

Table des Matières

PRÉAMBULE.....	1
CHAPITRE I : INTRODUCTION.....	4
I. <u>La compensation de dose</u>.....	5
a. <u>Compensation de dose entre les chromosomes sexuels</u>	6
b. <u>Compensation de dose avec les autosomes</u>	7
II. <u>L'inactivation du chromosomes X chez <i>Mus musculus</i></u>.....	8
a. <u>Inactivation méiotique des chromosomes sexuels</u>	8
b. <u>L'inactivation empreintée suite à la fécondation</u>	10
c. <u>Réactivation du chromosome X dans les blastocystes</u>	12
d. <u>L'inactivation aléatoire du chromosomes X</u>	13
i. Le comptage.....	14
ii. Le choix.....	15
e. <u>Réactivation du Xi dans la lignée germinale</u>	17
III. <u>Les Modèles Cellulaires</u>.....	19
IV. <u>Les caractéristiques du X inactif</u>.....	24
a. <u>L'ARN non-codant <i>Xist</i></u>	25
i. Propagation de <i>Xist</i> et l'extinction transcriptionnelle.....	26
ii. <i>Xist</i> au fil du temps.....	29
b. <u>Suite de l'extinction transcriptionnelle</u>	33
c. <u>Les modifications d'histones</u>	33
i. Généralités.....	33
ii. Modifications d'histones sur le Xi.....	36
d. <u>Autres modifications épigénétiques liées à l'Xi</u>	39
i. MacroH2A.....	39
ii. Méthylation de l'ADN.....	40
iii. Divers.....	44

iv.	Réplication asynchrone.....	44
v.	Conclusions sur les différentes marques épigénétiques.....	45
 V. Les principaux acteurs de l'inactivation du chromosome X : <i>Xist</i> et son antisens <i>Tsix</i>.....		
a.	<u>Le <i>XIC</i></u>	47
i.	Définition Cytogénétique et transgénèse.....	48
ii.	Unités fonctionnelles du <i>XIC</i>	49
iii.	"Point chaud" de modifications épigénétiques en 5' de <i>Xist</i>	50
b.	<u><i>Xist</i>, <i>Tsix</i> : généralités</u>	51
i.	<i>Xist</i>	51
ii.	<i>Tsix</i>	56
iii.	<i>DxPas34</i>	59
iv.	<i>Xite</i>	63
 VI. Régulation de <i>Xist</i>.....		
 Revue : Andrew Oldfield et Claire Rougeulle, <u>X-plications sur le contrôle de l'inactivation : la régulation du gène <i>Xist</i></u>. Biotur, n°304, Novembre 2009.....		
a.	<u>Régulation de <i>Xist</i> par <i>Tsix</i></u>	72
i.	Régulation post-transcriptionnelle.....	72
ii.	ARNi.....	74
iii.	Modification du paysage épigénétique au promoteur de <i>Xist</i>	78
iv.	Modification du paysage épigénétique le long du locus <i>Xist/Tsix</i>	83
v.	Effet de <i>Xist</i> sur <i>Tsix</i>	87
b.	<u>Couplage entre la régulation de l'expression de <i>Xist</i> et la pluripotence</u>	88
c.	<u>Régulation en trans de <i>Xist</i> par divers facteurs</u>	92
i.	Les répresseurs.....	92
ii.	Les activateurs.....	94
 CHAPITRE II : PROJET DE THÈSE.....		
		101

CHAPITRE III : RÉSULTATS ET DISCUSSIONS.....	104
I. L'étude du réseau transcriptionnel régulant l'expression de <i>Xist</i> chez la souris et sa conservation chez l'homme.....	105
a. <u>CTCF</u>	106
i. Fonction activatrice/répressive.....	108
ii. Fonction d'insulateur.....	109
iii. Rôle dans la méthylation de l'ADN.....	110
iv. Formation de structures tridimensionnelles.....	112
b. <u>YY1</u>	114
i. Protéines PcG.....	117
ii. Formations de structures tridimensionnelles.....	118
iii. Rôle dans la méthylation de l'ADN.....	119
c. <u>Résultats</u>	119
i. Fixation de CTCF le long du locus <i>Xist/Tsix</i>	119
ii. Fixation de YY1 en 5' de <i>Xist</i>	123
iii. Corrélation entre fixation de CTCF et YY1 et surexpression de <i>Xist</i> en cours de différenciation.....	125
iv. Un rôle pour les autres pics de CTCF ?.....	128
v. Conservation des fixations de CTCF, YY1 et Rad21 chez l'homme...134	
d. <u>Discussion</u>	138
i. Nos données sont corrélées par la littérature.....	139
ii. Situation dans les cellules ES non-différenciées.....	142
iii. Différents mutants de <i>Tsix</i> ont différents effets sur l'expression de <i>Xist</i>	144
iv. Interaction entre CTCF & YY1 ?.....	145
v. De quelle manière l'activation de <i>Xist</i> est-elle obtenue ?.....	146
vi. CTCF et YY1 impliquées dans la méthylation de l'ADN du promoteur de <i>Xist</i> ?.....	147
vii. Recrutement de Rad21.....	149
viii. Interactions 3D ?.....	150
ix. Un rôle dans le <i>kissing</i> ?.....	153
x. Un rôle pour le pic CTCF/Cohésine situé en amont de <i>Tsix</i>	154
xi. Conservation.....	159
xii. Conclusion.....	159

II. La caractérisation fonctionnelle des sites de fixation pour les protéines CTCF et YY1 dans la région promotrice de *Xist* chez la souris

.....	162
a. <u>Introduction</u>	162
b. <u>Le vecteur</u>	166
c. <u>L'électroporation</u>	168
d. <u>Le criblage</u>	169
e. <u>Le bilan</u>	170
f. <u>Perspectives</u>	172

III. L'étude du réseau transcriptionnel régulant *Xist* au sein de cellules germinales embryonnaires de souris

a. <u>Introduction</u>	175
b. <u>Résultats</u>	176
i. Expression de <i>Xist</i> , <i>Tsix</i> , Oct4 et Nanog dans les EGs.....	190
ii. Mécanisme de répression de <i>Xist</i> dans les EGs.....	177
iii. Fixation de CTCF et YY1 le long de <i>Xist/Tsix</i>	179
c. <u>Discussion</u>	179

