

Université de Montréal

**Perturbation des profils épigénétiques suite à une perte temporaire du maintien de la méthylation de l'ADN dans les cellules embryonnaires**

par

Virginie Bertrand-Lehouillier

Département de biochimie et médecine moléculaire

Faculté de Médecine

Mémoire présenté à la Faculté des études supérieures  
en vue de l'obtention du grade de Maître ès sciences (M.Sc.)  
en biochimie option générale

Août 2018

© Virginie Bertrand-Lehouillier, 2018

Université de Montréal  
Faculté des études supérieures et postdoctorales

Ce mémoire s'intitule:

Perturbation des profils épigénétiques suite à une perte temporaire du maintien de la  
méthylation de l'ADN dans les cellules embryonnaires

Présenté par : Virginie Bertrand-Lehouillier

Évalué par un jury composé des personnes suivantes :

Hugo Wurtele (PhD.), président-rapporteur

Serge McGraw (PhD.), directeur de recherche

Noël J-M Raynal (PhD.), co-directeur

Julie Lessard (PhD.), membre du jury

## RÉSUMÉ

Chez l'embryon précoce, une vague de reprogrammation majeure survient et permet de réinitialiser les profils de méthylation d'ADN de l'ensemble du génome. Lors de cette reprogrammation, les régions différentiellement méthylées (DMRs) (i.e., gènes empreintes) doivent toutefois être protégées de la déméthylation par une action continue de DNMT1 (Méthyltransférase d'ADN 1) pour assurer le développement adéquat de l'épigénome du fœtus. Sachant que l'induction d'une perte temporaire d'expression de *Dnmt1* dans un modèle de cellules souches embryonnaires de souris entraîne la perte permanente des patrons de méthylation d'ADN aux régions DMRs et DMR-like, mon projet de recherche vise à comprendre pourquoi ces régions sont incapables de retrouver leurs patrons de méthylation d'ADN initiaux. Notre hypothèse est qu'une adaptation épigénétique (i.e. réarrangement erroné de certaines modifications d'histones) survient aux régions régulatrices de l'expression des gènes (promoteurs et enhanceurs) et empêche directement ou indirectement le retour au paysage épigénétique initial aux régions affectées. L'objectif du projet est donc de précisément définir comment la perte temporaire de *Dnmt1* remodèle le paysage épigénétique aux régions promotrices (*H3K4me3*, *H3K27me3*, *H3K27ac*, *H3K4me1*, *H3K9me3*, méthylation d'ADN) et comment les adaptations épigénétiques sont associées avec des changements de l'expression des gènes (ex : gènes des régions DMRs et DMRs-like).

Mots-clés :

Épigénétique, Préimplantation, Embryon, Modifications d'histones, Méthylation d'ADN, DNMT1, Expression génique, Promoteurs, Enhancers

## ABSTRACT

In early embryos, a major reprogramming wave occurs and permits to reset DNA methylation profiles genome-wide. During the reprogramming wave, differentially methylated regions (DMRs) (imprinted genes) must be protected from demethylation by the continuous action of DNMT1 (DNA Methyltransferase 1) to ensure the proper development of the foetal epigenome. As the induction of a temporary loss of *Dnmt1* expression in a mouse embryonic stem cell model leads to permanent losses of DNA methylation at DMR and DMR-like regions, my project aims to understand why those regions are unable to re-establish their initial DNA methylation patterns. Our hypothesis is that an epigenetic adaptation (erroneous rearrangement of certain histone modifications) occurs at regulatory regions controlling gene expression (promoters and enhancers) and impede directly or indirectly the affected regions to return to their initial epigenetic landscape. The goal of this project is thus to define how the temporary loss of *Dnmt1* remodels the epigenetic landscape at promoter regions (*H3K4me3*, *H3K27me3*, *H3K27ac*, *H3K4me1*, *H3K9me3*, DNA methylation) and how the epigenetic adaptations are associated with changes in gene expression (ex: genes in DMR and DMR-like regions).

Keywords:

Epigenetics, Preimplantation, Embryo, Histone Modifications, DNA Methylation, DNMT1, Gene expression, Promoters, Enhancers

# TABLE DES MATIÈRES

RÉSUMÉ.....	i
ABSTRACT.....	ii
TABLE DES MATIÈRES.....	iii
LISTE DES TABLEAUX.....	vi
LISTE DES FIGURES.....	vii
LISTE DES ABBRÉVIATIONS.....	ix
REMERCIEMENTS.....	xi
MISE EN CONTEXTE.....	15
1. STRUCTURE DE LA CHROMATINE.....	15
1.1 <i>Le génome</i> .....	15
1.2 <i>Nucléosomes et histones</i> .....	18
1.3 <i>Euchromatine et hétérochromatine</i> .....	19
2. MODIFICATIONS ÉPIGÉNÉTIQUES.....	20
2.1 <i>Modification des histones</i> .....	20
2.2 <i>Méthylation de l'ADN et DNMTs</i> .....	23
2.2.1 <i>DNMT1 et ses interactions</i> .....	25
2.2.2 <i>Déméthylation</i> .....	27
2.2.3 <i>Gènes empreintes et « imprinted-like »</i> .....	28
3. INTERACTION ENTRE LES MODIFICATIONS ÉPIGÉNÉTIQUES.....	30
3.1 <i>Modifications épigénétiques et expression génique : l'oeuf ou la poule?</i> .....	30
4. REPROGRAMMATION ÉPIGÉNÉTIQUE DANS LE DÉVELOPPEMENT EMBRYONNAIRE PRÉCOCE.....	32
4.1 <i>Embryogenèse et développement embryonnaire précoce (préimplantation) murin</i> .....	32
4.2 <i>Reprogrammation des patrons de méthylation dans la préimplantation</i> .....	33
4.3 <i>Distribution des modifications d'histones dans la préimplantation</i> .....	35
5. PERTURBATION DU PROGRAMME ÉPIGÉNÉTIQUE LORS DE LA PRÉIMPLANTATION.....	36
5.1 <i>Erreurs épigénétiques développementales</i> .....	37
5.2 <i>Désordres engendrés par des erreurs épigénétiques dans le programme foetal</i> .....	38
6. MODÉLISATION D'UNE PERTE D'EXPRESSION DE <i>DNMT1</i> DANS LES CELLULES EMBRYONNAIRES: EMBRYONS VS CELLULES SOUCHES EMBRYONNAIRES.....	39
6.1 <i>Modèle murin (embryons)</i> .....	39
6.2 <i>Cellules souches</i> .....	40
7. PROBLÉMATIQUE.....	41

8. HYPOTHÈSES.....	42
9. OBJECTIFS .....	43
9.1 Modélisation de la dérégulation épigénétique héréditaire embryonnaire avec un modèle de cellules souches <i>Dnmt1<sup>tet/tet</sup></i> .....	43
9.2 Établissement des profils d'expression génique, de méthylation d'ADN ainsi que des distributions de marques d'histones .....	43
9.3 Annotation des régions promotrices par bio-informatique .....	44
9.4 Déterminer la réponse adaptative de chacun des facteurs épigénétiques étudiés et leur dynamique .....	44
10. MATÉRIEL ET MÉTHODES .....	44
10.1 Modèle cellulaire <i>Dnmt1<sup>tet/tet</sup></i> .....	44
10.2 Culture et récolte des cellules <i>Dnmt1<sup>tet/tet</sup></i> .....	46
10.3 Validation de l'expression de <i>Dnmt1</i> par qPCR .....	47
10.4 Immunoprécipitation de la chromatine (ChIP-Seq) couplée au séquençage de haut-débit .....	48
10.4.1 Principe .....	48
10.4.2 Préparation des immunoprécipitations et validation de l'enrichissement .....	49
10.4.3 Préparation des librairies et séquençage .....	51
10.5 Séquençage de l'ARN à haut-débit.....	51
10.6 Séquençage au bisulfite à représentation réduite (RRBS).....	52
10.6.1 Principe et accession des données.....	52
10.7 Traitement bioinformatique des données brutes de séquençage .....	52
10.7.1 Traitement et analyse des données brutes de séquençage de ChIP.....	52
10.7.2 Traitement et analyse des données brutes de séquençage de l'ARN.....	54
10.7.3 Traitement et analyse des données brutes de séquençage RRBS .....	54
10.8 Analyse du profil épigénétique aux promoteurs .....	55
10.8.1 Définition des annotations pour les promoteurs .....	55
11. RÉSULTATS.....	58
11.1 Validation du modèle cellulaire ES <i>Dnmt1<sup>tet/tet</sup></i> .....	58
11.1.1 Perte d'expression de <i>Dnmt1</i> .....	58
11.2 Établissement des profils d'expression génique, de méthylation d'ADN ainsi que des distributions de marques d'histones .....	59
11.2.1 Distribution des modifications d'histones .....	59
11.3 Annotation des régions promotrices par bio-informatique .....	60

<i>11.4 Caractérisation de la réponse adaptative de chacun des facteurs épigénétiques étudiés aux régions promotrices</i> .....	64
<i>11.4.1 Promoteurs actifs</i> .....	66
<i>11.4.2 Promoteurs bivalents</i> .....	71
<i>11.4.3 Promoteurs réprimés</i> .....	76
<i>11.4.4 Promoteurs déplétés</i> .....	81
<i>11.5 Adaptations épigénétiques des gènes empreintes et imprinted-like (Cluster Xlr)</i> .....	86
<i>11.5.1 Gènes empreintes</i> .....	86
<i>11.5.2 Gènes imprinted-like</i> .....	88
12. DISCUSSION .....	92
<i>12.1 Différences d'adaptations épigénétiques entre les types de promoteurs</i> .....	92
<i>12.2 Adaptation épigénétique des gènes empreintes</i> .....	97
<i>12.3 Adaptation épigénétique des gènes de la famille Xlr (imprinted-like)</i> .....	98
<i>12.4 Limitations de l'étude</i> .....	98
13. PERSPECTIVES.....	100
14. CONCLUSIONS .....	102
15. CONTRIBUTIONS.....	103
16. BIBLIOGRAPHIE.....	104

## LISTE DES TABLEAUX

Tableau I. Séquences des amorces utilisées pour valider l'enrichissement des immunoprécipitations

Tableau II. Nombre de lectures de séquençage brutes obtenues pour chaque duplicata des données de ChIP-Seq

Tableau III. Nombre de lectures de séquençage alignées obtenues pour chaque duplicata de bibliothèques de RNA-Seq

Tableau IV. Nombre de lectures de séquençage alignées obtenues par ligne de séquençage pour les données de RNA-Seq



## LISTE DES FIGURES

Figure 1. Structure détaillée des chromosomes dans le noyau de la cellule : des constituants de la chromatine jusqu'aux paires de bases de l'ADN

Figure 2. Localisation et structure d'un promoteur et d'un enhancer

Figure 3. Structure du nucléosome : histones et variants d'histones impliqués

Figure 4. Représentation schématique de l'hétérochromatine et de l'euchromatine

Figure 5. Sites de modifications post-traductionnelles sur les queues des histones

Figure 6. Apposition de la méthylation sur les brins d'ADN hémi-méthylés par Dnmt1 après la réplication

Figure 7. Facteurs interagissant avec Dnmt1 à la fourche de réplication pour maintenir la méthylation d'ADN sur les brins hémi-méthylés

Figure 8. Représentation des étapes du développement embryonnaire précoce chez la souris

Figure 9. Schéma de la vague de reprogrammation épigénétique survenant dans le développement embryonnaire précoce

Figure 10. Distribution temporelle globale de la méthylation d'ADN et de modifications d'histones tout au long du développement embryonnaire précoce de la fertilisation à l'implantation dans l'utérus

Figure 11. Construction et validation du modèle ES *Dnmt1<sup>tet/tet</sup>*

Figure 12. Récolte des cellules *Dnmt1<sup>tet/tet</sup>* à différents temps expérimentaux

Figure 13. Représentation schématique de la technique d'immunoprécipitation de la chromatine couplée au séquençage à haut-débit (ChIP-Seq)

Figure 14. Proportion de promoteurs retenus pour les analyses

Figure 15. Définition des annotations de promoteurs

Figure 16. Induction de la perte temporaire d'expression de *Dnmt1* dans les cellules ES *Dnmt1<sup>tet/tet</sup>*

Figure 17. Corrélation entre les comptes de lectures du chromosome 1 des bibliothèques de ChIP-Seq analysées

Figure 18. Densité des modifications d'histones observée pour chaque catégorie de promoteurs

Figure 19. Proportions des types de promoteurs dans notre annotation maison et leur expression

Figure 20. Proportions des promoteurs totaux ayant subi un changement de la distribution d'H3K4me3 entre les conditions Ctl et d0 ainsi qu'entre Ctl et d21

Figure 21. Adaptation épigénétique (modifications d'histones, méthylation d'ADN et expression) aux promoteurs actifs entre les conditions Ctl et d21 dans les cellules ES *Dnmt1<sup>tet/tet</sup>*

Figure 22. Évaluation de la méthylation et de la densité en CpGs aux promoteurs actifs ainsi que l'expression globale des gènes associés

Figure 23. Adaptation épigénétique (modifications d'histones, méthylation d'ADN et expression) aux promoteurs bivalents entre les conditions Ctl et d21 dans les cellules ES *Dnmt1<sup>tet/tet</sup>*

Figure 24. Évaluation de la méthylation et de la densité en CpGs aux promoteurs bivalents ainsi que l'expression globale des gènes associés

Figure 25. Adaptation épigénétique (modifications d'histones, méthylation d'ADN et expression) aux promoteurs réprimés entre les conditions Ctl et d21 dans les cellules ES *Dnmt1<sup>tet/tet</sup>*

Figure 26. Évaluation de la méthylation et de la densité en CpGs aux promoteurs réprimés ainsi que l'expression globale des gènes associés

Figure 27. Adaptation épigénétique (modifications d'histones, méthylation d'ADN et expression) aux promoteurs déplétés entre les conditions Ctl et d21 dans les cellules ES *Dnmt1<sup>tet/tet</sup>*

Figure 28. Évaluation de la méthylation et de la densité en CpGs aux promoteurs déplétés ainsi que l'expression globale des gènes associés

Figure 29. Adaptations épigénétiques aux régions de contrôle des empreintes (ICRs) des gènes empreintes

Figure 30. Adaptation de l'expression des gènes de la famille *Xlr*

Figure 31. Représentation des pics de modifications d'histones, de méthylation d'ADN et d'expression avec IGV (Integrative Genomics Viewer) pour les conditions Ctl, d0 et d21 dans les régions promotrices et géniques des gènes *Xlrs*

Figure 32. Représentation schématique du paysage épigénétique selon la condition expérimentale (Contrôle, Inactivation de *Dnmt1*, Réactivation de *Dnmt1*) pour les promoteurs actifs et bivalents

Figure 33. Représentation schématique du paysage épigénétique selon la condition expérimentale (Contrôle, Inactivation de *Dnmt1*, Réactivation de *Dnmt1*) pour les promoteurs réprimés et déplétés

## LISTE DES ABBRÉVIATIONS

5-mC : 5-méthylcytosine

5-hmC : 5-hydroxyméthylcytosine

Ac : Acétylation

ADCA-DN : Ataxie cérébelleuse autosomique dominante de type 1 - surdité et narcolepsie

ADN : Acide désoxyribonucléique

ADNc: ADN complémentaire

ARN : Acide Ribonucléique

gARN: ARN guide

ncARNs : ARNs non-codants

rRNA: ARNs ribosomaux

miRNA: microARNs

snoRNA: Small nucleolar RNA

snRNA : Small Nuclear RNA

tARNs : ARNs de transfert

BWA : Burrows-Wheeler Aligner

ChIP : Chromatin Immunoprecipitation

CpG : dinucléotide cytosine-phosphate-guanine

CRISPR : Clustered Regularly Interspaced Short Palindromic Repeats

DOX: doxycycline

DMR : Région différenciellement méthylée

DNMT : Méthyltransférase d'ADN

EPI : Épiblaste

ES : Cellules embryonnaires souches

FBS: Sérum bovin fœtal

FPKM : Fragments per Kilobase per Million Mapped Reads

HATs : Histones acétyltransférases

HDACs : Histones déacétylases

HMTs : Histones méthyltransférases  
HDMs : Histone déméthylases  
HSAN1E : Neuropathie héréditaire sensorielle et autonome de type 1E  
HSC: Cellules souches hématopoïétiques  
IC : Centre d’empreinte  
ICLs: Interstrand Crosslinks  
ICM: Masse cellulaire interne  
ICR : Région de contrôle de l’empreinte  
IGV: Integrative Genome Viewer  
Kb : Kilobases  
KO : Knock-Out  
Mb : Mégabase  
MC-Seq: Methyl-capture Sequencing  
Me : Méthylation  
MEFs: Mouse Embryonic Fibroblasts  
metADN: Méthylation d’ADN  
Mt: Maternel  
Nm : nanomètre  
PE : Endoderme primitif  
PGCs : Cellules Primordiales Germinales  
RRBS : Reduced Representation Bisulfite Sequencing  
SAM : S-Adenosyl méthionine  
Seq: Séquençage  
TE: Trophoctoderme  
TET : Ten-Eleven Translocation Méthylcytosines dioxygénases  
Tet-Off : Tetracycline Off  
TSAF : Trouble du Spectre de l’Alcoolémie Fœtale  
TSS : Site d’initiation de la transcription  
UCSC : University of California Santa Cruz (Browser)  
WGBS: Whole Genome Bisulfite Sequencing

## REMERCIEMENTS

Je souhaite tout d'abord remercier mon directeur de recherche, Serge McGraw, qui m'a guidée et supportée dans mes travaux et qui s'est toujours montré très compréhensif. Merci pour ton aide, tes conseils et ta passion! Et surtout merci pour le climat de labo agréable!

Je voudrais aussi remercier mon co-directeur de recherche, Noël Raynal, et son laboratoire pour l'aide et les conseils reçus au cours des deux dernières années.

Je remercie aussi évidemment tous mes collègues du labo McGraw qui ont chacun contribué à leur façon à mon succès : Lisa (my mate, best mate!), Mélanie (grande héritière du ChIP), Elizabeth (gardienne de nos bébés cellules et traductrice en chef), Anthony (Dr. Tony, super-ninja-bioinformaticien-en-herbe), Hicham et notre précieux alumni (Alba, Anne-Frédérique, Christine). Merci pour vos encouragements et tous les rires qu'on a partagé.

Un immense merci à Maxime et Anthony, nos bio-informaticiens senior et junior, sans qui mon mémoire n'existerait pas. Merci de vous être cassé la tête pour mon projet!

Merci aussi à mes sources de réconfort et de rires quotidiens : le labo Girard, Elodie, Gabrielle et Émilie. Et merci aux gens du 5<sup>e</sup> étage du centre de recherche pour la belle ambiance!

Et les derniers (et non les moindres!) que je souhaite remercier sont les membres de ma famille et mes amis pour leur amour et soutien inconditionnel. Merci de vous être intéressés à mon univers même si c'était très nébuleux pour vous. Et merci tout particulièrement à ma « sisternelle » moitié qui a dû m'endurer dans les moments plus difficiles et qui n'a jamais raté l'occasion de célébrer mes petites victoires. Merci tellement!

## MISE EN CONTEXTE

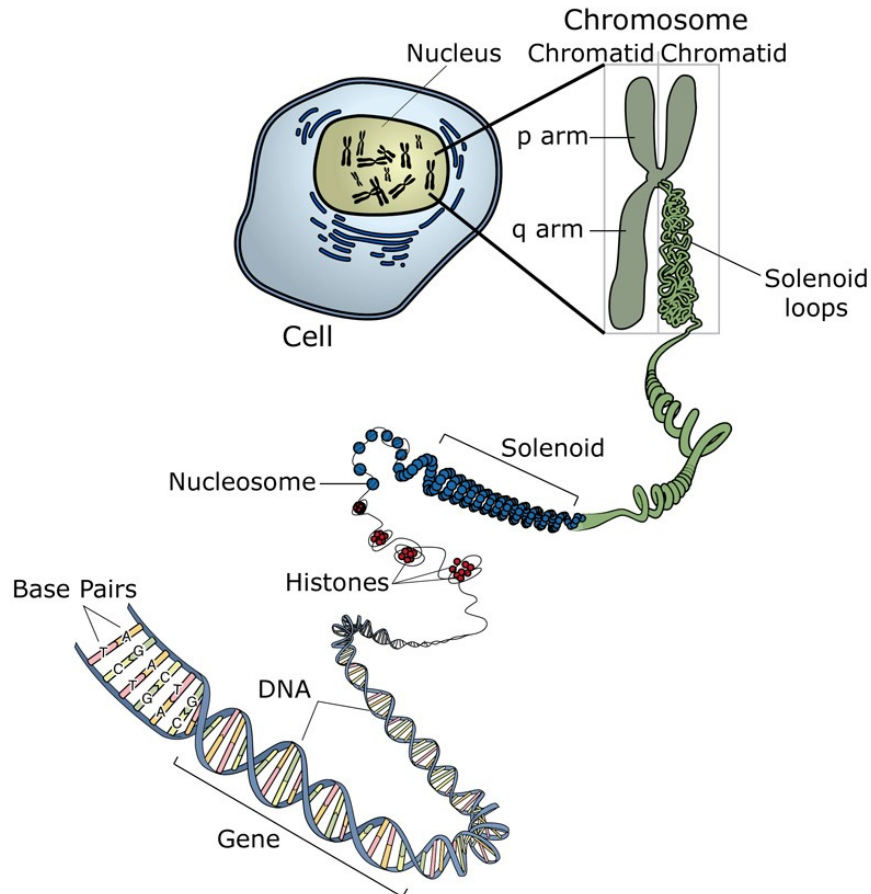
Comment un seul ovule fertilisé peut engendrer un organisme complexe ayant des cellules de phénotypes complètement différents? Cette question a su mettre les chercheurs d'autrefois sur les premières pistes qui allaient mener vers un domaine tout nouveau : l'épigénétique. L'épigénétique a tout d'abord été définie de façon très large jusqu'aux années 1950 comme l'ensemble des événements développementaux de régulation survenant à partir du matériel génétique pour donner un organisme complexe [1]. En 1996, la définition de l'épigénétique a été précisée par Riggs et Porter comme étant des changements héréditaires de la fonctionnalité des gènes qui se transmettent à travers les processus mitotiques et méiotiques et qui n'affectent pas la séquence codante de l'ADN [2]. Aujourd'hui, ces changements épigénétiques sont reconnus pour leur implication dans diverses maladies et phénomènes, telles que les troubles neuro-développementaux, le vieillissement, l'athérosclérose et le cancer, et ils représentent une hypothèse plausible pour expliquer les mécanismes de ségrégation se cachant derrière d'autres maladies à transmission non-Mendélienne [3]. Le rôle de l'épigénétique se révèle particulièrement crucial lors du développement embryonnaire chez l'humain et diverses autres espèces vu que des perturbations épigénétiques ont un impact sur le développement de troubles cognitifs et physiques se révélant dès l'enfance jusqu'à l'âge adulte. La façon dont les modifications épigénétiques ségréguent chez l'embryon est encore peu comprise et la façon dont ces changements se transmettent de cellule en cellule l'est encore moins (revu dans [4]). Il est donc primordial d'approfondir nos connaissances quant aux dérèglements épigénétiques survenant dès les premiers jours embryonnaires et comment ils mènent à des perturbations épigénétiques permanentes afin d'obtenir des réponses aux questions fondamentales concernant l'apparition de troubles et maladies tout au long de la vie d'un individu.

## 1. STRUCTURE DE LA CHROMATINE

### *1.1 Le génome*

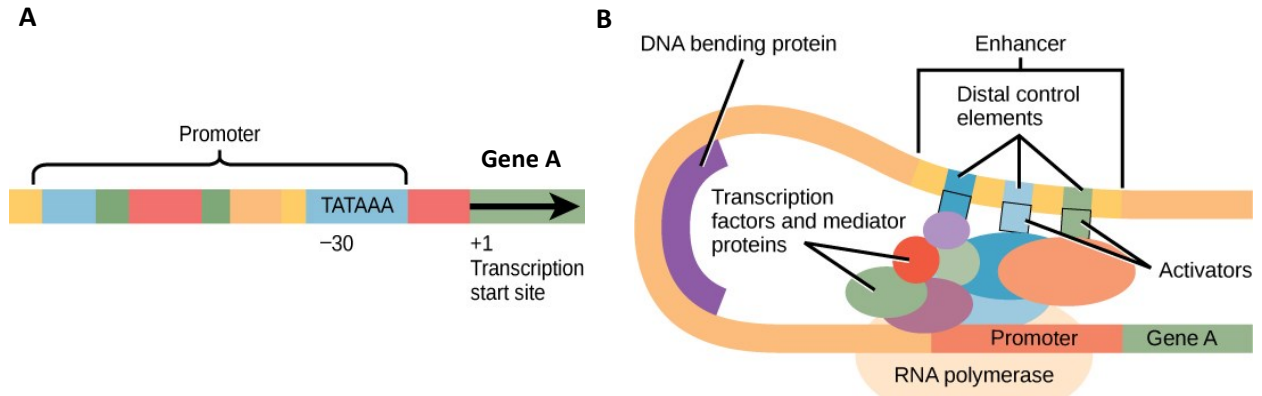
L'unité sur laquelle s'assemble la chromatine est la séquence d'ADN, constituée de paires de bases adénine (A), thymine (T), cytosine (C) et guanine (G). Cette série de bases constitue une séquence précise qui encode des gènes qui seront ou non exprimés en

fonction du repliement de la séquence et de la liaison de molécules (e.g. facteurs de transcription liaison ou modifications épigénétiques) dans la région génique. Cette séquence d'ADN constituée des régions codantes (gènes) et non-codantes (intergéniques) a un brin complémentaire, ce qui fait de l'ADN une double hélice chargée négativement. La chromatine est donc constituée de cette double hélice d'ADN ainsi que de protéines chargées positivement qui s'y lient et possède une structure stable et compactée dans le noyau des cellules eucaryotes (Figure 1). L'ensemble du matériel retrouvé dans le noyau d'une cellule constitue le génome de l'individu chez qui on le retrouve, c'est-à-dire que tout le matériel génétique chez l'homme se résume à 2 paires de 23 chromosomes sur lesquels on retrouve les gènes.



**Figure 1. Structure détaillée des chromosomes dans le noyau de la cellule : des constituants de la chromatine jusqu'aux paires de bases de l'ADN** (Courtesy: National Human Genome Research Institute <https://www.genome.gov/> [5]).

Chaque gène possède un promoteur en amont du site d'initiation de la transcription (TSS) qui permet d'activer la transcription du gène (Figure 2A). La série de facteurs constituant la machinerie de transcription s'installe à cet endroit précis pour débiter le processus. La majorité des gènes possèdent également des régions « enhancers » qui peuvent se trouver à quelques Kilobases (Kb) jusqu'à 1 Mégabase (Mb) en amont ou en aval du gène [6-8]. Ces régions peuvent avoir un impact positif sur l'expression d'un ou plusieurs gènes en se liant à la machinerie de transcription installée au promoteur du ou des gènes via le repliement de l'ADN (Figure 2B). Entre les régions transcriptionnelles, on retrouve donc des régions intergéniques composées entre autres de régions non-codantes et enhancers [9] (Revu dans [10]).



**Figure 2. Localisation et structure d'un promoteur (A) et d'un enhancer (B).** A : Le promoteur est situé en amont du TSS du gène « A » pour permettre le recrutement de la machinerie transcriptionnelle. B : L'enhancer, contenant des éléments de contrôle distaux, peut lier les facteurs recrutés à la région promotrice du gène « A » pour activer les facteurs de transcription et stimuler le processus (adaptée de Open Stax cnx : Gene regulation [11]).

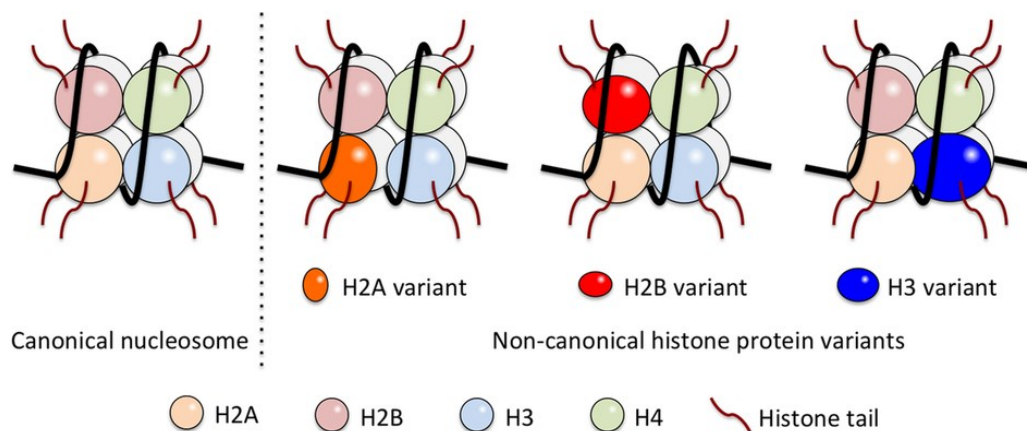
À ce jour, les coordonnées chromosomiques des régions géniques et intergéniques sont bien définies pour beaucoup d'espèces. Pour ce qui est des promoteurs et enhancers, il existe quelques banques de données (davantage pour les promoteurs) qui spécifient les coordonnées chromosomiques chez différentes lignées cellulaires. Les régions promotrices sont un peu plus faciles à identifier comme elles se trouvent à proximité du gène qu'elles régulent et que les régions TSS comportent des motifs spécifiques, c'est-à-dire de courtes parties de séquences qui se répètent dans certaines régions du génome. Pour les enhancers,



la situation est un peu plus complexe comme ceux-ci peuvent être éloignés de leur cible de régulation et comme ils peuvent être associés à plus d'un gène. De plus, pour de nombreux enhancers, les gènes cibles n'ont pas encore été identifiés à ce jour. Pour permettre de mieux déterminer où se trouvent ces régions, il est possible d'utiliser les sites de liaison de certains facteurs de transcription et modifications d'histones.

### 1.2 Nucléosomes et histones

Le premier degré de compaction de l'ADN est le nucléosome. Celui-ci est composé d'un octamère d'histones de types H2A, H2B, H3 et H4 (histones canoniques ou « histone cores »); des protéines interagissant avec l'ADN vu leur caractère basique [12]. L'ADN s'enroule autour de ces octamères, conférant déjà une compaction de sept fois plus serrée à l'ADN, puis l'histone H1 s'ajoute au complexe pour stabiliser la structure et permettre une meilleure compaction, donnant ainsi une fibre de chromatine de 30nm (Figure 3) [13-15].



**Figure 3. Structure du nucléosome : histones et variants d'histones impliqués.** Le nucléosome canonique contient les histones ordinaires, soit H2A/H2B/H3/H4 alors que les nucléosomes non-canoniques contiennent un variant d'histone remplaçant une histone de base [16].

La localisation des histones est dynamique; celles-ci subissent des échanges et sont remplacées par de nouvelles modifications d'histones. Ce phénomène rend la chromatine

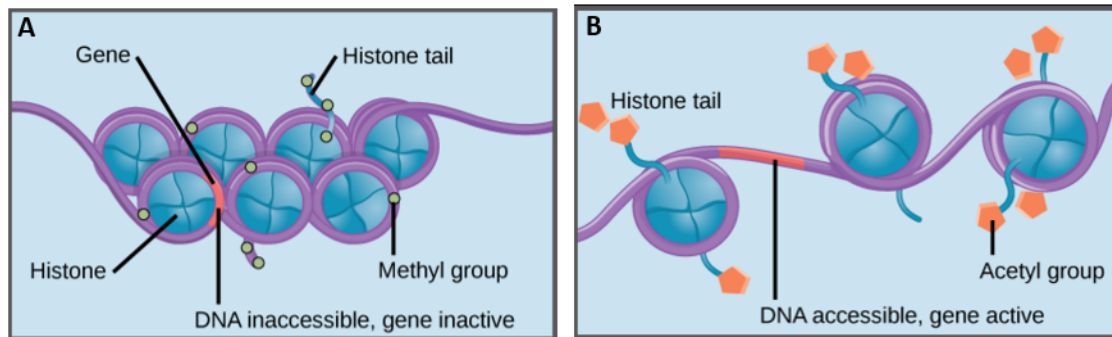
hautement muable et modifie l'accès de protéines, comme par exemple la Polymérase II, à certaines régions génomiques, empêchant ou favorisant la transcription [17-19] (Revu dans [20]). Il existe également plusieurs variants d'histones; ce sont des histones canoniques présentant des changements d'acides aminés ou une addition de domaines. Ces variants d'histones se distinguent des histones canoniques sur quelques points. Par exemple, au contraire des histones canoniques qui sont exprimées à la phase de réplication (S) du cycle cellulaire comme elles sont encodées par des groupes de gènes dépendants de la réplication, les variants d'histones sont encodés par des gènes réplication-indépendants; ces derniers ont des introns et leurs ARNm sont régulés par des queues Poly-A [21]. Les variants d'histones sont donc exprimés tout au long du cycle cellulaire. De plus, l'incorporation ou l'éjection des variants d'histones dans la chromatine est facilitée par des protéines chaperones, tout comme pour les histones, qui lient toutefois spécifiquement certains variants (Revu dans [22]). À ce jour, il existe plusieurs variants pour toutes les histones, excepté H4 (Revu dans [23]). L'histone H2A est celle qui possède le plus grand nombre de variants et ceux-ci ont tous une fonction différente. Par exemple, le variant d'histone H2A.Z est connu pour son anti-corrélation avec la méthylation d'ADN, c'est-à-dire qu'il lie préférentiellement les régions euchromatiques, alors que le variant H2A.X est impliqué dans la réparation de l'ADN suite à un dommage [24, 25].

Contrairement au premier niveau de compaction impliquant les histones, variants d'histones et nucléosomes, les plus hauts niveaux de compaction sont quant à eux moins bien définis. La formation du chromosome métaphasique implique un repliement de 10 000 à 20 000 fois et serait due à la présence de protéines échafauds qui lient des séquences spécifiques de l'ADN [26].

### *1.3 Euchromatine et hétérochromatine*

On retrouve deux types de chromatine : l'euchromatine, forme ouverte qui est majoritairement associée à l'activité génique, et l'hétérochromatine, forme fermée qui est majoritairement associée à la répression génique [27]. Dans l'euchromatine, on retrouve des nucléosomes beaucoup plus espacés sur la séquence d'ADN ce qui fait en sorte que cette dernière est plus accessible pour les facteurs de liaison (Figure 4B). De l'autre côté,

quand l'hétérochromatine est compactée (Figure 4A), les nucléosomes sont beaucoup plus rapprochés, ce qui fait en sorte que les séquences d'ADN se retrouvent majoritairement cachées par ces structures. Les promoteurs des gènes sont donc inaccessibles pour les facteurs de liaison et la transcription est réprimée.

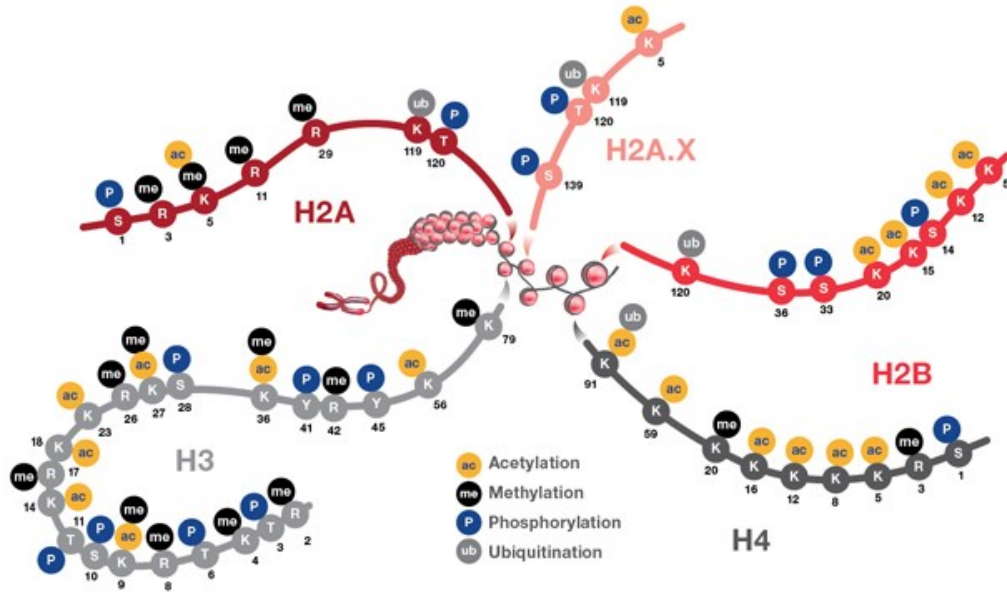


**Figure 4. Représentation schématique de l'hétérochromatine (A) et de l'euchromatine (B).** (Adaptée de Open Oregon : Eukaryotic epigenetic regulation [28])

## 2. MODIFICATIONS ÉPIGÉNÉTIQUES

### 2.1 Modification des histones

La modification post-traductionnelle des histones est catalysée par diverses enzymes spécifiques qui permettent l'ajout de groupements chimiques sur les acides aminés de la queue N-terminale des histones, tel que représenté à la figure 5. Les modifications les mieux décrites sont la phosphorylation, la méthylation (me), l'acétylation (ac), l'ubiquitination et la sumoylation. Ainsi, la nomenclature des modifications d'histones spécifie l'histone modifiée, l'acide aminé où la modification a lieu, la modification en question et le nombre de modifications s'il y a lieu. Par exemple, la modification d'histone H3K4me3 signifie que 3 groupements méthyles (me3) ont été ajoutés à la lysine 4 (K4) de l'histone 3 (H3).



