

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة 8 ماي 1945 قالمة

Université 8 Mai 1945 Guelma

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



## Mémoire En Vue de l'Obtention du Diplôme de Master

**Domaine :** Sciences de la Nature et de la Vie

**Filière :** Sciences Biologiques

**Spécialité/Option :** Biochimie Appliquée

**Département :** Biologie

### Thème

---

## Étude bio-informatique de l'expression des gènes chez les patients atteints du cancer de la prostate

---

**Présenté par :**

-Mme. NAILI CHERIFI Narimane \_ -Melle. LACHAAL Rania

**Devant le jury composé de :**

**Président : Dr. BOUDEN I.**

**M.C.B**

**Université de Guelma**

**Examineur : Dr. BAALI S.**

**M.A.A**

**Université de Guelma**

**Encadreur : Dr. OUMEDDOUR A.**

**M.C.A**

**Université de Guelma**

**Juin 2023**

## **REMERCIEMENTS**

*En préambule à ce mémoire nous remerciant ALLAH qui nous a donné la patience et le courage durant ces longues années d'étude.*

*Nous remercions Monsieur «Dr. BOUDEN I», Maître de conférences à l'Université De Guelma le 8 mai 1945, d'avoir accepté de présider le jury.*

*Nous remercions également Monsieur «Dr. BAALI S», maître assistant à L'Université de Guelma, le 8 mai 1945, d'avoir accepté d'examiner cet ouvrage.*

*Ce travail ne serait pas aussi riche et n'aurait pas pu avoir le jour sans l'aide et L'encadrement de notre promoteur «Dr. OUMEDDOUR A», nous la remercions pour Toute sa patience et sa disponibilité dont elle a fait preuve à notre égard. Ses conseils Et remarques constructives nous ont permis d'améliorer grandement la qualité de notre Travail et de notre mémoire.*

*Enfin, on remercie toutes les personnes et tout le département de biologie qui a participé De près ou de loin à l'exécution de ce travail.*

## **DEDICACE**

*J'ai l'honneur de dédie ce travail tous d'abord de mes très chers parents « Kamel, Wassila » qui m'ont soutenus et encourages durant toute la période de mes études et à qui je souhaite une longue et heureuse vie, que dieu les protège et me donne la force pour que je puisse leurs rendre un petit peu de leurs bien faits.*

*A mes très chers frères et sœurs CHERIFI ; Yasser ; Nedjmeddine et Ikram et Nourhane surtout mon mari « Abdallah »*

*A toute la famille 'LAID et 'CHERIFI surtout mes grands-parents.*

*A mon très cher binôme et sœur Rania et sa famille.*

*A mes chers amies Sofia, Imen, Wissal, Fatma Zohra, Meriem et ses familles.*

*En fin, a tous ceux qui m'aiment.*

**Narimane**

## **DEDICACE**

*A mes chers parents « Madjid et Dalila » pour tous leurs sacrifices, leurs tendresse, leurs amours et leurs prières depuis ma naissance et non seulement lors de la réalisation de ce travail.*

*A mon cher frère Bassem et ses princesses Ikhlas et Israa*

*A ma chère sœur Houda et son ange Racim*

*A toute ma famille et surtout ma belle tante Rachida ....merci d'être toujours là pour moi.*

*A ma chère binôme et sœur Narimane pour les agréables moments passés durant la réalisation de ce travail.*

*A mes chers amis Sofia, Imene, Mouna, Souad, Noussa pour leur bonne amitié.*

*A tous ceux qui, par un mot m'ont donné la force de continuer ....*

**Rania**

## **Résumé**

Le but de notre étude était d'identifier les gènes clés communs associés au cancer de la prostate et à son diagnostic. Les données des trois profils ont été téléchargées à partir de la base de données *GEO* et analysées à l'aide de plusieurs outils bio-informatiques. Les gènes à expression différentiels communs (DEGs) ont été déterminés à l'aide de la base de données *Venn diagram*. Des analyses *Gene Ontology* (GO) et *Kyoto Encyclopedia of Genes and Genomes* (KEGG) ont été effectuées via le site *DAVID*, et des réseaux d'interaction protéine-protéine (PPI) ont été construits via le site *STRING* et visualisés avec le logiciel *Cytoscape*. Enfin, nous avons vérifié la survie globale et l'expression des gènes clés à l'aide de la base de données *GEPIA2*.

On a obtenu 178 DEGs, dont 65 gènes surexprimés et 113 sous-exprimés. Les analyses GO et KEGG ont montré que ces DEGs ont été enrichi dans plusieurs processus biologiques (*Biological process* : BP), cellulaires (*Cellular component* : CC) et moléculaires (*Molecular function* : MF) associés au cancer de la prostate. De plus, 8 gènes clés ont été sélectionnés, dont 2 étaient associés à une diminution de la survie globale des patients et étaient significativement exprimés dans les tissus cancéreux par rapport aux tissus normaux.

En résumé, les deux gènes (*NRAS* : *NRAS proto-oncogene*, *SERPINB5* : *serpin family B member 5*) peuvent être excellents biomarqueurs pour le diagnostic du cancer prostatique et la thérapie génique ciblée.

**Mots clés** : cancer de la prostate, bio-informatique, les gènes à expression différentiels communs, développement du cancer, adénocarcinome.

## **Abstract**

The purpose of our study was to identify key common genes associated with prostate cancer and its diagnosis. The data from the three profiles were downloaded from the *GEO* database and analysed using several bioinformatics tools. Common differential expression genes (GDEs) were determined using the *Venn diagram* database. *Gene Ontology* (GO) and *Kyoto Encyclopedia of Genes and Genomes* (KEGG) analyses were performed via the *DAVID* site, and protein-protein interaction networks (PPI) were built via the *STRING* site and visualized with *Cytoscape* software. Finally, we checked the *overall survival* and expression of key genes using the *GEPIA2* database.

178 GDEs were obtained, of which 65 genes were over-expressed and 113 were under-expressed. GO and KEGG analyses showed that these GDEs were enriched in several biological processes (*Biological process: BP*), cellular (*Cellular component: CC*) and molecular (*Molecular function: MF*) associated with prostate cancer. In addition, 8 key genes were selected, 2 of which were associated with decreased overall patient survival and were significantly expressed in cancer tissue relative to normal tissue.

In summary, both genes (NRAS: *NRAS proto-oncogene*, SERPINB5: *serpin family B member 5*) can be excellent biomarkers for prostate cancer diagnosis and targeted gene therapy.

**Keywords** : prostate cancer, bioinformatics, common differential expression genes, cancer development, adenocarcinoma.

## ملخص

كان الغرض من دراستنا هو تحديد الجينات الرئيسية الشائعة المرتبطة بسرطان البروستاتا وتشخيصه. وتم تنزيل البيانات المستمدة من الموجزات الثلاثة من قاعدة بيانات GEO وتحليلها باستخدام عدة أدوات للمعلوماتية الحيوية. تم تحديد جينات ذات التعبير التفاضلي المشترك (GDEs) باستخدام قاعدة بيانات *Venn diagram*. تم إجراء تحليلات *KEGG* *Kyoto Encyclopedia of Genes and Genomes* (*GO*) *Gene Ontologie* عبر موقع *DAVID*، وتم بناء شبكات تفاعل البروتين والبروتين (*PPI*) عبر موقع *STRING* وتم تصورها باستخدام برنامج *Cytoscape*. أخيرًا، قمنا بفحص البقاء العام والتعبير عن الجينات الرئيسية باستخدام قاعدة بيانات *GEPIA2*.

تم الحصول على 178 DEGs، منها 65 جينًا تم التعبير عنها بشكل زائد و113 جينًا تم التعبير عنها بشكل ناقص. أظهرت تحليلات *GO* و *KEGG* أن هذه DEGs ثرية في العديد من العمليات البيولوجية (*Biological process*): *BP*، الخلية (*Cellular component*: *CC*) والجزئية (*Molecular function*: *MF*) المرتبطة بسرطان البروستاتا. بالإضافة إلى ذلك، تم اختيار 8 جينات رئيسية، 2 منها مرتبطة بانخفاض بقاء المريض بشكل عام وتم التعبير عنها بشكل كبير في أنسجة السرطان بالنسبة للأنسجة الطبيعية.

باختصار، يمكن أن يكون كلا الجينين (*NRAS: NRAS proto-oncogene*، *SERPINB5: serpin family B member 5*) مؤشرات حيوية ممتازة لتشخيص سرطان البروستاتا والعلاج الجيني المستهدف.

**الكلمات الرئيسية:** سرطان البروستاتا، المعلوماتية الحيوية، جينات التعبير التفاضلي المشترك، تطور السرطان، السرطان الغدي.

## Liste des figures

<b>Figure 1</b> : Anatomie des systèmes reproducteur et urinaire masculins. ....	2
<b>Figure 2</b> : anatomie zonale de la prostate. ....	3
<b>Figure 3</b> : technique de la toucher rectale. ....	4
<b>Figure 4</b> : Technique de la biopsie prostatique. ....	5
<b>Figure 5</b> : développement de cancer de la prostate. ....	6
<b>Figure 6</b> : les principaux stades de cancer de la prostate. ....	6
<b>Figure 7</b> : Les tracés volcaniques ( <i>Volcano plots</i> ) des trois ensembles de données GEO, le bleu représente les gènes sous-exprimés, le rouge représente les gènes surexprimés ( $P < 0 :05$ et $\log  FC  > 1$ ). ....	11
<b>Figure 8</b> : DEGs communs aux trois ensembles de données GEO. (A) 113 gènes sous-exprimés. (B) 65 gènes surexprimés. ....	12
<b>Figure 9</b> : Le réseau PPI créé par le site web <i>STRING</i> , les gènes surexprimés en rose, les gènes sous-exprimés en bleu. ....	16
<b>Figure 10</b> : Les huit principaux gènes clés classés par les algorithmes topologiques <i>Maximal Clique Centrality</i> (MCC) et <i>Degree</i> . ....	17
<b>Figure 11</b> : L'analyse de la survie globale des huit gènes clés par la base de données GEPIA2....	18
<b>Figure 12</b> : L'analyse de l'expression des gènes clés par la base de données GEPIA2. Les tissus cancéreux en rouge, les tissus normaux en bleu. (PRAD : <i>Prostate adenocarcinoma</i> ). ....	19

## Liste des tableaux

<b>Tableau 1</b> : Les facteurs de risques de cancer de la prostate. ....	7
<b>Tableau 2</b> : Les traitements selon les stades de cancer de la prostate. ....	8
<b>Tableau 03</b> : Liste des DEGs communs dans le cancer de la prostate. ....	12
<b>Tableau 04</b> : Analyse du GO des DEGs communs dans le cancer de la prostate. ....	14
<b>Tableau 05</b> : Analyse KEGG des DEGs courants dans le cancer de la prostate. ....	15

## Liste des abréviations

**PaC** : Cancer de la prostate.

**ADN**: Acide Désoxyribonucléique.

**PIN** : Le néoplasie intraépithélial de la prostate

**BP**: *Biological Process.*

**CC**: *Cellular Component.*

**DAVID**: *Database for Annotation Visualization and Integrated Discovery.*

**DEGs**: *Differentially Expressed Genes.*

**FC**: *Fold change.*

**GEO**: *Gene Expression Omnibus.*

**GEPIA2**: *Gene Expression Profiling Interaction Analysis.*

**GO**: *Gene Ontology.*

**GPL**: *Gene Platform.*

**GSE**: *Gene Sample Expression.*

**KEGG**: *Kyoto Encyclopedia of Genes and Genomes.*

**MCC**: *Maximal Clinique Centrality.*

**MCODE**: *Molecular Complex Detection.*

**MF**: *Molecular Function.*

**OS**: *Overall Survival.*

**HR** : *Hazard ratio.*

**P**: Valeur de probabilité.

**PPI**: *Protein-Protein Interaction.*

**HBP** : Hyperplasie bénigne de la prostate

**PRAD** : *Prostate adenocarcinoma*

**TLR4** : *Toll-Like Receptor 4*

**CAV1** : Cavéoline-1

**SERPINB5** : *Serpin family B member 5.*

**SFN** : *Stratifin.*

**SP1** : *Sp1 transcription factor.*

**GATA3** : *GATA binding protein 3.*

**NRAS** : *NRAS proto-oncogene.*

**KRT5** : *Keratin 5.*

**NF-κB**: *Nuclear Factor Kappa lightchain-enhancer of activated B cells.*

**Mp-MRI** : *Multiparametric-magnetic resonance imaging.*

**TRUS** : *Transrectal ultrasound.*

**NFAT** : *Nuclear factor of activated T-cells.*

**CDC20** : *Cell division cycle 20 homolog.*

**D-box** : *Destruction Box.*

**TNM**: *Tumeur, Ganglion Lymphatique, Métastase.*

**GSTM1** : *Gluthathione S-Transferase Mu 1.*

**GSTT1** : *Gluthathione S-Transferase Theta 1.*

**GEF** : *Guanine nucleotide exchange factor.*

**NF1** : *Neurofibromatosis type 1.*

**RTK** : *Tyrosine Kinase receptor.*

**TFIID** : *Transcription factor II D.*

**TBP** : *TATA binding protein.*

**OCT-1** : *Octamer transcription Factor 1.*

**VHL** : *Von Hippel-Lindau.*

**VEGF** : *Vascular Endothelial Growth Factor.*

**HNSCC** : *Head and Neck squamous cell carcinoma.*

**HDAC** : *Histone Deacetylase.*

**RTU-P** : Résection trasuréctale de la prostate.

**TR** : Toucher Rectal.

**APS** : *Prostate cancer antigen.*

**ETR** : *Ecographie transrectale.*

# Sommaire

## REMERCIEMENTS

## DEDICACE

### Résumé

### Abstract

### ملخص

### Liste des figures

### Liste des tableaux

### Liste des abréviations

## INTRODUCTION

### **Introduction.....1**

#### **Partie théorique    Chapitre I : cancer de la prostate**

1. La prostate.....	2
1.1. Anatomie de la prostate.....	2
1.2. Anatomie zonale de la prostate.....	2
1.3. Cancer de la prostate.....	3
1.4. Diagnostic du cancer de la prostate.....	3
1.4.1.    Symptômes.....	3
1.4.2.    Dépistage.....	4
1.4.3.    Diagnostiques complets de cancer.....	4
1.4.4.    Biopsie prostatiques.....	5
1.5. Origine et développement de cancer de la prostate.....	5
1.6. Les stades de cancer de la prostate.....	6
1.7. Les facteurs de risque.....	7
1.8. Les traitements.....	7

#### **Partie pratique    Chapitre II : Méthodes**

2. Méthodes.....	9
2.1. L'objectif de travail.....	9
2.2. Méthodes d'analyses.....	9
2.2.1.    Source de données et informations sur les puces à ADN.....	9
2.2.2.    Traitement des données et identification des DEGs.....	9
2.2.3.    Analyse de l'ontologie des gènes et de l'enrichissement des voies.....	10
2.2.4.    Analyse du réseau et des modules PPI.....	10

2.2.5.	Sélection des gènes clés .....	10
2.2.6.	Analyse de survie globale et validation de l'expression des gènes clés .....	10

**Partie pratique      Chapitre III : Résultats et Discussion**

3.	Résultats et Discussion.....	11
3.1.	Résultats .....	11
3.1.1.	Traitement des données et identification des DEGs .....	11
3.1.2.	Analyse de l'ontologie des gènes et de l'enrichissement des voies .....	13
3.1.3.	Analyse de Kyoto Encyclopedia of Genes and Genomes (KEGG) .....	15
3.1.4.	Analyse de l'interaction protéine-protéine .....	16
3.1.5.	Sélection des gènes clés .....	17
3.1.6.	Analyse de survie globale et validation de l'expression des gènes clés .....	18
3.2.	Discussion.....	20

**CONCLUSION**

<b>Conclusion.....</b>	<b>25</b>
------------------------	-----------

**Références Bibliographiques**

<b>Références bibliographiques.....</b>	<b>26</b>
<b>Glossaire.....</b>	<b>32</b>

# **INTRODUCTION**

## **Introduction**

Le cancer est une maladie causée par une croissance cellulaire incontrôlée dans les tissus corporels normaux. En raison de la nature de la tumeur primaire, il existe un risque de développer des métastases au fur et à mesure que la maladie progresse (1).

