



République Algérienne Démocratique et
Populaire Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université ABOUBEKR BELKAID – TLEMCEEN

Faculté des Sciences de la Nature et de la Vie, et Sciences de la Terre et de l'Univers

Département de Biologie



MÉMOIRE

Présenté par

Benchabane Amaria

Belkaid Adnane

En vue de l'obtention du

Diplôme de MASTER

En Immunologie

**Analyse *in silico* des SNPs à conséquences fonctionnelles sur CTLA-4 et SRC dans le cancer colorectal :
Prédiction des modifications post traductionnelles et modélisation moléculaire**

Thème

Soutenu le, devant le jury composé de:

Président	Aribi Mourad	Professeur	Université de Tlemcen
Encadrante	Hadjidj Zeyneb	MCB	Université de Tlemcen
Examinatrice	El Mezouar Chahrazed	MCA	Université de Tlemcen

Année universitaire 2022/2023



République Algérienne Démocratique et
Populaire Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université ABOUBEKR BELKAID – TLEMCEM



Faculté des Sciences de la Nature et de la Vie, et Sciences de la Terre et de l'Univers

Département Biologie

MÉMOIRE

Présenté par

Benchabane Amaria

Belkaid Adnane

En vue de l'obtention du

Diplôme de MASTER

En Immunologie

**Analyse *in silico* des SNPs à conséquences fonctionnelles sur CTLA-4 et Src dans le cancer colorectal :
Prédiction des modifications post traductionnelles et modélisation moléculaire**

Thème

Soutenu le, devant le jury composé de:

Président	Aribi Mourad	Professeur	Université de Tlemcen
Encadrante	Hadjidj Zeyneb	MCB	Université de Tlemcen
Examinatrice	El Mezouar Chahrazed	MCA	Université de Tlemcen

Année universitaire 2022/2023

Dédicace

Grace à dieu tout puissant et en signe de reconnaissance à tous les sacrifices consentis pour notre réussite et la volonté pour mener à bien ce modeste travail qu'on dédie :

A ma mère et mon père en témoignage de leur dévouement, leur patience, leur soutien, leur protection et leur encouragement depuis ma naissance. Vous étiez toujours présent quand j'avais besoin de vous, je n'aurais pu achever ce travail sans votre générosité et votre affection. Recevez à travers ce travail, toute ma gratitude et mes profonds sentiments.

A ma sœur KHEDIDJA et ma petite princesse DOUDOUS aussi mes frères ABDELKARIM et DJAWED pour leurs encouragements permanents, et leur soutien morale pour leur appui et leur efforts envers moi.

A mes grands parents et ma cousine GHIZLENE et MARIA qu'elle viennent de nous quitter Miséricorde de Dieu, vous étiez et vous êtes toujours dans notre mémoire esprit. Reposez en paix.

A toute la famille BENCHABANE et YAHI

A mes amies : AMINA, SANAA, IKRAM, ASSOUMA

Un remerciement infini pour mon binôme « BELKAID Adnane » pour votre vivacité et votre bienveillance

BENCHABANE AMARIA

À ma mère Zerga Rabia, ma source constante de soutien et d'inspiration, je voudrais te remercier du fond du cœur pour ta présence dans ma vie. Tu as toujours été là pour m'écouter, me comprendre et me guider lorsque j'en avais besoin. Tes mots encourageants et ton affection inconditionnelle ont été le moteur de ma confiance et de ma croissance personnelle.

À mon père Nouredinne, mon roc dans les moments difficiles, je veux exprimer toute ma gratitude pour ta force et ton dévouement.

Miséricorde de dieu pour ma tante Samia qui nous a quittée à cause du Cancer Colorectal et aussi ma tante Zahira.

Pour ma sœur Bouchra et son mari Hacem, Tu as été mon soutien inconditionnel, me rappelant constamment que je peux surmonter tous les défis qui se présentent à moi. Ta présence rassurante et tes conseils éclairés m'ont permis de garder la tête haute même lorsque les temps étaient durs.

À mes frères Amine et sa femme Wafaa, Seyf-edinne et sa femme Yasmine, Khaled mes oncles d'amitié et de bonheur, je tiens à vous dire combien vous avez illuminé ma vie. Votre présence joyeuse, votre esprit plein de positivité ont fait de chaque moment passé ensemble une véritable bénédiction, d'avoir apporté une telle joie à ma vie.

Sans oublier mon neveu Rayane et mes nièces Alae, Israe, Rihem et notre petite princesse Anfel.

À mes amis nos aventures, nos fous rires et nos confidences ont créé des souvenirs inestimables qui resteront gravés dans ma mémoire pour toujours, Chacun d'entre vous a laissé une empreinte unique dans mon cœur et dans ma mémoire.

Mes dédicaces à toute la famille Belkaid et Zerga.

Merci à mon binôme « Benchabane Amaria » pour ton soutien et ta gratitude.

BELKAID Adnane

Résumé

Introduction : Dans cette étude, nous avons réalisé des analyses *in silico* des Single Nucleotide Polymorphisms (SNPs) ayant des conséquences fonctionnelles sur les protéines CTLA-4 et SRC dans le contexte du cancer colorectal. Nous avons réalisé des prédictions des modifications post-traductionnelles associées aux SNPs identifiés comme potentiellement délétères, ainsi que la modélisation moléculaire des structures tridimensionnelles des protéines pour une visualisation plus précise des altérations.

Matériels et méthodes : Nous avons exploité plusieurs outils bioinformatiques tels que GPS6.0, MusiteDeep, NETPhos 3-1, GPS-MSP, GPS-UBER, PRmePred, GPS.PBS, CarSite-II, GPS-YNO2 et GPS-SNO 1.0 pour prédire les modifications post-traductionnelles. Ensuite, nous avons utilisé les logiciels Phyre2 et SWISS-MODEL pour réaliser la modélisation moléculaire de ces protéines mutées et ERRAT/PROCHECK pour évaluer la qualité de nos résultats.

Résultats : Nos résultats indiquent que certaines variations dans les protéines CTLA-4 (T72I, T147A, N145S) ainsi que c-SRC (P510S, P136R, A296S, T460K), peuvent potentiellement induire diverses modifications post-traductionnelles telles que la phosphorylation, la méthylation, l'ubiquitination ou la carbonylation, aussi induit un changement dans la structure de la protéine.

Conclusions : Notre étude *in silico* fournit une base préliminaire intéressante concernant l'impact possible de certains SNPs sur CTLA-4 et SRC dans le contexte du cancer colorectal. Ces résultats pourront être utilisés comme base de départ pour des études expérimentales visant à valider les effets de ces SNPs sur les protéines et leur rôle dans le développement de cette maladie.

Mots clés : CTLA-4, cancer colorectal, modélisation des protéines, modifications post traductionnelles, nsSNP, SRC.

Abstract

Background: In this study, we conducted in silico analyses of Single Nucleotide Polymorphisms (SNPs) with functional consequences on the CTLA-4 and SRC proteins in the context of colorectal cancer. We used prediction of post-translational modifications associated with SNPs identified as potentially deleterious, as well as three-dimensional molecular modeling of protein structures for a more precise visualization of alterations.

Materials and methods: We exploited several bioinformatics tools such as GPS6.0, MusiteDeep, NETPhos 3-1, GPS-MSP, GPS-UBER, PRmePRed, GPS.PBS, CarSite-II, GPS-YNO2 and GPS-SNO 1.0 to predict post-translational modifications. Then we used Phyre2 and SWISS-MODEL software to perform molecular modeling of these mutated proteins and ERRAT/PROCHECK to evaluate the quality of our results.

Results: Our results indicate that certain variations in coding regions of CTLA-4 proteins (T72I,T147A,N145S), as well as c-SRC (P510S,P136R,A296S,T460K), can potentially induce various post-translational modifications such as phosphorylation, méthylation ubiquitination or carbonylation for example, also induces a change in the structure of the protein.

Conclusions: our in silico study provides an interesting preliminary basis regarding the possible impact of some SNPs on CTLA-4 and SRC in the context of colorectal cancer. These results could be used as a starting point for experimental studies aimed at validating the effects of these SNPs on proteins and their role in disease development.

Keywords: colorectal cancer, CTLA-4, nsSNP, post translational modifications, protein modeling, Src.

ملخص

مقدمه في هذه الدراسة، قمنا بإجراء تحليلات حاسوبية (*in silico*) لتحديد اختلافات وراثية في نوكلبيوتيد واحد (Single Nucleotide Polymorphisms - SNPs) التي لها عواقب وظيفية على بروتينات CTLA-4 و SRC في سياق سرطان القولون. استخدمنا تحليل التغيرات المشاركة في مرحلة مابعد التَّرْجَمَة المُصَادَقَة بهذه SNPs ، بالإضافة إلى نمذجة هذه المستضدات ثلاثية الأبعاد لتصور أكثر دقة للاختلاف.

المواد و الأساليب لقد استغلنا العديد من أساليب حاسوب علم الأحياء ك- GPS6.0, MusiteDeep, NETPhos 3-1, GPS, MSP, GPS-UBER , PRmePRed, GPS.PBS, CarSite-II, GPS-YNO2 and GPS-SNO 1.0 لتحديد التغيرات المشاركة في مرحلة مابعد التَّرْجَمَة. ثم قُمْنَا باستخدام برامج Phyre2 و SWISS-MODEL لتنفيذ نمذجة مولكولارية لهذه البروتينات المُحدثة، و ERRAT/PROCHECK لتقييم جودة النتائج.

النتائج أظهرت نتائجنا أن بعض التغيرات في مناطق تشفير بروتين CTLA-4 (T72I,T147A,N145S) و c-SRC (P510S,P136R,A296S,T460K) يُمكن اِحْتِمَالًا ظهور تغيرات في مرحلة مابعد التَّرْجَمَة كالفَسْفَرِيلة، والإزالة البُورِيَّة، وانضمام رآديكلوب .

الخاتمة خلاصة هذه الدراسة *in silico* هو إعطاء قاعدة أساسية جادّة حول التأثير المُحْتَمَل لبعض SNPs على بروتينات CTLA-4 و SRC في سياق سرطان القولون. يُستخدم هذه الِنتائج كبادئ ذى بدء لإجراء دراسات تجاربية بهدف التأكّد من آثار هذه SNPs على البروتينات و دورها في تطور المرض.

REMERCIEMENTS

Avant tout, on remercie Allah de nous avoir aidées à accomplir ce travail.

On adresse toute nos reconnaissances à l'encadrant Madame HADJIDJ Zeyneb, pour sa patience, sa disponibilité et surtout ses judicieux conseils, qui ont contribué à alimenter nos réflexions.

Un remerciement infini à M. ARIBI Mourad professeur à la faculté de la science de la nature et de la vie de Tlemcen et directeur de laboratoire BIOMOLIM, pour ses précieux conseils

Nous voudrions remercier les membres de notre jury d'avoir accepté d'évaluer ce travail et de nous avoir consacré du temps, en l'occurrence Pr. Mourad ARIBI, et Dr. El Mezouar Chahrazed (université de Tlemcen).

A nos parents (nos piliers) qui nous ont soutenus durant les moments de doute et d'abandon, qui ont plus cru en nous

On voudrait exprimer notre reconnaissance aussi envers les amis et collègues qui nous ont apportés leur soutien moral et intellectuel tout au long de notre démarche.

MERCI INFINIMENT

LISTE DES FIGURES

- Figure1.1** : Nouveaux cas de cancer colorectal en 2020 pour les 02sexes et tous les âges
- Figure 1.2** : Mécanismes moléculaires du cancer colorectal
- Figure 1.3** : Structure tridimensionnelle du CTLA-4
- Figure1.4** : La configuration structurale de CTLA-4
- Figure1.5** : Organisation structurale de la protéine c-Src
- Figure1.6** : La structure 3D du c-SRC
- Figure1.7** : Signalisation liée à la protéine SRC dans une cellule normale et une cellule cancéreuse
- Figure2.1** : Plan de l'étude.
- Figure2.2** : Recherche des séquences sur UniProt
- Figure 2.3** : Prédiction des sites de phosphorylation par GPS 6.0
- Figure 2.4** : Prédiction des sites de phosphorylation par Netphos 3.1
- Figure 2.5** : Prédiction des sites de phosphorylation, de méthylation et d'ubiquitination par MusiteDeep
- Figure 2.6** : Prédiction des sites de méthylation par GPS-MSP
- Figure 2.7** : Prédiction des sites de méthylation par PRmePRed
- Figure 2.8** : Prédiction des sites d'ubiquitination par GPS-UBER
- Figure 2.9** : Prédiction des domaines de liaison aux phosphoprotéines par GPS.PBS
- Figure 2.10** : Prédiction des sites de carbonylation par CarSite-II
- Figure 2.11** : Prédiction des sites de nitration des tyrosines par GPS-YNO2
- Figure 2.12** : prédiction des sites denitration des tyrosines par DeepNitro
- Figure 2.13** : Prédiction des sites de S-nitrosylation des protéines par GPS-SNO 1.0
- Figure 2.14** : prédiction des sites de S-nitrosylation par DeepNitro
- Figure 2.15** : Modélisation des structures tridimensionnelles des protéines par Phyre2
- Figure 2.16** : Modélisation des structures tridimensionnelles des protéines par SWISS-Model
- Figure 2.17** : exemple d'une réalisation sur ERRAT
- Figure2.18** : Évaluations de la qualité des structures des protéines par ERRAT et PROCHECK
- Figure2.19**: alignement des structures des protéines par TM Align
- Figure 2.20** : exemple des résultats sur Chimera

LISTE DES TABLEAUX

Tableau 3.1 : Résultats de prédiction des sites de phosphorylation

Tableau 3.2 : Résultats de prédiction des sites de Méthylation

Tableau 3.3 : Résultats de prédiction des Ubiquitinations

Tableau 3.4 : Résultats de prédiction des domaines de liaison aux phosphoprotéines

Tableau 3.5 : Résultats de prédiction de Tyrosine nitration sites et S-nitrosylation

Tableau 3.6 : Résultats de l'évaluation de qualité des structures modélisées par SWISS-MODEL on utilisant TM-align, ERRAT et PROCHECK

Tableau 3.7 : Visualisation et superposition des protéines modélisées par SWISS-MODEL

Tableau 3.8 : Estimations de qualités des protéines modélisées par SWISS-MODEL

Tableau 3.9 : Alignement des structures tridimensionnelles modélisées par Phyre2 en utilisant TM align

Tableau 7.1 : les séquences FASTA mutées et sauvage

LISTE DES ABREVIATIONS

AA : acide aminé

AC : Anticorps

ADN : Acide désoxyribonucléique

CD : Cluster de Différenciation

APC: Polypose adénomateuse colique

CCR : Cancer colorectal

CTL : Lymphocyte T cytotoxique

SNP : Polymorphisme nucléotidique simple

nsSNP : Polymorphisme nucléotidique simple non synonyme

DbSNP: Base des données des polymorphismes nucléotidique simple

MAPK: Mitogen-activated protein kinases

NK: Natural killer cell

P53: Protéine 53

PI3K:Phosphoinositide 3-kinases

SSP : Polype dentelé sessile

CPA : Cellule présentatrice d'antigènes

CTLA-4 : Antigène 4 du lymphocytes T cytotoxique

ERK: Extracellular signal-regulated kinase

JNK: c-JUN N-terminal kinase.

LAT: Linker Activated T cells.

MAGE-3: Melanoma-associated antigen 3

MDSC: Myeloïd-derived suppressor cells

MUC-1: Mucin 1

mTOR: Mammalian target of rapamycin

NCBI: National Center for Biotechnology Information

NF-κB: Nuclear factor-kappa B

NY-ESO-1: New York esophageal squamous cell carcinoma 1

PAF : Polypose adénomateuse familiale

SRC: Sarcoma

TGF β : Transforming growth factor

TNM: Tumor, Nodes, Metastasis

VH : Chaîne lourde

VL : Chaîne légère

ZAP-70 : Protéine zêta-associée de 70kDa

FIT : test de sang occulte fécal

Plan de travail

Dédicace	III
Résumé	IV
Abstract.....	V
ملخص	VII
REMERCIEMENTS.....	VIII
LISTE DES FIGURES.....	VII
LISTE DES TABLEAUX	IX
LISTE DES ABREVIATIONS.....	X
Introduction.....	2
I. Chapitre I : Revue de la littérature	5
1. Cancer colorectal.....	5
1.1. Généralités :	5
1.2. Du polype au cancer	6
1.2.1. Cellules dysplasiques	6
1.2.2. Adénocarcinome.....	6
1.3. Facteurs de risque.....	6
1.4. Symptômes et diagnostic du cancer colorectal	7
1.4.1. Symptômes	7
1.4.2. Diagnostic	7
1.5. Traitement du cancer colorectal.....	8
1.6. Principaux mécanismes de la carcinogenèse colorectale.....	8
1.6.1. Lésions colorectales précancéreuses :.....	8
1.6.2. Bases moléculaires de la carcinogenèse colorectale.....	9
1.6.2.1. Instabilité chromosomique (CIN)	9
1.6.2.2. Instabilité microsatellitaire (MSI)	9
1.6.2.3. Le phénotype méthylateur des îlots CpG (CIMP).....	10
1.7. Immunoédition du cancer	11
1.8. Voies de signalisation impliquées dans le cancer colorectal	12
1.8.1. La voie P53 :.....	12
1.8.2. La voie TGFβ/SMAD.....	12
1.8.3. La Voie Wnt/APC/β-catenine	12
1.8.4. Voie PIK3/AKT.....	13
1.8.5. Voie RAS/PAF/MAPK	13
1.8.6. Cyclooxygénase-2 (COX-2)	14
1.8.7. Tyrosine kinase cytoplasmique (Src).....	14
1.8.8. Voie de l'insuline et O-GlcNAcylation.....	14
2. CTLA-4.....	14

2.1.1.	Régulation de l'expression du gène CTLA-4	14
2.2.	Structure du CTLA-4.....	15
2.3.	Rôle de CTLA-4 dans la réponse immunitaire	15
2.4.	Implication du CTLA-4 dans le cancer colorectal.....	16
2.5.	Inhibition de CTLA-4 dans le traitement du cancer colorectal.....	16
3.	c-Src.....	17
3.1.	Structure de la protéine c-Src humaine	17
3.2.	Les voies de signalisation associées à la protéine c-Src	18
3.3.	Implication de Src dans le cancer colorectal	19
3.4.	Signalisation de c-Src dans le cancer colorectal	20
3.5.	Inhibition de c-Src dans le traitement du cancer colorectal	20
4.	Polymorphisme nucléotidique	21
4.1.	Généralités	21
4.2.	Classification des SNP	21
4.2.1.	SNPs des régions non-codantes.....	21
4.2.2.	SNP des régions codantes	21
	SNP non-Synonyme.....	22
4.3.	Évaluation des facteurs de risque associés aux SNP	22
4.4.	Intérêt de l'étude du polymorphisme nucléotidique	22
4.5.	Rôle potentiel des SNP dans la prédisposition et le diagnostic du cancer colorectal	23
4.6.	Rôle des polymorphismes nucléotidiques dans le traitement du cancer colorectal	23
4.7.	Analyse <i>in silico</i> des SNP.....	23
4.7.1.	Définitions	23
4.7.2.	Importances des analyses <i>in silico</i> des SNP dans les maladies humaines	24
5.	Modifications post-traductionnels.....	24
5.1.	Effets des SNPs sur les modifications post traductionnelles	24
5.2.1.	La phosphorylation :	25
5.2.2.	La Méthylation.....	25
5.2.3.	L'ubiquitination.....	25
6.	Problématique.....	26
II-	Chapitre II : Matériel et méthodes.....	29
1.	Plan de l'étude	29
2.	Collecte de données et sélection de séquences de protéines	30
3.	Prédiction des modifications post traductionnelles.....	31
3.1.	Prédictions des sites de phosphorylation.....	31
3.1.1.	GPS 6.0	31
3.1.2.	NETPHOS.....	32
3.1.3.	MusiteDeep.....	33
3.2.	Prédiction des sites de méthylation	34

3.2.1.	GPS-MSP	34
3.2.2.	PRmkePRed	34
3.3.	Prédiction des sites d'ubiquitination	35
3.4.	Prédiction des domaines de liaison aux phosphoprotéines	35
3.5.	Prédiction des sites de carbonylation	36
3.6.	Prédiction des sites de nitration des tyrosines	37
3.6.1.	GPS-YNO2	37
3.6.2.	DeepNitro	38
3.7.	Prédiction des sites de S-nitrosylation.....	38
3.7.1.	GPS-SNO 1.0.....	38
3.7.2.	DeepNitro	38
4.	Modélisation moléculaire	39
4.1.	Modélisation par Phyre2.....	39
4.2.	Modélisation par SWISS-MODEL.....	40
5.	Évaluations de la qualité des structures des protéines modélisées	41
5.1.	ERRAT.....	41
5.2.	PROCHECK.....	42
6.	Comparaison entre les structures tridimensionnelles des protéines	43
6.1.	TM-align.....	43
6.2	Chimera.....	43
III-	Chapitre III : Résultats	46
1.	Résultats de Prédiction des modifications post-traductionnelles	46
1.1	Prédiction des sites de phosphorylation	46
1.2.	Résultats de prédiction des Méthylation	47
1.3.	Résultats de prédiction des Ubiquitinations	48
1.4.	Résultats de prédiction des domaines de liaison aux phosphoprotéines.	49
1.5.	Résultats de prédiction de Tyrosine nitration sites et S-nitrosylation	49
2.	Modélisation moléculaire	50
2.2.	Modélisation par SWISS-MODEL.....	50
2.2	Modélisation par Phyre2 :.....	67
IV-	DISCUSSION	69
V-	Chapitre V : Conclusions et perspectives.....	72
VI-	Chapitre VI : La bibliographie.....	74

INTRODUCTION

Introduction

Le cancer colorectal (CCR) est un problème de santé majeur qui est largement répandu à l'échelle mondiale. Il se classe au deuxième rang en termes d'incidence et de mortalité parmi toutes les affections malignes à travers le monde (**Belhamidi et al. 2018**). Il s'agit d'une croissance massive anormale des cellules épithéliales glandulaires du côlon et/ou le rectum (**Hossainet al., 2022**).

Les difficultés reconnues pour détecter cette maladie la rendent plus dangereuse, car elle ne présente pas de symptômes spécifiques jusqu'à un stade avancé où les chances de guérison deviennent faibles (**Vega et al., 2015**). Ceci provoque un retard de diagnostic et réduit les chances de traitement (**Loubna Allam, 2020**). Par conséquent, il est essentiel de mettre en œuvre des méthodes préventives telle que l'identification de variations génétiques liées à un risque accru de cancer colorectal. Ainsi, les personnes identifiées comme étant prédisposés au cancer colorectal peuvent bénéficier d'un suivi médical plus étroit, de dépistages plus fréquents et d'interventions préventives appropriées afin de réduire le risque de développer la maladie.

Le système immunitaire joue un rôle majeur dans la défense contre les tumeurs, Toutefois, les cellules cancéreuses peuvent élaborer des stratégies d'évasion immunitaire qui impactent l'amorçage, l'activation et la fonction des lymphocytes spécifiques au cancer, Le CCR ne fait pas exception à ce phénomène, Le microenvironnement tumoral a une influence négative sur l'activité cytolytique des lymphocytes T réactifs aux cellules cancéreuses, tandis que la présence de lymphocytes T "non effecteurs" avec un profil régulateur ou anergique favorise la progression du cancer colorectal (**Niccolai et al., 2017**).

Le CTLA-4 (Cytotoxic T-Lymphocyte Antigen 4) et le SRC (proto-oncogène tyrosine-protein kinase Src) sont deux molécules qui jouent un rôle clé dans la régulation du système immunitaire et la transmission de signaux cellulaires. Le dysfonctionnement de CTLA-4 et SRC peut contribuer au développement et à la progression des cancers (**Zou et al., 2018; Jin, 2020**). CTLA-4 est un récepteur principalement exprimé à la surface des lymphocytes T (**Van Coilliet al., 2020**). C'est un point de contrôle important qui se lie aux molécules CD80 et CD86 exprimés par les cellules présentatrices d'antigènes (CPA) permettant l'inhibition des lymphocytes (**Xu, Jie, éd., 2020**). D'autre part, SRC est reconnu comme l'un des premiers oncogènes connus, avec un rôle essentiel dans de nombreuses voies cellulaires clés, notamment celles impliquées dans la régulation, la prolifération, la différenciation, la survie,

la mobilité et l'angiogénèse(Finn, 2008).

Les SNPs correspondent à des modifications d'ADN pouvant survenir à l'intérieur ou à l'extérieur des gènes(Stanton, 1999). Ces SNPs sont généralement considérés comme le résultat d'une seule mutation, telle qu'une substitution, une insertion ou une suppression(Abd El-Fattahet *al.*, 2018), Le nsSNP (non-synonymous single nucleotide polymorphism) représentent un type de variation génétique qui altère un seul nucléotide dans une séquence d'ADN, entraînant ainsi l'incorporation d'un acide aminé différent dans la protéine (Malik *et al.*, 2020).

L'identification de ces SNP peut-être utile pour comprendre le fonctionnement de certaines maladies et développer des traitements thérapeutiques. C'est leur principal avantage(Kim and Misra, 2007). Les techniques expérimentales visant à identifier l'effet des nsSNP sont très coûteuses, laborieuses et chronophages. Par conséquent, les approches computationnelles dites "*in silico*" peuvent être utilisées en tant qu'outils préliminaires pour analyser l'impact des nsSNP à haut risque ou qui peuvent affecter des cibles médicamenteuses importantes(Venkata Subbiahet *al.*, 2020)(Mah *et al.*, 2011).

En 2022, Megnafi Hayat et Boukenadel Amel ont réalisé une analyse *in silico* du polymorphisme du CTLA-4 et SRC afin d'identifier les variations délétères à risque élevé qui peuvent contribuer à l'installation d'un cancer colorectal ou d'autres pathologies. Dans ce travail, nous allons étudier *in silico* l'effet de ces variations sur la structure tridimensionnelle des protéines et leurs modifications post traductionnelles.

CHAPITRE I :
REVUE DE LA
LITTERATURE

I. Chapitre I : Revue de la littérature

1. Cancer colorectal

1.1. Généralités :

Le CCR est l'un des cancers les plus fréquents dans le monde, avec près de 1,8 million de nouveaux cas diagnostiqués chaque année (Yue *et al.*, 2021). Il est également la deuxième cause de mortalité liée au cancer dans le monde entier (Tehamiet *al.*, 2023). En 2015, le cancer colorectal a représenté un pourcentage significatif des nouveaux cas diagnostiqués en Algérie, soit environ 11,2 % avec une moyenne d'âge au moment du diagnostic se situant entre 65 et 70, selon les données de GLOBACAN (Abeset *al.*, 2015).

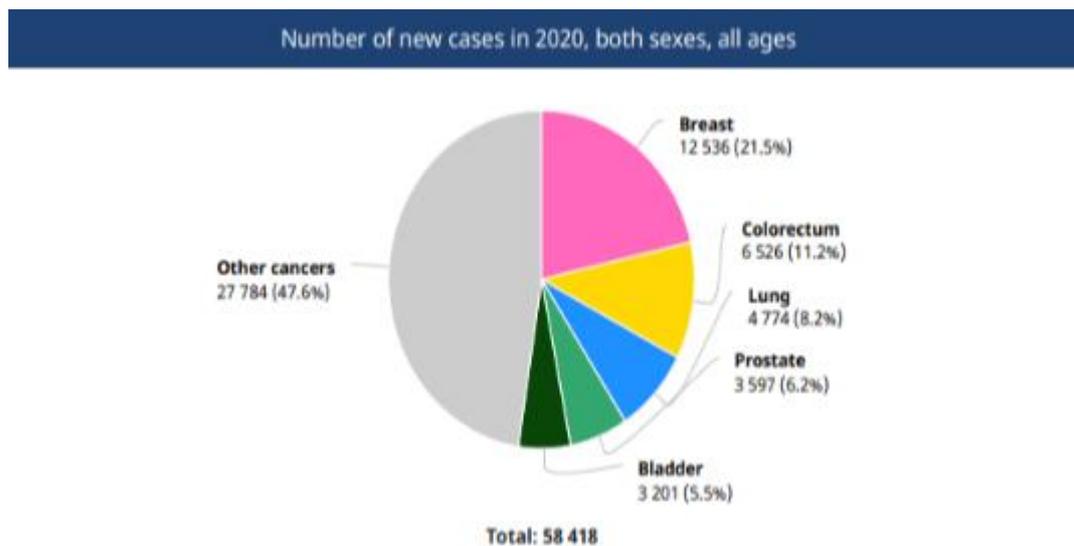


Figure 1.1 : profil d'incidence du cancer colorectal en Algérie l'année 2020 correspondant les 02 sexes et tous les âges (GLOBACAN)

On observe une prévalence plus élevée de CCR chez les hommes par rapport aux femmes, puisque le cancer du sein domine dans cette dernière catégorie, et une incidence 3 à 4 fois supérieure dans les pays développés en comparaison avec les pays sous-développés (Tehamiet *al.*, 2023). Il est possible que le CCR ne présente pas de symptômes dans ses stades précoces, et cela se produit dans environ 20 % des cas puisqu'il nécessite généralement 10 à 15 ans pour qu'un polype se transforme en tumeur maligne (Keum and Giovannucci, 2019). Dans de tels cas, le cancer est généralement détecté lorsque la maladie s'est déjà propagée. Toutefois, la détection précoce du CCR et un traitement rapide peuvent considérablement réduire le risque de mortalité (Misselwitz, 2017).

Il se présente sous forme de masse cancéreuse solide, où La formation d'une tumeur cancéreuse dans le côlon commence toujours par une excroissance d'un polype bénin sur la muqueuse colique(**Stintzing, 2014**). En général, le cancer du côlon débute par un adénocarcinome qui se forme à partir des glandes de Lieberkühn, responsables de la production de mucus dans la paroi du côlon(**Picard-Croguennec, 2018**). Bien que tous les polypes ne se transforment pas en cancer (seulement 1%), leur risque de dégénérescence augmente proportionnellement avec leur taille (un sur dix pour les polypes de plus de 1 cm)(**Stintzing, 2014**).

1.2. Du polype au cancer

1.2.1. Cellules dysplasiques

Lors du processus initial, les cellules saines du côlon subissent une conversion en cellules dysplasiques inoffensives, donnant ainsi naissance aux polypes, qui peuvent être catégorisés en adénomes présentant des anomalies de différenciation de faible ou de haut grade(**Picard-Croguennec, 2018**).

1.2.2. Adénocarcinome

Les différentes parois du côlon, comprenant la muqueuse, la sous-muqueuse, la couche musculaire et la séreuse, sont infiltrées par les cellules cancéreuses, et lorsque celles-ci pénètrent les vaisseaux sanguins et lymphatiques de la sous-muqueuse, le cancer devient infiltrant, augmentant ainsi le risque de métastases ; les ganglions à proximité de la tumeur sont les premiers à être touchés par les métastases, qui peuvent ensuite se propager au foie et aux poumons(**Picard-Croguennec, 2018**).

1.3. Facteurs de risque

Le risque individuel de cancer colorectal est déterminé par des facteurs inhérents qui ne peuvent pas être modifiés, tels que l'âge, le sexe et les antécédents familiaux. Toutefois, ce risque peut être également influencé par des facteurs de risque modifiables, tels que l'alimentation, le mode de vie, ou l'exposition à certaines substances(**Platzet al., 2000**). En effet, une consommation importante de viandes rouges, charcuteries et saucisses a été liée à un risque accru de cancer colorectal, tandis qu'une consommation élevée de poisson, de légumes frais et de légumes cuits a été associée à une réduction du risque de cette maladie(**Imad et al., 2020**). D'autre part, l'obésité (caractérisée par un indice de masse corporelle (IMC) égal ou supérieur à 25 kg/m²) (**Desjardin, 2014**) et une mauvaise santé bucco-dentaire (**Amal Idrissi Janati, 2014**) peuvent aussi contribuer à l'augmentation du risque de cancer colorectal.

