



HAL
open science

Impact des dérégulations de l'épigénome sur les maladies chez l'homme : études sur des modèles de la maladie d'Alzheimer et de cancers SWI/SNF-dépendants

Marina Nocente

► To cite this version:

Marina Nocente. Impact des dérégulations de l'épigénome sur les maladies chez l'homme : études sur des modèles de la maladie d'Alzheimer et de cancers SWI/SNF-dépendants. Génomique, Transcriptionnelle et Protéomique [q-bio.GN]. Université Paris-Saclay, 2021. Français. NNT : 2021UPASL053 . tel-04213025

HAL Id: tel-04213025

<https://theses.hal.science/tel-04213025>

Submitted on 21 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Impact des dérégulations de l'épigénome sur
les maladies chez l'homme - études sur des
modèles de la maladie d'Alzheimer et de
cancers SWI/SNF-dépendants

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 577
Structure et Dynamique des Systèmes Vivants (SDSV)
Spécialité de doctorat: Sciences de la Vie et de la Santé
Unité de recherche : Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology
of the Cell (I2BC), 91198, Gif-sur-Yvette, France.
Réfèrent : Faculté des sciences d'Orsay

Thèse présentée et soutenue à Paris-Saclay,
le 20/09/2021, par

Marina NOCENTE

Composition du Jury

Fabienne MALAGNAC Professeure, Université Paris Saclay	Présidente
Irwin DAVIDSON Directeur de recherche, IGBMC, Strasbourg	Rapporteur & Examineur
Saadi KHOCHBIN Directeur de recherche, IAB, Grenoble	Rapporteur & Examineur
Claire FRANCASTEL Directrice de recherche, Université Paris Diderot	Examinatrice
Amélie LEFORESTIER Directrice de recherche, LPS, Orsay	Examinatrice

Direction de la thèse

Matthieu GERARD Directeur de recherche, I2BC, CEA Saclay	Directeur de thèse
--	--------------------

Remerciements

Faire une thèse nécessite des qualités scientifiques, humaines et personnelles comme la passion pour son sujet, la détermination et la persévérance. Cette expérience professionnelle et personnelle marquante ne serait rien sans le soutien et les conseils des nombreuses personnes rencontrées durant ces quatre années.

Tout d'abord, un très grand merci à Matthieu Gérard mon directeur de thèse pour m'avoir choisi et fait confiance sur un sujet de recherche très ambitieux. Matthieu, merci pour ta patience, ta gentillesse, ta grande disponibilité, ton encadrement et tes précieux conseils tout au long de mon stage de M2 et de ma thèse. Je te remercie de m'avoir crue capable (peut-être même plus que moi au début) et de m'avoir incitée à me mettre à la bioinformatique dès le début de ma thèse. Le chemin a été long et parfois difficile mais j'en suis fière (même si le chemin continue). J'admire également ta sérénité, ta passion et tes nombreuses connaissances scientifiques et expérimentales. Le meilleur conseil que tu m'aies donné pendant ma thèse est probablement qu'il ne faut pas mettre tous ses œufs dans le même panier. Tu avais entièrement raison, nos projets l'ont démontré ! J'ai été fière d'être ta doctorante et je suis honorée de poursuivre la recherche à tes côtés pour quelques mois supplémentaires.

Je souhaite également remercier les membres de mon jury de thèse qui m'ont fait l'honneur de lire et d'évaluer mon travail : Irwin Davidson, Saadi Khochbin, Claire Francastel, Fabienne Malagnac et Amélie Leforestier.

Je remercie aussi Laurette Boutillier et Pierre-Antoine Defossez pour le temps qu'ils m'ont consacré lors de mes comités de thèse, pour leurs conseils bienveillants et pour nos discussions intéressantes qui ont aidé à faire progresser nos projets.

Les membres de mon équipe sont des personnes essentielles qui ont été présentes à mes côtés pour me soutenir et m'aider tout au long de ma thèse. Je vous en remercie chaleureusement :

Hélène, merci beaucoup pour ta bonne humeur quotidienne, pour tes encouragements et pour ta confiance. Partager un bureau avec toi pendant plus de quatre ans a été un véritable plaisir. Je te remercie de m'avoir transmis tous tes « trucs et astuces » et de m'avoir appris tout ce que tu sais sur la culture cellulaire, ta petite « Bubulle » a beaucoup progressé depuis le début de son M2 !

Émilie, je te remercie particulièrement de m'avoir guidée, aidée et soutenue (surtout lors du DU) dans cet univers obscur qu'était la bioinformatique à mes débuts. J'espère que tu es fière des progrès que ta « bébé bioinfo » a fait. Maintenant je comprends les choses étranges que tu racontes parfois. Merci également pour les nombreuses discussions partagées autour de la danse, du jardinage, de la cuisine...

Anida, thank you very much for your kindness, your help and our discussions during my PhD. With you, I improved a lot my English and I'm less afraid to speak in English. I would also like to thank you for teaching me the cloning and for answering to my numerous questions ("Anida, concerning my cloning, ..."). You are the best partner to make gradients, our arms still remember it! I wish you all the best for your professional and personal life with baby Emile.

Je souhaite également remercier nos collaboratrices, Sophie Chantalat et ses deux techniciennes Cécile et Florence, pour avoir préparé les *libraries* de nos très nombreux échantillons avec beaucoup de soin. Merci également à Cécile d'avoir partagé avec moi son savoir-faire concernant les expériences de CUT&RUN.

Je voudrais remercier tous les collègues du bâtiment pour leur sympathique présence et particulièrement quelques-uns : Gwen et Jean-Yves pour vos conseils sur les expériences de western-blot, Jean-Christophe pour ton aide sur R lors de mon DU, et surtout merci à toi Régis pour ta gentillesse et ta grande disponibilité, que ce soit pour des questions concernant les manip ou pour passer nos commandes.

Merci également aux animaliers : Anne-Sophie, Patrick, Jean-Charles et Sylvain d'avoir pris soin de mes souris tout au long du projet.

Je voudrais aussi remercier notre secrétaire Catherine qui est toujours prête à rendre service et à nous aider dans les nombreuses démarches administratives relatives au CEA.

Je souhaite également remercier mes différents professeurs : mes professeurs de fac, particulièrement ceux de Biologie moléculaire et d'Épigénétique, qui m'ont transmis leur passion et les différents professeurs m'ayant appris la Bioinformatique, que ce soit à la fac d'Orsay, à l'École de Bioinformatique de Roscoff ou à l'université Paris Diderot lors du Diplôme Universitaire de Bioinformatique intégrative. Merci également à Thibault Dayris et Marc Deloger pour leur accueil sur la plateforme de Bioinformatique de Gustave Roussy, pour leur aide et leurs bons conseils.

Enfin, je souhaite remercier ma famille qui a, en quelque sorte, participé à cette thèse : mes grands-parents, mes parents et Julien. C'est aussi grâce à vous si j'en suis là aujourd'hui. Merci pour votre amour et votre soutien qui m'ont accompagnée tout au long de ces années. Je voudrais particulièrement te remercier mamie pour ton dévouement, ton aide et ton soutien indéfectible depuis toujours.

Table des matières

REMERCIEMENTS	1
TABLE DES FIGURES ET TABLEAUX	
LISTE DES ABREVIATIONS	
PREAMBULE	
INTRODUCTION	1
CHAPITRE 1 : L'ORGANISATION DE LA CHROMATINE	1
I. <i>Les nucléosomes canoniques</i>	1
II. <i>Les structures non canoniques des nucléosomes et les particules subnucléosomales</i>	4
1) Détection de particules subnucléosomales et de nucléosomes fragiles.....	4
2) Description des particules nucléosomales et subnucléosomales existantes	7
a) Les octasomes.....	8
b) Les hexasomes	10
c) Le tétrasome	13
d) L'hémisome et le <i>split</i> nucleosome	14
CHAPITRE 2 : LES ELEMENTS CIS-REGULATEURS DE LA TRANSCRIPTION CHEZ LES MAMMIFERES ET LES DIFFERENTS FACTEURS QUI S'Y LIENT POUR REGULER L'EXPRESSION DES GENES.....	16
I. <i>La fonction des éléments promoteurs et des facteurs associés dans la régulation de la transcription mammifère</i>	16
1) Caractéristiques et structure des promoteurs	17
2) Différents facteurs participent à l'initiation de la transcription au niveau des promoteurs.....	20
a) L'ARN polymérase II	20
b) Les facteurs généraux de transcription sont des composants du complexe de pré-initiation de la transcription	21
c) Certains TFs se lient au niveau des promoteurs pour réguler la transcription	23
II. <i>La fonction des éléments enhancers dans la régulation de la transcription</i>	24
1) Caractéristiques générales des enhancers.....	24
2) Le profil épigénétique des enhancers et l'activité transcriptionnelle définissent différents types d'enhancer	24
3) Les TFs qui se lient aux enhancers en cellules ES	26
III. <i>Fonction des complexes modificateurs de la chromatine au niveau des éléments cis-régulateurs</i>	27
1) Les complexes Polycomb répriment la transcription	27
2) Les autres acteurs de la transcription	30
a) Le médiateur.....	30
b) Le complexe CBP/p300	31
c) Les complexes ATAC et SAGA	32

IV. Le fonctionnement des éléments cis-régulateurs de la transcription dans le contexte de la chromatine chez les mammifères	33
1) Définition et caractéristiques des facteurs pionniers	33
2) La liaison des facteurs pionniers OCT4 et SOX2 au nucléosome.....	36
CHAPITRE 3 : LES REMODELEURS DE LA CHROMATINE SONT IMPLIQUES DANS L'ORGANISATION DE LA CHROMATINE ET DANS LA REGULATION DE LA TRANSCRIPTION	40
I. Caractéristiques et fonctions des remodeleurs de la chromatine	40
1) Les remodeleurs de chromatine ATP-dépendants.....	40
2) Différents mécanismes recrutent les remodeleurs sur les nucléosomes	41
3) Les principaux mécanismes d'action des remodeleurs.....	42
II. Les remodeleurs de la chromatine se lient aux nucléosomes au niveau de différentes régions.....	44
1) Les remodeleurs interagissent au niveau du patch acide et de la queue H4 du nucléosome	44
2) Les remodeleurs se lient au niveau du SHL 2 ou SHL 6 sur l'ADN nucléosomal.....	46
III. Mécanisme de translocation de l'ADN.....	46
IV. La fonction des remodeleurs à l'échelle du génome	49
1) Présentation des remodeleurs étudiés pendant ma thèse.....	49
2) L'organisation nucléosome au niveau des promoteurs mammifères.....	49
3) Distribution génomique et fonction des remodeleurs dans la régulation de la transcription.....	50
V. Les remodeleurs interagissent avec les facteurs pionniers pour l'ouverture de la chromatine.....	53
VI. Le complexe de remodelage BAF impliqué dans les cancers humains.	57
1) Évolution et composition des différents complexes BAF.....	57
a) Évolution du complexe BAF	57
b) Les différents complexes BAF chez les mammifères	57
2) Détails de l'interaction entre le complexe BAF humain et son nucléosome cible	58
3) Le complexe BAF est nécessaire au maintien des caractéristiques fonctionnelles et de la transcription dans les cellules ES.....	60
4) Implication du complexe BAF dans les cancers humains.....	62
CHAPITRE 4 : IMPLICATION D'UN ACTEUR EPIGENETIQUE DANS LA MALADIE D'ALZHEIMER	66
I. Préambule sur la maladie d'Alzheimer.....	66
II. Rôle de l'histone désacétylase 2 (HDAC2) dans la maladie d'Alzheimer	68
MATERIEL ET METHODES.....	71
PARTIE A : EXPERIENCES	71
I. Projet portant sur l'organisation nucléosomale des enhanceurs et la fonction des remodeleurs de la chromatine.....	71
1) Culture de cellules ES de souris.....	71
2) Construction d'un plasmide et génération de la lignée cellulaire Brg1-mAID	72
3) Construction de vecteurs shRNA ciblant BRG1.....	73
4) Déplétion de remodeleurs d'intérêt par transfection d'un shRNA ou par incubation avec de l'auxine puis préparation de la chromatine.....	74
5) Vérification de la déplétion des remodeleurs et test des anticorps par western blot.....	75

6)	Vérification de la qualité de la digestion MNase de la chromatine sur gel d'agarose	77
7)	Centrifugation sur gradient de sucrose.....	79
8)	Expériences de CHIP-seq	80
9)	Expériences de CHIP-seq séquentiel : CHIP-tandem.....	83
<i>II. Projet portant sur l'implication de HDAC2 dans la régulation de la transcription de modèles murins de la maladie d'Alzheimer</i>		<i>87</i>
1)	Caractéristiques des lignées de souris utilisées	87
2)	Génotypage des souris par qPCR	88
3)	Dissection des hippocampes de souris et préparation en vue des expériences de western blot et de RNA-seq	89
4)	Expériences de western blot à partir de protéines extraites d'hippocampe	90
5)	Quantification des protéines HDAC2 après western-blot	91
6)	Expériences de RNA-seq	92
7)	Expérience de CHIP-seq séquentiel : tandem-CHIP	93
PARTIE B : ANALYSES BIOINFORMATIQUES		94
<i>I. Indications sur la reproductibilité bioinformatique de nos données</i>		<i>94</i>
<i>II. Analyses bioinformatiques des données de CHIP-seq.....</i>		<i>95</i>
1)	Contrôle qualité des données et nettoyage.....	95
2)	L'alignement des <i>reads</i> sur le génome murin	96
3)	Les différents processus post-mapping.....	97
4)	La visualisation des résultats.....	99
<i>III. Analyses bioinformatiques des données de RNA-seq.....</i>		<i>100</i>
1)	Contrôle qualité des données	100
2)	Alignement sur le génome et élimination des ARN ribosomiaux	100
3)	Vérification du génotype de nos souris.....	101
4)	Pseudo-mapping ou <i>mapping</i> et comptage de l'expression des gènes	102
RESULTATS		104
PARTIE A : COMMENT LES REMODELEURS DE LA CHROMATINE REGULENT L'ORGANISATION NUCLEOSOMALE AU NIVEAU DES ENHANCERS EN CELLULES ES DE SOURIS.....		104
<i>I. Article Nocente et al. : "The Brg1 ATPase generates subnucleosomal particles targeted by pioneer transcription factors at enhancers"</i>		<i>104</i>
<i>II. Résumé des résultats de l'article :</i>		<i>143</i>
PARTIE B : ÉTUDE DE LA FONCTION DE HDAC2 DANS LA MALADIE D'ALZHEIMER		147
<i>I. Les différents modèles de souris utilisés au cours de la thèse.....</i>		<i>147</i>
<i>II. HDAC2 est modérément surexprimée dans un modèle de souris de la maladie d'Alzheimer</i>		<i>150</i>
1)	Modèle 3xTg-AD.....	150
2)	Modèle NLGF.....	156
<i>III. Caractérisation du profil de liaison de HDAC2 sur le génome du cerveau adulte</i>		<i>158</i>
<i>IV. Analyse par RNA-seq du contrôle transcriptionnel exercé par HDAC2 dans le contexte de la maladie d'Alzheimer.....</i>		<i>160</i>

1) Impact du modèle de souris 3xTg-AD de la maladie d'Alzheimer sur le transcriptome murin.....	160
2) Influence de la perte de fonction de HDAC2 sur le transcriptome des souris dans le contexte sain et dans des modèles de la maladie d'Alzheimer	162
3) Influence du modèle NLGF de la maladie d'Alzheimer sur le transcriptome des souris.....	171
DISCUSSION.....	176
PARTIE A : CARACTERISATION DE L'ARCHITECTURE DES PARTICULES NUCLEOSOMALES AU NIVEAU DES ENHANCERS ET IMPACT DU REMODELEUR BRG1 SUR LA FORMATION DE PARTICULES SUBNUCLEOSOMALES	
	176
PARTIE B : HDAC2 CONTRIBUE TRES MODESTEMENT A LA MALADIE D'ALZHEIMER.....	
	180
PERSPECTIVES.....	184
I. APPROFONDIR LA CARACTERISATION DU ROLE DE BRG1 DANS LA FORMATION DES PARTICULES SUBNUCLEOSOMALES AUX ENHANCERS	
	184
II. VISUALISATION DES PARTICULES SUBNUCLEOSOMALES PAR MICROSCOPIE ELECTRONIQUE	
	186
III. EXPLORATION DE LA COMPOSITION EN HISTONES DES PARTICULES SUBNUCLEOSOMALES.....	
	186
CONCLUSION	188
BIBLIOGRAPHIE.....	189

Table des Figures et Tableaux

Figure 1 : Structure des nucléosomes canoniques.....	2
Figure 2 : Mise en évidence de particules nucléosomales non canoniques.....	5
Tableau 1: Les noms traditionnels et suggérés par Zlatanova et al des particules nucléosomales et subnucléosomales existantes.....	9
Figure 3 : Modélisation de la dynamique des nucléosomes canoniques et des particules subnucléosomales existantes.....	11
Figure 4 : Les différents types de <i>core</i> promoteurs actifs chez les mammifères	18
Figure 5 : Implication des GTFs dans la formation du PIC et dans l'initiation de la transcription	23
Tableau 2 : Le niveau de transcription et l'enrichissement en marque d'histone définissent trois catégories d'enhancers typiques : les enhancers actifs, <i>primed</i> et <i>poised</i>	26
Figure 6 : Différents facteurs influencent le recrutement de PRC2 et module son activité catalytique.....	29
Figure 7 : La méthode NCAP-SELEXA permet d'identifier le mode de liaison des TFs pionniers au nucléosome	34
Figure 8 : Les trois fonctions principales des remodeleurs de la chromatine ATP-dépendants	43
Figure 9 : Mécanisme de translocation de l'ADN le long d'un nucléosome par un remodeleur de la sous-famille SWI/SNF pouvant conduire à l'éjection d'un nucléosome.....	48
Figure 10 : Les remodeleurs de la chromatine régulent l'expression de différentes catégories de gènes dans les cellules ES de souris.	53
Figure 11 : Modèle dynamique de coopération entre le remodeleur RSC et les facteurs pionniers GRF pour le remodelage des nucléosomes fragiles chez la levure.....	56
Figure 12 : Modèle de la structure du complexe BAF humain endogène contenant de nombreuses sous-unités mutées dans les cancers.	61
Figure 13 : La progression des plaques séniles extracellulaires et des enchevêtrements neurofibrillaires au cours de la maladie d'Alzheimer.....	67
Tableau 3 : Les différentes lignées de cellules ES étiquetées utilisées au cours de ma thèse	71
Figure 14 : Vérification de l'efficacité de la déplétion de BRG1 par western blot	77

Figure 15 : Profils de digestion d'échantillons de chromatine digérés par la MNase analysés sur gel d'agarose	78
Figure 16 : Analyse des différentes fractions du gradient de sucrose issues de chromatine contrôle ou déplétée en BRG1	80
Tableau 4 : Anticorps utilisés dans les expériences de MNase-ChIP-seq et de ChIP-tandem. 86	
Tableau 5 : Caractéristiques des modèles de souris 3xTg-AD et App ^{tm3 (NL-G-F)}	87
Figure 17 : Vérification de la présence de la mutation du gène <i>Psen1</i> dans les souris modèles 3xTg-AD et 3xTg-AD ; <i>Hdac2</i> ^{-/-}	102
Tableau 6 : Les différentes lignées de souris utilisées dans ce projet.....	150
Figure 18 : Suivi de l'expression de la protéine HDAC2 au cours du temps chez les souris contrôles et les souris 3xTg-AD modèles de la maladie d'Alzheimer	152
Figure 19 : Suivi de l'expression de la protéine HDAC2 au cours du temps chez les souris mâles et femelles contrôles et 3xTg-AD.	155
Figure 20 : Suivi de l'expression de la protéine HDAC2 au cours du temps chez les souris contrôles et les souris NLGF modèles de la maladie d'Alzheimer	157
Figure 21 : Liaison de la protéine HDAC2 au niveau du promoteurs de ces gènes cibles dans le cerveau de souris	159
Figure 22 : Analyse en composante principale et calcul de corrélation entre les souris 3xTg-AD et contrôles.	161
Figure 23 : Vérification de la délétion de la partie C-terminale des transcrits <i>Hdac2</i> dans les souris <i>Hdac2</i> ^{-/-} et 3xTg-AD ; <i>Hdac2</i> ^{-/-}	163
Figure 24 : Analyse en composante principale et calcul de corrélation entre les souris 3xTg-AD ; <i>Hdac2</i> ^{-/-} et 3xTg-AD.	164
Figure 25 : Analyse en composante principale et calcul de corrélation entre les souris <i>Hdac2</i> ^{-/-} et <i>Hdac2</i> ^{+/+}	166
Figure 26 : Analyse de <i>Gene Ontology</i> révélant les différentes voies biologiques dans lesquelles les gènes dérégulés sont exprimés en absence de HDAC2.	167
Tableau 7 : Nombre de gènes différentiellement exprimés issus des analyses de RNA-seq pour des comparaisons d'intérêt chez les souris modèles 3xTg-AD et NLGF.	168
Figure 27 : Nombre de gènes communs entre différentes listes de gènes dérégulés	170
Figure 28 : Analyse en composante principale des échantillons NLGF et contrôles.	171

Figure 29 : Analyse en composante principale des échantillons NLGF et contrôles selon différentes comparaisons.....	172
Figure 30 : Analyse de <i>Gene Ontology</i> révélant les différentes voies biologiques dans lesquelles les gènes dérégulés sont exprimés.	175

Liste des abréviations

ACP (ou *PCA* en anglais) : Analyse en Composante Principale

ARNm : ARN messenger

ARNr : ARN ribosomal

ATAC : *Ada-Two-A-Containing*

BRG1 : *Brahma-related gene 1*

CBP : *CREB binding protein*

CHD1 : *Chromodomain Helicase DNA Binding Protein 1*

CHD4 : *Chromodomain Helicase DNA Binding Protein 4*

ChIP : *Chromatin ImmunoPrecipitation* ; immunoprécipitation de la chromatine

CHD8 : *Chromodomain Helicase DNA Binding Protein 8*

CTD : *C-Terminal domain* ; domaine C-terminal (fait référence à la Pol II)

CTRL : contrôle

DMEM : *Dulbecco's Modified Eagle Medium*

DNase I : désoxyribonucléase I

EP400 : *E1A-binding protein p400*

ES : *Embryonic stem* ; souches embryonnaires, se référant à cellules souches embryonnaires

GRF : *General Regulation Factor*

GTF : *General Transcription Factor* ; facteurs généraux de transcription

HD : homéodomaine

HDAC2 : histone desacetylase 2

HF : HA-Flag

HMG : *Sry-related High Mobility Group*

H2AK119ub : mono-ubiquitination de la lysine 119 de l'histone H2A

H3K4me1 : monométhylation de la lysine 4 de l'histone H3

H3K4me2 : diméthylation de la lysine 4 de l'histone H3

H3K4me3 : triméthylation de la lysine 4 de l'histone H3

H3K27ac : acétylation de la lysine 27 de l'histone H3

H3K27me3 : triméthylation de la lysine 27 de l'histone H3

Inr : élément *Initiator*

IP : ImmunoPrécipitation ; ImmunoPrécipiter

Kb : kilobase

KI : *Knock-In*

mAID : mini Auxin-Inducible Degron

MNase : endonucléase micrococcale

MSC : *Multiple Cloning Site* ; site de clonage multiple

NFR : *Nucleosome Free Region* ; région dépourvue en nucléosome

ON : OverNight ; toute la nuit

PIC : *PreInitiation Complex* ; complexe de pré-initiation de la transcription

pb : paire de base

POU : *Pit-Oct-Unc*

Pol II : ARN polymérase II

RT : Room Temperature ; température ambiante

SAGA : *Spt-Ada-Gcn5-acetyltransferase*

SHL : *SuperHelix Location* ; emplacement de super hélice

TAF : *TBP Associated Factor*

TBP : *TATA binding protein*

TFs : *Transcription Factors* ; facteurs de transcription

TFII : *Transcription Factor for polymerase II*

TSS : *Transcription Start Site* ; site d'initiation de la transcription

2D : deux dimensions

3D : tridimensionnel

3FTH : (Flag)₃-TEV-HA

Préambule

Mon sujet de thèse a consisté à explorer les mécanismes moléculaires qui contrôlent l'organisation de la chromatine et la transcription chez les mammifères. Je me suis intéressée à ces mécanismes en lien avec deux pathologies très différentes liées à mon sujet de thèse. D'une part, les cancers qui sont favorisés par les mutations dans les sous-unités du complexe SWI-SNF. Mon but a été de définir dans le détail la fonction de la sous-unité ATPase de ce complexe, le remodelleur BRG1, sur l'ouverture de la chromatine.

D'autre part, j'ai étudié la contribution de l'histone desacetylase 2 (HDAC2) dans la régulation transcriptionnelle et son rôle possible dans la maladie d'Alzheimer.

Mon manuscrit de thèse comporte une introduction axée sur l'organisation de la chromatine et les mécanismes moléculaires qui la contrôlent ainsi que la transcription. Je décris ensuite mes résultats sur la fonction des remodelleurs de chromatine sous la forme de notre article. Enfin, dans la dernière partie, je décris mes résultats sur la fonction de HDAC2 dans la maladie d'Alzheimer.

Introduction

Chapitre 1 : L'organisation de la chromatine

Au sein des cellules, l'ADN est naturellement associé avec des protéines pour former une structure, appelée chromatine. La chromatine est elle-même composée de nucléosomes. Les nucléosomes, unités structurantes de la chromatine, s'enchaînent les uns après les autres et forment une structure en collier de perles. Les génomes humain et murin sont composés d'environ 3 milliards de paires de base (pb) d'ADN, ce qui correspond à environ 15 millions de nucléosomes.

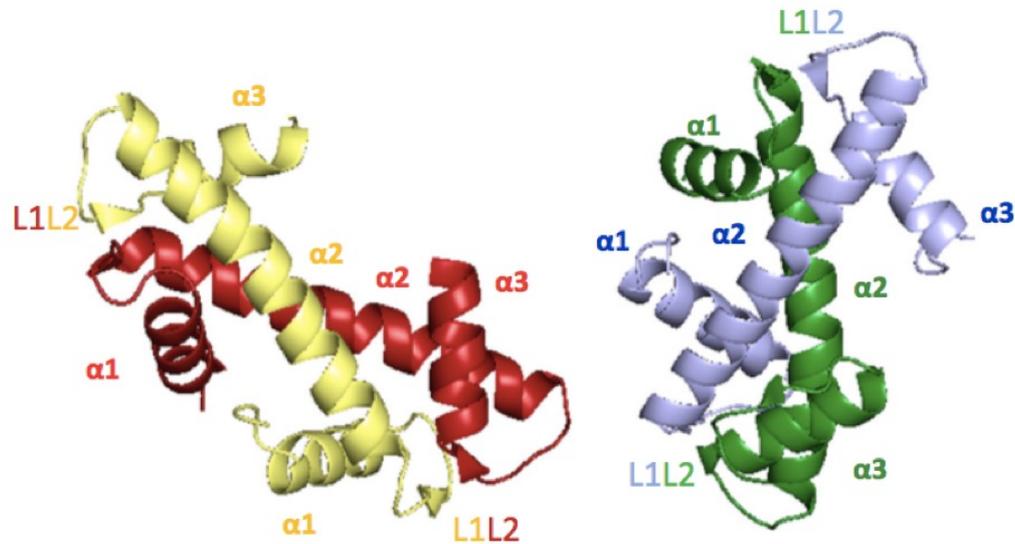
I. Les nucléosomes canoniques

La chromatine des génomes eucaryotes est constituée de nucléosomes espacés d'environ 200 pb (+/- 40 pb) (McGhee et Felsenfeld 1980). La structure du nucléosome a été résolue en 1997 à une résolution de 2,8 Å par cristallographie au rayons X (K. Luger et al. 1997). Les nucléosomes sont constitués de 147 pb d'ADN enroulé autour d'un octamère d'histone. L'ADN forme 1,65 tour dans une conformation de super hélice gauche autour des histones. L'octamère d'histone est composé de deux histones H2A, deux histones H2B, deux histones H3 et deux histones H4 (K. Luger et al. 1997). Chaque histone est organisée en trois hélices alpha (α_1 , α_2 et α_3) séparées par deux courtes boucles (L1 et L2) (**Figure 1a**) (Markert et Luger 2021). Cet enchainement d'hélices et de boucles est appelé un domaines *histone-fold* (**Figure 1a**). Ces domaines *histones-fold* (α_1 -L1- α_2 -L2- α_3) favorisent la formation des hétérodimères H3-H4 et des hétérodimères H2A-H2B (**Figure 1a**) (Markert et Luger 2021).

C'est au niveau de ces domaines *histones-fold* qu'ont lieu les contacts entre les histones et environ 120 pb d'ADN (K. Luger et al. 1997). Les 27 pb d'ADN restantes environ sont répartis aux deux extrémités et liés par l'hélice alpha N-terminale unique à H3 (K. Luger et al. 1997).

Au cœur de l'octamère d'histone, les deux dimères H3-H4 forment un hétérotétramère symétrique grâce à une structure à quatre hélices entre les deux histones H3 (**Figure 1b**). Les deux dimères H2A-H2B peuvent se lier au tétramère H3-H4 grâce à des interactions H2B-H4 (lu dans (Markert et Luger 2021)).

a.



b.

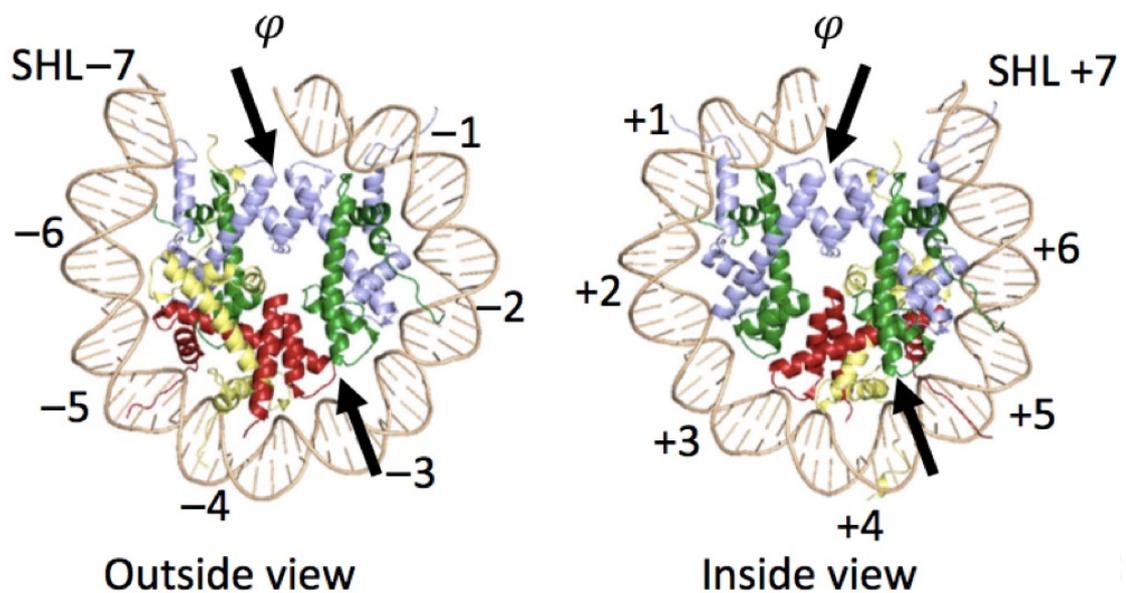


Figure 1 : Structure des nucléosomes canoniques

Un nucléosome canonique est constitué de 147 pb d'ADN enroulés autour d'un octamère d'histone comprenant deux histones H2A, deux histones H2B, deux histones H3 et deux histones H4. a. Les quatre histones possèdent des motifs *histone-fold* ($\alpha 1$ -L1- $\alpha 2$ -L2- $\alpha 3$) qui leur permettent de s'associer en hétérodimères H2A-H2B et en hétérodimères H3-H4. Les histones sont représentées par différentes couleurs : jaune pour H2A, rouge pour H2B, bleu clair pour H3 et vert pour H4. Sur ce genre de représentation, les hélices α du motif histone-fold sont représentées par des rubans hélicoïdales et les boucles L par de simples lignes.

b. Représentation de la structure d'un nucléosome vue de l'intérieur (*inside*) et de l'extérieur (*outside*) du nucléosome. Le centre du nucléosome, appelé la dyade, est désigné par la lettre ϕ et définit l'emplacement SHL 0. Les flèches désignent les contacts entre histones via quatre hélices. D'après (Markert et Luger 2021).

Au niveau du nucléosome, lorsque le sillon principal de l'ADN est en face de l'octamère d'histone, on désigne ces emplacements de sillon comme des « emplacements de super-hélice » ou SHL (K. Luger et al. 1997) (**Figure 1b**). Ces SHL sont numérotés de 0 à +/- 7 (+ pour une direction et – pour l'autre direction) et on dénombre 14 SHL (**Figure 1b**). Le SHL 0 est au niveau de la dyade, c'est-à-dire le centre du nucléosome (centrée sur le tétramère H3-H4), et le SHL +/- 7 est la dernière portion d'ADN lié à l'histone (K. Luger et al. 1997) (**Figure 1b**). L'ADN qui n'est pas en contact avec les histones est appelé ADN linker.

De par leurs nombreuses interactions avec l'ADN et leur structure, on pourrait penser que les nucléosomes sont très stables partout sur le génome. Cependant, *a minima* au niveau des éléments régulateurs, la structure des nucléosomes est dynamique afin de permettre la liaison des facteurs de transcriptions (TFs) et autres facteurs. Cette dynamique des nucléosomes se traduit par une composition en histone qui peut varier en nombre et en contenu (incorporation de variants d'histones), par une conformation différente des nucléosomes ou par différentes modifications post-traductionnelles des résidus au niveau des histones. De petites variations de séquences et d'interactions au niveau des 14 SHL vont être aussi importantes pour que des facteurs, comme les remodeleurs de la chromatine, puissent exercer leurs fonctions (Markert et Luger 2021).

La structure établie par le Dr Luger définit le nucléosome dit canonique. Ce nucléosome canonique se distingue d'une population de particules hétérogènes qui peuvent différer à différents niveaux de la structure canonique.

II. Les structures non canoniques des nucléosomes et les particules subnucléosomales

1) Détection de particules subnucléosomales et de nucléosomes fragiles

En 2011, le laboratoire du Dr Henikoff a couplé des expériences de digestion à la MNase avec un séquençage des reads *paired-end* (Henikoff et al. 2011). La MNase, pour endonucléase micrococcale, permet de cartographier de façon précise la position des nucléosomes sur le génome mammifère. La MNase va préférentiellement digérer les régions d'ADN accessibles au sein de la chromatine, notamment les régions qui ne sont pas protégées par des TFs, des nucléosomes ou d'autres acteurs (Noll 1974). La MNase va ainsi digérer préférentiellement l'ADN linker, situé entre les nucléosomes, et le séquençage des fragments d'ADN issus d'une telle digestion MNase (MNase-seq) va permettre de définir la distribution des nucléosomes au niveau du génome (Henikoff et al. 2011). Grâce à cette technique, Henikoff *et al.* ont pu mettre en évidence chez la levure l'existence des différentes particules nucléosomales à une résolution proche de la pb (**Figure 2**) (Henikoff et al. 2011). En effet, quelle que soit la durée de la digestion MNase, ils ont détecté des fragments d'ADN d'une taille comprise entre 150 et 165 pb, correspondant à des nucléosomes canoniques, et des plus petits fragments d'ADN, correspondant à des particules subnucléosomales ou à des protections de TFs (**Figure 2A**) (Henikoff et al. 2011).

Les particules subnucléosomales, avec des tailles de fragments d'ADN plus petits que 150 pb, sont majoritairement retrouvées dans les régions intergéniques et parfois au niveau de régions promotrices déplétées en nucléosome canonique (**Figures 2B et 2C**) (Henikoff et al. 2011).

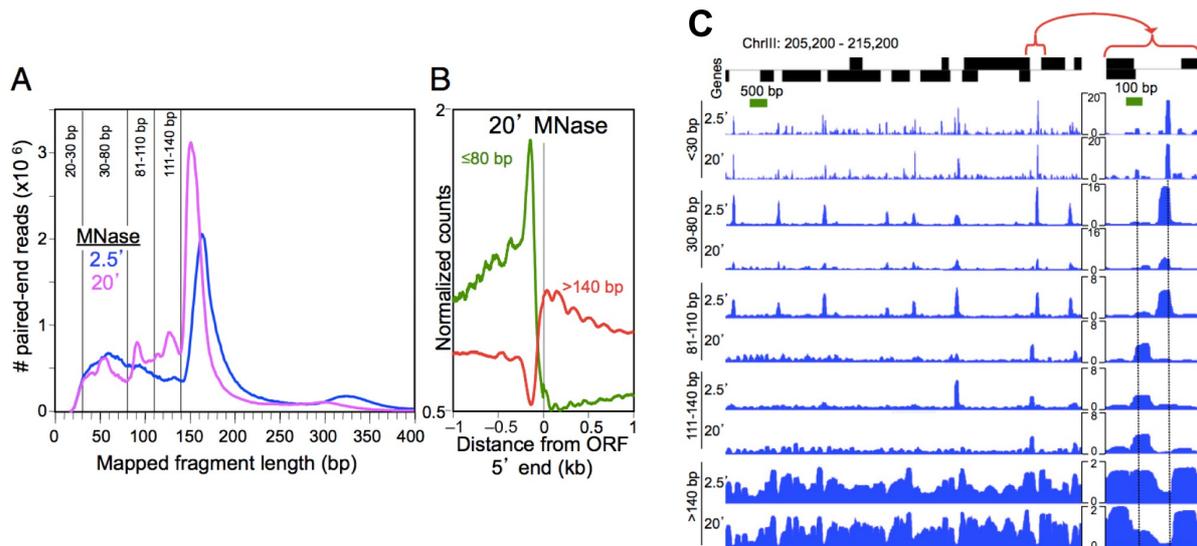


Figure 2 : Mise en évidence de particules nucléosomales non canoniques.

A. Henikoff *et al.* ont testé deux temps de digestion par la MNase : une digestion courte des noyaux de levure pendant 2,5 min (courbe bleue) et une digestion longue pendant 20 min (courbe rose). Cette figure représente le nombre de reads *paired-end* alignés sur différentes tailles de fragments d'ADN. Ils distinguent cinq catégories de taille de fragments : 20-30 pb, 30-80 pb, 81-110 pb, 111-140 pb et supérieur à 140 pb. Des fragments d'ADN de 30 à 80 pb, ainsi que de 150 et 165 pb ont été retrouvés dans les deux conditions de digestion MNase (Henikoff et al. 2011). Lors de la digestion MNase longue, ils ont également retrouvé des fragments d'ADN d'environ 90 pb et d'environ 130 pb (Henikoff et al. 2011). **B.** Cartographie du nombre de reads localisés sur des fragments d'ADN de taille inférieure ou égale à 80 pb (courbe verte) ou sur des fragments d'ADN de taille supérieure à 140 pb (courbe rouge) au niveau de l'extrémité 5' du cadre de lecture (ORF) (Henikoff et al. 2011). **C.** Les reads *paired-end* s'alignant sur les cinq catégories de taille de fragments d'ADN précédemment définis sont représentés au niveau d'un locus génomique du chromosome III de levure (Henikoff et al. 2011). On distingue les deux conditions de digestion (digestion courte en haut (2.5') et digestion longue en bas (20') de chaque catégorie de taille de fragments). La droite de la figure correspond à un agrandissement d'une région intergénique. Au niveau de cette région intergénique, les deux nucléosomes canoniques sont positionnés de part et d'autre des lignes pointillées et sont visibles dans la catégorie de taille de fragments supérieure à 140 pb. On distingue également une particule de 20 pb et une particule de 30 pb correspondant aux deux lignes pointillées. Entre les deux particules de 20 et 30 pb, une particule est visible sur la droite dans les catégories de taille de fragments de 30-80 pb et 81-110 pb. Une autre particule, vraisemblablement un nucléosome si on se base sur la taille, est visible sur la gauche dans les grandes catégories de taille de fragments dont dans la catégorie supérieure à 140 pb (Henikoff et al. 2011). D'après (Henikoff et al. 2011).

Henikoff *et al.* ont identifié trois types de particules de petite taille présentes dans la région intergénique : des particules d'environ 25 pb résistantes à la MNase correspondant aux régions d'ADN protégées par la liaison d'un ou plusieurs TFs (particules définies par les deux lignes pointillées visibles dans la catégorie de taille inférieure à 30 pb (Figure 2C)), une particule de plus grande taille sensible à la digestion MNase (particule visible au niveau de la

ligne pointillée à droite dans les catégories de 30 à 110 pb (**Figure 2C**)) et une particule détectée dans les grandes catégories de taille de fragments (visible au niveau de la ligne pointillée à gauche) correspondant à un nucléosome partiellement déroulé que la MNase a partiellement digéré (**Figure 2C**) (Henikoff et al. 2011). Ils concluent donc que les sites de liaison à l'ADN des TFs sont plus efficacement protégés de la digestion MNase que les particules subnucléosomales de tailles intermédiaires et que les particules nucléosomales partiellement déroulées (Henikoff et al. 2011).

Les auteurs ont également mis au point une représentation, appelée V-plot, permettant la visualisation de nucléosomes, de particules subnucléosomales et de TFs (Henikoff et al. 2011). La cartographie des différents fragments d'ADN au niveau de la région UAS du gène Gal4 chez la levure a permis de mettre en évidence des particules nucléosomales canoniques (fragments d'ADN supérieurs à 150 pb) de part et d'autre de l'UAS Gal4 et une particule subnucléosomale contenant environ 110 pb d'ADN au centre de cet UAS (Henikoff et al. 2011).

La limite de cette étude menée par le laboratoire du Dr Henikoff est qu'elle basée sur la taille des fragments d'ADN ; aucune expérience de ChIP-seq utilisant des anticorps anti-histone n'a été réalisée pour caractériser la nature des protéines associées à ces particules. Des ChIP-seq anti-histones auraient permis de déterminer avec précision la nature des particules qu'ils observent en se basant sur le contenu en histones et pas uniquement sur la taille de ces particules nucléosomales et subnucléosomales.

D'autres études basées sur des essais de digestion à la MNase (Xi et al. 2011; Henikoff et al. 2011; Kubik et al. 2015) ont mis en avant l'existence de particules subnucléosomales, qualifiées de « nucléosomes fragiles », qui sont plus abondantes en cas de digestion incomplète (courte digestion) qu'en cas de digestion complète (longue digestion) de la chromatine par la MNase (Xi et al. 2011). Les nucléosomes fragiles sont préférentiellement situés au niveau des promoteurs des gènes, au niveau des régions dépourvues en nucléosomes (NFR) (Xi et al. 2011; Henikoff et al. 2011). Ils représenteraient environ 5% des nucléosomes totaux chez les cellules de levure (Xi et al. 2011). Kubik *et al.* ont mis en évidence la présence de différentes catégories de particules sensibles à la MNase entre les nucléosomes

canoniques -1 et +1 autour du TSS (Kubik et al. 2015). Ils démontrent que les nucléosomes fragiles sont des particules nucléosomales atypiques qui incorporent environ 100 pb d'ADN contrairement au nucléosome canonique -1 stable qui en incorpore environ 140 (Kubik et al. 2015). Concernant la nature et la composition des régions NFR, Kubik *et al.* précisent que les régions véritablement dépourvues de nucléosomes doivent être inférieures à 150 pb, sinon un nucléosome canonique ou fragile peut s'y former (Kubik et al. 2015).

La fragilité de ces nucléosomes peut s'expliquer en partie par leur composition en variants d'histone, par l'intervention de différents acteurs, comme les remodeleurs de la chromatine qui favorisent la formation d'intermédiaires de remodelage au niveau desquels l'ADN est plus accessible à la MNase, mais aussi par la présence de séquences d'ADN enrichies en motifs polyA et G/C-rich (Xi et al. 2011; Kubik et al. 2015; Kubik, Bruzzone, et Shore 2017).

Kubik, Bruzzone, et Shore proposent que les nucléosomes fragiles soient produits à partir de particules nucléosomales dynamiquement remodelées (Kubik, Bruzzone, et Shore 2017). Ces nucléosomes fragiles auraient une conformation altérée de l'octamère d'histone ou de l'enroulement de l'ADN autour de l'octamère d'histone. L'incorporation de variants d'histone pourrait favoriser ces altérations. Les altérations de composition en histone sont susceptibles d'aboutir à la formation de particules subnucléosomales de type hexasome, tétrasome ou hémisome (Kubik, Bruzzone, et Shore 2017). Cependant, le lien entre nucléosomes fragiles et ces structures n'a pas encore été expérimentalement prouvé (Kubik, Bruzzone, et Shore 2017).

2) Description des particules nucléosomales et subnucléosomales existantes

Comme expliqué dans la revue de 2009 de Zlatanova *et al.*, malgré des preuves *in vivo* de la dynamique des nucléosomes autant dans leur composition que dans leur conformation, c'est une vision stable et inchangée du nucléosome qui a perduré dans le temps chez les scientifiques (Zlatanova et al. 2009). Or, le nucléosome est une « famille de particules dynamiques », avec des variations de composition en histones et de conformation de l'ADN (Zlatanova et al. 2009).

Dans ma thèse, je nomme « particules subnucléosomales » toutes les particules qui ont un poids moléculaire inférieur au nucléosome canonique et qui ont une variation dans le contenu en histones par rapport au nucléosome canonique.

Dans la revue Zlatanova *et al.*, les auteurs proposent de renommer les différentes particules nucléosomales et subnucléosomales décrites dans la littérature afin de clarifier leurs noms par rapport à leur contenu en histones et au sens d'enroulement de l'ADN autour des histones (**Tableau 1**). Pour ne pas trop perturber les lecteurs scientifiques, le terme historique de « nucléosome » peut continuer à être utilisé pour désigner l'octasome (Zlatanova et al. 2009). On distingue donc les octasomes (ou nucléosomes), les hexasomes, les tétrasomes, les hémisomes et les *split* nucléosomes (**Tableau 1**) (Zlatanova et al. 2009). Je vais détailler les connaissances actuelles sur chacune de ces particules de la plus grosse (octasome) à la plus petite (tétrasome), en terminant par les plus exotiques (hémisome et *split* nucléosome).

a) Les octasomes

La particule nucléosomale canonique, comme unité répétée de la chromatine, a été découverte en 1974 et sa structure à haute résolution a été décrite en 1997 (K. Luger et al. 1997). Le nucléosome canonique, dans lequel les huit histones s'organisent au sein de l'octamère en un tétramère H3/H4 et deux dimères H2A/H2B, est également appelé octasome (**Tableau 1 ; Figure 3b**) (Zlatanova et al. 2009). La structure tridimensionnelle (3D) du nucléosome montre que 147 pb d'ADN sont enroulés au sein du nucléosome (K. Luger et al. 1997).

Cependant il y a de nombreuses preuves expérimentales qui montrent que l'octamère d'histone peut être enroulé par 100 à 170 pb d'ADN (lu dans (Zlatanova et al. 2009)). Par exemple, il a été montré qu'entre 118 pb et 130 pb d'ADN sont enroulés autour des octamères contenant le variant d'histone H2A.Bbd (présent chez l'Homme et la souris) (Bao et al. 2004; Doyen et al. 2006).

Noms traditionnels	Noms suggérés	Composition/stœchiométrie en histones
nucléosome	nucléosome ou octasome	H2A/H2B (H3/H4) ₂ H2A/H2B
hexasome	hexasome	H2A/H2B (H3/H4) ₂ ou (H3/H4) ₂ H2A/H2B
tétrasome	tétrasome	(H3/H4) ₂
<i>split</i> nucléosome	<i>split</i> nucléosome	H2A/H2B/H3/H4 + H3/H4/H2A/H2B
<i>half</i> -nucléosome	hémisome	H2A/H2B/H3/H4
nucléosome centromérique	hémisome centromérique	H2A/H2B/CenH3/H4

Tableau 1 : Les noms traditionnels et suggérés par Zlatanova et al des particules nucléosomales et subnucléosomales existantes.

La première colonne comporte les noms traditionnels des particules nucléosomales et subnucléosomales. La deuxième colonne comporte les noms suggérés par (Zlatanova et al. 2009). La troisième colonne comporte la composition en histones H2A, H2B, H3 et H4 des différentes particules. En gris, les particules contenant 8 histones, en vert les particules contenant 6 histones, en bleu les particules contenant 4 histones (2 H3 et 2 H4) et en orange les particules contenant un exemplaire des 4 histones (H2A, H2B, H3 et H4). L'hémisome centromérique ne comporte pas l'histone H3 mais un variant, l'histone CenH3. D'après (Zlatanova et al. 2009).

Bien que le variant d'histone H2A.Bbd partage environ 48% d'homologie de séquence avec l'histone H2A, H2A.Bbd se distingue de H2A essentiellement au niveau du domaine qui lie le tétramère H3-H4 au dimère H2A-H2B, mais aussi par l'absence de queue C-terminal (lu dans (Bao et al. 2004)). Il a été montré qu'en présence de fortes concentrations en sel, les octasomes contenant le variant H2A.Bbd ont une conformation plus facilement changeante que les octasomes canoniques (Eirín-López, Ishibashi, et Ausió 2008). Zlatanova *et al.* proposent que l'enroulement de l'ADN nucléosomal autour de l'octamère d'histone, et donc la stabilité du nucléosome, puisse être contrôlé par les variants d'histone.

Des changements de conformation de l'ADN nucléosomal peuvent également venir d'un déroulement partiel (ouverture) ou de mouvements de respiration de l'ADN nucléosomal (**Figure 3a**) (lu dans (Zlatanova et al. 2009)). L'ouverture ou la respiration rend accessible des sites de liaison pour différents facteurs qui étaient au préalable inaccessibles car empaquetés dans le nucléosome (Zlatanova et al. 2009). Les phénomènes d'ouverture et de respiration sont des phénomènes spontanés d'ouverture de l'ADN au niveau du point d'entrée et de sortie du nucléosome. Il s'agit d'une dynamique naturelle qui ne fait pas intervenir d'enzyme *a priori*. L'ADN est rendu plus accessible à ces points d'entrée et de sortie qu'au centre du nucléosome.

Un autre type de dynamique nucléosomale a été prédite. Il s'agit de la « *gaping* » transition (**Figure 3b**) (Mozziconacci et Victor 2003). Contrairement à l'ouverture et à la respiration qui impliquent des ruptures entre ADN et histones, le phénomène de *gaping* implique une rupture de liaison entre histones. En effet, cette rupture de liaison aurait lieu entre l'histone H2A et H3/H4 (**Figure 9a**) (Mozziconacci et Victor 2003). Cette rupture s'accompagnerait par une ouverture d'environ 30° au niveau des deux résidus Cystéine 110 des deux histones H3 qui interagissent dans le tétramère (Mozziconacci et Victor 2003).

b) Les hexasomes

Un hexasome est une particule subnucléosomale contenant un tétramère H3/H4 et un dimère H2A/H2B (**Tableau 1 ; Figure 3b**). Les hexasomes possèdent 30 à 40 pb d'ADN nucléosomal qui ne sont pas empaquetés du côté du dimère H2A/H2B manquant (**Figure 3b**) (lu dans (Brehove et al. 2019)). Cette particule subnucléosomale a été décrite *in vivo* et *in vitro* et serait formée pendant la transcription (lu dans (Zlatanova et al. 2009)). Les hexasomes peuvent aussi exister *in vivo* au cours de la réplication, de la réparation et en tant que produits transitoires de l'échange incessant des dimères (lu dans (Lavelle et Prunell 2007)).

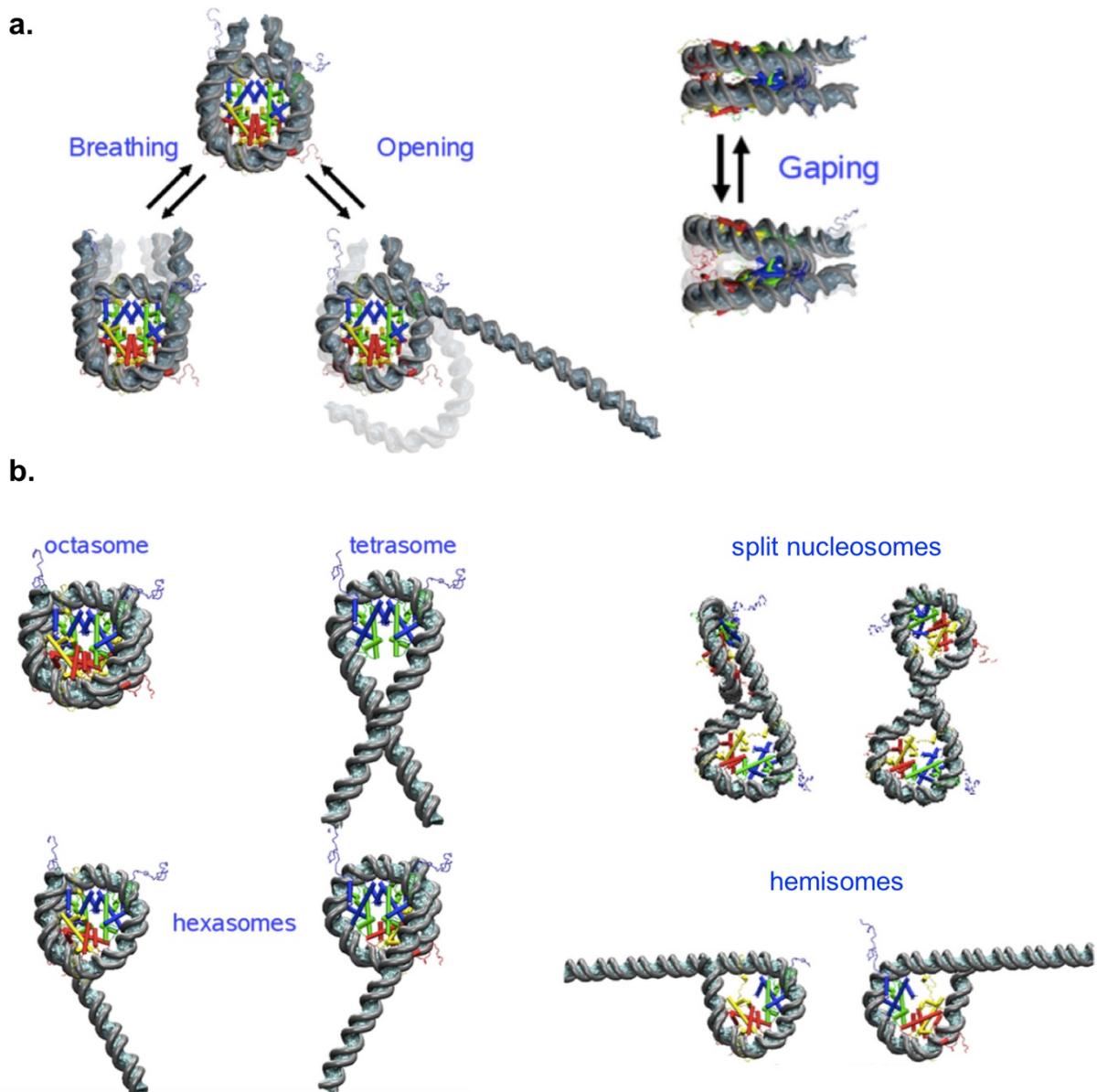


Figure 3 : Modélisation de la dynamique des nucléosomes canoniques et des particules subnucléosomales existantes

La double hélice d'ADN est représentée par une hélice grise et turquoise. Les histones sont représentées par différentes couleurs : jaune pour H2A, rouge pour H2B, bleu pour H3 et vert pour H4.

a. Modélisation de la dynamique des nucléosomes canoniques suivant les mécanismes l'ouverture (« opening »), de respiration (« breathing ») ou de *gapping*. **b.** Modélisation de différentes particules subnucléosomales : un octasome, deux hexasomes, un tétrasome, deux *split* nucléosomes (deux vues différentes) et deux hémisomes. Les histones sont toutes dans la même orientation. Ces modélisations, réalisées grâce à la représentation « all-atom » de la PDB, utilisent comme base la structure du nucléosome canonique avec 147 pb d'ADN obtenue au rayon X à une résolution de 1,9 Å (Davey et al. 2002) où des délétions d'histone et des changements de conformation d'ADN ont été autorisées (Zlatanova et al. 2009). D'après (Zlatanova et al. 2009).

La dynamique d'échange du dimère H2A/H2B a été montrée dans des cellules vivantes grâce à des expériences de *pulse/chase* des histones (Jackson 1990). Cette dynamique d'échange serait en partie dépendante de l'activité de l'ARN Pol II ((Jackson 1990) et lu dans (Zlatanova et al. 2009)).

Grâce au microscope de force atomique, Bintu *et al.* ont obtenu des images de la Pol II en complexe avec un nucléosome en amont, pendant et en aval de la transcription (Bintu et al. 2011). Ces images mettent en évidence la présence d'hexasomes qui sont des produits de remodelage de l'octasome (Bintu et al. 2011). La fraction d'hexasome présente dépend du taux d'élongation de la transcription (Bintu et al. 2011). Les cinétiques de compétition entre l'élongation de la transcription, le transfert d'histones et la dissociation des histones ont été intégrées pour proposer un modèle (Bintu et al. 2011). Dans ce modèle de compétition, les histones des nucléosomes remodelés au cours de l'élongation de la transcription sont exposées. Lorsque la Pol II se lie à l'ADN, cela entraîne une pliure de l'ADN qui va positionner les histones exposées près de l'ADN en amont de la Pol II et va donc entraîner le transfert d'histone sur cette portion d'ADN via la formation d'une boucle d'ADN (Bintu et al. 2011). Si les histones exposées ne sont pas positionnées près d'un fragment d'ADN, il peut alors y avoir une dissociation d'un dimère d'histones de l'octasome qui va alors devenir un hexasome (Bintu et al. 2011).

Au cours de la transcription, des chaperons d'histone jouent également un rôle important dans l'assemblage et le désassemblage des nucléosomes. Le chaperon FACT, constitué des sous-unités SPT16 and SSRP1, a été complexé *in vitro* avec le dimère H2A/H2B, puis avec le tétramère H3/H4 et 79 pb d'ADN afin de caractériser la structure de ce complexe FACT-particule subnucléosomale (Y. Liu et al. 2020). FACT a plusieurs actions : en absence d'ADN, il maintient le tétramère d'histone au cours des premières étapes de l'assemblage du nucléosome et il est lié au dimère H2A/H2B grâce au domaine C-terminal de sa sous-unité SPT16 (Y. Liu et al. 2020). Lorsque l'ADN s'enroule autour du tétramère H3/H4, FACT forme de nouvelles interactions avec H3/H4 qui favorisent l'incorporation du dimère H2A/H2B au niveau du tétrasome pour former un hexasome (Y. Liu et al. 2020). Cette particule subnucléosomale est stabilisée par le domaine C-terminal de SPT16 qui maintient les

interactions entre l'ADN et le dimère H2A/H2B (Y. Liu et al. 2020). La sous-unité SSRP1 va ajouter un deuxième dimère H2A/H2B sur le côté distal du nucléosome (Y. Liu et al. 2020).

FACT pourrait également intervenir dans le désassemblage du nucléosome via la réversion des étapes précédentes (Y. Liu et al. 2020). Le mécanisme d'action de FACT proposé par l'équipe de Luger est cohérent avec les observations en absence des sous-unités SPT16 ou SSRP1 dans les cellules de drosophile (Ramachandran, Ahmad, et Henikoff 2017). La déplétion de la sous-unité SPT16 entraîne une augmentation de 15% de la perte du dimère H2A/H2B proximal par rapport à la Pol II et la déplétion de la sous-unité SSRP1 entraîne une augmentation de 15% de la perte du dimère H2A/H2B distal par rapport à la Pol II (Ramachandran, Ahmad, et Henikoff 2017). Les auteurs ont également montré que la déplétion de chacune des sous-unités de FACT, qui interagissent avec les dimères H2A/H2B opposés de la particule nucléosomale, a pour conséquence une augmentation de la présence de particules de 90 pb au niveau du nucléosome +1 (Ramachandran, Ahmad, et Henikoff 2017). Ces particules de 90 pb traduisent la perte de dimère H2A/H2B conduisant à la formation d'hexasome (Ramachandran, Ahmad, et Henikoff 2017). FACT est donc nécessaire au maintien de la structure du nucléosome au cours de la transcription.

Enfin, en 2019, Brehove *et al.* ont réussi à orienter l'hexasome *in vitro* et ils ont pu ainsi démontrer que la région d'ADN libérée par le dimère H2A/H2B dans l'hexasome est beaucoup plus accessible pour la liaison des TFs (Brehove et al. 2019).

c) Le tétrasome

Le tétrasome est une particule subnucléosomale contenant deux dimères H3/H4 (**Tableau 1 ; Figure 3b**). Contrairement à l'octasome, l'ADN enroulé autour du tétrasome représente moins d'un super tour d'hélice (**Figure 3b**) (Alilat et al. 1999).

Au cours du processus de réplication de l'ADN, les protéines chaperons, comme CAF-1, assemblent les nucléosomes sur l'ADN après le passage de la polymérase réplivative. Les chaperon assemblent en premier deux dimères H3/H4 avec environ 60 pb d'ADN pour former

le tétrasome (Karolin Luger, Dechassa, et Tremethick 2012). D'autres chaperons ajoutent ensuite à ce tétramère deux dimères d'histone H2A/H2B pour former l'octamère autour duquel s'enroulent 147 pb d'ADN (Karolin Luger, Dechassa, et Tremethick 2012).

d) L'hémisome et le *split* nucleosome

En 1976, Weintraub *et al.* ont pour la première fois proposé que le nucléosome puisse s'ouvrir en deux demi-nucléosomes contenant chacun un exemplaire des quatre histones (H. Weintraub, Worcel, et Alberts 1976). Cela faisait référence aux hémisomes et aux *split* nucléosomes.

Les hémisomes sont des particules subnucléosomales uniques contenant un exemplaire de chaque histone (H3, H4, H2A et H2B) (**Tableau 1**). Des particules de type hémisome (CenH3/H4/H2A/H2B) ont aussi été détectées au niveau des centromères (**Tableau 1**). En effet, en 2007, une particule subnucléosomale tétramérique contenant Cid, le variant d'histone H3 des centromères chez la Drosophile, H4, H2A et H2B a été identifié par l'équipe du Dr S. Henikoff (Dalal, Wang, et al. 2007). Ces particules contenant le variant de H3 sont au maximum des tétramères, tandis que les particules contenant l'histone H3 sont au maximum des octamères (Dalal, Wang, et al. 2007). Le poids des particules contenant le variant d'histone centromérique, est d'environ 68 kDa, ce qui correspond bien à un tétramère hétérotypique (Dalal, Wang, et al. 2007). D'après (Dalal, Wang, et al. 2007), ces hémisomes centromériques ne résultent pas de la division d'un octamère canonique en deux tétramères, mais de l'assemblage d'un dimère CenH3/H4 et d'un dimère H2A/H2B. Enfin, les hémisomes centromériques existeraient dans d'autres organismes comme la levure *S.pombe* et l'Homme (Dalal, Furuyama, et al. 2007).

Les *split* nucléosomes sont des particules subnucléosomales formées par deux hémisomes (deux particules contenant H3/H4/H2A/H2B) reliés entre eux par de l'ADN linker (**Tableau 1 ; Figure 3b**) (Zlatanova et al. 2009).

En 1983, Prior *et al.* ont identifié des particules situées au niveau de l'ADN ribosomal actif transcriptionnellement chez les Physarum qui sont des organismes eucaryotes unicellulaires

contenant de nombreux répliquats de leur noyau (Prior et al. 1983). Ces particules identifiées sont constituées de 145 pb d'ADN enroulées autour des quatre histones présentes en quantité équivalentes et elles ont un coefficient de sédimentation plus petit que les particules nucléosomales canoniques (5S *versus* 11S) (Prior et al. 1983). Ces particules apparaissent en microscopie électronique comme deux masses symétriques séparées par 40 à 50 pb avec chacune des deux masses ayant un poids moléculaire d'environ 100 kDa, ce qui représente environ la moitié d'un nucléosome canonique (Prior et al. 1983).

En 1996, l'équipe de Allfrey ont purifié ces particules subnucléosomales sur colonnes d'affinité de mercure en utilisant les propriétés des groupements sulfhydryl de H3 puis la microscopie électronique a ensuite permis de visualiser la forme de ces particules nucléosomales (Bazett-Jones et al. 1996). Leurs données suggèrent un changement de conformation de l'ADN nucléosomal qui serait compatible avec la formation de la particule et une séparation de l'interface H3/H3 (Bazett-Jones et al. 1996).

Le mécanisme de formation des *split* nucléosomes est inconnu. Sa formation implique la dissociation des interactions au niveau de l'interface H3/H3, qui stabilisent le tétramère H3/H4. On peut penser que des enzymes tels que les remodeleurs de la chromatine puissent être responsables d'une telle activité. Par exemple, il est connu que le remodeleur EP400 peut catalyser l'incorporation du variant H3.3 au niveau de la région promotrice des gènes, ce qui implique une disruption du tétramère (H3/H4)₂ au sein de l'octasome (Pradhan et al. 2016).

Chapitre 2 : Les éléments cis-régulateurs de la transcription chez les mammifères et les différents facteurs qui s'y lient pour réguler l'expression des gènes

Il existe plus de 200 types de cellules distinctes dans le corps humain, dont l'identité est définie et maintenue grâce au contrôle de l'expression des gènes. Pour assurer ce contrôle, les cellules disposent d'environ 23 000 gènes, qui sont répartis sur notre génome, et environ 1 600 TFs pour les réguler (Lambert et al. 2018). A proximité de chaque gène, l'ADN comporte aussi des régions régulatrices. Ces régions régulatrices vont être reconnues et liées par des TFs au niveau de séquences spécifiques pour déterminer où et quand exprimer chaque gène. En effet, une fois liés à leur élément régulateur cible, les TFs facilitent le recrutement de l'ARN polymérase II (Pol II) et des facteurs généraux de transcription associés qui vont permettre la transcription des gènes.

Cependant, lorsque l'ADN est enroulé sous forme de nucléosome, il est généralement inaccessible aux TFs. Pour que les TFs puissent se lier à l'ADN dans le but de réguler l'expression des gènes, le nucléosome doit être transitoirement dissocié par des enzymes telles que les remodeleurs de la chromatine, qui peuvent réguler la présence et le positionnement des nucléosomes.

Dans cette partie, je vais présenter la fonction des différents éléments régulateurs et des facteurs qui s'y lient. Je présenterai aussi le rôle des complexes Polycomb et de quelques complexes co-activateurs dans la régulation de la transcription. Enfin j'expliquerai comment la chromatine influe sur le fonctionnement des éléments cis-régulateurs de la transcription.

I. La fonction des éléments promoteurs et des facteurs associés dans la régulation de la transcription mammifère

La transcription est un processus biologique au cours duquel la séquence nucléotidique d'un gène va être lue par une ARN polymérase qui va la convertir en ARN. La transcription est contrôlée par trois types de polymérases ; dans ma thèse, je vais me focaliser sur l'ARN

polymérase II qui permet la production d'un ARN messager (ARNm) à partir de la séquence d'un gène codant.

Plusieurs étapes se succèdent dans l'initiation de la transcription : la reconnaissance du promoteur au niveau duquel le complexe de pré-initiation de la transcription (PIC) s'assemble et l'échappement du promoteur.

1) Caractéristiques et structure des promoteurs

Les promoteurs sont des séquences cis-régulatrices situées en amont du site d'initiation de la transcription (TSS) des gènes (Dillon et Sabbattini 2000). Les éléments promoteurs sont des régions connues pour être hypersensibles à la désoxyribonucléase I (DNase I), une enzyme utilisée dans des essais pour tester l'accessibilité de l'ADN, *in vivo* et donc soumises à un important remodelage de la chromatine qui les rend accessibles pour les différents facteurs (Thurman et al. 2012). La fonction principale des promoteurs est de recruter la Pol II, des TFs, ainsi que des enzymes modifiant les histones (Dillon et Sabbattini 2000; Heintzman et al. 2009; Vernimmen et Bickmore 2015).

Chez les eucaryotes, le promoteur est constitué de plusieurs éléments. Le cœur du promoteur, ou *core promoter* en anglais, est une courte séquence nucléotidique située autour du TSS qui recrute la Pol II et des facteurs généraux de transcription pour former le PIC et initier la transcription (Roeder 1996). Le *core* promoteur, aussi appelé promoteur minimal, est suffisant pour contribuer à l'initiation de la transcription (Kadonaga 2012), mais il a un faible niveau d'activité basal (Spitz et Furlong 2012; Zabidi et Stark 2016). Son activité dépend d'éléments cis-régulateurs additionnels tels que les enhancers.

Les *core* promoteurs actifs ont été classés en deux catégories chez les métazoaires : les *core* promoteurs dits « focalisés » et les *core* promoteurs dits « dispersés » (Carninci et al. 2006). Comme tous les promoteurs actifs, ils partagent des marques d'histone communes : la triméthylation de la lysine 4 de l'histone H3 (H3K4me3) et l'acétylation de la lysine 27 de l'histone H3 (H3K27ac) ([Tableau 2](#)).

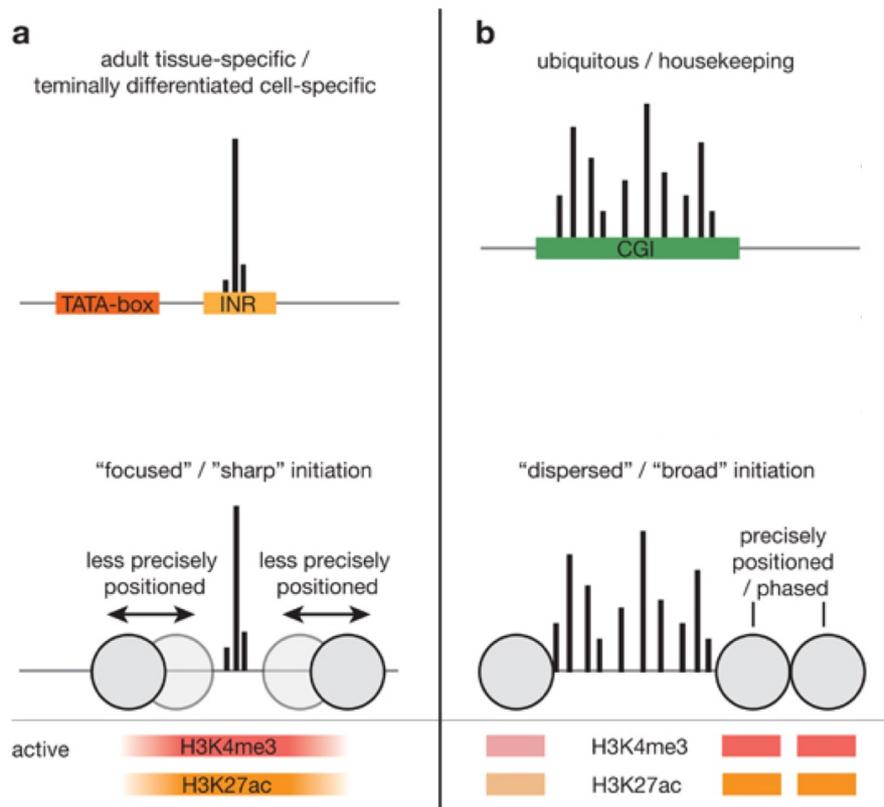


Figure 4 : Les différents types de *core* promoteurs actifs chez les mammifères

Les différents types de *core* promoteurs actifs sont définis par plusieurs critères comme le type cellulaire, la composition de la séquence, le positionnement des nucléosomes et l'enrichissement en marques d'histone. **a.** *Core* promoteurs présents dans les tissus adultes spécialisés et **b.** *Core* promoteurs présents au niveau des gènes de ménage ou des gènes à expression ubiquitaire. Partie supérieure de la figure : INR : élément Initiator, CGI : îlot CpG. Partie inférieure de la figure : les cercles gris représentent les nucléosomes. Les différents rectangles représentent le positionnement de l'enrichissement en marque d'histone, rouge pour H3K4me3, orange pour H3K27ac et bleu pour H3K27me3. D'après et adapté de (Haberle et Stark 2018).

La distinction entre les différentes classes de *core* promoteurs actifs reposent sur plusieurs critères comme la composition de la séquence, la configuration de la chromatine et la fonction des gènes dont ils contrôlent la transcription (Lenhard, Sandelin, et Carninci 2012; Haberle et Stark 2018).

Les *core* promoteurs focalisés sont caractérisés par la présence d'un TSS unique et bien défini, par des nucléosomes positionnés de façon peu précise et par la présence des motifs des éléments TATA-box et *Initiator* (Figure 4a) (Carninci et al. 2006; Lenhard, Sandelin, et Carninci 2012). Ces *core* promoteurs contrôlent généralement la transcription de gènes spécifiques d'un type cellulaire donné avec des patterns d'expression restreints, notamment dans les

cellules différenciées des tissus adultes (Carninci et al. 2006; Lenhard, Sandelin, et Carninci 2012).

Les *core* promoteurs dispersés sont eux caractérisés par la présence de plusieurs TSS, généralement utilisés avec une fréquence similaire, par un appauvrissement en nucléosomes, mais entouré de part et d'autre par des nucléosomes positionnés avec précision et enfin par un chevauchement d'îlots CpG individuels chez les mammifères (Carninci et al. 2006; Lenhard, Sandelin, et Carninci 2012) (**Figure 4b**). Ces *core* promoteurs sont présents en amont des gènes de ménage ou de gènes exprimés de façon ubiquitaire (Carninci et al. 2006; Lenhard, Sandelin, et Carninci 2012; Haberle et Stark 2018). Dans les cellules souches embryonnaires (ES) mammifères, on les retrouve également au niveau des gènes bivalents qui codent fréquemment des TFs impliqués dans le développement et la morphogenèse (Lenhard, Sandelin, et Carninci 2012; Haberle et Stark 2018). Les gènes bivalents sont des gènes qui portent à la fois les marques activatrices (H3K4me3) et les marques répressives (triméthylation de la lysine 27 de l'histone H3 : H3K27me3) de la transcription. La présence de ces deux catégories de marque d'histone reflète le recrutement au niveau des mêmes promoteurs de machineries enzymatiques qui contrôlent d'une part la triméthylation de la lysine 4, et d'autre part des complexes de type Polycomb, notamment PRC2, qui dépose H3K27me3. La présence des Polycomb a pour résultat un niveau d'expression relativement faible des gènes bivalents.

On distingue cinq éléments majeurs appartenant au cœur du promoteur humain : l'élément TATA, l'élément *Initiator*, l'élément DPE, l'élément BRE et l'élément DCE. La TATA box est le premier élément *core* promoteur qui a été identifié chez les eucaryotes. La TATA box est une séquence nucléotidique riche en AT localisée entre 25 et 30 pb en amont du TSS des *core* promoteurs focalisés (lu dans (Haberle et Stark 2018)). Alors qu'on a longtemps cru que la TATA box était indispensable à l'initiation de la transcription, on s'est ensuite rendu compte que seulement 10 à 20% des gènes humains contiennent une TATA box au niveau des promoteurs des gènes transcrits (Gershenson et Ioshikhes 2005; S. J. Cooper et al. 2006). La TATA box est également le site d'assemblage PIC et le site privilégié de liaison de la protéine TBP (*TATA binding protein*) (Patikoglou et al. 1999; Venters et Pugh 2009). L'élément *Initiator* (Inr), situé autour du TSS de *core* promoteurs focalisés, peut initier de façon précise la

transcription seul ou en partenariat avec la TATA box (Venters et Pugh 2009). L'élément Inr est lié par les facteurs TAF1 (*TBP Associated Factor 1*) et TAF2 (*TBP Associated Factor 2*) (Chalkley 1999).

2) Différents facteurs participent à l'initiation de la transcription au niveau des promoteurs

a) L'ARN polymérase II

La structure de la Pol II de levure a été établie par diffraction de rayons X sur cristaux (Cramer, Bushnell, et Kornberg 2001; Armache, Kettenberger, et Cramer 2003). La Pol II est un complexe multiprotéique de 550 kDa, composée de 12 sous-unités (Armache, Kettenberger, et Cramer 2003).

