

**Thèse présentée en vue de l'obtention du grade de
Docteur de l'Université de Strasbourg**

Discipline : Sciences de la Vie et de la Santé
Spécialité : Aspects moléculaires et cellulaires de la biologie

par **Angèle Santenard**

**Régulation de la transcription et organisation de la
chromatine dans les ovocytes et l'embryon
préimplantatoire de souris**

Soutenue publiquement le 30 Juin 2011 devant le jury composé de :

Directeur de thèse : **Maria-Elena TORRES-PADILLA**, IGBMC, Illkirch

Rapporteur Interne : **Michel LABOUESSE**, IGBMC, Illkirch

Rapporteur Externe : **Tilman BORGGREFE**, MPI, Freiburg

Rapporteur Externe : **Robert FEIL**, IGMM, Montpellier

Examineur : **Edith HEARD**, Institut Curie, Paris

Membre invité : **Didier DEVYS**, IGBMC, Illkirch

*“We used to think that our fate was in the stars.
Now we know that (...) our fate is in our genes.”*

James Watson

A mes parents...

REMERCIEMENTS

Je tiens tout d'abord à remercier les membres de mon jury, le Dr Edith Heard, le Dr Tilman Borggreffe, le Dr Robert Feil, le Dr Michel Labouesse et le Dr Didier Devys pour avoir accepté de lire et de juger mon travail de thèse.

Je remercie très chaleureusement ma directrice de thèse, le Dr Maria-Elena Torres-Padilla pour m'avoir permis de mener à bien ma thèse dans de si bonnes conditions. Merci pour ta patience, pour tout le temps que tu as passé à répondre à mes questions, et surtout pour ton enthousiasme et ta capacité à positiver et rebondir dans toutes les situations. Merci pour les *bier/tequila party* dans ton bureau (oups... fallait pas le dire) et à l'extérieur pour les jours de déprime. Quelle chance d'avoir eu une chef comme toi !

Je remercie également le Dr László Tora de m'avoir accueillie dans son laboratoire pour mon M2 et le début de ma thèse et de m'avoir fait travailler avec Maria-Elena.

Je remercie tous les membres actuels et passés du laboratoire de Maria-Elena et de Laszlo pour leurs conseils, leurs idées, et leur aide. Merci Céline pour toutes les manip que tu as faites pour moi et pour ton dynamisme au sein du laboratoire. Un grand merci à Eli pour le temps et l'énergie que tu as passés sur le projet TBP2. Merci pour ton expertise et ta disponibilité. Merci Marjorie pour les expériences de MuDPIT et pour tes explications.

Je tiens également à remercier les services communs de l'IGBMC et en particulier William et Alex à l'animalerie, toute l'équipe du service d'imagerie, le service de culture cellulaire et de biopuces. Merci de rendre notre travail beaucoup plus simple et plus rapide. Merci Evelyne, Maïté et Sylviane pour votre efficacité et votre sourire.

Et parce que les collègues, ça ne sert pas qu'à injecter des souris le week-end... *Thank you Ana for your enthusiasm. It's a pity that you arrived so late in the lab, I would have really appreciated to work and enjoy Strasbourg with you!*

Petite Anne, on s'est rencontré devant le micro-onde au 4^{ème} et on ne s'est plus quitté ! Merci pour ton amitié, ton naturel, tes coups de gueule, et pour avoir pensé à moi tous les midis pour aller manger.

Merci David pour tes conseils avisés et tes analyses psycho-socio-analytico-personnelles dignes d'un magazine féminin ;) ! On se retrouve à Montpellier dans quelques années...

Merci Jacques d'avoir gardé ton âme et ton innocence d'enfant dans ce monde de brute.

Un grand merci à Fernande pour la bouffée d'air que tu représentes au premier étage. Merci pour tout ce que tu m'as rac' et qui, bien sûr, restera entre nous... Reste toi-même, c'est ta force et c'est pour ça qu'on t'aime.

Merci à Fredo pour le défouloir du Body Combat et pour avoir un peu les mêmes opinions que moi sur les gens, la vie... Et surtout... on lâche rien KKIIIAAAAA !!!

Merci Anthony, tu es un partenaire de salsa remarquable (mais oui, David, toi aussi !). Si tu te reconvertis, je viendrai dans ton bar à cocktails avec plaisir !

Dans la vie, il n'y a pas que la biologie, il y a les cailloux aussi... Merci à Flavia, Marie-Laure, Emilie, Edouard et tous les autres géo-whatever pour les soirées apéro qui changent les idées et qui sont toujours les bienvenues.

Et parce qu'il y a une vie hors du labo et de Strasbourg...

Merci à Coralie et Marine pour votre amitié de très longue date, et parce que même si on se voit très peu, on a toujours quelque chose à se dire.

Séverine et Justine, merci de m'apporter une vision du monde différente de celle de la recherche et de l'enseignement. On ne se voit pas assez souvent, mais c'est toujours avec grand plaisir... et une bière ;). Guille, nous n'aurons pas le Prix Nobel ensemble... tout repose sur toi maintenant, j'y crois... « *I'm a Believer* »

Cat, Mumu, Ben, Nico et Xav, merci d'être là pour les très bons moments comme pour les moins bons. Parce qu'on se comprend sans forcément se parler, parce qu'on reste très proches, malgré la distance... avoir des frères et sœur en étant fille unique, ce n'est pas donné à tout le monde !

Maman Corine et Papa Jean-Mi, merci pour cette bulle de douceur et de quiétude que vous arrivez à entretenir et que vous nous faites partager.

Frédérique et Hervé, beau-papa, belle-maman, merci pour votre joie de vivre et votre soutien.

A toute ma famille, même si ce que je fais reste pour vous très obscur, merci d'être une valeur sûre, un retour aux sources dont je ne pourrais pas me passer.

A ma Maman et mon Papa, un grand merci pour votre présence, votre soutien à toute épreuve et votre aide. Merci, tout simplement.

Enfin, Manu, merci de tout mon cœur. Merci de me supporter (presque) tous les jours et de me soutenir et m'accompagner dans mes choix quels qu'ils soient. Merci d'être là...

AVANT-PROPOS

La fécondation d'un ovocyte par un spermatozoïde, deux cellules différenciées, constitue la première étape du développement embryonnaire. Elle induit la formation d'une cellule totipotente, le zygote. Après de nombreuses divisions et plusieurs étapes de différenciation, le zygote permet la formation d'un organisme adulte composé de plusieurs milliards de cellules différentes, pour la plupart différenciées et spécialisées dans l'accomplissement d'une tâche spécifique au sein de l'organisme.

Le rôle des cellules qui composent un organisme adulte, ou qui l'ont composé au cours du développement, est assuré grâce à l'activation et la répression de groupes particuliers de gènes à partir d'un génome dont l'intégrité et la stabilité sont maintenues au cours des divisions. Le contrôle de l'expression des gènes a lieu principalement au niveau transcriptionnel, et plus particulièrement pendant la phase d'initiation grâce à l'action de facteurs de transcription spécifiques et à la régulation de l'état de condensation de la chromatine, matrice de la transcription. L'étude des mécanismes de régulation de la transcription est nécessaire à la compréhension des phénomènes de différenciation, et du passage de deux cellules différenciées à une cellule totipotente lors de la fécondation.

Ma thèse s'articule autour de deux thèmes principaux. Dans un premier temps, je me suis intéressée au rôle d'un facteur de transcription général, TBP2, spécifique des ovocytes chez la souris. L'étude du phénotype d'une souris n'exprimant pas TBP2 et du rôle de cette protéine sur la transcription *in vivo* dans les ovocytes a mené à une publication qui sera présentée dans la partie 2. Elle sera suivie de quelques résultats non publiés sur ce thème.

Dans un deuxième temps, j'ai travaillé sur la structure de la chromatine au cours de l'embryogenèse précoce de souris et plus particulièrement sur la formation de l'hétérochromatine péricentrique des chromosomes paternels après la fécondation. Une publication sur le rôle de H3.3, variant de H3 dans ce phénomène constituera le début de la partie 3. Elle sera également suivie de données non publiées sur l'importance de la localisation de l'hétérochromatine au sein du noyau lors du développement précoce.

Ces deux parties seront précédées d'une introduction générale commune présentant des données globales sur la transcription et la chromatine et montrant la spécificité de mes deux types cellulaires d'étude, l'ovocyte et le zygote.

Dans un souci de clarté, la section « matériels et méthodes » ainsi que la discussion et les perspectives de chacune de ces études figureront à la suite des résultats au sein de leurs parties respectives.

TABLE DES MATIERES

REMERCIEMENTS.....	1
AVANT-PROPOS.....	3
TABLE DES MATIERES.....	4
Liste des figures et tableaux.....	11
Liste des abréviations.....	14
Partie 1 INTRODUCTION.....	16
A. La transcription par l'ARN polymérase II.....	17
I. Les acteurs de la transcription de base.....	17
1. L'ARN polymérase II.....	17
a. Composition et structure.....	17
b. Le cycle de l'ARN polymérase II : les différentes étapes de la transcription.....	18
2. Les facteurs généraux de la transcription : exemple de TBP.....	20
a. Le promoteur minimal.....	20
b. TBP, découverte et structure.....	22
c. TBP au cœur de l'initiation de la transcription.....	23
II. Modulation autour de la transcription de base.....	25
1. ...grâce à des séquences promotrices diverses.....	25
a. Variations autour du promoteur minimal de base.....	25
b. Les séquences régulatrices.....	27
2. ... grâce à des complexes TFIID alternatifs.....	28
a. Variation autour des TAF.....	28
b. La transcription indépendante de TBP.....	29
B. La chromatine, un environnement modulable.....	33
I. Structure de la chromatine.....	33
1. Les différents niveaux de compaction de la chromatine.....	33
a. La fibre de 10 nm.....	33
b. Les niveaux de compaction supérieurs.....	34
2. Histones et nucléosome.....	34

3. Hétérochromatine, euchromatine et transcription.....	35
II. Modifications post-traductionnelles des histones et compaction de la chromatine.....	37
1. Généralités.....	37
a. Mise en place et modes d'action.....	37
b. La théorie du code histone.....	38
2. Un exemple : la méthylation des histones.....	39
a. Généralités.....	39
b. La méthylation de H3K4.....	40
c. La méthylation de H3K9.....	42
d. La méthylation de H3K27.....	44
e. Les domaines bivalents.....	46
III. Les variants d'histones au sein de la chromatine.....	47
1. Généralités.....	47
2. Les variants de H3.....	48
a. H3.1 et H3.2, les formes canoniques de H3.....	48
b. H3.3.....	49
c. H3t, un variant spécifique des testicules.....	50
d. CENP-A, un variant spécifique des centromères.....	51
IV. Régulation des modifications chromatinienne : ARN non codant et hétérochromatinisation.....	51
1. Les différents types d'hétérochromatine.....	52
a. L'hétérochromatine constitutive.....	52
b. L'hétérochromatine facultative.....	53
2. Hétérochromatine facultative et longs ARN non codants chez les mammifères.....	54
3. Hétérochromatine péricentrique et ARN interférent.....	55
a. Chez la levure.....	55
b. Chez les mammifères.....	56
V. Epigénétique et héritabilité des marques des histones.....	56
1. Qu'est ce que l'épigénétique?.....	56
2. Héritabilité des marques des histones.....	57
a. Propagation de marques existantes.....	58

b. Rôle des ARN.....	59
C. Cellules germinales et zygotes, des cellules particulières.....	60
I. Des cellules germinales au blastocyste.....	60
1. La gamétogenèse.....	60
a. L'ovogenèse.....	60
b. La spermatogenèse.....	62
2. Fécondation et stade zygote.....	63
3. Du stade 2-cellules au stade blastocyste.....	64
II. Transcription dans les cellules germinales et pendant le développement précoce.....	65
1. Des taux de transcription variables.....	65
a. Dans les gamètes.....	65
b. Dans le zygote.....	66
2. Utilisation de facteurs de transcription de base spécifiques : exemple des homologues de TBP.....	67
III. Des cellules particulières d'un point de vue épigénétique.....	69
1. Reprogrammation dans les cellules souches primordiales.....	70
2. Réorganisation de la chromatine dans les spermatogonies.....	71
3. Reprogrammation du génome dans le zygote.....	73
a. Evolution du contenu en histones.....	73
b. Evolution des modifications des histones.....	73
c. Evolution de la méthylation de l'ADN.....	74
IV. Les régions péricentromériques dans l'embryon, un exemple de réorganisation de la chromatine au cours du développement précoce de la souris.....	76
1. Evolution de la localisation des régions péricentromériques au cours du développement précoce de la souris.....	76
2. Asymétrie épigénétique.....	77
3. Transcription des <i>major satellites</i>	78
 Partie 2	
ETUDE DU ROLE DE TBP2 DANS LES OVOCYTES DE SOURIS	79

A. Problématique.....	80
B. Publication 1.....	81
I. Résumé de la publication 1.....	81
1. Etude des souris TBP2 ^{-/-}	81
a. Les souris TBP2 ^{-/-} sont stériles suite à un arrêt de la folliculogenèse...	81
b. La perte de TPB2 induit une diminution de la transcription et l'altération du profil d'expression des gènes spécifiques des ovocytes.....	81
2. Spécificité des rôles de TBP et TBP2 au cours de la gamétogenèse et du développement.....	82
a. TBP2 se lie aux promoteurs des gènes activement transcrits dans l'ovocyte.....	82
b. TBP n'est pas nécessaire à la maturation des ovocytes, à la ZGA et au développement précoce.....	82
c. La surexpression de TBP2 a un effet négatif sur le développement embryonnaire et la croissance cellulaire.....	83
II. Publication 1.....	83
C. Résultats non publiés.....	84
I. L'absence de TBP2 n'entraîne pas de défauts évidents de la prophase I de méiose.....	84
1. TBP2 n'est pas détectable dans les ovaires d'embryons à E18,5.....	84
2. La perte de TBP2 n'entraîne pas de défauts évidents dans les ovaires à E17,5 et E18,5.....	85
3. La perte de TBP2 n'entraîne pas de défauts de la recombinaison homologue.....	85
II. L'absence de TBP2 entraîne différents phénotypes ovariens chez les souris âgées.....	86
III. Etude des partenaires protéiques de TBP2 dans les ovocytes, mise au point expérimentale.....	87
1. Préparation des extraits nucléaires d'ovaires.....	88
2. Différentes mises au point de l'immunoprécipitation.....	88

D. Discussion et perspectives.....	90
I. Etude du phénotype macroscopique des souris TBP2^{-/-}	90
1. L’ovaire, seul organe touché par la perte de TBP2 ?.....	90
2. Le rôle des homologues de TBP est-il conservé au cours de l’évolution ?.	91
II. Rôles de la protéine TBP2 dans l’ovocyte.....	92
1. Rôle de TBP2 sur la transcription.....	92
2. Rôle de TBP2 sur la répression de la transcription.....	93
3. Rôle de TBP2 sur la structure chromatinienne.....	94
III. Pourquoi remplacer TBP par TBP2 ?.....	95
1. TBP et TBP2, deux protéines pour un même rôle ?.....	95
2. Quels sont les mécanismes sous-tendant cette spécificité ?.....	96
IV. TBP2 et la stérilité féminine.....	97
E. Matériel et méthodes.....	99
I. Préparation d’extraits nucléaires d’ovaires.....	99
II. Transfection et préparation d’extraits cellulaires	
totaux de cellules HEK.....	99
III. Liaison covalente entre les anticorps et les billes	
de sépharose couplées à la protéine G.....	100
IV. Immunoprécipitation.....	100
Partie 3 MISE EN PLACE ET EVOLUTION	
DE L’HETEROCHROMATINE PERICENTRIQUE	
DANS L’EMBRYON PRECOCE DE SOURIS.....	102
A. Problématique.....	103
B. Publication 2.....	104
I. Résumé de la publication 2.....	104
1. Les embryons exprimant H3.3K27R présentent	
des défauts lors du développement précoce.....	104
2. L’expression de H3.3K27R entraîne des changements	
globaux dans le niveau de modifications des histones.....	104

3. H3.3 et H3.1 sont incorporés à différents moments et à différents locus dans les pronoyaux paternel et maternel.....	105
4. HP1 β se lie à la chromatine péricentrique grâce à des modifications de H3.3K27 et suite à sa liaison avec un ARN.....	105
5. L'expression de H3.3K27R induit des défauts de condensation de la chromatine péricentrique paternelle.....	106
6. L'injection d'ARNdb des <i>major satellites</i> permet une relocalisation correcte de HP1 β et un développement normal des embryons exprimant H3.3K27R.....	106
7. Conclusion.....	107
II. Publication 2.....	107
C. Résultats non publiés	108
I. Chronologie de la localisation, de la condensation, et de l'acquisition des marques des histones de la chromatine péricentrique dans le zygote de souris.....	108
1. La protéine ZF est un outil de localisation des <i>major satellites</i> dans les cellules.....	108
2. Localisation et condensation des régions péricentriques au cours du développement du zygote.....	109
3. Evolution des marques des histones au niveau de la chromatine péricentrique au cours du stade zygote.....	110
II. Délocalisation des centromères dans l'embryon précoce de souris.....	111
1. Problématique et principe de l'expérience.....	111
2. L'expression de la protéine HA-ZF-émerine induit une relocalisation de la chromatine dans les cellules NIH3T3.....	112
3. La protéine HA-ZF-émerine est exprimée au stade 2-cellules et induit un faible rapprochement entre la membrane nucléaire et les centromères.....	113
4. L'expression de la protéine HA-ZF-émerine au stade 2-cellules induit un arrêt du développement précoce des embryons de souris.....	115

D. Discussion et perspectives.....	116
I. Rôle de H3.3 au cours de l'embryogenèse précoce chez la souris.....	116
1. H3.3 et H3.1, deux protéines aux rôles différents ?.....	116
a. Publication 3.....	116
b. Perspectives.....	117
2. H3.3, un variant important pour la formation de l'hétérochromatine ?.....	118
II. H3.3K27, HP1β et l'ARN des <i>major satellites</i>, trois acteurs de l'hétérochromatinisation des régions péricentromériques.....	119
1. Le zygote, une cellule particulière pour l'étude de la formation de l'hétérochromatine.....	119
2. H3.3K27, activation ou répression ?.....	120
3. Quel est le rôle de l'ARN des <i>major satellites</i> ?.....	122
III. Hétérochromatine et organisation du noyau.....	124
1. L'organisation du noyau influence-t-elle la transcription ?.....	124
2. Les marques chromatiniennes influencent-elles l'organisation du noyau ?.....	125
IV. Hétérochromatine constitutive et absence de H3K9me3.....	126
E. Matériel et méthodes.....	128
I. Transfection des cellules NIH3T3 et préparation de la chromatine pour le ChIP.....	128
II. Immunoprécipitation de la chromatine.....	129
III. Calcul des enrichissements.....	129
CONCLUSION.....	132
REFERENCES BIBLIOGRAPHIQUES.....	134

LISTE DES FIGURES ET TABLEAUX

Figure 1 :	Représentation schématique de la théorie fondamentale de la biologie moléculaire.....	17
Figure 2 :	Structure cristallographique de l'ARN polymérase II de <i>S. cerevisiae</i> en cours d'élongation à une résolution de 3,2 Å.....	18
Figure 3 :	Cycle de l'ARN polymérase II.....	19
Figure 4 :	Représentation schématique des éléments du promoteur minimal et des séquences régulatrices proximales et distales.....	21
Figure 5 :	Représentation schématique et comparaison des domaines C-terminaux des protéines TBP de différentes espèces et structure cristallographique du domaine C-terminal conservé de TBP (en bleu) associée ou non à une séquence nucléotidique (en blanc).....	22
Figure 6 :	Structure du complexe TFIID de levure résolue en microscopie électronique, en lien avec l'ADN et localisation de ses différentes sous-unités.....	23
Figure 7 :	Modèle d'assemblage du complexe de pré-initiation de la transcription avec l'ADN.....	24
Figure 8 :	Schématisation de la reconnaissance des promoteurs proximal et distal (enhancer) par des facteurs spécifiques et des conséquences sur la transcription.....	27
Figure 9 :	Répartition des différents facteurs de la famille de TBP selon les espèces et comparaison de leurs domaines C-terminaux.....	29
Figure 10 :	Représentation schématique et comparaison des domaines C-terminaux et domaines N-terminaux des protéines TBP2 de différentes espèces.....	31
Figure 11 :	Exemples des variations entre les facteurs de transcription de base utilisés dans différents types cellulaires.....	32
Figure 12 :	Représentation schématique des différents niveaux de compaction de l'ADN et photographies du "collier de perle" et de la fibre de 30 nm en microscopie électronique.....	33
Figure 13 :	Structures de la particule de cœur du nucléosome et des dimères H3/H4 et H2A/H2B liés à l'ADN résolues en cristallographie aux rayons X.....	34
Figure 14 :	Schématisation du phénomène de variégation par effet de position (PEV) chez la drosophile.....	36

Figure 15 :	Principales modifications post-traductionnelles des histones de cœur.....	37
Figure 16 :	Distribution des modifications des histones au sein du génome.....	40
Figure 17 :	Déposition et rôles de H3K4me3 à la chromatine.....	41
Figure 18 :	Déposition et rôles de H3K27me3 à la chromatine.....	46
Figure 19 :	Schéma des différents variants humains de l'histone H3.....	48
Figure 20 :	Représentation schématique des répétitions en tandem des régions péricentriques et centromériques.....	52
Figure 21 :	Exemples de rôles des longs ARNnc dans le recrutement des complexes polycomb en cis dans le cas de Xist, et en trans dans le cas de HOTAIR.....	54
Figure 22 :	Mécanisme d'ARN interférence à l'origine de l'hétérochromatinisation des régions péricentriques chez la levure <i>S. pombe</i>	55
Figure 23 :	Rétrocontrôles positifs permettant la propagation de H3K9me3 et H3K27me3 après réplication.....	58
Figure 24 :	Schéma récapitulatif des étapes de l'ovogenèse, de la folliculogenèse et du développement précoce de l'embryon.....	60
Figure 25 :	Comparaison entre les périodes de transition entre un contrôle maternel et embryonnaire du développement (MZT) chez différents organismes modèles.....	67
Figure 26 :	Schéma récapitulatif des profils d'expression de TBP2 (en rose, en haut) et TBP (en bleu, en bas) dans les ovocytes et les cellules folliculaires au cours de la folliculogenèse et dans les embryons précoces.....	68
Figure 27 :	Reprogrammation de la méthylation de l'ADN dans les cellules germinales primordiales et au cours du développement préimplantatoire de souris.....	71
Figure 28 :	Transition entre le génome des spermatides condensé autour d'histones et le génome des spermatozoides condensé autour de protamines.....	72
Figure 29 :	Evolution de la localisation et du niveau d'expression des répétitions péricentromériques au cours du développement précoce chez la souris.....	77
Figure 30 :	TBP2 n'est pas détectable dans les ovocytes au stade pachytène tardif/diplotène précoce.....	84
Figure 31 :	Les ovaires n'exprimant pas TBP2 ne présentent pas de défauts histologiques évidents.....	85
Figure 32 :	La localisation de γ H2AX et VASA n'est pas différente entre de ovocytes exprimant ou non TBP2.....	86

Figure 33 :	Les ovaires de souris et TBP2 ^{-/-} âgées de 16 mois ne contiennent plus de structures folliculaires mais seulement des structures anormales.....	87
Figure 34 :	Schématisation et résultats de différentes mises au point de l'expérience d'immunoprécipitation de TBP2 sur les extraits nucléaires d'ovaires de souris sauvages.....	88
Figure 35 :	Schématisation et résultats de différentes mises au point de l'expérience d'immunoprécipitation de TBP2 sur des cellules HEK exprimant transitoirement TBP2.....	89
Figure 36 :	La protéine Flag-ZF-GFP se localise aux régions de chromatine très condensée et lie spécifiquement les <i>major satellites</i> dans les cellules NIH3T3.....	108
Figure 37 :	Condensation de la chromatine péricentrique et localisation des <i>major satellites</i> au cours du développement du zygote.....	109
Figure 38 :	Localisation de HP1b, H3K27me1 et des <i>major satellites</i> au cours du développement zygotique.....	110
Figure 39 :	Localisation de H3K27me3 et des <i>major satellites</i> au cours du développement zygotique.....	110
Figure 40 :	La construction HA-ZF-émerine transfectée dans des cellules NIH3T3 ou injectée à des embryons sous forme d'ARNm reconnaît spécifiquement les <i>major satellites</i>	111
Figure 41 :	Les cellules NIH3T3 exprimant la protéine HA-ZF-émerine montrent une relocalisation globale de leur ADN à la périphérie du noyau et une diminution du marquage CENP-A.....	112
Figure 42 :	HA-ZF-émerine est exprimée et incorporée dans le noyau à partir du stade 2-cellules, HA-émerine n'est pas détectée dans le noyau.....	113
Figure 43 :	Les embryons exprimant HA-ZF-émerine présentent une faible délocalisation des centromères et une modification de la forme de leurs noyaux.....	114
Figure 44 :	L'expression de HA-ZF-émerine induit des défauts lors du développement préimplantatoire.....	115
Tableau 1 :	Quantités d'ARN totaux, ARN polyadénylés et ARN ribosomiaux (ARNr) à différents stades du développement.....	65

LISTE DES ABREVIATIONS

ac	acétylation
ADN	Acide DésoxyriboNucléique
ADNc	ADN complémentaire
ARC	<i>ARgonaute Chaperonne</i>
ARN	Acide RinoNucléique
ARNdb	ARN double brin
ARNi	ARN interférent
ARNm	ARN messenger
ARNnc	ARN non codant
ARN pol II	ARN polymérase II
ASF1	<i>Anti-Silencing Function 1</i>
ATRX	<i>Alpha Thalassemia/mental Retardation syndrome X-linked</i>
BRE	<i>TFIIB recognition element</i>
BrUTP	BromoUridine-TriPhosphate
CAF-1	<i>Chromatin Assembly Factor 1</i>
CAGE	<i>Cap Analysis Gene Expression</i>
ChIP	<i>Chromatin ImmunoPrecipitation</i>
CENP-A	<i>CENtromere Protein A</i>
Clr4	<i>C-type lectin-related protein 4</i>
CMV	CytoMégaloVirus
CTD	<i>Carboxy-Terminal Domain</i>
Da	Dalton
DAPI	4,6-DiAmino-2-PhénylIdine
DAXX	<i>Death domain Associated protein</i>
DCE	<i>Downstream Core Element</i>
DPE	<i>Downstream Promoter Element</i>
DTT	DiThioThreitol
E1,5	1,5ème jour embryonnaire
EDTA	<i>Ethylene Diamine Tetraacetic Acid</i>
EED	<i>Embryonic Ectoderm Development</i>
EGTA	<i>Ethylene Glycol Tetraacetic Acid</i>
ESET	<i>ERG-associated protein with SET domain</i>
EZH2	<i>Enhancer of Zeste Homolog 2</i>
FISH	<i>Fluorescent In Situ Hybridization</i>
FSH	<i>Follicle Stimulating Hormon</i>
GFP	<i>Green Fluorescent Protein</i>
GST	<i>Glutathion-S-Transferase</i>
GTF	<i>General Transcription Factor</i>
GV	<i>Germinal Vesicle</i>
GVBD	<i>Germinal Vesicle BreakDown</i>
H3K27	Lysine 27 de l'histone H3
HCG	<i>Human Chorionic Gonadotropin</i>
HIRA	<i>Hlstone cell cycle Regulation defective homolog A</i>
HMT	<i>Histone Methyl Transferase</i>
HOX	Homeobox
HP1	<i>Heterochromatin Protein 1</i>
h pHCG	heures post-HCG
IAP	<i>Intracisternal A Particle</i>

ICS	Institut Clinique de la Souris
Inr	Initiateur
IP	ImmunoPrécipitation
iPS	<i>induced Pluripotent Stem cells</i>
LH	Luteinizing Hormon
MCI	Masse Cellulaire Interne
me1/2/3	Mono-/Di-/Triméthylation
MEF	<i>Mouse Embryonic Fibroblast</i>
MLL	<i>Mixed-Lineage Leukaemia</i>
MTE	<i>Motif Ten Element</i>
MuDPIT	<i>MultiDimensional Protein Identification Technology</i>
MZT	<i>Maternal to Zygotic Transition</i>
NSN	<i>Non-Surrounded Nucleolus</i>
nt	nucléotide
NuRD	<i>Nucleosome Remodeling and histone Deacetylation</i>
pb	paire de base
PBS	<i>Phosphate Buffered Saline</i>
PCNA	<i>Proliferating Cell Nuclear Antigen</i>
PEV	<i>Position Effect Variegation</i>
PGC	<i>Primordial Germ Cell</i>
PHD	<i>Plant Homeo Domain</i>
PIC	<i>Pre-initiation Complex</i>
PMSF	<i>PhenylMethylSulfonyl Fluoride</i>
PN1	<i>ProNuclear stage 1</i>
polyA	poly-adénylé
PRC1/2	<i>Polycomb Repressive Complex 1/2</i>
RBBP4/7	<i>RetinoBlastoma Binding Protein</i>
RdRC	<i>RNA dependant RNA polymerase Complex</i>
RING	<i>Really Interesting New Gene</i>
RITS	<i>RNA induced transcriptional silencing</i>
RT-PCR	<i>Reverse Transcription Polymerase Chain Reaction</i>
SET	<i>Su(var)3-9, Enhancer of zeste, Trithorax</i>
SETDB2	<i>SET domain, bifurcated 1</i>
SINE	<i>Short INterspersed Element</i>
SN	<i>Surrounded nucleolus</i>
SUV39H1/2	<i>SUppressor of Variegation 3-9, Homolog 1/2</i>
SUZ12	<i>SUppressor of Zeste 12</i>
TAF	<i>TBP associated Factor</i>
TBP	<i>TATA Binding Protein</i>
TBP2	<i>TATA Binding Protein 2</i>
TF	<i>Transcription Factor</i>
TLF	<i>TBP-Like Factor</i>
TRF1	<i>TBP-Related Factor 1</i>
TRX	<i>Trithorax</i>
TSS	<i>Transcription Start Site</i>
Xi	Chromosome X inactif
Xist	<i>X inactivation specific transcript</i>
ZGA	<i>Zygotic Genome Activation</i>
ZF	<i>Zinc Finger, abréviation utilisée ici pour designer la protéine reconnaissant les major satellites</i>

Partie 1

INTRODUCTION

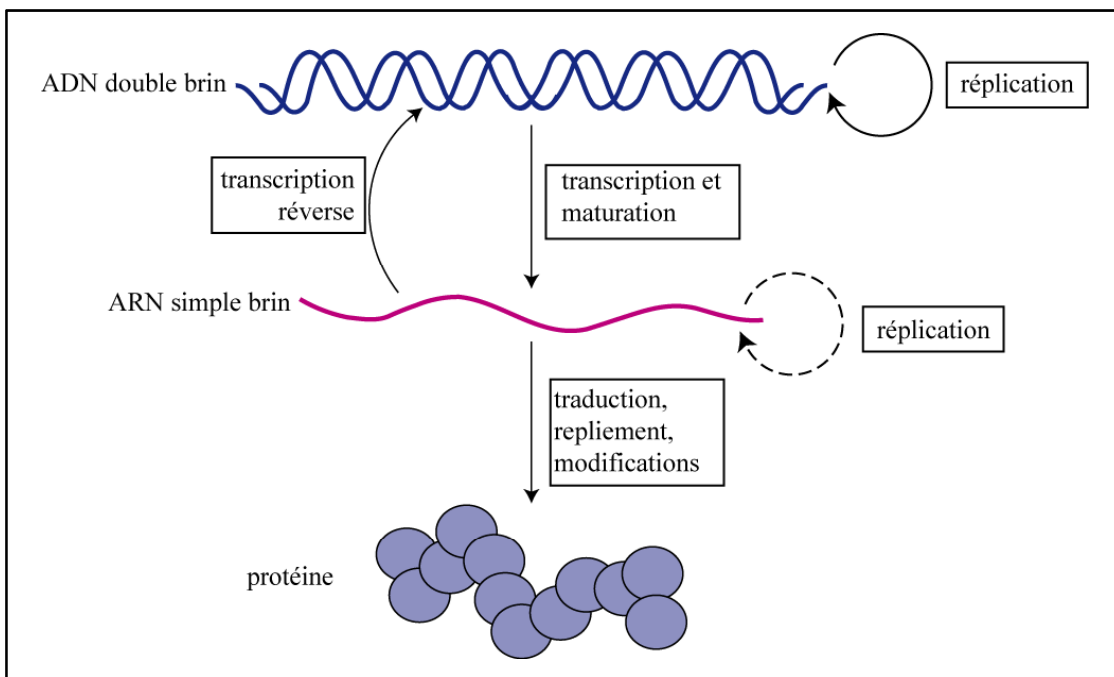


Figure 1 : Représentation schématique de la théorie fondamentale de la biologie moléculaire.

A. La transcription par l'ARN polymérase II

En 1958 Francis Crick affirme qu'il existe un flux d'information séquentielle, résidu par résidu, entre l'ADN et les protéines. Il ajoute que ce flux d'information fait intervenir une autre catégorie d'acide nucléique, les ARN. Il fait également état de la possibilité d'une réplication de l'ADN, de l'ARN et d'un transfert de l'ARN vers l'ADN, mais il stipule que des transferts de protéines vers protéines et de protéines vers ADN sont impossibles (Crick, 1970). Ce qui n'était à l'époque qu'une hypothèse a été établi par la suite comme le dogme central de la biologie moléculaire (figure 1) puis a été démontré et enrichi par des découvertes ultérieures.

Dans cette partie, nous nous intéresserons à la première étape de ce transfert d'information, la transcription de l'ADN en ARN, et sa régulation. Chez les eucaryotes, selon le type d'ARN transcrit, cette étape est assurée par trois différentes ARN polymérases, les ARN polymérases I, II, III. Une ARN polymérase IV a également été caractérisée chez les plantes (Herr et al., 2005) (Zhang et al., 2007). L'ARN polymérase II (ARN pol II) est responsable de la transcription des gènes codant pour des protéines ainsi que celle de certains petits ARN nucléaires, on parle de gènes de classe II. Nous détaillerons ici certains aspects de la transcription des gènes de classe II ainsi que certains facteurs responsables de sa régulation.

I. Les acteurs de la transcription de base

1. L'ARN polymérase II

a. Composition et structure

L'ARN pol II, très conservée entre tous les eucaryotes, est un complexe de 12 sous-unités d'un poids moléculaire total de 550 kDa. Les sous-unités, les protéines Rpb (*RNA polymerase B*) sont numérotées selon leur poids moléculaire, Rpb1 étant la plus grosse et Rpb12 la plus petite.

Des études cristallographiques de l'ARN pol II de levure libre et pendant l'élongation ont permis de comprendre et de visualiser certains mécanismes de la transcription (Cramer et al., 2000) (Cramer et al., 2001) (Gnatt et al., 2001). Le centre catalytique de l'enzyme est formé par les deux sous-unités Rpb1 et Rpb2. Elles forment un sillon chargé positivement

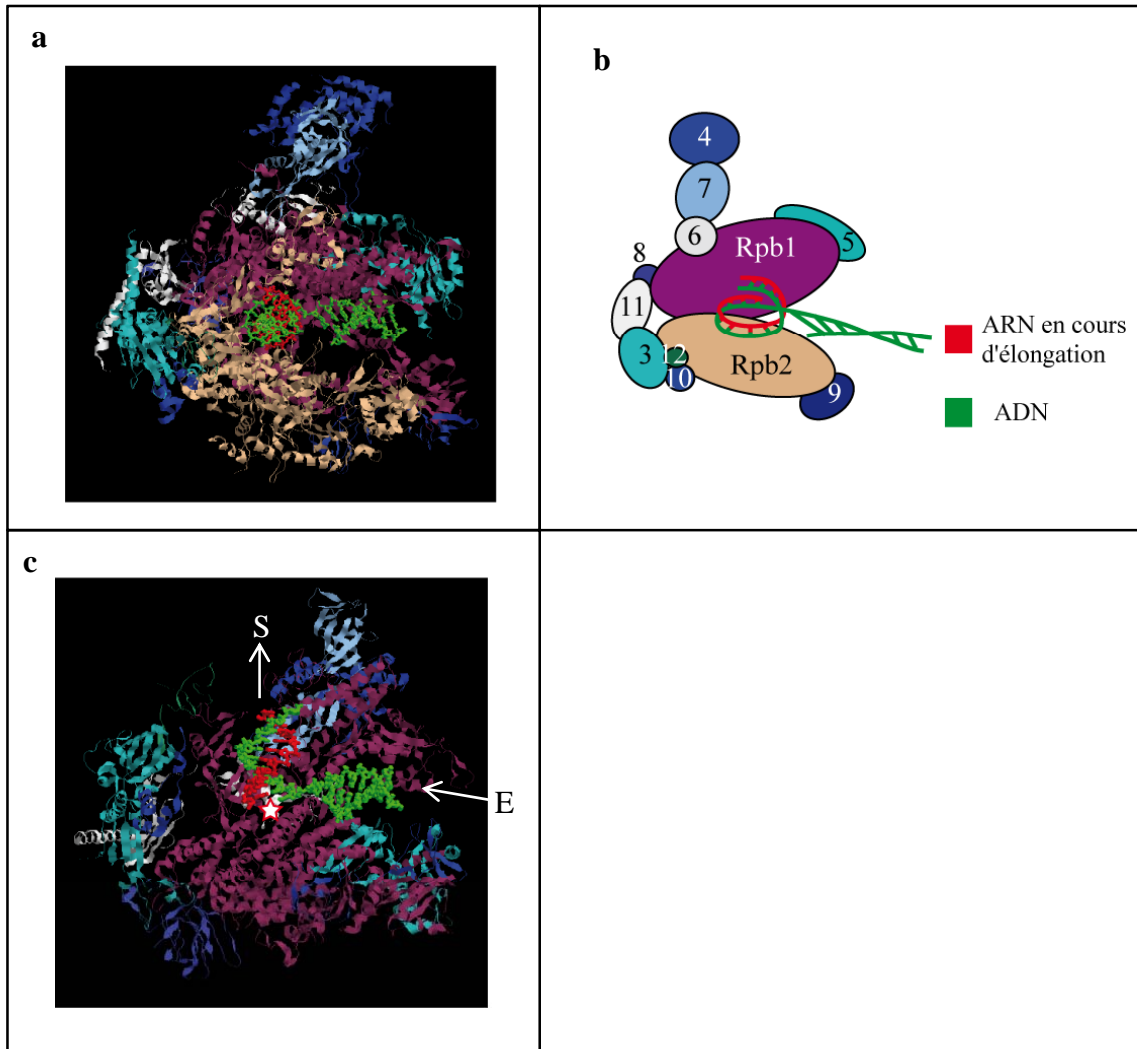


Figure 2 : Structure cristallographique de l'ARN polymérase II de *S. cerevisiae* en cours d'élongation à une résolution de 3,2 Å. **a** : Enzyme à une résolution de 3,2 Å. Chaque sous-unité de l'ARN pol II est représentée en mode ruban et par une couleur différente. L'ADN et l'ARN sont représentés en mode bâtonnets. Le numéro de chaque sous-unité est indiqué sur le schéma récapitulatif de la structure en **b**. **c** : Structure de l'ARN pol II vue du dessous par rapport à la vue en **a** et **b**. La sous-unité Rpb2 n'est pas représentée pour plus de clarté. Les flèches représentent les sites d'entrée (E) et de sortie (S) de l'ADN. Le site catalytique et la position de l'ion Mg^{2+} sont marqués par une étoile rouge et blanche. Le code couleur des sous-unités du complexe est le même qu'en **a** et **b**. La représentation a été réalisée d'après la structure de Sydow et al., 2009.

dans lequel vient se loger l'ADN. Au fond du sillon, au sein du site actif a lieu la formation d'un pont phosphodiester entre l'extrémité 3'-OH de l'ARN en croissance et un ribonucléotide triphosphate complémentaire de l'ADN matrice. Cette réaction nécessite le passage par un état de transition stabilisé par ion Mg^{2+} présent au site actif (figure 2). Pour permettre l'entrée des ribonucléotides triphosphates dans le sillon, un pore relie ce dernier à l'arrière de l'enzyme. La matrice ADN et le transcrit en cours d'élongation sont maintenus dans le sillon grâce à une pince formée par les extrémités N- et C-terminales de Rpb1 et par l'extrémité C-terminale de Rpb2. Cette pince, se refermant sur l'hétéroduplexe ADN/ARN pendant l'élongation, est à l'origine de la processivité de l'ARN pol II.

L'ARN pol II possède très peu de domaines non structurés. Un de ces domaines, l'extrémité C-terminale de la sous-unité Rpb1, joue un rôle important dans la régulation des différentes étapes de la transcription. Ce domaine C-terminal, nommé domaine CTD (*carboxy-terminal domain*), est composé de nombreuses répétitions d'une séquence de 7 acides aminés (Y-S-P-T-S-P-S). Il est hautement conservé de la levure à l'homme et le nombre de répétitions varie de 26 chez la levure à 52 chez les mammifères. L'importance de ce domaine réside dans le fait que l'heptapeptide répété peut être phosphorylé, de manière réversible, sur 5 de ses 7 acides aminés : la tyrosine (Y), la thréonine (T) et les sérines (S). Selon son degré de phosphorylation, l'ARN pol II peut alors recruter ou non certains facteurs de maturation de l'ARN, des facteurs d'élongation ou de terminaison permettant ainsi le déroulement de la transcription comme nous le verrons par la suite (pour une revue, Phatnani and Greenleaf, 2006).

b. Le cycle de l'ARN polymérase II : les différentes étapes de la transcription

- L'initiation

Pour initier la transcription, l'ARN pol II s'associe à d'autres complexes protéiques, les facteurs généraux de transcription (GTF : *general transcription factor*). Six complexes principaux font partie de ces GTF : TFIIA, TFIIB, TFIID, TFIIE, TFIIF et TFIIH, ils s'assemblent au niveau du promoteur minimal du gène, et y recrutent l'ARN pol II formant ainsi le complexe de pré-initiation de la transcription (PIC : *pre-initiation complex*). Au moment de la formation du PIC, l'ADN est sous forme double brin. Après l'assemblage des différents facteurs et grâce aux activités hélicase de TFIIH, l'ADN s'ouvre entre les positions -9 et +2 et l'ARN pol II peut catalyser la synthèse des premiers ponts phosphodiester formant alors un transcrit de quelques nucléotides.

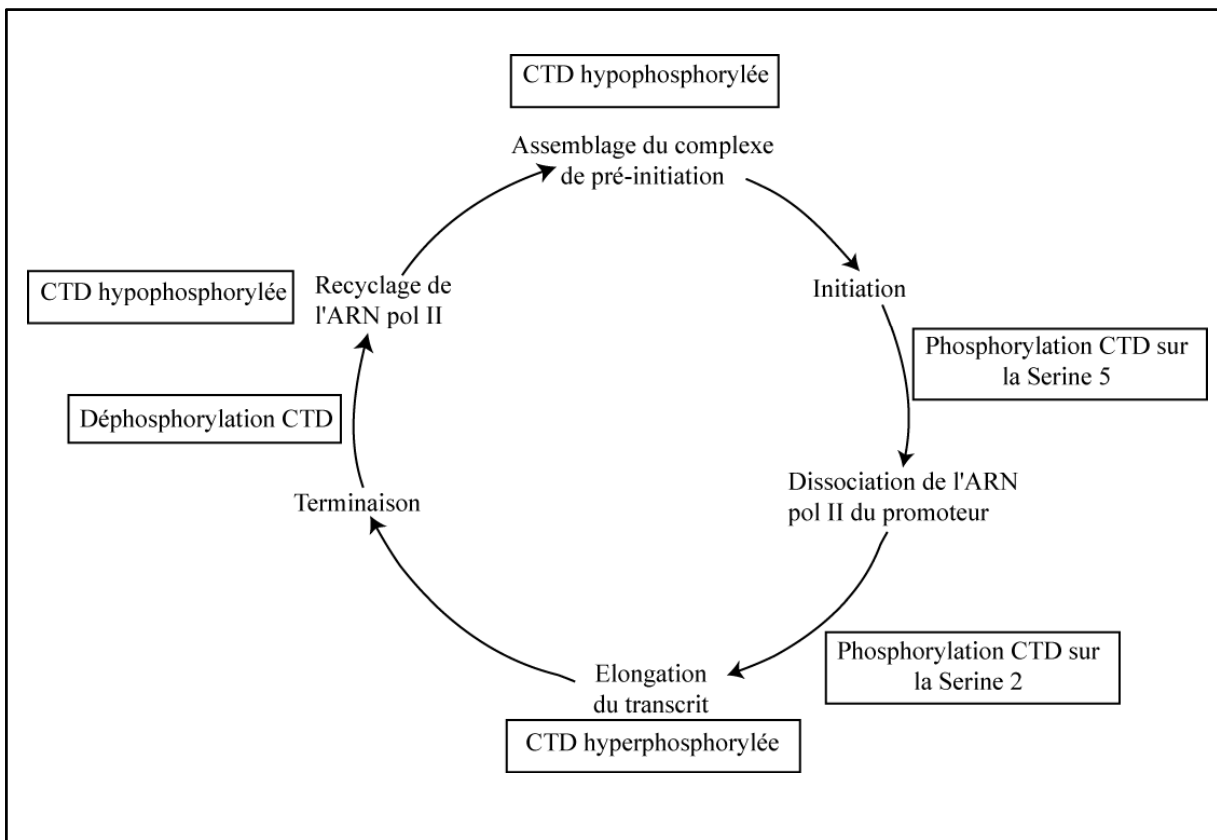


Figure 3 : Cycle de l'ARN polymérase II. Adapté d'après Svejstrup et al., 2004.

Au début de la transcription, le complexe d'élongation est instable et de nombreuses synthèses sont avortées. Lors de son recrutement dans le PIC, le domaine CTD de l'ARN pol II est hypophosphorylé. Il peut alors recruter le complexe Médiateur, un complexe protéique qui, entre autres, stimule l'activité kinase de TFIIF qui, grâce à sa sous-unité kinase CDK7, phosphoryle jusqu'à une centaine de positions du domaine CTD de l'ARN pol II.

- La dissociation du promoteur

Rapidement, l'ARN pol II doit s'éloigner du promoteur et donc se dissocier des GTF pour que la transcription soit efficace. Après la synthèse des premiers nucléotides et grâce à l'action du complexe Médiateur et de TFIIF, l'ARN pol II est hyperphosphorylée. Cette hyperphosphorylation du domaine CTD lui fait perdre son affinité pour les GTF et pour le complexe Médiateur, elle peut donc se dissocier du promoteur et commencer la phase d'élongation de la transcription.

- La phase d'élongation

Pendant le processus d'élongation, l'ARN pol II n'a pas une vitesse constante, elle peut être mise en pause pendant des périodes plus ou moins longues à cause de structures inhabituelles de l'ADN, ou de protéines liées à l'ADN comme pour permettre le passage de l'ADN polymérase lors de la réplication par exemple. En même temps que l'ARN pol II progresse sur l'ADN, l'ARN s'allonge et le transcrit subit une maturation co-transcriptionnelle qui consiste en l'épissage, retrait des introns et ligations des exons entre eux grâce à un complexe ribonucléoprotéique, et en l'acquisition de la coiffe à son extrémité 5'-phosphate, un résidu 5-méthylguanosine lié par un pont 5'-5'-triphosphate au premier ribonucléotide transcrit.

Encore une fois ici, l'état de phosphorylation du domaine CTD de l'ARN pol II est important : les protéines responsables de la formation de la coiffe interagissent avec le domaine CTD seulement si celui-ci est hyperphosphorylé, c'est-à-dire si l'initiation de la transcription a eu lieu correctement. Des expériences d'immunoprécipitation de la chromatine (ChIP : *chromatin immunoprecipitation*) ont montré une différence entre la phosphorylation de la sérine 2 (Ser2) et de la sérine 5 (Ser5) du motif répété au cours de la transcription : la Ser5 est principalement phosphorylée lors de la transition entre l'initiation et l'élongation alors que la Ser2 l'est au moment de l'élongation (Cho et al., 2001) (Palancade and Bensaude, 2003). Cela montre que le processus de phosphorylation du CTD est hautement régulé durant les différentes phases de la transcription.

- La phase de terminaison et le recyclage de l'ARN pol II

Suite à des signaux spécifiques sur l'ADN, dont le signal de polyadénylation (AAUAA) et des structures spécifiques de l'ADN au niveau du terminateur, la transcription s'arrête. Le transcrit est clivé et polyadénylé et l'ARN pol II peut être recyclée pour être utilisée dans un nouveau cycle de transcription. Pour que le recyclage de l'ARN pol II soit efficace, et pour que le complexe puisse à nouveau être réintégré dans un PIC, son domaine CTD doit être déphosphorylé. (pour des revues, Orphanides et al., 1996; Svejstrup, 2004) (figure 3)

L'ARN messenger (ARNm), correctement coiffé, polyadénylé, et épissé, peut alors être transféré dans le cytoplasme pour être traduit par les ribosomes.

Dans la suite de cette partie, je me consacrerai plus spécifiquement à l'étape d'initiation de la transcription et à ses acteurs. Je m'attarderai en particulier sur la fonction de la protéine TBP (*TATA binding protein*), la première protéine impliquée dans l'initiation de la transcription au côté de l'ARN pol II à avoir été découverte.

2. Les facteurs généraux de la transcription : exemple de TBP

Seule, l'ARN pol II peut transcrire une séquence d'ADN en ARN, mais elle n'est pas capable d'initier spécifiquement la transcription *in vitro*, i.e au niveau du site d'initiation de la transcription (TSS : *transcription start site*). L'étude d'extraits cellulaires totaux humains qui, ajoutés à la réaction, rendent la transcription *in vitro* spécifique (Weil et al., 1979), a permis de mettre en évidence l'existence des GTF et en particulier de la protéine TBP (Reinberg et al., 1987) (Reinberg and Roeder, 1987) (Cavallini et al., 1988). Lors de l'initiation de la transcription, les GTF s'assemblent donc sur l'ADN, formant le PIC au niveau d'une séquence proche du TSS, le promoteur minimal et y recrutent l'ARN pol II.

En nous intéressant plus en détail au cas de la protéine TBP au sein du complexe, nous verrons comment s'effectue l'assemblage du PIC au promoteur minimal.

a. Le promoteur minimal

Le promoteur minimal est défini comme le fragment d'ADN minimal, suffisant pour diriger l'initiation de la transcription à un taux basal, par l'ARN Pol II, *in vitro*, sur un ADN ne contenant un seul TSS connu. Selon cette définition, le promoteur minimal ne serait

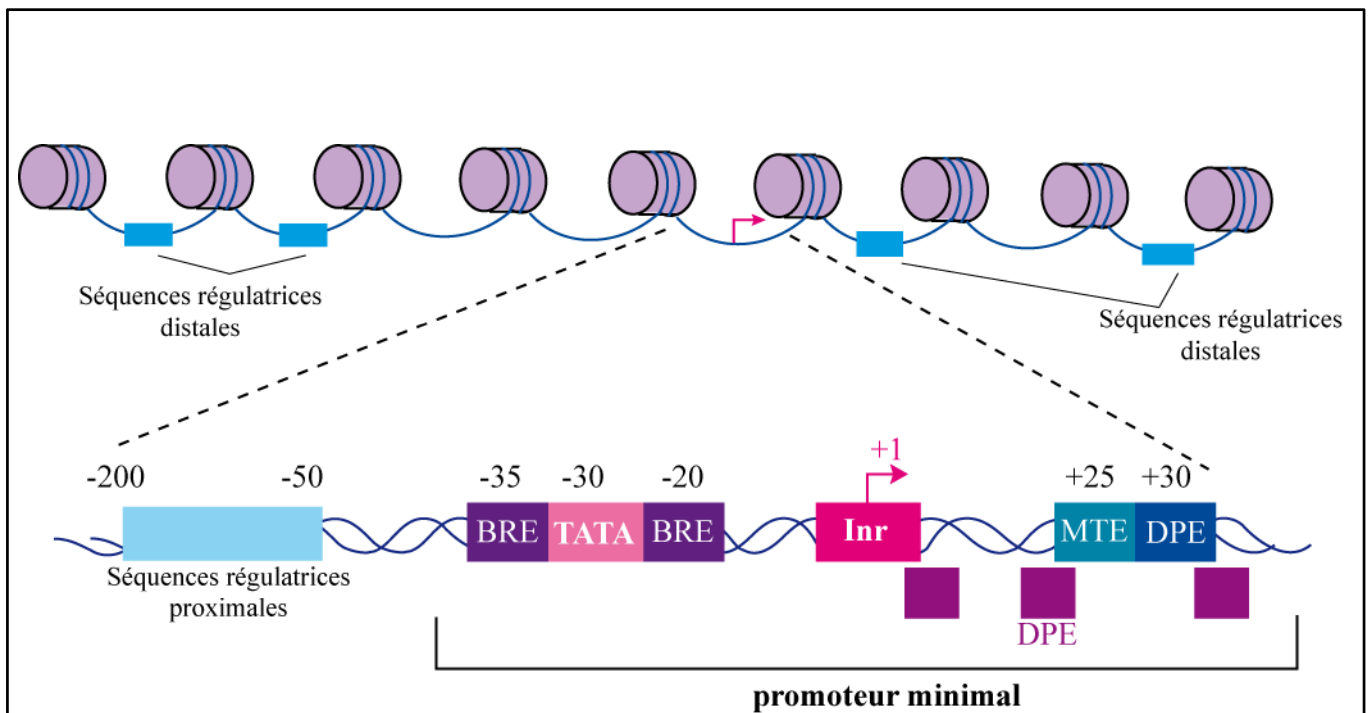


Figure 4 : Représentation schématique des éléments du promoteur minimal et des séquences régulatrices proximales et distales. Les séquences régulatrices les plus communes potentiellement retrouvées au sein d'un promoteur minimal sont présentées ici. On distingue au niveau du promoteur minimal la boîte TATA, l'élément BRE (élément de réponse à TFIIB), l'initiateur (Inr), l'élément MTE (*Motif Ten Element*), l'élément DPE (*Downstream Promoter Element*) et l'élément DCE (*Downstream Core Element*). Les positions des séquences régulatrices proximales et distales sont également notées. Les chiffres correspondent à la distance des séquences par rapport au TSS noté par une flèche.

composé que de deux séquences régulatrices, la boîte TATA et/ou l'élément initiateur (Inr). La boîte TATA, également nommée boîte de Goldberg-Hogness, a été le premier élément du promoteur minimal à être mis en évidence chez les eucaryotes, grâce à la comparaison de séquences en 5' de gènes codant pour des protéines de drosophile, de mammifères, et de virus. La séquence TATAA était présente entre 25 et 30 nucléotides (nt) en amont du TSS de nombreux gènes analysés (Breathnach and Chambon, 1981), et sa mutation entraînait la diminution de l'activité des promoteurs de ces gènes ou un changement de la localisation du TSS (Grosschedl and Birnstiel, 1980) (Wasylyk and Chambon, 1981). Chez les métazoaires, plus généralement, la séquence consensus de la boîte TATA est T-A-T-A-(A/T)-A-(A/T). Elle se situe à une trentaine de paires de bases en amont du site d'initiation et constitue le site de liaison de la protéine TBP au promoteur (Horikoshi et al., 1989) (Kim et al., 1993).

L'élément initiateur a, lui, été mis en évidence dans des promoteurs ne contenant pas de boîte TATA (Smale and Baltimore, 1989). Il s'agit d'un motif riche en pyrimidine dont la séquence consensus dans les cellules de mammifères est (C/T)-(C/T)-A₊₁-N-(T/A)-(C/T)-(C/T) (Lo and Smale, 1996), où l'adénine est le TSS. De nombreuses protéines semblent pouvoir se lier à l'initiateur comme par exemple des facteurs associés à TBP (TAF : *TBP associated factors*), TAF1 et TAF2 (Chalkley and Verrijzer, 1999), ainsi que l'ARN pol II (Carcamo et al., 1991).

In vivo, le promoteur minimal semble plus complexe, d'autres éléments pouvant lui être nécessaires. On peut alors définir le promoteur minimal comme une séquence d'approximativement une centaine de nucléotides centrée sur le TSS (de -50 à +50, +1 définissant la position du TSS). Les autres séquences régulatrices pouvant être présentes sont par exemple l'élément de reconnaissance de TFIIB (BRE), la séquence DPE (*Downstream Promoter Element*), le MTE (*Motif Ten Element*) et le DCE (*Downstream Core Element*) (pour une revue Juven-Gershon et al., 2006) pour les mieux définies (figure 4). Ces éléments, que nous ne détaillerons pas par ailleurs, ont tous une position précise par rapport au TSS et sont des sites de recrutement des facteurs généraux de transcription. L'ensemble des séquences régulatrices présentes au niveau du promoteur minimal définit la position du TSS et le sens de la transcription. Il est important de noter qu'aucun élément du promoteur minimal n'est universel, commun à tous les gènes.

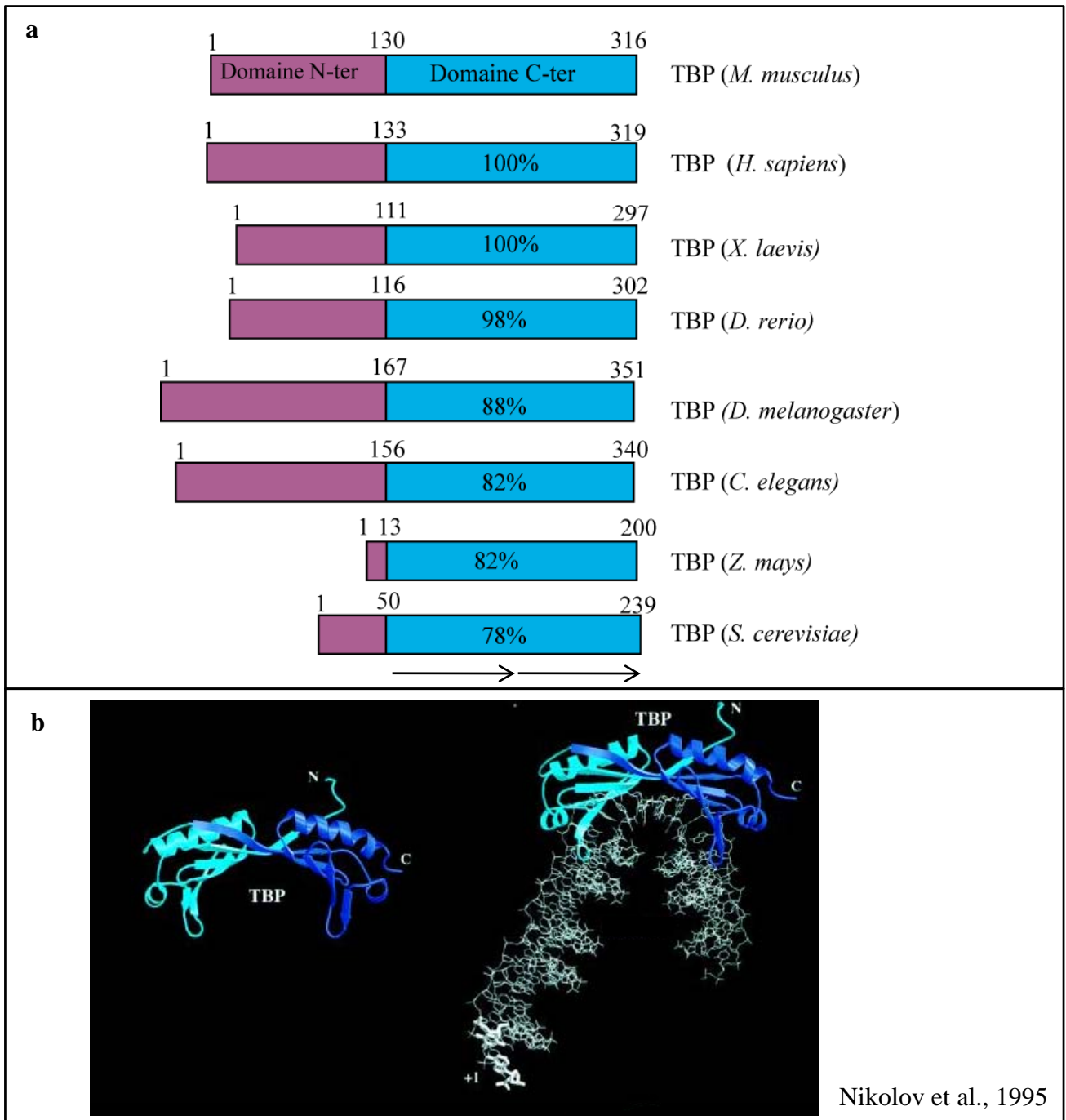


Figure 5 : Représentation schématique et comparaison des domaines C-terminaux des protéines TBP de différentes espèces (a) et structure cristallographique du domaine C-terminal conservé de TBP (en bleu) associée ou non à une séquence nucléotidique (en blanc) (b). a : Les nombres indiquent la position des acides aminés. Les domaines C-terminaux des protéines TBP, contenant le domaine de liaison à l'ADN, de chaque espèce sont comparés au même domaine chez la souris, le résultat est donné en pourcentage d'identité. La répétition directe qui forme la structure en selle de cheval du domaine conservé est indiquée par les flèches. b : La structure cristallographique de la protéine TBP humaine est réalisée à une résolution de 2,7 Å. La protéine TBP est représentée en mode ruban et les extrémités N- et C-terminales sont notées N et C respectivement. Le site d'initiation de la transcription est noté +1 sur l'ADN, la liaison de TBP à l'ADN se fait sur une boîte TATA. La répétition directe qui forme la structure en selle de cheval du domaine conservé est indiquée par les flèches.

b. TBP, découverte et structure

L'ARN pol II ne peut initier la transcription spécifiquement au TSS sans la présence de GTF. Lors de l'étude des extraits cellulaires, qui, ajoutés à l'ARN pol II rendaient la réaction de transcription *in vitro* spécifique (Weil et al., 1979), les protéines de ces extraits ont été séparées par chromatographie en 4 fractions, A, B, C et D (Matsui et al., 1980). La fraction D contenait un facteur, nommé alors TFIID, liant la boîte TATA (Nakajima et al., 1988). La purification, à partir de cette fraction, d'un complexe homogène présentant l'activité de TFIID s'est révélée impossible chez les vertébrés pendant très longtemps. Ce n'est que plus tard qu'une protéine de levure, TBP, capable de se substituer à l'activité du facteur TFIID dans un système de transcription *in vitro*, a été identifiée (Cavallini et al., 1988) (Cavallini et al., 1989). Grâce à la séquence de l'ADNc (ADN complémentaire) codant pour la protéine TBP de levure, des gènes codant pour TBP ont pu être identifiés chez différentes espèces, de l'homme aux archaebactéries (Horikoshi et al., 1989) (Eisenmann et al., 1989) (Cavallini et al., 1989).

TBP est un facteur de transcription universel, dont le poids moléculaire varie de 22 kDa chez *A. thaliana* à 38 kDa chez l'homme. La comparaison des séquences protéiques de TBP dans de nombreux organismes permet de subdiviser la protéine en deux domaines. (1) La partie C-terminale de la protéine TBP, le domaine de cœur, est hautement conservée entre les espèces, elle l'est à 100% entre l'homme et la souris, à 82% entre l'homme et le maïs, et à 74% entre l'homme et la levure *S. cerevisiae* (figure 5a). Elle contient deux copies imparfaites d'une séquence de 61-62 acides aminés. Ce domaine C-terminal est nécessaire et suffisant pour sa liaison avec l'ADN et pour l'assemblage de l'appareil de transcription de base. Il a une forme en selle de cheval avec deux boucles en forme d'étrier. Dix feuilletts *beta* antiparallèles forment la partie concave, zone de liaison à l'ADN, tandis que quatre hélices *alpha* permettent son interaction avec d'autres facteurs de transcription sur la partie convexe de la protéine. TBP reconnaît le petit sillon de l'ADN au niveau de la boîte TATA, l'ADN lié suit alors la courbure imposée par la concavité de la protéine, le petit sillon au niveau de la boîte TATA se trouvant partiellement déroulé, et le grand sillon replié sur lui-même (Nikolov et al., 1996) (figure 5b). Cette courbure importante de l'ADN lors de la liaison à TBP se fait dans la direction inverse de la direction d'enroulement de l'ADN autour des histones. Elle empêche donc l'ADN de se compacter autour des nucléosomes, le rendant accessible à la machinerie de transcription (Workman and Roeder, 1987). Il a également été suggéré que

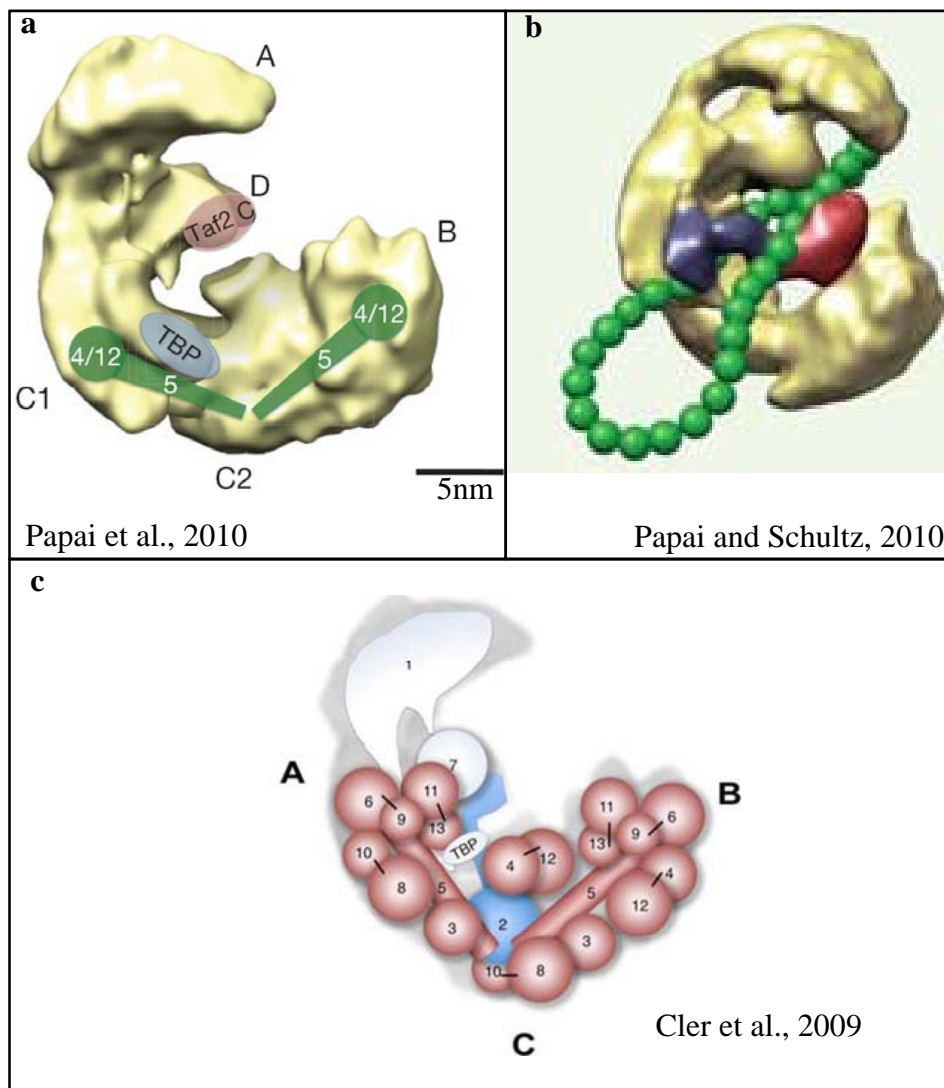


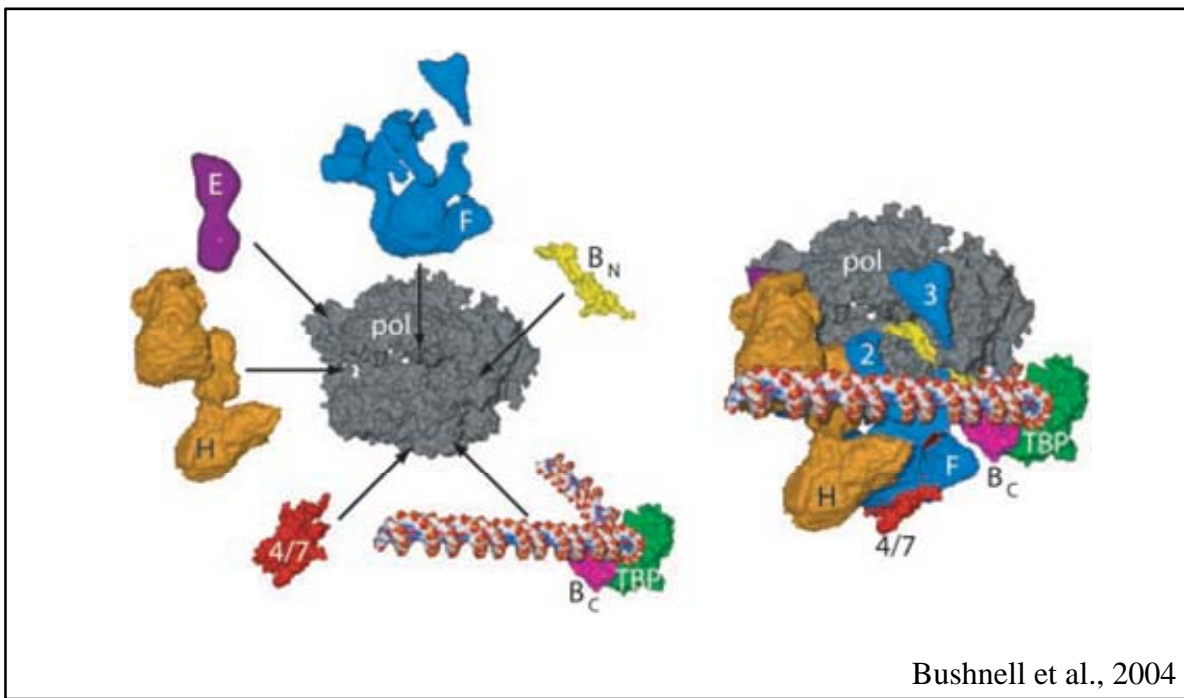
Figure 6 : Structure du complexe TFIID de levure résolue en microscopie électronique (a), en lien avec l'ADN (b) et localisation de ses différentes sous-unités (c). a : Quatre lobes majeurs (A, B, C1 et C2) de TFIID sont notés ici ainsi que la position des Taf2, Taf4, Taf5, Taf12 et TBP. b : TFIID est montré en lien avec un activateur, Rap1, en rouge et avec TFIIA, en bleu. Le complexe TFIID, TFIIA, Rap1 interagit avec l'ADN schématisé en vert. c : Représentation schématique de la localisation des TAF et de TBP dans le complexe TFIID de levure. Les positions approximatives ont été déterminées par immunomarquage. Les lignes noires représentent les interactions connues entre les TAF possédant un domaine « *histone fold* ». Les lobes A, B et C sont notés.

cette courbure pouvait être responsable du rapprochement de facteurs se situant en amont et en aval du promoteur, comme des facteurs de transcription par exemple. La partie convexe de TBP interagit avec les TAF, formant alors le complexe TFIID, qui interagit à son tour avec TFIIA, TFIIB et l'ARN pol II ainsi qu'avec d'autres facteurs modulant la transcription des gènes de classe II. (2) L'extrémité N-terminale de TBP est quant à elle très peu conservée d'une espèce à l'autre. Chez l'homme et chez la souris, elle contient une répétition de codons glutamine, motif commun à différentes protéines impliquées dans la transcription comme les protéines homeobox (HOX) par exemple. Le rôle de ces répétitions n'est pas connu mais un trop grand nombre de répétitions au sein de cette région peut être à l'origine de différentes maladies. Chez la levure, l'étude de la liaison de la protéine TBP, contenant ou non sa partie N-terminale, à la boîte TATA a révélé que ce domaine est responsable d'une modulation de l'efficacité de la liaison entre TBP et l'ADN (Gupta et al., 2010). Le domaine N-terminal peut, soit être replié dans la partie concave de TBP et déstabiliser ainsi la liaison du domaine C-terminal avec la boîte TATA, soit être exclu du domaine de liaison à l'ADN et ainsi créer une surface d'interaction supplémentaire avec d'autres protéines régulatrices (Khrapunov and Brenowitz, 2007). Chez la souris, la délétion de la partie N-terminale de TBP entraîne un fort taux de mortalité des embryons homozygotes, non à cause d'un problème de développement, mais à cause d'un rejet des embryons par le système immunitaire maternel (Hobbs et al., 2002). TBP, et plus particulièrement sa partie N-terminale, jouerait donc un rôle dans la transcription par le placenta de gènes permettant à l'embryon d'échapper au rejet immunitaire maternel. Les mécanismes sous-tendant ce rôle surprenant de la partie N-terminale de TBP ne sont pas connus.

c. TBP au cœur de l'initiation de la transcription

Bien que TBP soit suffisant à rendre une transcription *in vitro* spécifique, il n'est pas suffisant pour que la transcription réponde à l'action des activateurs, des protéines permettant d'augmenter spécifiquement le taux de transcription d'un gène (Pugh and Tjian, 1990) (Smale et al., 1990). *In vivo*, TBP n'agit donc pas seul au niveau du promoteur minimal pour initier la transcription. En effet, il fait partie du complexe TFIID en association avec d'autres protéines, les TAF (*TBP associated factors*) (figure 6).

Il existe 14 TAF principaux numérotés de 1 à 15 (TAF14 n'existant pas) en fonction notamment de leurs poids moléculaires de environ 210 kDa pour TAF1 à 15kDa pour TAF13 chez la souris (TAF15 étant plus gros, mais ayant été identifié plus tardivement à cause de son



Bushnell et al., 2004

Figure 7 : Modèle d'assemblage du complexe de pré-initiation de la transcription avec l'ADN. Les structures de TFIIE (E), TFIIIF (F) et TFIIH (H) ont été obtenues en microscopie électronique et les structures de TFIIB (B_N et B_C) et de l'ARN pol II (pol) ont été obtenues par cristallographie. 4/7 : Rpb 4/7. A gauche, les structures individuelles et les relations avec l'ARN pol II sont présentées. A droite, un modèle d'association des différents complexes précédents est présenté.

interaction avec seulement une sous-population de complexes TFIID). Les structures tridimensionnelles de TFIID humain et de levure ont été résolues en microscopie électronique et certaines sous-unités ont été immunolocalisées (Grob et al., 2006) (Leurent et al., 2002). Les TAF étant très conservés entre les espèces, de la levure à l'homme, les complexes humain et de levure présentent une structure très comparable caractérisée par la présence de 3 lobes, A, B et C. La protéine TBP se situe dans le lobe C, au cœur du complexe, et fait face à la cavité centrale du complexe, cavité reconnue comme étant le site majeur de liaison de TFIID à l'ADN. Outre leur rôle de liaison aux coactivateurs, les TAF ont également un rôle dans la reconnaissance et la liaison au promoteur minimal, TAF1 et TAF2 étant par exemple capables de lier l'Initiateur (Chalkley and Verrijzer, 1999). Certains TAF, font, de plus, partie de complexes régulateurs de la transcription ne contenant pas TBP, comme par exemple les complexes SAGA, un complexe à activité histone acétyl-transférase.

TFIID est le premier GTF à reconnaître le promoteur minimal lors de l'initiation de la transcription grâce à la liaison de TBP à la boîte TATA et grâce aux TAF. La liaison de TFIID au promoteur permet alors le recrutement des autres GTF pour former un PIC fonctionnel. On a longtemps pensé que l'assemblage du PIC au promoteur se faisait de manière séquentielle, chaque GTF arrivant un à un, comme recruté par le précédent. Dans ce modèle, le complexe TFIID-promoteur recruterait les facteurs TFIIA et TFIIB pour permettre l'ancrage du PIC et l'orientation de la transcription vers le site d'initiation (pour une revue, Deng and Roberts, 2007). TFIIA et TFIIB interagissent à la fois avec TBP et avec des séquences d'ADN (Geiger et al., 1996) (Tan et al., 1996). TFIIB lierait alors TFIIF et l'ARN pol II. Enfin, TFIIE viendrait se lier directement à l'ARN pol II et recruter le dernier élément du PIC, TFIIH. Ce dernier complexe possède des activités hélicases et kinases et permet alors l'ouverture de l'ADN au niveau du TSS ainsi que la phosphorylation du domaine CTD de l'ARN pol II (pour une revue, Orphanides et al., 1996).

Actuellement, ce modèle séquentiel de recrutement des GTF est très largement remis en cause *in vivo*. Il est suggéré que certains facteurs de base seraient liés à l'ARN pol II avant que celle-ci soit recrutée au promoteur. Ainsi, l'assemblage du PIC aurait lieu en deux étapes au moins, la première étant toujours le recrutement de TFIID (ou tout autre complexe contenant des TAF) au niveau du promoteur et la deuxième étant la fixation d'un sous-groupe de facteurs et, éventuellement, de co-régulateurs de transcription contenant l'ARN Pol II (Ranish et al., 1999). La transcription peut alors commencer (figure 7).

Chez la levure, la mutation de la protéine TBP affecte non seulement la transcription par l'ARN pol II mais aussi les transcriptions par les ARN pol I et ARN pol III (Cormack and

Struhl, 1992). Chez les métazoaires, la protéine TBP fait également partie des complexes d'initiation de la transcription par les trois ARN polymérases (Comai et al., 1992) (White et al., 1992).

II. Modulation autour de la transcription de base...

Les éléments de base de la transcription, ARN pol II et GTF, sont suffisants pour assurer un taux basal, faible, d'expression des gènes de classe II. Cependant, lors du développement, ainsi que dans un organisme adulte, le programme d'expression génique évolue de manière temporelle d'une part, au cours du temps, du cycle cellulaire par exemple, et de manière spatiale d'autre part, d'une cellule à une autre. L'ensemble des protéines exprimées ainsi que leur taux d'expression sont deux paramètres finement régulés. Toutes les étapes de la transcription, ainsi que de la traduction, sont soumises à régulation.

Nous verrons dans cette partie que l'étape d'initiation de la transcription par les facteurs de base peut être soumise à modulation selon le type de gène exprimé, la cellule dans laquelle on se place et selon les besoins de la cellule. Cette modulation met en jeu des promoteurs de base non canoniques, des GTF non canoniques, et d'autres protéines régulatrices interagissant avec les facteurs de base.

1. ...grâce à des séquences promotrices diverses

a. Variations autour du promoteur minimal de base

On a longtemps pensé que la boîte TATA était un élément commun à tous les promoteurs. L'augmentation du nombre de promoteurs analysés, grâce entre autre au développement de la technique de CAGE (*cap analysis of gene expression*) qui permet de séquencer la partie 5' des ADNc d'une cellule, de définir précisément la position des TSS des gènes dans différents types cellulaires et donc d'aligner les régions en amont de ce TSS, a permis de mettre en évidence que le promoteur minimal tel que nous l'avons présenté précédemment n'est présent que dans un nombre très limité de gènes. En effet, chez l'homme, la boîte TATA n'est présente que dans 10% des gènes par exemple, ce pourcentage diminue à moins de 3% si on ne considère que les boîtes TATA localisées à environ 30 nucléotides en amont du TSS. Il en est de même pour toutes les séquences du promoteur minimal citées ci-

dessus, il n'existe aucune séquence consensus universelle contenue dans tous les promoteurs minimaux, même au sein d'un seul et même organisme (FitzGerald et al., 2004). Il a été suggéré qu'en l'absence de boîte TATA, l'élément initiateur, présent à un plus grand nombre de promoteurs, pouvait recruter TFIID grâce à son interaction avec les TAF. On peut alors supposer que la présence ou l'absence d'un ou de plusieurs motifs peut dicter les différentes propriétés d'un promoteur minimal donné.

Des études systématiques ont permis de classer les promoteurs selon les séquences qu'ils contiennent et de mettre en évidence des groupes de gènes qui possèdent l'un ou l'autre des types de promoteurs (Bajic et al., 2006) (FitzGerald et al., 2004). Tout d'abord, les études de CAGE ont permis de mettre en évidence différents types de promoteurs en relation avec la présence ou non d'un TSS bien défini. Deux classes extrêmes de promoteurs peuvent être déterminées, les promoteurs possédant un seul TSS bien défini, les promoteurs à simple pic, et les promoteurs possédant plusieurs TSS distribués dans une fenêtre de 100 pb, les promoteurs à pic étendu. La première catégorie de promoteurs a été démontrée comme contenant plus souvent une boîte TATA ainsi que d'autres séquences consensus de liaison à des facteurs de transcription. Ils régulent le plus souvent des gènes dont l'expression est spécifique d'un tissu. Les promoteurs à pic étendu sont en revanche, en général, dépourvus de boîte TATA et sont associés à des gènes ubiquitaires comme des gènes de ménage par exemple. La mise en évidence d'un lien entre le type de gène régulé et une certaine catégorie de promoteurs renforce l'idée que selon le promoteur minimal qu'ils possèdent, deux gènes différents n'auront pas le même profil spatio-temporel d'expression.

En plus de posséder plusieurs TSS au sein d'une même région promotrice, certains gènes possèdent plusieurs promoteurs minimaux. Une étude chez les vertébrés suggère que 20 à 52% des gènes possèderaient des promoteurs alternatifs clairement séparés par plusieurs centaines de paires de bases (Kim et al., 2005) (Kimura et al., 2006). Le gène codant pour la protéine eIF-1a, impliquée dans l'initiation de la traduction possède deux promoteurs, le premier contenant une boîte TATA et le second, en 5' du premier, ne contenant pas de boîte TATA. Il a été montré que, dans l'ovocyte, 70% de la transcription de ce gène se fait sous la dépendance du promoteur contenant la boîte TATA, alors qu'après la fécondation, dans l'embryon, au moins 75% de la transcription de ce gène se fait sous la dépendance du promoteur en 5', ne contenant pas de boîte TATA (Davis and Schultz, 2000). L'utilisation de promoteurs minimaux alternatifs peut donc être un mécanisme à l'origine de l'expression

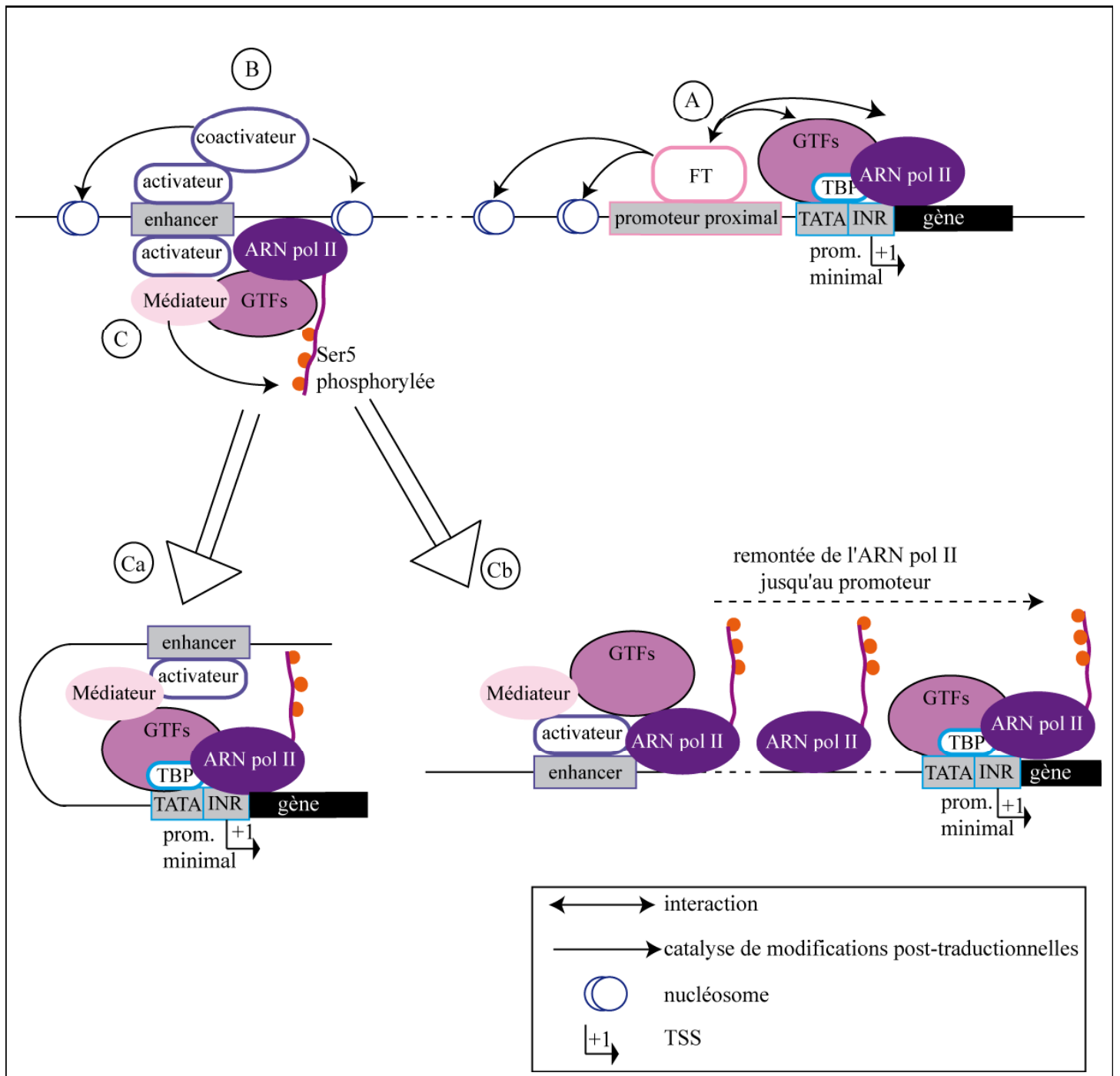


Figure 8 : Schématisation de la reconnaissance des promoteurs proximal et distal (« enhancer ») par des facteurs spécifiques. Le promoteur proximal d'un gène peut être reconnu par des facteurs de transcription (FT). Le facteur de transcription peut alors agir en activant ou inhibant le recrutement de l'ARN pol II et des GTF, ou en modifiant la chromatine par exemple (A). Un enhancer peut être reconnu par un activateur. L'activateur recrute alors un coactivateur qui peut par exemple agir sur les modifications de la chromatine (B). L'activateur peut recruter le médiateur qui phosphoryle la CTD de l'ARN pol II sur la sérine 5 (C). L'ARN pol II modifiée peut alors lier le promoteur minimal grâce à la formation d'une boucle dans l'ADN (Ca : phénomène de « looping ») ou en se déplaçant sur l'ADN (Cb : phénomène de « tracking »).

spécifique d'un gène dans un tissu ou à un stade de développement. Le choix entre un promoteur plutôt qu'un autre se fait grâce à l'interaction du promoteur minimal avec des séquences régulatrices en amont ou en aval du promoteur, via des protéines régulatrices spécifiques.

b. Les séquences régulatrices

Bien que tous les gènes ne possèdent pas le même promoteur minimal, voire en possèdent plusieurs, la diversité des promoteurs minimaux ne suffit pas à expliquer les différences d'expression géniques entre deux types cellulaires par exemple. En complément du promoteur minimal, les gènes possèdent des séquences régulatrices spécifiques qui autorisent le contrôle de leur expression grâce à leur liaison par des facteurs régulateurs. Immédiatement en amont du promoteur minimal, entre -50 et -200 pb par rapport au TSS, des séquences de reconnaissance spécifiques pour des facteurs de transcription peuvent être présentes, elles définissent le promoteur proximal. La liaison des facteurs de transcription au promoteur proximal, lorsque ceux-ci sont présents, permet alors une régulation de l'expression du gène en aval (figure 8A).

La lignée germinale constitue un excellent exemple de type cellulaire spécialisé. Ainsi, lors de la folliculogénèse précoce, de nombreux gènes spécifiques des ovocytes vont devoir être transcrits. *Oct4* et *Nobox*, deux facteurs de transcription, ainsi que *Gdf9*, un facteur de croissance, sont trois gènes dont l'expression est finement régulée au cours de la folliculogénèse. La délétion de la protéine NOBOX induit la formation d'ovocytes anormaux, dans lesquels *Gdf9* et *Oct4* sont sous-exprimés par rapport à des ovocytes normaux (Rajkovic et al., 2004). NOBOX est capable de se lier en amont des TSS de *Oct4* et *Gdf9* à respectivement 1 et 3 séquences, nommées éléments de liaison à NOBOX (NBE : NOBOX *binding elements*), et activer ainsi l'expression de ces gènes (Choi and Rajkovic, 2006). Les NBE se situent entre -426 et -1259 pb par rapport au TSS, au sein d'un promoteur plus éloigné, le promoteur distal. Au sein du promoteur distal, de nombreuses séquences régulatrices peuvent être présentes. Ces séquences sont des sites de liaison à des facteurs régulateurs. Les facteurs, donc les séquences, peuvent être activateurs ou inhibiteurs de la transcription. L'expression d'un facteur régulateur dans un type cellulaire donné entraînera l'expression adéquate des autres gènes de ce même type cellulaire. Parmi les séquences régulatrices du promoteur distal, les « *enhancers* » ont été bien caractérisés. Les « *enhancers* » sont des séquences régulatrices agissant en *cis* pour augmenter la transcription,

indépendamment de leur orientation et de leur distance au TSS (Banerji et al., 1981). Ces séquences peuvent être situées très loin du TSS, jusqu'à plusieurs dizaines de milliers, voire un million, de paires de bases en amont ou en aval du gène régulé, dans les introns, voire en aval du site de terminaison de la transcription (pour une revue, Blackwood and Kadonaga, 1998). L'expression d'un seul gène peut être régulée par plusieurs « *enhancers* » et différents « *enhancers* » peuvent être actifs dans différents types cellulaires, à différents moments, ou sous différents stimuli (Atchison, 1988). De plus, un même « *enhancer* » peut réguler l'expression de différents gènes. Ces éléments sont liés par des activateurs transcriptionnels qui pourront à leur tour recruter des coactivateurs permettant alors l'activation de la transcription en interagissant directement avec les facteurs de transcription de base (Mack et al., 1993) ou en augmentant l'accessibilité de la chromatine (voir partie chromatine) en induisant des modifications post-traductionnelles sur les histones par exemple (Demers et al., 2007) (figure 8B et C).

2. ... grâce à des complexes TFIID alternatifs

a. Variation autour des TAF

Le complexe TFIID composé de TBP et de l'ensemble des TAF a longtemps été considéré comme la seule forme présente pour toute transcription par l'ARN pol II. Or, il a été montré que les TAF peuvent être exprimés à des niveaux très différents dans des types cellulaires distincts. Grâce au développement d'anticorps contre TBP et les différents TAF, l'existence de complexes comparables au complexe TFIID mais ne possédant pas tous les TAF a pu être mise en évidence, suggérant que tous les TAF ne sont pas nécessaires à la transcription par l'ARN pol II dans tous les types cellulaires. La protéine TAF10, par exemple, semble être nécessaire à la transcription par l'ARN pol II dans l'embryon précoce de souris, alors que plus tard dans le développement, de nombreux gènes sont transcrits normalement en l'absence de TAF10 (Indra et al., 2005). Plus probable qu'un complexe TFIID composé de TBP et des 14 TAF, on pense actuellement qu'il existe un sous-complexe TFIID minimal stable, contenant seulement certains TAF et ne contenant pas TBP, qui peut être complété par d'autres TAF et par TBP selon le promoteur du gène transcrit et le type cellulaire (Wright et al., 2006) (Kaufmann et al., 1998) (Martinez et al., 1998).

En plus de pouvoir être ou non présents au sein de TFIID, certains TAF possèdent des orthologues exprimés spécifiquement dans certains types cellulaires. TAF4 possède, par

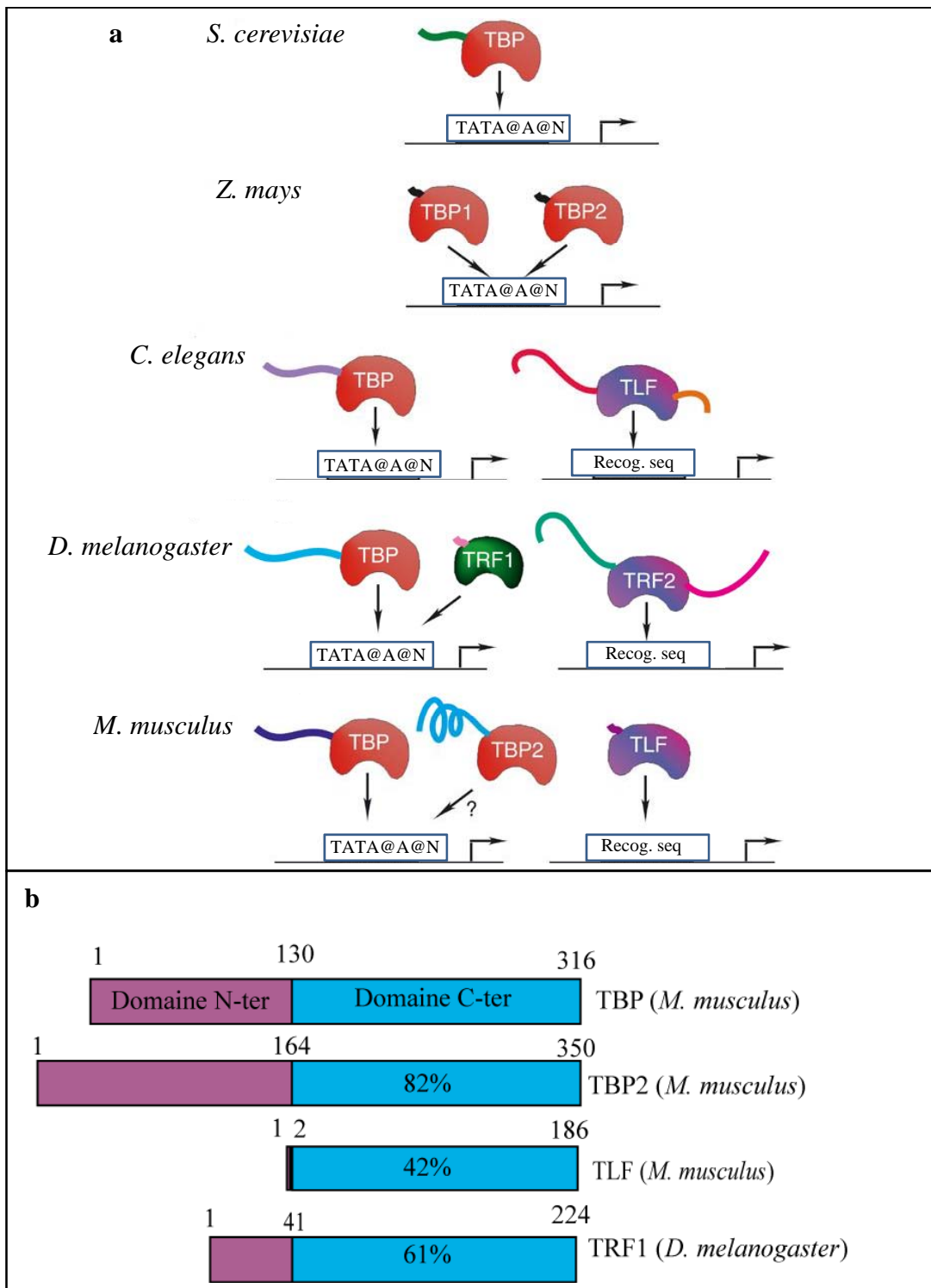


Figure 9 : Répartition des différents facteurs de la famille de TBP selon les espèces et comparaison de leurs domaines C-terminaux. a : Les domaines C-terminaux conservés des homologues de TBP sont représentés par une structure en selle de cheval, les domaines N-terminaux non conservés sont représentés par des traits. La cible de ces facteurs, une boîte TATA (TATA@A@N) ou un promoteur dépourvu de boîte TATA (Recog. Seq), est indiquée. La figure est adaptée d'après Müller et al., 2004. b : Les pourcentages d'identité entre les domaines C-terminaux des différents homologues de TBP et TBP sont présentés pour la souris. Le domaine C-terminal de TRF1 de drosophile est comparé à TBP de souris. Les nombres correspondent à la position des acides aminés.

exemple, un homologue TAF4b exprimé dans les ovocytes et les spermatozytes. Les souris femelles ne possédant pas TAF4b sont stériles (Falender et al., 2005) (Freiman et al., 2001) et il a été suggéré que TAF4b était à l'origine de la spécificité de la transcription de certains gènes dans les ovocytes (Geles et al., 2006).

Tous les TAF ne sont donc pas universels, mais ils peuvent agir comme des composants spécifiques du PIC relatifs à différents types cellulaires dans un organisme (figure 11).

b. La transcription indépendante de TBP

Bien qu'il existe un lien clair entre le degré d'occupation d'un promoteur par TBP et le taux d'expression d'un gène chez la levure (Kuras and Struhl, 1999), cette corrélation n'est pas toujours vérifiée chez les métazoaires. La surexpression de TBP augmente la transcription des gènes possédant une boîte TATA mais elle semble réprimer l'expression de ceux qui n'en possèdent pas (Colgan and Manley, 1992). Dans des embryons de xénope (Veenstra et al., 2000), de poisson zèbre (Muller et al., 2001), et de souris (Martianov et al., 2002b) ne contenant pas la protéine TBP, la transcription par l'ARN pol II ne cesse pas totalement et certains gènes sont alors transcrits aussi efficacement que dans des embryons normaux.

Toutes ces études montrent que TBP n'est pas nécessaire à toute transcription par l'ARN pol II et laissent supposer qu'il existe d'autres facteurs pouvant remplacer la fonction de TBP aux promoteurs. En recherchant, par alignement de séquences, des protéines similaires à TBP dans différents organismes, plusieurs paralogues de TBP ont été découverts chez les métazoaires, les levures n'en possédant pas. Les membres de la famille TBP ont une forte similarité avec la partie C-terminale de TBP et possèdent des acides aminés identiques ou chimiquement similaires à des positions importantes suggérant une structure et des caractéristiques fonctionnelles comparables à celles de TBP.

C. elegans possède un seul homologue de TBP, TLF. Le xénope et la souris possèdent, en plus de TLF, la protéine TBP2. Enfin, *D. melanogaster* possède non seulement les deux homologues précédents, mais également un membre qui lui est propre, TRF1 (figure 9).

- TRF1 (*TBP related factor 1*)

TRF1 est le seul membre de la famille de TBP à n'avoir été découvert que chez un seul organisme, la drosophile. TRF1 est très similaire au domaine de cœur de TBP et elle est

capable de lier la boîte TATA et d'interagir avec TFIIA et TFIIB pour former un PIC actif et ainsi se substituer à TBP *in vitro*.

In vivo, TRF1 est ubiquitaire pendant l'embryogenèse précoce de la drosophile puis son expression est restreinte aux cellules du système nerveux et aux gonades pendant le reste du développement (Crowley et al., 1993) (Hansen et al., 1997). Des sites de liaison de TRF1 à l'ADN ont été identifiés dans le génome de la drosophile. Il apparaît que 78% de ces sites de liaison colocalisent avec des sites de liaison du complexe d'initiation de la transcription par l'ARN pol III (Takada et al., 2000). TRF1 pourrait donc réguler la transcription par les ARN pol II et ARN pol III et ce, spécifiquement dans certains tissus.

- TLF (*TBP-like factor*, également nommée TRF2 TLP, TBPL1)

TLF est à 42% identique au domaine C-terminal de TBP. L'étude de sa séquence en comparaison avec celle de TBP a permis de suggérer que TLF était capable de lier TFIIA et TFIIB mais qu'il ne devait pas présenter de spécificité particulière pour lier la boîte TATA (Dantonel et al., 1999). La liaison à TFIIA a ensuite été montrée dans des cellules HeLa et des expériences de ChIP-on-chip n'ont pas permis de mettre en évidence de séquences consensus de liaison de TLF à l'ADN comparables à la boîte TATA. De plus, TLF ne peut pas se substituer TBP pour compléter TFIID dans des expériences de transcription *in vitro* (Teichmann et al., 1999) (Moore et al., 1999).

TLF est présent chez tous les métazoaires et son expression est nécessaire au développement embryonnaire de nombreuses espèces, *C. elegans* (Dantonel et al., 2000), le xénope (Veenstra et al., 2000), le poisson zèbre (Muller et al., 2001) et la drosophile (Kopytova et al., 2006) grâce à son implication spécifique dans l'expression de nombreux gènes. En effet, chez le xénope par exemple, 39% des gènes embryonnaires sont sous-exprimés en l'absence de TLF alors que seulement 2% le sont en l'absence de TBP. Chez la souris en revanche, les embryons ne possédant pas TLF sont viables mais les mâles sont stériles. En effet, TLF est très abondant dans les testicules et la régulation de son expression et de sa localisation est très différente de celle de TBP au cours de la spermatogenèse (Martianov et al., 2002a) (Martianov et al., 2001).

D'après ces observations, il a été suggéré que TLF pourrait, grâce à d'autres facteurs spécifiques, orienter la transcription lors de la spermatogenèse murine (Kimmins et al., 2004).

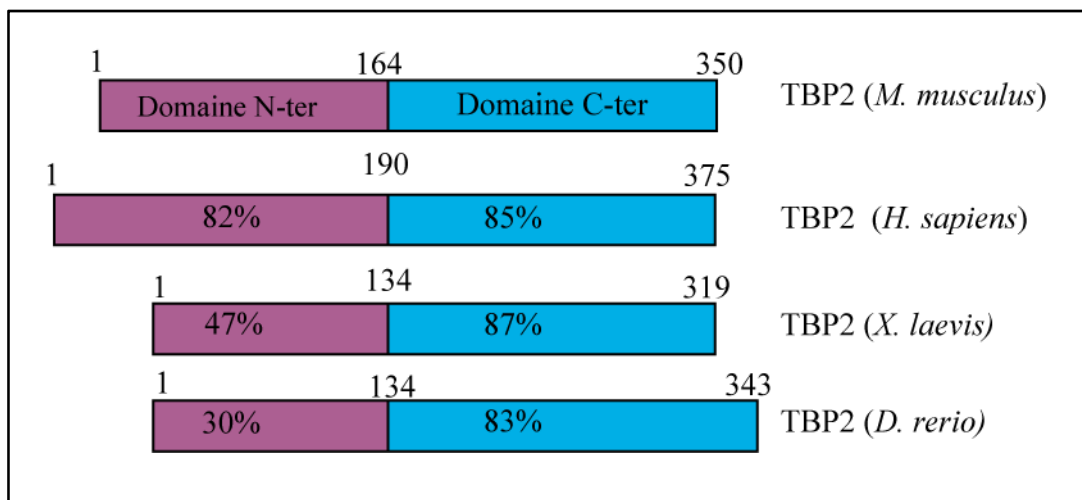


Figure 10 : Représentation schématique et comparaison des domaines C-terminaux et domaines N-terminaux des protéines TBP2 de différentes espèces. Le domaine C-terminal (C-ter) contient le domaine de liaison à l'ADN. Les nombres indiquent la position des acides aminés. Les domaines des protéines TBP2 de chaque espèce sont comparés au même domaine chez la souris, le résultat est donné en pourcentage d'identité.

- TBP2 (*TATA binding protein 2*, également nommée TRF3, TBPL2)

TBP2 est le dernier membre de la famille TBP à avoir été cloné après recherche de séquences similaires à celle de TBP. Chez l'homme, TBP2 possède les mêmes domaines que TBP : un domaine C-terminal identique à 93% à celui de TBP et un domaine N-terminal à 15% identique à TBP. TBP2 possède tous les résidus nécessaires aux interactions avec la boîte TATA, TFIIA et TFIIB (Persengiev et al., 2003). Ces résultats ont été confirmés peu après chez le poisson zèbre, où l'ADNc de la protéine TBP2 a été cloné, et son interaction avec la boîte TATA, TFIIA et TFIIB a été montrée *in vitro*. De plus, TBP2 est capable d'initier une réaction de transcription *in vitro*, aussi efficacement que TBP (Bartfai et al., 2004). TBP2 est présent exclusivement chez les vertébrés où son domaine C-terminal est conservé à plus de 80% (figure 10). Ces observations ont permis de faire l'hypothèse que le gène TBP2 serait apparu après un événement de duplication génique de TBP chez un ancêtre commun aux vertébrés (Bartfai et al., 2004) (Persengiev et al., 2003).

