

Université de Montréal

**Structure-Fonction de l'action pionnière de Pax7**

par

**Virginie Bascunana**

Programmes de Biologie Moléculaire

Faculté de Médecine avec le concours de la Faculté des arts et des sciences et de la Faculté de médecine  
vétérinaire.

Thèse présentée à la Faculté de Médecine

en vue de l'obtention du grade de Doctorat

en

Biologie Moléculaire option médecine cellulaire et moléculaire

Août, 2022

Université de Montréal  
Programmes de Biologie Moléculaire  
Faculté de Médecine avec le concours de la Faculté des arts et des sciences et de la Faculté de médecine  
vétérinaire.

Ce mémoire intitulé  
**Structure-Fonction de l'action pionnière de Pax7**  
Présenté par  
**Virginie Bascunana**

a été évalué par un jury composé des personnes suivantes :

Président-rapporteur  
**Michel Cayouette**

Directeur de recherche  
**Jacques Drouin**

Membres du jury  
**Trang Hoang**  
**Mathieu Lupien**  
**Serge McGraw**

## Résumé

Dans l'hypophyse, Pax7 spécifie l'état du lobe intermédiaire grâce à son action pionnière. Cette action implique dans un premier temps une liaison de cette protéine à certains sites d'hétérochromatine et, par la suite, le remodelage de cette hétérochromatine en euchromatine. Le remodelage de l'épigénome par Pax7 permet au facteur de transcription Tpit d'entraîner l'activation du transcriptome mélanotrope par opposition au destin hypophysaire corticotrope entraîné par Tpit en l'absence de Pax7. Pour étudier l'action pionnière de Pax7, différents membres du laboratoire Drouin ont utilisé un modèle de cellules corticotropes. L'expression ectopique du facteur pionnier Pax7 dans ces cellules permet leur transition d'un phénotype corticotrope à mélanotrope. En plus de l'isoforme prédominante de Pax7 exprimée dans l'hypophyse avec lequel les travaux de Budry et Mayran ont été faits, il existe trois autres isoformes de Pax7 qui diffèrent par quelques acides aminés (Q et GL) dans le domaine PAIRED liant l'ADN.

Nous voulions savoir quels étaient les domaines protéiques de Pax7 impliqués pour son activité pionnière.

Nous avons d'abord étudié les propriétés des différents isoformes naturelles de Pax7 présentant des variations dans le domaine PAIRED que ce soient pour leurs capacités de facteur transcription ou de facteur pionnier dans un modèle de cellules de corticotrope. Nous avons constaté que les quatre isoformes de Pax7, ainsi que Pax3, le facteur le plus homologue à Pax7 dans la famille Pax, présentent une capacité similaire à lier la chromatine fermée. Cependant, Pax3 et les isoformes de Pax7 (GL+) sont incapables d'ouvrir la chromatine sur un sous-groupe d'enhancers pionniers nécessaire à l'activation du transcriptome mélanotrope.

Par la suite, par analyse structure-fonction, nous avons identifié une région de Pax7 importante pour sa liaison à soixante pour cent des sites présents dans l'hétérochromatine qui sont ciblés par l'action pionnière et une région importante pour l'ouverture de chromatine aux sites des pionniers sensibles à la variation GL.

Collectivement, nos données ont montré que les acides aminés GL dans le domaine PAIRED en N-terminal et la partie C-terminale de la protéine Pax7 sont critiques pour l'ouverture de la chromatine à des enhancers spécifiques présents dans l'hétérochromatine dont l'activation est nécessaire au destin mélanotrope. D'autre part, les analyses de structure-fonctions nous ont permis d'identifier une région critique de Pax7 nécessaire à son interaction avec l'hétérochromatine sur ces mêmes enhancers.

Mot clé : différenciation, hypophyse, Pax7, action pionnière, isoformes, étude structure-fonction

## Summary

In the pituitary, Pax7 specifies the intermediate lobe state through a pioneer factor action that involves the binding of this protein to certain heterochromatin sites and, subsequently, the remodeling of this heterochromatin into euchromatin. Epigenome remodeling by Pax7 thus allows the transcription factor Tpit to determine the melanotrope fate as opposed to the corticotrope pituitary fate that is driven by Tpit in absence of Pax7. To study the pioneering action of Pax7, Budry et al, and Mayran et al, used a model of corticotrope cells which, following the expression of Pax7, transit from a corticotrope phenotype to melanotic phenotype thanks to its pioneering action. In addition to the predominant Pax7 isoform expressed in pituitary with which the work of Budry and Mayran were done, there are three other Pax7 isoforms that differ by a few amino acids (Q and GL) in the DNA-binding paired domain.

We wanted to know what protein domains of Pax7 was involved in its pioneering activity.

We have investigated the properties of the different Pax7 isoforms with respect to their transcriptional and pioneer capabilities in a model of corticotrope cells. We found that all four Pax7 isoforms together with the closely related factor Pax3, exhibit a similar ability to bind closed chromatin. However, Pax3 and the isoforms of Pax7 (GL+) failed to open the chromatin at a subset of enhancer necessary to activate the melanotrope transcriptome.

In addition, we have identified by structure-function analysis a region of Pax7 that appears important for Pax7 binding at 60 percent of heterochromatin sites that are targeted for pioneering and one region important for chromatin opening to the pioneer sites GL-sensible.

Collectively, our data suggest that the amino acids GL in the paired domain and the C-terminal part are critical for the chromatin opening to heterochromatin sites necessary to melanotrope transcriptome. On the other hand, structure-functions analyses allowed us to identify a critical region of Pax7 required for the interaction with heterochromatin on these same enhancers.

Keywords: differentiation, pituitary gland, Pax7, pioneer activity, isoforms, structure function study

## Table des matières

Résumé.....	3
Summary.....	4
Table des matières .....	5
Liste des tableaux .....	8
Liste des Figures .....	9
Liste des abréviations .....	11
Remerciements.....	13
Chapitre 1 : Introduction .....	14
1.1 Différentiation cellulaire au cours du développement : de la génétique et de l'épigénétique .....	14
1.1.1 Organogenèse : du zygote a un organisme multicellulaire fonctionnel.....	14
1.1.2 Constante génétique : rôle des facteurs de transcription dans l'identité cellulaire .....	14
1.1.3 Compaction de la chromatine et accessibilité de l'ADN .....	16
1.2 Le rôle des facteurs pionniers .....	23
1.2.1 Les facteurs pionniers : origine des modulations épigénétiques nécessaires à la transition cellulaire .....	23
1.2.2 Les facteurs de pluripotence .....	24
1.2.3 Les facteurs pionniers.....	25
1.2.4 Domaines protéiques et acteurs associés à l'action pionnière .....	28
1.3 Pax7 : un facteur pionnier dans l'hypophyse .....	31
1.3.1 Hypophyse et différenciation cellulaire.....	31
1.3.2 Le facteur pionnier Pax7.....	35
1.3.3 Pax7 : un membre de la famille Pax .....	37
1.3.4 Chronologie de l'action pionnière de Pax7 .....	41
1.3.5 Les isoformes de Pax7 .....	43
1.3.6 Alteration de Pax7 et conséquences .....	45
Objectifs et hypothèses du projet de thèse .....	47
Chapitre 2-Résultats .....	49
2.1 Des isoformes Pax7 aux capacités pionnières différentes .....	49
2.1.1 Les isoformes GL+ sont incapables d'activer le transcriptome mélanotrope .....	49
2.1.2 Les isoformes Pax7 ont un recrutement similaire sur le génome .....	56
2.1.3 Pax7 Q+GL+ prime les enhanceurs mélanotropes au lieu de les activer.....	60

2.2 Pax7 (GL-) et Pax3 (GL-) n'ont pas la même capacité pionnière dans l'hypophyse .....	69
2.2.1 Pax3 (GL-) ne parvient pas à ouvrir les enhancers spécifiques des mélanotropes sensibles au GL pour Pax7. ....	69
2-3 Étude structures fonction de la fonction pionnière de Pax7 .....	72
2.3.1 Les acides aminés 431 à 465 de Pax7 sont nécessaires pour son recrutement sur 60% des enhancers pionniers.....	72
2.3.2 Le mutant Pax7465Ala459-461 lie mais n'ouvre pas la chromatine sur un sous-ensemble d'enhancers pionniers.....	76
2-4 Des régions des parties N et C-terminales de Pax7 sont nécessaires pour l'ouverture de la chromatine des enhancers spécifiques des mélanotropes .....	81
2.4.1 Un même sous-ensemble d'enhancers est affecté par des modifications en N et C-terminal de la protéine Pax7 .....	81
2.4.2 Des enhancers mélanotropes aux caractéristiques particulières .....	83
2-5 Conclusions.....	87
2-6-Des mécanismes différents pour expliquer la perte d'ouverture observée dans l'isoforme Pax7 (GL+) et la construction Pax7465Ala459-461 .....	89
Chapitre 3 : Matériels et Méthodes .....	94
Culture cellulaire .....	94
Essais de luciférase .....	94
ChIPseq, ATACseq.....	94
Extraits nucléaires .....	95
Marquage des sondes .....	95
Essai de retard sur gel (EMSA).....	95
Électrophorèse .....	96
RT-qPCR .....	96
ChIP-qPCR.....	96
Analyse des données NGS .....	97
Analyse des motifs.....	98
RIME (Rapid immunoprecipitation mass spectrometry of endogenous proteins) .....	98
Chapitre 4 : Discussions.....	99
4.1 Des isoformes aux fonctions pionnières différentes.....	99
4.2 Le cas de Pax3.....	102
4.3 La partie C-terminal de Pax7 est nécessaire aux différentes étapes de l'action pionnière. ....	103
4.4 Quels mécanismes moléculaires pour expliquer la perte d'ouverture des sites pionniers spécifiques à la différenciation cellulaire des mélanotropes ? .....	105

4.5 Un mécanisme pionnier différent sur les enhancers GL-sensibles mélanotropes.....	106
4.5 Conclusion .....	107
Références.....	109

## Liste des tableaux

Tableau 1- Anticorps utilisés en ChIPs et Western blots.....	95
Tableau 2- Amorces utilisées pour les RT-qPCRs .....	96
Tableau 3- Amorces utilisées pour les CHIP-qPCRs.....	97

## Liste des Figures

### Chapitre 1

Figure 1.1 Différenciation et plasticité cellulaire.....	15
Figure 1.2 Structure de nucléosome.....	17
Figure 1.3 Les différents types de chromatine présent dans le noyau.....	19
Figure 1.4 Localisation des marques d’histone H3 associées à l’activation transcriptionnelle.....	22
Figure 1.5 La différenciation cellulaire des lignées endocrines de l’hypophyse.....	32
Figure 1.6 Clivage différentiel du précurseur hormonal POMC.....	35
Figure 1.7 Action de Pax7 dans l’hypophyse.....	37
Figure 1.8 Catégorisation des protéines de la famille Pax.....	39
Figure 1.9 Structure des facteurs de transcription Pax.....	40
Figure 1.10 Protéines Co-régulatrices interagissant avec les protéines Pax.....	41
Figure 1. 11 Chronologie de l’action pionnière de Pax7.....	43
Figure 1.12 Épissage alternatif du transcrit de Pax7 et isoformes.....	44

### Chapitre 2

Figure 2.1 Mise en place du projet.....	51
Figure 2.2 Les isoformes Pax7 GL+ ne réussissent pas à activer le transcriptome mélanotrope.....	53
Figure 2.3 Toutes les isoformes ont globalement une activité transcriptionnelle similaire.....	55
Figure 2.4 Reproductivité adéquate des ChIPseqs Pax7.....	57
Figure 2.5 Toute les isoformes de Pax7 ont un recrutement génomique similaire.....	59
Figure 2.6 L’ouverture de la chromatine propre à l’isoforme Pax7 Q+GL+ est corrélée avec la perte du transcriptome du mélanotrope.....	62
Figure 2.7 L’isoforme Q+GL+ prime les enhancers mélanotropes au lieu de les activer.....	65
Figure 2.8 Les remodeleurs de chromatine Ash2 et BRG1 sont diminués mais non perdus sur les enhancers GL-sensible (H3K4me1 symétrique) .....	68
Figure 2.9 Pax3 échoue à activer le transcriptome mélanotrope comme Pax7 Q+GL+.....	71
Figure 2.10 Des délétions en C-terminal de la protéine Pax7 impactent sa capacité à se lier aux sites pionniers.....	74
Figure 2.11 Les acides aminés Pax7 431-465 sont nécessaires pour l’activité pionnière.....	78

Figure 2.12 Les acides aminés 459-461 avec un C-terminal complet n'affecte pas la liaison de Pax7 aux sites pionniers.....	80
Figure 2.13 Des variations en C et N-terminal de Pax7 ont des effets similaires sur l'activation du transcriptome du mélanotrope.....	82
Figure 2.14 Les enhancers pionniers spécifiques des mélanotropes présentent des niveaux de H3K9me2 plus élevés que les autres sites pionniers.....	84
Figure 2.15 Les enhancers pionniers spécifiques des mélanotropes gardent un nucléosome central comprenant l'histone H2AZ.....	86
Figure 2.16 Conclusions.....	88
Figure 2.17 Possibles cofacteurs impliqués dans la perte de l'ouverture de la chromatine pour l'isoforme Pax7 Q+GL+ et le mutant Pax7-465Ala459-461.....	92

## Liste des abréviations

ACTH Adreno Corticotropic Hormone

ADN Acide Désoxyribonucléique

AL Anterior Lobe

ARMS Alveolar Rhabdomyosarcoma

ARN Acide Ribonucléique

ATACseq Assay for Transposase of Accessible Chromatin followed by high throughput sequencing

ATP Adenosine Triphosphate

bHLH basic Helix–Loop–Helix

ChIPseq Chromatin Immunoprecipitation followed by high throughput sequencing

DBD DNA Binding Domain

DNA Deoxyribonucleic Acid

EMSA Electrophoretic mobility shift assay

FSH Follicle-Stimulating Hormone

GH Growth Hormone

GR Glucocorticoid Receptor

HAT Histone acetyltransferase

HD Homeodomain

HDAC Histone désacétylase

HMT Histone methyltransferase

HTH Helix-Turn-Helix

IL Intermediate Lobe

LH Luteinizing Hormone

MNase-seq Micrococcal Nuclease followed by high throughput sequencing

MSH alpha-Melanocyte Stimulating Hormone

NGS Séquençage nouvelle génération

OAR Otp/ aristaless/ Rax

OP Octapeptide

Pax Paired box

Pb Paire de base

PD Paired domain

Pitx Pituitary Homeobox

POMC Pro-Opiomelanocortin

Prl Prolactin

RIME Rapid immunoprecipitation mass spectrometry of endogenous proteins

RNA Ribonucleic Acid

TF Transcription Factor

WGBS Whole genome bisulfite sequencing

## Remerciements

Je tiens d'abord à remercier les Drs Michel Cayouette, Trang Hoang, Mathieu Lupien et Serge McGraw, pour avoir accepté d'être rapporteur pour cette thèse. Merci au Dr Jacques Drouin qui m'a encadré tout au long de mes études graduées. Merci de m'avoir donné la chance de faire mon chemin et de développer mon autonomie. Merci à tous les membres du laboratoire du Dr Jacques Drouin, anciens et présents, avec qui j'ai évolué durant le doctorat et qui ont collaboré, de près ou de loin aux travaux de recherches pour cette thèse. Merci à Aurélio et Yves pour leur aide tout au long de ma formation. Merci à Arthur, Aurélio et Amandine pour leur aide en bio-informatique. Merci à Jacques, Aurélio, Alexandre, Audrey et Arthur dont les travaux et réflexions ont permis d'améliorer le contenu de cette thèse. Merci à Ryhem, Juliette, Konstantin, Kevin, Jonathan, Bastien, Vincent et Valérie pour leurs réflexions sur ce projet.

Je leur suis également reconnaissante pour leur soutien au cours de mon doctorat. Avec eux, j'ai vécu des joies, des peines, des coups de gueules, des rires, des hontes qui m'ont fait grandir en tant que personne et je ne l'oublierai pas. Merci à Roy, Aurélie, Dhara, Venetta, Carolina, Thomas, Daniel, Neera, Sophie pour tous ces moments qui ont fait que la route était belle. Merci à ma famille pour leur soutien tout au long de ce chemin. Merci d'avoir cru en moi quoi qu'il arrive.

## Chapitre 1 : Introduction

### 1.1 Différentiation cellulaire au cours du développement : de la génétique et de l'épigénétique

#### 1.1.1 Organogenèse : du zygote a un organisme multicellulaire fonctionnel

La biologie du développement a pour but de comprendre comment, à partir d'une seule cellule issue de la fusion des cellules germinales paternelles et maternelles, on obtient le zygote et on se retrouve neuf mois plus tard avec un humain adulte comprenant des centaines de types cellulaires différents [1].

Pendant le développement d'un organisme, la cellule unique du zygote va se multiplier et ses cellules filles vont se spécifier au cours des divisions. Cette spécification passe par la perte de capacité lors des transitions vers les différentes lignées cellulaires. Ainsi, le zygote et ses dérivatifs totipotents vont donner des cellules pluripotentes qui, elles même, vont se multiplier et se différencier dans différents états fonctionnels. Cette multitude de cellules plus ou moins spécialisées s'agencent et interagissent pour former nos tissus et organes.

Pour aboutir à leurs positions et états finals, les cellules vont être guidées par le microenvironnement cellulaire. Elles vont suivre des signaux moléculaires par interaction directe avec des molécules de surface présentes sur des cellules adjacentes ou par interaction indirecte par sécrétion de morphogènes, d'hormones ou d'autre ARNs ou protéines [2]. Ces processus développementaux suivent une régulation spatio-temporelle précise qui est nécessaire pour une mise en place adéquate des différentes structures de l'organisme.

#### 1.1.2 Constante génétique : rôle des facteurs de transcription dans l'identité cellulaire

Au sein des cellules, les facteurs de transcription (TF) sont une classe de protéines qui lient l'ADN pour activer ou réprimer la transcription des gènes. Il y a une coordination de la régulation menée par l'ensemble des TFs présents dans le noyau cellulaire ; c'est la cascade de régulation transcriptionnelle [3]. Celle-ci se déroule par étape : d'abord, un TF se lie à des régions régulatrices (promoteurs ou enhancers) et active ou réprime l'expression d'autres TFs. Ces nouveaux TFs vont eux même interagir avec de nouvelles régions du génome permettant l'expression d'un réseau complexe de gènes donnant l'identité finale de la cellule hôte. Le parcours menant au destin cellulaire est parsemé de « décisions » qui poussent la cellule à s'engager vers une voie de différenciation particulière. À chaque étape de cet engagement, des contrôles internes sont opérés pour maintenir l'état de différenciation cellulaire et éviter qu'il y ait un retour en arrière ou une permutation de l'identité de la cellule [4] (Figure 1.1).

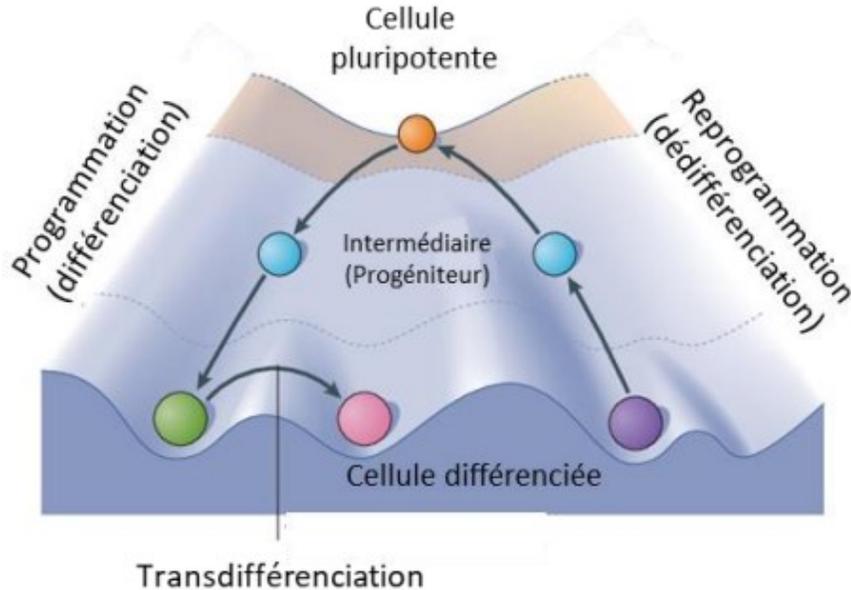


Figure 1.1 Différenciation et plasticité

Adapté de Takahashi & Yamanaka [5]

L'activité des facteurs de transcription est modulée à plusieurs niveaux incluant l'efficacité de transcription du gène qui les encode, l'épissage alternatif de l'ARNm, le temps d'accumulation et de dégradation de la protéine ainsi que des modifications post-traductionnelles [6, 7].

Au cours du développement embryonnaire, l'identité positionnelle des cellules est déterminée par les gènes sélecteurs, des TFs dont l'expression est restreinte à des compartiments particuliers et qui fournissent une identité précoce.

Chez les vertébrés, les gènes sélecteurs des familles Hox et Pax jouent un rôle pour le positionnement cellulaire des organes [8, 9] [10]. Les TFs à homéodomaine (HD) Pax3 et Pax7 confèrent une identité dorsale au tube neural qui est la structure embryonnaire à l'origine du système nerveux central [11], tandis que Pax6 établit des populations distinctes des cellules progénitrices de la zone sous-ventriculaire du cerveau [12]. L'expression aberrante des gènes sélecteurs ou la perturbation d'éléments nécessaires à ces gènes sélecteurs causent un développement pathologique et des malformations congénitales [13-15].

Les mécanismes moléculaires à la base de l'action des gènes sélecteurs sont encore à ce jour mal définis.

Jusqu' à peu, il était proposé que l'action combinatoire de plusieurs TFs spécifiques pourrait expliquer l'expression des gènes définissant la différenciation cellulaire [16]. Cependant, avec la prise en compte de l'environnement épigénétique cellulaire, le modèle à l'origine de l'initiation de nouveaux programmes génétiques s'est complexifié. En effet, l'accessibilité des TFs à l'ADN est dépendant de l'état chromatinien de régions régulatrices cibles.

Pour programmer la transition vers un nouveau type cellulaires, des TFs aux capacités particulières, dit « pionniers », doivent être capables de lier des régions régulatrices présentes dans de la chromatine compactée et donc inaccessibles aux TFs ne possédant la capacité pionnière. Les gènes sélecteurs ou les facteurs impliqués dans la programmation pourraient correspondre au concept de facteurs pionniers [17, 18]. La liaison initiale de facteurs pionniers dans des régions de chromatine compactée permet d'ouvrir localement la chromatine et la rendre facilement accessible pour la liaison des autres TFs non-pionnier ce qui, à terme, entrainerait l'activation du transcriptome nécessaire à la différenciation terminale de la cellule.

### **1.1.3 Compaction de la chromatine et accessibilité de l'ADN**

L'organisation du génome dans le noyau et la structure de la chromatine ont une influence sur le transcriptome exprimé dans une cellule. Les différents niveaux de compaction de l'ADN limitent l'accessibilité au double brin pour la machinerie protéique qui régule la transcription [19].

#### Le nucléosome

Chez les eucaryotes, l'unité de base qui forme la chromatine est le nucléosome constitué de 146 paires de base (pb) d'ADN enroulées sur un octamère protéique comprenant deux molécules de chacune des histones H2A, H2B, H3 et H4 [20] (Figure1.2).

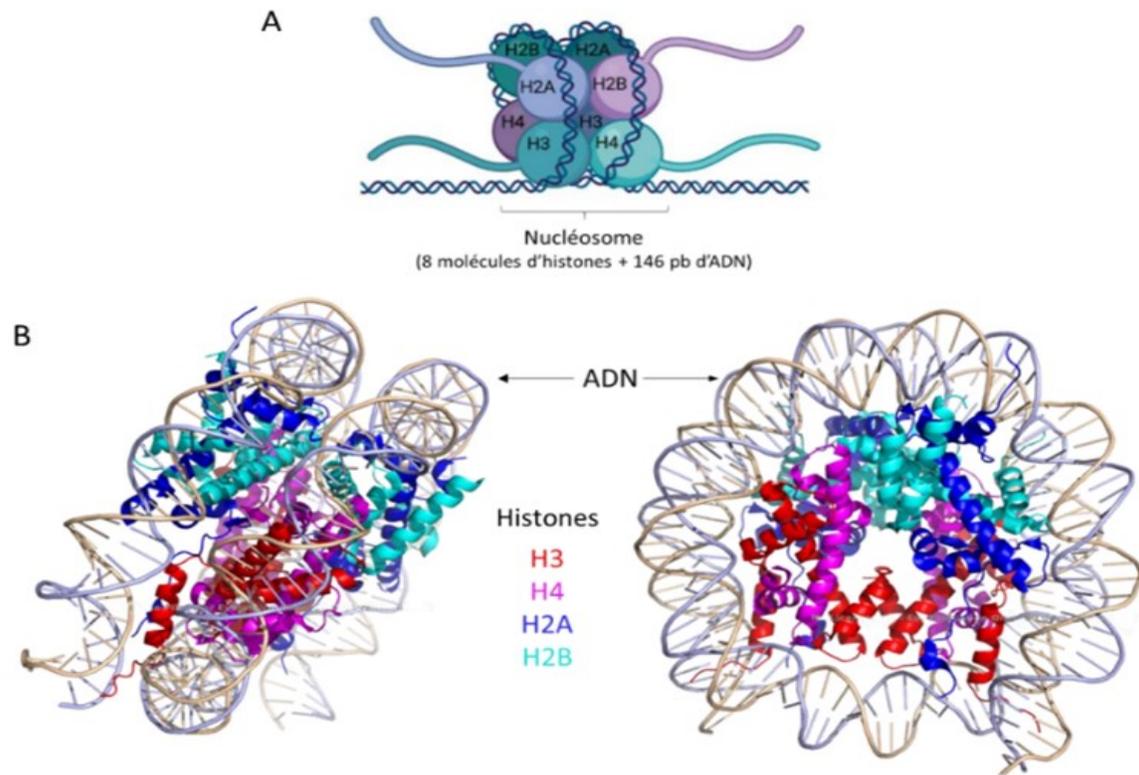


Figure 1.2 Structure de nucléosome

A) Représentation schématique du nucléosome créé à l'aide de BioRender

B) Structure de la particule cœur du nucléosome humain tiré de Tate et al, 2005 [21] modifié à l'aide de PyMOL par Audrey Pelletier

L'enroulement de l'ADN autour du nucléosome a pour conséquence d'empêcher la liaison de protéine à certaines séquences de l'ADN [22]. L'assemblage, la mobilité et la stabilité du nucléosome sont influencés par la séquence nucléotidique de l'ADN [23]. Les protéines du noyau d'histones ont une plus grande affinité pour les séquences d'ADN à haute teneur en GC et pour des intervalles d'environ 10 paires de bases (pb) de dinucléotides AA/TT/TA ce qui favorise la formation de nucléosomes [24-26]. À l'opposé, les séquences riche en poly (dAdT) sont connues pour déstabiliser les nucléosomes [27]. L'enchaînement des nucléosomes est nommé nucléofilament.

Un niveau supérieur de compaction de la chromatine fait intervenir l'ADN internucléosomal et l'histone H1. Leur interaction permet le repliement du nucléofilament en fibre de 30 nm [28]. Le nucléofilament s'organise en boucles de 250 nm pendant l'interphase du cycle cellulaire et le niveau maximal de compaction est le chromosome métaphasique de 850 nm.

Il existe aussi des variants d'histones dont l'incorporation est essentielle dans des domaines spécifiques du génome ou pour certains processus développementaux. Par exemple, H2A-Z modifie la structure de la chromatine afin de réguler la transcription. Ce variant est retrouvé aux promoteurs des gènes [29] et occupe souvent les deux nucléosomes entourant une région libre de nucléosome [30].

### Les différents types de chromatine et leurs effets sur la transcription

Il existe trois grandes catégories de structure d'ordre supérieur de la chromatine: l'euchromatine, l'hétérochromatine facultative et l'hétérochromatine constitutive. Ces structures ont des positions précises au sein du noyau.

L'euchromatine est une forme décondensée localisée au centre du nucléoplasme. C'est la chromatine accessible qui peut être activée par les TFs pour contribuer à l'expression du transcriptome de la cellule.

L'hétérochromatine est une forme de chromatine compacte et transcriptionnellement silencieuse, généralement située en périphérie du noyau. L'hétérochromatine est globalement désacétylée et l'ADN est méthylé, contrairement à l'euchromatine.

L'hétérochromatine constitutive se trouve majoritairement aux niveaux des régions structurales de la cellule comme les centromères et les télomères. Elle contient peu de régions régulatrices et est marquée par l'histone H3K9me3.

L'hétérochromatine facultative est impliquée dans la régulation génique spécifique à un type cellulaire donné. Marquée par l'histone H3K27me2 et H3K9me2, elle permet de réprimer les régions régulatrices de gènes non appropriés pour une identité cellulaire précise [31]. C'est ce type d'hétérochromatine qui peut être lié et modulé positivement par des facteurs pionniers pour aboutir à une transition d'identité cellulaire [32-34] (Figure 1.3).

### Les différents types de chromatine dans le noyau

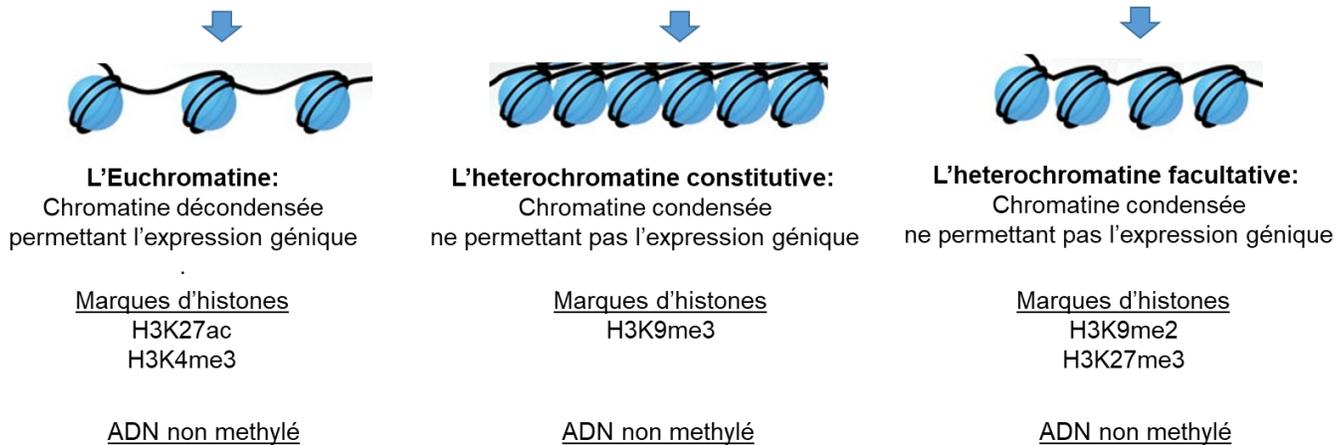


Figure 1.3 Les différents types de chromatine présent dans le noyau

Les marqueurs épigénétiques connus qui affectent la transcription sont la position ou le nombre de nucléosomes sur une zone de régulation génique (modulation de l'accessibilité de l'ADN pour les protéines impliquées dans la transcription), les modifications post-traductionnelles des histones et la méthylation de l'ADN. L'organisation et la maintenance de l'état de la chromatine sont assurées par trois types de facteurs : les facteurs de remodelage de la chromatine, les enzymes de modifications post-traductionnelles des histones et les enzymes modulant la méthylation de l'ADN. Ils peuvent induire des changements conformationnels au niveau du nucléosome aussi bien qu'au niveau de larges domaines de chromatine.

Pour maintenir l'identité cellulaire, les domaines de la chromatine et leurs marques épigénétiques doivent être conservées lorsque la cellule duplique son génome pour la division cellulaire [35].

#### Le remodelage de la chromatine

Les facteurs de remodelage de la chromatine influent sur l'activation ou sur la répression de certains gènes [36]. Pour activer la transcription, ils sont capables de déplacer l'octamère d'histones pour que les régions régulatrices de l'ADN soient accessibles [37]. Parmi les complexes de remodelage de la chromatine, nous avons les familles SWI/SNF (switch/sucrose non-fermentable), ISWI (imitation SWI) et CHD (chromodomain helicase DNA-binding). Ils dissocient l'ADN de la surface de l'octamère d'histone, permettant aux nucléosomes de glisser le long de l'ADN ou d'être dissociés ou déplétés. Pour cela, ils rompent les interactions entre les acides nucléiques et les protéines grâce à l'énergie libérée par l'hydrolyse de l'ATP. Une fois l'ADN rendu accessible, les TFs ont accès à leurs sites de liaison à l'ADN et y recrutent la machinerie transcriptionnelle [38]. Cependant, associés à certains complexes

répressifs, ces remodeleurs de chromatine peuvent réprimer la transcription des gènes. Par exemple le complexe NuRD/Mi-2/CHD entraîne la répression transcriptionnelle qui est requis pour le maintien de la pluripotence des cellules embryonnaires. CHD3/4 sert d'échafaudage pour le recrutement d'autres protéines liées à la répression telles que les histones désacétylases HDAC1/2, les protéines liant l'ADN méthylé MBD2/3, l'histone déméthylase LSD1 ou encore les protéines RBBP4 et RBBP7 qui agissent comme des chaperonnes d'histones [39].

#### Les modifications post-traductionnelles des histones

Les histones sont des protéines très conservées au cours de l'évolution. Elles sont modulées au niveau de leurs extrémités N-terminales qui sont situées hors de l'unité nucléosomale et donc accessibles aux modifications post-traductionnelles. Ces modifications affectent leur charge ce qui peut moduler l'interaction histone-histone ou histone-ADN. Ces changements de conformation histone-ADN peuvent modifier l'accessibilité à l'ADN pour les protéines de la machine de transcription. Les modifications d'histones servent aussi de marqueurs pour recruter différents partenaires pouvant entraîner d'autres changements de conformation chromatinien et/ou des acteurs de la transcription génique. Il existe plusieurs marques d'histones et la multitude de combinaisons possibles explique la diversité des activités de la chromatine dans le noyau [40]. Au cours du développement, la modulation de ces modifications post-traductionnelles est nécessaire pour la différenciation cellulaire et est impliquée dans le maintien de l'identité cellulaire [41] [42].

Un exemple connu est trouvé dans les cellules souches embryonnaires (ES). Elles possèdent des promoteurs bivalents présentant à la fois des marques activatrices (H3K4me3) et répressives (H3K27me3) au niveau des gènes de contrôles développementaux [43] [44]. Ces domaines bivalents formeraient un équilibre pour les gènes dont une activation précoce influencerait la décision du destin cellulaire [45].

Les modifications post-traductionnelles qui peuvent être ajoutées ou enlevées dynamiquement par voie enzymatique comprennent l'acétylation, la méthylation, la phosphorylation, l'ubiquitylation, la sumoylation et l'ADP-ribosylation. Les protéines reconnaissant ces modifications post-traductionnelles ont des domaines protéiques particuliers appelés modules « reader ». Les enzymes qui modifient les histones sont appelées « writers » si elles ajoutent des marques ou « erasers » si elles en suppriment.

Les modifications post-traductionnelles les plus connues pour la modulation de la transcription sont l'acétylation pour l'activation transcriptionnelle et la méthylation qui, en fonction du résidu modifié et du nombre de méthyle ajouté, peut être une marque activatrice ou répressive de la transcription. Ce sont ces types de marques chromatiniennes qui seront étudiées dans mon projet de thèse.

L'acétylation des histones est modulée par les enzymes histone-acétyltransférase (HAT) impliquées dans l'activation génique et les enzymes histone-désacétylases (HDAC) impliquées dans la répression génique [46, 47]. Ces enzymes agissent le plus souvent au sein de complexes protéiques impliqués dans l'activation ou la répression génique.

Les enzymes HAT ajoutent un groupement acétyle aux protéines d'histone, ce qui réduit l'affinité de celles-ci à l'ADN, facilitant l'accès aux TFs [46]. Les HAT les plus étudiées sont CBP et p300. Elles marquent les enhanceurs actifs avec l'acétylation de l'histone H3 à la lysine 27 (H3K27ac) et interagissent avec une multitude d'acteurs de la machinerie transcriptionnelle (Figure 1.4).

A l'opposée, les enzymes HDAC sont capable de retirer un groupement acétyle aux histones ce qui resserre l'ADN autour de l'histone et empêche les TFs d'accéder à leurs sites de liaisons. Les histone-désacétylases les plus connues, HDAC1/2, sont impliquées dans le maintien de la pluripotence cellulaire des cellules embryonnaires. Leur perte déstabilise les complexes corépresseurs Sin3A, NuRD, and CoREST, et entraîne une augmentation globale de l'acétylation des histones. La perte de HDAC1 notamment permet l'activation de la transcription génique associée à la différenciation cellulaire [39, 48].

L'acétylation des queues d'histone peut aussi moduler la transcription en permettant le recrutement de protéines ayant un « reader » nommé bromodomaine. Par exemple, les complexes SWI/SNIF contenant les ATPases apparentées Brg1 ou Brm se lient aux histones acétylées afin de remodeler la chromatine [49] [50].

La méthylation des histones est modulée par les histones méthyltransférase (HMTs). Ces enzymes sont capables d'ajouter d'un à trois groupements méthyles aux résidus lysines (KMTs) ou arginines (PRMTs) au niveau notamment de l'histone H3. Ces marques peuvent être permissives ou répressives [51]. Les HMT sont des composants de complexes protéiques activateurs ou répressifs.

Parmi les complexes protéique HMT responsables de l'activation génique, les complexes MLL (mixed-lineage leukemia) sont importants pour la mise en place de la marque d'enhancer actif H3K4me1 (monométhylation de H3K4) et la marque de promoteur actif H3K4me3 (tri-méthylation de H3K4) [52]. Le complexe SETD2 permet aussi l'activation génique en déposant la marque H3K36me3 au niveau des gènes (Figure 1.4).

Pour les complexes protéiques HMT, responsables de la répression génique, le complexe Ehmt1/Ehmt2 permet la di-méthylation de l'histone H3K9 (H3K9me2) et les complexes polycomb PRC1/2 permettent la déposition de la marque répressive H3K27me3 (tri-méthylation de H3K27) principalement au niveau des promoteurs [53-55].

Le complexe Suv-39h permet aussi de réprimer de larges domaines intergéniques par la déposition de l'histone H3K9me3 (tri-méthylation de H3K9), une marque d'hétérochromatine constitutive. Cette marque est considérée comme un gardien de la stabilité génomique en réprimant notamment les régions non codantes telles que les centrosomes et les régions de séquences répétées. Dans les cellules souches embryonnaires murines et au cours de la reprogrammation cellulaire, H3K9me3 agit comme un obstacle à la transition d'un type cellulaire vers un autre type cellulaire [56].

Les groupements méthyles ajoutés aux histones sont reconnus par des protéines ayant un « reader » nommé chromodomaine. Elles permettent le recrutement d'autres complexes protéiques tel que les complexes SWI-SNF ou PRC1/2. Une des protéines à chromodomaine connue est la protéine HP1alpha. Cette protéine reconnaît la queue d'histone H3K9me et est impliquée dans la répression génique. Elle peut aussi entraîner une séparation de phase et condenser l'ADN localement [57, 58].

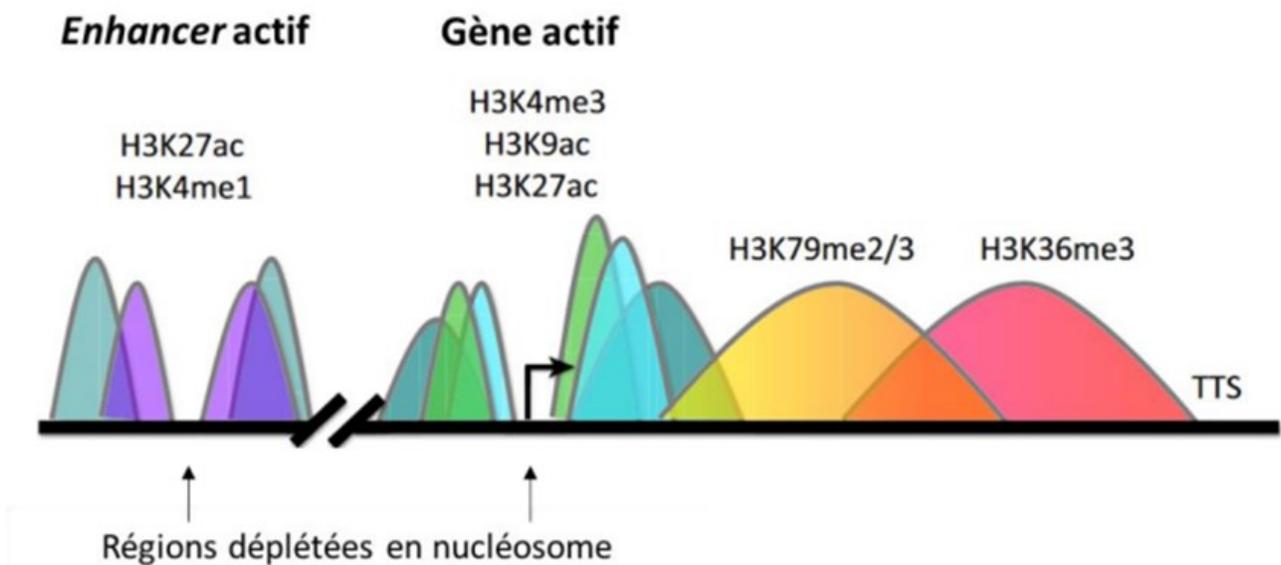


Figure 1.4 Localisation des marques d'histone H3 associées à l'activation transcriptionnelle

Adapté de Gates, Foulds et O'Malley, 2017 [59]

### La méthylation de l'ADN

La méthylation de l'ADN est une marque épigénétique qui reflète la mémoire cellulaire ; elle joue un rôle actif dans la répression géniques [60]. Durant le développement de l'embryon, les patrons de méthylation de l'ADN sont établis par des vagues de méthylation et de déméthylation [61]. Des profils uniques de méthylation de l'ADN ont été observés dans les cellules souches et dans différentes lignées cellulaires [62].