La sous-unité RPB1 de la Pol II contient un domaine C-terminal (CTD) composé d'une série de répétitions d'un heptapeptide : YSPTSPS (Tyrosine-Sérine-Proline-Thréonine-Sérine-Proline-Sérine) qui peut être phosphorylé (Stephen Buratowski 2009) et qui est impliqué dans les processus d'initiation et la transition vers l'élongation de la transcription par la Pol II. Le nombre de répétitions de l'heptapeptide varie entre 26 chez la levure *Saccharomyces cerevisiae* (Allison et al. 1985) à 52 chez les mammifères (Corden et al. 1985). La phosphorylation du CTD au niveau de la Sérine n°5 est trouvée principalement au niveau des promoteurs et est associée à l'initiation de la transcription (Venters et Pugh 2009).

Plusieurs fonctions de la queue CTD de la Pol II ont été décrites : elle servirait de plate-forme de liaison pour les protéines jouant un rôle dans la transcription, elle serait requise pour le processus efficace de maturation des ARNm (coiffage en 5', clivage et polyadénylation en 3' et épissage des introns), ainsi que pour certaines modifications des histones associées à la transcription (lu dans (Stephen Buratowski 2009)).

b) Les facteurs généraux de transcription sont des composants du complexe de pré-initiation de la transcription

L'ARN Pol II est incapable d'initier seule la transcription chez les eucaryotes. L'initiation de la transcription nécessite la présence de facteurs de transcription généraux (GTF). Ces facteurs vont positionner correctement la Pol II au niveau du promoteur, puis vont ouvrir les deux brins d'ADN pour initier la transcription, puis libérer l'ARN polymérase du promoteur pour permettre l'élongation de la transcription. Cet échafaudage de GTFs en association avec l'ARN Pol II au niveau du promoteur est appelé le PIC (*PreInitiation Complex*).

Ces GTF sont au nombre de six : TFIIA, TFIIB, TFIID, TFIIE, TFIIF et TFIIH. Ils comportent tous dans leur nom « TFII » qui signifie *Transcription Factor for polymerase II*. Ces différents facteurs se lient avec une chorégraphie bien précise au niveau des éléments TATA-box et Inr des *core* promoteurs (Zabidi et Stark 2016).

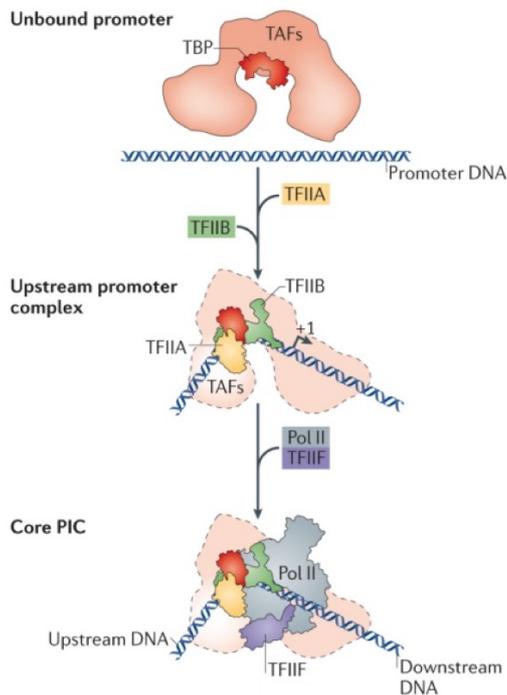
Le premier GTF à intervenir est TFIID (**Figure 5a**). Des études permettant de résoudre la structure des complexes protéiques ont montré que TFIID est le principal TF qui reconnaît le *core* promoteur (Cianfrocco et al. 2013; Sainsbury, Bernecky, et Cramer 2015; Louder et al. 2016). TFIID est composé de la sous-unité TBP (pour *TATA Binding Protein*) et de 12 à 14 sous-unités TAFs (pour *TBP-Associated Factors*). TBP est un facteur de transcription eucaryote universel (Hernandez 1993) qui, lorsqu'il est libre, a une structure en forme de selle (Chasman et al. 1993). TFIID, et plus particulièrement TBP, va se lier au niveau de la TATA box non loin du promoteur du gène à transcrire (Nakajima, Horikoshi, et Roeder 1988; Hoffmann et al. 1990). TBP semble nécessaire pour initier la transcription au niveau de tous promoteurs (Sainsbury, Bernecky, et Cramer 2015). La fixation de TFIID entraîne une forte distorsion de l'ADN au niveau de la TATA box, ce qui permet de rapprocher les différents facteurs pour poursuivre l'assemblage (**Figure 5a**).

TFIIA et TFIIB sont ensuite recrutés (**Figure 5a**). TFIIA n'est pas requis pour la transcription basale, mais il peut permettre la stabilisation de TBP sur l'ADN au niveau de la TATA box (Imbalzano, Zaret, et Kingston 1994) et la stabilisation des interactions ADN-TAFs (Geiger et al. 1996; Tan et al. 1996). TFIIB va permettre le recrutement d'un complexe formé par TFIIF associé avec l'ARN Pol II (Ha et al. 1993). TFIIB, via son domaine N-terminal, est requis pour le

recrutement de l'ARN Pol II au niveau du promoteur. Le domaine C-terminal de TFIIB stabilise l'interaction de TBP avec l'ADN au niveau du promoteur (S. Buratowski et Zhou 1993). TBP, TFIIB, TFIIF, l'ARN Pol II et l'ADN du promoteur forment le cœur du PIC (Mühlbacher et al. 2014) (**Figure 5a**).

TFIIF va déstabiliser les interactions ADN - Pol II non-spécifiques (Conaway et al. 1991). TFIIF va par contre stabiliser TFIIB au sein du PIC et il va recruter THIE et TFIH (Čabart et al. 2011). L'hétérodimère TFIE se lie à la Pol II (Flores, Maldonado, et Reinberg 1989) et facilite le recrutement de TFIH au niveau du PIC, en faisant une sorte de pont entre la Pol II et TFIH (**Figure 5b**) (Holstege, van der Vliet, et Timmers 1996). Les différentes sous-unités de TFIH possèdent différentes activités : ses activités hélicase et ATPase permettent notamment l'ouverture de l'ADN au niveau du promoteur (Guzmán et Lis 1999) au moment de l'initiation de la transcription et lors de la réparation par excision de nucléotides de l'ADN endommagé (Coin, Oksenysh, et Egly 2007). TFIH porte aussi une activité kinase qui permet la phosphorylation du CTD de la Pol II et le démarrage de la transcription (lu dans (Sainsbury, Bernecky, et Cramer 2015)). TFIE régule également les activités ATPase et kinase de TFIH (Ohkuma et Roeder 1994). TFIE, TFIH et le cœur du PIC forment le PIC complet (**Figure 5b**).

a.



b.

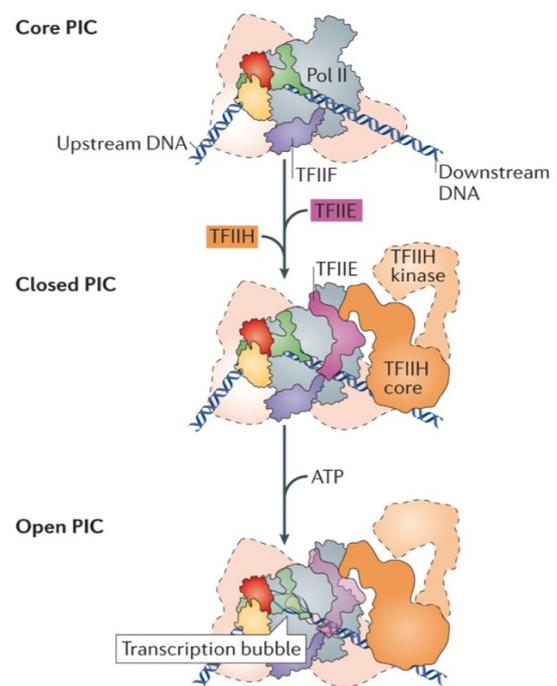


Figure 5 : Implication des GTFs dans la formation du PIC et dans l'initiation de la transcription

Les différentes étapes de l'assemblage du PIC : de la fixation du premier facteur TFIID au cœur du PIC (a.) et du cœur du PIC jusqu'à l'ouverture du PIC (b.).

Le facteur TFIID est représenté par une forme de couleur rouge brique. TFIIA en jaune, TFIIB en vert, Pol II en gris bleu, TFIIF en violet, TFIIE en rose, TFIIH en orange. D'après (Sainsbury, Bernecky, et Cramer 2015).

c) Certains TFs se lient au niveau des promoteurs pour réguler la transcription

En 2008, Chen *et al* ont décrit dans les cellules ES comment les éléments cis-régulateurs du génome sont liés par différentes catégories de TFs (X. Chen et al. 2008a).

Certains TFs se lient préférentiellement au niveau des promoteurs des gènes. C'est le cas des TFs c-MYC, KLF4 et E2F1 (X. Chen et al. 2008a). Les facteurs KLF4 et c-MYC ont été identifiés, avec les TFs OCT4 et SOX2, comme étant essentiels à la reprogrammation de cellules somatiques en cellules pluripotentes (Takahashi et Yamanaka 2006). Les TFs paralogues à KLF4, KLF2 et KLF5, sont redondants et ces trois TFs participent à l'auto-renouvellement des cellules ES (Jiang et al. 2008). Le TF E2F1 joue un rôle dans la prolifération cellulaires des ES.

II. La fonction des éléments enhancers dans la régulation de la transcription

1) Caractéristiques générales des enhancers

Dans les cellules ES mammifères, les enhancers sont des séquences cis-régulatrices situées le plus souvent dans des régions non-codantes du génome (Dillon et Sabbattini 2000). Ils sont localisés à une distance plus ou moins grande du promoteur des gènes dont ils régulent l'expression. Les enhancers sont majoritairement tissus-spécifiques (Vernimmen et Bickmore 2015).

Les enhancers sont la cible de nombreux TFs et de complexes capables de modifier la chromatine. Les enhancers des cellules ES sont liés par des TFs associés à la pluripotence, par le médiateur et la cohésine (X. Chen et al. 2008a; Kagey et al. 2010; Whyte et al. 2013). Ils présentent aussi généralement un enrichissement en marques d'histone H3K4me1 (monométhylation de la lysine 4 de l'histone H3) et H3K27ac (Heintzman et al. 2009; Vernimmen et Bickmore 2015). Les enhancers, comme les promoteurs ([Tableau 2](#)), peuvent dans certains cas recruter la Pol II qui peut transcrire de façon bidirectionnelle les enhancers (Core, Waterfall, et Lis 2008; Vernimmen et Bickmore 2015). Toutefois, contrairement à la transcription initiée aux promoteurs, la transcription des enhancers ne produit pas de transcrits stables.

Plusieurs approches complémentaires permettent de détecter les enhancers : hypersensibilité à la DNase I (DNase-seq), ChIP-seq ciblant les marques d'histone H3K27ac et H3K4me1, le médiateurs ou des TFs spécifiques (X. Chen et al. 2008a).

2) Le profil épigénétique des enhancers et l'activité transcriptionnelle définissent différents types d'enhancer

Différents travaux suggèrent l'existence de différentes catégories d'enhancers selon la présence de la machinerie de transcription et l'enrichissement pour certaines marques d'histone. Les enhancers dits « typiques », qui ont une taille d'environ 500 pb et qui vont lier

des TFs spécifiques, peuvent être divisés en trois sous-catégories : les enhancers actifs, *primed* et *poised*.

Les enhancers dits actifs sont fortement enrichis en marques H3K27ac et on peut détecter la présence de la Pol II et des GTF sur ces enhancers (**Tableau 2**) (Rada-Iglesias et al. 2011).

Les enhancers qualifiés de « *poised* » et « *primed* » sont des enhancers qui présentent certaines, mais pas toutes les caractéristiques des enhancers actifs. En effet, les *primed* enhancers sont différents des enhancers actifs par un enrichissement plus modéré en H3K27ac et un faible niveau de transcription (**Tableau 2**) (Rada-Iglesias et al. 2011). Les *poised* enhancers sont enrichis en marque d'histone H3K27me3 déposée par le système Polycomb, et dépourvus de H3K27ac et ils ne sont pas transcrits (**Tableau 2**) (Rada-Iglesias et al. 2011). Bien que les études fonctionnelles de ces éléments soient peu nombreuses, les données actuelles suggèrent que les enhancers *primed* et *poised* sont potentiellement des éléments inactifs en cellules ES, mais prêts à activer la transcription lorsque les cellules ES amorcent leur différenciation.

Enfin, une catégorie d'enhancers un peu particulière a été décrite. Il s'agit des super-enhancers. Un super-enhancer est défini comme un cluster de plusieurs enhancers associés (ressemblant chacun à un enhancer typique) pouvant occuper plusieurs dizaines de kilobases (kb) (Whyte et al. 2013). Comparativement aux enhancers classiques, les super-enhancers ont une taille supérieure et de plus forts enrichissements en médiateur et en marque d'histone H3K27ac (Whyte et al. 2013). Le contenu en TFs est toutefois très similaire à celui des enhancers typiques. Les travaux du laboratoire de R. Young tendent à démontrer que les super-enhancers ont une fonction particulière dans le contrôle de la transcription des gènes contrôlant le phénotype cellulaire (Whyte et al. 2013). Cependant, d'autres études suggèrent que les enhancers typiques et les super-enhancers ont des rôles semblables dans le contrôle de la transcription des gènes en cellules ES (Moorthy et al. 2017).

	Enhancers actif	Enhancers primed	Enhancers poised	Promoteurs actifs
ARN Pol II	+	- ou faible	-	+
eRNA	+	- ou faible	-	-
H3K27ac	fort	modéré	-	+
H3K4me3	-	-	-	+
H3K4me1	+	+	+	modéré
H3K27me3	-	-	+	-
Hypersensibilité à la DNase I	+	+	+	+
TFs associés à la pluripotence	+	+	+	+/-

Tableau 2 : Le niveau de transcription et l'enrichissement en marque d'histone définissent trois catégories d'enhancers typiques : les enhancers actifs, *primed* et *poised*.

La couleur verte et le « + » indiquent la présence de la marque d'histone au niveau des éléments régulateurs, tandis que la couleur rouge et le « - » indiquent l'absence de marque d'histone. La couleur orange indique une absence ou une faible présence de la Pol II et des transcrits ARN des enhancers. Concernant la marque d'histone H3K27ac, on peut détecter un enrichissement fort (vert foncé), modéré (vert clair) ou absent (rouge) en fonction de l'enhancer observé. Pour comparaison, les données au niveau des promoteurs actifs ont été ajoutées. La couleur vert clair et les « +/- » désignent une présence modérée des TFs associés à la pluripotence au niveau des promoteurs actifs par rapport à une présence plus forte aux enhancers (« + » et vert plus foncé).

H correspond à une histone, K à la lysine présence au niveau de la queue des histones, ac à l'acétylation, me1 à la mono-méthylation et me3 à la tri-méthylation. Le tableau est basé sur les données des laboratoire de (X. Chen et al. 2008a; Rada-Iglesias et al. 2011).

3) Les TFs qui se lient aux enhancers en cellules ES

Au cours de leur étude sur la distribution de TFs à grande échelle sur le génome, Chen *et al* ont montré que certains TFs associés à la pluripotence se lient de manière groupée au niveau des enhancers (X. Chen et al. 2008a). Les TFs OCT4 et SOX2 sont connus pour former un hétérodimère qui se lie sur le motif consensus 5'-CATTGTTATGCAAAT-3' (X. Chen et al. 2008a). Le TF NANOG a un motif de fixation très similaire à celui d'OCT4 et SOX2, ce qui suggère que NANOG se lie fréquemment aux sites OCT4 et SOX2 (X. Chen et al. 2008a). Les TFs OCT4, SOX2 et NANOG sont moins abondants au niveau des promoteurs (X. Chen et al. 2008a).

Chen *et al* ont montré que la distribution de CTCF, qui est un TF ayant un domaine de liaison à l'ADN à doigts de zinc, diffère de la distribution des TFs précédents liés au niveau des promoteurs ou des enhancers (X. Chen et al. 2008a). CTCF est enrichi sur de nombreux sites de liaison sur le génome, qui sont parfois incorporés aux promoteurs et aux enhancers, mais CTCF se lie également sur des éléments distincts des promoteurs et des enhancers. Dans certains cas, les sites de liaison de CTCF peuvent avoir des propriétés isolatrices. Ce type de sites de liaison de CTCF sont orientés de façon convergente pour partitionner le génome via la formation de boucles 3D de l'ADN (de Wit et al. 2015). La fonction de CTCF dans le partitionnement du génome 3D ne sera pas développée dans ma thèse.

Le facteur YY1 a également été identifié comme un TF à doigts de zinc (Shi et al. 1991) essentiel pour la viabilité cellulaire chez l'embryon et l'adulte (Donohoe et al. 1999). YY1 est à la fois présent aux promoteurs et aux enhancers actifs (A. S. Weintraub et al. 2017). Des expériences de ChIA-PET, qui détectent les interactions de la chromatine à l'échelle du génome via des protéines spécifiques, ont permis d'étudier les interactions de YY1 avec différentes régions de chromatine. Ainsi, ces expériences de ChIA-PET ont montré, au niveau du locus *Raf1*, que les interactions YY1-YY1 font uniquement intervenir des contacts enhancers-promoteurs, alors qu'au locus *Klf9*, les interactions YY1-YY1 impliquent des contacts entre les super-enhancers et leurs promoteurs cibles (A. S. Weintraub et al. 2017).

III. Fonction des complexes modificateurs de la chromatine au niveau des éléments cis-régulateurs

1) Les complexes Polycomb répriment la transcription

Le système Polycomb repose essentiellement sur deux familles de complexes multiprotéiques : le complexe PRC1 (*Polycomb Repressive Complex 1*) et le complexe PRC2 (*Polycomb Repressive Complex 2*) (Aranda, Mas, et Di Croce 2015). Les sous-unités catalytiques RING1A et RING1B du complexe PRC1 vont ubiquitiner la lysine 119 de l'histone H2A (H2AK119ub) et la méthyltransférase EZH2 avec les sous-unités catalytiques SUZ12 et EED du complexe PRC2 vont mono, di ou tri-méthyle la lysine 27 de l'histone H3 (H3K27me1,

H3K27me2 ou H3K27me3) (Chittock et al. 2017; Laugesen, Højfeldt, et Helin 2019). Le complexe PRC2 contient également les sous-unités RBBP4 et RBBP7 qui lui permettent d'interagir avec les histones (Murzina et al. 2008).

La diméthylation de H3K27 par PRC2 est une marque abondante retrouvée sur 50 à 70 % des histones H3, tandis que la mono et la triméthylation sont retrouvées sur 5 à 10 % des histones H3 (Ferrari et al. 2014; Højfeldt et al. 2018). La marque H3K27me1 est enrichie au niveau du corps des gènes transcrits, la marque H3K27me2 est enrichie au niveau des régions inter et intragéniques et enfin la marque H3K27me3 est enrichie au niveau des promoteurs liés par PRC2, et en plus faible mesure à d'autres positions sur le génome (Jung et al. 2010; Ferrari et al. 2014; Højfeldt et al. 2018).

PRC2 se lie sur des îlots CpG situés au niveau des promoteurs de gènes silencieux et ces îlots CpG sont caractérisés par des conformations d'ADN particulières, un faible niveau de transcription et sont également liés par le complexe PRC1 et par la marque H2AK119ub (Tanay et al. 2007; Ku et al. 2008; Laugesen, Højfeldt, et Helin 2019). La liaison du complexe PRC1 chevauche donc la région de liaison du complexe PRC2 (Boyer et al. 2006).

Plusieurs facteurs sont impliqués dans le recrutement de PRC2 et dans la régulation de son activité catalytique (**Figure 6**). Une sous-unité supplémentaire PLC du complexe PRC2 permet le recrutement de PRC2 au niveau des îlots CpG des promoteurs (Choi et al. 2017; Haojie Li et al. 2017). Le recrutement de PRC2 serait également guidé par des ARN. En effet, un mécanisme de compétition ou d'interférence entre les petits ARN et le complexe PRC2 a été proposé au niveau des gènes transcrits qui empêcherait la fixation de PRC2 (**Figure 6**) (Beltran et al. 2016; X. Wang et al. 2017). Les marques d'histone influencent également la liaison et l'activité de PRC2, comme les marques d'histone H3K36me2/me3 et H3K4me3 qui empêchent la triméthylation de H3K27 à des endroits inappropriés (**Figure 6**) (Schmitges et al. 2011; Haojie Li et al. 2017).

On a pendant longtemps pensé que PRC2 se liait en premier sur sa cible et que la méthylation de H3K27 recrutait le complexe PRC1 (Francis, Kingston, et Woodcock 2004). Toutefois, il a été montré qu'au moins une partie des complexes PRC1 qui déposent la marque H2AK119ub1 sont indépendants du recrutement de PRC2 (Tavares et al. 2012). La marque H2AK119ub1 semble capable de réguler la liaison et l'activité catalytique du complexe PRC2 grâce à son

interaction avec la sous-unité supplémentaire JARID2 (Figure 6) (Kalb et al. 2014; Blackledge et al. 2014; S. Cooper et al. 2016). Enfin différentes sous-unités supplémentaires du complexe PRC2 favorisent la liaison et l'activité de PRC2 au niveau de ces cibles (Figure 6). Par exemple, en plus de la liaison de la partie N-terminale de la sous-unité SUZ12, JARID2, les PLC1-3 et les RBBP4/7 interviennent dans la liaison de PRC2 au niveau de l'ADN (S. Cooper et al. 2016; Haojie Li et al. 2017; Højfeldt et al. 2018). Les sous-unités supplémentaires EPOP et PALI stimuleraient *in vitro* l'activité catalytique de PRC2 (Z. Zhang et al. 2011; Conway et al. 2018).

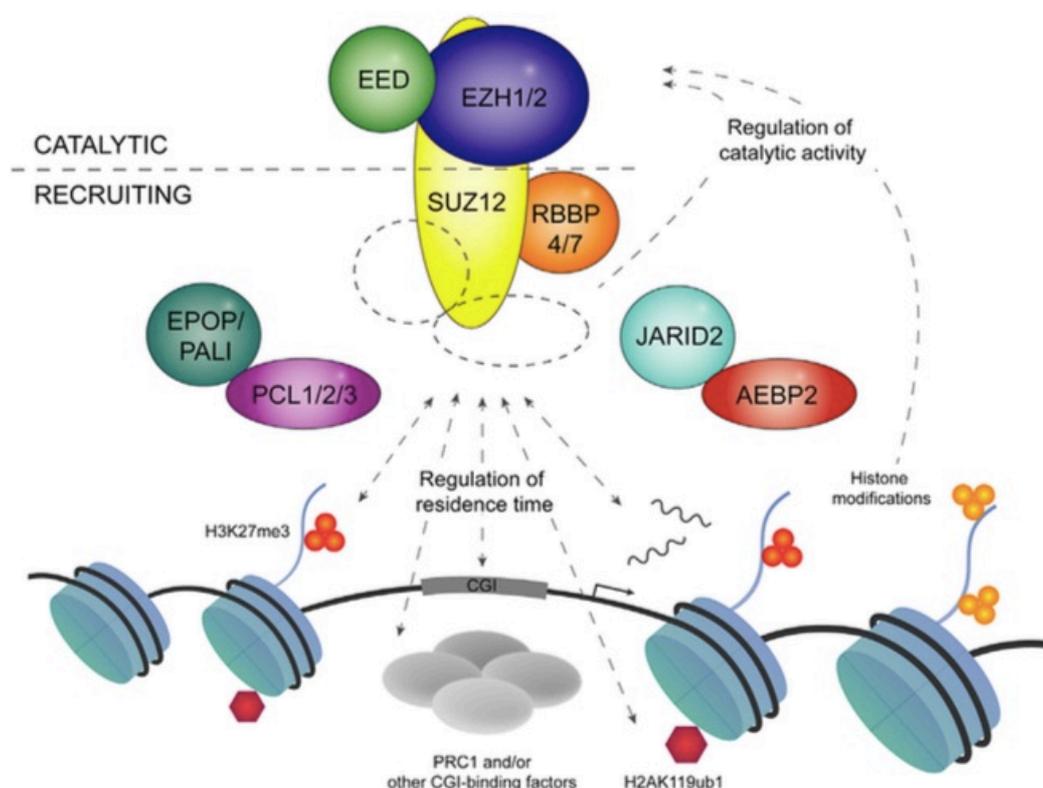


Figure 6 : Différents facteurs influencent le recrutement de PRC2 et module son activité catalytique.

Le cœur du complexe PRC2 est composé des sous-unités SUZ12 (en jaune), EZH1 et EZH2 (en violet) et EED (en vert). Les sous-unités additionnelles RBBP4 et RBBP7 (en orange) permettent la liaison aux histones, les PLC (en rose) permettent la liaison aux îlots CpG et JARID2 (en bleu) permet la liaison à la marque H2AK119ub1. Toutes les sous-unités du cœur ou supplémentaires jouent un rôle dans l'activité catalytique de PRC2.

Le recrutement de PRC2 est modulé par les marques H3K27me3 (les trois cercles rouges) et H2AK119ub1 (hexagone rouge), par la liaison d'autres facteurs comme PRC1 au niveau des îlots CpG, par la présence d'ARN et par les sous-unités supplémentaires.

L'activité catalytique de PRC2 est modulée par les modifications d'histone et par les sous-unités supplémentaires. D'après (Laugesen, Højfeldt, et Helin 2019).

Dans les cellules ES, les complexes Polycomb répriment les gènes exprimés lors du développement en se liant sur leurs promoteurs bivalents. Certains de ces gènes ont également leurs enhancers maintenus dans un état *poised* (Rada-Iglesias et al. 2011). Ce phénomène est cependant restreint car le nombre d'enhancers triméthylés au niveau de K27 est réduit (environ 2 000 enhancers *poised* en cellules ES).

2) Les autres acteurs de la transcription

De nombreux complexes régulateurs de la transcription interviennent au niveau des éléments régulateurs en plus des acteurs décrits précédemment. Dans le cadre de ma thèse, j'ai focalisé mes études sur les facteurs de remodelage de la chromatine et sur les TFs. Je vais très brièvement présenter dans cette partie trois complexes co-activateurs qui participent à la régulation de la transcription : le médiateur, le complexe CBP/p300 et le complexe ATAC/SAGA.

a) Le médiateur

Le médiateur est un complexe co-activateur de la transcription des gènes métazoaires d'environ 30 sous-unités (Soutourina 2018). Le médiateur a pour fonction de coordonner les signaux entre les enhancers et la machinerie générale de transcription grâce à son interaction avec la Pol II et les TFs (Spitz et Furlong 2012).

Chez les mammifères, le médiateur semble se lier de manière abondante au niveau des enhancers et des promoteurs (Kagey et al. 2010; Whyte et al. 2013; Lovén et al. 2013).

Kagey *et al.* ont réalisé des expériences de ChIP-seq qui ont révélées que les sous-unités du médiateur Med1 et Med12 occupent les promoteurs d'au moins 60% des gènes transcrits dans les cellules ES (Kagey et al. 2010). Ils ont aussi montré que le médiateur se lie au niveau des sites occupés par la Pol II et TBP au niveau des promoteurs et au niveau des enhancers liés par des TFs comme OCT4, SOX2 et NANOG (Kagey et al. 2010). Le médiateur favoriserait le contact entre promoteur et enhancer chez les mammifères via la formation d'une boucle d'ADN (Kagey et al. 2010; Lai et al. 2013). Cette boucle d'ADN activerait la transcription des gènes par des mécanismes qui sont mal connus.

Chez les mammifères, le médiateur est également impliqué dans l'initiation de la transcription via le recrutement et la stabilisation des GTF composants le PIC et via la régulation de la phosphorylation de la sérine 5 du CTD de la Pol II, qui permet le début de l'élongation de la transcription (lu dans (Soutourina 2018)).

b) Le complexe CBP/p300

CBP, pour *CREB binding protein*, et p300 sont des acétyltransférases responsables de l'acétylation d'H3K27 au niveau des éléments cis-régulateurs du génome (Heintzman et al. 2009). CBP et p300 partagent une forte identité de séquence au niveau des domaines enzymatiques histone acétyltransférase et des bromodomaines (Zeng et al. 2008; Xin Liu et al. 2008), cependant, ils interagissent avec différents partenaires protéiques (Dancy et Cole 2015).

L'expression de CBP et p300 est régulée au cours du développement pendant l'embryogenèse chez la souris (Partanen, Motoyama, et Hui 1999). Au cours de l'organogenèse, les protéines CBP et p300 sont exprimées dans différents types cellulaires (Partanen, Motoyama, et Hui 1999).

Martire *et al.* ont mis en évidence dans les cellules ES de souris que p300 est responsable de l'acétylation d'H3K27 et du maintien de cette marque au niveau des enhancers, et que p300 et CBP favorisent le maintien de l'expression des gènes dépendants de l'acétylation de leur promoteur (Martire et al. 2020).

La marque H3K27ac déposée par CBP/p300 est considérée comme une marque qui participe à l'activation dans la transcription (Jin et al. 2011), cependant ce complexe co-activateur peut également être associé dans certain cas à de la répression transcriptionnelle (Holmqvist et Mannervik 2013; Kasper et al. 2014). Il a été montré dans la littérature que certaines régions liées par CBP/p300 sont dépourvues de la marque H3K27ac, ce qui suggère que l'activité acétyltransférase de CBP/p300 est inactive au niveau de ces régions (Rada-Iglesias et al. 2011; Holmqvist et Mannervik 2013). Ortega *et al.* ont démontré que la dimérisation et l'activation de TFs comme IRF3 et STAT1 entraîne la trans-autoacétylation de p300 au niveau d'une boucle d'auto-inhibition riche en lysine, qui permet l'activation de la fonction acétyltransférase de

p300 (Ortega et al. 2018). L'activité acétyltransférase semble liée à l'activation des TFs, ce qui suggère un lien entre signalisation cellulaire, liaison du complexe CBP/p300 à l'ADN et transcription des gènes (Ortega et al. 2018).

c) Les complexes ATAC et SAGA

Les complexes mammifères SAGA, pour *Spt-Ada-Gcn5-acetyltransferase*, et ATAC, pour *Ada-Two-A-Containing*, sont composés de nombreuses sous-unités, dont plusieurs communes aux deux complexes, comme la sous-unité GCN5 et son paralogue PCAF qui sont responsables de l'activité histone acétyltransférase (Gamper, Kim, et Roeder 2009; Y.-L. Wang et al. 2008). Les protéines ADA réguleraient l'activité catalytique de GCN5 ou PCAF présentes dans le module histone acétyltransférase des complexes SAGA et ATAC (Gamper, Kim, et Roeder 2009). Ces deux complexes se distinguent par quelques caractéristiques qui leur confèrent des fonctions différentes : ATAC possède une autre sous-unité portant un domaine histone acétyltransférase (Suganuma et al. 2008; Guelman et al. 2009) et SAGA possède un domaine permettant la désubiquitination des histones H2A et H2B (Y. Zhao et al. 2008).

Dans les cellules humaines, le complexe SAGA se lie préférentiellement au niveau des promoteurs de ses gènes cibles tandis que le complexe ATAC se lie au niveau des promoteurs et des enhancers (Krebs et al. 2011). Les deux complexes se lient au niveau de gènes activement exprimés et tissu-spécifiques (Krebs et al. 2011). Les auteurs ont également mis en évidence qu'un sous-ensemble d'enhancers lié par ATAC, mais indépendant de p300, serait régulé par ATAC (Krebs et al. 2011).

Dans les cellules ES et différenciées de souris et dans des cellules humaines, le complexe ATAC et le médiateur forment un complexe co-activateur stable (Krebs et al. 2010). Ce complexe va réguler la transcription d'ARN non-codants par la Pol II dans les cellules ES (Krebs et al. 2010). Enfin, il a été montré dans des cellules ES de souris et dans des cellules HeLa que le médiateur joue un rôle essentiel dans l'assemblage du PIC, tandis que le complexe SAGA permet l'initiation de la transcription après l'assemblage du PIC (X.-F. Chen et al. 2012).

IV. Le fonctionnement des éléments cis-régulateurs de la transcription dans le contexte de la chromatine chez les mammifères

1) Définition et caractéristiques des facteurs pionniers

Les nucléosomes sont un obstacle majeur pour la liaison des TFs sur leurs sites cibles. Cependant, les facteurs pionniers peuvent se lier sur leur séquence cible au niveau de l'ADN qui est engagé dans un nucléosome et dont l'accessibilité est limitée par ses contacts avec les histones (Zaret et Carroll 2011; Spitz et Furlong 2012; Iwafuchi-Doi et Zaret 2016).

Zhu *et al.* ont mis au point une nouvelle approche expérimentale, appelée NCAP-SELEXA et basée sur de l'enrichissement de séquences, qui permet de détecter pour chaque TF dans quelles conditions il peut se lier au nucléosome (Zhu *et al.* 2018). Le NCAP-SELEXA repose sur la mise en contact de ligands d'ADN contenant des adaptateurs avec des octamères d'histone dans lesquels l'histone H2A est étiquetée (**Figure 7a**). Les ligands d'ADN qui se sont enroulés autour des octamères d'histone sont sélectionnés et mis en contact avec des TFs également étiquetés (**Figure 7a**). La purification des TFs liés aux nucléosomes et aux ligands d'ADN se fait par double affinité pour l'étiquette de l'histone H2A et pour l'étiquette du TF. Les ligands d'ADN liés au TF sont séparés en fonction de la présence ou de l'absence de liaison au nucléosomes avant d'être séquencés (**Figure 7a**) (Zhu *et al.* 2018). A l'aide de ce système expérimental, ils ont exploré les préférences de liaison de 220 TFs sur l'ADN empaqueté dans le nucléosome.

Les auteurs ont pu observer que les motifs de liaison des TFs sur l'ADN libre ou sur l'ADN empaqueté dans le nucléosome sont semblables, bien que les motifs sur l'ADN nucléosomal peuvent être dans une position et une orientation différentes sur l'ADN libre à cause de l'encombrement stérique et de la présence des histones (Zhu *et al.* 2018). En fonction des différents facteurs, ils ont identifié différents types de liaison aux nucléosomes (**Figure 7b**). Certains facteurs se lient préférentiellement au niveau de motifs présents sur les deux tours de l'ADN nucléosomal, alors que d'autres facteurs se lient sur un motif présent sur un seul des deux tours de l'ADN au niveau de la dyade du nucléosome, comme les facteurs de la famille SOX (**Figure 7b**) (Zhu *et al.* 2018). D'autres facteurs se lient préférentiellement au niveau des

extrémités de l'ADN nucléosomal, au niveau de positions périodiques sur l'ADN nucléosomal, ou au niveau de motifs présents dans une orientation particulière (Figure 7b) (Zhu et al. 2018).

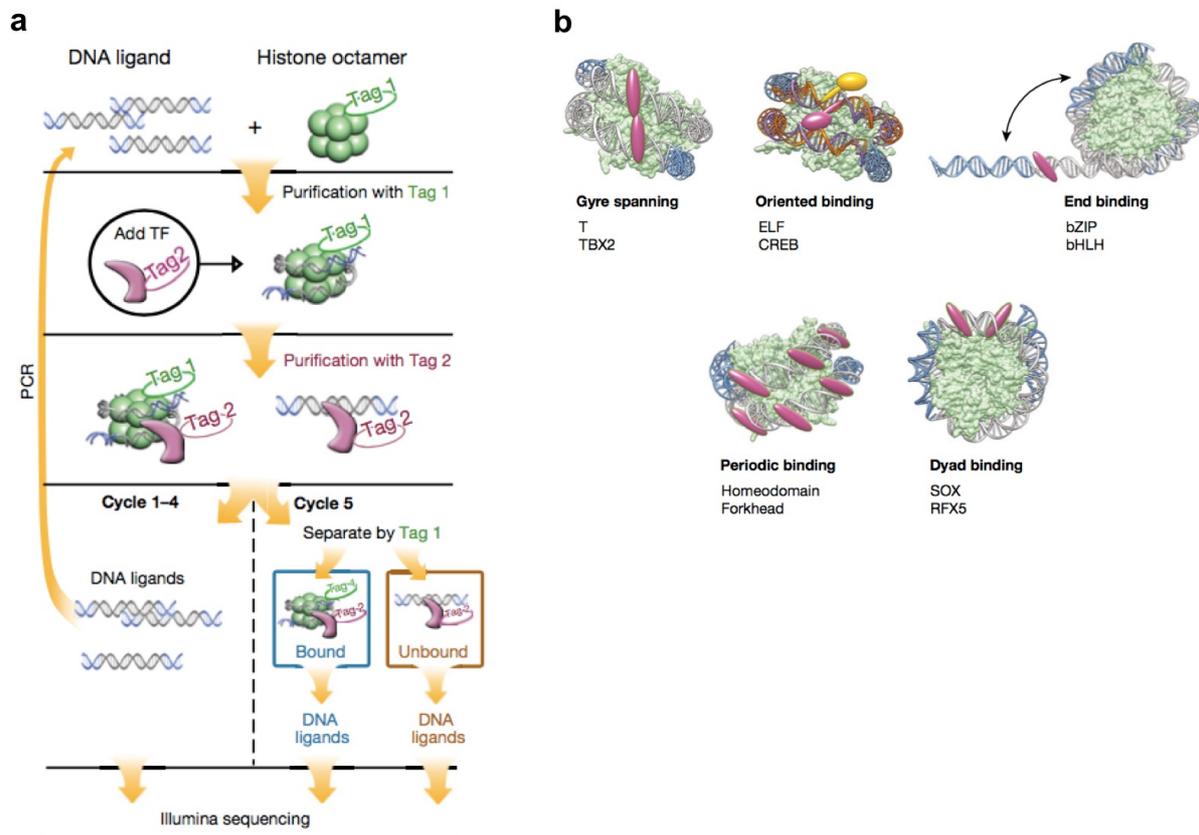


Figure 7 : La méthode NCAP-SELEXA permet d'identifier le mode de liaison des TFs pionniers au nucléosome

a. La méthode NCAP-SELEXA permet la mise en contact de ligand d'ADN contenant des adaptateurs (en bleu) avec un octamère d'histone dont l'histone H2A est étiquetée par le Tag 1 (en vert). Les ligands d'ADN liés à l'octamère sont sélectionnés par purification avec le Tag 1 et mis en interaction avec des TFs étiquetés par un Tag 2 (en rose). Les ligands d'ADN liés à l'octamère et au TF sont sélectionnés par purification avec le Tag2. Dans le dernier cycle de PCR, les ligands d'ADN liés au TF sont séparés en deux catégories selon le Tag 1: les ligands d'ADN liés à l'octamère d'histone et au TF (cadre bleu) et les ligands d'ADN uniquement liés au TF (cadre marron). Les ligands d'ADN sont ensuite séquencés.

b. Cinq modes d'interaction principaux entre TFs et nucléosome ont été identifiés. L'octamère d'histone est représenté en vert et les TFs par des ovales rose ou jaune. D'après (Zhu et al. 2018)

Les facteurs pionniers se lient à l'ADN nucléosomal via leur domaine de liaison en courtes hélices alpha, comme les domaines à doigts de zinc, les domaines bHLH et les homéodomains (Fernandez Garcia et al. 2019). Les facteurs qui se lient faiblement aux nucléosomes ou ceux qui se lient à l'ADN libre le font via des domaines de liaison non structurés, des domaines avec une grande hélice alpha ou des domaines avec des feuilletts beta (Fernandez Garcia et al. 2019). Il a également été démontré que sur la petite portion de 10 pb à laquelle est liée le facteur pionnier, il y a une compétition entre la force de liaison du facteur pionnier à l'ADN qui va être localement plus forte que la force de liaison des histones à l'ADN, ce qui va entraîner une déstabilisation du nucléosome (Zhu et al. 2018). Cependant, de rares facteurs pionniers, comme ceux à domaines T-Box (ex : TBX2), semblent stabiliser le nucléosome lors de leur liaison à l'ADN nucléosomal (Zhu et al. 2018). Le mode et la position de liaison du facteur pionnier sur l'ADN nucléosomal influencent donc la stabilité du nucléosome (Zhu et al. 2018).

Les interactions entre facteurs pionniers et nucléosomes semblent dépendre de l'affinité du facteur pour le motif de liaison (Meers, Janssens, et Henikoff 2019). Un motif de liaison peut être localisé soit sur de l'ADN empaqueté dans le nucléosome, soit sur de l'ADN libre. Le laboratoire du Dr Henikoff a observé que l'affinité du motif est anti-corrélée avec la liaison du facteur pionnier à l'ADN nucléosomal (Meers, Janssens, et Henikoff 2019). D'après leur modèle, les facteurs qui ont une faible affinité pour leurs motifs se lient à la fois sur l'ADN libre et sur l'ADN empaqueté dans le nucléosome, tandis que les facteurs qui ont une forte affinité pour leur motifs se lient uniquement sur l'ADN accessible transitoirement lors de la décompaction (Meers, Janssens, et Henikoff 2019).

Enfin, les facteurs pionniers pourraient jouer un rôle majeur dans l'ouverture de la chromatine, en particulier au niveau des enhanceurs. En effet, il a été montré dans la littérature que la liaison précoce et stable d'un facteur pionnier va contribuer au recrutement de remodeleurs de la chromatine et de TFs « classiques » au niveau d'un enhanceur, ce qui va contribuer à l'activation de l'enhancer (Zaret et Carroll 2011; Spitz et Furlong 2012; Iwafuchi-Doi et al. 2016). Les facteurs pionniers (comme SOX2) sont également nécessaires à différents processus développementaux comme la pluripotence des cellules ES, le développement

embryonnaire, la différenciation cellulaire et la reprogrammation cellulaire ((Cirillo et al. 2002; Iwafuchi-Doi et al. 2016) et lu dans (Dodonova et al. 2020)).

2) La liaison des facteurs pionniers OCT4 et SOX2 au nucléosome

Des expériences *in vivo* et *in vitro* ont montré que les facteurs pionniers OCT4 et SOX2 peuvent se lier à l'ADN nucléosomal (Soufi et al. 2015; Meers, Janssens, et Henikoff 2019). Le facteur OCT4 contient un domaine de liaison à l'ADN de type POU (*Pit-Oct-Unc*) divisé en deux parties séparées par de l'ADN linker : POU_S en N-terminal et POU_{HD} (homéodomaine) en C-terminal (Soufi et al. 2015). Le facteur SOX2 contient un domaine de liaison à l'ADN de type HMG (*Sry-related High Mobility Group*).

Plusieurs études ont permis de définir la façon dont les facteurs pionniers OCT4 et SOX2 se lient à l'ADN nucléosomal. Soufi *et al.* ont mis en évidence qu'OCT4 et SOX2 reconnaissent leur motif canonique complet (des octamères) lorsqu'ils se fixent sur l'ADN libre, alors que lorsqu'ils se fixent au niveau du nucléosome, les deux domaines de liaison de OCT4 reconnaissent chacun un demi motif canonique (4 nucléotides reconnus) tandis que SOX2 reconnaît son motif dégénéré où le sixième nucléotide, un « G », est absent (Soufi et al. 2015). Dodonova *et al.* ont utilisé une séquence d'ADN de 147 pb liée par le facteur SOX11 avec une haute affinité, pour définir la structure 3D du facteur SOX2 lié à un nucléosome (Zhu et al. 2018; Dodonova et al. 2020). Les motifs de liaison à l'ADN de SOX2 et SOX11 présentent 83% de similarité de séquence (Badis et al. 2009). Dans le contexte d'expériences réalisées *in vitro*, Dodonova *et al.* ont montré que SOX2 se lie au niveau du SHL +2 sur l'ADN nucléosomal en élargissant le petit sillon de l'ADN et en éloignant l'ADN des histones (Dodonova et al. 2020). La fixation des facteurs de la famille SOX2 et la déformation de l'ADN nucléosomal qu'elle implique vont entraîner des perturbations de la liaison de l'ADN aux histones au niveau des positions SHL -5, -6 et -7 et donc faciliter l'accès de cet ADN à d'autres facteurs non pionniers (Dodonova et al. 2020). Ils ont également observé que les facteurs de la famille SOX peuvent se lier au niveau de leur site de liaison sans détacher le deuxième tour que l'ADN fait autour du nucléosome (Dodonova et al. 2020).

Michael *et al.* ont mis au point la technique du SeEN-seq afin d'identifier toutes les interactions potentielles des facteurs OCT4 et SOX2 avec un nucléosome (Michael et al. 2020). Le SeEN-seq consiste à former *in vitro* des nucléosomes à partir d'une séquence positionnant les nucléosomes dans laquelle sont insérées des motifs consensus SOX2/OCT4 à toutes les positions possibles sur le nucléosome. Ces nucléosomes sont ensuite incubés avec les TFs OCT4 et SOX2 (Michael et al. 2020).

Les nucléosomes liés par les TFs sont séparés des non-liés sur gel, puis séquencés pour révéler la position spécifique de l'enrichissement des TFs sur le nucléosome (Michael et al. 2020). Ils ont ainsi montré que le facteur OCT4 se lie préférentiellement au niveau des sites d'entrée et de sortie de l'ADN nucléosomal, ce qui suggère que sa liaison est favorisée par le phénomène de « respiration » aux sites d'entrée/sortie de l'ADN (Michael et al. 2020). L'hétérodimère OCT4-SOX2 se lie également préférentiellement au niveau des sites d'entrée et de sortie du nucléosome, au niveau de SHL -6 ou de SHL +6, avec une affinité largement plus forte que OCT4. La liaison de OCT4-SOX2 est facilitée par l'orientation du motif (Michael et al. 2020). Lorsqu'OCT4 et SOX2 sont liés au niveau de leur site sur l'ADN nucléosomal à SHL -6, l'ADN est déformé par SOX2, et perd son interaction avec SOX2 tout en restant lié par OCT4 (Michael et al. 2020). Seul le domaine POU_S d'OCT4 interagit avec l'ADN dans le contexte de la liaison sur le nucléosome (le domaine POU_{HD} n'est impliqué dans l'interaction avec sa partie du site consensus que lors de la liaison sur l'ADN nu). Sans interagir directement avec les histones, OCT4 et SOX2 coopèrent pour détruire certaines liaisons de l'ADN avec les histones H2A et H3. Lorsqu'OCT4 et SOX2 se lient au niveau de l'ADN nucléosomal à SHL +6 l'ADN est courbé par SOX2 mais n'est pas détaché des histones, comme à SHL -6 (Michael et al. 2020). Enfin, ils ont montré *in vivo* que le domaine POU_S d'OCT4 est suffisant pour amorcer, en coopération avec SOX2, l'ouverture de la chromatine (Michael et al. 2020).

Egalement dans le but d'étudier les sites de liaison d'OCT4 sur le nucléosome, Echigoya *et al.* ont reconstitué un nucléosome à partir de 162 pb d'ADN comprenant la région distale de l'enhancer *LIN28B* humain qui comporte trois sites de liaison d'OCT4 potentiels (Echigoya et al. 2020). Grâce à ce système expérimental, ils ont mis en évidence que le site de liaison n°1 d'OCT4 se situe au niveau de SHL 7 sur l'ADN nucléosomal, le site de liaison n°2 se situe au niveau de SHL 4.5 et le site n°3 au niveau de SHL 1.5 (Echigoya et al. 2020). En combinant leur

système expérimental avec des expériences de spectrométrie de masse avec crosslinking, ils ont observé qu'OCT4 se lie préférentiellement au niveau du site d'entrée et de sortie du nucléosome (son site n°1) et plus particulièrement au niveau de la région N-terminale de l'histone H3, mais aussi en moindre mesure au niveau de petites régions N-terminales des histones H2A et H4 (correspondant aux sites n°2 et n°3 de liaison d'OCT4 sur l'ADN nucléosomal) (Echigoya et al. 2020).

Très récemment, le laboratoire du Dr Krebs a adapté la technique du *single-molecule footprinting*, précédemment développée chez la Drosophile, aux cellules mammifères (Sönmezer et al. 2021). Cette technique repose sur la méthylation par une enzyme méthyltransférase des régions non protégée de l'ADN, c'est-à-dire les régions de l'ADN qui n'interagissent pas avec des TFs ou qui ne sont pas occupées par des nucléosomes (Krebs et al. 2017). Cette technique permet donc de détecter et de quantifier *in vivo* les contacts entre les TFs, la Pol II et l'ADN et le positionnement des nucléosomes sur une seule molécule d'ADN au niveau des éléments régulateurs du génome (Krebs et al. 2017; Sönmezer et al. 2021). Sönmezer *et al.* ont combiné les techniques de *single-molecule footprinting* et de CHIP-seq en cellules ES mammifères pour détecter et quantifier les multiples événements de liaison des TFs sur des molécules d'ADN uniques (Sönmezer et al. 2021).

Les auteurs ont observé différents patterns de liaison au motif entre les différents TFs. En effet, Sönmezer *et al.* ont démontré que certains facteurs ont une capacité de liaison à l'ADN stable, probablement due à une cinétique de dissociation lente, alors que d'autres facteurs ont un temps d'interaction court avec leur cible sur l'ADN (Sönmezer et al. 2021). Ils ont ainsi pu observer que le facteur CTCF occupe une majorité de ses molécules cibles, tandis que l'hétérodimère de facteurs pionniers OCT4-SOX2 occupe qu'environ 10% de ses molécules d'ADN cibles (Sönmezer et al. 2021).

Les analyses de *single-molecule footprinting* couplées aux CHIP-seq permettent de déterminer la fréquence de co-occupation des TFs au niveau des éléments régulateurs du génome (Sönmezer et al. 2021). Ils ont montré que la liaison simultanée des TFs est indépendante de l'identité du TF et ne requiert pas de stéréotype d'organisation des motifs de liaison (Sönmezer et al. 2021). De plus, la liaison simultanée de TFs est prévalant au niveau de régions qui ont une forte densité nucléosomale (Sönmezer et al. 2021). Les auteurs proposent que la

co-occupation de plusieurs TFs augmente l'efficacité d'élimination du nucléosome et que cela confère un meilleur accès aux sites de liaison cibles. Il est probable que les remodeleurs de la chromatine soient impliqués dans ce mécanisme (Sönmezer et al. 2021).

Chapitre 3 : Les remodeleurs de la chromatine sont impliqués dans l'organisation de la chromatine et dans la régulation de la transcription

Le remodelage de la chromatine a été défini en 2004 par Flaus et Owen-Hughes comme une « transition dans la structure de la chromatine qui peut inclure des altérations de la composition des histones et des variants d'histone au sein des nucléosomes (histones H3.3, H2AZ), des modifications post-traductionnelles des histones, des méthylations de l'ADN, des modifications du contenu protéique non-histone de la chromatine et des altérations structurales des nucléosomes » (Flaus et Owen-Hughes 2004).

Dans ce chapitre, je vais me focaliser sur la fonction des remodeleurs de la chromatine ATP-dépendants. Ces facteurs sont impliqués dans certaines maladies chez l'Homme, et je présenterai l'implication du complexe BAF dans les cancers humains.

I. Caractéristiques et fonctions des remodeleurs de la chromatine

1) Les remodeleurs de chromatine ATP-dépendants

Les facteurs de remodelage de la chromatine ATP-dépendants, ou remodeleurs, font partie de complexes multi-protéiques possédant une sous-unité catalytique qui utilise l'énergie produite par l'hydrolyse de l'ATP pour altérer les liaisons histones-ADN au sein du nucléosome. L'altération des liaisons histones-ADN va modifier la position, la structure et parfois la composition de l'octamère d'histones vis à vis de l'ADN (Flaus et Owen-Hughes 2004).

On distingue quatre sous-familles parmi les remodeleurs ATP-dépendants: la sous-famille SWI/SNF, la sous-famille CHD, la sous-famille INO80 et la sous-famille ISWI. Ces quatre sous-familles sont définies en fonction des ressemblances et des divergences partagées par leur sous-unité catalytique et par leurs autres domaines (Flaus et al. 2006). Les sous-unités ATPase retrouvées dans ces quatre sous-familles sont toutes constituées de deux lobes RecA-like séparés par une séquence d'insertion plus ou moins longue qui forment le domaine ATPase-translocase (Clapier et al. 2017). Les différentes sous-familles se distinguent par la longueur

de la séquence d'insertion du domaine ATPase-translocase et par la nature de leurs domaines additionnels : les sous-familles ISWI, CHD et SWI/SNF contiennent une courte séquence d'insertion dans le domaine ATPase-translocase, tandis que la sous famille INO80 contient une longue séquence d'insertion (Clapier et al. 2017). Les sous-familles ISWI et CHD contiennent toutes deux un domaine NegC et ISWI contient en plus les domaines d'autoinhibition N-terminal (AutoN) et HSS, et CHD contient des domaines tandem CHD et des domaines de liaison à l'ADN (DBD). Les sous-familles SWI/SNF et INO80 contiennent toutes deux les domaines HSA et post-HSA et SWI/SNF contient en plus différents domaines dont un bromodomaine (Clapier et al. 2017).

Au cours de ma thèse, je me suis intéressée à plusieurs remodeleurs de la chromatine appartenant à différentes sous-familles : les remodeleurs BRG1 (ou SMARCA4, sous-famille SWI/SNF), CHD1, CHD4 et CHD8 (sous-famille CHD), EP400 (sous-famille type INO80) et SMARCA5 (sous-famille ISWI).

2) Différents mécanismes recrutent les remodeleurs sur les nucléosomes

Les mécanismes de recrutement des remodeleurs sur leur nucléosome cible sont encore mal caractérisés. Il existe au moins deux mécanismes principaux de recrutement.

Le premier mode de recrutement se fait par reconnaissance de modifications d'histone (Flaus et Owen-Hughes 2004). En effet, le bromodomaine de certains remodeleurs de la chromatine, dont la sous-famille SWI/SNF, peut reconnaître les acétylations des histones au niveau du nucléosome et ainsi favoriser le recrutement du remodeleur sur sa cible nucléosomale (lu dans (Bowman et Poirier 2015)). *In vitro*, il a été montré que la marque H3K4me3 semble favoriser le recrutement du remodeleur CHD1 humain au niveau de ses nucléosomes cibles et active le remodelage de la chromatine (lu dans (Bowman et Poirier 2015)). A l'inverse, le complexe NuRD de remodelage de la chromatine nécessite que la lysine 4 de l'histone H3 ne soit pas méthylée pour pouvoir se lier à la queue de l'histone H3 du nucléosome cible (Zegerman et al. 2002).

Le second mode de recrutement implique des interactions entre remodeleurs et TFs (Flaus et Owen-Hughes 2004). Des études protéomiques réalisées en cellules ES ont démontré que des remodeleurs de la chromatine ATP-dépendants interagissent de façon privilégiée avec des TFs associés à la pluripotence (van den Berg et al. 2010). Plus précisément, les remodeleurs CHD4, EP400 et BRG1 interagissent avec les TFs OCT4 et ESRRB (van den Berg et al. 2010).

3) Les principaux mécanismes d'action des remodeleurs

Le domaine ATPase de la sous-unité catalytique et les activités de translocation de l'ADN sont des caractéristiques des remodeleurs ATP-dépendants (Clapier et al. 2017). Dans leur revue de 2017, Clapier *et al.* ont comparé le domaine ATPase des remodeleurs à un « moteur qui utilise la translocation de l'ADN pour rompre les contacts histone-ADN et propulser l'ADN le long de la surface des histones » (Clapier et al. 2017). Ces auteurs proposent que les sous-unités catalytiques des différents remodeleurs utilisent un mécanisme commun pour réaliser des fonctions diverses.

Les remodeleurs de la chromatine peuvent effectuer trois fonctions principales (avec des préférences pour chaque famille de remodeleurs) :

- l'assemblage du nucléosome et l'organisation nucléosomale,
- la régulation de l'accès à la chromatine en déplaçant ou en enlevant des nucléosomes,
- l'édition de nucléosomes (régulation de la composition).

Les remodeleurs qui ont une fonction d'assemblage de la chromatine et d'organisation nucléosomale interviennent essentiellement après la réplication de l'ADN (Clapier et al. 2017). En effet, juste après la synthèse d'un nouveau brin d'ADN, les nucléosomes doivent être reformés pour éviter une dérégulation de l'expression des gènes. Ainsi, des complexes d'histones, sous forme de tétramères H3-H4 et de dimères H2A-H2B, sont libérés par les chaperons d'histone au niveau du brin d'ADN néo-synthétisé (Gurard-Levin, Quivy, et Almouzni 2014). Les remodeleurs des sous-familles ISWI et CHD vont avoir deux actions : (1) la maturation des pré-nucléosomes en nucléosomes canoniques contenant un octamère d'histones et 147 pb d'ADN enroulé autour (Fei et al. 2015), et (2) l'espacement régulier des nucléosomes (**Figure 8**) (lu dans (Clapier et al. 2017).

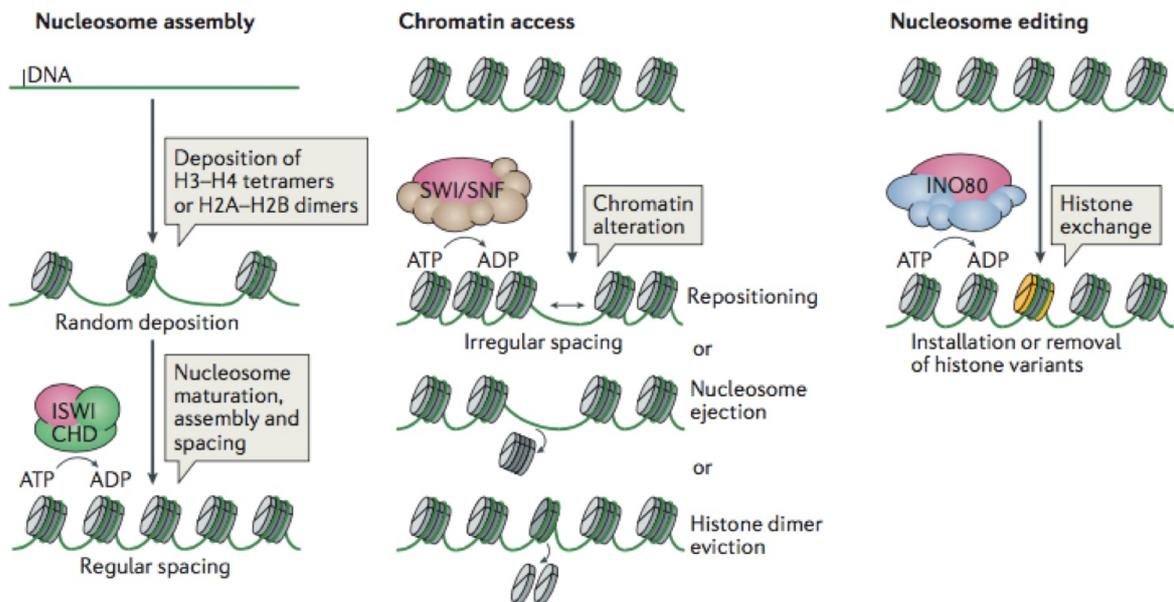


Figure 8 : Les trois fonctions principales des remodeleurs de la chromatine ATP-dépendants

Les remodeleurs de la chromatine ATP-dépendants ont trois fonctions principales : l'assemblage des nucléosomes, l'accès à la chromatine et l'édition de nucléosomes. La sous-unité catalytique de chacun des complexes de remodelage de la chromatine est représentée en rose, les autres sous-unités sont en vert pour ISWI et CHD, marron pour SWI/SNF et bleu pour INO80. Les dimères d'histones formant les nucléosomes sont représentés par les cylindres gris et l'ADN enroulé autour par la ligne verte. Le nucléosome représenté en jaune est un nucléosome contenant un variant d'histone comme H2A.Z. D'après (Clapier et al. 2017).

Afin de réguler l'expression des gènes, différents acteurs doivent avoir accès aux séquences régulatrices empaquetées dans les nucléosomes. Les remodeleurs, principalement de la sous-famille SWI/SNF, ont une fonction de régulateur d'accès à la chromatine (Clapier et al. 2017). Ces remodeleurs peuvent avoir trois modes d'action différents (Figure 8). Premièrement, ils peuvent repositionner un nucléosome en faisant glisser l'octamère le long de l'ADN, ce qui crée des espacements de taille anormale entre les nucléosomes. Deuxièmement, ils peuvent expulser un nucléosome entier, ou troisièmement, ils peuvent expulser seulement une partie du nucléosome, comme les dimères H2A-H2B. Les différents acteurs de la transcription, comme les activateurs, les répresseurs et les TFs, peuvent donc se lier directement sur de l'ADN libéré de tout nucléosome, ou sur de l'ADN encore enroulé autour du nucléosome mais accessible au niveau de leurs sites de liaison situés préférentiellement au niveau des enhancers et des promoteurs (Clapier et al. 2017).

Enfin, certains remodeleurs, ceux appartenant à la sous-famille INO80, peuvent faire de l'édition de nucléosome (**Figure 8**) (Clapier et al. 2017). Indépendamment de la réplication, ces remodeleurs peuvent échanger une ou plusieurs histones canoniques (H2A, H2B ou H3) par une ou plusieurs histones non-canoniques ou par des variants d'histone (H2AZ, H3.3, cenH3...) (Clapier et al. 2017).

II. Les remodeleurs de la chromatine se lient aux nucléosomes au niveau de différentes régions

1) Les remodeleurs interagissent au niveau du patch acide et de la queue H4 du nucléosome

Deux régions précises du nucléosome, le patch acide et la queue de l'histone H4, sont impliquées dans les interactions nucléosome-nucléosomes, dans les interactions remodeleurs-nucléosomes et dans la régulation de la fonction des remodeleurs (Markert et Luger 2021).

Le patch acide est une région chargée négativement située au niveau des dimères H2A-H2B (K. Luger et al. 1997). Il est constitué de six acides aminés de l'histone H2A (les glutamates en position 56, 61, 64, 91 et 92 et l'aspartate en position 90 chez l'Homme) et deux acides aminés de l'histone H2B (les glutamates en position 105 et 113 chez l'Homme) ((K. Luger et al. 1997) et lu dans (Zhou, Gaullier, et Luger 2019)). La queue de l'histone H4 se trouve au niveau des régions SLH 1 et SHL 2 de l'ADN nucléosomal et peut supporter de nombreuses modifications post-traductionnelles (Markert et Luger 2021).

Concernant le rôle du patch acide et de la queue de l'histone H4 dans les interactions entre nucléosomes, il a été montré que des mutations dans les acides aminés composant le patch acide ont des conséquences sur la compaction de la chromatine (Gordon, Luger, et Hansen 2005). De plus, dans les structures cristallisées, la queue de l'histone H4 d'un nucléosome peut interagir avec le patch acide du nucléosome voisin (K. Luger et al. 1997). Dans un contexte physiologique, les interactions entre nucléosomes sont stabilisées par la queue de l'histone

H4, tandis que le dépôt de la marque épigénétique H4K16ac empêche la compaction de la chromatine (Shogren-Knaak et al. 2006).

Concernant le rôle du patch acide et de la queue de l'histone H4 dans les interactions entre remodeleurs et nucléosomes et dans la régulation de la fonction des remodeleurs, il a été mis en évidence que des remodeleurs peuvent se lier sur le patch acide ou sur la queue de l'histone H4 au niveau des nucléosomes canoniques. Certains remodeleurs, comme CHD4 et BRG1, dépendent de la présence de la charge négative du patch acide pour leur activité (Dann et al. 2017). D'autres remodeleurs comme CHD1 et SMARCA5 subissent une diminution de leur activité de remodelage de la chromatine en absence de la queue de l'histone H4 (Dang, Kagalwala, et Bartholomew 2006; Ferreira, Flaus, et Owen-Hughes 2007; Mueller-Planitz et al. 2013; Racki et al. 2014), tandis que l'absence des queues des autres histones n'altère pas l'activité de ces remodeleurs. Ces résultats, ainsi que l'emplacement de la queue de l'histone H4 au niveau d'une région où de nombreux remodeleurs se lient, c'est-à-dire au niveau de SHL 2, suggèrent que la queue de l'histone H4 pourrait être importante pour le mécanisme de recrutement des remodeleurs (Dang, Kagalwala, et Bartholomew 2006; Ferreira, Flaus, et Owen-Hughes 2007).

Enfin, le patch acide et la queue de l'histone H4 pourrait avoir une fonction commune sur les remodeleurs de la chromatine. En effet, en absence de liaison à un nucléosome cible, certains remodeleurs de la chromatine peuvent être dans un état d'auto-inhibition ((Clapier et Cairns 2012) et lu dans (Markert et Luger 2021). L'interaction de domaines des remodeleurs avec le patch acide ou la queue de l'histone H4 d'un nucléosome va permettre à ces remodeleurs de sortir de cet état d'auto-inhibition et d'exercer leur fonction de remodelage. Par exemple, le domaine ATPase du remodeleur ISWI va interagir avec la queue de l'histone H4, ce qui va permettre au domaine catalytique de s'éloigner du domaine d'auto-inhibition (He et al. 2020), et les deux domaines d'auto-inhibition du remodeleur SMARCA5 peuvent interagir directement avec le patch acide (Gamarra et al. 2018).

Cependant, les remodeleurs de la chromatine qui interagissent avec les nucléosomes reconnaissent aussi d'autres régions en dehors du patch acide et de la queue de l'histone H4.

2) Les remodeleurs se lient au niveau du SHL 2 ou SHL 6 sur l'ADN nucléosomal

Dans la plupart des interactions entre remodeleurs et nucléosomes, le domaine catalytique des remodeleurs entre en contact avec l'ADN du nucléosome au niveau de SHL +/-2 (Markert et Luger 2021). C'est notamment le cas pour les remodeleurs appartenant aux sous-familles SWI/SNF (comme SNF2 chez la levure et BRG1 chez les mammifères), CHD (comme CHD4, CHD1 et CHD8) et ISWI (comme SMARCA5). Par exemple, grâce à des expériences de cryo-microscopie électronique chez la levure permettant d'étudier la structure du remodeleur CHD1 lié au nucléosome, on sait que le domaine ATPase de CHD1 se lie à l'ADN nucléosomal au niveau SHL +2, ainsi qu'à la queue de l'histone H4 (Farnung et al. 2017; Chittori et al. 2019). CHD1 peut également se lier au niveau de SHL +1 et de SHL -7 sur l'ADN nucléosomal (Farnung et al. 2017; Chittori et al. 2019).

Les domaines ATPase des remodeleurs de la sous-famille INO80 (comme INO80 et EP400) font exception car ils interagissent préférentiellement avec l'ADN nucléosomal au niveau du SHL +/-6 (Ayala et al. 2018; Markert et Luger 2021).

III. Mécanisme de translocation de l'ADN

Une fois qu'un remodeleur est lié au nucléosome, il peut effectuer son action de translocation de l'ADN. Je vais présenter le mécanisme général de translocation de l'ADN par les remodeleurs de la sous-famille SWI/SNF pouvant conduire à l'éjection d'un nucléosome. Clapier *et al.* ont proposé deux modèles dans lesquels la translocation de l'ADN peut entraîner l'éjection d'un nucléosome (**Figure 9**) (Clapier et al. 2017). Dans le premier modèle, le remodeleur de la sous-famille SWI/SNF va se lier au nucléosome au niveau de son site de liaison aux histones via son domaine ATPase-translocase et va transloquer 1 à 2 pb d'ADN le long du nucléosome (**Figure 9a**) (Clapier et al. 2017). En cas d'activité ATPase et d'une combinaison des domaines ATPase-translocase faibles ou moyennes, le déplacement de l'ADN est faible et les contraintes découlant de ce déplacement sur le domaine de translocation sont faibles et peuvent être résolues par glissement continu de plusieurs pb d'ADN (**Figure 9a**

haut). Par contre, en cas d'activité ATPase et d'une combinaison ATPase-translocase forte, la forte translocation de l'ADN entraîne de fortes contraintes sur le domaine de translocation qui sont résolues par l'éjection de tout ou partie des histones de l'octamère (**Figure 9a bas**) (Clapier et al. 2017). Dans le deuxième modèle, le glissement continu de quelques pb de l'ADN le long de l'octamère d'histone en cas de tensions d'ADN faibles à intermédiaires entraîne le rapprochement du nucléosome suivant, puis la collision avec ce nucléosome, ce qui conduit au décollement de l'ADN du nucléosome adjacent et à l'élimination de tout ou partie des histones du nucléosome adjacent (**Figure 9b**) (Clapier et al. 2017).

Récemment, il a été montré que les remodeleurs qui se lient à SHL 2 (comme ceux appartenant aux sous-familles CHD et SWI/SNF) peuvent tirer une pb d'ADN supplémentaire dans cette région, ce qui favorise une translocation efficace de l'ADN (Winger et al. 2018). L'ADN à transloquer par le remodeleur est donc stocké au niveau du domaine SHL2, qui forme une sorte de « puits » pour accueillir cet ADN (Markert et Luger 2021). L'ADN poussé vers la dyade et l'hydrolyse de l'ATP permettent la suite de la translocation d'ADN (lu dans (Markert et Luger 2021)).

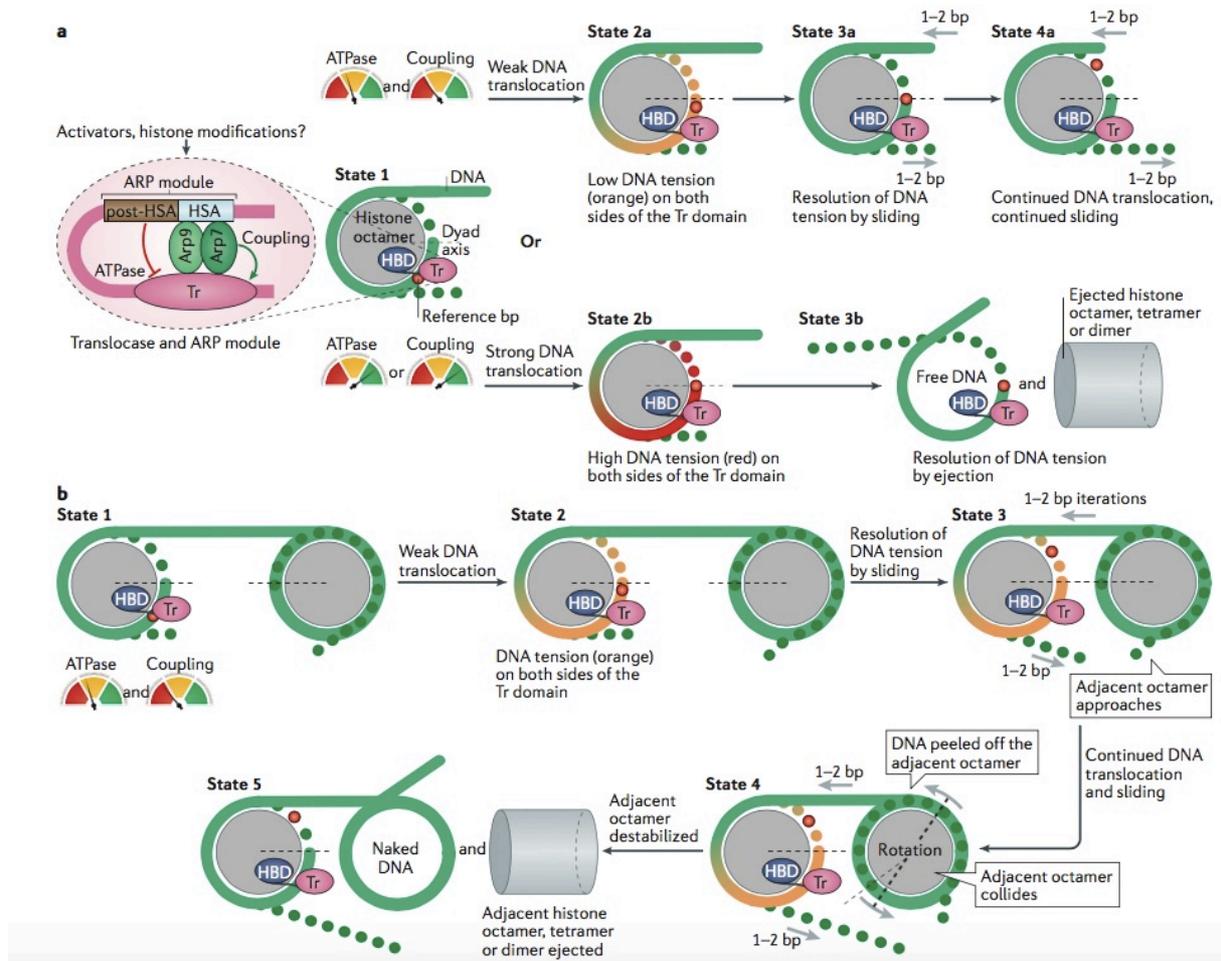


Figure 9 : Mécanisme de translocation de l'ADN le long d'un nucléosome par un remodeleur de la sous-famille SWI/SNF pouvant conduire à l'éjection d'un nucléosome

Les remodeleurs de la sous-famille sont constitués d'un module ARP, d'un module ATPase et d'un module translocase. Le niveau d'activité de l'ATPase et le niveau de couplage des domaines ATPase-translocase (Tr, en rose) sont régulés par le module ARP.

Dans les deux modèles (a. et b.), le domaine Tr interagit avec le domaine de liaison aux histones (HBD, en bleu) sur le nucléosome. Les différentes pb d'ADN impliquées dans la translocation sont représentées par des points vert et la pb de référence est représenté par un point rouge. On peut suivre la translocation de ce point rouge au cours des différentes étapes. Les cadrans rouges-orange-verts représentent le niveau de l'activité ATPase et le niveau de couplage des domaines ATPase-translocase (rouge pour faible, orange pour intermédiaire et vert pour fort). Les tensions exercées par la translocation sur l'ADN peuvent être faibles (représentées en orange sur l'ADN) ou bien forte (représentées en rouge). Dans le premier modèle (a.), les tensions de l'ADN du nucléosome concernés sont résolues par glissement si elles sont faibles ou bien par éjection d'histones si elles sont fortes. Dans le deuxième modèle (b.), les faibles tensions de l'ADN du nucléosome concernés tendent à éliminer tout ou partie des histones du nucléosome voisin dont le nucléosome concerné s'est rapporté avec les glissements continus. D'après (Clapier et al. 2017).

IV. La fonction des remodeleurs à l'échelle du génome

1) Présentation des remodeleurs étudiés pendant ma thèse

Les remodeleurs CHD4 (*Chromodomain Helicase DNA Binding Protein 4*), aussi appelé Mi2beta, CHD1 (*Chromodomain Helicase DNA Binding Protein 1*) et CHD8 (*Chromodomain Helicase DNA Binding Protein 8*) appartiennent à la sous-famille CHD.

CHD4 est le composant principal du complexe NuRD, qui a à la fois une activité de remodelage de la chromatine et une activité de désacétylation des histones. CHD4, qui se lie au niveau des enhancers et des promoteurs, est requis pour assurer l'auto-renouvellement des cellules ES murines (H. Zhao et al. 2017).

Le remodeleur BRG1 (*Brahma-related gene 1*), aussi appelé SMARCA4, appartient à la sous-famille SWI/SNF. BRG1 est la sous-unité catalytique principale du complexe BAF (SWI/SNIF). BRG1 se lie au niveau des promoteurs et des enhancers en cellules ES avec une préférence pour les enhancers (de Dieuleveult et al. 2016).

EP400 (*E1A-binding protein p400*) est un remodeleur typique de ceux appartenant à la sous-famille INO80. EP400 est un composant du complexe histone acétyltransférase Tip60/Ep400. Ce complexe est impliqué dans l'activation transcriptionnelle de gènes par remodelage et acétylation des histones. EP400 est présent au niveau des enhancers et des promoteurs, mais contrairement à BRG1, EP400 se lie préférentiellement au niveau des promoteurs actifs (de Dieuleveult et al. 2016).

2) L'organisation nucléosome au niveau des promoteurs mammifères

Certaines méthodes expérimentales permettent de caractériser l'organisation et l'architecture nucléosomale du génome. Il s'agit notamment du MNase-seq (présenté dans le chapitre 1) et du DNase-seq. Le DNase-seq est une technique basée sur la digestion par la DNase I (présentée dans le chapitre 1) des régions d'ADN non-engagées de façon permanente dans des interactions avec des protéines.

Au niveau des promoteurs des cellules ES de souris, une zone dépourvue en nucléosome, la NFR, a été définie. La taille des NFR peut être très variable en fonction des promoteurs : d'une dizaine à 100 pb autour du TSS (NFR courtes) ou de 150 pb à plus d'un kb autour du TSS (NFR longues) (de Dieuleveult et al. 2016).