**Figure 5. Sites de modifications post-traductionnelles sur les queues des histones.** Chaque histone ou variant d’histone peut être modifié sur sa queue N-Terminale ou C-terminale (non illustré sur l’image) sur des résidus précis, comme ceux indiqués sur l’image ci-haut. Par exemple, l’histone H3 peut être modifiée par des groupements méthyles à la lysine K4 en N-terminal (© 2018 Thermo Fisher Scientific Inc. Used under permission. [www.thermofisher.com](http://www.thermofisher.com) [29])

Comme l’ajout et le retrait des modifications d’histones se fait selon un processus dynamique, il existe des enzymes qui ajoutent et retirent spécifiquement ces groupements. Par exemple, les histones acétyltransférases (HATs) permettent l’ajout du groupement acétyle et les histones déacétylases (HDACs) le retire. De la même façon, les histones méthyltransférases (HMTs) ajoutent le ou les groupements méthyles et les histones déméthylases (HDMs) les enlèvent. Afin d’assurer leurs propriétés catalytiques, les enzymes peuvent nécessiter la liaison de co-facteurs. Par exemple, les HATs nécessitent un coenzyme (Acetyl-CoA) pour transférer le groupement acétyl sur la protéine. Les modifications d’histones sont impliquées dans la régulation de la transcription, dans la réparation des dommages à l’ADN, dans la recombinaison ainsi que dans la réplication. Au niveau transcriptionnel, les modifications des histones peuvent agir sous différents mécanismes : la modification d’histone elle-même peut modifier la conformation de la chromatine et donc l’activité génique, la modification peut être reconnue par des domaines

de protéines qui favoriseront l'initiation de la transcription, et la modification peut entraver la liaison de facteurs qui permettent la transcription. Parmi les modifications qui permettent le changement de conformation direct de la chromatine, on retrouve l'acétylation et la phosphorylation. Ces modifications réduisent la charge positive des histones, ce qui engendre une diminution des interactions et une décompaction de la chromatine permettant à la machinerie transcriptionnelle de s'installer. L'acétylation survient exclusivement sur les lysines. Ces lysines acétylées peuvent être reconnues et liées spécifiquement par les protéines à bromodomains, ce qui favorise les interactions protéines-protéines (i.e. le recrutement de complexes protéiques) et peut donc avoir un impact sur la transcription. Par exemple, certaines acétyltransférases (ex : PCAF) sont des protéines à bromodomains et peuvent donc lier les lysines acétylées sur la chromatine pour acétyler les lysines avoisinantes et favoriser la transcription (Revu dans [30]). Quant aux modifications qui interfèrent avec la liaison de molécules effectrices (ex : la méthylation des histones), la compatibilité du domaine de liaison du facteur en question avec la modification est ce qui permet la transcription. Par exemple, il existe plusieurs domaines de liaison des lysines méthylées comme les domaines PHD, les chromodomains ou encore les domaines Tudor. Un exemple concret est la liaison des protéines *Inhibitors of Growth* (ING), possédant un domaine PHD, aux modifications d'histones H3K4me<sub>3</sub>; ces protéines permettent le recrutement de molécules qui remodelent la chromatine comme les HATs et HDACs [31, 32] (Revu dans [33]). Parmi les marques d'histones les plus étudiées, H3K9me<sub>3</sub> et H3K27me<sub>3</sub> sont associées avec la répression de l'expression (hétérochromatine) tandis que H3K4me<sub>1</sub> et H3K4me<sub>3</sub> sont associées plutôt à l'activation de la transcription (euchromatine).

### 2.1.1 Répartition des modifications d'histones dans le génome

Comme les différentes modifications d'histones peuvent être associées avec l'activation ou la répression de l'expression, celles-ci peuvent donc être retrouvées en grande proportion au niveau des régions promotrices (e.g., H3K4me<sub>3</sub>, H3K27me<sub>3</sub>) et enhancers (e.g., H3K4me<sub>1</sub>, H3K27ac, H3K27me<sub>3</sub>). On peut aussi retrouver ces modifications d'histones aux corps des gènes (e.g., H3K36me<sub>3</sub>) ou sur les introns. Les régions intergéniques, quant à elles, peuvent être hétérochromatiques, soient pourvues de marques répressives telles que la méthylation d'ADN ou H3K9me<sub>3</sub> (ex : régions hétérochromatiques péri-centriques) ou

dépourvues de marques à certains endroits [34, 35]. Comme mentionné précédemment, les modifications d'histones pouvant constamment être retirées ou ajoutées, la chromatine est donc hautement muable aux régions sujettes à la transcription ou à la promotion de la transcription.

## 2.2 Méthylation de l'ADN et DNMTs

La méthylation de l'ADN est impliquée dans la stabilité et l'intégrité du génome, dans la répression de la transcription, entre autres pour les gènes à empreintes, ainsi que dans l'inactivation du chromosome X. Elle se caractérise par l'ajout de groupements méthyles sur les cytosines des dinucléotides CpG (cytosine-guanine) de la séquence d'ADN; on appelle cette modification 5-méthylcytosine (5mC). Les CpGs se regroupent à certains endroits dans le génome pour former ce qu'on appelle des îlots CpGs (Revu dans [36]). Ces îlots CpGs sont souvent retrouvés aux promoteurs de gènes, certains de ces gènes sont dits « housekeeping », c'est-à-dire qu'ils sont communs à toutes les cellules puisqu'ils participent à des fonctions essentielles [37, 38]. Toutefois, on retrouve aussi des îlots CpGs dans les régions intra et intergéniques et ceux-ci peuvent parfois être associés à de nouveaux promoteurs (i.e. des sites d'initiation de la transcription) (Revu dans [39]). Le génome humain contient ~30 millions de CpGs, dont 60 à 80% sont méthylés et le génome murin en contiendrait ~20 millions [40]. Cette réaction de méthylation est connue pour engendrer une répression de l'expression des gènes soit en inhibant la liaison de certains facteurs de transcription, soit en recrutant les *Methyl CpG binding proteins* (MBPs) qui modifient directement la structure de la chromatine (Revu dans [41]). Outre la modification 5-méthylcytosine (5mC), il existe aussi une forme oxydée appelée 5-hydroxyméthylcytosine (5hmC). Celle-ci est obtenue via l'activité des enzymes TET (*Ten-eleven translocation methylcytosines dioxygenases*) qui est forte lors du développement embryonnaire. La fonction de la modification 5hmC est peu caractérisée, mais elle semble être impliquée dans le recrutement de complexes modificateurs de chromatine [42]. Il existe aussi d'autres variants de méthylation d'ADN (5fC, 5caC, etc.) qui sont impliqués dans l'une ou l'autre des fonctions connues pour la méthylation de l'ADN comme mentionné plus haut.

La réaction de méthylation de l'ADN est catalysée par des enzymes appelées méthyltransférases d'ADN (DNMTs). Il existe trois classes de DNMTs : DNMT1, DNMT2 et DNMT3. Les isoformes de DNMTs ont chacun leur particularité pour la partie N-terminale de la protéine (domaines ATRX, CXXC, ADD, etc.) mais leur domaine catalytique en C-terminal est très similaire (un peu moins pour DNMT3L) [43].

Les DNMT3 se sous-divisent en plusieurs types. Tout d'abord, DNMT3a et b sont des enzymes responsables de la méthylation *de novo*, c'est-à-dire l'ajout de méthylation où il n'y en a sur aucun des 2 brins, des cytosines de l'ADN dans le développement embryonnaire précoce, dans les lignées germinales parentales et dans les cellules somatiques. La protéine DNMT3L, quant à elle, serait plutôt considérée comme un co-facteur pour DNMT3a et b comme elle n'a pas d'activité catalytique; sa nécessité pour le processus de méthylation *de novo* est toujours controversée [44, 45] . Très récemment, la méthyltransférase d'ADN DNMT3c a été découverte et serait impliquée également dans la méthylation *de novo* mais au promoteur de rétrotransposons dans la lignée germinale mâle [46]. Les différents isoformes de DNMT3 sont tous cruciaux pour un développement embryonnaire adéquat. Il a déjà été démontré qu'une inactivation de *Dnmt3a* fait en sorte que les souris meurent quelques semaines suivant leur naissance. Parallèlement, un modèle de souris déficient en *Dnmt3b* causera la mort embryonnaire au jour 9.5, soit à la mi-gestation [47]. Quant à *Dnmt3c*, la délétion chez les animaux entraîne la stérilité chez les mâles [46], ce qui est aussi observé lors d'une délétion de *Dnmt3L*, tandis que chez les femelles l'absence de *Dnmt3L* est létale à mi-gestation [48].

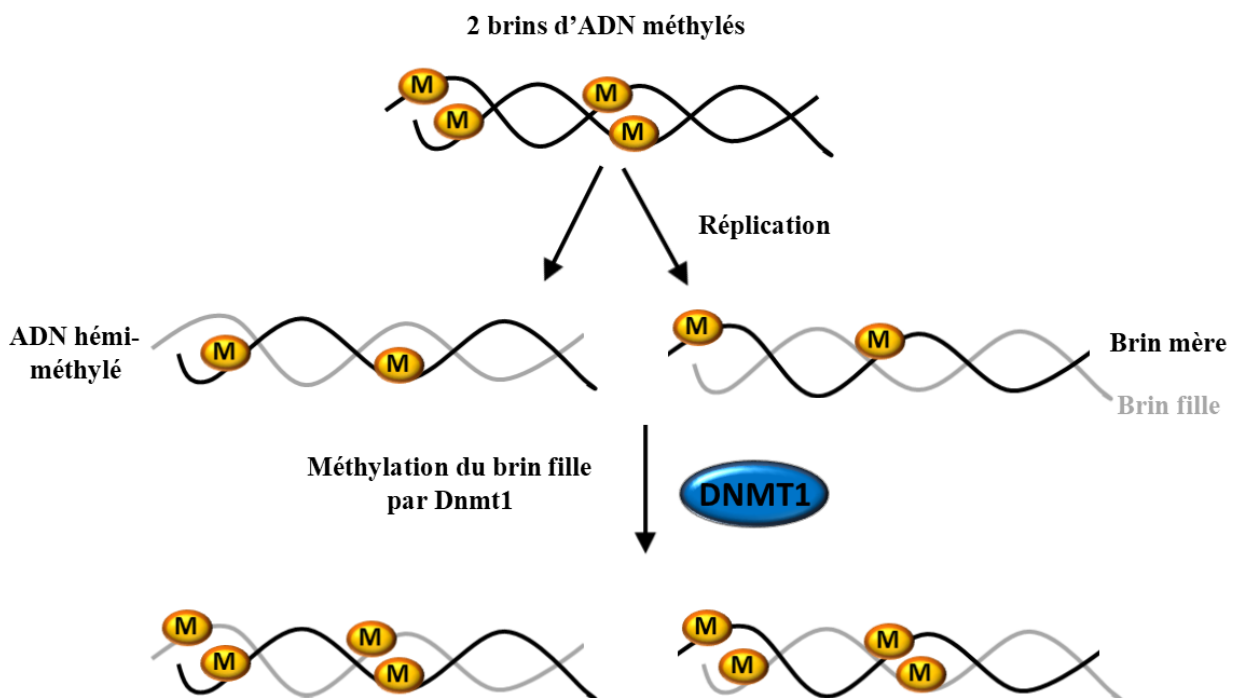
La méthyltransférase d'ADN 2 (DNMT2) est quant à elle beaucoup moins connue. Toutefois, on sait que son activité est plus faible et qu'elle méthyle un résidu spécifique des ARNs de transfert (tARNs), permettant une meilleure reconnaissance des codons environnants et assurant ainsi une bonne synthèse polypeptidique [49, 50] . Malgré sa forte homologie de séquences avec les autres DNMTs, DNMT2 n'a toutefois pas de domaine de régulation et son activité de méthylation d'ADN demeure floue [51, 52]. Contrairement aux isoformes de *Dnmt3*, une déficience en *Dnmt2* donne des souris viables et fertiles [50].

Finalement, DNMT1 est la méthyltransférase d'intérêt pour mon projet de maîtrise puisqu'elle est responsable du maintien de la méthylation dans les cellules embryonnaires

et somatiques [53]. Deux isoformes de DNMT1 existent : DNMT1o est synthétisé chez l'ovocyte seulement alors que DNMT1s est synthétisé chez l'ovocyte, chez l'embryon pré-implantatoire ainsi que dans les cellules somatiques [54]. Pour les pages subséquentes, la nomenclature DNMT1 sera utilisée afin de désigner l'isoforme DNMT1s.

### 2.2.1 DNMT1 et ses interactions

Le gène *Dnmt1* peut assurer son rôle de maintien de la méthylation, car il est spécifiquement transcrit pendant la phase S du cycle cellulaire, soit lors de la réplication de l'ADN [55]. Le domaine catalytique de *Dnmt1* permet ensuite de lier spécifiquement l'ADN hémi-méthylé, c'est-à-dire le brin d'ADN mère méthylé qui sera répliqué, et de méthyler le brin fille (Figure 6) [56].



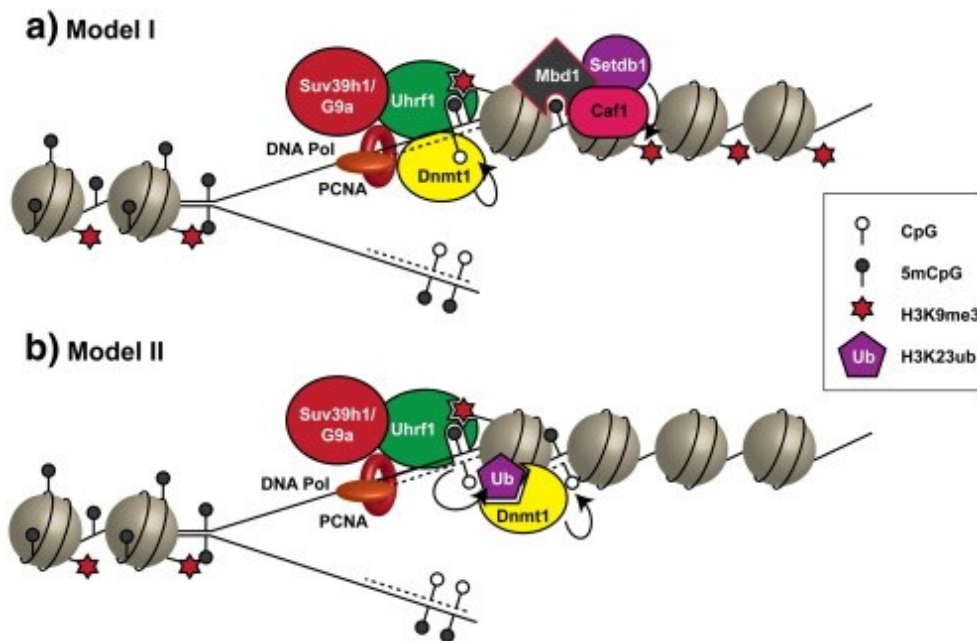
**Figure 6. Apposition de la méthylation sur les brins d'ADN hémi-méthylés par DNMT1 après la réplication.** Les brins d'ADN issus de la réplication sont hémi-méthylés (méthylés sur le brin mère seulement). DNMT1 appose des groupements méthyles spécifiquement sur les brins filles afin de maintenir la méthylation d'ADN.



DNMT1 a donc une fonction de méthylation comme les autres méthyltransférases, mais spécifique à l'ADN hémi-méthylé. En effet, le domaine CXXC en N-terminal de DNMT1 lie les CpGs non méthylés du brin mère et empêche le site actif en C-terminal de la protéine de méthyler un CpG non méthylé sur le brin fille [57]. Le processus est donc auto-régulé pour empêcher de générer des erreurs de méthylation sur les brins filles. Le rôle principal de DNMT1 est donc le maintien de la méthylation d'ADN à la fourche de réplication pour assurer que les marques de méthylation d'ADN seront transmises malgré la réplication. Toutefois, DNMT1 est aussi impliqué dans d'autres fonctions telles que la réparation des brisures double-brins à l'ADN ou des « Interstrand crosslinks » (ICLs) en interagissant indirectement avec des protéines impliquées dans le processus de réparation, par exemple avec la pince 9-1-1 PCNA-like (brisures double-brins) ou avec UHRF1 (ICLs) [58-60]. Il a aussi été démontré que DNMT1 joue un rôle crucial pour l'auto-renouvellement des cellules souches hématopoïétiques (HSCs) ainsi que dans la différenciation des cellules lymphoïdes [61, 62].

Pour permettre son recrutement à la fourche de réplication et son activité de maintien de la méthylation d'ADN, DNMT1 agit de concert avec une multitude de différents facteurs (Revu dans [63]). Entre autres, il est bien établi que PCNA et UHRF1 sont impliqués dans le recrutement de DNMT1. PCNA (*Proliferating Cell Nuclear Antigen*) se lie aux brins d'ADN nouvellement répliqués alors que UHRF1 (*Ubiquitin-like, with PHD and RING finger domains 1*), une protéine grandement impliquée dans la régulation épigénétique, a des domaines d'affinité pour DNMT1, pour l'ADN hémi-méthylé et pour H3K9me3, modification d'histone répressive enrichie aux régions hétérochromatiques constitutives (constituées de répétitions) [64-66]. DNMT1 lie également G9a, une histone méthyltransférase qui catalyse la mono et di méthylation de H3K9. En plus d'interagir avec DNMT1 et UHRF1, G9a s'inscrit dans un complexe qui contient LSH, interagissant avec HDAC1 et HDAC2, suggérant un lien avec les complexes de remodelage d'histones [67]. HDAC1 lie directement DNMT1 alors que HDAC2 se lie au complexe DNMT1/DMP1 qui recrute aussi PCNA (Revu dans [68]). La figure 7 illustre deux modèles possibles d'interaction des facteurs à la fourche de réplication. Dans le modèle I, UHRF1 s'assemble avec PCNA et H3K9me2/3 à la fourche de réplication en liant l'ADN hémi-méthylé; ce complexe recrute ensuite DNMT1 et les méthyltransférases d'H3K9 sur les brins filles pour

assurer le maintien de la méthylation d'ADN. Le complexe MBD1/SETDB1/CAF1 s'occupe ensuite de maintenir le niveau d'H3K9me3 sur ces nouveaux brins d'ADN. Dans le modèle II, il est suggéré que l'ubiquitination d'H3K23 par UHRF1 (en complexe avec PCNA et H3K9me2/3 à la fourche de réplication) est ce qui permet le recrutement de DNMT1 [69, 70]. En somme, les interactions impliquées dans le processus de recrutement de DNMT1 sur la chromatine est hautement complexe et les acteurs et étapes instigatrices du processus demeurent à être clairement définis.



**Figure 7. Facteurs interagissant avec DNMT1 à la fourche de réplication pour maintenir la méthylation d'ADN sur les brins hémi-méthylés (Adapté de [63]).**

### 2.2.2 Déméthylation

Contrairement aux DNMTs qui apposent les marques de méthylation sur l'ADN, la classe d'enzymes méthylcytosines dioxygénases TET sont responsables pour la déméthylation active sur l'ADN. Il existe 3 enzymes TET (1, 2 et 3) qui oxydent les 5mC en d'autres formes. Comme mentionné brièvement plus tôt, ces enzymes sont fortement exprimées dans le développement embryonnaire; *Tet1* pendant la préimplantation alors que *Tet2* et *Tet3* sont plutôt actives suivant la préimplantation jusqu'à la différenciation (Revu dans [71]).

La déméthylation passive, quant à elle, est le phénomène de déméthylation graduel qui survient lorsque l'ADN se réplique et que DNMT1 n'est pas en mesure de maintenir la méthylation d'ADN sur les brins nouvellement synthétisés. La déméthylation passive se trouve donc à être, en quelque sorte, la dilution des marques de méthylation par la réplication. L'application des deux types de déméthylation (active et passive) dans le développement embryonnaire précoce sera traitée à la section 4 [72, 73] (Revu dans [74]).

### 2.2.3 Gènes empreintes et « imprinted-like »

Les gènes empreintes sont des gènes dont l'expression est propre à un seul allèle, maternel ou paternel, puisqu'une des deux copies est réprimée par des marques de méthylation d'ADN dans la région régulatrice du gène ou par la liaison d'ARNs non-codants (ncRNAs) ou d'enzymes modifiant les histones [75]. Chaque groupe (ou « cluster ») de gènes empreintes est régulé par un même élément appelé région de contrôle de l'empreinte (ICR) ou centre d'empreinte (IC). Ces ICRs sont en vaste majorité des régions différenciellement méthylées (DMRs) (parfois des ncRNAs). En fait, on appelle les DMRs contrôlant spécifiquement des gènes empreintes des ICRs. Ces ICRs sont méthylés *de novo* et de façon différentielle sur les allèles dans les cellules germinales lors de la gamétogénèse. La méthylation de l'ADN ainsi installée contrôle l'expression monoallélique de ces gènes à empreinte qui acquièrent une méthylation allèle-spécifique uniquement dans la lignée germinale mâle ou femelle [76]. Dans d'autres cas, comme au locus *Kcnq1*, l'allèle maternel est réprimé par la méthylation de l'ICR alors que l'allèle paternel est réprimé via l'expression du ncRNA *Kcnq1ot1*. On remarque également un enrichissement de la modification d'histone H3K27me3 et de l'enzyme modifiant les histones, PRC2, aux régions chevauchant le « cluster » *Kcnq1*, suggérant qu'un remodelage allèle-spécifique des modifications d'histones, via l'interaction avec leurs régulateurs, peut être impliqué dans le phénomène d'empreinte [75, 77]. Il existe à ce jour plus de 100 gènes empreintes chez la souris, plus de 70 chez l'humain et environ 50 en commun entre les deux espèces [78, 79]. Par exemple, chez l'humain et la souris, le gène empreinte *H19* est normalement méthylé sur l'allèle paternel et le gène *Igf2* est, lui, méthylé sur l'allèle maternel. La modulation de l'expression d'*H19* et *Igf2* est enclenchée par une région enhancer tissu-dépendante, située en aval de *H19*, qui interagit avec les promoteurs *H19* et *Igf2* via la

liaison du facteur de transcription CTCF, conférant ainsi à *H19* une certaine activité de régulateur transcriptionnel [80]. Si une perte de méthylation (hypométhylation) survient à l'ICR paternel contrôlant l'expression de *H19*, la région enhancer pourra s'y lier via CTCF, empêchant la liaison au promoteur *Igf2*. L'expression d'*Igf2*, gène qui stimule la croissance cellulaire, sera donc inhibée, ce qui mènera au développement du syndrome Silver-Russel [81-83]. Un mécanisme similaire survient pour les autres désordres à empreinte tels que les syndromes de Prader-Willi, Angelman et Beckwith-Wiedemann mais à des loci différents.

L'absence de l'activité de *Dnmt1* est associée à une perte de méthylation sur les gènes à empreinte et mène à un génome dépourvu en méthylation aux DMRs [84, 85]. Il est aussi connu que la méthylation d'H3K4 (me1/me2) protège certains loci de la méthylation d'ADN *de novo* pendant l'oogenèse, démontrant une implication des modifications d'histones pour l'établissement d'empreintes maternelles [86]. Ainsi, en plus de la méthylation d'ADN, les ICRs possèdent aussi une distribution particulière de modifications d'histones allèle-spécifique qui doit être maintenue pour assurer la stabilité de l'expression monoallélique conférée par les cellules germinales [87].

Les gènes « imprinted-like », quant à eux, sont des gènes qui semblent agir selon un patron similaire aux gènes empreintes, mais sans en avoir toutes les caractéristiques. Les gènes « imprinted-like » peuvent être bialléliques ou monoalléliques en fonction du tissu où les gènes sont exprimés. Tout comme les gènes empreintes, leur profil de méthylation d'ADN doit être maintenu par une activité DNMT1-dépendante pendant la vague de reprogrammation épigénétique du jeune embryon [88]. À ce jour, les gènes « imprinted-like » demeurent peu étudiés et l'implication du maintien de leurs profils de méthylation pendant la vague de reprogrammation demeure peu caractérisée. Les travaux de McGraw et al. ont démontré qu'une perte temporaire du maintien de la méthylation d'ADN par DNMT1 dans un modèle de cellules souches embryonnaires de souris entraîne une perte de méthylation d'ADN non seulement aux régions empreintes, mais aussi aux régions dites « imprinted-like » [88]. Parmi ces gènes « imprinted-like », on retrouve la famille de gènes *X-linked lymphocyte regulated (Xlr)*. Ces gènes murins (*Xlr3a/b/c*, *Xlr4a/b/c*, *Xlr5a/b/c*) sont situés sur le chromosome X et seraient exprimés lors de la maturation des lymphocytes

B, dans les ovocytes et dans les spermatocytes primaires. Quant à l'expression de ces gènes, il a été démontré que *Xlr3b* et *Xlr4b/c* auraient une empreinte paternelle dans le cerveau alors que les autres isoformes auraient une expression biallélique [89-96]. Nous n'en savons toutefois pas plus sur leurs caractéristiques « imprinted-like » ni sur leur rôle à ce jour.

### 3. INTERACTION ENTRE LES MODIFICATIONS ÉPIGÉNÉTIQUES

Malgré le fait que chacune des modifications épigénétiques puisse agir indépendamment dans la régulation de mécanismes cellulaires, l'action conjointe, l'interdépendance ou l'antagonisme de différentes modifications épigénétiques peut également être nécessaire pour affiner certains de ces mécanismes de régulation. Dans le cas des histones, des interactions entre les modifications sur un même résidu ou une même histone (régulation en cis), sur différentes histones à l'intérieur d'un même nucléosome (régulation en trans) et même entre les nucléosomes (« nucleosome crosstalk ») peuvent avoir lieu [33, 97]. Les modifications des histones partagent aussi de nombreux liens fonctionnels avec la méthylation de l'ADN. Certaines modifications comme H3K27me3 sont compatibles et chevauchent la méthylation de l'ADN dans certaines circonstances, tandis que d'autres comme H3K4me3 semblent empêcher la méthylation d'ADN dans d'autres contextes [98].

#### *3.1 Modifications épigénétiques et expression génique : l'oeuf ou la poule?*

En effet, la présence de méthylation a longtemps été associée uniquement à une répression de l'expression des gènes. Toutefois, on sait maintenant que pour une petite proportion du génome les effets de la méthylation d'ADN sur l'expression ne sont pas si simples. En effet, il a été démontré que près de 30% des îlots CpGs méthylés peuvent activer la transcription (Revu dans [41]). Entre autres, dans les cellules souches ou les cellules somatiques, les promoteurs ayant un contenu CpG pauvre mais méthylé peuvent tout de même permettre la liaison de Pol II et l'activation de la transcription. Ainsi, la méthylation est répressive lorsque le contenu en CpG est riche (îlots CpG) et que le niveau de 5mC est élevé au promoteur d'un gène donné [99, 100].

En ce qui a trait aux modifications d'histones, chaque marque a une relation particulière avec la transcription des gènes. Par exemple, la modification H3K4me3 est retrouvée dans les régions du génome qui sont actives (i.e. hétérochromatine), particulièrement aux promoteurs et corps des gènes transcrits. À l'opposé, la modification H3K27me3 est une marque répressive retrouvée aux promoteurs et corps de gènes qui sont transcriptionnellement réprimés. D'autres marques telles que H3K4me1 et H3K27ac sont généralement retrouvées dans les régions enhancers et un peu aux promoteurs des gènes et permettent de moduler l'expression des gènes de façon indirecte. Aussi, certains domaines des DNMTs et autres enzymes modifiant les histones ont la capacité de se lier et d'interagir dans des complexes pour réguler la transcription [101]. Ainsi, la régulation de l'expression génique par les modifications épigénétiques s'avère un processus bien complexe vu le nombre d'acteurs ou d'intermédiaires et vu que certaines associations sont indirectes.

Dans les dernières années, la question de l'oeuf ou la poule, c'est-à-dire qu'elle modification épigénétique intervient en premier, a été soulevée. Une première hypothèse est que la méthylation intervient en premier lieu, suivie par l'apposition des modifications d'histones, ce qui entraîne l'expression ou la répression. Par exemple, certains facteurs de transcription sont maintenant reconnus comme des lecteurs de la méthylation d'ADN, c'est-à-dire que la méthylation d'ADN contribue à l'installation de la machinerie transcriptionnelle dans certaines régions. Toutefois, une bonne majorité des facteurs de transcription ne s'associent pas aux marques de méthylation [102]. La méthylation d'ADN semble avoir un effet inhibiteur sur la modification de l'histone H3K4, ce qui permet de maintenir l'état répressif [103]. De plus, les marques de méthylation sont liées par des MBPs comme MeCP2 qui, elles, permettent de remodeler la chromatine environnante en recrutant ou modifiant les histones appropriées [104].

Une deuxième hypothèse suppose aussi que les histones orchestrent le processus de régulation de l'expression. En effet, la présence de certaines modifications d'histones à elles seules dans les régions promotrices, enhancers ou encore dans le corps d'un gène peut induire des changements d'expression en recrutant la machinerie transcriptionnelle via la *Bromodomain-containing Protein 4* (BRD4) [105]. Des études supposent même que les modifications d'histones peuvent jouer un rôle dans l'apposition de la méthylation de

l'ADN, ce qui laisse présumer un rôle des histones en amont du processus. Par exemple, le domaine ADD des enzymes DNMT3a/L reconnaît les lysines 4 des histones 3 (H3K4) non modifiées, ce qui permet aux DNMTs de catalyser la réaction de méthylation uniquement aux régions contenant ces histones non-modifiées, alors que la modification d'H3K4 inhibe ce processus [106-109].

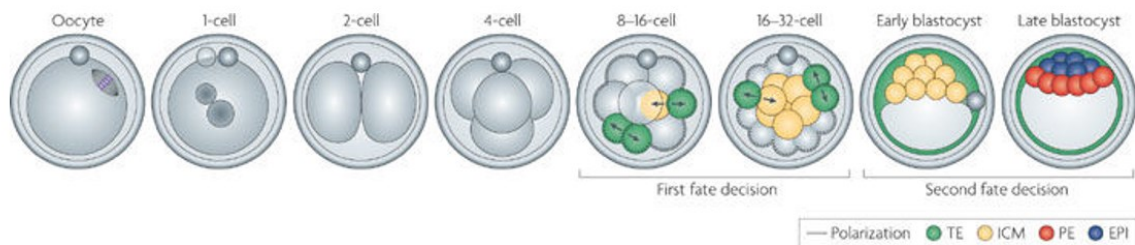
Une troisième hypothèse semble regrouper les deux hypothèses précédentes et nuancer le tout. Par exemple, l'équipe de Du et al. soutient que le processus de méthylation d'ADN peut être dépendant des modifications d'histones dans certains contextes pour la régulation de l'expression et que le repliement de la chromatine peut mener à du feedback positif, renforçant l'idée d'une coopération entre la méthylation et les modifications d'histones [110]. De plus en plus d'auteurs s'entendent sur le fait que le processus de régulation épigénétique est trop complexe pour être décrit par un simple modèle de cause à effet et que des recherches extensives sont nécessaires pour la compréhension de ces mécanismes [33, 36, 111-113].

## 4. REPROGRAMMATION ÉPIGÉNÉTIQUE DANS LE DÉVELOPPEMENT EMBRYONNAIRE PRÉCOCE

### *4.1 Embryogenèse et développement embryonnaire précoce (préimplantation) murin*

Le développement embryonnaire débute lorsque la fertilisation d'un ovule haploïde par un spermatozoïde haploïde se produit pour former un zygote diploïde; c'est le début de la phase préimplantatoire. L'ovule fertilisé est totipotent, c'est-à-dire qu'il a le potentiel de se diviser et d'évoluer pour former tous les types cellulaires du corps humain (Revu dans [114]). Ce zygote 1-cellule entre en processus actif de divisions cellulaires par mitose; celui-ci subit d'abord une division cellulaire symétrique en 2 cellules (appelées blastomères). Les divisions progressent de façon exponentielle et les blastomères deviennent de plus en plus petits. Au stade morula (8-16 cellules), les blastomères se compactent, rendant la différenciation plus difficile, puis une cavité se crée, ce qui mène à la formation du blastocyste. À ce stade, c'est-à-dire après 3.5 jours embryonnaires (E3.5) chez la souris, l'embryon possède une couche de cellules externes appelée trophectoderme (TE) et des cellules de la masse interne (ICM). Les cellules de l'ICM sont pluripotentes,

c'est-à-dire qu'elles sont des cellules souches pouvant se différencier en la majorité des types cellulaires du corps (endoderme, mésoderme ou ectoderme) sans toutefois pouvoir générer un organisme en entier. Si le blastocyste réussit à s'implanter dans l'utérus, c'est-à-dire si les hormones sécrétées sont adéquates pour permettre la réceptivité de l'utérus, le développement de l'embryon se poursuivra en phase postimplantatoire où les cellules deviendront multipotentes (pouvant donner quelques types cellulaires) puis différenciées en un type cellulaire fixe (Revu dans [114]). La 2<sup>e</sup> décision de devenir cellulaire permettra à ce moment de former l'endoderme primitif (PE) et l'épiblaste (EPI) à partir des cellules en surface ou en profondeur de l'ICM du blastocyste, ce qui donnera naissance respectivement au tissu extra-embryonnaire secondaire et au tissu fœtal (Figure 8). L'organogenèse et la gamétogenèse seront ensuite initiées et poursuivies tout au long du développement post-natal de l'organisme jusqu'à sa maturité sexuelle (Revu dans [115] [116]).



**Figure 8. Représentation des étapes du développement embryonnaire précoce chez la souris.** Une fois l'ovule fertilisé, l'embryon se divise jusqu'à ce que le destin cellulaire soit attribué et que les différents tissus forment le blastocyste qui s'implantera dans l'utérus (Trophoblaste TE, Cellules de la masse interne ICM, endoderme primitif PE, Épiblaste EPI) (Adapté de [117]).

#### 4.2 Reprogrammation des patrons de méthylation dans la préimplantation

Au cours du développement embryonnaire, il est crucial que les patrons de méthylation soient reprogrammés afin d'établir des patrons de méthylation différentielle et de permettre aux cellules de passer d'un état pluripotent (grand potentiel de différenciation, ex : cellule souche) à différencié (destin cellulaire fixe, ex : neurone). Deux événements épigénétiques cruciaux sont responsables du bon développement de l'embryon. Le premier événement survient lors du développement des gonades chez l'embryon (jours E8.5 à E10.5) où

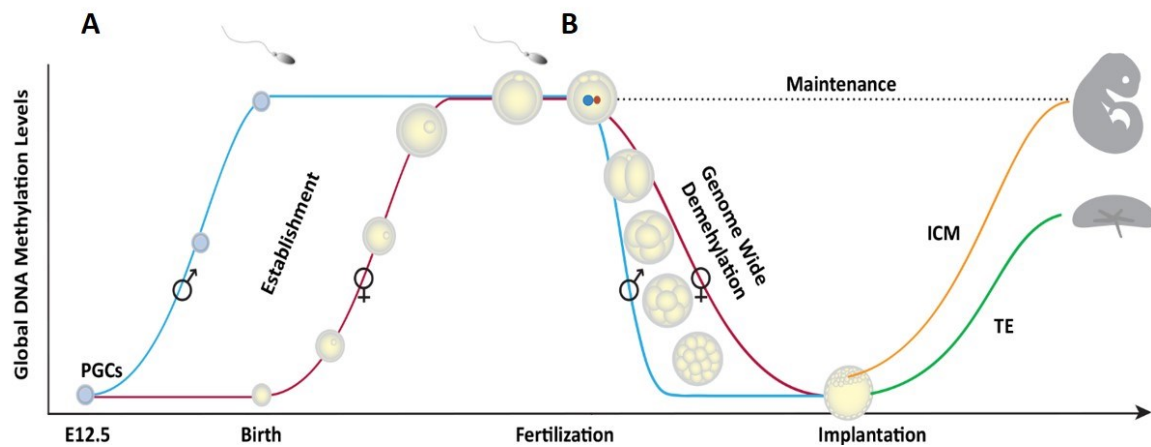


quelques cellules d'épiblaste (dérivées de la masse interne) se différencient en cellules primordiales germinales (PGCs) et migrent vers la crête génitale pour se diviser par mitose et former les gonades et les gamètes. Lors de cette migration, les PGCs subissent une première vague de déméthylation active sur tout le génome afin d'effacer les patrons de méthylation et d'établir de nouveaux patrons grâce à DNMT3a/b au jour E12,5 pour les gonades mâles et suivant la naissance pour les gonades femelles (Figure 9A) [115] (Revu dans [118]). C'est donc à cette première vague que les empreintes de méthylation qui pourront être transmises à un futur embryon sont établies. À partir de leur établissement, il est ensuite crucial que cette méthylation soit maintenue pour le développement adéquat de l'épigénome fœtal, c'est-à-dire tout au long de la deuxième vague de reprogrammation [119] [74].

Le deuxième événement épigénétique survient lors des premiers jours embryonnaires suivant la fertilisation de l'ovule par le spermatozoïde (Figure 9B). Une déméthylation globale de l'ADN survient avant même la première division cellulaire du zygote dans les heures suivant la fertilisation, d'abord pour le génome paternel puis suivi de près par le génome maternel. Le génome maternel est déméthylé de façon passive, c'est-à-dire que la méthylation est perdue graduellement et lentement au fil des divisions cellulaires. Pour ce qui est de la méthylation du génome paternel, elle est perdue rapidement de façon active, soit par l'activité de l'enzyme TET3 qui oxyde la forme 5mC. Cette vague de déméthylation autant paternelle que maternelle se produit pour l'ensemble du génome à l'exception des régions DMRs associées aux gènes empreintes où la méthylation est maintenue par DNMT1 (DNMT1s et DNMT1o) pendant cette période. Les régions se retrouvent donc protégées de la déméthylation puisque si la méthylation est perdue à ses régions, elle ne sera pas ré-établie par les méthyltransférases *de novo*. L'établissement adéquat des marques de méthylation suite à la reprogrammation épigénétique post-fertilisation est essentielle pour le bon déroulement de la différenciation subséquente des cellules qui formeront le fœtus (Revu dans [120]) [74].

La déméthylation persiste pendant quelques divisions cellulaires, soit du zygote 1-cellule (après la fertilisation) jusqu'au stade blastocyste (implantation), puis DNMT3a et

DNMT3b procèdent à la méthylation *de novo* afin d'assurer le rétablissement complet de la méthylation dans le génome (Figure 9B) [115].

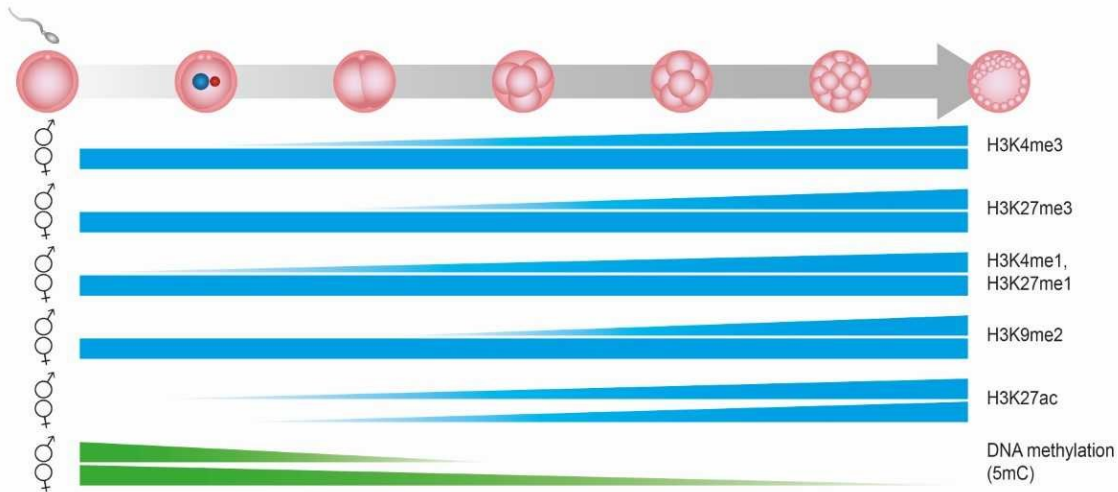


**Figure 9. Schéma de la vague de reprogrammation épigénétique survenant dans le développement embryonnaire précoce.** Les patrons de méthylation d'ADN sont initialement établis dans les cellules primordiales germinales (PGCs) puis, suivant la fertilisation, une vague de déméthylation et de reméthylation globale survient, excepté pour les gènes dont la méthylation d'ADN est maintenue par DNMT1 à cette période (Adapté de [121])

#### 4.3 Distribution des modifications d'histones dans la préimplantation

En ce qui concerne les modifications d'histones suivant la fertilisation, un remodelage est aussi observé. Comme une dynamique associe la méthylation d'ADN et les modifications d'histones, la perte de méthylation de la vague de reprogrammation survenant à cette période entraîne d'autres changements au passage. Le génome maternel semble beaucoup moins affecté que le génome paternel par ce remodelage. Seule H3K27ac augmente graduellement jusqu'au stade blastocyste alors que la méthylation d'ADN s'estompe sur le génome maternel, à l'exception des régions empreintes et « imprinted-like » qui sont protégées de la déméthylation de l'ADN. Pour ce qui est des autres marques les plus caractérisées (e.g. H3K4me3, H3K27me3, H3K9me2 etc.) une augmentation graduelle de leur niveau sur le génome paternel est observée entre les stades 1-cellule et 4-cellules jusqu'au stade blastocyste (Figure 10). Pour ce qui est de la distribution des modifications d'histones aux régions empreintes, le niveau demeure relativement stable tout au long du développement embryonnaire. Par exemple, le niveau des modifications H3K9me2,

H3K27me3 et H4K12ac varie peu au locus *KvDMR1* (impliqué dans le syndrome Silver-Russel) entre les stades 4-cellules, morula et blastocyste [122].