IV. Le couplage moléculaire de la régulation de *Tsix* et de la pluripotence cellulaire

a. <u>Introduction</u>	184
------------------------------	-----

Article : P. Navarro, Oldfield A., Legoupi J., Festucia N., Dubois A., Attia M., Rougeulle C., Chambers I., and Philip Avner. Molecular coupling of *Tsix* regulation and pluripotency. Accepté dans Nature.....186

b. <u>Discussion</u>	223
----------------------------	-----

CHAPITRE IV : ÉPILOGUE

Bibliographie	233
----------------------------	-----

Annexes	267
----------------------	-----

Index des figures

- Figure 1** : Différents mécanismes de compensation de dose. **p6**
- Figure 2** : Le cycle d'inactivation et de réactivation du chromosome X au cours du développement embryonnaire. **p9**
- Figure 3** : Représentation schématique de la localisation génique par rapport au compartiment hétérochromatique (principalement formé par les séquences répétées) lors de la mise en place de l'inactivation empreintée. **p11**
- Figure 4** : Expression de *Xist*/*Tsix* dans les cellules ES. **p14**
- Figure 5** : Le « *pairing* » au sein de cellules ES femelles. **p17**
- Figure 6** : Développement des cellules germinales primordiales au sein de l'embryon murin. **p18**
- Figure 7** : L'origine des cellules souches dans l'embryon mammifère. **p19**
- Figure 8** : Modèles cellulaires ex-vivo et inactivation du chromosome X. **p21**
- Figure 9** : Les séquences répétées de *Xist*. **p26**
- Figure 10** : Modèle du comportement des LINEs et de leur actions au cours de l'inactivation du chromosome X. **p28**
- Figure 11** : Mise en évidence de l'existence d'une mémoire chromosomique de l'état inactif induite par *Xist*. **p30**
- Figure 12** : Modèle proposant que la localisation périnucléolaire est requise pour la maintenance de l'état réprimé du chromosome X inactif. **p32**
- Figure 13** : Cinétique de l'inactivation aléatoire du chromosome X. **p34**
- Figure 14** : Marques épigénétiques associées à l'inactivation du chromosome X. **p37**
- Figure 18** : Tableau récapitulatif du taux de réactivation du Xi lors de diverses mutations et traitements chimiques. **p46**
- Figure 19** : Éléments fonctionnels au sein du centre d'inactivation du chromosome X. **p50**
- Figure 20** : La structure du gène *Xist* et sa conservation chez l'homme. **p53**
- Figure 21** : Comparaison de l'agencement des gènes *XIST/Xist* et *TSIX/Tsix* et de leur îlots CpG entre homme et souris. **p58**

Figure 22 : Schéma descriptif des éléments régulateurs de <i>Tsix</i> .	p59
Figure 23 : Schéma de l'analyse transcriptionnelle de cellules ES mâles tronquées ou pas pour <i>Tsix</i> en cours de différenciation.	p73
Figure 24 : Remodelage chromatinien de la région promotrice de <i>Xist</i> , comprise entre deux sites de fixation de CTCF, par <i>Tsix</i> .	p79
Figure 25 : Modèle de la régulation de la région <i>Xist/Tsix</i> au cours du développement embryonnaire précoce murin.	p84
Figure 26 : Modèle d'action de Rnf12 dans l'initiation de l'inactivation du chromosome X.	p98
Figure 27 : La structure de CTCF, ainsi que ses modifications et partenaires connus.	p107
Figure 28 : Les deux rôles possibles pour CTCF en tant qu'isolateur.	p109
Figure 29 : CTCF organise des contacts chromatiniens aux loci empreintés.	p111
Figure 30 : Schéma modélisant la structure du complexe de la cohésine.	p113
Figure 31 : Diagramme schématique de YY1 et de ses partenaires.	p116
Figure 32 : Tableau récapitulatif des différents interacteurs de YY1 et le but de leur coopération.	p117
Figure 33 : Récapitulatif des sites de fixation de CTCF et de YY1 le long du locus <i>Xist/Tsix</i> déjà publiés.	p120
Figure 34 : Distribution de CTCF dans une région de 50kb du XIC murin au sein de cellules ES et fibroblastiques femelles.	p121
Figure 35 : Étude de la fixation de CTCF le long du locus <i>Xist/Tsix</i> dans des cellules ES et fibroblastiques mâles.	p123
Figure 36 : Étude de la fixation de YY1 dans la région promotrice de <i>Xist</i> dans des cellules ES et somatiques mâles et femelles.	p125
Figure 37 : Schéma Récapitulatif des différentes mutations génétiques présentes au sein de cellules mâles qui ont servi à l'étude.	p126
Figure 38 : Expression de <i>Xist</i> au sein des différentes lignées cellulaires étudiées.	p127

Figure 39 : Fixation de CTCF et de YY1 à la région promotrice sur le Xa ou le Xi en cours de différenciation. p129

Figure 40 : Étude de la fixation de Rad21 le long du locus *Xist/Tsix* dans des cellules ES et fibroblastiques femelles. p130

Figure 41 : Étude de la fixation de Rad21 le long du locus *Xist/Tsix* dans des cellules ES et fibroblastiques mâles. p131

Figure 42 : Fixation de Rad21 le long du locus *Xist/Tsix* sur le Xa ou le Xi en cours de différenciation. p133

Figure 43 : Structure et conservation du gène CTCF. p134

Figure 44 : Fixation de CTCF, YY1 et Rad21 au sein de la région promotrice du *XIST* humain dans des cellules fibroblastiques mâles et femelles. p135

Figure 45 : Conservation des sites consensus de fixation pour CTCF, YY1 et Rad21 dans la région promotrice de *Xist* chez les euthériens. p138

Figure 46 : Les sites d'hypersensibilité à la DNase le long du locus *Xist/Tsix*. p141

Figure 47 : Activité transcriptionnelle d'un groupe de sites YY1 au promoteur P2 de *Xist*. p142

Figure 48 : Fixation de multiples facteurs de transcription au promoteur de Rnf12 dans les cellules ES de souris. p144

Figure 49 : Analyse de la méthylation de l'ADN au sein de la région promotrice dans Ma1L et Ma2L en cours de différenciation. p148

Figure 50 : Interactions tridimensionnelles d'un fragment du XIC avec la région promotrice de *Xist*. p151

Figure 51 : Modèle de la fixation cycle-cellulaire dépendante de la cohésine à la chromatine. p155

Figure 52 : Les insulateurs comme organisateurs de domaines de chromatine indépendants au sein du noyau. p156

Figure 53 : Corrélation entre la présence de séquences potentielles de MARs et la fixation de CTCF au locus *Xist/Tsix*. p157

Figure 54 : L'implication de SatB1 la régulation épigénétique des gènes. p158

Figure 55 : Modèle schématique représentant les différents phénomènes au sein du locus *Xist/Tsix* influant sur la régulation de *Xist*. p160

Figure 56 : Effet d'une réduction partielle de l'expression de *Tsix* sur celle de *Xist*. p164

- Figure 57** : Stratégie générale de *Recombinase-Mediated Cassette Exchange*. **p165**
- Figure 58** : Schéma du vecteur servant à introduire les sites de recombinaison au locus endogène. **p167**
- Figure 59** : Schéma des régions du locus *Xist/Tsix* impliquées dans la construction. **p168**
- Figure 60** : Tableau récapitulant les différents protocoles d'électroporation utilisés. **p168**
- Figure 61** : Schéma du protocole utilisé pour générer des mutations spécifiques au sein des sites de fixation des protéines d'intérêt. **p173**
- Figure 62** : Statut transcriptionnel de plusieurs gènes clés à notre sujet d'étude au sein de cellules EGs. **p177**
- Figure 63** : Fixation des facteurs de pluripotence Oct4, Nanog et Sox2 au sein de l'intron 1 et du minisatellite DxPas34 dans les EGs femelles. **p178**
- Figure 64** : Fixation de CTCF et YY1 le long du locus *Xist/Tsix* au sein de cellules EGs. **p180**
- Figure 65** : Moyennes des variances de l'expression log-transformée de gènes d'intérêt à nos travaux dans différentes lignées de cellules EGs et ES. **p181**
- Figure 66** : Réseau d'interactions entre différents facteurs de transcription étant, ou pouvant être, impliqués dans la régulation de l'expression de *Tsix* au sein des cellules ES indifférenciées. **p225**