Les MICI, qui sont des maladies inflammatoires chroniques de l'intestin comme la colite ulcéreuse ou la maladie de Crohn, figurent également parmi les facteurs de risque (Thicoipé, 2015) et c'est probablement causé d'un taux des mutations sporadiques élevé et de l'inflammation chronique du mucus où la conséquence d'un renouvellement cellulaire est accrue (Y Peterson, 2015).

Les antécédents familiaux représente un facteur de risque classique qui représentent 25% des cas positifs par cette maladie issue globalement des parents au premier degrés (Wilschut *et al.*, 2011).

La polypose adénomateuse familiale (PAF) et le cancer colorectal sans polypose héréditaire (le syndrome de Lynch), sont causés par des mutations génétiques de la lignée germinale et ont à un risque accru chez les membres de la famille, mais représentent moins de 5% de tous les cancers (Johnson *et al.*, 2013).

En résumé, Environ 70% des cas de cancer colorectal sont sporadiques, tandis que seulement 5% sont liés à des conditions héréditaires connues. Les cas restants présentent une prédisposition familiale apparente sans mutations germinales identifiables (Curraiset *et al.*, 2022).

1.4. Symptômes et diagnostic du cancer colorectal

1.4.1. Symptômes

Il est possible que les premiers stades du CCR ne présentent pas de signes ou symptômes visibles (Burnett-Hartman *et al.*, 2021).

Les patients présentent principalement des signes cliniques tels que des problèmes de transit (diarrhée, constipation ou alternance), des rectorragies, une perte de poids, des douleurs abdominales, un syndrome rectal caractérisé par des empreintes, des ténésmes, des faux besoins et des évacuations incomplètes, ainsi que des vomissements et une gêne en position assise (El Housse *et al.*, 2015).

1.4.2. Diagnostic

Il a été reconnu depuis longtemps une méthode simple qui détermine 70 % des cancers du rectum et 30 % des CRC c'est l'examen de rectum (Świdarska *et al.* 2014)

Maintenant et au développement de la science et des méthodes de diagnostic plusieurs d'autres approches sont mises en œuvre telle que l'endoscopie qui regroupe 02 méthodes : la coloscopie qui permet de donner une image entière du colon à l'intérieur avec une haute sensibilité et spécificité ; La sigmoïdoscopie autorise l'obtention d'une image de la partie inférieure du colon ou du rectum (Cieśla, Cieśla, et Cieślak *s. d.*). En outre, Le test de sang occulte fécal (FIT) est un test immunochimique qui est à la recherche des hémoglobines dans les selles car cela indique un

saignement du tractus gastro-intestinal (un test Fit positif n'est pas toujours rapporter avec CCR puisque un polypes de 1 à 2cm peut donner les mêmes résultats) (**Levin et al. 2008**).

Les examens de diagnostiques peuvent aussi attribuer au l'Antigène carcinoembryonnaire (ou ACE) et l'antigène carbohydate (AC 19-9) qui sont les marqueurs tumoraux sanguins ; mais les résultats ne sont pas tellement fiable puisque cet examen biologique et moins sensible(**Song and Li, 2016**)

1.5. Traitement du cancer colorectal

Les options de traitement du CCR dépendent de divers facteurs tels que l'âge du patient, son état de santé général, le stade du cancer et le type de cancer. La chirurgie demeure l'élément central du traitement curatif, et les techniques opératoires ont connu une évolution considérable au cours des dernières décennies (**Nikiteaset al., 2016**). La chirurgie la paroscopique est désormais couramment utilisée pour traiter le cancer du côlon (**Nikiteaset al., 2016**).

En plus de la chirurgie, d'autres options de traitement peuvent inclure l'ablation par radiofréquence, la cryochirurgie, la chimiothérapie, la radiothérapie et la thérapie ciblée. D'autre part, la médecine a connu une véritable avancée avec l'introduction d'immunothérapie en complément de ces méthodes classiques. Cette approche renforce le potentiel du système immunitaire et ouvre de nouvelles perspectives, en offrant des traitements plus ciblés et potentiellement moins invasifs (**Aparicio, 2018**).

1.6. Principaux mécanismes de la carcinogenèse colorectale

La carcinogenèse est un processus complexe en plusieurs étapes impliquant des modifications et des mutations des gènes qui régulent la fonction cellulaire normale, y compris les processus de contrôle de la croissance cellulaire(**Klaunig, 2020**).

1.6.1. Lésions colorectales précancéreuses :

Il existe deux types principaux de polypes avec un haut potentiel de malignité, les adénomes et les polypes sessiles serrés (SSPs)(**Simon, 2016**). La grande majorité des cancers colorectaux sporadiques (non-héréditaire) se développent à partir de polypes adénomateux colorectaux(**Conteduca et al., 2013**). Les adénomes se forment lorsque les mécanismes habituels de réparation de l'ADN et de prolifération cellulaire sont perturbés, Les polypes adénomateux augmentent de taille, présentent des caractéristiques dysplasiques (modifications anormales de croissance) au fil du temps et peuvent éventuellement devenir invasifs(**Nguyen et al., 2020**).La

plupart des adénomes ont une histologie tubulaire avec des glandes petites et atypiques, mais peuvent développer des zones d'architecture filamenteuse en croissance, rapportées en pathologie comme villiforme ou tubulovilliforme.

En revanche, les polypes sessiles serrés (SSPs) sont plats et ont une apparence de tapis, avec des glandes dentelées ou en forme de scie. Les SSPs comprennent les adénomes sessiles serrés, les adénomes traditionnels serrés et les polypes mixtes, qui ont tous été associés au développement du CCR(Simon, 2016).

1.6.2. Bases moléculaires de la carcinogenèse colorectale

Le cancer colorectal est une maladie hétérogène tant du point de vue de la carcinogenèse moléculaire que des voies morphologiques en plusieurs étapes. Trois processus moléculaires impliqués dans la cancérogenèse ont été identifiés : l'instabilité chromosomique (CIN), l'instabilité des microsatellites (MSI) et le phénotype méthylateur des îlots CpG (CIMP)(Ren and Tao, 2018).

1.6.2.1. Instabilité chromosomique (CIN)

Présente dans 65% à 70% des tumeurs colorectales sporadiques, se caractérise par des changements chromosomiques tels que les altérations du nombre de copies somatiques (SCNAs) causées par l'aneuploïdie, les délétions, les insertions, les amplifications ou la perte d'hétérozygotie. Les tumeurs qui empruntent cette voie sont considérées comme non hyper mutées en raison du faible nombre de mutations de paires de bases dans les séquences codantes. Les mécanismes sous-jacents à la CIN sont généralement liés à des défauts dans la ségrégation chromosomique, la sénescence cellulaire perturbée due au raccourcissement des télomères et à la réorganisation génomique, le dysfonctionnement de la réponse aux dommages de l'ADN et la perte d'hétérozygotie d'un gène suppresseur de tumeur(Nguyen *et al.*, 2020).

1.6.2.2. Instabilité microsatellitaire (MSI)

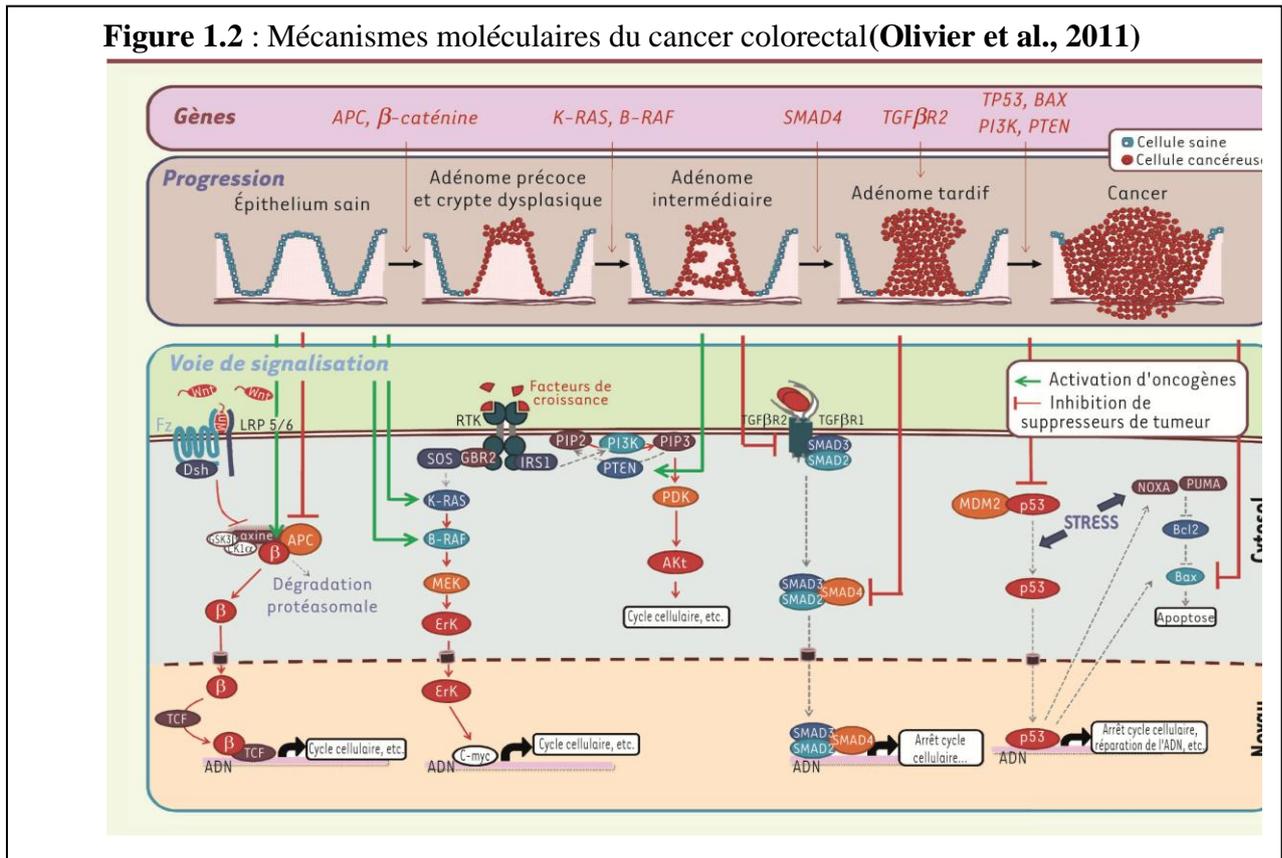
L'instabilité microsatellitaire(MSI) est un phénomène génétique caractérisé par une augmentation de la fréquence des mutations causées par un dysfonctionnement de la réparation des mésappariements de l'ADN (MMR), qui est responsable de corriger les erreurs lors de la réplication de l'ADN. Les cellules présentant une MMR défectueuse accumulent des erreurs et génèrent de nouveaux fragments de microsatellites. La MSI est fréquemment observée dans environ 15% des cancers colorectaux sporadiques et dans la quasi-totalité des cancers colorectaux associés au syndrome de Lynch. Les cancers colorectaux avec MSI peuvent également présenter

des altérations dans d'autres gènes tels que EGFR, BAX et TGFbetaRII, mais sont généralement indépendants des mutations dans KRAS ou TP53. Le statut de BRAF constitue un important facteur pronostique chez les patients atteints de MSI, avec une meilleure survie sans maladie et une survie globale améliorée chez ceux qui ne présentent pas de mutation de BRAF. Les cancers colorectaux avec MSI sont généralement dépourvus d'anomalies chromosomiques(**Ren and Tao, 2018**).

1.6.2.3. Le phénotype méthylateur des îlots CpG (CIMP)

La méthylation de l'ADN est un processus épigénétique qui régule la transcription des gènes en ajoutant un groupe méthyle aux bases cytosine lorsqu'elles sont appariées avec guanosine (CpG)(**Worthley, 2007**). Les altérations épigénétiques désignent des modifications de l'expression ou de la fonction d'un gène sans modifier sa séquence d'ADN spécifique. CIMP se manifeste par la présence simultanée d'une hyperméthylation de plusieurs gènes(**Al-Sohaily et al., 2012**).

Figure 1.2 : Mécanismes moléculaires du cancer colorectal(Olivier et al., 2011)



la transition d'un épithélium sain vers un adénome puis à un cancer c'est à travers la condensation progressive des altérations génétiques au niveau des cellules muqueuses colique et rectale (Benamrouz 2012).

La figure ci dessus montre les étapes d'acheminement du CCR et les bases moléculaires concernés. Un épithélium sain progresse une action d'inhibition sur les APC (Adenomatous Polyposis Coli) et activation sur les gènes suppresseurs de tumeurs : β-caténine (Benamrouz 2012) dont 90% des CCR approuvés la voie Wnt / β-caténine était l'élément initiateur dont il participe à l'embryogenèse et le renouvellement de l'épithélium des cryptes intestinales (Reya et Clevers 2005).

L'oncogenèse commence à la mutation d'APC et/ ou β caténine provoquant la naissance d'un adénome (Huels et Sansom 2015) ensuite une progression vers un adénome intermédiaire à l'aide du pro-oncogène KRAS (Kirsten rat sarcoma viral oncogene homologue) (Huang et al. 2021), et le gène B-RAF (Ciombor et al. 2022) qui favorisent la prolifération cellulaire ; après l'altération de fonction des gènes SMAD4 et TGFβ-R2 la propagation à un adénome tardif et puis une mutation tardive s'installe dans le gène suppresseur de tumeur TP53 et donc le stage métastatique (Benamrouz 2012) les mutations séquentielles de APC, KRAS, TP53, et SMAD4 permettent la signalisation oncogène et d'établir les caractéristiques du cancer. (Sell et al., 2023)

1.7. Immunoédition du cancer

Le CCR est capable d'établir une tolérance immunitaire en utilisant divers mécanismes d'immunoédition et d'évasion immunitaire, lui permettant ainsi d'échapper à la surveillance du système immunitaire (Forster and Radpour, 2020). C'est un processus par lequel le système immunitaire peut détecter les cellules cancéreuses transformées et déclencher une réponse antitumorale, ce qui peut entraîner la sélection de cellules tumorales moins immunogènes (Grasso et al., 2018).

Ce processus se compose de trois étapes distinctes. Tout d'abord, une phase préliminaire vise à éliminer les cellules cancéreuses naissantes en raison de leur immunogénicité élevée. Ensuite,

une période d'équilibre se met en place, durant laquelle la réponse immunitaire est maintenue. Enfin, les cellules cancéreuses évoluent en développant des mécanismes génétiques ou épigénétiques qui leur permettent de croître progressivement chez un hôte immunocompétent, ce qui conduit à leur évasion (**Grassoet al., 2018**).

1.8. Voies de signalisation impliquées dans le cancer colorectal

Diverses études ont été réalisées afin d'améliorer la compréhension des mécanismes de signalisation associés à ce genre de cancer. Dans la signalisation maligne les voies de signalisation majeure sont : mdm2/p53, Wnt/APC/ β caténine, TGF β /SMAD, RAS/MAPK, PI3k/AKT, Voie de l'insuline et O-GlcNAcylation, Tyrosine kinase cytoplasmique (Src), et la Cyclooxygénase-2 (COX-2) ils sont impliqués non seulement dans la cancérogenèse coliques mais aussi plusieurs d'autres cancer (**Agnes s. d.**) :

1.8.1. La voie P53 :

C'est un gène situé sur le chromosome 17p et se compose de 11 exons et 10 introns. L'un des gènes suppresseurs de tumeurs les plus communément étudiés car son vaste incorporation dans les phénomènes biologiques : l'arrêt de cycle cellulaire, la sénescence et l'apoptose en présence d'un stress cellulaire (**Li et al. 2015**)

La mutation de p53 dans le CCR se produit dans 34% des tumeurs du côlon proximal et dans 45% des tumeurs colorectales distales (**Li et al. 2015**). L'activation de P53 peut provoquer à la fois la voie intrinsèque et la voie extrinsèque. Les altérations génétiques du gène TP53 jouent un rôle crucial dans la transition adénome-carcinome au cours du développement de la pathologie tumorale(**Li et al. 2015**)

1.8.2. La voie TGF β /SMAD

La voie de signalisation TGF β participe dans plusieurs action vital dans la cellule : prolifération, la différenciation, la mobilité, l'adhésion et la mort cellulaire (**Buc 2012**) donc il semble être un cible meilleur pour la cancérogenèse. Une mutation activatrice du gène codant pour TGF β RII est retrouvée dans plus de 30% de l'ensemble des CCR provoquant l'augmentation de la prolifération cellulaire (**Buc 2012**)

En outre, Smad4 est mutée dans 16 à 25% des CCR et une altération de Smad2 a été identifiée dans 6% des CCR. En fait Smad2 et Smad4 sont tous les deux situés sur le chromosome 18q, une région très fréquemment délaitée dans le CCR (**Buc 2012; Grady et Pritchard 2014**)

1.8.3. La Voie Wnt/APC/ β -caténine

La voie Wnt est une voie de signalisation cruciale dans l'embryogenèse et même la carcinogenèse. Dans le cadre de l'activation normal, les récepteurs Frizzled par la fixation du

ligand Wnt fournissent la phosphorylation de la protéine Dishevelled qui, par son association à l'Axine, empêche la GSK3B de phosphoryler ses substrats dont la β -caténine, qui sera alors transloquée dans le noyau ou elle s'associe aux facteurs de transcription TCF/LEF pour activer la transcription de gènes cibles (C-myc, Cycline D1...etc). En l'absence du ligand Wnt, la GSK3B phosphoryle la β -caténine dans le complexe Axine/APC qui sera ubiquitinylée et dégradée par le protéasome (van Amerongen et Nusse 2009)

APC est un gène suppresseur de tumeur après sa mutation il empêche la phosphorylation efficace de la β -caténine, ce qui entraîne son accumulation et puis la β -caténine se translate au noyau, conduisant la transcription du gène médiés par LEF / TCF (Najdi, Holcombe, et Waterman 2011)

1.8.4. Voie PIK3/AKT

Cette voie peut être activée, soit directement par activation de récepteurs à activité tyrosine kinase (EGFR ou autre), soit par l'intermédiaire de la protéine RAS, Il existe donc un lien étroit entre la voie de signalisation RAS/MAPK et la voie PI3K/AKT, par l'intermédiaire de RAS (Danielsen et al, 2015), Les mutations les plus fréquemment observées dans le CCR sont la surexpression d'IGF2 et les mutations de PIK3CA, ces mutations sont retrouvées dans environ 40% des tumeurs du gros intestin (Zoratto et al. 2014)

1.8.5. Voie RAS/PAF/MAPK

La voie de signalisation MAPK est l'une des voies les plus importantes pour la prolifération cellulaire. La mutation de la protéine RAS permet de déclencher la MAPK, elle joue un rôle dans la transmission des signaux extracellulaires provenant des récepteurs membranaires vers le noyau (Buc 2012)

Cependant, Les protéines RAS font partie de la famille des GTPases, et jouent un rôle primordial dans la transmission de signaux extracellulaires provenant des récepteurs membranaires vers le noyau pour fournir la régulation de la prolifération, de la survie, de la différenciation et de la migration cellulaire, ainsi que de l'angiogène (Cully et al. 2006)

la signalisation des MAPK commence par le recrutement de la protéine RAF-1 à la membrane après activation de RAS, qui booste une cascade de phosphorylation des protéines MEK (MAPK-ERK-Kinase) et ERK (Extracellular signal-Regulated Kinase) ; cette dernière subit une translocation vers le noyau et entraîne l'expression de gènes codant pour des facteurs de transcription (c-FOS, c-MYC, c-JUN ou JUNB) qui, à leur tour, stimulent l'expression d'un grand nombre de gènes, en particulier ceux de la cycline D1 et de cdk6 ayant un rôle majeur dans l'initiation du cycle cellulaire en phase G1 (Buc 2012)

Dans le cancer colorectal, les mutations de KRAS (codon 12,13 et plus rarement 61) entraînent l'activation constitutive de la voie et sont retrouvées dans 50% des tumeurs de phénotype CIN, alors que les mutations de B-RAF spécifient plutôt les tumeurs de phénotype MIN. L'activation de la voie RAS/MAPK peut également être associée à la surexpression de

récepteurs de facteurs de croissance comme l'EGFR, surexprimé dans 30 à 70% des CCR (Neuville, 2010)

1.8.6. Cyclooxygénase-2 (COX-2)

Les COX jouent un rôle crucial dans la régulation de l'angiogenèse, de l'inflammation et de la carcinogenèse. COX-2 est une enzyme à durée de vie courte qui est impliquée dans la production de prostaglandines, des substances qui jouent un rôle important dans l'inflammation (Sheng *et al.*, 2020). La surexpression de la COX-2 dans les lignées cellulaires du CCR entraîne des modifications des gènes régulateurs favorisant l'angiogenèse et la formation de nouveaux vaisseaux sanguins nécessaires à la croissance tumorale (Koehne and Dubois, 2004).

1.8.7. Tyrosine kinase cytoplasmique (Src)

La tyrosine kinase cytoplasmique Src joue un rôle crucial en tant que régulateur de la signalisation induite par les facteurs de croissance et les intégrines. Elle est impliquée dans la croissance tumorale et la formation de métastases (Naudin, 2012).

1.8.8. Voie de l'insuline et O-GlcNAcylation

La O-GlcNAcylation modifie les protéines présentes dans le cytosol, le noyau et les mitochondries, La O-GlcNAcylation contribue à la désensibilisation de la voie de l'insuline observée dans de nombreux syndromes métaboliques et impacte un grand nombre d'oncogènes et de suppresseurs de tumeurs (Olivier *et al.*, 2011).

2. CTLA-4

CTLA-4 est une protéine impliquée dans la régulation du système immunitaire, en particulier dans la suppression des réponses immunitaires. Il agit en se liant aux molécules B7.1 et B7.2 avec une affinité supérieure à celle de CD28, conduisant à une réduction de la production d'IL-2 et à une inhibition de l'activation des lymphocytes T (Rosiket *et al.*, 2021).

2.1.1. Régulation de l'expression du gène CTLA-4

Plusieurs mécanismes régulent l'expression du gène CTLA-4, dont le facteur NFAT (Nuclear Factor of Activated T-cells) qui joue un rôle crucial dans son induction dans les lymphocytes. Les inhibiteurs de NFAT sont capables de supprimer l'expression du gène CTLA-4 en perturbant l'activation des lymphocytes T, ce qui peut avoir des implications dans le contexte de la réponse immunitaire antitumorale (Hosseini *et al.*, 2020).

Le CTLA-4 n'est pas exprimé par les lymphocytes T en état de repos, mais est induit après les premières étapes de l'activation des lymphocytes T (Perkins *et al.*, 1996).

2.2. Structure du CTLA-4

CTLA-4 est une protéine du système immunitaire formée de 223 acides aminés, avec une séquence de peptide signal de 35 acides aminés. Elle est présente sous forme appariée et pèse entre 41 et 43 kDa (Teffet *al.*, 2006). CTLA-4 est composé d'un domaine en forme de V flanqué de deux régions hydrophobes, dont l'une présente une structure suggérant un ancrage membranaire (Brunet *et al.*, 1987).



Figure 1.3 : Structure tridimensionnelle du CTLA-4

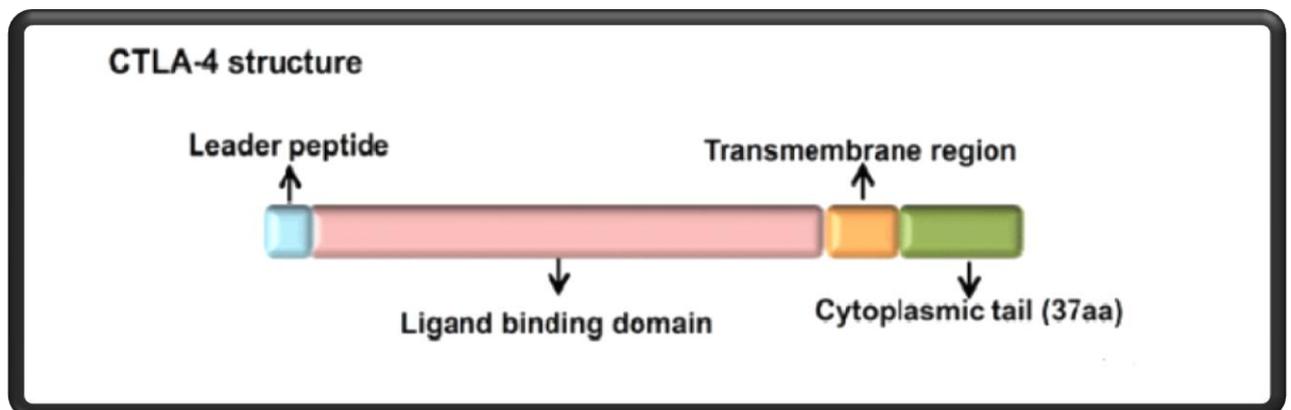


Figure 1.4 : la configuration structurale de CTLA-4

2.3. Rôle de CTLA-4 dans la réponse immunitaire

L'expression et la fonction biologique de CTLA-4 sont importantes pour la régulation

négative des réponses des cellules T. Par conséquent, l'élimination de CTLA-4 peut entraîner la rupture de la tolérance immunitaire et le développement de plusieurs maladies telles que les maladies auto-immunes(**Hosseini et al., 2020**).

CTLA-4 peut entrer en compétition avec son homologue stimulant CD28 pour se lier au ligand de CD80 et CD86, La co-stimulation de CD28 est nécessaire pour l'activation des lymphocytes T, tandis que CTLA-4 inhibe la réponse des lymphocytes T en s'opposant aux actions de la co-stimulation médiée par CD28(**Dumet et al., 2022**).

2.4. Implication du CTLA-4 dans le cancer colorectal

Le CTLA-4, qui est un inhibiteur du contrôle immunitaire, peut être présent dans les lymphocytes infiltrant la tumeur ainsi que dans les cellules du cancer colorectal (CCR) (**Derakhshani et al., 2021**). Ce point de contrôle immunitaire a la capacité d'affaiblir les réponses immunitaires anti-tumorales et de favoriser la croissance et la propagation des tumeurs (**Derakhshani et al., 2021**). Les médicaments qui ciblent CTLA-4, des inhibiteurs de point de contrôle immunitaire, contribuent à rétablir la réponse immunitaire contre les cellules cancéreuses colorectales (**Cancer.org**). Ces médicaments peuvent être administrés aux individus chez qui les cellules cancéreuses colorectales ont été diagnostiquées avec des changements génétiques spécifiques, tels qu'un niveau élevé d'instabilité microsatellite (MSI-H) ou des altérations dans l'un des gènes de réparation des appariements incorrects (MMR) (**Cancer.org**).

Le blocage du CTLA-4 agit par compétition avec la protéine CD28 pour se lier à leurs ligands B7 communs ou par un effet inhibiteur direct sur les molécules de signalisation intracellulaires (**Zhao et al., 2018**).

2.5. Inhibition de CTLA-4 dans le traitement du cancer colorectal

L'immunothérapie par blocage des anticorps CTLA-4 permet d'inverser la tolérance des lymphocytes T vis-à-vis des cellules cancéreuses en empêchant l'interaction de CTLA-4 avec ses ligands. Cette approche favorise l'activation et la cytotoxicité des lymphocytes T contre les cellules tumorales (**Chikuma, 2017**). Les anticorps tels que l'ipilimumab et le tremelimumab sont couramment utilisés pour traiter différentes formes de cancer car ils stimulent la réponse immunitaire antitumorale et améliorent la survie globale des patients atteints de cancer (**Zhao et al., 2018**).

Les inhibiteurs de points de contrôle immunitaire (ICIs) bloquent le récepteur CTLA-4, un point de contrôle immunitaire inhibiteur qui régule négativement la fonction des lymphocytes T anti-tumoraux. Ces inhibiteurs renforcent ainsi la réponse immunitaire antitumorale en augmentant

l'activation et la prolifération des lymphocytes T CD8⁺ et CD4⁺(**Fiegleet al., 2019**). Bien que les ICIs, comme les anticorps anti-CTLA-4, soient utilisés pour traiter le cancer colorectal (CRC), leur efficacité est limitée en raison d'une variété de facteurs qui restreignent leur action thérapeutique (**Shan et al., 2022**).

3. c-Src

La protéine c-Src est membre d'une famille de 11 kinases tyrosines non récepteurs connue sous le nom de kinases de la famille Src (SFKs) (**Jin, 2020**). Elle est présente dans toutes les cellules vertébrées, mais elle est largement surexprimée dans le cerveau, les ostéoclastes et les plaquettes. Les niveaux de cette protéine dans ces tissus sont de 5 à 200 fois plus élevés que dans la plupart des autres types cellulaires (Hanahan, 2022).

Dans les fibroblastes, c-Src est liée aux endosomes, aux membranes périnucléaires, aux vésicules sécrétoires et à la face cytoplasmique de la membrane plasmique où elle peut interagir avec une variété de facteurs de croissance, d'intégrines et de protéines G couplées aux récepteurs et servir d'intermédiaire essentiel dans la transduction de signal (**Jin, 2020**).

L'expression des niveaux élevés de c-Src dans les plaquettes (cellules anucléées) et dans les neurones (qui sont postmitotiques) indique que Src participe à des processus autres que la division cellulaire. Les Src kinases, y compris c-Src, catalysent la réaction suivante :



3.1. Structure de la protéine c-Src humaine

Le gène humain SRC prend la localisation chromosomique 20q11.23 et code pour 536 acides aminés(Roskoski, 2004).

La protéine c-Src est composé d'un domaine SH1, qui est la région catalytique de la kinase et qui contient le site de liaison du substrat et de l'ATP, ainsi que le site d'autophosphorylation (Tyr419 chez l'homme). Il possède également une courte queue C-terminale, qui contient le résidu de tyrosine impliqué dans la régulation négative (Tyr530 chez l'homme) du c-Src(**Boggon and Eck, 2004**), un domaine d'homologie 2 (SH2) qui a la capacité de se lier spécifiquement aux résidus de phosphotyrosine, en particulier dans le contexte des acides aminés acides(**Gul et al., 2023**), un domaine de l'homologie 3 (SH3), qui lie les motifs polyprolines(**Gul et al., 2023**), le domaine de l'homologie Src 4 (SH4) de la région amino terminale qui contient des séquences consensuelles pour la myristoylation et/ou la palmitoylation(**Resh, 1993a**), le domaine de l'homologie Src carboxyl-terminale 1, qui contient la région catalytique et a un queue court

carboxyle-terminale contenant la tyrosine réglementaire majeure (Seykora *et al.*, 2002).

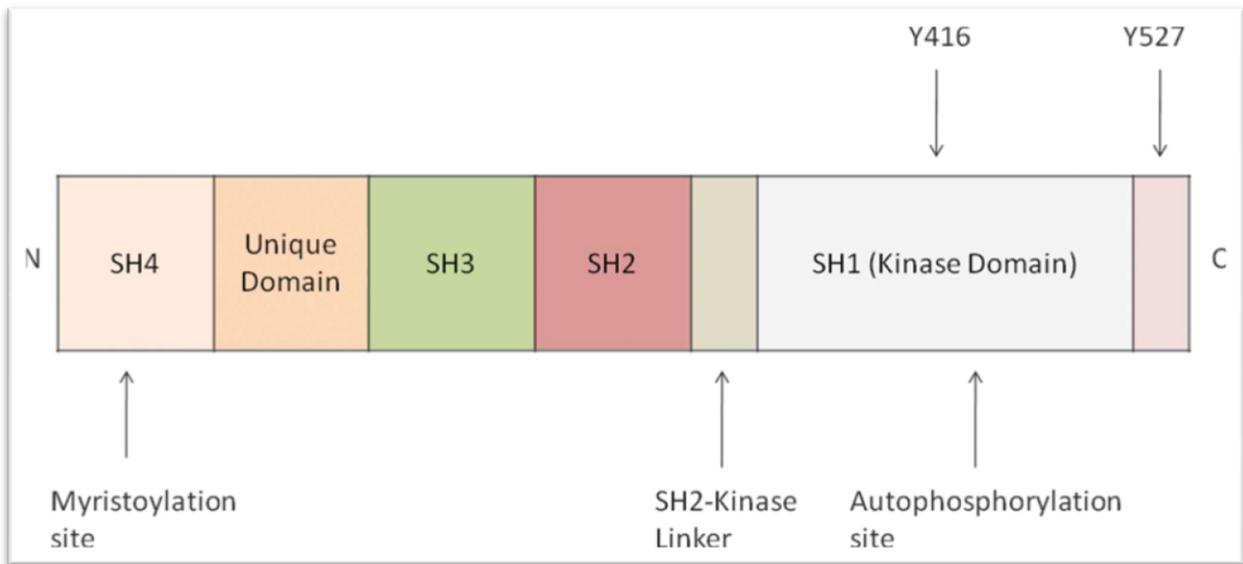


Figure 1.5 : Organisation structurale de la protéine c-Src (Chen *et al.*, 2015)



Figure 1.6 : la structure 3D du c-SRC (parmi nos résultats sur SWISS MODEL, Chimera)

3.2. Les voies de signalisation associées à la protéine c-Src

La protéine c-Src catalyse la phosphorylation de résidus tyrosine sur une grande variété de protéines, ce qui permet de réguler leur activité. Plusieurs récepteurs ont été identifiés comme étant des substrats potentiels pour c-Src, notamment les intégrines (Playford and Schaller, 2004), les récepteurs liés à la protéine G (Luttrell and Luttrell, 2004) et les récepteurs stéroïdes et tyrosine kinases tels que le récepteur du facteur de croissance dérivé des plaquettes (PDGFR), la famille du récepteur du facteur de croissance épidermique (EGF-R), le récepteur du facteur de

croissance fibroblastique (FGF-R), le récepteur du facteur de croissance 1 insulinique (IGF-1R), c-Met, le récepteur du facteur des cellules souches hématopoïétiques (LCR-1R) et le récepteur du facteur des cellules souches (SCF-R)(**Shupnik, 2004**).