La plupart des cancers de la prostate proviennent de lésions précancéreuses de l'épithélium de la prostate. Le néoplasie intraépithélial de la prostate (PIN) est reconnu comme une lésion précurseur du cancer de la prostate (PaC). Ce sont des anomalies cytoplasmiques nucléaires présentes dans les cellules monocouches des glandes prostatiques et des tubules rénaux (2).

L'incidence du cancer de la prostate (PAC) varie selon la région et la population. En 2018, 1 276 106 nouveaux cas de cancer de la prostate ont été enregistrés dans le monde, ce qui représente 7,1 % de tous les cancers masculins. Les taux de normalisation par âge les plus élevés se trouvaient en Océanie (79,1 pour 100 000 habitants) et en Amérique du Nord (73,7), suivis de l'Europe (62,1). À l'inverse, les taux d'incidence en Afrique et en Asie sont inférieurs à ceux des pays développés (respectivement 26,6 et 11,5). La différence de prévalence était de 190 fois entre le groupe à prévalence la plus élevée (France, Guadeloupe, 189,1) et le groupe à prévalence la plus faible (Bhoutan, 1,0) (3,4).

Selon *Globocan 2020*, il y a eu 58 418 cas de différents types de cancer en Algérie en 2020. Parmi ceux-ci, le cancer de la prostate représentait 3 597 cas, 6,2 % de tous les cas, et 21 635 décès, soit 45,5 %. Le cancer de la prostate arrive au quatrième rang après le cancer du sein (21,5 %), le cancer du poumon (8,2 %) et enfin le cancer du côlon (6,5 %). A noter que l'âge moyen des patients atteints d'un cancer de la prostate en Algérie est de 71 ans (5).

Notre étude est basée sur l'utilisation de multiples outils et bases de données bio-informatiques pour extraire et analyser un ensemble de données afin d'identifier les gènes communs exprimés différemment dans les tissus atteints de cancer de la prostate. Ce mémoire se divise en trois parties :

- La partie théorique est spécifique au cancer de la prostate : Anatomie du cancer de la prostate et anatomie zonale, diagnostique, origine et développement, facteurs de risque, stades et traitement.
- La première partie de la pratique (Méthodes) est consacrée à l'analyse bio-informatique des données des patients atteints d'un cancer de la prostate.
- Et se termine avec la deuxième partie de la pratique (Résultats et Discussion)

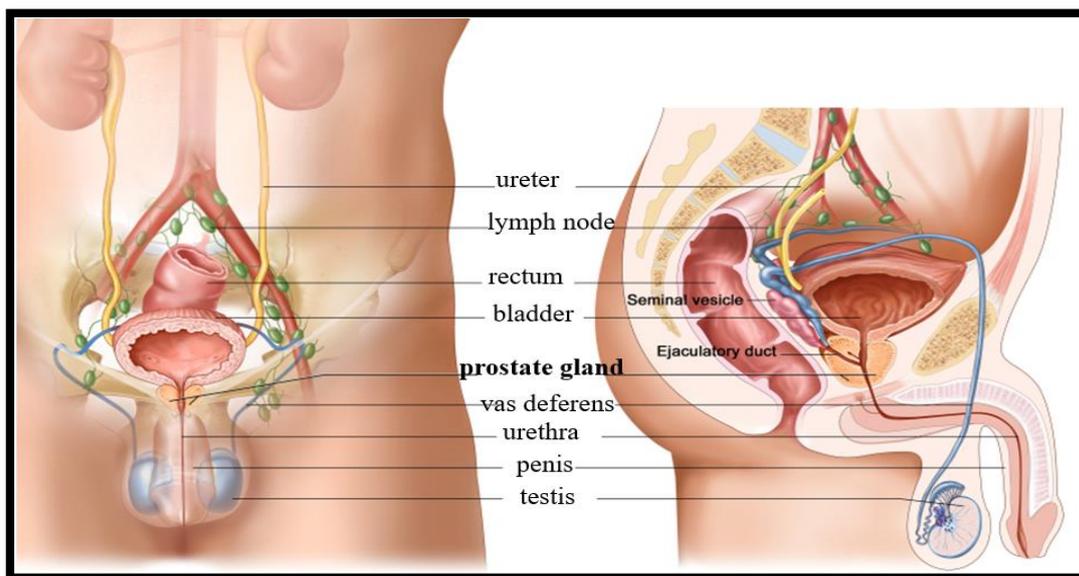
## **Partie théorique**

### **Chapitre I : cancer de la prostate**

## 1. La prostate

### 1.1. Anatomie de la prostate

La prostate est une glande exocrine située sous la vessie et devant le rectum. Il entoure le début de l'urètre (Figure 1), le tube par lequel l'urine et le sperme s'écoulent. Une série de fibres musculaires qui forment le sphincter urétral convergent sous la prostate au niveau du col de la vessie et se contractent ou se détendent pour contrôler le passage de l'urine, permettant le contrôle de la continence. La prostate est entourée d'une capsule qui la sépare des autres organes pelviens (6).



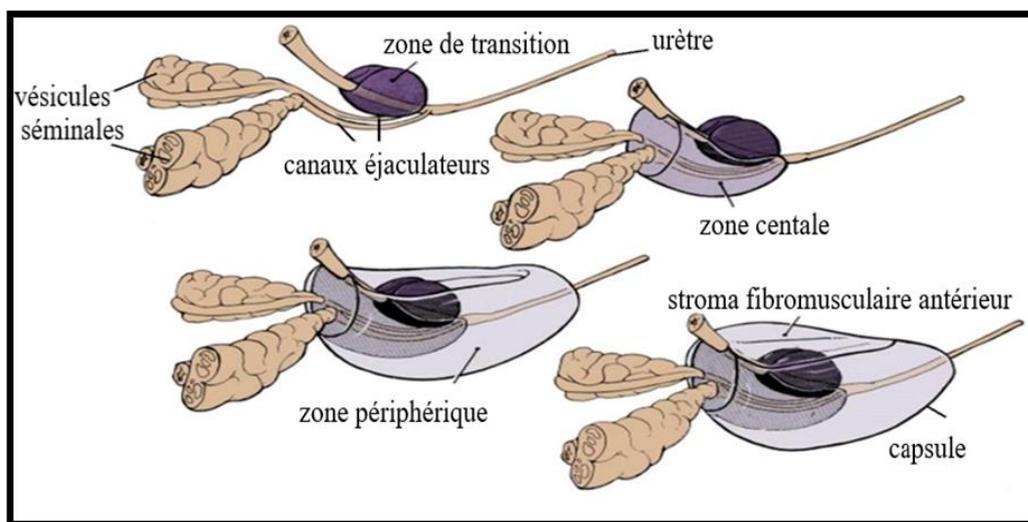
**Figure 1:** Anatomie des systèmes reproducteur et urinaire masculins (7).

### 1.2. Anatomie zonale de la prostate

Selon le modèle anatomique proposé par Mac Neal en 1988, la prostate est composée de quatre régions distinctes (8).

- a) **La zone de transition :** située au milieu de la prostate. Il entoure l'urètre et représente environ 5% de la prostate à 40 ans. Avec l'âge, cette zone se développe dans la majeure partie de la prostate. Ceci est appelé adénome prostatique (ou hyperplasie bénigne de la prostate) et est très fréquent chez les hommes de plus de 70 ans.
- b) **Zone centrale :** située à la racine et entourant le canal éjaculateur. Il représente 20% de la prostate.
- c) **Zone périphérique :** proche du rectum. Cette proximité facilite la palpation lors du toucher rectal. Il occupe la plus grande surface de la prostate à environ 75%. La plupart des tumeurs surviennent dans cette zone environnante.

- d) **Zone antérieure** : composée de stroma fibromusculaire et dépourvue de glandes. Il est en continuité avec le sphincter lisse de l'urètre. Non soumis à des processus pathologiques (Figure 2).



**Figure 2** : Anatomie zonale de la prostate (9).

### 1.3. Cancer de la prostate

Le cancer de la prostate (PAC) est une maladie des cellules de la prostate. Des organes génitaux masculins. Ce cancer commence dans les cellules se transforme en une forme normale, chaotique et proliférative de tumeurs Intelligent. Au fil du temps, les tumeurs peuvent se propager localement. Cellule cancéreuse, Il peut se diviser et se déplacer vers d'autres tissus et organes, où il peut former de nouvelles tumeurs appelées métastases. La maladie progresse souvent lentement au fil des ans (10).

### 1.4. Diagnostic du cancer de la prostate

#### 1.4.1. Symptômes

Le problème des maladies des voies urinaires chez les hommes âgés est le plus souvent dû à une hyperplasie bénigne de la prostate, qui est fréquente chez ces patients et peut coexister avec le cancer. Au fur et à mesure que la tumeur se développe, les symptômes les plus fréquemment observés sont l'hématurie (la présence de sang dans l'urine) ou l'obstruction des voies urinaires. Les métastases tumorales à distance peuvent provoquer un œdème des jambes en raison d'une obstruction locale des ganglions lymphatiques ou de douleurs osseuses lors de la métastase. Ces symptômes apparaissent très tardivement dans l'histoire du patient (11).

### 1.4.2. Dépistage

#### a) Dosage de PSA

Il consiste en une simple prise de sang. Le PSA est principalement sécrété uniquement par les cellules prostatiques. La majeure partie de ce PSA est excrété dans le sperme. Cependant, seule une petite partie de celui-ci pénètre dans le sang. La plupart des hommes en bonne santé ont un taux de PSA sanguin inférieur à 1.4 ng/ml. Les cellules tumorales sécrètent plus de PSA que les cellules saines. Le développement du cancer est généralement associé à des niveaux élevés de PSA dans le sang, mais ce n'est pas toujours le cas (10 à 15 patients atteints de cancer ont des niveaux de PSA normaux) (12).

#### b) Toucher rectale

Examen au cours duquel le médecin insère un doigt ganté dans le rectum du patient et palpe la prostate (Figure 3). Cela permet le volume, la consistance et la texture de surface de la prostate. Bien qu'il s'agisse d'un test moins sensible (les tumeurs sont suffisamment développées pour être palpables), il peut détecter certains types de cancer lorsque le PSA2 normal est présent (12).

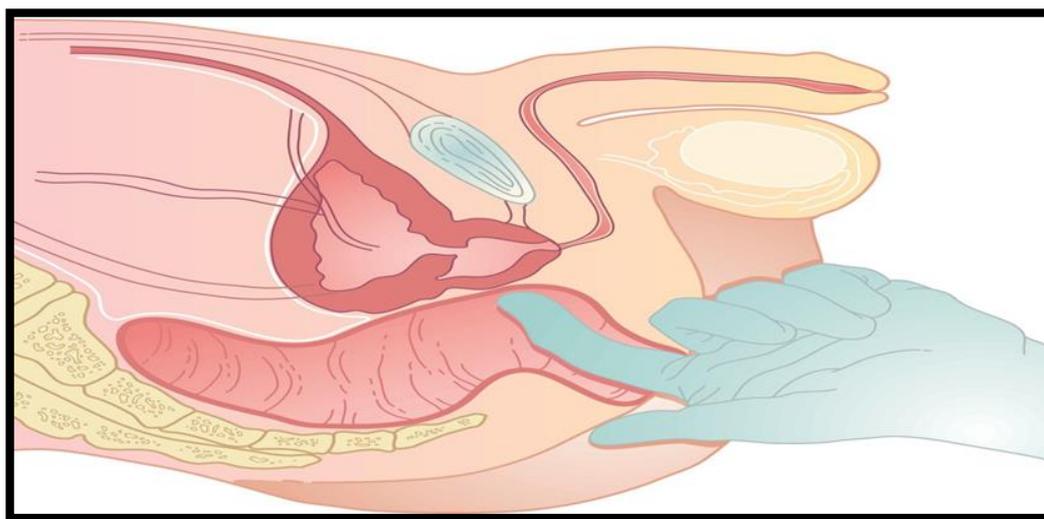


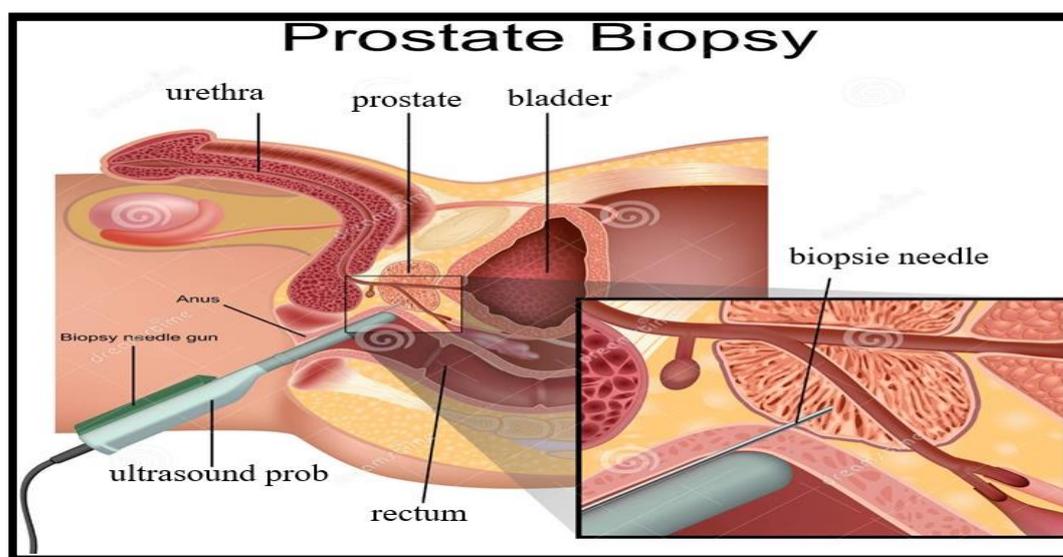
Figure 3 : Technique de la toucher rectale (13).