**Figure 10. Distribution temporelle globale de la méthylation d'ADN et de modifications d'histones tout au long du développement embryonnaire précoce de la fertilisation à l'implantation dans l'utérus.** Les barres bleues et vertes représentent les niveaux relatifs des modifications d'histones spécifiques et de la méthylation d'ADN (Adapté de [123])

## 5. PERTURBATION DU PROGRAMME ÉPIGÉNÉTIQUE LORS DE LA PRÉIMPLANTATION

La vague de déméthylation et de reméthylation de l'ADN du génome entier pendant la phase préimplantatoire de l'embryogenèse est essentielle pour le développement des mammifères et pour l'intégrité génomique. En temps normal, les régions différenciellement méthylées (DMRs), incluant les gènes empreintes et « imprinted-like », échappent toutefois à cette vague de reprogrammation épigénétique grâce à une méthylation soutenue par DNMT1. Ce maintien de la méthylation d'ADN est crucial puisque ces régions peuvent être méthylées *de novo* par DNMT3a/b uniquement lors de la vague de reprogrammation épigénétique survenant chez les cellules primordiales germinales (PGCs) qui constituent les gamètes. Une fois la formation des gamètes complétée, une perte de méthylation aux

régions empreintes et « imprinted-like » persiste dans le temps. Une absence du maintien de méthylation d'ADN adéquate sur ces DMRs peut donc engendrer des dérèglements épigénétiques héréditaires, c'est-à-dire des erreurs transmises de cellule en cellule, pouvant mener à diverses anomalies telles que les désordres à empreintes comme mentionné plus tôt. La permanence de ces modifications devra être adressée à plus long-terme pour la compréhension de ce mécanisme de dérégulation.

### *5.1 Erreurs épigénétiques développementales*

Une perturbation du programme épigénétique peut avoir diverses causes. Tout facteur néfaste de l'environnement maternel peut potentiellement mener à des troubles dans le programme développemental du fœtus. Par exemple, une exposition aigüe (de type « binge ») à l'alcool lorsque l'embryon n'est pas encore implanté peut être associée avec le développement de Troubles du Spectre de l'Alcoolémie Foetale (TSAF) qui atteint de 1 à 5% des nouveaux-nés. En effet, le sous-métabolite de l'alcool, l'acétaldéhyde, est une substance qui interfère directement avec le cycle de l'acide folique et avec le donneur universel de groupements méthyles (SAM, S-adenosyl methionine), ce qui affecte directement les processus d'établissement des marques épigénétiques [124, 125] (Revu dans [123]). Les DNMTs possèdent un domaine de liaison pour leur cofacteur SAM. Ainsi, un facteur néfaste peut nuire à l'action des DNMTs en affectant directement la quantité de donneurs de méthyl disponible [126]. Outre l'exposition à l'alcool, beaucoup d'autres facteurs ont été associés avec des changements de la méthylation d'ADN et sont en cours d'investigation pour mieux décrire le mécanisme qui survient. Parmi ces facteurs, on retrouve l'exposition à la cigarette (primaire et secondaire), certains médicaments ou traitements, un niveau de stress élevé, une diète mal équilibrée, etc. [127-130]. Une autre cause sous investigation à l'heure actuelle est l'utilisation des technologies de reproduction assistées qui démontre des changements significatifs au niveau de la méthylation d'ADN chez l'embryon qui se répercutent en désordres divers à l'âge adulte. En effet, la culture d'embryons (milieu, pH etc.) et la fertilisation *in vitro* sont des stressseurs préimplantatoires et ces derniers ont été liés à des effets néfastes à long-terme chez des modèles animaux et chez des embryons humain [131-135]. Ainsi, comme le début de la grossesse représente une fenêtre de vulnérabilité pour l'embryon, des facteurs néfastes divers dans

l'environnement maternel peuvent être à l'origine de défauts épigénétiques qui peuvent avoir des répercussions à court ou long-terme.

### 5.2 Désordres engendrés par des erreurs épigénétiques dans le programme foetal

À ce jour, des dérèglements épigénétiques héréditaires impliquant DNMT1 ont été associés avec des désordres tels que le cancer, le vieillissement prématuré, divers troubles neuro-développementaux ainsi que les désordres d'empreinte. Une absence temporaire de l'activité de DNMT1 dès les premiers jours embryonnaires peut être à l'origine d'absence de méthylation d'ADN ou d'hypométhylation à certaines régions, telles que celles qui contrôlent l'expression des gènes à empreintes. L'utilisation des technologies de reproduction assistées serait également responsable pour l'augmentation de l'incidence sporadique de ces désordres sévères chez l'enfant [121, 136]. Pour l'ensemble de ces désordres, les changements de méthylation sont assez bien caractérisés. Toutefois, on en sait moins concernant le remodelage des modifications d'histones et leur implication dans ces désordres. Quelques études se penchent vers cette direction. Une étude récente sur le syndrome d'Angelman suggère qu'une perte de fonction de la copie maternelle du gène *Ube3a* chez la souris mène à une diminution de l'acétylation des histones H3 et H4 ainsi qu'à une augmentation de l'activité des HDAC1/2 [137]. Quant au syndrome Beckwith-Wiedemann, des corrélations entre la perte de méthylation d'ADN et la modification de H3K9 au DMR *Lit1* ont été observées et suggèrent une synergie entre la méthylation et les modifications d'histones dans le mécanisme derrière ce désordre à empreinte [138]. Pour ce qui est du cancer et du vieillissement prématuré, l'accumulation d'erreurs de méthylation d'ADN (hypométhylation) et de changements dans la distribution des modifications d'histones engendrée par l'absence de DNMT1 est associée à un taux plus élevé de mutations, un arrêt du cycle cellulaire, à l'apoptose, et à une grande instabilité chromosomique pouvant être à l'origine de beaucoup des cas de cancers [139-142]. Cette hypométhylation et cette absence de DNMT1 peuvent avoir un impact sur l'expression d'oncogènes ou de gènes suppresseurs de tumeurs ainsi que sur la réparation des dommages à l'ADN [143]. De plus, comme les rétrotransposons (LTRs et LINEs) sont réprimés par la méthylation d'ADN dans les cellules somatiques et germinales afin d'empêcher l'accumulation de mutations, une perte de méthylation à ces régions engendrera une grande instabilité chromosomique [144]. Quant à la génétique, une mutation dans le gène *Dnmt1*

est responsable pour deux désordres neurologiques sévères, soit la Neuropathie héréditaire sensorielle et autonome de type 1E (HSAN1E) et l'Ataxie cérébelleuse autosomique dominante de type 1 - surdit  et narcolepsie (ADCA-DN) [43, 145-147]. Pour les patients HSAN1E, une  tude propose que la perte de m thylation d'ADN engendr e par une mutation dans le domaine de clivage des s quences de *Dnmt1* survient dans des r gions o  on retrouve les modifications d'histones r pressives H3K27me3 et H3K9me3 [148].

## 6. MOD LISATION D'UNE PERTE D'EXPRESSION DE *DNMT1* DANS LES CELLULES EMBRYONNAIRES: EMBRYONS VS CELLULES SOUCHES EMBRYONNAIRES

Sachant qu'un stress de l'environnement maternel peut alt rer l'activit  de DNMT1 et que cette alt ration peut avoir des impacts consid rables sur le d veloppement du f etus, l'utilisation de mod les induisant la perte de *Dnmt1* permet de pr cis ment comprendre le r le fonctionnel de cet acteur dans l' tablissement du programme  pig n tique f tal dans la pr implantation. Puisque la famille de g nes *Dnmt* est essentielle pour le bon d veloppement de l'embryon, les options de mod les d' tude viables sont donc limit es. En effet, comme une mutation homozygote dans un des g nes *Dnmt* est l thale au cours de la gestation (mi-gestation pour la majorit  d'entre eux chez la souris et l'humain), la possibilit  de g n rer un mod le animal au stade embryonnaire vivant et d ficient en *Dnmts* est  limin e [47, 149].

L'utilisation d'inhibiteurs de m thylation d'ADN tels que le 5-azacitidine et la d c tabine n'est pas id ale non plus pour des  tudes visant   inhiber sp cifiquement et directement *Dnmt1* pour cibler les r gions *Dnmt1*-d pendantes pour le maintien de la m thylation. Ces inhibiteurs ciblent tout le g nome pour la d m thylation en bloquant le site actif de toutes les m thyltransf rases et non seulement DNMT1 [150, 151]. Aucun inhibiteur sp cifique de *Dnmt1* n'existe   ce jour.

### 6.1 Mod le murin (*embryons*)

Comme mentionn  plus haut, une mutation homozygote entra nant une perte d'activit  de la prot ine DNMT1 engendre un d lai de d veloppement, donnant des embryons beaucoup

plus petits que la normale, ce qui cause la mort embryonnaire à mi-gestation. Une hypométhylation a été observée dans les tissus embryonnaires déficients en *Dnmt1*. Une mutation hétérozygote de *Dnmt1* chez la souris donne, quant à elle, un phénotype normal, et ce, même si une hypométhylation globale du génome est observée [149, 152]. Toutefois, la déficience en *Dnmt1* semblerait ne pas affecter les tissus extraembryonnaires [153, 154].

Chez un modèle « Knock-Out » (KO) spécifique pour *Dnmt1 $\alpha$* , la forme spécifique à l'ovocyte exprimée uniquement au stade 8-cellules, les embryons subissent également une perte de méthylation aux régions empreintes et meurent avant la naissance. Ainsi, même une absence de DNMT1 à un stade spécifique affecte le maintien des patrons de méthylation d'ADN et mène à des désordres chez l'embryon [84, 155].

### 6.2 Cellules souches

Afin d'étudier le rôle de DNMT1 dans une cellule embryonnaire pluripotente, le modèle le plus simple demeure donc d'utiliser les cellules souches embryonnaires (ES) comme celles-ci peuvent survivre en absence des DNMTs jusqu'à l'induction de leur différenciation, où elles mourront par apoptose. En effet, des cellules ES déficientes en *Dnmts* ne démontrent pas de changement sur leur croissance ou leur morphologie, mais elles ont un niveau de méthylation d'ADN plus faible [149]. Les cellules ES peuvent s'auto-renouveler et proliférer malgré cette hypométhylation d'ADN, contrairement aux cellules somatiques, ce qui leur confère un avantage important pour nos travaux [156, 157]. Les cellules souches embryonnaires sont issues des cellules de la masse interne (ICM) de l'embryon en formation, au stade blastocyste, et sont maintenues en culture avec des facteurs qui permettent le maintien de leur état de pluripotence. Cet état de pluripotence est caractérisé par l'expression des marqueurs *Oct4*, *Sox2* et *Nanog* spécifiques aux cellules embryonnaires non différenciées (Revu dans [158] [159]). Plusieurs études jusqu'à présent ont utilisé des modèles divers de cellules souches murins et humains pour étudier les mécanismes impliqués dans divers défauts de développement car elles sont facilement manipulables génétiquement.

Les cellules souches peuvent être modifiées, entre autres, en utilisant un système de recombinaison Cre-lox pour générer des double ou triple KO des *Dnmts 3a*, *3b* et *1*. Dans l'article de King et al., ils ont utilisé ces modèles de cellules TKO (*Dnmt1*, *Dnmt3a/b*) et

DKO (*Dnmt3a/b* seulement) pour ensuite reconstituer l'expression des *Dnmts* et observer les changements épigénétiques qui se sont produits [160]. Ils ont donc pu caractériser ce qui se passe en absence des *Dnmts* et après le retour d'expression. Ils n'ont toutefois pas évalué l'impact de l'absence de DNMT1 uniquement. Le modèle utilisé pour mon projet, *Dnmt1<sup>tet/tet</sup>*, fonctionne selon un système Tetracycline-Off (Tet-Off), c'est-à-dire qu'au contact d'une tetracycline dans le milieu de culture, les cellules perdront complètement l'expression du gène ciblé par le système. Les cellules *Dnmt1<sup>tet/tet</sup>* conservent leur morphologie et demeurent indifférenciées pendant la prolifération et le traitement à la doxycycline [85]. Lorsque le traitement des cellules est arrêté et le milieu de culture est changé, les cellules sont en mesure de retrouver l'expression de *Dnmt1* après quelques jours (voir détails à la section Matériel et Méthodes). Ce modèle permet donc de jumeler l'étude de l'absence de DNMT1 en plus du « rescue » de l'expression.

Au niveau épigénétique, la perte d'expression de *Dnmt1* dans les cellules souches et les cellules somatiques (fibroblastes de souris) est bien caractérisée; cette perte est associée avec une diminution drastique de méthylation d'ADN et un remodelage des modifications d'histones H3K4me3 et H3K27me3 [161]. La marque active H3K4me3 est davantage retrouvée aux promoteurs des gènes lors d'une perte de *Dnmt1*, indiquant une augmentation de l'expression des gènes et la marque H3K27me3, antagoniste, est moins présente aux régions promotrices [161]. Cette perte de méthylation se produit majoritairement aux régions empreintes ou « imprinted-like » (DMRs et DMRs-like), mais aussi un peu hors de ces régions. Le retour d'expression de *Dnmt1* après une perte transitoire ne permet pas aux DMRs et DMRs-like de retrouver leurs patrons de méthylation d'ADN initiaux alors que les autres régions le peuvent. Le mécanisme qui empêche les DMRs de retrouver leur état initial est toujours sous investigation.

## 7. PROBLÉMATIQUE

Comme discuté précédemment, dans le développement embryonnaire précoce dit normal, on observe une vague de reprogrammation des patrons de méthylation d'ADN de l'ensemble du génome à l'exception des régions DMRs et DMRs-like sur lesquelles la



méthylation d'ADN est maintenue par l'action de DNMT1. Il est crucial que l'établissement de l'épigénome soit effectué de cette façon pour le bon développement de l'embryon. Si une absence d'expression de *Dnmt1* survient pendant la préimplantation, les patrons de méthylation et la distribution des modifications d'histones sera altérée entraînant des changements de conformation de la chromatine et possiblement aussi des changements dans l'expression des gènes, menant à des défauts sévères dans l'établissement du programme épigénétique fœtal [88]. Si l'absence de *Dnmt1* est temporaire, les régions DMRs et DMRs-like du génome embryonnaire seront incapables de retrouver leurs patrons de méthylation d'ADN initiaux malgré le retour d'expression de *Dnmt1*. Pourquoi ces régions ne sont pas en mesure de recouvrir leur état épigénétique initial après une perte temporaire d'expression de *Dnmt1* n'est toujours pas bien défini. Il est primordial de se pencher sur les mécanismes derrière ces erreurs du programme épigénétique foetal afin de définir la réponse adaptative transcriptomique et épigénétique qui survient dans l'ensemble du génome, mais aussi précisément aux régions promotrices et enhancers qui régulent la transcription des gènes. Ces recherches permettront de mieux prévenir des conséquences néfastes chez les futurs nouveau-nés.

## 8. HYPOTHÈSES

Notre hypothèse est qu'une adaptation erronée, c'est-à-dire des réarrangements et/ou des substitutions, de marques d'histones spécifiques suivant la perte temporaire de *Dnmt1* survient aux régions promotrices et enhancers, contribuant à empêcher les DMRs et DMRs-like à retrouver leurs niveaux initiaux de méthylation d'ADN. Plus spécifiquement, sachant qu'une perte temporaire de *Dnmt1* entraîne une augmentation globale du niveau d'H3K4me3 et une diminution globale d'H3K27me3, nous pouvons nous attendre à retrouver certains changements dans la distribution de ces marques, logiquement aux éléments régulateurs des régions qui ne peuvent retrouver leur méthylation d'ADN après le retour d'expression de *Dnmt1*. Nous nous attendons également à observer des changements de modifications d'histones qui ne sont pas corrélés aux changements de méthylation. Par exemple, certaines études soutiennent qu'H3K9me3, associée aux régions

hétérochromatiques constitutives, devrait se montrer plutôt indépendante des changements de méthylation d'ADN induits par la perte de *Dnmt1* alors que d'autres soutiennent qu'une synergie existe bien entre la méthylation et H3K9me3 [162]. Il sera donc intéressant de valider si sa distribution est affectée par la perte de méthylation d'ADN ou non et, si oui, comment.

## 9. OBJECTIFS

Utilisant notre modèle de cellules souches embryonnaires *Dnmt1<sup>tet/tet</sup>*, le principal objectif de ma maîtrise était donc de précisément définir, suite à la perte transitoire de *Dnmt1*, la réponse adaptative de l'ensemble de l'épigénome. Pour ce faire, des profils d'expression génique, de méthylation d'ADN ainsi que de distribution de modifications d'histones ont été établis pour trois conditions expérimentales (ctl, *Dnmt1* inactivé, *Dnmt1* réactivé; voir explication dans le Matériel et Méthodes). Le projet a été divisé en plusieurs sous-objectifs, soit la modélisation des dérèglements épigénétiques héréditaires qui surviennent lors d'une perte temporaire d'expression de *Dnmt1*, l'établissement des profils globaux d'expression génique, de méthylation d'ADN ainsi que des distributions de modifications d'histones, l'annotation des régions promotrices par bioinformatique et la détermination de la réponse adaptative de chacun des facteurs épigénétiques étudiés aux régions promotrices.

### *9.1 Modélisation de la dérégulation épigénétique héréditaire embryonnaire avec un modèle de cellules souches *Dnmt1<sup>tet/tet</sup>**

Le modèle de cellules ES *Dnmt1<sup>tet/tet</sup>* utilisé permet une inhibition temporaire de l'expression de *Dnmt1*. J'ai donc procédé à la récolte des cellules à trois différents moments : avant le traitement (cellules « contrôles », ctl), après un traitement à la doxycycline de 6 jours induisant la perte d'expression de *Dnmt1* (d0, *Dnmt1* inactivé), puis 21 jours après avoir enlevé la doxycycline du milieu (d21, *Dnmt1* réactivé).

### *9.2 Établissement des profils d'expression génique, de méthylation d'ADN ainsi que des distributions de marques d'histones*

Pour établir les profils transcriptomique et épigénétique, nous avons utilisé les techniques de séquençage de l'ARN (RNA-Seq), de séquençage au bisulfite par représentation réduite

(RRBS) et d'immunoprécipitation de la chromatine couplée à du séquençage à haut débit (ChIP-Seq) (décrites dans la section Matériel et méthode). Les marques d'histones ciblées pour la technique de ChIP-Seq sont toutes des marques ayant été associées soit dans le positionnement de DNMT1 (H3K9me3), le maintien de la méthylation (H3K27me3), l'activité génique (H3K4me3, H3K27ac) ou encore utiles pour la définition des enhancers qui seront analysés dans la prochaine étape du projet (H3K4me1). Les résultats pour H3K4me1 seront présentés tout de même pour les promoteurs pour déterminer si cette marque joue un rôle aux régions promotrices en plus des régions enhancers dans notre modèle cellulaire de perte transitoire de DNMT1.

### *9.3 Annotation des régions promotrices par bio-informatique*

À partir des données de ChIP-Seq obtenues et de bases de données de ChIP-Seq publiques, nous avons déterminé des critères pour annoter les régions promotrices et enhancers dans l'ensemble du génome. Les modifications H3K4me3 et H3K27me3 ont été utilisées dans le but de définir la position des régions promotrices dans nos données. Ces critères ont ensuite été utilisés pour établir les coordonnées d'annotations des promoteurs.

### *9.4 Déterminer la réponse adaptative de chacun des facteurs épigénétiques étudiés et leur dynamique*

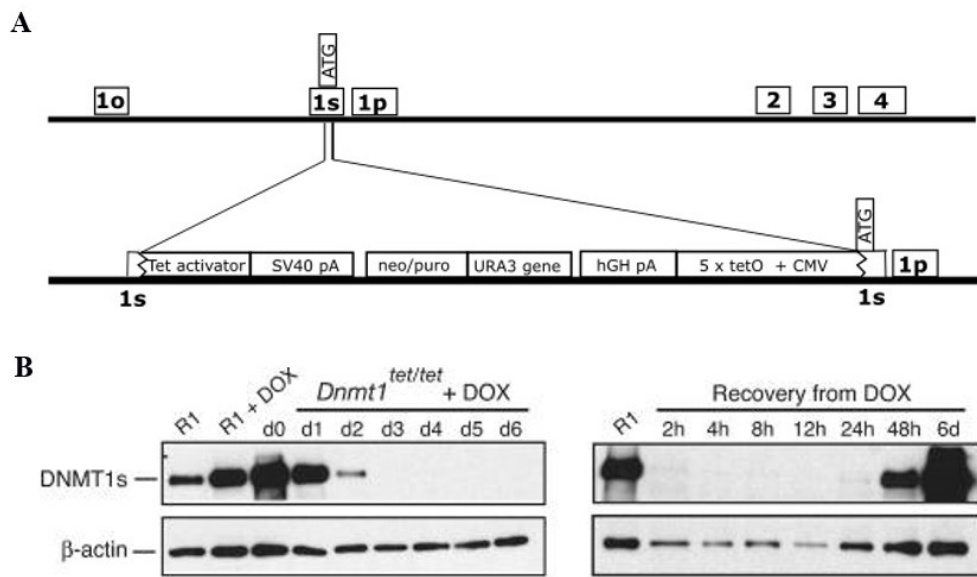
Les données brutes ont été traitées, annotées et alignées avec un pipeline d'analyse informatique afin de déterminer si des réarrangements de marques d'histones surviennent aux régions promotrices et enhancers dans l'ensemble du génome. De cette façon, il était possible de déterminer si une classe de promoteurs particulière était associée aux gènes les plus significativement affectés dans notre modèle de perte temporaire de *Dnmt1*. Les promoteurs actifs, bivalents, déplétés et réprimés ont été définis et étudiés. Des pics de distributions d'histones ont été générés pour tout le génome ainsi que des valeurs quantitatives d'expression et de méthylation pour, respectivement, les régions codantes et les régions CpG-riches du génome.

## 10. MATÉRIEL ET MÉTHODES

### *10.1 Modèle cellulaire *Dnmt1<sup>tet/tet</sup>**

Pour étudier l'impact d'une perturbation du maintien de la méthylation d'ADN sur l'épigénome du fœtus, nous avons utilisé un modèle cellulaire embryonnaire souche de souris modifié, *Dnmt1<sup>tet/tet</sup>*, provenant d'un collaborateur (Dr. John Richard Chaillet,

University of Pittsburgh) avec lequel il est possible d'induire une inhibition transitoire de l'expression du gène *Dnmt1*. Une cassette de régulation de type Tet-Off a été insérée par « knock-in » dans l'exon 1 des deux allèles de *Dnmt1* dans des cellules ES R1 (Figure 11A). Les cellules ES R1 sont dérivées d'un embryon murin à 3.5 jours, soit au stade blastocyste. La construction insérée dans les cellules comprend un activateur Tet (tTA, Tet activator), un signal de polyadénylation SV40 (SV40 pA), deux gènes de résistance, la néomycine (neo) et la puromycine (puro), le gène URA3 qui code pour une enzyme essentielle chez *S. Cerevisiae* ainsi qu'un signal de polyadénylation du gène Human Growth Hormone (hGH pA). La cassette contient aussi 5 opérateurs Tet (TetO) ainsi qu'un promoteur de cytomégalovirus (CMV) en amont du promoteur du gène *Dnmt1* (isoforme S). À l'ajout de 2ug/ml de doxycycline (DOX), le tTA se lie à la DOX et ne peut donc pas se lier aux opérateurs Tet afin d'activer l'expression du gène *Dnmt1*. L'expression de *Dnmt1* est considérée complètement perdue après 6 jours de traitement puisqu'une méthylation résiduelle de moins de 2% est observée, les cellules subissent environ huit divisions cellulaires sans expression de *Dnmt1* et une absence totale de la protéine est observée dès 3 jours de traitement DOX (d3) (Figure 11B) [85, 163]. Le modèle est ensuite réversible; lorsque le traitement DOX est arrêté et que le milieu de culture des cellules ES est changé, on détecte la présence de la protéine DNMT1 après 48h à un niveau semblable aux cellules contrôles (R1), confirmant le retour d'expression de *Dnmt1*.

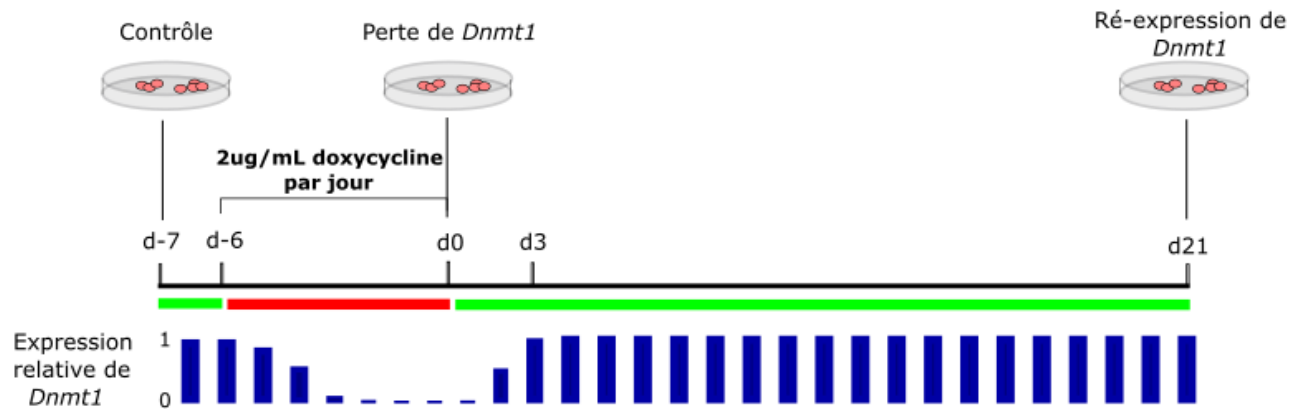


**Figure 11. Construction et validation du modèle ES *Dnmt1*<sup>tet/tet</sup>.** (A) Représentation de la construction insérée dans les cellules ES R1 pour générer la lignée de cellules *Dnmt1*<sup>tet/tet</sup>. L'insertion du système a été réalisée dans l'exon 1 de l'isoforme *Dnmt1s* avant le codon d'initiation de la transcription ATG. (B) Validation de l'expression de la protéine DNMT1 par Western Blot par l'équipe de Borowczyk et al.. R1 : Cellules contrôles, DOX : Traitement à la doxycycline, d0 : début de l'exposition à la DOX, d6 : 6<sup>e</sup> jour d'exposition à la DOX, 6d : 6 jours suivant l'arrêt du traitement DOX. La section de gauche démontre une perte complète de la protéine DNMT1 après 3 jours et la section de droite démontre un retour complet après 48h alors que la protéine  $\beta$ -actine demeure stable (Adapté de [85]).

### 10.2 Culture et récolte des cellules *Dnmt1*<sup>tet/tet</sup>

Les cellules embryonnaires souches de souris modifiées *Dnmt1*<sup>tet/tet</sup> ont été cultivées dans du milieu DMEM supplémenté en pyruvate et glucose avec les additifs suivants : 15% sérum bovin fœtal pour cellules ES, 2% L-glutamine 100X, 2% NEAA 100X, 2%  $\beta$ -mercaptoethanol, 2% Penicilline/Streptomycine, 2% Nucleosides 100X, Facteurs inhibiteurs de leucémie (LIFs) de souris (5 $\mu$ l à 10<sup>5</sup>U pour 50mL de milieu). Les cellules souches ont besoin d'un substrat supplémentaire au plastique pour adhérer et pour ne pas se différencier. Une couche de gélatine ou une couche de fibroblastes embryonnaires de souris (MEFs) inactivés répartis à confluence sur le fond du pétri peuvent être utilisés. Nos cellules ont été cultivées sur gélatine 0,1% comme l'utilisation de gélatine est plus simple que les MEFs et que les résultats sont tout aussi bons [164]. Les cellules ont été passées à confluence de 70-80%, soit environ aux 2 jours, avec de la trypsine 0,25%. Le milieu de culture a été changé à tous les jours. Les cellules ont été mises en culture et ont été récoltées une fois suffisamment amplifiées pour la condition contrôle (d-7), à l'exception d'un pétri de cellules qui a été utilisé pour la suite des temps expérimentaux. La perte d'expression de *Dnmt1* a ensuite été induite dans le pétri restant par un traitement DOX de 6 jours consécutifs (2 $\mu$ g/mL par jour) dans le milieu de culture des cellules *Dnmt1*<sup>tet/tet</sup>; le premier jour de traitement étant indiqué par d-6 et le dernier par d0 pour mes expériences. Le traitement a été effectué sur 6 jours malgré la perte complète de DNMT1 dès le 3<sup>e</sup> jour de traitement afin que quelques divisions cellulaires soient effectuées en absence de la protéine (temps de dédoublement d'environ 12h). À d0, les cellules ont été récoltées et un pétri de cellules a encore une fois été conservé en culture pour la réexpression de *Dnmt1*. Pour ce pétri restant, le milieu de culture des cellules a été changé pour du milieu sans

doxycycline. Les cellules ont été conservées en culture 21 jours suivant le retrait de la doxycycline pour permettre aux cellules de rétablir non seulement l'expression de *Dnmt1*, mais aussi les modifications épigénétiques, afin de ne conserver que les modifications du paysage épigénétique qui sont persistantes après 21 jours (soit environ 42 divisions cellulaires). Puis, les cellules ont été récoltées pour le dernier temps expérimental (d21) (Figure 12).



**Figure 12. Récolte des cellules *Dnmt1*<sup>tet/tet</sup> à différents temps expérimentaux.** Les cellules ont été récoltées avant le traitement (contrôle, d-7), après un traitement à la doxycycline de 6 jours (d0) où la protéine DNMT1 est absente et 21 jours après le rétablissement de l'expression de *Dnmt1* et du paysage épigénétique (d21). Le niveau d'expression relative de *Dnmt1* tout au long des temps expérimentaux est représenté par les barres bleues.

### 10.3 Validation de l'expression de *Dnmt1* par qPCR

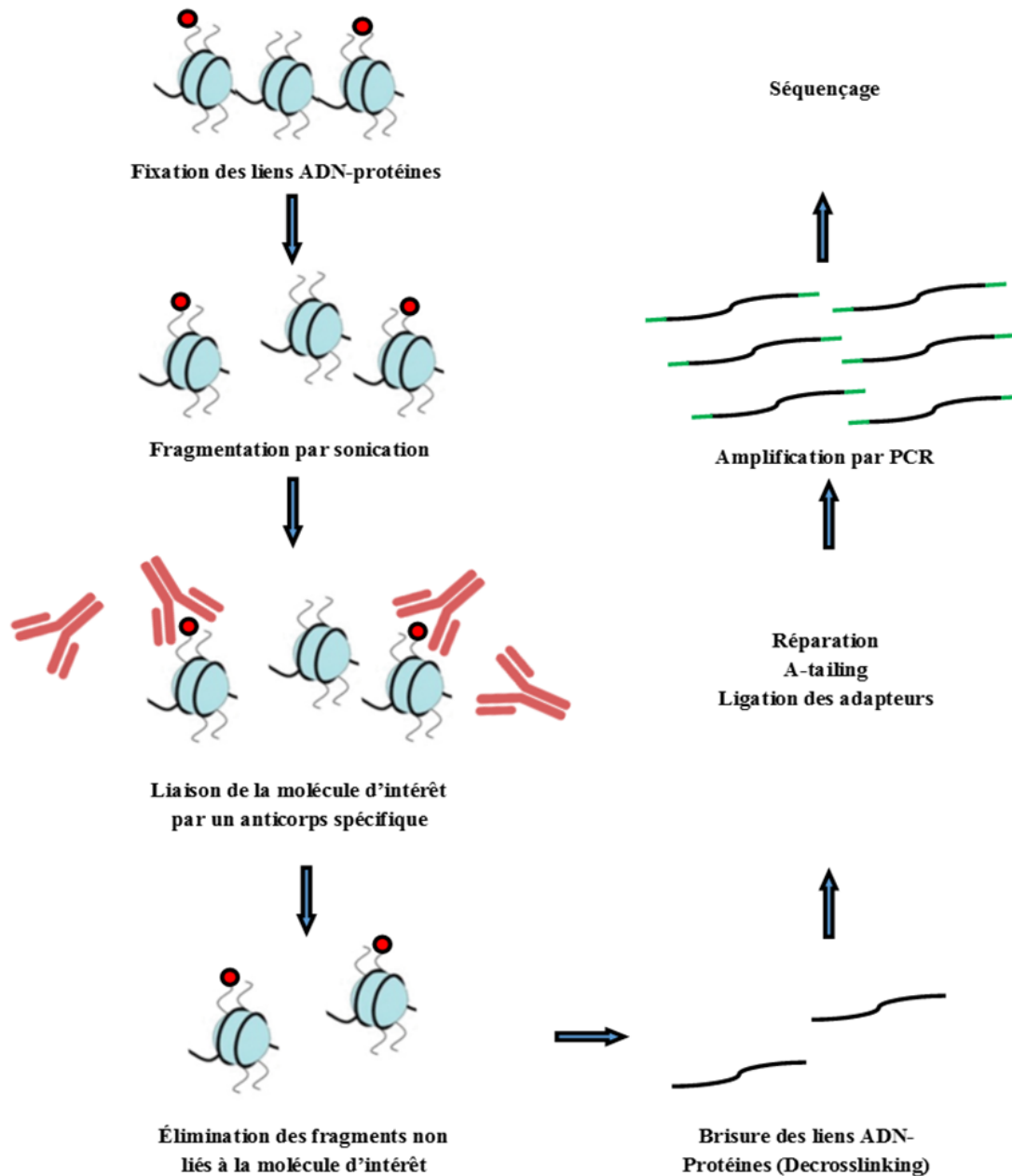
La technique de PCR quantitative en temps réel (qPCR) a été utilisée pour valider la perte d'expression de *Dnmt1* dans la lignée cellulaire *Dnmt1*<sup>tet/tet</sup>. Pour ce faire, des amorces ont été préparées pour *Dnmt1* et un gène contrôle, ici *Sdha*, un gène dont l'expression est stable dans notre lignée. L'ARN a été extrait des cellules contrôles, d0 et d21 ainsi que de deux conditions intermédiaires, soit d4 et d10, avec le RNeasy Mini Kit (Qiagen #74104). L'ARN a été converti en ADN complémentaire (ADNc) avec l'enzyme SuperScript IV (ThermoFisher Scientific #18090-050) puis la qPCR a été effectuée avec 1 ng d'ADNc par condition sur l'appareil Roche LightCycler 96 en triplicatas. L'enzyme utilisée pour la

qPCR est la TransStart Top Green qPCR Supermix (Civic Bioscience #AQ131). Les données d'expression obtenues ont été traitées selon la méthode de Delta CT en utilisant les valeurs du gène contrôle *Sdha* [165].

#### *10.4 Immunoprécipitation de la chromatine (ChIP-Seq) couplée au séquençage de haut-débit*

##### *10.4.1 Principe*

La technique d'immunoprécipitation de la chromatine (ChIP) consiste à fixer les interactions ADN-protéines de la chromatine pour isoler les fragments d'ADN liés à une protéine d'intérêt (histone/modification d'histone ou facteur de transcription) et ensuite séquencer ces fragments pour connaître la distribution de la protéine à travers le génome. Pour ce faire, la chromatine est d'abord fixée au formaldéhyde dans les cellules. La chromatine est ensuite extraite des cellules par lyse puis elle est soniquée pour donner de courts fragments. Les fragments d'ADN liés au facteur d'intérêt seront isolés par la liaison d'un anticorps spécifique. À l'aide de billes magnétiques (Protéine G ChIP-Grade, Cell Signaling #9006), les fragments liés par l'anticorps, et donc par le facteur d'intérêt, seront conservés et élués puis les liens ADN-protéine seront brisés en vue de la préparation de bibliothèques de séquençage. Les bouts d'ADN seront réparés et préparés pour la ligation (End-Repair, A-tailing) puis les adaptateurs, de courtes séquences d'ADN permettant l'identification des fragments, seront liés en vue de l'amplification et du séquençage (Figure 13).



**Figure 13. Représentation schématique de la technique d'immunoprécipitation de la chromatine couplée au séquençage à haut-débit (ChIP-Seq)**

#### 10.4.2 Préparation des immunoprécipitations et validation de l'enrichissement

Les cellules *Dnmt1<sup>tet/tet</sup>* récoltées aux différents temps (aliquots de 10 millions de cellules) ont d'abord été fixées avec du formaldéhyde 13% afin de maintenir les interactions ADN-protéines en place. Puis, les immunoprécipitations ont été réalisées selon un protocole



maison (pour H3K4me3, H3K27me3, H3K27ac) ou selon un protocole commercial modifié avec le IDEal ChIP-Seq Histone Kit (C01010051, Diagenode) sur le robot IP-Star (B03000002, Diagenode) (pour H3K9me3 et H3K4me1). Les cellules ont été préalablement lysées pendant 45 mins dans un tampon 10% SDS (protocole maison, Svotelis et al. 2009) ou 30 mins dans un tampon 1% IGEPAL pour la lyse membranaire puis 30 mins dans un tampon 1% SDS pour la lyse des noyaux (protocole commercial modifié). Puis, les cellules ont été soniquées au Bioruptor (Pico, Diagenode) pour 15 cycles 30 secondes ON/ 30 secondes OFF pour obtenir des fragments entre 150-400pb pour l'immunoprécipitation. Des anticorps validés pour le ChIP ont été utilisés et les quantités ont été ajoutées selon les recommandations du manufacturier (3ug d'H3K4me3 #39159 Active motif, 2ug d'H3K27me3 #07-449 Millipore, 2ug d'H3K27ac #ab4729 Abcam, 2ug d'H3K9me3 #ab8898 Abcam, 1ug d'H3K4me1 #CST5326 Cell Signaling). Les anticorps et la chromatine (volume total de 200uL) ont été incubés O/N à 4°C puis les billes magnétiques ont été ajoutées pour 2h avant de procéder aux lavages et à l'élution. Pour le protocole commercial modifié, ces étapes ont été réalisées sur le robot IP-Star avec des billes magnétiques (Protéine A) alors que pour le protocole maison, ces étapes ont été réalisées à la main avec des billes magnétiques (Protéine G).

