact directly with its ligand flavin adenine dinucleotide (FAD).

Our results suggest that these polymorphisms may alter the function of the NOX4 gene by modifying its structure. This alteration could significantly affect the expression of the NOX4 gene and the function of the protein that encodes this gene.

Key words: NOX4, polymorphism, SNP, diabetes, *in silico*, bioinformatics.

المخلص

لوقاية منه وفهم علم وظائف الأعضاء لمرض السكري تم اتباع عدة طرق للدراسة. في الواقع، على مدى العقود القليلة الماضية، أتاحت دراسة تعدد الأشكال SNP (تعدد أشكال النكليوتيدات المفردة) التي يحملها الجين تقدير ارتباطه بمرض السكري بالإضافة إلى أمراض أخرى .

يقوم جين NOX4 بتنشيف عضو من عائلة إنزيمات NOX التي تعمل كوحدة فرعية تحفيزية لمركب NADPH أوكسيداز. هذا الجين الموجود على الكروموسوم 11 في الموضع 14 من الذراع الطويلة (q1411) معروف بوظيفته الأساسية وهي توليد أنواع الأوكسجين التفاعلية (ROS) ، ولكن له أدوار فسيولوجية أخرى مثل تنظيم سلسلة إشارات الأنسولين.

الهدف من دراستنا هو دراسة ارتباط تعدد الأشكال لجين NOX4 بمرض السكري بطريقة (in silico). يهدف هذا الأخير إلى تحديد الإشارات القصاصة "rs" لتعدد أشكال النكليوتيدات المفردة (SNPs) المرتبطة بمرض السكري. للقيام بذلك، استخدمنا طريقة المعلوماتية الحيوية باستخدام العديد من قواعد البيانات مفتوحة المصدر بما في ذلك NCBI و Pub Med و Ensembl و UCSC و SIFT و PolyPhen-2. تم جمع البيانات وفرزها ثم تحليلها.

تظهر نتائجنا بوضوح وجود ارتباط بين تعدد الأشكال لجين NOX4 ومرض السكري. في الواقع، لقد حددنا أربعة "rs" (rs3017887 (5'UTR) و rs553635 (Promoteur) ؛ rs490934 (Intron2) ؛ rs614128 (Intron 2) و rs10765219 (DN) بالإضافة إلى ذلك، نحدد أيضاً ثلاثة "rs" rs3913535 (Intron 16) و rs11018670 (Intron 12) و rs1836882 (Promoteur) على أنه مرتبط بشكل كبير بـ (NODAT) ظهور مرض السكري الجديد بعد زرع الكلى.

في تحليل سيليكو لأكثر من سبعة وأربعين ألف SNPs من NOX4 حددت أربعة rs NOX4 إضافية ذات تأثير ضار : rs13931533 ، rs56061986 ، rs373220761 و rs148405952 لم يتم استكشافها في دراسات مراقبة الحالة المرتبطة بمرض السكري ، في حين كشفت دراسة Docking الجزيئية عن الأحماض الأمينية الثلاثة (TRP1305 و HIS1319 و PHE1321) لموقع ربط بروتين NOX4 الأصلي التي تتفاعل مباشرة مع مركب فلافين ثنائي نيوكليوتيد الأدينين (FAD).

تشير نتائجنا إلى أن هذه الأشكال المتعددة قد تغير وظيفة جين NOX4 عن طريق تعديل بنيته. يمكن أن يؤثر هذا التغيير بشكل كبير على التعبير عن جين NOX4 ووظيفة البروتين الذي يرمز لهذا الجين.

الكلمات المفتاحية : NOX4، تعدد الأشكال، تعدد أشكال النكليوتيدات المفردة، السكري، ان سيليكو، المعلوماتية الحيوية.

INTRODUCTION

Introduction

Il est bien établi que la prévalence du diabète de type 1 (DT1) augmente dans le monde entier et varie selon les pays. En 2021, le diabète affectait plus de 537 millions de personnes dans le monde (soit 1 personne sur 10), dont 61 millions en Europe avec 6, 7 millions de décès en 2021 (**Sun et al., 2022**).

La NADPH oxydase 4 (Nox4) est un enzyme membranaire codé par le gène Nox4 est situé sur le chromosome 11 en position 14 du bras long (11q14), appartenant à la classe des oxydoréductases a été identifiée pour la première fois par Shiose et al. En tant que nouveau Nox non phagocytaire dans le rein en 2001. C'est la seule isoforme de la famille Nox pour la production constitutive de peroxyde d'hydrogène (H₂O₂) (**Shiose et al., 2001**). D'abord été appelée RENOX car elle est fortement exprimée dans le cortex rénal. C'est une oxydase importante dans les cellules rénales humaines embryonnaires et adultes (**Gill & Wilcox, 2006**).

L'identification des variations polymorphiques SNPs de certains gènes, dans les populations du Maghreb arabo-berbère entre dans le cadre de la prévention primaire des risques liés au diabète. Elle vise à une meilleure compréhension des mécanismes moléculaires génétiques qui sont à l'origine de plusieurs pathologies (**Eddaikra, 2018**).

Plusieurs preuves sur l'association à l'échelle du génome du polymorphisme SNP (polymorphismes à nucléotide unique) de la NOX4 aux diabètes ont suggéré son implication dans la rétinopathie sévère et néphropathie diabétique (**Meng et al., 2018**), (**Lim et al., 2009**).

En effet, les SNPs du gène NOX4 jouent un rôle important dans notre susceptibilité à la maladie, la gravité de la maladie et la façon dont notre corps réagit au traitement (**He et al., 2018**). Ces données bibliographiques suggèrent que les variations polymorphiques mononucléotidiques 'SNPs' codant pour le gène NOX4 pourraient être impliquées dans l'apparition de diabète.

Dans ce contexte, nous nous sommes intéressées à identifier *in silico* les polymorphismes mononucléotidiques SNPs du gène NOX4 humain afin de comprendre la relation qui existe entre les polymorphismes de ce gène et le diabète. Cette étude « *in silico* » à fait recourt à l'utilisation d'une méthode bioinformatique pour identifier :

- Les différents variantes polymorphiques références snip « rs » qui peuvent changer l'expression du gène NOX4, la stabilité tridimensionnelle et le fonctionnement de la protéine qui code pour le gène NOX4 dans le cas du diabète.

Pour pouvoir couvrir tous les aspects de ce thème, la rédaction de notre mémoire a été structuré en Quatre chapitres :

- Dans le premier chapitre nous allons représenter les généralités sur le gène NOX4 et leur relation avec la physiopathologie du diabète.
- Le deuxième chapitre présente la méthodologie de travail en décrivant les outils bioinformatique et les différentes étapes utilisées pour répondre à notre problématique.
- Le troisième chapitre rapporte les résultats obtenus présenté par des figures et des tableaux.

Introduction

- Le quatrième et le dernier chapitre qui va rapporter la discussion des résultats et l'analyse des données.

Et enfin les résultats sont récapitulés sous forme d'une conclusion générale avec quelques recommandations pour les futurs travaux de recherches.

DONNEES
BIBLIOGRAPHIQUES

Chapitre1 : Données bibliographiques

1.1 Gène NOX4

1.1.1 Structure et localisation

Le gène Nox4 (NADPH OXYDASE 4 ; NADPH-OXYDASE RÉNALE ; RÉNOX) humain a été cartographié sur le chromosome 11q14.2, Positions 89167185 a 8968718, avec une longueur de 193.84kb contient 18 exons. Ce gène code pour un membre de la famille d'enzymes NOX qui fonctionne comme la sous-unité catalytique du complexe NADPH oxydase (**Guo & Chen, 2015**). La protéine codée est localisée dans les cellules non phagocytaires où elle agit comme un capteur d'oxygène et catalyse la réduction de l'oxygène moléculaire en diverses espèces réactives de l'oxygène (ROS) (**National Center for Biotechnology Information (NCBI), 2022**).

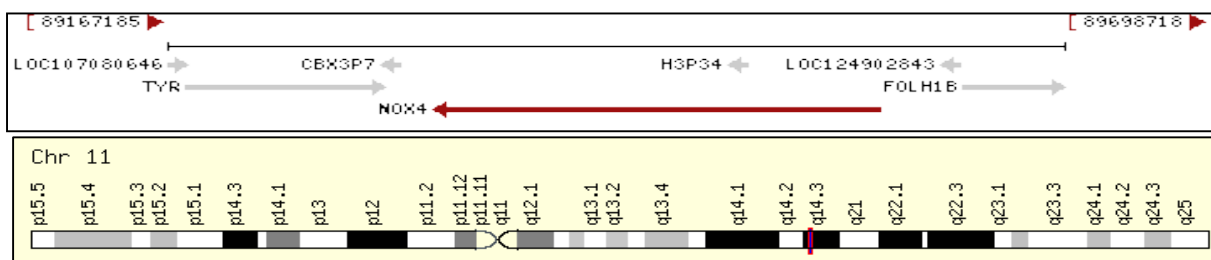


Figure 1: la localisation chromosomique de gène Nox4 (chromosome 11 q14.2)

<https://www.ncbi.nlm.nih.gov/gene/50507>

1.1.2 La protéine de Nox4

La séquence primaire du domaine catalytique de Nox4 comprend 578 acides aminés avec une masse moléculaire de 28 kDa, tandis que la protéine Nox4 complète est de 66,5 kDa (**National Center for Biotechnology Information (NCBI), 2022**). En général, Nox4 est localisée dans la mitochondrie et la membrane du réticulum endoplasmique avec une partie transmembranaire et extracytosolique, respectivement, tandis que dans les cellules endothéliales vasculaires humaines, il a été constaté que Nox4 avec le domaine transmembranaire est situé dans la membrane nucléaire. Comme les autres membres de la famille Nox, Nox4 comprend six hélices α de portée membranaire reliées par cinq boucles, à savoir les boucles A-E, et un domaine déshydrogénase (DH) qui contient les sites de liaison pour les deux groupes prosthétiques FAD et NADPH, ainsi qu'une extrémité C-terminale cytoplasmique relativement longue (**Guo & Chen, 2015**).

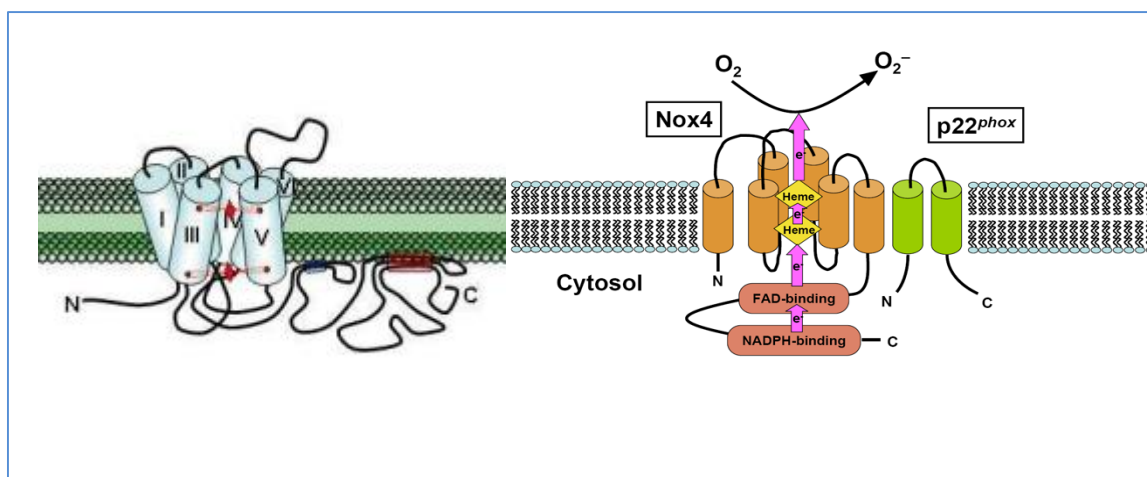


Figure 2: structure de la protéine du gène Nox4 (Nguyen, 2008).

1.1.3 Interaction entre Nox4 et P22phox

P22 (phox) est une protéine ubiquitaire codée par le gène CYBA (Cytochrome B-245 Alpha Chain) situé sur le bras long du chromosome 16 en position 24, contenant six exons et s'étendant sur 8,5 kb. P22 (phox) est un composant essentiel des nicotinamides adénine dinucléotide phosphate (NADPH) oxydases (NOX) génératrices de superoxyde. Il est associé à (NOX4 pour former le cytochrome b exprimé principalement dans les reines), P22 (phox) est également le partenaire de NOX4, une association qui est absolument requise pour qu'il soit fonctionnel (Stasia, 2016).

1.1.4 Le rôle physiologique de Nox4

Les Nox sont les principaux médiateurs de la production des ROS. L'anion O_2^- qui résulte de la réduction mono-électronique de l'oxygène, conduit à la formation successive, électron par électron, du peroxyde d'hydrogène (H_2O_2) et du radical hydroxyle ($^{\circ}OH$) (Stasia, 2016).

Le NADPH oxydase constitutive qui génère du superoxyde intracellulaire lors de la formation d'un complexe avec CYBA/p22phox. Régule les cascades de signalisation probablement par inhibition des phosphatases. Peut fonctionner comme un capteur d'oxygène régulant le canal potassique KCNK3/TASK-1 et l'activité HIF1A. Peut réguler la cascade de signalisation de l'insuline. Peut jouer un rôle dans l'apoptose, la résorption osseuse et l'activation médiée par les lipopolysaccharides du NF κ B. Peut produire du superoxyde dans le noyau et jouer un rôle dans la régulation de l'expression des gènes lors de la stimulation cellulaire.

<https://hmdb.ca/proteins/HMDBP08280> .

Tableau I: Distribution Tissulaire et Fonctions de la Nox4.

Distribution tissulaire	Fonctions	Physiopathologie
-Rein. -Cellules Endothéliales et Musculaires Lisses. -Cœur. -Pancréas. -Tissu Adipeux. -Placenta. -Muscle Squelettique. -Ovaire et Testicule. -Ostéoclastes, Fibroblastes, Astrocytes. -Nuerons.	-Senseur de l'O ₂ dans le rein. -Signalisations de l'insuline. -L'apoptose. -Interaction TLR4. -Inhibitions des phosphatases -La résorption osseuse. -Régule le canal potassique KCNK3/TASK-1.	-Diabète. -Maladies Neuro-dégénératives. -Obésité. -Maladies Cardio-vasculaires. -Athérosclérose. -Maladies Auto-immunes. -Cancer. -Maladies Inflammatoires. -Maladies Pulmonaires.

[\[https://www.ncbi.nlm.nih.gov/gene/50507\]](https://www.ncbi.nlm.nih.gov/gene/50507)

1.1.5 La famille de Nox

La famille des NADPH oxydases, Nox, chez l'homme, est composée de sept isoenzymes, cinq Nox et deux Duox (dual oxydase), [Nox1, Nox2, Nox3, Nox4, Nox5, Duox1, Duox2], La famille Nox Les NADPH oxydases (Nox) est largement connue pour leur fonction biologique de catalyse de l'oxygène pour produire des espèces réactives de l'oxygène (ROS) (Stasia, 2016).

Toutes les isoformes de NOX possèdent six hélices alpha transmembranaires avec des extrémités N et C cytosoliques. Elles sont exprimées et régulées différemment dans divers tissus et ont des localisations subcellulaires différentes, et même des produits ROS différents, c'est-à-dire super-oxyde contre peroxyde d'hydrogène (Stasia, 2016). NOX1, NOX2, et NOX5 semblent produire principalement du superoxyde NOX4, principalement du H₂O₂. Toutes les isoformes de NOX ont été signalées comme se liant à une ou plusieurs protéines membranaires et/ou cytosoliques. p22phox apparaît comme un partenaire de liaison général pour les NOX1-4 dans la membrane (Altenhöfer et al., 2012).

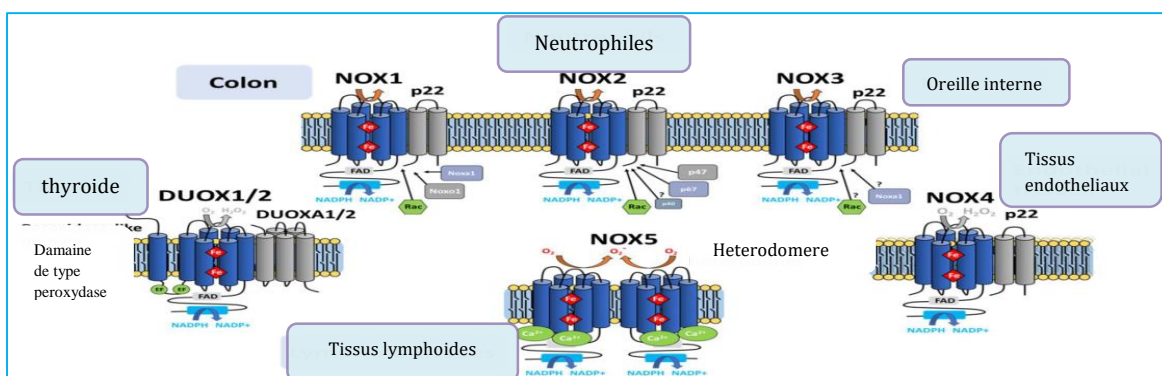


Figure 3: structure de la famille des NADPH Oxydase.

<https://www.mdpi.com/2076-3921/10/6/890>

1.1.6 Isoformes de Nox4

Il a été suggéré que Nox4 possède 4 autres isoformes (Nox4A, Nox4B, Nox4C, Nox4D, Nox4E) (Goyal, 2005).

La Nox4B diffère de la Nox4A par l'épissage de l'exon 14 qui entraîne une délétion du premier domaine de liaison de la NADPH, ce qui entraîne l'absence du premier site de liaison du NADPH dans la protéine. La Nox4C est le résultat d'un épissage de l'exon 9-11, qui génère un décalage de cadre dans l'ARNm, et donc un nouveau codon stop au début de l'exon 12. Par conséquent, il manque à la protéine Nox4C l'extrémité C-terminale complète, y compris les domaines de liaison FAD et NADPH. L'effet inhibiteur similaire sur la génération de ROS entre les variantes B et C indique que le premier site de liaison NADPH semble avoir un impact important sur l'activité de la NADPH oxydase et ne peut être entièrement compensé par les autres sites de liaison NADPH et FADPH (Altenhöfer et al., 2012).

Nox4D et Nox4E sont tous deux une conséquence de l'épissage des exons 3-11, tandis que la variante E présente un épissage supplémentaire de l'exon 14. Par conséquent, Nox4D possède uniquement le premier domaine transmembranaire et manque en outre de sites de liaison du FAD et du NADPH, tandis que Nox4E n'a que le premier domaine transmembranaire et manque en outre d'un site de liaison du NADPH, comme la variante B. (Lambeth, 2004) (Altenhöfer et al., 2012).

1.2 Polymorphismes

Le polymorphisme génétique est défini comme l'hérédité d'un trait contrôlé par un seul locus génétique à deux allèles, dans lequel l'allèle le moins commun a une fréquence d'environ 1 % ou plus (Ismail & Essawi, 2012).

Le polymorphisme génétique est une différence de séquence d'ADN entre les individus, les groupes ou les populations. Les sources comprennent les polymorphismes mononucléotidiques (SNP), les séquences répétées, les insertions, les délétions et les recombinaisons (par exemple, un polymorphisme génétique peut donner lieu à des yeux bleus par rapport à des yeux bruns, ou à des cheveux lisses par rapport à des cheveux bouclés). Les polymorphismes génétiques peuvent être le résultat de processus aléatoires ou peuvent avoir été induits par des agents externes tels que des virus ou des radiations (Ismail & Essawi, 2012).

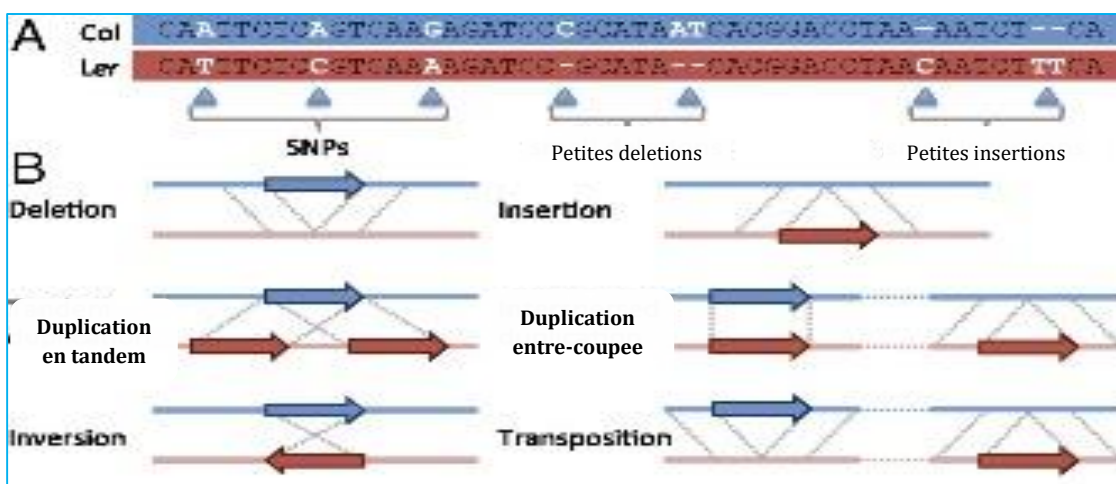


Figure 4:Types de polymorphisme (Qi et al., 2014).

Les polymorphismes résultent de mutations. La mutation peut être due à un passage d'un type de nucléotide à un autre, une insertion ou une délétion, ou un réarrangement de nucléotides. Une fois formé, un polymorphisme peut être hérité comme n'importe quelle autre séquence d'ADN, ce qui permet de suivre son héritage de parent à enfant (Ismail & Essawi, 2012).

1.2.1 Polymorphisme d'un seul nucléotide.

polymorphisme mononucléotidiques (SNP), variation d'une séquence génétique qui n'affecte qu'un seul des éléments constitutifs de base - l'adénine (A), la guanine (G), la thymine (T) ou la cytosine (C) - dans un segment d'une molécule d'ADN (Britannica, 2019).

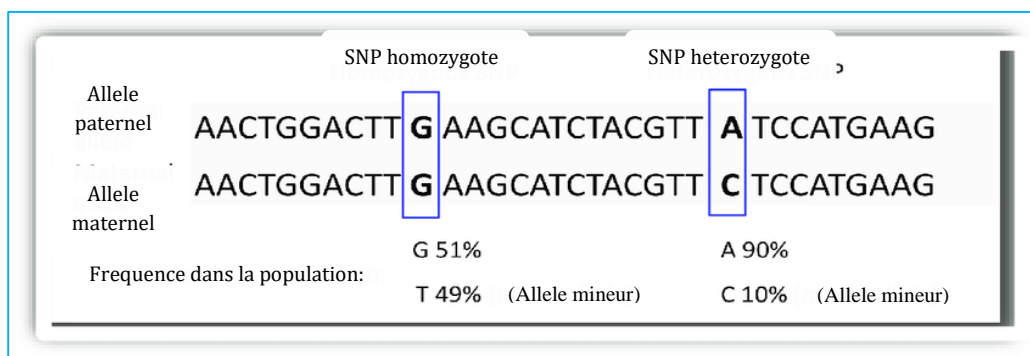


Figure 5: Polymorphisme de nucléotide unique (SNP) dans le génome individuel et dans la population (Popova et al., 2013).

Les polymorphismes mononucléotidiques (SNP) se trouvent dans toutes les 200 à 300 paires de bases du génome humain et servent de marqueurs génétiques (Lee et al., 2005). Environ 0,5 million de SNP sont présents dans la région codante du génome humain (Rajasekaran et al., 2008). La substitution d'acides aminés dans les régions conservées de la protéine peut exercer un effet sur la structure, la stabilité et la fonction de la protéine (George Priya Doss et al., 2008), (Jia et al., 2014).

Les polymorphismes (SNP), qui sont des substitutions d'un seul nucléotide à une position spécifique du génome survenant chez plus de 1 % de la population humaine, sont au cœur de ces études, car ils constituent la majorité des variations génétiques chez l'homme. Une identification fiable des SNP pathogènes devrait permettre un diagnostic précoce, des traitements personnalisés et la conception de médicaments ciblés (Bhatti et al., 2006) (Jia et al., 2014).

Chapitre 1 : Données bibliographiques

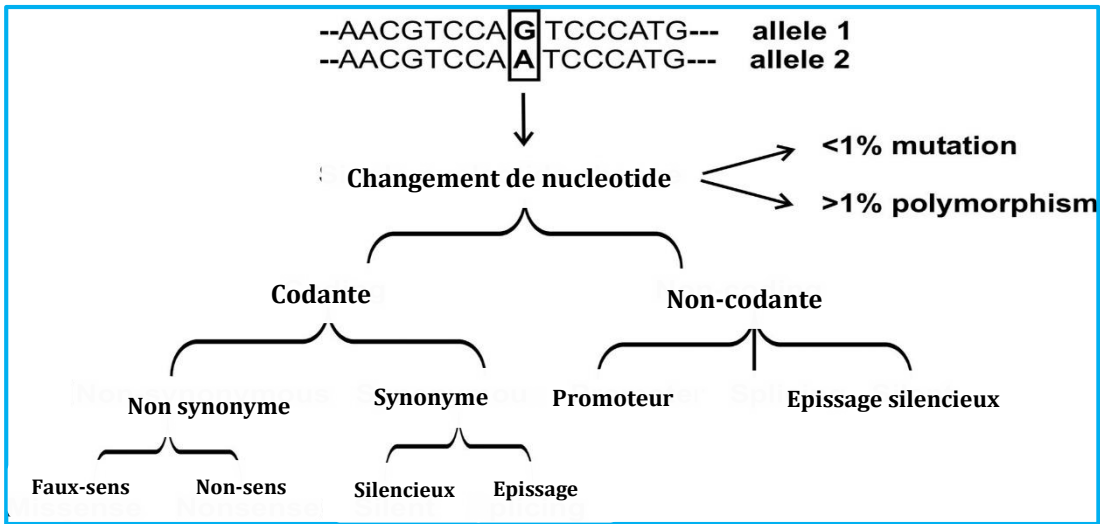


Figure 6: Modifications génomique d'un seul nucléotide (Pollex & Hegele, 2007)

Les polymorphismes synonymes sont ceux qui n'ont aucun effet sur l'organisme et sont dits sélectivement neutres car la substitution ne provoque aucun changement d'acide aminé dans la protéine produite. C'est ce qu'on appelle aussi une mutation silencieuse. Une substitution non synonyme entraîne une altération de l'acide aminé codé. Une mutation faux-sens modifie la protéine en provoquant une modification du codon. Une mutation non-sens entraîne un codon de terminaison mal placé. La moitié de tous les SNP de séquence codante entraînent des changements de codons non synonymes (<http://genomics.phrma.org/lexicon>)

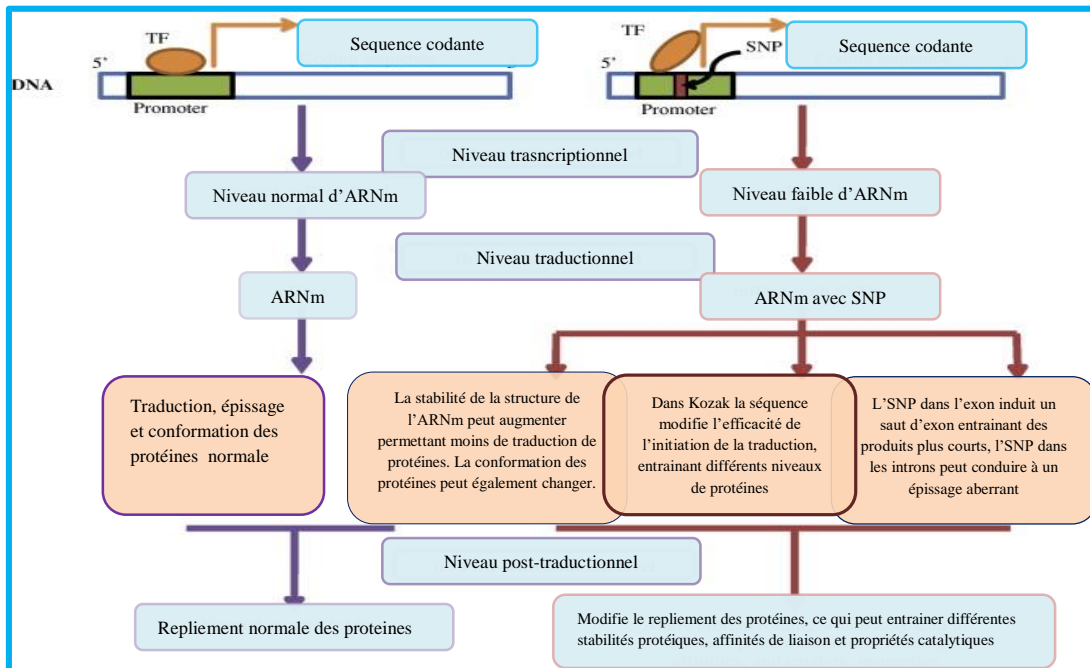


Figure 7: Conséquences possibles qui peuvent résulter des SNP (Fareed & Afzal, 2013).

1.2.2 Polymorphisme du gène NOX4

En utilisant la séquence du gène standard Nox4 humain à des fins de comparaison, environ 2385 sites de polymorphisme nucléotidique unique (SNP) ont été trouvés dans la région d'ADN génomique de Nox4 et 45 sites SNP dans la région codante du gène. L'un de ces sites a conféré une mutation de décalage de cadre. Les variations de séquence décrites pour Nox4 peuvent affecter la réplication ou la transcription des gènes, et éventuellement la fonction enzymatique et la production de ROS. Par conséquent, il est possible qu'ils contribuent à l'apparition et au développement de la maladie. Cependant, les corrélations entre les polymorphismes génétiques de Nox4 dans différentes populations et les variations liées à la maladie dans l'activité et/ou l'expression de Nox4 sont mal comprises. De telles associations seraient facilitées par des études génétiques étudiant la relation entre l'expression et la fonction de Nox4 dans les maladies liées à des niveaux élevés de ROS (**Chen et al., 2012**).

NOX4 a été signalé comme étant impliqué dans de multiples pathogénèses, notamment l'hypertension, le diabète sucré, les maladies coronariennes, les maladies du foie, les maladies neurodégénératives, les maladies métaboliques et les maladies cardiovasculaires.

1.3 Diabète

Le diabète sucré est un trouble métabolique complexe dont la principale caractéristique clinique et diagnostique est l'hyperglycémie (**Inzucchi et al., 2012**).

Le diabète a atteint des proportions épidémiques, touchant environ 387 millions de personnes dans le monde. Au cours des 20 prochaines années, sa prévalence devrait doubler et plus d'un demi-milliard de personnes seront touchées (**International Diabetes Federation, 2015**). La prévalence régionale estimée du diabète varie de 5,1 % en Afrique à 11,4 % en Amérique du Nord et dans les Caraïbes, avec plus de 75 % des sujets vivant dans des pays à revenu faible ou intermédiaire (**International Diabetes Federation, 2015**).

Le diabète (en particulier le DT2) double approximativement le risque d'un large éventail de maladies cardiovasculaires, y compris les maladies coronariennes et les accidents vasculaires cérébraux.⁶ De plus, le DT2 est également associé à un large éventail de maladies non vasculaires, notamment le cancer, les troubles du système mental et nerveux, infections et maladies du foie (**Rao Kondapally Seshasai et al., 2011**). De même, le diabète sucré de type 1 (T1DM ; ~10 % de tous les cas de diabète) est associée à un risque accru de complications vasculaires et non vasculaires (**Lind et al., 2014**) (**Livingstone et al., 2015**). Par conséquent, une meilleure compréhension des mécanismes qui entraînent l'hyperglycémie pourrait aider à identifier des cibles thérapeutiques potentielles pour freiner le diabète et ses complications associées (**Zaccardi et al., 2016**).