La méthylation de l'ADN consiste en l'addition d'un groupement méthyle à la cytosine du dinucléotide CpG dans l'ADN pour former du 5-méthylcytosine. Chez l'humain, les sites CpG sont peu abondants et inégalement répartis. Soixante pourcents des promoteurs sont riches en sites CpG (îlot CpG) bien que ceux-ci échappent en générale à la méthylation lorsqu'ils sont associés à un gène actif [63-65]. Il existe aussi des sites CpG situés dans les régions répétées, les transposons et les régions intergéniques. Ces régions sont en général moins denses en sites CpG mais ceux-ci sont hyperméthylés [64, 66]. La méthylation de ces régions contribue à la stabilité chromosomique et le contrôle de l'expression des gènes.

Les enzymes responsables pour le dépôt du groupement méthyle sur l'ADN sont les DNA (cytosine-5)-méthyltransférase Dnmt3a , Dnmt3b et dans une moindre mesure Dnmt2 [67]. Cette méthylation de l'ADN est maintenue lors de la division cellulaire par une autre enzyme, Dnmt1 [68]. Les enzymes TET peuvent enlever le méthyle de la méthylcytosine par catalyse enzymatique [69]. Les méthylcytosines sont aussi reconnues spécifiquement par des protéines contenant des domaines MBP (methyl binding protein). Elles médient l'interaction de l'ADN méthylé avec d'autres acteurs de modifications chromatinienne et participent à la formation de structures répressives de la chromatine [70] [71].

## **1.2 Le rôle des facteurs pionniers**

### **1.2.1 Les facteurs pionniers : origine des modulations épigénétiques nécessaires à la transition cellulaire**

Un facteur pionnier est un facteur de transcription capable d'entraîner des changements épigénétiques qui ouvrent la chromatine au niveau des régions régulatrices. En plus détaillé, les facteurs pionniers ont la capacité de lier leurs séquences d'ADN dans l'hétérochromatine, d'affecter la structure chromatinienne, d'entraîner l'ouverture de la chromatine et de maintenir une mémoire épigénétique de ces changements par la déméthylation de l'ADN [72].

Les changements résultant de l'action d'un facteur pionnier permettent l'accès à un nouvel éventail de régions régulatrices pour la transcription génique. Dans un contexte développemental, ils permettent donc la différenciation cellulaire [17, 18, 37, 73, 74]. Les facteurs pionniers semblent plus servir à la spécification que la différenciation terminale des cellules. Pour l'activation du transcriptome final de la cellule, plusieurs facteurs pionniers reconnus nécessitent la coopération d'autres facteurs de transcription pionniers ou non pionniers pour aboutir à la différenciation terminale de la cellule.

Méthode d'analyse permettant d'étudier l'action pionnière

Diverses techniques de séquençage à haut débit offrent la possibilité d'analyser l'environnement chromatinien sur tout le génome. L'accessibilité à l'ADN est évaluée par ATACseq (Assay for Transposon-Accessible Chromatin) et la cartographie des nucléosomes par MNase-seq (Micrococcal Nuclease) [75, 76]. Les différentes marques d'histones sont cartographiées par CHIPseq (immunoprécipitation de la chromatine suivie de séquençage massif). Enfin, l'étude de la méthylation de l'ADN se fait par le traitement au bisulfite de sodium suivi de séquençage massif (BS-seq ou Whole Genome Bisulfite Sequencing – WGBS) [77]. C'est sur ces profils que reposent la compréhension de l'action des facteurs pionniers, dans le contexte du développement et de la différenciation cellulaire

Cependant, tous les facteurs dit pionniers ne sont pas équivalents. Les facteurs pionniers ont été divisés en deux groupes : les facteurs pluripotents qui peuvent entraîner des changements épigénétiques au delà de leur site de liaison et les facteurs pionniers qui permettent des changements épigénétiques restreints à leurs sites de liaisons.

### **1.2.2 Les facteurs de pluripotence**

Les facteurs de pluripotence ont la capacité de reprogrammer les cellules somatiques en cellules souches pluripotentes (induced pluripotent stem - iPS) [78]. Ils sont aux nombres de quatre : Oct4, Sox2 et Klf4 et Myc. Oct4, Sox2 et Klf4 (OSK) sont suffisants pour la dédifférenciation cellulaire tandis que Myc augmente l'efficacité et la vitesse de reprogrammation cellulaire en stabilisant le recrutement des trois autres facteurs aux loci critiques [79]. Les OSKM entraînent des changements épigénétiques sur des mégabases d'ADN; ils effacent l'identité somatique et établissent un programme pluripotent stable. Les changements à l'origine de la transition de l'identité cellulaire somatique à pluripotente sont relativement lents et semblent s'effectuer par vagues de liaison des facteurs OSKM et de changements épigénétiques. Ils se font en deux phases : une phase initiale où les régions régulatrices du transcriptome somatique se referme et une phase finale où les régions régulatrices des capacités pluripotentes sont rendues accessibles et activées [80].

Les régions régulatrices responsables du transcriptome somatique ne présentent pas de motif OSK [81]. Le mécanisme pour les réprimer implique Sap30, une protéine du complexe corépresseur HDAC qui limite le dépôt de la marque H3K27ac au niveau de ces loci [80].

A l'opposé, les régions régulatrices responsables du transcriptome pluripotent sont enrichies en motif OSK. Les OSK lient des régions régulatrices dans la chromatine fermée dont l'ADN est méthylé, puis rendent cette chromatine accessible pour d'autres TFs et co-facteurs [82] [83]. Les facteurs OSK peuvent reconnaître des motifs partiels sur l'ADN nucléosomal [84], ont une forte affinité pour les cibles d'ADN méthylés [85], et une forte synergie entre eux [81, 86]. Aucun profil de marques d'histone particulier n'a toutefois pu être associé à cette activité [86].

Il est à noter que la liaison précoce des facteurs pionniers OSKM ne garantit pas l'activation immédiate des gènes cibles qui dépend également du contexte [87]. Ils peuvent avoir besoin d'interagir entre eux mais aussi avec d'autres facteurs de transcription liés à la pluripotente tel que ESRRB et NANOG, dont l'induction est plus tardive dans le processus de transition [81].

### **1.2.3 Les facteurs pionniers**

Les facteurs pionniers entraînent des changements épigénétiques restreints à leurs sites de liaisons. Ils permettent la spécification des différentes lignées cellulaires pendant le développement. Ils interagissent souvent avec d'autres facteurs de transcription pour aboutir à la différenciation terminale cellulaire.

Ci-dessous se retrouvent la littérature sur les différents facteurs reconnus comme pionniers qui ont eu un impact dans le développement cellulaire. Une grande partie de ces découvertes ont été répertoriées dans la revue de Drouin et al en 2014 [17].

#### Le destin hépatique

Durant le développement des dérivés de l'endoderme, les facteurs pionniers FOXA1 et FOXA2 ainsi que les facteurs GATA4 et GATA6 sont requis pour le développement du foie et du pancréas [88].

La conversion des fibroblastes embryonnaires en progéniteurs endodermiques nécessite le facteur pionnier FOXA1 et le support du facteur non-pionnier HNF4A. Ils agiraient de manière séquentielle avec l'action de FOXA1 suivie de celle HNF4A [89-92]. Ces conclusions reposent sur la base de la liaison et de l'ouverture de la chromatine après liaison des facteurs concernés. L'état de la chromatine initiale par des marqueurs d'histones post-traductionnelles n'a pas été étudié dans ce contexte. FOXA2 et GATA6, quant à eux sont essentiels pour l'ouverture des enhanceurs dirigeant l'activation des gènes du pancréas [93].

#### Le destin hématopoïétique

La spécification des lignées cellulaires hématopoïétiques est un processus complexe qui implique plusieurs facteurs de transcription, dont certains sont impliqués dans la différenciation des différentes lignées. La spécification séquentielle et la restriction de l'identité cellulaire conduisent à l'établissement de cellules différenciées des lignées myéloïdes et lymphoïdes. PU.1 entraîne les progéniteurs hématopoïétiques vers les lignées cellulaires myéloïdes et macrophages et agit en coopération avec le pionnier C/EBP $\alpha$ 11 [16, 94, 95]. Ces deux pionniers peuvent également reprogrammer les cellules pré-B et les fibroblastes en cellules de type macrophage [96].

Le facteur pionnier EBF1 est responsable de la mise en place des lymphocytes B. Il est nécessaire pour l'expression du facteur de transcription PAX5 qui coopère avec EBF1 dans l'engagement de la lignée et le maintien de l'identité de la cellule B. La combinaison des facteurs de transcription requis pour la différenciation et la maintenance des cellules B comprend également le facteur bHLH E2A et FOXO1 [97-100].

Le pionnier TCF1 ouvre la chromatine au niveau d'enhancers qui sont nécessaires pour établir la lignée de cellules T. Les autres facteurs de transcription impliqués dans l'établissement de la lignée de cellules T avec TCF1 comprennent GATA3 et RUNX [101-104].

#### Le destin mélanotrope

Dans l'hypophyse, le facteur pionnier Pax7 spécifie l'identité des mélanotropes, les cellules sécrétrices de l'hormone mélatonine présent dans le lobe intermédiaire de l'hypophyse. PAX7 permet l'ouverture de la chromatine à quelques milliers d'enhancers qui dirigent le destin des mélanotropes et réprime un sous-ensemble de gènes qui sont responsables de l'identité corticotrope [74, 105]. Tpit est le facteur qui effectue la différenciation terminale des cellules corticotropes et mélanotropes. Tpit n'est pas capable de lier la chromatine fermée mais il joue un rôle essentiel dans l'action pionnière de Pax7. Pax7 permet la mise en place d'un tout nouvel éventail d'enhancers accessibles. Tpit qui, par la suite, peut fixer ces enhancers et activer les gènes responsables de l'identité mélanotrope. De plus, Tpit a été montré nécessaire à l'ouverture de la chromatine au niveau des nouveaux enhancers acquis grâce à Pax7 dans l'hypophyse [106].

#### Le destin cardiaque

L'établissement du destin cardiaque dépend de nombreux facteurs de transcription agissant de façon coordonnée tel que TBX5 and NKX2-5 mais aussi le facteur pionnier GATA4 [107]. ISL1 est aussi requis pour le développement et la fonction cardiaque. Isl1 reconnaît son motif de liaison à l'ADN sur l'ADN enroulé autour des nucléosomes. Il interagit aussi avec BRG1, un composant du complexe SWI-SNF et sa perte entraîne une perte d'ouverture au niveau des régions cibles de ISL1. N'ayant pas l'état de la chromatine et donc son accessibilité avant liaison de ISL1, il est difficile de statuer clairement sur la capacité pionnière de ISL1 [108].

#### Le destin neuronal

La capacité pionnière des facteurs de transcription qui contrôlent la neurogenèse est principalement soutenue par des expériences de gain de fonction, à savoir par leur capacité de reprogrammation. Trois facteurs pionniers ASCL1, NEUROD1 et NEUROD2, ont la capacité d'initier la différenciation neurogénique. ASCL1 peut initier la différenciation neurogénique dans les cellules souches embryonnaires ou dans les fibroblastes. NEUROD1 peut

initier la différenciation neurogénique dans les cellules souches embryonnaires et NEUROD2 peut initier la différenciation neurogénique dans les cellules de carcinomes embryonnaires. Le pionnier ASCL1 est essentiel pour reprogrammer les fibroblastes dans des cellules neuronales, et le facteur de transcription BRN2 (également connu sous le nom de POU3F2) contribue à ce processus [109-111]. La capacité pionnière de NEUROD2 a été corrélée avec des changements dans la méthylation de l'ADN au niveau des enhanceurs neuronaux pendant le développement cérébral normal ; la déméthylase TET2 agit aux enhanceurs ciblés par NEUROD2, suggérant que la déméthylation de ces régions régulatrices est activement initiée par le recrutement de TET2 [112].

### Le destin myogénique

En plus des réseaux de régulation des gènes qui sous-tendent la spécification et l'engagement des cellules du muscle squelettique, des études récentes sur l'organisation et l'architecture de la chromatine ont révélée d'autres acteurs régulant l'expression des gènes myogéniques. PAX3 et le PAX7 sont capables de remodeler l'accessibilité de la chromatine aux loci associés à la fonction myogénique dans les cellules de Schwann périssynaptique (CSP) [113, 114]. Dans le cas de PAX3, l'analyse des cellules mésodermiques dérivées des cellules de souris surexprimant PAX3 a montré que celui-ci augmente l'accessibilité de la chromatine dans les régions contenant ses sites de liaison et que cela se produit en coopération avec SIX4 et un membre de la famille de domaines TEA 2 (TEAD2) pour un engagement myogénique robuste [114]. Chez les souris, la surexpression du PAX7 entraîne une augmentation de l'accessibilité à la chromatine et l'acquisition de modifications d'histones actives chez les progéniteurs musculaires et les myoblastes engagés [113]. Bien qu'un certain nombre de gènes associés à la détermination myogénique, conservent leurs caractéristiques chromatiniennes si l'activité de PAX7 est éliminée, d'autres groupes de gènes ont besoin de la présence continue de PAX7 pour soutenir les modifications histones et l'accessibilité de l'ADN [113]. PAX7 se lie à plusieurs super-enhancers qui contribuent à la différenciation myogénique [115]. En plus des protéines PAX3/7, MYOD peut également réguler l'architecture chromatinienne. Les régions régulatrices des gènes de différenciation myogénique dépendent de l'interaction de MYOD avec l'hélicase à chromodomaine CHD2 pour moduler l'incorporation du variant d'histone H3.3 au niveau des nucléosomes présents dans les régions qui régulent la transcription myogénique [116]. L'incorporation du variant d'histone H3.3 par l'histone chaperonne HIRA dans les régions régulatrices de gènes myogéniques, comme Pax7, a également été démontrée et contribue au maintien de l'identité des cellules myogéniques [117].

### 1.2.4 Domaines protéiques et acteurs associés à l'action pionnière

#### Liaison à l'hétérochromatine et décompaction de l'ADN

Les domaines protéiques impliqués dans la liaison à l'hétérochromatine sont ceux qui ont été le plus étudiés chez les protéines dites pionnières. Ceci peut s'expliquer par le fait que cette étape soit la seule qui peut être attribuer à la protéine pionnière seule. A l'inverse l'ouverture de la chromatine, les modifications post-traductionnelles des histones et la déméthylation de l'ADN requièrent le recrutement d'autres facteurs ayant des capacités de remodelage de la chromatine.

Pour Oct4 et Sox2, il a été montré dans les cellules ES que ces deux protéines coopèrent pour reconnaître un motif composite, provenant de la juxtaposition de leurs sites de liaison canonique individuels [118]. La protéine Sox2 interagit avec l'ADN via son domaine HMG (high-mobility group box) composé de trois hélices alpha reconnaissant spécifiquement un motif de 6 à 7 pb. Les domaines HMG sont uniques car ils interagissent avec le sillon mineur de l'hélice d'ADN et induisent une courbure importante de la molécule d'ADN [119]. La structure par cryo-microscopie électronique des facteurs SOX (SOX2 et son homologue SOX11) a montré leurs liaisons aux nucléosomes à deux tours d'hélice d'ADN suivant l'axe de dyade du nucléosome. Les facteurs Sox se lient ainsi à l'ADN nucléosomal et déforment localement la double hélice. La région liée par Sox2 est plus sensible à des déformations et au détachement des histones menant au réarrangement et à la translocation de l'ADN lorsque l'octamère subit des changements conformationnels [120, 121]. Cependant, l'analyse des sites de liaison de OCT4 et SOX2 in vivo a révélé que ces facteurs lient préférentiellement leurs cibles près des sites d'entrée et de sortie de l'ADN sur le nucléosome [122]. Les structures générées par cryo-microscopie électronique du nucléosome et des protéines OCT4 et SOX2 liés à ces sites ont mis en lumière l'aspect coopératif de la liaison à l'ADN nucléosomale [122]. SOX2 crée une distorsion de l'ADN à ces sites cibles ce qui permet un détachement local des histones et a pour effet d'éloigner l'ADN des histones à l'entrée/sortie du nucléosome. L'interaction allostérique entre les protéines OCT4 et SOX2 est médiée par l'ADN puisque la courbure de la double hélice induite par SOX2 augmenterait l'affinité de liaison à l'ADN de OCT4 [123]. Les facteurs SOX augmentent ainsi l'accessibilité de l'ADN en facilitant le détachement de la région terminale de l'ADN nucléosomal de l'octamère d'histone [124]. Il a aussi été proposé que la liaison d'un facteur SOX au nucléosome, qui permet un repositionnement de la queue N-terminale de l'histone H4, serait incompatible avec l'empilement de nucléosomes lors de la condensation de la chromatine et faciliterait ainsi l'ouverture de la chromatine.

La liaison à la chromatine fermée des facteurs FoxA est médiée par leurs DBDs qui lient leurs sites cible ainsi que par leur domaines C-terminal qui interagissent avec les histones H3 et H4. Cette double interaction est critique pour l'ouverture de la chromatine in vitro [91]. L'analyse de leurs structures cristallines a montré que le DBD des

protéines FoxA se replie en un motif hélice-tour-hélice (HTH) flanqué de chaque côté par des « wings » de chaîne polypeptidique moins ordonné [125, 126]. Ce DBD en hélice « winged » des protéines FoxA présente des similarités avec l'histone H1, ce dernier se liant près de l'axe de dyade du nucléosome [127]. Par conséquent, la liaison de FoxA sur ses cibles dans la chromatine compactée pourrait peut-être déplacer les histones H1 de la chromatine, expliquant partiellement l'augmentation de l'accessibilité à l'ADN [128]. Les régions « wings » pourraient aussi contribuer à la stabilisation de la liaison sur le site cible d'ADN puisqu'elles ont le potentiel d'interagir avec le sillon mineur de la double hélice d'ADN, flanquant les interactions du motif HTH avec le sillon majeur [129]. Pour les protéines FoxA, les domaines « wings » interagissent de manière stable avec leurs sites d'ADN cibles de haute affinité, mais moins au niveau de leurs sites d'ADN cibles d'affinité inférieure, où seul le motif HTH semble impliqué [129]. Le facteur FoxA1 reconnaît des motifs de séquences dégénérés dans le contexte du nucléosome. La liaison de FoxA1 au nucléosome se fait donc grâce à des motifs de faible affinité. La liaison de FoxA1 à des motifs optimaux et de haute affinité serait donc plutôt opportuniste et non nucléosome spécifique [130]. Suite à ces études, il a été proposé que la mobilité de FoxA1 dans un contexte de chromatine fermée reposerait sur des interactions spécifiques et non spécifiques avec la chromatine [131, 132].

Pelletier et al, ont montré que les domaines de liaison à l'ADN PD et HD de la protéine PAX7 sont nécessaires à la liaison à ses sites pionniers. L'étude a montré que ces domaines de liaisons peuvent lier leurs séquences cibles dans l'ADN méthylé et que les sites génomiques non accessibles dit pionniers présentent en général plus de trois ou quatre motifs reconnus par Pax7 comparés aux sites génomiques accessibles dit transcriptionnels qui n'en comptabilise en général qu'un ou deux [133].

Ceci montre qu'il existe divers modes de liaison à l'ADN nucléosomal par les facteurs pionniers. Cependant, ce qui est commun à tous les facteurs pionniers connus et compatible avec la liaison nucléosomique, est la reconnaissance de l'ADN via une hélice alpha courte d'un module HTH (Oct4, Sox2, FoxA, Pax7), à doigt de zinc (Klf4) ou de type « scissor-like » (Ascl1) [134]. Ces propriétés de liaison à l'ADN peuvent être importantes pour la capacité des facteurs à reconnaître des séquences d'ADN spécifiques dans la chromatine nucléosomique condensée.

#### Les modifications post-traductionnelles d'histone et ouverture de la chromatine.

Il a été montré pour plusieurs facteurs pionniers que leur liaison était suivie par une augmentation de l'accessibilité de l'ADN déterminée par les méthodes de DNase1, MNase ou ATAC. Cette ouverture est suivie par l'apparition ou l'augmentation des marques d'enhancers actifs H3K4me1 et H3K27ac avec un profil bimodal.

Plusieurs exemples dans la littérature ont montré que les facteurs pionniers pouvaient recruter des remodeleurs de chromatine tel que BRG1 et des modificateurs d'histone tels que les protéines MLL, LSD1 et ASH2 [Gouhier et al, 2022, en préparation]. Cependant, les domaines spécifiques des protéines pionnières responsables du recrutement des enzymes nécessaires au remodelage chromatinien reste à ce jour peu étudiés.

Parmi les différents travaux traitant ce sujet dans la littérature, Takaku et al ont montré que la protéine GATA3 délétée de son domaine de transactivation (TAD) permet la liaison de cette protéine aux sites d'hétérochromatine mais échoue à ouvrir la chromatine. Ils ont montré que la présence de ce domaine TAD était importante pour le recrutement de BRG1, un membre du complexe de remodelage SWI-SNF [135].

Pour EBF1, Boller et al ont montré que le domaine C-terminal de EBF1 (CTD) était superflu pour la liaison initiale de la protéine EBF1 mais est requis pour l'ouverture de la chromatine et le recrutement de BRG1 [99, 136]. Ils montrent aussi qu'un domaine prion like présent dans le CTD médie l'interaction de EBF1 avec des protéines de liaison de l'ARN de la famille FET et la protéine Brg1. Ce domaine permet également à EBF1 d'accomplir des séparations de phase et la formation de condensats dans le noyau [136].

Enfin, Rudnický et al ont identifié l'arginine méthyltransférase Carm1 comme une protéine interagissant avec Pax7 et ont découvert que Carm1 méthyle spécifiquement plusieurs arginines dans l'extrémité N-terminale de Pax7 où se trouve le domaine PAIRED. Pax7 méthylé lie directement les protéines du trithorax MLL1/2, ce qui entraîne le recrutement du complexe ASH2L:MLL1/2:WDR5:RBBP5 (histone H3K4 méthyltransférase) au niveau de ces enhancers et du promoteur proximal Myf5 [137]. Cependant ce mécanisme n'a pas été démontré comme nécessaire au mécanisme pionnier.

#### La méthylation de l'ADN

Vanzan et al, ont démontré récemment que la plupart des facteurs pionniers induisait la déméthylation de l'ADN via les enzymes TET, c'est le cas notamment de FOXA1 [138] et KLF4 [139, 140]. Cependant, pour certains dont Sox2, la déméthylation de l'ADN se fait de manière passive via DNMT1. Harris et al (en préparation) démontrent ce mécanisme de manière convaincante pour Pax7. Pax7, via son domaine PAI (sous partie du domaine de liaison PAIRED de Pax7) interagit le domaine de la protéine UHRF1 interagissant avec la protéine DNMT1. La liaison de Pax7 avec UHRF1 empêche l'interaction entre UHRF1 et DNMT1. L'interaction de UHRF1 et DNMT1 est essentielle pour la maintenance de la méthylation de l'ADN lors de la réplication cellulaire. Pax7, au niveau de son site de liaison, entrainerait une déméthylation passive de l'ADN en empêchant le complexe UHRF1 et DNMT1 d'accomplir son activité de maintenance de la méthylation.

## 1.3 Pax7 : un facteur pionnier dans l'hypophyse

### 1.3.1 Hypophyse et différenciation cellulaire

#### Développement de l'hypophyse

L'hypophyse ou glande pituitaire est située à la base du cerveau et est constituée de différentes cellules sécrétrices d'hormones. Les hormones hypophysaires agissent sur plusieurs organes dont principalement les glandes endocrines pour réguler des fonctions corporelles telles que la croissance, le métabolisme et la fertilité [141, 142].

L'hypophyse est divisée en trois lobes : le lobe antérieur, le lobe intermédiaire et le lobe postérieur. Si le lobe intermédiaire est bien développé chez certains animaux tel que les rongeurs et les chiens, l'humain voit cette structure seulement au stade fœtal, puis celle-ci régresse à partir de la quinzième semaine de gestation [143].

Chez l'embryon murin, le lobe antérieur se développe par une invagination épithéliale du plafond de l'ectoderme orale, tandis que le lobe postérieur provient d'une extension ventrale du diencephale [144]. L'invagination de l'ectoderme oral est aussi appelée poche de Rathke et son développement requiert des contacts soutenus avec l'épithélium neural (Figure 1.5).

L'hypophyse est reliée à l'hypothalamus par la tige hypophysaire composée principalement de vaisseaux sanguins et de prolongations de cellules nerveuses. La fonction des lobes antérieur et intermédiaire est modulée par des hormones passant par les vaisseaux sanguins qui proviennent de l'hypothalamus, tandis que le lobe postérieur est sujet à ses influx nerveux.

Plusieurs facteurs sont critiques pour le développement de l'hypophyse. Des étapes fondamentales à la formation de la poche de Rathke ont été identifiées chez la souris. Notamment, l'expression de Bmp4 (Bone morphogenetic protein 4) dans la partie ventrale du diencephale au jour embryonnaire 9,5 (e9,5), suivie par l'expression de Fgf8 (Fibroblast growth factor 8) au jour e10 [145]. Le facteur Shh (Sonic hedgehog) est aussi essentiel au développement de l'hypophyse et il est exprimé dans tout l'ectoderme oral, excluant la poche de Rathke. Cette dernière voit l'expression transitoire de Hesx1 entre e9,5 et e11,5 pendant que la structure forme rapidement une glande fermée par la séparation du lien avec l'ectoderme oral. Très tôt dans la formation de la poche de Rathke, la structure est marquée par l'expression des TFs Pitx1 et Pitx2 (Pituitary homeobox 1 et 2), puis des facteurs à homeodomaine Lim-homeo Lhx3, Lhx4 et Isl1 (Islet1). Entre les jours e12 et e15, l'hypophyse subit une expansion importante dû à la prolifération cellulaire; toutefois l'expression de marqueurs de différenciation cellulaire est plus tardive. La différenciation des cellules hypophysaires est progressive et séquentielle entre les jours e11,5 et e16,5. C'est seulement au jour e17,5 que toutes les cellules spécialisées de l'hypophyse expriment leurs hormones respectives, et acquièrent leurs identités cellulaires [146] (Figure 1.5).

Le contrôle de la prolifération cellulaire est critique lors de l'organogenèse afin de déterminer la taille des organes et la composition cellulaire des tissus. Les cellules progénitrices de l'hypophyse, indifférenciées et marquées par l'expression de Sox2, prolifèrent activement autour de la lumière situé entre le lobe antérieur et intermédiaire [147]. La prolifération des cellules progénitrices diminue au fur et à mesure que l'organogenèse progresse de telle sorte que, chez l'adulte, les cellules Sox2-positives quiescentes se trouvent principalement au niveau de la fente hypophysaire [148, 149]. Cette fente qui sépare le lobe intermédiaire du lobe antérieur constitue le reste de la lumière d'origine du primordium hypophysaire. Les cellules Sox2-positives ont la capacité de former des « pituisphères » en culture, qui peuvent se développer et former des sphères secondaires [148, 150]. Les pituisphères peuvent être induites pour se différencier vers tous les types de cellules hypophysaires.

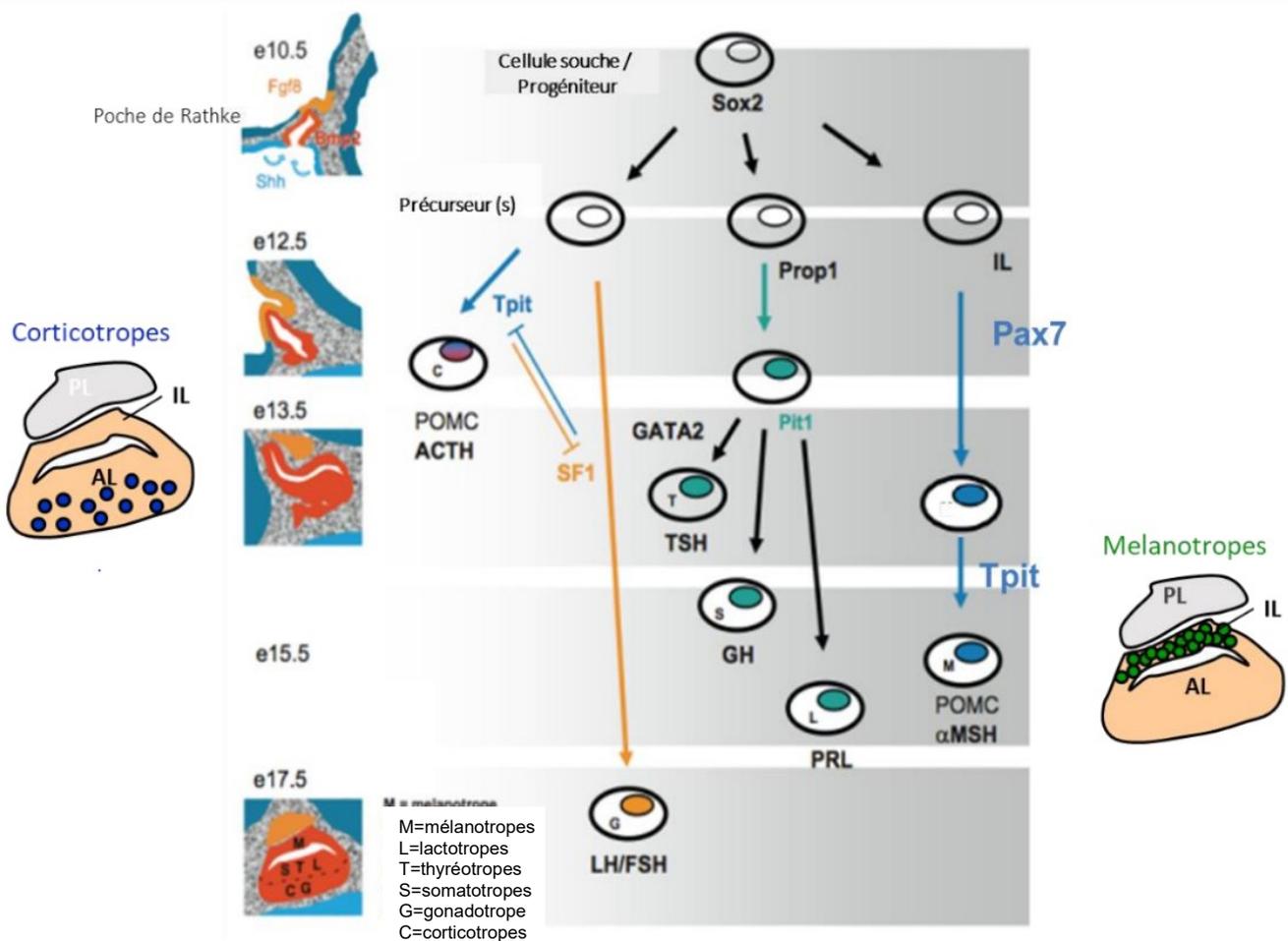


Figure 1.5 La différenciation cellulaire des lignées endocrines de l'hypophyse

Adapté de Drouin, 2022 [142]

### Cellules hypophysaires et facteurs de différenciation terminale

Les gènes codants pour les hormones caractéristiques de chaque lignée ont servi à identifier les TFs qui sont impliqués dans l'identité des cellules hypophysaires. Par conséquent, ce sont surtout les facteurs responsables de la différenciation terminale qui sont les mieux connus; ceux-ci sont généralement exprimés de 12 à 24 heures avant le gène de l'hormone lui-même [142].

Le lobe postérieur, ou neurohypophyse, est composé de cellules nerveuses reliées à l'hypothalamus. Ces cellules ne fabriquent pas d'hormone mais stockent les hormones antidiurétique (ADH) et ocytocines produites par les cellules neurosécrétrices de l'hypothalamus. Ces hormones sont nécessaires à la réabsorption de l'eau par les reins, à la stimulation de la contraction de l'utérus lors de l'accouchement et au réflexe de sécrétion du lait par la glande mammaire. Ces hormones sont relâchées de la neurohypophyse suite à des influx nerveux en provenance de l'hypothalamus.

Le lobe antérieur est composé de cinq types de cellules endocrines : les somatotropes, les lactotropes, les thyrotropes, les gonadotropes et les corticotropes (Figure1.5).

#### Le facteur de différenciation terminale Pit1

Le facteur de transcription Pit1 (POU homeodomain protein) activé par Prop1 (prophet of Pit1), observable au jour e12,5, est essentiel pour la différenciation des lignées thyrotropes, lactotropes et somatotropes [151-155].

Les somatotropes sont les cellules les plus abondantes de l'hypophyse et constitue environ 50% du lobe antérieur adulte chez le mâle. Elles produisent l'hormone de croissance (GH) nécessaire à la croissance des os, des muscles et le bon fonctionnement du métabolisme.

Les lactotropes constituent 15 à 20 % du lobe antérieur chez le male et près de 50% chez la femelle. Elles produisent la prolactine (PRL) nécessaire à la stimulation du lait dans la glande mammaire.

Les thyrotropes représentent 1-5% des cellules du lobe antérieur et produisent la thyrostimuline (TSH) qui stimule la formation d'hormones thyroïdes dans la glande thyroïdienne.

#### Le facteur de différenciation terminale SF1

Le facteur de différenciation terminale SF1 (Steroidogenic Factor 1) détectable au jour e13,5 [156, 157] est essentiel pour la différenciation des gonadotropes. Au jour e16.5, les gonadotropes agissent sur les ovaires et les testicules en produisant l'hormone folliculo-stimulante (FSH) qui stimule la production d'ovules et de spermatozoïdes et l'hormone lutéinisante (LH) qui stimule l'ovulation et production de testostérone.

Un antagonisme existe entre SF1 et le facteur de différenciation terminale des corticotropes et mélanotropes Tpit. Tpit est un régulateur négatif du destin gonadotrope et SF1 est un régulateur négatif du destin corticotrope. Cela suggère que ces lignées pourraient provenir de précurseurs communs [158, 159]. C'est en partie le résultat d'un mécanisme de transrépression entre les deux TFs spécifiques à leur lignée qui s'exerce sur leurs cibles génétiques respectives [160].

#### Le facteur de différenciation terminal Tpit

Le facteur de différenciation terminal Tpit (Pituitary Tbox factor) est essentiel pour les lignées cellulaires corticotropes (lobe antérieur) et mélanotropes (lobe intermédiaire). Ces types cellulaires dont les hormones proviennent du même précurseur POMC (proopiomélanocortine) [161] sont au cœur de mon projet de thèse.

Tpit est un TF à boîte T qui a d'abord été identifié comme un partenaire transcriptionnel du facteur Pitx1 agissant sur le promoteur du gène POMC. La différenciation terminale des deux lignées POMC est menée par Tpit. Les souris déficientes en Tpit ne parviennent pas à exprimer POMC et la différenciation des mélanotropes et des corticotropes n'aboutit pas [160].

Dans le lobe antérieur, les corticotropes expriment Tpit à partir du jour e11,5 et celui-ci est critique pour la différenciation terminale et la maintenance des corticotropes [160]. Les corticotropes clivent le précurseur POMC en hormone adrénocorticotrope (ACTH) grâce à la proprotéine-convertase 1 (PC1, codée par le gène Pcsk1) [162] (Figure 1.6). L'ACTH stimule la production de glucocorticoïdes par les glandes surrénales. Les glucocorticoïdes interagissent avec le récepteur cellulaire aux glucocorticoïdes (GR), exprimé dans de nombreux organes, afin de réguler notamment le métabolisme énergétique. Parmi les gènes marqueurs des corticotropes sont le récepteur des glucocorticoïdes (GR), l'hormone de libération de la corticotropine (corticotropin-releasing-hormone receptor 1 - Crhr1) et le récepteur de la vasopressine (vasopressin receptor 1B - Avpr1b) [163].

Dans le lobe intermédiaire, les mélanotropes clivent le précurseur POMC en hormone mélanocortines ( $\alpha$ MSH) grâce à la proprotéine-convertase 2 (PC2, codée par le gène PCSK2) [162] (Figure 1.6). Chez les mammifères, l' $\alpha$ MSH favorise la pigmentation de la peau, des cheveux et des muqueuses. Chez l'homme, où le lobe intermédiaire est absent, la pigmentation est régulée localement au niveau de la peau en réponse à l'exposition aux UV et à la production locale d' $\alpha$ MSH.

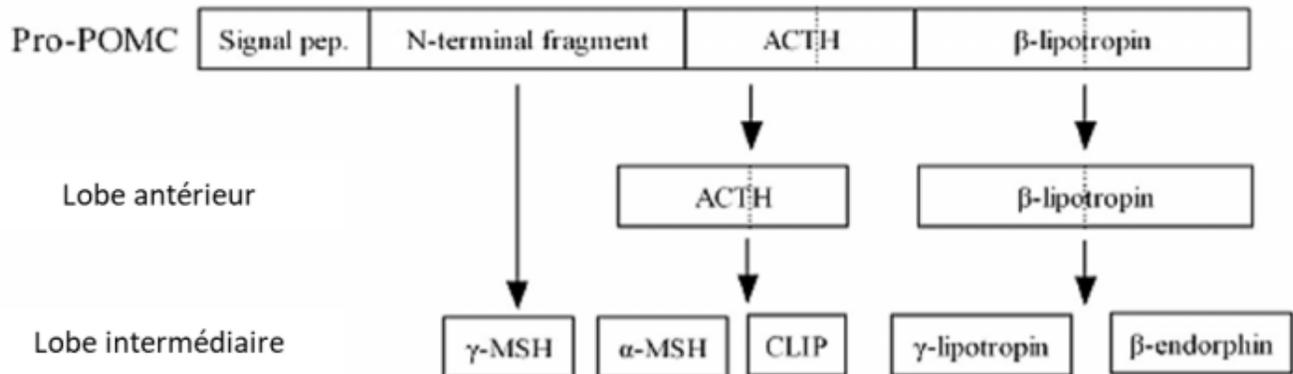


Figure 1.6 Clivage différentiel du précurseur hormonal POMC

Adapté de Wu, 2012 [164]

Durant le développement de l'hypophyse, l'expression de l'hormone POMC dans le lobe intermédiaire commence vers le jour e15,5 et est précédée par l'expression de Tpit [161]. En l'absence de Tpit, environ 10% des cellules du lobe intermédiaire changent de destin et deviennent gonadotropes. Elles sont SF1-positives et n'expriment aucun marqueur mélanotrope en accord avec l'antagonisme observé entre Tpit et SF1 [165]. D'autres cellules du lobe intermédiaire restent bloquées dans un état non prolifératif et non différencié exprimant p57 (du gène CDKN1C codant pour l'inhibiteur de kinase dépendant de la cycline 1C). p57 est normalement exprimé par des cellules progénitrices qui viennent de sortir de prolifération et qui n'expriment pas encore de marqueur connu [166].

Parmi les marqueurs des cellules mélanotropes, nous avons le gène PCSK2, le gène récepteur à la dopamine D2 (DRD2), l'absence du récepteur des glucocorticoïdes (GR), le gène HES1 un transducteur de la voie de signalisation Notch et le gène PAX7. Notre laboratoire a montré que Pax7 est un facteur pionnier qui spécifie le destin mélanotrope avec l'aide du facteur non-pionnier Tpit [105, 106, 167].

### 1.3.2 Le facteur pionnier Pax7

Les cellules mélanotropes sont marquées par l'expression de Pax7, un TF de type paired box [105]. Pax7 est essentiel pour l'activation des gènes spécifiques des mélanotropes et la répression des gènes corticotropes. Ainsi, Pax7 est un gène sélecteur du lobe intermédiaire et détermine le destin mélanotrope dans l'une des deux lignées POMC dépendantes de Tpit pour la différenciation terminale. Pax7 est exprimé dans le lobe intermédiaire juste avant Tpit et son expression chevauche de façon transitoire celle du marqueur de cellules souches hypophysaires Sox2. Pax7 amène les progéniteurs hypophysaires à se différencier notamment grâce à son action pionnière. Pax7

cible les enhancers spécifiques aux mélanotropes présents dans l'hétérochromatine et augmente leurs accessibilités permettant le recrutement d'autres facteurs de transcription tel que Tpit [167]. En absence de Pax7 souris, les cellules du lobe intermédiaire changent d'identité vers un destin corticotrope. Toutefois, Pax7 n'est pas suffisant pour la différenciation des mélanotropes car celle-ci est menée par Tpit. En ce sens, les cellules Pax7-positives du lobe intermédiaire de souris déficientes pour Tpit sont engagées dans la voie de différenciation, mais nécessitent Tpit pour terminer le processus [105]. Tpit exerce un effet positif sur l'expression de Pax7 et la coopération de ces deux facteurs permet de mettre en place le programme d'expression génique spécifique aux mélanotropes [106]. Dans les cellules AtT-20, une lignée cellulaire dérivée de corticotropes, l'expression ectopique de Pax7 conduit à un changement de l'identité cellulaire vers une signature génétique mélanotrope [105]. Pax7 peut reprogrammer ces cellules pour exprimer des marqueurs mélanotropes et ce changement d'identité implique un remodelage de la chromatine. Suite à l'arrivée de Pax7, les régions régulatrices mélanotropes spécifiques perdent les marques répressives H3K9me2 [Gouhier et al,2022, en préparation] et arborent une décompaction de la chromatine avec des marques d'histones d'enhancers actifs telles que les marques bimodales de H3K4me1 et H3K27ac [167]. L'action pionnière de Pax7 aux enhancers mélanotrope-spécifiques donne lieu à de nombreux nouveaux sites de liaison pour le TF Tpit, notamment aux loci Pcsk2 et DRD2 caractéristiques des mélanotropes. Ces nouveaux sites Tpit contiennent des motifs de liaison à l'ADN canoniques, TpitREpal, qui ne peuvent pas être distingués des autres sites de liaison de Tpit actifs dans les cellules AtT-20 [105]. L'action pionnière de Pax7 a un effet majeur sur l'accessibilité au niveau d'un répertoire d'enhancers spécifiques aux mélanotropes et crée donc un environnement de chromatine permissif pour la liaison de Tpit [167](Figure 1.7).