### 1.4.3. Diagnostiques complets de cancer

Si un examen rectal numérique pour les anomalies du PSA ou du PSA suggère un cancer, des tests supplémentaires sont nécessaires pour confirmer le diagnostic. Une façon de confirmer la présence d'un cancer est une biopsie prostatique. Une IRM est généralement effectuée avant ce test pour identifier les sites possibles. Une biopsie ciblée est effectuée pour augmenter les chances de trouver un cancer (11).

#### 1.4.4. Biopsie prostatiques

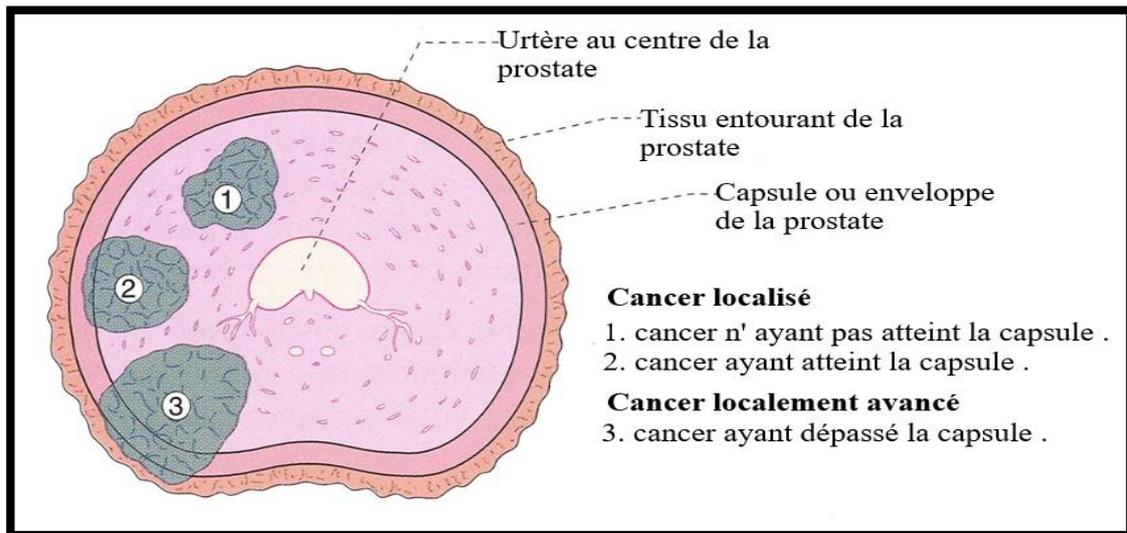
Elle consiste à ingérer au moins 12 « carottes » de tissu prostatique, qui sont ensuite examinées au microscope pour détecter la présence éventuelle de cellules cancéreuses. Cela se fait dans le rectum où le médecin insère une sonde à ultrasons dans le rectum. Cela permet au médecin de voir la prostate. Une biopsie est réalisée à l'aide d'une aiguille sous guidage échographique. Il réalise plusieurs biopsies "à l'aveugle" (Figure 4). C'est-à-dire que les noyaux sont sélectionnés au hasard dans différentes régions de la prostate, mais grâce à l'analyse IRM, des régions prédéfinies sont de plus en plus spécifiquement examinées (14).



**Figure 4 :** Technique de la biopsie prostatique (15).

#### 1.5. Origine et développement de cancer de la prostate

La prostate est un organe qui subit plusieurs stades de croissance tout au long de la vie. Très peu Début à la naissance et début brutal à la puberté sous l'effet de la testostérone. Et A 20 ans, le poids est d'environ 20 g et reste relativement stable jusqu'à 20 ans. Après 40 ans, la deuxième période de croissance commence, cette croissance est à l'origine d'une pathologie bénigne qui touche presque tous les hommes Année: Hyperplasie bénigne de la prostate (HBP). En effet, cette pathologie touche les personnes autour de 50 ans Hommes 51-60 et 90 ans 80+ Hommes 11 ans (Figure 5). Elle se caractérise par une déviation Voies urinaires causées par la compression de l'urètre par une prostate surdimensionnée (16,17).



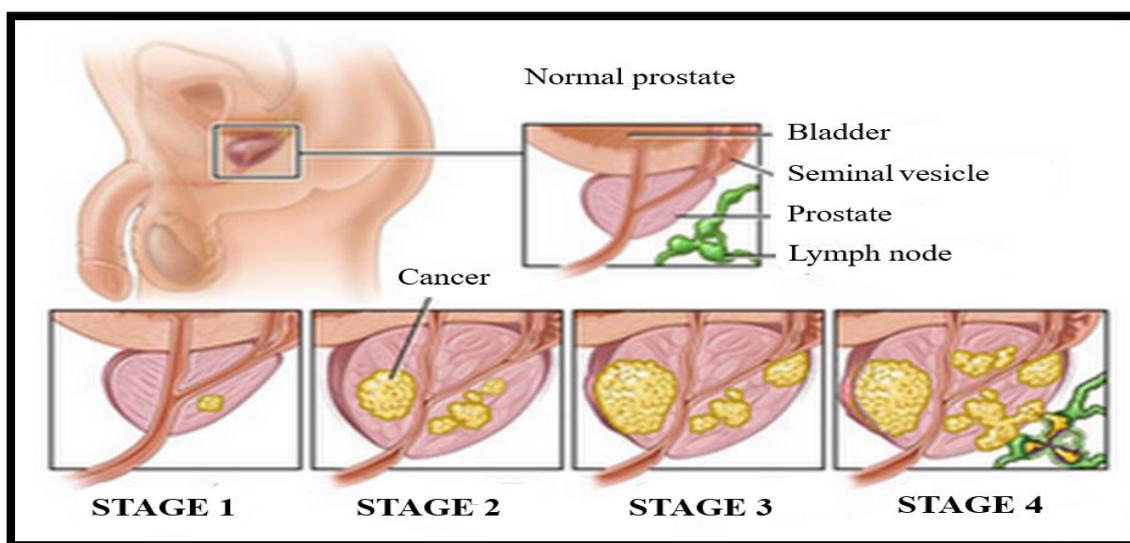
**Figure 5 :** Développement de cancer de la prostate (18).

### 1.6. Les stades de cancer de la prostate

Le cancer de la prostate a des stades principaux (Figure 6), et dépend de chacun La croissance et la propagation des cellules cancéreuses dans le corps. Connaître l'état d'avancement du cancer est très important pour déterminer le meilleur traitement contre le cancer de la prostate(19).

Le stade du cancer de la prostate dépend de trois facteurs :

- ✓ Taille de la tumeur.
- ✓ Localisation de la tumeur (à l'intérieur ou à l'extérieur de la prostate).
- ✓ La tumeur se propage à l'extérieur de la prostate et affecte cet organe (organes proches ou éloignés de la prostate) (20).



**Figure 6 :** Les principaux stades de cancer de la prostate (21).

## 1.7. Les facteurs de risque

Le cancer de la prostate n'a que quatre facteurs de risque bien établis. Il est spécifique au patient et ne peut donc pas être modifié (Tableau 1). Il s'agit de l'âge, de l'origine ethnique, Antécédents familiaux et prédisposition génétique au cancer de la prostate (22 ,23).

**Tableau 1** : Les facteurs de risques de cancer de la prostate.

Les facteurs de risque	Les descriptions
<b>L'âge</b>	Le cancer de la prostate est principalement un cancer lié à l'âge car il est si rare. Les patients diagnostiqués $\leq 50$ ans et $\geq 85$ ans ont été diagnostiqués chez des patients âgés de 55 à 85 ans. L'introduction du dépistage de l'APS a augmenté l'âge médian au diagnostic, ces derniers 66 ans restants (24).
<b>L'origine ethnique</b>	Les hommes asiatiques sont moins susceptibles de contracter la maladie, tandis que les hommes africains et afro-américains sont plus à risque (1,6 fois plus que les hommes blancs). Les différences ne sont pas encore entièrement comprises et sont probablement dues à de multiples facteurs. Ils reflètent des différences dans la génétique, l'exposition aux facteurs de risque et la détection (25).
<b>Antécédents familiaux</b>	Les antécédents familiaux jouent également un rôle important. En effet, avoir un parent au premier degré atteint d'un cancer de la prostate double votre risque de développer un cancer de la prostate. 20 cas de cancer de la prostate soient associés à des antécédents familiaux, dans la plupart des cas, cet héritage est polygénique, avec plus de 100 mutations génétiques associées à un risque accru de cancer de la prostate (26).
<b>prédisposition génétique</b>	Plus de 100 locus génétiques responsables du développement du cancer de la prostate ont été identifiés et confirmés dans des populations de diverses origines ethniques. Les 105 locus identifiés à ce jour sont responsables d'un tiers de l'héritabilité du cancer de la prostate. De nombreuses recherches se concentrent sur les facteurs environnementaux, en particulier les habitudes alimentaires, qui peuvent favoriser ou prévenir le développement du cancer de la prostate (26,27).

## 1.8. Les traitements

Le cancer de la prostate implique de multiples altérations moléculaires et génétiques qui surviennent et s'accumulent au cours de la maladie. Par conséquent, le traitement dépend du stade du cancer (Tableau2),

de la dépendance aux androgènes et de l'état d'avancement du cancer. Les traitements prodigués sont curatifs si le cancer est encore localisé et palliatifs si le cancer est avancé ou s'est propagé (28).

**Tableau 2 :** Les traitements selon les stades de cancer de la prostate (29).

Stades	Définition	Traitement	Survie
A1 ou T1a	Présence de cellules cancéreuses dans moins de 5 % des fragments provenant d'une RTU-P	Observation Traitement conservateur à moins que le cancer n'évolue	Plus de 10 ans
A2 ou T1b	Présence de cellules cancéreuses dans plus de 5 % des fragments provenant d'une RTU-P; aucun nodule détecté au TR	Prostatectomie ou radiothérapie	Plus de 10 ans
B0 ou T1c	Cancer détecté par ÉTR effectuée après constatation d'un taux d'APS élevé; pas de nodule détecté au TR	Prostatectomie ou radiothérapie	Plus de 10 ans
B1 ou T2a	Nodule cancéreux occupant plus de la moitié d'un lobe prostatique	Prostatectomie ou radiothérapie	Plus de 10 ans
B2 ou T2b	Nodule cancéreux occupant plus de la moitié d'un lobe prostatique	Radiothérapie ou, dans certains centres, prostatectomie	plus de 10 ans
B3 ou T2c	Tumeur touchant les lobes prostatiques	Prostatectomie ou radiothérapie	10 ans
C ou T3	Tumeur non confinée à la capsule prostatique envahissant les vésicules séminales du bassin	Prostatectomie ou radiothérapie	5 à 10 ans
D1 ou N+	Atteinte des ganglions lymphatiques pelviens	Hormonothérapie (précoce ou retardée)	5 ans
D2 ou M1	Métastases éloignées (osseuses)	Hormonothérapie Biphosphonates	2 ans
D3	Rechute consécutive à l'hormonothérapie	Soins palliatifs Chimiothérapie palliative Biphosphonates Protocoles de recherche	6 à 12 mois

**Partie pratique**  
**Chapitre II : Méthodes**

## 2. Méthodes

### 2.1. L'objectif de travail

L'objectif de cette étude est d'analyser l'expression différentielle des gènes associés au cancer de la prostate à l'aide des outils bio-informatiques intégrés.

### 2.2. Méthodes d'analyses

#### 2.2.1. Source de données et informations sur les puces à ADN

Les données des puces à ADN utilisées dans cette étude ont été téléchargées à partir du référentiel public de données de génomique fonctionnelle : la base de données *Gene Expression Omnibus* (GEO) (<https://www.ncbi.nlm.nih.gov/gds/>) Trois profils d'expression génique (GSE200879, GSE46602, GSE55945) étudiant le PC et des tissus normaux ont été choisis. Les données de matrice de GSE200879 consistaient en 115 tissus de PAC et 9 tissus normaux basés sur la plate-forme GPL32170 ([HTA-2\_0] *Affymetrix Human Transcriptome Array 2.0 [Genosplice]*). Les données de la matrice de GSE46602 referment 40 tissus de PAC et 10 tissus normaux basés sur la plate-forme de GPL570 ([HG-U133\_Plus\_2] *Affymetrix Human Genome U133 Plus 2.0 Array*). Les données de la matrice de GSE55945 comprennent 13 tissus PAC et 8 tissus normaux, plate-forme : GPL570 ([HG-U133\_Plus\_2] *Affymetrix Human Genome U133 Plus 2.0 Array*).

#### 2.2.2. Traitement des données et identification des DEGs

Les trois profils d'expression d'ARNm obtenus à partir de la base de données GEO ont été analysés par les outils en ligne GEO2R (<https://www.ncbi.nlm.nih.gov/geo/geo2r/>). Le  $|\log_2\text{-fold change (FC)}| > 1$  et une valeur  $P < 0,05$  ont été utilisés comme critères de sélection pour l'identification des gènes différentiellement exprimés (DEGs). Ensuite, les données brutes au format TXT ont été traitées par l'outil de *Venn Diagram* (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) pour obtenir les DEGs communs entre les trois profils. Les DEGs avec un  $\log FC > 0$  ont été considérés comme des gènes surexprimés (*up-regulated genes*), tandis que les DEGs avec  $\log FC < 0$  ont été considérés comme des gènes sous-exprimés (*down-regulated genes*).

### 2.2.3. Analyse de l'ontologie des gènes et de l'enrichissement des voies

L'analyse de l'ontologie des gènes (GO) est une approche couramment utilisée dans les études fonctionnelles du transcriptome ou des données génomiques à haut débit. L'analyse d'enrichissement des voies GO (BP, CC et MF) et (KEGG) est un outil de bio-informatique en ligne du logiciel *DAVID* (version 6.8, <https://david.ncifcrf.gov/>).  $P < 0,05$  a été fixé comme critère de sélection.