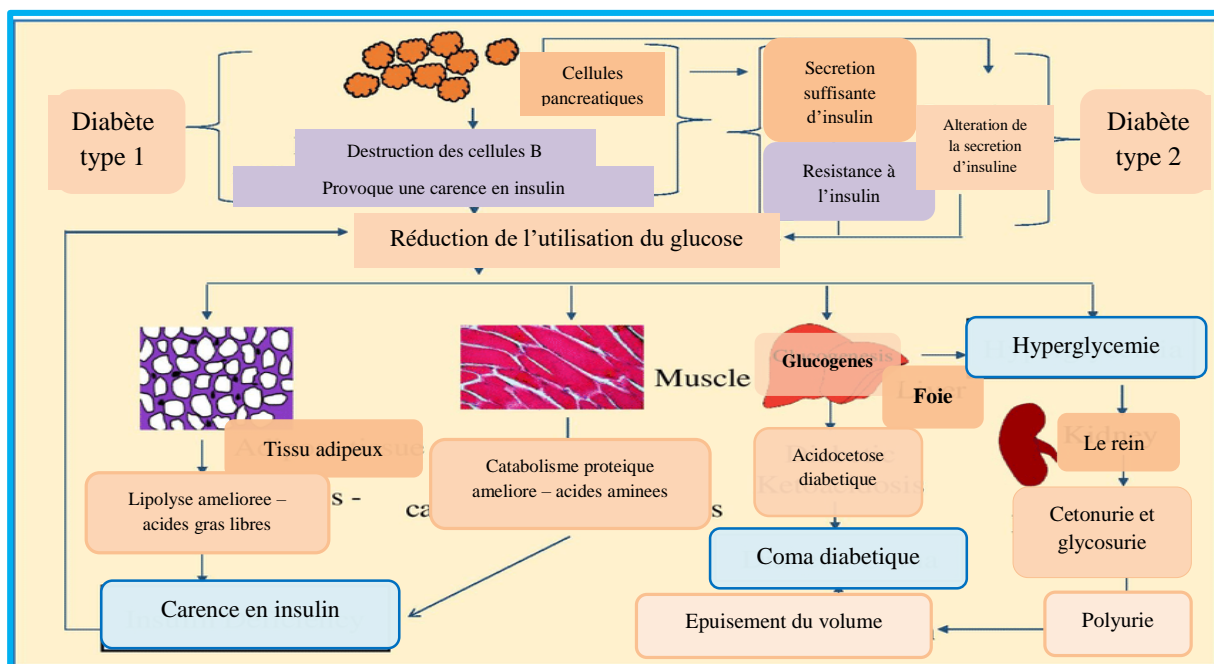


Figure 8: Physiopathologie du diabète sucré type 1 et 2. (Wei et al., 2021)

1.3.1 Diabète de type 1

Il est maintenant bien reconnu que le DT1 est une maladie auto-immune caractérisée par la destruction des cellules β pancréatiques productrices d'insuline (Atkinson et al., 2014). L'immunité humorale et cellulaire est impliquée dans la pathogenèse du DT1. Les lymphocytes T (qui mûrissent dans le thymus et jouent un rôle central dans l'immunité à médiation cellulaire) sont prédominants dans les lésions des îlots, avec des concentrations plus faibles d'autres cellules immunologiques, telles que les macrophages, les lymphocytes B et les plasmocytes (Willcox et al., 2009). La présence de l'immunité humorale, en revanche, a été reconnue il y a plus de 40 ans, lorsque des auto-anticorps dirigés contre les îlots pancréatiques ont été détectés chez des sujets atteints de DT1 (Bottazzo et al., 1974).

1.3.2 Diabète de type 2

Il a fallu attendre les années 1980 et 1990, avec la standardisation et l'évolution des techniques de mesure de l'élimination du glucose par l'insuline (Shen et al., 1970), (DeFronzo et al., 1979) que les chercheurs ont réalisé des expériences indépendantes confirmant la présence d'une résistance à l'insuline dans le diabète de « maturité » (DeFronzo et al., 1979) (Reaven, 1983).

Les sujets à risque de DT2 (sujets obèses et apparentés au premier degré) présentent un état initial d'insulino-résistance compensée par une hypersécrétion d'insuline des cellules β (hyperinsulinémie). Cependant, une telle réserve « fonctionnelle » pancréatique est finalement incapable de faire face à la sécrétion d'insuline requise. Au moment où le diabète est diagnostiqué, les cellules β ne sont plus capables de sécréter suffisamment d'insuline, ce qui

entraîne une hyperglycémie manifeste (hypoinsulinémie hyperglycémique) (**Jallut et al., 1990**).

1.3.3 Complications du diabète et NOX4

Le diabète est de plus en plus reconnu comme la principale cause d'insuffisance rénale chronique, et la majorité des patients évoluent vers une insuffisance rénale terminale (IRT) nécessitant une dialyse à vie ou une greffe de rein (**Rosen et al., 2018**). La néphropathie diabétique (DN) est maintenant cliniquement identifiée par une albuminurie protéique progressivement élevée et une baisse subséquente du taux de filtration glomérulaire (**Rosen et al., 2018**).

Dans le rein, l'isoforme majeure Nox4 est responsable de la génération de ROS, il est impliquée dans les effets néfastes de l'HG qui contribuent de manière significative aux complications microvasculaires associées au diabète (**Rosen et al., 2018**). Les données de modèles animaux diabétiques ont montré que l'hyperglycémie (HG) améliorent de manière significative l'accumulation de matrice extracellulaire dans les cellules mésangiales (**Shi et al., 2020**).

Dans une étude menée par Quian Shi et al en 2020, ils ont montré que HG favorise l'activation / l'expression de HuR et que HuR est nécessaire pour l'expression de la protéine Nox4 induite par Hyperglycémie (HG) / la traduction de l'ARNm, la génération de ROS et les lésions fibrotiques MC ultérieures (**Shi et al., 2020**).

La protéine HuR (Human antigen R) appartient à la famille protéique ELAV, qui joue un rôle essentiel dans la fonction et dans la destinée des ARN sur lesquels elle se fixe, les protéines HuR sont impliquées dans diverses étapes de la régulation post-transcriptionnelle à la fois dans le noyau et dans le cytoplasme (**Himman et Lou 2008**).

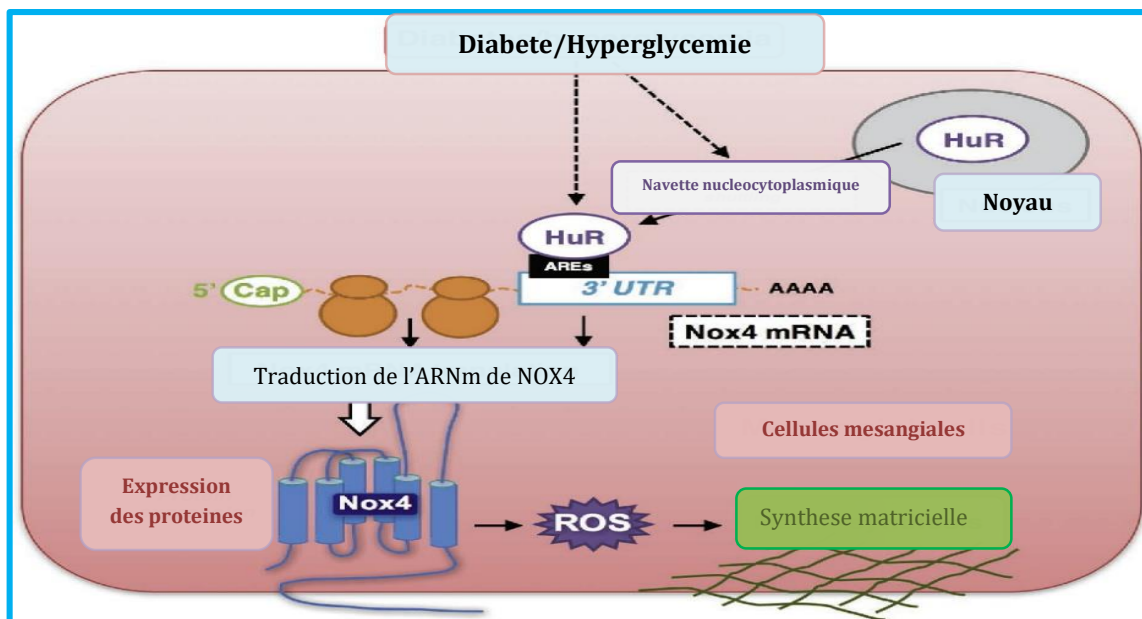


Figure 9: Interaction HuR/NOX4 dans la néphropathie diabétique. (Shi et al., 2020).

La rétinopathie diabétique (RD) est l'une des principales causes de perte de vision et de cécité dans le monde, avec ≈ 93 millions de personnes atteintes de RD et 29 millions de personnes présentant une RD menaçant la vision (Yau et al., 2012). Le stress oxydatif, la production excessive d'espèces réactives de l'oxygène (ROS), est un contributeur majeur aux complications du diabète, y compris la RD (Marco et al., 2015). L'excès de ROS favorise la production de facteurs angiogéniques et inflammatoires (Schröder, 2019), (Forrester et al., 2018) qui causent l'augmentation de la vasculopathie et de l'inflammation qui se développe dans la RD (Li et al., 2010) (Mohamed et al., 2012). Nox (NADPH oxydase) présente un intérêt en tant que cible de traitement de la RD car la seule fonction de cette famille d'enzymes est la production de ROS. Il existe 7 isoformes connues de Nox dont Nox1, Nox2 et Nox4 sont impliquées dans la pathogenèse du dysfonctionnement vasculaire et de l'inflammation dans diverses maladies rétinienne (Peng et al., 2019).

Dans des conditions hyperglycémiques, les isoformes NOX (NOX1-4) sont activées pour produire du superoxyde (O_2^-) ou du peroxyde d'hydrogène (H_2O_2). L'excès d'espèces réactives de l'oxygène (ROS) provoque des lésions oxydatives de la rétine (telles que la sénescence, l'apoptose, l'hyperperméabilité et les dommages mitochondriaux) ou agit comme des molécules de signalisation pour initier la néovascularisation rétinienne, qui contribuent toutes à la pathogenèse de la rétinopathie diabétique (Peng et al., 2019).

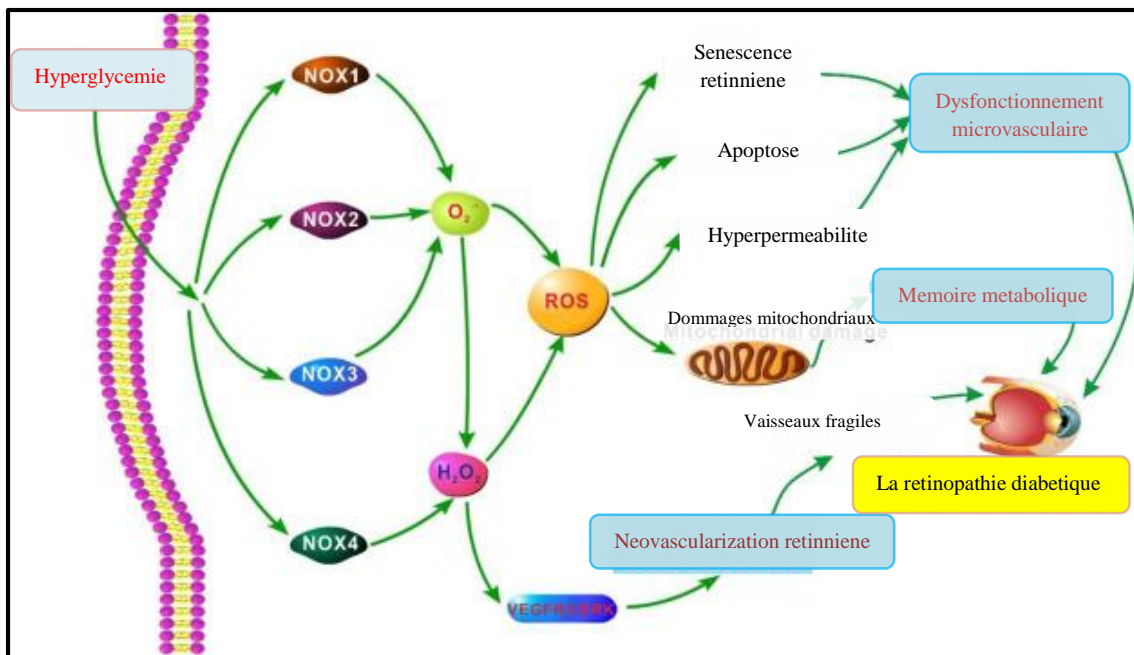


Figure 10:Rôle des isoformes NOX dans la pathogenèse de la néphropathie diabétique. (Peng et al., 2019)

*MATERIEL ET
METHODES*

Chapitre 2 : Matériel et méthodes

La partie pratique de notre étude s'est étalée sur une période de cinq (5) mois (Février-juin 2022). Notre objectif principal est d'identifier le polymorphisme du gène NOX4 et son association à l'apparition du diabète chez l'homme par méthode « *in Silico* » en utilisant différents outils de bio-informatique avec différentes bases de données génomiques et protéiques.

2. Matériel et Méthodes

2.1 Matériel

2.1.1 Matériel biologique

Les identifiants (numéros d'accessions) et les sources des séquences nucléotidiques transcrites du gène « NOX4 » Reference sequence (Refseq) Transcriptions et séquençage de l'ARN (RNAseq) sont représentés dans le tableau ci-dessous (tableau II). Ces données sont stockées dans des bases de données à accès libre NCBI.

Tableau II: Description du gène Nox4.

Gène	Nombre de pair de bases	Nombre d'exon	Emplacement	Refseq NCBI	Gène ID	Refseq protéique
NOX4	4 269 pb	18-26	11q14.3	NM_016931.5	50507	NP_058627.2

2.1.2 Matériel non biologique

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 bio-informatiques sont répertoriés dans les tableaux III et IV.

Tableau III: Les Bases des données utilisé dans la recherche

Bases des données	But d'utilisation	url
NCBI	L'analyse des données moléculaires et génomiques	https://www.ncbi.nlm.nih.gov/
UCSC	Etude de Gene +Recherches SNPs	https://genome.ucsc.edu/
Ensemble	Recherche bibliographique + recherche de séquence +SNPs+ Les Exons	https://www.ensembl.org/index.html
PubMed	Recherche bibliographique	https://pubmed.ncbi.nlm.nih.gov/
GeneCard	Donne accès à 150 autres sources de recherche	https://www.genecards.org/
Uniprot	La séquence protéique	https://www.uniprot.org/
SMART	L'identification et l'analyse des domaines protéiques	http://smart.embl-heidelberg.de/
SWISS-Model	Structure protéique+ligand	https://swissmodel.expasy.org/

Chapitre 2 : Matériel et Méthodes

Blast	Régions de similarité entre des séquences	https://blast.ncbi.nlm.nih.gov/Blast.cgi
ORF-Finder	Recherche les cadres de lecture ouverts (ORF) dans la séquence d'ADN	https://www.ncbi.nlm.nih.gov/orffinder/
SNPinfo	Recherche des SNPs+ récupération des informations sur l'effet du rs dans la séquence ADN	https://snpinfo.niehs.nih.gov/
SIFT	Identification des SNP délétères	https://sift.bii.a-star.edu.sg/
Polyphen-2	Identification des SNP délétères	http://genetics.bwh.harvard.edu/pph2/
MutPred-2	Impact sur le domaine protéique	http://mutpred.mutdb.org/
LitVar	Récupération des informations de littérature relative aux variantes	https://www.ncbi.nlm.nih.gov/CBBresearch/Lu/Demo/LitVar/#!?query
PHD-SNP	Prédire l'effet possible du changement d'acides aminés	https://snps.biofold.org/phd-snp/phd-snp.html
SNAP2	Prédire l'effet possible du changement d'acides aminés	https://www.rostlab.org/services/SNAP/
I-mutant	L'effet sur la stabilité de la protéine.	https://folding.biofold.org/i-mutant/i-mutant2.0.html
HOPE Project	Visualisation 3D de la protéine mutante et wild-type	https://www3.cmbi.umcn.nl/hope/input/#

Tableau IV: Les bases des données et les logiciels utilisés dans le dorking de la molécule.

Basse des données ou logiciel utilisé	But d'utilisation	Site web
PubChem	Recherche des ligands	https://pubchem.ncbi.nlm.nih.gov/
CASTp	Recherche des sites actifs	http://sts.bioe.uic.edu/castp/index.html?j_62877290304dc
Open Babel	Conversion des fichiers	Logiciel
PDB	La forme moléculaire en 3D	https://www.rcsb.org/
Autodock tools	Modélisation moléculaire	Logiciel
PyMOL	Modélisation moléculaire	Logiciel
Discovery Studio Visualizer	Traitement des résultats	Logiciel

2.2 Méthodes

La méthode que nous avons utilisée pour l'identification des polymorphismes du gène du Nox4 est une méthode « *in Silico* » représentée par l'organigramme ci-après (figure 11)

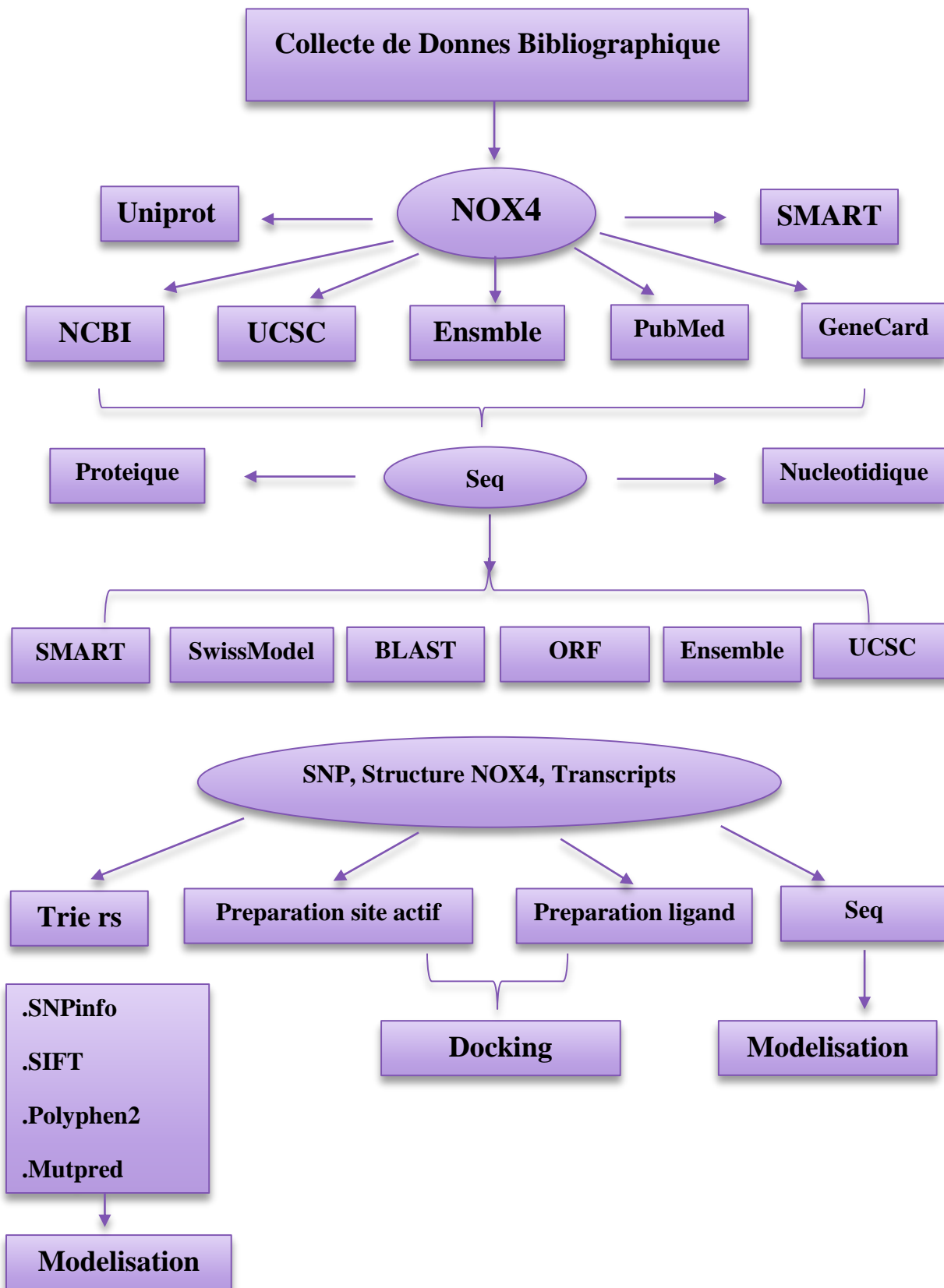


Figure 11: méthode d'étude et d'analyse du polymorphisme du gène NOX4.

2.2.1 Etude de gène et recherche des SNP

2.2.1.1 NCBI

Principe Le National Center for Biotechnology Information (NCBI) a été créé en novembre 1988 à la National Library of Medicine (NLM) aux États-Unis, La mission du NCBI est de développer de nouvelles technologies de l'information pour aider à la compréhension des phénomènes moléculaires et génétiques fondamentaux la santé et la maladie , cette source de données nous permet d'initier la récolte d'information sur le gène « NOX4» (**Jenuth, 2000**)

Méthode

Tous d'abord il faut consulter le site (<https://www.ncbi.nlm.nih.gov/>) et insérer le mot clé caractérisé par le symbole du gène « **NOX4** », en s'assurant que le type d'information que l'on cherche est sur « **GENE** », puis et on clique sur « **search** ».

La recherche nous donne une multitude d'ID et nom de gène avec différents, ont choisi le nom du gène de l'espèce homo sapiens (humain). Une autre page s'ouvre, indiquant toutes les informations sur le gène NOX4 avec les divers outils pour la suite des recherches (Figure 12),(Tableau V). (annexe1fig 54)

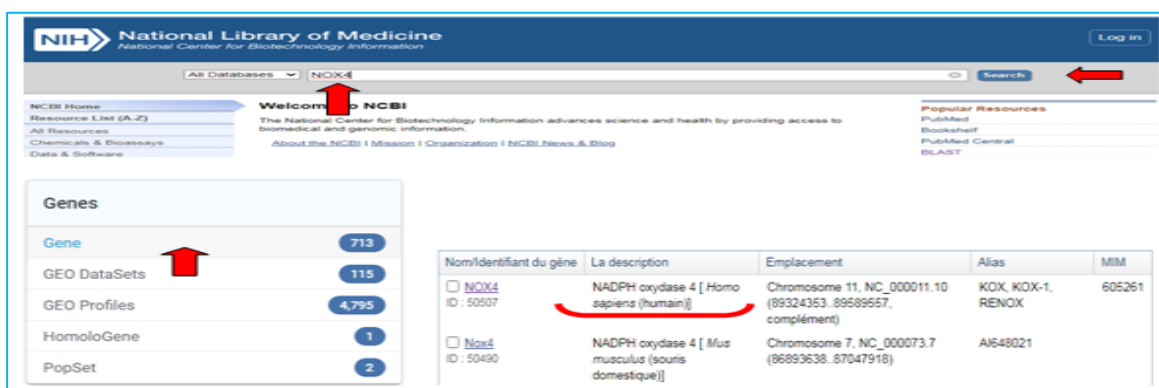


Figure 12: Les étapes de recherche sur NCBI.

Tableau V: Références GenBank NCBI du gène NOX4.

NCBI	
Nom	NADPH Oxydase-NOX4
Synonyme	KOX, KOX-1, Renox
Nature	Nucleotide
Organisme	Homo sapiens
Séquence de référence	NM_016931.5
Paires de bases	4269
Cytogénétique localisation	11q14.3
Gene groupe	Enzyme catalytique
Le nombre d'exons	26
N° des transcrites	8+1 non codant
N° des SNPs	103124

2.2.1.2 Base de données PubMed

Principe PubMed est une ressource gratuite soutenant la recherche et la récupération de la littérature biomédicale et des sciences de la vie dans le but d'améliorer la santé, à la fois globalement et personnellement. La base de données PubMed contient plus de 34 millions de citations et de résumés de littérature biomédicale. Il n'inclut pas les articles de revues en texte intégral. Accessible au public en ligne depuis 1996, PubMed a été développé et est maintenu par le [National Center for Biotechnology Information \(NCBI\)](https://pubmed.ncbi.nlm.nih.gov/), à la [US National Library of Medicine \(NLM\)](https://pubmed.ncbi.nlm.nih.gov/), situé dans les [National Institutes of Health \(NIH\)](https://pubmed.ncbi.nlm.nih.gov/).

Les citations dans PubMed proviennent principalement des domaines de la biomédecine et de la santé, et des disciplines connexes telles que les sciences de la vie, les sciences du comportement, les sciences chimiques et la bio-ingénierie (Misra & Ravindran, 2022).

Méthode

Pour préparer la recherche tous d'abord l'accès au site PubMed : <https://pubmed.ncbi.nlm.nih.gov/>, Lancier la rechercher des documents par l'insertion des mots clés (polymorphisme, NOX4, diabète) dans la barre de recherche, l'apparition d'une nouvelle page des résultats et Obtenir des documents (Article). En suite en va Filtrer et Sélectionner des documents pertinents au sujet.

The image shows two screenshots of the PubMed website. The top screenshot shows the search page with the search bar containing the text "Polymorphism, NOX4, Diabete" and a green "Search" button. Below the search bar, there is a brief description of PubMed: "PubMed® comprises more than 34 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full text content from PubMed Central and publisher web sites." The bottom screenshot shows the search results page. The search bar still contains "Polymorphism, NOX4, Diabete". Below the search bar, there are options for "Advanced", "Create alert", and "Create RSS". The results are sorted by "Best match" and there are 9 results. The first result is titled "NOX4 rs11018628 polymorphism associates with a decreased risk and better short-term recovery of ischemic stroke." and the second result is titled "Association between the CYBA and NOX4 genes of NADPH oxidase and its relationship with metabolic syndrome in non-alcoholic fatty liver disease in Brazilian population." The results page also includes a "RESULTS BY YEAR" bar chart showing a peak in 2022, and options for "TEXT AVAILABILITY" (Abstract, Free full text, Full text) and "ARTICLE ATTRIBUTES".

Figure 13: Les étapes de recherche sur PubMed.

2.2.1.3 Outil GeneCards

Principe GeneCards est une base de données interrogeable et intégrative qui fournit des informations complètes et conviviales sur tous les gènes humains annotés et prédits. La base de connaissances intègre automatiquement des données centrées sur les gènes provenant d'environ 150 sources Web, y compris des informations génomiques, transcriptomiques, protéomiques, génétiques, cliniques et fonctionnelles (Safran et al., 2003).

Méthode

En consultant le lien (<https://www.genecards.org/>), nous introduisons le mots clé « NOX4 » sur la page d'accueil de GeneCards puis on clique sur « Rechercher » pour lancer la recherche.

La page « Résultat de la recherche » affiche une liste formée de mini-cartes, indiquant le symbole du gène, la description, la catégorie, le GCID et un score de pertinence. Pour ouvrir la minicarte, cliquez sur le plus à gauche du symbole du gène sur la minicarte appropriée.

Sélectionné le gène NADPH Oxydase 4, ensuite une nouvelle page apparait qui donne un aperçu des informations disponibles au niveau du gène, ainsi que le cartographique génétique et la famille la fonction les domaines de la protéine l'empalment subcellulaire les voie de signalisation et les protéines interagissant avec le gène NOX4 (figure 14), voir annexe (annexe5, fig61)

The screenshot displays the GeneCards website interface. At the top, there is a navigation menu and a search bar containing the text 'NOX4'. Below the search bar, the website's header and a 'GeneCards Data Licensing' banner are visible. The main content area shows search results for 'NOX4', with a table listing 25 of 3,744 results. The table has columns for Symbol, Description, Category, GPF5, GC id, and Score. The first result is NOX4, NADPH Oxidase 4, with a score of 52.79. Red arrows highlight the search bar and the first result in the table.

Symbol	Description	Category	GPF5	GC id	Score
NOX4	NADPH Oxidase 4	Protein Coding	40	GC11M089324	52.79
NOX1	NADPH Oxidase 1	Protein Coding	38	GC03M100843	8.03
CYBB	Cytochrome B-245 Beta Chain	Protein Coding	43	GC03P037780	7.89
TYR	Tyrosinase	Protein Coding	45	GC11P089177	6.99
CYBA	Cytochrome B-245 Alpha Chain	Protein Coding	43	GC11M088643	6.60
TGFB1	Transforming Growth Factor Beta 1	Protein Coding	46	GC19M041301	5.55
FOLH1B	Folate Hydrolase 1B	Protein Coding	26	GC11P089638	5.23
ENSG00000255429	Novel Transcript	RNA Gene	6	GC11M089546	4.97
NOX5	NADPH Oxidase 5	Protein Coding	34	GC15P114661	4.49

Figure 14: Les étapes de recherche sur GeneCards.

2.2.1.4 Outil LitVar

Principe LitVar normalise la requête pour trouver la variante la mieux adaptée dans sa base de données, ainsi que des résultats de désambiguïsation alternatifs. Ensuite, LitVar renvoie une liste de publications contenant cette variante, classée par date de publication. Ce processus présente deux caractéristiques principales. Premièrement, pour chaque résultat, LitVar renvoie un ou plusieurs extraits mettant en évidence la variante recherchée ainsi que d'autres entités (par exemple, des maladies, des produits chimiques et d'autres variantes) qui apparaissent le plus souvent dans la même phrase que la variante recherchée. Elles sont détectées dans les publications lors du prétraitement et liées à chaque phrase dans notre base de données. Ceci est particulièrement utile pour détecter les relations potentielles (Allot et al., 2018).

Méthode

-En consultant le lien litVar <https://www.ncbi.nlm.nih.gov>, on introduit le nom de gène « NOX4 » sur la page d'accueil de LitVar dans la barre de recherche puis on clique sur le bouton « Rechercher » pour lancer la recherche. Nous avons par la suite, l'apparition de nouvelle page des résultats (figure 15)



Figure 15: la base des données LitVar.

2.2.1.5 SNPinfo

Principe : Le serveur SNPinfo est un ensemble de divers outils de sélection basés sur le Web, notamment Gene pipe, Genome pipe, Linkage pipe, Taq SNP, Func Pred, SNPseq, qui ont été utilisés pour sélectionner des SNP fonctionnels pour des études d'association génétique (<http://snpinfo.niehs.nih.gov/>). Cet algorithme prédit l'impact fonctionnel des SNP codants et non codants sur la base GWAS donnant des informations sur les gènes candidats (Xu & Taylor, 2009). L'organigramme des étapes de notre recherche est représenté par la figure 17.

La recherche se fait soit par 'SNP Function Prediction (FuncPred)' en insérant la liste des rs, soit par 'SNP Information in DNA Sequence (SNPSeq)' et par la recherche des rs un par un (figure 16).

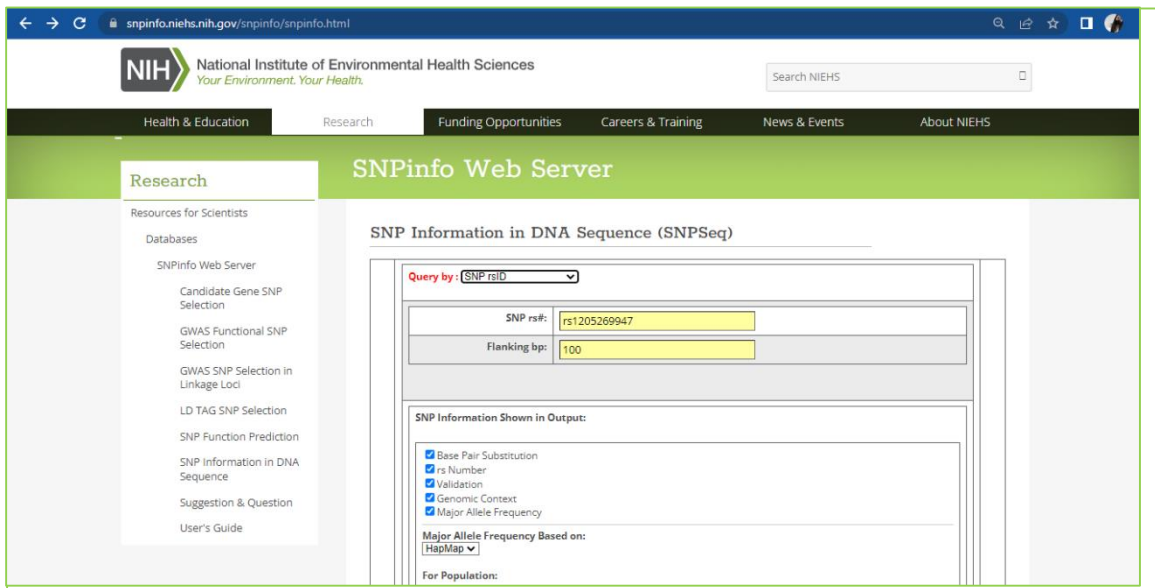


Figure 16: Les étapes de recherche sur SNPinfo.

