



الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique Et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère De L'enseignement Supérieur Et De La Recherche Scientifique



Université Constantine 1 Frères Mentouri
Faculté des Sciences de la Nature et de la Vie

جامعة قسنطينة 1 الإخوة منتوري
كلية علوم الطبيعة و الحياة

Département : Biologie Appliquée

قسم : البيولوجيا التطبيقية

Mémoire présenté en vue de l'obtention du Diplôme de Master

Domaine : Sciences de la Nature et de la Vie

Filière : Sciences Biologiques

Spécialité : Bioinformatique

N° d'ordre :

N° de série :

Intitulé :

Relation entre épigénétique et infertilité féminine : Approche *in Silico*

Présenté par : ZAIBET Amira

Le : 10/06/2024

Jury d'évaluation :

Président : MEDJROUBI Med Larbi (MCB U Constantine 1 Frères Mentouri).

Encadrant : BENZAADA Mostfa (MCA – U Constantine 1 Frères Mentouri).

Examinatrice : MEZIANI Y (MAB- U Constantine 1 Frères Mentouri).

Année universitaire
2023 – 2024

Remerciements

Premièrement, je tiens à exprimer ma profonde gratitude à Dieu Tout-Puissant de m'avoir accordé la force et la patience nécessaires pour surmonter toutes les difficultés et mener à bien ce travail.

Je tiens à exprimer ma profonde gratitude à mon encadrant, Monsieur **BESAADA Mostefa**, pour sa guidance, ses précieux conseils et son soutien constant tout au long de cette recherche. Son expertise et sa disponibilité ont été déterminantes pour l'aboutissement de ce projet.

Je remercie également chaleureusement les membres du jury, Monsieur **MEDJROUBI Med Larbi** et Madame **MEZIANI** pour avoir examiné et évalué mon travail avec attention.

Mes sincères remerciements vont également à tous nos enseignants, qui ont contribué à ma formation ainsi qu'à celle de mes camarades, partageant ainsi leur savoir et leur expérience, et nous permettant de progresser dans ce domaine.

Je souhaite remercier tout particulièrement le doctorant **BEN DAHMANE Abdelhafedh** pour son aide précieuse et son soutien. Ses conseils avisés et sa disponibilité ont été d'une grande aide pour moi.

Je n'oublie pas de remercier mes collègues de cette spécialité pour leur camaraderie, leur entraide et les moments partagés durant ces années d'études. Leur soutien a été précieux.

Enfin je souhaite exprimer toute ma reconnaissance et mon amour à mes parents. Leur soutien inconditionnel, leur amour et leurs sacrifices m'ont permis d'arriver là où je suis aujourd'hui. Merci pour votre confiance et vos encouragements constants.

Dédicaces

À mes chers parents,

Qui m'ont donné la vie, m'ont nourri d'amour et de sagesse, et m'ont toujours soutenu dans mes projets et mes ambitions.

Mon Cher Papa «Abdelaziz», Mon âme et mon seul et premier ami de toute ma vie. Je ne trouve pas les mots pour exprimer ma gratitude pour ton amour, ta sagesse et ton soutien inconditionnel. Ils ont été mes guides et ma lumière tout au long de ma vie. Je suis tellement fière d'avoir un papa comme toi, un véritable héros à mes yeux, le seul qui me comprend et me soutient infailliblement.

Ma Maman chérie, «Rahima», Au cœur généreux et à l'esprit si fort, tu as toujours été mon exemple de patience et de force, et tu le resteras toujours. Merci d'être à mes côtés, de me guider et de me soutenir dans toutes mes épreuves, jusqu'à mon arrivée à bon port.

À mon frère adoré, «Mouhamed Al Amine», Le meilleur frère que l'on puisse rêver, merci pour ton soutien indéfectible dans toutes les étapes de ma vie, et plus particulièrement cette année qui a été si difficile. Tu es un pilier pour moi, et je sais que je peux toujours compter sur toi.

À ma «Manoula», la plus belle des sœurs, au cœur d'or et à l'âme pure, qui m'a toujours offert un sourire réconfortant et une oreille attentive. Merci d'être toujours à mes côtés, prête à me soutenir et à me réconforter dans les moments difficiles comme dans les moments de joie. Je t'aime le plus dans le monde.

À celles qui font de ma vie une aventure merveilleuse, à mes sœurs et non seulement des amies,

«Imene» et «Malak» chacune d'entre vous a marqué ma vie d'une empreinte unique. Votre soutien et votre amitié ont été une lumière et une source de joie inestimables.

À la personne la plus importante de ma vie : moi-même.

Merci de n'avoir jamais renoncé

Amira

Résumé

L'étude des mécanismes épigénétiques est cruciale pour comprendre la régulation des gènes impliqués dans la reproduction féminine. Ce travail vise à explorer comment les îlots CpG dans les promoteurs de 23 gènes influencent l'infertilité féminine. À l'aide de méthodes *in silico* et de bioinformatique avancée, nous avons identifié et analysé les motifs CpG au sein des promoteurs des gènes sélectionnés. Nos analyses révèlent des modifications épigénétiques significatives dans plusieurs gènes, suggérant leur rôle potentiel dans la régulation de la fertilité. Ces résultats offrent de nouvelles perspectives pour la recherche et le diagnostic de l'infertilité féminine, soulignant l'importance des approches bioinformatiques pour l'étude des mécanismes épigénétiques.

Mots Clés : Épigénétique, Infertilité féminine, *in silico*, îlots CpG, bioinformatique, reproduction féminine.

Abstract

Studying epigenetic mechanisms is crucial for understanding the regulation of genes involved in female reproduction. This work aims to explore how CpG islands in the promoters of 23 genes influence female infertility. Using *in silico* methods and advanced bioinformatics, we identified and analyzed CpG motifs within the promoters of the selected genes. Our analyses reveal significant epigenetic modifications in several genes, suggesting their potential role in regulating fertility. These findings offer new perspectives for research and diagnosis of female infertility, highlighting the importance of bioinformatic approaches in studying epigenetic mechanisms.

Keywords: Epigenetic, female infertility, *in silico*, CpG islands, Bioinformatics, female reproduction.

ملخص

دراسة آليات علم الوراثة فوق الجينية مهمة لفهم تنظيم الجينات المشاركة في التكاثر الأنثوي. يهدف هذا العمل إلى استكشاف كيف تؤثر جزر CpG في محفزات 23 جيناً على العقم الأنثوي. باستخدام الاساليب الحاسوبية ونهج المعلوماتية الحيوية المتقدمة، قمنا بتحديد وتحليل أنماط CpG داخل محفزات الجينات المختارة. تكشف تحليلاتنا عن تعديلات جينية كبيرة في عدة جينات، مما يشير إلى دورها المحتمل في تنظيم الخصوبة. تقدم هذه النتائج أفقاً جديدة للبحث والتشخيص في مجال العقم الأنثوي، مما يبرز أهمية النهج المعلوماتي الحيوي في دراسة الآليات الجينية.

الكلمات المفتاحية : علم الوراثة فوق الجينية, العقم الانثوي, الاساليب الحاسوبية, جزر CpG, المعلوماتية الحيوية, التكاثر الانثوي.

Liste des figures:

Figure 1 Séquence d'un dinucléotide CpG d'un brin d'ADN comparé à une paire de bases C-G sur des brins complémentaires	13
Figure 2 Exemple de la recherche de gènes dans la littérature scientifique	17
Figure 3 Interface de recherche PubMed	18
Figure 4 Portail Nucleotide NCBI	21
Figure 5 Identification et sélection de variants du gène FOXL2 chez Homo sapiens.....	21
Figure 6 Représentation d'une séquence dans Sublime Text	22
Figure 7 Localisation de la région promotrice en amont de la CDS du gène FOXL2	22
Figure 8 Alignement multiple des régions promotrices du gène FOXL2.....	23
Figure 9 Alignement multiple des régions promotrices de tous les gènes.....	23
Figure 10 Script Python pour l'analyse du contenu en CG et des motifs CG	24
Figure 11 La base de données DBCAT	25
Figure 12 Caractérisation structurale des îlots CpG par DBCAT	25
Figure 13 Modèles de substitution de nucléotides par maximum de vraisemblance	26
Figure 14 Sélection du modèle optimal par BIC	26
Figure 15 L'outil WebLogo.....	27
Figure 16 Interface du logiciel Dotlet avec la matrice identité et la taille de la fenêtre.....	28
Figure 17 Analyse phylogénétique des gènes liés à l'infertilité féminine.....	30
Figure 18 AMS des gènes FSHB et WNT4 révèle des régions de similarité importantes	31
Figure 19 Exploration des pourcentages en CG et des motifs CG du gène FOXL2	32
Figure 20 Distribution des motifs CG et GC dans le promoteur du gène FOXL2	32
Figure 21 Représentation par logo de l'alignement des régions promotrices	34
Figure 22 Représentation par Logo de l'alignement des gènes du groupe 1.....	35
Figure 23 Représentation par Logo de l'alignement des gènes du groupe 2.....	35
Figure 24 Représentation par Logo de l'alignement des gènes du groupe 3.....	35

Liste des tableaux :

Tableau 1 Caractéristiques du matériel informatique utilisé.....	18
Tableau 2 Outils bioinformatique utilisés et leurs applications	18
Tableau 3 Résultats d'analyse des ilots CpG dans la région promotrice	33

Liste des abréviations :

ACOG: Collège Américain Des Obstétriciens Et Gynécologues

AMP: Assistance Médicale à la Procréation

AMS: Alignement Multiple des Séquences

AND: Acide Désoxyribonucléique

ARN: Acide Ribonucléique

BIC: Critère D'information Bayésien

CpG: Cytosine-phosphate-Guanine

DBCAT: Base de Données des Îlots CpG et Outils d'Analyse

DNMTS: DNA Méthyltransférases

FIV: Fécondation In Vitro

FSH: Hormones Folliculostimulantes

FT: Facteur De Transcription

HAT: Histones Acétyltransférases

HDAC: Histones Désacétyltransférases

IMC: Indice De Masse Corporelle

IOP: Insuffisance Ovarienne Prématuration

LH: Lutéinisantes

NCBI: Centre Américain Pour Les Informations Biotechnologiques

OMS: Organisation Mondiale De La Santé

Pb: Paire De Base

PE: Perturbateur Endocrinien

PRMT: Protein Arginine Méthyltransférases

RP: Région Promotrice

SAM: S-Adénosyl-Méthionine

SET: Su [Var] 3-9, Enhancer Of Zeste, Trithorax

SOPK: Syndrome Des Ovaires Polykystiques

Table des matières

Résumé	iii
Liste des figures:	vi
Liste des tableaux :	vii
Liste des abréviations :	viii
Introduction Générale	1
CHAPITRE 01 :	2
Synthèse Bibliographique	2
Introduction.....	5
1 Infertilité féminine:	5
2 Les causes de l'infertilité féminine :	5
2.1 Causes physiologiques:.....	5
2.2 Causes génétiques:.....	5
2.3 Causes épigénétiques:	6
3 Epigénétique et contrôle de l'expression des gènes :	6
3.1 Mécanismes épigénétiques.....	6
3.1.1 Méthylation de l'ADN :	7
3.1.2 Modification des histones:.....	7
3.1.3 Les ARN non codants :	8
4 La génétique de l'infertilité :	8
4.1 Les catégories fonctionnelles des gènes impliqués dans la reproduction féminine:	9
4.2 L'épigénétique : Un facteur clé de la fertilité au-delà du contrôle génétique:	9
5 Infertilité et épigénétique :	9
5.1 Influence de l'épigénétique sur l'infertilité féminine :	10
5.1.1 Facteurs environnementaux et mode de vie :	10
5.2 Empreinte parentale : un aspect de l'épigénétique.....	12
5.3 L'apport de l'étude de l'épigénétique :	12
6 Les îlots CpG :	12
6.1 Rôle dans la régulation de l'expression :	13
6.2 Caractéristiques structurales :	13
6.3 Lien avec l'infertilité féminine :	13
7 Approche in silico :	14
7.1 Principe	14
7.2 Approche in silico pour l'identification des îlots CpG	14

CHAPITRE 02 :.....	41
Matériel et Méthodes	39
Matériel :.....	17
1 Données biologiques :.....	17
1.1 Choix des gènes :.....	17
2 Outils bio-informatique :.....	18
Méthodes:	21
1 Collecte de données:	21
1.1 Téléchargement des séquences:	21
1.2 Stockage des séquences :.....	22
2 Préparation de données :	22
2.1 Identification de la région promotrice :	22
2.2 Alignement des séquences :	23
2.3 Calcul du CG% et du nombre de motifs CG:.....	24
2.4 Détection des îlots CpG:	24
3 Analyse visuelle de données:.....	26
3.1 Construction d'un arbre phylogénétique:	26
3.2 Construction des Logo:.....	27
3.3 Prédiction des motifs CG:.....	27
CHAPITRE 03 :.....	39
Résultats et Discussion	39
Résultats.....	30
1 Résultats de l'analyse phylogénétique des gènes:.....	30
1.1 Analyse des différents clades identifiés dans l'arbre phylogénétique :	30
2 Résultats d'analyse des îlots CpG:	31
3 Résultats d'analyse des Logo:.....	34
Discussion :.....	37
Conclusion Et Perspectives	39
Liste de références bibliographiques	41
Annexes	44

Introduction Générale

Introduction Générale

« Si je n'ai pas d'enfant, c'est comme si j'étais invitée dans la vie ». Cette douloureuse phrase a été lancée par une patiente au centre FIV au Caire.