Liste des Abréviations

ADN: acide désoxyribonucléique

ADNc:ADN complémentaire

ARN: acide ribonucléique

ARNi: voie d'interférence à ARN

CBP/p300: protéine de liaison à CREB, enzyme d'acétylation des histones

ChIP: immunoprécipitation de chromatine

CpG: dinucléotide cytosine-guanine

CTCF: facteur de transcription fixant les motifs CCCTC

Dicer: enzyme d'hydrolyse des ARNs double-brin

DNAse: enzyme d'hydrolyse de l'ADN

Dnmt: enzyme de méthylation des CpG

DXPas34: mini-satellite constitué d'un motif de 34 paires de bases répété, localisé sur la bande D du chromosome X, séquencé à l'Institut Pasteur

Eed: protéine de développement de l'ectoderme embryonnaire

ES: cellules embryonnaires souches

FISH: "fluorescence in-situ hybridization"

G2: phase Gap2 du cycle cellulaire

H3K14: Lysine 14 de l'histone H3

H3K27: Lysine 27 de l'histone H3

H3K4: Lysine 4 de l'histone H3

H3K9: Lysine 9 de l'histone H3

H4K20: Lysine 20 de l'histone H4

H4R3: Arginine 3 de l'histone H8

HAT: enzyme d'acétylation des histones

HDAC: enzyme de désacétylation des histones

HP1: protéine d'hétérochromatine 1

ICM: masse cellulaire interne du blastocyste

Igf2: facteur de croissance dépendant de l'insuline

ISWI: "imitation of SWI" complexe de remodelage ATP-dépendant

kb: kilobase

Kcnq1: gène codant un canal au potassium voltage dépendant

Kcnq1ot1: antisens de Kcnq1

kDa: kilodalton

LINE: "long interspersed nuclear element"

Mb: mégabase

MBD: protéine à domaine de liaison aux CpG méthylés

MeCP2: protéine de liaison au CpG méthylés
MEF: fibroblaste embryonnaire de souris
miRNA: micro ARN
PCAF: facteur associé à CBP et p300
PCR: réaction en chaîne de la polymérase
PHD: domaine protéique découvert chez des protéines homéotiques des plantes
PIC: complexe de pré-initiation de la transcription
PRC2: complexe polycomb 2
PRC1: complexe polycomb 1
PRMT: enzyme de méthylation des arginines
Q-PCR: PCR quantitative
RISC: "RNA induced silencing complex"
RITS: "RNA induced transcriptional silencing"
RT: transcription réverse
SET: domaine catalytique de méthylation des histones retrouvé dans les protéines Suvar, Enhancer of Zeste, et Trithorax
siRNA: "short interfering RNA"
STOP: signal de terminaison de la transcription
SuVar: supresseur de variéation, enzyme de méthylation de H3K9
SWI/SNF: "mating type Switching-Sucrose Non-Fertile"
TBP: "TATA-binding proteine"
TFII: facteur de transcription basal de l'ARN Polymérase II
TS: cellules souches du trophoctoderme
TSA: Trichostatine A
Tsix: antisens à Xist
Xa: chromosome X actif
XEN: cellules souches de l'endoderme extra-embryonnaire
Xi: chromosome X inactif
Xic: centre d'inactivation du chromosome X
Xist: "X-inactif specific transcript"
Xite: X intergenic transcription element
Xm: chromosome X maternel
XO: cellule femelle ayant perdu un chromosome X
Xp: chromosome X paternel
XX: cellule femelle
XY: cellule mâle
YY1: facteur de transcription Yin-Yang

Préambule

La transcription des gènes, essentielle à toute forme de vie, est le mécanisme à l'origine de l'identité de l'être vivant. L'expression d'un génome devient possible en rendant le code génétique (l'ADN) de l'organisme accessible aux facteurs chargés de la fabrication des unités constitutives de l'individu (les protéines). Ceci est réalisé en transcrivant l'ADN en ARN messenger qui, de par sa petite taille, peut franchir l'enveloppe nucléaire en passant par les pores nucléaires (chose impossible pour les chromosomes, trop volumineux) et ainsi permettre la traduction de ce message en macromolécules protéiques qui rempliront une infinité de tâches au sein de la cellule.

À travers la multitude de découvertes engendrées quotidiennement par la communauté scientifique, la complexité du phénomène de transcription s'avère être exponentielle, ses causes, ses conséquences, ses acteurs et ses ramifications s'accroissant à chaque instant. Depuis maintenant quelque temps, en étudiant la diversité cellulaire, on s'aperçoit que chaque cellule a un rôle et une identité qui lui est propre et que cette identité est gérée par son transcriptome (ensemble de transcrits présents à un moment donné au sein de la cellule). Ce transcriptome peut varier fortement d'un type cellulaire à un autre, différents gènes étant exprimés à différents instants et à différentes intensités afin de remplir diverses fonctions ; les cellules ne pouvant exprimer continuellement l'ensemble de leurs gènes à tout moment et prétendre avoir une panoplie de fonctions bien distinctes les unes des autres.

Il y a maintenant près de 50 ans, François Jacob et Jacques Monod ont décrit le premier exemple de régulation transcriptionnelle chez *Escherichia coli* (Jacob & Monod, 1961). En révélant au grand jour le fonctionnement de l'opéron lactose, ils ont permis la description du premier facteur de transcription, celui-ci se trouvant être un répresseur transcriptionnel. Par la suite il aura fallu attendre une dizaine d'années avant de se rendre compte de l'existence d'activateurs (Riggs *et al.*, 1971), les gènes n'étant exprimés qu'à un niveau basal en leur absence.

De nos jours, nos connaissances en matière de régulation transcriptionnelle se sont beaucoup approfondies, et cela nous a permis de nous rendre compte que l'orchestration d'un transcriptome se doit d'être quelque chose d'extrêmement complexe, dynamique, et de très finement régulé afin de pouvoir répondre aux besoins en matière de diversité cellulaire de l'organisme ainsi qu'entre les espèces. Au cours de l'évolution, la cellule a donc mis en place

bon nombre de processus lui permettant d'obtenir une identité qui lui est propre, en régulant quantitativement ainsi que temporellement l'expression de sa panoplie génique.

Sous la tutelle de signaux extérieurs indiquant à la cellule vers quelle voie elle doit s'engager, la machinerie basale de transcription (dont les sous-unités varient de manière exhaustive), l'organisation des facteurs de transcription, le contexte chromatinien péri-génique, ainsi que plusieurs modes de régulation post-transcriptionnelle, jouent tous des rôles prépondérants dans la mise en place d'une véritable identité transcriptionnelle de la cellule.

Les questions qui se posent alors sont : comment la cellule arrive-t-elle à gérer l'ensemble de son transcriptome ? Qu'est-ce qui fait qu'un gène, parmi des dizaines de milliers, va être exprimé à un temps et à une intensité donnés et qui lui sont propres ?