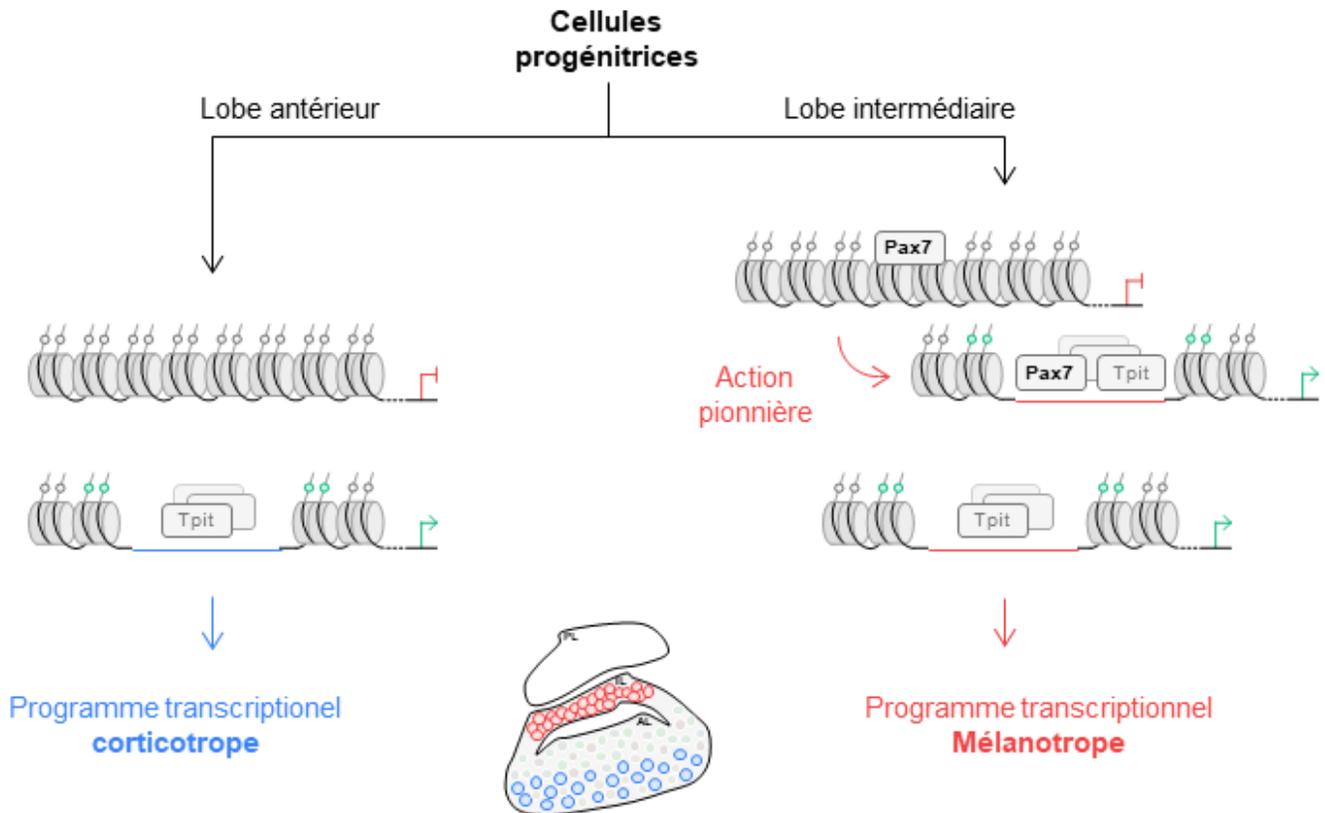


Figure 1.7 Action de Pax7 dans l'hypophyse

Créé par Arthur Gouhier et inspiré par Budry et al. 2012 [105].

### 1.3.3 Pax7 : un membre de la famille Pax

Pax7 fait partie de la famille de gènes Pax qui ont plusieurs rôles importants dans le développement et l'organogenèse.

Les gènes de la famille paired (PAX) codent une famille de facteurs de transcription très conservés trouvés chez les vertébrés et les invertébrés. Les gènes PAX codent pour une famille de facteurs de transcription qui possèdent un domaine de liaison à l'ADN paired (PD) contenant 128 acides aminés conservés [168]. Le PD est constitué de deux sous-domaines à motif hélix-tour-hélix (HTH), PAI et RED [169-174] qui interagisse avec le sillon majeur de l'ADN [175]. La capacité du PAI et du RED à interagir avec l'ADN entraîne la nature bipartite de la liaison protéine-ADN des Pax [171]. Des modifications post-traductionnelle du PD peuvent nuire à sa capacité de lier l'ADN par l'intermédiaire du sous-domaine PAI [176, 177]. Souvent appelées « régulateurs maîtres », les protéines PAX orchestrent le développement des tissus et des organes lors de la différenciation cellulaire. Elles sont essentielles à la structure et au fonctionnement des tissus par le maintien de l'identité cellulaire. Les mutations dans les gènes PAX sont associées à une myriade de maladies humaines (p. ex., microphthalmie, anophtalmie, colobome,

hypothyroïdie, leucémie lymphoblastique aiguë) (Figure 1.8). La régulation transcriptionnelle par les protéines PAX est, en partie, modulée par l'épissage différentiel de leur transcrits [178].

Les mammifères possèdent neuf protéines PAX, qui sont divisées en quatre groupes. La présence d'un lien octapeptidique (OP) et/ou d'un homéodomaine partiel ou complet (HD) détermine la classification des sous-groupe [168].

Les protéines PAX des groupes II, III et IV contiennent un domaine de liaison ADN HD partiel ou complet, qui leur permet de lier et de réguler des sous-ensembles uniques de cibles génétiques. Les protéines PAX des groupes I, II et III contiennent également un domaine ou un dérivé du domaine octapeptide pouvant être lié par des cofacteurs qui régulent la transactivation médiée par les protéines PAX des groupes II et III [179-182]. Les protéines PAX du groupe I n'ont pas de HD, ce qui indique qu'un HD n'est pas nécessaire pour leur fonction de liaison à l'ADN. Il a été montré que le HD influence la liaison protéine-ADN en interagissant avec le PD [174] (Figure 1.8). Dans l'hypophyse, Pelletier et al ont montré que les domaines PD et HD sont nécessaires à l'action pionnière de Pax7 notamment pour sa liaison de Pax7 à la chromatine fermée [133].

Sous-groupe	Gène Pax	Localisation sur le chromosome		Structure				Expression pendant le développement	Syndromes /maladies associés aux gènes PAX/Pax	
		Human	Mouse	PD	OP	HD	TD		Syndromes /maladies humaines	Souris knock-out
I	1	20p-11	2	+	+	-	+	Sclérome, thymus, Squelette	Klippel — syndrome de Feil, syndrome de Jarcho-Levin, tumeur de la glande salivaire	Squelettogénèse perturbée
	9	14q 12-13	12	+	+	-	+	Sclérome, squelette, cranio-facial, dents, thymus	Pas de thymus, pas de glandes parathyroïdes, pas de dents, malformations craniofaciales et des membres	Oligodontie, carcinome oesophagien, syndrome de Jarcho-Levin
	2	10q 25	19	+	+	Tronqué	+	Système nerveux central, rein, oeil, oreille, glande mammaire	Syndrome du colobome rénal (syndrome papillorénal), carcinome rénal, tumeur de Wilms, cancer du sein, sarcome de Kaposi	Syndrome du colobome rénal
II	5	9p 13	4	+	+	Tronqué	+	Système nerveux central, lymphoïde B, testicules	Lymphome à grandes cellules, leucémie lymphocytaire, médulloblastome, neuroblastome, astrocytome	Pas de cellules B, défauts cérébraux
	8	2q 12-14	2	+	+	Tronqué	+	Système nerveux central, rein, thyroïde	Dysplasie thyroïdienne, carcinome folliculaire thyroïdien, tumeur de Wilms, cancer du placenta, tumeurs séreuses ovariennes	Hypothyroïdie, malformation de la crête neurale
III	3	2q 35	1	+	+	Complet	+	Système nerveux central, crête neurale, muscle	Syndrome de Waardenburg, RMS, sarcome d'Ewing	Sp, Spr, Spd, Sp1H, Sp2H, Sp4H
	7	1p 36.2	4	+	+	Complet	+	Système nerveux central, crête neurale, muscle	rhabdomyosarcome, sarcome d'Ewing, mélanome, carcinome pulmonaire épidermoïde	Anomalie de la crête neurale
	4	7q 32	6	+	-	Complet	+	Système nerveux central, pancréas, hypophyse	Syndrome de Silver-Russell, syndrome de Wolcott-Rallison, diabète, insulinoïme	Pas de cellules β et δ du pancréas
IV	6	11 p13	2	+	-	Complet	+	Système nerveux central, yeux, nez	Aniridie, cataracte, glioblastome multiforme, glioblastome anaplasique, gliome astrocytaire	Petit oeil, pas de cellules α pancréatiques, malformations cérébrales

Figure 1.8 Catégorisation des protéines de la famille Pax

Tiré de Wang et al, 2008 [183]

Bien que toutes les protéines PAX aient une partie C-terminale avec un domaine de transactivation riche en sérine, proline, thréonine (TD), Pax7 est unique dans la famille pour la présence d'une séquence d'acides aminés en C-terminal, l'OAR (Otp/aristaless /Rax), qui se trouve dans plus de trente protéines HD, mais dont la fonction est encore mal définie.

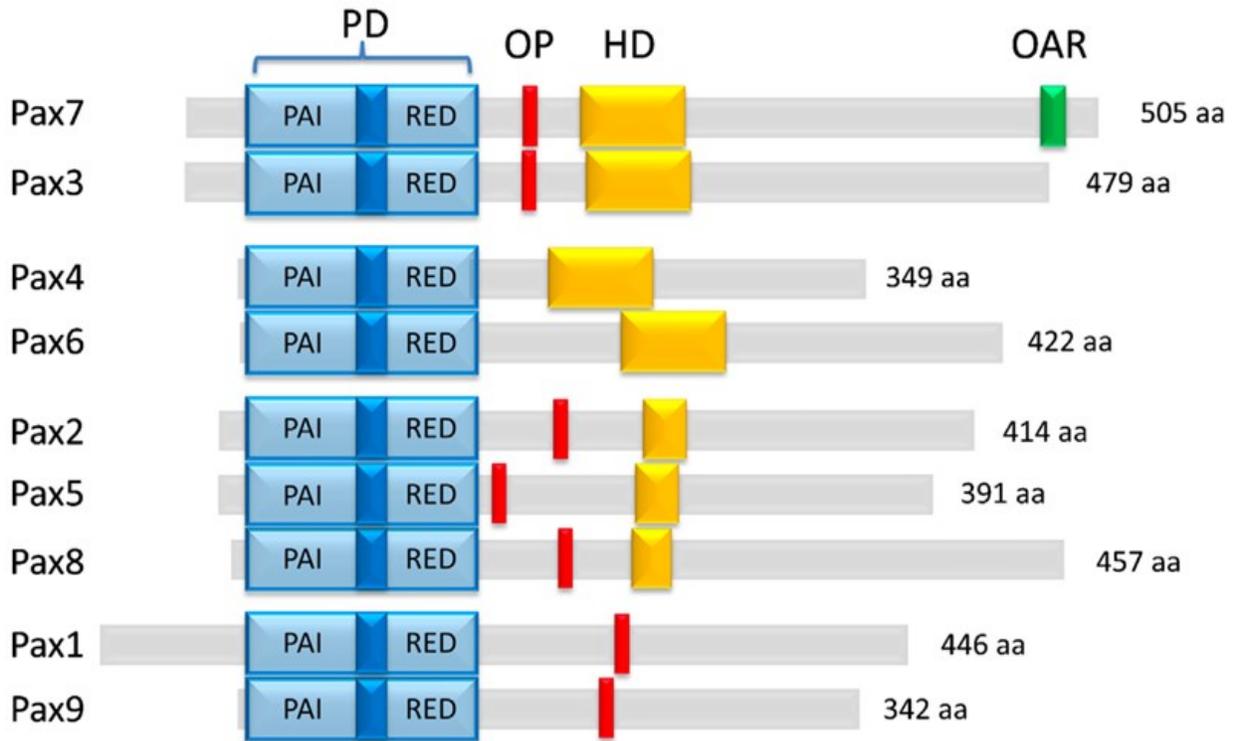


Figure 1.9 Structure des facteurs de transcription Pax

tiré de Mayran et Pelletier 2015 [184]

Les protéines Pax peuvent interagir avec des remodeleurs de chromatine (voir tableau ci-dessous) mais à part Pax7, l'action pionnière n'a été clairement démontrée chez aucune des autres protéines de la famille Pax. Cela pourrait être expliqué par le manque de modèle pratique pour étudier les propriétés pionnières de ces facteurs.

Facteurs Pax	HD	OP	Co-Activateurs	Co-Represseurs	Facteurs de remodelage de la chromatine
2	1/2	+	PTIP	Grg4 PPM1B	MII3/4, Prmt5
5	1/2	+	PTIP, Daxx, CBP	Grg4, NCor1 Daxx	Brg1
8	1/2	+	Rb		
3	+	+	Pax3/7BP Taz, Tap65 HDAC10	HIRA ,Daxx KAP1, HP1	Tif1b HP1
7	+	+	Pax3/7BP Carm1 (Prmt4)		H3K4/HMT MII1/2:Wdr5
4	+	-			
6	+	-	Tip60		Brg1 Snf2h

Figure 1.10 Protéines Co-régulatrices interagissant avec les protéines Pax

tiré de Mayran et Pelletier 2015 [184]

### 1.3.4 Chronologie de l'action pionnière de Pax7

Pax7 est un facteur de transcription et ne contient aucun domaine enzymatique permettant le remodelage de la chromatine. Il est donc difficile d'imaginer qu'il entraîne seul tous les changements épigénétiques cités plus haut. La coopération avec d'autres facteurs de transcription ou enzymes de remodelage chromatinien lui sont donc nécessaire pour aboutir au processus pionnier [72].

En 2018, Mayran et al montrent avec un système de Pax7 inducible dans les cellules corticotropes AtT20 une première chronologie de l'action pionnière [167]. Ils établissent que la liaison de Pax7 aux sites pionniers prend 30 minutes mais n'est maximale et stable qu'à partir de 12 heures. L'ouverture de la chromatine mesurée par ATACseq prend 24 à 48 heures et l'activation des gènes sous le contrôle d'enhancers pionniers prennent 3 à 5 jours pour être complètement activés. A l'inverse, l'activation des enhancers accessibles dit « transcriptionnels » ou « constitutifs » de Pax7 est beaucoup plus rapide, 30 minutes pour une liaison stable, une décompaction de la

chromatine déjà présente au départ et une activation des gènes sous le contrôle de ces enhancers transcriptionnels en moins de 24 heures.

En 2022, en utilisant le même modèle de travail, Gouhier et al montrent que l'action pionnière se déroule en deux phases dont la plus tardive dépend de la réplication cellulaire.

Dans la première phase indépendante de la réplication cellulaire, Pax7 enclenche progressivement une perturbation de l'état chromatinien en parallèle avec la stabilisation de son recrutement (12 heures). Cette perturbation est marquée par une augmentation marginale de l'accessibilité de l'ADN et l'apparition d'un faible pic monomodale de H3K4me1 ainsi que la perte de l'histone H1 (marque de condensation de la chromatine) et de la marque répressive H3K9me2.

Pendant cette phase, Pax7 recrute différents facteurs de transcriptions et remodelateurs de la chromatine, tel que Tpit, Ash2 et Lsd1.

Dans la deuxième phase ( $t_{1/2}=24h$ ) dépendante de la réplication cellulaire, PAX7 entraîne l'ouverture de la chromatine marquée par la perte d'occupation du nucléosome mesurée par l'histone H3 et le gain d'accessibilité de l'ADN, la transition du profil de l'histone H3K4me1 de monomodal à bimodal, l'ajout de la modification d'histone H3K27ac et la perte de la marque Lamin B1. Durant cette phase, Pax7 permet le recrutement de BRG1 (ATPase du complexe de remodelage SWI/SNF) favorisant l'ouverture de la chromatine et le recrutement de l'enzyme P300 permettant le dépôt de la modification H3K27ac.

Gouhier explique le retard de cette deuxième phase dit d'ouverture de la chromatine par le positionnement des enhancers pionniers dans le noyau qui serait différent lors de la première et de la deuxième phase de l'action pionnière.

Pendant la première phase, les enhancers pionniers de Pax7 se retrouvent au niveau de la périphérie du noyau, et sont donc inaccessible aux facteurs de remodelage de la chromatine présent dans le centre du noyau. Pendant la réplication, ces enhancers pionniers se détachent de la périphérie du noyau et n'y retourneront qu'après la réplication grâce à la marque H3K9me2 qui s'associe avec la LaminB1 présente à la périphérie nucléaire.

Les sites pionniers liés par Pax7 après la réplication ne présentent plus H3K9me2 et donc ne peuvent plus retourner à la périphérie du noyau. Ils deviennent donc accessibles et sensibles à l'action des facteurs de remodelage tel que BRG1 ou P300 (Figure 1.11).

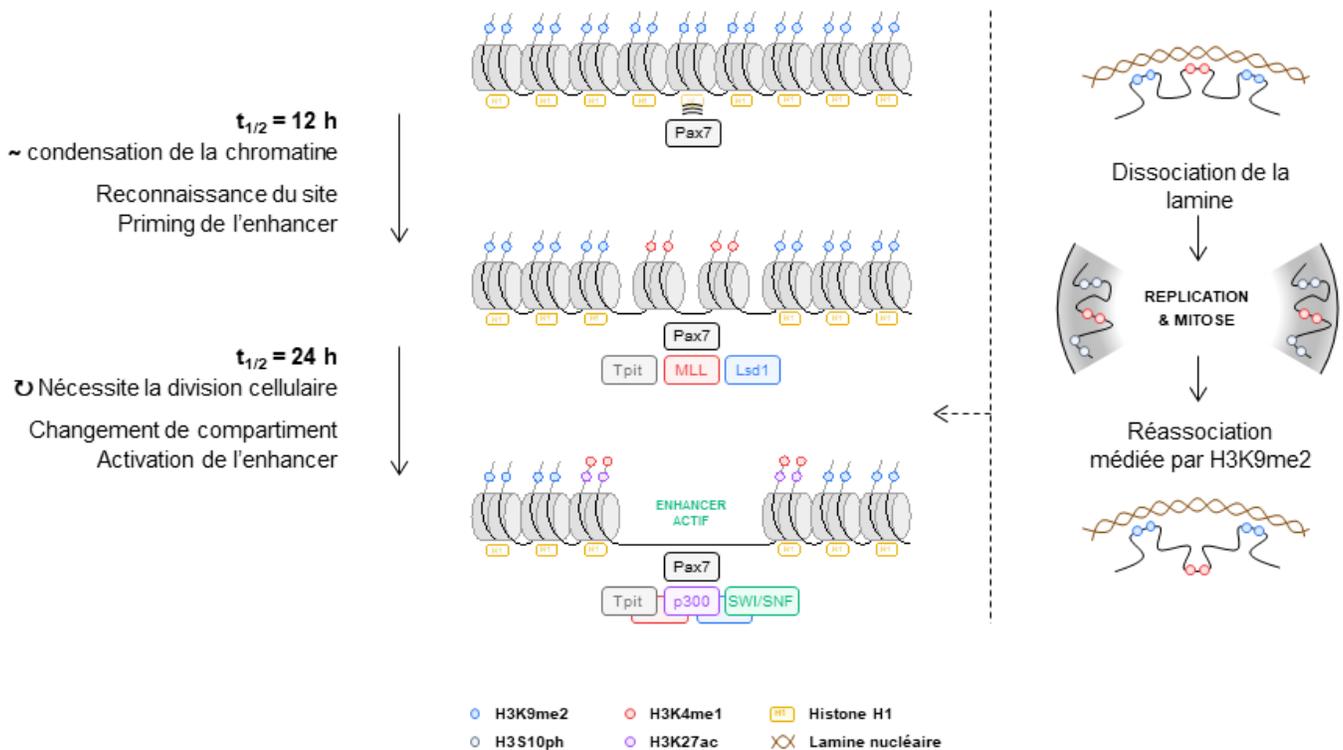


Figure 1.11 Chronologie de l'action pionnière de Pax7

Tiré de Gouhier et al, 2022 (en préparation)

Il est à noter que Pax7 partage des traits communs d'actions avec d'autres facteurs pionniers. Par exemple, Oct4 et GATA3 ont montré, eux aussi, avoir besoin du complexe SWI-SNF pour ouvrir la chromatine [135, 185]. Cependant, le mécanisme pionnier de Pax7 présente aussi des différences avec d'autres facteurs pionniers. FoxA2, par exemple, n'as pas besoin de la réplication pour entrainer l'ouverture de la chromatine contrairement à Pax7[186].

### 1.3.5 Les isoformes de Pax7

Différents isoformes de Pax7 ont été identifiées dans divers tissus murins et humains. L'épissage de transcrits alternatifs se produit soit à la jonction de l'intron 2 et de l'exon 3 (trinuécléotide) ou à la jonction de l'intron 3 et de l'exon 4 (hexanucléotide). Les quatre isoformes en résultant diffèrent par l'inclusion ou l'exclusion de résidus glutamine (Q) et/ou glycine et leucine (GL) dans le domaine PD. Plus précisément, le résidu glutamine est situé dans la région linker des sous-domaines du PD, à proximité du sous-domaine RED, et le dipeptide GL est situé en N-terminal de la troisième hélice alpha du RED (Figure 1.12). Ces isoformes appelées Pax7 Q-GL+, Pax7 Q+GL+,

Pax7 Q-GL- et Pax7 Q+GL- ont des caractéristiques structurales différentes qui pourraient altérer la liaison à l'ADN.

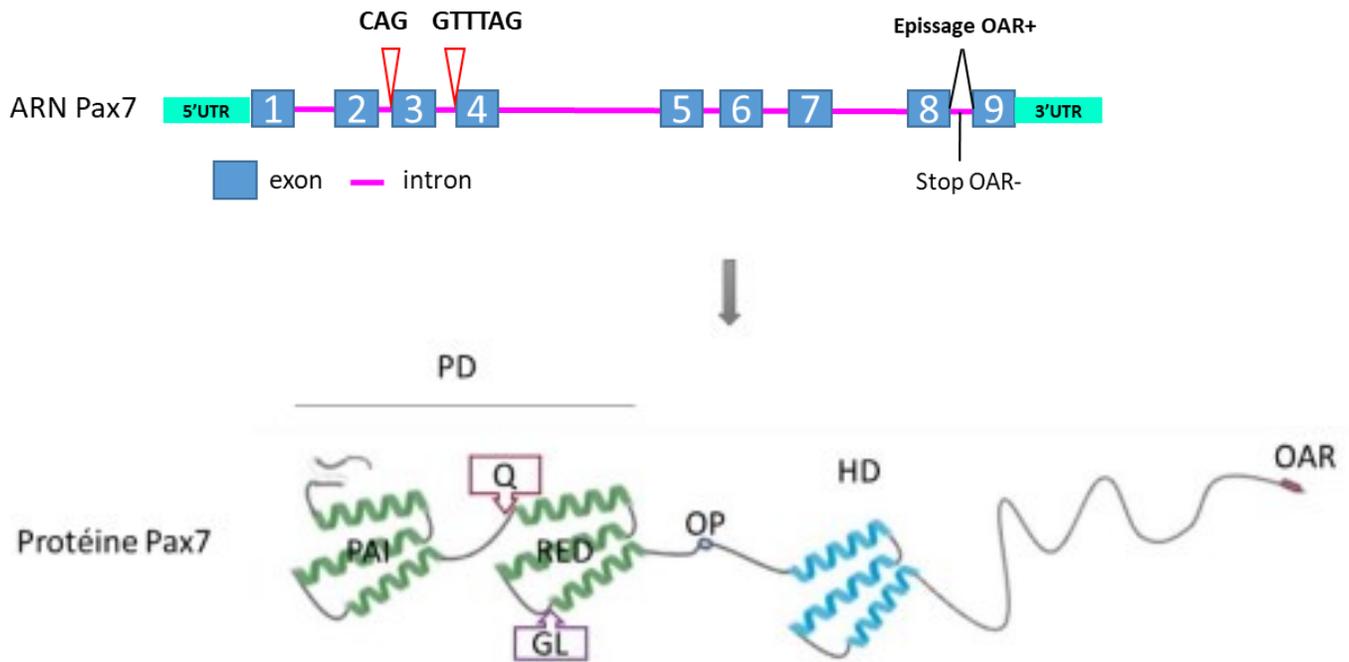


Figure 1.12 Épissage alternatif du transcrit de Pax7 et isoformes

Représentation schématique créé à l'aide de BioRender par Audrey Pelletier

Pour Pax3, l'homologue le plus semblable à Pax7, un variant d'épissage qui présente la forme Q- a une affinité de liaison plus forte aux sites PD et une activité transcriptionnelle plus élevée que la forme Q+ [187, 188]. Il a été proposé que l'absence de Q qui se trouve dans la région reliant les sous-domaines PAI et RED crée une proximité plus étroite des sous-domaines, ce qui augmenterait l'affinité de liaison. Pour le dipeptide GL, des prédictions de structure secondaire indiquent qu'il aurait une influence importante sur la conformation de la protéine puisque la troisième hélice alpha du sous-domaine RED serait allongée [189]. Il a été montré dans la littérature que l'interdépendance des DBDs des facteurs Pax3 et 7 et la variation structurale dans le domaine PD sont responsables des différentes spécificités de liaison à l'ADN de ces protéines. Il est donc vraisemblable que l'existence de plusieurs transcrits de Pax7 différents au niveau du domaine PD dans un organisme confère à cette protéine une diversité fonctionnelle supplémentaire [187, 188, 190]. Cette diversité aurait le potentiel de médier des fonctions distinctes dans l'embryon en développement. Les quatre transcrits de Pax7 possèdent des profils d'expression différentiels selon les lignées cellulaires et même selon les souches de souris [191]. Ziman et al ont comparé l'abondance de chacun des transcrits dans les tissus musculaires et cérébraux [189, 192]. Quoique les quatre transcrits soient détectables dans les cellules musculaires, Pax7 Q+GL- est abondant et Pax7 Q+GL+ semble

majoritaire. Pour ce qui est du cerveau, les transcrits Pax7 Q+GL+ et Pax7 Q+GL- sont tous deux enrichis et les transcrits Pax7 Q-GL+ et Pax7 Q-GL- sont peu ou pas représentés. Les isoformes présentes dans le lobe intermédiaire de l'hypophyse de souris sont majoritairement l'isoforme Pax7 Q+GL- (75%), et dans une moindre mesure Pax7 Q+GL+ (25%).

Au niveau fonctionnel, l'isoforme Pax7 Q+GL+ permet la différenciation des cellules pluripotentes P19 vers une identité neuronale et cette isoforme, active plus efficacement l'expression du récepteur du facteur neurotrophique ciliaire (Cntfr) que les autres isoformes Pax7 [193, 194]. Dans les cellules souches isolées du muscle, cette même isoforme active efficacement le programme myogénique et augmente aussi l'expression de Pax7 endogène (et donc possiblement d'autres isoformes) [195].

Chez l'humain, une forme supplémentaire de Pax7 a été trouvée dans différents tissus. Il s'agit du résultat d'une terminaison transcriptionnelle différentielle dans l'exon 8 au lieu de l'exon 9 [196] [197]. Cette troncation de l'extrémité C-terminale mène à la perte du domaine OAR (Pax7 OAR-) et pourrait présenter des propriétés de transactivation différentes. Cette forme semble présente faiblement dans plusieurs tissus (excepté les tissus du cerveau fœtal et adulte), tandis que la forme Pax7 pleine longueur (OAR+) est exprimée abondamment et de manière spécifique dans les tissus du cerveau, des muscles et des testicules [196].

### 1.3.6 Alteration de Pax7 et conséquences

Chez la souris, les mutants déficients pour Pax7 sont viables jusqu'à deux à trois semaines après la naissance avec des retards de développement et une faiblesse musculaire [198]; [199, 200]. L'inactivation du gène PAX7 cause une absence complète de cellules satellites dans le muscle squelettique [201] et des défauts de développement de structures cranio-faciales dérivées de la crête neurale [199]. Chez l'humain, des études d'association pangénomique (genome-wide association studies - GWAS) ont révélé que des mutations du gène PAX7 constituent un facteur de risque pour des malformations congénitales. Plusieurs variants rares et potentiellement pathogènes situés dans ou à proximité du gène PAX7 ont été détectés chez des patients présentant des fentes orofaciales (soit des fentes labiales ou palatines) [202-205]. Les maladies associées à PAX7 incluent la myopathie congénitale progressive avec scoliose (MYOSCO) où les variants PAX7 homozygotes associés à ces maladies sont concentrés dans le domaine PD [206]. Récemment des mutations dans les exons 6 à 8 de Pax7 sont ressortis chez neuf personnes d'une cohorte de patients présentant une scoliose congénitale [207]. Une mutation de PAX7 a aussi été associée à un syndrome de retard de croissance, d'hypotonie et de retard neurodéveloppemental global. Cette mutation cible un site d'épissage dans PAX7 qui empêche l'expression de l'isoforme PAX7 OAR+ [196]. De plus, les TFs PAX3 et PAX7 sont impliqués dans la pathogenèse du rhabdomyosarcome alvéolaire (ARMS), un cancer pédiatrique des tissus mous lié à la lignée musculaire striée [190, 208]. Les translocations chromosomiques

caractéristiques du ARMS génèrent des fusions PAX3-FKHR et PAX7-FKHR qui contiennent le DBD de PAX3/PAX7 et le domaine d'activation transcriptionnelle FKHR, ayant pour conséquence la surexpression des cibles des protéine PAX [209].

En définitive, les mécanismes de l'action pionnière montrent leur implication majeure dans la spécification du destin cellulaire, la reprogrammation cellulaire et le développement de pathologie. Les facteurs de transcription pionniers constituent toutefois un groupe hétérogène, qui, contrairement à d'autres TFs, sont capables de reconnaître et lier leurs séquences d'ADN cibles dans la chromatine fermée puis d'organiser le recrutement de facteurs de remodelage pour décompacter la chromatine et entraîner l'apparition d'un nouveau groupe de régions régulatrices actives sur l'ADN.

## Objectifs et hypothèses du projet de thèse

Mon projet est de comprendre quels sont les domaines clés de Pax7 pour les différentes étapes de l'action pionnière nécessaire à l'établissement du transcriptome mélanotrope.

Pour cela, j'ai commencé par étudier les isoformes naturelles de Pax7. L'idée ici est de savoir si les quatre isoformes Pax7 présentant des variations dans le domaine PAIRED ont toutes la capacité pionnière observée avec l'isoforme Q+GL- dans l'hypophyse [74, 105]. Comme cité dans l'introduction, ces isoformes peuvent être exprimées préférentiellement dans certains organes et à différents temps dans le développement. Le fait que certaines de ces isoformes puissent avoir l'action pionnière alors que d'autres non, pourrait nous aider dans la compréhension du rôle de ces facteurs lors de la mise en place de certains organes au cours du développement.

J'ai aussi étudié la capacité pionnière de Pax3. C'est la protéine la plus semblable à Pax7 dans la famille Pax. Pax3 ne présente pas le domaine OAR caractéristique de la protéine Pax7. Si Pax3 n'est pas exprimé dans l'hypophyse, il est présent dans le muscle et le cerveau où Pax7 est aussi exprimé. Ces deux protéines ont des fonctions distinctes bien que certaines de leurs capacités soit redondantes [210, 211]. Le fait de distinguer lesquelles de ces protéines ont l'action pionnière ou non permettrait de mieux comprendre leur rôle.

Enfin, j'ai fait une étude structure fonction de la protéine Pax7 basée sur ses zones d'homologie avec la protéine Pax3. Là encore, nous voulons déterminer quels sont les domaines qui participent à l'action pionnière de Pax7.

Pour clarifier notre hypothèse, rappelons-nous de l'introduction de cette thèse où a été expliqué le travail de Arthur Gouhier en 2022 ; il a montré que l'action pionnière se déroule en deux phases. Une première phase, indépendante de la réplication, pendant laquelle la marque répressive H3K9me2 est enlevée et la marque d'enhancer H3K4me1 est mise en place. Une deuxième phase, dépendante de la réplication, où un nucléosome (histone H3) est déplacé ou enlevé entraînant une ouverture de la chromatine. Le profil de H3K4me1 passe d'un profil monomodal à bimodal caractéristique des enhanceurs actifs. Enfin, la marque activatrice H3K27ac est mise en place par l'histone acétyltransférase P300.

Nous voulons savoir quels domaines de Pax7 sont impliqués dans l'une ou l'autre de ces phases et quel sont les tenants moléculaires de cette implication.

Il sera utile de comprendre quels sont le ou les domaines qui régissent les différentes étapes de l'action pionnière pour permettre de moduler cette action et de l'utiliser à terme pour la transdifférenciation de cellules de patient ex vivo.

Ce travail n'a pas été fait seul. Bien qu'ayant fait la plupart des expériences, j'ai été aidé par des collègues formidables. Audrey Pelletier a notamment eu la charge de faire les gels à retardement et les transfections des isoformes Pax7 et Pax3 pour tester respectivement les capacités de liaison et de transcription de ces protéines. Arthur Gouhier, Aurelio Balsalobre et Amandine Bemmo ont été d'une grande aide en bio-informatique. Les autres personnes dans mon laboratoire Ryhem, Juliette, Kevin, Konstantin, Bastien, Alexandre, Jonathan et mon directeur de laboratoire Jacques Drouin ont été une grande source d'inspiration et de discussion pour l'aboutissement de ce projet et je les en remercie tout autant.

## Chapitre 2-Résultats

### 2.1 Des isoformes Pax7 aux capacités pionnières différentes

#### 2.1.1 Les isoformes GL+ sont incapables d'activer le transcriptome mélanotrope

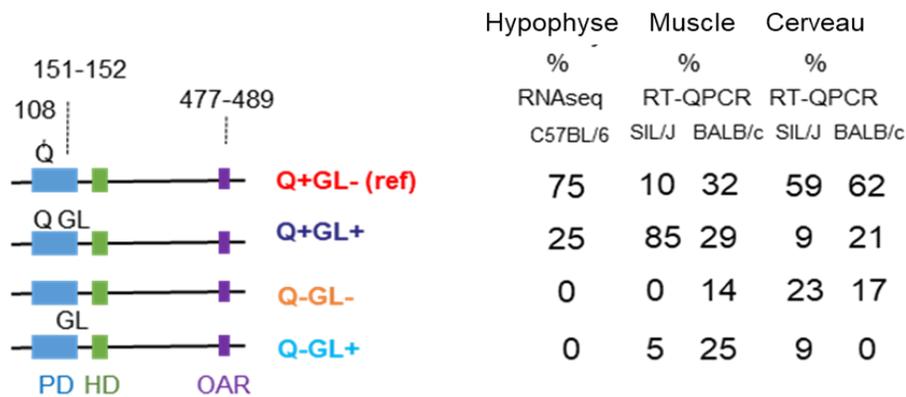
Nous avons étudié les isoformes de Pax7 pour comprendre leur pertinence dans le développement. Il est répertorié chez la souris et l'homme quatre isoformes majeures pour la protéine Pax7. Elles résultent de l'épissage de transcrits alternatifs au niveau de la jonction de l'intron 2 et de l'exon 3 (inclusion ou exclusion du glutamine) ou/et au niveau de la jonction de l'intron 3 et de l'exon 4 (inclusion ou exclusion d'une glycine et leucine) (Figure 2.1 a). Ces insertions ou excisions d'un seul ou d'une paire d'acides aminés se retrouvent toutes deux dans le domaine de liaison d'ADN PAIRED (PD) (Figure 2.1 a). Ces isoformes ainsi créées sont appelées Pax7 Q-GL+, Pax7 Q+GL+, Pax7 Q-GL- et Pax7 Q+GL-. Au vu de la position de ces variations, nous avons supposé que les caractéristiques structurelles différentes de ces isoformes pourraient entraîner des spécificités dans leur fonction notamment au niveau de leur recrutement sur l'ADN [187, 188, 190]. Cette diversité de fonction pourrait s'avérer pertinente lorsque la protéine Pax7 est appelée à exercer des fonctions distinctes à des moments précis de développement ou dans des organes spécifiques. Chez la souris, des travaux antérieurs ont défini l'expression relative des quatre isoformes de Pax7 dans des cellules musculaires et neuronales par RT-QPCR [189] et nous avons évalué par ARNseq l'expression relative des isoformes de Pax7 dans l'hypophyse. Que ce soit dans le muscle, le cerveau ou l'hypophyse, on remarque que les isoformes Pax7 Q- sont celles qui sont le moins exprimées. Dans le cerveau, l'isoforme Pax Q+GL- est prédominante. Dans le muscle, selon le type de souris étudié, l'isoforme Pax7 Q+GL+ peut être prédominante ou également exprimée avec l'isoforme Q+GL-. Les isoformes présentes dans le lobe intermédiaire de l'hypophyse de souris sont majoritairement Pax7 Q+GL- (75%), et dans une moindre mesure Pax7 Q+GL+ (25%) (Figure 2.1 a).

Nous savons que, dans l'hypophyse de souris, l'isoforme Q+GL- est un facteur pionnier permettant la spécification des mélanotropes [105, 167]. Suite à l'arrivée de Pax7, les régions régulatrices spécifiques aux mélanotropes perdent la marque répressive H3K9me2 et arborent une décompaction de la chromatine avec des marques d'histones d'enhancers actifs telles que les marques bimodales de H3K4me1 et H3K27ac [167]. Cela a pu être reproduit partiellement dans un système de gain de fonction précédemment développé dans les cellules de souris AtT-20. Suite à l'expression ectopique du facteur pionnier Pax7 Q+GL-, ces cellules corticotropes subissent une reprogrammation en cellules mélanotropes (Figure 2.1 b) [105].

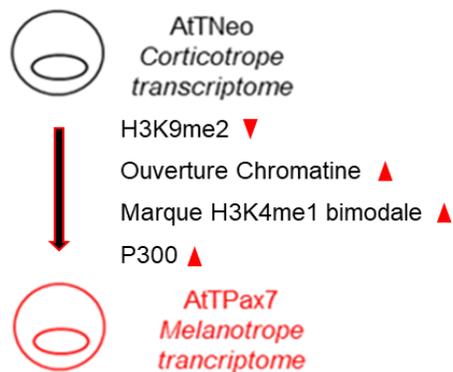
Dans ce modèle de gain de fonction, le cDNA de Pax7 Q+GL- est fusionné à une séquence marqueur appelée Flag en N et C-terminal. Le Pax7 Q+GL- ainsi flaggué a été placé sous le contrôle de séquences rétrovirales (promoteur

cytomégalovirus) dans un plasmide présentant un gène de résistance à la néomycine. Ce plasmide a ensuite été inséré dans les cellules AtT20 par infection. Les cellules exprimant la protéine Pax7 ont été sélectionnées avec de la néomycine et l'expression de Pax7 dans ce groupe de cellules infectées a été évaluée par western blot. La liaison de Pax7 a été définie par CHIPseq avec un anticorps reconnaissant la séquence FLAG. Enfin, les changements épigénétiques entraînés par la protéine Pax7 Q+GL- ont été déterminés par CHIPseq pour la marque répressive H3K9me2, la marque d'enhancer H3K4me1 et la protéine P300. Enfin, l'accessibilité de l'ADN a été mesurée par ATACseq (Figure 2.1 b) [105, 167].

## a) Isoformes Pax7



## b) Modele de travail



## c) Western blot

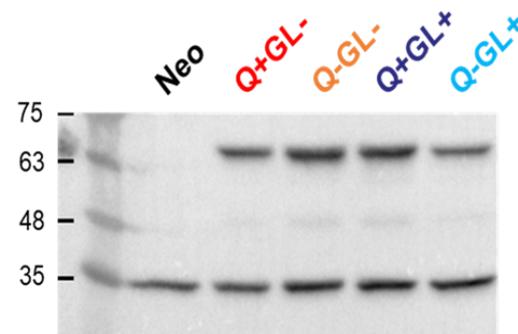


Figure 2.1 Mise en place du projet

a) (Gauche) Schéma des différentes isoformes de Pax7 présentant des variations dans le domaine PAIRED. (Droite) Résumé bibliographique résumant la proportion quantitative des quatre isoformes Pax7 dans l'hypophyse, le muscle et le cerveau.

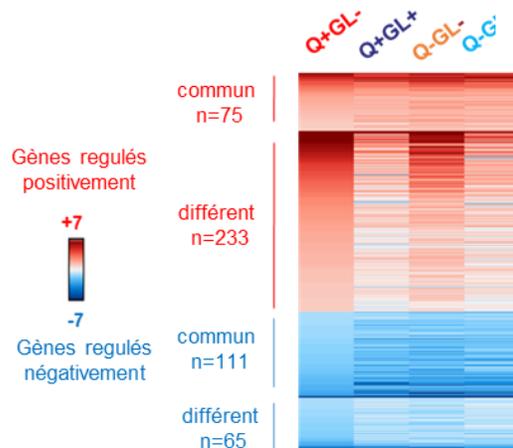
b) Modèle résumant le travail de Mayran et al [167] sur le rôle pionnier de Pax7 dans le modèle de cellules corticotropes AtT20.

c) Western blot de la protéine Pax7 et de la protéine contrôle PCNA pour les différents groupes de cellules exprimant chacun une isoforme de Pax7.