Chez le poisson zèbre et le xénope, TBP2 est exprimé pendant le développement embryonnaire, ainsi que très fortement dans les gonades adultes. TBP2 semble avoir un rôle dans la différenciation dorso-ventrale des embryons de xénope et de poisson zèbre. Il est exprimé plus fortement du côté ventral, et l'expression de nombreux gènes ventraux dépend de la présence de TBP2 (Bartfai et al., 2004) (Jacobi et al., 2007). L'inactivation de TBP2 chez le poisson zèbre entraîne des défauts lors de la gastrulation, de la formation du mésoderme, et de l'hématopoïèse. Les problèmes d'hématopoïèse seraient dus exclusivement à une sous-expression d'une protéine impliquée dans la spécification du mésoderme, *Mespa*, et cible directe de TBP2. TBP2 est par ailleurs très abondante dans les ovocytes où TBP est très faiblement exprimé, il est recruté aux promoteurs (Jallow et al., 2004) et est capable d'y initier la transcription *in vivo* (Akhtar and Veenstra, 2009). Ces résultats caractérisent TBP2 comme un facteur de remplacement de TBP dans les ovocytes de poisson zèbre et de xénope, et suggèrent encore une fois que ce facteur peut être à l'origine d'une transcription spécifique dans des types cellulaires particuliers.

Chez la souris, le profil d'expression et le rôle de TBP2 sont soumis à controverse. Les premières études rapportent une expression ubiquitaire de la protéine dans de nombreux tissus et lignées cellulaires murins (Persengiev et al., 2003). En revanche, selon une autre étude, l'ARNm de TBP2 ne serait détectable que dans les ovaires et plus précisément dans les ovocytes, remettant en cause la présence de la protéine dans les autres tissus (Xiao et al., 2006). Cependant, en étudiant des cellules C2C12, lignée cellulaire de myoblastes murins, une autre équipe a montré que pour permettre la différenciation des myoblastes en myotubes,

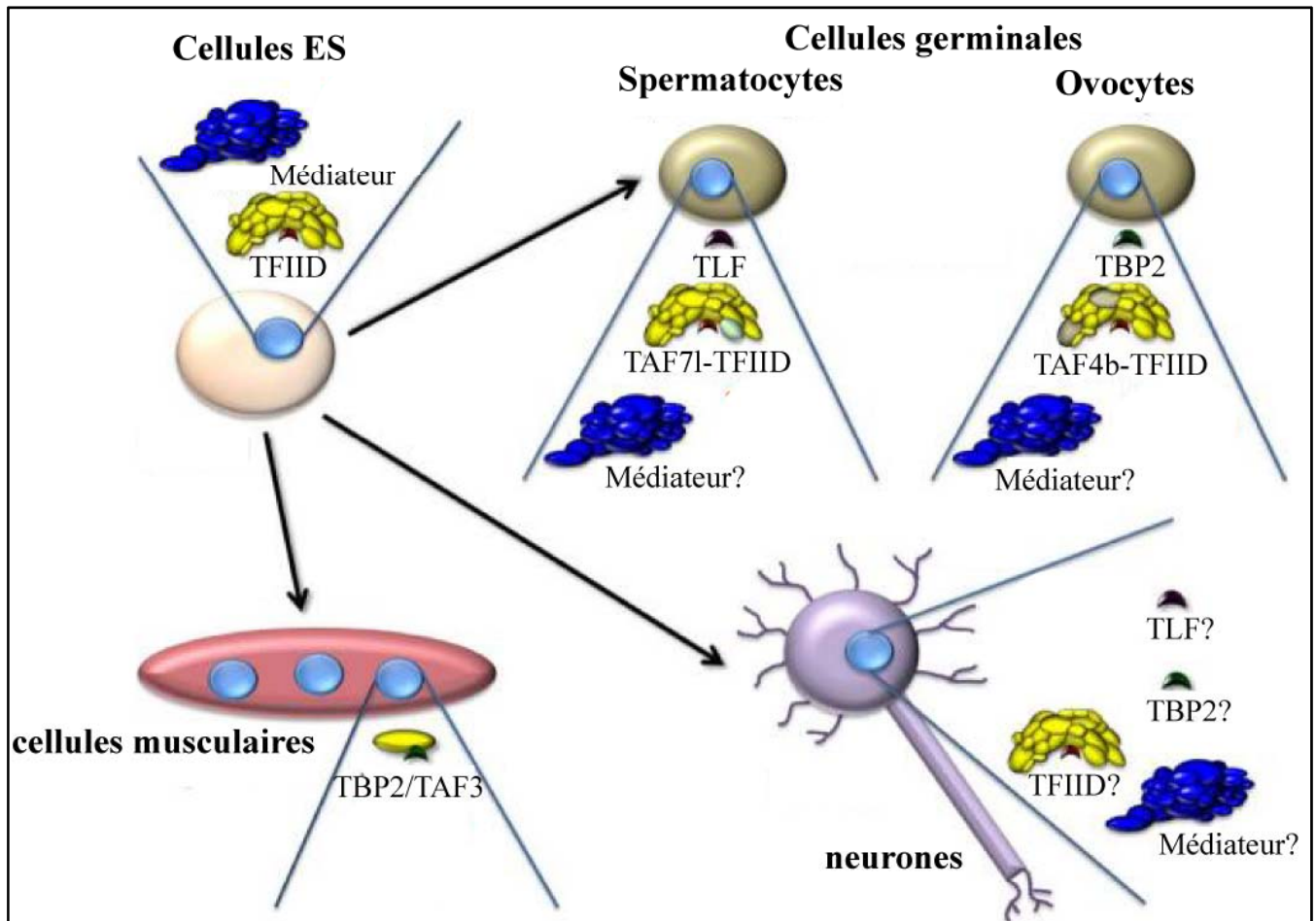


Figure 11 : Exemples des variations entre les facteurs de transcription de base utilisés dans différents types cellulaires. Les cellules ES contiennent un complexe TFIID et un complexe méiateur canoniques. Les formes de remplacement utilisées spécifiquement dans les différents types cellulaires sont listées. “Méiateur ?” signifie que la composition du méiateur n’est pas connue dans le type cellulaire considéré. La figure est adaptée d’après D’Alessio et al., 2009

TFIID est remplacé par un complexe contenant TAF3 et TBP2 qui se lierait au promoteur de la myogénine, facteur indispensable à la différenciation, pour induire son expression (Deato and Tjian, 2007) (Deato et al., 2008).

Toutes ces données suggèrent que TBP2 pourrait être à l'origine de la formation d'un complexe comparable à TFIID, mais permettant la transcription d'un ensemble de gènes spécifiques d'un type cellulaire donné. Bien que des expériences de chromatographie par filtration montrent que TBP2 fait partie d'un complexe d'environ 200kDa chez l'homme (TFIID faisant environ 1MDa) (Persengiev et al., 2003), la nature, la fonction et les mécanismes d'action d'un tel complexe restent à élucider.

Les rôles des différents homologues de TBP illustrent que des facteurs généraux de la transcription peuvent jouer le rôle de régulateurs de l'expression génique dans différents types cellulaires et, en particulier ici, pendant le développement et dans les gonades (pour une revue, Torres-Padilla and Tora, 2007) (figure 11).

La régulation de l'expression génique peut donc se faire au niveau de l'initiation de la transcription grâce à des variations autour de la machinerie de base de la transcription et à des facteurs spécifiques recrutant cette machinerie de base. Ces mécanismes, s'ils sont suffisants pour activer la transcription *in vitro*, sur un ADN nu, ne le sont pas *in vivo*, où l'ADN est fortement compacté au sein du noyau autour de protéines, les histones. Pour activer ou inhiber la transcription, la cellule, et en particulier les facteurs et les cofacteurs spécifiques dont nous avons parlé précédemment, vont également devoir et pouvoir moduler l'état de compaction de la chromatine, permettant ou non à la machinerie de base de la transcription d'accéder au gène à transcrire.

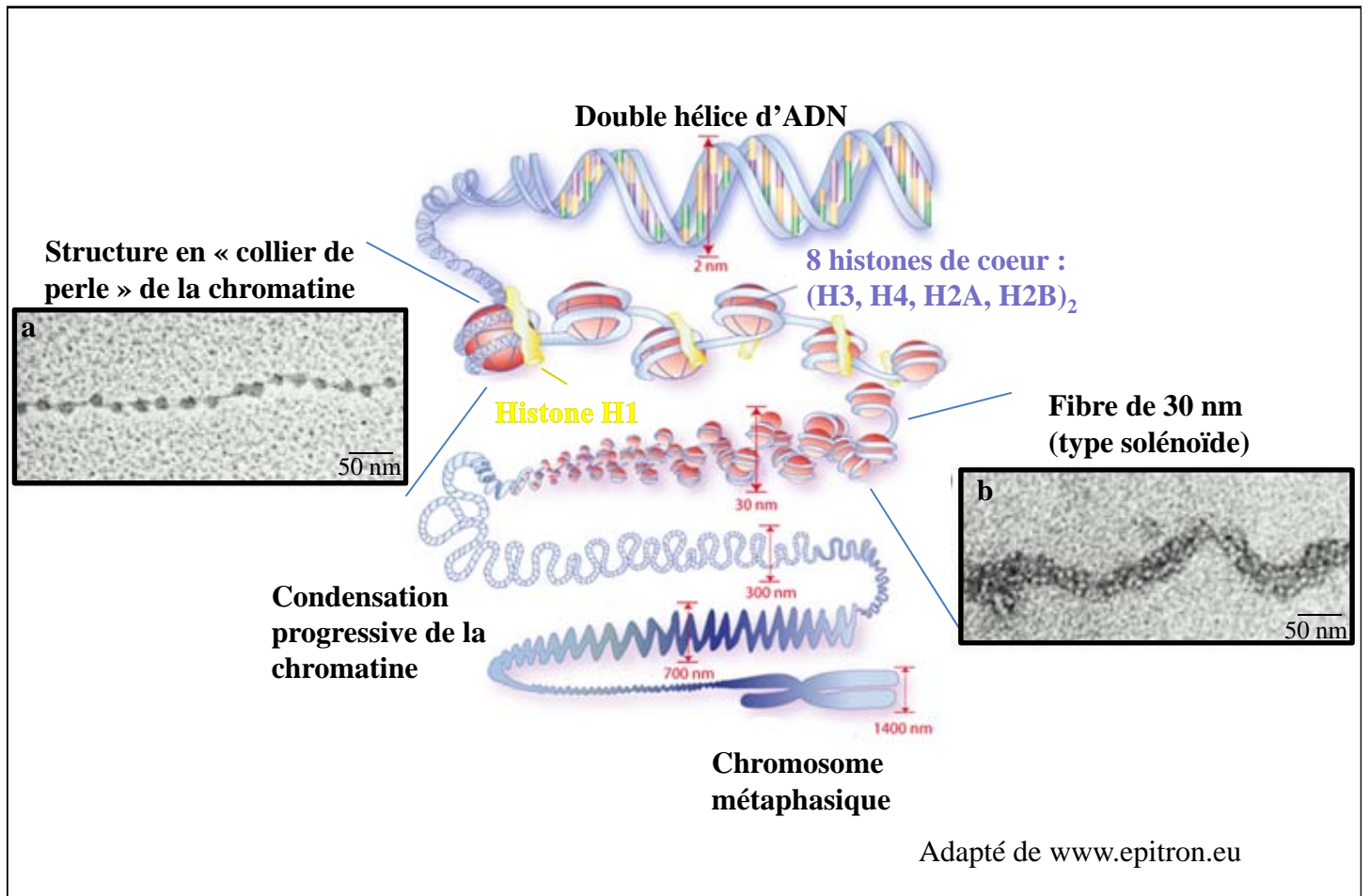


Figure 12 : Représentation schématique des différents niveaux de compaction de l'ADN et photographies du « collier de perle » et de la fibre de 30 nm en microscopie électronique. Le schéma est réalisé d'après les données de (Hansen, 2002). Les différents niveaux de compaction sont représentés à partir de l'ADN nu, en haut, à la chromatine la plus condensée sous sa forme mitotique, en bas. Le diamètre de la fibre chromatinienne est indiqué pour chaque étape. **a** : Photographie en microscopie électronique d'un segment de chromatine expérimentalement déroulé après isolement pour en montrer les nucléosomes. **b** : Photographie en microscopie électronique d'un segment de chromatine isolé directement d'un noyau en interphase. Les photographies a et b sont issues de Albert et al., 2004.

Adapté de www.epitron.eu

B. La chromatine, un environnement modulable

Dans le noyau des cellules eucaryotes, le matériel génétique est organisé en une structure complexe de nature nucléo-protéique, la chromatine. Elle est composée de l'ADN et de petites protéines basiques, les histones. L'enroulement du matériel génétique autour des histones permet un repliement très hiérarchisé des 1,8 mètres d'ADN qui composent notre génome, nécessaire à son empaquetage dans le compartiment restreint de 6µm qu'est le noyau. La chromatine interfère avec l'accessibilité des séquences régulatrices décrites précédemment, ainsi, la modulation de sa structure constitue un niveau de régulation de la transcription.

I. Structure de la chromatine

1. Les différents niveaux de compaction de la chromatine

a. La fibre de 10 nm

Les premières observations de la chromatine en microscopie électronique ont révélé la présence d'une structure régulière, en forme de collier de perles, d'un diamètre d'environ 10 nm, communément appelée fibre de 10 nm (Olins and Olins, 1974) (Oudet et al., 1975) (figure 12a). Ce premier niveau d'organisation correspond à l'enroulement d'ADN autour d'un octamère d'histones formant ainsi un nucléosome, unité fondamentale de la chromatine, la perle du collier (Kornberg, 1974). La formation de la fibre de 10 nm a deux rôles principaux. Elle permet d'une part d'apporter un premier degré de compaction du génome au sein du noyau, et d'autre part, elle neutralise les charges négatives de l'ADN grâce à son interaction avec des histones chargées positivement.

La fibre de 10 nm constitue le substrat sur lequel la transcription peut avoir lieu. La chromatine peut également adopter des niveaux de condensation supérieurs, rendant ainsi l'ADN inaccessible à la machinerie de transcription.

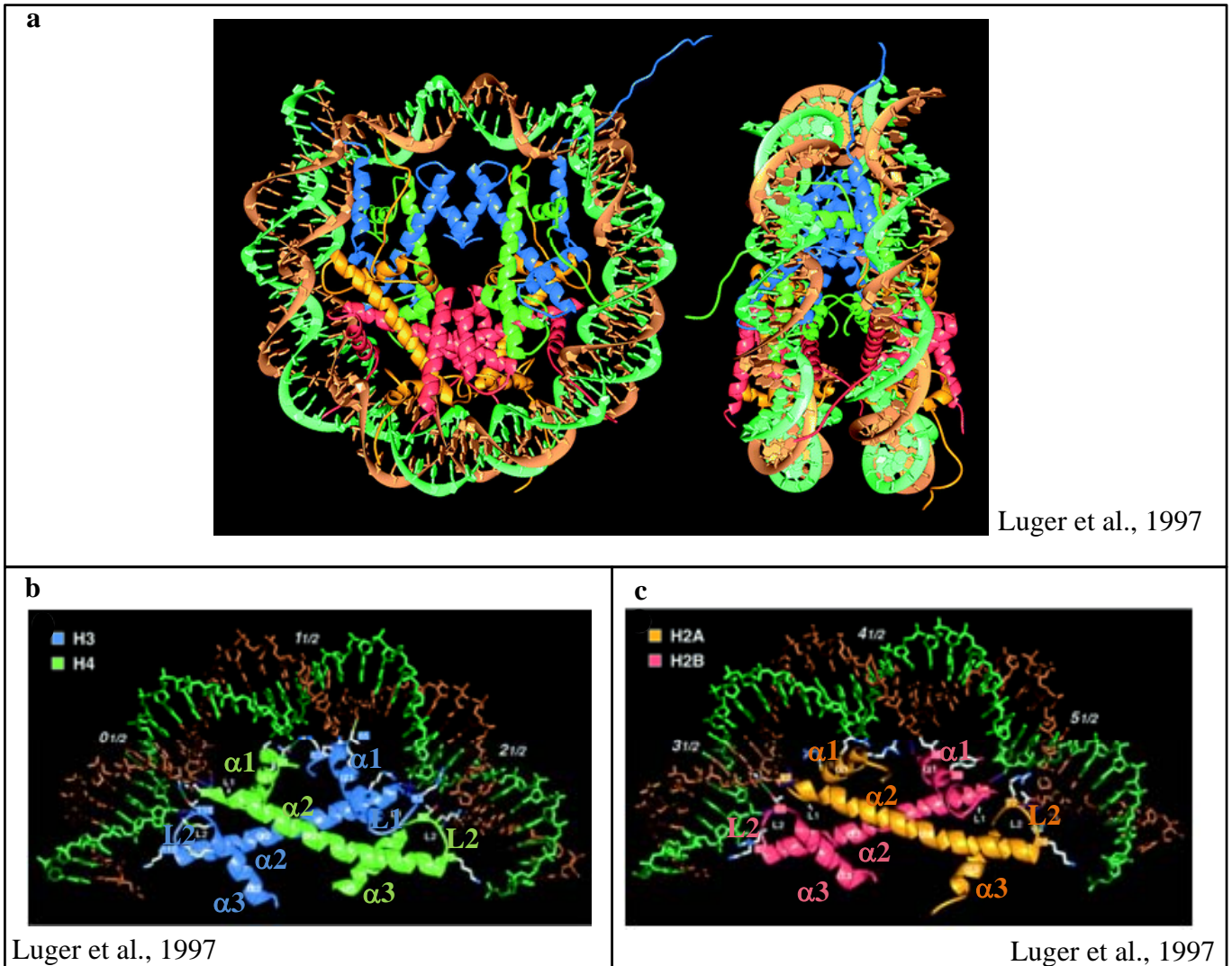


Figure 13 : Structures de la particule de cœur du nucléosome et des dimères H3/H4 et H2A/H2B liés à l'ADN résolues en cristallographie aux rayons X. **a** : Représentation en ruban du squelette phosphodiester des 146 pb d'ADN composant le cœur du nucléosome (marron et turquoise), et des chaînes principales des huit histones de cœur (en bleu pour H3, vert pour H4, orange pour H2A et rose pour H2B). Les vues sont par dessous l'axe de la superhélice d'ADN pour la représentation de gauche, et perpendiculaire à cet axe pour la représentation de droite. **b** : Domaines « *histone fold* » de H3 (en bleu) et H4 (en vert) dans leur position au sein du nucléosome. Les éléments structuraux $\alpha 1$ -L1- $\alpha 2$ -L2- $\alpha 3$ sont représentés. **c** : Domaines « *histone fold* » de H2A (en orange) et H2B (en rose) dans leur position au sein du nucléosome. Les éléments structuraux $\alpha 1$ -L1- $\alpha 2$ -L2- $\alpha 3$ sont représentés.

b. Les niveaux de compaction supérieurs

Des données de microscopie électronique et de diffraction aux rayons X supportent l'idée que la fibre de 10 nm pourrait s'enrouler sur elle-même pour se condenser en une fibre de 30 nm (figure 12b). Plusieurs modèles de repliement ont été proposés, les deux plus importants étant le repliement de type solénoïde où la fibre de 10 nm s'enroule autour d'un axe de symétrie central (Finch and Klug, 1976), et le modèle d'arrangement en zig-zag où les nucléosomes forment une hélice à partir de deux sites d'initiation (Horowitz et al., 1994) (Woodcock et al., 1984). L'existence de la fibre de 30 nm est actuellement remise en cause et il est suggéré que la fibre de 10 nm serait le seul arrangement régulier de la chromatine, les autres niveaux de compaction se feraient alors par des courbures plus ou moins fréquentes et régulières de la fibre de 10 nm (pour une revue, Fussner et al., 2011).

La formation de la fibre de 10 nm permet de condenser six fois l'ADN nu, et celle de la fibre de 30 nm le permet soixante fois. La chromatine peut encore adopter des niveaux de compaction bien supérieurs, avec un repliement de la super hélice de 30 nm en une fibre de 100 à 300 nm de diamètre constituée de boucles de plusieurs centaines de milliers de paires de bases (pour une revue, Cook, 1995). La compaction extrême de ces boucles de 300 nm donnera ensuite naissance aux chromosomes mitotiques (Belmont, 2006) (Fraser and Bickmore, 2007) (figure12).

2. Histones et nucléosome

Le nucléosome est l'unité de base de la chromatine dans presque toutes les cellules eucaryotes, à l'exception, par exemple, des spermatozoïdes où l'ADN est majoritairement condensé autour de protamines. La structure du nucléosome a été résolue précisément par cristallographie aux rayons X (Davey et al., 2002) (Luger et al., 1997). Il est formé d'une particule de cœur contenant un octamère des quatre histones de cœur, H2A, H2B, H3 et H4, autour duquel sont enroulées 146 pb d'ADN en une superhélice gauche (figure 13a).

Les histones du nucléosomes, H2A, H2B, H3 et H4 sont des petites protéines basiques très conservées au cours de l'évolution. Elles possèdent un domaine central globulaire à repliement caractéristique de type « *histone fold* » où trois hélices α , α_1 , α_2 et α_3 , sont reliées entre elles par des boucles flexibles, L1 et L2, formant une sorte de Z (Luger et al., 1997) (figure 13b et 13c). Certaines interactions entre les différentes histones et les interactions entre les histones et l'ADN se font au niveau de ce domaine central. Les extrémités N- et C-

terminales des histones ne sont pas structurées, mais n'en sont pas moins les régions les plus conservées des histones entre les espèces. Elles se projettent en dehors du cylindre formé par les régions centrales des histones entourées par l'ADN, et jouent le rôle de domaine d'interaction à la fois des histones entre elles et des histones avec d'autres protéines. Ces extrémités non structurées, et en particulier l'extrémité N-terminale, sont le site de nombreuses modifications post-traductionnelles qui sont impliquées dans la régulation de la transcription comme nous le verrons par la suite.

En plus des quatre histones de cœur, il existe une cinquième histone possédant la même structure de base que les autres mais ne faisant pas partie du nucléosome au sens strict, l'histone H1. Elle est en contact avec l'ADN au niveau de son point d'entrée et de sortie du nucléosome et sert de clampe pour stabiliser le nucléosome et le rendre moins mobile. L'histone H1 n'est pas présente au niveau de chaque nucléosome, et son enrichissement est plutôt corrélé à une chromatine condensée (Thomas, 1984). Tout comme les histones de cœur, H1 est la cible de plusieurs modifications post-traductionnelles que nous n'aborderons pas ici (pour une revue, Woodcock et al., 2006).

3. Hétérochromatine, euchromatine et transcription

Le lien entre le degré de compaction de la chromatine et le degré d'expression des gènes avait déjà été supposé au début du XX^{ème} siècle suite à des études cytologiques. En marquant la chromatine et visualisant les noyaux sous microscope, Emil Heitz a noté que la chromatine des cellules eucaryotes peut être divisée en deux catégories selon son apparence microscopique : l'hétérochromatine et l'euchromatine (Heitz, 1928). L'hétérochromatine est alors définie comme des régions de chromatine condensée dans le noyau, fortement marquées par la coloration, et ne se décondensent pas au cours de l'interphase. L'euchromatine est au contraire la portion de chromatine qui se décondense après la sortie de la mitose. Emil Heitz a alors supposé que l'hétérochromatine correspondait aux régions inactives du génome alors que l'euchromatine représentait le génome activement transcrit (Passarge, 1979). Ces domaines sont aujourd'hui de mieux en mieux définis d'un point de vue moléculaire : les régions hétérochromatiques sont en effet plus condensées (Weintraub and Groudine, 1976), contiennent peu de gènes et répliquent tardivement pendant la phase S (Holmquist, 1987). A l'opposé, l'euchromatine est plus ouverte, sa structure correspond à celle de la fibre de 10 nm. Elle est riche en gènes, et se réplique au début de la phase S (Gilbert et al., 2004).

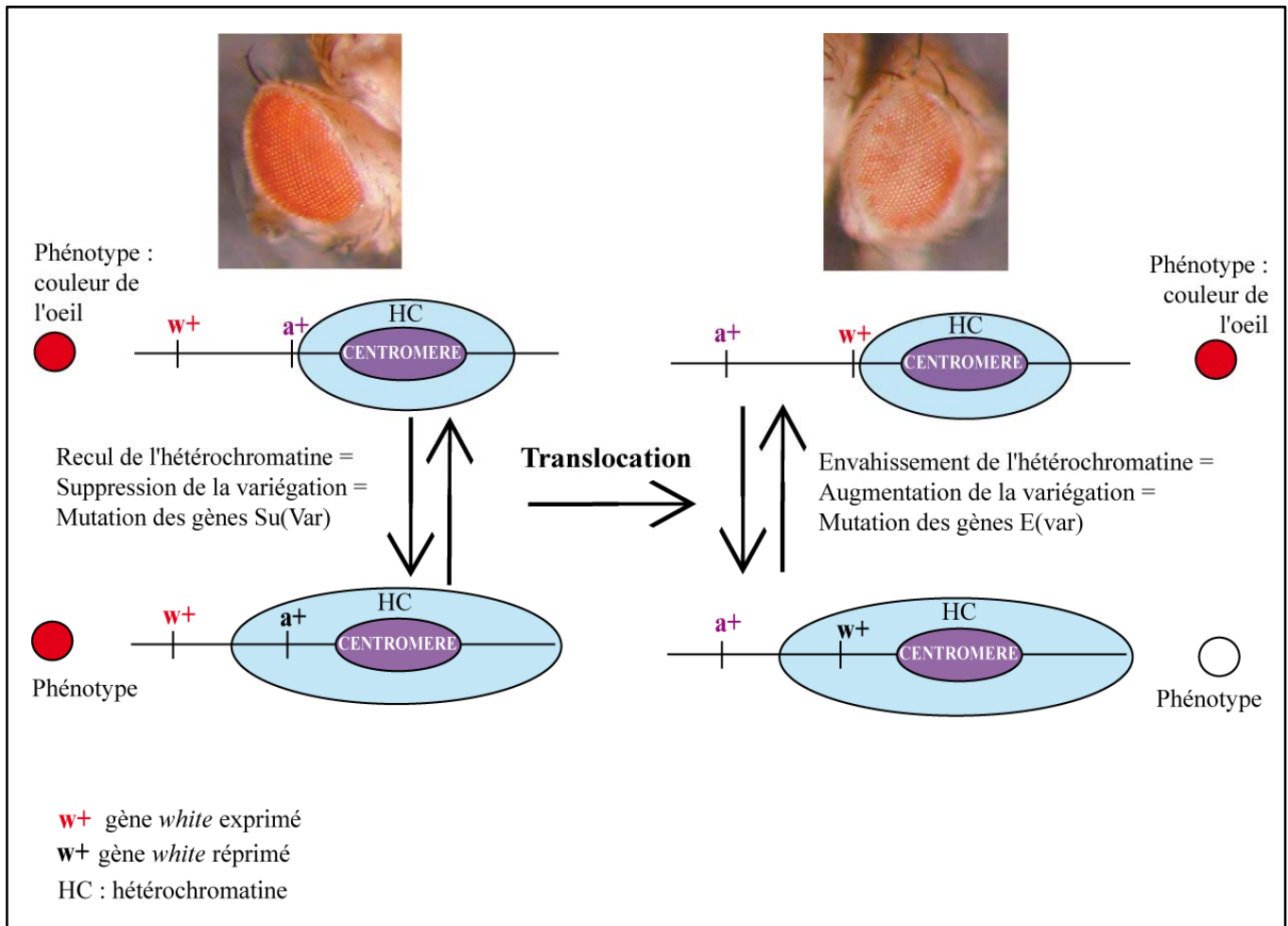


Figure 14 : Schématisation du phénomène de variégation par effet de position (PEV) chez la drosophile. Chez les drosophiles sauvages (en haut à gauche), le locus du gène *white* (w^+) est très éloigné des centromères, sa transcription n'est pas affectée par l'envahissement de l'hétérochromatine (HC) dû à une mutation des gènes *E(var)* (en bas à gauche). La drosophile a donc les yeux entièrement rouges. Après translocation (à droite), le gène *white* se trouve à proximité de l'hétérochromatine centromérique. La transcription du gène *white* peut être inhibée par un envahissement de l'hétérochromatine, donnant aux cellules concernées une couleur blanche (en bas à droite).

La première preuve de l'influence du degré de compaction de la chromatine sur l'activité transcriptionnelle des gènes provient de la découverte du phénomène d'extinction par effet de position, PEV (*Position Effect Variegation*), chez la drosophile (Tartof et al., 1984). Au cours du PEV, des gènes de l'euchromatine se retrouvent placés dans ou à proximité de l'hétérochromatine suite à des translocations chromosomiques et peuvent être réprimés par envahissement du domaine d'hétérochromatine. Des drosophiles mutantes présentent alors une expression en mosaïque du gène *white*, responsable normalement de la pigmentation rouge des yeux. Des études génétiques chez la drosophile ont permis de mieux comprendre ce phénomène et d'identifier les facteurs suppresseurs (Su(var)s) et activateurs, « *enhancer* », (E(var)s) du PEV (pour une revue, Girton and Johansen, 2008) (figure 14).

Ce phénomène de répression par recrutement d'un gène au niveau de l'hétérochromatine péricentromérique a aussi été décrit par la suite dans différents organismes, de la levure aux mammifères (Dillon and Festenstein, 2002).

La chromatine est donc la matrice sur laquelle la transcription va avoir lieu. Le degré de compaction de la chromatine joue un rôle sur l'accessibilité des facteurs de transcription à l'ADN et est donc un facteur de régulation de l'initiation de la transcription. Trois paramètres principaux influencent l'ouverture ou la fermeture de la chromatine : les modifications post-traductionnelles des histones, l'incorporation de variants d'histones spécifiques à la chromatine, et la méthylation des cytosines au sein de l'ADN. Seule l'importance des modifications post-traductionnelles des histones et des variants d'histones sur l'initiation de la transcription sera abordée ici. Bien que la méthylation de l'ADN joue un rôle très important sur la condensation de la chromatine et la régulation de la transcription, ses mécanismes de mise en place et d'action ne seront pas abordés dans cette introduction.

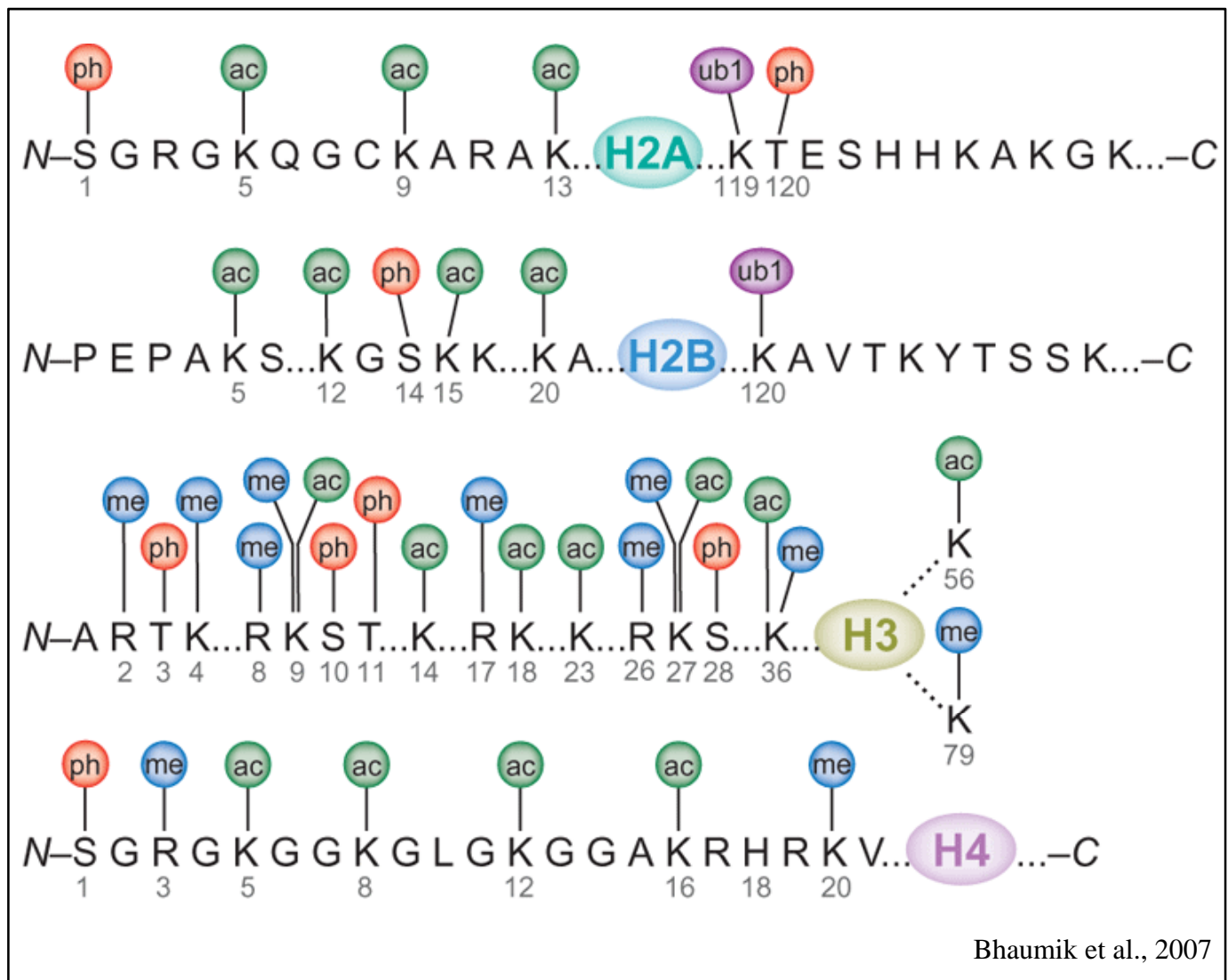


Figure 15 : Principales modifications post-traductionnelles des histones de cœur. Les modifications post-traductionnelles les plus étudiées incluant phosphorylation (ph), acétylation (ac), monoubiquitination (ub1), méthylation (me) sont schématisées ici. Le numéro du résidu modifié est renseigné en gris. Les domaines globulaires, domaines « histone fold » des histones de cœur sont représentés par des formes ovales.

II. Modifications post-traductionnelles des histones et compaction de la chromatine

1. Généralités

a. Mise en place et modes d'action

Comme d'autres protéines, les histones peuvent être la cible de diverses modifications post-traductionnelles. Au moins neuf types de modifications distincts peuvent être présents sur les histones (méthylation, acétylation, phosphorylation, ubiquitination, sumoylation, ADP ribosylation, déimination, citrullination et isomérisation des prolines), chaque modification pouvant apparaître sur un ou plusieurs acides aminés (méthylation des lysines et des arginines, acétylation des lysines, phosphorylation des sérines et des thréonines par exemple). Avec l'utilisation d'anticorps spécifiques ou suite à des études par spectrométrie de masse, plus de soixante sites de modifications ont été recensés sur les quatre histones de cœur ainsi que sur l'histone H1 (figure 15) (pour une revue, Kouzarides, 2007). Les sites de modifications des histones se trouvent majoritairement au niveau des extrémités N-terminales non structurées, mais quelques rares modifications ont également été décrites sur le domaine central et l'extrémité C-terminale.

Ces modifications sont catalysées par des enzymes spécifiques selon le type de modification, le résidu et l'histone mis en jeu. Leurs fonctions au sein de la chromatine sont assurées selon deux types de mécanismes. (1) Les modifications des histones peuvent, grâce à leurs propriétés chimiques intrinsèques, stabiliser ou déstabiliser les interactions ADN/histones ou les interactions entre les différents nucléosomes et ainsi agir sur le degré de compaction de l'ADN. Entre toutes les modifications connues, l'acétylation des lysines est la plus représentative de ce phénomène. En effet, elle neutralise la charge basique de ce résidu, diminuant les interactions à l'intérieur et entre les nucléosomes, rendant alors l'ADN plus accessible pour d'autres facteurs et facilitant donc la transcription par exemple (Hong et al., 1993) (Li et al., 1993b). (2) Les modifications post-traductionnelles des histones peuvent également permettre le recrutement de protéines non-histones ayant une affinité pour une modification spécifique. Les protéines recrutées peuvent soit posséder une activité enzymatique permettant d'autres modifications en cascade de la chromatine pour en induire

son ouverture ou sa condensation, soit servir d'intermédiaire pour le recrutement d'autres protéines comme l'ARN pol II par exemple (Kouzarides, 2007).

Par leur importance sur l'état de compaction de la chromatine et le recrutement de nombreux facteurs, les modifications post-traductionnelles des histones jouent des rôles majeurs dans la régulation et le bon déroulement de nombreux mécanismes cellulaires comme la transcription, la réparation de l'ADN et la réplication par exemple et permettent également le maintien de l'intégrité du génome en régulant la structure de la chromatine à certaines régions clés comme nous le verrons par la suite pour les centromères par exemple.

Nous nous intéresserons par la suite essentiellement aux rôles des modifications des histones sur la condensation de la chromatine en lien avec la transcription.

b. La théorie du code histone

Il a été proposé que la modification des histones soit un moyen pour la cellule de stocker de l'information, indépendamment de la séquence d'ADN, permettant alors à différentes cellules d'un même organisme, contenant le même ADN, d'exprimer différents gènes. Il a alors été supposé que la lecture de cette information devait suivre un code, le « code histone » de la même manière que la lecture de la séquence d'ADN s'effectue via le code génétique (Turner, 1993) (Strahl and Allis, 2000) (Jenuwein and Allis, 2001). La notion de code supposait qu'une modification donnée, ou une association de modifications, (le signe du code) devait entraîner une conséquence prévisible dans la cellule (la signification du signe) grâce à un système de lecture intermédiaire, les protéines de liaison aux histones modifiées. Avec l'augmentation du nombre de modifications identifiées, et la meilleure compréhension des rôles de ces modifications, il apparaît aujourd'hui que le code histone, dans la simplicité et la dualité avec laquelle il avait été décrit, une modification d'histones entraînant une réponse cellulaire, n'est pas réaliste (Turner, 2007). Par exemple, la triméthylation de la lysine 9 de l'histone H3 (H3K9me3), modification d'abord découverte pour son rôle au sein de l'hétérochromatine silencieuse, peut également être trouvée sur des gènes actifs (Vakoc et al., 2005). Le lien entre les modifications des histones et la transcription semble beaucoup plus complexe qu'il n'y paraissait au départ, chaque modification post-traductionnelle pouvant jouer un rôle différent selon les autres modifications auxquelles elle est associée, selon sa localisation sur le génome, ou selon les cellules étudiées.

Cependant, malgré la complexité apparente de leur lecture, l'importance de l'information véhiculée par les modifications des histones n'est en aucun cas remise en cause.

2. Un exemple : la méthylation des histones

a. Généralités

La méthylation des histones est la modification la plus caractérisée à l'heure actuelle. Elle peut être déposée sur de nombreux résidus, lysine ou arginine, des histones H3 et H4. La diversité des modifications possibles est augmentée par le fait que chaque résidu peut être méthylé de trois manières différentes, les lysines pouvant être mono-, di- ou triméthylées, et les arginines étant soit monométhylées, soit diméthylées de façon symétrique ou asymétrique.

Les méthylations sont déposées par des méthyltransférases d'histones (HMT : *histone methyltransferase*). L'activité de la plupart des HMT est assurée par un domaine catalytique très conservé chez les eucaryotes, le domaine SET. Ce nom est l'acronyme des trois premières protéines partageant ce domaine identifiées chez la drosophile : Su(var)3-9, E(z), Trx, trois HMT dont nous verrons le rôle par la suite (Jenuwein et al., 1998). Les HMT sont des enzymes très spécifiques, une même modification peut être déposée par des enzymes différentes selon le contexte cellulaire et génomique, et inversement, une enzyme peut être responsable de plusieurs modifications, mono-, di- ou triméthylation, généralement sur le même résidu.

Par analogie avec le rôle des acétylations, on pensait que la méthylation des histones devait affecter la conformation, l'hydrophobicité ou l'affinité pour l'ADN des histones méthylées. Des données structurales suggèrent par exemple que la triméthylation de H4K20 augmente la stabilité et la condensation des nucléosomes *in vitro* (Lu et al., 2008). Néanmoins, il est maintenant clair que le rôle majeur des méthylations est de créer des sites de reconnaissance pour d'autres protéines. Les histones méthylées sont reconnues par des protéines possédant des domaines particuliers : chromodomaine, domaine Tudor et domaine MBT appartenant aux domaines de la famille Royal, et domaine PHD.

La méthylation des histones a longtemps été considérée comme une marque non seulement stable, mais surtout irréversible. Si la mise en place de certaines modifications d'histones est nécessaire à l'établissement d'un programme transcriptionnel spécifique au cours du développement et de la différenciation cellulaire, il est assez intuitif de penser que la suppression et le remplacement d'autres marques doivent aussi faire partie du processus. Ce n'est que récemment que les premières déméthylases ont été découvertes, nous en verrons par la suite quelques exemples. Ces enzymes sont importantes au cours du développement, de la

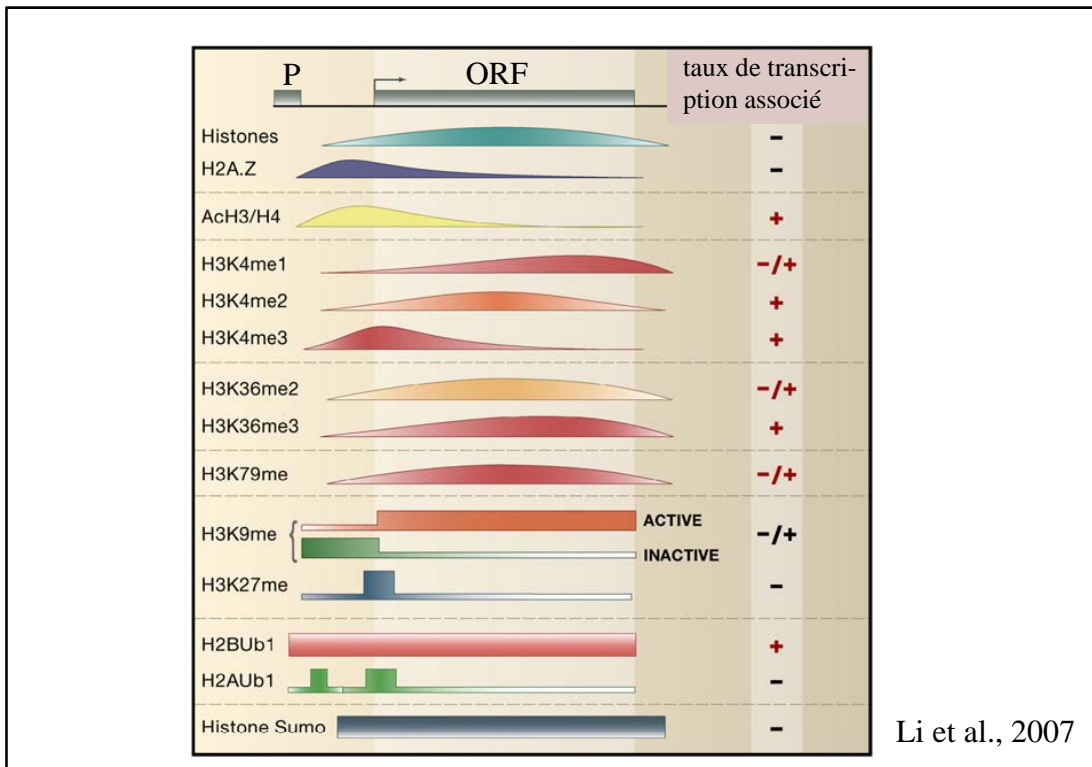


Figure 16 : Distribution des modifications des histones au sein du génome. La distribution est notée relativement au TSS d'un gène arbitraire (noté $\overrightarrow{\text{P}}$), à son cadre ouvert de lecture (ORF) et à son promoteur (P), en lien avec leur rôle sur la transcription. (-) : rôle répresseur; (+) : rôle activateur. Les courbes représentent les modèles de distribution déterminés par des études à l'échelle du génome. Les carrés indiquent les données obtenues à partir d'un petit nombre d'études seulement. A l'exception des données sur H3K9 et H3K27, la plupart des données sont basées sur des gènes types de levure.

différenciation cellulaire et sont donc impliquées dans l'apparition de certaines maladies comme certains cancers (pour une revue, Agger et al., 2008).

Bien que le rôle précis des méthylation des histones dépende de beaucoup de paramètres, ces modifications, comme toutes les modifications des histones, peuvent être classées de manière simplifiées en deux groupes : les modifications permissives caractéristiques d'une chromatine ouverte et d'une transcription active, et les modifications répressives caractérisant plutôt une chromatine fermée. Les principales méthylation permissives sont les triméthylations des lysines 4, 36 et 79 de l'histone H3. La triméthylation des lysines 9 et 27 de l'histone H3, ainsi que celle de la lysine 20 de l'histone H4 sont en revanche connues comme des mutations répressives.

Nous nous intéresserons ici particulièrement aux méthylation des lysines 4, 9 et 27 de H3 pouvant chacune être présentes sous trois formes, mono-, di- et triméthylation.

b. La méthylation de H3K4

- Mise en place

La première enzyme de méthylation de H3K4 découverte est la protéine Set1 chez la levure *S. cerevisiae*. L'homologue de Set1 chez la drosophile est une protéine connue de longue date, la protéine *Trithorax* (TRX) (Nislow et al., 1997). Sa mutation entraîne, chez la drosophile, des problèmes au cours du développement qui se traduisent, entre autres, par l'apparition de trois paires d'ailes au lieu d'une seule (Ingham, 1985). L'étude de TRX a permis de démontrer qu'elle est un régulateur positif de l'expression des gènes *HOX* grâce à son activité HMT (Ingham, 1985). Plus généralement, la méthylation de la lysine 4 de H3 est une marque de la chromatine ouverte et de la transcription active dans de nombreux organismes comme la levure, la drosophile et les mammifères. Des études à haut débit ont montré une évolution du degré de méthylation de H3K4 le long de la phase codante des gènes, H3K4me3 étant particulièrement abondante au niveau du TSS des gènes actifs, H3K4me2 étant plus fréquente au cœur de la région codante des gènes transcrits, et H3K4me1 plutôt présente à la fin de la phase codante (Li et al., 2007) (figure 16).

Les mammifères possèdent six gènes de la famille de TRX : *Set1a*, *Set1b*, *Mll1*, *Mll2*, *Mll3* et *Mll4*. Ils possèdent également quatre autres HMT spécifiques de H3K4, ASH1L, SET7/9, SMYD1, SMYD3 et PRDM9. Les HMT de la famille de Set1 agissent au sein de complexes contenant tous les protéines WDR5, RbBP5 et ASH2L (Dou et al., 2006). La déplétion de RbBP5 ou de WDR5 induit une diminution importante du taux global de

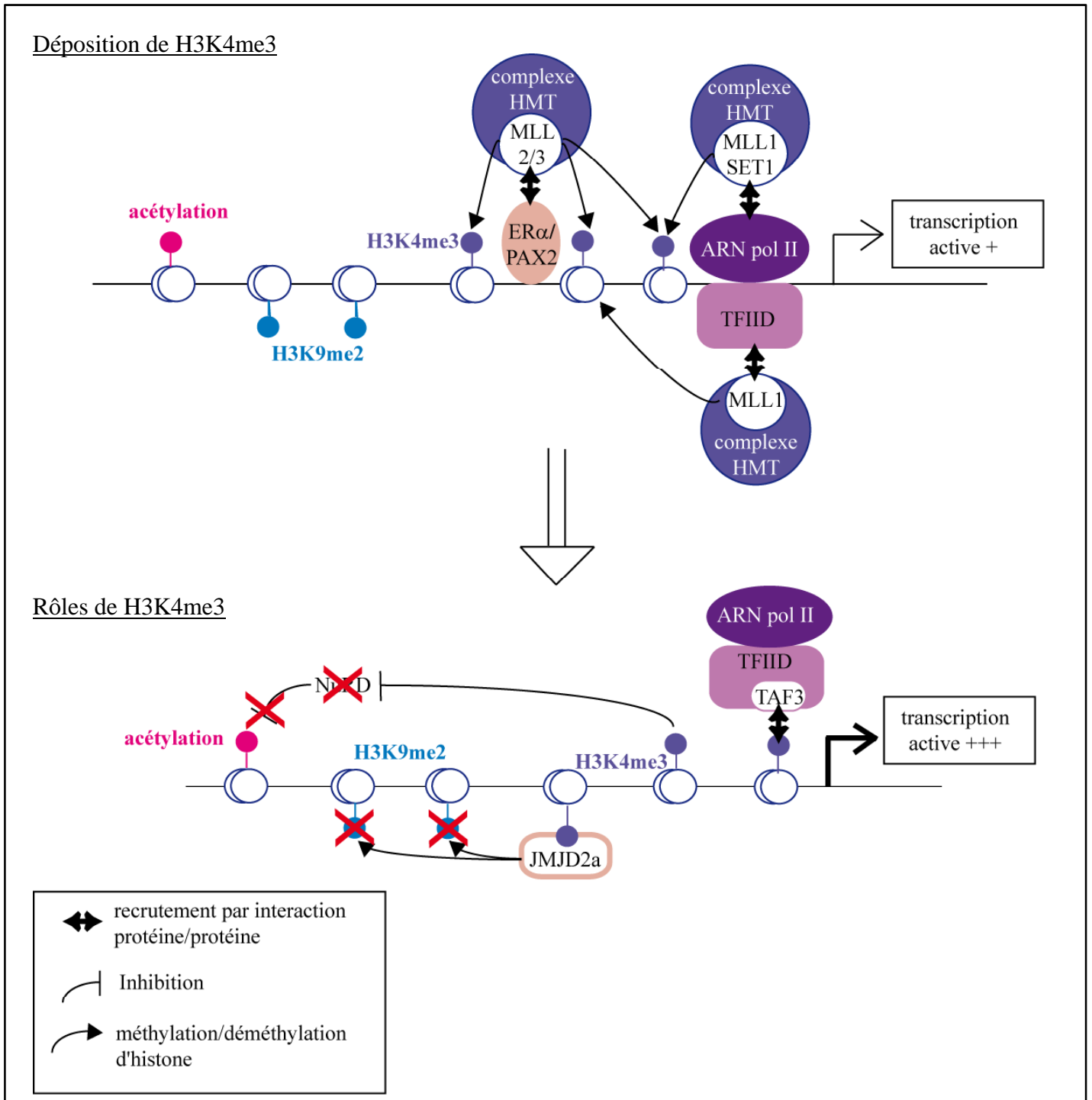


Figure 17 : Déposition et rôles de H3K4me3 à la chromatine. Des exemples de mécanismes de déposition et d'action de H3K4me3 dans le renforcement d'une transcription active sont schématisés ici. Les protéines sont représentées par des formes vides et les complexes par des formes pleines. Les complexes responsables de la méthylation de H3K4 peuvent être recrutés par leur interaction avec des facteurs de transcription ou avec la machinerie de transcription de base. La présence de H3K4me3 permet à son tour le recrutement de la machinerie de transcription de base, la déméthylation de H3K9 et le maintien de l'état acétylé de la chromatine.

H3K4me2 et H3K4me3, et la déplétion de ASH2L réduit seulement le taux global de H3K4me3 (Dou et al., 2006). Les résultats de ces expériences suggèrent que les protéines de la famille Set1 sont responsables de la majorité des di- et triméthylation de H3K4 dans les cellules. Elles sont également capables de catalyser la monométhylation de H3K4.

Les protéines de la famille Trithorax peuvent être recrutées à la chromatine de différentes manières. SET1a, SET1b et MLL1 font partie de complexes interagissant avec l'ARN pol II sous sa forme hypophosphorylée (Milne et al., 2005) (Lee and Skalnik, 2008). MLL1 interagit également avec des composants de TFIID et directement avec la chromatine grâce à l'ADN riche en CpG non méthylés (Birke et al., 2002). Les protéines MLL2 et MLL3 sont en revanche amenées à la chromatine par des facteurs de transcription, ER α (Mo et al., 2006) et PAX2 (Patel et al., 2007) respectivement, par exemple. Grâce à ces interactions, les complexes de méthylation de H3K4 peuvent être ciblés au promoteur et TSS des gènes transcrits. La méthylation de H3K4 semble donc marquer des gènes où la transcription est active pour potentiellement maintenir ou augmenter les niveaux d'expression.

- Rôles à la chromatine

Ensuite, la présence de H3K4me à la chromatine peut avoir deux conséquences principales. (1) Elle peut être reconnue par des protéines à activité enzymatique et induire leur recrutement à la chromatine, comme par exemple la protéine JMJD2A qui lie H3K4me grâce à son domaine Tudor et entraîne la déméthylation de H3K9 (Huang et al., 2006). Il a également été montré que le complexe TFIID pouvait interagir directement avec H3K4me3 grâce au domaine PHD de TAF3, permettant donc le recrutement de la machinerie de transcription de base (Vermeulen et al., 2007). (2) La présence de cette marque peut également inhiber l'association d'autres complexes à la chromatine. Elle prévient en effet la liaison du complexe de déacétylation des histones, NuRD, maintenant alors un état acétylé de la chromatine, permettant donc l'accès aux protéines nécessaires à la transcription par exemple (Zegerman et al., 2002) (figure 17).

- Suppression de la marque

En plus de la machinerie de méthylation de H3K4, il existe également une machinerie de déméthylation de H3K4 très conservée de la levure à l'humain. LSD1 (*Lysine Specific Demethylase 1*) est la première enzyme déméthylant les lysines des histones à avoir été découverte. Elle convertit H3K4me1 et H3K4me2 en H3K4 non métylée par une réaction de déméthylation oxydative (Shi et al., 2004). L'activité enzymatique de LSD1 peut être modulée

par son association avec d'autres protéines comme le complexe CoREST (*co-repressor for RE1 silencing transcription factor*) par exemple (Shi et al., 2005). Plus tardivement, les protéines de la famille JARID1 ont également été identifiées comme déméthylant H3K4 et plus particulièrement H3K4me3. L'activité catalytique des protéines de la famille JARID1 est assurée par un domaine hautement conservé entre de nombreuses déméthylases, le domaine JmjC (Christensen et al., 2007) (Klose et al., 2007) (Lee et al., 2007a) (Tahiliani et al., 2007). En plus du domaine JmjC, les protéines de la famille JARID1 contiennent plusieurs domaines remarquables, le domaine JmjN, important pour le maintien de l'intégrité structurale du domaine JmjC (Chen et al., 2006), un domaine de liaison à l'ADN (Wilsker et al., 2002) et plusieurs domaines PHD. Tout comme LSD1, l'activité des protéines de la famille JARID1 peut être modulée par des interactions avec d'autres facteurs (pour une revue, Benevolenskaya, 2007).

c. La méthylation de H3K9

La méthylation de la lysine 9 de H3 est une des marques de la chromatine condensée la plus étudiée. La méthylation de H3K9 est associée à la répression des gènes de l'euchromatine, répression qui peut être ponctuelle dans le temps et dans l'espace, ainsi qu'à la condensation de la chromatine à plus grande échelle, dans les régions hétérochromatiques constitutives.

- Mise en place

Chez les mammifères, il existe au moins 5 enzymes catalysant la méthylation de H3K9, SUV39H1, SUV39H2, G9a en complexe avec KMT1D, ESET et SETDB2. Le gène *Su(var)3-9* de la drosophile et son homologue *Clr4* chez *S. pombe*, ont été identifiés les premiers par génétique inverse, en étudiant des mutants présentant une suppression du phénomène de PEV. Des études biochimiques de leurs homologues humain et murin, ont permis de révéler que cette protéine a une activité enzymatique spécifique de méthylation de H3K9 (Rea et al., 2000).

Chez la souris, la mutation de *Suv39h1* ou *Suv39h2* n'entraîne pas de phénotype particulier, les souris sont viables et fertiles. En revanche les souris doublement mutées pour *Suv39h1* et *Suv39h2* montrent des problèmes développementaux à partir de E12,5 (12,5^{ème} jour embryonnaire), une diminution de la viabilité associée à un risque élevé de développer des tumeurs et des défauts de la méiose mâle (Peters et al., 2001). Ces résultats sont à mettre

en lien avec une perte spécifique de H3K9me3, au niveau de l'hétérochromatine péricentrique. Cependant, dans les mutants, le niveau de condensation de la chromatine péricentrique ne semble pas impacté, et la perte de H3K9me3 semble être remplacée par une augmentation de H3K9me1 et H3K27me3 à ces mêmes régions, suggérant que différentes enzymes de modification de H3 peuvent agir pour compenser la perte d'une autre et maintenir la stabilité du génome (Peters et al., 2003).

SETDB2 triméthyle H3K9 au niveau des éléments répétés dispersés et des répétitions centromériques. Sa délétion induit une perte des protéines spécifiques du centromère (Falandry et al., 2010).

La mutation de la protéine G9a induit, chez la souris, des retards de croissance et une mortalité à E8,5. Contrairement à la déplétion de l'activité des protéines SUV39H, la perte de G9a entraîne la diminution spécifique de H3K9me3 au niveau de l'euchromatine, et non au niveau de l'hétérochromatine (Tachibana et al., 2002).

La protéine ESET semble être également responsable de la méthylation de H3K9 dans les régions euchromatiques. Sa déplétion entraîne la mort des embryons entre E3,5 et E5,5 (Dodge et al., 2004). Toutes ces données *in vivo* montrent que la méthylation de H3K9 à des régions génomiques différentes sont assurées par des HMT spécifiques.

- Rôles à la chromatine, recrutement des protéines de la famille HP1

La répression de la transcription par H3K9me implique le recrutement de la protéine HP1 au promoteur des gènes réprimés. La protéine HP1 a été identifiée chez la drosophile sous le nom de Su(var)2-5. Quand elle est mutée, Su(var)2-5 induit une suppression du phénomène de PEV, c'est-à-dire une diminution de l'envahissement de l'hétérochromatine (Eissenberg et al., 1990). Chez la souris et l'homme, HP1 possède trois isoformes présentant une localisation sub-nucléaire spécifique dans les cellules somatiques : HP1 α est principalement associée aux régions d'hétérochromatine péricentrique, HP1 γ montre une localisation plus euchromatique, et HP1 β présente une répartition intermédiaire (Nielsen et al., 1999). Ces trois protéines montrent une organisation structurale commune composée d'un chromodomaine N-terminal, responsable de la liaison à H3K9me_{2/3} et au domaine central de H3 (Bannister et al., 2001) (Lachner et al., 2001) (Nielsen et al., 2001), d'un domaine *chromoshadow* C-terminal responsable des interactions protéine-protéine (Nielsen et al., 2001), et d'un domaine *Hinge*, central permettant l'association de HP1 avec des acides nucléiques (Maison et al., 2011) (Muchardt et al., 2002).

Le dogme selon lequel H3K9me serait associé exclusivement à un rôle répresseur de la transcription a été mis à mal par la mise en évidence de H3K9me3 et de HP1 γ dans la région codante de gènes actifs dépendamment de la présence de la forme hyperphosphorylée de l'ARN pol II (Vakoc et al., 2005). Ceci suggèrerait un rôle de H3K9me3 différent selon s'il se localise dans la région codante ou au promoteur des gènes, mais cette hypothèse n'a pas encore été démontrée à l'échelle du génome.

- Suppression de la marque

H3K9 peut être déméthylée par LSD1, tout comme H3K4, conférant à cette protéine à la fois un rôle dans l'activation et la répression de la transcription selon le complexe dans lequel elle se trouve (Metzger et al., 2005). Des déméthylases possédant un domaine JmjC peuvent également déméthyliser H3K9. Les protéines de la famille JMJD2 déméthylent H3K9me2 et H3K9me3, (Fodor et al., 2006) (Cloos et al., 2006) (Whetstine et al., 2006) elles possèdent, en plus du domaine JmjC catalytique, un domaine JmjN, des domaines PHD et des domaines Tudor. H3K9me1 peut être déméthylée par JHDM2A (Yamane et al., 2006), protéine de la famille JHDM2, également caractérisée par la présence d'un domaine JmjC, et par la présence de domaines en doigts de zinc (pour une revue, Klose et al., 2006).

d. La méthylation de H3K27

- Mise en place

En parallèle avec la découverte des gènes de la famille Trithorax chez la drosophile, une autre famille de gènes a été identifiée, la famille Polycomb. Les complexes Polycomb, à l'opposé des complexes Trithorax, ont un rôle sur la répression de la transcription (pour une revue, Ringrose and Paro, 2004) et sont responsables de la méthylation de H3K27 chez la drosophile comme chez les mammifères (Czermin et al., 2002) (Kirmizis et al., 2004). La protéine EZH2 (*Enhancer of zeste 2*) possède un domaine SET et est responsable de la méthylation de H3K27 (Cao et al., 2002) (Kuzmichev et al., 2002). Elle fait partie du complexe PRC2 (*Polycomb repressive complex 2*) en association avec quatre autres protéines : SUZ12 (Suppressor of zeste 12), responsable de la stabilisation du complexe (Pasini et al., 2004), EED (Embryonic ectoderm development) conférant la spécificité du complexe pour H3K27 (Kuzmichev et al., 2004), et deux protéines de liaison aux histones RBBP7 et RBBP4. Alors que EED est requise pour les trois états de méthylation possibles de H3K27 (Montgomery et al., 2005), la protéine EZH2 permet essentiellement les di- et

triméthylations de H3K27. Il a été suggéré qu'EZH1, homologue de EZH2, pourrait, en association avec EED et SUZ12, être responsable de la monométhylation de H3K27 (Shen et al., 2008), mais ces données restent controversées (Margueron et al., 2008). Chez les plantes, la monométhylation de H3K27 semble être assurée par un complexe complètement différent du complexe PRC2, ATRX1 et 2 (Jacob et al., 2009).

Les embryons n'exprimant pas EZH2 cessent de se développer juste après l'implantation, ou au plus tard avant la fin de la gastrulation, entre E7,5 et E10,5 (O'Carroll et al., 2001) (Puschendorf et al., 2008). De plus, l'étude de la déplétion de la contribution maternelle pour EZH2 dans des zygotes, tout en conservant la contribution paternelle commençant au stade 4-cellules, montre des conséquences à long terme sur le développement avec des retards de croissance suggérant un rôle important de la protéine dans les premières heures de développement (Erhardt et al., 2003). De même, les mutations de *Eed* et *Suz12* sont létales avant E10,5 et les embryons mutés montrent une diminution globale du taux de H3K27me, suggérant que ces trois protéines sont nécessaires au fonctionnement correct du complexe PRC2 (Faust et al., 1995) (Pasini et al., 2004).

Chez la drosophile, des protéines liant l'ADN permettent de recruter le complexe PRC2 à des séquences régulatrices spécifiques, les éléments de réponse à Polycomb (PRE) (Schwartz et al., 2006). Bien que l'existence de PRE ait également été suggérée chez les mammifères (Woo et al., 2010), leur abondance et leur rôle n'est pas aussi évident que chez la drosophile. En revanche, il est maintenant démontré que le recrutement des complexes Polycomb à la chromatine peut se faire via leur interaction avec des ARN non codants (ARNnc). Nous détaillerons deux exemples par la suite. De plus, la protéine EED est capable de lier des histones déjà méthylées sur H3K27 et permet ainsi un enrichissement de PRC2 à des régions riches en H3K27me.

- Rôles à la chromatine, recrutement du complexe PRC1

L'action de H3K27me sur l'état de compaction de la chromatine se fait grâce à sa liaison par le deuxième complexe Polycomb mammifère, PRC1. A cause de la duplication de plusieurs gènes *Polycomb*, différents complexes PRC1 fonctionnels existent chez les mammifères, dépendamment du type cellulaire étudié et du stade de développement (Otte and Kwaks, 2003) (Whitcomb et al., 2007). Cinq protéines CBX (CBX2/4/6/7/8) possèdent des chromodomaines conservés mais d'affinités différents pour H3K27me3 et H3K9me3 (Bernstein et al., 2006b). Les protéines RING1a et RNF2, toutes deux faisant partie de PRC1, possèdent un domaine RING à activité E3 ubiquitine ligase (de Napoles et al., 2004) (Wang et

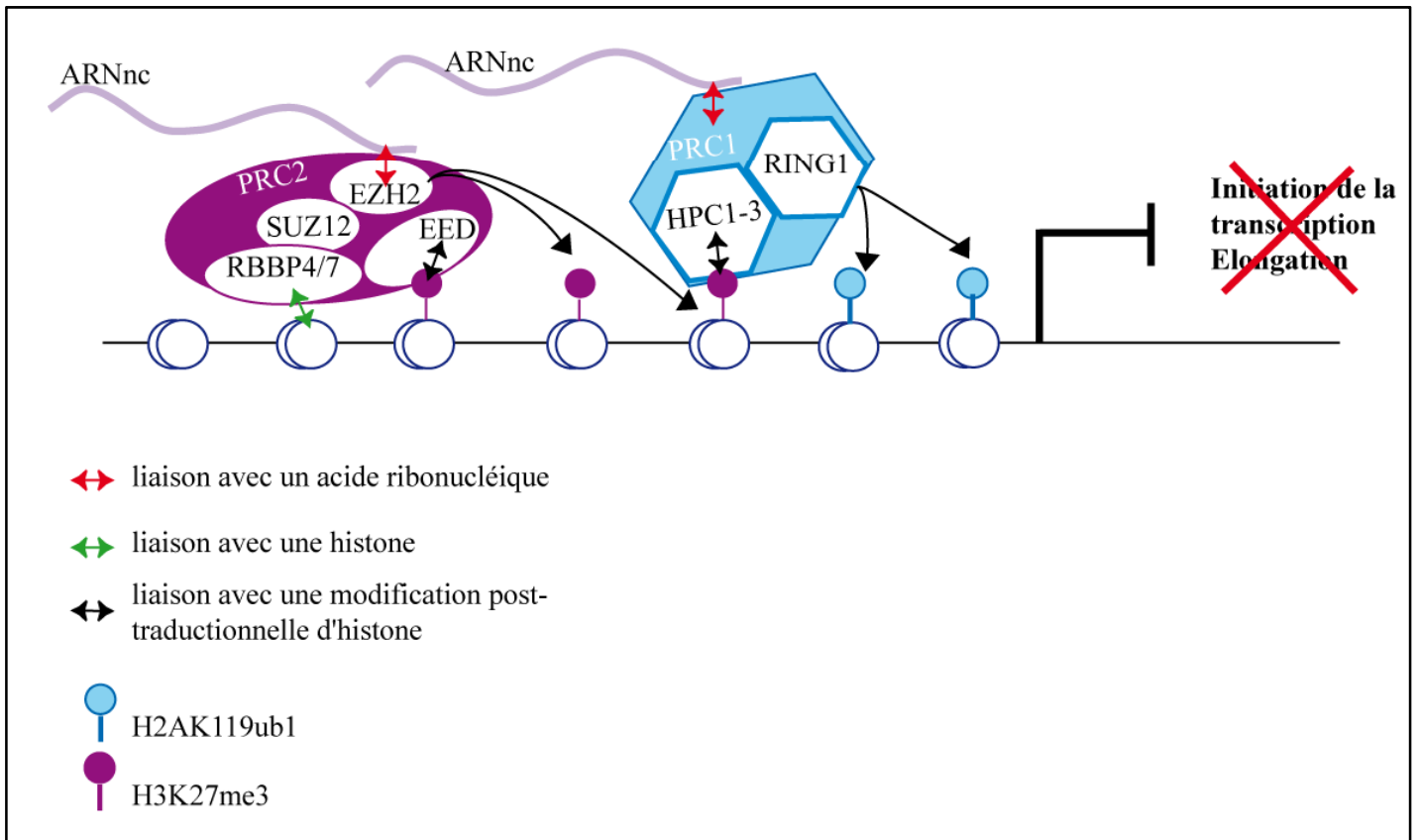


Figure 18 : Dépôt et rôles de H3K27me3 à la chromatine. Les complexes sont représentés par des formes pleines et les protéines par des formes vides. Les interactions sont schématisées par des doubles flèches. Les flèches simples représentent la catalyse des modifications des histones par les enzymes. Le recrutement de PRC2 à la chromatine peut se faire par liaison aux histones, par liaison à H3K27me3 ou par liaison d'un ARNnc. PRC2 permet alors la di- et tri-méthylation de H3K27 reconnues par PRC1 qui monoubiquitinylyle H2AK119, inhibant l'initiation de la transcription et l'élongation des transcrits. PRC1 peut également être recruté directement à la chromatine par interaction avec un ARNnc.

al., 2004). Suite à la méthylation de H3K27 par PRC2, le complexe PRC1 est recruté à la chromatine et induit la monoubiquitination de H2AK119. *In vitro*, PRC1 peut induire la compaction de nucléosomes d'une manière qui ne nécessite pas la queue des histones (Shao et al., 1999). *In vivo*, le lien entre la fonction de PRC1 et la répression des gènes n'est pas connue dans les détails, mais il semble que ce n'est pas l'accès de la machinerie de transcription à la chromatine qui est affectée, mais plutôt son activité, l'initiation de la transcription n'ayant pas lieu en présence de PRC1 (Dellino et al., 2004). Une étude plus récente montre que la présence de H2AK119ub empêcherait le recrutement de facteurs d'élongation, inhibant la transcription (Zhou et al., 2008). Il est important de noter que bien que H3K27me₃, et donc l'activité de PRC2, soit nécessaire au recrutement de PRC1 à la chromatine à certains gènes (Boyer et al., 2006), cette dépendance n'est pas une généralité. En effet, lors du développement préimplantatoire de la souris, après le stade de zygote, l'association globale de PRC1 à la chromatine n'est pas affectée par l'absence de EZH2 (Puschendorf et al., 2008). De même, lors de l'initiation de l'inactivation du X chez les mammifères femelles, le recrutement de PRC1 est indépendant de la fonction de EED, donc de PRC2 (Schoeftner et al., 2006) (figure 18). De façon similaire, l'activité E3 ligase de PRC1 ne semble pas être absolument requise pour la compaction de la chromatine sur certains loci (Eskeland et al., 2010).

- Suppression de la marque

La déméthylation de H3K27me₃ et H3K27me₂ en H3K27me₁ peut être effectuée par deux histones déméthylases spécifiques, UTX et JMJD3 (pour une revue, Agger et al., 2007). Elles possèdent toutes les deux un domaine JmjC responsable de leur activité. Ces deux protéines interagissent avec des composants des complexes Trithorax, responsables de la méthylation de H3K4 (De Santa et al., 2007) (Lee et al., 2007b). La déméthylation de H3K27 est donc couplée à une méthylation de H3K4, rendant l'activation de la transcription plus efficace.

e. Les domaines bivalents

Dans les cellules ES, de nombreuses régions du génome sont marquées à la fois par H3K27me₃, une marque répressive, et H3K4me₃, une marque permissive. Ces régions sont donc qualifiées de domaines bivalents (Bernstein et al., 2006a). Les gènes au sein des domaines bivalents sont en grande majorité réprimés indiquant que la marque répressive de

H3K27me3 serait plus « forte » que la marque permissive H3K4me3. Après différenciation, les régions promotrices bivalentes deviennent pour la plupart marquées exclusivement soit par l'une ou l'autre des marques précédentes. Il a donc été supposé que les domaines bivalents permettaient de réprimer des gènes du développement, de différenciation, tout en les maintenant dans un état d'attente d'une future activation ou répression rapide au cours du développement. Il est important de noter que la présence de domaines bivalents n'est pas spécifique des cellules ES mais existe également dans des cellules progénitrices et différenciées. Pendant la différenciation cellulaire, des domaines bivalents existants disparaissent alors que d'autres semblent se former (Mohn et al., 2008).

Une explication plus fonctionnelle a été suggérée pour expliquer la présence de marques permissives en association avec H3K27me3. Comme nous le verrons par la suite, le recrutement de PRC2 à la chromatine peut se faire grâce à la présence d'ARNnc, transcrits en *cis* ou en *trans*. On peut donc supposer que la marque permissive H3K4me3 permettrait la transcription des ARNnc, responsables du recrutement de PRC2 et donc de la méthylation de H3K27 aux mêmes domaines (Kanhere et al., 2010).

Ces deux hypothèses expliquant l'existence de domaines bivalents ne sont pas exclusives et des domaines bivalents différents peuvent très bien répondre à des mécanismes de formation et des rôles différents.

III. Les variants des histones au sein de la chromatine

1. Généralités

Les gènes codant pour les formes majeures des histones, formes dites canoniques, sont présents, pour la plupart, en de multiples copies, souvent rassemblées en clusters, au sein du génome. Les ARN correspondants ne sont pas polyadénylés et leur traduction est régulée par la liaison de protéines spécifiques sur une structure particulière en boucle formée par la région 3' de leurs ARN (Arnold et al., 2008). Ils sont exprimés massivement au cours de la phase S et incorporés au sein de la chromatine de manière dépendante de la réplication (Osley, 1991).

En plus des formes canoniques des histones, des variants de remplacement dont la séquence primaire varie de celle des histones canoniques de quelques acides aminés à de larges domaines, ont été découverts pour H2A, H2B et H3 (Zweidler, 1984). Les gènes codant pour les formes de remplacement des histones sont le plus souvent uniques ou représentés

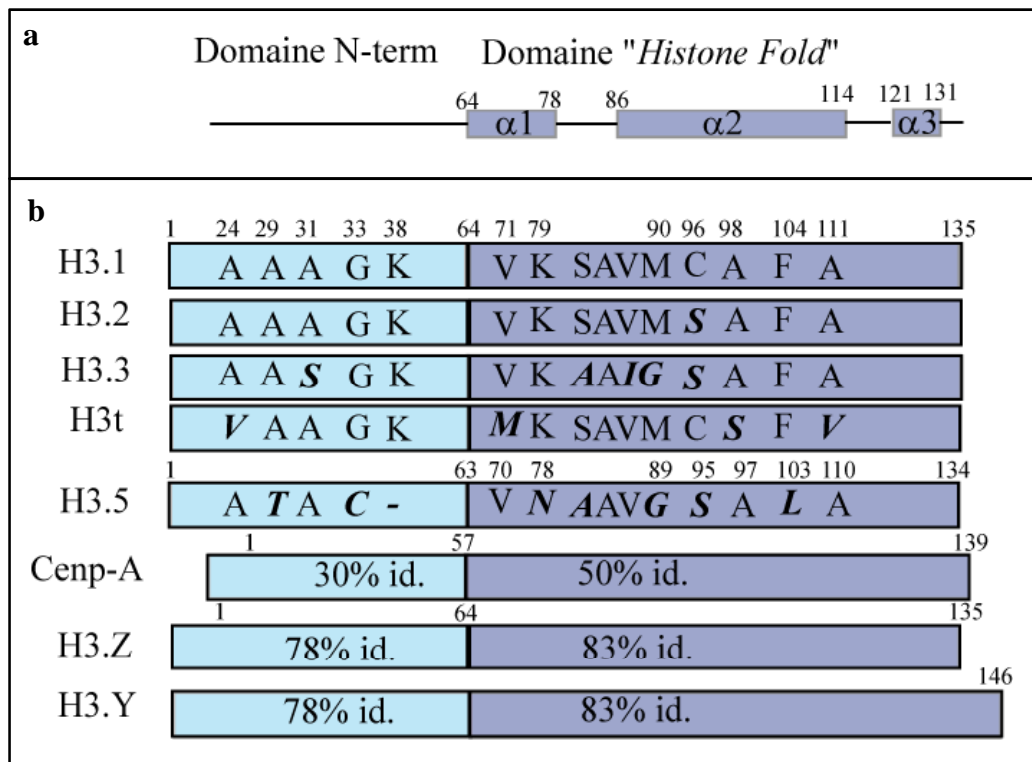


Figure 19 : Schéma des différents variants humains de l'histone H3. a : Structure de H3.1. **b :** Les acides aminés qui diffèrent entre les variants, ainsi que leur position, sont indiqués pour les variants H3.1, H3.2 H3.3 et H3t et H3.5. Lorsqu'ils diffèrent de la séquence de H3.1, les acides aminés sont notés en gras italique. Pour Cenp-A, H3.Z et H3.Y, les pourcentages d'identité avec H3.1 sont rapportés pour chaque domaine.

seulement par un faible nombre de copies dans le génome, ils codent de plus pour des ARNm polyadénylés. L'expression et l'incorporation des histones de remplacement ont lieu tout au long du cycle cellulaire. Ils ont un patron d'expression qui leur est propre et des propriétés au sein de la chromatine différentes de celles de l'histone majoritaire. L'incorporation spécifique de différents variants d'histones au sein du nucléosome contribue donc à la régulation du niveau de condensation de la chromatine.

Nous nous intéresserons ici seulement aux variants de l'histone H3, spécifiquement chez les mammifères.

2. Les variants de H3 (figure 19)

En plus des formes canoniques de H3, il existe chez les mammifères trois variants de remplacement de H3 : le variant spécifique des centromères CENP-A (existant sous le nom de CenH3 dans les autres organismes), le variant spécifique des testicules, H3t, et un variant ubiquitaire H3.3. Peuvent être ajoutés à cette liste trois variants caractérisés récemment, H3.X et H3.Y, spécifiques des primates (Wiedemann et al., 2010) et H3.5, présent exclusivement dans les tubes séminifères chez l'homme (Schenk et al., 2011). Nous aborderons ici seulement les caractéristiques des variants canoniques de H3, de H3.3, H3t et CENP-A.

a. H3.1 et H3.2, les formes canoniques de H3

Contrairement aux autres histones de cœur, il existe deux formes canoniques de l'histone H3, H3.1 et H3.2. Ces deux protéines ne diffèrent que par un seul acide aminé en position 96 (substitution Cystéine-Sérine). Elles sont toutes deux exprimées au moment de la phase S, et incorporées au sein de la chromatine de manière dépendante de la réplication, associée à la phase S (Ahmad and Henikoff, 2002) ou lors du processus de réparation (Polo et al., 2006). Dans les cellules HeLa, la déposition de H3.1 et H3.2 au sein de la chromatine fait intervenir la protéine chaperonne ASF-1 et le complexe CAF-1 (Tagami et al., 2004). CAF-1 interagit avec la protéine PCNA, présente à la fourche de réplication et ainsi permet la déposition de H3.1/2, en dimère avec H4, aux sites de synthèse d'ADN (Akey and Luger, 2003) (Shibahara and Stillman, 1999).

Du fait de leur très grande similitude, très peu d'études ont consisté à mettre en évidence des différences potentielles entre ces deux variants de H3. Néanmoins, suite à des données de spectrométrie de masse montrant que H3.2 serait enrichi en des marques

répressives comme H3K27me_{2/3} et H3K9me₃ par rapport à H3.1, il a été suggéré que H3.2 serait plutôt associé à une chromatine condensée (Hake et al., 2006). Cependant, les mécanismes moléculaires permettant à la cellule de différencier ces deux variants restent obscures.

b. H3.3

H3.3 est une des protéines les plus conservées, présente chez tous les eucaryotes (Malik and Henikoff, 2003). Elle est transcrite, chez la souris, à partir de deux gènes H3.3A et H3.3B. H3.3A est exprimé ubiquitairement, sa mutation est létale chez 50% des porteurs homozygotes, et les mâles survivants sont subfertiles (Couldrey et al., 1999). Ce phénotype suggère que H3.3 joue un rôle particulier au cours du développement et de la gamétogenèse, rôle que nous aborderons dans le chapitre suivant. Seulement cinq acides aminés différencient H3.3 de H3.1 (en position 31, dans la région N-terminale, et en positions 87, 89, 90 et 96 au sein du domaine « *histone fold* »). Cette différence de séquence, aussi faible soit-elle, est suffisante pour que H3.3 possède des propriétés différentes de H3.1 et H3.2 au sein de la chromatine.

Tout d'abord, des données de localisation de H3.3 au sein du génome dans des cellules somatiques montrent un enrichissement spécifique de H3.3 dans le corps des gènes actifs (Chow et al., 2005) (Delbarre et al., 2010), aux promoteurs de gènes actifs et inactifs ainsi qu'à des régions régulatrices géniques et intergéniques (Jin et al., 2009). Des études au spectromètre de masse ont montré que les modifications post-traductionnelles majoritairement associées à H3.3 sont les marques d'une transcription active. En effet, H3.3 est soumis à de nombreuses acétylations et est souvent méthylé sur la lysine 4 (Hake et al., 2006). De plus, les nucléosomes contenant H3.3 sont moins stables que ceux contenant la forme canonique de H3, en particulier lorsqu'ils contiennent également H2A.Z, un variant de l'histone H2A (Jin and Felsenfeld, 2007) (Jin et al., 2009). Il a également été observé chez la drosophile que la présence de H3.3 au sein d'un nucléosome le rendait moins favorable à interagir avec l'histone H1 donc à être présent dans des structures chromatiniennes très condensées (Braunschweig et al., 2009). Toutes ces données sont donc en faveur d'un rôle important de H3.3 au sein d'une chromatine active. Cependant, des cellules ES de souris déficientes pour un enrichissement de H3.3 au sein des gènes ne présentent pas de changements dramatiques de leurs niveaux de transcription en comparaison avec des cellules ES sauvages (Goldberg et al., 2010). De plus, dans les cellules ES et des fibroblastes embryonnaires de souris, H3.3

s'accumule également dans des régions très condensées du génome, les télomères (Goldberg et al., 2010).

La protéine HIRA est la première protéine chaperonne spécifique de H3.3 à avoir été découverte grâce à des études de spectrométrie de masse et par comparaison des complexes associés à H3.1 et H3.3 dans des cellules HeLa (Tagami et al., 2004). Plus récemment, suite à l'observation de locus (les télomères et certains sites de liaison à des facteurs de transcription) où la déposition de H3.3 n'est pas dépendante de la présence de HIRA dans des cellules souches embryonnaires et des précurseurs neuronaux, la protéine DAXX a été identifiée comme une autre chaperonne spécifique de H3.3 (Goldberg et al., 2010). Des résultats similaires ont ensuite été obtenus *in vitro* et dans des fibroblastes de souris (MEF : *mouse embryonic fibroblast*) (Drane et al., 2010). La déposition de H3.3 serait donc sous la dépendance d'au moins deux complexes distincts et indépendants, le premier contenant HIRA associée à ASF1 et une sous-unité du complexe CAF-1 et l'autre DAXX associée à la protéine ATRX. Ces deux complexes semblent réguler la déposition de H3.3 à différents sites dans le génome : la présence de H3.3 au corps des gènes actifs ou réprimés serait plutôt sous la dépendance de HIRA, alors que la présence de H3.3 aux télomères, aux répétitions péri-centriques et à certains sites de liaison à des facteurs de transcription, globalement aux régions régulatrices intergéniques, serait plutôt sous la dépendance de DAXX (Lewis et al., 2010) (Elsaesser and Allis, 2010). Une troisième protéine, la protéine DEK, a été suggérée comme pouvant être une chaperonne de H3.3 chez la drosophile (Sawatsubashi et al., 2010). Chez les mammifères, bien que DEK semble également pouvoir jouer le rôle de chaperonne d'histone (Kappes et al., 2011), sa spécificité pour H3.3 n'a pas été démontrée.

c. H3t, un variant spécifique des testicules

Le variant H3t possède quatre acides aminés de différence avec H3.1. Bien qu'apparemment exprimé faiblement dans les cellules somatiques, il est considéré comme un variant spécifique des testicules (Witt et al., 1996) (Govin et al., 2005). Il est déposé au sein de la chromatine de manière dépendante de la réplication par l'histone chaperonne NAP2, spécifique des testicules (Tachiwana et al., 2008). De récentes études structurales ont permis de montrer que des nucléosomes contenant le variant H3t sont moins stables que ceux contenant H3.1, cette instabilité pouvant ainsi être favorable à la réorganisation de la chromatine qui a lieu pendant la spermatogenèse (Tachiwana et al., 2010).

d. CENP-A, un variant spécifique des centromères

CENP-A est un variant de H3 spécifique de la chromatine centromérique. CENP-A est le variant de H3 qui diverge le plus de la forme canonique. Il possède une région N-terminale unique d'environ 50 acides aminés, domaine d'interaction avec ses partenaires spécifiques et une région C-terminale identique à plus de 60% avec H3.1 contenant le domaine « *histone fold* » (Sullivan et al., 1994). L'incorporation de ce variant au niveau des centromères et son association avec le complexe CENP-A^{NAP} (*CENP-A Nucleosome Associated Complex*) crée une structure particulière très repliée de la chromatine, et permet l'assemblage du complexe kinétochorien et la résistance aux forces appliquées par les fuseaux de microtubules au moment de la méiose et de la mitose (pour une revue, Black and Bassett, 2008).

IV. Régulation des modifications chromatiniennes : ARN non codant et hétérochromatinisation

Les modifications des histones ainsi que la présence de différents variants permettent donc de réguler le degré de condensation de la chromatine, le recrutement de certains facteurs et par conséquent la transcription. Si les modifications chromatiniennes jouent un rôle sur la régulation de la transcription, la mise en place de ces modifications peut elle-même être régulée grâce à la transcription. Un nombre de plus en plus important d'études montre que de nombreux ARN non codants sont responsables de l'établissement de certaines modifications d'histones, de la méthylation de l'ADN et de la formation et du maintien de l'hétérochromatine.

Après avoir défini les deux types d'hétérochromatine caractérisés chez les eucaryotes, nous verrons comment les ARNnc sont impliqués dans l'établissement de ces domaines chromatiniens.

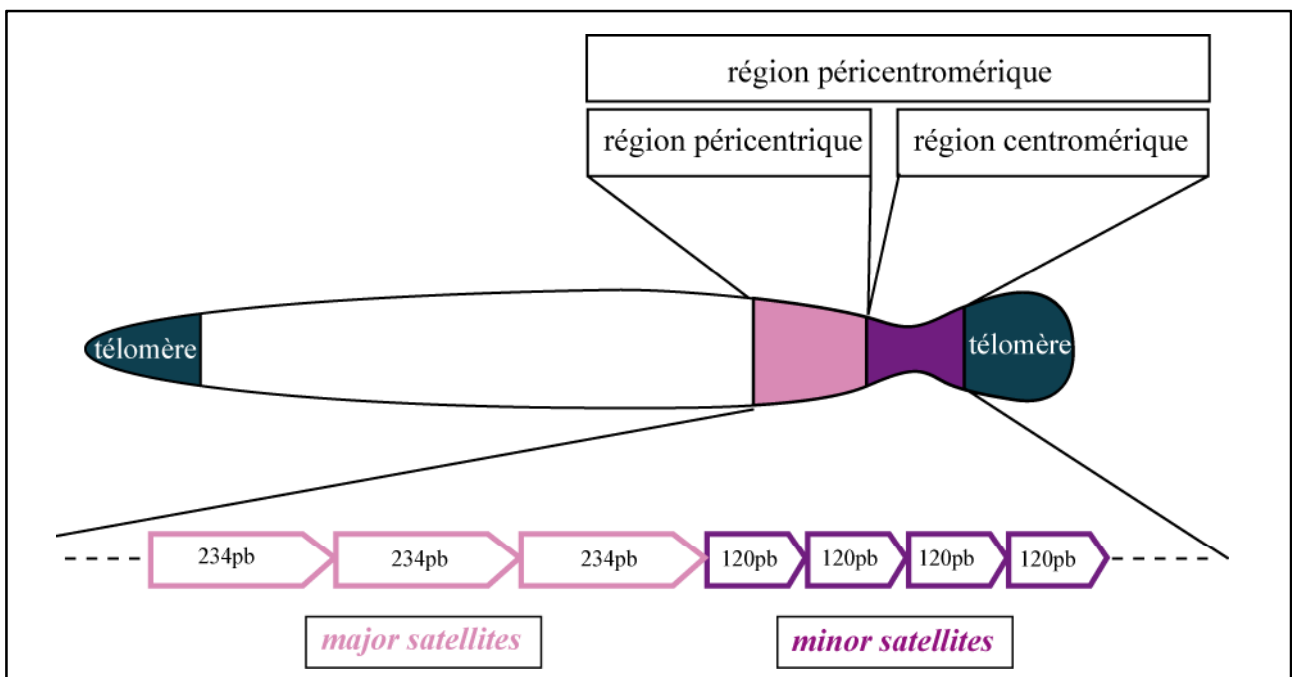


Figure 20: Représentation schématique des répétitions en tandem des régions péricentriques et centromériques. Les régions péricentriques et centromériques ainsi que les centromères sont localisés sur un chromosome acrocentrique de souris. Les répétitions péricentromériques en tandem sont schématisées par des flèches représentant les transcrits sens de ces répétitions.

1. Les différents types d'hétérochromatine

a. L'hétérochromatine constitutive

- Caractéristiques

L'hétérochromatine constitutive caractérise les régions du génome que l'on retrouve fortement condensées quels que soient le type cellulaire et l'organisme étudiés. Elle compose environ un tiers du génome humain et est représentée principalement par les régions péri-centriques, centromériques et télomériques. Ces régions sont pauvres en gènes mais riches en séquences répétées et leur condensation est nécessaire à la bonne ségrégation des chromosomes au cours de la mitose, à la protection de leurs extrémités et, plus généralement, au maintien de l'intégrité de la structure des chromosomes dans la cellule.

Suite à l'étude des marques chromatiniennes présentes aux régions péri-centriques, les marques caractéristiques de l'hétérochromatine constitutive ont été définies comme étant une hypoacétylation globale des histones et la présence de H3K9me3, H3K27me1, H4K20me3 et H3K64me3 (Peters et al., 2003) (Schotta et al., 2004) (Martens et al., 2005) (Daujat et al., 2009).

- Exemple des centromères

Les centromères sont caractérisés par la présence de séquences répétées en tandem. Les régions péri-centriques, périphériques, correspondent chez la souris à environ 6 Mb d'ADN où se succèdent des séquences de 234 pb répétées en tandem, les *major satellites*. Les régions centromériques, centrales, sont des répétitions d'une séquence de 120 pb, les *minor satellites*, sur environ 600 kb (figure 20) (Lehnertz et al., 2003). (Dans la suite de ce manuscrit, nous qualifierons de région péri-centromérique la région qui comprend les *major* et *minor satellites*).

Bien que la présence de séquences répétées aux régions péri-centromériques soit commune à tous les eucaryotes, leur séquence nucléotidique n'est pas conservée au cours de l'évolution. Chez tous les organismes en revanche, la fonction des centromères est d'assurer la ségrégation des chromosomes pendant la mitose. Pendant cette phase, un complexe protéique, le complexe kinétochorien s'assemble aux régions centromériques, lie les microtubules et permet une migration et une séparation efficace des chromatides sœurs. L'assemblage du complexe kinétochorien n'est pas régulé par la reconnaissance d'une

séquence conservée, mais plutôt par une organisation et une condensation particulière de la chromatine dans cette région.

Chez la souris, dans les cellules somatiques, l'ADN péricentrique est méthylé sur les sites CpG (Lewis et al., 1992) et il est enroulé autour d'histones hypoacétylées et enrichi en H3K9me3 (Peters et al., 2001), H4K20me3, H3K27me1 (Martens et al., 2005) (Peters et al., 2003) et H3K64me3 (Daujat et al., 2009). D'autres protéines non histones sont spécifiquement enrichies au niveau de l'hétérochromatine péricentrique comme c'est le cas de HP1 α et HP1 β liant H3K9me3 (Guenatri et al., 2004). Dans des fibroblastes de souris n'exprimant ni SUV39H1 ni SUV39H2, une diminution spécifique de H3K9me3 a été observée au niveau de la chromatine péricentrique, entraînant une perte de HP1 α . Les cellules présentent par ailleurs une augmentation de H3K9me1 et de H3K27me3 à ces mêmes régions (Peters et al., 2003). La condensation des centromères n'étant pas altérée dans ces mutants, ces expériences suggèrent que la méthylation de H3K9 et le recrutement de HP1 ne sont pas les seuls mécanismes permettant de maintenir la stabilité de l'hétérochromatine péricentrique.

Les régions centromériques sont également caractérisées par la présence de variants d'histones particuliers. La présence de CENP-A, variant de H3, aux régions centromériques est essentielle à l'organisation du kinétochore et à la ségrégation des chromosomes.

b. L'hétérochromatine facultative

La deuxième catégorie d'hétérochromatine correspond à l'hétérochromatine facultative. Comme l'hétérochromatine constitutive, elle est transcriptionnellement silencieuse mais peut se convertir en euchromatine et ainsi permettre la transcription selon des contextes temporels différents, stade du cycle cellulaire ou stade de différenciation, des contextes spatiaux différents, changement de localisation de la région au sein du noyau, ou des contextes parentaux différents, expression monoallélique d'un gène.

Les données moléculaires concernant les caractéristiques de l'hétérochromatine facultative sont majoritairement basées sur l'étude du phénomène d'inactivation d'un chromosome X chez les mammifères femelles. Au cours du développement préimplantatoire de la souris femelle, son chromosome X paternel est inactif puis réactivé au stade blastocyste dans les cellules de la masse cellulaire interne. Avant la gastrulation, un des deux chromosomes X, choisi au hasard, sera sujet à une condensation globale de sa chromatine, condensation transmise aux cellules filles et maintenue tout au long de la vie (Heard and Disteche, 2006) (Okamoto and Heard, 2009).

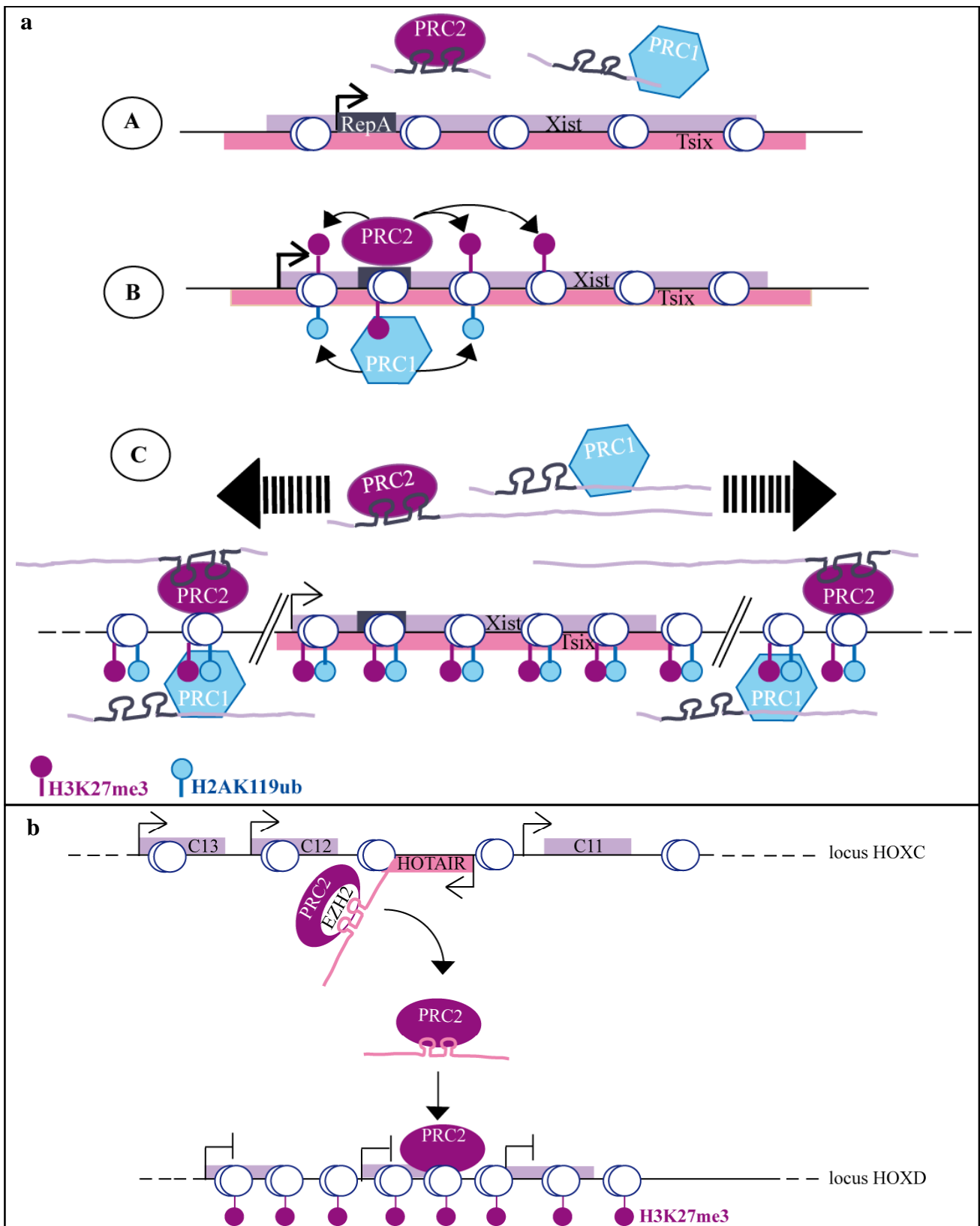


Figure 21 : Exemples de rôles des longs ARNnc dans le recrutement des complexes polycomb en *cis* dans le cas de Xist (a), et en *trans* dans le cas de HOTAIR (b). a : La transcription de RepA permet la formation d'un ARN formant une structure tige-boucle et recrutant PRC1 et PRC2 (A) en *cis* et permettant la méthylation de H3K27 et l'ubiquitination de H2AK119 (B). L'ARN de Xist transcrit par la suite, va se propager, couvrir le Xi et y recruter PRC1 et PRC2 (C). b : La transcription de HOTAIR au locus HOXC permet le recrutement de PRC2 en *trans* au gène HOXD et l'inhibition de sa transcription grâce à la méthylation de H3K27.

L'hétérochromatine facultative est donc caractérisée par une hypoacétylation des histones, et par la présence de H3K9me2, H3K27me3 et des protéines du groupe Polycomb (Bernstein et al., 2006b) (Rougeulle et al., 2004).

Malgré leur état condensé, une faible transcription des régions hétérochromatiques constitutives et facultatives est possible et dans certains cas nécessaire à leur mise en place.

2. Hétérochromatine facultative et longs ARN non codants chez les mammifères

Le génome des mammifères code pour un très grand nombre de longs ARNnc dont la séquence est conservée au cours de l'évolution. Chez l'homme, environ 20% de ces ARNnc ont été montrés comme interagissant avec le complexe PRC2, supposant un rôle de ces ARN dans le recrutement de complexes de modification des histones à la chromatine (Khalil et al., 2009). Nous aborderons ici deux exemples de longs ARNnc, Xist et HOTAIR.

Chez les mammifères, la formation d'hétérochromatine facultative au niveau du chromosome X inactif (Xi) chez la femelle dépend de la transcription d'un long ARNnc, Xist, *en cis*, par le Xi (Penny et al., 1996). Xist peut être immunoprécipité avec les composants du complexe PRC2 grâce à la présence d'un petit motif, nommé RepA, replié en structures particulières en tige-boucle (Zhao et al., 2008). La transcription de Xist au chromosome X permet donc le recrutement de PRC2, via ses sous-unités EZH2 et SUZ12, et la méthylation de H3K27 *en cis* (Zhao et al., 2008) (Kanhere et al., 2010). De plus, le complexe PRC1 peut également être directement recruté au Xi via sa sous-unité RING1b, de manière dépendante de l'ARN, mais indépendamment de la méthylation de H3K27 (Schoeftner et al., 2006) (figure 21a). La délétion du motif RepA de Xist n'empêchant pas le recrutement de PRC1 à la chromatine, RepA ne semble pas être le domaine de liaison de Xist à PRC1 (Schoeftner et al., 2006), ce domaine de liaison n'est pas connu. Ensemble, la transcription de Xist et le recrutement de PRC1 et PRC2 contribuent à l'inhibition de la transcription au Xi, à sa condensation et à la transmission de cet état inactif au cours des générations cellulaires.

Le recrutement des complexes Polycomb à la chromatine grâce aux longs ARNnc peut également se faire *in trans*. HOTAIR est un long ARNnc de 2,2kb permettant la répression *in trans* du locus HOXD via le recrutement de PRC2 et la méthylation de H3K27 chez l'homme (Rinn et al., 2007). Le recrutement de PRC2 se fait également via la sous-unité EZH2 grâce à

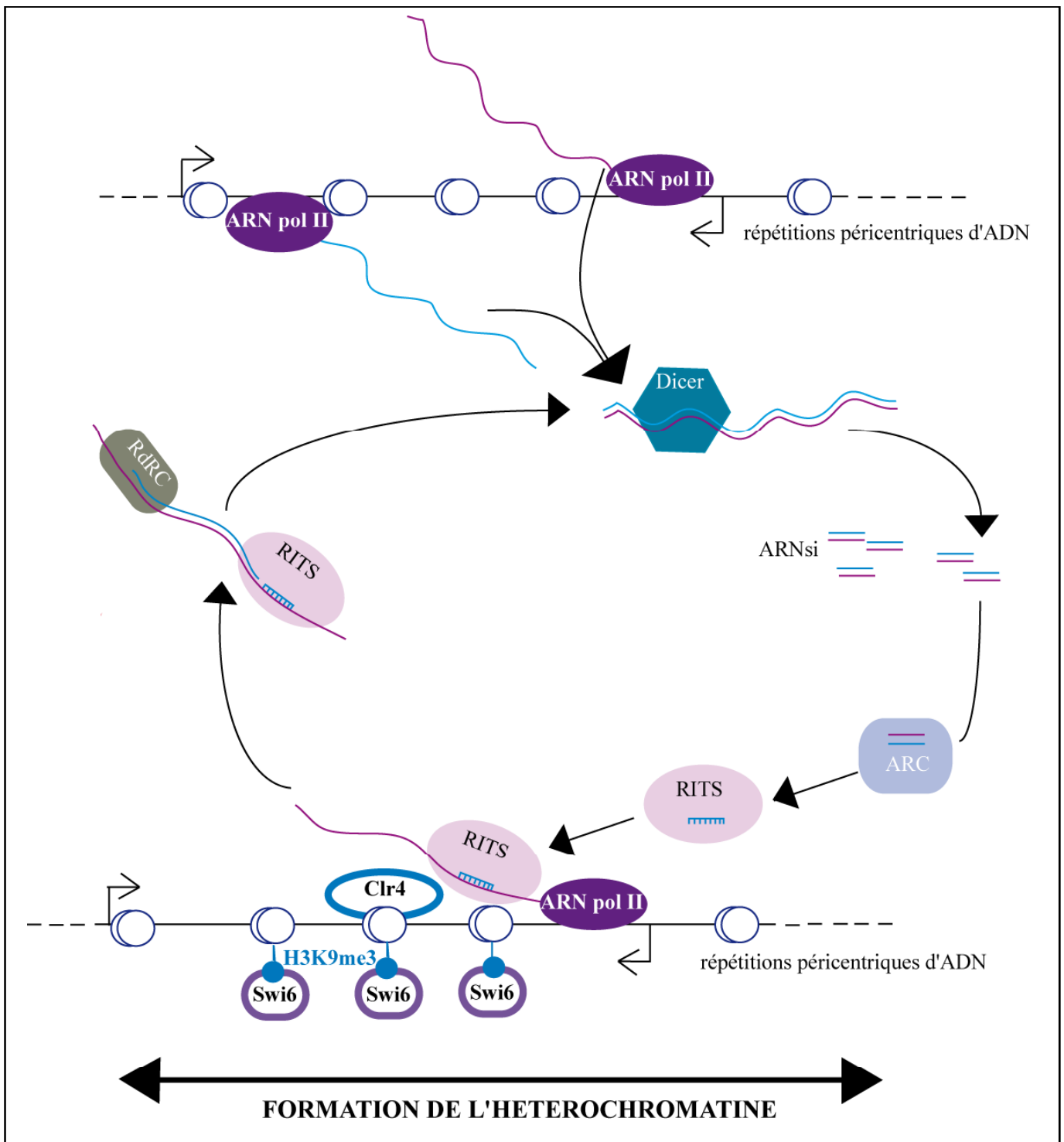


Figure 22 : Mécanisme d'ARN interférence à l'origine de l'hétérochromatinisation des régions péricentriques chez la levure *S. pombe*. La transcription des séquences péricentriques sens et antisens permet la formation d'un ARN double brin qui sera clivé en petits ARN interférents (ARNsi) par Dicer. Les petits ARN interférents, double brin, sont reconnus par le complexe ARC (*AR*gonaute *CH*aperone), et un seul brin est chargé dans le complexe RITS. Par complémentarité des bases, l'ARN simple brin permet le ciblage du RITS au lieu de transcription des régions péricentriques. Le complexe RITS recrute alors Clr4 responsable de la triméthylation de H3K9. Le complexe RITS peut également s'associer avec un complexe comprenant une ARN polymérase ARN dépendante (RdRC) pour augmenter la synthèse d'ARN double-brin et donc propager l'hétérochromatinisation.

la présence sur l'ARN d'une structure en tige-boucle comparable à celle de RepA (Kaneko et al., 2010). L'ARNnc HOTAIR peut également lier la protéine LSD1 responsable de la déméthylation de H3K4, assurant ainsi la répression de ses gènes cibles de deux manières différentes (Tsai et al., 2010). La surexpression de HOTAIR dans des cas de cancer du sein chez l'homme induit une mauvaise localisation du complexe PRC2 à plus de 850 gènes (Gupta et al., 2010) mais les mécanismes de ciblage des séquences régulées par HOTAIR ne sont pas encore connus (figure 21b).