Près d'un couple sur dix en âge de procréer est confronté à l'infertilité féminine, un problème de santé publique mondial aux lourdes conséquences. L'Organisation mondiale de la santé (OMS) définissant l'infertilité comme l'incapacité d'obtenir une grossesse après un an de rapports sexuels non protégés (Roy, 2009). Affectant environ 17,5% des couples en âge de procréer dans les pays développés et 1 couple sur 7 consultent pour des problèmes de reproduction (Steiner *et al.*, 2020). En Algérie, ce chiffre est estimé à 12% des couples, soit 7% des couples en âge de procréer (Li *et al.*, 2015). Cela représente des millions de femmes à travers le monde confrontées à l'incapacité de concevoir un enfant, ce qui peut avoir un impact profond sur leur santé mentale et physique, ainsi que sur leur vie de couple.

Comprendre les causes de l'infertilité féminine est crucial pour développer des traitements plus efficaces et soutenir les femmes qui vivent cette épreuve. Il est important de souligner que l'infertilité est une expérience complexe et subjective ; chaque femme la vit différemment et il n'y a pas de réponse unique à ce problème.

L'exploration des mécanismes biologiques impliqués dans l'infertilité féminine, avec un focus particulier sur l'épigénétique, constitue le fil conducteur de ce mémoire. L'épigénétique se réfère aux modifications héréditaires de l'expression des gènes qui ne modifient pas la séquence d'ADN elle-même. Ces modifications peuvent être influencées par des facteurs environnementaux, le mode de vie et l'âge.

L'épigénétique joue un rôle crucial dans la fertilité féminine, influençant le développement folliculaire, la maturation des ovocytes, la fertilisation et l'implantation embryonnaire. Des anomalies épigénétiques ont été associées à divers troubles de la fertilité féminine, tels que le syndrome des ovaires polykystiques (SOPK), l'endométriose, les fausses couches à répétition et la ménopause précoce (Zhu *et al.*, 2018).

De nombreuses études ont identifié des gènes spécifiques impliqués dans l'infertilité féminine, pouvant être directement liés à la fonction reproductive ou jouer un rôle indirect dans des processus tels que la régulation hormonale ou le développement folliculaire. L'épigénétique peut influencer l'expression de ces gènes, modulant leur activité et impactant la fertilité.

Les îlots CpG sont des régions riches en cytosine-guanine (CG) dans l'ADN et des cibles privilégiées pour les modifications épigénétiques, telles que la méthylation, qui peut influencer l'expression des gènes (Bird, 2002). L'analyse des profils de méthylation des îlots CpG peut fournir des informations précieuses sur les mécanismes épigénétiques impliqués dans la fertilité féminine.

La question que nous posons est la suivante : Comment l'environnement, et plus précisément l'épigénétique, influencent-ils la fertilité féminine ? C'est dans ce cadre que nous nous proposons d'apporter notre contribution en explorant la relation entre l'épigénétique et la fertilité féminine avec une approche *in silico*.

L'approche *in silico*, basée sur des simulations bioinformatique, offre un outil puissant pour décrypter les interactions complexes entre l'épigénétique et la fertilité féminine. Cette approche permet d'analyser de vastes ensembles de données génomiques et épigénomiques, d'identifier des gènes clés et de modéliser les processus biologiques impliqués dans la fertilité.

Notre manuscrit est organisé en trois chapitres :

Chapitre 01 : Ce chapitre présente une revue de la littérature existante sur l'infertilité féminine, en se concentrant sur les aspects génétiques et épigénétiques. Il couvrira une définition du problème, identification des causes, décryptage des mécanismes épigénétiques impliqués et découvrirons les approches *in silico* pour les étudier.

Chapitre 02 : Ce chapitre décrit les outils bioinformatiques utilisés pour l'identification des îlots CpG, ainsi que les méthodes d'utilisation de ces outils pour atteindre l'objectif.

Chapitre 03 : Ce chapitre présente les résultats obtenus au cours de l'identification des îlots CpG, les comparer avec d'autres résultats publiés et propose une discussion approfondie de l'interprétation de ces résultats.

CHAPITRE 01 :

Synthèse

Bibliographique

Introduction

L'infertilité fait du projet d'enfant un parcours semé d'embûches. Infertilité, *infertilis*, désigne l'état de ce qui est infertile (Le Littré, 1877), c'est-à-dire de ce qui ne peut pas produire, ni porter. Il existe trois grands types d'infertilité : L'infertilité féminine, l'infertilité masculine et l'infertilité mixte. 30% des cas d'infertilité est d'origine féminine (Inserm, 2019).

1 Infertilité féminine :

Une femme est considérée comme infertile lorsque qu'elle ne parvient pas à tomber enceinte au bout de 2 ans, malgré des rapports sexuels réguliers et en absence de toute contraception. On distingue deux types d'infertilité féminine : l'infertilité dite absolue (primordiale) lorsque la femme n'a jamais été enceinte et l'infertilité dite relative (secondaire et accessoire), lorsque la femme a déjà mené une grossesse à terme ou connu une ou plusieurs fausses couches. D'après l'OMS, 40 % de couples sont atteints par l'infertilité primordiale et 60 % par l'infertilité secondaire (Evers, 2002).

2 Les causes de l'infertilité féminine :

Pour comprendre l'ampleur de ce phénomène, il est nécessaire de faire un focus sur les divers facteurs qui le causent.

2.1 Causes physiologiques :

Chez les femmes, le stock d'ovocytes est déterminé avant la naissance et diminue progressivement jusqu'à la ménopause. Par conséquent, avec l'âge dès 30 ans le taux de fécondité féminin chute (Ameli, 2023). Ainsi la fécondabilité par cycle est estimée à 25 % vers 20-30 ans, mais à seulement 12 % à 35 ans et à 6 % à 40 ans (Hamam.S, Beriloux.S, 2020). L'infertilité est très souvent liée à des causes médicales. Elle peut avoir une origine mécanique, elle peut résulter de deux types de pathologies : les pathologies tubaires, et les pathologies utérines. Elle peut aussi être d'origine hormonale (Hamam.S, Beriloux.S, 2020).

2.2 Causes génétiques :

Ainsi les causes génétiques incluent les problèmes d'insuffisance ovarienne. En effet, une insuffisance ovarienne prématurée (IOP) non syndromique ainsi que les anomalies chromosomiques peuvent être causées par des mutations génétiques, telles que celles affectant

le gène GDF9. De plus, le syndrome de Turner (45, X), 47, XXX, la délétion du chromosome X et les translocations X-autosomiques, conjugués aux gènes autosomiques et à l'IOP, représentent 10 à 13% des cas d'IOP. Par ailleurs, il existe des risques génétiques pour l'endométriose et le syndrome des ovaires polykystiques (SOPK), deux maladies qui peuvent affecter la fertilité. Cependant, ces facteurs génétiques sont liés aux origines génétiques de l'infertilité féminine.

2.3 Causes épigénétiques :

L'influence de facteurs environnementaux sur la fertilité des individus a également été mise en évidence c'est ce qu'on appelle "l'exposome" qui désigne l'ensemble des expositions à des facteurs environnementaux (c'est-à dire non génétiques) que subit un organisme de sa conception à la fin de vie en passant par le développement in utero, complétant l'effet du génome (Hamam.S, Beriloux.S, 2020). En plus d'être considéré comme un agent physiologique, l'âge peut également être classé avec les facteurs épigénétiques affectant l'infertilité féminine. Le poids, les troubles alimentaires, le stress, le manque de sommeil, l'exposition aux toxines et aussi les troubles liés à l'usage de substances psychoactives (tabac, cannabis, alcool, drogues mais aussi le café) sont les principaux facteurs environnementaux pourraient avoir un impact sur les gamètes en termes quantitatif et qualitatif.

3 Epigénétique et contrôle de l'expression des gènes :

L'épigénétique, un domaine en plein essor, offre une perspective nouvelle sur l'impact de l'environnement sur la fertilité féminine. Le terme épigénétique désigne les processus moléculaires permettant de moduler l'expression des gènes, mais qui ne sont pas fondés sur des changements dans la séquence de l'ADN (Baird, 2007, Bourc'his, 2010, Gabory, 2011). De plus l'épigénétique peut être causée par des facteurs environnementaux, elles peuvent également être héritée de génération en génération. Notre mode de vie pourrait influencer non seulement l'expression de nos propres gènes, mais aussi celle de nos enfants et nos petits-enfants (E. Gnansia, 2017)

3.1 Mécanismes épigénétiques

La régulation épigénétique implique une multitude de mécanismes qui modulent l'expression des gènes sans altérer la séquence d'ADN elle-même

3.1.1 Méthylation de l'ADN :

Il s'agit-là du premier mécanisme épigénétique qui impacte directement la structure de la chromatine. L'expression « méthylation de l'ADN » signifie, en effet, que des méthyltransférases ajoutent, sur des cytosines, des groupements méthyl qui deviennent alors 5-méthylcytosines. Ces derniers sont adjacents à un résidu Guanine formant un dinucléotide cytosine-phosphate-guanine (îlots CpG) (M. Aber, 2008). Le catalyse de la réaction est réalisée par une famille d'enzymes dites DNA méthyltransférases (DNMTs) via le substrat S-adénosyl-méthionine (SAM) qui donne le groupement méthyl. La méthylation de l'ADN permet de réguler l'expression des gènes en fonction de leur niveau de méthylation. Une hyperméthylation de leurs promoteurs, une forte méthylation conduit à l'inactivation des gènes car ceux-ci vont inhiber l'expression génique ; inversement, une hypométhylation, une faible méthylation favorise l'activation génique car elle induit une forte expression des gènes.

3.1.2 Modification des histones :

L'unité fondamentale de compactage de l'ADN est le nucléosome. Le nucléosome est formé de deux tours d'ADN à double hélice enroulés autour d'un octomère constitué lui-même de deux copies des histones H2A, H2B, H3 et H4 (Kowalski and Pałyga, 2012). Les modifications post-synthèse traductionnelle des histones font partie des modifications épigénétiques. Celles-ci comprennent notamment : l'acétylation, la méthylation, la phosphorylation et l'ubiquitylation).

▪ L'acétylation :

Il s'agit de l'ajout d'un groupement acétyle sur les résidus lysine, qui neutralise leurs charge positive. L'acétylation est réalisée par des histones acétyltransférases (HAT), elle joue un rôle important dans l'accessibilité de l'ADN aux facteurs de transcriptions (Smith and Meissner, 2013 ; Woodcock, 2006). La désacétylation par les histones Désacétyltransférases (HDAC) : elle désacétyle l'histone.

▪ La méthylation :

La méthylation des histones est plus complexe, elle peut se produire sur les lysines ou sur les arginines. Effectivement, il est possible que les lysines soient mono-, di- ou tri-méthylées, tandis que les arginines peuvent être mono- ou di-méthylées. Les protéines à domaine SET

ont la capacité de méthyler les arginines, tandis que les PRMT (protein arginine méthyltransférases) ont la capacité de méthyler les lysines (Lacoste, N. & Côté, J. 2003).

- **La phosphorylation :**

La phosphorylation des histones se produit lorsque les enzymes kinases ajoutent des groupements phosphate aux résidus de sérine, de thréonine ou de tyrosine des histones. Elle exerce des effets stimulants ou inhibiteurs sur la transcription génétique. Elle participe également à la restauration de l'ADN après des cassures doubles brins. Les phosphatases ont la capacité d'inverser la phosphorylation.

- **L'ubiquitylation :**

Il s'agit d'incorporer une ou plusieurs copies d'ubiquitine sur une protéine ciblée. Un groupe d'enzymes, une enzyme d'activation E1, une enzyme de conjugaison E2 et une ligase E3, réalisent ce processus. L'ubiquitylation joue un rôle crucial dans plusieurs processus cellulaires comme : la régulation de la dégradation des protéines, la signalisation cellulaire et le trafic intracellulaires.

3.1.3 Les ARN non codants :

Les ARN non codants (ARNnc) sont des molécules d'acide ribonucléique (ARN) issues de la transcription de l'ADN, mais qui ne sont pas traduites en protéines. Ils interagissent avec d'autres types d'ARN, tels que les ARN messagers (ARNm), les ARN de transfert (ARNt) et les ARN ribosomiques (ARNr), pour remplir leurs fonctions (Brown, C. J. D. et *al.*, 2016). Les ARNnc se classent en deux catégories principales en fonction de leur taille et de leur fonction : les microARN (miARN) et les longs ARN non codants (lncARN) (Yao, F. et *al.*, 2019). Loin d'être de simples transcrits passifs, les ARN non codants (ARNnc) s'affirment comme des acteurs majeurs de la régulation épigénétique. En effet, ils peuvent moduler l'expression des gènes à la fois au niveau transcriptionnel et post-transcriptionnel (Weinhold, B. et *al.*, 2009)

4 La génétique de l'infertilité :

La génétique de l'infertilité est déterminée par un ensemble de gènes qui influencent la reproduction féminine, favorisant ainsi une reproduction optimale. Des recherches incessantes révèlent que plus de 200 gènes sont impliqués dans la reproduction féminine (Yatsenko.SA et Rajkovic. A, 2019).

