

République Algérienne Démocratique et Populaire
Ministère de l'enseignement Supérieur et de la Recherche Scientifique
Université SAAD DAHLEB -BLIDA1



Faculté des Sciences de la nature et de la Vie
Département de Biologie et Physiologie Cellulaire
Mémoire de fin d'études en vue de l'obtention du
Diplôme de Master II
Options : Bioinformatique

Thème

**ETUDE PAR METHODE BIO-INFORMATIQUE DES
INTERACTIONS ENTRE LES GENES ANTIOXYDANTS CAT,
SOD1 , SOD2 , GPx1 ET LE GENE DE L'INSULINE DANS LE
DIABETE TYPE1**

Soutenu le : 08 / 10 /2017

Présenté par :

Mlle BENMOUSSA SARAH

Mlle MECHOUCHE SALIMA

Devant le jury composé de :

Mme Zerouti k	MAA	USUB1	Présidente
Mme GUESSAIBIA	MCB	USUB1	Examinatrice
Mme EDDAIKRA ATIKA	MAA	USUB1	Promotrice

Année universitaire 2016-2017

Remerciements

« Ce travail n'aurait jamais pu voir le jour sans le courage et la volonté que nous a donné le bon **DIEU** »

Nous tenons à adresser nos sincères remerciements et profond respect à notre promotrice **Mme EDDAIKRA ATIKA**, chargé de cours et Maitre Assistante à la faculté des sciences de la nature de l'université de Blida. Nous tenons à lui témoigné notre profonde gratitude de nous avoir proposé la thématique de notre sujet de fin d'étude, nous la remercions pour sa patience, sa compréhension, sa disponibilité, son savoir faire et pour toute l'aide qu'elle nous a rapporté.

Un grand merci également aux membres de ce jury qui ont accepté d'évaluer notre travail :

Mme GUESSAIBIA Nadia, Maitre de conférence à la faculté des sciences de la nature de l'université de Blida, en qualité d'examinatrice. et un grand merci aussi pour la présidente **Madame ZERROUTI KHadidja**

Nous exprimons notre profonde gratitude à Monsieur **SENINET Mohamed El-Hadi** en qualité de co-promoteur et qui nous a accompagnés dans le départ de ce travail.

Nous exprimons notre profonde gratitude à Mme la Chef de département d'avoir été à notre écoute pendant toute la durée de notre formation.

Nous remercions nos familles, tous les amis pour leurs encouragements.

Nous remercions aussi tous ceux qui ont contribué de près de loin à la mise en œuvre de ce travail. »

Dédicaces

Avant toute chose je remercie Allah le tous puissant de m'avoir donné la santé, la patience et le courage pour réaliser ce travail.

J'ai l'honneur de dédie ce modeste travail à mes chers parents, qui m'avez dirigé et suivi pondent toute mes années d'étude et surtout MA Mère NOURA la plus belle mère dans la vie, pour toutes tes efforts fournée, toutes tes sacrifices pour toute la confiance que vous m'avez donné afin de me motivé dans mes études et me voir satisfaite et heureuse.

*A mon père Mahfoud pour sa patience, sa confiance et son respect de mes choix rien au monde ne vont les efforts fournis jour et nuit pour mon éducation et mon bien être

* Je vous offre toutes mes années d'étude ainsi que mon diplôme. Je ne pourrai jamais oublier d'exprimer ma profonde gratitude à : Ma grande -mère et mon grande père que dieu la garde pour nous.

* Mon frères : Abed el hak

* Mes très chères sœurs : Ikram et farah

* Mes camarades : Mohamed, Dalal , amina , Asma, Salima , Hadjar Et bien sur tous mes enseignants depuis la période de primaire jusqu'à l'université et spécialement Mme **EDDAIKRA ATIKA** et **MOHAMED SAID Ramdan** ET Mme **GUESSAIBIYA Nadia**.

sarah

Dédicaces

Je dédie ce travail à mes chers parents .merci papa merci maman, merci de m'avoir soutenue tout au long de mes études, de m'encourager dans les moments difficiles c'est grâce a vous si je suis arrivé jusqu'à ici les mots ne suffisent pas à vous exprimer ma gratitude.

Merci à mes chère sœurs et amies Zohra et aicha , sans vous la vie ne serait pas aussi belle .

Merci à mes chers frères, Mohamed et sa femme Aicha et ses fils Amira , Sohila et Abde El Rahman , DJamal.et spécialement Hassen et sa femme Nora et son fils Abd El Moamin ,Omar et Hamza , merci de m'avoir soutenu.

à mes chères amies Naima , Hamnane ,Ahlam ,Zakia , Saja , pour tous les bon moment .

à mes copine Meriame ,Asmaa , Hajar ,Djamila,

Et je dédie spécialement à et à ma collègue de ce travail Sarah .

SALIMA

Résumé

Le stress oxydatif joue un rôle important dans le développement de diabète type1 (DT1) et les enzymes antioxydants réduisent ce stress oxydatif. Les variations génétiques des gènes antioxydants pourraient donc être impliquées dans la pathogénie de DT1. Notre étude in silico repose sur l'utilisation des outils bioinformatique pour étudier les interactions entre les gènes antioxydants : superoxyde dismutase1 et 2 : (SOD1 , SOD2), la glutathion peroxydase (GPx1) et la catalase (CAT). Aussi de comprendre la relation de ces derniers avec le gène d'insuline et le DT1.

Les séquences protéiques ont été récupérés et blastés pour rechercher des similitudes de régions avec d'autres gènes, puis alignés entre eux par l'outil UNIPROT. L'obtention des arbres phylogénique sa été faite par l'outil **Cluster Omega**. Pour la visualisation des interactions entre nos séquences protéiques, nous avons utilisé les outils STRING, GeneMania et Bio GRID. La recherche d'association entre le diabète type1 et les variants polymorphes des enzymes antioxydantes par l'outil db SNP et UCSC.

Nos résultats confirment l'association des enzymes antioxydantes dans la physiopathologie du diabète type1 grâce aux interactions moléculaires des gènes antioxydant et la présence de variants. Ces résultats peuvent nous aider aussi, à comprendre les voies et la prédiction des fonctions des enzymes antioxydantes et leurs rôles dans le DT1.

Les mots clés : Le stress oxydatif , diabète type1, interaction , gènes antioxydants,outil bioinformatique.

Abstract

Oxidative stress plays an important role in the development of type 1 diabetes (DT1) and antioxidant enzymes reduce this oxidative stress. Genetic variations of antioxidant genes could therefore be implicated in the pathogenesis of T1D. Our in silico study is based on the use of bioinformatics tools to study the interactions between antioxidant genes: superoxide dismutase1 and 2: (SOD1, SOD2), glutathione peroxidase (GPx1) and catalase (CAT). Also to understand the relationship of these with the insulin gene and DT1.

The protein sequences were recovered and blasted to look for similarities of regions with other genes and then aligned with each other by the UNIPROT tool. Obtaining the phylogenetic trees was made by the Omega Cluster tool. To visualize the interactions between our protein sequences, we used STRING, GeneMania and Bio GRID tools. The search for association between diabetes type1 and polymorphic variants of antioxidant enzymes by the db SNP tool and UCSC.

Our results confirm the association of antioxidant enzymes in the pathophysiology of type 1 diabetes thanks to the molecular interactions of the antioxidant genes and the presence of variants. These results can also help us understand the pathways and the prediction of the functions of antioxidant enzymes and their roles in DT1

Key words: Oxidative stress, type 1 diabetes, interaction, antioxidant genes, bioinformatics tool.

ملخص

الإجهاد التأكسدي يلعب دورا هاما في تطوير داء السكري من النوع 1 (DT1) والإنزيمات المضادة للأكسدة تقليل هذا الأكسدة الإجهاد وبالتالي يمكن أن تتورط الاختلافات الجينية للجينات المضادة للأكسدة في التسبب في DT1 ويستند لدينا في دراسة سيليكو على استخدام أدوات المعلوماتية الحيوية لدراسة التفاعلات بين الجينات المضادة للأكسدة (SOD1 و SOD2) و بين (CAT و gpX1) أيضا لفهم العلاقة مع هذه الجينات الأنسولين و DT1.

تم استعادة تسلسل البروتين وانتقد للبحث عن أوجه التشابه بين المناطق مع الجينات الأخرى ومن ثم الانحياز مع بعضها البعض من قبل أداة UNIPROT تم الحصول على أشجار النشوء والتطور من قبل الأداة Cluster Omega

لتصور التفاعلات بين تسلسل البروتين لدينا، استخدمنا الأدوات .STRING .GeneMania و Bio GRID. البحث عن ارتباط بين نوع السكري 1 ومتغيرات متعددة الأشكال من الإنزيمات المضادة للأكسدة من قبل الأداة db SNP و UCSC .

نتائجنا تؤكد ارتباط الإنزيمات المضادة للأكسدة في الفيزيولوجيا المرضية من مرض السكري من النوع 1 وذلك بفضل التفاعلات الجينية للجينات المضادة للأكسدة وجود المتغيرات.

هذه النتائج يمكن أن تساعدنا أيضا على فهم مسارات والتنبيؤ وظائف الإنزيمات المضادة للأكسدة وأدوارها في DT1.

الكلمات الرئيسية: الإجهاد التأكسدي ، داء السكري من النمط الأول، التفاعل، الجينات المضادة للأكسدة، أداة المعلومات البيولوجية.

Liste des Tableaux

Liste des Tableaux

Numéro de page

Tableau I : Description des outils bioinformatiques.....	22
Tableau II : Description des molécules biologiques.....	24
Tableau III : les identifications des enzymes dans UNIPROT.....	28
Tableau IV : Résultats du blast X de la glutathion peroxydase.....	30
Tableau V : Résultats du blast X de l'insuline.....	31
Tableau VI : Résultats du blast X de Catalase.....	32
Tableau VII : Résultats de la totalité des alignements trouvés par UNIPROT entre les Différentes séquences.....	35
Tableau VIII : Cadres de lecture de gène	37
Tableau IX : Enrichissement fonctionnel.....	38
Tableau X : les fréquences allyliques génotypiques du polymorphisme C1167T et C (-262) T de la CAT chez les DT1.....	43
Tableau XI : les fréquences allyliques génotypiques du polymorphisme de la GPX1 dans Différentes pathologies.....	44
Tableau XII : les fréquences allyliques génotypiques majeur et mineur du polymorphisme de la SOD2 dans différentes pathologies chez cinq population différentes.....	45

Liste des Figures

Liste des Figures

Numéro de la page

Figure1 : Mode d'action des principaux systèmes enzymatiques antioxydants et de leurs cofacteurs métalliques.....	4
Figure2 : Localisation du gène de la catalase au niveau du chromosome11.....	6
Figure3 : Histogramme représentatif du niveau d'expression des RNA-Seq de la catalase dans différents tissus.....	6
Figure4 : Localisation du gène GPx1 au niveau du chromosome3.....	7
Figure 5 : Histogramme représentatif du niveau d'expression des RNA-Seq de la GPx1 dans différents tissus.....	8
Figure6 : Histogramme représentatif du niveau d'expression des RNA-Seq de la GPx1 dans différents tissus.....	9
Figure7 : Histogramme représentatif du niveau d'expression des RNA-Seq de la SOD1 dans différents tissus.....	10
Figure 8 : Localisation du gène SOD2 au niveau du chromosome 6.....	11
Figure9 : Histogramme représentatif du niveau d'expression des RNA-Seq de la SOD2 dans différents tissus.....	11
Figure10 : structure du gène de l'insuline.....	13
Figure11 : Histogramme représentatif du niveau d'expression des RNA-Seq du INS dans différents tissus.....	14
Figure12 : Organisation génomique du gène humain SOD1.....	15
Figure13 : La région promoteur du gène de l'insuline.....	16
Figure14 : SNPs et Haplotypes.....	19
Figure16 : l'organisation du protocole pour nos étude.....	23

Liste des Figures

Figure18 : les résultats d'alignement pour la séquence nucléotidique du glutathion peroxydase.....	29
Figure19 : les résultats d'alignement pour la séquence nucléotidique de l'insuline.....	30
Figure20 : les résultats d'alignement pour la séquence nucléotidique de Catalase.....	31
Figure21 : Alignement des quatre séquences peptidiques (SOD1, SOD2, CAT, GPX1).....	32
Figure22 : Alignement des séquences peptidiques de la SOD1 et SOD2 par UNIPROT	32
Figure23 : des séquences peptidique de GPX1et CAT par UNIPROT.....	33
Figure24 : Arbre phylogénique des quatre enzymes antioxydants par CLUSTER Omega....	34
Figure25 : Arbre phylogénétique des quatre enzymes du stress avec l'insuline avec CLUSTER Omega.....	35
Figure26 : interaction pour les gène des enzymes antioxydants avec l'insuline.....	36
Figure27 : représentation des réseaux d'interaction entre les protéines antioxydants CAT, GXP1, SOD1,SOD2.....	36
Figure28 : Les résultats d'interactions par l'outil Gene Mania.....	37
Figure29 : Resaux d'ineraction entes les gènes anitiaxydantepar BioGird	38
Figure30 : Représentation des variantes et des exons de SOD1 par UCSC.....	39
Figure31 : Représentation des variantes et des exons de SOD2 par UCSC.....	39
Figure32 : Représentation des variants et des exons de la CAT par UCSC.....	40
Figure33 : Représentation des variants et des exons de GPX1 par UCSC.....	40

Abréviations :

Aa: Acide aminé

ADN : Acide Désoxyribonucléique

Ap2 : Activating Protein 2

BLAST: Basic Local Aligement Search Tool

BLASTn : Basic Local Aligement Search Tool nucleotides

BLASTp : Basic Local Aligement Search Tool protein

CAT : Catalase

DT1 : diabète type 1.

Gpx : Glutathionne peroxydase cellulaire

H⁺ : Ion hydrogen

H₂O₂ :Peroxyde d' hydrogène

MnSOD : Manganèse superoxyde dismutase

NF- κ B : Nuclear factor-kappa B

OH : L'hydroxyle

O²⁻ :Anion superoxyde

ROS :Radicaux oxygénés libres (Reactive oxygen species)

SMS : Séquence Manipulation Suite

SNP : Single Nucleotide Polymorphism

SOD1 : Superoxyde dismutase cytosolique

SOD2 : Superoxyde dismutase mitochondriale

SWISS-PROT: Swiss Protein Data Bank

TFD: Transcription Factor Database

SOMMAIRE

	Numéro de page
Résumé	
Abréviations	
Liste des tableaux	
Liste des figures	
Introduction	1
.....	2
Chapitre I : Généralités	
I-Diabète	3
I.1- Diabète de type 1	3
I.1.1- Ethiopathologie du diabète de type 1.....	3
I.1.2- Terrain génétique de susceptibilité	3
II-Relation diabète type 1 et stress oxydant	4
II.1-Les antioxydants	4
III-Gènes et biochimie des enzymes antioxydantes	5
III.1-Catalase	5
III.2-La glutathionne peroxydase (GPx1)	8
III.3- Superoxyde Dismutase1 (SOD1)	10
III.4-Superoxyde dismutase mitochondrial (SOD2)	11
III.5- Gène d'insuline	13
IV- Facteurs de transcription	
IV.1-Facteurs de transcription de la SOD1.....	15
IV.2-Facteurs de transcription de l'insuline.....	17
IV.3-Le facteur de la transcription NF-κB.....	18
V- Polymorphisme	18
Chapitre II : Matériel et Méthodes	
II.1.Matériel	20
II.1.1. Matériel non biologique.....	20
II.1.1.1.Outils bio- informatiques.....	2
II.1.2. Matériel biologique	23
II.2. Méthode d'étude.....	25
II.2.1.Récolte des données moléculaires	26
II.2.2.Traitement des données moléculaires.....	26
II.2.3 Les ORF s des nos séquences antioxydants	26
II.2.4.Outils Blast	27
II.2.5. l'alignement et la comparaison des séquences nucléotidique et peptidique	27
II.2.5.1.UNIPROT.....	28
II.2.6.Etude des Interactions entre les gènes des enzymes antioxydants	28

II.2.6. 1. Outils STRING.....	28
II.2.6. 2 .Outils Gene MANIA	29
II.2.7.La recherche des variantes de ces gènes antioxydants	29
II.2.7.1.Le programme Ensemble	29
Chapitre III : Résultats et discussion.....	30
III.1. 1. Etudes des similitudes par l’outil Blast	30
III.1.2. Alignement des séquences par UNIPROT.....	33
III.1.3.Interactions entre les enzymes antioxydant et le gène d’insuline	37
III.1.4.Les résultats d’ outil UCSC.....	41
III.1.5. Recherche des mutations polymorphique des gènes antioxydants	43
III.1.2 Discussion.....	44
Chapitre IV : Conclusion.....	45
Références bibliographiques	
Biblio –Web	
Annexe	

Introduction

Le stress oxydatif est impliqué dans de nombreuses conditions physiopathologiques. Les outils pour étudier le stress oxydatif sont maintenant en expansion et résulte de plusieurs efforts fournis pour l'étude du génome humain. (*Forsberg et al, 2001*)

En effet, les progrès de la biologie moléculaire et le projet du génome humain « human genome project » ont élargi nos outils pour étudier les mécanismes patho-génétique des maladies humaines multifactoriel tel que le diabète type1.

Le diabète insulino-dépendant ou diabète de type 1 (DT1) est une maladie chronique touchant l'enfant, le jeune adulte et l'adulte ayant une prédisposition génétique. Il en résulte une destruction des cellules β des îlots de Langerhans aboutissant à une insulino-déficience (*Kukko et al, 2003*). Cet état d'apoptose des cellules bêta pancréatiques est une conséquence d'un état d'hyperglycémie chronique conduisant à un stress oxydant, c'est-à-dire un déséquilibre entre pro-oxydants et antioxydants au profit des premiers. Il en résulte une oxydation accrue des cibles cellulaires et/ou une baisse des systèmes de défense antioxydants (*BONNEFONT-ROUSSELOT et al., 2004*). Parmi les espèces réactives oxygénées (ROS), on trouve le peroxyde d'oxygène (H_2O_2) qui peut être converti en oxygène O_2 et en eau H_2O par la catalase. La SOD et le GPx viennent renforcer l'élimination de l'excès des radicaux libres et de la défense oxydative (*BONNEFONT-ROUSSELOT et al., 2004*). En effet, l'état d'hyperglycémie est une conséquence d'une cascade d'interaction moléculaire.

En plus de ces interactions, il existe une susceptibilité génétique que l'on classe parmi les variantes génétiques, principalement sous la forme de polymorphismes à un seul nucléotide appelé SNPs (single nucléotide polymorphisme) et qui sont pertinents pour l'étude par une approche moléculaire et épidémiologique du stress dans plusieurs pathologies et chez différentes populations.

Notre travail s'intéresse à l'étude de quelques gènes impliqués dans la défense antioxydante **CAT**, **SOD1**, **SOD2**, **GPx1** chez le diabétique de type1 et plus particulièrement à l'étude des interactions moléculaires de ces gènes antioxydant aux gènes de l'insuline (protéine hypoglycémisante) afin de prédire leurs rôles et fonctions métaboliques et de comprendre la relation qui peut exister entre ces gènes et le diabète type1.

Pour cette étude nous avons utilisé des outils bioinformatiques pour répondre aux questions suivantes :

Ya t-il des régions similaires entre les différentes régions codante et non codantes entre ces gènes. si oui quel est leurs rôle dans la régulation métabolique.

Quel sont les différents variantes polymorphiques (SNPs) qui peuvent changer l'expression de ces gènes et inhiber la fonction de ces protéines.

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

- ✓ Le premier chapitre va représenter les généralités décrivant la relation entre physiopathologie du diabète type1, les gènes de la défense oxydative et une présentation de quelques bases de données utilisée comme outils bioinformatique pour traiter les données biologiques.
- ✓ Le deuxième chapitre va décrire les outils bioinformatiques utilisé et les différentes étapes utilisés pour répondre a notre problématique.
- ✓ Le troisième chapitre va rapporter les résultats obtenus ainsi qu'une discussion à lumière des données bibliographiques.

Enfin, on terminera ce mémoire par une conclusion générale ou nous dégagerons les principaux résultats avec quelques perspectives.

I-Diabète

Le terme de diabète regroupe plusieurs maladies, ayant en commun l'hyperglycémie (*Grimaldi, 2012*). La nouvelle classification des diabètes proposée par l'OMS en 1997 distingue quatre types de diabète : le diabète de type 1 (anciennement diabète insulino-dépendant), le diabète de type 2 (diabètes non insulino-dépendants), les « autres diabètes spécifiques » (ou diabètes secondaires) et le diabète gestationnel (*Guillausseau, 1997*).

I.1- Diabète de type 1

Le diabète de type 1 est dû à une destruction auto-immune des cellules β des îlots pancréatiques. La réaction auto-immune survient sur un terrain de susceptibilité génétique à la suite des facteurs déclenchants et peut être dépistée avant l'apparition de l'hyperglycémie par des dosages sanguins d'auto-anticorps (*Grimaldi., 2000; Rodier., 2001*).