La transcription de longs ARNnc semble donc réguler la condensation de la chromatine, et donc la transcription, *en cis* ou *en trans*, à plusieurs endroits du génome.

3. **Hétérochromatine péricentrique et ARN interférent**

a. Chez la levure

Chez *S. pombe* où les régions péricentriques et centromériques ont des caractéristiques chromatiniennes comparables à celles des mammifères, il a été montré que la formation de l'hétérochromatine péricentrique et plus particulièrement la méthylation de H3K9 par Clr4, l'homologue de SUV39H, est nécessaire au recrutement de CENP-A et à la formation des kinétochores au niveau centromérique (Folco et al., 2008). De plus, il a été montré que Clr4 est ciblée aux régions centromériques grâce au processus d'ARN interférence. La transcription bidirectionnelle des séquences péricentriques répétées, pendant la phase S du cycle cellulaire, entraîne la formation d'ARN double brin (ARNdb) (Chen et al., 2008) (Volpe et al., 2002). Cet ARNdb est ensuite reconnu par la protéine Dicer qui le clive en de petits ARN interférents (ARNi) de 21nt qui s'associent avec la protéine argonaute (Ago1) au sein du complexe RITS (*RNA induced transcriptional silencing*). Grâce à la complémentarité de séquence entre les petits ARNi et les régions péricentriques, le complexe RITS est ciblé à la chromatine péricentrique où il induit la méthylation de H3K9 et l'hétérochromatinisation grâce au recrutement de Clr4 et Swi6, les homologues de SUV39H et HP1 respectivement. Il existe, de plus, chez la levure des ARN polymérase ARN dépendantes (Rdrp) permettant de synthétiser le brin complémentaire d'un ARN pour générer encore plus d'ARNdb et ainsi amplifier le processus de répression de la transcription et de condensation de la chromatine (Verdel et al., 2004) (Zhang et al., 2008) (Moazed, 2009) (figure 22). En résumé, chez la levure, la transcription des séquences péricentriques répétées et le recrutement de la

machinerie d'ARN interférence induisent l'établissement de l'hétérochromatine et le recrutement de CENP-A et des protéines du kinétochore aux centromères pour permettre le maintien de l'intégrité du génome au cours de la mitose.

b. Chez les mammifères

Chez les mammifères, des transcrits correspondant aux *major satellites* ont été identifiés dans les deux sens de transcription (Lehnertz et al., 2003) (Rudert et al., 1995) (Lu and Gilbert, 2008) et des traitements à la RNase détruisent l'organisation chromatinienne des centromères ainsi que le recrutement de HP1 α (Maison et al., 2002). D'après ces données, la possibilité d'une implication d'une machinerie d'ARN interférence dans la formation de l'hétérochromatine péricentrique chez les mammifères est donc envisageable. Cependant, l'absence de protéine RDRP chez les mammifères ne permet pas de faire un parallèle complet avec les mécanismes de répression de la chromatine péricentrique mis en œuvre chez la levure. De plus, la déplétion de DICER n'induit qu'une faible augmentation des transcrits correspondant aux *major satellites* et n'a quasiment aucun effet sur la structure de l'hétérochromatine (Kanellopoulou et al., 2005) (Murchison et al., 2005). Un lien plus direct entre l'hétérochromatinisation et la transcription a néanmoins été suggéré par le fait que les protéines HP1 α et HP1 β peuvent lier directement une molécule d'ARN grâce à son domaine *Hinge* (Muchardt et al., 2002). De plus, une spécificité de liaison entre les transcrits sens des *major satellites* et le domaine *Hinge* sumoylé de HP1 α a été démontrée très récemment, suggérant très fortement un rôle de l'ARN des *major satellites* dans le ciblage spécifique de HP1 au niveau de l'hétérochromatine péricentrique (Maison et al., 2011).

V. Héritabilité des marques des histones et épigénétique

1. Qu'est ce que l'épigénétique ?

Les modifications de la chromatine, par modification post-traductionnelle des histones, incorporation de variants d'histones ou méthylation de l'ADN, peuvent être divisées en deux catégories selon leur dynamique. (1) Les modifications ponctuelles dans le temps, mises en place et supprimées au cours d'un même cycle cellulaire, et n'influençant pas la génération suivante. C'est le cas par exemple des modifications permettant la condensation de

la chromatine au cours de la mitose. (2) Les modifications héritées d'une cellule mère à une cellule fille, nécessitant des mécanismes de transmission de l'information pendant plusieurs cycles cellulaires. Ces modifications héréditaires sont caractérisées de modifications épigénétiques.

La définition de l'épigénétique a beaucoup évolué dans le temps. Dans sa définition proche de l'actuelle, elle a été utilisée tout d'abord en 1942 par Waddington pour parler d'une branche de la biologie étudiant les mécanismes par lesquels les gènes et leurs produits sont associés à l'environnement pour définir un phénotype. Puis, en 1987, Holliday a fait évoluer cette définition pour parler de situations dans lesquelles des changements de méthylation de l'ADN induisent des modifications dans l'expression des gènes. Aujourd'hui, l'épigénétique est définie comme l'étude des modifications stables et héréditaires de l'expression du génome n'impliquant pas de modifications dans la séquence d'ADN. Pour ainsi modifier l'expression des gènes, plusieurs mécanismes sont possibles : la méthylation de l'ADN, l'incorporation de variants d'histones et de modifications post-traductionnelles des histones. D'autres facteurs tels que les ARN non codants et l'architecture nucléaire sont parfois aussi considérés comme des mécanismes épigénétiques. Pour qu'une modification soit caractérisée d'épigénétique, trois critères principaux doivent être respectés : elle doit pouvoir se propager, il doit exister des évidences de sa transmission entre les générations et elle doit avoir un effet sur l'expression des gènes.

Bien que le terme de modification épigénétique soit employé aujourd'hui avec plus de souplesse caractérisant parfois toute modification de la chromatine, nous nous efforcerons de l'utiliser seulement selon sa définition précise.

2. Héritabilité des marques des histones

Pour que les modifications d'histones soient considérées comme des marques épigénétiques, il faut qu'elles soient transmissibles de la cellule mère aux cellules filles à travers la réplication de l'ADN et la mitose. Lors du passage de la fourche de réplication, les histones sont dissociées de l'ADN et, d'après des expériences de réplication *in vitro*, la chromatine se reforme peu après avec à la fois à des histones parentaux et des histones nouvellement synthétisés. Que les histones parentaux soient redéposés de manière aléatoire sur la chromatine (Jackson and Chalkley, 1985), ou qu'elles le soient selon un modèle semi-conservatif impliquant une distribution de la moitié des nucléosomes parentaux sur chaque double brin d'ADN (Tagami et al., 2004), le message porté par la chromatine est dilué après

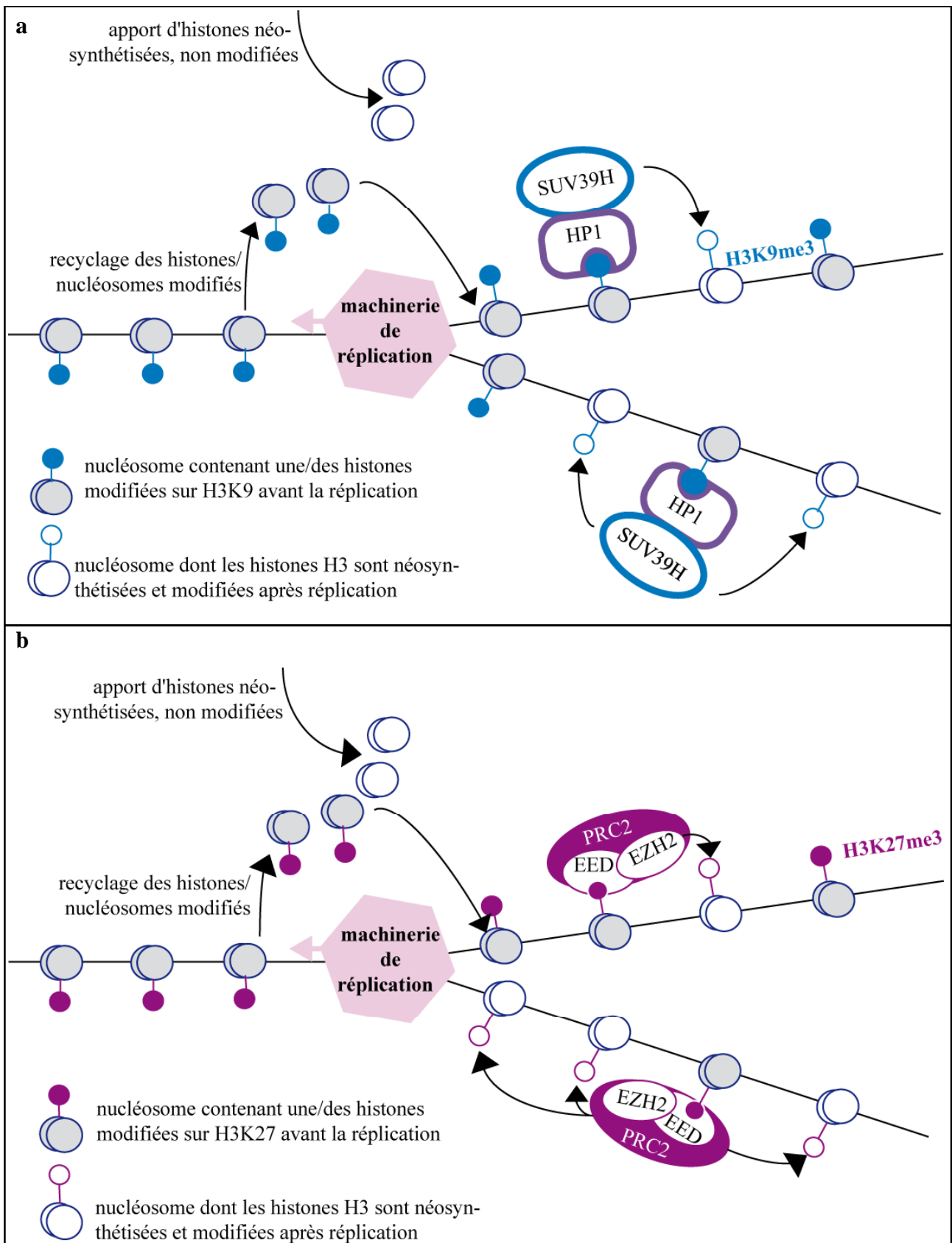


Figure 23 : Rétrocontrôles positifs permettant la propagation de H3K9me3 (a) et H3K27me3 (b) après réplication. La question de la conservation des tétramères H3/H4 ou du passage d'un tétramère à deux dimères H3/H4 lors de la réplication n'est pas abordée ici.

la réplication et doit être remis en place. Les histones nouvellement synthétisées sont modifiées sur certains résidus avant leur incorporation dans la chromatine : l'acétylation de H4K5 et H4K12 est par exemple nécessaire à leur déposition (Sobel et al., 1995), et chez l'homme une petite fraction des histones H3 cytoplasmiques est monométhylée sur H3K9 (Loyola et al., 2006). Néanmoins, à part ces quelques modifications déjà présentes sur les histones nouvellement synthétisées, ces dernières devront acquérir la plupart des marques, en adéquation avec leur environnement d'insertion.

Bien que les mécanismes précis de duplication et donc de conservation des marques des histones à travers la réplication ne soient pas connus, plusieurs pistes ont été suggérées, nous en aborderons deux par la suite.

a. Propagation de marques existantes

On peut supposer que dans des régions enrichies pour certaines modifications, les histones parentales, déjà modifiées, peuvent permettre la propagation de leurs marques aux histones environnantes. Ceci nécessite une coopérativité entre les marques et les enzymes qui les catalysent. De tels mécanismes de rétrocontrôles positifs ont été mis en évidence pour la propagation de H3K9me3 et H3K27me3.

- Propagation de H3K9me3

Le chromodomaine de HP1 reconnaît H3K9me2 et H3K9me3 (Bannister et al., 2001) (Lachner et al., 2001). HP1 interagit également directement avec SUV39H1 et SUV39H2, responsables de la di- et triméthylation de H3K9, grâce à son domaine *chromoshadow* (Stewart et al., 2005). Ainsi, H3K9me2/3 peut se propager grâce à la reconnaissance de la marque par HP1, et au recrutement ultérieur, grâce à HP1, des méthyltransférases spécifiques de H3K9 (figure 23a).

- Propagation de H3K27me3

En ce qui concerne H3K27me3, EED, peut lier spécifiquement les marques répressives H1K26me3, H3K9me3, H4K20me3 et H3K27me3 grâce à une poche aromatique. Tout particulièrement, la liaison avec H3K27me3 induit une stimulation allostérique du complexe PRC2, et donc une augmentation de l'activité histone méthyltransférase sur H3K27 (Margueron et al., 2009). La présence de H3K27me3 au sein d'un nucléosome induit donc la triméthylation de H3K27 des nucléosomes adjacents (figure 23b).

b. Rôle des ARN

Les enzymes catalysant les modifications des histones peuvent donc d'une part être recrutées par les histones modifiées elles-mêmes, mais elles peuvent d'autre part être ciblées à la chromatine par facteurs agissant en *trans*. Comme nous l'avons vu dans la partie précédente, de nombreux ARNnc longs et courts, agissent sur la régulation de l'expression des gènes en recrutant plus ou moins directement des enzymes de modification de la chromatine. Ces ARN sont possiblement transmis d'une génération de cellule à l'autre et peuvent ainsi jouer un rôle quant à l'héritabilité des marques épigénétiques, en particulier pour la formation et le maintien des régions chromatiniennes condensées. Dans le cas où les ARN, agissant en *cis*, ne seraient pas transmis physiquement, l'ouverture de la chromatine et la dérèpression des gènes entraînerait la nouvelle synthèse d'ARNnc et donc à nouveau la condensation de ces régions.

Les modifications de la chromatine permettent donc de réguler le programme d'expression génique d'une cellule de manière réversible mais également de manière potentiellement transmissible de génération cellulaire en génération cellulaire. Au cours de ma thèse, j'ai abordé le phénomène de régulation de la transcription et d'organisation de la chromatine dans deux types cellulaires particuliers, les ovocytes et les zygotes. Ces cellules sont l'emblème, *in vivo*, du passage d'un état différencié à un état totipotent, et tout ce qui est démontré dans les cellules somatiques ne s'applique peut-être pas systématiquement à ces deux types cellulaires en pleine transition. Nous verrons, dans le dernier chapitre de cette introduction, comment est accommodée la transition entre cellules germinales et zygote d'un point de vue transcriptionnel et épigénétique.

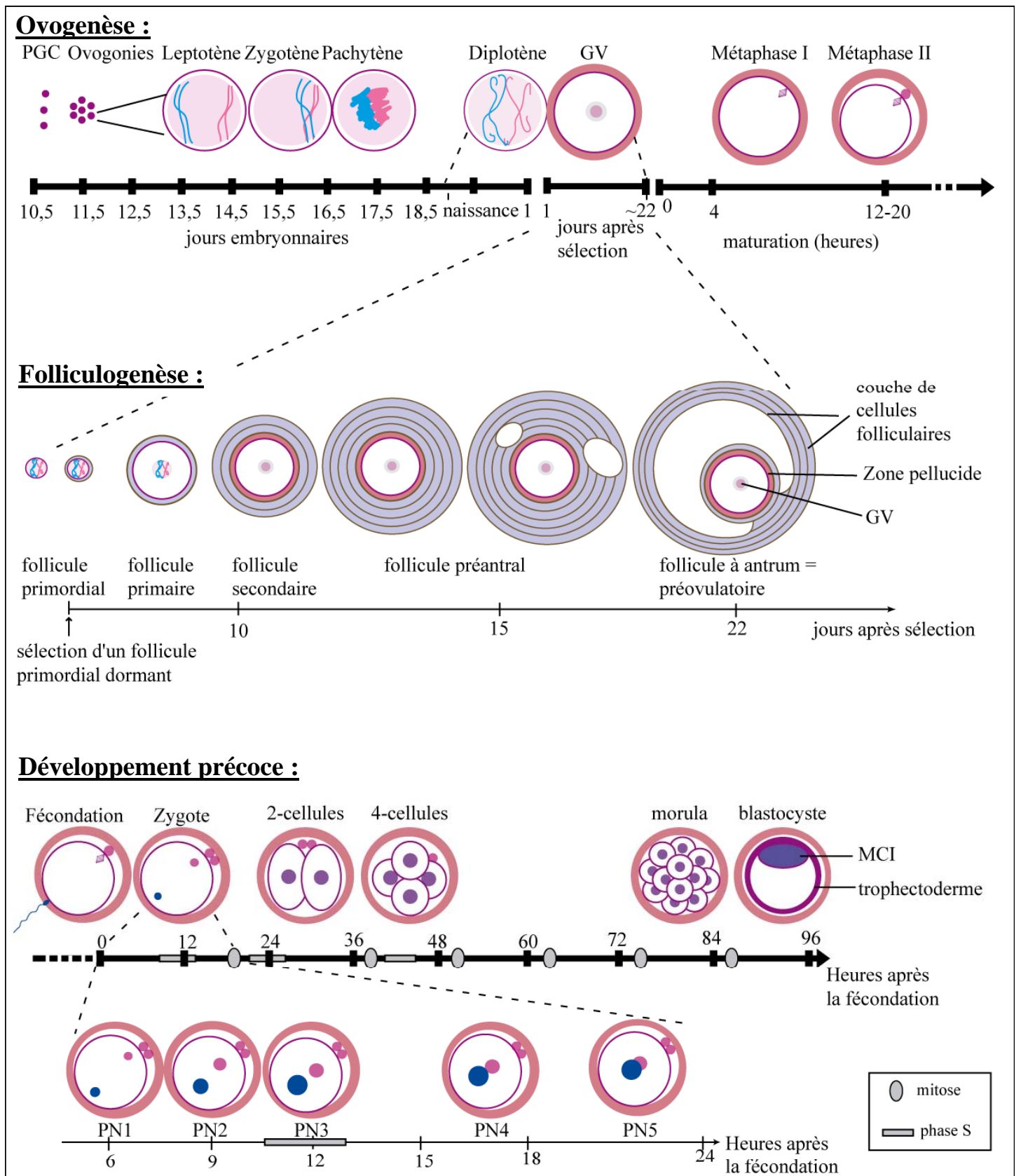


Figure 24 : Schéma récapitulatif des étapes de l'ovogenèse, de la folliculogenèse et du développement précoce de l'embryon. Pendant les premières phases de la prophase I de méiose, la position relative de deux chromosomes homologues est schématisée. En phase GV, le noyau, central, est montré en mauve. Les différents stades du stade zygote sont détaillés et la position relative des pronoyaux paternel (en bleu) et maternel (en rose) est schématisée. Le schéma est adapté de Martin et al., 2006, Adénot et al., 1997, et Racki et al., 2006.

C. Cellules germinales et zygotes, des cellules particulières

Ovocyte et spermatozoïde sont des cellules différenciées. Leur fusion lors de la fécondation induit la formation d'une cellule totipotente, le zygote. Le passage d'un état différencié à un état totipotent nécessite des changements importants, des reprogrammations, du programme d'expression génique et épigénétique de ces cellules. Après avoir exposé les étapes de formation des cellules germinales et le passage d'un ovocyte et d'un spermatozoïde haploïdes à un zygote puis un blastocyste diploïdes (figure 24), nous verrons en quoi ces cellules sont particulières d'un point de vue transcriptionnel et épigénétique.

I. Des cellules germinales au blastocyste

1. La gamétogenèse

a. L'ovogenèse

Chez la souris, l'ensemble des transformations qui vont mener à la formation d'un gamète femelle mature, l'ovogenèse, peut être divisée en deux phases. La première phase commence pendant la vie fœtale, les ovocytes de souris subissent le début de la prophase I de la méiose et s'arrêtent au stade diplotène après la naissance. La deuxième phase correspond à une sélection continue, pendant la période reproductive de la souris, d'ovocytes pour entrer en phase de croissance et potentiellement atteindre l'ovulation au bout d'une vingtaine de jours (Gilbert, 2010).

- **Phase prénatale de l'ovogenèse :**

Au cours du développement, les cellules germinales primordiales (PGC : *primordial germ cell*), des cellules unipotentes différenciées à partir de cellules de l'épiblaste proximal, migrent dans les crêtes génitales entre E7,5 et E11,5. Après reprogrammation, les cellules germinales subissent de nombreuses mitoses et forment de nombreuses ovogonies. A environ E13, les ovogonies commencent à se différencier en ovocytes et subissent les premières étapes de la prophase I de méiose. Au cours du stade leptotène, les chromosomes se condensent et les télomères se lient à la membrane nucléaire. Puis, les chromosomes homologues s'alignent côte à côte lors du stade zygotène et s'apparient grâce à l'action du

complexe synaptonémal formant une tétrade (= 4 chromatides) ou un bivalent (= 2 chromosomes). Lorsque les chromosomes sont entièrement appariés, les chromatides s'entremêlent et certains segments sont échangés entre les chromosomes homologues, entraînant un brassage intra-chromosomique du matériel génétique au cours du stade pachytène. Les échanges se font par un mécanisme de recombinaison homologue, ils font intervenir des cassures double brin dans l'ADN, cassures marquées par la présence du variant de H2A, H2AX, phosphorylé, reconnues et réparées par la machinerie de réparation de l'ADN classique de la cellule (pour une revue, Richardson et al., 2004). Enfin, au stade diplotène, les chromosomes homologues se séparent partiellement, leurs points de contact étant les sites où la recombinaison homologue a eu lieu. Les ovocytes arrêtent alors leur développement au stade diplotène jusqu'à la maturité sexuelle des souris.

- Follicules et folliculogénèse

La suite du développement de l'ovocyte s'effectue au sein d'un follicule, en coordination avec le développement de celui-ci. Les follicules ovariens sont composés d'un ovocyte, de cellules de la granulosa ou cellules folliculaires enveloppant l'ovocyte et sécrétant la progestérone et des cellules de la thèque sécrétrices d'oestrogènes. La thèque est richement vascularisée, mais reste séparée des cellules folliculaires par une lame basale déconnectant totalement l'ovocyte et la granulosa des vaisseaux sanguins. Une communication bidirectionnelle entre l'ovocyte et les cellules de la granulosa est nécessaire à la prolifération et à la différenciation des cellules folliculaires d'une part et à la production d'un gamète femelle fonctionnel d'autre part.

Dans les ovaires de nouveau-nés, les ovocytes sont au stade diplotène de prophase I, ils se sont entourés de quelques cellules somatiques, qui deviendront les cellules de la granulosa, formant ainsi des follicules primordiaux. Le stock de follicules primordiaux est généré pendant la phase prénatale de l'ovogénèse et aucun nouveau follicule primordial n'est formé au cours de la vie de la souris. A ce stade, les follicules primordiaux sont dormants et certains seront sélectionnés de manière non synchronisée, à tout moment de la vie reproductive de la souris, pour subir la folliculogénèse. La grande majorité des follicules primordiaux n'entre pas en phase de croissance et dégénère spontanément.

Au sein des follicules primordiaux sélectionnés, les cellules de la granulosa prolifèrent, prennent une forme cuboïdale et entourent l'ovocyte d'une couche unique de cellules. Les follicules sont alors au stade de follicule primaire. L'ovocyte du follicule primaire croît jusqu'à augmenter 100 fois son volume initial, devenant alors l'une des plus grosses cellules

du corps. Au stade follicule secondaire, les cellules de la granulosa continuent de proliférer pour entourer l'ovocyte en croissance de deux couches de cellules. Les cellules folliculaires prolifèrent alors de plus en plus rapidement pour former jusqu'à 6 ou 7 couches de cellules autour de l'ovocyte ayant atteint sa taille maximale, formant un follicule préantral. Une cavité remplie de liquide se forme ensuite au sein du follicule, la présence d'une telle cavité caractérise alors un follicule à *antrum*. A partir de ce stade, alors que l'ovocyte est mature, son noyau, au stade diplotène, porte le nom de vésicule germinale (GV : *germinal vesicle*), on parle alors d'ovocyte au stade GV.

Toutes ces premières étapes de la folliculogénèse sont indépendantes de l'action des hormones hypophysaires, FSH (*follicle stimulating hormone*) et LH (*luteinizing hormone*). Le follicule à *antrum* précoce est alors sensible à une hausse de la sécrétion de FSH par l'hypophyse. Lorsqu'il est soumis à une quantité d'hormones appropriée, la cavité s'étend fortement autour de l'ovocyte dans le follicule devenu préovulatoire. Le pic de LH et de FSH induit alors l'ovulation et la reprise de la méiose par l'ovocyte qui entre en métaphase I après la dissociation de la membrane nucléaire (phase GVBD : *germinal vesicle breakdown*), puis en anaphase I et télophase I. La première division méiotique induit la formation d'un premier corps polaire possédant un chromosome de chaque paire, et ne subissant pas de deuxième division méiotique. L'ovocyte, possédant également un chromosome de chaque paire à deux chromatides, commence la deuxième division méiotique et s'arrête à nouveau en métaphase II jusqu'à la fécondation. Ces dernières étapes de reprise de la méiose caractérisent l'étape de maturation de l'ovocyte.

L'ovocyte finit sa deuxième division de méiose au moment de la fécondation et forme alors un deuxième corps polaire.

b. La spermatogénèse

Contrairement à l'ovogénèse, la spermatogénèse est un processus continu qui débute après la naissance des souris et produit des spermatozoïdes matures à partir de cellules germinales tout au long de la vie reproductive du mâle. Après leur migration dans les crêtes génitales pendant la vie embryonnaire, les PGC arrêtent leurs mitoses jusqu'à la maturité sexuelle des mâles où elles donneront des spermatogonies. Les spermatogonies subissent des mitoses successives en périphérie des tubules séminifères, puis se différencient en spermatocytes subissant la méiose. Les étapes de la méiose sont les mêmes que pour l'ovogénèse à ceci près qu'elles se déroulent sans arrêt majeur. Au stade pachytène de la

prophase I, les spermatocytes subissent une recombinaison homologue comparable à celle des ovocytes. La seconde division méiotique des spermatocytes résulte en la production de spermatides haploïdes. Les spermatides entrent alors dans une phase de maturation post-méiotique, la spermiogenèse. Pendant cette phase, les spermatides subissent de nombreuses transformations morphogéniques comme la compaction de leur ADN que nous reverrons par la suite, l'évacuation de la majorité de leur cytoplasme, un réarrangement des mitochondries et la formation d'un acrosome et d'un flagelle. Pendant toute la spermatogenèse, les spermatocytes migrent vers la lumière des tubules séminifères puis dans l'épididyme pour la fin de la maturation. Après toutes ces modifications, les spermatozoïdes sont prêts à être éjaculés. Ils subiront une dernière étape de maturation les rendant aptes à la fécondation dans le tractus génital femelle, la capacitation.

Les différentes étapes de la spermatogenèse induisant la formation d'un spermatozoïde hautement différencié à partir de cellules souches diploïdes unipotentes se font grâce à un dialogue paracrine entre les cellules de Sertoli, cellules des tubes séminifères, et les cellules germinales sous la dépendance de LH et de FSH sécrétées par l'hypophyse (Gilbert, 2010).

2. Fécondation et stade zygote

Le développement de l'embryon commence par la fécondation de l'ovocyte en métaphase II par le spermatozoïde mature. La fusion des deux gamètes induit la formation d'un embryon au stade 1-cellule, un zygote. Le zygote contient deux groupes de chromatine, la chromatine maternelle et la chromatine paternelle, coexistant sous forme d'entités nucléaires séparées, les pronoyaux, pendant la totalité du premier cycle cellulaire (Mayer et al., 2000b). Le pronoyau paternel se forme un peu plus tardivement que le pronoyau maternel. Après la fécondation, les pronoyaux sont petits et se situent de part et d'autre de la cellule. Au cours des 12 premières heures de développement, les pronoyaux migrent vers le centre de la cellule, et leur volume augmente, le plus gros étant le pronoyau mâle. Une classification des stades de développement du zygote, de PN1 à PN5 (PN : *Pronuclear stage*), s'appuyant sur la taille et la position relative des pronoyaux a été proposée (Adenot et al., 1997). Nous suivrons cette nomenclature dans la suite de ce manuscrit (figure 24). L'évolution coordonnée des pronoyaux mâles et femelles se termine par la condensation synchrone des chromosomes pour former une plaque métaphasique unique avant la première division mitotique. Bien que les génomes parentaux se situent sein du même noyau à partir du stade 2-cellules, ils resteront

séparés et occuperont des zones différentes dans les noyaux jusqu'au stade 4-cellules (Mayer et al., 2000b).

La fécondation permet le passage de deux cellules haploïdes hautement différenciées, l'ovocyte et le spermatozoïde, à une cellule diploïde totipotente. La formation du zygote constitue le paroxysme de la totipotence à cause de son habilité à produire tous les types cellulaires d'un organisme ainsi que tous les tissus extra-embryonnaires.

3. Du stade 2-cellules au stade blastocyste

Les divisions successives de l'embryon s'enchaînent sans que celui-ci ne change significativement de volume, induisant donc une diminution de la taille du cytoplasme des cellules jusqu'au stade blastocyste. A la fin du stade morula, stade 16-cellules, une cavité liquidienne se creuse dans l'embryon alors appelé blastocyste. A ce stade apparaissent les premiers signes morphologiques évidents de différenciation cellulaire. En effet, le blastocyste contient deux types cellulaires différents, les cellules du trophoctoderme, externes, à l'origine des tissus extra-embryonnaires, et les cellules de la masse cellulaire interne (MCI) à l'origine des différents tissus qui composent l'embryon et d'un autre tissu extra-embryonnaire, l'endoderme primitif.

Le stade à partir duquel les blastomères de l'embryon commencent à être différents les uns des autres n'est pas encore précisément établi. Au stade 2-cellules, après une séparation mécanique des deux blastomères et leur réimplantation dans des femelles pseudo-gestantes, chacune des deux cellules permet le développement d'une souris viable (Tarkowski, 1959). A des stades plus tardifs, la dérivation de jumeaux à partir de blastomères isolés n'est plus possible (Tarkowski and Wroblewska, 1967) (Rossant, 1976). Néanmoins, quand un blastomère issu d'un embryon aux stades 4-cellules ou 8-cellules est ajouté à des blastomères « receveurs », permettant la présence d'un nombre de cellules assez important, les cellules dérivées de ce premier blastomère peuvent contribuer à la formation de tous les tissus embryonnaires (Kelly, 1977) (Tarkowski et al., 2005). Jusqu'au stade 8-cellules, les blastomères sont donc considérés comme totipotents, mais ils ne semblent pas pour autant être complètement identiques (Fujimori et al., 2003) (Piotrowska-Nitsche et al., 2005). A la fin du stade 8-cellules, une polarité apparaît au sein de chaque blastomère, chaque cellule présentant un pôle apical, vers l'extérieur de l'embryon, et un pôle baso-latéral vers le centre de l'embryon. Cette polarité est visible au niveau de l'organisation du cytosquelette, du cytoplasme, et de la membrane. Lors de la division des embryons entre les stades 8-cellules et

	ARN totaux (ng)	ARN polyA (pg)	ARNr (ng)
ovocyte phase GV	ND	0,95 ± 0,18	ND
ovocyte non fécondé	0,35 ± 0,03	0,70 ± 0,07	ND
Zygote	ND	0,83 ± 0,07	0,22
2-cellules	0,24 ± 0,04 (0,12/cellule)	0,26 ± 0,01 (0,13/cellule)	0,17 (0,085/cellule)
8/16-cellules	0,69 ± 0,07 (0,09-0,04/cellule)	0,44 ± 0,04 (0,055-0,028/cellule)	0,4 (0,05-0,025/cellule)

Tableau 1 : Quantités d'ARN totaux, ARN polyadénylés et ARN ribosomaux (ARNr) à différents stades du développement. Les quantités sont exprimées en moyenne ± erreur standard. Les moyennes sont faites sur au minimum 3 expériences. Les ovocytes et embryons sont disséqués à partir de femelles superovulées et leur ARN est extrait au phénol. La quantité totale d'ARN est déterminée par spectrophotométrie. La quantité d'ARN polyadénylé (polyA) est déterminée par hybridation avec une sonde poly(U) triciée. La quantité d'ARN ribosomal (ARNr) est estimée à partir du comptage du nombre de ribosomes présents sur une section d'embryon observée en microscopie électronique. Les données sont extraites de Piko et al., 1982.

16-cellules deux populations de cellules distinctes vont alors apparaître, une population interne, n'ayant aucun contact avec l'extérieur de l'embryon, et une population externe. Bien que ces cellules internes et externes soient encore capables de changer leur devenir dans le cas d'une transplantation à un autre endroit de l'embryon par exemple, elles constituent une première étape à la différenciation des cellules embryonnaires (pour une revue, Johnson and McConnell, 2004). Les cellules internes se divisent et différencient en cellules de la masse cellulaire interne (MCI), alors que les cellules externes se différencient en cellules du trophoctoderme au stade blastocyste. La perte de la totipotence des cellules de l'embryon devient complète à la fin du stade blastocyste où les cellules de la MCI n'ont plus le potentiel de s'intégrer à des dérivés du trophoctoderme après transplantation (Rossant and Lis, 1979)

II. Transcription dans les cellules germinales et pendant le développement précoce

1. Des taux de transcription variables

Lorsqu'on observe les quantités d'ARN totaux et ARN polyadénylés avant et après la fécondation, on remarque que ce taux n'évolue pas significativement entre l'ovocyte non fertilisé et le zygote alors qu'il diminue drastiquement aux stades 2- et 8-cellules (tableau 1). Bien que la quantité de transcrits ne soit pas le reflet direct de l'activité transcriptionnelle d'une cellule, ces données illustrent le fait qu'il existe une transcription très importante dans l'ovocyte, que les ARN formés sont transmis au zygote puis sont dégradés pendant les stades suivants (Piko and Clegg, 1982).

a. Dans les gamètes

Le taux de transcription dans les ovocytes est faible, voire nul, jusqu'au début du stade pachytène. Il augmente ensuite modérément pendant le stade pachytène, puis beaucoup plus fortement pendant la phase de croissance de l'ovocyte où la synthèse d'ARN est très rapide (Moore et al., 1974).

Au cours de la folliculogénèse, deux populations d'ovocytes sont détectables selon leur organisation chromatinienne et leur niveau de transcription. Une population précoce est caractérisée par l'absence de chromatine condensée autour de la structure en anneau du

nucléole, on parle de type « non-entouré » (NSN : *non-surrounding nucleolus*). Les ovocytes de type NSN sont transcriptionnellement actifs. A l'inverse, une population plus tardive, transcriptionnellement silencieuse est définie par la présence d'un anneau de chromatine périnucléolaire très condensé, on parle de type « entouré » (SN : *surrounding nucleolus*). Le niveau global de transcription est alors quasiment nul lors de la rupture de la vésicule germinale à la fin de la prophase I (Bouniol-Baly et al., 1999) (Moore et al., 1974). Les transcriptions par les ARN pol I et ARN pol II sont concernées par cette répression.

Il a été montré qu'une grande partie des ARNm produits lors de la phase de croissance est accumulée, stockée dans l'ovocyte jusqu'à la fécondation (Jahn et al., 1976). Grâce à des régions 3'UTR spécifiques et à la régulation de la taille de la queue polyA en 3', ces ARNm, d'une très grande stabilité ne seront traduits qu'après la fécondation, dans le zygote (Vassalli et al., 1989) (pour une revue, Bettogowda and Smith, 2007). Ils contiennent les messages pour la synthèse des protéines de ménage, des protéines impliquées dans la différenciation embryonnaire précoce et une grande quantité de transcrits codant des séquences répétées. La régulation post-transcriptionnelle des ARNm maternels pour qu'ils soient conservés et transcrits seulement après la fécondation est nécessaire au bon déroulement de l'embryogenèse précoce. Après la fécondation, 30 à 40% des ARN maternels sont dégradés avant le stade 2-cellules (Bachvarova, 1985) (Kaplan et al., 1982).

Durant la spermatogenèse, le pic de transcription coïncide avec la deuxième division méiotique. La transcription diminue ensuite pendant la spermiogenèse jusqu'à être presque totalement éteinte dans les spermatozoïdes matures (Grunewald et al., 2005) (MacLaughlin and Turner, 1973). De manière surprenante pour une cellule transcriptionnellement silencieuse et surtout pour une cellule dont la quasi-totalité du cytoplasme a été supprimée, on sait aujourd'hui que les spermatozoïdes matures contiennent de nombreux ARNm qu'ils peuvent traduire *de novo* (Gur and Breitbart, 2006) ou transmettre au zygote (Ostermeier et al., 2004). Le rôle des ARN paternels au cours du développement préimplantatoire n'est pas encore élucidé.

b. Dans le zygote

Dans la plupart des espèces, le développement embryonnaire débute par une période transcriptionnellement silencieuse contrôlée exclusivement par des transcrits provenant du cytoplasme de l'ovocyte donc d'un héritage maternel (la contribution des ARN paternels au

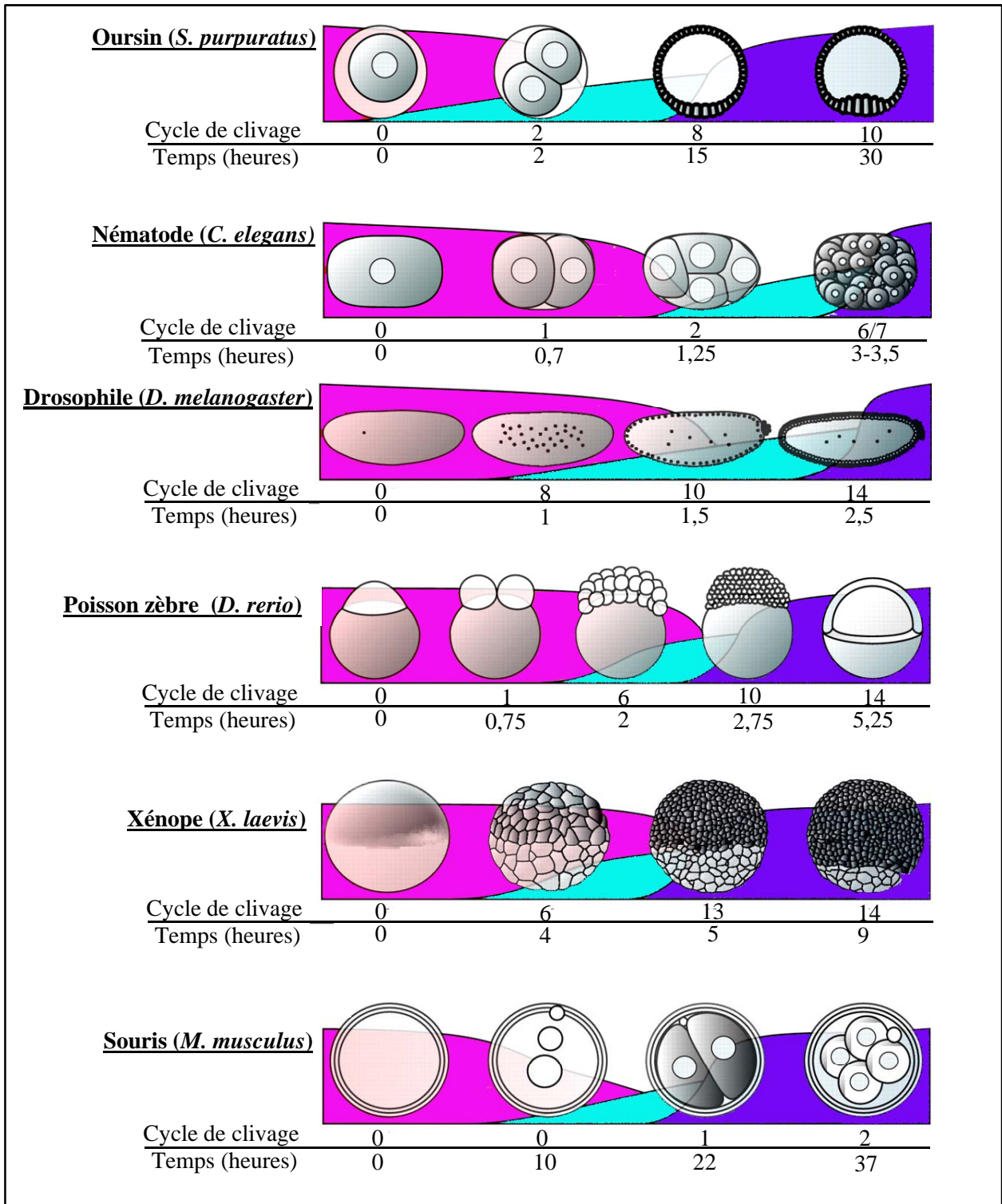


Figure 25 : Comparaison des périodes de transition entre un contrôle maternel et embryonnaire du développement (MZT) chez différents organismes modèles. Les stades embryonnaires clés sont schématisés au dessus du cycle de clivage et du temps d'apparition après la fécondation correspondants. Les courbes roses représentent le profil de dégradation des transcrits maternels. Les courbes vertes et bleu foncé illustrent les vagues mineures et majeures d'activation du génome embryonnaire respectivement. Le dernier stade embryonnaire présenté montre le premier moment du développement qui requiert majoritairement des transcrits embryonnaires. La figure est adaptée d'après Tadros et al., 2009.

développement est soumis à débat, (pour une revue, Lalancette et al., 2008). La période de transition entre un contrôle maternel et un contrôle embryonnaire du développement varie énormément entre les espèces (figure 25).

Chez la souris, les premières expériences de sensibilité à l' α -amanitine, un inhibiteur de la transcription, ont montré que la transcription à partir du génome de l'embryon au stade 2-cellules est nécessaire à son développement. Il a alors été supposé que l'activation du génome de l'embryon avait lieu au stade 2-cellules (Flach et al., 1982). Cependant, des expériences d'incorporation de BrUTP dans des embryons ont mis en évidence une première activation du génome zygotique (*ZGA : zygotic genome activation*) au stade 1-cellule (Bouniol et al., 1995). La transcription commence donc dans l'embryon de souris, à un niveau faible, pendant la première phase S au stade zygote, on parle de *ZGA* mineure. A ce stade, la transcription dans le pronoyau mâle est quatre à cinq fois plus importante que dans le pronoyau femelle. Elle s'intensifie ensuite pendant le stade 2-cellules (Aoki et al., 1997) lors de la *ZGA* majeure. La *ZGA* est une des deux étapes nécessaires à la transition entre un contrôle maternel et un contrôle embryonnaire du développement (*MZT : maternal to zygotic transition*). Couplée à la transcription du génome embryonnaire, une dégradation des ARNm maternels a lieu dans l'embryon. Ces étapes de la *MZT* sont finement régulées. La dégradation des ARN maternels est d'abord régulée par des facteurs maternels puis par des protéines et des petits ARN exprimés à partir des gènes embryonnaires (Giraldez et al., 2006). De plus, au sein des premières protéines synthétisées à partir du génome de l'embryon, des activateurs vont permettre une augmentation de la transcription embryonnaire. Le résultat de ces régulations est le transfert du contrôle du développement du génome maternel au génome embryonnaire entre la fécondation et le stade 4-cellules.

En plus de taux de transcription spécifiques, les cellules germinales et les embryons vont posséder des patrons d'expression des gènes très différents. La mise en place du programme d'expression spécifique de l'ovocyte et des spermatozoïdes met en jeu différents mécanismes de régulation de la transcription dont nous allons détailler l'exemple de l'utilisation différentielle d'homologues de TBP.

2. Utilisation de facteurs de transcription de base spécifiques : exemple des homologues de TBP

Pour assurer leur fonction biologique, les ovocytes et les spermatozoïdes doivent établir des programmes d'expression génique spécialisés. Les ARN produits sont, en

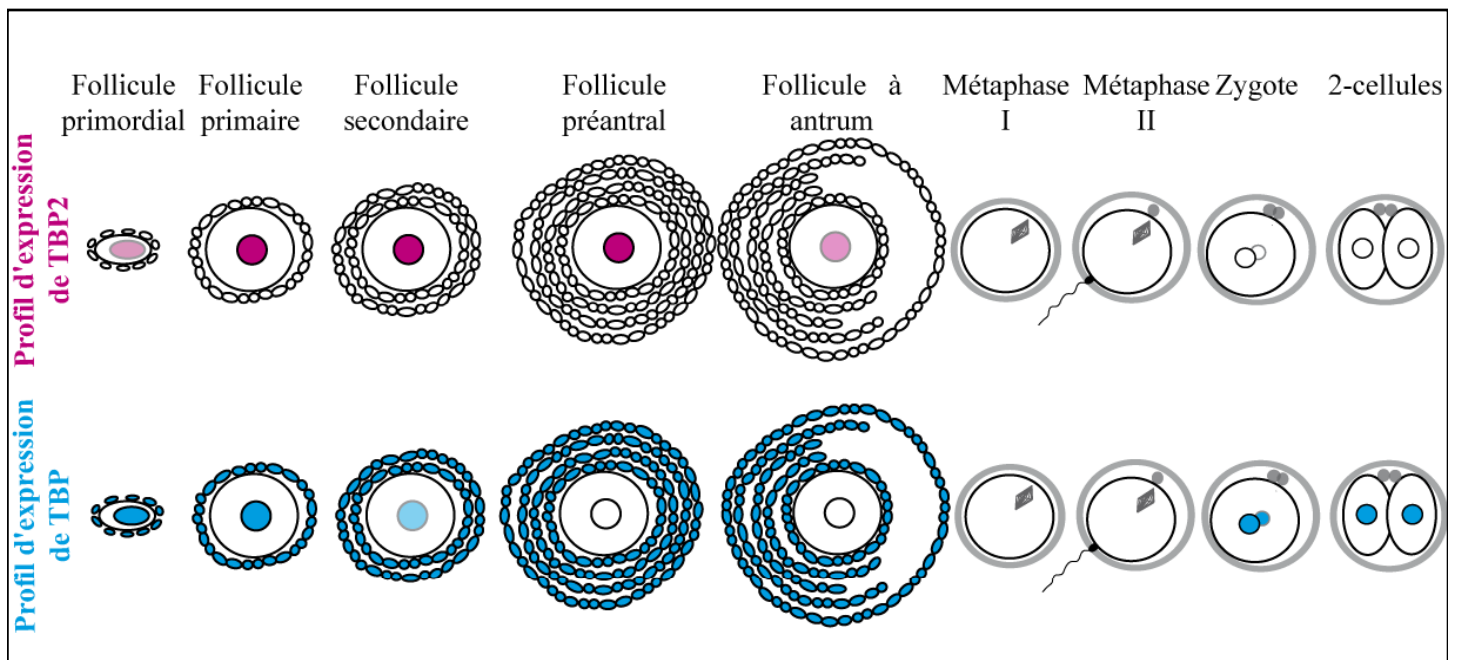


Figure 26 : Schéma récapitulatif des profils d'expression de TBP2 (en rose, en haut) et TBP (en bleu, en bas) dans les ovocytes et les cellules folliculaires au cours de la folliculogénèse et dans les embryons précoces. Les follicules sont schématisés en accord avec leur stade de développement noté en haut. L'intensité de la couleur dans les noyaux correspond à la quantité relative de TBP ou TBP2 par rapport à une quantité maximale. Le schéma a été réalisé d'après les données de Gazdag et al., 2007.

beaucoup de points, différents de ceux produits dans les cellules somatiques concernant leur taux de production, les isoformes produites, l'utilisation de TSS alternatifs. Ces observations suggèrent que les mécanismes moléculaires de l'initiation transcriptionnelle doivent être spécifiques aux cellules germinales. Des études ont montré que les cellules germinales utilisent des formes alternatives de facteurs de transcription de base et que la transcription est initiée à partir de promoteurs spécifiques, silencieux dans les cellules somatiques (pour une revue, DeJong, 2006). Nous nous intéresserons ici aux homologues de TBP mis en jeu lors de la gamétogenèse.

TBP est exprimée dans les ovocytes de souris au début de la folliculogenèse, au stade follicule primordial, puis sa quantité diminue au stade follicule primaire et la protéine devient indétectable à partir du stade préantral. TBP est à nouveau présente dans le noyau après la fécondation. En opposition, la protéine TBP2, homologue de TBP, est fortement exprimée dans les ovocytes du stade follicule primordial à l'ovulation. Elle diminue ensuite, après l'ovulation, pour devenir indétectable après la fécondation (Gazdag et al., 2007) (figure 26). Ce profil de localisation en miroir de TBP et TBP2 suggère que ces deux protéines pourraient jouer des rôles différents dans l'établissement de programmes d'expression génique spécialisés pendant l'ovogenèse et le début du développement.

Chez le xénope, ce profil d'expression en miroir de TBP et TBP2 est aussi présent dans les ovocytes. Dans cet organisme, il a de plus été montré que TBP2 est recruté aux promoteurs de certains gènes dans l'ovocyte et que TBP était capable de remplacer fonctionnellement TBP2 au promoteur de ces gènes (Akhtar and Veenstra, 2009). Cette dernière observation soulève la question de l'apport spécifique de TBP2 dans l'ovocyte en opposition à TBP.

Une autre protéine homologue à TBP, la protéine TLF, semble être nécessaire à la formation des spermatozoïdes matures chez la souris. En effet, une inactivation de TLF entraîne la stérilité des souris mâles à cause d'un arrêt de la spermiogenèse. En l'absence de TLF, la majorité des spermatides entrent en apoptose. Bien que des données montrent une modification du profil d'expression de certains gènes en l'absence de TLF dans les spermatides, il n'existe aucune évidence d'une contribution directe de TLF à l'expression de ces gènes (Zhang et al., 2001). De plus, dans les spermatides à un stade précoce de la spermiogenèse, l'hétérochromatine centromérique de chaque chromosome est organisée en une structure unique appelée le chromocentre. Les spermatides TLF^{-/-} possèdent 3 ou 4

structures hétérochromatiniennes, suggérant un rôle pour TLF dans la structuration de l'hétérochromatine au sein du noyau lors de la spermiogenèse (Martianov et al., 2002a).

D'un point de vue évolutif, il est intéressant de constater que, alors que TLF et TBP2 sont nécessaires au développement normal de l'embryon de xénope et de poisson zèbre (Bartfai et al., 2004) (Muller et al., 2001) (Veenstra et al., 2000), ces deux homologues de TBP semblent n'être présents ou importants qu'au cours de la gamétogenèse chez la souris et non plus après la fécondation (Gazdag et al., 2007) (Martianov et al., 2001) (Zhang et al., 2001).

L'utilisation différentielle d'homologues de TBP au cours de la gamétogenèse pourrait être un moyen pour la cellule de mettre en place un programme d'expression génique spécifique, différent des cellules somatiques. Bien que cette hypothèse paraisse attrayante, la réalité et les mécanismes d'un tel phénomène n'ont pas encore été mis en évidence.

III. Des cellules particulières d'un point de vue épigénétique

Comme nous l'avons vu dans la partie précédente, le programme d'expression génique spécifique d'une cellule peut être contrôlé par des modifications de sa chromatine. Chaque cellule ayant sa propre signature chromatinienne en lien avec son type cellulaire et son transcriptome. Pour la plupart des types cellulaires du corps, ces marques chromatinienne restent fixes une fois la différenciation terminée ou le cycle cellulaire arrêté. Cependant, au cours du développement, entre les gamètes et le zygote, le passage de deux cellules différenciées à une cellule totipotente impose une reprogrammation épigénétique majeure comprenant le retrait global des marques épigénétiques existantes suivi par l'établissement de nouvelles marques. La première reprogrammation majeure intervient dans les PGC où les empreintes génétiques parentales sont effacées, puis une deuxième reprogrammation globale a lieu dans le zygote. Cette deuxième reprogrammation commence en fait, pour le génome paternel, dans les spermatogonies par le remplacement de la quasi-totalité des histones par des protamines. Ces deux vagues de reprogrammation majeures sont nécessaires à l'acquisition de la totipotence par le zygote.

1. Reprogrammation dans les cellules souches primordiales

Chez les mammifères placentaires et les marsupiaux, certains gènes sont exprimés de manière monoallélique selon leur origine parentale, on parle de gènes soumis à l’empreinte. Ainsi, certains gènes tels que *H19* ne sont exprimés qu’à partir de l’allèle maternel, tandis que d’autres gènes tels que *Igf2* ne sont exprimés qu’à partir de l’allèle paternel (Delaval and Feil, 2004). Le nombre de gènes soumis à l’empreinte est aujourd’hui estimé à une centaine chez la souris ou l’humain, environ la moitié étant exprimé à partir de l’allèle paternel et l’autre moitié à partir de l’allèle maternel. Certains de ces gènes sont impliqués dans la prolifération cellulaire et la croissance fœtale, d’autres jouent un rôle important dans des processus neurologiques et dans le comportement (Arnaud and Feil, 2005). Dans les cellules somatiques, la répression des gènes soumis à l’empreinte fait intervenir la méthylation de l’ADN au niveau de régions régulatrices nommées régions de contrôle de l’empreinte (Li et al., 1993a). L’étude des marques des histones présentes aux gènes soumis à l’empreinte dans des cellules somatiques, dans différents tissus, révèle que l’allèle exprimé, non méthylé, est associé à H3K4me₂, et H3K9ac (Fournier et al., 2002). L’allèle réprimé, méthylé, est en revanche caractérisé par la présence de modifications répressives comme H3K9me_{2/3} (Fournier et al., 2002), H4K20me₃ (Delaval et al., 2007) et H3K27me₃ (Umlauf et al., 2004).

Dans les cellules germinales primordiales, avant la méiose, l’empreinte établie précédemment est enlevée et remise en place selon le sexe de l’animal. Ce changement implique une reprogrammation épigénétique majeure du génome. La première étape de cette reprogrammation est une déméthylation active et globale du génome dans les PGC à E11,5, elle permet une perte de la méthylation de l’ADN au niveau des allèles réprimés des gènes soumis à l’empreinte (Tada et al., 1998) (Hajkova et al., 2002). La déméthylation globale du génome est suivie d’une perte de certaines modifications répressives de l’histone H3 comme H3K9me₃ et H3K27me₃, ainsi que d’une perte de l’histone H1. Ces modifications s’accompagnent d’un desserrement de la structure chromatinienne caractérisé par une augmentation de la taille du noyau et une disparition des marquages intenses de DAPI (Hajkova et al., 2008). Ces changements induisent une perte de l’empreinte parentale ainsi que la réactivation du X inactif chez les femelles (Chuva de Sousa Lopes et al., 2008). De plus, pendant cette période de reprogrammation, la protéine chaperonne HIRA est enrichie dans le noyau, suggérant l’incorporation de H3.3 et on observe une perte du variant H2A.Z à la chromatine. Il a donc été suggéré que la disparition des marques répressives peut être la

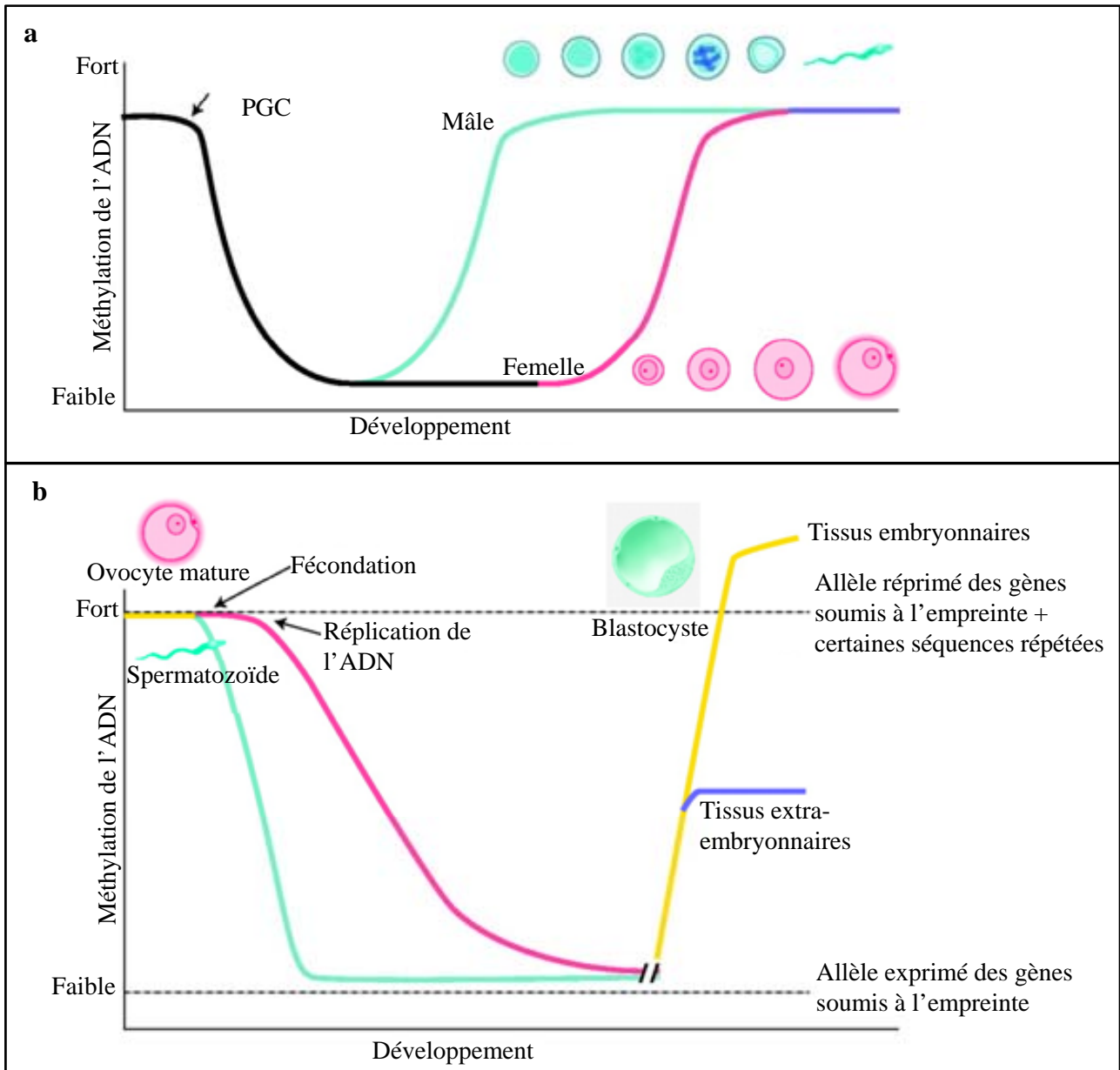


Figure 27 : Reprogrammation de la méthylation de l'ADN dans les cellules germinales primordiales et au cours du développement préimplantatoire de souris. **a** : L'ADN des PGC est déméthylé au cours du développement à E11,5 chez la souris. La reméthylation commence à E16 dans les spermatogonies et après la naissance dans les ovocytes en croissance. **b** : Dans le zygote, le génome paternel est déméthylé activement, rapidement après la fécondation. Le génome maternel est déméthylé par un mécanisme passif dépendant de la réplication de l'ADN. L'ADN de l'embryon est reméthylé avant l'implantation de manière différente entre les tissus embryonnaires et extra-embryonnaires. Les allèles réprimés des gènes soumis à l'empreinte et certaines séquences répétées ne sont pas déméthylés (ligne pointillée supérieure). Les allèles exprimés des gènes soumis à l'empreinte ne sont pas méthylés (ligne pointillée inférieure). La figure est adaptée d'après Reik et al., 2001.

conséquence d'un remplacement global des histones au sein de la chromatine des PGC à E11,5 (Hajkova et al., 2008).

La diminution du taux de modification des histones n'est que transitoire et des marques réapparaissent à partir de E12,5 en même temps que le marquage intense en DAPI (Hajkova et al., 2008). Ces observations ont été faites sur des taux globaux au sein du noyau et on peut penser que les modifications enlevées puis remises peuvent être associées à des loci totalement différents au sein du génome entre E11,5 et E12,5 et ainsi modifier totalement le programme d'expression génique de ces cellules, l'orientant vers un programme spécifique des cellules germinales. Alors que la perte des modifications d'histones n'est que transitoire, la reméthylation de l'ADN ne s'effectue qu'à des stades plus tardifs, dans l'ovocyte en croissance après la naissance chez les femelles et dans les spermatogonies et les spermatocytes I chez les mâles (Lucifero et al., 2002) (Kato et al., 2007) (figure 27a).

Les mécanismes responsables de la déméthylation active de l'ADN ne sont pas connus précisément. Des données récentes suggèrent que des mécanismes de réparation de l'ADN, comme l'excision et réparation d'une base, pourraient être mis en jeu dans les PGC au moment de la déméthylation globale du génome (Hajkova et al., 2010).

2. Réorganisation de la chromatine dans les spermatogonies

Pendant la spermiogenèse, les spermatides subissent un remplacement presque total de leurs histones en protamines, des protéines fortement basiques et de petit poids moléculaire. Entre le retrait des histones et l'incorporation des protamines, la chromatine est associée à 90% avec un troisième type de protéines basiques, les protéines de transition. Chez la souris, il existe deux protéines de transition majeures, TNP1 et TNP2 (pour une revue, Meistrich et al., 2003). La déplétion de TNP1 et TNP2 entraîne une diminution de la fertilité des mâles à cause d'une compaction insuffisante du génome des spermatozoïdes, mais elle n'abolit pas le remplacement des histones par les protamines (Zhao et al., 2004). Les protéines de transition ne sont donc pas nécessaires à l'éviction des histones dans les spermatides. Ces protéines sont ensuite remplacées par les protamines PRM1 et PRM2. Le remplacement des histones par des protamines permet une compaction du génome du spermatozoïde six fois supérieure à la compaction du génome somatique (Ward and Coffey, 1991) et évite l'apparition de dommages sur l'ADN avant la fécondation (Cho et al., 2003). Cette éviction globale des histones induit également une suppression des marques épigénétiques portées par les histones

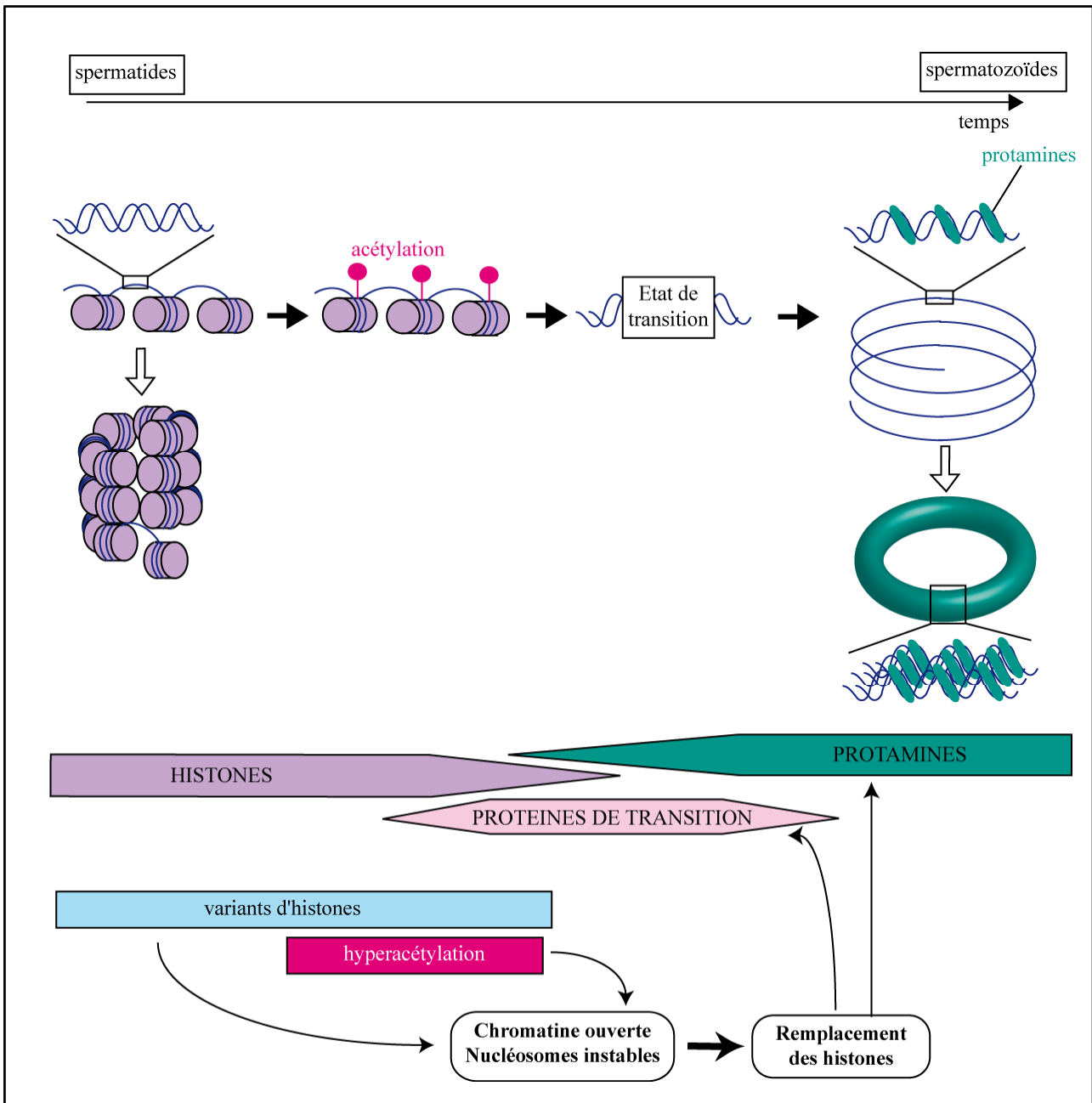


Figure 28 : Transition entre le génome des spermatides condensé autour d'histones et le génome des spermatozoïdes condensé autour de protamines. Dans les spermatides, le nucléosome est l'unité de base de la chromatine et permet une condensation plus ou moins importante du génome, schématisée ici par un repliement de type solénoïde. Dans les spermatozoïdes, le génome est hautement compacté en structures en forme de « *doughnuts* » grâce à son interaction avec des protamines. La suppression des histones se fait grâce à une déstabilisation des nucléosomes suite à leur hyperacétylation et à l'incorporation de certains variants d'histones. Les histones sont d'abord remplacées par des protéines de transition, puis par les protamines. La figure est adaptée de Ward et al., 1993 et Gaucher et al., 2009.

et est donc la première étape d'une reprogrammation épigénétique du génome paternel qui se poursuit dans l'embryon avec l'établissement de nouvelles marques.

Les mécanismes moléculaires permettant l'incorporation puis le retrait des protéines de transition ainsi que l'incorporation des protamines ne sont pas complètement élucidés. Des données suggèrent que des chaperonnes, et plus particulièrement la chaperonne HSP70.2, pourraient jouer un rôle dans cette réorganisation (Govin et al., 2006). Il est également possible que le remplacement des histones se fasse par simple compétition des différentes protéines pour la liaison à l'ADN, compétition gagnée directement par les protamines dans des souris ne possédant pas de protéines de transition. Le désassemblage de la chromatine est précédé par une vague globale d'acétylation des histones (Hazzouri et al., 2000) et par l'incorporation au sein de la chromatine de variants des histones spécifiques des testicules dont H1t, H2BL1/2, H2AL1/2/3 ainsi que H3.3 et H2AZ. L'acétylation globale et la présence de ces variants d'histones diminue la stabilité des nucléosomes, facilitant le déplacement et l'éviction des histones de la chromatine (pour une revue, Caron et al., 2005 Gaucher et al., 2010) (figure 28).

Le remplacement des histones par des protamines a longtemps été considéré comme total. Quelques études montrent aujourd'hui que chez la souris et chez l'homme, respectivement 1% et 15% de l'ADN des spermatozoïdes est lié à des nucléosomes (Bench et al., 1996) (Tanphaichitr et al., 1978) contenant entre autres les variants H3.3, H2A.Z et CENP-A (Gatewood et al., 1990) (Palmer et al., 1990). Ces nucléosomes peuvent présenter des modifications post-traductionnelles et ils ne semblent pas répartis aléatoirement sur le génome. Ils semblent majoritairement liés à des locus importants pour le développement de l'embryon comme les gènes soumis à l'empreinte, les gènes HOX, des gènes codant pour des facteurs de transcription exprimés au cours du développement (Hammoud et al., 2009) ainsi qu'aux centromères (Palmer et al., 1990). Bien qu'il a été montré que ces histones dérivées des spermatozoïdes contribuent à la formation de la chromatine de l'embryon (van der Heijden et al., 2008), aucun rôle précis n'a pu leur être attribué. Des hypothèses ont cependant été émises concernant ce rôle : ils pourraient permettre une transcription rapide des gènes marqués dès la fécondation voire dans le spermatozoïde mature, la transmission d'une information épigénétique du spermatozoïde à l'embryon, ou pourraient servir de zone de nucléation pour le remodelage de la chromatine après la fécondation.

3. Reprogrammation du génome dans le zygote

Après la fécondation, les génomes parentaux sont à différents stades du cycle cellulaire, et possèdent des structures chromatiniennes et des marques épigénétiques différentes. Bien qu'ils aient en théorie accès aux mêmes facteurs maternels, les deux pronoyaux vont subir des programmes de remodelage de la chromatine différents, induisant de nombreuses asymétries entre les génomes maternel et paternel pendant le stade zygote.

a. Evolution du contenu en histones

Le génome paternel est condensé majoritairement autour de protamines, alors que le génome maternel, encore en métaphase II de méiose, est condensé autour d'histones. Après la fécondation, alors que le génome maternel termine sa deuxième division méiotique, les protamines sont remplacées par des histones dans le pronoyau mâle, elles sont alors indétectables 80 minutes après la fécondation (van der Heijden et al., 2005). Les histones H3 et H4 sont traduites à partir des ARNm maternels alors que les histones H2A et H2B sont directement apportées sous forme de protéines dans le cytoplasme de l'ovocyte. La présence de la protéine HIRA couplée à des données d'expériences d'immunofluorescence dans le pronoyau paternel au moment de l'assemblage de la chromatine montrent que H3.3 est le seul variant de H3 à être incorporé aux nucléosomes paternels juste après la fécondation (van der Heijden et al., 2005) (Torres-Padilla et al., 2006). L'absence d'incorporation de H3.1 et H3.2 peut être expliquée par le fait que ce remplacement global est réalisé indépendamment de la synthèse d'ADN, il ne peut donc faire intervenir que le variant non canonique de H3. Du fait que la chromatine maternelle contienne, au moins en partie, H3.1 et H3.2, la présence exclusive de H3.3 dans la chromatine paternelle fait apparaître une première asymétrie majeure entre les deux pronoyaux. Cette asymétrie globale est maintenue jusqu'à la réplication, où les variants canoniques de H3 sont incorporés au sein de la chromatine paternelle.

b. Evolution des modifications des histones

Une deuxième asymétrie évidente vient du fait que, alors que des modifications chromatiniennes sont déjà abondantes dans le pronoyau femelle, le génome mâle doit réacquies toutes les marques post-traductionnelles sur ses histones nouvellement incorporées,

et ce, aux loci appropriés. Ces marques doivent permettre la transcription des gènes adéquats lors de la ZGA, la répression des gènes indésirables, et l'établissement des structures chromatinienne nécessaires à la première mitose comme les centromères par exemple. La modification progressive des histones de la chromatine paternelle tend à l'acquisition d'un état chromatinien proche de celui du noyau femelle, permettant aux deux génomes de subir une mitose commune. Les histones incorporées dans le pronoyau mâle sont plus fortement acétylées que les histones du noyau femelle. Rapidement après l'incorporation des histones, H3K4me1, H4K20me1, H3K9me1 et H3K27me1 sont détectées dans le pronoyau mâle (van der Heijden et al., 2005) (Santos et al., 2005). Les degrés de méthylation supérieurs de H3K4 et H3K27, bien que présents abondamment dans le pronoyau femelle, n'apparaissent que faiblement dans le pronoyau mâle qu'en fin de stade PN3, début de stade PN4 (Santos et al., 2005). Les marques liées à l'hétérochromatine, H3K9me3 et H4K20me2/3, restent absentes de la chromatine paternelle au cours du stade zygote. A cause d'une absence d'activité de SUV39H dans l'embryon jusqu'au stade 4-cellules, peut-être due à la présence d'inhibiteurs, aucune triméthylation de H3K9 n'est détectée au sein de la chromatine paternelle jusqu'à ce stade (Liu et al., 2004). Malgré l'absence de H3K9me3, l'hétérochromatine péricentrique paternelle est marquée par une faible quantité de HP1 β (Arney et al., 2002). HP1 β est la seule marque hétérochromatique constitutive à être présente à la chromatine paternelle dans le zygote, le génome paternel est donc considéré comme ne possédant pas de réelle hétérochromatine avant l'apparition de H3K9me3 au stade 4-cellules.

Alors que des changements dynamiques des marques et des variants d'histones interviennent au niveau du pronoyau mâle, la chromatine du pronoyau maternel est plus stable et les marques chromatinienne de l'ovocyte ne subissent pas de modifications aussi drastiques que celles du spermatozoïde. Seulement certains changements ont lieu, majoritairement, concernant les marques euchromatiques en lien direct avec la réplication ou l'activation transcriptionnelle comme l'acétylation de H3 et H4 et la méthylation des arginines (Sarmiento et al., 2004) (Adenot et al., 1997).

c. Evolution de la méthylation de l'ADN

Une dernière asymétrie entre les pronoyaux mâles et femelles est révélée par différents patrons globaux de méthylation de l'ADN et plus précisément par la dynamique globale de déméthylation de l'ADN qui suit la fécondation (figure 27b). Alors que le pronoyau paternel est déméthylé très rapidement après la fécondation et en partie avant la première réplication

de l'ADN, donc certainement via un mécanisme actif, le pronoyau maternel n'est que passivement déméthylé au cours des cycles de réplication successifs (Mayer et al., 2000a) (Oswald et al., 2000). Comme dans les PGC, la déméthylation active de l'ADN pourrait être la conséquence d'un mécanisme de réparation de l'ADN (Hajkova et al., 2010). Une autre étude récente a mis en évidence l'importance d'une sous-unité du complexe Elongator dans la déméthylation active du génome paternel (Okada et al., 2010). Le rôle précis joué par ce complexe, connu pour interagir avec l'ARN pol II au cours de l'élongation de la transcription, sur la déméthylation des cytosines n'est pas défini.

La déméthylation de l'ADN n'est pas uniforme sur tout le génome, pour les gènes soumis à l'empreinte parentale par exemple, la méthylation de l'allèle du parent réprimé est maintenue après la fécondation et au cours du développement préimplantatoire grâce aux protéines DNMT1o et DNMT1s (Tremblay et al., 1997) (Hirasawa et al., 2008). De plus, alors que l'ADN péricentromérique maternel est déméthylé comme le reste du génome, ces mêmes régions au sein du génome paternel gardent leur méthylation (Rougier et al., 1998). Suite à des expériences montrant que la méthylation de H3K9 pouvait induire la méthylation de l'ADN (Jackson et al., 2002) (Lehnertz et al., 2003), il a été suggéré que la présence de H3K9me2/3 devait protéger le génome maternel d'une déméthylation active, l'absence de H3K9me2/3 au sein du génome paternel la rendant possible spécifiquement dans le pronoyau mâle (Santos et al., 2005) (Burton and Torres-Padilla, 2010). Cette déméthylation globale du génome après la fécondation fait partie du processus de reprogrammation du génome zygotique. La déméthylation plus précoce du génome paternel peut être mise en lien avec le fait qu'il sera également plus fortement transcrit au stade zygote. Cette observation est appuyée par l'observation que les espèces présentant une ZGA précoce sont aussi sujettes à une déméthylation du génome zygotique plus importante que les espèces où la ZGA est plus tardive (Beaujean et al., 2004) (Memili and First, 2000). Néanmoins, le fait que la déméthylation soit la cause ou la conséquence de la ZGA précoce n'est pas établi.

La déméthylation globale des génomes parentaux continue jusqu'en fin de stade morula et est suivie par une méthylation *de novo* spécifiquement dans les cellules de la MCI (Santos et al., 2002). Cette méthylation de l'ADN *de novo* dans la MCI induit une asymétrie épigénétique entre les deux premiers types cellulaires différenciés au cours du développement.

IV. Les régions péricentromériques dans l'embryon, un exemple de réorganisation de la chromatine au cours du développement précoce de la souris.

Les régions péricentromériques sont déjà sous forme d'hétérochromatine dans le génome maternel au moment de la fécondation alors qu'elles doivent acquérir une structure compactée compatible avec la réalisation de la première mitose dans le génome paternel. Les mécanismes responsables de l'hétérochromatinisation des régions péricentromériques paternelles ne sont pas connus, mais ils sont accompagnés d'une évolution bien caractérisée de la localisation et de la transcription de ces régions dans les deux pronoyaux. Nous verrons ici que ces régions constituent un exemple d'étude des asymétries existant entre les génomes mâle et femelle et des réorganisations subies par la chromatine au cours du développement préimplantatoire de la souris.

1. Evolution de la localisation des régions péricentromériques au cours du développement précoce de la souris

Comme nous l'avons vu dans le chapitre précédent, les régions péricentromériques sont composées de séquences répétées, les *minor satellites* centromériques et les *major satellites* péricentriques. Des études de FISH sur l'ADN ont permis de mettre en évidence une évolution de la localisation des satellites dans l'embryon entre la fécondation et la fin du stade 2-cellules (Probst et al., 2010) (Probst et al., 2007) (figure 29a).

Dans les ovocytes, les *major satellites* sont groupés au sein du noyau et associés aux *minor satellites*. Après la fécondation, au stade PN1, les *major* et *minor satellites* maternels s'individualisent et présentent une structure proche des chromocentres des cellules somatiques où les *major satellites* forment des structures ponctuelles associées jusqu'à trois groupements de *minor satellites*. A ce stade, une structure particulière en anneau se forme au sein des pronoyaux maternels, le précurseur non fonctionnel du nucléole (NLB : *nucleolar like body*) (Kopecny et al., 1995). Les *major satellites* s'alignent alors de manière discontinue autour du NLB unique, certains *minor satellites* s'insérant dans les discontinuités. Au stade PN2, 70% des zygotes possèdent cette organisation en anneau des régions péricentromériques dans le pronoyau maternel (Probst et al., 2007).

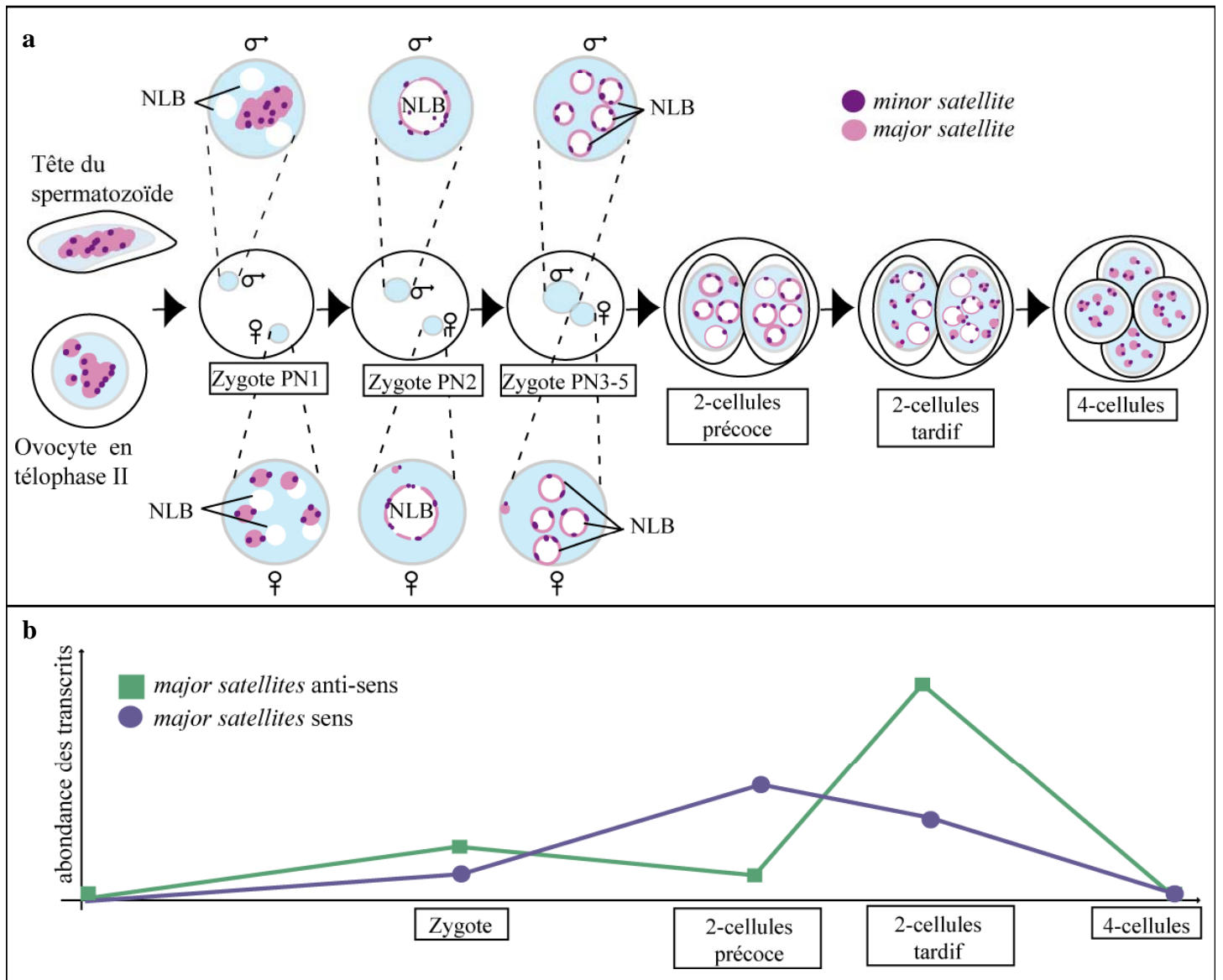


Figure 29: Evolution de la localisation et du niveau d'expression des répétitions péricentromériques au cours du développement précoce chez la souris. **a** : La localisation des *major satellites* et des *minor satellites* au sein des noyaux dans les gamètes et dans l'embryon précoce de souris a été déterminée par du FISH sur ADN. Les noyaux sont schématisés en bleu clair, dans les embryons, un grossissement des pronoyaux mâles (en haut) et femelle (en bas) est réalisé pour plus de clarté. La taille relative entre les cellules et les noyaux et entre les cellules entre elles n'est pas représentative de la réalité. **b** : Les quantités relatives entre les transcrits sens (en bleu) et antisens (en vert) des *major satellites* a été déterminée par RT-PCR à différents stades du développement. Les schémas sont réalisés d'après les données de Probst et al., 2008 et Probst et al., 2011.

Dans les spermatozoïdes en revanche, les *major satellites* occupent de manière diffuse le cœur du noyau et les *minor satellites* y sont associés sous forme de points. Après la fécondation, la chromatine mâle se décondense et, alors qu'un NLB se forme dans le pronoyau à partir du stade PN1, les *major satellites* s'y localisent au stade PN2 sous forme d'un anneau continu. Les *minor satellites* sont associés de part et d'autre de cet anneau (figure 29a).

A partir de la fin du stade PN2, la localisation des régions péricentromériques est comparable entre les deux pronoyaux. Alors que la maturation du zygote avance, le nombre de NLB par pronoyau augmente. Les répétitions péricentromériques y restent associées dans les deux pronoyaux, à l'exception d'en moyenne un ou deux chromosomes par noyau.

Après la mitose, au début du stade 2-cellules, la majorité des satellites centromériques parentaux est toujours associée aux NLB encore présents dans les pronoyaux. Entre 24h et 36h après la fécondation (36-48h phCG), les séquences répétées péricentromériques subissent une nouvelle réorganisation. Les *major satellites* perdent leur organisation en anneaux et forment des chromocentres. L'organisation des régions péricentromériques sous forme de chromocentres est ensuite conservée dans la plupart des cellules somatiques.

2. Asymétrie épigénétique entre les régions péricentromériques mâles et femelles

Comme nous l'avons vu précédemment, il existe une asymétrie importante entre les marques hétérochromatiques portées par les génomes paternel et maternel au stade zygote. Alors qu'à PN3-5, les anneaux formés par les *major satellites* sont fortement enrichis en H3K9me3 et H4K20me3 dans le pronoyau femelle, ces marques sont absentes du pronoyau mâle. La chromatine péricentrique mâle est en revanche enrichie en H3K9me1 et H3K27me3. L'acquisition des marques hétérochromatiques constitutives par la chromatine paternelle s'effectue au stade 4-cellules et les chromatines péricentromériques parentales deviennent similaires à partir du stade 8-cellules (Probst et al., 2007) (Santos et al., 2005).

Malgré l'absence de H3K9me3 à la chromatine paternelle, une colocalisation claire entre HP1 β et les *major satellites* apparaît dans les deux pronoyaux. Il a été suggéré que le recrutement de HP1 β aux régions péricentriques paternelles pourrait se faire grâce à H3K9me1 et que la présence de cette protéine serait une marque pour cibler la triméthylation de H3K9 pendant des stades de développement plus tardifs (Santos et al., 2005). Le

mécanisme par lequel HP1 β pourrait, seul, permettre la condensation de la chromatine péricentromérique paternelle et l'assemblage du kinétochore nécessaire à la ségrégation des chromatides lors de la première division méiotique chez la souris n'a pas été étudié.

3. Evolution de la transcription aux régions péricentriques

L'organisation des régions péricentriques en anneaux puis en chromocentres au début du développement de la souris est accompagnée de la transcription des séquences répétées qui les composent selon différentes orientations (figure 29b).

Des expériences de FISH sur les ARN des *major satellites* ont mis en évidence, au moment de la première ZGA, au stade PN3-4, une augmentation de la transcription des *major satellites* dans le zygote majoritairement à partir du génome paternel. Par RT-PCR spécifique d'un brin, il a été possible de montrer que les transcrits présents à ce stade sont pour la plupart d'orientation anti-sens. Entre le début du stade 2-cellules et la fin de la phase S de ce même stade, une augmentation de l'expression du transcrit sens est détectée à partir du génome paternel. Suite à cette augmentation, et alors que les majors satellites se réorganisent en chromocentres, le transcrit anti-sens est surexprimé à partir des deux génomes jusqu'à la fin du stade 2-cellules. Les transcrits sens et anti-sens coexistent donc dans l'embryon pendant la phase G2 du stade 2-cellules permettant la formation potentielle d'ARNdb. L'expression des *major satellites* diminue ensuite fortement pour devenir presque indétectable au stade 8-cellules. L'inhibition de l'expression des *major satellites* est alors en accord avec l'état fortement condensé de ces régions hétérochromatiques (Probst et al., 2010). Il est important de noter que ces résultats sur la transcription bidirectionnelle des *major satellites* n'ont été publiés que tard au cours de ma thèse, ces données n'étaient donc pas disponibles lors de la réalisation de la publication 2.

Partie 2

ETUDE DU ROLE DE

TBP2 DANS LES

OVOCYTES DE SOURIS

A. Problématique

Dans les ovocytes de souris, l'expression de la protéine TBP cesse à partir du stade follicule secondaire alors que la transcription est encore active. En revanche, l'homologue de TBP spécifique des vertébrés, la protéine TBP2, est présente tout au long de la croissance et de la maturation des ovocytes. Suite à ces observations, l'objectif de mon travail de thèse était donc d'essayer de comprendre l'importance et le rôle spécifique de la protéine TBP2 au cours de la folliculogenèse *in vivo*. Alors qu'un rôle direct de la protéine TBP2 sur la transcription a été mis en évidence chez le poisson zèbre et chez le xénope, très peu de données fonctionnelles étaient disponibles chez la souris. Nous voulions donc savoir si TBP2 joue un rôle dans la transcription des gènes spécifiques de l'ovocyte chez la souris. Pour répondre à ces questions, nous avons, en collaboration avec une autre étudiante du laboratoire du Dr Tora, effectué la caractérisation phénotypique macroscopique de souris déficientes pour la protéine TBP2, générées au laboratoire. Une étude plus approfondie des ovaires de ces souris a également été réalisée. En parallèle, le rôle de la protéine TBP2 dans des ovaires de souris sauvages a été analysé.

De plus, une controverse existe concernant la localisation de TBP2 chez la souris adulte. En effet, bien que l'ARNm de la protéine TBP2 n'ait été détecté que dans les ovocytes, certaines études rapportent un rôle de cette protéine dans d'autres types cellulaires comme les myotubes et les cellules C2C12, des cellules myoblastiques (Deato et al., 2008) (Deato and Tjian, 2007). L'étude des souris TBP2^{-/-} a donc également permis de déterminer, en partie, si TBP2 est nécessaire à d'autres processus que celui de la folliculogenèse.

B. Publication 1

I. Résumé de la publication 1

Note : les figures appelées dans le texte sont celles de la publication 1

Cette étude a été effectuée en suivant deux axes principaux. Dans un premier temps, la caractérisation phénotypique des souris déficientes pour TBP2, souris TBP2^{-/-}, a été réalisée. Puis, dans un deuxième temps, les rôles de TBP2 et de TBP ont été analysés dans des ovocytes et des embryons de souris sauvages.

1. Etude des souris TBP2^{-/-}

a. Les souris TBP2^{-/-} sont stériles suite à un arrêt de la folliculogénèse

Une lignée de souris, supposée exprimer une protéine TBP2 ne possédant pas de domaine de liaison à l'ADN, a été générée. L'expression de la protéine TBP2 s'est révélée être complètement abolie dans les ovaires des souris homozygotes que nous caractériserons donc de lignée TBP2^{-/-} (figure 1A et B).

Les souris TBP2^{-/-} sont viables et ne présentent pas de défauts anatomiques apparents. En revanche, les femelles TBP2^{-/-} sont stériles et aucun ovocyte ou follicule mature n'est présent dans leurs ovaires. Plus précisément, nous avons mis en évidence un arrêt de la folliculogénèse entre les stades follicule primaire et follicule secondaire chez les mutants (figure 1C, D et E, et figure 2).

Ces résultats montrent que TBP2 est nécessaire à la croissance des ovocytes et au développement des follicules.

b. La perte de TBP2 induit une diminution de la transcription et l'altération du profil d'expression des gènes spécifiques des ovocytes

Le taux de transcription active a ensuite été analysé dans les ovocytes mutants par immunohistochimie avec des anticorps reconnaissant H3K4me3 et la phosphorylation de la

sérine 2 de la CTD de l'ARN pol II. Ces expériences révèlent une diminution de ces deux marques de transcription active dans les ovocytes mutants, suggérant que TBP2 est, au moins en partie, nécessaire à l'activité transcriptionnelle globale des ovocytes (figure 3). Pour préciser cette observation, une analyse transcriptomique des ovaires TBP2^{-/-} a été effectuée. Les résultats montrent la dérégulation d'un certain nombre de gènes par rapport aux souris sauvages, la moitié des gènes dérégulés étant surexprimée et l'autre moitié sous-exprimée. 30% des gènes sous-exprimés en l'absence de TBP2 sont des gènes spécifiques des ovocytes et les promoteurs des gènes sous-exprimés sont spécifiquement enrichis en régions riches en adénines et thymines incluant la boîte TATA (figure 4). De plus, en l'absence de TBP2, la chromatine des ovocytes mutants semble moins compactée que celle des ovocytes sauvages (figure 3D).

Ces données démontrent que la suppression de TBP2 induit des changements majeurs du transcriptome des ovaires ainsi que du degré de compaction de la chromatine dans les ovocytes. Ils suggèrent également que TBP2 pourrait jouer un rôle direct sur la transcription des gènes spécifiques des ovocytes.

2. Spécificité des rôles de TBP et TBP2 au cours de la gaméto-genèse et du développement

a. TBP2 se lie aux promoteurs des gènes activement transcrits dans l'ovocyte

Des expériences de ChIP ont ensuite été réalisées avec un anticorps spécifique de TBP2 sur des ovaires de souris sauvages. TBP2 se lie aux promoteurs de *Zp3* et *Gdf9*, deux gènes spécifiques des ovocytes, sous-exprimés dans le mutant, ainsi qu'au promoteur d'un gène de l'histone H2B (figure 5).

Ces résultats confirment le rôle direct de TBP2 sur la transcription dans l'ovocyte.

b. TBP n'est pas nécessaire à la maturation des ovocytes, à la ZGA et au développement précoce

Alors que TBP2 est exprimée tout au long de la folliculogenèse, TBP n'est présente que jusqu'au stade follicule secondaire. Pour déterminer l'importance de TBP au cours de la folliculogenèse, des souris femelles n'exprimant pas TBP spécifiquement dans leurs ovocytes

ont été générées (figure 6A). Les ovocytes TBP^{-/-} ne présentent pas de problèmes de maturation et sont fécondés dans les mêmes proportions que les ovocytes sauvages. Après fécondation des ovocytes mutants par des spermatozoïdes sauvages, le développement des embryons, ne possédant donc pas l'allèle maternel de TBP, suit celui des embryons sauvages et les nouveau-nés sont viables (figure 6B et C). Contrairement à TBP2, TBP n'est donc pas indispensable à la maturation des ovocytes et la contribution maternelle de TBP n'est pas nécessaire pour que l'embryon atteigne le stade blastocyste, voire pour le développement embryonnaire.

c. La surexpression de TBP2 a un effet négatif sur le développement embryonnaire et la croissance cellulaire

Alors que TBP réapparaît dans le zygote, TBP2 devient indétectable à ce stade. Pour déterminer si TBP et TBP2 peuvent agir en parallèle dans l'embryon, des expériences de surexpression de TBP2 ont été menées dans le zygote. Les embryons exprimant TBP2 cessent de se développer entre le stade 2-cellules et 8-cellules, montrant que l'expression de TBP2 est délétère au cours du développement, et confirmant que TBP et TBP2 jouent donc des rôles distincts sur la transcription chez la souris (figure 6E et F).

Cette étude nous a donc permis de mettre en évidence le rôle direct de TBP2 sur la transcription des gènes de l'ovocyte. Son expression dans les ovocytes, contrairement à celle de TBP, est nécessaire à leur croissance et leur maturation. De plus, la disparition de TBP2 à partir du stage zygote est indispensable au développement embryonnaire précoce.

II. Publication 1

TBP2 is essential for germ cell development by regulating transcription and chromatin condensation in the oocyte

Emese Gazdag,^{1,4} Angèle Santenard,^{1,2,4} Céline Ziegler-Birling,^{1,2} Gioia Altobelli,^{1,3} Olivier Poch,³ László Tora,^{1,5} and Maria-Elena Torres-Padilla^{1,2,6}

¹Department of Functional Genomics, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), UMR 7104 CNRS, Uds, INSERM U964, BP 10142, F-67404 Illkirch Cedex, CU de Strasbourg, France; ²Department of Developmental and Cell Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), UMR 7104 CNRS, Uds, INSERM U964, BP 10142, F-67404 Illkirch Cedex, CU de Strasbourg, France; ³Bioinformatics and Integrative Biology Laboratory, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), UMR 7104 CNRS, Uds, INSERM U964, BP 10142, F-67404 Illkirch Cedex, CU de Strasbourg, France

Development of the germline requires consecutive differentiation events. Regulation of these has been associated with germ cell-specific and pluripotency-associated transcription factors, but the role of general transcription factors (GTFs) remains elusive. TATA-binding protein (TBP) is a GTF involved in transcription by all RNA polymerases. During ovarian folliculogenesis in mice the vertebrate-specific member of the TBP family, TBP2/TRF3, is expressed exclusively in oocytes. To determine TBP2 function in vivo, we generated TBP2-deficient mice. We found that *Tbp2*^{-/-} mice are viable with no apparent phenotype. However, females lacking TBP2 are sterile due to defective folliculogenesis, altered chromatin organization, and transcriptional misregulation of key oocyte-specific genes. TBP2 binds to promoters of misregulated genes, suggesting that TBP2 directly regulates their expression. In contrast, TBP ablation in the female germline results in normal ovulation and fertilization, indicating that in these cells TBP is dispensable. We demonstrate that TBP2 is essential for the differentiation of female germ cells, and show the mutually exclusive functions of these key core promoter-binding factors, TBP and TBP2, in the mouse.

[*Keywords:* TBPL2; TRF3; TBP; Zp3; RNA Polymerase II; ovary]

Supplemental material is available at <http://www.genesdev.org>.

Received April 16, 2009; revised version accepted July 27, 2009.

Regulation of transcription initiation by RNA polymerase II (Pol II) is central to any developmental process. Development of the germline is essential for transmitting information between generations and for sexual reproduction. Female germ cells develop during oogenesis, leading to the formation of a highly differentiated and specialized cell, the oocyte. Oogenesis comprises an orderly number of events, each of which must be regulated. Although some work has been done to understand how specific transcription factors (Rajkovic et al. 2004; Choi and Rajkovic 2006) regulate oocyte-specific gene expression, little is known about the role of general transcription factors (GTFs) in regulation of female germ cell development.

The general RNA Pol II transcription factor TFIID is composed of TATA-binding protein (TBP) and 14 TBP-

associated factors (TAFs) (Burley and Roeder 1996). TBP plays a crucial role in preinitiation complex assembly by nucleating the binding of GTFs to the core promoters of genes. TBP also plays a role in Pol I and Pol III transcription. TBP possesses two regions: an N-terminal domain and a highly conserved saddle-like C-terminal DNA-binding domain (Hernandez 1993). The N termini of the different TBP proteins are divergent among species and seem to have a role in regulation of Pol II and Pol III transcription (Muller and Tora 2004). In contrast to the generally accepted idea that TBP is a universal transcription factor, recent studies documented TBP-independent Pol II transcription (Muller and Tora 2004). Indeed, in the absence of TBP, TBP paralogs are able to mediate Pol II and Pol III transcription of developmentally important genes in vertebrate embryos (Takada et al. 2000; Veenstra et al. 2000; Muller et al. 2001; Martianov et al. 2002a; Ferg et al. 2007; Jacobi et al. 2007). TBP-like factor (TLF, also known as TBPL1, TRF2, or TRP) has been found in different metazoan species. TLF shares ~45% identity with the core domain of TBP and can interact with TFIIA and TFIIB, but cannot bind the TATA box (Muller and

⁴These authors contributed equally to this work.

Corresponding authors.

⁵E-MAIL laszlo@igbmc.u-strasbg.fr; FAX 33-3-88-65-32-01.

⁶E-MAIL metp@igbmc.fr; FAX 33-3-88-65-32-01.

Article is online at <http://www.genesdev.org/cgi/doi/10.1101/gad.535209>.

Tora 2004). In contrast, another vertebrate-specific TBP-related factor, TBP2 (also known as TBPL2 or TRF3) shows ~92% conservation in its DNA-binding core domain compared with TBP. Indeed, TBP2 is able to bind the TATA box, interact with TFIIA and TFIIB, and mediate Pol II transcription initiation in vitro (Bartfai et al. 2004; Jallow et al. 2004; Deato et al. 2008). Moreover, a role for TBP2 in directing cell type-specific transcription using an ex vivo model of myotube differentiation has been reported (Deato and Tjian 2007). In mice, the expression of *Tbp2* mRNA has been detected exclusively in the ovary, specifically in the oocytes (Bartfai et al. 2004; Xiao et al. 2006). Enrichment of TBP2 in the ovary is conserved in zebrafish and *Xenopus*, further suggesting a fundamental role for TBP2 in the vertebrate ovary (Bartfai et al. 2004; Jallow et al. 2004).

Most oocytes enter meiosis during the embryonic life or during the early neonatal period. At birth, ovaries contain clusters of oocytes without surrounding granulosa cells that are in transitory stages of prophase (pachytene to early diplotene) or in late diplotene (Pedersen and Peters 1968). From day 3 after birth, some primordial follicles form, undertake a growth phase, and enter folliculogenesis (Matzuk and Lamb 2002). Oocytes within the primary follicle start to increase their size and undergo growth during the preantral follicular stages. Later, in antral, preovulatory follicles, the fully grown germinal vesicle (GV) stage oocyte is ready for ovulation. Maturation follows with the resumption of meiosis induced by ovulatory stimulus and proceeds until the metaphase II, at which stage the oocyte awaits fertilization. Genes specifically expressed in the oocyte are necessary for either growth or communication with follicular cells. These include *Zona Pellucida 3 (Zp3)* (Bleil and Wassarman 1980), the growth factors *Bmp15* (bone morphogenetic protein 15) and *Gdf9* (growth differentiation factor 9) (Yan et al. 2001), and transcription factors specific to the oocyte such as *Nobox* (Rajkovic et al. 2004). Null females for all these genes are sterile (Andreu-Vieyra et al. 2006).

The identification of gonad-specific variants of core promoter-binding factors suggests that, in the gonads of different metazoan organisms, specific transcription initiation mechanisms have evolved (Muller and Tora 2004). The oocyte-specific expression of TBP2 (Xiao et al. 2006; Gazdag et al. 2007) prompted us to determine the role of TBP2 in vivo. We show that *Tbp2*^{-/-} mice are viable and display no obvious morphological phenotype. However, females lacking TBP2 are sterile due to defective folliculogenesis. Ovaries from *Tbp2*^{-/-} females show defects in the formation of secondary follicles. *Tbp2*^{-/-} females lack fully grown GV stage oocytes and do not ovulate. Transcriptome analysis of ovaries from *Tbp2*-deficient mice revealed a high number of oocyte-specific genes severely deregulated. TBP2 binds to some of their promoters, and this suggests a direct regulation. In contrast to TBP2, we show that TBP is dispensable for oocyte maturation and fertilization. Finally, when TBP2 is mis-expressed in mouse embryos, it has a negative effect on cell proliferation, leading to developmental arrest. Thus,

TBP2 plays a critical and specialized role in mammalian female germ cell development.

Results

Tbp2^{-/-} mice are viable without obvious abnormalities or apparent anatomical aberration

The dynamic nature and specific expression of TBP2 during oogenesis (Xiao et al. 2006; Gazdag et al. 2007) suggests that it may play a role in female germ cell development. To investigate the in vivo role of TBP2, the *Tbp2* (*Tbpl2*) gene was targeted by homologous recombination and a Cre-loxP strategy. Embryonic stem (ES) cells were generated in which exon 4 of *Tbp2* was flanked by two loxP sites (Supplemental Fig. S1A). To obtain *Tbp2*^{+/-} mice, we deleted floxed exon 4 by breeding mice carrying the recombined allele with CMV-Cre mice (White et al. 1997), which led to the generation of the null allele (Supplemental Fig. S1A,B). Deletion of exon 4 is expected to produce a shorter mRNA that, if stable, would give rise to a TBP2 protein truncated at amino acid 206 in which the core domain is absent, thereby eliminating its DNA-binding domain. *Tbp2*^{+/-} mice were viable and indistinguishable from their wild-type littermates. To generate *Tbp2*-null mice, we performed *Tbp2*^{+/-} crosses. *Tbp2*^{-/-} mice were born at normal Mendelian ratios, indicating no embryonic lethality (Table 1). Expression of TBP2 mRNA was strongly reduced in *Tbp2*^{-/-} ovaries (Fig. 1A) and TBP2 protein was not detected in extracts from these ovaries, demonstrating that we generated a null allele (Fig. 1B). *Tbp2* heterozygotes showed reduced levels of TBP2 mRNA, but protein levels appeared largely unaffected compared with wild-type littermates (Fig. 1A,B). *Tbp2*^{-/-} mice developed to adulthood normally, were of normal size and weight, and showed no external abnormalities or apparent anatomical aberrations. Hence, TBP2 is not essential for mouse viability.