Bien que certains intermédiaires, tels que Ral-GTPase (**Martin et al., 2011**), aient été impliqués dans l'activation de Src kinase, les récepteurs mentionnés précédemment jouent également un rôle important. Ils se lient aux domaines SH2 ou SH3 de Src et perturbent les interactions intramoléculaires, ce qui active Src kinase. D'autres facteurs peuvent également contribuer à cette activation incluant des facteurs de transcription comme Stat3(**Playford and Schaller, 2004**); des protéines adaptatrices telles que Shc activent la cascade de signalisation Ras/Raf/Erk(**Lennartsson et al., 1999**); d'autres kinases telles que PI3Ks(Papadatos-Pastos *et al.*, 2015), MAPK ou Akt(**Fang and Richardson, 2005**); des protéines membranaires impliquées dans la communication cellulaire telle que connexin43 (**Sorgen et al., 2004**)ou pannexin1(**DeLalio et al., 2019**); des composants du cytosquelette comme FAK, p130 CAS, cortactine, paxilline ou encore p190 Rho-GTPase-activatingprotein (GAP) ainsi qu'autres enzymes métaboliques(**Guarino, 2010**).

Les phosphorylations directes induites par c-Src peuvent affecter des protéines cibles et des voies de signalisation secondaires, telles que la cycline D1 ou HIF-1 α (**Malek et al., 2002**). En tant qu'effecteur et régulateur d'une multitude de récepteurs, c-Src facilite la diaphonie entre différentes voies de signalisation. Ainsi, cette protéine kinase est un nœud de communication dans un réseau complexe de protéines en interaction, qui peut influencer divers événements cellulaires tels que la prolifération, la différenciation cellulaire, la survie cellulaire, la migration cellulaire, l'organisation du cytosquelette, l'adhésion cellulaire ainsi que les processus métaboliques(**Parsons and Parsons, 2004**).

3.3.Implication de Src dans le cancer colorectal

La famille de kinases Src (SFK) joue un rôle essentiel dans le développement du CCR et sa progression (**Jin, 2020**), L'activation des SFK favorise la multiplication des cellules cancéreuses colorectales, leur propagation vers d'autres organes et leur résistance aux traitements chimiothérapeutiques(**Jin, 2020**), En conséquence, l'inhibition de Src pourrait être bénéfique dans le traitement du cancer colorectal. Cependant, il est important de noter que l'inhibition de Src peut également affaiblir les réponses immunitaires essentielles à l'élimination des cellules cancéreuses (**Chen et al., 2014**).

3.4. Signalisation de c-Src dans le cancer colorectal

Initialement, la participation de c-Src dans la signalisation oncogénique a été établie suite à la découverte du gène transformant v-SRC, présent dans le rétrovirus aviaire RSV. Bien que les mutations spécifiques du gène SRC soient rares, une surexpression et une hyperactivité du SRC ont été démontrées dans plusieurs types de cancers, y compris le cancer colorectal (Irby and Yeatman, 2000).

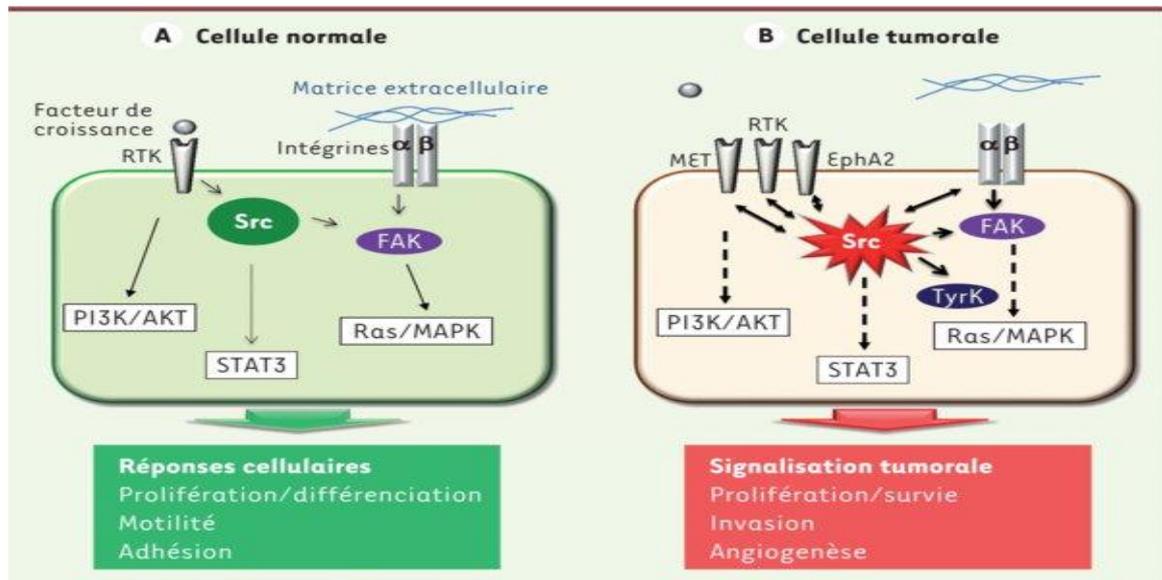


Figure 1.7 : Signalisation liée à la protéine SRC dans une cellule normale et une cellule cancéreuse (Sirvent *et al.*, 2014)

A la liaison du facteur de croissance au récepteur RTK (récepteur à tyrosine kinase) et les intégrines une succession d'événements s'installe ; dans le cadre de la signalisation normale c-Src provoque l'excitation de RAF puis une cascade de signalisation PI3K/AKT, STAT3, Ras/MAPK pour aboutir soit la différenciation, prolifération, adhésion... ; contrairement au stade oncogénique qui est caractérisé par hypersensibilisation des gènes SRC finalement une surproduction des protéines c-Src dont les résultats une prolifération cellulaire hiérarchique non contrôlée provoque un état de métastase (Sirvent *et al.* 2010)

3.5. Inhibition de c-Src dans le traitement du cancer colorectal

L'inhibition de Src est une cible thérapeutique potentielle pour le traitement du CCR.

Des études ont révélé que l'efficacité du traitement du cancer colorectal pourrait être améliorée en combinant l'inhibition des enzymes MEK et Src. Des expériences menées sur des cultures organotypiques *ex vivo* ont permis d'identifier des réponses spécifiques chez les patients atteints de CCR à cette combinaison d'inhibiteurs de MEK et de Src (Gavert *et al.*, 2022).

De plus, il a été démontré que l'activation de la tyrosine kinase c-Src est impliquée en tant que mécanisme de résistance à la chimiothérapie dans le cas du cancer colorectal métastatique (Ruiz-Saenz *et al.*, 2023).

4. Polymorphisme nucléotidique

4.1. Généralités

Les SNPs (single nucleotide polymorphism) correspondent à des variations génétiques se manifestant par la différence d'une paire de bases dans l'ADN entre différents individus (Morris et al., 2015).

La plupart de ces anomalies sont délétères par l'action suppressive du système immunitaire mais d'autres d'entre eux persiste et provoque la biodiversité donc ils ont pas seulement des conséquences néfastes mais aussi positifs (Ng and Henikoff, 2006)

Une analyse a révélé la présence de 45 SNP associés à une augmentation moyenne de 1,14 fois du risque de cancer colorectal par allèle, avec une plage allant de 1,05 à 1,53. De manière significative, il a été observé que les personnes présentant le plus grand nombre d'allèles à risque, soit celles se trouvant dans le quintile supérieur, ont un risque de cancer colorectal 1,81 fois supérieur à celui de la population moyenne (Jenkins et al., 2016).

La classification des polymorphismes nucléotidiques (SNP) peut se baser sur leur emplacement sur les chromosomes ainsi que leur association avec le risque de cancer colorectal (CRC). Des exemples incluent des SNP situés sur les chromosomes 1q41, 3q26.2, 6p21, 8q23.3, 8q24.21, 10p14, 11q13, 11q23.1, 12q13.13, 14q22, 14q22.2, 15q13.3, 16q22.1, 18q21.1, 19q13.11, 20p12, 20p12.3, 20q13.33 et Xp22, qui sont associés au risque de développer un cancer colorectal (Morris et al., 2015).

4.2. Classification des SNP

Les SNP ont un pouvoir de se positionner dans les séquences codantes ou non codantes des gènes même dans les parties intergéniques (entre les gènes) ; et selon ça on distingue :

4.2.1. SNPs des régions non-codantes

S'appelle aussi SNP synonyme ou silencieuse puisqu'elle s'entraîne pas une variation d'acides aminés (McCarthy et al., 2017) ; mais elle participe dans la structuration de la chromatine (Schaub et al., 2012), peuvent toucher la conformation de l'ARNm (Lu et al., 2015) et aussi la régulation des gènes (Molineris et al., 2013).

4.2.2. SNP des régions codantes

SNP synonyme font la substitution d'un codon par un autre donnant le même AA cela signifie donc que ni la fonction ni la conformation est délétères (Kimchi-Sarfaty et al., 2007) ;

mais d'autres études ont montré que les SNP synonymes exigent un effet sur l'épissage de l'ARNm (Cartegni *et al.*, 2002) et même sa stabilité (Nackley *et al.*, 2006) ainsi que le repliement et la traduction des protéines (Kimchi-Sarfaty *et al.*, 2007).

SNP non-Synonyme

Faux-sens c'est une transversion (changement purine/pyrimidine ou l'inverse) ou une transition (changement pyrimidine/pyrimidine ou purine/purine) qui booste la modification des AA donc changement de la protéine et puis une maladie

non-sens c'est par la génération d'un codon stop dit prématuré ou un codon non-sens dans l'ARNm transcrit et le produit protéique obtenu est incomplet souvent non fonctionnelle (Cordovado *et al.*, 2012)

4.3. Évaluation des facteurs de risque associés aux SNP

Les risques relatifs aux SNP (Polymorphismes Mononucléotidiques) sont influencés par leur impact sur la fonction des gènes et la régulation de l'expression des gènes. Certains SNP ont été liés à une probabilité accrue de développer diverses pathologies, y compris le cancer colorectal (Korzeniewski *et al.*, 2013).

Un exemple de cela est le SNP rs17042479, qui a été corrélé à un stade plus avancé de diagnostic du cancer colorectal, suggérant un potentiel de développement plus agressif de la maladie chez les patients qui portent ce SNP (Olsson *et al.*, 2022).

Des études ultérieures ont identifié des SNP liés à une probabilité accrue de développer un cancer colorectal en utilisant des données génomiques et épigénétiques comme sources d'information. (Tak *et al.*, 2014).

4.4. Intérêt de l'étude du polymorphisme nucléotidique

Les polymorphismes mononucléotidiques (SNP) sont largement utilisés dans différents domaines tels que la médecine, la génétique et la pharmacologie, en raison de leur polyvalence. Voici quelques-unes des principales applications sur les SNPs :

- **Marqueurs biologiques** Les SNP peuvent être utilisés comme marqueurs biologiques pour la localisation de gènes liés aux maladies et pour l'identification de variations génétiques qui contribuent à certaines conditions médicales (MedlinePlus Genetics).

- **Prédiction de la réponse aux médicaments** Les SNP peuvent être utilisés pour prédire la réponse individuelle à certains médicaments, ce qui facilite l'adaptation des traitements en fonction du profil génétique du patient (imWechsel.com).

- Détection des mutations Les méthodes basées sur la PCR (réaction en chaîne par polymérase) sont fréquemment employées pour détecter les SNP et les mutations, offrant des indications sur la susceptibilité aux maladies, la pathogenèse des maladies et l'efficacité de médicaments spécifiques(Matsuda, 2017).

- Recherche en génétique Les SNP sont utilisés en recherche pour étudier la diversité génétique et la structure des populations, et pour identifier les gènes impliqués dans les troubles multifactoriels(Syvänen, 2001).

4.5. Rôle potentiel des SNP dans la prédisposition et le diagnostic du cancer colorectal

De nombreuses études ont identifié des SNP associés à un risque accru de cancer colorectal. Par exemple, une étude menée au sein de l'UK Biobank a examiné la capacité de 45 SNP connus pour leur implication dans la susceptibilité au cancer colorectal à prédire le risque de développer cette maladie.(Gafniet *al.*, 2021), Des études ont évalué les interactions entre les SNP et le risque de résultats cliniques chez les patients atteints de cancer colorectal en utilisant des techniques de réduction de la dimensionnalité telles que la réduction multifactorielle (MDR)(Curtis *et al.*, 2022).

4.6. Rôle des polymorphismes nucléotidiques dans le traitement du cancer colorectal

Plusieurs polymorphismes génétiques ont été étudiés pour leur application clinique dans le traitement du cancer colorectal. L'UGT1A1 (UDP-glucuronosyltransférase 1A1), une enzyme impliquée dans la biotransformation des médicaments, est un exemple de gène présentant des polymorphismes qui peuvent affecter la réponse au traitement chez les patients atteints de cancer colorectal(Aielloet *al.*, 2011).

4.7. Analyse *in silico* des SNP

4.7.1. Définitions

La recherche *in silico* de SNP implique l'utilisation d'outils de bioinformatique pour prédire les conséquences des variations nucléotidiques sur la structure et la fonction des protéines. Ces techniques permettent de détecter les SNP potentiellement nuisibles qui pourraient altérer significativement la fonctionnalité des protéines et, par conséquent, affecter la santé et la réponse aux traitements médicamenteux (Mah *et al.*, 2011).

Les méthodes informatiques employées pour prédire les effets phénotypiques d'un SNP non-

synonyme (nsSNP) reposent sur l'analyse de diverses données telles que l'environnement structural et les alignements de séquences multiples. Ces outils permettent d'évaluer si le nsSNP pourrait avoir un impact sur la fonctionnalité de la protéine concernée (**Dabhi and Mistry, 2014**).

4.7.2. Importances des analyses *in silico* des SNP dans les maladies humaines

L'utilisation de l'analyse *in silico* des SNP est avantageuse pour identifier les SNP les plus prometteurs dans le cadre de la recherche de médicaments et pour réduire le temps et les efforts nécessaires pour mener des expériences en laboratoire. Cette approche permet de cibler plus efficacement les SNP susceptibles d'affecter la fonctionnalité des protéines et d'optimiser ainsi la recherche de molécules thérapeutiques (**Mah et al., 2011**).

5. Modifications post-traductionnels

Les modifications post-traductionnelles (MPT) sont des altérations chimiques qui surviennent sur les protéines après leur synthèse. Les MPT jouent un rôle essentiel dans la régulation de la fonction protéique, de la stabilité, de la localisation et des interactions avec d'autres protéines (Santos and Lindner, 2017)

5.1. Effets des SNPs sur les modifications post traductionnelles

Les PTM sont responsable de la variété globale des protéines au sein de l'organisme après ou pendant la traduction, et les SNP sont la forme la plus courante de variations trouvées dans le génome humain. Des études ont montré que nsSNP au site PTM modifie l'AA qui convient à une chaîne des protéines traduite et il va être le responsable d'un fonctionnement anormal de la protéines aboutissant à l'apparition de la maladie (**Gulzar et al., 2017**) comme la maladie d'Alzheimer et la maladie de Parkinson (**Gulzar et al., 2017**)

5.2. Prédiction *in silico* des modifications post traductionnelles

Les protéines au cours de synthetisation sont subissent des modifications (PTM) qui peut être une phosphorylation, ubiquitination, méthylation, nitrosylation pour le but de perfectionner les protéines et les joindrais à la circulation (**Vatsa, 2022**)

A l'avancement de la science il est possible de détecter le type de modification que la protéines à subi en utilisant des outils computationnelles qui fonctionne par des approches mathématiques et chimiques (**Vatsa, 2022**). La Prédiction *in silico* des modifications post traductionnelles peut être fondées en trois catégories : fondées sur la séquence, fondées sur la

structure et fondées sur la physique (ed. **WR Strohl, and LM Strohl, 2012**)

5.2.1. La phosphorylation :

La phosphorylation est une modification post-traductionnelle qui implique l'ajout d'un groupe phosphate à une protéine ou à une autre molécule organique. Cette modification peut avoir différents effets sur la fonction de la protéine, notamment la régulation de l'activité enzymatique par activation ou désactivation en modifiant sa structure tridimensionnelle ou en régulant son interaction avec d'autres protéines, les changements dans les interactions protéine-protéine et les modifications de la localisation protéique (**Wilson, 2017**).

Des recherches ont démontré que la phosphorylation peut également jouer un rôle dans la pathogenèse de diverses maladies, notamment le cancer et le diabète. Par exemple, la dérégulation de la phosphorylation d'Akt/protéine kinase B (PKB) est impliquée dans la pathogenèse du cancer et du diabète (**Sarbassovet *al.*, 2005**).

5.2.2. La Méthylation

C'est la modification enzymatique par l'ajout du groupement méthyle de S-adénosylméthionine, La méthylation est une action irréversible d'addition des esters méthyliques sur les groupements et qui peut modifier la fonction de la protéine

D'autre part, les réactions de méthylation survenant sur les atomes d'azote dans les positions N-terminal et side-chain sont généralement irréversibles. Ces réactions créent de nouveaux types de résidus d'acides aminés et peuvent élargir le répertoire de la chimie qu'une protéine peut effectuer(**Grillo and Colombatto, 2005**)

5.2.3. L'ubiquitination

L'ubiquitination est une modification post-traductionnelle qui implique l'ajout d'ubiquitine, une petite protéine, à une protéine cible. Cette modification peut entraîner différentes conséquences sur la fonction de la protéine, telles que la dégradation, des modifications dans les interactions avec d'autres protéines et des changements de localisation (**Ray *et al.*, 2020**).

L'ubiquitination joue un rôle crucial dans de nombreux processus cellulaires, notamment la dégradation des protéines, la transcription, les interactions protéine-protéine, l'assemblage des signalosomes et la localisation subcellulaire (**Song and Li, 2021**). De plus, elle participe à la régulation du système immunitaire et à l'activation de facteurs de transcription qui régulent la prolifération cellulaire, l'apoptose et l'infection virale (**Song and Li, 2021; Roberts *et al.*, 2022**).

5.2.4. La carbonylation

C'est l'addition des dérivés carbonyles (*reactivecarbonyl compounds* ou RCC) sur les protéines (**Grimsrud et al., 2008**) qui nécessite à leurs formation des mécanisme oxydatifs (**Wehr and Levine, 2013**) le mieux connu est la peroxydation lipidique.

Les RCC peuvent se fixer directement sur les chaînes latérales des acides aminés (surtout des résidus de lysine, cystéine et histidine) *via* la formation d'une base de Schiff (**Irla et al., 2008**) ou suite à une addition de Michael (**Irla et al., 2008**). Comme la plupart des RCC dérivent de la peroxydation lipidique, les produits terminaux issus de la réaction de carbonylation sont fréquemment appelés *advancedlipoxidation end-products* (ALE). Les ALE les plus couramment décrits sont la 4-hydroxy-2-nonéal-lysine, la malondialdéhyde-lysine, la N^ε-carboxyéthyl-lysine et des composés plus complexes comme des liaisons croisées de type pyrrole (**Sacks et al., 2002**)

5.2.5. S-nitrosylation

La S-nitrosylation est un procédé chimique général pour l'introduction d'un groupe nitro (-NO₂) dans un composé chimique (**Turko and Murad, 2002**), il constitue une grande partie de l'influence omniprésente de l'oxyde nitrique sur la transduction des signaux cellulaires et l'accumulation de preuves indique des rôles importants pour la S-nitrosylation à la fois dans la physiologie normale et dans un large spectre de maladies humaines (**Foster et al., 2009**).

5.2.6. tyrosine nitration

Dans le cas des protéines, plusieurs acides aminés sont préférentiellement nitrés, tels que la tyrosine (Y), le tryptophane (W), la cystéine (C) et la méthionine (M) (**Turko and Murad, 2002**). Cependant, la plupart des études concernent la nitration de tyrosine, qui consiste à ajouter un groupe nitro à l'un des deux ortho-carbones équivalents du cycle aromatique des résidus de tyrosine (**Turko and Murad, 2002**). Cette modification provoque la transformation de la tyrosine en une fraction de nitrotyrosine hydrophile à charge négative.

la nitration de la tyrosine est considérée comme un processus sélectif, et les protéines ont habituellement environ 3 à 4 mol% de Tyr, mais seulement une ou deux de ces tyrosines peuvent devenir préférentiellement nitrées (**Corpas et al., 2009**)

6. Problématique

Comment les variations génétiques identifiées dans les protéines CTLA-4 et SRC peuvent affecter leur structure tridimensionnelle et leur interaction avec d'autres protéines impliquées dans le développement du cancer colorectal ?

En 2022, Megnaïhayet et Boukenadel Amel ont mené une recherche, sous la direction du Dr. Hadjidj Zeyneb, pour étudier l'impact des variations génétiques dans les gènes CTLA-4 et SRC sur la fonction immunitaire normale et le risque de cancer. Des outils d'analyse *in silico* ont été utilisés pour identifier les SNP potentiellement délétères et prédire leur impact sur les propriétés de la protéine CTLA-4 et d'autres protéines liées, en particulier la protéine SRC. Cette étude a identifié 4 variations probablement délétères dans CTLA-4 (T72I, N145H, N145S, T147A) et 14 dans SRC (P136R, D407N, T460K, P487L, R509W, P510A, P210S, G173V, F223L, A296S, A330D, M344L, T432M, M498L). Ces variations pourraient être des candidats importants pour le cancer colorectal.

Dans ce travail, l'étude se poursuit pour approfondir la compréhension de l'impact de la variation identifiée en utilisant des approches *in silico* pour la prédiction des modifications post traductionnelles et la modélisation des structures tridimensionnelles des protéines mutées.

- **Objectif**

Explorer les effets des variations génétiques sur la structure tridimensionnelle et les modifications post traductionnelles dans les protéines CTLA-4 et Src.

- **But**

Identifier les mutations responsables des variations de structure et des modifications post-traductionnelles dans les protéines CTLA-4 et Src, qui sont potentiellement impliquées dans le développement du cancer colorectal.

CHAPITRE II :
MATÉRIEL ET
MÉTODES

II- Chapitre II : Matériel et méthodes

1. Plan de l'étude

Les travaux de mémoire de fin d'études réalisés en 2022 par Boukenadel Amel et Megnafi Hayet, sous la direction du Dr. Hadjidj Zeyneb, ont permis d'identifier quatre variations potentiellement délétères dans CTLA-4 (T72I, N145H, N145S et T147A) ainsi que quatorze autres dans SRC (P136R, D407N, T460K, P487L, R509W, P510A, P210S, G173V, F223L, A296S, A330D, M344L, T432M, M498L).

Ce travail se poursuit actuellement avec deux axes principaux : la recherche des modifications post-traductionnelles engendrées ou supprimées par ces variations et la modélisation tridimensionnelle des variant protéiques afin d'identifier les différences structurelles susceptibles d'être causées par ces variations.

Ces travaux ont été menés au sein du Laboratoire Biologie Moléculaire Appliquée et Immunologie dirigé par le Professeur ARIBI Mourad.

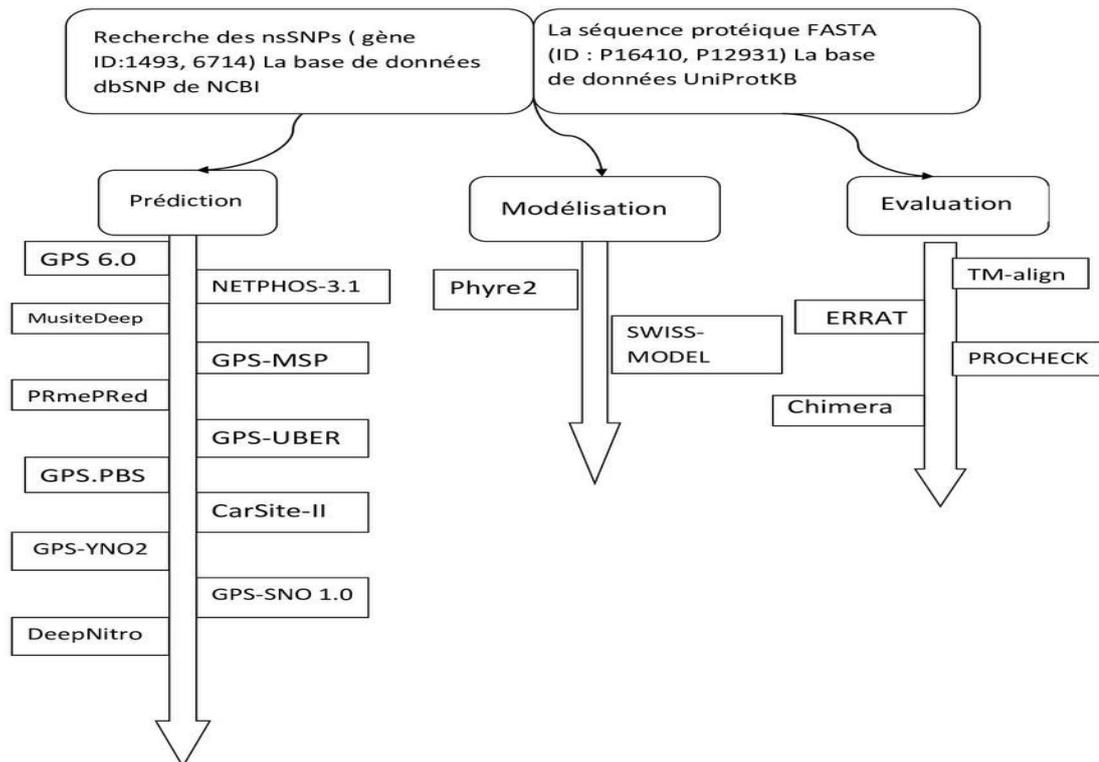


Figure 2.1 : plan de l'étude.

2. Collecte de données et sélection de séquences de protéines

Les séquences primaires des protéines, correspondant aux gènes CTLA-4 et SRC, ont été obtenues à partir de la base de données UniProt (Sherry *et al.*, 2001; EuropeanMolecularBiologyLaboratory, 2017). Les séquences en acides aminés FASTA des protéines sont disponibles sous les numéros d'accès Uniprot P16410 pour CTLA-4 et P12931 pour SRC.

Afin d'obtenir les séquences mutées des protéines étudiées, nous avons tout d'abord obtenu les séquences primaires à partir de la base de données UniProt (<https://www.uniprot.org>). Ensuite, ces séquences ont été modifiées pour inclure les mutations spécifiques que nous souhaitons étudier. Les acides aminés variant ont ainsi été substitués par leur équivalent muté dans chaque protéine.

Figure 2.3 : exemple d'une Recherche des séquences sur **UniProt**

3. Prédiction des modifications post traductionnelles

Différents outils bio-informatiques, disponibles en ligne, ont été utilisés pour effectuer les prédictions des modifications post-traductionnelles. Ce sont des logiciels de prédiction des modifications qui utilisent différents algorithmes bioinformatiques pour analyser ces séquences et identifier les sites potentiels de modification post-traductionnelle. Ces prédictions peuvent être basées sur l'analyse de motifs spécifiques dans la séquence, sur une comparaison avec une base de données connue ou encore sur un modèle statistique entraîné à partir d'un ensemble de données connues.

Les séquences des protéines CTLA-4 et c-Src sauvages et muté sont été soumis aux logiciels de prédiction en format FASTA.

3.1. Prédictions des sites de phosphorylation

3.1.1. GPS 6.0

GPS 6.0 (Group-based Prediction System Version 6.0) (<https://gps.biocuckoo.cn/>) est un outil dédié à la prédiction des modifications post-traductionnelles (PTM) des protéines, en se concentrant particulièrement sur les sites potentiels de phosphorylation des résidus de sérine, thréonine et tyrosine. Il offre la possibilité de choisir les protéines kinases responsables des phosphorylations potentielles et de personnaliser les paramètres de prédiction en fonction des besoins spécifiques, tels que le seuil de score à partir duquel un site modifié sera considéré

comme potentiellement vrai ou faux. Le résultat est fourni sous forme de liste de sites de phosphorylation potentiels avec leurs scores de confiance associés(Chen et al. 2023).