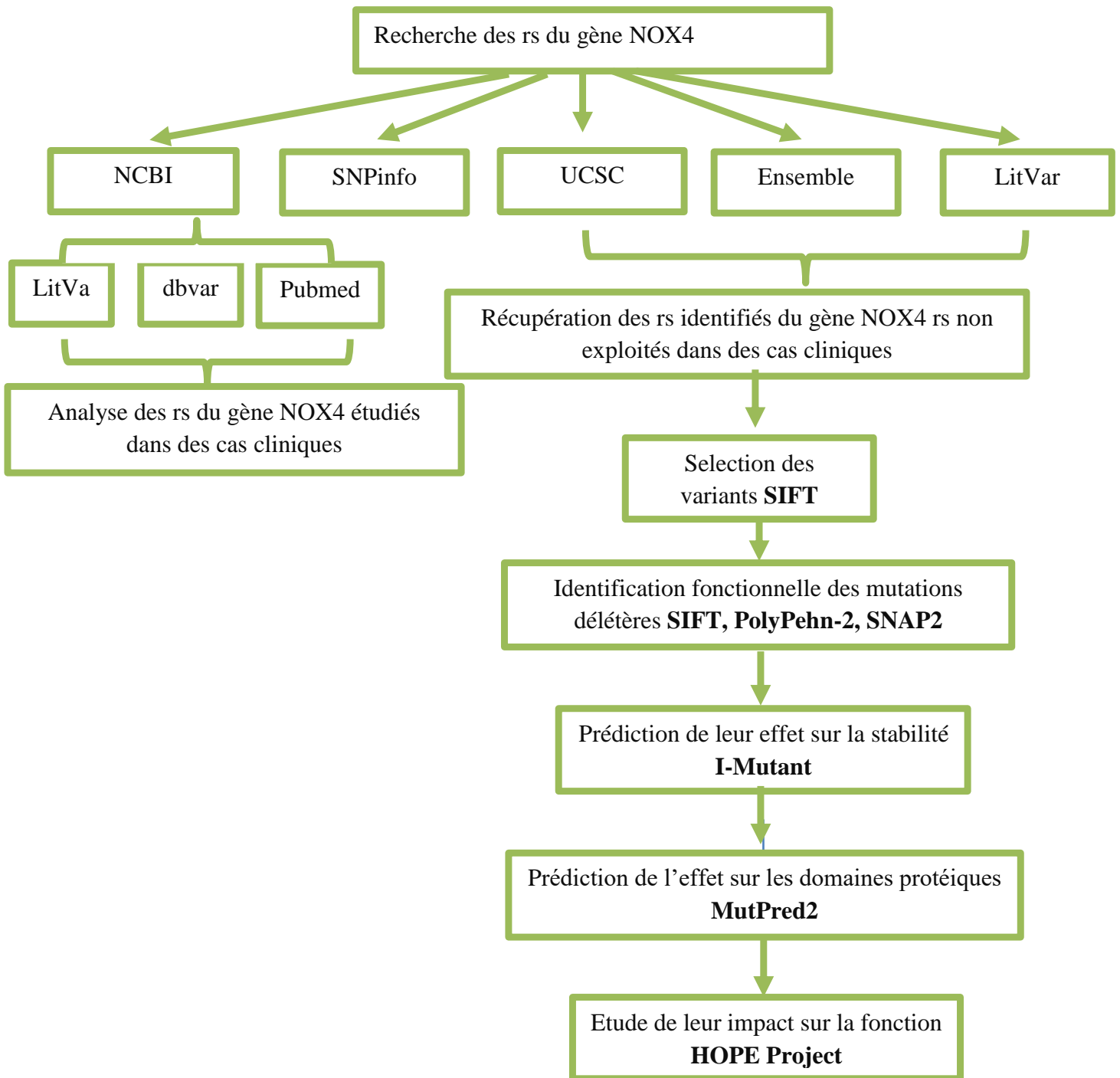


Figure 17: Méthode de recherche et d'étude des rs du gène NOX4, exploités et non exploités dans des cas cliniques « études cas-témoins »

2.2.2 Identification des rs SNPs du gène de NOX4

2.2.2.1 Outil UCSC

Principe Le navigateur de génome UCSC est un navigateur de génome en ligne et téléchargeable hébergé par l' Université de Californie à Santa Cruz (UCSC) Le 22 juin 2000.Ce navigateur affiche les contigs d'assemblage et les lacunes, les alignements d'ARNm et de marqueurs de séquences exprimées, les prédictions de gènes multiples, les homologies entre espèces, les polymorphismes de nucléotides simples, les sites marqués par des séquences, les données d'hybridation de rayonnement, les répétitions de transposons..etc (**Simon & Gomez, 2021**)

Méthode :

Le lien <https://genome.ucsc.edu> va nous guider directement vers la plateforme d'UCSC, ou on clique sur « **Genome Browser** », ce qui va nous menez a une nouvelle page ou on met l'espèce « **homo sapiens** », l'assemblage humain « **dec.2013(GRCh38/hg38)** » et le nom du gène « **NOX4** ». Par la suite une multitude de résultats apparaissent, on va 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**» dans : Common SNPs \ ALL SNPs \ Flagged SNPs pour variation, et Old UCSC genes \ UCSC all événements pour Gènes. Puis on clique sur actualiser (refresh). Ensuite, sur « **Tools** » on tape sur « **table browser** » pour trier les « rs », puis « **output** », « **done with selections** » et « **send query to galaxy** »

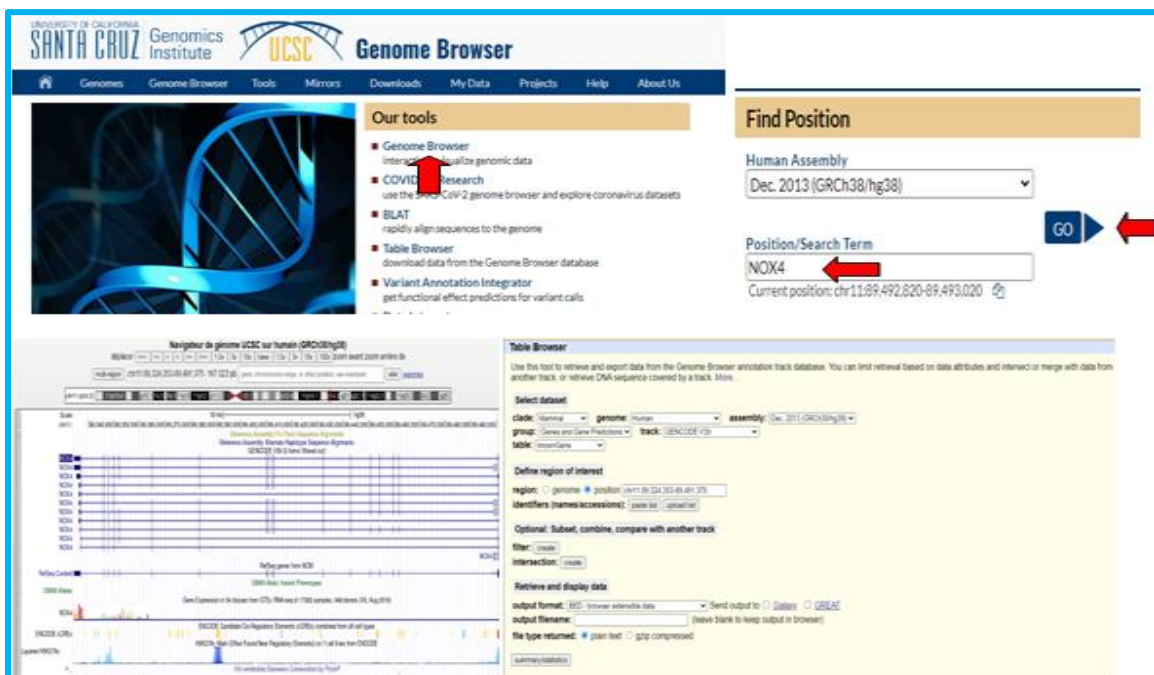
The image shows a screenshot of the UCSC Genome Browser website. At the top, there's a navigation bar with 'Genomes', 'Genome Browser', 'Tools', 'Mirrors', 'Downloads', 'My Data', 'Projects', 'Help', and 'About Us'. Below this, there's a 'Find Position' section with a dropdown menu for 'Human Assembly' set to 'Dec. 2013 (GRCh38/hg38)' and a text input field for 'Position/Search Term' containing 'NOX4'. A blue 'GO' button is to the right of the input field. Below the 'Find Position' section, there's a 'Table Browser' section with various options for selecting datasets, defining regions of interest, and retrieving data. The main content area shows a genomic track for the NOX4 gene, with various annotations and tracks visible.

Figure 18: Les étapes de recherche sur UCSC.

2.2.2.2 Outil Ensembl

Principe Ensembl fournit un navigateur génomique qui agit comme un point d'accès unique aux génomes annotés pour les espèces principalement vertébrées, qui soutient la recherche en génomique, évolution, variation de séquence et régulation transcriptionnelle. Ensemble annote les gènes, calcule plusieurs alignements, prédit la fonction de régulation et collecte des données sur les maladies. Les outils d'Ensembl incluent BLAST, BLAT, BioMart et le Variant Effect Predictor (VEP) pour toutes les espèces prises en charge, et des informations sur les domaines protéiques, la variation génétique, l'homologie, les régions synténiques et les éléments régulateurs (Fernández & Birney, 2010).

Méthode :

Tous d'abord on va vers plateforme Ensembl (<http://www.ensembl.org>), on retrouve une barre bleu claire ou on introduit le gène « **NOX4** » et On choisit l'espèce « **homo sapiens** »

Une multitude de résultat apparaissent, on clique sur « **NOX4(Human Gene)** », ensuite une nouvelle page apparaît qui donne un aperçu des informations disponibles au niveau du gène, ainsi que le tableau de transcription qui montre chaque variante d'épissage d'un gène (les transcrits codant et non codant pour les protéines) et le résumé avec des liens vers des bases de données externes.

Ensuite, on se dirige vers « **sequence** » trouvé en haut de « **Gene-based displays** », la séquence génomique du gène « **NOX4** » apparaît, on peut télécharger la séquence en cliquant sur « **download sequence** » puis on appuis sur « **configure this page** » et on a fait les changements. (annexe2,fig55),voir annexe (annexe3,fig56,57)

The screenshot shows the Ensembl genome browser interface. At the top, there are navigation links for 'Tools', 'BioMart', 'BLAST/BLAT', and 'Variant Effect Predictor'. The search bar contains 'Human' and 'NOX4'. Below the search bar, there are options for 'All genomes' and 'Favourite genomes'. The search results show 'NOX4 (Human Gene)' with a red circle around the gene name. The interface includes a sidebar on the right with information about Ensembl releases and rapid releases.

Figure 19: Les étapes de recherche sur la base des données Ensembl.

Gene: NOX4 ENSG0000009091

Description: NADPH oxidase 4 [Source:HGNC Symbol;Acc:HGNC:73214]

Location: Chromosome 11: 89 324 353-89 498 187 reverse strand
GRCh38: chr11:89324353-89498187

About this gene: This gene has 15 transcripts (coding isoforms), 182 orthologues and 6 paralogs.

Transcript ID	Name	bp	Protein	IsoType	CCDS	UniProt Match	RefSeq Match	Flags
ENST00000263317.9	NOX4-201	4269	278aa	Protein coding	CCDS44054P	Q9N725-1P	NM_018311.2P	MANE Select v0.85, Ensembl Canonical, GENCODE basic, APPRIS P1, TSL:1
ENST00000242118.5	NOX4-205	4184	254aa	Protein coding	CCDS44054P	Q9N725-1P	-	GENCODE basic, TSL:5
ENST00000343722.9	NOX4-202	3347	254aa	Protein coding	CCDS44054P	Q9N725-1P	-	GENCODE basic, TSL:2
ENST00000272956.5	NOX4-210	2979	254aa	Protein coding	CCDS44054P	Q9N725-1P	-	GENCODE basic, TSL:6
ENST00000347113	NOX4-215	2156	239aa	Protein coding	CCDS44054P	Q9N725-1P	-	GENCODE basic, TSL:1
ENST00000326241.5	NOX4-214	2039	214aa	Protein coding	CCDS73362P	Q9N725-1P	-	GENCODE basic, TSL:2
ENST00000263413	NOX4-211	1699	153aa	Protein coding	CCDS73362P	Q9N725-1P	-	GENCODE basic, TSL:2
ENST00000275205.5	NOX4-209	1774	201aa	Protein coding	ENST00000275205.5	-	-	GENCODE basic, TSL:2
ENST00000241196.5	NOX4-207	1555	242aa	Protein coding	ENST00000241196.5	-	-	GENCODE basic, TSL:1
ENST00000378979.7	NOX4-203	816	271aa	Protein coding	Q9N725-1P	-	-	GENCODE basic, TSL:1
ENST0000031342.5	NOX4-213	696	231aa	Protein coding	Q9N725-1P	-	-	GENCODE basic, TSL:1
ENST0000022022.2	NOX4-204	560	249aa	Protein coding	Q9N725-1P	-	-	GENCODE basic, TSL:1
ENST0000029842.9	NOX4-212	1403	214aa	Nonsense mediated decay	Q9N725-1P	-	-	TSL:1
ENST00000242181.1	NOX4-208	1309	No protein	Retained intron	-	-	-	TSL:2
ENST00000244731.1	NOX4-206	755	No protein	Retained intron	-	-	-	TSL:2

Figure 20: tableaux des transcrits de gène NOX4

Séquence balisée

Télécharger la séquence | BLAST cette séquence

Exons: NOX4 exons | All exons in this region

Markup loaded

```
>chromosome:GRCh38:11:89323753:89498757:-1
TGCCAAAAGTGTGACAAATTAATGCCAAAAGTGCATTCAACTA GCATGCCACAAACATAGCAA
ATATGCTTAAGAATCCCCTGTGATGGGAACATTTCATATCCCCTCGGCCCTCCTGCCGAA
TTCTTTCTTCCATATTAATAGTCATTCTGACAGTTATTTTTCCTATGGGATCCAAAAGG
TACCCTAATATATGCTTCCAAAGTGGAAAATAGGGTACAGAAATCAGTTATTTGCAGTTT
TTCTATTTTTCATTCGGTGAACAAGTTTCAGCAATTCAACTTTATAGCAATTAATAAACC
TCAACATACCTTCGGTGAACAAGTTTCAGCAATTCAACTTTATAGCAATTAATAAACC
TGTTAAGTTTATTCGGACATTTGGATTTGACGTTGAGGTTGGCAAGTAAATTTTACTGAC
AGCAACTAATGATGTTTATAAATTTTATCATATAATAGATCTATCTTTTACTATA
CTTTCTTAGTGTGTCACCAATGCAAAATACATGTTTFTTAAFGTATCTATGCTTGTGCTG
TGCCATATAGTTTGTATTAATAATACATTAACAACAACATAAATCTTTTCTCAAAG
TACTTCTAGTCTAGTGGAGAAAACAACCTGAGAAAATGCACCAACAATAATGGGGCACAATAA
CTACTCTGAGAGAGAGTGGCACACAGTGGACGCAAGAAGAGGTTCTCTGATGCAGCTGGG
AAGAATCTACATGGCCCTCTAGCAGGGAAGATCTGCTTGAGTTTGGAGACTAAGTA
AGCATCTCCAG - GTAATGCTATGAAACCAAGTAAAGCAACACATAAATGGAAATGTG
AGAAAACAGGATATGTTGGGAATAGAAAGATTAAGAAGTACAGAATATCATATTAG
TGTAAAAAGAAACTGTCAATGTAATTAATTTTGAAGAAGTACAGGAATTTAGAT
TGCTCTACTGTCTTTTCATTTTCATTAABAACCATTACTGAGACCCTACTGATGTTCCA
```

Figure 21: la séquence nucléotidique de gène NOX4

2.2.2.3 Outil SIFT (Sorting Intolerant from Tolerant).

Principe SIFT (<https://sift.bii.a-star.edu.sg/>) (figure 22) utilise une méthode basée sur l'homologie de séquence pour spéculer sur les SNP délétères et tolérés afin que les changements phénotypiques et fonctionnels sur la molécule de protéine dus aux substitutions d'acides aminés puissent être analysés. Dans notre étude, nous avons soumis des rsID extraits d'ensemble en tant que requête pour faire des prédictions. La prédiction des SNP délétères par le serveur SIFT a ensuite été utilisée pour déterminer leur effet sur la structure et la fonction du gène NOX4 (Bromberg et al., 2008).

SIFT prédit que les substitutions avec des scores inférieurs à 0,05 sont délétères. Certains utilisateurs de SIFT ont constaté que les substitutions avec des scores inférieurs à 0,1 offrent une meilleure sensibilité pour détecter les SNP délétères (Leabman et al., 2003).

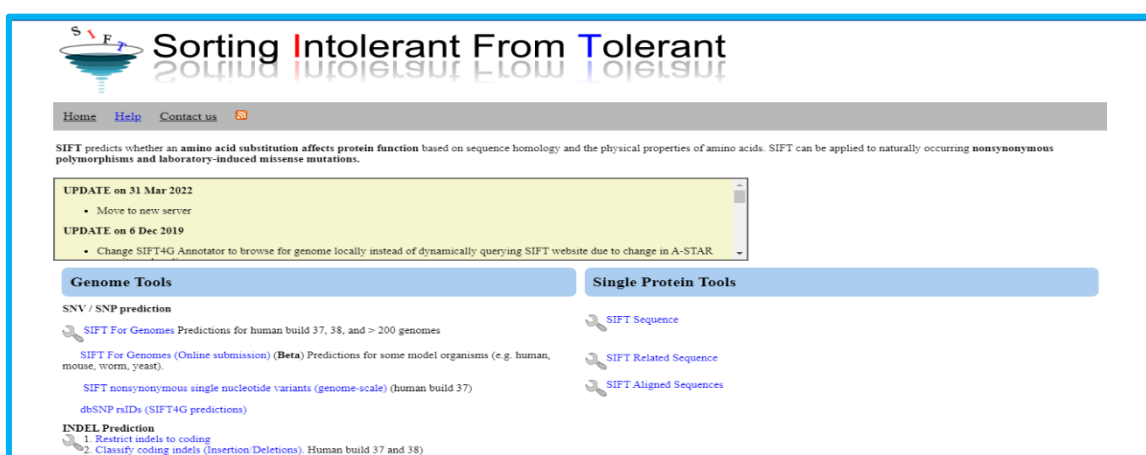


Figure 22: Les étapes de recherche sur SIFT.

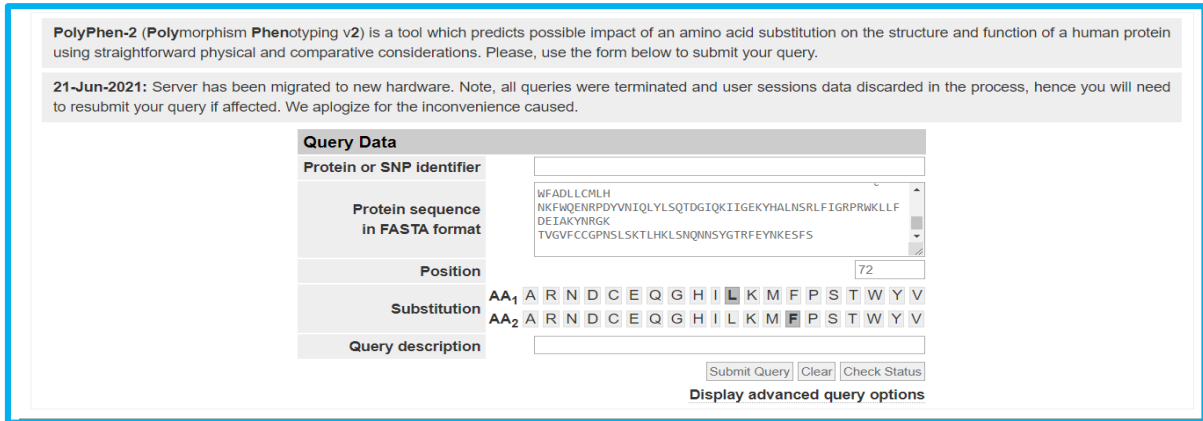
2.2.2.4 Outil PolyPhène2 (Polymorphisme Phénotypage v2)

Principe PolyPhen-2 prédit l'effet possible de la substitution d'acides aminés sur la fonction et la structure d'une protéine sur la base de plusieurs critères tels que la phylogénétique, les informations structurelles et la séquence de la protéine. Le score PolyPhen-2 varie de 0,0 (toléré) à 1,0 (délétère), de sorte que les valeurs plus proches de 1,0 sont prédites avec plus de confiance comme étant délétères (I. Adzhubei et al., 2013).

PolyPhen-2 (<http://genetics.bwh.harvard.edu/pph2/>) a été utilisé pour rechercher l'effet possible d'une substitution d'acide aminé sur la structure et la fonction de la protéine NOX4. La séquence protéique, l'ID de la base de données/le numéro d'accèsion et les détails de la substitution des acides aminés sont utilisés comme options d'entrée pour PolyPhen-2 (I. A. Adzhubei et al., 2010).

Méthode :

Le lien de l'outil Polyphen-2 <http://genetics.bwh.harvard.edu/pph2/>, nous a permis de coller la séquence protéique sous le format « FASTA » dans la zone de texte. Puis on clique sur «Submit» (figure 23).



The screenshot shows the PolyPhen-2 web interface. At the top, there is a description of the tool: "PolyPhen-2 (Polymorphism Phenotyping v2) is a tool which predicts possible impact of an amino acid substitution on the structure and function of a human protein using straightforward physical and comparative considerations. Please, use the form below to submit your query." Below this, a notice states: "21-Jun-2021: Server has been migrated to new hardware. Note, all queries were terminated and user sessions data discarded in the process, hence you will need to resubmit your query if affected. We apologize for the inconvenience caused." The main form is titled "Query Data" and contains several fields: "Protein or SNP identifier" (empty), "Protein sequence in FASTA format" (containing a FASTA sequence), "Position" (set to 72), "Substitution" (showing a comparison between AA1 and AA2), and "Query description" (empty). At the bottom right, there are buttons for "Submit Query", "Clear", and "Check Status", along with a link to "Display advanced query options".

Figure 23: Les étapes de recherche sur PolyPhen-2.

2.2.2.5 Outil SNAP2 (Screening for Nonacceptable Polymorphisms).

Principe L'outil SNAP2 (<https://www.rostlab.org/services/SNAP/>) (figure 24) prédit les effets fonctionnels des mutations. En tenant compte d'une variété de caractéristiques de séquence et de variantes, il différencie les variantes à effet des variantes neutres. Le format FASTA des séquences de protéines est utilisé comme option d'entrée pour SNAP2 (Bromberg et al., 2008).

SNAP2 répartit ses prédictions dans un large intervalle de scores, en plus de prédire un résultat binaire réduit (effet/neutre). L'échelle va de -100 (SAV fortement prédit comme neutre) à +100 (SAV fortement prédit comme effet - soit délétère, soit bénéfique). Pour une prédiction binaire, SNAP2 est optimisé pour le seuil = 0 (ie neutre : $-100 \leq \text{SNAP2-score} \leq 0$ et effet : $0 < \text{SNAP2-score} \leq 100$). À ce seuil par défaut, 78 % des SAV neutres connus et 79 % des SAV à effet connu sont prédits correctement. Des valeurs absolues plus élevées des scores impliquent des prédictions plus fiables (Mahlich et al., 2017).

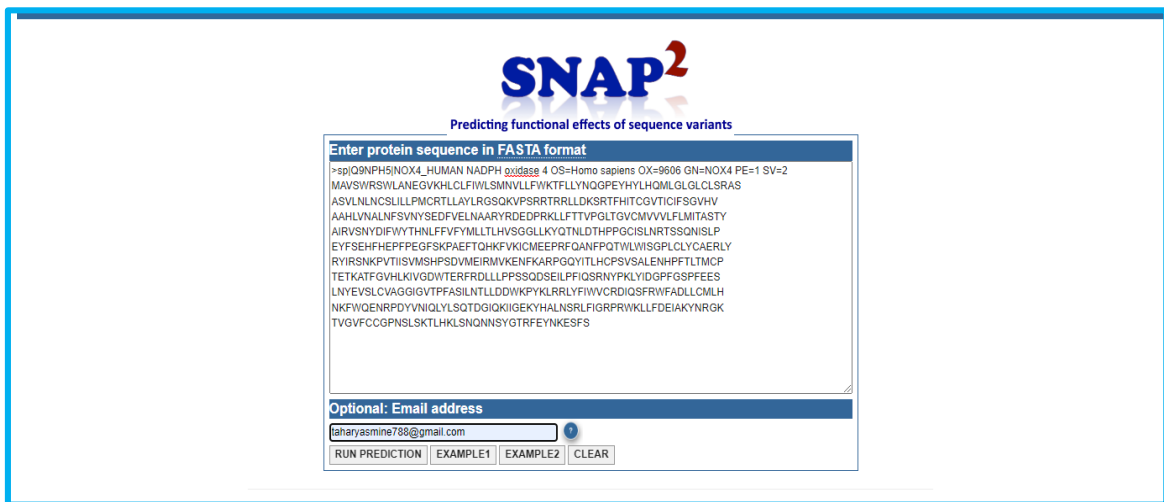


Figure 24: Les étapes de recherche sur SNAP2.

2.2.2.6 Outil PhD-SNP : Predictor of human Deleterious Single Nucleotide Polymorphisms

Principe polymorphismes (PhD-SNP) PhD-SNP est un outil Web disponible à (<http://snps.biofold.org/phd-snp/phd-snp.html>2017) (figure 25) . Il prédit si le nouveau phénotype dérivé d'un SNP est lié à la maladie (effet) ou non lié à la maladie (neutre). Dans cette étude, la séquence protéique obtenue auprès d'Uniprot a été soumise au programme après avoir fourni la position de la mutation et le nouveau résidu d'acide aminé (Bava et al., 2004).

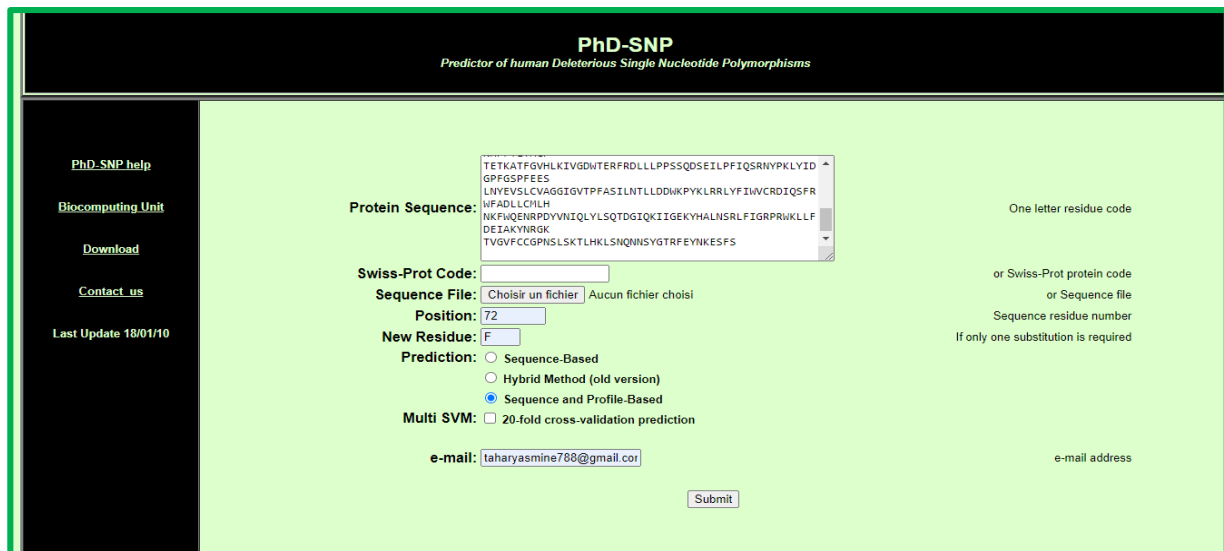


Figure 25: Les étapes de recherche sur PhD-SNP.

2.2.2.7 Outil I-Mutant2.0

Principe I-Mutant2.0(<http://folding.biofold.org/i-mutant/i-mutant2.0.html>) (figure 26) est un serveur Web basé sur un réseau de neurones pour la prédiction automatique de la stabilité des protéines changements sur les SNP délétères. I-Mutant2.0 peut classer les signes des changements de stabilité des protéines dus à une variation et il peut prédire le changement relatif de l'énergie sans Gibbs (ΔG) à une température donnée (Capriotti et al., 2005).

Dans cette version améliorée, la méthode permet de prédire si une mutation peut largement déstabiliser la protéine ($DDG < -0,5$ Kcal/mol) ou largement stabiliser ($DDG > 0,5$ Kcal/mol) ou avoir un effet faible ($-0,5 \leq DDG \leq 0,5$ Kcal/mole). Pour cette tâche particulière, le I-MutantDDG atteint une précision de 68 % lorsque la structure tridimensionnelle est connue et de 64 % lorsque seule la séquence protéique est disponible (<http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi>) .

Figure 26: les étapes de recherche sur I-Mutant2.0.

2.2.2.8 Outil MutPred2

MutPred2 est un outil basé sur l'apprentissage automatique méthode et progiciel qui intègre des données génétiques et moléculaires pour raisonner de manière probabiliste sur la pathogénicité des substitutions d'acides aminés (Cooper & Shendure, 2011).

Actuellement, ce site Web fournit MutPred2.0. Il nécessite des séquences de protéines au format FASTA, une liste de substitutions d'acides aminés dans les en-têtes FASTA correspondants (séparés uniquement par des espaces) et une adresse e-mail valide. L'ID de protéine ne peut pas contenir d'espaces, de points-virgules ou de virgules.

Le résultat de MutPred2 consiste en un score général (g), c'est-à-dire la probabilité que la substitution d'acides aminés soit pathogène. Ce score est la moyenne des scores de tous les réseaux de neurones dans MutPred2. S'il est interprété comme une probabilité, un seuil de score de 0,50 suggérerait une pathogénicité. (<http://mutpred.mutdb.org/#qform>) (figure 27).

MutPred2 web server

The web server can be used for small-scale data sets (at most 100 substitutions, with no input sequence of length > 35,000 residues). Currently, this web site provides MutPred2.0. It requires protein sequences in the FASTA format, a list of amino acid substitutions in the corresponding FASTA headers (separated by spaces only), and a valid email address. The protein ID cannot contain spaces, semi-colons or commas.