### 2.2.4. Analyse du réseau et des modules PPI

Pour explorer les interactions entre les DEGs, nous avons utilisé une base de données en ligne, *STRING* (version 11.5, <https://string-db.org/>). La confiance élevée (0,400) a été choisie comme score d'interaction minimum requis. Ensuite, après avoir supprimé les nœuds déconnectés, un réseau PPI a été visualisé à l'aide du logiciel *Cytoscape* 3.9.0. De plus, le plug-in *Molecular Complex Detection* (MCODE) du logiciel *Cytoscape* a été utilisé pour cribler les modules du réseau PPI.

### 2.2.5. Sélection des gènes clés

Les gènes ont été extraits à l'aide du plug-in *cytoHubba* du logiciel *Cytoscape*. Ensuite, nous avons sélectionné les huit principaux gènes clés classés par les algorithmes topologiques *Maximal Clique Centrality* (MCC) et *Degree*.

### 2.2.6. Analyse de survie globale et validation de l'expression des gènes clés

L'outil *Gene Expression Profiling Interactive Analysis* GEPIA2 (<http://gepia2.cancer-pku.cn/>), a été utilisé pour évaluer l'effet d'huit gènes clés sur la survie globale (Overall survival : OS) des patients atteints de PAC. Pour estimer le taux d'OS des patients atteints de PAC, les échantillons ont été divisés en groupes d'expression élevée et faible en fonction de l'expression médiane. Le rapport de risque (*Hazard ratio* : HR), les intervalles de confiance à 95 % correspondants et les valeurs P du *log-rank* ont été calculés et affichés sur le graphique. En plus, nous avons utilisé la même outil pour analyser les données d'expression de séquençage d'ARN des patients atteints de PAC en comparaison avec des échantillons témoins.

## **Partie pratique**

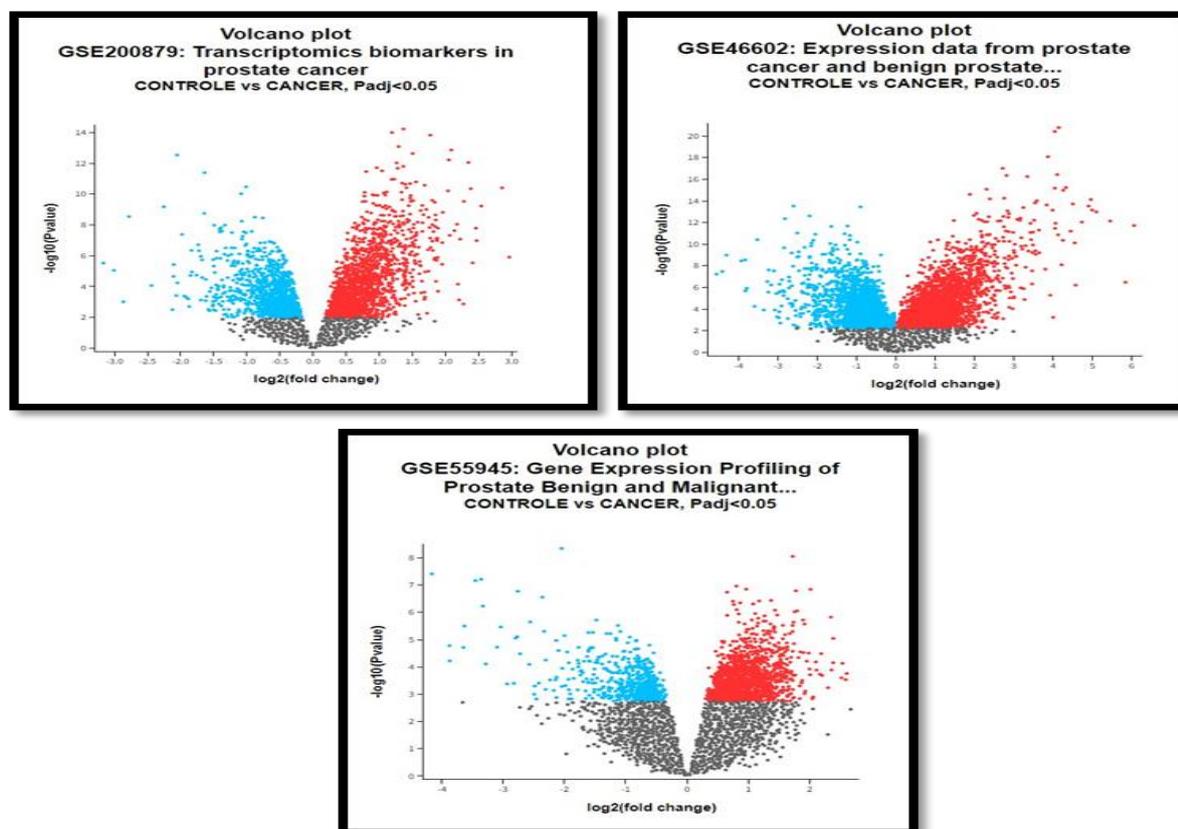
### **Chapitre III : Résultats et Discussion**

### 3. Résultats et Discussion

#### 3.1. Résultats

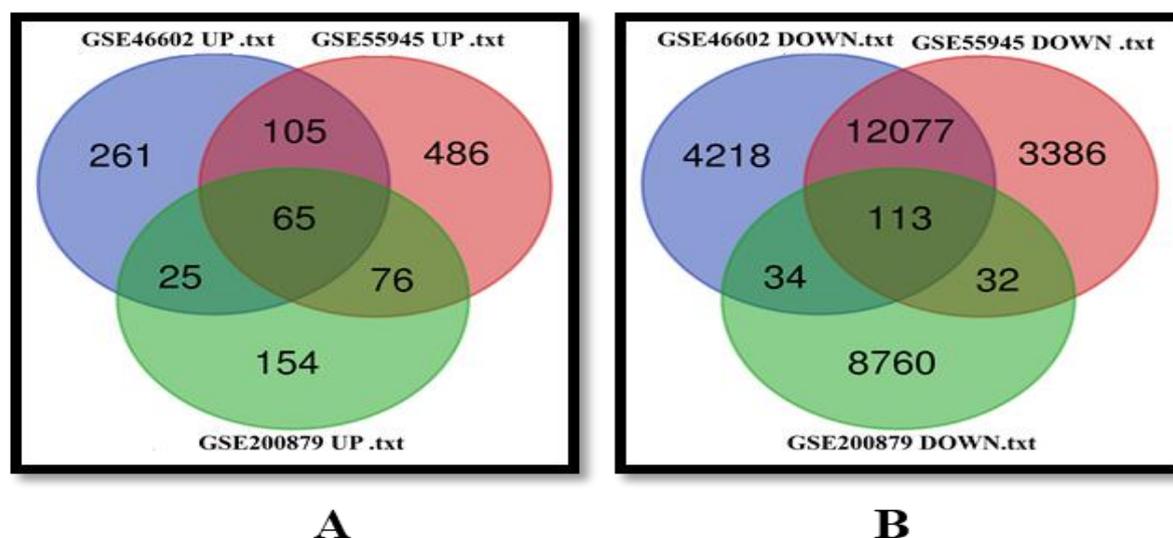
##### 3.1.1. Traitement des données et identification des DEGs

Les trois profils d'expression d'ARNm obtenus à partir de la base de données GEO ont été analysés par l'outil en ligne GEO2R avec une valeur  $P < 0,05$  et  $|\log_2\text{-fold change (FC)}| > 1$ . Il y avait 4479 DEGs dans l'ensemble de données GSE46602, dont 261 gènes surexprimés et 4218 gènes sous-exprimés. 3872 DEGs dans l'ensemble de données GSE55945, dont 486 gènes surexprimés et 3386 gènes sous-exprimés. 8914 DEGs dans l'ensemble de données GSE200879, dont 154 gènes surexprimés et 876 gènes sous-exprimés. Les traces volcaniques (*Volcano Plot*) des trois ensembles de données GEO sont illustrés dans la figure 07.



**Figure 07:** Les tracés volcaniques (*Volcano plots*) des trois ensembles de données GEO, le bleu représente les gènes sous-exprimés, le rouge représente les gènes surexprimés ( $P < 0,05$  et  $\log |FC| > 1$ ).

Les gènes ont été déterminés à l'aide d'une base de données de VENN de diagramme Expression différentielle commune (DEGs) entre trois profils. Un total de 178 DEGs communs, dont 65 gènes surexprimés et 113 gènes sous-exprimés. Ils ont été identifiés dans les trois ensembles de données GEO, comme le montre la figure 08.



**Figure 08 :** DEGs communs aux trois ensembles de données GEO. (A) 65 gènes surexprimés. (B) 113 gènes sous-exprimés.

La liste des DEGs communs surexprimés et sous-exprimés dans les trois *datasets* sont montrés dans le tableau 03.

**Tableau 03 :** Liste des DEGs communs exprimés dans les trois *datasets*.

DEGs communs	Symbole de gène
<b>Les gènes surexprimés 65 Gènes</b>	HPN HSPB6 SLC2A5 SERPINB5 KCNK3 ATF3 PCDH7 TRIM29 VSNL1 ZNF655 GSTP1 RLN1 NDNF PCA3 PPARGC1A CYP3A5 MYOF FGFR2 AOC1 MS4A8 BICD1 WIF1 PTP4A3 SFN RND3 SLC7A11 L3MBTL4 BEND4 MAGI2-AS3 LSAMP CACNA1D DPT PRDM8 TMEM45B DSC3 NEK5 GATA3 ALOX15B ID4 STIL THBS4 HOXD10 LGALS3 KCNG3 TMC5 CXCL12 RRM2 CAV1 SEMA6D SMOC1 WFDC2 KRT7 FLRT3 OR51E2 F5 SIM2 GPX3 GPM6B ADAMTS9-AS2 KRT5 CTHRC1 DLX1 AOX1 CYP4B1 KRT23
<b>Les gènes sous-exprimés 113 Gènes</b>	BRIX1 BANK1 TAOK3 POLR2J NENF LINC00161 CEP63 KLK3 ZNF461 DEDD RREB1 SP1 LRRC40 MRPL49 FAH C18orf25 TSPAN6 WDSUB1 SLC35D2 USP4 PIGB TRIP11 ZNF493 SFXN1 SYPL1 RNF181 POLA2 PSMA3 CDK6 TEAD2 PTK7 TYRO3 PARP1 ADPRH HNRNPM CTDSPL FAM122C PPP2R3C ARL13B NUCB1 SMIM1 TRAK2 STAB1 SEMG2 CEP192 INTS2 RNH1 CDC23 ARIH2 CACUL1 NT5C3A ZNF777 IRX2 SETDB2 SLC41A2 TLR4 ENOSF1 RPLP1 P4HA3 MED25 RECQL5 MRI1 COPG1 COG2 TYSND1 ST3GAL2 MBD6 KIAA1468 TJP2 BTF3 DPH7 PLEKHH1 DEPDC4 NDUFA5 RBBP9 PRR14 TNRC6C ARHGAP6 TNRC6B ZNF606 ETV1 SPTLC1 ATP5I VAPB USP25 ADK C21orf58 WHSC1L1 THAP1 IMPAD1 PACS1 CDC37 OGFOD1 AXIN1 UQCRB UACA SLX4 ATPAF2 SAR1B MARK3 ITGA6 PRSS27 PRKACB VASH2 FLVCR1 PSMC2 NRAS CEP89 RANBP3 SAP30L PLXDC2 SDHC TRIM8

### 3.1.2. Analyse de l'ontologie des gènes et de l'enrichissement des voies

L'analyse d'enrichissement des voies de GO a été réalisée pour les DEGs communs (gènes surexprimés et sous-exprimés). Le critère de sélection pour l'analyse était la valeur  $P < 0,05$ . L'analyse comprend (BP), (CC) et (MF).

Dans les BP, les gènes surexprimés étaient principalement enrichis dans les voies suivantes : *cell adhesion, axon guidance, blood circulation, prostate gland morphogenesis, negative regulation of cell cycle et positive regulation of thyroid hormone generation*. Les gènes sous-exprimés étaient enrichis dans les voies suivantes : *ubiquitin-dependent protein catabolic process, transcription from RNA polymerase II promoter, MAPK cascade, ER to Golgi vesicle-mediated transport, aerobic respiration et cell division*.

Dans les CC, les gènes surexprimés étaient plus enrichis dans les voies suivantes : *extracellular region, extracellular space, extracellular exosome cytoplasmic vesicle, plasma membrane et cell surface*. Les gènes sous-exprimés étaient plus enrichis dans les voies : *nucleoplasm, cytosol, fibrillar center chromosome, Golgi apparatus et mediator complex*.

Dans les MF, les gènes surexprimés étaient en grande partie enrichis dans les voies suivantes : *heparin binding, calcium ion binding, iron ion binding protein homodimerization activity et glutathione peroxidase activity chemorepellent activity*. Les gènes sous-exprimés étaient surtout enrichis dans les voies suivants: *protein binding, R-SMAD binding et ubiquitin protein ligase binding*.

L'analyse GO des DEGs commun est présentée dans le tableau 04.