Concernant les promoteurs à NFR courtes, ils sont flanqués par deux nucléosomes de part et d'autre de la NFR (**Figure 10a**) (de Dieuleveult et al. 2016). Le nucléosome en amont de la NFR est appelé nucléosome -1 et le nucléosome en aval de la NFR est appelé nucléosome +1 (**Figure 10**). Ces deux nucléosomes, et surtout celui en 5' (le +1), sont hypersensibles à la DNase I. Ces nucléosomes sont donc dynamiques : parfois le nucléosome est positionné et l'ADN enroulé autour de ce nucléosome est détecté par la technique du MNase-seq et parfois le nucléosome n'est pas positionné et l'ADN est donc digéré par la DNase I (de Dieuleveult et al. 2016).

Au niveau des promoteurs à NFR longues, une petite région en 5' de la NFR incluant le TSS est hypersensible à la DNase I, tandis que le reste de la NFR localisée en aval du TSS est résistante à la DNase I (de Dieuleveult et al. 2016). Cela signifie que l'ADN est protégé par différentes interactions. Des variants d'histones, des îlots CpG et des modifications épigénétiques des histones ont été détectées à ce niveau (**Figure 10b**) (de Dieuleveult et al. 2016). Cependant, l'ADN au niveau de cette longue NFR ne semble pas être empaqueté autour de nucléosomes canoniques car cet ADN est toujours sensible à la MNase (**Figure 10b**) (de Dieuleveult et al. 2016).

L'organisation nucléosomale dans les cellules ES de souris n'a pas été bien caractérisée au niveau des enhancers. Il s'agit d'un des objectifs de ma thèse.

3) Distribution génomique et fonction des remodeleurs dans la régulation de la transcription

Mon laboratoire de thèse a précédemment réalisé une étude pour caractériser la liaison sur le génome des cellules ES des remodeleurs BRG1, EP400, CHD1, CHD4 et CHD8 grâce à des expériences de CHIP-seq dans des cellules ES de souris (de Dieuleveult et al. 2016).

Ils ont mis en évidence qu'au niveau du promoteur d'un gène activement transcrit, là où les acteurs de la transcription sont liés comme la Pol II et le TBP, on retrouve souvent la présence de remodeleurs de la chromatine (de Dieuleveult et al. 2016). Quel que soit le niveau d'expression des gènes, les remodeleurs CHD4 et BRG1 lient des positions similaires au niveau des nucléosomes, tandis que CHD1 et EP400 se lient préférentiellement au niveau de gènes avec un fort niveau d'expression (de Dieuleveult et al. 2016).

Au niveau des promoteurs à NFR courtes, les remodeleurs EP400 et CHD4 se lient fortement au niveau du nucléosome -1 (**Figure 10a**). Un pic d'hypersensibilité à la DNase I a été détecté au niveau de ce nucléosome -1, ce qui suggère que CHD4 et EP400 pourraient participer à le dissocier et à le reformer fréquemment (de Dieuleveult et al. 2016). Les remodeleurs EP400, CHD4, CHD1, CHD8 et BRG1 se lient également au niveau du nucléosome +1 (**Figure 10a**) (de Dieuleveult et al. 2016). On trouve au niveau de ce nucléosome +1 des variants d'histone et les marques d'histone H3K4me3 et H3K27ac (**Figure 10a**). La présence de nombreux remodeleurs au niveau de ce nucléosome +1 pourraient faciliter le passage de la Pol II à travers ce nucléosome et donc faciliter le démarrage de la transcription.

Au niveau des promoteurs à NFR longues, les remodeleurs EP400, CHD4, CHD8 et BRG1 sont tous localisés préférentiellement au niveau du nucléosome -1 (**Figure 10b**). Seul CHD1 est localisé au niveau du nucléosome +1 (**Figure 10b**) (de Dieuleveult et al. 2016). Au niveau de ces NFR longues, les variants d'histone H3.3, H2AZ et les marques H3K4me3 et H3K27ac sont majoritairement détectés dans la région de chromatine non-canonique en aval du TSS (**Figure 10b**) (de Dieuleveult et al. 2016).

Cette étude a permis de mettre à jour trois principaux modes de régulation des gènes par les remodeleurs en cellules ES de souris (de Dieuleveult et al. 2016).

Au niveau des gènes actifs avec des promoteurs à NFR courtes entourées de nucléosomes flanquant, CHD4 et EP400 ont un rôle d'activateur de la transcription, tandis que BRG1 a un modeste rôle répresseur (**Figure 10a**). Malgré leur liaison sur les nucléosomes flanquant la NFR, la fonction des remodeleurs CHD1 et CHD8 est moins claire car ils régulent un petit nombre de gènes (de Dieuleveult et al. 2016).

Au niveau des gènes actifs avec des promoteurs à NRF longues, BRG1 se lie au niveau du nucléosome -1 et agit comme un activateur de la transcription (**Figure 10b**) (de Dieuleveult et al. 2016).

Enfin, au niveau des promoteurs de gènes bivalents, caractérisés par une longue NRF enrichie en marque H3K4me3, en variants d'histone et en îlots CpG et par la présence de la marque H3K27me3 au nucléosome +1 déposée par les Polycomb, BRG1 agit comme un activateur de la transcription et les remodeleurs EP400 et CHD4 comme des répresseurs (**Figure 10c**) (de Dieuleveult et al. 2016).

On peut donc conclure de cette étude qu'un remodeleur peut avoir un effet d'activateur de la transcription sur une certaine catégorie de gènes et un effet répresseur sur d'autres. De plus, au sein de la même catégorie de gènes, un remodeleur qui a une fonction d'activateur de la transcription peut être empêché par un remodeleur avec une fonction répresseur. Les remodeleurs coopèrent donc pour réguler l'expression des gènes (de Dieuleveult et al. 2016).

Le laboratoire du Dr Schübeler a récemment démontré que la régulation de l'espacement des nucléosomes en cellules ES dépend de SMARCA5 (aussi appelé SNF2H) (Barisic et al. 2019). Ils ont également montré que le remodeleur SMARCA5 régule la liaison du facteur CTCF sur ces sites CTCF (Barisic et al. 2019). Le remodeleur SMARCA5 est l'une des deux sous-unités catalytiques du complexe ISWI.

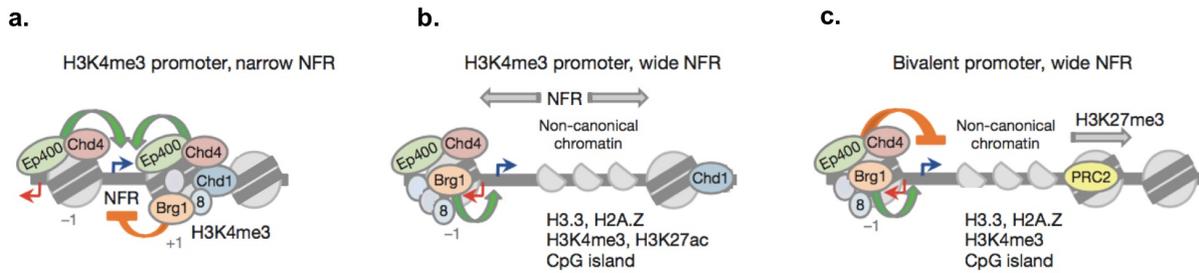


Figure 10 : Les remodeleurs de la chromatine régulent l'expression de différentes catégories de gènes dans les cellules ES de souris.

Cette figure présente un modèle de régulation des régions promotrices de trois catégories de gènes : les gènes ayant des promoteurs enrichis en H3K4me3 et une NFR courte (**a.**), les gènes ayant des promoteurs enrichis en H3K4me3 et une NFR longue (**b.**) et les gènes bivalents ayant des NLF longues (**c.**). Les nucléosomes canoniques sont symbolisés par des cercles gris clair, les nucléosomes non-canoniques par des demi-cercles gris clair, le TSS par une flèche bleue horizontale et le site d'initiation de la transcription antisens par une flèche rouge horizontale. La liaison des remodeleurs de la chromatine est représentée par des cercles de couleurs (Ep400 en vert, Chd4 en rouge, Brg1 en orange, Chd1 en bleu et Chd8 en gris-mauve). Les flèches vertes indiquent une action activatrice du remodeleur, tandis que les flèches orange indiquent une action répressive. L'enrichissement en marques épigénétiques et en variants d'histone ont été écrits à l'endroit de leur détection. Adaptée d'après (de Dieuleveult et al. 2016).

V. Les remodeleurs interagissent avec les facteurs pionniers pour l'ouverture de la chromatine

En 2017, King et Klose ont étudié comment le facteur pionnier OCT4 se lie *in vivo* au niveau de ses séquences cibles en cellules ES de souris. Ils ont confirmé que le facteur OCT4 se lie préférentiellement au niveau des enhancers, où il permet le maintien de la chromatine accessible et la liaison de TFs associés à la pluripotence comme SOX2 et NANOG (King et Klose 2017). Grâce à une combinaison d'expériences de CHIP-seq et d'ATAC-seq, les auteurs ont également démontré que le remodeleur BRG1 se lie préférentiellement aux régions régulatrices du génome dépendantes d'OCT4 pour leur accessibilité (King et Klose 2017). BRG1 et OCT4 sont tous deux nécessaires au maintien de l'accessibilité de la chromatine au niveau des régions cibles d'OCT4 et plus particulièrement, c'est le facteur pionnier OCT4 qui permet le recrutement du remodeleur BRG1 dans les cellules ES afin de favoriser l'accessibilité de la

chromatine et la liaison d'autres TFs sur leurs sites cibles (King et Klose 2017). Enfin, les résultats de King & Klose suggèrent que les facteurs associés à la pluripotence OCT4, SOX2, NANOG peuvent reconnaître leurs sites de liaison en absence de BRG1, ce qui est cohérent avec la reconnaissance partielle de leur motif (Soufi et al. 2015), mais ils ne sont pas capable d'établir une liaison solide ni d'exercer leurs fonctions sans BRG1 au niveau des régions régulatrices du génome (King et Klose 2017).

Très récemment, *Iurlaro et al.* se sont intéressés à la dynamique de l'activité du remodeleur BRG1 au niveau des régions régulatrices du génome, à savoir si son action était continuellement ou ponctuellement requise pour l'accessibilité à la chromatine en cellules ES (Iurlaro et al. 2021). Leurs résultats montrent que l'activité du remodeleur BRG1 est continuellement requise au niveau de ces cibles pour permettre l'accès à la chromatine et pour le maintien de programmes transcriptionnels spécifiques en cellules ES (Iurlaro et al. 2021). Afin de savoir si la liaison des TFs à leurs cibles nécessite également une activité continue de BRG1, ils ont étudié les changements d'accessibilité aux sites liés par les TFs en absence de BRG1 (Iurlaro et al. 2021). Ils ont ainsi trouvé que l'accessibilité à leurs sites de liaison et la liaison des facteurs dépendants de BRG1 pour leur liaison, comme l'hétérodimère OCT4/SOX2, REST et ESRRB, sont fortement réduites en absence d'une activité constante de BRG1 (Iurlaro et al. 2021). Enfin, les auteurs ont conclu que la perte d'accessibilité dans une région donnée en absence de BRG1 se traduit par une diminution de la liaison des TFs et une augmentation des nucléosomes dans cette région, et cette situation est rapidement réversible en présence de BRG1 (Iurlaro et al. 2021).

Chez la levure, certains facteurs sont également capables de se lier directement sur les nucléosomes fragiles.

Les facteurs GRF, pour *General Regulation Factor*, de la levure ont des propriétés de TF pionnier. Ces facteurs pionniers sont préférentiellement liés au niveau des promoteurs avec de larges régions NFR contenant des nucléosome fragiles (Kubik et al. 2015). Kubik *et al.* ont montré que ces facteurs pionniers sont impliqués dans l'espacement entre les nucléosomes -1 et +1 favorisant ainsi la formation de nucléosomes fragiles (Kubik et al. 2015). Au niveau des promoteurs contenant un nucléosome fragile, les auteurs ont détecté des motifs G/C-rich et

polyA en plus du motif de liaison des GRF, qui vont permettre le recrutement du remodelleur RSC (Kubik et al. 2015). Les auteurs proposent que les GRF pourraient aider au recrutement du remodelleur RSC en favorisant son accès aux deux motifs polyA et *G/C-rich* et/ou en interagissant directement avec RSC (Kubik et al. 2015). La liaison de RSC, un homologue du complexe SWI/SNF mammifère, va provoquer le déplacement de nucléosome par *sliding* et augmenter la distance entre les nucléosomes canoniques -1 et +1, et favoriser la formation d'un nucléosome fragile (Kubik et al. 2015).

Dans une autre étude chez la levure, Brahma et Henikoff ont vérifié que les facteurs GRF et le remodelleur RSC se lient sur des nucléosomes fragiles au niveau des promoteurs à NFR larges (Brahma et Henikoff 2019). Ils ont proposé un modèle dynamique de coopération entre RSC et les GRF sur les nucléosomes fragiles au niveau des promoteurs à larges NFR (**Figure 11**) (Brahma et Henikoff 2019). Grâce à ses deux motifs cibles (polyA et *C/G-rich*), RSC se fixe sur un nucléosome situé au niveau d'une large NFR et va partiellement remodeler ce nucléosome en utilisant l'énergie de l'hydrolyse de l'ATP pour rendre une partie de l'ADN nucléosomal accessible (**Figure 11 : étapes 1 et 2**) (Brahma et Henikoff 2019). Les GRF vont ensuite plus facilement venir se lier au niveau de leurs motifs cibles situés sur l'ADN nucléosomal partiellement accessible grâce à RSC, et la liaison conjointe des GRF et de RSC va aboutir à l'éjection totale du nucléosome et à la stabilisation de la liaison des GRF à ses motifs sur l'ADN désormais libre (**Figure 11 : étapes 3 et 4**) (Brahma et Henikoff 2019). De façon similaire à ce qui a été démontré chez les mammifères avec la coopération du facteur pionnier OCT4 et du remodelleur BRG1 pour l'accessibilité à la chromatine au niveau des éléments régulateurs, les GRF et le remodelleur RSC coopèrent également chez la levure pour éliminer les nucléosomes fragiles des régions promotrices.

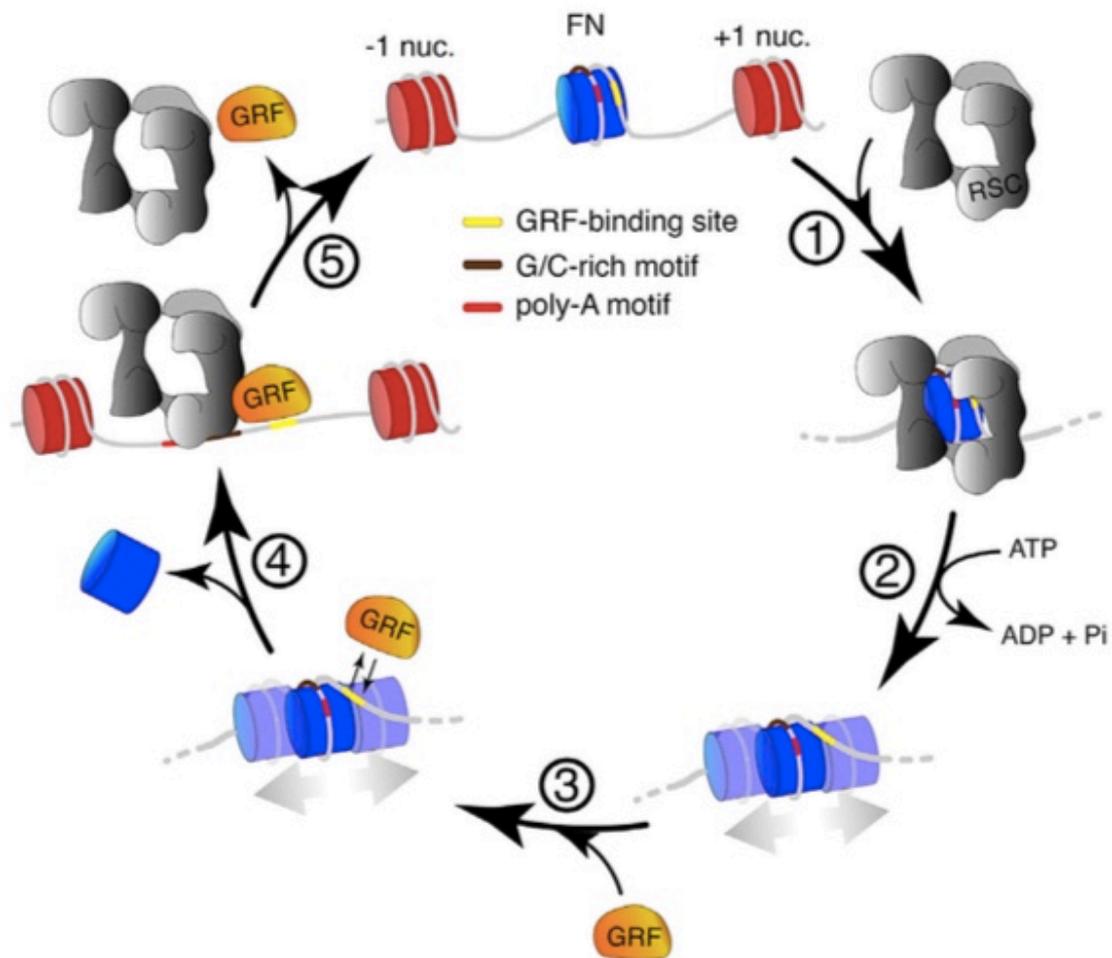


Figure 11 : Modèle dynamique de coopération entre le remodeleur RSC et les facteurs pionniers GRF pour le remodelage des nucléosomes fragiles chez la levure

Les nucléosomes canoniques sont représentés par des cylindres rouge, le nucléosome fragile par un cylindre bleu, le remodeleur RSC par la forme grise et les GRF par l'ovale orange. Les rectangles jaune, marron et rouge sur l'ADN nucléosomal et libre correspondent respectivement aux sites de liaison pour les GRF, aux motifs *G/C-rich* et aux motifs polyA.

Dans l'étape 1, RSC va se fixer sur le nucléosome fragile au niveau de ses motifs cibles, puis, grâce à l'énergie fournie par l'hydrolyse de l'ATP, RSC va exercer sa fonction de remodelage et va exposer les sites de liaison des GRF dans l'étape 2. Les facteurs pionniers GRF vont venir se fixer à leurs sites cibles dans l'étape 3. L'action conjointe de RSC et GRF va expulser les histones. RSC et/ou GRF peuvent être lié à l'ADN libre dans l'étape 4. Puis lorsqu'ils se détachent, un nucléosome fragile peut se reformer dans l'étape 5. Le remodeleur RSC est présent des étapes 1 à 4 et les GRF sont présents des étapes 3 à 4. D'après (Brahma et Henikoff 2019).

VI. Le complexe de remodelage BAF impliqué dans les cancers humains.

1) Évolution et composition des différents complexes BAF

a) Évolution du complexe BAF

Le complexe BAF ou SWI/SNF est un complexe qui a évolué au cours des millénaires (Kadoch et Crabtree 2015). Le premier complexe BAF contenant 9 sous-unités a été identifié chez la levure (lu dans (Kadoch et Crabtree 2015)). Chez les drosophiles, le complexe SWI/SNF a évolué en complexe BAP contenant quelques changements majeurs : le complexe BAP a perdu quatre sous-unités du complexe SWI/SNF et en a gagné trois nouvelles (Kadoch et Crabtree 2015). L'acquisition de la pluricellularité et de l'histone H1, ainsi que la nécessité de réprimer des gènes via le système Polycomb seraient à l'origine de ces modifications (Kadoch et Crabtree 2015). Chez les vertébrés, le complexe BAP a évolué en complexe BAF contenant neuf nouvelles sous-unités essentielles sans homologues chez les levures et drosophiles probablement dû à la diversité des fonctions, à la plus grande taille et à la méthylation du génome (lu dans (Kadoch et Crabtree 2015)). Enfin, chez les mammifères, un nouveau complexe BAF, le complexe nBAF, a été identifié (**Figure 12a**). Ce complexe contient quatre nouvelles sous-unités spécifiques des neurones allant probablement de pair avec l'apparition d'un système nerveux complexe (Lessard et al. 2007; Staahl et al. 2013)

Malgré son évolution, le complexe BAF a conservé au cours du temps la structure de la sous-unité catalytique (SWI2 ou SNF2 chez la levure, BRM chez la drosophile, BRG1 (ou SMARCA4) et BRM (SMARCA5) chez les vertébrés et les mammifères) et la fonction de remodelage (Kadoch et Crabtree 2015).

b) Les différents complexes BAF chez les mammifères

Le complexe BAF chez les mammifères, également appelé mSWI/SNF, est formé par l'expression combinatoire de 29 gènes (lu dans (Michel et al. 2018)). Le complexe BAF existe sous trois formes finales : le complexe BAF canonique (cBAF) (**Figure 12a**), le complexe polybromo-associés BAF (PBAF) et le complexe BAF non-canonique (ncBAF) (lu dans (Michel et al. 2018)). Ces trois complexes se distinguent par l'incorporation de différentes sous-unités (Michel et al. 2018) :

- ARID1A/ARID1B et DPF2 dans le complexe cBAF (**Figure 12a**),
- PBRM1, ARID2 et BRD7 dans le complexe PBAF,
- GLTSCR1/GLTSCR1L et BRD9 dans le complexe ncBAF.

Une cartographie par CHIP-seq des différents complexes BAF mammifères a permis de mettre en évidence leur localisation sur le génome (Michel et al. 2018). Les trois complexes ont une distribution comprenant de nombreux sites génomiques partagés, et des sites d'enrichissement préférentiels. En effet, le complexe cBAF est majoritairement présent au niveau des enhancers, le complexe PBAF est majoritairement présent au niveau des promoteurs actifs et dans le corps des gènes, tandis que le complexe ncBAF est localisé au promoteur, mais surtout au niveau des sites CTCF (Michel et al. 2018).

2) Détails de l'interaction entre le complexe BAF humain et son nucléosome cible

En 2020, un complexe BAF humain composé de la sous-unité ATPase SMARCA4 et de neuf sous-unités auxiliaires purifié a été reconstitué *in vitro* et sa structure a été résolue en complexe avec un nucléosome par cryo-électromicroscopie à une résolution de 3,7 Å (He et al. 2020). En parallèle, la purification du complexe BAF humain endogène lié par un nucléosome a également été réalisée par des études de structure par cryo-électromicroscopie et masse spectrométrie crosslinking (Mashtalir et al. 2020).

Ces deux études ont révélé que le complexe BAF humain est constitué de trois modules : le module ATPase, le module de protéine liée à l'actine (ARP) et le module de base (**Figure 12a**) (He et al. 2020; Mashtalir et al. 2020). Les domaines hélicase de SMARCA4 sont le constituant majoritaire du module ATPase. Ces domaines interagissent au niveau de la partie supérieure du nucléosome et au niveau de l'emplacement SHL 2 sur l'ADN nucléosomal (**Figure 12a**) (Xiaoyu Liu et al. 2017; Mashtalir et al. 2020).

L'hétérodimère ACTL6A-ACTB et l'hélice alpha du domaine HSA de SMARCA4 forment le module ARP. Ce module ne se lie pas directement au nucléosome, mais il interagit avec le nucléosome en servant de pont entre le module ATPase et le module de base (**Figure 12a**) (He

et al. 2020; Mashtalir et al. 2020). Les auteurs ont pu observer que le module ARP maintient l'hélice du domaine HSA de SMARCA4 dans une conformation rigide et qu'il permet probablement au module ATPase et au module de base de voir leurs mouvements couplés/coordonnés au cours du remodelage de la chromatine (He et al. 2020).

Le module de base est composé du domaine pré-HSA de SMARCA4 et de sept sous-unités auxiliaires qui représentent environ 80% de la masse moléculaire totale du complexe BAF : ARID1A et DPF2 que l'on trouve uniquement dans le complexe BAF, et cinq autres sous-unités présentes à la fois dans BAF et PBAF : SMARCB1, SMARCD1, SMARCE1 et deux copies de SMARCC (**Figure 12a**) (lu dans (He et al. 2020)). Une sous-partie du module de base, le sous-module de tête, se lie au niveau de l'octamère d'histone sur le nucléosome grâce à la sous-unité SMARCB1 qui a reconnu le nucléosome au niveau du patch acide, tandis qu'une autre sous-partie du module de base, le pont, se lie au module ARP (**Figure 12a**) (He et al. 2020).

La sous-unité ARID1A, qui est la plus grande sous-unité du complexe BAF, sert de base structurelle pour lier SMARCA4 et stabiliser toutes les autres sous-unités du module de base (He et al. 2020). Les deux sous-unités SMARCC quant à elles servent d'échafaudage au module de base en liant toutes ses sous-unités (He et al. 2020; Mashtalir et al. 2020).

Concernant l'interaction entre complexe BAF humain et nucléosome, les deux études de structure ont apporté différentes informations.

He *et al.* ont pu déterminer que le complexe de remodelage BAF humain prend en sandwich le nucléosome d'un côté avec la sous-unité ATPase, et de l'autre côté par SMARCB1, qui interagit avec le patch acide du nucléosome (**Figure 12a**) (He et al. 2020). Les auteurs ont démontré qu'en présence d'un ATP, le module ATPase change de conformation et s'engage avec l'ADN nucléosomal au niveau de SHL 2 (He et al. 2020). L'hydrolyse de l'ATP entraîne la translocation de l'ADN (He et al. 2020).

L'étude de Mashtalir *et al.* a permis de montrer que la liaison du complexe BAF au nucléosome entraîne des changements conformationnels importants dans le complexe BAF (Mashtalir et al. 2020). En effet, en absence de nucléosome, le module ATPase du complexe BAF se trouve au-dessus du module basal, ce qui rend les sous-unités SMARCA4 et SMARCB1 très proches (Mashtalir et al. 2020). En présence d'un nucléosome, le module ATPase se détache du module

basal et tourne autour du module ARP. Cela crée un espace en forme de pince dans lequel le nucléosome vient se fixer au niveau des sous-unités SMARCA4 et SMARCB1 (**Figure 12a**) (Mashtalir et al. 2020). La fixation d'un nucléosome entraîne aussi des modifications de liaison entre les sous-unités des modules ATPase et ARP (Mashtalir et al. 2020).

Ils ont également démontré qu'il y aurait deux mouvements principaux des modules ATPase et ARP par rapport au module basal en absence de nucléosome : un mouvement de pivotement de 40° et un mouvement de basculement de 50° (Mashtalir et al. 2020). Le dynamisme de ces deux modules facilite probablement l'engagement des nucléosomes et le remodelage de la chromatine (Mashtalir et al. 2020). Ils proposent un modèle en trois étapes pour la fixation d'un nucléosome au complexe BAF. Tout d'abord, le complexe BAF cherche un nucléosome à recruter grâce aux mouvements du module ATPase et du module basal (Mashtalir et al. 2020). Puis lorsqu'un nucléosome a été reconnu, les domaines de liaison de l'ATPase et de DPF2 vont se fixer au niveau des queues acétylées des histones H3 (Mashtalir et al. 2020). Le nucléosome est ensuite amarré dans le complexe BAF au niveau de la pince formée par les modules ATPase et basal, reliés par le module ARP (Mashtalir et al. 2020). Les domaines SnAc /post-SnAC de SMARCA4 et le domaine C-terminal de SMARCB1 interagissent tous les deux avec le patch acide des dimères H2A/H2B de chaque côté du nucléosome (**Figure 12a**) (Mashtalir et al. 2020). Le nucléosome est ensuite remodelé grâce à une rotation d'environ 35° du nucléosome dans le complexe et grâce aux mouvements de glissement des modules ATPase et ARP par rapport au module de base ce qui permet la translocation de l'ADN (Mashtalir et al. 2020).

3) Le complexe BAF est nécessaire au maintien des caractéristiques fonctionnelles et de la transcription dans les cellules ES

Les complexes BAF ont des compositions différentes en fonction des types cellulaires.

Ainsi, dans les cellules ES, le complexe esBAF est formé entre autre de la sous-unité catalytique BRG1 (SMARCA4), de deux sous-unités BAF155 (SMARCC1) et des sous-unités BAF60a (SMARCD1) et BAF53b (ACTL6B) (Ho et al. 2009). esBAF ne contient pas les sous-unités spécifiques aux neurones, BRM (SMARCA2), BAF170 (SMARCC2), BAF60c (SMARCD3), BAF45b (DPF1), BAF45c (DPF3) et BAF53b (ACTL6B) (Ho et al. 2009; Michel et al. 2018).

Une déplétion des sous-unités BAF155 et BRG1 du complexe esBAF entraîne une diminution de la prolifération cellulaire et une augmentation de l'apoptose dans les cellules ES (Ho et al. 2009). BRG1 et BAF155 coopèrent pour maintenir le phénotype des cellules ES. Le complexe esBAF est également essentiel au maintien de l'auto-renouvellement et de la pluripotence en cellules ES mammifères (Ho et al. 2009).

Enfin, il a été montré que les TFs associés à la pluripotence comme OCT4 et SOX2 interagissent avec le complexe esBAF, ce qui suggère que esBAF pourrait se lier à ces facteurs pour réguler l'expression des gènes en cellules ES (Ho et al. 2009).

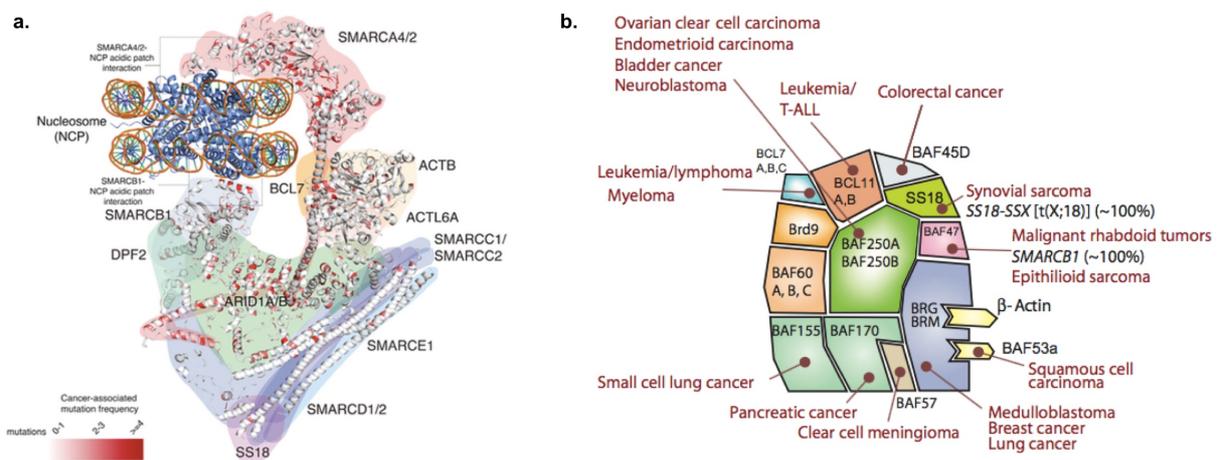


Figure 12 : Modèle de la structure du complexe BAF humain endogène contenant de nombreuses sous-unités mutées dans les cancers.

a. Représentation du modèle structural du complexe BAF humain endogène lié à un nucléosome d'après (Mashtalir et al. 2020). Le complexe BAF est constitué de trois modules : le module ATPase représenté en rose, le module basal représenté en vert et bleu et le module ARP représenté en orange. L'octamère d'histone, via ces hélices et boucles, est représenté en bleu foncé et l'ADN nucléosomal enroulé autour de l'octamère est représenté en orange. Les sous-unités SMARCA4 et SMARCB1 interagissent toutes deux avec le patch acide au niveau du nucléosome. La fréquence de mutation associée à des cancers est représentée par l'échelle de couleur rouge. Plus un résidu est muté dans un cancer et plus sa couleur rouge est intense. **b.** Représentation schématique du complexe mammifère BAF, d'après (Kadoch et Crabtree 2015), dont les sous-unités sont fréquemment mutées dans les cancers humains. Le cancer dans lequel chacune des sous-unités est mutée est précisé en rouge. Les sous-unités SS18-SSX et SMARCB1 sont mutées dans tous les sarcomes synoviaux et dans toutes les tumeurs rhabdoïdes respectivement.

4) Implication du complexe BAF dans les cancers humains

Le séquençage du génome de cellules cancéreuses a mis en évidence des mutations fréquentes de la chromatine dans de nombreux cancers (Helming, Wang, et Roberts 2014). Il a plus particulièrement été montré que des gènes codant les sous-unités du complexe BAF sont mutés dans 20% de tous les cancers humains (**Figure 12b**) (Kadoch et al. 2013; Helming, Wang, et Roberts 2014). Les premières découvertes d'un lien entre complexe BAF et cancer datent des années 90 avec l'identification de la mutation de la sous-unité SMARCB1 (BAF47) dans les tumeurs rhabdoïdes (**Figure 12b**) (lu dans (Helming, Wang, et Roberts 2014)). Au moins une dizaine de gènes codant des sous-unités du complexe BAF ont été trouvés mutés dans les cancers humains (**Figure 12b**) (Helming, Wang, et Roberts 2014). Par exemple, le gène codant la sous-unité ARID1A (BAF250A) est muté dans les cancers ovariens et de l'endomètre, ARID1B (BAF250B) dans les mélanomes et les cancers hépatiques, SMARCA2 (BRM) dans les cancers du cerveau et du colon, SMARCA4 (BRG1) dans les cancers du poumon et les médulloblastomes, SMARCC2 (BAF170) dans les cancers du pancréas et BCL11A/B dans les leucémies (**Figure 12b**) (Helming, Wang, et Roberts 2014; Kadoch et Crabtree 2015). Ces différentes mutations seraient à l'origine d'une perte de fonction du complexe BAF ou bien de la formation d'un complexe BAF aberrant, entraînant une vulnérabilité et une prédisposition aux cancers. Ces données mettent en évidence un rôle suppresseur de tumeur du complexe BAF (Kadoch et al. 2013).

Une analyse de séquençage d'exons a mis en évidence que le complexe BAF est plus souvent muté dans les cancers humains que d'autres complexes comme TIP80, NuRD, ISWI, INO80 et les deux complexes Polycomb (Kadoch et al. 2013). Cette même étude a permis de mettre en évidence chez les mammifères les sous-unités BCL7A, BCL7B, BCL7C, BCL11A, BCL11B, SS18 et BRD9 (**Figure 12b**). Ces nouvelles sous-unités de BAF sont spécifiques, stables et non-échangeables (Kadoch et al. 2013). Ils se sont également intéressés aux fréquences de mutation des différentes sous-unités dans les différents types de cancer. Ils trouvent que la sous-unité ARID1A (BAF250A) est la sous-unité la plus fréquemment mutée (par exemple, dans environ 45% des cancers de l'ovaire et 19% des cancers gastriques) (Kadoch et al. 2013). Enfin, ils ont mis en évidence que des translocations chromosomales au niveau de gènes codant des sous-unités du complexe BAF peuvent également entraîner des cancers humains

(Kadoch et al. 2013). En effet, la translocation du gène de la sous-unité SS18 du complexe BAF est présente dans presque 100% des sarcomes synoviaux (**Figure 12b**) (Kadoch et al. 2013; Kadoch et Crabtree 2015).

Le sarcome synovial, qui est un cancer qui atteint les tissus mous, se caractérise par une translocation chromosomique produisant l'oncoprotéine de fusion SS18–SSX (**Figure 12b**). Cette oncoprotéine qui appartient au complexe cBAF déstabilise SMARCB1 (Kadoch et Crabtree 2013). Les tumeurs rhabdoïdes malignes, atteignant l'encéphale et la moelle épinière des enfants, sont dues à la perte bi-allélique du gène SMARCB1 (**Figure 12b**) (lu dans (Michel et al. 2018)).

Une étude du laboratoire du Dr Kadoch a mis en évidence une perturbation des sous-unités BRD9 et GLTSCR1, appartenant toutes deux au complexe ncBAF, ainsi que de la sous-unité SMARCD1, partagée par les trois complexes BAF, dans les sarcomes synoviaux et des tumeurs rhabdoïdes malignes (Michel et al. 2018). Ainsi, ils ont proposé un modèle dans lequel l'expression des gènes est maintenue par les complexes ncBAF (grâce aux sous-unités BRD9, GLTSCR1 et SMARCD1) au niveau des promoteurs et des sites CTCF en présence d'un complexe cBAF perturbé (par la translocation chromosomique dans la sous-unité SS18 et par la perturbation des sous-unités SMARCB1, SMARCE1 and ARID1A/B) (Michel et al. 2018). Lorsque le complexe ncBAF subit des perturbations chimiques ou des mutations de ses sous-unités BRD9, GLTSCR1 et SMARCD1, cela entraîne l'arrêt du maintien de l'expression des gènes du cBAF et donc des cancers dans lesquels le cBAF est perturbé, comme le sarcome synovial et les tumeurs rhabdoïdes malignes (Michel et al. 2018).

Nakayama *et al.* ont étudié plus en profondeur la sous-unité SMARCB1 (BAF47), présente dans les complexes cBAF et PBAF, dont la fonction est perdue dans les sarcomes synoviaux (Nakayama et al. 2017). Ils ont déterminé que la perte de SMARCB1 altère très probablement la fonction de BAF et sa capacité de liaison aux nucléosomes car la perte de BAF diminue son affinité pour la chromatine (Nakayama et al. 2017). Après une reexpression de SMARCB1, ils ont observé dans des lignées cellulaires, dont les lignées des tumeurs rhabdoïdes malignes, un fort gain d'occupation du complexe BAF sur la chromatine (Nakayama et al. 2017).

Ils ont aussi démontré que SMARCB1 est nécessaire à l'activation des enhancers (typiques et super-enhancers, uniquement via le complexe cBAF) et des promoteurs bivalents (via cBAF et PBAF), ce qui permet de s'opposer à la répression maintenue par les complexes Polycomb (Nakayama et al. 2017).

Il a été montré que SMARCA4 est la sous-unité du complexe BAF qui contient le plus de points de substitutions conduisant à des changements d'acides aminés dans les cancers humains (Kadoch et Crabtree 2015). Mashtalir *et al.* ont cartographié les fréquences de mutation des sous-unités SMARCA4 / SMARCA2 dans les cancers humains sur la structure de l'ATPase SMARCA4 (**Figure 12a**) (Mashtalir et al. 2020). Ils ont ainsi mis en évidence que les résidus les plus fréquemment mutés se situent au niveau des interactions entre le domaine HSA de SMARCA4 et la région répétée ARM des sous-unités ARID1A/B (**Figure 12a**), mais également au niveau des domaines de l'hélicase catalytique, au niveau du domaine HSA de SMARCA4 vers la région d'interaction avec le module ARP, au niveau du site d'insertion du module basal et au niveau des régions d'ancrage, les régions pré-HSA (**Figure 12a**) (Mashtalir et al. 2020). Enfin ils ont montré que les mutations R408W/Q (interface avec le pré-HSA) et R441W/Q (interface avec ARID1A) au niveau de SMARCA4 entraînent une diminution de l'activité de remodelage de la chromatine opérée par le complexe BAF (Mashtalir et al. 2020). Ainsi, Mashtalir *et al.* suggèrent que le positionnement de la sous-unité ATPase et son mouvement dynamique au niveau du module basal doivent être coordonnés et sont nécessaires pour une capacité maximale de remodelage de la chromatine par le complexe BAF (Mashtalir et al. 2020).

Enfin, Mashtalir *et al.* ont identifié des mutations de SMARCA4/2 qui ne sont pas localisées au niveau de l'hélicase. Par exemple, ils ont trouvé des mutations dans les sous-unités SMARCC2 et ARID1A (**Figure 12a**) (Mashtalir et al. 2020). Ils ont également classé l'ensemble des mutations en fonction de leur localisation au niveau de différentes interfaces : entre différentes sous-unités d'un même module de BAF, entre différents modules de BAF ou entre sous-unités de BAF et nucléosome (Mashtalir et al. 2020). Ils ont trouvé que 70% des mutations conduisant à un changement d'acide aminé ont lieu aux interfaces entre différentes sous-unités du module basal, 19% des mutations ont lieu au niveau des interfaces entre sous-

unités de BAF et nucléosome, et 11% des mutations ont lieu à l'interface entre le module ATPase et le module ARP (Mashtalir et al. 2020).

Chapitre 4 : Implication d'un acteur épigénétique dans la maladie d'Alzheimer

I. Préambule sur la maladie d'Alzheimer

La maladie d'Alzheimer est une maladie neurodégénérative incurable qui touche environ 6 millions de personnes dans l'Union européenne, dont 900 000 en France (ce qui représente environ 1,5% de la population). Chaque année, on dénombre environ 225 000 nouveaux cas. Parmi les personnes atteintes de la maladie d'Alzheimer, seulement environ 1% des cas sont héréditaires, c'est-à-dire que la maladie est transmise génétiquement, les autres cas sont sporadiques. Ces chiffres proviennent de la Fondation pour la Recherche sur Alzheimer et de l'Inserm.

La maladie d'Alzheimer se définit de façon histologique par la présence de :

- plaques séniles extracellulaires constituées de peptides beta-amyloïdes (**figure 13a**),
- enchevêtrements neurofibrillaires constitués de la protéine TAU hyperphosphorylée (**figure 13a**),
- une mort neuronale.

La première région à être touchée par cette maladie est l'hippocampe. Le cortex est également touché par la suite. On dénombre plusieurs stades de la maladie d'Alzheimer au cours desquels les plaques séniles et la dégénérescence neurofibrillaire s'étendent à différentes régions du cerveau de façon parallèle et asynchrone, augmentant ainsi la sévérité de la maladie (**figure 13b**).

Bien qu'il s'agisse de la maladie neurodégénérative la plus répandue dans le monde, il n'existe, à ce jour, aucun traitement curatif. Les traitements actuels des déficits sont essentiellement palliatifs. Des résultats d'essais cliniques sur la maladie d'Alzheimer datant de 2013 et 2014 suggèrent que la réduction des niveaux de plaques séniles constituées de peptides β -amyloïde pourrait ne pas être suffisante pour améliorer les résultats cognitifs (Doody et al. 2013; 2014; Salloway et al. 2014).

Une série de publications qui a exploré le lien entre formation de la mémoire et acétylation des histones a proposé que la déficience cognitive pourrait résulter d'une dérégulation épigénétique précoce et maintenue de manière stable. L'histone désacétylase 2 (HDAC2), une enzyme impliquée dans le contrôle épigénétique de l'expression génique, jouerait un rôle central dans cette dérégulation.

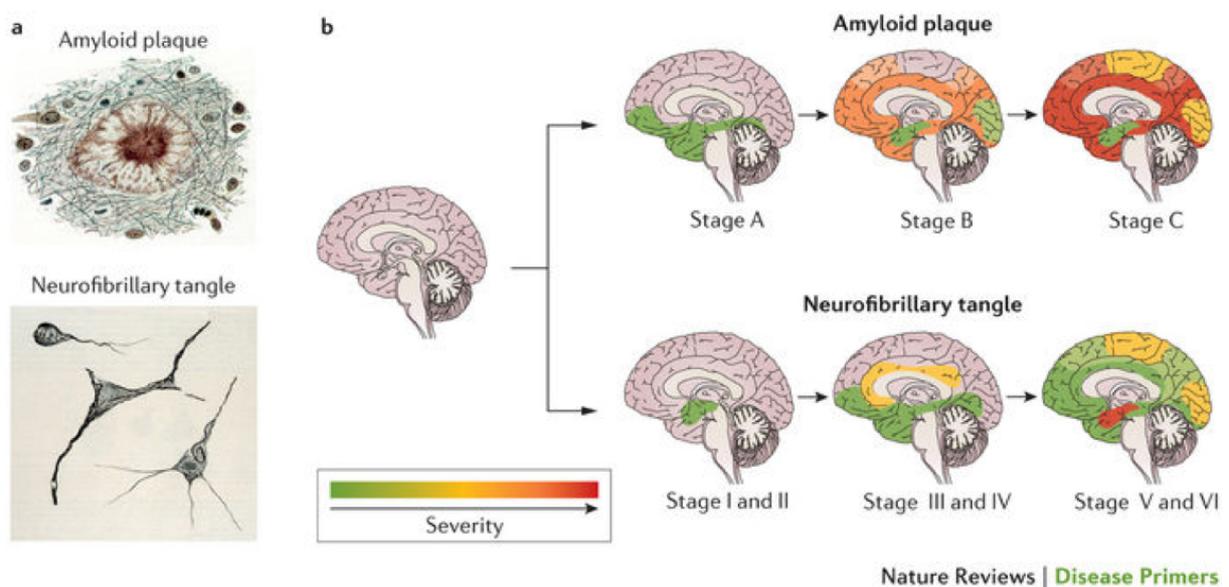


Figure 13 : La progression des plaques séniles extracellulaires et des enchevêtrements neurofibrillaires au cours de la maladie d'Alzheimer

(a) La maladie d'Alzheimer est caractérisée par la présence de plaques séniles extracellulaires constituées de peptides beta-amyloïdes et par un enchevêtrement neurofibrillaire constitué de la protéine TAU hyperphosphorylée. (b) Aux stades précoces de la maladie, les deux défauts sont centralisés dans une région précise puis, au cours de l'évolution de la maladie, les défauts gagnent d'autres régions du cerveau. La couleur représente la sévérité des défauts et donc de la maladie. Plus c'est rouge et plus les défauts sont importants. D'après (Masters et al. 2015).

II. Rôle de l'histone désacétylase 2 (HDAC2) dans la maladie d'Alzheimer

HDAC2, comme toutes les histone desacétylases, a pour fonction principale de supprimer le groupement acétyl des lysines situées au niveau N-terminal des queues des histones. La désacétylation des histones entraîne une compaction de la chromatine, qui favorise la répression de la transcription.

En 2012, le laboratoire du groupe du Dr Tsai à l'université de Harvard a démontré que la protéine HDAC2 est surexprimée de façon précoce dans le cortex et l'hippocampe des cerveaux des patients atteints de la maladie d'Alzheimer (Gräff et al. 2012). Grâce à des expériences d'immunofluorescence, la surexpression de HDAC2 a également été mise en évidence dans le cortex et l'hippocampe de différents modèles murins de cette maladie, dont le modèle CK-p25 (Bie et al. 2014; Gräff et al. 2012). Le modèle de souris CK-p25 est caractérisé par une surexpression aberrante de la protéine p25 dans le cerveau « antérieur » (prosencéphale), ce qui entraîne une neurodégénérescence, une pathologie amyloïde et tau, une déficience cognitive et un dysfonctionnement synaptique (Cruz et al. 2003).

Gräff *et al.* ont également montré que la surexpression de HDAC2 dans le modèle CK-p25 est en corrélation avec une augmentation de la liaison de HDAC2 au promoteur des gènes impliqués dans la formation de la mémoire, de l'apprentissage et de la plasticité synaptique (Gräff et al. 2012). L'accumulation d'HDAC2 au niveau de ces promoteurs entraîne une diminution de l'acétylation des histones et une réduction des niveaux d'ARNm pour les gènes transcrits à partir de ces promoteurs. La surexpression de HDAC2 contribuerait donc à la perte progressive des fonctions cognitives par un effet répresseur sur la transcription de gènes impliqués dans l'apprentissage et la plasticité synaptique (Gräff et al. 2012). Une autre étude a montré que la surexpression de HDAC2 entraîne aussi une diminution de la densité des épines dendritiques, du nombre de synapses, de la plasticité synaptique et de la formation de la mémoire (Guan et al. 2009).

La surexpression de HDAC2 est sous le contrôle transcriptionnel d'un élément de réponse au récepteur des glucocorticoïdes (Gräff et al. 2012). L'activation de cette voie est liée à la pathologie amyloïde (Gräff et al. 2012). En 2017, il a été proposé que le facteur de

transcription SP3, qui interagit avec HDAC2, pourrait être responsable de la répression des gènes impliqués dans l'apprentissage et la plasticité synaptique (Yamakawa et al. 2017).

Le laboratoire du Dr Tsai a également montré que les déficits cognitifs et le dysfonctionnement synaptique des souris CK-p25 peuvent être inversés par knock-down de *Hdac2* au sein des hippocampes (Gräff et al. 2012). Grâce à ce knockdown de *Hdac2* dans l'hippocampe de ces souris, les activités transcriptionnelles et synaptiques, ainsi que la cognition, sont restaurées (Gräff et al. 2012). Le phénotype est donc réversible. De plus, il a également été montré dans d'autres modèles de souris que le knock-down de *Hdac2* rétablit la fonction cognitive (Bie et al. 2014; Singh et al. 2015).

Ces résultats incitent à utiliser des inhibiteurs de HDAC2 pour traiter les déficits cognitifs dans la maladie d'Alzheimer. Plusieurs études ont montré que les inhibiteurs d'HDAC à large spectre restaurent les performances cognitives dans les modèles de rongeurs de la maladie d'Alzheimer (Ricobaraza et al. 2009; Kilgore et al. 2010; Govindarajan et al. 2011; Sung et al. 2013; Bie et al. 2014).

Cependant, à cause de leur faible spécificité et de l'expression ubiquitaire de HDAC2, il n'est pas envisageable d'administrer les inhibiteurs HDAC existants actuellement sur le long terme. Il existe en effet de nombreux effets secondaires de ces inhibiteurs allant de la fatigue, des nausées et vomissements, une neutropénie (Marsoni, Damia, et Camboni 2008) à des cas de toxicité au niveau du tractus gastro-intestinal et du cœur (Lakshmaiah et al., 2014). La très grande distribution tissulaire et la grande variété de processus physiologiques dans lesquels HDAC2 est impliquée va rendre difficile le développement d'inhibiteurs plus sélectifs efficaces et sans effets secondaires graves. Par conséquent, il est donc essentiel d'identifier des cibles thérapeutiques alternatives.

Au cours de ma thèse, j'ai exploré à l'échelle du génome entier l'hypothèse selon laquelle HDAC2 contribue à la maladie d'Alzheimer en provoquant un blocage épigénétique. J'ai utilisé pour cela deux modèles de souris de la maladie d'Alzheimer, les modèles 3xTg-AD (Oddo et al. 2003) et *App*^{tm3 (NL-G-F)} (Saito et al. 2014).

Mon but initial sur ce projet consistait à identifier les cibles de HDAC2 sur le génome par différentes approches : les gènes liés par HDAC2 au niveau de leur promoteur grâce à des expériences de CHIP-seq et les gènes dont l'expression est altérée par HDAC2 grâce à des expériences de RNA-seq.

Matériel et Méthodes

Partie A : expériences

I. Projet portant sur l'organisation nucléosomale des enhancers et la fonction des remodeleurs de la chromatine

1) Culture de cellules ES de souris

Au cours de ma thèse, j'ai utilisé différentes lignées de cellules ES (les cellules 46C, E14, E14Tg2a) et d'autres lignées construites à partir de ces trois lignées primaires (**Tableau 3**). Les cellules ES croissent dans du milieu complet Dulbecco's Modified Eagle (DMEM ; Sigma) additionné de 15 % de sérum de veau fœtal (Invitrogen), 1x d'acides aminés non-essentiels (Invitrogen), 1x de pénicilline/streptomycine (Invitrogen), 0,2 % de β -mercaptoéthanol (Sigma), et une quantité saturante de LIF. Les cellules sont cultivées à 37 °C, 5 % de CO₂, sur des fibroblastes embryonnaires de souris inactivés à la mitomycine C (M4287, Sigma). Quand la confluence cellulaire est d'environ 60 %, les cellules sont remises en suspension par trypsinisation (T4174, Sigma) et étalées à nouveau sur boîte.

Nom de la lignée cellulaire	Remodeleurs étiquetés	Étiquettes utilisées	Lignée de cellules ES d'origine	Références
Brg1-3FTH	Brg1 = Smarca4	(Flag) ₃ -TEV-HA	46C	(de Dieuleveult et al. 2016)
Chd1-3FTH	Chd1	(Flag) ₃ -TEV-HA	46C	(de Dieuleveult et al. 2016)
Chd4-3FTH	Chd4	(Flag) ₃ -TEV-HA	46C	(de Dieuleveult et al. 2016)
Chd8-3FTH	Chd8	(Flag) ₃ -TEV-HA	46C	(de Dieuleveult et al. 2016)
Ep400-3FTH	Ep400	(Flag) ₃ -TEV-HA	46C	(de Dieuleveult et al. 2016)
Brg1-mAID	Smarca4 = Brg1	mAID	E14Tg2a	Nocente et al. 2021

Tableau 3 : Les différentes lignées de cellules ES étiquetées utilisées au cours de ma thèse

Ce tableau contient le nom des différentes lignées cellulaires avec le remodeleur étiqueté, l'étiquette utilisée et la lignée cellulaire d'origine. Ces lignées cellulaires ont pour la première fois été publiées dans l'article présenté dans ma thèse (Nocente et al.) ou dans (de Dieuleveult et al. 2016).

3FTH = (Flag)₃-TEV-HA ; mAID = mini Auxin-Inducible Degron

2) Construction d'un plasmide et génération de la lignée cellulaire Brg1-mAID

Pour l'étiquetage des deux allèles de *Smarca4*, nous avons adopté la stratégie de double sélection par les marqueurs Néomycine et Hygromycine décrite dans (Natsume et al. 2016). Le vecteur Brg1-mAID a été assemblé par une série de modifications à partir des vecteurs pMK287 (mAID-Hygro, Addgene #72825) et pL452 (Néo, (de Dieuleveult et al. 2016)). Tout d'abord, nous avons introduit deux adaptateurs, en amont et en aval de la région codante pour mAID et la résistance Hygro dans le vecteur pMK287. L'adaptateur en amont contient, en plus du site de clonage multiple (MSC), un linker (GGGS)₃ (avant la séquence codant le mAID). La région contenant la séquence linker-mAID a ensuite été amplifiée par PCR grâce à la polymérase Phusion (référence M0531S, NEB) à partir du plasmide pMK287-mAID-Hygro modifié et introduit par clonage dans le vecteur pL452 qui porte la cassette de résistance à la Néomycine sous le contrôle du promoteur PGK.

Les bras d'homologie gauche et droit (400-500 pb chacun) du dernier exon et du 3'UTR de *Brg1* ont été amplifiés par PCR en utilisant de l'ADN génomique E14Tg2A comme modèle. Les oligonucléotides pour l'amplification des bras d'homologie ont été construits pour introduire une mutation silencieuse après intégration de la cassette mAID en aval du locus du gène *Brg1* dans le but d'empêcher l'enzyme Cas9 de recouper.

Au cours de nouvelles étapes de clonage, les produits de PCR des bras d'homologie droit et gauche ont été introduits séquentiellement en utilisant le système de restriction-digestion dans les deux vecteurs (pMK287-mAID-Hygro and pL452-mAID-Neo). Les plasmides finaux pour électroporation ont été préparés avec le kit Nucleobond Maxi kit (Macherey Nagel). Par la suite, les constructions ont été linéarisées et concentrées par précipitation à l'éthanol.

Les vecteurs CRISPR/Cas9 ont été assemblés par hybridation des paires d'oligonucléotides portant les séquences des ARN guides et par clonage dans le plasmide pSpCas9(BB)-2A-Puro (PX459) de Feng Zhang (Addgene #62988) en suivant le protocole décrit dans (Ran et al. 2013). Deux ARN guide distincts ciblant *Brg1* ont été désignés pour cibler une séquence proche du dernier exon de *Brg1*. Ils ont ensuite été clonés séparément par hybridation des oligonucléotides caccgTTGGCTGGGACGAGCGCCTC et aaacGAGGCGCTCGTCCCAGCCAAc pour le premier ARN guide et caccgGCGCCTCGGGGTCAGGACTC et aaacGAGTCCTGACCCCGAGGCGC pour le deuxième ARN guide.

Pour générer des lignées de cellules ES stables, les deux vecteurs codant pour Cas9 et les ARN guides (2,5 µg chacun) et les deux plasmides linéarisés (pMK287, pL452, 10 µg chacun), portant les bras d'homologie de *Brg1* et l'étiquette mAID, ont été introduits par électroporation en utilisant le kit Neon (MPK10096, ThermoFisher) dans des cellules E14Tg2a homozygotes pour un *knock-in* (KI) de la cassette Tir1-2A-puro au locus *Tigre* (Nora et al. 2017). Une double sélection avec Néomycine (250 µg/mL, G-418, Sigma) et Hygromycine (300 µg/mL, Thermo Fisher Scientific) a débuté 24 h après électroporation. Après 6 jours de sélection, les clones uniques ont été repiqués à la main, cultivés dans des plaques de 24 puits et génotypés par PCR et séquençage de la région *knock-in*. Les clones homozygotes pour l'insertion de la cassette mAID au locus *Brg1* ont été sélectionnés après validation par des expériences de western blot.

Toutes les étapes conduisant à la production de cette nouvelle lignée de cellules *Brg1*-mAID ont été réalisées par la post-doctorante du laboratoire, Anida Mesihovic ([Tableau 3](#)).

3) Construction de vecteurs shRNA ciblant BRG1

À mon arrivée dans mon laboratoire de thèse, un important répertoire de shRNA ciblant les remodeleurs de la chromatine CHD1, CHD4, CHD8, EP400 et BRG1 était disponible. Au cours de ma thèse deux nouveaux shRNA ciblant les ARNm de *Brg1* (O5 et O7) ont été construits pour avoir différentes efficacités de déplétion de BRG1 selon le même protocole décrit dans (de Dieuleveult et al. 2016). Les séquences du brin sens du shRNA ciblant chaque remodeleur sont les suivantes :

Chd1 shRNA : 5'-GCAAAGACGGCGACTAGAAGA-3' (de Dieuleveult et al. 2016) ;

Chd4 shRNA D2: 5'- GGACGACGATTTAGATGTAGA-3' (de Dieuleveult et al. 2016) ;

Chd8 shRNA : 5'-GCAGT TACTGACGTCTACA-3' (de Dieuleveult et al. 2016) ;

Ep400 shRNA R1: 5'-GGTAAAGAGTCCAGATTAAAG-3' (de Dieuleveult et al. 2016) ;

Smarca4 shRNA O1: 5'-ACTTCTTGATAGAATTCTACC-3' (de Dieuleveult et al. 2016) ;

Smarca4 shRNA O5: 5'-GCTCCAGTAAAGATGTCTACT-3' ;

Smarca4 shRNA O7: 5'-GAGCGAATGCGGAGGCTTATG-3'.

Le vecteur pHYPER dans lequel est inséré la séquence codant le shRNA d'intérêt comporte également une cassette de résistance à la puromycine servant de marqueur de sélection. Nous disposons également d'un vecteur vide contrôle, le vecteur pLinker.

4) Déplétion de remodeleurs d'intérêt par transfection d'un shRNA ou par incubation avec de l'auxine puis préparation de la chromatine

Les vecteurs pHYPER exprimant les shRNA (utilisés pour la déplétion du remodeleur cible) et les vecteurs contrôle pLinker sont transfectés dans 2 millions de cellules ES, dans lesquelles le remodeleur cible était étiqueté par le tag (Flag)₃-TEV-HA (3FTH) (**Tableau 3**), par électroporation (15 µg de plasmide par électroporation) avec le kit Amaxa nucleofector (référence VVPH-1001, Lonza) ou le kit Néon (référence MPK10096, ThermoFisher). Après 24 h de transfection débute une sélection à la puromycine (2 µg/mL) pendant 30 h pour les cellules transfectées avec les vecteurs shRNA ciblant EP400 et CHD4 ou pendant 48 h pour les cellules transfectées avec le vecteur shRNA ciblant BRG1, CHD1 et CHD8. L'efficacité de déplétion de tous les shRNA a été vérifiée par western blot en utilisant des anticorps monoclonaux anti-Flag (M2, Sigma F1804) ou anti-HA (H7, Sigma H3663) (voir paragraphe I. 5) (de Dieuleveult et al. 2016).

Pour la déplétion des protéines BRG1 en présence d'auxine (IAA, indole-3-acetic acid, Cayman Chemical), 1 mM d'auxine est ajouté au milieu de culture des cellules Brg1-mAID pendant 3 h ou 20 h. L'efficacité de la dégradation des protéines BRG1 a été analysée par western blot en utilisant un anticorps anti-Brg1 (ab110641, rabbit monoclonal, Abcam) (voir paragraphe I. 5) (**Figure 14**).

Quelle que soit la méthode utilisée pour induire la perte de fonction des remodeleurs, les cellules sont incubées 30 min sur boîtes gélatinées afin d'éliminer les fibroblastes embryonnaires. Les cellules sont ensuite fixées dans le milieu de culture comportant 1 % (à volume égal) de formaldéhyde. Après 10 min d'incubation, la réaction est stoppée par ajout de 125 mM (concentration finale) de Glycine et les cellules sont lavées avec du PBS 1X et comptées sur lame de Malassez. Les cellules sont ensuite perméabilisées dans la solution I (15

mM Tris-HCl pH 7,5, 0,3 M Sucrose, 60 mM KCl, 15 mM NaCl, 5 mM MgCl₂, 0,1 mM EGTA et 1x d'inhibiteurs complets de protéase) puis lysées dans la solution II (solution I + 0,8% IGEPAL CA-630 (Sigma)). Après incubation de 15 min sur la glace, les cellules lysées sont diluées dans deux volumes de tampon MNase (20 mM Tris-HCl pH 7,5, 20 mM NaCl, 2 mM CaCl₂, 4 mM MgCl₂, 15 mM KCl, 0,1 M sucrose) pour stopper la réaction, puis centrifugées. La chromatine de 1 million de cellules est digérée avec 6 unités de MNase (New England Biolabs : 200 unités/μl) pour une digestion modérée à 37 °C pendant 10 min. La digestion est stoppée par ajout de 10 mM EDTA. La qualité et la reproductibilité de la digestion MNase ont été vérifiées sur gel d'agarose 1,7 % (voir paragraphe I. 6)) (Figure 15).

13 aller-retours sont faits à la main à travers une aiguille de 26G (7/8) dans une seringue avec la chromatine destinée à des expériences d'immunoprécipitation de la chromatine (ChIP-seq) avec des anticorps anti-histone. La chromatine destinée à des expériences de ChIP-seq anti-TFs ou à des ChIP-tandem est soniquée selon un programme de 4 cycles de 20 s ON et 40 s OFF sur l'appareil Biorupteur à intensité médium. Les aller-retours à travers l'aiguille ou la légère sonication permettent d'améliorer la solubilité de la chromatine. Après centrifugation pour éliminer les débris cellulaires (10 min, 4°C, 14 000 rpm), la chromatine digérée en mono-di-ou tri-nucléosomes est prête pour les expériences ChIP. La chromatine est stockée à -80°C.

5) Vérification de la déplétion des remodeleurs et test des anticorps par western blot

Une petite quantité de cellules déplétées en remodeleur et de cellules contrôles étiquetées 3FTH ont préalablement été mises de côté après l'élimination des fibroblastes embryonnaires sur boîte de gélatine et avant la fixation au formaldéhyde.

Ces cellules sont lavées au PBS 1X froid et centrifugées. Le culot de cellules est incubé 1 h à 4 °C sous agitation dans du tampon RIPA (50 mM Tris pH 7,5, 150 mM NaCl, 1 % NP40, 0,5 % deoxycholate, 0,1 % SDS) + complete 1X (Complete Protease Inhibitor Cocktail, référence 11836145001, Roche) + PMSF 1X (P7626, Sigma). Le lysat est ensuite centrifugé 15 min à 12 000 rpm, 4°C, et le surnageant contenant les protéines est conservé. Les protéines sont quantifiées par un dosage de Bradford (Protein Assay Dye reagent concentrate, #500-0006, Bio-Rad).

Environ 40 µg de protéines sont dénaturées dans du tampon Laemmli (62,5 mM TrisHCl pH 7, 2 % SDS, 10 % glycérol, 0,01 % bleu de Bromophénol, 5 % beta-mercaptoéthanol) pendant 5 à 7 min à 95 °C, puis elles sont déposées dans les puits de cassettes NuPAGE (Gel NuPAGE 4-12 % Bis-Tris, référence NP0335BOX, ThermoFisher) dans une cuvette contenant du tampon de migration 1X NuPAGE (MOPS SDS Running Buffer, référence NP0001, ThermoFisher) pour une migration sur gel entre 90 et 110 V. Après électrophorèse, le gel et la membrane (de nitrocellulose ou PVDF selon la taille des protéines) sont incubés dans du tampon de transfert 1X NuPAGE (Transfert buffer, référence NP0006, ThermoFisher) + 10 % méthanol pendant 15 min à température ambiante (RT) sous agitation. Le transfert du gel contenant les protéines sur la membrane a été réalisé grâce à un système semi-sec (Fast-blot de Biomerta) entre des papiers Whatman, pendant 30 (petites protéines) à 45 min (grandes protéines) à 10 W. La qualité du transfert est ensuite vérifiée grâce à une coloration au rouge Ponceau. Les membranes portant les protéines sont ensuite incubées dans du tampon de blocage, contenant soit (½ tampon Odyssey (Intercept (PBS) blocking buffer, référence 927-70001, Li-Cor) + ½ (PBS 1X), soit PBS 1X + 5 % Régilait), pendant 1 h à RT sous agitation. Après blocage, les membranes sont incubées avec les anticorps dilués toute la nuit (ON, *overnight*) à 4°C. Le lendemain, les membranes sont lavées cinq fois 5 min avec du tampon contenant du PBS 1X + Tween 0,1 % à RT sous agitation. Elles sont ensuite incubées 1 h à RT sous agitation avec l'anticorps secondaire (Odyssey anti-anticorps primaire IRDye, Li-Cor) dilué couplé avec des fluorophores et lavées de nouveau. Les protéines sont révélées grâce aux fluorophores sur une machine Odyssey CLx Li-Cor ou par le système ECL (**Figure 14**).

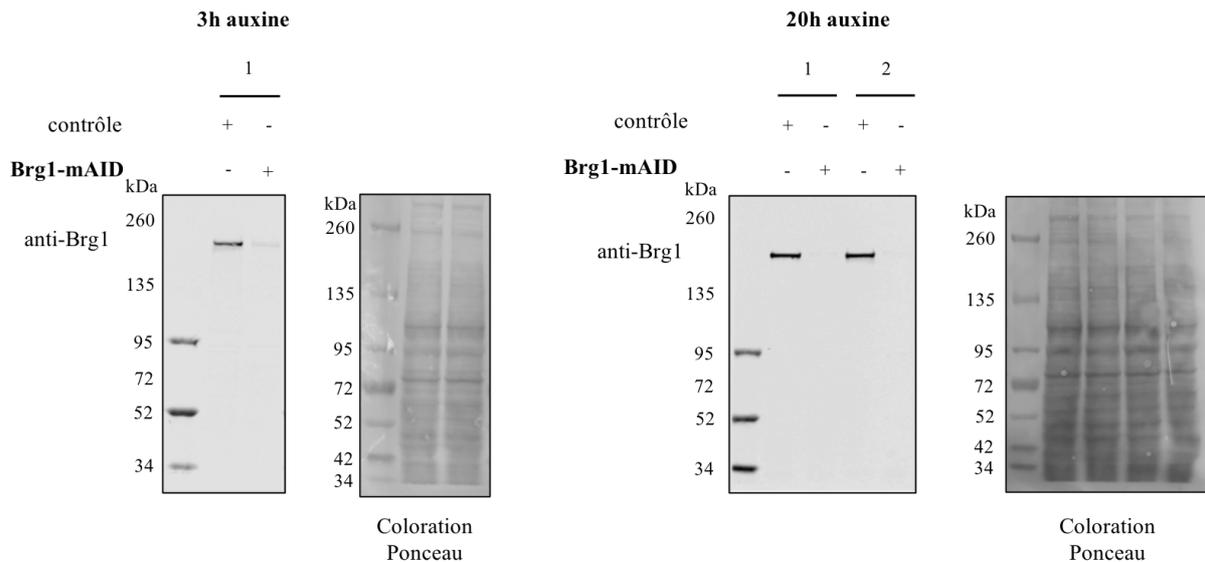


Figure 14 : Vérification de l'efficacité de la déplétion de BRG1 par western blot.

Des cellules ES exprimant les protéines BRG1 contrôles et des cellules ES exprimant BRG1 étiquetées avec mAID (lignée Brg1-mAID) ont été soumises pendant 3h ou 20h à un traitement à l'auxine (1 mM). 40 µg de protéines ont été déposés dans chaque puits du gel et une coloration au Rouge Ponceau a été réalisée après transfert semi-sec sur membrane de nitrocellulose afin de vérifier l'homogénéité des dépôts protéiques et l'efficacité du transfert. Les membranes ont été incubées avec l'anticorps anti-Brg1 (ab110641, dilution 1/2000, rabbit monoclonal) puis lavées, incubées avec l'anticorps secondaire anti-rabbit (IRDye goat anti-rabbit IgG secondary antibody, #926-32211, dilution 1/10000, Li-Cor). Les protéines sont révélées grâce aux fluorophores contenus dans l'anticorps secondaire sur la machine Odyssey.

Taille attendue du remodeleur BRG1 : 181 kDa. Marquer de poids moléculaire : Spectra Multicolor Broad Range Protein Ladder, référence 26623, ThermoFisher.

6) Vérification de la qualité de la digestion MNase de la chromatine sur gel d'agarose

Un petit aliquot de chromatine (correspondant à environ 1 million de cellules) est prélevé après digestion MNase et sonication ou passage dans la seringue afin de vérifier la qualité et la reproductibilité de la digestion MNase.

Le volume de l'aliquot de chromatine ajusté à 300 µL avec de l'eau est incubé 1 h à 65 °C en présence de 15 µg de RNase A (référence R4642, Sigma). Après élimination des ARN, l'aliquot de chromatine est incubé avec 200 µL d'une solution de protéinase K (0,5 mg/mL protéinase K (référence 03115852001, Roche), 100 mM Tris pH 8, 20 mM NaCl, 0.27 % SDS, 10 mM EDTA) pendant 2 h à 55 °C afin de digérer les protéines. L'aliquot de chromatine est ensuite incubé

ON à 65 °C en présence de 200 mM final de NaCl pour détruire la liaison entre les protéines et l'ADN (*reverse crosslink*).

Le lendemain, un volume de Phénol/Chloroforme est ajouté à l'aliquot de chromatine, puis il est vortexé et centrifugé 10 min, à 13 000 rpm, à RT. La phase aqueuse est ensuite collectée précautionneusement et 30 µg de glycogène (référence 10901393001, Roche) sont ajoutés à l'ADN afin de visualiser le culot dans les étapes suivantes. L'ADN est précipité en présence de 1/10^{ème} de volume de NaAc 3M pH 5,2 et de 2 volumes d'éthanol 100 % froid pendant au moins 1 h à -20 °C. Après précipitation, les échantillons sont centrifugés 40 min à 4°C, 14 000 rpm. Le culot ADN-glycogène visible au fond du tube est lavé avec 500 µL d'éthanol 70 %, puis séché à RT. L'ADN est resuspendu dans 20 µL d'eau. La concentration de l'ADN est dosée au Qubit (Invitrogen Qubit dsDNA HS Assay Kit, référence Q32854/ 10616763, Fisher Scientific) et le profil de digestion MNase est analysé sur gel d'agarose 1,7 % (Figure 15).

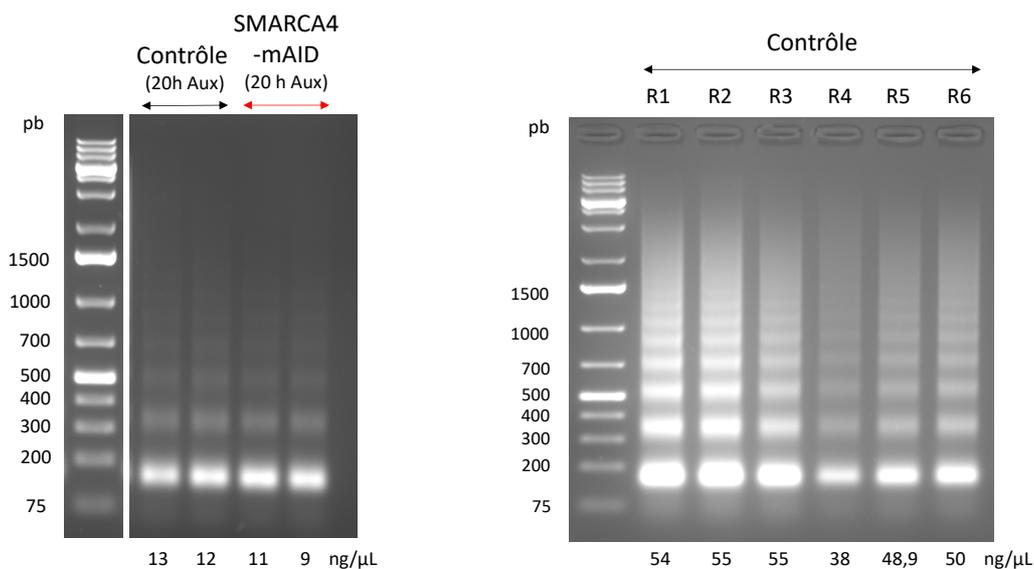


Figure 15 : Profils de digestion d'échantillons de chromatine digérés par la MNase analysés sur gel d'agarose

Les échantillons de chromatine ont tous été digérés avec une dose modérée de MNase (6 unités de MNase pour 1 million de cellules). À gauche : profils de digestion typiques d'échantillons de chromatine passés à travers une aiguille. À droite : profils de digestion typiques d'échantillons de chromatine faiblement soniqués. Les profils de digestion MNase observés sont classiques et très reproductibles entre les différents échantillons (2 échantillons contrôles et 2 échantillons dont la chromatine est déplétée en Brg1 (=Smarca4) passés sur aiguille et 6 échantillons contrôles (R1 à R6, R pour réplicat). 20 µL de chaque échantillon d'ADN ayant des concentrations similaires (mesurées au Qubit) et 7 µL de marqueur (O'GenRuler 1 kb Plus DNA Ladder) sont déposés sur un gel d'agarose 1,7 %. Le gel d'agarose est révélé sous une lampe à UV.

7) Centrifugation sur gradient de sucrose

Afin de séparer les différentes espèces nucléosomales (*footprints* d'ADN, particules subnucléosomales, mononucléosomes, dinucléosomes, etc), des expériences de centrifugation sur gradient de sucrose ont été réalisées. Les solutions TEEP80 30 % sucrose (10 mM TrisHCl pH 7,5, 1 mM EDTA, 1 mM EGTA, 80 mM NaCl et 150 g de sucrose pour 500 mL final) et TEEP80 0 % sucrose (10 mM TrisHCl pH 7,5, 1 mM EDTA, 1 mM EGTA, 80 mM NaCl pour 500 mL final) sont préparées, puis les 11 fractions du gradient sont constituées : 10 % à 30 % de sucrose en incrémentant tous les 2 % à partir des solutions TEEP80 0 % et 30 % de sucrose. 1X final de complete filtré est ajouté à chaque fraction maintenue sur la glace. La première étape de cette expérience consiste à couler très délicatement à la main les différentes fractions du gradient dans des tubes spéciaux Thinwall (référence 331372, Beckman Coulter) en commençant par 900 µL de la fraction 30 %, puis 900 µL de la fraction 28 % et ainsi de suite. Les échantillons de chromatine à déposer sur gradient ont préalablement été digérés et passés à travers une aiguille. Le profil de la digestion MNase et la déplétion du remodeleur BRG1 ont bien été vérifiés, comme décrit auparavant.

De la chromatine contrôle ou déplétée en BRG1, additionnée de 1X de complete final et filtrée afin d'éliminer les petits précipités dus à la congélation, est ensuite déposée dans chaque tube. Quatre gradients sont réalisés en même temps : deux gradients avec de la chromatine contrôle et deux gradients avec de la chromatine déplétée en BRG1. Chaque échantillon de chromatine déposé sur gradient comporte environ 38 millions de cellules. Les tubes sont ensuite centrifugés dans le rotor SW41.2 d'une ultracentrifugeuse Beckman Coulter Optima LE-80K pendant 16 h, à 4 °C, à une vitesse de 40 000 rpm.

Après centrifugation, 20 fractions de 500 µL de chromatine sont collectées (fractions n°1 à n°20) et conservées sur glace puis à -80 °C. Un aliquot de chaque fraction est prélevé afin de visualiser la taille de l'ADN de chaque fraction de chromatine sur un gel de 4 % agarose (**Figure 16**) et afin de préparer de l'ADN Input qui sera directement séquencé.