L'enrichissement des immunoprécipitations a été validé par PCR quantitatif (qPCR) sur de l'ADN de ChIP pour au moins une paire d'amorces positives (régions enrichies) et négatives (régions dépourvues) pour les modifications d'histones ciblées préalablement à la préparation des bibliothèques. Les séquences des amorces de validation utilisées pour les différentes immunoprécipitations sont indiquées au Tableau I (H3K9me3 (*Gnas*, *H19*, *Desert2*), H3K4me1 (*Gapdh*, *Desert2*), H3K27ac (*Gapdh*, *Myod1*, *Inter1*), H3K27me3 (*Gapdh*, *Myod1*, *Inter1*), H3K4me3 (*Gapdh*, *IAP*, *Inter1*)). Les qPCR ont été réalisées sur le Roche LightCycler 96 avec l'enzyme TransStart Top Green qPCR Supermix (Civic Bioscience AQ131). Nous avons utilisé un calcul d'enrichissement par rapport à l'input (pourcentage). L'input est un aliquot de chromatine non-immunoprécipité qui permet de déterminer quel pourcentage de chromatine a été isolé par l'immunoprécipitation. Les signaux obtenus pour le ChIP sont donc divisés par la valeur du signal de l'input (qui représente 5% de la chromatine de départ pour l'immunoprécipitation) (Lin et al. 2012).

**Tableau I. Séquences des amorces utilisées pour valider l'enrichissement des immunoprécipitations**

<b>Amorce</b>	<b>Séquence <i>Forward</i></b>	<b>Séquence <i>Reverse</i></b>
<b>Gapdh F/R</b>	CCTCCTCCCTCTCTTTGGAC	TCCTATCCTGGGAACCATCA
<b>Myod1 F4/R2</b>	CCAGGACACGACTGCTTTCT	GCTCCATATCCCAGTTCCTG
<b>IAP up2/low2c</b>	AAGCAGCAATCACCCACTTTGG	CAATCATTAGATGCGGCTGCCAAG
<b>Desert2 F/R</b>	TTGGTGGCTCAGTGTCAGAG	CATAAGCTGGTGACCCCTGT
<b>H19 F1/R1</b>	GACTCGGACTCCCAAATCAA	GGCAACCCTGGTCTTTACAC
<b>Gnas F1/R1</b>	AAGGGCTTGAGAGAGGAAATGA	CCGAGACTCGCATCTTCCA
<b>IntergenicF1/R1</b>	CCCCTTTTGTGAGTCAGTGAA	CCAATCAATCATGGTGGGTA

L'enrichissement pour toutes les immunoprécipitations semble plus élevé pour les régions positives (18 à 57% de l'input pour toutes les modifications d'histones) comparativement aux enrichissements des régions négatives, qui varient entre 0 et 5% de l'input.

#### *10.4.3 Préparation des bibliothèques et séquençage*

Des duplicatas de bibliothèques ont été préparés en vue du séquençage selon le protocole de Rodrigue et al. 2010 pour le protocole maison et avec le Kapa Hyper Prep Kit pour le protocole commercial modifié (KK8502, Kapa Biosystems) [166]. Les bibliothèques du protocole maison (H3K4me3, H3K27me3, H3K27ac) ainsi que les bibliothèques du protocole commercial modifié ont été séquencées en Hi-Seq 2500 PE 50bp.

#### *10.5 Séquençage de l'ARN à haut-débit*

L'ARN a été extrait des cellules à partir d'un kit d'extraction ADN/ARN (Qiagen # 80284). Des bibliothèques d'ARN ont été préparées à partir des cellules aux trois points de récolte (ctl, d0 et d21). Un traitement à la DNase a été effectué préalablement pour éliminer les ADN contaminants puis le Kit TruSeq Stranded total RNA (# 20020596, Illumina) a été utilisé pour préparer les bibliothèques à partir de l'ARN total. Le kit Ribo-Zero Gold (#MRZG12324, Illumina) a été utilisé pour éliminer les ARN ribosomiaux (rRNA) des bibliothèques [167]. Les

librairies ont ensuite été multiplexées, c'est-à-dire combinées en un échantillon pour éviter les effets de « batch », et séquencées en Hi-Seq 4000 sur 2 flow-cell (5 lignes pour la flow-cell 1, 5 lignes sur la flow-cell 2), Paired-End 75bp.

### *10.6 Séquençage au bisulfite à représentation réduite (RRBS)*

#### *10.6.1 Principe et accession des données*

Le séquençage au bisulfite à représentation réduite (RRBS) consiste à convertir l'ADN extrait des cellules par un traitement au bisulfite, qui convertit les cytosines non méthylées en uracils alors que les cytosines méthylées sont protégées contre le traitement. Une librairie d'ADN converti au bisulfite est ensuite préparée pour être séquencée à haut-débit. Les librairies ont été préparées et séquencées sur Hi-Seq 2000 en PE 100bp par l'équipe de McGraw et al. selon les procédures décrites dans leur article [88]. Les données de méthylation utilisées pour nos analyses ont été tirées de European Nucleotide Archive ENA (Project Accession Number PRJEB6698).

### *10.7 Traitement bioinformatique des données brutes de séquençage*

Les données de séquençage brutes ont d'abord été traitées pour les trois techniques en collaboration avec deux bioinformaticiens, Maxime Caron (CHUSJ, McGill) et Anthony Lemieux (CHUSJ).

#### *10.7.1 Traitement et analyse des données brutes de séquençage de ChIP*

Les fichiers Bam ont d'abord été convertis en fastq puis les lectures (reads) ont été coupées via Trimmomatic pour éliminer les amorces de séquençage et les bases de mauvaise qualité [168]. Les données ont été alignées au génome de référence grâce à Burrows-Wheeler Aligner (BWA) puis le « Peak Calling » a été réalisé avec MACS2 [169, 170]. Puis, Flagstat et Samtools ont été utilisés pour générer des statistiques de contrôle qualité (ex : % alignement) et le nombre de lectures total [171]. Les nombres de lectures bruts et filtrés par échantillon ont ensuite été combinés dans un même fichier.

**Tableau II. Nombre de lectures de séquençage brutes obtenues pour chaque duplicata des données de ChIP-Seq**

Échantillon	Nombre de lectures / ligne de séquençage	Nombre de lectures
H3K9me3_Ctl 1	364 305 492	42 214 653
H3K9me3_Ctl 2		49 421 031
H3K9me3_d0 1		48 543 461
H3K9me3_d0 2		41 733 970
H3K9me3_d21 1		45 077 713
H3K9me3_d21 2		36 044 313
Input_d0		52 446 746
H3K4me1_Ctl 1	292 532 241	44 162 151
H3K4me1_Ctl 2		49 212 654
H3K4me1_d0 1		58 440 996
H3K4me1_d0 2		44 226 398
H3K4me1_d21 1		47 004 842
H3K4me1_d21 2		49 485 200
Input_Ctl	209 984 128	25 704 845
H3K27me3_Ctl 1		20 905 509
H3K27me3_d0 1		18 762 702
H3K27me3_d21 1		22 015 024
H3K27ac_Ctl 1		19 120 058
H3K27ac_d0 1		21 644 571
H3K27ac_d21 1		21 862 103
H3K4me3_Ctl 1		24 235 464
H3K4me3_d0 1		16 170 676
H3K4me3_d21 1		19 563 176
Input_d21	204633002	21 548 757
H3K27me3_Ctl 2		22 568 954
H3K27me3_d0 2		23 901 922
H3K27me3_d21 2		17 771 319
H3K27ac_Ctl 2		17 653 062
H3K27ac_d0 2		21 984 419
H3K27ac_d21 2		23 054 519
H3K4me3_Ctl 2		21 684 462
H3K4me3_d0 2		18 193 317
H3K4me3_d21 2		16 272 271

### 10.7.2 Traitement et analyse des données brutes de séquençage de l'ARN

Les séquences des librairies ont été alignées sur la version du génome de souris mm10 en utilisant STAR v2.4.2 [172]. Les nombres de lectures de séquençage alignées pour chaque librairie ainsi que les nombres de lectures par ligne de séquençage sont indiqués aux tableaux II et III. Les comptes d'expression ont été obtenus en utilisant htseq-count v0.6.1p1 sur les annotations de la version 81 d'Ensembl [173]. Les décomptes de lectures d'expression génique normalisés ont été obtenus avec deseq2 v1.6.3 sur 40,236 gènes (les gènes avec biotypes miRNA, Mt\_rRNA, Mt\_tRNA, snoRNA, snRNA et rRNA ont été retirés) et ont été transformés avec un log régularisé [174].

**Tableau III. Nombre de lectures de séquençage alignées obtenues pour chaque duplicata de librairies de RNA-Seq**

<b>Échantillon</b>	<b>Nombre de lectures alignées/ échantillon</b>
RNA_Ctl 1	65 423 458
RNA_Ctl 2	87 967 769
RNA_d0 1	84 102 535
RNA_d0 2	63 168 265
RNA_d21 1	77 994 180
RNA_d21 2	73 189 403

**Tableau IV. Nombre de lectures de séquençage alignées obtenues par ligne de séquençage pour les données de RNA-Seq**

<b>Flow-cell 1 (lignes)</b>	<b>Nombre de lectures alignées/ ligne</b>	<b>Flow-cell 2 (lignes)</b>	<b>Nombre de lectures alignées/ ligne</b>
<b>2</b>	21 087 775	<b>4</b>	50 664 548
<b>3</b>	50 070 981	<b>5</b>	47 713 244
<b>4</b>	50 086 860	<b>6</b>	46 590 460
<b>5</b>	49 128 112	<b>7</b>	44 256 615
<b>6</b>	50 686 096	<b>8</b>	41 560 919

### 10.7.3 Traitement et analyse des données brutes de séquençage RRBS

Les lectures ont été alignées sur mm10 en utilisant bsmmap v2.74 et les nombres de lectures de méthylation ont été obtenus par le même logiciel [175]. Pour chaque réplicat à chaque

condition expérimentale, une somme des nombres de lectures de méthylation a été calculée. Puis, une valeur de méthylation moyenne pour tous les réplicats a été calculée en conservant les régions génomiques où on retrouvait au moins 10 lectures. Finalement, seuls les CpGs suffisamment couverts pour les trois conditions expérimentales ont été utilisés pour les analyses subséquentes. La méthylation aux CpGs dans l'ensemble du génome a été comparée pour les conditions contrôle, d0 et d21.

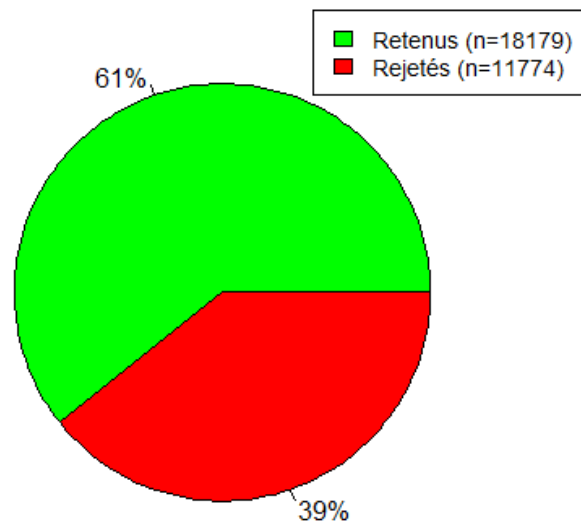
### *10.8 Analyse du profil épigénétique aux promoteurs*

Afin de savoir comment une perturbation du paysage épigénétique peut affecter la régulation de l'expression des gènes, nous avons orienté les analyses vers les promoteurs. Pour ce faire, nous avons d'abord généré des annotations, c'est-à-dire un fichier contenant toutes les localisations chromosomiques des régions promotrices. À partir de ces annotations, les données globales générées à la section 10.7 ont pu être alignées pour associer les changements de méthylation d'ADN, d'expression et de distribution de modifications d'histones à ces régions précisément.

#### *10.8.1 Définition des annotations pour les promoteurs*

Les annotations pour les promoteurs ont été déterminées à partir des données d'Encode pour H3K4me3 pour tous les tissus de souris disponibles (mm9) (GSM1000086, GSM918416, GSM769017, GSM1000132, GSM769027, GSM1000095, GSM769012, GSM769014, GSM1000083, GSM769016, GSM1000075, GSM1000079, GSM769021, GSM1000101, GSM769008, GSM1000123, GSM769029, GSM1000087, GSM769036, GSM1000128, GSM1000065, GSM769026) (ENCODE Project Consortium, 2012) [176]. Les données de pics de ChIP-Seq résultantes pour tous les tissus ont été fusionnées lorsque les pics étaient à plus ou moins 1Kb les uns des autres à l'aide de l'outil mergeBed de BEDTools [177]. Le centre de ces régions fusionnées a été déterminé et 2Kb ont été ajoutés de chaque côté du centre pour créer une zone tampon pour les promoteurs. Les positions chromosomiques théoriques des promoteurs de UCSC ont également été utilisées (n= 29953). L'intersection entre les coordonnées des pics issus d'Encode et les coordonnées des promoteurs de la base de données UCSC (promoteurs aussi définis à plus ou moins 2Kb du TSS) a été effectuée. Cette intersection a été fusionnée en régions de 1Kb et les coordonnées des promoteurs de UCSC non superposés avec les promoteurs définis selon

les données d'Encode ont été conservés. Une fusion de l'intersection entre UCSC et Encode avec les promoteurs de UCSC non superposés a été réalisée pour obtenir l'annotation finale de promoteurs. Pour les gènes ayant plus d'un TSS, le plus long isoforme a été conservé pour définir les annotations. Une seule région promotrice a été assignée par gène en conservant celle qui se trouve le plus proche du TSS. Dans le cas où deux annotations étaient superposées, elles ont été retirées excepté lorsque les annotations étaient près de 2 gènes différents. L'annotation résultante (61% des promoteurs théoriques de UCSC, n= 18179) a été transférée (Liftover) de mm9 à mm10 avec un ratio de 0.95, c'est-à-dire que 95% est demeuré identique entre les 2 versions (Figure 14).



**Figure 14. Proportion de promoteurs retenus pour les analyses.** Sur un total de 29953 promoteurs, 61% des promoteurs totaux ont été retenus (n=18179) alors que 39% ont été rejetés (n=11774) selon les critères de sélection et d'exclusion énoncés à la section 10.8.1.

L'annotation de promoteurs résultante a ensuite été sous divisée en différents types de promoteurs. Pour ce faire, nos données de ChIP-Seq pour les modifications d'histones H3K4me3 et H3K27me3, marques présentes aux régions promotrices, ont été alignées avec le fichier d'annotation de promoteurs. Ainsi, les promoteurs ont été divisés en promoteurs actifs (H3K4me3 +/ H3K27me3-), bivalents (H3K4me3 +/ H3K27me3 +), déplétés (H3K4me3 -/ H3K27me3 -) et réprimés (H3K27me3 +/ H3K4me3-) en fonction du niveau de modifications d'histones (> ou < 1 FPKM) retrouvées à chaque promoteur du fichier

d'annotation [160] (Figure 15). Le seuil de présence ou absence significative pour les modifications d'histones a été établi à 1 FPKM; la valeur normalisée en FPKM prend en considération la profondeur du séquençage (nombre de « reads ») et la longueur des fragments et permet donc de comparer les pics de modification d'histones entre eux [178]. Les modifications d'histones H3K4me1, H3K27ac et H3K9me3 n'ont pas été utilisées pour définir l'emplacement des promoteurs dans le génome, mais peuvent également être présentes à certains d'entre eux [160, 179]. Ces modifications seront traitées dans la section « Résultats ».

	H3K4me1	H3K4me3	H3K27me3	H3K27ac	H3K9me3
Actif					
Déplété					
Bivalent					
Réprimé					

**Figure 15. Définition des annotations de promoteurs.** Les modifications d'histones utilisées pour définir l'annotation sont indiquées en rouge (présence significative, seuil  $\geq 1$  FPKM), en bleu (absence significative, seuil  $< 1$  FPKM) ou en blanc (non utilisée pour les annotations) pour chaque catégorie de promoteurs.

Pour tous les promoteurs, un fichier a été généré pour regrouper les données de méthylation d'ADN, d'expression et des modifications d'histones étudiées pour les trois conditions expérimentales. Ce fichier nous a permis d'analyser en détail l'impact de la perte de *Dnmt1* sur le paysage épigénétique global dans nos cellules ES *Dnmt1<sup>tet/tet</sup>* et ainsi d'établir la réponse adaptative de l'expression, de la méthylation d'ADN et des modifications d'histones étudiées.

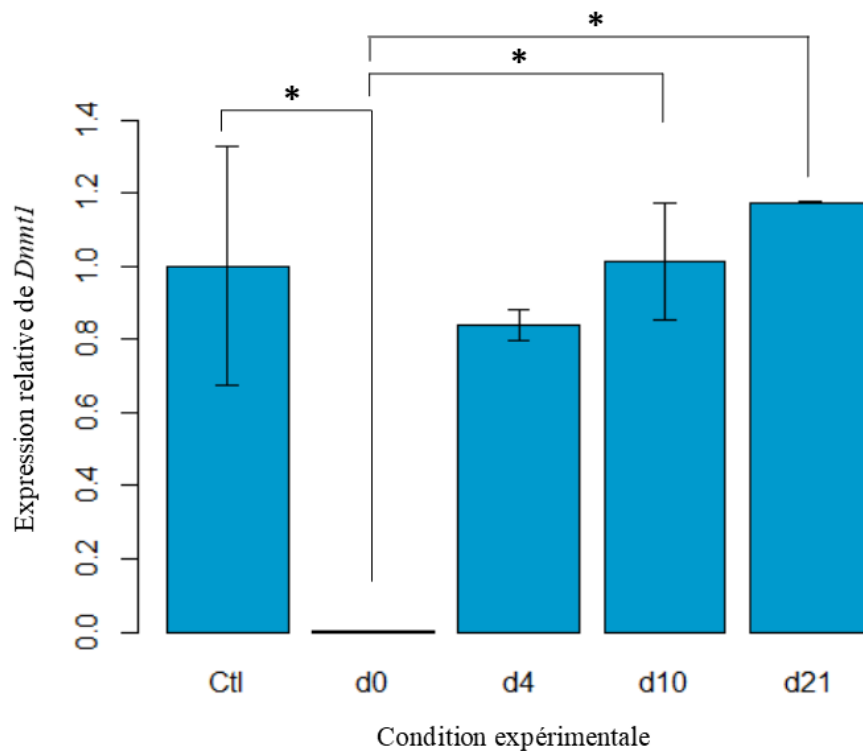


## 11. RÉSULTATS

### 11.1 Validation du modèle cellulaire ES *Dnmt1*<sup>tet/tet</sup>

#### 11.1.1 Perte d'expression de *Dnmt1*

Avant de procéder à toute analyse, il était crucial de valider le modèle cellulaire avec lequel les adaptations épigénétiques survenant pendant l'inactivation et suivant la réactivation de *Dnmt1* allaient être investiguées. Pour ce faire, nous avons récolté des cellules aux trois conditions expérimentales (Ctl : avant le traitement, d0 : après 6 jours de traitement DOX pour inactiver *Dnmt1*, d21 : période de réactivation de *Dnmt1* de 21 jours) et à deux temps intermédiaires lors du passage des cellules (d4 : 4 jours suivant la réactivation de *Dnmt1*, d10 : 10 jours suivant la réactivation de *Dnmt1*). Une extraction de l'ARN pour les 5 temps de récolte a été performée, suivie d'une conversion en ADNc puis d'une PCR quantitative (qPCR) pour valider le niveau d'expression de *Dnmt1* à chaque étape. La validation de l'induction de la perte de *Dnmt1* dans notre modèle cellulaire est représenté à la Figure 16. La validation de la perte de *Dnmt1* a pu être confirmée puisqu'une perte significative d'expression a été observée à d0 et que le niveau d'expression a retrouvé un niveau normal après l'arrêt temporaire (d10 et d21).



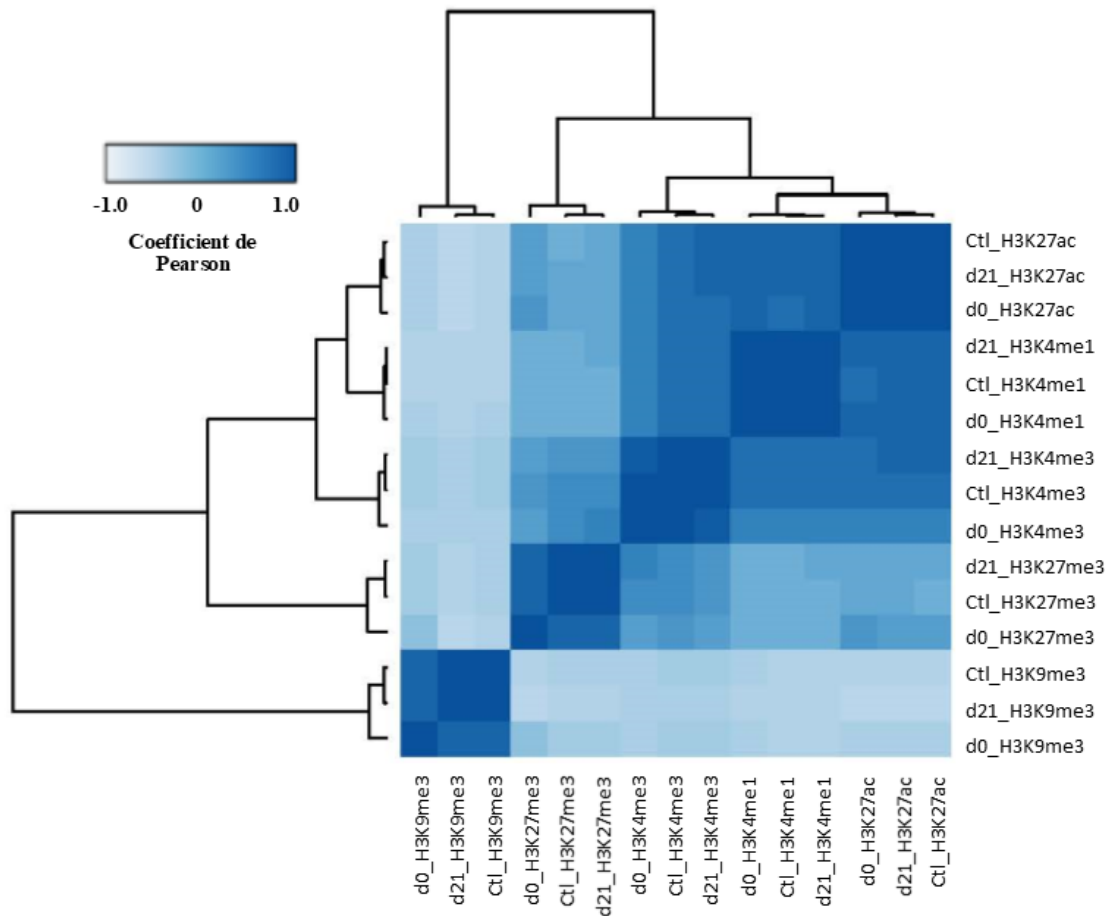
**Figure 16. Induction de la perte temporaire d'expression de *Dnmt1* dans les cellules ES *Dnmt1<sup>tet/tet</sup>*.** Le niveau d'expression relatif de *Dnmt1* (normalisé par l'expression du gène *Sdha*) est semblable pour les cellules contrôles (Ctl) avant le traitement et 4, 10 et 21 jours après la perte temporaire d'expression de *Dnmt1* alors que la perte est complète au jour 0 (d0). Un test des étendues de Tukey (One-way Anova) démontre que la perte d'expression s'est montrée significative pour les conditions Ctl-d0 (P-value = 0.0275), d10-d0 (P-value = 0.0301) et d21-d0 (P-value = 0.0173) (n=2).

### *11.2 Établissement des profils d'expression génique, de méthylation d'ADN ainsi que des distributions de marques d'histones*

Sachant que la perte temporaire de *Dnmt1* induit des perturbations de la méthylation d'ADN et un remaniement de certaines modifications d'histones chez les cellules ES, nous avons cherché à identifier si des perturbations similaires sont observées après un retour d'expression complet de *Dnmt1* dans notre modèle cellulaire ES *Dnmt1<sup>tet/tet</sup>*. Pour ce faire, nous avons procédé à l'analyse de la distribution de modifications d'histones pour ensuite établir si certaines perturbations de ces modifications sont associées à la présence ou l'absence de méthylation d'ADN ainsi qu'à des changements de l'expression des gènes.

#### *11.2.1 Distribution des modifications d'histones*

Une fois les bibliothèques de ChIP préparées et séquencées, nous avons cherché à valider le « clustering » des bibliothèques. Les comptes de lectures pour toutes les bibliothèques ont été comparés selon une corrélation de Pearson (Figure 17). Les différentes conditions pour une même modification d'histone (Ctl, d0, d21) corrélaient très fortement ensemble. Les conditions Ctl et d21 corrélaient davantage ensemble qu'avec la condition d0, ce qui indique un changement plus drastique des modifications d'histones à d0 qu'à d21. Également, le « clustering » nous permet de voir que les modifications d'histones actives ou retrouvées dans des régions similaires corrélaient davantage entre elles qu'avec H3K27me3 et H3K9me3, par exemple. Même entre les deux types de méthodes de préparation des bibliothèques de ChIP, on peut remarquer que les bibliothèques H3K4me1 (protocole commercial modifié) et H3K27ac (protocole maison) corrélaient fortement. Ce résultat démontre donc les corrélations attendues pour l'ensemble des bibliothèques de ChIP.



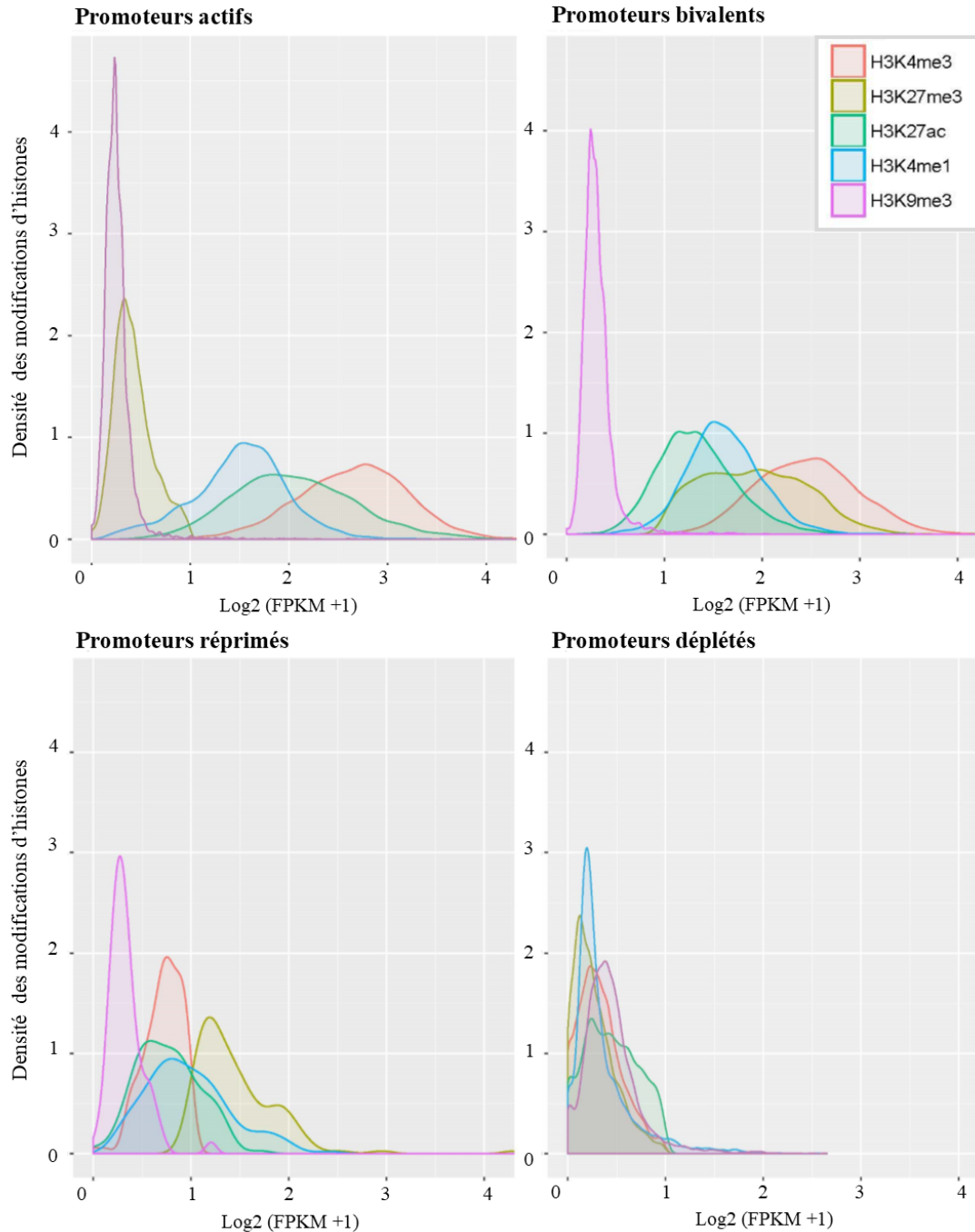
**Figure 17. Corrélacion entre les comptes de lectures du chromosome 1 des librairies de ChIP-Seq analysées.** Le chromosome 1 a été divisé en fragments de 10Kb et tous les échantillons ont été comparés entre eux selon leur distribution génomique. Une valeur du coefficient de corrélation de Pearson a été émise pour chaque comparaison et est représentée sur la « Heatmap ».

### 11.3 Annotation des régions promotrices par bio-informatique

Afin d'obtenir plus d'informations sur les effets d'une perte temporaire de *Dnmt1* dans notre modèle ES, nous avons d'abord ciblé l'analyse des profils épigénétiques aux régions promotrices afin de permettre des associations avec l'altération de l'expression de certains gènes. La littérature décrit des façons diverses d'annoter les régions promotrices, autant en se basant sur des modifications d'histones que sur des facteurs de transcription [180]. Des modifications d'histones largement utilisées pour la définition des annotations sont

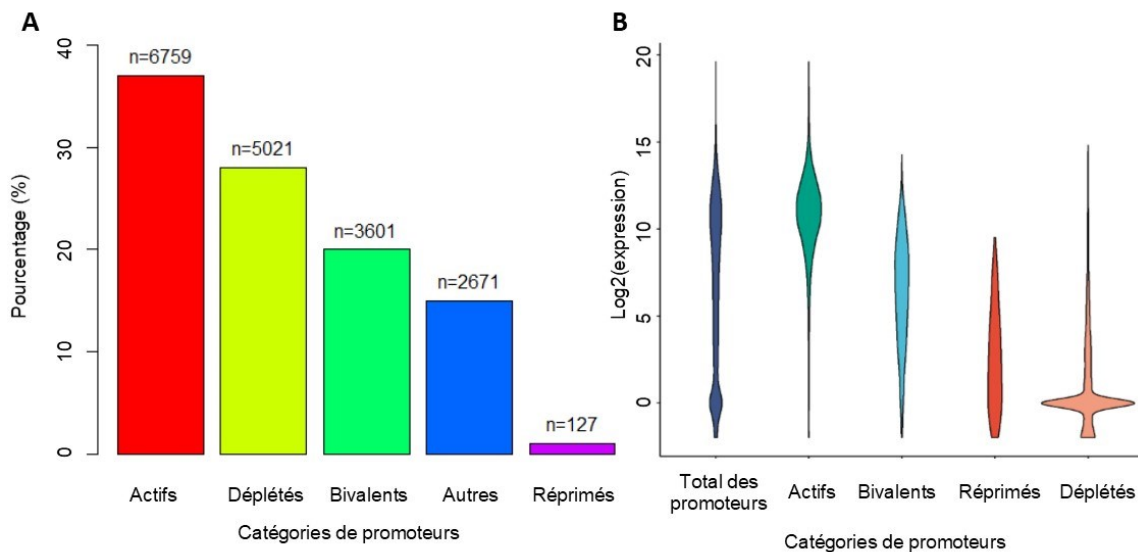
H3K4me3 et H3K27me3 pour les promoteurs. Nous nous sommes basés sur des modifications d'histones uniquement pour nos annotations.

Comme mentionné plus haut, nous avons défini et annoté les promoteurs selon différentes classes (actifs, bivalents, réprimés et déplétés) en vue de nos analyses. Pour chaque catégorie de promoteurs, les modifications d'histones et leurs densités respectives sont représentées à la figure 18. Un promoteur actif est caractérisé par la présence d'H3K4me3 et l'absence d'H3K27me3. Donc, on suppose que ces promoteurs activeront l'expression des gènes qu'ils contrôlent. Un promoteur bivalent est caractéristique des cellules souches puisqu'il possède les marques antagonistes H3K4me3 et H3K27me3 à une densité élevée et que le promoteur permettra l'activation ou la répression du gène en fonction du contexte (ex : induction de la différenciation). Les promoteurs réprimés sont, quant à eux, plus riches en H3K27me3 et les promoteurs déplétés ne possèdent pas de modifications ou très peu; ces régions promotrices sont donc plus fortement méthylées que les régions actives et bivalentes chez les cellules ES [160].



**Figure 18. Densité des modifications d'histones observée pour chaque catégorie de promoteurs.** Les courbes de densité de chaque modification d'histone sont illustrées en superposition ci-dessus pour chaque catégorie de promoteurs (actifs, bivalents, déplétés, réprimés). Le  $\text{log}_2(\text{FPKM}+1)$  fourni un indicatif de présence de la modification d'histone. Plus la valeur du  $\text{log}_2(\text{FPKM}+1)$  est élevée, plus la modification est présente.

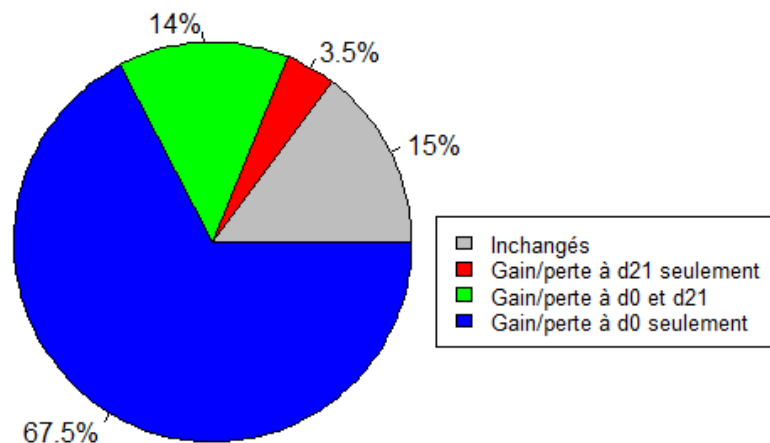
Il est intéressant d'observer que H3K27ac et H3K4me1 sont aussi des modifications d'histones présentes assez fortement aux promoteurs actifs et bivalents. Le pourcentage de promoteurs pour chaque catégorie (actif, bivalent, réprimé, déplété) est représenté à la figure 19A. On remarque une très faible proportion de promoteurs réprimés, soit près de 1% (n=127 sur 18179 promoteurs totaux). Les promoteurs actifs, déplétés et bivalents sont, quant à eux, présents respectivement à des pourcentages de 37%, 27% et 20%. Les promoteurs « Autres » sont des promoteurs qui ne se classent dans aucune des autres catégories dû à la présence de plusieurs marques ou dû à un seuil de présence trop faible (< 1 FPKM). La distribution générale de l'expression pour les promoteurs actifs, bivalents, réprimés et déplétés est représentée à la figure 19B. On remarque un niveau d'expression plus élevé pour les promoteurs actifs, intermédiaire pour les promoteurs bivalents, bas pour les promoteurs réprimés et presque nul pour l'ensemble des promoteurs déplétés. Les promoteurs ayant été précisément définis, la suite des analyses pourra ensuite être orientée vers les adaptations épigénétiques survenant à ces régions précises.



**Figure 19. Proportions des types de promoteurs dans notre annotation maison et leur expression.** **A)** Sur un nombre de promoteurs total de n=18179, 37% sont actifs, 28% sont déplétés, 20% sont bivalents, 15% sont autres et 1% sont réprimés. **B)** Distribution de l'expression (valeur en log régularisé, Deseq2) pour tous les promoteurs puis séparément pour les promoteurs actifs, bivalents, réprimés et déplétés.

#### 11.4 Caractérisation de la réponse adaptative de chacun des facteurs épigénétiques étudiés aux régions promotrices

Afin d'obtenir un portrait global des adaptations épigénétiques survenant à ces régions promotrices précédemment décrites et définies, les données des expériences de ChIP-Seq (H3K9me3, H3K4me1, H3K4me3, H3K27me3 et H3K27ac) ont été analysées en parallèle avec les données de RNA-Seq ainsi qu'avec des résultats de méthylation par séquençage au bisulfite à représentation réduite (RRBS). Globalement, pour les promoteurs, la distribution d'H3K4me3, modification principale pour la définition des annotations, demeure inchangée suite à la réactivation de *Dnmt1* pour la grande majorité des promoteurs. En effet, la figure 20 représente le pourcentage des promoteurs dont la distribution d'H3K4me3 change selon un patron similaire. Les sections en gris (15%) et bleu (67.5%) représentent la proportion des promoteurs qui n'ont pas subi d'adaptation à d21, soit un total de 82.5% des promoteurs. En vert (14%) et rouge (3.5%), ce sont les 17.5% de promoteurs qui subissent des gains ou des pertes en H3K4me3 à d0 et d21 ou à d21 seulement. Les analyses seront dirigées vers cette petite proportion de promoteurs affectés en vert et rouge suite à la réactivation de *Dnmt1* puisque nous sommes intéressés par les régions qui sont affectées de façon permanente dans notre modèle malgré le retour d'expression de *Dnmt1*.



**Figure 20. Proportions des promoteurs totaux ayant subi un changement de la distribution d'H3K4me3 entre les conditions Ctl et d0 ainsi qu'entre Ctl et d21. 15% des promoteurs totaux sont demeurés inchangés en H3K4me3 au cours de la perte et réactivation de *Dnmt1*. 67.5% des promoteurs ont subi un changement en H3K4me3 lors de la perte de *Dnmt1*, mais ont retrouvé leur niveau initial suite au retour de *Dnmt1*. 14%**

des promoteurs ont aussi subi un changement à d0 en plus de d21; le gain ou la perte s'est donc maintenu pendant l'inactivation de *Dnmt1* et après sa réactivation. 3.5% des promoteurs ont été affectés uniquement à d21, soit après la perte et la réactivation de *Dnmt1*.