**Tableau 04** : Analyse du GO des DEGs communs dans les trois *datasets*.

<i>Categorie</i>	<i>Term</i>	<i>N°</i>	<i>%</i>	<i>P-Value</i>
<b>Les gènes surexprimés</b>				
GOTERM_BP_DIRECT	<i>cell adhesion</i>	8	12,3	1,60E-03
GOTERM_BP_DIRECT	<i>axon guidance</i>	5	7,7	2,80E-03
GOTERM_BP_DIRECT	<i>blood circulation</i>	3	4,6	5,00E-03
GOTERM_BP_DIRECT	<i>prostate gland morphogenesis</i>	2	3,1	6,20E-03
GOTERM_BP_DIRECT	<i>negative regulation of cell cycle</i>	3	4,6	1,20E-02
GOTERM_BP_DIRECT	<i>positive regulation of thyroid hormone generation</i>	2	3,1	1,20E-02
GOTERM_CC_DIRECT	<i>extracellular region</i>	17	26,2	2,20E-04
GOTERM_CC_DIRECT	<i>extracellular space</i>	15	23,1	9,80E-04
GOTERM_CC_DIRECT	<i>extracellular exosome</i>	16	24,6	1,20E-03
GOTERM_CC_DIRECT	<i>cytoplasmic vesicle</i>	5	7,7	1,30E-02
GOTERM_CC_DIRECT	<i>plasma membrane</i>	24	36,9	1,90E-02
GOTERM_CC_DIRECT	<i>cell surface</i>	6	9,2	3,60E-02
GOTERM_MF_DIRECT	<i>heparin binding</i>	5	7,7	2,30E-03
GOTERM_MF_DIRECT	<i>calcium ion binding</i>	8	12,3	9,40E-03
GOTERM_MF_DIRECT	<i>iron ion binding</i>	4	6,2	1,10E-02
GOTERM_MF_DIRECT	<i>protein homodimerization activity</i>	7	10,8	2,70E-02
GOTERM_MF_DIRECT	<i>glutathione peroxidase activity</i>	2	3,1	6,80E-02
GOTERM_MF_DIRECT	<i>chemorepellent activity</i>	2	3,1	8,50E-02
<b>Les gènes sous-exprimés</b>				
GOTERM_BP_DIRECT	<i>ubiquitin-dependent protein catabolic process</i>	8	7	8,40E-04
GOTERM_BP_DIRECT	<i>transcription from RNA polymerase II promoter</i>	6	5,2	9,40E-03
GOTERM_BP_DIRECT	<i>MAPK cascade</i>	5	4,3	9,80E-03
GOTERM_BP_DIRECT	<i>ER to Golgi vesicle-mediated transport</i>	4	3,5	3,60E-02
GOTERM_BP_DIRECT	<i>aerobic respiration</i>	3	2,6	5,20E-02
GOTERM_BP_DIRECT	<i>cell division</i>	6	5,2	5,70E-02
GOTERM_CC_DIRECT	<i>nucleoplasm</i>	37	32,2	1,60E-04
GOTERM_CC_DIRECT	<i>Cytosol</i>	45	39,1	4,40E-04
GOTERM_CC_DIRECT	<i>fibrillar center</i>	5	4,3	7,10E-03
GOTERM_CC_DIRECT	<i>chromosome</i>	6	5,2	9,50E-03
GOTERM_CC_DIRECT	<i>Golgi apparatus</i>	13	11,3	1,20E-02
GOTERM_CC_DIRECT	<i>mediator complex</i>	3	2,6	2,00E-02
GOTERM_MF_DIRECT	<i>protein binding</i>	92	80	6,40E-06
GOTERM_MF_DIRECT	<i>R-SMAD binding</i>	3	2,6	7,90E-03
GOTERM_MF_DIRECT	<i>ubiquitin protein ligase binding</i>	5	4,3	9,30E-02

### 3.1.3. Analyse de Kyoto Encyclopedia of Genes and Genomes (KEGG)

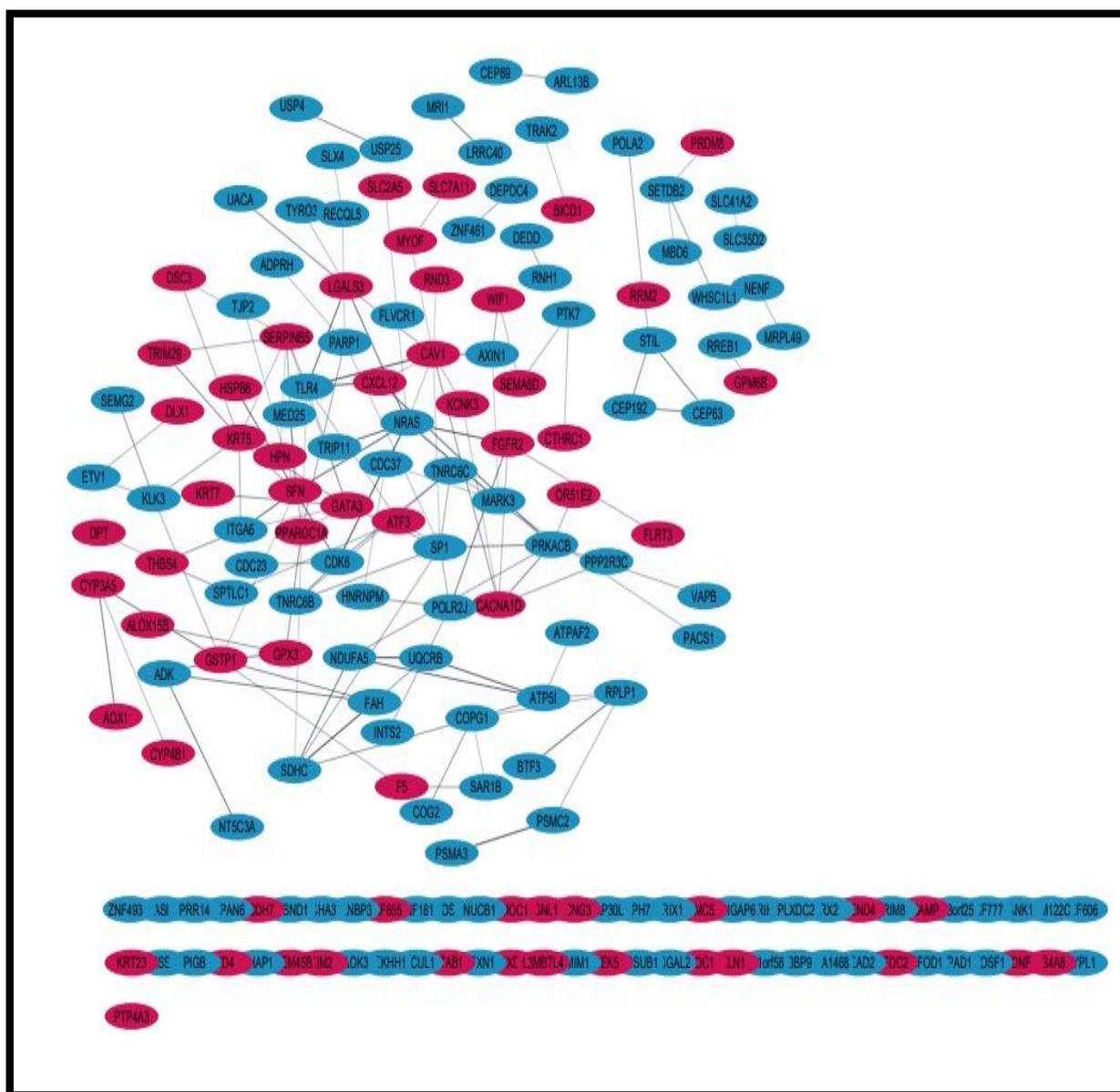
Le critère de sélection de l'analyse KEGG était la valeur  $P < 0,05$ . Les résultats montrent dix voies de signalisation identifiées. Les voies de signalisation des DEGs communs étaient principalement enrichies dans : *Glutathione metabolism*, *Drug metabolism - cytochrome P450*, *p53 signaling pathway*, Pour les gènes surexprimés et *Huntington disease*, *Pathways of neurodegeneration - multiple diseases*, *Parkinson disease*, *Prion disease*, *Alzheimer disease*, *Diabetic cardiomyopathy*, pour les gènes sous-exprimés. L'analyse de la voie KEGG des DEGs commun est présentée dans le tableau 05.

**Tableau 05 :** Analyse KEGG des DEGs courants dans le cancer de la prostate.

<i>Categorie</i>	<i>Term</i>	<i>N°</i>	<i>%</i>	<i>P-Value</i>
<b>Les gènes surexprimés</b>				
KEGG_PATHWAY	<i>Glutathione metabolism</i>	3	4,6	1,60E-02
KEGG_PATHWAY	<i>Drug metabolism - cytochrome P450</i>	3	4,6	2,50E-02
KEGG_PATHWAY	<i>p53 signaling pathway</i>	3	4,6	2,50E-02
<b>Les gènes sous-exprimés</b>				
KEGG_PATHWAY	<i>Huntington disease</i>	7	6,1	6,80E-03
KEGG_PATHWAY	<i>Pathways of neurodegeneration - multiple diseases</i>	8	7	1,60E-02
KEGG_PATHWAY	<i>Parkinson disease</i>	6	5,2	1,60E-02
KEGG_PATHWAY	<i>Prion disease</i>	6	5,2	1,80E-02
KEGG_PATHWAY	<i>Alzheimer disease</i>	7	6,1	1,90E-02
KEGG_PATHWAY	<i>Diabetic cardiomyopathy</i>	5	4,3	2,60E-02

### 3.1.4. Analyse de l'interaction protéine-protéine

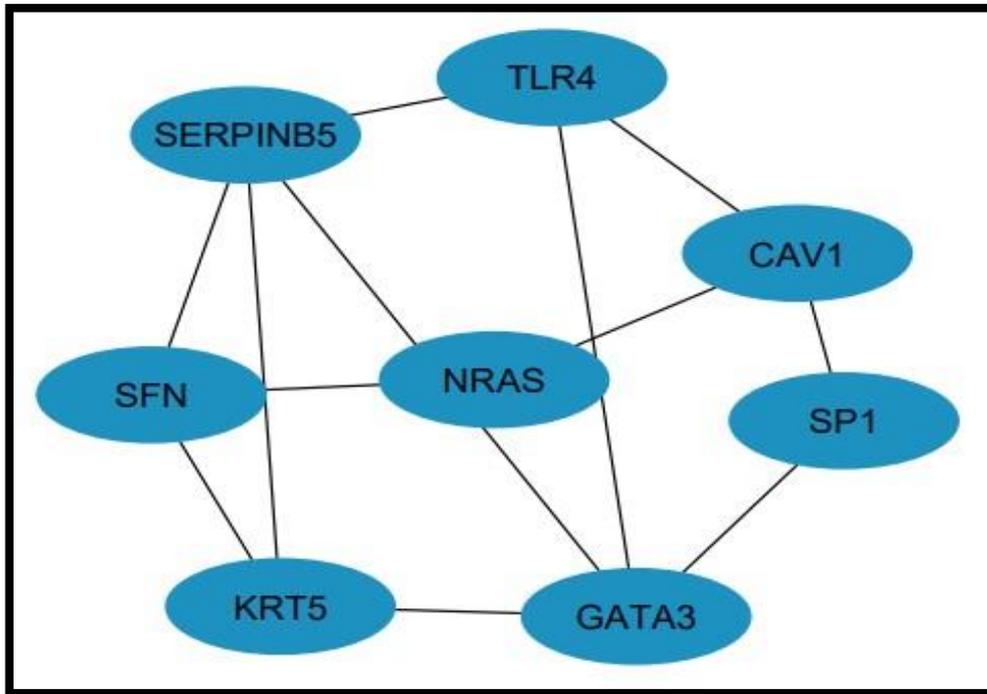
Pour mieux évaluer les interactions entre les DEGs communs identifiés, nous avons effectué ce qui suit: Créez un réseau PPI à l'aide du site Web *STRING*. Le réseau PPI de DEGs communs est illustré à la Figure 09. Le réseau PPI tracé est composé de 62 gènes surexprimés et de 112 gènes sous-exprimés. Ensuite, par l'utilisation le *plug-in MCODE* du *Cytoscape*, nous avons identifié un ensemble de 8 gènes (TLR4, CAV1, SP1, GATA3, KRT5, SFN, NRAS, SERPINB5).



**Figure 09 :** Le réseau PPI créée par la base de données *STRING*, les gènes surexprimés sont marqués en rose, les gènes sous-exprimés en bleu.

### 3.1.5. Sélection des gènes clés

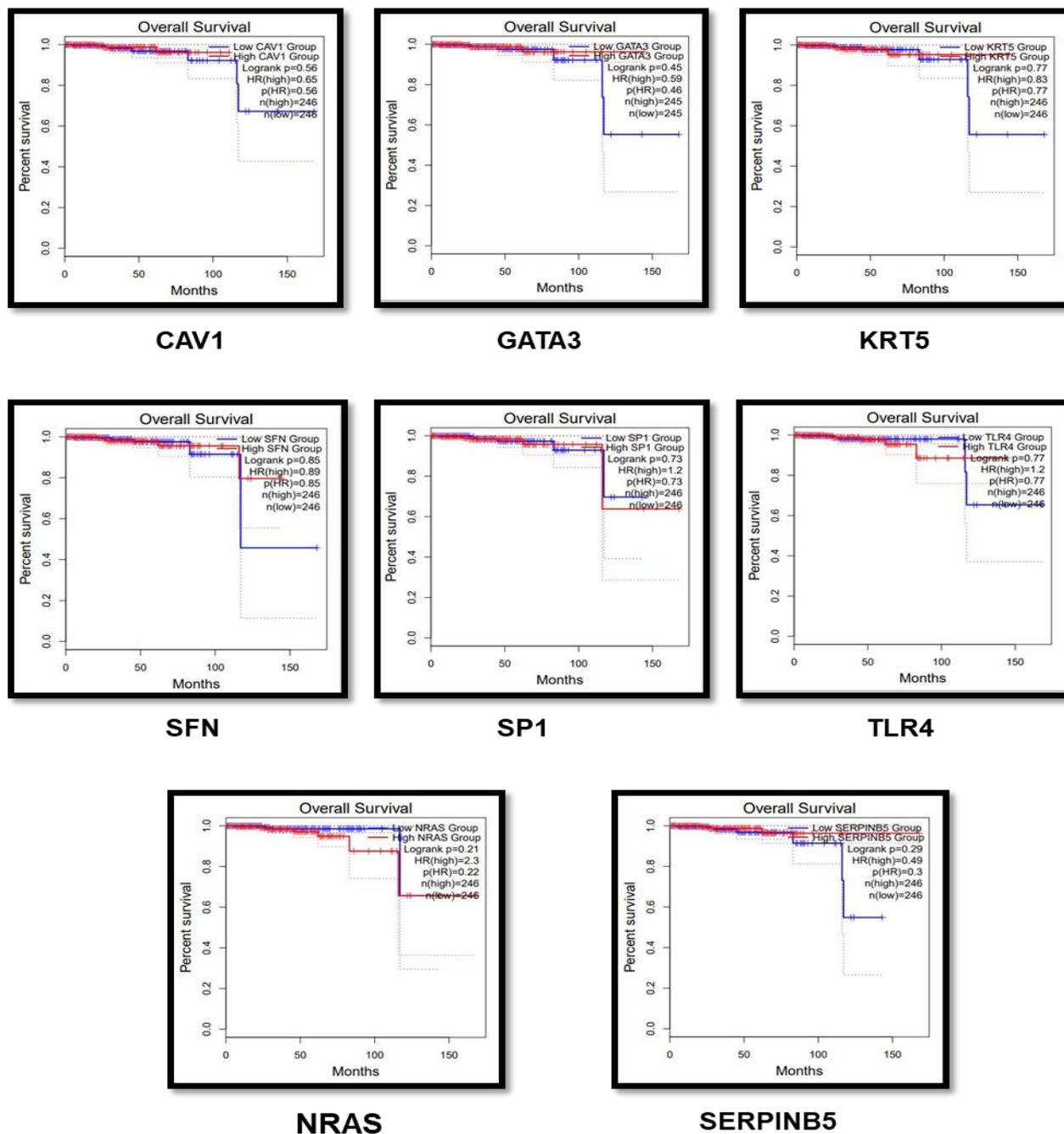
A partir *cytoHubba*, Nous avons trouvé une unité fonctionnelle du réseau PPI. Contenait 8 gènes : TLR4, CAV1, SP1, GATA3, KRT5, SFN, NRAS, SERPINB5. Ces huit gènes sont considérés comme les gènes clés associés au cancer de la prostate (figure 10).