I.1.1- Ethio-pathologie du diabète de type 1

Le diabète de type 1 est dû à une destruction auto-immune des cellules β des îlots pancréatiques. La réaction auto-immune survient sur un terrain de susceptibilité génétique à la suite des facteurs déclenchants et peut être dépistée avant l'apparition de l'hyperglycémie par des dosages sanguins d'auto-anticorps (*Grimaldi., 2000; Rodier., 2001*).

I.1.2- Terrain génétique de susceptibilité

Le diabète de type 1 représente une maladie hétérogène, se développe sur un terrain génétique particulier : il s'agit d'une susceptibilité plurigénique avec au moins 10 gènes en cause. Ce caractère héréditaire se traduit par un risque accru de la maladie chez les apparentés d'un sujet diabétique de type 1 (*Eisenbarth, 1994*). Le risque pour une mère diabétique insulino-dépendante d'avoir un enfant diabétique est environ 2 % alors que le risque est de 4 à 5 % lorsque c'est le père qui est diabétique insulino-dépendant. Les différents risques sont résumés dans le tableau I. (*Grimaldi., 2000*).

Le principal gène concerné se situe sur le chromosome 6 au niveau des gènes du système HLA de classe II avec un risque relatif de 3 à 5, lorsqu'il existe un antigène HLA DR3 ou DR4. Le risque relatif atteint 20 à 40 lorsque les deux antigènes DR3 et DR4 sont associés, ce qui veut dire que l'association DR3-DR4 est fréquente dans la population diabétique alors qu'elle est exceptionnelle dans la population non-diabétique

II-Relation diabète type 1 et stress oxydant :

Le stress oxydatif est défini comme un déséquilibre entre les prooxydants et antioxydants, et entre la production d'espèces réactives d'oxygène et les défenses antioxydantes (*Betteridge., 2000*)

Cependant, certaines preuves suggèrent que le stress oxydatif est également impliqué dans l'origine du diabète de type 1, en particulier par l'apoptose de cellules β pancréatiques, qui se caractérisent par la faible teneur en enzymes antioxydantes par rapport aux autres tissus (*BONNEFONT-ROUSSELOT et al. , 2004*)

II.1-Les antioxydants :

Les défenses antioxydantes éliminent les excès de Peroxydes et radicaux libres d'oxygène, avec une diminution du déséquilibre entre la production et le piégeage des radicaux libres dans le but de réduire le stress oxydatif dans l'organisme. Plusieurs études suggèrent que en présence d'un diabète de type 1 il existe une faible efficacité de la défense oxydative, à la fois enzymatique (superoxyde dismutase, catalase, Glutathion peroxydase) et non enzymatique (vitamine A, C et E) (figure1). Ce qui suggère que ces enzymes antioxydantes jouent un rôle protecteur contre le stress oxydatif dans le diabète de type 1 Et leurs gènes pourraient contribuer à la susceptibilité à cette maladie auto-immune (*Chistiakov et al., 2004*)

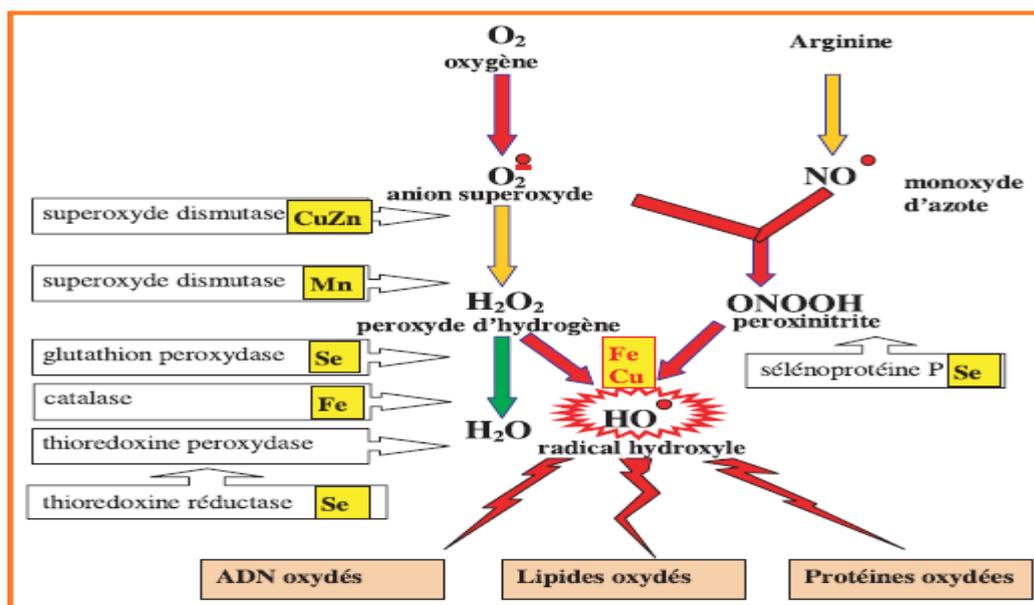


Figure 1 : Mode d'action des principaux systèmes enzymatiques antioxydants et de leurs cofacteurs métalliques (*Favier., 2003*)

III-Gènes et biochimie des enzymes antioxydantes

III.1-Catalase

Le catalase est un antioxydant endogène impliqué dans les voies de neutralisation des espèces réactive de l'oxygène (ROS) et dans les mécanismes contre le stress oxydatif.

La catalase est une hémoprotéine tétramère qui catalyse la conversion du peroxyde d'hydrogène (H_2O_2) en H_2O et en oxygène (O_2), empêchant ainsi les lésions cellulaires par les ROS.

-Le gène CAT humaine, est codé par le chromosome nucléaire 11p13, composé de 13 exons et 12 introns.

Des clones de phage recombinant contenant le gène de la catalase humaine ont été isolés et caractérisés. Le gène a une longueur de 34 kb et est divisé en 13 exons. On a put prédire ainsi que la région en amont du gène ne contient pas de boîte TATA. Cette région est riche en GC (67%) et contient plusieurs séquences CCAAT et GGGCGG qui peuvent faire partie du promoteur. La traduction de l'ARNm de la catalase semble commencer immédiatement en amont du résidu Ala amino-terminal de la catalase (*Quan et al., 1968*)

Selon la littérature, les enzymes (protéase) peuvent avoir des effets néfastes sur l'expression ou la fonction de la CAT, ce qui peut entraîner une diminution de l'activité enzymatique de la CAT activité et une plus grande sensibilité aux ROS. Par conséquent, les variations alléliques dans la CAT peuvent modifier la désintoxication des ROS et augmenter le stress oxydatif, impliquant des lésions oxydatives de l'ADN impliquant des lésions oxydatives de l'ADN et modulant le risque aux maladies (*Goth et al., 2004; Forsberg et al., 2001*)

Une série de les polymorphismes à un seul nucleotide (SNP) dans le gène CAT ont ont été identifiés. Les polymorphismes génétiques situés dans la région du promoteur pourrait influencer sur les taux de transcription, ce qui dans une faible activité CAT. La mutation la plus étudiée est la rs1001179 SNP, qui est situé dans le 50-UTR, la base 262 paires du site de début de la transcription du gène CAT (*Goth et al., 1997 ; Jiang et al., 2001*)

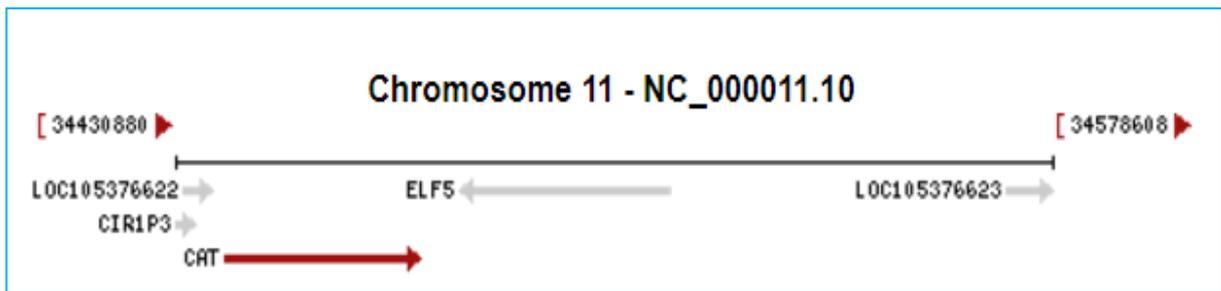


Figure 2: Localisation du gène de la catalase au niveau du chromosome 11.

(URL : <https://www.ncbi.nlm.nih.gov/gene/847>)

Grace a la technologie NGS (Next Generation Sequencing) des RNA-seq ont été isolé a partir des échantillons de tissus provenant de 95 individus humains représentant 27 tissus différents afin de déterminer la spécificité tissulaire de tous les gènes codant pour la protéine (figure 3)

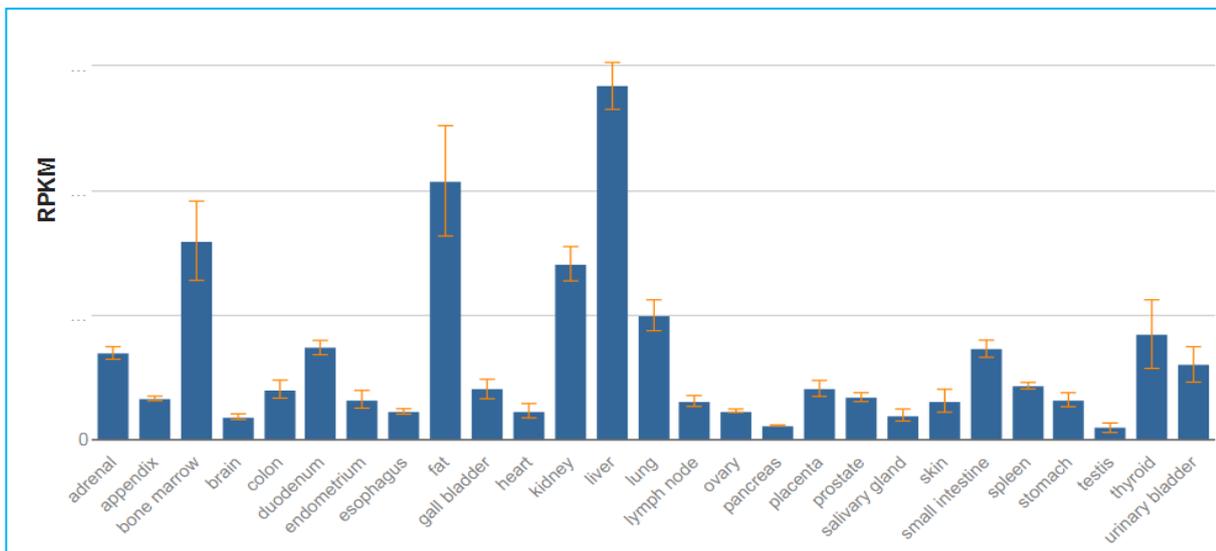


Figure 3: Histogramme représentatif du niveau d’expression des RNA-Seq de la catalase dans différents tissus.

(URL : <https://www.ncbi.nlm.nih.gov/gene/847>)

➤ Catalase et diabète type1

Des travaux ont montré que les concentrations élevées d' H_2O_2 ont a montré que la peroxydation lipidique entraîne sur l'ADN par des produits de peroxydation lipidique, en particulier dans le cas du diabète de type 1 (*Collins et al., 2001*). Parmi les ROS, le peroxyde d'hydrogène contribue dans la signalisation faisant intervenir le récepteur de l'insuline. Ainsi, le H_2O_2 , peut jouer un rôle clé dans la modulation des voies de transduction du signal régulées par l'insuline. Par conséquent, l'inactivation du peroxyde d'hydrogène par la catalase pourrait représenter une étape cruciale dans l'élimination des ROS du milieu intracellulaire dans les cellules productrices d'insuline jouant ainsi, un rôle de protection contre l'apparition précoce du diabète de type 1 (*Cavarape et al., 2001*)

III.2-La glutathionne peroxydase (GPx1)

Le gène GPX1 cellulaire humain a été localisé au niveau de la région 3p 21.31 du chromosome 3, GPx1 a une longueur de 1183 bases et contient 3 introns et 2 exons (*Chada et al.,1990*)

La protéine codée par ce gène appartient à la famille de la glutathionne peroxydase, dont la fonction est de catalyser la réduction des hydroperoxydes organiques et du peroxyde d'hydrogène (H_2O_2) par le glutathion en protégeant ainsi les cellules contre les dommages oxydatifs.

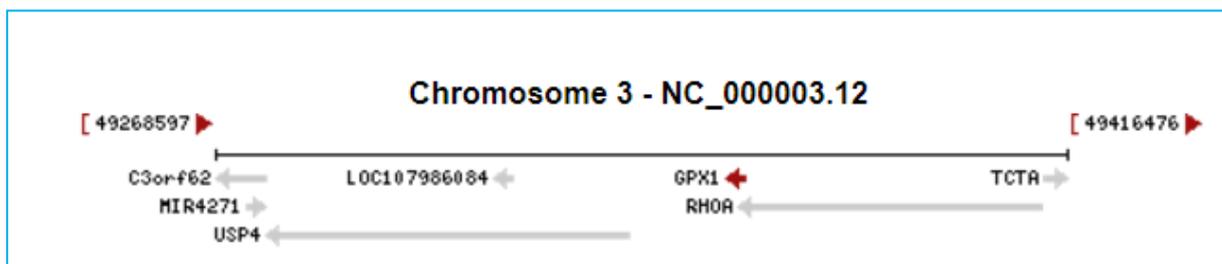


Figure 4: Localisation du gène GPx1 au niveau du chromosome3.

(URL : <https://www.ncbi.nlm.nih.gov/gene/2876>)

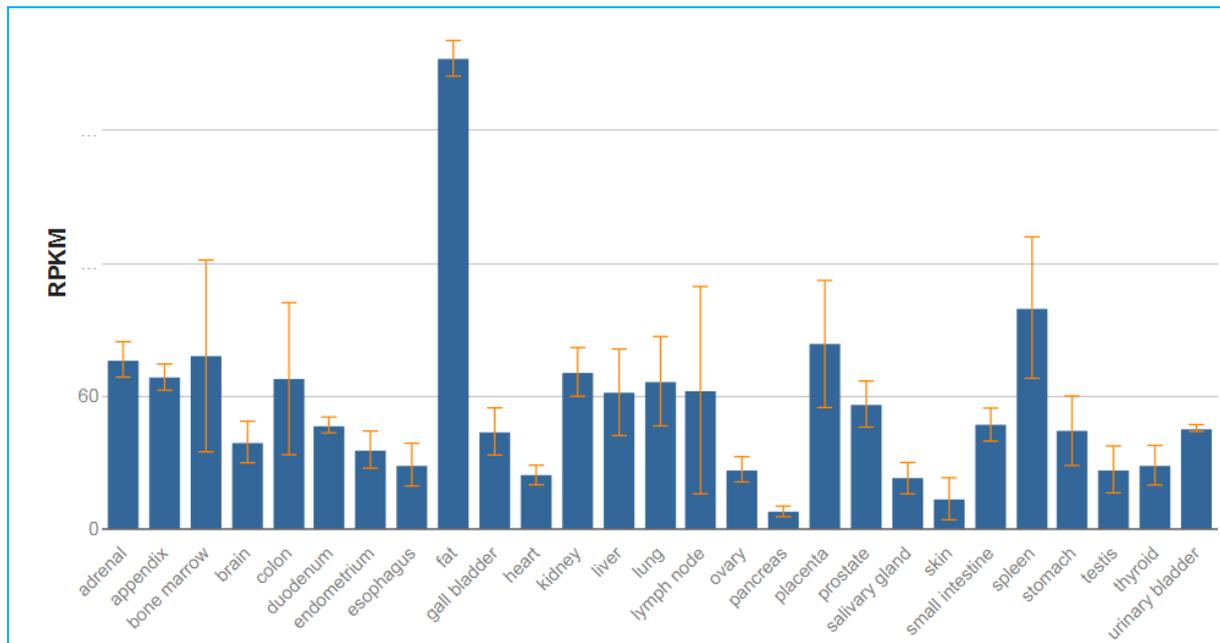


Figure 5: Histogramme représentatif du niveau d'expression des RNA-Seq de la GPx1 dans différents tissus.

(URL : <https://www.ncbi.nlm.nih.gov/gene/2876>)

D'autres études indiquent que H_2O_2 est également essentiel pour la transduction du signal médiée par le facteur de croissance, la fonction mitochondriale et le maintien de l'équilibre redox thiol; Par conséquent, en limitant l'accumulation de H_2O_2 , les glutathion peroxydases sont également impliquées dans la modulation de ces processus.

Plusieurs isozymes de cette famille de gènes existent chez les vertébrés, qui varient selon la localisation cellulaire et spécificité du substrat. Cet isozyme est le plus abondant, est omniprésent. Il a été localisé dans le cytoplasme, et dont le substrat préféré est le peroxyde d'hydrogène. Sur son site actif, on retrouve, une sélénoprotéine contenant l'acide rare sélénocystéine (Sec). Sec est codé par le codon UGA, qui normalement est un signal la fin de la traduction. Les 3' UTR des ARNm de la sélénoprotéine contiennent une structure de tige en boucle conservée, désignée par un élément de séquence d'insertion Sec (SECIS), qui est nécessaire pour la reconnaissance d'UGA comme codon Sec, plutôt que comme un signal d'arrêt. Ce gène contient une répétition de tri nucléotide GCG dans le cadre dans la région codante. Également, il a été trouvés une répétition des trois allèles 4, 5 ou 6 dans la population humaine. L'allèle avec 4 répétitions GCG a été significativement associé au risque de cancer du sein chez les femmes pré ménopausées. Alternativement, des variantes de transcription

épissées ont été trouvées pour ce gène. Les pseudogènes de ce locus ont été identifiés sur les chromosomes X et 21. (<https://www.ncbi.nlm.nih.gov/gene/2876>, 2017)

(Fourni par Ref Seq2017)

Des études ont publié des résultats cohérents avec l'implication de GPX1 dans le mécanisme de protection rénale contre le stress oxydatif chez les patients diabétiques de type 1.

(*Mohammedi et al., 2015*).

III. 3- Superoxyde Dismutase1 (SOD1)

Le gène SOD1 est un gène de 11 Kb de longueur. Il comprend 5 exons et 4 introns, il a été localisé dans le chromosome 21 au niveau de la région 21 q22. (*Levanon et al., 1985*)

La protéine codée par ce gène lie les ions cuivre et zinc. C'est l'une des deux isozymes responsables de la destruction des radicaux libres dans le corps. Le premier isozyme est une protéine cytoplasmique soluble (SOD1), agissant comme un homodimère pour convertir ROS en oxygène moléculaire et en peroxyde d'hydrogène. L'autre isozyme est une protéine mitochondriale (SOD2). Les mutations dans ce gène ont été impliquées comme des causes de plusieurs pathologies dont le diabète. Des variantes de transcription rares ont été signalées pour ce gène.

Le gène humain SOD1 se trouve sur le chromosome 21q22.11. Il code pour un monomère polypeptidique SOD1 (153 acides aminés, poids moléculaire 16 kDa). Plus précisément, ce gène est situé à partir de la paire de bases 33,031,935 à la paire de bases 33,041,241 avec une taille génomique de 9307 pb, selon UCSC Genome Browser

Nombreux polymorphismes ont été identifiés dans le gène SOD1, principalement distribué dans les différentes régions chromosomiques, y compris le promoteur

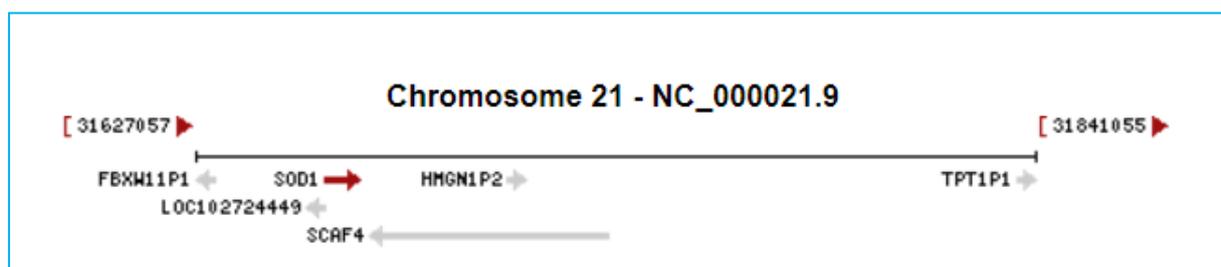


Figure 6: Localisation du gène SOD1 au niveau du chromosome 21.
(URL : <https://www.ncbi.nlm.nih.gov/gene/6647>)

Le niveau d'expression des RNA-Seq de la SOD1 dans différents tissus a montré que la SOD1 à un niveau élevé au niveau du foie et du rein (figure 7).