Les données ont été recueillies et assemblées pour les promoteurs ayant subi un changement significatif en H3K4me3 (gain ou perte) pour les conditions Ctl-d0 et Ctl-d21 afin d'obtenir les promoteurs ayant été significativement changés à d0 (perte de *Dnmt1*) et dont le changement s'est maintenu à d21 (réactivation de *Dnmt1*) (en vert sur la figure 20). Ces promoteurs affectés sont représentés par les nuages de points de couleurs verte (gain) ou rouge (perte) dans les Figures 21, 23, 25 et 27A pour les 4 différents types de promoteurs. La modification d'histone H3K27me3 a plutôt été utilisée pour les promoteurs réprimés comme H3K4me3 est absente à ces régions. Pour chaque type de promoteurs, une « heatmap » combinant les changements (« fold change ») pour toutes les modifications d'histones ainsi que pour l'expression et la méthylation d'ADN permet d'avoir une vue globale de l'adaptation des différents facteurs épigénétiques étudiés. Ainsi, tous les promoteurs ont d'abord été classés selon s'ils avaient un changement ou non de distribution en H3K4me3 (ou H3K27me3 pour les promoteurs réprimés) à d0 et d21 dans la première colonne des « heatmaps » des figures 21, 23, 25 et 27 B. Un gain a été indiqué en vert, une absence de changement, en noir, et une perte, en rouge. En respectant l'ordre des promoteurs de la première colonne (H3K4me3 ou H3K27me3), la valeur de « fold change » pour toutes les autres modifications d'histones ainsi que pour la méthylation d'ADN et l'expression a été attribuée pour chaque promoteur. Les modifications d'histones sont représentées dans les mêmes couleurs alors que des couleurs différentes ont été utilisées pour la méthylation (Gain = Jaune, Perte = Bleu, Pas de changement= Noir, Donnée non disponible= Gris) et l'expression (Gain = Rouge foncé, Perte = Bleu Foncé, Pas de Changement= Blanc).

Par la suite, les sous-groupes de gènes les plus pertinents pour chaque catégorie de promoteurs ont été investigués pour émettre des hypothèses quant à l'adaptation épigénétique survenant à d21. Les analyses détaillées pour chaque type de promoteurs sont présentées aux sections 11.4.1 à 11.4.4 ci-dessous.

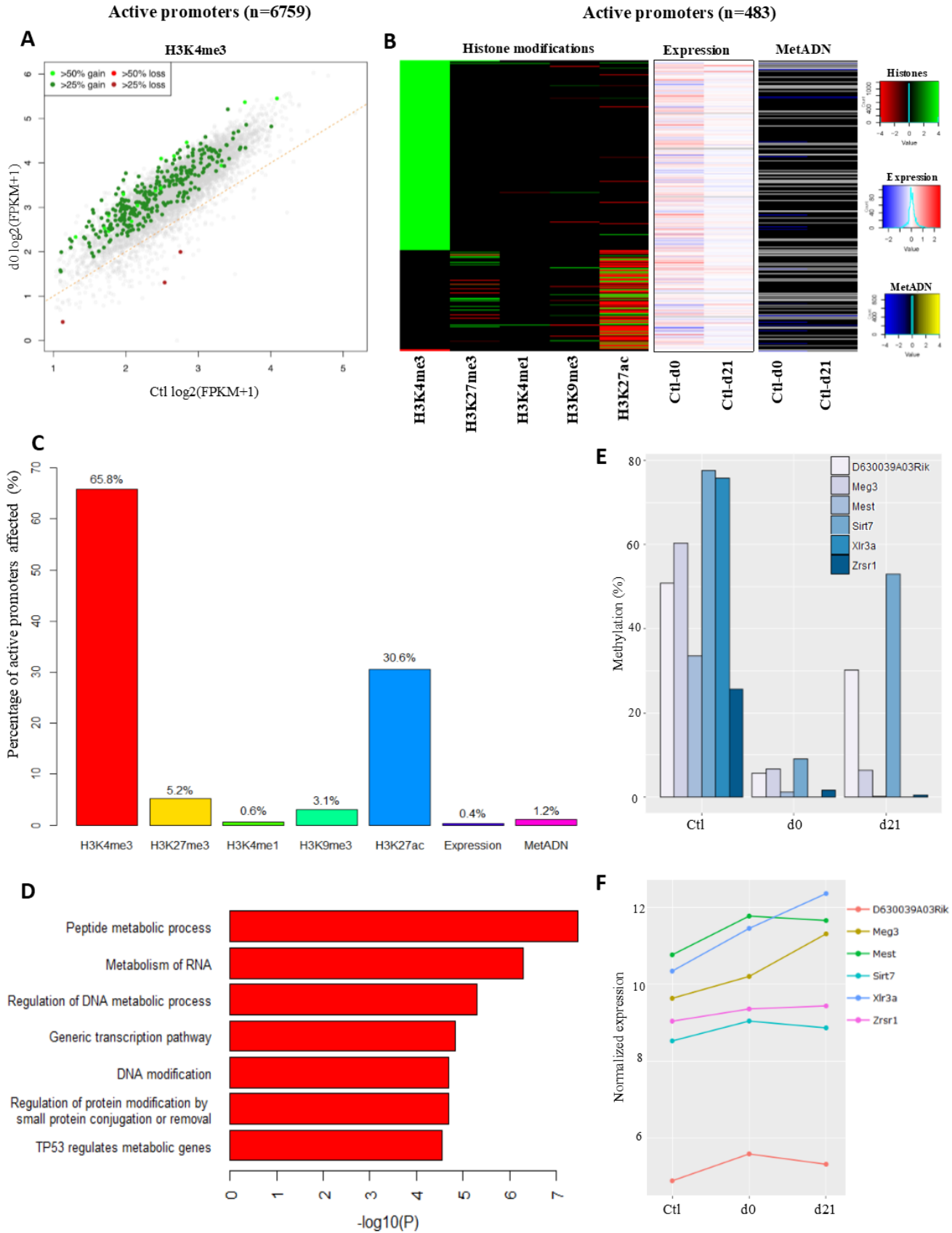


#### 11.4.1 Promoteurs actifs

Comme mentionné plus haut, nous avons débuté par analyser la modification d'histone H3K4me3 puisque plusieurs promoteurs ont subi des changements en densité de cette marque. Des valeurs quantitatives de perte, gain ou non-changement ( $\log_2(\text{FPKM}+1)$ ) ont été calculées pour déterminer le changement observé en H3K4me3 à d0 et à d21, comparativement à la condition contrôle. Les valeurs résultantes à d0 et à d21 ont été corrélées aux valeurs de la condition contrôle (Figure 21A). Parmi tous les promoteurs actifs (n=6759), la majorité des changements observés en H3K4me3 est une augmentation (points verts, n=315) alors qu'une proportion négligeable d'entre eux subissent une perte (points rouges, n=3). Parmi les gènes ayant un gain d'H3K4me3 de 50% ou plus à d0 et d21 (points vert pâle), on retrouve les promoteurs actifs de *Xlr3a* (gène décrit comme « imprinted-like ») et *Meg3* (gène empreinte). Les promoteurs ayant eu un gain de 25% et plus sont, quant à eux, beaucoup plus nombreux. Parmi ceux-ci, on retrouve les promoteurs des gènes empreintes *Snurf* et *Zrsr1*. Au total, ce sont près de 66% des promoteurs actifs (318 de n=483) qui subissent un changement en H3K4me3 à d0 qui est maintenu à d21 (Figure 21C).

En ce qui concerne les autres modifications d'histones, il est à noter que les changements en H3K27me3, H3K9me3 et H3K27ac observés pour les promoteurs inchangés en H3K4me3 touchent des promoteurs différents, c'est-à-dire que rares sont les promoteurs actifs partageant une augmentation de plus d'une de ces marques à d0 et d21 (Figure 21B). La deuxième modification d'histone la plus affectée au niveau des promoteurs actifs est H3K27ac; elle représente 30.6% des promoteurs actifs affectés à d0 et d21 (Figure 21C). La majorité des changements en H3K27ac sont des pertes significatives de la marque (rouge), mais des augmentations (vert) sont observées pour quelques promoteurs également (Figure 21 B, colonne H3K27ac). Des analyses d'ontologie génique par Metascape [181] ont été réalisées pour les promoteurs ayant eu un gain significatif en H3K4me3 (promoteurs en vert dans la première colonne de la « heatmap », figure 21B). Près de 10% de ces gènes se sont avérés impliqués dans le métabolisme peptidique (*Apobec1*, *Rpl6*, *Sox4*, *Upf1*, etc.) et près de 3% des gènes sont impliqués dans la modification de la chromatine ou de l'ADN (*Kdm4b*, *Meg3*, *Prmt5*, *Parp1*, etc.) (Figure 21D).

En ce qui concerne l'expression et la méthylation d'ADN, seuls quelques gènes ont été affectés significativement (différence d'expression de 2 « fold change » et différence de méthylation d'au moins 20%). Les valeurs d'expression et de méthylation différentielles ont été déterminées pour Ctl-d0 et Ctl-d21 (Figure 21 B) (Méthylation d'ADN : Gain = Jaune, Perte = Bleu, Pas de changement = Noir, Donnée non disponible = Gris ; Expression : Gain = Rouge foncé, Perte = Bleu Foncé, Pas de Changement = Blanc). Ainsi, un changement de méthylation significatif entre Ctl-d0 (bleu) qui n'est pas maintenu entre Ctl-d21 (noir) représente un promoteur qui a perdu sa méthylation lors de l'inhibition de DNMT1, mais qui a retrouvé son niveau de méthylation initial. Peu de promoteurs actifs ont subi un changement de méthylation maintenu à d21; ces promoteurs sont représentés en bleu à la figure 21 B dans la colonne MetADN Ctl-d21. Ces promoteurs ayant perdu leur méthylation à d21 sont les promoteurs des gènes *Meg3*, *Mest*, *Sirt7*, *Xlr3a*, *Zrsr1* et *D630039A03Rik*. Leurs pourcentages de méthylation sont détaillés à la figure 21E. Les niveaux de méthylation de ces promoteurs varient entre 25 et 77% en condition contrôle, entre 0 et 9% lorsque *Dnmt1* est inactivé (d0) et entre 0 et 52% après réactivation (d21), ce qui veut dire que certains de ces promoteurs ne réacquièrent pas des groupements méthyles entre d0 et d21 alors que d'autres oui, mais pas suffisamment pour retrouver leur niveau initial. À noter que ces promoteurs ont tous un contenu en CpG moyen (entre 2 à 4%) excepté *Zrsr1* qui a un contenu de 6%. Parmi les gènes dont la méthylation au promoteur est affectée, seuls *Xlr3a* (d0 et d21) et *Meg3* (d21), deux gènes hautement exprimés en condition contrôle, ont une augmentation significative d'expression. Les autres gènes ont une expression différentielle moindre, ce qui ne corrèle pas nécessairement avec les pertes de méthylation d'ADN observées (Figure 21F).



**Figure 21. Adaptation épigénétique (modifications d’histones, méthylation d’ADN et expression) aux promoteurs actifs entre les conditions Ctl et d21 dans les cellules ES *Dnmt1<sup>tet/tet</sup>*.**

**A. Changements significatifs de la distribution de H3K4me3 aux promoteurs actifs entre les conditions Ctl-d0 et Ctl-d21.** Chaque point corrèle la valeur normalisée de FPKM de la condition contrôle avec la valeur de la condition d0. Si un changement significatif est observé à d0 ainsi qu’à d21, les points seront colorés en vert ou en rouge. Les points verts représentent les promoteurs ayant un gain de 25% ou plus, les points rouges représentent une perte de 25% ou plus et les points gris sont les promoteurs inchangés.

**B. Adaptation de la distribution des modifications d’histones H3K27me3, H3K4me1, H3K9me3 et H3K27ac, de la méthylation d’ADN et de l’expression pour les promoteurs actifs affectés ayant une augmentation, une perte ou pas de changement en H3K4me3.** Chaque colonne représente la valeur normalisée différentielle de présence d’une modification d’histone différente (à d0 et d21). Chaque ligne représente un promoteur. Les promoteurs ayant une augmentation de plus de 25% du « fold change » sont représentés en vert, ceux ayant une diminution de plus de 25% sont représentés en rouge et ceux étant demeurés stables sont représentés en noirs. Les promoteurs de gènes n’ayant subi aucun changement pour toutes les modifications d’histones ont été retirés des « heatmaps » (actifs = 93%). Pour la méthylation d’ADN, un changement significatif (gain ou une perte) a été évalué à 20%. Les couleurs représentées sur la « heatmap » sont les suivantes (Gain = Jaune, Perte = Bleu, Pas de changement= Noir, Donnée non disponible= Gris). En ce qui a trait à l’expression, le seuil de significativité a été attribué à 2 « fold change » pour les gains et les pertes d’expression et les couleurs sont les suivantes (Gain = Rouge foncé, Perte = Bleu Foncé, Pas de Changement= Blanc). Pour la méthylation d’ADN et l’expression, les valeurs différentielles entre Ctl-d0 et Ctl-d21 sont comparées.

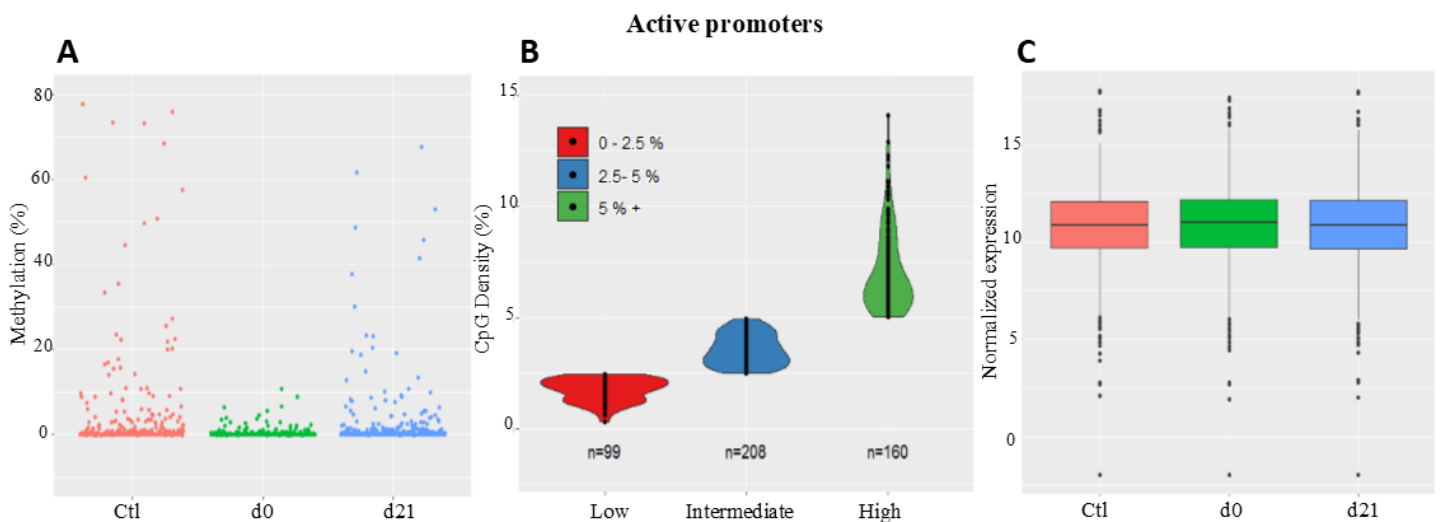
**C. Pourcentages de promoteurs actifs affectés (total n=483) à d21 pour les modifications épigénétiques étudiées.**

**D. Ontologie génique des promoteurs actifs ayant subi une hausse d’H3K4me3 à d21.**

**E. Pourcentage de méthylation des 6 promoteurs actifs ayant subi une perte de méthylation d’ADN significative à d21 pour les conditions Ctl, d0 et d21.**

**F. Valeurs d’expression des gènes sous le contrôle des 6 promoteurs actifs ayant subi une perte de méthylation d’ADN significative à d21.** Les valeurs d’expression représentées sont une normalisation des valeurs des duplicatas de bibliothèques de RNA-Seq.

Globalement, les promoteurs actifs sont caractérisés par une faible méthylation pour permettre l'expression des gènes sous leur contrôle. Comme les promoteurs actifs dans nos analyses ont des pourcentages de méthylation majoritairement très bas (moins de 20%), ceci peut en partie expliquer pourquoi peu de changements significatifs de méthylation sont observés (Figure 22A). Le contenu en CpGs de ces promoteurs est plutôt élevé; la majorité des promoteurs ont un contenu intermédiaire (2.5-5%) ou élevé (plus de 5%) (Figure 22B). L'expression globale des gènes sous le contrôle de ces promoteurs est aussi relativement élevée, comme attendu, aux conditions Ctl, d0 et d21 (Figure 22C).



**Figure 22. Évaluation de la méthylation et de la densité en CpGs aux promoteurs actifs ainsi que l'expression globale des gènes associés.** **A)** Chaque point correspond au pourcentage de méthylation d'un promoteur actif en condition contrôle, d0 et d21. **B)** La densité en CpGs de chaque promoteur actif a été divisé en classes de faible (0 à 2,5% du promoteur est constitué de CpGs), moyen (2.5-5%) et élevé (plus de 5% du contenu). **C)** La valeur normalisée d'expression pour chaque gène a été représentée pour chaque condition (ctl, d0, d21). La valeur moyenne est représentée au centre des boîtes.

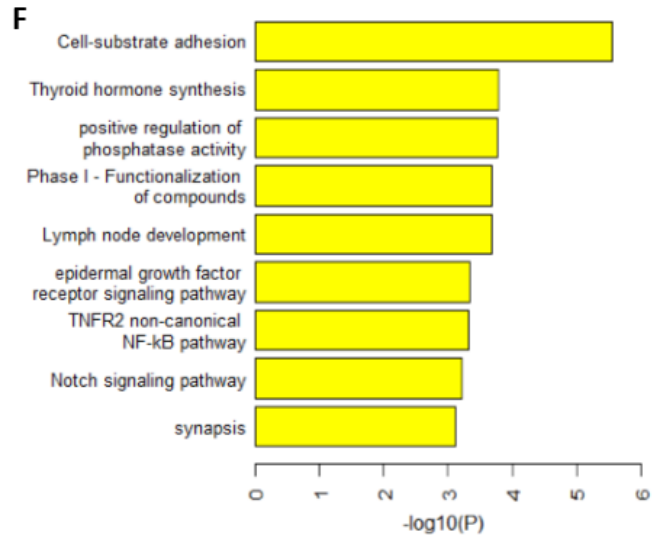
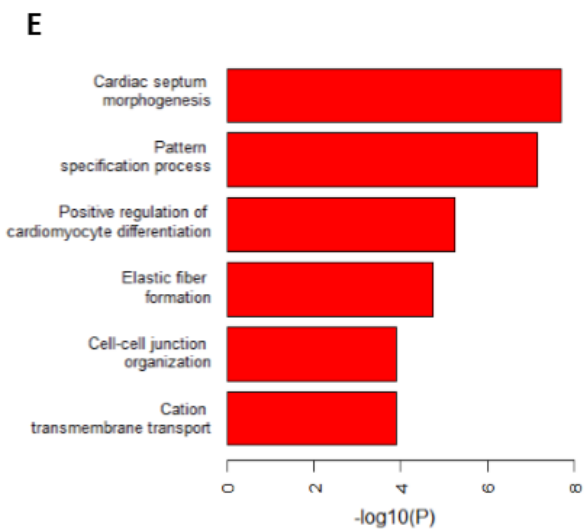
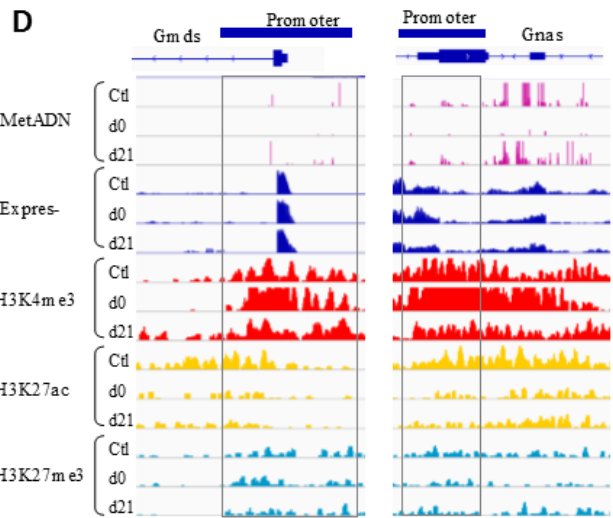
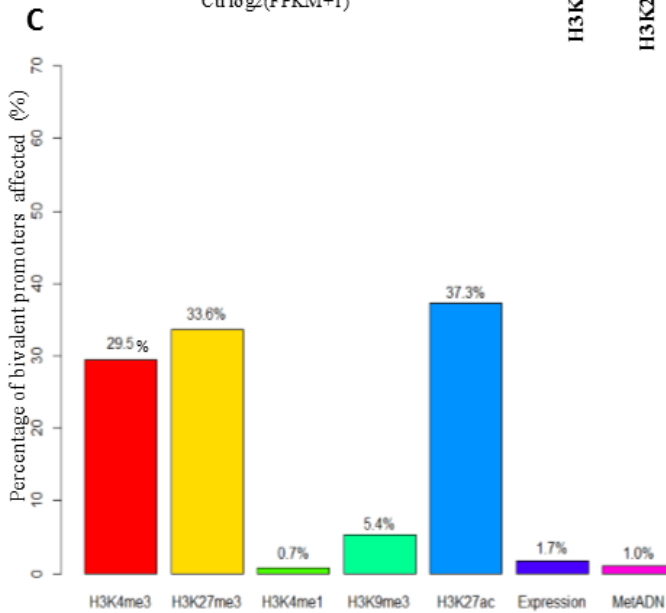
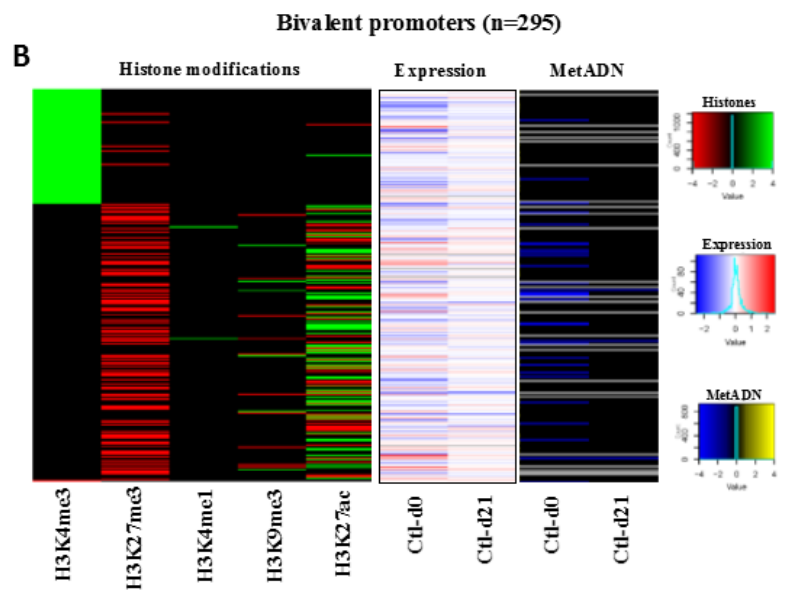
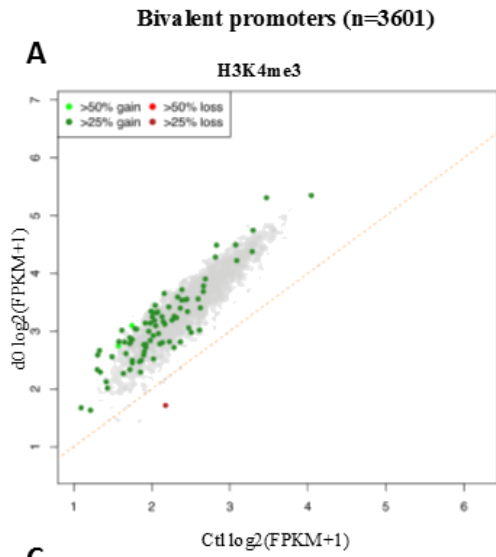
En somme, on remarque que les promoteurs actifs affectés dans notre modèle cellulaire ES *Dnmt1*<sup>tet/tet</sup> sont peu nombreux sur le nombre total de promoteurs actifs annotés (n=6759). Les adaptations épigénétiques observées sont majoritairement des augmentations de la densité en H3K4me3 et des pertes d'H3K27ac. Selon King et al., la méthylation d'ADN

agirait en amont d'H3K27ac alors qu'H3K4me3 serait plutôt indépendante [160]. Ici, ces adaptations ont été peu associées à des changements majeurs d'expression ou de méthylation d'ADN. Ceci peut toutefois suggérer que les deux modifications d'histones seront affectées différemment par la perte de maintien de la méthylation par DNMT1.

#### 11.4.2 Promoteurs bivalents

Des analyses similaires ont été effectuées pour les promoteurs bivalents afin de caractériser l'adaptation épigénétique entraînée par une absence du maintien de la méthylation par DNMT1 et son effet sur les gènes dont l'expression est méticuleusement contrôlée lors du développement embryonnaire. Le nuage de points en Figure 23A démontre qu'une petite population de promoteurs bivalents gagnent de la marque H3K4me3 (points verts) suite à l'inactivation (d0) et la réactivation de *Dnmt1* (d21). La « heatmap » en B démontre également la présence de nombreux changements en H3K27me3 (perte) ainsi qu'en H3K27ac (gains et pertes). Davantage de perturbations de la méthylation à d0 semblent être observées pour les promoteurs qui auront un changement en H3K27me3 et H3K27ac à d21. De tous les promoteurs bivalents affectés par la perte temporaire de *Dnmt1* à d21, 29.5% ont eu un changement en H3K4me3, 33.6% en H3K27me3 et 37.3% en H3K27ac. Deux exemples de promoteurs ayant subi à la fois une perte de H3K4me3 et une perte de H3K27me3 sont démontrés à la Figure 23D. Pour le promoteur de *Gnas*, par exemple, on observe un niveau plus élevé d'H3K4me3 en conditions Ctl et d0 comparativement à d21. Pour H3K27me3, la perte se produit à d0 et d21 et le niveau résiduel de la modification d'histone tombe sous le seuil de présence significative (1 FPKM). Ainsi, le promoteur de *Gnas* a perdu son titre de bivalent comme il ne reste qu'H3K4me3. Le promoteur serait donc considéré « actif » sans être lié à un changement de l'expression du gène ni à un changement de méthylation d'ADN au promoteur, possiblement dû au niveau résiduel faible d'H3K27me3 (<1 FPKM, non significatif) à cette région. Pour tenter de récupérer davantage d'informations concernant les gènes des promoteurs bivalents affectés, des analyses d'ontologie génique ont été réalisées ainsi que des analyses de recherches de motifs de facteurs de transcription (résultats non-significatifs). La figure 23E présente les « GO Terms » identifiés pour les gènes ayant eu un gain en H3K4me3. Fait intéressant, plusieurs des fonctions obtenues semblent liées au développement cardiaque (*Nkx2-5*,

*Tbx5*, *Lefty1*, etc.). D'ailleurs, certains de ces facteurs cruciaux pour le développement cardiaque sont des facteurs de transcription impliqués dans le recrutement d'enzymes de remodelage de la chromatine [182]. Ainsi, la perte de DNMT1 pourrait potentiellement induire un remodelage de la chromatine via ces promoteurs bivalents. La même approche a été utilisée pour les gènes dont les promoteurs ont perdu en H3K27me3. Les fonctions étaient, cette fois, beaucoup plus variées et générales comme l'adhésion cellule-substrat ou encore la synthèse d'hormones. Il ne semble donc pas y avoir de lien direct entre les promoteurs affectés par une perte en H3K27me3. Il est à noter que la majorité de ces promoteurs ont perdu de la marque H3K27me3 sans pour autant perdre leurs caractéristiques bivalentes, c'est-à-dire que le niveau d'H3K27me3 est toujours au-dessus du seuil de significativité de 1 FPKM malgré la perte.





**Figure 23. Adaptation épigénétique (modifications d'histones, méthylation d'ADN et expression) aux promoteurs bivalents entre les conditions Ctl et d21 dans les cellules ES *Dnmt1*<sup>tet/tet</sup>.**

**A. Changements significatifs de la distribution de H3K4me3 aux promoteurs entre les conditions Ctl-d0 et Ctl-d21.** Chaque point (n=3601) corrèle la valeur normalisée de FPKM de la condition contrôle avec la valeur de la condition d0. Si un changement significatif est observé à d0 ainsi qu'à d21, les points seront colorés en vert ou en rouge. Les points verts représentent les promoteurs ayant un gain de 25% ou plus, les points rouges représentent une perte de 25% ou plus et les points gris sont les promoteurs inchangés.

**B. Adaptation de la distribution des modifications d'histones H3K27me3, H3K4me1, H3K9me3 et H3K27ac, de la méthylation d'ADN et de l'expression pour les promoteurs bivalents affectés ayant une augmentation, une perte ou pas de changement en H3K4me3.** Chaque colonne représente la valeur normalisée différentielle de présence d'une modification d'histone différente (à d0 et d21). Chaque ligne représente un promoteur bivalent. Les promoteurs ayant une augmentation de plus de 25% du « fold change » sont représentés en vert, ceux ayant une diminution de plus de 25% sont représentés en rouge et ceux étant demeurés stables sont représentés en noirs. Les promoteurs de gènes n'ayant subi aucun changement pour toutes les modifications d'histones ont été retirés des « heatmaps » (bivalents = 92%). Pour la méthylation d'ADN, un changement significatif (gain ou une perte) a été évalué à 20%. Les couleurs représentées sur la « heatmap » sont les suivantes (Gain = Jaune, Perte = Bleu, Pas de changement = Noir, Donnée non disponible = Gris). En ce qui a trait à l'expression, le seuil de significativité a été attribué à 2 « fold change » pour les gains et les pertes d'expression et les couleurs sont les suivantes (Gain = Rouge foncé, Perte = Bleu Foncé, Pas de Changement = Blanc). Pour la méthylation d'ADN et l'expression, les valeurs différentielles entre Ctl-d0 et Ctl-d21 sont comparées.

**C. Pourcentages de promoteurs bivalents affectés (total n=295) à d21 pour les modifications épigénétiques étudiées.**

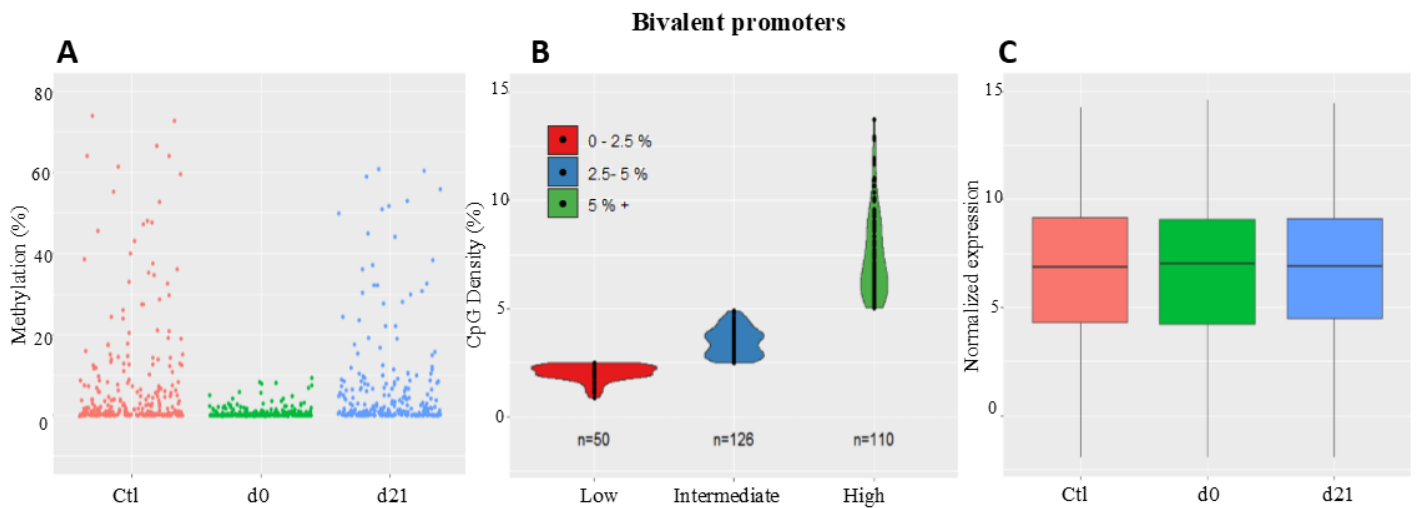
**D. Représentation des pics de modifications d'histones, de méthylation d'ADN et d'expression avec IGV (Integrative Genomics Viewer) pour les conditions Ctl, d0 et d21 dans les régions promotrices et géniques de *Gmfs* et *Gnas*.**

**E. Ontologie génique des promoteurs bivalents affectés ayant subi une hausse d'H3K4me3 à d21.**

**F. Ontologie génique des promoteurs bivalents affectés ayant subi une perte d'H3K27me3 à d21.**

En ce qui a trait à la méthylation d'ADN, comme mentionné brièvement précédemment, des pertes de méthylation sont observées pour certains des promoteurs bivalents à d0. Par contre, seulement 3 promoteurs ont une perte de méthylation significative à d21. Ces promoteurs, soit *Grp2*, *Nespas* et *Tktl2*, ne subissent pas de gain d'expression malgré la perte de méthylation d'ADN observée. Pour ce qui est des changements de l'expression des gènes, seuls *Arl4c*, *Cdx2*, *Tacstd2*, *Tmprss2* et *Vgll3* subissent une perte d'expression significative à d21. Parmi ceux-ci, *Cdx2* et *Vgll3* sont des gènes impliqués dans le développement embryonnaire [183, 184]. Aucun des promoteurs de ces 5 gènes n'a subi de changement significatif de la méthylation d'ADN entre Ctl et d21; *Vgll3* et *Arl4c* ont subi un gain en H3K4me3, *Cdx2* et *Tacstd2*, une perte en H3K27ac et *Tmprss2*, une perte en H3K27me3.

De façon similaire aux promoteurs actifs, le pourcentage de méthylation aux promoteurs bivalents est relativement faible, excepté pour quelques-uns ayant un pourcentage de plus de 40% (Figure 24). Le contenu en CpGs est aussi majoritairement modéré à élevé. Toutefois, on remarque que le niveau moyen d'expression des gènes sous le contrôle des promoteurs bivalents est plus faible que pour les gènes actifs pour les trois conditions expérimentales. Ceci suggère donc encore une fois que peu de pertes de méthylation d'ADN sont survenues dû au fait qu'il y a peu de méthylation au départ à ces promoteurs.



**Figure 24. Évaluation de la méthylation et de la densité en CpGs aux promoteurs bivalents ainsi que l'expression globale des gènes associés.** A) Chaque point correspond au pourcentage de méthylation d'un promoteur bivalent en condition contrôle, d0 et d21. B) La densité en CpGs de chaque promoteur bivalent a été divisé en classes de faible (0 à 2,5% du promoteur est constitué de CpGs), moyen (2.5-5%) et élevé (plus de 5% du contenu). C) La valeur normalisée d'expression pour chaque gène a été représentée pour chaque condition (ctl, d0, d21). La valeur moyenne est représentée au centre des boîtes.

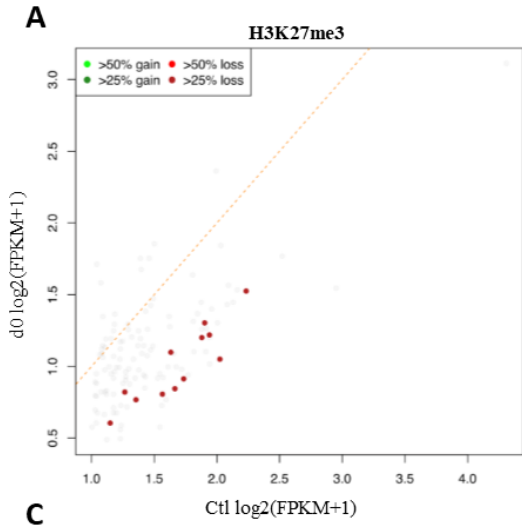
En somme, peu d'altérations de méthylation d'ADN et d'expression surviennent à long-terme (d21) suivant une perte temporaire d'expression de *Dnmt1* pour les promoteurs bivalents. Toutefois, un remaniement des modifications d'histones H3K4me3, H3K27me3 et H3K27ac est observé à d0 et d21 pour près de 8% des promoteurs bivalents totaux annotés. L'impact de ce remodelage de la chromatine sur le rôle des cellules embryonnaires est toutefois inconnu pour le moment, mais il sera intéressant à investiguer comme beaucoup des gènes sous le contrôle des promoteurs bivalents semblent importants dans le processus de développement embryonnaire.