Tbp2^{-/-} females are sterile due to defective folliculogenesis

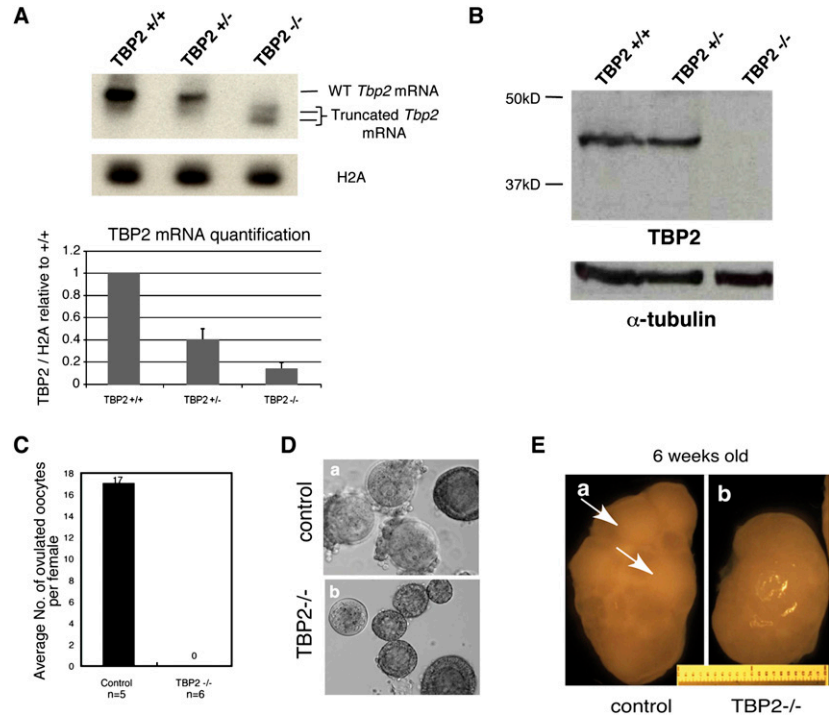
Given the specific expression of TBP2 in the female gonads (Xiao et al. 2006; Gazdag et al. 2007), we next asked whether *Tbp2*^{-/-} mice exhibited normal reproductive success. Wild-type, heterozygous, and mutant females were caged with wild-type males over a breeding period of 6 mo. While *Tbp2*^{+/+} and *Tbp2*^{+/-} females gave birth repeatedly and with normal litter sizes, *Tbp2*-null females never gave birth when caged with several fertile males (Table 2). This has been observed for *Tbp2*^{-/-} females for a continued breeding period of 12 mo, indicating that *Tbp2*^{-/-} females are sterile. In contrast,

Table 1. Mice bearing a deletion of TBP2 are viable and born in normal Mendelian ratios

Genotype	+/+	+/-	-/-	Total
Number of pups (%)	47 (22.2%)	116 (54.7%)	49 (23.1%)	212 (100%)

Genotype of pups derived from *TBP2*^{+/-} crosses.

Figure 1. Characterization of *Tbp2*-deficient mice. (A) RNA from ovaries derived from females of the indicated genotype was reverse-transcribed and amplified with primers specific for TBP2. PCR products were transferred and hybridized with a radioactive *Tbp2*-specific probe to confirm the specificity of the products and to allow quantification. Histone H2A was used as internal control. Quantification of TBP2 mRNA levels in *Tbp2*^{+/+}, *Tbp2*^{+/-}, and *Tbp2*^{-/-} ovaries is shown on the graph. Data are normalized to H2A expression and calculated relative to TBP2 mRNA expression in wild-type ovaries. Average \pm SEM of two ovaries per genotype are represented. (B) TBP2 protein is absent in *Tbp2*^{-/-} ovaries. Western blot analysis on total extracts from ovaries of females of the indicated genotype using the α -TBP2 antibody. α -Tubulin was used as a loading control. Shown is a representative experiment of three biological replicates. (C) *Tbp2*^{-/-} females do not ovulate upon hormonal stimulation. Control and *Tbp2*^{-/-} 4- to 6-wk-old females were induced for superovulation with intraperitoneal injection of PMSG followed by administration of hCG 48 h later. The number of ovulated oocytes was determined after dissection for each female. The average number of oocytes ovulated per female is indicated. (D) Increased number of immature follicles and abnormal oocytes in *Tbp2*^{-/-} females. Examples of follicles dissected from control and *Tbp2*^{-/-} ovaries observed under DIC microscopy. While GV stage (preovulatory) oocytes are easily distinguishable in control females, *Tbp2*^{-/-} oocytes contain mainly immature follicles or oocytes of abnormal morphology and devoid of zona pellucida. (E) Gross morphology of ovaries from 6-wk-old control and *Tbp2*^{-/-} females. While control ovaries show clear mature follicles (arrows), such structures are absent in the *Tbp2*^{-/-} ovaries.



Tbp2^{-/-} males were fertile and produced normal litters (Table 2). This female sterility phenotype is consistent with the ovary-restricted expression of TBP2 in the mouse (Bartfai et al. 2004; Xiao et al. 2006; Gazdag et al. 2007), and demonstrates that TBP2 is absolutely required for female fertility.

In order to get insight into the mechanisms underlying the ovarian failure elicited by TBP2 loss, we first assessed whether the sterility phenotype could be overridden by hormonal stimulation. We superovulated *Tbp2*^{+/+} and *Tbp2*^{-/-} female littermates and determined the number of ovulated oocytes per female. While all control females responded to hormonal stimulation and ovulated mature oocytes (17 oocytes per female, SD = 14, n = 5), we did not

find any ovulated oocytes in *Tbp2*^{-/-} females (Fig. 1C). In agreement with this, superovulated *Tbp2*^{-/-} females never gave birth (data not shown). Dissection of control and *Tbp2*^{-/-} ovaries revealed an increased number of immature follicles and abnormal oocytes (Fig. 1D, panel b). In contrast, oocytes dissected from *Tbp2*^{+/+} littermates included mature, dictyate GV stage oocytes in antral follicles, indicative of normal ovarian physiology (Fig. 1D, panel a). These observations suggest that TBP2-null females are devoid of mature fully grown oocytes.

To elucidate the defect in oogenesis of *Tbp2*^{-/-} mice, we performed detailed ovarian morphological analyses. Ovaries from null and wild-type 2-wk-old females (e.g., prepuberty) were of similar size and morphological

Table 2. Females lacking TBP2 are sterile

Genotype	Mated with	Number of mice analyzed	Total number of productive pregnancies	Average litter size
Female				
+/+	Wild-type B6 male	3	9	8
+/-	Wild-type B6 male	4	10	7.9
-/-	Wild-type B6 male	6	0	0
Male				
+/+	Wild-type B6 female	4	7	7.3
+/-	Wild-type B6 female	6	12	6.8
-/-	Wild-type B6 female	5	13	7.3

Number of productive pregnancies and litter size of females and males of the indicated genotype mated with either male or female C57/Bl6, accordingly. Mating period analyzed covered 6 mo except for the TBP2^{-/-} females, where it spanned >10 mo.

appearance (data not shown). However, ovaries of 6-wk-old mutant mice were significantly smaller than wild-type ones (Fig. 1E). Moreover, while mature follicles are clearly visible in control ovaries (Fig. 1E, panel a, arrow), ovaries from *Tbp2*^{-/-} females lack such structures (Fig. 1E, panel b).

We next analyzed the histology of 2-wk-old and 6-wk-old ovaries in *Tbp2*^{-/-} mice. In 2-wk-old *Tbp2*^{-/-} ovaries, follicle development takes place up to the secondary follicle stage, and this appeared to occur normally when compared with wild-type ovaries (Fig. 2A, panels a–d). However, a detailed analysis of the number of oocytes in each stage on individual sections revealed an increased number of oocytes that persisted in the primordial follicle stage and a reduced number of oocytes at the secondary follicle stage (Fig. 2C). This presumably reflects the limited recruitment of follicles for growth and results in reduced numbers of preantral and antral follicles (Fig. 2A,C). Adult ovaries from *Tbp2*^{-/-} females were remarkably different from the control ovaries (Fig. 2A, panels e–h). At 6 wk of age there was a gradual depletion of

follicles in *Tbp2*^{-/-} ovaries and abnormal oocyte development became pronounced (Fig. 2A, panels e–h). Null ovaries showed greater numbers of small follicles in the cortex that were delayed in their development and that consisted mainly of abnormal secondary follicles (Fig. 2A). The oocytes within the secondary follicles with antrum showed signs of degeneration and irregularly shaped nuclei (Fig. 2A,B). These oocytes lost their symmetrical rounded shape and the zona pellucida was absent (Fig. 2B, panel h, arrow). Phenotypical changes upon TBP2 loss were further evident in the number of small (late primary and secondary) follicles containing irregularly shaped oocytes and the emergence of cyst-containing (empty) structures as well as interstitial cell proliferation. Anovular small, empty follicles with disarranged granulosa cells were also frequent in the mutant compared with the wild type, and no signs of corpus luteum (CL) formation were detected (Fig. 2A,B). Overall, adult *Tbp2*^{-/-} ovaries showed a strong decrease in secondary follicles (Fig. 2C), suggesting that folliculogenesis is impaired before or

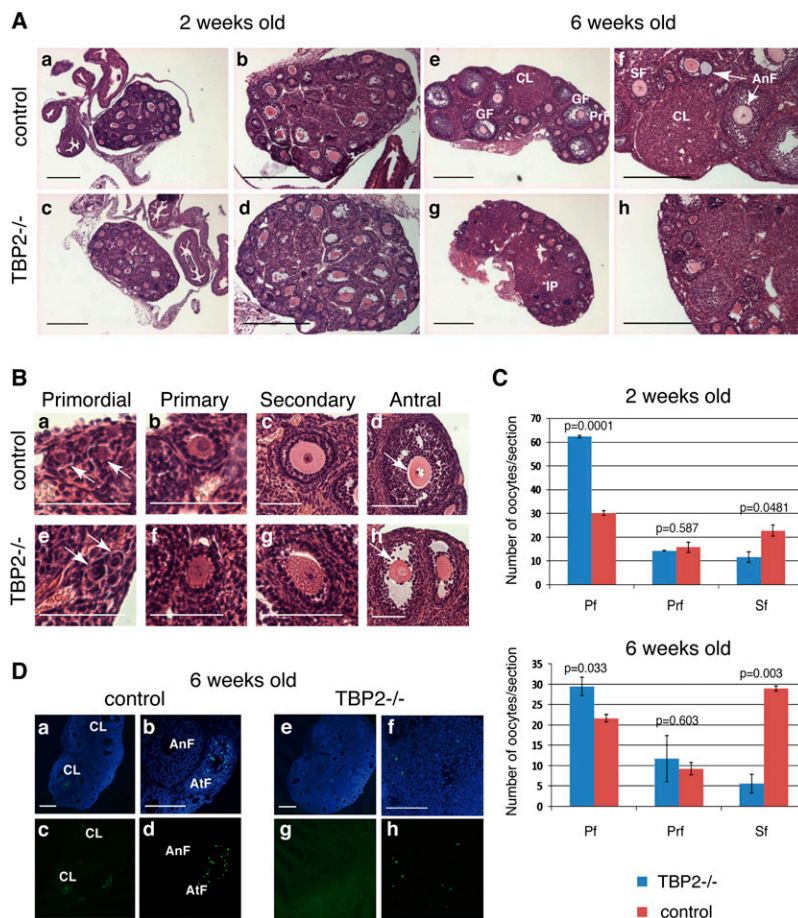


Figure 2. TBP2 loss results in ovarian failure due to defective folliculogenesis. (A–C) Histological analysis of ovaries from *Tbp2*^{-/-} mice reveals a block in folliculogenesis around the primary and secondary follicular stage. In A, sections from 2-wk-old and 6-wk-old control (panels a–b,e–f) and *Tbp2*^{-/-} (panels c–d,g–h) ovaries stained with hematoxylin-eosin are shown at two different magnifications. In control 6-wk-old ovaries, primary (Prf), secondary (SF), and antral follicles (AnF) as well as corpus lutei (CL) are indicated. *Tbp2*^{-/-} oocytes lack antral secondary follicles and large graafian follicles; instead, interstitial proliferation (IP) is visible in panel g. Data are representative of three biological replicates. Bar, 100 μm. (B) Higher magnifications of oocytes at different stages of follicular development from 6-wk-old control (panels a–d) and *Tbp2*^{-/-} (panels e–h) ovaries. (Panels a,e) Arrows point to two different primordial follicles, which appear normal when compared with control ovaries. Primary and secondary follicles can be found in null ovaries, albeit in reduced numbers (see C). (Panel h) Following the secondary follicle stage, the few *Tbp2*^{-/-} oocytes that survive fail to undergo antrum formation and instead show disorganized follicular structure. Note the absence of zona pellucida in these oocytes as seen from the lack of regular white space surrounding the oocyte (arrow). (C) Increased number of underdeveloped oocytes in *Tbp2*^{-/-} females. Analysis of oocyte counts comparing *Tbp2*^{-/-} and control 2-wk-old and 6-wk-old ovaries. Oocytes within primordial (PF), primary (PrF), and secondary (SF) follicles were counted on ovarian sections. Numbers represent

average of counts of three sequential sections from serially sectioned ovaries. Antral oocytes are only present in the controls and were therefore not included in the counts. *P* values were calculated using an unpaired *t*-test for triplicates. (D) TUNEL assay on sections from control and *Tbp2*^{-/-} ovaries. Apoptotic nuclei were labeled using a modified TUNEL protocol with fluorescein detection. DNA was stained with DAPI. In 6-wk-old control sections, an atretic follicle (AtF) and nuclei within the CL show normal signs of apoptosis. A healthy antral follicle (AnF) is also depicted. Representative merge (panels a,b,e,f) and green (panels c,d,g,h) channel images of lower (panels a,c,e,g) and higher (panels b,d,f,h) magnifications are shown.

during secondary follicle stages. Given the above-described degeneration of oocytes, we next looked for apoptosis in *Tbp2*^{-/-} ovaries by TUNEL analysis. We did not detect increased apoptosis in ovaries from *Tbp2*-null females at 2 wk of age (data not shown). In control 6-wk-old ovaries, we observed apoptotic granulosa cells in ovulated follicles that can be distinguishable because of the presence of the CL (Fig. 2D). In contrast, these apoptotic granulosa cells were missing in the *Tbp2*^{-/-} mice (Fig. 2D, panels e,f), further demonstrating that TBP2-deficient oocytes do not reach a fully grown ovulatory stage. These data indicate that TBP2 is essential for proper oocyte growth and follicular development and its absence leads to infertility.

TBP2 is essential for transcription and correct chromatin organization in the developing oocytes

Because TBP2 is able to mediate transcription initiation in vitro (Bartfai et al. 2004; Deato et al. 2008), we asked whether TBP2 can regulate transcription in vivo. To this end, we examined the transcriptional competence of *Tbp2*^{-/-} oocytes by visualizing the Ser2 phosphorylation state of the C-terminal domain (CTD) of the largest subunit of Pol II, an indicator of Pol II activity (Bellier et al. 1997). Phosphorylation of the CTD is highest at the primary and preantral stages of folliculogenesis. It then decreases gradually until it is almost undetectable at the fully grown preovulatory stage, when the oocyte is transcriptionally silent (Bachvarova 1985). Primordial follicles displayed similar levels of CTD phosphorylation in control and *Tbp2*^{-/-} mice (Fig. 3A). However, reduction in the Ser2 phosphorylation of the CTD was observed in TBP2-deficient oocytes from 6-wk-old females, mainly at the primary follicle stage (Fig. 3A). We next asked whether the reduction of Ser2 phosphorylation of the Pol II CTD was due to a reduction in the levels of Pol II protein by using an antibody that recognizes all Pol II CTD repeats. No reduction in the levels of Pol II CTD was detected at any of the stages of oocyte development analyzed, when comparing TBP2-deficient and control oocytes (Fig. 3B). Therefore, despite the fact that Pol II levels remain stable, it appears that transcription is perturbed mainly at the primary follicle stage as a result of TBP2 loss, at a stage where wild-type oocytes exhibit extensive transcriptional activity (Bachvarova 1985).

As another marker of active transcription, we examined global levels of histone H3K4 trimethylation (H3K4me3) in *Tbp2*-null ovaries. Consistent with the reduction of levels of Pol II phospho-Ser2, oocytes lacking TBP2 showed diminished levels of H3K4me3 from the primary follicle stage onward (Fig. 3C). This observation further indicates that the transcription-related chromatin organization is impaired in *Tbp2*^{-/-} oocytes. While control oocytes showed the normal chromatin maturation process from the so-called nonsurrounded nucleolus to a surrounded nucleolus chromatin configuration during folliculogenesis (Zuccotti et al. 2005), *Tbp2*^{-/-} oocytes displayed a disorganized chromatin (Fig. 3D). Instead of forming one prominent nucleolar-like body,

dispersed small DAPI-rich regions were formed (Fig. 3D), indicating that from the primary follicle stage the chromatin organization of the mutant oocytes is abnormal. Hence, these data indicate that TBP2 is required for correct chromatin organization.

TBP2 loss leads to altered transcriptional profile of oocyte-specific genes

To identify genes misregulated in *Tbp2*^{-/-} ovaries (see the Discussion), we performed a transcriptome analysis from 2-wk-old females. At the very stringent false discovery rate (FDR) of 0.12 we identified 156 misregulated probes, which correspond to 153 genes, among which half (76) were up-regulated and half (77) were down-regulated upon TBP2 loss (Fig. 4A; Supplemental Tables S1, S2). When a standard FDR value of 0.27 was used, a greater number of misregulated genes was identified, with 744 probes, equivalent to 714 genes (Fig. 4A). Thus, as expected, the number of probes increases as the FDR stringency is lowered. Using the 0.27 FDR, we identified more genes in the up-regulated class as compared with the down-regulated one (Fig. 4A). However, at any given FDR value, we found more genes with high negative fold changes in the down-regulated class than in the up-regulated one (Fig. 4A), suggesting that the down-regulated genes may be direct TBP2 target genes.

We next validated our microarray analysis by randomly choosing seven up-regulated and seven down-regulated genes that included genes with both a high and a low fold change and an FDR ranging between 0.12 and 0.27 and assessed their expression by RT-qPCR (Fig. 4B). All genes analyzed within this FDR range behaved as predicted by the microarray (Fig. 4B). We next asked whether misregulated genes in TBP2-depleted ovaries are involved in specific functional categories. We used the 153 genes with an FDR of 0.12 for this analysis. Among the down-regulated genes, the most enriched functional categories were reproductive and developmental processes, tissue development, and cellular growth and proliferation (Fig. 4C, left). The significant enrichment of genes ($P = 0.002$) belonging to the reproductive system category is in full agreement with an oocyte-specific role of TBP2 and the sterility phenotype of *Tbp2*-null females. The up-regulated class of genes also shows enrichment in various developmental and tissue differentiation pathways, although the statistical significance is lower ($P = 0.049$). Among these, the hematological system is highly overrepresented, suggesting impaired or delayed vascular development in *Tbp2*^{-/-} ovaries.

We further compared misregulated genes in *Tbp2*^{-/-} ovaries with an oocyte-specific set of genes that we built from the Novartis Tissue Expression Atlas (<http://expression.gnf.org>). This set contained genes expressed in oocytes and genes coding for maternally inherited transcripts in the cytoplasm of the fertilized zygote. We compared this oocyte-specific gene list, which contains 851 genes, to the up-regulated and down-regulated genes identified in *Tbp2*^{-/-} ovaries (Fig. 4D). Intersection of the down-regulated genes revealed that 29% (23 out of 77) of the genes

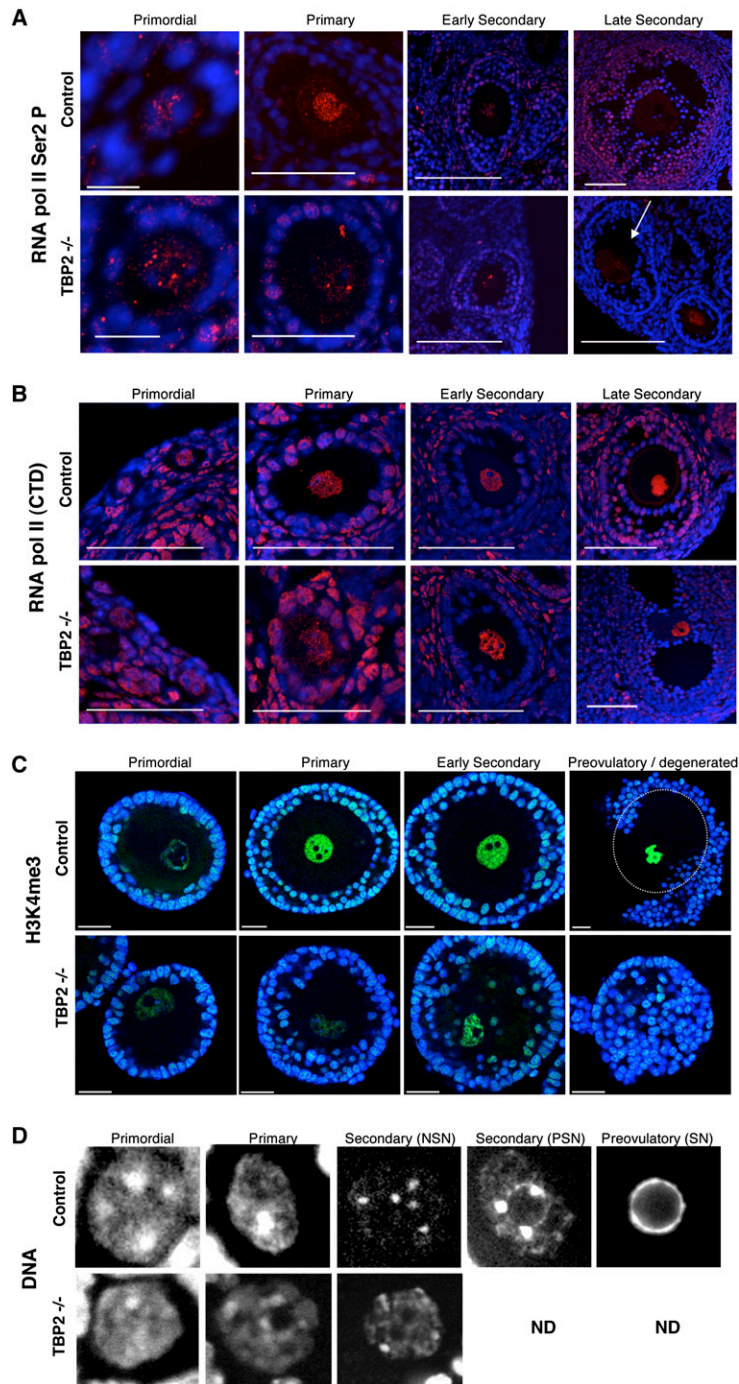


Figure 3. RNA Pol II activity and chromatin structure are altered in oocytes lacking TBP2. (*A,B*) Immunofluorescence analysis of control and *Tbp2*^{-/-} ovarian sections from 6-wk-old females using antibodies recognizing the Ser2-phosphorylated CTD of Pol II (*A*) or all forms of the CTD of Pol II (*B*). Pol II staining is shown in red. DNA (blue) was counterstained with DAPI. Follicular stages are indicated. (*A*) The arrow points to the follicle at the stage indicated at the *top*. Bars: for primordial stages, 10 μm ; for later stages, 50 μm . (*B*) Bar, 50 μm . (*C*) *Tbp2*^{-/-} oocytes show reduced H3K4me3 levels from the primary follicular stage. Immunostaining analysis for H3K4me3 in oocytes at different stages of follicular development from 6-wk-old females. DNA (blue) was stained with DAPI. Bar, 10 μm . Confocal acquisition was done using identical parameters to allow comparison. Hence, H3K4me3 levels in preovulatory stage oocyte in the control appear saturated. Note that *Tbp2*^{-/-} oocytes degenerate after the secondary follicular stage, as seen from the presence of follicular cells within the follicle. The dashed white line demarcates the oocyte membrane. (*D*) Altered chromatin configuration in *Tbp2*^{-/-} oocytes. Sections from control and *Tbp2*^{-/-} mice were stained with DAPI and analyzed under confocal microscopy. (NSN) Nonsurrounded nucleolus; (PSN) partially surrounded nucleolus; (SN) surrounded nucleolus. Representative sections are shown. (ND) Not determined.

are specific to the oocyte ($P = 1.34e-15$), and may thus be direct targets of TBP2 (Supplemental Table S3). In contrast, the up-regulated category did not show a significant enrichment for oocyte-specific genes ($P = 0.17$) (Fig. 4D).

We next aimed to identify putative common signatures within the regulatory elements of the misregulated genes. We searched for common transcription factor-binding motifs in the regions spanning from -200 to $+50$ base pairs (bp) relative to their transcription start sites (TSSs) using pScan (Zambelli et al. 2009). This analysis revealed that promoters of the down-regulated genes are signifi-

cantly enriched in AT-rich binding motifs including the TATA box. This pattern was not found in the up-regulated promoters, for which no significant motif appeared. The nucleotide composition between the two classes of promoters was also significantly different, with promoters of genes positively regulated by TBP2 displaying higher AT content, in particular in the -30 , $+1$, and $+60$ to $+80$ regions relative to the TSS (Fig. 4E). Thus, the two groups of genes can be distinguished at the promoter level on the basis of their composition, suggesting different mechanisms of transcriptional regulation for

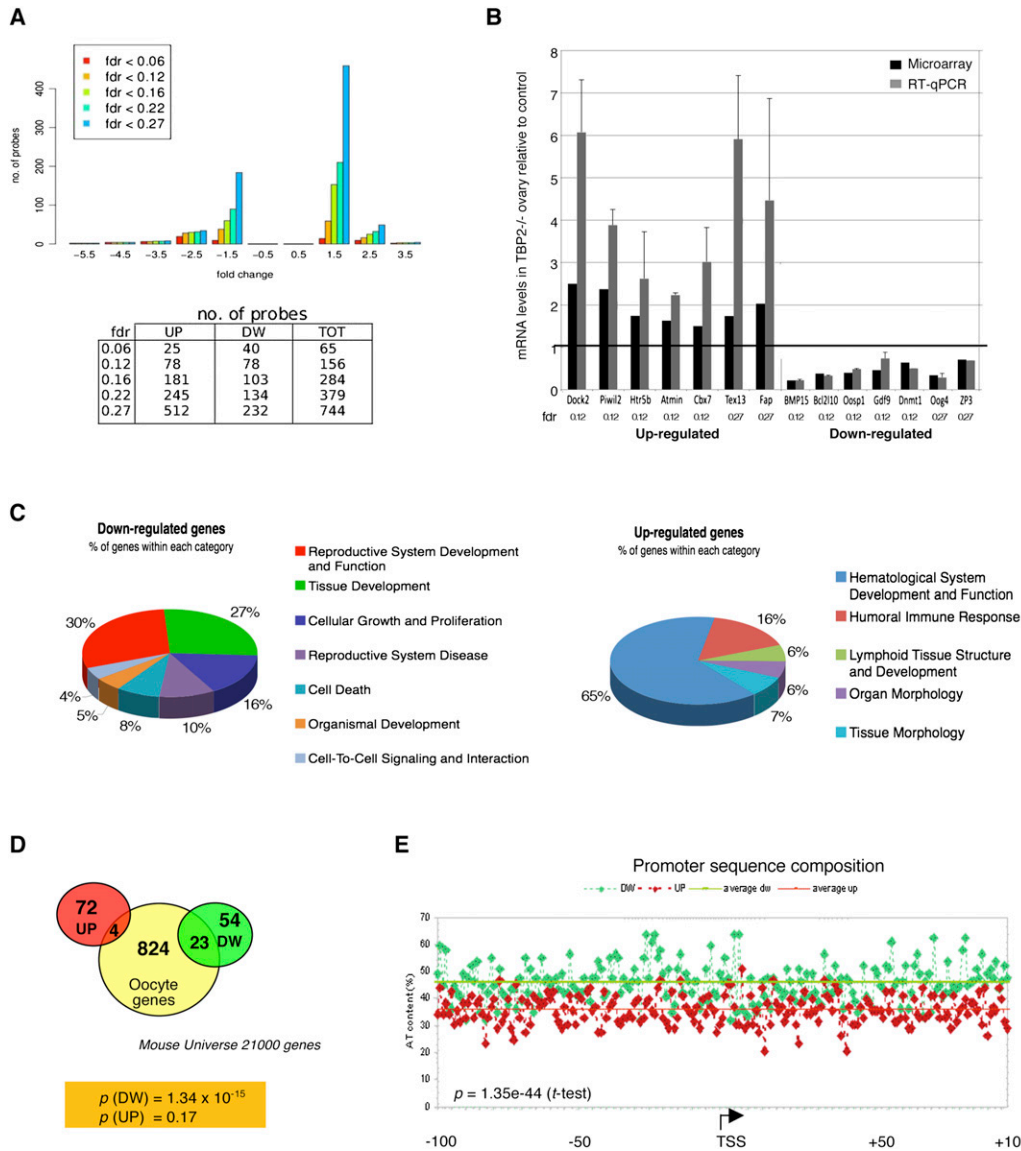


Figure 4. TBP2 ablation results in altered oocyte-specific transcriptional profile in 2-wk-old ovaries. (A) Distribution of the number of probes on the microarray with respect to their FDR and fold change in 2-wk-old the *Tbp2*^{-/-} ovaries compared with the control. (Top panel) In the histogram, bars of different colors represent the number of probes differentially hybridized at the indicated FDR threshold. (Bottom panel) The exact number of probes is displayed in the table. At FDR = 0.12 (yellow bars), there are as many probes in the down-regulated class (negative fold change) as there are in the up-regulated one (positive fold change). (B) Microarray validation in 2-wk-old *Tbp2*^{-/-} and *Tbp2*^{+/+} ovaries. Up-regulated and down-regulated genes with different FDR and fold change values (ranging from -0.22-fold to 2.5-fold) were randomly selected and their mRNA levels were analyzed by RT-qPCR. mRNA levels are shown relative to 18S RNA and are expressed relative to the control, for which mRNA levels were set at 1 (black line). Microarray (black bars) and RT-qPCR (gray bars) fold changes are average of three independent biological replicates, respectively. Error bars are SD. (C) Most enriched functional categories among the down-regulated and up-regulated genes according to Ingenuity Pathways. (D) Genes down-regulated upon TBP2 loss are enriched in oocyte-specific genes. Comparison of the up-regulated and down-regulated genes with an FDR of 0.12 with genes that are specifically expressed in the oocyte. P values for the intersection of the up-regulated and down-regulated genes are shown at the bottom. Note that only the intersection of down-regulated genes is statistically significant. (E) Sequence composition of the promoters belonging to the two classes of genes according to their nucleotide content. The percentage of AT content of the up-regulated (red) and down-regulated (green) gene promoters within a region from -100 to +100 bp around TSS is shown. The average AT content for the two promoter populations is also indicated. Comparison of the two groups of genes using a Student's *t*-test indicates that the two populations of promoters differ significantly. The TSS positions follow the current NCBI annotation.

up-regulated and down-regulated genes. Taken together, our data demonstrate that ablation of TBP2 results in major changes in the ovary transcriptome, a significant

number of which are genes specifically expressed in the oocyte, substantiating an oocyte-specific transcriptional regulatory function of TBP2.

TBP2 binds promoters of actively transcribed genes in the oocytes

Because TBP2 is required for the correct expression of genes in the oocyte, we next asked whether endogenous TBP2 directly regulates such genes. To this end, we tested whether TBP2 binds to the promoter of transcribed genes in the oocyte. We analyzed the promoters of genes encoding two secreted growth factors (*Bmp15* and *Gdf9*), *Zp3*, and two transcription factors (*Oct4/Pou5f1* and *Nobox*). These genes are expressed to different extents in the oocyte (Supplemental Fig. S2) and were either down-regulated (*Bmp15*, *Gdf9*, and *Zp3*) or up-regulated (*Oct4* and *Nobox*) in *Tbp2*^{-/-} ovaries (Supplemental Tables S1, S2). Furthermore, because histone genes are highly transcribed in the oocyte, we also included histone H2B promoter in our analysis. We performed chromatin immunoprecipitation (ChIP) with ovaries from 2-wk-old females, which contain mainly oocytes at early stages of folliculogenesis (Kerr et al. 2006), when TBP2 accumulation is highest (Gazdag et al. 2007). We used primers encompassing the TSS for each gene and two primer pairs for an intergenic region (IR) as negative control. We detected a good and reproducible enrichment of TBP2 at *Zp3* and the *H2bc* promoters (Fig. 5A). The *Gdf9* promoter displayed significant, albeit weak, binding of TBP2. ChIP with a H3K4me3 antibody revealed a good enrichment of H3K4me3 at the promoter regions of *H2bc*, *Zp3*, *Gdf9*, *Oct4*, and *Nobox*, but not on the *Bmp15* promoter, further indicating that these genes are actively transcribed in the oocyte (Fig. 5B). Thus, our data demonstrate

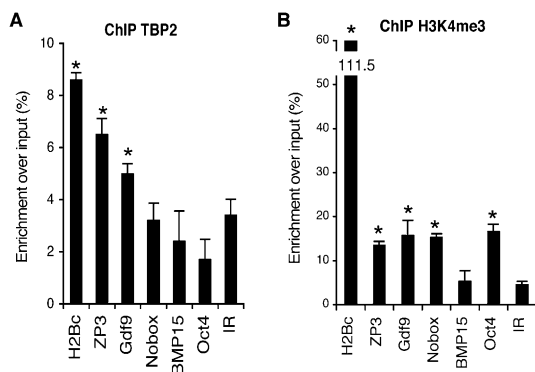


Figure 5. TBP2 binds to promoters of active genes in the oocyte. (A–B) ChIP using anti-TBP2 or anti-H3K4me3 antibodies as indicated and promoter regions of oocyte-specific genes were analyzed by real-time qPCR. Data is represented as the enrichment over the input (in percentage). IRs are shown as negative controls. ChIP with a mouse or rabbit IgG antibody showed no enrichment above background levels. Shown are representative results of at least three independent experiments. Asterisk (*) indicates significant enrichment ($P \leq 0.05$). Note that although *Bmp15* mRNA accumulation in the oocyte is high (Supplemental Fig. S2), H3K4me3 levels on its promoter are not significant, suggesting that H3K4me3 regulation on *Bmp15* promoter is different or that *Bmp15* mRNA is relatively stable and that actual rate of active transcription of this gene is low.

that endogenous TBP2 is able to bind to promoters of genes that are actively transcribed in the oocyte; in particular, to those showing a high level of expression and whose expression is severely down-regulated upon TBP2 loss. These data, together with the transcriptome analysis, indicate that TBP2 plays an essential and direct role in mediating oocyte-specific transcription during folliculogenesis.

TBP is dispensable for oocyte maturation, zygotic genome activation, and subsequent development

We established that TBP2 is a functional oocyte transcription factor in vivo. The distribution of TBP2 and TBP at the protein level is mutually exclusive during oocyte development, TBP2 being predominant in the oocyte during folliculogenesis (Gazdag et al. 2007). Thus, we next asked whether loss of TBP has an effect on oocyte development and subsequent fertilization. Depletion of TBP in the oocyte was achieved by a genetic approach using a *Zp3*-Cre transgenic line (Fig. 6A). Males homozygous for the *Zp3*-Cre transgene (de Vries et al. 2000) were mated with females homozygous for the floxed *Tbp* allele (*Tbp*^{lox/lox}) (Martianov et al. 2002b). *Zp3* promoter-driven Cre expression is restricted to the oocyte, thereby resulting in maternal depletion of floxed *Tbp* allele. As a result of Cre recombination, 50% of the haploid oocytes ovulated from the females derived from these crosses will inherit the deleted *Tbp* allele and 50% will inherit the wild-type allele (Fig. 6A). In order to analyze whether inheritance of *Tbp* in the oocyte is required for ovulation and/or embryonic development, we mated females issued from *Tbp*^{lox/lox}; *Zp3*-Cre^{tg/tg} crosses with wild-type males. As a reciprocal control, littermate males issued from the same *Tbp*^{lox/lox}; *Zp3*-Cre^{tg/tg} crosses were mated with wild-type females (Fig. 6A). We first asked whether *Tbp*⁻ haploid oocytes could be fertilized in normal ratios: We observed no difference between the oocytes carrying *Tbp*⁻ or *Tbp*⁺ (wild-type) alleles (Fig. 6B). We further assessed whether embryos inheriting the maternal *Tbp*⁻ allele develop until the blastocyst stage in the same proportions as those inheriting the *Tbp*⁺ allele. Both groups of embryos developed equally well and with the same timing (Fig. 6B–C). Thus, developmental progression was unaffected upon loss of the maternal *Tbp* allele. Finally, we analyzed the number of pups born that carried the maternal *Tbp*⁻ allele versus those carrying the maternal *Tbp*⁺ allele (Fig. 6B). Pups that inherited the *Tbp*-deleted allele from their mothers were born at 60% frequency, whereas those inheriting the *Tbp*⁺ wild-type allele were found at 40%. We verified that TBP protein was efficiently depleted by analyzing zygotes immediately after fertilization by immunostaining using an anti-TBP antibody. As shown in Figure 6D, TBP was undetectable in the pronuclei of 50% of the zygotes analyzed ($n = 9$). Single embryo genotyping following immunostaining confirmed that zygotes displaying no TBP staining were *Tbp*^{-/+}, and hence inherited the deleted allele from their mother (Fig. 6D). Collectively, these data suggest that, in contrast to TBP2, maternal

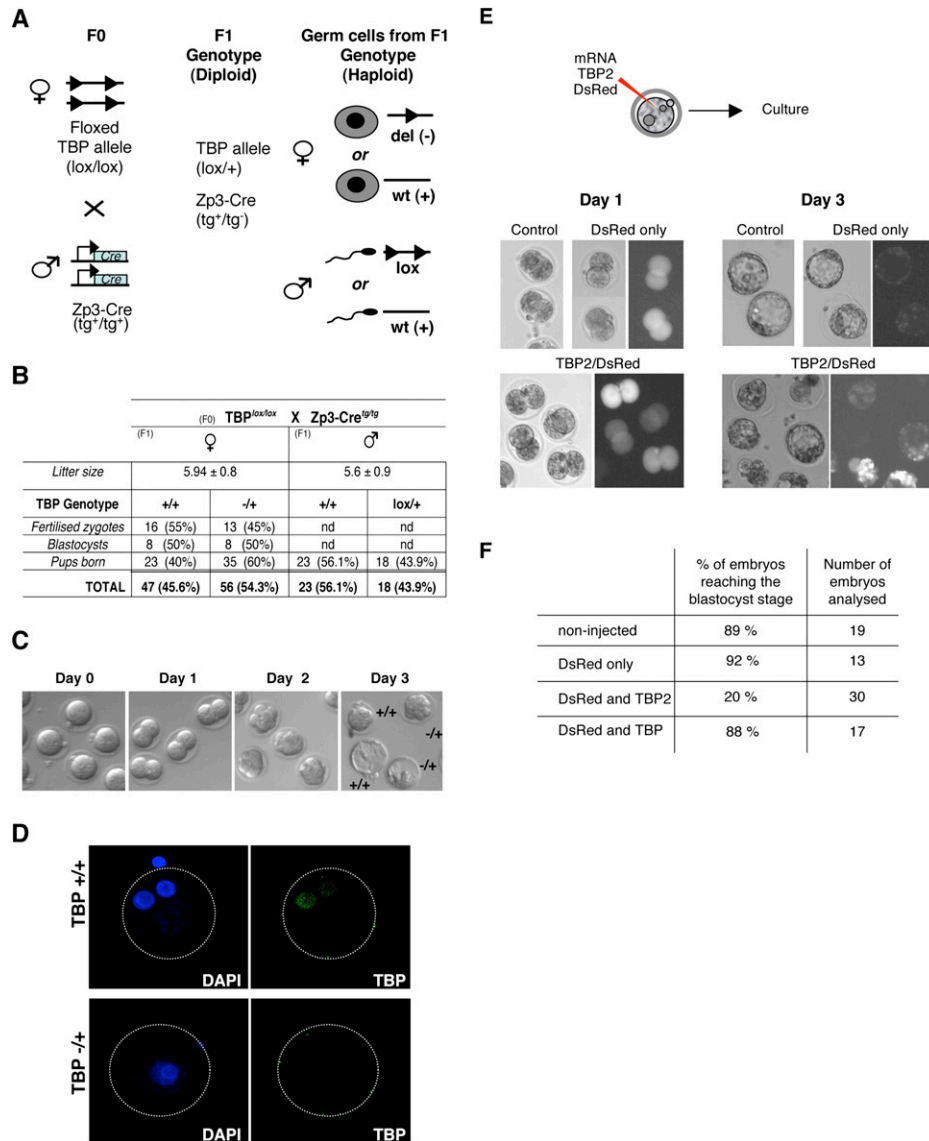


Figure 6. TBP is dispensable for oocyte maturation, ovulation, and subsequent fertilization. (A) Genetic approach used to delete TBP in the maternal germline. Homozygous females for the floxed *Tbp* allele were mated to homozygous males for the *ZP3-Cre* transgene. Females express Cre in their oocytes. Haploid genotype of 50% of their oocytes is thus *Tbp*⁻ (deleted) and the other 50% have a *Tbp*⁺ (wild-type) genotype. To determine the competence of these oocytes, females were mated with wild-type males as indicated. Male littermates of the F1 were used as controls. (B) Table summarizing the results obtained from the crosses between transgenic F1 females (*Tbp*^{lox/+}; *ZP3-Cre*^{tg/+}) or males and their respective wild-type partner. (nd) Not determined. (C) Oocytes carrying the maternal *Tbp*⁻ (deleted) allele can be fertilized in vivo and develop normally until the blastocyst stage. Zygotes derived from F1(*Tbp*^{lox/+}; *ZP3-Cre*^{tg/+}) females mated with wild-type males were collected at fertilization and were cultured until the blastocyst stage. Single embryo genotyping was performed subsequently. (D) Efficient deletion of TBP in oocytes ovulated from F1(*Tbp*^{lox/+}; *ZP3-Cre*^{tg/+}) females upon Cre expression driven from the *Zp3* promoter. Zygotes were collected from F1 X wild-type male crosses immediately after fertilization, stained with anti-TBP antibody (3G3), and analyzed using confocal microscopy. Individual embryos were genotyped after acquisition. Shown are representative embryos derived from oocytes inheriting the deleted *Tbp* allele maternally (*Tbp*^{-/+}) and those inheriting the *Tbp* wild-type allele (*Tbp*^{+/+}). Note that in the *Tbp*^{+/-} embryo, the two pronuclei mask each other. (E) TBP2 overexpression results in altered embryonic development and in reduced cell proliferation. Experimental design for TBP2 overexpression is shown on the top. Zygotes were microinjected with mRNA for DsRed alone (control) or in combination with mRNA for TBP2 and cultured until control embryos reached the blastocyst stage. (F) Zygotes overexpressing TBP2 cleave to the two-cell stage normally, but show reduced development until the blastocyst stage.

TBP is not required for oocyte maturation, fertilization, and subsequent development. More generally, our results show that, in oocytes, TBP2 is the major factor mediating

Pol II transcription during female germ cell development, and that TBP is dispensable in this process during oocyte maturation.

TBP2 overexpression has a negative effect on progression of embryonic development and cell growth

We established that, in primary and secondary follicular stage oocytes, TBP2 is the main TBP-type factor required for Pol II transcription initiation. However, TBP2 protein levels decline upon ovulation to become undetectable after fertilization (Gazdag et al. 2007). The mutually exclusive roles described for TBP2 and TBP in the oocytes are the opposite following fertilization, when only TBP becomes abundant concomitant with zygotic genome activation. Thus, to test whether TBP2 can function in the zygote in parallel with TBP, we used a gain-of-function approach. We microinjected zygotes with mRNA for TBP2 together with mRNA of the fluorescent DsRed protein as a marker for injection. As controls, we used noninjected embryos, embryos injected with mRNA for TBP and DsRed, and embryos injected with DsRed mRNA only (Fig. 6E,F). Injection of TBP and TBP2 mRNA led to efficient protein translation (data not shown). Zygotes overexpressing TBP2 cleaved to the two-cell stage normally and at the same time as the control groups (Fig. 6E, day 1). Noninjected, TBP + DsRed, and DsRed-only embryos reached the blastocyst stage after 3 d. In contrast, TBP2-expressing embryos displayed a low development rate: Only 20% reached the blastocyst stage (Fig. 6E [day 3], F). Among these, 40% ceased development at the two-cell stage, 6.6% ceased development at the two- to four-cell stage, and 33% ceased development between the four- and eight-cell stage (data not shown). Thus, misexpression of TBP2 in the embryo is deleterious to development, further suggesting that its role is restricted to the highly specialized female germ cell transcription program.

Discussion

TBP2 is an oocyte-specific TBP-type factor

Recent data support the idea that an alternative form of transcription initiation machinery has evolved to control the expression of genes required for differentiation and development (Deato and Tjian 2007; Jones 2007; Muller et al. 2007; Reina and Hernandez 2007). Determining the importance of homolog transcription factors in a given species in mediating initiation of transcription by Pol II at different developmental stages and in different cell types is central to understanding how transcription is regulated in a cell-specific context. The two core promoter-binding factors TBP2 and TBP share an almost identical DNA-binding domain and differ in their N termini. Although TBP2 has been shown to interact with TFIIA and TFIIB, to bind a consensus TATA box, and to mediate transcription initiation in vitro, the role of TBP2 in mammalian cells has been only partially addressed. Here we show that (1) TBP2-deficient mice are viable, but females are sterile; (2) TBP2 is the main essential TBP-type transcription factor during follicular female germ cell differentiation; (3) TBP is not required during the later stages of folliculogenesis and subsequent fertilization; (4) TBP2 inhibits cell proliferation when misexpressed in cells that express TBP;

and (5) TBP2 loss results in major changes in transcription profile, demonstrating the ability of TBP2 to regulate gene expression globally. Finally we show that in vivo TBP2 is able to bind to promoters of genes that are expressed in the oocyte, suggesting that TBP2 is directly involved in the regulation of oocyte-specific gene expression. Thus, our data indicate for the first time that TBP2 is a central basal transcriptional regulator that drives specialized processes and establishes cell type-specific features in the oocyte.

TBP2 regulates chromatin organization and oocyte-specific gene expression pathways

We found that TBP2 is able to bind directly to certain genes transcribed in the oocyte. Consistent with a positive regulatory function of TBP2 in transcription initiation, oocyte development does not proceed beyond the secondary follicle stage in females lacking TBP2. Indeed, the TBP2-dependent genes we identified begin to be transcribed in primary follicles (Supplemental Fig. S3). This suggests that in oocytes at the primary and early secondary stages, TBP2-dependent Pol II transcription is essential to reach the transcriptionally quiescent preovulatory follicle stage (Bachvarova 1985) in which the chromatin has reached a maximum degree of condensation (Zuccotti et al. 2005). Indeed, we found that TBP2 loss results in disorganization of the chromatin in nuclei of oocytes and that in the TBP2^{-/-} oocytes chromatin never reaches a fully condensed status. These observations together suggest that TBP2 affects both transcription of specific genes and global chromatin organization. In this sense, note that mouse TLF, which is essential for spermatogenesis, has been shown to participate in the organization of the chromocenter during spermiogenesis and is therefore conceived as a factor with a dual role: as a classical transcription factor and as a structural factor (Martianov et al. 2002a). Whether TBP2 directly regulates these chromatin condensation steps or acts on a particular set of genes involved in chromatin remodeling remains to be investigated. However, it is remarkable that genes important for proper chromatin configuration such as the oocyte-specific linker histone H1, *H1foo*, and the DNA methyltransferase *Dnmt1* are down-regulated in *Tbp2*^{-/-} ovaries.

The identified set of misregulated genes upon TBP2 loss is in full agreement with the sterility phenotype that we report. Among these there are several genes (i.e., *Gdf9*, *Zp3*, *Bmp15*, *Rfpl4*, *H1foo*, *Zar1*, and *Dnmt1*) (see Supplemental Table S1) that are necessary for oocyte development and function. This group of misregulated genes in the *Tbp2*^{-/-} ovary is strikingly similar to those identified in mice lacking other oocyte-specific transcription factors such as NOBOX or FIGLA. Mice lacking either of these genes display a complete female sterility phenotype and reduced expression of *Gdf9*, *Bmp15*, *Rfpl4*, *H1foo*, *Zar1*, and *Dnmt1o* (Rajkovic et al. 2004; Joshi et al. 2007). Within the ovary, the oocyte has a central role in regulating granulosa cell metabolism (Su et al. 2009). The full complement of GDF9 and BMP15 is necessary for follicular function and female fertility

(Yan et al. 2001; Su et al. 2004). We observed reduced expression of *Gdf9* and *Bmp15* upon TBP2 loss, which suggests that oocyte-granulosa cell communication is altered in TBP2-deficient ovaries (Yan et al. 2001; Su et al. 2004). Indeed, we observed down-regulation of *Kif23*, which is a *Gdf9* downstream target and is involved in the formation of the intercellular bridges in the ovary (Greenbaum et al. 2009), further suggesting that GDF9 pathway is indeed seriously affected.

We performed the microarray analysis on whole ovaries because we sought to identify any possible effect of TBP2 loss on the communication between the oocyte and the follicular cells. However, this type of analysis may mask higher fold changes of gene expression in the oocytes, specially if a gene is expressed also in other cells of the ovary. Nevertheless, the robustness of our microarray analysis is supported by the fact that we identified many oocyte-specific genes and also genes involved in the oocyte/granulosa cell communication. Moreover, we also observe that (1) the down-regulated class of genes is enriched in both oocyte-specific genes (tissue specificity analysis) and functions related to developmental pathways (ontology analysis); (2) the promoters of the two classes of genes differ significantly in terms of their AT percentage content (promoter analysis), which is noticeably higher in the down-regulated gene pool; and (3) some of these down-regulated genes turned out to be TBP2 direct targets as determined by ChIP. Thus, our results together indicate that the folliculogenesis program of the oocyte is severely dysregulated in the absence of TBP2.

TBP2, but not TBP, is an essential transcription factor in the developing oocyte

We showed previously that TBP is not expressed during oocyte development from the primary follicular stage (Gazdag et al. 2007). In agreement, we show that deletion of the *Tbp* gene in oocytes, where TBP is not detected, does not affect oocyte maturation and subsequent fertilization. Why is it important to replace TBP function with that of TBP2 in the developing oocytes? While TBP has a positive role in cell proliferation in somatic cells (Johnson et al. 2003), we found that TBP2 has a negative effect on proliferation in the early embryo. Thus, it is possible that TBP2, in contrast to TBP, has acquired a highly specialized role in promoting cell growth, but not cell proliferation. This is one of the main features of the oocytes, which are among the largest cells with a single nucleus in vertebrates, and therefore require a long growing period without cell division. TBP2 could then be able to block (directly or indirectly) the cell division program, but at the same time promote the transcription of genes whose products accumulate in the growing oocyte.

TBP2/TRF3 has been suggested to play a role in orchestrating muscle differentiation in an ex vivo model (Deato and Tjian 2007). However, we find that, in vivo, TBP2 expression is undetectable in tissues other than the ovary or in other cell differentiation models (Supplemental Fig. S4A,B). Consistent with this restricted pattern of TBP2

expression, skeletal muscle in *Tbp2*^{-/-} mice is undistinguishable from that of control animals, and the expression of key myogenic factors involved in muscle differentiation and integrity is unaffected (Supplemental Fig. S4C,D).

Because of the high identity between the core domain of TBP and TBP2, it is conceivable that TBP2 can potentially assemble with the same TAFs as TBP. The high divergence on the N terminus between TBP2 and TBP raises the interesting question of whether and how the N-terminal domain of TBP2 can diversify its function from that of TBP. One possibility would be that the N-terminal domain of TBP2 could determine the association of TBP2 with TAFs or TAF-like proteins. Similar to *Drosophila* germ cells (Hiller et al. 2004; Chen et al. 2005), tissue-specific TAF homologs may be expressed in mouse oocytes, thereby allowing association of TBP2 with a TFIID-like complex.

Interestingly, the genes that require TBP2 for their transcription in oocytes have a high frequency of AT nucleotides around their TSSs. The high homology between the DNA-binding core domains of TBP2 and TBP and our results showing that TBP2 can bind to oocyte-specific genes further suggest that TBP2 could recognize AT-rich sequences that may resemble the canonical TATA box.

Groups of genes clustered near the telomeres and specifically expressed in the oocyte have been identified recently (Hamatani et al. 2008). Importantly, our microarray analysis revealed an overrepresentation of genes within such clusters (e.g., *Obox1*, *Obox2*, *Obox5*, *Obox3*, *Nlrp9c*, *Nlrp9b*, *Nlrp4a*, and *Nlrp4c*) among the down-regulated genes in *Tbp2*^{-/-} ovaries (Supplemental Table S1). Analysis of the core promoter sequences of the oocyte-specific homeobox cluster *Obox* revealed a consensus for a TATA box, further suggesting that TBP2 can bind directly to TATA boxes and regulate these genes.

Specialized TBP-type factors as mediators of germline-specific transcription

Our results indicate that TBP2 is essential for female germ cell differentiation in mice. In a mirroring situation, differentiation of male germ cells in the mouse requires another TBP-type factor: TLF. Indeed, absence of TLF results in increased germ cell apoptosis and male infertility (Martianov et al. 2001). Why have different TBP-type factors such as TBP2 and TLF acquired such a specialized germ cell-specific function? The beginning of TBP2 accumulation in the oocyte nuclei or TLF accumulation in male germ cell nuclei coincides with the phase of meiosis I. The abundance of TBP2 in oocytes during meiosis I therefore suggests that it may be involved in the control of expression of genes in a meiotic context. Interestingly, the murine TLF was not detected in mitotically dividing spermatogonia. Thus, contrary to TBP2 in oocytes, and to TLF in forming spermatocytes, TBP expression is largely restricted to mitotically dividing cells and drops below the detection limit in primary germ cells that are not proliferating. Interestingly, loss of TBP2 resulted in the induction of expression of several genes

that are usually only expressed in the male gonad (*Tex11*, *Tex13*, *Tex14*, *Tex19*, *Piwil2*, and *Tekt2*) (Wang et al. 2001). This suggests that TBP2 may have a role in repressing the spermatogenesis transcriptional program in the female germline. These findings together further lead to the concept that in the germ cells of opposite sexes transcription initiation, with either TBP2 in oocytes or with TLF in spermatocytes, has evolved to carry out specialized functions, and that these respective functions cannot be performed by other TBP-type factors.

Finally, it will be important to determine whether TBP2 is linked to cases of female infertility in humans. Increased infertility in humans has attracted considerable attention in recent years, with ~20% of couples of reproductive age being infertile. Thus, the identification of new animal models and molecules involved in sterility phenotypes, such as TBP2, will aid in tackling the genetic and cellular base of human infertility.

Materials and methods

Generation of TBP2 mice

The *Tbp2/Trf3* conditional mutant mouse line was generated at the Mouse Clinical Institute/Institut Clinique de la Souris (<http://www.mci.u-strasbg.fr>). The targeting vector was constructed to obtain a floxed fragment comprising exon 4, and was verified by sequencing. After electroporation into 129S2/SvPas ES cells, recombined clones were analyzed by Southern blot with Neo and 5' and 3' external probes. A positive ES clone was microinjected into C57BL/6J blastocysts, and male chimaeras with germline transmission were obtained. Genotyping was performed with tail DNA. To verify TBP2 mRNA levels in *Tbp2^{-/-}* and *Tbp2^{+/-}* mice, total RNA was extracted from ovaries of 2-wk-old females with Trizol (Invitrogen), reverse-transcribed, and analyzed by PCR using specific primers for TBP2. PCR products were transferred onto an Amersham Hybond-XL membrane and probed with a radioactive probe specific for TBP2. Autoradiography was quantified using SynGene. ZP3-Cre mice were obtained from the Jackson laboratory. TBP^{lox/lox} mice have been described previously (Martianov et al. 2002b).

Embryo culture and microinjection

Zygotes were microinjected with 1–2 pL of a mixture containing 300 ng/μL mRNA for mouse TBP2 or TBP and 200 ng/μL mRNA for DsRed, which were capped and transcribed in vitro from the pRN3P plasmid. Embryos were cultured in KSOM medium under a 5% CO₂ atmosphere at 37°C and their development was monitored daily.

Immunostaining of cells and embryos

Oocytes or embryos were washed and fixed as described (Torres-Padilla et al. 2006), and incubated with the TBP (1:800) (Gazdag et al. 2007) or the H3K4me3 (Abcam) antibodies and the corresponding Alexa488-coupled secondary antibody (Molecular Probes). Embryos were mounted in Vectashield with DAPI (Vector Laboratories) and analyzed under confocal microscopy on a Leica SP2 AOBs confocal microscope.

Western blot

Ovaries from *Tbp2^{+/+}*, *Tbp2^{+/-}*, or *Tbp2^{-/-}* 2-wk-old females were homogenized in 20 μL of loading blue containing SDS and

separated by 8% PAGE. Western blotting was performed using the α-TBP2 2B12 monoclonal antibody, and α-tubulin was used as loading control.

Histology and immunohistochemistry

Ovaries from *Tbp2^{-/-}* or control 2-wk-old and 6-wk-old females were fixed in 4% paraformaldehyde, dehydrated in ethanol, and embedded in paraffin. Sections (5 μm thick) were analyzed using the following antibodies: mouse monoclonal anti-TBP2 2B12 antibody (Gazdag et al. 2007), anti-CTD antibody 7G5 (kind gift from M. Vigneron), and H5 clone recognizing RNA Pol II phosphorylated at Ser2 (Covance). Sections were mounted in Vectashield with DAPI (Vector Laboratories) and analyzed using confocal microscopy.

ChIP and real-time PCR

Ovaries from 12- to 15-d-old CD1 females were fixed in 1% formaldehyde for 15 min and subsequently homogenized in 50 mM sodium phosphate buffer (pH 7.5), 10% glycerol, 10 mM β-mercaptoethanol, 0.1 mM phenylmethylsulfonyl fluoride, and protease inhibitors with a pestle. Fifteen micrograms of DNA were used for each immunoprecipitation with the mouse anti-TBP2 2B12 or the H3K4me3 antibodies (Abcam). Immunoprecipitated chromatin samples were washed and eluted from the antibody/beads complex by adding 0.1% SDS and 0.1 M NaHCO₃. DNA was purified after cross-linking reversal, and the enrichment of specific DNA sequences in the immunoprecipitated samples (bound fractions) compared with the input samples (input fractions) was quantified by qPCR using SYBR Green JumpStart Taq ReadyMix (Sigma). Primer sequences are available upon request. The enrichment of each sequence is given by the ratio of the bound to input content (B/I).

Microarray analysis and RT-qPCR

RNA extraction from ovaries of 2-wk-old females was performed with the PicoPure RNA Kit (Arcturus). Total mRNA was amplified in two rounds, and three independent biological replicates were hybridized with Cy3 or Cy5 dyes onto the Affymetrix GeneChip Mouse Gene 1.0 ST microarrays (<http://www-microarrays.u-strasbg.fr>). Misregulated genes were analyzed using RMA (robust multiarray) normalization. Statistical analysis was performed with the Limma package within R-Bioconductor and the empirical Bayes method. Microarray data were validated in at least two independent biological replicates by real-time qPCR using SYBR Green JumpStart Taq ReadyMix (Sigma). PCR reactions were carried out in triplicate. To ensure proper biological cross-comparison, expression levels were compared between littermates.

Acknowledgments

We are grateful to I. Davidson and I. Martianov for the *Tbp^{lox/lox}* mice; to M. Boeglin and D. Hensch for advice; to M. Vigneron for antibodies; to A. Fadloun, F. Müller, and B. Murphy for critical reading of the manuscript; and to A. Buj-Bello for advice. The TBP2 mutant mice were generated at the Mouse Clinical Institute (MCI/ICS) in the Targeted Mutagenesis and Transgenesis Department. E.G. was supported by a fellowship from the European Community (HPRN-CT-00504228). This work was supported by funds from CNRS, INSERM, Université de Strasbourg, FRM, ANR (05-BLAN-0396-01; Regulome), and European Community (HPRN-CT 00504228 and STREP LSHG-CT-2004-502950;

EUTRACC LSHG-CT-2007-037445) to L.T., and PNRRE/INSERM number P006427 and AVENIR to M.E.T.P.

References

- Andreu-Vieyra C, Lin YN, Matzuk MM. 2006. Mining the oocyte transcriptome. *Trends Endocrinol Metab* **17**: 136–143.
- Bachvarova R. 1985. Gene expression during oogenesis and oocyte development in mammals. *Dev Biol* **1**: 453–524.
- Bartfai R, Balduf C, Hilton T, Rathmann Y, Hadzhiev Y, Tora L, Orban L, Muller F. 2004. TBP2, a vertebrate-specific member of the TBP family, is required in embryonic development of zebrafish. *Curr Biol* **14**: 593–598.
- Bellier S, Chastant S, Adenot P, Vincent M, Renard JP, Bensaude O. 1997. Nuclear translocation and carboxyl-terminal domain phosphorylation of RNA polymerase II delineate the two phases of zygotic gene activation in mammalian embryos. *EMBO J* **16**: 6250–6262.
- Bleil JD, Wassarman PM. 1980. Structure and function of the zona pellucida: Identification and characterization of the proteins of the mouse oocyte's zona pellucida. *Dev Biol* **76**: 185–202.
- Burley SK, Roeder RG. 1996. Biochemistry and structural biology of transcription factor IID (TFIID). *Annu Rev Biochem* **65**: 769–799.
- Chen X, Hiller M, Sancak Y, Fuller MT. 2005. Tissue-specific TAFs counteract Polycomb to turn on terminal differentiation. *Science* **310**: 869–872.
- Choi Y, Rajkovic A. 2006. Characterization of NOBOX DNA binding specificity and its regulation of Gdf9 and Pou5f1 promoters. *J Biol Chem* **281**: 35747–35756.
- Deato MD, Tjian R. 2007. Switching of the core transcription machinery during myogenesis. *Genes & Dev* **21**: 2137–2149.
- Deato MD, Marr MT, Sottero T, Inouye C, Hu P, Tjian R. 2008. MyoD targets TAF3/TRF3 to activate myogenin transcription. *Mol Cell* **32**: 96–105.
- de Vries WN, Binns LT, Fancher KS, Dean J, Moore R, Kemler R, Knowles BB. 2000. Expression of Cre recombinase in mouse oocytes: A means to study maternal effect genes. *Genesis* **26**: 110–112.
- Ferg M, Sanges R, Gehrig J, Kiss J, Bauer M, Lovas A, Szabo M, Yang L, Straehle U, Pankratz MJ, et al. 2007. The TATA-binding protein regulates maternal mRNA degradation and differential zygotic transcription in zebrafish. *EMBO J* **26**: 3945–3956.
- Gazdag E, Rajkovic A, Torres-Padilla ME, Tora L. 2007. Analysis of TATA-binding protein 2 (TBP2) and TBP expression suggests different roles for the two proteins in regulation of gene expression during oogenesis and early mouse development. *Reproduction* **134**: 51–62.
- Greenbaum MP, Iwamori N, Agno JE, Matzuk MM. 2009. Mouse TEX14 is required for embryonic germ cell intercellular bridges but not female fertility. *Biol Reprod* **80**: 449–457.
- Hamatani T, Yamada M, Akutsu H, Kuji N, Mochimaru Y, Takano M, Toyoda M, Miyado K, Umezawa A, Yoshimura Y. 2008. What can we learn from gene expression profiling of mouse oocytes? *Reproduction* **135**: 581–592.
- Hernandez N. 1993. TBP, a universal eukaryotic transcription factor? *Genes & Dev* **7**: 1291–1308.
- Hiller M, Chen X, Pringle MJ, Suchorolski M, Sancak Y, Viswanathan S, Bolival B, Lin TY, Marino S, Fuller MT. 2004. Testis-specific TAF homologs collaborate to control a tissue-specific transcription program. *Development* **131**: 5297–5308.
- Jacobi UG, Akkers RC, Pierson ES, Weeks DL, Dagle JM, Veenstra GJ. 2007. TBP paralogs accommodate metazoan and vertebrate-specific developmental gene regulation. *EMBO J* **26**: 3900–3909.
- Jallow Z, Jacobi UG, Weeks DL, Dawid IB, Veenstra GJ. 2004. Specialized and redundant roles of TBP and a vertebrate-specific TBP paralog in embryonic gene regulation in *Xenopus*. *Proc Natl Acad Sci* **101**: 13525–13530.
- Johnson SA, Dubeau L, Kawalek M, Dervan A, Schonthal AH, Dang CV, Johnson DL. 2003. Increased expression of TATA-binding protein, the central transcription factor, can contribute to oncogenesis. *Mol Cell Biol* **23**: 3043–3051.
- Jones KA. 2007. Transcription strategies in terminally differentiated cells: Shaken to the core. *Genes & Dev* **21**: 2113–2117.
- Joshi S, Davies H, Sims LP, Levy SE, Dean J. 2007. Ovarian gene expression in the absence of FIGLA, an oocyte-specific transcription factor. *BMC Dev Biol* **7**: 67. doi: 10.1186/1471-213X-7-67.
- Kerr JB, Duckett R, Myers M, Britt KL, Mladenovska T, Findlay JK. 2006. Quantification of healthy follicles in the neonatal and adult mouse ovary: Evidence for maintenance of primordial follicle supply. *Reproduction* **132**: 95–109.
- Martianov I, Fimia GM, Dierich A, Parvinen M, Sassone-Corsi P, Davidson I. 2001. Late arrest of spermiogenesis and germ cell apoptosis in mice lacking the TBP-like TLF/TRF2 gene. *Mol Cell* **7**: 509–515.
- Martianov I, Brancorsini S, Gansmuller A, Parvinen M, Davidson I, Sassone-Corsi P. 2002a. Distinct functions of TBP and TLF/TRF2 during spermatogenesis: Requirement of TLF for heterochromatic chromocenter formation in haploid round spermatids. *Development* **129**: 945–955.
- Martianov I, Viville S, Davidson I. 2002b. RNA polymerase II transcription in murine cells lacking the TATA binding protein. *Science* **298**: 1036–1039.
- Matzuk MM, Lamb DJ. 2002. Genetic dissection of mammalian fertility pathways. *Nat Cell Biol* **4**: S41–S49. doi: 10.1038/ncb-nm-fertility41.
- Muller F, Tora L. 2004. The multicoloured world of promoter recognition complexes. *EMBO J* **23**: 2–8.
- Muller F, Lakatos L, Dantonel J, Strahle U, Tora L. 2001. TBP is not universally required for zygotic RNA polymerase II transcription in zebrafish. *Curr Biol* **11**: 282–287.
- Muller F, Demeny MA, Tora L. 2007. New problems in RNA polymerase II transcription initiation: matching the diversity of core promoters with a variety of promoter recognition factors. *J Biol Chem* **282**: 14685–14689.
- Pedersen T, Peters H. 1968. Proposal for a classification of oocytes and follicles in the mouse ovary. *J Reprod Fertil* **17**: 555–557.
- Rajkovic A, Pangas SA, Ballow D, Suzumori N, Matzuk MM. 2004. NOBOX deficiency disrupts early folliculogenesis and oocyte-specific gene expression. *Science* **305**: 1157–1159.
- Reina JH, Hernandez N. 2007. On a roll for new TRF targets. *Genes & Dev* **21**: 2855–2860.
- Su YQ, Wu X, O'Brien MJ, Pendola FL, Denegre JN, Matzuk MM, Eppig JJ. 2004. Synergistic roles of BMP15 and GDF9 in the development and function of the oocyte-cumulus cell complex in mice: Genetic evidence for an oocyte-granulosa cell regulatory loop. *Dev Biol* **276**: 64–73.
- Su YQ, Sugiura K, Eppig JJ. 2009. Mouse oocyte control of granulosa cell development and function: Paracrine regulation of cumulus cell metabolism. *Semin Reprod Med* **27**: 32–42.
- Takada S, Lis JT, Zhou S, Tjian R. 2000. A TRF1:BRF complex directs *Drosophila* RNA polymerase III transcription. *Cell* **101**: 459–469.
- Torres-Padilla ME, Bannister AJ, Hurd PJ, Kouzarides T, Zernicka-Goetz M. 2006. Dynamic distribution of the replacement

- histone variant H3.3 in the mouse oocyte and preimplantation embryos. *Int J Dev Biol* **50**: 455–461.
- Veenstra GJ, Weeks DL, Wolffe AP. 2000. Distinct roles for TBP and TBP-like factor in early embryonic gene transcription in *Xenopus*. *Science* **290**: 2312–2315.
- Wang PJ, McCarrey JR, Yang F, Page DC. 2001. An abundance of X-linked genes expressed in spermatogonia. *Nat Genet* **27**: 422–426.
- White JK, Auerbach W, Duyao MP, Vonsattel JP, Gusella JF, Joyner AL, MacDonald ME. 1997. Huntingtin is required for neurogenesis and is not impaired by the Huntington's disease CAG expansion. *Nat Genet* **17**: 404–410.
- Xiao L, Kim M, DeJong J. 2006. Developmental and cell type-specific regulation of core promoter transcription factors in germ cells of frogs and mice. *Gene Expr Patterns* **6**: 409–419.
- Yan C, Wang P, DeMayo J, DeMayo FJ, Elvin JA, Carino C, Prasad SV, Skinner SS, Dunbar BS, Dube JL, et al. 2001. Synergistic roles of bone morphogenetic protein 15 and growth differentiation factor 9 in ovarian function. *Mol Endocrinol* **15**: 854–866.
- Zambelli F, Pesole G, Pavesi G. 2009. Pscan: Finding over-represented transcription factor binding sites motifs in sequences of co-regulated and co-expressed genes. *Nucleic Acids Res* **37**: W247–W252. doi: 10.1093/nar/gkp464.
- Zuccotti M, Garagna S, Merico V, Monti M, Alberto Redi C. 2005. Chromatin organisation and nuclear architecture in growing mouse oocytes. *Mol Cell Endocrinol* **234**: 11–17.

Supplemental Figures and Tables

TBP2 is essential for germ cell development by regulating transcription and chromatin condensation in the oocyte

**Emese Gazdag^{1,4}, Angèle Santenard^{1,2,4}, Céline Ziegler-Birling^{1,2}, Gioia Altobelli^{1,3},
Olivier Poch³, László Tora^{1*} and Maria-Elena Torres-Padilla^{1,2*}**

¹Department of Functional Genomics, ²Department of Developmental and Cell Biology and ³Bioinformatics and Integrative Biology Laboratory, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), UMR 7104 CNRS, Uds, INSERM U964, BP 10142, F-67404 ILLKIRCH Cedex, CU de Strasbourg, France.

⁴ Equal first authors

*Corresponding authors

laszlo@igbmc.u-strasbg.fr

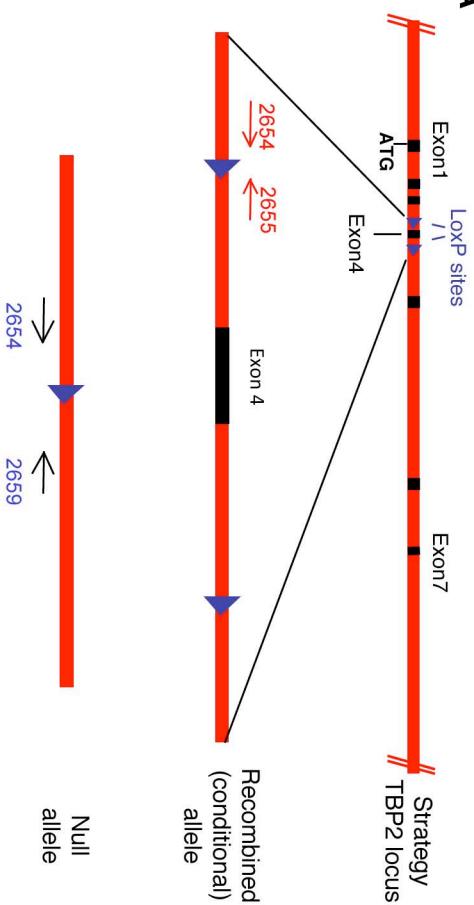
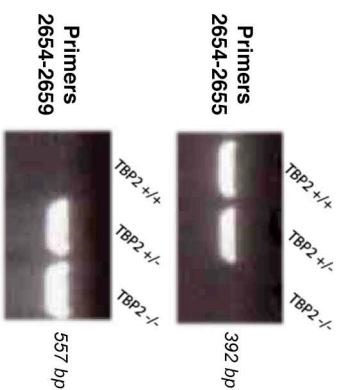
Tel. + 33(0) 3 88 65 34 44

Fax. + 33(0) 3 88 65 32 01

metp@igbmc.fr

Tel. + 33(0) 3 88 65 33 60

Fax. + 33(0) 3 88 65 32 01

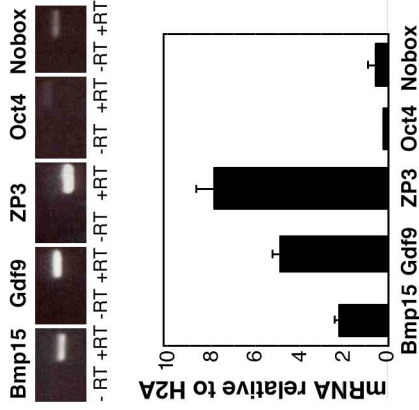
A**B**

Supplementary Figure 1.

Strategy for disruption of the *Tbp2* gene by homologous recombination.

A. Diagram of the mouse *Tbp2* gene. The position of the LoxP sites surrounding exon 4 is shown (blue arrowheads). Representation of the null allele upon deletion of exon 4 (bottom). Positions of the primers used for the genotyping by PCR is indicated.

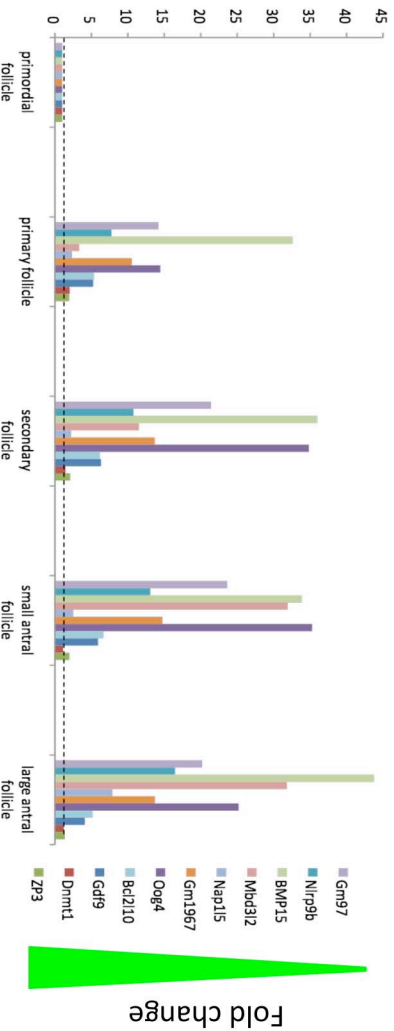
B. Genotyping results of using the primers indicated in (A) for *Tbp2*^{+/+}, *Tbp2*^{+/-} and *Tbp2*^{-/-} mice. The corresponding genotype of each mouse is indicated above each lane.



Supplementary Figure 2.

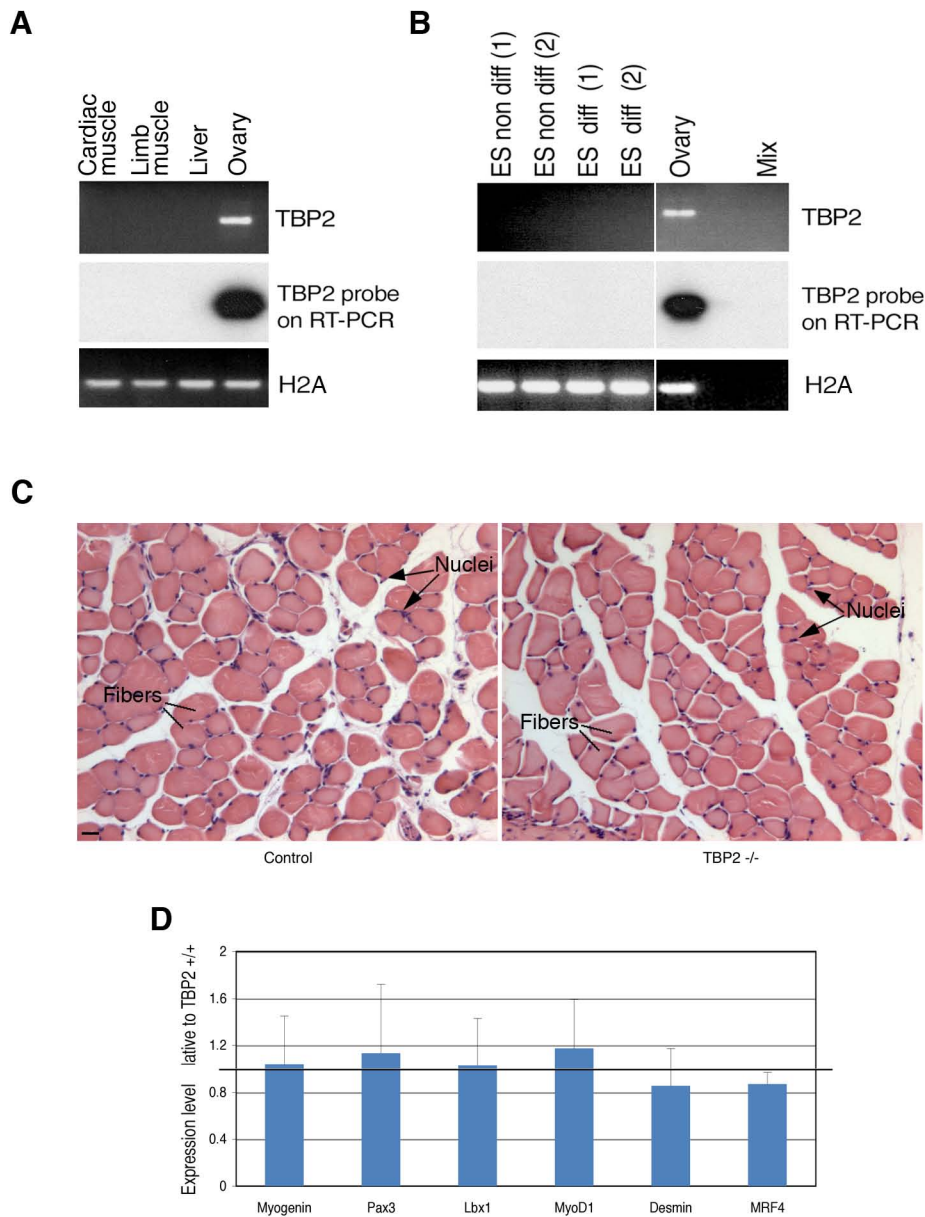
Expression levels of genes analysed by ChIP in wild type ovaries.

Total RNA was extracted from 2-weeks old ovaries and processed by RT-qPCR. cDNA was analysed with specific primers for indicated genes and analysed by electrophoresis (top). Real-Time PCR analysis of mRNA levels of each of the genes indicated relative to histone H2A mRNA are shown in the graph. Shown is a representative of at least two independent experiments.



Supplementary figure 3.

Genes down-regulated upon TBP2 loss are those that start being expressed from the primary and secondary follicle stages. Expression pattern of down-regulated genes in *Tbp2*^{-/-} ovaries throughout folliculogenesis in wild type follicles. Eleven down-regulated genes upon TBP2 loss were randomly chosen from the microarray analysis. The expression pattern during folliculogenesis of these genes was collected from Gene Expression Omnibus (GEO) Profile on the NCBI website (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=gco>). mRNA levels are expressed relatively to their mRNA level in primordial follicles. The results were plotted on a graph and the genes were ordered in accordance to their fold change on the array as shown on the right.



Supplementary figure 4.

A-B. Expression of *Tbp2* in cardiac muscle, limb skeletal muscle, liver, ovary and ES cells. Total RNA from the indicated tissues was prepared, retrotranscribed and analysed with primers specific for *Tbp2*. To validate the identity of the amplified PCR product and to increase the sensitivity of the PCR, products were transferred onto a membrane and probed with a radiolabelled *Tbp2*-specific probe, which does not crossreact with *Tbp*. H2A was used as internal control. For ES cells, two replicates (1 and 2) for either non-differentiated or differentiated cells were used.

C. Analysis of skeletal muscle from control and *Tbp2*^{-/-} mice. Transverse sections from 6-week-old mice were prepared from the tibialis anterior muscle and stained with Hematoxylin-Eosin. Shown are sections representative of two biological replicates. Note that muscle fibers are equally distributed in both *Tbp2*^{-/-} and control mice and that no signs of necrosis or inflammation are observed. Fibers and their nuclei are indicated.

D. Quantitative RT-PCR analysis of muscle specific genes in adult *Tbp2*^{-/-} and control skeletal muscle. Total RNA was prepared from tibialis anterior muscle from either control or *Tbp2*^{-/-} mice, retrotranscribed and analysed by Q-PCR with primers for the indicated genes. Data are expressed as relative to the expression in control muscle and normalised to *Gapdh* mRNA levels. Data are representative of two biological replicates and PCRs done in triplicate. Genes analysed include genes required for skeletal muscle determination and acquisition of myoblast precursor fate (*MyoD1*, *MRF4*), for terminal skeletal muscle differentiation (*Myogenin*, *MyoD1*, *MRF4*), for the regulation of the entry of cells into the myogenic program (*Pax3*), for the survival and delamination of muscle progenitor cells (*Lbx1*) and for muscle fiber integrity (*Desmin*).

Supplementary Table I

List of genes down-regulated in *Tbp2*^{-/-} ovaries.

affyID	Gene.Accession	Gene.Symbol	mRNA.Accession	fdr	FC
1,1E+07	NM_194058	Nlrp9b	NM_194058	0,0015	-5,342
1,1E+07	NM_001008428	Fbxw12	NM_001008428	0,0015	-4,245
1,1E+07	NM_009757	Bmp15	NM_009757	0,0029	-4,369
1E+07	NM_178763	Zfp750	NM_178763	0,0067	-2,293
1E+07	BC139082	E330034G19Rik	BC139082	0,0067	-3,42
1E+07	NM_001033283	Gm97	NM_001033283	0,0067	-5,67
1,1E+07	NM_029662	Mfsd2	NM_029662	0,0067	-2,549
1,1E+07	BC099944	E330021D16Rik	BC099944	0,0067	-4,843
1,1E+07	NM_013479	Bcl2l10	NM_013479	0,0067	-2,606
1,1E+07	NM_177069	E330009P21Rik	NM_177069	0,0067	-4,914
1E+07	NM_177471	Ccdc69	NM_177471	0,0153	-2,532
1E+07	BC145922	Gm813	BC145922	0,0153	-3,281
1E+07	NM_001033452	Gm1967	NM_001033452	0,0153	-3,457
1,1E+07	NM_173773	Oog4	NM_173773	0,0153	-2,737
1,1E+07	BC147090	Gm839	BC147090	0,0153	-2,298
1,1E+07	NM_015793	Fbxw14	NM_015793	0,0153	-2,858
1,1E+07	NM_199036	Fbxw15	NM_199036	0,0153	-3,747
1E+07	NM_010267	Gdap1	NM_010267	0,0279	-1,796
1E+07	NM_145977	Slc45a3	NM_145977	0,0279	-2,22
1E+07	NM_008110	Gdf9	NM_008110	0,0279	-2,154
1E+07	NM_009580	Zp1	NM_009580	0,0279	-2,265
1,1E+07	NM_138954	Rfpl4	NM_138954	0,0279	-2,327
1,1E+07	NM_177598	Fbxw13	NM_177598	0,0279	-3,723
1E+07	NM_080844	Serpinc1	NM_080844	0,0379	-2,094
1E+07	NM_026251	Patl2	NM_026251	0,0379	-2,04
1,1E+07	BC147526	Gm1965	BC147526	0,0379	-1,762
1,1E+07	NM_008872	Plat	NM_008872	0,0379	-1,99
1,1E+07	BC132114	EG382109	BC132114	0,0379	-2,601
1E+07	---	---	---	0,0571	-1,859
1E+07	DQ279479	Eif4e1b	DQ279479	0,0571	-1,741
1E+07	BC094241	D10Wsu102e	BC094241	0,0571	-1,65
1E+07	NM_013773	Tcl1b1	NM_013773	0,0571	-2,046
1E+07	NM_010701	Lect1	NM_010701	0,0571	-1,771
1E+07	ENSMUST00000100435	ENSMUSG00000075545	ENSMUST00000100435	0,0571	-2,187
1,1E+07	BC053719	C87499	BC053719	0,0571	-2,528
1,1E+07	ENSMUST00000101176	Gm1965	ENSMUST00000101176	0,0571	-1,879
1,1E+07	NM_145709	Obox5	NM_145709	0,0571	-2,023
1,1E+07	NM_010066	Dnmt1	NM_010066	0,0571	-1,591
1,1E+07	NM_177703	Fbxw19	NM_177703	0,0571	-2,574
1,1E+07	BC139081	C85627	BC139081	0,0571	-3,683
1E+07	NM_145856	Il17f	NM_145856	0,0897	-1,577
1E+07	NR_003564	OTTMUSG00000001305	NR_003564	0,0897	-1,863
1E+07	NM_001024135	Kbtbd7	NM_001024135	0,0897	-1,609
1E+07	NM_019977	Miox	NM_019977	0,0897	-2,118
1E+07	BC137768	Dppa2	BC137768	0,0897	-2,094
1E+07	NM_133353	Oosp1	NM_133353	0,0897	-2,336
1E+07	NM_016889	Insm1	NM_016889	0,0897	-1,496
1,1E+07	---	---	ENSMUST00000083804	0,0897	-1,663
1,1E+07	NM_138311	H1foo	NM_138311	0,0897	-1,958
1,1E+07	NM_001004762	Pla2g4c	NM_001004762	0,0897	-1,826
1,1E+07	ENSMUST00000098469	ENSMUSG00000074134	ENSMUST00000098469	0,0897	-2,239
1,1E+07	NM_001007462	Pabpn1	NM_001007462	0,0897	-1,896
1,1E+07	NM_025274	Dppa5a	NM_025274	0,0897	-1,765
1E+07	NM_001113187	Khdc1b	NM_001113187	0,0975	-1,92
1E+07	NR_003564	OTTMUSG00000001305	NR_003564	0,0975	-1,62
1E+07	NM_199299	Phf15	NM_199299	0,0975	-1,651
1E+07	NM_013775	Tcl1b2	NM_013775	0,0975	-2,042

1E+07	NM_013774	Tcl1b4	NM_013774	0,0975	-1,952
1E+07	NM_011877	Ptpn21	NM_011877	0,0975	-1,436
1E+07	NM_029001	Elovl7	NM_029001	0,0975	-2,033
1E+07	ENSMUST00000100166	ENSMUSG00000075396	ENSMUST00000100166	0,0975	-1,755
1,1E+07	NM_001033226	Calr4	NM_001033226	0,0975	-2,299
1,1E+07	NM_172710	2310045A20Rik	NM_172710	0,0975	-1,754
1,1E+07	NM_177561	Usp46	NM_177561	0,0975	-1,518
1,1E+07	---	---	ENSMUST00000113394	0,0975	-1,556
1,1E+07	NM_018791	Zfp108	NM_018791	0,0975	-1,412
1,1E+07	---	---	AF357388	0,0975	-1,552
1,1E+07	NR_001277	Snord33	NR_001277	0,0975	-1,655
1,1E+07	NM_007555	Bmp5	NM_007555	0,0975	-2,164
1E+07	ENSMUST00000094774	Sp110 // MGC118608	ENSMUST00000094774	0,1186	-1,576
1E+07	NM_146250	Gpr1	NM_146250	0,1186	-1,747
1E+07	NM_001033269	Eif4e1b	NM_001033269	0,1186	-1,697
1E+07	NM_001034882	EG432879	NM_001034882	0,1186	-1,639
1E+07	NM_175303	Sall4	NM_175303	0,1186	-1,377
1E+07	NM_025723	4921515J06Rik	NM_025723	0,1186	-2,38
1,1E+07	NM_025350	Cpa1	NM_025350	0,1186	-1,479
1,1E+07	---	---	ENSMUST00000089622	0,1186	-1,419
1,1E+07	ENSMUST00000094774	Sp110 // MGC118608	ENSMUST00000094774	0,1186	-1,576
1E+07	---	---	---	0,1554	-1,406
1E+07	NM_133815	Lbr	NM_133815	0,1554	-1,364
1E+07	NM_177368	Tmtc2	NM_177368	0,1554	-1,427
1E+07	NM_008759	Sebox	NM_008759	0,1554	-1,665
1E+07	BC137602	3830431G21Rik	BC137602	0,1554	-1,422
1E+07	NM_009472	Unc5c	NM_009472	0,1554	-1,465
1,1E+07	NM_174877	Zar1	NM_174877	0,1554	-1,515
1,1E+07	NM_001003953	Fbxl10	NM_001003953	0,1554	-1,374
1,1E+07	NM_001042612	Nlrp9c	NM_001042612	0,1554	-1,604
1,1E+07	BC080701	EG382106	BC080701	0,1554	-2,168
1E+07	NM_175328	Slc6a15	NM_175328	0,1567	-3,434
1E+07	NM_178657	Oog1	NM_178657	0,1567	-1,6
1E+07	NM_178657	Oog1	NM_178657	0,1567	-1,635
1E+07	NM_008145	Gnrh1	NM_008145	0,1567	-1,553
1E+07	NM_028615	Dppa2	NM_028615	0,1567	-1,751
1,1E+07	NM_174877	Zar1	NM_174877	0,1567	-1,315
1,1E+07	XR_034094	OTTMUSG00000025963	XR_034094	0,1567	-1,466
1,1E+07	ENSMUST00000070778	XM_359419	ENSMUST00000070778	0,1567	-1,595
1,1E+07	BC024822	AW146242	BC024822	0,1567	-1,516
1,1E+07	NM_028238	Rab38	NM_028238	0,1567	-1,482
1,1E+07	NR_000002	Snord32a	NR_000002	0,1567	-1,49
1,1E+07	NM_001033632	Ifitm6	NM_001033632	0,1567	-2,735
1,1E+07	NM_001005863	Mtus1	NM_001005863	0,1567	-1,409
1,1E+07	NM_001014395	EG382156	NM_001014395	0,1567	-1,986
1,1E+07	BC100377	AU022751	BC100377	0,1567	-1,443
1E+07	NM_175397	Sp110	NM_175397	0,1714	-1,399
1,1E+07	---	---	AF357363	0,1714	-1,423
1E+07	NM_183322	Khdc1a	NM_183322	0,1793	-1,715
1E+07	AK016212	4931428L18Rik	AK016212	0,1857	-1,448
1E+07	NM_212457	Bex4	NM_212457	0,1857	-1,9
1E+07	NM_178715	Tmem30b	NM_178715	0,1857	-1,277
1E+07	NM_007940	Ephx2	NM_007940	0,1857	-1,41
1E+07	NM_033037	Cdo1	NM_033037	0,1857	-1,509
1,1E+07	NM_178657	Oog1	NM_178657	0,1857	-1,924
1,1E+07	NM_153574	D430015B01Rik	NM_153574	0,1857	-1,892
1,1E+07	NM_008217	Has3	NM_008217	0,1857	-1,433
1E+07	---	---	---	0,2116	-1,557
1E+07	NM_177388	Slc41a2	NM_177388	0,2116	-1,379
1E+07	NM_010927	Nos2	NM_010927	0,2116	-1,605
1E+07	NR_003492	2210409E12Rik	NR_003492	0,2116	-1,405
1E+07	ENSMUST00000091466	Ttc37	ENSMUST00000091466	0,2116	-1,3
1E+07	NM_001033441	Alg10b	NM_001033441	0,2116	-1,31

1E+07	NM_175549	Robo2	NM_175549	0,2116	-1,333
1E+07	ENSMUST00000097502	ENSMUSG00000073518	ENSMUST00000097502	0,2116	-1,424
1E+07	NM_172858	Pak7	NM_172858	0,2116	-1,513
1E+07	NM_007378	Abca4	NM_007378	0,2116	-1,322
1,1E+07	---	---	ENSMUST00000083163	0,2116	-1,361
1,1E+07	NM_025829	Eif4e3	NM_025829	0,2116	-1,369
1,1E+07	NM_027802	Obox1	NM_027802	0,2116	-1,637
1,1E+07	NM_172896	Nlrp4a	NM_172896	0,2116	-1,714
1,1E+07	NM_145707	Obox3	NM_145707	0,2116	-1,635
1,1E+07	NR_000003	Rnu35a	NR_000003	0,2116	-1,325
1,1E+07	NR_002173	449631	NR_002173	0,2116	-1,372
1,1E+07	---	---	mmu-mir-202	0,2116	-1,669
1,1E+07	NM_146367	Olfr976	NM_146367	0,2116	-2,451
1,1E+07	BC083100	EG546387	BC083100	0,2116	-1,476
1E+07	---	---	---	0,2279	-1,3
1E+07	NM_172557	Rufy1	NM_172557	0,2279	-1,323
1E+07	NM_001034865	Nek10	NM_001034865	0,2279	-1,406
1E+07	NM_007914	Ehf	NM_007914	0,2279	-1,629
1E+07	NM_013822	Jag1	NM_013822	0,2279	-1,379
1,1E+07	NM_011107	Pla2g1b	NM_011107	0,2279	-1,52
1,1E+07	NR_003269	Obox2	NR_003269	0,2279	-1,61
1,1E+07	NR_003269	Obox2	NR_003269	0,2279	-1,61
1E+07	---	---	---	0,2310	-1,351
1E+07	---	---	---	0,2310	-1,388
1E+07	---	---	---	0,2310	-1,387
1E+07	NM_010262	Gbx2	NM_010262	0,2310	-1,401
1E+07	NM_027968	Fbxo30	NM_027968	0,2310	-1,283
1E+07	NM_001081432	Ptprq	NM_001081432	0,2310	-1,414
1E+07	---	---	AF357373	0,2310	-1,505
1E+07	ENSMUST00000101450	ENSMUSG00000073103	ENSMUST00000101450	0,2310	-1,39
1E+07	NM_172953	Serpina5	NM_172953	0,2310	-2,593
1E+07	NM_013772	Tcl1b3	NM_013772	0,2310	-1,492
1E+07	NM_030174	Mctp1	NM_030174	0,2310	-1,693
1E+07	NM_172595	Arl15	NM_172595	0,2310	-1,276
1E+07	NM_019460	Sfmbt1	NM_019460	0,2310	-1,384
1E+07	---	---	ENSMUST00000082834	0,2310	-1,788
1E+07	NM_177307	Cyp4f39	NM_177307	0,2310	-1,379
1E+07	NM_001110832	Nfya	NM_001110832	0,2310	-1,244
1E+07	ENSMUST00000097295	AU016765	ENSMUST00000097295	0,2310	-1,407
1E+07	NM_201362	Ccdc68	NM_201362	0,2310	-1,397
1E+07	NM_001081213	Ermp1	NM_001081213	0,2310	-1,277
1E+07	---	---	GENSCAN00000009239	0,2310	-1,462
1E+07	NM_175462	Kcnt1	NM_175462	0,2310	-1,317
1E+07	ENSMUST00000074761	ENSMUSG00000063611	ENSMUST00000074761	0,2310	-1,32
1E+07	ENSMUST00000016530	Mpped2	ENSMUST00000016530	0,2310	-1,776
1E+07	AF440737	Cst12	AF440737	0,2310	-1,544
1E+07	NM_013735	Trp53bp1	NM_013735	0,2310	-1,235
1,1E+07	NM_009899	Clca1	NM_009899	0,2310	-1,335
1,1E+07	BC084594	OTTMUSG00000010009	BC084594	0,2310	-2,079
1,1E+07	BC158124	OTTMUSG00000007332	BC158124	0,2310	-1,426
1,1E+07	---	---	ENSMUST00000083370	0,2310	-1,565
1,1E+07	NM_011776	Zp3	NM_011776	0,2310	-1,45
1,1E+07	NM_178612	Cnpy4	NM_178612	0,2310	-1,258
1,1E+07	NM_007984	Fscn1	NM_007984	0,2310	-1,641
1,1E+07	NM_009193	Slbp	NM_009193	0,2310	-1,363
1,1E+07	---	---	U34827	0,2310	-1,685
1,1E+07	NM_001109749	Cntrn4	NM_001109749	0,2310	-1,916
1,1E+07	NM_026820	lfitm1	NM_026820	0,2310	-1,284
1,1E+07	---	---	ENSMUST00000083703	0,2310	-1,27
1,1E+07	NM_144934	Mbd3l2	NM_144934	0,2310	-3,376
1,1E+07	BC023444	Onecut1	BC023444	0,2310	-1,261
1,1E+07	NM_001081322	Myo5c	NM_001081322	0,2310	-1,408
1,1E+07	NM_024245	Kif23	NM_024245	0,2310	-1,323

1,1E+07 BC052196	E330016A19Rik	BC052196	0,2310	-1,371
1,1E+07 NM_212457	Bex4	NM_212457	0,2310	-1,766
1,1E+07 NM_177353	Slc9a7	NM_177353	0,2310	-1,314
1E+07 ---	---	---	0,2530	-1,28
1E+07 ---	---	---	0,2530	-1,474
1E+07 ---	---	---	0,2530	-1,408
1E+07 NM_010475	Hsd17b1	NM_010475	0,2530	-2,179
1E+07 NR_003649	EG432649	NR_003649	0,2530	-1,558
1E+07 ---	---	BC137768	0,2530	-1,699
1E+07 NM_001007460	Zdhhc23	NM_001007460	0,2530	-1,356
1E+07 NM_007552	Bmi1	NM_007552	0,2530	-1,237
1E+07 NM_016904	Cks1b	NM_016904	0,2530	-1,258
1,1E+07 ---	---	GENSCAN00000026443	0,2530	-1,62
1,1E+07 NM_175217	Mmd2	NM_175217	0,2530	-1,484
1,1E+07 NM_172893	Parp12	NM_172893	0,2530	-1,357
1,1E+07 NM_053096	Cml2	NM_053096	0,2530	-1,509
1,1E+07 ENSMUST00000113384	A530040E14Rik	ENSMUST00000113384	0,2530	-1,368
1,1E+07 ENSMUST00000113384	A530040E14Rik	ENSMUST00000113384	0,2530	-1,368
1,1E+07 NM_026480	Ooep	NM_026480	0,2530	-1,68
1,1E+07 ---	---	ENSMUST00000075518	0,2530	-1,279
1,1E+07 NM_212444	Gyk	NM_212444	0,2530	-1,325
1,1E+07 NM_026662	Prps2	NM_026662	0,2530	-1,536
1E+07 ---	---	---	0,2657	-1,464
1E+07 ---	---	---	0,2657	-1,328
1E+07 ---	---	---	0,2657	-1,521
1E+07 NM_183322	Khdc1a	NM_183322	0,2657	-1,501
1E+07 NM_001004173	Sgpp2	NM_001004173	0,2657	-1,247
1E+07 ENSMUST00000113384	A530040E14Rik	ENSMUST00000113384	0,2657	-1,366
1E+07 BC100303	A530040E14Rik	BC100303	0,2657	-1,424
1E+07 ENSMUST00000113384	A530040E14Rik	ENSMUST00000113384	0,2657	-1,366
1E+07 NM_009061	Rgs2	NM_009061	0,2657	-1,215
1E+07 NM_011129	Sept4	NM_011129	0,2657	-1,281
1E+07 NM_181071	Tanc2	NM_181071	0,2657	-1,221
1E+07 NM_134246	Acot3	NM_134246	0,2657	-1,605
1E+07 AK008011	2010001K21Rik	AK008011	0,2657	-1,348
1E+07 NM_023118	Dab2	NM_023118	0,2657	-1,265
1E+07 NM_008474	Krt84	NM_008474	0,2657	-1,251
1E+07 BC099383	E330017A01Rik	BC099383	0,2657	-1,549
1E+07 NM_178774	Prr18	NM_178774	0,2657	-1,254
1E+07 NM_001003919	Ddx11	NM_001003919	0,2657	-1,468
1E+07 NM_022311	Tcte2	NM_022311	0,2657	-1,37
1E+07 ENSMUST00000097557	ENSMUSG00000073551	ENSMUST00000097557	0,2657	-1,452
1E+07 ENSMUST00000048109	2700046A07Rik	ENSMUST00000048109	0,2657	-1,257
1E+07 NM_011824	Grem1	NM_011824	0,2657	-1,31
1,1E+07 ENSMUST00000098020	8030463A06Rik	ENSMUST00000098020	0,2657	-1,636
1,1E+07 NM_175523	Ppm1k	NM_175523	0,2657	-1,246
1,1E+07 ENSMUST00000108482	Psg20 // EG434540	ENSMUST00000108482	0,2657	-1,317
1,1E+07 NM_001033983	7420426K07Rik	NM_001033983	0,2657	-1,278
1,1E+07 NM_176807	Folr4	NM_176807	0,2657	-1,39

Supplementary Table II

List of genes up-regulated in *Tbp2*^{-/-} ovaries.

affyID	Gene.Accession	Gene.Symbol	mRNA.Accession	fdr	FC
10605044	NM_001111293	Xlr5b	NM_001111293	0,0067	3,324
10554900	NM_011807	Dlg2	NM_011807	0,0153	1,943
10605018	NM_001045539	Xlr5a	NM_001045539	0,0153	3,003
10357288	NM_010483	Htr5b	NM_010483	0,0279	1,744
10584604	NM_023655	Trim29	NM_023655	0,0279	2,482
10341885	---	---	---	0,0379	1,74
10347226	NM_175564	Tmem169	NM_175564	0,0379	1,803
10368947	NM_172393	Aim1	NM_172393	0,0379	2,377
10385118	NM_033374	Dock2	NM_033374	0,0379	2,53
10449631	NM_027060	Btbd9	NM_027060	0,0379	1,591
10487987	NM_177303	Lrrn4	NM_177303	0,0379	1,99
10526169	NM_181045	Caln1	NM_181045	0,0379	2,107
10547471	NM_198884	B4galnt3	NM_198884	0,0379	1,737
10364262	NM_008404	Itgb2	NM_008404	0,0571	1,729
10392910	BC024617	C630004H02Rik	BC024617	0,0571	1,828
10421325	NM_021308	Piwil2	NM_021308	0,0571	2,414
10430649	NM_144811	Cbx7	NM_144811	0,0571	1,496
10432661	NM_172451	Galnt6	NM_172451	0,0571	2,31
10439583	NM_198034	Sidt1	NM_198034	0,0571	2,322
10549804	NM_031389	Nlrp4c	NM_031389	0,0571	2,587
10556067	NM_021889	Syt9	NM_021889	0,0571	1,589
10570516	BC080858	Kbtbd11	BC080858	0,0571	1,625
10575745	NM_177700	Atmin	NM_177700	0,0571	1,63
10587604	NM_027100	Rwdd2a	NM_027100	0,0571	2,719
10608697	---	---	NM_016928.2	0,0571	1,613
10427369	NM_008800	Pde1b	NM_008800	0,0897	1,536
10428619	NM_015744	Enpp2	NM_015744	0,0897	1,464
10435793	NM_007877	Drd3	NM_007877	0,0897	1,723
10477557	NM_199304	Zfp341	NM_199304	0,0897	1,575
10488565	NM_001114677	OTTMUSG00000015829	NM_001114677	0,0897	2,424
10489440	NM_183271	Wfdc15a	NM_183271	0,0897	1,835
10497622	NM_027941	Lrrc34	NM_027941	0,0897	3,35
10507961	NM_008636	Mtf1	NM_008636	0,0897	1,526
10524515	ENSMUST00000031566	Myo1h	ENSMUST00000031566	0,0897	1,574
10533869	NM_144819	Ccdc92	NM_144819	0,0897	1,505
10545096	NM_028025	2410003J06Rik	NM_028025	0,0897	2,205
10555039	NM_001039039	Kctd21	NM_001039039	0,0897	1,753
10338794	---	---	---	0,0975	1,768
10368101	NM_001033258	D10Bwg1379e	NM_001033258	0,0975	1,566
10371159	NM_019757	Fzr1	NM_019757	0,0975	1,47
10376832	NM_007413	Adora2b	NM_007413	0,0975	1,476
10383511	AF285590	Tex19.1	AF285590	0,0975	1,62
10397736	NM_146037	Kcnk13	NM_146037	0,0975	1,665
10402283	NM_172584	Itpk1	NM_172584	0,0975	1,729
10413517	NM_175343	Chdh	NM_175343	0,0975	1,45
10463739	NM_177342	Taf5	NM_177342	0,0975	1,724
10476617	XR_034950	OTTMUSG00000015606	XR_034950	0,0975	1,512
10500011	NM_011656	Tuft1	NM_011656	0,0975	1,452
10511984	NM_027041	1700003M02Rik	NM_027041	0,0975	1,602
10516335	NM_011902	Tekt2	NM_011902	0,0975	1,538
10536908	NM_146173	Tspan33	NM_146173	0,0975	1,775
10537458	NM_001004182	EG434008	NM_001004182	0,0975	1,722
10546421	NM_001081146	Prickle2	NM_001081146	0,0975	1,618