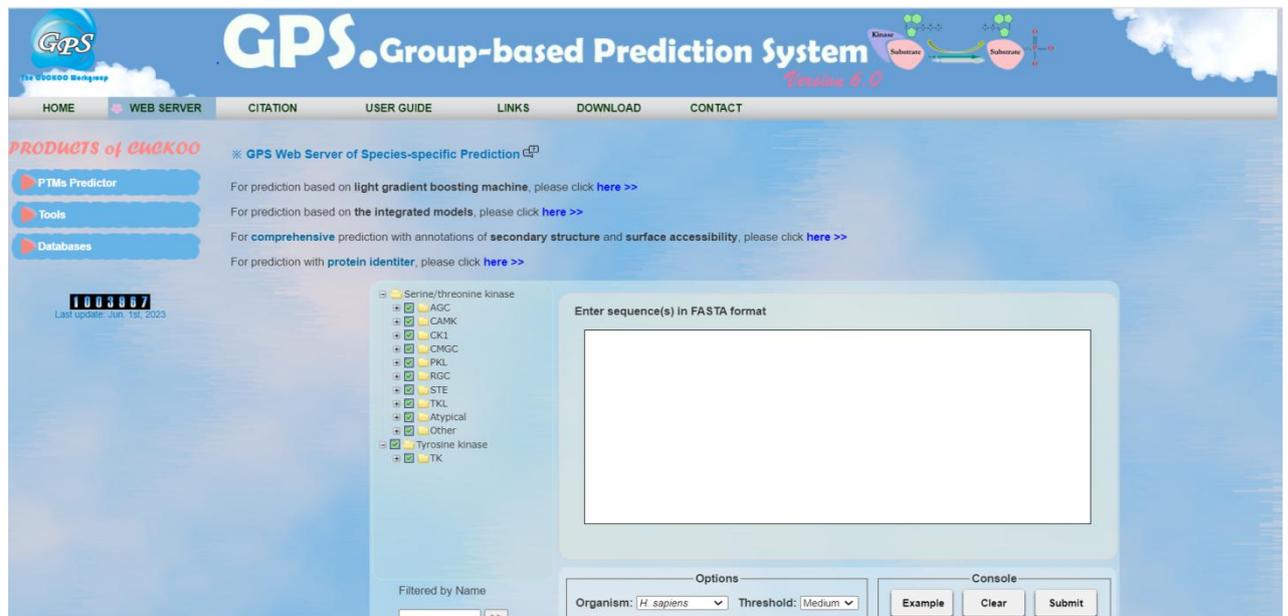


Figure 2.3 : Prédictions des sites de phosphorylation par GPS 6.0

3.1.2. NETPHOS

NETPHOS(<https://services.healthtech.dtu.dk/services/NetPhos-3.1>) est un logiciel basé sur des réseaux neuronaux qui permet de prédire les sites potentiels de phosphorylation des résidus de sérine, thréonine ou tyrosine dans les séquences protéiques (Biswas, Noman, et Sikder 2010). Il permet de générer une liste de sites de phosphorylation prédits, accompagnée de leurs scores de confiance respectifs.

jobid=6492301A00000FB2305C7A19&wait=20 Server Output - DTU Health Tech

```
>sp_P12931_SRC_HUMAN 536 amino acids
#
# netphos-3.1b prediction results
#
# Sequence      # x Context      Score Kinase Answer
#-----
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.901 PKC YES
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.458 GSK3 .
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.430 CaM-II .
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.419 CKI .
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.406 cdc2 .
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.349 DNAPK .
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.348 PKG .
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.327 p38MAPK .
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.322 RSK .
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.274 CKII .
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.266 ATM .
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.200 cdk5 .
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.136 PKA .
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.116 PKB .
# sp_P12931_SRC_HUMAN 3 S --MGSNKS 0.008 unsp .
#
# sp_P12931_SRC_HUMAN 6 S GSNKSPKD 0.542 unsp YES
# sp_P12931_SRC_HUMAN 6 S GSNKSPKD 0.538 PKC YES
# sp_P12931_SRC_HUMAN 6 S GSNKSPKD 0.442 CaM-II .
# sp_P12931_SRC_HUMAN 6 S GSNKSPKD 0.441 cdc2 .
# sp_P12931_SRC_HUMAN 6 S GSNKSPKD 0.438 GSK3 .
# sp_P12931_SRC_HUMAN 6 S GSNKSPKD 0.391 CKI .
# sp_P12931_SRC_HUMAN 6 S GSNKSPKD 0.390 RSK .
# sp_P12931_SRC_HUMAN 6 S GSNKSPKD 0.358 DNAPK .
# sp_P12931_SRC_HUMAN 6 S GSNKSPKD 0.316 p38MAPK .
# sp_P12931_SRC_HUMAN 6 S GSNKSPKD 0.305 CKII .
# sp_P12931_SRC_HUMAN 6 S GSNKSPKD 0.288 PKG .
# sp_P12931_SRC_HUMAN 6 S GSNKSPKD 0.275 ATM .
# sp_P12931_SRC_HUMAN 6 S GSNKSPKD 0.183 cdk5 .
```

Activer Windows
Accédez aux paramètres de l'ordinateur pour

Figure 2.4 :Prédictions des sites de phosphorylation par Netphos 3.1

MusiteDeep (<https://www.musite.net/>) a été utilisé pour prédire le site de phosphorylation, de méthylations et d'ubiquitinations. C'est un outil qui permet de prédire les sites de phosphorylation généraux et spécifiques aux kinases. Il permet d'analyser les données de séquence brute en entrée pour identifier avec précision les sites potentiels modifiés. Cette plateforme utilise des réseaux neuronaux convolutifs combinés à un mécanisme d'attention bidimensionnelle unique afin d'améliorer la qualité globale des prédictions.

MusiteDeep offre une amélioration relative significative de plus de 50 % dans l'évaluation Aire Sous la Courbe (AUC) précision-rappel pour la prédiction générale des sites de phosphorylation par rapport à d'autres outils bien connus. De plus, il produit également des résultats compétitifs dans la prédiction spécifique aux kinases sur les données référentielles (Wang et al. 2017).

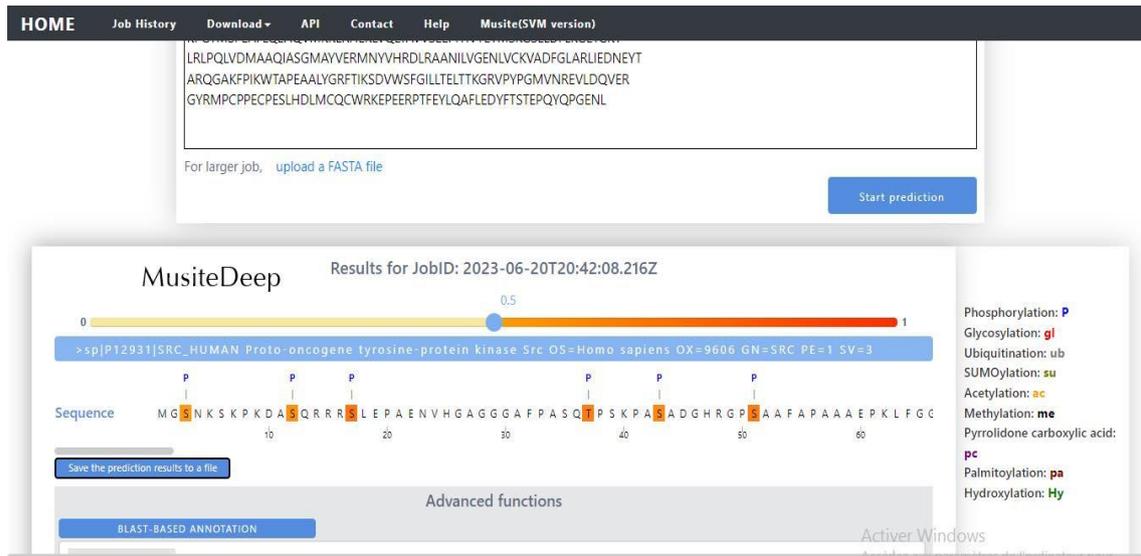


Figure 2.5 : Prédictions des sites de phosphorylation, de méthylation et d’ubiquitinations par MusiteDeep

3.2. Prédiction des sites de méthylation

3.2.1. GPS-MSP

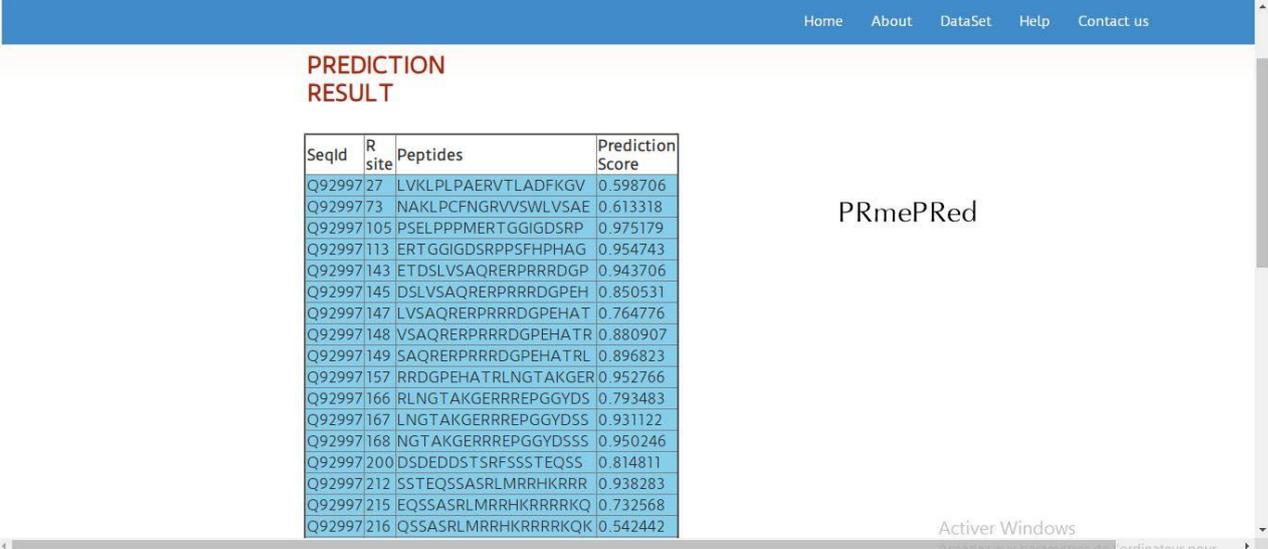
L’outil GPS-MSP (<http://msp.biocuckoo.org/>) a été utilisé pour prédire les sites de méthylation. Cet outil est spécifiquement conçu pour prédire les sites de méthylation de la lysine et de l’arginine dans les protéines. Il utilise une approche basée sur le groupe, qui repose sur l’identification de motifs de séquences conservées dans des groupes de protéines similaires, afin de prédire les sites de méthylation (Wankun Deng, Yongbo Wang, and Dr. YuXue).



Figure 2.6 : Prédiction des sites de méthylation par GPS-MSP

3.2.2. PRmkePRed

L'outil PRmePRed(<http://bioinfo.icgeb.res.in/PRmePRed/>) a été utilisé pour prédire les sites de méthylation. C'est un outil de prédiction basé sur les machines à vecteurs de support (SVM) pour prédire les sites de méthylation de l'arginine dans les protéines.



PRmePRed

SeqId	R site	Peptides	Prediction Score
Q92997	27	LVKLLPLPAERVTLADFKGV	0.598706
Q92997	73	NAKLPCFNQGRVSVLVSVAE	0.613318
Q92997	105	PSELPPMERTGGIGDSRP	0.975179
Q92997	113	ERTGGIGDSRPPSFHPHAG	0.954743
Q92997	143	ETDSLVAQRERPRRRDGP	0.943706
Q92997	145	DSLVAQRERPRRRDGP	0.850531
Q92997	147	LVAQRERPRRRDGP	0.764776
Q92997	148	VSAQRERPRRRDGP	0.880907
Q92997	149	SAQRERPRRRDGP	0.896823
Q92997	157	RRDGPEHATRLNGTAKGER	0.952766
Q92997	166	RLNGTAKGERRRREP	0.793483
Q92997	167	LNGTAKGERRRREP	0.931122
Q92997	168	NGTAKGERRRREP	0.950246
Q92997	200	DSDEDDSTSRFSSSTEQSS	0.814811
Q92997	212	SSTEQSSASRLMRRHKRRR	0.938283
Q92997	215	EQSSASRLMRRHKRRR	0.732568
Q92997	216	QSSASRLMRRHKRRR	0.542442

Figure 2.7 : Prédiction des sites de méthylation par PRmePRed

3.3. Prédiction des sites d'ubiquitination

L'outil GPS-UBER(GPS-Ubiquitin-protéine ligase Enzymes-substrate relation prédiction)<http://gpsuber.biocuckoo.cn/> a été utilisé pour prédire les sites d'ubiquitination.



GPS-UBER
Ubiquitin-protein ligase enzymes-substrate relationship prediction

HOME WEB SERVER USER GUIDE LINKS CONTACT

products of cuckoo * GPS-Uber Web Service

Download Return

Result has 9 items!

ID	Position	Code	E3 enzymes	Peptide	Score	Cutoff
sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3	106	K	General	ESRTETDLSEFKGERLQIVNN	0.2669	0.2655
sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3	155	K	General	SIQAEWYFGKITRRESEKLL	0.2960	0.2655
sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3	252	K	General	HRLTVCPTSPKPTGGLAEDA	0.6137	0.2655

Figure 2.8 : Prédiction des sites d'ubiquitination par GPS-UBER

3.4. Prédiction des domaines de liaison aux phosphoprotéines

L'outil GPS-PBS<http://pbs.biocuckoo.cn/index.php> a été utilisé pour prédire les domaines de

liaison aux phosphoprotéines. C'est un cadre d'apprentissage profond pour prédire les sites de phosphorylation qui interagissent spécifiquement avec des domaines de liaison aux phosphoprotéines (Guo et al. 2020).

The screenshot shows the GPS-PBS web service interface. At the top, there is a logo for GPS (The CUCKOO Workgroup) and the title "GPS-PBS. Prediction of PPBD-specific binding p-sites" with "Version 1.0" on the right. Below the title, there are navigation links: HOME, WEB SERVER, USER GUIDE, LINKS, and CONTACT. The main content area is titled "GPS-PBS Web Service" and contains a "Download" and "Return" button. Below the buttons, it says "Result has 3 items!". A table displays the results:

ID	Position	Code	PPBD	Peptide	Score	Cutoff	Source	Logo
RAF1_HUMAN	257	S	14-3-3/YWHA8	SLSQRGRSTSTFNVH	0.1863	0.1041	Pred.	
RAF1_HUMAN	259	S	14-3-3/YWHA8	SQRGRSTSTFNVHMV	0.4709	0.1041	Exp.	

Figure 2.9 : Prédiction des domaines de liaison aux phosphoprotéines par GPS.PBS

3.5. Prédiction des sites de carbonylation

L'outil CarSite-II (<http://47.100.136.41:8081/>) a été utilisé pour prédire les sites de carbonylation sur des résidus d'acides aminés K, P, R et T dans les protéines d'intérêt. Il possède un algorithme de classification intégré, basé sur la technique de sous-échantillonnage par similarité K-means et de suréchantillonnage synthétique des données minoritaires. Cette méthode permet l'identification efficace des sites carbonylés dans les protéines étudiées.

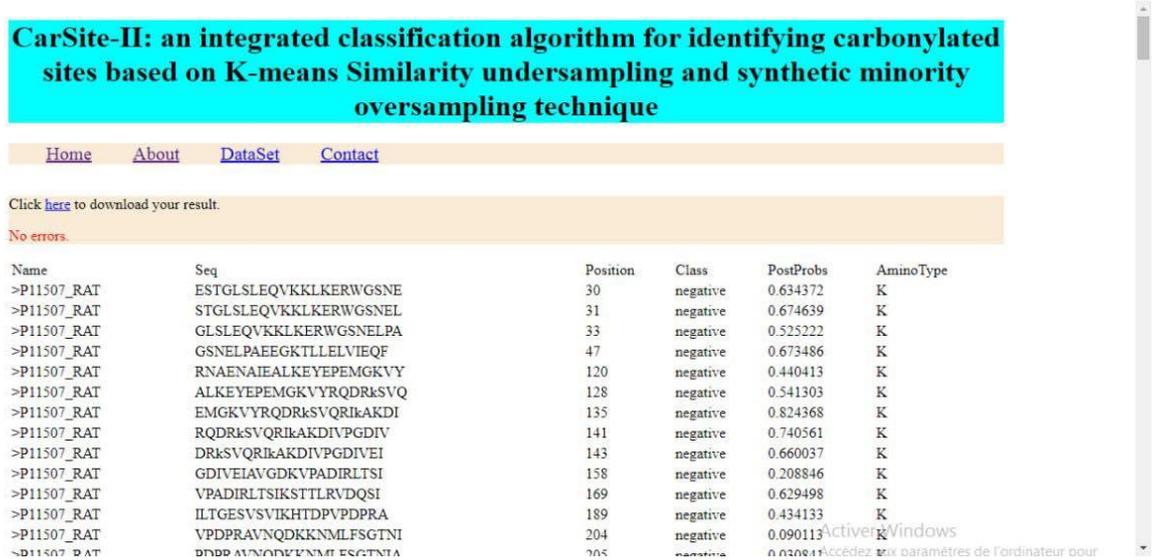


Figure 2.10 : Prédiction des sites de carbonylation par CarSite-II

3.6. Prédiction des sites de nitration des tyrosines

3.6.1. GPS-YNO2

GPS-YNO2 (<http://yno2.biocuckoo.org/>) a été utilisé pour prédire les sites de nitration des tyrosines. Ce programme a une précision prometteuse de 76,51 %, ainsi qu'une sensibilité et une spécificité respectives de 50,09 % et 80,18 %, obtenues lors d'une validation leave-one-out rigoureuse (Liu et al. 2011).

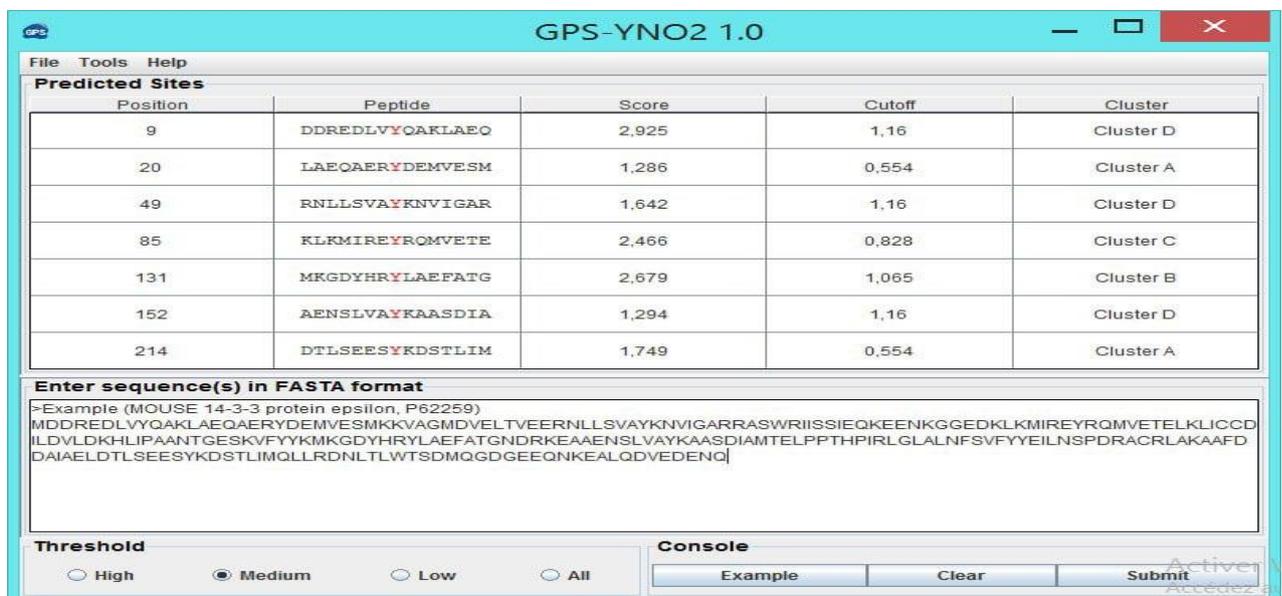


Figure 2.11 : Prédiction des sites de nitration des tyrosines par GPS-YNO2

3.6.2. DeepNitro

DeepNitro a été utilisé pour prédire les sites de nitrations des tyrosines. Il présente une performance supérieure pour prédire la nitration des protéines par rapport aux outils existants. Sa capacité de prédiction est améliorée principalement par les nouvelles méthodes de codage, en particulier la méthode PSSM (Position-Specific Scoring Matrix) (Xie et al. 2018).

3.7. Prédiction des sites de S-nitrosylation

3.7.1. GPS-SNO 1.0

Le programme GPS-SNO 1.0 <http://yno2.biocuckoo.org/> a été utilisé pour prédire les sites de S-nitrosylation des protéines. Il s'agit d'un outil informatique qui utilise un algorithme modifié basé sur le système GPS (Xue et al. 2010).

Position	Peptide	Score	Cutoff	Cluster
10	EELVLERCDLELETN	21,803	20,743	Cluster C
27	DHHTADLCREKLVVR	3,478	2,454	Cluster B
98	TVVDQDCTLSLQLT	24,343	20,743	Cluster C
143	ILLFNAWCPADAVYL	3,348	2,454	Cluster B
230	VVSGMVNCDNDDQGVL	3,114	2,454	Cluster B
370	EKSEGTYCCGPVPVR	1,76	1,67	Cluster A
505	NTAEEYVCRLLLCAR	22,635	20,743	Cluster C
524	NGILGPECGTKYLLN	1,88	1,67	Cluster A
620	LPVALEGCTFVEGA	3,19	2,454	Cluster B

Enter sequence(s) in FASTA format

>Example (Human Protein- glutamine gamma- glutamyltransferase, P21980)
MAEELVLERCDLELETNGRDHHTADLCREKLVRRGQPFWLT LHFEGRNYEASVDSLTF SVWTPAPSQEQAGTKARFPLRDAVEEGDWTATVDDQQ
DCTLSLQLTTPANAPIGLYRLSLEASTGYQGSSFLGHFILLFNAWCPADAVYLDSEEEERQEYVLTQQGFIYQGSQSAKFIKNIPWNFGQFEDGILDICLILL
DVNPKFLKNAGRDCSRRSSPVYVGRVYSGMVNCDNDDQGVLLGRWDNNYGDGVSPMSWIGSVDILRRWKNHGCCQRVKYGGCCWVFAVAECTVLRCL
GIPTRVVTNYSNADQNSNLLIEYFRNEFGEIQGDKSEMIWNFHCWVESWMTFRPDLQPGYEGWQALDPTPQEKSEGTCCGFPVPVRAIKEGDLSTK
YDAPFVFAEVNADVVDWIQQDDG SVHKSINRSLIVGLKISTKSVGRDEREDITHYKYPEGSSSEEREAFTRANHLNKLAEKEETGMAMRIRVQSMNM
GSDFDVF AHITNNTAEEYVCRLLLCARTVSYNGILGPECGTKYLLNLEPFSEKSVPLCLILYKEYRDCLETESNLIKVRALLVEPVINSYLLAERDLYLEN
PEIKIRILGEPKQKRKLVAEVS LQNPLPVALEGCTFVEGAGL TEEQKTVEIPDPVEAGEEVKVRMDLLPLHMGHLKLVNLFESDKLKAVKGFNRNVIIGPA

Threshold

High Medium Low All

Console

Figure 2.13 :Prédiction des sites de S-nitrosylation des protéines par GPS-SNO 1.0

3.7.2. DeepNitro

DeepNitro, a également été utilisé pour prédire les sites de S-nitrosylation. Il présente une performance supérieure pour la prédiction de s-nitrosylation des protéines par rapport aux outils existants grâce à la méthode PSSM (Position-Specific Scoring Matrix) (Xie et al. 2018).

Home Download **Web Server** Database Citation Help Contact

Result Form

Download ▾ ▾

ID	Position	Peptide	Score	Cutoff	Cluster
Example (HUMAN Phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase, Q6XPS3)	267	HYRVYNLC S ERAYDP	0.471	0.37	S-Nitrosylation
Example (HUMAN Phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase, Q6XPS3)	411	YSIRGDV C DLKVVV	0.468	0.37	S-Nitrosylation
Example (HUMAN Phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase, Q6XPS3)	506	NPHKQ K AWKIYPPEF	0.778	0.575	Tryptophan Nitration
Example (HUMAN Phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase, Q6XPS3)	509	KQKAWKI Y PPEFAVE	0.563	0.295	Tyrosine Nitration

Showing 1 to 4 of 4 rows

Visualize

Activer Windows
Accédez aux paramètres de l'ordinateur pour activer Windows.

Figure 2.14 : prédiction des sites de S-nitrosylation par DeepNitro

4. Modélisation moléculaire

4.1. Modélisation par Phyre2

Nous avons utilisé **Phyre2** (**P**rotein **H**omology/analog**Y** **R**ecognition **E**ngine **V**) 2.0 (<http://www.sbg.bio.ic.ac.uk/~phyre2>) pour générer des modèles de structurations tridimensionnelles des protéines sauvages mutées sélectionnées dans cette étude. C'est un outil bioinformatique en ligne basé sur la détection d'homologie pour construire des modèles 3D.

Une fois les résultats obtenus, nous pourrions les utiliser pour comparer la structure protéique mutante avec celle du modèle protéique sauvage en utilisant un logiciel approprié.

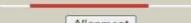
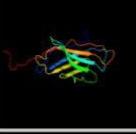
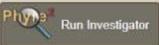
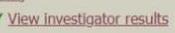
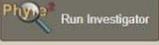
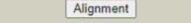
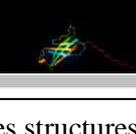
#	Template	Alignment Coverage	3D Model	Confidence	% i.d.	Template Information
1	d1ah1a	 Alignment		100.0	99	Fold: Immunoglobulin-like beta-sandwich Superfamily: Immunoglobulin Family: V set domains (antibody variable domain-like) PDB entry: PDBe RCSB PDBj 
2	d1i8lc	 Alignment		100.0	98	Fold: Immunoglobulin-like beta-sandwich Superfamily: Immunoglobulin Family: V set domains (antibody variable domain-like) PDB entry: PDBe RCSB PDBj 
3	d1d0ta	 Alignment		100.0	65	Fold: Immunoglobulin-like beta-sandwich Superfamily: Immunoglobulin Family: V set domains (antibody variable domain-like) PDB entry: PDBe RCSB PDBj 
4	c4n8pA	 Alignment		100.0	56	PDB header: immune system Chain: A: PDB Molecule: uncharacterized protein; PDB Title: crystal structure of a strand swapped cta-4 from duckbill platypus2 [psi-nysgrc-012711] PDB Entry: PDBe RCSB PDBj

Figure 2.16 : Modélisation des structures tridimensionnelles des protéines par Phyre2

4.2. Modélisation par SWISS-MODEL

Pour la modélisation des structures protéiques nous avons aussi utilisés **SWISS-MODEL** (<https://swissmodel.expasy.org/>). C'est est un serveur de modélisation de structures protéiques entièrement automatisé, accessible via le serveur web ExPasy. Il permet de générer des modèles en ligne en réalisant des correspondances avec plusieurs templates et bases de données établies. Il offre trois types de requêtes de modélisation (mode automatique, mode d'alignement, mode de projet) qui varient en termes d'intervention de l'utilisateur. La modélisation est basée sur les meilleurs templates disponibles en utilisant PROMOD-II et MODELLER. Les modèles sont construits en se basant sur l'alignement entre la séquence cible et le Template, et leur qualité est évaluée à l'aide de la fonction de score QMEAN, qui estime la qualité globale et locale du modèle. Des pondérations spécifiques pour les différents critères QMEAN ont été optimisés pour améliorer les performances du serveur (**Boominathan et Ferreira 2016**).



Figure 2.16 : Modélisation des structures tridimensionnelles des protéines par SWISS-Model

5. Évaluations de la qualité des structures des protéines modélisées

5.1. ERRAT

Pour évaluer la qualité des modèles générés, nous avons utilisé le logiciel **ERRAT** (<https://saves.mbi.ucla.edu/>), qui est un programme conçu pour vérifier les structures des protéines. Il utilise une fenêtre coulissante de neuf résidus pour calculer les valeurs d'erreur en fonction de leur position. Ces erreurs sont déterminées en se basant sur les statistiques des interactions atomiques non liées présentes dans la structure étudiée, comparées à une base de données de structures de haute résolution considérées comme fiables. Ainsi, ERRAT permet d'évaluer objectivement la qualité de la structure protéique et de détecter d'éventuelles erreurs ou divergences par rapport aux normes attendues (**ERRAT – UCLA-DOE Institute s. d.**).

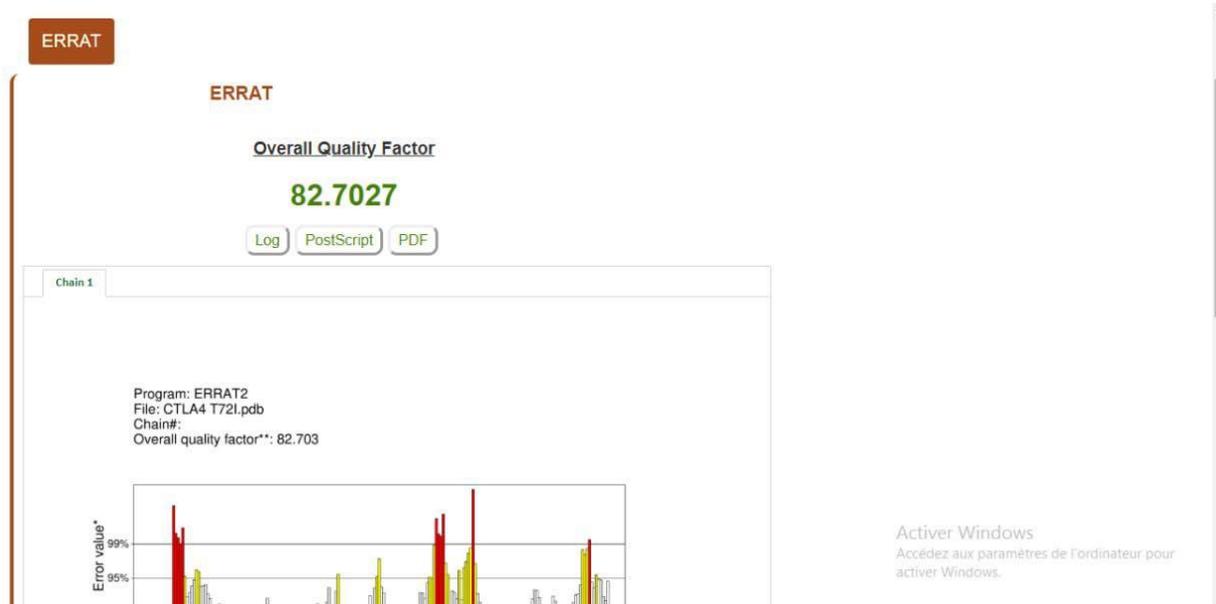


Figure 2.17 : exemple d'une réalisation sur ERRAT

5.2.PROCHECK

Et on a utilisé après le logiciel **PROCHECK**, Le programme effectue des calculs de plusieurs paramètres stéréochimiques pour le modèle protéique fourni, puis les compare à des valeurs "idéales" provenant d'une base de données de structures protéiques bien raffinées et de haute résolution dans la Protein Data Bank (**Laskowski, MacArthur, et Thornton 2006**).

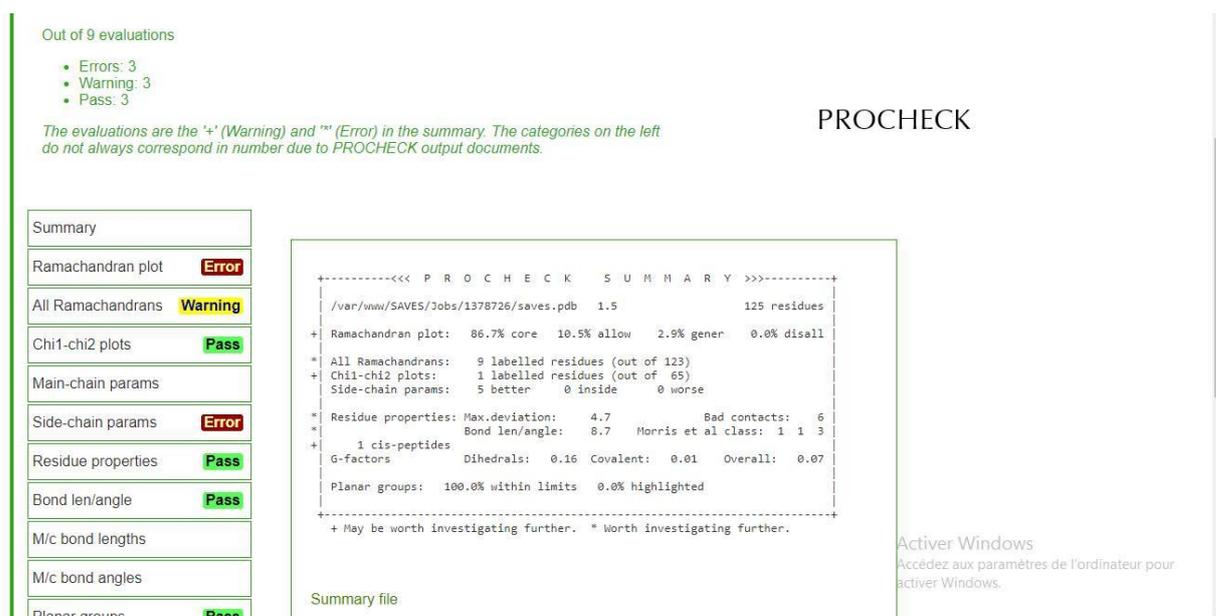


Figure 2.18 : Évaluations de la qualité des structures des protéines par PROCHECK

6. Comparaison entre les structures tridimensionnelles des protéines

6.1. TM-align

Pour comparer entre les structures des protéines mutées et des protéines sauvages, nous avons utilisé l'outil **TM Align** (<https://zhanggroup.org/TM-align/>). Cette méthode utilise le TM-score pour aligner les structures protéiques. Le TM-score est une mesure de similitude entre deux structures tridimensionnelles qui évalue leur ressemblance à partir d'une programmation dynamique. En prenant en compte les différences et variations entre ces structures, TM Align optimise cet alignement afin d'obtenir la correspondance la plus précise possible (**Zhang et Skolnick 2005**).

```

TM-align Results

*****
*                               TM-align (Version 20190822)                               *
* An algorithm for protein structure alignment and comparison                          *
* Based on statistics:                                                            *
*   0.0 < TM-score < 0.30, random structural similarity                            *
*   0.5 < TM-score < 1.00, in about the same fold                                *
* Reference: Y Zhang and J Skolnick, Nucl Acids Res 33, 2302-9 (2005)             *
* Please email your comments and suggestions to: zhang@umich.edu                  *
*****

Name of Chain_1: A232593
Name of Chain_2: B232593
Length of Chain_1: 129 residues
Length of Chain_2: 125 residues

Aligned length= 125, RMSD= 0.00, Seq_ID=n_identical/n_aligned= 0.992
TM-score= 0.96899 (if normalized by length of Chain_1)
TM-score= 1.00000 (if normalized by length of Chain_2)
(You should use TM-score normalized by length of the reference protein)

(":" denotes aligned residue pairs of d < 5.0 A, "." denotes other aligned residues)
AHHVAQPAVVLASSRGIASFVCEYASPGKATEVRVTVLRQADSQVTEVCAATYMMGNELFLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVELIYPPPYLGIQNGTQTVYVIDPEPCPDSQEPK
.....
AHHVAQPAVVLASSRGIASFVCEYASPGKATEVRVTVLRQADSQVTEVCAATYMMGNELFLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVELIYPPPYLGIQNGAQTIVYVIDPEPCPDSQ-----

```

Visualization (Protein-1 in blue and Protein-2 in red)

Superposition of two proteins Superposition of two proteins with ligands and solvents (when available)

Application loaded

Figure 2.19: alignement des structures des protéines par TM Align

6.2 Chimera

UCSF Chimera est un programme très extensible pour la visualisation interactive et l'analyse des structures moléculaires et des données connexes, y compris les cartes de densité, les assemblages supramoléculaires, les alignements de séquence, les résultats d'amarrage, les trajectoires et les ensembles conformationnels. Des images et des animations de haute qualité peuvent être générées. Disponible à télécharger ici [<https://www.cgl.ucsf.edu/chimera/download.html>].