4.1 Les catégories fonctionnelles des gènes impliqués dans la reproduction féminine :

- **Gènes impliqués dans la gamétogenèse** responsable de production d'ovules ;
- **Gènes impliqués dans la fonction hormonale** qui régulent des hormones sexuelles telles que les hormones FSH, LH et œstrogènes ;
- **Gènes impliqués dans la structure et la fonction de l'appareil reproducteur féminin**, ces gènes jouent un rôle dans le développement et fonctionnement de l'utérus, des trompes de Fallope et des ovaires.

Des mutations dans ces gènes peuvent entraîner une production d'ovules de qualité inférieure, perturber le cycle menstruel et altérer la capacité des organes reproductifs, affectant la fertilisation. Il est important de souligner que l'infertilité est un problème complexe et multifactoriel, et que les gènes ne sont qu'un des éléments.

4.2 L'épigénétique : Un facteur clé de la fertilité au-delà du contrôle génétique :

Des recherches récentes révèlent un nouveau chapitre dans l'histoire de l'infertilité : l'épigénétique. Plus de 1000 gènes peuvent être affectés par des modifications épigénétiques et jouer un rôle dans l'infertilité féminine (Institut Pasteur, 2003). Ces gènes ne causent pas nécessairement l'infertilité et ils ne sont pas soumis à l'épigénétique de la même manière : certains gènes sont plus sensibles aux modifications épigénétiques que d'autres (Defert, O. et *al.*, 2011).

5 Infertilité et épigénétique :

Les différents mécanismes épigénétiques tels que la méthylation de l'ADN, les modifications post traductionnelles des histones et les ARN non codants, jouent un rôle important dans le contrôle des gènes qui interviennent dans la fertilité féminine. Ces modifications épigénétiques peuvent agir indépendamment ou en coordination sur la compaction de la chromatine et donc sur le niveau de transcription des gènes impliqués dans la folliculogenèse, la maturation des ovocytes et l'implantation (CHEN Q et *al.*, 2016). Ces dysfonctionnements sont délétères car ils perturbent les régulations physiologiques de la fertilité qui peuvent à l'origine du développement de certaines maladies d'infertilité. Si les mécanismes épigénétiques sont mieux définis dans la connaissance des pathologies d'infertilité féminine, ils peuvent représenter de nouvelles cibles thérapeutiques et diagnostiques.

5.1 Influence de l'épigénétique sur l'infertilité féminine :

La santé reproductive féminine peut être influencée par les mécanismes épigénétiques, sensibles aux modifications de l'environnement et du mode de vie.

5.1.1 Facteurs environnementaux et mode de vie :

De nombreux facteurs environnementaux, comme la nutrition, le style de vie et l'exposition à des polluants, peuvent influencer la capacité à procréer. Parmi ces facteurs, le surpoids et l'obésité, en augmentation croissante chez les femmes en âge de procréer, constituent un facteur non négligeable de baisse de la fertilité naturelle (Robker et al, 2009).

- **Les substances psychoactives :**

Le tabagisme affecte les gamètes femelles et induit une diminution de fertilité en perturbant la synthèse des hormones stéroïdiennes ainsi qu'en diminuant la qualité des ovocytes. Les risques de fausses couches et de grossesses extra-utérines augmentent selon la dose de tabac consommée. De même, la consommation d'alcool peut avoir un impact négatif sur la fertilité. Cependant, il est bien établi qu'une consommation d'alcool, même modérée, durant la grossesse, a un impact délétère sur le développement fœtal. Des récepteurs spécifiques aux cannabis ont été identifiés dans les organes reproducteurs et les glandes endocrines des femmes, ce qui confirme l'impact négatif de la consommation du cannabis sur la reproduction.

- **Le poids :**

Une nouvelle étude menée par des chercheurs met en évidence une association entre le poids et l'infertilité, notamment entre l'indice de masse corporelle (IMC) et les problèmes de fertilité. En ce qui concerne l'obésité, les femmes obèses normo-ovulantes (ayant un IMC supérieur à 30) sont plus susceptibles d'avoir des risques accrus pour les fonctions reproductives que les femmes ayant un poids santé. Le surpoids peut altérer la qualité ovocytaire et endométriale, et augmenter le risque de fausses couches lorsqu'il s'accompagne de résistance à l'insuline (Pesant et al, 2010). En ce qui concerne la maigreur, les femmes ayant un IMC inférieur à 18 peuvent également avoir des problèmes de fertilité. Le sous-poids peut provoquer un arrêt des cycles d'ovulation ce qui peut entraîner difficultés à tomber enceinte.

- **Les autres troubles alimentaires :**

Des chercheurs ont montré que la nutrition est capable d'influencer la disposition des marques épigénétiques des cellules sexuelles. La malnutrition perturbe le métabolisme ce qui

affecte le bon fonctionnement de l'appareil reproducteur : lors de la production des ovules, les gènes du métabolisme établissent de nouvelles marques épigénétiques (Pembrey ME, Bygren LO, Kaati G, et al. 2006; E.Gnansia, 2017). Chez la femme, une consommation élevée d'aliments gras et une faible consommation de fruits est associée à des troubles de l'ovulation (Chavarro et al, 2007 ; Grieger et al, 2018). En revanche, le régime méditerranéen a souvent été associé à de meilleures chances de grossesse (Hamamah & Berlioux, 2022). De même, une association positive a été observée entre la consommation de produits laitiers et les chances de naissance vivante chez les femmes de plus de 35 ans. Ainsi, les femmes qui suivent un régime "profertilité" avaient plus de chances de donner naissance à un enfant vivant après avoir eu recours à des techniques d'AMP, contrairement aux femmes dont l'alimentation était pauvre en fruits et riche en "Fast Food" et en boissons sucrées (Gaskins et al ; Grieger et al. 2018)

- **Le stress :**

Le stress psychologique est considéré comme un facteur environnemental pouvant influencer négativement la fertilité féminine. Il peut agir de manière indirecte en favorisant des troubles comme les perturbations du sommeil. En outre, un sommeil insuffisant ou de mauvaise qualité peut affecter l'ovulation chez la femme (Goldstein & Smith, 2016) cela peut également affecter les hormones reproductives telles que l'œstrogène et la progestérone, essentielles à la fertilité. De plus, le stress peut entraîner des troubles du cycle menstruel rendant plus long ou plus court (Harlow & Matanoski, 1991; Fenster et al. 1999). De même, le stress maternel pendant la grossesse pourrait induire des modifications négatives transmissibles aux enfants et petits-enfants. Il est important de noter que le stress provoque l'infertilité féminine. En effet, des études selon le collège américain des obstétriciens et gynécologues (ACOG), montrent que les patientes qui subissent une fausse couche répondent aux critères du trouble de stress post-traumatique.

- **Exposition aux toxines :**

Selon la définition de l'OMS (2002), « Un perturbateur endocrinien (PE) désigne une substance ou un mélange qui altère les fonctions du système endocrinien et de ce fait induit des effets néfastes dans un organisme intact, chez sa progéniture ou au sein de (sous)-populations. » L'exposition à des polluants, à des PE qui sont souvent présents dans des produits d'hygiène et de beauté (p.ex., le parabène et les phtalates,...) et à d'autres substances toxiques dans l'environnement peuvent favoriser une puberté précoce, qui a fréquemment une relation avec des anomalies de l'ovaire et l'endométriose, le cancer du sein et d'autres

anomalies morphologiques de l'ovaire ou des fausses couches à répétition. Les PE affectent donc la fertilité des femmes adultes, mais également la fertilité de leurs enfants. Ils sont connus pour entraîner des modifications épigénétiques du génome (Hamamah & Berlioux, 2022). La pollution atmosphérique ou la mauvaise qualité de l'air est un autre facteur reprotoxique, associée aux deux types d'infertilité féminine (primaire et secondaire). En effet, plus une femme est exposée à ces facteurs environnementaux, plus elle aura tendance à diminuer sa fréquence de rapports sexuels (Bodenmann et *al*, 2010).

5.2 Empreinte parentale : un aspect de l'épigénétique

L'empreinte parentale est un mode de contrôle de l'expression des gènes, jouant un rôle essentiel dans le développement des mammifères. Certains gènes, s'expriment de manière monoparentale (par l'un des deux allèles parentaux) au cours du développement embryonnaire. A ce jour, des centaines de gènes ont été identifiés comme étant soumis à l'empreinte parentale. Des anomalies de l'empreinte parentale sont associées à un certain nombre de troubles, dont l'infertilité féminine. Ces anomalies peuvent affecter divers aspects de la reproduction, tels que le dysfonctionnement ovarien, la fertilisation, l'implantation embryonnaire et la croissance fœtale (E. Jullien, 2008).

5.3 L'apport de l'étude de l'épigénétique :

Au contraire de la génétique, l'épigénétique offre la possibilité de modifier les facteurs de style de vie, l'alimentation, l'activité physique et d'autres conduites de façon à influencer potentiellement l'expression génique impliquée dans l'infertilité féminine. Cette capacité à agir sur l'expression génique par des interventions environnementales ou comportementales ouvre des perspectives intéressantes pour la prise en charge de l'infertilité féminine et souligne l'importance de comprendre en profondeur les mécanismes épigénétiques dans ce contexte.

6 Les îlots CpG :

Les îlots CpG (5'-Cytosine-phosphate-Guanine-3') sont des régions d'ADN riches en cytosine et en guanine (CG), principalement localisées dans les promoteurs des gènes (Bird, 2002) et occupant environ 1 % de la surface totale du génome humain (Antequera et *al.*, 2004). Les îlots CpG sont composés de dinucléotide "CG" dans le sens 5' vers 3' et de "GC" dans le sens inverse (Figure 1).

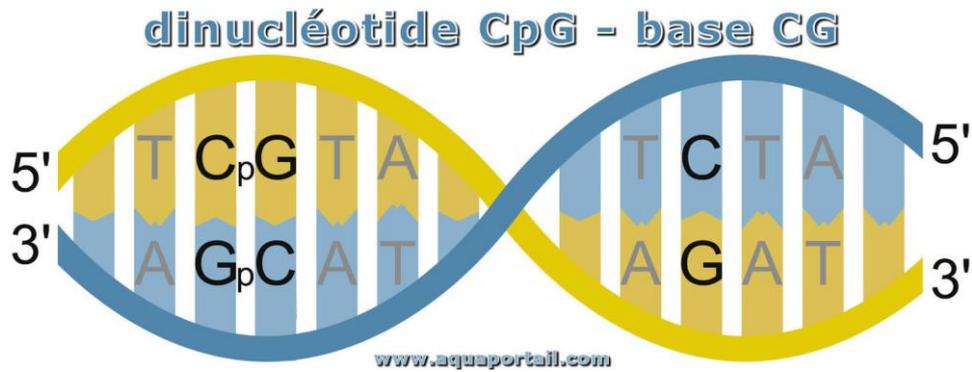


Figure 1 : Séquence d'un dinucléotide CpG d'un brin d'ADN comparé à une paire de bases C-G sur des brins complémentaires (F. Fortier, 2021)

6.1 Rôle dans la régulation de l'expression :

Les promoteurs sont des séquences d'ADN situées en amont des gènes et contrôlent leur transcription. La localisation des îlots CpG dans cette région des promoteurs leur confère un rôle important dans la régulation de l'expression des gènes. Dans les cellules saines, les cytosines des îlots CpG ne sont généralement pas méthylées. Cela permet aux facteurs de transcription de se lier à l'ADN et d'initier la transcription (Alberts et *al.*, 2002). En revanche, dans les cellules cancéreuses ou celles présentant des anomalies de méthylation, les CpG peuvent être hyperméthylés. Cette hyperméthylation empêche la liaison des facteurs de transcription et réprime l'expression des gènes (Bird, 2002).

6.2 Caractéristiques structurales :

Les îlots CpG présentent plusieurs caractéristiques structurales distinctives. En effet, ils varient en longueur allant de 200 pb (Bird, 2002). Ils présentent une teneur en CG (cytosine-guanine) nettement plus élevée que le reste du génome, généralement supérieure à 50 % (Antequera et *al.*, 2004). De plus, les îlots CpG sont souvent enrichis en motifs CG et GC, qui sont des séquences répétitives de ces dinucléotides.

6.3 Lien avec l'infertilité féminine :

La méthylation des îlots CpG des gènes impliqués dans la reproduction féminine peut perturber la fertilité en affectant divers processus clés : implantation embryonnaire, maturation des ovules, et le développement placentaire, et en diminuant leurs capacités de se féconder (Haffner et *al.*, 2015). L'âge maternel avancé constitue un facteur de risque important pour l'hyperméthylation des CpG dans les gènes impliqués dans la reproduction, ce qui contribue à la diminution de la fertilité avec l'âge (Bayraktar et *al.*, 2012).