D'après des estimations récentes (Messina *et al.*, 2004 ; Babu & Teichmann, 2004), il existerait environ de 2000 à 3000 facteurs de transcription chez l'homme, chacun ayant un rôle particulier dans un contexte particulier. De plus, la quasi-totalité d'entre eux se lient à des centaines voir à des dizaines de milliers de sites à travers le génome et sont connus pour interagir avec bon nombre d'autres protéines impliquées dans la régulation transcriptionnelle ou bien le paysage chromatinien. En fonction de son site de reconnaissance à l'ADN, du contexte chromatinien ainsi que de ses partenaires à un site donné, un facteur donné peut avoir des effets radicalement opposés sur la transcription d'un gène (par exemple : Carroll *et al.*, 2006). Ceci laisse entrevoir le fait que la régulation d'un simple gène est quelque chose d'extrêmement complexe, qui dépend d'un ensemble de conditions, mais qui semble pourtant être orchestré de main de maître par la cellule.

Ces dernières années, l'avènement de techniques de laboratoire adaptées à l'étude de la fixation de facteurs de transcription, comme l'ImmunoPrécipitation de Chromatine (ChIP), la ChIP couplée à une analyse sur puce à ADN (ChIP-CHIP) ou bien encore la ChIP suivie d'un séquençage des produits immunoprécipités (ChIP-Seq) (voir les Annexes 1 et 2, p268-269), utilisées en parallèle de techniques de biologie moléculaire classiques, nous facilitent grandement la tâche quant à la découverte de leurs modes de fonctionnement, tout en complexifiant l'interprétation de leur fixation à l'échelle d'un génome entier.

Au cours de mon doctorat, je me suis tout particulièrement intéressé à la manière dont est régulé un gène non-codant et son anti-sens : *Xist* et *Tsix*, respectivement, et ce, principalement dans le contexte de cellules embryonnaires murines. Ces deux gènes sont impliqués dans

l'inactivation du chromosome X (la forme de compensation de dose présente chez les mammifères placentaires).

En employant majoritairement la méthode de ChIP suivie d'une analyse par PCR-quantitative (Réaction en Chaîne de la Polymérase-quantitative) couplée à une approche de biologie moléculaire classique, j'ai pu établir un modèle simplifié sur la manière dont l'activation transcriptionnelle de ce locus (*Xist/Tsix*) se met en place au sein de différents contextes cellulaires.

Avant de détailler mes résultats, dans un premier chapitre introductif je décrirai ce qu'est la compensation de dose, en me focalisant sur ce phénomène chez les mammifères placentaires. Je m'attacherai ensuite à expliquer les différents types d'inactivation du chromosome X chez la souris. Puis nous verrons les particularités des différents composants génétiques du centre d'inactivation (région située sur le chromosome X responsable de l'inactivation) suivies des différents modes de régulation de son acteur principal : le gène *Xist*.

Chapitre I : Introduction

I. La compensation de dose

Comme nous venons de le voir en préambule, l'intensité de la transcription ainsi que le moment où celle-ci se déroule sont primordiaux pour obtenir le transcriptome souhaité, ceci étant vrai quel que soit le sexe de l'individu. Or, chez les métazoaires, la différence entre les sexes se fait à travers la présence de gonosomes hétéromorphes soit chez les mâles (par exemple : XY chez l'homme) soit chez les femelles (ZW chez les oiseaux), le sexe opposé ayant deux gonosomes homologues (XX chez la femme, et ZZ chez l'oiseau mâle). Le fait qu'une paire de chromosomes soit différente entre les sexes engendre des complications relatives au transcriptome : certains gènes devraient alors se trouver deux fois plus exprimés chez les individus du sexe possédant des chromosomes sexuels homologues. De plus, comme nous le verrons par la suite, il est aussi nécessaire d'avoir un équilibre transcriptionnel entre gonosomes et autosomes.

À travers l'évolution, l'ensemble des métazoaires a donc trouvé diverses manières de réguler cette différence du nombre de transcrits engendrée par les différences de sexe (Figure 1), ces régulations transcriptionnelles au niveau des chromosomes sexuels sont regroupées sous le terme de compensation de dose.

a. Compensation de dose entre les chromosomes sexuels

Chez les mammifères la détermination sexuelle est indépendante du processus de compensation de dose, le choix du sexe se déroulant au moment de la fécondation : à savoir si le spermatozoïde qui participe à la création du fœtus porte un chromosome X ou Y.

Les euthériens ont un système de compensation de dose, qui diverge des autres métazoaires, étant donné que parmi les deux chromosomes X, l'un d'entre eux (le choix est fait de manière théoriquement aléatoire) va être inactivé autour de 5jpc (Jours *Post-Coïtum*) dans l'embryon. Cette inactivation est principalement médiée par un gène se trouvant lui-même sur le chromosome X : le gène *XIST/Xist* (« *X-Inactive Specific Transcript* »). *XIST/Xist* est un transcrit non-codant dont l'expression est spécifique au Xi (chromosome X Inactif) et qui va recouvrir la quasi-intégralité de ce chromosome en *cis* (Brown *et al.*, 1992), induisant son extinction transcriptionnelle quasi-totale. Cette inactivation est exercée dans toutes les

cellules, donnant lieu à un individu mosaïque quant à l'origine du chromosome X inactivé et elle est maintenue tout au long de la vie de l'individu.

Afin de savoir combien de chromosomes elle doit inactiver, la cellule effectue donc un comptage, comme chez les autres espèces, afin d'établir un ratio X:A. Ainsi, la cellule ne maintient actif qu'un chromosome X par lot d'autosomes (Lyon, 1962 ; Ohno *et al.*, 1964) ce phénomène étant particulièrement observable dans le cas des maladies génétiques comme le syndrome de Klinefelter où mâles (XXY) et femelles (XXX) ne gardent qu'un seul chromosome X actif.