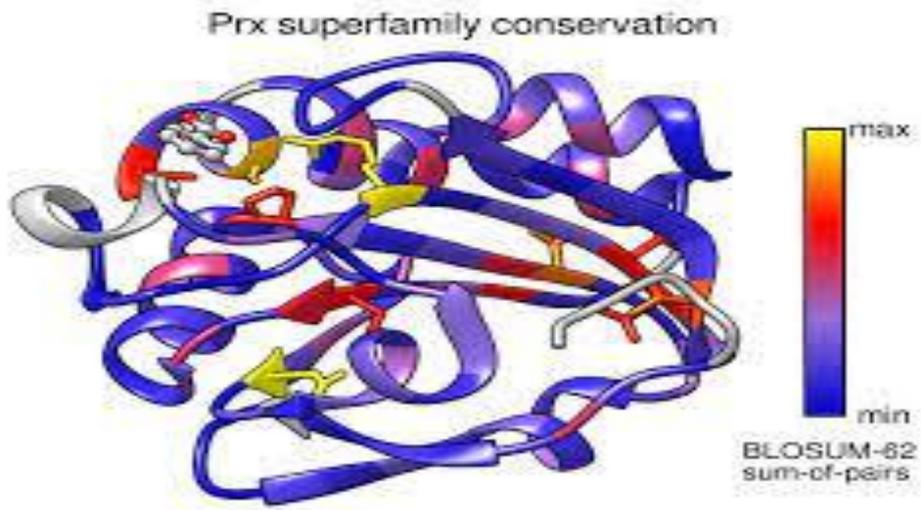


Figure 2.20 : exemple des résultats sur Chimera

CHAPITRE III :
RESULTATS
ET
INTERPRÉTA
TIONS

III- Chapitre III : Résultats

1. Résultats de Prédiction des modifications post-traductionnelles

1.1 Prédiction des sites de phosphorylation

Les séquences sauvages des protéines CTLA-4 et c-SRC ont été extraites de la base de données Uniprot, puis analysées pour identifier les sites de phosphorylation. En parallèle à cela, nous avons également examiné les protéines mutantes CTLA-4 (T72I, N145S, T147A) et c-SRC (A296S, T432M, T460K, P510S).

Nous avons utilisé des logiciels tels que GPS 6.0 pour prédire les sites potentiels de phosphorylation des résidus de sérine, thréonine et tyrosine. Les résultats ont été rapportés sous forme d'un score, qui indique le potentiel du résidu à être phosphorylé (plus la valeur est élevée, plus le potentiel est grand), avec un seuil fixé pour éviter les faux positifs. Si la valeur était nulle, cela signifiait qu'aucune phosphorylation intéressante n'a été détectée.

Nos recherches ont révélé que la protéine CTLA-4 sauvage présentait une perte significative de phosphorylation au niveau des positions Thréonine 72 et Thréonine 147. Cela suggère que ces sites ont été altérés par mutation. En revanche, rien d'anormal n'a été signalé dans c-SRC.

Nous avons utilisé d'autres outils en ligne pour prédire les sites potentiels de phosphorylation des résidus de sérine, thréonine ou tyrosine dans les séquences protéiques CTLA-4 et SRC sauvages. Le seuil était fixé à 0,4 pour Netphos 3.1 et 0,05 pour le serveur Musitedeep, qui permettent tous deux la prédiction des sites généraux et spécifiques aux kinases.

Les résultats obtenus par ces logiciels ont confirmé ceux du premier logiciel utilisé. En ce qui concerne l'état muté de SRC, nous avons observé une augmentation significative de la phosphorylation à la position Sérine 510. Cette observation peut indiquer une hyperactivation possible de l'activité kinase de SRC.

Tableau 3.1 : Résultats de prédiction des sites de phosphorylation

Phosphorylation										
Protéine	Position	Sauvage			SNP ID	Position Mutée	Mutée			Résultats
		GPS 6.0	MusiteDeep (Thershold 0.05)	Netphos (Thershold 0.4)			GPS 6.0	MusiteDeep (Thershold 0.05)	Netphos (Thershold 0.4)	
CTLA-4	T 72	0.0115 Cutoff : 0.0109	0.097	0.461	rs781579729	T72I	Nul	Nul	Nul	Positif
	N 145	Nul	Nul	Nul	rs1204026047	N 145S	0.0102	0.053	0.478	Négatif
	T 147	0.4362 Cutoff : 0.0222	0.051	0.591	rs1466152724	T 147A	Nul	Nul	Nul	Positif
SRC	A296	Nul	Nul	Nul	Rs750534584	A296S	0	0.147	0.506	Négatif
	T432	0	0.119	0.651	Rs1490969287	T432M	Nul	Nul	Nul	Négatif
	T460	0	0.079	0.667	Rs1171125750	T460K	Nul	Nul	Nul	Négatif
	P510	Nul	Nul	Nul	Rs1346524802	P510S	0.2283 (Cutoff : 0.04)	0.082	0.532	Positif

1.2. Résultats de prédiction des Méthylation

Nous avons effectué des tests de prédiction de méthylation des acides aminés Arginine et Lysine en utilisant GPS-MSP, MusiteDeep et PRmePRed. Aucune méthylation n'a été détectée dans les séquences sauvages de CTLA-4 et c-SRC. Cependant, dans l'état muté de SRC au site P136R, les résultats obtenus par GPS-MSP avec un seuil fixé à 0,3 ont révélé une prédiction positive pour la méthylation des résidus d'arginine. Cette observation a été validée par MusiteDeep et PRmePRed.

Tableau 3.2 : Résultats de prédiction des sites de Méthylation

Méthylation (R-K)										
Protéine	Position	Sauvage			SNP ID	Position Mutée	Mutée			
		GPS-MS P	Musite Deep (Thers hold 0.05)	PRmePR ed			GPS-MSP (theresh old 0.3)	Musite deep (Thers hold 0.05)	PRmePR ed	résultats
CTLA-4		Aucune Mutation qui peut causer Une méthylation								
SRC	P136	Nulle	Nul	Nul	Rs14126 26612	P136R	0.3807	0.157	0.598806	Positif
	T460	Nulle	Nul	Nul	Rs11711 25750	T460K	0.1881	0.085	Nul	Négatif
	R509	0.1823	0.028	Nul	Rs13053 388441	R509W	Nulle	Nulle	Nul	Négatif

1.3. Résultats de prédiction des Ubiquitinations

Dans l'état sauvage de la protéine SRC, les prédictions ne montrent pas de site d'ubiquitination potentiel.

Dans l'état muté de la protéine, les résultats de prédiction confirment également l'absence de sites d'ubiquitination.

Tableau 3.3 Résultats de prédiction des Ubiquitinations

Ubiquitination (K)									
Protéine	Position	Sauvage		SNP ID	Position mutée	Mutée		résultats	
		GPS.Uber	MusiteDeep			GPS.Uber	MusiteDeep		
CTLA4		Aucune site potentiel d'ubiquitination							
SRC	T460	Nul	Nul	Rs11711 25750	T460K	0	0.22	Négatif	

1.4. Résultats de prédiction des domaines de liaison aux phosphoprotéines.

Nous avons utilisé le logiciel GPS-PBS pour prédire les domaines de liaison aux phosphoprotéines. Pour la séquence sauvage de CTLA-4, les résultats ont permis d'identifier la présence de deux domaines potentiels de liaison aux phosphoprotéines au niveau des positions **Thréonine 72 et Thréonine 147**. En revanche, dans la séquence mutée de CTLA-4, les prédictions ont révélé une altération potentielle de ces deux domaines de liaison aux phosphoprotéines ainsi que l'apparition d'un nouveau site (145) potentiellement impliqué dans la liaison avec les phosphoprotéines.

Concernant la séquence sauvage c-SRC, les résultats obtenus étaient négatifs et suggèrent qu'il n'y a pas un domaine identifiable impliqué. Cependant, pour la séquence mutée c-SRC, l'introduction d'une mutation au niveau du site A296S a créé un nouveau domaine potentiellement impliqué dans la liaison avec les phosphoprotéines.

Tableau 3.4 : Résultats de prédiction des domaines de liaison aux phosphoprotéines

Domaines de liaison aux phosphoprotéines (S-T-Y)						
Protéine	Position	Sauvage	SNP ID	Position Mutée	Mutée	Résultats
		GPS-PBS			GPS-PBS	
CTLA-4	T72	Score : 0.0272 Cutoff : 0.0197	rs78157 9729	T72I	Nul	Positif
	N145	Nul	rs12040 26047	N145S	Score : 0.1640 Cutoff : 0.0406	Positif
	T147	Score : 0.9836 Cutoff : 0.9007 (PBD)	rs14661 52724	T147A	Nul	Positif
SRC	A296	Nul	Rs75053 4584	A296S	Score : 0.0833 Cutoff : 0.0197	Positif
	T432	Nul	Rs14909 69287	T432M	Nul	Négatif
	T460	Nul	Rs11711 25750	T460K	Nul	Négatif
	P510	Nul	Rs13465 24802	P510S	Nul	Négatif

1.5. Résultats de prédiction de Tyrosine nitration sites et S-nitrosylation

Nous avons utilisé les outils bioinformatiques GPS-YNO2 et DeepNitro pour effectuer des prédictions des sites de nitration des Tyrosine et S-nitrosylation. Cependant, nous n'avons pas pu exécuter cette opération car il n'y avait aucune mutation présentant des acides aminés susceptibles d'être nitrés (tyrosine).

Tableau 3.5 : Résultats de prédiction de Tyrosine nitration sites et S-nitrosylation

Tyrosine nitration sites							
On n'a pas une mutation qui touche Tyrosine							
Protéine	Position	Sauvage		Position mutée	Mutée		Résultats
		GPS-YNO2	DeepNitro		GPS-YNO2	DeepNitro	
S-nitrosylation sites							
On n'a pas une mutation qui touche Cystéine							
Protéine	Position	Sauvage		Position mutée	Mutée		Résultats
		GPS-SNO 1.0	DeepNitro		GPS-SNO 1.0	DeepNitro	

2. Modélisation moléculaire

2.2. Modélisation par SWISS-MODEL

Nous avons obtenu des résultats de modélisation moléculaire pour les séquences protéiques CTLA-4 et c-SRC en utilisant des modèles de référence (Templates), afin d'optimiser les résultats et d'obtenir des structures fiables. Le template utilisé pour CTLA-4 était 3osk.1.B. Pour c-SRC, le template choisi était 2h8h.1.A.

SWISS-Model

SWISS-Model a été utilisé pour modéliser les structures protéiques, et les scores obtenus sont les suivants :

- GMQE Score : Pour CTLA4 et SRC, nous avons obtenu un score compris entre 0.78 et 0.97, ce qui suggère une bonne qualité globale de nos modèles basés sur de bons alignements.
- QMEAN Z-Score : Dans CTLA-4 et dans le SNP T72I nous avons obtenus un Z-score de 0,53 ce qui indique que l'énergie de notre modèle de séquences mutées est légèrement plus élevée que la moyenne des structures similaires utilisées comme référence, dans les autres mutations, on a obtenu un z-score positif qui suggère une bonne concordance énergétique entre Nos modèle et les structures de référence.

Dans SRC, nous avons obtenu un Z-score compris entre -0.76 et -0.02 indique que l'énergie de nos modèles des séquences mutées est plus élevée que la moyenne des structures similaires utilisées comme référence. Ramachandran value : Nos modèles ont présenté des scores compris entre 96% et 97% pour nos séquences mutées indiquant que la grande majorité des angles de torsion des résidus se situe dans les régions autorisées du diagramme de Ramachandran.

PROCHECK :

Régions du diagramme de Ramachandran: nos résultats suggèrent une bonne qualité de la conformation des résidus de nos modèles de séquences mutées par rapport au diagramme de Ramachandran.

TM-Align :

TM-Align a été utilisé pour comparer entre les structures des protéines mutées et des protéines sauvages :

TM Score : Pour CTLA4 et SRC, nous avons obtenu un TM Score de 0,99, ce qui indique une bonne similarité entre notre modèle et la structure de référence. Cela suggère que notre modèle capture bien les caractéristiques structurelles essentielles de la protéine mutée.

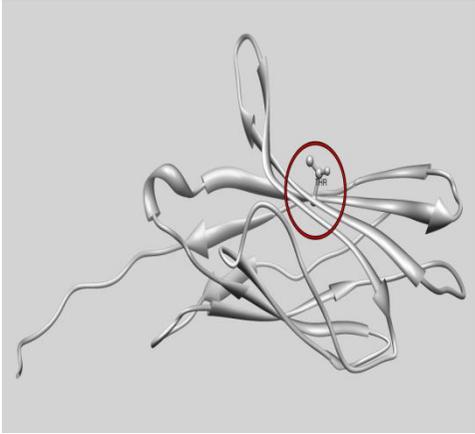
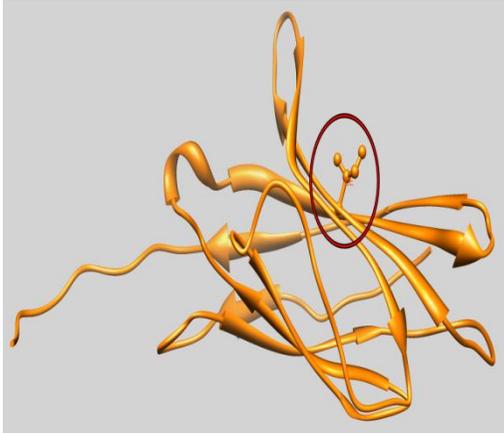
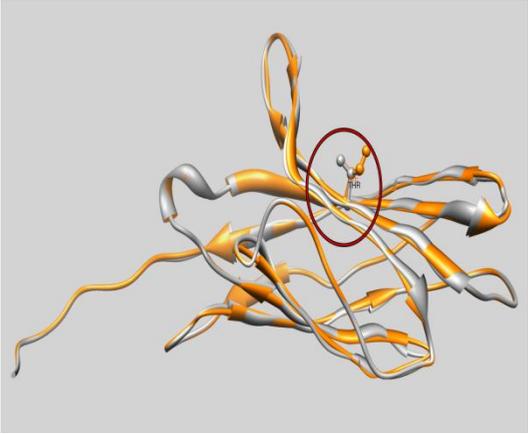
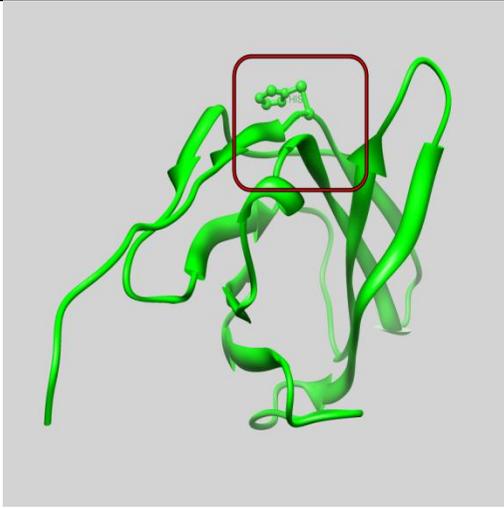
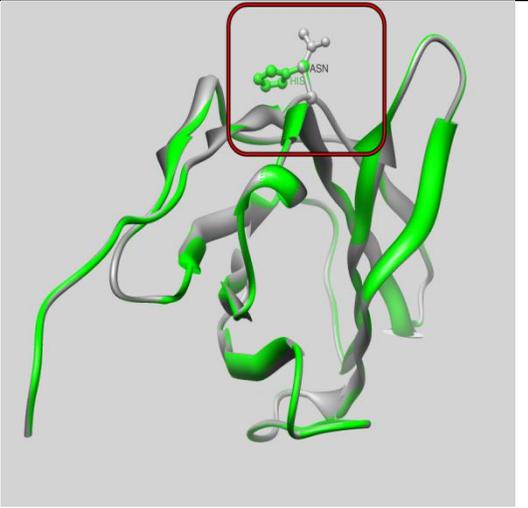
RMSD : Pour CTLA4 et SRC, nous avons obtenu un RMSD compris entre 0,08 et 0,11. Cela indique que les positions des atomes dans notre modèle diffèrent en moyenne de seulement 0,08 à 0,11 par rapport à la structure sauvage.

Tableau 3.6 : Résultats de l'évaluation de qualité des structures modélisées par SWISS-MODEL on utilisant TM-align, ERRAT et PROCHECK

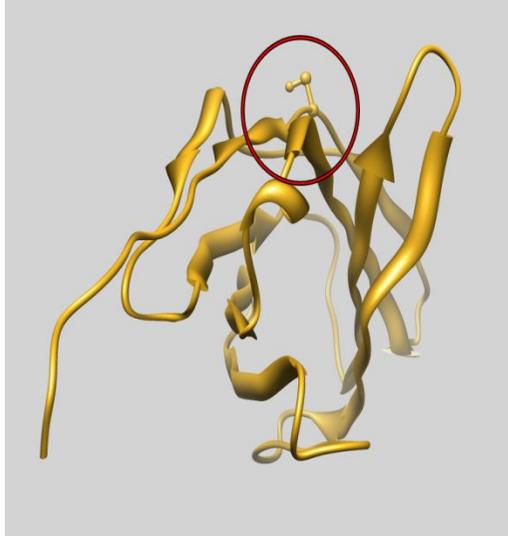
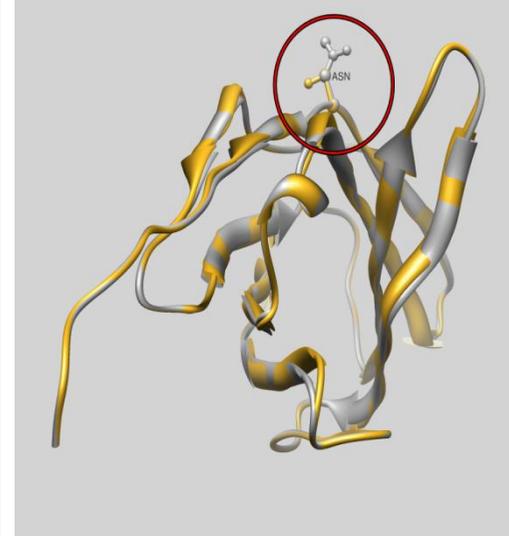
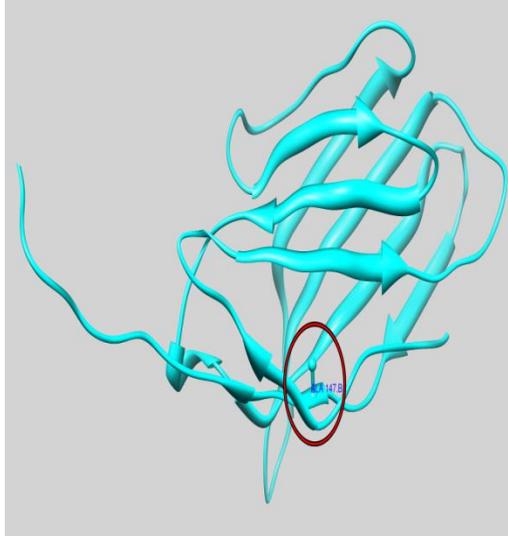
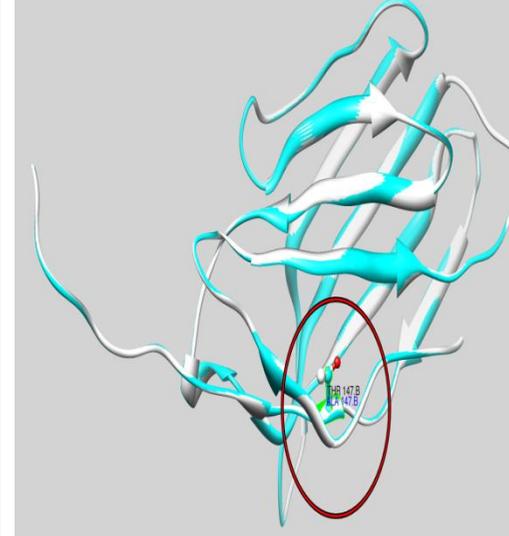
SNP Id	Mutation	TM-align		SWISS-MODEL			ERRAT	PROCHECK (Ramachandran plot analysis)			
		Tm Score	RM SD	GMQE score	QMEAN Z-score	Ramachandran value	valeur	Residues in most favoured regions	Residues in additional allowed regions	Residues in generously allowed regions	Residues in disallowed regions
Modélisation de CTLA-4 en utilisant le template 3osk.1.B											
rs781579729	T72I	0.99966	0.08	0.97	0.53	95.83%	88.3721	187 (91.2%)	17 (8.3)	1 (0.5%)	0
rs1356678649	N145H	0.99958	0.08	0.80	0.06	95.04%	84.6154	92 (89.3%)	11(10.7)	0	0
rs1204026047	N145S	0.99956	0.09	0.80	0.20	95.04%	84.6154	92 (89.3%)	11(10.7)	0	0
rs1466152724	T147A	0.99956	0.09	0.78	-0.02	95.04%	84.6154	92 (89.3%)	11(10.7)	0	0
Modélisation de c-Src en utilisant le template 2h8h.1.A											
Rs1412626612	P136R	0.99759	0.10	0.82	-0.12	96.84%	85.6512	465(86.7%)	51 (9.6%)	11 (2%)	9(1.7%)
Rs186207963	D407N	0.99760	0.09	0.83	-0.02	96.91%	87.1965	465(86.7%)	51 (9.6%)	11 (2%)	9(1.7%)
Rs1171125750	T460K	0.99763	0.08	0.82	-0.29	96.84%	86.5342	465(86.7%)	51 (9.6%)	11 (2%)	9(1.7%)
Rs756665933	P487L	0.99763	0.08	0.83	-0.38	96.84%	86.755	465(86.7%)	51 (9.6%)	11 (2%)	9(1.7%)
Rs13053388441	R509W	0.99758	0.10	0.83	-0.11	96.61%	88.3002	465(86.7%)	51 (9.6%)	11 (2%)	9(1.7%)
Rs1346524802	P510A	0.99763	0.08	0.83	-0.16	96.84%	86.9757	465(86.7%)	51 (9.6%)	11 (2%)	9(1.7%)
Rs1346524802	P510S	0.99758	0.10	0.83	-0.14	96.84%	86.9757	465(86.7%)	51 (9.6%)	11 (2%)	9(1.7%)
Rs1366542593	G173V	0.99755	0.11	0.82	-0.76	95.94%	86.5342	462 (86.3%)	54 (10%)	11 (2%)	9(1.7%)
Rs761890358	F223L	0.99759	0.10	0.82	-0.17	96.84%	87.638	465(86.7%)	51 (9.6%)	11 (2%)	9(1.7%)
Rs750534584	A296S	0.99758	0.10	0.82	-0.08	97.07%	87.1965	465(86.7%)	51 (9.6%)	11 (2%)	9(1.7%)

Dans le tableau qui suit résultat de modélisation moléculaire de toute les séquences protéiques CTLA-4 sauvages(disposer dans un templates) VS mutés (T72I, N145H, N145S, T147A) et c-Src (P136R, D407N, T460K, P487L, R509W, P510A, P210S, G173V, F223L, A296S, A330D, M344L, T432M, M498L) par le serveur SWISS-Model afin de visualiser les sites de mutation sur les protéines et vérifier s'il y'a des changements significatifs. Dans le produit obtenu la détection que toutes les mutations déjà annoncer ci-dessus affecte la configuration structurelle des protéines qu'ils appartiennent, résultat confirmer et visualiser par la méthode de superposition configurer par l'outil bioinformatique Chimera.

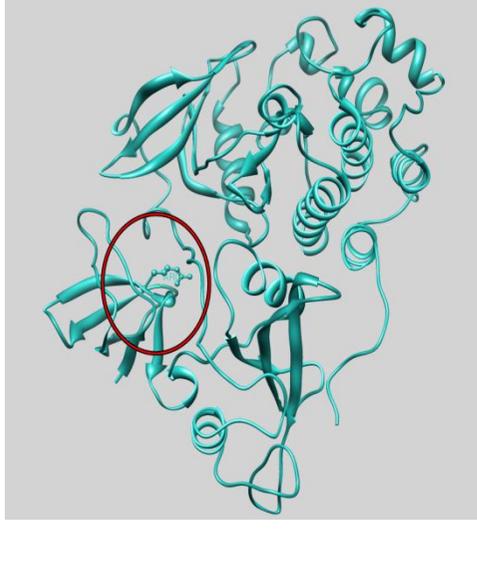
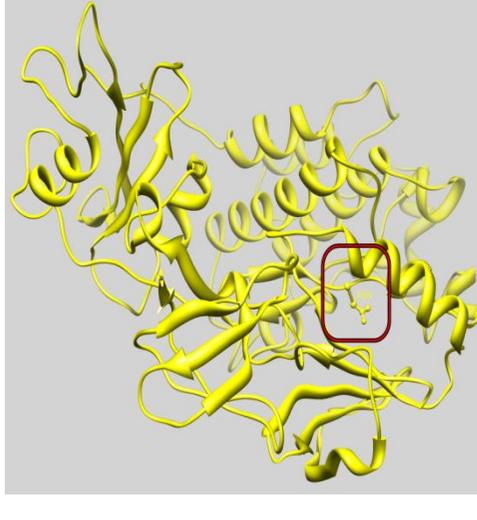
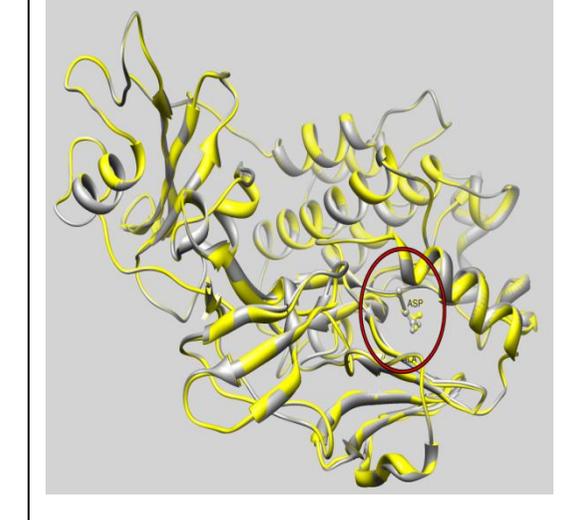
Tableau 3.7 : Visualisation et superposition des protéines modélisées par SWISS-MODEL

SNP Id	Mutation	Site Sauvage	Site Mutée	Merged
Modélisation de CTLA-4 en utilisant le template 3osk.1.B				
rs7815797 29	T72I			
rs1356678 649	N145H			

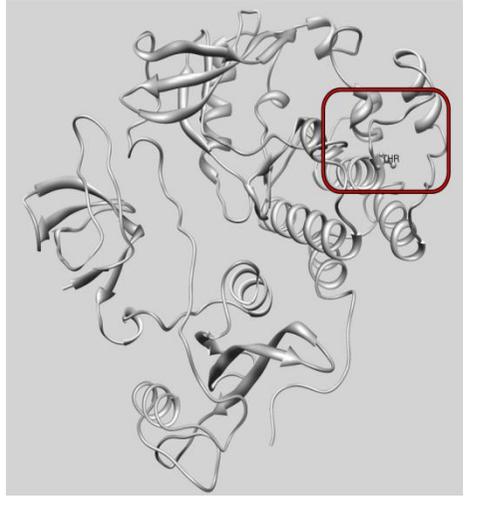
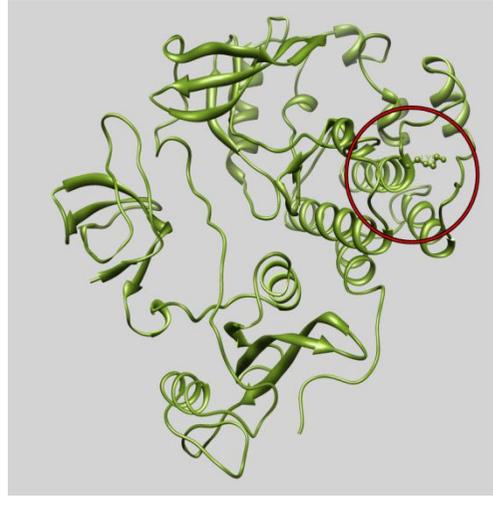
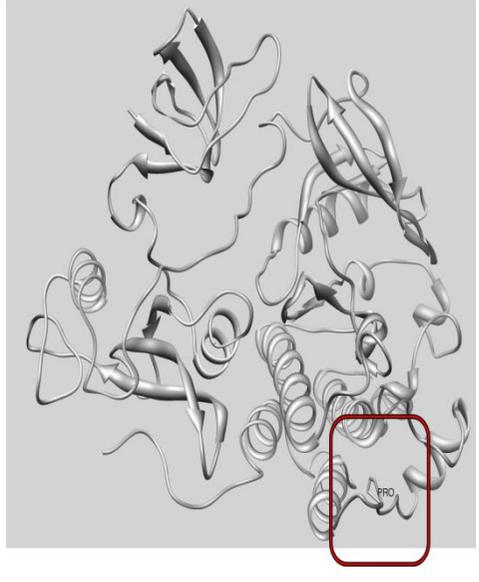
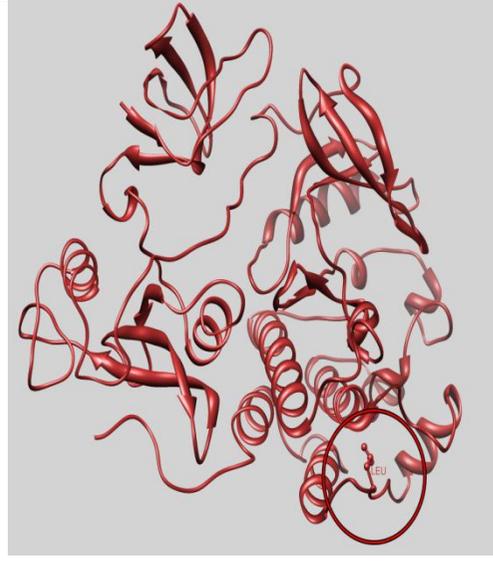
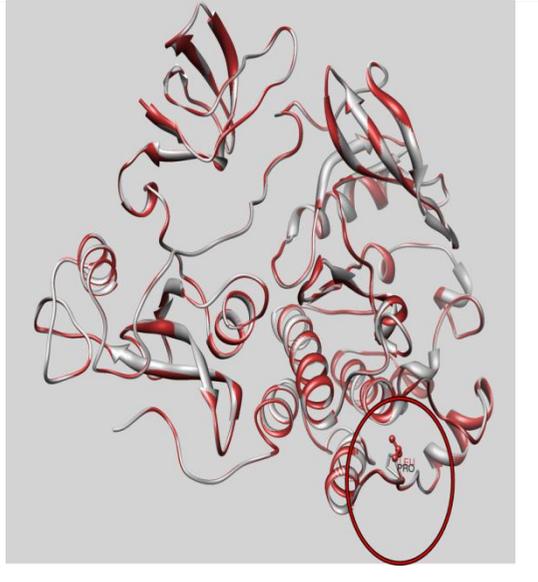
RESULTATS

rs1204026 047	N145S			
rs1466152 724	T147A			
Modélisation de c-Src en utilisant le template 2h8h.1.A				