10556749	NM_178414	Acsm4	NM_178414	0,0975	1,508
10580469	NM_019626	Cbln1	NM_019626	0,0975	1,531
10584634	NM_198092	Usp2	NM_198092	0,0975	1,482
10597427	NM_001042503	Trim71	NM_001042503	0,0975	1,707
10604984	NM_021365	Xlr4b	NM_021365	0,0975	1,895
10605034	NM_183094	Xlr4c	NM_183094	0,0975	2,028
10351380	ENSMUST00000111416	LOC100039795	ENSMUST00000111416	0,1186	1,517
10355152	NM_022721	Fzd5	NM_022721	0,1186	1,474
10358787	NM_008485	Lamc2	NM_008485	0,1186	1,591
10366597	NM_028736	Grip1	NM_028736	0,1186	1,53
10372583	NM_001003950	Rab3ip	NM_001003950	0,1186	1,471
10382449	NM_021411	Rab37	NM_021411	0,1186	1,744
10392440	NM_001029842	Slc16a6	NM_001029842	0,1186	1,941
10425814	NM_172610	Mpped1	NM_172610	0,1186	1,754
10429100	NM_019457	Lrrc6	NM_019457	0,1186	2,229
10447383	NM_008532	Tacstd1	NM_008532	0,1186	2,37
10455761	NM_001033281	Prdm6	NM_001033281	0,1186	1,38
10466735	NM_001114174	Gm967	NM_001114174	0,1186	1,569
10490502	NM_178254	Tcfl5	NM_178254	0,1186	2,018
10540075	NM_144919	Hdac11	NM_144919	0,1186	1,452
10548745	NM_008157	Gpr19	NM_008157	0,1186	1,553
10570018	NM_033622	Tnfsf13b	NM_033622	0,1186	1,675
10579731	NM_027485	Med26	NM_027485	0,1186	1,554
10600131	NM_021365	Xlr4b	NM_021365	0,1186	1,905
10605542	NM_001113734	2410003J06Rik	NM_001113734	0,1186	2,164
10350297	NM_001081258	Kif14	NM_001081258	0,1554	1,616
10354003	NM_173870	Mgat4a	NM_173870	0,1554	2,137
10362490	NM_030203	Tspyl4	NM_030203	0,1554	1,386
10365891	NM_172051	Tmcc3	NM_172051	0,1554	1,492
10366712	NM_001110218	Ppm1h	NM_001110218	0,1554	1,716
10372831	ENSMUST00000064107	Tbc1d30	ENSMUST00000064107	0,1554	1,564
10376959	NM_023479	Elac2	NM_023479	0,1554	1,589
10378367	NM_145099	Trpv3	NM_145099	0,1554	1,494
10382998	NM_001012273	Birc5	NM_001012273	0,1554	1,662
10386628	NM_025294	Gtlf3b	NM_025294	0,1554	1,389
10390075	NM_001105561	OTTMUSG00000002043	NM_001105561	0,1554	1,319
10394054	NM_009854	Cd7	NM_009854	0,1554	1,608
10396454	NM_172804	Syt16	NM_172804	0,1554	1,731
10399376	ENSMUST00000062469	9930038B18Rik	ENSMUST00000062469	0,1554	1,44
10400967	NM_009189	Six1	NM_009189	0,1554	1,457
10408266	NM_175663	Hist1h2ba	NM_175663	0,1554	1,543
10408638	ENSMUST00000021851	1700026J04Rik	ENSMUST00000021851	0,1554	1,358
10409118	NM_029361	Wnk2	NM_029361	0,1554	1,9
10427015	NM_007395	Acvr1b	NM_007395	0,1554	1,407
10434094	NM_145479	Klhl22	NM_145479	0,1554	1,357
10434577	NM_001081366	Vps8	NM_001081366	0,1554	1,8
10440849	ENSMUST00000023696	Synj1	ENSMUST00000023696	0,1554	1,54
10458052	NM_013512	Epb4.114a	NM_013512	0,1554	1,354
10474096	NM_178725	Lrrc4c	NM_178725	0,1554	1,643
10475804	AB257502	A530057A03Rik	AB257502	0,1554	1,827
10507032	NM_026279	2310026E23Rik	NM_026279	0,1554	1,336
10507551	BC022764	2610528J11Rik	BC022764	0,1554	1,538
10508249	ENSMUST00000094712	OTTMUSG00000009332	ENSMUST00000094712	0,1554	1,513
10511901	NM_001012450	Ankrd6	NM_001012450	0,1554	1,497
10512640	NM_015828	Gne	NM_015828	0,1554	1,367
10522577	NM_177187	D5Ertd577e	NM_177187	0,1554	1,744
10525677	NM_145070	Hip1r	NM_145070	0,1554	1,606
10527870	NM_013823	Kl	NM_013823	0,1554	1,521

10532574	XM_912851	Myo18b	XM_912851	0,1554	1,507
10535739	NM_011669	Usp12	NM_011669	0,1554	1,561
10536159	BC099967	AU018829	BC099967	0,1554	1,744
10549794	NR_003622	EG272350	NR_003622	0,1554	1,646
10550772	NM_001034904	EG545929	NM_001034904	0,1554	2,891
10554156	BC034069	Fam174b	BC034069	0,1554	1,477
10560807	NM_001034904	EG545929	NM_001034904	0,1554	2,891
10561153	NM_001081148	Cyp2b23	NM_001081148	0,1554	1,457
10575052	NM_009864	Cdh1	NM_009864	0,1554	1,515
10575993	BC027126	6430548M08Rik	BC027126	0,1554	1,378
10583056	NM_008605	Mmp12	NM_008605	0,1554	1,768
10590983	NM_019482	Panx1	NM_019482	0,1554	1,326
10594001	NM_019689	Arid3b	NM_019689	0,1554	1,434
10595953	NM_177775	D9Ertd280e	NM_177775	0,1554	1,616
10597377	NM_178624	Fbxl2	NM_178624	0,1554	1,482
10602677	NM_010951	Gpr143	NM_010951	0,1554	2,537
10338400	---	---	---	0,1567	1,503
10352110	BC147679	EG545391	BC147679	0,1567	1,391
10353192	NM_010164	Eya1	NM_010164	0,1567	1,596
10357261	NM_010766	Marco	NM_010766	0,1567	1,462
10365974	NM_007833	Dcn	NM_007833	0,1567	1,324
10372503	NM_010195	Lgr5	NM_010195	0,1567	2,365
10373846	NM_001029937	Sec14l3	NM_001029937	0,1567	1,352
10376127	ENSMUST00000094193	4930404A10Rik	ENSMUST00000094193	0,1567	1,536
10376803	NM_178618	2310040C09Rik	NM_178618	0,1567	1,364
10383767	NM_152818	Osbp2	NM_152818	0,1567	1,332
10386123	XR_034114	LOC676530	XR_034114	0,1567	1,333
10391301	NM_213659	Stat3	NM_213659	0,1567	1,631
10400023	NM_025359	Tspan13	NM_025359	0,1567	1,688
10401343	NM_177395	Map3k9	NM_177395	0,1567	1,505
10402440	NM_010351	Gsc	NM_010351	0,1567	1,434
10411680	NM_001038602	Marveld2	NM_001038602	0,1567	2,005
10411728	NM_021886	Cenph	NM_021886	0,1567	1,656
10413461	NM_177814	Erc2	NM_177814	0,1567	1,793
10430255	NM_001017983	Foxred2	NM_001017983	0,1567	1,406
10430289	NM_025931	Rabl4	NM_025931	0,1567	1,666
10448016	NM_011560	Tcte3	NM_011560	0,1567	1,369
10448023	NM_011560	Tcte3	NM_011560	0,1567	1,352
10450154	NM_010378	H2-Aa	NM_010378	0,1567	1,381
10455942	BC075669	A730017C20Rik	BC075669	0,1567	1,45
10474181	NM_178890	Abtb2	NM_178890	0,1567	1,439
10491613	ENSMUST00000057272	4932438A13Rik	ENSMUST00000057272	0,1567	1,428
10492720	NM_001081230	Mtap9	NM_001081230	0,1567	1,428
10501963	NM_011674	Ugt8a	NM_011674	0,1567	1,804
10512714	ENSMUST00000061986	Shb	ENSMUST00000061986	0,1567	1,353
10517401	NM_001013756	Grhl3	NM_001013756	0,1567	1,508
10519028	NM_027504	Prdm16	NM_027504	0,1567	1,404
10522573	NM_177187	D5Ertd577e	NM_177187	0,1567	1,717
10522581	NM_177187	D5Ertd577e	NM_177187	0,1567	1,744
10523312	NM_177187	D5Ertd577e	NM_177187	0,1567	1,705
10524909	NM_008712	Nos1	NM_008712	0,1567	1,934
10525473	NM_001039723	Tmem120b	NM_001039723	0,1567	1,855
10536021	NM_177187	D5Ertd577e	NM_177187	0,1567	1,717
10536155	ENSMUST00000111414	AU018829	ENSMUST00000111414	0,1567	1,744
10543791	NM_013723	Podxl	NM_013723	0,1567	1,381
10544547	NM_001081382	Zfp777	NM_001081382	0,1567	1,33
10544829	NM_173406	Jazf1	NM_173406	0,1567	1,515
10548097	BC147705	9630033F20Rik	BC147705	0,1567	1,359

10549473	NM_181541	Caprin2	NM_181541	0,1567	1,729
10552725	NM_133945	Vrk3	NM_133945	0,1567	1,473
10557060	ENSMUST00000033166	4933427G17Rik	ENSMUST00000033166	0,1567	1,29
10558333	NM_025850	Fank1	NM_025850	0,1567	1,639
10560764	NM_001034904	EG545929	NM_001034904	0,1567	2,753
10560777	NM_001034904	EG545929	NM_001034904	0,1567	2,753
10560791	NM_001034904	EG545929	NM_001034904	0,1567	2,655
10566161	NM_198012	Trim68	NM_198012	0,1567	1,388
10582196	BC027666	1190005I06Rik	BC027666	0,1567	1,501
10591706	NM_010487	Elavl3	NM_010487	0,1567	1,511
10604996	NM_001110784	Xlr3a	NM_001110784	0,1567	1,758
10608709	---	---	NM_025998.1	0,1567	1,343
10379321	NM_175543	Rab11fip4	NM_175543	0,1714	1,51
10397882	NM_007693	Chga	NM_007693	0,1714	1,407
10405964	BC119515	1700001L19Rik	BC119515	0,1714	1,354
10494978	NM_008979	Ptpn22	NM_008979	0,1714	1,567
10359801	NM_175296	Mael	NM_175296	0,1793	2,284
10366683	NM_177335	D930020B18Rik	NM_177335	0,1857	2,003
10405876	ENSMUST00000079272	ENSMUSG00000071302	ENSMUST00000079272	0,1857	2,869
10424750	NM_001081065	Zfp707	NM_001081065	0,1857	1,362
10429564	NM_010738	Ly6a	NM_010738	0,1857	1,833
10433145	ENSMUST00000023129	Gtsf1	ENSMUST00000023129	0,1857	1,955
10491605	ENSMUST00000057272	4932438A13Rik	ENSMUST00000057272	0,1857	1,368
10505145	NM_001037127	Musk	NM_001037127	0,1857	1,86
10536037	NM_001122668	E330014E10Rik	NM_001122668	0,1857	1,696
10536048	NM_177187	D5Ertd577e	NM_177187	0,1857	1,696
10536147	BC099967	AU018829	BC099967	0,1857	1,7
10555303	NM_027629	Pgm2l1	NM_027629	0,1857	1,466
10575867	NM_019966	Mlycd	NM_019966	0,1857	1,519
10596200	NM_178638	Tmem108	NM_178638	0,1857	1,666
10599554	NM_173376	Rbmx2	NM_173376	0,1857	1,385
10341077	---	---	---	0,2116	1,612
10355711	NM_029342	Nhej1	NM_029342	0,2116	1,355
10380025	NM_031386	Tex14	NM_031386	0,2116	1,507
10382462	NM_012030	Slc9a3r1	NM_012030	0,2116	1,366
10386238	AY603754	Obscn	AY603754	0,2116	1,404
10389601	BC020005	1200011M11Rik	BC020005	0,2116	1,334
10394060	NM_026907	Sectm1b	NM_026907	0,2116	1,414
10397257	NM_007701	Vsx2	NM_007701	0,2116	1,343
10397966	NM_026580	Otub2	NM_026580	0,2116	1,435
10400141	NM_181412	Zbed4	NM_181412	0,2116	1,336
10410237	ENSMUST00000091541	2410141K09Rik	ENSMUST00000091541	0,2116	1,685
10426425	ENSMUST00000035399	Pdzrn4	ENSMUST00000035399	0,2116	1,767
10431124	NM_029427	Pnpla5	NM_029427	0,2116	1,68
10434815	NM_001127259	Trp63	NM_001127259	0,2116	1,552
10448694	NM_011522	Syngn3	NM_011522	0,2116	1,43
10457333	NM_001081393	Armc4	NM_001081393	0,2116	1,361
10459183	NM_007885	Slc26a2	NM_007885	0,2116	1,402
10463737	AK144917	Ina	AK144917	0,2116	1,988
10479975	---	---	ENSMUST00000083890	0,2116	1,375
10488567	NM_001114677	OTTMUSG00000015829	NM_001114677	0,2116	2,186
10490061	NM_029815	Bcas1	NM_029815	0,2116	1,38
10490129	NM_007557	Bmp7	NM_007557	0,2116	1,45
10494536	NM_021517	Pdzk1	NM_021517	0,2116	1,623
10500559	NM_013821	Hsd3b6	NM_013821	0,2116	2,007
10518201	NM_029948	Pramef12	NM_029948	0,2116	1,615
10522554	NM_177187	D5Ertd577e	NM_177187	0,2116	1,697
10522569	NM_177187	D5Ertd577e	NM_177187	0,2116	1,697

10523350	NM_177187	D5Ertd577e	NM_177187	0,2116	1,654
10531497	NM_177187	D5Ertd577e	NM_177187	0,2116	1,694
10533734	NM_019875	Abcb9	NM_019875	0,2116	1,387
10536114				0,2116	1,654
10538214	BC055818	D330028D13Rik	BC055818	0,2116	1,446
10544941	NM_011054	Pde1c	NM_011054	0,2116	1,348
10547075	BC060267	BC060267	BC060267	0,2116	1,334
10548043	NM_145983	Kcna5	NM_145983	0,2116	1,373
10550222	NM_001099277	Zfp541	NM_001099277	0,2116	1,315
10550786	NM_001122682	EG667283	NM_001122682	0,2116	2,391
10560754	BC099438	EG667283	BC099438	0,2116	2,367
10561335	M94632	Prkcz	M94632	0,2116	1,506
10561819	NM_177889	Zfp82	NM_177889	0,2116	1,491
10573332	BC075658	4930432K21Rik	BC075658	0,2116	1,519
10600122	NM_001081643	Xlr3b	NM_001081643	0,2116	1,848
10603768	NM_028916	Efhc2	NM_028916	0,2116	1,635
10604961	NM_008067	Gabra3	NM_008067	0,2116	1,787
10605007	NM_011727	Xlr3c	NM_011727	0,2116	1,789
10339691	---	---	---	0,2279	1,592
10340303	---	---	---	0,2279	1,437
10346410	NM_023617	Aox3	NM_023617	0,2279	1,415
10356177	NM_152915	Dner	NM_152915	0,2279	1,814
10391378	NM_007970	Ezh1	NM_007970	0,2279	1,336
10393955	NM_026824	Dus1l	NM_026824	0,2279	1,374
10395039	NM_020557	Cmpk2	NM_020557	0,2279	1,257
10400006	NM_013464	Ahr	NM_013464	0,2279	1,292
10402142	NM_026681	Ccdc88c	NM_026681	0,2279	1,662
10406777	AK157788	Gm73	AK157788	0,2279	1,395
10416437	NM_008879	Lcp1	NM_008879	0,2279	1,809
10454235	ENSMUST00000062608	Asxl3	ENSMUST00000062608	0,2279	1,432
10480869	NM_001001714	Sohlh1	NM_001001714	0,2279	1,62
10490053	NM_001033299	Zfp217	NM_001033299	0,2279	1,437
10491603	ENSMUST00000057272	4932438A13Rik	ENSMUST00000057272	0,2279	1,322
10508883	NM_001083916	1810019J16Rik	NM_001083916	0,2279	1,779
10522530	NM_001122733	Kit	NM_001122733	0,2279	1,556
10534152	NM_145218	Wbscr17	NM_145218	0,2279	1,507
10546349	NM_009531	Xpc	NM_009531	0,2279	1,299
10550762	NM_001034904	EG545929	NM_001034904	0,2279	2,629
10550775				0,2279	2,629
10550778	NM_001122682 // NM_C	EG667283 // EG667283	NM_001122682	0,2279	2,476
10560767	---	---	GENSCAN00000047966	0,2279	2,476
10575034	NM_001037809	Cdh3	NM_001037809	0,2279	1,267
10588767	NM_011634	Traip	NM_011634	0,2279	1,358
10600688	NM_025729	Map3k7ip3	NM_025729	0,2279	1,308
10600777	---	---	BC141215	0,2279	1,548
10338996	---	---	---	0,2310	1,453
10339667	---	---	---	0,2310	1,369
10339751	---	---	---	0,2310	1,398
10341986	---	---	---	0,2310	1,422
10342227	---	---	---	0,2310	1,24
10342944	---	---	---	0,2310	1,349
10344050	---	---	---	0,2310	2,288
10344390	---	---	---	0,2310	1,889
10345482	NM_033570	Cnm4	NM_033570	0,2310	1,513
10345840	NM_177084	Slc9a4	NM_177084	0,2310	1,307
10347583	NM_010043	Des	NM_010043	0,2310	1,57
10350286	NM_026500	Ddx59	NM_026500	0,2310	1,391
10350710	BC113157	1700012A16Rik	BC113157	0,2310	1,834

10353754	NM_133817	Zfp451	NM_133817	0,2310	1,302
10354141	NM_001029878	Lonrf2	NM_001029878	0,2310	1,9
10354563	ENSMUST00000069293	Dnahc7b	ENSMUST00000069293	0,2310	1,315
10357488	NM_010016	Cd55	NM_010016	0,2310	1,71
10357705	NM_177129	Cntn2	NM_177129	0,2310	1,298
10358272	NM_001042577	Lhx9	NM_001042577	0,2310	1,378
10361023	NM_008937	Prox1	NM_008937	0,2310	1,422
10361846	NM_009048	Reps1	NM_009048	0,2310	1,259
10362379	NM_025855	Echdc1	NM_025855	0,2310	1,295
10365384	NM_001024918	Rfx4	NM_001024918	0,2310	1,341
10369774	NM_001036293	Nrbf2	NM_001036293	0,2310	1,242
10375265	NM_176999	Atp10b	NM_176999	0,2310	1,58
10377889	NM_001045959	Mink1	NM_001045959	0,2310	1,348
10382565	NM_025305	Mrps7	NM_025305	0,2310	1,3
10384930	NM_001100394	4930505A04Rik	NM_001100394	0,2310	1,484
10385175	NM_170779	Wwc1	NM_170779	0,2310	1,249
10385466	NM_011891	Sgcd	NM_011891	0,2310	1,555
10388194	BC025823	Spns2	BC025823	0,2310	1,265
10388310	NM_001015046	Garnl4	NM_001015046	0,2310	1,28
10390050	NM_177752	Eme1	NM_177752	0,2310	1,588
10396671	NM_153804	Plekhg3	NM_153804	0,2310	1,303
10398224	NM_010010	Cyp46a1	NM_010010	0,2310	1,336
10398795	BC126962	A530050D06Rik	BC126962	0,2310	1,272
10401443	BC062132	9830169C18Rik	BC062132	0,2310	1,501
10402473	NM_053155	Clnn	NM_053155	0,2310	2,074
10404737	XM_001472337	LOC637277	XM_001472337	0,2310	1,646
10404941	NM_172262	Aof1	NM_172262	0,2310	1,544
10407591	NM_033269	Chrm3	NM_033269	0,2310	1,454
10410248	ENSMUST00000079229	Zfp640	ENSMUST00000079229	0,2310	3,001
10411668	NM_008756	Ocln	NM_008756	0,2310	1,751
10417144	NM_001037744	Timm8a2	NM_001037744	0,2310	1,507
10418868	---	---	GENSCAN0000002628E	0,2310	1,367
10419170	NM_172597	Txndc16	NM_172597	0,2310	1,355
10419729	NM_001036293	Nrbf2	NM_001036293	0,2310	1,268
10419731	NM_001036293	Nrbf2	NM_001036293	0,2310	1,268
10420596	NM_013869	Tnfrsf19	NM_013869	0,2310	1,422
10425421	NM_145986	AW544981	NM_145986	0,2310	1,327
10426093	NM_181412	Zbed4	NM_181412	0,2310	1,308
10427704	NM_030046	Dnajc21	NM_030046	0,2310	1,441
10427816	NM_001081064	Pdzd2	NM_001081064	0,2310	1,365
10428943	NM_031378	Gsdmc1	NM_031378	0,2310	1,282
10430113	BC060637	D15Wsu169e	BC060637	0,2310	1,302
10430425	NM_025622	Lgals2	NM_025622	0,2310	1,494
10430693	NM_001082536	Mkl1	NM_001082536	0,2310	1,331
10431113	NM_013873	Sult4a1	NM_013873	0,2310	1,417
10441270	NM_023663	Ripk4	NM_023663	0,2310	1,389
10447860	BC158033	Pnlcd1	BC158033	0,2310	1,597
10456254	NM_001114386	Nedd4l	NM_001114386	0,2310	1,401
10456639	NM_028948	Ccdc11	NM_028948	0,2310	1,537
10456653	NM_201600	Myo5b	NM_201600	0,2310	1,517
10460645	NM_022012	Map3k11	NM_022012	0,2310	1,257
10466659	NM_010266	Gda	NM_010266	0,2310	1,669
10469622	NM_001004761	Gpr158	NM_001004761	0,2310	1,367
10471904	NM_028809	Arpc5l	NM_028809	0,2310	1,46
10474902	NM_011234	Rad51	NM_011234	0,2310	1,564
10475653	NM_011978	Slc27a2	NM_011978	0,2310	2,207
10476740	NM_053195	Slc24a3	NM_053195	0,2310	1,718
10476952	NM_172117	Entpd6	NM_172117	0,2310	1,721

10478778	NM_001085495	Arfgef2	NM_001085495	0,2310	1,291
10479230	NM_028806	Phactr3	NM_028806	0,2310	1,472
10481772	NM_178888	Garnl3	NM_178888	0,2310	1,341
10483081	NM_007986	Fap	NM_007986	0,2310	2,04
10483871	NM_011652	Ttn	NM_011652	0,2310	1,744
10484894	NM_008982	Ptprj	NM_008982	0,2310	1,619
10486172	NM_001033136	1200015F23Rik	NM_001033136	0,2310	1,265
10486241	NM_172857	Exdl1	NM_172857	0,2310	1,541
10487787	NM_020014	Gfra4	NM_020014	0,2310	1,345
10489055	NM_029983	Sla2	NM_029983	0,2310	1,338
10491621	ENSMUST00000057272	4932438A13Rik	ENSMUST00000057272	0,2310	1,355
10491629	BC079623	4932438A13Rik	BC079623	0,2310	1,371
10497894	NM_145825	Cetn4	NM_145825	0,2310	1,382
10498907	NM_010298	Glrp	NM_010298	0,2310	1,478
10500406	NM_018767	Cd160	NM_018767	0,2310	1,591
10503170	NM_001081417	Chd7	NM_001081417	0,2310	1,313
10503218	NM_001081417	Chd7	NM_001081417	0,2310	1,42
10504534	NM_001081172	Frmpr1	NM_001081172	0,2310	1,45
10505270	NM_025286	Slc31a2	NM_025286	0,2310	1,4
10505299	NM_138653	Bspry	NM_138653	0,2310	1,41
10507379	NM_001029912	Zswim5	NM_001029912	0,2310	1,409
10510861	ENSMUST00000030897	Megf6	ENSMUST00000030897	0,2310	1,318
10514779	NM_178143	Prkaa2	NM_178143	0,2310	1,3
10517373	NM_022980	Rcan3	NM_022980	0,2310	1,384
10523341	NM_177187	D5Ert577e	NM_177187	0,2310	1,6
10523563	NM_173370	Cds1	NM_173370	0,2310	1,697
10523856	---	---	ENSMUST00000094541	0,2310	1,392
10526113	NM_007761	Crcp	NM_007761	0,2310	1,317
10526153	NM_001015876	Tyw1	NM_001015876	0,2310	1,387
10526781	BC038925	BC038925	BC038925	0,2310	1,233
10527565	NM_008814	Pdx1	NM_008814	0,2310	1,335
10532711	NM_008153	Cmklr1	NM_008153	0,2310	1,269
10533007	NM_001080808	Ccdc64	NM_001080808	0,2310	1,266
10534021	NM_001081388	Rimbp2	NM_001081388	0,2310	1,314
10535759	NM_080795	Lnx2	NM_080795	0,2310	1,361
10540923	NM_023290	Mkrn2	NM_023290	0,2310	1,343
10542460	NM_172733	Dera	NM_172733	0,2310	1,358
10543802	NM_175750	Plxna4	NM_175750	0,2310	1,384
10544648	NM_018769	Dfna5h	NM_018769	0,2310	1,442
10544875	NM_027268	Scrn1	NM_027268	0,2310	1,457
10546217	NM_025351	Chchd6	NM_025351	0,2310	1,235
10546272	ENSMUST00000101153	Iqsec1	ENSMUST00000101153	0,2310	1,36
10546706	NM_019743	Rybp	NM_019743	0,2310	1,366
10547015	NM_026894	1500001M20Rik	NM_026894	0,2310	1,331
10550972	NM_001102613	Phldb3	NM_001102613	0,2310	1,441
10550980	NM_133743	Lypd3	NM_133743	0,2310	1,241
10556828	NM_028085	Anks4b	NM_028085	0,2310	1,536
10559207	NM_019391	Lsp1	NM_019391	0,2310	1,485
10560760	NM_001034904	EG545929	NM_001034904	0,2310	2,461
10560773	NM_001034904	EG545929	NM_001034904	0,2310	2,461
10560803	NM_001034904	EG545929	NM_001034904	0,2310	2,461
10568115	NM_080638	Mvp	NM_080638	0,2310	1,451
10569618	NM_178642	Ano1	NM_178642	0,2310	1,475
10572733	BC049349	BC049349	BC049349	0,2310	1,531
10574384	NM_145602	Ndrp4	NM_145602	0,2310	1,689
10581088	NM_180958	Ccdc79	NM_180958	0,2310	2,066
10582152	BC030874	4632415K11Rik	BC030874	0,2310	1,315
10585555	NM_011193	Pstpip1	NM_011193	0,2310	1,518

10588429	NM_001081309	Pik3r4	NM_001081309	0,2310	1,331
10591643	NM_031874	Rab3d	NM_031874	0,2310	1,535
10593799	NM_001081341	Scaper	NM_001081341	0,2310	1,342
10596403	AM748257	EG665033	AM748257	0,2310	2,917
10599792	NM_133251	Vgll1	NM_133251	0,2310	1,456
10599917	NM_174993	Fmr1nb	NM_174993	0,2310	1,444
10600159	NM_009566	Zfp92	NM_009566	0,2310	1,317
10603289	NM_016691	Cicn5	NM_016691	0,2310	1,269
10604765	NM_133197	Mcf2	NM_133197	0,2310	1,434
10605917	ENSMUST00000101369	OTTMUSG00000018066	ENSMUST00000101369	0,2310	1,388
10605954	NM_031384	Tex11	NM_031384	0,2310	2,806
10606060	XM_136015	EG212753	XM_136015	0,2310	2,171
10607366	NM_172441	Shroom2	NM_172441	0,2310	1,35
10343350	---	---	---	0,2530	1,29
10345752	NM_010555	Il1r2	NM_010555	0,2530	1,405
10354031	NM_207228	Tsga10	NM_207228	0,2530	1,887
10354418	NM_028696	Obfc2a	NM_028696	0,2530	1,522
10354506	NM_133829	2210010L05Rik	NM_133829	0,2530	1,386
10362829	NM_172416	Ostm1	NM_172416	0,2530	1,402
10363921	NM_023115	Pcdh15	NM_023115	0,2530	1,739
10366446	NM_146010	Tspan8	NM_146010	0,2530	1,543
10370559	NM_027422	Mier2	NM_027422	0,2530	1,252
10372891	NM_001081037	Srgap1	NM_001081037	0,2530	1,55
10378857	NM_139128	Coro6	NM_139128	0,2530	1,236
10383853	NM_029458	Hormad2	NM_029458	0,2530	1,583
10389590	NM_025638	Gdpd1	NM_025638	0,2530	1,573
10392787	ENSMUST00000063993	4932435O22Rik	ENSMUST00000063993	0,2530	1,252
10403428	NM_172585	Larp5	NM_172585	0,2530	1,271
10405355	NM_153131	Unc5a	NM_153131	0,2530	1,252
10405872	AY149175	Zfp640	AY149175	0,2530	2,664
10411882	NM_029447	Nln	NM_029447	0,2530	1,305
10418341	NM_019583	Il17rb	NM_019583	0,2530	1,479
10423770	NM_026496	Grhl2	NM_026496	0,2530	1,287
10430993	ENSMUST00000100370	1700001L05Rik	ENSMUST00000100370	0,2530	1,402
10469609	ENSMUST00000100373	OTTMUSG00000011664	ENSMUST00000100373	0,2530	1,475
10474606	NM_029232	4930563P21Rik	NM_029232	0,2530	1,665
10478355	NM_008652	Mybl2	NM_008652	0,2530	1,431
10487011	NM_025961	Gatm	NM_025961	0,2530	1,57
10491597	---	---	ENSMUST00000039841	0,2530	1,374
10491617	ENSMUST00000057272	4932438A13Rik	ENSMUST00000057272	0,2530	1,321
10491625	ENSMUST00000057272	4932438A13Rik	ENSMUST00000057272	0,2530	1,363
10491835	NM_001040399	Larp2	NM_001040399	0,2530	1,446
10497079	NM_011196	Ptger3	NM_011196	0,2530	1,483
10498972	XM_001477059	Rbm46	XM_001477059	0,2530	1,635
10503184	NM_001081417	Chd7	NM_001081417	0,2530	1,343
10516765	NM_172702	Serinc2	NM_172702	0,2530	1,221
10516950	NM_201368	Xkr8	NM_201368	0,2530	1,225
10518857	NM_133879	Zbtb48	NM_133879	0,2530	1,247
10525365	NM_001042489	Hvcn1	NM_001042489	0,2530	1,424
10534395	NM_009903	Cldn4	NM_009903	0,2530	1,265
10540233	NM_182808	C630007B19Rik	NM_182808	0,2530	1,3
10543369	NM_153163	Cadps2	NM_153163	0,2530	1,7
10545538	NM_009819	Ctnna2	NM_009819	0,2530	1,543
10548735	NM_130447	Dusp16	NM_130447	0,2530	1,39
10571329	XM_146277	Gm501	XM_146277	0,2530	1,334
10575799	NM_172285	Plcg2	NM_172285	0,2530	1,52
10576610	NM_033620	Pard3	NM_033620	0,2530	1,386
10577954	NM_001080813	Rab11fip1	NM_001080813	0,2530	1,297

10594404	NM_016769	Smad3	NM_016769	0,2530	1,24
10596409	AM748257	EG665033	AM748257	0,2530	1,979
10599822	NM_023774	4930550L24Rik	NM_023774	0,2530	1,431
10602805	NM_001081124	Mtap7d2	NM_001081124	0,2530	2,105
10338845	---	---	---	0,2657	1,345
10342198	---	---	---	0,2657	1,253
10345411	NM_183019	Arhgef4	NM_183019	0,2657	1,328
10352267	NM_001103182	Lin9	NM_001103182	0,2657	1,319
10365933	NM_001001932	Eea1	NM_001001932	0,2657	1,421
10370242	NM_021568	Pcbp3	NM_021568	0,2657	1,537
10378420	ENSMUST00000108463	OTTMUSG00000006163	ENSMUST00000108463	0,2657	1,505
10380303	NM_028296	Car10	NM_028296	0,2657	1,334
10381445	NM_144830	Tmem106a	NM_144830	0,2657	1,23
10382532	NM_001080934	Slc16a5	NM_001080934	0,2657	1,336
10404218	NM_008156	Gpld1	NM_008156	0,2657	1,31
10418604	NM_027949	Phf7	NM_027949	0,2657	1,444
10420780	NM_021480	Tdh	NM_021480	0,2657	1,294
10425120	NM_145929	Gga1	NM_145929	0,2657	1,315
10425559	NM_145993	L3mbtl2	NM_145993	0,2657	1,323
10432640	---	---	BC127021	0,2657	1,769
10434396	NM_013852	Abcf3	NM_013852	0,2657	1,362
10447467	NM_023630	Gtf2a1l	NM_023630	0,2657	1,508
10447521	NM_146074	Tfb1m	NM_146074	0,2657	1,502
10472034	NM_177139	Lypd6	NM_177139	0,2657	1,359
10474936	NM_016907	Spint1	NM_016907	0,2657	1,402
10484888	ENSMUST00000116469	OTTMUSG00000014243	ENSMUST00000116469	0,2657	1,561
10486833	NM_145973	Ell3	NM_145973	0,2657	1,757
10491609	ENSMUST00000057272	4932438A13Rik	ENSMUST00000057272	0,2657	1,35
10494043	NM_028307	Tdrkh	NM_028307	0,2657	1,637
10495371	ENSMUST00000062935	4921525H12Rik	ENSMUST00000062935	0,2657	1,356
10503168	NM_001081417	Chd7	NM_001081417	0,2657	1,284
10503176	NM_001081417	Chd7	NM_001081417	0,2657	1,432
10510836	NM_201226	Lrrc47	NM_201226	0,2657	1,239
10510886	BC031402	Megf6	BC031402	0,2657	1,297
10512904	NM_144903	Aldob	NM_144903	0,2657	1,25
10517312	NM_025382	Tmem57	NM_025382	0,2657	1,236
10524234	NM_198306	Galnt9	NM_198306	0,2657	1,232
10532767	NM_198109	Ssh1	NM_198109	0,2657	1,28
10536931	NM_021414	Ahcyl2	NM_021414	0,2657	1,24
10555848	NM_001013616	Trim6	NM_001013616	0,2657	1,392
10563820	ENSMUST00000098414	6620401M08Rik	ENSMUST00000098414	0,2657	1,219
10570513	ENSMUST00000058886	2900016B01Rik	ENSMUST00000058886	0,2657	1,439
10577782	NM_001081187	Htra4	NM_001081187	0,2657	1,227
10581650	NM_011998	Chst4	NM_011998	0,2657	1,405
10589494	NM_013884	Cspg5	NM_013884	0,2657	1,271
10594948	NM_001081153	Unc13c	NM_001081153	0,2657	1,79
10595298	NM_001081243	Filip1	NM_001081243	0,2657	1,356
10598240	NM_001040459	Shroom4	NM_001040459	0,2657	1,366
10602367	NM_011022	Ott	NM_011022	0,2657	1,579
10602385	NM_008824	Pfkfb1	NM_008824	0,2657	1,397
10606997	NM_031381	Tex13	NM_031381	0,2657	1,65
10607336	NM_025660	Ribc1	NM_025660	0,2657	1,407
10607848	NM_019397	Egfl6	NM_019397	0,2657	1,353

Supplementary Table III.

List of the 23 down-regulated genes contained within the intersection of oocyte-specific genes. This intersection displayed a significant p value of 1.34×10^{-5} .

Down-regulated genes

Gene symbol

1. 4921515J06Rik
2. Bcl2l10
3. Bmp15
4. C87499
5. D10Wsu102e
6. E330009P21Rik
7. E330021D16Rik
8. E330034G19Rik
9. Fbxw14
10. Fbxw15
11. Gdf9
12. Gpr1
13. H1foo
14. Oog4
15. Oosp1
16. Pabpn1
17. Plat
18. Ptpn21
19. Rfp14
20. Tcl1b2
21. Tcl1b4
22. Zfp108
23. Zp1

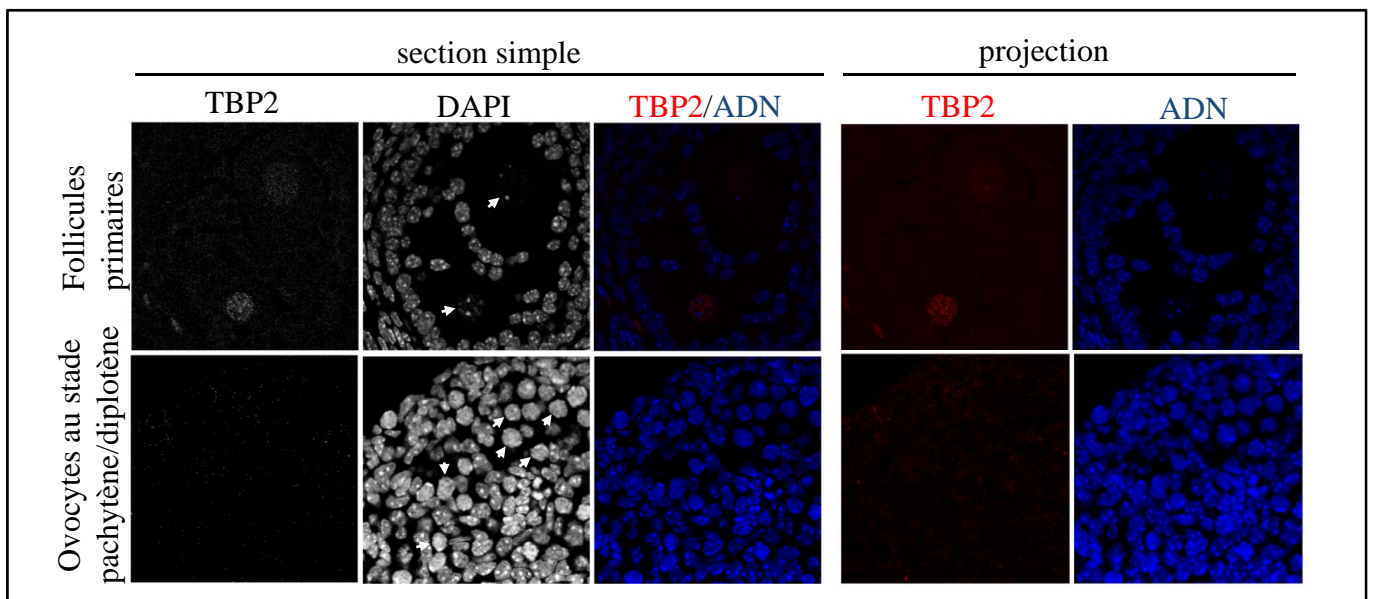


Figure 30 : TBP2 n'est pas détectable dans les ovocytes au stade pachytène tardif/diplotène précoce. Analyse par immunohistochimie de coupes d'ovaires de souris sauvages de deux semaines et à E18,5 avec un anticorps anti-TBP2 (2B12, IGBMC, Gazdag et al. 2007). L'ADN est marqué au DAPI. Les ovaires de souris de 2 semaines contiennent des ovocytes au sein de follicules au stade primaire. Les noyaux de ces ovocytes sont pointés par des flèches. Les ovaires de souris à E18,5 contiennent des ovocytes au stade pachytène tardif et diplotène précoce caractérisés par un noyau rond. Certains de ces ovocytes à ces stades sont pointés par des flèches.

C. Résultats non publiés

Nous avons ensuite continué l'étude de la protéine TBP2 selon les deux mêmes axes que précédemment. Nous avons d'une part poursuivi la caractérisation du phénotype des souris mutantes. Nous avons essayé de comprendre les causes de l'arrêt de la folliculogénèse en testant l'hypothèse selon laquelle des défauts lors de la recombinaison homologe pourraient en être responsables. Puis, nous avons caractérisé le phénotype histologique des ovaires de souris âgées. D'autre part, nous avons voulu déterminer les partenaires protéiques de TBP2 dans des ovaires sauvages. Malheureusement, malgré les multiples mises au point effectuées, aucun résultat concluant n'a été obtenu.

I. L'absence de TBP2 n'entraîne pas de défauts évidents de la prophase I de méiose

Parmi la liste des gènes dérégulés dans les ovaires de souris TBP2^{-/-}, de nombreux gènes de la famille *Xlr* font partie des gènes les plus surexprimés en l'absence de TBP2. Les protéines XLR sont exprimées dans les ovocytes au cours de la prophase I de méiose et il a été supposé qu'elles pouvaient jouer un rôle au cours de la recombinaison homologe au stade zygotène de méiose (Escalier et al., 2002). De plus, les ovocytes présentant des défauts de recombinaison homologe, et en particulier des défauts de formation des cassures double brin nécessaires à la recombinaison, meurent au cours de la folliculogénèse (Di Giacomo et al., 2005). Nous avons donc émis l'hypothèse que des défauts au cours de la recombinaison homologe pourraient être à l'origine de l'arrêt du développement des ovocytes au stade follicule secondaire, et des problèmes observés de la conformation de la chromatine des ovocytes à ce stade chez les souris TBP2^{-/-}.

1. TBP2 n'est pas détectable dans les ovaires d'embryons à E18,5

Pour déterminer si TBP2 est nécessaire au bon déroulement de la méiose et plus particulièrement de la recombinaison, nous avons d'abord vérifié si TBP2 est présent dans des ovocytes avant leur incorporation au sein de follicules. Nous avons pour cela réalisé des

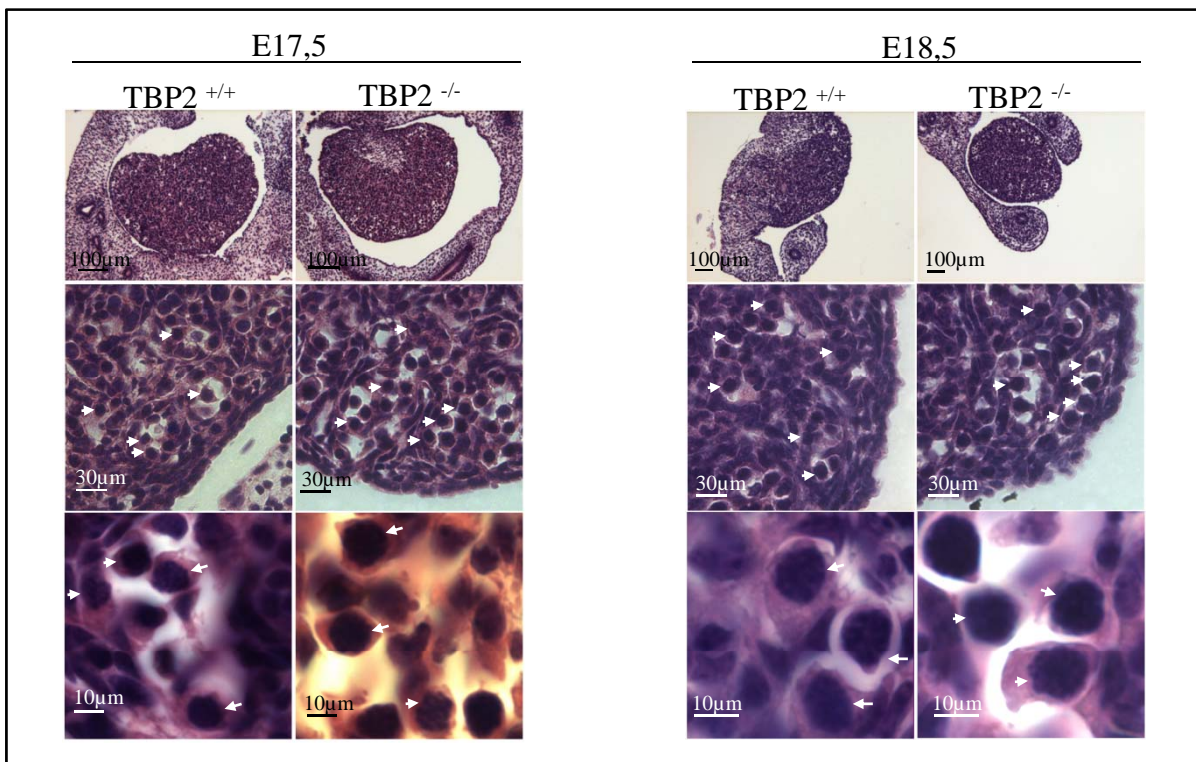


Figure 31 : Les ovaires n'exprimant pas TBP2 ne présentent pas de défauts histologiques évidents avant la naissance. Un marquage hématoxyline-éosine a été réalisé sur des coupes d'ovaires de souris mutantes et sauvages à E17,5 et E18,5. Les images ont été prises au microscope optique avec différents grossissements. Sur ces images, l'ADN est marqué en violet foncé et le cytoplasme en violet clair. Certains ovocytes au stade pachytène tardif/diplotène précoce sont pointés par des flèches.

coupes d'ovaires d'embryons à E17,5 et E18,5 que nous avons marquées en immunohistochimie avec un anticorps anti-TBP2 (anticorps 2B12, Gazdag et al., 2007).

Alors que TBP2 est détectée dans des ovocytes au stade follicule primaire dans des ovaires de souris âgées de 2 semaines, aucun marquage n'est visible dans les ovocytes en fin de stade pachytène, début de stade diplotène dans des ovaires d'embryons à E17,5 (données non montrées) et à E18,5 (figure 30a). L'expression de la protéine TBP2 semble donc débiter dans les ovocytes après le stade pachytène, avant ou au cours du stade follicule primordial (Gazdag et al., 2007). Cependant, étant donné le faible marquage observé dans les follicules primaires, stade auquel TBP2 doit être le plus fortement exprimé (Gazdag et al., 2007), il n'est pas possible d'exclure que la protéine est exprimée pendant les phases précoces de la méiose mais de manière trop faible pour être détectée en immunohistochimie.

2. La perte de TBP2 n'entraîne pas de défauts évidents dans les ovaires à E17,5 et E18,5

Afin de vérifier si TBP2, même exprimée en très faible quantité, est ou non nécessaire à la réalisation de la méiose, nous avons réalisé une analyse morphologique et histologique des ovaires de souris TBP2^{+/+} et TBP2^{-/-} à E17,5 et E18,5 (figure 31). Les ovaires de souris sauvages et mutantes ont une taille similaire. Ils présentent également tous deux des groupements d'ovocytes en cours de prophase I de méiose, ovocytes qui ne sont pas encore entourés de cellules folliculaires. La chromatine des ovocytes possède une structure similaire dans les deux groupes montrant que toutes les cellules sont au même stade méiotique. Les coupes d'ovaires n'étant pas toutes réalisées au même niveau dans les ovaires, un comptage du nombre d'ovocytes par coupe ne serait pas approprié.

La perte de TBP2 ne semble donc pas entraîner de retards ni de défauts évidents de la méiose dans les ovocytes.

3. La perte de TBP2 n'entraîne pas de défauts de la recombinaison homologue

Une des étapes importantes de la méiose est la recombinaison homologue qui a lieu au stade zygotène. Cette étape commence par la formation de cassures double brin dans les ovocytes, sites reconnaissables grâce à l'incorporation d'un variant de l'histone H2A phosphorylé, γ H2A.X. En l'absence de cassures double brin, la recombinaison homologue n'a

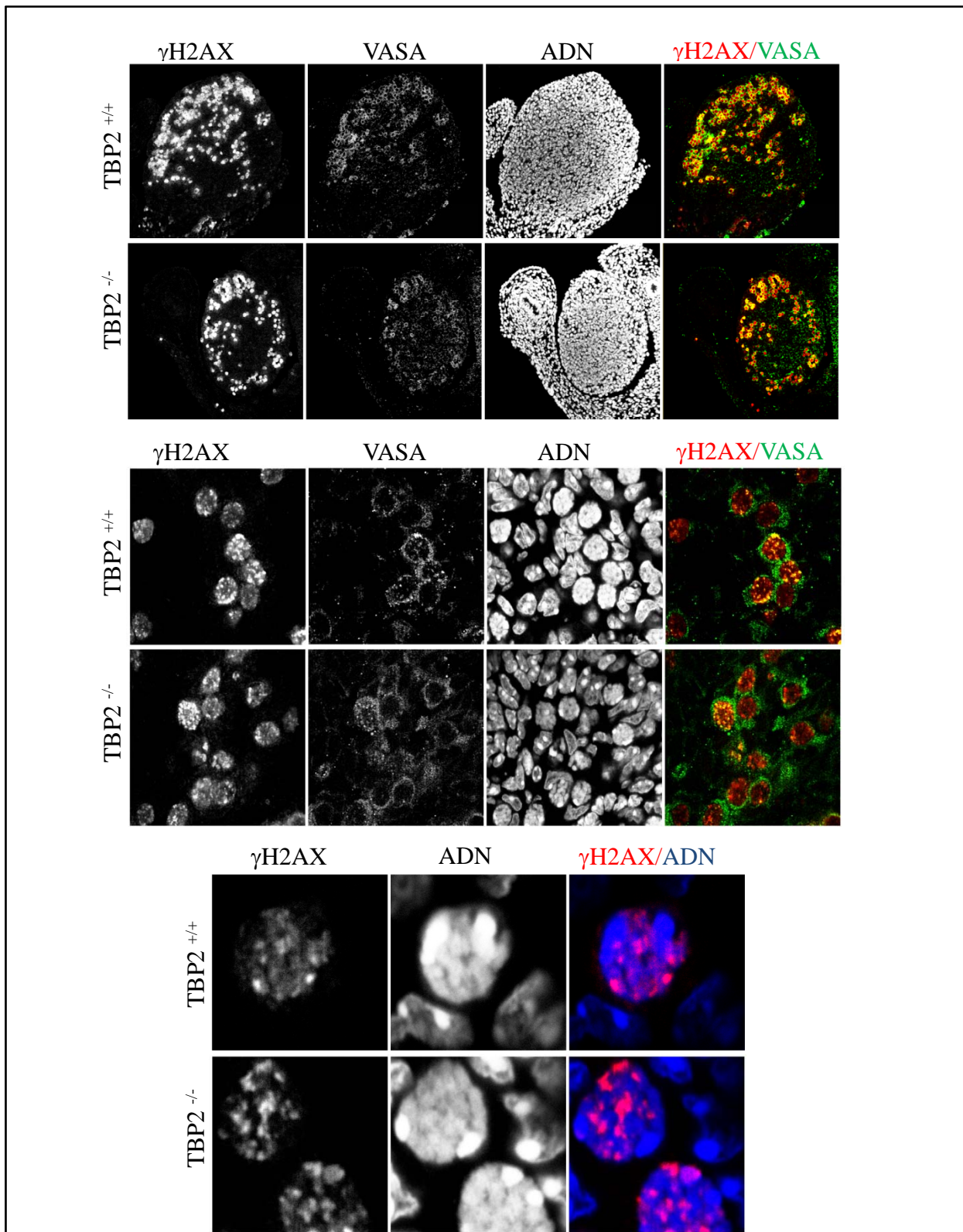


Figure 32 : Le profil de localisation de γ H2AX n'est pas différent entre de ovocytes exprimant ou non TBP2. Analyse par immunohistochimie de coupes d'ovaires de souris à E18,5 avec un anticorps anti- γ H2AX (Upstate) et un anticorps anti-VASA (don de P. Sassone-Corsi, Kotaja et al., 2006). La protéine VASA est spécifique de la lignée germinale et a été utilisé pour identifier les cellules germinales. L'ADN est marqué au DAPI. Les images ont été prises au microscope confocal à plusieurs grossissements, et des sections simples sont montrées ici.

pas lieu et la ségrégation des chromosomes au cours de la première division méiotique est très souvent incomplète, entraînant des défauts chromosomiques dans les ovocytes formés (Di Giacomo et al., 2005). D'autre part, un trop grand nombre de cassures double brin peut également entraîner potentiellement des problèmes de réparation et la formation d'ovocytes défectueux.

Pour vérifier la présence et l'abondance des cassures double brin au cours de la méiose, en l'absence de TBP2, nous avons effectué un marquage avec un anticorps spécifique de γ H2A.X dans les ovaires d'embryons sauvages et mutants à E17,5 (données non montrées) et E18,5 (figure 32), alors que la recombinaison homologue doit être terminée. Ce marquage a été complété par un marquage de la protéine VASA, exprimée dans les PGC puis dans les ovocytes au cours de la méiose avant la folliculogénèse, afin de localiser ces cellules. Aucune différence évidente n'est observée entre les ovaires TBP2^{+/+} et les ovaires TBP2^{-/-}, les deux groupes possédant γ H2A.X dans le noyau des cellules exprimant VASA. Les taux de γ H2A.X étant très différents d'un ovocyte à l'autre au sein d'un même ovaire, aucune différence n'a pu être mise en évidence quant à l'abondance de ce marqueur entre les différents ovaires.

La protéine TBP2 ne semble donc pas jouer un rôle sur la formation des cassures double brin dans les ovocytes, première étape de la recombinaison homologue nécessaire à la réalisation de la méiose.

L'arrêt de la folliculogénèse au stade follicule secondaire dans les ovocytes n'exprimant pas TBP2 ne semble donc pas dû à des défauts de réalisation des stades précoces de la méiose.

II. L'absence de TBP2 entraîne différents phénotypes ovariens chez les souris âgées

L'étude des souris TBP2^{-/-} nous a permis de mettre en évidence une stérilité totale des animaux femelles. TBP2 étant également exprimée chez l'homme, il est envisageable que cette protéine puisse jouer un rôle dans la fertilité féminine et que sa mutation chez la femme soit la cause de certains cas de stérilité. Cependant, aucune donnée n'est aujourd'hui disponible sur une potentielle mutation de TBP2 chez des femmes stériles. Dans le but de cibler des femmes possédant un phénotype proche de celui des souris TBP2^{-/-} grâce à des biopsies d'ovaires, il est nécessaire de définir le phénotype morphologique et histologique des

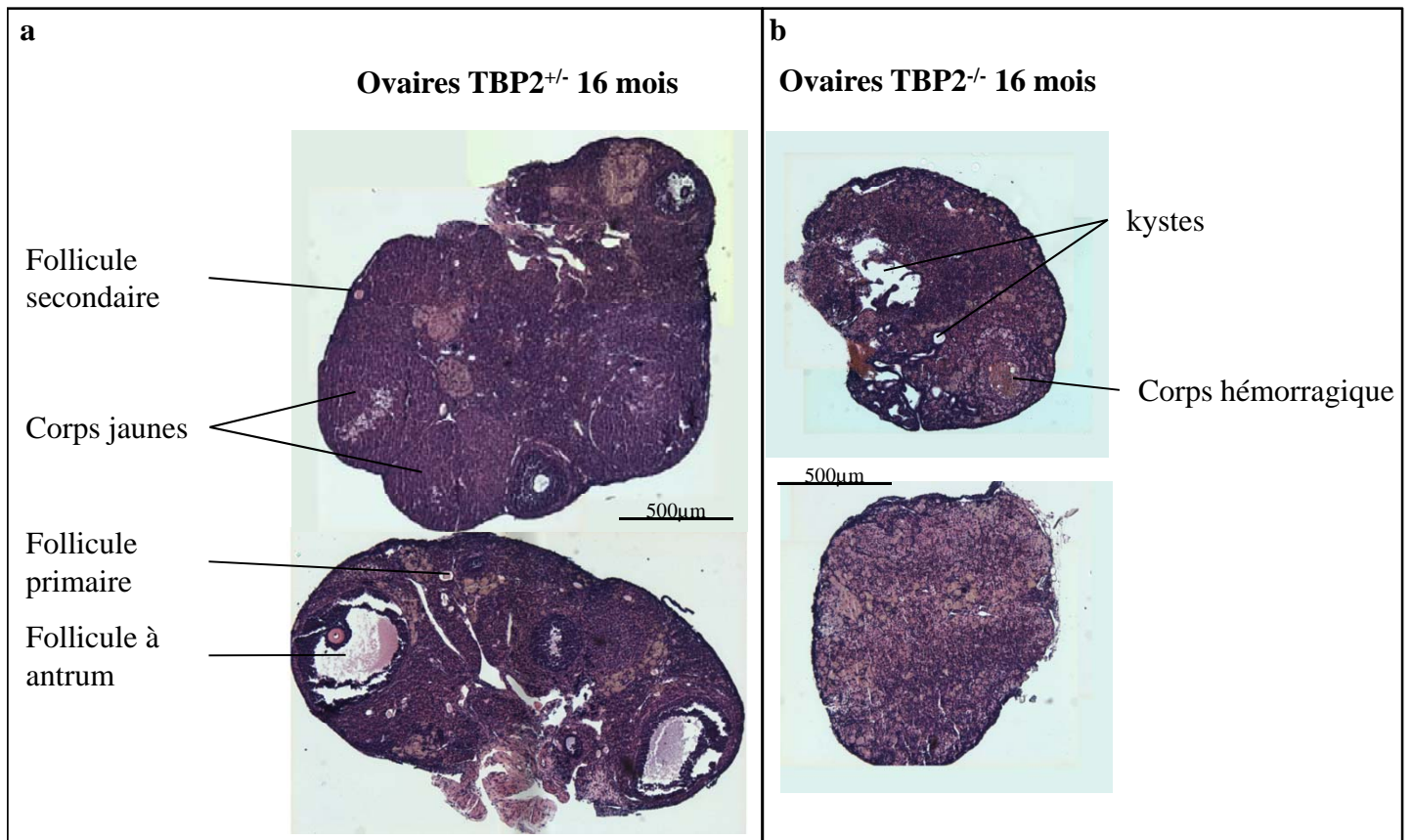


Figure 33 : Les ovaires de souris et TBP2^{-/-} âgées de 16 mois ne contiennent plus de structures folliculaires mais seulement des structures anormales. Un marquage hématoxyline-éosine a été réalisé sur des coupes d'ovaires de deux souris mutantes hétérozygotes (**a**) et deux souris homozygotes (**b**) âgées de 16 mois. Les ovaires présentés ici correspondent aux ovaires de quatre souris différentes. Les images ont été prises au microscope optique. L'ADN est marqué en violet foncé et le cytoplasme en violet clair. Les structures identifiables de chaque ovaire sont annotées.

ovaires des souris TBP2^{-/-} à un stade plus avancé que le stade 6 semaines, correspondant à un stade jeune adulte (la puberté des souris C57/BL6 ayant lieu à environ 4 semaines).

Pour déterminer les défauts à long terme engendrés par la perte de TBP2, nous avons réalisé des coupes d'ovaires de deux souris TBP2^{+/-} (utilisées comme contrôles) et deux souris TBP2^{-/-} âgées de 16 mois. Comme à 6 semaines, les ovaires de souris TBP2^{-/-} sont trois à quatre fois plus petits que ceux des souris contrôles. Les coupes ont ensuite été colorées à l'hématoxiline-éosine et leur histologie a été analysée (figure 33). Alors que tous les stades folliculaires sont encore représentés dans les ovaires des deux souris TBP2^{+/-}, aucune structure folliculaire n'est présente dans les ovaires n'exprimant pas TBP2. L'absence de corps jaune sur toutes les coupes TBP2^{-/-} analysées confirme le fait qu'aucune ovulation n'a eu lieu chez ces souris. Alors que les ovaires d'une des deux souris mutantes ne présentent aucune structure remarquable, les ovaires de la deuxième souris contiennent des kystes ainsi qu'un corps hémorragique. Ils peuvent provenir de la transformation de follicules qui auraient dû dégénérer.

Les caractéristiques qui pourraient être recherchées chez des femmes âgées, potentiellement porteuses d'une mutation de TBP2 sont donc la présence de kystes, hémorragiques ou non. Cependant, ce phénotype ne semble pas être pénétrant à 100%. En revanche, il est important de noter qu'aucune tumeur n'a été détectée, non seulement dans les ovaires des souris analysées pour cette expérience, mais aussi chez aucune souris, tout âge confondu, analysée au cours de ma thèse. La mutation de la protéine TBP2 ne semble donc pas avoir de lien avec l'apparition des cancers de l'ovaire.

III. Etude des partenaires protéiques de TBP2 dans les ovocytes, mises au point expérimentales

Afin de comprendre plus précisément les mécanismes d'action de la protéine TBP2 ainsi que les différences fonctionnelles entre TBP et TBP2, nous avons voulu définir les partenaires protéiques de TBP2 *in vivo*, dans l'ovocyte afin de les comparer aux sous-unités du complexe TFIID canonique. L'étude du ou des complexes contenant TBP2 nécessite la mise au point d'une expérience d'immunoprécipitation (IP) de protéines sur des extraits nucléaires d'ovaires, contenant une très faible quantité de matériel. Cette étude est réalisée en collaboration avec l'équipe du Dr Lászlò Tora à l'IGBMC et les nombreuses IP ont été effectuées par Elisabeth Scheer. J'ai pour ma part réalisé les extraits nucléaires d'ovaires de

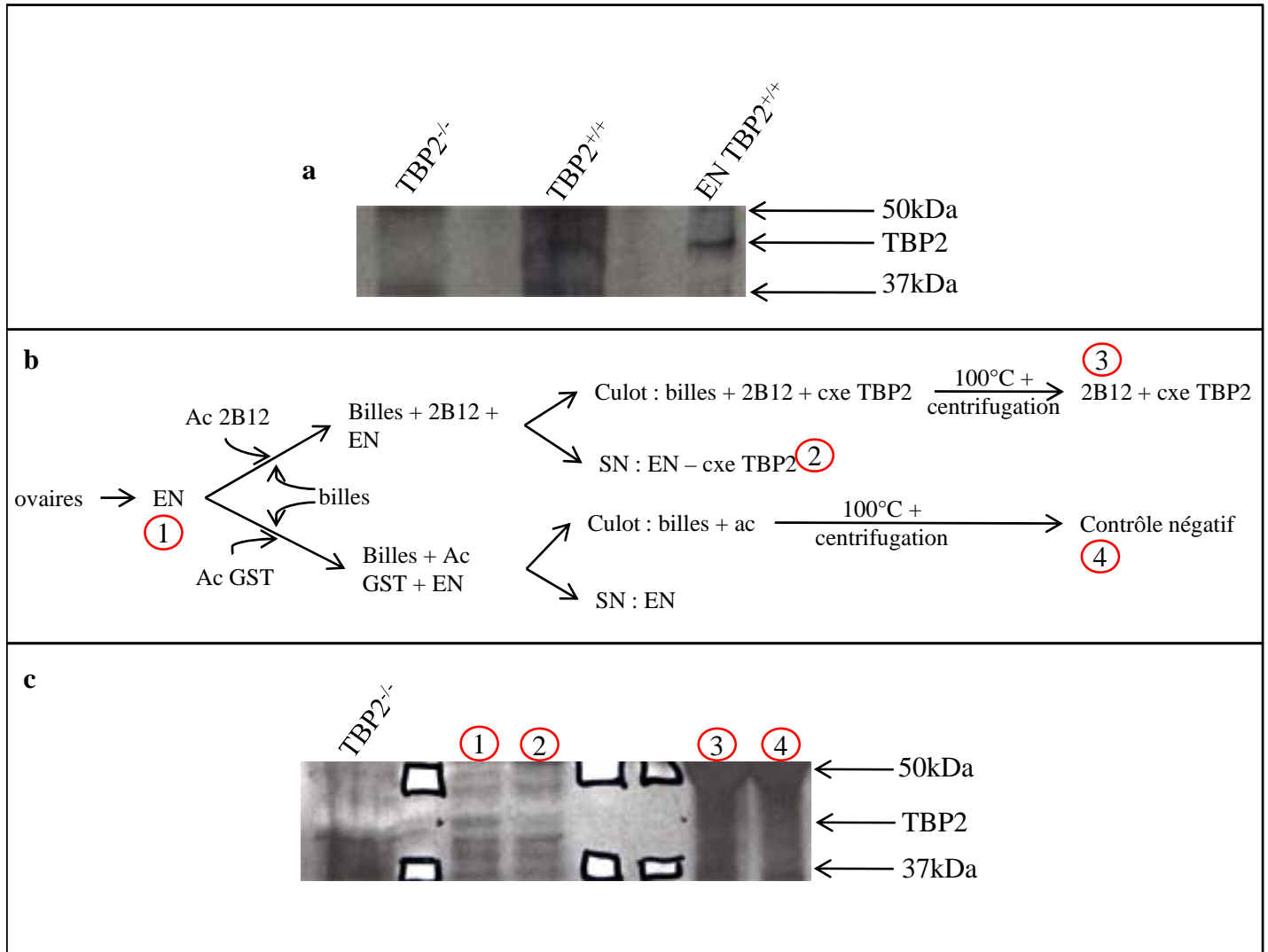


Figure 34 : Schématisation et résultats de différentes mises au point de l'expérience d'immunoprécipitation de TBP2 sur les extraits nucléaires d'ovaires de souris sauvages. **a** : La protéine TBP2 est présente dans les extraits nucléaires d'ovaires de souris sauvages. Analyse par western blot d'extraits totaux de 3 ovaires de souris $TBP2^{-/-}$ (ovaires $TBP2^{-/-}$), de 2 ovaires de souris $TBP2^{+/+}$ (ovaires $TBP2^{+/+}$) et d'un extrait nucléaire d'ovaires de souris $TBP2^{+/+}$ (le dépôt correspond à 2 ovaires) (EN ovaires $TBP2^{+/+}$). La protéine TBP2 est détectée avec l'anticorps monoclonal anti-TBP2 2B12, et la bande correspondant à TBP2 est fléchée. **b**. Schématisation de l'expérience d'immunoprécipitation de TBP2 dans des extraits nucléaires d'ovaires. La moitié de l'extrait nucléaire (EN) est utilisée pour une expérience d'IP avec l'anticorps anti-TBP2 2B12, l'autre moitié est utilisée pour un contrôle négatif, une IP avec un anticorps reconnaissant la GST (Ac GST). Après rinçages, les billes liées à l'anticorps et au complexe (billes+2B12+cxe) sont récupérées et mises à bouillir pour que l'anticorps et le complexe s'en détachent (2B12 + cxe). **c**. Après IP anti-TBP2, l'extrait nucléaire d'ovaires est dépleted en protéine TBP2. Western blot anti-TBP2 sur des extraits totaux d'ovaires $TBP2^{-/-}$ ($TBP2^{-/-}$), sur 10% des extraits nucléaires d'ovaires sauvages utilisés pour l'IP (1), sur ces mêmes extraits après immunoprécipitation avec un anticorps anti-TBP2 (2), sur le produit d'IP retenu sur les billes (3) et sur le contrôle négatif, IP avec un anticorps anti-GST (4).

souris nécessaires aux différentes expériences et les contrôles de la qualité de ces extraits. Malheureusement, ce projet ambitieux n'a donné aucun résultat concluant pour le moment. Cependant, je vais résumer, dans cette dernière sous-partie, les différentes grandes étapes de mise au point réalisées pour cette expérience.

1. Préparation des extraits nucléaires d'ovaires

Afin d'obtenir une quantité suffisante de complexes contenant TBP2, des extraits nucléaires ont été réalisés à partir d'ovaires de souris sauvages âgées de 10 à 15 jours. A cet âge, les ovaires possèdent de nombreux follicules primaires, stade auquel la quantité de TBP2 est la plus importante dans les ovocytes (Gazdag et al., 2007). Les protéines nucléaires sont extraites selon une méthode adaptée à des petites quantités de tissus (inférieures à 100 mg). Les premières extractions ont été effectuées avec 50 à 80 ovaires, donnant un rendement, selon l'âge des souris, compris entre 100 et 200 μg de protéines par extraction. Les extractions suivantes, utilisées pour la spectrométrie de masse, ont été réalisées sur 200 ovaires, plus jeunes, donnant entre 200 μg et 400 μg de protéines par extraction. La présence de TBP2 a été vérifiée dans les extraits nucléaires par western blot (figure 34a).

2. Différentes mises au point de l'immunoprécipitation

La première expérience d'IP a été mise au point directement sur les extraits nucléaires d'ovaires avec l'anticorps monoclonal 2B12 anti-TBP2 caractérisé au laboratoire (Gazdag et al., 2007). Cet anticorps est capable d'immunoprécipiter TBP2 mais ne peut pas être élué des billes après immunoprécipitation. Les billes doivent donc être bouillies en présence de DTT puis centrifugées pour récupérer le surnageant contenant à la fois les complexes retenus par l'anticorps mais également les anticorps eux-mêmes et de nombreuses protéines retenues aspécifiquement sur les billes (figure 34b). Après plusieurs expériences de mise au point une déplétion de la protéine TBP2 a pu être mise en évidence dans le surnageant après IP (figure 34c). Cependant, sur le western blot, après que les billes ont été bouillies, le bruit de fond trop important rend presque invisible la bande correspondant à TBP2. Malgré l'importance du bruit de fond, cette expérience a été répétée sur 150 μg de protéines et les produits immunoprécipités ont été envoyés pour une analyse en spectrométrie de masse. TBP2 n'a pas été détectée parmi les protéines immunoprécipitées dans cette expérience. Une expérience similaire avait été menée en parallèle avec un anticorps anti-TBP, connu pour être très

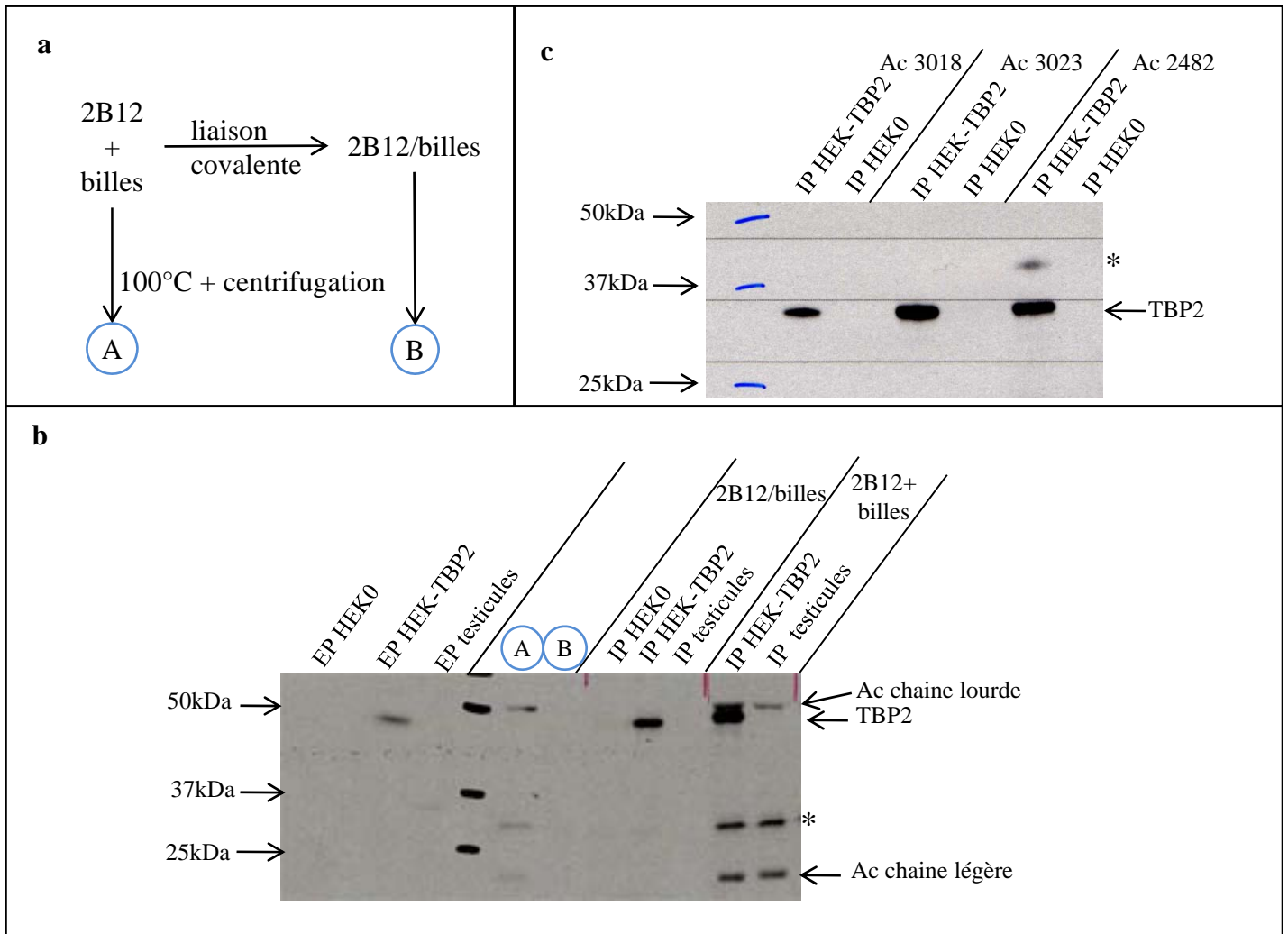


Figure 35 : Schématisation et résultats de différentes mises au point de l'expérience d'immunoprécipitation de TBP2 sur des cellules HEK exprimant transitoirement TBP2. **a.** Schématisation de la réalisation d'une liaison covalente entre les billes et l'anticorps anti-TBP2 2B12. Après ébullition, les anticorps liés de manière covalente aux billes ne s'en détachent pas (A sur figure **b**) alors que les anticorps non liés sont retrouvés dans le surnageant (B sur figure **b**). **b.** Une IP est réalisée sur des extraits totaux de cellules HEK non transfectées (HEK0), de cellules HEK exprimant transitoirement TBP2 (HEK-TBP2), et de testicules (testicules). TBP2 est immunoprécipitée grâce à l'anticorps anti-TBP2 lié (2B12/billes) ou non (2B12 + billes) de manière covalente aux billes. La présence de TBP2 dans les extraits protéiques de départ (EP) et dans les produits d'IP après ébullition (IP) est testée par western blot. **c.** Des IP suivies d'éluion sont réalisées sur des extraits nucléaires de cellules HEK non transfectées (HEK0) ou de cellules HEK exprimant transitoirement TBP2 (HEK-TBP2) avec 3 anticorps polyclonaux anti-TBP2 (Ac 3018, Ac 3023, Ac 2482, IGBMC). La présence de la protéine TBP2 dans les produits d'éluion des IP est testée par western blot avec l'anticorps anti-TBP2 2B12. * bande aspécifique

efficace en IP. Les résultats de cette IP anti-TBP étant également très mauvais, TBP n'apparaissant que très tard dans la liste des protéines détectées, nous en avons déduit que le problème pouvait venir d'un bruit de fond trop important, cachant le signal spécifique.

Afin de réduire le bruit de fond, l'anticorps 2B12 a été lié de manière covalente aux billes, l'empêchant ainsi de s'en détacher lorsque les billes sont bouillies (figure 35a et b). Les mises au point de cette expérience ont été réalisées sur des extraits cellulaires totaux de cellules HEK exprimant transitoirement une protéine TBP2 fusionnée avec une étiquette Flag (HEK-TBP2) et sur des cellules HEK non transfectées (HEK0) comme contrôle. L'étape de liaison des anticorps aux billes permet bien la disparition des anticorps du produit obtenu après ébullition à la fin de l'IP (figure 35b). Alors que les expériences de mise au point ont été réalisées sur 1,5 mg de protéines totales extraites des cellules HEK, l'IP finale anti-TBP2 a été réalisée sur les extraits nucléaires de 120 ovocytes correspondant à 200 µg de protéines. Après analyse des résultats en spectrométrie de masse, TBP2 ne faisait toujours pas partie des protéines immunoprécipitées dans l'expérience sur les ovocytes.

A ce stade, nous avons décidé de changer l'anticorps avec lequel les IP sont réalisées. Des nouveaux anticorps polyclonaux anti-TBP2 réalisés au laboratoire ont été testés sur des extraits totaux de cellules HEK-TBP2 et HEK0. Deux anticorps pouvant être élués après l'IP ont été retenus, les anticorps 2482 et 3023 (figure 35c). Deux IP ont été effectuées avec ces nouveaux anticorps sur des extraits nucléaires d'ovaires contenant 300 µg de protéines et les produits d'éluion ont été analysés en MudPIT (multidimensional protein identification technology). Dans la première expérience, alors que des peptides correspondant à TBP2 ont été identifiés dans les produits d'éluion de l'IP, ils ont également été identifiés dans le contrôle négatif (IP réalisée avec un anticorps anti-GST). Cette première expérience n'est donc pas analysable. La deuxième expérience, quant-à elle, n'a encore une fois pas mis en évidence la protéine TBP2 dans les produits d'éluion de l'IP.

Aucun partenaire protéique de TBP2 n'a donc pu être mis en évidence dans les ovocytes de souris. Les pistes qui pourraient être envisagées pour définir ces partenaires seront détaillées dans la partie discussion.

D. Discussion et perspectives

L'ensemble de ce travail de thèse nous a permis de commencer à comprendre le rôle de la protéine TBP2 *in vivo* chez la souris. Le profil d'expression très limité de cette protéine ainsi que sa faible abondance dans les ovocytes rendent son étude très difficile. De nombreux points concernant ses mécanismes d'action et ses rôles précis au sein de la cellule restent encore à élucider. Dans cette partie, je discuterai les résultats que nous avons obtenus au vu des données disponibles dans la littérature sur les rôles et les mécanismes d'action connus de TBP2 et de ses homologues dans différents organismes et je proposerai des expériences pour répondre à certaines questions encore posées sur cette protéine.

I. Etude du phénotype macroscopique des souris TBP2^{-/-}

1. L'ovaire, seul organe touché par la perte de TBP2 ?

Afin de comprendre le rôle de la protéine TBP2 *in vivo* chez les mammifères, une souris exprimant une forme tronquée de TBP2 a été générée au laboratoire. La stérilité totale des souris TBP2^{-/-} a été mise en évidence. Elle s'explique par une absence de maturation des ovocytes après le stade follicule secondaire et donc une absence d'ovulation. Ces résultats sont en accord avec le profil d'expression de TBP2 dans les ovaires de souris publié précédemment (Gazdag et al., 2007). En effet, TBP2 est détectable dans les ovocytes du stade follicule primordial au stade follicule à *antrum*. Bien que TBP soit également présent dans ces cellules jusqu'au stade follicule secondaire, nous avons montré que TBP2 était le seul à être nécessaire au déroulement correct de la folliculogenèse et à la fertilité des souris femelles.

Cependant, d'autres études chez la souris suggèrent que l'ovocyte ne serait pas le seul type cellulaire exprimant TBP2, mais que cette protéine serait également responsable de la transcription de certains gènes spécifiques des cellules de la lignée musculaire et en particulier des myotubes (Deato et al., 2008) (Deato and Tjian, 2007). Nous avons donc quantifié l'expression de certains gènes dans les muscles des souris mutantes. Aucune différence significative n'a pu être mise en évidence entre la quantité d'ARNm de certains gènes spécifiques des cellules musculaires, en particulier de la myogénine montrée comme étant une cible directe de TBP2 (Deato et al., 2008), entre les souris mutantes et les souris sauvages. De

plus, les souris mutantes ne semblent pas présenter de problèmes de motricité et sont morphologiquement indistinguables des souris sauvages. La plupart des conclusions sur le rôle de TBP2 dans les cellules musculaires ont été tirées à partir d'expériences réalisées dans les cellules C2C12, des myoblastes de souris. Ces cellules sont dérivées à partir de cellules musculaires satellites adultes, et sont utilisées en général comme modèles pour comprendre le développement des cellules musculaires squelettiques. Les cellules satellites sont des cellules quiescentes entourant les cellules musculaires squelettiques adultes et capables de proliférer rapidement et de se différencier en cellules musculaires pour régénérer un muscle squelettique lésé (Appell et al., 1988) (Charge and Rudnicki, 2004). Il est possible que le rôle de TBP2 mis en évidence dans les cellules C2C12 ne soit pas nécessaire à la différenciation des cellules musculaires pendant le développement mais plutôt au cours d'un processus de régénération chez l'adulte. Dans le but de confirmer ou d'infirmer cette hypothèse, des expériences de lésion musculaire sont en cours de réalisation sur les souris TBP2^{-/-} à l'Institut Clinique de la Souris (ICS) à l'IGBMC. La capacité des cellules satellites à se réactiver et à se différencier en l'absence de TBP2 sera alors testée.

De plus, afin de mettre en évidence d'autres problèmes moins évidents que la stérilité chez les souris TBP2^{-/-}, de nombreux tests (cognitifs, moteurs, sanguins) sont en cours de réalisation à l'ICS. Les résultats nous permettront de savoir si les ovaires, seuls organes où l'ARNm de TBP2 a été détecté par RTPCR (Xiao et al., 2006), sont réellement les seuls organes où cette protéine, même en très faible quantité, est essentielle.

2. Le rôle des homologues de TBP est-il conservé au cours de l'évolution ?

Le phénotype des mutants TBP2^{-/-} a précédemment été analysé chez le xénope et le poisson zèbre. Alors que les souris TBP2^{-/-} ne présentent aucun défaut évident de développement, la délétion de TBP2 chez le xénope et le poisson zèbre est létale au stade embryonnaire (Bartfai et al., 2004) (Jallow et al., 2004). La différence entre la souris et les autres vertébrés est d'autant plus importante que chez la souris, l'expression de TBP2 dans les embryons est délétère pour leur développement. Ces résultats sont en accord avec des profils d'expression différents de TBP2 dans les trois organismes étudiés, TBP2 étant exprimé dans les embryons de xénope et de poisson zèbre alors qu'elle disparaît au moment de la fécondation chez la souris (Gazdag et al., 2007).

Un parallèle peut être fait entre ces résultats et les conséquences de la déplétion de la protéine TLF dans ces mêmes organismes. Alors que la déplétion de TLF entraîne un arrêt

précoce du développement au cours de l'embryogenèse du xénope et du poisson zèbre (Muller et al., 2001) (Veenstra et al., 2000), la plupart des souris TLF^{-/-} sont viables et les souris mâles TLF^{-/-} sont stériles (Martianov et al., 2001). La réduction du champ d'activité des facteurs de la famille de TBP à la lignée germinale chez les mammifères ne semble donc pas être spécifique à TBP2.

La question de la pertinence de ce phénomène peut être posée. Est-elle à mettre en relation avec l'apparition ou la disparition d'autres facteurs généraux de la transcription dans ces mêmes types cellulaires au cours de l'évolution ? Dans les embryons de xénope et de poisson zèbre, la MZT est beaucoup plus tardive que chez la souris. Alors que la ZGA majeure a lieu au stade 2-cellules chez la souris, elle a lieu après plus de 10 cycles de division chez le poisson zèbre et le xénope. Existe-t-il un lien entre l'importance des facteurs de la famille de TBP au cours du développement et le moment de la transition entre un contrôle maternel et embryonnaire de la transcription dans l'embryon ? Dans ce cas, qu'en est-il de l'importance de la protéine TBP2 chez l'homme où la ZGA majeure a lieu au stade 8-cellules (Schultz and Heyner, 1992)? Répondre à ces questions permettrait de comprendre les changements du profil d'expression et de l'importance de ces protéines au cours de l'évolution et peut-être d'en déduire certains de leurs mécanismes d'action.

II. Rôles de la protéine TBP2 dans l'ovocyte

1. Rôle de TBP2 sur la transcription

Le domaine C-terminal de TBP2 est identique à plus de 80% à celui de TBP et TBP2 est capable de lier les complexes TFIIA et TFIIB ainsi que la boîte TATA *in vitro* (Bartfai et al., 2004). Cependant, les données concernant un rôle direct de TBP2 sur la transcription *in vivo* sont rares voire inexistantes chez la souris. Grâce à une expérience de ChIP sur des ovaires de souris sauvages, nous avons pu mettre en évidence la liaison de TBP2 au promoteur de deux gènes spécifiques des ovocytes, gènes sous-exprimés dans les mutants TBP2^{-/-}. Ces résultats suggèrent donc que TBP2 doit jouer un rôle direct dans la transcription active de gènes spécifiques de l'ovocyte. Cependant, une analyse à l'échelle du génome serait nécessaire pour confirmer et généraliser ces résultats. L'expérience de ChIP sur les ovaires de souris âgées de 2 semaines a demandé de nombreuses mises au point et de nombreux points restent limitant dans l'optique d'une étude à l'échelle du génome. Tout d'abord, la quantité

de matériel disponible pour le ChIP est très faible. Les études concernant le nombre d'ovocytes par ovaire de souris donnent des résultats très différents mais elles estiment entre 2000 et 10000 le nombre de follicules primordiaux et entre 1000 et 3000 le nombre de follicules primaires par ovaire de souris au début de la vie reproductive (Myers et al., 2004). TBP2 étant le plus abondant au stade follicule primaire et les expériences de ChIP étant réalisées sur une moyenne de 10 ovaires, on peut considérer qu'environ 30000 cellules sont utilisées par ChIP. Pour réaliser des ChIP sur un petit nombre de cellules, une technique consiste à utiliser d'autres cellules, ne contenant pas la protéine d'intérêt, servant d'entraîneur lors de la précipitation (O'Neill et al., 2006). Dans ce but, et également pour des raisons pratiques, les ovaires ont été utilisés entiers dans les expériences de ChIP. Cependant, bien que TBP2 ne soit exprimée que dans les ovocytes, le grand nombre de cellules somatiques utilisées pour les ChIP induisent un bruit de fond rendant impossible l'étude à l'échelle du génome. De plus, l'anticorps anti-TBP2 utilisé pour le ChIP n'est pas assez efficace sur une faible quantité de matériel. Alors que l'efficacité de cet anticorps a été prouvée sur des IP sur des extraits de cellules HEK-TBP2, elle s'est avérée nulle sur les ovaires. Pour les expériences de ChIP, la faible efficacité de l'anticorps m'a contrainte à refaire les expériences de nombreuses fois pour obtenir un résultat reproductible et fiable. Le problème posé par la faible quantité de matériel peut difficilement être pallié étant donné qu'augmenter le nombre d'ovaires utilisé par expérience revient également à en augmenter le bruit de fond. Pour augmenter l'efficacité de la précipitation de la protéine TBP2 et donc de la chromatine associée, il faudrait donc utiliser un anticorps très efficace. Alors qu'aucun anticorps anti-TBP2 testé ne semble être très efficace en immunoprécipitation sur les ovaires, une possibilité serait de générer une lignée de souris où la protéine TBP2 serait fusionnée à une étiquette HA par exemple, l'anticorps anti-HA étant très efficace en ChIP et IP. Ce projet, à long terme, pourrait permettre de mettre en évidence les sites de liaison de TBP2 à l'ADN, et donc non seulement de connaître ses cibles directes et de confirmer son rôle direct sur la transcription mais également de définir une potentielle séquence consensus de liaison. La comparaison de cette séquence consensus de TBP2 avec la boîte TATA, séquence consensus de liaison de TBP, permettrait de comprendre un peu mieux la spécificité d'action de ces protéines.

2. Rôle de TBP2 sur la répression de la transcription

L'étude du transcriptome des ovaires de souris TBP2^{-/-} révèle que parmi les gènes dérégulés, la moitié est surexprimée et l'autre moitié sous-exprimée. Un tel résultat peut

paraître étrange pour un facteur supposé être impliqué dans l'initiation de la transcription. De plus, de nombreux gènes spécifiques de la lignée germinale mâles sont surexprimés en l'absence de TBP2. Ces observations laissent penser que, en plus du rôle que peut avoir TBP2 dans l'initiation de la transcription de gènes spécifiques des ovocytes, elle pourrait également être responsable, directement ou indirectement, de la répression des gènes spécifiques d'autres types cellulaires. Des résultats similaires sont obtenus lors de la diminution de l'expression de TBP dans des embryons de poisson zèbre (Ferg et al., 2007). Parmi les gènes dérégulés dans les embryons exprimant peu de TBP, la moitié est surexprimée et l'autre moitié sous-exprimée. De plus, des expériences menées avec un gène rapporteur *Gfp* (*green fluorescent protein*) sous la dépendance de promoteurs de gènes dérégulés en l'absence de TBP montrent que TBP peut être, de manière directe, un activateur ou un répresseur de la transcription selon le promoteur considéré. Les mécanismes responsables de l'action répressive de TBP ne sont pas connus. Un rôle similaire de TBP2 pourrait donc être envisageable dans les ovocytes et des expériences *in vitro* d'expression d'un gène rapporteur en présence de TBP2 pourraient être menées, de la même manière que pour TBP, dans le but de le démontrer. De même, TLF semble pouvoir jouer un rôle d'activateur et d'inhibiteur de la transcription (Tanaka et al., 2007). Des données complètes de CHIP à l'échelle du génome permettraient, en lien avec les résultats de l'étude transcriptomique obtenues dans les ovaires mutants, de conclure quant au caractère activateur ou inhibiteur de TBP2.

3. Rôle de TBP2 sur la structure chromatinienne

Dans les ovocytes sauvages, la diminution progressive de la transcription dans les ovocytes avant la phase GVBD est accompagnée d'une condensation de la chromatine passant d'un ovocyte de type NSN à un ovocyte de type SN silencieux transcriptionnellement (Bachvarova, 1985) (Zuccotti et al., 2005). Nous avons montré que les ovocytes n'exprimant pas TBP2 n'acquièrent pas le degré maximum de condensation. Encore une fois, le caractère direct ou indirect de ce phénotype n'a pas pu être déterminé. Cependant, il est intéressant de noter les spermatocytes n'exprimant pas TLF présentent des défauts de réorganisation de l'hétérochromatine péricentromérique pendant leur maturation. Dans les ovocytes, les régions concernées par cette absence de condensation n'ont pas été déterminées. Des études par FISH sur ADN pourraient être envisagées afin de vérifier la localisation des régions péricentromériques en présence et en l'absence de TBP2 dans les ovocytes. Les marques

chromatiniennes associées à la condensation comme H3K9me3 ou HP1 pourraient également être analysées par immunohistochimie pour mieux comprendre ce phénotype.

Ces défauts de condensation peuvent, de plus, être mis en relation avec la surexpression d'un grand nombre de gènes observée dans les ovaires mutants. Ces deux processus peuvent présenter une cause commune, ou on peut aussi penser que l'absence de condensation induit des défauts de répression de certains gènes dans l'ovocyte.

III. Pourquoi remplacer TBP par TBP2 ?

In vitro, TBP2 semble avoir des propriétés proches de celles de TBP. De plus, chez le xénope, l'expression de TBP est capable de remplacer celle de TBP2 dans des ovocytes (Akhtar and Veenstra, 2009). La protéine TBP2 possède-t-elle les mêmes caractéristiques, les mêmes mécanismes d'action que TBP ? Quelles sont les spécificités de TBP2 par rapport à TBP ? Le phénomène de remplacement de TBP par TBP2 dans les ovocytes étant commun à tous les vertébrés, il doit donc être important. Pourquoi est-ce important de remplacer TBP par TBP2 dans les ovocytes ?

1. TBP et TBP2, deux protéines pour un même rôle ?

Tout d'abord, l'équivalence entre les protéines TBP et TBP2 n'est pas totale. Nous avons en effet montré que, alors que TBP est requise pour le développement embryonnaire et que sa surexpression n'entraîne pas de défauts de développement jusqu'au stade blastocyste, la surexpression de TBP2 après la fécondation entraîne un arrêt du développement embryonnaire entre les stades 2-cellules et 8-cellules. TBP2 ne peut donc pas se substituer à TBP au cours du développement précoce. Une expérience de sauvetage des ovocytes TBP2^{-/-} par la surexpression de TBP pourrait être envisagée pour tester le potentiel de TBP à se substituer à TBP2 chez la souris. L'ADNc de TBP devrait donc être inséré dans le génome sous la dépendance du promoteur de *Tbp2*. L'absence de protéine TBP dans l'ovocyte étant régulée traductionnellement (Xiao et al., 2006), les régions 5'UTR et 3'UTR de TBP devront également être remplacées par celles de TBP.

Si on compare le profil d'expression de TBP avec celui de TBP2 et de TRF2 chez la souris, on remarque que TBP régule la transcription des cellules se divisant par mitose alors que TBP2 et TLF sont exprimées dans des cellules en cours de méiose, les cellules

germinales, ou dans des cellules ne se divisant plus, les myotubes. Nous avons montré que TBP2 n'est pas nécessaire aux premiers stades de la méiose dans les ovocytes suggérant donc que TBP2 n'est pas responsable de la transition entre des cellules en mitose et des cellules en méiose lors de la formation des ovogonies. Le plus probable reste donc que TBP2 permette de mettre en place ou de maintenir un programme d'expression génique spécifique d'un type cellulaire particulier, comme suggéré par l'analyse transcriptomique des ovaires TBP2^{-/-}.

2. Quels sont les mécanismes sous-tendant cette spécificité ?

Si TBP2 est capable d'induire l'expression de gènes spécifiques, les mécanismes sous-tendant cette spécificité ne sont pas connus. Une hypothèse serait que TBP2 permet l'assemblage d'un complexe TFIID particulier, reconnaissant les promoteurs des gènes spécifiques des ovocytes et activant leur transcription. Cette hypothèse est celle retenue dans les myotubes où TBP2 en association avec TAF3 seulement serait responsable de la transcription du gène *myogénine* (Deato et al., 2008). On peut également supposer que TBP2 peut être recrutée aux promoteurs de certains gènes grâce à l'action de facteurs de transcription, ou d'autres facteurs activateurs spécifiques des ovocytes.

Pour permettre de mettre en évidence des différences entre les protéines associées à TBP et les protéines associées à TBP2, la finalisation de l'IP des complexes contenant TBP2 est nécessaire. Les problèmes rencontrés lors de la mise au point de cette expérience sont les mêmes que ceux rencontrés lors de la mise au point du ChIP : une trop faible quantité de matériel associée à un problème d'efficacité des anticorps testés. La réalisation de souris exprimant une version étiquetée de TBP2 pourrait également permettre d'améliorer cette expérience. Cependant, des précautions doivent être prises pour vérifier que l'étiquette fusionnée à TBP2 n'empêche ni son interaction avec l'ADN, ni son interaction avec ses partenaires protéiques.

Une manière détournée de rechercher les partenaires potentiels de TBP2 dans l'ovocyte serait d'abord de vérifier l'expression des TAF canoniques et de leurs variants connus dans les ovocytes, à différents stades folliculaires et de définir leur colocalisation avec TBP2 par immunohistochimie. TAF4 est par exemple remplacée dans les ovocytes par son homologue TAF4b . Cette protéine est donc un bon candidat pour être un partenaire de TBP2. Cependant, le phénotype des souris TAF4b^{-/-} a été déterminé, et bien que stériles, les souris ne présentent pas les mêmes défauts que les souris TBP2^{-/-} au cours de la folliculogénèse suggérant que ces deux protéines n'interviennent pas exclusivement dans la régulation des

mêmes gènes (Falender et al., 2005). Une fois les partenaires potentiels de TBP2 déterminés et dans le cas où des modifications post-traductionnelles de ces protéines ne sont pas nécessaires, leur interaction avec une protéine TBP2 recombinante pourrait être testée *in vitro*.

TBP2, associée à des protéines particulières, pourrait alors reconnaître des séquences génomiques, des promoteurs spécifiques et ainsi induire ou réprimer l'initiation de la transcription de gènes spécifiques, non régulés par TBP.

TBP2 peut également, même associée aux mêmes TAF que TBP, avoir une séquence de liaison à l'ADN spécifique, différente de la boîte TATA canonique. Bien que TBP2 soit capable de lier la boîte TATA *in vitro*, la séquence consensus de liaison de TBP2 à l'ADN *in vivo* devra être déterminée par ChIP dans les ovocytes.

IV. TBP2 et la stérilité féminine

Le gène *Tbp2* est présent chez l'homme. Son profil d'expression n'a été publié que par une seule étude où TBP2 est présentée comme une protéine ubiquitaire chez l'homme et la souris (Persengiev et al., 2003). Avec ce que l'on sait actuellement de TBP2 chez la souris, les données présentées par cette étude peuvent être remises en cause. Cependant, il a également été montré que l'ARN de TBP2 est présent dans l'ovocyte humain (Di Pietro et al., 2008). Pour affiner le profil d'expression temporel et surtout confirmer la présence de cette protéine dans les ovaires, l'étude de biopsies d'ovaires humains serait nécessaire.

Chez l'homme, aujourd'hui, environ 20% des couples en âge de reproduction sont stériles. Les causes de la stérilité restent encore souvent mal connues. Il est donc important de déterminer si TBP2 est lié à des cas de stérilité chez la femme. Dans le but de délimiter le phénotype des femmes pouvant potentiellement être porteuses de mutations sur le gène *Tbp2*, nous avons commencé à déterminer le phénotype histologique des souris TBP2^{-/-}. Nous avons montré que la perte de TBP2 est associée à l'apparition de kystes et de corps hémorragiques, mais n'avons trouvé aucun cas où l'absence de TBP2 induisait à l'apparition de tumeurs. Cependant, les coupes histologiques n'ont été réalisées que sur deux femelles et les deux paires d'ovaires analysées présentent des structures très différentes. Une analyse plus systématique des ovaires de souris TBP2^{-/-} âgées doit donc être envisagée pour définir plus précisément les conséquences à long terme de la perte de cette protéine. Ensuite, des biopsies d'ovaires de femmes fertiles présentant un phénotype comparable à celui des souris TBP2^{-/-} pourront être analysées en collaboration avec le Pr Stéphane Viville au CHU de Strasbourg.

Ces analyses permettront de déterminer si TBP2 peut être lié à un phénotype de stérilité chez la femme, et donc si cette protéine joue un rôle similaire chez l'homme et chez la souris. Les souris TBP2^{-/-} pourront alors servir de modèles dans la recherche de potentiels traitements contre la stérilité.

E. Matériel et méthodes

Je ne présenterai dans cette section que les protocoles utilisés pour la détermination des partenaires protéiques de TBP2. Les autres protocoles utilisés sont détaillés dans la publication 1.

I. Préparation d'extraits nucléaires d'ovaires

Entre 50 et 200 ovaires sont disséqués à partir de souris CD1 âgées de 10 à 15 jours. Les extraits nucléaires sont réalisés selon la méthode de Carl Wu adaptée à des petites quantités de tissus (Wu, 1989). Les cellules sont dissociées et lysées à 4°C avec un *Dounce B* dans 5mL de tampon A (60 mM KCl, 15 mM NaCl, 5 mM MgCl₂, 0,1 mM EGTA, 15 mM Tris pH 7,4, 0,5 mM DTT, 0,1 mM PMSF, 300 mM saccharose). Le tout est alors filtré sur de la gaze puis centrifugé à 4500 g, 4°C pendant 5 minutes. Les noyaux présents dans le culot sont alors resuspendus dans 2 mL de tampon A et déposés sur un coussin de 5 mL de tampon B (60 mM KCl, 15 mM NaCl, 5 mM MgCl₂, 0,1 mM EGTA, 15 mM Tris pH 7,4, 0,5 mM DTT, 0,1 mM PMSF, 1,7 M saccharose). L'interface entre les deux tampons est mélangée pour former un gradient grossier de saccharose. Les noyaux sont centrifugés à 20000g, 4°C pendant 1 heure puis resuspendus dans 100 µL de tampon IP100 (25 mM Tris pH7,9, 5 mM MgCl₂, 0,1% NP40, 0,3 mM DTT, 100 mM KCl, inhibiteurs de protéases). L'ADN des noyaux est digéré pendant 5 minutes à température ambiante en présence de 3 U de micrococcale (SIGMA) et 1 mM CaCl₂. La réaction est arrêtée avec 10 mM final d'EDTA et les noyaux sont lysés dans un sonicateur Bioruptor® pendant 5 minutes à puissance maximale (cycles 30 s ON/30 s OFF). Pour éliminer les débris, l'extrait nucléaire est centrifugé 15 minutes à 11000 g à 4°C.

II. Transfection et préparation d'extraits cellulaires totaux de cellules

HEK

Les cellules HEK sont transfectées au jetPEI avec un plasmide pSG5 contenant l'ADNc de TBP2 fusionné à une étiquette Flag sous la dépendance d'un promoteur SV40, selon le protocole du fabricant (Polyplus transfection). 48h après transfection, les cellules sont

lavées 2 fois avec du PBS (*phosphate buffered saline*), puis resuspendues dans 300 μL de tampon d'extraction (20 mM Tris pH 7,5, 2 mM DTT, 20% glycérol, 400 mM KCl, inhibiteurs de protéases) et soumises à 4 cycles de congélation/décongélation. Après une centrifugation de 30 minutes à 4°C, le surnageant contenant l'extrait cellulaire total est récupéré.

III. Liaison covalente entre les anticorps et les billes de sépharose couplées à la protéine G

Des billes de sépharose couplées à la protéine G (Protein G Sepharose, Dominique Dutscher SA) sont lavées 5 fois avec 5 à 10 volumes de PBS. Un volume de billes est ensuite resuspendu dans 1 volume de PBS et mis en présence de 0,55 à 1,25 mg/mL d'anticorps final (anti-TBP2 2B12, ou anti-GST) pendant 30 minutes à température ambiante. Les complexes billes-anticorps sont lavés 5 fois avec 5 volumes de PBS et resuspendus dans 2 fois leur volume de PBS. La formation d'une liaison covalente entre les billes et les anticorps est induite avec 1,25 mM à 2 mM final de DSS (disuccinimidyl suberate, Fisher Scientific) ajouté goutte à goutte (la poudre de DSS est resuspendue extemporanément à 20 mM dans du DMSO) pendant 1 heure à température ambiante. Les billes sont lavées 3 fois dans du PBS, bloquées avec soit 5 mg/mL d'insuline ou 10 mg/mL de BSA (*Bovine Serum Albumine*, Biolabs) ou 1% de gélatine de poisson (*Gelatin from cold water fish skin*, Sigma, G-7765). Après 3 lavages dans le tampon IP500 (25 mM Tris pH 7,9, 5 mM MgCl_2 , 0,1% NP40, 0,3 mM DTT, 500 mM KCl, inhibiteurs de protéases) puis 2 lavages dans le tampon IP100, les billes sont resuspendues dans une fois leur volume de tampon IP100.