#### 11.4.3 Promoteurs réprimés

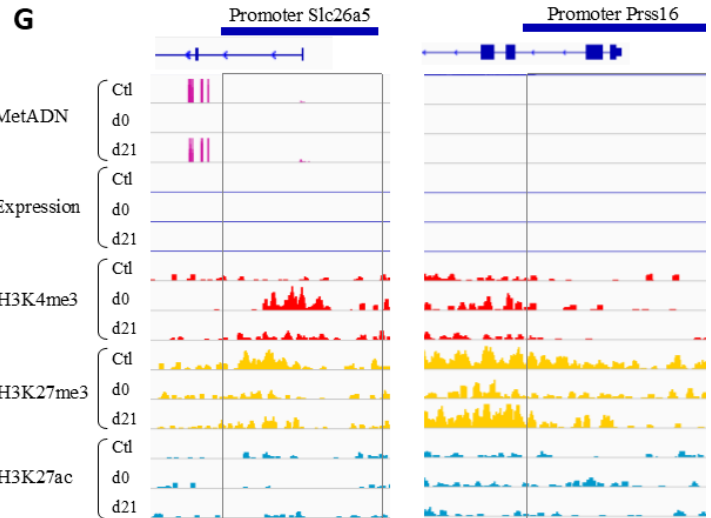
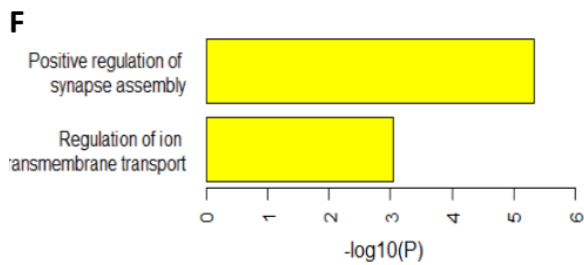
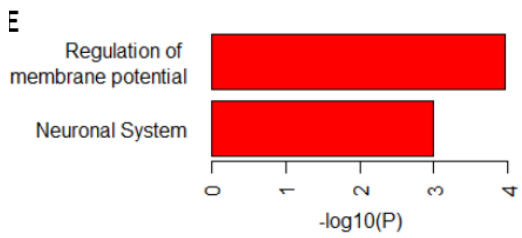
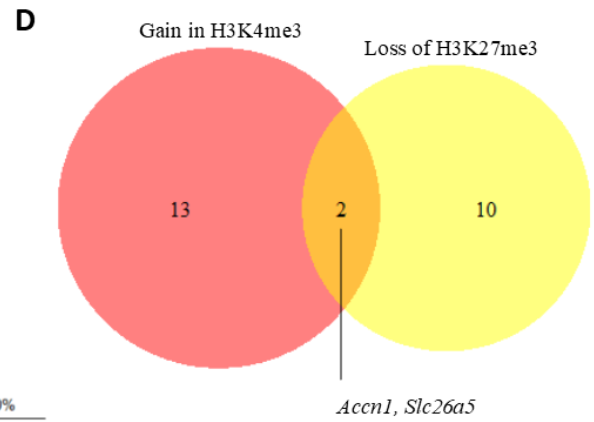
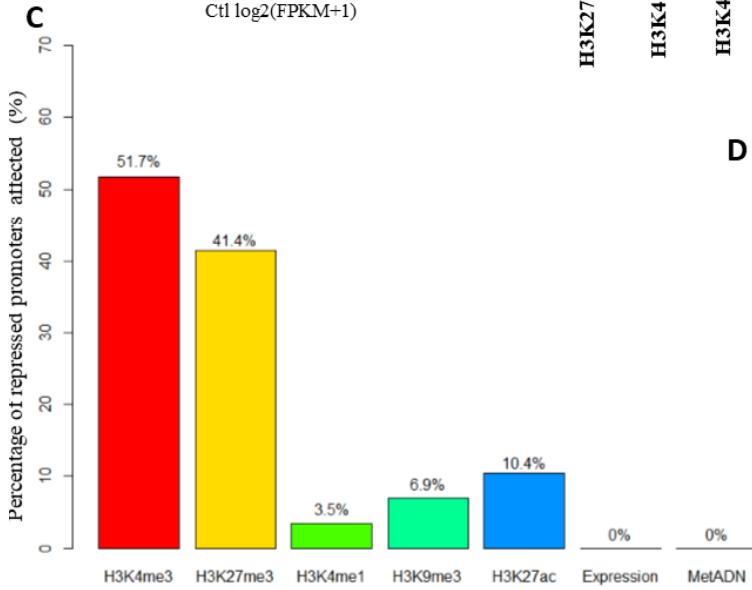
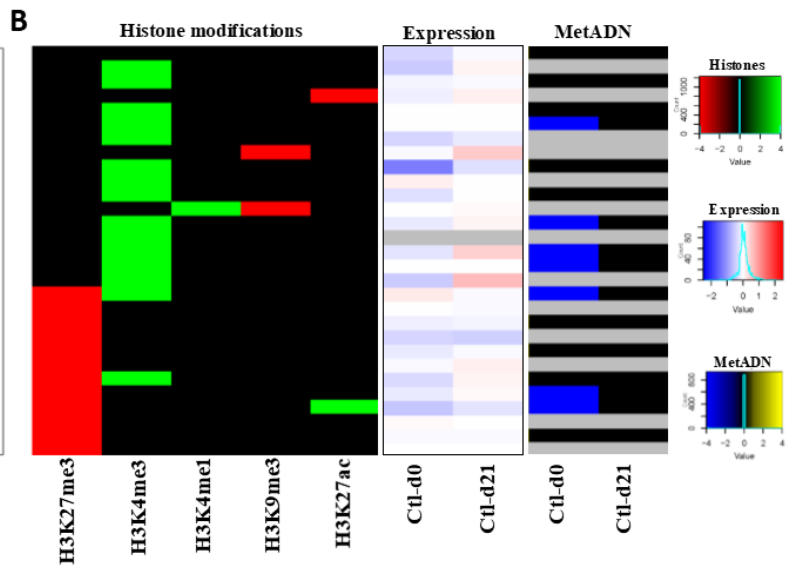
Pour l'analyse des adaptations épigénétiques survenant aux promoteurs réprimés, la modification d'histone H3K27me3 a cette fois été utilisée comme point de départ comme les promoteurs réprimés sont définis par une présence significative de cette modification d'histone et d'une absence d'H3K4me3. Un remaniement de la distribution en H3K27me3 a effectivement été observé; 12 promoteurs réprimés sur un total de 127 ont subi une perte significative de la distribution en H3K27me3 (Figure 25A). Aucun des promoteurs réprimés n'a subi de gain en H3K27me3, ce qui suggère qu'aucun n'est devenu davantage réprimé. Au total, 29 promoteurs réprimés ont été affectés pour au moins une des modifications épigénétiques à l'étude (Figure 25B). Parmi ces 29 promoteurs, 41.4 % ont subi une perte de la marque répressive H3K27me3 et 51.7% ont subi un gain en H3K4me3 (Figure 25C). Fait intéressant, seuls deux promoteurs ont subi une perte de H3K27me3 en simultané avec un gain en H3K4me3 (Figure 25D). Ces promoteurs (*Accn1* et *Slc26a5*) n'ont toutefois pas été affectés au niveau de leur méthylation d'ADN et *Slc26a5* a été en mesure de retrouver son niveau d'expression initial suite à l'inactivation et réactivation de

*Dnmt1*. En ce qui concerne *Accn1*, l'expression tend à augmenter suite à la perte et réactivation de *Dnmt1*, mais le changement demeure sous le seuil significatif (gain de 0.7fois). Ainsi, ces promoteurs réprimés ne sont pas devenus des promoteurs activant l'expression génique. Par contre, un nombre important de ces promoteurs réprimés ayant gagné en H3K4me3 ressemblent maintenant davantage à des promoteurs bivalents. Une analyse d'ontologie génique a révélé que 23% des gènes associés à ces promoteurs nouvellement bivalents (gain en H3K4me3 et maintien en H3K27me3) partagent des fonctions dans le système neuronal ( $\text{Log}_{10}(P)=3.17$ ) (*Kcnj11*, *Kcnh6*, *Adcy6*). Pour ce qui est de l'ensemble des promoteurs ayant gagné en H3K4me3, sans tenir compte des autres modifications d'histones, l'ontologie génique donne également des fonctions neuronales en plus de la régulation du potentiel de membrane (Figure 25E). Ceux ayant perdu leur marque répressive sont quant à eux liés à la régulation de la formation des synapses et du transport ionique transmembranaire (Figure 25F). Ainsi, il semble y avoir un enrichissement de fonctions liées au développement cérébral pour les gènes contrôlés par un promoteur réprimé. Deux exemples de ces promoteurs sont illustrés en figure 25G; *Prss16* et *Slc26a5*. Pour les deux promoteurs, on observe une absence de pics d'expression et même de méthylation d'ADN pour *Prss16*. La modification d'histone H3K27me3 a un niveau significatif uniquement en condition contrôle.

Repressed promoters (n=127)



Repressed promoters (n=29)



**Figure 25. Adaptation épigénétique (modifications d'histones, méthylation d'ADN et expression) aux promoteurs réprimés entre les conditions Ctl et d21 dans les cellules ES *Dnmt1*<sup>tet/tet</sup>.**

**A. Changements significatifs de la distribution de H3K27me3 aux promoteurs réprimés entre les conditions Ctl-d0 et Ctl-d21.** Chaque point (n=127) corrèle la valeur normalisée de FPKM de la condition ctl avec la valeur de la condition d0. Si un changement significatif est observé à d0 ainsi qu'à d21, les points seront colorés en vert ou en rouge. Les points verts représentent les promoteurs ayant un gain de 25% ou plus, les points rouges représentent une perte de 25% ou plus et les points gris sont les promoteurs inchangés.

**B. Adaptation de la distribution des modifications d'histones H3K4me3, H3K4me1, H3K9me3 et H3K27ac, de la méthylation d'ADN et de l'expression pour les promoteurs réprimés affectés ayant une augmentation, une perte ou pas de changement en H3K27me3.** Chaque colonne représente la valeur normalisée différentielle de présence d'une modification d'histone différente (à d0 et d21). Chaque ligne représente un promoteur. Les promoteurs ayant une augmentation de plus de 25% du « fold change » sont représentés en vert, ceux ayant une diminution de plus de 25% sont représentés en rouge et ceux étant demeurés stables sont représentés en noirs. Les promoteurs de gènes n'ayant subi aucun changement pour toutes les modifications d'histones ont été retirés des « heatmaps » (réprimés = 77%). Pour la méthylation d'ADN, un changement significatif (gain ou une perte) a été évalué à 20%. Les couleurs représentées sur la « heatmap » sont les suivantes (Gain = Jaune, Perte = Bleu, Pas de changement = Noir, Donnée non disponible = Gris). En ce qui a trait à l'expression, le seuil de significativité a été attribué à 2 « fold change » pour les gains et les pertes d'expression et les couleurs sont les suivantes (Gain = Rouge foncé, Perte = Bleu Foncé, Pas de Changement = Blanc). Pour la méthylation d'ADN et l'expression, les valeurs différentielles entre Ctl-d0 et Ctl-d21 sont comparées.

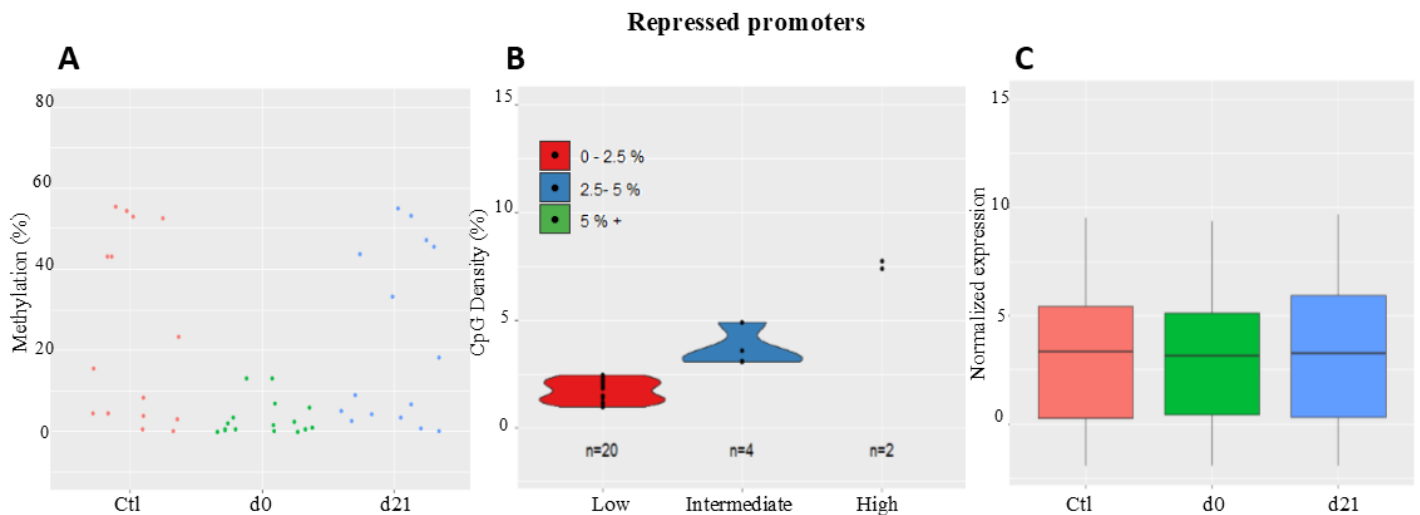
**C. Pourcentages de promoteurs réprimés affectés (total n=29) à d21 pour les modifications épigénétiques étudiées.**

**D. Diagramme de Venn représentant le croisement entre les promoteurs réprimés ayant un gain en H3K4me3 et une perte en H3K27me3.** Le promoteur *Slc26a5* a perdu ses caractéristiques réprimées tandis que *Accn1* a conservé un niveau d'H3K27me3 au-dessus du seuil de significativité de 1 FPKM.

**E et F. Ontologie génique des promoteurs réprimés affectés ayant subi une hausse d'H3K4me3 à d21 (E) et ayant subi une perte d'H3K27me3 à d21 (F).**

**G. Représentation des pics de modifications d'histones, de méthylation d'ADN et d'expression avec IGV (Integrative Genomics Viewer) pour les conditions Ctl, d0 et d21 dans les régions promotrices et géniques de *Slc26a5* et *Prss16*.**

Suite à la perte et au retour d'expression de *Dnmt1*, aucune perte permanente de méthylation d'ADN (d21) n'a été observée. En ce qui concerne l'expression, plusieurs des promoteurs réprimés des cellules ES *Dnmt1<sup>tet/tet</sup>* semblent avoir une tendance à perdre de l'expression à d0 et à en gagner suite au retour d'expression de *Dnmt1*. Toutefois, l'ensemble des changements d'expression observés à d0 et d21 sont sous le seuil significatif (< 1 fold change). La figure 26 démontre que beaucoup de promoteurs réprimés ont un niveau initial de méthylation et un contenu en CpGs faible. Toutefois, quelques-uns d'entre eux ont un pourcentage de méthylation de plus de 40% qui demeure aussi haut suivant le retour d'expression de *Dnmt1* (Figure 26A et B). L'expression des gènes réprimés est aussi à un niveau très bas comparativement aux gènes actifs et bivalents (Figure 26C).



**Figure 26. Évaluation de la méthylation et de la densité en CpGs aux promoteurs réprimés ainsi que l'expression globale des gènes associés.** **A)** Chaque point correspond au pourcentage de méthylation d'un promoteur réprimé en condition contrôle, d0 et d21. **B)** La densité en CpGs de chaque promoteur réprimé a été divisé en classes de faible (0 à 2,5% du promoteur est constitué de CpGs), moyen (2.5-5%) et élevé (plus de 5% du contenu). **C)** La valeur normalisée d'expression pour chaque gène a été représentée pour chaque condition (ctl, d0, d21). La valeur moyenne est représentée au centre des boîtes.

Pour résumer, de façon similaire aux promoteurs actifs et bivalents, les promoteurs réprimés ont subi une adaptation de certaines modifications d'histones après la perte et le

retour d'expression de *Dnmt1*. Comme les promoteurs réprimés ayant subi des réarrangements de modifications d'histones semblent avoir des implications dans les fonctions cérébrales selon les analyses d'ontologie génique, il sera intéressant d'approfondir l'analyse et de comprendre pourquoi ces gènes auraient tendance à être réprimés dans les cellules ES. Une hypothèse possible serait que ces gènes sont réprimés dans les cellules pluripotentes et activés lors de la différenciation en cellules neuronales et que la perte de DNMT1 perturbe des facteurs de transcription qui assurent le maintien de cet état réprimé lorsque les cellules sont indifférenciées [185-187].

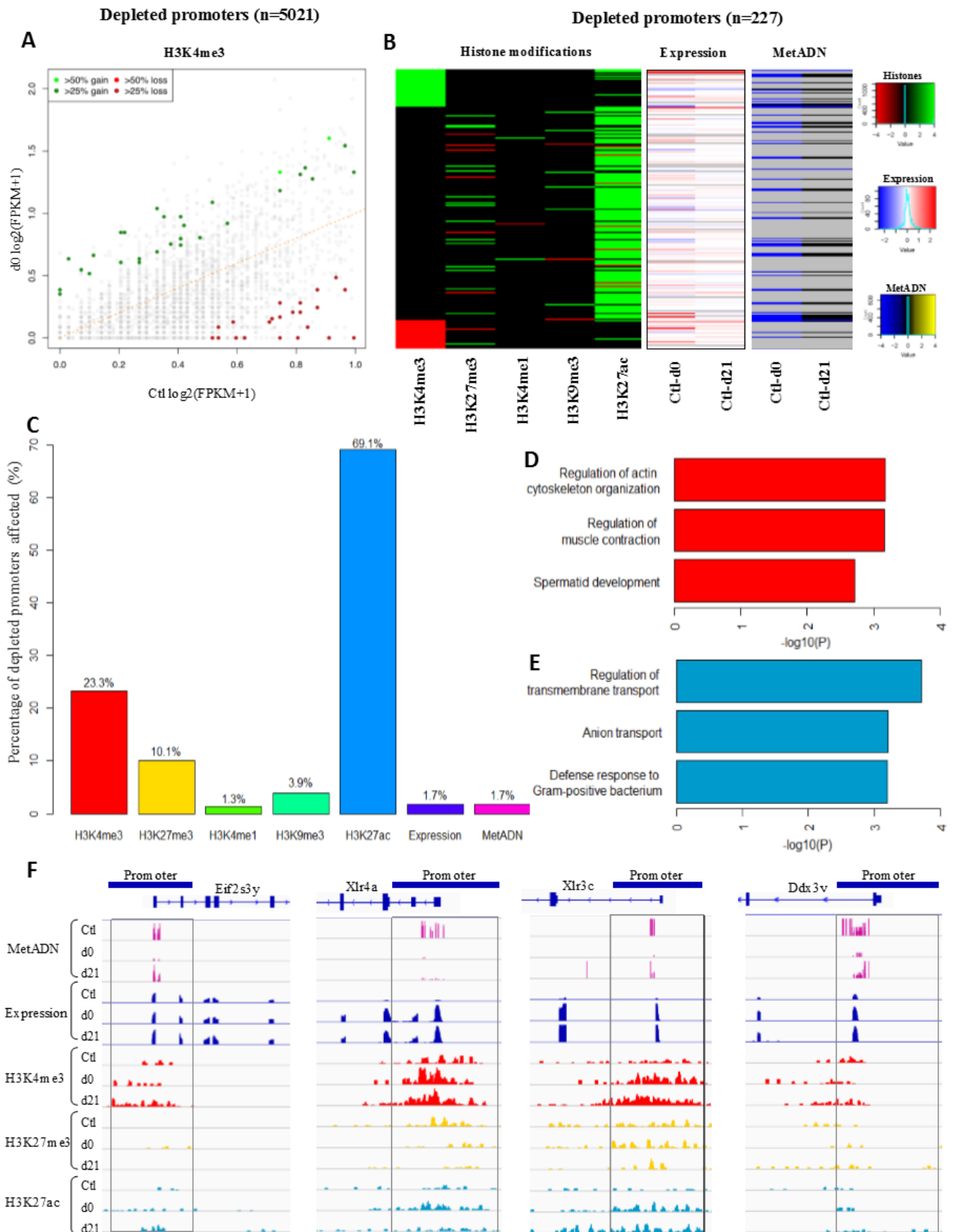
#### 11.4.4 Promoteurs déplétés

Finale­ment, nous avons porté notre attention sur les promoteurs déplétés, c'est-à-dire les promoteurs dépourvus en marques épigénétiques en condition contrôle dans les cellules ES. Le même type d'analyse que précédemment décrit a été réalisé. Le nuage de points de la Figure 27A permet d'apprécier autant un gain qu'une perte en H3K4me3 aux promoteurs déplétés. La perte en H3K4me3 est d'abord surprenante, mais elle est expliquée par le fait que certains des promoteurs étaient légèrement sous le seuil qui détermine une présence significative de H3K4me3, soit 1 FPKM, en condition contrôle. Ces promoteurs ont donc perdu le peu d'H3K4me3 qu'ils possédaient pour devenir encore plus déplétés suite à la perte et la réactivation de *Dnmt1*. Parmi les promoteurs déplétés ayant un gain d'H3K4me3 de 50% ou plus à d0 et d21 (points vert pâle), on retrouve les promoteurs déplétés de *Xlr4a* et *Xlr3c*. Parmi ceux ayant eu un gain de 25% et plus, on retrouve *Xlr4c*. En ce qui concerne les promoteurs déplétés n'ayant pas eu de changement en H3K4me3, on remarque une augmentation marquée en H3K27ac. De tous les promoteurs affectés, la hausse en H3K27ac est majoritaire et représente 69.1% de ces promoteurs affectés. Des analyses d'ontologie génique ont été réalisées pour les promoteurs ayant une hausse en H3K4me3 (Figure 27D) et une hausse en H3K27ac (Figure 27E). Les gènes liés aux promoteurs ayant une hausse en H3K4me3 suite à la réactivation de *Dnmt1* (d21) démontrent un enrichissement pour des fonctions telles que la régulation du cytosquelette d'actine et des fibres musculaires ainsi que le développement des spermatides. De l'autre côté, les gènes liés aux promoteurs enrichis en H3K27ac à d21 sont plutôt associés au transport transmembranaire (8% des gènes) ou au transport d'anions (7% des gènes). Les promoteurs



ayant perdu H3K4me3 après la perte et le retour d'expression de *Dnmt1* n'ont pas pu être liés à des fonctions géniques communes.

N'en sachant pas davantage quant aux différences entre les promoteurs déplétés ayant subi un gain d'H3K4me3 ou un gain d'H3K27ac seulement, la méthylation d'ADN et l'expression ont été investigués pour tenter de définir différents patrons d'adaptation suite à l'inactivation et réactivation de *Dnmt1*. Parmi les gènes ayant subi une perte de méthylation d'ADN significative à leur promoteur déplété à d21 (figure 27B, metADN colonne Ctl-d21), on retrouve *Eif2s3y*, *Ddx3y* ainsi que *Xlr4a* et *Xlr4c*. *Xlr3c* semble également avoir subi une perte de méthylation, mais un trop faible taux de couverture lors du séquençage empêche de confirmer. L'ensemble des *Xlrs* ont également subi un gain d'expression et un gain en H3K4me3 (et H3K27ac pour *Xlr4a*), *Ddx3y* a subi un gain d'expression seulement alors que *Eif2s3y* a conservé une expression stable et a eu un gain en H3K27ac (Figure 27F). Dans l'ensemble, aucun des gènes affectés significativement au niveau de leur expression ou de leur méthylation d'ADN n'a présenté de changement en H3K4me1 ni H3K27me3 et seul *Ddx3y* a démontré une perte en H3K9me3. Plusieurs changements significatifs d'expression ont été observés lors de la perte de *Dnmt1*, mais seuls *Xlr4a*, *Xlr3c*, *Xlr4c* et *Ddx3y* ont un gain d'expression de plus d'un « fold change » qui est maintenu à d21.



**Figure 27. Adaptation épigénétique (modifications d’histones, méthylation d’ADN et expression) aux promoteurs déplétés entre les conditions Ctl et d21 dans les cellules ES *Dnmt1*<sup>tet/tet</sup>.**

**A. Changements significatifs de la distribution de H3K4me3 aux promoteurs déplétés entre les conditions Ctl-d0 et Ctl-d21.** Chaque point (n=5021) corrèle la valeur normalisée de FPKM de la condition contrôle avec la valeur de la condition d0. Si un changement significatif est observé à d0 ainsi qu’à d21, les points seront colorés en vert ou en rouge. Les points verts représentent les promoteurs ayant un gain de 25% ou plus, les points rouges représentent une perte de 25% ou plus et les points gris sont les promoteurs inchangés.

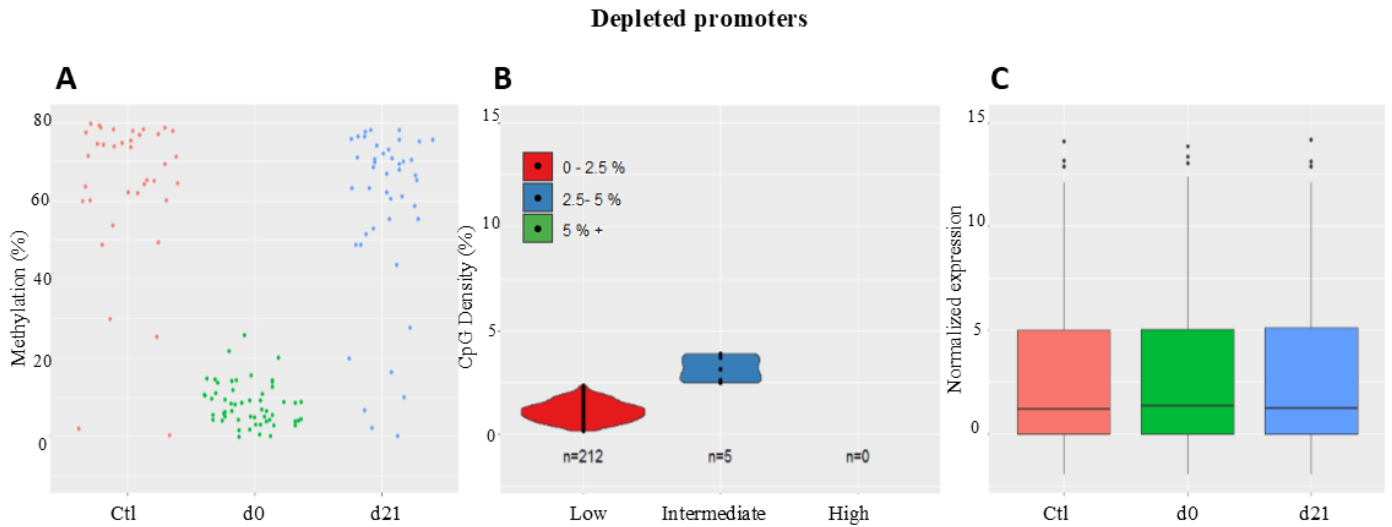
**B. Adaptation de la distribution des modifications d’histones H3K27me3, H3K4me1, H3K9me3 et H3K27ac, de la méthylation d’ADN et de l’expression pour les promoteurs déplétés affectés ayant une augmentation, une perte ou pas de changement en H3K4me3.** Chaque colonne représente la valeur normalisée différentielle de présence d’une modification d’histone différente (à d0 et d21). Chaque ligne représente un promoteur. Les promoteurs ayant une augmentation de plus de 25% du « fold change » sont représentés en vert, ceux ayant une diminution de plus de 25% sont représentés en rouge et ceux étant demeurés stables sont représentés en noirs. Les promoteurs de gènes n’ayant subi aucun changement pour toutes les modifications d’histones ont été retirés des « heatmaps » (déplétés = 95%). Pour la méthylation d’ADN, un changement significatif (gain ou une perte) a été évalué à 20%. Les couleurs représentées sur la « heatmap » sont les suivantes (Gain = Jaune, Perte = Bleu, Pas de changement= Noir, Donnée non disponible= Gris). En ce qui a trait à l’expression, le seuil de significativité a été attribué à 2 « fold change » pour les gains et les pertes d’expression et les couleurs sont les suivantes (Gain = Rouge foncé, Perte = Bleu Foncé, Pas de Changement= Blanc). Pour la méthylation d’ADN et l’expression, les valeurs différentielles entre Ctl-d0 et Ctl-d21 sont comparées.

**C. Pourcentages de promoteurs déplétés affectés (total n=227) à d21 pour les modifications épigénétiques étudiées.**

**D. Ontologie génique des promoteurs déplétés affectés ayant subi une hausse d’H3K4me3 à d21.**

**E. Ontologie génique des promoteurs déplétés affectés ayant subi une hausse d’H3K27ac à d21.**

**F. Représentation des pics de modifications d’histones, de méthylation d’ADN et d’expression avec IGV (Integrative Genomics Viewer) pour les conditions Ctl, d0 et d21 dans les régions promotrices et géniques de *Eif2s3y*, *Xlr4a*, *Xlr3c* et *Ddx3y*.**



**Figure 28. Évaluation de la méthylation et de la densité en CpGs aux promoteurs déplétés ainsi que l'expression globale des gènes associés.** **A)** Chaque point correspond au pourcentage de méthylation d'un promoteur déplété en condition contrôle, d0 et d21. **B)** La densité en CpGs de chaque promoteur déplété a été divisé en classes de faible (0 à 2,5% du promoteur est constitué de CpGs), moyen (2.5-5%) et élevé (plus de 5% du contenu). **C)** La valeur normalisée d'expression pour chaque gène a été représentée pour chaque condition (ctl, d0, d21). La valeur moyenne est représentée au centre des boîtes.

La catégorie de promoteurs déplétés est celle qui se dénote davantage du lot par rapport aux autres catégories puisque les promoteurs déplétés présentent un haut niveau de méthylation en condition contrôle, malgré leur faible contenu en CpGs. On remarque que la méthylation de ces promoteurs est grandement affectée à d0 (perte de *Dnmt1*), mais qu'elle est majoritairement rétablie après la réactivation de *Dnmt1*, excepté pour les promoteurs de *Xlr4a*, *Xlr4c*, *Baiap2l2* et *Eif2s3y*. L'expression est, quant à elle, très basse pour la majorité des gènes comme nous nous attendions, ce qui explique probablement pourquoi très peu de changements significatifs d'expression sont observés à d0 et d21 car si le niveau d'expression est déjà faible, il est possible qu'une perte survienne mais ne soit pas significative.

En somme, les promoteurs déplétés sont les promoteurs étant le plus dépourvus en modifications d'histones, mais les plus méthylés. Les changements observés chez ces

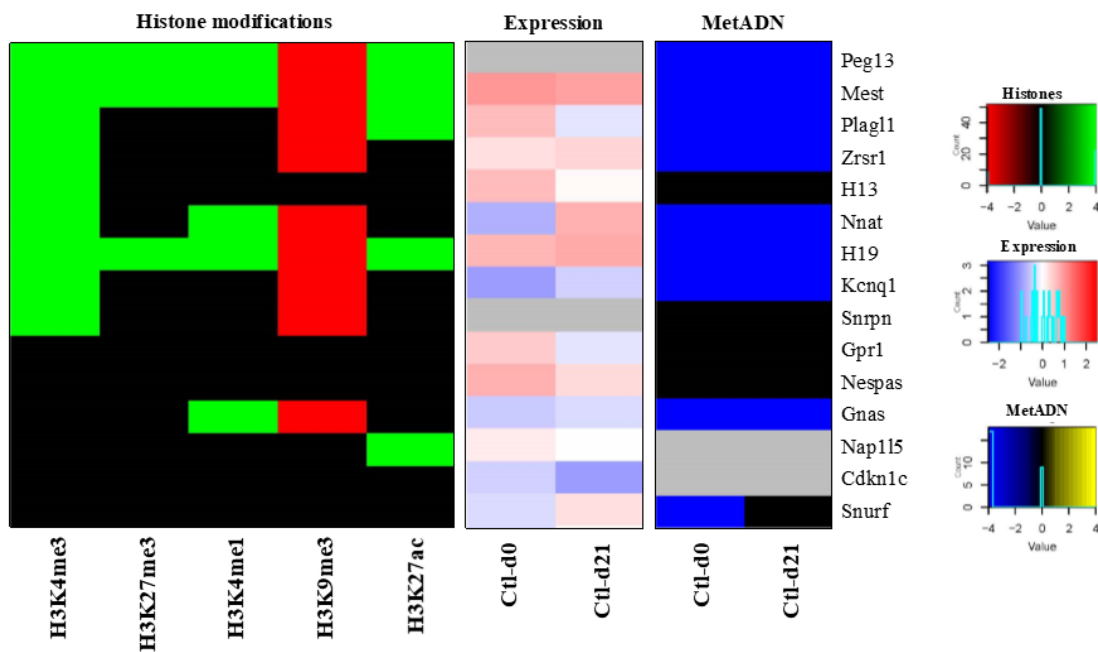
promoteurs sont donc majoritairement des gains des modification d'histones actives, soit H3K4me3 et H3K27ac.

*11.5 Adaptations épigénétiques des gènes empreintes et « imprinted-like » (Cluster *Xlr*)*  
Jusqu'à maintenant, les analyses ont été concentrées sur les adaptations épigénétiques survenant aux régions promotrices. Comme discuté précédemment, des perturbations majeures de l'épigénome lors de la perte temporaire de *Dnmt1* (d0) et des perturbations moins nombreuses suite au retour d'expression de *Dnmt1* (d21) étaient attendues et ont été observées aux régions promotrices. Certains promoteurs de gènes empreintes et « imprinted-like », particulièrement le cluster de gènes *Xlr*, ont démontré une adaptation. Toutefois, comme l'expression des gènes empreintes dépend de la région de contrôle de l'empreinte (ICR), nous avons effectué une analyse des adaptations épigénétiques survenant à ces régions pour plusieurs des gènes. Pour ce faire, nous avons utilisé la même approche qu'à la section 11.4 pour les gènes empreintes. Pour ce qui est des gènes « imprinted-like », nous avons mis l'emphase exclusivement sur le cluster *Xlr*; les données de méthylation d'ADN et de modifications d'histones ont été regroupées pour les promoteurs et les corps de gènes comme la région génomique les contrôlant n'est pas connue et nous avons utilisé le logiciel Integrative Genome Browser (IGV) pour visualiser leur adaptation épigénétique.

#### *11.5.1 Gènes empreintes*

Pour les régions empreintes, les valeurs de densités d'histones et de méthylation d'ADN ont été comparées pour les trois temps expérimentaux (Ctl, d0, d21). Pour les régions de contrôle des empreintes ayant subi un changement épigénétique significatif lors de la perte temporaire de *Dnmt1* et après le retour d'expression de *Dnmt1*, une valeur de perte ou de gain a été assignée et représentée sur la « heatmap ». Pour les modifications d'histones, un gain a été indiqué en vert, une absence de changement, en noir, et une perte, en rouge. En respectant l'ordre des régions géniques de la première colonne (H3K4me3), la valeur de « fold change » pour toutes les autres modifications d'histones ainsi que pour la méthylation d'ADN et l'expression a été attribuée pour chaque région. Les modifications d'histones sont représentées dans les mêmes couleurs alors que des couleurs différentes ont été utilisées pour la méthylation (Gain = Jaune, Perte = Bleu, Pas de changement = Noir,

Donnée non disponible= Gris) et l'expression (Gain = Rouge foncé , Perte = Bleu Foncé, Pas de Changement= Blanc) (Figure 29). La figure 29 illustre l'adaptation épigénétique globale pour les ICRs des gènes empreintes. Un peu plus de la moitié des régions de contrôle des empreintes semblent subir des réarrangements de leur paysage de modifications d'histones (gains en H3K4me3, H3K27me3, H3K4me1 et H3K27ac, perte en H3K9me3). On remarque également que près de la moitié des régions, qui correspondent majoritairement à celles ayant subi des changements en modifications d'histones, ont subi une perte de méthylation d'ADN significative en absence de *Dnmt1* qui s'est maintenue malgré la réactivation de *Dnmt1*. L'expression des gènes empreintes semble toutefois demeurer relativement stable au cours des trois temps expérimentaux. Malgré le fait que l'expression de certains des gènes empreintes semble tendre à augmenter ou diminuer à d0 et/ou d21, les valeurs se situent sous le seuil de significativité de 1 « fold change » pour l'ensemble d'entre eux.

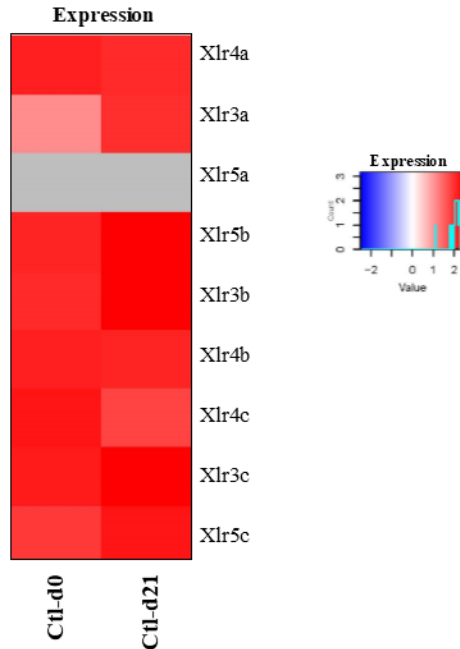


**Figure 29. Adaptations épigénétiques aux régions de contrôle des empreintes (ICRs) des gènes empreintes.** Chaque colonne représente la valeur normalisée différentielle de présence d'une modification d'histone différente (à d0 et d21) ou pour l'expression ou la méthylation d'ADN (Ctl-d0 et Ctl-d21). Chaque ligne représente une région de contrôle d'empreinte spécifique; les gènes empreintes sont indiqués sur la figure. Les valeurs représentées en gris pour l'expression et la méthylation d'ADN représentent des données indisponibles.

En somme, nous avons pu confirmer que la majorité des gènes empreintes sont affectés via leur ICR lors de la perte temporaire de *Dnmt1* au niveau de leurs modifications d'histones et/ou de la méthylation d'ADN. Par contre, les changements de l'expression de ces gènes sont moins grands qu'attendu.

#### 11.5.2 Gènes « imprinted-like »

Tel que mentionné à la section 2.2.3, les gènes « imprinted-like » sont des gènes qui peuvent agir de façon similaire à des gènes empreintes selon le tissu et dont le maintien de la méthylation d'ADN DNMT1-dépendant semble primordial. Les gènes de la famille *Xlr* ont été identifiés comme faisant partie de cette famille dans les travaux de McGraw et al. en 2015 [88]. Comme peu est connu quant aux caractéristiques et aux fonctions développementales des gènes de la famille *Xlr*, mais que ces gènes semblent drastiquement affectés suivant la perte et le retour d'expression de *Dnmt1*, nous avons analysé les modifications épigénétiques aux corps de gènes en plus des promoteurs. En utilisant la même méthode décrite plus haut, nous avons pu apprécier que les valeurs de changement à d21 attribuées pour toutes les modifications d'histones sont de 0. En effet, tous les *Xlrs* sont déplétés et stables en modifications d'histones à leurs corps de gènes à l'exception d'*Xlr3a* qui est actif (présence significative d'H3K4me3) (Résultat non montré). Les *Xlrs* semblent également dépourvus de méthylation d'ADN aux corps de gènes, cependant les données de méthylation sont indisponibles pour certains *Xlrs* et devront donc être confirmées. Toutefois, il est intéressant de voir que les *Xlrs* sont des gènes qui ont tous un gain d'expression d'au moins 2 « fold change » après la perte et la réactivation de *Dnmt1* (Figure 30).