De plus, les facteurs environnementaux, tels que l'exposition aux toxines comme les phtalates et le bisphénol A, peuvent également induire des altérations de la méthylation et la déméthylation des îlots CpG dans les tissus reproductifs de la femme, augmentant le risque d'infertilité et d'autres complications reproductives (Haffner et *al.*, 2015).

7 Approche *in silico* :

Le terme "*in silico*" dérive du latin "*in silice*", signifiant "en silice" ou "en pierre", et a été inventé en 1989 (P. Abélard, 1989). Il ne fait pas référence au silicium en tant que matériau, mais plutôt à l'idée de simuler un processus à l'aide d'un ordinateur.

7.1 Principe

La recherche "*in silico*", également connue sous le nom de biologie computationnelle, utilise des modèles biomathématiques et des bases de données provenant d'expériences *in vivo* ou *in vitro* pour analyser et simuler des systèmes biologiques complexes à l'aide de logiciels informatiques (Gallezot, 2002).

7.2 Approche *in silico* pour l'identification des îlots CpG

L'une des applications majeures des méthodes "*in silico*" est la recherche de motifs spécifiques dans les gènes, comme les îlots CpG dans les gènes impliqués dans l'infertilité féminine. Cette recherche s'appuie sur des bases de données et commence par la collecte de données précises, en particulier les régions d'intérêts (Bernstein et *al.*, 2002). La définition de critères précis pour les motifs à rechercher est une étape cruciale pour une recherche plus affinée (Marín-Esteban et *al.*, 2009). Ces critères peuvent être calculés ou collectés grâce à l'utilisation d'algorithmes (Pires et *al.*, 2014). L'annotation bioinformatique permet de préciser la position des motifs, leurs proximités avec d'autres motifs ou leurs associations (Croset et *al.*, 2016). Des bases de données spécifiques existent pour la détection des îlots CpG et leurs caractéristiques, avec une représentation graphique des résultats (Wei et *al.*, 2018).

En résumé, nous pouvons dire que l'infertilité féminine, phénomène complexe et multifactoriel, est influencée par des dysfonctionnements épigénétiques, notamment l'hyperméthylation des îlots CpG, altérant l'expression des gènes impliqués dans les processus reproductifs. L'exploration des interactions entre l'épigénétique, l'environnement et la fertilité féminine, en s'appuyant sur les mécanismes épigénétiques, ouvre de nouvelles perspectives pour le diagnostic et la prise en charge de cette pathologie.

CHAPITRE 02 :

Matériel et Méthodes

Matériel :

Cette section décrit en détail les différents éléments utilisés pour mener à bien les analyses et obtenir les résultats présentés dans notre travail.

1 Données biologiques :

Nous avons utilisé un ensemble de séquences nucléiques au format FASTA de 23 gènes impliqués dans l'infertilité féminine. Ces séquences correspondent aux variants Homo sapiens de chaque gène.

1.1 Choix des gènes :

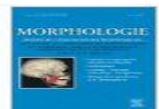
La sélection des gènes a été guidée par l'objectif de la recherche, en privilégiant ceux ayant un lien direct et avéré avec l'infertilité féminine et soumis à la régulation épigénétique. Le choix des gènes soumis à l'épigénétique a été motivé par le fait que les îlots CpG, que nous visons à identifier, sont particulièrement sensibles à la méthylation de l'ADN, qui constitue le principal mécanisme épigénétique (M. Waber, 2008).

Pour réaliser cette recherche :

Nous avons consulté des revues scientifiques et des publications spécialisées afin d'identifier des gènes potentiels liés à l'infertilité féminine et à la régulation épigénétique (figure 2).



Morphologie
Volume 106, Issue 354, Supplement, September 2022, Pages S19-
S20



PF-12-CHEC

Validation fonctionnelle d'un variant homozygote du gène de l'AMH chez une patiente avec insuffisance ovarienne débutante

Figure 2 Exemple de la recherche de gènes dans la littérature scientifique (S.Jaillard, 2022)

Ensuite, nous avons exploité des bases de données génétiques réputées telles que PubMed (figure 3) et GenBank pour obtenir des descriptions détaillées et des résumés fonctionnels de

ces gènes. Cette approche méthodique nous a permis d'évaluer l'intérêt de chaque gène candidat et de sélectionner ceux les plus pertinents pour notre étude.

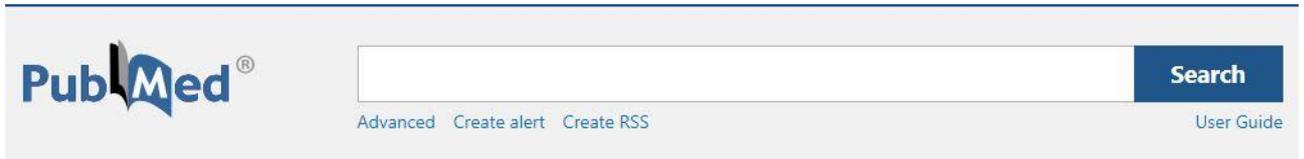


Figure 3 Interface de recherche PubMed

2 Outils bio-informatique :

Les outils bioinformatiques décrits dans le tableau 2 ont été exécutés sur un ordinateur dont les caractéristiques sont présentées dans le tableau 1.

Tableau 1 Caractéristiques du matériel informatique utilisé

Ordinateur	Caractéristiques
Processeur	Intel(R) Core (TM) i3-4005U CPU @ 1.70GHz 1.70 GHz
Mémoire installée RAM	4,00 Go
Système d'exploitation	Windows 10 Famille
Type de système	Système d'exploitation 64 bits, processeur x64
Version du système	22H2

Tableau 2 Outils bioinformatique utilisés et leurs applications

Outil	Utilité
GenBank	Téléchargement des séquences
Sublime Texte	Stockage de données
dbCAT	Identification des ilots CpG et leurs structures
MEGA 11	Alignement multiple des séquences et construction de l'arbre phylogénétique
WebLogo	Visualisation de l'alignement des séquences
Dotlet	Prédiction des régions répétées à l'aide Dotblot
Python	Analyse du contenu en CG et identifier les motifs CG.

Voici des descriptions détaillées de chaque logiciel employé :

NCBI

National Center for Biotechnology Information est un institut national pour l'information biologique moléculaire, fondé aux Etats-Unis en 1988. Il conduit des recherches Bioinformatiques pour l'analyse des données du génome, et abrite plusieurs bases de données, comme GenBank, PubMed, etc...(J.Helden, 2021)

GenBank

Une base de données de séquences d'ADN et ARN créée par NCBI, contenant toutes les séquences nucléiques et protéiques de plus de 100000 organismes différents (J.Helden, 2021).

FASTA

Un format standard pour le stockage des données nucléiques et protéiques. Les séquences sont identifiées par le caractère ">" et un en-tête spécifique (Jsobel, 2014).

Sublime Text

Sublime Text est un éditeur de texte et de code source reconnu pour sa rapidité, sa flexibilité et sa large gamme de fonctionnalités. Grâce à sa lisibilité et sa facilité d'utilisation, il peut également être utilisé comme un outil de stockage de données FASTA (L.Brian, 2023).

Python

Un langage de programmation fonctionnel multiplateforme, puissant, polyvalent et accessible, est largement utilisé dans divers domaines, y compris l'apprentissage automatique. Ce qui en fait un choix populaire pour les développeurs de tous niveaux.

DBCAT

DBCAT (Database of CpG Islands and Analytical Tools) est une ressource bioinformatique complète dédiée à l'analyse des îlots CpG et de la méthylation de l'ADN, permettant ainsi d'explorer la régulation épigénétique dans le contexte des maladies (J. Comput Biol, 2011).

MEGA 11

Un logiciel puissant pour l'analyse évolutive moléculaire. Offrant plusieurs fonctionnalités essentielles telles que : l'alignement des séquences, la construction des arbres phylogénétiques, etc...(K.Tamura et *al.*, 2016).

ClustalW

Un algorithme d'alignement multiple utilisé pour comparer les séquences biologiques afin d'identifier leurs similarités et leurs niveaux de conservation (L.Bruna et *al.*, 2017).

WebLogo

WebLogo est une application Web conçue pour rendre la génération de logos de séquence aussi simple et conviviale que possible (E.Escrocs et *al.*, 2004).

Dotlet JS

Un logiciel bioinformatique dédié à l'analyse de séquences nucléiques et protéiques. Il utilise une méthode appelée diagramme diagonal pour visualiser les similitudes et les différences entre les séquences, permettant ainsi l'identification de mutations et la détection de motifs répétitifs (JM.Claverie, 2007).

Méthodes :

Cette section expose de manière méthodique les étapes mises en œuvre pour identifier les îlots CpG dans les régions promotrices de nos gènes, qui se décomposent en trois étapes principales.

1 Collecte de données :

La présente section présente la méthodologie employée pour collecter les données.

1.1 Téléchargement des séquences :

Nous avons téléchargé les séquences nucléiques des gènes impliqués dans l'infertilité féminine à partir du portail Nuclotide (GenBank) dans NCBI (figure 4).

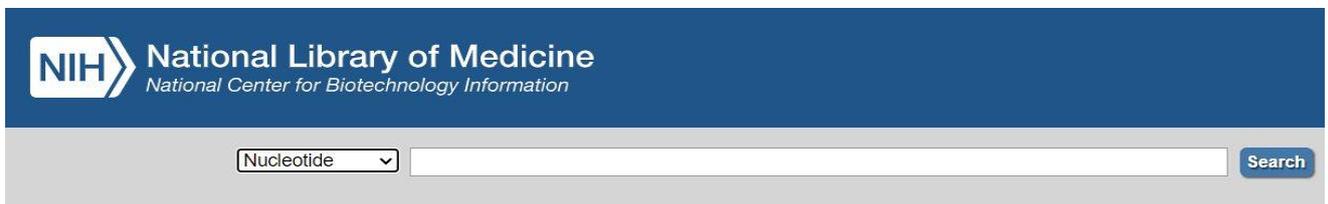


Figure 4 Portail Nucleotide NCBI

Nous avons mené des analyses sur des gènes Homo sapiens et sélectionné quelques

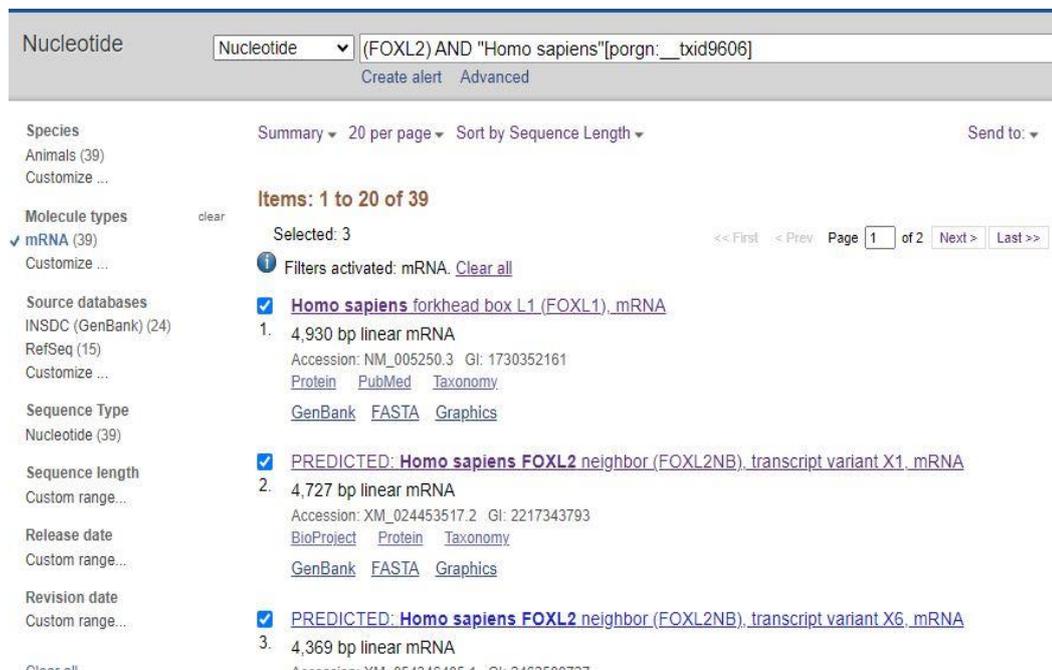


Figure 5 Identification et sélection de variants du gène FOXL2 chez Homo sapiens variants pour chaque gène (figure 5).

Ensuite, nous avons téléchargé ces variants au format FASTA.

1.2 Stockage des séquences :

Les séquences FASTA sont stockées à l'aide du logiciel Sublime Text (figure 6).



Figure 6 Représentation d'une séquence dans Sublime Text

2 Préparation de données :

La préparation est constituée d'un ensemble d'étapes permettant de nettoyer les séquences en ne conservant que la région d'intérêt et d'effectuer les opérations requises pour identifier les îlots CpG dans cette région.

2.1 Identification de la région promotrice :

On retrouve fréquemment les îlots CpG dans les régions promotrices des gènes. C'est pourquoi nous avons extrait ces régions. La localisation de la séquence codante CDS, accessible sur GenBank, peut servir d'indice pour identifier le promoteur, qui se situe généralement en amont de la CDS (figure 7).