IV. Immunoprécipitation

Pour chaque expérience, la moitié des extraits nucléaires est utilisée pour une IP avec un anticorps anti-TBP2 et l'autre moitié est utilisée comme contrôle négatif pour une IP avec un anticorps anti-GST (15TF2 1D10, IGBMC). Les anticorps anti-TBP2 utilisés sont l'anticorps monoclonal 2B12 et les anticorps polyclonaux 2482, 3023 et 3018 (IGBMC). Les extraits protéiques sont mis en présence soit de 40 μL de protéine G-sépharose (pour les anticorps monoclonaux, ou protéine A-sépharose pour les anticorps polyclonaux) et de 2 μL à

4 μ L d'ascite ou de 2 μ L à 8 μ L de sérum de l'anticorps adéquat, soit de 40 μ L de protéine G-sépharose liée de manière covalente à l'anticorps, sur la nuit à 4°C. La réaction est réalisée dans un volume final de 250 μ L pour les IP sur les extraits nucléaires d'ovaires et dans 1mL pour les extraits de cellules HEK. Les dilutions sont effectuées dans le tampon IP100. Les complexes sont ensuite lavés 2 fois dans du tampon IP500 puis 3 fois dans le tampon IP100. Dans le cas où l'élution est possible, elle est réalisée en présence de 2 mg/mL de peptide correspondant à l'anticorps dans le tampon IP100 pendant 1h à 4°C, le surnageant est ensuite utilisé pour un western blot ou conservé pour analyse en MuDPIT. Dans le cas où l'élution est impossible, les billes sont mises à bouillir pendant 10 minutes dans du bleu de dénaturation de protéines (200 mM Tris pH 6,8, 8% SDS, 40% glycérol, 400 mM DTT, 0,1% bleu de bromophénol) et le surnageant est chargé sur un gel de polyacrylamide puis soumis à western blot ou conservé pour analyse au spectromètre de masse.

Partie 3

MISE EN PLACE ET

EVOLUTION DE

L'HETEROCHROMATINE

PERICENTRIQUE DANS

L'EMBRYON PRECOCE

DE SOURIS

A. Problématique

Plusieurs asymétries caractérisent les pronoyaux paternel et maternel dans le zygote de souris. Alors que le pronoyau mâle contient exclusivement le variant de remplacement H3.3 juste après la fécondation, le pronoyau femelle contient, au moins en partie, H3.1 et H3.2 (Torres-Padilla et al., 2006). De plus, alors que les modifications post-traductionnelles de H3K4 et H3K27 sont héritées du noyau de l'ovocyte au pronoyau femelle, le pronoyau male, condensé autour de protamines dans le spermatozoïde, acquiert les différents degrés de méthylation sur ces deux résidus entre autres, au cours des 24 premières heures de développement (Santos et al., 2005). L'abondance des différentes modifications post-traductionnelles sur les différents variants de H3 a été analysée par spectrométrie de masse dans des cellules somatiques (Hake et al., 2006), mais aucune donnée n'est disponible dans l'embryon. Nous nous sommes donc demandé si les modifications post-traductionnelles des lysines 4 et 27 de H3 peuvent jouer des rôles différents selon si elles sont portées par H3.1 ou H3.3. Ces deux variants jouent-ils un rôle similaire au cours du développement ?

La chromatine paternelle doit acquérir, en très peu de temps, quasiment toutes les modifications post-traductionnelles des histones qui sont nécessaires à l'initiation de la transcription d'un groupe de gènes adapté au développement du zygote ou à l'établissement de domaines chromatinien spécifiques. Les chromosomes paternels doivent, par exemple, former les structures d'hétérochromatine constitutive, dont l'hétérochromatine péricentromérique, permettant le maintien de leur intégrité au cours des divisions cellulaires. De plus, la chromatine péricentromérique subit de nombreuses relocalisations dans les noyaux des embryons avant d'acquérir une structure somatique en chromocentres. Plusieurs questions se posent alors. Comment s'effectue l'hétérochromatinisation *de novo* des régions péricentriques paternelles ? Existe-t-il un lien entre la localisation des régions péricentromériques au sein du noyau et le degré de condensation ? La condensation est-elle la cause ou la conséquence de la relocalisation ?

Pour répondre à ces questions, nous avons voulu utiliser des embryons exprimant des protéines exogènes à partir du stade zygote. Aucune transcription n'ayant lieu dans l'embryon au début du développement, l'induction d'une expression transitoire de protéines est possible grâce à la microinjection d'ARNm dans le zygote au moment de la fécondation. La protéine est alors traduite et détectable dans l'embryon en environ 30 minutes.

B. Publication 2

I. Résumé de la publication 2

Note : les figures appelées dans le texte sont celles de la publication 2.

La question première posée avant la réalisation de cette publication était de savoir si les modifications post-traductionnelles des lysines 4 et 27 de H3 jouent des rôles différents au cours du développement si elles sont portées par le variant canonique H3.1 ou par le variant de remplacement H3.3. Pour répondre à cette question, nous avons réalisé des injections d'ARNm codant pour les protéines H3.1 et H3.3 où la lysine 4 (H3.1K4R et H3.3K4R) ou la lysine 27 (H3.1K27R ou H3.3K27R) a été remplacée par une arginine, dans des embryons sauvages au moment de la fécondation (figure 1a).

1. Les embryons exprimant H3.3K27R présentent des défauts lors du développement précoce

Les embryons surexprimant les formes sauvages de H3.1 et H3.3 ainsi que ceux exprimant les protéines dont la lysine 4 est mutée atteignent le stade blastocyste dans les mêmes proportions que les embryons sauvages. En revanche, la mutation de la lysine 27, spécifiquement sur H3.3 entraîne un arrêt du développement des embryons entre les stades 2-cellules et 8-cellules (figure 1 et figure 2).

Ces résultats montrent que des modifications de la lysine 27, spécifiquement sur le variant H3.3, sont nécessaires au développement préimplantatoire des embryons.

2. L'expression de H3.3K27R entraîne des changements globaux dans le niveau de modifications des histones

Des analyses en immunofluorescence des embryons exprimant H3.3K27R ont révélé une diminution de plus de 50% des taux globaux de H3K27me1 et H3K27me3 au stade 2-cellules par rapport à des embryons sauvages. L'expression de H3.3K27R induit donc des changements importants de méthylation de H3K27 au stade 2-cellules, changements potentiellement responsables de l'arrêt du développement (figure 4).

3. H3.3 et H3.1 sont incorporés à différents moments et à différents loci dans les pronoyaux paternel et maternel

Pour tenter de comprendre pourquoi seule la mutation de H3.3 entraîne un arrêt du développement, nous avons analysé les différences d'incorporation temporelles et spatiales de H3.3-GFP et H3.1-GFP dans le zygote. Au moment de la formation du pronoyau paternel, seule H3.3 est déposée à la chromatine. H3.1 apparaît dans le pronoyau mâle seulement environ 1 heure plus tard. Dans le pronoyau femelle, H3.3 et H3.1 sont en revanche déposées simultanément environ 1 heure après la formation du pronoyau (figure 3 et tableau 1). Au tout début du développement, H3.3 est donc le seul variant présent à la chromatine paternelle (en considérant que H3.2 sera déposée en même temps que H3.1). Ces résultats, déjà publiés précédemment, nous permettent cependant de confirmer que nos protéines exogènes sont incorporées selon la même chronologie que les histones endogènes.

De plus, après la réplication, alors que les deux variants H3.3 et H3.1 sont présents dans les deux pronoyaux, H3.3-GFP est présente au niveau des anneaux de chromatine condensée autour des NLB dans le pronoyau paternel alors qu'elle n'est pas enrichie au niveau de ces structures dans le pronoyau maternel (figure 5a, b et c). Cette observation permet d'émettre l'hypothèse que H3.3 joue un rôle spécifique à la chromatine péricentromérique paternelle, rôle qu'elle ne jouerait pas à la chromatine maternelle.

4. HP1 β se lie à la chromatine péricentrique suite à des modifications de H3.3K27 et à sa liaison avec un ARN

L'étude phénotypique des embryons au stade 2-cellules exprimant H3.3K27R nous a permis de mettre en évidence des problèmes d'incorporation de l'ADN au sein des noyaux (figure 6a). Ce phénotype pouvant être la conséquence de problèmes de ségrégation de la chromatine au moment de la première mitose, et étant donné la localisation de H3.3 au niveau des centromères paternels, nous avons décidé d'analyser ces régions plus en détails dans les embryons exprimant H3.3K27R.

L'expression de H3.3K27R induit une relocalisation des centromères au stade 2-cellules. En effet, un nombre plus important de centromères ne se localise pas aux NLB au stade 2-cellules dans les embryons exprimant H3.3K27R que dans les embryons sauvages (figure 5d et e).

La protéine HP1 β , normalement enrichie aux régions péricentromériques, présente une localisation diffuse dans le nucléoplasme des embryons exprimant H3.3K27R. De plus, une forme mutante de HP1 β , ne possédant pas le domaine *Hinge* et étant donc incapable de se lier aux acides nucléiques, présente le même profil de localisation dans des embryons sauvages au stade 2-cellules (figure 6b et c). Le traitement des embryons avec de la RNase A induisant un profil de relocalisation de HP1 β similaire au profil observé lors de la mutation du domaine *Hinge*, nous en avons déduit que la localisation de HP1 β aux régions péricentromériques devait se faire via sa liaison à un ARN. La modification de H3.3K27, ainsi que la liaison de HP1 β à des acides ribonucléiques, sont donc tous deux nécessaires à une localisation correcte de HP1 β aux régions péricentromériques dans l'embryon. De plus, *in vitro*, HP1 β peut lier l'ARN transcrit à partir des *major satellites*, séquences répétées en tandem et qui caractérisent la chromatine péricentrique.

5. L'expression de H3.3K27R induit des défauts de condensation de la chromatine péricentrique paternelle

Dans les embryons exprimant H3.3K27R, les *major satellites* sont surexprimés au stade 2-cellules suggérant que la chromatine péricentromérique n'est pas suffisamment condensée pour que la transcription soit inhibée (figure 6d).

Alors que la chromatine péricentromérique maternelle est déjà sous forme d'hétérochromatine au moment de la fécondation, la chromatine paternelle est la seule à devoir subir une hétérochromatinisation rapide avant la première mitose. Nous avons montré que les ARN des *major satellites* présents au stade zygote ont une origine paternelle. De plus, dans nos embryons mutants, les défauts de ségrégation des chromosomes au moment de la première mitose concernent des chromosomes paternels (figure 6e).

Ces résultats montrent que la modification de H3.3K27 est nécessaire à la formation de l'hétérochromatine péricentrique paternelle.

6. L'injection d'ARN double brin des *major satellites* permet une localisation correcte de HP1 β et un développement normal des embryons exprimant H3.3K27R

Pour confirmer le lien potentiel entre la modification de H3.3K27 et l'hétérochromatinisation de la chromatine péricentrique grâce à la transcription des *major satellites*, nous avons injecté à la fois l'ARNm codant pour H3.3K27R et l'ARNdb des *major satellites* dans des zygotes au moment de la fécondation (figure 7a). Les zygotes doublement injectés se développent jusqu'au stade blastocyste dans les mêmes proportions que des zygotes sauvages et la localisation de HP1 β est normale dans ces embryons au stade 2-cellules (figure 7b et c). Ces résultats confirment qu'il existe un lien fonctionnel entre la modification de H3.3K27 et les ARNdb des *major satellites* dans l'embryon précoce de souris. Ils confirment également que le phénotype développemental observé chez les embryons exprimant H3.3K27R est bien la conséquence de la mauvaise localisation de HP1 β , donc certainement de la mauvaise hétérochromatinisation des régions péricentromériques paternelles.

7. Conclusion

En conclusion, nous avons tout d'abord montré que l'importance des modifications post-traductionnelles de la lysine 27 de H3 au cours du développement dépend du variant de H3 par lequel elles sont portées. En effet, seule la mutation de H3.3K27 entraîne un arrêt du développement préimplantatoire des embryons à cause de défauts de formation de l'hétérochromatine péricentrique paternelle. Nos résultats mettent en évidence trois pistes qui doivent être prises en compte pour expliquer l'hétérochromatinisation *de novo* qui a lieu aux régions péricentriques paternelles. (1) Les modifications de H3.3K27 sont nécessaires à la condensation et à la localisation des régions péricentromériques paternelle dans le noyau des embryons. (2) Le recrutement de HP1 β aux régions péricentromériques s'effectue grâce aux modifications de H3.3K27 et via sa liaison avec un ARN pouvant potentiellement être transcrit à partir des *major satellites*. (3) L'ARN des *major satellites* joue un rôle sous forme d'ARNdb sur la localisation de HP1 β et la condensation de l'hétérochromatine péricentrique.

Les liens entre ces différentes voies ainsi que les mécanismes précis de l'hétérochromatinisation *de novo* chez les mammifères restent encore à déterminer.

II. Publication 2

Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3

Angèle Santenard¹, Céline Ziegler-Birling¹, Marc Koch¹, László Tora¹, Andrew J. Bannister² and Maria-Elena Torres-Padilla^{1,3}

In mammals, oocyte fertilization by sperm initiates development. This is followed by epigenetic reprogramming of both parental genomes, which involves the *de novo* establishment of chromatin domains. In the mouse embryo, methylation of histone H3 establishes an epigenetic asymmetry and is predominant in the maternal pronucleus. However, the roles of differential incorporation of histone H3 variants in the parental chromatin, and of modified residues within specific histone variants, have not been addressed. Here we show that the histone variant H3.3, and in particular lysine 27, is required for the establishment of heterochromatin in the mouse embryo. H3.3 localizes to paternal pericentromeric chromatin during S phase at the time of transcription of pericentromeric repeats. Mutation of H3.3 K27, but not of H3.1 K27, results in aberrant accumulation of pericentromeric transcripts, HP1 mislocalization, dysfunctional chromosome segregation and developmental arrest. This phenotype is rescued by injection of double-stranded RNA (dsRNA) derived from pericentromeric transcripts, indicating a functional link between H3.3K27 and the silencing of such regions by means of an RNA-interference (RNAi) pathway. Our work demonstrates a role for a modifiable residue within a histone-variant-specific context during reprogramming and identifies a novel function for mammalian H3.3 in the initial formation of dsRNA-dependent heterochromatin.

Fertilization of the oocyte by the sperm constitutes the first event of embryogenesis and results in the formation of the zygote. The creation of such a totipotent cell from two differentiated ones involves epigenetic reprogramming of the parental genomes. Throughout the complete first cell cycle, the male and female pronuclei behave as two distinct sets of chromatin that coexist as separate nuclear entities. Both pronuclei evolve differently, showing different chromatin signatures, histone marks, and replication and transcription timings^{1–7}. How the chromatin is assembled, specified and reprogrammed after fertilization remains a central question in biology. Before fertilization, the sperm nucleus is condensed sixfold more than a somatic cell nucleus, and most of its histones are replaced by protamines⁸. Immediately after the entry of the sperm nucleus into the oocyte cytoplasm, protamines are removed from the sperm and replaced by maternally provided histones⁹. In mice, histones H3 and H4 are translated from maternal messenger RNAs stored in the oocyte, whereas histones H2A and H2B are already present as proteins in the oocyte¹⁰.

Apart from the canonical histones, which are synthesized exclusively during S phase, histone variants can be incorporated into chromatin throughout the cell cycle. The replication-independent H3 variant H3.3 is preferentially incorporated into the male pronucleus after fertilization⁵, whereas the replication-dependent H3.1/2 variants are found predominantly in the female

pronucleus¹¹, suggesting a role of differential incorporation of histone variants in the formation of embryonic chromatin. However, the precise time at which chromatin is formed, or how newly incorporated histone variants and the residues within them contribute to development is unknown. H3.3 has been associated with active transcription in somatic cells^{12,13}, and H3 variants differ in their relative abundances of various modifications^{14,15}. For example monomethylation on Lys 27 (K27me1) and trimethylation on Lys 4 (K4me3) are more abundant in H3.3 than in H3.1. Changes in the methylation of histone H3 occur after fertilization, and there is a differential distribution of histone modifications between male and female pronuclei^{2–5}. Of these modifications, methylation of H3K4, H3K9 and H3K27 are less abundant in the male pronucleus, and the paternal chromatin only gradually acquires chromatin signatures during the first cell cycle (Supplementary Information, Figs S1–S4).

RESULTS

Embryos expressing H3.3K27R are compromised in their development

To address the contribution of specific residues within the histone H3 variants H3.1 and H3.3 to the establishment and subsequent

¹Institut de Génétique et de Biologie Moléculaire et Cellulaire, Centre national de la recherche scientifique/Institut National de la Santé et de la Recherche Médicale U964, Université de Strasbourg, F-67404 Illkirch, France. ²The Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, UK.

³Correspondence should be addressed to M.-E.T.-P. (e-mail: metp@igbmc.fr)

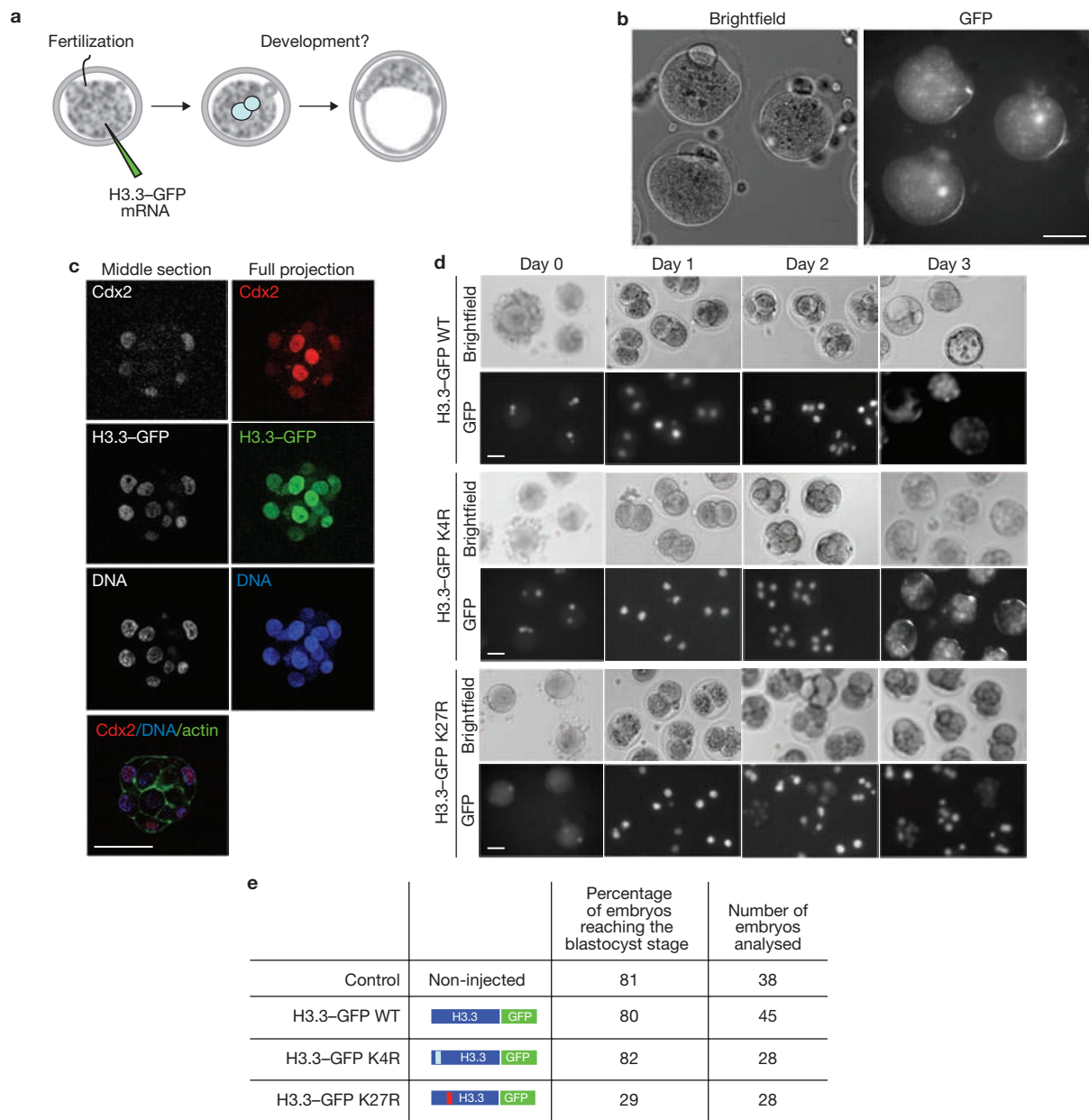


Figure 1 Embryos expressing H3.3-GFP K27R are compromised in their development. **(a)** Diagram showing the experimental design. Zygotes were collected and microinjected at the fertilization cone stage (17–18 h after injection of human chorionic gonadotropin) with *in vitro*-transcribed mRNA and cultured until the blastocyst stage. **(b)** Efficient incorporation of H3.3-GFP in the forming male pronucleus after fertilization on microinjection of mRNA. Brightfield and GFP images of representative zygotes, showing GFP fluorescence in the male pronucleus. Scale bar, 50 μ m. **(c)** Overexpression of H3.3-GFP does not alter development or blastocyst formation. Zygotes were injected with mRNA for H3.3-GFP WT as in **a**, cultured until the blastocyst stage, fixed and processed for immunostaining with a Cdx2 antibody ($n = 8$). Note that H3.3 expression persists until the blastocyst stage in all cells. Shown are a middle section (left) and a full maximal-intensity projection (right) of stack z-images

acquired every 2 μ m. Cortical actin labelled with phalloidin demarcates cell boundaries. DNA (blue) was stained with DAPI. Scale bar, 50 μ m. **(d)** Mutation of K27 of H3.3 compromises development to the blastocyst stage. Zygotes were microinjected as in **a** with the indicated mRNAs, cultured, monitored daily and scored for their developmental stage. Among the H3.3 K27R-expressing embryos showing developmental arrest, 39% arrested at the 2-cell stage, 41% between the 3-cell and 4-cell stages and 20% at the 8-cell stage. Shown are representative brightfield and fluorescence images of embryos imaged daily. Scale bar, 50 μ m. **(e)** Summary of development of control, non-injected embryos or embryos expressing H3.3-GFP WT, H3.3-GFP K4R or H3.3-GFP K27R from six independent experiments. Statistical analysis was performed by analysis of variance; $P = 0.000001$. Note that embryos injected with only GFP mRNA developed to the blastocyst in 87% of cases ($n = 18$, not shown).

reprogramming of chromatin after fertilization, we expressed these variants harbouring point mutations in zygotes and assessed the development of these embryos. Because high expression of exogenous histones can cause intra-S-phase checkpoint activation, potentially

eliciting non-specific developmental defects¹⁶, we first established conditions under which the expression of H3 did not alter normal development by titrating mRNA concentrations of wild-type (WT) H3.3 (Supplementary Information, Fig. S5). We injected mRNA for

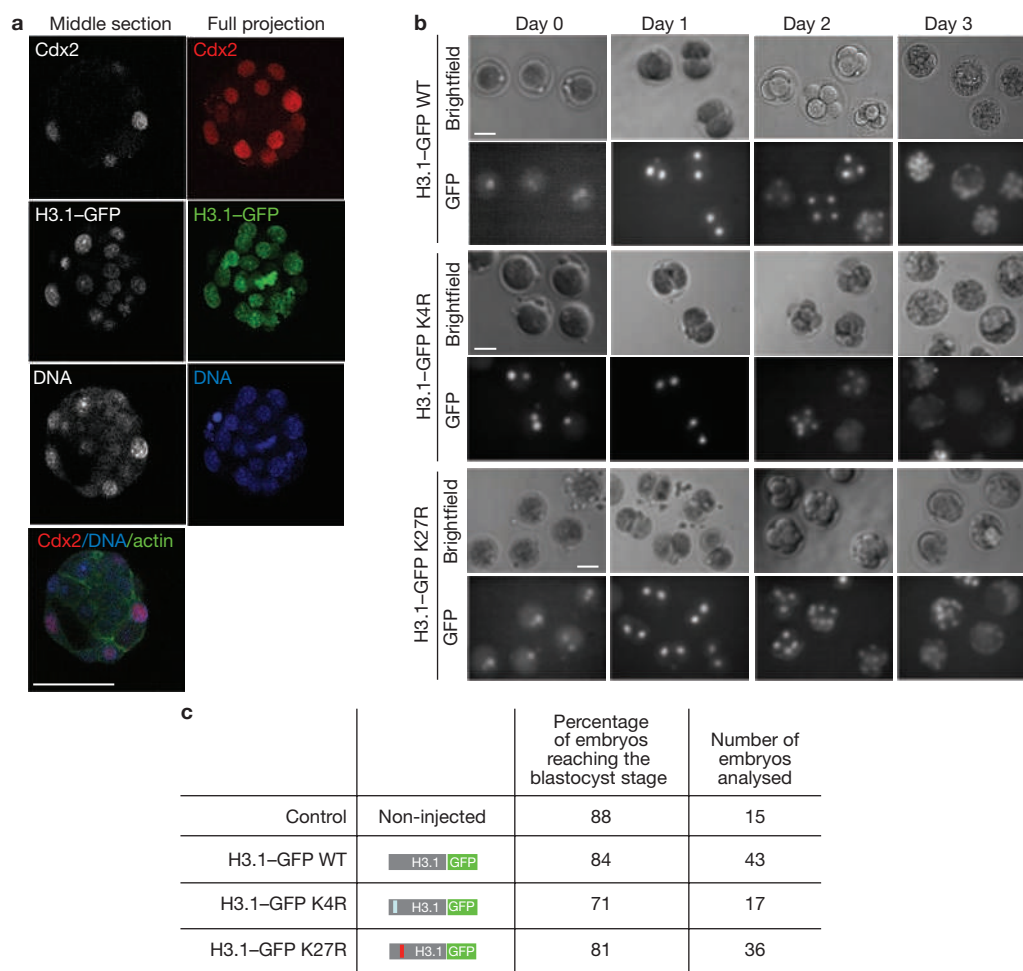


Figure 2 Developmental defects elicited on mutation of K27 are H3.3-specific. (a) Overexpression of H3.1-GFP does not alter development or blastocyst formation. Zygotes were injected with mRNA for H3.1-GFP WT as in Fig. 1, cultured until they reached the blastocyst stage, fixed, and analysed with the Cdx2 antibody. Confocal sections were taken every 2 μm . Shown are a middle section (left) and a full projection (right) of one representative

blastocyst of 12 analysed. Scale bar, 50 μm . (b) Mutation of K27 in H3.1 does not alter developmental progression. Zygotes were microinjected as in Fig. 1a and their developmental stage was monitored daily. Scale bar, 50 μm . (c) Summary of development of embryos expressing H3.1-GFP WT, H3.1-GFP K4R or H3.1-GFP K27R from five independent experiments. Statistical analysis was performed by analysis of variance; $P = 0.063$.

green fluorescent protein (GFP)-tagged H3.3 WT into zygotes at the fertilization cone stage before formation of the pronucleus (Fig. 1a). This resulted in efficient accumulation of H3.3-GFP in the forming male pronucleus (Fig. 1b), in line with our previous observations⁵. H3.3-GFP was still present in the embryo even after several cell divisions (Fig. 1c, d). H3.3-GFP WT-expressing embryos developed to the blastocyst stage in normal ratios compared with non-injected embryos (Fig. 1d, e) and showed normal expression patterns of the trophectoderm and inner-cell-mass markers *Cdx2* and *Nanog*, respectively, as determined by immunostaining (Fig. 1c, and not shown). We then used these parameters and expressed H3.3-GFP K4R or K27R mutants and followed the development of these embryos (Fig. 1d). Whereas embryos injected with H3.3 WT and H3.3 K4R reached the blastocyst stage at the same time and in similar ratios to those of the non-injected controls, embryos expressing H3.3 K27R exhibited a reduced rate of development and only 29% of the embryos reached the blastocyst stage ($P = 0.000001$) (Fig. 1d, e). We ruled out the possibility that H3.3-GFP K27R had effects on developmental progression as a result of higher levels of expression than H3.3 WT or H3.3 K4R by measuring

the GFP fluorescence intensity in embryos expressing each of the three proteins. This revealed no difference in GFP levels among the three groups ($P = 0.75$). Thus, of the residues analysed on H3.3, only K27 seems to be important for early embryonic development.

Developmental defects elicited on mutation of K27 are H3.3 specific

Modifications of K27 of H3, mainly methylation, have been involved in developmental processes including the establishment of a chromatin environment associated with pluripotency¹⁷. However, it is not known whether the role of K27 modification is pertinent to a particular H3 variant, because antibodies recognizing K27 marks do not distinguish between different variants. To determine whether the defects of H3.3-GFP K27R that we observed were H3 variant-specific, we repeated the same experiments as above, but with H3.1-GFP. Embryos expressing H3.1-GFP WT developed to the blastocyst stage in normal ratios compared with non-injected control embryos and showed correct expression of *Cdx2* and *Nanog* (Fig. 2a–c, and not shown). When we expressed H3.1 K4R or K27R mutants in the zygote, both groups of embryos

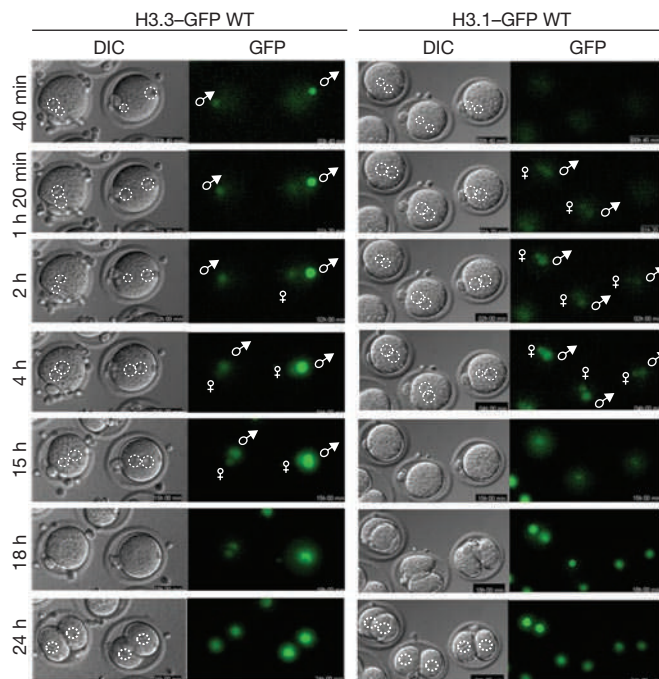


Figure 3 H3.3 and H3.1 are incorporated at different times in the two pronuclei. H3.3 is first incorporated in the male pronucleus, whereas H3.1 is only incorporated later and in both pronuclei simultaneously. Zygotes were injected as in Fig. 1 with mRNAs for H3.3-GFP WT or H3.1-GFP WT and were subsequently grown under time-lapse microscopy along z-series covering both pronuclei. Shown are DIC and GFP channels of images captured at the indicated times; the times are arbitrarily taken from the beginning of the time-lapse experiment and are not related to the time of formation of pronuclei. Where relevant, male and female pronuclei are indicated. The white dashed line demarcates the position of the nuclei. For both groups of embryos nuclear envelope breakdown occurred $13\text{ h }34\text{ min} \pm 4\text{ h }6\text{ min}$ after pronucleus formation, and cytokinesis followed $1\text{ h }46 \pm 11\text{ min}$ later (errors represent s.d.). This was similar to non-injected embryos, confirming that our conditions for expression of H3 variants do not significantly affect S-phase progression or mitosis.

developed normally and showed no differences in their development from H3.1 WT or non-injected controls ($P = 0.63$) (Fig. 2b, c). These results indicate that the developmental defects elicited by mutation of K27 are H3.3 specific.

H3.3 and H3.1 are incorporated at different times in the maternal and paternal pronuclei

Next, we wished to determine the basis for the variant-specific difference that we observed after mutation of K27. We examined whether the timing of incorporation of H3.1 and H3.3 would be relevant for such effects. In somatic cells, H3.1 deposition, contrary to H3.3 deposition, is replication-dependent^{12,18}. It is currently unknown whether the same mechanisms regulate the deposition of H3 variants in the zygote. We performed time-lapse analyses to determine the dynamics of appearance of H3.3 and H3.1 in the forming male and female pronuclei, which were easily distinguishable under differential interference contrast optics (Fig. 3). Zygotes injected as above with H3.3 or H3.1-GFP WT mRNA were imaged every 10 min along z planes covering both pronuclei. We found that H3.3 is found on the paternal chromatin concomitant with formation of the pronucleus (Fig. 3 and Table 1). However, in the female pronucleus H3.3 was not detected until 1 h after formation of

the pronucleus (Fig. 3 and Table 1). In contrast with H3.3, H3.1 was deposited almost simultaneously in both pronuclei (Table 1) at the time when bromodeoxyuridine (BrdU) incorporation was first observed in the zygote (data not shown, and ref. 7), suggesting that H3.1 is deposited similarly to somatic cells during S phase. Thus, H3.3 appears first in the male chromatin concomitantly with pronuclear formation, and its incorporation in the female pronucleus occurs only about 1 h later, whereas H3.1 is deposited in both pronuclei simultaneously at the onset of DNA replication.

We speculated that the time difference between the incorporation of H3.3 and H3.1 could be relevant for the variant-specific effects of K27 mutation. Because this time corresponds to the onset of the first S phase, we considered two possibilities to explain the developmental defects after expression of H3.3 K27R: that the effect of H3.3 K27R either is related to a function of H3.3 before or concomitant with S phase and/or is related to its preferential incorporation to a specific genomic location.

Expression of H3.3K27R leads to global changes in specific epigenetic marks

First, we analysed whether embryos overexpressing H3.3 K27R show defects in epigenetic marks. Quantification of global levels of H3K27me3 in embryos expressing H3.3-GFP K27R revealed methylation levels that were decreased to about 65% of those in control embryos (Fig. 4a; Supplementary Information, Fig. S5). Expression of H3.3-GFP K27R further resulted in an altered distribution and global decrease in H3K27me1 levels to about 54% of those of the controls (Fig. 4b; Supplementary Information, Fig. S5). In contrast, we did not detect any obvious change in the distribution of euchromatic marks such as H3K4me3, nor indeed in H3K9me3, in embryos expressing either H3.3 WT or the K27R mutant in comparison with control embryos (Fig. 4c, d).

H3.3 localises to paternal pericentromeric chromatin

Second, we analysed H3.3 localization in zygotes during S phase. H3.3 is present in both pronuclei, with fluorescence levels slightly higher in the male pronucleus (Fig. 5a), similarly to endogenous H3.3 (ref. 5). We found that H3.3 localizes to the heterochromatic 4',6-diamidino-2-phenylindole (DAPI)-intense rings surrounding the nucleolar-like bodies (NLBs) in the male pronucleus but not in the female one (Fig. 5a, b). The same pattern was observed for the H3.3 K4R and K27R mutants (not shown). H3.1 was distributed throughout both pronuclei and followed the DAPI distribution, with no discernibly specific pattern (Fig. 5c). These dense ring-like structures around NLBs in the zygote are formed by the chromocentres that contain pericentromeric and centromeric chromatin. In the male pronucleus, these regions gradually acquire methylation of H3K27 (but not of H3K9 or H3K4) after the mid-pronuclear stages (PN3–PN4), concomitant with the progression of S phase (Supplementary Information, Figs S2–S4; see also ref. 4). We next analysed the localization of the chromocentres in our mutants. Expression of H3.3-GFP K27R induces a spatial relocalization of chromocentres at the 2-cell stage (Fig. 5d, e). This effect is specific for H3.3, because embryos expressing H3.1-GFP K27R had a normal pattern of chromocentre localization (not shown). Because the association of chromocentres to NLBs has been suggested to reflect their heterochromatic nature¹⁹, their mislocalization in H3.3 K27R-expressing embryos suggests defects in heterochromatin.

Table 1 Time (in min) of GFP appearance in the male or the female pronucleus for either H3.1 or H3.3 relative to the time of pronuclear formation

	GFP in male PN (relative to PN formation)	GFP in female PN (relative to PN formation)	Time difference in appearance of GFP between 2 PN (Female PN–Male PN)	Number of zygotes analysed
H3.3–GFP	–7 ± 12	79 ± 37	86 ± 33	10
H3.1–GFP	67 ± 42	82 ± 44	15 ± 24	10
	$P = 0.0004$	$P = 0.87$	$P < 0.0001$	

PN, pronucleus. Error indicates s.d.

Expression of H3.3K27R leads to mislocalisation of HP1 β and de-repression of pericentromeric chromatin

A major architectural component of chromocentres is the HP1 protein, and its appropriate localization is essential for chromocentre function. Furthermore, in *Schizosaccharomyces pombe* the HP1 homologue Swi6 is required for normal chromosome segregation and its deletion leads to chromosome lagging, a feature that we observed in 2-cell embryos expressing H3.3 K27R (Fig. 6a). Considering all these facts together, we reasoned that an analysis of HP1 localization might shed light on the mechanism(s) underlying the embryonic defects caused by the expression of H3.3 K27R. We found that expression of H3.3–GFP K27R led to mislocalization of HP1 β , which was no longer enriched in the DAPI-rich regions around the NLBs in comparison with control embryos and was instead spread out in the nucleoplasm (Fig. 6b; Supplementary Information, Fig. S5). This suggests that recruitment of HP1 β to pericentric chromatin is impaired in H3.3 K27R-expressing embryos. To explore further the involvement of HP1 β in the observed effects we expressed three HP1 β constructs: first, a tagged full-length HP1 β WT, second, a similar construct containing a mutation within its chromodomain that abolishes H3K9me binding (HP1 β CD), and third, a similar construct lacking the HP1 β Hinge region (HP1 β Δ Hinge). Tagged full-length HP1 β WT showed the same localization as that of endogenous HP1 β in both H3.3–GFP WT-expressing embryos and H3.3–GFP K27R-expressing embryos (Fig. 6c). Mutation of the chromodomain did not alter HP1 β enrichment around the NLBs in the H3.3 WT background (Fig. 6c). This is consistent with the fact that the male pronucleus is devoid of detectable H3K9me3 (ref. 4). In contrast, deletion of the Hinge region resulted in a similar HP1 β localization to that observed with the H3.3 K27R mutation (Fig. 6c). Furthermore, a significant proportion of embryos expressing HP1 β Δ Hinge showed chromosome lagging (25%, $n = 12$), which did not occur in embryos expressing HP1 β WT or HP1 β CD (0 out of 25 and 0 out of 16, respectively). Thus, it seems that the HP1 β Hinge region is necessary for the pericentric enrichment of HP1 β in 2-cell embryos, but a functional chromodomain is dispensable. The HP1 Hinge region is essential for its interaction with RNA²⁰, suggesting that the effects that we observed may be due to an interaction between HP1 β and chromatin-associated RNA molecules. Indeed, evidence for this stems from our observation that treatment of the zygote with RNase A results in mislocalization of HP1 β and that HP1 β binds to pericentromeric RNA transcripts *in vitro* (Supplementary Information, Fig. S6).

So far, our data suggest a defect in chromocentre remodelling in H3.3 K27R-expressing embryos involving RNA and HP1 β . In *S. pombe*, transcription of centromeric chromatin during S phase is essential for preservation of the heterochromatic nature of the centromere by means of a mechanism involving RNA, the HP1 homologue Swi6 and the RITS (RNA-induced transcriptional silencing) complex^{21,22}. Heterochromatin assembly of the surrounding repetitive elements also requires transcription²³. Although it is unclear whether these processes are conserved in full²⁴, major

satellites are also transcribed predominantly during S phase in mammalian cells²⁵. Because it is the chromocentres that contain the minor (centromeric) and major (pericentromeric) satellite repeats^{4,19}, we wondered whether H3.3 K27R expression affects the accumulation of transcripts derived from major satellites. Indeed, expression of H3.3–GFP K27R resulted in a modest but significant increase in the abundance of major satellite transcripts at the 2-cell stage, suggestive of derepression or lack of silencing ($P = 0.004$) (Fig. 6d). Importantly, overexpression of H3.3–GFP WT did not alter the levels of transcripts derived from major satellites compared with those in non-injected embryos ($P = 0.74$) (not shown). Furthermore, because H3.3 associates with active genomic regions in somatic cells, we tested whether H3.3 accumulation around NLBs in the male pronucleus might be indicative of major satellites being transcribed from the paternal genome. We assessed this by performing crosses with mice of different subspecies and found that it is indeed the paternal major satellites that are preferentially transcribed (Supplementary Information, Fig. S7a), in agreement with previous RNA fluorescence *in situ* hybridization data²⁶. Furthermore, extrapolating these observations led us to speculate that the lagging chromatin in 2-cell-stage embryos expressing H3.3 K27R (Fig. 6a) must be of paternal origin, which we confirmed (Fig. 6e). We also observed that H3K9me3 levels — a mark detected only in the female chromatin — seemed unaffected (Fig. 4d, e). Notably, the defects in chromosome segregation occurred during the first mitosis (Fig. 6e). Taken together, these data demonstrate that mutation of H3.3 K27 results in the accumulation of major satellite transcripts after the first mitotic division, suggesting defects in silencing and/or paternal heterochromatin formation.

Pericentromeric dsRNA rescues HP1 β localization and the developmental defects of H3.3K27R embryos

The similarity between the lagging phenotype and mislocalization of HP1 β obtained after deletion of the Hinge region or following H3.3 K27R expression, together with our observations that mutation of H3.3 K27 leads to improper silencing of major satellite repeats, prompted us to explore a link between these events. In *S. pombe*, centromeric silencing is mediated by an RNAi pathway that depends on the generation of double-stranded RNA (dsRNA) from centromeric repeats. Moreover, the HP1 homologue Swi6 binds to centromeric transcripts — as determined by RNA immunoprecipitation — and this association is necessary to enable RNAi-mediated silencing²⁷. Our own observations also show that HP1 β can bind major satellite transcripts *in vitro* (Supplementary Information, Fig. S6). We showed above that H3.3 K27R mutation led to an accumulation of major satellite transcripts. Importantly, this analysis was performed at the early 2-cell stage, a time point after the stage at which the repeats are normally silenced by inclusion into heterochromatin (Fig. 7d). This indicated that mutation of H3.3 K27 led to derepression, or lack of silencing, of these repeats. We hypothesized that in the zygote (that is, before the 2-cell stage) H3.3 K27R leads to inefficient production of

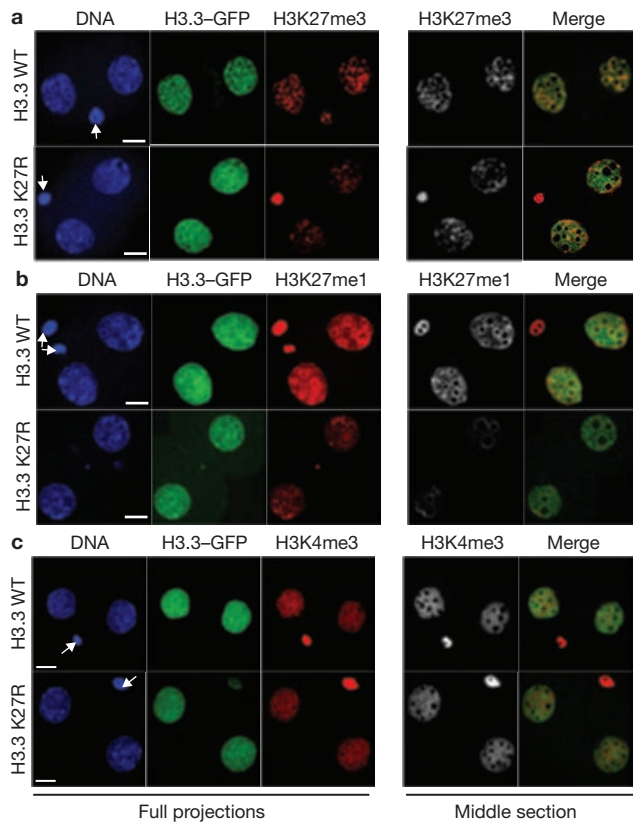
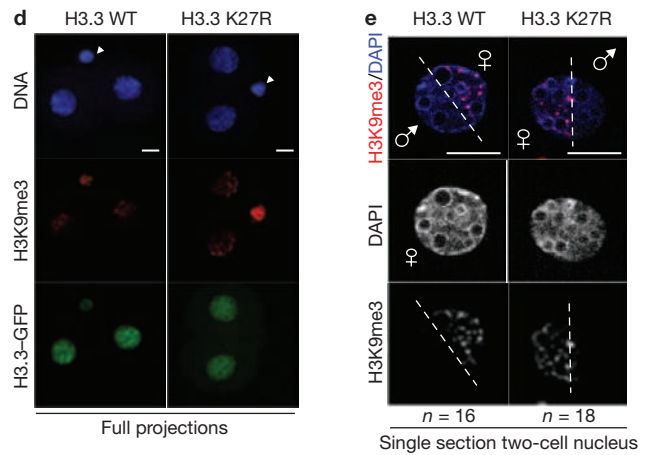


Figure 4 Expression of H3.3-GFP K27R leads to changes in some epigenetic marks. **(a, b)** Changes in H3K27me3 **(a)** and H3K27me1 **(b)** in 2-cell-stage embryos expressing either H3.3-GFP WT or H3.3-GFP K27R. Shown are full projections (left) or middle sections (right) of images taken along the z axis every 0.6 μm . For each antibody, embryos were processed and analysed in parallel with identical confocal acquisition parameters. Data shown are representative of 16 (H3K27me1) and 19 (H3K27me3) embryos analysed. Quantification of fluorescence levels was performed throughout individual sections with ImageJ. DNA was stained with DAPI (blue) and the GFP signal is shown in green. Where visible, polar bodies are arrowed. Non-injected controls are shown in Supplementary Information, Fig. S5. Scale bar, 14 μm . **(c)** H3.3-GFP WT or H3.3-GFP K27R has no apparent effect on the global distribution of H3K4me3 at the 2-cell stage. Embryos were processed as in

major satellite dsRNA, which in turn would lead to an inability to initiate effective heterochromatinization of the repeats. We therefore tested whether artificial reintroduction of dsRNA from major satellites would overcome the developmental phenotype elicited by mutation of H3.3 K27R. To address this we injected embryos with either H3.3 WT or H3.3 K27R mRNAs as before, and left the zygotes to progress into S phase (about 6 h later). Then we microinjected dsRNA for major satellites (MajSat), or LacZ as control, into the paternal pronucleus (Fig. 7a). As above, we monitored development daily and scored the number of embryos that reached the blastocyst stage. Introduction of dsRNA for MajSat or LacZ had no effect on embryos expressing H3.3 WT (Fig. 7b). In contrast, we found that embryos expressing H3.3-GFP K27R developed in normal ratios after the introduction of dsRNA for MajSat, but not dsRNA for LacZ, into the paternal pronucleus (Fig. 7b). We further addressed whether the reintroduction of dsRNA for MajSat had an effect on HP1 β localization in H3.3-GFP K27R-expressing embryos. Although H3.3 K27R-expressing embryos injected with dsRNA for LacZ showed the same diffuse pattern of localization of HP1 β as above, we found that HP1 β relocated to



a and b and analysed with an antibody against H3K4me3. Representative embryos of at least seven embryos analysed per group are shown. Embryos were processed in parallel and acquisition was performed with identical confocal parameters. Similar results were obtained for H3.1 WT and H3.1 K27R injections. Scale bar, 10 μm . **(d, e)** Analysis of H3K9me3 following expression of H3.3 K27R. Zygotes were processed as above and stained with an antibody against H3K9me3. Shown in **d** are maximal projections of z-stack sections taken every 0.5 μm of representative 2-cell-stage embryos ($n = 8$ and $n = 9$, respectively). In **e**, single sections of individual nuclei of 2-cell-stage embryos are shown. The maternal chromatin is clearly distinguishable from the paternal chromatin by H3K9me3 staining (red). The white dashed line indicates the approximate 'limit' between the two parental chromatins. Scale bar, 10 μm .

DAPI-rich regions after the reintroduction of dsRNA for MajSat (Fig. 7c). Single-stranded RNA from MajSat (ssMajSat) did not rescue the localization of HP1 β (Fig. 7c), suggesting that the RNAi machinery — or at least the need for a dsRNA template — is involved in this process. These results also suggest that the phenotype elicited by H3.3 K27R expression is due primarily to the defects in transcriptional silencing of pericentromeric transcripts. Moreover, it suggests that production of dsRNA from the satellites might be limiting in the silencing process in the H3.3 K27R mutants. Most significantly, the addition of dsRNA, specifically that of major satellite transcripts, rescues the developmental phenotype elicited by the expression of H3.3 K27R, demonstrating a functional link between H3.3 and pericentromeric RNA in this developmental context.

DISCUSSION

Heterochromatic repeats need to be transcribed to generate dsRNA and propagate their heterochromatic state. We suggest that in wild-type embryos the low levels of K27me and the presence of H3.3 provide a window of opportunity for such transcription in the male pronucleus (Fig. 7d).

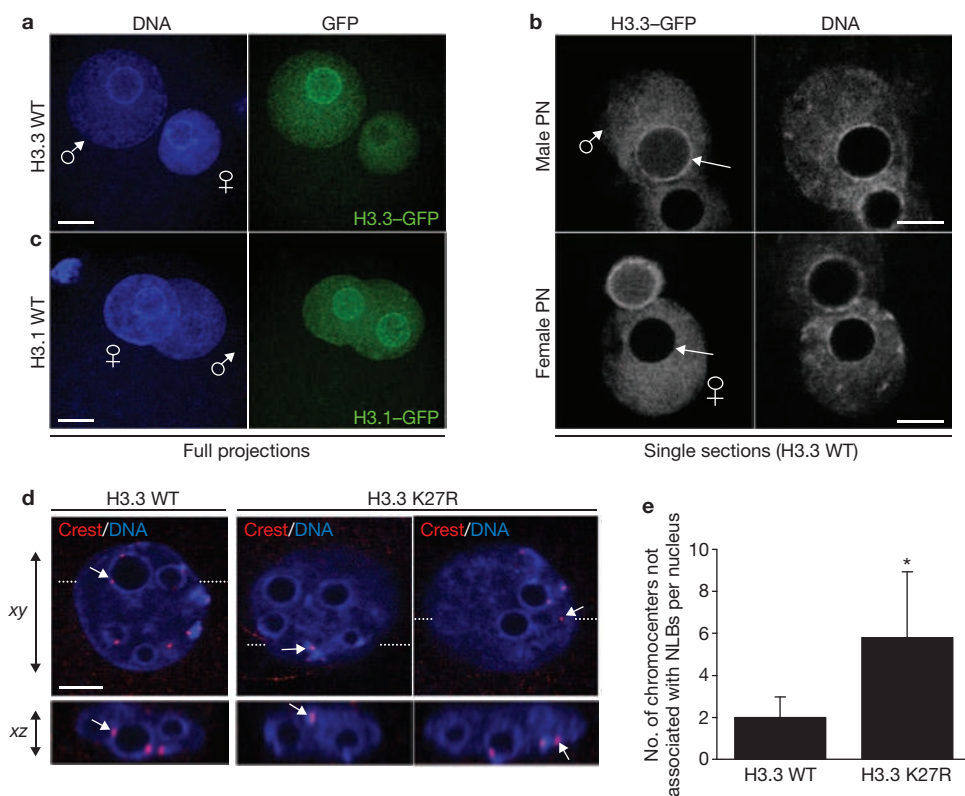


Figure 5 H3.3 localizes to pericentromeric domains surrounding NLBs in the male pronucleus only. (**a, b**) Zygotes injected with mRNA for H3.3-GFP WT were analysed under confocal microscopy ($n = 73$). Shown are full maximal intensity projections (**a**) or single sections (**b**) of the GFP and the DAPI channels in which the diameter of the male (top) or female (bottom) pronucleus (PN) is maximal. Note that H3.3 localizes to the pericentromeric chromatin surrounding the NLBs in the male pronucleus but not in the female pronucleus (arrows). (**c**) H3.1 is distributed throughout both male and female pronuclei. Zygotes injected with mRNA for H3.1-GFP WT were fixed at the same time as the zygote shown in **a**. Shown are full maximal-intensity projections of a zygote (representative of 27 zygotes). Male and female pronuclei are indicated (arrows). (**d, e**) Chromocentre localization is affected by H3.3 K27R expression. (**d**) Confocal microscopy acquisition along the x - y and x - z planes allowed us to determine the spatial

positioning of individual chromocentres (labelled with a Crest antiserum). Most chromocentres occupy a restricted spatial localization around the NLBs at the 2-cell stage, with only one or two chromocentres per nucleus showing no association with NLBs¹⁹. (**e**) In agreement, we found 2 ± 1 non-NLB-associated chromocentres per nucleus in non-injected embryos or in embryos expressing H3.3-GFP WT ($n = 10$). In contrast, spatial localization of chromocentres was affected by the expression of H3.3-GFP K27R, with an increased number of chromocentres not associated with NLBs (6 ± 2 ; $n = 16$). Examples of x - y and x - z planes of one nucleus from H3.3-GFP WT-expressing embryos and two nuclei from H3.3-GFP K27R-expressing embryos. The x - z section at the bottom of **d** corresponds to the section cut at the level of the white dashed line on the x - y plane. Arrows point to the same Crest spots seen from two optical planes. Asterisk, $P = 0.0001$. Error indicates s.d. Scale bar, 5 μ m.

H3.3K27me1 might support transcription of these regions. Although this certainly deserves further investigation, this is in line with the observed correlation between monomethylation of K27 and active transcription¹. Further dimethylation and trimethylation of K27 in the paternal chromatin and recruitment of HP1 β , mediated at least in part through the Hinge region, could ensure subsequent heterochromatinization of pericentromeric chromatin in the absence of H3K9me3 (Fig. 7d). Our findings that H3.3K9R-expressing embryos develop in normal ratios (Supplementary Information, Fig. S7c, d) further support this notion. HP1 β might then act as an architectural component necessary for chromosome segregation and/or for the propagation of heterochromatin after the first mitosis. Reinforcement of additional histone modifications in the second cell cycle will then facilitate heterochromatin maintenance. In this sense, the mouse embryo might be particularly sensitive, because 'naked paternal DNA' will assemble into newly formed chromatin in the moments after fertilization. Indeed, we would not expect such a marked effect would be seen after similar manipulations in, for example, a somatic cell, in which the heterochromatin structure is already established and needs only to

be maintained. Although we cannot formally rule out the possibility that acetylation of K27 is involved in the developmental heterochromatic defects that we describe, this possibility seems unlikely given the pattern of localization of acetylated H3K27 in the zygote (excluded from pericentromeric chromatin) and the sharp decrease to almost undetectable levels at the early 2-cell stage (Supplementary Information, Fig. S8a). Moreover, H3K27 has been reported to be hypoacetylated in centromeres²⁸.

Our results suggest that H3.3 has an essential role during zygotic S phase in the transcription of pericentromeric domains that trigger their silencing after the first cell cycle. The acquisition of K27 methylation by the male pronucleus coincides with S phase (Supplementary Information, Figs S2, S3 and S8b; ref. 4). Transcription during S phase has been shown to be critical for the propagation of heterochromatic structures²⁴. Indeed, S phase could be the first time when newly assembled potential heterochromatin in the male pronucleus will be primed to give rise to specialized chromatin regions such as pericentromeric chromatin. Thus, we have uncovered a novel function for H3.3 in the initial establishment of heterochromatin in the mouse embryo. □

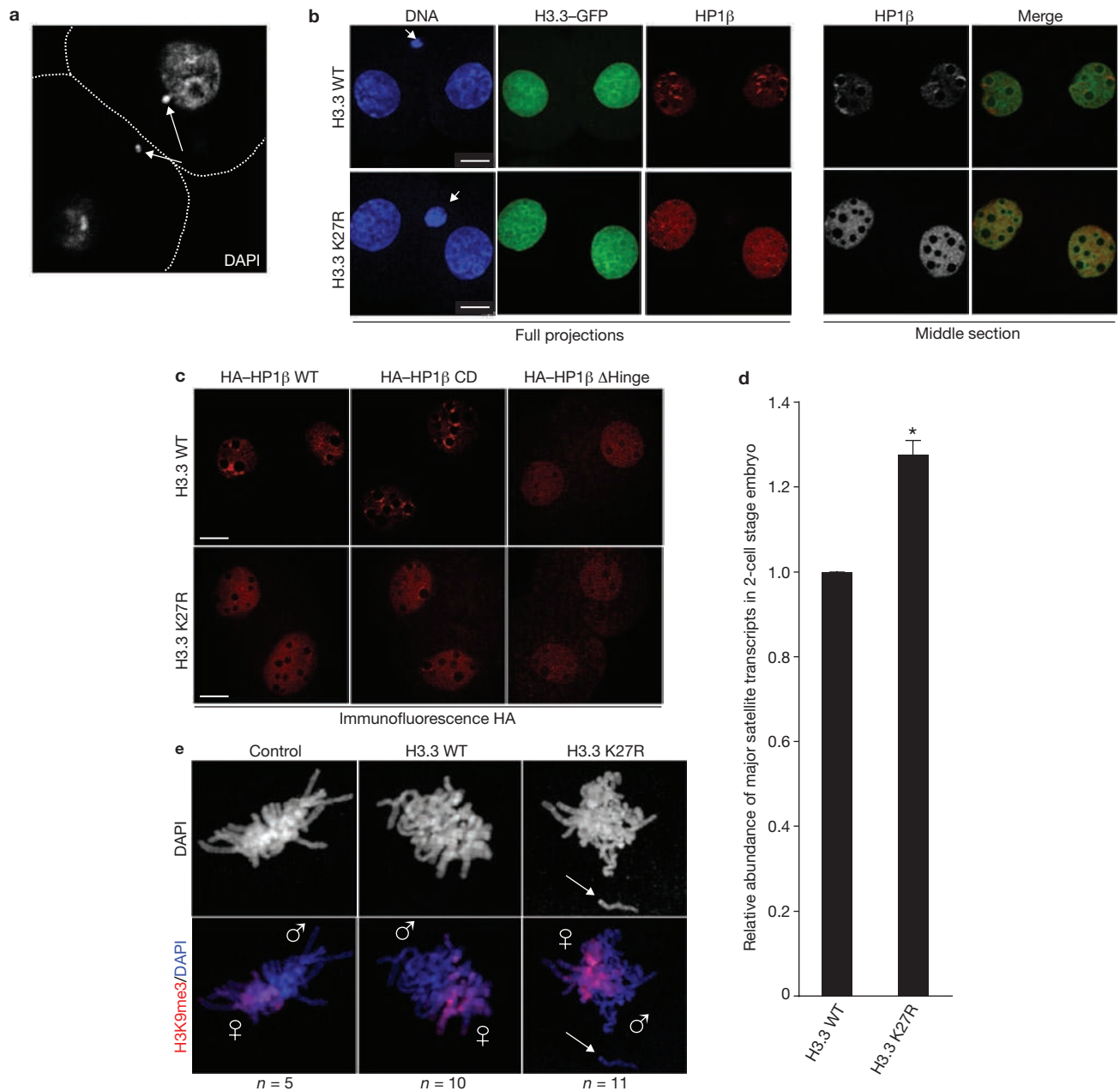


Figure 6 Expression of H3.3-GFP K27R in the zygote leads to altered pericentromeric transcription, defects in chromosome segregation and mislocalization of HP1 β . **(a)** Embryos expressing H3.3 K27R show chromosome segregation defects (24%; $n = 57$) in comparison with embryos expressing H3.3 WT or H3.1 K27R or with non-injected embryos (5%, $n = 30$; 8%, $n = 26$; 5%, $n = 21$, respectively). Shown is a single section of a 2-cell-stage embryo expressing H3.3-GFP K27R stained with DAPI. The white dashed line marks the cell membrane. Arrows indicate DNA fragments that were not incorporated into the nuclei. **(b)** HP1 β is mislocalized in 2-cell-stage embryos expressing H3.3-GFP K27R. Zygotes injected with H3.3-GFP WT ($n = 19$) or H3.3-GFP K27R ($n = 18$) mRNAs were stained with an HP1 β antibody. Shown are full projections (left) or middle sections (right) of images taken along the z axis every 0.6 μm . Embryos were processed in parallel and with identical confocal acquisition parameters. Non-injected embryos are shown in Supplementary Information, Fig. S5. Polar bodies are indicated by arrowheads. Scale bar, 14 μm . **(c)** Deletion of the Hinge domain in HP1 β results in mislocalization of HP1 β similarly to that observed on H3.3 K27R mutation. The indicated

HP1 β mRNA was co-injected with either H3.3-GFP WT or K27R mRNA as in Fig. 1. Embryos were analysed with an antibody against haemagglutinin (HA). Shown are single confocal sections of representative 2-cell-stage embryos. Scale bar, 14 μm . **(d)** Increased accumulation of major satellite transcripts following expression of H3.3 K27R. Zygotes were injected with mRNA for H3.3-GFP WT or H3.3-GFP K27R and cultured until the early 2-cell stage. Total RNA was retrotranscribed and analysed by PCR with specific primers for major satellites. Data were normalized to H2A mRNA levels and are presented as means \pm s.d. for three independent biological replicates. We observed no difference in major satellite transcription between H3.3 WT and non-injected embryos (not shown). Asterisk, $P = 0.0042$. **(e)** Paternal chromatin shows segregation defects during the first mitosis in embryos expressing H3.3 K27R. Non-injected (control), H3.3 WT-expressing or H3.3 K27R-expressing embryos were cultured until they reached mitosis; they were then fixed and stained with an antibody against H3K9me3 (which at this stage marks exclusively maternal chromatin). Paternal chromosomes (those lacking H3K9me3) show severe division defects (arrows).

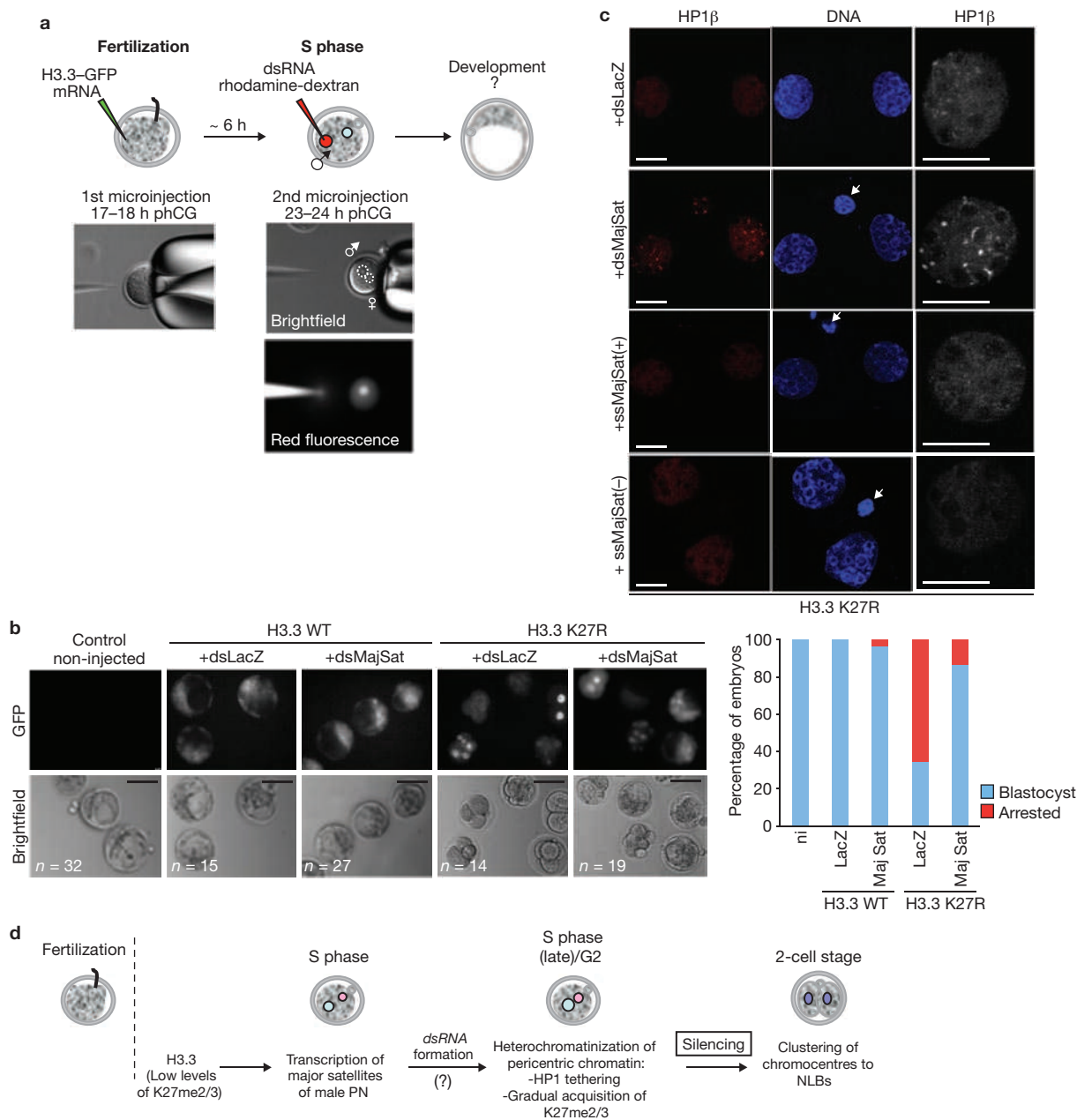


Figure 7 Addition of dsRNA from major satellites in the zygote rescues the developmental phenotype of H3.3 K27R mutants. **(a)** Experimental design for injection of dsRNA in embryos expressing H3.3 WT or H3.3 K27R. The first injection of H3.3-GFP WT or H3.3-GFP K27R was done as in Fig. 1, 17–18 h after injection with human chorionic gonadotropin. Embryos were then cultured for about 6 h until S phase, at which time the male pronucleus was microinjected with dsRNA for LacZ (control; dsLacZ) or for the major satellite repeat (dsMajSat). To control for pronuclear injection, dsRNA was mixed with rhodamine-coupled dextran. **(b)** The developmental phenotype elicited by H3.3 K27R is rescued by the addition of dsRNA of major satellites. Representative embryos after three days of development are shown for each experimental group. The graph on the right depicts the percentage of embryos in each group that reached the blastocyst stage (blue), relative to the control, non-injected (n.i.) embryos, which was set at 100%. The percentages of embryos showing defective development are shown in red. Data were compiled from four independent experiments. **(c)** HP1 β relocates

to DAPI-rich regions after injection of dsRNA for major satellites in H3.3 K27R-expressing embryos. Embryos were injected as in **a**, fixed at the two-cell stage and analysed with an antibody against HP1 β . Shown are representative of ten embryos analysed per group in two independent experiments. For the single-stranded RNA experiments, results shown are for the sense (+) and the antisense (–) RNA (*n* = 13 and *n* = 8, respectively). The right panels show higher magnifications of one of the nuclei shown on the left. Scale bar, 10 μ m. **(d)** Model for heterochromatin establishment at pericentromeric repeats. The presence of H3.3 and low levels or absence of K27 methylation in wild-type embryos provide a chromatin environment for transcription of pericentromeric chromatin during the first S phase in the male pronucleus. Both the transcription of these domains and the gradual accumulation of K27 methylation would subsequently lead to their heterochromatinization and correct nuclear spatial positioning around NLBs at the two-cell stage. Reinforcement of histone modifications in the second cell cycle will then allow heterochromatin maintenance.

METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/naturecellbiology/>

Note: Supplementary Information is available on the Nature Cell Biology website.

ACKNOWLEDGEMENTS

We thank Kami Ahmad for providing H3.3 constructs, Florence Cammas and Tony Kouzarides for HP1 β cDNAs, Natacha Dreumont and James Stevenin for advice on RNA-shift, and Olivier Pourquie for critical reading of the manuscript. A.J.B. acknowledges funding from Cancer Research U.K.; M.E.T.-P. acknowledges funding from Programme National de Recherche Reproduction Endocrinologie (PNRRE) and the Avenir programme from INSERM, and ANR-09-Blanc-0114.

AUTHOR CONTRIBUTIONS

A.S. designed, performed and analysed most of the experiments in this study. C.Z.-B. and M.K. performed experiments. L.T. provided support during the initial phase of this work. A.J.B. contributed to the design of the project. M.E.T.-P. conceived the project, designed and supervised the study and performed experiments. A.S., A.J.B. and M.-E.T.-P. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/naturecellbiology>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>

- Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).
- Arney, K. L., Bao, S., Bannister, A. J., Kouzarides, T. & Surani, M. A. Histone methylation defines epigenetic asymmetry in the mouse zygote. *Int. J. Dev. Biol.* **46**, 317–320 (2002).
- Lepikhov, K. & Walter, J. Differential dynamics of histone H3 methylation at positions K4 and K9 in the mouse zygote. *BMC Dev. Biol.* **4**, 12 (2004).
- Santos, F., Peters, A. H., Otte, A. P., Reik, W. & Dean, W. Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev. Biol.* **280**, 225–236 (2005).
- Torres-Padilla, M. E., Bannister, A. J., Hurd, P. J., Kouzarides, T. & Zernicka-Goetz, M. Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos. *Int. J. Dev. Biol.* **50**, 455–461 (2006).
- Adenot, P. G., Mercier, Y., Renard, J. P. & Thompson, E. M. Differential H4 acetylation of paternal and maternal chromatin precedes DNA replication and differential transcriptional activity in pronuclei of 1-cell mouse embryos. *Development* **124**, 4615–4625 (1997).
- Bouniol-Baly, C., Nguyen, E., Besombes, D. & Debey, P. Dynamic organization of DNA replication in one-cell mouse embryos: relationship to transcriptional activation. *Exp. Cell Res.* **236**, 201–211 (1997).
- Wouters-Tyrou, D., Martinage, A., Chevillier, P. & Sautiere, P. Nuclear basic proteins in spermiogenesis. *Biochimie* **80**, 117–128 (1998).
- Wright, S. J. Sperm nuclear activation during fertilization. *Curr. Top. Dev. Biol.* **46**, 133–178 (1999).
- Wiekowski, M., Miranda, M., Nothias, J. Y. & DePamphilis, M. L. Changes in histone synthesis and modification at the beginning of mouse development correlate with the establishment of chromatin mediated repression of transcription. *J. Cell Sci.* **110**, 1147–1158 (1997).
- van der Heijden, G. W. *et al.* Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech. Dev.* **122**, 1008–1022 (2005).
- Ahmad, K. & Henikoff, S. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* **9**, 1191–1200 (2002).
- Schwartz, B. E. & Ahmad, K. Transcriptional activation triggers deposition and removal of the histone variant H3.3. *Genes Dev.* **19**, 804–814 (2005).
- Hake, S. B. *et al.* Expression patterns and post-translational modifications associated with mammalian histone H3 variants. *J. Biol. Chem.* **281**, 559–568 (2006).
- McKittrick, E., Gafken, P. R., Ahmad, K. & Henikoff, S. Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc. Natl Acad. Sci. USA* **101**, 1525–1530 (2004).
- Polo, S. E., Roche, D. & Almouzni, G. New histone incorporation marks sites of UV repair in human cells. *Cell* **127**, 481–493 (2006).
- Azuara, V. *et al.* Chromatin signatures of pluripotent cell lines. *Nat. Cell Biol.* **8**, 532–538 (2006).
- Tagami, H., Ray-Gallet, D., Almouzni, G. & Nakatani, Y. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* **116**, 51–61 (2004).
- Probst, A. V., Santos, F., Reik, W., Almouzni, G. & Dean, W. Structural differences in centromeric heterochromatin are spatially reconciled on fertilisation in the mouse zygote. *Chromosoma* **116**, 403–415 (2007).
- Muchardt, C. *et al.* Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1 α . *EMBO Rep.* **3**, 975–981 (2002).
- Chen, E. S. *et al.* Cell cycle control of centromeric repeat transcription and heterochromatin assembly. *Nature* **451**, 734–737 (2008).
- Djupedal, I. *et al.* RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. *Genes Dev.* **19**, 2301–2306 (2005).
- Verdel, A. *et al.* RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**, 672–676 (2004).
- Grewal, S. I. & Elgin, S. C. Transcription and RNA interference in the formation of heterochromatin. *Nature* **447**, 399–406 (2007).
- Lu, J. & Gilbert, D. M. Proliferation-dependent and cell cycle regulated transcription of mouse pericentric heterochromatin. *J. Cell Biol.* **179**, 411–421 (2007).
- Puschendorf, M. *et al.* PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat. Genet.* **40**, 411–420 (2008).
- Motamedi, M. R. *et al.* HP1 proteins form distinct complexes and mediate heterochromatic gene silencing by nonoverlapping mechanisms. *Mol. Cell* **32**, 778–790 (2008).
- Mellone, B. G. *et al.* Centromere silencing and function in fission yeast is governed by the amino terminus of histone H3. *Curr. Biol.* **13**, 1748–1757 (2003).

METHODS

Embryo collection and microinjection. Zygotes were collected at the fertilization cone stage from F₁ (C57BL/6 × CBA/H) crosses on superovulation about 17 h after injection of human chorionic gonadotropin. H3.1 (histone cluster 1, H3e), H3.3, HP1β WT, HP1β CD (Lys 41-Trp 42 to Ala-Ala) and HP1β ΔHinge (Δ86–116) mRNAs were transcribed *in vitro* from the pRN3P plasmid with the mMACHINE mMESSAGE Kit (Ambion). All cDNAs were subcloned to include identical 5' untranslated region (UTR) (including KOZAK) and 3' UTR sequences to ensure equivalent expression levels. Zygotes were microinjected with 1–2 pl of the corresponding mRNAs (120 ng μl⁻¹), and their development was monitored at regular intervals. Because it is well established that expression of exogenous histones at levels in excess of just 10–15% of the endogenous can elicit detrimental effects¹⁶, we first experimentally titrated the amount of exogenous histones that embryos could tolerate. For this, we injected various amounts of H3.3 WT mRNA to determine the maximum amount of exogenous mRNA that the embryos could tolerate without altering development (Supplementary Information, Fig. S5a). Arbitrary fluorescence units of H3–GFP in the nucleus were quantified comparatively among the groups of embryos with the use of Metamorph (Universal Imaging). Single-stranded RNA and dsRNA of the 234-base-pair repeat of mouse major satellite or the LacZ sequence was transcribed with the MaxiScript Kit (Ambion)²⁹. Pictures were acquired with a CoolSnap camera attached to a Leica DMIRE2 inverted microscope.

Immunostaining and fluorescence microscopy. Immunostaining and confocal analysis was performed as described⁵ on a Leica TCS SP2/AOBS inverted confocal microscope with a Plan Apo CS (numerical aperture 1.4) 63× objective. Antibodies used were AlexaFluor635-phalloidin (Molecular Probes), anti-Cdx2 (BioGenex), anti-H3K27me1 and anti-H3K27me3 (Upstate-Millipore), anti-H3K9me3 (Upstate-Millipore), anti-HP1β (IGBMC) and anti-Crest antisera (Europa Bioproducts). All other antibodies are described in the legends to figures in the Supplementary Information. Time-lapse imaging was performed with a Plan Apo CS (numerical aperture 0.7) 20× objective on an inverted Leica DMIRE2

microscope and analysed with Metamorph. Images were captured every 10 or 20 min at multiple *z* planes (covering 60 μm) to include both pronuclei.

Chromocentre localization. For chromocentre localization analysis, embryos expressing H3.3–GFP WT or H3.3–GFP K27R mutants were stained with Crest antiserum and DAPI and analysed by confocal microscopy through *z*-sections. Three-dimensional reconstructions and orthogonal views were analysed with Metamorph from *z*-series acquired every 0.6 μm. The number of chromocentres not associated with NLBs per nucleus was determined after three-dimensional reconstitution.

Quantification of fluorescence levels. For quantification of H3K27me levels, pixel intensity values (grey-value intensity of the unprocessed eight-bit images) were determined throughout individual confocal sections covering the complete nucleus (about 19 sections per nucleus) with ImageJ. Values from all sections were added for each nucleus to obtain the total content of H3K27me per nucleus. Controls were analysed under identical conditions and in parallel. For example, in a typical experiment the intensity values of H3K27me3 in H3.3–GFP WT-expressing embryos (*n* = 4) were 1,414.8 ± 25 units per nucleus and those for the H3.3–GFP K27R embryos were 932.2 ± 372 (*n* = 8; error indicates s.d.). Similarly, for H3K27me1, values per nucleus were 372 ± 185 for H3.3–GFP WT embryos (*n* = 6) and 201 ± 49 for H3.3–GFP K27R embryos (*n* = 10).

RT-PCR. Total mRNA was extracted with the PicoPure RNA extraction kit (Arcturus), and RT-PCR analysis for major satellites was performed from pooled embryos as described³⁰ and normalized to H2A mRNA levels.

Statistical analysis. Statistical analysis was performed after a *t*-test or analysis of variance according to the size of the sample groups.

29. Wianny, F. & Zernicka-Goetz, M. Specific interference with gene function by double-stranded RNA in early mouse development. *Nat. Cell Biol.* **2**, 70–75 (2000).

30. Lehnertz, B. *et al.* Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.* **13**, 1192–1200 (2003).

DOI: 10.1038/ncb2089

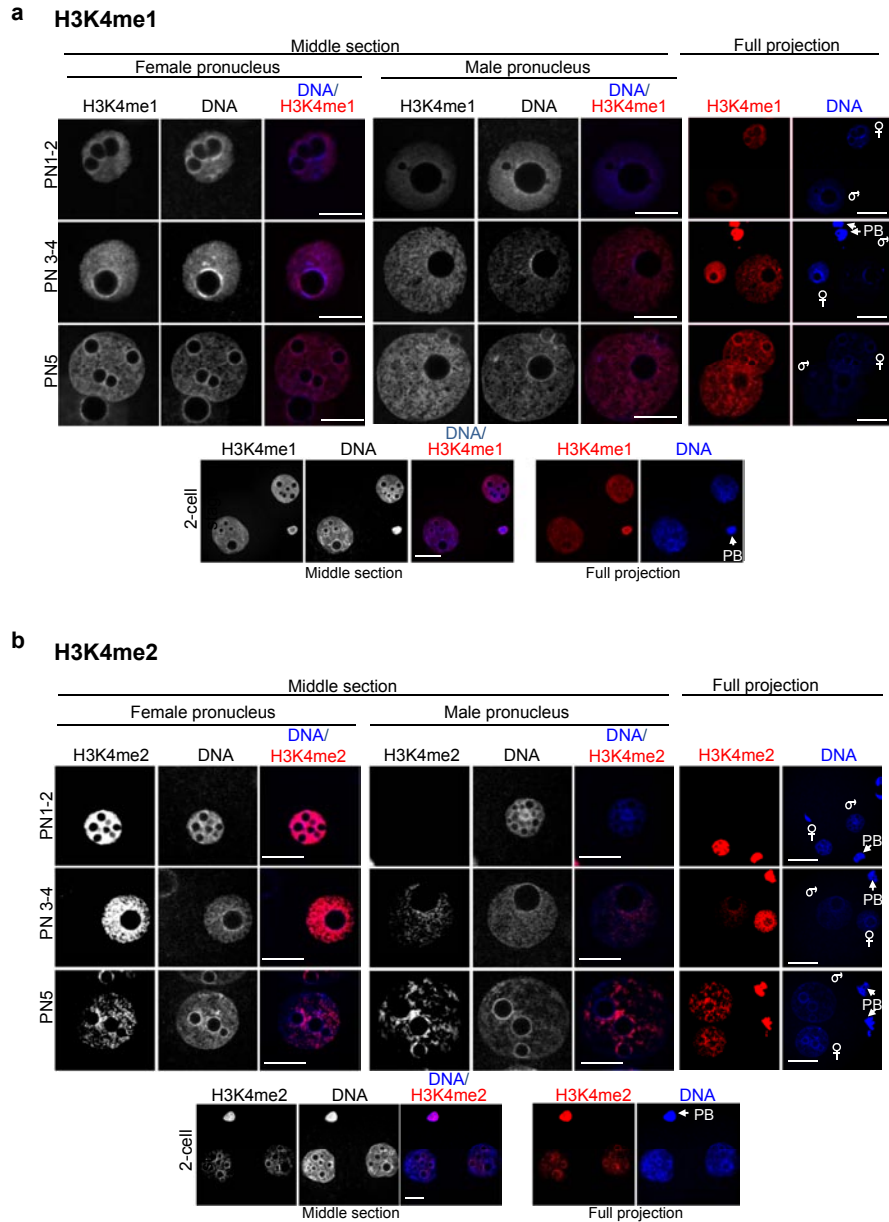


Figure S1 Pattern of localisation of H3K4me1 (a) and H3K4me2 (b) during zygotic development and in 2-cell stage embryos. Zygotes at different pronuclear (PN) stages or 2-cell stage embryos were collected from CD1 X CD1 crosses and processed immediately for immunostaining with a H3K4me1 antibody (abcam ab8895) (a) or a H3K4me2 antibody (abcam ab7766) (b). In all figures, PN stages were determined according to Adenot et al.¹. Note that PN1-2 are zygotic stages before DNA replication; S-phase takes place mostly during PN3 and PN4, and PN5 is the latest pronuclear stage. Shown are representative zygotes at the PN1-2, PN3-4

and PN5 stages and a 2-cell stage embryo. Single sections where female (left) and male (right) pronuclei diameter was maximal are shown (Middle section). Full (maximal) projections of Z-sections acquired every 1 µm are also shown on the right. The confocal parameters used for the acquisition were identical and fluorescence levels are therefore comparable. Male and Female pronuclei are indicated in the full projection panel. Scale bar is 10 µm. Where visible, the position of the polar body (PB) is indicated. Embryos shown are representative of three independent experiments with at least 16 embryos analysed.

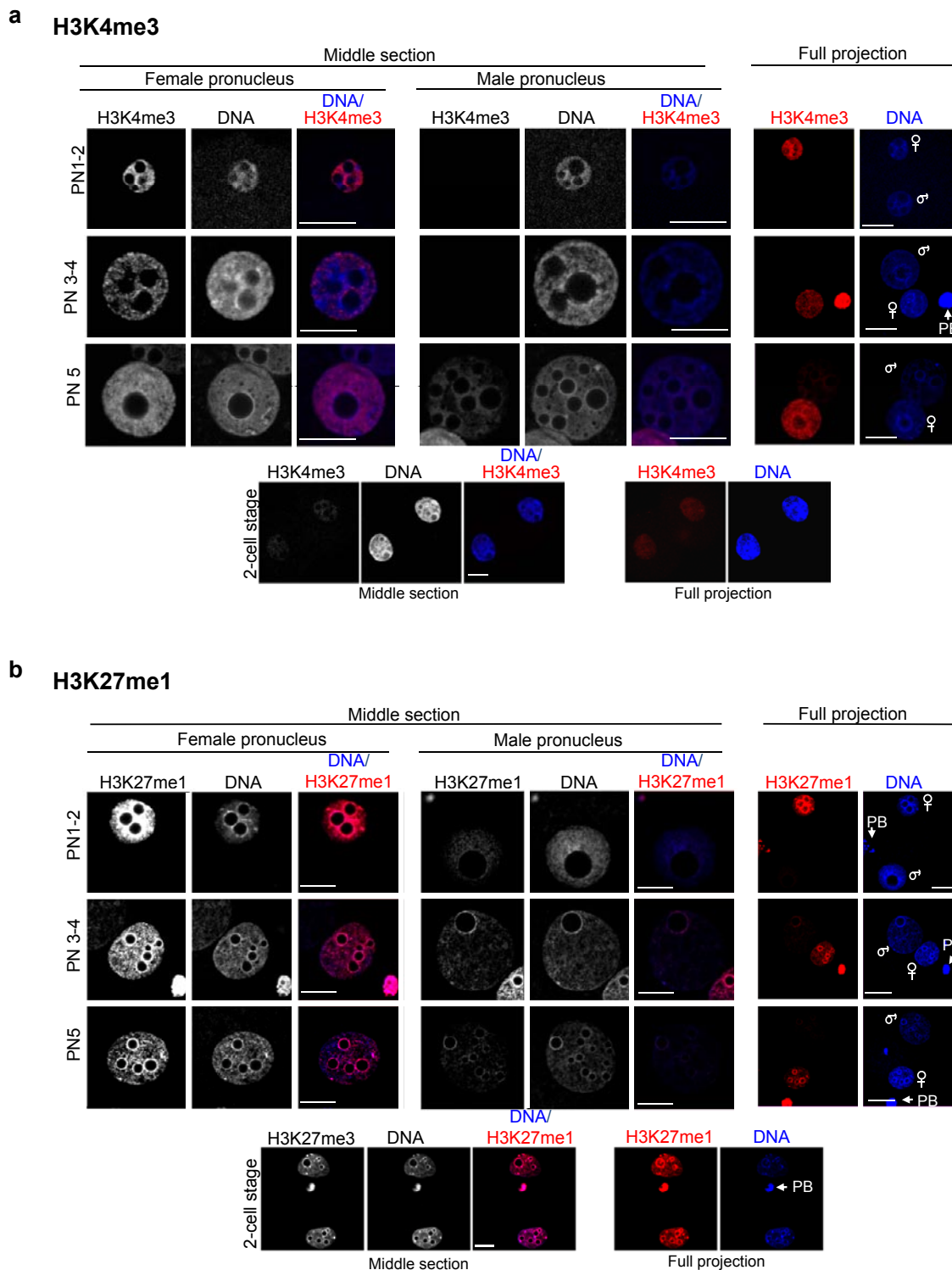


Figure S2 Pattern of localisation of H3K4me3 (a) and H3K27me1 (b) during zygotic development and in 2-cell stage embryos. Zygotes at different pronuclear (PN) stages or 2-cell stage embryos were collected from CD1 X CD1 crosses and processed immediately for immunostaining with a H3K4me3 antibody (Upstate/Millipore 04-745) (a) or a H3K27me1 antibody (Upstate/Millipore 07-448) (b). Shown are representative zygotes at the PN1-2, PN3-4 and PN5 stages and 2-cell stage. Single sections where female (left) and

male (right) pronuclei diameter was maximal are shown (Middle section). Full projections of Z-sections acquired every 1 μm are also shown on the right. The confocal parameters used for the acquisition were identical and fluorescence levels are therefore comparable. Male and Female pronuclei are indicated in the full projection panel. Scale bar is 10 μm . Where visible, the position of the polar body (PB) is indicated. Embryos shown are representative of three independent experiments with at least 17 embryos analysed.

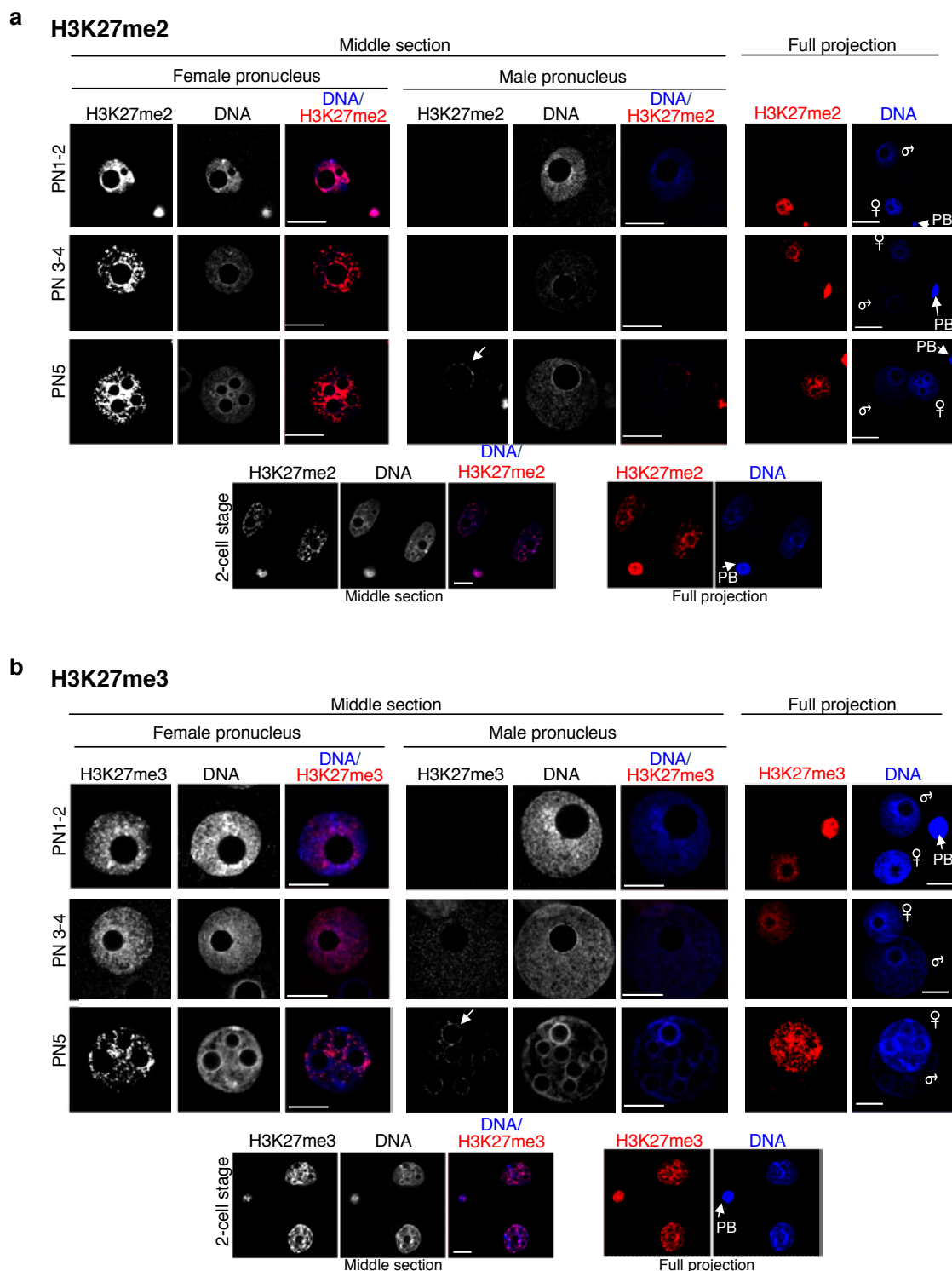


Figure S3 Accumulation of H3K27me2 (a) and H3K27me3 (b) during zygotic development and in 2-cell stage embryos. Zygotes at different pronuclear (PN) stages or 2-cell stage embryos were collected from CD1 crosses and processed immediately for immunostaining with a H3K27me2 antibody (abcam ab24684) (a) or a H3K27me3 antibody (Upstate/Millipore 07-449) (b). Shown are representative zygotes at the PN1-2, PN3-4 and PN5 stages and a 2-cell stage embryo. Single sections where female (left) and

male (right) pronuclei diameter is maximal are shown (Middle section). Full projections of Z-sections acquired every 1 μm are also shown on the right. The confocal parameters used for acquisition were identical and fluorescence levels are therefore comparable. Male and Female pronuclei are indicated in the full projection panels. Scale bar is 10 μm . Where visible, the position of the polar body (PB) is indicated. Embryos shown are representative of two independent experiments with at least 14 embryos analysed.

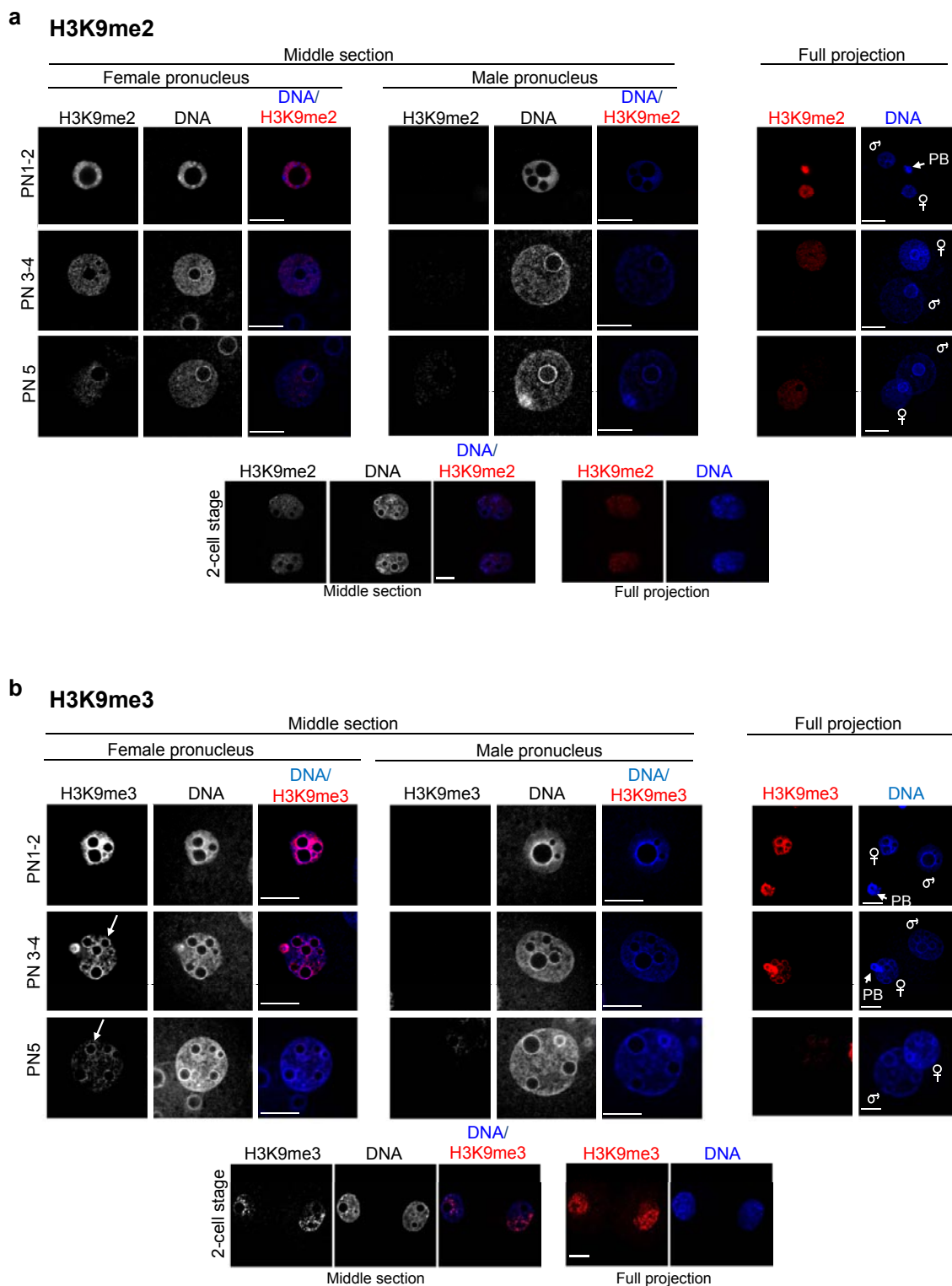


Figure S4 Pattern of localisation of H3K9me2 (a) or H3K9me3 (b) during zygotic development and in 2-cell stage embryos. Zygotes at different pronuclear (PN) stages or 2-cell stage embryos were collected from CD1 X CD1 crosses and processed immediately for immunostaining with a H3K9me2 antibody (Upstate/Millipore 07-441) or a H3K9me3 antibody (Upstate/Millipore 07-442). Shown are representative zygotes at the PN1-2, PN3-4 and PN5 stages and a 2-cell stage embryo. Note that the fact that the two parental genomes do not mix can be seen at the 2-cell stage in (b), where H3K9me3 stains strongly only half nucleus (e.g. maternal chromatin). Single sections where female (left) and

male (right) pronuclei diameter was maximal are shown (Middle section). Full projections of Z-sections acquired every 1 μ m are also shown on the right. The confocal parameters used for the acquisition were identical and fluorescence levels are therefore comparable. Male and Female pronuclei are indicated in the full projection panel. Note that H3K9me3 is enriched around the nucleolar-like bodies NLBs in the female pronucleus (arrow)(See also ref 2), but is undetectable in the male pronucleus. Scale bar is 10 μ m. Where visible, the position of the polar body (PB) is indicated. Embryos shown are representative of three independent experiments with at least 18 embryos analysed.

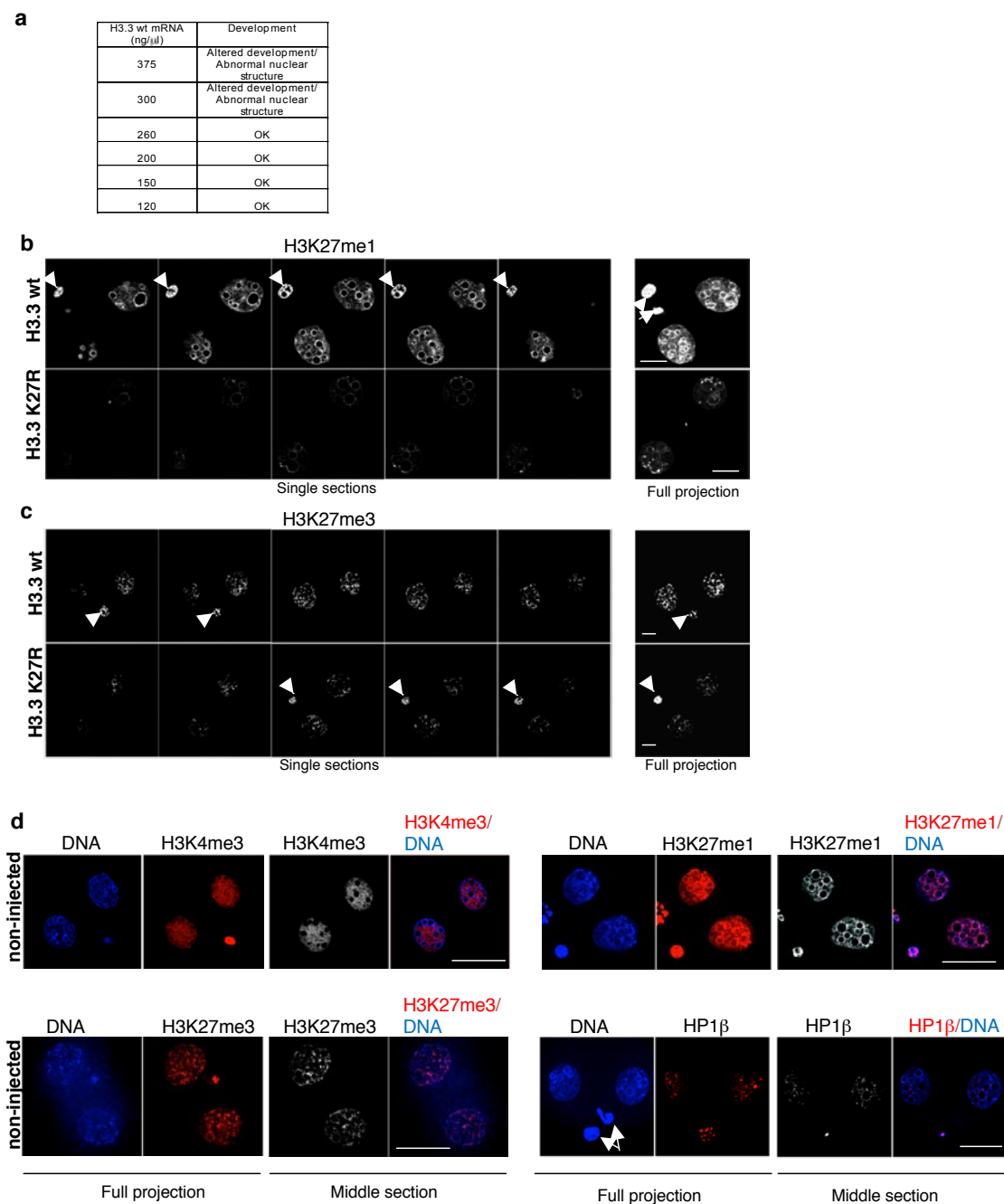


Figure S5 a. Establishment of conditions for expressing exogenous histones in early embryos. Because it is well established that expression of exogenous histones at levels in excess of just 10–15% of the endogenous can elicit detrimental effects (for example, see ref 3), we first titrated experimentally the amount of exogenous histones that embryos can tolerate. For this, we injected various amounts of H3.3 wt mRNA to determine the maximum amount of exogenous mRNA that the embryos could tolerate without altering development. We found that embryos displayed detrimental effects when greater than 260ng/ μ l mRNA was injected. We observed no developmental defects in multiple experiments when 260ng/ μ l or less was injected. Although our approach does not provide a direct quantitation of the amount of exogenous histone relative to endogenous, it does indicate the 'safe' level of histone mRNA that can be injected without eliciting detrimental effects. **b-c.** Effects of H3.3 -GFP wild type or K27R mutant on global H3K27me1 and H3K27me3. Zygotes were microinjected at the fertilisation cone stage with mRNA for H3.3 -GFP wild type or K27R mutant, cultured until the 2-cell stage, fixed

and analysed using a H3K27me1 (b) or a H3K27me3 antibody (c). Shown are middle sections (left) or full projections (right) of confocal Z-series of representative embryos shown in Figure 4a-b. For each antibody, embryos were processed in parallel and acquisition was performed under identical confocal parameters. Where visible, the polar body is depicted by an arrowhead. Scale bar is 10 μ m. **d.** Distribution of H3K4me3, H3K27me1, H3K27me3 and HP1 β in non-injected embryos at the 2-cell stage. Zygotes were collected at the fertilisation cone stage and were either microinjected with the corresponding mRNAs (see main text) or left non-injected. All groups of embryos were cultured until the 2-cell stage at which time they were fixed and analysed with HP1 β , H3K27me3, H3K27me1 or H3K4me3 antibodies as indicated (red). Note that quantification of H3K27me1 and H3K27me3 levels did not reveal any difference between non-injected controls and embryos expressing H3.3-GFP wt. Shown are full projections (left) or a middle section (right) of Z-series taken every 0.5 μ m of representative embryos of at least 18 embryos analysed for each antibody. The polar body is depicted by an arrow. Scale bar is 10 μ m.

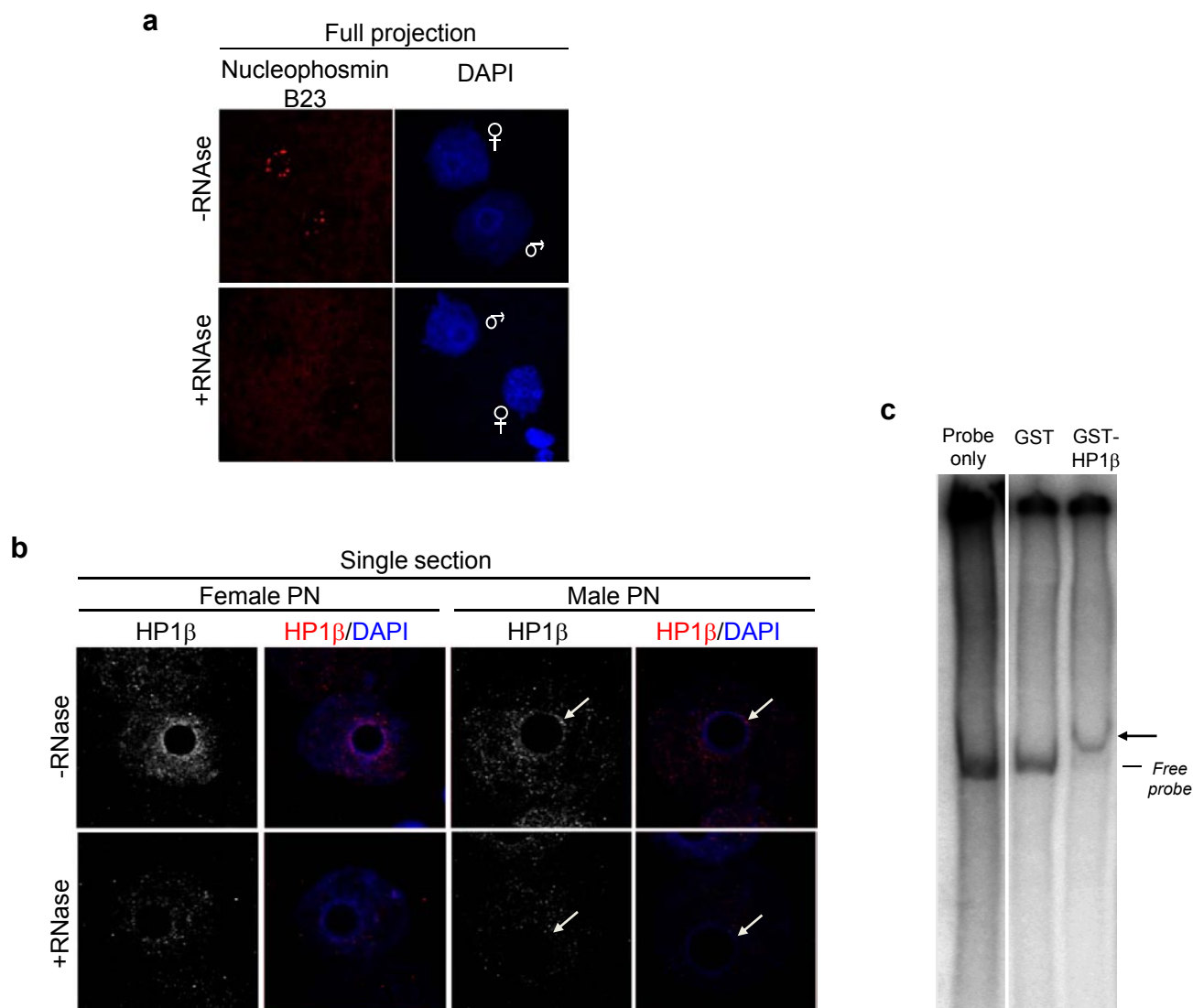


Figure S6 RNAse treatment in zygotes results in mislocalisation of HP1β and HP1β is able to bind major satellite transcripts. **a.** RNAse treatment in the zygote leads to mislocalisation of Nucleophosmin B23. Shown are full projections of images acquired every 1 μm of representative embryos of 6 analysed. DNA was stained with DAPI. Zygotes were permeabilised in Triton 0.1% PBS, treated with RNAse, fixed, processed for immunostaining with a Nucleophosmin B23 antibody (kindly provided by M. Oulad-Abdelghani)(a) or with the HP1β antibody (b) and analysed under confocal microscopy. Because the size of the embryo is significantly bigger than routinely used cultured cells, we first established conditions to achieve an efficient RNAse treatment using the nucleolar protein NucleophosminB23 as control (a), as it is known

that Nucleophosmin B23 localisation depends on RNA. Embryos with each antibody were analysed in parallel and using identical confocal parameters. **b.** HP1β localisation is altered upon RNAse treatment. Shown are single confocal sections of two PN5 zygotes where female and male pronuclei diameter are maximal. DNA was stained with DAPI. Embryos shown are representative of 15 embryos analysed and three independent experiments. Note that the permeabilisation treatment results in high background and the quality of the images is compromised. Arrows in the male pronucleus point to the NLBs. **c.** HP1β binds to pericentromeric transcripts. RNA-shift assay was performed with recombinant GST-HP1β and major satellite RNA that was transcribed and radioactively labelled in vitro. GST-HP1β-RNA shift is indicated by an arrow.