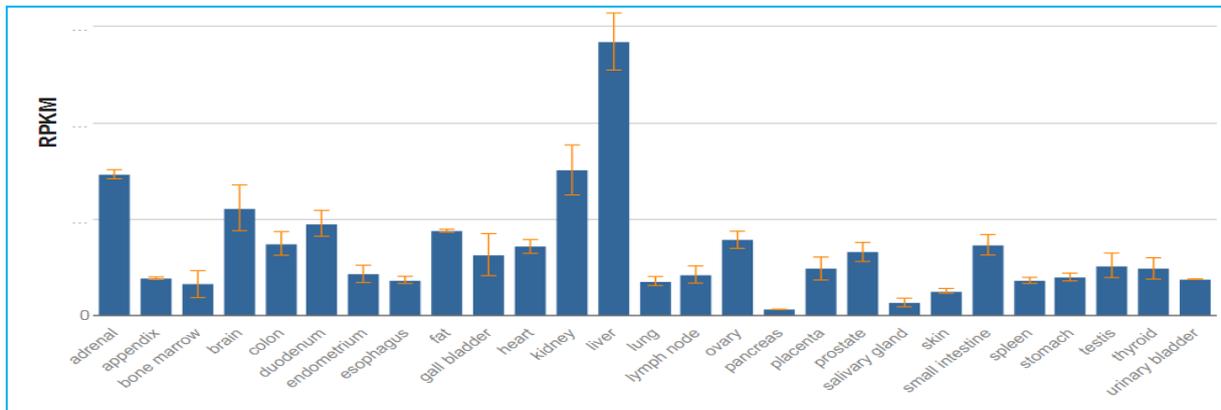


Figure 7: Histogramme représentatif du niveau d'expression des RNA-Seq de la SOD1 dans différents tissus.

(URL: <https://www.ncbi.nlm.nih.gov/gene/6647>)

D'autres travaux ont montrés l'implication du stress oxydatif dans la pathophysiologie de la néphropathie diabétique montrant ainsi le rôle majeur des enzymes antioxydantes comme mécanisme de protection rénale. Ces mêmes travaux ont montré une association des variations allélique de la SOD1 aux complications rénale chez les diabétiques de type1

(Mohamedi et al , 2011)

III.4-Superoxyde dismutase mitochondrial (SOD2)

Le gène de SOD2 humain est un gène de 15 Kb de longueur et formé de 5 exons et 4 introns localisé à la région 6q25 du chromosome 6. Ce gène fait partie de la famille des superoxyde dismutase fer / manganèse.

Il code pour une protéine mitochondriale qui forme un homotetramer et lie un ion manganèse par des sous-unités. Cette protéine se lie aux sous-produits superoxyde de la phosphorylation oxydante et les convertit en peroxyde d'hydrogène et en oxygène diatomique.

Les mutations de ce gène ont été associées à la cardiomyopathie idiopathique (IDC), au vieillissement prématuré, à la maladie des neurones moteurs sporadiques et au cancer. L'épissage alternatif de ce gène aboutit à de multiples variantes de transcription

(fourni par Ref Seq ,Avril2016)

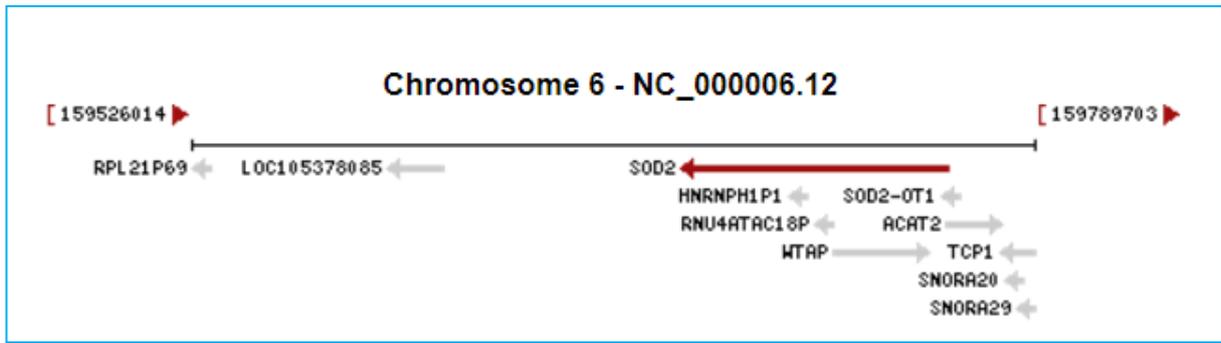


Figure 8: Localisation du gène SOD2 au niveau du chromosome 6.

URL (<https://www.ncbi.nlm.nih.gov/gene/6648>)

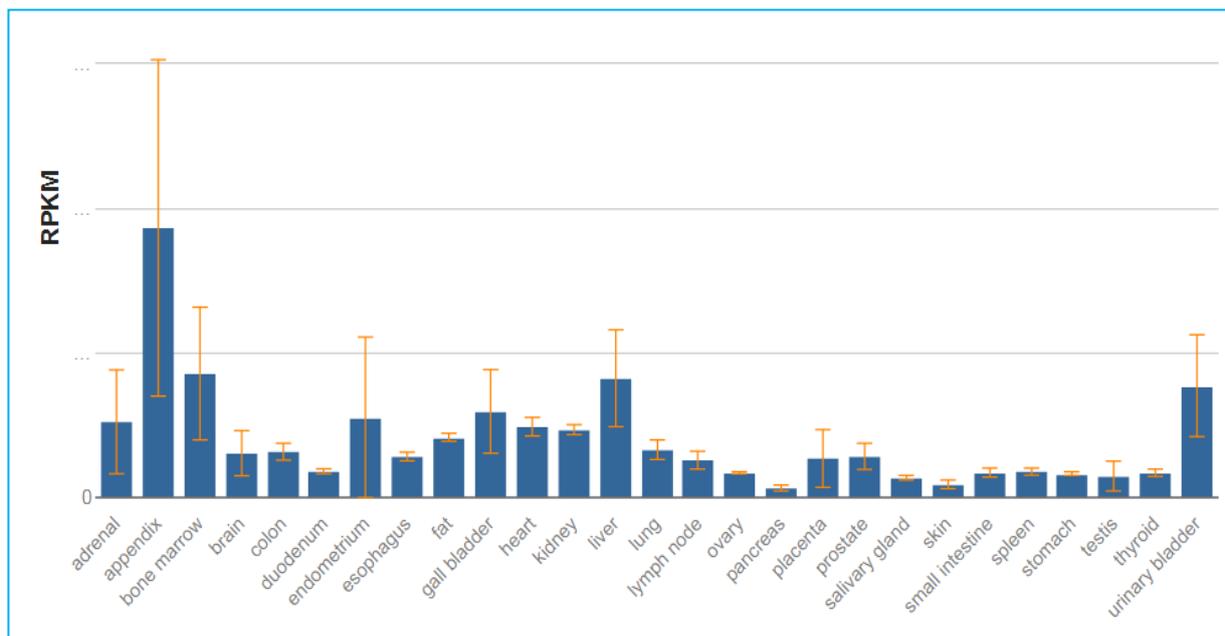


Figure9 : Histogramme représentatif du niveau d’expression des RNA-Seq de la SOD2 dans différents tissus

(<https://www.ncbi.nlm.nih.gov/gene/6648>)

III.5- Gène de l'insuline

L'insuline est une hormone polypeptidique produite par le pancréas, qui permet au glucose de pénétrer dans les cellules de l'organisme où il est transformé en énergie. Elle contribue à maintenir constant le taux de la glycémie, cette action est largement utilisée en thérapeutique dans le traitement du diabète.

La molécule d'insuline est un polypeptide d'un poids moléculaire de 6 KDA. C'est un hétérodimère constitué de deux chaînes polypeptidiques, la chaîne α à 21 acides aminés et la chaîne β à 30 acides aminés, reliées entre elles par deux ponts disulfures (*Idelman et Verdetti., 2000*).

L'insuline est codée par le gène de l'insuline situé sur le chromosome 11p15.5 (*Ohneda et al, 2000*). Il se trouve à proximité du gène de l'insulin-like growth factor 2 (IGF-2).

L'expression de l'insuline est limitée à la cellule β et la transcription de l'insuline est contrôlée par le promoteur de l'insuline, et en particulier une région de 340 pb fortement conservée située immédiatement en amont du site de départ de l'initiation de la transcription (*Poitout et al.,2006*).

Ce gène contrôle la synthèse d'un précurseur de haut poids moléculaire, la pré-pro-insuline. Il est long de 1355 paires de bases et sa région codante comporte trois exons séparés de deux introns qui comportent respectivement 179 et 786 paires de bases. Le premier exon, long de 42 paires de bases, contrôle la synthèse du peptide signal. Le deuxième exon (187 paires de bases) code pour la chaîne B de l'insuline et une partie du peptide connecteur, le troisième exon dont la taille est de 220 paires de bases, code pour la chaîne B et le reste du peptide connecteur.

5'UTR et 3' UTR sont des séquences non traduites. L'extrémité 5' est flanquée d'une région à très haut degré de variabilité dont les trois allèles (I,II et III) peuvent être distinguées en fonction de leur taille (fig 10)

L'organisation de la partie proximale (-340 pb à +1 pb) du promoteur d'insuline comprenant des éléments critiques d'activation de la transcription.

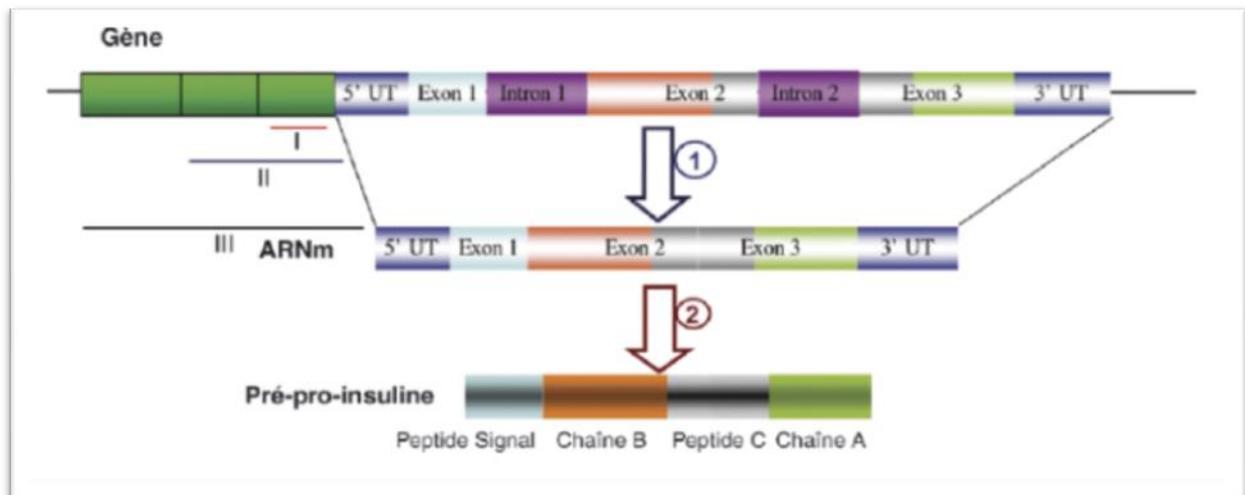


Figure 10 : structure du gène de l'insuline (*Steiner et al . ,1985*)

1 : Transcription/épissage ; 2 : traduction ; 3 :ARNm (acide ribonucléique messager)

L'insuline est produite dans les cellules β qui constituent 75% des îlots de Langerhans du pancréas. La biosynthèse de l'insuline s'amorce dans le noyau des cellules β , à partir de l'information contenue dans le code génétique, située sur le chromosome 11 chez l'homme (*Langlois, 2008*). La transcription du gène aboutit à un ARN messager qui est traduit en pré pro insuline, une protéine de 11,5 KDA, le pré pro insuline en cours d'élongation est rapidement déversée dans le réticulum endoplasmique où les enzymes protéolytiques clivent la séquence finale, formant ainsi la pro-insuline qui est un peptide de 9 KDA contenant les chaînes α et β de l'insuline connectées entre elles par le peptide C. Cette étape dure entre 10 et 20 minutes. Après son passage au réticulum endoplasmique, la pro-insuline est transportée dans des microvésicules intermédiaires vers l'appareil de Golgi (*Baalbaki, 2012*). C'est dans cet organe que la proinsuline est clivée pour donner le peptide-C (31 acides aminés ayant un poids moléculaire de 3 KDA) et un peptide bicaténaire ; l'insuline d'une taille finale de 51 acides aminés et de poids moléculaire de 6KDA. Ces deux peptides sont stockés dans des granules jusqu'à sécrétion (*Auberval , 2010*).

Le niveau d'expression des RNA-Seq de l'insuline (INS) dans différents tissus est plus élevé au niveau du pancréas (Figure 11)

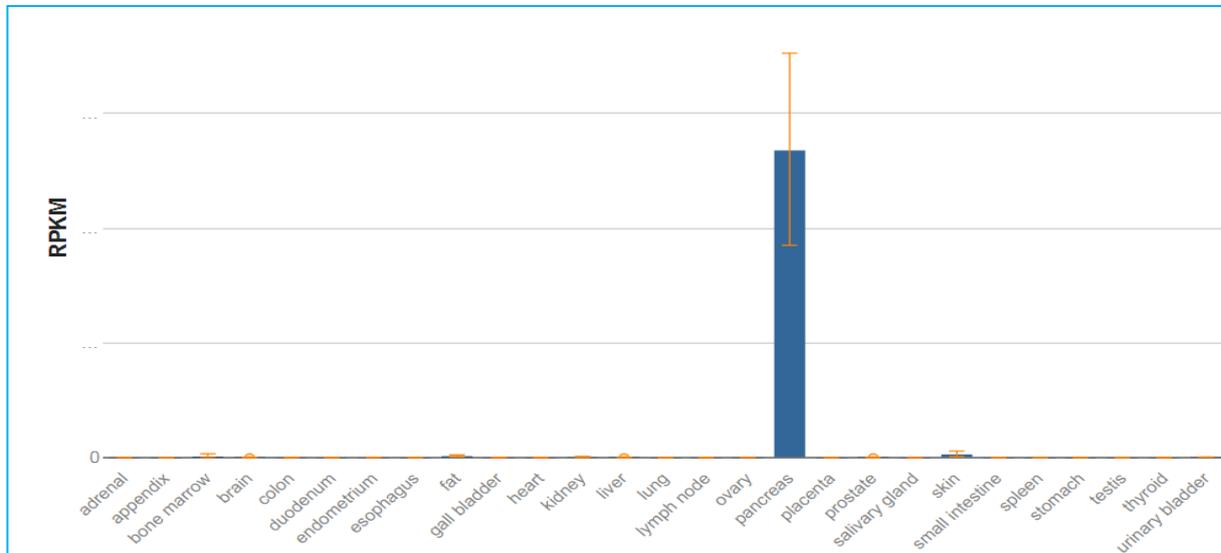


Figure 11: Histogramme représentatif du niveau d'expression des RNA-Seq de l'INS dans différents tissus

(<https://www.ncbi.nlm.nih.gov/gene/3630>)

IV- Facteurs de transcription

IV.1-Facteurs de transcription de la SOD1

Le promoteur proximal du gène SOD1 humain, impliqué dans la transcription basale, a été bien étudié, et il contient non seulement la boîte TATA, mais aussi une boîte CCAAT et une région riche en GC, reconnues par les protéines de liaison CCAAT / enhancer (C / EBP) et la protéine de spécificité 1 (Sp1), respectivement. D'autres sites de liaison pour les facteurs de transcription Egr1, AP1, AHR, Nrf2, NF- κ B et TR ont également été identifiés et vérifiés par des études fonctionnelles (Figure12). Ces facteurs de transcription sont impliqués dans la régulation de l'expression génique inducible de SOD1 dans des conditions extra-intracellulaires spécifiques. (Yoo *et al.*,1999).

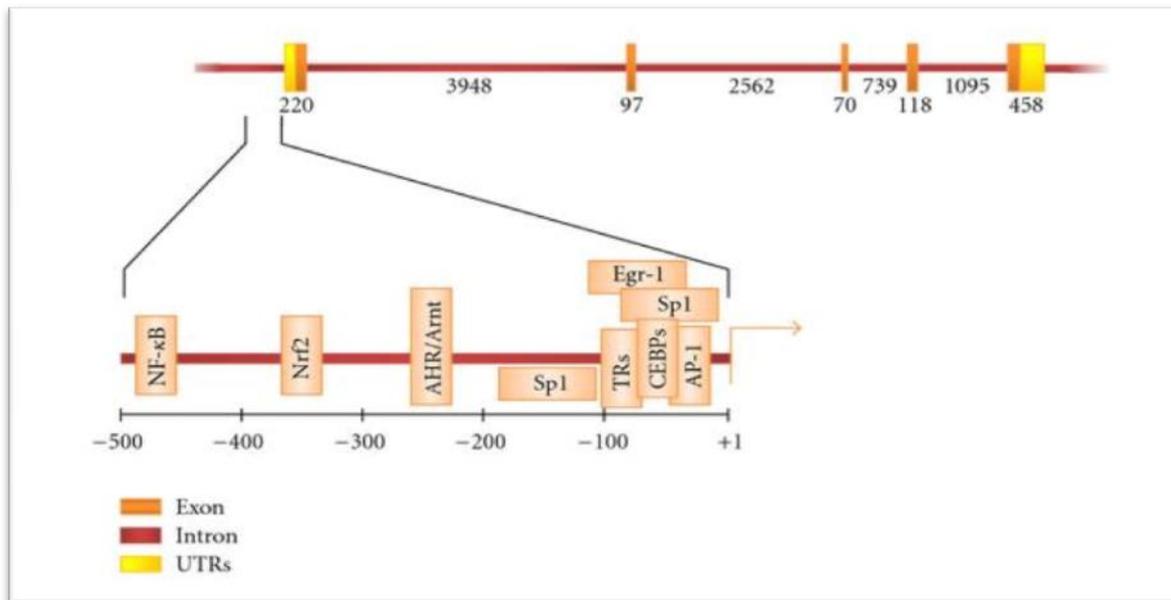


Figure12 : Organisation génomique du gène humain SOD1 (*Yoo et al.,1999*)

IV.2-Facteurs de transcription de l'insuline

Les facteurs de transcription qui reconnaissent des petites séquences d'ADN localisées en amont du site de démarrage de la transcription (comme le facteur SP-1 qui reconnaît les boîtes GC), sont ubiquitaires. Ils augmentent l'efficacité de la transcription des gènes contenant les GC dans leurs promoteurs.

- **PDX-1 :** Le facteur de transcription de l'insuline le plus important est le facteur de transcription de l'homéodomaine pathétique pancréatique / duodénale homeobox-1 (PDX-1). Dans l'îlot adulte, PDX-1 est seulement exprimé dans les cellules β pancréatiques et joue un rôle important dans la transcription du gène de l'insuline stimulée par le glucose. Le PDX-1 est également essentiel pour la maintenance du phénotype des cellules β et du développement du pancréas. (*Ohneda et al.,2000*)

Le PDX-1 se lie à la boîte A3 du promoteur d'insuline. Le PDX-1 a peu d'activité, mais il devient un facteur puissant lorsqu'il interagit avec l'hétérodimère des protéines basiques de hélice-boucle-hélice (bHLH) qui se lie à la boîte E1. Ces hétérodimères de BHLH sont constitués d'une protéine BHLH de Classe A qui est exprimée de manière omniprésente dans de nombreux tissus différents et une protéine BHLH de Classe B exceptionnellement exprimée dans les cellules β .

- **MafA** : Un autre facteur important de transcription de l'insuline est le MafA qui appartient à la famille de facteurs de transcription Maf. Cette famille de facteurs de transcription contient un motif basique suivi d'une fermeture à glissière de leucine. En outre, MafA contient également un domaine acide qui agit comme un domaine d'initiation de la transcription (*Zhang et al., 2005*). MafA se lie au site C1 du promoteur d'insuline. MafA est un activateur transcriptionnel spécifique de la cellule β et du pancréas régulés par le glucose pour le gène de l'insuline (*Kataoka et al., 2004*) qui produit une activation synergique avec PDX-1 ou BETA2. En plus de PDX1, E47 / BETA2 / NeuroD et MafA, d'autres facteurs influencent la transcription de l'insuline. Les histones et les protéines liées à l'ADN du groupe de haute mobilité (HMG) augmentent la liaison des hétérodimères PDX-1 et bHLH aux sites A et E. Une protéine HMG importante est HMG 1 (Y) qui se lie au site A3 / A4 du promoteur de l'insuline (*Ohneda et al., 2000*)

D'autres protéines qui contribuent à la transcription de l'insuline comprennent des membres des facteurs nucléaires hépatiques et des familles PAX .

Le promoteur de l'insuline est responsable de la régulation métabolique spécifique des tissus du gène de l'insuline. Les éléments d'activation de la transcription les plus critiques du promoteur d'insuline sont les sites A3, C1 et E1 (fig13).