E-Mail

taharyasmine75@gmail.com

Sequence and substitution information (FASTA format)

```
>NOX4_HUMAN L7ZF
MAVSWRSWLANEGVKHLCLEFIWLSMNVLLFWKTFLLYNOGPEYHYLHOMLGLGLCLSRAS
ASVNLNLCSLLLPMCRITLALYLRGSKVPSRRTRRLDKSRTHFITCGVTICIFSGVHV
AAHLNVALNFSVNYGDFVELIAARYRDEDRKLLFTTVPGLTGVGMVWFLMITASTY
AIRVSNYDFWYTHNLFFVYMLLTLHVSGLLKYQTNLDTHPPGCSLNRSSQNSLIP
EYFSEHFHEPFEGFSKPAEFTQHKFKVCMEEPRFOANPQTWLVWISGPLECYCAERLY
RYIRSNKPVTVISVMSHPSDVMEIRMVKENFKARPGQVYTLHCPVSALENHPFTLMCP
TETKATFGVILKVGQWTERFRDQLLFPSSQSEILPFGSRVYKLYIDGPPGSPFEES
LNYEVSVCVAGGIVTFFASILNLLDDWPKYKLRRLYFVWCRDIQSFRWFADLLCMLH
NKFVQENRPDYVNIQLYLSQTDGIGIKIKEYHALNSRLFIGRPRWLLFDEIAKYNRGK
```

OR
Upload FASTA file

Choisir un fichier Aucun fichier choisi

P-value threshold (optional)

0.05

Submit

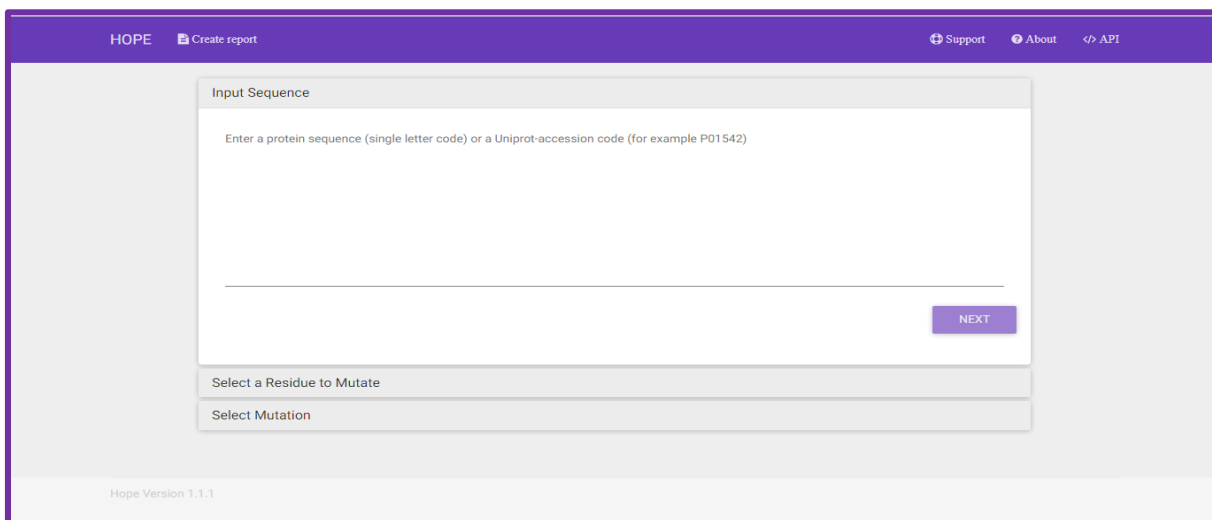
Figure 27: Les étapes de recherche sur MutPred2.

2.2.2.9 Outil HOPE Project

Principe Le projet HOPE permet la visualisation structurale en 3D des protéines mutées et donne les résultats en utilisant UniProtKB et les prédictions des serveurs DAS. La séquence FASTA de protéine entière et la sélection de variantes mutantes sont considérées comme une option d'entrée, la sortie est basée sur la variation structurale entre les résidus mutants et les résidus de type sauvage «wild-type» (Venselaar et al., 2010).

Méthode :

HOPE Project a révélé la structure 3D des protéines avec son nouveau résidu. En outre, il a décrit la réaction et les propriétés physicochimiques de ces candidats. Propriétés de ces candidats. Pour lancer la recherche sur HOPE Project, d'abord l'accédez à la page d'accueil par le site <https://www3.cmbi.umcn.nl/hope/input/#> (figure 28) ,après on va coller la séquence protéique ou le code d'accèsion à UniProt., En suite, sélectionnez un résidu à muter.et Sélectionnez la Mutation et lancer la recherche.



The screenshot shows the HOPE project web interface. At the top, there is a navigation bar with 'HOPE', 'Create report', 'Support', 'About', and 'API'. The main content area is titled 'Input Sequence' and contains a text input field with the placeholder text 'Enter a protein sequence (single letter code) or a Uniprot-accession code (for example P01542)'. Below the input field is a 'NEXT' button. Underneath the input field, there are two more sections: 'Select a Residue to Mutate' and 'Select Mutation'. At the bottom left of the interface, it says 'Hope Version 1.1.1'.

Figure 28: les étapes de recherche avec HOPE project.

2.2.2.10 Outil UniProt

Principe Cet outil en ligne (<https://www.uniprot.org/>) « UniProt », pour Universal Protein Resource, permet d'étudier les protéines dans un cadre évolutionniste. Il vous permet en particulier de :

1. Rechercher une protéine donnée, dans la banque, pour connaître sa séquence.
2. Etudier en détail l'organisation moléculaire d'une protéine.
3. Comparer une même protéine mais provenant d'espèces animales ou végétales différentes dans le but d'établir le degré de parenté entre espèce.
4. Rechercher dans la banque, le nom de protéine à partir d'une séquence d'acides aminés connue. (**Consortium, 2021**)

Méthode

La page d'accueil d'UniProt à l'[adresse http://www.uniprot.org/](http://www.uniprot.org/), permet d'introduire notre requête « nox4 » dans l'onglet recherche après avoir sélectionné UniProtKB. Ainsi, une nouvelle page apparaît donnant un aperçu sur les informations disponibles au niveau de la base UniProt, on peut finalement récupérer la séquence protéique de la NOX4 (figure 29, 30, 31) voir annexe (annexe 3, fig58, tableau XVII).

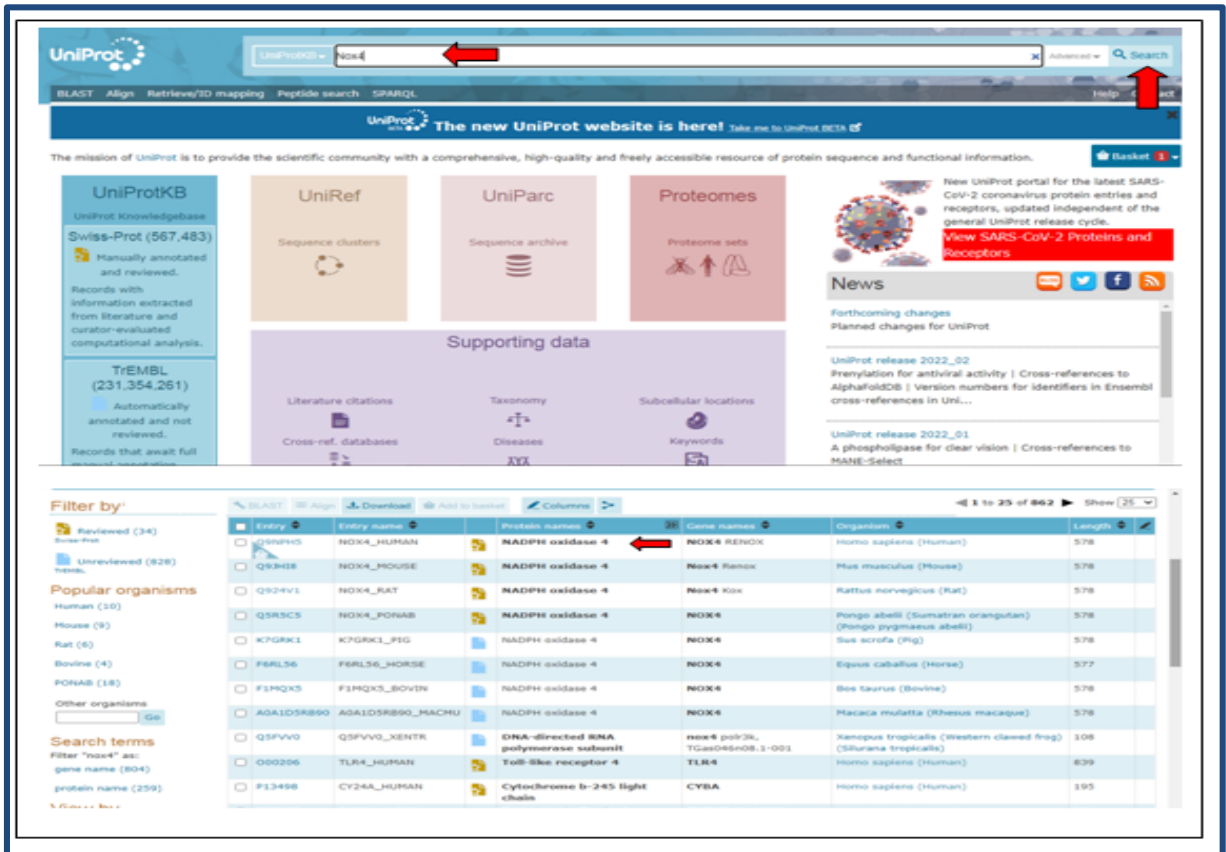


Figure 29: Les étapes de recherche sur UniProt.

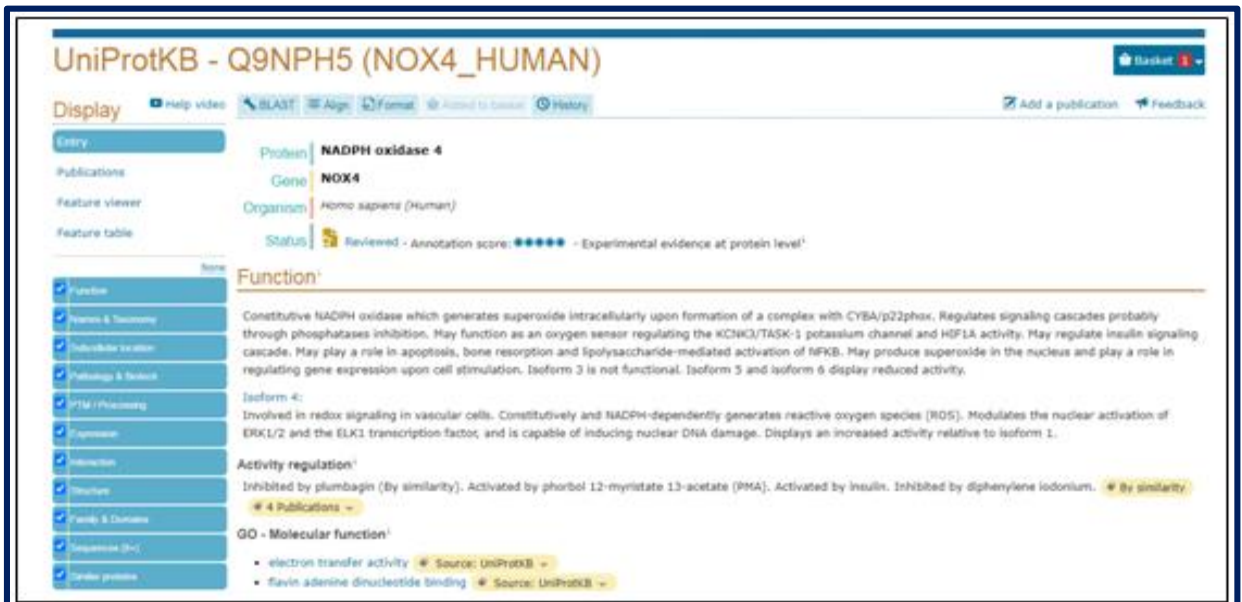


Figure 30: Les différentes informations fournis par UniProt sur le gène NOX4.

2.2.2.11 Outil SMART

Principe Simple Modular Architecture Research Tool (SMART), est une base de données biologique qui est utilisée dans l'identification et l'analyse des domaines protéiques dans les séquences de protéines. SMART utilise des modèles de Markov à profil caché construits à partir d'alignements de séquences multiples pour détecter les domaines protéiques dans les séquences de protéines (Schultz et al., 2000). SMART nous permet de obtenir plusieurs informations sur les domaines fonctionnelle de a protéine et bien comprendre la fonction de cette dernière.

Méthode

le lien : <http://smart.embl-heidelberg.de/> permet de visualiser la page d'accueil de l'outil SMART. Par la suite on clique sur « smart mode normale » puis on colle notre séquence sous format FASTA dans la zone de texte et on clique sur « **sequence smart** ». Une nouvelle page apparait donnant un aperçu les informations disponibles sur les domaines fonctionnelles de la protéine NOX4 et leurs positions présentées par schéma (figure31) et tableau VI.

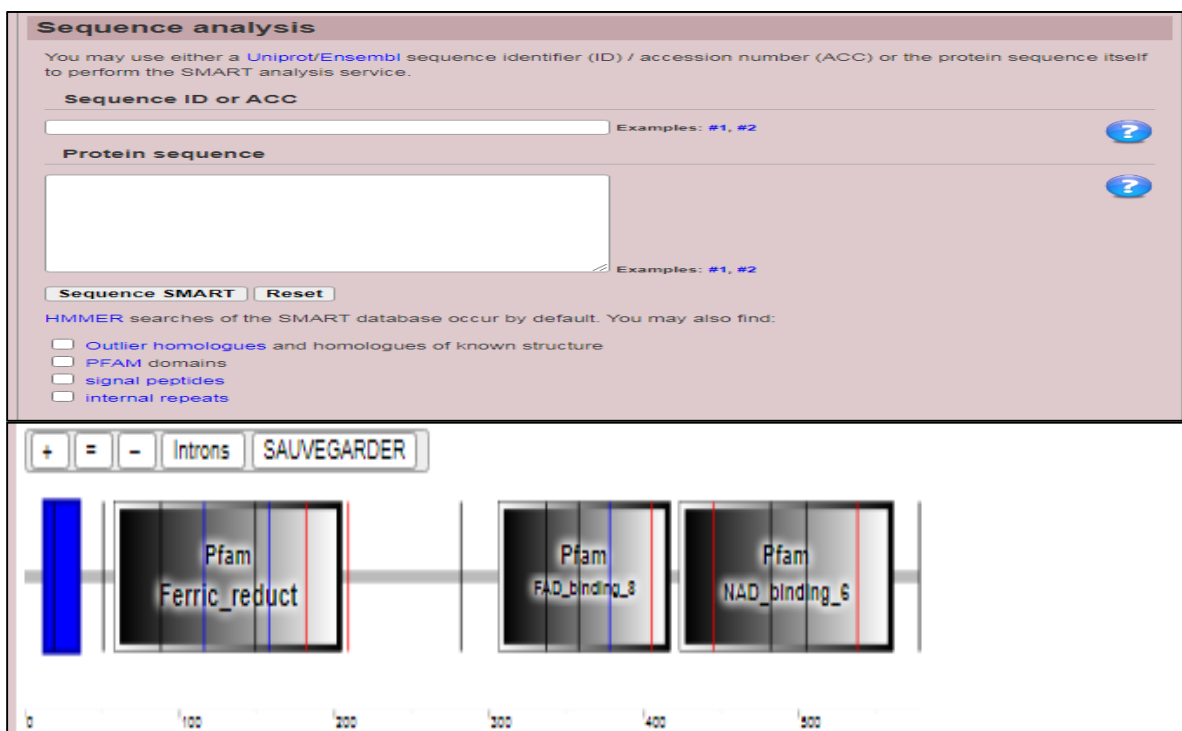


Figure 31: Les étapes de recherche sur SMART.

Tableau VI: les domaines de la protéine NOX4 (SMART).

Domaine	Description
Région transmembranaire	-région d'hélice transmembranaire. -La région commence à la position 13 et se termine à la position 35.
Domaine Ferrique	- 58 à 205 -Ce domaine représente une région commune dans les protéines transmembranaires de la chaîne lourde du cytochrome B-245 de mammifère
Domaine FAD	-306 à 417 -Ce domaine de liaison FAD est associé aux protéines de liaison ferrique réductase NAD et à la chaîne lourde du cytochrome b-245.
domaine NAD	- 423 à 561 -Cette entrée contient des protéines de liaison ferrique réductase NAD.

2.2.2.12 Outil ORF-Finder

Principe L'ORF Finder (Open Reading Frame Finder) est un outil d'analyse graphique qui trouve tous les cadres de lecture ouverts d'une taille minimale sélectionnable dans une séquence utilisateur ou dans une séquence déjà présente dans la base de données. Cet outil identifie tous les cadres de lecture ouverts en utilisant les codes génétiques standards ou alternatifs. La séquence d'acides aminés déduite peut être enregistrée dans divers formats et recherchée dans la base de données de séquences à l'aide du serveur de l'outil de recherche d'alignement local de base (BLAST). L'ORF Finder devrait être utile pour préparer des soumissions de séquences complètes et précises. Il est également fourni avec le logiciel de soumission de séquence Sequin (analyseur de séquence) (Sieber et al., 2018). Cependant, il existe plusieurs méthodes pour la recherche des ORFs, Nous avons choisi NCBI ORF finder.

Méthode :

La d'accueil d'ORF Finder, permet de voir un champ de texte qui permet la recherche des ORFs de notre séquence requête, sous format FASTA. Par la suite on clique sur le bouton de recherche ORF pour obtenir le résultat. Le résultat montre les cadres de lecture possibles présents dans la requête de séquence saisie. Par la suite en recherche l'ORF le plus long (figure 32 et 33) voir annexe (annexe4,fig59,60).

Figure 32: Les étapes de recherche des ORF (NCBI ORF-finder).

Label	Strand	Frame	Start	Stop	Length (nt aa)
ORF12	+	3	1521	1850	330 109
ORF6	+	2	2465	2755	291 96
ORF21	-	1	204	28	177 58
ORF1	+	1	2671	2844	174 57
ORF31	-	3	1084	914	171 56
ORF2	+	1	3943	4104	162 53
ORF11	+	3	1092	1253	162 53
ORF5	+	2	1669	1729	162 53
ORF9	+	3	354	505	153 50
ORF15	+	3	3402	3554	153 50
ORF14	+	3	3222	3371	150 49

Figure 33: les nombre des ORF (31) obtenues et l'ORF le plus long dans le gène NOX4.

2.2.2.13 Outil SWISS-MODEL

Principe C'est un serveur pour la modélisation comparative automatisée de structures protéiques tridimensionnelles (3D). Il a été le pionnier dans le domaine de la modélisation automatisée à partir de 1993 et est aujourd'hui l'installation de modélisation automatisée gratuite basée sur le Web la plus largement utilisée, SWISS-MODEL propose plusieurs niveaux d'interaction avec l'utilisateur via son interface World Wide Web : en « mode première approche », seule une séquence d'acides aminés d'une protéine est soumise pour construire un modèle 3D. La sélection de modèles, l'alignement et la construction de modèles sont entièrement automatisés par le serveur. Dans le « mode d'alignement », le processus de modélisation est basé sur un alignement cible-modèle défini par l'utilisateur. Les tâches de modélisation complexes peuvent être gérées avec le "mode projet" à l'aide de DeepView (Swiss-PdbViewer), un atelier intégré séquence-structure (Waterhouse et al., 2009).

Méthode :

Pour lancer la recherche sur SWISS-MODEL, nous insérons la séquence protéique ou bien le numéro d'accession et nous lançons la recherche.

Lorsque la recherche de modèles est terminée une page s'affiche avec une liste des modèles disponibles classés en fonction de la qualité attendue des modèles résultats évalués par GMQE (Global Model Quality Estimation) qui varie entre 0-1 Ce score de qualité nous permet de choisir le ou les bons modèles. Nous pouvons observer alors qu'ils partagent tous une identité de séquence élevés avec la cible requête. Ces modèles sont affectés dans la visionneuse 3D en cliquant sur « Build a model ». Ces modèles sélectionnés représentent la modélisation de la protéine cible (figure 34,35)

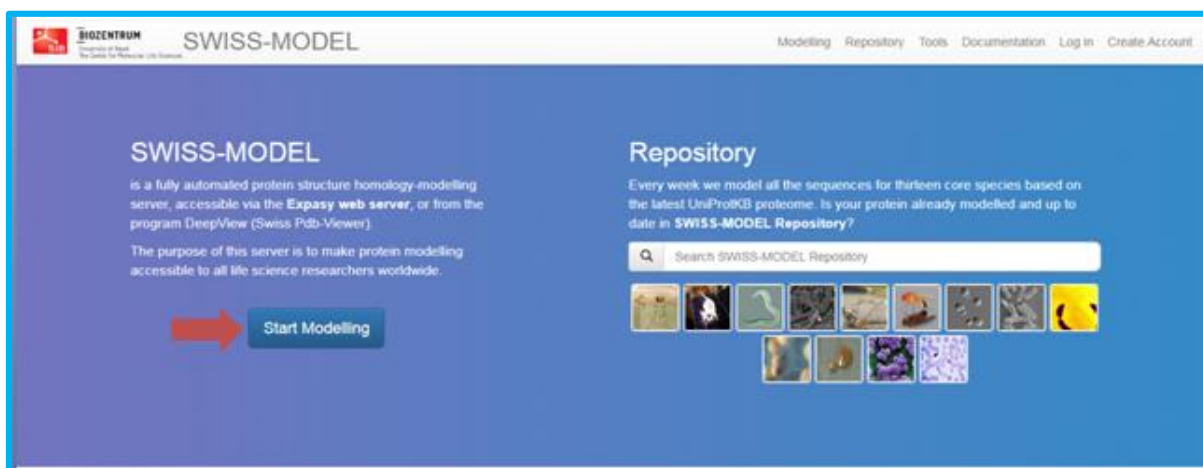


Figure 34: la modélisation dans la base de données SWISS-MODEL.

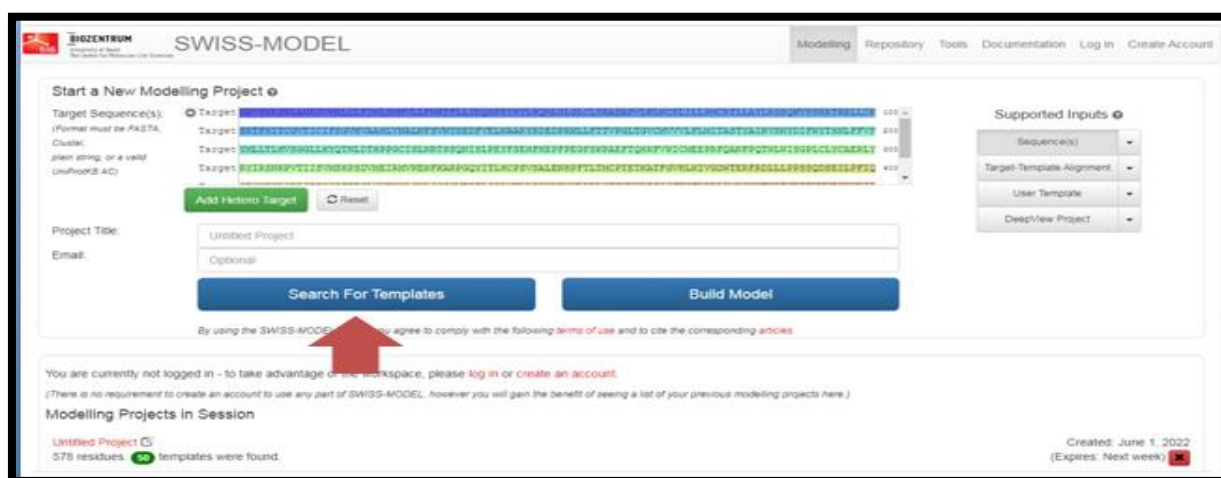


Figure 35: Les étapes de la modélisation.

2.2.3 Docking Moléculaire

Le Docking moléculaire *in silico* vise à prédire la structure d'un complexe moléculaire à partir des molécules isolées, ce qui est considérablement plus facile à mettre en œuvre, moins cher et plus rapide que l'utilisation des méthodes expérimentales *in vitro*. Les logiciels de docking sont donc des outils très utiles en biologie, pharmacie et médecine, car la plupart des principes actifs sont de petites molécules (ligand) qui interagissent avec une cible biologique d'intérêt thérapeutique, généralement protéique (récepteur), afin d'influencer le mécanisme dans lequel cette protéine est impliquée (**Kitchen et al., 2004**). La méthode et les étapes du docking moléculaire est représentée par la figure 36.

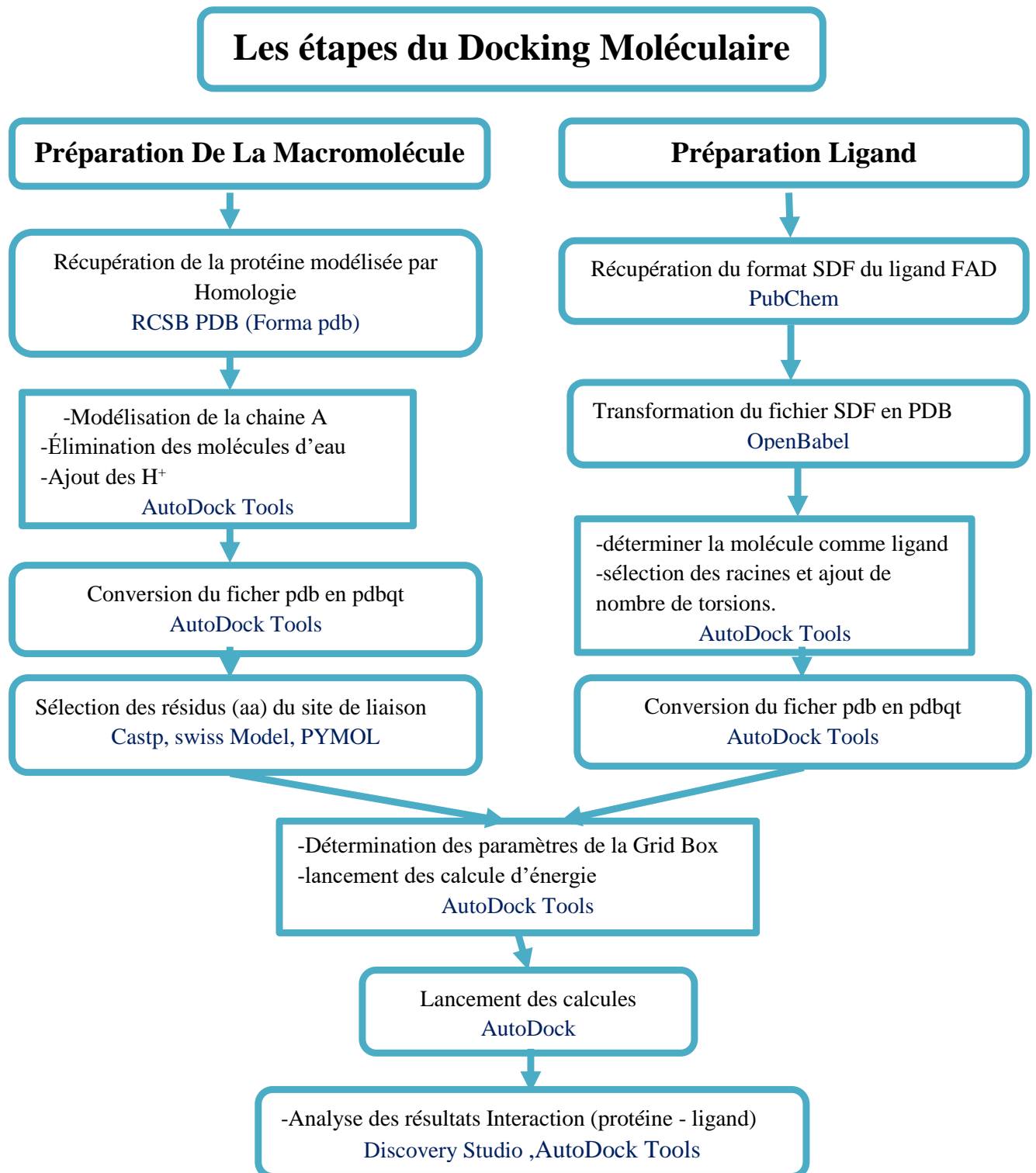


Figure 36: Méthode du Docking moléculaire (Amarrage moléculaire).

2.2.3.1 Outil PDB

La structure de l'enzyme NOX4 provient de la PDB « Protein Data Bank ». Cette dernière, est la plus grande archive de données structurales de macromolécules biologiques, comme les protéines et les acides nucléiques. Les données générées par cette base sont principalement obtenues par cristallographie aux rayons X et résonance magnétique nucléaire (RMN).

Méthode

le lien <https://www.rcsb.org/> va nous guider directement vers la plateforme PDB, Interdire le nom de la protéine « 7d3e » dans la barre de recherche puis appuies sur entré pour lance la recherche La page Résultat de la recherche affiche la forme 3D de la protéine cherché et une description sur la structure de cette dénaire et des autre information en peut aussi téléchargé la forme PDB de cette protéine en clique simplement sur téléchargé des fichiers après sur forma PDB (figure 39 et 40).

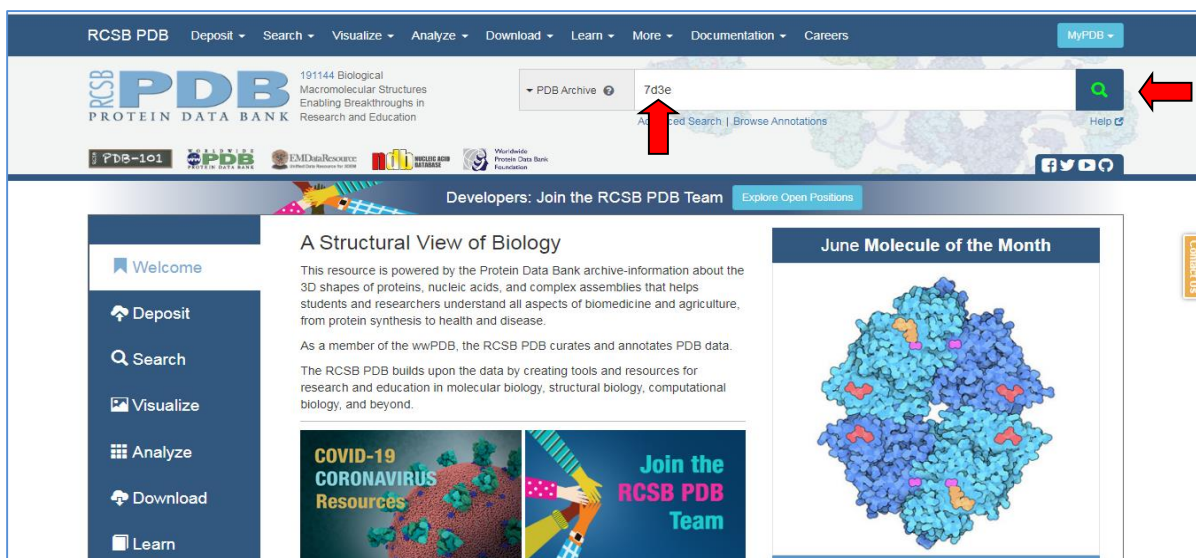


Figure 37: La base de données PDB.