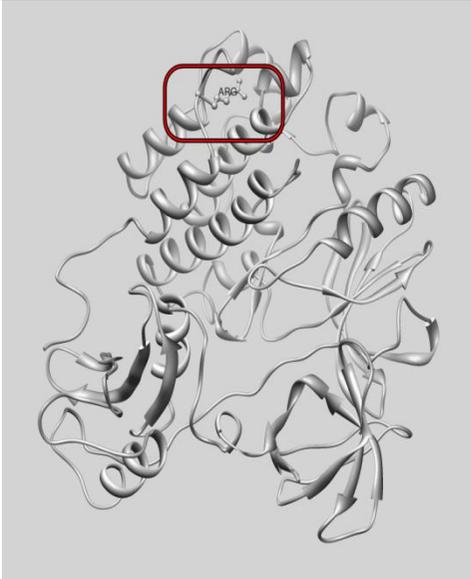
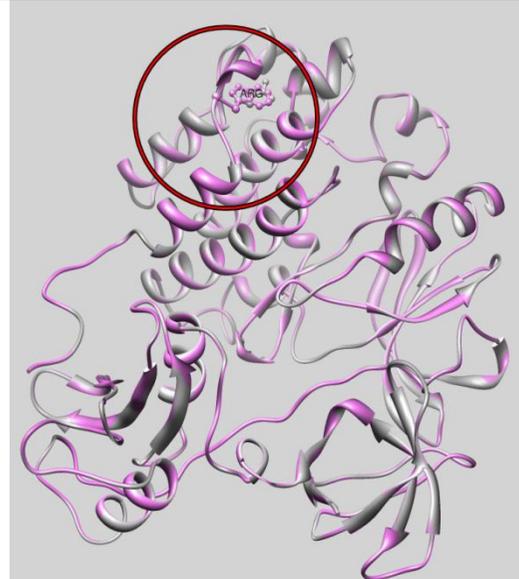
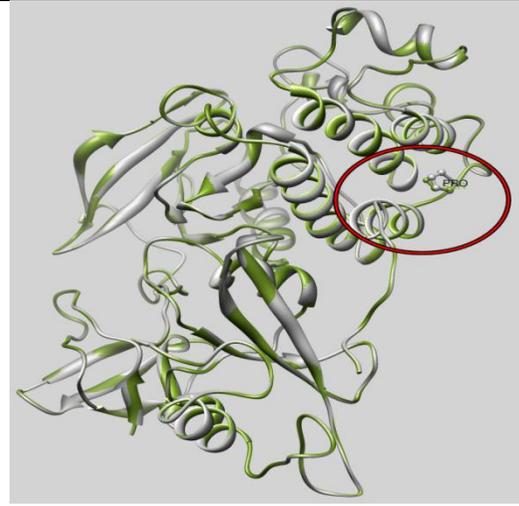
RESULTATS

Rs141262 6612	P136R			
Rs186207 963	D407N			

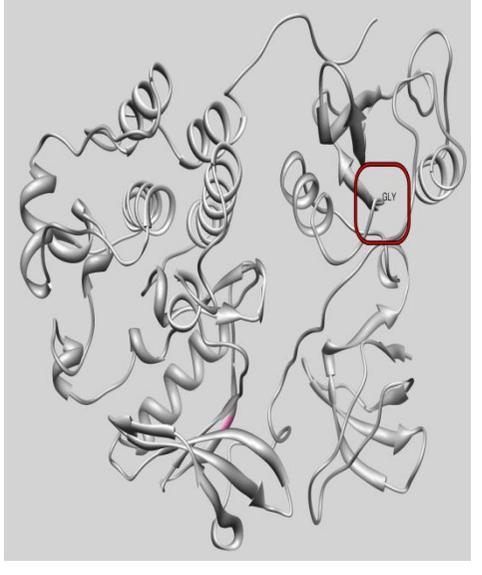
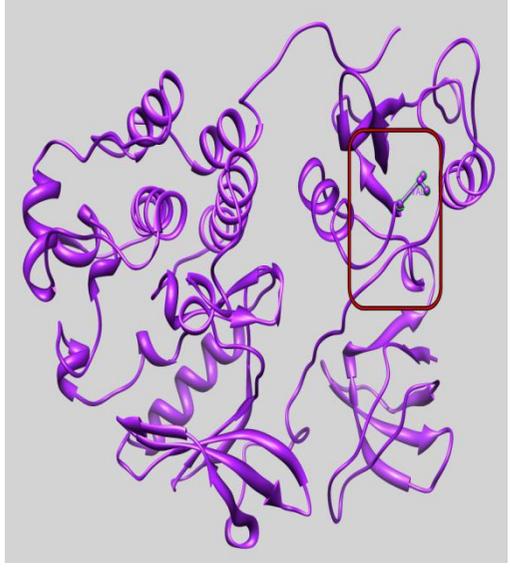
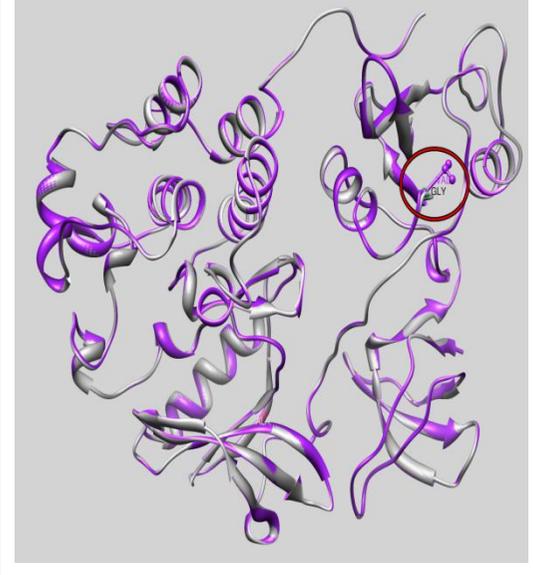
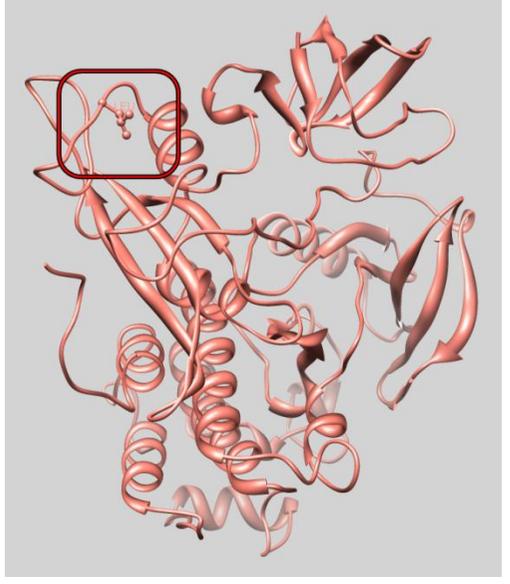
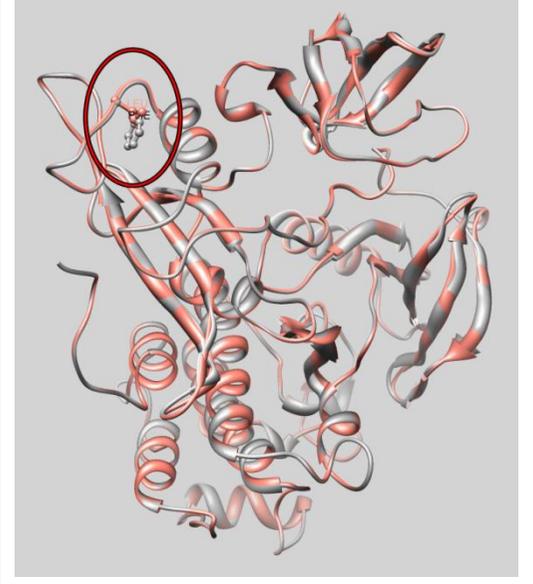
RESULTATS

Rs117112 5750	T460K			
Rs756665 933	P487L			

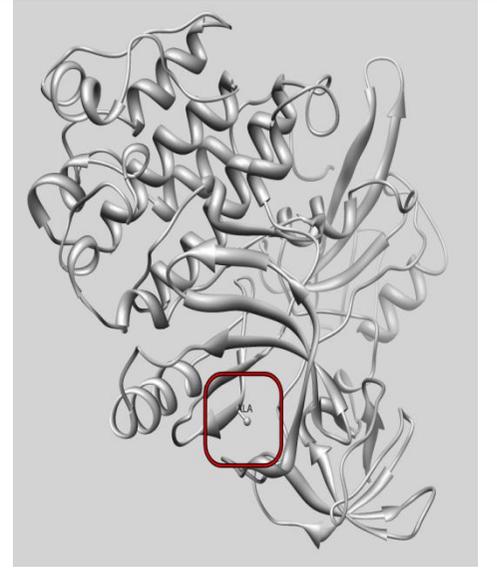
RESULTATS

Rs130533 88441	R509W			
Rs134652 4802	P510S			

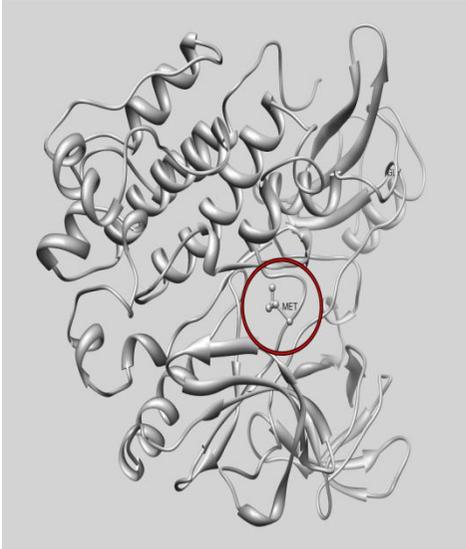
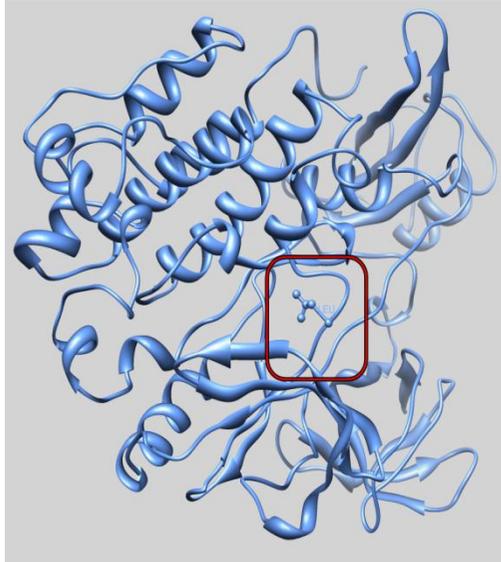
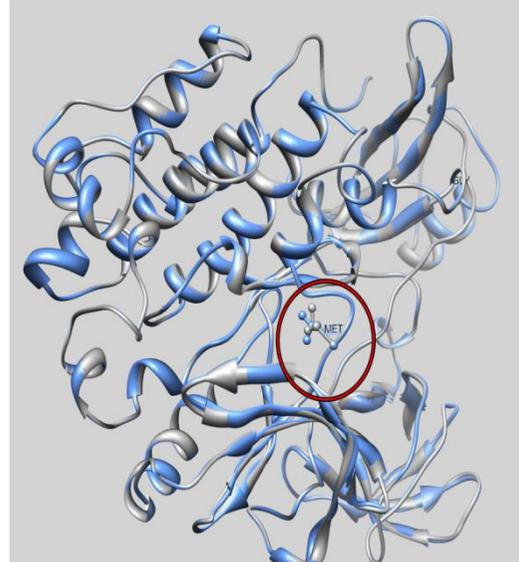
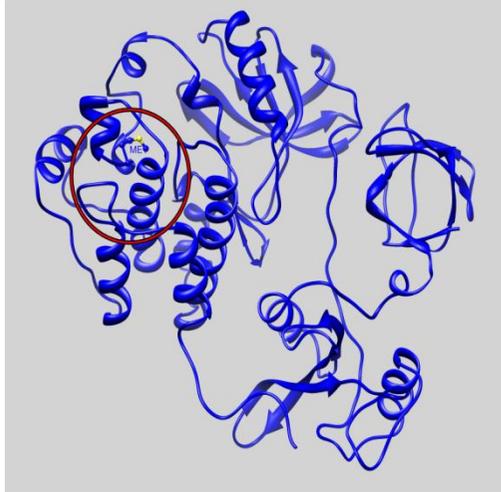
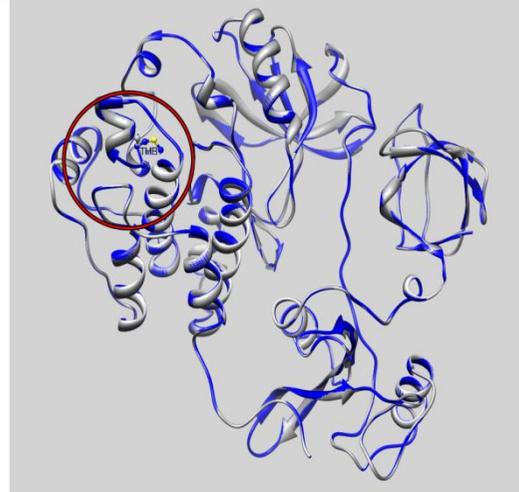
RESULTATS

Rs136654 2593	G173V	 <p>3D ribbon diagram of a protein structure. The mutation G173V is highlighted in a red box. The protein is shown in a light gray color.</p>	 <p>3D ribbon diagram of a protein structure. The mutation G173V is highlighted in a red box. The protein is colored in purple.</p>	 <p>3D ribbon diagram of a protein structure. The mutation G173V is highlighted in a red circle. The protein is colored in purple.</p>
Rs761890 358	F223L	 <p>3D ribbon diagram of a protein structure. The mutation F223L is highlighted in a red box. The protein is shown in a light gray color.</p>	 <p>3D ribbon diagram of a protein structure. The mutation F223L is highlighted in a red box. The protein is colored in orange.</p>	 <p>3D ribbon diagram of a protein structure. The mutation F223L is highlighted in a red circle. The protein is colored in orange.</p>

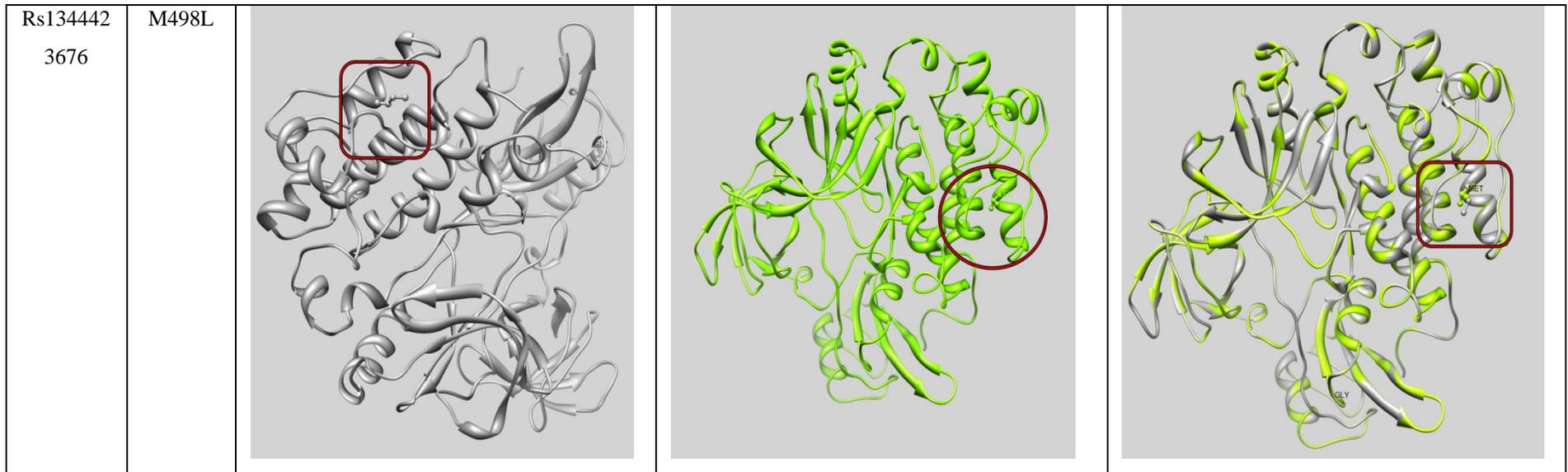
RESULTATS

Rs750534 584	A296S			
Rs127432 2227	A330D			

RESULTATS

Rs754968 435	M344L			
Rs149096 9287	T432M			

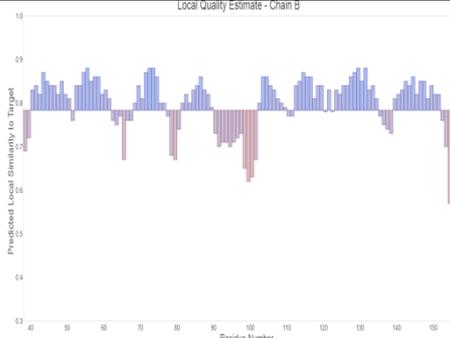
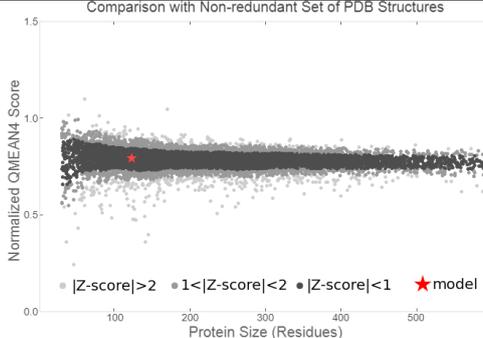
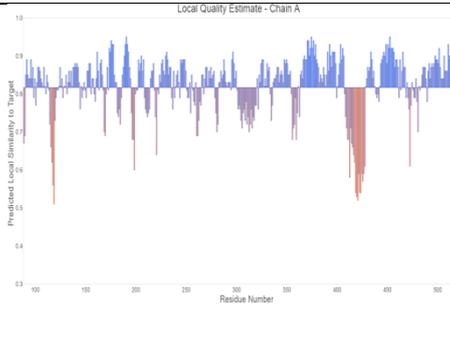
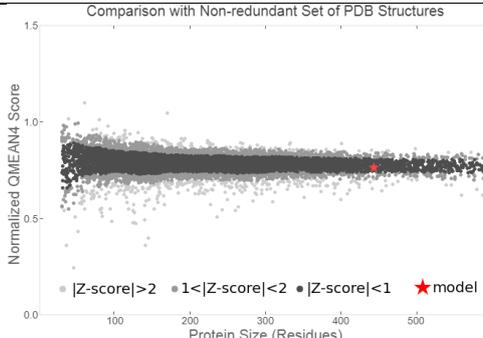
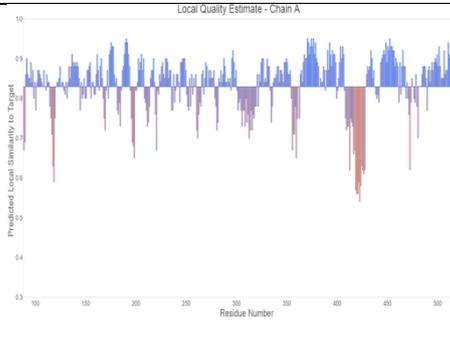
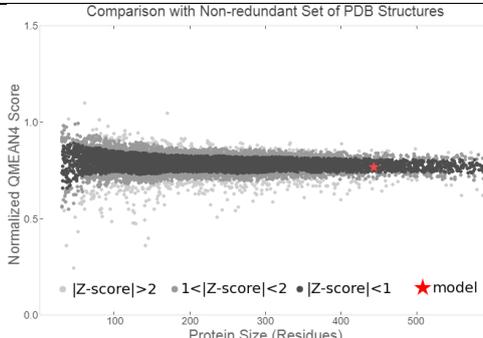
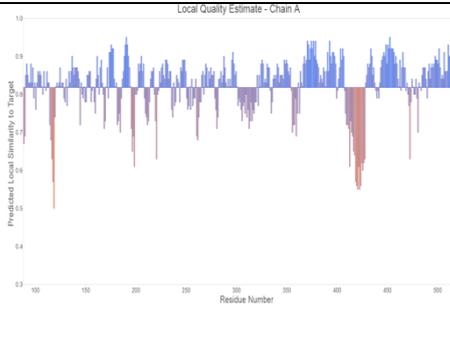
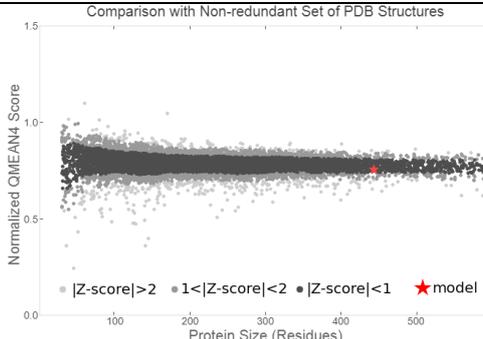
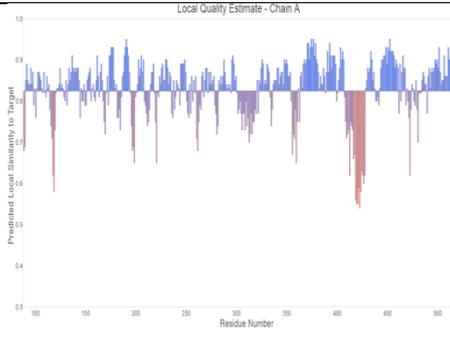
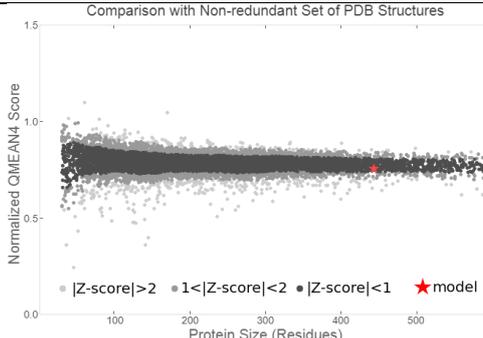
RESULTATS

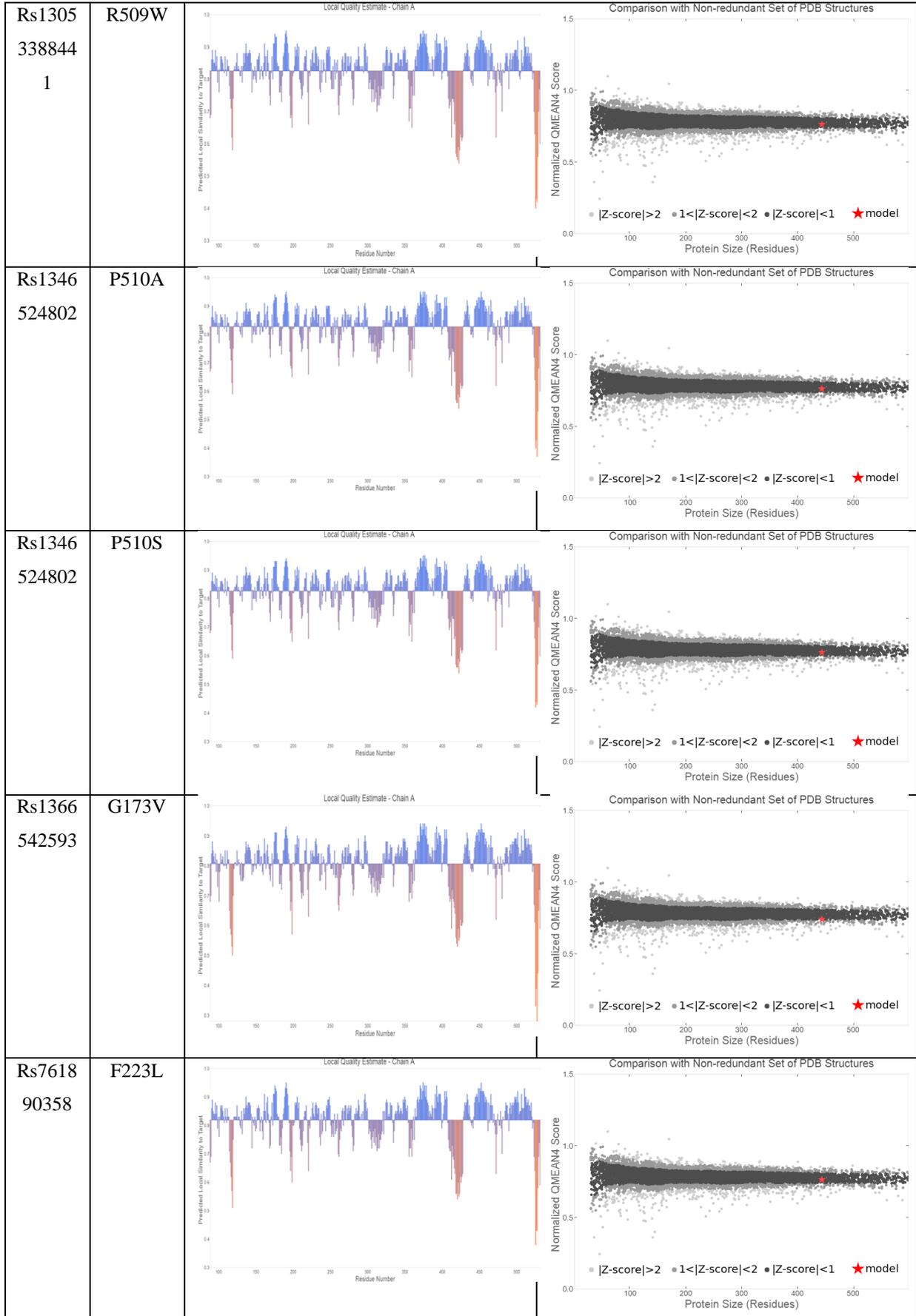


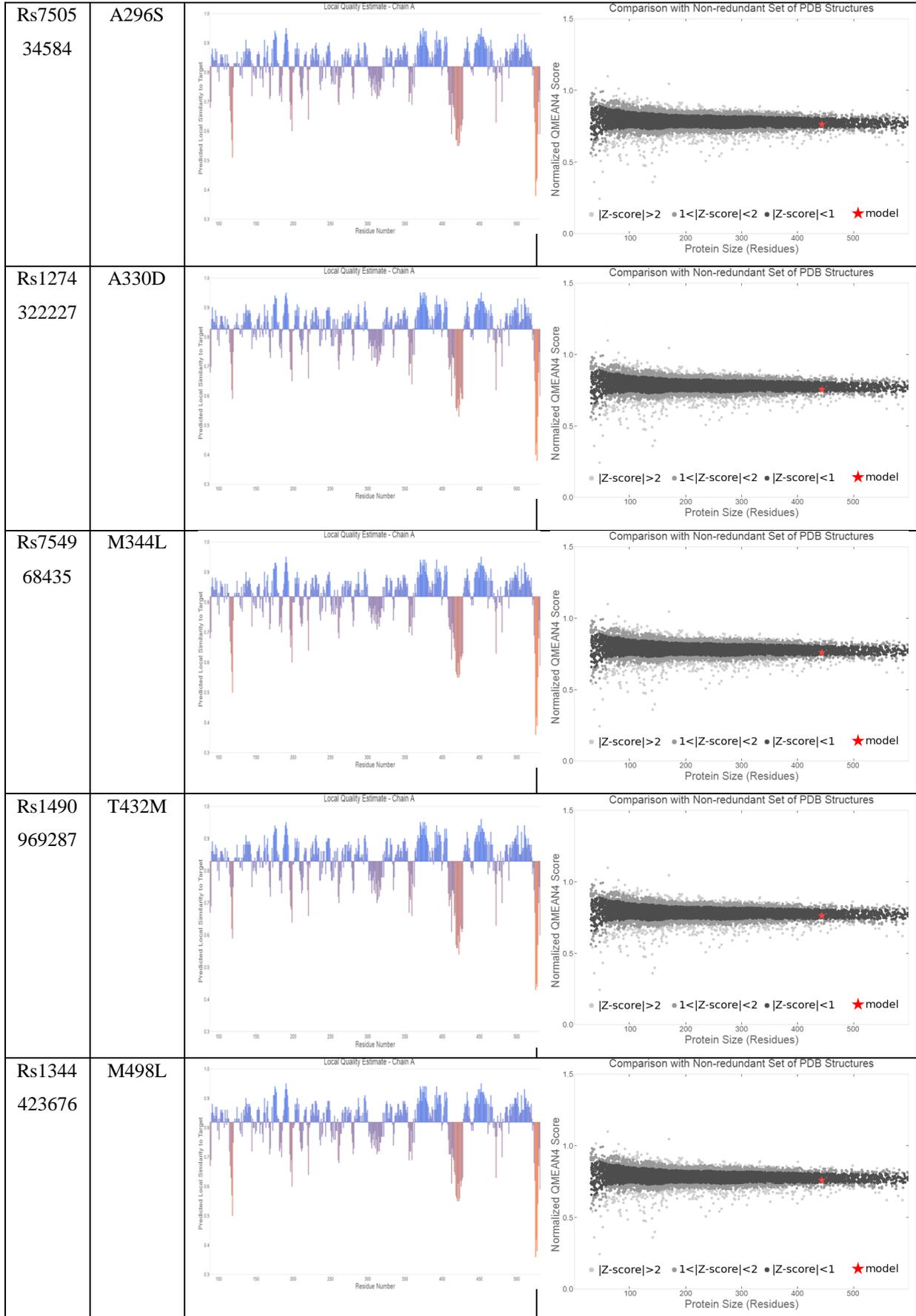
Le graphe « local quality estimate » évalue nos séquences comme étant très proche de la valeur maximale. Cela suggère que ces structures ont été bien modélisées et que sa position et son environnement chimique dans la structure sont conformes aux valeurs expérimentales ou aux règles géométriques attendues pour un résidu de cette nature. Le deuxième graphe « comparison with redundant Set of PDB structures » illustre la comparaison avec un ensemble redondant de structures PDB obtenues par des méthodes expérimentales. La qualité globale et par résidu du modèle a été évaluée à l'aide du score QMEAN.

Tableau 3.8 : Estimations de qualités des protéines modélisées par SWISS-MODEL

SNP Id	Mutation	Local Qualité	Comparaison de la qualité
Modélisation de CTLA-4 en utilisant le template 3osk.1.B			
rs7815 79729	T72I		
rs1356 678649	N145H		
rs1204 026047	N145S		

<p>rs1466 152724</p>	<p>T147A</p>		
<p>Modélisation de c-Src en utilisant le template 2h8h.1.A</p>			
<p>Rs1412 626612</p>	<p>P136R</p>		
<p>Rs1862 07963</p>	<p>D407N</p>		
<p>Rs1171 125750</p>	<p>T460K</p>		
<p>Rs7566 65933</p>	<p>P487L</p>		





2.2 Modélisation par Phyre2 :

Après avoir modélisé les structures tridimensionnelles sauvages et mutées de CTLA-4 et c-SRC, nous avons réalisé des alignements entre les structures mutées et sauvages de chaque protéine à l'aide de l'outil TM align.