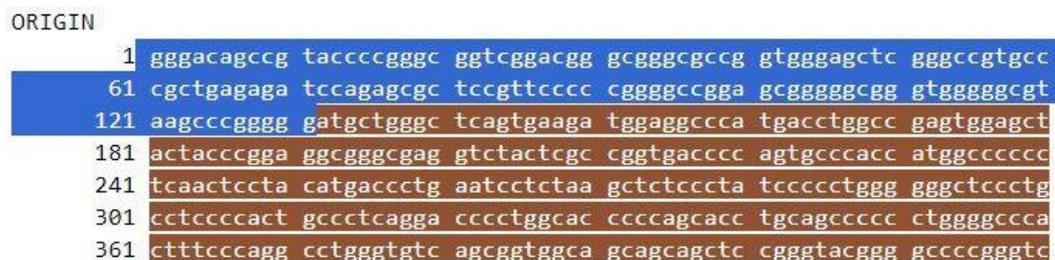


Figure 7 Localisation de la région promotrice en amont de la CDS du gène FOXL2

La partie en marron correspond à la CDS, ce qui indique que la partie en bleu représente la région promotrice.

2.2 Alignement des séquences :

Après avoir conservé la région promotrice en supprimant la partie de notre séquence allant de la CDS à la fin, nous avons réalisé un alignement multiple des séquences restantes à l'aide du logiciel MEGA X version 11 et de l'algorithme d'alignement multiple ClustalW (représenté par numéro 1 dans la figure 8).

L'alignement est divisé en deux phases :

- **Première phase :**

Alignement des séquences promotrices des variants d'un même gène.



Figure 8 Alignement multiple des régions promotrices du gène FOXL2 à l'aide de ClustalW

- **Deuxième phase :**

Nous avons ensuite aligné les séquences des régions promotrices de tous les gènes, afin d'identifier les régions conservées et les variations entre les gènes (figure 9).

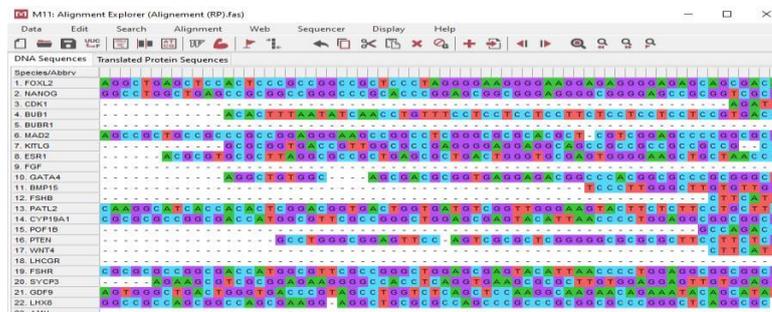


Figure 9 Alignement multiple des régions promotrices de tous les gènes.

Nous avons exporté chaque alignement au format FASTA et au format MEGA pour des raisons d'utilisation ultérieure.

2.3 Calcul du CG% et du nombre de motifs CG:

En utilisant un script Python, nous avons calculé le nombre de motifs "CG" dans les régions promotrices, ainsi que le pourcentage de CG pour chacune d'entre elles (figure 10).

Le calcul des motifs CG et du CG% des promoteurs constitue une étape essentielle dans la recherche des îlots CpG et la compréhension de la régulation de l'expression des gènes.

```
def count_gc_cg(sequence):
    sequence = sequence.upper()
    count_gc = 0
    count_cg = 0

    for i in range(len(sequence) - 1):
        if sequence[i:i+2] == 'GC':
            count_gc += 1
        elif sequence[i:i+2] == 'CG':
            count_cg += 1

    return count_gc, count_cg

|
sequence = input("Entrez votre séquence : \n").upper()
G= sequence.count("G")
C= sequence.count("C")
L = len(sequence)
PGC = ((G+C)/L)*100
gc_count, cg_count = count_gc_cg(sequence)
print("Nombre de motifs GC :", gc_count)
print("Nombre de motifs CG :", cg_count)
print("Nombre total = :", cg_count+gc_count)
print("le % GC = :", PGC,"%")
```

Figure 10 Script Python pour l'analyse du contenu en CG et des motifs CG

2.4 Détection des îlots CpG :

En utilisant la base de données dbCAT (figure 11), nous avons identifié les îlots CpG et leur structure au sein de la région promotrice de chaque gène. L'analyse des motifs CG et du CG% des promoteurs est une étape cruciale dans l'identification des îlots CpG et la compréhension de la régulation de l'expression des gènes. En effet, les îlots CpG se caractérisent par un pourcentage CG supérieur à 50 % et sont généralement riches en motifs CG (GC dans le brin anti sens).

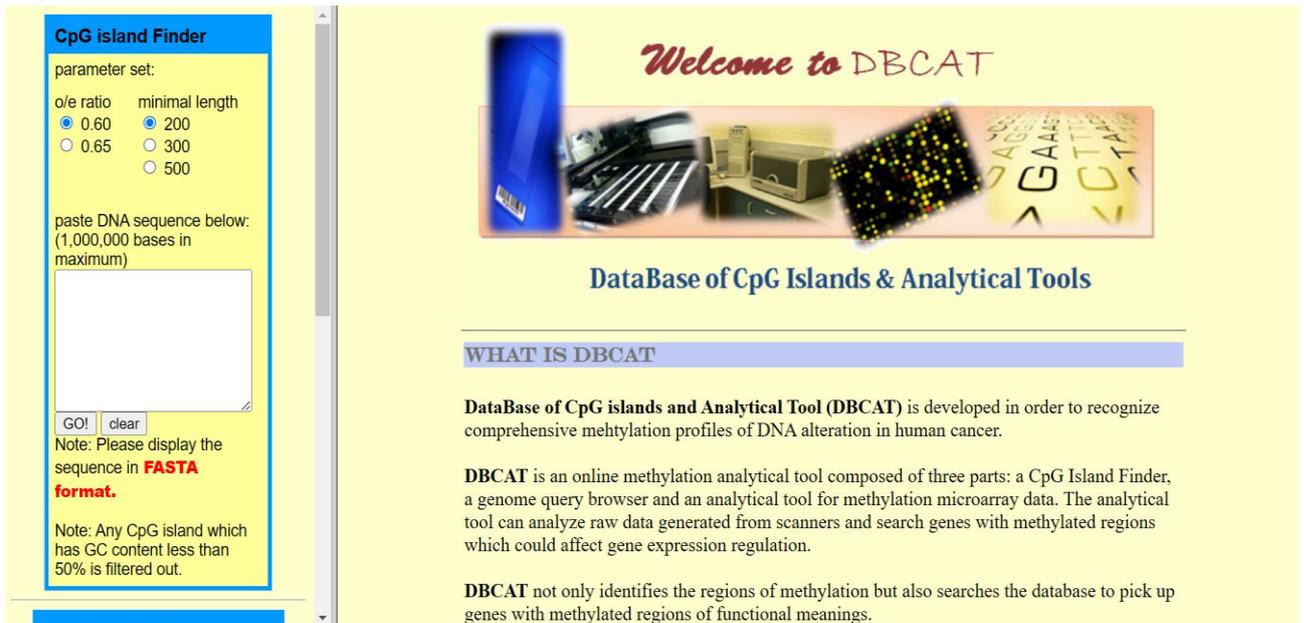


Figure 11 La base de données DBCAT

Cette base de données nous a également permis de détecter les séquences contenant des îlots CpG, en précisant leur taille et leur motif spécifique. La structure des îlots CpG varie en fonction de leur taille et de leur contenu en C G (figure 12).

Possible CpG islands shown in graph:

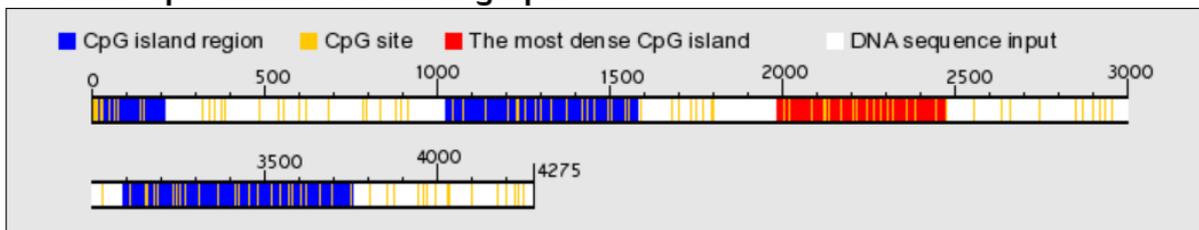


Figure 12 Caractérisation structurale des îlots CpG par DBCAT

Les régions en jaune représentent les sites CpG ou les CpG isolés (longueur inférieure à 200 pb). Les régions en bleu représentent les CpG intermédiaires ou de taille modérée (supérieure à 200 pb). Les régions en rouge représentent les CpG les plus condensés (en clusters) de grande taille.

Afin de synthétiser les résultats principaux issus des deux étapes méthodologiques précédentes, un tableau récapitulatif a été élaboré. Ce tableau présente la taille des régions promotrices, le pourcentage de CG (%CG), le nombre de motifs CG et GC, ainsi que la structure et la taille des îlots CpG identifiés à l'aide de la base de données DBCAT (voir tableau 3).

3 Analyse visuelle de données :

L'objectif de la visualisation complète est d'améliorer la compréhension des relations entre les gènes, de détecter les motifs conservés et les îlots CpG par des moyens visuels.

3.1 Construction d'un arbre phylogénétique :

Nous avons construit un arbre phylogénétique à partir du fichier d'AMS exporté au format "MEGA". L'arbre phylogénétique est un outil essentiel pour la visualisation de relations évolutives des espèces ou des séquences sous forme de schéma graphique.

A partir de l'option "Models" de MEGA 11, nous avons sélectionné le meilleur modèle parmi les 24 modèles de substitutions nucléotidiques par maximum de vraisemblance disponibles, celui qui correspond le mieux à nos données (figure 13).

Model	Parameters	BIC	AICc	lnL	(+I)	(+G)	R	r(A)	r(T)	r(C)	r(G)	r(AT)	r(AC)	r(AG)	r(TA)	r(TC)	r(TG)	r(CA)	r(CT)	r(CG)	r(GA)	r(GT)	r
T92	45	7411,753	7141,332	-3524,988	n/a	n/a	0,80	0,150	0,150	0,350	0,350	0,039	0,090	0,170	0,039	0,170	0,090	0,039	0,073	0,090	0,073	0,039	r
T92+G	46	7419,452	7143,051	-3524,818	n/a	31,66	0,81	0,150	0,150	0,350	0,350	0,038	0,089	0,172	0,038	0,172	0,089	0,038	0,074	0,089	0,074	0,038	r
T92+I	46	7419,462	7143,061	-3524,823	0,00	n/a	0,80	0,150	0,150	0,350	0,350	0,039	0,090	0,170	0,039	0,170	0,090	0,039	0,073	0,090	0,073	0,039	r
HKY	47	7426,360	7143,982	-3524,252	n/a	n/a	0,80	0,149	0,151	0,335	0,364	0,039	0,086	0,178	0,038	0,164	0,093	0,038	0,074	0,093	0,073	0,039	r
T92+G+I	47	7427,276	7144,898	-3524,710	0,00	35,07	0,81	0,150	0,150	0,350	0,350	0,038	0,089	0,172	0,038	0,172	0,089	0,038	0,074	0,089	0,074	0,038	r
HKY+I	48	7433,886	7145,532	-3523,996	0,00	n/a	0,80	0,149	0,151	0,335	0,364	0,039	0,086	0,178	0,038	0,164	0,093	0,038	0,074	0,093	0,073	0,039	r
HKY+G	48	7434,101	7145,747	-3524,103	n/a	34,09	0,81	0,149	0,151	0,335	0,364	0,038	0,085	0,179	0,038	0,165	0,093	0,038	0,074	0,093	0,073	0,038	r
TN93	48	7434,386	7146,032	-3524,245	n/a	n/a	0,80	0,149	0,151	0,335	0,364	0,039	0,086	0,181	0,038	0,162	0,093	0,038	0,073	0,093	0,074	0,039	r
HKY+G+I	49	7441,784	7147,456	-3523,925	0,00	39,63	0,81	0,149	0,151	0,335	0,364	0,038	0,085	0,179	0,038	0,165	0,093	0,038	0,074	0,093	0,073	0,038	r
TN93+I	49	7441,923	7147,594	-3523,994	0,00	n/a	0,80	0,149	0,151	0,335	0,364	0,039	0,086	0,179	0,038	0,163	0,093	0,038	0,073	0,093	0,073	0,039	r
TN93+G	49	7442,163	7147,835	-3524,114	n/a	29,06	0,82	0,149	0,151	0,335	0,364	0,038	0,085	0,181	0,038	0,164	0,092	0,038	0,074	0,092	0,074	0,038	r
TN93+G+I	50	7449,794	7149,492	-3523,910	0,00	42,16	0,81	0,149	0,151	0,335	0,364	0,038	0,085	0,180	0,038	0,164	0,093	0,038	0,074	0,093	0,074	0,038	r
GTR	51	7452,095	7145,821	-3521,041	n/a	n/a	0,79	0,149	0,151	0,335	0,364	0,068	0,082	0,170	0,068	0,158	0,088	0,036	0,071	0,080	0,069	0,036	r
GTR+I	52	7459,779	7147,534	-3520,863	0,00	n/a	0,79	0,149	0,151	0,335	0,364	0,069	0,082	0,169	0,068	0,159	0,087	0,036	0,071	0,080	0,069	0,036	r
GTR+G	52	7459,832	7147,587	-3520,890	n/a	30,60	0,80	0,149	0,151	0,335	0,364	0,068	0,081	0,170	0,068	0,160	0,088	0,036	0,072	0,078	0,070	0,037	r
GTR+G+I	53	7467,626	7149,412	-3520,767	0,00	38,48	0,80	0,149	0,151	0,335	0,364	0,068	0,081	0,169	0,068	0,160	0,088	0,036	0,072	0,079	0,069	0,036	r
JC	43	7620,437	7361,977	-3637,370	n/a	n/a	0,50	0,250	0,250	0,250	0,250	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	r
JC+G	44	7628,185	7363,743	-3637,224	n/a	39,49	0,50	0,250	0,250	0,250	0,250	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	r
JC+I	44	7628,439	7363,998	-3637,351	0,00	n/a	0,50	0,250	0,250	0,250	0,250	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	r
JC+G+I	45	7636,204	7365,783	-3637,214	0,00	43,60	0,50	0,250	0,250	0,250	0,250	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	0,083	r

Figure 13 Modèles de substitution de nucléotides par maximum de vraisemblance

Le modèle avec la valeur BIC (Bayesian Information Criterion) la plus basse (présentée par chiffre 2 dans la figure 14) est considéré comme le meilleur.