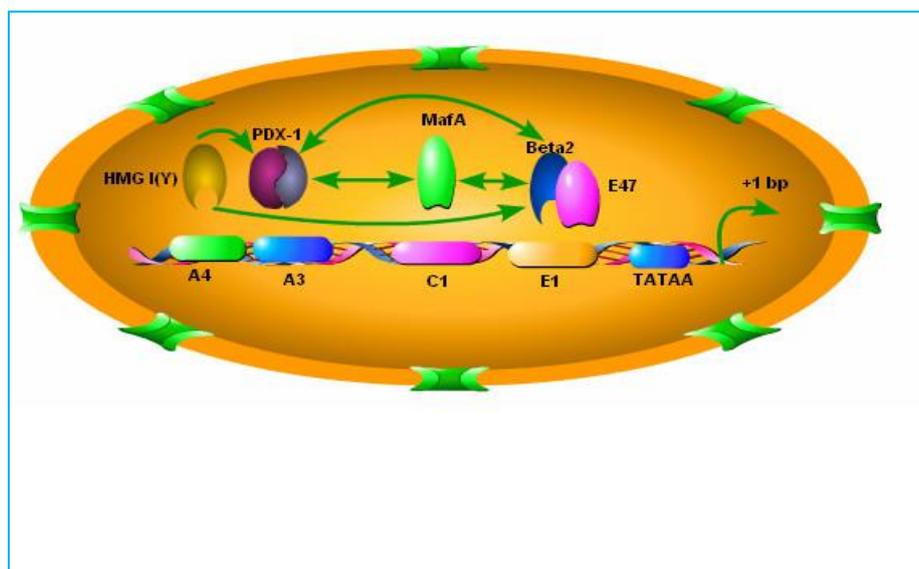


Figure13: La région promoteur du gène de l'insuline (*German et al., 1995*)

IV.3-Le facteur de la transcription NF- κ B :

Le terme NF- κ B désigne une famille de cinq facteurs de transcription structurellement liés (p50, p52, RelA / p65, c-Rel et RelB), tous contenant le domaine d'homologie Rel (RHD) dans l'extrémité N-terminale et agissant comme homodimères et des complexes hétérodimères de liaison à l'ADN . Leur fonctionnalité et leur localisation nucléaire sont contrôlés par une famille de protéines inhibitrices, appelées I κ Bs (I κ B). Dans les cellules non stimulées, les dimères NF- κ B sont liés aux protéines I κ B inhibitrices et sont ainsi séquestrés dans le cytoplasme en tant que complexes inactifs. Plusieurs études ont montré que l'activité NF- κ B est induite dans la plupart des types de cellules en réponse à une grande variété de stimuli, allant des cytokines, des rayonnements et du stress oxydatif (comme l'exposition au H₂O₂), avec des rôles majeurs dans la coordination de l'immunité innée et adaptative, l'activation et la prolifération des cellules, la survie, le développement et l'apoptose .

Le NF- κ B a été l'un des premiers facteurs de transcription démontrés comme régulation redox. En effet, Rojo et ses collègues ont montré que le traitement cellulaire avec H₂O₂ déclenche les cascades PI3K / Akt, qui participe à l'activation NF- κ B et à l'induction transcriptionnelle de la SOD1 subséquente. En effet, les auteurs ont identifié un site de liaison p65-NF- κ B dans le promoteur SOD1 humain (GGTAAGTCCC), et ils ont démontré que le NF- κ B activé par Akt présente une liaison accrue à cette séquence, en médiant la régulation positive de l'expression de SOD1 (*O'Dea et Hoffmann., 2000*) .

V- Polymorphisme

Le génome humain contient environ 25 000 gènes codant pour des protéines soit environ moins du (1,5%) du génome. Le projet de séquençage humain, le séquençage individuel de génomes humains, et l'étude des maladies génétiques ont démontré la présence de nombreuses variations du génome d'un individu à l'autre. Ces variations ne sont pas appelées mutations mais polymorphismes car elles ne provoquent pas de maladie a priori.

Un polymorphisme peut se localiser n'importe où sur le génome que ce soit dans un gène ou non avec une fréquence variable dans les populations (différences ethniques et géographiques).

Un polymorphisme existe si la fréquence de celui-ci est supérieure ou égale à 1 %. Si la fréquence de cette variation est inférieure à 1 %, on parle de variant.

Le polymorphisme est dit bi-allélique si une base (A, C, G ou T) peut être remplacée par une autre base sur un gène.

Les anglosaxons définissent le polymorphisme bi-allélique par le terme *single nucleotide polymorphism* (SNP), terme souvent repris en français. Les SNPs sont distribués partout dans le génome humain et sont souvent utilisés dans les études d'association de maladies avec des gènes (notamment pour les maladies multifactorielles). On estime que le génome de chaque individu contient environ dix millions de SNPs. L'association de SNPs dans une région du génome située sur un même chromosome s'appelle un haplotype. Certains SNPs proches physiquement sur le génome ont tendance à être transmis en bloc (ensemble). On dit qu'ils sont en déséquilibre de liaison (qui peut être partiel ou total)

Les polymorphismes de répétition sont des séquences répétées en tandem de nombreuses fois, à partir de motifs de longueur variable (de quelques, à plusieurs centaines de paires de bases définissant en fonction de la taille les Microsatellites, Minisatellites, Satellites, et Mégasatellites).

Connaître les seuls SNPs ne permet d'identifier que quelques fréquences du génome. La connaissance des haplotypes permet une économie considérable, il suffit d'un SNP pour identifier tout un groupe de SNPs et de gènes (*Amorilic et al., 2008*)

- Les CNVs (Copy Number Variations): Il s'agit de variation du nombre d'exemplaires contigus de grands segments génomiques (perte ou gain de fragments de quelques kb à plusieurs Mb). A ce jour, des CNVs ont été identifiés dans environ 15% du génome humain. Les CNVs sont référencés dans la base de données Database for Genomic Variants : (<http://projects.tcag.ca/variation/>) ; avec plus de 65000 CNVs différents rapportés à ce jour).

Bases moléculaires des mutations et bases moléculaires du mode de transmission des maladies génétiques sont représentés dans la (fig 14)

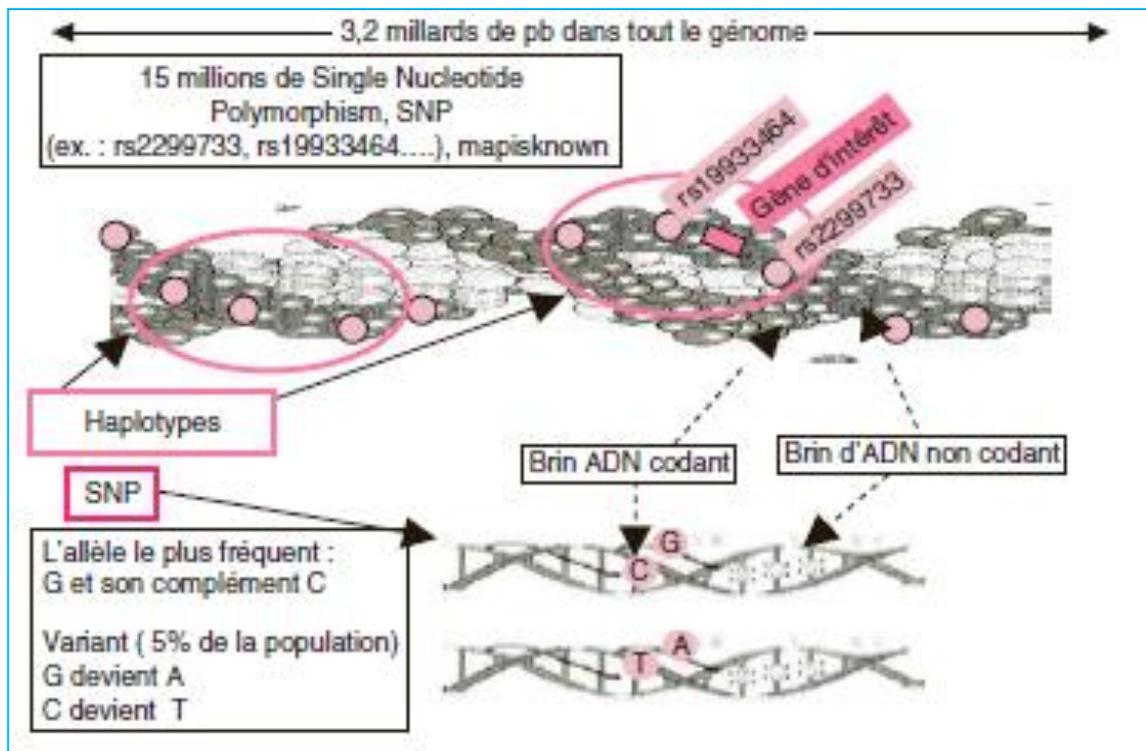


Figure 14 : SNPs et Haplotypes (Noll et al., 2003)

Le présent travail a été réalisé au département de physiologie et biologie cellulaire(BPC) de la faculté des sciences de la nature et de la vie (SNV) de l'université SAAD DAHLAB de BLIDA, pendant une période de 6 mois: du mois de Mai 2017 au mois de Octobre 2017

N'ayant pas de plateforme de séquençage à haut débit en Algérie, nous sommes focaliser dans notre travail à une étude *in silico* en utilisant les données des bases de données international et stockées et publiées dans plusieurs bases de données.

Rappelons que l'objectif de notre travail s'intéresse à l'étude de quelques gènes impliqués dans la défense antioxydante **CAT, SOD1, SOD2, GPx1** chez le diabétique de type1 et plus particulièrement à l'étude des interactions moléculaires de ces gènes antioxydant aux gènes de l'insuline (protéine hypoglycémiant) afin de prédire leurs rôle et fonctions métabolique et de comprendre la relation qui peut exister entre ces gènes et le diabète type1

II.1. Matériel

II.1.1. Matériel non biologique

Rappelons que notre étude est une étude *in Silico*, qui nécessite l'utilisation du matériel informatique (ordinateur portable) et d'une connexion internet à haut débit.

II.1.1.1.Outils bio-informatiques

Les progrès de la biologie moléculaire et du grand projet du génome humain ont élargi nos outils pour étudier les mécanismes pathogéniques des maladies humaines et la prédiction de certaines fonctions biologiques. Ainsi, nous avons utilisé des outils bio-informatiques (description en annexes) pour le traitement de nos échantillons moléculaire et qui sont des bases de données à accès libre (Open Source) nécessitant des interface web, des logiciels, des softwares et des algorithmes et qui sont répertorié dans le **tableau I**.

Tableau I: Description des outils bioinformatiques

Base de données	Commentaires	URL
NCBI	C'est une base de données qui intègre des informations spécifiques aux gènes à partir de sources de données multiples.	https://www.ncbi.nlm.nih.gov
BLAST	Le logiciel blast permet de comparer une séquence, nucléique ou protéique, dite requête, à une banque de séquences, nucléiques ou protéiques. Il existe essentiellement cinq types de comparaison possibles : Blast n ,Blast p, blast x ,tblast n, tbalstx	https://blast.ncbi.nlm.nih.gov
UNIPROT	UniProt est une base de données de séquences de protéines, C'est une base de données ouverte, stable et accessible en ligne	http://www.uniprot.org/

II.1.2. Matériel biologique

Pour répondre à la problématique de notre étude *in Silico*, nous avons utilisé des échantillons stockés dans des bases de données à accès libre (Open Source) comme NCBI. Ces échantillons ont été obtenus après une étude moléculaire qui avait nécessité une expérimentation à partir desquelles des molécules biologiques ont été obtenus puis traités par des plateformes à l'échelle internationale utilisant la technologie de séquençage à haut débit comme les NGS (Next Generation Sequencing). En effet, cette dernière, nous permet d'obtenir des séquences d'ADN ou de transcriptome et ceci pour répondre à plusieurs problématiques ou thématiques dans le souci de comprendre les mécanismes d'interactions moléculaires, les voies cellulaires et biologiques associées à des pathologies qui dans notre cas est le diabète type 1, ou aux diversités des espèces vivantes ou à des phénomènes biologiques. (fig 15, annexe)

Après une sélection des molécules biologiques que nous avons utilisé pour répondre à notre Problématique, nous avons retenu les molécules qui sont répertoriés dans le tableau suivant :

Tableau II: Description des molécules biologiques

Nature des séquences	Numéro d'accèsion	ID : identifiant	Nature de expérience	Nature du fichier	Format
ADN Catalase	NC_000011	ID: 847	Polymorphisme	SNP	Fasta
ADN SOD1	NC_000021	ID:6647	Polymorphisme	SNP	Fasta
ADN SOD2	NC_000006	ID: 6648	Polymorphisme	SNP	Fasta
ADN GPx1	NC_000003	ID:2876	Polymorphisme	SNP	Fasta

II.2. Méthode d'étude

L'organigramme de notre étude (figure16)

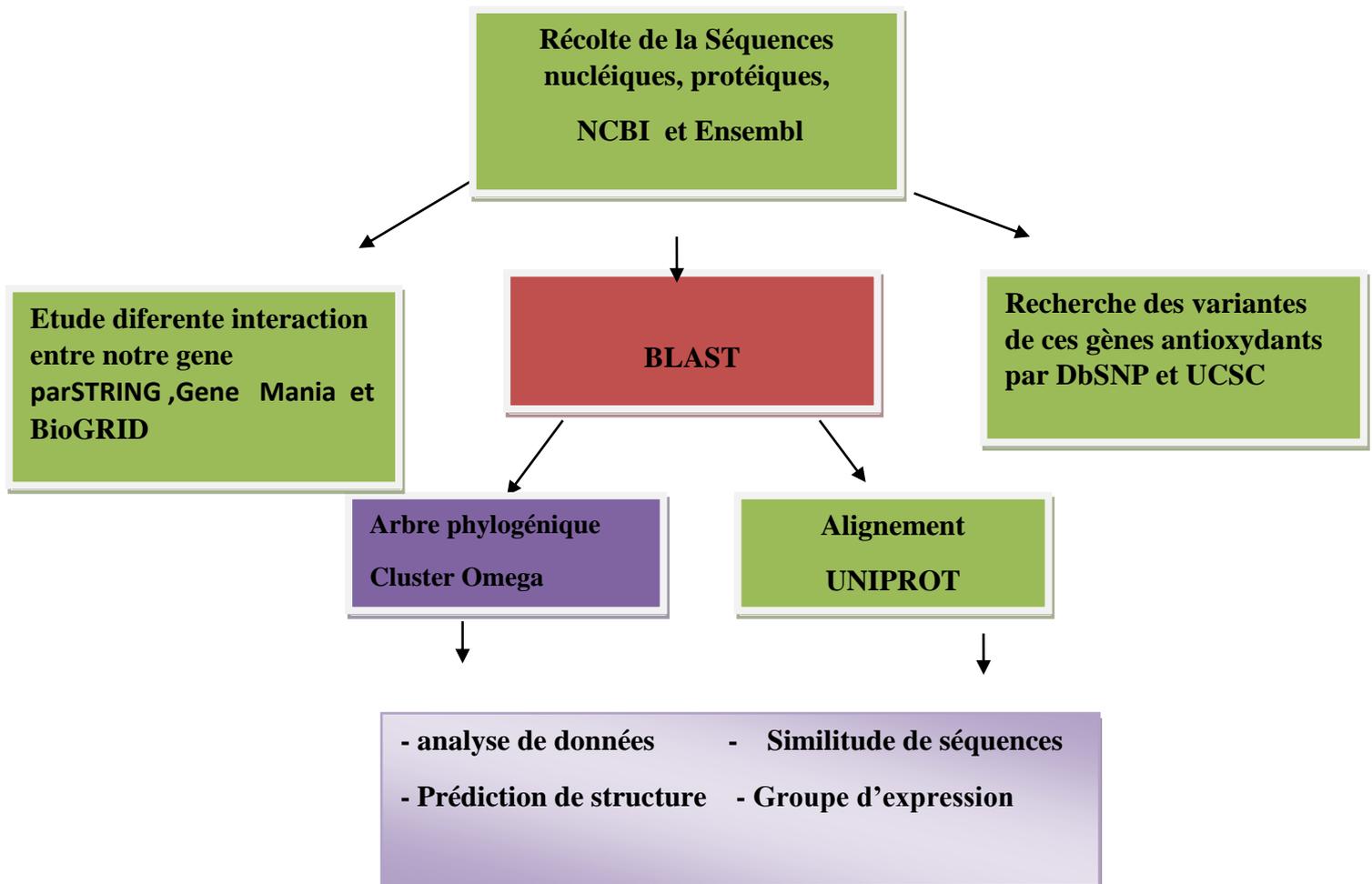


Figure16: protocole de notre étude.

II.2.1. Récolte des données moléculaires

- Séquences nucléiques, protéiques des gènes antioxydants CAT, SOD1, SOD2, GPx1 dans la banque de données NCBI (voir annexe).

II.2.2. Traitement des données moléculaires

Pour l'étude d'interaction entre les protéines a été réalisé par l'outil suivant :

II.2.3 : Les ORFs des nos séquences antioxydants

➤ Protocole

- Ouvrir le site SMS (Séquence Manipulation Suite) version 2

<http://www.bioinformatics.org/sms2/index.html>

- ensuite allez vers Analysis et choisir ORF Finder, puis sur sur Clear

- Puis on copie la séquence d'ADN sous sa forme fasta et on la colle dans l'espace réservé aux séquences

- Puis on va choisir dans l'ORFs canbeginwith : any codon

- puis l'onglet Recherche for ORFs in reading frame puis choisir 1,2 and 3

- Et ensuite choisir en premier l'onglet direct puis reverse

- puis choisir " Only return ORFs that are at least" puis choisir le nombre 30 codons longs .

- et choisir le code génétique standard (1)

- Et puis on clique sur Submit

- Les résultats d'ORF vont apparaitre puis on refait le travail dans l'autre sens

- On va choisir ensuite la séquence protéique la plus longue

II .2.4. Outil Blast :

➤ Protocole :

- Sur le site BLAST <https://blast.ncbi.nlm.nih.gov/Blast.cgi>

- Cliqué sur nucléotides blast

- Copier la séquence sous sa forme fasta

- Coller la séquence dans l'espace de blastx

- Et cliqué sur BLAST en bas à gauche de la page

- Attendre quelques secondes

- Il va apparaitre un tableau avec les séquences les plus proches à la nôtre par classement décroissant classé selon le pourcentage de ressemblance et l'E-value

- Tout à fait à droite il y'a une colonne d'accèsion qui contient des détails sur la séquence la plus proche de la nôtre.

II.2.5: l'alignement et la comparaison des séquences nucléotidique et peptidique

Après la récupération des différents détails de notre gène on passe à l'étape d'alignement et à la comparaison de notre gène. Ainsi pour effectuer cette étape on utilise les outils suivants:

II.2.5. 1.UNIPROT**➤ Le protocole :**

Pour comparer les séquences peptidiques de nos gènes enzymes antioxydantes avec l'insuline

-On cherche nos enzymes par le nom dans ce programme et on va les mettre dans le champ baskets et cela pour les aligner

On clique sur le champ baskets et on coche les gènes que nous voulons aligner

-Puis cliquer sur le bouton aligne et on aura l'alignement et la proposition de l'arbre phylogène.

Tableau III : Les identifications des enzymes par UNIPROT

P04040	CATA_HUMAN	Catalase	CAT
P07203	GPX1_HUMAN	Glutathionne peroxydase 1	GPX1
P01308	INS_HUMAN	Insuline	INS
P04179	SODM_HUMAN	Superoxide dismutase [Mn], mitochon	SOD2
P00441	SODC_HUMAN	Superoxide dismutase[Cu- Zn]	SOD1

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

II.2.6. Etude des Interactions entre les gènes des enzymes antioxydants

II.2.6. 1. Outil STRING

- **Principe** : vise à simplifier l'accès à des informations en fournissant une collection complète, mais contrôlée par la qualité, d'associations protéines-protéines pour un grand nombre d'organismes

- **Méthode**
 - Rendez-vous au programme string génome annotation <http://string.embl.de/>.
 - Cliquez sur le bouton multiple protéine
 - Insérez les symboles de notre gène SOD1, SOD2, CAT, GPx1.
 - Choisissez dans le champ des organismes (homo sapiens)
 - Cliquez sur rechercher .
 - Il va apparaître un schéma qui explique les interactions entre les enzymes antioxydants
 - Tout en bas il y a une colonne accession qui contient des détails sur les interactions

II.2.6. 2. Outil Gene MANIA

- **Principe** : On utilise ce programme pour la recherche des différentes interactions entre nos gènes, les interactions physiques, co-expression, co-localisation

- **Méthode** :

Rendez-vous au programme **GeneMANIA** (<http://www.genemania.org>)

- Insérez les symboles de nos gènes SOD1, SOD2, GPX1, CAT dans le champ de recherche
- Puis choisissez dans le champ les paramètres des interactions que nous voulons pour notre travail
- Puis lancez la recherche. On trouve alors les différentes interactions entre nos gènes et avec d'autres gènes

II.2.7. La recherche des variantes de ces gènes antioxydants

Dans cet étape on va explorer les différents polymorphisme des gènes antioxydants et les identifier au niveau des : exons, le promoteurs ,3'UTR,5'UTR , position , sur le gène. Identifier les variations des fréquences alléliques et génotypiques, l'allèle ancestrale ,les fréquence génotypique et phénotypique ,qui peuvent avoir une relation avec le diabète type 1 et on va les organiser sous forme de, tableaux .