Nous avons obtenu des scores TM supérieurs à 0,9 pour CTLA-4 et SRC, ce qui indique que les structures sont considérées comme ayant une correspondance presque parfaite. De plus, le RMSD était inférieur à 1 angstrom, ce qui suggère que les structures sont très similaires avec une correspondance précise au niveau atomique.

Tableau 3.9 : Alignement des structures tridimensionnelles modélisées par Phyre2 en utilisant TM align

SNP Id	Mutation	TMalign	
		Tm Score	RMSD
d1ah1a_ Immunoglobulin-like beta-sandwich			
rs781579729	T72I	0.96899	0
rs1356678649	N145H	0.96899	0
rs1204026047	N145S	0.96899	0
rs1466152724	T147A	0.96899	0
c2h8hA_ src kinase in complex with a quinazoline inhibitor			
Rs1412626612	P136R	0.99535	0.09
Rs186207963	D407N	0.99535	0.09
Rs1171125750	T460K	0.99524	0.12
Rs756665933	P487L	0.99535	0.09
Rs13053388441	R509W	0.99535	0.09
Rs1346524802	P510A	0.99428	0.32
Rs1346524802	P510S	0.99535	0.09
Rs1366542593	G173V	0.99535	0.09
Rs761890358	F223L	0.99535	0.09
Rs750534584	A296S	0.99535	0.09
Rs1274322227	A330D	0.99428	0.32
Rs754968435	M344L	0.99428	0.32
Rs1490969287	T432M	0.99535	0.09
Rs1344423676	M498L	0.99524	0.12

Chapitre IV :

Discussion

IV- DISCUSSION

Les lymphocytes T sont des cellules importantes pour la surveillance et l'intervention lors de réactions infectieuses et tumorales. Pour optimiser les signaux transmis, ces cellules possèdent des récepteurs de haute spécificité ainsi que des molécules de co-stimulation telles que CTLA4 (Al-Harbi *et al.*, 2023). Cette molécule joue un rôle clé dans la régulation du système immunitaire et est considérée comme une molécule inhibitrice de fonction anti-tumorale (Olive *et al.*, 2011). De même, SRC est connu comme l'un des premiers oncogènes identifiés avec un rôle essentiel dans plusieurs voies cellulaires clés liées à la prolifération, la différenciation, la survie, la mobilité et l'angiogenèse (Finn, 2008).

Les molécules CTLA-4 et SRC jouent un rôle crucial dans la régulation du système immunitaire et la transmission de signaux cellulaires. En raison de leur importance, ces molécules sont susceptibles d'être affectées par des mutations pouvant entraîner une dysfonction. Dans notre étude, nous avons particulièrement étudié les mutations de type nsSNP qui ont le pouvoir d'induire la cancérogenèse colique.

L'année dernière, BOUKENADEL et MEGNAFI ont identifié plusieurs SNP à conséquences fonctionnelles dans les gènes CTLA-4 (T72I, N145H, N145S, T147A) et SRC (P136R, D407N, T460K, P487L, R509W, P510A, P510S, G173V, F223L, A296S, A330D, M344L, T432M, M498L).

Dans cette étude présentée ici-même, nous avons effectué une analyse *in silico* des modifications post-traductionnelles qui peuvent être engendrées ou supprimées par ces variations ainsi que sur la modélisation moléculaire de ces protéines

Nous avons utilisé 11 outils bioinformatiques pour prédire les modifications post-traductionnelles des protéines d'intérêt. Nos analyses ont fourni les informations suivantes : la phosphorylation a été détectée dans les séquences sauvages de CTLA-4 aux positions T72 et T147, mais ces phosphorylations ont disparu après les mutations (T72I et T147A). En revanche, SRC a montré une phosphorylation dans la position P510 après la mutation P510S. Nous avons également détecté l'apparition d'une méthylation à la position P136 R dans c-SRC après mutation.

Pour les domaines de liaison aux phosphoprotéines quatre modifications ont été détectées : les mutations dans CTLA-4 ont été responsables de la disparition des sites de liaison aux phosphoprotéines (T72, T147) et l'apparition d'un site de liaison aux phosphoprotéines (N145S). Pour c-SRC, une seule mutation a été responsable de l'apparition d'un site de liaison aux phosphoprotéines (A296S). La carbonylation a été localisée sur la protéine CTLA-4 sauvage en

position T72, ainsi que sur c-SRC mutés en position T460K.

Cependant, aucune ubiquitination, tyrosine nitration ou S-nitrosylation significative n'a été observée. Ces PTM peuvent influencer considérablement l'activité régulatrice et structurelle des protéines étudiées.

Pour la modélisation moléculaire, nous avons choisi deux serveurs bioinformatiques : SWISS-Model et Phyre2 pour obtenir les structures 3D de nos protéines sauvages et mutées en utilisant des modèles structurels "templates" pour modéliser nos structures sur la même base. Pour CTLA-4 nous avons choisis le template 3osk.1.B et pour c-SRC nous avons choisis le template 2h8h.1.A. Les templates ont été choisis en fonction des scores dans but d'avoir une modélisation de qualité. Nous avons utilisé ERRAT et PROCHECK pour évaluer la qualité de nos modélisations. Ensuite nous avons utilisé TM-align pour comparer entre les structures tridimensionnelles des protéines sauvages et mutées. Nos résultats ont montré que la mutation T72I pouvait influencer plusieurs PTM tels que la phosphorylation, les domaines de liaison aux phosphoprotéines ainsi que la carbonylation.

En fin, le programme Chimera extensible a été utilisée pour visualiser et analyser interactivement les structures moléculaires. Après superposition des structures mutées sur celles sauvages, une variation structurelle significative a été observée dans les sites fonctionnels des protéines. Cela confirme l'impact significatif des nsSNP dans ces séquences de CTLA-4 (T72I, N145H, N145S et T147A) ainsi que c-SRC (P136R, D407N, T460K, P487L, R509W, P510A, P210S, G173V, F223L, A296S, A330D, M344L, T432M, M498L) sur la configuration structurelle et fonctionnelle de la protéine, ce qui renforce l'hypothèse que ces mutations pourraient altérer les fonctions des protéines CTLA-4 et c-SRC.

Chapitre V : Conclusions et perspectives

V- Chapitre V : Conclusions et perspectives

L'étude actuelle a examiné les effets des SNPs fonctionnels associés aux gènes CTLA-4 et SRC humain par des méthodes computationnelles en raison de la relation de ces gènes avec le cancer colorectal, dont le but était de prédire les modifications post traductionnelle et la réalisation des modélisations moléculaires pour valoriser notre recherche.

Dans un total de 04 nsSNP sur le gène CTLA-4 et 14 sur c-SRC nos outils bioinformatiques de prédiction des modifications poste-traductionnelles (PTM) ont pu détecter 07 nsSNP a effet significatif sur nos protéines d'intérêt dans un total de 02 nsSNP dans CTLA-4 et 05 dans c-SRC. Et on a modélisé les 18 séquences des nsSNP avec un score de qualité positif, ces séquences induit un changement dans la structure après une superposition avec la séquence sauvage.

Notre travaille ouvre un autre intervalle de recherche pour les autres PTM no encore cités, pour les autres mutations des 02 séquences protéiques puisqu'il est un domaine de recherche vaste non limité

Ces résultats fournissent des informations utiles pour une meilleure compréhension de la pathogenèse du cancer colorectal et pourraient éventuellement conduire au développement de nouvelles stratégies thérapeutiques ciblées basées sur ces cibles spécifiques.

Chapitre VI :

La bibliographie

VI- Chapitre VI : La bibliographie

Abd El-Fattah, AA, Sadik, NAH, Shaker, OG, and Mohamed Kamal, A (2018). Single Nucleotide Polymorphism in SMAD7 and CHI3L1 and Colorectal Cancer Risk. *Mediators of Inflammation* 2018, 1–23.

Abes, F, Adjaouti, F, and Belaidi, M (2015). Etude rétrospective et prospective du cancer colorectal en général et de la thérapie ciblée impliquée dans son traitement en particulier.

Agnes, N Carcinogénèse colique : sous-typage moléculaire des étapes précoces du cancer du côlonde phénotype non-MIN.

Aiello, M, Vella, N, Cannavò, C, Scalisi, A, Spandidos, DA, Toffoli, G, Buonadonna, A, Libra, M, and Stivala, F (2011). Role of genetic polymorphisms and mutations in colorectal cancer therapy (Review). *Mol Med Rep* 4, 203–208.

Al-Harbi, N, Abdulla, M-H, Vaali-Mohammed, M-A, Bin Traiki, T, Alswayed, M, Al-Obeed, O, Abid, I, Al-Omar, S, and Mansour, L (2023). Evidence of Association between CTLA-4 Gene Polymorphisms and Colorectal Cancers in Saudi Patients. *Gènes (Basel)* 14, 874.

Al-Sohaily, S, Biankin, A, Leong, R, Kohonen-Corish, M, and Warusavitarne, J (2012). Molecular pathways in colorectal cancer. *Journal of Gastroenterology and Hepatology* 27, 1423–1431.

Amal Idrissi Janati (2014). La santé orale compromise comme facteur de risque potentiel pour le cancer colorectal sporadique (COLDENT) : une étude de faisabilité | Semantic Scholar. Available at: <https://www.semanticscholar.org/paper/La-sant%C3%A9-orale-compromise-comme-facteur-de-risque-%3A-Janati/3f080167a6ad395ee058236a864b5e68e6010870>. Accessed April 24, 2023.

van Amerongen, R, and Nusse, R (2009). Towards an integrated view of Wnt signaling in development. *Development* 136, 3205–3214.

Aparicio, T (2018). Traitement du cancer colorectal métastatique : ciblage moléculaire. *Colon Rectum* 12, 178–183.

Belhamidi, MS, Sinaa, M, Kaoukabi, A, Krimou, H, Menfaa, M, Sakit, F, and Choho, A (2018). Profil épidémiologique et anatomopathologique du cancer colorectal: à propos de 36 cas. *Pan Afr Med J* 30.

Benamrouz, S (2012). Infection par *Cryptosporidium* spp. du modèle souris SCID traité à la dexaméthasone : caractérisation cellulaire et moléculaire du processus de cancérisation des épithéliums digestifs.

Biswas, AK, Noman, N, and Sikder, AR (2010). Machine learning approach to predict protein phosphorylation sites by incorporating evolutionary information. *BMC Bioinformatics* 11, 273.

Boggon, TJ, and Eck, MJ (2004). Structure and régulation of Src family kinases. *Oncogene* 23, 7918–7927.

Boominathan, VP, and Ferreira, T (2016). Predicting tbx22 Zebrafish Protein Structure Using Multi-Level Prédiction Tools and Demonstration of Conserved Structural Domains in Relation to Orthologous tbx22 Proteins in Humans. *JBM* 04, 79–92.

- Brown, MT, and Cooper, JA (1996). Régulation, substrates and functions of src. *Biochim Biophys Acta* 1287, 121–149.
- Brunet, J-F, Denizot, F, Luciani, M-F, Roux-Dosseto, M, Suzan, M, Mattei, M-G, and Golstein, P (1987). A new member of the immunoglobulin superfamily—CTLA-4. *Nature* 328, 267–270.
- Buc, E (2012). Piste infectieuse et carcinogenèse colique : implication des *Escherichia coli* associés à la muqueuse. phdthesis. Université d’Auvergne - Clermont-Ferrand I.
- Burnett-Hartman, AN, Lee, JK, Demb, J, and Gupta, S (2021). An update on the epidemiology, molecular characterization, diagnosis, and screening strategies for early-onset colorectal cancer. *Gastroenterology* 160, 1041–1049.
- Cancer.org Immunotherapy for Colorectal Cancer | Immunotherapy for Rectal Cancer. Available at: <https://www.cancer.org/cancer/types/colon-rectal-cancer/treating/immunotherapy.html>. Accessed May 24, 2023.
- Cartegni, L, Chew, SL, and Krainer, AR (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 3, 285–298.
- Chen, J, Elfiky, A, Han, M, Chen, C, and Saif, MW (2014). The role of Src in colon cancer and its therapeutic implications. *Clin Colorectal Cancer* 13, 5–13.
- Chen, M et al. (2023). GPS 6.0: an updated server for prédiction of kinase-specific phosphorylation sites in proteins. *Nucleic Acids Research*, gkad383.
- Chen, Q, Zhou, Z, Shan, L, Zeng, H, Hua, Y, and Cai, Z (2015). The importance of Src signaling in sarcoma (Review). *Oncology Letters* 10, 17–22.
- Chikuma, S (2017). CTLA-4, an Essential Immune-Checkpoint for T-Cell Activation. *Curr Top Microbiol Immunol* 410, 99–126.
- Cieśla, P, Cieśla, B, and Cieślak, H naczenie profilaktyki wtórnej w nowotworze jelita grubego.
- Ciombor, KK, Strickler, JH, Bekaii-Saab, TS, and Yaeger, R (2022). BRAF-Mutated Advanced Colorectal Cancer: A Rapidly Changing Therapeutic Landscape. *J Clin Oncol* 40, 2706–2715.
- Cohen, GB, Ren, R, and Baltimore, D (1995). Modular binding domains in signal transduction proteins. *Cell* 80, 237–248.
- Conteduca, V, Sansonno, D, Russi, S, and Dammacco, F (2013). Precancerous colorectal lesions (Review). *International Journal of Oncology* 43, 973–984.
- Cordovado, SK, Hendrix, M, Greene, CN, Mochal, S, Earley, MC, Farrell, PM, Kharrazi, M, Hannon, WH, and Mueller, PW (2012). CFTR mutation analysis and haplotype associations in CF patients. *Mol Genet Metab* 105, 249–254.
- Corpas, FJ, Chaki, M, Leterrier, M, and Barroso, JB (2009). Protein tyrosine nitration. *Plant Signal Behav* 4, 920–923.
- Currais, P, Rosa, I, and Claro, I (2022). Colorectal cancer carcinogenesis: From bench to bedside. *World J Gastrointest Oncol* 14, 654–663.

- Curtis, A, Yu, Y, Carey, M, Parfrey, P, Yilmaz, YE, and Savas, S (2022). Examining SNP-SNP interactions and risk of clinical outcomes in colorectal cancer using multifactor dimensionality reduction based methods. *Front Genet* 13, 902217.
- Dabhi, B, and Mistry, KN (2014). In silico analysis of single nucleotide polymorphism (SNP) in human TNF- α gene. *Meta Gene* 2, 586–595.
- DeLalio, LJ et al. (2019a). Constitutive SRC-mediated phosphorylation of pannexin 1 at tyrosine 198 occurs at the plasma membrane. *J Biol Chem* 294, 6940–6956.
- DeLalio, LJ et al. (2019b). Constitutive SRC-mediated phosphorylation of pannexin 1 at tyrosine 198 occurs at the plasma membrane. *J Biol Chem* 294, 6940–6956.
- Derakhshani, A et al. (2021). Cytotoxic T-Lymphocyte Antigen-4 in Colorectal Cancer: Another Therapeutic Side of Capecitabine. *Cancers (Basel)* 13, 2414.
- Desjardin, M (2014). Facteurs de risque et impact clinique de l'hépatotoxicité des chimiothérapies dans le cancer colorectal métastatique.
- Dum, D et al. (2022). Semi-automated validation and quantification of CTLA-4 in 90 different tumor entities using multiple antibodies and artificial intelligence. *Lab Invest* 102, 650–657.
- El Housse, H, Ajbara, W, Amsaguine, S, El Amrani, N, Drissi, H, Ahallat, M, and Radallah, D (2015). Profils épidémiologique et anatomoclinique d'une population marocaine atteinte de cancer colorectal. *J Afr Cancer* 7, 95–99.
- Fang, JY, and Richardson, BC (2005). The MAPK signalling pathways and colorectal cancer. *Lancet Oncol* 6, 322–327.
- Fiegle, E, Doleschel, D, Koletnik, S, Rix, A, Weiskirchen, R, Borkham-Kamphorst, E, Kiessling, F, and Lederle, W (2019). Dual CTLA-4 and PD-L1 Blockade Inhibits Tumor Growth and Liver Metastasis in a Highly Aggressive Orthotopic Mouse Model of Colon Cancer. *Neoplasia* 21, 932–944.
- Finn, RS (2008). Targeting Src in breast cancer. *Annals of Oncology* 19, 1379–1386.
- Forster, S, and Radpour, R (2020). Molecular Immunotherapy: Promising Approach to Treat Metastatic Colorectal Cancer by Targeting Resistant Cancer Cells or Cancer Stem Cells. *Front Oncol* 10, 569017.
- Foster, MW, Hess, DT, and Stamler, JS (2009). Protein S-nitrosylation in health and disease: a current perspective. *Trends Mol Med* 15, 391–404.
- Gafni, A, Dite, GS, Spaeth Tuff, E, Allman, R, and Hopper, JL (2021). Ability of known colorectal cancer susceptibility SNPs to predict colorectal cancer risk: A cohort study within the UK Biobank. *PLoS One* 16, e0251469.
- Gavert, N et al. (2022). Ex vivo organotypic cultures for synergistic therapy prioritization identify patient-specific responses to combined MEK and Src inhibition in colorectal cancer. *Nat Cancer* 3, 219–231.
- Gerolin, M (2017). Modélisation de la dynamique dans les protéines au niveau moléculaire,

approches théoriques et méthodologiques dans une perspective expérimentale.

Grady, WM, and Pritchard, CC (2014). Molecular alterations and biomarkers in colorectal cancer. *Toxicol Pathol* 42, 124–139.

Grandval, P (2020). Diagnostic et surveillance des patients à très haut risque génétique du cancer colo rectal. *GASTRO ENTÉROLOGIE*.

Grasso, C et al. (2018). Abstract B011: Immunoediting in untreated mismatch repair deficient colorectal cancer. *Molecular Cancer Therapeutics* 17, B011.

Grillo, MA, and Colombatto, S (2005). S-adenosylmethionine and protein methylation. *Amino Acids* 28, 357–362.

Grimsrud, PA, Xie, H, Griffin, TJ, and Bernlohr, DA (2008). Oxidative stress and covalent modification of protein with bioactive aldehydes. *J Biol Chem* 283, 21837–21841.

Guarino, M (2010). Src signaling in cancer invasion. *Journal of Cellular Physiology* 223, 14–26.

Gul, M, Navid, A, and Rashid, S (2023). Structural basis of constitutive c-Src kinase activity due to R175L and W118A mutations. *Journal of Biomolecular Structure and Dynamics* 41, 634–645.

Gulzar, N, Dingerdissen, H, Yan, C, and Mazumder, R (2017). Impact of Nonsynonymous Single-Nucleotide Variations on Post-Translational Modification Sites in Human Proteins. *Methods Mol Biol* 1558, 159–190.

Guo, Y, Ning, W, Jiang, P, Lin, S, Wang, C, Tan, X, Yao, L, Peng, D, and Xue, Y (2020). GPS-PBS: A Deep Learning Framework to Predict Phosphorylation Sites that Specifically Interact with Phosphoprotein-Binding Domains. *Cells* 9, 1266.

Hanahan, D (2022). Hallmarks of Cancer: New Dimensions. *Cancer Discovery* 12, 31–46.

Hossain, MS et al. (2022). Colorectal Cancer: A Review of Carcinogenesis, Global Epidemiology, Current Challenges, Risk Factors, Preventive and Treatment Strategies. *Cancers* 14, 1732.

Hosseini, A, Gharibi, T, Marofi, F, Babaloo, Z, and Baradaran, B (2020). CTLA-4: From mechanism to autoimmune therapy. *International Immunopharmacology* 80, 106221.

Huang, L, Guo, Z, Wang, F, and Fu, L (2021). KRAS mutation: from undruggable to druggable in cancer. *Signal Transduct Target Ther* 6, 386.

Huels, DJ, and Sansom, OJ (2015). Stem vs non-stem cell origin of colorectal cancer. *Br J Cancer* 113, 1–5.

Imad, FE, Drissi, H, Tawfiq, N, Bendahhou, K, Benider, A, and Radallah, D (2020). Facteurs de risque alimentaires du cancer colorectal au Maroc: étude cas témoin. *Pan Afr Med J* 35, 59.

Imazeki, H, Ogiwara, Y, Kawamura, M, Boku, N, and Kudo-Saito, C (2021). Original research: CD11b+CTLA4+ myeloid cells are a key driver of tumor evasion in colorectal cancer. *Journal for Immunotherapy of Cancer* 9.

imWechsel.com Quest-ce qu'un polymorphisme nucléotidique unique (SNP) et quel est son impact

sur la pharmacothérapie? | Wechsel.

Irby, RB, and Yeatman, TJ (2000). Role of Src expression and activation in human cancer. *Oncogene* 19, 5636–5642.

Irla, M et al. (2008). Autoantigen-Specific Interactions with CD4+ Thymocytes Control Mature Medullary Thymic Epithelial Cell Cellularity. *Immunity* 29, 451–463.

Jenkins, MA et al. (2016). Quantifying the utility of single nucleotide polymorphisms to guide colorectal cancer screening. *Future Oncology* 12, 503–513.

Jin, W (2020a). Regulation of Src Family Kinases during Colorectal Cancer Development and Its Clinical Implications. *Cancers (Basel)* 12, 1339.

Jin, W (2020b). Regulation of Src Family Kinases during Colorectal Cancer Development and Its Clinical Implications. *Cancers (Basel)* 12, 1339.

Johnson, CM, Wei, C, Ensor, JE, Smolenski, DJ, Amos, CI, Levin, B, and Berry, DA (2013). Meta-analyses of colorectal cancer risk factors. *Cancer Causes Control* 24, 1207–1222.

K. S. Salama, F. Hodi (2011). Molecular Pathways Cytotoxic T-Lymphocyte – Associated Antigen-4.

Keum, N, and Giovannucci, E (2019). Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat Rev Gastroenterol Hepatol* 16, 713–732.

Kim, S, and Misra, A (2007). SNP genotyping: technologies and biomedical applications. *Annu Rev Biomed Eng* 9, 289–320.

Kimchi-Sarfaty, C, Oh, JM, Kim, I-W, Sauna, ZE, Calcagno, AM, Ambudkar, SV, and Gottesman, MM (2007). A “silent” polymorphism in the MDR1 gene changes substrate specificity. *Science* 315, 525–528.

Klaunig, JE (2020). Carcinogenesis. In: *An Introduction to Interdisciplinary Toxicology*, Elsevier, 97–110.

Koehne, C-H, and Dubois, RN (2004). COX-2 inhibition and colorectal cancer. *Seminars in Oncology* 31, 12–21.

Korzeniewski, S, Hofman, P, and Brest, P (2013). Des polymorphismes silencieux plutôt bruyants. *Med Sci (Paris)* 29, 124–126.

La Vecchia, S, and Sebastián, C (2020). Metabolic pathways regulating colorectal cancer initiation and progression. *Seminars in Cell & Developmental Biology* 98, 63–70.

Laskowski, RA, MacArthur, MW, and Thornton, JM (2006). PROCHECK: validation of protein-structure coordinates. urn:isbn:978-0-7923-6857-1. Available at: <https://onlinelibrary.wiley.com/iucr/itc/Fa/ch25o2v0001/sec25o2o6/>. Accessed June 19, 2023.

Lennartsson, J, Blume-Jensen, P, Hermanson, M, Pontén, E, Carlberg, M, and Rönstrand, L (1999a). Phosphorylation of Shc by Src family kinases is necessary for stem cell factor receptor/c-kit mediated activation of the Ras/MAP kinase pathway and c-fos induction. *Oncogene* 18, 5546–

5553.

Lennartsson, J, Blume-Jensen, P, Hermanson, M, Pontén, E, Carlberg, M, and Rönstrand, L (1999b). Phosphorylation of Shc by Src family kinases is necessary for stem cell factor receptor/c-kit mediated activation of the Ras/MAP kinase pathway and c-fos induction. *Oncogene* 18, 5546–5553.

Levin, B et al. (2008). Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology* 134, 1570–1595.

Li, X-L, Zhou, J, Chen, Z-R, and Chng, W-J (2015). p53 mutations in colorectal cancer- molecular pathogenesis and pharmacological reactivation. *World J Gastroenterol* 21, 84–93.

Lindsten, T, Lee, KP, Harris, ES, Petryniak, B, Craighead, N, Reynolds, PJ, Lombard, DB, Freeman, GJ, Nadler, LM, and Gray, GS (1993). Characterization of CTLA-4 structure and expression on human T cells. *The Journal of Immunology* 151, 3489–3499.

Liu, S et al. (2020). CTLA4 has a profound impact on the landscape of tumor-infiltrating lymphocytes with a high prognosis value in clear cell renal cell carcinoma (ccRCC). *Cancer Cell International* 20, 519.

Liu, Z, Cao, J, Ma, Q, Gao, X, Ren, J, and Xue, Y (2011). GPS-YNO2: computational prediction of tyrosine nitration sites in proteins. *Mol BioSyst* 7, 1197.

Loubna Allam (2020). Conception et évaluation des molécules anticancéreuses par approches in silico.

Lu, Y-F, Mauger, DM, Goldstein, DB, Urban, TJ, Weeks, KM, and Bradrick, SS (2015). IFNL3 mRNA structure is remodeled by a functional non-coding polymorphism associated with hepatitis C virus clearance. *Sci Rep* 5, 16037.

Luttrell, DK, and Luttrell, LM (2004). Not so strange bedfellows: G-protein-coupled receptors and Src family kinases. *Oncogene* 23, 7969–7978.

Mah, JTL, Low, ESH, and Lee, E (2011). In silico SNP analysis and bioinformatics tools: a review of the state of the art to aid drug discovery. *Drug Discov Today* 16, 800–809.

Malek, RL et al. (2002a). Identification of Src transformation fingerprint in human colon cancer. *Oncogene* 21, 7256–7265.

Malek, RL et al. (2002b). Identification of Src transformation fingerprint in human colon cancer. *Oncogene* 21, 7256–7265.

Malik, A, Pande, K, Kumar, A, Vemula, A, Madhuri, R, and Chandramohan, V (2020). Finding Pathogenic nsSNP's and their structural effect on COPS2 using Molecular Dynamic Approach, Biophysics.

Manning, G, Whyte, DB, Martinez, R, Hunter, T, and Sudarsanam, S (2002). The protein kinase complement of the human genome. *Science* 298, 1912–1934.

- Martin, TD, Samuel, JC, Routh, ED, Der, CJ, and Yeh, JJ (2011). Activation and involvement of Ral GTPases in colorectal cancer. *Cancer Res* 71, 206–215.
- Matsuda, K (2017). PCR-Based Detection Methods for Single-Nucleotide Polymorphism or Mutation. In: *Advances in Clinical Chemistry*, Elsevier, 45–72.
- Mayer, BJ, and Eck, MJ (1995). SH3 domains. Minding your p's and q's. *Curr Biol* 5, 364–367.
- McCarthy, C, Carrea, A, and Diambra, L (2017). Bicodon bias can determine the role of synonymous SNPs in human diseases. *BMC Genomics* 18, 227.
- MedlinePlus Genetics What are single nucleotide polymorphisms (SNPs)? : MedlinePlus Genetics. Available at: <https://medlineplus.gov/genetics/understanding/genomicresearch/snp/>. Accessed May 14, 2023.
- Misselwitz, B (2017). Le cancer colorectal. 2017 17:44 17, 943–952.
- Molineris, I, Schiavone, D, Rosa, F, Matullo, G, Poli, V, and Provero, P (2013). Identification of functional cis-regulatory polymorphisms in the human genome. *Hum Mutat* 34, 735–742.
- Morris, EJA, Penegar, S, Whiffin, N, Broderick, P, Bishop, DT, Northwood, E, Quirke, P, Finan, P, and Houlston, RS (2015). A Retrospective Observational Study of the Relationship between Single Nucleotide Polymorphisms Associated with the Risk of Developing Colorectal Cancer and Survival. *PLoS ONE* 10, e0117816.
- Nackley, AG, Shabalina, SA, Tchivileva, IE, Satterfield, K, Korchynskiy, O, Makarov, SS, Maixner, W, and Diatchenko, L (2006). Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. *Science* 314, 1930–1933.
- Najdi, R, Holcombe, RF, and Waterman, ML (2011). Wnt signaling and colon carcinogenesis: beyond APC. *J Carcinog* 10, 5.
- Narayanan, V, Pavithra, V, Dhanapal, D, Sundaram, S, and Narayanan, CD (2022). Role of CTLA4 immunohistochemistry in the diagnosis of colon cancers. *Indian J Pathol Microbiol* 65, 791–795.
- Naudin, C (2012). Régulation de la signalisation oncogénique de Src par l'adaptateur SLAP dans les cellules de cancer colorectal. These de doctorat. Montpellier 2.
- Ng, PC, and Henikoff, S (2006). Predicting the Effects of Amino Acid Substitutions on Protein Function. *Annu Rev Genom Hum Genet* 7, 61–80.
- Nguyen, LH, Goel, A, and Chung, DC (2020). Pathways of Colorectal Carcinogenesis. *Gastroenterology* 158, 291–302.
- Niccolai, E et al. (2017). The Different Functional Distribution of “Not Effector” T Cells (Treg/Tnull) in Colorectal Cancer. *Front Immunol* 8, 1900.
- Nikiteas, NI, Weber, M, Antoniou, A, and Dimitroulis, D (2016). *Advances in Colorectal Cancer*. *Biomed Res Int* 2016, 7347648.
- Olive, D, Thi, S le, Xerri, L, Hirsch, I, and Nunès, JA (2011). Rôle de CTLA-4 dans la

- cosignalisation négative du système immunitaire. *Med Sci (Paris)* 27, 842–849.
- Olivier, S, Mir, A-M, Michalski, J-C, and Lefebvre, T (2011). Signalisation et prédispositions métaboliques liées au cancer colorectal. *Med Sci (Paris)* 27, 514–520.
- Olsson, JB et al. (2022). Colorectal cancer-associated SNP rs17042479 is involved in the regulation of NAF1 promoter activity. *PLoS One* 17, e0274033.
- Papadatos-Pastos, D, Rabbie, R, Ross, P, and Sarker, D (2015). The role of the PI3K pathway in colorectal cancer. *Crit Rev Oncol Hematol* 94, 18–30.
- Parsons, SJ, and Parsons, JT (2004a). Src family kinases, key regulators of signal transduction. *Oncogene* 23, 7906–7909.
- Parsons, SJ, and Parsons, JT (2004b). Src family kinases, key regulators of signal transduction. *Oncogene* 23, 7906–7909.
- Perkins, D et al. (1996). Regulation of CTLA-4 expression during T cell activation. *J Immunol* 156, 4154–4159.
- Picard-Croguennec, M (2018). Le cancer colorectal, physiopathologie et principaux symptômes. *Actualités Pharmaceutiques* 57, 22–23.
- Platz, EA, Willett, WC, Colditz, GA, Rimm, EB, Spiegelman, D, and Giovannucci, E (2000). Proportion of colon cancer risk that might be preventable in a cohort of middle-aged US men. *Cancer Causes and Control* 11, 579–588.
- Playford, MP, and Schaller, MD (2004a). The interplay between Src and integrins in normal and tumor biology. *Oncogene* 23, 7928–7946.
- Ray, P, Raghunathan, K, Ahsan, A, Allam, US, Shukla, S, Basrur, V, Veatch, S, Lawrence, TS, Nyati, MK, and Ray, D (2020). Mutant EGFR is a preferred SMURF2 substrate of ubiquitination: role in enhanced receptor stability and TKI sensitivity. 2020.04.02.022012.
- Ren, Z, and Tao, Z (2018). Molecular Basis of Colorectal Cancer: Tumor Biology. In: *Surgical Treatment of Colorectal Cancer*, ed. NK Kim, K Sugihara, and J-T Liang, Singapore: Springer Singapore, 23–34.
- Resh, MD (1993). Interaction of tyrosine kinase oncoproteins with cellular membranes. *Biochim Biophys Acta* 1155, 307–322.
- Reya, T, and Clevers, H (2005). Wnt signalling in stem cells and cancer. *Nature* 434, 843–850.
- Rickles, RJ, Botfield, MC, Zhou, XM, Henry, PA, Brugge, JS, and Zoller, MJ (1995). Phage display selection of ligand residues important for Src homology 3 domain binding specificity. *Proc Natl Acad Sci U S A* 92, 10909–10913.
- Roberts, JZ, Crawford, N, and Longley, DB (2022). The role of Ubiquitination in Apoptosis and Necroptosis. *Cell Death Differ* 29, 272–284.
- Roshan, A, and Byrne, MF (2022). Artificial intelligence in colorectal cancer screening. *CMAJ* 194, E1481–E1484.