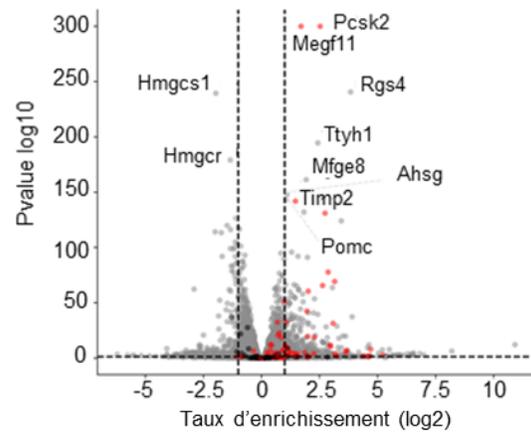
Dans le but de déterminer si toutes les isoformes de Pax7 sont capables d'avoir l'action pionnière responsable de la spécification des mélanotropes, nous avons créé des groupes de cellules AtT-20 exprimant chacune une

isoforme Pax7 différente. Nous avons par la suite montré que le niveau d'expression des isoformes Pax7 mesurée dans nos nouveaux groupes de cellules AtT20 était similaire à celui retrouvé dans le groupe de cellules AtT20 exprimant notre isoforme référence PAX7 Q+GL- (Figure 2.1 c). Des transcriptomes ont été obtenus par ARNseq pour chaque groupe de cellules exprimant une isoforme Pax7 spécifique. Nous avons évalué la capacité de chacune des isoformes des Pax7 à activer le transcriptome mélanotrope dans les cellules corticotropes AtT20 comme l'isoforme de référence Pax7 Q+GL- (Figure 2.2 a, b, c, d). Cette comparaison révèle que les isoformes GL+ affectent l'expression de beaucoup moins de gènes que les isoformes sans insertion de GL. La présence de Q, quant à elle, a un effet marginal sur le transcriptome activé et réprimé par l'isoforme Pax7 Q+GL- (Figure. 2.2 a). De façon significative, la représentation des gènes mélanotropes en rouge et corticotropes en noir exprimés de façon différentielle entre les isoformes Q+GL- et Q+GL+ révèle que la plupart (91 %) des gènes mélanotropes précédemment identifiés [105, 167] sont activés préférentiellement (supérieur de plus de 2 fois avec une p. value inférieur à 0.05) dans les cellules Q+GL- par rapport aux cellules Q+GL+ (Figure 2.2 b, c, d). La capacité de Pax7 à mettre en œuvre ce programme mélanotrope semble donc dépendre principalement des isoformes GL- et est empêchée par l'insertion de GL dans le domaine PD de l'isoforme GL+.

a) Transcriptome des isoformes Pax7



b) Volcanoplot ARNseq Pax7 Q+GL+ vs Pax7 Q+GL-

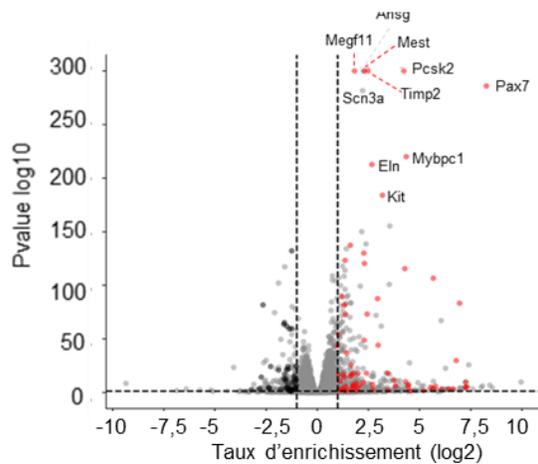


Gènes mélanotropes (n=72)

Gènes corticotropes (n=46)

Autres gènes

c) Volcanoplot ARNseq Pax7 Q+GL- vs Neo

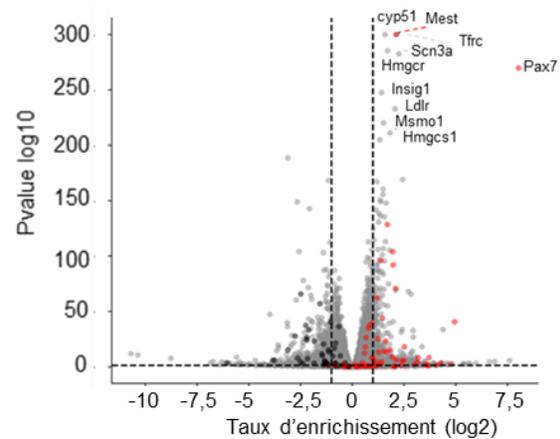


Gènes mélanotropes (n=72)

Gènes corticotropes (n=46)

Autres gènes

d) Volcanoplot ARNseq Pax7 Q+GL+ vs Neo



Gènes mélanotropes (n=72)

Gènes corticotropes (n=46)

Autres gènes

Figure 2.2 Les isoformes Pax7 GL+ ne réussissent pas à activer le transcriptome mélanotrope

a) Représentation des ARNseq des différents groupes de cellules exprimant chacun une isoforme spécifique de Pax7 en heatmap.

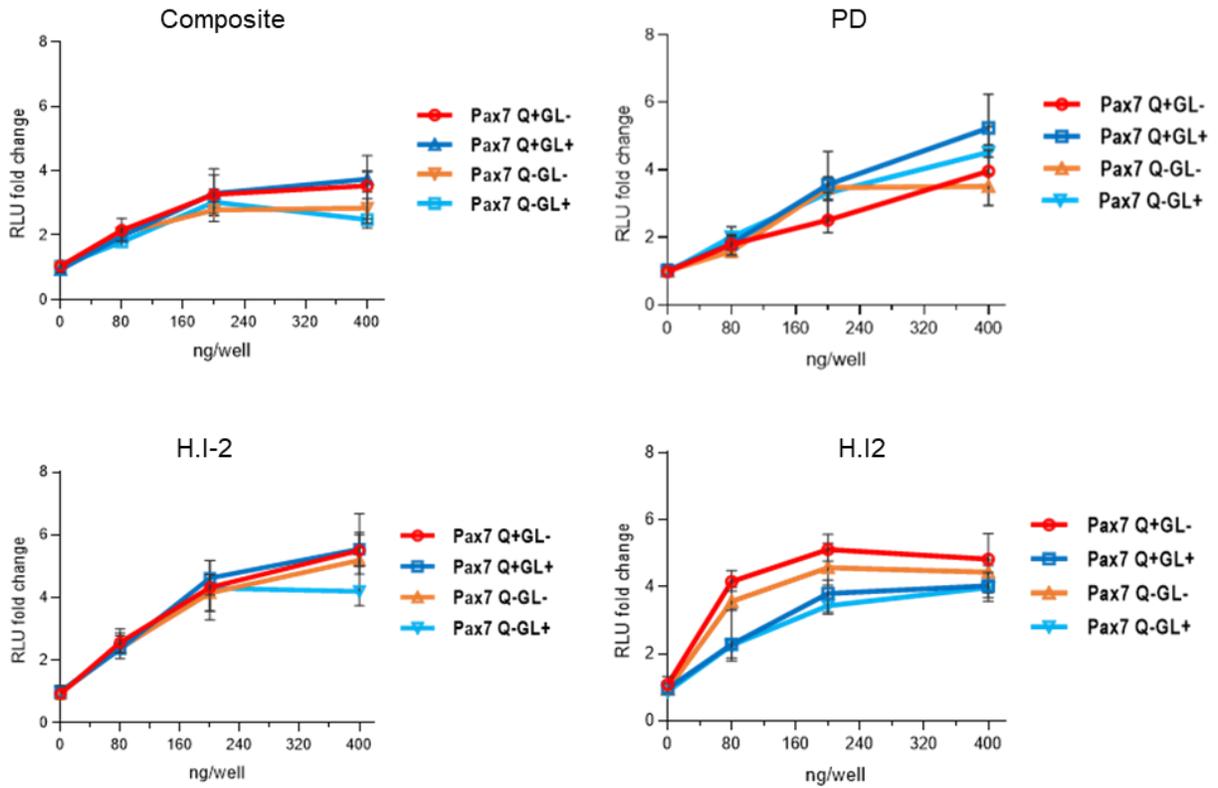
b) Volcanoplot représentant les gènes différentiellement exprimés entre les cellules exprimant Pax7 Q+GL- ou Pax7 Q+GL+.

c) Volcanoplot représentant les gènes différentiellement exprimés entre les cellules exprimant Pax7 Q+GL- ou le plasmide de référence (Néo)

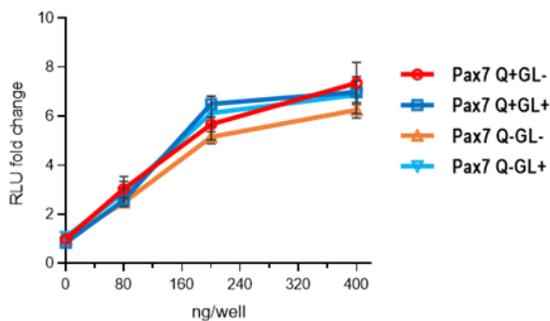
d) Volcanoplot représentant les gènes différentiellement exprimés entre les cellules exprimant Pax7 Q+GL+ ou le plasmide de référence (Néo)

Nous nous sommes demandé si cette expression différentielle des gènes pouvait être due à des différences dans la capacité d'activation transcriptionnelle des isoformes GL+ par rapport aux isoformes GL-. Nous avons évalué cela directement par transfection. Plus précisément, nous avons co-transfecté un plasmide contenant une des isoforme Pax7 avec un plasmide rapporteur contenant le gène luciférase dans les cellules AtT20. Ce gène rapporteur était soit sous le contrôle d'enhancers artificiels contenant des multimères de sites de liaison d'ADN reconnus spécifiquement par Pax7 (COMPOSITE, PAIRED, HI-2, HI2) [133] (Figure 2.3 a) , soit sous le contrôle d'un enhancer naturel reconnu par Pax7 tel que l'enhancer du gène spécifique des mélanotropes PCSK2 [105] (Figure 2.3 b). Ces cellules ont été ensuite lysées et on a ajouté à ce lysat de la luciférine qui produit de la lumière mesurée en RLU (intensité de lumière relative) au contact de la protéine luciférase. La capacité de chaque isoforme Pax7 à activer le gène luciférase sous le contrôle de son enhancer associé est une mesure directe de l'activité transcriptionnelle de l'isoforme. Nous avons constaté que les différents isoformes de Pax7 ont une capacité transcriptionnelle très similaire dans tous ces systèmes (Figure 2.3 a, b). Nous en avons conclu que ce n'est pas la capacité de transcription qui explique les propriétés uniques des isoformes observées par ARNseq. Un exemple précis est montré avec les niveaux de l'ARNm (ARN messenger) du gène PCSK2. Pour ce gène, les isoformes GL- expriment des niveaux significativement plus élevés d'ARNm PCSK2 que les isoformes GL+. La plus grande différence étant entre les deux isoformes exprimées dans l'hypophyse : les isoformes Pax7 Q+GL- et Pax7 Q+GL+ (Figure 2.3 c). Ceci est d'autant plus intéressant que l'expression de Pcsk2 avait été précédemment montrée comme dépendante de la capacité pionnière de Pax7 [212].

## a) Activation d'enhancers artificiels composés de motifs reconnus par Pax7 (transfection)



## b) Activation de l'enhancer Pcsk2 (transfection)



## c) Pcsk2 expression (RNAseq)

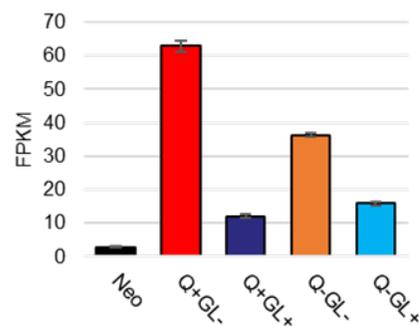


Figure 2.3 Toutes les isoformes ont globalement une activité transcriptionnelle similaire.

a) Transfection des isoformes Pax7 avec un rapporteur sous le contrôle d'enhancer présentant plusieurs copies de la séquence composite (en haut à gauche), la séquence PD (en haut à droite), la séquence H.I-2 (au milieu à gauche), la séquence H.I2 (au milieu à droite).

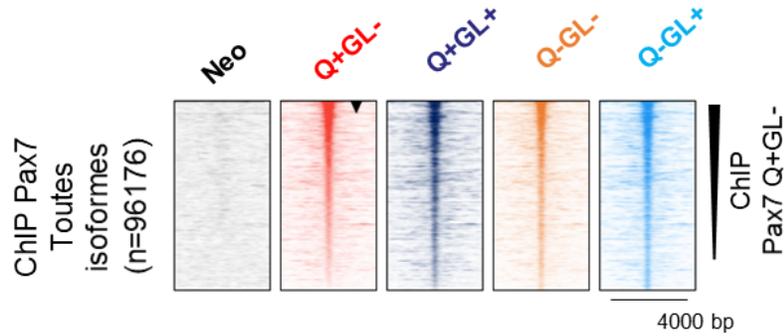
b) Transfection des isoformes Pax7 avec un rapporteur comprenant l'enhancer naturel de Pcsk2.

c) Données RNAseq pour l'expression du gène Pcsk2

### 2.1.2 Les isoformes Pax7 ont un recrutement similaire sur le génome

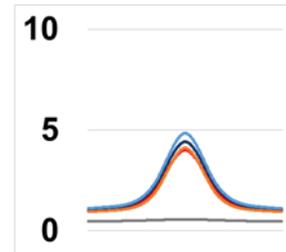
Puisque l'activité transcriptionnelle intrinsèque des isoformes de Pax7 n'explique pas les différences de transcriptome entre nos différents groupes de cellules, nous nous sommes demandé si la différence pouvait être expliquée par une capacité de recrutement sur l'ADN différente pour chaque isoforme Pax7. Nous avons donc évalué la capacité de chaque isoforme Pax7 à être recrutée sur l'ADN par ChIPseq (Figure 2.4 a). Ces expériences de ChIPseq ont été faites deux fois (réplicas biologiques) et présentent une corrélation satisfaisante avec un coefficient de corrélation (R) supérieur à 0.7 (Figure 2.4 b). Globalement, toutes les isoformes ont montré un recrutement semblable (Figure 2.4 a, c). Nous avons décidé de regarder plus spécifiquement la capacité de recrutement des différentes isoformes Pax7 aux différentes catégories d'enhancers impactant le transcriptome des AtT20. Parmi ces enhancers impactant le transcriptome, nous avons les catégories suivantes : les enhancers constitutifs (ouverts et actifs avant et après l'expression de Pax7), les enhancers activés (gains de signaux ATAC, H3K4me1 après la liaison de Pax7 avec des signaux faibles observés antérieurement et un gain de p300 qui n'était pas présent antérieurement) et les enhancers pionniers (pas de marques d'enhancers actifs ATAC, H3K4me1 et p300 avant Pax7 mais ces marques sont gagnées après) (Figure 2.5 a). Les profils moyens de la force de liaison des isoformes Pax7 sur ces différentes catégories d'enhancers montrent que les isoformes GL+ se retrouvent avantagées pour le recrutement sur les sites constitutifs mais désavantagées pour le recrutement sur les sites pionniers. On remarque aussi que globalement l'isoforme Q-GL-semble avoir une liaison similaire à l'isoforme Q+GL-, la variation Q semble donc avoir un faible impact sur le recrutement de Pax7 à l'ADN in vivo (Figure 2.4 d).

a) ChIPseq: recrutement des isoformes Pax7

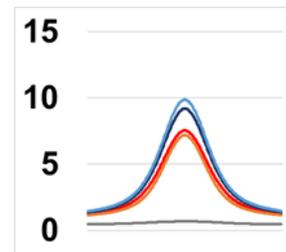


d) Profil moyen de chaque isoformes Pax7 sur les différentes catégories d'enhancers ou est recruté Pax7 (ref)

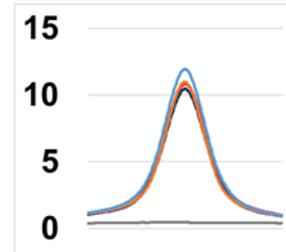
Sites de toute les isoformes (n=96176)



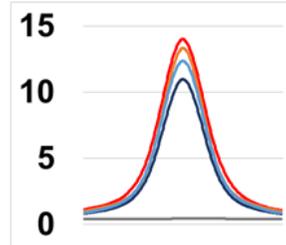
Sites constitutifs commun (n=9205)



Sites activés commun (n=633)

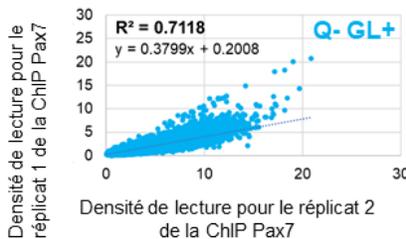
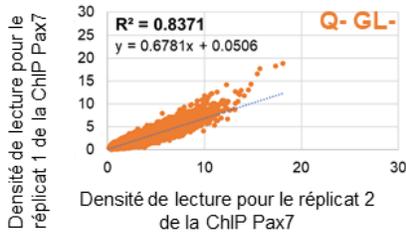
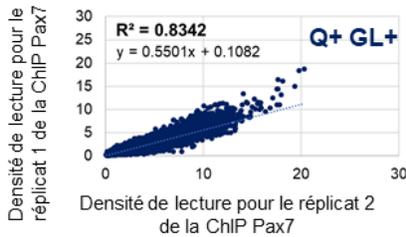
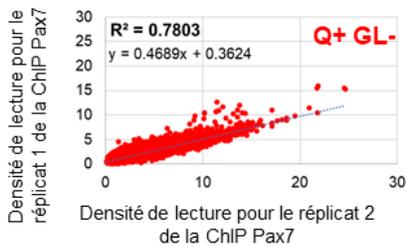


Sites pionniers commun (n=735)



Q+ GL-  
Q+ GL+  
Q- GL-  
Q- GL+

b) Reproductibilité des réplicats de ChIPseq



c) Comparaison des réplicats 1

Courbe de tendance linéaire:  
Y=AX+B  
R2=COEFFICIENT DE DETERMINATION

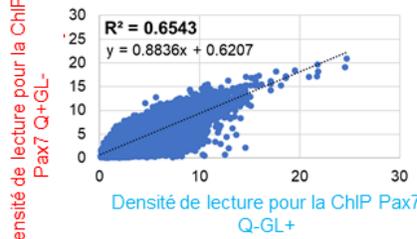
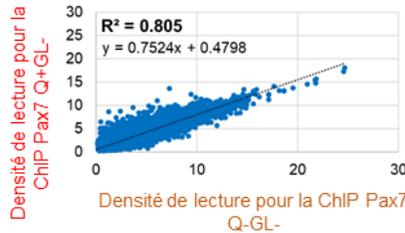
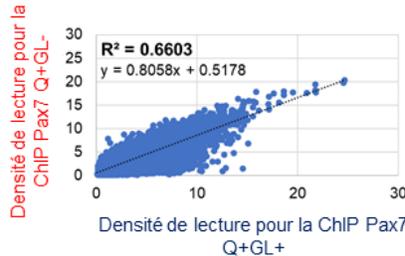


Figure 2.4 Reproductivité adéquate des ChIPseqs Pax7

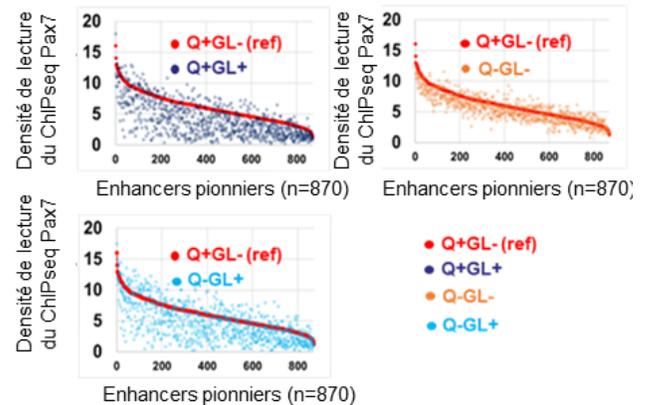
- a) ChIPseqs des isoformes Pax7 faits avec l'anticorps Flag et représentés en heatmap ordonné de manière croissante en fonction de la force de liaison du Pax7 Q+GL- référence (flèche noire).
- b) Diagramme de corrélation des réplicas de ChIPseq pour chaque isoforme Pax7.
- c) Diagramme de corrélation entre le réplica 1 du ChIPseq de Pax7 Q+GL- (référence) et le réplica 1 des ChIPseq des autres isoformes Pax7
- d) Profils représentant le pic moyen des isoformes Pax7 pour chaque catégorie d'enhancers décrits dans [133] liés par Pax7 Q+GL- (référence)

Le recrutement est similaire pour toutes les isoformes Pax7 aux enhancers constitutifs et activés avec environ 90% des sites occupés pour chacune des isoformes (Figure 2.5 a). On note aussi que, pour les isoformes GL+, la force de liaison est diminuée mais non perdue aux enhancers pionniers par rapport à l'isoforme Q+GL- de référence (Figure 2.5 b) à l'exception d'un petit groupe de 115 sites pionniers (14 %) où le recrutement est perdu, décrit comme GL-sensible (liaison) pour la suite de la thèse (Figure 2.5 a, b). Lorsqu'on regarde l'impact de la force de liaison des isoformes Pax7 GL+ sur l'ensemble des sites pionniers liés par Pax7 Q+GL-, on se rend compte que cette perte n'est pas homogène comparé à la force de liaison de l'isoforme Pax7 Q+GL-. Nous supposons que cette perte de liaison serait donc dépendante du site regardé et de la protéine étudiée (Figure 2.5 c). Nous nous sommes demandé si cette hétérogénéité de force de liaison aux différents sites pionniers était due à une déficience des isoformes GL+ pour certains motifs d'ADN reconnus par Pax7 (COMPOSITE, PAIRED, HI-2, HI2) avec l'un et/ou l'autre de ses domaines de liaison d'ADN (PD et HD) [133]. Nous avons donc évalué, par une expérience de retard sur gel, la liaison des isoformes de Pax7 à des sondes comprenant les motifs reconnus par Pax7 (COMPOSITE, PAIRED, HI-2, HI2). L'expérience de retard sur gel consiste à observer le retard de migration sur gel d'une sonde d'ADN marquée radioactivement lorsque la protéine reconnaissant le motif ADN de cette sonde s'y lie. Il est à noter que, plus l'affinité entre la protéine et la sonde est grande, plus la bande retardée est large et foncée à la révélation du gel. Ces expériences ont montré que la capacité de liaison in vitro aux différents motifs d'ADN reconnus par Pax7 est similaire pour les différentes isoformes bien qu'avec une liaison légèrement plus faible des isoformes GL+ par rapport aux isoformes de GL- pour toutes les sondes étudiées (Figure 2.5 d). Nous en avons conclu que le recrutement plus faible observé pour les isoformes GL+ dans les expériences ChIPseq pourrait être en partie dû à une affinité de liaison intrinsèque plus faible aux différents motifs reconnus par Pax7 dans l'ADN. Cependant, il est difficile d'expliquer comment la faible perte de liaison homogène observée sur tous les motifs de liaison de PAX7 in vitro (Figure 2.5 d) entraîne l'hétérogénéité de perte de recrutement observée pour les isoformes GL+ sur l'ensemble des sites pionniers in vivo (Figure 2.5 c).

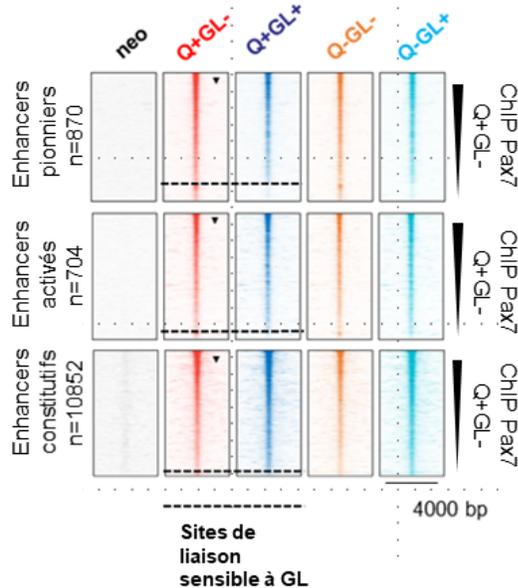
a) Occupation des isoformes Pax7 aux différents types d'enhancers (ChIPseq)

	ATAC		H3K4me		p300		Nombre de pics comparé au Pax7 référence (%)			
	neo	Pax7	neo	Pax7	neo	Pax7	Pax7 Q+GL-	Pax7 Q+GL+	Pax7 Q-GL-	Pax7 Q-GL+
Enhancers pionniers (n=870)	-	++	-	++	-	++	100	86	98	89
Enhancers activés (n=704)	+	++	+	++	-	++	100	94	97	95
Enhancers constitutifs (n=10852)	++	++	++	++	++	++	100	96	87	96

c) Force de liaison des isoformes Pax7 aux sites pionniers



b) Heatmaps des isoformes Pax7 aux différents types d'enhancers (ChIPseq)



d) Liaison in vitro des isoformes Pax7

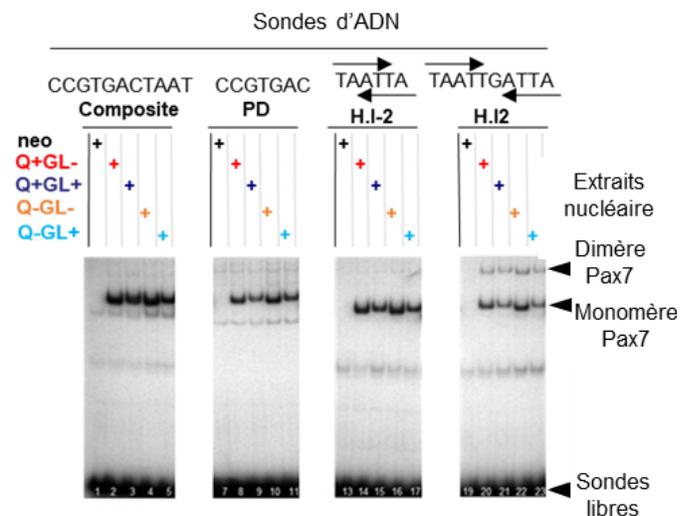


Figure 2.5 Toute les isoformes de Pax7 ont un recrutement génomique similaire

a) Occupation des isoformes de Pax7 dans les différentes catégories d'enhancers actifs reconnues par Pax7 Q+GL. La présence/absence de différentes marques utilisées pour définir les catégories d'enhancers avant et après l'arrivée de Pax7 est indiquée dans le tableau à gauche.

b) Représentation par heatmap du recrutement des isoformes de Pax7 aux catégories d'enhancer citées en a) et ordonné de manière croissante par rapport à la force de liaison de l'isoforme Pax7 référence Q+GL-. Un petit sous-ensemble d'environ 11 % des sites pionniers n'est plus lié par les isoformes GL+, comme l'indique la ligne pointillée en bas des heatmaps.

- c) Comparaison de l'efficacité de recrutement des isoformes Pax7 sur les sites pionniers.
- d) Évaluation par essai de gel à retardement (EMSA) de la liaison in vitro des isoformes de Pax7 à diverses sondes d'ADN (séquence Composite, PAIRED (PD) ou HD (H.I-2, H.I2)). Les positions des monomères et des dimères Pax7, ainsi que des sondes libres, sont indiquées sur la droite.

### 2.1.3 Pax7 Q+GL+ prime les enhancers mélanotropes au lieu de les activer

Puisque la liaison des isoformes au génome semble difficilement expliquer la perte du transcriptome mélanotropes observée Figure 2.2 a, nous avons évalué la capacité de l'isoforme Pax7 Q+GL+ à initier l'ouverture de la chromatine. Pour cela, nous avons réalisé des ATACseq sur les groupes de cellules exprimant les isoformes hypophysaires Pax7 Q+GL- et Q+GL+. Des signaux ATACseq similaires sont observés aux sites constitutifs pour ces deux isoformes. Cependant l'isoforme Q+GL+ est incapable d'ouvrir un sous-ensemble de sites dans les sous-catégories pionnières et activées (Figure 2.6 a). Environ 45% des sites pionniers ne montrent pas de signal ATACseq significatif en présence de Q+GL+ par rapport à l'isoforme Q+GL- : nous avons nommé ce sous-ensemble de 345 enhancers GL-sensible (ATAC). Les 418 sites pionniers restants qui présentent un signal ATAC ont été nommés GL-insensible (ATAC). Nous avons voulu évaluer si la liaison stable de Pax7 entre les deux sous-ensembles d'enhancers pionniers peut expliquer la différence d'ouverture de chromatine révélée par ATACseq. Pour cela, nous évalué la différence de force de liaison entre les isoformes Pax7 Q+GL+ et Q+GL- sur les sites GL-sensibles et GL-insensibles (ATAC). Bien que la liaison PAX7 ait été plus touchée pour le sous-ensemble de sites sensibles au GL que pour le sous-ensemble insensible au GL, une liaison importante est demeurée pour la plupart des sites dans l'ensemble des 345 enhancers sensibles au GL (ATAC) (Figure 2.6 b, d). Ceci explique difficilement la perte de la capacité d'ouverture de la chromatine révélée par l'ATACseq qui est significativement différente entre les deux sous-ensembles (Figure 2.6 a, e).

Nous avons ensuite voulu évaluer si les motifs de liaison ADN sous-jacents dans chaque sous-ensemble d'enhancers pouvaient expliquer la sensibilité de l'ouverture de la chromatine à l'insertion des résidus GL dans le domaine PD. Nous avons d'abord fait des recherches de motif denovo et n'avons retrouvé aucun motif particulier qui pourrait être considéré comme spécifique pour l'un ou l'autre de nos deux sous ensemble (résultats non montrés). Nous avons ensuite voulu évaluer si certains sites reconnus par Pax7 étaient plus représentés dans l'un de nos sous ensemble. Ni la fréquence des motifs individuels (Figure 2.6 f) ni le nombre total de motifs par sites Pax7 (Figure 2.6 g) ne peuvent expliquer la différence d'ouverture de la chromatine entre les deux sous-ensembles de site pionniers. Cependant, nous avons découvert que les sites GL-sensibles (liaison) présentent moins de sites Pax7 (<3) que les autres sous-groupes.

Nous avons aussi noté que le sous-ensemble d'enhancers qui n'a pas montré de signaux ATACseq significatifs dans les cellules exprimant Pax7 Q+GL+ montre un petit gain en H3K4me1 qui se présente principalement comme un seul pic par rapport à la distribution bimodale H3K4me1 généralement observée aux enhancers pleinement actifs (Figure 2.6 c).

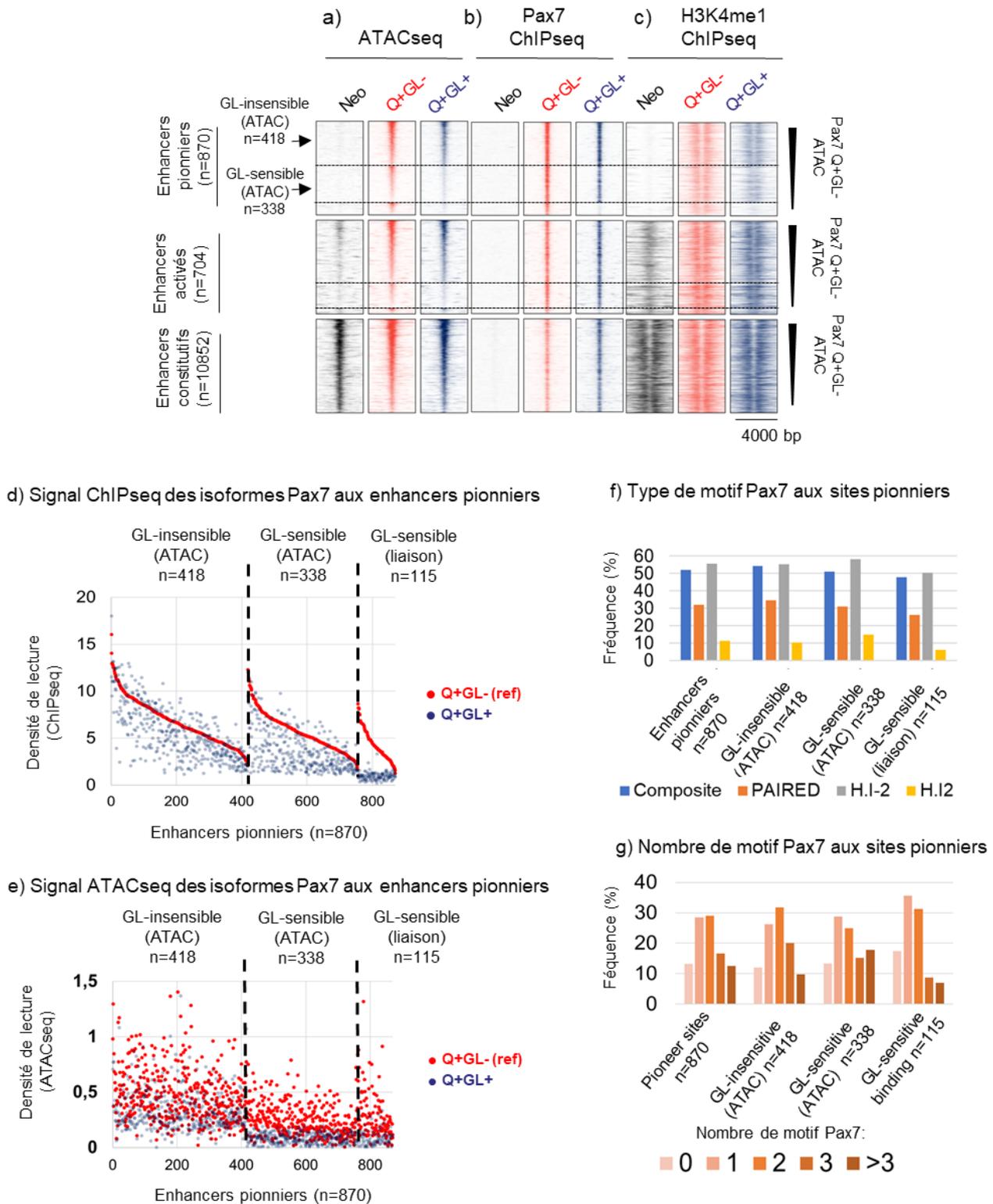


Figure 2.6 L'ouverture de la chromatine propre à l'isoforme Pax7 Q+GL+ est corrélée avec la perte du transcriptome du mélanotrope

a-c) Capacité pionnière des deux isoformes Pax7 exprimées dans l'hypophyse évaluées par ATACseq (a), ChIPseq Pax7 (b) et ChIPseq H3K4me1 (c). Ces données sont représentées par heatmap pour les différentes catégories d'enhancers liées par l'isoforme référence Pax7 Q+GL-. Ces heatmaps sont ordonnées de façon croissante en fonction de l'ATAC de Pax7 Q+GL- et sous-divisées en sous-ensembles insensibles au GL (ATAC), sensibles au GL (ATAC) et sensibles au GL (liaison) basés sur les profils ATAC et ChIPseq de Pax7 Q+GL- et Pax7 Q+GL+.

d) Diagramme représentant les valeurs du ChIPseq Pax7 (densité de lecture sous le sommet du pic Pax7 à plus ou moins 200 paires de bases) pour chacun de 870 sites pionniers. Ce diagramme est ordonné de façon croissante en fonction du ChIPseq de Pax7 Q+GL- et est sous-divisé en sous-ensembles insensibles au GL (ATAC), sensibles au GL (ATAC) et sensibles au GL (liaison) basés sur les profils ATAC et ChIPseq de Pax7 Q+GL- et Pax7 Q+GL+.

e) Diagramme représentant les valeurs d'ATACseq (densité de lecture sous le sommet du pic Pax7 à plus ou moins 200 paires de bases) pour chacun de 870 sites pionniers. Ce diagramme est ordonné de façon croissante en fonction du ChIPseq de Pax7 Q+GL- et est sous-divisé en sous-ensembles insensibles au GL (ATAC), sensibles au GL (ATAC) et sensibles au GL (liaison) basés sur les profils ATAC et ChIPseq de Pax7 Q+GL- et Pax7 Q+GL+.

f) Représentation quantitative des motifs spécifiques de Pax7 dans les sous-ensembles insensibles au GL (ATAC), sensibles au GL (ATAC) et sensibles au GL (liaison).

g) Représentation quantitative du nombre de motif spécifique de Pax7 dans les sous-ensembles insensibles au GL (ATAC), sensibles au GL (ATAC) et sensibles au GL (liaison).

Pour confirmer cette différence, nous avons réorganiser les sites pionniers en fonction du profil bimodal H3K4me1 obtenu dans les cellules Pax7 Q+GL- par rapport au Pax7 Q+GL+ (Figure2.7 a, b, c). Basés sur ce critère, les enhancers pionniers ont été séparés en trois sous ensemble. Le premier comportant 242 sites pionniers est considéré comme insensible à la variation GL (H3K4me1). Les enhancers présentent un profil bimodal après expression de l'isoforme Pax7 GL- ou GL+. Le deuxième sous-ensemble de 512 sites est considéré comme sensible à la variation GL. Ces sites présentent une altération de l'aspect bimodal de la marque H3K4me1 obtenue après l'expression de l'isoforme Pax7 Q+GL- lorsque l'isoforme exprimée est Pax7 Q+GL+. Ces sites présentent un profil de pic unique de H3K4me1 typique des enhancers décrits comme « primés » par Mayran et al [74] . Ces sites présentent un début de remodelage de la chromatine avec un faible niveau d'ATAC, l'apparition d'un pic monomodal H3K4me1 et pas de signe d'activation transcriptionnelle marqué par l'enzyme p300 [74]. Des

exemples de ce deuxième sous ensemble sont présentés pour deux enhancers (Figure 2.7 d). Ce sous ensemble est séparé en 3 groupes en fonction de la symétrie du profil bimodal obtenue dans les cellules exprimant Pax7 Q+GL-. Un groupe de 151 sites présentent un profil bimodal avec un pic H3K4me1 proéminent à droite, un groupe de 105 sites présentent un profil bimodal avec un pic H3K4me1 proéminent à gauche et enfin le dernier groupe de 259 sites présentent un profil bimodal symétrique. L'ensemble des sites contenant ces trois groupes perdent leur aspect bimodal quand l'isoforme Q+GL+ est exprimé au lieu de l'isoforme Q+GL-. Enfin, le dernier sous-ensemble de 115 sites ne présente pas de liaison de l'isoforme Q+GL+ comparé par l'isoforme Q+GL+ et de manière logique ne présente pas de remodelage au niveau de ces sites (Figure 2.7 b, 2.8 a).

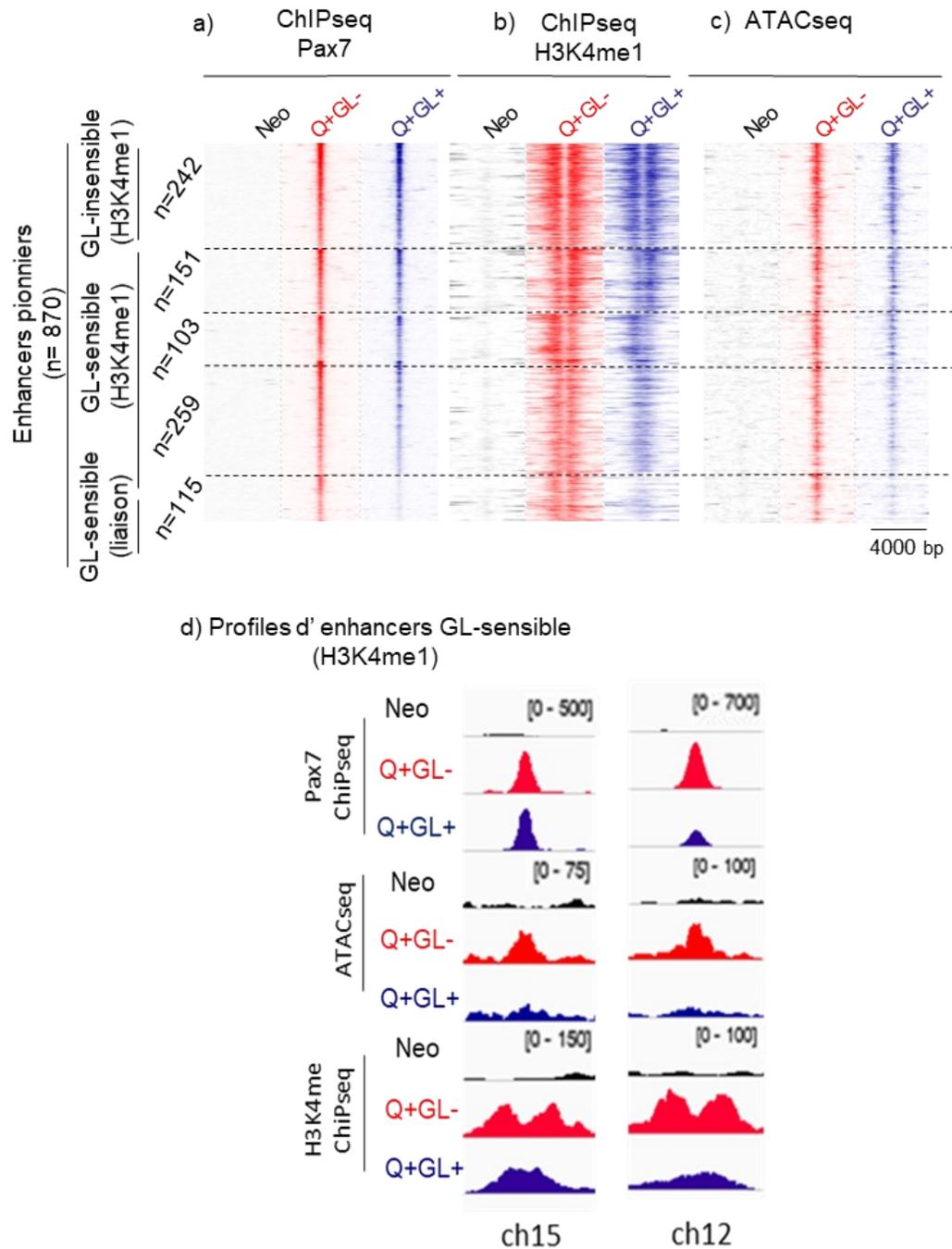


Figure 2.7 L'isoforme Q+GL+ prime les enhancers mélanotropes au lieu de les activer

a-c) Capacités pionnières des deux isoformes Pax7 exprimées dans l'hypophyse évaluées par ChIPseq Pax7 (a), ChIPseq H3K4me1 (b) et ATACseq (c). Ces données sont représentées par heatmap pour les différentes catégories d'enhancers liées par l'isoforme Pax7 référence Q+GL-. Ces heatmaps sont ordonnées de façon

croissante en fonction du ChIPseq Pax7 Q+GL- et sous-divisées en sous-ensembles insensibles au GL (H3K4me1), sensibles au GL (H3K4me1 asymétrique à droite), sensibles au GL (H3K4me1 asymétrique à gauche), sensibles au GL (H3K4me1 symétrique) et sensibles au GL (liaison) basés sur les profils ChIPseq H3K4me1 de Pax7 Q+GL- et Pax7 Q+GL+.