- Pour cette étape on utilise La base de données des données polymorphique nucléotidique (dbSNP)

- **Principe :** le programme **Ensembl** est utiliser pour la recherche des variantes, également les références snip ou « rs » pour chaque variant SNP.
- **Méthode :**
 - Rendez-vous à la page <http://www.ensembl.org> puis choisir l'espèce humaine dans le champ des espèces
 - Puis on insère le symbole du gène et on lance la recherche
 - Ensuite on clique sur le champ gène pour préciser le gène.

III.1 .Les résultats

Les résultats concernant l'analyse des interactions les séquences protéiques du stress CAT, SOD1, SOD2, GPx1 et insuline par les outils bioinformatique se présentent comme suit :

III.1. 1. Etudes des similitudes par l'outil Blast

L'annotation génomique par l'outil blast pour les 5 séquences : INS, CAT- GPx1, SOD1, SOD2 débutant par la récolte de la séquence FASTA, puis Blast X , nous a donné un champ vaste des résultats bruts. Les résultats sont répertoriés dans le **tableau IV**.

Tableau VI : Résultats du blast X de la glutathion peroxydase

Nom de la séquence	Identification	E-value	Score totale	Numéro d'accesion	Longueur	Matches
glutathioneperoxidase 1 isoform 5 [Homo sapiens]	100%	3e-79	249	NP_001316384.1	157	1
glutathioneperoxidase 1 isoform 1 [Homo sapiens]	100%	1e-78	378	NP_000572.2	203	2
glutathioneperoxidase [Homo sapiens]	100%	1e-78	377	CAB37833.1	201	2
GPX1 protein [Homo sapiens]	99%	3e-76	299	AAH70258.1	145	2

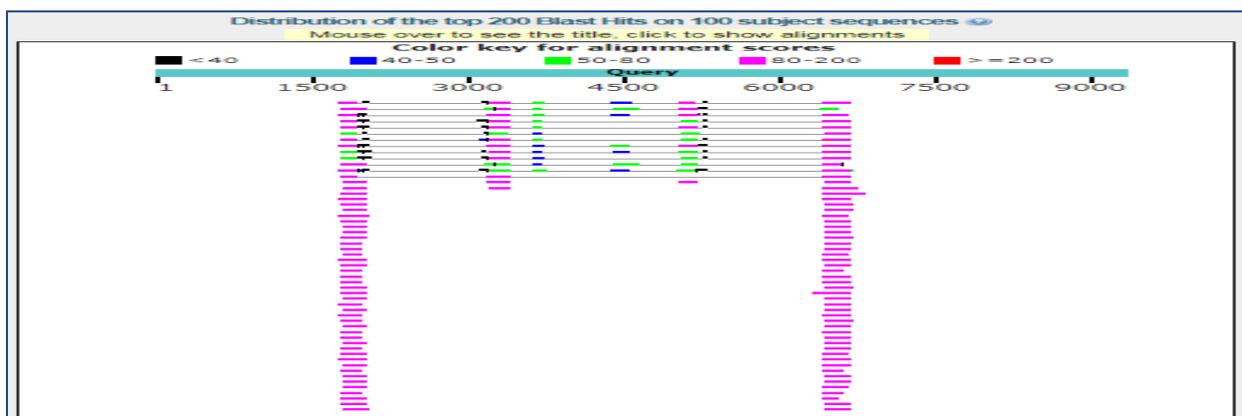


Figure 18: les résultats d'alignement pour la séquence nucléotidique du glutathion peroxydase

Tableau V : Résultats du blast X de l'insuline

Nom de la séquence	Identification	E-value	Score totale	Numéro d'accèsion	Longueur	Matches
Insulin [Homo sapiens]	98%	3e-21	182	AEG19452.1	107	2
Insulin [Homo sapiens]	88%	2e-21	177	AAA59179.1	94	2
Insulin preproprotein [Homo sapiens]	98%	2e-21	182	NP_000198.1	110	2

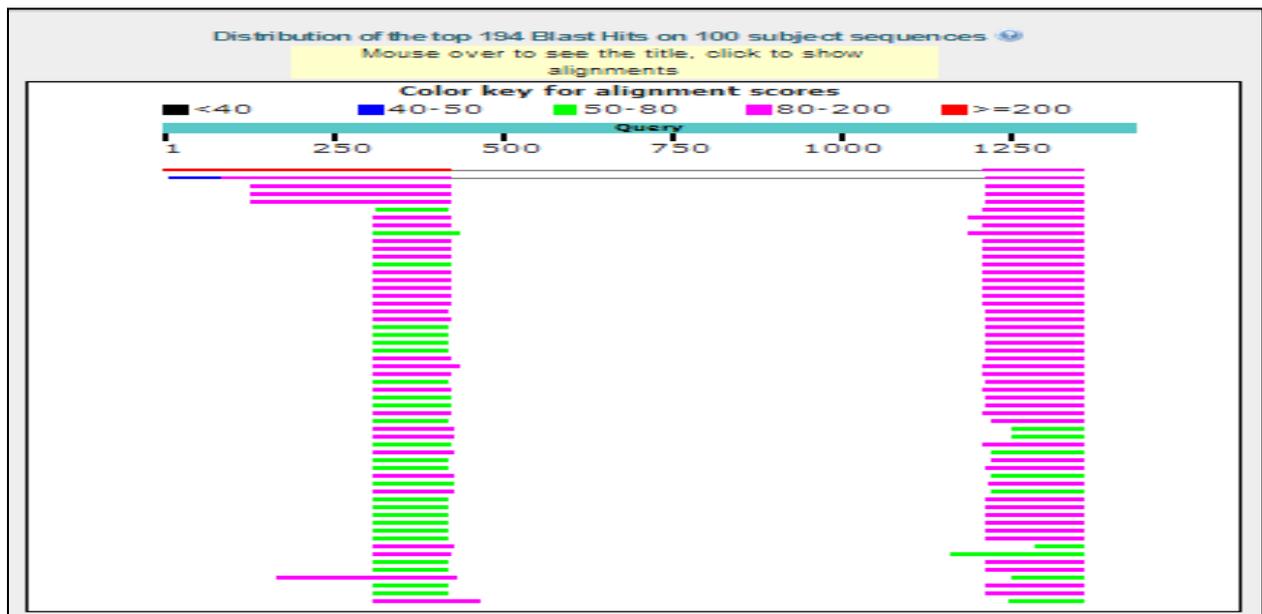
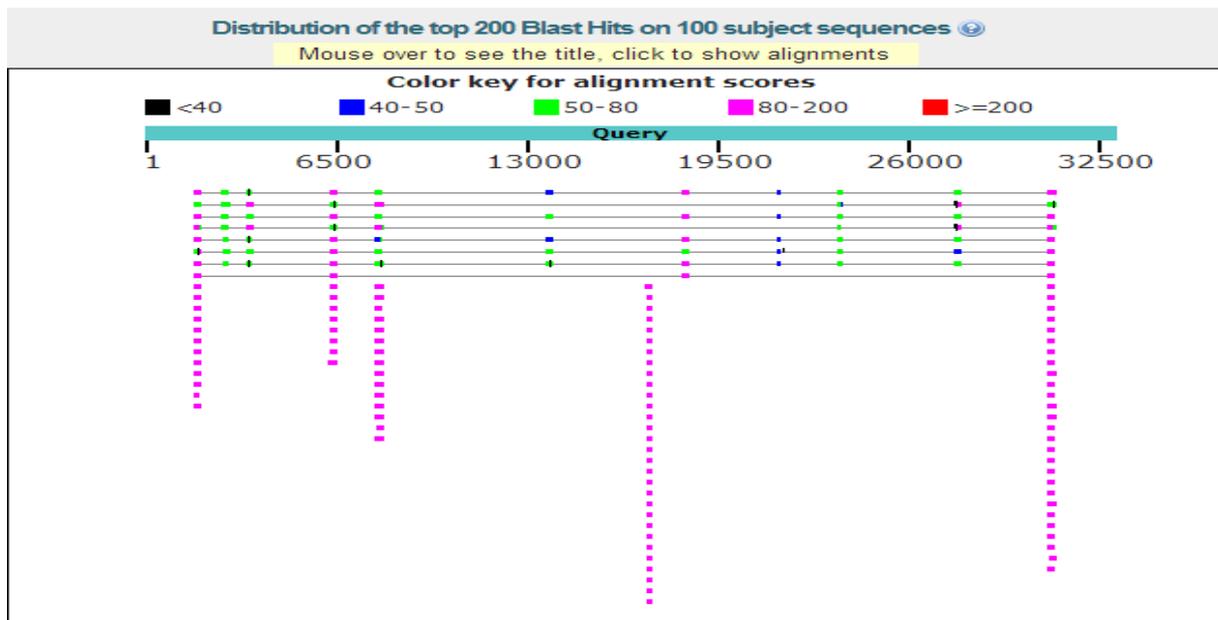
**Figure 19** : les résultats d'alignement pour la séquence nucléotidique de l'insuline

Tableau VI: Résultats du blast X de la catalase

Nom de séquence	Identifiant	E- valu	Scor totale	Numéro d'cession	longueur	Matches
catalase [Homo sapiens]	100%	3e-32	993	AAB59522.1	451	11
catalase [Homo sapiens]	100%	1e-29	993	AAB59522.1	451	11
catalase [Homo sapiens]	100%	3e-29	1169	NP_001743.1	527	13
catalase [Homo sapiens]	100%	9e-32	1169	NP_001743.1	527	13

**Figure20:** les résultats d'alignement pour la séquence nucléotidique de Catalase

III.1.2. Alignement des séquences par UNIPROT

Les séquences peptidiques des enzymes antioxydantes ont été obtenue suite à l’alignement de séquences. Grace aux outils d’alignement se trouvant dans l’interface UNIPROT nous avons obtenu les résultats représentés par la Figure 18.

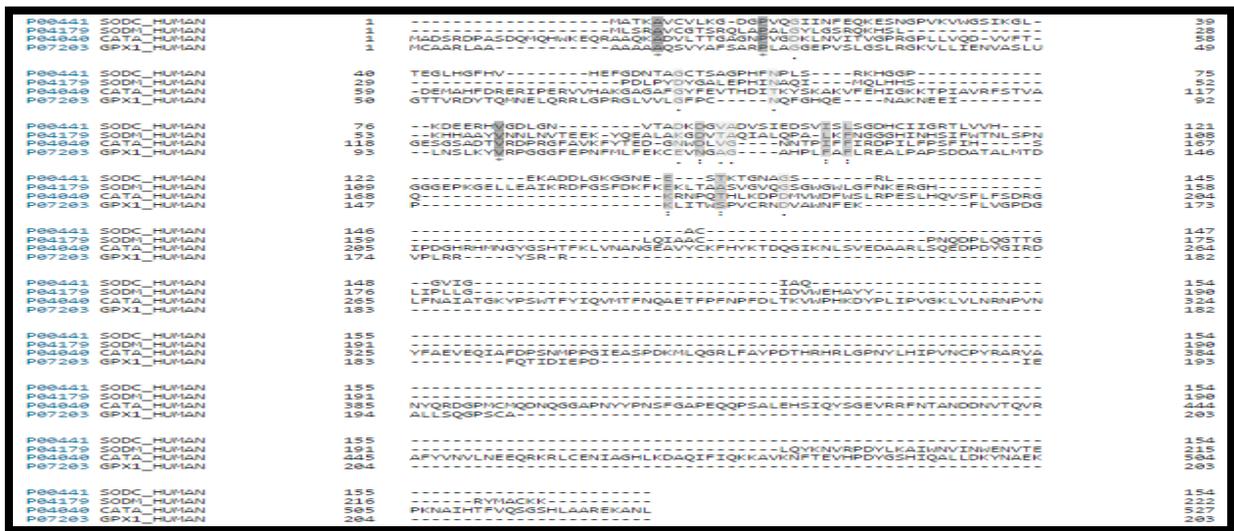


Figure 21 : Alignement des quatre séquences peptidiques (SOD1,SOD2,CAT, GPX1)

Nous constatons que l’alignements représente un faible pourcentage de similarité, avec 3 positions identique et 12 position similaires entre elles

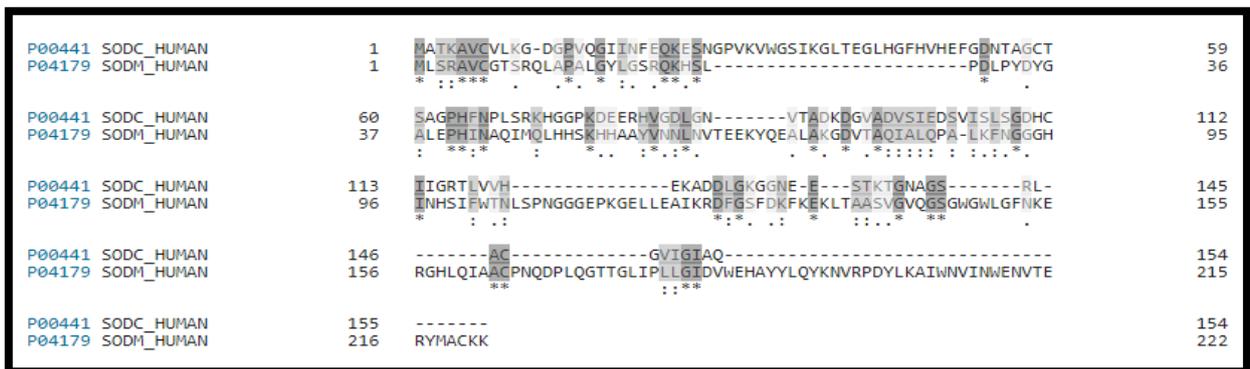


Figure 22 : Alignement des séquences peptidiques de la SOD1 et SOD2 par UNIPROT (<http://www.uniprot.org/>)

Les résultats de cet l’alignement montrent un pourcentage de similarité de 12.551% avec 47 positions similaires et 31 positions identiques entre les deux enzymes SOD1 et SOD2.

Tableau VII : Résultats de la totalité de l'alignement trouvé par UNIPROT entre les différentes séquences

L'alignement Des gènes des séquences peptidiques	Identité	Position de similarité	position d'identité	Dure de réalisation
Cat/gpx1/sod1/sod2/ins	0%	0	0	31.7 s
Cat/gpx1/sod1/sod2	0.533%	12	3	20s
CAT/GPX1	7.86%	65	42	25.6s
SOD1/SOD2	12.551%	47	31	13.9s
CAT/INS	6.78%	25	16	12.9s

L'alignement multiple grâce à l'outil CLUSTER Omega montre un arbre phylogénique de nos 4 gènes visualisé dans la (figure 24)

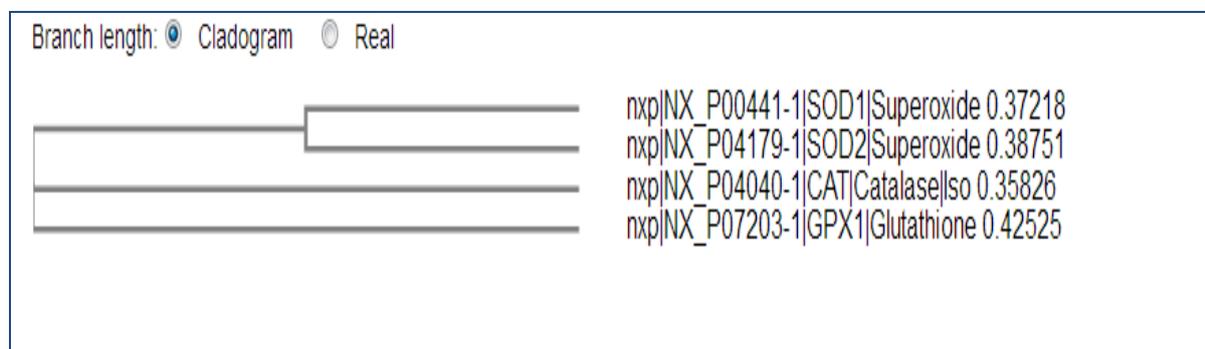


Figure 24:Arbre phylogénique des quatre enzymes antioxydants par CLUSTER Omega

Cette figure 24 montre la relation phylogénique entre nos quatre gènes. On peut remarquer facilement l'approchement qui se trouve entre les gènes SOD1 et SOD2. Ce rapprochement suggère qu'il Ya un ancêtre commun entre ces deux gènes.

Au niveau des gènes GPx1 et CAT, on remarque qu'il ya une certaine distance entre ces deux gènes et l'arbre ne montre pas de relation ancestrale entre ces deux gènes ce qui montre encore une fois la grande différence qui se trouve entre GPx1 et la CAT

Cette Arbre phylogénique confirme les résultats trouvés dans notre étude en montrant, d'une part, la grande ressemblance entre les deux gènes SOD1 et SOD2 et les différences significatives entre les deux gènes GPXI et CAT, d'une autre part.

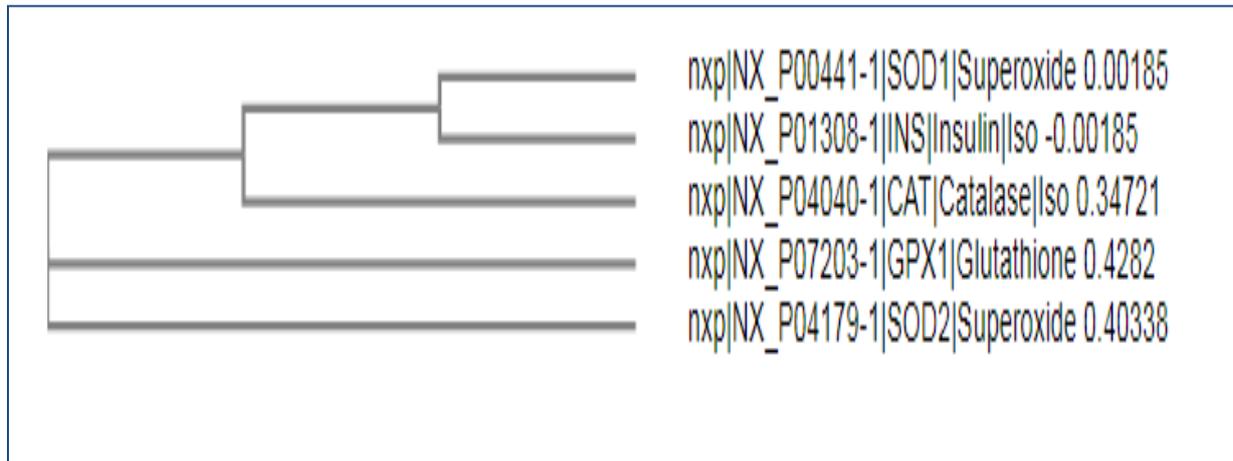


Figure25 : Arbre phylogénétique des quatre enzymes du stress avec l'insuline avec CLUSTER Omega

Après la comparaison des quatre gènes deux par deux (SOD1 et SOD2 puis GPXI et CAT) on a essayé de faire une comparaison multiple des quatre gènes pour voir la relation entre les quatre gènes en utilisant un programme de comparaison multiple **CLUSTER Omega**.

Ce programme permet de comparer les quatre gènes ensemble. Nous constatons dans cette étape, qu'il n'y a pas beaucoup de régions communes entre les quatre gènes, ce qui montre qu'il n'y a pas de relation importante entre les quatre gènes au niveau de leurs structures nucléotidiques.

Tableau VIII: Cadres de lecture de gène :

Gène	Sens	N Cadre de lecture	N ORF	La taille	début	fin
CAT	revers	3	21	521	32613	33134.
GPX1	directe	2	2	518	8	526.
SOD1	revers	1	8	692	8581	9273
SOD2	revere	3	23	1097	27162	28259.
INS	directe	9	1	551	2	553.