- Rosik, J, Szostak, B, Machaj, F, and Pawlik, A (2021). The Role of CTLA4 and Its Polymorphisms in Solid Organ and Haematopoietic Stem Cell Transplantation. *Int J Mol Sci* 22, 3081.
- Roskoski, R (2004). Src protein-tyrosine kinase structure and regulation. *Biochem Biophys Res Commun* 324, 1155–1164.
- Ruiz-Saenz, A et al. (2023). A reversible SRC-relayed COX2 inflammatory program drives resistance to BRAF and EGFR inhibition in BRAFV600E colorectal tumors. *Nat Cancer* 4, 240–256.
- Sacks, DB, Bruns, DE, Goldstein, DE, Maclaren, NK, McDonald, JM, and Parrott, M (2002). Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. *Clin Chem* 48, 436–472.
- Salama, KS, and Hodi, F (2011). Molecular Pathways Cytotoxic T-Lymphocyte – Associated Antigen-4 April.
- Santos, AL, and Lindner, AB (2017). Protein Posttranslational Modifications: Roles in Aging and Age-Related Disease. *Oxid Med Cell Longev* 2017, 5716409.
- Sarbassov, DD, Guertin, DA, Ali, SM, and Sabatini, DM (2005). Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* 307, 1098–1101.
- Schaller, MD, Bouton, AH, Flynn, DC, and Parsons, JT (1993). Identification and characterization of novel substrates for protein tyrosine kinases. *Prog Nucleic Acid Res Mol Biol* 44, 205–227.
- Schaub, MA, Boyle, AP, Kundaje, A, Batzoglou, S, and Snyder, M (2012). Linking disease associations with regulatory information in the human genome. *Genome Res* 22, 1748–1759.
- Sell, T, Klotz, C, Fischer, MM, Astaburuaga-García, R, Krug, S, Drost, J, Clevers, H, Sers, C, Morkel, M, and Blüthgen, N (2023). Oncogenic signaling is coupled to colorectal cancer cell differentiation state. *Journal of Cell Biology* 222, e202204001.
- Seykora, JT, Mei, L, Dotto, GP, and Stein, PL (2002). 'Srcasm: a novel Src activating and signaling molecule. *J Biol Chem* 277, 2812–2822.
- Shan, J, Han, D, Shen, C, Lei, Q, and Zhang, Y (2022). Mechanism and strategies of immunotherapy resistance in colorectal cancer. *Front Immunol* 13, 1016646.
- Sheng, J, Sun, H, Yu, F-B, Li, B, Zhang, Y, and Zhu, Y-T (2020). The Role of Cyclooxygenase-2 in Colorectal Cancer. *Int J Med Sci* 17, 1095–1101.
- Sherry, ST, Ward, MH, Kholodov, M, Baker, J, Phan, L, Smigielski, EM, and Sirotkin, K (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 29, 308–311.
- Shupnik, MA (2004a). Crosstalk between steroid receptors and the c-Src-receptor tyrosine kinase pathways: implications for cell proliferation. *Oncogene* 23, 7979–7989.
- Simon, K (2016). Colorectal cancer development and advances in screening. *Clinical Interventions in Aging* 11, 967–976.

- Sirvent, A, Bénistant, C, Pannequin, J, Veracini, L, Simon, V, Bourgaux, J-F, Hollande, F, Cruzalegui, F, and Roche, S (2010). Src family tyrosine kinases-driven colon cancer cell invasion is induced by Csk membrane delocalization. *Oncogene* 29, 1303–1315.
- Sirvent, A, Urbach, S, and Roche, S (2014). Analyse *in vivo* de la signalisation tumorale induite par les tyrosine-kinases par protéomique quantitative. *Med Sci (Paris)* 30, 558–566.
- Sobhani, N, Tardiel-Cyril, DR, Davtyan, A, Generali, D, Roudi, R, and Li, Y (2021). CTLA-4 in Regulatory T Cells for Cancer Immunotherapy. *Cancers* 13, 1440.
- Song, K, and Li, S (2021). The Role of Ubiquitination in NF- κ B Signaling during Virus Infection. *Viruses* 13.
- Song, L-L, and Li, Y-M (2016). Current noninvasive tests for colorectal cancer screening: An overview of colorectal cancer screening tests. *World J Gastrointest Oncol* 8, 793–800.
- Songyang, Z, Shoelson, SE, Chaudhuri, M, Gish, G, Pawson, T, Haser, WG, King, F, Roberts, T, Ratnofsky, S, and Lechleider, RJ (1993). SH2 domains recognize specific phosphopeptide sequences. *Cell* 72, 767–778.
- Sorgen, PL, Duffy, HS, Sahoo, P, Coombs, W, Delmar, M, and Spray, DC (2004a). Structural changes in the carboxyl terminus of the gap junction protein connexin43 indicates signaling between binding domains for c-Src and zonula occludens-1. *J Biol Chem* 279, 54695–54701.
- Sorgen, PL, Duffy, HS, Sahoo, P, Coombs, W, Delmar, M, and Spray, DC (2004b). Structural changes in the carboxyl terminus of the gap junction protein connexin43 indicates signaling between binding domains for c-Src and zonula occludens-1. *J Biol Chem* 279, 54695–54701.
- Stanton, VP (1999). Single-nucleotide polymorphisms can cause different structural folds of mRNA. *Proc Natl Acad Sci U S A* 96, 7871–7876.
- Stintzing, S (2014). Management of colorectal cancer. *F1000Prime Rep* 6, 108.
- Syvänen, A-C (2001). Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nat Rev Genet* 2, 930–942.
- T, G, M, S, Z, L, Da, F, Sg, K, and La, F (2000). An EGF receptor/Ral-GTPase signaling cascade regulates c-Src activity and substrate specificity. *The EMBO Journal* 19.
- Tak, YG, Yao, L, Berman, BP, and Farnham, PJ (2014). Functional annotation of colon cancer risk SNPs. *Nat Commun* 5, 5114.
- Teft, WA, Kirchhof, MG, and Madrenas, J (2006). A MOLECULAR PERSPECTIVE OF CTLA-4 FUNCTION. *Annu Rev Immunol* 24, 65–97.
- Tehami, W, Nani, A, Khan, NA, and Hichami, A (2023). New Insights Into the Anticancer Effects of p-Coumaric Acid: Focus on Colorectal Cancer. *Dose Response* 21, 15593258221150704.
- Thicoïpé, A (2015). Maladies Inflammatoires Chroniques Intestinales et Cancer colorectal: facteurs de risque et résultats ontologiques à long terme.
- Thomas, SM, and Brugge, JS (1997). Cellular functions regulated by Src family kinases. *Annu*

Rev Cell Dev Biol 13, 513–609.

Tian, X, Yan, L, Zhang, D, Guan, X, Dong, B, Zhao, M, and Hao, C (2016). PTK7 overexpression in colorectal tumors: Clinicopathological correlation and prognosis relevance. *Oncology Reports* 36, 1829–1836.

Turko, IV, and Murad, F (2002). Protein nitration in cardiovascular diseases. *Pharmacol Rev* 54, 619–634.

The UniProt Consortium (2017). UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 45, D158–D169.

Van Coillie, S, Wiernicki, B, and Xu, J (2020). Molecular and Cellular Functions of CTLA-4. In: *Regulation of Cancer Immune Checkpoints*, ed. J Xu, Singapore: Springer Singapore, 7–32.

Vatsa, S (2022a). In silico prediction of post-translational modifications in therapeutic antibodies. *MAbs* 14, 2023938.

Vatsa, S (2022b). In silico prediction of post-translational modifications in therapeutic antibodies. *MAbs* 14, 2023938.

Vega, P, Valentín, F, and Cubiella, J (2015). Colorectal cancer diagnosis: Pitfalls and opportunities. *World J Gastrointest Oncol* 7, 422–433.

Venkata Subbiah, H, Ramesh Babu, P, and Subbiah, U (2020). In silico analysis of non-synonymous single nucleotide polymorphisms of human DEFB1 gene. *Egyptian Journal of Medical Human Genetics* 21, 66.

Waldner, M, Schimanski, CC, and Neurath, MF (2006). Colon cancer and the immune system: The role of tumor invading T cells. *World J Gastroenterol* 12, 7233–7238.

Wang, D, Zeng, S, Xu, C, Qiu, W, Liang, Y, Joshi, T, and Xu, D (2017). MusiteDeep: a deep-learning framework for general and kinase-specific phosphorylation site prediction. *Bioinformatics* 33, 3909–3916.

Wehr, NB, and Levine, RL (2013). Quantification of protein carbonylation. *Methods Mol Biol* 965, 265–281.

Wilschut, JA, Steyerberg, EW, van Leerdam, ME, Lansdorp-Vogelaar, I, Habbema, JDF, and van Ballegooijen, M (2011). How much colonoscopy screening should be recommended to individuals with various degrees of family history of colorectal cancer? *Cancer* 117, 4166–4174.

Wilson, DF (2017). Oxidative phosphorylation: regulation and role in cellular and tissue metabolism. *J Physiol* 595, 7023–7038.

Worthley, DL (2007). Colorectal carcinogenesis: Road maps to cancer. *WJG* 13, 3784.

Xie, Y et al. (2018). DeepNitro: Prediction of Protein Nitration and Nitrosylation Sites by Deep Learning. *Genomics Proteomics Bioinformatics* 16, 294–306.

Xu, Jie, éd. (2020). *1248 Regulation of Cancer Immune Checkpoints: Molecular and Cellular Mechanisms and Therapy*. Singapore: Springer Singapore.

Xue, Y, Liu, Z, Gao, X, Jin, C, Wen, L, Yao, X, and Ren, J (2010). GPS-SNO: Computational Prediction of Protein S-Nitrosylation Sites with a Modified GPS Algorithm. *PLoS One* 5, e11290.

Y Peterson, C (2015). Risk Factors for Colorectal Cancer. *MOJS* 2.

Yu, CL, Meyer, DJ, Campbell, GS, Lerner, AC, Carter-Su, C, Schwartz, J, and Jove, R (1995). Enhanced DNA-binding activity of a Stat3-related protein in cells transformed by the Src oncoprotein. *Science* 269, 81–83.

Yue, T, Chen, S, Zhu, J, Guo, S, Huang, Z, Wang, P, Zuo, S, and Liu, Y (2021). The aging-related risk signature in colorectal cancer. *Aging (Albany NY)* 13, 7330–7349.

Zhang, H, Lang, Z, and Zhu, J-K (2018). Dynamics and function of DNA methylation in plants. *Nat Rev Mol Cell Biol* 19, 489–506.

Zhang, M, Song, J, Yuan, W, Zhang, W, and Sun, Z (2021). Roles of RNA Methylation on Tumor Immunity and Clinical Implications. *Front Immunol* 12, 641507.

Zhang, Y, and Skolnick, J (2005). TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* 33, 2302–2309.

Zhao, Y, Yang, W, Huang, Y, Cui, R, Li, X, and Li, B (2018). Evolving Roles for Targeting CTLA-4 in Cancer Immunotherapy. *Cell Physiol Biochem* 47, 721–734.

Zoratto, F, Rossi, L, Verrico, M, Papa, A, Basso, E, Zullo, A, Tomao, L, Romiti, A, Lo Russo, G, and Tomao, S (2014). Focus on genetic and epigenetic events of colorectal cancer pathogenesis: implications for molecular diagnosis. *Tumour Biol* 35, 6195–6206.

Zou, C, Qiu, H, Tang, W, Wang, Y, Lan, B, and Chen, Y (2018a). CTLA4 tagging polymorphisms and risk of colorectal cancer: a case–control study involving 2,306 subjects. *Onco Targets Ther* 11, 4609–4619.

Zou, C, Qiu, H, Tang, W, Wang, Y, Lan, B, and Chen, Y (2018b). CTLA4 tagging polymorphisms and risk of colorectal cancer: a case–control study involving 2,306 subjects. *Onco Targets Ther* 11, 4609–4619.

(2012). 16 - Development issues: antibody stability, developability, immunogenicity, and comparability. In: *Therapeutic Antibody Engineering*, ed. WR Strohl, and LM Strohl, Woodhead Publishing, 377–595.

(2021). Cancers du côlon : comment est établi le diagnostic ? - Cancer du côlon. Available at: <https://www.e-cancer.fr/Patients-et-proches/Les-cancers/Cancer-du-colon/Le-diagnostic>. Accessed April 24, 2023.

ERRAT – UCLA-DOE Institute.

Chapitre VII :

ANNEXES

VII- Annexes :

Un tableau avec toutes les séquences mutées en format FASTA

Tableau 7.1 : les séquences FASTA mutées et sauvage

SNP ID	Mutation	Séquence mutée en format FASTA
CTLA4 Sauvage		>sp P16410 CTLA4_HUMAN Cytotoxic T-lymphocyte protein 4 OS=Homo sapiens OX=9606 GN=CTLA4 PE=1 SV=3 MACLGFQRHKAQLNLATRTWPCTLLFFLLFI PVFCKAMHVAQPAVVLASSRGIASFVCEYA SPGKATEVRVTVLRQADSQVTEVCAATYM MGNELTFLDDSICTGTSSGNQVNLTIQGLRA MDTGLYICKVELMYPPPYLIGINGTQIYVI DPEPCPDSDFLLWILAAVSSGLFFYSFLLTAV SLSKMLKKRSPLTTGVYVKMPPTPEPECEKQF QPYFIPIN
rs781 57972 9	T72I	>sp P16410 CTLA4_HUMAN Cytotoxic T-lymphocyte protein 4 OS=Homo sapiens OX=9606 GN=CTLA4 PE=1 SV=3 MACLGFQRHKAQLNLATRTWPCTLLFFLLFI PVFCKAMHVAQPAVVLASSRGIASFVCEYA SPGKATEVRVIVLRQADSQVTEVCAATYMM GNELTFLDDSICTGTSSGNQVNLTIQGLRAM DTGLYICKVELMYPPPYLIGINGTQIYVIDP EPCPDSDFLLWILAAVSSGLFFYSFLLTAVSL SKMLKKRSPLTTGVYVKMPPTPEPECEKQFQ PYFIPIN
rs135 66786 49	N145H	>sp P16410 CTLA4_HUMAN Cytotoxic T-lymphocyte protein 4 OS=Homo sapiens OX=9606 GN=CTLA4 PE=1 SV=3 MACLGFQRHKAQLNLATRTWPCTLLFFLLFI PVFCKAMHVAQPAVVLASSRGIASFVCEYA SPGKATEVRVTVLRQADSQVTEVCAATYM MGNELTFLDDSICTGTSSGNQVNLTIQGLRA MDTGLYICKVELMYPPPYLIGHGTQIYVI DPEPCPDSDFLLWILAAVSSGLFFYSFLLTAV SLSKMLKKRSPLTTGVYVKMPPTPEPECEKQF QPYFIPIN
rs120 40260	N145S	>sp P16410 CTLA4_HUMAN Cytotoxic T-lymphocyte protein 4 OS=Homo sapiens OX=9606 GN=CTLA4 PE=1 SV=3

47		MACLGFQRHKAQLNLATRTWPCTLLFFLLFI PVFCKAMHVAQPAVVLASSRGIASFVCEYA SPGKATEVRVTVLRQADSQVTEVCAATYM MGNELTFLDDSICTGTSSGNQVNLIQGLRA MDTGLYICKVELMYPPPYLGIIGSGTQIYVI DPEPCPDSDFLLWILAAVSSGLFFYSFLLTAV SLSKMLKKRSPLTTGVYVKMPPTPEPECEKQF QPYFIPIN
rs146 61527 24	T147A	>sp P16410 CTLA4_HUMAN Cytotoxic T- lymphocyte protein 4 OS=Homo sapiens OX=9606 GN=CTLA4 PE=1 SV=3 MACLGFQRHKAQLNLATRTWPCTLLFFLLFI PVFCKAMHVAQPAVVLASSRGIASFVCEYA SPGKATEVRVTVLRQADSQVTEVCAATYM MGNELTFLDDSICTGTSSGNQVNLIQGLRA MDTGLYICKVELMYPPPYLGIIGNGAQIYVI DPEPCPDSDFLLWILAAVSSGLFFYSFLLTAV SLSKMLKKRSPLTTGVYVKMPPTPEPECEKQF QPYFIPIN
SRC Sauvage		>sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3 MGSNKSKPKDASQRRRSLEPAENVHGAGGG AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDTVTSPQRAGPLAGGVTTFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTGYIPSNYVAPSDSIQAEWYF GKITRRESERLLLNAENPRGTFLVRESETTKG AYCLSVSDFDNAKGLNVKHYKIRKLDSSGGF YITSRTQFNSLQQLVAYYSKHADGLCHRLTT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMGTWNGTTRVAIKTLKPGTMS PEAFLQEAQVMKKLRHEKLVQLYAVVSEEP IYIVTEYMSKGSLLDFLKGETGKYLRLPQLV DMAAQIASGMAYVERMNYVHRDLRAANIL VGENLVCKVADFGLARLIEDNEYTARQGAK FPIKWTAPEAALYGRFTIKSDVWSFGILLTEL TTKGRVPYPGMVNREVLQDQVERGYRMP CPP ECPESLHDLMCQCWRKEPEERPTFEYLQAFL EDYFTSTEPQYQGENL
Rs141 26266 12	P136R	>sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3 MGSNKSKPKDASQRRRSLEPAENVHGAGGG

		<p>AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDTVTSPQRAGPLAGGVTTTFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTGYIRSNYVAPSDSIQAEEWYF GKITRRESERLLLNAENPRGTFLVRESETTKG AYCLSVSDFDNAKGLNVKHYKIRKLDSSGGF YITSRTQFNSLQQLVAYYSKHADGLCHRLLT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMGTWNGTTRVAIKTL KPGTMSPEAFLQEAQVMKKLRHEKLVQLY AVVSEEPYIVTEYMSKGSLLDFLKGETGKY LRPQLVDMAAQIASGMAYVERMNYVHRD LRAANILVGENLVCKVADFGLARLIEDNEYT ARQGAKFPIKWTAPEAALYGRFTIKSDVWSF GILLTELTTKGRVPYPGMVNREVLDQVERG YRMPCPPECPESLHDLMCQCWRKEPEERPTF EYLQAFLEDYFTSTEPQYQPGENL</p>
Rs186 20796 3	D407N	<p>>sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3 MGSNKSKPKDASQRRRSLEPAENVHGAGGG AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDTVTSPQRAGPLAGGVTTTFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTGYIPSNYVAPSDSIQAEEWYF GKITRRESERLLLNAENPRGTFLVRESETTKG AYCLSVSDFDNAKGLNVKHYKIRKLDSSGGF YITSRTQFNSLQQLVAYYSKHADGLCHRLLT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMGTWNGTTRVAIKTLKPGTMS PEAFLQEAQVMKKLRHEKLVQLYAVVSEEP IYIVTEYMSKGSLLDFLKGETGKYLRPQLV DMAAQIASGMAYVERMNYVHRDLRAANIL VGENLVCKVANFGLARLIEDNEYTARQGAK FPIKWTAPEAALYGRFTIKSDVWSFGILLTEL TTKGRVPYPGMVNREVLDQVERGYRMPCPP ECPESLHDLMCQCWRKEPEERPTFEYLQAFLE DYFTSTEPQYQPGENL</p>
Rs117 11257 50	T460K	<p>>sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3 MGSNKSKPKDASQRRRSLEPAENVHGAGGG AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDTVTSPQRAGPLAGGVTTTFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTGYIPSNYVAPSDSIQAEEWYF GKITRRESERLLLNAENPRGTFLVRESETTKG</p>

		<p> AYCLSVSDFDNAKGLNVKHYKIRKLDSSGGF YITSRTQFNSLQQLVAYYSKHADGLCHRLTT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMGTWNGTTRVAIKTLKPGTMS PEAFLQEAQVMKKLRHEKLVQLYAVVSEEP IYIVTEYMSKGSLLDFLKGETGKYLRPLQLV DMAAQIASGMAYVERMNYVHRDLRAANIL VGENLVCKVADFGFLARLIEDNEYTARQGAK FPIKWTAPEAALYGRFTIKSDVWSFGILLTEL TTKGRVPYPGMVNREVLDQVERGYRMPCL PECPELHDLMCQCWRKEPEERPTFEYLQAF LEDYFTSTEPQYQPGENL </p>
Rs756 66593 3	P487L	<p> >sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3 MGSNKSKPKDASQRRRSLEPAENVHGAGGG AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDVTVSPQRAGPLAGGVTTFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTYIPSNYVAPSDSIQAEWYF GKITRRESERLLLNAENPRGTFLVRESETTKG AYCLSVSDFDNAKGLNVKHYKIRKLDSSGGF YITSRTQFNSLQQLVAYYSKHADGLCHRLTT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMGTWNGTTRVAIKTLKPGTMS PEAFLQEAQVMKKLRHEKLVQLYAVVSEEP IYIVTEYMSKGSLLDFLKGETGKYLRPLQLV DMAAQIASGMAYVERMNYVHRDLRAANIL VGENLVCKVADFGFLARLIEDNEYTARQGAK FPIKWTAPEAALYGRFTIKSDVWSFGILLTEL TTKGRVPYPGMVNREVLDQVERGYRMPCL PECPELHDLMCQCWRKEPEERPTFEYLQAF LEDYFTSTEPQYQPGENL </p>
Rs130	R509W	<p> >sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens </p>

53388 441		OX=9606 GN=SRC PE=1 SV=3 MGSNKSKPKDASQRRRSLEPAENVHGAGGG AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDTVTSPQRAGPLAGGVTTTFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTGYIPSNYVAPSDSIQAEWYF GKITRRESERLLLNAENPRGTFLVRESETTKG AYCLSVSDFDNAKGLNVKHYKIRKLDSSGGF YITSRTQFNSLQQLVAYYSKHADGLCHRLTT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMGTWNGTTRVAIKTLKPGTMS PEAFLQEAQVMKKLRHEKLVQLYAVVSEEP IYIVTEYMSKGSLLDFLKGETGKYLRPLQLV DMAAQIASGMAYVERMNYVHRDLRAANIL VGENLVCKVADFGLARLIEDNEYTARQGAK FPIKWTAPEAALYGRFTIKSDVWSFGILLTEL TTKGRVPYPGMVNREVLDQVERGYRMPCPP ECPESLHDLMCQCWRKEPEEWPTFEYLQAF LEDYFTSTEPQYQPGENL
Rs134 65248 02	P510A	>sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3 MGSNKSKPKDASQRRRSLEPAENVHGAGGG AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDTVTSPQRAGPLAGGVTTTFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTGYIPSNYVAPSDSIQAEWYF GKITRRESERLLLNAENPRGTFLVRESETTKG AYCLSVSDFDNAKGLNVKHYKIRKLDSSGGF YITSRTQFNSLQQLVAYYSKHADGLCHRLTT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMGTWNGTTRVAIKTLKPGTMS PEAFLQEAQVMKKLRHEKLVQLYAVVSEEP IYIVTEYMSKGSLLDFLKGETGKYLRPLQLV

		DMAAQIASGMAYVERMNYVHRDLRAANIL VGENLVCKVADFGLARLIEDNEYTARQGAK FPIKWTAPEAALYGRFTIKSDVWSFGILLTEL TTKGRVPYPGMVNREVLDQVERGYRMPCPP ECPESLHDLMCQCWRKEPEERATFEYLQAF LEDYFTSTEPQYQPGENL
Rs134 65248 02	P510S	>sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3 MGSNKSKPKDASQRRRSLEPAENVHGAGGG AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDTVTSPQRAGPLAGGVTTTFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTYIPSNYVAPSDSIQAEEWYF GKITRRESERLLLNAENPRGTFLVRESETTKG AYCLSVSDFDNAKGLNVKHYKIRKLDSGGF YITSRTQFNSLQQLVAYYSKHADGLCHRLTT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMTWNGTTRVAIKTLKPGTMS PEAFLQEAQVMKKLRHEKLVQLYAVVSEEP IYIVTEYMSKGSLLDFLKGETGKYLRPLQLV DMAAQIASGMAYVERMNYVHRDLRAANIL VGENLVCKVADFGLARLIEDNEYTARQGAK FPIKWTAPEAALYGRFTIKSDVWSFGILLTEL TTKGRVPYPGMVNREVLDQVERGYRMPCPP ECPESLHDLMCQCWRKEPEERSTFEYLQAF EDYFTSTEPQYQPGENL
Rs136 65425 93	G173V	>sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3 MGSNKSKPKDASQRRRSLEPAENVHGAGGG AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDTVTSPQRAGPLAGGVTTTFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTVYIPSNYVAPSDSIQAEEWYF

		<p>GKITRRESERLLLNAENPRVTFVRESETTKG AYCLSVSDFDNAKGLNVKHYKIRKLDSGGF YITSRTQFNSLQQLVAYYSKHADGLCHRLTT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMGTWNGTTRVAIKTLKPGTMS PEAFLQEAQVMKKLRHEKLVQLYAVVSEEP IYIVTEYMSKGSLLDFLKGETGKYLRPLQLV DMAAQIASGMAYVERMNYVHRDLRAANIL VGENLVCKVADFGLARLIEDNEYTARQGAK FPIKWTAPEAALYGRFTIKSDVWSFGILLTEL TTKGRVPYPGMVNREVLDQVERGYRMPCPP ECPESLHDLMCQCWRKEPEERPTFEYLQAFL EDYFTSTEPQYQPGENL</p>
Rs761 89035 8	F223L	<p>>sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3 MGSNKSQPKDASQRRRSLEPAENVHGAGGG AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDTVTSPQRAGPLAGGVTTFFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTYIPSNYVAPSDSIQAEWYF GKITRRESERLLLNAENPRGTFLVRESETTKG AYCLSVSDFDNAKGLNVKHYKIRKLDSGGF YITSRTLNSLQQLVAYYSKHADGLCHRLTT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMGTWNGTTRVAIKTLKPGTMS PEAFLQEAQVMKKLRHEKLVQLYAVVSEEP IYIVTEYMSKGSLLDFLKGETGKYLRPLQLV DMAAQIASGMAYVERMNYVHRDLRAANIL VGENLVCKVADFGLARLIEDNEYTARQGAK FPIKWTAPEAALYGRFTIKSDVWSFGILLTEL TTKGRVPYPGMVNREVLDQVERGYRMPCPP ECPESLHDLMCQCWRKEPEERPTFEYLQAFL EDYFTSTEPQYQPGENL</p>

Rs750 53458 4	A296S	>sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3 MGSNKSKPKDASQRRRSLEPAENVHGAGGG AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDTVTSPQRAGPLAGGVTTFFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTGYIPSNYVAPSDSIQAEWYF GKITRRESERLLLNAENPRGTFLVRESETTKG AYCLSVSDFDNAKGLNVKHYKIRKLDSSGGF YITSRTQFNSLQQLVAYYSKHADGLCHRLLT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMGTWNGTTRVSIKTLKPGTMSP EAFLQEAQVMKKLRHEKLVQLYAVVSEEPI YIVTEYMSKGSLLDFLKGETGKYLRPLQLVD MAAQIASGMAYVERMNYVHRDLRAANILV GENLVCKVADDFGLARLIEDNEYTARQGAKF PIKWTAPEAALYGRFTIKSDVWSFGILLTELT TKGRVPYPGMVNREVLDQVERGYRMPCPE CPESLHDLMCQCWRKEPEERPTFEYLQAFLE DYFTSTEPQYQGENL
Rs127 43222 27	A330D	>sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3 MGSNKSKPKDASQRRRSLEPAENVHGAGGG AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDTVTSPQRAGPLAGGVTTFFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTGYIPSNYVAPSDSIQAEWYF GKITRRESERLLLNAENPRGTFLVRESETTKG AYCLSVSDFDNAKGLNVKHYKIRKLDSSGGF YITSRTQFNSLQQLVAYYSKHADGLCHRLLT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMGTWNGTTRVAIKTLKPGTMS PEAFLEAQAQVMKKLRHEKLVQLYDVVSEEP IYIVTEYMSKGSLLDFLKGETGKYLRPLQLV DMAAQIASGMAYVERMNYVHRDLRAANIL VGENLVCKVADDFGLARLIEDNEYTARQGAK FPIKWTAPEAALYGRFTIKSDVWSFGILLTEL TTKGRVPYPGMVNREVLDQVERGYRMPCPP ECPELHDLMCQCWRKEPEERPTFEYLQAFLE

		EDYFTSTEPQYQPGENL
Rs754 96843 5	M344L	>sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3 MGSNKSKPKDASQRRRSLEPAENVHGAGGG AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDTVTSPQRAGPLAGGVTTTFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTGYIPSNYVAPSDSIQAEWYF GKITRRESERLLLNAENPRGTFLVRESETTKG AYCLSVSDFDNAKGLNVKHYKIRKLDSSGGF YITSRTQFNLSLQQLVAYYSKHADGLCHRLTT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMGTWNGTTRVAIKTLKPGTMS PEAFLQEAQVMKKLRHEKLVQLYAVVSEEP IYIVTEYLSKGSLLDFLKGETGKYLRPLQLV DMAAQIASGMAYVERMNYVHRDLRAANIL VGENLVCKVADFGLARLIEDNEYTARQGAK FPIKWTAPEAALYGRFTIKSDVWSFGILLTEL TTKGRVPYPGMVNREVLDQVERGYRMPCPP ECPESLHDLMCQCWRKEPEERPTFEYLQAFL EDYFTSTEPQYQPGENL
Rs149 09692 87	T432M	>sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3 MGSNKSKPKDASQRRRSLEPAENVHGAGGG AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDTVTSPQRAGPLAGGVTTTFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTGYIPSNYVAPSDSIQAEWYF GKITRRESERLLLNAENPRGTFLVRESETTKG AYCLSVSDFDNAKGLNVKHYKIRKLDSSGGF YITSRTQFNLSLQQLVAYYSKHADGLCHRLTT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMGTWNGTTRVAIKTLKPGTMS

		<p>PEAFLQEAQVMKKLRHEKLVQLYAVVSEEP IYIVTEYMSKGSLLDFLKGETGKYLRPQLV DMAAQIASGMAYVERMNYVHRDLRAANIL VGENLVCKVADFGLARLIEDNEYTARQGAK FPIKWMAPEAALYGRFTIKSDVWSFGILLTE LTTKGRVPYPGMVNREVLQVERGYRMPC PPECPESLHDLMCQCWRKEPEERPTFEYLQA FLEDYFTSTEPQYQPGENL</p>
Rs134 44236 76	M498L	<p>>sp P12931 SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens OX=9606 GN=SRC PE=1 SV=3 MGSNKSKPKDASQRRRSLEPAENVHGAGGG AFPASQTPSKPASADGHRGPSAAFAPAAAEP KLFGGFNSSDTVTSPQRAGPLAGGVTTFVAL YDYESRTETDLSFKKGERLQIVNNTTEGDWW LAHSLSTGQTGYIPSNYVAPSDSIQAEWYF GKITRRESERLLLNAENPRGTFLVRESETTKG AYCLSVSDFDNAKGLNVKHYKIRKLDSSGGF YITSRTQFNSLQQLVAYYSKHADGLCHRLTT VCPTSKPQTQGLAKDAWEIPRESLRLEVKLG QGCFGEVWMGTWNGTTRVAIKTLKPGTMS PEAFLQEAQVMKKLRHEKLVQLYAVVSEEP IYIVTEYMSKGSLLDFLKGETGKYLRPQLV DMAAQIASGMAYVERMNYVHRDLRAANIL VGENLVCKVADFGLARLIEDNEYTARQGAK FPIKWTAPEAALYGRFTIKSDVWSFGILLTEL TTKGRVPYPGMVNREVLQVERGYRMPCPP ECPESLHDLLCQCWRKEPEERPTFEYLQAFL EDYFTSTEPQYQPGENL</p>