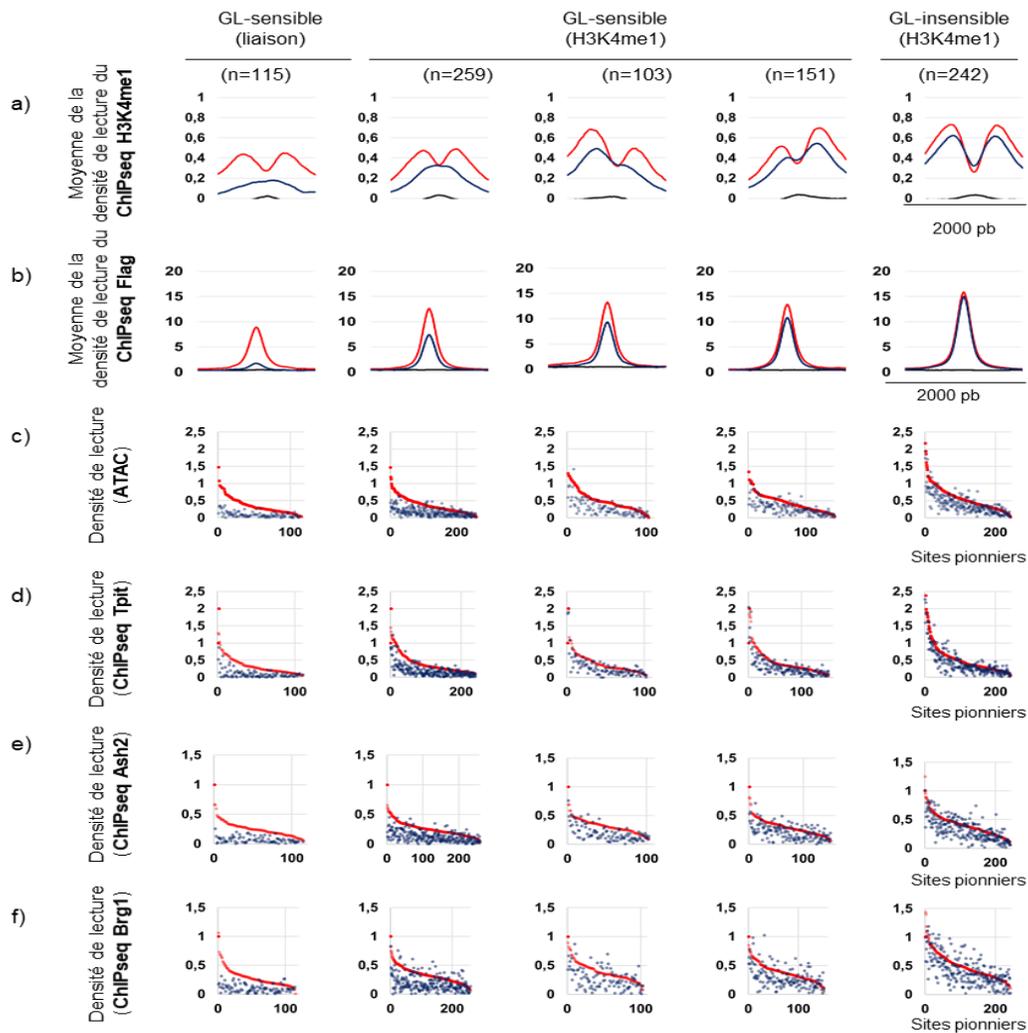
b) Profil IGV (integrative genome view) de sites GL-sensibles (H3K4me1) comprenant le ChIPseq de Pax7, l'ATACseq et le ChIPseq H3K4me1 pour les cellules contrôles AtT-20 (Neo) et les cellules AtT-20 exprimant l'isoforme Pax7 Q+GL- ou Pax7 Q+GL+.

Là encore, nous avons évalué dans quelle mesure la liaison stable de Pax7 peut différer entre les sous-ensembles d'enhancers pionniers (Figure 2.8 a, b) et si cela peut expliquer la différence d'ouverture de la chromatine révélée par ATACseq (Figure 2.8 c). Bien que la liaison PAX7 soit plus touchée pour les sous-ensembles sensibles au GL que pour le sous-ensemble insensible au GL (H3K4me1), la liaison demeure significative pour la plupart des sites à l'intérieur des sous-ensembles d'enhancers sensibles au GL (H3K4me1) (Figure 2.8 b, c). Cependant, la perte de capacité d'ouverture de la chromatine révélée par l'ATACseq est significativement différente entre les différents sous-ensembles (Figure 2.8 b, c).

Nous avons ensuite comparé l'environnement protéique des enhancers sensibles et insensibles à GL (H3K4me1). Pour cela, nous avons fait des ChIPseqs pour différentes protéines impliquées dans le remodelage de la chromatine par Pax7 [Gouhier et al, en préparation] [106]. Parmi elles; la protéine Tpit, un facteur non pionnier nécessaire pour l'ouverture de la chromatine dans les mélanotropes initiée par Pax7 [106]. La protéine BRG1, une ATPase du complexe de remodelage de la chromatine SWI-SNF. La protéine ASH2, un membre du complexe MLL nécessaire à la mise en place de la marque H3K4me1. Alors que le recrutement du facteur de transcription Tpit et des protéines complexes de remodelage Ash2 et BRG1, est similaire pour Q+GL+ et Q+GL- au sous-ensemble d'enhancers insensibles au GL (n=242), leur recrutement est diminué de moitié pour le sous-ensemble un profil symétrique (n=259) (Figure 2.8 d, e, f). Les groupes d'enhancers au profil asymétrique sensible à GL quant à eux présentent des profils Ash2, BRG1 et Tpit qui corrélerent avec la liaison de Pax7 Q+GL+, et qui correspondent à 70% de la liaison Pax7 Q+GL. Bien que ces diminutions de recrutement soient corrélées à une plus faible ouverture de chromatine (ATAC), elles sont plus particulièrement associées à un échec de l'éviction nucléosomale (un seul pic plutôt que des profils bimodaux H3K4me1) et des enhancers qui semblent bloqués à l'état « primé ». Nous avons essayé d'établir des corrélations entre la présence du facteur de transcription Pax7 Q+GL+ et les facteurs Tpit, Ash2, et BRG1. On note que les facteurs Ash2 et Brg1 ont une corrélation modérée ( $R^2 \approx 0.48$ ) avec l'occupation de l'isoforme Pax7 Q+GL+ suggérant que leurs pertes pourraient être expliquée en partie par la perte du recrutement qu'on observe pour cette isoforme de Pax7. A l'inverse, la présence de Tpit n'est pas corrélée avec

la présence de Pax7 Q+GL+ ( $R^2=0.27$ ). De manière cohérente, la présence de Tpit n'explique pas directement le recrutement de BRG1 et Ash2 ( $R^2\approx 0.25$ ) (Figure 2.8 g).

Pour conclure, nous avons établi dans ce chapitre que les isoformes de Pax7 ne sont pas équivalentes. Les isoformes Pax7 présentant la variation GL+ dans le domaine de liaison à l'ADN PD présentent une activité pionnière altérée causant la perte du transcriptome mélanotrope activé par les isoformes Pax7 GL- dans les cellules AtT20. Les isoformes Pax7 GL+ semblent échouer à compléter la deuxième partie de l'action pionnière qui est caractérisée par une ouverture conséquente de la chromatine mesurée par ATAC, une éviction de l'histone H3 résultant à l'apparition d'un profil bimodal de la marque d'enhancer H3K4me1 et à la mise en place de la marque d'enhancer actif H3K27ac. Ceci pourrait être dû à une occupation de ces sites par Pax7 trop peu stable empêchant le recrutement effectif ou l'activation des remodeleurs de chromatine par Pax7 sur ces sites. Il nous reste à déterminer quels seraient le ou les paramètres pour expliquer l'échec de la mise en place de la seconde phase de l'action pionnière sur ces sites particuliers. On peut notamment faire l'hypothèse que ces sites présentent un environnement répressif plus important que ce soit par des niveaux de marques répressives plus élevés ou l'occupation de ces sites par des protéines répressives entrant en compétition avec la protéine Pax7 pour le recrutement ou l'activation des complexes de remodelage de la chromatine.



g) Corrélation entre la liaison des facteurs de transcription et les remodeleurs de chromatine

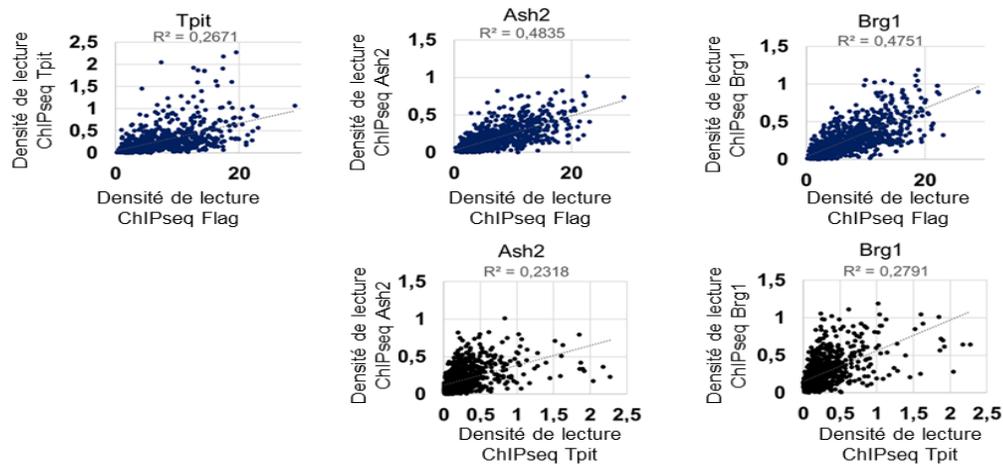


Figure 2.8 Les remodeleurs de chromatine Ash2 et BRG1 et Tpit sont diminués mais non perdus sur les enhancers GL-sensible (H3K4me1 symétrique)

a, b) Profils moyens du ChIPseq H3K4me1 (a) et Pax7 (b) pour les différents sous-groupes insensibles au GL (H3K4me1), sensibles au GL (H3K4me1 asymétrique à droite), sensibles au GL (H3K4me1 asymétrique à gauche), sensibles au GL (H3K4me1 symétrique) et sensibles au GL (liaison).

c-f) Représentation des mesures d'ATACseq (c) et des ChIPseqs Tpit (d), Ash2 (e) et Brg1 (f) ordonnée de manière croissante pour chaque marque mesurée dans les cellules exprimant Pax7 Q+GL+ et séparées en fonction des différents sous-groupes insensibles au GL (H3K4me1), sensibles au GL (H3K4me1 asymétrique à droite), sensibles au GL (H3K4me1 asymétrique à gauche), sensibles au GL (H3K4me1 symétrique) et sensibles au GL (liaison).

g) Corrélation entre les données du ChIPseqs de Pax7 Q+GL+ et les données pour Pax7 Q+GL+ du ChIPseq Tpit (en haut à gauche), du ChIPseq Ash2 (en haut au milieu), du ChIPseq BRG1 (en haut à droite). Dans les cellules exprimant Pax7 Q+GL+, corrélation des données du ChIPseq de Tpit avec les données du ChIPseq Ash2 (en bas au milieu) ou du ChIPseq BRG1 (en bas à droite)

## 2.2 Pax7 (GL-) et Pax3 (GL-) n'ont pas la même capacité pionnière dans l'hypophyse

### 2.2.1 Pax3 (GL-) ne parvient pas à ouvrir les enhancers spécifiques des mélanotropes sensibles au GL pour Pax7.

Le facteur de transcription Pax le plus similaire à Pax7 est Pax3. Dans le muscle squelettique, ces deux facteurs peuvent avoir des actions redondantes mais leurs fonctions sont clairement différentes dans ce tissu [210, 211]. Bien que Pax3 ne soit pas exprimé dans l'hypophyse, il est intéressant de comparer la capacité pionnière de ces deux facteurs au regard de leur co-expression dans d'autres tissus.

Pax3 a donc été introduit dans les cellules AtT-20 et exprimé à un niveau similaire à celui de notre Pax7 référence (Figure 2.9b). Les séquences de Pax3 et Pax7 sont très conservées dans leurs domaines de liaison à l'ADN et sont relativement conservés dans leurs régions C-terminales. Il est à noter que Pax3 n'a pas la séquence en C-terminus de Pax7 qui contient le domaine OAR (Figure 2.9 a). Avant de commencer l'évaluation de l'activité pionnière de Pax7, nous avons vérifié si Pax7 et Pax3 étaient capables de reconnaître les motifs d'ADN reconnus par Pax7 avec une expérience de retard sur gel. Nous avons constaté que Pax7 et Pax3 ont une capacité de liaison à l'ADN in vitro similaire sur tous les sites reconnus par Pax7 (Figure 2.9 b). Cependant, les cellules AtT-20 exprimant Pax3 échouent à activer le transcriptome mélanotrope de la même manière que les cellules exprimant l'isoforme Pax7 Q+GL+. De plus, lorsque le transcriptome Pax3 (GL-) est comparé au transcriptome de l'isoforme Pax7 Q+GL+, on remarque que c'est le même sous-ensemble de gènes qui est affecté par l'isoforme GL+ de Pax7 et Pax3 qui est lui-même Q+GL- (Figure 2.9 c, d). La comparaison des profils ChIPseqs de Pax7 et H3K4me1 et le profil ATACseq de Pax3 par rapport aux isoformes Pax7 montre clairement que Pax3 se comporte de la même manière que

l'isoforme Pax7 Q+GL+ bien que Pax3 soit GL- (Figure 2.9 e, f, g, h). Par conséquent, Pax3 est également déficient pour la capacité d'ouverture de la chromatine (Figure 2.9 f) malgré l'impact limité sur son recrutement étudié par CHIPseq au sous-ensemble sensible au GL (ATAC) (Figure 2.9 e) avec un profil de monomodal H3K4me1 typique des enhanceurs primés (Figure 2.9 g).

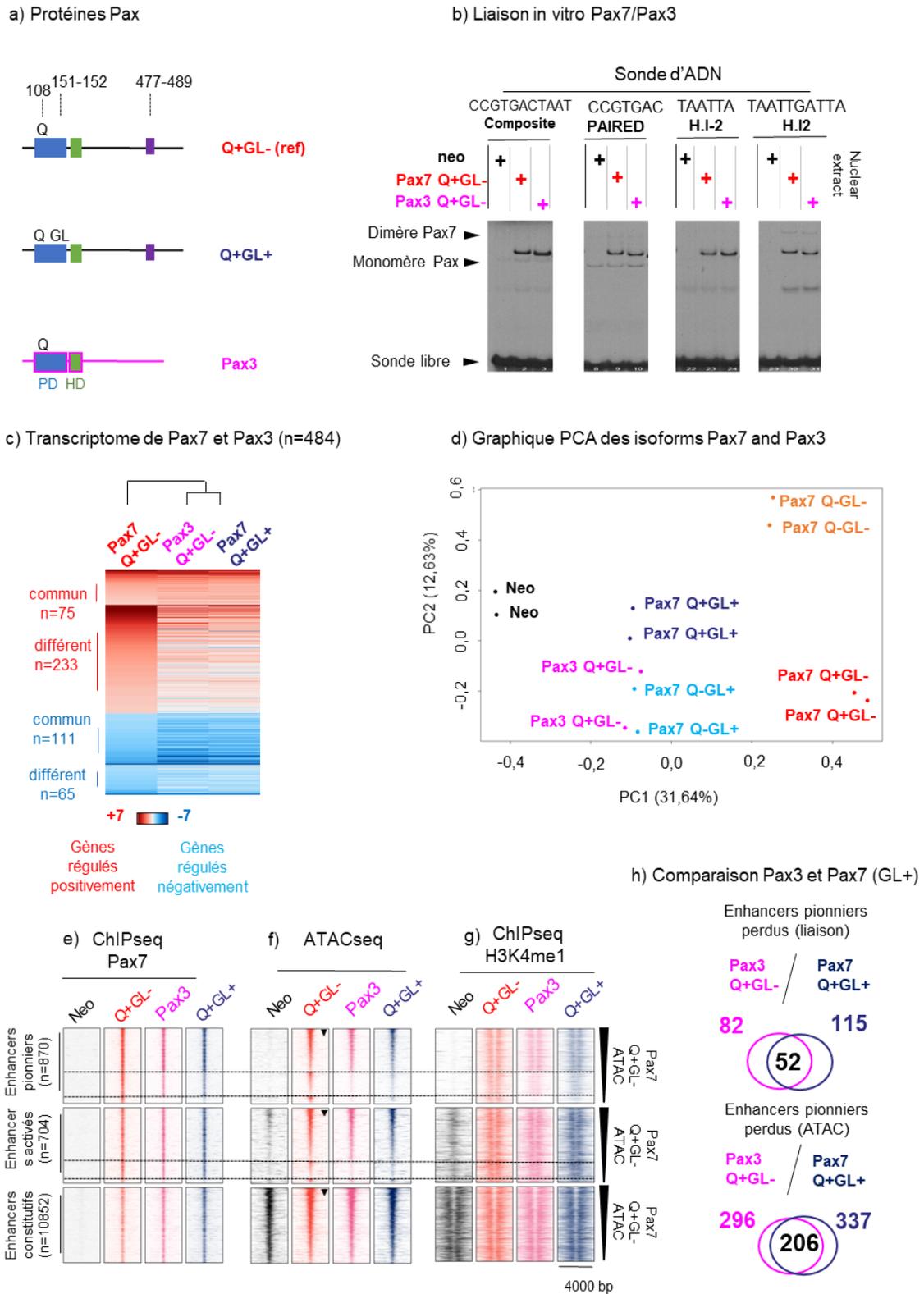


Figure 2.9 Pax3 échoue à activer le transcriptome mélanotrope comme Pax7 Q+GL+

- a) Schéma des isoformes de Pax7 Q+GL- et Q+GL+ ainsi que la protéine Pax3.
- b) Évaluation par EMSA de la liaison in vitro de Pax7 Q+GL- et Pax3 à diverses sondes d'ADN (séquences Composite, PAIRED (PD) ou HD (H.I-2, H.I2)). Les positions des monomères et des dimères Pax7, ainsi que des sondes libres, sont indiquées sur la gauche.
- c) Représentation heatmap des ARNseq des différents groupes de cellules exprimant les isoformes Pax7 Q+GL-, Pax7 Q+GL+ et Pax3.
- d) Analyse en composantes principales des analyses ARNseq des isoformes Pax7 et Pax3.
- e, f, g) Capacités pionnières des deux isoformes Pax7 exprimées dans l'hypophyse et Pax3 évaluées par ChIPseq Pax7 (e), ATACseq (f) et ChIPseq H3K4me1 (g). Ces données sont représentées en heatmaps pour les différentes catégories d'enhancers liées par l'isoforme référence Pax7 Q+GL-. Ces heatmaps sont ordonnées de façon croissante en fonction de l'ATAC de Pax7 Q+GL- et sous-divisées en sous-ensembles insensibles au GL (ATAC), sensibles au GL (ATAC) et sensible au GL (liaison) basés sur les profils ATAC et ChIPseq de Pax7 Q+GL- et Pax3.
- h) Comparaison des sites pionniers non-liés par l'isoforme Pax7 Q+GL+ et Pax3 (en haut) et des sites pionniers non ouverts par l'isoforme Pax7 Q+GL+ et Pax3 (en bas)

En résumé, nous avons montré dans ce chapitre que Pax3 (GL-) et Pax7 (GL+) sont tous deux incapables d'entraîner la deuxième partie de l'action pionnière. Le fait que Pax3 ne possède pas la variation GL mais aboutit au même phénotype que l'isoforme Pax7 Q+GL+ laisse à penser qu'une autre région de la protéine serait responsable de cette déficience. Cette région pourrait être autonome du domaine de liaison PD de la protéine Pax7 en participant par exemple au recrutement d'une autre protéine compris dans un même complexe de remodelage de la chromatine ou participant à l'ouverture de la chromatine d'une autre façon. Elle pourrait aussi agir de manière complémentaire au domaine PAIRED de Pax7 pour stabiliser une protéine nécessaire à l'ouverture de la chromatine par repliement de la protéine Pax7.

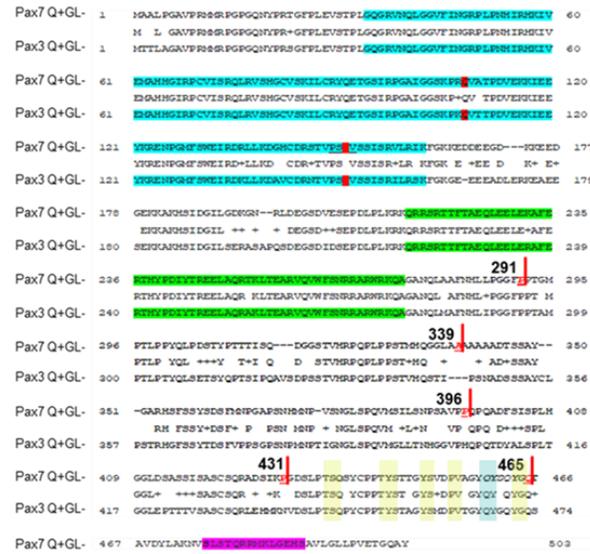
## **2-3 Étude structures fonction de la fonction pionnière de Pax7**

### **2.3.1 Les acides aminés 431 à 465 de Pax7 sont nécessaires pour son recrutement sur 60% des enhancers pionniers**

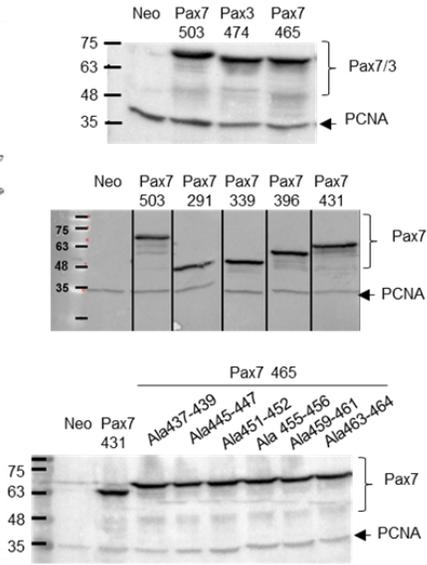
Afin de définir les domaines Pax7 impliqués dans l'activité pionnière, nous avons construit des mutants de Pax7 qui sont délétés dans son domaine C-terminal (Figure 2.10 a et 2.11 a). Le choix des régions C-terminales à supprimer a été guidé par l'alignement des séquences entre Pax7 et Pax3. Nous voulions supprimer les régions présentant le plus d'homologie entre ces deux protéines (Figure 2.10 a).

Des groupes de cellules AtT-20 exprimant les protéines Pax7 délétées à des niveaux similaires (Figure 2.10 b) ont d'abord été évalués pour leur effet sur le recrutement de Pax7 par ChIP-qPCR et sur l'expression des gènes cibles par RT-qPCR. Nous avons noté que la délétion des acides aminés 431-503 empêche le recrutement sur les enhancers pionniers mais pas sur les enhancers constitutifs. Un exemple d'enhancer pionnier et constitutif est montré à la Figure 2.10 c avec l'enhancer du gène PCSK2 et l'enhancer constitutif PDE2A. De manière concordante, l'expression du gène *Pcsk2* dépendant de l'action pionnière est perdue alors que l'expression du gène transcriptionnel *Lmcd1* est maintenue avec la construction Pax7 431. *Lmcd1* est gène activé rapidement après l'expression de Pax7 suggérant que celui-ci est sous le contrôle d'un enhancer dans une chromatine ouverte où Pax7 joue seulement son rôle de facteur de transcription (Figure 2.10 d). A l'inverse, on note que la délétion des acides aminés 465-503 comprenant le domaine OAR de la protéine Pax7 n'a pas grand effet sur la liaison de la protéine à l'enhancer du gène PCSK2 et à l'enhancer constitutif PDE2A. La perte de ces acides aminés n'affecte pas non plus l'expression du gène pionnier *Pcsk2* et du gène transcriptionnel *Lmcd1*. Nous en avons conclu que les acides aminés 431 à 465 devait avoir un impact important pour la liaison de Pax7 aux sites pionniers. L'intervalle critique entre les positions 431 et 465 est intéressant car il est très conservé en comparaison à Pax3 et parce qu'il contient plusieurs sites putatifs de phosphorylation (Figure 2.10 a). Afin d'identifier les résidus critiques dans cet intervalle pour le recrutement de Pax7 sur les sites pionniers, nous avons effectué une série de mutations remplaçant les résidus phosphorylables conservés entre Pax7 et Pax3 par des résidus neutres alanines et avons créé des groupes de cellules avec les mutants correspondants. L'expression des différents mutants dans nos groupes de cellules a été vérifié par western blot (Figure 2.10 b). Nous avons par la suite évalué si ces mutants ont toujours les fonctions pionnières par ChIP-qPCR et RT-qPCR. Un exemple de nos résultats sur l'enhancer pionnier PCSK2 et sur l'enhancer constitutif PDE2A est illustré dans la Figure 2.10 c. La mutagenèse des résidus 459-461 ou 463-464 en alanines dans la protéine Pax7 délété des acides aminés 465 à 503 a des effets délétères sur le recrutement de Pax7 au site pionnier de PCSK2 (Figure 2.10 c) mais pas sur le site constitutif PDE2A (Figure 2.10 c). Dans le même esprit, les niveaux d'ARNm de PCSK2 sont sévèrement affectés pour les mutants Pax7465Ala459-461 et Pax7465Ala463-464 (Figure 2.10 d) alors que ces mutations ont des effets marginaux sur le gène *Lmcd1* transcriptionnellement activé par Pax7 (Figure 2.10 d).

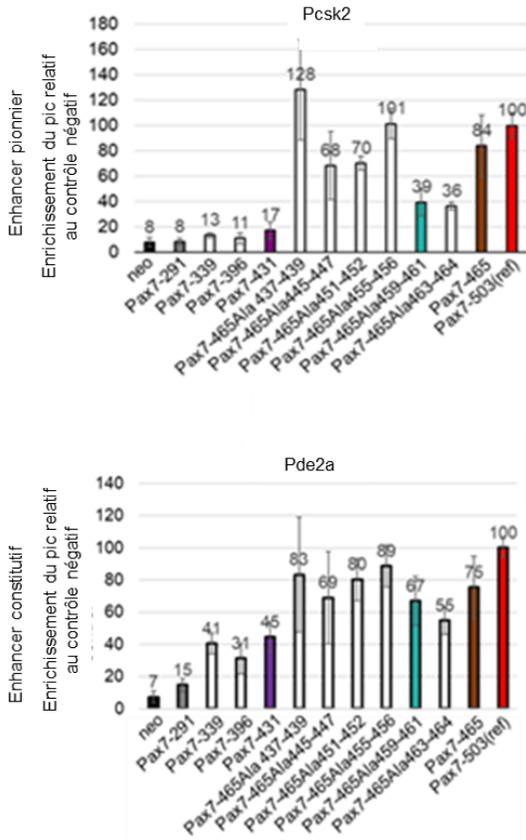
a) Comparaison des sequences Pax7/Pax3 (GL-)



b) Western blot



c) ChIP-qPCR



d) RT-qPCR

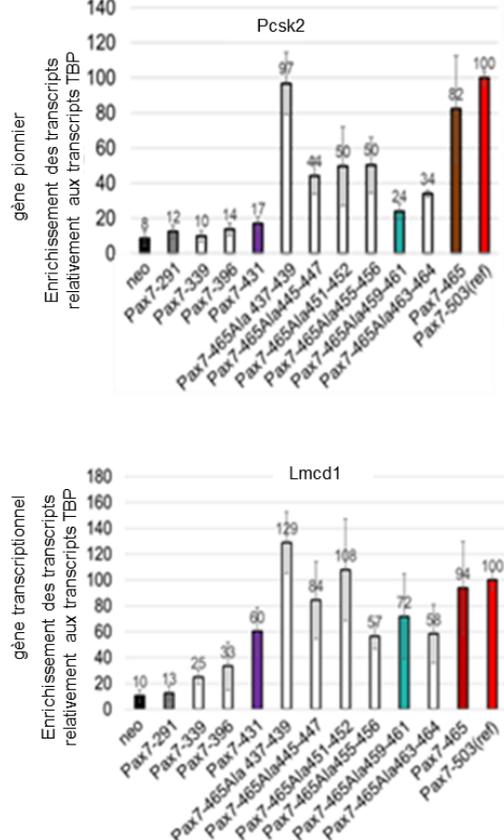


Figure 2.10 Des délétions en C-terminal de la protéine Pax7 impactent sa capacité à se lier aux sites pionniers

a) Comparaison des séquences de souris Pax3 avec Pax7. Les domaines PAIRED et HD sont mis en surbrillance respectivement en bleu et vert et le domaine OAR est en violet. Les délétions de Pax7 en C-terminal sont indiquées par un trait rouge et les mutations faites dans la protéine Pax7-465 sont indiquées par des boîtes jaunes ou vertes.

b) Western blot de la protéine Flag et de la protéine contrôle PCNA sur les groupes de cellules AtT20 exprimant les mutants de Pax7 ainsi que Pax3.

c, d) Évaluation des différents mutants Pax7 pour l'activité pionnière et transcriptionnelle lors de leur expression dans les cellules AtT-20. Le recrutement génomique des différents mutants a été évalué par ChIP-qPCR sur l'enhancer pionnier PCSK2 et l'enhancer constitutif PDE2A (c) et l'expression des gènes cibles a été évaluée par RT-qPCR au gène pionnier *Pcsk2* et au gène transcriptionnel *Lmcd1* (d).

Nous avons décidé d'approfondir notre étude sur les mutants Pax7 qui nous semblaient clés pour l'action pionnière (Figure 2.11 a). Parmi eux, Pax7-291, une construction courte contenant essentiellement les domaines de liaison à l'ADN de la protéine Pax7, Pax7-431, Pax7-465 et Pax7-465 Ala459-461.

Nous avons dans un premier temps testé la capacité de liaison à l'ADN *in vitro* de ces différentes constructions. Ces expériences sur gel à retardement ont été faites avec les différentes sondes reconnues par l'un ou l'autre des domaines PD et HD, et montrent une activité de liaison *in vitro* similaire à notre protéine Pax7 référence pour tous les mutants de délétion (Figure 2.11 b).

Par la suite, nous avons voulu étudier la fonction pionnière de ces différentes constructions sur l'ensemble du génome. Dans ce but, nous avons évalué la liaison de ces constructions par ChIPseq, leur capacité à ouvrir la chromatine par ATACseq et leur capacité à permettre le dépôt de la marque H3K4me1 par ChIPseq.

Le plus petit délétant en C-terminal, Pax7-465 ne contient plus la région comprenant le domaine OAR unique à Pax7 dans la famille Pax. Il présente une force de liaison représentant 70% de celle du Pax7-503 référence sur les sites pionniers (Figure 2.11 c). Il entraîne une accessibilité à l'ADN mesurée par l'ATACseq (Figure 2.11 d) et un dépôt de la marque bimodale H3K4me1 similaire à ceux que l'on peut observer après l'expression de la protéine Pax7-503 référence sur les sites pionniers (Figure 2.11 e). Les recrutements des facteurs Tpit, Ash2 et Brg1 sont similaires entre le Pax7-465 et le Pax7-503 (Figure 2.11 f). Nous en avons conclu que le domaine OAR n'est pas nécessaire pour la capacité pionnière.

A l'opposé, le mutant avec une délétion presque complète du C-terminus à la position 291 ne lie plus les sites pionniers et lie faiblement les sites constitutifs (Figure 2.11 c, d, e, f) démontrant que les domaines de liaison à l'ADN de Pax7 ne sont pas suffisants pour permettre une liaison même instable aux sites pionniers.

La suppression des acides aminés 431 à 465 conduit à la perte du recrutement de Pax7 sur environ 60% des sites pionniers (Figure 2.11 c), mais pas aux sites constitutifs. Sur les sites pionniers liés par cette construction, la force de liaison de Pax7-431 est diminuée de moitié par rapport à celle de Pax7-503. L'ouverture de la chromatine mesurée par ATACseq au niveau de ces sites est diminuée de plus de moitié et le pic H3K4me1 présente un profil monomodal au lieu de bimodal (Figure 2.11 d, e). De manière cohérente avec la liaison de Pax7-431, le recrutement de Tpit, Brg1 et Ash2 est réduit de 50% comparé à son recrutement avec le Pax7-503 référence. On suppose ici que c'est la diminution de la liaison stable de Pax7-431 qui entraîne une perte de stabilité des complexes de remodelage sur ces sites. Une hypothèse serait que Pax7-431, incapable de recruter de manière stable les acteurs nécessaires à l'ouverture de la chromatine, entrainerait l'amorçage de l'action pionnière avec la mise en place de l'histone H3K4me1 mais échouerait à la deuxième étape de l'action pionnière pendant laquelle nous avons l'éviction de H3, une amélioration conséquente de l'accessibilité à l'ADN mesurée par ATACseq, le changement de profil de H3K4me1 de monomodal à bimodal et la mise en place de la marque d'enhancer active H3K27ac via la protéine p300.

### **2.3.2 Le mutant Pax7465Ala459-461 lie mais n'ouvre pas la chromatine sur un sous-ensemble d'enhancers pionniers**

Pour le mutant 465Ala459-461, les conséquences sont plus complexes.

Cette protéine lie les mêmes sites que la protéine Pax7 431 (sous ensemble 1). Lorsqu'il est évalué par ChIPseq, le sous-ensemble 1 des sites pionniers présente un recrutement de Pax7 465Ala459-461 avec une force de liaison similaire à celle de Pax7-465, le signal ATAC est légèrement diminué et H3K4me1 a un profil bimodal diminué par rapport à celui de Pax7-503 (Figure 2.11 c, d, e, f). De manière étonnante, malgré une liaison similaire entre Pax7 465 et Pax7 465Ala459-461 sur ce sous-groupe, Tpit et les remodeleurs de chromatine Ash2 et Brg1 présentent une diminution de leur recrutement cohérente avec le recrutement de Pax7 465Ala459-461 contrairement au Pax7-465, qui pour un même niveau de recrutement est capable de recruter ces protéines aussi bien que Pax7-503.

Pax7 465Ala459-461, comme Pax7 431, présente une perte de liaison sur 20 % des sites pionniers (sous ensemble 3) (Figure 2.11 c, d, e, f).

Pour finir, Pax7-465Ala459-461 récupère une liaison significative sur un ensemble de sites non liés par Pax7-431 (sous ensemble 2). La force de recrutement de Pax7-465Ala459-461 sur ces sites est deux fois moins forte comparé à celle que l'on peut observer pour le Pax7 503 référence. Au niveau de ce sous ensemble, Pax7-465Ala459-461 entraîne des signaux ATACseq réduits et un pic unique et faible de H3K4me1 pour Pax7-465Ala459-461 (Figure 2.11 c, d, e, f). Les recrutements des facteurs Tpit, Ash2 et Brg1 sont cohérents avec le recrutement de Pax7 465Ala459-461.

Ici, nous pouvons faire l'hypothèse que les acides aminés 459-461 sont nécessaires à la stabilité de Pax7 sur certains sites pionniers. La stabilité du recrutement de Pax7 aux enhanceurs présents dans la chromatine fermée entraînerait des conséquences sur le recrutement stable des remodelleurs de chromatine nécessaire à l'accessibilité et/ou à l'activité de ces derniers.

Nous avons ensuite évalué si les motifs de liaison ADN sous-jacents dans chaque sous-ensemble d'enhancers pouvaient expliquer les différences de recrutement et la sensibilité de l'ouverture de la chromatine observées pour les différentes délétions de Pax7 en C-terminal. Nous n'avons trouvé aucun motif particulier qui pourrait être considéré comme spécifique pour nos différents sous-ensembles (résultats non montrés). Enfin, ni la fréquence des motifs individuels de Pax7 (Figure 2.11 g) ni le nombre total de motifs par site Pax7 (Figure 2.11 h) ne peuvent expliquer les différences de comportement observées chez nos différents mutants.

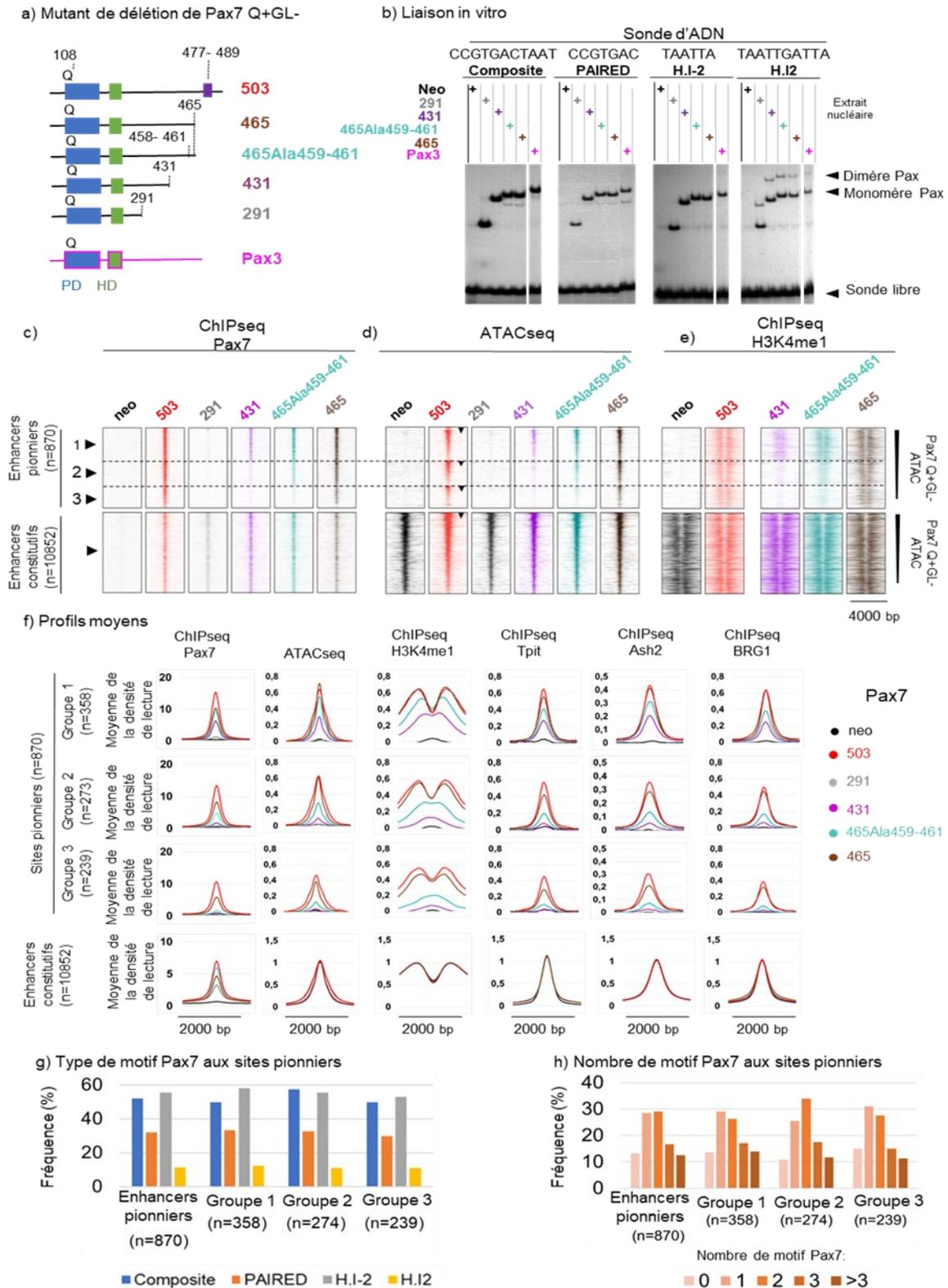
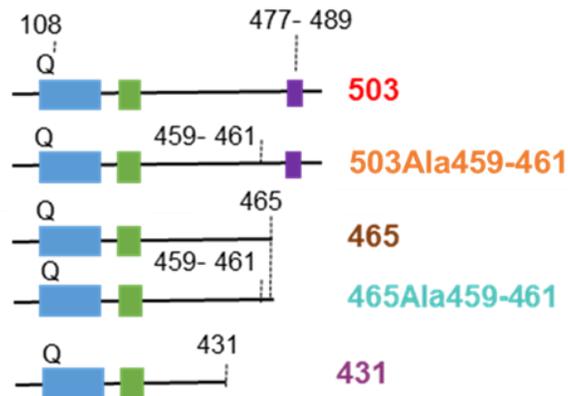


Figure 2.11 Les acides aminés 431-465 sont nécessaires pour l'activité pionnière de Pax7

- a) Représentation schématique des mutants de délétion en C-terminal de Pax7 évalués pour la capacité pionnière.
- b) Liaison de Pax7 à l'ADN in vitro des mutants de délétion en C-terminal de Pax7 évalués par gel à retardement à l'aide des sondes ADN indiquées.
- c-e) Analyse de la capacité pionnière des mutants de délétion en C-terminal évaluée par CHIPseq Pax7 (c), ATACseq (d) et CHIPseq H3K4me1 (e) représentée par heatmap. Les heatmap sont ordonnées en fonction du signal d'ATAC de Pax7 Q+GL- et séparées en fonction de la perte de recrutement graduelle observée pour les mutants Pax7-431 et Pax7-465Ala459-461.
- f) Profils moyens du CHIPseq Pax7, ATACseq, CHIPseq H3K4me1, CHIPseq Tpit, CHIPseq Ash2 et CHIPseq BRG1 pour les sous-ensembles de sites pionniers affectés par les mutants de délétion en C-terminal.
- g, h) Fréquences et nombre total de motifs d'ADN de liaison de Pax7 présents dans les différents sous-ensembles de sites pionniers définis par les mutants de délétion en C-terminal.