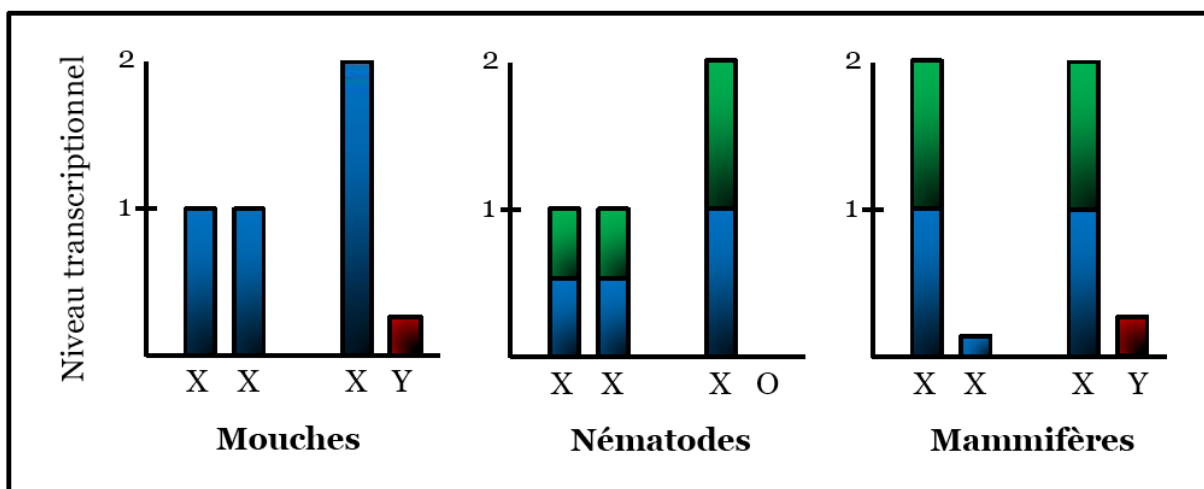


Figure 1 : Différents mécanismes de compensation de dose. Chez *Drosophila melanogaster*, l'expression génique provenant du chromosome X est doublée chez le mâle (XY) (pour une revue, voir Gelbart & Kuroda, 2009 ; Ilik & Akhtar, 2009). Chez *Caenorhabditis elegans*, la transcription est diminuée d'un facteur 2 à partir des deux chromosomes X de l'hermaphrodite (pour une revue, voir Meyer, 2005). Chez les mammifères, un des deux chromosomes X est inactivé chez les femelles. En bleu l'expression due à la compensation de dose gonosomale, en vert l'expression après prise en compte de la compensation de dose entre autosome et gonosome, et en rouge l'expression du chromosome Y.

Le phénomène de compensation de dose chez les mammifères femelles passe donc par l'inactivation transcriptionnelle d'un des deux chromosomes X. Cependant, cette inactivation n'est pas totale et la quantité de gènes échappant à celle-ci peut varier d'une espèce à une autre. Ainsi, chez les souris, il a été estimé qu'environ 3% des gènes du chromosome X échappaient à l'inactivation (Yang *et al.*, 2010), tandis que chez l'homme ce phénomène serait beaucoup plus répandu en atteignant 15% auxquels viendraient s'ajouter 10% de gènes dont l'état actif/inactif varierait en fonction du tissu ou de l'individu étudié (Carrel & Willard, 2005) (pour revue voir Prothero *et al.*, 2009, Heard & Disteche, 2006, Disteche *et al.*, 2002).

b. Compensation de dose avec les autosomes

Comme nous venons de le voir, le phénomène de compensation de dose a donc comme but de réduire les inégalités en terme de quantité de transcrits dues à la présence d'un système à gonosomes hétéromorphes. Bien que l'objectif principal soit atteint, la mise en place de ce mécanisme engendre l'apparition d'un nouveau déséquilibre transcriptionnel : une différence de la quantité de transcrits entre les gonosomes et les autosomes chez les mammifères et les nématodes, les rendant donc fonctionnellement aneuploïdes par rapport aux autosomes (ratio X:AA ou XX:AA de 0,5).

Pour pallier à cette différence transcriptionnelle, induite soit par l'hémizygotie (chez les mâles) soit par la compensation de dose (chez les femelles), les espèces abaissant leur niveau de transcrits chez les femelles compensent en sur-exprimant le chromosome X chez les mâles, (le chromosome X actif chez les mammifères ou bien les deux chromosomes X partiellement réprimés chez le nématode) d'un facteur deux, afin que les gonosomes atteignent dans tous les cas le niveau transcriptionnel d'une paire d'autosomes (soit un ratio X:AA et XX:AA égal à 1) (Gupta *et al.*, 2006 ; Nguyen & Distèche, 2006). Bien que complexifiant considérablement l'idée que l'on se fait de la compensation de dose, cette adaptation par rapport aux autosomes est indispensable pour la survie de l'individu, l'haploïdie pour seulement 1% du génome réduisant considérablement la viabilité et 3% induisant la léthalité de l'organisme (Lindsley *et al.*, 1972).

Le mode d'action de cette sur-expression est encore inconnu, mais il pourrait s'agir d'un mécanisme actif impliquant des modifications épigénétiques similaires au système présent chez la mouche, et pourrait être impliqué soit dans l'accessibilité des promoteurs pour l'ARN polymérase, soit dans l'élongation des transcrits soumis à compensation de dose (Smith *et al.*, 2001).

L'inactivation d'un des deux chromosomes X au sein de toute cellule femelle ainsi que la sur-activation transcriptionnelle du X actif (aussi bien chez les femelles que chez les mâles) est donc la manière dont les euthériens parviennent à compenser leur nombre de transcrits émanant des gonosomes mais aussi par rapport au reste des chromosomes non-sexuels.

Les mécanismes gouvernant l'inactivation du chromosome X sont très complexes, mais je vais tout de même tenter de donner au lecteur une appréciation globale du fonctionnement de ce paradigme épigénétique à travers ces prochains chapitres. Dans ce dessein nous nous

focaliserons sur le modèle murin, espèce, chez qui ce phénomène a été le plus étudié et le mieux décrit bien qu'il y existe encore de très nombreuses zones d'ombres.

II. L'Inactivation du chromosome X chez *Mus musculus*

Comme nous venons de l'entreapercevoir, l'inactivation du chromosome X est le fruit de son extinction transcriptionnelle quasi-globale lors des phases précoces du développement embryonnaire afin d'atteindre un équilibre transcriptionnel entre les deux sexes.

Cependant chez la souris (*Mus musculus*), cette inactivation existe sous plusieurs formes qui s'établissent au cours de vagues successives et de manières différentes en fonction du type cellulaire étudié. Dans un premier temps, nous allons donc voir de quelle manière l'inactivation est établie à différents stades cellulaires permettant ainsi d'avoir un aperçu préliminaire de l'importance du développement au cours de ce processus particulier (Figure 2). Ensuite, nous nous attacherons à définir la façon dont la cellule femelle arrive à compter puis à choisir le chromosome X à inactiver, ainsi que les caractéristiques épigénétiques de cet X inactif. Et enfin, nous nous pencherons dans le détail sur les deux ARNs non-codants au centre de la mise en place de l'inactivation : *Xist* et son anti-sens *Tsix*.

a. Inactivation méiotique des chromosomes sexuels

Le premier type d'inactivation des chromosomes sexuels a lieu avant même la formation du zygote : dans les spermatocytes (pour une revue voir Turner, 2007). Les chromosomes sexuels chez le mâle ne possèdent que très peu de régions homologues (les régions pseudo-autosomales) ce qui empêche un appariement correct de ceux-ci lors de la méiose ; la cellule intervient alors en inactivant les chromosomes sexuels par un processus nommé MSCI (*meiotic sex-chromosome inactivation*) (voir cadre bleuté de la Figure 2 ; Lifschytz et Lindsley, 1972; McCarrey & Dilworth, 1992; Monesi, 1965) formant ainsi ce qui est couramment qualifié de « vésicule sexuelle » (*sex-body*) (Handel, 2004). Ce phénomène a donc comme objectif d'empêcher les recombinaisons inappropriées ou délétères au sein des régions chromosomiques non-appariées tout en les stabilisant.