**Figure 10** : Les huit principaux gènes clés classés par les algorithmes topologiques *Maximal Clique Centrality* (MCC) et *Degree*.

### 3.1.6. Analyse de survie globale et validation de l'expression des gènes clés

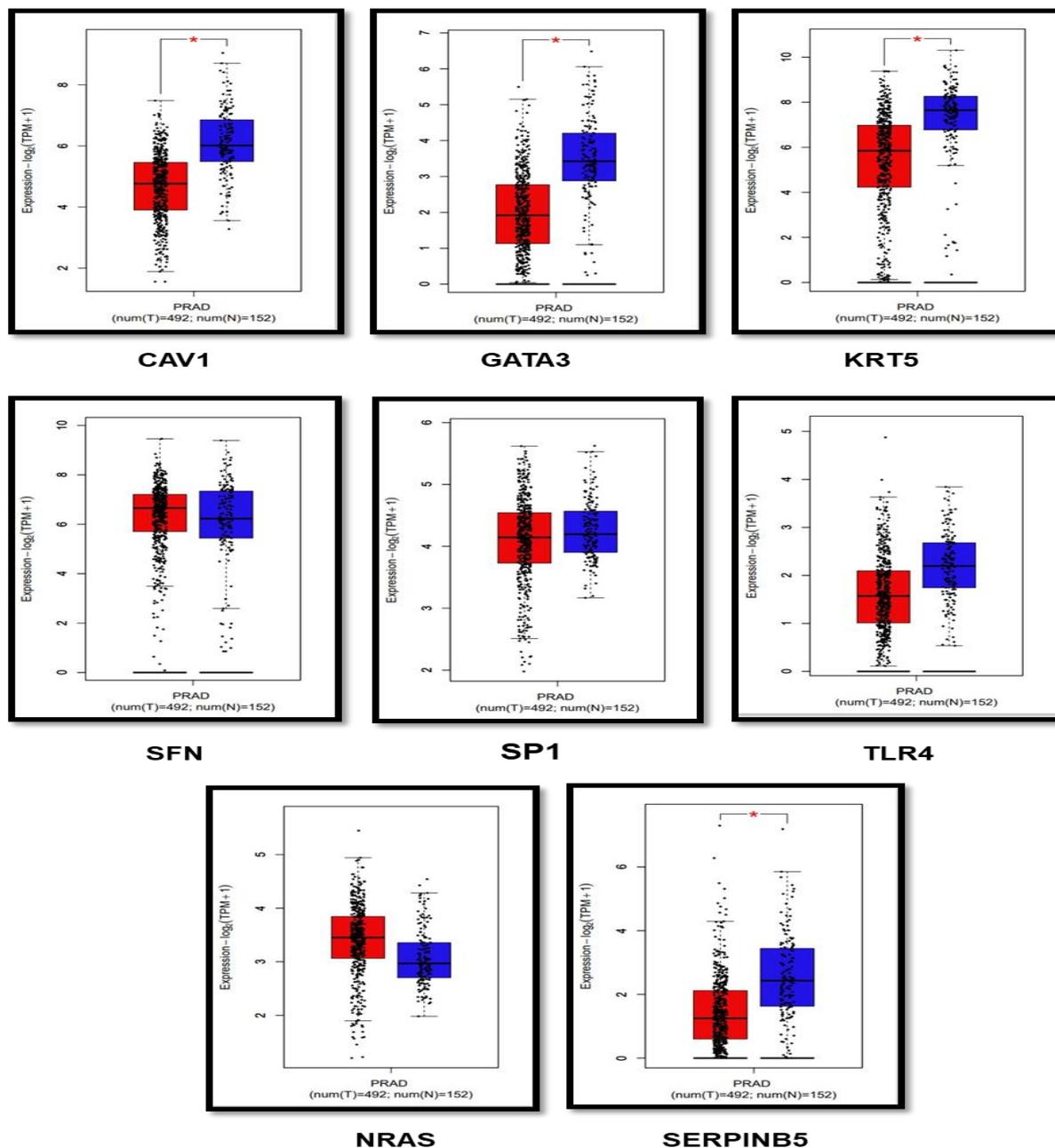
L'analyse de la survie totale des huit gènes (Figure 11) montre que l'expression de deux gènes (NRAS, SERPINB5) est associée à la mauvaise survie des patients atteints de cancer de la prostate.



**Figure 11 :** L'analyse de la survie globale des huit gènes clés par la base de données GEPIA2.

La base de données GEPIA2 (basé sur la base de données TCGA) et la base de données HPA ont été utilisés pour analyser et vérifier l'expression des 8 gènes clés dans les échantillons de PCa (Figure 12). Sur la base de l'ensemble de données PRAD dans GEPIA, six gènes sur huit (TLR4, CAV1, SP1,

GATA3, KRT5, SFN) ont été significativement sous exprimés (valeur  $P < 0,001$ ), et NRAS, SERPINB5 étaient significativement sur-exprimés dans le tissu du PC (valeur  $P < 0,001$ ).



**Figure 12** : L'analyse de l'expression des gènes clés par la base de données GEPIA2. Les tissus cancéreux sont marqués en rouge, les tissus normaux en bleu. (PRAD : *Prostate adenocarcinoma*).

### 3.2. Discussion

Le cancer de la prostate (PaC) présente les deuxièmes taux de morbidité et de mortalité les plus élevés chez les hommes dans le monde, après le cancer du poumon (79). En 2020, environ 1 414 259 nouveaux cas de PaC et 375 304 décès associés sont survenus dans le monde (80). D'ici 2030, le nombre de nouveaux cas de PaC dans le monde devrait passer à 1 700 000 et entraîner environ 500 000 décès (79).

Actuellement, les méthodes de diagnostic clinique typiques pour PaC comprennent l'examen rectal numérique (*DRE*), la mesure du niveau d'antigène spécifique de la prostate (PSA) sérique, l'imagerie par résonance magnétique multiparamétrique (*mpMRI*) et la biopsie guidée par échographie transrectale (*TRUS*) (81). Cependant, chacune de ces méthodes ne permet d'identifier qu'une partie des cancers. Pour une plus grande efficacité de diagnostic, ces méthodes sont généralement utilisées en combinaison (82).

Avec le développement rapide de la technologie de séquençage à haut débit, l'analyse bio-informatique est devenue un outil puissant dans le domaine biomédical pour prédire les gènes associés aux maladies, les sous-types de maladies et le traitement des maladies. La recherche de gènes liés à la tumeur et de leur mécanisme moléculaire associé a largement impliqué l'utilisation d'analyses de profil d'expression génique dans le but de découvrir des bio marqueurs spécifiques à la tumeur, des cibles thérapeutiques médicamenteuses et des prédicteurs de pronostic(83).

Dans ce travail, nous avons analysé et identifié l'expression différentielle des gènes clés associés au cancer de la prostate par l'utilisation des outils et méthodes bio-informatiques. D'abord, 178 DEGs communs ont été identifiés parmi les trois ensembles de données GEO, dont 65 gènes surexprimés et 113 gènes sous-exprimés. L'analyse GO sur les DEGs communs selon les BP et CC et MF a montré que les gènes surexprimés étaient principalement enrichis en *cell adhesion*, *axon guidance*, *blood circulation*, *prostate gland morphogenesis*, *negative regulation of cell cycle*, *positive regulation of thyroid hormone generation*, *extracellular region*, *extracellular space*, *extracellular exosome*, *cytoplasmic vesicle*, *plasma membrane*, *cell surface*, *heparin binding*, *calcium ion binding*, *iron ion binding*, *protein homodimerization activity*, *glutathione peroxidase activity* *chemorepellent activity*. Ces résultats sont cohérents avec le concept que le rôle du  $Ca^{2+}$  dans la majorité des voies de signalisation cellulaire impliquées dans la carcinogenèse est bien établi. L'homéostasie calcique, conséquence de la signalisation calcique, est un état d'équilibre entre l'influx, l'efflux et le stockage de  $Ca^{2+}$ . D'un point de vue physiologique, la signalisation  $Ca^{2+}$  est impliquée dans la manifestation du phénotype cellulaire, la prolifération, la différenciation, l'apoptose, et dans les activités cellulaires telles que la contraction ou la sécrétion ou l'excitabilité cellulaire. Ainsi, chaque phénotype cellulaire, qu'il soit normal ou pathologique, est caractérisé par une « signature calcique » particulière reflétant sa cinétique, son amplitude et la localisation subcellulaire des signaux calciques. En effet, si les oscillations du calcium cytosolique stimulent la prolifération cellulaire via l'activation du co-facteur de transcription  $Ca^{2+}$ -dépendant, NFAT (84) une élévation soutenue de la concentration de  $Ca^{2+}$  cytosolique induit l'apoptose

des cellules cancéreuses (85). Parce que le problème de l'homéostasie du  $\text{Ca}^{2+}$  dans les cellules cancéreuses est trop vaste, même en ce qui concerne les cellules PCa (86,39-42,30). En revanche les gènes sous-exprimés étaient plus enrichis en *ubiquitin-dependent protein catabolic process*, *transcription from RNA polymerase II promoter*, *MAPK cascade*, *ER to Golgi vesicle*, *mediated transport*, *aerobic respiration*, *cell division*, *nucleoplasm*, *cytosol*, *fibrillar center*, *chromosome*, *Golgi apparatus*, *mediator complex*, *protein binding*, *R-SMAD binding*, *ubiquitin protein ligase binding*. Après ces résultats on confirme que le cycle de division cellulaire 20 (CDC20) (31), un régulateur des points de contrôle du cycle cellulaire qui active la polypose adénomateuse colique (APC) (32). CDC20 est une clé E3 ligase, une protéine contenant le domaine de répétition WD40 qui se lie à APC et reconnaît les substrats *D-box* ou *KEN box* pour favoriser la dégradation du protéasome. APC/CDC20 exerce sa fonction cancérogène pendant la métaphase à l'anaphase en perturbant les régulateurs clés du cycle cellulaire (33).

L'analyse KEGG a montré que les DEGsal communs étaient principalement enrichies dans : *Glutathione metabolism*, *Drug metabolism - cytochrome P450*, *p53 signaling pathway*, *Huntington disease*, *Pathways of neurodegeneration - multiple diseases*, *Parkinson disease*, *Prion disease*, *Alzheimer disease*, *Diabetic cardiomyopathy*. Pour les gènes sur et sous-exprimés. donc Les glutathion S - transférases (GST) sont impliquées dans le métabolisme de nombreux cancérogènes potentiels de la prostate. Des délétions homozygotes courantes de la lignée germinale existent dans les gènes qui codent pour GST- $\mu$  (GSTM1) et GST- $\theta$  (GSTT1) et empêchent l'expression enzymatique. Pour évaluer si GSTM1 et/ou GSTT1 contribuent à l'étiologie du cancer de la prostate (CaP), l'étude de 237 cas incidents de CaP et 239 témoins appariés selon l'âge et la race. La probabilité d'avoir un CaP était augmentée chez les hommes qui avaient des génotypes non supprimés (fonctionnels) au GSTT1 (odds ratio, 1,83 ; intervalle de confiance à 95 %, 1,19-2,80) mais pas au GSTM1 (rapport de cotes, 1,07 ; intervalle de confiance à 95 %, 0,74-1,55). Aucune interaction de ces gènes dans l'étiologie du CaP n'a été observée. La GST- $\theta$  est fortement exprimée dans la prostate et peut produire des effets génotoxiques lors d'une exposition à des cancérogènes spécifiques. Ces résultats suggèrent que GSTT1 est associé au risque de CaP (34).

Ensuite, Dans cette étude aussi nous avons effectué une analyse de module sur le réseau PPI construit les huit gènes clé Grâce à la base de données GEPIA2 nous obtenons que les gènes (TLR4, CAV1, SP1, GATA3, KRT5, SFN, NRAS, SERPINB5) étaient significativement exprimés de manière différentielle entre les tissus tumoraux et les tissus normaux. L'analyse de l'OS A montré que les expressions des deux gènes clés (NRAS, SERPINB5) peuvent indiquer que les chances de survie des patients atteints de cancer de la prostate sont faibles.

TLR4, L'activation de TLR pourrait jouer un rôle différent dans la croissance tumorale du cancer de la prostate. L'expression de TLR4 dans le cancer de la prostate a été démontrée dans plusieurs modèles

animaux. Des études ont révélé une expression constitutive de TLR4 dans les cellules épithéliales de la prostate ventrale de rat ainsi que dans une lignée cellulaire d'adénocarcinome de rat et dans des cellules de culture primaire de prostate (35,36). TLR4 est également exprimé dans DU-145, PC3 et la prostate normale dans le stroma et l'épithélium (37). De plus, il a également été démontré que TLR4 est exprimé dans des échantillons cliniques de cancer de la prostate. A fonction et l'importance biologique des TLR dans le cancer de la prostate semblent complexes. Peut-être que les voies de signalisation TLR distinctes et non identifiées sont activées dans les cellules cancéreuses ou les cellules immunitaires innées au cours de la progression tumorale ; ou, la première activation du TLR dans les cellules cancéreuses ou les cellules immunitaires innées affecte de manière marquée l'activation ultérieure et les effecteurs induits. Le mystère sera approfondi et affectera le potentiel des agonistes ou antagonistes du TLR en tant qu'agents thérapeutiques anti-tumoraux (38).

NRAS, les membres fondateurs de la superfamille RAS des petites GTPases (43), sont situés au centre d'un réseau de signalisation très complexe qui contrôle de nombreux aspects des processus cellulaires normaux fondamentaux, y compris la différenciation cellulaire, la survie et la prolifération (44,45). Les protéines RAS sont également des contributeurs clés à l'oncogenèse humaine (45, 46,47). En particulier, des mutations ponctuelles qui entraînent une activation chronique du RAS se retrouvent dans environ 30 % de tous les cancers humains. Normalement, les protéines RAS sont régulées positivement par les facteurs d'échange de nucléotides guanine (GEF) tels que SOS1 qui facilitent l'échange de GDP contre GTP pour favoriser l'état actif (48), et régulées négativement par les protéines activant la GTPase (GAP) telles que p120GAP et NF1 qui catalysent l'hydrolyse du GTP en GDP et favorisent ainsi la restauration de l'état de repos (49). Les mutations oncogènes rendent généralement les protéines RAS relativement indépendantes des GEF et/ou insensibles aux GAP qui modulent les fonctions des RAS de type sauvage. L'activité du RAS peut également être améliorée dans les cancers par amplification du RAS oncogène ou de type sauvage et par l'échappement du RAS de type sauvage de la régulation suite au gain d'apports en amont tels que les récepteurs tyrosine kinases (RTK) ou les GEF, ou à la perte de régulateurs négatifs en aval tels que les GAP ou de la rétroaction des protéines *SPRY/SPRED*.