III.1.3. Interactions entre les enzymes antioxydant et le gène d'insuline

Pour visualiser les interactions entre les différentes molécules de notre étude, nous avons utilisé l'outil (STRING: functional protein association networks) qui signifie une chaîne de molécules

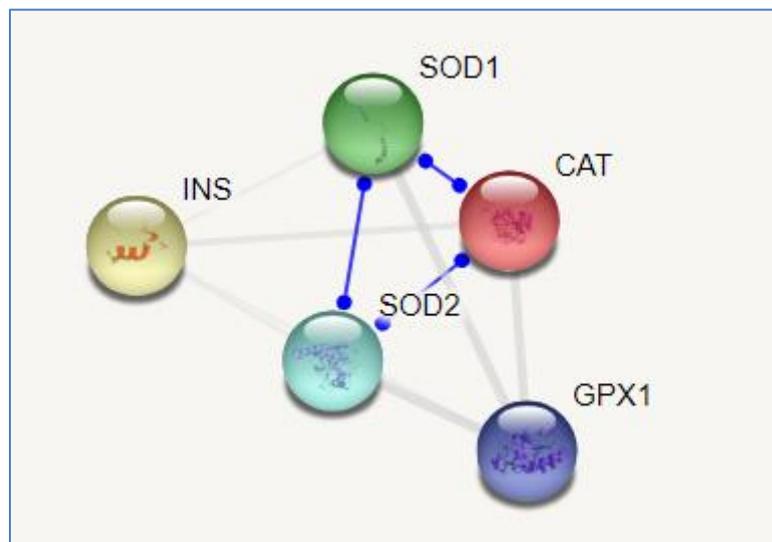


Figure 26: L'interaction pour les gènes des enzymes antioxydants avec l'insuline

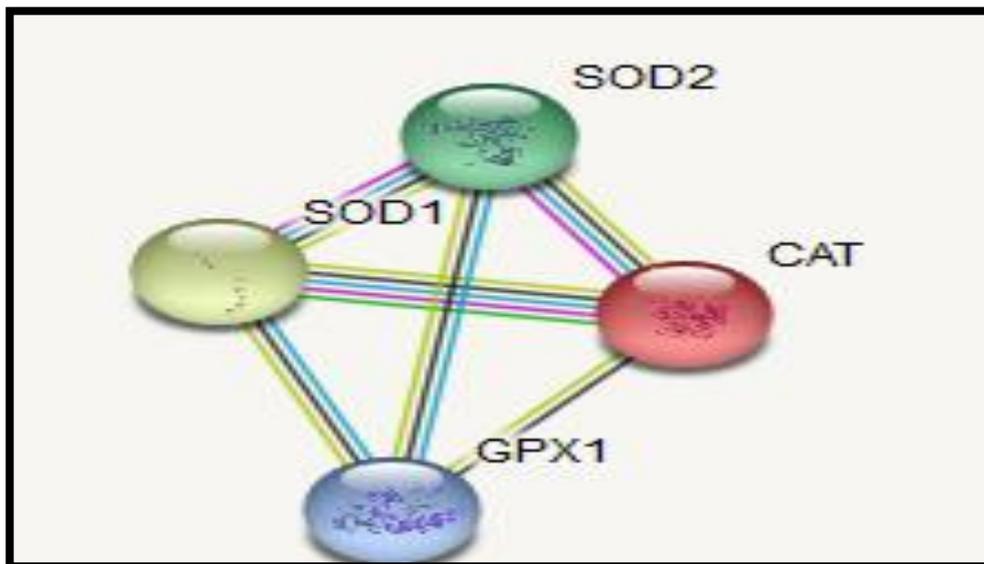
(<https://string-db.org/cgi/network.pl>)

les processus biologiques par l'enrichissement fonctionnel associé à ces interactions sont regroupés dans le tableau ci-dessous.

Tableau IX : Enrichissement fonctionnel

Biological Process (GO)			
<i>pathway ID</i>	<i>pathway description</i>	<i>count in gene set</i>	<i>false discovery rate</i>
GO:1902175	regulation of oxidative stress-induced intrinsic apoptotic signaling pathway	4	1.41e-07
GO:1902176	negative regulation of oxidative stress-induced intrinsic apoptotic signaling pathway	3	2.2e-05
GO:0042311	vasodilation	3	3.94e-05
GO:0042743	hydrogen peroxide metabolic process	3	3.94e-05
GO:0043066	negative regulation of apoptotic process	5	8.72e-05
GO:0007568	aging	4	9.17e-05
GO:1902531	regulation of intracellular signal transduction	5	0.000549
GO:0072593	reactive oxygen species metabolic process	3	0.000874
GO:0009650	UV protection	2	0.00161
GO:0019430	removal of superoxide radicals	2	0.00161
GO:0009636	response to toxic substance	3	0.0022
GO:0071451	cellular response to superoxide	2	0.0023
GO:0000302	response to reactive oxygen species	3	0.00244
GO:0043523	regulation of neuron apoptotic process	3	0.00295

Les interactions entre les différentes séquences protéiques des quatre enzymatique est illustré par la figure 27.

**Figure27**: représentation des réseaux d'interaction entre les protéines antioxydants

CAT, GXP1, SOD1, SOD2

(<https://string-db.org/cgi/network.pl>)

La probabilité d'interaction est évaluée par un enrichissement ($p\text{-value} = 1.21e^{-05}$.)

Nos résultats montrent qu'il n'existe pas de Co-expression entre la SOD2 et la CAT alors les homologues putatifs sont co-exprimés dans d'autres espèces avec un score (0.056).

De même, nous constatons qu'il n'existe pas d'interaction entre la CAT et le GPX1. Nous suggérons ainsi, qu'il n'existe pas de d'interaction de fusion et pas de Cooccurrence entre les gènes. En revanche, nous constatons qu'il existe des homologues putatifs qui sont coexprimés dans d'autres espèces avec un score de (0.160)

Pour amplifier notre réseau d'interaction physique, les colorisations, et les voies de Co-expression entre les gènes du stress oxydant et le gène d'insuline, nous avons utilisé l'outil Gene Mania qui nous a permis de visualiser les interactions représentées dans la figure 28.

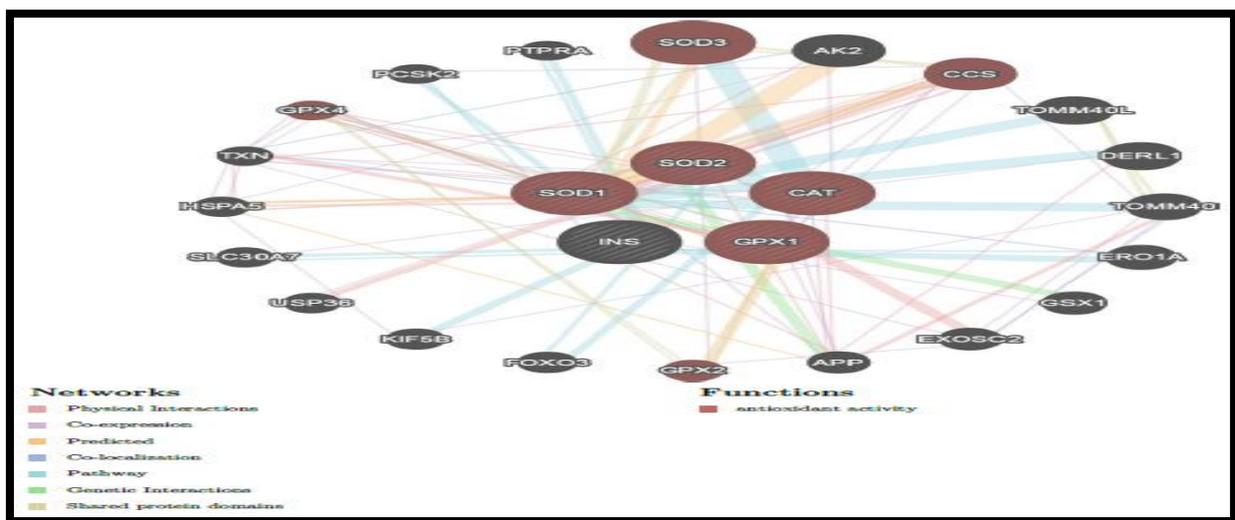


Figure28: Les résultats d'interactions par l'outil Gene Mania

(<http://genemania.org/>)

Le troisième outil utilisé pour la visualisation des interactions est l'outil Biogrid : (<https://thebiogrid.org/>). (Figure 29)

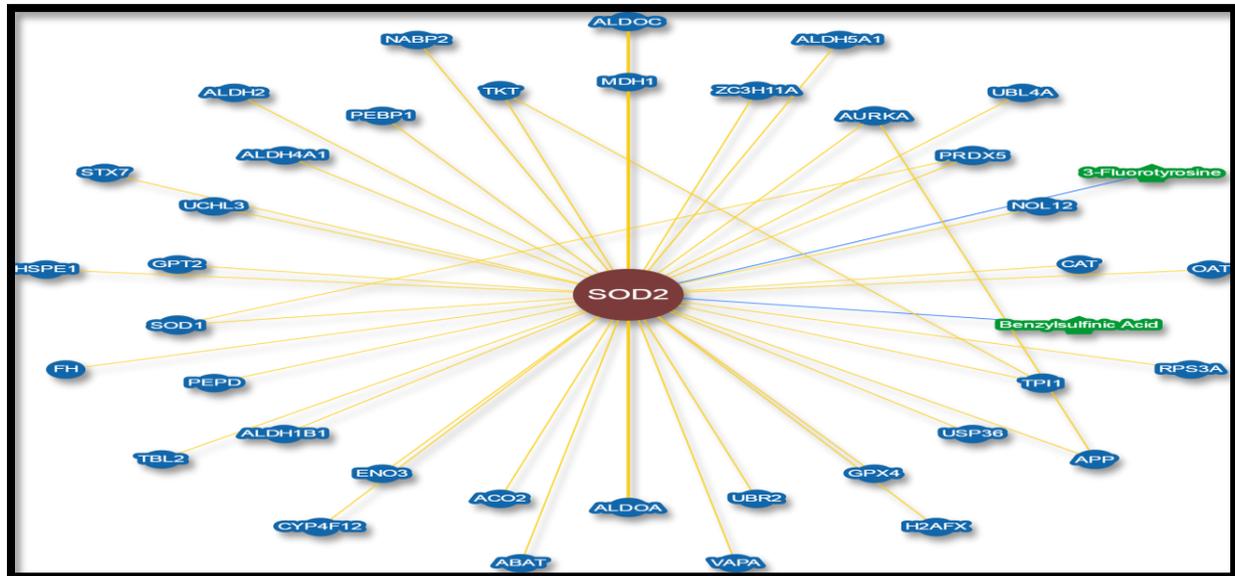


Figure 29: Réseaux d'interaction entre les gènes antioxydants par BioGrid (<https://thebiogrid.org/>)

La figure 29 représente un réseau d'interaction de 43 bordures, 40 nœuds d'interaction, 41 bordures physiques et 2 nœuds chimiques, cependant il n'existe pas d'interaction physique et chimique annotée.

Selon les résultats générés par l'outil BioGrid, nous constatons qu'il n'existe pas de co-expression entre la GPX1 et la SOD1. En revanche, nos résultats montrent qu'il existe des homologues putatifs coexprimés dans d'autres espèces avec un score de 0,098.

Aussi, nous observons aucune co-expression entre la SOD1 et SOD2. Cependant, les homologues putatifs sont coexprimés dans d'autres espèces avec un score de 0.056.

Nous constatons que l'utilisation de l'outil BIOGRID donne un réseau avec beaucoup plus d'interactions. Un tel enrichissement indique que les protéines sont au moins partiellement liées biologiquement, en tant que groupe.

III.1.4. l'outil UCSC

Pour rechercher et visualiser les variants et les exons de la SOD1, nous avons utilisé l'outil UCSC (The University of California Santa Cruz) Genome Browser Database.

Les différents résultats des enzymes antioxydantes et l'insuline sont représentés par la (figure 30)

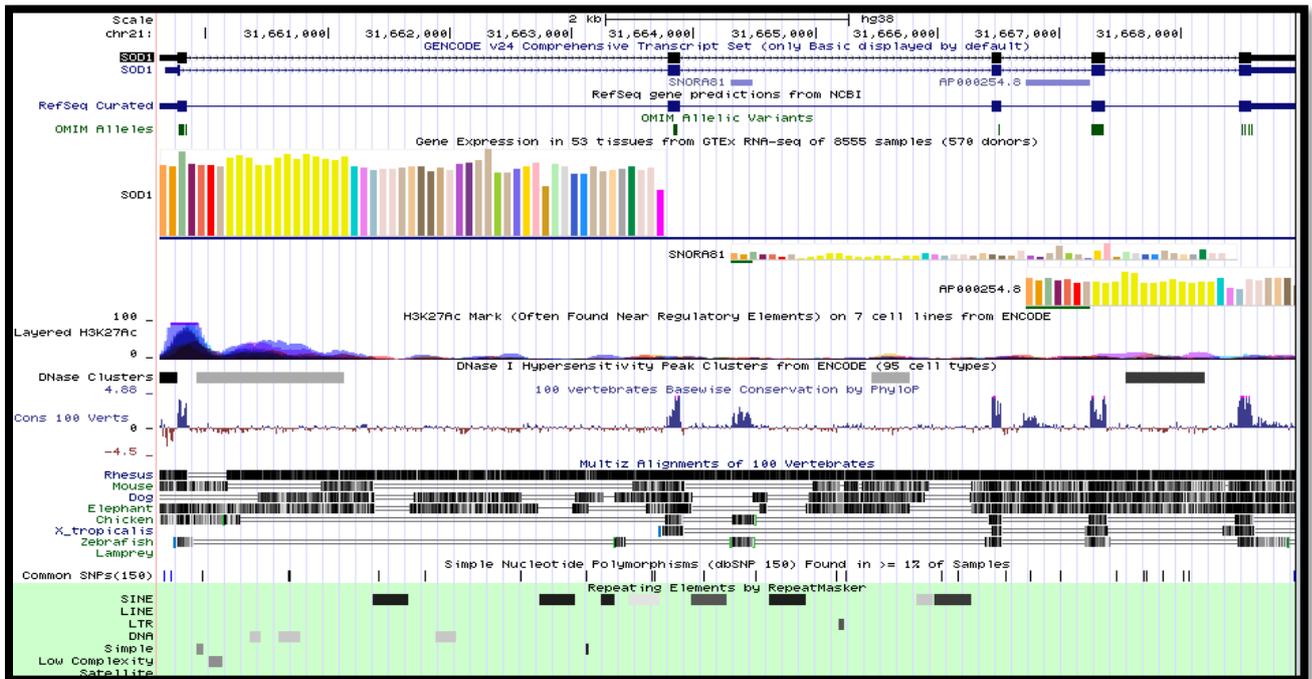


Figure 30: Représentation des variants et des exons de SOD1 par UCSC

(<http://genome.ucsc.edu/>)

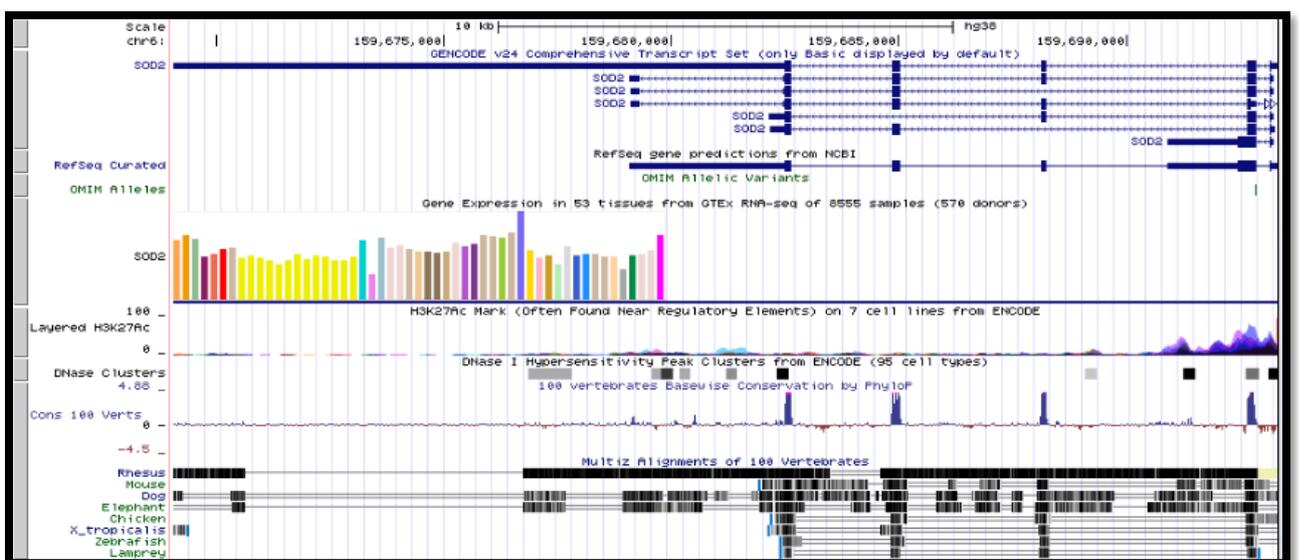


Figure31 : Représentation des variants et des exons de SOD2 par UCSC

(<http://genome.ucsc.edu/>)

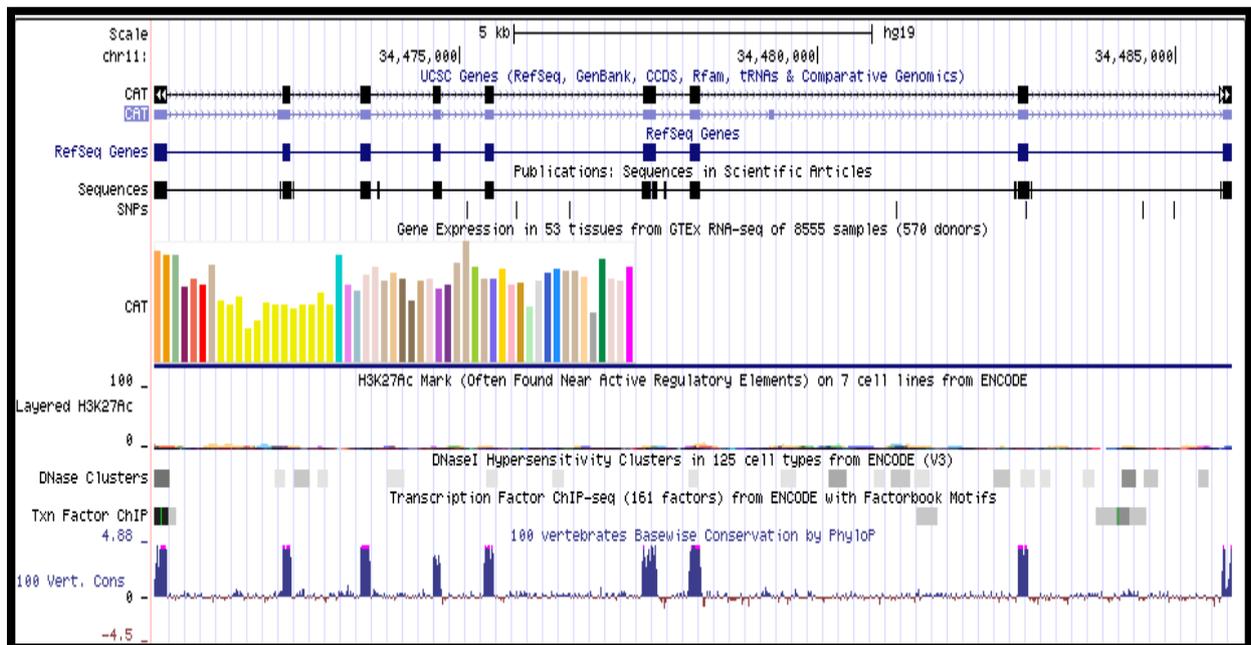


Figure 32: Représentation des variants et des exons de la CAT par UCSC

(<http://genome.ucsc.edu/>)

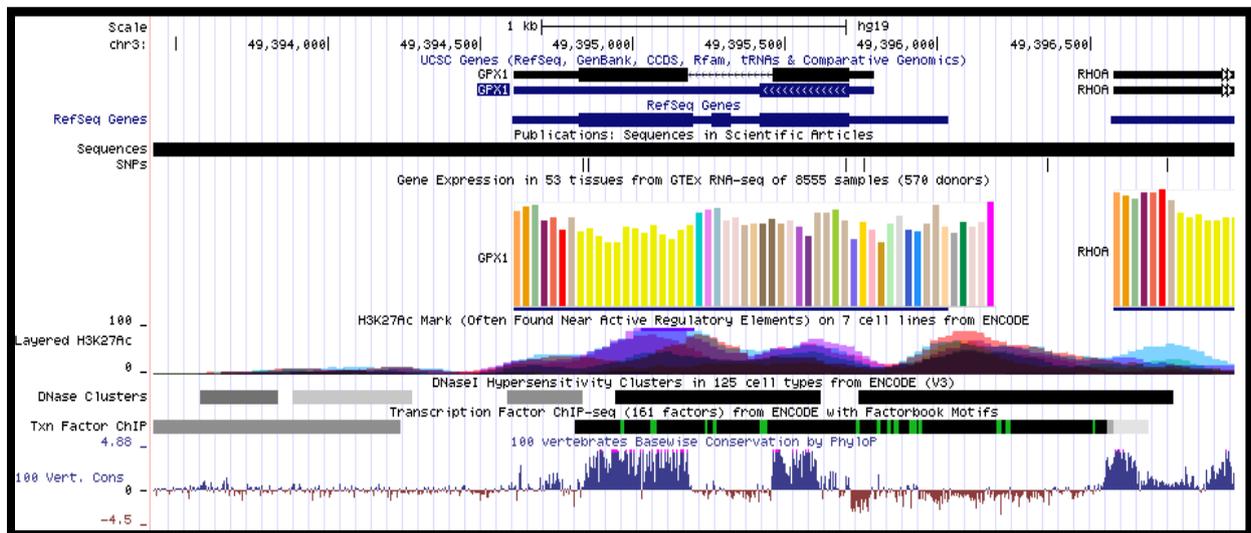


Figure33 : Représentation des variantes et des exons de GPX1 par UCSC

(<http://genome.ucsc.edu/>)

III.1.5. Recherche des mutations polymorphisme des gènes antioxydants

En utilisant les données des publications récolté dans les sites des données bibliographiques

« Pub Med » et « ENSEMBL » nous avons pu répertoriés les informations suivante :

➤ Cas du polymorphisme du gène de la Catalase

Le tableau ci-dessous représente deux polymorphisme majeurs retrouvés chez les diabétiques de type1. Les résultats de **Pasket al en 2006 montrent** en utilisant la régression logistique, jusqu'à 3530 cas et 3930 contrôles qui a pour but de faire ressortir les associations entre ces deux polymorphismes et le DT1.