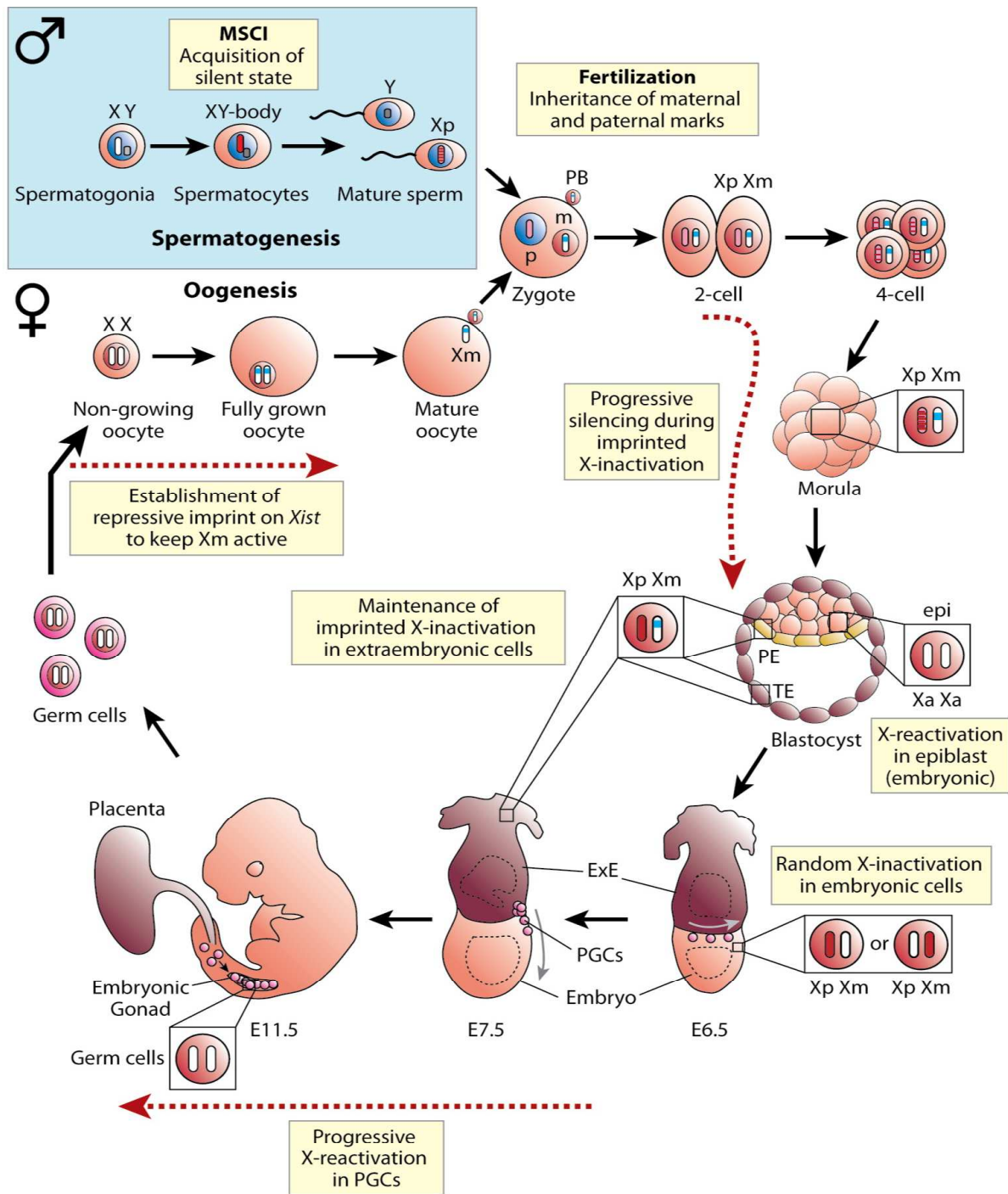


Figure 2 : Le cycle d'inactivation et de réactivation du chromosome X au cours du développement embryonnaire. Ce diagramme montre les différentes étapes de la régulation transcriptionnelle du chromosome X chez la femelle, à part le MSCI qui à lieu dans la lignée germinale chez le mâle (rectangle bleu). L'activité transcriptionnelle du chromosome X est représentée en blanc (majoritairement actif), en rose (partiellement actif/inactif) et en rouge (majoritairement inactif). La ligne bleue sur l'*X* actif symbolise une empreinte répressive sur *Xist*, afin que le X maternel reste actif. L'ombrage rose du Xp symbolise l'héritage d'informations épigénétiques à partir de la lignée germinale mâle. Epi : Épiblaste, ExE : Ectoderme extraembryonnaire, m : Pronoyau maternel, p : Pronoyau paternel, PE : Endoderme primitif, PGCs : Cellules germinales primordiales, PB : Corps polaire, TE : Trophectoderme, Xa : X actif, Xm : X maternel, Xp : X paternel. (Adaptée de Payer & Lee, 2008)

De manière intéressante, plusieurs études ont démontré que bien qu'exprimés faiblement, *Xist* et son anti-sens *Tsix*, deux ARNs non-codants essentiels dans la mise en place de l'inactivation du chromosome X dans les cellules embryonnaires (comme nous le verrons par la suite), ne sont pas indispensables à la mise en place du MSCI, la mutation de *Xist* n'empêchant pas le recrutement des marques répressives (Turner, 2002 ; McCarrey *et al.*, 2002). Ceci étant vrai malgré le fait que l'ARN de *Xist* recouvre les chromosomes sexuels dans les spermatocytes sauvages (Ayoub *et al.*, 1997).

Par la suite, l'environnement répressif semble être plus ou moins maintenu (appelé PMSC, pour : « *Post-Meiotic Sex Chromatin* ») (Costanzi & Pehrson, 1998 ; Greaves *et al.*, 2006 ; Turner *et al.*, 2006 ; Namekawa *et al.*, 2006), bien que nous assistons à une réactivation partielle du gonosome unique à partir du stade spermatide rond puis allongé (Nguyen & Disteche, 2006, Wang *et al.*, 2005, Khalil *et al.*, 2004).

b. L'inactivation empreintée suite à la fécondation

Suite à la fécondation ainsi qu'à l'activation zygotique du génome (au stade 2-cellules) dans les cellules femelles, on assiste de manière très précoce à une nouvelle forme d'inactivation du chromosome X : l'inactivation empreintée. L'empreinte est définie comme une marque épigénétique basée sur l'origine parentale de l'élément en question.

Cette forme particulière d'inactivation se caractérise par le fait que le chromosome X d'origine paternelle (Xp) sera toujours celui subissant l'inactivation, le chromosome X d'origine maternelle restant actif et étant même sur-exprimé pour accomplir la compensation de dose vis-à-vis des autosomes (*idem* chez le mâle) (Nguyen & Disteche, 2006). Chez la souris, cette inactivation est mise en place lors des premières divisions cellulaires de l'embryon puis est maintenue au cours du développement des tissus extraembryonnaires comme le placenta et une partie des membranes extraembryonnaires (Huynh & Lee, 2003 ; Mak *et al.*, 2004 ; Okamoto *et al.*, 2004 ; Takagi & Sasaki, 1975), provenant du trophoctoderme et de l'endoderme.