CAV1 est une protéine membranaire cancérogène associée à l'endocytose, au tissu de la matrice extracellulaire, à la distribution du cholestérol, à la migration cellulaire et à la transduction du signal. Des études antérieures ont montré que CAV1 est impliqué dans le cancer du foie, le cancer du côlon, le cancer du sein, le cancer du rein, le cancer du poumon et le cancer de la peau, etc., et agit comme un promoteur ou un inhibiteur du cancer selon le type de cancer et sa progression (50,51). Plusieurs agents endogènes et exogènes, tels que *Chrysotobenzyl*, *Cordycepin* et *Giantol*, ont été utilisés pour moduler l'expression de CAV-1 afin de réguler la progression du cancer du poumon (52,53).

Sp1 est un facteur de transcription exprimé de manière omniprésente qui appartient à une famille de doigts de zinc (57,58). On pense que Sp1 régule la transcription des gènes en se liant à un élément riche

en GC (ou boîte GC) dans le promoteur des gènes cibles. Sp1 a récemment été montré non seulement pour interagir directement avec les facteurs de transcription basaux, y compris TFIID (59) et TBP (60), mais aussi pour s'associer physiquement et coopérer fonctionnellement avec un certain nombre de facteurs de transcription spécifiques à la séquence tels que NF- $\kappa$ B (61,62), Oct-1 (63) et protéine GATA-1 (64). D'autre part, certains des facteurs de transcription peuvent altérer l'activité transcriptionnelle médiée par Sp1. Par exemple, la protéine de la leucémie promyélocytaire réprime la transcription médiée par Sp1 en l'empêchant de se lier à l'ADN (65). La protéine suppresseur de tumeur *von Hippel-Lindau* (VHL) interagit avec Sp1 et inhibe l'activité du promoteur du facteur de croissance de l'endothélium vasculaire (VEGF) (66).

GATA3 fait partie d'une famille de régulateurs transcriptionnels capables de fonctionner dans la détermination de l'identité cellulaire (71). Son expression est associée à la spécification du type de cellule dans le système immunitaire, où son action fait partie intégrante de la différenciation de plusieurs types de cellules (72). Dans la glande mammaire, GATA3 est exprimé dans les cellules épithéliales luminales différenciées tapissant les structures canalaire du sein, où il est estimé être le facteur de transcription le plus exprimé (73,74).

KRT5 est l'une des protéines de la kératine humaine, principalement exprimée dans les kératinocytes basaux épidermiques (54). *Cimpéan et al* ont rapporté que le niveau d'expression de KRT5 est en corrélation avec le pronostic et le stade TNM dans les carcinomes épidermoïdes de la tête et du cou (HNSCC) (55). Et *Ricciardelli et al* ont fondé que la surexpression de K5 dans le cancer séreux de l'ovaire est associée à la récurrence et à la résistance à la chimiothérapie (56). IL est étroitement associé aux gènes liés au cancer de la prostate ; en outre, des preuves ont été accumulées pour suggérer que KRT5 participe à l'incidence de nombreux types de cancer. Par exemple, *Breyer et al* (67) ont démontré qu'il existait une forte corrélation entre l'expression de KRT5 et la récurrence et les métastases du cancer de la vessie non invasif musculaire. Une autre étude a révélé que KRT5 pourrait réguler la formation du cytosquelette et l'invasion veineuse du cancer (68,69). De plus, *Ricciardelli et al* (70) ont vérifié la fonction de suppresseur de tumeur de KRT5 dans la réduction de la récurrence et de la résistance chimiothérapeutique du cancer de l'ovaire. Toutes ces découvertes suggèrent un rôle positif de KRT5 dans le cancer.

Le SFN agit comme un inhibiteur de l'histone désacétylase (HDAC). Ceci est cliniquement pertinent, car la modification épigénétique par l'augmentation de l'activité HDAC est impliquée dans la tumorigenèse et la progression du cancer. En raison de la nature réversible de l'altération épigénétique, l'utilisation d'inhibiteurs d'HDAC pourrait fournir le moyen d'induire les cellules malignes à revenir à leur phénotype épithélial sain d'origine. La possibilité que le SFN en tant qu'inhibiteur naturel de l'HDAC puisse prévenir, retarder ou contrecarrer les altérations épigénétiques du cancer par la consommation d'un régime enrichi en SFN est incontestablement attrayante. Néanmoins, le rôle du SFN

dans le traitement du cancer de la prostate n'est pas encore clair et les études cliniques disponibles sont controversées. Cette étude a été conçue pour étudier l'effet du SFN sur la croissance et la prolifération des cellules tumorales résistantes aux androgènes, PC3 et DU-145 (75).

L'inhibiteur de la sérine peptidase (SERPINB5), également connu sous le nom de maspine, est un membre de la famille des serpins des inhibiteurs de la sérine protéase, et structurellement, il est homologue à l'inhibiteur de l'activateur du plasminogène 1 (SERPIN1) (76). Il est proposé que SERPINB5 soit un suppresseur de tumeur. Cependant, le rôle moléculaire de SERPIN1 reste à élucider (76). Il a récemment été démontré que SERPINB5 se lie directement au collagène, une interaction qui peut contribuer à l'adhésion cellulaire (77). Il a été démontré que SERPINB5 est régulé de manière différentielle dans la progression de nombreux types de tumeurs solides, telles que la prostate, le sein et les poumons (78).

# CONCLUSION

## Conclusion

Les patients atteints de cancer de la prostate n'ont pas de manifestations cliniques évidentes, qui ne peuvent être diagnostiquées que lors d'un examen histologique, et sont souvent confinées aux organes. L'apparence clinique des patients symptomatiques atteints de cancer de la prostate dépend également de la sévérité de la maladie et de ses manifestations cliniques. Elles peuvent aller d'une urination fréquente à des douleurs sévères causées par des métastases.

Pour assurer le succès de la thérapie génique, il est essentiel de sélectionner le gène de traitement approprié afin de maximiser l'efficacité tout en minimisant la toxicité. Pour l'instant, on se concentre sur l'identification de nouveaux gènes qui sont exprimés de manière différentielle dans les cellules cancéreuses et qui peuvent modifier les effets de la tumeur.

Avec le développement des bases de données biologiques, la bio-informatique, en particulier la progression des méthodes d'exploration des données, a changé la façon dont les cibles sont découvertes en associant des idées biologiques à des outils informatiques ou à des méthodes statistiques pour extraire ou filtrer les cibles utiles. On utilise ces méthodes dans nos études pour identifier les gènes clé impliqué dans le développement et l'évolution du cancer de la prostate.

Cette étude nous a donné l'occasion d'accéder aux résultats suivants :

- Identification de 178 gènes communs entre les trois profils d'expression des patients, dont 65 gènes surexprimés et 113 gènes sous-exprimés.
- Les analyses d'enrichissement fonctionnel (GO) et (KEGG) ont montré que les gènes étaient enrichis dans de nombreux processus biologiques (BP), cellulaires (CC) et moléculaires (MF) associés au cancer de la prostate.
- 8 gènes ont été trouvés : (TLR4, CAV1, SP1, GATA3, KRT5, SFN, NRAS, SERPINB5).
- Deux gènes clés (NRAS, SERPINB5) ont été confirmés comme étant exprimés de manière différentielle entre les tissus cancéreux et les tissus normaux
- L'expression de ces deux gènes (NRAS, SERPINB5) est associée à la mauvaise survie des patients atteints de cancer de la prostate.

Ces résultats suggèrent que deux gènes clés pourraient être des bio-marqueurs potentiels pour le diagnostic du cancer de la prostate. De plus, ils pourraient jouer un rôle thérapeutique efficace dans le futur traitement du cancer de la prostate.

## **Références Bibliographiques**

**Références bibliographiques**

1. Audrey Monge (2008) Conséquences Fonctionnelles Et Structurales De L'association De Deux Mutations Du Récepteur Des Androgènes Dans Le Cancer De La Prostate ; 43.
2. Vandebos F., Chevallier D., Michiels J.F. (2004). Cancers De La Prostate : Classification Histopathologique. *Oncologie* 6 : 385 – 393.
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 68 (6):394- 424.
4. Ferlay J EM, Lam F, Colombet M, Mery L, Pineros M, Znaor A, Soerjomataram I, et al (2019). Global cancer observatory: cancer today. Lyon, France: International Agency for Research on Cancer.
5. Quinn M, Babb P (2002). Patterns and trends in prostate cancer incidence, survival, prevalence and mortality. Part I: international comparisons. *BJU Int.* 90(2):162-173.
6. Rio-Schmidt F- X., 2012, Cancer de la prostate: revue d'informations pour le pharmacien d'officine, Thèse de Doctorat, Université de Nantes, 110p.
7. <https://www.teresewinslow.com/> Male Genitourinary Anatomy © 2005 Terese Winslow Created for the National Cancer Institute.
8. Bastien L., Fourcade R-O., Makhoul B., Meria P., et al, 2011, Hyperplasie bénigne de la prostate, EMC (Elsevier Masson SAS, Paris). *Urologie*, 22 :14-29.
9. <https://www.semanticscholar.org/paper/PGC-1%CE%B1-r%C3%A9gule-un-programme-onco-m%C3%A9tabolique-capable>.
10. HAS (Haute Autorité De Santé), INCa (Institut Nationale Du Cancer)., 2010, La prise en charge du cancer de la prostate, Guide-ALD,1-8.
11. Haute Autorité de Santé - ALD n° 30 - Cancer de la prostate. Available at: [https://www.has.sante.fr/portail/jcms/c\\_725257/fr/ald-n-30-cancer-de-la-prostate](https://www.has.sante.fr/portail/jcms/c_725257/fr/ald-n-30-cancer-de-la-prostate). (Accessed: 17th July 2018).
12. Synthèse sur les bénéfices et les risques d'un dépistage du cancer de la prostate par dosage du PSA - Ref : ETATPROSTATE15 | Institut National Du Cancer. Available at: <http://www.e-cancer.fr/Expertises-et-publications/Catalogue-des-publications/Synthese-sur-les-benefices-et-les-risques-d-un-depistage-du-cancer-de-la-prostate-par-dosage-du-PSA>. (Accessed: 17th July 2018).
13. <https://www.urofrance.org/lafu-academie/formation-ducollege/referentiel-du-college-durologie-5eme-edition/chapitre-10-hypertrophie-benigne-de-la-prostate>.
14. Rozet, F. et al. [CCAFU french national guidelines 2016-2018 on prostate cancer]. *Prog. Urol.* 27 Suppl 1, S95–S143 (2016).
15. <https://fr.dreamstime.com/image-libre-droits-biopsie-prostate>.

16. Hayward, S. W. & Cunha, G. R. The prostate: development and physiology. *Radiol. Clin. North Am.* 38, 1–14 (2000).
17. Berry, S. J., Coffey, D. S., Walsh, P. C. & Ewing, L. L. The development of human benign prostatic hyperplasia with age. *J. Urol.* 132, 474–479 (1984).
18. <https://www.cancer-de-la-prostate.fr/comprendre-cancer-prostate/developpement-cancer-objectif-chirurgical.html>.
19. Belabed Z., Bouamama A., 2015. Cancer de la prostate (Etude Epidémiologique), mémoire de master, Université des Frères Mentouri Constantine, 50p.
20. FRM (Fondation Recherche Médicale). 2007, Soutenir la Fondation pour la Recherche Médicale, recherche & santé, N° 8117. ISSN 0241-0338.
21. <https://www.actuscimed.com/2011/11/denosumab-et-survie-sans-metastases.html>.
22. Grönberg, H. Prostate cancer epidemiology. *Lancet* 361, 859–864 (2003).
23. Leitzmann, M. F. & Rohrmann, S. Risk factors for the onset of prostatic cancer: age, location, and behavioral correlates. *Clin Epidemiol* 4, 1–11 (2012).
24. SEER Cancer Stat Facts. Available at: <https://seer.cancer.gov/statfacts/>. (Accessed: 4th July 2018).
25. Amundadottir, L. T. et al. A common variant associated with prostate cancer in European and African populations. *Nat. Genet.* 38, 652–658 (2006).
26. Al Olama, A. A. et al. A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. *Nat. Genet.* 46, 1103–1109 (2014).
27. Hoffmann, T. J. et al. A large multiethnic genome-wide association study of prostate cancer identifies novel risk variants and substantial ethnic differences. *Cancer Discov* 5, 878–891 (2015).
28. Al Nakouzi N., 2011, Etablissement d'un nouveau modèle préclinique de cancer de la prostate et identification de biomarqueurs de résistance au docetaxel, Thèse de Doctorat, Université Paris Sud-Paris XI. 245p.
29. Fred Saad, et Paul Perotte, 2003. Mise à jour sur le cancer de prostate, *Le clinicien*.
30. Santella L, Ercolano E, Nusco GA. Le cycle cellulaire : une nouvelle entrée dans le domaine de la signalisation Ca<sup>2+</sup>. *Cell Mol Life Sci* 2005.
31. H. Huang, Q. Zhang, C. Ye, J. Lv, X. Liu, L. Chen, et al. Identification de marqueurs pronostiques du cancer de la prostate de haut grade par une approche bioinformatique intégrée *J. Cancer Res. Clin*, 143 (12) (2017), p. 2571 – 2579.
32. J. Weinstein , FW Jacobsen , J. Hsu-Chen , T. Wu , LG Baum Une nouvelle protéine de mammifère, p55CDC, présente dans les cellules en division est associée à une activité de protéine kinase et présente une homologie avec les protéines du cycle de division cellulaire de *Saccharomyces cerevisiae* Cdc20 et Cdc4 *Mol Cell Biol* , 14 (5) (1994) , pages 3350 – 3363.