Tableau X : les fréquences alléliques génotypiques du polymorphisme C1167T et C (-262) T de la CAT chez les DT1

		CAT	Cases N(%)		Control		OR (95%CI)		P
CAT C1167T	(rs769217)	Allèle T	1573	(22.6)	1719	22.3	1.00		-
		Allèle C	5391	(77.4)	5999	77.7	1.01	(0.93-1.01)	79
Génotype		TT	179	(5.1)	189	(4.9)	1.00		-
		TC	1215	(34.9)	1341	(34.7)	1.01	(0.85-1.27)	0.92
		CC	2088	(60.0)	2329	(60.4)	1.02	(0.82-1.28)	0.84
Cat C262T	(rs1001179)	T	1507	(21.8)	1694	(22.1)	1.00		-
		C	5395	(78.2)	5978	77.9	0.98	(0.90-1.06)	0.55
Génotype		TT	177	(5.1)	201	5.2	1.00		-
		TC	1153	(33.4)	1292	(33.7)	0.93(0.74-1.17)		0.55
		CC	2121	(61.5)	2343	(61.5)	0.92(0.74-1.15)		0.47

Ces résultats montrent qu'il n'existe pas une association entre le polymorphisme des gènes de la catalase C1167T et C (-262) T allèle et le DT1 pour les génotypes homozygotes et hétérozygotes

(<https://www.ensembl.org/>)

➤ **Cas du polymorphisme du gène du GPX1**

Les fréquences alléliques et génotypique du variant de la GPX1 trouver par l’outil Ensemble de déférente population est répertorié dans le tableau ci-dessous

Tableau XI : les fréquences alléliques génotypiques du polymorphisme de la GPX1 dans différentes pathologies

Population	Fréquence alléliques (% ,n)		Fréquence génotypique (% ,n)		
	G	A	GG	AA	GA
ALL	G: 0.799 (4001)	A: 0.201 (1007)	G G: 0.643 (1610)	A A: 0.045 (113)	A G: 0.312 (781)
AFR	G: 0.789 (1043)	A: 0.211 (279)	G G: 0.625 (413)	A A: 0.047 (31)	A G: 0.328 (217)
AMR	G: 0.829 (575)	A: 0.171 (119)	G G: 0.683 (237)	A A: 0.026 (9)	A G: 0.291 (101)
EAS	G: 0.928 (935)	A: 0.072 (73)	G G: 0.857 (432)	A A: 0.002 (1)	A G: 0.141 (71)
EUR	G: 0.664 (668)	A: 0.336 (338)	G G: 0.429 (216)	A A: 0.101 (51)	A G: 0.469 (236)
SAS	G: 0.798 (780)	A: 0.202 (198)	G G: 0.638 (312)	A A: 0.043 (21)	A G: 0.319 (156)

(<https://www.ensembl.org/>)

➤ **Cas du polymorphisme du gène de la SOD2**

Les fréquences alléliques et génotypique du variant de la SOD2 trouver par l'outil **Ensembl** de déférente population est répertoriée dans le tableau ci-dessous.

Tableau XII : les fréquences alléliques génotypiques majeur et mineur du polymorphisme de la SOD2 dans différentes pathologies chez cinq population différentes

SNP	Population	Fréquence allélique		Fréquence génotypique		
		Majeur	Mineur	Majeur	Mineur	
rs4880 5'UTR	AFR	A: 0.576	G: 0.424 (560)	A G:0.487 (322)	A A:0.333 (220)	G G:0.180 (119)
	AMR	G: 0.584 (405)	A: 0.416 (289)	A G:0.476 (165)	A A:0.179 (62)	G G:0.346 (120)
	EAS	A: 0.875 (882)	G: 0.125 (126)	A A:0.764 (385)	A G:0.222 (112)	G G:0.014 (7)
	EUR	A: 0.534 (537)	G: 0.466 (469)	A G:0.479 (241)	A A:0.294 (148)	G G:0.227 (114)
	SAS	G: 0.508 (497)	A: 0.492 (481)	A G:0.493 (241)	A A:0.245 (120)	G G:0.262 (128)

(<https://www.ensembl.org/>)

III.2. Discussion

Il faut rappeler que l'organisme humain a développé un système de défense contre les agents agressifs qui sont les enzymes antioxydantes. Par conséquent, le rôle des enzymes antioxydantes est crucial dans la régulation du métabolisme.

Rappelons aussi, que notre étude a été réalisée pour étudier les interactions entre les gènes des enzymes antioxydantes en utilisant l'outil bioinformatique. Le choix de nos molécules repose sur les enzymes antioxydantes impliquées directement dans la défense de plusieurs pathologies dont le diabète type 1 (DT1). Plusieurs études ont montré l'implication de ces gènes dans la défense antioxydantes

Concernant l'étude des interactions entre les enzymes antioxydantes, Seules les études phylogéniques ont été annotées au niveau des bases de données spécifiquement pour les espèces bactériennes.

Selon **Flekac et al en 2008**, les espèces réactives de l'oxygène (ROS) générées suite à une hyperglycémie modifient la structure et la fonction des lipides, protéines et d'autres molécules générant ainsi des changements vasculaires chez les diabétiques. De plus, une activité basse des enzymes antioxydantes a été observée chez plusieurs patients diabétiques. Ainsi, la présence d'un stress oxydant pourrait empêcher l'action du rôle protecteur de ces enzymes.

A travers l'étude du polymorphisme de la SOD et la CAT, Flekac et collaborateurs, confirment que le polymorphisme des enzymes antioxydantes est une cause possible dans l'inactivation de ces enzymes suite à une hyperglycémie chronique.

En effet, au cours de notre travail, nous avons observé grâce à l'utilisation des outils bioinformatiques et par les données générées par les big DATA, un panel de données concernant des études de polymorphisme de plusieurs populations qui pourrait par une étude statistique d'association donner des interprétations prometteuses de l'association de ces gènes aux pathologies.

Au fur et à mesure que la quantité de séquences génomiques dans les bases de données publiques augmente, l'annotation devient un élément essentiel dans l'analyse et l'interprétation des données (*Karolchik et al., 2003*)

Les méthodes sans alignement pour comparer les séquences de protéines se sont révélées être des alternatives viables avec des approches qui reposent spécifiquement sur un alignement des séquences à comparer. Cependant, elles partagent peut-être de similitudes (*Nojoomi et Koehl., 2017*)

ont proposé récemment une méthode sans alignement basée sur le concept des noyaux de chaîne (SeqKernel) qui sont des Méthodes basées sur l'alignement standard, ses applications sont potentiellement limitées, en raison des biais dues principalement aux longueurs de séquences. Ce n'est pas le cas pour notre étude puisque toutes les séquences étaient relativement courtes.

Dans notre travail, l'utilisation des trois outils de visualisation des interactions semble être utile pour le choix des gènes dans le but d'une annotation impliqué dans le niveau d'expression, la régulation post transcriptionnelle, les modifications post traductionnelle, la formation des complexes protéiques, l'activation ou l'inhibition d'un gène par rapport à un autre, aussi l'étude de leurs participations à des voies de signalisations.

Selon Kanehisa et al., chaque interaction importée ou prédite, STRING associe un score d'un jeu de données de référence composé d'association réel bien documentées et extraites à partir des groupements fonctionnels maintenus par KEGG (*Kanehisa et al., 2008*). En effet dans le cas de notre étude chaque paire de protéines, il existe un score d'interaction « combiné » reflétant le, ou les types d'interactions que l'on souhaite, ou non prendre en compte lors de la construction et la visualisation d'une partie d'interactions fourni par le site web de String.

CONCLUSION

Conclusion :

Le diabète est une maladie métabolique caractérisée par un niveau de sucre trop élevé dans le sang et l'altération de la production d'insuline, lié ou stress oxydantes produire de radicaux libres responsables de l'oxydation des cellules et de leurs composés et de la perturbation du système de défense antioxydant. les cellules ne contrôlent plus de l'existence des espèces oxygénées.

Notre étude est faite sur l'enzymes antioxydants le plus importante qui agissent directement sur les radicaux libres .notre objectif de travail est l'étude d'interaction entre les gènes antioxydant et le gène d'insuline dans les diabètes type 1 pour ce faire on utilise programme bioinformatique.

Premièrement l'annotation des 5 séquences nucléotidiques CAT ,GPX1 ,SOD1, SOD2,INS, par l'outil blast pour rechercher des séquences similaires à nos séquences .

Ensuite l'alignement des séquences par l'outil UniProt pour identifier les régions similaires et identiques entre la structure des séquences peptidiques des gènes CAT, GPX1,SOD1,SOD2,INS. les résultats de cet alignement montrent que il n'y a pas de ressemblance de la structure des séquences peptidiques entre les 5 séquences (0%) d'identification , on trouve un faible pourcentage de ressemblance de l'alignement des séquences peptidiques de CAT ,GPX 1 ,SOD1,SOD 2 ce qui montre qu'il n'existe pas de similarité structurale entre ces protéines et 12,5% d'identité entre SOD1 et 2 et 6,7 entre GPX1 et CAT ce qui nous permet de dire qu'il y a une certaine ressemblance .

Pour la recherche des interactions entre les enzymes antioxydantes et les gènes d'insuline on utilise trois outils string , BioGrid et GeneMania pour chercher les différentes interactions géniques , physiques , co expression .

On a terminé notre étude par la recherche des mutations polymorphiques des gènes antioxydants CAT ,GPX1,SOD1 SOD2 par l'utilisation des données des publications recueillies dans le site bibliographique Pub Med et Ensemble on a trouvé les différents polymorphismes majeurs qui sont associés avec le diabète type 1, Flekac et collaborateurs, confirment que le polymorphisme des enzymes antioxydantes est une cause possible dans l'inactivation de ces enzymes suite à une hyperglycémie chronique .

Plusieurs études ont été faites sur les polymorphismes des gènes antioxydants montrant qu'il existe des polymorphismes dans ces gènes associés au diabète type 1.

Comme on a indiqué en plusieurs reprises, durant notre étude on a utilisé des outils purement bio-informatiques, seront très intéressantes pour mieux comprendre cette relation qu'on a déterminé entre les gènes antioxydants et le diabète type 1 .

Références bibliographique

-**Baalbaki L**, (2012). Les traitements innovants du diabète de type 1 : Focus sur la greffe des îlots de Langerhans (son historique, son optimisation et ses défis réglementaires) ; thèse de doctorat en pharmacie ; Université Joseph Fourier : p 24

- **Betteridge DJ**¹What is oxidative stress? *Metabolism*. 2000 p3-8

-**Bonnefont-Rousselot D**¹.The role of antioxidant micronutrients in the prevention of diabetic complications. *Treat Endocrinol*. 2004;3(1):41-52.

-**Cavarape, A., F. Feletto, F. Mercuri, L. Quagliaro, G. Daman, A. Ceriello A.** (2001) High-fructose diet decreases catalase mRNA levels in rat tissues. *J Endocrinol Invest* 24:838–845.

-**Chada S, Le Beau MM, Casey L, Newburger PE** (Feb 1990). "Isolation and chromosomal localization of the human glutathione peroxidase gene". *Genomics*. 6 (2): 268–71

- **Chistiakov1 D. A. Savost'anov K. V. Turakulov R. I. Titovich Zilberman E. V. L. I. . Kuraeva., T. L Dedov., I. I. Nosikov V. V.** A new type 1 diabetes susceptibility locus containing the catalase gene (chromosome 11p13) in a Russian population. *Diabetes Metab Res Rev* 2004; 20: 219–224.

-**Collins AR, Horvathova E.** Oxidative DNA damage, antioxidants and DNA repair: applications of the comet assay. *Biochem Soc Trans* 2001; 29: 337–341

-**Dedov., I. I. Nosikov V. V.** A new type 1 diabetes susceptibility locus containing the catalase gene (chromosome 11p13) in a Russian population. *Diabetes Metab Res Rev* 2004; 20: 219–224.

-**Deisseroth, A., D ounce, A. L.** (1970) Catalase: physical and chemical properties, mechanism of catalysis, and physiological role. *Physiol. Rev.* 50, 319-375.

-**D. Levanon, J. Lieman-Hurwitz, N. Dafni et al.**, “Architecture and anatomy of the chromosomal locus in human chromosome 21 encoding the Cu/Zn superoxide dismutase,” *The EMBO Journal*, vol. 4, no. 1, pp. 77–84, 1985.

-**Eisenbarth GS, Gianani R, Pugliese A, Verge CF, Pietropaolo M.** Prediction and prevention of type I diabetes. *Transplant Proc.* 1994 Apr;26(2):361–362

- **E. O'Dea and A. Hoffmann**, “The regulatory logic of the NF-kappaB signaling system.,” *Cold Spring Harbor Perspectives in Biology*, vol. 2, no. 1, Article ID a000216, 2010.

-**Estácio SG, Leal SS, Cristóvão JS, Faisca PF, Gomes CM** (February 2015). "Calcium binding to gatekeeper residues flanking aggregation-prone segments underlies non-fibrillar amyloid traits in superoxide dismutase 1 (SOD1)". *Biochimica et Biophysica Acta*. 1854 (2): 118–26. *PMID 25463043*. *doi:10.1016/j.bbapap.2014.11.005*

- **Favier.,A.** Intérêt conceptuel et expérimental dans la compréhension des mécanismes des maladies et potentiel thérapeutique .l'actualité chimique . 2003. 108-115

-**Forsberg, L., Lyrenäs, L., de Faire, U., Morgenstern, R.** (2001) A common functional C-T substitution polymorphism in the promoter region of the human catalase gene influences transcription factor binding, reporter gene transcription and is correlated to blood catalase levels. *Free Radic. Biol.Med.* 30, 500-505.

-**Forbes, J.M., Coughlan, M.T., Cooper, M.E.** Oxidative stress as a major culprit in kidney disease in diabetes. *Diabetes*. 2008;57:1446–1454.

--**German M, Ashcroft S, Docherty K, Edlund H, Edlund T, Goodison S,** et al. The insulin gene promoter, a simplified nomenclature. *Diabetes* 1995;44:1002–4.

-**Góth, L., Rass, P., Páy, A.** (2004) Catalase enzyme mutations and their association with diseases. *Mol. Diagn.* 8, 141-149

-**Grimaldi.A.** Diabétologie Questions d'internat .Université Pierre et Marie Curie1999 – 2000

-**Grimaldi A** How to help the patient motivate himself? *Diabetes Metab.*2012

-**Giugliano D¹, Ceriello A, Paolisso G.**Oxidative stress and diabetic vascular complications. *Diabetes Care*. 1996 . p 257-67.

-**Guillausseau p.-j.** classification et critères diagnostiquésdu diabète : propositions de l'ada et de l'oms *diabetes metabolism vol 23, n° 5novemb1997p. 454*

-**Idelman S, Verdetti J.** (2000). Le pancreas endocrine. *Endocrinologie et communications cellulaires, Les Ulis: EDP science, 277-327*

- **Johansen JS, Harris AK, Rychly DJ, Ergul A:** Oxidative stress and the use of antioxidants in diabetes: linking basic science to clinical practice. *Cardiovasc Diabetol.* 2005, 4: 5-10.1186/1475-2840-4-5.

-**Kataoka K¹, Suzuki Y, Kitada M, Hashimoto T, Chou H, Bai H, Ohta M, Wu S, Suzuki K, Ide C.** Alginate enhances elongation of early regenerating axons in spinal cord of young rats. *Tissue Eng.* 2004 10(3-4):493-504.

-Kukko M, Kimpimaki T, Kupila A, Korhonen S, Kulmala P, Savola K, Simell T, Keskinen P, Ilonen J, Simell O, Knip M. 2003. Signs of beta-cell autoimmunity and hla defined diabetes susceptibility in the finnish population: The sib cohort from the type 1 diabetes prediction and prevention study. *Diabetologia*, 46:65-70.

-Levanon D, Liemann-Hurwitz J, Dafni N, Wigderson M, Sherman L, Bernstein Y, Laver-Rudich Z, Danciger E, Stein O, Groner Y (1985) Architecture and anatomy of the chromosomal locus in human chromosome 21 encoding the Cu/Zn superoxide peroxidase. *EMBO J* 4:77–84

-Liu X, Xu W, Pan Y, Du E.. Liu . suspect . (2015) may have underestimated dissolved organic nitrogen (N) but overestimated total particulate N in wet deposition in China.

- Magnan .C., Ktorza. , A. Production et sécrétion de l'insuline par la cellule b pancréatique Production and secretion of insulin by the pancreatic b-cell. Elsevier SAS. 2005 .241–264

-Melloul D, Marshak S, Cerasi E. Regulation of insulin gene transcription. *Diabetologia* 2002;45:309–26.

-Miller,,Eugene W. Myers et David Lipman (1990), Basic Local Alignment Search Tool, *Journal of Molecular Biology*, 215, 3, 403410.

- Mohammedi A. Patente. ; Bellili-Muñoz N. , Driss F ;Hervé Le Nagard ;Frédéric Fumeron., Ronan Roussel ., Hadjadj S .,Lúcia Corrêa-GiannellaM ., Marre M ., Velho G.(2015)Glutathione peroxidase-1 gene (GPX1) variants, oxidative stress and risk of kidney complications in people with type 1 diabetes.

-Mohammedi K, Maimaitiming S, Emery N, Bellili-Munoz N, Roussel R, Fumeron F, Hadjadj S, Marre M, Velho G. Allelic variations in superoxide dismutase-1 (SOD1) gene are associated with increased risk of diabetic nephropathy in type 1 diabetic subjects. *Mol. Genet. Metab.* 2011;104:654–660.

- **Noll H.** The digital origin of human language-a synthesis. *BioEssay* 2003, 25, 489-500

-Ohneda K., Ee H., German M. Regulation of insulin gene transcription. *Semin. Cell Dev. Biol.* 2000;11:227–233

-Poitout V., Hagman D., Stein R., Artner I., Robertson R.P., Harmon J.S. Regulation of the insulin gene by glucose and fatty acids. *J. Nutr.* 2006;136:873–876 15 Steiner DF, Chan SJ, Welsh JM, Kwok SC. Structure and evolution of the insulin gene. *Annu Rev Genet* 1985;19: 463–84.

-Rodier M. Définition et classification du diabète. *Médecine Nucléaire - Imagerie fonctionnelle et métabolique* - 2001 - vol.25 - n°2.p 91-93

-Schroeder, L.C., Boggs, D.H., Dome, G., Halberstam, I.M., Jones, W.L., Pierson, W.J. and Wentz, F.J. (1982). The relationship between wind vector and normalized radar cross section used to derive SEASAT-A satellite scatterometer winds

-Yoo HY, et al. (1999) Overexpression and simple purification of human superoxide dismutase (SOD1) in yeast and its resistance to oxidative stress. *J Biotechnol* 68(1):29-35

-Zhang C, Moriguchi T , Kajihara.M, Esaki.R, Harada A, Shimohata H, Oishi H, Hamada.M, Morito.N,¹ Hasegawa.K,¹ Kudo.T,¹ Engel.J,² Yamamoto.M, Takahashi.S. MafA Is a Key Regulator of Glucose-Stimulated Insulin Secretion. *Mol Cell Biol.* 2005 Jun; 25(12): 4969–4976.