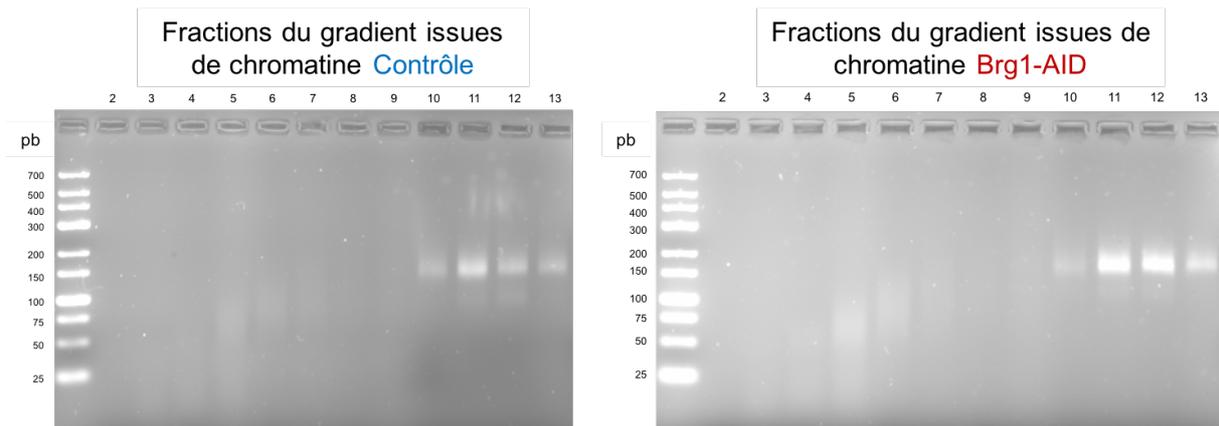


Figure 16 : Analyse des différentes fractions du gradient de sucrose issues de chromatine contrôle ou déplétée en BRG1

L'ADN issu de chromatine contrôle (à gauche) ou de chromatine déplétée en BRG1 (à droite) a été préparé comme précédemment décrit et déposé sur un gel d'agarose 4 %. Différentes tailles de fragments d'ADN sont observées : des fragments d'ADN autour de 150 pb correspondant à des mononucléosomes sont visibles dans les fractions 10 à 13, des fragments d'ADN compris entre 50 et 100 pb et des fragments d'ADN inférieurs à 50 pb.

8) Expériences de ChIP-seq

Différentes expériences de ChIP-seq ont été réalisées dans des buts différents :

1. Cartographier la liaison des TFs au niveau des éléments régulateurs du génome, et particulièrement au niveau des enhanceurs,
2. Cartographier la présence d'histones au niveau des enhanceurs en présence et en absence de remodeleur,
3. Déterminer la nature des différentes espèces nucléosomales et subnucléosomales présentes dans les différentes fractions de chromatine après centrifugation sur gradient de sucrose en présence et en absence du remodeleur Brg1.

Pour les ChIP-seq réalisés sur de la chromatine de cellules ES 46C ou E14 dans le but 1. (cartographier la liaison des TFs sur les enhanceurs), deux étapes de pré-traitement (*preclearing* et blocage) des billes de protéine G-agarose et de la chromatine sont réalisées afin de réduire le bruit de fond des expériences de ChIP-seq. Les ChIP-seq réalisés sur de la chromatine contrôle ou déplétée en remodeleur avec ou sans centrifugation sur gradient dans les buts 2.

et 3. comportent uniquement l'étape de pré-traitement (seulement le blocage) des billes de protéine G-agarose.

Préparation des billes de protéine G – Agarose bloquées BSA/tRNA :

Les billes de protéines G-agarose (référence 11243233001, Sigma) prévues pour l'IP sont lavées trois fois dans du tampon CHIP (20 mM Tris-HCl pH 7,5, 20 mM NaCl et 4 mM MgCl₂), centrifugées 3 min, 800 g, à 4 °C et préincubées pendant 1 h 30 à 4 °C sous agitation avec un mélange de BSA (200 µg/mL, référence B6917, Sigma) et ARNt de levure (50 µg/mL, référence AM7119, ThermoFisher) dans du tampon CHIP. Cette préincubation permet de saturer de BSA/ARNt d'éventuels sites non-spécifiques de la résine qui pourraient lier la chromatine. Après lavage dans du tampon CHIP et centrifugation, les billes sont reprises dans du tampon CHIP supplémenté avec du complete 1X et conservées à 4 °C.

Pré-clearing de la chromatine (uniquement pour les ChIP fait dans le but 1.) :

Les échantillons de chromatine sont préparés afin d'avoir 10 millions de cellules par échantillon à immunoprécipiter (IP). Ces échantillons de chromatine sont ensuite déposés sur des filtres de 0,45 µM, 13 mm (référence SLHVX13N, Merck) afin d'éliminer des précipités formés lors de la congélation. Le filtre est rincé après le passage de la chromatine avec 50 µL de tampon CHIP. 40 µL de billes de protéine G-agarose bloquées BSA/ARNt sont ajoutées aux échantillons de chromatine filtrés et incubés 1 h 30 à 4 °C sous agitation. Ce pré-clearing de la chromatine permet de se débarrasser des régions de la chromatine susceptibles d'avoir une affinité non-spécifique pour la résine et donc de « parasiter » l'interprétation du ChIP-seq. Après incubation, les échantillons sont centrifugés 3 min à 800 g à 4 °C, la chromatine « propre » est prête l'IP et les billes sont jetées.

Un petit aliquot de chromatine est prélevé à cette étape afin de servir d'échantillon Input pour le séquençage.

Incubation anticorps-chromatine :

Si la chromatine n'a pas subi d'étape de *pré-clearing*, les différents échantillons doivent être préparés. Dans le but 2., des échantillons de chromatine contenant environ 1,2 millions de cellules et passés sur aiguille que l'on filtre comme précédemment ont été utilisés. Dans le

but 3., nous avons utilisé des groupes des fractions 2-3-4, 5-6 et 7-8-9 issus des expériences de centrifugation sur gradient de sucrose que l'on filtre comme précédemment. Les échantillons sont ajustés après filtration au même volume avec du tampon MNase + 10 mM EDTA.

Les échantillons de chromatine issus du *pré-clearing* (but 1.) ou non (buts 2. et 3.) sont ensuite incubés ON à 4 °C sous agitation avec 5 µg d'anticorps anti-histone ou 20 µg d'anticorps anti-TF ([Tableau 4](#)).

La suite des étapes des CHIP-seq sont communes à tous les buts (1., 2., et 3.).

Incubation anticorps-chromatine avec les billes de protéine G-agarose :

Après l'incubation ON des échantillons de chromatine avec l'anticorps, 50 µL de billes de protéine G-agarose bloquées sont ajoutées et incubées pendant 4 h à 4 °C sous agitation.

Lavages des billes après IP :

Les billes auxquelles sont fixées les complexes (sub)nucléosomes-TFs/histones sont soumises à une série de trois lavages consécutifs avec 10 mL de tampon TEN (20 mM Tris-HCl pH 7,5, 150 mM NaCl, 3 mM MgCl₂, 0,1 mM EDTA et 0,01 % IGEPAL), puis quatre lavages consécutifs avec 10 mL de tampon WBLiCl (50 mM Hepes pH 7,5, 500 mM LiCl, 1 mM EDTA, 1 % IGEPAL, 0,7 % Na-désoxycholate). Entre deux lavages, les échantillons sont centrifugés 3 min, 800 g, à 4°C. Les billes sont reprises dans 3 mL de tampon TEN et centrifugés 3 min, 800 g, à 4 °C.

Élution des échantillons :

Les billes sont mises en suspension dans 500 µL de tampon d'éluion (50 mM Tris-HCl pH 8, 10 mM EDTA et 1 % SDS 10%) préchauffé à 65°C et éluées sur thermomixer pendant 15 min, à 65 °C, sous agitation. Le surnageant est ensuite récupéré après passage sur colonne Pierce (référence 89868, ThermoFisher) et les éluats, contenant les complexes (sub)nucléosomes-TFs/histones, peuvent être conservés à -20 °C.

Extraction d'ADN des échantillons de chromatine IP et Input :

Les échantillons de chromatine IP et Input sont préparés comme décrits précédemment :

- Échantillons IP élués : digestion des protéines avec 0,2 mg/mL final de protéinase K, destruction des liaisons ADN-protéines avec 200 mM final de NaCl ON.
- Échantillons Input : digestion des ARN avec 15 µg de RNaseA, digestion des protéines avec 0,2 mg/mL final de protéinase K, destruction des liaisons ADN-protéines avec 200 mM final de NaCl ON.

L'ADN est ensuite extrait au phénol-chloroforme, puis précipité à l'éthanol, comme décrit précédemment. L'ADN, venant du CHIP ainsi purifié, est quantifié par un dosage Picogreen (Quant-iT PicoGreen dsDNA Reagent, référence P7581, Invitrogen).

Préparation des libraries d'ADN en vue du séquençage :

Pour chaque échantillon d'ADN, une banque a été préparée au CNRGH d'Evry : il s'agit d'un ajout d'oligonucléotides par ligation suivi d'une amplification par PCR avec un nombre modéré de cycle. Ces banques sont produites avec les kits « MicroPlex Library Preparation Kit v2 (12 indexes) » (référence C05010013, Diagenode) et plus récemment avec le kit « 24 Dual indexes for MicroPlex Kit v3 /48 rxns » (référence C05010003, Diagenode). Elles sont ensuite séquencées principalement sur Illumina NextSeq 500 avec le kit « NextSeq 500 High Output v2 kit (75 cycles) » (référence FC-404-2005, Illumina).

Toutes les expériences de MNase-ChIP-seq sont réalisées en duplicat ou en triplicat.

9) Expériences de ChIP-seq séquentiel : ChIP-tandem

Deux protocoles de ChIP-tandem existent dans mon laboratoire de thèse. Le premier protocole développé en 2014 est disponible ici : <http://dx.doi.org/10.1038/protex.2014.040> (DOI : 10.1038/protex.2014.040, Protocol exchange). Brièvement ces expériences de ChIP-tandem reposent sur l'utilisation d'un très grand nombre de cellules (au moins 250 millions) par échantillon de chromatine digérés par la MNase et préparés comme expliqué dans la partie 2). Les cellules utilisées pour cette expérience ont le gène d'intérêt étiqueté avec (Flag)₃-TEV-HA. Le ChIP-tandem comprend deux étapes d'IP, la première IP avec des billes couplées à un anticorps anti-HA (ou anti-FLAG), puis une étape d'éluion des complexes nucléosomes-remodeleurs ou nucléosomes-TFs soit par compétition avec un peptide, soit par

clivage de la protéase TEV, enfin la deuxième IP avec des billes couplées à un anticorps anti-FLAG (ou anti-HA). Les complexes sont ensuite élués et l'ADN est extrait.

Au cours de ma thèse, j'ai adapté ce protocole de CHIP-tandem afin de pouvoir s'affranchir de l'étiquette (Flag)₃-TEV-HA couplée au gène d'intérêt dans les cellules ES. Mon but était de mettre en évidence une éventuelle colocalisation entre le TF OCT4 et l'histone H2B.

Pour cela, j'ai utilisé des échantillons de 243 millions de cellules contrôle E14 chacun dans lesquels la chromatine a été digérée par la MNase (digestion vérifiée sur gel de 1,7 % agarose) et légèrement soniquée.

Préparation des billes de protéine G-agarose :

Même protocole que pour le CHIP-seq classique (voir précédemment).

Préparation des échantillons de chromatine et première IP :

On travaille avec 243 millions de cellules par échantillon en réplicat.

Même protocole de filtration que pour le CHIP-seq classique (voir précédemment). Les échantillons de chromatine sont incubés avec 45 µg d'anticorps anti-Oct4 ou 60 µg d'anticorps non-spécifique au total, ON à 4 °C sous agitation ([Tableau 4](#)).

Incubation avec les billes de protéine G – agarose et lavages :

50 µL de billes de protéine G – agarose bloquées sont ajoutées à chaque échantillon de chromatine et incubées 4 h à 4 °C sous agitation.

Après incubation, les différents tubes du même génotype sont mélangés, centrifugés 2 min, 800 g, à 4 °C et lavés 7 fois avec 8 mL de tampon TEN. Les billes sont ensuite remises en suspension dans 75 µL de tampon TEN pour avoir un volume final de 150 µL de *slurry* (billes en suspension dans le tampon).

Élution par le peptide OCT4 :

Le peptide OCT4 (séquence en C-ter d'OCT4 : H-Cys-Lys-Lys-Lys-Lys-Pro-Ser-Val-Pro-Val-Thr-Ala-Leu-Gly-Ser-Pro-Met-His-Ser-Asn-OH, GeneCust) est solubilisé dans du tampon TEN à une concentration finale de 2 µg/µL. Le pH est bien situé autour de 7,5 (vérifié sur papier pH).

Aux 150 µL de billes en suspension, on ajoute 150 µL de peptide OCT4 à 2 µg/µL et 1X final de complete, puis les échantillons sont incubés 2 h à RT sous agitation. Après 5 min de centrifugation à 4 °C, 1 000 rpm, le surnageant correspondant à l'éluat n°1 est récupéré. Une élution de 2 h, à RT, sous agitation, avec 300 µL de peptide OCT4 à 1 µg/µL est appliquée aux billes. Après centrifugation, l'éluat n°2 est récupéré. Les billes sont ensuite éluées ON à 4 °C sous agitation avec 300 µL de peptide OCT4 à 1 µg/µL. Le lendemain, après centrifugation, l'éluat n°3 est collecté, mélangé avec les éluats n°1 et n°2 et filtré sur colonne Pierce. Les échantillons élués contenant les protéines OCT4 liées à leurs cibles sont prêts pour la suite du protocole.

Les billes restantes sont lavées avec du tampon TEN et incubées dans 200 µL de tampon d'élution (préchauffé à 65 °C) pendant 15 min, à 65 °C, sous agitation. Cet éluat contient la chromatine qui n'est pas liée par OCT4.

Deuxième IP :

La chromatine éluée, qui était liée par OCT4, est incubée avec l'anticorps anti-H2B ([Tableau 4](#)) ON à 4 °C sous agitation.

Les étapes suivantes sont les mêmes que pour un protocole de CHIP-seq classique :

- Les billes de protéine G – agarose sont préparées comme précédemment pour la suite de la deuxième IP
- les échantillons de chromatine – anticorps sont incubés avec les billes de protéine G – agarose bloquées.
- Après les 4 h d'incubation, les billes sont lavées, puis éluées avec le tampon d'élution.
- Les protéines sont ensuite éliminées des échantillons élués et les ADN sont extraits, puis quantifiés grâce à un dosage Picogreen. Les fragments d'ADN obtenus à la fin du protocole ont été à la fois liés par OCT4 et par l'histone H2B.
- Les bibliothèques d'ADN sont préparées, puis séquençage.

Les expériences de CHIP-tandem sont réalisées en duplicat.

Protéine ciblée / nom de l'anticorps	Référence	Type d'anticorps	Marque	Quantité utilisée	Expérience et but
anti-H2A	ab18255	Rabbit polyclonal	Abcam	5 µg	MNase-ChIP-seq ; but 3.
anti-H2B	ab1790	Rabbit polyclonal	Abcam	5 µg	MNase-ChIP-seq ; buts 2. et 3. + ChIP-tandem
anti-H3	ab1791	Rabbit polyclonal	Abcam	5 µg	MNase-ChIP-seq ; buts 2. et 3.
anti-H4	ab7311	Rabbit polyclonal	Abcam	5 µg	MNase-ChIP-seq ; but 3.
anti-H3K27ac	ab4729	Rabbit polyclonal	Abcam	5 µg	MNase-ChIP-seq ; but 1.
anti-CTCF	07-729	Rabbit polyclonal	Millipore	20 µg	MNase-ChIP-seq ; but 1.
anti-Esrrib	PP-H6705-00	Mouse monoclonal	PPMX	20 µg	MNase-ChIP-seq ; but 1.
anti-Nanog	RCAB-001P	Rabbit polyclonal	Reprocell	10 µg	MNase-ChIP-seq ; but 1.
anti-Oct4	sc-8628	Goat polyclonal	Santa Cruz Biotechnology	20 µg	MNase-ChIP-seq ; but 1.
anti-Oct4	sc-5279	Mouse monoclonal	Santa Cruz Biotechnology	20 µg	MNase-ChIP-seq ; but 3.
anti-Oct4	MBS470786	Rabbit polyclonal	myBioSource	45 µg	tandemChIP
anti-Pol II	sc-55492	Mouse monoclonal	Santa Cruz Biotechnology	20 µg	MNase-ChIP-seq ; but 1.
anti-Sox2	sc-17320	Goat polyclonal	Santa Cruz Biotechnology	20 µg	MNase-ChIP-seq ; but 1.
anti-TBP	ab51841	Mouse monoclonal	Abcam	20 µg	MNase-ChIP-seq ; but 1.
Isotype control	ab37415	Rabbit polyclonal	Abcam	60 µg	tandem-ChIP

Tableau 4 : Anticorps utilisés dans les expériences de MNase-ChIP-seq et de ChIP-tandem.

Les couleurs correspondent aux expériences et aux buts dans lesquels les anticorps sont utilisés : vert pour les anticorps utilisés en MNase-ChIP-seq qui vont permettre de cartographier les TFs sur les enhancers (but 1.), bleu pour les anticorps utilisés en MNase-ChIP-seq qui vont permettre de déterminer la nature des particules subnucléosomales (but 3.), jaune pour les anticorps utilisés en MNase-ChIP-seq qui vont permettre soit de cartographier les histones au niveau des enhancers en présence et en absence de remodeleurs (but 2.), soit de déterminer la nature des particules subnucléosomales (but 3.), orange pour les anticorps utilisés en MNase-ChIP-seq qui vont permettre soit de cartographier les histones au niveau des enhancers en présence et en absence de remodeleurs (but 2.), soit de déterminer la nature des particules subnucléosomales (but 3.), soit de vérifier la colocalisation entre OCT4 et H2B par ChIP-tandem et gris pour les ChIP-tandem.

II. Projet portant sur l'implication de HDAC2 dans la régulation de la transcription de modèles murins de la maladie d'Alzheimer

1) Caractéristiques des lignées de souris utilisées

Deux modèles de souris de la maladie d'Alzheimer ont été utilisés au cours de ma thèse : les souris 3xTg-AD et les souris App^{tm3 (NL-G-F)} (Oddo et al. 2003; Saito et al. 2014). Les caractéristiques de ces deux modèles de souris sont résumées dans le **Tableau 5**.

Modèle de souris Alzheimer	Transgène ou Gène ^{mutation}	Pathologie
3xTg-AD	<p>APP^{K670N/M671L}</p> <p>MAPT^{P301L}</p> <p>PS1^{M146V} KI</p>	<ul style="list-style-type: none"> • Hyperphosphorylation de Tau et enchevêtrements neurofibrillaires. • Plaques β-amyloïdes • Déficits cognitifs
App ^{tm3 (NL-G-F)}	<p>APP^{KM670/671NL}</p> <p>(Swedish)</p> <p>APP^{I716F}</p> <p>(Beyreuther/Iberian)</p> <p>APP^{E693G} (Arctic)</p>	<ul style="list-style-type: none"> • PAS de tauopathie ni d'enchevêtrements neurofibrillaires • PAS de perte neuronale, mais perte de synapse • Plaques β-amyloïdes • Déficits cognitifs (mémoire spatiale et apprentissage)

Tableau 5 : Caractéristiques des modèles de souris 3xTg-AD et App^{tm3 (NL-G-F)}.

Les modèles 3xTg-AD et App^{tm3 (NL-G-F)} sont deux modèles murins différents de la maladie d'Alzheimer. Les transgènes (en bleu) ou les mutations (en rouge) conférant les défauts de la maladie d'Alzheimer sont indiqués pour les deux modèles. Les défauts associés à la maladie d'Alzheimer humaine et possédés par ces deux modèles de souris sont également indiqués dans la colonne « pathologie ». D'après (Oddo et al. 2003; Saito et al. 2014).

Dans le but d'étudier un éventuel blocage épigénétique médié par HDAC2 dans le modèle 3xTg-AD, plusieurs autres lignées cellulaires ont été produites et utilisées : une lignée de souris saines dans laquelle HDAC2 est déplétée, une lignée de souris 3xTg-AD dans laquelle HDAC2

est déplétée et une lignée de souris dans laquelle *Hdac2* est étiquetée par HA-Flag sur un seul de ses allèles

Ces différentes lignées de souris sont présentées plus en détail dans la partie Résultats B.I. ([Tableau 6](#)).

2) Génotypage des souris par qPCR

Les souris *Hdac2*^{-/-} sont génotypées par qPCR en utilisant des amorces ciblant le gène rapporteur *LacZ* qui sert de marqueur de sélection pour les cassettes de *gene-trap* bien insérées et qui conduisent à la perte de fonction de HDAC2.

LacZF1 : 5' GTTTCGTCAGTATCCCCGTTT 3' ;

LacZR1 : 5' GGGTTGCCGTTTTTCATCATA 3'.

La distinction entre homozygotes et hétérozygotes se fait par un dosage par rapport au gène de la préséniline 1 (*Psen1*) sauvage :

PS1Fwt2 : 5' GCCATCATGATCAGTGTCATTGTCATTA 3' ;

PS1R2 : 5' ACTGCCTCGTGCTCACCTTGTAG 3'.

Les souris 3xTg-AD ; *Hdac2*^{-/-} sont des triples homozygotes pour le modèle 3xTg-AD (homozygotes pour les gènes *App*, *Psen1* et *Tau*). Elles sont génotypées par qPCR en suivant le gène *LacZ*. Pour distinguer les souris homozygotes des hétérozygotes, un dosage par rapport au gène de la préséniline 1 (*Psen1*) mutant est effectué :

PS1mutF2 : 5' CATGATCAGTGTCATTGTCGTGG 3' ;

PS1R2 : 5' ACTGCCTCGTGCTCACCTTGTAG 3'.

Les souris *Hdac2*^{HF/+}, comme les souris 3xTg-AD ; *Hdac2*^{HF/+}, sont génotypées par qPCR en suivant les deux allèles du gène *Hdac2*.

Amorces pour l'allèle sauvage :

H2HFwtF : 5' GAACAACCTCAGCAACCCTTGAA 3'

H2HFwtR : 5' CAGTCCATGCCAAAACAGTATAAAA 3'

Amorces pour l'allèle étiqueté :

H2HFmutF2 : 5' CAACAGAGCCAAGTCAGAACA 3'

H2HFmutR2 : 5' GTCGTCGTCATCCTTGTAGTCT 3'

Les souris App^{tm3 (NL-G-F)} sont génotypées par qPCR grâce au couple d'amorces :

Ricken F : 5' CTCCTTGTGGCTGGCGGTCACAC 3'

Ricken R : 5' CTATCGTGGACCGAGAATGGTCATG 3'

3) Dissection des hippocampes de souris et préparation en vue des expériences de western blot et de RNA-seq

Dissection des hippocampes :

Les souris sont euthanasiées par dislocation cervicale, puis le cerveau est prélevé et déposé dans une boîte de Pétri sur glace avant d'être lavé avec du Hank's froid (référence 14175053, ThermoFisher). Le cerveau est coupé en deux dans le sens sagittal, le cervelet et les lobes oculaires sont éliminés. Les deux hémisphères du cerveau sont lavés dans une nouvelle boîte de Pétri avec du Hank's froid et les hippocampes sont ensuite prélevés.

Préparation des hippocampes pour les expériences de western-blot :

Les hippocampes disséqués sont broyés dans du Hank's, puis du PBS 1X est ajouté. Les protéines des hippocampes sont extraites dans 100 µL de tampon RIPA, 1X complete et 1X PMSF après incubation pendant 1 h à 4 °C sous agitation et centrifugation. Les protéines extraites sont ensuite quantifiées à l'aide d'un dosage de Bradford et conservées à -80 °C.

Préparation des hippocampes pour les expériences de RNA-seq :

Les hippocampes d'une souris fraîchement disséqués sont déposés dans 300 µL de Qiazol (provenant du kit RNeasy Lipid Tissue Mini Kit, référence 74804, Qiagen), avant d'être broyés à l'aide d'un piston sur la glace. Le matériel broyé est conservé à -80°C pour un temps bref. Pour les RNA-seq, il est très important de toujours disséquer les souris contrôles en même temps que les souris « test », que la dissection soit faite toujours au même moment de la journée et par le même expérimentateur. Cela permet de limiter les effets « batch », ou effet confondants, dans les RNA-seq. Les effets « batch » sont des effets indésirables provenant de

l'environnement (expérimentateur, moment de la journée, température, état de la souris) qui vont perturber et introduire des biais dans les expériences de RNA-seq.

4) Expériences de western blot à partir de protéines extraites d'hippocampe

Les protéines précédemment extraites sont dénaturées par ajout de tampon Laemmli et par la chaleur (5 min à 95 °C).

Le protocole de western blot est sensiblement le même que celui décrit dans la partie I.5) avec quelques exceptions ou précisions :

- Les protéines sont transférées sur une membrane de nitrocellulose
- La membrane portant les protéines est incubée avec les anticorps anti-HA (clone 3F10, rat monoclonal, dilution : 1/1000, référence 11867423001, Merck) et anti-HDAC2 (clone 3F3, mouse monoclonal, dilution : 0,5 µg/µL, référence 05-814, Millipore) pour les souris 3xTg-AD ; *Hdac2^{HF/+}* et contrôles (*Hdac2^{HF/+}*). La membrane portant les protéines extraites des souris NLGF et contrôles est uniquement incubée avec l'anticorps anti-HDAC2.
- Les protéines HDAC2 sont révélées grâce à l'appareil Odyssey.

Au total 95 souris ont été impliquées dans ces études de western blot :

- 7 souris contrôles (*Hdac2^{HF/+}*, 4 mâles et 3 femelles) et 7 souris 3xTg-AD ; *Hdac2^{HF/+}* (4 mâles et 3 femelles) âgées de 2 mois,
- 7 souris contrôles (*Hdac2^{HF/+}*, 4 mâles et 3 femelles) et 7 souris 3xTg-AD ; *Hdac2^{HF/+}* (4 mâles et 3 femelles) âgées de 3 mois,
- 4 souris mâles contrôles (*Hdac2^{HF/+}*) et 4 souris mâles 3xTg-AD ; *Hdac2^{HF/+}* âgées de 5 mois,
- 6 souris contrôles (*Hdac2^{HF/+}*, 3 mâles et 3 femelles) et 6 souris 3xTg-AD ; *Hdac2^{HF/+}* (2 mâles et 4 femelles) âgées de 9 mois,
- 12 souris femelles contrôles (C57) et 12 souris femelles *App^{tm3 (NL-G-F)}* âgées de 3 mois,
- 12 souris femelles contrôles (C57) et 12 souris femelles *App^{tm3 (NL-G-F)}* âgées de 9 mois.

5) Quantification des protéines HDAC2 après western-blot

Après la révélation de la protéine d'intérêt sur la membrane, il est nécessaire, pour la quantification de cette protéine d'intérêt, de colorer toutes les protéines de la membrane afin qu'elles soient quantifiées dans un but de normalisation. Pour cela, la membrane est rincée à l'eau, puis incubée pendant 5 min à RT sous légère agitation avec 5 mL de la solution commerciale Revert 700 Total Protein Stain provenant du kit « Revert 700 Total Protein Stain » (référence 926-11016, Li-Cor), puis lavée deux fois avec la solution commerciale Revert 700 Wash Solution. Les protéines peuvent désormais être détectées sur la machine Odyssey (dans le canal 700 nm).

La quantification se fait grâce à ImageStudio, l'application de l'Odyssey. Les photos des membranes prises avec l'Odyssey sont importées afin d'effectuer la quantification totale et la quantification de la protéine d'intérêt. Pour cela, des rectangles sont dessinés autour des pistes (quantification totale) ou des bandes d'intérêt (quantification de la bande d'HDAC2) qui vont automatiquement donner une mesure de signal et d'aire. Dans un fichier Excel, le ratio Signal/Area pour chaque piste (quantification totale) et pour chaque bande (quantification d'une protéine sur une membrane) est calculé.

Pour la quantification des protéines totales de chaque piste, le « Lane normalization Factor » est calculé :

$$\text{Facteur de normalisation par piste} = \frac{(\text{Signal/Area de la piste})}{(\text{Signal/Area le plus fort de la membrane})}$$

Le signal normalisé par le nombre de protéines total dans chaque piste est ensuite calculé :

$$\text{Signal normalisé} = \frac{(\text{Signal/Area de la piste})}{(\text{Lane normalization Factor de la même piste})}$$

Pour la quantification des protéines sur une membrane, une 2ème normalisation du signal est calculée par le nombre de protéines dans l'échantillon contrôle :

$$\text{Signal normalisé final} = \frac{(\text{Signal normalisé par nombre total de protéines 3xTg - AD})}{(\text{Signal normalisé par nombre total de protéines contrôle})}$$

Des graphiques sont ensuite réalisés (avec Excel et un script R créé par moi-même, calculs de moyenne, erreur standard).

La significativité des résultats est établie avec le test statistique des rangs signés de Wilcoxon pour échantillons appariés. En effet, moins de 30 échantillons sont disponibles, la loi que suivent les données est inconnue (test non paramétrique) et l'hypothèse que la quantification des données est pairée entre le contrôle et le mutant d'une même membrane est posée (car les quantités sont révélées en même temps (même temps d'exposition...)). Les calculs sont réalisés avec le logiciel R. L'hypothèse H0 qu'il n'y a pas de différence entre la condition 3xTg-AD (3xTg-AD ; *Hdac2*^{HF/+}) et contrôle (*Hdac2*^{HF/+}) concernant l'expression de la protéine est faite (j'ai mélangé les résultats des membranes anti-HDAC2 et anti-HA pour le modèle 3xTg-AD car cela permet de suivre les deux allèles de *Hdac2*).

Pour tester si la différence d'expression de la protéine HDAC2 entre les souris 3xTg-AD ; *Hdac2*^{HF/+} mâles et femelles est significative, le test statistique de Mann-Whitney est utilisé. Il s'agit d'un test non paramétrique, pour les échantillons de petite taille et les échantillons qui ne sont pas appariés. J'ai fait l'hypothèse H0 qu'il n'y a pas de différence entre les mâles et les femelles 3xTg-AD ; *Hdac2*^{HF/+} concernant l'expression de HDAC2 (les quantifications des membranes anti-HDAC2 et anti-HA pour le modèle 3xTg-AD ont été mélangées car cela permet de suivre les deux allèles de *Hdac2*). Si la p-value est supérieure à 0,05, il n'y a pas de différence significative entre les mâles et les femelles.

6) Expériences de RNA-seq

Les tissus préparés au préalable sont décongelés sur glace. Le kit RNeasy Lipid Tissue Mini Kit (référence 74804, Qiagen) est utilisé pour l'extraction des ARN totaux. Les échantillons d'hippocampe sont incubés 5 min à température ambiante avec l'association thiocyanate de guanidium/phénol-chloroforme, puis 3 min avec du chloroforme. Les échantillons sont ensuite centrifugés à 12 000 g pendant 15 min à 4°C afin de favoriser la séparation des phases. Les protéines dénaturées s'accumulent dans la phase organique phénol-chloroforme alors que l'ARN est dans la phase aqueuse. Un volume d'éthanol 70% est ajouté à la phase aqueuse contenant les ARN totaux afin de les précipiter. L'échantillon est ensuite transféré sur colonne RNeasy Mini spin, puis lavé avec du tampon RW1, contenant un sel guanidine et de l'éthanol.

Les ARN de plus de 200 pb restent accrochés à la colonne. Par la suite, les échantillons sont lavés deux fois avec du tampon de lavage RPE, contenant de l'éthanol, qui permet d'éliminer les traces de sels. Les ARN sont resuspendus dans de l'eau RNase-free et une librairie est préparée avec le kit « TruSeq stranded total RNA library prep Gold (48 samples) » (Référence 20020598, Illumina) avant d'être séquencée.

Les expériences de RNA-seq ont impliqué au total 52 souris :

- 6 souris *Hdac2*^{+/+} âgées de 3 à 6 mois (5 femelles et 1 mâle),
- 3 souris *Hdac2*^{-/-} âgées de 3 à 5 mois (2 femelles et 1 mâle),
- 5 souris femelles 3xTg-AD ; *Hdac2*^{+/+} âgées de 6 à 9 mois,
- 6 souris femelles 3xTg-AD ; *Hdac2*^{-/-} âgées de 6 à 9 mois,
- 8 souris femelles contrôle (C57) et 8 souris femelles *App*^{tm3 (NL-G-F)} âgées de 3 mois,
- 8 souris femelles contrôle (C57) et 8 souris femelles *App*^{tm3 (NL-G-F)} âgées de 9 mois.

7) Expérience de ChIP-seq séquentiel : tandem-ChIP

Dans le but de cartographier la liaison de HDAC2 au niveau des promoteurs de gènes cibles, des souris *Hdac2*^{HF/+} ont été disséquées, les tissus ont été fixés au formaldéhyde, puis la chromatine a été préparée. Un ChIP-tandem « classique » a été réalisé sur cette chromatine avec une première IP utilisant des billes couplées à un anticorps anti-Flag, puis une élution avec le peptide FLAG, et une seconde IP utilisant des billes couplées à un anticorps anti-HA.

Partie B : analyses bioinformatiques

I. Indications sur la reproductibilité bioinformatique de nos données

Les données issues du séquençage produites au cours de ma thèse ont dans la très grande majorité été analysées par l'ingénieure bioinformaticienne de mon équipe, Emilie Drouineau. Grâce à différentes formations suivies tout au long de ma thèse en bioinformatique et grâce au Diplôme Universitaire en Bioinformatique intégrative que j'ai obtenu en 2020, je suis capable de comprendre chaque étape des analyses bioinformatiques réalisées pendant ma thèse. Je comprends également la fonction de chacun des outils utilisés et je suis capable de créer des scripts et des pipelines pour des analyses de CHIP-seq et de RNA-seq.

Les analyses bioinformatiques menées pendant ma thèse respectent autant que possible le principe du FAIR en bioinformatique. Le FAIR repose sur quatre qualités : Facile à trouver, Accessible, Interopérable et Réutilisable.

En tant que biologiste ayant été un peu formée à la bioinformatique, je me suis astreinte tout au long de ma thèse à avoir des fichiers de suivi les plus complets et détaillés possibles concernant chaque échantillon de nos données afin d'avoir des métadonnées et des fichiers de *design* complets et faciles à retrouver.

Concernant l'accessibilité de nos données, elles sont hébergées sur différents supports au CEA et les données nécessaires à la reproductibilité seront rendues accessibles lors de la publication de l'article.

Afin d'être reproductible et de limiter les problèmes liés à l'environnement, le gestionnaire de paquets Conda a été utilisé (lien : <https://docs.conda.io/en/latest/>). Il permet de gérer le parc logiciel qui nous est nécessaire en simplifiant l'installation et l'utilisation des outils et d'utiliser toujours les mêmes versions d'outils que ce soit sur le cluster de l'I2BC ou sur nos ordinateurs. Les grandes étapes de nos analyses ont été réalisées grâce à des *pipelines* (chaîne de traitement de calcul) Snakemake (lien : <https://snakemake.readthedocs.io/en/stable/>, <https://academic.oup.com/bioinformatics/article-abstract/28/19/2520/290322>). Ce dernier est un gestionnaire de flux (*workflow*) qui permet de définir simplement les différentes étapes

au sein des analyses, de plus facilement paralléliser des tâches et de reproduire fidèlement les différentes étapes de l'analyse avec les mêmes paramètres définis dans notre pipeline pour un type de données précis. Cet outil, quand il est utilisé conjointement avec Conda, permet d'optimiser la reproductibilité.

Des *notebooks*, qui sont des équivalents de cahier de laboratoire bioinformatique en langage Markdown ou RMarkdown, ont été utilisés pour décrire le protocole bioinformatique, insérer des morceaux de code, discuter des résultats et exporter des rapports dans différents formats.

II. Analyses bioinformatiques des données de ChIP-seq

Nos analyses de MNase-ChIP-seq, de ChIP-seq et de ChIP-tandem reposent sur une succession d'étapes similaires comprenant un contrôle qualité des données, l'alignement, différents processus post-alignement et des représentations graphiques.

1) Contrôle qualité des données et nettoyage

Une fois les données de séquençage récupérées, la conformité à notre demande mais aussi la qualité des lectures (*reads*) suffisante pour nos analyses doivent être vérifiées.

Différents contrôles ont été réalisés, allant de la taille de la librairie au nombre de duplicat. Pour ce faire deux logiciels ont été utilisés. Le premier logiciel est FastQC (version fastqc=0.11.9, paramètres par défaut). Il fournit des statistiques basiques sur les *reads* séquencés (nombre de reads, longueur, pourcentage en GC...), la qualité de séquençage pour chaque base, le contenu en GC sur la totalité du *read*, le contenu en base N que le séquenceur n'a pas pu identifier, le niveau de duplicat, les séquences surreprésentées et la contamination par les adaptateurs.

Le deuxième logiciel, fastqScreen (version fastq-screen=0.14.0, paramètres : `--aligner bowtie2`) permet de rechercher des contaminations possibles des *reads* par d'autres génomes ou par des rRNA (Wingett et Andrews 2018). Les *reads* sont alignés de manière indifférente sur différents génomes ou transcriptomes (par exemple : l'Homme, la souris, la bactérie *Escherichia coli*, la levure *Saccaromyces cerevisiae*, le phage PhiX, le blé...). Ainsi, la spécificité de nos *reads* sur le génome de la souris peut être vérifiée. Plus la spécificité à notre organisme de référence est grande, plus la confiance en nos données est importante. Le phage

PhiX est intégré car il est utilisé pour calibrer les séquenceurs, et est donc souvent retrouvé à l'état de trace. Si une quantité non négligeable de *reads* s'alignent sur d'autres génomes, cela veut dire qu'il y a contamination et que ces *reads* sont à supprimer.

Selon leur qualité, les *reads* ont ensuite été plus ou moins nettoyés avec l'outil de trimming `fastp` (version `fastp=0.20.1`, paramètres : `--length_required 25 --detect_adapter_for_pe --cut_front --cut_tail --cut_window_size 6 --cut_mean_quality 10 --unqualified_percent_limit 40 --n_base_limit 7 --average_qual 0 --overrepresentation_analysis --trim_poly_g`) (S. Chen et al. 2018). L'outil `fastp` élimine les *reads* de mauvaise qualité, tronque les *reads* en 5' et en 3' si la qualité est insuffisante, élimine automatiquement les adaptateurs de séquençage des *reads* et fournit un rapport.

2) L'alignement des *reads* sur le génome murin

L'étape suivant le contrôle qualité des *reads* est l'alignement des *reads* nettoyés sur le génome de référence de la souris mm9. Dans la première étape, le génome de référence de la souris mm9 a été indexé par l'outil `Bowtie 2` (version `bowtie2=2.4.2`, paramètres par défaut) (Langmead et Salzberg 2012). Puis, les *reads paired-end* nettoyés sont alignés sur le génome indexé également par l'outil `Bowtie 2` (version `bowtie2=2.4.2`, paramètres par défaut). `Bowtie 2` est un outil qui aligne les *reads* sur les génomes mammifères de façon rapide et efficace. Par défaut, il aligne les *reads* de « bout-en-bout » (*end-to-end*) en impliquant le plus de bases possibles du *read*. Les fichiers comportant les *reads* alignés sur le génome de la souris sont au format `.bam`. La qualité de l'alignement est vérifiée par l'outil `Samtools flagstat` (version `samtools=1.11`, paramètres par défaut) qui utilise en entrée les fichiers `.bam` et qui fournit des informations sur le nombre total de *reads*, le nombre d'alignements uniques ou multiples, le nombre de *reads* où les deux paires sont correctement alignés, etc (Heng Li et al. 2009).

3) Les différents processus post-mapping

Après l'alignement, de nombreuses opérations sur les fichiers sont réalisées afin de pouvoir analyser et visualiser les résultats. Il s'agit des processus post-mapping.

Après alignement, les fichiers .bam sont triés en fonction de leurs coordonnées génomiques à l'aide de Samtools sort (version samtools=1.11, paramètres par défaut). Par exemple, deux *reads* ayant les même coordonnées génomiques sont écrits l'un à la suite de l'autre dans le fichier .bam trié.

Les fichiers .bam contenant les *reads* triés et alignés ont été nettoyés des *reads* dupliqués. L'outil Picard Markduplicates (version picard=2.23.8, paramètres : `CREATE_INDEX=true` `VALIDATION_STRINGENCY=STRICT`) permet de détecter et de marquer les *reads* dupliqués des fichiers .bam de sortie ("Picard Toolkit." 2019. *Broad Institute, GitHub Repository*. <http://broadinstitute.github.io/picard/>; *Broad Institute*). Ces *reads* dupliqués marqués et les *reads* avec une faible qualité d'alignement sont éliminés par l'outil Samtools view (version samtools=1.11, paramètres : `-h -f 1 -F 3852 -q 30`). Le fichier de sortie est un fichier .bam nettoyé, trié et vierge de tout duplicat.

Plusieurs tests de qualité des données et plusieurs analyses sont possibles à partir de ce fichier .bam.

Tests de qualité des fichiers .bam :

L'efficacité de l'élimination des *reads* dupliqués dans le fichier de sortie, ainsi que la qualité de leur alignement peuvent être vérifiées grâce à Samtools flagstat (version samtools=1.11, paramètres : par défaut).

Calcul de la couverture génomique :

La profondeur de séquençage correspond au produit de la taille moyenne des *reads* par le nombre de *reads*, divisé par la taille du génome de référence. La couverture génomique, qui est directement corrélée à la profondeur de séquençage, correspond à la partie du génome réellement couverte par les *reads*. La couverture génomique est calculée grâce à l'outil Deeptools (version deeptools=3.5.0, paramètres : `--outFileFormat bigwig --binSize 1 --numberOfProcessors 5 --effectiveGenomeSize 2652783500 --normalizeUsing RPKM --MNase --minFragmentLength 10`

```
--extendReads --centerReads --maxFragmentLength 500 --  
ignoreForNormalization chrX chrY chrM --blackListFileName  
gencode_mm9/mm9-blacklist.bed.gz) (Ramírez et al. 2016).
```

Peak-calling :

Le fichier .bam nettoyé est utilisé pour faire du *peak-calling* grâce à l'outil MACS2 callpeak (version macs2=2.2.7.1, paramètres : `-f BAMPE -gsize 1.87e9`) (Y. Zhang et al. 2008).

Le *peak-calling* identifie les régions spécifiquement enrichies en *reads* suite aux ChIP-seq.

Vérification de la qualité des peaks obtenus :

L'outil phantompeakqualtools (version phantompeakqualtools=1.1, paramètres par défaut) permet de discriminer le *peak* correspondant à l'enrichissement en *reads* à une position génomique et le *peak* correspondant à la longueur du *read*.

Transition d'un fichier contenant des reads pairés en un fichier contenant des fragments :

Un fichier .bam, fichier contenant des *reads* pairés, est converti en un fichier .bed contenant des fragments d'ADN grâce à un script de la bioinformaticienne utilisant l'outil Bedtools bamtobed (version bedtools=2.29.2, paramètres : `-bedpe`) (Quinlan et Hall 2010).

Normalisation des données avec Deeptools :

Afin que nos échantillons soient comparables entre eux et pour réaliser des clichés (*heatmaps* ou IGV), une étape de normalisation (en RPKM) des données par rapport à la profondeur de séquençage est réalisée avec l'outil Deeptools (version deeptools=3.5.0, paramètres : `--binSize 10 --effectiveGenomeSize 2652783500 --normalizeUsing RPKM --enxtendReads --centerReads --maxFragmentLength 500 -ignore chrX chrY chrM --blackListFileName mm9-blacklist.bed.gz`)

Nos paramètres sont adaptés aux données traitées par la MNase grâce à l'utilisation des centres des fragments.

Détection des motifs avec HOMER :

Tout d'abord, les fichiers d'alignement des quatre échantillons Input contrôles et des quatre échantillons Input déplétés en BRG1 pour les groupes de fraction 2-3-4 ont été fusionnés grâce

à samtools merge (version samtools=1.10, utilisant htlib=1.10.2, paramètres par défaut). Du peak-calling sur les fichiers bam a été réalisé avec MACS2 callpeak (version macs2=2.2.7.1, `f BAMPE -g mm --min-length 40`). Sur les fichiers de peaks, les motifs ont été détectés avec l'outil HOMER findMotifsGenome.pl (version homer=4.11, paramètres : `findMotifsGenome.pl mm9 -size given`). Dans les fichiers de sortie de HOMER, j'ai choisi 31 motifs dont la p-value est significative d'après HOMER ($p\text{-value} < 10^{-50}$) dans tous les échantillons contrôles.

Les fichiers sont ensuite chargés dans R. Grâce à un script R que j'ai créé, le nombre de motifs CTCF pour l'échantillon contrôle est divisé par le nombre de motifs CTCF pour l'échantillon déplété en BRG1. Nous avons normalisé par le nombre de motifs CTCF car nous avons fait l'hypothèse que la liaison du facteur CTCF n'est pas impactée par la perte du remodelleur BRG1 (d'après des observations issues de la littérature : (Barisic et al. 2019)). Le coefficient de normalisation est appliqué aux 30 autres motifs dans l'échantillon déplété en BRG1. Par la suite, le ratio entre le nombre de séquences cibles déplétées en BRG1 et le nombre de séquences cibles contrôles est calculé. Une *heatmap* est ensuite produite pour visualiser les résultats.

4) La visualisation des résultats

Les fichiers .bed sont nécessaires pour la visualisation des résultats sous forme de graphe en deux dimensions (2D).

Les *heatmaps* et profils moyens sont obtenus en utilisant la suite de programme seqMINER (Ye et al. 2011a). Selon les paramètres définis, seqMINER divise un intervalle génomique en fenêtres dans lesquelles il va compter le nombre de *reads*. Un score est associé au nombre de *reads* dans chaque fenêtre et une matrice est générée. La matrice associe les fenêtres du génomes (les colonnes) à un élément précis sur le génome, par exemple des enhancers (les lignes). Plus le nombre de *reads* dans une fenêtre est élevé et plus l'intensité de la couleur sur la *heatmap* est élevée (Ye et al. 2011a).

Les résultats peuvent également être visualisés sous la forme de profils IGV (version 2.4.9). Sous IGV, le profil de couverture (via un fichier bigwig) et la localisation des *peaks* (via le fichier bed contenant la position des *peaks*) sur le génome sont visualisables.

Enfin notre bioinformaticienne a développé une représentation des résultats, un histogramme 2D, basée sur la représentation du Vplot du laboratoire du Dr Henikoff (Henikoff et al. 2011). Sur l'histogramme 2D, l'axe Y correspond à la taille des fragments d'ADN en pb, l'axe X correspond à la position relative des fragments d'ADN sur le génome. Chaque point correspond aux coordonnées du centre d'un fragment d'ADN et l'intensité de la couleur des points dépend du nombre de fragments : plus il y a de fragments à un emplacement donné et plus la couleur rouge est intense.

III. Analyses bioinformatiques des données de RNA-seq

Nos analyses de RNA-seq reposent sur différentes étapes comprenant un contrôle qualité et un nettoyage des données, l'alignement, l'estimation de l'expression des gènes et des représentations graphiques. Certaines étapes sont communes aux analyses de ChIP-seq.

1) Contrôle qualité des données

Comme pour les analyses de ChIP-seq, le contrôle qualité des données de séquençage a été réalisé avec les outils fastQC et fastqScreen et le nettoyage des *reads* avec l'outil fastp.

Une contamination assez importante de nos échantillons par des ARN ribosomaux (ARNr) a été détectée. Les *reads* alignés sur les ARNr ont dû être éliminés.

2) Alignement sur le génome et élimination des ARN ribosomaux

Les *reads* nettoyés sont alignés sur le génome de référence de la souris mm10 préalablement indexé grâce à l'outil STAR (version star=2.7.3a, paramètres : `sjdbOverhang 149`) (Dobin et al. 2013). Les fichiers de sortie, au format .bam, contiennent les *reads* alignés sur le génome

dont la qualité peut être contrôlée grâce à l'outil Samtools flagstat (version samtools=1.10, utilisant htslib=1.10.2 ; paramètres par défaut). Grâce à l'outil bedtools intersect (version bedtools=2.29.2 ; paramètres par défaut) et au fichier contenant les annotations du génome au format .gtf, les ARNr ont été éliminés du fichier de sortie contenant uniquement les *reads* alignés sur le génome ailleurs qu'au niveau des gènes codant les ARNr. Ces fichiers ont ensuite été reconvertis en fichier fastq nettoyés des ARNr grâce à samtools fastq.

3) Vérification du génotype de nos souris

Bien que tous les génotypes de nos souris aient été préalablement vérifiés expérimentalement par qPCR, nous avons tout de même souhaité les valider par des méthodes bioinformatiques. Tout d'abord, nous avons vérifié que toutes les souris des lignées 3xTg-AD et 3xTg-AD ; *Hdac2*^{-/-} utilisées en transcriptomique possédaient bien la mutation caractéristique du modèle 3xTg-AD : une mutation de la Méthionine en Valine en position 146 de la séquence protéique de la préséniline.

L'outil samtools mpileup (version samtools=1.1 , paramètres : `--count-orphans --no-BAQ --positions psen1.bed`) a détecté les bases mutées au niveau du gène *Psen1* entre les *reads* qui s'alignent et le génome de référence. L'outil de détection de variant VarScan2 (version VarScan2=2.4.4 ; paramètres : `--p-value 0.05 --variants --output-vcf v`) a jugé probable les différentes mutations en considérant le risque alpha de 0,05 (Koboldt et al. 2012). VarScan2 a fourni un fichier de sortie contenant les bases mutées, la nature de la mutation (dans le cas d'une mutation Met vers Val dans la séquence protéique, il s'agit d'une mutation A vers G dans la séquence nucléotidique), et sa fréquence. Tous nos *reads* des lignées de souris d'intérêt comportaient bien la mutation A vers G avec une p-value significative.

L'outil SnpEff (version 5.0e (build 2021-03-09 06:01), by Pablo Cingolani; paramètres par défaut) a traité le fichier fourni par Varscan2. En étudiant le contexte de la mutation, les connaissances sur cette mutation, les mutations possibles d'acides aminés et les modifications de fonction connues de la protéine, SnpEff a bien annoté la mutation de Met en Val en position 146 de la Préséniline.

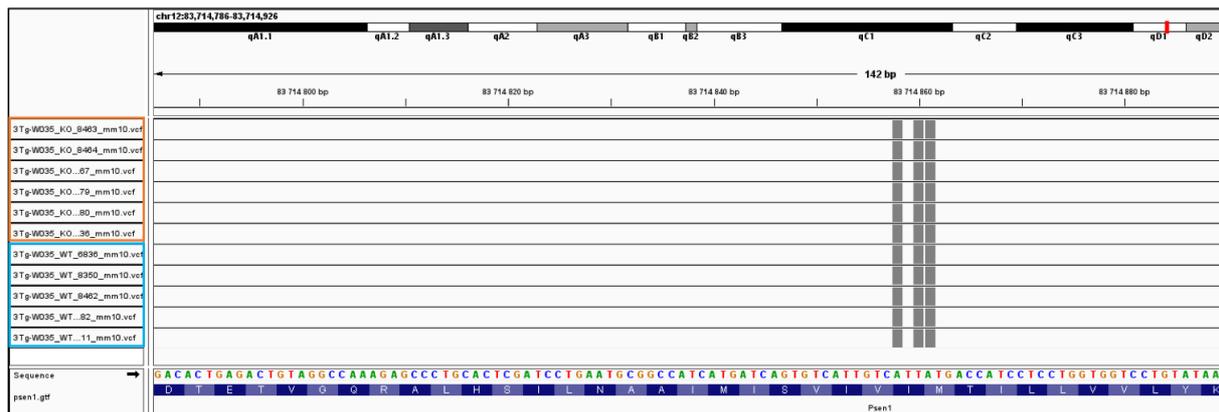


Figure 17 : Vérification de la présence de la mutation du gène *Psen1* dans les souris modèles 3xTg-AD et 3xTg-AD ; *Hdac2*^{-/-}

Mutation de la Méthionine en Valine en position 146 de la séquence protéique du gène de la préséniline (*Psen1*) dans les échantillons 3xTg-AD (rectangle bleu) et 3xTg-AD ; *Hdac2*^{-/-} (rectangle orange). Mutation représentée par les rectangles gris.

Nous avons également voulu vérifier que toutes les souris des lignées *Hdac2*^{-/-} et 3xTg-AD-*Hdac2*^{-/-} qui ont été utilisées en transcriptomique possèdent bien des transcrits tronqués de *Hdac2*. Pour cela, les fichiers .bam et le fichier d'annotation du gène *Hdac2* (fichier .gtf encode mm10) ont été chargés dans l'outil IGV qui a affiché sur paramétrage les quatre transcrits de *Hdac2*, ainsi que les jonctions entre exons. Nous avons bien observé que les transcrits sont tronqués après l'exon 7 (Figure 23).

4) Pseudo-mapping ou mapping et comptage de l'expression des gènes

Pour traiter les nouveau fichiers fastq nettoyés des ARNr, l'outil Salmon (version salmon=0.14 ; paramètres Salmon quant : `--numBootstraps 100 --validateMappings --gcBias --seqBias --posBias`) a été choisi pour les échantillons 3xTg-AD et dérivés (Patro et al. 2017). Les échantillons NLGF et leurs contrôles ont été alignés avec l'outil STAR (version star=2.7.3a, paramètres : `sjdbOverhang 149`). Tous les *reads* ont été alignés sur la version gencode mm10 du génome de la souris. L'outil Salmon fait du *pseudo-mapping* des *reads* sur les différents isoformes des ARNm plutôt que sur les gènes. En effet, les génomes eucaryotes comportent des introns et subissent l'épissage qui produit différents isoformes de transcrits pour un gène donné. L'outil STAR fait du *mapping* de *reads* sur le génome de la souris.

La quantification de l'expression des gènes a été faite par Salmon et le calcul du différentiel d'expression des gènes a été réalisé par DESeq2 (version DESeq2=1.32.0 ; paramètres par défaut) (Love, Huber, et Anders 2014).

Dans le cas d'un alignement avec STAR, l'outil featureCounts est nécessaire pour quantifier l'expression des gènes (version featureCounts=v2.0.0, paramètres : `-p -s 2`) (Liao, Smyth, et Shi 2014). L'analyse du différentiel d'expression des gènes entre deux conditions est ensuite réalisée grâce au package R SARTools (versions r-sartools=1.7.2 r-base=3.6.3 ; paramètres par défaut) utilisant l'outil DESeq2 et est présentée sous forme de différents graphes (Varet et al. 2016).

Les analyses de *Gene Ontology* réalisées à l'aide du package R clusterProfiler (version 3.14.3, R version 3.6.3 (2020-02-29)) et de scripts R sont présentées sous forme de barplot (Yu et al. 2012).

Résultats

Partie A : Comment les remodeleurs de la chromatine régulent l'organisation nucléosomale au niveau des enhancers en cellules ES de souris

- I. Article Nocente *et al.* : “The Brg1 ATPase generates subnucleosomal particles targeted by pioneer transcription factors at enhancers”

The Brg1 ATPase generates subnucleosomal particles targeted by pioneer transcription factors at enhancers

Marina Nocente^{1,3}, Anida Mesihovic^{1,3}, Emilie Drouineau^{1,3}, Cécile Pignat², Florence Ribierre², Hélène Picaud¹, Françoise Ochsenbein¹, Sophie Chantalat², Matthieu Gérard¹

1. Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France
2. Institut de biologie François Jacob, Centre National de Recherche en Génomique Humaine
3. These authors contributed equally

e-mail: matthieu.gerard@i2bc.paris-saclay.fr

Abstract

The mammalian SWI/SNF complexes (mSWI/SNF) are ATP-dependent chromatin remodelling factors with tumour suppressor properties. mSWI/SNF were shown to be essential for chromatin opening at enhancers, but the detailed mechanism of this activity remains unclear. We show in mouse embryonic stem (ES) cells that Brg1 (Smarca4), the catalytic subunit of mSWI/SNF, is required *in vivo* for the generation of subnucleosomal particles that protect genomic DNA fragments of 50-80 bp in length from micrococcal nuclease (MNase) digestion. These particles, which contain all four histones H3, H4, H2B, H2A, sediment in sucrose gradients with an apparent molecular weight markedly inferior to that of canonical nucleosomes. Our refined analysis of enhancer nucleosomal and subnucleosomal organization suggests that these particles arise from the remodelling of fragile nucleosomes into hemisomes. We demonstrate that, *in vivo*, these subnucleosomal particles are targeted by Oct4 and other pioneer transcription factors more efficiently than the canonical nucleosomes present within the enhancer. Moreover, we show that Oct4 enrichment onto these subnucleosomal particles is an order of magnitude above its enrichment onto accessible, histone-free DNA. We propose that a critical activity of mSWI/SNF at enhancers is to generate subnucleosomal particles that act as anchors for pioneer and other transcription factors during ES cell self-renewal.

In mES cells, enhancers are bound by a series of pluripotency-associated transcription factors (TFs) including Oct4 (Pou5f1), Sox2, Nanog and Esrrb (Chen et al. 2008). The DNA binding sites for these TFs are located in relative proximity to each other, without any constrained distribution pattern (Spitz et Furlong 2012; Rao, Ahmad, et Ramachandran 2021; Sönmezer et al. 2021). Enhancers are also characterized by their hypersensitivity to DNase I, the presence of the mediator complex, and, for active enhancers, RNA pol2 and TBP, as well as H3K27ac (Kagey et al. 2010; Koch et al. 2011; Thurman et al. 2012). Some elements of nucleosomal organization at enhancers could be defined, yet their chromatin organization is less defined than that of promoters and CTCF binding sites (Teif et al. 2012; Carone et al. 2014; Ishii, Kadonaga, et Ren 2015; Iwafuchi-Doi et al. 2016; Voong et al. 2016). In this work, we investigated nucleosomal organization at enhancers using MNase-seq combined with chromatin immunoprecipitation (ChIP) with antibodies against histones. Formaldehyde-fixed ES cells were permeabilized and incubated with MNase, using mild digestion conditions which allowed us to examine the distribution of canonical and fragile nucleosomes, as well as subnucleosomal particles. K-means clustering of enhancers based on the distribution patterns of both, canonical nucleosomes and subnucleosomal particles, led to the identification of three groups of enhancers having distinct but highly related nucleosomal organization (Extended Data Fig. 1).

To visualize nucleosomal organization, we used the V-plot representation (Henikoff et al. 2011), which shows in Figure 1 the distribution and density of DNA fragments immunoprecipitated with various antibodies as a function of the position of their mid-point on the genome (x axis), and their length (y axis). In the three groups of enhancers, we detected in the central region several clusters of short (50-120 bp) DNA fragments associated with histone H3, which revealed the presence of subnucleosomal particles precisely positioned across the enhancers (Fig. 1 and Extended Data Figs. 1, 2). This central region coincides in position with the peak of hypersensitivity to DNase I (Thurman et al. 2012). These subnucleosomal particles alternate in position with one or more full size nucleosomes (i.e. histones associated with 150-180 bp long DNA fragments), their number depending on the distance that separates the two well positioned nucleosomes that flank the centre of the enhancer (Fig. 1a and Extended Data Figs. 1, 2a). Enhancers groups 1 to 3 differ from each other according to the distance between these two flanking nucleosomes (defined hereafter

as -1 and +1 nucleosomes), with group 1 having the narrowest (350-380 bp) and group 3 the widest (540-630 bp) distance (Extended Data Fig. 1). In enhancer group 1, a single centrally located nucleosome is present between -1 and +1 nucleosomes (Fig. 1a). This central nucleosome is highly sensitive to MNase dosage, as it completely disappears when ES cells were incubated with a four-fold excess of MNase (Fig. 1b), and we will thereafter refer to this central particle as a fragile nucleosome (Iwafuchi-Doi et al. 2016). In contrast, the flanking nucleosomes -1 and +1 were only slightly affected by the high-dose MNase treatment, revealing their canonical nature. Enhancer group 2 is characterized by the presence of fragile nucleosomes at two main positions instead of one in the central interval between nucleosomes -1 and +1 (Extended Data Fig. 2a, b). We observed that in group 1, the single central fragile nucleosome is symmetrically flanked on each side of its dyad axis by a subnucleosomal particle protecting from MNase digestion a minimum DNA length of 54 bp, as assessed by the length of the DNA fragments present at the vertex of the “V” shape signal (Fig. 1a). In group 2, each of the two fragile nucleosomes of the central interval are flanked by subnucleosomal particles in a pattern reminiscent of group 1 (Extended Data Fig. 2a). Enhancer group 3 shows a variation of these patterns, with a wider central interval accommodating a larger number of positioned subnucleosomal particles (Extended Data Fig. 1). Examination of individual enhancers not belonging to the three groups isolated by k-means clustering confirmed their similar organization in combinations of fragile nucleosomes and subnucleosomal particles, flanked by canonical nucleosomes (data not shown). Thus, our analysis revealed that enhancers have a modular chromatin organization in which two positioned canonical nucleosomes flank a central region containing a variable number of basic modules, composed each of a fragile nucleosome flanked on both sides of its dyad axis by subnucleosomal particles. In the basic module, the fragile nucleosome and the two subnucleosomal particles occupy precisely the same genomic interval, and thus are likely alternating chromatin configurations.

We next examined the distribution of TFs and epigenetic features that contribute to enhancer function, using the same MNase ChIP-seq protocol described above for mapping nucleosomes (Fig. 1f-l and Extended Data Fig. 2f-l). ChIP-seq enrichment for all the tested TFs revealed that they are mainly bound within the central interval of the enhancer in which subnucleosomal

particles are intermixed with fragile nucleosome(s). Even the TFs known to have pioneer functions, such as Oct4 and Sox2, were preferentially enriched onto DNA fragments spanning 50 to 120 bp in length rather than the 150-180 bp typical of canonical nucleosomes (Fig. 1f, g and Extended Data Fig. 2f, g). Intriguingly, the subnucleosomal particles were not acetylated at H3K27, in contrast to their flanking nucleosomes -1 and +1 (Fig. 1j and Extended Data Fig. 2j). General transcription factors, as well RNA polymerase II, were also detected within the same central interval of the enhancer, associated with 40-120 bp DNA fragments (Figure 1i, l and Extended Data Fig. 2i, l). In contrast, CTCF was not bound at any location within the three groups of enhancers analysed, thus excluding the possibility that it might contribute to organize nucleosome positioning (data not shown).

We afterwards wondered how this enhancer-specific nucleosomal organization is controlled in mES cells. Prime candidates for the regulation of this organization are ATP-dependent chromatin remodellers (Narlikar, Sundaramoorthy, et Owen-Hughes 2013; de Dieuleveult et al. 2016; Clapier et al. 2017; Markert et Luger 2021). In particular, the Brg1 (Smarca4) ATPase of the mammalian SWI/SNF complexes was shown in several recent studies to be required for the generation of accessible chromatin at enhancers (Schick et al. 2021; Iurlaro et al. 2021). The nature of this accessible chromatin remains unclear, and we thus tested whether Brg1 might be required for the establishment of the well-ordered nucleosomal and subnucleosomal organization at enhancers groups 1 and 2. We first tested this hypothesis by depleting Brg1 using an shRNA targeting its mRNA, and performing chromatin immunoprecipitation with an antibody against histone H3 followed by sequencing (ChIP H3 MNase-seq). Extensive depletion of Brg1 resulted in striking alterations of both nucleosomal and subnucleosomal organization at enhancers. First, the subnucleosomal particles were almost completely absent at enhancers in cells depleted of Brg1, compared to control cells (Fig. 1d and Extended Data Fig. 2d). Second, the positioning of canonical nucleosomes was severely altered, resulting in a random redistribution of canonical nucleosomes across the enhancer. All these alterations were reproduced with a second shRNA targeting a different region of Brg1 mRNA (Extended Data Fig. 3a). Moreover, identical defects in nucleosomal organization were also detected when we used an antibody against histone H2B to immunoprecipitate the chromatin (Extended Data Fig. 3a). In these experiments with shRNA vectors, we analysed chromatin organization after a relatively long term (48h) depletion of Brg1, which might result in

secondary effects. We thus turned to the auxin-induced degradation (AID) system to test the function of Brg1 at enhancers with an experimental system compatible with short term analysis (Natsume et al. 2016). Depletion of Brg1 using the AID system resulted in similar defects in chromatin organization as observed with the shRNA vectors (Fig. 1e and Extended Data Fig. 2e). The perturbations were slightly less extensive than with shRNAs, which is likely due to the remaining presence of low amounts of Brg1 using the AID system, compared to the more extensive depletion we obtained with shRNAs (Extended Data Fig. 4). Importantly, alterations in chromatin organization at enhancers were almost identical after 3h or 20h of Brg1 depletion, showing that the defects we observed in nucleosomal organization at enhancers with either shRNAs or the AID system are due to Brg1 depletion and not due to secondary effects (Extended Data Fig. 3).

We then examined how Brg1 binds to nucleosomes and subnucleosomal particles at enhancers using MNase-ChIP-seq. This experiment revealed that Brg1 is enriched both at the level of the centrally located fragile nucleosome(s) in enhancers group 1 and 2, and at the level of the subnucleosomal particles (Fig. 1c and Extended Data Fig. 2c). This binding pattern is compatible with a scenario in which Brg1 binds the central nucleosome to convert it into a fragile nucleosome, and into the smaller subnucleosomal particles flanking the central nucleosome. We next tested the possibility that other chromatin remodellers might be involved in the control of nucleosomal and subnucleosomal organization at enhancers. Using the same experimental approach, we tested how depletion of Chd1, Chd4, or Smarca5 could affect chromatin organization (Extended Data Fig. 5). Depletion of these remodelers in ES cells did not affect the generation of subnucleosomal particles at enhancers (Extended Data Fig. 5). We concluded from these experiments that Brg1 is essential for the establishment of the highly specific nucleosomal architecture of enhancers, which involves the generation of subnucleosomal particles and the control of the positioning of both canonical and fragile nucleosomes, as well as of subnucleosomes.

In the next step, we investigated the nature of the subnucleosomal particles present at enhancers. We made the hypothesis that Brg1 activity could convert the central fragile nucleosome(s) into smaller particles, such as hexasomes [(H3-H4)₂; H2-H2B] or hemisomes containing one copy of each of the histones H3, H4, H2A, H3B (Zlatanova et al. 2009). However, it was also possible that Brg1 could alter the structure of the central fragile

nucleosome in a way, resulting in the extrusion of a DNA loop in the region of the dyad axis, which would be highly sensitive to MNase digestion, and yield a histone octasome associated with two short DNA fragments instead of the 150-180 bp fragment. To test these two non-exclusive hypotheses, we set up an experimental strategy based on the separation of nucleosomal and subnucleosomal particles by centrifugation through sucrose gradients (Fig. 2). Chromatin was prepared from either control or Brg1-depleted cells, using the AID system described above, and loaded onto a 10-30% sucrose gradient. In a sucrose gradient, nucleosomal and subnucleosomal particles sediment as a function of their molecular weight and shape. After centrifugation, individual fractions were collected and first tested onto an agarose gel. This analysis confirmed that we could isolate from the gradient a series of genomic DNA fragments ranging from 30 to 120 bp, which sedimented more slowly than canonical nucleosomes (Fig. 2a). To identify which of these fractions contained subnucleosomal particles, we realized ChIP-seq experiments with an antibody against histone H3 (Fig. 2b). We found that the 50-80 bp-long subnucleosomal particles in gradient fractions 5-6, which sedimented much slower than canonical mononucleosomes (fractions 11-13) were associated with histone H3 at group 1 enhancers (Fig. 2b). In contrast, the particles that sedimented in fractions 3-4 were not efficiently immunoprecipitated, suggesting that they contain a population of DNA fragments devoid of histone H3. Subnucleosomal particles associated with increasing DNA lengths were detected in gradient fractions 7-8 and 9-10. The fragile nucleosomes located in the central region of the enhancers sedimented in fractions 9-10, revealing sedimentation properties slightly distinct from those of canonical nucleosomes (Fig. 2b).

We further investigated the nature of the subnucleosomal particles associated with 50-80 bp DNA fragments that sedimented in fractions 5-6. Their slow sedimentation properties, compared to canonical mononucleosomes, is not compatible with the hypothesis that these particles correspond to octasomes associated with two 50-80 bp DNA fragments resulting from the digestion of nucleosomal DNA at the level of the dyad, because such particles would have approximately the same molecular weight as canonical nucleosomes and would thus be expected to have similar sedimentation properties. To define the histone content of the 50-80 bp subnucleosomal particles of fractions 5-6, we performed ChIP-seq experiments with antibodies against the remaining three histones H4, H2A and H2B. We found that all of them

efficiently immunoprecipitated the 50-80 bp particles present in fractions 5-6, but not the particles sedimenting slower in gradient fractions 2-3-4 (Fig. 3). This result excludes the possibility that the 50-80 bp particles might correspond to tetrasomes composed of two dimers of histone H3 and H4, which are nucleosome assembly intermediates detected during DNA replication (Clément et Almouzni 2015). The presence of the four histone suggests that these particles correspond to either hexasomes or hemisomes (Zlatanova et al. 2009; Ramachandran, Ahmad, et Henikoff 2017). However, the length of the DNA protected from MNase digestion by these particles is difficult to reconcile with the hexasome hypothesis, for which MNase protection yielded DNA fragments ranging from 90 to 130 pb (Ramachandran, Ahmad, et Henikoff 2017). In our experimental setting, the subnucleosomal particles associated with 90-130 bp DNA fragments were detected in fractions 7-8 and 9-10, suggesting a histone content different from the 50-80 bp particles that sedimented in fractions 5-6. Furthermore, the positions of the DNA fragments centres of these 50-80 bp particles are symmetrically distributed on each side of the dyad axis of the central fragile nucleosome, and the distance spanning the sum of the lengths of the two consecutive particles (140-180 bp) is equivalent to the length of DNA wrapped around an octasome. This striking positioning of the 50-80 bp particles, their histone content, and their low-speed sedimentation properties suggest that these particles correspond to hemisomes, generated from Brg1-mediated remodelling events targeting the central fragile nucleosome (Extended Data Fig. 11).

We next tested which of the different categories of subnucleosomal particles detected in the sucrose gradient experiments were dependent on Brg1 activity. We repeated the sucrose gradient sedimentation experiments using chromatin prepared from either Brg1-depleted or control cells, and analysed the distribution and abundance of subnucleosomal particles. ChIP-seq experiments with antibodies against each of the four histones H3, H4, H2A and H2B revealed that Brg1 depletion results in a strong reduction of 50-80 bp subnucleosomal particles at group 1 enhancers, thus confirming that their production is dependent on Brg1 (Fig. 3). In addition, the amount of 90-130 bp subnucleosomal particles was also reduced in Brg1-depleted cells, showing that Brg1 activity is required for the generation of all types of subnucleosomal particles at enhancers (Extended Data Fig. 6).

We also analysed how Brg1 depletion affects the generation of histone-free DNA particles present in the first fractions of the gradients, which are expected to correspond to low

molecular weight DNA-TF complexes (Henikoff et al. 2011). Deep sequencing of the DNA prepared from these fractions revealed that Brg1-depletion resulted in a marked reduction of enhancer genomic fragments, revealing that Brg1 is also required for efficient production of accessible, histone-free DNA (Extended Data Fig. 7a). Analysis of TF motifs revealed that Oct4, Sox2, Esrrb, Klf4 and Nanog motifs were depleted in the pool of remaining histone-free DNA templates in Brg1-depleted cells, whereas CTCF binding motifs were unaffected (Extended Data Fig. 7b). Altogether, our results demonstrate that Brg1 activity is required at enhancers for the generation of all categories of subnucleosomal particles and for the production of histone-free DNA templates bound by pluripotency-associated TFs in ES cells.

Intriguingly, the Oct4, Sox2, or Nanog MNase ChIP-seq experiments (Fig. 1f-h) revealed that the sizes of the DNA fragments immunoprecipitated by antibodies against these TFs were more compatible with that of the 50-80 bp subnucleosomal particles than with the average size of histone-free DNA fragments (Fig. 3a left). This pattern of interaction with genomic DNA, which was reproduced with three distinct antibodies against Oct4 (Fig. 1f, 4 and data not shown), suggests that *in vivo*, the subnucleosomal particles might represent a better binding substrate for TFs than histone-free DNA fragments. To investigate this possibility, we performed additional sucrose gradient experiments and analysed by ChIP-seq the binding of Oct4 to the subnucleosomal particles and to histone-free DNA. While the Oct4 antibody successfully immunoprecipitated both histone-free DNA fragments (gradient fractions 2-3-4) and 50-80 bp subnucleosomal particles (gradient fractions 5-6), Oct4 enrichment was consistently higher in the latter (Fig. 4a and Extended Data Fig. 8, 9). This was not due to a lesser efficiency of library preparation linked to the lower average length of histone-free DNA fragments, since sequencing of the input DNA from this fraction (without Oct4 immunoprecipitation) yielded high quality signal and control metrics (Extended Data Fig. 7 and Methods). Oct4 binding to both histone-free DNA and 50-80 bp subnucleosomal particles was strongly dependent on Brg1 (Fig. 4a and Extended Data Fig. 8, 9)

To ultimately verify that *in vivo*, Oct4 is bound onto subnucleosomal-size genomic DNA in complex with histones, we performed sequential ChIP experiments. Chromatin prepared from control ES cells was first immunoprecipitated with an Oct4 antibody. Elution of Oct4-bound chromatin was performed using a competitor peptide, and the eluted chromatin was then immunoprecipitated with a second antibody against histone H2B. Deep-sequencing of the

DNA associated with both, Oct4 and histone H2B, revealed DNA-fragments of subnucleosomal size, thus confirming that Oct4 is preferentially associated with 50-80 bp subnucleosomal particles *in vivo* (Fig. 4b).

Afterwards, we addressed how Oct4 might bind to subnucleosomal particles. We first analysed the distribution of Oct4 consensus binding sites, in relation with the positioning of subnucleosomal particles, as well as of fragile and canonical nucleosomes. This analysis revealed that Oct4 consensus motifs sites are strongly enriched within a ≈ 40 bp interval centred on the position of the dyad axis of the central fragile nucleosome (Extended Data Fig. 10). Our ChIP-seq analysis of Oct4 binding to histone-free DNA fragments revealed that it binds precisely at the level of consensus motifs (Extended Data Fig. 8, 9). In contrast, Oct4 binding pattern to DNA within subnucleosomal particles did not correlate with the position of consensus binding sites (Extended Data Fig. 8, 9), suggesting that Oct4 might interact with partial motifs, as previously shown on the nucleosome (Soufi et al. 2015; Michael et al. 2020). However, partial Oct4 motifs were not specifically enriched within the genomic interval spanning the subnucleosomes, compared to neighbouring regions (data not shown). Thus, it is the context of the subnucleosome which seems to facilitate or stabilize Oct4 interactions with DNA. A previous analysis of Oct4 interactions with nucleosomes by crosslinking experiments revealed contacts between Oct4 and histone H3, and to a lesser extent with histones H2A and H4 (Echigoya et al. 2020). Whether interactions between Oct4 and histones might contribute to the strong Oct4 enrichment onto subnucleosomes remains to be determined.

In this study, we described the modular nucleosomal organization of enhancers, in which two positioned canonical nucleosomes flank a variable number of basic modules (Extended Data Figure 11). The basic module is composed of a fragile nucleosome targeted by Brg1 to produce on each side of its dyad axis a subnucleosomal particle associated with 50-80 bp of DNA. Presence of the four histones H3, H4, H2A, H2B, and their slow sedimentation speed compared to canonical nucleosomes, suggest that these particles correspond to hemisomes. We established that these subnucleosomal particles are the main genomic regions of enrichment for Oct4, a major regulator of pluripotency, as well as for other TFs essential for the maintenance of ES cell phenotype. We further demonstrated that Brg1 is required for the proper positioning of both canonical and fragile nucleosomes across the enhancer, and that

Brg1 activity also releases accessible, histone-free DNA fragments, which are bound by Oct4, but with an enrichment well below that of subnucleosomal particles. This last finding is coherent with a recent study based on single molecule footprinting experiments, which revealed that Oct4 interactions with its motif onto naked DNA are low frequency events in ES cells (Sönmezer et al. 2021). The 50-80 bp subnucleosomal particles could be intermediates in the conversion of nucleosomal DNA to accessible, histone-free DNA. However, their status of TF-best enriched substrate suggest that they have a function of their own, which could be to serve as anchors for pioneer and other classes of TFs, in order to maintain enhancer function during ES cell self-renewal.