Model	Parameters	BIC	AICc	lnL	(
T92	45	7411,753	7141,332	-3524,988	r
T92+G	46	7419,452	7143,051	-3524,818	r
T92+I	46	7419,462	7143,061	-3524,823	0
HKY	47	7426,360	7143,982	-3524,252	r
T92+G+I	47	7427,276	7144,898	-3524,710	0
HKY+I	48	7433,886	7145,532	-3523,996	0
HKY+G	48	7434,101	7145,747	-3524,103	r
TN93	48	7434,386	7146,032	-3524,245	r
HKY+G+I	49	7441,784	7147,456	-3523,925	0
TN93+I	49	7441,923	7147,594	-3523,994	0

Figure 14 Sélection du modèle optimal par BIC

Le critère BIC est un outil précieux pour la sélection de modèles de substitution. Il permet d'identifier le modèle optimal parmi plusieurs candidats en s'appuyant sur un équilibre entre la complexité du modèle et sa capacité d'adaptation aux données (Sharma et *al.*, 2021).

En privilégiant le modèle avec la valeur BIC la plus basse, on maximise les chances de choisir un modèle performant et robuste. Dans l'option "Phylogeny", nous avons sélectionné le modèle choisi et construit l'arbre. Ensuite, nous avons modifié l'arbre en utilisant l'outil "Subtrees" et le "Format Subtree". Cette modification nous a permis de regrouper les gènes (voir figure 17).

3.2 Construction des Logo :

Nous avons effectué cette opération à partir du fichier FASTA d'alignement exporté, en utilisant l'outil WebLogo.

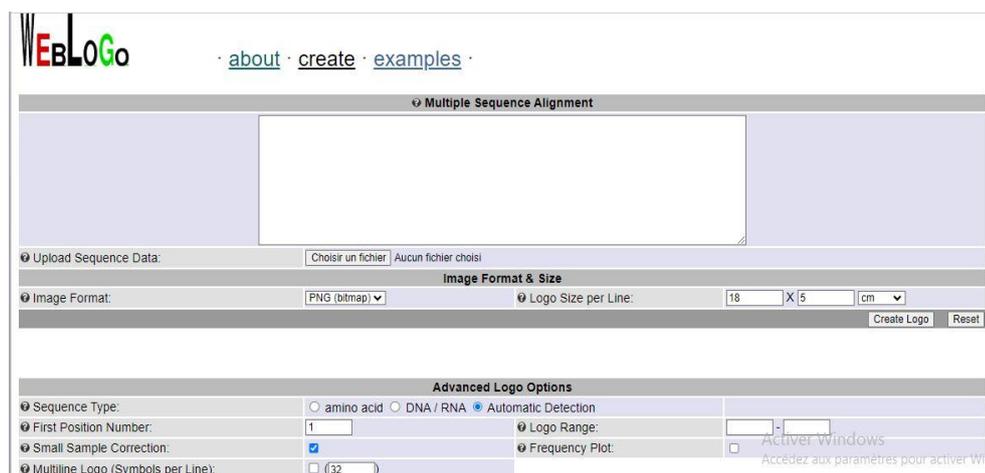


Figure 15 L'outil WebLogo

Nous avons construit un logo pour l'alignement global des régions promotrices. De plus, pour chaque groupe phylogénétique, nous avons construit un logo spécifique afin de mieux comprendre les relations entre les gènes au sein de chaque groupe. Ces logos facilitent la visualisation des résultats d'alignement, les rendant plus lisibles et permettant l'extraction de toutes les informations pertinentes liées aux îlots CpG.

3.3 Prédiction des motifs CG :

En utilisant l'outil Dotlet en mode identité avec une fenêtre de taille 2 (représenté par chiffre 3 dans la figure 16), nous avons détecté les motifs CG dans les régions promotrices des gènes.

L'utilisation de la matrice d'identité pour la visualisation des motifs CG répétés dans Dotblot propose une approche simple, claire visuellement et efficace pour l'analyse de ces motifs.

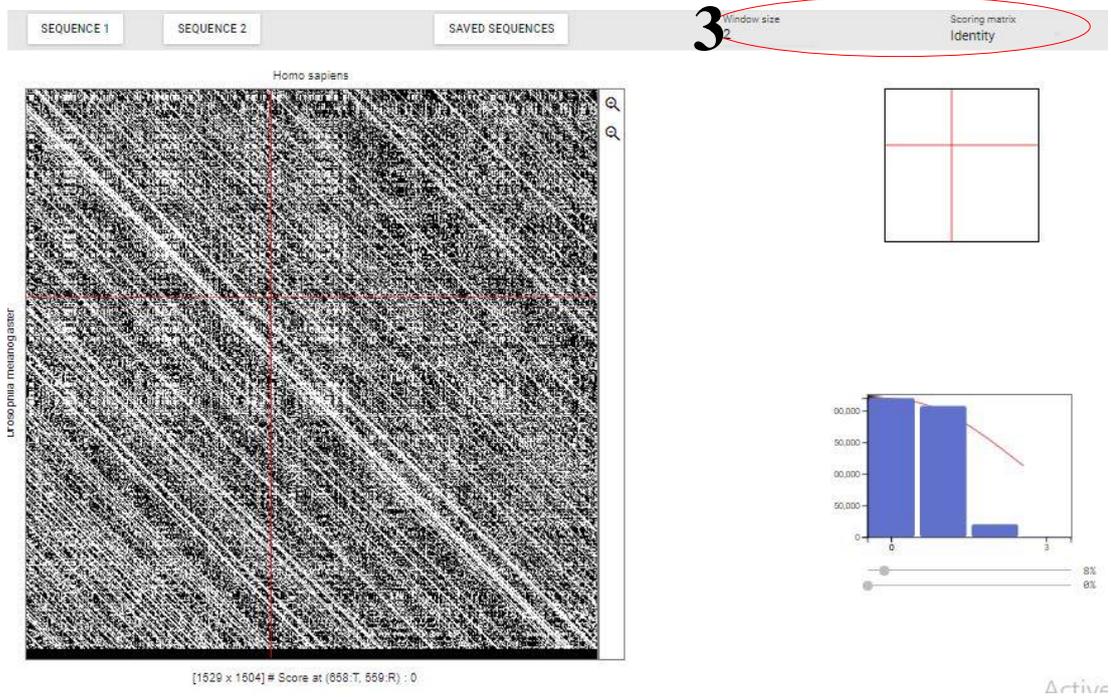


Figure 16 Interface du logiciel Dotlet avec la matrice identité et la taille de la fenêtre

La visualisation des motifs CG par dot blot nous a permis de localiser précisément les régions riches en CG, et par conséquent, les îlots CpG.

CHAPITRE 03 :

Résultats et

Discussion

Résultats

Nous présentons dans cette partie les résultats d'analyse et de la visualisation de la région promotrice des gènes.

1 Résultats de l'analyse phylogénétique des gènes :

L'arbre phylogénétique présenté dans la figure ci-dessous représente les relations évolutives entre les 23 espèces. L'arbre a été construit à partir de l'AMS de gènes promoteurs.

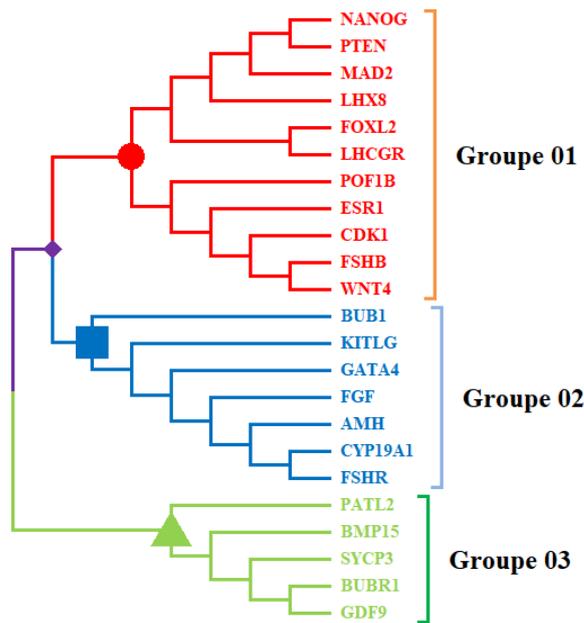


Figure 17 Analyse phylogénétique des gènes liés à l'infertilité féminine, révélant des groupes distincts.

L'arbre phylogénétique est divisé en plusieurs branches et chacune représente un gène. Les gènes les plus proches sur l'arbre sont ceux qui ont les séquences de promoteurs les plus similaires.

1.1 Analyse des différents clades identifiés dans l'arbre phylogénétique :

L'arbre phylogénétique montre que les gènes ont été regroupés en 3 groupes partagés sur 2 clades principaux. Le Groupe 1 et le Groupe 2 ont un ancêtre commun (présenté en couleur violet sur la figure 17), indiquant une relation évolutive plus étroite.

Groupe 1 (en rouge dans l'arbre phylogénétique) :

Ce groupe est composé de 11 gènes : **FOXL2, CDK1, NANOG, MAD2, ESR1, FSHB, POF1B, PTEN, WNT4, LHCGR et LHX8.**

L'analyse des séquences promotrices révèle une homologie remarquable entre ces gènes, indiquant une origine évolutive commune. Cette homologie est particulièrement frappante pour les gènes FSHB et WNT4, situés dans le dernier clade du groupe 1, qui présentent une identité de séquence complète dans la région promotrice (figure 18).

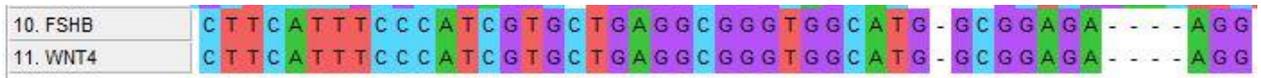


Figure 18 AMS des gènes FSHB et WNT4 révèle des régions de similarité importantes

Groupe 2 (en bleu dans l'arbre phylogénétique):

Ce groupe est composé de 7 gènes : **BUB1, KITLG, FGF, GATA4, CYP19A1, AMH.**

Les gènes du groupe 2, unis par une ascendance commune, suggèrent une stratégie évolutive partagée.

Groupe 3 (en vert dans l'arbre phylogénétique):

Ce groupe est composé de 5 gènes : **BUBR1, BMP15, PATL2, SYCP3.**

Le groupe 3 présente une divergence marquée par rapport aux 2 autres groupes. En effet, les deux premiers groupes partagent un ancêtre commun au sein d'un même clade, tandis que le groupe 3 se situe dans un clade distinct. Cette divergence phylogénétique suggère que les gènes du groupe 3 ont subi une évolution plus importante au cours du temps.

2 Résultats d'analyse des îlots CpG :

Dans cette section, nous présentons les résultats de l'analyse des îlots CpG dans les promoteurs des gènes étudiés. L'analyse a porté sur :

- a) Les résultats du calcul de pourcentage CG et la recherche de motifs CG et GC sont illustrés par 2 figures :

La figure 19 présente le CG% et le nombre des motifs CG et GC dans les promoteurs analysés.

```

Entrez votre séquence :
TGGGCGGACTCCCCATGGCCAGAGGCTGAGCTCCACTCCCGCCGGCCGCTCCCTAGGGGAAGGGG
...
Nombre de motifs GC : 36
Nombre de motifs CG : 24
Nombre total = : 60
le % GC = : 70.46413502109705 %
    
```

Figure 19 Exploration des pourcentages en CG et des motifs CG du gène FOXL2 à l'aide de Python

La figure 20 présente la distribution des motifs CG et GC (représentés en jaune) dans un promoteur spécifique, tels que prédits par l'analyse graphique par le Dotplot.