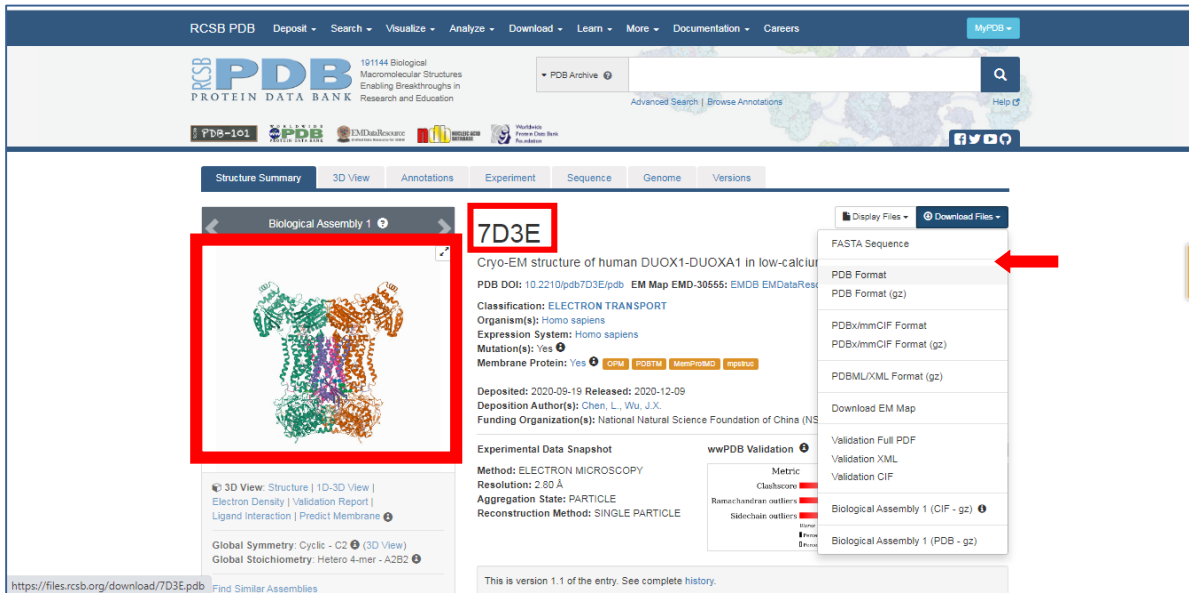


Figure 38: Récupération du format PDB de la protéine.

2.2.3.2 Outil SWISS-MODEL.

A partir de la liste des ligands citée par SWISS-MODEL, on peut choisir et télécharger les ligands à partir de deux bases des données PDB et PubChem. Avec un simple clic sur le ligand cité par SWISS-MODEL. Ce qui va nous menez à une nouvelle page dans la base des données PDB présentant le ligand et les informations chimiques comme la formule chimique, la structure 2D et 3D, le poids moléculaire etc. On peut le télécharger la structure 3D sous format SDF (figure 39 et 40)



Figure 39: Recherche des ligands dans SWISS-MODEL.

The screenshot shows the RCSB PDB website interface. At the top, there is a navigation bar with options like 'Deposit', 'Search', 'Visualize', 'Analyze', 'Download', 'Learn', 'More', 'Documentation', and 'Careers'. The main header features the RCSB PDB logo and the tagline '101144 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education'. A search bar is visible with 'PDB Archive' selected. The main content area displays the search results for 'FAD' (FLAVIN-ADENINE DINUCLEOTIDE). It includes two 3D ball-and-stick models of the molecule, one in a vertical orientation and one in a horizontal orientation. Below the models are two tables: 'Chemical Component Summary' and 'Chemical Details'.

Chemical Component Summary	
Name	FLAVIN-ADENINE DINUCLEOTIDE
Identifiers	[[2R,3S,4R,5R]-5-(6-aminopurin-9-yl)-3,4-dihydroxyoxolan-2-yl]methyl [[(2R,3S,4S)-5-(7,8-dimethyl-2,4-dioxo-benzo[pteridin-10-yl]-2,3,4-trihydroxy-pentoxyl)-hydroxy-phosphoryl] hydrogen phosphate
Formula	C ₂₇ H ₃₃ N ₉ O ₁₅ P ₂

Chemical Details	
Formal Charge	0
Atom Count	86
Chiral Atom Count	9
Bond Count	91
Aromatic Bond Count	16

Figure 40: Recherche du ligand Flavin Adenine Dinucleotide (FAD) de la protéine NOX4.

2.2.3.3 Outil PubChem

Principe PubChem répertorie plusieurs millions de composés en mettant gratuitement en ligne, pour chaque substance, une grande quantité de données de divers ordres : chimique, biochimique, pharmacologique, production, toxicologique, environnemental, etc.

Méthode :

Le lien <https://pubchem.ncbi.nlm.nih.gov/> nous guide directement à la base de données PubChem, on insère dans la barre de recherche le nom de ligand. La page de résultats apparaîtra. On choisit le bon ligand et on le télécharge en format SDF. (figure 41)

The image is a composite of three screenshots from the PubChem website. The top screenshot shows the search bar with the text 'Flavin adenine dinucleotide (FAD)' and a dropdown menu listing several related terms. The middle screenshot shows the 'Compound Summary' page for Flavin adenine dinucleotide, featuring a 3D ball-and-stick model of the molecule, its molecular formula $C_{15}H_{22}N_{10}O_{17}P_2$, and a table of contents. A red arrow points to the 'Download' button in the table of contents. The bottom screenshot shows a 'DOWNLOAD' dialog box with options to save data in JSON, XML, SDF, and ASNT formats, and a section for downloading a PDF of the page.

Figure 41: La récupération du ligand FAD du site PubChem.

2.2.3.4 Outil PrankWeb

Principe PrankWeb s'appuie sur P2Rank, une méthode basée sur l'apprentissage automatique pour la prédiction des sites de liaison des ligands à partir de la structure des protéines. (Prédiction du site de liaison de ligands)

Méthode

Après avoir consulté le site <https://prankweb.cz/>, nous avons inséré le mot clé caractérisé par le symbole de la protéine « 7D3E » puis nous avons lancé la recherche. La recherche nous donne les sites de liaisons aux ligands pour télécharger le paquet d'informations (figure 42).

The image displays two screenshots of the PrankWeb interface. The top screenshot shows the input form where the PDB Code '7D3E' is entered, and the 'Use original structure' and 'Use conservation' options are checked. The bottom screenshot shows the results, including a protein structure visualization with binding sites highlighted in various colors. A text box with the French text 'Les sites de liaison colorée' (Colored binding sites) has a blue arrow pointing to the highlighted binding sites on the protein structure. The right sidebar contains a 'Tools' section with a 'Download' button, a 'Source' section with 'Experimental structure 7D3E', and a 'Pockets' section with a table of pocket statistics.

Property	Value
Pocket rank	1
Pocket score	28.26
Probability score	0.906
AA count	33
Conservation	0.000

Figure 42: l'identification du site de liaison avec PrankWeb.

2.2.4 Méthode du Docking moléculaire

Principe : Moléculaire Docking est le processus dans lequel l'interaction intermoléculaire entre 2 molécules a été étudiée en In-silico. Dans ce processus, la macromolécule est le récepteur de la protéine. La micromolécule est la molécule de ligand qui peut jouer le rôle d'inhibiteur.

Méthode

Ainsi, le processus d'amarrage implique les étapes suivantes :

Étape I - préparation de la protéine : la structure tridimensionnelle de la protéine doit être extraite de la banque de données sur les protéines (PDB), ensuite, la structure récupérée doit être prétraitée. Cela doit admettre l'élimination des molécules d'eau de la cavité, stabilisation des charges, remplissage des résidus manquants (**Chaudhary & Mishra, 2016**), sélection des chaînes impliquées dans l'interaction etc. selon les paramètres disponibles, voire (annexe 6,fig62)

Étape II – prédiction du site actif : après la préparation de la protéine, le site actif de la protéine doit être prédit. Le récepteur peut posséder de nombreux sites actifs, mais celui de la préoccupation doit être choisi. La plupart du temps, les molécules d'eau et les hétéroatomes sont éliminés s'ils sont présents (**McMartin & Bohacek, 1997**) (**Schnecke & Kuhn, 2000**). Dans notre cas le site actif a été prédit avec castp, SWISS Model et PYMOL, voire (annexe6,fig63)

Étape III - préparation du ligand : les ligands peuvent être récupérés à partir de plusieurs bases de données telles que ZINC, Pub Chem ou peuvent être esquissés à l'aide de l'outil d'esquisse Chem., voire(annexe 7,fig64)

Étape IV - amarrage : le ligand est ancré contre la protéine et les interactions sont analysées. La fonction de score donne un score sur la base du meilleur complexe de ligand amarré qui est choisi (**Chaudhary & Mishra, 2016**). Le ligand choisi dans notre cas est Flavin adenine dinucleotide FAD.

RESULTATS

Chapitre 3 : Résultats

3. Chapitre 3 : Résultats

Rappelons que notre travail est une étude *in silico* qui utilise des outils bio-informatiques pour comprendre et étudier l'association du polymorphisme du gène NOX4 dans le développement du diabète.

3.1 Résultats de l'OUTIL NCBI

Les résultats de notre recherche dans les bases de données NCBI « dbVar, LitVar », Ensembl, UCSC sont représentés par l'histogramme (Figure 43) montrant les nombres des variants polymorphiques du gène NOX4.

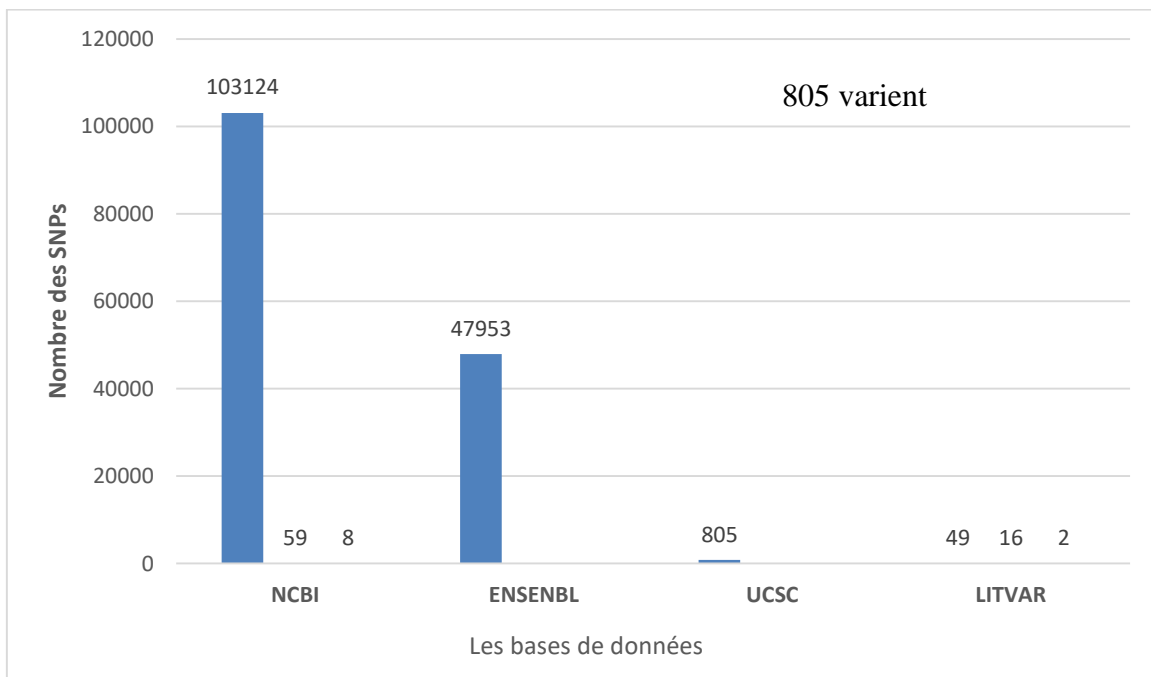


Figure 43: Nombre des variants polymorphiques récupérés à partir des différentes plateformes.

Les résultats de notre recherche montrent que le gène NOX4 chez l'homme est caractérisé par un profil polymorphique de 103124 variants identifié par NCBI. 49 SNP du gène NOX4 ont été identifié par l'outil LitVar dont seulement 8 SNP sont liés aux diabètes. 805 SNP ont été identifié par l'outil UCSC. Alors que 47953 ont été identifié par l'outil Ensembl.

Chapitre 3 : Résultats

3.2 Cas cliniques de l'association des « rs » du gène NOX4

Les résultats de l'association du polymorphisme des références snip « rs » au diabète décrit par les études cas témoins cliniques sont répertoriés par le tableau ci-après (tableau VII)

Tableau VII: Résultats des « rs » du gène de la NOX4 associés dans les cas cliniques.

IDrs	Position (Chr)	Position	Allèle mineure	Contrôle	Cases	Fréquence allélique mineure (cases)	Fréquence allélique mineure (témoin)	P value	OR (95% CI)	Population	Les maladies associées	Les références
rs1836882	chr11:89498993	Promoteur	C	848	254	0.28	0.27	0.588	1.06 (0.85-1.33)	coréenne	NODAT	(Hyeon Seok Hwang et al 2019)
rs614128	chr11:89480946	2-3	G	445	487	-	-	0.005	2.05	Chinoise	Néphropathie diabétique	(S. C. Lim et al 2009)
rs490934	chr11:89490448	2-3	C	445	487	-	-	0.005	2.05	Chinoise	Néphropathie diabétique	(S. C. Lim et al 2009)
rs553635	chr11:89496263	promoteur	A	445	487	-	-	0.005	2.05	Chinoise	Néphropathie diabétique	(S. C. Lim et al 2009)
rs3017887	chr11:89492920	5'UTR	A	445	487	-	-	0.005	2.05	Chinoise	Néphropathie diabétique	(S. C. Lim et al 2009)
rs10765219	chr11:89621110	13-14 31 kb to NOX4	T	4106	5060	0.424	0.529	7,41. 10-8	1.54 (1.31-1.80)	Ecosses (Taysique)	Rétinopathie diabétique	(Weihua Meng et al, 2018)
rs11018670	chr11:89623460	12-13 33 kb to NOX4	G	4106	5060	0.427	0.534	1,23. 10-8	1.55 (1.33-1.80)	Ecosses (Taysique)	Rétinopathie diabétique	(Weihua Meng et al, 2018)
rs3913535	chr11:89363589	16-17	C	4106	5060	0.411	0.518	4,05. 10-9	1.55 (1.34-1.79)	Ecosses (Taysique)	Rétinopathie diabétique	(Weihua Meng et al, 2018)

NODAT: New onset diabetes after transplantation.

L'interrogation des différents navigateurs, concernant les études cas cliniques « cas-témoins » a montré trois (3) études d'associations du polymorphisme du gène de la NOX4 avec le diabète sucré et ses complications. Parmi ces études 8 rs de NOX4 ont été identifiés.

La première étude cas témoins faite sur une population coréenne a étudié le rs1836882 dans le cas de l'apparition du diabète après transplantation rénale. Les résultats de ce dernier, n'ont pas montré de différence significative ($p=0.588$; $\alpha=0,05$) ni une susceptibilité à développer la maladie.

La deuxième étude concerne une analyse par puce à ADN de plusieurs gènes candidats et protéines plasmatiques associées à la néphropathie secondaire au diabète de type 2 chez la population Chinoise. Parmi les 13 SNP étudiés, quatre snip (rs614128G>C, rs490934G>C, rs3017887C>A et rs553635C>T) sont regroupés pour former un bloc d'haplotype de NOX4 à l'extrémité 5'. Les résultats ont été très significatifs ($P=0.005$; $\alpha=0,05$) et ont montré une forte susceptibilité et association à la néphropathie diabétique (OR= 2.05) (tableaux 1, 2).

Une autre étude d'association à l'échelle du génome (GWAS) s'est intéressée à explorer l'association du gène NADPH oxydase 4 (NOX4) avec la rétinopathie diabétique sévère chez les diabétiques de type 2. Cette dernière, est une étude cas témoins regroupant cinq mille soixante (5060) cas de DT2 génotypés et quatre mille cent six (4106) témoins chez la population écossaise. Les résultats de dernière étude ont montré une forte association à la rétinopathie diabétique pour les trois « rs » étudiés, rs10765219 (OR= 1.54) ; rs11018670 (OR=1.55) et rs3913535 (OR=1.55) en plus les valeurs ont été très significatives $p=7,41 \times 10^{-8}$, $P=1,23 \times 10^{-8}$, $P=4,05 \times 10^{-9}$ respectivement.

3.3 Résultats de l'outil Ensembl

Le total de 47954 variants de diverses classes a été trouvé dans le gène NOX4 à l'aide de l'outil Ensembl. Parmi ces mutations, nous retrouvons plus fréquemment des SNPs (n = 41 912 ; 87%), des mutations type indel (n = 3143 ; 4%), des mutations type COSMIC (n = 2024 ; 7%), type délétion (n = 654 ;1%), type insertion (n = 215 ; %) et substitution (n = 9 ; 0%) (Figure 44).

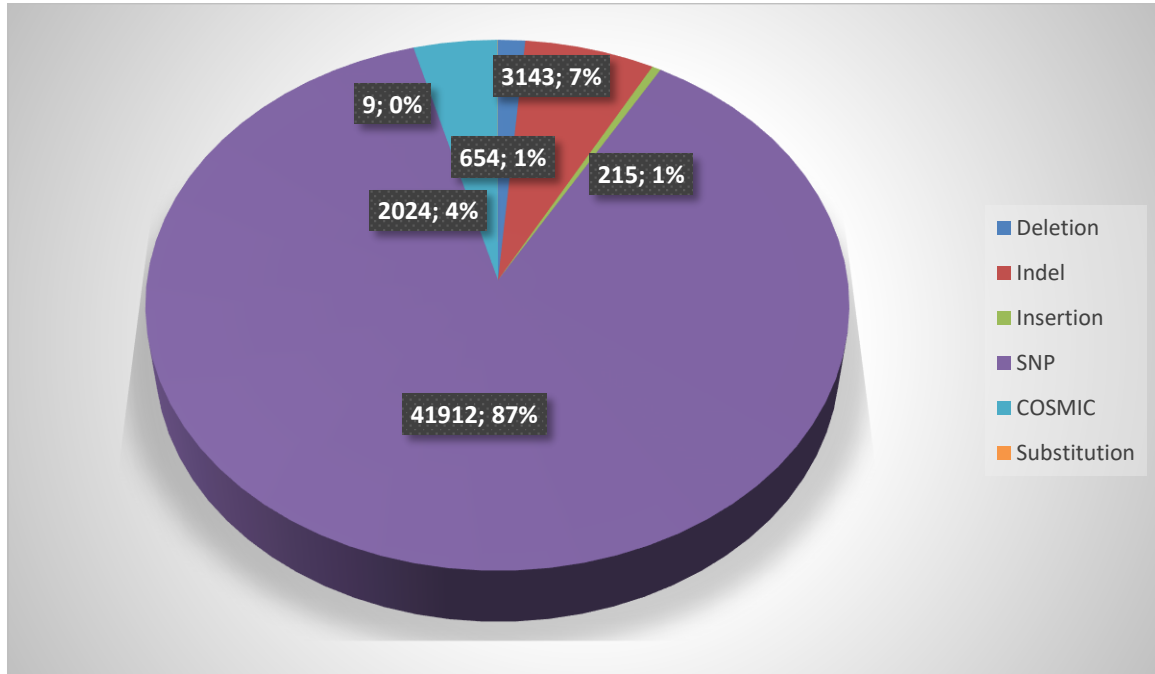


Figure 44:Diagramme circulaire représentant les différents types de variations de NOX4 selon le navigateur Ensembl.

La base de données Ensembl, nous a permis d'identifier 47953 SNPs associés aux pathologies et classés selon des scores de tolérances ou de pathogénicités. Citons le cas de l'outil PolyPhen et SIFT (annexe 2, fig 55).

3.4 Résultats de l'OUTIL UCSC

Le navigateur UCSC est un outil conçu pour permettre la visualisation d'un très grand nombre de snip avec leurs numéros « rs » (<http://genome.ucsc.edu>) .

La figure 45 présente une piste des « rs » montrant quatre types de SNPs : des SNPs codants non synonymes (rouge), des SNPs codants synonymes (verts), des SNPs des régions introniques (noir) et des régions non traduites (bleu). Ces SNPs sont distribuées sur toute la longueur du gène « NOX4 ». Le navigateur UCSC nous a permis aussi d'observer la cartographie du gène NOX4, ainsi que son emplacement dans le chromosome 11.

Le paramétrage de la piste des variants nous a permis d'identifier diverses pistes de « rs » telles que : 153 dbsnp, 151 SNPs commun. Le navigateur UCSC nous a également aidé à obtenir plus d'informations sur chaque SNP en cliquant simplement dessus. Par la suite, le navigateur nous amène vers une autre page qui nous fournit différents détails sur chaque « rs » tels que la position chromosomique, la classe du variant (single , indel..), et la fonction (5'UTR, intronic..)

Chapitre 3 : Résultats



Figure 45: Affichage des données dans le navigateur UCSC.

<http://genome.ucsc.edu/cgi-bin/hgTracks>

la figure 45 nous a permis de visualiser 805 SNPs communs dont 3 rs codants synonymes (vert), 7rs codants non-synonymes (rouge), 3 rs des régions non traduites (bleu) et 780 rs des régions introniques.

3.5 Résultats de l'OUTIL LitVar

LitVar nous a permis de visualiser et voir la relation entre les variantes polymorphe et les pathologies associées. Nous avons identifié 49 polymorphismes du gène NOX4 dont 2 SNPs ont une association au diabète « rs11018628 et rs1836882 » (Tableau VIII). Selon les auteurs s.c.Lin 2009 et Hyeon 2019.

Tableau VIII: Polymorphismes de NOX4 associés au diabète selon LitVar.

SNP	GENE	POSITION Ch	ALLELE	LOCALISATION	MALADIES
rs11018628	NOX4	11:89473343	T>C / T>G	Intron 2-3	INFARCTUS CEREBRAL DIABETE SUCRE HYPERTENSION SYNDROME HEPATOPULMONAIRE INSUFFISANCE RENALE CHRONIQUE
rs1836882	NOX4	11:89498993	T>C		LESIONS HEPATIQUES INDUITES PAR DES PRODUITS CHIMIQUES ET DES MEDICAMENTS MALADIES DU FOIE DIABETE SUCRE

3.6 Résultat de la sélection des variants (SNPs).

Dans notre étude nous nous sommes intéressées à identifier les « rs » non exploités dans des cas cliniques ou dans des études cas-témoins liée au diabète. L'analyse « *in silico* » par l'outil SIFT, nous a permis de prédire les changements d'acides aminés qui affectent la fonction de la protéine NOX4.

Parmi 41912 SNP introduit dans le navigateur SIFT, 130 SNPs ont abouti à des « rs » candidats avec des variants polymorphiques affectant les différentes régions du gène (figure 46).

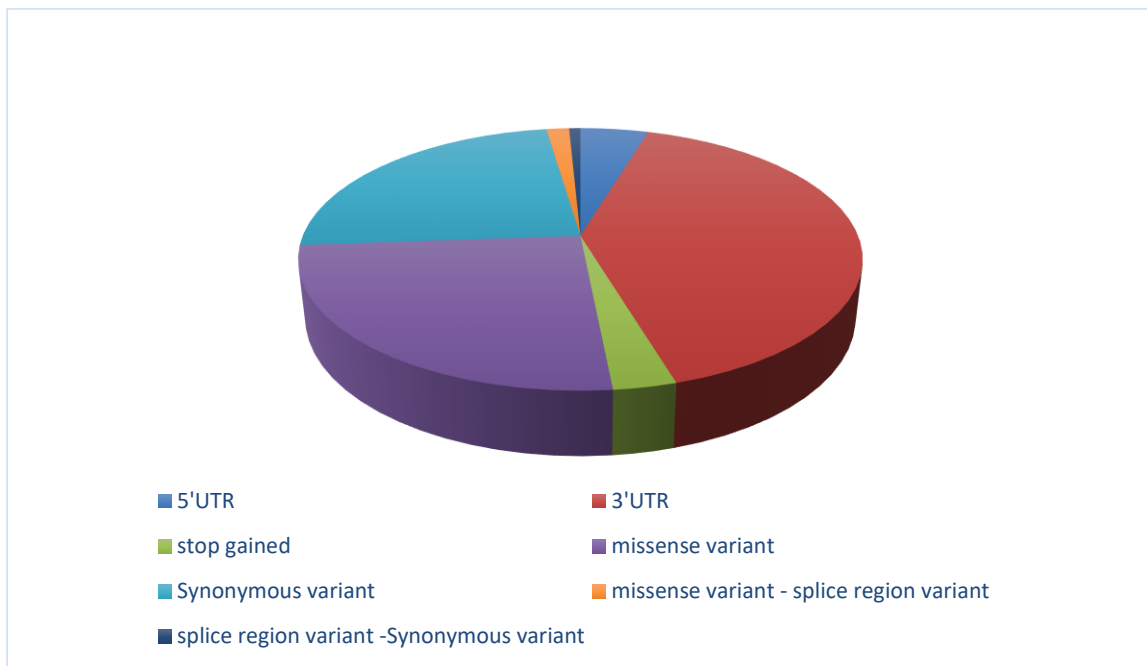


Figure 46: Diagramme circulaire représentant les variations nucléotidiques du gène NOX4 détectées par SIFT.

les SNPs de la région 3'UTR étaient les plus fréquents (n=53 ; 41%) par rapport au variants mi-sens (faux-sens) (n= 33 ; 25%), variants synonymes (n=31 ; 24%), gain stop (n= 4 ; 3%), variant mi-sens de la région d'épissage (n=2 ; 2%) et variant synonyme de la région d'épissage (n=1 ;1%).

Ensuite, après consultation de nos résultats on a sélectionné quatre SNPs pour :

- Identifier le type de substitution d'acide aminée qui va être apporté à la protéine (l'outil SIFT).
- Prédire l'effet délétère ou neutre de cette dernière (SIFT, PolyPhen-2, SNAP2, PhD-SNP).
- Explorer la conséquence de cette substitution sur l'augmentation ou la diminution de la stabilité de la protéine (I-Mutant).
- Et enfin, prédire les effets probables de ces polymorphismes sur la fonction de la protéine NOX4 en utilisant MutPred et HOPE Project.

3.6.1 Identification fonctionnelle des mutations délétères

Les résultats de l'identification fonctionnelle des mutations délétères est représenté par le tableau IX. Nos résultats montrent 4 « rs » : dont 3 rs mi-sens (rs139341533, rs56061986, rs373220761) et 1 rs synonyme (rs148405952).

- Le rs139341533 est une substitution au niveau protéique se traduit pour l'acide aminé de la leucine en phénylalanine en positions 72 et 118.
- Le rs56061986 est une substitution au niveau protéique se traduit pour l'acide aminé la serine en glycine en position 91, tous les deux sont localisé dans l'exon 4.
- Le rs373220761 localisé dans l'exon 16, c'est une substitution au niveau protéique se traduit pour l'acide aminé la tyrosine en cystéine en position 512 et 184,

Le rs synonyme est situé dans l'exon 5 du gène NOX4. C'est une substitution d'asparagine en lysine en position 126. Ces rs été traitées par les outils SIFT, PolyPhen-2, SNAP2 pour explorer leur effet délétère ou toléré (tableau IX).

Tableau IX: Prédictions de SIFT, PolyPhen-2, SNAP2 pour les rs de NOX4.

rs ID	Changement De AA	SIFT	score	PolyPhen-2	score	SNAP2	score
rs139341533	L72F	Toléré	0.177	Probablement dommageable	1.000	Neutre	-22
rs56061986	S91G	Toléré	0.55	Bénigne	0.007	Neutre	-65
rs373220761	Y512C	Délétère	0.003	Eventuellement dommageable	0.647	Neutre	-19
rs148405952	N126K	Toléré	0.156	Probablement dommageable	0.993	Ayant un effet	68

(Prédiction SIFT : score délétère $\leq 0,05$; score toléré $> 0,05$, Prédiction PolyPhen2 : bénigne = 0,000–0,004 ; éventuellement dommageable = 0,401–0,800 ; probablement dommageable = 0,801–1,000. SNAP2 : -100 (SAV fortement prédit comme neutre) à +100 (SAV fortement prédit comme effet - soit délétère, soit bénéfique), SAV : Single amino acide variant.

Selon les résultats, 3 de ces polymorphismes se sont avérés tolérés (score $> 0,05$) et un était délétère (score $\leq 0,05$). Avec polyPhen-2 deux mutations été considéré probablement dommageable (0,801–1,000), une éventuellement dommageable (= 0,401–0,800) et une bénigne (0,000–0,004). Avec SNAP2 trois mutations étaient prédits come neutre et une ayant un effet.

3.6.2 Prédiction de l'effet de substitution d'acides aminés sur la stabilité de la protéine

L'effet de la substitution d'acides aminés dans le cas des quatre « rs » a fait l'objet d'une analyse par l'outil PhD-SNP (prédicteur des SNP délétères) puis I-Mutant. Cette analyse nous a permis d'explorer leurs impacts sur la stabilité de la protéine (tableau X).

Chapitre 3 : Résultats

Tableau X:Prédiction des SNPs délétères et analyse de leurs effets sur la stabilité de la protéine.

rs ID	Changement De AA	PhD-SNP	I-Mutant	DDG
rs139341533	L118F	*Neutral	¹ Decrease Stability	-0.23
rs56061986	S91G	Neutral	² Increase stability	0.08
rs373220761	Y184C	**Disease	Neutral stability	0.35
rs148405952	N126K	Neutral	-	-

*Neutral : neutre, **Disease : polymorphisme pathologique ; ¹Diminution de la stabilité ; ²Augmentation de la stabilité ; DDG<0 : Diminuer la stabilité ; DDG>0 : Augmenter la stabilité.

L'outil PhD-SNP classe la mutation comme un polymorphisme pathogène ou neutre. Parmi les SNPs analysés, Y184C était prédit comme pathogène et les autres substitutions étaient prédites comme neutres (tableau X). Les résultats du test de stabilité des 4 mutations concernant les valeurs d'énergie libre selon le serveur I-Mutant 2.0 a révélé que les 4 modèles mutants ont montré un changement d'énergie libre (DDG/DDG).

3.6.3 Prédiction des effets sur les domaines protéiques par MutPred.

Les quatre mutations été analysé avec plus de détails par MutPred, toutes les mutations avait des scores délétères et ont marqués des changements de mécanismes moléculaires (tableau XI).

Tableau XI:Prédiction du serveur MutPred pour l'effet des substitutions d'acides aminés de NOX4.

Substitution	Score MutPred	Mécanisme moléculaire prédit (valeur P)
L72F	0.635	Perte d'hélice (P=0.01)
S91G	0.533	Gain de brin (P=0.03) Une altération de l'interface désordonnée (P=0.03)
Y512C	0.860	Protéine transmembranaire altérée (P=9.7e-05) Altération de l'interface désordonnée (P=0.03) Altération de la liaison métallique (P=0.02) Perte du site allostérique à Y512 (P=0.02) Perte de l'acétylation à K511 (P=0.02)
N126K	0.904	Une altération de la fixation des métaux (P=6.2e-03) Perte du site catalytique à H123 (P=0,03) Altération de la protéine transmembranaire (P=0,04) Gain de la glycosylation N-liée à N129 (P=0,03)

(Probabilité de score de mutation délétère $\geq 0,500$).

Le score de mutation délétère probable ($\geq 0,500$) détecté pour les substitutions L72F, S91G, Y512C, N126K était respectivement de 0.635, 0.533, 0.860, 0.904. Le tableau V montre que la prédiction des mécanismes moléculaires avec leur P-valeur a démontré que la mutation L72F provoque une perte d'hélice. Nos résultats montrent que la mutation Y512C peut altérer la protéine transmembranaire et avoir également un effet sur la perte d'acétylation à K511. Cette mutation peut aussi altérer la liaison métallique. Alors que, la mutation N162K provoque une perte du site catalytique à H123 et un gain de glycosylation à N129.

3.6.4 Résultat Hope Project

L'outil Hope Project a pour but de prédire les effets probables des polymorphismes (rs139341533 ; rs56061986 ; rs373220761 ; rs148405952) sur la fonction de la protéine NOX4.

Rappelons que pour visualiser avec plus de détails l'effet structural de ces quatre polymorphismes, nous avons utilisé le navigateur HOPE, qui nous a fourni plusieurs informations concernant les caractéristiques physiques et chimiques du résidu mutant et de type sauvage, Qui en conséquence, nous a permis de connaître les prédictions sur l'effet de ces substitutions sur la fonction de la NOX4 qui sont les suivantes (figure 47).

- En position 72, la substitution C/A du gène NOX4 «rs139341533 » cause la mutation de leucine en phénylalanine
- En position 126, la substitution A/C du gène NOX4 «rs148405952 » cause la mutation de l'asparagine en lysine.
- En position 512, la substitution T/C du gène NOX4 «rs373220761 » cause la mutation d'une Tyrosine en Cystéine
- En position 91, la substitution T/C «rs56061986 » du gène NOX4 cause la mutation d'une Sérine en Glycine.