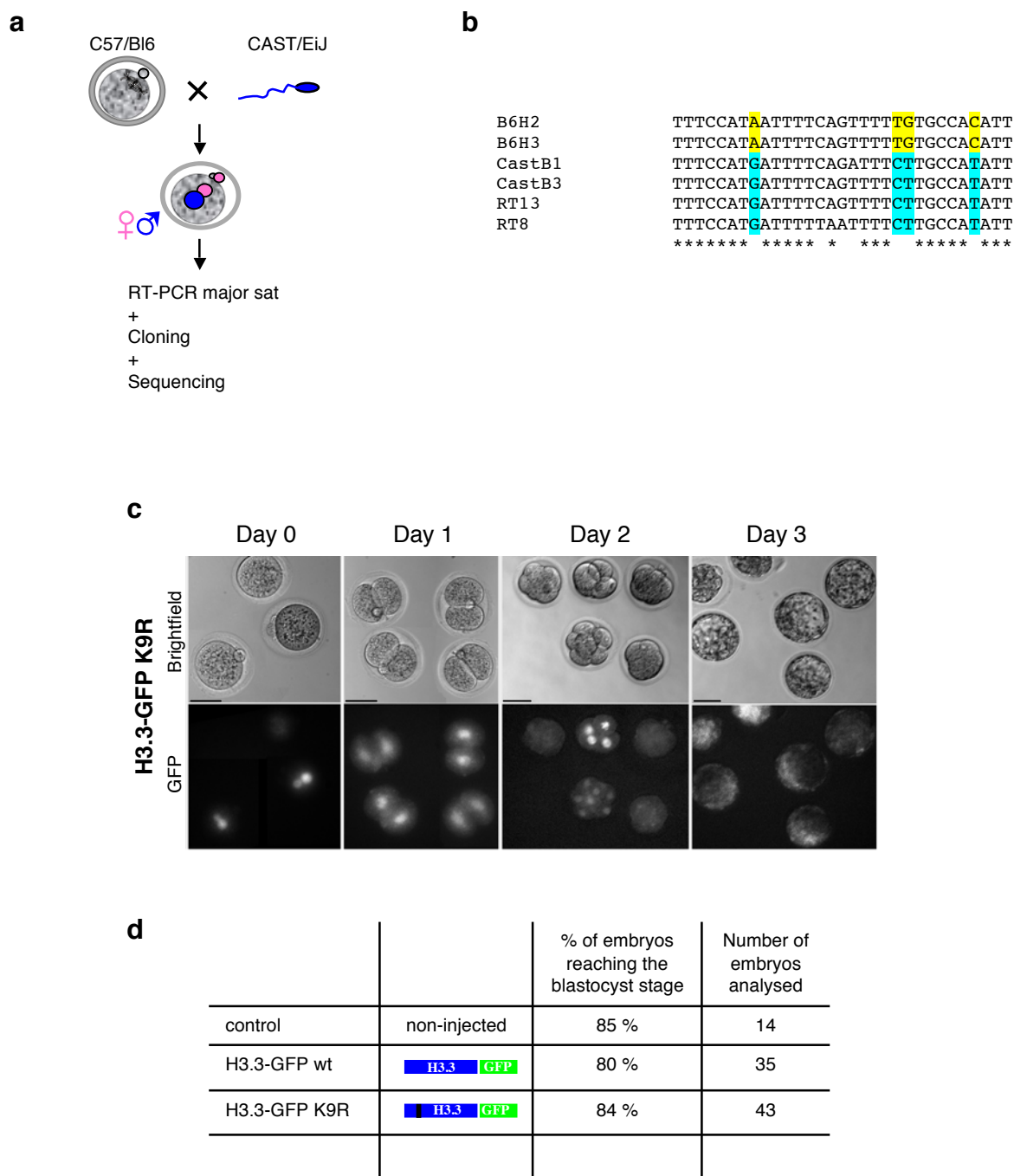


Figure S7 a-b. Major satellites are primarily transcribed from the paternal genome in the mouse zygote. **a.** RNA from zygotes derived from C57/B16mat X CAST/EiJpat crosses was extracted, reversed-transcribed and amplified with primers specific for major satellites⁴. PCR products were cloned and individual clones were sequenced. Note that transcription from major satellite repeats can generate dsRNA in ES cells^{5L2}. **b.** Alignment of a fragment of genomic *Mus musculus domesticus* (C57B16) and *Mus musculus castaneus* (CAST/EiJ) major satellites sequences compared to the sequences obtained from cDNA from mouse zygotes. Shown on the alignment are two representative sequences obtained from each of the three groups analysed: genomic DNA from C57B16 (B6H2-H3), genomic DNA from CAST/EiJ (CastB1-B3) or from two representative of 31 clones obtained after RT-PCR and cloning from mouse

zygotes. All RT-PCR clones sequenced were of the Cast/EiJ background. The blue and yellow color code highlights the nucleotides shared by the CAST/EiJ background and those by the C57B16 one, respectively. **c-d.** Development of embryos expressing H3.3 - GFP K9R. **c.** Mutation of K9 of H3.3 does not appear to compromise development to the blastocyst stage. Zygotes were collected and microinjected with in vitro transcribed mRNA and cultured till the blastocyst stage as indicated in Figure 1. Zygotes were monitored daily and scored for their developmental stage. Shown are representative brightfield and fluorescence images of embryos imaged daily. Scale bar is 50 μ m. **d.** Summary of development of control, non-injected embryos or embryos expressing H3.3 - GFP wt or K9R of 2 independent experiments. The total number of embryos analysed per group is indicated.

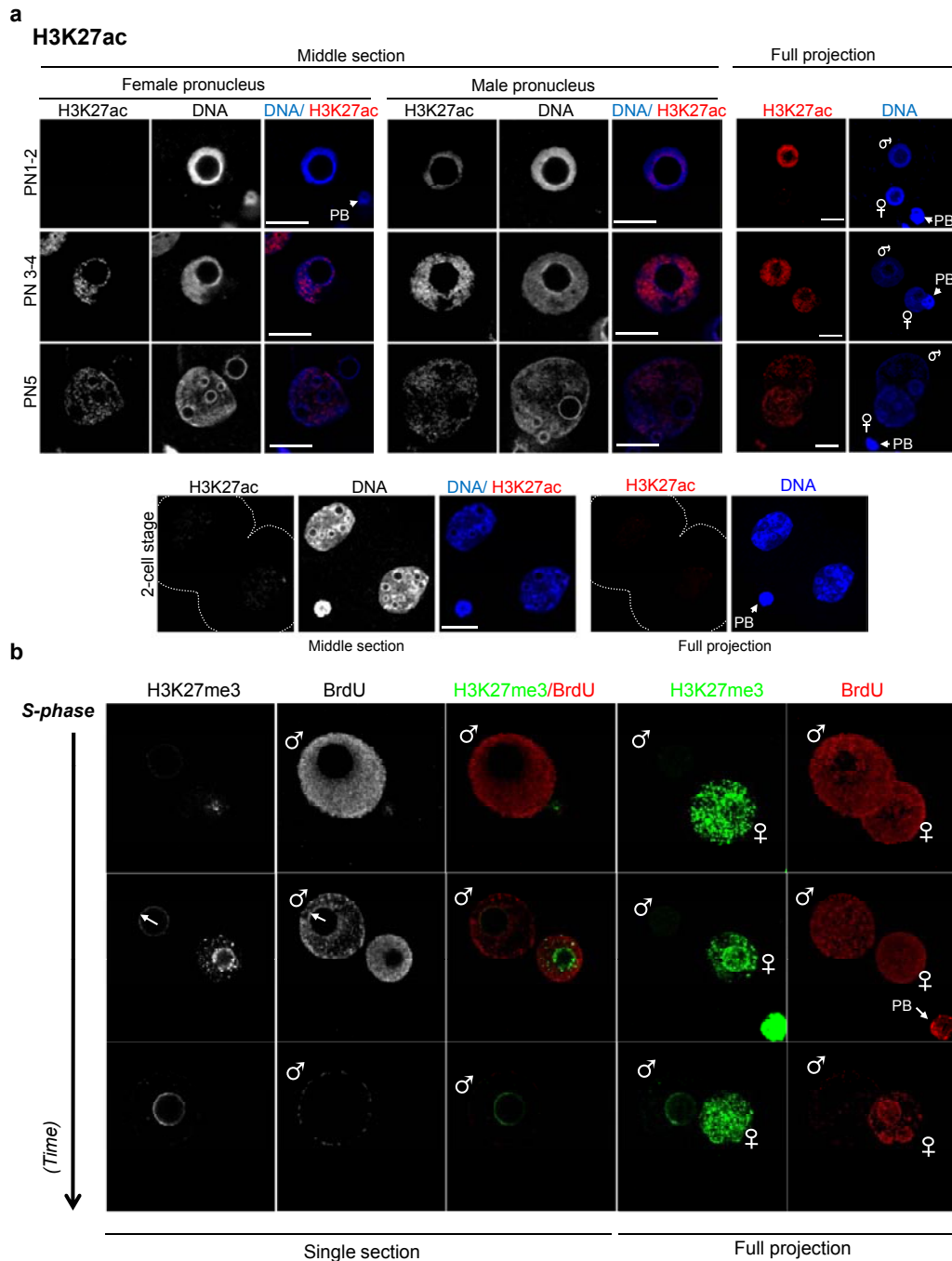


Figure S8 a. Pattern of H3K27ac in zygotes and 2-cell stage embryos. Zygotes or 2-cell stage embryos were collected from CD1 crosses and fixed immediately for immunostaining with a H3K27ac antibody (Active Motif AM39133). Shown are representative zygotes at the PN2, PN3-4 and PN5 stages (top) or 2-cell stage (bottom). Single sections where female (left) and male (right) pronuclei diameter is maximal are shown (Middle section). Full projections of Z-sections acquired every 0.6 μm are also shown on the right. The confocal parameters used for acquisition were identical and fluorescence levels are therefore comparable. Male and Female pronuclei are indicated in the full projection panels. Scale bar is 10 μm . Where visible, the position of the polar body (PB) is indicated. The white dashed line in the 2-cell stage embryo delineates the cell membrane. Embryos shown are representative of two independent experiments with at least 16 embryos analysed. **b.** Relationship between acquisition of H3K27me3 and S-phase progression. Zygotes were subjected to a BrdU pulse

of 30 minutes during mid- and late S-phase and processed for immunostaining with an anti-BrdU and an anti-H3K27me3 antibody. Representative zygotes of 20 analysed at different stages of S-phase are shown chronologically from the top to the bottom. The zygote shown in the top panel has not undergone replication of the satellites sequences around the NLBs (pattern A). In the middle panel, BrdU labelling is detected around the NLB in the male pronucleus (arrow), as it is H3K27me3 (arrow). Note that in the bottom panel, the male pronucleus has undergone replication almost completely (pattern E), with only the peripheral nuclear regions labelled with BrdU. These patterns are according to the replication patterns described by Bouniol-Baly et al⁶. Shown are maximal projections of Z-stack sections taken every 0.5 μm (right panels) or single sections (left panels) where the diameter of the male pronucleus is maximal. Male and female pronuclei are indicated. Where visible, the 2nd polar body (PB) is indicated.

References

1. Adenot, P. G., Mercier, Y., Renard, J. P. & Thompson, E. M. Differential H4 acetylation of paternal and maternal chromatin precedes DNA replication and differential transcriptional activity in pronuclei of 1-cell mouse embryos. *Development* **124**, 4615-25 (1997).
2. Santos, F., Peters, A. H., Otte, A. P., Reik, W. & Dean, W. Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev Biol* **280**, 225-36 (2005).
3. Polo, S. E., Roche, D. & Almouzni, G. New histone incorporation marks sites of UV repair in human cells. *Cell* **127**, 481-93 (2006).
4. Lehnertz, B. et al. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* **13**, 1192-200 (2003).
5. Martens, J. H. et al. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *Embo J* **24**, 800-12 (2005).
6. Bouniol-Baly, C., Nguyen, E., Besombes, D. & Debey, P. Dynamic organization of DNA replication in one-cell mouse embryos: relationship to transcriptional activation. *Exp Cell Res* **236**, 201-211 (1997).

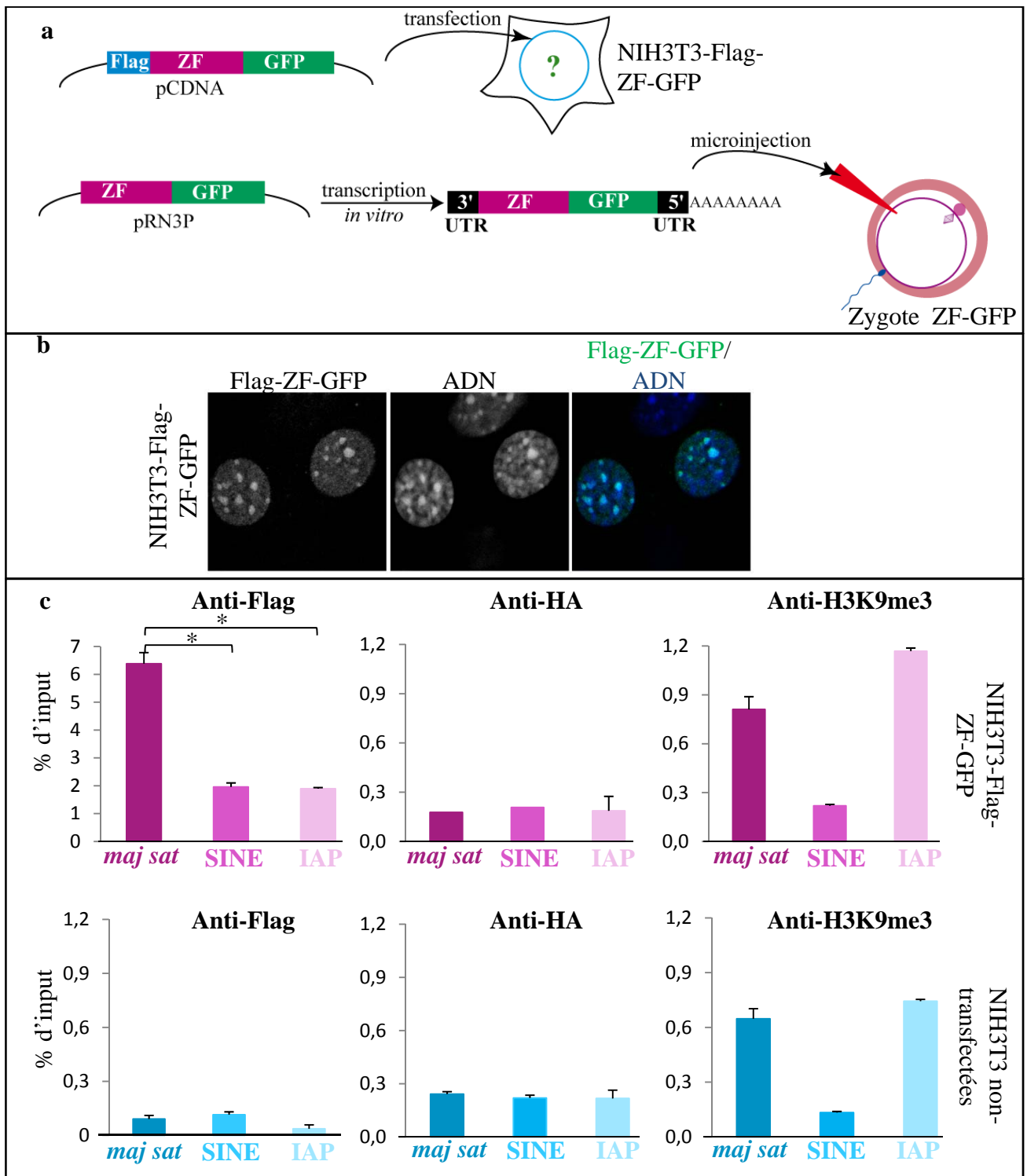


Figure 36 : La protéine Flag-ZF-GFP se localise aux régions de chromatine très condensée et lie spécifiquement les *major satellites* dans les cellules NIH3T3. **a** : Des cellules NIH3T3 sont transfectées avec un plasmide codant pour la construction Flag-ZF-GFP. Des embryons sont injectés au moment de la fécondation avec l'ARNm codant pour la protéine ZF-GFP transcrit *in vitro*. La construction est ensuite localisée dans le noyau des cellules et des embryons. **b** : La protéine Flag-ZF-GFP a été localisée dans les cellules transfectées comme montré en **a** grâce à la fluorescence de la GFP. L'ADN est marqué au DAPI. Les acquisitions ont été faites sur un microscope confocal et une section simple est présentée ici. **c** : Des ChIP contre les protéines Flag (gauche), HA (milieu) et H3K9me3 (droite) ont été réalisés sur des cellules NIH3T3 transfectées (NIH3T3-ZF-GFP-Flag, en haut) ou non (NIH3T3 non transfectées, en bas) comme en **a**. * : p-value < 0,005

C. Résultats non publiés

I. Chronologie de la localisation, de la condensation, et de l'acquisition des marques des histones de la chromatine péricentrique dans le zygote de souris.

Afin de préciser le rôle que peuvent jouer la lysine 27 de H3 et HP1 β au cours de la condensation de la chromatine péricentrique dans le pronoyau mâle, nous avons voulu établir une chronologie d'acquisition de ces marques au niveau de la chromatine péricentrique au cours du stade zygote. Pour localiser la chromatine péricentrique, nous avons utilisé une protéine contenant trois doigts de zinc reconnaissant une séquence de neuf nucléotides spécifique des *major satellites* (Lindhout et al., 2007). Après avoir vérifié la spécificité de cette protéine (notée ZF) pour les *major satellites* par ChIP dans des cellules NIH3T3, nous avons suivi sa localisation, ainsi que l'apparition de différentes marques de la chromatine, dans le zygote au cours de son développement (figure 36a).

1. La protéine ZF est un outil de localisation des *major satellites* dans les cellules

Le plasmide contenant l'ADNc codant pour la protéine ZF fusionnée au marqueur fluorescent GFP et à une étiquette Flag (notée Flag-ZF-GFP) sous la dépendance d'un promoteur CMV nous a généreusement été envoyé par le Dr Van der Zaal. Afin de vérifier l'efficacité de cette construction, nous avons transfecté des cellules NIH3T3 avec ce plasmide et avons localisé la fluorescence de la GFP dans les cellules (figure 36b). Dans toutes les cellules transfectées, la GFP se situe dans le noyau et est enrichie dans les régions montrant un marquage dense en DAPI, régions où sont localisés les *major satellites*. Cependant, un bruit de fond de fluorescence assez intense dans tout le noyau nous a amené à nous poser la question de la spécificité de cette protéine.

Afin de déterminer si les *major satellites* sont reconnus spécifiquement par la construction Flag-ZF-GFP, nous avons effectué une expérience de ChIP avec un anticorps anti-Flag sur les cellules NIH3T3 transfectées ou non avec le plasmide. Des expériences de ChIP avec des anticorps anti-H3K9me3 et anti-HA ont été utilisées comme contrôles positifs

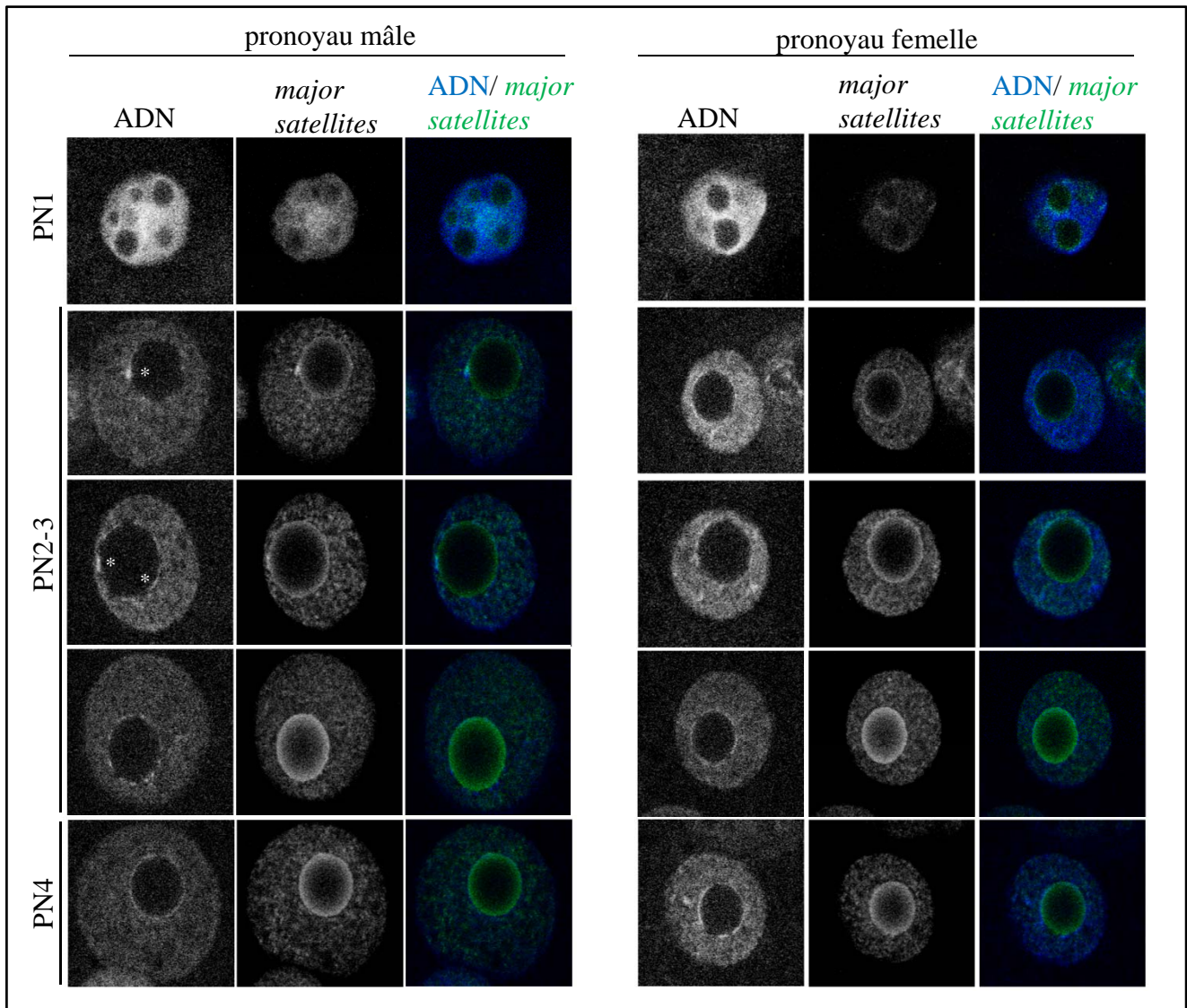


Figure 37 : Condensation de la chromatine péracentrique et localisation des *major satellites* au cours du développement du zygote. Des zygotes ont été injectés comme sur la figure 36a. Les embryons ont été fixés à différents temps après la fécondation et la localisation de ZF-GFP a été analysée en microscopie confocale. L'ADN est marqué au DAPI. Les régions considérées comme condensées sont notées avec une étoile sur les deux premiers pronoyaux mâles au stade PN2-3. Des sections simples sont présentées ici et le stade de développement des zygotes est précisé. Sur une même ligne, les pronoyaux mâle et femelle appartiennent au même embryon.

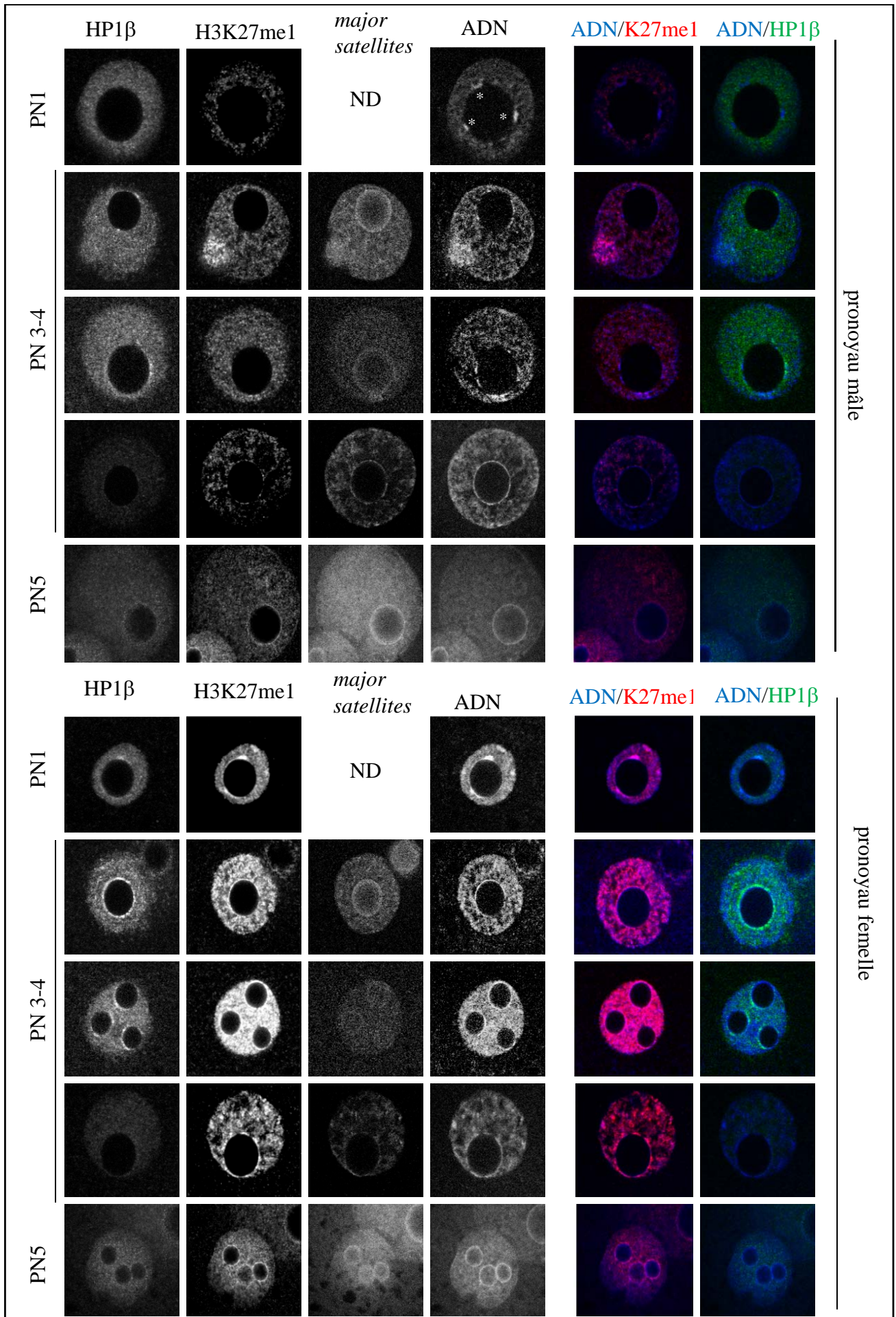
et négatifs respectivement (figure 36c). La liaison de ces protéines aux *major satellites* ainsi qu'à d'autres séquences répétées présentes dans le génome, SINE et IAP, a été testée. Les *major satellites* sont liés de manière significativement plus importante que les autres séquences testées par la protéine Flag-ZF-GFP. La construction lie donc spécifiquement les régions péricentriques de la chromatine. Néanmoins, si on compare les enrichissements observés entre les cellules non transfectées et les cellules transfectées, la construction HA-ZF-émerine semble également lier, de manière moins importante, les autres séquences testées. Ceci expliquerait le bruit de fond observé en fluorescence.

2. Localisation et condensation des régions péricentriques au cours du développement du zygote

Afin d'exprimer la protéine ZF-GFP (dépourvue d'étiquette Flag) dans l'embryon, l'ADNc codant pour la protéine ZF-GFP est transcrit *in vitro* et l'ARNm produit est injecté à des zygotes au moment de la fécondation (figure 36a). La localisation de la protéine ZF-GFP est ensuite analysée par fluorescence au cours du développement et comparée à la localisation des régions présentant un marquage intense en DAPI (figure 37). Etant donné que le marquage intense en DAPI ne colocalise pas, à tout moment, avec le signal des *major satellites*, nous en avons déduit qu'il n'était pas seulement dû à la présence de régions d'ADN riches en adénine et thymine, mais que l'intensité du marquage au DAPI pouvait être également liée au degré de condensation de la chromatine. Nous avons donc considérées comme condensées les régions présentant un marquage intense en DAPI.

Comme précédemment publié, nous avons observé qu'à PN1, les *major satellites* se localisent au centre du pronoyau mâle alors qu'ils se situent préférentiellement autour des NLB en formation dans le pronoyau femelle. A partir du stade PN2, les *major satellites* forment un anneau, ici vu comme continu, autour du NLB dans les pronoyaux mâle et femelle. Dans le pronoyau femelle, le marquage au DAPI est intense sur tout l'anneau formé par les *major satellites* à l'exception de quelques segments. Les *major satellites* maternels sont donc condensés dès le début du stade zygote. Dans le pronoyau mâle en revanche, avant le stade PN4, le marquage au DAPI autour des NLB n'est pas continu mais ponctuel. La condensation des anneaux formés par l'ADN péricentrique n'est donc pas complète dans le pronoyau mâle avant le stade PN4 alors qu'elle l'est dans le pronoyau femelle.

D'après ces observations, nous pouvons donc conclure que la localisation de la chromatine péricentromérique mâle autour des NLB précède sa condensation.



3. Evolution des marques des histones au niveau de la chromatine péricentrique au cours du stade zygote

Pour essayer de déterminer qui, de la méthylation de H3K27 et de l'acquisition de HP1 β , pouvait être le premier à l'origine de la condensation des régions péricentriques maternelles, nous avons voulu établir la chronologie d'acquisition de H3K27me1, H3K27me3 et HP1 β en fonction du niveau de condensation de ces régions au cours du stade zygote. Pour cela, nous avons effectué des marquages immunofluorescents avec des anticorps anti-H3K27me1, anti-H3K27me3 et anti-HP1 β dans des embryons injectés avec l'ARNm codant pour la protéine ZF-GFP (figure 38, figure 39).

Dans le pronoyau femelle, H3K27me1, H3K27me3 et HP1 β sont localisées au niveau de l'anneau d'ADN péricentrique condensé à partir de PN2. HP1 β semble ensuite disparaître des anneaux aux stades PN4-5. Dans le pronoyau mâle, en revanche, H3K27me3 commence à être détectée au niveau du NLB seulement à PN5 quand tout l'anneau est condensé. HP1 β se localise au niveau de l'anneau mâle de manière ponctuelle à partir de PN2-3 mais ne semble pas colocaliser systématiquement avec des régions condensées. Un enrichissement en H3K27me1 apparaît également ponctuellement aux anneaux maternels à partir de PN2-3, il ne colocalise jamais avec les sites enrichis en HP1 β , et seulement parfois avec l'ADN condensé. Au stade PN5, H3K27me1 est enrichie tout autour des NLB.

Ces données montrent que H3K27me1, H3K27me3, et HP1 β apparaissent après les premières étapes de condensation de la chromatine péricentrique maternelle. De plus, H3K27me1 et HP1 β ne colocalisant pas, il semble que l'acquisition de ces deux marques soit indépendante. Il est cependant possible qu'après initiation de la condensation de la chromatine péricentrique maternelle, H3K27me et HP1 β agissent selon deux voies parallèles pour propager et maintenir un état condensé favorable à la réalisation de la première mitose.

Figure 38 : Localisation de HP1 β , H3K27me1 et des major satellites au cours du développement zygotique. Des zygotes sont injectés comme dans la figure 36a. Les embryons sont fixés à différents temps après injection et soumis à une expérience d'immunocytochimie avec des anticorps anti HP1 β (IGBMC) et anti H3K27me1 (Millipore). La localisation de ZF-GFP est analysée grâce à la fluorescence de la GFP. L'ADN est marqué au DAPI. Les portions de l'anneau autour du NLB considérées comme condensées sont notées avec une étoile. Les images sont prises grâce à un microscope confocal. Des sections simples sont présentées ici et le stade de développement des zygotes est précisé. Les pronoyaux mâle et femelle d'une même ligne appartiennent au même embryon.

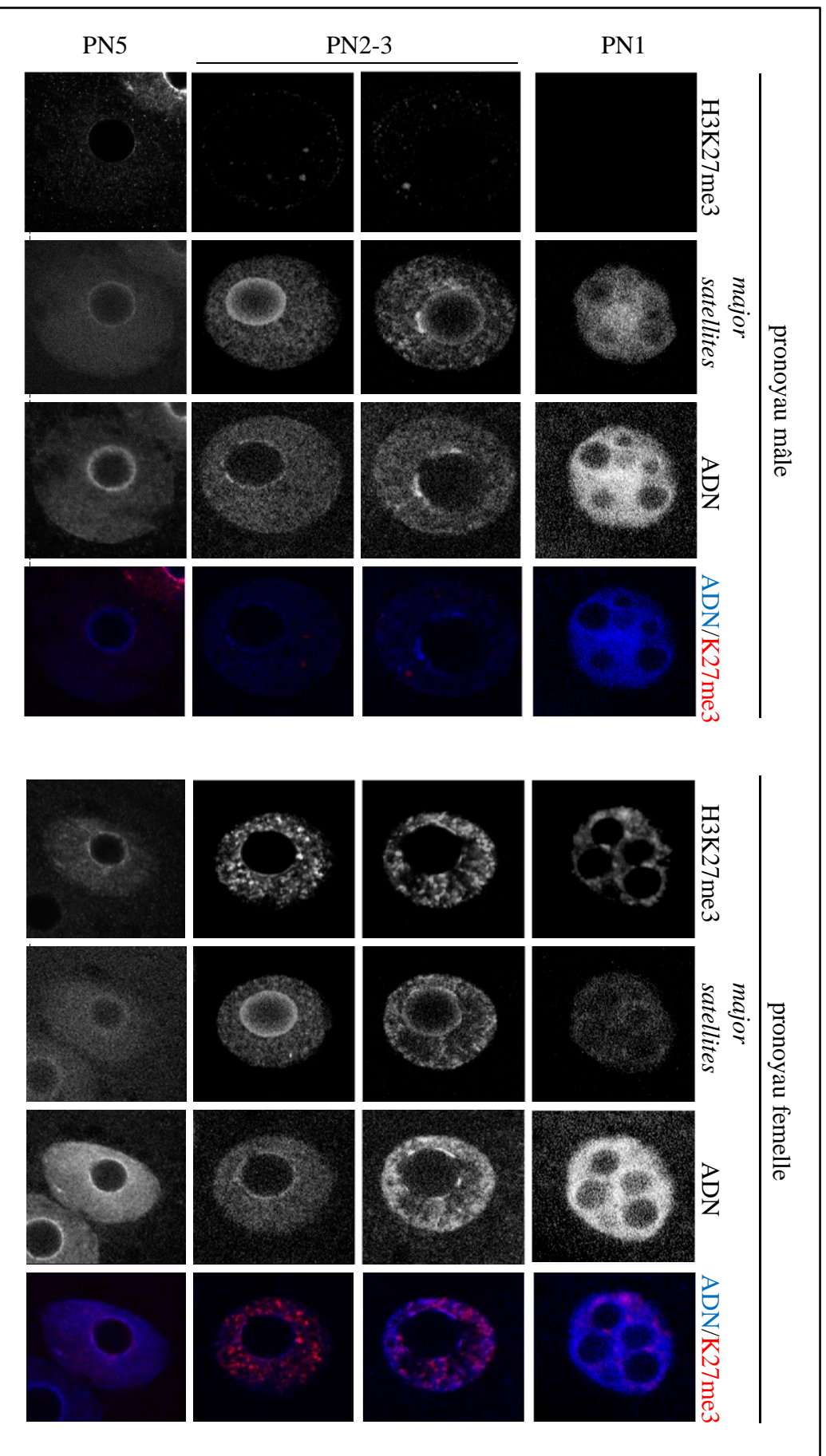


Figure 39 : Localisation de H3K27me3 et des *major satellites* au cours du développement zygotique. Des zygotes sont injectés avec l'ARNm codant pour la protéine ZF-GFP au moment de la fécondation. Les embryons sont fixés à différents temps après injection et soumis à une expérience d'immunocytochimie avec des anticorps anti H3K27me3 (Millipore). La localisation de ZF-GFP est analysée grâce à la fluorescence de la GFP. L'ADN est marqué au DAPI. Les images sont prises grâce à un microscope confocal. Des sections simples sont présentées ici et le stade de développement des zygotes est précisé. Les pronoyaux mâle et femelle d'une même ligne appartiennent au même embryon. Les stades de développement des zygotes sont renseignés à gauche.

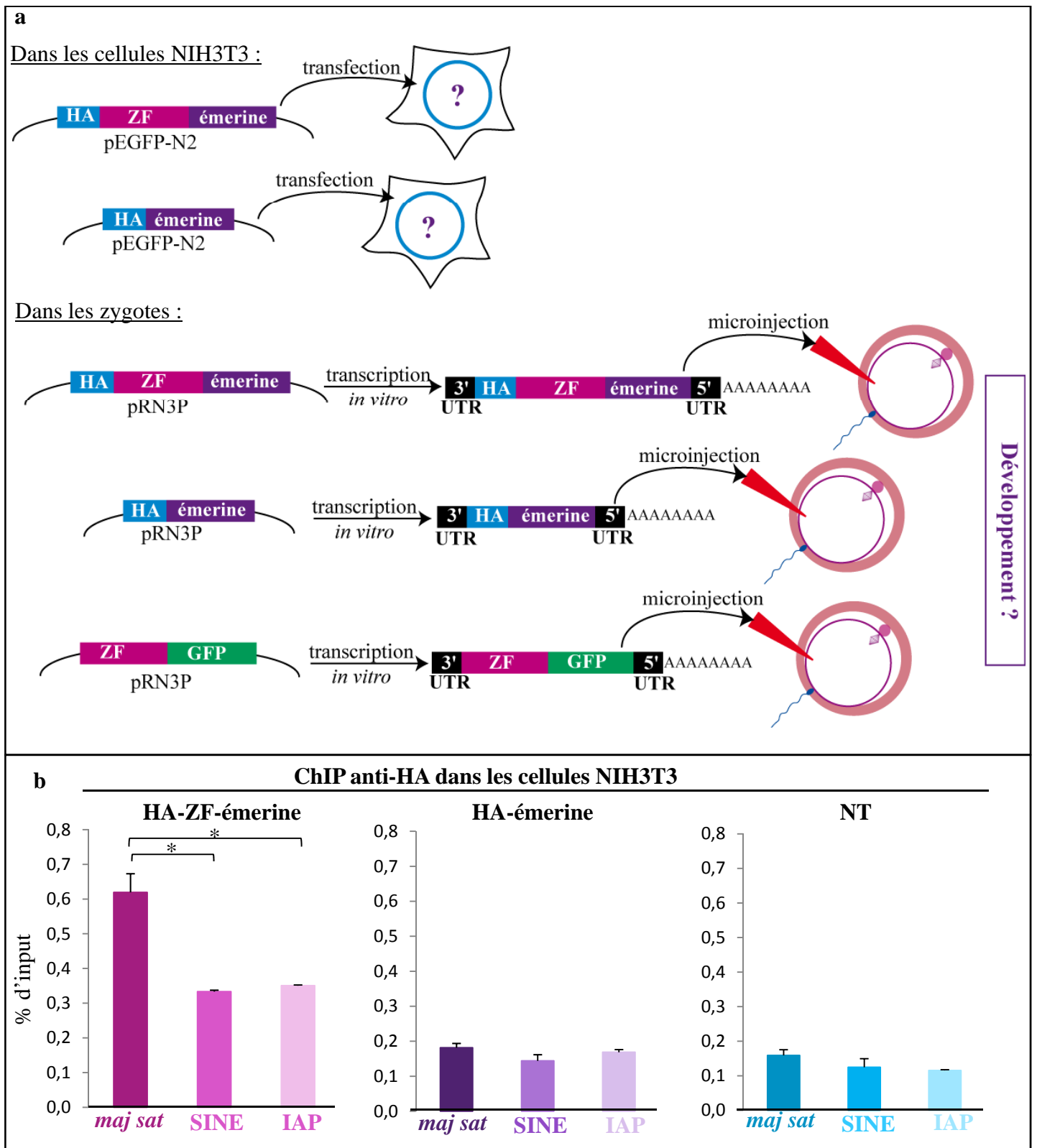


Figure 40 : La construction HA-ZF-émerine transfectée dans des cellules NIH3T3 ou injectée à des embryons sous forme d'ARNm reconnaît spécifiquement les *major satellites*. **a** : Le plasmide pEGFP-N2 contenant les constructions HA-ZF-émerine ou HA-émerine est transfecté dans des cellules NIH3T3. Les séquences codant pour ces mêmes constructions ont été clonées dans un plasmide pRN3P et transcrites *in vitro*. Les ARNm sont ensuite injectés dans des embryons au moment de la fécondation. **b** : Un ChIP anti-HA a été réalisé sur des cellules NIH3T3 non transfectées (NT) ou exprimant HA-émerine ou HA-ZF-émerine. La liaison de ces protéines aux *major satellites* et aux séquences SINE et IAP a été testée. Les résultats sont montrés en enrichissement des séquences par rapport à l'input. * : p-value<0,001

II. Délocalisation des centromères dans l'embryon précoce de souris

1. Problématique et principe de l'expérience

Comme nous l'avons abordé en introduction et comme nous l'avons observé dans la partie précédente, les *major satellites* subissent deux changements de localisation majeurs au cours du développement précoce. Le premier changement a lieu au début du stade zygote. Entre les stades PN1 et PN2, la chromatine péricentromérique paternelle acquiert, à partir d'une position centrale dans le noyau, une structure en anneau autour des NLB. A ce même stade, la chromatine péricentromérique maternelle perd sa structure en chromocentres pour se localiser également en anneau autour des NLB. Cette première relocalisation semble précéder l'acquisition du marquage intense en DAPI autour des NLB. La deuxième relocalisation a lieu au stade 2-cellules tardif. La structure en anneau de la chromatine péricentromérique est alors perdue et les *major satellites* et *minor satellites* forment des chromocentres, structure conservée dans la plupart des cellules somatiques. De plus, des liens temporels existent entre la transcription des *major satellites* et la réorganisation de leur localisation au sein des noyaux (figure 29b) (Probst and Almouzni, 2008) (Probst et al., 2010). Dans la publication 2, nous avons également montré qu'en présence de la protéine H3.3K27R, les centromères présentent des défauts de localisation au stade 2-cellules, suggérant un lien entre le niveau de condensation des régions péricentriques et leur localisation aux NLB. Suite à ces observations, nous nous sommes demandé si la localisation des régions péricentromériques aux NLB au début du développement, puis leur structuration en chromocentres sont nécessaires à leur transcription, leur condensation et donc au développement des embryons. Pour répondre à cette question, nous avons voulu modifier la localisation de la chromatine péricentrique dans le zygote et d'en observer les effets sur le développement.

Dans le but de modifier la localisation des séquences péricentriques dans le noyau, nous avons choisi de cibler ces régions à la périphérie du noyau. Pour ceci, nous avons réalisé une construction (notée HA-ZF-émerine) où la protéine ZF, reconnaissant les *major satellites*, est fusionnée à l'émerine, une protéine intramembranaire de la face interne de la membrane nucléaire liant les lamines A et B (Fairley et al., 1999) (Manilal et al., 1996) (figure 40a). La présence d'une étiquette HA permet la révélation de la protéine en immunocytochimie. L'efficacité et la spécificité de cette construction ont d'abord été testées dans des cellules NIH3T3. Puis l'ARNm codant pour la protéine HA-ZF-émerine a été injecté dans des

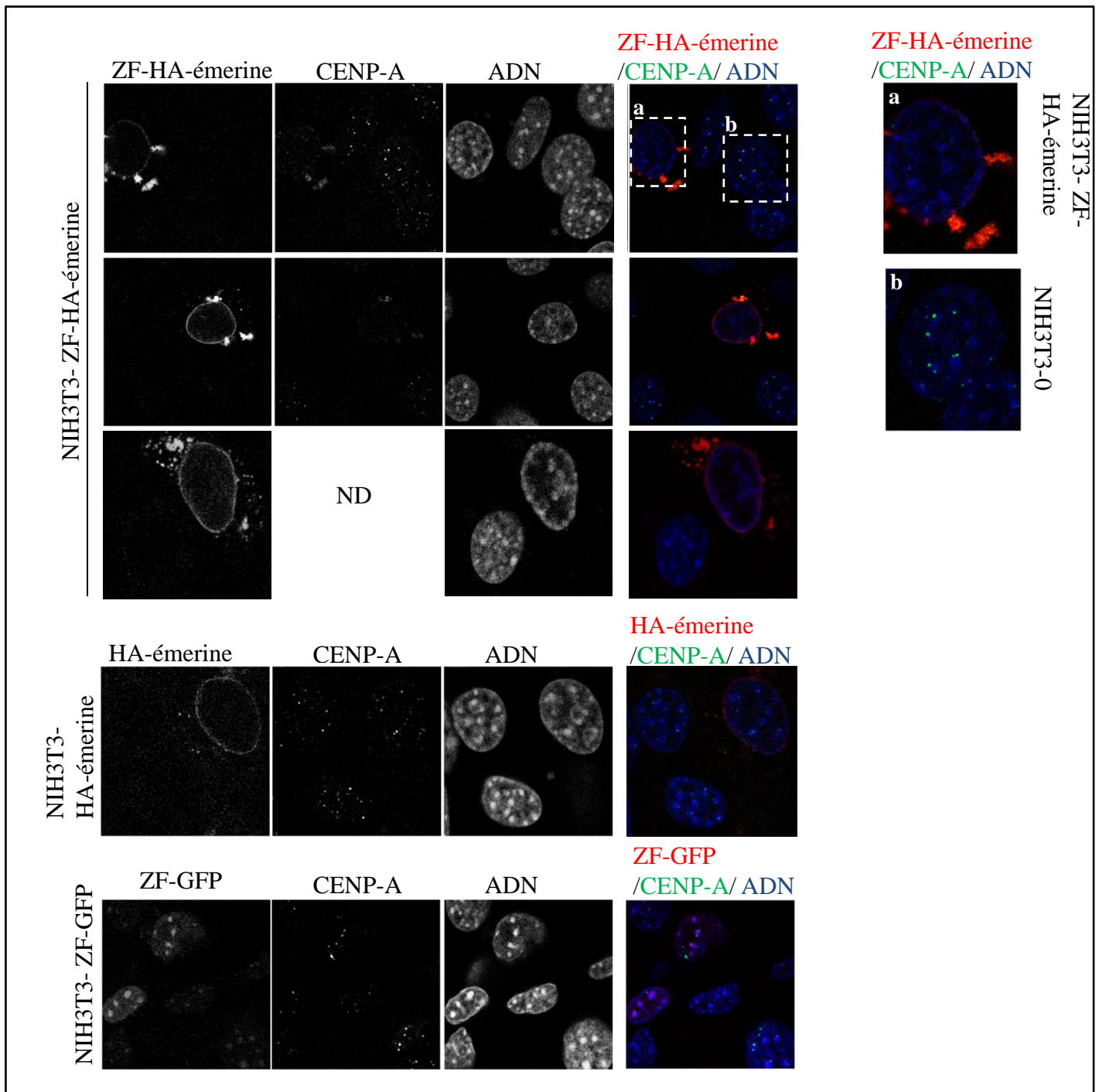


Figure 41 : Les cellules NIH3T3 exprimant la protéine HA-ZF-émerine montrent une relocalisation globale de leur ADN à la périphérie du noyau et une diminution du marquage CENP-A. Des cellules NIH3T3 sont transfectées comme présenté sur la figure 40a. Après fixation, les cellules sont marquées avec un anticorps anti-HA (Roche) et un sérum anti-CENP-A (antisérum anti-Crest, Europa Bioproducts) et analysées sous un microscope confocal. L'ADN est marqué au DAPI. Les images présentées ici sont des sections prises selon un plan unique. **a** et **b** représentent un grossissement de ce qui est présenté dans les rectangles pointillés. **a** montre une cellule transfectée et **b** une cellule non transfectée (NIH3T3-0). ND : non déterminé.

embryons au moment de la fécondation, avant la formation des pronoyaux. Pour ces expériences, deux constructions contrôles ont été utilisées : la protéine ZF-GFP précédemment décrite et une protéine fusion entre l'émerine et une étiquette HA, HA-émerine. Les résultats présentés dans cette partie sont des résultats préliminaires et plusieurs expériences restent à faire pour qu'une conclusion fiable soit tirée des résultats, elles seront abordées dans la partie discussion et perspectives.

2. L'expression de la protéine HA-ZF-émerine induit une relocalisation de la chromatine dans les cellules NIH3T3

Nous avons tout d'abord voulu vérifier que la protéine ZF, même fusionnée à l'émerine, est toujours capable de reconnaître spécifiquement les *major satellites*. Nous avons donc transfecté des cellules NIH3T3 avec les plasmides codant pour les constructions HA-ZF-émerine (NIH3T3-HA-ZF-émerine) et HA-émerine (NIH3T3-HA-émerine), et nous avons réalisé un ChIP avec un anticorps anti-HA sur ces cellules (figure 40b). Comme précédemment, les répétitions SINE et IAP sont utilisées comme contrôles. La construction HA-ZF-émerine reconnaît toujours spécifiquement les *major satellites*. Tout comme la protéine ZF-GFP, HA-ZF-émerine semble reconnaître également les autres répétitions testées mais ceci de manière significativement moins importante. Le contrôle négatif HA-émerine ne lie en revanche pas l'ADN.

Nous avons ensuite voulu savoir si notre construction permet la relocalisation des régions péricentriques dans les cellules NIH3T3. Pour localiser les régions péricentromériques et tester la présence de nos constructions, nous avons réalisé un co-marquage immunocytochimique avec un anticorps reconnaissant CENP-A, un variant de l'histone H3 spécifique des centromères, d'une part, et un anticorps anti-HA d'autre part (figure 41). Alors qu'un marquage de CENP-A est clairement visible dans les cellules non transfectées et les cellules transfectées avec les constructions contrôles, le marquage est beaucoup plus diffus dans les cellules NIH3T3-HA-ZF-émerine. Certains points correspondant à des centromères sont cependant visibles en périphérie du noyau. De plus, en analysant la répartition de l'ADN au sein des noyaux, on observe que l'ADN forme une bordure à la périphérie des noyaux de la totalité des cellules NIH3T3-HA-ZF-émerine. Dans des cas extrêmes, ne correspondant pas aux cellules les plus fortement transfectées, une très grande proportion de l'ADN est relocalisée en périphérie du noyau, le marquage au DAPI

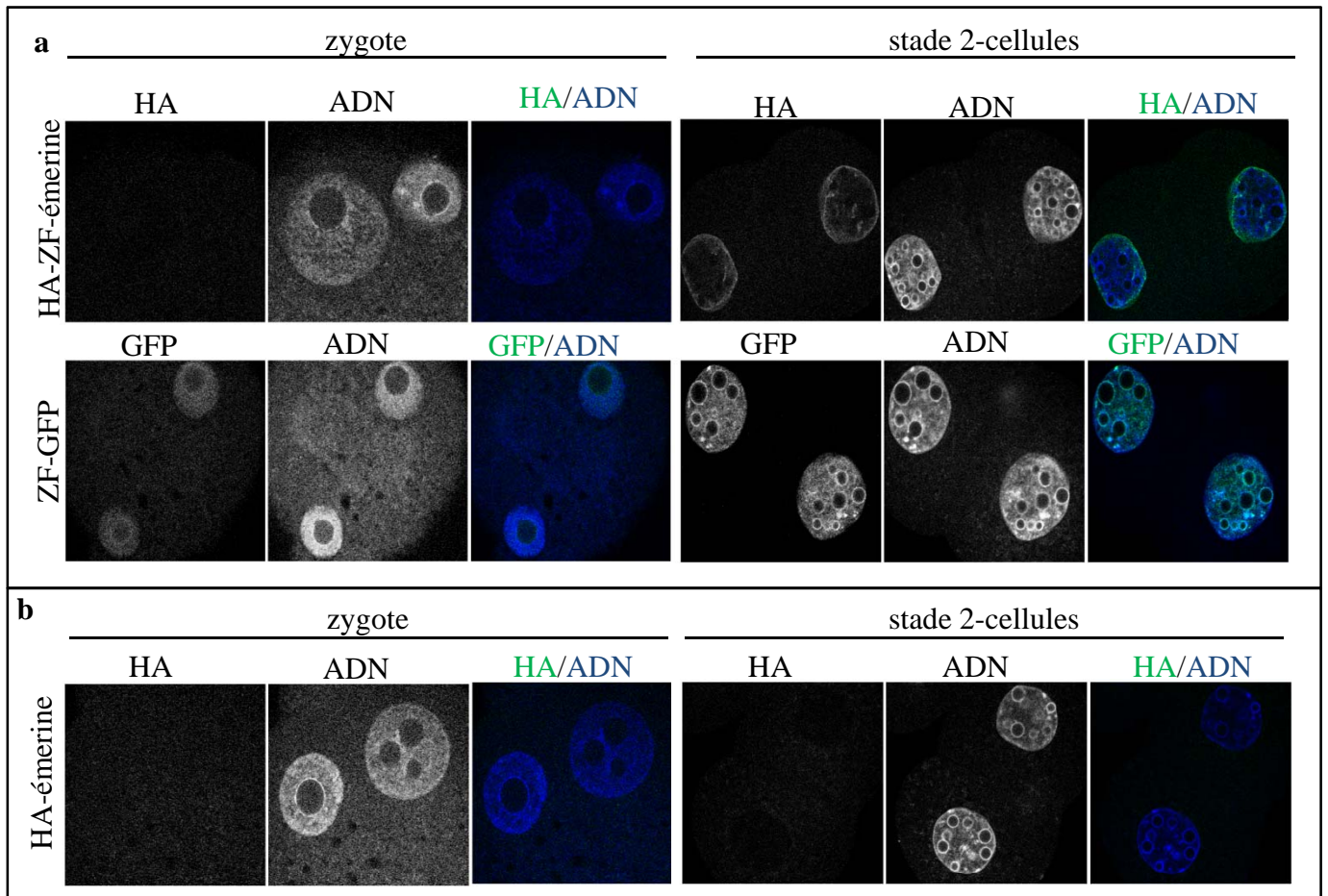


Figure 42 : HA-ZF-émerine est exprimée et incorporée dans le noyau à partir du stade 2-cellules (a), HA-émerine n'est pas détectée dans le noyau (b). Des zygotes sont injectés comme dans la figure 40a puis fixés 6h ou 30h plus tard et marqués avec l'anticorps anti-HA. L'ADN est marqué au DAPI. L'acquisition des images a été réalisée en microscopie confocale et des sections simples sont présentées ici. **a** : La présence de la protéine HA-ZF-émerine est détectée en périphérie du noyau au stade 2-cellules. La construction ZF-GFP est exprimée et localisée dans le noyau à partir du stade zygote. **b** : La construction HA-émerine n'est pas détectable dans le noyau après injection.

n'étant que faiblement détecté dans la partie centrale. La relocalisation de l'ADN à la périphérie du noyau n'est pas observée dans les cellules contrôles.

La construction HA-ZF-émerine entraîne donc une relocalisation globale de la chromatine des cellules NIH3T3 transfectées en périphérie du noyau. De plus, l'expression de cette protéine semble altérer l'incorporation de CENP-A aux centromères, ou au moins sa reconnaissance par l'anticorps.

3. La protéine HA-ZF-émerine est exprimée au stade 2-cellules et induit un faible rapprochement entre la membrane nucléaire et les centromères

Après transcription *in vitro*, les ARNm codant pour les protéines HA-ZF-émerine, HA-émerine, et ZF-GFP sont injectés dans des zygotes au moment de la fécondation, avant la formation des pronoyaux. L'expression des protéines est ensuite testée par immunocytochimie avec un anticorps anti-HA et grâce à la fluorescence de la GFP (figure 42). Au stade 2-cellules, la protéine HA-ZF-émerine se localise, comme attendu, à la périphérie du noyau, et la protéine ZF-GFP est exprimée, comme précédemment, intensément au niveau de certaines régions marquées fortement en DAPI. En revanche, la protéine HA-émerine n'est pas détectable dans le noyau. De plus, malgré des essais d'injection des ARNm très précoces, aucune construction contenant l'émerine n'est détectée dans le noyau au stade zygote.

Seules les constructions HA-ZF-émerine et ZF-GFP sont donc exprimées et localisées dans le noyau à partir du stade 2-cellules. Avec la protéine HA-ZF-émerine, nous pouvons donc essayer de déterminer l'importance de la localisation des régions péricentriques à partir du stade 2-cellules. Cette même étude au stade zygote, stade de condensation de la chromatine péricentromérique paternelle, nécessiterait la mise au point d'une autre construction.

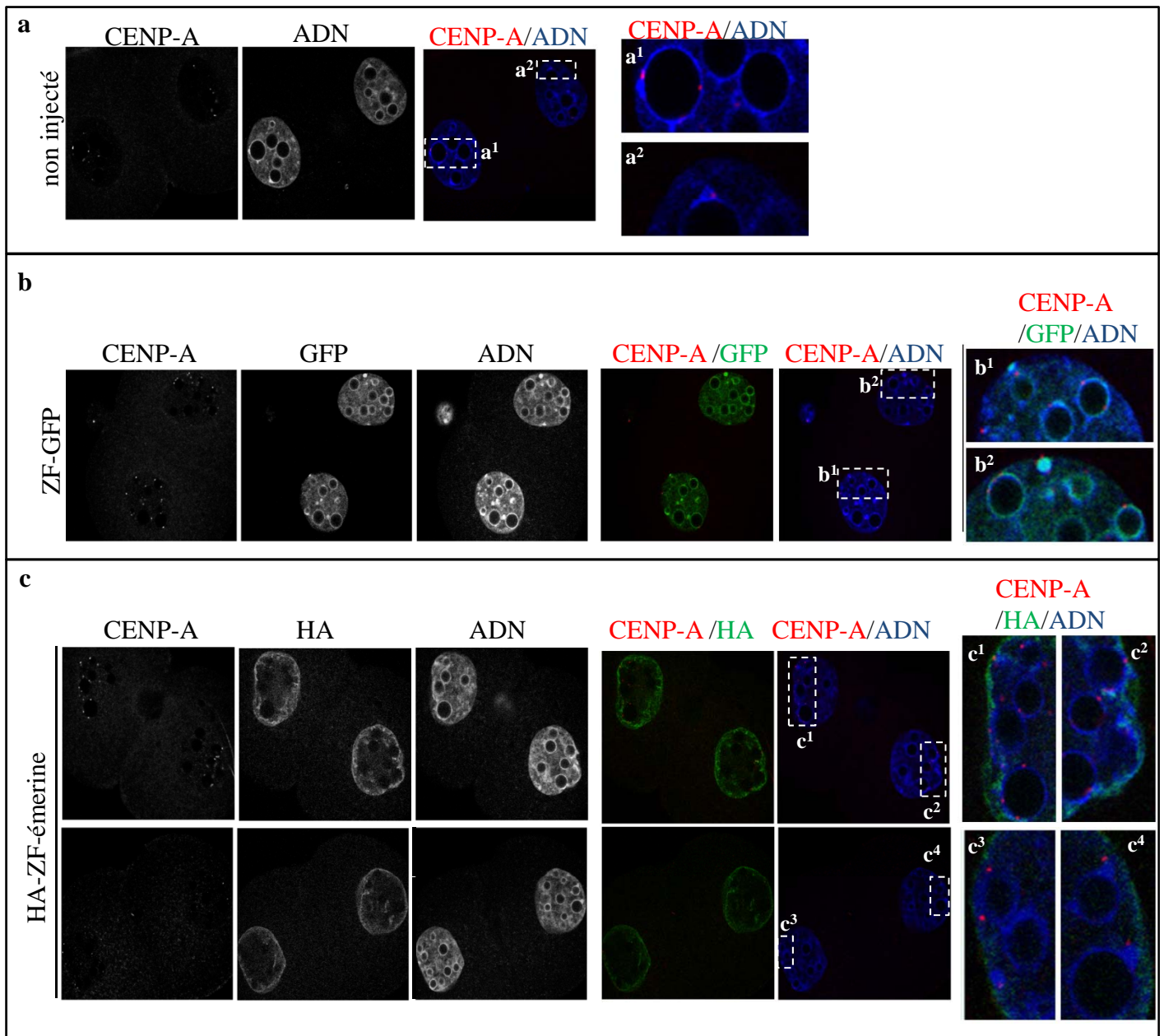


Figure 43 : Les embryons exprimant HA-ZF-émerine présentent une faible délocalisation des centromères et une modification de la forme de leurs noyaux au stade 2-cellules. Des zygotes sont injectés ou non (a) comme dans la figure 40a avec les ARNm codant pour ZF-GFP (b) ou HA-ZF-émerine (c) puis fixés 24h plus tard et marqués avec l'anticorps anti-HA et le sérum anti-CENP-A. L'ADN est marqué au DAPI. L'acquisition des images a été réalisée en microscopie confocale et des sections simples sont présentées ici. a¹, a², b¹, b², c¹, c², c³, c⁴ : grossissement des zones présentées dans les rectangles pointillés.

Pour déterminer si la protéine HA-ZF-émerine permet une relocalisation des régions péricentriques au stade 2-cellules, un marquage anti-CENP-A a été réalisé dans les embryons injectés avec les ARNm de HA-ZF-émerine et ZF-GFP (figure 43). Aucun embryon injecté ne présente de relocalisation globale de l'ADN aussi importante que celle observée dans les cellules NIH3T3-HA-ZF-émerine. Dans les embryons contrôles, les centromères se localisent majoritairement au niveau des NLB ou de régions marquées fortement en DAPI. Dans les embryons exprimant HA-ZF-émerine, alors que la plupart des centromères semble avoir une localisation normale, certains autres semblent décalés par rapport aux anneaux. Dans plusieurs embryons, nous avons observé des images où les centromères semblent entraîner la chromatine condensée des NLB vers la périphérie du noyau. De plus, dans les embryons exprimant HA-ZF-émerine, on observe parfois un repliement de la membrane nucléaire vers l'intérieur du noyau à des endroits où le marquage HA colocalise avec des centromères. Notre construction n'induit donc pas une réorganisation massive des centromères dans l'embryon au stade 2-cellules. Elle semble cependant permettre des changements de localisation des centromères par rapport aux NLB. De plus, si ce ne sont pas les centromères qui sont déplacés vers la membrane, il semble que la membrane soit rapprochée des centromères. L'étude d'un plus grand nombre d'embryons ainsi que la quantification de ces résultats seront nécessaires à l'établissement d'une conclusion fiable. De plus, un contrôle n'exprimant que l'émerine sera nécessaire pour démontrer la spécificité du phénotype.

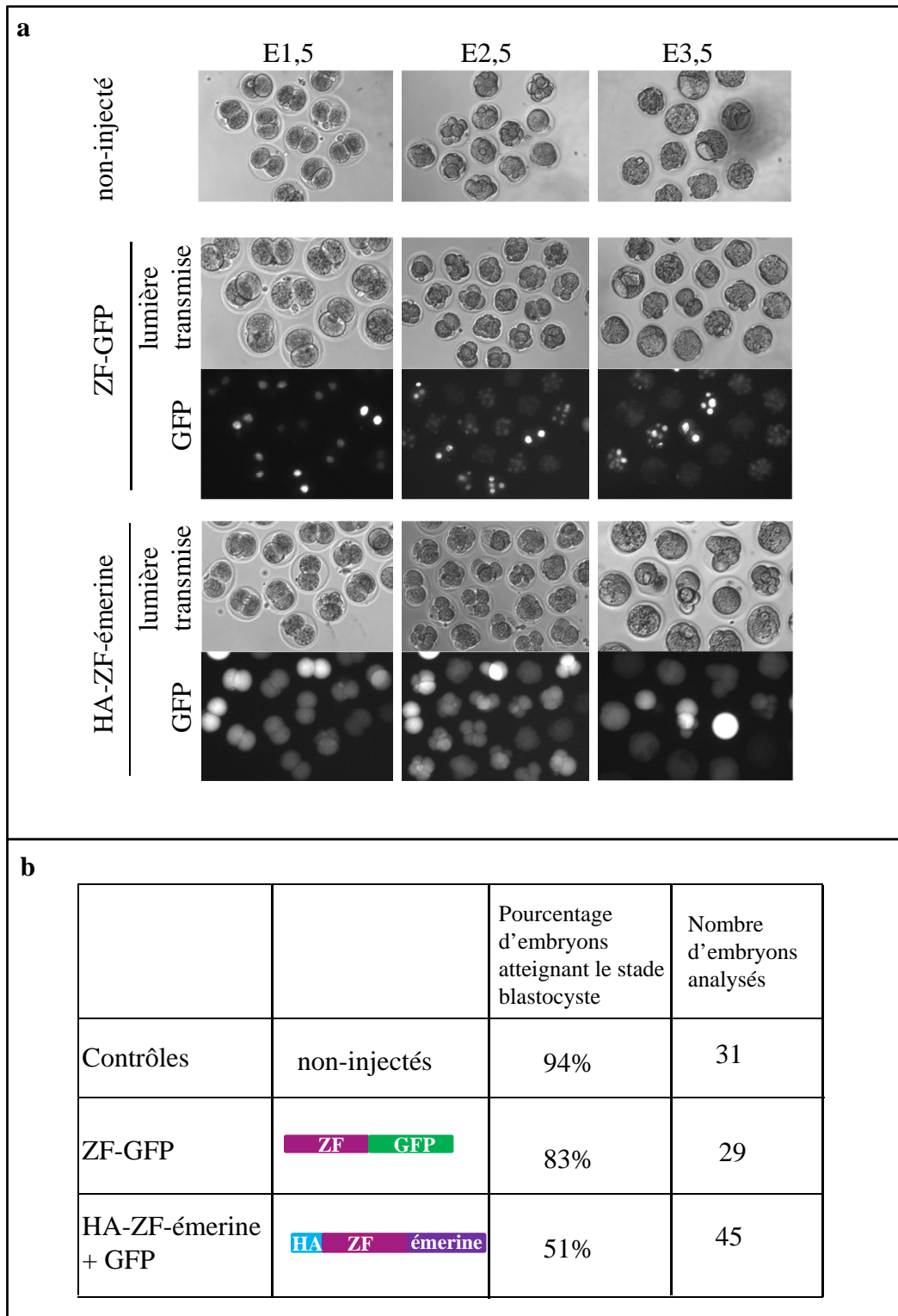


Figure 44 : L'expression de HA-ZF-émerine induit des défauts lors du développement préimplantatoire. **a** : Des zygotes ont été injectés comme montré sur la figure 40a et leur stade de développement a été analysé tous les jours. **b** : Résumé du développement des embryons exprimant ZF-GFP et HA-ZF-émerine à partir de 4 expériences indépendantes.

4. L'expression de la protéine HA-ZF-émerine au stade 2-cellules induit un arrêt du développement précoce des embryons de souris

Enfin, nous avons voulu savoir si la relocalisation des centromères par rapport à la membrane nucléaire observée dans les embryons exprimant HA-ZF-émerine induit des défauts lors du développement précoce. Nous avons donc suivi le développement des embryons injectés avec les ARNm codant pour HA-ZF-émerine et ZF-GFP (contrôles) au moment de la fécondation jusqu'au stade blastocyste (figure 44). Alors que 94% des embryons non injectés et 83% des embryons injectés avec l'ARNm codant pour ZF-GFP se développent jusqu'au stade blastocyste, seulement 51% des embryons injectés avec l'ARNm codant pour HA-ZF-émerine atteignent ce stade. Le développement de ces embryons semble s'arrêter entre les stades 3-cellules et 8-cellules. L'expression de la protéine HA-ZF-émerine est donc délétère pour environ un embryon sur deux.

La construction HA-ZF-émerine est donc exprimée dans les embryons à partir du stade 2-cellules. Son expression induit un rapprochement faible entre la membrane nucléaire et les centromères ainsi que des défauts du développement précoce des embryons. Ces résultats préliminaires, bien que prometteurs, doivent être répétés et quantifiés. Enfin, les conséquences de la relocalisation des centromères sur l'acquisition des marques chromatiniennes et sur le niveau de transcription des répétitions péricentromériques devront être recherchées pour comprendre l'arrêt du développement et la mort des embryons qui a lieu aux stades 3- à 8-cellules. Les expériences envisageables seront détaillées dans la partie discussion et perspectives.

D. Discussion et perspectives

I. Rôle de H3.3 au cours de l'embryogenèse précoce chez la souris

1. H3.3 et H3.1, deux protéines aux rôles différents ?

Notre étude a permis de montrer que H3.3 et H3.1 jouent des rôles différents au cours de l'embryogenèse précoce chez la souris. Des hypothèses concernant la spécificité de ces deux variants au cours du développement préimplantatoire, ainsi que dans la reprogrammation du génome au cours de la gamétogenèse et dans le zygote sont discutées dans la publication 3. Cette publication mentionne également l'importance des variants de H2A au cours de la reprogrammation du génome.

a. Publication 3

Point of View

Epigenetic reprogramming in mammalian reproduction

Contribution from histone variants

Angèle Santenard and Maria-Elena Torres-Padilla*

Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC); CNRS/INSERM U964; Illkirch, France

Key words: histone, H3.3, mouse embryo, histone variants, chromatin, development

Development of the mammalian embryo is, by definition, epigenetic. At the level of the nucleosome, the building block of the chromatin, changes in chromatin structure can be regulated through histone content. Apart from the canonical histones whose synthesis is restricted to S-phase, different histone variants have been identified. Histone variants can help to establish specialised chromatin regions and to regulate developmental and cell differentiation processes. While the role of histone variants has been extensively explored in differentiated cells, less is known in germ cells and embryos. Increasing lines of evidence suggest that the functions and/or properties of histone variants in embryos might be different to those in somatic cells. During reprogramming, histone variants such as H3.3 or H2A.Z are candidates to play potential important roles. We suggest that H3.3 has an important role in setting up a 'transition' signature, and provides the possibility to infer changes in chromatin architecture independent of DNA replication. This should confer flexibility during important developmental processes. The specific pathways through which H3.3 could regulate different chromatin conformations at different loci and the identification of specific proteins responsible for this deposition are an important challenge for future investigation. Lastly, the set of variants incorporated within the nucleosome can have important consequences in the regulation of epigenetic mechanisms during development.

Eukaryotic gene expression is deeply influenced by changes in chromatin structure. When these changes are independent from the DNA sequence per se and are inherited through cell division, they can constitute epigenetic signatures. Apart from DNA methylation, most often these changes occur at the level of the building block of the chromatin, the nucleosome. The nucleosome is formed by two copies of each of the core histones: H3, H4, H2A and H2B, around which a ~146 bp piece of DNA is wrapped. The so called canonical histones are incorporated into the chromatin exclusively

during S-phase, their expression is cell cycle regulated and they are transcribed from multiple genes often distributed in clusters.¹⁻⁴ However, for all histones—apart from H4—alternative histone variants exist. These histone variants are also expressed and incorporated outside of S-phase and are transcribed from single-copy genes.⁵⁻⁷ Importantly, some of these variants have acquired specialised functions in developmental processes such as fertilization and X-chromosome inactivation. Here we discuss and present some hypotheses on the contribution of two types of variants: those of H3 and of H2A as well as their potential role in directing epigenetic information in mammalian development.

Histone H3 Variants: Small Differences Can be Big

Most studies concerning histone variants have focused on variants of H3. In mammals, there are two canonical histones of H3: H3.1 and H3.2 as well as three other H3 variants: H3.3, a testis specific H3 variant (H3t) and centromeric H3 variants (CenH3s). We will focus on the three variants H3.1, H3.2 and H3.3, as they have been studied throughout key developmental processes and in cell differentiation systems, including mouse embryos and ES cells. Only one amino acid differs between H3.1 and H3.2, located at position 96 (a Cys-Ser substitution) and five amino acids are different between H3.1 and H3.3 at positions 31, 87, 89, 90 and 96 (Ala-Ser, Ser-Ala, Val-Ile, Met-Gly, Cys-Ser substitutions respectively).⁸ The differences in amino acid sequence amongst them, albeit surprisingly small, lead to different pathways of incorporation into the chromatin for H3.1/H3.2 on the one hand and H3.3 on the other hand.^{5,9,10} In somatic cells, it has been well established that the canonical variants H3.1 and H3.2 are synthesised only during S-phase and then deposited into the chromatin through a Replication Coupled (RC) pathway.^{9,11,12} Recently, it was further shown that these two variants can also be incorporated in the context of DNA repair.¹³ More generally, the canonical variants of H3 are deposited into chromatin concomitant and/or in a pathway that involves DNA synthesis. On the contrary, H3.3 is synthesised throughout the cell cycle and is incorporated into chromatin by a Replication Independent (RI) pathway, hence its name of 'replacement variant.'

As a general rule, H3 is deposited as H3-H4 tetramers.¹⁰⁻¹² RI and RC deposition of H3 into the chromatin are carried out by different histone chaperones. In HeLa cells, the RC deposition complex for H3.1/2 contains the histone chaperone ASF-1 and the entire CAF-1 complex, while the RI deposition complex of H3.3

*Correspondence to: Maria-Elena Torres-Padilla; Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC); CNRS/INSERM U964; BP 10142; Illkirch Cedex, CU de Strasbourg F-67404 France; Tel.: +33.0.3.88.65.33.60; Fax: +33.0.3.88.65.32.01; Email: metp@igbmc.fr

Submitted: 09/17/08; Accepted: 01/14/09

Previously published online as an *Epigenetics* E-publication:
<http://www.landesbioscience.com/journals/epigenetics/article/7838>

contains ASF-1, the smallest subunit of the CAF-1 complex (p48) and a specific histone chaperone HIRA.^{10,14} CAF-1 can interact with PCNA and promote subsequent deposition of H3.1/2-H4 heterodimers to sites of DNA synthesis, allowing for the assembly of new nucleosomes at the replication fork.¹⁵ The absence of deposition of H3.1/2 outside of the S-phase cannot be explained exclusively by the fact that these variants are not available during the entire cell cycle. Indeed, in *Drosophila*, the pathway of assembly for H3 canonical variants is specified by the residues located in positions 87, 89 and 90, which differ among H3.1/H3.2 and H3.3 variants.^{5,9}

Once incorporated into nucleosomes, H3.1, H3.2 and H3.3 will have different effects on chromatin-mediated processes. Indeed, studies in mammalian cells show that upon transcriptional induction, H3.3 is recruited at highly expressed loci,¹⁶⁻¹⁸ while H3.1 and H3.2 are potentially linked with both active and inactive genes. Despite their high degree of similarity, mass-spectrometric analysis of histone H3 variants has revealed that the post-translational modifications (PTMs) occurring on the three variants differ: H3.3 is mostly associated with marks of transcriptionally active chromatin, H3.2 is enriched in repressive marks and H3.1 contains PTMs of active and inactive chromatin.¹⁹⁻²¹ It is important to note that, to date, this is the only evidence that H3.1 and H3.2 could have different functions. It has also been shown that when H3.3 is present in the chromatin, the surrounding nucleosomes, even if containing H3.1, will have predominantly marks of active transcription.²² So, it appears that, in somatic cells, H3.1 and H3.2 are deposited rather broadly on the genome during replication and could play roles both in active or inactive genomic regions, most likely due to their content of different PTMs. On the contrary, deposition of H3.3 could be more regulated, and perhaps targeted to specific genomic regions, as it is the case of sites of active transcription.

There are further considerations for H3.3 deposition outside of S-phase. It has been suggested that the deposition of H3.3 at transcribed genes is the result of a neutral replacement, in a replication-independent manner, of nucleosomes that have been evicted by the transcription machinery.²³ However, other studies suggest that H3.3 by itself, even in the absence of post-translational modifications, could act as a marker for active transcription.¹⁶ These conclusions are based on findings that show conservation of H3.3 on the chromatin within a highly active transgene through mitosis, which is expected to promote gene activation immediately after exit from mitosis.¹⁶ However, the pathways responsible for controlling specific deposition of H3.3 to active genes and the pathways of recognition of H3.3 by the transcription machinery still have to be elucidated. An important feature is that nucleosomes containing H3.3 appear to be less stable than those containing H3.1 or H3.2,²⁴ and could therefore be more easily removed, thereby allowing transcription and/or local chromatin remodelling.

H3 Variants in Reprogramming During Mammalian Development

The roles played by the different variants of H3 in germ cells and during preimplantation development are less well described compared to what is known in somatic cells, most likely due to the technical limitations of both systems. The current lack of antibodies specific for the different variants does not facilitate this task. However, recent works have begun to highlight the fact that H3 variants could have

roles in sexual reproduction that are different from their roles in somatic cells. Whether the three H3 variants display a similar pattern of PTMs in embryos and in somatic cells is not known, but most likely the repertoire of modifications in each of the three variants differ in the two situations. For example, we have previously reported a dicotomy between H3.3 enrichment in the male pronucleus and a lack of the active marks H3K4me2 and H3K4me3 following fertilization.²⁵

In germ cells and embryos, two major events of large-scale replacement of chromatin marks (DNA methylation, histone PTMs, histone content) occur. In mouse primordial germ cells (PGCs), parental imprints are erased and re-established in a sex-specific manner. Monoallelic parent-of-origin specific expression of imprinted genes is regulated by epigenetic modifications. The reprogramming of PGCs occurs between embryonic day (E) 11.5 and E12.5. During this reprogramming event, it has been inferred, because of the transient presence of HIRA in the nucleus of the PGCs between E11.5 and E12.5, that H3.3 could play a role in this phenomenon.²⁶ Another event involving huge reprogramming occurs immediately following fertilization when two highly differentiated cells, the oocyte and the sperm, give rise to a totipotent cell, the zygote. In the sperm most histones are replaced by protamines and other nuclear basic proteins during spermiogenesis.²⁷ After the sperm entry into the oocyte cytoplasm, protamines are removed and replaced by nucleosomal histones. It has been shown that H3.3 accumulates preferentially in newly assembled chromatin on the paternal genome.^{25,27,28} This constitutes the only event of genome wide deposition of H3.3 in the life of an organism. As a consequence of protamine exchange and the global acquisition of newly incorporated histones on the paternal genome, the paternal chromatin must acquire most epigenetic signatures, while the maternal genome has already epigenetic marks, which were acquired mainly during folliculogenesis.²⁹

Histone H3 Variants: Time-Window for a Transition

We can suggest different hypotheses concerning a potential role for H3.3 during reprogramming, perhaps through facilitating this process at the nucleosomal level. First, because in both PGCs and in the zygote reprogramming takes place—at least initially— independently of DNA replication, H3.3 deposition in PGCs and in the paternal chromatin might occur only because H3.1/2 is not available. This is however unlikely, as we have performed microinjections of H3.1 in zygotes before pronuclear formation and have not observed incorporation of H3.1 before the S-phase (our unpublished data). Thus, in the mouse zygote, as it is in *Drosophila*,³⁰ the non-incorporation of H3.1 might not only be due to the absence of the H3.1 protein, but rather to the absence of functional H3.1 deposition machinery outside of the S-phase. If we consider that H3.3 is deposited on the genome only because H3.1 cannot be deposited, then one could predict that H3.3 will play no particular role by itself but will just have a “neutral” or “default” filling-in role. Broadly speaking, this should be true if the deposition of H3 variants is genome-wide and not ‘loci-specifically’ regulated. However, we have recently found that in the mouse zygote, H3.3 and H3.1 show different distributions along specific genomic regions in the two pronuclei (Santenard A, et al. submitted). This new data suggests that H3.3 could have specific roles in the establishment of new epigenetic signatures in the male pronucleus after fertilization. If we consider that H3.3

can have both a “neutral” effect at some loci and specific effects on reprogramming of other loci, this would predict that H3.3 has to be recognised by different complexes at these different specific locations. If the genomic regions where specific variants become incorporated are important, the relevant question is how is this targeted deposition regulated? Is this achieved through the same HIRA complex? or Would there be a specific deposition complex, maybe different from the HIRA complex, in specialised genomic regions? And if so, would this putative complex recognise transcription factor sites or would it work based on RNA-interactions and/or bridging proteins? Of note, some histone chaperones can physically interact with specific transcription factors: HIRA has been shown to interact with Pax3 and Pax7, and NAP1 with E2F.^{31,32} This suggests that deposition at given target genes can be potentially achieved through these interactions.

There is some evidence of H3.3 being enriched in specific genomic locations in developmentally regulated processes. During spermatogenesis, the X and Y chromosomes partially synapse to form the sex body in a process called Meiotic Sex Chromosome Inactivation (MSCI). During this process X and Y chromosomes become transcriptionally silenced.³³ In mouse and human, immediately after the initiation of MSCI at the early pachytene stage of meiotic prophase, nucleosomes present at the sex chromosomes are removed and replaced by new ones. Because this incorporation of new nucleosomes occurs independently of DNA synthesis, it is exclusively H3.3 that is incorporated in a chromosome-wide manner.³⁴ These observations have interesting consequences: firstly, that deposition of H3.3 can be specifically regulated and secondly, that H3.3 can also be associated with silenced, heterochromatic genomic regions or, at least, that this variant does not interfere with the heterochromatic state of the sex chromosomes. Whether H3.3 deposition is potentially involved in silencing is unclear. During MSCI, the incorporation of H3.3 and presumably newly assembled nucleosomes into chromatin further coincide with a lack of detection of almost all of the histone PTMs.³⁴ This also occurs during reprogramming of PGCs, where there is a loss of almost all PTMs between E11.5 and E12.5.^{26,35} At this precise time, the replacement variant of H2A, H2A.Z, is also detected on the chromatin.²⁶ Thus, although it is currently impossible to show whether an individual nucleosome contains both H3.3 and H2A.Z variants, there is a time window at ~E11.5 during which H2A.Z and H3.3 are presumably both present in the nucleosomes. Of note, nucleosomes containing H2A.Z and H3.3 are less stable than those containing the canonical H2A and H3.1.²⁴ Because H2A.Z-H3.3-containing nucleosomes are less stable, the cell will thus need less energy to remove these. Thus, replacing the existing nucleosomes with their PTMs by less stable ones could be an efficient way to quickly remove old epigenetic marks and rapidly allow a huge reprogramming of the genome. Non-stable nucleosomes could also be a signal for the cell to ‘realise’ that something is ongoing at these loci and that some proteins have to come to replace these non-stable octamers by another ones that should thereafter remain as a long-term message. For reprogramming, H3.3 could then be a mark of transient state while H3.1/2 a long-term mark. Thus, another hypothesis concerning the role of H3.3 incorporation during reprogramming relates to the flexibility imparted by its potential ‘transient’ or ‘less stable’ state.

Histone H2A Variants: Nucleosomal Partners for H3.3?

It is important to ask which are the nucleosomal partners of H3.3 during these developmental processes, in particular because it has been suggested that the identity of the partners within the nucleosome might give rise to different effects on the chromatin.²⁴ Variants of H2A are also involved in wide-scale chromatin remodelling events and have acquired specialised functions during development. Among these, phosphorylation of the H2A variant H2A.X was implicated in the initiation of the MSCI.³⁶ H2A.X is the most abundant H2A variant in *Xenopus* eggs, and the ability to remodel the sperm nucleus to form a paternal pronucleus after fertilization is directly associated with its phosphorylation status.³⁷ Moreover, unusually high levels of phosphorylated H2A.X are present throughout mouse preimplantation development (Ziegler-Birling et al., in press), opening the possibility for a role of this variant in chromatin assembly and remodelling during these early stages.⁵⁸ MacroH2A is an H2A variant possessing a long C-terminal domain and is conserved in vertebrates.³⁸ MacroH2A is enriched in the inactive X chromosome and becomes enriched in X and Y chromosomes during MSCI.^{39,40} It is also peculiar because it shows a very tight temporal specific pattern of expression in mouse development. MacroH2A is present in the chromatin of developing and mature oocytes but it is immediately lost from the maternal chromatin in the zygote following fertilization and reappears only after the 8-cell stage.⁴¹ Because macroH2A inhibits chromatin remodelling—and transcription⁴²—it is possible that its rapid disappearance following fertilization is necessary to render the zygotic chromatin permissive for remodelling and epigenetic reprogramming.

We have mentioned above a third variant of H2A, H2A.Z, which is associated with both euchromatin and facultative heterochromatin.⁴³ H2A.Z has been implicated in apparently opposing functions such as prevention of gene silencing in *S. cerevisiae* and the establishment of heterochromatin in *Drosophila*.^{44,45} Like H3.3, H2A.Z is not uniformly distributed throughout the genome. H2A.Z is enriched at promoters of developmentally regulated genes in chicken erythroid cells and in mouse ES cells.^{46,47} Interestingly, H2A.Z localization is different in pluripotent versus differentiated cells.⁴³ In mouse ES cells, H2A.Z is required for differentiation and is functionally linked to Polycomb-induced gene silencing.⁴⁷ These findings are in full agreement with the early embryonic lethality resulting from loss of H2A.Z, as mouse blastocysts lacking H2A.Z fail to survive beyond implantation.⁴⁸ As we have already stated, a direct relationship between H2A.Z-H3.3-containing nucleosomes and local changes in chromatin structure have been documented. There are some interesting links between H3.3 and H2A.Z (H2Av in *Drosophila*). Nucleosome eviction during transcription can result in loss of H3.3/H2A.Z nucleosomes.⁴⁹ Given the low stability conferred by the presence of both H3.3 and H2A.Z to the nucleosome compared to H2A- or H3-containing nucleosomes, it has been suggested that H3.3 possesses a specific regulatory role when coupled with H2A.Z.²⁴ Another level of regulation by H2A variants arises from the interrelation of H2A.Z with DNA methylation. The presence of H2A.Z and DNA-methylation seem to be mutually exclusive in plants, and most importantly, loss of H2A.Z deposition following SWR1 complex knock-down leads to genomic hypermethylation. This led to the suggestion that H2A.Z protects genes from DNA

methylation.⁵⁰ Thus, it is clear that histone variants have an important role in establishing functional chromatin domains *in vivo*.

Because of all the above, we can suggest that a different complement of variants within the nucleosome core particle and the localised distribution of such variants are key elements to regulate epigenetic phenomena during development. Insight into the mechanisms governing histone variant deposition will therefore be central to our understanding of mammalian development. In particular the role for chaperones or deposition complexes with specificity towards H3.3 such as Chd1,⁵¹ and of Chtz1/SWR1 for H2A.Z⁵² in early development and whether there is a functional relation between the two types of deposition complexes.

Can Histone Variants Leave a Mark on the Transmission of Epigenetic Information?

In human sperm, around 15% of the DNA remains wrapped around histones.⁵³ The same applies to the mouse sperm, albeit with a lower proportion (~4%). Concerning H3 variants, H3.1, H3.2 and H3.3 are all three present in human sperm.⁵⁴ Nothing is known about the specific differences of incorporation or the potential role that could be played by the variants in sperm. However, it has been shown that nucleosomal histones do not localise to random genomic regions.⁵⁵ It was demonstrated for the β -globin domain that loci bound by histones within this domain are preferentially those, which will be transcribed during early development and that silenced genes will be preferentially bound by protamines.⁵⁶ More recently, it was shown that sperm derived histones in human—at least H3.1 and H3.2—will contribute to the formation of nucleosomes in the paternal genome. If the observations of the β -globin locus can be generalised, we can think that histones are retained at genes that will be the first to be transcribed during zygotic genome activation. This is probably because the chromatin is less condensed with histones than with protamines, and as a consequence, the genes will be more easily, and more quickly transcribed after fertilization. Another role for these inherited histones could be to mark imprinted genes. After fertilization, when nearly the entire paternal genome is wrapped around H3.3-containing nucleosomes, remaining H3.1/2-containing nucleosomes could constitute a mark recognised by the cell. This has also been suggested in plants, where the paternal variant of H3 is carried by the sperm and will be maintained in the paternal genome in the endosperm for several cell divisions. The authors suggest that this phenomenon contributes to the parental genomic imprinting that occurs in the endosperm.⁵⁷ While all these hypotheses appear attractive, work is still to be done to test them directly in the mouse embryo.

Conclusion

While deposition pathways and stability features of the different H3 variants are, up to now, assumed to be the same between somatic cells, germ cells and embryos, different roles for these variants in different cell types start to emerge. First, in somatic cells, H3.3 was described to be specifically linked with active transcription, while in germ cells and embryos, the replacement variant could be present as well in heterochromatin. Second, in somatic cells, H3.1/2 are known to be the most abundant variants, but in germ cells and embryos, because of a large scale replacement of nucleosomes that occurs

independently of replication, H3.3 will probably be more broadly present and will play an important role in reprogramming. The fact that H3.3 participates in the formation of less stable nucleosomes compared to H3.1 could be an advantage in germ cells and embryos, not only to allow transcription as it is in somatic cells, but also to allow quick and efficient removal of existing epigenetic marks, acquisition of new ones and a more open and permissive chromatin conformation. Thus, H3.1/2 would rather play a role in the propagation of a stable epigenetic state, a mark for example, through mitosis. We can also suggest that, in an H3.3 background in the paternal genome in zygotes, the presence of H3.1/2 could act as ‘marks’ by themselves, as H3.3 does in specific (e.g., transcribed) loci in a broad H3.1/2 background in somatic cells. If we assume that the same protein has different roles in a same organism, this implies that there should be different complexes that will interact with each of the variants depending on its localization or on its interacting proteins. Up to now, only one complex containing the histone chaperone HIRA has been described to interact with and deposit H3.3 but further investigations are needed to understand the pathways through which H3.3 could regulate different chromatin conformations at different loci. Thus, we suggest that H3.3 has an important role in setting up a ‘transition’ signature in germ cells and embryos, and provides the possibility to infer changes on the chromatin independent of one of the most basic DNA process, DNA replication. This should confer flexibility during important developmental processes to support changes mediated via alterations in chromatin composition.

Acknowledgements

Work in M.-E.T.-P. laboratory is supported by the PNRRE and AVENIR programs from the INSERM.

References

- Zweidler A. Core histone variants of the mouse: primary structure and differential expression. Wiley 1984.
- Hereford LM, Osley MA, Ludwig TRn, McLaughlin CS. Cell cycle regulation of yeast histone mRNA. *Cell* 1981; 24:367-75.
- Sierra F, Lichter A, Marashi F, Rickles R, Van Dyke T, Clark S, et al. Organization of human histone genes. *Proc Natl Acad Sci USA* 1982; 79:1795-9.
- Stein G, Park W, Thrall C, Mans R, Stein J. Regulation of cell cycle-specific transcription of histone genes from chromatin by non-histone chromosomal proteins. *Nature* 1975; 257:764-7.
- Ahmad K, Henikoff S. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* 2002; 9:1191-200.
- Wu RS, Tsai S, Bonner WM. Patterns of histone variant synthesis can distinguish G₀ from G₁ cells. *Cell* 1982; 31.
- Kamakaka RT, Biggins S. Histone variants: deviants? *Genes Dev* 2005; 19:295-310.
- Franklin SG, Zweidler A. Non-allelic variants of histones 2a, 2b and 3 in mammals. *Nature* 1977; 266:273-5.
- Ahmad K, Henikoff S. Histone H3 variants specify modes of chromatin assembly. *Proc Natl Acad Sci USA* 2002; 99:16477-84.
- Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* 2004; 116:51-61.
- Akey CW, Luger K. Histone chaperones and nucleosome assembly. *Curr Opin Struct Biol* 2003; 13:6-14.
- Polo SE, Almouzni G. Chromatin assembly: a basic recipe with various flavours. *Curr Opin Genet Dev* 2006; 16:104-11.
- Polo SE, Roche D, Almouzni G. New histone incorporation marks sites of UV repair in human cells. *Cell* 2006; 127:481-93.
- Ray-Gallet D, Quivy JP, Scamps C, Martini EM, Lipinski M, Almouzni G. HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. *Mol Cell* 2002; 9:1091-100.
- Shibahara K, Stillman B. Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. *Cell* 1999; 96:575-85.
- Chow CM, Georgiou A, Szutorisz H, Maia e Silva A, Pombo A, Barahona I, et al. Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division. *EMBO Rep* 2005; 6:354-60.

17. Daury L, Chailleux C, Bonvallet J, Trouche D. Histone H3.3 deposition at E2F-regulated genes is linked to transcription. *EMBO Rep* 2006; 7:66-71.
18. Janicki SM, Tsukamoto T, Salghetti SE, Tansey WP, Sachidanandam R, Prasanth KV, et al. From Silencing to Gene Expression: Real-Time Analysis in Single Cells. *Cell* 2004; 116:683-98.
19. Garcia BA, Thomas CE, Kelleher NL, Mizzen CA. Tissue-specific expression and post-translational modification of histone H3 variants. *J Proteom Res* 2008; 7:4225-36.
20. Hake SB, Garcia BA, Duncan EM, Kauer M, Dellaire G, Shabanowitz J, et al. Expression patterns and post-translational modifications associated with mammalian histone H3 variants. *J Biol Chem* 2006; 281:559-68.
21. McKittrick E, Gafken PR, Ahmad K, Henikoff S. Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc Natl Acad Sci USA* 2004; 101:1525-30.
22. Loyola A, Bonaldi T, Roche D, Imhof A, Almouzni G. PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Mol Cell* 2006; 24:309-16.
23. Li B, Carey M, Workman JL. The role of chromatin during transcription. *Cell* 2007; 128:707-19.
24. Jin C, Felsenfeld G. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev* 2007; 21:1519-29.
25. Torres-Padilla ME, Bannister AJ, Hurd PJ, Kouzarides T, Zernicka-Goetz M. Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos. *Int J Dev Biol* 2006; 50:455-61.
26. Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Cesari F, et al. Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* 2008; 452:877-81.
27. Wouters-Tyrou D, Martinage A, Chevaillier P, Sautiere P. Nuclear basic proteins in spermiogenesis. *Biochimie* 1998; 80:117-28.
28. van der Heijden GW, Dieker JW, Derijck AA, Muller S, Berden JH, Braat DD, et al. Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech Dev* 2005; 122:1008-22.
29. Kageyama S, Liu H, Kaneko N, Ooga M, Nagata M, Aoki F. Alterations in epigenetic modifications during oocyte growth in mice. *Reproduction* 2007; 133:85-94.
30. Loppin B, Bonnefoy E, Anselme C, Laurencon A, Karr TL, Couble P. The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus. *Nature* 2005; 437:1386-90.
31. Magnaghi P, Roberts C, Lorain S, Lipinski M, Scambler PJ. HIRA, a mammalian homologue of *Saccharomyces cerevisiae* transcriptional co-repressors, interacts with Pax3. *Nature Genetics* 1998; 20:74-7.
32. Rehtanz M, Schmidt HM, Warthorst U, Steger G. Direct interaction between nucleosome assembly protein 1 and the papillomavirus E2 proteins involved in activation of transcription. *Mol Cell Biol* 2004; 24:2153-68.
33. Turner BM. Meiotic sex chromosome inactivation. *Development* 2007; 134:1823-31.
34. van der Heijden GW, Derijck AA, Postfai E, Giele M, Pelczar P, Ramos L, et al. Chromosome-wide nucleosome replacement and H3.3 incorporation during mammalian meiotic sex chromosome inactivation. *Nature Genetics* 2007; 39:251-8.
35. Hajkova P, Erhardt S, Lane N, Haaf T, El-Maari O, Reik W, et al. Epigenetic reprogramming in mouse primordial germ cells. *Mech Dev* 2002; 117:15-23.
36. Fernandez-Capetillo O, Mahadevaiah SK, Celeste A, Romanienko PJ, Camerini-Otero RD, Bonner RF, et al. H2AX is required for chromatin remodelling and inactivation of sex chromosomes in male mouse meiosis. *Dev Cell* 2003; 4:497-508.
37. Dimitrov S, Dasso MC, Wolffe AP. Remodeling sperm chromatin in *Xenopus laevis* egg extracts: the role of core histone phosphorylation and linker histone B4 in chromatin assembly. *J Cell Biol* 1994; 126:591-601.
38. Pehrson JR, Fried VA. MacroH2A, a core histone containing a large nonhistone region. *Science* 1992; 257:1398-400.
39. Costanzi C, Stein P, Worrall DM, Schultz RM, Pehrson JR. Histone macroH2A1 is concentrated in the inactive X chromosome of female preimplantation mouse embryos. *Development* 2000; 127:2283-9.
40. Hoyer-Fender S, Costanzi C, Pehrson JR. Histone macroH2A1.2 is concentrated in the XY-body by the early pachytene stage of spermatogenesis. *Exp Cell Res* 2000; 258:254-60.
41. Chang CC, Ma Y, Jacobs S, Tian XC, Yang X, Rasmussen TP. A maternal store of macroH2A is removed from pronuclei prior to onset of somatic macroH2A expression in preimplantation embryos. *Dev Biol* 2005; 278:367-80.
42. Angelov D, Molla A, Perche PY, Hans F, Coté J, Khochbin S, et al. The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. *Mol Cell* 2003; 11:1033-41.
43. Sarcinella E, Zuzarte PC, Lau PN, Draker R, Cheung P. Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin. *Mol Cell Biol* 2007; 27:6457-68.
44. Meneghini MD, Wu M, Madhani HD. Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. *Cell* 2003; 112:725-36.
45. Swaminathan J, Baxter EM, Corces VG. The role of histone H2Av variant replacement and histone H4 acetylation in the establishment of *Drosophila* heterochromatin. *Genes Dev* 2005; 19:65-76.
46. Bruce K, Myers F, Mantouvalou E, Lefevre P, Greaves I, Bonifer C, et al. The replacement histone H2A.Z in a hyperacetylated form is a feature of active genes in the chicken. *Nucleic Acids Res* 2005; 33:5633-9.
47. Creighton MP, Markoulaki S, Levine SS, Hanna J, Lodato MA, Sha K, et al. H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment. *Cell* 2008; 135:649-61.
48. Faast R, Thonglairoam V, Schulz TC, Beall J, Wells JR, Taylor H, et al. Histone variant H2A.Z is required for early mammalian development. *Curr Biol* 2001; 11:1183-7.
49. Henikoff S, Henikoff JG, Sakai A, Loeb GB, Ahmad K. Genome-wide profiling of salt fractions maps physical properties of chromatin. *Genome Res* 2008; Epub ahead of print.
50. Zilberman D, Coleman-Derr D, Ballinger T, Hennikoff S. Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* 2008; 456:125-9.
51. Konev AY, Tribus M, Park SY, Podhraski V, Lim CY, Emelyanov AV, et al. CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin in vivo. *Science* 2007; 317:1087-90.
52. Zhang H, Roberts DN, Cairns BR. Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell* 2005; 123:219-31.
53. Tanphaichitr N, Sobhon P, Taluppeth N, Chalermisaraichai P. Basic nuclear proteins in testicular cells and ejaculated spermatozoa in man. *Exp Cell Res* 1978; 117:347-56.
54. Gatewood JM, Cook GR, Balhorn R, Schmid CW, Bradbury EM. Isolation of four core histones from human sperm chromatin representing a minor subset of somatic histones. *J Biol Chem* 1990; 265:20662-6.
55. Gatewood JM, Cook GR, Balhorn R, Bradbury EM, Schmid CW. Sequence-specific packaging of DNA in human sperm chromatin. *Science* 1987; 236:962-4.
56. Gardiner-Garden M, Ballesteros M, Gordon M, Tam PP. Histone- and protamine-DNA association: conservation of different patterns within the beta-globin domain in human sperm. *Mol Cell Biol* 1998; 18:3350-6.
57. Ingouff M, Hamamura Y, Gourgues M, Higashiyama T, Berger F. Distinct dynamics of HISTONE3 variants between the two fertilization products in plants. *Curr Biol* 2007; 17:1032-7.
58. Ziegler-Birling C, Helmrich A, Tora L, Torres-Padilla ME. Distribution of p53 binding protein 1 (53BP1) and phosphorylated H2A.X during mouse preimplantation development in the absence of DNA damage. *Int J Dev Biol* 2009; In Press.

b. Perspectives

Pour que les variants de H3 jouent des rôles différents dans l'embryon précoce, des protéines reconnaissant spécifiquement ces différents variants et leurs modifications post-traductionnelles doivent exister. Alors que plusieurs études ont permis de mettre en évidence des protéines se liant spécifiquement à l'un ou l'autre des variants dans les cellules somatiques (Tagami et al., 2004) (Lewis et al., 2010), aucune expérience d'immunoprécipitation de complexes contenant H3.3 ou H3.1 n'a été réalisée dans des embryons préimplantatoires certainement par manque de matériel.

Dans le but de mettre en évidence des partenaires spécifiques des différents variants de H3 dans les zygotes de souris, nous envisageons de réaliser une expérience de double hybride où l'appât serait une partie des protéines H3.3 ou H3.1 et où la banque d'ADNc utilisée pour générer les proies serait une banque spécifique des zygotes. Tout d'abord, aucune banque d'ADNc spécifique du zygote n'existe à l'heure actuelle. Pour générer ce genre de banques avec des protocoles classiques, un minimum de 50µg d'ARN total ou de 2µg d'ARN polyadénylé est requis (Crative Bioloabs, Dualsystems Biotech). Si l'on considère qu'un zygote contient 0,83pg d'ARN polyadénylé (tableau 1) (Piko and Clegg, 1982), $2,5 \times 10^6$ zygotes sont nécessaires à la réalisation de la banque. Après discussion avec les laboratoires réalisant ce type de librairies, nous avons convenu que la synthèse d'une banque d'ADNc pourrait être envisagée avec 200ng d'ARN total. La récolte d'embryons à E0,5 est en cours pour former cette librairie. Ensuite, la surexpression de protéines histones dans les levures au cours de l'expérience de double hybride fait apparaître un deuxième problème majeur. En effet, l'expression d'une trop grande quantité de protéines histones dans des cellules entraîne une très forte compaction de leur ADN et l'arrêt du cycle cellulaire. De plus, il est possible que de nombreux faux-positifs soient générés par cette liaison aspécifique des histones à l'ADN. Pour éviter ce problème, nous envisageons de n'exprimer qu'une forme tronquée des histones, où la boucle L2 et l'hélice $\alpha 3$ ne sont pas présentes. Cependant, il est fortement probable qu'en l'absence de l'extrémité C-terminale, les appâts correspondant à H3.3 et H3.1 ne soient pas repliés correctement, et donc ne soient pas reconnus par certains de leurs partenaires cellulaires. Malgré tous ces inconvénients, la réalisation de cette expérience est sérieusement envisagée au laboratoire.