Aknowledgments:

We thank A. Goldar, A. Leforestier, J.M. Victor and N. Gilbert for discussions. This work was supported by grants from the INCA (2017-1-PL BIO-02-CEA-1) and the ANR (ANR-17-CE12-0001-01) and Investissement d'Avenir ANR (Revive Labex ANR-10- LABX-73).

Authors contributions:

M.N. designed, performed and analysed experiments, and bioinformatic analyses.

A.M. designed, performed and analysed experiments.

E. D. performed the majority of bioinformatic analyses.

C.P. and F.R. performed experiments.

H.P. assisted with experiments.

F.O. contributed to data interpretation.

S.C. designed and analysed experiments.

M.G. conceived the project, designed and analysed experiments, and wrote the manuscript.

References:

- Carone, Benjamin R., Jui-Hung Hung, Sarah J. Hainer, Min-Te Chou, Dawn M. Carone, Zhiping Weng, Thomas G. Fazio, et Oliver J. Rando. 2014. « High-Resolution Mapping of Chromatin Packaging in Mouse Embryonic Stem Cells and Sperm ». *Developmental Cell* 30 (1): 11-22. <https://doi.org/10.1016/j.devcel.2014.05.024>.
- Chen, Xi, Han Xu, Ping Yuan, Fang Fang, Mikael Huss, Vinsensius B. Vega, Eleanor Wong, et al. 2008. « Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells ». *Cell* 133 (6): 1106-17. <https://doi.org/10.1016/j.cell.2008.04.043>.

- Clapier, Cedric R., Janet Iwasa, Bradley R. Cairns, et Craig L. Peterson. 2017. « Mechanisms of Action and Regulation of ATP-Dependent Chromatin-Remodelling Complexes ». *Nature Reviews. Molecular Cell Biology* 18 (7): 407-22. <https://doi.org/10.1038/nrm.2017.26>.
- Clément, Camille, et Geneviève Almouzni. 2015. « MCM2 Binding to Histones H3-H4 and ASF1 Supports a Tetramer-to-Dimer Model for Histone Inheritance at the Replication Fork ». *Nature Structural & Molecular Biology* 22 (8): 587-89. <https://doi.org/10.1038/nsmb.3067>.
- Dieuleveult, Maud de, Kuangyu Yen, Isabelle Hmitou, Arnaud Depaux, Fayçal Bousouar, Daria Bou Dargham, Sylvie Jounier, et al. 2016. « Genome-Wide Nucleosome Specificity and Function of Chromatin Remodellers in ES Cells ». *Nature* 530 (7588): 113-16. <https://doi.org/10.1038/nature16505>.
- Echigoya, Kenta, Masako Koyama, Lumi Negishi, Yoshimasa Takizawa, Yuka Mizukami, Hideki Shimabayashi, Akari Kuroda, et Hitoshi Kurumizaka. 2020. « Nucleosome Binding by the Pioneer Transcription Factor OCT4 ». *Scientific Reports* 10 (1): 11832. <https://doi.org/10.1038/s41598-020-68850-1>.
- Henikoff, Jorja G., Jason A. Belsky, Kristina Krassovsky, David M. MacAlpine, et Steven Henikoff. 2011. « Epigenome Characterization at Single Base-Pair Resolution ». *Proceedings of the National Academy of Sciences of the United States of America* 108 (45): 18318-23. <https://doi.org/10.1073/pnas.1110731108>.
- Ishii, Haruhiko, James T. Kadonaga, et Bing Ren. 2015. « MPE-Seq, a New Method for the Genome-Wide Analysis of Chromatin Structure ». *Proceedings of the National Academy of Sciences of the United States of America* 112 (27): E3457-3465. <https://doi.org/10.1073/pnas.1424804112>.
- Iurlaro, Mario, Michael B. Stadler, Francesca Masoni, Zainab Jagani, Giorgio G. Galli, et Dirk Schübeler. 2021. « Mammalian SWI/SNF Continuously Restores Local Accessibility to Chromatin ». *Nature Genetics* 53 (3): 279-87. <https://doi.org/10.1038/s41588-020-00768-w>.
- Iwafuchi-Doi, Makiko, Greg Donahue, Akshay Kakumanu, Jason A. Watts, Shaun Mahony, B. Franklin Pugh, Dolim Lee, Klaus H. Kaestner, et Kenneth S. Zaret. 2016. « The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation ». *Molecular Cell* 62 (1): 79-91. <https://doi.org/10.1016/j.molcel.2016.03.001>.
- Kagey, Michael H., Jamie J. Newman, Steve Bilodeau, Ye Zhan, David A. Orlando, Nynke L. van Berkum, Christopher C. Ebmeier, et al. 2010. « Mediator and Cohesin Connect Gene Expression and Chromatin Architecture ». *Nature* 467 (7314): 430-35. <https://doi.org/10.1038/nature09380>.
- Koch, Frederic, Romain Fenouil, Marta Gut, Pierre Cauchy, Thomas K. Albert, Joaquin Zacarias-Cabeza, Salvatore Spicuglia, et al. 2011. « Transcription Initiation Platforms and GTF Recruitment at Tissue-Specific Enhancers and Promoters ». *Nature Structural & Molecular Biology* 18 (8): 956-63. <https://doi.org/10.1038/nsmb.2085>.
- Markert, Jonathan, et Karolin Luger. 2021. « Nucleosomes Meet Their Remodeler Match ». *Trends in Biochemical Sciences* 46 (1): 41-50. <https://doi.org/10.1016/j.tibs.2020.08.010>.
- Michael, Alicia K., Ralph S. Grand, Luke Isbel, Simone Cavadini, Zuzanna Kozicka, Georg Kempf, Richard D. Bunker, et al. 2020. « Mechanisms of OCT4-SOX2 Motif Readout on

- Nucleosomes ». *Science* (New York, N.Y.) 368 (6498): 1460-65. <https://doi.org/10.1126/science.abb0074>.
- Narlikar, Geeta J., Ramasubramanian Sundaramoorthy, et Tom Owen-Hughes. 2013. « Mechanisms and Functions of ATP-Dependent Chromatin-Remodeling Enzymes ». *Cell* 154 (3): 490-503. <https://doi.org/10.1016/j.cell.2013.07.011>.
- Natsume, Toyoaki, Tomomi Kiyomitsu, Yumiko Saga, et Masato T. Kanemaki. 2016. « Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors ». *Cell Reports* 15 (1): 210-18. <https://doi.org/10.1016/j.celrep.2016.03.001>.
- Ramachandran, Srinivas, Kami Ahmad, et Steven Henikoff. 2017. « Transcription and Remodeling Produce Asymmetrically Unwrapped Nucleosomal Intermediates ». *Molecular Cell* 68 (6): 1038-1053.e4. <https://doi.org/10.1016/j.molcel.2017.11.015>.
- Rao, Satyanarayan, Kami Ahmad, et Srinivas Ramachandran. 2021. « Cooperative Binding between Distant Transcription Factors Is a Hallmark of Active Enhancers ». *Molecular Cell* 81 (8): 1651-1665.e4. <https://doi.org/10.1016/j.molcel.2021.02.014>.
- Schick, Sandra, Sarah Grosche, Katharina Eva Kohl, Danica Drpic, Martin G. Jaeger, Nara C. Marella, Hana Imrichova, et al. 2021. « Acute BAF Perturbation Causes Immediate Changes in Chromatin Accessibility ». *Nature Genetics* 53 (3): 269-78. <https://doi.org/10.1038/s41588-021-00777-3>.
- Sönmezer, Can, Rozemarijn Kleinendorst, Dilek Imanci, Guido Barzaghi, Laura Villacorta, Dirk Schübeler, Vladimir Benes, Nacho Molina, et Arnaud Regis Krebs. 2021. « Molecular Co-Occupancy Identifies Transcription Factor Binding Cooperativity In Vivo ». *Molecular Cell* 81 (2): 255-267.e6. <https://doi.org/10.1016/j.molcel.2020.11.015>.
- Soufi, Abdenour, Meilin Fernandez Garcia, Artur Jaroszewicz, Nebiyu Osman, Matteo Pellegrini, et Kenneth S. Zaret. 2015. « Pioneer Transcription Factors Target Partial DNA Motifs on Nucleosomes to Initiate Reprogramming ». *Cell* 161 (3): 555-68. <https://doi.org/10.1016/j.cell.2015.03.017>.
- Spitz, François, et Eileen E. M. Furlong. 2012. « Transcription Factors: From Enhancer Binding to Developmental Control ». *Nature Reviews. Genetics* 13 (9): 613-26. <https://doi.org/10.1038/nrg3207>.
- Teif, Vladimir B., Yevhen Vainshtein, Maiwen Caudron-Herger, Jan-Philipp Mallm, Caroline Marth, Thomas Höfer, et Karsten Rippe. 2012. « Genome-Wide Nucleosome Positioning during Embryonic Stem Cell Development ». *Nature Structural & Molecular Biology* 19 (11): 1185-92. <https://doi.org/10.1038/nsmb.2419>.
- Thurman, Robert E., Eric Rynes, Richard Humbert, Jeff Vierstra, Matthew T. Maurano, Eric Haugen, Nathan C. Sheffield, et al. 2012. « The Accessible Chromatin Landscape of the Human Genome ». *Nature* 489 (7414): 75-82. <https://doi.org/10.1038/nature11232>.
- Voong, Lilien N., Liqun Xi, Amy C. Sebeson, Bin Xiong, Ji-Ping Wang, et Xiaozhong Wang. 2016. « Insights into Nucleosome Organization in Mouse Embryonic Stem Cells through Chemical Mapping ». *Cell* 167 (6): 1555-1570.e15. <https://doi.org/10.1016/j.cell.2016.10.049>.
- Zlatanova, Jordanka, Thomas C. Bishop, Jean-Marc Victor, Vaughn Jackson, et Ken van Holde. 2009. « The Nucleosome Family: Dynamic and Growing ». *Structure* (London, England: 1993) 17 (2): 160-71. <https://doi.org/10.1016/j.str.2008.12.016>.

Materials and methods:

Culture conditions of mouse ES cells

Mouse ES cells (46C, E14Tg2a and their derivatives) were grown on mitomycin C-inactivated mouse embryonic fibroblasts at 37 °C, 5 % CO₂, in DMEM (Sigma) which was supplemented by a saturating amount of leukaemia inhibitory factor (LIF), 1X non-essential amino acids (Invitrogen), 15% fetal bovine serum (Invitrogen), 0.2 % β-mercaptoethanol (Sigma) and 1X penicillin/streptomycin (Invitrogen) (Tessarollo 2001).

Plasmid construction

pHYPER shRNA vectors targeting CHD1, CHD4, CHD8, BRG1 and EP400 chromatin remodelers were previously described by de Dieuleveult, *et al.* 2016. Two new shRNAs targeting Brg1 were designed (O5 and O7) to have a better efficiency for Brg1 depletion using the same protocol published in de Dieuleveult, *et al.* 2016. Sequences of the sense strand are: *Smarca4* shRNA O5: 5'-GCTCCAGTAAAGATGTCTACT-3'; *Smarca4* shRNA O7: 5'-GAGCGAATGCGGAGGCTTATG-3'.

For the targeting of both SMARCA4 alleles we adopted the two-selection marker strategy using Neomycin and Hygromycin as described by (Natsume et al. 2016). The SMARCA4-mAID targeting vector was assembled by a serial modification of the base vectors pMK287 (mAID-Hygro, Addgene #72825) and pL452 (Neo,(de Dieuleveult et al. 2016)). First, we introduced two adapters, upstream and downstream of the region coding for mAID and Hygro resistance into the pMK287 vector. The upstream adapter contained in addition to the MCS a (GGGS)₃ linker (in front of the mAID coding sequence). The region containing the linker-mAID sequence was then PCR amplified using Phusion polymerase (NEB, M0531S) from the modified pMK287-mAID-Hygro plasmid and subcloned into the pL452 vector which harbors the Neomycin resistance cassette under the control of a PGK promoter. Left and right homology arms (400-500 bp each) to the last exon and 3'UTR of *Smarca4* were PCR amplified from E14Tg2A genomic DNA. Oligos for the amplification of homology arms were designed to introduce silent mutations after integration of the mAID cassette downstream of the *Smarca4* gene locus in order to prevent re-cutting by the Cas9 enzyme. In the next cloning steps, PCR products of left and right homology arms were introduced sequentially using restriction-

digestion into both vectors (pMK287-mAID-Hygro and pL452-mAID-Neo). Final plasmids for electroporation were prepared using the Nucleobond Maxi kit (Macherey Nagel). Afterwards, the constructs were linearized and concentrated using ethanol precipitation.

CRISPR/Cas9 vectors were assembled by annealing pairs of oligos carrying the sgRNA sequences and cloning into pSpCas9(BB)-2A-Puro (PX459) from Feng Zhang (Addgene #62988) following the protocol described in (Ran et al. 2013). Two distinct Smarca4-targeting sgRNAs were cloned separately by annealing oligos caccgTTGGCTGGGACGAGCGCCTC and aaacGAGGCGCTCGTCCCAGCCAAC for the first sgRNA and caccgGCGCCTCGGGGTCAGGACTC and aaacGAGTCCTGACCCCGAGGCGCc for the second sgRNA. The sgRNAs were designed to target the area close to the last exon of Brg1 using the crispr.mit.edu tool.

Generation of the Brg1-mAID cell line

To generate stable cell lines, we used the E14-TIR1 cell line which was generated by targeting the *Tigre* locus of E14Tg2a cells homozygously with the Tir1-2A-puro cassette (cell line #6, (Nora et al. 2017)). This cell line was transfected by electroporation using the Neon kit (MPK10096, ThermoFisher) with the 2 vectors encoding for the Cas9 enzyme and sgRNAs (2.5 µg each) as well as with the 2 linearized plasmids (pMK287, pL452, 10 µg each) carrying the Brg1 homology arms and the mAID tag. Neomycin (250 µg/mL, G-418, Sigma) and Hygromycin B (300 µg/mL, Thermo Fisher Scientific) selection was started after 24 h. After 6 days of selection, single clones were manually picked, expanded in 24-well plates and subjected to genotyping by PCR and sequencing of the knock-in region. Selected clones with a homozygous insertion of the mAID cassette at the Brg1 locus were confirmed by Western blotting of protein extracts from auxin-treated cells.

Chromatin remodeler depletion and chromatin preparation

Two million of (Flag)₃-TEV-HA-tagged (FTH) ES cells (published in (de Dieuleveult et al. 2016)) were transfected by electroporation (Amaxa nucleofactor kit (VVPH-1001, Lonza) or Neon kit) with 15 µg of pHYPER-linker control plasmid or a pHYPER plasmid containing remodeler-specific shRNA. 24h after electroporation, cells were under a puromycin selection during 30h (Ep400-3FTH, Chd4-3FTH) or 48h (Chd1-3FTH, Chd8-3FTH, Brg1-3FTH) to keep only the cells which had incorporated the plasmid. The efficiency of remodeler depletion was validated by

western blot experiments using monoclonal antibodies anti-Flag (M2, F1804, mouse monoclonal, Sigma), or anti-HA (HA-11, MMS-101P, mouse monoclonal, Biolegend) epitopes.

For auxin-mediated protein depletion, the cell culture medium of Brg1 mAID-tagged cells was supplemented with 1 mM auxin (IAA, indole-3-acetic acid, Cayman Chemical) during 3 h or 20 h. The efficiency of Brg1 deletion was validated by western blot using an anti-Brg1 antibody (ab110641, Abcam).

Whatever the loss-of-function technique, mouse embryonic fibroblasts were separated from ES cells by 30 min incubation on a gelatin coated plate. Cells were fixed in culture medium with 1 % formaldehyde (Sigma) during 10 min, then quenched with 125 mM Glycine and washed three times with 1X PBS. Cells were counted on a Malassez counting chamber followed by permeabilization and lysis in equal parts of solution I (Tris-HCl pH 7.5 15 mM, NaCl 15 mM, MgCl₂ 5 mM, EGTA 0.1 mM, KCl 60 mM, Sucrose 0.3 M) and solution II (Solution I with 0.8% IGEPAL CA-630 (I3021, Sigma)). After 15 min of incubation on ice the reaction was stopped using two volumes of MNase buffer (Tris-HCl pH 7.5 20 mM, NaCl 20 mM, CaCl₂ 2 mM, MgCl₂ 4 mM, KCl 15 mM, 0.1 M sucrose). The chromatin was digested using 6 units of MNase (New England Biolabs: 200 U/μl) per 1 million cells at 37 °C during 10 min in 500 μL MNase buffer. MNase digestion was stopped by adding 10 mM EDTA to the solution.

For ChIP-seq anti-histones, chromatin was passed through 26G (7/8) syringe (Terumo, AGANI) 13 times up/down to liberate the maximum of chromatin. For ChIP-seq anti-TFs or tandem-ChIP, chromatin was slightly sonicated (Bioruptor, 4 cycles 20 s ON / 40 s OFF, medium intensity). Finally, chromatin samples were centrifuged 10 min, 4°C, 14 000 rpm and supernatants containing chromatin ready for ChIP-seq experiments were stored at -80°C.

The quality and reproducibility of MNase digestion patterns were validated on a 1.7 % agarose gel after crosslink reversal and DNA extraction of chromatin aliquots.

Fractionation of MNase digested chromatin on a sucrose gradient

Digested and filtered (0,45 μM filters (SLHVX13N, Merck)) chromatin samples supplemented with 1X complete (Complete Protease Inhibitor Cocktail, 11836145001, Roche) of around 38

million cells were loaded on 11 home-made 10 % to 30 % (incremented by 2 %) sucrose gradients fractions containing TrisHCl pH 7.5 10 mM, EDTA 1 mM, EGTA 1 mM, NaCl 80 mM, complete 1X and the corresponding percentage of sucrose. After a centrifugation of 16h, 40 000 rpm, 4°C on a Beckman Coulter Optima LE-80K (COL00K24) centrifuge, chromatin fractions of 500 µL were collected (fractions n°1 to n°20). An aliquot of each chromatin fraction was incubated for 1 h at 65°C with 15 µg RNase A (reference), followed by incubation at 55°C with 10 µg proteinase K (03115852001, Roche) for 2 h. Afterwards, the samples were subjected to reverse crosslinking ON at 65°C with 200 mM NaCl, followed by phenol–chloroform DNA extraction and ethanol precipitation. The DNA was quantified using the Qubit system (Thermo). Part of this DNA was analyzed on a 4% agarose gel in order to verify the DNA fragments size and another part was directly sequenced as the DNA input samples.

MNase-ChIP-seq experiments

Different pools of chromatin samples from sucrose gradient experiments were prepared: pool of gradient fractions n°2-3-4, pool of gradient fractions n°5-6 and pool of gradient fractions n°7-8-9. These three pools of chromatin as well as regular whole cell chromatin samples of 1.2 million cells (for ChIP-seq anti-histones on total chromatin) or 10 million cells (for ChIP anti-TFs on total chromatin) were filtered and incubated with antibodies (20 µg of antibody against TFs and Pol II and 5 µg of antibody against histones) overnight (ON) at 4°C. Protein G-agarose beads (ref 11243233001, Sigma) were washed three times with ChIP buffer (Tris-HCl pH 7.5 20 mM, NaCl 20 mM and MgCl₂ 4mM), then blocked 1 h 30 at 4 °C with the blocking buffer (ChIP buffer, BSA 0,2 mg/mL, tRNA 50 µg/mL final) to saturate any non-specific sites on beads which could bind chromatin.

For ChIP-seq anti-TFs chromatin samples of 10 million cells for were first incubated with 40 µL of blocked protein G-agarose beads during 1 h 30 at 4 °C. This chromatin pre-clearing makes it possible to eliminate the chromatin regions which may have an unspecific affinity for the resin. After centrifugation, supernatants containing the cleared chromatin were incubated ON at 4°C with the corresponding antibodies.

Chromatin-antibody complexes were incubated with 50 µL of blocked protein G-agarose beads during 4h at 4°C. Then, chromatin-antibody-beads complexes were centrifuged to eliminate the unbound chromatin and washed three times with TEN buffer (Tris-HCl pH 7.5 20

mM, NaCl 150 mM, MgCl₂ 3 mM, EDTA 0.1 mM, Igepal 0,01 %), four times with WBLiCl buffer (Hepes pH 7.5 50 mM, LiCl 500 mM, EDTA 1 mM, 1 % IGEPAL, 0.7 % Na-deoxycholate) and one time with TEN buffer. Between each washing, beads were centrifuged 3 min, 800 g, 4°C. Finally, beads were resuspended in elution buffer (Tris-HCl pH 8 50 mM, EDTA 10 mM, SDS 1 %) and incubated 15 min at 65°C. Beads were eliminated on Pierce columns (ref 89868, ThermoFisher). The eluted chromatin was incubated for 2 h at 55°C in the presence of 10 µg proteinase K. Afterwards, the samples were subjected to reverse crosslinking ON at 65°C in the presence of 200 mM NaCl. Then, DNA was extracted using Phenol/Chlorophorm (ref 77617, Sigma), precipitated using ethanol and quantified by a Picogreen dosage (Quant-iT PicoGreen dsDNA Reagent, reference P7581, Invitrogen) and a Bioanalyzer (Agilent) dosage. DNA libraries were processed for sequencing with a moderate number of PCR cycles thanks to the kit « MicroPlex Library Preparation Kit v2 (12 indexes) » (C05010013, Diagenode) or the kit « 24 Dual indexes for MicroPlex Kit v3 /48 rxns » (C05010003, Diagenode). Sequencing was mainly done on NextSeq 500 Illumina with the kit « NextSeq 500 High Output v2 kit (75 cycles) » (FC-404-2005, Illumina).

For each ChIP-seq experiment at least two replicate samples were sequenced.

Sequential ChIP-seq

The sequential ChIP-seq experiment protocol has been adapted from the previously published Tandem affinity purification protocol for mESCs:

<http://dx.doi.org/10.1038/protex.2014.040> (DOI: 10.1038/protex.2014.040, Protocol exchange).

Around 243 million formaldehyde fixed cells with MNase digested chromatin were used for each sample. The chromatin was filtered and incubated ON at 4 °C with 45 µg of antibody against OCT4 (MBS420786, myBioSource) or 60 µg of IgG (ab37415, Abcam) as a control. Blocked protein G-agarose beads were added to the chromatin-Antibody complex and incubated 4h at 4°C. After seven washes with TEN buffer, beads were incubated with OCT4 peptide (C-ter of OCT4 sequence: H-Cys-Lys-Lys-Lys-Lys-Pro-Ser-Val-Pro-Val-Thr-Ala-Leu-Gly-Ser-Pro-Met-His-Ser-Asn-OH, GeneCust) for an elution by competition. Three elutions were done: first elution during 2 h at room temperature (RT) with 2 mg/mL peptide, second elution during 2 h at RT with 1 mg/mL peptide and a third elution ON at 4 °C with 1 mg/mL peptide.

Then, the three elution fractions were pooled for the second immunoprecipitation step, for which 5 µg of antibody against H2B was used ON at 4 °C. After a second incubation with protein G-agarose blocked beads during 4h at 4 °C, washes and elution, as described in the “MNase-ChIP-seq experiments” part, the DNA was extracted and prepared for sequencing. As a negative control, 60 µg of non-specific antibody was used in the first immunoprecipitation step to define the background of the experiment. Two replicates were prepared for each condition (control and tested colocalization between Oct4-H2B).

Western blot

Proteins from a small amount ($0.5 - 1 \times 10^6$) of either control or chromatin remodeler depleted cells were washed with 1X PBS, extracted with RIPA buffer (Tris pH 7.5 50 mM, NaCl 150 mM, NP40 1 %, sodium deoxycholate 0.5 % and SDS 0.1 %) + 1X complete + 1X PMSF (P7626, Sigma) during 1h at 4°C. Lysates were centrifuged 15 min at 12 000 rpm, 4°C and supernatants containing proteins were quantified by a Bradford dosage. Around 40 µg of protein denatured for 7 min at 95 °C in Laemmli buffer (TrisHCl pH 7 62.5 mM, SDS 2 %, glycerol 10 %, Bromophenol blue 0.01 %, beta-mercaptoethanol 5 %) were loaded on 4-12 % Bis-Tris NuPAGE gel (NP0335BOX, ThermoFisher) for migration in 1X NuPAGE migration buffer (NP0001, ThermoFisher).

After electrophoresis, a nitrocellulose membrane (GE10600001, Amersham) and gel were incubated for 15 min at RT in 1X NuPAGE transfert buffer (NP0006, ThermoFisher) with 10 % methanol. Protein transfer was done in a semi-dry system (Fast-blot Biomerta) between Whatman papers during 30 min (small proteins) to 45 min (big proteins) at 10 W. Membranes were blocked in blocking buffer (Odyssey buffer (927-70001, Li-Cor) and PBS 1X + Tween 0.1 % at 1:1 ratio) during 1 h at RT. Afterwards, membranes were incubated with the corresponding antibody dilution ON at 4°C. Membranes were washed with PBS 1X + Tween 0.1 % and incubated with the secondary antibody coupled with a fluorophore (Odyssey IRDye anti-primary antibody, Li-Cor) during 1h at RT, followed by washing with PBS 1X. Proteins were revealed on an Odyssey CLx Li-Cor machine.

Table of antibodies

Target protein/name	Reference	Source and clonality	Company	Quantity used	Experiment
anti-H2A	ab18255	Rabbit polyclonal	Abcam	5 µg	MNase-ChIP-seq
anti-H2B	ab1790	Rabbit polyclonal	Abcam	5 µg	MNase-ChIP-seq + tandemChIP
anti-H3	ab1791	Rabbit polyclonal	Abcam	5 µg	MNase-ChIP-seq
anti-H3K27ac	ab4729	Rabbit polyclonal	Abcam	5 µg	MNase-ChIP-seq
anti-H4	ab7311	Rabbit polyclonal	Abcam	5 µg	MNase-ChIP-seq
anti-CTCF	07-729	Rabbit polyclonal	Millipore	20 µg	MNase-ChIP-seq
anti-Esr1b	PP-H6705-00	Mouse monoclonal	PPMX	20 µg	MNase-ChIP-seq
anti-Nanog	RCAB-001P	Rabbit polyclonal	Reprocell	10 µg	MNase-ChIP-seq
anti-Oct4	sc-8628	Goat polyclonal	Santa Cruz Biotechnology	20 µg	MNase-ChIP-seq
anti-Oct4	sc-5279	Mouse monoclonal	Santa Cruz Biotechnology	20 µg	MNase-ChIP-seq
anti-Oct4	MBS470786	Rabbit polyclonal	myBioSource	45 µg	Tandem ChIP
anti-Pol II	sc-55492	Mouse monoclonal	Santa Cruz Biotechnology	20 µg	MNase-ChIP-seq
anti-Sox2	sc-17320	Goat polyclonal	Santa Cruz Biotechnology	20 µg	MNase-ChIP-seq
anti-TBP	ab51841	Mouse monoclonal	Abcam	20 µg	MNase-ChIP-seq
anti-Brg1	ab110641	Rabbit monoclonal	Abcam	1:2000	Western blot
anti-Flag M2	F1804	Mouse monoclonal	Sigma-Aldrich	1:1000	Western blot
anti-HA (HA.11)	MMS-101P	Mouse monoclonal	Biologend	1:2000	Western blot
Isotype control	ab37415	Rabbit polyclonal	Abcam	60 µg	tandem-ChIP

Table of cell lines

Name	Gene	Tag	Parental cell line	Reference
Brg1-3FTH	Brg1 (Smarca4)	(Flag) ₃ -TEV-HA	46C	(de Dieuleveult et al. 2016)
Chd1-3FTH	Chd1	(Flag) ₃ -TEV-HA	46C	(de Dieuleveult et al. 2016)
Chd4-3FTH	Chd4	(Flag) ₃ -TEV-HA	46C	(de Dieuleveult et al. 2016)
Chd8-3FTH	Chd8	(Flag) ₃ -TEV-HA	46C	(de Dieuleveult et al. 2016)
Ep400-3FTH	Ep400	(Flag) ₃ -TEV-HA	46C	(de Dieuleveult et al. 2016)
TIR1	OstTIR1	-	E14Tg2a	(Nora et al. 2017)
Smarca4-mAID	Brg1 (Smarca4)	mAID	TIR1	This study

Table of plasmid constructs

Name	Reference
pHYPER-shChd1	(de Dieuleveult et al. 2016)
pHYPER-shChd4	(de Dieuleveult et al. 2016)
pHYPER-shChd8	(de Dieuleveult et al. 2016)
pHYPER-shEp400	(de Dieuleveult et al. 2016)
pHYPER-Linker	(de Dieuleveult et al. 2016)
pHYPER-shSmarca4-O5	This study
pHYPER-shSmarca4-O7	This study
pMK287	Addgene #72825, (Natsume et al. 2016)
pL452	(de Dieuleveult et al. 2016)
pMK287-Smarca4-mAID	This study
pL452-Smarca4-mAID	This study
pSpCas9(BB)-2A-Puro (PX459)	Addgene #62988, (Ran et al. 2013)
pSpCas9(BB)-2A-Puro-Smarca4-sgRNA1	This study
pSpCas9(BB)-2A-Puro-Smarca4-sgRNA2	This study

Data analysis:

The whole pipeline is running using Snakemake (Mölder et al. 2021), Conda and in-house python script.

From quality controls to mapping

Quality of sequencing data was controlled by FastQC (fastqc=0.11.9, default parameters, developed by Babraham Institute) and fastqScreen (fastq-screen=0.14.0, parameter: `--aligner bowtie2`) tools (Wingett et Andrews 2018). Depending of their quality, sequencing files are trimmed with fastp tool (fastp=0.20.1, parameters: `--length_required 25 --detect_adapter_for_pe --cut_front --cut_tail --cut_window_size 6 --cut_mean_quality 10 --unqualified_percent_limit 40 --n_base_limit 7 --average_qual 0 --overrepresentation_analysis --trim_poly_g`) (Chen et al. 2018).

Trimmed paired-end reads were mapped on the mm9 mouse genome with the Bowtie 2 tool (bowtie2=2.4.2, default parameters) (Langmead et Salzberg 2012). The quality of mapped reads was verified by Samtools flagstat tool (samtools=1.11, default parameters) (Li et al. 2009).

Post-mapping process

After mapping, reads were sorted by genomic coordinates using Samtools sort (samtools=1.11, default parameters) before marking and eliminating duplicated reads with Picard Markduplicates (picard=2.23.8, parameters: `CREATE_INDEX=true VALIDATION_STRINGENCY=STRICT`) ("Picard Toolkit." 2019. *Broad Institute, GitHub Repository*. <http://broadinstitute.github.io/picard/>; Broad Institute) and Samtools view (samtools=1.11, parameters: `-h -f 1 -F 3852 -q 30`).

Peak-calling was made on cleaned mapping files with the MACS2 callpeak tool (macs2=2.2.7.1, parameters: `-f BAMPE -gsize 1.87e9`) (Zhang et al. 2008).

The genomic coverage was calculated thanks to Deeptools (deeptools=3.5.0, parameters: `bamCoverage --outFileFormat bigwig --binSize 1 --numberOfProcessors 5 --effectiveGenomeSize 2652783500`

```
normalizeUsing RPKM --MNase --minFragmentLength 10 --extendReads
--centerReads --maxFragmentLength 500 --ignoreForNormalization
chrX chrY chrM --blackListFileName gencode_mm9/mm9-
blacklist.bed.gz) (Ramírez et al. 2016).
```

MultiQC (MultiQC=1.9, default parameters) is used to summarize all the quality controls we are doing through the pipeline (Ewels et al. 2016).

IGV visualizations and 2D histograms (V-plots)

Peaks and normalized coverage profiles were visualized (via a bigwig file) as IGV screen-shots and heatmaps.

2D graphs (V-plots) are based on bed files. A home-made script using bedtools bamtobed (bedtools=2.29.2, parameters: `-bedpe`) allows the conversion of a bam file containing paired-end mapped reads to a bed file containing DNA fragments (Quinlan et Hall 2010).

These fragments can be visualized as a heatmap using seqMINER software (Ye et al. 2011) and as a 2D graph thanks to a home-made script. In the plot, the Y axis corresponds to DNA fragment size in bp and the X axis to relative position of DNA fragments on the genome. Each dot corresponds to the center coordinates of a DNA fragment and color intensity reflects the number of DNA fragments. To get comparable 2D histograms, the signal is divided by the ratio of difference of the library size. The color scale is fixed to the highest value for all the selected samples that we want to compare.

Motif detection and analysis

Fusion of Input files (quadruplets) for gradient fractions control and BRG1-depleted samples was performed thanks to Samtools merge (samtools=1.10, using htlib=1.10.2, default parameters). Peak-calling was made on paired-end mapped files with MACS2 callpeak (macs2=2.2.7.1, parameters: `-f BAMPE -g mm --min-length 40`). Then, motifs were detected with HOMER findMotifsGenome.pl (homer=4.11, parameters: `findMotifsGenome.pl mm9 -size given`). In motifs detection files, we choose 31 motifs which are significant (p -value $< 10^{-50}$) in all control samples. Thanks to a home-made R script, we divided the CTCF motif number for the control sample by the CTCF motif number for Brg1-depleted sample. We normalized according to the CTCF motifs number because we

assumed that CTCF binding is not impacted by the loss-of-function of the BRG1 remodeler. The normalization coefficient is applied to 30 other motifs in all BRG1-depleted samples. Then, the ratio between motifs in BRG1-depleted cells vs control cells is calculated. A heatmap is produced to visualize the results.

For motifs analysis, we worked with the Meme suite (v5.3.0) (Bailey et al. 2009). To detect *de novo* motifs, streame was running with the parameters `--dna --minw 3`. To find known motifs as Oct4 in a list of reference peaks, Fimo was selected (the threshold value was adapted depending on the motifs size). The Oct4 motif matrix was taken from the Jaspas database (Fornes et al. 2020).

Materials and methods references:

- Bailey, Timothy L., Mikael Boden, Fabian A. Buske, Martin Frith, Charles E. Grant, Luca Clementi, Jingyuan Ren, Wilfred W. Li, et William S. Noble. 2009. « MEME SUITE: Tools for Motif Discovery and Searching ». *Nucleic Acids Research* 37 (Web Server issue): W202-208. <https://doi.org/10.1093/nar/gkp335>.
- Chen, Shifu, Yanqing Zhou, Yaru Chen, et Jia Gu. 2018. « Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor ». *Bioinformatics (Oxford, England)* 34 (17): i884-90. <https://doi.org/10.1093/bioinformatics/bty560>.
- Dieuleveult, Maud de, Kuangyu Yen, Isabelle Hmitou, Arnaud Depaux, Fayçal Boussouar, Daria Bou Dargham, Sylvie Jounier, et al. 2016. « Genome-Wide Nucleosome Specificity and Function of Chromatin Remodellers in ES Cells ». *Nature* 530 (7588): 113-16. <https://doi.org/10.1038/nature16505>.
- Ewels, Philip, Måns Magnusson, Sverker Lundin, et Max Källér. 2016. « MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report ». *Bioinformatics (Oxford, England)* 32 (19): 3047-48. <https://doi.org/10.1093/bioinformatics/btw354>.
- Fornes, Oriol, Jaime A. Castro-Mondragon, Aziz Khan, Robin van der Lee, Xi Zhang, Phillip A. Richmond, Bhavi P. Modi, et al. 2020. « JASPAR 2020: Update of the Open-Access Database of Transcription Factor Binding Profiles ». *Nucleic Acids Research* 48 (D1): D87-92. <https://doi.org/10.1093/nar/gkz1001>.
- Langmead, Ben, et Steven L. Salzberg. 2012. « Fast Gapped-Read Alignment with Bowtie 2 ». *Nature Methods* 9 (4): 357-59. <https://doi.org/10.1038/nmeth.1923>.
- Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, et 1000 Genome Project Data Processing Subgroup. 2009. « The Sequence Alignment/Map Format and SAMtools ». *Bioinformatics (Oxford, England)* 25 (16): 2078-79. <https://doi.org/10.1093/bioinformatics/btp352>.
- Mölder, Felix, Kim Philipp Jablonski, Brice Letcher, Michael B. Hall, Christopher H. Tomkins-Tinch, Vanessa Sochat, Jan Forster, et al. 2021. « Sustainable Data Analysis with Snakemake ». *F1000Research* 10: 33. <https://doi.org/10.12688/f1000research.29032.2>.

- Natsume, Toyoaki, Tomomi Kiyomitsu, Yumiko Saga, et Masato T. Kanemaki. 2016. « Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors ». *Cell Reports* 15 (1): 210-18. <https://doi.org/10.1016/j.celrep.2016.03.001>.
- Nora, Elphège P., Anton Goloborodko, Anne-Laure Valton, Johan H. Gibcus, Alec Uebersohn, Nezar Abdennur, Job Dekker, Leonid A. Mirny, et Benoit G. Bruneau. 2017. « Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization ». *Cell* 169 (5): 930-944.e22. <https://doi.org/10.1016/j.cell.2017.05.004>.
- Quinlan, Aaron R., et Ira M. Hall. 2010. « BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features ». *Bioinformatics (Oxford, England)* 26 (6): 841-42. <https://doi.org/10.1093/bioinformatics/btq033>.
- Ramírez, Fidel, Devon P. Ryan, Björn Grüning, Vivek Bhardwaj, Fabian Kilpert, Andreas S. Richter, Steffen Heyne, Friederike Dündar, et Thomas Manke. 2016. « DeepTools2: A next Generation Web Server for Deep-Sequencing Data Analysis ». *Nucleic Acids Research* 44 (W1): W160-165. <https://doi.org/10.1093/nar/gkw257>.
- Ran, F. Ann, Patrick D. Hsu, Jason Wright, Vineeta Agarwala, David A. Scott, et Feng Zhang. 2013. « Genome Engineering Using the CRISPR-Cas9 System ». *Nature Protocols* 8 (11): 2281-2308. <https://doi.org/10.1038/nprot.2013.143>.
- Tessarollo, L. 2001. « Manipulating Mouse Embryonic Stem Cells ». *Methods in Molecular Biology (Clifton, N.J.)* 158: 47-63. <https://doi.org/10.1385/1-59259-220-1:47>.
- Wingett, Steven W., et Simon Andrews. 2018. « FastQ Screen: A Tool for Multi-Genome Mapping and Quality Control ». *F1000Research* 7: 1338. <https://doi.org/10.12688/f1000research.15931.2>.
- Ye, Tao, Arnaud R. Krebs, Mohamed-Amin Choukrallah, Celine Keime, Frederic Plewniak, Irwin Davidson, et Laszlo Tora. 2011. « SeqMINER: An Integrated ChIP-Seq Data Interpretation Platform ». *Nucleic Acids Research* 39 (6): e35. <https://doi.org/10.1093/nar/gkq1287>.
- Zhang, Yong, Tao Liu, Clifford A. Meyer, Jérôme Eeckhoute, David S. Johnson, Bradley E. Bernstein, Chad Nusbaum, et al. 2008. « Model-Based Analysis of ChIP-Seq (MACS) ». *Genome Biology* 9 (9): R137. <https://doi.org/10.1186/gb-2008-9-9-r137>.

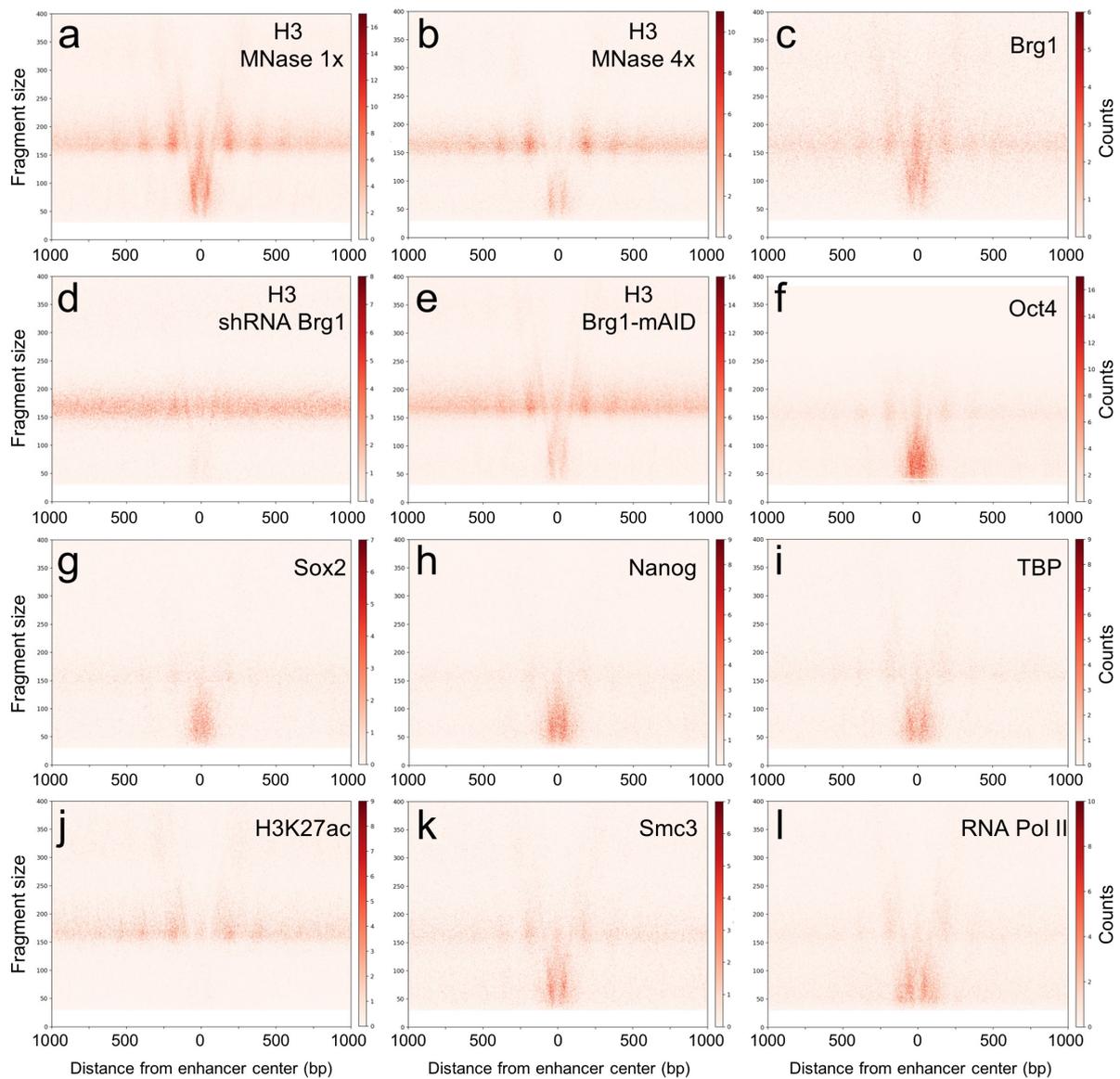


Figure 1: Nucleosomal and subnucleosomal organization of enhancers is controlled by Brg1.

a, b. V-plots of histone H3 ChIP-seq spanning +/- 1000 bp from group 1 enhancer centre, using either the standard (1x) MNase dose (**a**), or a four-fold excess (**b**). Red dots coordinates on the x axis indicate the genomic position of the centre of immunoprecipitated DNA fragments, and their length in bp on the y axis. The color scale corresponds to the number of DNA fragments.

c. V-plot of Brg1 ChIP-seq. **d, e.** Histone H3 ChIP-seq fragments distribution was analysed at group 1 enhancers following depletion of Brg1 using either an shRNA (**d**), or the mAID system (**e**). **f-l.** V-plots of Oct4, Sox2, Nanog, TBP, H3K27ac, Smc3 and RNA Pol II at group 1 enhancers.

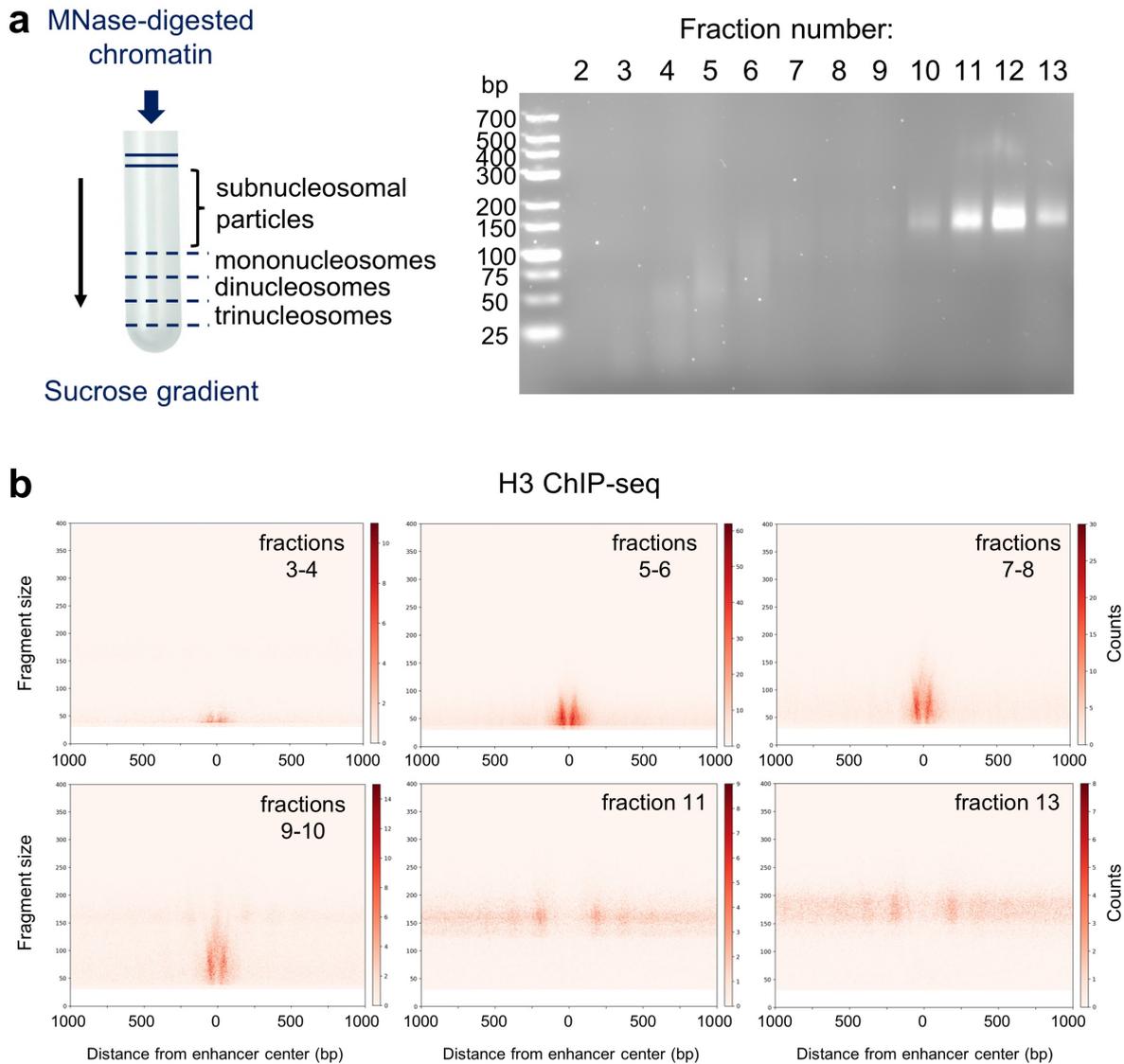


Figure 2: Sucrose gradient centrifugation experiments highlight the existence of subnucleosomal particles associated with DNA fragments of different size at enhancers.

a. Centrifugation of MNase-digested chromatin on a sucrose gradient separates molecules according to their shape and molecular weight. After centrifugation, the different fractions are collected and the DNA analyzed onto a 4% agarose gel.

b. H3 ChIP-seq enrichment at group 1 enhancers coming from chromatin gradient fractions: pool 3-4, pool 5-6, pool 7-8, pool 9-10, fraction 11 and fraction 13 from control cells. See the legend of Figure 1 for the description of the V-plot representation.

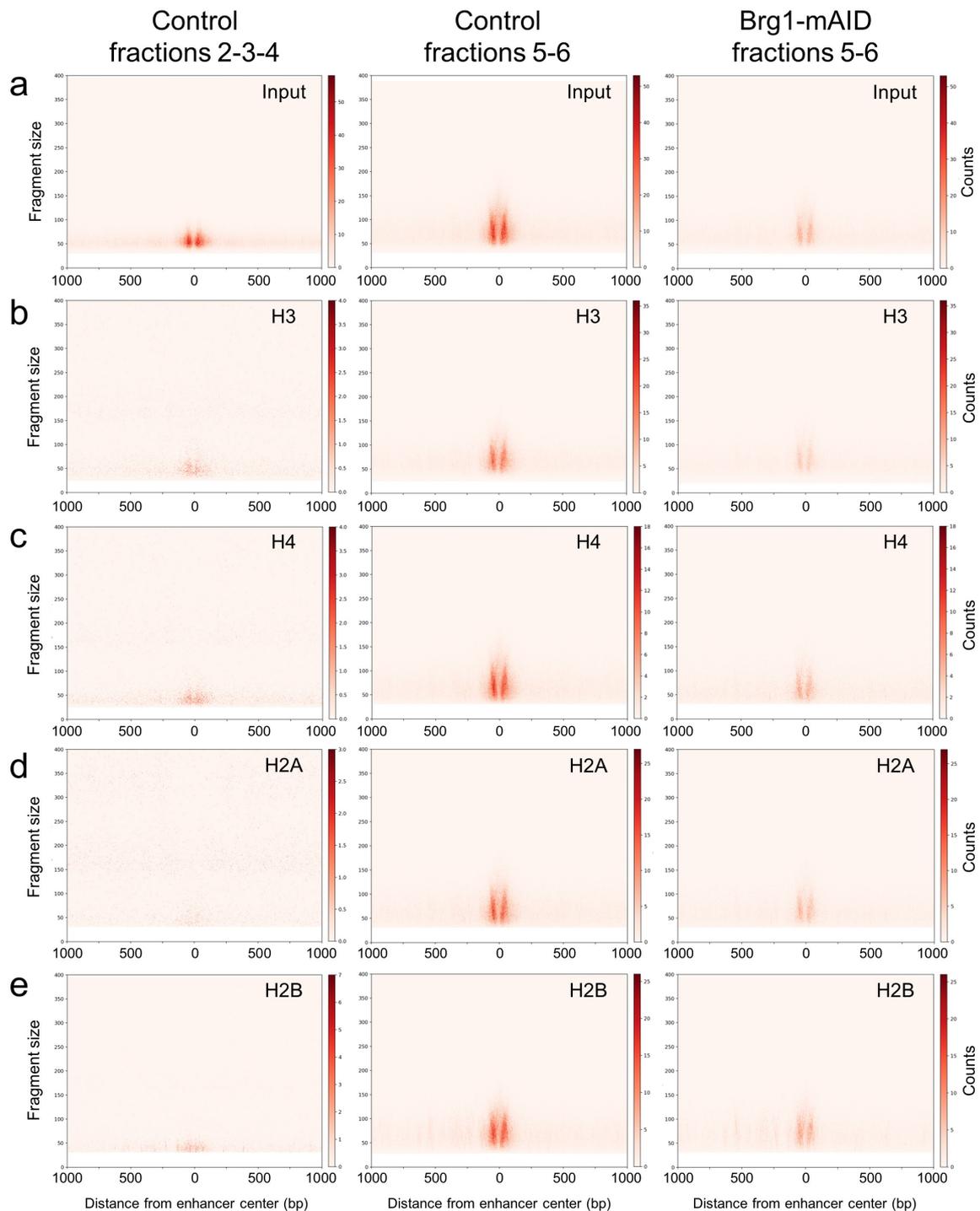


Figure 3: The SWI/SNF remodeler ATPase Brg1 is involved in the generation of subnucleosomal particles at enhancer elements.

V-plots of input DNA sequenced from selected gradient fractions, or of ChIP-seq experiments performed with histone antibodies. 1000 bp are shown on each side of group 1 enhancer centres. Chromatin was prepared either from control ES cell chromatin (left and middle) or from Brg1-depleted cells (20 h Aux treatment of Brg1-mAID cells, right). ChIP-seq was performed either on a pool of gradient fractions 2-3-4 (left) or 5-6 (middle and right).

a. Input DNA. **b.-e.** ChIP against histone H3 (**b.**), H4 (**c.**), H2A (**d.**) and H2B (**e.**).

See the legend of Figure 1 for the description of the V-plot representation.

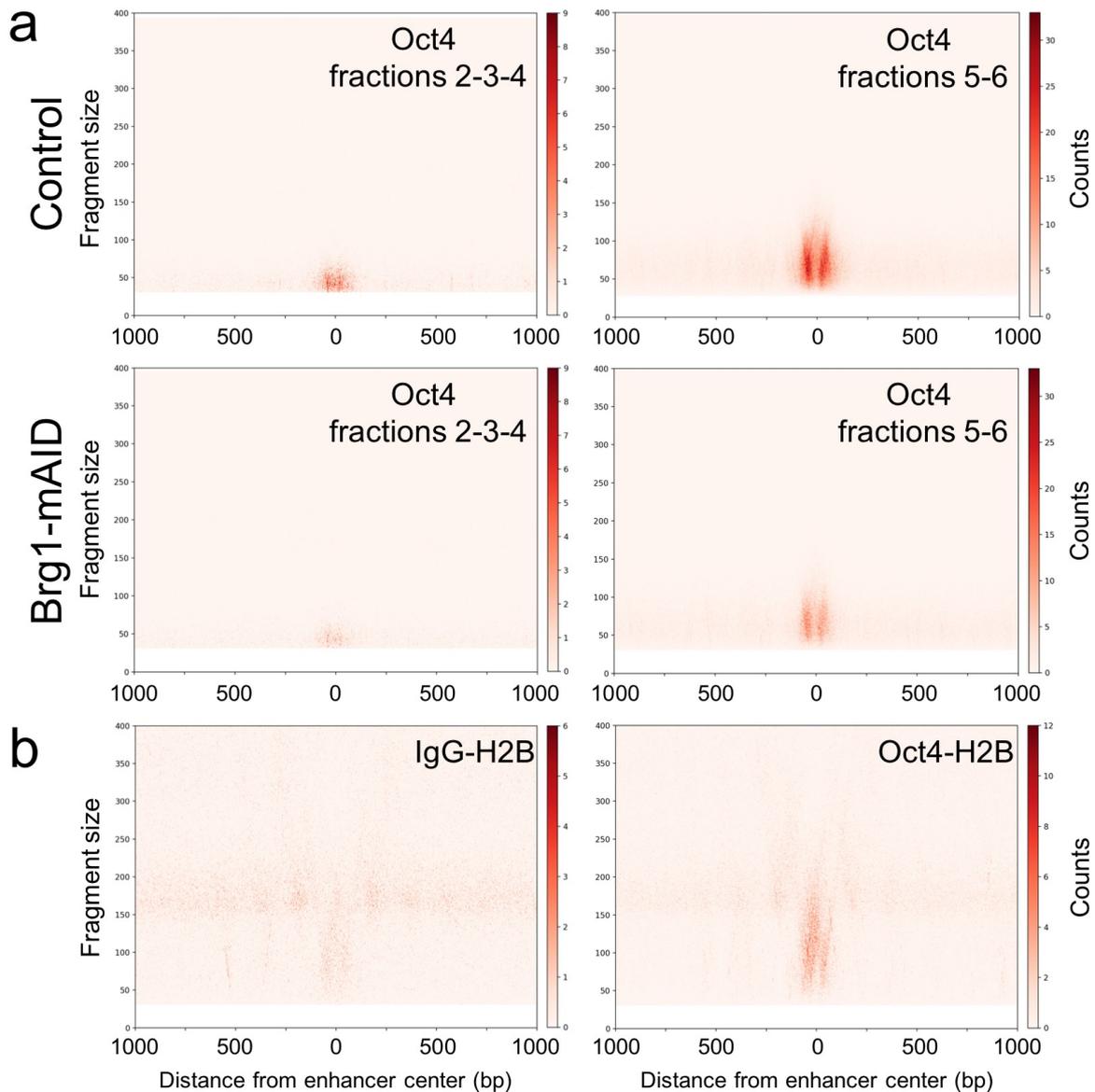
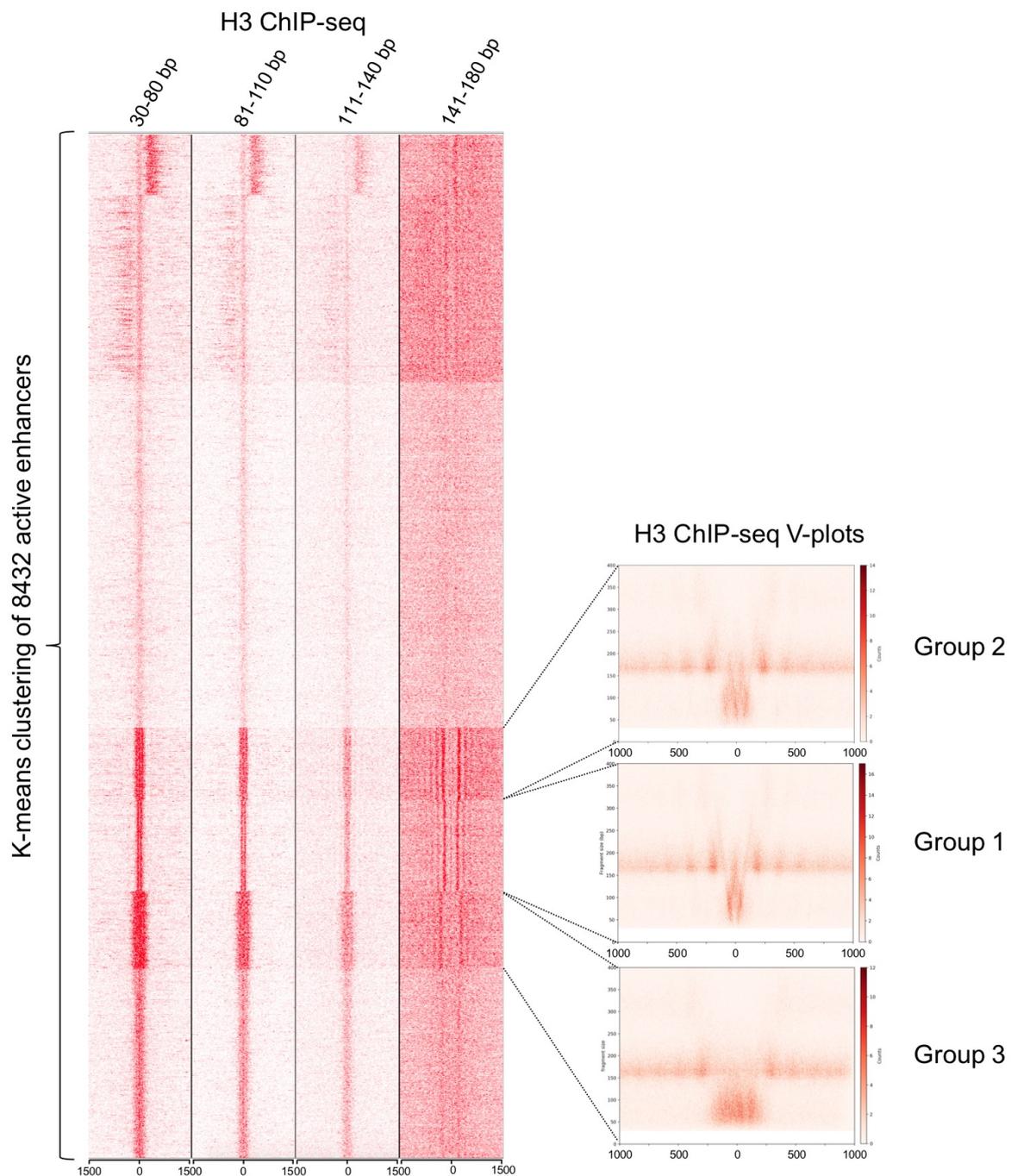


Figure 4: Oct4 binding to ES cell genome occurs predominantly at the level of the subnucleosomal particles generated by Brg1.

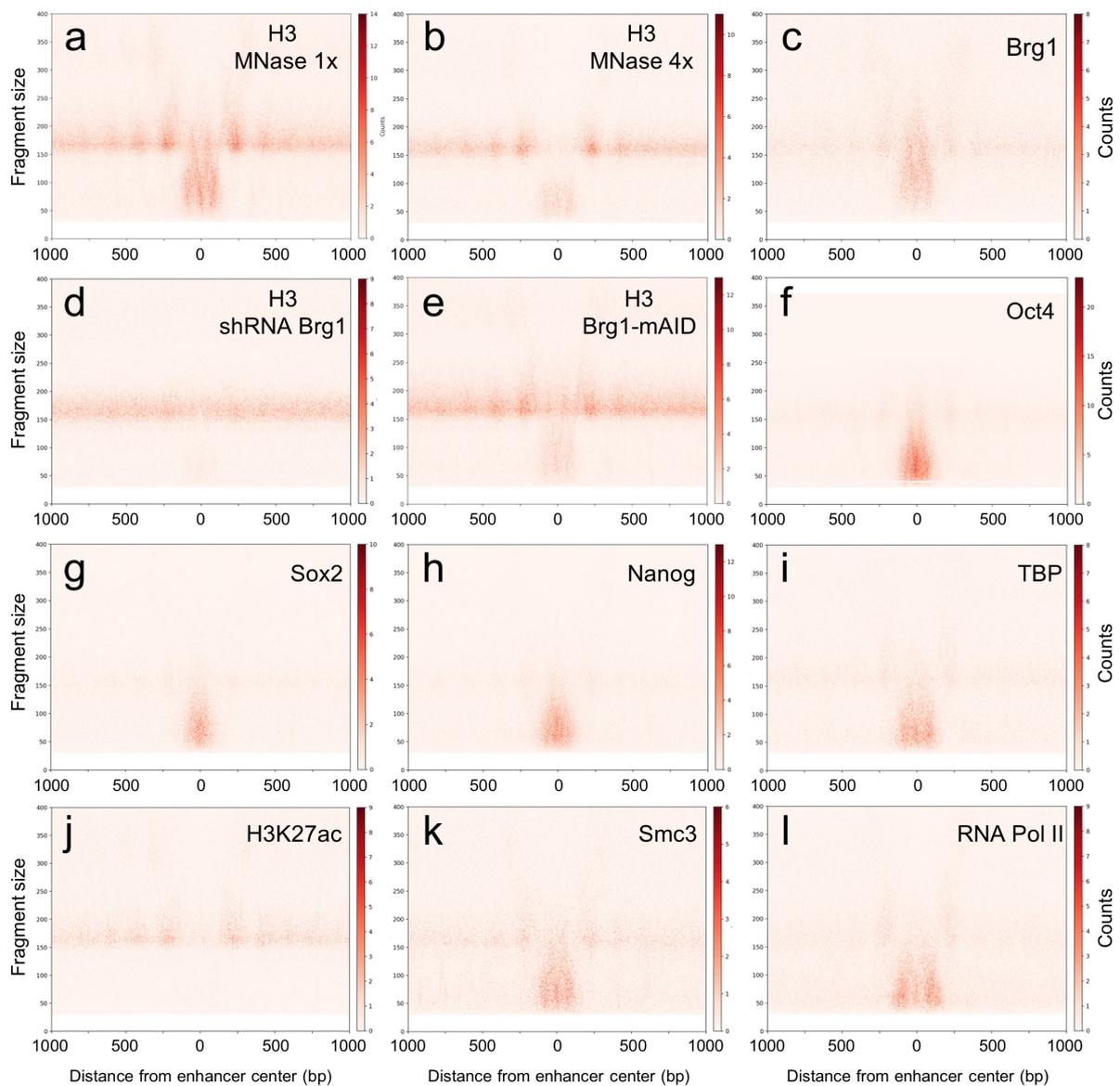
a. V-plot of Oct4 ChIP-seq spanning +/- 1000 bp from group 1 enhancer centre. ChIP experiments were performed on selected sucrose gradient fractions obtained after centrifugation of control and Brg1-depleted (20 h Aux treatment of Brg1-mAID cells) chromatin. Fractions 2-3-4 contain mainly histone-free DNA, whereas fractions 5-6 contain the 50-90 bp subnucleosomal particles.

b. Sequential ChIP-seq experiment performed on chromatin prepared from control ES cells. Right panel : chromatin was first immunoprecipitated with an Oct4 antibody targeting the C-terminal 20 amino-acids. Elution was performed by peptide competition, and the eluted fraction was subjected to a second round of immunoprecipitation with an antibody against H2B. In the left panel, chromatin was first incubated with a control, non-specific IgG, antibody. All other steps were identical to the ChIP performed with Oct4 as the first antibody. See the legend of Figure 1 for the description of the V-plot representation.



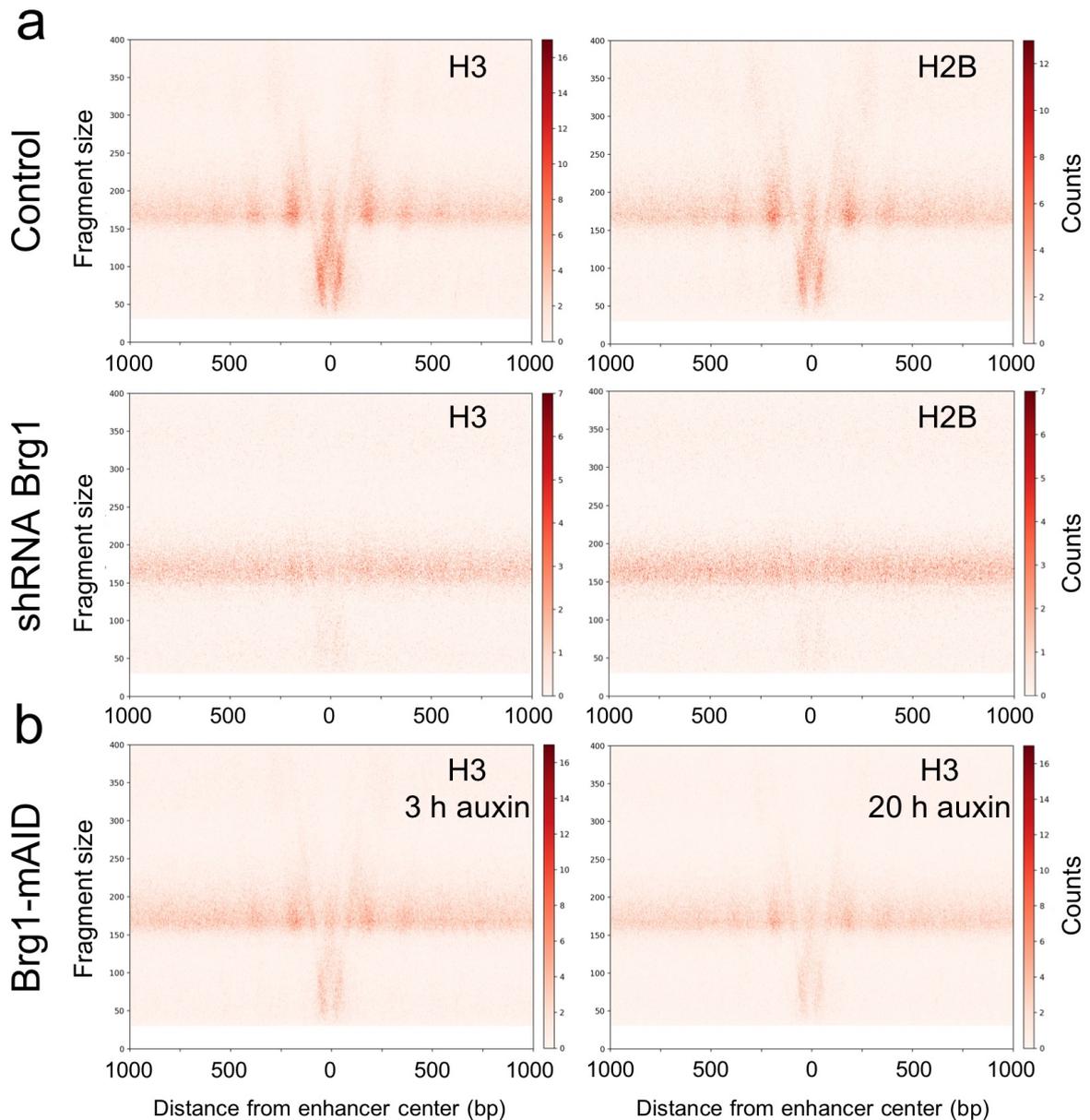
Extended data Figure 1: K-means clustering of mESC-enhancers based on their H3 enrichment in mESCs.

Histone H3 enrichment at 8432 active enhancers was determined by MNase-ChIP-seq. Using k-means clustering, 10 groups of enhancers were obtained (left panel). Groups 1, 2 and 3, which show specific patterns of nucleosomal and subnucleosomal organization were chosen for V-plot analyses (right panels).



Extended data Figure 2: Nucleosomal and subnucleosomal organization of group 2 enhancers.

a, b. V-plots of histone H3 ChIP-seq spanning +/- 1000 bp from group 2 enhancer centre, using either the standard (1x) MNase dose (**a**), or a four-fold excess (**b**). **c.** V-plot of Brg1 ChIP-seq. **d, e.** Histone H3 ChIP-seq fragments distribution was analysed at group 2 enhancers following depletion of Brg1 using either an shRNA (**d**), or the mAID system (**e**). **f-l.** V-plots of Oct4, Sox2, Nanog, TBP, H3K27ac, Smc3 and RNA Pol II at group 2 enhancers.

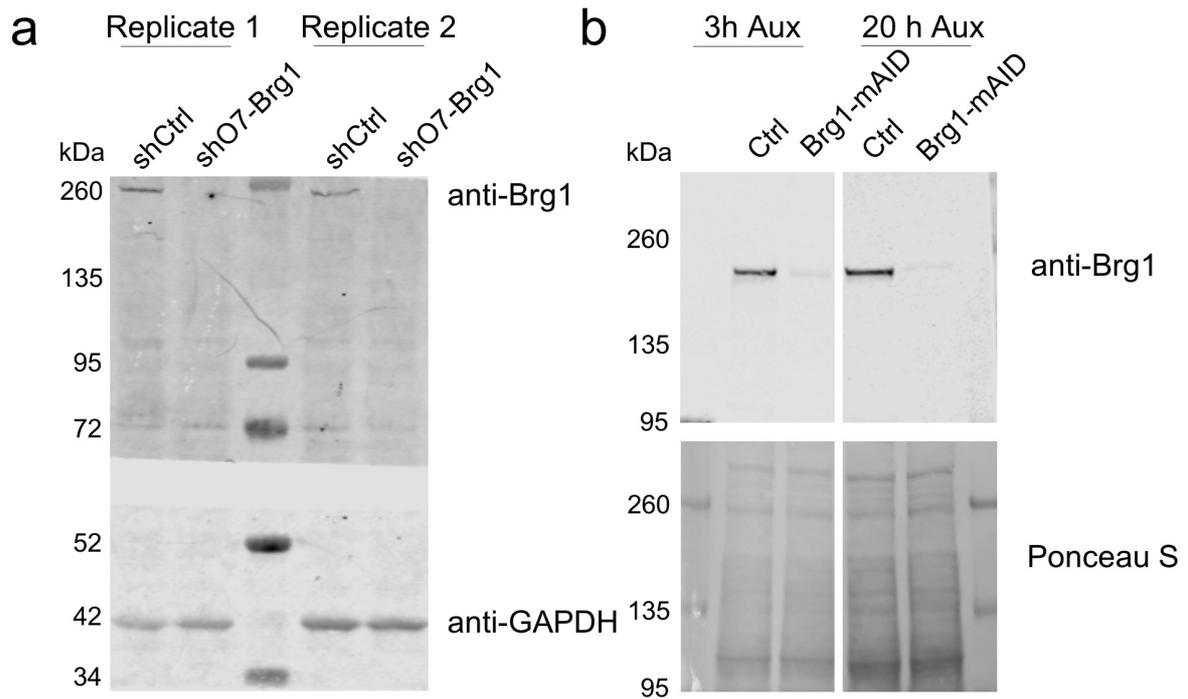


Extended data Figure 3: Brg1 depletion induces a loss of subnucleosomal particles and a perturbation of nucleosome phasing at enhancers.

a. H3 and H2B ChIP-seq was performed on MNase-digested control chromatin or chromatin obtained from cells after shRNA-depletion of Brg1 during 72 h.

b. H3 ChIP-seq was performed on MNase-digested chromatin of Brg1-mAID cells after a 3 h or 20h-long treatment with 1 mM auxin.

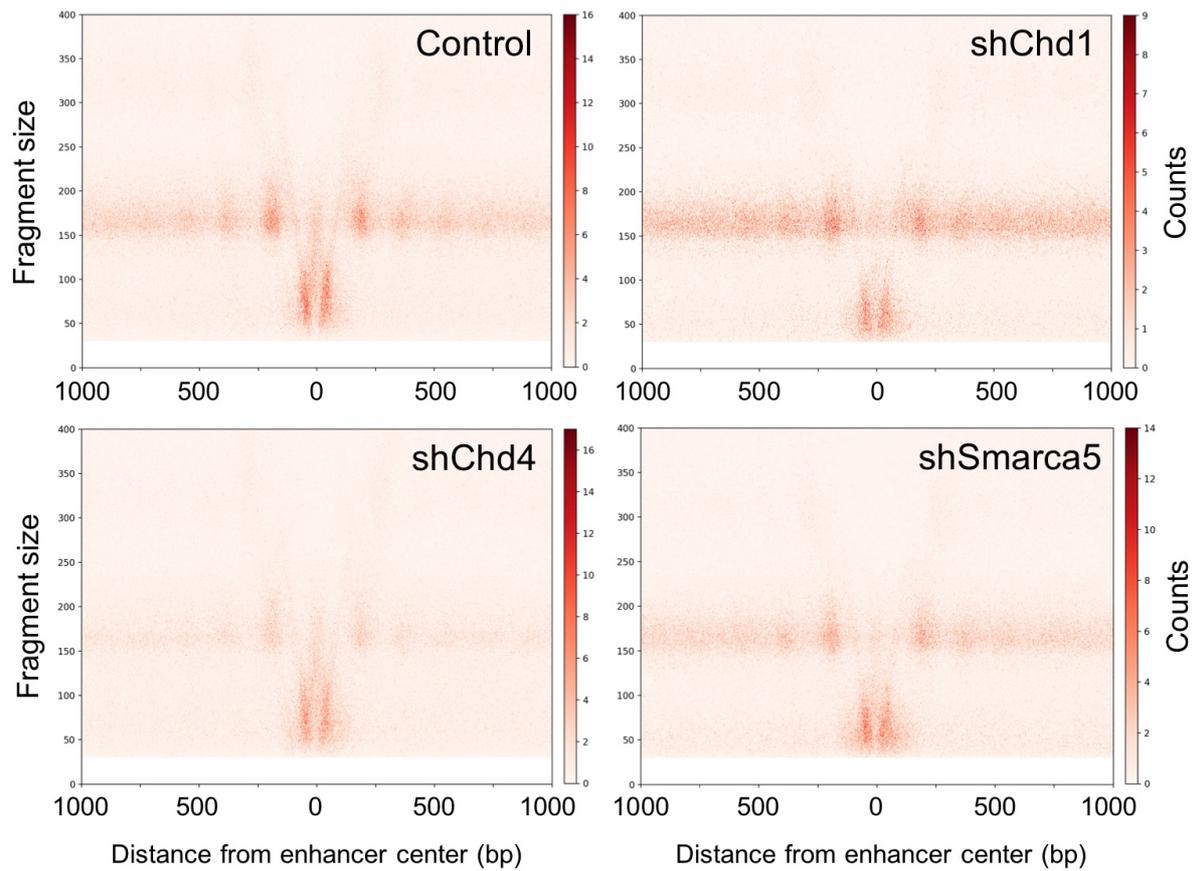
See the legend of Figure 1 for the description of the V-plot representation.



Extended data Figure 4: shRNA- and auxin-mediated depletion of the Brg1 ATPase.