Enfin, nous avons voulu savoir si, dans le contexte de la protéine Pax7 entière non délétée des acides aminés 465-503, la mutation Ala459-461 était nécessaire à l'action pionnière. Après avoir créé un groupe de cellules exprimant ce mutant de manière similaire au groupe de cellules exprimant notre Pax7 référence (Figure 2.12 a, résultats non montrés), nous avons testé la capacité de cette construction à lier des enhanceurs pionniers (MYBPC1) et constitutifs (TCFEB) par CHIP-qPCR. Nous avons constaté que, dans le contexte de la protéine entière, la mutation des acides aminés 459-461 n'a pas ou peu d'effet sur le recrutement de Pax7 au site pionnier MYBPC1 (Figure 2.12 b) laissant supposer une redondance de fonction entre la région comprenant les acides aminés 459-461 et la région 465 à 503 de la protéine Pax7.

## a) Mutants et délétants de la protéine Pax7



## b) ChIP-qPCR

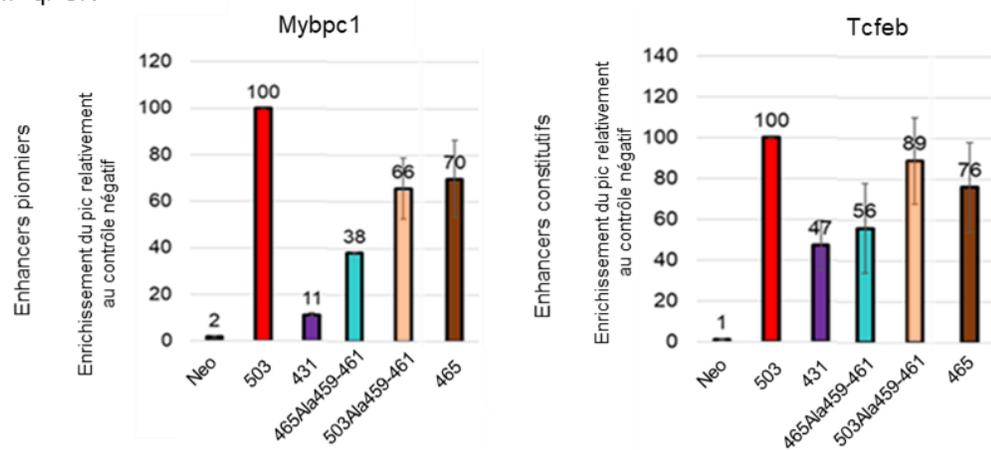


Figure 2.12 Les acides aminés 459 à 461 avec un C-terminal complet n'affectent pas la liaison de Pax7 aux sites pionniers

a) Représentation schématique des mutants présentant des délétions en C-terminal de Pax7 évalués pour la capacité pionnière.

b) Évaluation de différents mutants Pax7 pour l'activité pionnière et transcriptionnelle lors de leur expression dans les cellules AtT-20. Le recrutement génomique des différents mutants a été évalué par ChIP-qPCR sur l'enhancer pionnier MYBPC1 et l'enhancer constitutif TCFEb.

Notre étude structure fonction de Pax7 a permis d'établir que les domaines de liaison à l'ADN de Pax7 ne sont pas suffisants pour le recrutement de Pax7 aux sites pionniers et que le domaine OAR n'est pas nécessaire aux

différentes étapes de l'action pionnière de Pax7. Elle a aussi permis d'établir que la région C-terminale de Pax7 compris entre les acides aminés 431 et 465 est nécessaire au recrutement de Pax7 sur plus de 50% des sites pionniers (Figure 2.11).

Enfin, on a montré que la protéine Pax7-465 Ala459-461 présente un recrutement plus faible par ChIPseq sur certains sites pionniers. Sur ces sites, on observe une réduction du signal ATACseq et un pic monomodal de H3K4me1. Les caractéristiques chromatiniennes de ces sites sont similaires à celles observées sur les sites pionniers affectés par l'insertion de GL dans l'isoforme Pax7 Q+GL+ (Figure 2.8 et Figure 2.11).

## **2-4 Des régions des parties N et C-terminales de Pax7 sont nécessaires pour l'ouverture de la chromatine des enhancers spécifiques des mélanotropes**

### **2.4.1 Un même sous-ensemble d'enhancers est affecté par des modifications en N et C-terminal de la protéine Pax7**

La comparaison des sous-ensembles de sites pionniers sensibles en liaison pour le Pax7-431 ou en ouverture pour le Pax7-465Ala459-461 avec ceux qui sont sensibles en ouverture due à l'insertion du GL dans le domaine PAIRED de Pax7 montrent qu'ils sont majoritairement identiques (Figure 2.13 a, b, c). Bien que Pax7-465Ala459-461 ait une liaison diminuée en comparaison avec Pax7 Q+GL+ (Figure 2.13 d, e), les profils ATAC et H3K4me1 entraînés par l'expression de l'isoforme Pax7 Q+GL+ et de la construction Pax7-465Ala459-461 se superposent de manière quasi-parfaite (Figure 2.13 d, g). Ces deux constructions de Pax7 présentent des recrutements du facteur Ash2 et similaires sur les sites GL-sensibles (H3K4me1). Les recrutements de Tpit et BRG1 sur ces mêmes sites sont légèrement plus faibles pour Pax7-465Ala459-461 que pour l'isoforme Pax7 Q+GL+(Figure 2.13 g, h, i).

Nous concluons donc que les enhancers mélanotropes qui présentent une ouverture de chromatine altérée à la suite de l'insertion GL+ dans le domaine PD dépendent également des séquences C-terminales comprises entre les acides aminés 459 et 503. Le fait qu'un même sous-ensemble d'enhancers est dépendant de ces deux caractéristiques structurales de Pax7 pour l'action pionnière, suggère que la chromatine sur ces enhancers présente des propriétés uniques et que le C et N-terminale de Pax7 sont nécessaires à sa décompaction.

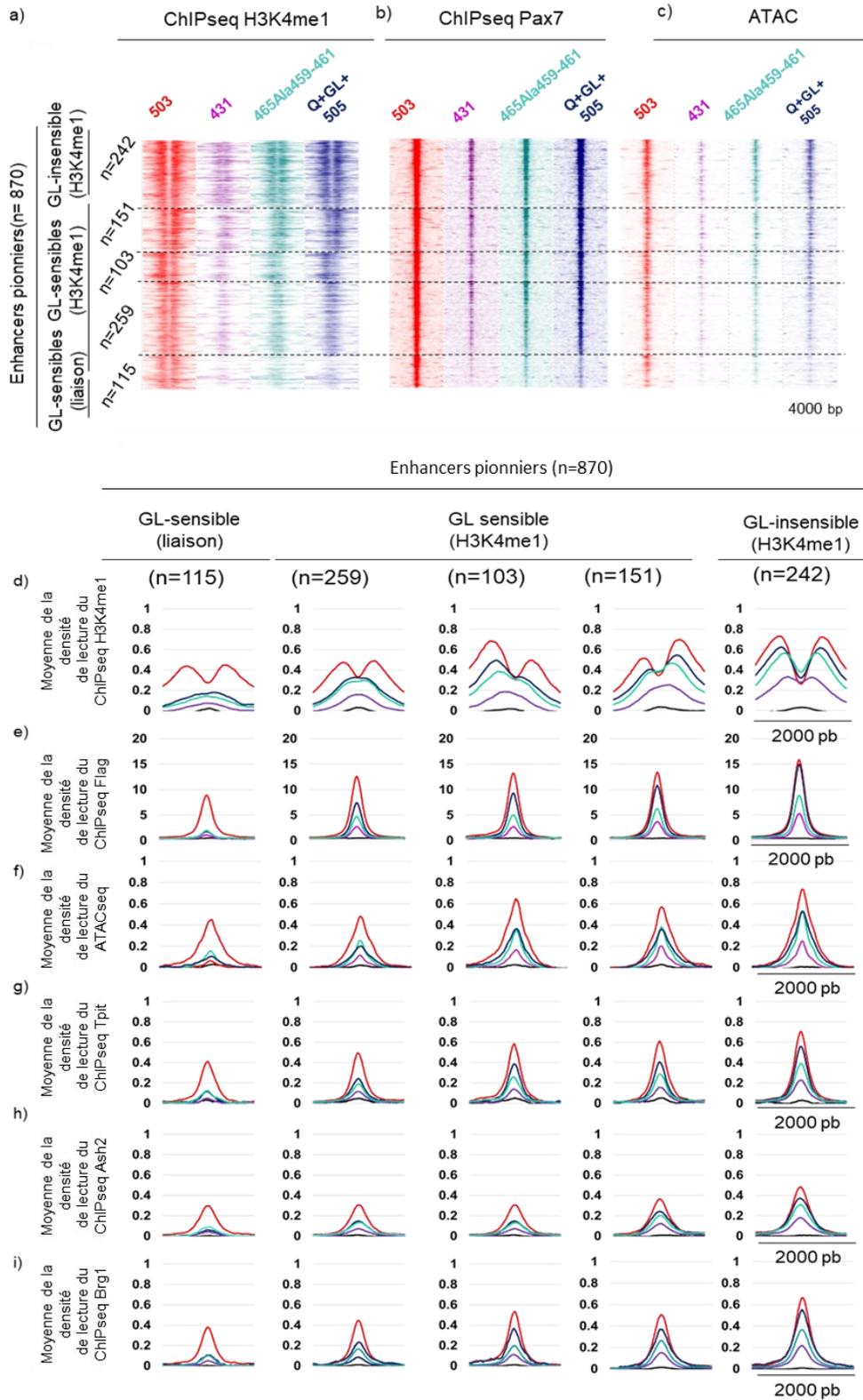


Figure 2.13 Des variations en N et C-terminal ont des effets similaires sur l'activation du transcriptome du mélanotrope

a-c) Comparaison des effets délétères des délétions de Pax7 en C-terminal et de l'insertion de GL dans le domaine de liaison à l'ADN PAIRED de Pax7 sur les sites pionniers. Pour chacune des constructions Pax7, les ChIPseqs de la marque H3K4me1 (a), de la protéine Pax7 (b) et l'ATACseq (c) sur les sites pionniers ont été représentés par heatmaps. Ces heatmaps sont ordonnées de façon décroissante en fonction du ChIPseq de Pax7 Q+GL- (503) et divisées en sous-ensembles insensibles au GL (H3K4me1), sensibles au GL (H3K4me1 asymétrique à droite), sensibles au GL (H3K4me1 asymétrique à gauche), sensibles au GL (H3K4me1 symétrique) et sensibles au GL (liaison) basés sur les profils ChIPseq H3K4me1 de Pax7 Q+GL- et Pax7 Q+GL+.

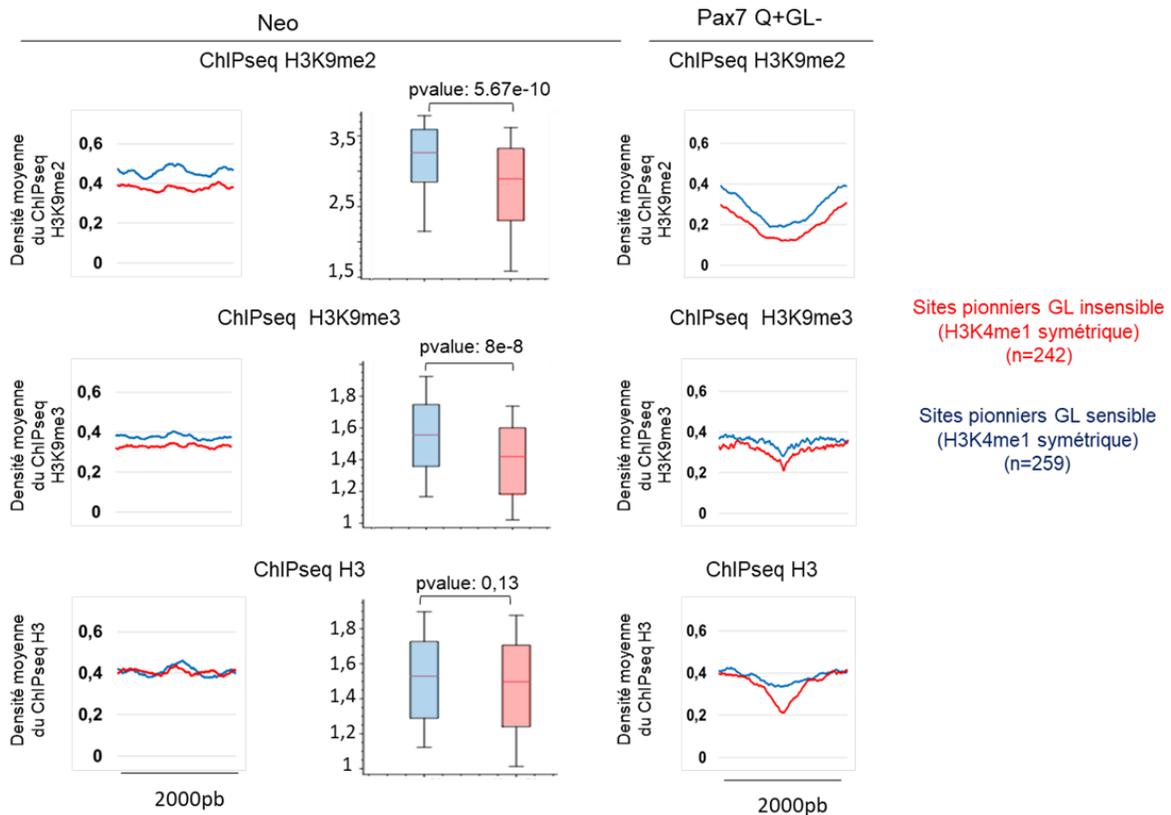
d-i) Profils moyens du ChIPseq H3K4me1 (d), Pax7 (e), ATAC (f), Tpit (g), Ash2(h), Brg1(i) pour les différents sous-groupes insensibles au GL (H3K4me1), sensibles au GL (H3K4me1 asymétrique à droite), sensibles au GL (H3K4me1 asymétrique à gauche), sensibles au GL (H3K4me1 symétrique) et sensibles au GL (liaison).

#### 2.4.2 Des enhanceurs mélanotropes aux caractéristiques particulières

N'ayant pas trouvé de différences marquées (motifs de novo, distribution des motifs Pax7) entre les sites pionniers GL-insensibles et GL-sensibles (H3K4me1) expliquant pourquoi la deuxième phase de l'action pionnière semble être compromise pour les constructions Pax7 affectées en N et C-terminal, nous nous sommes intéressés à la première phase de l'action pionnière représentée majoritairement par la perte H3K9me2 découlant de la liaison de Pax7 à la chromatine fermée [Gouhier et al.,2022, en préparation]. En étudiant les marques répressives H3K9me2, H3K9me3 et l'histone H3 présentent au niveau des sites pionniers avant et après l'expression de la protéine Pax7 Q+GL-, on remarque que les sites pionniers GL-sensibles (H3K4me1) présentent plus de H3K9me2 et de H3K9me3 que les sites GL-insensibles (H3K4me1) avant et après l'expression du Pax7 référence Q+GL-. Les sites GL-sensibles présentent aussi une diminution moindre du niveau d'histone H3 comparés aux sites pionniers GL-insensibles après l'expression Pax7 Q+GL-, et ceci, malgré des niveaux de l'histone H3 similaires entre les deux types de sites avant son expression (Figure 2.14 a). Nous supposons qu'au niveau des sites GL-sensibles et GL-insensibles, le paysage chromatinien natif serait différent. Nos résultats supportent que les sites GL-sensibles présenteraient une chromatine plus compactée avec des niveaux de marques répressives plus élevés. La marque H3K9me2 permet de maintenir les régions régulatrices du génome dans la lamine où elles sont inaccessibles aux facteurs de remodelage de la chromatine. Nous nous sommes demandé si la perte d'ouverture au niveau des sites pionniers GL-sensibles pouvait être due à une incapacité des protéines Pax7 Q+GL+ et Pax7-465Ala459-461 à enlever la marque répressive H3K9me2 sur ces sites. Dans ce contexte, ces sites ne seraient donc plus accessibles aux facteurs de remodelage de la chromatine. Pour répondre à cette interrogation, nous avons évalué par ChIP-qPCR si Pax7 Q+GL+ et Pax7-465Ala459-461 avaient la capacité d'entraîner la perte de la marque H3K9me2 dans

les sous-ensembles pionniers GL-insensibles (H3K4me1) et GL-sensibles (H3K4me1) de manière similaire à Pax7 Q+GL-. Nous pouvons voir que Pax7 Q+GL+ et Pax7-465Ala459-461 ont toutes deux la capacité d'entraîner la perte de la marque H3K9me2 (Figure 2.14 b) nous laissant supposer que ce n'est pas l'absence de cette perte d'H3K9me2 qui empêche l'ouverture de la chromatine à ces sites mais un autre paramètre qu'il reste à découvrir.

a) Profil moyen des marques repressives avant et après l'expression de Pax7 Q+GL-



b) ChIP-qPCR H3K9me2 aux sites pionniers GL insensibles et sensibles (H3K4me1 symétrique)

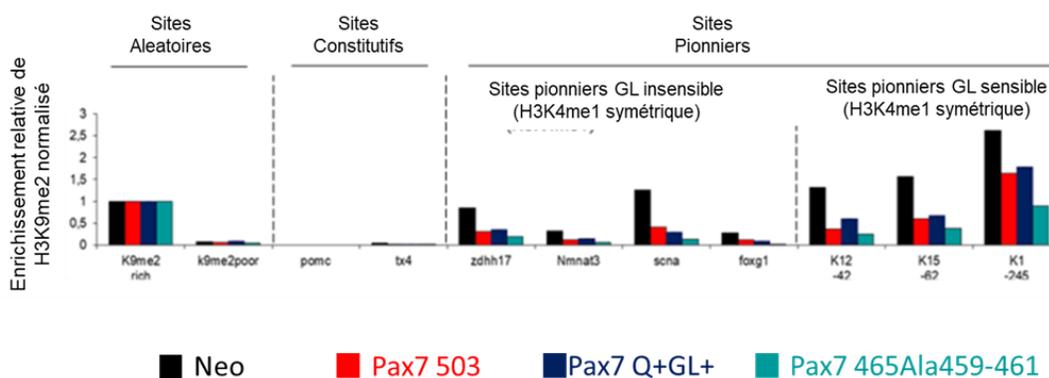


Figure 2.14 Les enhancers pionniers spécifiques des mélanotopes présentent des niveaux de H3K9me2 plus élevés que les autres sites pionniers

a) Représentations des niveaux des marques répressives H3K9me2 (au haut) et H3K9me3 (au milieu) et de l'histone H3 (en bas) sur les sites GL-sensibles H3K4me1-symétriques représentés en bleu et les sites GL-insensibles H3K4me1-symétriques représentés en rouge avant (à gauche) et après l'expression du Pax7 référence Q+GL- (à droite).

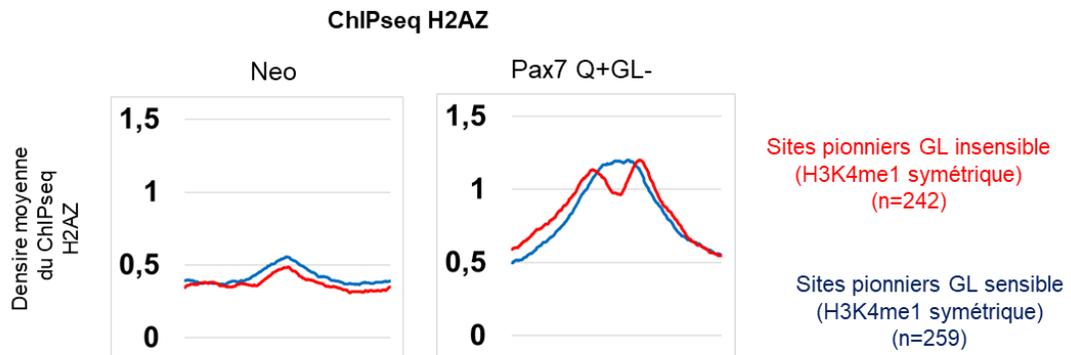
b) ChIP-qPCRs de la marque répressive H3K9me2 dans les cellules Néo n'exprimant pas Pax7 ainsi que dans les cellules exprimant les isoformes Pax7 hypophysaires et le mutant Pax7-465Ala459-461. Ces ChIPs ont été faites sur des sites aléatoires non liés par Pax7 présentant la marque H3K9me2, des sites constitutifs liés par Pax7 ne présentant pas la marque H3K9me2 ainsi que sur des sites pionniers GL-sensibles et des sites pionniers GL-insensibles (H3K4me1 symétrique).

Par la suite, nous avons aussi étudié le comportement de la marque H2AZ. L'incorporation de cette marque dans la chromatine compacte permet une déstabilisation des nucléosomes et donc une décompaction de la chromatine [213]. Nous avons constaté qu'avant l'expression de Pax7 Q+GL-, le niveau de H2AZ est faible et quantitativement similaire entre les sites pionniers GL-insensible (H3K4me1) et GL-sensible (H3K4me1). Cependant, après l'expression de Pax7 Q+GL-, le profil d'H2AZ arbore un profil monomodal dans les sites GL-sensibles alors que son profil est bimodal dans les sites GL-insensibles (Figure 2.15 a).

Le profil monomodal de la marque H2AZ associé à un niveau central de la marque H3K4me1 plus élevé (Figure 2.15 a,b) et une diminution de l'histone H3 plus faible (Figure 2.14 a) dans les sites pionniers GL-sensibles après l'expression de Pax7 Q+GL- laissent suggérer que, au niveau de ces sites spécifiques nécessaires au transcriptome mélanotrope, la décompaction de la chromatine suivrait un ou des mécanismes différents de celui ou ceux qui se produisent aux niveaux des sites pionniers GL-insensibles.

Le maintien d'un nucléosome mobile aux sites GL-sensibles pourrait permettre une régulation transcriptionnelle plus complexe de ces sites importants pour la destinée cellulaire des mélanotropes.

a) Profil moyen de la marque H2AZ avant et après l'expression de Pax7 Q+GL-



b) Profil moyen de la marque H3K4me1 avant et après l'expression de Pax7 Q+GL-

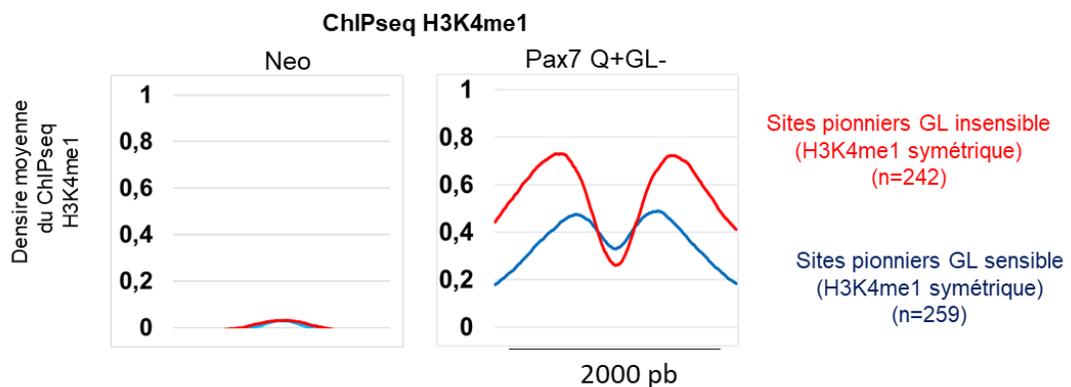


Figure 2.15 Les enhancers pionniers spécifiques des mélanotropes gardent un nucléosome central comprenant l'histone H2AZ.

a) Profil moyen de la marque H2AZ (ChIPseq) sur les sites GL-sensibles H3K4me1-symétrique représentés en bleu et les sites GL-insensibles H3K4me1-symétrique représentés en rouge avant (à gauche) et après l'expression du Pax7 référence Q+GL- (à droite)

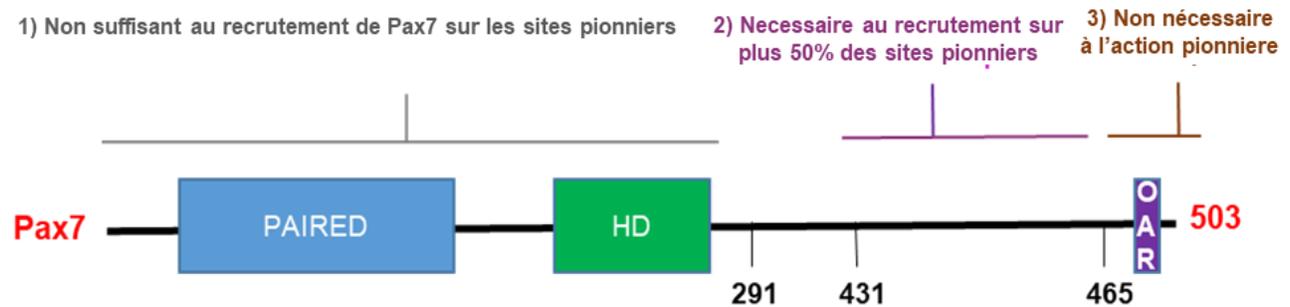
b) Profil moyen de la marque H3K4me1 (ChIPseq) sur les sites GL-sensibles H3K4me1-symétrique représentés en bleu et les sites les sites GL-insensibles H3K4me1-symétrique représentés en rouge avant (à gauche) et après l'expression du Pax7 référence Q+GL- (à droite)

## 2-5 Conclusions

En conclusion, nous avons établi dans cette étude une cartographie des domaines Pax7 nécessaires à l'action pionnière. Notre étude a permis dans un premier temps de caractériser les grands domaines de la protéine Pax7 et leur impact sur l'action pionnière de Pax7 dans l'hypophyse. Nous avons montré que les domaines de liaison à l'ADN de Pax7 ne sont pas suffisants pour le recrutement aux sites pionniers dans l'hétérochromatine bien qu'ils permettent un faible recrutement aux enhancers constitutifs dans l'euchromatine. Les acides aminés 465 à 503 en C-terminale de Pax7 comprenant le domaine OAR ne sont pas nécessaires à l'action pionnière de Pax7. Enfin, les acides aminés 431 à 465 sont nécessaires à la liaison de Pax7 sur plus de 60 % des sites pionniers (Figure 2.16 a).

Dans un second temps, nous avons démontré que l'isoforme Q+GL+ est capable d'entraîner la première étape de l'action pionnière (perte de la marque H3K9me2 et mise en place de la marque H3K4me1 avec un profil monomodal) mais incapable de mettre en place la deuxième phase de l'action pionnière de Pax7 caractérisée par une ouverture conséquente de la chromatine, l'éviction du nucléosome H3 et la mise en place de la marque d'enhancer actif H3K4me1 avec un profil bimodal sur les enhancers spécifiques au transcriptome mélanotrope. Le même phénotype est obtenu pour la construction Pax7 délété en C-terminal Pax7-465Ala459-461 nous permettant de suggérer que les domaines N et C-terminal de la protéine Pax7 agissent de façon complémentaire pour les étapes nécessaires à la deuxième phase de l'action pionnière de Pax7 Q+GL- sur des sites pionniers spécifiques nécessaire à l'activation du transcriptome mélanotrope (Figure 2.16 b).

## a) Role des grands domaines de la protéine Pax7



## b) Des changement subtils de Pax7 en N-and C-terminal subtil affectent l'ouverture de la chromatine au niveau des enhancers melanotropes

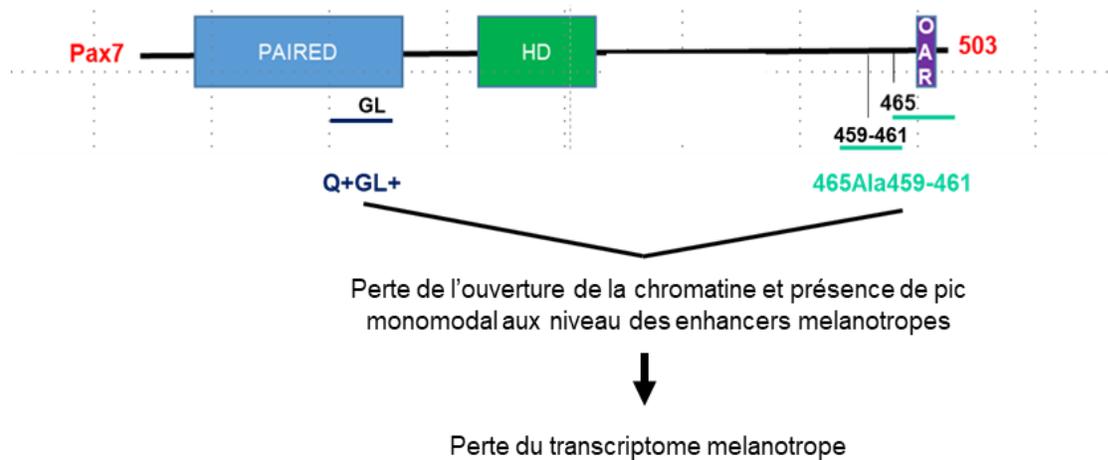


Figure 2.16 Conclusions

- Schéma représentant le rôle des grands domaines de Pax7 sur son action pionnière.
- Schéma représentant les conclusions sur l'impact de l'isoforme Pax7 Q+GL+ et du mutant Pax7-465Ala459-461 sur l'action pionnière de Pax7

## **2-6-Des mécanismes différents pour expliquer la perte d'ouverture observée dans l'isoforme Pax7 (GL+) et la construction Pax7465Ala459-461**

Nous avons voulu savoir si le même mécanisme était affecté pour expliquer la perte d'ouverture de la chromatine observée lors de l'expression de l'isoforme Pax7 Q+GL+ et de la construction Pax7465Ala459-461.

Nous avons émis l'hypothèse que la perte d'ouverture de la chromatine est due à une inaptitude de ces protéines Pax7 à recruter des protéines nécessaires à l'ouverture de la chromatine ou à la stabilité de leurs liaisons sur les sites pionniers.

Pour répondre à cette question, nous avons fait des RIMES (Rapid immunoprecipitation mass spectrometry of endogenous proteins) des isoformes Pax7 Q+GL-, Pax7 Q+GL+ et du mutant Pax7465Ala459-461. L'expérience de RIME permet d'identifier les interacteurs d'une protéine d'intérêt. Lors de nos expériences de RIME, notre protéine cible Pax7 et son environnement protéique ont été immunoprécipités sur chacun de ces sites de recrutement sur le génome. Par la suite, cet immunoprécipitat a été traité en spectrophotométrie de masse pour identifier les protéines dans l'environnement proche de Pax7 sur le génome. Cette expérience est à prendre avec précaution car Pax7 est généralement plus présent dans la chromatine ouverte sur le génome que sur la chromatine fermée. Les sites pionniers représentant environ 10 % des sites de recrutement de Pax7. Les protéines associées à l'action pionnière ne seront donc pas forcément les protéines les plus abondantes de notre immunoprécipitat. De plus, un interacteur de Pax7 pourrait être important à la fois sur les sites pionniers et sur les sites transcriptionnels de Pax7.

Cela étant dit, nous avons donc porté notre attention sur les protéines gagnées ou perdues dans l'isoforme Pax7 Q+GL+ et le mutant Pax7465Ala459-461 par rapport à notre Pax7 Q+GL- référence. Nous voulions savoir si l'insertion du GL en N terminal et la mutation des acides aminés 459 à 461 dans le Pax7-465 affectaient le recrutement de protéines similaires et si la présence de ces protéines était modulée dans le même sens par ces variations.

Nous constatons malheureusement que cela n'est pas le cas laissant supposer que la perte d'ouverture de la chromatine observé chez ces deux formes de Pax7 est causée par des mécanismes différents (Figure 2.17a et 2.17b).

Plus précisément, nous observons que, comparé à la protéine référence Pax7 Q+GL-, l'isoforme Pax7 Q+GL+ et le mutant Pax7465Ala459-461 présentent des variations significatives (de plus de deux fois avec une p-value inférieur à 0.05) de protéines impliquées dans la régulation de la transcription et dans l'organisation de la chromatine (Figure 2.17 a).

On retrouve, dans les protéines affectées par les variations et mutations de Pax7, des protéines liées à la répression transcriptionnelle, à l'hétérochromatine facultative et au complexe SWI-SNF. Parmi les protéines identiques qui se retrouvent affectées dans l'isoforme Pax7 Q+GL+ et le mutant Pax7465Ala459-461, on retrouve les protéines Ncor1 et Trim33. La protéine Ncor1 fait partie d'un complexe répressif associé à la lamine et peut jouer un rôle dans la compartimentation de l'hétérochromatine au niveau lamine à la périphérie du noyau et de l'euchromatine à l'intérieur du noyau ou cette dernière serait accessible aux facteurs protéiques de remodelage et de transcription [214]. Cette protéine est aussi associée au HDACs, protéine dont les fonctions désacétylases peuvent empêcher la déstabilisation du nucléosome H3 et par conséquent empêcher l'ouverture de la chromatine. Cependant si cette hypothèse est plausible pour l'isoforme Q+GL+ (Figure 2.17 a) qui recrute ses protéines en plus grande quantité comparé à la protéine Pax7 Q+GL- référence, il est à noter que Pax7-465Ala459-461 (Figure 2.17 b) perd ces protéines suggérant qu'un autre mécanisme pour la perte de l'ouverture de la chromatine se passe pour ce mutant.

Parmi les protéines perdues pour ce mutant, nous retrouvons des protéines impliquées dans la stabilité des protéines pionnières aux sites de chromatine fermée comme CBFA2T2 et PRDM6 (un homologue de la protéine PRDM14) [215]. Nous retrouvons aussi le facteur RNF20, une protéine responsable de l'ubiquitination de l'histone H2B. Lorsque cette protéine est perdue, cela entraîne la perte de marque d'histone active, H2Bub1, la perte de la marque d'histone H3K4me1, et la perte de l'hyperacétylation de H3, H4, et H2A.Z [216]. La perte de RNF20 pourrait être un mécanisme pour expliquer la perte de l'ouverture observée lorsque la construction Pax7-465Ala459-461 est exprimée.



b) Protéines perdues par Pax7465Ala459-461 comparé aux Pax7 Q+GL- référence

Term	RT	%	P-Value
negative regulation of transcription, DNA-templated	10	27,8	8,2E-7
chromatin organization	7	19,4	1,4E-5
RNA splicing	6	16,7	8,0E-5

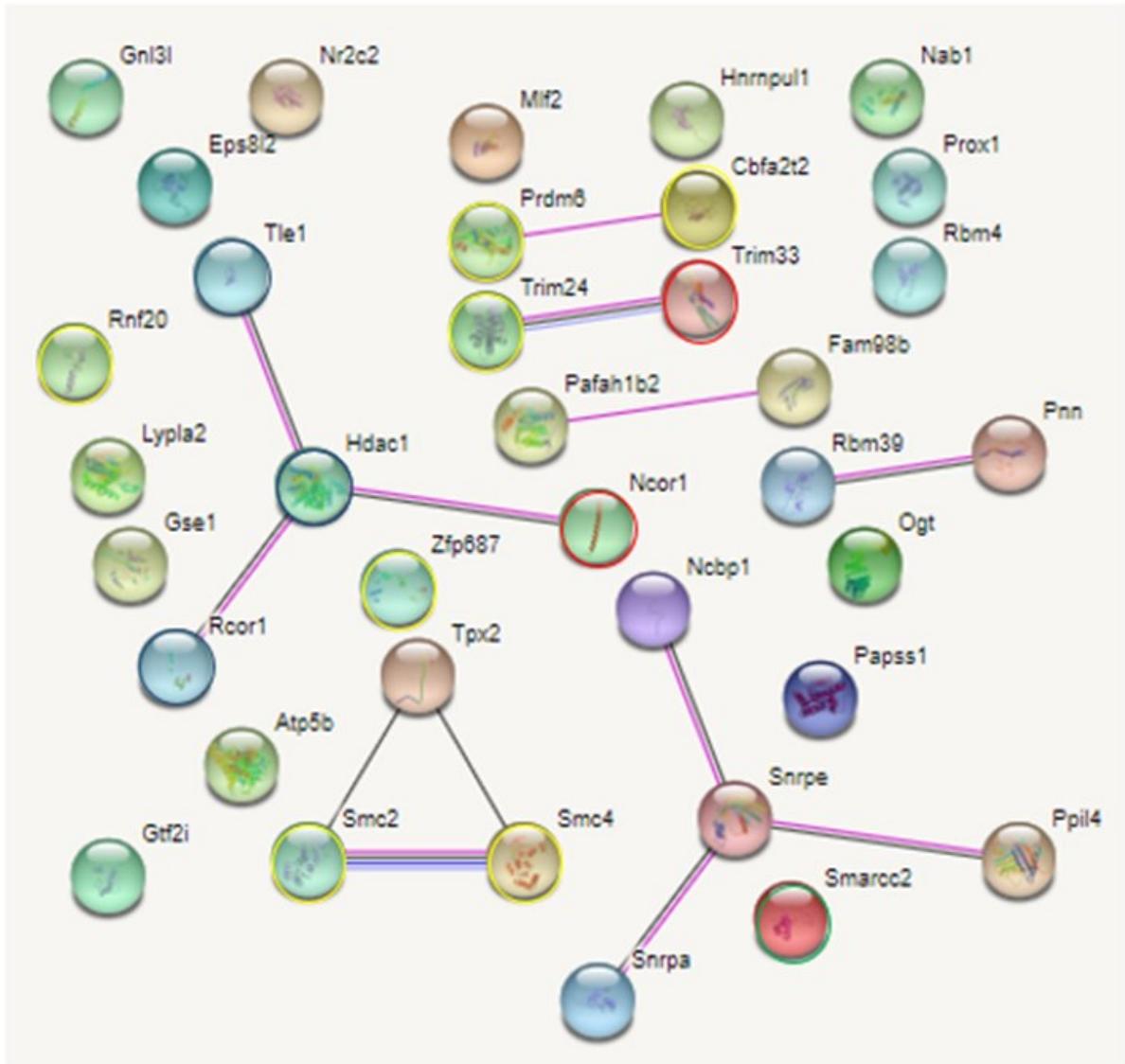


Figure 2.17 Possibles cofacteurs impliqués dans la perte de l'ouverture de la chromatine pour l'isoforme Pax7 Q+GL+ et le mutant Pax7-465Ala459-461

- a) Représentation avec le logiciel String des protéines exprimées plus de deux fois avec une pvalue de 0.05 entre l'isoforme Pax7 Q+GL+ et Pax7 Q+GL-.
- b) Représentation avec le logiciel String des protéines exprimées moins de deux fois avec une pvalue de 0.05 entre le mutant Pax7-465Ala459-461 et Pax7 Q+GL-.

## Chapitre 3 : Matériels et Méthodes

### Culture cellulaire

Des cellules AtT-20/D16v-F2 ont été cultivées dans le milieu modifié d'Eagle de Dulbecco, complété par du sérum bovin fœtal à 10 % et des antibiotiques (pénicilline/streptomycine). Pour générer des populations de cellules AtT-20 Pax7 transgéniques stables, les rétrovirus ont été emballés à l'aide de la lignée cellulaire d'emballage rétroviral Platinum-E (Cell Biolabs, catalogue #RV-101) et les infections ont été exécutées comme décrit [217]. La sélection des populations de cellules infectées par le rétrovirus a été réalisée avec 400 µg/mL de Geneticin (Gibco, #11811-031). Des colonies résistantes ont été regroupées pour générer des populations de centaines de colonies indépendantes.

### Essais de luciférase

Les cellules AtT20 ont été cultivées dans des plaques de 12 puits (4 x 10<sup>5</sup> cellules par puits) et les transfections réalisées le lendemain à l'aide de PEI 25K (Polysciences #23966-1) en triplicat. L'ADN transfecté (total de 2 µg dans 150 µL de DMEM) a été ajouté à la solution de PEI 25K (4 µL par 150 µL de DMEM) et incubé à température ambiante pendant 30 minutes avant d'ajouter 100 µL du mélange à chaque puits. Le milieu a été changé 24 heures après la transfection et l'activité de luciférase a été évaluée 48 heures après la transfection. On a retiré le milieu des cellules transfectées et ajouté 200 µL de tampon de lyse (Tris 100 mM, NP40 0,5 % et DTT 0,001M) à chaque puits. Après 10 minutes d'agitation vigoureuse, 100 µL de surnageant de lysat cellulaire ont été utilisés pour l'analyse à l'aide du luminomètre GloMax Navigator (Promega) en utilisant le protocole suivant (injection de 100 µL, vitesse de 100 µL/s, temps d'intégration de 10 s). La signification statistique a été évaluée au moyen du test de combinaisons multiples ANOVA et Tukey.

### ChIPseq, ATACseq

Les expériences de ChIPseq ont été réalisées dans des cellules AtT-20 comme décrit précédemment [167, 218]. Les ATACseq ont été réalisés comme décrit précédemment [167]. Les bibliothèques et Flow cells ont été préparées par le plateau de biologie moléculaire de l'IRCM suivant les recommandations d'Illumina (Illumina, San Diego, Californie). Les ChIP ont été séquencées sur le séquenceur Illumina Hi-Seq 2000. Le tableau 1 fournit des détails sur les anticorps utilisés pour les ChIPs. Les profils moyens de ChIP et d'ATACseq de Pax7 ont été générés par la fonction de signal moyen d'Easeq (<http://easeq.net>) et les données ont été exportées vers Microsoft Excel pour créer des graphiques. Pour visualiser les heatmaps de ChIPseq et ATACseq, nous avons utilisé l'outil Integrative Genome Viewer de easeq [219] ou un programme personnalisé en Python [133]. Toutes les ChIPs et les ATACs à

l'exception de la ChIP Flag ont été normalisés en fonction de sites positifs constitutifs et de sites négatifs aléatoires.