2. H3.3, un variant important pour la formation de l'hétérochromatine ?

Alors que la plupart des études de H3.3 dans les cellules somatiques de différents organismes associe la présence de cette protéine à la formation d'une chromatine ouverte, transcrite (Chow et al., 2005) (Mito et al., 2005), les résultats de la publication 2 montrent que H3.3, et plus précisément des modifications sur sa lysine 27, est nécessaire à la condensation de l'hétérochromatine péricentrique dans le zygote. Il a également été montré que l'incorporation de H3.3 aux télomères est nécessaire au maintien de leur intégrité dans des cellules ES de souris (Wong et al., 2009). Ces résultats suggèrent donc que H3.3 peut également être associée à des régions hétérochromatiques, voire responsable de leur formation à certaines étapes du développement, ou à certains stades de différenciation.

Plusieurs hypothèses peuvent être émises pour expliquer le rôle de H3.3 au niveau de l'hétérochromatine. Tout d'abord, en lien avec l'importance démontrée de H3.3 au sein de la chromatine active, H3.3 pourrait, au niveau des télomères et des centromères, permettre la transcription de certains ARN régulateurs nécessaires à l'hétérochromatinisation de ces régions. Aux régions péricentriques, on peut envisager que l'acétylation ou la monométhylation de H3.3K27 permette la transcription des *major satellites* (Kouzarides, 2007) (Barski et al., 2007). Cependant, d'après nos immunomarquages contre H3K27ac et H3K27me1, il semble que ces modifications ne soient pas très abondantes au niveau des NLB du pronoyau mâle au stade PN3, début de la transcription des *major satellites*. Des expériences de FISH sur les ARN des *major satellites* associées à des immunomarquages de H3K27ac et H3K27me1 seraient nécessaires pour mettre en évidence une colocalisation et donc suggérer un lien entre modifications de H3K27 et transcription des *major satellites*. Il est également possible que H3.3 puisse jouer des rôles différents, sur l'activation de la transcription ou l'hétérochromatinisation, selon le locus considéré. Pour que cela soit possible, H3.3 doit être reconnue par des protéines différentes selon sa localisation. Les protéines chaperonnes de H3.3 pourraient être de bons candidats pour expliquer de telles différences. En effet, on connaît actuellement deux protéines chaperonnes de H3.3 chez les mammifères, DAXX, associée à ATRX (Drane et al., 2010), et HIRA (Tagami et al., 2004). Alors que HIRA semble être responsable de la présence de H3.3 au promoteur et dans le corps des gènes activement transcrits dans les cellules ES, l'incorporation de H3.3 aux régions péricentriques et télomériques semble être sous la dépendance de DAXX et ATRX (Goldberg et al., 2010). De plus, l'absence d'ATRX chez les mammifères entraîne des défauts de ségrégation des chromosomes au cours de la mitose (Ritchie et al., 2008) et de la méiose et

la perte de la contribution maternelle de ATRX induit une instabilité des centromères dans les ovocytes et le zygote (Baumann et al., 2010). Néanmoins, l'incorporation de H3.3 n'a pas été analysée dans les mutants ATRX, le lien entre H3.3 et le phénotype observé n'est donc pas certain. On peut donc émettre l'hypothèse que selon le complexe de déposition de H3.3, donc selon le locus de déposition, les protéines recrutées par H3.3 et sa chaperonne seront différentes et entraîneront soit une condensation soit une ouverture de la chromatine. Selon les données disponibles actuellement, on peut supposer que si H3.3 est déposée par DAXX, elle serait favorable à la condensation de la chromatine, alors que H3.3 déposée par HIRA induirait plutôt l'ouverture de la chromatine.

Pour définir l'importance de ces complexes de déposition dans l'embryon précoce, l'étude de la formation de l'hétérochromatine péricentrique dans des embryons n'exprimant pas DAXX peut être envisagée. Les embryons de souris n'exprimant pas HIRA présentent des défauts de gastrulation (Roberts et al., 2002), mais la chromatine ne semble pas avoir été analysée dans les embryons précoces, ce qui pourrait également être fait. De plus, des mutants conditionnels de HIRA dans la lignée germinale femelle pourraient être étudiés pour comprendre l'importance de cette protéine avant la première ZGA chez la souris. Les protéines recrutées par HIRA et DAXX dans l'embryon pourraient également être analysées pour mettre en évidence un potentiel recrutement d'enzymes de modification de la chromatine différentes selon la chaperonne étudiée.

II. H3.3K27, HP1 β et l'ARN des *major satellites*, trois acteurs de l'hétérochromatinisation des régions péricentromériques

1. Le zygote, une cellule particulière pour l'étude de la formation de l'hétérochromatine

L'hétérochromatine constitutive est caractérisée, dans les cellules somatiques par la présence de plusieurs modifications d'histones dont H3K9me3 et H4K20me3 (Martens et al., 2005). Dans le zygote, aucune triméthylation *de novo* de H3K9 ni de H4K20 n'est observée avant les stades 4-cellules et blastocyste respectivement (Liu et al., 2004, données non publiées). L'environnement chromatinien des zygotes est donc très particulier et les caractéristiques de l'hétérochromatine constitutive décrites dans les cellules somatiques ne

sont pas directement applicables au zygote. L'étude de l'hétérochromatine péricentrique au début du développement de l'embryon est donc importante.

De plus, lors de la fécondation, la chromatine paternelle est empaquetée majoritairement autour de protamines remplacées très rapidement par des histones (van der Heijden et al., 2005). Juste après la fécondation, les chromosomes paternels doivent acquérir *de novo*, les structures chromatiniennes nécessaires à la première mitose, en particulier l'hétérochromatine péricentrique. La condensation et la répression des régions péricentriques de la chromatine paternelle doivent donc être rapides et l'hétérochromatine doit être mise en place *de novo*, à partir d'un ADN quasiment nu. Le zygote, et plus spécifiquement le pronoyau mâle, est donc le seul modèle où les premières étapes de la formation de l'hétérochromatine péricentrique peuvent être étudiées *in vivo*.

Notre étude a donc permis de mettre en évidence trois acteurs importants pour la formation de l'hétérochromatine, *de novo*, dans le pronoyau mâle : H3.3K27, HP1 β et l'ARN des *major satellites*. A cause de l'environnement chromatinien spécifique du pronoyau mâle dans le zygote, il est fortement possible que l'ensemble de ces acteurs soit seulement nécessaire à l'initiation de la formation de l'hétérochromatine péricentrique et non à son maintien au cours des divisions cellulaires ultérieures. Les mécanismes d'action de ces facteurs dans le zygote sont donc potentiellement différents de ceux qui ont été et seront décrits dans les cellules somatiques.

2. H3.3K27, activation ou répression ?

D'après nos résultats, des modifications post-traductionnelles de H3.3K27 sont nécessaires à la condensation de la chromatine péricentrique paternelle au cours du stade zygote. Comme nous l'avons déjà discuté précédemment, nous ne savons pas si ces modifications sont nécessaires pour l'activation précoce de la transcription des *major satellites* ou pour la répression et la condensation de ces régions en l'absence de H3K9me3, ces deux rôles n'étant pas exclusifs. Les immunomarquages réalisés au cours du stade zygote montrent que les différents degrés de méthylation de H3K27 apparaissent après le début de la condensation de ces régions et ne colocalisent pas, avant les stades PN4-PN5, avec les régions condensées autour des NLB. Si la méthylation de H3.3K27 est impliquée dans la condensation des régions péricentromériques, il semble donc qu'elle n'intervienne pas dans les étapes très précoces d'initiation de cette condensation. La deuxième hypothèse, déjà envisagée précédemment, serait que l'acétylation ou la monométhylation de H3.3K27 jouent

un rôle dans l'expression précoce des *major satellites*. Cependant, les données immunohistochimiques ne semblent pas non plus soutenir cette hypothèse. La modification de H3.3K27 ne joue donc peut-être qu'un rôle relativement tardif, mais néanmoins nécessaire, dans l'établissement de l'hétérochromatine péricentrique paternelle.

Un moyen de savoir si le rôle majoritaire de H3.3K27 repose sur l'activation ou la répression de la transcription, serait de savoir quelle(s) modification(s) est la plus importante dans la formation de l'hétérochromatine. La mutation du gène codant pour EZH2 entraîne un arrêt du développement des embryons de souris après l'implantation ou au cours de la gastrulation. De plus, l'absence de contribution maternelle pour l'expression d'EZH2 n'entraîne pas d'arrêt du développement des embryons avant le stade 8-cellules (Puschendorf et al., 2008). EZH2 étant responsable des di- et triméthylation de H3K27, et si on considère qu'il n'existe pas d'autre protéine pouvant effectuer ces mêmes modifications, on peut donc affirmer que l'absence de di- et de triméthylation de H3K27 n'est pas suffisante pour expliquer l'arrêt du développement des embryons exprimant H3.3K27R entre les stades 2-cellules et 8-cellules. La monométhylation et/ou l'acétylation de H3.3K27 doivent donc également jouer un rôle dans l'établissement de l'hétérochromatine péricentrique. Pour confirmer ces hypothèses, nous aimerions injecter des zygotes avec des histones H3.3 mono-, di- ou triméthylées sur K27, ne pouvant pas être méthylées à un degré supérieur, et tester leur développement jusqu'au stade blastocyste. De même, il serait possible de tester si l'injection des zygotes exprimant H3.3K27R avec ces protéines ou avec une protéine H3.3K27ac induit un sauvetage du phénotype développemental observé chez les embryons mutants. Dans le but de mettre au point des expériences d'injection de protéines histones modifiées dans des embryons, j'ai réalisé des essais préliminaires au cours de ma thèse. Après expression et purification de protéines H3.1, H3.3 et H4 recombinantes couplées à une étiquette HA, de nombreux essais d'injections ont été effectués. Les protéines ont été soit injectées sous forme de monomères, soit sous forme de dimères H3/H4, à des concentrations variées, dans des zygotes au moment de la fécondation. Aucune incorporation des protéines recombinantes au sein de la chromatine des embryons n'a pu être mise en évidence quel que soit le stade d'injection et la quantité d'histones injectée. Dans la cellule, les histones H3 et H4 sont acétylées avant leur incorporation dans la chromatine (Sobel et al., 1995), l'absence de modification des histones recombinantes pourrait expliquer le fait qu'elles ne sont pas incorporées à la chromatine. Des essais d'injection d'histones pré-acétylées devront donc être mis en place pour tenter d'obtenir l'incorporation des histones injectées dans la chromatine embryonnaire.

3. Quel est le rôle de l'ARN des *major satellites* ?

Nos données suggèrent que l'ARN des *major satellites* joue un rôle dans l'hétérochromatinisation des régions péricentromériques sous la forme d'ARN double brin. De plus, selon des données plus récentes, le blocage de la transcription des *major satellites* au stade zygote entraîne un arrêt du développement des embryons au stade 2-cellules (Probst et al., 2010). Les *major satellites*, potentiellement transcrits en sens et antisens, sont donc importants pour la condensation des régions péricentriques et le développement précoce de l'embryon. Les mécanismes moléculaires par lesquels ces ARN induisent la formation d'hétérochromatine chez les mammifères ne sont pas connus, plusieurs pistes, non exclusives, peuvent cependant être envisagées.

Tout d'abord, l'importance de l'ARNdb des *major satellites* chez la souris rappelle le phénomène d'ARN interférence chez la levure. Chez *S. pombe*, la formation d'hétérochromatine péricentrique requiert la formation d'un ARN double brin à partir des transcrits sens et antisens des *major satellites* (Volpe et al., 2002). Cet ARNdb, clivé en petits ARNi par Dicer recrute le complexe RITS et induit la triméthylation de H3K9 par Clr4, homologue de SUVAR39 (figure 23) (Zhang et al., 2008) (Verdel et al., 2004). Bien qu'aucune méthylation *de novo* de H3K9 ne soit détectable dans les zygotes de souris (Santos et al., 2005), il est envisageable qu'un mécanisme comparable à un mécanisme d'ARN interférence soit mis en place dans l'embryon et entraîne par exemple la méthylation de H3K27 ou le recrutement direct de HP1. La déplétion maternelle de DICER entraîne la mort des embryons au stade zygote. Ce phénotype a été expliqué par des défauts de formation du fuseau mitotique (Bernstein et al., 2003). Cependant, l'état de compaction ainsi que de la localisation des centromères pourraient être analysées dans les zygotes mutants pour démontrer la potentielle importance des petits ARN dans la formation de l'hétérochromatine péricentrique. De plus, un immunomarquage de zygotes avec des anticorps anti-DICER et anti-AGO2 peut être envisagé pour mettre en évidence une potentielle colocalisation de ces protéines avec la chromatine péricentrique. La présence de ces protéines peut également être quantifiée dans les embryons exprimant H3.3K27R pour vérifier si un tel mécanisme est altéré ou non par la mutation.

Les ARN des *major satellites* pourraient également permettre un recrutement direct de HP1 β par son domaine *Hinge* comme le suggèrent nos expériences. Dans ce cas, aucun mécanisme intermédiaire de type ARN interférence ne serait nécessaire à la condensation de la chromatine péricentrique chez les mammifères. Une interaction a de plus été démontrée très

récemment dans des cellules somatiques de souris entre HP1 α et l'ARN transcrit à partir des *major satellites* (Maison et al., 2011). Ces ARN pourraient donc permettre le recrutement direct des protéines HP1 à la chromatine péricentrique à la fois dans l'embryon et les cellules somatiques pour en induire la condensation et la répression.

Enfin, des ARNnc et des modifications de H3K27 étant nécessaires à la condensation de la chromatine péricentrique en l'absence de H3K9me3, un mécanisme s'apparentant au mécanisme de condensation et de répression de l'hétérochromatine facultative, du Xi par exemple, peut être envisagé pour la formation de l'hétérochromatine constitutive à ce stade (Zhao et al., 2008). On peut en effet émettre l'hypothèse que l'ARN des *major satellites* serait capable, tout comme de nombreux ARNnc (Kanhere et al., 2010), de recruter EZH2 et donc le complexe PRC2, d'induire la méthylation de H3K27, la condensation de la chromatine péricentrique ainsi que la répression de la transcription à ces régions. D'autres études montrent que le complexe PRC1 est spécifiquement enrichi à l'hétérochromatine péricentrique paternelle et ce, indépendamment de la méthylation de H3K27 (Puschendorf et al., 2008). L'ARN des *major satellites* pourrait donc également recruter PRC1 aux régions péricentromériques, comme c'est le cas pour Xist au Xi (Schoeftner et al., 2006). Les ARNnc pouvant interagir avec PRC2 semblent posséder une structure spécifique en tige-boucle (Maenner et al., 2010). La présence d'une telle structure au sein de l'ARN des *major satellites* pourrait être analysée en utilisant des logiciels de prédiction de structures secondaires. De plus, des tests d'interaction entre les ARN des *major satellites* et EZH2 ou une protéine du complexe PRC1 pourraient être menés *in vitro*.

Ces trois hypothèses, de mécanisme d'ARN interférence ou de recrutement direct de HP1 β et de PRC2, ne sont pas exclusives. Il est en effet possible que plusieurs voies parallèles soient mises en place par l'embryon pour qu'une condensation rapide et efficace de la chromatine péricentrique paternelle ait lieu avant la première mitose. L'absence de colocalisation entre les marquages de HP1 β et de H3K27me au début du stade zygote suggère en effet que ces deux marques sont la conséquence de deux voies de signalisations différentes dans la cellule. Des expériences de FISH sur ARN couplées aux immunomarquages contre HP1 β et H3K27me sont nécessaires pour définir un lien entre ces marques et la transcription des *major satellites*.

III. Hétérochromatine et organisation du noyau

Aux stades zygote et 2-cellules, alors que l'hétérochromatine péricentromérique est en cours de condensation, elle s'organise en anneau autour des NLB. Plus tard au cours du développement, ainsi que dans la plupart des cellules somatiques, les régions péricentromériques s'organisent en chromocentres au sein du noyau (Probst and Almouzni, 2008). Nos données montrent qu'en présence du mutant H3.3K27R, lorsque la transcription des *major satellites* n'est pas complètement inhibée, un certain nombre de centromères ne se localise pas aux NLB au stade 2-cellules. Un lien semble donc exister entre la localisation de la chromatine et son état de condensation. La localisation de la chromatine aux NLB est-elle nécessaire à sa condensation ? A l'inverse, la condensation de la chromatine péricentromérique induit-elle son recrutement aux NLB ?

1. L'organisation du noyau influence-t-elle la transcription ?

Dans la plupart des cellules somatiques, les régions pauvres en gènes se localisent préférentiellement à proximité de la membrane nucléaire alors que les régions riches en gènes sont plutôt présentes au centre du noyau (Cremer et al., 2001). Expérimentalement, le recrutement de gènes à la périphérie du noyau dans des cellules somatiques entraîne la répression de leur transcription (Reddy et al., 2008). Cependant, si l'expression du gène est inductible par un facteur exogène, l'induction de la transcription reste possible de manière similaire à celle observée dans la région centrale du noyau (Kumaran and Spector, 2008).

Nous avons donc voulu tester si une telle relocalisation des régions péricentriques à la périphérie du noyau, à des stades où ces régions doivent être localisées autour des NLB, pouvait altérer leur transcription, leur hétérochromatinisation et le développement des embryons. L'injection de HA-ZF-émerine semble modifier la localisation des centromères et, même si cette relocalisation est très faible, semble arrêter le développement de la moitié des embryons. Pour confirmer la relocalisation des centromères, les expériences devront encore être répétées et la distance entre les centromères et les NLB ou la membrane nucléaire devra être quantifiée. Bien que la localisation des centromères nous renseigne sur la localisation des régions péricentromériques, des expériences de FISH sur l'ADN des *major satellites* devront également être mises au point dans les embryons injectés pour localiser précisément ces répétitions. De plus, le problème d'expression du contrôle HA-émerine dans les embryons

devra être résolu. En effet, l'injection d'une trop grande quantité d'émerine pourrait rigidifier le noyau, rendre sa dissociation ou sa reformation impossible lors de la mitose et donc être à l'origine des défauts de développement observés dans les embryons exprimant HA-ZF-émerine. Il est étonnant que la construction HA-émerine ne soit pas exprimée et intégrée dans les embryons alors que cette même construction est détectable après transfection de cellules NIH3T3. L'hypothèse la plus probable est que les ARNm transcrits *in vitro* ont été dégradés. De nouveaux ARNm codant pour la construction HA-émerine vont donc être transcrits *in vitro* et testés dans les embryons.

Une fois formés, les anneaux autour des NLB semblent être des structures très compactes, très difficilement mobilisables et surtout difficilement dissociables, empêcher leur formation à partir du stade zygote donnerait certainement des résultats plus facilement visibles et interprétables. Dans les expériences menées dans les cellules somatiques, l'intégration des protéines exogènes à la membrane nucléaire et donc la relocalisation des régions d'intérêt nécessite un passage par une mitose (Kumaran and Spector, 2008). Nous pensions qu'en injectant nos constructions avant la formation des pronoyaux, la protéine HA-ZF-émerine serait intégrée à partir du stade zygote. Il s'avère que nous n'avons pas réussi à détecter l'incorporation de cette protéine au stade zygote. Nous voulons maintenant essayer de réaliser et d'injecter une nouvelle construction où les *major satellites* ne seraient pas relocalisés à la membrane grâce à l'émerine mais grâce au récepteur de la lamine B (LBR : *Lamin B Receptor*).

Ensuite, dans les embryons injectés, nous devons tester si la faible relocalisation des centromères au stade 2-cellules induit des défauts d'acquisition de H3K9me3 au stade 4-cellules et 8-cellules, des défauts de formation des chromocentres et d'inhibition de la transcription des *major satellites*. Pour cela, nous réaliserons, aux stades 4-cellules et 8-cellules, des immunomarquages contre H3K9me3 et H3K27me3 dans les embryons exprimant HA-ZF-émerine. Les *major satellites* en cours de transcription seront localisés à ces mêmes stades par du FISH sur l'ARN des *major satellites*. Les expériences de FISH sur l'ADN et l'ARN des *major satellites* ont déjà été mises au point au laboratoire.

2. Les marques chromatiniennes influencent-elles l'organisation du noyau ?

Il est possible que l'état de condensation de la chromatine influence sa position dans le noyau. Cette hypothèse pourrait être testée dans les cellules somatiques où une construction formée de plusieurs répétitions de la séquence LacO pourraient être intégrée dans le génome

(Tumbar et al., 1999). Des modificateurs de la chromatine, SUVAR39, EZH2, ou MLL pourraient être fusionnés à la protéine LacI, et donc ciblées aux répétitions LacO pour mimer une hétérochromatinisation constitutive, facultative ou euchromatinisation respectivement. La localisation des répétitions LacO pourra ensuite être testée dans ces cellules pour déterminer si l'acquisition de certaines marques chromatiniennes influence la localisation de l'ADN dans le noyau.

IV. Hétérochromatine constitutive et absence de H3K9me3

L'hétérochromatine péricentrique des embryons précoces de souris, du stade zygote au stade 4-8 cellules, est caractérisée par la présence de H3K27me, de HP1 β et de H3K9me1 (Santos et al., 2005). Dans les cellules somatiques, l'hétérochromatine constitutive est caractérisée par la présence de H3K9me3 et H4K20me3, deux marques absentes de l'hétérochromatine péricentrique paternelle jusqu'aux stades 4-cellules et blastocyste respectivement (Liu et al., 2004, données non publiées). Lors d'expériences de transfert de noyau de cellule somatique dans des ovocytes de bovins, le fort taux d'embryons n'atteignant pas le stade blastocyste a été corrélé à une même proportion d'embryons présentant une hyperméthylation de H3K9 (Santos et al., 2003). L'absence de marques hétérochromatiques constitutives, à l'exception de HP1 β , semble donc être cruciale pour le développement préimplantatoire des embryons. Pourquoi l'embryon possède-t-il, lors des premières divisions, une chromatine moins fermée, moins verrouillée, que les cellules somatiques ? Aucune réponse définitive n'existe à l'heure actuelle. Il est possible que cette organisation particulière des régions hétérochromatiques soit le reflet de la plasticité du génome et de la totipotence de ces cellules jusqu'au stade 4-8 cellules. A partir de ces cellules totipotentes, tous les programmes d'expression génétique doivent pouvoir être établis. La chromatine doit donc être modifiable rapidement pour mettre en place des profils d'expression héréditaires spécifiques d'un type cellulaire plus ou moins différencié. Pour déterminer l'importance de l'absence de H3K9me3 dans les embryons, les ARNm de SUVAR39 ou de G9a pourraient être injectés dans des zygotes au moment de la fécondation, voire directement dans le pronoyau paternel après sa formation. Le développement des embryons injectés pourra alors être analysé. Cependant, il a été suggéré que l'absence de méthylation de H3K9 dans l'embryon serait la conséquence de la présence d'un inhibiteur des méthyltransférases spécifiques de H3K9 plutôt que de l'absence de ces enzymes (Liu et al., 2004). L'injection des méthyltransférases

de H3K9 dans l'embryon pourra donc être sans conséquence sur le niveau de H3K9me. De plus, la formation d'une chromatine structurée autour de NLB, comme l'absence de H3K9me3, est considérée comme une marque de totipotence des cellules. En effet, en l'absence de NLB, le développement des embryons s'arrête après seulement quelques clivages (Ogushi et al., 2008), et lors d'une expérience de transfert de noyau de cellule somatique dans un ovocyte énucléé, la reprogrammation du noyau somatique commence par la formation de NLB (Martin et al., 2006b). Le lien entre l'acquisition des NLB et l'absence de H3K9me3 pourrait donc être étudié. Grâce à une fusion avec la protéine ZF spécifique des *major satellites*, les protéines de méthylation de H3K9 pourraient être ciblées aux *major satellites* au stade zygote. La formation des NLB puis des chromocentres dans l'embryon pourra alors être testée pour mettre en évidence un potentiel lien entre l'absence de H3K9me3 dans les embryons et la formation des structures particulières que sont les NLB.

E. Matériel et méthodes

Seule l'expérience de ChIP sur les cellules transfectées sera développée ici. Les autres protocoles utilisés dans cette partie sont les mêmes que ceux de la publication 2.

I. Transfection des cellules NIH3T3 et préparation de la chromatine pour le ChIP

Les cellules NIH3T3 sont transfectées à la lipofectamineTM 2000 (Invitrogène) selon le protocole du fabricant. 24 heures après transfection, les cellules sont fixées avec 1% de formaldéhyde dilué directement dans le milieu de culture pendant 10 minutes à température ambiante. La fixation est arrêtée par ajout de 200 mM de glycine. Les cellules sont alors décollées mécaniquement et le culot de cellules est rincé deux fois dans du PBS froid. Les cellules sont ensuite lysées dans le tampon de lyse (10 mM Tris-HCl pH 7,5, 10 mM NaCl, 3 mM MgCl₂, 0,5% NP-40, inhibiteurs de protéases) pendant 10 minutes à 4°C, puis lavées deux fois dans ce même tampon. Entre chaque lavage, les noyaux sont récupérés par centrifugation à 400 g à 4°C pendant 4 minutes. Le culot est ensuite resuspendu dans le tampon de sonication (2x10⁷ cellules pour 300 µL de tampon dans un tube de 1,5 mL) et la chromatine est soniquée dans un Bioruptor® en deux étapes. Après une première sonication de 5 minutes (30 secondes ON/ 30 secondes OFF) à puissance maximale, la chromatine est centrifugée pendant 10 minutes à 13000 g. Le surnageant est conservé à 4°C (surnageant 1) et le culot correspondant à la chromatine insoluble est resuspendu dans 300 µL de tampon de sonication et soniqué pendant 10 minutes à puissance maximale. La chromatine est alors centrifugée pendant 10 minutes à 13000 g et le surnageant est conservé (surnageant 2). 50 µL sont prélevés dans les surnageants 1 et 2 pour vérifier l'efficacité de la sonication et quantifier l'ADN (tubes V1 et V2). Le reste de la chromatine est conservé à -80°C. Après incubation pendant 4h à 65°C en présence de 0,3 µg/µL de protéinase K (Sigma), l'ADN des tubes V1 et V2 est extrait au phénol-chloroforme. La quantité d'ADN est ensuite quantifiée sur un Nanodrop® et la taille des fragments est vérifiée sur un gel d'agarose. Les fragments doivent avoir une taille comprise entre 300 pb et 1000 pb. Si la taille des fragments d'ADN des tubes V1 et V2 est correcte, les chromatines des surnageants 1 et 2 sont mélangées. Un volume

correspondant à 2,5 µg d'ADN est prélevé et conservé pour représenter 10% de la chromatine de départ utilisée pour chaque CHIP notée 10% d'input.

II. Immunoprécipitation de la chromatine

Une quantité de chromatine correspondant à 25 µg d'ADN est utilisée pour chaque immunoprécipitation. Le volume des différents échantillons est ajusté avec le tampon de sonication puis chaque échantillon est dilué 10 fois dans le tampon de dilution (1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl pH 8). La chromatine est incubée avec un mélange de billes de sépharose couplées à la protéine A et à la protéine G pendant 2 heures à 4°C (les billes ont été préalablement bloquées avec 22 µg/mL d'ADN de sperme de saumon soniqué et 20 µg/mL de BSA pendant 2 heures à température ambiante). Le surnageant est ensuite récupéré et mis en présence de l'anticorps adéquat sur la nuit à 4°C (2µg d'anticorps anti-H3K9me3, Diagénodé, 5µL d'ascite de l'anticorps anti-HA 12CA5, IGBMC). Pour le CHIP contre l'étiquette Flag, la chromatine est mise en présence de 60µL de billes agarose couplées à un anticorps anti-Flag (ANTI-FLAG M2 Affinity Gel, Sigma) sur la nuit à 4°C. 20 µL de mélange de billes sépharose couplées à la protéine A et à la protéine G, bloquées, sont ajoutés aux CHIP anti-HA et anti H3K9me3 pendant 2 h à 4°C. Les billes sont ensuite lavées 3 fois dans du tampon de lavage (0,1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl pH 8), puis une fois dans du tampon final (0,1% SDS, 1% Triton X-100, 2 mM EDTA, 500 mM NaCl, 20 mM Tris-HCl pH 8). La chromatine est éluée avec 450 µL de tampon d'éluion (1% SDS, 100 mM NaHCO₃) pendant 15 minutes à température ambiante. Le surnageant, ainsi que la chromatine dite input, sont ensuite incubés pendant 4 heures à 65°C en présence de 0,3 µg/µL de protéinase K, et l'ADN est extrait au phénol-chloroforme. Le culot d'ADN est resuspendu dans 20 µL de tampon TE (10 mM Tris HCl pH 8,1, 1 mM EDTA pH 8).

III. Calcul des enrichissements

L'enrichissement en séquences d'ADN spécifiques dans les produits d'éluion par rapport à l'input est quantifié par PCR quantitative en utilisant le LightCycler® 480 SYBR Green I Master (Roche) sur un instrument LightCycler® 480 selon les recommandations du fabricant. Les réactions de PCR ont été réalisées en triplicats en utilisant 1µL de produit

d'élution de chaque ChIP et 1 μ L de l'ADN 10% input dilué 10 fois (correspondant alors à 1% de l'input réel). Une courbe standard est réalisée avec des quantités croissantes d'ADN génomique extrait à partir de souris C57BL/6, elle est utilisée pour déterminer la concentration de la séquence d'ADN dans les fractions éluées et input. L'enrichissement de chaque séquence est donné par le ratio (fraction éluée)/(input).

Les séquences des amorces utilisées sont les suivantes :

IAP : - sens : CCCC GTCCCTTTT TAGGAGA

- antisens : CTCCATGTGCTCTGCCTTCC

Sine B2 : - sens : GAGCACCTGACTGCTCTTCC

- antisens : ACACACCAGAAGAGGGCATC

Major satellites : - sens : GACGACTTGAAAAATGACGAA

- antisens : CATATTCCAGGTCCTTCAGTGTGC

CONCLUSION

CONCLUSION

Cellules germinales et zygote sont des cellules particulières d'un point de vue transcriptionnel. Alors que la transcription ne commence dans le zygote qu'au stade PN3, les ARN et protéines nécessaires aux premières étapes du développement embryonnaire sont synthétisés dans l'ovocyte au cours de sa maturation. De plus, après la fécondation, les chromatines paternelle et maternelle, organisées différemment dans les cellules germinales, vont subir des modifications épigénétiques nécessaires à leur collaboration au sein d'un même organisme.

Mes travaux de thèse ont permis de mettre en évidence, d'une part, la contribution de la protéine TBP2 à la mise en place d'un programme d'expression génique spécifique de l'ovocyte. TBP2, homologue du facteur de transcription de base TBP exprimé seulement dans les ovocytes chez la souris, est nécessaire à la maturation de ces cellules.

D'autre part, j'ai démontré un nouveau rôle du variant d'histone H3.3, et plus particulièrement de ses modifications sur la lysine 27, dans l'établissement *de novo* de l'hétérochromatine péricentrique dans le zygote de souris. Cette hétérochromatinisation requiert également la transcription des *major satellites* en ARNnc qui pourraient agir selon une voie comparable à celle des petits ARNi chez la levure. De plus, j'ai commencé un nouveau projet qui permettra de définir l'importance de l'organisation nucléaire spécifique des embryons au stade 1-cellule et 2-cellules sur l'acquisition de différentes marques épigénétiques et sur le développement.

Ces deux sujets m'ont permis de contribuer à démontrer l'importance des facteurs de transcription de base et des modifications chromatiniennes sur la mise en place de programmes d'expression génique spécifiques. Grâce à mes deux modèles d'études, j'ai abordé le phénomène de reprogrammation génétique et épigénétique nécessaire à l'acquisition de la totipotence, à la fois dans l'ovocyte et le zygote. Cependant, de nombreux mécanismes moléculaires ainsi que les protéines nécessaires au passage d'un état différencié à un état totipotent restent à être déterminés. Une caractérisation moléculaire de l'état totipotent pourrait, à long terme, permettre la reprogrammation *in vitro* de cellules somatiques en cellules totipotentes induites, iTS, à l'image de ce qui se fait actuellement pour les cellules pluripotentes induites, les iPS.

“There is no gene for the human spirit.”

GATTACA, Andrew Niccol

REFERENCES
BIBLIOGRAPHIQUES

- Adenot, P. G., Mercier, Y., Renard, J. P., and Thompson, E. M. (1997). Differential H4 acetylation of paternal and maternal chromatin precedes DNA replication and differential transcriptional activity in pronuclei of 1-cell mouse embryos. *Development* 124, 4615-4625.
- Agger, K., Christensen, J., Cloos, P. A., and Helin, K. (2008). The emerging functions of histone demethylases. *Curr Opin Genet Dev* 18, 159-168.
- Agger, K., Cloos, P. A., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A. E., and Helin, K. (2007). UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* 449, 731-734.
- Ahmad, K., and Henikoff, S. (2002). Histone H3 variants specify modes of chromatin assembly. *Proc Natl Acad Sci U S A* 99 Suppl 4, 16477-16484.
- Akey, C. W., and Luger, K. (2003). Histone chaperones and nucleosome assembly. *Curr Opin Struct Biol* 13, 6-14.
- Akhtar, W., and Veenstra, G. J. (2009). TBP2 is a substitute for TBP in *Xenopus* oocyte transcription. *BMC Biol* 7, 45.
- Albers, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2004). *Biologie moléculaire de la cellule*, Quatrième édition, Flammarion Médecine Sciences edn).
- Aoki, F., Worrall, D. M., and Schultz, R. M. (1997). Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. *Dev Biol* 181, 296-307.
- Appell, H. J., Forsberg, S., and Hollmann, W. (1988). Satellite cell activation in human skeletal muscle after training: evidence for muscle fiber neof ormation. *Int J Sports Med* 9, 297-299.
- Arnaud, P., and Feil, R. (2005). Epigenetic deregulation of genomic imprinting in human disorders and following assisted reproduction. *Birth Defects Res C Embryo Today* 75, 81-97.
- Arney, K. L., Bao, S., Bannister, A. J., Kouzarides, T., and Surani, M. A. (2002). Histone methylation defines epigenetic asymmetry in the mouse zygote. *Int J Dev Biol* 46, 317-320.
- Arnold, D. R., Francon, P., Zhang, J., Martin, K., and Clarke, H. J. (2008). Stem-loop binding protein expressed in growing oocytes is required for accumulation of mRNAs encoding histones H3 and H4 and for early embryonic development in the mouse. *Dev Biol* 313, 347-358.
- Atchison, M. L. (1988). Enhancers: mechanisms of action and cell specificity. *Annu Rev Cell Biol* 4, 127-153.
- Bachvarova, R. (1985). Gene expression during oogenesis and oocyte development in mammals. *Dev Biol (N Y)* 105, 453-524.
- Bajic, V. B., Tan, S. L., Christoffels, A., Schonbach, C., Lipovich, L., Yang, L., Hofmann, O., Kruger, A., Hide, W., Kai, C., *et al.* (2006). Mice and men: their promoter properties. *PLoS Genet* 2, e54.
- Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27, 299-308.
- Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120-124.
- Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823-837.

- Bartfai, R., Balduf, C., Hilton, T., Rathmann, Y., Hadzhiev, Y., Tora, L., Orban, L., and Muller, F. (2004). TBP2, a vertebrate-specific member of the TBP family, is required in embryonic development of zebrafish. *Curr Biol* *14*, 593-598.
- Baumann, C., Viveiros, M. M., and De La Fuente, R. (2010). Loss of maternal ATRX results in centromere instability and aneuploidy in the mammalian oocyte and pre-implantation embryo. *PLoS Genet* *6*.
- Beaujean, N., Hartshorne, G., Cavilla, J., Taylor, J., Gardner, J., Wilmut, I., Meehan, R., and Young, L. (2004). Non-conservation of mammalian preimplantation methylation dynamics. *Curr Biol* *14*, R266-267.
- Belmont, A. S. (2006). Mitotic chromosome structure and condensation. *Curr Opin Cell Biol* *18*, 632-638.
- Bench, G. S., Friz, A. M., Corzett, M. H., Morse, D. H., and Balhorn, R. (1996). DNA and total protamine masses in individual sperm from fertile mammalian subjects. *Cytometry* *23*, 263-271.
- Benevolenskaya, E. V. (2007). Histone H3K4 demethylases are essential in development and differentiation. *Biochem Cell Biol* *85*, 435-443.
- Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., *et al.* (2006a). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* *125*, 315-326.
- Bernstein, E., Duncan, E. M., Masui, O., Gil, J., Heard, E., and Allis, C. D. (2006b). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Mol Cell Biol* *26*, 2560-2569.
- Bernstein, E., Kim, S. Y., Carmell, M. A., Murchison, E. P., Alcorn, H., Li, M. Z., Mills, A. A., Elledge, S. J., Anderson, K. V., and Hannon, G. J. (2003). Dicer is essential for mouse development. *Nat Genet* *35*, 215-217.
- Bettegowda, A., and Smith, G. W. (2007). Mechanisms of maternal mRNA regulation: implications for mammalian early embryonic development. *Front Biosci* *12*, 3713-3726.
- Bhaumik, S. R., Smith, E., and Shilatifard, A. (2007). Covalent modifications of histones during development and disease pathogenesis. *Nat Struct Mol Biol* *14*, 1008-1016.
- Birke, M., Schreiner, S., Garcia-Cuellar, M. P., Mahr, K., Titgemeyer, F., and Slany, R. K. (2002). The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic Acids Res* *30*, 958-965.
- Black, B. E., and Bassett, E. A. (2008). The histone variant CENP-A and centromere specification. *Curr Opin Cell Biol* *20*, 91-100.
- Blackwood, E. M., and Kadonaga, J. T. (1998). Going the distance: a current view of enhancer action. *Science* *281*, 60-63.
- Bouniol-Baly, C., Hamraoui, L., Guibert, J., Beaujean, N., Szollosi, M. S., and Debey, P. (1999). Differential transcriptional activity associated with chromatin configuration in fully grown mouse germinal vesicle oocytes. *Biol Reprod* *60*, 580-587.
- Bouniol, C., Nguyen, E., and Debey, P. (1995). Endogenous transcription occurs at the 1-cell stage in the mouse embryo. *Exp Cell Res* *218*, 57-62.
- Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K., *et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* *441*, 349-353.
- Braunschweig, U., Hogan, G. J., Pagie, L., and van Steensel, B. (2009). Histone H1 binding is inhibited by histone variant H3.3. *Embo J* *28*, 3635-3645.
- Breathnach, R., and Chambon, P. (1981). Organization and expression of eucaryotic split genes coding for proteins. *Annu Rev Biochem* *50*, 349-383.

- Burton, A., and Torres-Padilla, M. E. (2010). Epigenetic reprogramming and development: a unique heterochromatin organization in the preimplantation mouse embryo. *Brief Funct Genomics* 9, 444-454.
- Bushnell, D. A., Westover, K. D., Davis, R. E., and Kornberg, R. D. (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocystal at 4.5 Angstroms. *Science* 303, 983-988.
- Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 298, 1039-1043.
- Carcamo, J., Buckbinder, L., and Reinberg, D. (1991). The initiator directs the assembly of a transcription factor IID-dependent transcription complex. *Proc Natl Acad Sci U S A* 88, 8052-8056.
- Caron, C., Govin, J., Rousseaux, S., and Khochbin, S. (2005). How to pack the genome for a safe trip. *Prog Mol Subcell Biol* 38, 65-89.
- Cavallini, B., Faus, I., Matthes, H., Chipoulet, J. M., Winsor, B., Egly, J. M., and Chambon, P. (1989). Cloning of the gene encoding the yeast protein BTF1Y, which can substitute for the human TATA box-binding factor. *Proc Natl Acad Sci U S A* 86, 9803-9807.
- Cavallini, B., Huet, J., Plassat, J. L., Sentenac, A., Egly, J. M., and Chambon, P. (1988). A yeast activity can substitute for the HeLa cell TATA box factor. *Nature* 334, 77-80.
- Chalkley, G. E., and Verrijzer, C. P. (1999). DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. *Embo J* 18, 4835-4845.
- Charge, S. B., and Rudnicki, M. A. (2004). Cellular and molecular regulation of muscle regeneration. *Physiol Rev* 84, 209-238.
- Chen, E. S., Zhang, K., Nicolas, E., Cam, H. P., Zofall, M., and Grewal, S. I. (2008). Cell cycle control of centromeric repeat transcription and heterochromatin assembly. *Nature* 451, 734-737.
- Chen, Z., Zang, J., Whetstine, J., Hong, X., Davrazou, F., Kutateladze, T. G., Simpson, M., Mao, Q., Pan, C. H., Dai, S., *et al.* (2006). Structural insights into histone demethylation by JMJD2 family members. *Cell* 125, 691-702.
- Cho, C., Jung-Ha, H., Willis, W. D., Goulding, E. H., Stein, P., Xu, Z., Schultz, R. M., Hecht, N. B., and Eddy, E. M. (2003). Protamine 2 deficiency leads to sperm DNA damage and embryo death in mice. *Biol Reprod* 69, 211-217.
- Cho, E. J., Kobor, M. S., Kim, M., Greenblatt, J., and Buratowski, S. (2001). Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes Dev* 15, 3319-3329.
- Choi, Y., and Rajkovic, A. (2006). Characterization of NOBOX DNA binding specificity and its regulation of Gdf9 and Pou5f1 promoters. *J Biol Chem* 281, 35747-35756.
- Chow, C. M., Georgiou, A., Szutorisz, H., Maia e Silva, A., Pombo, A., Barahona, I., Dargelos, E., Canzonetta, C., and Dillon, N. (2005). Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division. *EMBO Rep* 6, 354-360.
- Christensen, J., Agger, K., Cloos, P. A., Pasini, D., Rose, S., Sennels, L., Rappsilber, J., Hansen, K. H., Salcini, A. E., and Helin, K. (2007). RBP2 belongs to a family of demethylases, specific for tri- and dimethylated lysine 4 on histone 3. *Cell* 128, 1063-1076.
- Chuva de Sousa Lopes, S. M., Hayashi, K., Shovlin, T. C., Mifsud, W., Surani, M. A., and McLaren, A. (2008). X chromosome activity in mouse XX primordial germ cells. *PLoS Genet* 4, e30.

- Cler, E., Papai, G., Schultz, P., and Davidson, I. (2009). Recent advances in understanding the structure and function of general transcription factor TFIID. *Cell Mol Life Sci* 66, 2123-2134.
- Cloos, P. A., Christensen, J., Agger, K., Maiolica, A., Rappsilber, J., Antal, T., Hansen, K. H., and Helin, K. (2006). The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. *Nature* 442, 307-311.
- Colgan, J., and Manley, J. L. (1992). TFIID can be rate limiting in vivo for TATA-containing, but not TATA-lacking, RNA polymerase II promoters. *Genes Dev* 6, 304-315.
- Comai, L., Tanese, N., and Tjian, R. (1992). The TATA-binding protein and associated factors are integral components of the RNA polymerase I transcription factor, SL1. *Cell* 68, 965-976.
- Cook, P. R. (1995). A chromomeric model for nuclear and chromosome structure. *J Cell Sci* 108 (Pt 9), 2927-2935.
- Cormack, B. P., and Struhl, K. (1992). The TATA-binding protein is required for transcription by all three nuclear RNA polymerases in yeast cells. *Cell* 69, 685-696.
- Couldrey, C., Carlton, M. B., Nolan, P. M., Colledge, W. H., and Evans, M. J. (1999). A retroviral gene trap insertion into the histone 3.3A gene causes partial neonatal lethality, stunted growth, neuromuscular deficits and male sub-fertility in transgenic mice. *Hum Mol Genet* 8, 2489-2495.
- Cramer, P., Bushnell, D. A., Fu, J., Gnatt, A. L., Maier-Davis, B., Thompson, N. E., Burgess, R. R., Edwards, A. M., David, P. R., and Kornberg, R. D. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 288, 640-649.
- Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* 292, 1863-1876.
- Cremer, M., von Hase, J., Volm, T., Brero, A., Kreth, G., Walter, J., Fischer, C., Solovei, I., Cremer, C., and Cremer, T. (2001). Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. *Chromosome Res* 9, 541-567.
- Crick, F. (1970). Central dogma of molecular biology. *Nature* 227, 561-563.
- Crowley, T. E., Hoey, T., Liu, J. K., Jan, Y. N., Jan, L. Y., and Tjian, R. (1993). A new factor related to TATA-binding protein has highly restricted expression patterns in *Drosophila*. *Nature* 361, 557-561.
- Czernin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). *Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* 111, 185-196.
- D'Alessio, J. A., Wright, K. J., and Tjian, R. (2009). Shifting players and paradigms in cell-specific transcription. *Mol Cell* 36, 924-931.
- Dantoni, J. C., Quintin, S., Lakatos, L., Labouesse, M., and Tora, L. (2000). TBP-like factor is required for embryonic RNA polymerase II transcription in *C. elegans*. *Mol Cell* 6, 715-722.
- Dantoni, J. C., Wurtz, J. M., Poch, O., Moras, D., and Tora, L. (1999). The TBP-like factor: an alternative transcription factor in metazoa? *Trends Biochem Sci* 24, 335-339.
- Daujat, S., Weiss, T., Mohn, F., Lange, U. C., Ziegler-Birling, C., Zeissler, U., Lappe, M., Schubeler, D., Torres-Padilla, M. E., and Schneider, R. (2009). H3K64 trimethylation marks heterochromatin and is dynamically remodeled during developmental reprogramming. *Nat Struct Mol Biol* 16, 777-781.
- Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W., and Richmond, T. J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. *J Mol Biol* 319, 1097-1113.

- Davis, W., Jr., and Schultz, R. M. (2000). Developmental change in TATA-box utilization during preimplantation mouse development. *Dev Biol* 218, 275-283.
- de Napoles, M., Mermoud, J. E., Wakao, R., Tang, Y. A., Endoh, M., Appanah, R., Nesterova, T. B., Silva, J., Otte, A. P., Vidal, M., *et al.* (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev Cell* 7, 663-676.
- De Santa, F., Totaro, M. G., Prosperini, E., Notarbartolo, S., Testa, G., and Natoli, G. (2007). The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. *Cell* 130, 1083-1094.
- Deato, M. D., Marr, M. T., Sottero, T., Inouye, C., Hu, P., and Tjian, R. (2008). MyoD targets TAF3/TRF3 to activate myogenin transcription. *Mol Cell* 32, 96-105.
- Deato, M. D., and Tjian, R. (2007). Switching of the core transcription machinery during myogenesis. *Genes Dev* 21, 2137-2149.
- DeJong, J. (2006). Basic mechanisms for the control of germ cell gene expression. *Gene* 366, 39-50.
- Delaval, K., and Feil, R. (2004). Epigenetic regulation of mammalian genomic imprinting. *Curr Opin Genet Dev* 14, 188-195.
- Delaval, K., Govin, J., Cerqueira, F., Rousseaux, S., Khochbin, S., and Feil, R. (2007). Differential histone modifications mark mouse imprinting control regions during spermatogenesis. *Embo J* 26, 720-729.
- Delbarre, E., Jacobsen, B. M., Reiner, A. H., Sorensen, A. L., Kuntziger, T., and Collas, P. (2010). Chromatin environment of histone variant H3.3 revealed by quantitative imaging and genome-scale chromatin and DNA immunoprecipitation. *Mol Biol Cell* 21, 1872-1884.
- Dellino, G. I., Schwartz, Y. B., Farkas, G., McCabe, D., Elgin, S. C., and Pirrotta, V. (2004). Polycomb silencing blocks transcription initiation. *Mol Cell* 13, 887-893.
- Demers, C., Chaturvedi, C. P., Ranish, J. A., Juban, G., Lai, P., Morle, F., Aebersold, R., Dilworth, F. J., Groudine, M., and Brand, M. (2007). Activator-mediated recruitment of the MLL2 methyltransferase complex to the beta-globin locus. *Mol Cell* 27, 573-584.
- Deng, W., and Roberts, S. G. (2007). TFIIB and the regulation of transcription by RNA polymerase II. *Chromosoma* 116, 417-429.
- Di Giacomo, M., Barchi, M., Baudat, F., Edelmann, W., Keeney, S., and Jasin, M. (2005). Distinct DNA-damage-dependent and -independent responses drive the loss of oocytes in recombination-defective mouse mutants. *Proc Natl Acad Sci U S A* 102, 737-742.
- Di Pietro, C., Vento, M., Ragusa, M., Barbagallo, D., Guglielmino, M. R., Maniscalchi, T., Duro, L. R., Tomasello, L., Majorana, A., De Palma, A., *et al.* (2008). Expression analysis of TFIID in single human oocytes: new potential molecular markers of oocyte quality. *Reprod Biomed Online* 17, 338-349.
- Dillon, N., and Festenstein, R. (2002). Unravelling heterochromatin: competition between positive and negative factors regulates accessibility. *Trends Genet* 18, 252-258.
- Dodge, J. E., Kang, Y. K., Beppu, H., Lei, H., and Li, E. (2004). Histone H3-K9 methyltransferase ESET is essential for early development. *Mol Cell Biol* 24, 2478-2486.
- Dou, Y., Milne, T. A., Ruthenburg, A. J., Lee, S., Lee, J. W., Verdine, G. L., Allis, C. D., and Roeder, R. G. (2006). Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol* 13, 713-719.
- Drane, P., Ouararhni, K., Depaux, A., Shuaib, M., and Hamiche, A. (2010). The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev* 24, 1253-1265.

- Eisenmann, D. M., Dollard, C., and Winston, F. (1989). SPT15, the gene encoding the yeast TATA binding factor TFIID, is required for normal transcription initiation in vivo. *Cell* 58, 1183-1191.
- Eissenberg, J. C., James, T. C., Foster-Hartnett, D. M., Hartnett, T., Ngan, V., and Elgin, S. C. (1990). Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 87, 9923-9927.
- Elsaesser, S. J., and Allis, C. D. (2010). HIRA and Daxx Constitute Two Independent Histone H3.3-Containing Predeposition Complexes. *Cold Spring Harb Symp Quant Biol*.
- Erhardt, S., Su, I. H., Schneider, R., Barton, S., Bannister, A. J., Perez-Burgos, L., Jenuwein, T., Kouzarides, T., Tarakhovsky, A., and Surani, M. A. (2003). Consequences of the depletion of zygotic and embryonic enhancer of zeste 2 during preimplantation mouse development. *Development* 130, 4235-4248.
- Escalier, D., Eloy, L., and Garchon, H. J. (2002). Sex-specific gene expression during meiotic prophase I: Xlr (X linked, lymphocyte regulated), not its male homologue Xmr (Xlr related, meiosis regulated), is expressed in mouse oocytes. *Biol Reprod* 67, 1646-1652.
- Eskeland, R., Leeb, M., Grimes, G. R., Kress, C., Boyle, S., Sproul, D., Gilbert, N., Fan, Y., Skoultchi, A. I., Wutz, A., and Bickmore, W. A. (2010). Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination. *Mol Cell* 38, 452-464.
- Fairley, E. A., Kendrick-Jones, J., and Ellis, J. A. (1999). The Emery-Dreifuss muscular dystrophy phenotype arises from aberrant targeting and binding of emerin at the inner nuclear membrane. *J Cell Sci* 112 (Pt 15), 2571-2582.
- Falandry, C., Fourel, G., Galy, V., Ristriani, T., Horard, B., Bensimon, E., Salles, G., Gilson, E., and Magdinier, F. (2010). CLLD8/KMT1F is a lysine methyltransferase that is important for chromosome segregation. *J Biol Chem* 285, 20234-20241.
- Falender, A. E., Shimada, M., Lo, Y. K., and Richards, J. S. (2005). TAF4b, a TBP associated factor, is required for oocyte development and function. *Dev Biol* 288, 405-419.
- Faust, C., Schumacher, A., Holdener, B., and Magnuson, T. (1995). The eed mutation disrupts anterior mesoderm production in mice. *Development* 121, 273-285.
- Ferg, M., Sanges, R., Gehrig, J., Kiss, J., Bauer, M., Lovas, A., Szabo, M., Yang, L., Straehle, U., Pankratz, M. J., *et al.* (2007). The TATA-binding protein regulates maternal mRNA degradation and differential zygotic transcription in zebrafish. *Embo J* 26, 3945-3956.
- Finch, J. T., and Klug, A. (1976). Solenoidal model for superstructure in chromatin. *Proc Natl Acad Sci U S A* 73, 1897-1901.
- FitzGerald, P. C., Shlyakhtenko, A., Mir, A. A., and Vinson, C. (2004). Clustering of DNA sequences in human promoters. *Genome Res* 14, 1562-1574.
- Flach, G., Johnson, M. H., Braude, P. R., Taylor, R. A., and Bolton, V. N. (1982). The transition from maternal to embryonic control in the 2-cell mouse embryo. *Embo J* 1, 681-686.
- Fodor, B. D., Kubicek, S., Yonezawa, M., O'Sullivan, R. J., Sengupta, R., Perez-Burgos, L., Opravil, S., Mechtler, K., Schotta, G., and Jenuwein, T. (2006). Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells. *Genes Dev* 20, 1557-1562.
- Folco, H. D., Pidoux, A. L., Urano, T., and Allshire, R. C. (2008). Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. *Science* 319, 94-97.

- Fournier, C., Goto, Y., Ballestar, E., Delaval, K., Hever, A. M., Esteller, M., and Feil, R. (2002). Allele-specific histone lysine methylation marks regulatory regions at imprinted mouse genes. *Embo J* 21, 6560-6570.
- Fraser, P., and Bickmore, W. (2007). Nuclear organization of the genome and the potential for gene regulation. *Nature* 447, 413-417.
- Freiman, R. N., Albright, S. R., Zheng, S., Sha, W. C., Hammer, R. E., and Tjian, R. (2001). Requirement of tissue-selective TBP-associated factor TAFII105 in ovarian development. *Science* 293, 2084-2087.
- Fujimori, T., Kurotaki, Y., Miyazaki, J., and Nabeshima, Y. (2003). Analysis of cell lineage in two- and four-cell mouse embryos. *Development* 130, 5113-5122.
- Fussner, E., Ching, R. W., and Bazett-Jones, D. P. (2011). Living without 30nm chromatin fibers. *Trends Biochem Sci*.
- Gatewood, J. M., Cook, G. R., Balhorn, R., Schmid, C. W., and Bradbury, E. M. (1990). Isolation of four core histones from human sperm chromatin representing a minor subset of somatic histones. *J Biol Chem* 265, 20662-20666.
- Gaucher, J., Reynoird, N., Montellier, E., Boussouar, F., Rousseaux, S., and Khochbin, S. (2010). From meiosis to postmeiotic events: the secrets of histone disappearance. *Febs J* 277, 599-604.
- Gazdag, E., Rajkovic, A., Torres-Padilla, M. E., and Tora, L. (2007). Analysis of TATA-binding protein 2 (TBP2) and TBP expression suggests different roles for the two proteins in regulation of gene expression during oogenesis and early mouse development. *Reproduction* 134, 51-62.
- Geiger, J. H., Hahn, S., Lee, S., and Sigler, P. B. (1996). Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science* 272, 830-836.
- Geles, K. G., Freiman, R. N., Liu, W. L., Zheng, S., Voronina, E., and Tjian, R. (2006). Cell-type-selective induction of c-jun by TAF4b directs ovarian-specific transcription networks. *Proc Natl Acad Sci U S A* 103, 2594-2599.
- Gilbert, N., Boyle, S., Fiegler, H., Woodfine, K., Carter, N. P., and Bickmore, W. A. (2004). Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. *Cell* 118, 555-566.
- Gilbert, S. F. (2010). *Developmental Biology*, ninth edition: SINAUER ASSOCIATES).
- Giraldez, A. J., Mishima, Y., Rihel, J., Grocock, R. J., Van Dongen, S., Inoue, K., Enright, A. J., and Schier, A. F. (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* 312, 75-79.
- Girton, J. R., and Johansen, K. M. (2008). Chromatin structure and the regulation of gene expression: the lessons of PEV in *Drosophila*. *Adv Genet* 61, 1-43.
- Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A., and Kornberg, R. D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* 292, 1876-1882.
- Goldberg, A. D., Banaszynski, L. A., Noh, K. M., Lewis, P. W., Elsaesser, S. J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X., *et al.* (2010). Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* 140, 678-691.
- Govin, J., Caron, C., Escoffier, E., Ferro, M., Kuhn, L., Rousseaux, S., Eddy, E. M., Garin, J., and Khochbin, S. (2006). Post-meiotic shifts in HSPA2/HSP70.2 chaperone activity during mouse spermatogenesis. *J Biol Chem* 281, 37888-37892.
- Govin, J., Caron, C., Rousseaux, S., and Khochbin, S. (2005). Testis-specific histone H3 expression in somatic cells. *Trends Biochem Sci* 30, 357-359.
- Grob, P., Cruse, M. J., Inouye, C., Peris, M., Penczek, P. A., Tjian, R., and Nogales, E. (2006). Cryo-electron microscopy studies of human TFIID: conformational breathing in the integration of gene regulatory cues. *Structure* 14, 511-520.

- Grosschedl, R., and Birnstiel, M. L. (1980). Identification of regulatory sequences in the prelude sequences of an H2A histone gene by the study of specific deletion mutants in vivo. *Proc Natl Acad Sci U S A* 77, 1432-1436.
- Grunewald, S., Paasch, U., Glander, H. J., and Anderegg, U. (2005). Mature human spermatozoa do not transcribe novel RNA. *Andrologia* 37, 69-71.
- Guenatri, M., Bailly, D., Maison, C., and Almouzni, G. (2004). Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J Cell Biol* 166, 493-505.
- Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., Tsai, M. C., Hung, T., Argani, P., Rinn, J. L., *et al.* (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464, 1071-1076.
- Gur, Y., and Breitbart, H. (2006). Mammalian sperm translate nuclear-encoded proteins by mitochondrial-type ribosomes. *Genes Dev* 20, 411-416.
- Hajkova, P., Ancelin, K., Waldmann, T., Lacoste, N., Lange, U. C., Cesari, F., Lee, C., Almouzni, G., Schneider, R., and Surani, M. A. (2008). Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* 452, 877-881.
- Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., and Surani, M. A. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mech Dev* 117, 15-23.
- Hajkova, P., Jeffries, S. J., Lee, C., Miller, N., Jackson, S. P., and Surani, M. A. (2010). Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* 329, 78-82.
- Hake, S. B., Garcia, B. A., Duncan, E. M., Kauer, M., Dellaire, G., Shabanowitz, J., Bazett-Jones, D. P., Allis, C. D., and Hunt, D. F. (2006). Expression patterns and post-translational modifications associated with mammalian histone H3 variants. *J Biol Chem* 281, 559-568.
- Hammoud, S. S., Nix, D. A., Zhang, H., Purwar, J., Carrell, D. T., and Cairns, B. R. (2009). Distinctive chromatin in human sperm packages genes for embryo development. *Nature* 460, 473-478.
- Hansen, J. C. (2002). Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. *Annu Rev Biophys Biomol Struct* 31, 361-392.
- Hansen, S. K., Takada, S., Jacobson, R. H., Lis, J. T., and Tjian, R. (1997). Transcription properties of a cell type-specific TATA-binding protein, TRF. *Cell* 91, 71-83.
- Hazzouri, M., Pivot-Pajot, C., Faure, A. K., Usson, Y., Pelletier, R., Sele, B., Khochbin, S., and Rousseaux, S. (2000). Regulated hyperacetylation of core histones during mouse spermatogenesis: involvement of histone deacetylases. *Eur J Cell Biol* 79, 950-960.
- Heard, E., and Disteché, C. M. (2006). Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev* 20, 1848-1867.
- Heitz, E. (1928). Das heterochromatin der moose. *I JahrbWiss Botanik* 69, 762-818.
- Herr, A. J., Jensen, M. B., Dalmay, T., and Baulcombe, D. C. (2005). RNA polymerase IV directs silencing of endogenous DNA. *Science* 308, 118-120.
- Hirasawa, R., Chiba, H., Kaneda, M., Tajima, S., Li, E., Jaenisch, R., and Sasaki, H. (2008). Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes Dev* 22, 1607-1616.
- Hobbs, N. K., Bondareva, A. A., Barnett, S., Capecchi, M. R., and Schmidt, E. E. (2002). Removing the vertebrate-specific TBP N terminus disrupts placental beta2m-dependent interactions with the maternal immune system. *Cell* 110, 43-54.
- Holmquist, G. P. (1987). Role of replication time in the control of tissue-specific gene expression. *Am J Hum Genet* 40, 151-173.

- Hong, L., Schroth, G. P., Matthews, H. R., Yau, P., and Bradbury, E. M. (1993). Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. *J Biol Chem* 268, 305-314.
- Horikoshi, M., Wang, C. K., Fujii, H., Cromlish, J. A., Weil, P. A., and Roeder, R. G. (1989). Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box. *Nature* 341, 299-303.
- Horowitz, R. A., Agard, D. A., Sedat, J. W., and Woodcock, C. L. (1994). The three-dimensional architecture of chromatin in situ: electron tomography reveals fibers composed of a continuously variable zig-zag nucleosomal ribbon. *J Cell Biol* 125, 1-10.
- Huang, Y., Fang, J., Bedford, M. T., Zhang, Y., and Xu, R. M. (2006). Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science* 312, 748-751.
- Indra, A. K., Mohan, W. S., 2nd, Frontini, M., Scheer, E., Messaddeq, N., Metzger, D., and Tora, L. (2005). TAF10 is required for the establishment of skin barrier function in foetal, but not in adult mouse epidermis. *Dev Biol* 285, 28-37.
- Ingham, P. W. (1985). A clonal analysis of the requirement for the trithorax gene in the diversification of segments in *Drosophila*. *J Embryol Exp Morphol* 89, 349-365.
- Jackson, J. P., Lindroth, A. M., Cao, X., and Jacobsen, S. E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 416, 556-560.
- Jackson, V., and Chalkley, R. (1985). Histone segregation on replicating chromatin. *Biochemistry* 24, 6930-6938.
- Jacob, Y., Feng, S., LeBlanc, C. A., Bernatavichute, Y. V., Stroud, H., Cokus, S., Johnson, L. M., Pellegrini, M., Jacobsen, S. E., and Michaels, S. D. (2009). ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. *Nat Struct Mol Biol* 16, 763-768.
- Jacobi, U. G., Akkers, R. C., Pierson, E. S., Weeks, D. L., Dagle, J. M., and Veenstra, G. J. (2007). TBP paralogs accommodate metazoan- and vertebrate-specific developmental gene regulation. *Embo J* 26, 3900-3909.
- Jahn, C. L., Baran, M. M., and Bachvarova, R. (1976). Stability of RNA synthesized by the mouse oocyte during its major growth phase. *J Exp Zool* 197, 161-171.
- Jallow, Z., Jacobi, U. G., Weeks, D. L., Dawid, I. B., and Veenstra, G. J. (2004). Specialized and redundant roles of TBP and a vertebrate-specific TBP paralog in embryonic gene regulation in *Xenopus*. *Proc Natl Acad Sci U S A* 101, 13525-13530.
- Jenuwein, T., and Allis, C. D. (2001). Translating the histone code. *Science* 293, 1074-1080.
- Jenuwein, T., Laible, G., Dorn, R., and Reuter, G. (1998). SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cell Mol Life Sci* 54, 80-93.
- Jin, C., and Felsenfeld, G. (2007). Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev* 21, 1519-1529.
- Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat Genet* 41, 941-945.
- Johnson, M. H., and McConnell, J. M. (2004). Lineage allocation and cell polarity during mouse embryogenesis. *Semin Cell Dev Biol* 15, 583-597.
- Juven-Gershon, T., Cheng, S., and Kadonaga, J. T. (2006). Rational design of a super core promoter that enhances gene expression. *Nat Methods* 3, 917-922.

- Kaneko, S., Li, G., Son, J., Xu, C. F., Margueron, R., Neubert, T. A., and Reinberg, D. (2010). Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. *Genes Dev* 24, 2615-2620.
- Kanellopoulou, C., Muljo, S. A., Kung, A. L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D. M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev* 19, 489-501.
- Kanhere, A., Viiri, K., Araujo, C. C., Rasaiyaah, J., Bouwman, R. D., Whyte, W. A., Pereira, C. F., Brookes, E., Walker, K., Bell, G. W., *et al.* (2010). Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2. *Mol Cell* 38, 675-688.
- Kaplan, G., Abreu, S. L., and Bachvarova, R. (1982). rRNA accumulation and protein synthetic patterns in growing mouse oocytes. *J Exp Zool* 220, 361-370.
- Kappes, F., Waldmann, T., Mathew, V., Yu, J., Zhang, L., Khodadoust, M. S., Chinnaiyan, A. M., Luger, K., Erhardt, S., Schneider, R., and Markovitz, D. M. (2011). The DEK oncoprotein is a Su(var) that is essential to heterochromatin integrity. *Genes Dev* 25, 673-678.
- Kato, Y., Kaneda, M., Hata, K., Kumaki, K., Hisano, M., Kohara, Y., Okano, M., Li, E., Nozaki, M., and Sasaki, H. (2007). Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Hum Mol Genet* 16, 2272-2280.
- Kaufmann, J., Ahrens, K., Koop, R., Smale, S. T., and Muller, R. (1998). CIF150, a human cofactor for transcription factor IID-dependent initiator function. *Mol Cell Biol* 18, 233-239.
- Kelly, S. J. (1977). Studies of the developmental potential of 4- and 8-cell stage mouse blastomeres. *J Exp Zool* 200, 365-376.
- Khalil, A. M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., Thomas, K., Presser, A., Bernstein, B. E., van Oudenaarden, A., *et al.* (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A* 106, 11667-11672.
- Khrapunov, S., and Brenowitz, M. (2007). Influence of the N-terminal domain and divalent cations on self-association and DNA binding by the *Saccharomyces cerevisiae* TATA binding protein. *Biochemistry* 46, 4876-4887.
- Kim, T. H., Barrera, L. O., Zheng, M., Qu, C., Singer, M. A., Richmond, T. A., Wu, Y., Green, R. D., and Ren, B. (2005). A high-resolution map of active promoters in the human genome. *Nature* 436, 876-880.
- Kim, Y., Geiger, J. H., Hahn, S., and Sigler, P. B. (1993). Crystal structure of a yeast TBP/TATA-box complex. *Nature* 365, 512-520.
- Kimmins, S., Kotaja, N., Davidson, I., and Sassone-Corsi, P. (2004). Testis-specific transcription mechanisms promoting male germ-cell differentiation. *Reproduction* 128, 5-12.
- Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., *et al.* (2006). Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res* 16, 55-65.
- Kirmizis, A., Bartley, S. M., Kuzmichev, A., Margueron, R., Reinberg, D., Green, R., and Farnham, P. J. (2004). Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. *Genes Dev* 18, 1592-1605.
- Klose, R. J., Kallin, E. M., and Zhang, Y. (2006). JmjC-domain-containing proteins and histone demethylation. *Nat Rev Genet* 7, 715-727.

- Klose, R. J., Yan, Q., Tothova, Z., Yamane, K., Erdjument-Bromage, H., Tempst, P., Gilliland, D. G., Zhang, Y., and Kaelin, W. G., Jr. (2007). The retinoblastoma binding protein RBP2 is an H3K4 demethylase. *Cell* *128*, 889-900.
- Kopecny, V., Landa, V., and Pavlok, A. (1995). Localization of nucleic acids in the nucleoli of oocytes and early embryos of mouse and hamster: an autoradiographic study. *Mol Reprod Dev* *41*, 449-458.
- Kopytova, D. V., Krasnov, A. N., Kopantceva, M. R., Nabirochkina, E. N., Nikolenko, J. V., Maksimenko, O., Kurshakova, M. M., Lebedeva, L. A., Yerokhin, M. M., Simonova, O. B., *et al.* (2006). Two isoforms of Drosophila TRF2 are involved in embryonic development, premeiotic chromatin condensation, and proper differentiation of germ cells of both sexes. *Mol Cell Biol* *26*, 7492-7505.
- Kornberg, R. D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* *184*, 868-871.
- Kotaja, N., Bhattacharyya, S. N., Jaskiewicz, L., Kimmins, S., Parvinen, M., Filipowicz, W., and Sassone-Corsi, P. (2006). The chromatoid body of male germ cells: similarity with processing bodies and presence of Dicer and microRNA pathway components. *Proc Natl Acad Sci U S A* *103*, 2647-2652.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* *128*, 693-705.
- Kumaran, R. I., and Spector, D. L. (2008). A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *J Cell Biol* *180*, 51-65.
- Kuras, L., and Struhl, K. (1999). Binding of TBP to promoters in vivo is stimulated by activators and requires Pol II holoenzyme. *Nature* *399*, 609-613.
- Kuzmichev, A., Jenuwein, T., Tempst, P., and Reinberg, D. (2004). Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. *Mol Cell* *14*, 183-193.
- Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev* *16*, 2893-2905.
- Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* *410*, 116-120.
- Lalancette, C., Miller, D., Li, Y., and Krawetz, S. A. (2008). Paternal contributions: new functional insights for spermatozoal RNA. *J Cell Biochem* *104*, 1570-1579.
- Lee, J. H., and Skalnik, D. G. (2008). Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. *Mol Cell Biol* *28*, 609-618.
- Lee, M. G., Norman, J., Shilatifard, A., and Shiekhatar, R. (2007a). Physical and functional association of a trimethyl H3K4 demethylase and Ring6a/MBLR, a polycomb-like protein. *Cell* *128*, 877-887.
- Lee, M. G., Villa, R., Trojer, P., Norman, J., Yan, K. P., Reinberg, D., Di Croce, L., and Shiekhatar, R. (2007b). Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science* *318*, 447-450.
- Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A. H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* *13*, 1192-1200.
- Leurent, C., Sanders, S., Ruhlmann, C., Mallouh, V., Weil, P. A., Kirschner, D. B., Tora, L., and Schultz, P. (2002). Mapping histone fold TAFs within yeast TFIID. *Embo J* *21*, 3424-3433.

- Lewis, J. D., Meehan, R. R., Henzel, W. J., Maurer-Fogy, I., Jeppesen, P., Klein, F., and Bird, A. (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* 69, 905-914.
- Lewis, P. W., Elsaesser, S. J., Noh, K. M., Stadler, S. C., and Allis, C. D. (2010). Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proc Natl Acad Sci U S A* 107, 14075-14080.
- Li, B., Carey, M., and Workman, J. L. (2007). The role of chromatin during transcription. *Cell* 128, 707-719.
- Li, E., Beard, C., and Jaenisch, R. (1993a). Role for DNA methylation in genomic imprinting. *Nature* 366, 362-365.
- Li, W., Nagaraja, S., Delcuve, G. P., Hendzel, M. J., and Davie, J. R. (1993b). Effects of histone acetylation, ubiquitination and variants on nucleosome stability. *Biochem J* 296 (Pt 3), 737-744.
- Lindhout, B. I., Fransz, P., Tessadori, F., Meckel, T., Hooykaas, P. J., and van der Zaal, B. J. (2007). Live cell imaging of repetitive DNA sequences via GFP-tagged polydactyl zinc finger proteins. *Nucleic Acids Res* 35, e107.
- Liu, H., Kim, J. M., and Aoki, F. (2004). Regulation of histone H3 lysine 9 methylation in oocytes and early pre-implantation embryos. *Development* 131, 2269-2280.
- Loyola, A., Bonaldi, T., Roche, D., Imhof, A., and Almouzni, G. (2006). PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Mol Cell* 24, 309-316.
- Lu, J., and Gilbert, D. M. (2008). Cell cycle regulated transcription of heterochromatin in mammals vs. fission yeast: functional conservation or coincidence? *Cell Cycle* 7, 1907-1910.
- Lu, X., Simon, M. D., Chodaparambil, J. V., Hansen, J. C., Shokat, K. M., and Luger, K. (2008). The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nat Struct Mol Biol* 15, 1122-1124.
- Lucifero, D., Mertineit, C., Clarke, H. J., Bestor, T. H., and Trasler, J. M. (2002). Methylation dynamics of imprinted genes in mouse germ cells. *Genomics* 79, 530-538.
- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.
- Mack, D. H., Vartikar, J., Pipas, J. M., and Laimins, L. A. (1993). Specific repression of TATA-mediated but not initiator-mediated transcription by wild-type p53. *Nature* 363, 281-283.
- MacLaughlin, J., and Ternier, C. (1973). Ribonucleic acid synthesis by spermatozoa from the rat and hamster. *Biochem J* 133, 635-639.
- Maenner, S., Blaud, M., Fouillen, L., Savoye, A., Marchand, V., Dubois, A., Sanglier-Cianferani, S., Van Dorsselaer, A., Clerc, P., Avner, P., *et al.* (2010). 2-D structure of the A region of Xist RNA and its implication for PRC2 association. *PLoS Biol* 8, e1000276.
- Maison, C., Bailly, D., Peters, A. H., Quivy, J. P., Roche, D., Taddei, A., Lachner, M., Jenuwein, T., and Almouzni, G. (2002). Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat Genet* 30, 329-334.
- Maison, C., Bailly, D., Roche, D., de Oca, R. M., Probst, A. V., Vassias, I., Dingli, F., Lombard, B., Loew, D., Quivy, J. P., and Almouzni, G. (2011). SUMOylation promotes de novo targeting of HP1alpha to pericentric heterochromatin. *Nat Genet*.

- Malik, H. S., and Henikoff, S. (2003). Phylogenomics of the nucleosome. *Nat Struct Biol* 10, 882-891.
- Manilal, S., Nguyen, T. M., Sewry, C. A., and Morris, G. E. (1996). The Emery-Dreifuss muscular dystrophy protein, emerin, is a nuclear membrane protein. *Hum Mol Genet* 5, 801-808.
- Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J., Drury, W. J., 3rd, Voigt, P., Martin, S. R., Taylor, W. R., De Marco, V., *et al.* (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* 461, 762-767.
- Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C. L., Dynlacht, B. D., and Reinberg, D. (2008). Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol Cell* 32, 503-518.
- Martens, J. H., O'Sullivan, R. J., Braunschweig, U., Opravil, S., Radolf, M., Steinlein, P., and Jenuwein, T. (2005). The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *Embo J* 24, 800-812.
- Martianov, I., Brancorsini, S., Gansmuller, A., Parvinen, M., Davidson, I., and Sassone-Corsi, P. (2002a). Distinct functions of TBP and TLF/TRF2 during spermatogenesis: requirement of TLF for heterochromatic chromocenter formation in haploid round spermatids. *Development* 129, 945-955.
- Martianov, I., Fimia, G. M., Dierich, A., Parvinen, M., Sassone-Corsi, P., and Davidson, I. (2001). Late arrest of spermiogenesis and germ cell apoptosis in mice lacking the TBP-like TLF/TRF2 gene. *Mol Cell* 7, 509-515.
- Martianov, I., Viville, S., and Davidson, I. (2002b). RNA polymerase II transcription in murine cells lacking the TATA binding protein. *Science* 298, 1036-1039.
- Martin, C., Beaujean, N., Brochard, V., Audouard, C., Zink, D., and Debey, P. (2006a). Genome restructuring in mouse embryos during reprogramming and early development. *Dev Biol* 292, 317-332.
- Martin, C., Brochard, V., Migne, C., Zink, D., Debey, P., and Beaujean, N. (2006b). Architectural reorganization of the nuclei upon transfer into oocytes accompanies genome reprogramming. *Mol Reprod Dev* 73, 1102-1111.
- Martinez, E., Ge, H., Tao, Y., Yuan, C. X., Palhan, V., and Roeder, R. G. (1998). Novel cofactors and TFIIA mediate functional core promoter selectivity by the human TAFII150-containing TFIID complex. *Mol Cell Biol* 18, 6571-6583.
- Matsui, T., Segall, J., Weil, P. A., and Roeder, R. G. (1980). Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *J Biol Chem* 255, 11992-11996.
- Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000a). Demethylation of the zygotic paternal genome. *Nature* 403, 501-502.
- Mayer, W., Smith, A., Fundele, R., and Haaf, T. (2000b). Spatial separation of parental genomes in preimplantation mouse embryos. *J Cell Biol* 148, 629-634.
- Meistrich, M. L., Mohapatra, B., Shirley, C. R., and Zhao, M. (2003). Roles of transition nuclear proteins in spermiogenesis. *Chromosoma* 111, 483-488.
- Memili, E., and First, N. L. (2000). Zygotic and embryonic gene expression in cow: a review of timing and mechanisms of early gene expression as compared with other species. *Zygote* 8, 87-96.
- Metzger, E., Wissmann, M., Yin, N., Muller, J. M., Schneider, R., Peters, A. H., Gunther, T., Buettner, R., and Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* 437, 436-439.
- Milne, T. A., Dou, Y., Martin, M. E., Brock, H. W., Roeder, R. G., and Hess, J. L. (2005). MLL associates specifically with a subset of transcriptionally active target genes. *Proc Natl Acad Sci U S A* 102, 14765-14770.

- Mito, Y., Henikoff, J. G., and Henikoff, S. (2005). Genome-scale profiling of histone H3.3 replacement patterns. *Nat Genet* 37, 1090-1097.
- Mo, R., Rao, S. M., and Zhu, Y. J. (2006). Identification of the MLL2 complex as a coactivator for estrogen receptor alpha. *J Biol Chem* 281, 15714-15720.
- Moazed, D. (2009). Small RNAs in transcriptional gene silencing and genome defence. *Nature* 457, 413-420.
- Mohn, F., Weber, M., Rebhan, M., Roloff, T. C., Richter, J., Stadler, M. B., Bibel, M., and Schubeler, D. (2008). Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell* 30, 755-766.
- Montgomery, N. D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S. J., Otte, A. P., and Magnuson, T. (2005). The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Curr Biol* 15, 942-947.
- Moore, G. P., Lintern-Moore, S., Peters, H., and Faber, M. (1974). RNA synthesis in the mouse oocyte. *J Cell Biol* 60, 416-422.
- Moore, P. A., Ozer, J., Salunek, M., Jan, G., Zerby, D., Campbell, S., and Lieberman, P. M. (1999). A human TATA binding protein-related protein with altered DNA binding specificity inhibits transcription from multiple promoters and activators. *Mol Cell Biol* 19, 7610-7620.
- Muchardt, C., Guilleme, M., Seeler, J. S., Trouche, D., Dejean, A., and Yaniv, M. (2002). Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1alpha. *EMBO Rep* 3, 975-981.
- Muller, F., Lakatos, L., Dantonel, J., Strahle, U., and Tora, L. (2001). TBP is not universally required for zygotic RNA polymerase II transcription in zebrafish. *Curr Biol* 11, 282-287.
- Muller, F., and Tora, L. (2004). The multicoloured world of promoter recognition complexes. *Embo J* 23, 2-8.
- Murchison, E. P., Partridge, J. F., Tam, O. H., Cheloufi, S., and Hannon, G. J. (2005). Characterization of Dicer-deficient murine embryonic stem cells. *Proc Natl Acad Sci U S A* 102, 12135-12140.
- Myers, M., Britt, K. L., Wreford, N. G., Ebling, F. J., and Kerr, J. B. (2004). Methods for quantifying follicular numbers within the mouse ovary. *Reproduction* 127, 569-580.
- Nakajima, N., Horikoshi, M., and Roeder, R. G. (1988). Factors involved in specific transcription by mammalian RNA polymerase II: purification, genetic specificity, and TATA box-promoter interactions of TFIID. *Mol Cell Biol* 8, 4028-4040.
- Nielsen, A. L., Ortiz, J. A., You, J., Oulad-Abdelghani, M., Khechumian, R., Gansmuller, A., Chambon, P., and Losson, R. (1999). Interaction with members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family. *Embo J* 18, 6385-6395.
- Nielsen, A. L., Oulad-Abdelghani, M., Ortiz, J. A., Remboutsika, E., Chambon, P., and Losson, R. (2001). Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins. *Mol Cell* 7, 729-739.
- Nikolov, D. B., Chen, H., Halay, E. D., Hoffman, A., Roeder, R. G., and Burley, S. K. (1996). Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc Natl Acad Sci U S A* 93, 4862-4867.
- Nikolov, D. B., Chen, H., Halay, E. D., Usheva, A. A., Hisatake, K., Lee, D. K., Roeder, R. G., and Burley, S. K. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* 377, 119-128.

- Nislow, C., Ray, E., and Pillus, L. (1997). SET1, a yeast member of the trithorax family, functions in transcriptional silencing and diverse cellular processes. *Mol Biol Cell* 8, 2421-2436.
- O'Carroll, D., Erhardt, S., Pagani, M., Barton, S. C., Surani, M. A., and Jenuwein, T. (2001). The polycomb-group gene *Ezh2* is required for early mouse development. *Mol Cell Biol* 21, 4330-4336.
- O'Neill, L. P., VerMilyea, M. D., and Turner, B. M. (2006). Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations. *Nat Genet* 38, 835-841.
- Ogushi, S., Palmieri, C., Fulka, H., Saitou, M., Miyano, T., and Fulka, J., Jr. (2008). The maternal nucleolus is essential for early embryonic development in mammals. *Science* 319, 613-616.
- Okada, Y., Yamagata, K., Hong, K., Wakayama, T., and Zhang, Y. (2010). A role for the elongator complex in zygotic paternal genome demethylation. *Nature* 463, 554-558.
- Okamoto, I., and Heard, E. (2009). Lessons from comparative analysis of X-chromosome inactivation in mammals. *Chromosome Res* 17, 659-669.
- Olins, A. L., and Olins, D. E. (1974). Spheroid chromatin units (v bodies). *Science* 183, 330-332.
- Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes Dev* 10, 2657-2683.
- Osley, M. A. (1991). The regulation of histone synthesis in the cell cycle. *Annu Rev Biochem* 60, 827-861.
- Ostermeier, G. C., Miller, D., Huntriss, J. D., Diamond, M. P., and Krawetz, S. A. (2004). Reproductive biology: delivering spermatozoan RNA to the oocyte. *Nature* 429, 154.
- Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. *Curr Biol* 10, 475-478.
- Otte, A. P., and Kwaks, T. H. (2003). Gene repression by Polycomb group protein complexes: a distinct complex for every occasion? *Curr Opin Genet Dev* 13, 448-454.
- Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* 4, 281-300.
- Palancade, B., and Bensaude, O. (2003). Investigating RNA polymerase II carboxyl-terminal domain (CTD) phosphorylation. *Eur J Biochem* 270, 3859-3870.
- Palmer, D. K., O'Day, K., and Margolis, R. L. (1990). The centromere specific histone CENP-A is selectively retained in discrete foci in mammalian sperm nuclei. *Chromosoma* 100, 32-36.
- Papai, G., Schultz, P. (2010). Régulation de la transcription par le coactivateur TFIID. *médecine/sciences* 26, 1018-1019.
- Papai, G., Tripathi, M. K., Ruhlmann, C., Layer, J. H., Weil, P. A., and Schultz, P. (2010). TFIIA and the transactivator Rap1 cooperate to commit TFIID for transcription initiation. *Nature* 465, 956-960.
- Pasini, D., Bracken, A. P., Jensen, M. R., Lazzerini Denchi, E., and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *Embo J* 23, 4061-4071.
- Passarge, E. (1979). Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. *Am J Hum Genet* 31, 106-115.
- Patel, S. R., Kim, D., Levitan, I., and Dressler, G. R. (2007). The BRCT-domain containing protein PTIP links PAX2 to a histone H3, lysine 4 methyltransferase complex. *Dev Cell* 13, 580-592.

- Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. *Nature* *379*, 131-137.
- Persengiev, S. P., Zhu, X., Dixit, B. L., Maston, G. A., Kittler, E. L., and Green, M. R. (2003). TRF3, a TATA-box-binding protein-related factor, is vertebrate-specific and widely expressed. *Proc Natl Acad Sci U S A* *100*, 14887-14891.
- Peters, A. H., Kubicek, S., Mechtler, K., O'Sullivan, R. J., Derijck, A. A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., *et al.* (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol Cell* *12*, 1577-1589.
- Peters, A. H., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schofer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., *et al.* (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* *107*, 323-337.
- Phatnani, H. P., and Greenleaf, A. L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev* *20*, 2922-2936.
- Piko, L., and Clegg, K. B. (1982). Quantitative changes in total RNA, total poly(A), and ribosomes in early mouse embryos. *Dev Biol* *89*, 362-378.
- Piotrowska-Nitsche, K., Perea-Gomez, A., Haraguchi, S., and Zernicka-Goetz, M. (2005). Four-cell stage mouse blastomeres have different developmental properties. *Development* *132*, 479-490.
- Polo, S. E., Roche, D., and Almouzni, G. (2006). New histone incorporation marks sites of UV repair in human cells. *Cell* *127*, 481-493.
- Probst, A. V., and Almouzni, G. (2008). Pericentric heterochromatin: dynamic organization during early development in mammals. *Differentiation* *76*, 15-23.
- Probst, A. V., Okamoto, I., Casanova, M., El Marjou, F., Le Baccon, P., and Almouzni, G. (2010). A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Dev Cell* *19*, 625-638.
- Probst, A. V., Santos, F., Reik, W., Almouzni, G., and Dean, W. (2007). Structural differences in centromeric heterochromatin are spatially reconciled on fertilisation in the mouse zygote. *Chromosoma* *116*, 403-415.
- Pugh, B. F., and Tjian, R. (1990). Mechanism of transcriptional activation by Sp1: evidence for coactivators. *Cell* *61*, 1187-1197.
- Puschendorf, M., Terranova, R., Boutsma, E., Mao, X., Isono, K., Brykczynska, U., Kolb, C., Otte, A. P., Koseki, H., Orkin, S. H., *et al.* (2008). PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat Genet* *40*, 411-420.
- Racki, W. J., and Richter, J. D. (2006). CPEB controls oocyte growth and follicle development in the mouse. *Development* *133*, 4527-4537.
- Rajkovic, A., Pangas, S. A., Ballow, D., Suzumori, N., and Matzuk, M. M. (2004). NOBOX deficiency disrupts early folliculogenesis and oocyte-specific gene expression. *Science* *305*, 1157-1159.
- Ranish, J. A., Yudkovsky, N., and Hahn, S. (1999). Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev* *13*, 49-63.
- Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., and Jenuwein, T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* *406*, 593-599.
- Reddy, K. L., Zullo, J. M., Bertolino, E., and Singh, H. (2008). Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* *452*, 243-247.

- Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. *Science* 293, 1089-1093.
- Reinberg, D., Horikoshi, M., and Roeder, R. G. (1987). Factors involved in specific transcription in mammalian RNA polymerase II. Functional analysis of initiation factors IIA and IID and identification of a new factor operating at sequences downstream of the initiation site. *J Biol Chem* 262, 3322-3330.
- Reinberg, D., and Roeder, R. G. (1987). Factors involved in specific transcription by mammalian RNA polymerase II. Purification and functional analysis of initiation factors IIB and IIE. *J Biol Chem* 262, 3310-3321.
- Richardson, C., Horikoshi, N., and Pandita, T. K. (2004). The role of the DNA double-strand break response network in meiosis. *DNA Repair (Amst)* 3, 1149-1164.
- Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu Rev Genet* 38, 413-443.
- Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Bruggmann, S. A., Goodnough, L. H., Helms, J. A., Farnham, P. J., Segal, E., and Chang, H. Y. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129, 1311-1323.
- Ritchie, K., Seah, C., Moulin, J., Isaac, C., Dick, F., and Berube, N. G. (2008). Loss of ATRX leads to chromosome cohesion and congression defects. *J Cell Biol* 180, 315-324.
- Roberts, C., Sutherland, H. F., Farmer, H., Kimber, W., Halford, S., Carey, A., Brickman, J. M., Wynshaw-Boris, A., and Scambler, P. J. (2002). Targeted mutagenesis of the Hira gene results in gastrulation defects and patterning abnormalities of mesoendodermal derivatives prior to early embryonic lethality. *Mol Cell Biol* 22, 2318-2328.
- Rossant, J. (1976). Postimplantation development of blastomeres isolated from 4- and 8-cell mouse eggs. *J Embryol Exp Morphol* 36, 283-290.
- Rossant, J., and Lis, W. T. (1979). Potential of isolated mouse inner cell masses to form trophectoderm derivatives in vivo. *Dev Biol* 70, 255-261.
- Rougeulle, C., Chaumeil, J., Sarma, K., Allis, C. D., Reinberg, D., Avner, P., and Heard, E. (2004). Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol Cell Biol* 24, 5475-5484.
- Rougier, N., Bourchis, D., Gomes, D. M., Niveleau, A., Plachot, M., Paldi, A., and Viegas-Pequignot, E. (1998). Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev* 12, 2108-2113.
- Rudert, F., Bronner, S., Garnier, J. M., and Dolle, P. (1995). Transcripts from opposite strands of gamma satellite DNA are differentially expressed during mouse development. *Mamm Genome* 6, 76-83.
- Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol* 241, 172-182.
- Santos, F., Peters, A. H., Otte, A. P., Reik, W., and Dean, W. (2005). Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev Biol* 280, 225-236.
- Santos, F., Zakhartchenko, V., Stojkovic, M., Peters, A., Jenuwein, T., Wolf, E., Reik, W., and Dean, W. (2003). Epigenetic marking correlates with developmental potential in cloned bovine preimplantation embryos. *Curr Biol* 13, 1116-1121.
- Sarmiento, O. F., Digilio, L. C., Wang, Y., Perlin, J., Herr, J. C., Allis, C. D., and Coonrod, S. A. (2004). Dynamic alterations of specific histone modifications during early murine development. *J Cell Sci* 117, 4449-4459.
- Sawatsubashi, S., Murata, T., Lim, J., Fujiki, R., Ito, S., Suzuki, E., Tanabe, M., Zhao, Y., Kimura, S., Fujiyama, S., *et al.* (2010). A histone chaperone, DEK, transcriptionally coactivates a nuclear receptor. *Genes Dev* 24, 159-170.

- Schenk, R., Jenke, A., Zilbauer, M., Wirth, S., and Postberg, J. (2011). H3.5 is a novel hominid-specific histone H3 variant that is specifically expressed in the seminiferous tubules of human testes. *Chromosoma*.
- Schoeftner, S., Sengupta, A. K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T., and Wutz, A. (2006). Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *Embo J* 25, 3110-3122.
- Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* 18, 1251-1262.
- Schultz, G. A., and Heyner, S. (1992). Gene expression in pre-implantation mammalian embryos. *Mutat Res* 296, 17-31.
- Schwartz, Y. B., Kahn, T. G., Nix, D. A., Li, X. Y., Bourgon, R., Biggin, M., and Pirrotta, V. (2006). Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*. *Nat Genet* 38, 700-705.
- Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J. R., Wu, C. T., Bender, W., and Kingston, R. E. (1999). Stabilization of chromatin structure by PRC1, a Polycomb complex. *Cell* 98, 37-46.
- Shen, X., Liu, Y., Hsu, Y. J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G. C., and Orkin, S. H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Mol Cell* 32, 491-502.
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119, 941-953.
- Shi, Y. J., Matson, C., Lan, F., Iwase, S., Baba, T., and Shi, Y. (2005). Regulation of LSD1 histone demethylase activity by its associated factors. *Mol Cell* 19, 857-864.
- Shibahara, K., and Stillman, B. (1999). Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. *Cell* 96, 575-585.
- Smale, S. T., and Baltimore, D. (1989). The "initiator" as a transcription control element. *Cell* 57, 103-113.
- Smale, S. T., Schmidt, M. C., Berk, A. J., and Baltimore, D. (1990). Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. *Proc Natl Acad Sci U S A* 87, 4509-4513.
- Sobel, R. E., Cook, R. G., Perry, C. A., Annunziato, A. T., and Allis, C. D. (1995). Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proc Natl Acad Sci U S A* 92, 1237-1241.
- Stewart, M. D., Li, J., and Wong, J. (2005). Relationship between histone H3 lysine 9 methylation, transcription repression, and heterochromatin protein 1 recruitment. *Mol Cell Biol* 25, 2525-2538.
- Strahl, B. D., and Allis, C. D. (2000). The language of covalent histone modifications. *Nature* 403, 41-45.
- Sullivan, K. F., Hechenberger, M., and Masri, K. (1994). Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. *J Cell Biol* 127, 581-592.
- Svejstrup, J. Q. (2004). The RNA polymerase II transcription cycle: cycling through chromatin. *Biochim Biophys Acta* 1677, 64-73.
- Sydow, J. F., Brueckner, F., Cheung, A. C., Damsma, G. E., Dengl, S., Lehmann, E., Vassilyev, D., and Cramer, P. (2009). Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Mol Cell* 34, 710-721.

- Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., and Shinkai, Y. (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev* *16*, 1779-1791.
- Tachiwana, H., Kagawa, W., Osakabe, A., Kawaguchi, K., Shiga, T., Hayashi-Takanaka, Y., Kimura, H., and Kurumizaka, H. (2010). Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. *Proc Natl Acad Sci U S A* *107*, 10454-10459.
- Tachiwana, H., Osakabe, A., Kimura, H., and Kurumizaka, H. (2008). Nucleosome formation with the testis-specific histone H3 variant, H3t, by human nucleosome assembly proteins in vitro. *Nucleic Acids Res* *36*, 2208-2218.
- Tada, T., Tada, M., Hilton, K., Barton, S. C., Sado, T., Takagi, N., and Surani, M. A. (1998). Epigenotype switching of imprintable loci in embryonic germ cells. *Dev Genes Evol* *207*, 551-561.
- Tadros, W., and Lipshitz, H. D. (2009). The maternal-to-zygotic transition: a play in two acts. *Development* *136*, 3033-3042.
- Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* *116*, 51-61.
- Tahiliani, M., Mei, P., Fang, R., Leonor, T., Rutenberg, M., Shimizu, F., Li, J., Rao, A., and Shi, Y. (2007). The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation. *Nature* *447*, 601-605.
- Takada, S., Lis, J. T., Zhou, S., and Tjian, R. (2000). A TRF1:BRF complex directs *Drosophila* RNA polymerase III transcription. *Cell* *101*, 459-469.
- Tan, S., Hunziker, Y., Sargent, D. F., and Richmond, T. J. (1996). Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature* *381*, 127-151.
- Tanaka, Y., Nanba, Y. A., Park, K. A., Mabuchi, T., Suenaga, Y., Shiraishi, S., Shimada, M., Nakadai, T., and Tamura, T. A. (2007). Transcriptional repression of the mouse *wee1* gene by TBP-related factor 2. *Biochem Biophys Res Commun* *352*, 21-28.
- Tanphaichitr, N., Sobhon, P., Taluppeth, N., and Chalermisarachai, P. (1978). Basic nuclear proteins in testicular cells and ejaculated spermatozoa in man. *Exp Cell Res* *117*, 347-356.
- Tarkowski, A. K. (1959). Experiments on the development of isolated blastomers of mouse eggs. *Nature* *184*, 1286-1287.
- Tarkowski, A. K., Ozdzenski, W., and Czolowska, R. (2005). Identical triplets and twins developed from isolated blastomeres of 8- and 16-cell mouse embryos supported with tetraploid blastomeres. *Int J Dev Biol* *49*, 825-832.
- Tarkowski, A. K., and Wroblewska, J. (1967). Development of blastomeres of mouse eggs isolated at the 4- and 8-cell stage. *J Embryol Exp Morphol* *18*, 155-180.
- Tartof, K. D., Hobbs, C., and Jones, M. (1984). A structural basis for variegating position effects. *Cell* *37*, 869-878.
- Teichmann, M., Wang, Z., Martinez, E., Tjernberg, A., Zhang, D., Vollmer, F., Chait, B. T., and Roeder, R. G. (1999). Human TATA-binding protein-related factor-2 (hTRF2) stably associates with hTFIIA in HeLa cells. *Proc Natl Acad Sci U S A* *96*, 13720-13725.
- Thomas, J. O. (1984). The higher order structure of chromatin and histone H1. *J Cell Sci Suppl* *1*, 1-20.
- Torres-Padilla, M. E., Bannister, A. J., Hurd, P. J., Kouzarides, T., and Zernicka-Goetz, M. (2006). Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos. *Int J Dev Biol* *50*, 455-461.

- Torres-Padilla, M. E., and Tora, L. (2007). TBP homologues in embryo transcription: who does what? *EMBO Rep* 8, 1016-1018.
- Tremblay, K. D., Duran, K. L., and Bartolomei, M. S. (1997). A 5' 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development. *Mol Cell Biol* 17, 4322-4329.
- Tsai, M. C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., Shi, Y., Segal, E., and Chang, H. Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 329, 689-693.
- Tumbar, T., Sudlow, G., and Belmont, A. S. (1999). Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain. *J Cell Biol* 145, 1341-1354.
- Turner, B. M. (1993). Decoding the nucleosome. *Cell* 75, 5-8.
- Turner, B. M. (2007). Defining an epigenetic code. *Nat Cell Biol* 9, 2-6.
- Umlauf, D., Goto, Y., Cao, R., Cerqueira, F., Wagschal, A., Zhang, Y., and Feil, R. (2004). Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat Genet* 36, 1296-1300.
- Vakoc, C. R., Mandat, S. A., Olenchok, B. A., and Blobel, G. A. (2005). Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol Cell* 19, 381-391.
- van der Heijden, G. W., Dieker, J. W., Derijck, A. A., Muller, S., Berden, J. H., Braat, D. D., van der Vlag, J., and de Boer, P. (2005). Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech Dev* 122, 1008-1022.
- van der Heijden, G. W., Ramos, L., Baart, E. B., van den Berg, I. M., Derijck, A. A., van der Vlag, J., Martini, E., and de Boer, P. (2008). Sperm-derived histones contribute to zygotic chromatin in humans. *BMC Dev Biol* 8, 34.
- Vassalli, J. D., Huarte, J., Belin, D., Gubler, P., Vassalli, A., O'Connell, M. L., Parton, L. A., Rickles, R. J., and Strickland, S. (1989). Regulated polyadenylation controls mRNA translation during meiotic maturation of mouse oocytes. *Genes Dev* 3, 2163-2171.
- Veenstra, G. J., Weeks, D. L., and Wolffe, A. P. (2000). Distinct roles for TBP and TBP-like factor in early embryonic gene transcription in *Xenopus*. *Science* 290, 2312-2315.
- Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S. I., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* 303, 672-676.
- Vermeulen, M., Mulder, K. W., Denissov, S., Pijnappel, W. W., van Schaik, F. M., Varier, R. A., Baltissen, M. P., Stunnenberg, H. G., Mann, M., and Timmers, H. T. (2007). Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* 131, 58-69.
- Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I., and Martienssen, R. A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* 297, 1833-1837.
- Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R. S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. *Nature* 431, 873-878.
- Ward, W. S. (1993). Deoxyribonucleic acid loop domain tertiary structure in mammalian spermatozoa. *Biol Reprod* 48, 1193-1201.
- Ward, W. S., and Coffey, D. S. (1991). DNA packaging and organization in mammalian spermatozoa: comparison with somatic cells. *Biol Reprod* 44, 569-574.

- Wasylyk, B., and Chambon, P. (1981). A T to A base substitution and small deletions in the conalbumin TATA box drastically decrease specific in vitro transcription. *Nucleic Acids Res* 9, 1813-1824.
- Weil, P. A., Luse, D. S., Segall, J., and Roeder, R. G. (1979). Selective and accurate initiation of transcription at the Ad2 major late promoter in a soluble system dependent on purified RNA polymerase II and DNA. *Cell* 18, 469-484.
- Weintraub, H., and Groudine, M. (1976). Chromosomal subunits in active genes have an altered conformation. *Science* 193, 848-856.
- Whetstine, J. R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., and Shi, Y. (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* 125, 467-481.
- Whitcomb, S. J., Basu, A., Allis, C. D., and Bernstein, E. (2007). Polycomb Group proteins: an evolutionary perspective. *Trends Genet* 23, 494-502.
- White, R. J., Rigby, P. W., and Jackson, S. P. (1992). The TATA-binding protein is a general transcription factor for RNA polymerase III. *J Cell Sci Suppl* 16, 1-7.
- Wiedemann, S. M., Mildner, S. N., Bonisch, C., Israel, L., Maiser, A., Matheisl, S., Straub, T., Merkl, R., Leonhardt, H., Kremmer, E., *et al.* (2010). Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y. *J Cell Biol* 190, 777-791.
- Wilsker, D., Patsialou, A., Dallas, P. B., and Moran, E. (2002). ARID proteins: a diverse family of DNA binding proteins implicated in the control of cell growth, differentiation, and development. *Cell Growth Differ* 13, 95-106.
- Witt, O., Albig, W., and Doenecke, D. (1996). Testis-specific expression of a novel human H3 histone gene. *Exp Cell Res* 229, 301-306.
- Wong, L. H., Ren, H., Williams, E., McGhie, J., Ahn, S., Sim, M., Tam, A., Earle, E., Anderson, M. A., Mann, J., and Choo, K. H. (2009). Histone H3.3 incorporation provides a unique and functionally essential telomeric chromatin in embryonic stem cells. *Genome Res* 19, 404-414.
- Woo, C. J., Kharchenko, P. V., Daheron, L., Park, P. J., and Kingston, R. E. (2010). A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell* 140, 99-110.
- Woodcock, C. L., Frado, L. L., and Rattner, J. B. (1984). The higher-order structure of chromatin: evidence for a helical ribbon arrangement. *J Cell Biol* 99, 42-52.
- Woodcock, C. L., Skoultchi, A. I., and Fan, Y. (2006). Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. *Chromosome Res* 14, 17-25.
- Workman, J. L., and Roeder, R. G. (1987). Binding of transcription factor TFIID to the major late promoter during in vitro nucleosome assembly potentiates subsequent initiation by RNA polymerase II. *Cell* 51, 613-622.
- Wright, K. J., Marr, M. T., 2nd, and Tjian, R. (2006). TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. *Proc Natl Acad Sci U S A* 103, 12347-12352.
- Wu, C. (1989). Analysis of hypersensitive sites in chromatin. *Methods Enzymol* 170, 269-289.
- Xiao, L., Kim, M., and DeJong, J. (2006). Developmental and cell type-specific regulation of core promoter transcription factors in germ cells of frogs and mice. *Gene Expr Patterns* 6, 409-419.
- Yamane, K., Toumazou, C., Tsukada, Y., Erdjument-Bromage, H., Tempst, P., Wong, J., and Zhang, Y. (2006). JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell* 125, 483-495.

- Zegerman, P., Canas, B., Pappin, D., and Kouzarides, T. (2002). Histone H3 lysine 4 methylation disrupts binding of nucleosome remodeling and deacetylase (NuRD) repressor complex. *J Biol Chem* 277, 11621-11624.
- Zhang, D., Penttila, T. L., Morris, P. L., Teichmann, M., and Roeder, R. G. (2001). Spermiogenesis deficiency in mice lacking the Trf2 gene. *Science* 292, 1153-1155.
- Zhang, K., Mosch, K., Fischle, W., and Grewal, S. I. (2008). Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nat Struct Mol Biol* 15, 381-388.
- Zhang, X., Henderson, I. R., Lu, C., Green, P. J., and Jacobsen, S. E. (2007). Role of RNA polymerase IV in plant small RNA metabolism. *Proc Natl Acad Sci U S A* 104, 4536-4541.
- Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J., and Lee, J. T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 322, 750-756.
- Zhao, M., Shirley, C. R., Hayashi, S., Marcon, L., Mohapatra, B., Suganuma, R., Behringer, R. R., Boissonneault, G., Yanagimachi, R., and Meistrich, M. L. (2004). Transition nuclear proteins are required for normal chromatin condensation and functional sperm development. *Genesis* 38, 200-213.
- Zhou, W., Zhu, P., Wang, J., Pascual, G., Ohgi, K. A., Lozach, J., Glass, C. K., and Rosenfeld, M. G. (2008). Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. *Mol Cell* 29, 69-80.
- Zuccotti, M., Garagna, S., Merico, V., Monti, M., and Alberto Redi, C. (2005). Chromatin organisation and nuclear architecture in growing mouse oocytes. *Mol Cell Endocrinol* 234, 11-17.
- Zweidler, A. (1984). Core histone variants of the mouse : Primary structure and differential expression, In *Histone Genes*, J. L. S. G. S. Stein, W. F. Marzuff, ed., pp. 339.