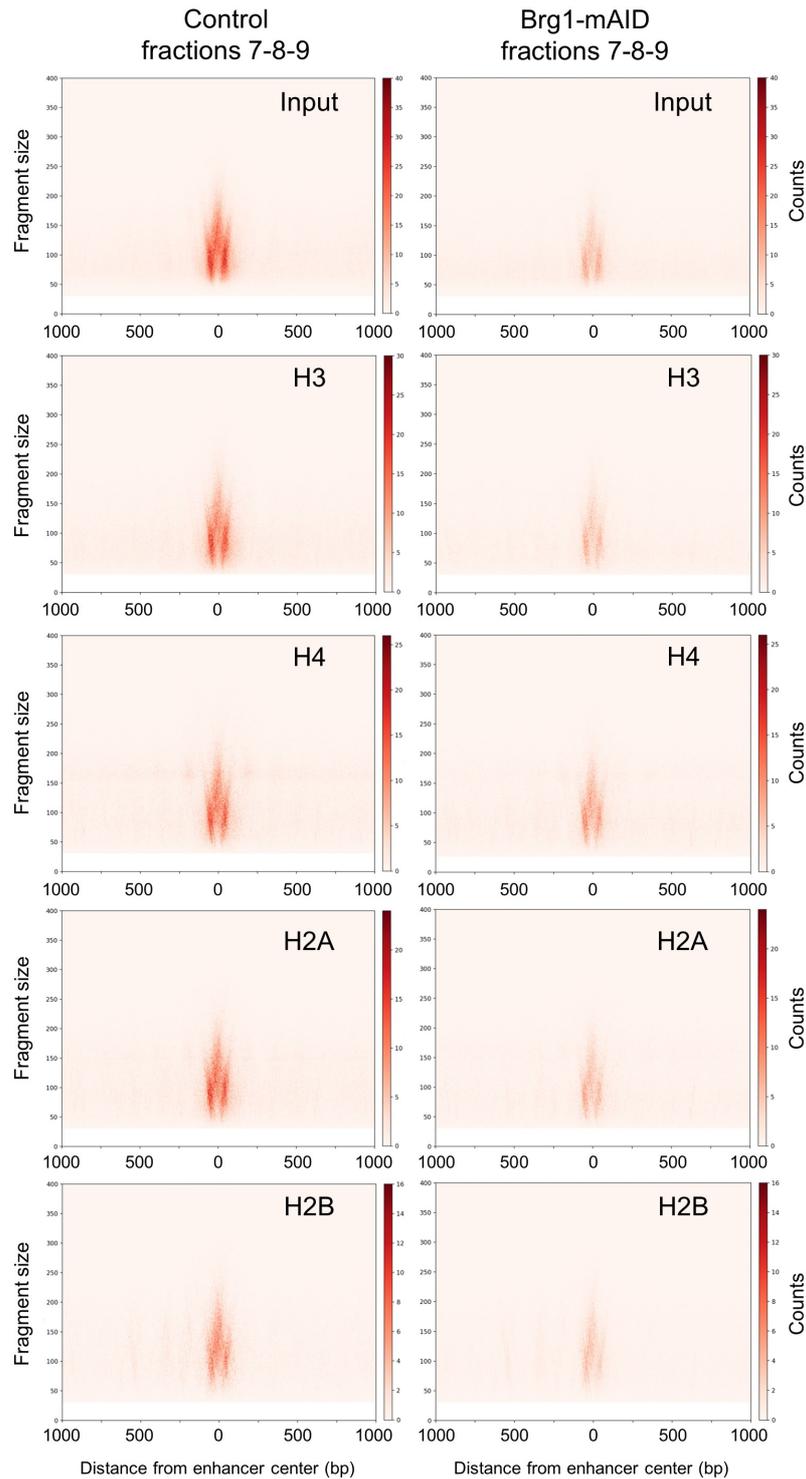
a. Western blot of protein extracts obtained from mESCs transfected with a control vector or shRNA O7 targeting Brg1, during 72 h. Two replicates are shown. GAPDH protein levels serve as a loading control.

b. Western blot of protein extracts obtained from Brg1-mAID mESCs treated with 1 mM auxin (Aux) during 3 h or 20 h. Ponceau staining of the membrane is shown as a loading control



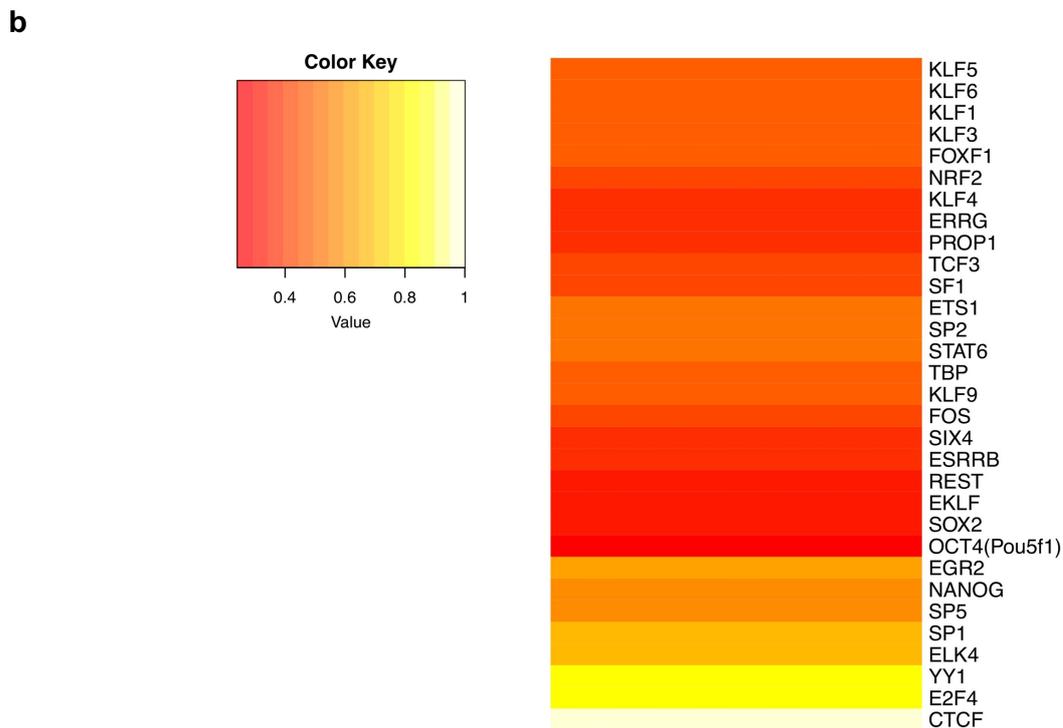
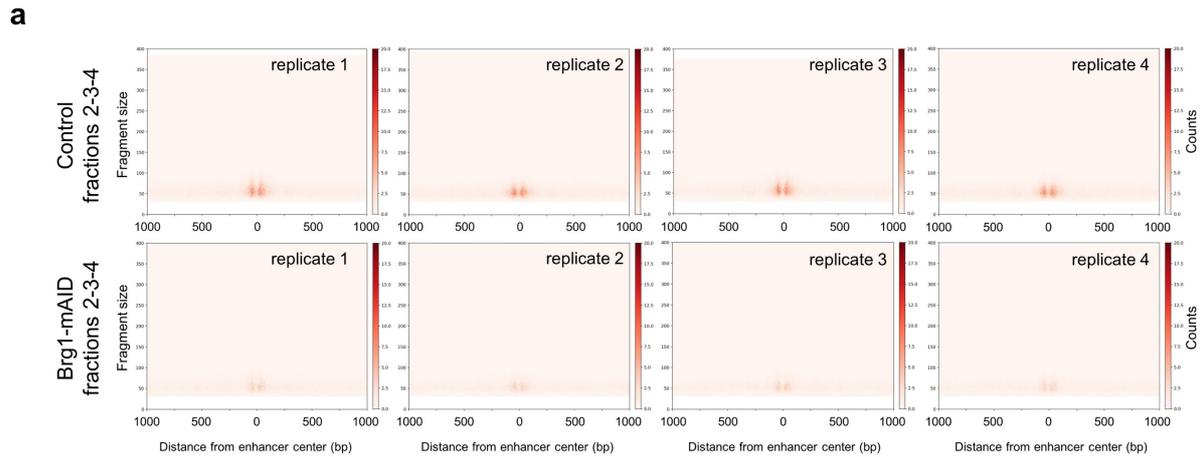
Extended data Figure 5: Analysis of nucleosomal and subnucleosomal organization at group 1 enhancers in Chd1, Chd4 and Smarca5-depleted ES cells.

H3 ChIP-seq was performed on MNase-digested control or remodeler-depleted chromatin. Chd1, Chd4 and Smarca5 chromatin remodelers were depleted using the corresponding shRNA constructs. See the legend of Figure 1 for the description of the V-plot representation.



Extended data Figure 6: Brg1 depletion affects the generation of 80-130 subnucleosomal particles at enhancers.

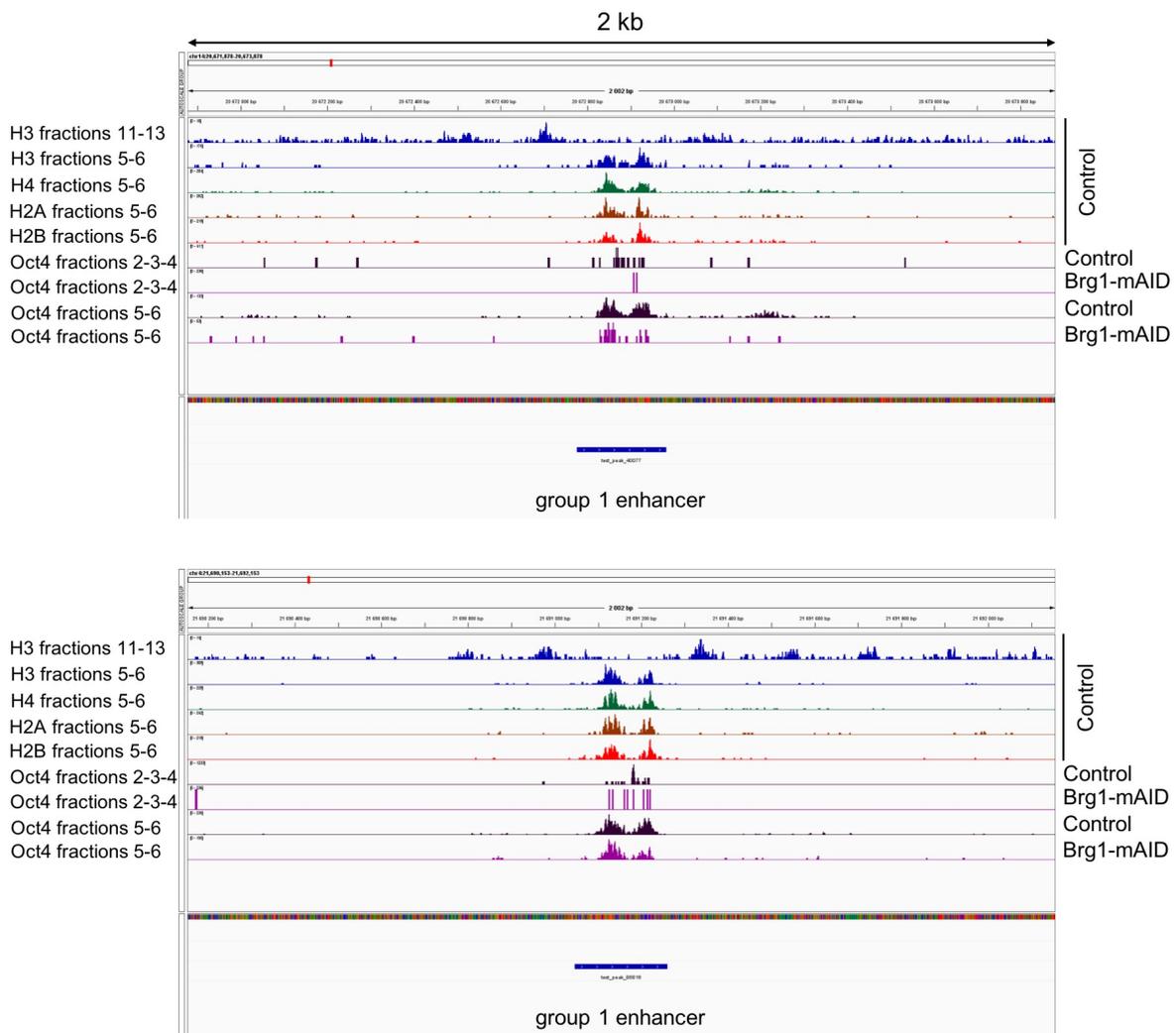
V-plots of sequenced input DNA or ChIP-seq experiments performed with antibodies against histones H3, H4, H2A and H2B. Chromatin was prepared from either control cells (left) or from Brg1-depleted cells (20 h Aux treatment of Brg1-mAID cells, right), and centrifugated through a 10-30% sucrose gradient. Fractions 7-8-9 were pooled and used as input chromatin for the ChIP experiments, or for direct sequencing (Input DNA). The V-plot span +/- 1000 bp from group 1 enhancer centre.



Extended data Figure 7: Brg1 is required for the production of histone-free, accessible DNA at enhancers.

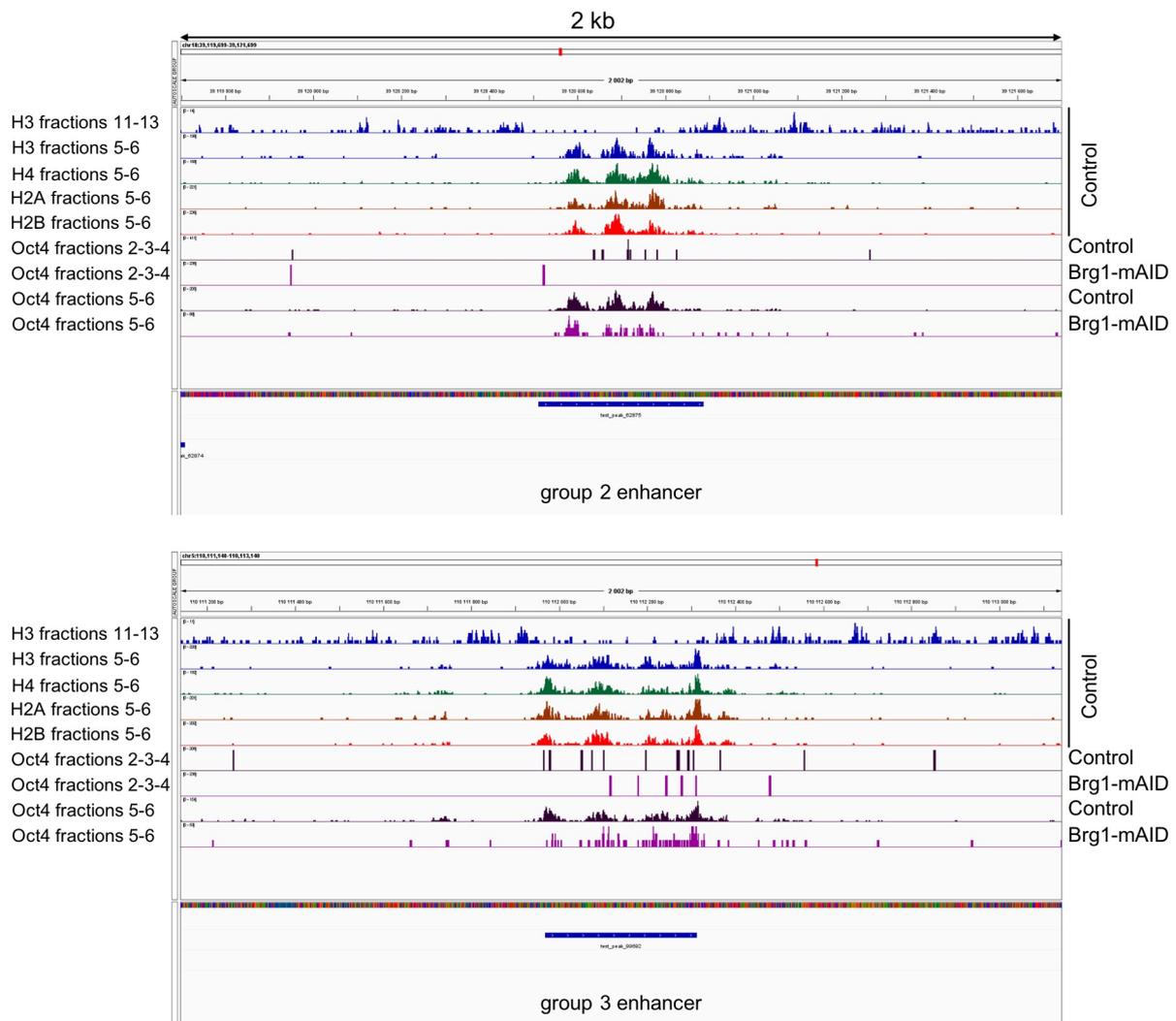
a. Deep sequencing of the DNA fragments which sedimented in fractions 2-3-4 was performed after centrifugation of MNase-treated chromatin prepared from either control ES cells or ES cells depleted of Brg1 (20 h Aux treatment of Brg1- mAID cells, lower panel).

b. Motif detection with the HOMER tool on the DNA fragments which sedimented in fractions 2-3-4 described in **a**. A ratio between the number of CTCF motifs (which are unaffected by Brg1-depletion) in control versus Brg1-depleted samples was calculated and applied to motifs of the Brg1-depleted sample. A second ratio between the number of motifs in Brg1-depleted and control cells was calculated and visualized as a heatmap.



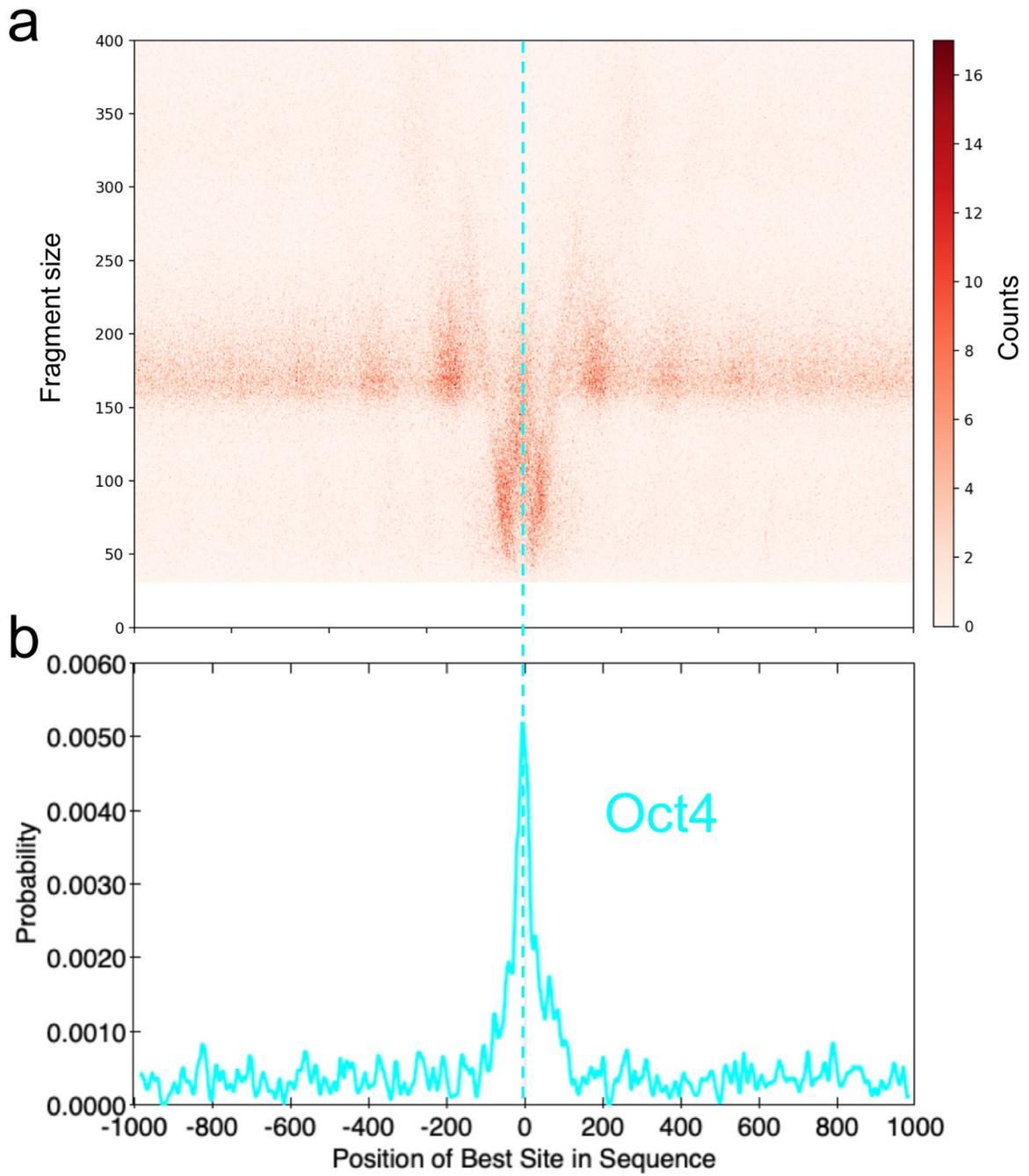
Extended data Figure 8: IGV screenshots of histones and Oct4 ChIP-seq enrichments at two different representative enhancers from cluster group 1.

The top lane of each panel shows the distribution of canonical nucleosomes isolated by sucrose gradient centrifugation (fractions 11 and 13). The next four lanes show histone H3, H4, H2A and H2B ChIP-seq patterns of a pool of sucrose gradient fractions 5-6, which contain subnucleosomal particles. The following four lanes correspond to Oct4 ChIP-seq of either a pool of sucrose gradient fractions 2-3-4 (lanes 6 and 7), or on a pool of fractions 5-6 (lanes 8 and 9). Brg1-depleted ES cells (20 h, 1mM auxin) were used for the experiments shown in lanes 7 and 9. All other lanes correspond to experiments realized with control (WT) ES cells expressing Brg1.



Extended data Figure 9: IGV screenshots of histones and Oct4 ChIP-seq enrichments at two different representative enhancers from groups 2 (top) and 3 (bottom).

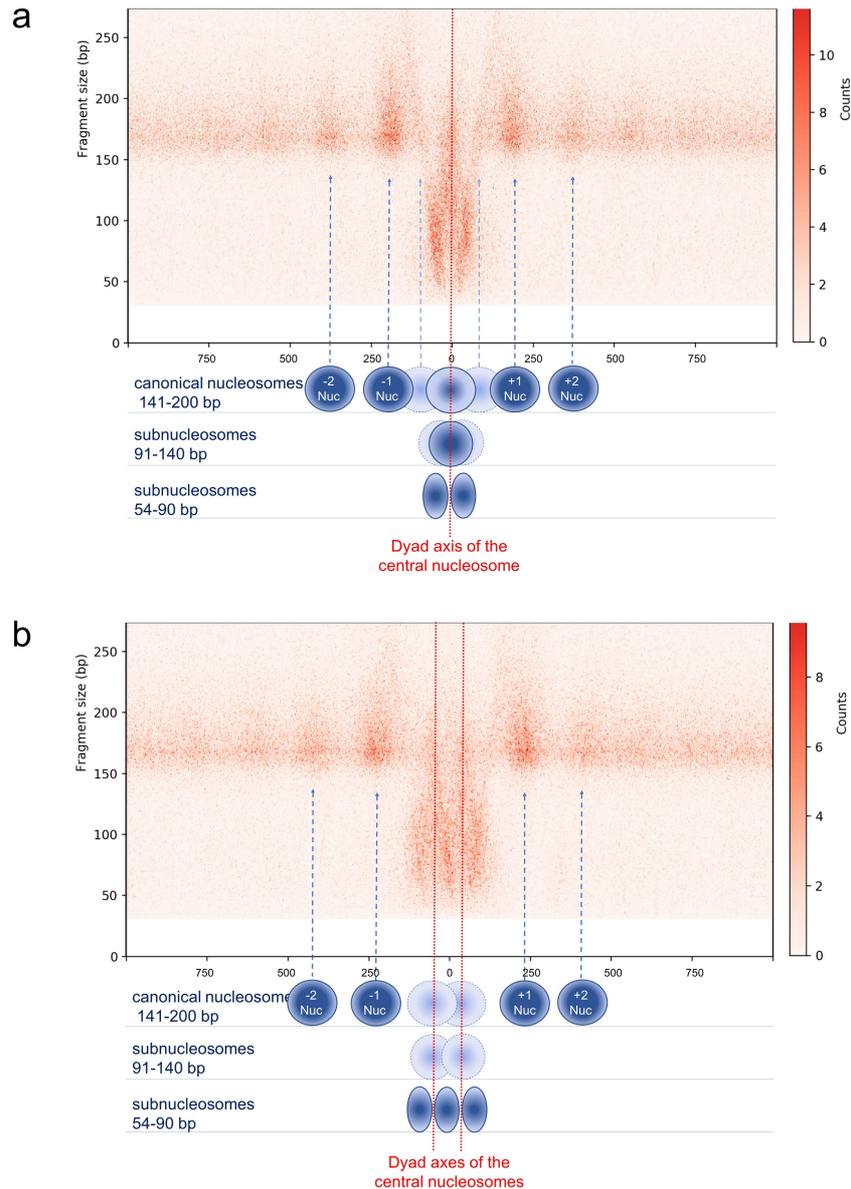
The top lane of each panel shows the distribution of canonical nucleosomes isolated by sucrose gradient centrifugation (fractions 11 and 13). The next four lanes show histone H3, H4, H2A and H2B ChIP-seq of a pool of sucrose gradient fractions 5-6, which contain subnucleosomal particles. The following four lanes correspond to Oct4 ChIP-seq of either a pool of sucrose gradient fractions 2-3-4 (lanes 6 and 7), or on a pool of fractions 5-6 (lanes 8 and 9). Brg1-depleted ES cells (20 h, 1mM auxin) were used for the experiments shown in lanes 7 and 9. All other lanes correspond to experiments realized with control (WT) ES cells expressing Brg1.



Extended data Figure 10: Oct 4 consensus motif sites enrichment coincides with the positioning of the central fragile nucleosome at group 1 enhancers.

(a) V-plot of histone H3 enrichment at group 1 enhancers.

(b) Enrichment of Oct4 consensus motif sites defined by the centrimo software.



Extended data Figure 11: Nucleosomal and subnucleosomal organization of group 1 and group 2 enhancers.

At the center of group 1 enhancers (**a**), the two canonical nucleosomes (Nuc-1 and Nuc +1) are positioned such that they are flanking the basic module of one fragile nucleosome (in light blue at the level of canonical nucleosomes), whereas enhancers from group 2 (**b**) contain two fragile nucleosomes at the center, between the well positioned nucleosomes -1 and +1. Fragile nucleosomes are targeted by the SWI/SNF ATPase BRG1, leading to their remodelling into subnucleosomal particles associated with DNA which is either 91-140 bp long or particles associated 54-90 bp DNA fragments, as well as shorter histone-free DNA fragments. Subnucleosomes containing 54-90 bp DNA are composed of all four histones (H3, H4, H2A and H2B) which leads to the hypothesis that those particles could be so called half nucleosomes or “hemisomes” composed of one copy of each histone protein. Furthermore, these particles are genome-wide the main regions of Oct4 enrichment and their positioning corresponds exactly to regions of enrichment of various TFs, highlighting their importance in enhancer function in mESC.

II. Résumé des résultats de l'article :

Dans ce manuscrit, nous décrivons comment les remodeleurs de la chromatine régulent l'organisation nucléosomale et subnucléosomale au niveau des enhancers dans les cellules ES de souris. Nous avons cartographié les nucléosomes et particules subnucléosomales grâce à des expériences de CHIP-seq utilisant des anticorps anti-histones sur de la chromatine digérée par une quantité modérée de MNase. Nous avons analysé les données de CHIP-seq d'une part sur *heatmaps* et d'autre part sur des représentations en deux dimensions (V-plots). Nous avons montré que les enhancers ont une région centrale sujette à des événements de remodelage de chromatine qui produisent des particules subnucléosomales de différentes tailles. Cette région centrale de l'enhancer est également caractérisée par la présence d'un nombre variable de nucléosomes fragiles, qui coïncident en positionnement sur le génome avec la localisation des particules subnucléosomales. Nos travaux ont permis de révéler une organisation modulaire des enhancers, chaque module étant composé d'un nucléosome fragile associé, avec un positionnement très précis, à deux particules subnucléosomales composées des quatre histones H3, H4, H2A et H2B et d'un fragment d'ADN de longueur comprise entre 50 et 90 pb. La région centrale de l'enhancer comporte donc un nombre variable de ces modules nucléosome fragile/subnucléosomes. Nous avons identifié trois groupes majeurs d'enhancers selon le nombre de modules. Le groupe 1 comporte un seul module, le groupe 2, deux modules, et le groupe 3, trois modules ou plus. Dans tous ces groupes d'enhancers, la région centrale modulaire est flanquée par deux nucléosomes canoniques bien positionnés, appelés nucléosomes -1 et +1. Le positionnement régulier des nucléosomes est visible au moins jusqu'aux nucléosomes adjacents (-2 et +2).

Nous avons analysé les profils de liaison des TFs associés à la pluripotence, ainsi que les profils des autres acteurs de la transcription typiquement présent au niveau des enhancers. Notre objectif consistait ici à définir comment ces TFs ou enzymes interagissent avec ces différentes particules nucléosomales et subnucléosomales, ainsi qu'au niveau de l'ADN dit accessible (qui n'est pas associé aux histones). Pour cela, nous avons réalisé des expériences de CHIP-seq sur de la chromatine digérée par la MNase en utilisant des anticorps ciblant ces différents TFs, la Pol II, H3K27ac et une sous-unité de la cohésine, SMC3. Nous observons que les TFs associés à la pluripotence (OCT4, SOX2, NANOG), ainsi que les composants de machinerie basale de

transcription (TBP, Pol II) se lie principalement dans la région centrale de l'enhancer au niveau des particules subnucléosomales.

Les facteurs pionniers OCT4 et SOX2 se lient principalement au niveau des particules subnucléosomales ayant une taille de fragments d'ADN comprise entre 50 et 120 pb, plutôt qu'au niveau des nucléosomes canoniques (fragments d'ADN compris entre 150 et 180 pb).

Nous avons ensuite analysé le rôle du remodelleur BRG1 dans la régulation de l'organisation nucléosomale aux enhancers. Grâce à des expériences de CHIP-seq utilisant des anticorps anti-H3 et anti-H2B dans des cellules déplétées en BRG1, nous avons démontré que la déplétion de BRG1 entraîne une forte réduction du signal des particules subnucléosomales (évalué relativement au signal des nucléosomes canoniques flanquant la région centrale de l'enhancer). Nous avons également observé une altération du positionnement des nucléosomes canoniques flanquant la région centrale de l'enhancer : en l'absence de BRG1, leur positionnement strict est perdu, et les nucléosomes couvrent l'ensemble de l'enhancer avec un pattern flou. Grâce à des expériences de MNase-CHIP-seq, nous avons révélé l'enrichissement de BRG1 au niveau des nucléosomes fragiles et des particules subnucléosomales, ce qui permet de proposer un modèle dans lequel BRG1 se lie au niveau du nucléosome central de l'enhancer pour le remodeler en nucléosome fragile, puis en particules subnucléosomales plus petites. Cette fonction de BRG1 aux enhancers est très spécifique, car la déplétion d'autres remodelleurs tels que CHD1, CHD4 ou SMARCA5 n'a pas d'effet apparent sur l'organisation nucléosomale aux enhancers.

Nous avons étudié plus précisément la nature des particules subnucléosomales en testant leurs propriétés de sédimentation dans des expériences de centrifugation sur gradient de sucrose en présence et en absence de BRG1. Dans ces expériences, les particules nucléosomales sédimentent selon leur forme et selon leur poids moléculaire. Plus une particule a une masse moléculaire élevée, plus elle sédimentera vite. La vitesse de sédimentation des particules subnucléosomales remodelées par BRG1 aux enhancers s'est révélée nettement inférieure à celle des nucléosomes canoniques. Nous avons analysé la littérature existante sur les particules subnucléosomales pour mieux caractériser leur nature. D'après leurs propriétés de sédimentation, ces particules subnucléosomales pouvaient

potentiellement correspondre à des hexasomes ((H3/H4)₂, H2A/H2B), des tétrasomes ((H3/H4)₂) ou des hémisomes (H3/H4/H2A/H2B).

Pour explorer ces hypothèses, nous avons commencé par réaliser des expériences de CHIP-seq sur les différentes fractions du gradient collectées après centrifugation, en utilisant des anticorps dirigés contre chacune des quatre histones H2A, H2B, H3 et H4. Ces expériences nous ont démontré que les particules subnucléosomales contenant des fragments d'ADN de 50 à 90 pb sont associées avec les quatre histones, ce qui exclue donc l'hypothèse du tétrasome. Par ailleurs, la taille des fragments d'ADN associés (50-90 pb) protégés de la digestion par la MNase n'était pas en phase avec les résultats obtenus par le groupe du Dr Henikoff, qui avait caractérisé les particules de type hexasome présentes aux promoteurs : ces travaux avaient permis de révéler des particules dont les histones protégeaient de la digestion par la MNase des fragments d'ADN de 90 à 130 pb. Or, nous observons également des particules de ce type aux enhanceurs, mais elles sédimentent plus rapidement dans le gradient de sucrose que les particules de 50-90 pb. L'analyse de la position du centre des fragments d'ADN des particules subnucléosomales de 50-90 pb révèle leur distribution symétrique de part et d'autre de l'axe de la dyade du nucléosome fragile central. De plus, lorsqu'on additionne la longueur d'ADN enroulée autour de chacune des deux particules subnucléosomales, on obtient la longueur d'ADN enroulée autour d'un octasome (130-180 pb). L'ensemble de ces résultats et analyses suggère que les particules subnucléosomales 50-90 bp correspondent à des hémisomes, qui proviennent du remodelage par BRG1 du nucléosome fragile central. Nos résultats sont tous compatibles avec un scénario dans lequel BRG1 se lierait au niveau du nucléosome central pour le convertir en deux particules de type hémisomes qui resteraient associées ensemble via un segment d'ADN de type *linker*, très vulnérable à la digestion par la MNase. Ce type de particules évoque les demi-nucléosomes proposés dans quelques publications des années 1970-1980, qui avait été détectées dans les régions actives au niveau transcriptionnel.

Nous avons ensuite analysé en détail comment le facteur pionnier OCT4 se lie *in vivo* sur ces particules subnucléosomales. Cette analyse nous a conduit à une conclusion très étonnante et inédite : *in vivo*, le facteur OCT4 interagit avec ces particules avec un enrichissement nettement meilleur qu'avec les nucléosomes canoniques. De plus, lorsqu'on compare le niveau d'enrichissement de OCT4 sur l'ADN accessible (sans histone associée) des enhanceurs,

nous avons constaté que OCT4 est plus abondant sur les particules subnucléosomales de 50-90 pb, que sur l'ADN nu. Nous avons mis en œuvre une expérience de CHIP séquentiel en deux étapes (CHIP tandem) qui a permis de confirmer par une approche supplémentaire la présence simultanée de OCT4 et des histones au sein des particules subnucléosomales de 50-90 pb.

Nos résultats ont progressé de manière importante la compréhension des mécanismes de remodelage de chromatine par les facteurs SWI/SNF. Nous proposons également un nouveau modèle d'organisation et de fonctionnement des enhancers dans lequel le remodeleur BRG1 cible spécifiquement certains nucléosomes de la région centrale de l'enhancer pour les convertir en particules subnucléosomales de type hémisomes. Ces particules subnucléosomales sont le substrat préférentiel pour la fixation de différents types de TFs, y compris les TFs ayant des propriétés de facteurs pionniers tels que OCT4.

Partie B : Étude de la fonction de HDAC2 dans la maladie d'Alzheimer

I. Les différents modèles de souris utilisés au cours de la thèse

Les travaux du groupe du Dr Li-Huei Tsai à l'Université de Harvard avaient mis en évidence une composante épigénétique contribuant au déclin cognitif des patients atteints de la maladie d'Alzheimer. D'une part, l'histone désacétylase HDAC2 avait été révélée comme un répresseur transcriptionnel des gènes impliqués dans la formation de la mémoire, dans l'hippocampe de la souris (Guan et al. 2009). D'autre part, sur la base de travaux menés avec deux modèles murins de la maladie d'Alzheimer, ce même groupe a révélé un phénomène de surexpression de HDAC2 dans l'hippocampe coïncidant avec la dégradation des capacités d'apprentissage des souris (Graff et al. 2012). Ces auteurs avaient conclu à l'existence d'un blocage épigénétique affectant la transcription dans le cerveau des patients atteints de la maladie d'Alzheimer.

Une partie de mon travail de thèse a consisté à caractériser à l'échelle du génome entier le blocage épigénétique HDAC2-dépendant dans le contexte de la maladie d'Alzheimer, en utilisant deux modèles de souris alternatifs de la maladie d'Alzheimer. Un des buts de ce projet consistait à identifier extensivement les gènes cibles de HDAC2 dans l'hippocampe de souris, avec pour perspective d'identifier les gènes susceptibles d'être utilisés comme cibles thérapeutiques pour contrecarrer le déclin cognitif.

Le premier modèle que j'ai utilisé est la lignée 3xTg-AD (Oddo et al. 2003). Ce modèle de souris, qui a souvent été utilisé comme modèle de la maladie d'Alzheimer, a été créé en introduisant les transgènes humains exprimant les allèles $APP^{K670N/M671L}$ et Tau^{P301L} dans des embryons de souris mutants (*Knock-in*) pour le gène de la préséniline ($Psen1^{M146V}$) (**Tableau 6**) (Oddo et al. 2003). Ces souris présentent une hyperphosphorylation des protéines TAU et des enchevêtrements neurofibrillaires, des plaques β -amyloïdes et des déficits cognitifs (Oddo et al. 2003).

Le second modèle correspond à la lignée $App^{tm3(NL-G-F)}$ (Saito et al. 2014), appelée NLGF par la suite. Ces souris combinent trois mutations dans le gène *APP* murin dont la séquence a été humanisée à certains endroits : la mutation suédoise $APP^{KM670/671NL}$, la mutation Beyreuther/ibérique APP^{I716F} et la mutation arctique APP^{E693G} (**Tableau 6**) (Saito et al. 2014).

Ces trois mutations ne se retrouvent pas naturellement présentes ensemble chez les mêmes patients, mais leur association chez ce modèle permet d'avoir chez des souris âgées de 9 mois une série d'altérations mimant celles des formes génétiques de la maladie d'Alzheimer chez l'Homme. Nous avons choisi ce modèle récent de la maladie d'Alzheimer car les protéines APP et TAU ne sont pas surexprimées, ce qui limite les effets secondaires non spécifiques de la maladie d'Alzheimer qui ont été détectés dans les premières générations de modèles murins (Saito et al. 2014). Ce modèle est caractérisé par des altérations plus proches des véritables défauts induits par la maladie d'Alzheimer chez l'Homme.

Les souris NLGF subissent des défauts cognitifs, notamment de la mémoire spatiale et de l'apprentissage, caractéristiques de la maladie d'Alzheimer et elles accumulent rapidement et fortement des dépôts de plaques β -amyloïdes et des réponses gliales (Saito et al. 2014; Masuda et al. 2016). Ces souris ne présentent pas d'hyperphosphorylation des protéines TAU, d'enchevêtrements neurofibrillaires, ni de perte neuronale (Saito et al. 2014).

Pour faciliter les expériences d'identification des gènes cibles de HDAC2 par ChIP-seq dans le cerveau de souris, nous avons généré au laboratoire une lignée de souris dans laquelle un des deux allèles du gène *Hdac2* a été étiqueté en C-terminal par les épitopes HA et Flag (HF) (de Dieuleveult et al. 2016). La séquence d'ADN codant une seule copie de chacune des étiquettes HA et FLAG a été introduite immédiatement en 3' de la séquence codante de *Hdac2* par recombinaison homologe en cellules ES. Ces cellules ES ont servi à dériver la lignée de souris *Hdac2*^{HF/+} (Tableau 6). Par croisement entre souris hétérozygotes, nous avons obtenu des souris homozygotes qui ont servi à réaliser des expériences de ChIP-seq pour identifier le pattern de liaison de HDAC2 sur ses gènes cibles dans le cerveau. La lignée *Hdac2*^{HF} a également été croisée avec des souris 3xTg-AD afin de suivre par ChIP-seq une éventuelle évolution de la distribution de HDAC2 sur le génome dans le contexte de ce modèle de la maladie d'Alzheimer. Plusieurs croisements consécutifs ont permis d'obtenir la lignée 3xTg-AD ; *Hdac2*^{HF/+}, qui est homozygote pour le *Knock-In* et les transgènes composant le modèle 3xTg-AD (Tableau 6). La lignée 3xTg-AD ; *Hdac2*^{HF/+} m'a également permis de suivre l'expression des deux allèles de *Hdac2* lorsque j'ai évalué le niveau de surexpression de cette enzyme chez le modèle 3xTg-AD.

Afin d'évaluer l'impact de la protéine HDAC2 sur le transcriptome des souris saines et chez le modèle 3xTg-AD, nous avons utilisé un allèle mutant de *Hdac2*. Cet allèle *Hdac2* mutant a été obtenu par insertion aléatoire dans le génome d'une cassette de *gene-trap* composée d'un site accepteur d'épissage situé en amont de la séquence du gène rapporteur *LacZ* fusionné au gène de résistance à la néomycine, suivi d'un signal de polyadénylation nécessaire pour terminer la transcription (Wurst et al. 1995). Cette cassette est insérée, dans notre lignée de souris, au niveau de l'intron 8 du gène *Hdac2*. On estime qu'il s'agit d'un allèle KO car une partie du domaine catalytique C-terminal est délétée. Les souris porteuses de cet allèle (*Hdac2*^{+/−}) ont été croisées avec les souris 3xTg-AD afin d'obtenir les souris 3xTg-AD ; *Hdac2*^{−/−} (Tableau 6).

Les génotypes de toutes nos souris (les deux modèles de la maladie d'Alzheimer et les autres lignées qui en découlent) ont été validés par qPCR avant de réaliser les différentes expériences, ainsi que par des méthodes bioinformatiques après analyses par RNA-seq.

Lignée de souris	gène <i>Hdac2</i>	gène <i>APP</i>	gène <i>Tau</i>	gène <i>Psen1</i>
3xTg-AD	<i>Hdac2</i> ^{+/+}	<i>APP</i> ^{K670N/M671L}	<i>Tau</i> ^{P301L}	<i>Psen1</i> ^{M146V}
<i>Hdac2</i> ^{-/-}	<i>Hdac2</i> ^{-/-}	<i>APP</i> ^{+/+}	<i>Tau</i> ^{+/+}	<i>Psen1</i> ^{+/+}
3xTg-AD ; <i>Hdac2</i> ^{-/-}	<i>Hdac2</i> ^{-/-}	<i>APP</i> ^{K670N/M671L}	<i>Tau</i> ^{P301L}	<i>Psen1</i> ^{M146V}
<i>Hdac2</i> ^{HF/+}	<i>Hdac2</i> ^{HF/+}	<i>APP</i> ^{+/+}	<i>Tau</i> ^{+/+}	<i>Psen1</i> ^{+/+}
3xTg-AD ; <i>Hdac2</i> ^{HF/+}	<i>Hdac2</i> ^{HF/+}	<i>APP</i> ^{K670N/M671L}	<i>Tau</i> ^{P301L}	<i>Psen1</i> ^{M146V}
C57	<i>Hdac2</i> ^{+/+}	<i>APP</i> ^{+/+}	<i>Tau</i> ^{+/+}	<i>Psen1</i> ^{+/+}
App ^{tm3} (NL-G-F)	<i>Hdac2</i> ^{+/+}	<i>APP</i> ^{KM670/671NL} (Swedish) <i>APP</i> ^{I716F} (Beyreuther/Iberian) <i>APP</i> ^{E693G} (Artic)	<i>Tau</i> ^{+/+}	<i>Psen1</i> ^{+/+}

Tableau 6 : Les différentes lignées de souris utilisées dans ce projet

Bilan des différentes lignées de souris utilisées pendant ma thèse avec leurs génotypes pour les gènes *Hdac2*, *APP*, *Tau* et *Psen1*. Pour les gènes *APP*, *Tau* et *Psen1*, les mutations sont indiquées en rouge et les transgènes en bleu.

II. HDAC2 est modérément surexprimée dans un modèle de souris de la maladie d'Alzheimer

Les travaux de l'équipe du Dr Tsai suggéraient un rôle clé de HDAC2 dans l'altération des fonctions cognitives, à travers une surexpression de HDAC2 dans l'hippocampe, liée à la progression de la maladie d'Alzheimer (Gräff et al. 2012). Nous avons commencé par déterminer le niveau d'expression de la protéine HDAC2 dans nos souris modèles de la maladie d'Alzheimer au cours du vieillissement.

1) Modèle 3xTg-AD

Dans un premier temps, les hippocampes de souris 3xTg-AD ; *Hdac2*^{HF/+} et de souris *Hdac2*^{HF/+} ont été prélevés à différents âges : 2 mois, 3 mois, 5 mois et 9 mois. Après extraction des protéines des hippocampes, des expériences de western blot ont permis de suivre l'expression d'une part de la protéine HDAC2 étiquetée HA-Flag, avec un anticorps anti-HA, et d'autre part son allèle non-étiqueté avec un anticorps anti-HDAC2.

Au total 48 souris ont été utilisées dans cette étude (**Figure 18**) :

- 14 souris âgées de 2 mois : 7 souris contrôles *Hdac2^{HF/+}* et 7 souris 3xTg-AD ; *Hdac2^{HF/+}*,
- 14 souris âgées de 3 mois : 7 souris contrôles *Hdac2^{HF/+}* et 7 souris 3xTg-AD ; *Hdac2^{HF/+}*,
- 8 souris âgées de 5 mois : 4 souris contrôles *Hdac2^{HF/+}* et 4 souris 3xTg-AD ; *Hdac2^{HF/+}*,
- 12 souris âgées de 9 mois : 6 souris contrôles *Hdac2^{HF/+}* et 6 souris 3xTg-AD ; *Hdac2^{HF/+}*.

Après la révélation des protéines exprimées par chaque allèle *Hdac2^{+/+}* ou *Hdac2^{HF/+}*, l'ensemble des protéines de la membrane a été coloré dans un but de normalisation (**Figure 18a panel du bas**).

Chez les souris 3xTg-AD âgées de 2 mois, les 7 souris testées surexpriment HDAC2 à des niveaux différents des contrôles lorsque l'allèle étiqueté est suivi et 6 souris sur 7 surexpriment HDAC2 lorsque l'allèle non-étiqueté est suivi (**Figure 18b haut gauche et droite**). L'expression des deux allèles codant HDAC2 est donc également altérée chez ce modèle. Les 7 souris 3xTg-AD âgées de 2 mois surexpriment HDAC2 de façon homogène par rapport aux souris contrôles.

Chez les souris 3xTg-AD âgées de 3 mois, 1 souris sur 7 souris surexpriment assez fortement HDAC2 par rapport aux souris contrôles lorsque l'allèle étiqueté est suivi et 4 souris sur 7 surexpriment HDAC2 plus fortement que les contrôles lorsque l'allèle non-étiqueté est suivi (**Figure 18b haut gauche et droite**). Il existe donc une variabilité inter-individu parmi les 7 souris 3xTg-AD âgées de 3 mois.

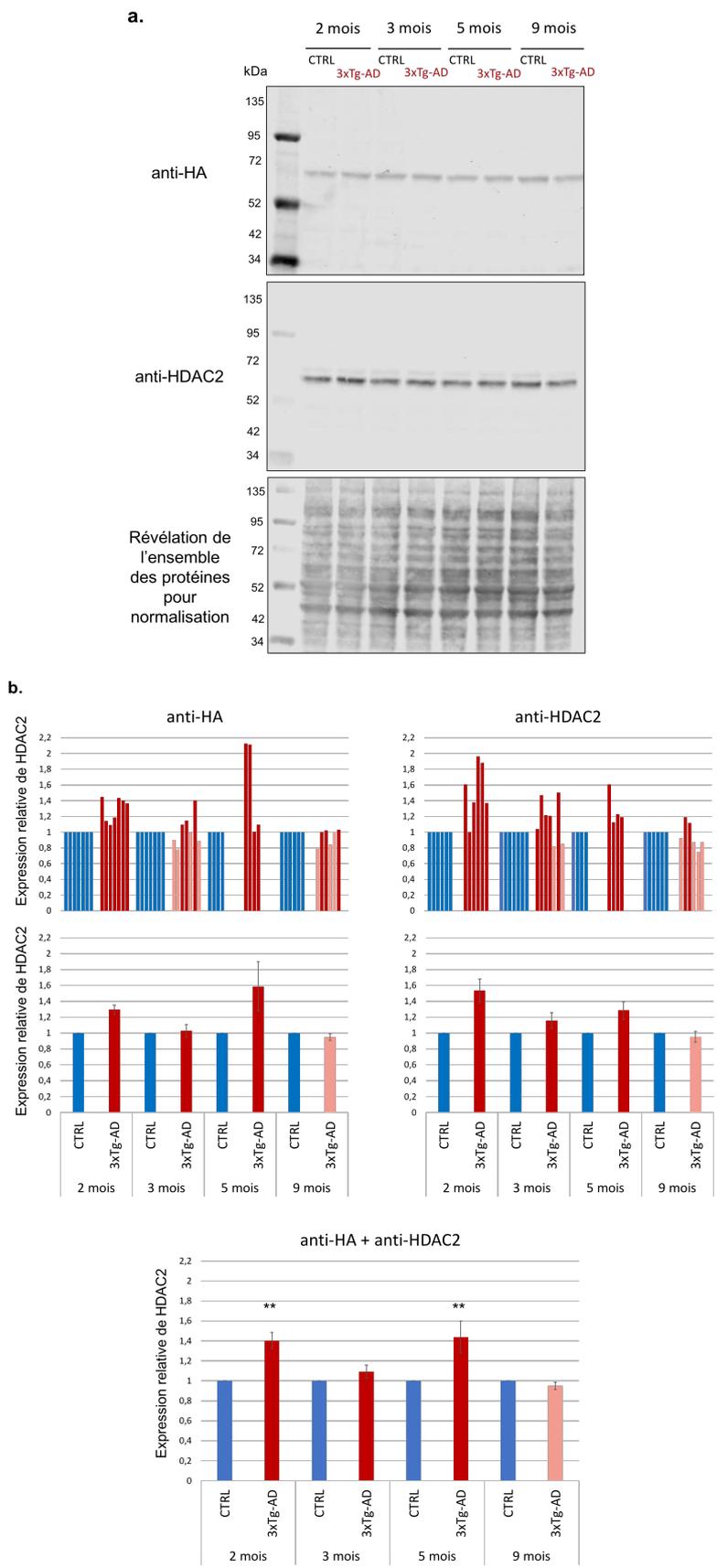


Figure 18 : Suivi de l'expression de la protéine HDAC2 au cours du temps chez les souris contrôles et les souris 3xTg-AD modèles de la maladie d'Alzheimer

Légende Figure 18 :

a. L'expression de l'allèle *Hdac2* étiqueté par HA-Flag est suivi par western blot avec un anticorps anti-HA (panel du haut) et l'allèle *Hdac2* non-étiqueté est suivi par western blot avec un anticorps anti-HDAC2 (panel du centre). L'expression de la protéine HDAC2 est suivie au cours du temps chez des souris *Hdac2*^{HF/+} contrôles (CTRL) et des souris 3xTg-AD ; *Hdac2*^{HF/+} (3xTg-AD) à 2 mois, 3 mois, 5 mois et 9 mois. Ces western blots montrent l'expression de HDAC2 chez une série de souris CTRL âgées de 2, 3, 5 et 9 mois et une série de souris 3xTg-AD âgées de 2, 3, 5 et 9 mois. L'ensemble des protéines est révélé par une coloration de la membrane pour la normalisation des données lors de la quantification (panel du bas).

b. L'expression de HDAC2 a été quantifiée pour les souris CTRL (n = 7) et 3xTg-AD (n = 7) âgées de 2 mois, les souris CTRL (n = 7) et 3xTg-AD (n = 7) âgées de 3 mois, les souris CTRL (n = 4) et 3xTg-AD (n = 4) âgées de 5 mois et les souris CTRL (n = 6) et 3xTg-AD (n = 6) âgées de 9 mois (panels de gauche). L'étude comprend au total 48 souris. Les souris contrôles sont représentées en bleu, les souris 3xTg-AD dont l'expression de HDAC2 est supérieure à une fois le niveau des souris contrôles sont représentées en rouge foncé et les souris 3xTg-AD dont l'expression de HDAC2 est inférieure à une fois le niveau des contrôles sont représentées en rose. Une moyenne de l'expression de HDAC2 entre individus de même génotype et de même âge a été faite pour les allèles anti-HA et anti-HDAC2 (panels du milieu). La moyenne des quantifications de l'expression de HDAC2 pour les allèles étiquetés et non-étiquetés a également été faite (panel de droite) et le test statistique des rangs signés de Wilcoxon pour échantillons appariés a été effectué pour déterminer si les différences d'expression de HDAC2 entre les souris CTRL et 3xTg-AD sont significatives ou non. ** : p-value < 0,01.

Chez les souris 3xTg-AD âgées de 5 mois, 2 souris sur 4 testées surexpriment fortement (supérieur à 2 fois) HDAC2 par rapport aux souris contrôles lorsque l'allèle étiqueté est suivi (**Figure 18b haut gauche**). Les 4 souris testées surexpriment HDAC2 à différents niveaux lorsque l'allèle non étiqueté est suivi (**Figure 18b haut droite**). Les 4 souris 3xTg-AD âgées de 5 mois surexpriment HDAC2 de façon homogène par rapport aux souris contrôles.

Chez les souris 3xTg-AD âgées de 9 mois, 4 souris sur 6 expriment HDAC2 avec la même intensité que les souris contrôles ou très légèrement plus lorsque l'allèle étiqueté est suivi et 2 souris sur 6 surexpriment légèrement HDAC2 lorsque l'allèle non-étiqueté est suivi (**Figure 18b haut gauche et droite**). Les souris 3xTg-AD âgées de 9 mois sont donc assez homogènes pour l'expression de HDAC2 par rapport aux souris contrôles.

La moyenne de l'expression de HDAC2 des souris de même génotype et de même âge a été calculée pour les deux allèles du gène *Hdac2*, ainsi que pour les deux allèles réunis. On remarque que les deux allèles de *Hdac2* varient globalement dans le même sens : on observe toujours une surexpression de HDAC2 à 2 mois et 5 mois chez les souris 3xTg-AD par rapport aux contrôles, une très légère surexpression à 3 mois et pas de surexpression de HDAC2 à 9

mois (**Figure 18b centre droite et gauche**). Le test statistique des rangs signés de Wilcoxon pour échantillons appariés a été effectué afin de déterminer si les légères surexpressions de HDAC2 sont significatives ou non.

Une surexpression significative d'environ 40 % de la protéine HDAC2 a été observée chez les souris modèles 3xTg-AD âgées de 2 mois et 5 mois par rapport aux souris contrôles (**Figure 18b bas**). La très légère surexpression de HDAC2 visible à 3 mois n'est pas statistiquement significative, probablement à cause de l'importante variabilité inter-individu. La protéine HDAC2 n'est pas significativement différentiellement exprimée chez les souris 3xTg-AD âgées de 9 mois par rapport aux souris contrôles (**Figure 18b bas**).

Comme notre étude comporte des souris mâles et femelles parmi les souris contrôles et les souris 3xTg-AD pour les âges de 2 mois, 3 mois et 9 mois, il me semblait important de vérifier si le sexe de la souris avait un impact sur l'expression de la protéine HDAC2 au cours du temps (**Figure 19**). Bien que les souris femelles 3xTg-AD semblent exprimer HDAC2 plus fortement que les souris mâles 3xTg-AD à 2 mois et à 3 mois, cette différence d'expression entre souris mâles et femelles n'est significative à aucun des âges testés (**Figure 19**). Le sexe de la souris n'a donc pas d'influence sur l'expression de la protéine HDAC2 au cours du temps chez les souris 3xTg-AD.

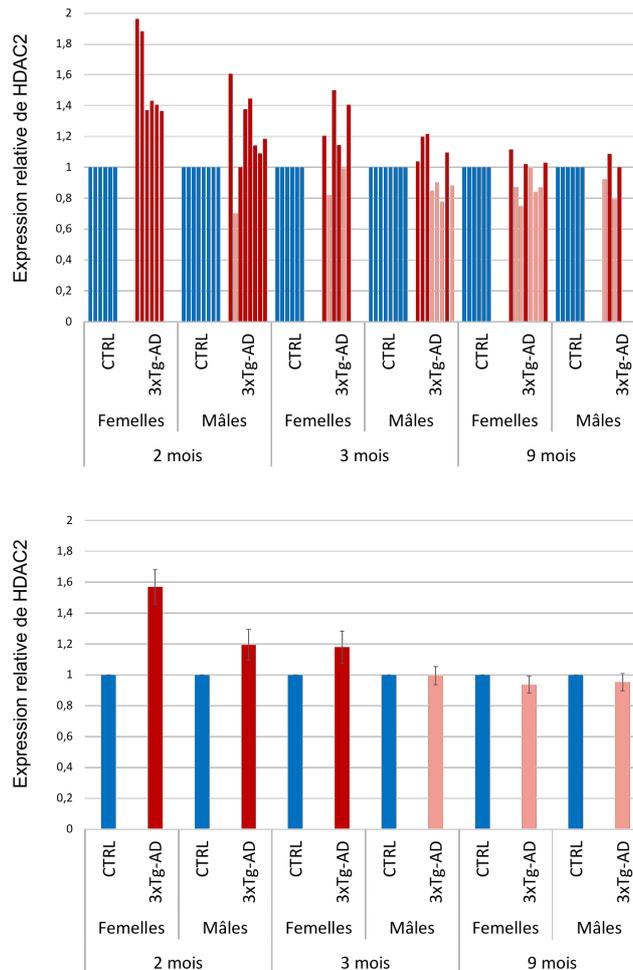


Figure 19 : Suivi de l'expression de la protéine HDAC2 au cours du temps chez les souris mâles et femelles contrôles et 3xTg-AD.

A partir des western blot réalisés sur les hippocampes de souris contrôles $Hdac2^{HF/+}$ et 3xTg-AD ; $Hdac2^{HF/+}$ (3xTg-AD) modèles de la maladie d'Alzheimer à 2 mois, 3 mois et 9 mois (voir figure 13), l'expression de la protéine HDAC2 a été quantifiée comme précédemment afin de comparer l'expression de HDAC2 entre les souris mâles et femelles 3xTg-AD. Cette étude comprend au total 40 souris : des souris femelles et mâles contrôles ($n_{femelles} = 3$; $n_{mâles} = 4$) et 3xTg-AD ($n_{femelles} = 3$; $n_{mâles} = 4$) âgées de 2 mois, des souris femelles et mâles contrôles ($n_{femelles} = 3$; $n_{mâles} = 4$) et 3xTg-AD ($n_{femelles} = 3$; $n_{mâles} = 4$) âgées de 3 mois, des souris femelles et mâles contrôles ($n_{femelles} = 3$; $n_{mâles} = 3$) et 3xTg-AD ($n_{femelles} = 4$; $n_{mâles} = 2$) âgées de 9 mois. Pour chacune des souris, les deux allèles de *Hdac2* (étiqueté et non étiqueté) sont suivis. Les souris contrôles sont représentées en bleu, les souris 3xTg-AD dont l'expression de HDAC2 est supérieure à une fois le niveau des souris contrôles sont représentées en rouge foncé et en rose pour une expression de HDAC2 inférieure à une fois le niveau des contrôles. L'expression de HDAC2, suivie grâce aux deux allèles, pour chacun des individus est présentée sur le panel du haut. Une moyenne de l'expression de HDAC2 entre individus de même génotype, de même âge et de même sexe a été faite pour les allèles étiquetés et non-étiquetés sur le panel du bas. Le test statistique de Mann-Whitney a été effectué pour déterminer si les différences d'expression de HDAC2 entre les souris 3xTg-AD mâles et femelles sont significatives ou non au cours du temps. Aucune différence d'expression de HDAC2 n'est significative (p -values > 0,37).

2) Modèle NLGF

Nous avons réalisé des expériences de western blot avec un anticorps anti-HDAC2 sur des extraits protéiques d'hippocampes de souris du modèle NLGF. Ces expériences de western blot ont permis de suivre l'expression de la protéine HDAC2 chez 24 souris femelles âgées de 3 mois (12 souris contrôles et 12 souris NLGF) et chez 24 souris femelles âgées de 9 mois (12 souris contrôles et 12 souris NLGF) (**Figure 20a**).

Que ce soit chez les souris NLGF âgées de 3 mois ou de 9 mois, le niveau d'expression de HDAC2 est assez hétérogène par rapport aux souris contrôles (**Figure 20b haut**). Une variabilité inter-individu significative est observée chez les souris NLGF.

L'expression moyenne de la protéine HDAC2 a été calculée pour les souris de même âge et de même génotype et le test statistique des rangs signés de Wilcoxon pour échantillons appariés a été effectué. Ces expériences n'ont mis en évidence aucune différence significative d'expression de la protéine HDAC2 chez les souris NLGF âgées de 3 mois et de 9 mois par rapport aux souris contrôles du même âge (**Figure 20b bas**).

Nous concluons qu'une surexpression sporadique et de faible amplitude de la protéine HDCA2 a été détectée dans l'hippocampe des souris modèles 3xTg-AD, mais aucune surexpression significative de HDAC2 associée à la progression de la maladie d'Alzheimer n'a été détectée chez les souris NLGF.

Or, le modèle NLGF est considéré comme un modèle de la maladie d'Alzheimer moins artificiel que le modèle 3xTg-AD ; en effet, le modèle NLGF est basé sur un KI de trois mutations au niveau du gène *APP*, tandis que le modèle 3xTg-AD comprend deux transgènes qui surexpriment à des niveaux non physiologiques les protéines APP et TAU pouvant conduire à des phénotypes artéfactuels (Oddo et al. 2003; Saito et al. 2014; Masuda et al. 2016).

La surexpression de la protéine HDAC2 pourrait être modèle-dépendante.

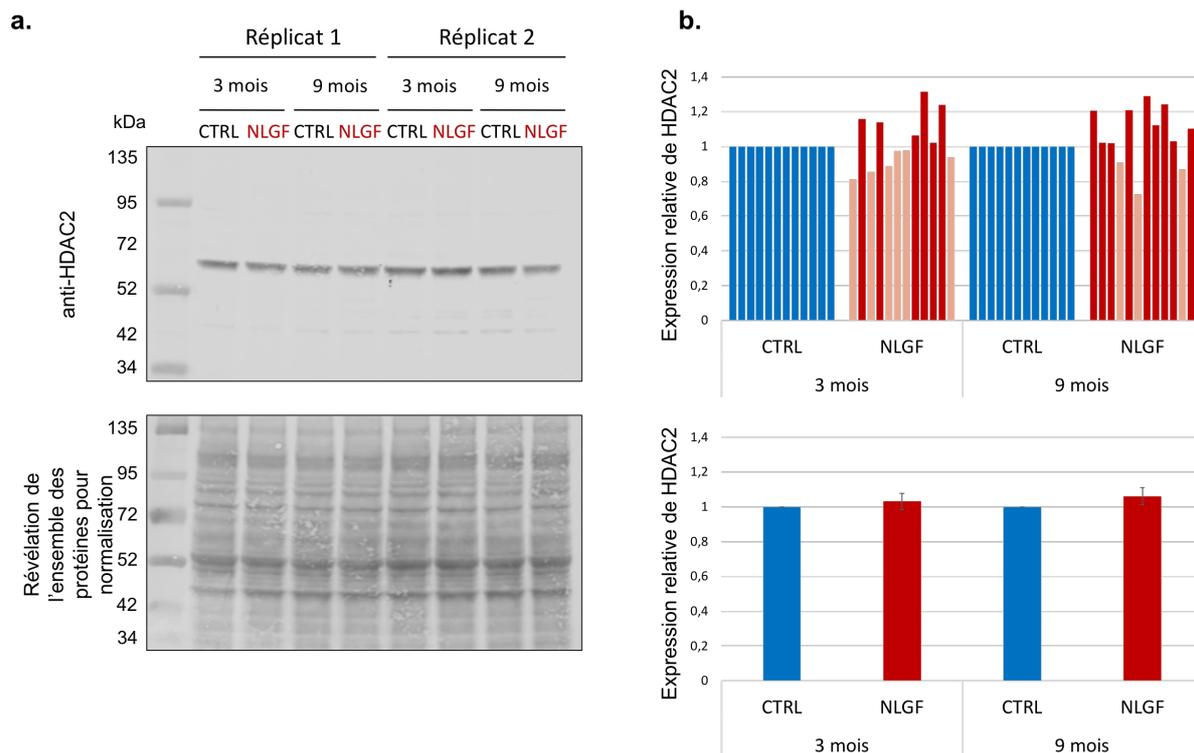


Figure 20 : Suivi de l'expression de la protéine HDAC2 au cours du temps chez les souris contrôles et les souris NLGF modèles de la maladie d'Alzheimer

a. L'expression de la protéine HDAC2 est suivie par western blot avec un anticorps anti-HDAC2 (panel du haut). L'expression de HDAC2 est suivie au cours du temps chez des souris C57 (CTRL) et des souris $App^{tm3(NL-G-F)}$ (NLGF) à 3 mois et 9 mois. Ce western blot présente deux répliquats du suivi de l'expression de HDAC2 chez une souris de chaque génotype et âge. Une coloration de toutes les protéines présentes sur la membrane (panel du bas) permet de normaliser les données lors que la quantification de l'expression de HDAC2.

b. L'expression de HDAC2 a été quantifiée pour les souris CTRL ($n_{3\text{ mois}} = 12$; $n_{9\text{ mois}} = 12$) et NLGF ($n_{3\text{ mois}} = 12$; $n_{9\text{ mois}} = 12$) âgées de 3 mois ou de 9 mois (panel du haut). L'étude comprend au total 48 souris. Les souris contrôles sont représentées en bleu, les souris NLGF dont l'expression de HDAC2 est supérieure à une fois le niveau des souris contrôles sont représentées en rouge foncé et celles dont l'expression de HDAC2 est inférieure à une fois le niveau des contrôles sont représentées en rose. Une moyenne de l'expression de HDAC2 entre individus de même génotype et de même âge a été faite (panel du bas) et le test statistique des rangs signés de Wilcoxon pour échantillons appariés a été effectué pour déterminer si les différences d'expression de HDAC2 entre les souris CTRL et NLGF sont significatives ou non. Aucune différence d'expression n'est significative (p-values > 0,30).

III. Caractérisation du profil de liaison de HDAC2 sur le génome du cerveau adulte

Dans le but d'évaluer la fonction potentielle de HDAC2 dans la régulation de la transcription dans le cerveau, nous avons étudié son profil de liaison sur le génome.

Dans le contexte sain, des expériences de CHIP-seq séquentiel (CHIP-tandem) sur des cerveaux de souris *Hdac2^{HF/+}* et sur des cerveaux de souris contrôles sans étiquetage du gène *Hdac2* (souris *Hdac2^{+/+}*) ont permis de suivre la liaison de la protéine HDAC2 au niveau de promoteurs de gènes dans le cerveau de souris (**Figure 21**). Le CHIP-tandem comporte deux étapes d'IP : une première IP avec l'anticorps anti-Flag, puis une élution par compétition avec le peptide FLAG et ensuite une deuxième IP avec l'anticorps anti-HA (de Dieuleveult et al. 2016).

Grâce à ces expériences, on constate que HDAC2 est cartographiée au niveau de promoteurs de gènes actifs dans le cerveau de souris (**Figure 21**). Le profil de liaison de HDAC2 au niveau des promoteurs s'étend parfois sur plusieurs kb dans la région codante des gènes sur lesquels HDAC2 est le plus liée (**Figure 21**).

Nous avons réalisé des expériences de CHIP-seq tandem identiques chez des animaux 3xTg-AD ; *Hdac2^{HF/+}* et 3xTg-AD ; *Hdac2^{HF/HF}*. Toutefois, ces expériences ne nous ont pas permis de détecter des différences de distribution significatives de HDAC2 sur le génome des souris 3xTg-AD et des souris contrôles. Nous avons conclu que la modeste surexpression de HDAC2 chez les souris 3xTg-AD ne s'accompagne pas d'une redistribution de HDAC2 sur le génome.

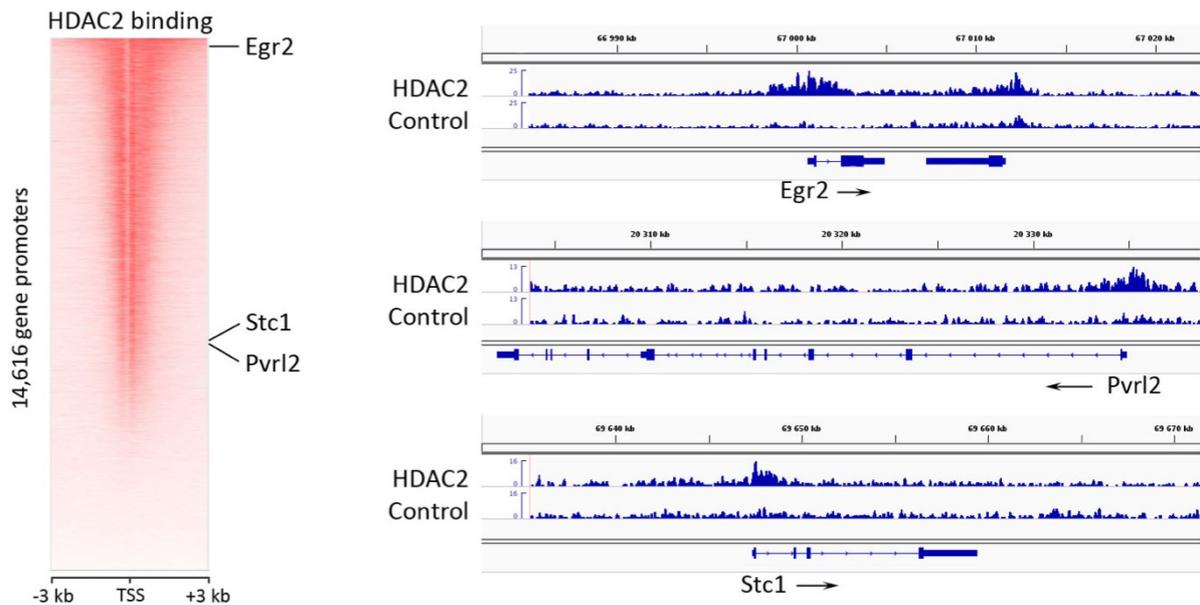


Figure 21 : Liaison de la protéine HDAC2 au niveau du promoteurs de ces gènes cibles dans le cerveau de souris

Des expériences de CHIP-seq séquentiel (tandem : IP anti-Flag, élution par compétition avec le peptide FLAG et IP anti-HA) ont été réalisées sur des cerveaux de souris *Hdac2*^{HF/+} âgées de 3 à 5 mois. La figure de gauche est une *heatmap* obtenue grâce au logiciel seqMINER (Ye et al. 2011). Chaque ligne de la *heatmap* représente un des 14 616 promoteurs de gènes et chaque colonne représente une petite fenêtre de l'intervalle génomique de 3 kb de part et d'autre du TSS. La couleur rouge est proportionnelle au nombre de *reads* présents au niveau d'un promoteur. Plus la couleur rouge est soutenue, plus la protéine HDAC2 est liée au niveau de ce promoteur. Les promoteurs ont été classés en ordre décroissant en fonction de leur niveau de liaison pour HDAC2. La figure de droite correspond à des représentations IGV de la liaison de la protéine HDAC2 au niveau du promoteur de trois gènes cibles dans des cerveaux de souris *Hdac2*^{HF/+} (HDAC2) et dans des cerveaux de souris *Hdac2*^{+/+} où la protéine HDAC2 n'est pas étiquetée (Control). L'axe Y indique le nombre de *reads* par million. Sur l'axe X, les flèches indiquent le sens de transcription des trois gènes sélectionnés pour exemple.

IV. Analyse par RNA-seq du contrôle transcriptionnel exercé par HDAC2 dans le contexte de la maladie d'Alzheimer

Dans le but de mesurer la contribution de la protéine HDAC2 à la régulation de la transcription dans l'hippocampe, nous avons réalisé des expériences de RNA-seq dans différents contextes génétiques.

Tout d'abord, nous avons étudié l'impact du modèle 3xTg-AD sur la transcription des gènes par rapport aux souris contrôles.

Puis, nous avons étudié les conséquences de la déplétion de la protéine HDAC2 sur le transcriptome dans le contexte sain et dans le contexte du modèle 3xTg-AD. Notre objectif était de tester si certaines dérégulations transcriptionnelles liées à la maladie d'Alzheimer pouvaient être provoquées par HDAC2.

Enfin, nous nous sommes intéressés aux répercussions de l'âge dans un second modèle murin de la maladie d'Alzheimer en comparant le transcriptome des souris modèles NLGF avec celui des souris contrôles.

1) Impact du modèle de souris 3xTg-AD de la maladie d'Alzheimer sur le transcriptome murin

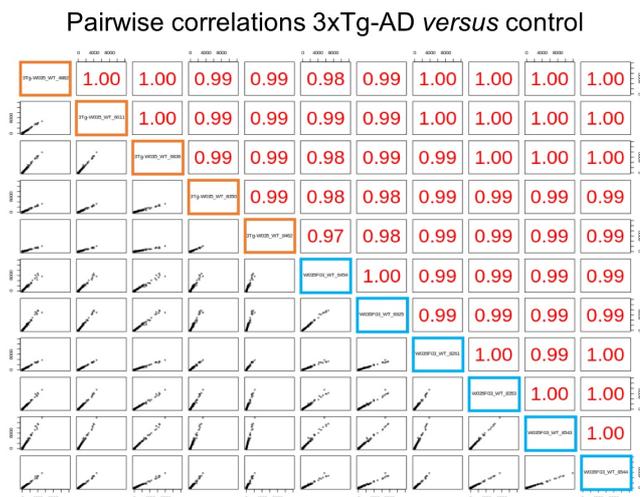
Pour étudier l'influence du modèle 3xTg-AD de la maladie d'Alzheimer sur l'expression des gènes, nous avons comparé 5 souris modèles 3xTg-AD avec 6 souris contrôles (*Hdac2*^{+/+}).

L'analyse des corrélations a permis de montrer que les différents échantillons de RNA-seq sont bien homogènes (**Figure 22a**). Sur l'analyse en composante principale (ACP ou PCA en anglais), l'axe 1 qui représente 92,17 % de la variance ne permet pas de séparer les différents échantillons selon leur génotype (3xTg-AD ou contrôle) (**Figure 22b**).

L'analyse de ces RNA-seq a révélé 989 gènes différentiellement exprimés (sur un total d'environ 30 000 gènes) dans le contexte du modèle 3xTg-AD de la maladie d'Alzheimer (**Tableau 7**). Le gène *Hdac2* n'a pas été identifié parmi les gènes différentiellement exprimés (*Fold-Change* = -0,25 et *p-value* = 0,25). Lorsque l'on choisit un seuil de dérégulation (*Fold-Change* \geq 1,5), il reste alors 685 gènes différentiellement exprimés (599 up-régulés et 86 down-régulés) (**Tableau 7**). Le fait d'observer 599 gènes up-régulés, contre seulement 86 gènes down-régulés, ne soutient pas l'hypothèse d'un blocage épigénétique lié à la

surexpression de HDAC2. En effet, selon cette hypothèse, seuls les 86 gènes down-régulés sont candidats à une répression médiée par HDAC2. Le modèle 3xTg-AD de la maladie d'Alzheimer entraîne donc une perturbation modérée de l'expression des gènes dans l'hippocampe du cerveau de souris.

a.



b.