Biblio –Web

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

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

STRING: <https://string-db.org/>

Gene Mania: <http://genemania.org/>

Ensembl genome browser 90 : <https://www.ensembl.org/>

DbSNP : <http://www.ncbi.nlm.nih.gov/SNP>

UNIPROT: <http://www.uniprot.org>

Cluster Omega: <http://www.ebi.ac.uk/Tools/msa/clustalo/>

-<https://www.ncbi.nlm.nih.gov/gene/847>

-<https://thebiogrid.org/>

-<http://genome.ucsc.edu/>

Annexes

Annexes1

Les séquences nucléotidiques :

La séquence nucléotidique du gène Gpx1 :

>NC_000003.12:c49358358-49357176 Homo sapiens chromosome 3, GRCh38.p7 Primary Assembly

CAGTTAAAAGGAGGCGCCTGCTGGCCTCCCCTTACAGTGCTTGTTCGGGGCGCTCCGCTGGCTTCTTGGA
CAATTGCGCCATGTGTGCTGCTCGGCTAGCGGCGGCGGCGGCGGCGGCCAGTCGGTGTATGCCTTCTCG
GCGCGCCCGCTGGCCGGCGGGGAGCCTGTGAGCCTGGGCTCCCTGCGGGGCAAGGTACTACTTATCGAGA
ATGTGGCGTCCCTCTGAGGCACCACGGTCCGGGACTACACCCAGATGAACGAGCTGCAGCGGCGCCTCG
ACCCCGGGGCTGGTGGTGTCTCGGCTTCCCGTGCAACCAGTTTGGGCATCAGGTGCGCCGGGCGGAGCGG
GGCGGGGCGGGGGCGGACGTGCAGTAGTGGCTGGGGGCGCCGGCGGTGTGCTGGTGGGTGCCGTCGGCTC
CATGCGCGGAGAGTCTGGCTACTCTCTCGTTTCTTTCTGTTGCTCGTAGCTGCTGAAATTCCTCTCCGC
CCTGGGATTGCGCATGGAGGGCAAATCCCGTGACTCATAGAAAATCTCCCTTGTTTGTGGTTAGAAC
GTTTCTCTCCTCCTCTTGACCCCGGTTCTAGCTGCCCTTCTCTCCTGTAGGAGAACGCCAAGAACGAAG
AGATTCTGAATTCCTCAAGTACGTCCGGCCTGGTGGTGGGTTTCGAGCCCAACTTCATGCTCTTCGAGAA
GTGCGAGGTGAACGGTGCGGGGGCGCACCTCTTTCGCCTTCTGCGGGAGGCCCTGCCAGCTCCCAGC
GACGACGCCACCGCCTTATGACCGACCCCAAGCTCATCACCTGGTCTCCGGTGTGTCGCAACGATGTTG
CCTGGAACTTTGAGAAGTTCCTGGTGGGCCCTGACGGTGTGCCCTACGCAGGTACAGCCGCCGCTTCCA
GACCATTGACATCGAGCCTGACATCGAAGCCCTGCTGTCTCAAGGGCCCAGCTGTGCCTAGGGCGCCCT
CCTACCCCGGCTGCTTGGCAGTTGCAGTGCTGCTGTCTCGGGGGGTTTTTCATCTATGAGGGTGTTCCT
CTAAACCTACGAGGGAGGAACACCTGATCTTACAGAAAATACCACCTCGAGATGGGTGCTGGTCCTGTTG
ATCCCAGTCTCTGCCAGACCAAGGCGAGTTTCCCCTAATAAAGTGCCGGGTGTCAGCAGAA

Annexes

La sequence nucléotidique du gene Insulin

>NC_000011.10:c2161209-2159779 Homo sapiens chromosome 11, GRCh38.p7 Primary Assembly

AGCCCTCCAGGACAGGCTGCATCAGAAGAGGCCATCAAGCAGGTCTGTTCCAAGGGCCTTTGCGTCAGGT
GGGCTCAGGATTCCAGGGTGGCTGGACCCCAGGCCCCAGCTCTGCAGCAGGGAGGACGTGGCTGGGCTCG
TGAAGCATGTGGGGGTGAGCCCAGGGGCCCAAGGCAGGGCACCTGGCCTTACGCCTGCCTCAGCCCTGC
CTGTCTCCAGATCACTGTCTTCTGCCATGGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGCTG
GCCCTCTGGGGACCTGACCCAGCCGCAGCCTTTGTGAACCAACACCTGTGCGGCTCACACCTGGTGGAAG
CTCTCTACCTAGTGTGCGGGGAACGAGGCTTCTTCTACACACCCAAGACCCGCCGGGAGGCAGAGGACCT
GCAGGGTGAGCCAACTGCCATTGCTGCCCTGGCCGCCCCAGCCACCCCTGCTCCTGGCGCTCCAC
CCAGCATGGGCAGAAGGGGGCAGGAGGCTGCCACCCAGCAGGGGGTCAGGTGCACTTTTTTAAAAAGAAG
TTCTTTGGTCACGTCCTAAAAGTGACCAGCTCCCTGTGGCCAGTCAGAATCTCAGCCTGAGGACGGTG
TTGGCTTCGGCAGCCCCGAGATACATCAGAGGGTGGGCACGCTCCTCCCTCCACTCGCCCCTAAACAAA
TGCCCCGAGCCATTCTCCACCCTCATTTGATGACCGCAGATTCAAGTGTTTTGTAAAGTAAAGTCTT
GGGTGACCTGGGGTCACAGGGTGCCCCACGCTGCCTGCCTCTGGGCGAACACCCCATCACGCCCGGAGGA
GGGCGTGGCTGCCTGCCTGAGTGGGCCAGACCCTGTGCGCCAGGCCTCACGGCAGCTCCATAGTCAGGAG
ATGGGGAAGATGCTGGGGACAGGCCCTGGGGAGAAGTACTGGGATCACCTGTTTCAGGCTCCCACTGTGAC
GCTGCCCCGGGGCGGGGGAAGGAGGTGGGACATGTGGGCGTTGGGGCCTGTAGGTCCACACCCAGTGTGG
GTGACCCCTCCCTCTAACCTGGGTCCAGCCCGCTGGAGATGGGTGGGAGTGCGACCTAGGGCTGGCGGGC
AGGGGGCACTGTGTCTCCCTGACTGTGTCTCCTGTGTCCCTCTGCCTCGCCGCTGTTCCGGAACCTGC
TCTGCGCGGCACGTCCTGGCAGTGGGGCAGGTGGAGCTGGGCGGGGGCCCTGGTGCAGGCAGCCTGCAGC
CCTTGGCCCTGGAGGGGTCCCTGCAGAAGCGTGGCATTGTGGAACAATGCTGTACCAGCATCTGCTCCCT
CTACCAGCTGGAGAATACTGCAACTAGACGCAGCCCGCAGGCAGCCCCACACCCGCCGCTCCTGCACC
GAGAGAGATGGAATAAAGCCCTTGAACCAGC

SOD1: >NC_000021.9:31659622-31668931 Homo sapiens chromosome 21, GRCh38.p7
Primary Assembly

SOD2: >NC_000006.12:c159762529-159679064 Homo sapiens chromosome 6, GRCh38.p7
Primary Assembly

Cat: NC_000011.10:34438925-34472060 Homo sapiens chromosome 11, GRCh38.p7
Primary Assembly

Annexes

Les séquences peptidiques des quatre enzymes antioxydants SOD 1, SOD2, GPx1 , Cat et INS

>nxp|NX_P040401|CAT|Catalase|Iso1

MADSRDPASDQMQHWKEQRAAQKADVLTTGAGNPVGDKLVITVGPRGPLLVQDVVFTDEMAHFDR
ERIPERVVHAKGAGAFGYFEVTHDITKYSKAKVFEHIGKKTPIAVRFSTVAGESGSADTVRDPRGFAVK
FYTEDGNWDLVGNNTPIFFIRDPILFPSFIHSQKRNPQTHLKDPMVWDFWSLRPESLHQVSFLFSDRGIP
DGHRHMNGYGSHTFKLVNANGEAVYCKFHYKTDQGIKNLSVEDAARLSQEDPDYGIRDLFNAIATGK
YPSWTFYIQVMTFNQAETFPFNPFDLTKVWPHKDYPLIPVGKLVNLRNPVNYFAEVEQIAFDPSNMPPGI
EASPDKMLQGRLFAYPDTHRHLGPNYLHIPVNCYPYRARVANYQRDGPMMQDNQGGAPNYYPNSF
GAPEQQPSALEHSIQYSGEVRRFNTANDDNVTQVRAFYVNVLNNEEQKRKRLCENIAGHLKDAQIFIQKKA
VKNFTEVHPDYGSHIQALLDKYNAEKPKNAIHTFVQSGSHLAAREKANL

>nxp|NX_P00441-1|SOD1|Superoxide dismutase [Cu-Zn]|Iso 1

MATKAVCVLKGDPVQGIINFEQKESNGPVKVVWGSIKGLTEGLHGFHVHEFGDNTAGCTSAGPHFNPL
SRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSDHCCIIGRTLVVHEKADDLGKGGNEEST
KTGNAGSRLACGVIGIAQ

>nxp|NX_P04179-1|SOD2|Superoxide dismutase [Mn], mitochondrial|Iso 1

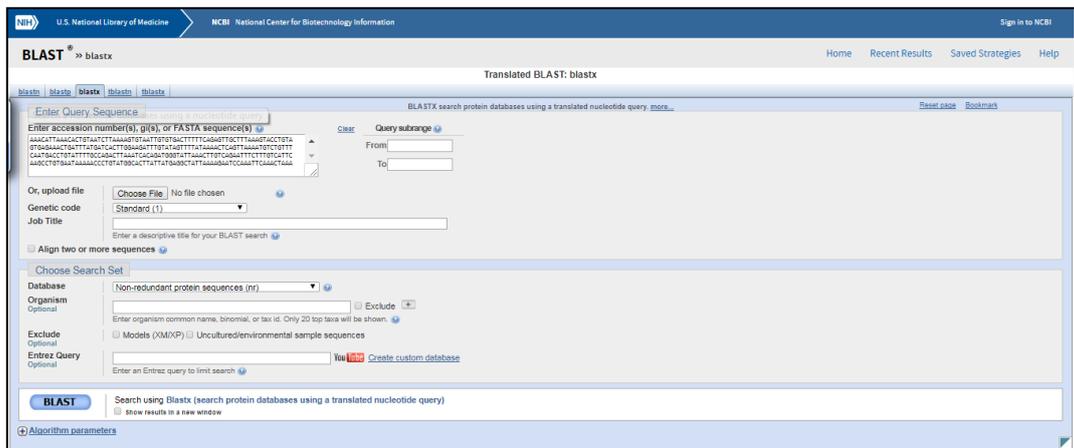
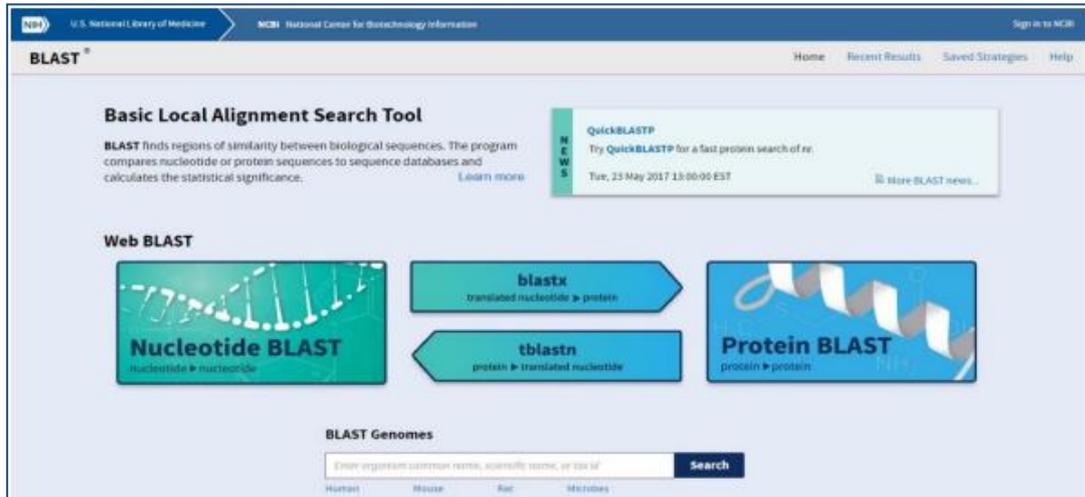
MLSRAVCGTSRQLAPALGYLGSRQKHSPLDLPYDYGALPHINAQIMQLHHSKHAAAYVNNLNVTEE
KYQEALAKGDVTAQIALQPALKFNGGGHINHSIFWTNLSPNGGEPKGELEAIKRDGFSDFKFEKLT
AASVGVQGSWGWLGFNKERGHLQIAACPNQDPLQGTGLIPLLIDVWEHAYYLQYKNVRPDYLKA
IWNVINWENVTERYMACKK

>nxp|NX_P07203-1|GPX1|Glutathione peroxidase 1|Iso 1

MCAARLAAAAAAQSVYAFSARPLAGGEPVSLGSLRGKVLLIENVASLUGTTVRDYTQMNELQRRLG
PRGLVVLGFPCNQFGHQENAKNEEILNSLKYVRPGGGFEPNFMLFEKCEVNGAGAHPLFAFLREALPAP
SDDATALMTDPKLITWSPVCRNDVAWNFEKFLVGPDGVPLRRYSRRFQTIDIEPDIEALLSQGPSCA

Annexes

Figure représenté l'outil Blast T:



Annexes

UNIPROT

The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

[UniProtKB](#) (UniProt Knowledgebase)
 Swiss-Prot (554,241) Manually annotated and reviewed.
 TrEMBL (84,827,567) Automatically annotated and not reviewed.

[UniRef](#) Sequence clusters
 UniParc Sequence archive
 Proteomes FHL

Supporting data
 Literature citations: 1
 Taxonomy: h
 Subcellular locations: 0
 Cross-ref. databases: 4
 Diseases: d
 Keywords: 0

Getting started
 1 Text search
 Our basic text search allows you to search all the resources available.

UniProt data
 Download latest release
 Get the UniProt data
 Statistics
 View Swiss-Prot and TrEMBL statistics

News
 Forthcoming changes
 Planned changes for UniProt
 UniProt release 2017_04
 Death (by insulin) in paradise
 UniProt release 2017_03
 Viral Short Message Service: peptide texting guides the outcome of infection
 UniProt release 2017_02
 News archive

Protein spotlight
 Between You And Me
 April 2017
 Communication has a purpose and is usually selfish. Humans have raised it to the level of entertainment in the form of...

UniProtKB results

Filter by: Reviewed (682) Swiss-Prot, Unreviewed (36,852) TrEMBL

Popular organisms: A. thaliana (28), Mouse (15), Rice (12), Human (9), E. coli K12 (7), Other organisms

Search terms: Filter "catalase" as: disease (1)

Entry	Entry name	Protein names	Gene	UniProtKB (5)	UniRef (0)	UniParc (0)	Remove
P04040	CATA_HUMAN	Catalase	CAT	<input checked="" type="checkbox"/>			<input type="checkbox"/>
P00441	SODC_HUMAN			<input checked="" type="checkbox"/>			<input type="checkbox"/>
P04179	SODM_HUMAN			<input checked="" type="checkbox"/>			<input type="checkbox"/>
P08294	SODE_HUMAN			<input checked="" type="checkbox"/>			<input type="checkbox"/>
P07203	GPX1_HUMAN			<input checked="" type="checkbox"/>			<input type="checkbox"/>
P24270	CATA_MOUSE	Catalase	Cat C				
P04762	CATA_RAT	Catalase	Cat C				
P00432	CATA_BOVIN	Catalase	CAT				
P13029	KATG_ECOLI	Catalase-peroxidase	katG b3942, JW3914				
Q42547	CATA3_ARATH	Catalase-3	CAT3 At1g20620, F2D10.40, FSM15.5				
F4JM86	F4JM86_ARATH	Catalase	CAT2 CATALASE, catalase 2, At4g35090, T12J5.2				
D5ER10	CATA2_ADATH	Catalase-2	CAT2 CAT, At4g35090, MAE13.140				

Alignment

How to print an alignment in color

Entry	Entry name	Protein names	Gene	Position	Sequence	Position
P04040	CATA_HUMAN	Catalase	CAT	1	-----MADSRDPASDQNHQKKEQRAAQKADLVTTGAGHPVSGDKLNVIITVGRPRGLVQD-	54
P00441	SODC_HUMAN			1	-----MADSRDPASDQNHQKKEQRAAQKADLVTTGAGHPVSGDKLNVIITVGRPRGLVQD-	54
P04179	SODM_HUMAN			1	MLSRVAVCTSRQLA-PALGY-----DSAEHIDRMVAVYV-----	19
P08294	SODE_HUMAN			1	-MLALLCCCLLAAGASDAM-----	19
P07203	GPX1_HUMAN			1	-----NCAARLAA-----AAAAAGSVYAFSARPLAGGEPVSLGSLRGKVLLEN	44

Annotation
 Helix
 Active site
 Natural variant
 Cross-link
 Turn
 Metal binding
 Lipidation
 Transit peptide
 Sequence conflict
 Glycosylation
 Non-standard residue
 Initiator methionine
 Disulfide bond
 Mutagenesis
 Chain
 Alternative sequence
 Beta strand
 Signal peptide
 Site
 Modified residue

Amino acid properties
 Similarity
 Hydrophobic
 Negative

Annexes

Outil STRING

Version: 10.5 LOGIN REGISTER

STRING Search Download Help My Data

Welcome to STRING

Protein-Protein Interaction Networks

ORGANISMS | PROTEINS | INTERACTIONS
2031 | 9.6 mio | 1380 mio

SEARCH

https://string-db.org/cgi/network.pl?block_structure_pics_in_bubbles=0&caller_identity=1&chemicalmode=0&have_user_input Rechercher

STRING Search Download Help My Data

The following proteins in Homo sapiens appear to match your input. Please review the list, then click 'Continue' to proceed to the association network.

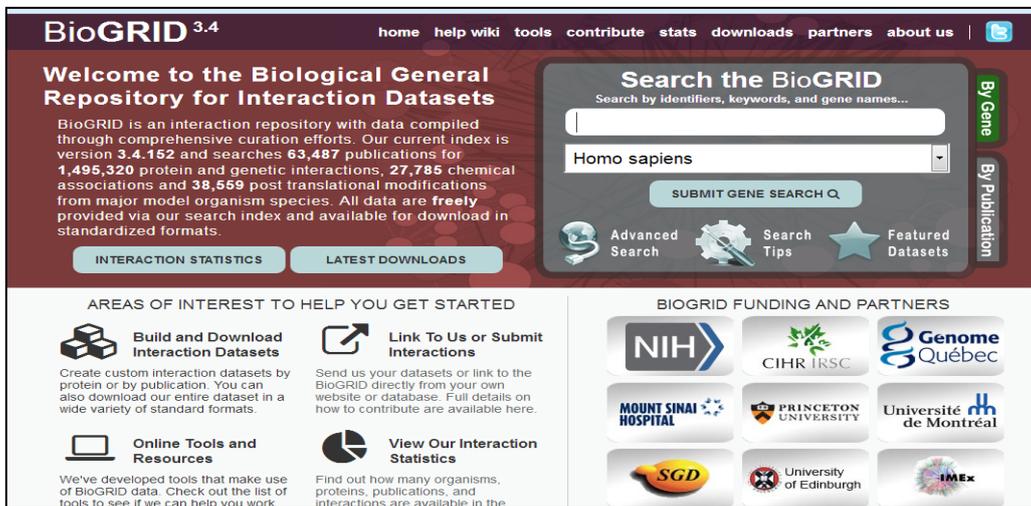
< BACK CONTINUE >

query sequence	STRING protein	annotation	identity	bitscore
NP_000445.1:	<input checked="" type="checkbox"/> SOD1	Superoxide dismutase 1, soluble; Destroys radicals which are normally produced within the cells and which are toxic to biological systems	100 %	310
	<input type="checkbox"/> CCS	Copper chaperone for superoxide dismutase; Delivers copper to copper zinc superoxide dismutase (SOD1)	50 %	137
	<input type="checkbox"/> SOD3	Superoxide dismutase 3, extracellular; Protect the extracellular space from toxic effect of reactive oxygen intermediates by converting superoxide radicals into hydrogen peroxide and oxygen	40 %	103
NP_001309749.1:	<input checked="" type="checkbox"/> SOD2	Superoxide dismutase 2, mitochondrial; Destroys superoxide anion radicals which are normally produced within the cells and which are toxic to biological systems	100 %	368
NP_001316384.1:	<input checked="" type="checkbox"/> GPX1	Glutathione peroxidase 1; Protects the hemoglobin in erythrocytes from oxidative breakdown	77 %	280
	<input type="checkbox"/> GPX2	Glutathione peroxidase 2 (gastrointestinal); Could play a major role in protecting mammals from the toxicity of ingested organic hydroperoxides. Tert-butyl hydroperoxide, cumene hydroperoxide and linoleic acid hydroperoxide but not phosphatidylcholine hydroperoxide, can act as acceptors	69 %	177
	<input type="checkbox"/> GPX5	Glutathione peroxidase 5 (epididymal androgen-related protein); Protects cells and enzymes from oxidative damage, by catalyzing the reduction of hydrogen peroxide, lipid peroxides and organic hydroperoxide, by glutathione. May constitute a glutathione peroxidase-like protective system against peroxide damage in sperm membrane lipids	50 %	118
	<input type="checkbox"/> GPX3	Glutathione peroxidase 3 (plasma); Protects cells and enzymes from oxidative damage, by catalyzing the reduction of hydrogen peroxide, lipid peroxides and organic hydroperoxide, by glutathione	44 %	107
	<input type="checkbox"/> GPX6	Glutathione peroxidase 6 (olfactory)	43 %	97.8
	NP_001743.1:	<input checked="" type="checkbox"/> CAT	Catalase; Occurs in almost all aerobically respiring organisms and serves to protect cells from the toxic effects of hydrogen peroxide. Promotes growth of cells including T-cells, B-cells, myeloid leukemia cells, melanoma cells, mastocytoma cells and normal and transformed fibroblast cells	100 %

Annexes



Figure : SMS (Séquence Manipulation Suite) programme ORF Finder



Annexes