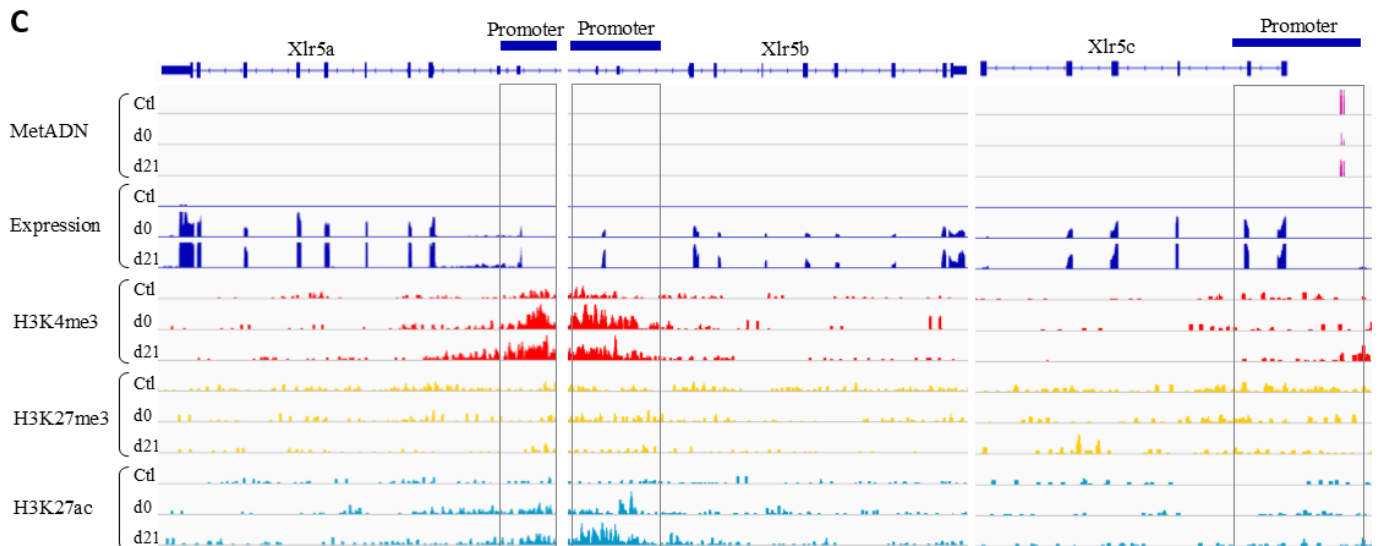
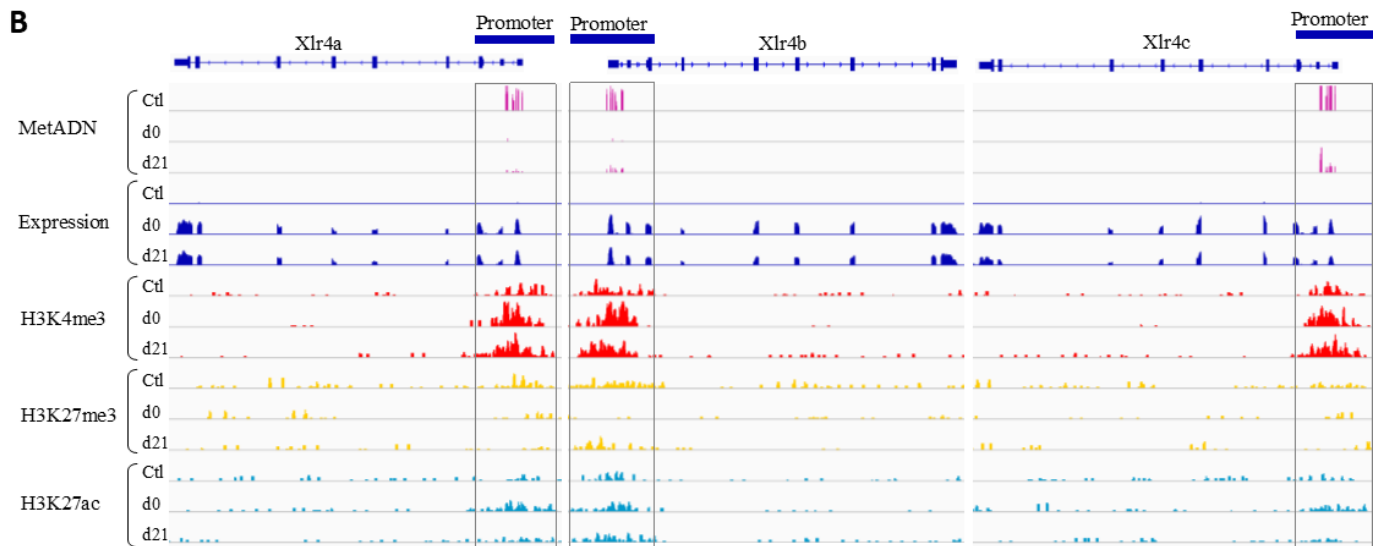
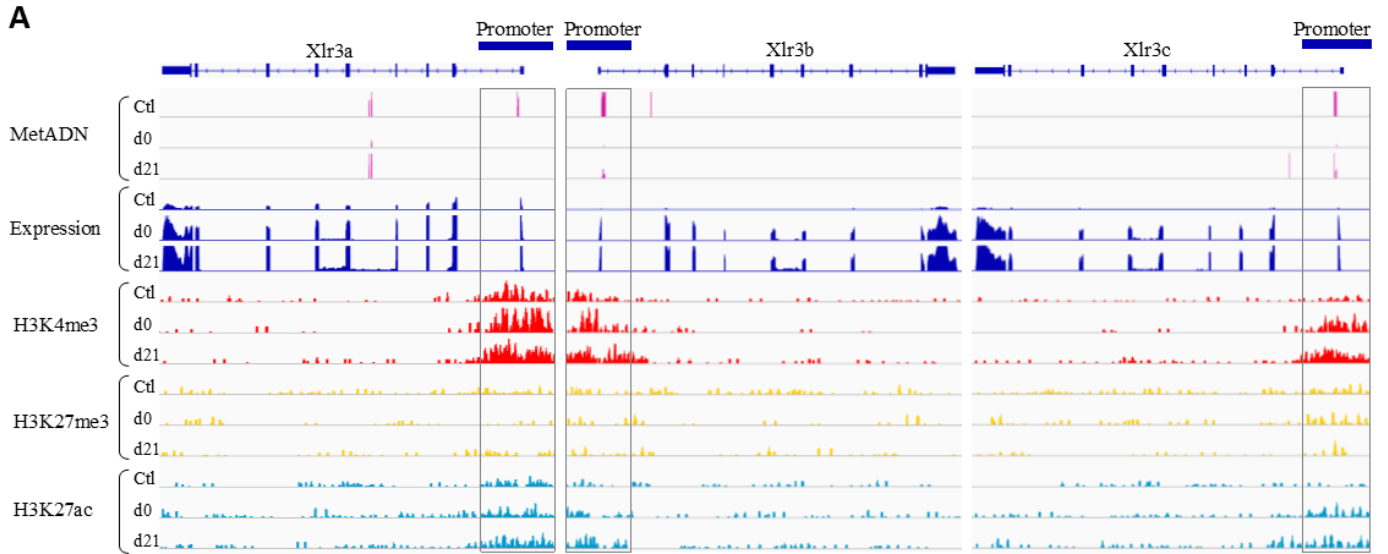


**Figure 30. Adaptation de l'expression des gènes de la famille *Xlr*.** Chaque colonne représente la valeur normalisée différentielle pour l'expression de chacun des gènes indiqués à droite de la « heatmap ». La valeur d'expression différentielle est indiquée par la légende de couleur (rouge : gain, bleu : perte, blanc : inchangé, gris : données indisponibles).

Comme il est possible de voir à la figure 31, ces gènes semblent tous perdre significativement leur méthylation d'ADN au promoteur également, excepté pour *Xlr5a* et *Xlr5b* pour lesquels les données sont indisponibles dû à un trop faible taux de couverture lors du séquençage au bisulfite. Il sera intéressant de valider que ces régions perdent toutes significativement leur méthylation après le retour d'expression de *Dnmt1*, ce qui expliquerait les gains d'expression persistants égaux ou supérieurs à 2 « fold change ». Il est intéressant aussi d'observer que les changements de distribution des modifications d'histones aux promoteurs déplétés de ces gènes (excluant *Xlr3a*) sont similaires. On observe un gain significatif d'H3K4me3 pour *Xlr3b/c* et *Xlr4a/c* et une tendance à augmenter pour les autres *Xlrs*. *Xlr3a*, quant à lui, conserve un niveau de modifications d'histones plutôt stable malgré la perte et la réactivation de *Dnmt1*, mais démontre tout de même un gain d'expression d'au moins 2 « fold change » et une perte complète de méthylation d'ADN. Ainsi, les corps de gènes des *Xlrs* n'ont pas été très informatifs pour



comprendre l'adaptation épigénétique survenant suite à l'inactivation de *Dnmt1*, mais les adaptations observées aux promoteurs sont significatives pour la méthylation d'ADN et la modification d'histone H3K4me3, corrélant avec le gain significatif d'expression. La fonction de ces gènes devra être investiguée pour vérifier s'ils ont un rôle important à jouer dans le développement embryonnaire et si leur perturbation peut avoir un impact d'ampleur.



**Figure 31. Représentation des pics de modifications d’histones, de méthylation d’ADN et d’expression avec IGV (Integrative Genomics Viewer) pour les conditions Ctl, d0 et d21 dans les régions promotrices et géniques des gènes *Xlrs*. Représentation des gènes *Xlr3a/b/c* (A), *Xlr4a/b/c* (B) et *Xlr5a/b/c* (C).**

## 12. DISCUSSION

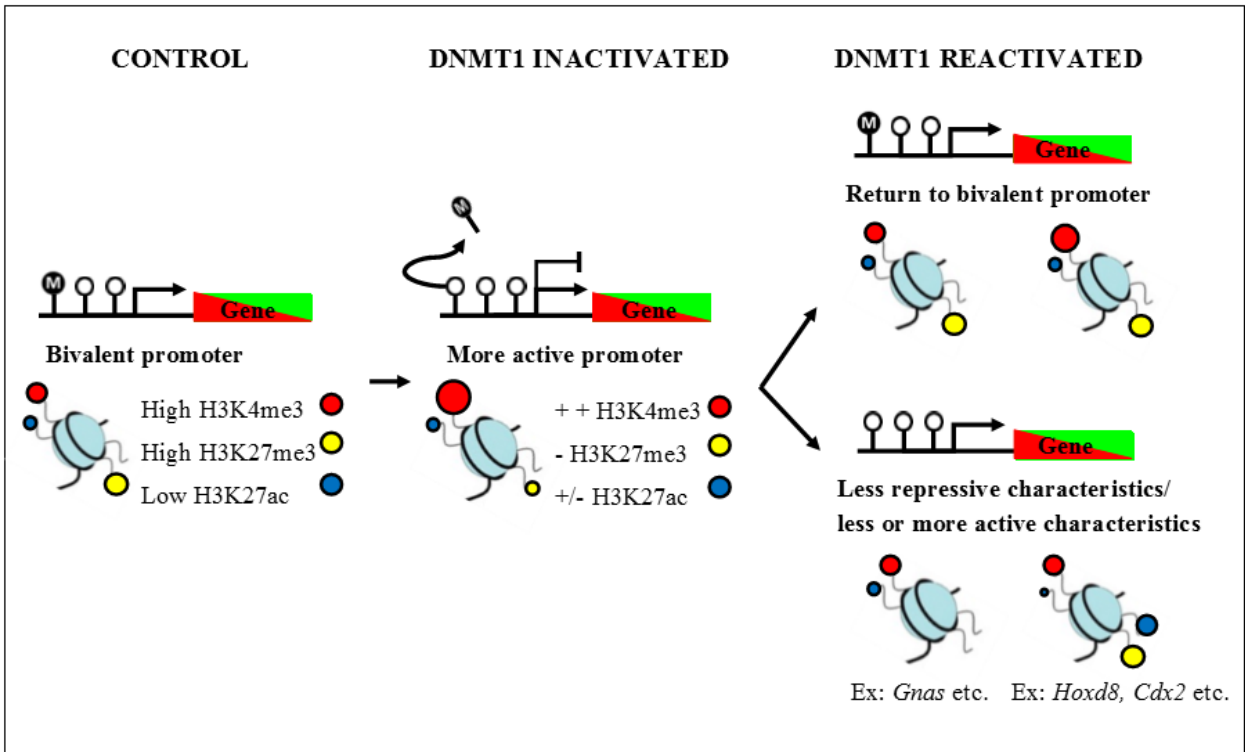
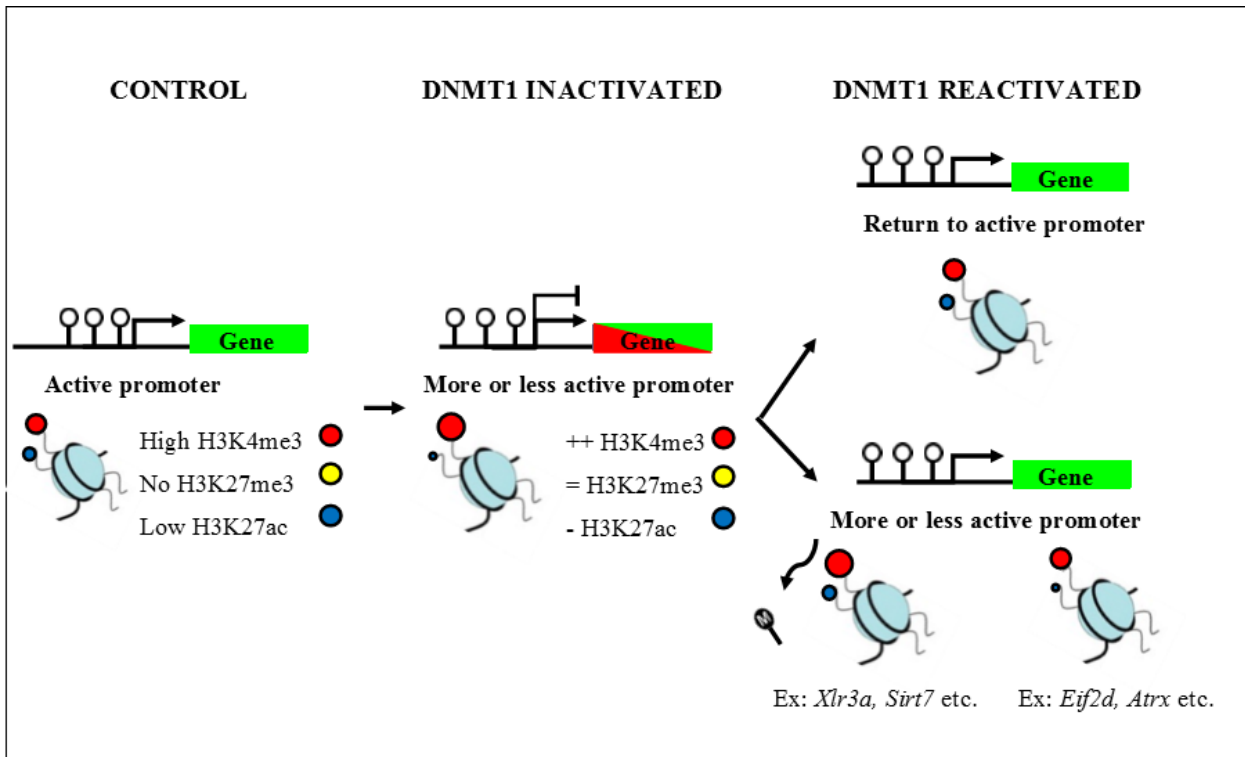
Avec cette étude, nous avons pu établir une vue globale des changements survenant dans l’épigénome avant, pendant et suivant une perte transitoire d’expression de *Dnmt1* dans un modèle embryonnaire souche. Nos résultats démontrent qu’il y a bien une adaptation erronée des modifications d’histones qui semble survenir pendant et après la perte temporaire de *Dnmt1* à certaines régions promotrices. Plusieurs altérations de méthylation d’ADN et d’expression des gènes observés lors de la perte temporaire de *Dnmt1* ont persisté suite au retour d’expression de *Dnmt1*. Toutefois, il semble y avoir peu de corrélation entre les altérations observées pour la méthylation d’ADN, l’expression des gènes et les modifications d’histones. Pour certains promoteurs, il existe possiblement une interdépendance entre certaines modifications épigénétiques qui suggère un mécanisme sous-jacent pour expliquer les altérations observées.

### *12.1 Différences d’adaptations épigénétiques entre les types de promoteurs*

Tel que démontré plus haut, les promoteurs actifs, bivalents, réprimés et déplétés ont tous à petite échelle été affectés à long-terme par une perte ou un gain d’une ou plusieurs modifications épigénétiques. Ces réarrangements épigénétiques détectés autant lors de la perte temporaire qu’après le retour d’expression de *Dnmt1*, se sont toutefois manifestés par des patrons d’adaptations un peu différents. Les figures 32 et 33 illustrent les changements épigénétiques globaux survenus respectivement pour les promoteurs actifs et bivalents puis pour les promoteurs réprimés et déplétés. Les perturbations sont majeures lors de l’inactivation de *Dnmt1* pour tous les types de promoteurs; tous semblent majoritairement gagner des caractéristiques actives telles qu’un gain d’expression, un gain en densité des marques d’histones actives (H3K4me3 et/ou H3K27ac) et une perte des marques répressives de méthylation d’ADN. Ceci est facilement expliqué par le fait que l’absence

de *Dnmt1* entraîne une dilution des patrons de méthylation d'ADN aux régions géniques répliquées, ce qui empêche donc l'épigénome de se perpétuer. Après la réactivation de *Dnmt1*, la majorité des promoteurs actifs, bivalents, réprimés et déplétés retrouvent leurs caractéristiques initiales ou presque, mais une petite proportion de promoteurs de chacune des catégories semble conserver des caractéristiques distinctes qui sont illustrées aux Figures 32 et 33. La perturbation du maintien de la méthylation d'ADN par DNMT1 engendre une augmentation d'H3K4me3 conjointement à une perte de méthylation permanente ou un maintien des caractéristiques actives avec une perte d'H3K27ac pour certains promoteurs actifs. Ceci suggère donc que la perturbation des gènes contrôlés par les promoteurs actifs est peu risquée pour le bon déroulement du développement embryonnaire puisque leur expression et méthylation au promoteur ne sont pas significativement entravés pour la grande majorité.

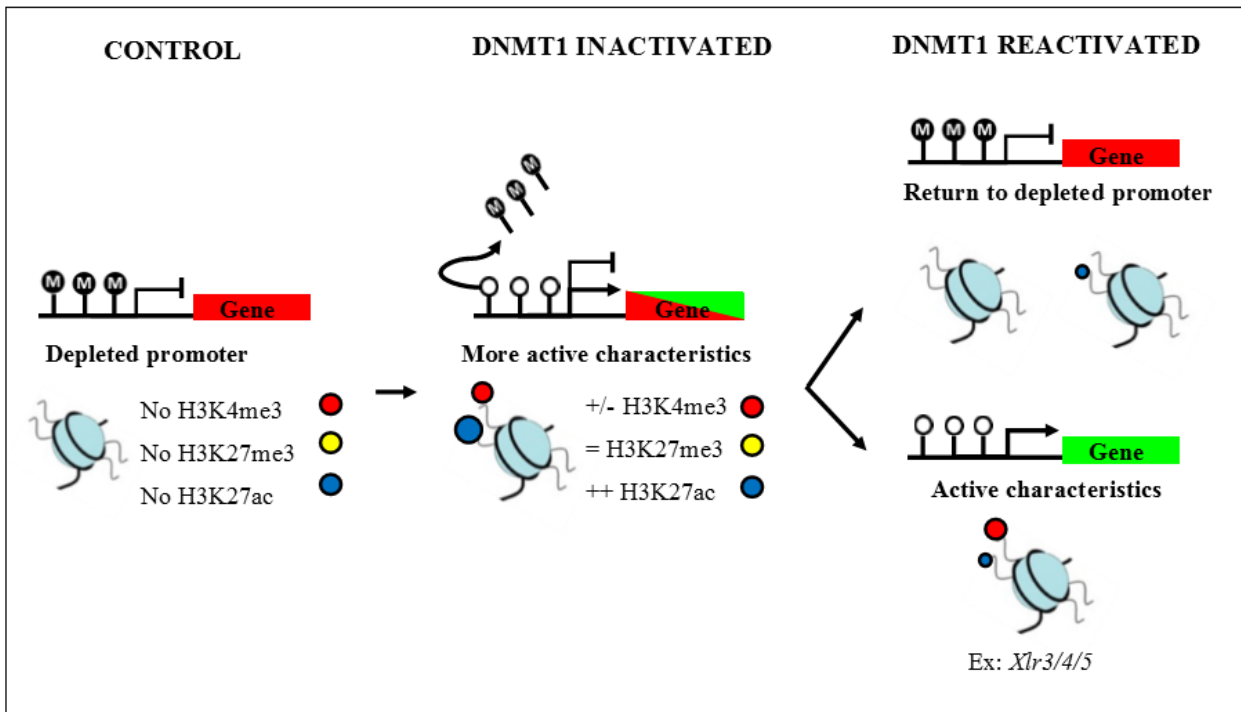
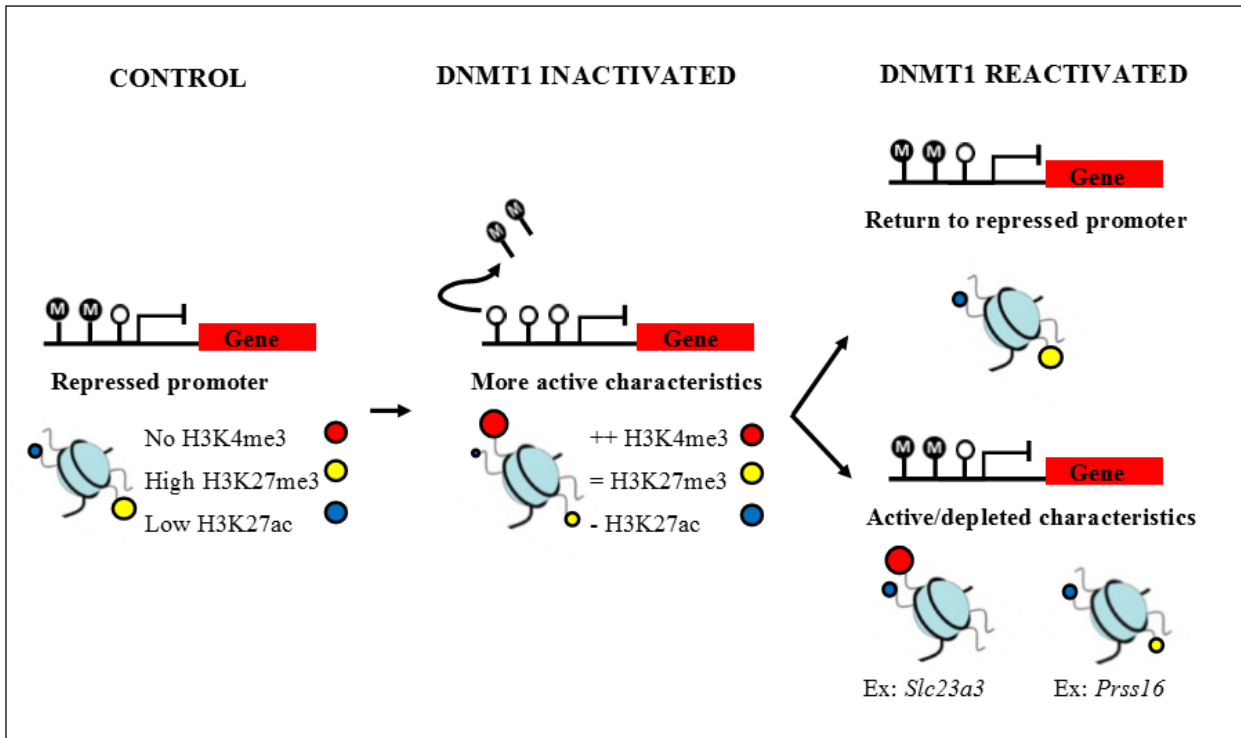
En ce qui a trait aux promoteurs bivalents, la perte et réactivation de *Dnmt1* semble permettre à certains promoteurs d'acquérir des caractéristiques actives (perte d'H3K27me3, gain d'H3K27ac) plutôt que répressives, ce qui peut suggérer une activation subséquente de gènes qui doivent généralement être maintenus dans un état répressif pour le développement adéquat de l'embryon (par exemple, des gènes codant pour des facteurs de transcription) [188]. Au stade indifférencié, cette adaptation épigénétique ne mène pas à des changements drastiques d'expression selon nos données. Toutefois, les changements de distribution des modifications d'histones observés pour les promoteurs bivalents ont possiblement davantage d'impact sur l'expression des gènes suivant la différenciation ou la spécialisation des cellules ES lors du développement embryonnaire. C'est à cette période que les promoteurs bivalents, par exemple, subissent un revirement épigénétique du côté actif ou répressif [189, 190]. Plusieurs recherches ont démontré une grande corrélation entre le statut épigénétique et le processus de différenciation [191-194]. Par exemple, il a été démontré qu'une inhibition de *Dnmt1* chez la souris a engendré des défauts de méthylation d'ADN à certaines régions enhancers se répercutant en défauts d'expression et en une différenciation retardée pour l'épithélium intestinal [195]. Ainsi, si une adaptation survient avant même l'induction de la différenciation, il est fort probable que la cellule ait une spécialisation erronée par la suite.



**Figure 32. Représentation schématique du paysage épigénétique selon la condition expérimentale (Contrôle, Inactivation de *Dnmt1*, Réactivation de *Dnmt1*) pour les promoteurs actifs et bivalents.**

Pour ce qui est des promoteurs réprimés, la tendance veut qu'il y ait un gain en H3K4me3 et/ou une perte en H3K27me3 pour la majorité de ceux qui sont affectés par la perte de maintien de la méthylation d'ADN par DNMT1 à d21. Par contre, de façon similaire aux promoteurs bivalents, les promoteurs de ces gènes ayant gagné des caractéristiques plus actives n'ont pas de gain d'expression ni de perte de méthylation drastiques. Il sera donc intéressant d'investiguer l'évolution de l'état épigénétique des promoteurs réprimés affectés et non-affectés lors du processus de différenciation.

Les promoteurs déplétés, quant à eux, ont surtout subi des gains des modifications d'histones actives (H3K4me3 et H3K27ac) dans l'ensemble. Il est intéressant de remarquer que les promoteurs déplétés les plus affectés au point de vue épigénétique (pertes de méthylation d'ADN et modification des histones) sont des promoteurs de gènes liés au sexe (ChrX ou Y) soit *Eif2s3y*, *Ddx3y* et les *Xlrs*. *Ddx3y* serait impliqué dans la différenciation neuronale puisqu'un knock-down entraîne l'apoptose et l'arrêt de la différenciation [196]. Quant à *Eif2s3y*, une implication cruciale dans la spermatogenèse a été démontrée; son absence entraîne l'infertilité chez les souris mâles [197]. Ainsi, une déstabilisation du profil épigénétique de ces gènes pourrait potentiellement entraîner des problèmes plus tard au cours du développement. En ce qui a trait aux gènes de la famille *Xlr*, le sujet sera abordé à la section 12.3.



**Figure 33. Représentation schématique du paysage épigénétique selon la condition expérimentale (Contrôle, Inactivation de *Dnmt1*, Réactivation de *Dnmt1*) pour les promoteurs réprimés et déplétés.**

Pour l'ensemble des promoteurs affectés, les résultats suggèrent que la différenciation des cellules ES en cellules somatiques permettrait de comprendre davantage l'impact de l'adaptation épigénétique sur la suite du développement embryonnaire. Il sera primordial de pousser les recherches au point de vue fonctionnel afin de comprendre comment ces adaptations épigénétiques actives vont remodeler le paysage épigénétique post-différenciation.

### *12.2 Adaptation épigénétique des gènes empreintes*

Il est connu dans la littérature qu'une perturbation de *Dnmt1* lors du développement embryonnaire précoce mène à une hypométhylation ou une perte complète de la méthylation d'ADN aux régions empreintes, accompagné d'un réarrangement des modifications d'histones H3K4me3 et d'une perte d'H3K27me3. Nous avons pu observer des résultats similaires pour certains des gènes empreintes avec l'ensemble de nos données. Après le retour d'expression de *Dnmt1*, nous avons obtenu un maintien de cette perte de méthylation d'ADN et une tendance de gain ou de perte d'expression, mais les changements d'expression ne se sont pas avérés significatifs. Les gènes empreintes ont généralement un pourcentage de méthylation élevé à leur ICR, représentant la méthylation d'un des deux allèles, soit autour de 30 à 60% dépendamment des régions et des types cellulaires [198]. Dans le cadre du projet, nous avons observé des pourcentages de méthylation d'ADN plus faibles en condition contrôle pour certains des gènes empreintes. La littérature suggère que l'hypométhylation est une caractéristique intrinsèque des cellules ES, contrairement aux cellules somatiques qui ont généralement un pourcentage de méthylation d'ADN plus élevé [199-202]. Ceci fait en sorte que certains d'entre eux n'ont pas pu démontrer une différence de méthylation significative comme la valeur différentielle de changement pour la méthylation d'ADN a été établie à une différence de 20%. Si la valeur de méthylation d'ADN était plus faible en condition contrôle avec notre modèle cellulaire, il est donc possible que les changements d'expression aient été influencés et soient donc considérés comme non significatifs. Il est également possible que les valeurs de méthylation ne permettent pas d'avoir un portrait complet de l'adaptation de méthylation puisque les analyses sont réalisées sur des tuiles de 100pb avec le RRBS, donc pas sur les régions entières. Ainsi, en prenant la tuile qui chevauchait davantage la région, il est possible que la tuile était un peu à côté de l'ICR, diminuant ainsi le pourcentage de



méthylation observé au départ. Certains articles soutiennent également que les gènes empreintes ont une susceptibilité différente aux perturbations épigénétiques et que certains démontrent donc une plus grande stabilité [203]. Il sera donc pertinent d'investiguer davantage la susceptibilité de ces régions aux perturbations afin de mieux discriminer les adaptations dans les cellules ES.

### *12.3 Adaptation épigénétique des gènes de la famille Xlr (« imprinted-like »)*

Dans le cas de la famille de gènes *Xlrs*, des gènes décrits comme « imprinted-like », nous avons observé des adaptations épigénétiques différentes de celles observées pour les gènes empreintes. En considérant les données aux promoteurs et aux corps de gènes de ceux-ci, on remarque des altérations de la méthylation d'ADN (perte) et d'expression génique (gain) pour tous les *Xlrs* (excepté quelques données non disponibles de méthylation d'ADN). L'équipe de Auclair et al. a également démontré que ces gènes, *Xlr3a/b/c* et *Xlr4a/b/c* sont fortement activés chez des embryons déficients en *Dnmt3b* [204], ce qui suggère que ces gènes sont sensibles aux perturbations de méthylation d'ADN dans le développement embryonnaire en général. Toutefois, le phénotype des embryons déficients en *Xlrs* ou ayant une surexpression des *Xlrs* n'est pas connu pour le moment. À ce jour, il est difficile également de confirmer pourquoi ces gènes ont tous une adaptation épigénétique très similaire. Sont-ils contrôlés par leurs promoteurs et ceux-ci sont affectés de façon similaire par la perte et réactivation de *Dnmt1*? Ou encore, est-il possible qu'un facteur de transcription soit activé et affecte les promoteurs de ces gènes qui sont tous situés dans la même région sur le chromosome X? Une autre hypothèse pouvant expliquer ce phénomène est qu'une région enhancer contrôle l'ensemble de ces gènes et est affectée par l'absence temporaire de *Dnmt1*. La suite de ce projet permettra donc de donner plus d'indications à ce sujet. Des essais fonctionnels seront aussi nécessaires pour être en mesure de mieux comprendre la fonction des gènes *Xlrs* et pourquoi ils semblent si susceptibles aux adaptations épigénétiques dans un contexte de développement embryonnaire.

### *12.4 Limitations de l'étude*

Une des limitations de ce projet concerne les résultats de méthylation d'ADN obtenus par RRBS. Une petite proportion des données de méthylation d'ADN n'étaient pas disponibles puisque la technique ne permet pas de couvrir tous les CpGs du génome. Ainsi, il est

possible que quelques adaptations de la méthylation d'ADN n'aient pas pu être observées. Toutefois, la technique nous permet d'avoir une bonne vision globale de l'adaptation qui survient pour la méthylation d'ADN. Afin de remédier à la situation, il serait possible d'utiliser une méthode d'analyse globale comme le Whole Genome Bisulfite Sequencing (WGBS) ou bien le pyroséquençage, afin d'aller chercher des informations plus spécifiques [205, 206].

Une autre limitation réside dans le modèle d'étude utilisé, c'est-à-dire les cellules souches embryonnaires. Comme la perte de *Dnmt1* est létale chez un modèle animal, les cellules souches sont la meilleure option pour nos études. Toutefois, il sera difficile d'obtenir des réponses aux questions fonctionnelles avec ces cellules souches comme il est impossible de les différencier sans qu'elles meurent. Un modèle de développement embryonnaire *in vivo* serait idéal pour obtenir ces informations à plus long-terme puisqu'un arrêt total de DNMT1, tel que réalisé *in vitro*, n'est probablement pas tout à fait représentatif d'une perturbation de DNMT1 *in vivo*. Toutefois, notre modèle cellulaire actuel nous permet d'obtenir suffisamment d'informations pour jeter les premières bases du projet et développer la suite. Pour la suite du projet, il sera important de considérer avoir un contrôle cellulaire qui nous permettra de valider que les effets adaptatifs observés sont les effets de la perte de DNMT1 et non les effets de la doxycycline. En effet, des évidences récentes suggèrent que les molécules de la famille des tétracyclines, dont fait partie la doxycycline, auraient des impacts significatifs sur la fonction mitochondriale dans plusieurs lignées cellulaires et chez la souris, pouvant engendrer des gains ou pertes d'expression de certains gènes mitochondriaux [207-209]. Il faudra donc prévoir cet ajout pour les expériences subséquentes avec ce modèle cellulaire. Jusqu'à maintenant, il nous a été impossible d'avoir un tel contrôle comme le niveau basal de DNMT1 dans nos cellules *Dnmt1<sup>tet/tet</sup>* est un peu plus élevé que dans les cellules « wildtype ». Il pourra être intéressant de générer un modèle animal *Dnmt1<sup>tet/tet</sup>* en modifiant le nombre d'opérateurs Tet (TetO, n=5) dans la cassette de mutagenèse afin d'obtenir une perte moins drastique de *Dnmt1*. Ce modèle nous permettrait d'avoir un contrôle adéquat (cellules ES R1 « wild type ») tel que décrit plus haut en plus de permettre l'induction de la perte temporaire de *Dnmt1* chez des cellules isolées puis différenciées *in vitro*, ou encore directement *in vivo*, afin d'observer les changements aux promoteurs et enhancers survenant lors de la différenciation.

Finalement, les résultats ont révélé que l'adaptation du paysage épigénétique des cellules ES non différenciées est beaucoup plus complexe que pensé initialement. Toutefois, ces résultats représentent une base solide afin de mieux comprendre les mécanismes survenant pendant et suivant la perte transitoire de *Dnmt1* et menant à des erreurs dans le programme embryonnaire. Il sera intéressant de définir si la perte de méthylation d'ADN ou un remaniement du paysage des modifications d'histones peut être corrigé et si cette correction sera suffisante pour contrôler la régulation de l'expression des gènes dans les cellules ES non différenciées (Revu dans [210]).

### 13. PERSPECTIVES

Tel que mentionné brièvement dans l'introduction, les promoteurs ne sont pas les seuls responsables pour la régulation de l'expression des gènes. En effet, les régions enhancers, plus distales et moins bien définies, ont aussi une forte influence sur un ou plusieurs gènes à la fois. Pour être en mesure de mieux comprendre les adaptations épigénomiques survenant suivant une perte transitoire de *Dnmt1*, les changements survenant aux régions enhancers devront aussi être adressés. Cette portion d'analyse permettra de déterminer comment la régulation indirecte de l'expression par les enhancers est affectée par une perte transitoire du maintien de la méthylation d'ADN DNMT1-dépendante. Notre hypothèse veut que plusieurs adaptations épigénétiques surviennent aux régions enhancers en plus des adaptations observées aux régions promotrices. Les résultats de ChIP-Seq pour les modifications d'histones qui serviront à la définition des régions enhancers (H3K4me1, H3K27me3, H3K27ac) étant déjà générés, l'annotation des régions a déjà pu être débutée et l'analyse des résultats sera réalisée dans les prochains mois. Nous serons donc en mesure d'obtenir un portrait encore plus précis des adaptations épigénétiques survenant dû à une absence de maintien de méthylation d'ADN par DNMT1.

Une fois ces informations recueillies, nous pourrons tenter de décortiquer le mécanisme d'adaptation en analysant directement la dépendance des modifications épigénétiques avec l'expression des gènes et la méthylation d'ADN. Pour ce faire, un système d'édition de l'épigénome pourra être utilisé pour spécifiquement rétablir la méthylation d'ADN aux

régions ayant perdu leurs patrons de méthylation (entre autres aux régions empreintes et « imprinted-like »). Le système permettant un tel « rescue » est le plasmide CRISPR dCas9-DNMT3A [211]. Ce plasmide peut être inséré dans les cellules ES par transfection avec des ARN guides (gARNs) permettant de cibler les régions d'intérêt ayant perdu leur méthylation. Utilisant une Cas9 désactivée couplée avec l'enzyme DNMT3A, des groupements méthyles seront ajoutés aux régions ciblées par les gARNs via ce plasmide, mais la Cas9 n'induit pas de coupure dans l'ADN ciblé. Des gARNs ont déjà été créés pour différentes régions telles que *H19*, *Peg13* ou les différents *Xlrs* et seront transfectés avec le plasmide dCas9-DNMT3A-EGFP (Addgene #71666). Cette technique permettra donc de récolter des cellules ES dans lesquelles la correction des patrons de méthylation d'ADN a été effectuée à des régions précises. À partir de ces cellules nous pourrions évaluer si le rétablissement de la méthylation d'ADN est suffisant pour rétablir l'expression des gènes et la distribution de modifications d'histones aux niveaux initiaux ou non. Cette information permettra donc d'avoir une meilleure idée du facteur qui intervient en premier dans le mécanisme, c'est-à-dire soit la méthylation d'ADN ou les modifications des histones.

Il sera aussi intéressant d'investiguer l'interaction de facteurs de liaison directs et indirects de DNMT1 puisqu'il est possible que la perte temporaire de *Dnmt1* entrave la liaison d'un autre facteur ou d'une modification d'histone qui interagit avec le complexe de recrutement de DNMT1. Ainsi, des facteurs tels que ZFP57, G9a et UHRF1, ayant des implications diverses dans le remodelage de la chromatine, le recrutement de DNMT1 ou encore la méthylation des histones, devront être étudiés davantage afin de mieux comprendre le mécanisme d'adaptation épigénétique et de contrer les troubles du développement pouvant être associés à ces dérèglements épigénétiques chez l'embryon précoce.

Il sera également intéressant d'investiguer l'adaptation de la distribution du variant d'histone H2A.Z dans notre modèle puisque ce variant a été démontré comme étant anti-corrélé avec la méthylation d'ADN et essentiel à la régulation épigénétique dans le développement embryonnaire [25, 212, 213]. Cette partie de projet présentement en cours permettra de valider si les régions ayant perdu des marques de méthylation d'ADN gagnent

du variant H2A.Z et si la présence du variant peut empêcher le retour au niveau initial de méthylation après rétablissement de l'expression de *Dnmt1*.

## 14. CONCLUSIONS

Le développement embryonnaire est un processus hautement dynamique et complexe d'un point de vue épigénétique et la moindre perturbation peut avoir des impacts faibles à fulgurants sur la santé d'un individu. Il est donc nécessaire de comprendre les mécanismes d'interactions de ces modifications épigénétiques afin de pouvoir prévenir des changements permanents néfastes. Pour ce faire, notre modèle ES *Dnmt1*<sup>tet/tet</sup> va nous permettre d'isoler les événements épigénétiques associés à un dérèglement héréditaire de méthylation d'ADN suite à une perte transitoire de *Dnmt1* dans les cellules embryonnaires. À ce jour, nous avons pu démontrer que les régions promotrices sont grandement remodelées lors de la perte de *Dnmt1* et que certaines des perturbations persistent à travers le temps. Ces dernières touchent autant les promoteurs actifs, bivalents, réprimés que déplétés et semblent contribuer à des caractéristiques d'activation aux régions promotrices. Comme il est même probable que les perturbations observées chez les cellules souches soient amplifiées ou différentes dans un contexte de différenciation subséquent, il est crucial de pousser les recherches pour acquérir de nouveaux indices quant à l'instigateur de ces altérations épigénomiques et sur l'impact de ces dernières à court et long-terme. Avec la suite de ce projet et l'évolution de notre modèle d'étude, nous allons donc être mieux positionnés pour répondre à ces questions mécanistiques jusqu'à maintenant irrésolues concernant le dérèglement héréditaire des régions empreintes et « imprinted-like » dans les cellules embryonnaires ainsi qu'accroître nos connaissances sur l'interaction existant entre la méthylation d'ADN et les modifications d'histones lors du processus de reprogrammation embryonnaire. Ces connaissances futures permettront de mieux comprendre la sensibilité épigénétique rencontrée aux régions régulatrices du génome lors du développement embryonnaire précoce.