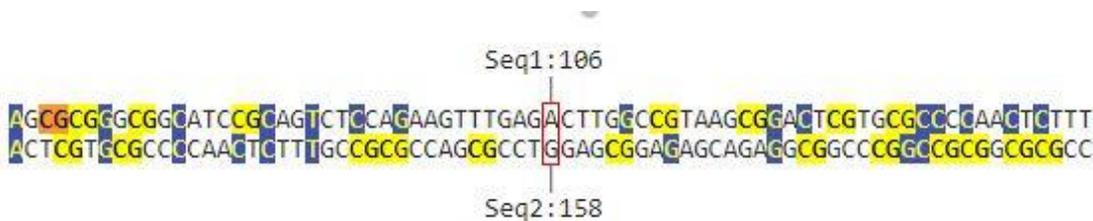


Figure 20 Distribution des motifs CG et GC dans le promoteur du gène FOXL2 révélée par Dotplot

L'analyse Dotplot Identité confirme la présence de motifs CG dans les régions promotrices des gènes étudiés.

b) Les résultats de la détection des îlots CpG par la base de données DBCAT.

Le tableau ci-dessous résume l'ensemble de ces résultats obtenus lors de l'analyse des îlots CpG dans les promoteurs des gènes étudiés.

Tableau 3 Résultats d'analyse des îlots CpG dans la région promotrice

Nom du gène	Taille de la région RP (pb)	CG% dans la RP	CpG Island		Structure et taille
			Nombre de motif CG	Nombre de motif GC	
AMH	81	79.01%	12	12	CpG isolés
BMP15	49	51.02%	0	3	CpG isolés
BUB1	206	71.36%	24	28	CpG isolés
BUBR1	81	59.26%	6	10	CpG isolés
CDK1	89	64.04%	10	11	CpG isolés
CYP19A1	221	79.18%	32	37	CpG isolés
ESR1	206	60.68%	11	25	CpG isolés
FGF	35	62.86%	3	3	CpG isolés
FOXL2	237	70.546%	24	36	1 CpG intermédiaires (205pb) CpG isolés
FSHB	44	61.36%	3	4	CpG isolés
FSHR	221	79.18%	32	37	CpG isolés
GATA4	106	71.7%	12	13	CpG isolés
GDF9	197	54.31%	5	16	CpG isolés
KITLG	188	84.04%	39	38	CpG isolés
LHCGR	53	66.03%	4	7	CpG isolés
LHX8	232	71.99%	27	39	CpG isolés
MAD2	670	70.6%	73	87	1 CpG intermédiaires (656pb) CpG isolés
NANOG	367	77.38%	37	49	1 CpG intermédiaires (312pb) CpG isolés
PATL2	223	56.95%	6	15	CpG isolés
POF1B	68	57.35%	0	6	CpG isolés
PTEN	179	78.77%	23	28	CpG isolés
SYCP3	129	58.14%	10	10	CpG isolés
WNT4	44	61.36%	3	4	CpG isolés

L'analyse des résultats dans le tableau ci-dessus révèle que le pourcentage de dinucléotide CG est supérieur à 50% dans les régions promotrices de tous les 23 gènes étudiés et une

variation du nombre de motifs CG, mais avec une tendance générale à un nombre élevé et une forte concentration de CpG dans ces régions dont toutes les 3 structures d'ilots CpG sont présentes.

3 Résultats d'analyse des Logo :

L'analyse des clusters CpG par Logo repose sur l'analyse de quatre logos :

- ✚ Un logo pour l'AMS des régions promotrices de tous les gènes.
 - ✚ Trois logos individuels, chacun dédié à l'alignement des gènes de chaque groupe phylogénétique.
- La figure ci-dessous représente le Logo global.

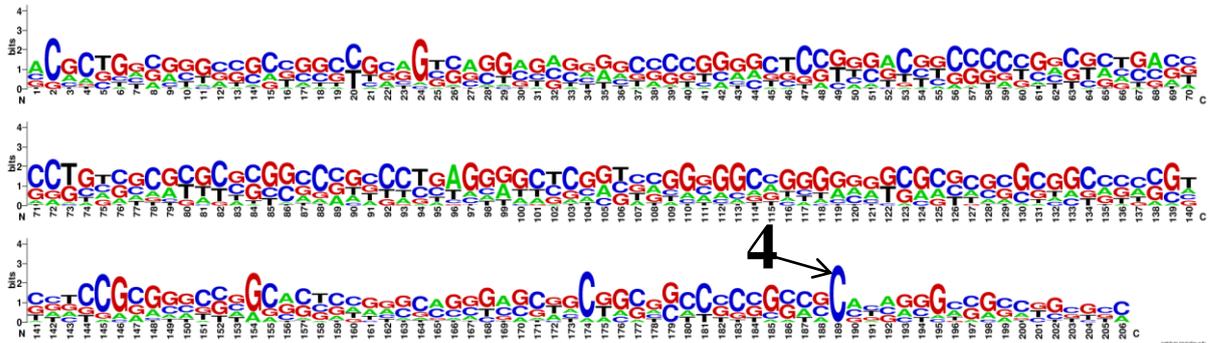


Figure 21 Représentation par logo de l'alignement des régions promotrices de tous les gènes étudiés

L'analyse visuel de (la figure 21) révèle que:

- La base "C" est conservée à la position 189 (présentées par chiffre 4 dans la figure 21) sur 10 des 23 gènes analysés. Les autres gènes présentent des gaps à cette position. L'arbre phylogénétique montre que ces 10 gènes avec la base "C" conservée sont regroupés en deux clades (groupes 1 et 2) qui partagent un ancêtre commun (voir figure 17).
 - Nous avons constaté la présence d'environ 16 palindromes de bases "C" et "G". Leurs tailles varient entre 3 et 13 pb. Ils contiennent près de 30 motifs CG dans le brin sens, en plus des motifs CG présents en dehors des palindromes (figure 21).
- La figure ci-dessous représente le logo du premier groupe

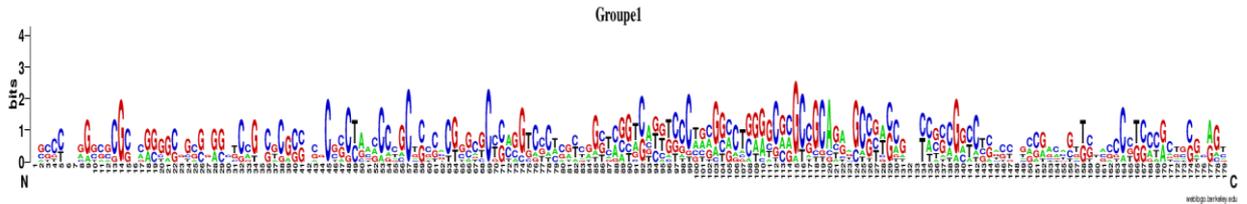


Figure 22 Représentation par Logo de l'alignement des gènes du groupe 1

L'analyse visuelle de la figure 22 révèle que :

- La séquence de la région promotrice de ce groupe de gènes est presque entièrement composée de motifs poly-C et poly-G en alternances tout au long des séquences.
- Plusieurs bases C et G ont été conservées dans diverses positions, en particulier dans les gènes NANOG, PTEN, MAD2, LHX8 et FOXL2. Notons que ces gènes appartiennent au même groupe phylogénétique.

➤ La figure ci-dessous représente le logo du deuxième groupe

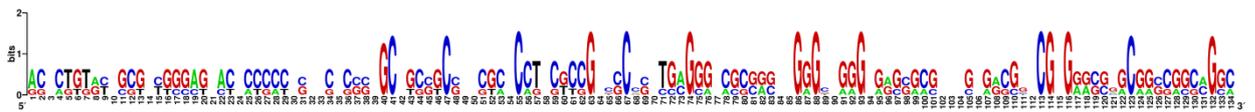


Figure 23 Représentation par Logo de l'alignement des gènes du groupe 2

L'analyse du Logo ci-dessus a révélé :

- Une proportion remarquablement élevée des séquences des gènes du groupe 2 est composée de bases C et G.
- Présence de motifs poly-C et poly-G.
- La conservation des bases C et G est particulièrement remarquable dans les gènes du groupe 2.

➤ La figure ci-dessous représente le Logo du troisième groupe

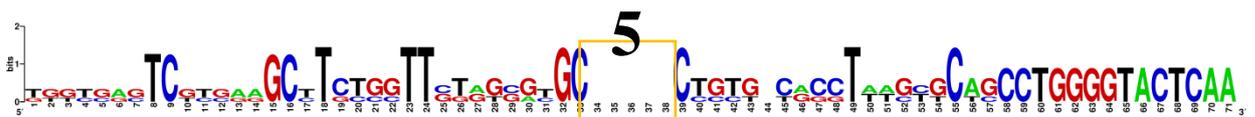


Figure 24 Représentation par Logo de l'alignement des gènes du groupe 3

L'analyse du Logo ci-dessus a révélé :

- Une conservation de bases plus élevées entre les séquences.
- La présence de poly-C, poly-G.
- La présence d'un motif conservé dans la région 34-38 (représentée par chiffre 5 dans la figure 24) du premier gène de l'alignement. Ce motif est absent ou remplacé par des gaps dans les autres gènes du groupe.

Discussion :

En analysant les régions promotrices de 23 gènes impliqués dans la fonction ovarienne et la fertilité, nous avons constaté la présence de groupes monophylétiques. Ces gènes sont regroupés en fonction de leurs relations évolutives et fonctionnelles, révélant que les gènes du groupe 1, essentiels à la différenciation ovarienne et à la fertilité, partagent des clusters de méthylation similaires. Le gène FOXL2 interagit directement et indirectement avec certains de ces gènes, suggérant des fonctions communes et donc influençant ainsi la fonction ovarienne et la santé reproductive. Chez les femmes infertiles, une hyperméthylation des sites CpG dans la région promotrice du gène FOXL2 a été observée, potentiellement influencée par les motifs poly-C/poly-G (Dong, X. et *al.*, 2013). De plus, la similarité des séquences promotrices des gènes FSHB et WNT4 souligne leur rôle crucial dans la régulation de la maturation folliculaire et de la libération des ovules. Les gènes du groupe 2, impliqués dans la croissance et la fonction gonadiques, partagent des caractéristiques avec le groupe 1, notamment une forte teneur en CG et des motifs poly-C/poly-G étendus, influençant la méthylation de l'ADN et la liaison des protéines. Les travaux sur le gène KITLG ont mis en lumière son implication dans l'infertilité masculine et féminine, soulignant l'importance des altérations de la méthylation des CpG dans ces troubles (Bayraktar, B. S. et *al.*, 2014). Le troisième groupe de gènes, essentiels à la fonction reproductrice et à la stabilité du génome, présente une forte conservation des bases C et G dans leurs séquences, avec des motifs conservés comme celui du gène BUBR1 qui pourrait jouer un rôle clé dans la régulation de l'expression épigénétique grâce à sa composition en motifs CG. L'analyse de la distribution des clusters CpG dans les régions promotrices des gènes étudiés a révélé une corrélation entre la localisation et la densité des CpG et les fonctions des gènes. Les gènes du groupe 1, avec des motifs poly-C/poly-G étendus, pourraient être plus sensibles à la méthylation. Pour les CpG dispersés des groupes 2, leur présence peut influencer la formation de clusters CpG dans le temps, potentiellement en favorisant la méthylation des sites CpG adjacents et la création de structures CpG denses. Alors que la conservation des séquences dans le groupe 2 pourrait protéger les CpG fonctionnels. Le groupe 3, avec une teneur en CG plus faible, pourrait être éventuellement régulé par d'autres mécanismes génétiques et/ou épigénétiques. D'après une analyse complémentaire des régions hors promoteur de nos gènes, nous avons observé une distribution hétérogène des îlots CpG. Des altérations de la méthylation de ces îlots CpG, notamment une hyperméthylation, peuvent affecter la régulation de l'expression des gènes impliqués dans des processus clés tels que la maturation ovocytaire, la fertilisation ou l'implantation embryonnaire, potentiellement contribuant ainsi à l'infertilité féminine.

La visualisation du schéma de la distribution des ilots CpG dans les régions promotrices a montré un profil similaire à celui de nos gènes étudiés, confirmant ainsi le rôle de l'épigénétique dans la régulation des gènes impliqués dans l'infertilité féminine.

*Conclusion Et
Perspectives*

Conclusion et perspectives :

L'épigénétique, par ses multiples mécanismes, s'impose comme un régulateur clé de l'expression des gènes impliqués dans la reproduction féminine, en ciblant spécifiquement les îlots CpG de leurs promoteurs. Cette modulation de l'expression génique est influencée par des facteurs environnementaux et un mode de vie inapproprié. L'analyse *in silico* nous a permis d'identifier et de caractériser les îlots CpG dans les promoteurs des 23 gènes étudiés. Les résultats obtenus révèlent des profils spécifiques de la distribution des îlots CpG, permettant d'inférer leur degré de méthylation potentiel, suggérant une régulation épigénétique différentielle. Ces résultats préliminaires ouvrent la voie à des recherches futures visant à décrypter les mécanismes précis par lesquels la méthylation des îlots CpG dans les promoteurs des gènes de la reproduction féminine peut conduire à l'infertilité.