Nom	Nom standard	Origine	Référence	Usage
Flag M2	FLAG M2	Drouin Lab, IRCM	N/A	ChIP, Western
Tpit	TPIT	Drouin Lab, IRCM	N/A	ChIP
Ash2	ASH2L	Bethyl Laboratories	A300-489A	ChIP
Brg1	SMARCA4	Abcam	ab110641	ChIP
H3K4me1	H3K4me1	Abcam	ab8895	ChIP
H3K9me2	H3K9me2	Abcam	ab1220	ChIP

Tableau 1- Anticorps utilisés en ChIPs et Western Blots

### Extraits nucléaires

Pour préparer des extraits nucléaires, les cellules AtT-20 d'une boîte de Pétri de 100 mm ont été lavées et recueillies dans le PBS. Les cellules ont été remises en suspension dans 800  $\mu$ L d'un tampon froid à faible teneur en sel (10 mM HEPES pH 7,9, 10 mM KCl, 0,1 mM EDTA, 0,1 mM EGTA, 1 mM DTT, 0,5 mM PMSF, 0,6  $\mu$ M aprotinine, 0,6  $\mu$ M leupeptin et 2  $\mu$ M pepstatine A). Après 15 min d'incubation sur glace, 100  $\mu$ L NP40 10 % ont été ajoutés. Les fractions cytosoliques ont été éliminées, et les protéines nucléaires extraites dans 100  $\mu$ L de tampon froid à haute teneur en sel contenant 20 mM HEPES pH 7,9, 400 mM KCl, 1 mM EDTA, 1 mM EGTA, 1 mM DTT, 1 mM PMSF, 0,6  $\mu$ M aprotinine, 0,6  $\mu$ M leupeptin et 2  $\mu$ M pepstatine A. La concentration en protéines a été quantifiée par essai Bradford.

### Marquage des sondes

Les séquences des sondes d'ADN à double brin ont été décrites dans Pelletier et al [133]. 1  $\mu$ M de chaque brin ont été associés à haute température dans le tampon TEN (10 mM Tris pH 8,0, 1 mM EDTA et 50 mM NaCl). Le marquage a été effectué à température ambiante pendant 20 minutes dans le tampon suivant : 50 mM Tris pH 7,5, 10 mM MgCl<sub>2</sub>, 1 mM DTT et 0,5 mg/mL BSA complétés par 0,25  $\mu$ M dGTP, dTP, 0,2  $\mu$ M  $\alpha$ -<sup>32</sup>P-dCTP (3 000 Ci/mmol, 10 mCi/mL, Perkin Elmer) et 1 unité de ADN polymérase I 90 Large (Klenow) fragment d'E. coli (ONÉ, no de catalogue M0210M). Les sondes ont ensuite été purifiées sur la colonne de chromatographie Tris Micro Bio-Spin P-30 (Bio-rad, numéro de catalogue 732-6223).

### Essai de retard sur gel (EMSA)

Les essais de retard sur gel ont été effectués en utilisant 1  $\mu$ g d'extraits de protéines nucléaires de différentes lignées cellulaires AtT-20 préincubées 10 min sur la glace avec 1  $\mu$ g d'ADN non spécifique poly(dI-dC) et poly(dA-dT) dans un tampon de liaison (25 mM HEPES pH 7,9, 84 mM KCl, 10 % glycérol, 5 mM DTT). Pour les polypeptides purifiés, on a utilisé 300 fmoles par réaction de liaison. Les sondes d'ADN radiomarquées (50000 cpm) ont ensuite

été ajoutées pour un volume total de 20µL et incubées 60 min sur la glace. Le cas échéant, 1 µg d'anticorps Flag M2 (F3165, SIGMA) a été ajouté au mélange au cours des 10 dernières minutes d'incubation.

### Électrophorèse

Des gels de polyacrylamide non dénaturant (40 mM Tris, 195 mM glycine, 0,08 % APS, 0,5 µL/mL TEMED, 5 % Acrylamide:Bisacrylamide dans un rapport de 19:1) ont été électrophorés dans du tampon Tris-glycine (40 mM Tris et 195 mM glycine) pendant 60 min à 300 V à 4 °C. Le mélange de réaction a ensuite été chargé, et soumis à l'électrophorèse pendant 3 h à 300 V à 4 °C. Les gels ont été séchés sur du papier Whatman de 3 MM dans un séchoir à gel à 80 °C pendant 60 min sous vide. Les films d'autoradiographie (HyBlot CL, catalogue E3018) ont ensuite été exposés avec les gels séchés pendant 10 à 48 heures à l'aide d'un écran intensificateur.

### RT-qPCR

L'ARN a été extrait des cellules AtT-20 à l'aide du kit RNeasy Mini (Qiagen #74104), et l'ADNc a été synthétisé à l'aide de 5 µg d'ARN et de la transcriptase inverse SuperScript III (Invitrogen, #18080-044) conformément aux recommandations du fabricant. Les ADNc résultants (5 µl de dilution 1/50) ont été analysés par qPCR à l'aide du réactif SYBRgreen (ThermoFisher 91 Scientific #A25741) complété par 500 nM de chaque paire d'amorces spécifiques au gène (fourni dans le tableau supplémentaire 2) dans un volume total de 10 µl. Au moins 2 répliques biologiques ont été analysés en double pour chaque condition à l'aide d'un dispositif de PCR en temps réel ViiA™ 7 (ThermoFisher scientific), et les résultats ont été analysés à l'aide du logiciel connexe. Les valeurs d'expression des gènes ont été normalisées par rapport à celles de TBP (Amorces fournies tableau 2).

Nom	Catégorie	Localisation	Séquence amorce sens	Séquence amorce antisens
TBP	Aleatoire	exons 4-5	CTACCGTGAATCTTGGCTGTAA	GTTGTCCGTGGCTCTCTTATT
Pcsk2	Pionnier	exons 8-9	ATGGGAAGACGGTTGATG	TGATGGCTGAGTTGATGG
Lmcd1	Transcriptionnel	exons 3-4	CCACCTTTGACACTATCAC	GATGAGTTCCATGTA CTGC

Tableau 2- Amorces utilisées pour les RT-qPCRs

### ChIP-qPCR

Les expériences de ChIPs ont été réalisées dans des cellules AtT-20 comme décrit précédemment [167]. Les ADN immunoprécipités ont été dilués dans 100 µL et ont été analysés par qPCR à l'aide du réactif SYBRgreen (ThermoFisher 91 Scientific #A25741) complété par 500 nM de chaque paire d'amorces spécifiques aux enhanceurs (fourni dans le tableau 2) dans un volume total de 10 µl. Au moins 2 répliques biologiques ont été analysés en double pour chaque condition à l'aide d'un dispositif de PCR en temps réel ViiA™ 7 (ThermoFisher scientific), et les résultats ont été analysés à l'aide du logiciel connexe. Les valeurs de liaison des protéines ont été normalisées

par rapport à celles de de deux régions négatives aléatoires et deux régions positives constitutives (Amorces fournies dans tableau 3).

Nom	Catégorie	Localisation	Séquence amorce sens	Séquence amorce antisens
H3K9me3-Poor	Aleatoire	chr12:16988808-16988894	TCTGCAGACAACAGTGTTTAG	TTGCTAAGTCCTAGCCAATTC
H3K9me3-Rich	Aleatoire	chr4:121618790-121618811	TGTGGTATGATGAAGTCTATC	CTTGTGAGTTTCTCCACAGTAT
POMC-7200	Constitutif	chr12:3947596-3947676	ATAAGGAGGCCATCTCTG	CATGGGAGGGTTTAATTGG
tx4	Constitutif	chr9:50396155-50396229	CTCCTGGATTCTGTAATTAAGCAGG	GGTGTGGAGTTGTTGCTCTA
Tfeb	Constitutif	chr17:47729736-47729832	ATACTGGGTGTGGTCTTG	TAGGATTCCAAGGCTCTTC
Pde2a	Constitutif	chr7:101466827-101466959	ACTCAGCATTGAGGGTTC	TGAATAAGGATGTGACAGGAG
Nmnat3	Pionnier	chr9:98329818-98329899	GGTCACGTGAGAGCTAAT	GCTGAGCTTACAGTTTAG
Pcsk2	Pionnier	chr2:143399750-143399829	GACGTGACACTGTATTC	TATTGGTCACGGTCAGTAG
Snca	Pionnier	chr6:60585672-60585768	TGTAACCCAGGAGGTAATG	CTCTGACTTGGAGAGACAA
Zdhh17	Pionnier	chr10:111024870-111024954	TTACCATTGAGCTTGGTCTC	GGAGCAGTGTGAAATTTGTTAG
K12-42	Pionnier	chr12:42716335-42716480	CTTCCATGTTGCAATGCCT	CTGTTGCCTTTGCTCAATC
K15-62	Pionnier	chr15:62959658-62959772	ACCATTGCTTGTCCACCATA	GGCTTCTGTGCCATAACA
K1-245	Pionnier	chr1:24591342-24591452	CCAGCTCCCTTGACATACAC	CAGACAGACAGGCTTGAAA
Mybpc1	Pionnier	chr10:88603126-88603205	CACAAACCAACCCITTCATC	TAATAAGTCAGGTGCTAGGC
Foxg1	Pionnier	chr12:49061394-49061475	CCATGTTATTGACACTGGAA	CCTACAGCCTAGTAGAAGAG

Tableau 3- Amorces utilisées pour les CHIP-qPCRs

### Analyse des données NGS

Les séquences des deux extrémités des fragments des CHIPseq et ATACseq ont été rognées à l'aide de la commande Trimmomatic/0,36 [220] et alignées sur le génome de référence de la souris mm10 à l'aide de la commande Bowtie/2.3.5 (---very-sensitive --no-mixed --no-unal) [221]. Les fichiers Bam ont été créés en utilisant le programme SAMtools/1.9 [222] et les séquences dupliquées ont été supprimées en utilisant MarkDuplicates de Picard/2.17.3 (Broad Institute). Les BAM ont été créés en utilisant la commande bamCoverage (-- normaliser RPKM -bs 10 -e) de deepTools/3.3.0 [223]. Enfin, les pics ont été déterminés par comparaison avec les niveaux de lectures obtenue dans cellules AtT20 (ATACseq) ou dans les cellules AtTNeo (CHIPseq) en utilisant la commande Peaks calling (-f BAMPE -p (1e-3) -g mm) de MACS/2.1.1 [224].

Les catégories de sites de Pax7 sont basées sur des listes déjà publiées [133]. Les pics de CHIPseq de Pax7 ont été catégorisés en utilisant l'absence ou la présence, avant ou après l'expression de Pax7, de la marque d'enhancer H3K4me1 par CHIPseq ; de la marque d'enhancer actif p300 par CHIPseq et de l'accessibilité de l'ADN par ATACseq. Les pics considérés comme significatifs en CHIPseq (Flag, Tpit, BRG1, ASH2) doivent avoir une p-value inférieur à 1e-5. Pour les ATAC, la p-value seuil est de 1e-3. Pour les catégories de pics H3K4me1, la mesure du pic a été quantifié par easeq sur une fenêtre de 200pb au sommet du pic Pax7 et à 500pb de chaque côté du sommet du pic Pax7. Par excel, nous avons déterminé pour quels sites les régions H3K4me1 présentes à 500pb du sommet Pax7 étaient égales, inférieures ou supérieures à la région centrale du pic H3K4me1 (sommet du pic Pax7).

### **Analyse des motifs**

Les motifs reconnus par Pax7 ont été décrits dans Pelletier et al [133]. La présence ou l'absence d'un motif précis à 100 paires de base du sommet sur chaque site pionnier a été déterminée grâce à la commande Homer findMotifsGenome.pl (<http://homer.ucsd.edu/homer/ngs/peakMotifs.html>). Le nombre de motifs Pax7 reconnu sur une fenêtre de 200 paires de bases autour du sommet du pic Pax7 sur chaque site pionnier a été analysée à l'aide d'un programme personnalisé Python [133].

### **RIME (Rapid immunoprecipitation mass spectrometry of endogenous proteins)**

Le RIME est fait selon le protocole de Mohammed et al, 2016 [225]. L'anticorps utilisé est l'anticorps Flag M2 (F3165, SIGMA) et la sonication de l'ADN s'est faite avec le sonicateur Covaris E220 pendant 420s avec une force de 10%. Nous avons analysé les protéines identifiées avec le logiciel Scaffold. Les spectres peptidiques pris en compte ont un FDR (False discovery rate) de 0.1, un minimum de peptide égal à 2 et ont une signifiante de 95 %. Les protéines prises en compte dans les échantillons Pax7 Q+GL+ et Pax7-465Ala459-461 sont cinq fois plus présentes dans ces échantillons que dans notre échantillon contrôle (AtTNeo) et exprimées deux fois plus ou deux fois moins que dans l'échantillon Pax7 Q+GL- avec une pvalue inférieur à 0.05.

## Chapitre 4 : Discussions

L'action pionnière de Pax7 dans l'hypophyse est nécessaire au transcriptome mélanotrope. Elle se déroule en deux phases. La première phase est indépendante de la réplication. Pendant celle-ci, on observe une augmentation marginale de l'accessibilité de l'ADN, l'apparition d'un faible pic monomodal de H3K4me1 ainsi qu'une perte de l'histone H1 (marque de condensation de la chromatine) et de la marque répressive H3K9me2. Durant cette phase, Pax7 recrute différents facteurs de transcriptions et remodeleurs de la chromatine, tels que le facteur de transcription Tpit, le membre du complexe MLL Ash2 et l'histone H3K9 déméthylase Lsd1. Dans la deuxième phase, dépendante de la réplication cellulaire, PAX7 entraîne l'ouverture de la chromatine. Cette ouverture est marquée par la perte d'occupation du nucléosome centrale mesurée par l'histone H3, le gain d'accessibilité de l'ADN, la transition du profil monomodal d'histone H3K4me1 en profil bimodal, l'ajout de la modification d'histone H3K27ac et la perte de la protéine LaminB1. Durant cette phase, nous observons aussi le recrutement de BRG1 et de l'enzyme P300 [Gouhier et al,2022, en préparation].

Nous avons cherché quel domaine dans la protéine Pax7 était nécessaire aux différentes étapes de l'action pionnière. Nous avons donc fait une étude structure-fonction de la protéine Pax7 et utilisé pour cette étude les isoformes naturelles de Pax7, la protéine Pax3 et des mutants Pax7 délétés en C-terminal.

Cette étude a permis d'identifier des domaines de Pax7 nécessaires à la liaison de Pax7 sur les sites pionniers et à l'ouverture de la chromatine sur ces mêmes sites. Elle a aussi permis de définir un type particulier d'enhancers pionniers impliqués dans l'activation du transcriptome mélanotrope. Ces enhancers présentent un plus fort taux de H3K9me2 et H3K9me3 avant le recrutement du facteur pionnier Pax7 et présentent un nucléosome mobile comprenant l'histone H2AZ au niveau des sites de liaisons de Pax7 après la décompaction de la chromatine. L'ouverture de ces sites est sensible à des altérations de Pax7 en N- et C-terminale.

### 4.1 Des isoformes aux fonctions pionnières différentes

Différentes isoformes de Pax7 ont été identifiées dans divers tissus murins et humains. Quatre isoformes résultent de l'épissage alternatif de transcrits alternatifs qui sont produits à la jonction soit de l'intron 2 et de l'exon 3 (trinuécléotide d'acides aminés codant pour le résidu glutamine Q), soit de l'intron 3 et de l'exon 4 (hexanucléotide d'acides aminés codant pour les résidus glycine et leucine GL). Le résidu glutamine est situé dans la région linker des sous-domaines du PD, à proximité du sous-domaine RED, et le dipeptide GL est situé en N-terminal de la troisième hélice alpha du RED. Nous voulions savoir si ces isoformes présentaient des fonctions différentes que ce soit au niveau de leurs fonctions de recrutement, de leurs fonctions transcriptionnelles ou de leurs fonctions

pionnières. Des prédictions de structure secondaire indiquent que l'insertion de GL aurait une influence importante sur la conformation de la protéine puisque la troisième hélice alpha du sous-domaine RED serait allongée [189]. Pelletier et al ont montré l'interdépendance des domaines de liaison PAIRED et HD du facteur Pax7 pour l'action pionnière [133]. Nous avons montré que les isoformes Pax7 GL+ lient en général un peu moins bien les sites d'ADN Pax7 spécifiques que les isoformes Pax7 GL- (Figure 2.5 d). Cependant, quand on regarde les capacités de transcription induites par ces différentes isoformes mesurées par transfection, on constate que la plus faible affinité aux séquences Pax7 des isoformes GL+ n'affectent pas leurs capacités transcriptionnelles (Figure 2.3 a). Ces résultats laissent supposer que l'affinité de Pax7 avec un motif n'est pas un bon marqueur de la capacité de Pax7 à activer la transcription via ce même motif. Il est à déterminer si cela est vrai pour les autres fonctions de Pax7. In vivo, dans les cellules AtT20, les isoformes GL+ sont incapables d'activer le transcriptome mélanotrope (Figure 2.2). Ceci n'est pas expliqué par leurs activités transcriptionnelles avec des hexamère de motifs reconnus par Pax7 ni avec l'enhancer naturel de Pcsk2. Le cas de PCSK2 est d'autant plus significatif quand on voit que toutes les isoformes sont capables d'activer cet enhancer par transfection de manière similaire alors que les résultats ARNseq montrent que l'activation du gène Pcsk2 est significativement amoindrie pour les isoformes Pax7 GL+. La différence d'expression de ce gène in vivo est d'autant plus intéressante car l'expression de ce gène est soumise à l'action pionnière de Pax7 (Figure 2.3). Nous nous sommes donc interrogés sur le recrutement des isoformes GL+ in vivo. Nous voulions savoir si le recrutement des isoformes GL+ était altéré sur certaines catégories d'enhancers présentant un profil d'enhancer actif après le recrutement de Pax7 référence Q+GL-. Nous avons tout d'abord constaté que le recrutement génomique des différentes isoformes Pax7 sur le génome était globalement similaire (Figure 2.3 et 2.4). En regardant précisément les différentes catégories d'enhancer activés par Pax7 (enhancers constitutifs, activés ou pionnier), on constate que le recrutement des isoformes GL+ est avantageux aux sites présents dans la chromatine ouverte (Figure 2.4), ce qui est, là encore, en opposition avec les résultats trouvés in vitro par gel à retardement. Cependant, les sites constitutifs sont la catégorie de sites activés par Pax7 présentant le moins de motif Pax7 (0 à 3) [133] et où la liaison de Pax7 serait donc plus dépendante de l'interaction de Pax7 avec des cofacteurs que de son affinité avec son site de liaison. On pourrait donc supposer au vu des résultats obtenus par ChIPseq (Figure 2.4 et 2.5) et des résultats de RIMEs (Figure 2.16) montrant un meilleur recrutement de différents cofacteurs protéiques que les isoformes GL+ sont avantagées au niveau des sites constitutifs car elles ont une configuration permettant des liaisons plus fortes avec les cofacteurs présents sur ces sites.

Ces mêmes isoformes Pax7 GL+ sont désavantagées sur les sites pionniers (Figure 2.4 et 2.5). Bien qu'elles soient recrutées sur environ 90 % des sites pionniers, la force de recrutement mesurée par ChIPseq sur chacun des 870 sites pionniers est réduite de manière hétérogène selon le site regardé. Là encore, cette hétérogénéité de liaison

suppose que c'est l'environnement chromatinien de chaque site, que ce soit les marques d'histones, l'ouverture de la chromatine ou la présence de cofacteurs, qui impacte la force de recrutement de Pax7 sur ces sites présents dans la chromatine fermée.

Lorsque nous avons étudié la faculté des isoformes hypophysaire Q+GL- (référence) et Q+GL+ à ouvrir la chromatine par ATACseq, nous avons constaté que l'isoforme Pax7 Q+GL+ perd la capacité d'ouvrir la chromatine sur 45 % des sites pionniers. Ces sites présentent une diminution de leur niveau d'ATAC accompagné d'une diminution de liaison de Pax7 Q+GL+ par rapport à Pax7 Q+GL- bien que là encore cette diminution est hétérogène selon le site regardé. Au niveau de ces sites, après le recrutement de Pax7 Q+GL+, la chromatine présente un simple pic H3K4me1 (Figure 2.6). Ces types de sites ont déjà été observés pour le Pax7 Q+GL de référence : ces sites nommés « primés » présentent un profil H3K4me1 monomodal, une liaison significative mais faible du Pax7 Q+GL- et une absence du facteur non pionnier Tpit connu pour être nécessaire à l'ouverture de la chromatine entraînée par Pax7 dans les mélanotropes [106]. Lorsque nous avons étudié la présence de cette protéine au niveau des sites présentant une perte d'ouverture de la chromatine, nous avons observé que Tpit est diminué mais non corrélé directement avec la liaison de Pax7. A l'inverse, la diminution de la liaison des facteurs de remodelages BRG1 et ASH2 présente une corrélation modérée avec la liaison de Pax7 ce qui suggère un recrutement direct de ces facteurs par Pax7 mais pas pour Tpit (Figure 2.8). Deux raisons peuvent expliquer cette absence de corrélation : la première est que Tpit est un facteur de transcription qui peut avoir des liaisons à d'autres protéines permettant sa stabilisation au niveau des pionniers, Pax7 serait donc le responsable direct de son recrutement mais ne serait pas la seule protéine à assurer sa stabilisation sur les sites pionniers. La seconde serait que Pax7 ne recrute pas Tpit directement. Tpit n'est pas un facteur pionnier et ne peut donc pas lier la chromatine fermée ; son recrutement serait donc la conséquence de la première décompaction de la chromatine provoquée par le recrutement de Pax7 associé à la perte de la marque H3K9me2. Cependant, le manque de recrutement de Tpit sur les sites primés par Pax7 Q+GL- contredit cette hypothèse (Figure 2.14). Le rôle de Tpit, bien que nécessaire pour l'ouverture de la chromatine reste obscur. Tpit ne semble pas, dans notre modèle de cellules AtT20, être responsable de l'arrivée du remodeleur de chromatine BRG1 (la corrélation entre ces deux facteurs étant très faible), ni être nécessaire à la déméthylation de la marque H3K9me2. Il est aussi difficile de savoir pour l'instant si la diminution du recrutement des facteurs Tpit, Ash2 et BRG1 est causale pour la perte d'ouverture aux sites sensibles à GL (H3K4me1 symétrique) ou si elle est une conséquence d'un environnement chromatinien non permissif à leur recrutement ou à leur activité ; que ce soit par des marques d'histone ou des facteurs protéiques répressifs. Dans l'article de Mayran et al [167], la marque répressive H3K9me3 et CTCF ont été anticorrélés à la liaison de Pax7 à ses motifs spécifiques suggérant que certaines marques répressives peuvent

empêcher certaines étapes de l'action pionnière de Pax7. Les sites GL-sensibles pourraient présenter un environnement répressif que Pax7 devrait enlever avant de pouvoir entraîner la mise en place et l'activation des enhancers pionniers. On a d'ailleurs montré dans cette thèse que ces sites GL-sensibles (H3K4me1 symétrique) présentent un plus haut taux de la marque répressive H3K9me2 et H3K9me3 laissant supposer que, au niveau de ces sites importants pour la différenciation mélanotrope, la chromatine présenterait une conformation différente. Bien que les RIMEs ne soient pas une expérience propre et que le nombre de sites pionniers représente environ 10% des sites liés par Pax7, la présence de la protéine Ncor1 qui, accompagnée des protéines HDACs, est reliée à la chromatine fermée et au processus de compartimentalisation des gènes transcrits (dans le noyau) et non transcrits (à la périphérie du noyau attaché à la lamine ) pourrait être une piste pour comprendre pourquoi la deuxième partie de l'action pionnière ne se fait pas avec l'isoforme Pax7 Q+GL+. Nous avons déterminé par CHIP-qPCR que Pax7 était capable d'entraîner les mécanismes moléculaires nécessaires à la déméthylation la marque répressive H3K9me2 mais est-elle suffisante pour enlever les protéines répressives présentes sur les sites pionniers et ces protéines peuvent-elles empêcher l'action des remodeleurs de la chromatine sur les sites pionniers ? Bien que cela soit très théorique, il serait intéressant de faire des CHIP-qPCR de la protéine Ncor1 sur des sites pionniers GL-sensibles et insensibles dans les cellules exprimant nos isoformes hypophysaires et le mutant Pax7-465Ala459-461. Cela permettrait de voir si des protéines de complexes répressifs tels que Ncor1 seraient présents sur les sites GL-sensibles (H3K4 symétrique) et pourraient être impliqués dans l'action pionnière comme barrière de cette dernière. Enfin, le fait que des modifications dans les domaines de liaison de Pax7 puissent affecter l'interaction de cette protéine avec d'autres cofacteurs a déjà été montré dans le travail de l'équipe de Rudnicky. Ils ont identifié l'arginine méthyltransférase Carm1 comme une protéine interagissant avec Pax7 et ont découvert que Carm1 méthyle spécifiquement plusieurs arginines dans l'extrémité N-terminale de Pax7 où se trouve le domaine PAIRED. Pax7 méthylé lie directement les protéines du trithorax MLL1/2, ce qui entraîne le recrutement du complexe ASH2L: MLL1/2: WDR5: RBBP5 (histone H3K4 méthyltransférase) au niveau de ces enhancers et du promoteur proximal Myf5 [137].

## 4.2 Le cas de Pax3

Pax3 et Pax7 sont les protéines les plus homologues dans la famille Pax. Elles ont un taux de similitudes de 86% chez l'humain [226]. Les principales différences de ces protéines se situent en C-terminal où Pax3 ne présente pas de domaine OAR (domaine dont la fonction est encore inconnue). Si Pax3 n'est pas exprimé dans l'hypophyse, Pax7 et Pax3 se retrouvent ensemble durant le développement crâniofacial [227] et le développement musculaire [211].

Pendant le développement crâniofacial, Pax3 et Pax7 jouent des rôles cruciaux et redondants. L'absence de Pax3 et Pax7 entraîne une augmentation significative de l'apoptose des cellules issues des crêtes neurales chez le double mutant [227]. Dans les progéniteurs musculaires, Pax7 et Pax3 ont des fonctions redondantes mais aussi distinctes. En l'absence de Pax3 et de Pax7, il y a un déficit majeur dans le muscle squelettique avec l'arrêt de la myogenèse au cours du développement embryonnaire et fœtal. Il a été observé chez les souris mutantes Pax7 une perte postnatale progressive de cellules satellites (cellules souches du muscle adulte) alors même que Pax3 est encore présent. Pax7 est donc nécessaire à la survie des cellules satellites musculaires et leur progression dans le cycle cellulaire, une fonction que son homologue Pax3 ne peut compenser [228]. Dans les cellules musculaires, de souris, il a été montré que Pax7 Q+GL+ était prédominant chez les souris adultes SIL/J qui possèdent des capacités de réparation du muscle squelettique forte et exprimé de la même manière que Pax7 Q+GL- chez la souris adulte BALB/c qui présente des capacités de réparation du muscle squelettique moindre. Basés sur les similarités et les différences entre les protéines Pax7 et Pax3 trouvés dans différents organes, nous avons voulu savoir si Pax3 Q+GL- avait les mêmes capacités pionnières que Pax7 Q+GL-. À notre grande surprise, les cellules exprimant Pax3 montrent le même phénotype que la protéine Pax7 Q+GL+. Elles échouent à activer le transcriptome mélanotrope et échouent à activer les enhanceurs pionniers. Pax3 prime les sites GL-sensible (ATAC) de la même manière que Pax7 Q+GL+ malgré le fait qu'il est GL- ce qui laisse supposer qu'une autre partie de la protéine Pax7, différente de Pax3, pourrait entrer en jeu dans l'ouverture et l'activation de ces sites pionniers.

### **4.3 La partie C-terminal de Pax7 est nécessaire aux différentes étapes de l'action pionnière.**

Nous avons pris en considération le fait que Pax3, malgré son échec à accomplir la deuxième phase de l'action pionnière, est capable d'accomplir une partie importante de la première étape de l'action pionnière : la liaison aux sites pionniers. Nous avons recherché quelles séquences étaient conservées entre les séquences homologues de Pax7 et Pax3 en C-terminal et les avons délétées de la protéine Pax7. L'étude de ces différents délétants pour l'ensemble de l'action pionnière nous ont permis de déterminer que les domaines de liaison à l'ADN de Pax7 ne sont pas suffisants pour la liaison des sites pionniers et compromettent fortement la stabilité de la liaison de Pax7 sur les sites constitutifs. Ce résultat supporte l'idée que les cofacteurs présents sur site sont importants pour la liaison d'une protéine à un environnement chromatinien particulier. À l'opposé, le mutant Pax7-465, qui ne comporte pas de domaine OAR, présente une action pionnière similaire au Pax7 référence bien qu'avec des marques d'histones légèrement diminuées. Ceci nous indique que le domaine OAR et les séquences qui l'entourent ne sont pas nécessaires à l'action pionnière (Figure 2.11). Ce résultat apporte de nouvelles informations sur le rôle du domaine OAR. En 2017, une mutation de PAX7 entraînant seulement l'expression de la protéine Pax7 sans le domaine OAR a été associée à un syndrome de retard de croissance, d'hypotonie et de

retard neurodéveloppemental global chez un patient [196]. Nos résultats suggèrent que nous pouvons maintenant enlever une absence de la capacité pionnière de Pax7 comme cause de ces différents phénotypes à moins que, dans le cerveau, d'autres cofacteurs soient impliqués dans le remodelage de la chromatine que ceux présents dans l'hypophyse.

Enfin, nous avons défini un intervalle entre les acides aminés 431 et 465 qui est important pour la liaison aux sites pionniers GL-sensibles et en général pour l'ouverture de la chromatine. Cette région est riche en résidus phosphorylables qui pourraient permettre de moduler les différentes fonctions de la protéine Pax7. Dans la littérature, on a identifié, chez Pax7, des mutations dans cette région chez neuf personnes d'une cohorte de patients présentant une scoliose congénitale [202, 207]. Une autre étude a montré que des modifications telles que l'acétylation de certains résidus de Pax7 changent sa fonctionnalité : elle réduit sa capacité de liaison via son domaine HD et inhibe l'activité transcriptionnelle qui découle de ces déficiences de recrutement [229]

Lorsque nous avons remplacé les résidus phosphorylables de l'intervalle 431-465 par des résidus neutres alanines, le remplacement des résidus 459-461 et 463-464 de Pax7-465 ont montré par CHIP-qPCR des impacts significatifs sur la liaison de Pax7 à certains sites pionniers. Ces résidus dans le contexte de la protéine Pax7 entière perdent leur impact laissant supposer un rôle redondant entre la partie C-terminale de cette protéine et ces résidus (Figure 2.10, 2.12 et résultats non montrés). Nous avons étudié l'action pionnière du Pax7-465Ala459-461 par CHIPseq, nous avons constaté que bien que la liaison de ce mutant soit affectée, elle n'est pas perdue sur une bonne partie des sites GL-sensibles. Sur les sites GL-insensibles (H3K4me1), ce mutant est capable d'activer ces sites de manière similaire à Pax7 Q+GL- (503) référence et Pax7 Q+GL+. Cependant, sur les sites GL-sensibles (H3K4me1 symétrique), Pax7-465Ala459-461 lie ces sites mais n'est pas capable de les ouvrir et de les activer. On se retrouve encore, pour ce mutant, avec des sites primés au lieu d'être activés. Cela nous suggère une nouvelle fois deux choses : la première est que les parties N- et C-terminales de Pax7 semblent jouer un rôle au niveau des sites GL-sensibles dans la deuxième étape de l'action pionnière de Pax7. Un impact de la partie C-terminale de Pax7 sur sa partie N-terminale a déjà été montré par Luan et al en 2013. Cette étude a montré que la SUMOylation de Pax7 en N-terminale est essentielle pour la fonction de Pax7 dans le développement de la crête neurale, la différenciation myogénique C2C12 et la transactivation transcriptionnelle. Ils démontrent que le domaine de transactivation de Pax7 est nécessaire à cette SUMOylation et que l'ajout du domaine de transactivation de Pax7 au domaine N-terminale de la protéine Pax3 lui permet d'acquérir la capacité de SUMOylation que cette protéine originellement. La seconde est que ces sites GL-sensibles sont particuliers et ont sûrement un environnement spécial qui explique leurs sensibilités aux variations /mutations de la protéine Pax7.

#### **4.4 Quels mécanismes moléculaires pour expliquer la perte d'ouverture des sites pionniers spécifiques à la différenciation cellulaire des mélanotropes ?**

Les régions en N et C-terminale ont des conséquences similaires sur des sites pionniers de la protéine Pax7 nommés GL-sensibles (H3K4me). La perte de l'activation de ces sites corrèle avec une perte du transcriptome mélanotrope nous suggérant que ces régions jouent un rôle important dans la différenciation cellulaire des cellules mélanotropes. Ces sites présentent un signal H3K9me2 plus élevé ce qui nous laisse penser que leur environnement chromatinien est différent. Au niveau de ces sites, l'isoforme Pax7 Q+GL+ et le délétant Pax7 465Ala459-461 entraînent la même perte d'ouverture de la chromatine et la mise en place d'un même pic monomodal de la protéine H3K4me1. Au niveau de ces sites, le recrutement des protéines Tpit, Ash2 et BRG1, que permettent ces protéines Pax7, est diminué de plus de moitié par rapport à celui obtenu avec l'isoforme Pax7 référence.

On pourrait dans un premier temps prendre le parti que la diminution de ces protéines aux sites GL-sensibles est suffisant pour expliquer la perte d'ouverture de la chromatine à ces sites.

Si le rôle dans l'action pionnière d'ASH2 et BRG1 semble clair, le rôle de Tpit dans l'ouverture de la chromatine reste obscur; quelques articles ont mis en avant l'interaction que pouvait avoir les facteur Tbox avec les modulateurs épigénétiques H3K27-Demethylase and H3K4-Methyltransferase par le biais de leur domaine de liaison [230]. Dans leur article publié en 2021, Yujin Kang et al ont délété le domaine SET des histones méthyltransférases MLL3 and MLL4 entraînant la perte de la marque H3K4me1. Cette perte a entraîné une diminution de la perte d'histone H3 aux enhancers putatifs et une diminution du recrutement des facteurs de remodelage de la chromatine BRG1 (SMARCA4), INI1 (SMARCB1) and BAF170 (SMARCC2), des sous unités du complexe SWI/SNF (BAF) ainsi que de SNF2h (SMARCA5), une sous unité du complexe de remodelage de la chromatine ACF [231]. Tpit jouerait donc un rôle dans le recrutement des facteurs MLL nécessaire à la mise en place de H3K4me1. La marque H3K4me1, elle-même, serait nécessaire à l'éviction de nucléosome H3 et au recrutement de divers remodeleurs de la chromatine nécessaire à l'ouverture de cette dernière aux enhancers. Cependant, la marque H3K4me1 est présente sur nos sites GL-sensibles, bien que fortement réduite avec un profil monomodal : ce ne serait donc pas cette étape qui pose un problème pour la perte de l'ouverture de la chromatine aux sites pionniers sur nos sites GL-sensibles.

Un autre article a montré que la protéine T-box BRACHYURY/T était capable d'interagir avec la protéine p300 et d'impacter la mise en place de la marque H3K27ac. Deux articles ont proposé que l'acétylation de l'histone H3 est nécessaire pour son éviction car ces modifications affecteraient la stabilité de l'histone [232, 233] .

La diminution de la liaison de Tpit n'est pas expliquée par la perte de force de liaison de la protéine Pax7 aux sites GL-sensibles. L'explication pour cette diminution reste encore à trouver.

Si on prend le parti qu'une perte partielle des protéines ASH2, BRG1 et Tpit ne peut expliquer la perte d'ouverture de la chromatine observée lors de l'expression des isoformes Pax7 GL+ et du mutant 465Ala459-461, les RIMES de ces deux protéines et de notre Pax7 référence sont une première étape pour caractériser les interacteurs de Pax7 impliqués dans cette perte d'ouverture.

Les RIMES de ces protéines semblent nous indiquer que des mécanismes différents sont impliqués pour expliquer la perte d'ouverture de la chromatine semblable observé pour ces deux protéines.

Le complexe Ncor1 comprenant les histones désacétylases HDAC1 et 2 est deux fois plus représentée dans les RIMES faits avec l'isoforme Pax7 Q+GL+ que dans les RIMES faits avec l'isoforme Pax7 Q+GL- (Figure 2.17). La présence de cette protéine sur les sites pionniers GL-sensibles pourrait contrer l'acétylation de l'histone H3K27ac par Tpit dont le recrutement est diminué à ces sites et donc empêcher l'éviction de l'histone H3 se déroulant dans la deuxième partie de l'action pionnière. Cette hypothèse est bien sûr à confirmer par CHIP.

Pour ce qui est de la protéine Pax7-465Ala459-461, l'expérience de RIME montre que ce délétant présentent des interactions amoindrit avec des protéines qui ont été montrées importante pour la stabilité des protéines pionnières aux sites de chromatine fermée : CBFA2T2 et PRDM6, un homologue de la protéine, PRDM14 [215]. Le facteur RNF20, une protéine responsable de l'ubiquitination de l'histone H2B (Figure 2.17) a, lui aussi, une interaction amoindrit avec le mutant Pax7-465Ala459-461. La perte de la protéine RNF20 entraîne la perte de marque d'histone active, H2Bub1 nécessaire à la méthylation de l'histone H3K4me1, et l'hyperacétylation de H3, H4, et H2A.Z [216] qui ont été montrés nécessaires pour l'éviction de l'histone H3. Cependant, il reste à confirmer que ces protéines soient perdues spécifiquement au niveau des sites GL-sensibles par CHIP dans les cellules exprimant Pax7-465Ala459-461.

#### **4.5 Un mécanisme pionnier différent sur les enhancers GL-sensibles mélanotropes**

On sait que l'isoforme Pax7 Q+GL+ et que la construction Pax7 465Ala 459-461 sont capables d'entraîner les mécanismes moléculaires nécessaires à la transition des sites primés (peu d'ATAC, faible pic H3K4me1 en profil monomodal, pas de P300) en sites complètement activés (fort ATAC, fort pic H3K4me1 en profil bimodale, présence de la protéine P300), sur un sous-ensemble de sites pionniers dits GL- insensibles (H3K4me1). Ceci laisse suggérer que c'est un paramètre spécifique des sites pionniers GL-sensibles qui cause la déficience d'ouverture sur ces sites et non une incapacité de la protéine à accomplir la deuxième étape de l'action pionnière. Nos résultats ont montré un plus fort niveau des marques répressives H3K9me2 et H3K9me3 sur les pionniers GL-sensibles

avant et après l'arrivée de Pax7 dans le contexte du Pax7 référence Q+GL-. De plus, dans le contexte de Pax7 Q+GL-, lorsqu'on compare l'état final de l'action pionnière sur les sites GL-sensibles et GL-insensibles, nous avons montré que les sites pionniers GL-sensibles semblent garder un nucléosome mobile comprenant l'histone H2AZ au niveau du site de liaison de Pax7. A l'opposé, les sites pionniers GL-insensibles ne présentent pas de nucléosomes au niveau du site de liaison de Pax7 suggérant que ce nucléosome est enlevé de l'ADN et non déstabilisé au niveau de ces sites après l'action pionnière. La différence dans le mécanisme pionnier au niveau des sites pionniers GL-sensibles et GL-insensibles pourrait venir de l'incorporation de l'histone H3.3 qui permettrait l'éviction totale du nucléosome sur les sites GL-insensibles et qui ne se ferait pas sur les sites pionniers GL-sensibles. Dans les RIMEs, la protéine RNF20 a été relié à la protéine H2AZ mais un lien de causalité de l'un sur la présence de l'autre reste à prouver [216].

#### 4.5 Conclusion

Les travaux de cette thèse ont permis de mettre en avant des domaines de Pax7 suffisants ou nécessaires aux différentes phases de l'action pionnière de Pax7. Ils ont aussi permis de mettre en lumière un type de sites pionniers nécessaire à l'activation du transcriptome mélanotrope présentant un signal H3K9me2 et H3K9me3 plus élevé avant et après l'action pionnière de Pax7, un nucléosome mobile contenant le variant d'histone H2AZ au niveau du site de liaison de Pax7 après l'action pionnière et dont l'ouverture de la chromatine est sensible aux mutations/variation en N et C-terminale de la protéine Pax7. Notre étude a montré d'ailleurs que, comme pour le facteur pionnier EBF1 [99, 136], la partie C-terminale de Pax7 est nécessaire à l'ouverture de la chromatine. La partie C-terminale des protéines EBF1 et Pax7 est non structurée et présente un fort potentiel pour le recrutement d'autres facteurs qui pourraient être nécessaires à la transcription ou l'activité pionnière de ces protéines. Pour Pax7, elle possède aussi un bon nombre de résidus modifiables qui pourrait affecter la structure de la protéine et permettre le recrutement de différentes combinaisons d'interacteurs nécessaire aux diverses fonctions de la protéine Pax7. Les mêmes phénotypes comprenant la perte d'ouverture de chromatine et un profil monomodal de la marque H3K4me1 observés pour des perturbations en N et C-terminal laissent à penser que ces régions pourraient agir ensemble lors de l'action pionnière. Nous sommes obligés d'accepter qu'il soit, pour l'instant, difficile d'expliquer de manière claire quels sont les mécanismes moléculaires empêchant la deuxième phase de l'action pionnière de se dérouler avec l'isoforme Pax7 Q+GL+ et le mutant Pax7-465Ala459-461. La perte d'affinité d'une protéine à son site de liaison n'est pas forcément signe de dysfonction de la protéine. Cela pose un doute sur l'explication simple que la diminution de liaison des mutants et isoformes de Pax7 aux sites pionniers entraînerait une déstabilisation des complexes de remodelage, qui n'étant pas suffisamment ancrés aux sites pionniers, ne pourraient assurer leurs fonctions. La seconde hypothèse impliquant la protéine Ncor1 ou la protéine RNF20 (expliqué plus haut) doit être testée par ChIPseq de ces protéines au niveau de la chromatine

native des AtT20 pour Ncor1 et dans la chromatine après le recrutement de Pax7 pour Ncor1 et RNF20 dans un premier temps.