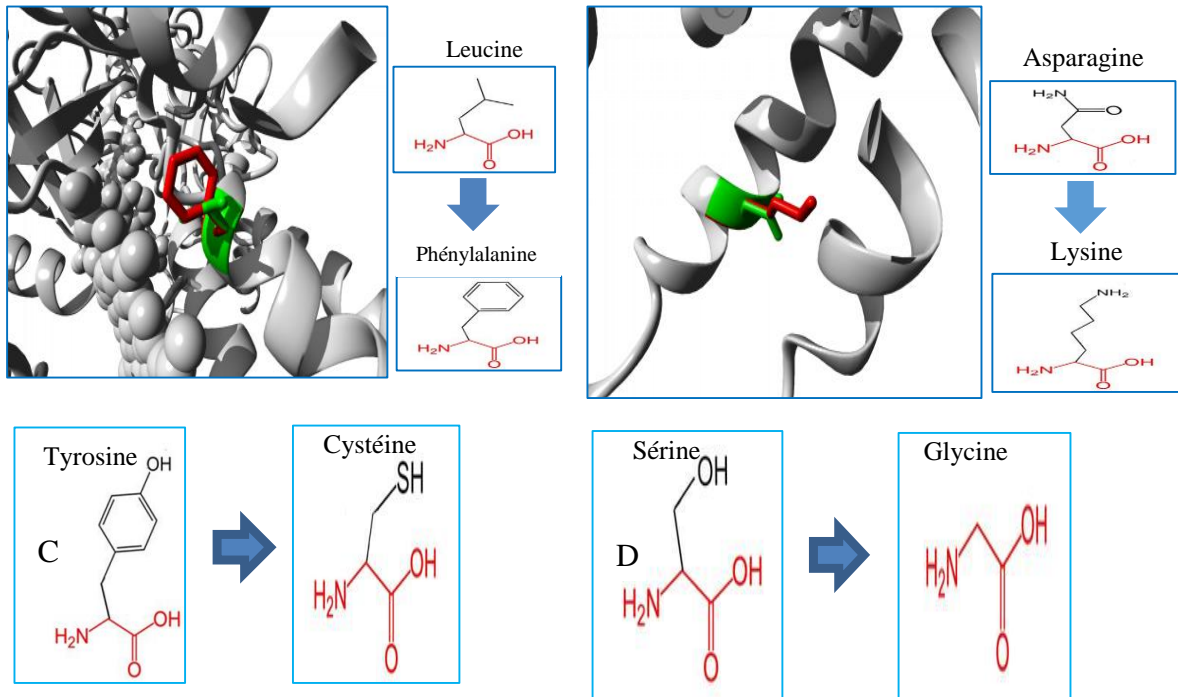


Figure 47: Close-up sur la mutation. La protéine est colorée en gris, les chaînes latérales du résidu de type sauvage et du résidu mutant sont représentées et colorées respectivement en vert et en rouge.

A : Close-up sur la mutation «rs139341533 » en position 72, la substitution C/A du gène NOX4; B : Close-up sur la mutation «rs148405952 » en position 126, la substitution A/C du gène NOX4 C : la structure schématique de l'acide aminé d'origine (à gauche) et du mutant (à droite).de la mutation «rs373220761 » en position 512, la substitution T/C du gène NOX4 ;D la structure schématique de l'acide aminé d'origine (à gauche) et du mutant (à droite).de la mutation «rs56061986 » en position 91, la substitution T/C du gène NOX4

3.7 Prédiction in silico de la structure 3D du gène NOX4

L'étude de la prédiction in silico de la structure 3D la protéine qui code pour le gène NOX4 et des ligands associés a fait de deux étapes indispensables : SUISSE MODEL et le Docking moléculaire.

3.7.1 Résultat de l'outil SUISSE MODEL

L'outil SUISSE MODEL permet de prédire in silico la structure 3D des protéines. La connaissance de la structure d'une protéine est une étape importante pour mieux comprendre la fonction de la protéine étudiée. Rappelons que l'outil SUISSE MODEL se base sur la modélisation comparative de la protéine qui code pour le gène NOX4 (figure 48.49). De plus, la modélisation en structure 3D de cette dernière avait pour but la connaissance des différents acides aminés qui sont directement liées à la fonction de la protéine « NOX4 » plus particulièrement au niveau du site actif.

Au total, 249 modèles ont été trouvés pour correspondre à la séquence cible (séquence de la protéine NOX4). Cette liste a été filtrée par une méthode heuristique jusqu'à 50. Le modèle que nous avons choisi est celui de la *DUOX1-DUOX1* (7D3E) Synthase avec un pourcentage de recouvrement de 29,47% par rapport à la séquence cible.

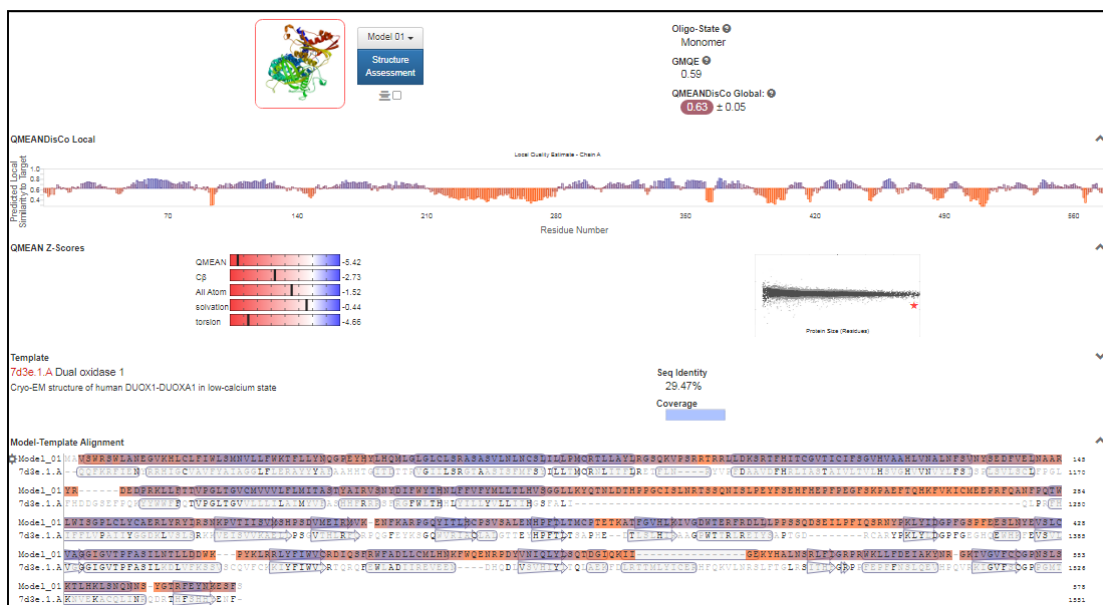


Figure 48:Modélisation comparative de la protéine NOX4 à l'aide de SWISS-MODEL.

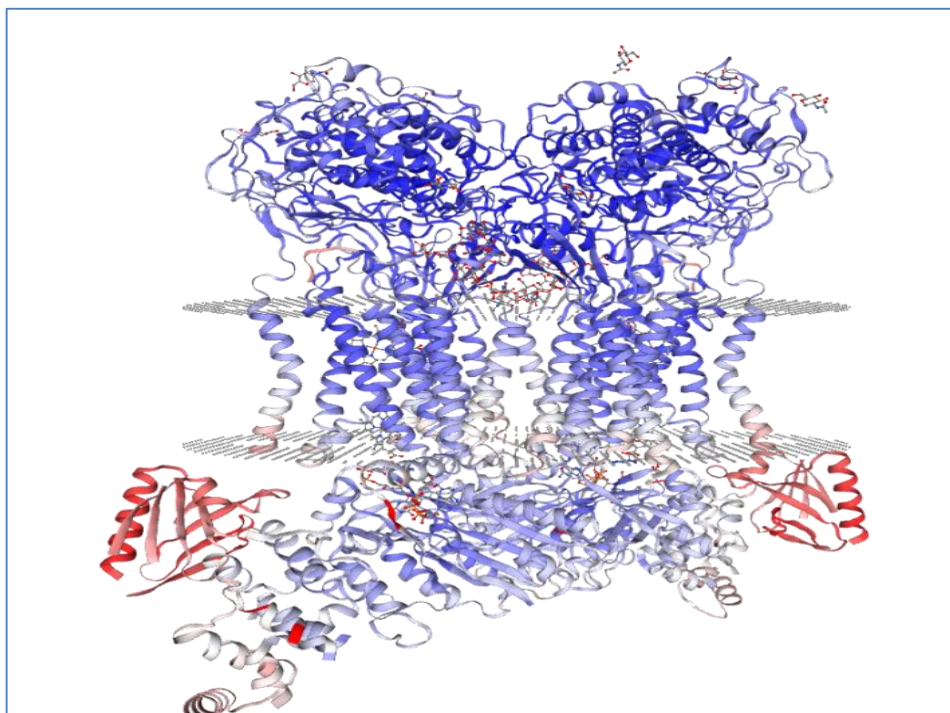


Figure 49: Structure 3D de la protéine 7D3E qui code pour le gène NOX4 présenté par SWISS-MODEL.

Afin d'étudier l'effet direct du polymorphisme sur les interactions de la NOX4 avec ces cibles, on doit tout d'abord étudier son comportement à état native (non muté). Dans cette partie, nous avons exploré par modélisation et Docking moléculaire les différentes interactions de la protéine NOX4 avec son ligand Flavin adenine dinucleotide (FAD).

3.8 Résultat du Docking Moléculaire

Le Docking moléculaire « ou amarrage moléculaire » a pour but de déterminer le mode d'interaction d'un complexe formé de deux ou plusieurs molécules, en cherchant des orientations dans l'espace et des conformations favorables pour la fixation d'un ligand à un récepteur. Ainsi, le Docking de chaque ligand génère un ensemble de modes de liaison appelés « poses » pour lesquelles un score est calculé. Ce score est une approximation de l'énergie de liaison du complexe protéine-ligand (Kcal/mol). Ainsi, les résultats du Docking moléculaire se présentent en trois étapes :

3.8.1 Préparation de la cible (enzyme)

Les résultats de la préparation de la cible (site actif) de la protéine (7D3E) de la NOX4 (humain), nous a permis d'obtenir le fichier pdb (Protein data bank) à partir du lien <https://www.rcsb.org/> de la base de données RCSB PDB.

En consultant les informations de la structure 3D de la protéine (7D3E) NOX4 sur la base de données RCSB PDB et SWISS Model, nous avons constaté que le ligand FAD interagit avec les chaînes A et C. ainsi, lors de la préparation de la protéine, nous avons modélisé les autres chaînes par l'outil PyMol pour visionner le site actif.

Cependant, les résultats des acides aminés impliqués dans le site actif a été confirmé par Castp (<http://sts.bioe.uic.edu/castp/index.html?1yys>). Le tableau XII ci-après montre que le le site actif du ligand FAD est formé de 21 acides aminés (tableau XII) (figure 50).

Tableau XII: Résidus de site actif de NOX4.

Potine	Chaîne	Nombre de AA	Ligand	Résidus
7D3E	A	21	FAD	R.1113, D.1124, V.1127, D.1128, R.1131, R.1214, W.1305, Y.1318, H.1319, P.1320, F.1321, T.1322, H.1335, I.1336, R.1337, A.1339, G.1340, P.1341, W.1342, T.1343, T.1393

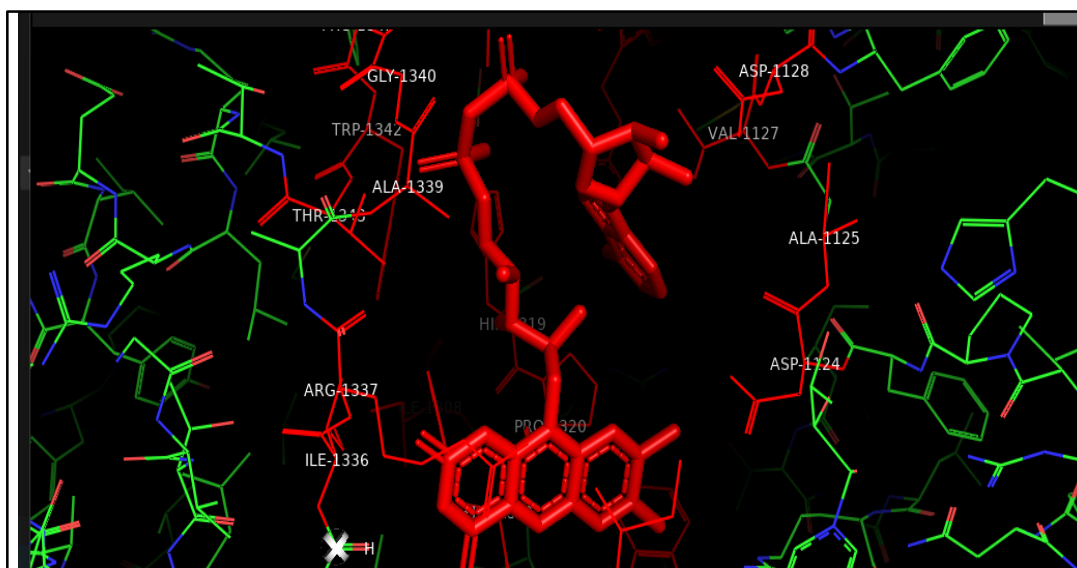


Figure 50: Acides aminés formant le site actif selon PyMOL.

3.8.2 Préparation du ligand (FAD)

Cette étape concerne la modélisation et l'optimisation de la géométrie du ligand par l'outil Autodock tools. Cette dernière vise à rechercher l'énergie minimale (énergie la plus faible) du ligand FAD qui a pour but de visualiser la dynamique moléculaire.

D'après la comparaison d'énergie de score des différentes « poses » du ligand, nous avons sélectionné la conformation 1 qui présente l'énergie minimale ayant pour SCORE = -3.03 Kcal/mol (tableau XIII). Selon les résultats obtenus, on note que les valeurs de RMSD sont considérées comme idéales (RMSD < 1.5) pour toutes les conformations du ligand FAD.

Tableau XIII: Résultats énergétiques de ligand de référence FAD.

Ligand	Score (kcal/mol)	Cluster RMSD (kcal/mol)	vdW + Hbond + desolv Energy	Inhibition Constant, Ki
FAD	-3.03	0.00	-7.33	6.0 M

RMSD : Root-Mean-Square Deviation (déviation de la racine de la moyenne des carrés) ;vdW :rayons de *van der Waals* des atomes ;

Les résultats concernant la position du site actif et le calcul d'énergie est représenté par le tableau XIV voire (annexe7, fig65). Rappelons que le paramétrage des coordonnées du site actif a été généré par l'outil Grid Box (Autodock tools). Le but de cet outil est de connaître les différentes conformations possibles pour le docking.

Chapitre 3 : Résultats

Tableau XIV : les coordonnées de Grid Box pour l'amarrage du NOX4-FAD.

Spacing (Angström)		1,000
Center Grid Box		<off set>
X center	0,983	1,611
Y centre	116,558	-0,333
Z centre	127,011	-1,722

Angstrom= 10^{-4} micromètre

3.8.3 Docking moléculaire

Après modélisation du ligand et de la cible, nous avons cherché les positions éligibles du ligand. Ainsi l'outil Pymol nous a permis de modéliser pour former le complexe (Protein-Ligand). Celui-ci va adopter la conformation la plus stable correspondant au niveau énergétique le plus faible en adoptant plusieurs poses. La conformation prend le RMSD (Root-Mean-Square Deviation) comme valeur inférieure à 2Å. C'est la conformation recherchée lors des simulations. Ces résultats sont répertoriés par le tableau XV.

Tableau XV: Résultats de SCORE et RMSD de docking moléculaire.

Conformation	Score (kcal/mol)	Cluster RMSD	Reference RMSD
1	-3.03	0.00	189.10
2	-1.34	0.00	189.29
3	-1.27	0.00	195.84
4	-1.26	0.00	190.33
5	-0.81	0.00	196.76
6	-0.69	0.00	186.11
7	-0.56	0.00	187.15
8	-0.16	0.00	196.13
9	+0.41	0.00	184.33
10	+0.66	0.00	194.66

Chapitre 3 : Résultats

Ainsi deux types de liaisons d'interactions direct ont été identifiés avec l'outil « Discovery studio visualizer » et des interactions classiques des liaisons hydrogène et des interactions Pi-cation ligand. Le tableau XVI montre les différents types d'interactions entre les acides aminés du site actif et le ligand comme les pond d'hydrogène, les liaisons électrostatiques, et les liaisons cationiques non covalente (π) (Pi-cation).

Tableau XVI: Rapport des interactions entre les résidus du site actif avec le ligand FAD.

Nom	Distance	Catégorie	Types	De	Donneur	à
A:PHE1321:HN - ligand:O	2,81427	Hydrogène Bond	Conventional Hydrogen Bond	A:PHE1321:HN	H-Donneur	ligand:O
ligand:O - A:HIS1319:O	3,1837	Hydrogène Bond	Conventional Hydrogen Bond	ligand:O	H-Donneur	A:HIS1319:O
ligand:O - A:PHE1321:O	3,24344	Hydrogène Bond	Conventional Hydrogen Bond	ligand:O	H-Donneur	A:PHE1321:O
ligand:O - A:TRP1305	4,0785	Electrostatic	Pi-Cation	ligand:O	Positive	A:TRP1305
ligand:N - A:TRP1305	4,23694	Electrostatic	Pi-Cation	ligand:N	Positive	A:TRP1305
ligand:N - A:TRP1305	3,59362	Electrostatic	Pi-Cation	ligand:N	Positive	A:TRP1305

Par la suite, l'analyse des résultats du docking avec Autodock Tools a montré que le ligand FAD interagit directement avec trois résidus TRP1305, HIS1319 et PHE 1321 de la protéine (7D3E) NOX4 (figure 51). De plus, l'analyse avec Discovery studio nous a permis d'identifier la nature de ces interactions (figure 52).

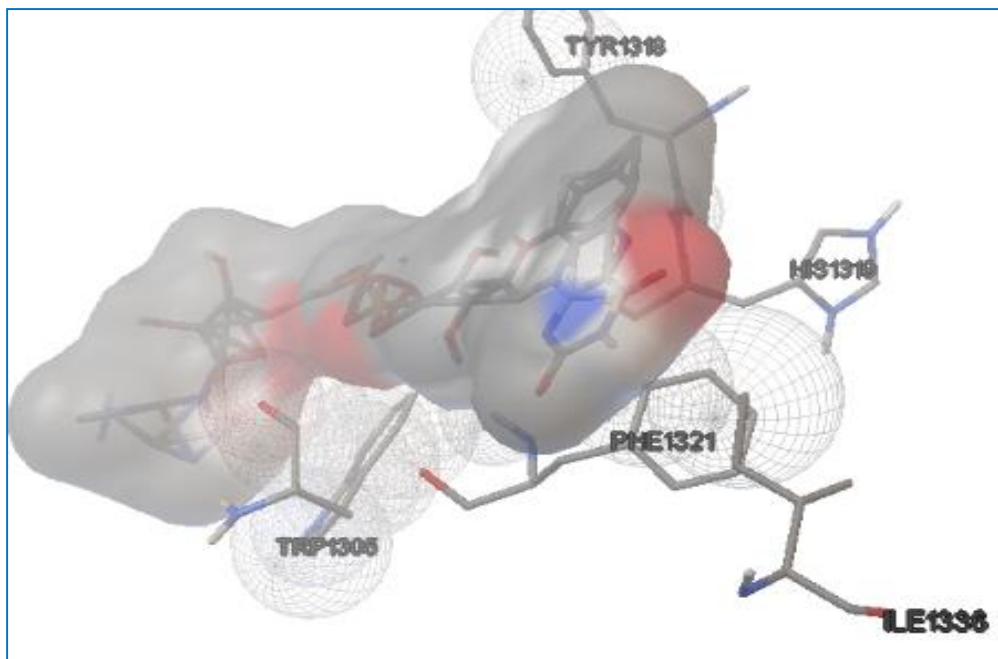


Figure 51:L'interaction entre ligand FAD et les résidus de site actif de NOX4 selon Autodock Tools.

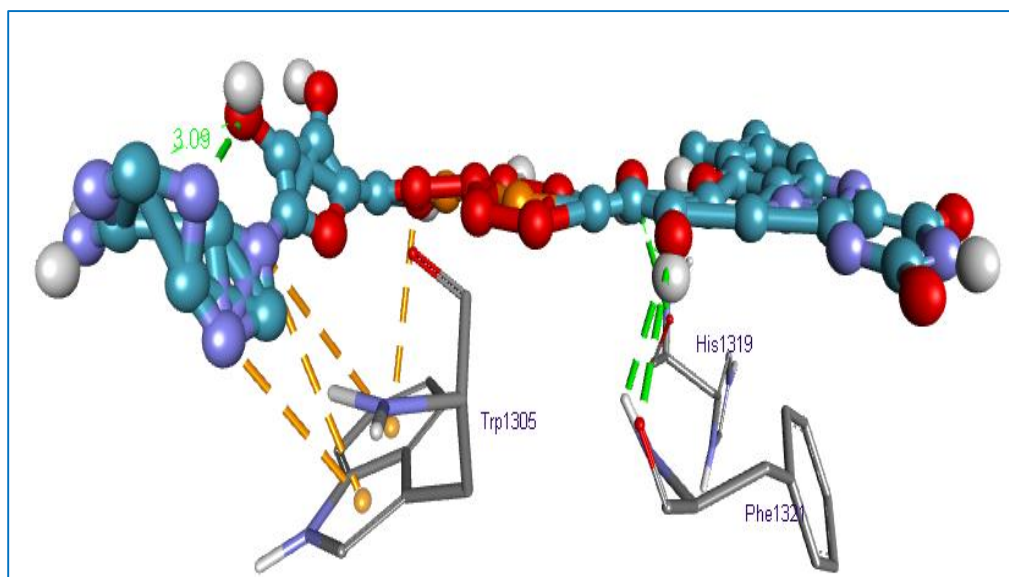


Figure 52:L'interaction entre ligand FAD et les résidus de site actif de NOX4 selon Discovery studio visualizer.

DISCUSSION

4. Chapitre 4 : Discussion

Le nicotinamide adénine dinucléotide phosphate oxydase 4 (NOX4) est un sous-type de la famille NOX, est un complexe enzymatique membranaire appartenant à la classe des oxydoréductases. C'est l'enzyme clé de l'explosion oxydative, où elle catalyse la réaction d'oxydation du NADPH par l'oxygène O₂, ce qui produit du NADP⁺, de l'H⁺ et de l'O₂⁻ (**Brown & Griendling, 2009**). Et sont la seule famille d'enzymes connue dont la seule fonction est de produire des ROS.

Rappelons que toutes les isoformes de NOX possèdent six hélices alpha transmembranaires avec des extrémités N et C-terminaux cytosoliques. Elles sont exprimées et régulées de manière différentielle dans divers tissus et ont des localisations subcellulaires différentes (**Altenhöfer et al., 2012**).

Il faut également prendre en considération que le stress oxydatif est considéré comme un risque de plusieurs pathologies comme le diabète les maladies cardiovasculaire les maladies rénales maladies pulmonaires des autres maladies inflammatoires. Les NADPH -oxydases de la famille Nox sont importants producteurs d'espèces réactives de l'oxygène ROS.

Nox4, l'isoforme majeure du NADPH dans le rein, Nox4 contribue aux processus redox impliqués dans la néphropathie diabétique, et d'autres maladies rénales en activant plusieurs voies de signalisation. La NOX4 se trouve dans une variété de types cellulaires y compris le rein, les cellules épithéliales, les podocytes, les cellules mésangiales, les cellules endothéliales et les fibroblastes, cellules musculaires lisses, cœur, pancréas tissu adipeux, placenta muscles squelettique, ostéoclastes, fibroblastes, astrocytes et neurones. Cependant, son rôle n'est pas clair et même controversé. Aussi, le NOX4 induit l'apoptose cellulaire, l'inflammation ou la fibrogènes. (**Yang et al., 2018**)

Selon Mahadev et coll en 2004, les voies de signalisations associées à NOX4 sont compliquées. Le NOX4 joue un rôle d'amplificateur des voies de signalisations de l'insuline par l'intermédiaire des Radicaux Libres Oxygène (FRO). Ces dernières inhibent l'activation de tyrosine phosphatases, notamment PTP1B impliquée dans la résistance à l'insuline (**Mahadev et al., 2004**), et la voie de signalisation AGE-RAGE dans les complications du diabète néphropathie diabétique (**Kanwar et al., 2011**).

Rappelons aussi que notre étude *in silico* été réalisé pour identifier l'association du polymorphisme du gène NADPH oxydase 4 (NOX4) au diabète chez l'homme. Les résultats que nous avons obtenus, nous ont permis d'identifier plusieurs types de polymorphisme associant le NOX4 et le diabète. Ainsi, d'après les résultats de NCBI jusqu'à aujourd'hui, seulement 59 études sont associées au gène NOX4 <https://www.ncbi.nlm.nih.gov/>.

Chapitre 4 : Discussion

Différentes études d'associations polymorphiques sur le diabète ont montré que le gène NOX4 joue un rôle dans l'apparition et le développement de cette maladie. Cependant, ces études ne se sont intéressées qu'à 8 SNPs qui montrent une association aux complications de diabète dont la néphropathie diabétique, la rétinopathie diabétique et le NODAT (Nouvelle apparition du diabète après une transplantation). Ces SNPs sont le rs1836882, rs614128, rs490934, rs3017887, rs553635, rs10765219, rs11018670, rs3913535. **(Hwang et al., 2019) (Lim et al., 2009) (Meng et al., 2018)**.

En outre, Hwang et coll en 2019, ont rapporté que le rs1836882 C>T était également associé à un risque de développement de la NODAT (OR 1.06, IC 95 % (0.85, 1.33); $p = 0.588$). Ce polymorphisme était associé avec d'autres gènes. En effet, les gènes ATP5F1P6, CELA2B, CASP9, NOX4 et INPP5A ont été identifiés comme des gènes à risque chez les transplantés rénaux caucasiens. Ces variantes génétiques ont été impliquées dans les voies d'apoptose cellulaire, mais pas dans la résistance à l'insuline, ce qui suggère que l'apoptose cellulaire est un élément essentiel de la pathogenèse de la NODAT **(Hwang et al., 2019)**.

De plus, l'analyse des haplotypes des différents SNPs (rs614128G>C, rs490934G>C, rs3017887C>A et rs553635C>T) ont permis d'identifier différents Loci : GGCC avec une fréquence de 0,776% ayant une susceptibilité hautement significative à une néphropathie diabétique ($p=0,0055$) entre hétérozygote (OR =2,05 ; IC₉₅ % = 1,04-4,06) et homozygote (OR=2,48 ; IC₉₅ % = 1,27-4,83). Rappelons que l'haplotype est un groupe d'allèles de différents loci situés sur un même chromosome et habituellement transmis ensemble) **(Lim et al., 2009)**.

Les SNP : rs10765219 G>T, rs11018670A>G, rs3913535A>C>G, Dans le gène de la NADPH oxydase 4 (NOX4) ont une association hautement significative au diabète avec $p = 4.05 \times 10^{-9}$ Aussi, deux SNP voisins, rs10765219 et rs11018670, ont également montré une association et une susceptibilité hautement significative.

Dans la méta-analyse utilisant plusieurs cohortes caucasiennes (à l'exclusion de GoDARTS), les rs10765219 et rs11018670 ont montré des associations hautement significatives pour la rétinopathie diabétique ($p = 0,003$ et $p=0,007$, respectivement). Alors que le rs3913535 l'association n'était pas significative ($p = 0,429$) **(Meng et al., 2018)**.

Les preuves que les isoformes Nox1, Nox2 et Nox4 ont un rôle causal dans divers troubles ont continué à se développer **(Touyz et al., 2019)**. Il a été rapporté que l'inactivation de Nox4 par l'ARNsi délivré par l'adénovirus réduisait la néovascularisation rétinienne dans l'OIR, et

Chapitre 4 : Discussion

l'inhibition de Nox1/4 atténuaient la néovascularisation rétinienne et les niveaux de VEGF dans ce modèle (**Wilkinson-Berka et al., 2014**). L'inhibition de Nox1/4 a de puissantes actions anti-inflammatoires dans la rétinopathie (**Wilkinson-Berka et al., 2014**) et d'autres contextes tels que l'athérosclérose et la néphropathie associées au diabète sucré (**Gray et al., 2017**), (**Di Marco et al., 2014**), (**Gorin et al., 2015**).

Les résultats de l'outil Ensembl, nous montrent non seulement le nom de chaque SNPs du gène NOX4 mais aussi sa localisation et le type de mutation en associant les résultats précédant présentés par LIT VAR et SIFT.

L'outil LitVar nous confirment les résultats précédents et montrent que le gène NOX4 est lié à de nombreuses pathologies (57 pathologies) de ces 57 maladies le diabète sucré.

<https://www.ncbi.nlm.nih.gov/CBBresearch/Lu/Demo/LitVar/#!?query=nox4>

Rappelons que dans notre étude nous nous sommes intéressées à identifier les « rs » non exploités dans des cas cliniques ou dans des études cas-témoins liée au diabète. L'analyse « *in silico* » par l'outil SIFT <https://sift.bii.a-star.edu.sg/>, nous a permis de prédire les changements d'acides aminés qui affectent la fonction de la protéine NOX4.

Après consultation de nos résultats on a sélectionné quatre SNPs 3 rs mi-sens (rs139341533, rs56061986, rs373220761) et 1 rs synonyme (rs148405952).

L'identification du type de substitution d'acide aminée nous a permis de prédire l'effet délétère ou neutre de la protéine, explorer la conséquence de cette substitution sur l'augmentation ou la diminution de la stabilité de la protéine, Et enfin, prédire les effets probables de ces polymorphismes sur la fonction de la protéine (NOX4)

Les résultats du test de stabilité des 4 mutations concernant les valeurs d'énergie libre selon le serveur I-Mutant 2.0 <https://folding.biofold.org/i-mutant/i-mutant2.0.html> a révélé que les 4 modèles mutants ont montré un changement d'énergie libre (DDG/DDG). La valeur DDG négative suggère que la substitution d'acides aminés est délétère pour la stabilité de la protéine. Les valeurs été comprises entre -0,23 et 0,35 kcal/Mol.

Concernant la prédiction de l'effet de substitution par l'outil I-Mutant, Il semble que L118F est susceptible de diminuer la stabilité de la protéine, S91G semble causé une stabilité élevée, alors Y184C semble causé une stabilité neutre pour former la structure protéique. En revanche, N126K ne présente aucun effet sur la protéine. En effet, les 4 SNPs cités précédemment suggèrent un effet de diminution ou d'augmentation de la stabilité de la protéine du gène NOX4. Qui en conséquence, peut mener à une perte de fonction.

Chapitre 4 : Discussion

Les résultats de HOPE Project <https://www3.cmbi.umcn.nl/hope/> nous ont permis de renforcer notre opinion sur la nécessité d'explorer ces quatre SNPs de NOX4 dans des cas cliniques, ainsi que leurs associations au diabète.

La mutation C/A « rs139341533 » de gène NOX4 est situé dans une région annotée dans la base de données UniProt en tant que domaine transmembranaire. Le résidu mutant est plus grand que le résidu type sauvage. Cette différence de taille peut affecter les contacts avec la membrane lipidique. La mutation est située dans un domaine, annoté dans UniProt comme Oxydoréductase ferrique, la mutation introduit un acide aminé aux propriétés différentes (www.uniprot.org)

La structure 3D, suggère que le résidu de type sauvage a des interactions avec un ligand annoté comme HEM, La différence de propriétés entre le type sauvage et la mutation peut facilement entraîner une perte d'interactions avec le ligand. Parce que la liaison au ligand est souvent importante pour la fonction de la protéine, cette fonction pourrait être perturbée par cette mutation qui peut perturber ce domaine et inhiber sa fonction.