Pour nos perspectives :

Nous proposons de développer des algorithmes de détection d'îlots CpG spécifiques aux tissus reproductifs féminins et utiliser le machine learning pour identifier leurs motifs discriminants, permettant une détection plus précise. Ensuite, créer des bases de données annotées dédiées aux tissus reproductifs féminins, stockant des informations sur chaque îlot CpG. Enfin, mettre en place des protocoles d'analyse de la méthylation des îlots CpG chez des patientes fertiles et infertiles pour comparer leurs profils de méthylation et identifier des différences significatives.

En outre, l'analyse des profils de méthylation des îlots CpG permettra d'identifier de nouveaux biomarqueurs épigénétiques potentiels pour le diagnostic et le suivi thérapeutique de l'infertilité féminine.

*Liste de références
bibliographiques*

Liste de références bibliographiques

1. Aimeur, S., Manseur, F., Krouk, Y., & Oukid, M. S. (2021). Contribution of microRNAs in infertility.
2. Allsworth, J. E., Clarke, J., Peipert, J. F., Hebert, M. R., CRNP, A. C., & Boardman, L. A. (2007). The influence of stress on the menstrual cycle among newly incarcerated women. *Women's Health Issues: Official Publication of the Jacobs Institute of Women's Health*, 17(4), 202-209. <https://doi.org/10.1016/j.whi.2007.02.002>
3. Antequera, F. (2003). Structure, function and evolution of CpG island promoters. *Cellular and Molecular Life Sciences: CMLS*, 60(8), 1647-1658. <https://doi.org/10.1007/s00018-003-3088-6>
4. Apicella, C., Ruano, C. S. M., Méhats, C., Miralles, F., & Vaiman, D. (2019). The Role of Epigenetics in Placental Development and the Etiology of Preeclampsia. *International Journal of Molecular Sciences*, 20(11), 2837. <https://doi.org/10.3390/ijms20112837>
5. Biason-Lauber, A. (2012). WNT4, RSPO1, and FOXL2 in sex development. *Seminars in Reproductive Medicine*, 30(5), 387-395. <https://doi.org/10.1055/s-0032-1324722>
6. Biol, J-C. (2011). DBCAT : Database of CpG islands and analytical tools for identifying comprehensive methylation profiles in cancer cells.
7. Bydlowski, M. (2019). Étude psychopathologique de l'infertilité féminine. *Le Carnet PSY*, 222(1), 24-34. <https://doi.org/10.3917/lcp.222.0024>
8. Cederroth, C. R., Vassalli, J.-D., & Nef, S. (2007). De l'épigénétique et du développement. *Revue Médicale Suisse*, 3(100), 528-532. <https://www.revmed.ch/revue-medicale-suisse/2007/revue-medicale-suisse-100/de-l-epigenetique-et-du-developpement>
9. Chamani, I. J., & Keefe, D. L. (2019). Epigenetics and Female Reproductive Aging. *Frontiers in Endocrinology*, 10, 473. <https://doi.org/10.3389/fendo.2019.00473>
10. Chatterjee, S., Modi, D., Maitra, A., Kadam, S., Patel, Z., Gokrall, J., & Meherji, P. (2007). Screening for FOXL2 gene mutations in women with premature ovarian failure : An Indian experience. *Reproductive Biomedicine Online*, 15(5), 554-560. [https://doi.org/10.1016/s1472-6483\(10\)60388-4](https://doi.org/10.1016/s1472-6483(10)60388-4)
11. Clémentine, S et Mélanie, L-C (2021). Les causes d'infertilité féminine. <https://www.fertility.com/fr-fr/causes-infertilité/feminine.html>.
12. DBCAT: Database of CpG Islands and Analytical Tools for Identifying Comprehensive Methylation Profiles in Cancer Cells | Journal of Computational Biology. (s. d.). Consulté 25 mai 2024, à l'adresse <https://www.liebertpub.com/doi/abs/10.1089/cmb.2010.0038?journalCode=cmb>
13. Deninotti, J. (2023). Représentations et ajustement psychosocial des femmes confrontées à l'infertilité. Université de Nîmes.
14. Fortier, J.-F. (s. d.). Dinucléotide CpG : Définition et explications. AquaPortail. Consulté 30 mai 2024, à l'adresse <https://www.aquaportail.com/dictionnaire/definition/15037/dinucléotide-cpg>
15. Fortin, I. (2020). Le vécu des hommes et des femmes souffrant ou ayant souffert de difficultés de procréation au Saguenay-Lac-Saint-Jean. Chicoutimi.
16. Ge, Z.-J., Schatten, H., Zhang, C.-L., & Sun, Q.-Y. (2015). Oocyte ageing and epigenetics. *Reproduction* (Cambridge, England), 149(3), R103-R114. <https://doi.org/10.1530/REP-14-0242>
17. Gicquel, C., El-Osta, A., & Le Bouc, Y. (2008). Epigenetic regulation and fetal programming. *Best Practice & Research. Clinical Endocrinology & Metabolism*, 22(1), 1-16. <https://doi.org/10.1016/j.beem.2007.07.009>
18. Gonon, F., & Moisan, M.-P. (2013). L'épigénétique, la nouvelle biologie de l'histoire individuelle ? : *Revue française des affaires sociales*, 1, 21-31. <https://doi.org/10.3917/rfas.125.0021>
19. Guibourdenche, M. (s. d.). Effets d'une exposition chronique maternelle à un pesticide et un régime obésogène sur les fonctions physiologiques digestives : Approches in vivo et in vitro.
20. Hamamah, S. (2023). Vers une stratégie nationale de lutte contre l'infertilité. In *Infertilité* (p. 439-444). Elsevier. <https://doi.org/10.1016/B978-2-294-78134-6.00049-8>
21. Heard, E. (2014). Épigénétique et mémoire cellulaire. *L'annuaire du Collège de France. Cours et travaux*, 113, Article 113. <https://doi.org/10.4000/annuaire-cdf.2420>
22. <http://intersono-ivf.com/fr/blog/comment-le-stress-affecte-la-capacite-de-tomber-enceinte>. Consulté 26 avril 2024.
23. <https://acces.ens-lyon.fr/biotic/procreat/amp/html/EmpreintParent.htm>. Consulté 27 avril 2024.
24. <https://dotlet.vital-it.ch/>. Consulté 25 mai 2024, à l'adresse <https://dotlet.vital-it.ch/>
25. <https://weblogo.berkeley.edu/logo.cgi>. Consulté 25 mai 2024.

26. <https://www.deuxiemeavis.fr/pathologie/infertilite-feminine>. Consulté 24 avril 2024.
27. <https://www.gyneco-online.com/fertilite/le-risque-epigenetique-en-assistance-medical-la-procreation-donnees-epidemiologiques>. Consulté 24 avril 2024.
28. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9225096/>. Consulté 27 avril 2024.
<https://www.sciencedirect.com/science/article/pii/S0001407919306545>
https://www.unboundmedicine.com/medline/citation/21214365/DBCAT:_database_of_CpG_islands_and_analytical_tools_for_identifying_comprehensive_methylation_profiles_in_cancer_cells_
29. Jullien, P. E., & Berger, F. (2008a). L’empreinte parentale, un phénomène épigénétique essentiel pour la reproduction des plantes. *médecine/sciences*, 24(8-9), Article 8-9. <https://doi.org/10.1051/medsci/20082489753>
30. Junien, C., Panchenko, P., Fneich, S., Pirola, L., Chriett, S., Amarger, V., Kaeffer, B., Parnet, P., Torrisani, J., Bolaños Jimenez, F., Jammes, H., & Gabory, A. (2016). Épigénétique et réponses transgénérationnelles aux impacts de l’environnement: Des faits aux lacunes. *médecine/sciences*, 32(1), 35-44. <https://doi.org/10.1051/medsci/20163201007>
31. National Center for Biotechnology Information. (s. d.). Consulté 25 mai 2024, à l’adresse <https://www.ncbi.nlm.nih.gov/>
32. Nedjma, C., & Amina, T. (s. d.). Contribution à l’étude des facteurs de risques de l’infertilité féminine : Myome chez la femme.
33. Pannetier, M., & Pailhoux, E. (2010). FOXL2, le gardien de l’identité ovarienne. *médecine/sciences*, 26(5), Article 5. <https://doi.org/10.1051/medsci/2010265470>
34. Pisarska, M. D., Chan, J. L., Lawrenson, K., Gonzalez, T. L., & Wang, E. T. (2018). Genetics and Epigenetics of Infertility and Treatments on Outcomes. *The Journal of Clinical Endocrinology and Metabolism*, 104(6), 1871-1886. <https://doi.org/10.1210/jc.2018-01869>
35. Reiner, A. (2016) Le facteur de transcription FOXL2 : Un acteur clé de la différenciation de l’ovaire, de son maintien et de la fertilité, 200(6), 115-1127.
36. Weber, M. (2008). Profils de méthylation de l’ADN dans les cellules normales et cancéreuses. *médecine/sciences*, 24(8-9), Article 8-9. <https://doi.org/10.1051/medsci/20082489731>
37. Yatsenko, S. A., & Rajkovic, A. (2019). Genetics of human female infertility†. *Biology of Reproduction*, 101(3), 549-566. <https://doi.org/10.1093/biolre/ioz084>

Annexes :

Fonctions et rôles des gènes clés de la fertilité

Gène	Fonction	Rôle dans la reproduction féminine
AMH	Hormone anti-müllérienne	Inhibe la croissance folliculaire
BMP15	Protéine morphogénique osseuse 15	Rôle dans la maturation folliculaire et la stéroïdogénèse
BUB1	Protéine de liaison à l'ubiquitine 1	Rôle dans la division méiotique
BUBR1	Protéine de liaison à l'ubiquitine 1, homologue 1	Rôle dans la division méiotique
CDK1	Kinase dépendante de la cycline 1	Rôle dans la progression du cycle cellulaire
CYP19A1	Aromatase	Convertit les androgènes en œstrogènes
ESR1	Récepteur aux œstrogènes 1	Rôle dans la maturation folliculaire, la prolifération endométriale et l'implantation embryonnaire
FGF	Facteur de croissance des fibroblastes	Rôle dans la croissance folliculaire, la maturation ovocytaire et l'angiogénèse
FOXL2	Facteur de transcription Forkhead box L2	Rôle dans le développement des ovaires
FSHR	Récepteur à l'hormone folliculo-stimulante	Rôle dans la croissance folliculaire et la production d'œstrogènes
FSHB	Hormone folliculo-stimulante bêta	Rôle dans la croissance folliculaire et la production d'œstrogènes
GATA4	Facteur de transcription GATA 4	Rôle dans le développement des gonades et la maturation folliculaire
GDF9	Facteur de croissance différentiel 9	Rôle dans la croissance folliculaire et la maturation ovocytaire
KITLG	Ligand du kit	Rôle dans la migration des cellules germinales et la maturation folliculaire
LHCGR	Récepteur à la LH/CG	Rôle dans l'ovulation et la production de progestérone
LHX8	Facteur de transcription LIM homeobox 8	Rôle dans le développement des ovaires
MAD2	Protéine de liaison à l'ubiquitine 2	Rôle dans la division méiotique
NANOG	Facteur de transcription de l'embryon	Rôle dans le maintien de la pluripotence des cellules souches embryonnaires
PATL2	Protéine de liaison aux acides gras périoxisomiques 2	Rôle dans le métabolisme des lipides
POF1B	Protéine folliculosténotine 1B	Rôle dans la fonction ovarienne et la fertilité
PTEN	Phosphatase et tensine homologue 1	Rôle dans la croissance cellulaire et la régulation de la signalisation PI3K
SYCP3	Protéine 3 du complexe synaptonémal	Rôle dans la recombinaison méiotique
WNT4	Protéine Wnt 4	Rôle dans la croissance folliculaire et la maturation ovocytaire

Relation entre épigénétique et infertilité féminine : Approche *in Silico***Mémoire pour l'obtention du diplôme de Master en Bioinformatique****Résumé**

L'étude des mécanismes épigénétiques est cruciale pour comprendre la régulation des gènes impliqués dans la reproduction féminine. Ce travail vise à explorer comment les îlots CpG dans les promoteurs de 23 gènes influencent l'infertilité féminine. À l'aide de méthodes *in silico* et de bioinformatique avancée, nous avons identifié et analysé les motifs CpG au sein des promoteurs des gènes sélectionnés. Nos analyses révèlent des modifications épigénétiques significatives dans plusieurs gènes, suggérant leur rôle potentiel dans la régulation de la fertilité. Ces résultats offrent de nouvelles perspectives pour la recherche et le diagnostic de l'infertilité féminine, soulignant l'importance des approches bioinformatiques pour l'étude des mécanismes épigénétiques.

Mots-clefs : Épigénétique, Infertilité féminine, *in silico*, îlots CpG, bioinformatique, reproduction féminine.

Président du jury : MEDJROUBI Med Elarbi (MCB - Université Frères Mentouri Constantine 1).

Encadrant : BENZAADA Mostfa (MCA – Université Frères Mentouri Constantine 1).

Examinatrice : MEZIANI (MAB - Université Frères Mentouri Constantine 1).