Nous retiendrons aussi de cette étude que les différentes isoformes de Pax7 et Pax3 ne partagent pas les mêmes capacités pionnières dans l'hypophyse. Pax3 et Pax7 présentent les mêmes variations dans le domaine PD chez l'humain et la souris. Les protéines Pax7 et Pax3 sont exprimées à différents temps au cours du développement et dans différents organes. Il serait intéressant de savoir quelles isoformes de ces protéines sont exprimées durant le développement. L'action pionnière étant un évènement nécessaire pour la mise en place de nouveaux enhancers modulant le transcriptome et entraînant la différenciation cellulaire, nous pourrions déterminer quand et où Pax3 et Pax7 utilisent cette capacité particulière pour entraîner une transition cellulaire lors du développement.

## Références

1. Bianconi, E., et al., *An estimation of the number of cells in the human body*. Ann Hum Biol, 2013. **40**(6): p. 463-71.
2. Hagolani, P.F., et al., *Cell signaling stabilizes morphogenesis against noise*. Development, 2019. **146**(20).
3. Bolouri, H. and E.H. Davidson, *Transcriptional regulatory cascades in development: initial rates, not steady state, determine network kinetics*. Proc Natl Acad Sci U S A, 2003. **100**(16): p. 9371-6.
4. Natoli, G., *Maintaining cell identity through global control of genomic organization*. Immunity, 2010. **33**(1): p. 12-24.
5. Mitsui, K., et al., *The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells*. Cell, 2003. **113**(5): p. 631-642.
6. Baralle, F.E. and J. Giudice, *Alternative splicing as a regulator of development and tissue identity*. Nat Rev Mol Cell Biol, 2017. **18**(7): p. 437-451.
7. Burow, D.A., et al., *Dynamic regulation of mRNA decay during neural development*. Neural Dev, 2015. **10**: p. 11.
8. Mallo, M., *Reassessing the Role of Hox Genes during Vertebrate Development and Evolution*. Trends Genet, 2018. **34**(3): p. 209-217.
9. Stoykova, A. and P. Gruss, *Roles of Pax-genes in developing and adult brain as suggested by expression patterns*. J Neurosci, 1994. **14**(3 Pt 2): p. 1395-412.
10. Monsoro-Burq, A.H., *PAX transcription factors in neural crest development*. Semin Cell Dev Biol, 2015. **44**: p. 87-96.
11. Mansouri, A. and P. Gruss, *Pax3 and Pax7 are expressed in commissural neurons and restrict ventral neuronal identity in the spinal cord*. Mech. Dev, 1998. **78**(1-2): p. 171-178.
12. Ericson, J., et al., *Pax6 controls progenitor cell identity and neuronal fate in response to graded Shh signaling*. Cell, 1997. **90**(1): p. 169-180.
13. Chalepakis, G., et al., *Pax: gene regulators in the developing nervous system*. J Neurobiol, 1993. **24**(10): p. 1367-84.
14. Quinonez, S.C. and J.W. Innis, *Human HOX gene disorders*. Mol Genet Metab, 2014. **111**(1): p. 4-15.
15. Smith, M. and P.L. Flodman, *Expanded Insights Into Mechanisms of Gene Expression and Disease Related Disruptions*. Front Mol Biosci, 2018. **5**: p. 101.
16. Heinz, S., et al., *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities*. Mol Cell, 2010. **38**(4): p. 576-589.
17. Drouin, J., *Minireview: pioneer transcription factors in cell fate specification*. Mol. Endocrinol, 2014. **28**(7): p. 989-998.
18. Zaret, K.S. and J.S. Carroll, *Pioneer transcription factors: establishing competence for gene expression*. Genes Dev, 2011. **25**(21): p. 2227-2241.
19. Wang, J., et al., *Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors*. Genome Res, 2012. **22**(9): p. 1798-1812.
20. Bowman, G.D. and M.G. Poirier, *Post-translational modifications of histones that influence nucleosome dynamics*. Chem Rev, 2015. **115**(6): p. 2274-95.
21. Tsunaka, Y., et al., *Alteration of the nucleosomal DNA path in the crystal structure of a human nucleosome core particle*. Nucleic Acids Res, 2005. **33**(10): p. 3424-34.
22. Kornberg, R.D. and Y. Lorch, *Primary Role of the Nucleosome*. Mol Cell, 2020. **79**(3): p. 371-375.
23. Huertas, J., et al., *Nucleosomal DNA Dynamics Mediate Oct4 Pioneer Factor Binding*. Biophys J, 2020. **118**(9): p. 2280-2296.
24. Eslami-Mossallam, B., H. Schiessel, and J. van Noort, *Nucleosome dynamics: Sequence matters*. Adv Colloid Interface Sci, 2016. **232**: p. 101-113.

25. Ioshikhes, I.P., et al., *Nucleosome positions predicted through comparative genomics*. Nat Genet, 2006. **38**(10): p. 1210-5.
26. Segal, E., et al., *A genomic code for nucleosome positioning*. Nature, 2006. **442**(7104): p. 772-8.
27. Anderson, J.D. and J. Widom, *Poly(dA-dT) promoter elements increase the equilibrium accessibility of nucleosomal DNA target sites*. Mol Cell Biol, 2001. **21**(11): p. 3830-9.
28. Bednar, J., et al., *Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin*. Proc Natl Acad Sci U S A, 1998. **95**(24): p. 14173-8.
29. Zhang, H., D.N. Roberts, and B.R. Cairns, *Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss*. Cell, 2005. **123**(2): p. 219-31.
30. Raisner, R.M., et al., *Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin*. Cell, 2005. **123**(2): p. 233-48.
31. Allshire, R.C. and H.D. Madhani, *Ten principles of heterochromatin formation and function*. Nat Rev Mol Cell Biol, 2018. **19**(4): p. 229-244.
32. Bonev, B., et al., *Multiscale 3D Genome Rewiring during Mouse Neural Development*. Cell, 2017. **171**(3): p. 557-572.e24.
33. Dixon, J.R., et al., *Chromatin architecture reorganization during stem cell differentiation*. Nature, 2015. **518**(7539): p. 331-6.
34. Phillips-Cremins, J.E., et al., *Architectural protein subclasses shape 3D organization of genomes during lineage commitment*. Cell, 2013. **153**(6): p. 1281-1295.
35. Escobar, T.M., et al., *Active and Repressed Chromatin Domains Exhibit Distinct Nucleosome Segregation during DNA Replication*. Cell, 2019. **179**(4): p. 953-963.e11.
36. Peterson, C.L. and J.L. Workman, *Promoter targeting and chromatin remodeling by the SWI/SNF complex*. Curr. Opin. Genet. Dev, 2000. **10**(2): p. 187-192.
37. Steger, D.J. and J.L. Workman, *Remodeling chromatin structures for transcription: what happens to the histones?* Bioessays, 1996. **18**(11): p. 875-84.
38. Clapier, C.R., et al., *Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes*. Nat Rev Mol Cell Biol, 2017. **18**(7): p. 407-422.
39. Dege, C. and J. Hagman, *Mi-2/NuRD chromatin remodeling complexes regulate B and T-lymphocyte development and function*. Immunol Rev, 2014. **261**(1): p. 126-40.
40. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications*. Nature, 2000. **403**(6765): p. 41-5.
41. Aloia, L., B. Di Stefano, and L. Di Croce, *Polycomb complexes in stem cells and embryonic development*. Development, 2013. **140**(12): p. 2525-34.
42. Sauvageau, M. and G. Sauvageau, *Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer*. Cell Stem Cell, 2010. **7**(3): p. 299-313.
43. Bernstein, B.E., et al., *A bivalent chromatin structure marks key developmental genes in embryonic stem cells*. Cell, 2006. **125**(2): p. 315-26.
44. Fakhouri, T.H., et al., *Dynamic chromatin organization during foregut development mediated by the organ selector gene PHA-4/FoxA*. PLoS. Genet, 2010. **6**(8).
45. Hoffmann, A., C.A. Zimmermann, and D. Spengler, *Molecular epigenetic switches in neurodevelopment in health and disease*. Front Behav Neurosci, 2015. **9**: p. 120.
46. Eberharter, A. and P.B. Becker, *Histone acetylation: a switch between repressive and permissive chromatin*. Second in review series on chromatin dynamics. EMBO Rep, 2002. **3**(3): p. 224-9.
47. Shen, Y., W. Wei, and D.X. Zhou, *Histone Acetylation Enzymes Coordinate Metabolism and Gene Expression*. Trends Plant Sci, 2015. **20**(10): p. 614-621.
48. Dovey, O.M., C.T. Foster, and S.M. Cowley, *Histone deacetylase 1 (HDAC1), but not HDAC2, controls embryonic stem cell differentiation*. Proc Natl Acad Sci U S A, 2010. **107**(18): p. 8242-7.

49. Morrison, E.A., et al., *DNA binding drives the association of BRG1/hBRM bromodomains with nucleosomes*. Nat Commun, 2017. **8**: p. 16080.
50. Raab, J.R., et al., *Co-regulation of transcription by BRG1 and BRM, two mutually exclusive SWI/SNF ATPase subunits*. Epigenetics Chromatin, 2017. **10**(1): p. 62.
51. Cui, J.Y., Z.D. Fu, and J. Dempsey, *The role of histone methylation and methyltransferases in gene regulation*, in *Toxicoepigenetics*. 2019, Elsevier. p. 31-84.
52. Herz, H.M., et al., *Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4*. Genes Dev, 2012. **26**(23): p. 2604-2620.
53. Di Croce, L. and K. Helin, *Transcriptional regulation by Polycomb group proteins*. Nat. Struct. Mol. Biol, 2013. **20**(10): p. 1147-1155.
54. Margueron, R., et al., *Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms*. Mol. Cell, 2008. **32**(4): p. 503-518.
55. Shen, X., et al., *EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency*. Mol. Cell, 2008. **32**(4): p. 491-502.
56. Nicetto, D., et al., *H3K9me3-heterochromatin loss at protein-coding genes enables developmental lineage specification*. Science, 2019. **363**(6424): p. 294-297.
57. Jacobs, S.A. and S. Khorasanizadeh, *Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail*. Science, 2002. **295**(5562): p. 2080-3.
58. Abdul-Latif, H., et al., *Growth hormone deficiency type 1B caused by cryptic splicing of the GH-1 gene*. J. Pediatr. Endocrinol. Metab, 2000. **13**(1): p. 21-28.
59. Gates, L.A., C.E. Foulds, and B.W. O'Malley, *Histone Marks in the 'Driver's Seat': Functional Roles in Steering the Transcription Cycle*. Trends Biochem Sci, 2017. **42**(12): p. 977-989.
60. Bernstein, B.E., A. Meissner, and E.S. Lander, *The mammalian epigenome*. Cell, 2007. **128**(4): p. 669-81.
61. Zeng, Y. and T. Chen, *DNA Methylation Reprogramming during Mammalian Development*. Genes (Basel), 2019. **10**(4).
62. Shipony, Z., et al., *Dynamic and static maintenance of epigenetic memory in pluripotent and somatic cells*. Nature, 2014. **513**(7516): p. 115-9.
63. Antequera, F., *Structure, function and evolution of CpG island promoters*. Cell Mol Life Sci, 2003. **60**(8): p. 1647-58.
64. Fernandez, A.F., et al., *A DNA methylation fingerprint of 1628 human samples*. Genome Res, 2012. **22**(2): p. 407-19.
65. Weber, M., et al., *Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome*. Nat Genet, 2007. **39**(4): p. 457-66.
66. Laurent, L., et al., *Dynamic changes in the human methylome during differentiation*. Genome Res, 2010. **20**(3): p. 320-31.
67. Hermann, A., H. Gowher, and A. Jeltsch, *Biochemistry and biology of mammalian DNA methyltransferases*. Cell Mol Life Sci, 2004. **61**(19-20): p. 2571-87.
68. Robert, M.F., et al., *DNMT1 is required to maintain CpG methylation and aberrant gene silencing in human cancer cells*. Nat Genet, 2003. **33**(1): p. 61-5.
69. Koivunen, P. and T. Laukka, *The TET enzymes*. Cell Mol Life Sci, 2018. **75**(8): p. 1339-1348.
70. Rottach, A., H. Leonhardt, and F. Spada, *DNA methylation-mediated epigenetic control*. J Cell Biochem, 2009. **108**(1): p. 43-51.
71. Bommarito, P.A. and R.C. Fry, *The role of DNA methylation in gene regulation*, in *Toxicoepigenetics*. 2019, Elsevier. p. 127-151.
72. Balsalobre, A. and J. Drouin, *Pioneer factors as master regulators of the epigenome and cell fate*. Nat Rev Mol Cell Biol, 2022. **23**(7): p. 449-464.
73. Zaret, K., *Developmental competence of the gut endoderm: genetic potentiation by GATA and HNF3/fork head proteins*. Dev Biol, 1999. **209**(1): p. 1-10.

74. Mayran, A. and J. Drouin, *Pioneer transcription factors shape the epigenetic landscape*. J Biol Chem, 2018. **293**(36): p. 13795-13804.
75. Ninova, M., K. Fejes Tóth, and A.A. Aravin, *The control of gene expression and cell identity by H3K9 trimethylation*. Development, 2019. **146**(19).
76. Stricker, S.H., A. Köferle, and S. Beck, *From profiles to function in epigenomics*. Nat Rev Genet, 2017. **18**(1): p. 51-66.
77. Bibikova, M., et al., *High density DNA methylation array with single CpG site resolution*. Genomics, 2011. **98**(4): p. 288-95.
78. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors*. Cell, 2006. **126**(4): p. 663-676.
79. Takahashi, K. and S. Yamanaka, *A decade of transcription factor-mediated reprogramming to pluripotency*. Nat Rev Mol Cell Biol, 2016. **17**(3): p. 183-93.
80. Li, D., et al., *Chromatin Accessibility Dynamics during iPSC Reprogramming*. Cell Stem Cell, 2017. **21**(6): p. 819-833.e6.
81. Chronis, C., et al., *Cooperative Binding of Transcription Factors Orchestrates Reprogramming*. Cell, 2017. **168**(3): p. 442-459.e20.
82. Chen, J., et al., *Hierarchical Oct4 Binding in Concert with Primed Epigenetic Rearrangements during Somatic Cell Reprogramming*. Cell Rep, 2016. **14**(6): p. 1540-1554.
83. Knaupp, A.S., et al., *Transient and Permanent Reconfiguration of Chromatin and Transcription Factor Occupancy Drive Reprogramming*. Cell Stem Cell, 2017. **21**(6): p. 834-845.e6.
84. Soufi, A., et al., *Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming*. Cell, 2015. **161**(3): p. 555-68.
85. Yin, Y., et al., *Impact of cytosine methylation on DNA binding specificities of human transcription factors*. Science, 2017. **356**(6337).
86. Soufi, A., G. Donahue, and K.S. Zaret, *Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome*. Cell, 2012. **151**(5): p. 994-1004.
87. Apostolou, E. and M. Stadtfeld, *Cellular trajectories and molecular mechanisms of iPSC reprogramming*. Curr Opin Genet Dev, 2018. **52**: p. 77-85.
88. Gualdi, R., et al., *Hepatic specification of the gut endoderm in vitro: cell signaling and transcriptional control*. Genes Dev, 1996. **10**(13): p. 1670-82.
89. Sekiya, S. and A. Suzuki, *Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors*. Nature, 2011. **475**(7356): p. 390-3.
90. Morris, S.A., et al., *Dissecting engineered cell types and enhancing cell fate conversion via CellNet*. Cell, 2014. **158**(4): p. 889-902.
91. Cirillo, L.A., et al., *Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4*. Mol. Cell, 2002. **9**(2): p. 279-289.
92. Horisawa, K., et al., *The Dynamics of Transcriptional Activation by Hepatic Reprogramming Factors*. Mol Cell, 2020. **79**(4): p. 660-676.e8.
93. Lee, K., et al., *FOXA2 Is Required for Enhancer Priming during Pancreatic Differentiation*. Cell Rep, 2019. **28**(2): p. 382-393.e7.
94. Carotta, S., L. Wu, and S.L. Nutt, *Surprising new roles for PU.1 in the adaptive immune response*. Immunol Rev, 2010. **238**(1): p. 63-75.
95. Barozzi, I., et al., *Coregulation of transcription factor binding and nucleosome occupancy through DNA features of mammalian enhancers*. Mol Cell, 2014. **54**(5): p. 844-57.
96. van Oevelen, C., et al., *C/EBPalpha Activates Pre-existing and De Novo Macrophage Enhancers during Induced Pre-B Cell Transdifferentiation and Myelopoiesis*. Stem Cell Reports, 2015. **5**(2): p. 232-47.
97. Gyory, I., et al., *Transcription factor Ebf1 regulates differentiation stage-specific signaling, proliferation, and survival of B cells*. Genes Dev, 2012. **26**(7): p. 668-82.

98. Treiber, T., et al., *Early B cell factor 1 regulates B cell gene networks by activation, repression, and transcription-independent poising of chromatin*. *Immunity*, 2010. **32**(5): p. 714-25.
99. Boller, S., et al., *Pioneering Activity of the C-Terminal Domain of EBF1 Shapes the Chromatin Landscape for B Cell Programming*. *Immunity*, 2016. **44**(3): p. 527-41.
100. Li, R., et al., *Dynamic EBF1 occupancy directs sequential epigenetic and transcriptional events in B-cell programming*. *Genes Dev*, 2018. **32**(2): p. 96-111.
101. Ren, G., et al., *Transcription factors TCF-1 and GATA3 are key factors for the epigenetic priming of early innate lymphoid progenitors toward distinct cell fates*. *Immunity*, 2022. **55**(8): p. 1402-1413.e4.
102. Shin, B., et al., *Runx1 and Runx3 drive progenitor to T-lineage transcriptome conversion in mouse T cell commitment via dynamic genomic site switching*. *Proc Natl Acad Sci U S A*, 2021. **118**(4).
103. Johnson, J.L., et al., *Lineage-Determining Transcription Factor TCF-1 Initiates the Epigenetic Identity of T Cells*. *Immunity*, 2018. **48**(2): p. 243-257.e10.
104. van de Wetering, M., et al., *Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box*. *Embo j*, 1991. **10**(1): p. 123-32.
105. Budry, L., et al., *The selector gene Pax7 dictates alternate pituitary cell fates through its pioneer action on chromatin remodeling*. *Genes & Development*, 2012. **26**(20): p. 2299-2310.
106. Mayran, A., et al., *Pioneer and nonpioneer factor cooperation drives lineage specific chromatin opening*. *Nature Communications*, 2019. **10**(1): p. 3807.
107. Luna-Zurita, L., et al., *Complex Interdependence Regulates Heterotypic Transcription Factor Distribution and Coordinates Cardiogenesis*. *Cell*, 2016. **164**(5): p. 999-1014.
108. Gao, R., et al., *Pioneering function of Isl1 in the epigenetic control of cardiomyocyte cell fate*. *Cell Res*, 2019. **29**(6): p. 486-501.
109. Wapinski, O.L., et al., *Rapid Chromatin Switch in the Direct Reprogramming of Fibroblasts to Neurons*. *Cell Rep*, 2017. **20**(13): p. 3236-3247.
110. Wapinski, O.L., et al., *Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons*. *Cell*, 2013. **155**(3): p. 621-35.
111. Pataskar, A., et al., *NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program*. *Embo j*, 2016. **35**(1): p. 24-45.
112. Hahn, M.A., et al., *Reprogramming of DNA methylation at NEUROD2-bound sequences during cortical neuron differentiation*. *Sci Adv*, 2019. **5**(10): p. eaax0080.
113. Lilja, K.C., et al., *Pax7 remodels the chromatin landscape in skeletal muscle stem cells*. *PLoS One*, 2017. **12**(4): p. e0176190.
114. Magli, A., et al., *Time-dependent Pax3-mediated chromatin remodeling and cooperation with Six4 and Tead2 specify the skeletal myogenic lineage in developing mesoderm*. *PLOS Biology*, 2019. **17**(2): p. e3000153.
115. Zhang, N., et al., *Muscle progenitor specification and myogenic differentiation are associated with changes in chromatin topology*. *Nat Commun*, 2020. **11**(1): p. 6222.
116. Harada, A., et al., *Chd2 interacts with H3.3 to determine myogenic cell fate*. *Embo j*, 2012. **31**(13): p. 2994-3007.
117. Esteves de Lima, J., et al., *Unexpected contribution of fibroblasts to muscle lineage as a mechanism for limb muscle patterning*. *Nat Commun*, 2021. **12**(1): p. 3851.
118. Chen, X., et al., *Integration of external signaling pathways with the core transcriptional network in embryonic stem cells*. *Cell*, 2008. **133**(6): p. 1106-1117.
119. Reményi, A., et al., *Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers*. *Genes Dev*, 2003. **17**(16): p. 2048-59.
120. Bilokapic, S., M. Strauss, and M. Halic, *Structural rearrangements of the histone octamer translocate DNA*. *Nat Commun*, 2018. **9**(1): p. 1330.

121. Bilokapic, S., M. Strauss, and M. Halic, *Histone octamer rearranges to adapt to DNA unwrapping*. *Nat Struct Mol Biol*, 2018. **25**(1): p. 101-108.
122. Michael, A.K., et al., *Mechanisms of OCT4-SOX2 motif readout on nucleosomes*. *Science*, 2020. **368**(6498): p. 1460-1465.
123. Merino, F., B. Bouvier, and V. Cojocaru, *Cooperative DNA Recognition Modulated by an Interplay between Protein-Protein Interactions and DNA-Mediated Allostery*. *PLoS Comput Biol*, 2015. **11**(6): p. e1004287.
124. Dodonova, S.O., et al., *Nucleosome-bound SOX2 and SOX11 structures elucidate pioneer factor function*. *Nature*, 2020. **580**(7805): p. 669-672.
125. Clark, K.L., et al., *Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5*. *Nature*, 1993. **364**(6436): p. 412-20.
126. Ramakrishnan, V., et al., *Crystal structure of globular domain of histone H5 and its implications for nucleosome binding*. *Nature*, 1993. **362**(6417): p. 219-23.
127. Goytisolo, F.A., L.C. Packman, and J.O. Thomas, *Photoaffinity labelling of a DNA-binding site on the globular domain of histone H5*. *Eur J Biochem*, 1996. **242**(3): p. 619-26.
128. Iwafuchi-Doi, M., et al., *The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation*. *Mol Cell*, 2016. **62**(1): p. 79-91.
129. Cirillo, L.A. and K.S. Zaret, *Specific interactions of the wing domains of FOXA1 transcription factor with DNA*. *J Mol Biol*, 2007. **366**(3): p. 720-4.
130. Meers, M.P., D.H. Janssens, and S. Henikoff, *Pioneer Factor-Nucleosome Binding Events during Differentiation Are Motif Encoded*. *Mol Cell*, 2019. **75**(3): p. 562-575.e5.
131. Sekiya, T., et al., *Nucleosome-binding affinity as a primary determinant of the nuclear mobility of the pioneer transcription factor FoxA*. *Genes Dev*, 2009. **23**(7): p. 804-9.
132. Zaret, K.S., J. Lerner, and M. Iwafuchi-Doi, *Chromatin Scanning by Dynamic Binding of Pioneer Factors*. *Mol Cell*, 2016. **62**(5): p. 665-7.
133. Pelletier, A., et al., *Pax7 Pioneer action requires both paired and homeo DNA binding domains*. *Nucleic Acids Research*, 2021: p. 1-13.
134. Fernandez Garcia, M., et al., *Structural Features of Transcription Factors Associating with Nucleosome Binding*. *Mol Cell*, 2019. **75**(5): p. 921-932.e6.
135. Takaku, M., et al., *GATA3-dependent cellular reprogramming requires activation-domain dependent recruitment of a chromatin remodeler*. *Genome Biol*, 2016. **17**: p. 36.
136. Wang, Y., et al., *A Prion-like Domain in Transcription Factor EBF1 Promotes Phase Separation and Enables B Cell Programming of Progenitor Chromatin*. *Immunity*, 2020. **53**(6): p. 1151-1167.e6.
137. Kawabe, Y., et al., *Carm1 regulates Pax7 transcriptional activity through MLL1/2 recruitment during asymmetric satellite stem cell divisions*. *Cell Stem Cell*, 2012. **11**(3): p. 333-45.
138. Yang, Y.A., et al., *FOXA1 potentiates lineage-specific enhancer activation through modulating TET1 expression and function*. *Nucleic Acids Res*, 2016. **44**(17): p. 8153-64.
139. Sardina, J.L., et al., *Transcription Factors Drive Tet2-Mediated Enhancer Demethylation to Reprogram Cell Fate*. *Cell Stem Cell*, 2018. **23**(5): p. 727-741.e9.
140. Vanzan, L., et al., *High throughput screening identifies SOX2 as a super pioneer factor that inhibits DNA methylation maintenance at its binding sites*. *Nat Commun*, 2021. **12**(1): p. 3337.
141. Drouin, J., *Pituitary Development*, in *The Pituitary*, S. Melmed, Editor. 2010, Elsevier-Academic Press. p. 3-19.
142. Drouin, J. and J. Brière, *Chap 1 - Pituitary Development*, in *The Pituitary (Fifth Edition)*, S. Melmed, Editor. 2022. p. 1-24.
143. Dubois, P.M., A. el Amraoui, and A.G. Heritier, *Development and differentiation of pituitary cells*. *Microsc. Res. Tech*, 1997. **39**(2): p. 98-113.
144. Rizzoti, K., *Genetic regulation of murine pituitary development*. *J Mol Endocrinol*, 2015. **54**(2): p. R55-73.

145. Budry, L., et al., *Related pituitary cell lineages develop into interdigitated 3D cell networks*. Proc. Natl. Acad. Sci. USA, 2011. **108**(30): p. 12515-12520.
146. Davis, S.W., A.H. Mortensen, and S.A. Camper, *Birthdating studies reshape models for pituitary gland cell specification*. Dev Biol, 2011. **352**(2): p. 215-227.
147. Ikeda, H. and T. Yoshimoto, *Developmental changes in proliferative activity of cells of the murine Rathke's pouch*. Cell Tissue Res, 1991. **263**(1): p. 41-47.
148. Fauquier, T., et al., *SOX2-expressing progenitor cells generate all of the major cell types in the adult mouse pituitary gland*. Proc. Natl. Acad. Sci. USA, 2008. **105**(8): p. 2907-2912.
149. Gremeaux, L., et al., *Activated phenotype of the pituitary stem/progenitor cell compartment during the early-postnatal maturation phase of the gland*. Stem Cells Dev, 2012. **21**(5): p. 801-813.
150. Chen, C., et al., *ERM is required for transcriptional control of the spermatogonial stem cell niche*. Nature, 2005. **463**(7053): p. 1030-1034.
151. Raetzman, L.T., R. Ward, and S.A. Camper, *Lhx4 and Prop1 are required for cell survival and expansion of the pituitary primordia*. Development, 2002. **129**(18): p. 4229-4239.
152. Cohen, L.E. and S. Radovick, *Molecular basis of combined pituitary hormone deficiencies*. Endocr. Rev, 2002. **23**(4): p. 431-442.
153. Li, S., et al., *Dwarf locus mutants lacking three pituitary cell types result from mutations in the POU-domain gene pit-1*. Nature, 1990. **347**: p. 528-533.
154. Ward, R.D., et al., *Role of PROP1 in pituitary gland growth*. Mol. Endocrinol, 2005. **19**(3): p. 698-710.
155. Potok, M.A., et al., *WNT signaling affects gene expression in the ventral diencephalon and pituitary gland growth*. Dev Dyn, 2008. **237**(4): p. 1006-20.
156. Barnhart, K.M. and P.L. Mellon, *The orphan nuclear receptor, steroidogenic factor-1, regulates the glycoprotein hormone alpha-subunit gene in pituitary gonadotropes*. Mol. Endocrinol, 1994. **8**(7): p. 878-885.
157. Zhao, L., et al., *Steroidogenic factor 1 (SF1) is essential for pituitary gonadotrope function*. Development, 2001. **128**(2): p. 147-54.
158. Pulichino, A.M., et al., *Mutations du facteur de transcription Tpit et différenciation hypophysaire*. médecine/sciences, 2004. **20**(11): p. 1009-1013.
159. Pulichino, A.M., S. Vallette-Kasic, and J. Drouin, *Transcriptional regulation of pituitary gland development: binary choices for cell differentiation*. Curr. Opin. Endocrinol. Diabetes, 2004. **11**(1): p. 13-17.
160. Pulichino, A.M., et al., *Tpit determines alternate fates during pituitary cell differentiation*. Genes Dev, 2003. **17**(6): p. 738-747.
161. Lamolet, B., et al., *A pituitary cell-restricted T-box factor, Tpit, activates POMC transcription in cooperation with Pitx homeoproteins*. Cell, 2001. **104**(6): p. 849-859.
162. Seidah, N.G., et al., *The subtilisin/kexin family of precursor convertases. Emphasis on PC1, PC2/7B2, POMC and the novel enzyme SKI-1*. Ann. NY Acad. Sci, 1999. **885**: p. 57-74.
163. Papadimitriou, A. and K.N. Priftis, *Regulation of the hypothalamic-pituitary-adrenal axis*. Neuroimmunomodulation, 2009. **16**(5): p. 265-271.
164. Wu, Y., *Sustained epigenetic programming of POMC by early life stress*. 2012, Imu.
165. Pulichino, A.M., et al., *Human and mouse Tpit gene mutations cause early onset pituitary ACTH deficiency*. Genes Dev, 2003. **17**(6): p. 711-716.
166. Bilodeau, S., A. Roussel-Gervais, and J. Drouin, *Distinct developmental roles of cell cycle inhibitors p57Kip2 and p27Kip1 distinguish pituitary progenitor cell cycle exit from cell cycle re-entry of differentiated cells*. Mol. Cell. Biol, 2009. **29**(7): p. 1895-1908.
167. Mayran, A., et al., *Pioneer factor Pax7 deploys a stable enhancer repertoire for specification of cell fate*. Nat Genet, 2018. **50**(2): p. 259-269

168. Blake, J.A. and M.R. Ziman, *Pax genes: regulators of lineage specification and progenitor cell maintenance*. Development, 2014. **141**(4): p. 737-751.
169. Xu, W., et al., *Crystal structure of a paired domain-DNA complex at 2.5 Å resolution reveals structural basis for Pax developmental mutations*. Cell, 1995. **80**(4): p. 639-650.
170. Czerny, T. and M. Busslinger, *DNA-binding and transactivation properties of Pax-6: three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5)*. Mol. Cell Biol, 1995. **15**(5): p. 2858-2871.
171. Czerny, T., G. Schaffner, and M. Busslinger, *DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site*. Genes Dev, 1993. **7**(10): p. 2048-2061.
172. Epstein, J., et al., *Identification of a Pax paired domain recognition sequence and evidence for DNA-dependent conformational changes*. J. Biol. Chem, 1994. **269**(11): p. 8355-8361.
173. Epstein, J.A., et al., *Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing*. Genes Dev, 1994. **8**(17): p. 2022-2034.
174. Jun, S. and C. Desplan, *Cooperative interactions between paired domain and homeodomain*. Development, 1996. **122**(9): p. 2639-2650.
175. Luscombe, N.M., et al., *An overview of the structures of protein-DNA complexes*. Genome Biol, 2000. **1**(1): p. Reviews001.
176. Cao, R., Y. Tsukada, and Y. Zhang, *Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing*. Mol. Cell, 2005. **20**(6): p. 845-854.
177. Codutti, L., et al., *The solution structure of DNA-free Pax-8 paired box domain accounts for redox regulation of transcriptional activity in the pax protein family*. J Biol. Chem, 2008. **283**(48): p. 33321-33328.
178. Thompson, B., et al., *Overview of PAX gene family: analysis of human tissue-specific variant expression and involvement in human disease*. Hum Genet, 2021. **140**(3): p. 381-400.
179. Eberhard, D., et al., *Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family*. EMBO J, 2000. **19**(10): p. 2292-2303.
180. Hollenbach, A.D., et al., *The Pax3-FKHR oncoprotein is unresponsive to the Pax3-associated repressor hDaxx*. EMBO J, 1999. **18**(13): p. 3702-3711.
181. Lechner, M.S. and G.R. Dressler, *Mapping of Pax-2 transcription activation domains*. J Biol Chem, 1996. **271**(35): p. 21088-93.
182. Muhr, J., et al., *Groucho-mediated transcriptional repression establishes progenitor cell pattern and neuronal fate in the ventral neural tube*. Cell, 2001. **104**(6): p. 861-73.
183. Wang, Q., et al., *Pax genes in embryogenesis and oncogenesis*. J Cell Mol Med, 2008. **12**(6A): p. 2281-2294.
184. Mayran, A., A. Pelletier, and J. Drouin, *Pax factors in transcription and epigenetic remodelling*. Semin Cell Dev Biol, 2015. **44**: p. 135-144.
185. King, H.W. and R.J. Klose, *The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells*. Elife, 2017. **6**.
186. Donaghey, J., et al., *Genetic determinants and epigenetic effects of pioneer-factor occupancy*. Nat Genet, 2018. **50**(2): p. 250-258.
187. Vogan, K.J. and P. Gros, *The C-terminal subdomain makes an important contribution to the DNA binding activity of the Pax-3 paired domain*. J Biol. Chem, 1997. **272**(45): p. 28289-28295.
188. Vogan, K.J., D.A. Underhill, and P. Gros, *An alternative splicing event in the Pax-3 paired domain identifies the linker region as a key determinant of paired domain DNA-binding activity*. Mol. Cell Biol, 1996. **16**(12): p. 6677-6686.
189. Ziman, M.R. and P.H. Kay, *Differential expression of four alternate Pax7 paired box transcripts is influenced by organ- and strain-specific factors in adult mice*. Gene, 1998. **217**(1-2): p. 77-81.

190. Du, S., et al., *Co-expression of alternatively spliced forms of PAX3, PAX7, PAX3-FKHR and PAX7-FKHR with distinct DNA binding and transactivation properties in rhabdomyosarcoma*. *Int. J. Cancer*, 2005. **115**(1): p. 85-92.
191. Ziman, M.R., S. Fletcher, and P.H. Kay, *Alternate Pax7 transcripts are expressed specifically in skeletal muscle, brain and other organs of adult mice*. *Int J Biochem Cell Biol*, 1997. **29**(7): p. 1029-36.
192. Ziman, M.R., et al., *A key role for Pax7 transcripts in determination of muscle and nerve cells*. *Exp Cell Res*, 2001. **268**(2): p. 220-9.
193. White, R.B. and M.R. Ziman, *Genome-wide discovery of Pax7 target genes during development*. *Physiol Genomics*, 2008. **33**(1): p. 41-49.
194. Ziman, M.R., et al., *Pax genes in development and maturation of the vertebrate visual system: implications for optic nerve regeneration*. *Histol Histopathol*, 2001. **16**(1): p. 239-49.
195. Seale, P., et al., *Pax7 is necessary and sufficient for the myogenic specification of CD45+:Sca1+ stem cells from injured muscle*. *PLoS Biol*, 2004. **2**(5): p. E130.
196. Proskorovski-Ohayon, R., et al., *PAX7 mutation in a syndrome of failure to thrive, hypotonia, and global neurodevelopmental delay*. *Hum Mutat*, 2017. **38**(12): p. 1671-1683.
197. Vorobyov, E. and J. Horst, *Expression of two protein isoforms of PAX7 is controlled by competing cleavage-polyadenylation and splicing*. *Gene*, 2004. **342**(1): p. 107-12.
198. Buckingham, M. and F. Relaix, *PAX3 and PAX7 as upstream regulators of myogenesis*. *Semin Cell Dev Biol*, 2015. **44**: p. 115-25.
199. Mansouri, A., et al., *Dysgenesis of cephalic neural crest derivatives in Pax7-/- mutant mice*. *Development*, 1996. **122**(3): p. 831-8.
200. Seale, P. and M.A. Rudnicki, *A new look at the origin, function, and "stem-cell" status of muscle satellite cells*. *Dev. Biol*, 2000. **218**(2): p. 115-124.
201. Seale, P., et al., *Pax7 is required for the specification of myogenic satellite cells*. *Cell*, 2000. **102**(6): p. 777-786.
202. Butali, A., et al., *Replication of genome wide association identified candidate genes confirm the role of common and rare variants in PAX7 and VAX1 in the etiology of nonsyndromic CL(P)*. *Am J Med Genet A*, 2013. **161a**(5): p. 965-72.
203. Gaczkowska, A., et al., *PAX7 nucleotide variants and the risk of non-syndromic orofacial clefts in the Polish population*. *Oral Dis*, 2019. **25**(6): p. 1608-1618.
204. Gowans, L.J., et al., *Association Studies and Direct DNA Sequencing Implicate Genetic Susceptibility Loci in the Etiology of Nonsyndromic Orofacial Clefts in Sub-Saharan African Populations*. *J Dent Res*, 2016. **95**(11): p. 1245-56.
205. Leslie, E.J., et al., *Identification of functional variants for cleft lip with or without cleft palate in or near PAX7, FGFR2, and NOG by targeted sequencing of GWAS loci*. *Am J Hum Genet*, 2015. **96**(3): p. 397-411.
206. Feichtinger, R.G., et al., *Biallelic variants in the transcription factor PAX7 are a new genetic cause of myopathy*. *Genet Med*, 2019. **21**(11): p. 2521-2531.
207. Wang, M., et al., *The identification of PAX7 variants and a potential role of muscle development dysfunction in congenital scoliosis*. *Cell Regen*, 2022. **11**(1): p. 16.
208. Barr, F.G., *Gene fusions involving PAX and FOX family members in alveolar rhabdomyosarcoma*. *Oncogene*, 2001. **20**(40): p. 5736-46.
209. Bennicelli, J.L., et al., *PAX3 and PAX7 exhibit conserved cis-acting transcription repression domains and utilize a common gain of function mechanism in alveolar rhabdomyosarcoma*. *Oncogene*, 1999. **18**(30): p. 4348-4356.
210. Kuang, S., et al., *Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis*. *J. Cell Biol*, 2006. **172**(1): p. 103-113.
211. Relaix, F., et al., *Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells*. *J. Cell Biol*, 2006. **172**(1): p. 91-102.

212. Budry, L., et al., *The Ets Factor Etv1 Interacts with Tpit Protein for Pituitary Pro-opiomelanocortin (POMC) Gene Transcription*. J Biol Chem, 2011. **286**(28): p. 25387-25396.
213. Pradhan, S.K., et al., *EP400 Deposits H3.3 into Promoters and Enhancers during Gene Activation*. Mol Cell, 2016. **61**(1): p. 27-38.
214. Padeken, J., S.P. Methot, and S.M. Gasser, *Establishment of H3K9-methylated heterochromatin and its functions in tissue differentiation and maintenance*. Nat Rev Mol Cell Biol, 2022. **23**(9): p. 623-640.
215. Tu, S., et al., *Co-repressor CBFA2T2 regulates pluripotency and germline development*. Nature, 2016. **534**(7607): p. 387-90.
216. Ma, M.K., et al., *Histone crosstalk directed by H2B ubiquitination is required for chromatin boundary integrity*. PLoS Genet, 2011. **7**(7): p. e1002175.
217. Rambaud, J., et al., *TIF1beta/KAP-1 is a coactivator of the orphan nuclear receptor NGFI-B/Nur77*. J. Biol. Chem, 2009. **284**(21): p. 14147-14156.
218. Langlais, D., C. Couture, and J. Drouin, *The Stat3/GR interaction code: predictive value of direct/indirect DNA recruitment for transcription outcome*. Molecular Cell, 2012. **47**(1): p. 38-49.
219. Thorvaldsdottir, H., J.T. Robinson, and J.P. Mesirov, *Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration*. Brief. Bioinform, 2013. **14**(2): p. 178-192.
220. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data*. Bioinformatics, 2014. **30**(15): p. 2114-20.
221. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome*. Genome Biol, 2009. **10**(3): p. R25.
222. Li, H., et al., *The Sequence Alignment/Map format and SAMtools*. Bioinformatics, 2009. **25**(16): p. 2078-9.
223. Ramírez, F., et al., *deepTools2: a next generation web server for deep-sequencing data analysis*. Nucleic Acids Res, 2016. **44**(W1): p. W160-5.
224. Zhang, Y., et al., *Model-based analysis of ChIPseq (MACS)*. Genome Biol, 2008. **9**(9): p. R137.
225. Mohammed, H., et al., *Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes*. Nat Protoc, 2016. **11**(2): p. 316-26.
226. Soleimani, V.D., et al., *Transcriptional dominance of pax7 in adult myogenesis is due to high-affinity recognition of homeodomain motifs*. Dev Cell, 2012. **22**(6): p. 1208-1220.
227. Zalc, A. and F. Relaix, *[Pax3 and Pax7 play essential safeguard functions against environmental stress-induced birth defects]*. Med Sci (Paris), 2015. **31**(8-9): p. 723-5.
228. Bajard, L., et al., *A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb*. Genes Dev, 2006. **20**(17): p. 2450-2464.
229. Sincennes, M.C., et al., *Acetylation of PAX7 controls muscle stem cell self-renewal and differentiation potential in mice*. Nat Commun, 2021. **12**(1): p. 3253.
230. Miller, S.A., et al., *Coordinated but physically separable interaction with H3K27-demethylase and H3K4-methyltransferase activities are required for T-box protein-mediated activation of developmental gene expression*. Genes Dev, 2008. **22**(21): p. 2980-2993.
231. Kang, Y., et al., *Histone H3K4me1 and H3K27ac play roles in nucleosome eviction and eRNA transcription, respectively, at enhancers*. Faseb j, 2021. **35**(8): p. e21781.
232. Devaiah, B.N., et al., *BRD4 is a histone acetyltransferase that evicts nucleosomes from chromatin*. Nat Struct Mol Biol, 2016. **23**(6): p. 540-8.
233. Luebben, W.R., N. Sharma, and J.K. Nyborg, *Nucleosome eviction and activated transcription require p300 acetylation of histone H3 lysine 14*. Proc Natl Acad Sci U S A, 2010. **107**(45): p. 19254-9.