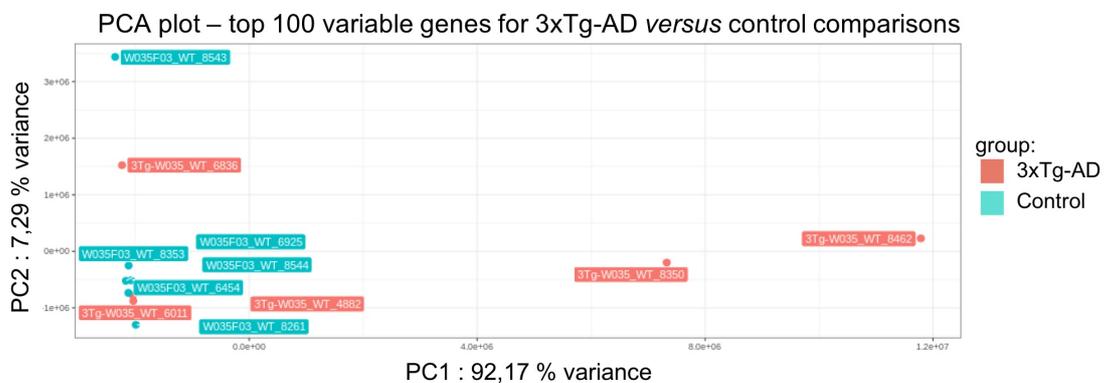


Figure 22 : Analyse en composante principale et calcul de corrélation entre les souris 3xTg-AD et contrôles.

- a. Calcul de corrélation entre les 5 échantillons RNA-seq issus de souris 3xTg-AD (cadres orange) et les 6 échantillons RNA-seq issus de souris contrôle (cadres bleus).
- b. Représentation ACP sur laquelle les différents individus 3xTg-AD (en rouge) et contrôles (en bleu) sont séparés selon l'axe 1 représentant 92,17 % de la variance et selon l'axe 2 représentant 7,29 % de la variance.

2) Influence de la perte de fonction de HDAC2 sur le transcriptome des souris dans le contexte sain et dans des modèles de la maladie d'Alzheimer

Nous avons également étudié les conséquences de la perte de fonction de HDAC2 dans le contexte du modèle 3xTgAD de la maladie d'Alzheimer. Pour cela, nous avons combiné les allèles des lignées 3xTgAD et *Hdac2* KO, comme décrit précédemment, pour comparer le transcriptome des hippocampes de souris 3xTgAD; *Hdac2*^{-/-} avec le transcriptome des hippocampes de souris 3xTgAD ; *Hdac2*^{+/+}. Notre objectif consistait à identifier les gènes qui seraient à la fois régulés par HDAC2 et dérégulés dans le contexte de la maladie d'Alzheimer : si la répression transcriptionnelle dans l'hippocampe des souris 3xTg-AD est médiée par HDAC2, alors nous devrions observer une levée de cette répression transcriptionnelle dans le contexte génétique 3xTgAD; *Hdac2*^{-/-}.

La première étape de nos analyses a été de vérifier de manière bioinformatique que les transcrits du gène *Hdac2* sont bien tronqués dans les souris *Hdac2*^{-/-} et 3xTg-AD ; *Hdac2*^{-/-} (**Figure 23**).

Nous avons également vérifié de manière bioinformatique que toutes les souris 3xTg-AD et 3xTg-AD ; *Hdac2*^{-/-} portent bien la mutation caractéristique du modèle 3xTg-AD : une mutation de la Méthionine en Valine en position 146 de la séquence protéique de la préséniline (**voir Matériel et Méthodes, partie B.III.3) et figure 17**).

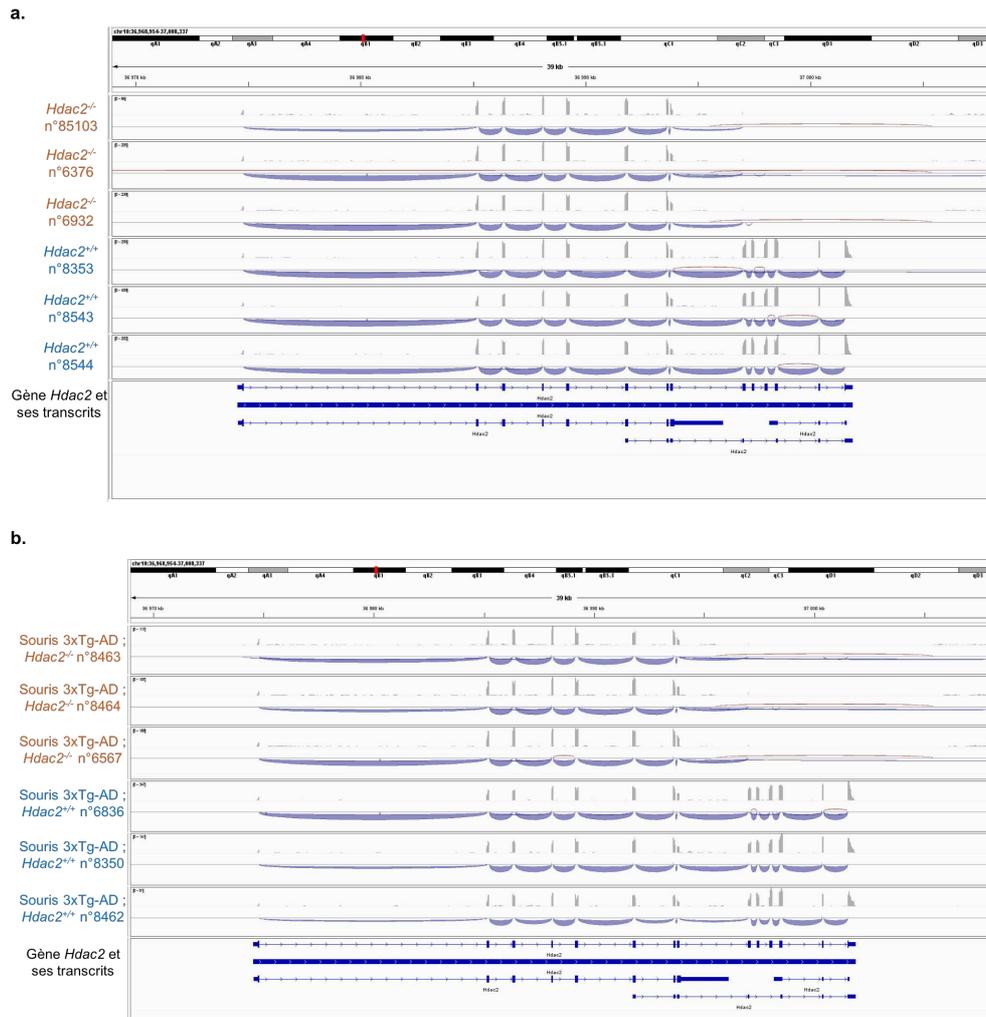
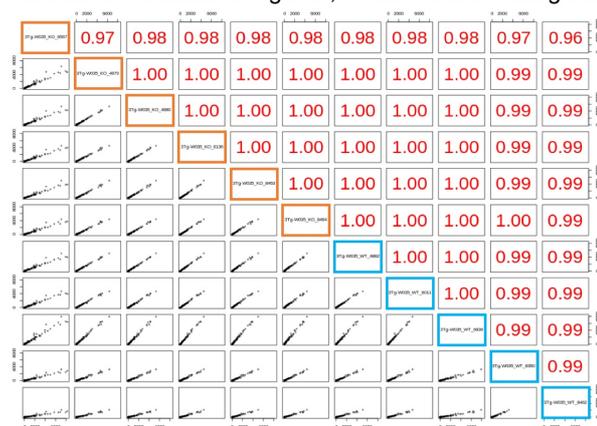


Figure 23 : Vérification de la délétion de la partie C-terminale des transcrits *Hdac2* dans les souris *Hdac2*^{-/-} et 3xTg-AD ; *Hdac2*^{-/-}

Le gène *Hdac2* (en trait plein bleu) possède 4 transcrits différents (lignes bleues). Les exons sont représentés par des pics gris et les jonctions entre exons par les vagues bleues grises. Plus la vague est épaisse et plus la jonction est présente. Les fines vagues rouges correspondent à de l'expression antisens. Pour 3 souris *Hdac2*^{-/-} (orange) versus 3 souris contrôles *Hdac2*^{+/+} (bleu) (a.) et 3 souris 3xTg-AD ; *Hdac2*^{-/-} (orange) versus 3 souris contrôles 3xTg-AD ; *Hdac2*^{+/+} (bleu) (b.) prises pour exemple, les six derniers exons des souris *Hdac2*^{-/-} et 3xTg-AD ; *Hdac2*^{-/-} ont bien été délétés.

a.

Pairwise correlations 3xTg-AD ; *Hdac2*^{-/-} versus 3xTg-AD



b.

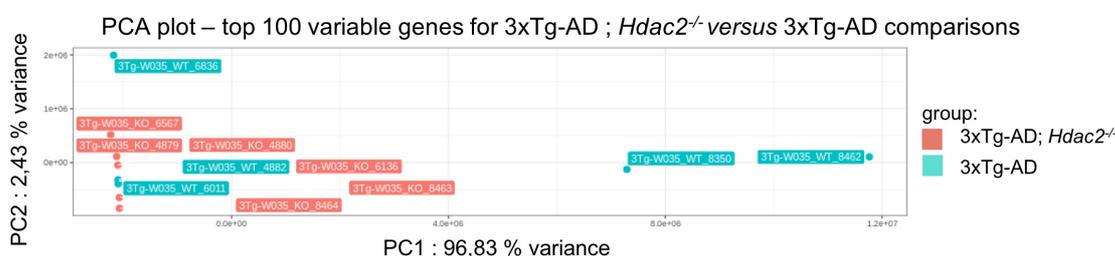


Figure 24 : Analyse en composante principale et calcul de corrélation entre les souris 3xTg-AD ; *Hdac2*^{-/-} et 3xTg-AD.

- a. Calcul de corrélation entre les 6 échantillons RNA-seq issus de souris 3xTg-AD ; *Hdac2*^{-/-} (cadres orange) et les 5 échantillons RNA-seq issus de souris 3xTg-AD (cadres bleus).
- b. Représentation ACP sur laquelle les différents individus 3xTg-AD ; *Hdac2*^{-/-} (en rouge) et 3xTg-AD (en bleu) sont séparés selon l'axe 1 représentant 96,83 % de la variance et selon l'axe 2 représentant 2,43 % de la variance.

Dans une deuxième étape, les échantillons RNA-seq issus de 6 souris 3xTg-AD ; *Hdac2*^{-/-} et 5 souris 3xTg-AD âgées de 6 à 9 mois ont été analysés. A ces âges, les défauts associés à la maladie d'Alzheimer sont bien marqués.

Les différents échantillons de RNA-seq sont bien homogènes d'après leurs coefficients de corrélation (Figure 24a). Sur l'ACP, l'axe 1 qui représente 96,83 % de la variance ne permet pas de séparer les différents échantillons selon leur génotype 3xTg-AD ; *Hdac2*^{-/-} ou 3xTg-AD (Figure 24b).

Nos analyses ont révélé 320 gènes différentiellement exprimés entre souris 3xTg-AD ; *Hdac2*^{-/-} et souris 3xTg-AD (Tableau 7). Lorsque l'on choisit un seuil de dérégulation (*Fold-Change* >= 1,5), il reste alors 240 gènes différentiellement exprimés (122 up-régulés et 118 down-régulés)

(**Tableau 7**). Ces 122 gènes up-régulés dans le contexte du modèle 3xTg-AD (qui sont homozygotes pour l'allèle mutant *Hdac2*) pourraient correspondre à des gènes cibles de HDAC2, réprimés au niveau transcriptionnel lorsque HDAC2 est présente.

Dans le cadre de cette étude, nous voulions également mesurer l'impact de la protéine HDAC2 sur la transcription globale des gènes dans l'hippocampe dans le contexte sain. L'objectif était de pouvoir comparer les gènes cibles de HDAC2 dans l'hippocampe sain et dans l'hippocampe du modèle 3xTg-AD. Cette comparaison est nécessaire pour évaluer si les gènes cibles de HDAC2 restent constants dans le contexte pathologique, par rapport au contexte sain, ou si au contraire HDAC2 était, du fait de sa surexpression, capable de réprimer de nouveaux gènes cibles dans le contexte pathologique de la maladie d'Alzheimer.

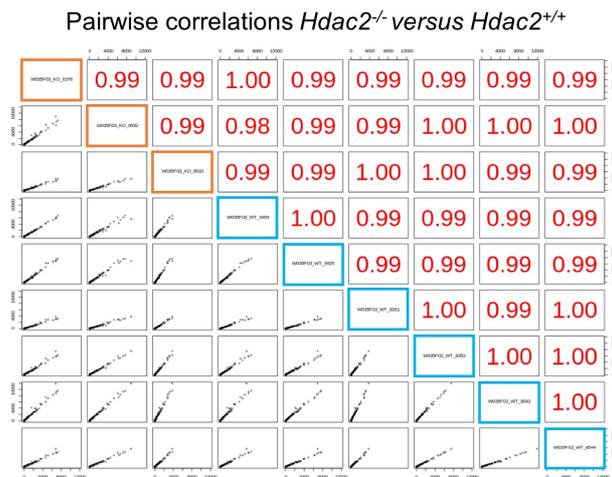
Dans ce but, nous avons comparé les données transcriptomiques de 3 souris *Hdac2*^{-/-} âgées de 3 à 6 mois et de 6 souris contrôles *Hdac2*^{+/+} du même âge. L'analyse des corrélations a montré l'homogénéité des différents échantillons de RNA-seq (**Figure 25a**).

Sur l'ACP, l'axe 1 représentant 93,74 % de la variance ne permet pas de séparer les différents échantillons selon leur génotype *Hdac2*^{-/-} ou *Hdac2*^{+/+} (**Figure 25b**).

Les analyses des transcriptomes entre souris *Hdac2*^{-/-} et souris contrôles ont mis en évidence 137 gènes différentiellement exprimés dans le contexte sain en absence de HDAC2 (**Tableau 7**). Comme attendu, le gène *Hdac2* est bien différentiellement exprimé (*Fold-Change* = -1,6 et *p-value* = 10⁻⁵⁰). Lorsqu'un seuil de dérégulation a été choisi (*Fold-Change* ≥ 1,5), il n'y a plus que 74 gènes retrouvés dérégulés (42 up-régulés et 32 down-régulés) (**Tableau 7**).

Sur les 137 gènes différentiellement exprimés dans les souris *Hdac2*^{-/-} par rapport aux souris contrôle, une analyse de *Gene Ontology* a été effectuée pour détecter d'éventuelles voies biologiques qui seraient significativement enrichies en gènes dérégulés par l'absence de HDAC2 (**Figure 26**). Les voies significativement enrichies en gènes dérégulés sont impliquées dans la gliogénèse et dans la mise en place de la gaine de myéline entourant les axones (**Figure 26**).

a.



b.

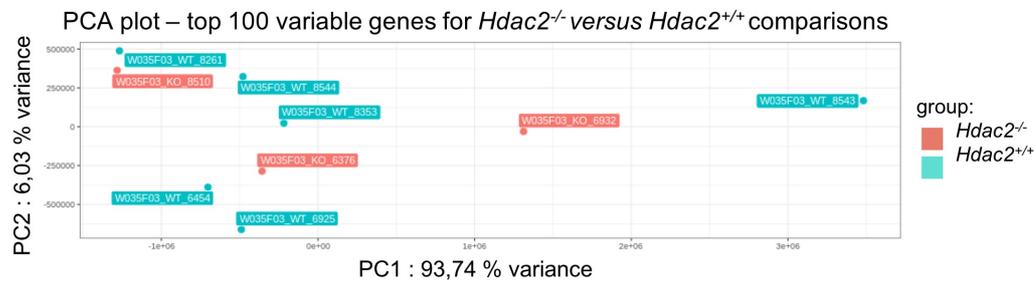


Figure 25 : Analyse en composante principale et calcul de corrélation entre les souris *Hdac2*^{-/-} et *Hdac2*^{+/+}.

a. Calcul de corrélation entre les 3 échantillons RNA-seq issus de souris *Hdac2*^{-/-} (cadres orange) et les 6 échantillons RNA-seq issus de souris contrôle *Hdac2*^{+/+} (cadres bleus).

b. Représentation ACP sur laquelle les différents individus *Hdac2*^{-/-} (en rouge) et contrôles (en bleu) sont séparés selon l'axe 1 représentant 93,74 % de la variance et selon l'axe 2 représentant 6,03 % de la variance.

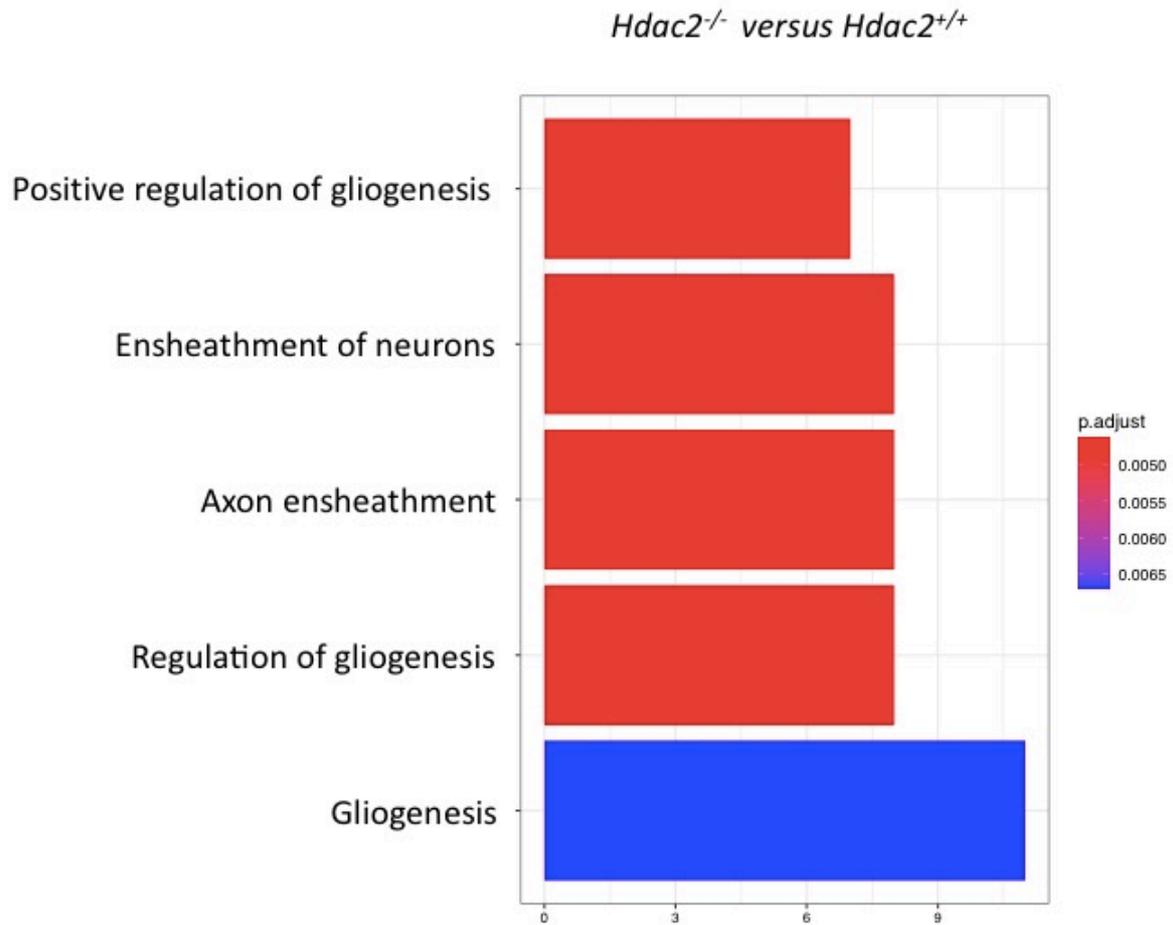


Figure 26 : Analyse de *Gene Ontology* révélant les différentes voies biologiques dans lesquelles les gènes dérégulés sont exprimés en absence de HDAC2.

Représentation en barplot des différentes voies biologiques (axes Y) dans lesquelles les gènes sont dérégulés dans les comparaisons impliquant des souris *Hdac2^{-/-}* versus des souris contrôles *Hdac2^{+/+}*. La taille des barres le long de l'axe X indique le nombre de gènes présents dans chaque voie. La couleur représente les p-valeurs ajustées correspondant aux scores d'enrichissement.

Comparaison	Nombre de gènes différentiellement exprimés		Nombre de gènes UP régulés		Nombre de gènes DOWN régulés	
sauvage (<i>Hdac2^{+/+}</i>) versus <i>Hdac2^{-/-}</i>	137	74	79	42	58	32
3xTg-AD (<i>Hdac2^{+/+}</i>) versus 3xTg-AD ; <i>Hdac2^{-/-}</i>	320	240	173	122	147	118
sauvage versus 3xTg-AD	989	685	682	599	307	86
sauvage 3 mois versus NLGF 3 mois	274	14	161	6	113	8
sauvage 9 mois versus NLGF 9 mois	1 117	421	771	388	346	33
NLGF 3 mois versus NLGF 9 mois	1 067	461	777	446	290	15
sauvage 3 mois versus sauvages 9 mois	67	24	35	15	32	9

Tableau 7 : Nombre de gènes différentiellement exprimés issus des analyses de RNA-seq pour des comparaisons d'intérêt chez les souris modèles 3xTg-AD et NLGF.

Le nombre total de gènes différentiellement exprimés (colonne 2), ainsi que le nombre de gènes UP-régulés (colonne 3) et DOWN-régulés (colonne 4) obtenu par des analyses RNA-seq sans filtre sur le *Fold-Change* (nombres en noir) ou avec un filtre sur le *Fold-Change* ≥ 1.5 (nombres en rouge) sont présentés dans ce tableau suite à différentes comparaisons.

Le laboratoire du groupe du Dr Li-Huei Tsai avait montré que des gènes impliqués dans la régulation de l'activité neuronale (*Bdnf*, *Egr1*, *Homer1*, *Fos* et *Cpg15*), dans la formation des synapses (*Nrxn1*, *Nrxn3*, *Syn2*, *Psd95* et *Agrin*) et dans la plasticité synaptique (*GluR1*, *NR2b*, *NR2a*, *Camk2a*, *Creb1*, *Crebbp*, *Cdk5r1*, *Cdk5*) sont dérégulés dans les hippocampes de souris déficientes en HDAC2 (souris *Hdac2^{-/-}*) (Guan et al. 2009). Sur ces 18 gènes dérégulés, seul le

gène *Syn2* est retrouvé down-régulé dans les analyses comparant les transcriptomes 3xTg-AD ; *Hdac2*^{-/-} et 3xTg-AD.

Dans une autre étude, le même laboratoire a montré que des gènes impliqués dans les processus d'apprentissage et de mémorisation (*Arc*, *Bdnf*, *Cdk5*, *Egr1*, *Homer1*) et de la plasticité synaptique (*GluR1*, *GluR2*, *Nfl*, *NR2a*, *NR2b*, *Syp* et *Syt1*) sont dérégulés dans le modèle de souris CK-p25. Sur ces 12 gènes, aucun n'est dérégulé dans l'hippocampe des souris chez le modèle 3xTg-AD par rapport aux souris contrôles.

Notre objectif initial était de caractériser à l'échelle du génome le blocage épigénétique exercé par HDAC2 sur la transcription des gènes dans le contexte de deux modèles de la maladie d'Alzheimer. Nos résultats obtenus sur le modèle 3xTg-AD ne semblent pas en phase avec ceux obtenus par le Laboratoire du Dr LH Tsai avec le modèle CK-p25 : les gènes dont la transcription est réprimée par HDAC2 ne sont pas communs entre les différents modèles de souris.

J'ai effectué différentes comparaisons avec les données RNA-seq obtenues avec nos différentes lignées de souris, pour réévaluer l'étendue du contrôle transcriptionnel exercé par HDAC2 dans l'hippocampe (**Figure 27**).

Si HDAC2 exerce sa fonction répressive sur la transcription des gènes dans le contexte sain, ces mêmes gènes devraient également être trouvés réprimés dans le contexte du modèle 3xTg-AD. Ainsi les gènes up-régulés par l'absence de HDAC2 devraient être retrouvés parmi les gènes up-régulés dans le contexte des souris modèles 3xTg-AD en absence de HDAC2.

Parmi les 42 gènes up-régulés dans les comparaisons *Hdac2*^{+/+} versus *Hdac2*^{-/-} et les 122 gènes up-régulés dans les comparaisons 3xTg-AD versus 3xTg-AD ; *Hdac2*^{-/-}, 19 gènes sont communs (**Figure 27**). Nos données montrent donc que seulement 19 gènes sont réprimés (de façon reproductible) par HDAC2 lorsque la version sauvage de cette enzyme est exprimée dans l'hippocampe. Le fait qu'on retrouve plus de gènes up-régulés par la perte de fonction de HDAC2 dans le contexte 3xTg-AD que dans le contexte sauvage est toutefois compatible avec la possibilité que HDAC2 réprime des gènes additionnels dans le contexte du modèle 3xTg-AD. Si cette hypothèse est vraie, cela implique que HDAC2 régule des gènes parmi les 86 gènes down-régulés dans le modèle 3xTg-AD. De nombreux gènes communs devraient donc être

trouvés entre les 86 gènes down-régulés dans le contexte du modèle 3xTg-AD et les 122 gènes up-régulés dans le contexte de la perte de fonction de HDAC2 dans le modèle 3xTg-AD. Cette comparaison a mis en évidence seulement 11 gènes communs (Figure 27). Notre hypothèse de blocage épigénétique médié par HDAC2 n'est donc potentiellement valide que pour ces 11 gènes.

On peut donc conclure de ces études que la protéine HDAC2 a un impact très modeste sur le transcriptome des souris modèles 3xTg-AD. De manière globale, ces résultats remettent en cause la solidité de l'hypothèse du bocage épigénétique causé par HDAC2 dans le contexte de la maladie d'Alzheimer. En effet, si les gènes cibles de HDAC2 varient selon le modèle choisi pour étudier la maladie, alors il devient difficile de se projeter sur des stratégies thérapeutiques sur la base de signatures transcriptionnelles spécifiques à chaque modèle.

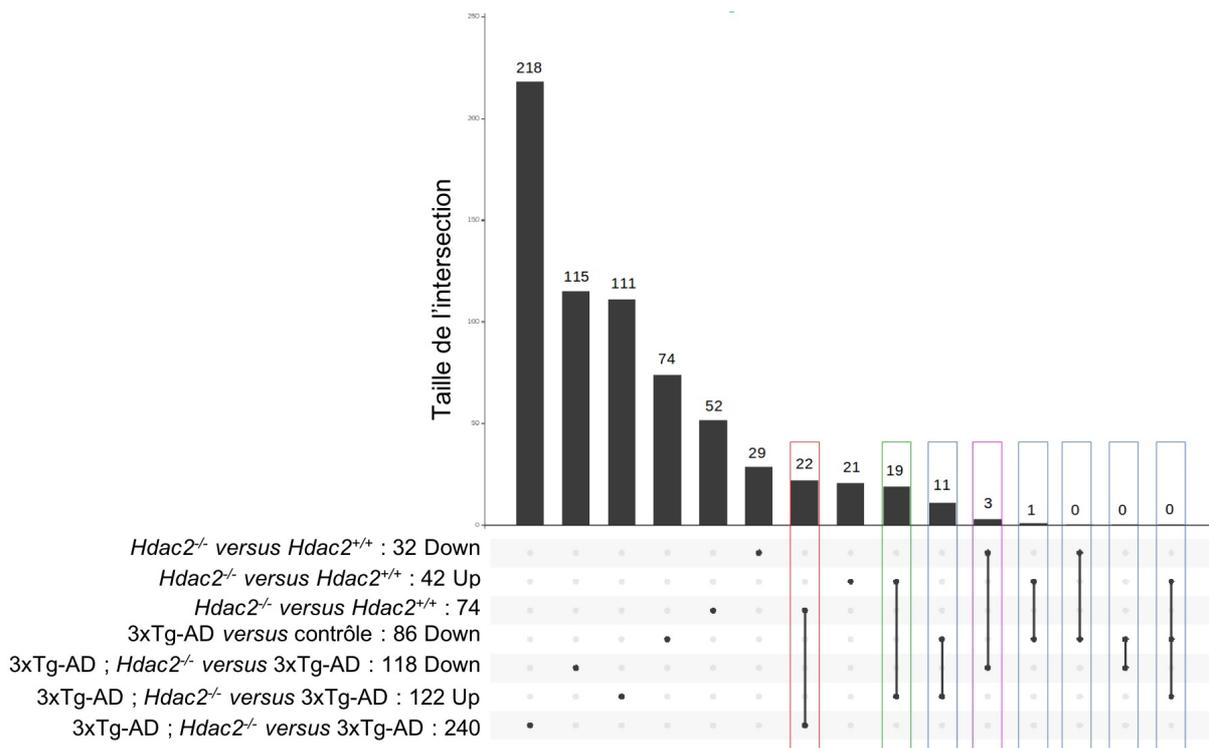


Figure 27 : Nombre de gènes communs entre différentes listes de gènes dérégulés

Upsetplot représentant le nombre de gènes en commun présents dans différentes liste de gènes dérégulés. Parmi les listes de gènes dérégulés en abscisse, on a les gènes down-régulés, les gènes up-régulés et le total des gènes dérégulés dans la comparaison *Hdac2*^{-/-} versus *Hdac2*^{+/+}, les gènes down-régulés dans la comparaison 3xTg-AD versus contrôle, les gènes up-régulés, down-régulés et dérégulés dans la comparaison 3xTg-AD ; *Hdac2*^{-/-} versus 3xTg-AD . En ordonnée, les points reliés par la ligne indique dans quelles listes les gènes communs sont cherchés et les barres noires indiquent le nombre de gènes trouvé en commun.

3) Influence du modèle NLGF de la maladie d'Alzheimer sur le transcriptome des souris

Comme expliqué précédemment, l'absence de surexpression de HDAC2 chez les souris NLGF nous a suggéré que HDAC2 n'aurait aucun rôle dans la progression des symptômes de la maladie. Nous nous sommes fixé comme objectif de repli d'étudier l'influence de la maladie d'Alzheimer sur l'expression des gènes dans un autre modèle murin. Nous avons comparé les transcriptomes de 16 souris modèles NLGF avec ceux de 16 souris contrôles. Les souris sont âgées de 3 ou de 9 mois. L'ACP de cette comparaison a mis en évidence une très bonne séparation des souris NLGF âgées de 9 mois par rapport aux souris NLGF âgées de 3 mois et aux souris contrôles de tout âge (Figure 28). Afin de prendre en compte cet effet âge visible chez les souris NLGF, des nouvelles analyses de RNA-seq ont été réalisées afin d'étudier l'influence de l'âge et de ce modèle de la maladie d'Alzheimer sur le transcriptome de souris.

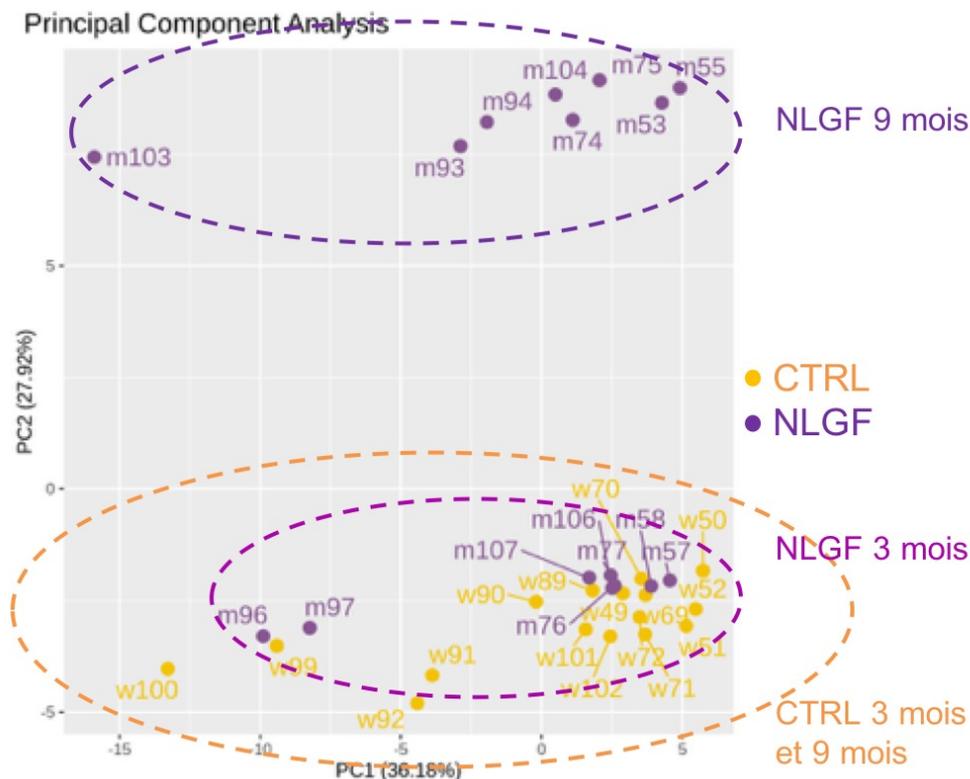


Figure 28 : Analyse en composante principale des échantillons NLGF et contrôles.

Représentation ACP sur laquelle les différents individus NLGF (8 souris âgées de 3 mois et 8 souris âgées de 9 mois ; en violet) et les individus contrôles (CTRL) (8 souris âgées de 3 mois et 8 souris âgées de 9 mois ; en jaune-orangé) sont séparés selon l'axe 1 représentant 36,18 % de la variance et selon l'axe 2 représentant 27,92 % de la variance. Des ellipses en pointillées ont été tracées pour montrer les groupes d'individus.

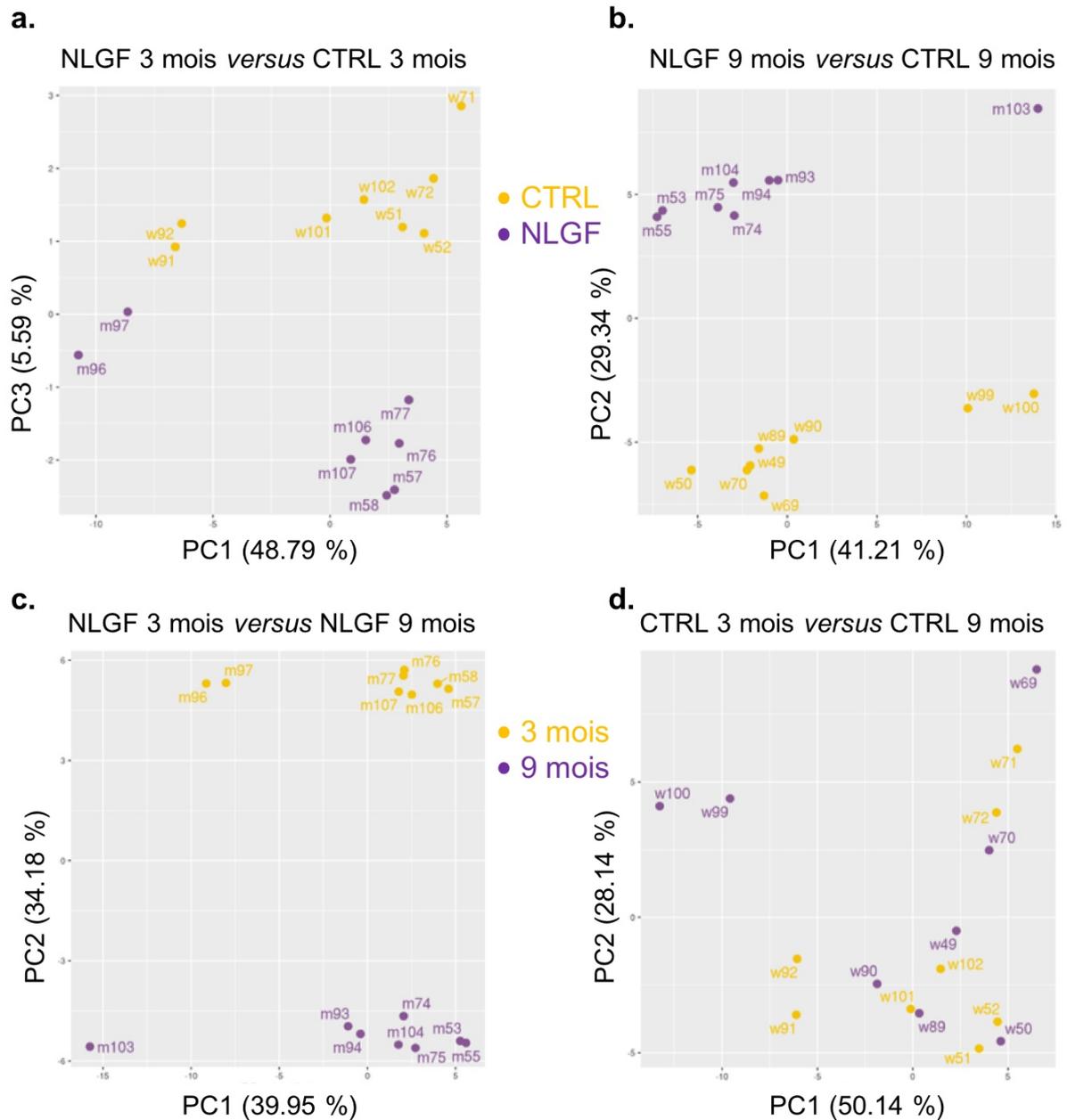


Figure 29 : Analyse en composante principale des échantillons NLGF et contrôles selon différentes comparaisons.

a. et b. Représentation ACP sur laquelle 8 souris NLGF (en violet) et 8 souris contrôle (CTRL ; en jaune) toutes âgées de 3 mois (**a.**) ou de 9 mois (**b.**) sont séparées selon l'axe 1 et l'axe 2 ou 3 qui représentent différents pourcentages de variance.

c. et d. Représentation ACP sur laquelle 8 souris âgées de 9 mois (en violet) et 8 souris âgées de 3 mois (en jaune) toutes de génotype NLGF (**c.**) ou CTRL (**d.**) sont séparées selon l'axe 1 et l'axe 2 qui représentent différents pourcentages de variance.

Les transcriptomes des souris NLGF et des souris contrôles ont été étudiés à 3 mois, stade pré-symptomatique de ce modèle de la maladie d'Alzheimer. Notre premier objectif était de savoir si une dérégulation de l'expression des gènes était déjà présente à 3 mois. Sur la représentation ACP, les échantillons des souris NLGF de 3 mois sont séparés des échantillons de souris contrôles du même âge (**figure 29a**). Les analyses transcriptomiques ont révélé 274 gènes différentiellement exprimés entre les souris NLGF et contrôles âgées de 3 mois (**Tableau 7**). Lorsque l'on choisit un seuil de dérégulation (*Fold-Change* $\geq 1,5$), il ne reste que 14 gènes différentiellement exprimés (6 up-régulés et 8 down-régulés) (**Tableau 7**). Le faible nombre de gènes différentiellement exprimés entre les souris NLGF et contrôle à 3 mois est cohérent avec le fait que les souris sont encore asymptomatiques à cet âge.

Les transcriptomes des souris NLGF et des souris contrôles ont également été étudiés à 9 mois, stade auquel les symptômes de la maladie d'Alzheimer sont bien présents.

Les résultats sur l'ACP montrent une bonne séparation des souris NLGF âgées de 9 mois et des souris contrôles du même âge selon l'axe 2 représentant 29,34 % de la variance (**figure 29b**). Les analyses RNA-seq ont révélé 1 117 gènes différentiellement exprimés entre les souris NLGF et contrôles âgées de 9 mois (**Tableau 7**). Après avoir choisi un seuil de dérégulation (*Fold-Change* $\geq 1,5$), il reste 421 gènes différentiellement exprimés (388 up-régulés et 33 down-régulés) (**Tableau 7**). L'expression des gènes est donc perturbée chez les souris de 9 mois dans le contexte du modèle NLGF de la maladie d'Alzheimer.

Pour caractériser l'effet de l'âge des souris sur la régulation de l'expression des gènes, nous avons comparé les souris NLGF âgées de 3 mois avec les NLGF âgées de 9 mois et les souris contrôles âgées de 3 mois avec les contrôles âgées de 9 mois. Alors que les échantillons de RNA-seq issus des souris contrôles âgées de 3 mois et 9 mois sont inséparables sur l'ACP (**figure 29d**), les échantillons de RNA-seq issus des souris NLGF âgées de 3 mois sont bien séparés des échantillons de souris NLGF âgées de 9 mois selon l'axe 2 représentant 34,18 % de la variance (**figure 29c**).

Concernant le nombre de gènes différentiellement exprimés, seuls 67 gènes sont différentiellement exprimés entre les souris contrôles de 3 mois et de 9 mois (24 gènes différentiellement exprimés après choix d'un seuil de *Fold-Change* $\geq 1,5$), tandis que 1 067

gènes sont différentiellement exprimés entre les souris NLGF de 3 mois et de 9 mois (461 gènes différentiellement exprimés après choix d'un seuil de *Fold-Change* $\geq 1,5$) ([Tableau 7](#)). Il n'est pas surprenant de trouver un faible nombre de gènes différentiellement exprimés chez les souris contrôles entre 3 mois et 9 mois car 9 mois n'est pas un âge très avancé pour les souris de laboratoire (espérance de vie jusqu'à 2 ans).

Nous pouvons donc conclure de ces analyses que l'âge exerce une influence sur la régulation de la transcription des gènes dans le contexte du modèle NLGF de la maladie d'Alzheimer.

Nous avons mis en évidence que l'âge et la maladie d'Alzheimer impactent la régulation de la transcription des gènes chez la souris.

Parmi les 12 gènes dérégulés dans le modèle CK-p25, seul le gène *Homer1* qui est impliqué dans l'activité neuronale a été trouvé up-régulé chez les souris NLGF âgées de 9 mois. Le gène *Creb1* (qui était dérégulé en absence de HDAC2 dans l'étude avec le modèle CK-p25) impliqué dans la plasticité synaptique est trouvé up-régulé dans les souris âgées de 9 mois dans le contexte du modèle NLGF. Aucun des autres gènes impliqués dans l'activité neuronale, la formation et la plasticité synaptique et dans les processus d'apprentissage et de mémorisation n'ont été retrouvés dérégulés dans le modèle NLGF.

Pour aller plus loin, nous avons identifié des voies biologiques significativement enrichies en gènes dérégulés au stade symptomatique dans les souris NLGF. Des analyses de *Gene Ontology* ont été effectuées sur les RNA-seq comparant les échantillons NLGF et contrôles à 9 mois et sur les RNA-seq comparant les échantillons NLGF à 3 mois et 9 mois ([figure 30](#)). Dans les deux comparaisons, les voies biologiques impliquées dans les processus de l'inflammation et de l'immunité sont significativement enrichies en gènes dérégulés ([figure 30](#)). Le modèle NLGF au stade symptomatique induit donc une dérégulation de l'expression des gènes qui semblent notamment impliqués dans l'action des cellules immunitaires.

Nous remarquons également que la majorité des gènes dérégulés dans le modèle NLGF sont up-régulés, ce qui suggère un effet inducteur du modèle NLGF sur l'expression des gènes.

Plus généralement, il ne semble pas y avoir un blocage épigénétique d'envergure dans le modèle NLGF car seulement 15 gènes sont down-régulés chez les souris symptomatiques ([Tableau 7](#)).

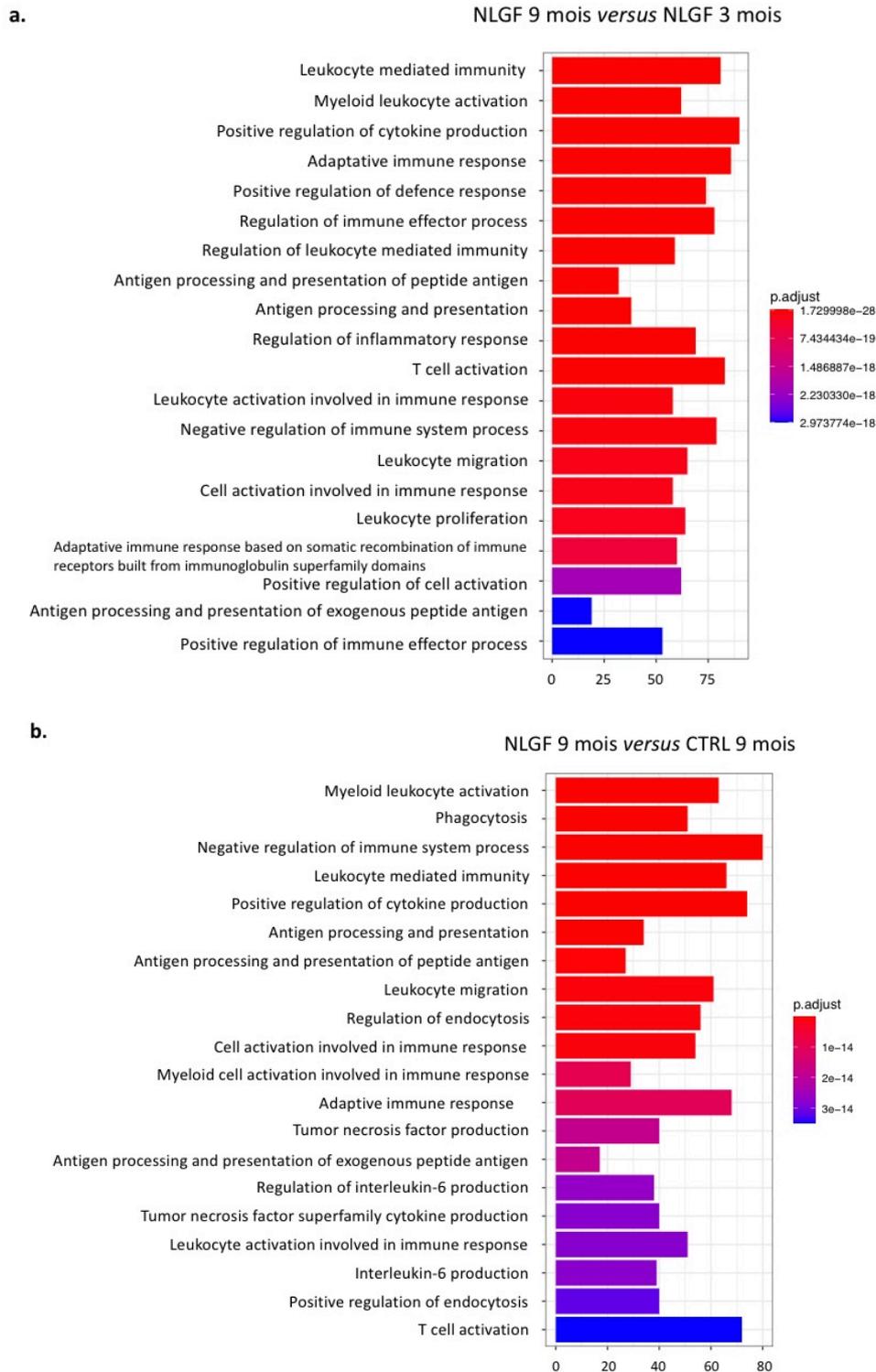


Figure 30 : Analyse de *Gene Ontology* révélant les différentes voies biologiques dans lesquelles les gènes dérégulés sont exprimés.

Représentation en barplot des différentes voies biologiques (axes Y) dans lesquelles les gènes sont dérégulés dans les comparaisons impliquant des souris NLGF 9 mois *versus* des souris NLGF 3 mois (**a.**) ou dans les comparaisons impliquant des souris NLGF 9 mois *versus* des souris CTRL 9 mois (**b.**). La taille des barres le long de l'axe X indique le nombre de gènes présents dans chaque voie. La couleur représente les p-valeurs ajustées correspondant aux scores d'enrichissement.

Discussion

Partie A : Caractérisation de l'architecture des particules nucléosomales au niveau des enhancers et impact du remodeleur BRG1 sur la formation de particules subnucléosomales

Jusqu'à présent, les enhancers étaient réputés pour être des éléments régulateurs peu structurés au niveau de l'organisation nucléosomale. Grâce à nos études basées sur des expériences de MNase-ChIP-seq, nous avons pu analyser de manière détaillée leur architecture en caractérisant séparément des sous-groupes d'enhancers présentant des variantes d'une même organisation. Au centre de l'enhancer, nous avons mis en évidence l'existence de nucléosomes fragiles et de particules subnucléosomales ayant la même position génomique que le nucléosome fragile. Cette région centrale est flanquée par deux nucléosomes canoniques bien positionnés. Les différents groupes d'enhancer que nous avons identifiés se distinguent par le nombre de nucléosomes fragiles présents entre les deux nucléosomes canoniques. La distance entre ces deux nucléosomes canoniques (appelés arbitrairement -1 et +1) est en effet variable selon les groupes d'enhancers, et plus l'intervalle entre les nucléosomes -1 et +1 est grand, plus le nombre de nucléosomes fragiles présents dans cet intervalle est important. Chaque nucléosome fragile est associé à deux particules subnucléosomales localisées de part et d'autre de l'axe de la dyade du nucléosome fragile, ce qui définit un module qui peut être présent en une seule copie, dans le cas le plus simple (groupe 1), en deux copies (groupe 2), ou en de multiples copies (3 à 6 copies) dans le groupe 3.

La présence de nucléosomes fragiles dans la région centrale des enhancers avait été décrite précédemment par le laboratoire du Dr Zaret (Iwafuchi-Doi et al. 2016). Nous avons caractérisé la nature des particules subnucléosomales avec des approches de MNase-ChIP-seq réalisées après séparation des espèces nucléosomales et subnucléosomales sur gradient de sucrose. Nous avons montré que la partie centrale des enhancers contient des particules subnucléosomales associées à des fragments d'ADN de différentes tailles. Nous nous sommes particulièrement intéressés aux particules subnucléosomales protégeant de la digestion par la MNase des fragments de 50 à 90 pb. Ces particules sont retrouvées dans les fractions du

gradient (5-6) bien séparées des nucléosomes canoniques (11-13), ce qui suggèrent qu'elles ont une masse moléculaire nettement inférieure à celle de l'octasome. Nous avons démontré que les 4 histones H3, H4, H2A et H2B sont présentes dans ces particules. La distribution symétrique de ces particules de part et d'autre de l'axe de la dyade du nucléosome fragile, ainsi que leur composition en histones et leur faible vitesse de sédimentation suggèrent que ces particules sont des hémisomes. Nous avons également montré que BRG1 est requis pour générer ces particules subnucléosomales. L'examen des espèces nucléosomales et subnucléosomales liées par BRG1 *in vivo* montre que ce remodelleur se lie préférentiellement d'une part sur les nucléosomes fragiles présents au centre de l'enhancer, et d'autre part sur les particules subnucléosomales de 50-90 pb. Ce pattern de liaison est compatible avec un scénario dans lequel BRG1 cible les nucléosomes du centre de l'enhancer pour les convertir en particules de type hémisomes. L'activité de BRG1 pourrait également être responsable de la sensibilité à la MNase du nucléosome fragile central.

Le facteur OCT4 est connu pour se fixer au niveau des enhancers sur l'ADN nu et au niveau des nucléosomes canoniques lorsqu'il exerce sa fonction de facteur pionnier. Nos expériences de MNase-ChIP-seq ont révélé un enrichissement du facteur OCT4 plus important au niveau des particules subnucléosomales de 50-90 pb qu'au niveau des nucléosomes canoniques et de l'ADN nu. Un plus faible enrichissement d'OCT4 au niveau de l'ADN nu qu'au niveau des particules subnucléosomales est cohérent avec une observation récente réalisée par *single-molecule footprinting* en cellules ES dans lesquelles les interactions d'OCT4 avec son motif cible sont des événements peu fréquents (Sönmezer et al. 2021).

Nous avons également montré que les motifs consensus d'OCT4 sont fortement enrichis au niveau du nucléosome fragile. Au niveau des particules subnucléosomales, OCT4 ne se lie que partiellement sur ses sites de liaison consensus. Cela pourrait suggérer que OCT4 se lie sur des motifs partiels au niveau des particules subnucléosomales comme c'est le cas au niveau des nucléosomes canoniques (Soufi et al. 2015; Michael et al. 2020). Cependant, les motifs partiels d'OCT4 ne sont pas spécifiquement enrichis au niveau des subnucléosomes. Des interactions entre les histones H3, et en moindre mesure H2A et H4, avec le facteur OCT4, démontrées *in vitro* (Echigoya et al. 2020), pourraient aider OCT4 à se lier sur les particules

subnucléosomales. Un autre TF ou une modification d'histone pourrait également aider OCT4 à se fixer aux particules subnucléosomales.

La capacité de OCT4 à interagir avec les nucléosomes n'est pas remise en cause par nos observations car au cours du processus d'auto-renouvellement présent dans les cellules ES, les éléments enhancers sont maintenus en activité. Dans le contexte de la reprogrammation cellulaire (Soufi, Donahue, et Zaret 2012), les enhancers sont éteints. La fonction pionnière d'OCT4, c'est-à-dire sa capacité à se lier aux nucléosomes canoniques pour permettre la réactivation de l'enhancer, est alors cruciale. La faible fréquence de reprogrammation cellulaire atteste que la propriété de facteur pionnier est inefficace lorsque OCT4 est exprimée artificiellement dans un type cellulaire non pluripotent. Dans les cellules ES, au contraire, les enhancers sont constamment remodelés par le complexe mSWI/SNF (Iurlaro et al. 2021; Schick et al. 2021) et on peut donc spéculer que les particules subnucléosomales servent d'ancres aux TFs pour aider à maintenir les enhancers actifs.

Aux enhancers, le remodeleur BRG1 est aussi nécessaire au maintien du positionnement correct des nucléosomes canoniques et fragiles et il permet de rendre l'ADN accessible pour la fixation de TFs. Nous proposons un modèle dans lequel les particules subnucléosomales sont un intermédiaire de remodelage de la chromatine qui permet de rendre l'ADN de l'enhancer plus accessible aux TFs et aux facteurs pionniers.

Lors de ma thèse, j'ai étudié en profondeur le rôle du remodeleur BRG1 au niveau des enhancers. Cependant, les différents remodeleurs de la chromatine ont différentes activités au niveau des éléments régulateurs. Le laboratoire du Dr Schübeler a comparé *in vivo* le rôle des remodeleurs BRG1 et SMARCA5, qui sont respectivement les sous-unités ATPase des complexes SWI/SNF et ISWI. Ces deux remodeleurs ont été délétés par *knockout* (Barisic et al. 2019). Afin d'étudier le positionnement des nucléosomes au niveau des éléments régulateurs, Barisic *et al.* ont réalisé plusieurs expériences : une digestion de la chromatine à la MNase suivie d'un séquençage des fragments d'ADN correspondant aux mononucléosomes, un séquençage au bisulfite révélant les patterns de méthylation sur le génome et donc la position des nucléosomes protégés de cette méthylation et enfin de l'ATAC-seq pour déterminer les

régions de chromatine accessibles et en déduire les régions de chromatine ouvertes ainsi que celles occupées par les nucléosomes (Barisic et al. 2019). La perte de fonction de SMARCA5 conduit à une réduction du *phasing*, c'est-à-dire du positionnement régulier, des nucléosomes majoritairement au niveau des régions distales hypersensibles à la DNase I, qui correspondent généralement à des sites CTCF (Barisic et al. 2019). Cette altération du *phasing* s'accompagne d'une augmentation de la taille de l'ADN *linker* entre nucléosomes (Barisic et al. 2019). Aucune différence dans la taille de l'ADN *linker* n'a été détectée en absence de BRG1 (Barisic et al. 2019). La régulation de l'espacement entre les nucléosomes au niveau des sites CTCF dépend donc du remodeleur SMARCA5 chez les mammifères, mais pas du remodeleur BRG1.

Les remodeleurs de la chromatine ont également différents impacts sur la liaison des TFs au niveau de leur site cible. Grâce à des expériences de CHIP-seq et d'ATAC-seq, Barisic *et al.* ont mis en évidence différents ensembles de TFs qui semblent dépendants de l'activité de SMARCA5 ou de BRG1 pour se lier à leur motif cible. En effet, la liaison du facteur CTCF au niveau de ses motifs cibles dépend de SMARCA5 mais pas de BRG1, tandis que la liaison du facteur REST à ses motifs dépend de BRG1 et non de SMARCA5 (Barisic et al. 2019). Enfin, probablement grâce à son implication dans la liaison de CTCF à ses motifs cibles, SMARCA5 joue un rôle essentiel dans la formation des boucles de chromatine et des domaines topologiques associés de l'ADN (Barisic et al. 2019).

Au cours de ma thèse, j'ai également analysé l'organisation nucléosomale et la fonction des remodeleurs de la chromatine au niveau d'autres éléments régulateurs. Ces différentes expériences révèlent la présence de particules subnucléosomales au niveau des promoteurs et des sites CTCF. Cependant, la distribution et la symétrie des particules subnucléosomales sont un peu différentes au niveau des sites CTCF et des promoteurs qu'au niveau des enhancers.

Concernant le rôle d'un éventuel remodeleur dans la formation des particules subnucléosomales au niveau des promoteurs, aucun des remodeleurs CHD1, CHD4, CHD8, EP400 ou BRG1 ne semble avoir d'effet majeur. Au niveau des sites CTCF, suite aux études du laboratoire du Dr Schübeler et à nos propres données, on peut penser que le remodeleur

SMARCA5 pourrait avoir un rôle dans la formation des particules subnucléosomales (Barisic et al. 2019). Cette hypothèse devra être testée expérimentalement.

Partie B : HDAC2 contribue très modestement à la maladie d'Alzheimer

Dans ce projet, nous avons testé à l'échelle du génome le rôle de répresseur transcriptionnel de l'histone déacétylase HDAC2 un niveau de l'hippocampe du cerveau de souris, dans le contexte de deux modèles de la maladie d'Alzheimer. Nous voulions évaluer la capacité de HDAC2 à induire un blocage épigénétique dans l'hippocampe, sur la base des publications du laboratoire du Dr Li-Huei Tsai. Ce blocage épigénétique entraînerait la répression transcriptionnelle de gènes impliqués dans l'apprentissage, la cognition et la fonction synaptique, et pourrait donc expliquer en partie la perte des fonctions cognitives au cours de la progression de la maladie d'Alzheimer. Il était important de tester cette hypothèse de la composante épigénétique de la maladie d'Alzheimer, car elle ouvre la voie à des traitements basés sur des inhibiteurs d'HDAC2.

Nos expériences de CHIP-seq ont révélé que HDAC2 se lie à la majorité des promoteurs des gènes dans le cerveau de souris, avec un enrichissement s'étendant généralement sur plusieurs kb au début de la région codante des gènes. On peut donc assumer que la protéine HDAC2 est capable de se lier à ses gènes cibles dans l'hippocampe de souris et qu'elle a le potentiel de les réguler.

Le laboratoire du Dr Tsai avait montré que la surexpression de HDAC2 dans le modèle de souris CK-p25 était corrélée à une augmentation de la liaison de HDAC2 au niveau des promoteurs des gènes impliqués dans les processus de formation de la mémoire et de plasticité synaptique (Gräff et al. 2012). Cette augmentation de liaison de HDAC2 entraînerait une réduction du niveau d'acétylation des histones au niveau des promoteurs et donc une répression de l'expression des gènes ciblés.

Nous avons vérifié la surexpression de la protéine HDAC2 dans le modèle de souris 3xTg-AD et dans le modèle App^{tm3 (NL-G-F)}, nommé NLGF dans ma thèse. Une surexpression modeste de

la protéine HDAC2 a uniquement été détectée dans l'hippocampe des souris 3xTg-AD. Aucune surexpression de HDAC2 n'a été détectée dans les souris NLGF.

La surexpression de HDAC2 a donc été détectée dans les modèles 3xTg-AD, CK-p25 et 5xFAD, mais pas dans le modèle NLGF, ce qui montre qu'il s'agit d'un phénomène qui dépend du modèle utilisé. Les modèles 3xTg-AD et CK-p25 sont des modèles plus artificiels que le modèle NLGF car ils sont basés sur la surexpression de transgènes. L'expression de ces transgènes miment des défauts de phosphorylation (hyperphosphorylation de la protéine TAU) dans les souris 3xTg-AD et CK-p25. Cette hyperphosphorylation, absente chez les souris NLGF, pourrait activer un élément de réponse aux récepteurs des glucocorticoïdes qui contrôle la surexpression de HDAC2 (Gräff et al. 2012).

Afin d'explorer la contribution possible de HDAC2 à un blocage épigénétique dans le modèle 3xTg-AD, j'ai réalisé des expériences de RNA-seq. Le principal objectif de ces expériences consistait à tester si HDAC2 contribuait à la répression transcriptionnelle de gènes down-régulés chez le modèle 3xTg-AD.

Chez le mutant homozygote *Hdac2*^{-/-}, un nombre assez restreint de gènes sont dérégulés dans l'hippocampe (74 gènes dérégulés, dont 47 gènes up-régulés avec un seuil de dérégulation de 1,5).

Dans le contexte 3xTg-AD en présence de deux allèles sauvages de *Hdac2*, 685 gènes sont dérégulés à 9 mois, un stade qui correspond à l'installation des symptômes chez ce modèle. Parmi ces gènes dérégulés, 86 sont down-régulés par rapport au contrôle (avec un seuil de dérégulation de 1,5). Ces 86 gènes sont susceptibles de correspondre à des cibles de HDAC2 qui sont réprimés dans le cadre du blocage épigénétique.

Pour identifier parmi ces 86 gènes ceux dont la répression transcriptionnelle dépend de HDAC2, nous avons introduit l'allèle mutant *Hdac2* à l'état homozygote dans le contexte 3xTg-AD. La comparaison des transcriptomes des hippocampes de souris 3xTg-AD avec ceux des souris 3xTg-AD ; *Hdac2*^{-/-} a permis d'identifier 240 gènes dérégulés, dont 122 up-régulés (avec un seuil de dérégulation de 1,5).

Pour tester si l'hypothèse du blocage épigénétique exercé par HDAC2 dans le contexte du modèle 3xTg-AD est vraie, nous avons recherché les gènes communs entre les 86 gènes down-régulés dans le contexte 3xTg-AD, et les 122 gènes up-régulés dans le contexte 3xTg-AD ;

Hdac2^{-/-}. Or on ne retrouve que 11 gènes communs parmi ces gènes. Cela signifie donc que l'hypothèse du blocage épigénétique médié par HDAC2 n'est potentiellement valide que pour ces 11 gènes. Parmi ces 11 gènes, nous n'avons pas retrouvé les différents gènes liés à la cognition, l'apprentissage et la fonction synaptique qui avaient été décrits dans l'article de Gräff et al. 2012. Nous pouvons conclure de notre étude que l'effet de HDAC2 sur la répression des gènes impliqués dans l'apprentissage et la mémoire est spécifique à l'hippocampe des souris modèles CK-p25.

Le très faible nombre de gènes dérégulés dans le modèle 3xTg-AD (un seul) retrouvé en commun avec les gènes dérégulés soit dans le contexte du modèle CK-p25, soit dans des souris ayant une perte de fonction de HDAC2, confirme cette dépendance au modèle de souris utilisé.

Globalement, nous rejetons l'hypothèse du blocage épigénétique médié par HDAC2 dans l'hippocampe des modèles 3xTg-AD et NLGF. La composante épigénétique conduite par HDAC2 ne semble donc pas avoir une grande contribution aux défauts caractéristiques de la maladie d'Alzheimer dans ces deux modèles de souris.

Les analyses transcriptomiques comparant des souris NLGF modèles de la maladie d'Alzheimer jeunes et âgées et les analyses transcriptomiques comparant des souris contrôles et NLGF âgées ont révélé une dérégulation des gènes. Les souris NLGF, qui sont un modèle moins artificiel de la maladie d'Alzheimer, ont donc permis de mettre en évidence des dérégulations transcriptionnelles indépendantes de HDAC2.

Ces gènes dérégulés dans le contexte du modèle NLGF et particulièrement dans les souris NLGF âgées de 9 mois sont significativement enrichis dans les voies biologiques ayant un rôle dans le processus de l'inflammation et de l'immunité. Ces observations sont cohérentes avec la littérature qui indique que le système immunitaire et le processus d'inflammation sont impliqués dans la maladie d'Alzheimer. Par exemple, la relation entre la pathologie tau et le système immunitaire a été étudiée chez un modèle de souris développant une tauopathie dans l'hippocampe et des déficits de mémoire spatiale (Laurent et al. 2017). Les auteurs ont mis en évidence que la tauopathie conduit à l'activation des cellules microgliales et des

astrocytes comme précédemment décrit, mais aussi au déclenchement d'une réaction des lymphocytes T de l'hippocampe (Laurent et al. 2017). Ces lymphocytes T entraînent une progression des défauts de la mémoire spatiale en favorisant les réponses neuroinflammatoires innées, sans perturbation du dépôt ou de la phosphorylation des protéines TAU (Laurent et al. 2017).

La maladie d'Alzheimer semble donc avoir des causes encore inconnues qui entraînent une perturbation de différents processus biologiques notamment au niveau du système immunitaire (inflammation...) ou du système circulatoire qui vont aggraver les symptômes de la maladie.

Perspectives

I. Approfondir la caractérisation du rôle de BRG1 dans la formation des particules subnucléosomales aux enhancers

D'après nos résultats, nous proposons que le remodelleur BRG1 favorise la formation d'un nucléosome fragile central et de particules subnucléosomales contenant les quatre histones au niveau des enhancers. Trois modes d'action de BRG1 peuvent être proposés pour former des particules subnucléosomales :

- (1) Si BRG1 agit sur un nucléosome canonique, il va le remodeler pour former deux hémisomes.
- (2) Si BRG1 agit sur un nucléosome fragile, il va le remodeler potentiellement en hexasome avec la perte d'un seul dimère H2A/H2B.
- (3) Si BRG1 agit sur une particule subnucléosomale déjà remodelée de type hexasome, il va remodeler cet hexasome pour former un hémisome et une particule plus petite contenant uniquement H3/H4.

Nos approches de type « bulk chromatin » (chromatine préparée à partir de 1 à plusieurs millions de cellules) nous permettent d'observer uniquement des phénomènes détectables sur un grand nombre de molécules d'ADN. Or, ces approches ne nous permettent pas de trancher entre les hypothèses précédentes. Ces hypothèses ne sont toutefois pas mutuellement exclusives, et il est possible que BRG1 soit capable de générer différents types de particules subnucléosomales.

Pour mieux comprendre la nature des particules subnucléosomales, nous envisageons de mettre en œuvre des expériences de *single-molecule footprinting* en présence et en absence de BRG1. Comme expliqué dans mon introduction, le *single-molecule footprinting* permet la détection et la quantification des contacts *in vivo* entre protéines (TFs, Pol II, nucléosomes) et ADN sur des molécules d'ADN uniques au niveau des éléments régulateurs (Krebs et al. 2017; Sönmezer et al. 2021). Le *single-molecule footprinting* repose sur la méthylation des régions d'ADN non protégées par des enzymes méthyltransférases recombinantes. Les régions

protégées par des TFs ou des nucléosomes ne sont pas méthylées et sont différenciables par leur taille.

Un des challenge de cette méthode est de choisir une enzyme méthyltransférase efficace pour avoir une bonne résolution dans les expériences de *single-molecule footprinting*. Or, le génome mammifère est déjà naturellement méthylé au niveau des nucléotides CG. Les noyaux des cellules ES pourront être incubés avec une enzyme méthyltransférase qui méthyle les nucléotides GC, comme la M.CviPI (résolution spatiale d'environ 14 pb) (Sönmezer et al. 2021). Par conséquent, si BRG1 remodèle un nucléosome canonique en hémisome, les nucléotides GC pourront être méthylés des deux côtés des hémisomes car l'ADN sera accessible. Si BRG1 remodèle un nucléosome fragile en hexasome, la méthylation des nucléotides GC ne pourra avoir lieu que du côté où le dimère H2A/H2B a été éliminé.

Un autre challenge est de réussir à capturer et séquencer les molécules d'ADN positionnées uniquement au niveau des régions d'intérêt. Le laboratoire du Dr Krebs utilise des sondes qui capturent uniquement les régions régulatrices du génome, ce qui représente environ 2 % du génome, afin de les séquencer (Sönmezer et al. 2021). Cependant, cette méthode permet d'obtenir une profondeur d'environ 70 molécules d'ADN par élément régulateur. Or, si nos particules subnucléosomales ne sont présentes que 10 % du temps en raison d'une forte dynamique, alors nous pourrions ne détecter que 7 molécules d'ADN par élément régulateur ce qui est insuffisant pour une analyse statistique.

Pour pallier à ce problème, nous nous focaliserons sur des enhancers précis et des *primers* seront *designés* pour cibler précisément ces régions enhancers. Les séquences codant ces régions seront amplifiées par PCR. Les amplicons obtenus seront séquencés à haut débit. Nous pourrions ainsi obtenir 1 000 molécules d'ADN sur nos régions d'intérêt et cela augmenterait nos chances de détecter les régions enhancers portant nos particules subnucléosomales.

Grâce à l'association d'expériences de *single-molecule footprinting* et d'expériences de ChIP-seq anti-histones ou anti-TFs, nous allons essayer de mieux comprendre et interpréter certaines informations obtenues avec la chromatine de plusieurs millions de cellules. Par exemple, nous pourrions déterminer si les deux particules subnucléosomales sont présentes simultanément ou non de part et d'autre de l'axe de la dyade du nucléosome fragile dans la

région centrale des enhanceurs. En effet, leur présence simultanée serait un fort indice de la séparation du nucléosome central en deux sous-particules de type hémisome.

II. Visualisation des particules subnucléosomales par microscopie électronique

Au cours de ma thèse, j'ai essayé de caractériser le contenu en histones, la distribution et la fonction des particules subnucléosomales présentes dans les fractions 5-6 issues des expériences de centrifugation sur gradient de sucrose. Cependant la structure précise de ces particules subnucléosomales reste à caractériser. Des expériences de microscopie électronique pourraient permettre de visualiser ces particules.

Une limite de la microscopie électronique pourrait être la capacité à retrouver nos particules subnucléosomales dans l'échantillon. Afin d'enrichir nos échantillons de chromatine fixée en particules subnucléosomales, nous pourrions réaliser un CHIP anti-OCT4 sur les fractions 5-6 des gradients. De cette manière, l'élution du CHIP anti-OCT4 serait enrichie en des particules subnucléosomales d'intérêt, du fait que OCT4 est principalement lié aux enhanceurs.

Des échantillons de chromatine fixée seront préparés en présence et en absence du remodeleur BRG1 afin de visualiser l'impact de BRG1 sur la structure des particules subnucléosomales.

III. Exploration de la composition en histones des particules subnucléosomales

Nos expériences de CHIP-seq sur les différentes fractions du gradient ont montré que les quatre histones sont efficacement immunoprécipitées. Ce résultat suggère que les particules subnucléosomales comportent chacune les quatre histones. Pour le démontrer, il faudrait être capable de réaliser des expériences de CHIP-seq séquentielles (CHIP-seq tandem) avec deux anticorps anti-histone sur de la chromatine issue des fractions 5-6 des gradients comportant les particules subnucléosomales.

Pour cela, nous envisageons de construire une lignée de cellules ES dans laquelle l'histone H3.3 est étiquetée par HA-Flag en C-terminal. Ces cellules seront utilisées dans des expériences de CHIP-tandem en deux étapes : enrichissement en fragments de chromatine portant H3.3 grâce à une immunoprécipitation (IP) anti-Flag, puis élution par compétition avec le peptide FLAG et deuxième enrichissement en fragments de chromatine portant H3.3 et une autre histone grâce à une IP anti-H2A, anti-H2B ou anti-H4.

Nous pourrions ainsi prouver la colocalisation entre plusieurs histones dans les particules subnucléosomales au niveau des enhanceurs.

Conclusion

Au cours de ma thèse, les nombreuses expériences de MNase-ChIP-seq réalisées ont permis de définir avec précision l'organisation des enhancers dans les cellules ES, ainsi que la caractérisation de la fonction de BRG1 dans le remodelage des particules nucléosomales au niveau des enhancers.

La perte de fonction du remodeleur BRG1, associée à des cas de cancers humains, entraîne une perturbation de l'organisation des particules nucléosomales et de la formation des particules subnucléosomales au niveau des enhancers.

De nombreuses hypothèses restent encore à explorer, notamment concernant la structure précise des particules subnucléosomales aux enhancers et leur existence au niveau des autres éléments régulateurs de la transcription. Cependant, un rétablissement de la fonction du remodeleur BRG1 et/ou un rétablissement de l'organisation nucléosomale aux enhancers sont des pistes thérapeutiques à explorer en vue de mieux comprendre et de guérir les cancers. En effet, rétablir une organisation structurée des enhancers permettrait de maintenir une régulation stable de l'expression des gènes.

Notre étude concernant l'implication de la protéine HDAC2 dans la maladie d'Alzheimer, précédemment identifiée comme un dérégulateur épigénétique capable de déclencher un blocage épigénétique, a révélé que HDAC2 a une fonction très modeste sur le transcriptome de l'hippocampe du cerveau des souris. Il sera donc important de publier notre travail afin que du temps ne soit pas perdu à développer des thérapeutiques ciblant HDAC2 ou les gènes cibles de HDAC2 dans le but de soigner la maladie d'Alzheimer.

La maladie d'Alzheimer semble avoir un impact sur la régulation de la transcription d'un certain nombre de gènes. D'autres acteurs épigénétiques pourraient éventuellement contribuer à cette dérégulation.

Notre étude a également révélé des disparités entre les différents modèles de souris de la maladie d'Alzheimer, ce qui pousse à préférer des modèles murins moins artificiels de la maladie d'Alzheimer.

Bibliographie

- Alilat, M., A. Sivolob, B. Révet, et A. Prunell. 1999. « Nucleosome Dynamics. Protein and DNA Contributions in the Chiral Transition of the Tetrasome, the Histone (H3-H4)₂ Tetramer-DNA Particle ». *Journal of Molecular Biology* 291 (4): 815-41. <https://doi.org/10.1006/jmbi.1999.2988>.
- Allison, L. A., M. Moyle, M. Shales, et C. J. Ingles. 1985. « Extensive Homology among the Largest Subunits of Eukaryotic and Prokaryotic RNA Polymerases ». *Cell* 42 (2): 599-610. [https://doi.org/10.1016/0092-8674\(85\)90117-5](https://doi.org/10.1016/0092-8674(85)90117-5).
- Aranda, Sergi, Gloria Mas, et Luciano Di Croce. 2015. « Regulation of Gene Transcription by Polycomb Proteins ». *Science Advances* 1 (11): e1500737. <https://doi.org/10.1126/sciadv.1500737>.
- Armache, K.-J., H. Kettenberger, et P. Cramer. 2003. « Architecture of Initiation-Competent 12-Subunit RNA Polymerase II ». *Proceedings of the National Academy of Sciences* 100 (12): 6964-68. <https://doi.org/10.1073/pnas.1030608100>.
- Ayala, Rafael, Oliver Willhoft, Ricardo J. Aramayo, Martin Wilkinson, Elizabeth A. McCormack, Lorraine Ocloo, Dale B. Wigley, et Xiaodong Zhang. 2018. « Structure and Regulation of the Human INO80-Nucleosome Complex ». *Nature* 556 (7701): 391-95. <https://doi.org/10.1038/s41586-018-0021-6>.
- Badis, Gwenaël, Michael F. Berger, Anthony A. Philippakis, Shaheynoor Talukder, Andrew R. Gehrke, Savina A. Jaeger, Esther T. Chan, et al. 2009. « Diversity and Complexity in DNA Recognition by Transcription Factors ». *Science (New York, N.Y.)* 324 (5935): 1720-23. <https://doi.org/10.1126/science.1162327>.
- Bao, Yunhe, Kasey Konesky, Young-Jun Park, Simona Rosu, Pamela N. Dyer, Danny Rangasamy, David J. Tremethick, Paul J. Laybourn, et Karolin Luger. 2004. « Nucleosomes Containing the Histone Variant H2A.Bbd Organize Only 118 Base Pairs of DNA ». *The EMBO Journal* 23 (16): 3314-24. <https://doi.org/10.1038/sj.emboj.7600316>.
- Barisic, Darko, Michael B. Stadler, Mario Iurlaro, et Dirk Schübeler. 2019. « Mammalian ISWI and SWI/SNF Selectively Mediate Binding of Distinct Transcription Factors ». *Nature* 569 (7754): 136-40. <https://doi.org/10.1038/s41586-019-1115-5>.
- Bazett-Jones, D. P., E. Mendez, G. J. Czarnota, F. P. Ottensmeyer, et V. G. Allfrey. 1996. « Visualization and Analysis of Unfolded Nucleosomes Associated with Transcribing Chromatin ». *Nucleic Acids Research* 24 (2): 321-29. <https://doi.org/10.1093/nar/24.2.321>.
- Beltran, Manuel, Christopher M. Yates, Lenka Skalska, Marcus Dawson, Filipa P. Reis, Keijo Viiri, Cynthia L. Fisher, et al. 2016. « The Interaction of PRC2 with RNA or Chromatin Is Mutually Antagonistic ». *Genome Research* 26 (7): 896-907. <https://doi.org/10.1101/gr.197632.115>.
- Berg, Debbie L. C. van den, Tim Snoek, Nick P. Mullin, Adam Yates, Karel Bezstarosti, Jeroen Demmers, Ian Chambers, et Raymond A. Poot. 2010. « An Oct4-Centered Protein Interaction Network in Embryonic Stem Cells ». *Cell Stem Cell* 6 (4): 369-81. <https://doi.org/10.1016/j.stem.2010.02.014>.
- Bie, Bihua, Jiang Wu, Hui Yang, Jijun J Xu, David L Brown, et Mohamed Naguib. 2014. « Epigenetic Suppression of Neuroligin 1 Underlies Amyloid-Induced Memory Deficiency ». *Nature Neuroscience* 17 (2): 223-31. <https://doi.org/10.1038/nn.3618>.