Il en est de même pour la mutation A/C «rs148405952» du gène NOX4 est située dans un domaine, annoté dans UniProt comme Oxydoréductase ferrique. La mutation introduit un acide aminé aux propriétés différentes, qui peut perturber ce domaine et inhiber sa fonction.

Nos résultats ont montré qu'il existe une différence de charge entre l'acide aminé de type sauvage et mutant. De plus, la charge de résidu de type sauvage était neutre et la charge de résidu mutant est positive. Il semble que le résidu mutant introduit une charge dans un résidu enfoui qui peut entraîner des problèmes de repliement des protéines. Les acides aminés de type sauvage et mutant diffèrent en taille, le résidu de type sauvage a été enfoui dans le noyau de la protéine. Le résidu mutant est plus gros et ne conviendra probablement pas.

Il en est de même aussi pour la mutation T/C «rs373220761» du gène NOX4 est située dans une étendue de résidus annotés dans UniProt comme une région spéciale : médie l'interaction avec TLR4. Ce qui suggère que les différences de propriétés des acides aminés peuvent perturber cette région et perturber sa fonction. Les acides aminés de type sauvage et mutant diffèrent en taille, le résidu mutant est plus petit, cela peut entraîner une perte d'interactions. L'hydrophobicité du résidu de type sauvage et mutant diffère, la mutation introduit un résidu plus hydrophobe à cette position. Cela peut entraîner une perte de liaisons hydrogène et/ou perturber le repliement correct.

Chapitre 4 : Discussion

La mutation T/C «rs56061986» du gène NOX4 est située dans un domaine, annoté dans UniProt comme Oxydoréductase ferrique (www.uniprot.org). La mutation introduit un acide aminé aux propriétés différentes, qui peut perturber ce domaine et abolir sa fonction, la mutation introduit une glycine à cette position. Les glycines sont très flexibles et peuvent perturber la rigidité requise de la protéine à cette position. Les acides aminés de type sauvage et mutant diffèrent en taille, le résidu mutant est plus petit, cela peut entraîner une perte d'interactions.

L'objet d'une modélisation en incluant un Docking moléculaire pour comprendre les effets direct et indirect sur les cibles liées à la protéine

Sachant que la NOX4 est une protéine transmembranaire qui assure plusieurs fonctions Complexes sa vraie structure jusqu'à maintenant reste inconnue. Rappelons que la modélisation de la NOX4 dans cette étude a été faite par homologie avec SWISS-MODEL. (<https://swissmodel.expasy.org/>). En conséquence, cela nous a permis de récupérer la structure homologue de la NOX4 (code PDB : 7D3E), sachant que le docking a été fait avec la forme native de la NOX4

L'outil Grid Box le paramétrage des coordonnées du site actif a été généré par l'outil Grid Box (Autodock tools). Dans le but de connaître les différentes conformations possibles pour le docking et donc plus il est négatif, meilleur est l'affinité prédite. Le ligand FAD représente alors souvent la référence idéale.

Calcul de RMSD D'après la comparaison d'énergie de score des différentes « poses » du ligand, nous avons sélectionné la conformation qui présente l'énergie minimale ayant pour SCORE = -3.03 Kcal/mol. Selon les résultats obtenus, on suggère que les valeurs de RMSD sont considérées comme idéales (RMSD < 1.5) pour toutes les conformations du ligand FAD.

D'après les rapports de docking précédents et les résultats obtenus, nous avons constaté que le ligand flavin adenine dinucleotide (FAD) étudié interagit réellement avec le site actif de la NOX4, plus précisément avec les résidus PHE1321, HIS1319, TRP1305.

À un niveau plus large, l'étude de docking révèle que la distance entre chaque ligand et les sites de liaison était comprise entre [2.81 Å et 4.23 Å]. D'après Anne Imbert et col, Les interactions ayant des distances comprises entre 2.5 Å et 3.1Å sont considérées comme fortes et celles comprises entre 3.1Å et 3.55Å sont supposées moyennes et quand leurs distances sont supérieures à 3.55Å elles sont considérées faibles. (Imberty et al., 1991)

Chapitre 4 : Discussion

Donc, selon ces résultats d'étude nous pouvons facilement constater que les distances entre les acides aminés du site active PHE1321, HIS1319 et le FAD ont une liaison forte d'interaction, 2.81 Å et 3.18 Å respectivement par rapport au résidu TRP 1305 qui a trois liaison faible d'interaction 3.24 Å , 4.07 Å et 4,32 Å successivement.

CONCLUSION

Conclusion

Le diabète est une maladie métabolique caractérisée par une élévation anormale chronique de la glycémie, causée par un dysfonctionnement de la sécrétion ou de l'action de l'insuline.

Le dysfonctionnement de l'enzyme NOX4 peut entraîner à des maladies chroniques comme le diabète et les complications associées aux diabètes.

Notre étude vise à identifier les différents variants du polymorphisme du gène NOX4 associés au diabète par méthode *in silico* en utilisant des outils bioinformatique.

Nos résultats confirment que certains variants polymorphiques du gène NOX4 peuvent entraîner une modification dans la structure de la protéine, pouvant aboutir à un dysfonctionnement et une perte de fonction.

Nos résultats montrent différents variants du gène NOX4 dont certains sont susceptibles d'être associés au diabète. Les outils SIFT et LIT VAR, nous ont permis d'identifier : 8 SNPs (rs1836882, rs614128, rs490934, rs3017887, rs553635, rs10765219, rs11018670, rs3913535). Ces derniers sont significativement associés aux diabètes plus précisément aux NODAT, à la néphropathie diabétique et la rétinopathie.

Nos résultats suggèrent aussi que le gène NOX4 peut jouer un rôle important dans la perturbation de la signalisation de l'insuline, à la résistance à l'insuline par la voie de signalisation AGE-RAGE.

Dans le cas des études cliniques ou dans des études cas-témoins liée au diabète. L'analyse « *in silico* » par l'outil SIFT, nous a permis de prédire les changements d'acides aminés qui affectent la fonction de la protéine NOX4, Nos résultats montrent 4 « rs » : dont 3 rs mi-séance (rs139341533, rs56061986, rs373220761), 1 rs synonyme (rs148405952) et Trois mutations été considéré probablement dommageable selon Polyphen-2.

Le résultat de Docking montre au niveau du site d'interaction protéine-ligand : 3 acides aminés liés à la fonction du NOX4 au niveau du site actif.

En perspective, nous proposons qu'il serait souhaitable de continuer ce travail en ciblant ces trois résidus TRP1305, HIS1319 et PHE 1321 de la protéine NOX4 et voir s'il existe un polymorphisme au niveau de ces sites, sachant qu'un polymorphisme au niveau de ces sites pourrait influencer considérablement la fonction de la protéine.

Conclusion

Aussi, les rs non cités dans les références bibliographiques peuvent faire l'objet d'une étude cas-témoins concernant la protéine nox4,

Jusqu'à présent il n'y a pas d'études sur le polymorphisme de NOX4 incluant les modifications au niveau de ces acides aminés, par contre il serait souhaitable de cibler ces trois résidus pour voir par le biais du Docking s'il existe une modification au niveau de la protéine incluant une mutation sachant qu'un polymorphisme dans ces trois acides aminés identifiés peut affecter considérablement la fonction de la protéine ce qui va permettre de développer de nouvelles cibles thérapeutiques.

Références bibliographiques

- Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, 7(4), 248–249.
- Adzhubei, I., Jordan, D. M., & Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. In *Current Protocols in Human Genetics* (Vol. 2, Issue SUPPL.76). <https://doi.org/10.1002/0471142905.hg0720s76>
- Allot, A., Peng, Y., Wei, C.-H., Lee, K., Phan, L., & Lu, Z. (2018). LitVar: a semantic search engine for linking genomic variant data in PubMed and PMC. *Nucleic Acids Research*, 46(W1), W530–W536.
- Altenhöfer, S., Kleikers, P. W. M., Radermacher, K. A., Scheurer, P., Hermans, J. J. R., Schiffers, P., Ho, H., Wingler, K., & Schmidt, H. H. H. W. (2012). The NOX toolbox: Validating the role of NADPH oxidases in physiology and disease. *Cellular and Molecular Life Sciences*, 69(14), 2327–2343. <https://doi.org/10.1007/s00018-012-1010-9>
- Atkinson, M. A., Eisenbarth, G. S., & Michels, A. W. (2014). Type 1 diabetes. *The Lancet*, 383(9911), 69–82.
- Bava, K. A., Gromiha, M. M., Uedaira, H., Kitajima, K., & Sarai, A. (2004). ProTherm, version 4.0: thermodynamic database for proteins and mutants. *Nucleic Acids Research*, 32(suppl_1), D120–D121.
- Belyaev, I., Dean, A., Eger, H., Hubmann, G., Jandrisovits, R., Kern, M., Kundi, M., Moshammer, H., Lercher, P., & Müller, K. (n.d.). *Lignes directrices 2016 de l'EUROPAEM pour la prévention, le diagnostic et le traitement des sujets atteints de problèmes de santé et de maladies en lien avec les champs électromagnétiques*.
- Bhatti, P., Church, D. M., Rutter, J. L., Struewing, J. P., & Sigurdson, A. J. (2006). Candidate single nucleotide polymorphism selection using publicly available tools: a guide for epidemiologists. *American Journal of Epidemiology*, 164(8), 794–804.
- Bottazzo, G., Florin-Christensen, A., & Doniach, D. (1974). Islet-cell antibodies in diabetes mellitus with autoimmune polyendocrine deficiencies. *The Lancet*, 304(7892), 1279–1283.
- Britannica, T. E. of E. (2019). single nucleotide polymorphism. In *Encyclopedia Britannica*. <https://www.britannica.com/science/single-nucleotide-polymorphism>
- Bromberg, Y., Yachdav, G., & Rost, B. (2008). SNAP predicts effect of mutations on protein function. *Bioinformatics*, 24(20), 2397–2398.
- Brown, D. I., & Griendling, K. K. (2009). Nox proteins in signal transduction. *Free Radical Biology and Medicine*, 47(9), 1239–1253.
- Capriotti, E., Fariselli, P., & Casadio, R. (2005). I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Research*, 33(suppl_2), W306–W310.
- Chaudhary, K. K., & Mishra, N. (2016). A review on molecular docking: novel tool for drug discovery. *Databases*, 3(4), 1029.
- Chen, F., Haigh, S., Barman, S., & Fulton, D. J. R. (2012). From form to function: the role of Nox4 in the cardiovascular system. *Frontiers in Physiology*, 3, 412.

Références bibliographiques

- Consortium, U. (2021). Bateman A., Martin M. J., Orchard S., Magrane M., Agivetova R., Ahmad S., Alpi E., Bowler-Barnett EH, Britto R., et Al. *UniProt: The Universal Protein Knowledgebase In*.
- Cooper, G. M., & Shendure, J. (2011). Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. *Nature Reviews Genetics*, *12*(9), 628–640.
- DeFronzo, R., Deibert, D., Hendler, R., Felig, P., & Soman, V. (1979). Insulin sensitivity and insulin binding to monocytes in maturity-onset diabetes. *The Journal of Clinical Investigation*, *63*(5), 939–946.
- Di Marco, E., Gray, S. P., Chew, P., Koulis, C., Ziegler, A., Szyndralewicz, C., Touyz, R. M., Schmidt, H. H. H. W., Cooper, M. E., Slattery, R., & Jandeleit-Dahm, K. A. (2014). Pharmacological inhibition of NOX reduces atherosclerotic lesions, vascular ROS and immune-inflammatory responses in diabetic Apoe(-/-) mice. *Diabetologia*, *57*(3), 633–642. <https://doi.org/10.1007/s00125-013-3118-3>
- Eddaikra, A. (2018). Association entre le polymorphisme des enzymes antioxydantes et la physiopathologie du diabète type1 chez la population Algérienne.
- Fareed, M., & Afzal, M. (2013). Single nucleotide polymorphism in genome-wide association of human population: a tool for broad spectrum service. *Egyptian Journal of Medical Human Genetics*, *14*(2), 123–134.
- Fernández, X. M., & Birney, E. (2010). Ensembl genome browser. In *Vogel and Motulsky's Human Genetics* (pp. 923–939). Springer.
- Forrester, S. J., Kikuchi, D. S., Hernandez, M. S., Xu, Q., & Griendling, K. K. (2018). Reactive oxygen species in metabolic and inflammatory signaling. *Circulation Research*, *122*(6), 877–902. <https://doi.org/10.1161/CIRCRESAHA.117.311401>
- George Priya Doss, C., Rajasekaran, R., Sudandiradoss, C., Ramanathan, K., Purohit, R., & Sethumadhavan, R. (2008). A novel computational and structural analysis of nsSNPs in CFTR gene. *Genomic Medicine*, *2*(1), 23–32.
- Gill, P. S., & Wilcox, C. S. (2006). NADPH oxidases in the kidney. *Antioxidants & Redox Signaling*, *8*(9–10), 1597–1607.
- Gorin, Y., Cavaglieri, R. C., Khazim, K., Lee, D.-Y., Bruno, F., Thakur, S., Fanti, P., Szyndralewicz, C., Barnes, J. L., Block, K., & Abboud, H. E. (2015). Targeting NADPH oxidase with a novel dual Nox1/Nox4 inhibitor attenuates renal pathology in type 1 diabetes. *American Journal of Physiology. Renal Physiology*, *308*(11), F1276-87. <https://doi.org/10.1152/ajprenal.00396.2014>
- Goyal, P. (2005). Weissmann N, Rose F, Grimminger F, Schafers HJ, Seeger W, Hanze J. *Identification of Novel Nox4 Splice Variants with Impact on ROS Levels in A549 Cells. Biochem Biophys Res Commun*, *329*, 32–39.
- Gray, S. P., Jha, J. C., Kennedy, K., van Bommel, E., Chew, P., Szyndralewicz, C., Touyz, R. M., Schmidt, H. H. H. W., Cooper, M. E., & Jandeleit-Dahm, K. A. M. (2017). Combined NOX1/4 inhibition with GKT137831 in mice provides dose-dependent reno- and atheroprotection even in established micro- and macrovascular disease. *Diabetologia*, *60*(5), 927–937. <https://doi.org/10.1007/s00125-017-4215-5>
- Guo, S., & Chen, X. (2015). The human Nox4: Gene, structure, physiological function and pathological significance. *Journal of Drug Targeting*, *23*(10), 888–896. <https://doi.org/10.3109/1061186X.2015.1036276>

Références bibliographiques

- He, W., Wang, Q., Gu, L., Zhong, L., & Liu, D. (2018). NOX4 rs11018628 polymorphism associates with a decreased risk and better short-term recovery of ischemic stroke. *Experimental and Therapeutic Medicine*, 16(6), 5258–5264.
- Health, N. I. of. (2019). What are single nucleotide polymorphisms (SNPs). *Genetics Home Reference-NIH. US National Library Of Medicine*. URL: <https://ghr.nlm.nih.gov/Primer/Genomicresearch/Snp>.
- Hinman, M. N., & Lou, H. (2008). Diverse molecular functions of Hu proteins. *Cellular and Molecular Life Sciences*, 65(20), 3168–3181.
- Hwang, H. S., Hong, K. W., Kim, J. S., Kim, Y. G., Moon, J. Y., Jeong, K. H., & Lee, S. H. (2019a). Validation of identified susceptible gene variants for new-onset diabetes in renal transplant recipients. In *Journal of Clinical Medicine* (Vol. 8, Issue 10). <https://doi.org/10.3390/jcm8101696>
- Hwang, H. S., Hong, K. W., Kim, J. S., Kim, Y. G., Moon, J. Y., Jeong, K. H., & Lee, S. H. (2019b). Validation of identified susceptible gene variants for new-onset diabetes in renal transplant recipients. *Journal of Clinical Medicine*, 8(10). <https://doi.org/10.3390/jcm8101696>
- Imberty, A., Hardman, K. D., Carver, J. P., & Perez, S. (1991). Molecular modelling of protein-carbohydrate interactions. Docking of monosaccharides in the binding site of concanavalin A. *Glycobiology*, 1(6), 631–642.
- International Diabetes Federation. (2015). *The International Diabetes Federation. Diabetes Atlas*.
- Inzucchi, S. E., Bergenstal, R. M., Buse, J. B., Diamant, M., Ferrannini, E., Nauck, M., Peters, A. L., Tsapas, A., Wender, R., & Matthews, D. R. (2012). Management of hyperglycemia in type 2 diabetes: a patient-centered approach: position statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes Care*, 35(6), 1364–1379.
- Ismail, S., & Essawi, M. (2012). Genetic polymorphism studies in humans. *Middle East Journal of Medical Genetics*, 1(2), 57–63.
- Jallut, D., Golay, A., Munger, R., Frascarolo, P., Schutz, Y., Jequier, E., & Felber, J. P. (1990). Impaired glucose tolerance and diabetes in obesity: a 6-year follow-up study of glucose metabolism. *Metabolism*, 39(10), 1068–1075.
- Jenuth, J. P. (2000). The ncbi. In *Bioinformatics methods and protocols* (pp. 301–312). Springer.
- Jia, M., Yang, B., Li, Z., Shen, H., Song, X., & Gu, W. (2014). Computational analysis of functional single nucleotide polymorphisms associated with the CYP11B2 gene. *PLoS One*, 9(8), e104311.
- Kanwar, Y. S., Sun, L., Xie, P., Liu, F. Y., & Chen, S. (2011). A glimpse of various pathogenetic mechanisms of diabetic nephropathy. *Annual Review of Pathology*, 6, 395–423. <https://doi.org/10.1146/ANNUREV.PATHOL.4.110807.092150>
- Kitchen, D. B., Decornez, H., Furr, J. R., & Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews Drug Discovery*, 3(11), 935–949.

Références bibliographiques

- Lambeth, J. D. (2004). NOX enzymes and the biology of reactive oxygen. *Nature Reviews Immunology*, 4(3), 181–189.
- Leabman, M. K., Huang, C. C., DeYoung, J., Carlson, E. J., Taylor, T. R., de La Cruz, M., Johns, S. J., Stryke, D., Kawamoto, M., & Urban, T. J. (2003). Natural variation in human membrane transporter genes reveals evolutionary and functional constraints. *Proceedings of the National Academy of Sciences*, 100(10), 5896–5901.
- Lee, J.-E., Choi, J. H., Lee, J. H., & Lee, M. G. (2005). Gene SNPs and mutations in clinical genetic testing: haplotype-based testing and analysis. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 573(1–2), 195–204.
- Li, J., Wang, J. J., Yu, Q., Chen, K., Mahadev, K., & Zhang, S. X. (2010). Inhibition of reactive oxygen species by Lovastatin downregulates vascular endothelial growth factor expression and ameliorates blood-retinal barrier breakdown in db/db mice: role of NADPH oxidase 4. *Diabetes*, 59(6), 1528–1538.
- Lim, S. C., Liu, J. J., Low, H. Q., Morgenthaler, N. G., Li, Y., Yeoh, L. Y., Wu, Y. S., Goh, S. K., Chionh, C. Y., Tan, S. H., Kon, Y. C., Soon, P. C., Bee, Y. M., Subramaniam, T., Sum, C. F., & Chia, K. S. (2009). Microarray analysis of multiple candidate genes and associated plasma proteins for nephropathy secondary to type 2 diabetes among Chinese individuals. *Diabetologia*, 52(7), 1343–1351. <https://doi.org/10.1007/s00125-009-1368-x>
- Lind, M., Svensson, A.-M., Kosiborod, M., Gudbjörnsdóttir, S., Pivodic, A., Wedel, H., Dahlqvist, S., Clements, M., & Rosengren, A. (2014). Glycemic control and excess mortality in type 1 diabetes. *New England Journal of Medicine*, 371(21), 1972–1982.
- Livingstone, S. J., Levin, D., Looker, H. C., Lindsay, R. S., Wild, S. H., Joss, N., Leese, G., Leslie, P., McCrimmon, R. J., & Metcalfe, W. (2015). Estimated life expectancy in a Scottish cohort with type 1 diabetes, 2008-2010. *Jama*, 313(1), 37–44.
- Mahadev, K., Motoshima, H., Wu, X., Ruddy, J. M., Arnold, R. S., Cheng, G., Lambeth, J. D., & Goldstein, B. J. (2004). The NAD (P) H oxidase homolog Nox4 modulates insulin-stimulated generation of H₂O₂ and plays an integral role in insulin signal transduction. *Molecular and Cellular Biology*, 24(5), 1844–1854.
- Mahlich, Y., Reeb, J., Hecht, M., Schelling, M., De Beer, T. A. P., Bromberg, Y., & Rost, B. (2017). Common sequence variants affect molecular function more than rare variants? *Scientific Reports*, 7(1), 1608. <https://doi.org/10.1038/s41598-017-01054-2>
- Marco, E. Di, Jha, J. C., Sharma, A., Wilkinson-Berka, J. L., Jandeleit-Dahm, K. A., & de Haan, J. B. (2015). Are reactive oxygen species still the basis for diabetic complications? *Clinical Science*, 129(2), 199–216.
- Mcmartin, C., & Bohacek, R. S. (1997). QXP: powerful, rapid computer algorithms for structure-based drug design. *Journal of Computer-Aided Molecular Design*, 11(4), 333–344.
- Meng, W., Shah, K. P., Pollack, S., Toppila, I., Hebert, H. L., McCarthy, M. I., Groop, L., Ahlqvist, E., Lyssenko, V., Agardh, E., Daniell, M., Kaidonis, G., Craig, J. E., Mitchell, P., Liew, G., Kifley, A., Wang, J. J., Christiansen, M. W., Jensen, R. A., ... Palmer, C. N. A. (2018). A genome-wide association study suggests new evidence for an association of the NADPH Oxidase 4 (NOX4) gene with severe diabetic retinopathy in type 2 diabetes. *Acta Ophthalmologica*, 96(7), e811–e819. <https://doi.org/10.1111/aos.13769>
- Misra, D. P., & Ravindran, V. (2022). An overview of the functionalities of PubMed. *Journal*

Références bibliographiques

of the Royal College of Physicians of Edinburgh, 52(1), 8–9.

- Mohamed, I. N., Soliman, S. A., Alhusban, A., Matragoon, S., Pillai, B. A., Elmarkaby, A. A., & El-Remessy, A. B. (2012). Diabetes exacerbates retinal oxidative stress, inflammation, and microvascular degeneration in spontaneously hypertensive rats. *Molecular Vision*, 18, 1457.
- Morel, F., & Vignais, P. V. (1984). Examination of the oxidase function of the b-type cytochrome in human polymorphonuclear leucocytes. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 764(2), 213–225.
- National Center for Biotechnology Information (NCBI). (2022). Bethesda (MD): National Library of Medicine (US). <https://doi.org/50507>
- Nguyen, M. V. C. (2008). NADPH oxydase Nox4 native et recombinante: composés quinoniques, éléments de régulation. Grenoble 1.
- Peng, J. J., Xiong, S. Q., Ding, L. X., Peng, J., & Xia, X. B. (2019). Diabetic retinopathy: Focus on NADPH oxidase and its potential as therapeutic target. *European Journal of Pharmacology*, 853(March), 381–387. <https://doi.org/10.1016/j.ejphar.2019.04.038>
- Pollex, R. L., & Hegele, R. A. (2007). Copy number variation in the human genome and its implications for cardiovascular disease. *Circulation*, 115(24), 3130–3138.
- Popova, T., Boeva, V., Manie, E., Rozenholc, Y., Barillot, E., & Stern, M.-H. (2013). *Analysis of somatic alterations in cancer genome: from SNP arrays to next generation sequencing*.
- Qi, J., Chen, Y., Copenhaver, G. P., & Ma, H. (2014). Detection of genomic variations and DNA polymorphisms and impact on analysis of meiotic recombination and genetic mapping. *Proceedings of the National Academy of Sciences*, 111(27), 10007–10012.
- Rajasekaran, R., Doss, C. G. P., Sudandiradoss, C., Ramanathan, K., Rituraj, P., & Rao, S. (2008). Computational and structural investigation of deleterious functional SNPs in breast cancer BRCA2 gene. *Chinese Journal of Biotechnology*, 24(5), 851–856.
- Rao Kondapally Seshasai, S., Kaptoge, S., Thompson, A., Di Angelantonio, E., Gao, P., Sarwar, N., Whincup, P. H., Mukamal, K. J., Gillum, R. F., & Holme, I. (2011). Diabetes mellitus, fasting glucose, and risk of cause-specific death. *The New England Journal of Medicine*, 364(9), 829–841.
- Reaven, G. M. (1983). Insulin resistance in noninsulin-dependent diabetes mellitus: does it exist and can it be measured? *The American Journal of Medicine*, 74(1), 3–17.
- Rosen, E. D., Kaestner, K. H., Natarajan, R., Patti, M.-E., Sallari, R., Sander, M., & Susztak, K. (2018). Epigenetics and epigenomics: implications for diabetes and obesity. *Diabetes*, 67(10), 1923–1931.
- Safran, M., Chalifa-Casp, V., Shmueli, O., Rosen, N., Benjamin-Rodrig, H., Ophir, R., Yanai, I., Shmoish, M., & Lancet, D. (2003). The GeneCards/spl trade/family of databases: GeneCards, GeneLoc, GeneNote and GeneAnnot. *Computational Systems Bioinformatics. CSB2003. Proceedings of the 2003 IEEE Bioinformatics Conference. CSB2003*, 427–428.
- Schnecke, V., & Kuhn, L. A. (2000). Virtual screening with solvation and ligand-induced

Références bibliographiques

- complementarity. In *Virtual Screening: An Alternative or Complement to High Throughput Screening?* (pp. 171–190). Springer.
- Schröder, K. (2019). Redox control of angiogenesis. *Antioxidants & Redox Signaling*, *30*(7), 960–971.
- Schultz, J., Copley, R. R., Doerks, T., Ponting, C. P., & Bork, P. (2000). SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Research*, *28*(1), 231–234.
- Sedek, M., Callera, G., Montezano, A., Gutsol, A., Heitz, F., Szyndralewicz, C., Page, P., Kennedy, C. R. J., Burns, K. D., & Touyz, R. M. (2010). Critical role of Nox4-based NADPH oxidase in glucose-induced oxidative stress in the kidney: implications in type 2 diabetic nephropathy. *American Journal of Physiology-Renal Physiology*, *299*(6), F1348–F1358.
- Shen, S.-W., Reaven, G. M., & FARQuHAR, J. W. (1970). Comparison of impedance to insulin-mediated glucose uptake in normal subjects and in subjects with latent diabetes. *The Journal of Clinical Investigation*, *49*(12), 2151–2160.
- Shi, Q., Lee, D. Y., Féliers, D., Abboud, H. E., Bhat, M. A., & Gorin, Y. (2020). Interplay between RNA-binding protein HuR and Nox4 as a novel therapeutic target in diabetic kidney disease. *Molecular Metabolism*, *36*(February), 100968. <https://doi.org/10.1016/j.molmet.2020.02.011>
- Shiose, A., Kuroda, J., Tsuruya, K., Hirai, M., Hirakata, H., Naito, S., Hattori, M., Sakaki, Y., & Sumimoto, H. (2001). A novel superoxide-producing NAD (P) H oxidase in kidney. *Journal of Biological Chemistry*, *276*(2), 1417–1423.
- Sieber, P., Platzer, M., & Schuster, S. (2018). The definition of open reading frame revisited. *Trends in Genetics*, *34*(3), 167–170.
- Simon, J. M., & Gomez, N. C. (2021). Epigenetic analysis in Ewing sarcoma. In *Ewing Sarcoma* (pp. 285–302). Springer.
- Stasia, M. J. (2016). CYBA encoding p22(phox), the cytochrome b558 alpha polypeptide: gene structure, expression, role and physiopathology. *Gene*, *586*(1), 27–35. <https://doi.org/10.1016/J.GENE.2016.03.050>
- Sun, H., Saedi, P., Karuranga, S., Pinkepank, M., Ogurtsova, K., Duncan, B. B., Stein, C., Basit, A., Chan, J. C. N., & Mbanya, J. C. (2022). IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Research and Clinical Practice*, *183*, 109119.
- Touyz, R. M., Anagnostopoulou, A., Rios, F., Montezano, A. C., & Camargo, L. L. (2019). NOX5: Molecular biology and pathophysiology. *Experimental Physiology*, *104*(5), 605–616. <https://doi.org/10.1113/EP086204>
- Venselaar, H., Te Beek, T. A. H., Kuipers, R. K. P., Hekkelman, M. L., & Vriend, G. (2010). Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics*, *11*(1), 1–10.
- Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), 1189–1191.
- Wei, M., Dong, L., Wang, F., Cui, K., Ma, D., Yang, N., & Li, Y. (2021). The Prevalence and

Références bibliographiques

- Control of Type 2 Diabetes Mellitus in Residents of a Rural Town, Shandong Province, China. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy*, 14, 4505.
- Wilkinson-Berka, J. L., Deliyanti, D., Rana, I., Miller, A. G., Agrotis, A., Armani, R., Szyndralewicz, C., Wingler, K., Touyz, R. M., Cooper, M. E., Jandeleit-Dahm, K. A., & Schmidt, H. H. H. W. (2014). NADPH oxidase, NOX1, mediates vascular injury in ischemic retinopathy. *Antioxidants & Redox Signaling*, 20(17), 2726–2740. <https://doi.org/10.1089/ars.2013.5357>
- Willcox, A., Richardson, S. J., Bone, A. J., Foulis, A. K., & Morgan, N. G. (2009). Analysis of islet inflammation in human type 1 diabetes. *Clinical & Experimental Immunology*, 155(2), 173–181.
- Xu, Z., & Taylor, J. A. (2009). SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. *Nucleic Acids Research*, 37(suppl_2), W600–W605.
- Yang, Q., Wu, F., Wang, J., Gao, L., Jiang, L., Li, H.-D., Ma, Q., Liu, X., Wei, B., & Zhou, L. (2018). Nox4 in renal diseases: An update. *Free Radical Biology and Medicine*, 124, 466–472.
- Yau, J. W. Y., Rogers, S. L., Kawasaki, R., Lamoureux, E. L., Kowalski, J. W., Bek, T., Chen, S.-J., Dekker, J. M., Fletcher, A., & Grauslund, J. (2012). Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care*, 35(3), 556–564.
- Zaccardi, F., Webb, D. R., Yates, T., & Davies, M. J. (2016). Pathophysiology of type 1 and type 2 diabetes mellitus: A 90-year perspective. *Postgraduate Medical Journal*, 92(1084), 63–69. <https://doi.org/10.1136/postgradmedj-2015-133281>