Deux études ont maintenant montré qu'il existe une forme d'inactivation empreintée indépendante de *Xist* (Kalantry *et al.*, 2009, Namekawa *et al.*, 2010), le fait qu'il s'agisse de la répression d'éléments répétés intergéniques paraissant l'hypothèse la plus probable. En tenant compte du fait que plusieurs autres études ont démontré la persistance du PMSC sur le chromosome X au cours de la spermatogénèse (Costanzi & Pehrson, 1998 ; Turner *et al.*,

2006 ; Namekawa *et al.*, 2006 ; Greaves *et al.*, 2006) malgré une réactivation partielle (Nguyen & Distèche, 2006 ; Wang *et al.*, 2005 ; Khalil *et al.*, 2004), il est tout à fait possible d'imaginer qu'une partie du chromosome X paternel soit déjà inactivé lors de la fécondation (en l'occurrence : les régions répétées). Ceci pourrait avoir l'avantage d'agir en tant que première empreinte parentale qui faciliterait éventuellement l'inactivation des parties géniques du chromosome X, actifs lors de la réactivation zygotique du génome (Patrat *et al.*, 2009 ; Namekawa *et al.*, 2010), par l'expression de *Xist*. L'expression de *Xist* a lieu dès le stade 2-cellules (Okamoto *et al.*, 2000, 2005 ; Huynh & Lee, 2003 ; Namekawa *et al.*, 2010 ; Kalantry *et al.*, 2009 ; Patrat *et al.*, 2009) et son ARN non-codant forme un domaine de taille croissante allant jusqu'à englober l'intégralité du chromosome X afin d'inactiver les régions géniques de celui-ci (Figure 3). Cette deuxième vague d'inactivation serait donc due à un second type d'empreinte parentale (peut être influencée par la première) qui permettrait la transcription du *Xist* paternel et donc la répression génique de son chromosome en *cis* à un moment variant entre le stade 4-cellules (Okamoto *et al.*, 2004 ; Patrat *et al.*, 2009) et le stade blastocyste précoce (Namekawa *et al.*, 2010 ; Patrat *et al.*, 2009) en fonction des gènes.

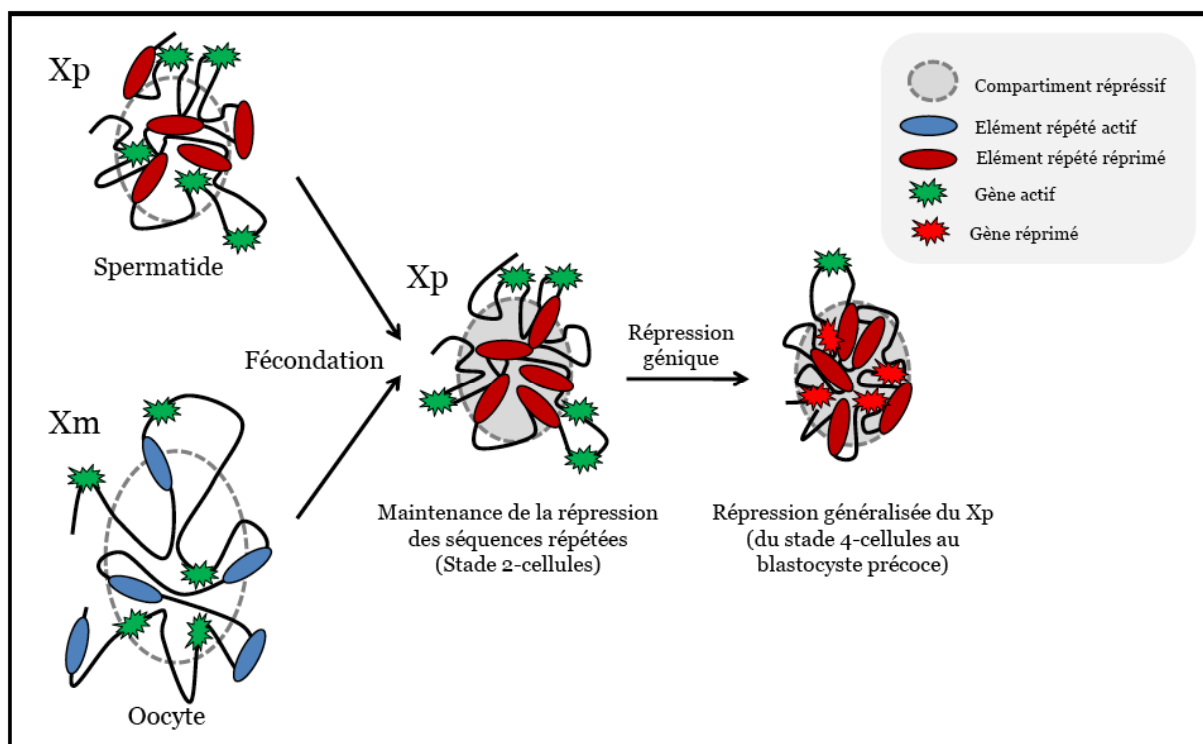


Figure 3 : Représentation schématique de la localisation génique par rapport au compartiment hétérochromatique (principalement formé par les séquences répétées) lors de la mise en place de l'inactivation empreintée. Avant la fécondation, le chromosome X présent a sein de l'ovocyte est transcriptionnellement actif. Dans les gamètes mâles, dans un premier temps (dans le spermatozoïde) la partie du génome du chromosome X se réactive partiellement suite au PMSC, tandis que les séquences répétées restent potentiellement réprimées. Dans un second temps

(dans le spermatozoïde) le génome entier est extrêmement compacté dû au remplacement des nucléosomes par les protamines, d'où une répression transcriptionnelle globale. Cependant, les gènes sont prêts à être exprimés dès la réactivation du génome post-fécondation, tandis que les séquences répétées seraient peut-être programmées à rester réprimées. Suite à la fécondation, on observe la formation d'un compartiment répressif principalement formé des séquences répétées autour du X d'origine paternelle (correspondant éventuellement à une forme d'empreinte). Puis au cours du développement pré-implantatoire, les séquences géniques vont progressivement se trouver réprimées lors de leur migration vers le centre du compartiment répressif.

Quant au chromosome X d'origine maternelle, qui pourrait être porteur d'une empreinte le protégeant de l'inactivation à ce stade (Lyon & Rastan, 1984 ; Shao & Takagi, 1990 ; Tada *et al.*, 2000), il est transcriptionnellement actif de la réactivation zygotique du génome jusqu'à l'initiation de l'inactivation aléatoire du chromosome X se déroulant dans les cellules de la masse interne du blastocyste tardif. Cette empreinte pourrait impliquer *Tsix*. Deux études ont démontré le rôle de l'antisens (par sa mutation) dans la protection de l'*Xm* de l'inactivation empreintée (Sado *et al.*, 2001 ; Lee, 2000), ceci pouvant tout simplement être dû au fait que le *Xm* n'arrive pas à réprimer correctement *Xist*, induisant ainsi l'inactivation sur les deux chromosomes (engendrant une nullisomie fonctionnelle et donc une mortalité cellulaire élevée).

Le fait que *Xist* soit exprimé uniquement à partir du chromosome X paternel lors du stade 2-cellules et qu'il initie l'inactivation dès la division cellulaire suivante, laisse supposer qu'il doit exister un moyen de répression secondaire afin d'empêcher son expression à partir du *Xm*, sachant que l'expression de *Tsix* n'est jamais observée avant le stade 8-cellules (Sado *et al.*, 2001).

c. Réactivation du chromosome X dans les blastocystes

Le chromosome X paternel est ensuite réactivé au sein des cellules de la masse interne de l'embryon lors des stades péri-implantatoires du développement (entre 3,5 et 4,5jpc) (Huynh & Lee, 2003, Mak *et al.*, 2004, Okamoto *et al.*, 2004). Il est devenu évident que cette étape (absente dans les tissus extra-embryonnaires) est nécessaire à l'effacement des marques hétérochromatiques engendrées par l'inactivation empreintée avant que la cellule puisse initier une deuxième vague d'inactivation du chromosome X : l'inactivation aléatoire.