## 15. CONTRIBUTIONS

J'ai moi-même procédé à la mise en culture du modèle cellulaire *Dnmt1<sup>tet/tet</sup>*, à la récolte des cellules et à la fixation de ces dernières, aux ChIPs d'H3K9me3 et d'H3K4me1 en plus de la préparation des librairies puis j'ai validé les enrichissements de tous les ChIPs et la perte de DNMT1 à d0. J'ai ensuite participé à la définition des promoteurs, j'ai procédé à l'analyse du profil épigénétiques des promoteurs à partir des données générées par les bioinformaticiens puis j'ai procédé à la génération des graphiques de résultats (Figures 14, 15, 16, 18-20, 21c/d/e/f, 22, 23c/d/e/f, 24, 25c/d/e/f, 26, 27c/d/e/f, 28, 31-33), au montage des figures et à la modification des graphiques de résultats bruts.

Lisa-Marie Legault a procédé aux expériences de ChIP pour H3K4me3, H3K27me3 et H3K27ac avant mon arrivée au laboratoire, mais n'a pas été impliquée dans l'analyse des résultats présentés dans mon mémoire ni dans le montage des figures. Ces expériences de ChIP ont servi subséquemment à l'analyse et l'élaboration des figures suivantes : 17, 18, 20, 21, 23, 25, 27, 29, 31.

Maxime Caron a procédé au traitement et à l'analyse des données brutes de RRBS et RNA-Seq à la base des figures 19b, 21e/f, 22, 24, 26, 28, 30. Il a aussi participé à la définition des promoteurs et au traitement des données de ChIP-Seq (Supervision d'Anthony Lemieux pour les figures 17, 18, 21a/b, 23a/b, 25a/b, 27a/b, 29, 30).

Anthony Lemieux (sous la supervision de M. Caron) a procédé au traitement et à l'analyse des données brutes de ChIP-Seq, a participé à la définition des promoteurs et a généré des graphiques de résultats (Figures 17, 18, 21a/b, 23a/b, 25a/b, 27a/b, 29, 30).

## 16. BIBLIOGRAPHIE

1. Waddington, C., *Epigenetics and evolution*. Symp Soc Exp Biol 1953. **7**: p. 186–199.
2. Riggs, A. and T. Porter, *Overview of epigenetic mechanisms*. Cold Spring Harbor Laboratory Press, 1996: p. pp. 29–45.
3. Weinhold, B., *Epigenetics: the science of change*. Environ Health Perspect, 2006. **114**(3): p. A160-7.
4. Cantone, I. and A.G. Fisher, *Epigenetic programming and reprogramming during development*. Nat Struct Mol Biol, 2013. **20**(3): p. 282-9.
5. NHGRI. *Genetic Illustrations*. 2018 01-06-2018]; Available from: [http://www.nhgri.nih.gov/DIR/VIP/Learning\\_Tools/genetic\\_illustrations.html](http://www.nhgri.nih.gov/DIR/VIP/Learning_Tools/genetic_illustrations.html)
6. Ernst, J., et al., *Mapping and analysis of chromatin state dynamics in nine human cell types*. Nature, 2011. **473**(7345): p. 43-9.
7. He, B., et al., *Global view of enhancer-promoter interactome in human cells*. Proc Natl Acad Sci U S A, 2014. **111**(21): p. E2191-9.
8. Aran, D., S. Sabato, and A. Hellman, *DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes*. Genome Biol, 2013. **14**(3): p. R21.
9. Muller, M.M., T. Gerster, and W. Schaffner, *Enhancer sequences and the regulation of gene transcription*. Eur J Biochem, 1988. **176**(3): p. 485-95.
10. Kim, T.K. and R. Shiekhattar, *Architectural and Functional Commonalities between Enhancers and Promoters*. Cell, 2015. **162**(5): p. 948-59.
11. Cnx.org. *Gene Regulation: Operon Theory*. 2018 01-06-2018]; Available from: <https://cnx.org/contents/5CvTdmJL@4.18:LeWYNHAI@3/Gene-Regulation-Operon-Theory>.
12. Kornberg, R.D. and Y. Lorch, *Chromatin structure and transcription*. Annu Rev Cell Biol, 1992. **8**: p. 563-87.
13. Finch, J.T. and A. Klug, *Solenoidal model for superstructure in chromatin*. Proc Natl Acad Sci U S A, 1976. **73**(6): p. 1897-901.
14. Widom, J., *Toward a unified model of chromatin folding*. Annu Rev Biophys Chem, 1989. **18**: p. 365-95.
15. Cutter, A.R. and J.J. Hayes, *A brief review of nucleosome structure*. FEBS Lett, 2015. **589**(20 Pt A): p. 2914-22.
16. Kato, S., T. Ishii, and A. Kouzmenko, *Point mutations in an epigenetic factor lead to multiple types of bone tumors: role of H3.3 histone variant in bone development and disease*. Bonekey Rep, 2015. **4**: p. 715.
17. Lorch, Y., J.W. LaPointe, and R.D. Kornberg, *Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones*. Cell, 1987. **49**(2): p. 203-10.
18. Zhou, W., et al., *Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation*. Mol Cell, 2008. **29**(1): p. 69-80.
19. Jimeno-Gonzalez, S., et al., *Defective histone supply causes changes in RNA polymerase II elongation rate and cotranscriptional pre-mRNA splicing*. Proc Natl Acad Sci U S A, 2015. **112**(48): p. 14840-5.

20. Venkatesh, S. and J.L. Workman, *Histone exchange, chromatin structure and the regulation of transcription*. Nat Rev Mol Cell Biol, 2015. **16**(3): p. 178-89.
21. Marzluff, W.F., E.J. Wagner, and R.J. Duronio, *Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail*. Nat Rev Genet, 2008. **9**(11): p. 843-54.
22. Mattioli, F., S. D'Arcy, and K. Luger, *The right place at the right time: chaperoning core histone variants*. EMBO Rep, 2015. **16**(11): p. 1454-66.
23. Turinetto, V. and C. Giachino, *Histone variants as emerging regulators of embryonic stem cell identity*. Epigenetics, 2015. **10**(7): p. 563-73.
24. Rogakou, E.P., et al., *Megabase chromatin domains involved in DNA double-strand breaks in vivo*. J Cell Biol, 1999. **146**(5): p. 905-16.
25. Zilberman, D., et al., *Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks*. Nature, 2008. **456**(7218): p. 125-9.
26. Paulson, J.R. and U.K. Laemmli, *The structure of histone-depleted metaphase chromosomes*. Cell, 1977. **12**(3): p. 817-28.
27. Woodcock, C.L. and R.P. Ghosh, *Chromatin higher-order structure and dynamics*. Cold Spring Harb Perspect Biol, 2010. **2**(5): p. a000596.
28. Openoregon.pressbooks.pub. *Principles of biology: Eukaryotic epigenetic regulation*. 2018 [12-10-2018]; Available from: <https://openoregon.pressbooks.pub/mhccmajorsbio/chapter/what-is-the-epigenome/>.
29. ThermoFisher. *Epigenetic antibodies: Histone modifications*. 2018 [01-06-2018]; Available from: <https://www.thermofisher.com/ca/en/home/life-science/antibodies/primary-antibodies/research-area-antibodies/epigenetic-antibodies.html>.
30. Josling, G.A., et al., *The role of bromodomain proteins in regulating gene expression*. Genes (Basel), 2012. **3**(2): p. 320-43.
31. Palacios, A., et al., *Solution structure and NMR characterization of the binding to methylated histone tails of the plant homeodomain finger of the tumour suppressor ING4*. FEBS Lett, 2006. **580**(30): p. 6903-8.
32. Shi, X., et al., *ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression*. Nature, 2006. **442**(7098): p. 96-9.
33. Bannister, A.J. and T. Kouzarides, *Regulation of chromatin by histone modifications*. Cell Res, 2011. **21**(3): p. 381-95.
34. Riddle, N.C., et al., *Plasticity in patterns of histone modifications and chromosomal proteins in Drosophila heterochromatin*. Genome Res, 2011. **21**(2): p. 147-63.
35. Muramatsu, D., et al., *Pericentric H3K9me3 Formation by HP1 Interaction-defective Histone Methyltransferase Suv39h1*. Cell Struct Funct, 2016. **41**(2): p. 145-152.
36. Cedar, H. and Y. Bergman, *Linking DNA methylation and histone modification: patterns and paradigms*. Nat Rev Genet, 2009. **10**(5): p. 295-304.
37. Bird, A.P., *CpG-rich islands and the function of DNA methylation*. Nature, 1986. **321**(6067): p. 209-13.
38. Antequera, F., *Structure, function and evolution of CpG island promoters*. Cell Mol Life Sci, 2003. **60**(8): p. 1647-58.



39. Deaton, A.M. and A. Bird, *CpG islands and the regulation of transcription*. *Genes Dev*, 2011. **25**(10): p. 1010-22.
40. Trasler, J.M., *Gamete imprinting: setting epigenetic patterns for the next generation*. *Reprod Fertil Dev*, 2006. **18**(1-2): p. 63-9.
41. Siegfried, Z. and I. Simon, *DNA methylation and gene expression*. *Wiley Interdiscip Rev Syst Biol Med*, 2010. **2**(3): p. 362-371.
42. Tahiliani, M., et al., *Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1*. *Science*, 2009. **324**(5929): p. 930-5.
43. Maresca, A., et al., *DNA methyltransferase 1 mutations and mitochondrial pathology: is mtDNA methylated?* *Front Genet*, 2015. **6**: p. 90.
44. Bourc'his, D., et al., *Dnmt3L and the establishment of maternal genomic imprints*. *Science*, 2001. **294**(5551): p. 2536-9.
45. Guenatri, M., et al., *Plasticity in Dnmt3L-dependent and -independent modes of de novo methylation in the developing mouse embryo*. *Development*, 2013. **140**(3): p. 562-72.
46. Barau, J., et al., *The DNA methyltransferase DNMT3C protects male germ cells from transposon activity*. *Science*, 2016. **354**(6314): p. 909-912.
47. Okano, M., et al., *DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development*. *Cell*, 1999. **99**(3): p. 247-57.
48. Webster, K.E., et al., *Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis*. *Proc Natl Acad Sci U S A*, 2005. **102**(11): p. 4068-73.
49. Hermann, A., S. Schmitt, and A. Jeltsch, *The human Dnmt2 has residual DNA-(cytosine-C5) methyltransferase activity*. *J Biol Chem*, 2003. **278**(34): p. 31717-21.
50. Tuorto, F., et al., *The tRNA methyltransferase Dnmt2 is required for accurate polypeptide synthesis during haematopoiesis*. *Embo j*, 2015. **34**(18): p. 2350-62.
51. Jeltsch, A., et al., *Mechanism and biological role of Dnmt2 in Nucleic Acid Methylation*. *RNA Biol*, 2017. **14**(9): p. 1108-1123.
52. Jeltsch, A., W. Nellen, and F. Lyko, *Two substrates are better than one: dual specificities for Dnmt2 methyltransferases*. *Trends Biochem Sci*, 2006. **31**(6): p. 306-8.
53. Hirasawa, R., et al., *Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development*. *Genes Dev*, 2008. **22**(12): p. 1607-16.
54. Cirio, M.C., et al., *Preimplantation expression of the somatic form of Dnmt1 suggests a role in the inheritance of genomic imprints*. *BMC Dev Biol*, 2008. **8**: p. 9.
55. Robertson, K.D., et al., *DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters*. *Nat Genet*, 2000. **25**(3): p. 338-42.
56. Bashtrykov, P., et al., *Specificity of Dnmt1 for methylation of hemimethylated CpG sites resides in its catalytic domain*. *Chem Biol*, 2012. **19**(5): p. 572-8.
57. Song, J., et al., *Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation*. *Science*, 2011. **331**(6020): p. 1036-40.
58. Ha, K., et al., *Rapid and transient recruitment of DNMT1 to DNA double-strand breaks is mediated by its interaction with multiple components of the DNA damage response machinery*. *Hum Mol Genet*, 2011. **20**(1): p. 126-40.

59. Jin, B. and K.D. Robertson, *DNA methyltransferases, DNA damage repair, and cancer*. Adv Exp Med Biol, 2013. **754**: p. 3-29.
60. Hashimoto, S., H. Anai, and K. Hanada, *Mechanisms of interstrand DNA crosslink repair and human disorders*. Genes Environ, 2016. **38**: p. 9.
61. Trowbridge, J.J., et al., *DNA methyltransferase 1 is essential for and uniquely regulates hematopoietic stem and progenitor cells*. Cell Stem Cell, 2009. **5**(4): p. 442-9.
62. Avgustinova, A. and S.A. Benitah, *Epigenetic control of adult stem cell function*. Nat Rev Mol Cell Biol, 2016. **17**(10): p. 643-58.
63. Rose, N.R. and R.J. Klose, *Understanding the relationship between DNA methylation and histone lysine methylation*. Biochim Biophys Acta, 2014. **1839**(12): p. 1362-72.
64. Rottach, A., et al., *The multi-domain protein Np95 connects DNA methylation and histone modification*. Nucleic Acids Res, 2010. **38**(6): p. 1796-804.
65. Liu, X., et al., *UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9*. Nat Commun, 2013. **4**: p. 1563.
66. Sharif, J., et al., *The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA*. Nature, 2007. **450**(7171): p. 908-12.
67. Myant, K. and I. Stancheva, *LSH cooperates with DNA methyltransferases to repress transcription*. Mol Cell Biol, 2008. **28**(1): p. 215-26.
68. Hervouet, E., et al., *Specific or not specific recruitment of DNMTs for DNA methylation, an epigenetic dilemma*. Clin Epigenetics, 2018. **10**: p. 17.
69. Harrison, J.S., et al., *Hemi-methylated DNA regulates DNA methylation inheritance through allosteric activation of H3 ubiquitylation by UHRF1*. Elife, 2016. **5**.
70. Vaughan, R.M., et al., *Comparative biochemical analysis of UHRF proteins reveals molecular mechanisms that uncouple UHRF2 from DNA methylation maintenance*. Nucleic Acids Res, 2018. **46**(9): p. 4405-4416.
71. Rasmussen, K.D. and K. Helin, *Role of TET enzymes in DNA methylation, development, and cancer*. Genes Dev, 2016. **30**(7): p. 733-50.
72. Mayer, W., et al., *Demethylation of the zygotic paternal genome*. Nature, 2000. **403**(6769): p. 501-2.
73. Oswald, J., et al., *Active demethylation of the paternal genome in the mouse zygote*. Curr Biol, 2000. **10**(8): p. 475-8.
74. Guo, F., et al., *Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote*. Cell Stem Cell, 2014. **15**(4): p. 447-459.
75. Weaver, J.R. and M.S. Bartolomei, *Chromatin regulators of genomic imprinting*. Biochim Biophys Acta, 2014. **1839**(3): p. 169-77.
76. Bartolomei, M.S. and S.M. Tilghman, *Genomic imprinting in mammals*. Annu Rev Genet, 1997. **31**: p. 493-525.
77. Terranova, R., et al., *Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos*. Dev Cell, 2008. **15**(5): p. 668-79.
78. Edwards, C.A. and A.C. Ferguson-Smith, *Mechanisms regulating imprinted genes in clusters*. Curr Opin Cell Biol, 2007. **19**(3): p. 281-9.

79. Ishida, M. and G.E. Moore, *The role of imprinted genes in humans*. Mol Aspects Med, 2013. **34**(4): p. 826-40.
80. Charalambous, M., et al., *An enhancer element at the Igf2/H19 locus drives gene expression in both imprinted and non-imprinted tissues*. Dev Biol, 2004. **271**(2): p. 488-97.
81. Cocchi, G., et al., *Silver-Russell syndrome due to paternal H19/IGF2 hypomethylation in a twin girl born after in vitro fertilization*. Am J Med Genet A, 2013. **161a**(10): p. 2652-5.
82. Wakeling, E.L., et al., *Diagnosis and management of Silver-Russell syndrome: first international consensus statement*. Nat Rev Endocrinol, 2017. **13**(2): p. 105-124.
83. Gronskov, K., et al., *Deletions and rearrangements of the H19/IGF2 enhancer region in patients with Silver-Russell syndrome and growth retardation*. J Med Genet, 2011. **48**(5): p. 308-11.
84. Howell, C.Y., et al., *Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene*. Cell, 2001. **104**(6): p. 829-38.
85. Borowczyk, E., et al., *Identification of a region of the DNMT1 methyltransferase that regulates the maintenance of genomic imprints*. Proc Natl Acad Sci U S A, 2009. **106**(49): p. 20806-11.
86. Ciccone, D.N., et al., *KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints*. Nature, 2009. **461**(7262): p. 415-8.
87. Zhang, T., et al., *G9a/GLP Complex Maintains Imprinted DNA Methylation in Embryonic Stem Cells*. Cell Rep, 2016. **15**(1): p. 77-85.
88. McGraw, S., et al., *Transient DNMT1 suppression reveals hidden heritable marks in the genome*. Nucleic Acids Res, 2015. **43**(3): p. 1485-97.
89. Cohen, D.I., et al., *Expression of an X-linked gene family (XLR) in late-stage B cells and its alteration by the xid mutation*. Nature, 1985. **314**(6009): p. 372-4.
90. Garchon, H.J., *The Xlr (X-linked lymphocyte regulated) gene family (a candidate locus for an X-linked primary immune deficiency)*. Immunodeficiency Rev, 1991. **2**(4): p. 283-302.
91. Allenet, B., D. Escalier, and H.J. Garchon, *A putative human equivalent of the murine Xlr (X-linked, lymphocyte-regulated) protein*. Mamm Genome, 1995. **6**(9): p. 640-4.
92. Escalier, D., L. Eloy, and H.J. Garchon, *Sex-specific gene expression during meiotic prophase I: Xlr (X linked, lymphocyte regulated), not its male homologue Xmr (Xlr related, meiosis regulated), is expressed in mouse oocytes*. Biol Reprod, 2002. **67**(5): p. 1646-52.
93. Davies, W., et al., *Xlr3b is a new imprinted candidate for X-linked parent-of-origin effects on cognitive function in mice*. Nat Genet, 2005. **37**(6): p. 625-9.
94. Raefski, A.S. and M.J. O'Neill, *Identification of a cluster of X-linked imprinted genes in mice*. Nat Genet, 2005. **37**(6): p. 620-4.
95. Qureshi, S.M., et al., *Identification of a novel imprinting mechanism at the X-linked imprinted locus, X-linked Lymphocyte Regulated 3/4 (Xlr3/4)*. Epigenetics Chromatin, 2013. **6** (Suppl 1): p. 68.
96. Foley, R.J., *Insight into the Function of X-linked Lymphocyte Regulated 3 (XLR3) and the Mechanism Regulating its Imprinted Expression*. 2015, University of Connecticut: Doctoral Dissertations. p. 134.

97. Izzo, A. and R. Schneider, *Chatting histone modifications in mammals*. Brief Funct Genomics, 2010. **9**(5-6): p. 429-43.
98. Balasubramanian, D., et al., *H3K4me3 inversely correlates with DNA methylation at a large class of non-CpG-island-containing start sites*. Genome Med, 2012. **4**(5): p. 47.
99. Mohn, F., et al., *Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors*. Mol Cell, 2008. **30**(6): p. 755-66.
100. Weber, M., et al., *Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome*. Nat Genet, 2007. **39**(4): p. 457-66.
101. Rountree, M.R., K.E. Bachman, and S.B. Baylin, *DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci*. Nat Genet, 2000. **25**(3): p. 269-77.
102. Zhu, H., G. Wang, and J. Qian, *Transcription factors as readers and effectors of DNA methylation*. Nat Rev Genet, 2016. **17**(9): p. 551-65.
103. Okitsu, C.Y. and C.L. Hsieh, *DNA methylation dictates histone H3K4 methylation*. Mol Cell Biol, 2007. **27**(7): p. 2746-57.
104. Jones, P.L., et al., *Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription*. Nat Genet, 1998. **19**(2): p. 187-91.
105. Mochizuki, K., et al., *Relationship between epigenetic regulation, dietary habits, and the developmental origins of health and disease theory*. Congenit Anom (Kyoto), 2017. **57**(6): p. 184-190.
106. Otani, J., et al., *Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain*. EMBO Rep, 2009. **10**(11): p. 1235-41.
107. Guo, X., et al., *Structural insight into autoinhibition and histone H3-induced activation of DNMT3A*. Nature, 2015. **517**(7536): p. 640-4.
108. Ooi, S.K., et al., *DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA*. Nature, 2007. **448**(7154): p. 714-7.
109. Eustermann, S., et al., *Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin*. Nat Struct Mol Biol, 2011. **18**(7): p. 777-82.
110. Du, J., et al., *DNA methylation pathways and their crosstalk with histone methylation*. Nat Rev Mol Cell Biol, 2015. **16**(9): p. 519-32.
111. Henikoff, S. and A. Shilatifard, *Histone modification: cause or cog?* Trends Genet, 2011. **27**(10): p. 389-96.
112. Lim, Y.C., et al., *A complex association between DNA methylation and gene expression in human placenta at first and third trimesters*. PLoS One, 2017. **12**(7): p. e0181155.
113. Marcho, C., W. Cui, and J. Mager, *Epigenetic dynamics during preimplantation development*. Reproduction, 2015. **150**(3): p. R109-20.
114. Mitalipov, S. and D. Wolf, *Totipotency, pluripotency and nuclear reprogramming*. Adv Biochem Eng Biotechnol, 2009. **114**: p. 185-99.
115. Wu, S.C. and Y. Zhang, *Active DNA demethylation: many roads lead to Rome*. Nat Rev Mol Cell Biol, 2010. **11**(9): p. 607-20.

116. Wang, H. and S.K. Dey, *Roadmap to embryo implantation: clues from mouse models*. Nat Rev Genet, 2006. **7**(3): p. 185-99.
117. Zernicka-Goetz, M., S.A. Morris, and A.W. Bruce, *Making a firm decision: multifaceted regulation of cell fate in the early mouse embryo*. Nat Rev Genet, 2009. **10**(7): p. 467-77.
118. Messerschmidt, D.M., B.B. Knowles, and D. Solter, *DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos*. Genes Dev, 2014. **28**(8): p. 812-28.
119. Seisenberger, S., et al., *The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells*. Mol Cell, 2012. **48**(6): p. 849-62.
120. Saitou, M., S. Kagiwada, and K. Kurimoto, *Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells*. Development, 2012. **139**(1): p. 15-31.
121. McGraw, S. and I.M. Trasler, *Oocyte epigenetics and the risks for imprinting disorders associated with assisted reproduction*, in *Biology and Pathology of the Oocyte: Role in Fertility, Medicine and Nuclear Reprograming*. 2013. p. 384.
122. Kim, J.M. and A. Ogura, *Changes in allele-specific association of histone modifications at the imprinting control regions during mouse preimplantation development*. Genesis, 2009. **47**(9): p. 611-6.
123. Legault, L.M., V. Bertrand-Lehouillier, and S. McGraw, *Pre-implantation alcohol exposure and developmental programming of FASD: an epigenetic perspective*. Biochem Cell Biol, 2018. **96**(2): p. 117-130.
124. Barak, A.J., H.C. Beckenhauer, and D.J. Tuma, *Betaine effects on hepatic methionine metabolism elicited by short-term ethanol feeding*. Alcohol, 1996. **13**(5): p. 483-6.
125. Hamid, A. and J. Kaur, *Long-term alcohol ingestion alters the folate-binding kinetics in intestinal brush border membrane in experimental alcoholism*. Alcohol, 2007. **41**(6): p. 441-6.
126. Castillo-Aguilera, O., et al., *DNA Methylation Targeting: The DNMT/HMT Crosstalk Challenge*. Biomolecules, 2017. **7**(1).
127. Knopik, V.S., et al., *The epigenetics of maternal cigarette smoking during pregnancy and effects on child development*. Dev Psychopathol, 2012. **24**(4): p. 1377-90.
128. Armstrong, D.A., et al., *Maternal smoking during pregnancy is associated with mitochondrial DNA methylation*. Environ Epigenet, 2016. **2**(3).
129. Emes, R.D., et al., *Antiepileptic drugs and the fetal epigenome*. Epilepsia, 2013. **54**(1): p. e16-9.
130. Geraghty, A.A., et al., *Nutrition During Pregnancy Impacts Offspring's Epigenetic Status-Evidence from Human and Animal Studies*. Nutr Metab Insights, 2015. **8**(Suppl 1): p. 41-7.
131. Canovas, S., et al., *DNA methylation and gene expression changes derived from assisted reproductive technologies can be decreased by reproductive fluids*. Elife, 2017. **6**.
132. Bloise, E., et al., *Impaired placental nutrient transport in mice generated by in vitro fertilization*. Endocrinology, 2012. **153**(7): p. 3457-67.

133. Delle Piane, L., et al., *Effect of the method of conception and embryo transfer procedure on mid-gestation placenta and fetal development in an IVF mouse model*. Hum Reprod, 2010. **25**(8): p. 2039-46.
134. Fernandez-Gonzalez, R., et al., *Long-term effect of in vitro culture of mouse embryos with serum on mRNA expression of imprinting genes, development, and behavior*. Proc Natl Acad Sci U S A, 2004. **101**(16): p. 5880-5.
135. Feuer, S. and P. Rinaudo, *Preimplantation stress and development*. Birth Defects Res C Embryo Today, 2012. **96**(4): p. 299-314.
136. Paoloni-Giacobino, A., *Implications of reproductive technologies for birth and developmental outcomes: imprinting defects and beyond*. Expert Rev Mol Med, 2006. **8**(12): p. 1-14.
137. Jamal, I., et al., *Rescue of altered HDAC activity recovers behavioural abnormalities in a mouse model of Angelman syndrome*. Neurobiol Dis, 2017. **105**: p. 99-108.
138. Higashimoto, K., et al., *Loss of CpG methylation is strongly correlated with loss of histone H3 lysine 9 methylation at DMR-LIT1 in patients with Beckwith-Wiedemann syndrome*. Am J Hum Genet, 2003. **73**(4): p. 948-56.
139. Gaudet, F., et al., *Induction of tumors in mice by genomic hypomethylation*. Science, 2003. **300**(5618): p. 489-92.
140. Jin, B., et al., *DNMT1 and DNMT3B modulate distinct polycomb-mediated histone modifications in colon cancer*. Cancer Res, 2009. **69**(18): p. 7412-21.
141. Li, D., et al., *Regulation of DNA methyltransferase 1 transcription in BRCA1-mutated breast cancer: a novel crosstalk between E2F1 motif hypermethylation and loss of histone H3 lysine 9 acetylation*. Mol Cancer, 2014. **13**: p. 26.
142. Wang, X. and B. Li, *DNMT1 regulates human endometrial carcinoma cell proliferation*. Onco Targets Ther, 2017. **10**: p. 1865-1873.
143. Sheaffer, K.L., E.N. Elliott, and K.H. Kaestner, *DNA Hypomethylation Contributes to Genomic Instability and Intestinal Cancer Initiation*. Cancer Prev Res (Phila), 2016. **9**(7): p. 534-46.
144. Walsh, C.P., J.R. Chaillet, and T.H. Bestor, *Transcription of IAP endogenous retroviruses is constrained by cytosine methylation*. Nat Genet, 1998. **20**(2): p. 116-7.
145. Wright, A. and P.J. Dyck, *Hereditary sensory neuropathy with sensorineural deafness and early-onset dementia*. Neurology, 1995. **45**(3 Pt 1): p. 560-2.
146. Melberg, A., et al., *Autosomal dominant cerebellar ataxia deafness and narcolepsy*. J Neurol Sci, 1995. **134**(1-2): p. 119-29.
147. Klein, C.J., et al., *Mutations in DNMT1 cause hereditary sensory neuropathy with dementia and hearing loss*. Nat Genet, 2011. **43**(6): p. 595-600.
148. Sun, Z., et al., *Aberrant signature methylome by DNMT1 hot spot mutation in hereditary sensory and autonomic neuropathy 1E*. Epigenetics, 2014. **9**(8): p. 1184-93.
149. Li, E., T.H. Bestor, and R. Jaenisch, *Targeted mutation of the DNA methyltransferase gene results in embryonic lethality*. Cell, 1992. **69**(6): p. 915-26.
150. Stresemann, C. and F. Lyko, *Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine*. Int J Cancer, 2008. **123**(1): p. 8-13.

151. Yoo, J., et al., *Molecular modeling of inhibitors of human DNA methyltransferase with a crystal structure: discovery of a novel DNMT1 inhibitor*. Adv Protein Chem Struct Biol, 2012. **87**: p. 219-47.
152. Lei, H., et al., *De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells*. Development, 1996. **122**(10): p. 3195-205.
153. Sado, T., et al., *X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation*. Dev Biol, 2000. **225**(2): p. 294-303.
154. Sakaue, M., et al., *DNA methylation is dispensable for the growth and survival of the extraembryonic lineages*. Curr Biol, 2010. **20**(16): p. 1452-7.
155. McGraw, S., et al., *Loss of DNMT1o disrupts imprinted X chromosome inactivation and accentuates placental defects in females*. PLoS Genet, 2013. **9**(11): p. e1003873.
156. Tsumura, A., et al., *Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b*. Genes Cells, 2006. **11**(7): p. 805-14.
157. Jackson, M., et al., *Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells*. Mol Cell Biol, 2004. **24**(20): p. 8862-71.
158. Martinez Arias, A., J. Nichols, and C. Schroter, *A molecular basis for developmental plasticity in early mammalian embryos*. Development, 2013. **140**(17): p. 3499-510.
159. Radziskeuskaya, A. and J.C. Silva, *Do all roads lead to Oct4? the emerging concepts of induced pluripotency*. Trends Cell Biol, 2014. **24**(5): p. 275-84.
160. King, A.D., et al., *Reversible Regulation of Promoter and Enhancer Histone Landscape by DNA Methylation in Mouse Embryonic Stem Cells*. Cell Rep, 2016. **17**(1): p. 289-302.
161. Reddington, J.P., et al., *Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes*. Genome Biol, 2013. **14**(3): p. R25.
162. Lehnertz, B., et al., *Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin*. Curr Biol, 2003. **13**(14): p. 1192-200.
163. Gossen, M. and H. Bujard, *Tight control of gene expression in mammalian cells by tetracycline-responsive promoters*. Proc Natl Acad Sci U S A, 1992. **89**(12): p. 5547-51.
164. Lin, S. and P. Talbot, *Methods for culturing mouse and human embryonic stem cells*. Methods Mol Biol, 2011. **690**: p. 31-56.
165. Bustin, S.A., *Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems*. J Mol Endocrinol, 2002. **29**(1): p. 23-39.
166. Rodrigue, S., et al., *Unlocking short read sequencing for metagenomics*. PLoS One, 2010. **5**(7): p. e11840.
167. Benes, V., J. Blake, and K. Doyle, *Ribo-Zero Gold Kit: improved RNA-seq results after removal of cytoplasmic and mitochondrial ribosomal RNA*. Nature Methods, 2011. **8**(11).

168. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data*. *Bioinformatics*, 2014. **30**(15): p. 2114-20.
169. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform*. *Bioinformatics*, 2009. **25**(14): p. 1754-60.
170. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS)*. *Genome Biol*, 2008. **9**(9): p. R137.
171. Li, H., et al., *The Sequence Alignment/Map format and SAMtools*. *Bioinformatics*, 2009. **25**(16): p. 2078-9.
172. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner*. *Bioinformatics*, 2013. **29**(1): p. 15-21.
173. Anders, S., P.T. Pyl, and W. Huber, *HTSeq--a Python framework to work with high-throughput sequencing data*. *Bioinformatics*, 2015. **31**(2): p. 166-9.
174. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. *Genome Biol*, 2014. **15**(12): p. 550.
175. Xi, Y. and W. Li, *BSMAP: whole genome bisulfite sequence MAPping program*. *BMC Bioinformatics*, 2009. **10**: p. 232.
176. Sloan, C.A., et al., *ENCODE data at the ENCODE portal*. *Nucleic Acids Res*, 2016. **44**(D1): p. D726-32.
177. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features*. *Bioinformatics*, 2010. **26**(6): p. 841-2.
178. Tamura, I., et al., *Genome-wide analysis of histone modifications in human endometrial stromal cells*. *Mol Endocrinol*, 2014. **28**(10): p. 1656-69.
179. Sharifi-Zarchi, A., et al., *DNA methylation regulates discrimination of enhancers from promoters through a H3K4me1-H3K4me3 seesaw mechanism*. *BMC Genomics*, 2017. **18**(1): p. 964.
180. Rye, M., et al., *Clustered ChIP-Seq-defined transcription factor binding sites and histone modifications map distinct classes of regulatory elements*. *BMC Biol*, 2011. **9**: p. 80.
181. Tripathi, S., et al., *Meta- and Orthogonal Integration of Influenza "OMICs" Data Defines a Role for UBR4 in Virus Budding*. *Cell Host Microbe*, 2015. **18**(6): p. 723-35.
182. Martinez, S.R., M.S. Gay, and L. Zhang, *Epigenetic mechanisms in heart development and disease*. *Drug Discov Today*, 2015. **20**(7): p. 799-811.
183. Strumpf, D., et al., *Cdx2 is required for correct cell fate specification and differentiation of trophoblast in the mouse blastocyst*. *Development*, 2005. **132**(9): p. 2093-102.
184. Simon, E., et al., *Vestigial-like 3 is a novel Ets1 interacting partner and regulates trigeminal nerve formation and cranial neural crest migration*. *Biol Open*, 2017. **6**(10): p. 1528-1540.
185. Symmank, J. and G. Zimmer, *Regulation of neuronal survival by DNA methyltransferases*. *Neural Regen Res*, 2017. **12**(11): p. 1768-1775.
186. Hutnick, L.K., et al., *DNA hypomethylation restricted to the murine forebrain induces cortical degeneration and impairs postnatal neuronal maturation*. *Hum Mol Genet*, 2009. **18**(15): p. 2875-88.
187. Loring, J.F., et al., *A gene expression profile of embryonic stem cells and embryonic stem cell-derived neurons*. *Restor Neurol Neurosci*, 2001. **18**(2-3): p. 81-8.



188. Bernstein, B.E., et al., *A bivalent chromatin structure marks key developmental genes in embryonic stem cells*. Cell, 2006. **125**(2): p. 315-26.
189. Attema, J.L., et al., *Epigenetic characterization of hematopoietic stem cell differentiation using miniChIP and bisulfite sequencing analysis*. Proc Natl Acad Sci U S A, 2007. **104**(30): p. 12371-6.
190. Xie, W., et al., *Epigenomic analysis of multilineage differentiation of human embryonic stem cells*. Cell, 2013. **153**(5): p. 1134-48.
191. D'Aiuto, L., et al., *Mouse ES cells overexpressing DNMT1 produce abnormal neurons with upregulated NMDA/NR1 subunit*. Differentiation, 2011. **82**(1): p. 9-17.
192. Schmidt, C.S., et al., *Global DNA hypomethylation prevents consolidation of differentiation programs and allows reversion to the embryonic stem cell state*. PLoS One, 2012. **7**(12): p. e52629.
193. Boland, M.J., K.L. Nazor, and J.F. Loring, *Epigenetic regulation of pluripotency and differentiation*. Circ Res, 2014. **115**(2): p. 311-24.
194. Kraushaar, D.C. and K. Zhao, *The epigenomics of embryonic stem cell differentiation*. Int J Biol Sci, 2013. **9**(10): p. 1134-44.
195. Sheaffer, K.L., et al., *DNA methylation is required for the control of stem cell differentiation in the small intestine*. Genes Dev, 2014. **28**(6): p. 652-64.
196. Vakilian, H., et al., *DDX3Y, a Male-Specific Region of Y Chromosome Gene, May Modulate Neuronal Differentiation*. J Proteome Res, 2015. **14**(9): p. 3474-83.
197. Matsubara, Y., et al., *TALEN-Mediated Gene Disruption on Y Chromosome Reveals Critical Role of EIF2S3Y in Mouse Spermatogenesis*. Stem Cells Dev, 2015. **24**(10): p. 1164-70.
198. Yagi, M., et al., *Derivation of ground-state female ES cells maintaining gamete-derived DNA methylation*. Nature, 2017. **548**(7666): p. 224-227.
199. Sato, S., et al., *Genome-wide DNA methylation profile of tissue-dependent and differentially methylated regions (T-DMRs) residing in mouse pluripotent stem cells*. Genes Cells, 2010. **15**(6): p. 607-18.
200. Isagawa, T., et al., *DNA methylation profiling of embryonic stem cell differentiation into the three germ layers*. PLoS One, 2011. **6**(10): p. e26052.
201. Li, X., et al., *A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints*. Dev Cell, 2008. **15**(4): p. 547-57.
202. Quenneville, S., et al., *In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions*. Mol Cell, 2011. **44**(3): p. 361-72.
203. Rugg-Gunn, P.J., A.C. Ferguson-Smith, and R.A. Pedersen, *Status of genomic imprinting in human embryonic stem cells as revealed by a large cohort of independently derived and maintained lines*. Hum Mol Genet, 2007. **16 Spec No. 2**: p. R243-51.
204. Auclair, G., et al., *Ontogeny of CpG island methylation and specificity of DNMT3 methyltransferases during embryonic development in the mouse*. Genome Biol, 2014. **15**(12): p. 545.
205. Habibi, E., et al., *Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells*. Cell Stem Cell, 2013. **13**(3): p. 360-9.

206. Harrington, C.T., et al., *Fundamentals of pyrosequencing*. Arch Pathol Lab Med, 2013. **137**(9): p. 1296-303.
207. Moullan, N., et al., *Tetracyclines Disturb Mitochondrial Function across Eukaryotic Models: A Call for Caution in Biomedical Research*. Cell Rep, 2015.
208. Chatzisprou, I.A., et al., *Tetracycline antibiotics impair mitochondrial function and its experimental use confounds research*. Cancer Res, 2015. **75**(21): p. 4446-9.
209. Ahler, E., et al., *Doxycycline alters metabolism and proliferation of human cell lines*. PLoS One, 2013. **8**(5): p. e64561.
210. Pandian, G.N. and H. Sugiyama, *Strategies to modulate heritable epigenetic defects in cellular machinery: lessons from nature*. Pharmaceuticals (Basel), 2012. **6**(1): p. 1-24.
211. Vojta, A., et al., *Repurposing the CRISPR-Cas9 system for targeted DNA methylation*. Nucleic Acids Res, 2016. **44**(12): p. 5615-28.
212. Faast, R., et al., *Histone variant H2A.Z is required for early mammalian development*. Curr Biol, 2001. **11**(15): p. 1183-7.
213. Coleman-Derr, D. and D. Zilberman, *Deposition of histone variant H2A.Z within gene bodies regulates responsive genes*. PLoS Genet, 2012. **8**(10): p. e1002988.