- Bintu, Lacramioara, Marta Kopaczynska, Courtney Hodges, Lucyna Lubkowska, Mikhail Kashlev, et Carlos Bustamante. 2011. « The Elongation Rate of RNA Polymerase Determines the Fate of Transcribed Nucleosomes ». *Nature Structural & Molecular Biology* 18 (12): 1394-99. <https://doi.org/10.1038/nsmb.2164>.
- Blackledge, Neil P., Anca M. Farcas, Takashi Kondo, Hamish W. King, Joanna F. McGouran, Lars L. P. Hanssen, Shinsuke Ito, et al. 2014. « Variant PRC1 Complex-Dependent H2A Ubiquitylation Drives PRC2 Recruitment and Polycomb Domain Formation ». *Cell* 157 (6): 1445-59. <https://doi.org/10.1016/j.cell.2014.05.004>.
- Bowman, Gregory D., et Michael G. Poirier. 2015. « Post-Translational Modifications of Histones That Influence Nucleosome Dynamics ». *Chemical Reviews* 115 (6): 2274-95. <https://doi.org/10.1021/cr500350x>.
- Boyer, Laurie A., Kathrin Plath, Julia Zeitlinger, Tobias Brambrink, Lea A. Medeiros, Tong Ihn Lee, Stuart S. Levine, et al. 2006. « Polycomb Complexes Repress Developmental Regulators in Murine Embryonic Stem Cells ». *Nature* 441 (7091): 349-53. <https://doi.org/10.1038/nature04733>.
- Brahma, Sandipan, et Steven Henikoff. 2019. « RSC-Associated Subnucleosomes Define MNase-Sensitive Promoters in Yeast ». *Molecular Cell* 73 (2): 238-249.e3. <https://doi.org/10.1016/j.molcel.2018.10.046>.
- Brehove, Matthew, Elan Shatoff, Benjamin T. Donovan, Caroline M. Jipa, Ralf Bundschuh, et Michael G. Poirier. 2019. « DNA Sequence Influences Hexasome Orientation to Regulate DNA Accessibility ». *Nucleic Acids Research* 47 (11): 5617-33. <https://doi.org/10.1093/nar/gkz283>.
- Buratowski, S., et H. Zhou. 1993. « Functional Domains of Transcription Factor TFIIB ». *Proceedings of the National Academy of Sciences of the United States of America* 90 (12): 5633-37. <https://doi.org/10.1073/pnas.90.12.5633>.
- Buratowski, Stephen. 2009. « Progression through the RNA Polymerase II CTD Cycle ». *Molecular Cell* 36 (4): 541-46. <https://doi.org/10.1016/j.molcel.2009.10.019>.
- Čabart, Pavel, Andrea Újvári, Mahadeb Pal, et Donal S. Luse. 2011. « Transcription Factor TFIIF Is Not Required for Initiation by RNA Polymerase II, but It Is Essential to Stabilize Transcription Factor TFIIB in Early Elongation Complexes ». *Proceedings of the National Academy of Sciences of the United States of America* 108 (38): 15786-91. <https://doi.org/10.1073/pnas.1104591108>.
- Carninci, Piero, Albin Sandelin, Boris Lenhard, Shintaro Katayama, Kazuro Shimokawa, Jasmina Ponjavic, Colin A. M. Semple, et al. 2006. « Genome-Wide Analysis of Mammalian Promoter Architecture and Evolution ». *Nature Genetics* 38 (6): 626-35. <https://doi.org/10.1038/ng1789>.
- Chalkley, G. E. 1999. « DNA binding site selection by RNA polymerase II TAFs: a TAFII250-TAFII150 complex recognizes the Initiator ». *The EMBO Journal* 18 (17): 4835-45. <https://doi.org/10.1093/emboj/18.17.4835>.
- Chasman, D. I., K. M. Flaherty, P. A. Sharp, et R. D. Kornberg. 1993. « Crystal Structure of Yeast TATA-Binding Protein and Model for Interaction with DNA ». *Proceedings of the National Academy of Sciences of the United States of America* 90 (17): 8174-78. <https://doi.org/10.1073/pnas.90.17.8174>.
- Chen, Shifu, Yanqing Zhou, Yaru Chen, et Jia Gu. 2018. « Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor ». *Bioinformatics (Oxford, England)* 34 (17): i884-90. <https://doi.org/10.1093/bioinformatics/bty560>.

- Chen, Xi, Han Xu, Ping Yuan, Fang Fang, Mikael Huss, Vinsensius B. Vega, Eleanor Wong, et al. 2008a. « Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells ». *Cell* 133 (6): 1106-17. <https://doi.org/10.1016/j.cell.2008.04.043>.
- Chen, Xiao-Fen, Lynn Lehmann, Justin J. Lin, Ajay Vashisht, Ryan Schmidt, Roberto Ferrari, Chengyang Huang, et al. 2012. « Mediator and SAGA Have Distinct Roles in Pol II Preinitiation Complex Assembly and Function ». *Cell Reports* 2 (5): 1061-67. <https://doi.org/10.1016/j.celrep.2012.10.019>.
- Chittock, Emily C., Sebastian Latwiel, Thomas C. R. Miller, et Christoph W. Müller. 2017. « Molecular Architecture of Polycomb Repressive Complexes ». *Biochemical Society Transactions* 45 (1): 193-205. <https://doi.org/10.1042/BST20160173>.
- Chittori, Sagar, Jingjun Hong, Yawen Bai, et Sriram Subramaniam. 2019. « Structure of the Primed State of the ATPase Domain of Chromatin Remodeling Factor ISWI Bound to the Nucleosome ». *Nucleic Acids Research* 47 (17): 9400-9409. <https://doi.org/10.1093/nar/gkz670>.
- Choi, Jeongyoon, Andreas Linus Bachmann, Katharina Tauscher, Christian Benda, Beat Fierz, et Jürg Müller. 2017. « DNA Binding by PHF1 Prolongs PRC2 Residence Time on Chromatin and Thereby Promotes H3K27 Methylation ». *Nature Structural & Molecular Biology* 24 (12): 1039-47. <https://doi.org/10.1038/nsmb.3488>.
- Cianfrocco, Michael A., George A. Kassavetis, Patricia Grob, Jie Fang, Tamar Juven-Gershon, James T. Kadonaga, et Eva Nogales. 2013. « Human TFIID Binds to Core Promoter DNA in a Reorganized Structural State ». *Cell* 152 (1-2): 120-31. <https://doi.org/10.1016/j.cell.2012.12.005>.
- Cirillo, Lisa Ann, Frank Robert Lin, Isabel Cuesta, Dara Friedman, Michal Jarnik, et Kenneth S. Zaret. 2002. « Opening of Compacted Chromatin by Early Developmental Transcription Factors HNF3 (FoxA) and GATA-4 ». *Molecular Cell* 9 (2): 279-89. [https://doi.org/10.1016/s1097-2765\(02\)00459-8](https://doi.org/10.1016/s1097-2765(02)00459-8).
- Clapier, Cedric R., et Bradley R. Cairns. 2012. « Regulation of ISWI Involves Inhibitory Modules Antagonized by Nucleosomal Epitopes ». *Nature* 492 (7428): 280-84. <https://doi.org/10.1038/nature11625>.
- Clapier, Cedric R., Janet Iwasa, Bradley R. Cairns, et Craig L. Peterson. 2017. « Mechanisms of Action and Regulation of ATP-Dependent Chromatin-Remodelling Complexes ». *Nature Reviews. Molecular Cell Biology* 18 (7): 407-22. <https://doi.org/10.1038/nrm.2017.26>.
- Coin, Frédéric, Valentyn Oksenyshyn, et Jean-Marc Egly. 2007. « Distinct Roles for the XPB/P52 and XPD/P44 Subcomplexes of TFIIH in Damaged DNA Opening during Nucleotide Excision Repair ». *Molecular Cell* 26 (2): 245-56. <https://doi.org/10.1016/j.molcel.2007.03.009>.
- Conaway, R. C., K. P. Garrett, J. P. Hanley, et J. W. Conaway. 1991. « Mechanism of Promoter Selection by RNA Polymerase II: Mammalian Transcription Factors Alpha and Beta Gamma Promote Entry of Polymerase into the Preinitiation Complex ». *Proceedings of the National Academy of Sciences of the United States of America* 88 (14): 6205-9. <https://doi.org/10.1073/pnas.88.14.6205>.
- Conway, Eric, Emilia Jerman, Evan Healy, Shinsuke Ito, Daniel Holoch, Giorgio Oliviero, Orla Deevy, et al. 2018. « A Family of Vertebrate-Specific Polycombs Encoded by the LCOR/LCORL Genes Balance PRC2 Subtype Activities ». *Molecular Cell* 70 (3): 408-421.e8. <https://doi.org/10.1016/j.molcel.2018.03.005>.

- Cooper, Sara J., Nathan D. Trinklein, Elizabeth D. Anton, Loan Nguyen, et Richard M. Myers. 2006. « Comprehensive Analysis of Transcriptional Promoter Structure and Function in 1% of the Human Genome ». *Genome Research* 16 (1): 1-10. <https://doi.org/10.1101/gr.4222606>.
- Cooper, Sarah, Anne Grijsenhout, Elizabeth Underwood, Katia Ancelin, Tianyi Zhang, Tatyana B. Nesterova, Burcu Anil-Kirmizitas, et al. 2016. « Jarid2 Binds Mono-Ubiquitylated H2A Lysine 119 to Mediate Crosstalk between Polycomb Complexes PRC1 and PRC2 ». *Nature Communications* 7 (novembre): 13661. <https://doi.org/10.1038/ncomms13661>.
- Corden, J. L., D. L. Cadena, J. M. Ahearn, et M. E. Dahmus. 1985. « A Unique Structure at the Carboxyl Terminus of the Largest Subunit of Eukaryotic RNA Polymerase II ». *Proceedings of the National Academy of Sciences of the United States of America* 82 (23): 7934-38. <https://doi.org/10.1073/pnas.82.23.7934>.
- Core, Leighton J., Joshua J. Waterfall, et John T. Lis. 2008. « Nascent RNA Sequencing Reveals Widespread Pausing and Divergent Initiation at Human Promoters ». *Science (New York, N.Y.)* 322 (5909): 1845-48. <https://doi.org/10.1126/science.1162228>.
- Cramer, P., D. A. Bushnell, et R. D. Kornberg. 2001. « Structural Basis of Transcription: RNA Polymerase II at 2.8 Angstrom Resolution ». *Science (New York, N.Y.)* 292 (5523): 1863-76. <https://doi.org/10.1126/science.1059493>.
- Cruz, Jonathan C., Huang-Chun Tseng, Joseph A. Goldman, Heather Shih, et Li-Huei Tsai. 2003. « Aberrant Cdk5 Activation by P25 Triggers Pathological Events Leading to Neurodegeneration and Neurofibrillary Tangles ». *Neuron* 40 (3): 471-83. [https://doi.org/10.1016/S0896-6273\(03\)00627-5](https://doi.org/10.1016/S0896-6273(03)00627-5).
- Dalal, Yamini, Takehito Furuyama, Danielle Vermaak, et Steven Henikoff. 2007. « Structure, Dynamics, and Evolution of Centromeric Nucleosomes ». *Proceedings of the National Academy of Sciences of the United States of America* 104 (41): 15974-81. <https://doi.org/10.1073/pnas.0707648104>.
- Dalal, Yamini, Hongda Wang, Stuart Lindsay, et Steven Henikoff. 2007. « Tetrameric Structure of Centromeric Nucleosomes in Interphase Drosophila Cells ». *PLoS Biology* 5 (8): e218. <https://doi.org/10.1371/journal.pbio.0050218>.
- Dancy, Beverley M., et Philip A. Cole. 2015. « Protein Lysine Acetylation by P300/CBP ». *Chemical Reviews* 115 (6): 2419-52. <https://doi.org/10.1021/cr500452k>.
- Dang, Weiwei, Mohamedi N. Kagalwala, et Blaine Bartholomew. 2006. « Regulation of ISW2 by Concerted Action of Histone H4 Tail and Extranucleosomal DNA ». *Molecular and Cellular Biology* 26 (20): 7388-96. <https://doi.org/10.1128/MCB.01159-06>.
- Dann, Geoffrey P., Glen P. Liszczak, John D. Bagert, Manuel M. Müller, Uyen T. T. Nguyen, Felix Wojcik, Zachary Z. Brown, et al. 2017. « ISWI Chromatin Remodellers Sense Nucleosome Modifications to Determine Substrate Preference ». *Nature* 548 (7669): 607-11. <https://doi.org/10.1038/nature23671>.
- Davey, Curt A., David F. Sargent, Karolin Luger, Armin W. Maeder, et Timothy J. Richmond. 2002. « Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9 Å Resolution ». *Journal of Molecular Biology* 319 (5): 1097-1113. [https://doi.org/10.1016/S0022-2836\(02\)00386-8](https://doi.org/10.1016/S0022-2836(02)00386-8).
- Dieuleveult, Maud de, Kuangyu Yen, Isabelle Hmitou, Arnaud Depaux, Fayçal Boussouar, Daria Bou Dargham, Sylvie Jounier, et al. 2016. « Genome-Wide Nucleosome Specificity and Function of Chromatin Remodellers in ES Cells ». *Nature* 530 (7588): 113-16. <https://doi.org/10.1038/nature16505>.

- Dillon, N., et P. Sabbattini. 2000. « Functional Gene Expression Domains: Defining the Functional Unit of Eukaryotic Gene Regulation ». *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 22 (7): 657-65. [https://doi.org/10.1002/1521-1878\(200007\)22:7<657::AID-BIES8>3.0.CO;2-2](https://doi.org/10.1002/1521-1878(200007)22:7<657::AID-BIES8>3.0.CO;2-2).
- Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, et Thomas R. Gingeras. 2013. « STAR: Ultrafast Universal RNA-Seq Aligner ». *Bioinformatics (Oxford, England)* 29 (1): 15-21. <https://doi.org/10.1093/bioinformatics/bts635>.
- Dodonova, Svetlana O., Fangjie Zhu, Christian Dienemann, Jussi Taipale, et Patrick Cramer. 2020. « Nucleosome-Bound SOX2 and SOX11 Structures Elucidate Pioneer Factor Function ». *Nature* 580 (7805): 669-72. <https://doi.org/10.1038/s41586-020-2195-y>.
- Donohoe, M. E., X. Zhang, L. McGinnis, J. Biggers, E. Li, et Y. Shi. 1999. « Targeted Disruption of Mouse Yin Yang 1 Transcription Factor Results in Peri-Implantation Lethality ». *Molecular and Cellular Biology* 19 (10): 7237-44. <https://doi.org/10.1128/mcb.19.10.7237>.
- Doody, Rachelle S., Rema Raman, Martin Farlow, Takeshi Iwatsubo, Bruno Vellas, Steven Joffe, Karl Kieburtz, et al. 2013. « A Phase 3 Trial of Semagacestat for Treatment of Alzheimer's Disease ». *New England Journal of Medicine* 369 (4): 341-50. <https://doi.org/10.1056/NEJMoa1210951>.
- Doody, Rachelle S., Ronald G. Thomas, Martin Farlow, Takeshi Iwatsubo, Bruno Vellas, Steven Joffe, Karl Kieburtz, et al. 2014. « Phase 3 Trials of Solanezumab for Mild-to-Moderate Alzheimer's Disease ». *New England Journal of Medicine* 370 (4): 311-21. <https://doi.org/10.1056/NEJMoa1312889>.
- Doyen, Cécile-Marie, Fabien Montel, Thierry Gautier, Hervé Menoni, Cyril Claudet, Marlène Delacour-Larose, Dimitri Angelov, et al. 2006. « Dissection of the Unusual Structural and Functional Properties of the Variant H2A.Bbd Nucleosome ». *The EMBO Journal* 25 (18): 4234-44. <https://doi.org/10.1038/sj.emboj.7601310>.
- Echigoya, Kenta, Masako Koyama, Lumi Negishi, Yoshimasa Takizawa, Yuka Mizukami, Hideki Shimabayashi, Akari Kuroda, et Hitoshi Kurumizaka. 2020. « Nucleosome Binding by the Pioneer Transcription Factor OCT4 ». *Scientific Reports* 10 (1): 11832. <https://doi.org/10.1038/s41598-020-68850-1>.
- Eirín-López, José María, Toyotaka Ishibashi, et Juan Ausió. 2008. « H2A.Bbd: A Quickly Evolving Hypervariable Mammalian Histone That Destabilizes Nucleosomes in an Acetylation-Independent Way ». *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 22 (1): 316-26. <https://doi.org/10.1096/fj.07-9255com>.
- Farnung, Lucas, Seychelle M. Vos, Christoph Wigge, et Patrick Cramer. 2017. « Nucleosome-Chd1 Structure and Implications for Chromatin Remodelling ». *Nature* 550 (7677): 539-42. <https://doi.org/10.1038/nature24046>.
- Fei, Jia, Sharon E. Torigoe, Christopher R. Brown, Mai T. Khuong, George A. Kassavetis, Hinrich Boeger, et James T. Kadonaga. 2015. « The Prenucleosome, a Stable Conformational Isomer of the Nucleosome ». *Genes & Development* 29 (24): 2563-75. <https://doi.org/10.1101/gad.272633.115>.
- Fernandez Garcia, Meilin, Cedric D. Moore, Katharine N. Schulz, Oscar Alberto, Greg Donague, Melissa M. Harrison, Heng Zhu, et Kenneth S. Zaret. 2019. « Structural Features of Transcription Factors Associating with Nucleosome Binding ». *Molecular Cell* 75 (5): 921-932.e6. <https://doi.org/10.1016/j.molcel.2019.06.009>.

- Ferrari, Karin J., Andrea Scelfo, Sriganesh Jammula, Alessandro Cuomo, Iros Barozzi, Alexandra Stützer, Wolfgang Fischle, Tiziana Bonaldi, et Diego Pasini. 2014. « Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity ». *Molecular Cell* 53 (1): 49-62. <https://doi.org/10.1016/j.molcel.2013.10.030>.
- Ferreira, Helder, Andrew Flaus, et Tom Owen-Hughes. 2007. « Histone Modifications Influence the Action of Snf2 Family Remodelling Enzymes by Different Mechanisms ». *Journal of Molecular Biology* 374 (3): 563-79. <https://doi.org/10.1016/j.jmb.2007.09.059>.
- Flaus, Andrew, David M. A. Martin, Geoffrey J. Barton, et Tom Owen-Hughes. 2006. « Identification of Multiple Distinct Snf2 Subfamilies with Conserved Structural Motifs ». *Nucleic Acids Research* 34 (10): 2887-2905. <https://doi.org/10.1093/nar/gkl295>.
- Flaus, Andrew, et Tom Owen-Hughes. 2004. « Mechanisms for ATP-Dependent Chromatin Remodelling: Farewell to the Tuna-Can Octamer? ». *Current Opinion in Genetics & Development* 14 (2): 165-73. <https://doi.org/10.1016/j.gde.2004.01.007>.
- Flores, O., E. Maldonado, et D. Reinberg. 1989. « Factors Involved in Specific Transcription by Mammalian RNA Polymerase II. Factors IIE and IIF Independently Interact with RNA Polymerase II ». *The Journal of Biological Chemistry* 264 (15): 8913-21.
- Francis, Nicole J., Robert E. Kingston, et Christopher L. Woodcock. 2004. « Chromatin Compaction by a Polycomb Group Protein Complex ». *Science (New York, N.Y.)* 306 (5701): 1574-77. <https://doi.org/10.1126/science.1100576>.
- Gamarra, Nathan, Stephanie L. Johnson, Michael J. Trnka, Alma L. Burlingame, et Geeta J. Narlikar. 2018. « The Nucleosomal Acidic Patch Relieves Auto-Inhibition by the ISWI Remodeler SNF2h ». *ELife* 7 (avril). <https://doi.org/10.7554/eLife.35322>.
- Gamper, Armin M., Jaehoon Kim, et Robert G. Roeder. 2009. « The STAGA Subunit ADA2b Is an Important Regulator of Human GCN5 Catalysis ». *Molecular and Cellular Biology* 29 (1): 266-80. <https://doi.org/10.1128/MCB.00315-08>.
- Geiger, J. H., S. Hahn, S. Lee, et P. B. Sigler. 1996. « Crystal Structure of the Yeast TFIIA/TBP/DNA Complex ». *Science (New York, N.Y.)* 272 (5263): 830-36. <https://doi.org/10.1126/science.272.5263.830>.
- Gershenson, Naum I., et Ilya P. Ioshikhes. 2005. « Synergy of Human Pol II Core Promoter Elements Revealed by Statistical Sequence Analysis ». *Bioinformatics (Oxford, England)* 21 (8): 1295-1300. <https://doi.org/10.1093/bioinformatics/bti172>.
- Gordon, Faye, Karolin Luger, et Jeffrey C. Hansen. 2005. « The Core Histone N-Terminal Tail Domains Function Independently and Additively during Salt-Dependent Oligomerization of Nucleosomal Arrays ». *The Journal of Biological Chemistry* 280 (40): 33701-6. <https://doi.org/10.1074/jbc.M507048200>.
- Govindarajan, Nambirajan, Roberto Carlos Agis-Balboa, Jonas Walter, Farahnaz Sananbenesi, et André Fischer. 2011. « Sodium Butyrate Improves Memory Function in an Alzheimer's Disease Mouse Model When Administered at an Advanced Stage of Disease Progression ». *Journal of Alzheimer's Disease: JAD* 26 (1): 187-97. <https://doi.org/10.3233/JAD-2011-110080>.
- Gräff, Johannes, Damien Rei, Ji-Song Guan, Wen-Yuan Wang, Jinsoo Seo, Krista M. Hennig, Thomas J. F. Nieland, et al. 2012. « An Epigenetic Blockade of Cognitive Functions in the Neurodegenerating Brain ». *Nature* 483 (7388): 222-26. <https://doi.org/10.1038/nature10849>.

- Guan, Ji-Song, Stephen J. Haggarty, Emanuela Giacometti, Jan-Hermen Dannenberg, Nadine Joseph, Jun Gao, Thomas J. F. Nieland, et al. 2009. « HDAC2 Negatively Regulates Memory Formation and Synaptic Plasticity ». *Nature* 459 (7243): 55-60. <https://doi.org/10.1038/nature07925>.
- Guelman, Sebastián, Kenji Kozuka, Yifan Mao, Victoria Pham, Mark J. Solloway, John Wang, Jiansheng Wu, Jennie R. Lill, et Jiping Zha. 2009. « The Double-Histone-Acetyltransferase Complex ATAC Is Essential for Mammalian Development ». *Molecular and Cellular Biology* 29 (5): 1176-88. <https://doi.org/10.1128/MCB.01599-08>.
- Gurard-Levin, Zachary A., Jean-Pierre Quivy, et Geneviève Almouzni. 2014. « Histone Chaperones: Assisting Histone Traffic and Nucleosome Dynamics ». *Annual Review of Biochemistry* 83: 487-517. <https://doi.org/10.1146/annurev-biochem-060713-035536>.
- Guzmán, E., et J. T. Lis. 1999. « Transcription Factor TFIID Is Required for Promoter Melting in Vivo ». *Molecular and Cellular Biology* 19 (8): 5652-58. <https://doi.org/10.1128/MCB.19.8.5652>.
- Ha, I., S. Roberts, E. Maldonado, X. Sun, L. U. Kim, M. Green, et D. Reinberg. 1993. « Multiple Functional Domains of Human Transcription Factor IIB: Distinct Interactions with Two General Transcription Factors and RNA Polymerase II ». *Genes & Development* 7 (6): 1021-32. <https://doi.org/10.1101/gad.7.6.1021>.
- Haberle, Vanja, et Alexander Stark. 2018. « Eukaryotic Core Promoters and the Functional Basis of Transcription Initiation ». *Nature Reviews Molecular Cell Biology* 19 (10): 621-37. <https://doi.org/10.1038/s41580-018-0028-8>.
- He, Shuang, Zihan Wu, Yuan Tian, Zishuo Yu, Jiali Yu, Xinxin Wang, Jie Li, Bijun Liu, et Yanhui Xu. 2020. « Structure of Nucleosome-Bound Human BAF Complex ». *Science (New York, N.Y.)* 367 (6480): 875-81. <https://doi.org/10.1126/science.aaz9761>.
- Heintzman, Nathaniel D., Gary C. Hon, R. David Hawkins, Pouya Kheradpour, Alexander Stark, Lindsey F. Harp, Zhen Ye, et al. 2009. « Histone Modifications at Human Enhancers Reflect Global Cell-Type-Specific Gene Expression ». *Nature* 459 (7243): 108-12. <https://doi.org/10.1038/nature07829>.
- Helming, Katherine C., Xiaofeng Wang, et Charles W. M. Roberts. 2014. « Vulnerabilities of Mutant SWI/SNF Complexes in Cancer ». *Cancer Cell* 26 (3): 309-17. <https://doi.org/10.1016/j.ccr.2014.07.018>.
- Henikoff, J. G., J. A. Belsky, K. Krassovsky, D. M. MacAlpine, et S. Henikoff. 2011. « Epigenome Characterization at Single Base-Pair Resolution ». *Proceedings of the National Academy of Sciences* 108 (45): 18318-23. <https://doi.org/10.1073/pnas.1110731108>.
- Hernandez, N. 1993. « TBP, a Universal Eukaryotic Transcription Factor? ». *Genes & Development* 7 (7B): 1291-1308. <https://doi.org/10.1101/gad.7.7b.1291>.
- Ho, Lena, Jehnna L. Ronan, Jiang Wu, Brett T. Staahl, Lei Chen, Ann Kuo, Julie Lessard, Alexey I. Nesvizhskii, Jeff Ranish, et Gerald R. Crabtree. 2009. « An Embryonic Stem Cell Chromatin Remodeling Complex, EsBAF, Is Essential for Embryonic Stem Cell Self-Renewal and Pluripotency ». *Proceedings of the National Academy of Sciences of the United States of America* 106 (13): 5181-86. <https://doi.org/10.1073/pnas.0812889106>.
- Hoffmann, Alexander, Eric Sinn, Tohru Yamamoto, Josephine Wang, Ananda Roy, Masami Horikoshi, et Robert G. Roeder. 1990. « Highly Conserved Core Domain and Unique N

- Terminus with Presumptive Regulatory Motifs in a Human TATA Factor (TFIID) ». *Nature* 346 (6282): 387-90. <https://doi.org/10.1038/346387a0>.
- Højfeldt, Jonas W., Anne Laugesen, Berthe M. Willumsen, Helene Damhofer, Lin Hedehus, Andrey Tvardovskiy, Faizaan Mohammad, Ole N. Jensen, et Kristian Helin. 2018. « Accurate H3K27 Methylation Can Be Established de Novo by SUZ12-Directed PRC2 ». *Nature Structural & Molecular Biology* 25 (3): 225-32. <https://doi.org/10.1038/s41594-018-0036-6>.
- Holmqvist, Per-Henrik, et Mattias Mannervik. 2013. « Genomic Occupancy of the Transcriptional Co-Activators P300 and CBP ». *Transcription* 4 (1): 18-23. <https://doi.org/10.4161/trns.22601>.
- Holstege, F. C., P. C. van der Vliet, et H. T. Timmers. 1996. « Opening of an RNA Polymerase II Promoter Occurs in Two Distinct Steps and Requires the Basal Transcription Factors IIE and IIH ». *The EMBO Journal* 15 (7): 1666-77.
- Imbalzano, A. N., K. S. Zaret, et R. E. Kingston. 1994. « Transcription Factor (TF) IIB and TFIIA Can Independently Increase the Affinity of the TATA-Binding Protein for DNA ». *The Journal of Biological Chemistry* 269 (11): 8280-86.
- Iurlaro, Mario, Michael B. Stadler, Francesca Masoni, Zainab Jagani, Giorgio G. Galli, et Dirk Schübeler. 2021. « Mammalian SWI/SNF Continuously Restores Local Accessibility to Chromatin ». *Nature Genetics* 53 (3): 279-87. <https://doi.org/10.1038/s41588-020-00768-w>.
- Iwafuchi-Doi, Makiko, Greg Donahue, Akshay Kakumanu, Jason A. Watts, Shaun Mahony, B. Franklin Pugh, Dolim Lee, Klaus H. Kaestner, et Kenneth S. Zaret. 2016. « The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation ». *Molecular Cell* 62 (1): 79-91. <https://doi.org/10.1016/j.molcel.2016.03.001>.
- Iwafuchi-Doi, Makiko, et Kenneth S. Zaret. 2016. « Cell Fate Control by Pioneer Transcription Factors ». *Development (Cambridge, England)* 143 (11): 1833-37. <https://doi.org/10.1242/dev.133900>.
- Jackson, V. 1990. « In Vivo Studies on the Dynamics of Histone-DNA Interaction: Evidence for Nucleosome Dissolution during Replication and Transcription and a Low Level of Dissolution Independent of Both ». *Biochemistry* 29 (3): 719-31. <https://doi.org/10.1021/bi00455a019>.
- Jiang, Jianming, Yun-Shen Chan, Yuin-Han Loh, Jun Cai, Guo-Qing Tong, Ching-Aeng Lim, Paul Robson, Sheng Zhong, et Huck-Hui Ng. 2008. « A Core Klf Circuitry Regulates Self-Renewal of Embryonic Stem Cells ». *Nature Cell Biology* 10 (3): 353-60. <https://doi.org/10.1038/ncb1698>.
- Jin, Qihuang, Li-Rong Yu, Lifeng Wang, Zhijing Zhang, Lawryn H. Kasper, Ji-Eun Lee, Chaochen Wang, Paul K. Brindle, Sharon Y. R. Dent, et Kai Ge. 2011. « Distinct Roles of GCN5/PCAF-Mediated H3K9ac and CBP/P300-Mediated H3K18/27ac in Nuclear Receptor Transactivation ». *The EMBO Journal* 30 (2): 249-62. <https://doi.org/10.1038/emboj.2010.318>.
- Jung, Hye Ryung, Diego Pasini, Kristian Helin, et Ole N. Jensen. 2010. « Quantitative Mass Spectrometry of Histones H3.2 and H3.3 in Suz12-Deficient Mouse Embryonic Stem Cells Reveals Distinct, Dynamic Post-Translational Modifications at Lys-27 and Lys-36 ». *Molecular & Cellular Proteomics: MCP* 9 (5): 838-50. <https://doi.org/10.1074/mcp.M900489-MCP200>.

- Kadoch, Cigall, et Gerald R. Crabtree. 2013. « Reversible Disruption of MSWI/SNF (BAF) Complexes by the SS18-SSX Oncogenic Fusion in Synovial Sarcoma ». *Cell* 153 (1): 71-85. <https://doi.org/10.1016/j.cell.2013.02.036>.
- . 2015. « Mammalian SWI/SNF Chromatin Remodeling Complexes and Cancer: Mechanistic Insights Gained from Human Genomics ». *Science Advances* 1 (5): e1500447. <https://doi.org/10.1126/sciadv.1500447>.
- Kadoch, Cigall, Diana C. Hargreaves, Courtney Hodges, Laura Elias, Lena Ho, Jeff Ranish, et Gerald R. Crabtree. 2013. « Proteomic and Bioinformatic Analysis of Mammalian SWI/SNF Complexes Identifies Extensive Roles in Human Malignancy ». *Nature Genetics* 45 (6): 592-601. <https://doi.org/10.1038/ng.2628>.
- Kadonaga, James T. 2012. « Perspectives on the RNA Polymerase II Core Promoter: Perspectives on the RNA Polymerase II Core Promoter ». *Wiley Interdisciplinary Reviews: Developmental Biology* 1 (1): 40-51. <https://doi.org/10.1002/wdev.21>.
- Kagey, Michael H., Jamie J. Newman, Steve Bilodeau, Ye Zhan, David A. Orlando, Nynke L. van Berkum, Christopher C. Ebmeier, et al. 2010. « Mediator and Cohesin Connect Gene Expression and Chromatin Architecture ». *Nature* 467 (7314): 430-35. <https://doi.org/10.1038/nature09380>.
- Kalb, Reinhard, Sebastian Latwiel, H. Irem Baymaz, Pascal W. T. C. Jansen, Christoph W. Müller, Michiel Vermeulen, et Jürg Müller. 2014. « Histone H2A Monoubiquitination Promotes Histone H3 Methylation in Polycomb Repression ». *Nature Structural & Molecular Biology* 21 (6): 569-71. <https://doi.org/10.1038/nsmb.2833>.
- Kasper, Lawryn H., Chunxu Qu, John C. Obenauer, Daniel J. McGoldrick, et Paul K. Brindle. 2014. « Genome-Wide and Single-Cell Analyses Reveal a Context Dependent Relationship between CBP Recruitment and Gene Expression ». *Nucleic Acids Research* 42 (18): 11363-82. <https://doi.org/10.1093/nar/gku827>.
- Kilgore, Mark, Courtney A. Miller, Daniel M. Fass, Krista M. Hennig, Stephen J. Haggarty, J. David Sweatt, et Gavin Rumbaugh. 2010. « Inhibitors of Class 1 Histone Deacetylases Reverse Contextual Memory Deficits in a Mouse Model of Alzheimer's Disease ». *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology* 35 (4): 870-80. <https://doi.org/10.1038/npp.2009.197>.
- King, Hamish W., et Robert J. Klose. 2017. « The Pioneer Factor OCT4 Requires the Chromatin Remodeller BRG1 to Support Gene Regulatory Element Function in Mouse Embryonic Stem Cells ». *ELife* 6 (mars). <https://doi.org/10.7554/eLife.22631>.
- Koboldt, Daniel C., Qunyuan Zhang, David E. Larson, Dong Shen, Michael D. McLellan, Ling Lin, Christopher A. Miller, Elaine R. Mardis, Li Ding, et Richard K. Wilson. 2012. « VarScan 2: Somatic Mutation and Copy Number Alteration Discovery in Cancer by Exome Sequencing ». *Genome Research* 22 (3): 568-76. <https://doi.org/10.1101/gr.129684.111>.
- Krebs, Arnaud R., Jeroen Demmers, Krishanpal Karmodiya, Nan-Chi Chang, Alice Chien Chang, et László Tora. 2010. « ATAC and Mediator Coactivators Form a Stable Complex and Regulate a Set of Non-Coding RNA Genes ». *EMBO Reports* 11 (7): 541-47. <https://doi.org/10.1038/embor.2010.75>.
- Krebs, Arnaud R., Dilek Imanci, Leslie Hoerner, Dimos Gaidatzis, Lukas Burger, et Dirk Schübeler. 2017. « Genome-Wide Single-Molecule Footprinting Reveals High RNA Polymerase II Turnover at Paused Promoters ». *Molecular Cell* 67 (3): 411-422.e4. <https://doi.org/10.1016/j.molcel.2017.06.027>.

- Krebs, Arnaud R., Krishanpal Karmodiya, Marianne Lindahl-Allen, Kevin Struhl, et László Tora. 2011. « SAGA and ATAC Histone Acetyl Transferase Complexes Regulate Distinct Sets of Genes and ATAC Defines a Class of P300-Independent Enhancers ». *Molecular Cell* 44 (3): 410-23. <https://doi.org/10.1016/j.molcel.2011.08.037>.
- Ku, Manching, Richard P. Koche, Esther Rheinbay, Eric M. Mendenhall, Mitsuhiro Endoh, Tarjei S. Mikkelsen, Aviva Presser, et al. 2008. « Genomewide Analysis of PRC1 and PRC2 Occupancy Identifies Two Classes of Bivalent Domains ». *PLoS Genetics* 4 (10): e1000242. <https://doi.org/10.1371/journal.pgen.1000242>.
- Kubik, Slawomir, Maria Jessica Bruzzone, et David Shore. 2017. « Establishing Nucleosome Architecture and Stability at Promoters: Roles of Pioneer Transcription Factors and the RSC Chromatin Remodeler ». *BioEssays* 39 (5): 1600237. <https://doi.org/10.1002/bies.201600237>.
- Kubik, Slawomir, Maria Jessica Bruzzone, Philippe Jacquet, Jean-Luc Falcone, Jacques Rougemont, et David Shore. 2015. « Nucleosome Stability Distinguishes Two Different Promoter Types at All Protein-Coding Genes in Yeast ». *Molecular Cell* 60 (3): 422-34. <https://doi.org/10.1016/j.molcel.2015.10.002>.
- Lai, Fan, Ulf A. Orom, Matteo Cesaroni, Malte Beringer, Dylan J. Taatjes, Gerd A. Blobel, et Ramin Shiekhataar. 2013. « Activating RNAs Associate with Mediator to Enhance Chromatin Architecture and Transcription ». *Nature* 494 (7438): 497-501. <https://doi.org/10.1038/nature11884>.
- Lambert, Samuel A., Arttu Jolma, Laura F. Campitelli, Pratyush K. Das, Yimeng Yin, Mihai Albu, Xiaoting Chen, Jussi Taipale, Timothy R. Hughes, et Matthew T. Weirauch. 2018. « The Human Transcription Factors ». *Cell* 172 (4): 650-65. <https://doi.org/10.1016/j.cell.2018.01.029>.
- Langmead, Ben, et Steven L. Salzberg. 2012. « Fast Gapped-Read Alignment with Bowtie 2 ». *Nature Methods* 9 (4): 357-59. <https://doi.org/10.1038/nmeth.1923>.
- Laugesen, Anne, Jonas Westergaard Højfeldt, et Kristian Helin. 2019. « Molecular Mechanisms Directing PRC2 Recruitment and H3K27 Methylation ». *Molecular Cell* 74 (1): 8-18. <https://doi.org/10.1016/j.molcel.2019.03.011>.
- Laurent, Cyril, Guillaume Dorothée, Stéphane Hunot, Elodie Martin, Yann Monnet, Marie Duchamp, Yuan Dong, et al. 2017. « Hippocampal T Cell Infiltration Promotes Neuroinflammation and Cognitive Decline in a Mouse Model of Tauopathy ». *Brain: A Journal of Neurology* 140 (1): 184-200. <https://doi.org/10.1093/brain/aww270>.
- Lavelle, Christophe, et Ariel Prunell. 2007. « Chromatin Polymorphism and the Nucleosome Superfamily: A Genealogy ». *Cell Cycle (Georgetown, Tex.)* 6 (17): 2113-19. <https://doi.org/10.4161/cc.6.17.4631>.
- Lenhard, Boris, Albin Sandelin, et Piero Carninci. 2012. « Metazoan Promoters: Emerging Characteristics and Insights into Transcriptional Regulation ». *Nature Reviews. Genetics* 13 (4): 233-45. <https://doi.org/10.1038/nrg3163>.
- Lessard, Julie, Jiang I. Wu, Jeffrey A. Ranish, Mimi Wan, Monte M. Winslow, Brett T. Staahl, Hai Wu, Ruedi Aebersold, Isabella A. Graef, et Gerald R. Crabtree. 2007. « An Essential Switch in Subunit Composition of a Chromatin Remodeling Complex during Neural Development ». *Neuron* 55 (2): 201-15. <https://doi.org/10.1016/j.neuron.2007.06.019>.
- Li, Haojie, Robert Liefke, Junyi Jiang, Jesse Vigoda Kurland, Wei Tian, Pujuan Deng, Weidi Zhang, et al. 2017. « Polycomb-like Proteins Link the PRC2 Complex to CpG Islands ». *Nature* 549 (7671): 287-91. <https://doi.org/10.1038/nature23881>.

- Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, et 1000 Genome Project Data Processing Subgroup. 2009. « The Sequence Alignment/Map Format and SAMtools ». *Bioinformatics (Oxford, England)* 25 (16): 2078-79. <https://doi.org/10.1093/bioinformatics/btp352>.
- Liao, Yang, Gordon K. Smyth, et Wei Shi. 2014. « FeatureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features ». *Bioinformatics (Oxford, England)* 30 (7): 923-30. <https://doi.org/10.1093/bioinformatics/btt656>.
- Liu, Xiaoyu, Meijing Li, Xian Xia, Xueming Li, et Zhucheng Chen. 2017. « Mechanism of Chromatin Remodelling Revealed by the Snf2-Nucleosome Structure ». *Nature* 544 (7651): 440-45. <https://doi.org/10.1038/nature22036>.
- Liu, Xin, Ling Wang, Kehao Zhao, Paul R. Thompson, Yousang Hwang, Ronen Marmorstein, et Philip A. Cole. 2008. « The Structural Basis of Protein Acetylation by the P300/CBP Transcriptional Coactivator ». *Nature* 451 (7180): 846-50. <https://doi.org/10.1038/nature06546>.
- Liu, Yang, Keda Zhou, Naifu Zhang, Hui Wei, Yong Zi Tan, Zhening Zhang, Bridget Carragher, Clinton S. Potter, Sheena D'Arcy, et Karolin Luger. 2020. « FACT Caught in the Act of Manipulating the Nucleosome ». *Nature* 577 (7790): 426-31. <https://doi.org/10.1038/s41586-019-1820-0>.
- Louder, Robert K., Yuan He, José Ramón López-Blanco, Jie Fang, Pablo Chacón, et Eva Nogales. 2016. « Structure of Promoter-Bound TFIID and Model of Human Pre-Initiation Complex Assembly ». *Nature* 531 (7596): 604-9. <https://doi.org/10.1038/nature17394>.
- Love, Michael I., Wolfgang Huber, et Simon Anders. 2014. « Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2 ». *Genome Biology* 15 (12): 550. <https://doi.org/10.1186/s13059-014-0550-8>.
- Lovén, Jakob, Heather A. Hoke, Charles Y. Lin, Ashley Lau, David A. Orlando, Christopher R. Vakoc, James E. Bradner, Tong Ihn Lee, et Richard A. Young. 2013. « Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers ». *Cell* 153 (2): 320-34. <https://doi.org/10.1016/j.cell.2013.03.036>.
- Luger, K., A. W. Mäder, R. K. Richmond, D. F. Sargent, et T. J. Richmond. 1997. « Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution ». *Nature* 389 (6648): 251-60. <https://doi.org/10.1038/38444>.
- Luger, Karolin, Mekonnen L. Dechassa, et David J. Tremethick. 2012. « New Insights into Nucleosome and Chromatin Structure: An Ordered State or a Disordered Affair? » *Nature Reviews. Molecular Cell Biology* 13 (7): 436-47. <https://doi.org/10.1038/nrm3382>.
- Markert, Jonathan, et Karolin Luger. 2021. « Nucleosomes Meet Their Remodeler Match ». *Trends in Biochemical Sciences* 46 (1): 41-50. <https://doi.org/10.1016/j.tibs.2020.08.010>.
- Marsoni, Silvia, Giovanna Damia, et Gabriella Camboni. 2008. « A Work in Progress: The Clinical Development of Histone Deacetylase Inhibitors ». *Epigenetics* 3 (3): 164-71. <https://doi.org/10.4161/epi.3.3.6253>.
- Martire, Sara, Jennifer Nguyen, Aishwarya Sundaresan, et Laura A. Banaszynski. 2020. « Differential Contribution of P300 and CBP to Regulatory Element Acetylation in ESCs ». *BMC Molecular and Cell Biology* 21 (1): 55. <https://doi.org/10.1186/s12860-020-00296-9>.

- Mashtalir, Nazar, Hiroshi Suzuki, Daniel P. Farrell, Akshay Sankar, Jie Luo, Martin Filipovski, Andrew R. D'Avino, et al. 2020. « A Structural Model of the Endogenous Human BAF Complex Informs Disease Mechanisms ». *Cell* 183 (3): 802-817.e24. <https://doi.org/10.1016/j.cell.2020.09.051>.
- Masters, Colin L., Randall Bateman, Kaj Blennow, Christopher C. Rowe, Reisa A. Sperling, et Jeffrey L. Cummings. 2015. « Alzheimer's Disease ». *Nature Reviews. Disease Primers* 1 (octobre): 15056. <https://doi.org/10.1038/nrdp.2015.56>.
- Masuda, Akira, Yuki Kobayashi, Naomi Kogo, Takashi Saito, Takaomi C. Saido, et Shigeyoshi Itohara. 2016. « Cognitive Deficits in Single App Knock-in Mouse Models ». *Neurobiology of Learning and Memory* 135 (novembre): 73-82. <https://doi.org/10.1016/j.nlm.2016.07.001>.
- McGhee, J. D., et G. Felsenfeld. 1980. « Nucleosome Structure ». *Annual Review of Biochemistry* 49: 1115-56. <https://doi.org/10.1146/annurev.bi.49.070180.005343>.
- Meers, Michael P., Derek H. Janssens, et Steven Henikoff. 2019. « Pioneer Factor-Nucleosome Binding Events during Differentiation Are Motif Encoded ». *Molecular Cell* 75 (3): 562-575.e5. <https://doi.org/10.1016/j.molcel.2019.05.025>.
- Michael, Alicia K., Ralph S. Grand, Luke Isbel, Simone Cavadini, Zuzanna Kozicka, Georg Kempf, Richard D. Bunker, et al. 2020. « Mechanisms of OCT4-SOX2 Motif Readout on Nucleosomes ». *Science (New York, N.Y.)* 368 (6498): 1460-65. <https://doi.org/10.1126/science.abb0074>.
- Michel, Brittany C., Andrew R. D'Avino, Seth H. Cassel, Nazar Mashtalir, Zachary M. McKenzie, Matthew J. McBride, Alfredo M. Valencia, et al. 2018. « A Non-Canonical SWI/SNF Complex Is a Synthetic Lethal Target in Cancers Driven by BAF Complex Perturbation ». *Nature Cell Biology* 20 (12): 1410-20. <https://doi.org/10.1038/s41556-018-0221-1>.
- Moorthy, Sakthi D., Scott Davidson, Virlana M. Shchuka, Gurdeep Singh, Nakisa Malek-Gilani, Lida Langroudi, Alexandre Martchenko, Vincent So, Neil N. Macpherson, et Jennifer A. Mitchell. 2017. « Enhancers and Super-Enhancers Have an Equivalent Regulatory Role in Embryonic Stem Cells through Regulation of Single or Multiple Genes ». *Genome Research* 27 (2): 246-58. <https://doi.org/10.1101/gr.210930.116>.
- Mozziconacci, Julien, et Jean-Marc Victor. 2003. « Nucleosome Gaping Supports a Functional Structure for the 30nm Chromatin Fiber ». *Journal of Structural Biology* 143 (1): 72-76. [https://doi.org/10.1016/s1047-8477\(03\)00102-3](https://doi.org/10.1016/s1047-8477(03)00102-3).
- Mueller-Planitz, Felix, Henrike Klinker, Johanna Ludwigsen, et Peter B. Becker. 2013. « The ATPase Domain of ISWI Is an Autonomous Nucleosome Remodeling Machine ». *Nature Structural & Molecular Biology* 20 (1): 82-89. <https://doi.org/10.1038/nsmb.2457>.
- Mühlbacher, Wolfgang, Sarah Sainsbury, Matthias Hemann, Merle Hantsche, Simon Neyer, Franz Herzog, et Patrick Cramer. 2014. « Conserved Architecture of the Core RNA Polymerase II Initiation Complex ». *Nature Communications* 5 (juillet): 4310. <https://doi.org/10.1038/ncomms5310>.
- Murzina, Natalia V., Xue-Yuan Pei, Wei Zhang, Mike Sparkes, Jose Vicente-Garcia, J. Venkatesh Pratap, Stephen H. McLaughlin, et al. 2008. « Structural Basis for the Recognition of Histone H4 by the Histone-Chaperone RbAp46 ». *Structure (London, England: 1993)* 16 (7): 1077-85. <https://doi.org/10.1016/j.str.2008.05.006>.
- Nakajima, N., M. Horikoshi, et R. G. Roeder. 1988. « Factors Involved in Specific Transcription by Mammalian RNA Polymerase II: Purification, Genetic Specificity, and TATA Box-Promoter Interactions of TFIID ». *Molecular and Cellular Biology* 8 (10): 4028-40. <https://doi.org/10.1128/mcb.8.10.4028>.

- Nakayama, Robert T., John L. Pulice, Alfredo M. Valencia, Matthew J. McBride, Zachary M. McKenzie, Mark A. Gillespie, Wai Lim Ku, et al. 2017. « SMARCB1 Is Required for Widespread BAF Complex-Mediated Activation of Enhancers and Bivalent Promoters ». *Nature Genetics* 49 (11): 1613-23. <https://doi.org/10.1038/ng.3958>.
- Natsume, Toyoaki, Tomomi Kiyomitsu, Yumiko Saga, et Masato T. Kanemaki. 2016. « Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors ». *Cell Reports* 15 (1): 210-18. <https://doi.org/10.1016/j.celrep.2016.03.001>.
- Noll, M. 1974. « Subunit Structure of Chromatin ». *Nature* 251 (5472): 249-51. <https://doi.org/10.1038/251249a0>.
- Nora, Elphège P., Anton Goloborodko, Anne-Laure Valton, Johan H. Gibcus, Alec Uebersohn, Nezar Abdennur, Job Dekker, Leonid A. Mirny, et Benoit G. Bruneau. 2017. « Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization ». *Cell* 169 (5): 930-944.e22. <https://doi.org/10.1016/j.cell.2017.05.004>.
- Oddo, Salvatore, Antonella Caccamo, Jason D. Shepherd, M. Paul Murphy, Todd E. Golde, Rakez Kaye, Raju Metherate, Mark P. Mattson, Yama Akbari, et Frank M. LaFerla. 2003. « Triple-Transgenic Model of Alzheimer's Disease with Plaques and Tangles: Intracellular Abeta and Synaptic Dysfunction ». *Neuron* 39 (3): 409-21. [https://doi.org/10.1016/s0896-6273\(03\)00434-3](https://doi.org/10.1016/s0896-6273(03)00434-3).
- Ohkuma, Y., et R. G. Roeder. 1994. « Regulation of TFIID ATPase and Kinase Activities by TFIIE during Active Initiation Complex Formation ». *Nature* 368 (6467): 160-63. <https://doi.org/10.1038/368160a0>.
- Ortega, Esther, Srinivasan Rengachari, Ziad Ibrahim, Naghmeh Hoghoughi, Jonathan Gaucher, Alex S. Holehouse, Saadi Khochbin, et Daniel Panne. 2018. « Transcription Factor Dimerization Activates the P300 Acetyltransferase ». *Nature* 562 (7728): 538-44. <https://doi.org/10.1038/s41586-018-0621-1>.
- Partanen, A., J. Motoyama, et C. C. Hui. 1999. « Developmentally Regulated Expression of the Transcriptional Cofactors/Histone Acetyltransferases CBP and P300 during Mouse Embryogenesis ». *The International Journal of Developmental Biology* 43 (6): 487-94.
- Patikoglou, G. A., J. L. Kim, L. Sun, S. H. Yang, T. Kodadek, et S. K. Burley. 1999. « TATA Element Recognition by the TATA Box-Binding Protein Has Been Conserved throughout Evolution ». *Genes & Development* 13 (24): 3217-30. <https://doi.org/10.1101/gad.13.24.3217>.
- Patro, Rob, Geet Duggal, Michael I. Love, Rafael A. Irizarry, et Carl Kingsford. 2017. « Salmon Provides Fast and Bias-Aware Quantification of Transcript Expression ». *Nature Methods* 14 (4): 417-19. <https://doi.org/10.1038/nmeth.4197>.
- Pradhan, Suman K., Trent Su, Linda Yen, Karine Jacquet, Chengyang Huang, Jacques Côté, Siavash K. Kurdistani, et Michael F. Carey. 2016. « EP400 Deposits H3.3 into Promoters and Enhancers during Gene Activation ». *Molecular Cell* 61 (1): 27-38. <https://doi.org/10.1016/j.molcel.2015.10.039>.
- Prior, C. P., C. R. Cantor, E. M. Johnson, V. C. Littau, et V. G. Allfrey. 1983. « Reversible Changes in Nucleosome Structure and Histone H3 Accessibility in Transcriptionally Active and Inactive States of RDNA Chromatin ». *Cell* 34 (3): 1033-42. [https://doi.org/10.1016/0092-8674\(83\)90561-5](https://doi.org/10.1016/0092-8674(83)90561-5).

- Quinlan, Aaron R., et Ira M. Hall. 2010. « BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features ». *Bioinformatics (Oxford, England)* 26 (6): 841-42. <https://doi.org/10.1093/bioinformatics/btq033>.
- Racki, Lisa R., Nariman Naber, Ed Pate, John D. Leonard, Roger Cooke, et Geeta J. Narlikar. 2014. « The Histone H4 Tail Regulates the Conformation of the ATP-Binding Pocket in the SNF2h Chromatin Remodeling Enzyme ». *Journal of Molecular Biology* 426 (10): 2034-44. <https://doi.org/10.1016/j.jmb.2014.02.021>.
- Rada-Iglesias, Alvaro, Ruchi Bajpai, Tomek Swigut, Samantha A. Brugmann, Ryan A. Flynn, et Joanna Wysocka. 2011. « A Unique Chromatin Signature Uncovers Early Developmental Enhancers in Humans ». *Nature* 470 (7333): 279-83. <https://doi.org/10.1038/nature09692>.
- Ramachandran, Srinivas, Kami Ahmad, et Steven Henikoff. 2017. « Transcription and Remodeling Produce Asymmetrically Unwrapped Nucleosomal Intermediates ». *Molecular Cell* 68 (6): 1038-1053.e4. <https://doi.org/10.1016/j.molcel.2017.11.015>.
- Ramírez, Fidel, Devon P. Ryan, Björn Grüning, Vivek Bhardwaj, Fabian Kilpert, Andreas S. Richter, Steffen Heyne, Friederike Dündar, et Thomas Manke. 2016. « DeepTools2: A next Generation Web Server for Deep-Sequencing Data Analysis ». *Nucleic Acids Research* 44 (W1): W160-165. <https://doi.org/10.1093/nar/gkw257>.
- Ran, F. Ann, Patrick D. Hsu, Jason Wright, Vineeta Agarwala, David A. Scott, et Feng Zhang. 2013. « Genome Engineering Using the CRISPR-Cas9 System ». *Nature Protocols* 8 (11): 2281-2308. <https://doi.org/10.1038/nprot.2013.143>.
- Ricobaraza, Ana, Mar Cuadrado-Tejedor, Alberto Pérez-Mediavilla, Diana Frechilla, Joaquin Del Río, et Ana García-Osta. 2009. « Phenylbutyrate Ameliorates Cognitive Deficit and Reduces Tau Pathology in an Alzheimer's Disease Mouse Model ». *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology* 34 (7): 1721-32. <https://doi.org/10.1038/npp.2008.229>.
- Roeder, R. 1996. « The Role of General Initiation Factors in Transcription by RNA Polymerase II ». *Trends in Biochemical Sciences* 21 (9): 327-35. [https://doi.org/10.1016/S0968-0004\(96\)10050-5](https://doi.org/10.1016/S0968-0004(96)10050-5).
- Sainsbury, Sarah, Carrie Bernecky, et Patrick Cramer. 2015. « Structural Basis of Transcription Initiation by RNA Polymerase II ». *Nature Reviews Molecular Cell Biology* 16 (3): 129-43. <https://doi.org/10.1038/nrm3952>.
- Saito, Takashi, Yukio Matsuba, Naomi Mihira, Jiro Takano, Per Nilsson, Shigeyoshi Itoharu, Nobuhisa Iwata, et Takaomi C Saido. 2014. « Single App Knock-in Mouse Models of Alzheimer's Disease ». *Nature Neuroscience* 17 (5): 661-63. <https://doi.org/10.1038/nn.3697>.
- Salloway, Stephen, Reisa Sperling, Nick C. Fox, Kaj Blennow, William Klunk, Murray Raskind, Marwan Sabbagh, et al. 2014. « Two Phase 3 Trials of Bapineuzumab in Mild-to-Moderate Alzheimer's Disease ». *The New England Journal of Medicine* 370 (4): 322-33. <https://doi.org/10.1056/NEJMoa1304839>.
- Schick, Sandra, Sarah Grosche, Katharina Eva Kohl, Danica Drpic, Martin G. Jaeger, Nara C. Marella, Hana Imrichova, et al. 2021. « Acute BAF Perturbation Causes Immediate Changes in Chromatin Accessibility ». *Nature Genetics* 53 (3): 269-78. <https://doi.org/10.1038/s41588-021-00777-3>.
- Schmitges, Frank W., Archana B. Prusty, Mahamadou Faty, Alexandra Stützer, Gondichatnahalli M. Lingaraju, Jonathan Aiwazian, Ragna Sack, et al. 2011. « Histone

- Methylation by PRC2 Is Inhibited by Active Chromatin Marks ». *Molecular Cell* 42 (3): 330-41. <https://doi.org/10.1016/j.molcel.2011.03.025>.
- Shi, Y., E. Seto, L. S. Chang, et T. Shenk. 1991. « Transcriptional Repression by YY1, a Human GLI-Krüppel-Related Protein, and Relief of Repression by Adenovirus E1A Protein ». *Cell* 67 (2): 377-88. [https://doi.org/10.1016/0092-8674\(91\)90189-6](https://doi.org/10.1016/0092-8674(91)90189-6).
- Shogren-Knaak, Michael, Haruhiko Ishii, Jian-Min Sun, Michael J. Pazin, James R. Davie, et Craig L. Peterson. 2006. « Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions ». *Science (New York, N.Y.)* 311 (5762): 844-47. <https://doi.org/10.1126/science.1124000>.
- Singh, Padmanabh, Arpita Konar, Ashish Kumar, Sweta Srivas, et Mahendra K. Thakur. 2015. « Hippocampal Chromatin-Modifying Enzymes Are Pivotal for Scopolamine-Induced Synaptic Plasticity Gene Expression Changes and Memory Impairment ». *Journal of Neurochemistry* 134 (4): 642-51. <https://doi.org/10.1111/jnc.13171>.
- Sönmezer, Can, Rozemarijn Kleinendorst, Dilek Imanci, Guido Barzaghi, Laura Villacorta, Dirk Schübeler, Vladimir Benes, Nacho Molina, et Arnaud Regis Krebs. 2021. « Molecular Co-Occupancy Identifies Transcription Factor Binding Cooperativity In Vivo ». *Molecular Cell* 81 (2): 255-267.e6. <https://doi.org/10.1016/j.molcel.2020.11.015>.
- Soufi, Abdenour, Greg Donahue, et Kenneth S. Zaret. 2012. « Facilitators and Impediments of the Pluripotency Reprogramming Factors' Initial Engagement with the Genome ». *Cell* 151 (5): 994-1004. <https://doi.org/10.1016/j.cell.2012.09.045>.
- Soufi, Abdenour, Meilin Fernandez Garcia, Artur Jaroszewicz, Nebiyu Osman, Matteo Pellegrini, et Kenneth S. Zaret. 2015. « Pioneer Transcription Factors Target Partial DNA Motifs on Nucleosomes to Initiate Reprogramming ». *Cell* 161 (3): 555-68. <https://doi.org/10.1016/j.cell.2015.03.017>.
- Soutourina, Julie. 2018. « Transcription Regulation by the Mediator Complex ». *Nature Reviews. Molecular Cell Biology* 19 (4): 262-74. <https://doi.org/10.1038/nrm.2017.115>.
- Spitz, François, et Eileen E. M. Furlong. 2012. « Transcription Factors: From Enhancer Binding to Developmental Control ». *Nature Reviews Genetics* 13 (9): 613-26. <https://doi.org/10.1038/nrg3207>.
- Staahl, Brett T., Jiong Tang, Wei Wu, Alfred Sun, Aaron D. Gitler, Andrew S. Yoo, et Gerald R. Crabtree. 2013. « Kinetic Analysis of NpBAF to NBAF Switching Reveals Exchange of SS18 with CREST and Integration with Neural Developmental Pathways ». *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 33 (25): 10348-61. <https://doi.org/10.1523/JNEUROSCI.1258-13.2013>.
- Suganuma, Tamaki, José L. Gutiérrez, Bing Li, Laurence Florens, Selene K. Swanson, Michael P. Washburn, Susan M. Abmayr, et Jerry L. Workman. 2008. « ATAC Is a Double Histone Acetyltransferase Complex That Stimulates Nucleosome Sliding ». *Nature Structural & Molecular Biology* 15 (4): 364-72. <https://doi.org/10.1038/nsmb.1397>.
- Sung, You Me, Taehee Lee, Hyejin Yoon, Amanda Marie DiBattista, Jung Min Song, Yoojin Sohn, Emily Isabella Moffat, et al. 2013. « Mercaptoacetamide-Based Class II HDAC Inhibitor Lowers A β Levels and Improves Learning and Memory in a Mouse Model of Alzheimer's Disease ». *Experimental Neurology* 239 (janvier): 192-201. <https://doi.org/10.1016/j.expneurol.2012.10.005>.
- Takahashi, Kazutoshi, et Shinya Yamanaka. 2006. « Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors ». *Cell* 126 (4): 663-76. <https://doi.org/10.1016/j.cell.2006.07.024>.

- Tan, S., Y. Hunziker, D. F. Sargent, et T. J. Richmond. 1996. « Crystal Structure of a Yeast TFIIA/TBP/DNA Complex ». *Nature* 381 (6578): 127-51. <https://doi.org/10.1038/381127a0>.
- Tanay, Amos, Anne H. O'Donnell, Marc Damelin, et Timothy H. Bestor. 2007. « Hyperconserved CpG Domains Underlie Polycomb-Binding Sites ». *Proceedings of the National Academy of Sciences of the United States of America* 104 (13): 5521-26. <https://doi.org/10.1073/pnas.0609746104>.
- Tavares, Lígia, Emilia Dimitrova, David Oxley, Judith Webster, Raymond Poot, Jeroen Demmers, Karel Bezstarosti, et al. 2012. « RYBP-PRC1 Complexes Mediate H2A Ubiquitylation at Polycomb Target Sites Independently of PRC2 and H3K27me3 ». *Cell* 148 (4): 664-78. <https://doi.org/10.1016/j.cell.2011.12.029>.
- Thurman, Robert E., Eric Rynes, Richard Humbert, Jeff Vierstra, Matthew T. Maurano, Eric Haugen, Nathan C. Sheffield, et al. 2012. « The Accessible Chromatin Landscape of the Human Genome ». *Nature* 489 (7414): 75-82. <https://doi.org/10.1038/nature11232>.
- Varet, Hugo, Loraine Brillet-Guéguen, Jean-Yves Coppée, et Marie-Agnès Dillies. 2016. « SARTools: A DESeq2- and EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data ». *PloS One* 11 (6): e0157022. <https://doi.org/10.1371/journal.pone.0157022>.
- Venters, Bryan J., et B. Franklin Pugh. 2009. « How Eukaryotic Genes Are Transcribed ». *Critical Reviews in Biochemistry and Molecular Biology* 44 (2-3): 117-41. <https://doi.org/10.1080/10409230902858785>.
- Vernimmen, Douglas, et Wendy A. Bickmore. 2015. « The Hierarchy of Transcriptional Activation: From Enhancer to Promoter ». *Trends in Genetics: TIG* 31 (12): 696-708. <https://doi.org/10.1016/j.tig.2015.10.004>.
- Wang, Xueyin, Richard D. Paucek, Anne R. Gooding, Zachary Z. Brown, Eva J. Ge, Tom W. Muir, et Thomas R. Cech. 2017. « Molecular Analysis of PRC2 Recruitment to DNA in Chromatin and Its Inhibition by RNA ». *Nature Structural & Molecular Biology* 24 (12): 1028-38. <https://doi.org/10.1038/nsmb.3487>.
- Wang, Yuan-Liang, Francesco Faiola, Muyu Xu, Songqin Pan, et Ernest Martinez. 2008. « Human ATAC Is a GCN5/PCAF-Containing Acetylase Complex with a Novel NC2-like Histone Fold Module That Interacts with the TATA-Binding Protein ». *The Journal of Biological Chemistry* 283 (49): 33808-15. <https://doi.org/10.1074/jbc.M806936200>.
- Weintraub, Abraham S., Charles H. Li, Alicia V. Zamudio, Alla A. Sigova, Nancy M. Hannett, Daniel S. Day, Brian J. Abraham, et al. 2017. « YY1 Is a Structural Regulator of Enhancer-Promoter Loops ». *Cell* 171 (7): 1573-1588.e28. <https://doi.org/10.1016/j.cell.2017.11.008>.
- Weintraub, H., A. Worcel, et B. Alberts. 1976. « A Model for Chromatin Based upon Two Symmetrically Paired Half-Nucleosomes ». *Cell* 9 (3): 409-17. [https://doi.org/10.1016/0092-8674\(76\)90085-4](https://doi.org/10.1016/0092-8674(76)90085-4).
- Whyte, Warren A., David A. Orlando, Denes Hnisz, Brian J. Abraham, Charles Y. Lin, Michael H. Kagey, Peter B. Rahl, Tong Ihn Lee, et Richard A. Young. 2013. « Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes ». *Cell* 153 (2): 307-19. <https://doi.org/10.1016/j.cell.2013.03.035>.
- Winger, Jessica, Ilana M. Nodelman, Robert F. Levandosky, et Gregory D. Bowman. 2018. « A Twist Defect Mechanism for ATP-Dependent Translocation of Nucleosomal DNA ». *ELife* 7 (mai). <https://doi.org/10.7554/eLife.34100>.

- Wingett, Steven W., et Simon Andrews. 2018. « FastQ Screen: A Tool for Multi-Genome Mapping and Quality Control ». *F1000Research* 7: 1338. <https://doi.org/10.12688/f1000research.15931.2>.
- Wit, Elzo de, Erica S. M. Vos, Sjoerd J. B. Holwerda, Christian Valdes-Quezada, Marjon J. A. M. Verstegen, Hans Teunissen, Erik Splinter, Patrick J. Wijchers, Peter H. L. Krijger, et Wouter de Laat. 2015. « CTCF Binding Polarity Determines Chromatin Looping ». *Molecular Cell* 60 (4): 676-84. <https://doi.org/10.1016/j.molcel.2015.09.023>.
- Wurst, W., J. Rossant, V. Prideaux, M. Kownacka, A. Joyner, D. P. Hill, F. Guillemot, S. Gasca, D. Cado, et A. Auerbach. 1995. « A Large-Scale Gene-Trap Screen for Insertional Mutations in Developmentally Regulated Genes in Mice ». *Genetics* 139 (2): 889-99.
- Xi, Yuanxin, Jianhui Yao, Rui Chen, Wei Li, et Xiangwei He. 2011. « Nucleosome Fragility Reveals Novel Functional States of Chromatin and Poises Genes for Activation ». *Genome Research* 21 (5): 718-24. <https://doi.org/10.1101/gr.117101.110>.
- Yamakawa, Hidekuni, Jemie Cheng, Jay Penney, Fan Gao, Richard Rueda, Jun Wang, Satoko Yamakawa, Oleg Kritskiy, Elizabetha Gjoneska, et Li-Huei Tsai. 2017. « The Transcription Factor Sp3 Cooperates with HDAC2 to Regulate Synaptic Function and Plasticity in Neurons ». *Cell Reports* 20 (6): 1319-34. <https://doi.org/10.1016/j.celrep.2017.07.044>.
- Ye, Tao, Arnaud R. Krebs, Mohamed-Amin Choukallah, Celine Keime, Frederic Plewniak, Irwin Davidson, et Laszlo Tora. 2011a. « SeqMINER: An Integrated ChIP-Seq Data Interpretation Platform ». *Nucleic Acids Research* 39 (6): e35. <https://doi.org/10.1093/nar/gkq1287>.
- Yu, Guangchuang, Li-Gen Wang, Yanyan Han, et Qing-Yu He. 2012. « ClusterProfiler: An R Package for Comparing Biological Themes among Gene Clusters ». *Omics: A Journal of Integrative Biology* 16 (5): 284-87. <https://doi.org/10.1089/omi.2011.0118>.
- Zabidi, Muhammad A., et Alexander Stark. 2016. « Regulatory Enhancer–Core-Promoter Communication via Transcription Factors and Cofactors ». *Trends in Genetics* 32 (12): 801-14. <https://doi.org/10.1016/j.tig.2016.10.003>.
- Zaret, Kenneth S., et Jason S. Carroll. 2011. « Pioneer Transcription Factors: Establishing Competence for Gene Expression ». *Genes & Development* 25 (21): 2227-41. <https://doi.org/10.1101/gad.176826.111>.
- Zegerman, Philip, Benito Canas, Darryl Pappin, et Tony Kouzarides. 2002. « Histone H3 Lysine 4 Methylation Disrupts Binding of Nucleosome Remodeling and Deacetylase (NuRD) Repressor Complex ». *The Journal of Biological Chemistry* 277 (14): 11621-24. <https://doi.org/10.1074/jbc.C200045200>.
- Zeng, Lei, Qiang Zhang, Guillermo Gerona-Navarro, Natalia Moshkina, et Ming-Ming Zhou. 2008. « Structural Basis of Site-Specific Histone Recognition by the Bromodomains of Human Coactivators PCAF and CBP/P300 ». *Structure (London, England: 1993)* 16 (4): 643-52. <https://doi.org/10.1016/j.str.2008.01.010>.
- Zhang, Yong, Tao Liu, Clifford A. Meyer, Jérôme Eeckhoute, David S. Johnson, Bradley E. Bernstein, Chad Nusbaum, et al. 2008. « Model-Based Analysis of ChIP-Seq (MACS) ». *Genome Biology* 9 (9): R137. <https://doi.org/10.1186/gb-2008-9-9-r137>.
- Zhang, Zhuo, Amanda Jones, Chiao-Wang Sun, Chao Li, Chia-Wei Chang, Heui-Yun Joo, Qian Dai, et al. 2011. « PRC2 Complexes with JARID2, MTF2, and EsPRC2p48 in ES Cells to Modulate ES Cell Pluripotency and Somatic Cell Reprogramming ». *Stem Cells (Dayton, Ohio)* 29 (2): 229-40. <https://doi.org/10.1002/stem.578>.
- Zhao, Haixin, Zhijun Han, Xinyuan Liu, Junjie Gu, Fan Tang, Gang Wei, et Ying Jin. 2017. « The Chromatin Remodeler Chd4 Maintains Embryonic Stem Cell Identity by Controlling

- Pluripotency- and Differentiation-Associated Genes ». *The Journal of Biological Chemistry* 292 (20): 8507-19. <https://doi.org/10.1074/jbc.M116.770248>.
- Zhao, Yue, Guillaume Lang, Saya Ito, Jacques Bonnet, Eric Metzger, Shun Sawatsubashi, Eriko Suzuki, et al. 2008. « A TFTC/STAGA Module Mediates Histone H2A and H2B Deubiquitination, Coactivates Nuclear Receptors, and Counteracts Heterochromatin Silencing ». *Molecular Cell* 29 (1): 92-101. <https://doi.org/10.1016/j.molcel.2007.12.011>.
- Zhou, Keda, Guillaume Gaullier, et Karolin Luger. 2019. « Nucleosome Structure and Dynamics Are Coming of Age ». *Nature Structural & Molecular Biology* 26 (1): 3-13. <https://doi.org/10.1038/s41594-018-0166-x>.
- Zhu, Fangjie, Lucas Farnung, Eevi Kaasinen, Biswajyoti Sahu, Yimeng Yin, Bei Wei, Svetlana O. Dodonova, et al. 2018. « The Interaction Landscape between Transcription Factors and the Nucleosome ». *Nature* 562 (7725): 76-81. <https://doi.org/10.1038/s41586-018-0549-5>.
- Zlatanova, Jordanka, Thomas C. Bishop, Jean-Marc Victor, Vaughn Jackson, et Ken van Holde. 2009. « The Nucleosome Family: Dynamic and Growing ». *Structure (London, England: 1993)* 17 (2): 160-71. <https://doi.org/10.1016/j.str.2008.12.016>.

Titre : Impact des dérégulations de l'épigénome sur les maladies chez l'homme – études sur des modèles de la maladie d'Alzheimer et de cancers SWI/SNF-dépendants.

Mots clés : nucléosome, éléments cis-régulateurs, remodeleur de la chromatine, épigénomique, cancer SWI/SNF-dépendant, maladie d'Alzheimer.

Résumé : Au niveau des éléments régulateurs du génome mammifère, les remodeleurs de la chromatine rendent accessible l'ADN empaqueté dans le nucléosome, permettant ainsi la liaison des facteurs de transcription (TFs). Au cours de ma thèse, j'ai exploré dans deux projets distincts les mécanismes moléculaires qui contrôlent l'organisation de la chromatine et la transcription :

(1) Dans le premier projet, j'ai re-exploré des travaux antérieurs suggérant que la surexpression de l'histone désacetylase 2 (HDAC2) dans l'hippocampe provoquait l'altération des fonctions cognitives dans la maladie d'Alzheimer. Nous avons détecté une surexpression de faible amplitude de HDAC2 dans l'hippocampe d'un des deux modèles murins de la maladie d'Alzheimer que nous avons choisi pour notre étude, mais pas dans le deuxième modèle. Le faible impact transcriptionnel de HDAC2 chez ces modèles murins suggère un rôle

mineur de HDAC2 dans la maladie d'Alzheimer.

(2) Dans le deuxième projet, j'ai étudié la fonction de la sous-unité BRG1 du complexe de remodelage SWI/SNF, qui est fréquemment muté dans les cancers humains. Nous avons tout d'abord défini l'organisation nucléosomale et subnucléosomale des enhancers, grâce à des approches d'immunoprécipitation de chromatine combinées à la technique du MNase-seq. Nous avons montré que BRG1 génère au niveau des enhancers des particules subnucléosomales qui sont la cible des différents TFs s'y liant in vivo. Nous avons également démontré que le facteur pionnier OCT4 se lie sur ces particules avec un enrichissement meilleur que sur les nucléosomes canoniques. Ces résultats nous ont permis de proposer un nouveau modèle d'organisation et de fonctionnement des enhancers.

Title : Impact of epigenome dysregulations on human diseases – study of Alzheimer's disease models and SWI/SNF-dependent cancer models.

Keywords : nucleosome, cis-regulatory elements, chromatin remodeller, epigenomics, SWI/SNF-dependent cancer, Alzheimer's disease

Abstract : Mammalian chromatin remodellers unwrap nucleosomal DNA to facilitate the binding of transcription factors (TFs). During my PhD, I explored in two distinct projects the molecular mechanisms which control chromatin organization and transcription:

(1) In the first project, I re-explored a previous work suggesting that overexpression of histone deacetylase 2 (HDAC2) in the hippocampus of Alzheimer's disease's patients results in alterations of cognitive function. We detected a low amplitude overexpression of HDAC2 in the hippocampus of only one of the two mouse models of Alzheimer's disease that we selected for our study. The low transcriptional impact of HDAC2 in these mouse models suggests a minor role of HDAC2 in Alzheimer's disease.

(2) In the second project, I studied the function of BRG1, the catalytic subunit of the SWI/SNF chromatin remodelling complex, which is frequently mutated in human cancers. First, we defined the nucleosomal and subnucleosomal organization of enhancers, using chromatin immunoprecipitation approaches combined to MNase-seq technique. We showed that at enhancers, BRG1 generates subnucleosomal particles which are targeted by different TFs in vivo. We also demonstrated that the OCT4 pioneer TF binds more efficiently to these subnucleosomal particles than to canonical nucleosomes. Based on these results, we propose a new model of organization and function for enhancers.