Les sites utiliser

<https://www.ncbi.nlm.nih.gov/>

<https://genome.ucsc.edu/>

<https://www.ensembl.org/index.html>

<https://pubmed.ncbi.nlm.nih.gov/>

<https://www.genecards.org/>

<https://www.uniprot.org/>

<http://smart.embl-heidelberg.de/>

<https://swissmodel.expasy.org/>

<https://blast.ncbi.nlm.nih.gov/Blast.cgi>

<https://www.ncbi.nlm.nih.gov/orffinder/>

<https://snpinfo.niehs.nih.gov/>

<https://sift.bii.a-star.edu.sg/>

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

<http://mutpred.mutdb.org/>

<https://www.ncbi.nlm.nih.gov/CBBresearch/Lu/Demo/LitVar/#!?query>

<https://snps.biofold.org/phd-snp/phd-snp.html>

<https://www.rostlab.org/services/SNAP/>

<https://folding.biofold.org/i-mutant/i-mutant2.0.html>

<https://www3.cmbi.umcn.nl/hope/input/#>

<https://pubchem.ncbi.nlm.nih.gov/>

http://sts.bioe.uic.edu/castp/index.html?j_62877290304dc

<https://www.rcsb.org/>

Annexes

Annexe 1

```

>NM_016931.5 Homo sapiens NADPH oxidase 4 (NOX4), transcript variant
1, mRNA
ACTCAGTCTTTGACCCTCGGTCTCGCTCAGCGGCCCGCGCAGGCCGCACAACCTGTAACCGCTGCCCGGCG
CGCCGCGCGCTCCTTCTCGGTCCGGCCGGGCACAGAGCGCAGCGCGGGCGGGCGGCGGCGATGGCTGTGTGTC
CTGGAGGAGCTGGCTCGCCAACGAAGGGTTAAACACCTCTGCCTGTTTCATCTGGCTCTCCATGAATGTC
CTGCTTTTCTGGAAAACCTTCTTGCTGTATAACCAAGGGCCAGAGTATCACTACCTCCACCAGATGTTGG
GGCTAGGATTTGTGCTAAGCAGAGCCTCAGCATCTGTCTTAACTCAACTGCAGCCTTATCCTTTTACC
CATGTGCCGAACACTCTTGGCTTACCTCCGAGGATCACAGAAGGTTCCAAGCAGGAGAACCAGGAGATTG
TTGGATAAAAAGCAGAACATTCATATTACCTGTGGTGTACTATCTGTATTTTCTCAGGCGTGCATGTGG
CTGCCCATCTGGTGAATGCCCTCAACTTCTCAGTGAATTACAGTGAAGACTTTGTTGAACTGAATGCAGC
AAGATACCGGAGATGAGGATCCTAGAAAACCTTCTTTCACAACCTGTTCCCTGGCCTGACAGGGGTCTGCATG
GTGGTGGTGTCTATCCATCATGATCACAGCCTCTACATATGCAATAAGAGTTTCTAACTATGATATCTTCT
GTATATACATAAACCTTCTTGTCTTCTACATGCTGCTGACGTGTGCATGTTTCAGGAGGGGTGTGAA
GTATCAAATAATTTAGATACCCACCTCCCGGCTGCATCAGTCTTAAACCGAACAGCTCTCAGAATATT
TCCTTACCAGAGTATTTCTCAGAACATTTTCATGAACCTTTCCCTGAAGGATTTTCAAACCCGCGAGAGT
TTACCCAGCACAAATTTGTGAAGATTTGTATGGAAGAGCCAGATTTCCAAGCTAATTTCCACAGACTTG
GCTTTGGATTTTGGACTTTTGTCCCTTGTCCCTGTGCGGAAAGACTTTTACAGGTATATCCGGAGCAATAAG
CCAGTCACCATCATTTCCGGTCTAGTGCATCCCTCAGATGTGCATGGAAATCCGAATGGTCAAAGAAAATTT
TTAAAGCAAGACTGGTCAGTATATTACTCTACATTTGCCAGTGTATCTGCATAGAAAATCATCCATT
TACCCTCACAATGTGTCAACTGAAACCAAAGCAACATTTGGGGTTTCACTTAAAAATAGTAGGAGACTGG
ACAGAACGATTTTGGACTTTTACTACTGCTCCATCTAGTCAAGACTCCGAAAATTTCTGCCCTTCACTCAAT
CTAGAAAATATCCCAAGCTGTATATTGATGGTCTTTTGGAAAGTCCATTTGAGGAATCACTGAATATGA
GGTCAGCCTCTGCGTGGCTGGAGGCATTTGGAGTAACTCCATTTGCATCAATACTCAACACCCCTGTTGGAT
GACTGGAAACCATACAAGCTTAGAAGACTATACTTTATTTGGGTATGCAGAGATATCCAGTCTTCCGTT
GGTTTGCAGATTTACTCTGTATGTTGCATAACAAGTTTGGCAAGAGAACAGACCTGACTATGTCAACAT
CCAGCTGTACCTCAGTCAAACAGATGGGATACAGAAGATAATTGGAGAAAAATATCATGCACTGAATTCA
AGACTGTTTATAGGACCTCCTCGGTGGAACCTTTGTTGATGAAATAGCAAAAATATAACAGAGGAAAAA
CAGTTGGTGTCTTCTGTTGTGGACCCAAATCACTATCCAAGACTCTTCATAAACTGAGTAACCAGAACAA
CTCATAATGGGACAAGATTTGAATACAATAAAGAGTCTTTTCAGCTGAAAACCTTTTGGCCATGAAGCAGGACT
CTAAAGAAGGAATGAGTGCAATTTCTAAGACTTTGAAACTCAGCGGAATCAATCAGCTGTGTTATGCCAA
AGAATAGTAAGGTTTTCTTATTTATGATTTTAAATGGAATGTGAGAATGTGGCAAGATGACCGTCA
CATTACATGTTAATCTGGAACCAAAGAGACCCTGAAGAATATTGATGTGATGATTTCACTTTTCAGTT
CTCATAATAAAAGAAAACCTTAGATGCACACTGTTGATTTTTCATGGTGGATTCAAGAATCCCTAGTGA
GGAGCTGAACTTGCTCAATCTAAGGCTGATTGTCGTGTTCCCTCTTAAATGTTTTTGGTTGAACAAATG
CAAGATTGAACAAAATTAATAATTCATTGAAGCTGAAATCCATTTTCTGTGTGTGTATAAACAGAGTA
GCTTTAATTTGAAGCACTCCAGGCAATATATAGATGTTTGAACACACAGCACAAGACTCTGTATTGA
GCTGGTACTTTTGTGTCAAATCTAATCGTCTCCACTACTTATGCTAATAACCTCTATTGTATCTGAA
ACTATATGCTAACTGAACCTTCCCTCAATGTTGTTATAGTATCTATTTTATATATTTTTTCTTTTTTAT
TCCTCTCTTAGGGAAATATAGCTTCCCTTAGCATGCATAGACATAATGATTTAATAGGTCCTTTTCAT
CTTCATTTAAATCTATCACTATTGCATGGTAATGAAAATATTCCTACTATAAATATAAAGGGATATATA
TATATGGATATATATATATATACACATATATATATACACACACACATATATATATATATATATATA
TATATACACATATACTAATAACTTTTCCCTTTTTTTCAGCATTTTGTCTCTATTATTATTTGTTTTTT
TCCCAGGTAGGTTTGTCTTAGGCTGTAGCCTTAAGGATAGTTAGTTAATTTGCACCTTTGAGACCAAAG
GACATCATGTGTGTCAGTAGGGACTGAAATATAAGATTTATCTCCTTTGCCACACATTTGGTTTTATGATGGA
GACATTTGAAGTCTAGTGCATATTCCTGAACAGTAAACCTGTATTTTACCTTTTAAAGTAAGAGGAAATAT
GATATTTCTTATCAAACCTTAAAGTTTGAATCCAGAAATATTAAGTGTGCACTTTTGGTATCCTGAGTTTC
CATAGGAACTATTGGGTTTAAAGTCAACGTTGGAACCTACACTGTGTGATCTTATAAACTTATGTTCCCT
GGCTATCATAATCTTGGCTCAGAACAATATTTCCCATTAATCTTAAAGAAATTAAGGCATTCATGGCTCA
CGCTGTAAATCCGGCACTTTGGGAGGCCAAGGCAGGTGGATCACAGGTCAGAGTTCAAGACCAGCCTG
ACCAAGTTGGTGAAACCCATCGCTACTAAAAATACAATAATAGCGCGGTGTTGGTGGCGGGTGCCTATA
ATCCAACTACTCGGGAGACTGAGGCAAGAATCGCTTGAACCCGGTGGCGGAGGTTGCAGTGAAGCAA
GATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGATTCATCTCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAATAAGCATTACACAGTTTAGTATTTGTTGTTAGTAGTCCGCTATCAATGCTATCAAAAT
ATAACACTGCTGAAATCAGCAGTGTGACTTACCTTGCCATTTGTTAAATGTTTACATAAAAAACATAACATGA
TAGATGCTAAGGCTTTTTTGTCTATAATTCACCAATAGCAATCAAGCATGCTAACCCATACTGAATGAT
ATTTACTTGTAGATATTTCTTCTTCTTCTTGAATTTCTCTTTCTATGAAAGAAGATGAACCCAAAAA
GTGATAGGAAATGTGGAATGCTCATGCAGATTTAGCTCTGAAGGCATATTTAAACTAGTATGTCTTGA
CAACAGCTTTTAGATTAATAAAGAAATTTTCACTGGAAACATTTAACAGAAAGAACTAGTAAAAAGACACTTT
GAGTTAGTCCAGGCTTAAATGTGCAATACTGACTCTATACTGATCATAATTTATTTATGCGATCATATTA
TAGACCTAATTTCAATTAACAGGTTAGAAGATTTTCAAAGAAAGATGATGTTTCAAAGCTGGTCTGCC
ATCTTAGATGACCTCCTTGTCTTATTTAAGTCCAGTAGGTGTTCAAATGTTAAATGTTAAACATAGGTC
ATCTTTGCTTCTGCAAGGCTTCATCTTGCATGTTTAAAGAGAACCTTTGTTTTATTTTGGGGATATTTCT
CTGGGATCATTTTATAATAACAGCCTTAAATCACTAATTTTAGTAGCAATAAATGTTAAATGAA

```

Figure 53: séquence nucléotidique (FASTA) du gène NOX4 "NM_016931.5".

Annexes

Annexe 2

Study ID	Organism	Study type	Number of variant regions	Publication
nsid102	human	Case-Set	76027	
nsid75	human	Case-Set	927	Waener et al. 2012
nsid11	human	Tumor vs. Matched-Normal	181	Walker et al. 2009
nsid180	human	Control Set	92934	
nsid214	human	Collection	89819	
nsid213	human	Somatic	5864	
nsid209	human	Collection	173332	
nsid212	human	Control Set	132198	Wu et al. 2021
nsid199	human	Collection	38027	Quan et al. 2021
nsid207	human	Collection	107990	Ebert et al. 2021

Study ID	Organism	Study type	Number of variant regions	Publication
nsid204	human	Collection	129165	Chan et al. 2021
nsid206	human	Control Set	152814	Almani et al. 2020
nsid200	human	Control Set	290092	Abal et al. 2020
nsid186	human	Control Set	304733	Collins et al. 2020
nsid183	human	Control Set	55675	
nsid173	human	Case-Control	23549	Zarrei et al. 2019
nsid187	human	Control Set	30218	Wang et al. 2019
nsid162	human	Collection	103985	Chaisson et al. 2019
nsid188	human	Collection	18424	Levy-Sakin et al. 2019
nsid182	human	Collection	99610	Audano et al. 2019

Study ID	Organism	Study type	Number of variant regions	Publication
nsid144	human , chimpanzee	Control Set	37798	Gardner et al. 2017
nsid132	human	Case-Set	6173	Walker et al. 2017
nsid140	human	Control Set	22644	Fan et al. 2017
nsid151	human	Case-Set	49339	Ruderfer et al. 2016
nsid130	human	Collection	39657	Lessa et al. 2016
nsid219	human	Control Set	68925	Sudmant et al. 2015
nsid214	human	Control Set	61678	1000 Genomes Project Consortium et al. 2015
nsid113	human	Case-Set	9010	Fedyak et al. 2015
nsid99	human	Control Set	11116	Tharaja et al. 2015
nsid217	human	Control Set	40141	Bessenbacher et al. 2015

Study ID	Organism	Study type	Number of variant regions	Publication
nsid212	human	Control Set	16676	Uddin et al. 2014
nsid100	human	Case-Control	70319	Gao et al. 2014
nsid224	human	Control Set	3570	Sukhtovat et al. 2014
nsid108	human	Control Set	39678	Almouzni et al. 2013
nsid92	human	Control Set	245	Forabono et al. 2013
nsid209	human	Control Set	471917	Fang et al. 2014
nsid80	human	Case-Control	1410	Widomson Sillfouat et al. 2013
nsid82	Sumatran orangutan , human , eastern gorilla , Bornean orangutan , chimpanzee , western gorilla , pygmy chimpanzee	Control Set	29214	Sudmant et al. 2013
nsid89	human	Case-Set	1512	Bosch et al. 2013
nsid71	human	Control Set	45084	Xu et al. 2011

Study ID	Organism	Study type	Number of variant regions	Publication
nsid54	human	Case-Control	81345	Cooper et al. 2011
nsid188	human	Control Set	60247	Pinto et al. 2011
nsid59	human	Control Set	228871	1000 Genomes Project Consortium et al. 2010
nsid180	human	Control Set	23887	Pang et al. 2010
nsid197	human	Control Set	232775	McKernan et al. 2009
nsid27	human	Control Set	13843	Itara et al. 2009
nsid194	human	Control Set	504912	Bentley et al. 2008
nsid192	human	Collection	61167	Forbes et al. 2008, Forbes et al. 2010
nsid30	human	Control Set	1428	Jakobsson et al. 2008
nsid22	human	Control Set	780358	Levy et al. 2007

Study ID	Organism	Study type	Number of variant regions	Publication
nsid24	human	Control Set	1773	de Smith et al. 2007
nsid9	human	Control Set	415434	Mills et al. 2006
nsid215	human	Control Set	21981	
nsid211	human	Collection	56732	Chuang et al. 2021
nsid203	human	Collection	140215	Borras-Monroy et al. 2021
nsid109	human	Collection	1014	Wildschutte et al. 2015
nsid94	human	Tumor vs. Matched-Normal	8568	Helman et al. 2014
nsid128	human	Collection	1328521	Mallick et al. 2016
nsid41	human	Control Set	255	Jafraie et al. 2004

Figure 54:les déférentes variants de gène NOX4 (NCBI dbvar).

Transcript ID	Name	bp	Protein	Biotype	CCDS	UniProt Match	RefSeq Match	Flags				
ENST00000263317.9	NOX4-201	4269	578aa	Protein coding	CCDS8285	Q9NPH5-1	NM_016931.5	MANE Select v0.95	Ensembl Canonical	GENCODE basic	APPRIS P1	TSL:1
ENST00000424319.5	NOX4-205	4184	554aa	Protein coding	CCDS44696	Q9NPH5-3	-			GENCODE basic		TSL:5
ENST00000343727.9	NOX4-202	3347	554aa	Protein coding	CCDS44696	Q9NPH5-3	-			GENCODE basic		TSL:2
ENST00000527956.5	NOX4-210	2379	554aa	Protein coding	CCDS44696	Q9NPH5-3	-			GENCODE basic		TSL:5
ENST00000534731.5	NOX4-215	2156	538aa	Protein coding	CCDS44695	Q9NPH5-6	-			GENCODE basic		TSL:1
ENST00000532825.5	NOX4-214	2039	514aa	Protein coding	CCDS73361	Q9NPH5-9	-			GENCODE basic		TSL:2
ENST00000528341.5	NOX4-211	1899	553aa	Protein coding	CCDS73382	E8PPP2	-			GENCODE basic		TSL:2
ENST00000527626.5	NOX4-209	1774	391aa	Protein coding		E8PR43	-			GENCODE basic		TSL:2
ENST00000525196.5	NOX4-207	1655	342aa	Protein coding		E8PI95	-			GENCODE basic		TSL:1
ENST00000375979.7	NOX4-203	816	271aa	Protein coding		Q9NPH5-4	-			GENCODE basic		TSL:1
ENST00000531342.5	NOX4-213	696	231aa	Protein coding		Q9NPH5-3	-			GENCODE basic		TSL:1
ENST00000393282.2	NOX4-204	580	54aa	Protein coding		Q9NPH5-7	-			GENCODE basic		TSL:1
ENST00000529343.5	NOX4-212	1403	224aa	Nonsense mediated decay		Q9NPH5-5	-					TSL:1
ENST00000525278.1	NOX4-208	1309	No protein	Retained intron		-	-					TSL:3
ENST00000524473.1	NOX4-206	755	No protein	Retained intron		-	-					TSL:2

Figure 55:les transcrits de gène NOX4 (Ensembl).

Annexes

Annexe 4

Tableau XVII: les informations sur la protéine NADPH Oxydase 4 (UniProt).

UniProt	
Protéine	NADPH Oxydase4, Oxydase rénale-1
Gène	NOX4
Organisme	Homo Sapiens
Fonction moléculaire	Oxydoréductase
Symbole	Q9NPH5
Ligand	NADP
Domains	Oxydoréductase ferrique, FAD
Isoforme	9 isoformes
Type	Transmembranaire
Famille	NOX1-5, Duox1-2
Interaction	CYBA/P22Phox.

Sequence
ORFs found: 31 Genetic code: 1 Start codon: ATG only

ORF12 (100 aa) Display ORF as... Mark

Label	Strand	Frame	Start	Stop	Length (nt aa)
ORF12	+	3	1521	1850	330 109
ORF6	+	2	2465	2755	291 96
ORF21	-	1	204	28	177 58
ORF1	+	1	2671	2844	174 57
ORF31	-	3	1084	914	171 56
ORF2	+	1	3843	4104	162 53
ORF11	+	3	1062	1253	192 63
ORF5	+	2	1568	1729	162 53
ORF9	+	3	354	508	153 50
ORF15	+	3	3402	3554	153 50
ORF14	+	3	3222	3371	150 49

BLAST Database: UniProtKB/Swiss-Prot (swissprot)

FOLLOW NCBI

Figure 59: recherche des ORF (ORF Finder)

```
>|c|ORF1
MAVSWRSLWANEGVKKHLCLFIWLSMNVLLFWKTFLLYNQGPPEYHYLHQMLGLGLCLSRASASVNLNCSLILLPMCRLLAYLRGSQKVPSR
RTRRLLDK

SRTFHITCGVTICIFSGVHVAAHLVNALNFSVNYSEDFVELNAARYRDEDPRKLLFTTVPGLTGVCMVVVLFMLITASTYAIRVSNYDIFWYTHN
LFFVF

YMLLTLHVSGLLKYQTNLDTHPPGCISLNRSSQNSLPEYFSEHFHEFPPEGFSKPAEFTQHKFVKICMEEPRFQANFPQTWLVISGPLCLYCAE
RLY

RYIRSNKPVTIISVMSPSDVMEIRMVKENFKARPGQYITLHCPSVSALENHPFTLTMCPPTETKATFGVHLKIVGDWTERFRDLLLPSSQDSEILP
FIQ

SRNYPKLYIDGPFPGSPFEESLNYEVSCLVAGGIGVTPFASILNLLDDWKPYKLRRLYFIWVCRDIQSFRWFADLLCMLHNKFWQENRPDYVNIQ
LYLSQ

TDGIQKIIEKGYHALNSRFLFIGRPRWKLDFDEIAKYNRGKTVGVFCGPNLSKTLHKLSNQNSYGRFEYNKESFS
```

Figure 60: L'ORF le plus long (ORF-Finder NCBI).

Annexe 5

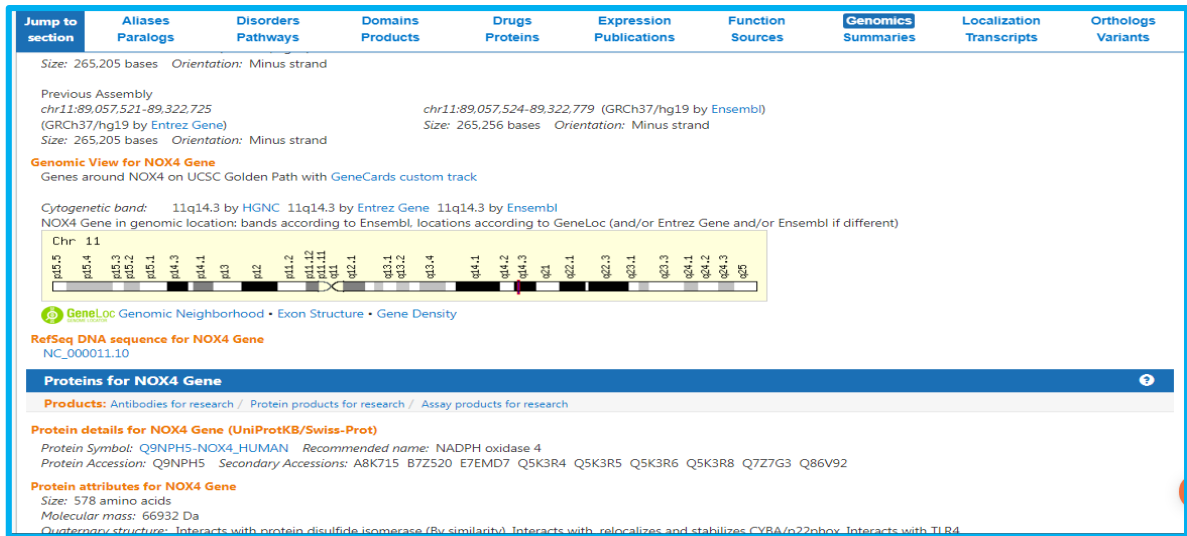


Figure 61: cartographie génétique du gène NOX4.

Annexe 6

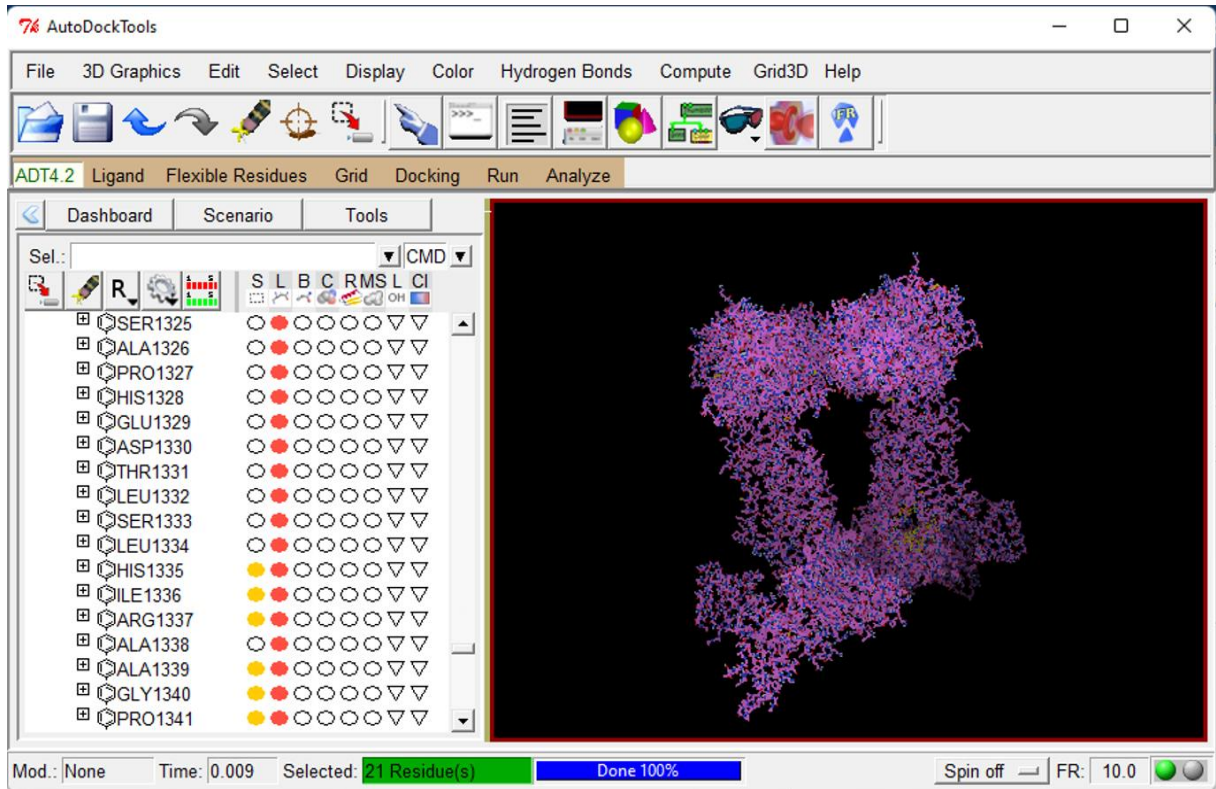


Figure 62: Préparation de la protéine de NOX4 (7D3E) (Autodock tools).

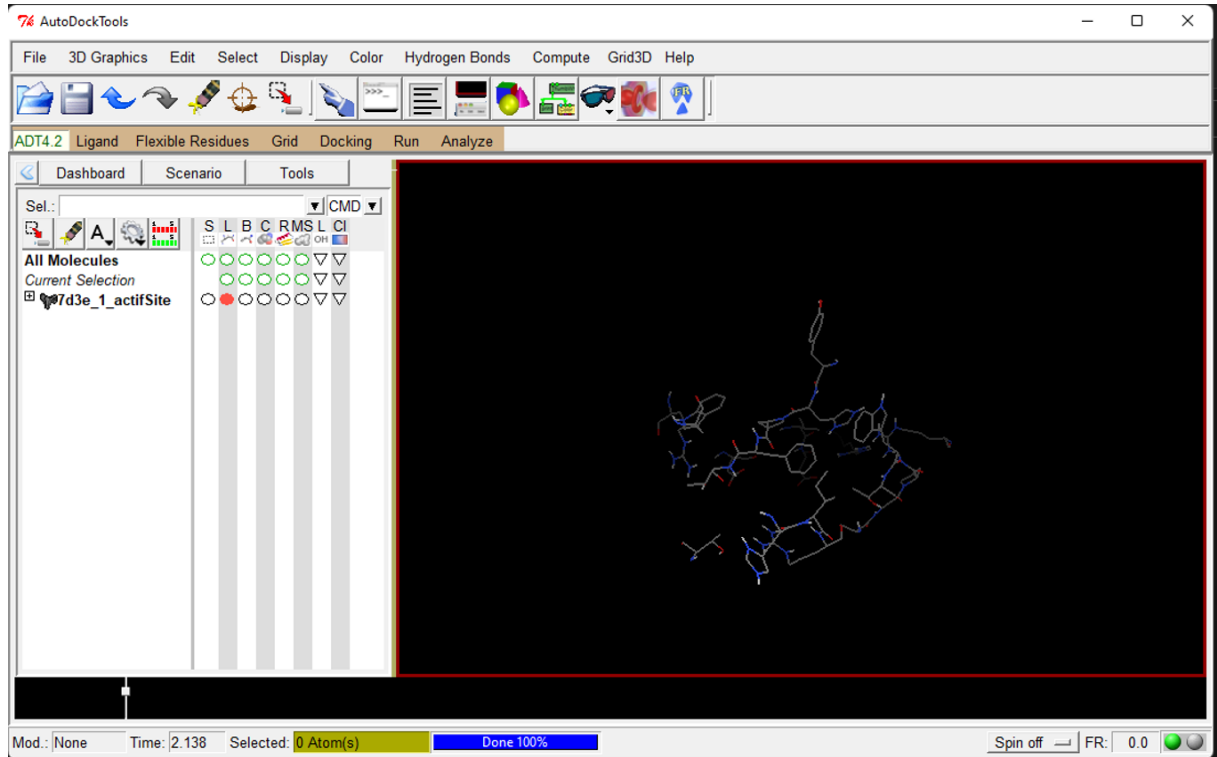


Figure 63: Préparation du site actif de la protéine NOX4 (7D3E) pour le Docking moléculaire (Autodock tools)

Annexe 7

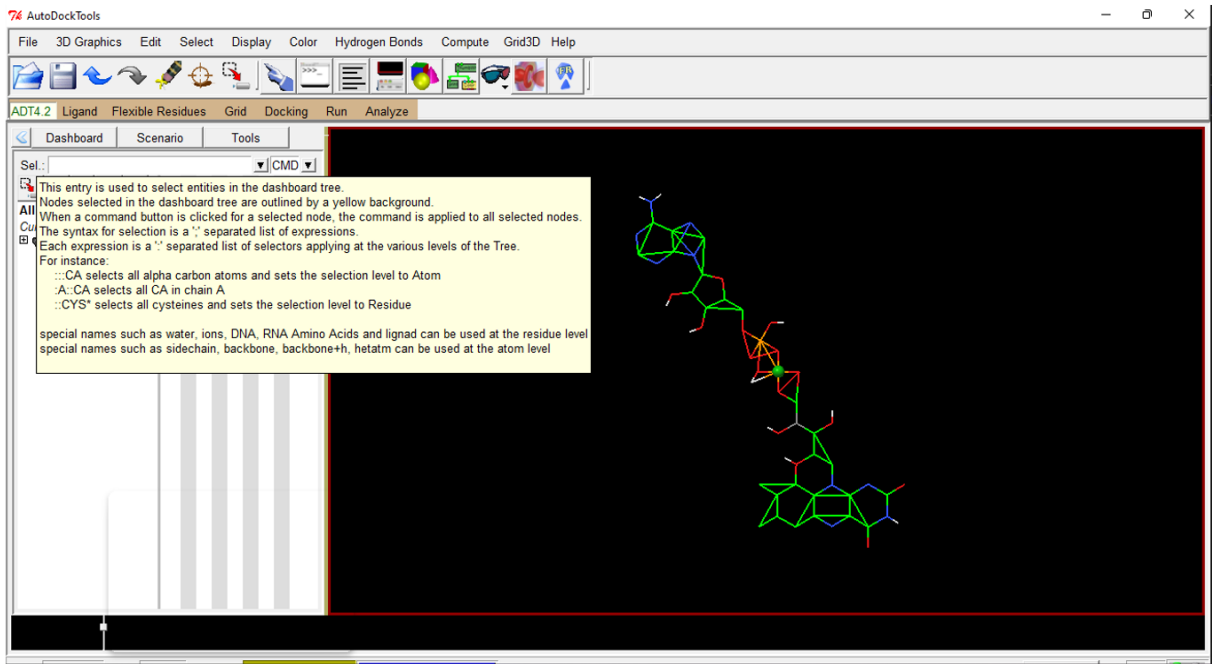


Figure 64: Préparation du ligand Flavin adenine dinucleotide (FAD) (Autodock tools) .

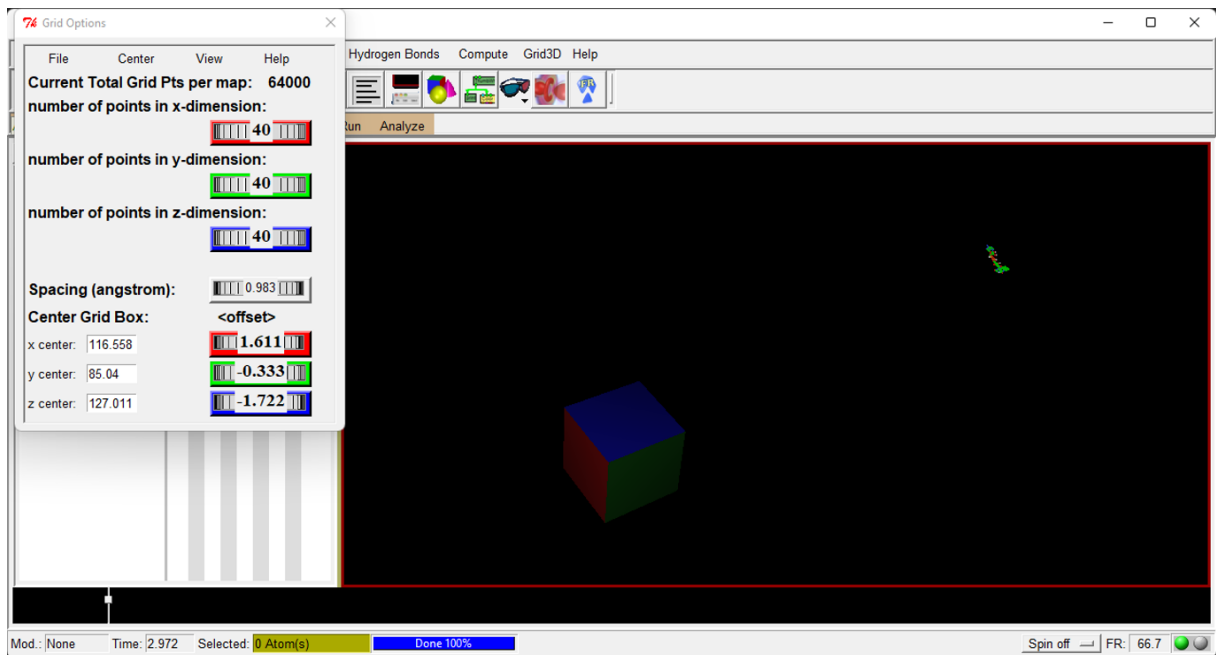


Figure 65: La configuration des coordonnées du Grid box pour le Docking moléculaire (Autodock tools).