33. H.Yu \_Cdc20 : un activateur WD40 pour une machine de dégradation du cycle cellulaire Mol. Cellule, 27 (1) (2007), p. 3 – 16.
34. Timothy R. Rebbeck ; Amy H. Walker ; Julie M. Jaffé ; David L. White ; Alan J. Wein ; S. Bruce Malkowicz, Génotypes du glutathion S -transférase- $\mu$  (GSTM1) et - $\theta$  (GSTT1) dans l'étiologie du cancer de la prostate, Cancer Epidemiol Biomarkers Prev (1999).
35. Quintar AA, Roth FD, De Paul AL, Aoki A, Maldonado CA. Récepteur Toll-like 4 dans la prostate du rat : modulation par la testostérone et infection bactérienne aiguë dans les cellules épithéliales et stromales. Biol Reprod (2006) 75 (5):664–72.
36. Gatti G, Rivero V, Motrich RD, Maccioni M. Les cellules épithéliales de la prostate peuvent agir comme des capteurs précoces d'infection en régulant à la hausse l'expression de TLR4 et les médiateurs pro-inflammatoires lors de la stimulation du LPS. J Leukoc Biol (2006) 79 (5):989–98.
37. Pei Z, Lin D, Song X, Li H, Yao H. La signalisation TLR4 favorise l'expression du VEGF et du TGF $\beta$ 1 dans les cellules PC3 épithéliales de la prostate humaine induites par le lipopolysaccharide. Cell Immunol (2008) 254 (1):20–7.
38. Gatti G, Quintar AA, Andreani V, Nicola JP, Maldonado CA, Masini-Repiso AM, et al. Expression du récepteur de type péage 4 dans la glande prostatique et son association avec la gravité du cancer de la prostate. Prostate (2009) 69 (13):1387–97.
39. Orrenius S, Zhivotovsky B, Nicotera P. Régulation de la mort cellulaire : le lien calcium-apoptose. Nat Rev Mol Cell Biol 2003.
40. Rizzuto R, Pinton P, Ferrari D, Chami M, Szabadkai G, Magalhaes PJ et al. Calcium et apoptose : faits et hypothèses. Oncogène 2003.
41. Lipskaia L, Lompre AM. Altération de la cinétique temporelle de la signalisation Ca<sup>2+</sup> et contrôle de la croissance et de la prolifération. Cellule Biol 2004.
42. Munaron L, Antoniotti S, Lovisolo D. Signaux calciques intracellulaires et contrôle de la prolifération cellulaire : combien de mécanismes ? J Cell Mol Med 2004.
43. AM Kidger et al. La régulation de la signalisation oncogène Ras/ERK par des protéines kinases phosphatases (MKP) activées par un mitogène à double spécificité, Sémin. Cellule Dév. Biol. (2016).
44. CW Brennan et al. Le paysage génomique somatique du glioblastome Cellule(2013).
45. RG Verhaak et al. L'analyse génomique intégrée identifie les sous-types cliniquement pertinents de glioblastome caractérisés par des anomalies dans PDGFRA, IDH1, EGFR et NF1 Cellule cancéreuse(2010).
46. IG Kolfschoten et al. Un dépistage génétique identifie PITX1 comme un suppresseur de l'activité RAS et de la tumorigénicité Cellule(2005).
47. DF Calvisi et al. L'inactivation des protéines activatrices de Ras GTPase favorise l'activité illimitée de Ras de type sauvage dans le cancer du foie humain J. Hépatol. (2011).

48. E. Grabocka et al. Les H- et N-Ras de type sauvage favorisent la tumorigenèse induite par les K-Ras mutants en modulant la réponse aux dommages à l'ADN Cellule cancéreuse (2014).
49. JC Wolfman et al. Le cN-Ras endogène fournit un signal anti-apoptotique à l'état d'équilibre. *J. Biol. Chim.* (2000).
50. Senetta R, Stella G, Pozzi E, et al. La cavéoline-1 en tant que promoteur de la propagation tumorale : quand, comment, où et pourquoi. *J Cell Mol Med* 2013.
51. Shi YB, Li J, Lai XN, et al. Rôles multiformes de la cavéoline-1 dans le cancer du poumon : une nouvelle enquête axée sur l'apparition, le développement et le traitement des tumeurs. *Cancers (Bâle)* 2020.
52. Petpiroon N, Bhummaphan N, Tungsukruthai S, et al. Inhibition par le chrysotobibenzyle de la migration des cellules cancéreuses du poumon par la médiation dépendante de la cavéoline-1 du commutateur d'intégrine et la sensibilisation des cellules cancéreuses du poumon à l'apoptose médiée par le cisplatine. *Phytomédecine* 2019.
53. Charoenrungruang S, Chanvorachote P, Sritularak B, et al. Le gigantol, un bibenzyle de *Dendrobium draconis*, inhibe le comportement migratoire des cellules cancéreuses pulmonaires non à petites cellules. *J Nat Prod* 2014.
54. Karantza V. Les kératines dans la santé et le cancer : plus que de simples marqueurs de cellules épithéliales. *Oncogène* 2011.
55. Cimpean AM, Balica RA, Doros IC, et al. Récepteur du facteur de croissance épidermique (EGFR) et kératine 5 (K5) : acteurs clés polyvalents définissant les sous-classes pronostiques et thérapeutiques des carcinomes épidermoïdes de la tête et du cou. *Cancer Génomique Protéomique* 2016.
56. Ricciardelli C, Lokman NA, Pyragius CE, et al. La surexpression de la kératine 5 est associée à la récurrence sévère du cancer de l'ovaire et à la résistance à la chimiothérapie. *Oncocible* 2017.
57. Berg, JM (1992) Sp1 et la sous-famille des protéines à doigts de zinc avec des sites de liaison riches en guanine. *Proc. Natl Acad. Sci. Etats-Unis*, 89, 11109–11110.
58. Philipsen, S. et Suske, G. (1999) Une histoire de trois doigts : la famille des facteurs de transcription Sp/XKLF chez les mammifères. *Nucleic Acids Res.*, 27, 2991–3000.
59. Gill, G., Pascal, E., Tseng, ZH et Tjian, R. (1994) Un patch hydrophobe riche en glutamine dans le facteur de transcription Sp1 entre en contact avec le composant dTAfIII10 du complexe *Drosophila* TFIID et médie l'activation transcriptionnelle. *Proc. Natl Acad. Sci. Etats-Unis*, 91, 192–196.
60. Emili, A., Greenblatt, J. et Ingles, CJ (1994) Interaction spécifique à l'espèce du domaine d'activation riche en glutamine de Sp1 avec la protéine de liaison à la boîte TATA. *Mol. Cellule Biol.*, 14, 1582–1593.

61. Perkins, ND, Edwards, NL, Duckett, CS, Agranoff, AB, Schmid, RM et Nabel, GJ (1993). Une interaction coopérative entre NF-kappa B et Sp1 est nécessaire pour l'activation de l'activateur du VIH-1. *EBO J.*, 12, 3551 –3558.
62. Perkins, ND, Agranoff, AB, Pascal, E. et Nabel,GJ ( 1994 ) Une interaction entre les domaines de liaison à l'ADN de RelA(p65) et Sp1 médie l'activation du gène du virus de l'immunodéficience humaine. *Mol. Cellule. Biol.*, 14, 6570 –6583.
63. Strom, AC, Forsberg, M., Lillhager, P. et Westin, G. (1996) Les facteurs de transcription Sp1 et Oct-1 interagissent physiquement pour réguler l'expression du gène humain U2 snRNA. *Nucleic Acids Res.* , 24, 1981 –1986.
64. Mérিকা, M. et Orkin, SH (1995) Synergie fonctionnelle et interactions physiques du facteur de transcription érythroïde GATA-1 avec les protéines de la famille Krüppel Sp1 et EKLF. *Mol. Cellule. Biol.*, 15, 2437 –2447.
65. Vallian, S., Chin, KV et Chang, KS (1998) La protéine de la leucémie promyélocytaire interagit avec Sp1 et inhibe sa transactivation du promoteur du récepteur du facteur de croissance épidermique. *Mol. Cellule. Biol.*, 18, 7147 –7156.
66. Mukhopadhyay, D., Knebelmann, B., Cohen, HT, Ananth, S. et Sukhatme, vice-président (1997) Le produit du gène suppresseur de tumeur de von Hippel-Lindau interagit avec Sp1 pour réprimer l'activité du promoteur du facteur de croissance endothélial vasculaire. *Mol. Cellule. Biol.*, 17, 5629 –5639.
67. Ricciardelli C, Lokman NA, Pyragius CE, Ween MP, Macpherson AM, Ruszkiewicz A, Hoffmann P, Oehler MK. La surexpression de la kératine 5 est associée à la récurrence sévère du cancer de l'ovaire et à la résistance à la chimiothérapie. *Oncologie*. 2017.
68. Breyer J, Wirtz RM, Erben P, Rinaldetti S, Pire TS, Stoehr R, Eckstein M, Sikic D, Denzinger S, Burger M, Hartmann A, Otto W. La surexpression de FOXM1 est associée à des résultats indésirables et prédit la réponse à traitement par instillation dans le cancer de la vessie non invasif sur le plan musculaire de stade pT1. *BJU Int*. 2019.
69. Berens EB, Sharif GM, Schmidt MO, Yan G, Shuptrine CW, Weiner LM, Glasgow E, Riegel AT, Wellstein A. La protéine associée à la kératine 5-5 contrôle la fonction cytosquelettique et l'invasion vasculaire des cellules cancéreuses. *Oncogène*. 2017.
70. Wang CC, Bajikar SS, Jamal L, Atkins KA, Janes KA. Un circuit de régulation TGFBR3-JUND-KRT5 dépendant du temps et de la matrice dans les cellules épithéliales mammaires uniques et les prémaligènes de type basal. *Nat Cell Biol*. 2014.
71. Rothenberg EV. Pilotes transcriptionnels du programme de lignée des lymphocytes T. *Curr Opin Immunol* 2012.
72. Tindemans I, Serafini N, Di Santo JP, Hendriks RW. Fonction GATA-3 dans l'immunité innée et adaptative. *Immunité* 2014

73. Kouros-Mehr H, Werb Z. Régulateurs candidats de la morphogénèse de ramification mammaire identifiés par analyse de transcription à l'échelle du génome. *Dev Dyn* 2006.
74. Kouros-Mehr H, Slorach EM, Sternlicht MD, Werb Z. GATA-3 maintient la différenciation du destin des cellules luminales dans la glande mammaire. *Cellule* 2006.
75. Jochen Rutz , Sarah Thalier, Sébastien Maxeiner, Félix K.-H. Chun, Roman A. Blaheta. Département d'urologie, Université Goethe, 60323 Francfort-sur-le-Main, Allemagne 2020.
76. Ellis V, Murphy G. Stratégies cellulaires pour le ciblage protéolytique pendant la migration et l'invasion. *FEBS Lett.* 2001.
77. Blacque OE, Worrall DM. Preuve d'une interaction directe entre la serpine suppresseur de tumeur, la maspine et le collagène de types I et III. *J Biol Chem.* 2002.
78. Dzinic SH, Kaplun A, Li X, Bernardo M, Meng Y, Dean I, Krass D, Stemmer P, Shin N, Lonardo F, Sheng S. Identification d'un déterminant intrinsèque critique pour la localisation et la fonction subcellulaires de maspin. *PLoS One.* 2013.
79. Ferlay J, Soerjomataram I, Dikshit R, et al. Incidence et mortalité du cancer dans le monde : sources, méthodes et tendances majeures dans GLOBOCAN 2012. *Int J Cancer* 2015.
80. Zhu Y, Mo M, Wei Y, et al. Épidémiologie et génomique du cancer de la prostate chez les hommes asiatiques. *Nat Rev Urol* 2021.
81. Noureldin M, Eldred-Evans D, Khoo CC, et al. Article de synthèse : Biopsies ciblées par IRM pour le diagnostic et la prise en charge du cancer de la prostate. *World J Urol* 2021.
82. Grozescu T, Popa F. Cancer de la prostate entre pronostic et traitement adéquat/approprié. *J Med Life* 2017.
83. Li A, He J, Zhang Z, et al. L'analyse bioinformatique intégrée révèle des gènes marqueurs et des cibles thérapeutiques potentielles pour l'hypertension artérielle pulmonaire. *Gènes (Bâle)* 2021.
84. Thebault S, Roudbaraki M, Sydorenko V, Shuba Y, Lemonnier L, Slomianny C et al. Les récepteurs alpha1-adrénérgiques activent les canaux cationiques perméables au Ca (<sup>2+</sup>) dans les cellules épithéliales du cancer de la prostate. *J Clin Invest* 2003.
85. Vanoverberghe K, Mariot P, Vanden Abeele F, Delcourt P, Parys JB, Prevarskaya N. Mécanismes de la signalisation calcique induite par l'ATP et de l'arrêt de la croissance dans les cellules cancéreuses de la prostate humaine. *Calcium cellulaire* 2003.
86. Hajnoczky G, Davies E, Madesh M. Signalisation calcique et apoptose. *Biochem Biophys Res Commun* 2003.

## Glossaire

**Cytoscape** : est un logiciel libre utilisé en bio-informatique pour la visualisation de réseaux d'interaction moléculaire. Plus généralement, il permet l'analyse et la visualisation des réseaux complexes.

**DAVID (*Database for Annotation Visualization and Integrated Discovery*)**: fournit un ensemble complet d'outils d'annotation fonctionnelle permettant aux chercheurs de comprendre la signification biologique derrière de grandes listes de gènes.

**FC (*Fold Change*)** : est une mesure décrivant à quel point une quantité change entre une mesure originale et une mesure ultérieure. Il est défini comme le rapport entre les deux grandeurs ; pour les quantités A et B, le pli de B par rapport à A est  $B/A$ .

**GEO (*Gene Expression Omnibus*)** : est un référentiel public international qui archive et distribue librement les puces à ADN, le séquençage de nouvelle génération et d'autres formes d'ensembles de données génomiques fonctionnelles à haut débit.

**GO (*Gene Ontology*)** : est une initiative bio-informatique majeure visant à unifier la représentation des attributs des gènes et des produits génétiques dans toutes les espèces.

**Adénocarcinome** : est une tumeur maligne développée aux dépens d'un épithélium glandulaire. Le terme est à distinguer de celui d'adénome qui désigne une tumeur développée aux dépens d'un épithélium glandulaire, mais bénigne.

**KEGG (*Kyoto Encyclopedia of Genes and Genomes*)** : est une collection de bases de données traitant des génomes, des voies biologiques, des maladies, des médicaments et des substances chimiques.

**NCBI (*National Center for Biotechnology Information*)** : une série de bases de données pertinentes pour la biotechnologie et la biomédecine et constitue une ressource importante pour les outils et services bio-informatiques. Les principales bases de données comprennent *GenBank* pour les séquences d'ADN et PubMed, une base de données bibliographique pour la littérature biomédicale.

**OS (*Overall Survival*)** : Il s'agit du pourcentage de personnes dans une étude ou un groupe de traitement encore en vie pendant une période donnée après le diagnostic. C'est une méthode de description du pronostic dans certaines conditions de la maladie.

**P (Valeur de probabilité)** : est utilisé dans le contexte du test d'hypothèse nulle afin de quantifier la signification statistique d'un résultat.

**Venn diagram** : est un style de diagramme largement utilisé qui montre la relation logique entre les ensembles, popularisé par John Venn dans les années 1880.

***Volcano plot*** : est un type de nuage de points utilisé pour identifier rapidement les changements dans de grands ensembles de données composés de données répliquées.