Ce processus est sans doute effectué de manière active au vu du peu de nombre de divisions cellulaires nécessaires à l'obtention d'une réactivation quasi-totale des *Xi* au sein de l'épiblaste. L'absence de méthylation de l'ADN, au cours de l'inactivation empreintée facilite

sûrement cette transition. De manière concomitante à cette réactivation, on observe l'apparition d'une expression de *Tsix* à partir des deux allèles du chromosome X. La question de savoir si l'expression de l'antisens est responsable de la répression de *Xist* sur le Xp ou s'il s'agit d'autres mécanismes (par exemple, liés à la pluripotence cellulaire) est encore ouverte à débat.

Chez l'homme, après plusieurs années d'incertitude, une étude récente a finalement prouvé de manière fiable qu'il n'existe pas d'inactivation empreintée chez l'homme (au sein des tissus extraembryonnaires) (Moreira de Mello *et al.*, 2010). L'hypothèse d'une corrélation entre l'absence d'inactivation et l'absence d'une homologie de *TSIX* entre humains et souris paraît très intéressante, la découverte de la nature des empreintes parentales gouvernant ce processus chez la souris y apportera sûrement des éclaircissements précieux.

d. L'inactivation aléatoire du chromosome X

Suite à la réactivation du Xp au sein de l'épiblaste, ces cellules pluripotentes possèdent deux chromosomes X actifs. À partir du moment où l'embryon s'implante dans l'utérus, les cellules épiblastiques du blastocyste tardif commencent à subir une différenciation lignage-spécifique qui donnera lieu à tous les tissus constituant l'individu à proprement parler et parallèlement à cela, inactivent de manière aléatoire l'un de leurs deux chromosomes X (au sein des cellules femelles uniquement) (Monk & Harper, 1979 ; Tan & Tam *et al.*, 1993) (Figure 4). Ce processus d'inactivation aléatoire peut aussi être récapitulé *in vitro* lors de la rentrée en différenciation de cellules pluripotentes dérivées, comme les carcinomes embryonnaires et les cellules souches embryonnaires (ES) (Martin & Yatziv *et al.*, 1978 ; Rastan & Robertson, 1985), en faisant des outils particulièrement utiles pour disséquer les rouages génétiques et moléculaires du processus d'inactivation aléatoire du chromosome X.

Précédant l'inactivation aléatoire d'un des deux chromosomes X, la cellule doit d'abord se soumettre à deux processus particuliers : le comptage du nombre de chromosomes X dans la cellule et leur comparaison avec le nombre d'autosomes, ainsi que le choix de quel chromosome va être inactivé (Xm ou Xp), deux processus génétiquement distincts (Lee *et al.*, 1999).

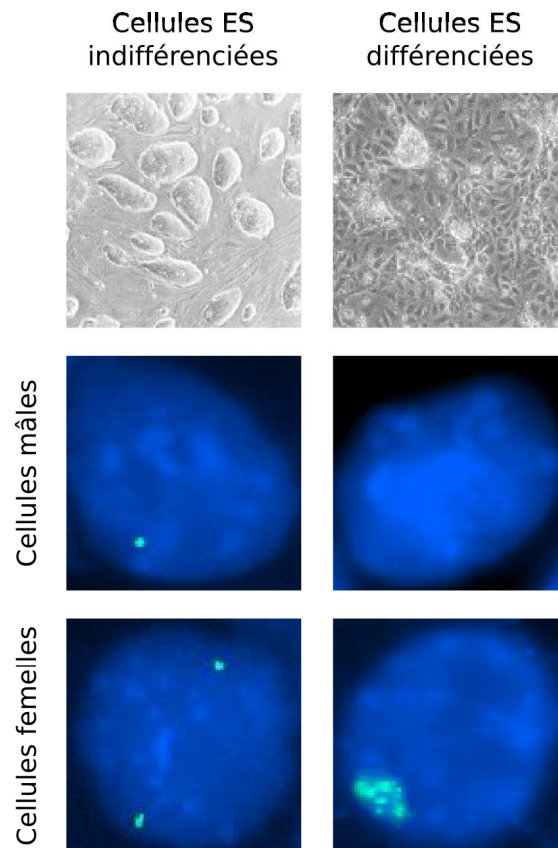


Figure 4 : Expression de *Xist/Tsix* dans les cellules ES. Panneaux du haut: apparence en microscopie à transmission de cultures de cellules ES (avant ou après différenciation à l'acide rétinoïque pendant 4 jours). Panneaux du milieu et du bas: Marquage par RNAFISH (en vert) des ARN issus du locus *Xist/Tsix*. Les noyaux sont colorés en bleu au DAPI. À gauche: cellules indifférenciées. Les transcrits *Tsix* sont quantitativement prépondérants par rapport aux transcrits *Xist* (Sun *et al.*, 2006). À droite: cellules différenciées. Dans les cellules femelles (en bas), l'expression de *Tsix* est perdue sur les deux chromosomes X, les transcrits visibles correspondent donc à *Xist* qui s'accumule au niveau du chromosome X inactif pour former un domaine coïncidant avec le territoire chromosomique (Chaumeil *et al.*, 2006).

i. Le Comptage

Tandis que lors de l'inactivation empreintée la cellule semble inactiver un chromosome X uniquement en fonction de son origine parentale (plus précisément de son empreinte parentale), au cours de l'inactivation aléatoire la cellule éteint un chromosome X en fonction du ratio X:A.

De manière globale, il est accepté que la cellule inactive tous les chromosomes X qu'elle contient, sauf un, par lot diploïde d'autosomes de manière à obtenir un ratio d'X actifs / nombre d'autosomes = $\frac{1}{2}$ (Rastan, 1983; Giannelli, 1963; Grumbach *et al.*, 1963).

Deux processus de comptage sont donc nécessaires à la cellule pour une inactivation aléatoire correcte : combien la cellule contient-elle de chromosomes X ? Ainsi que, combien la cellule possède-t-elle de lots d'autosomes par rapport au nombre de chromosomes X ?

Divers modèles ont été établis au fil des années, postulant l'existence d'un facteurs bloquant et de facteurs de compétence à l'inactivation du chromosome X (Starmer & Magnuson, 2009) mais, à ce jour, aucun de ces modèles n'a été complètement validé biologiquement, l'existence même de ces facteurs étant encore sujette à débat.

ii. Le Choix

Une fois la cellule ayant compté le nombre de chromosomes X présents au sein de son noyau, au moment où la différenciation de l'épiblaste se met en place chez la femelle, un des deux chromosomes X doit être choisi afin d'être inactivé. Théoriquement ce choix est aléatoire, 50% des X inactifs étant d'origine maternelle et 50% d'origine paternelle, les deux X étant au préalable, *a priori*, indiscernables l'un de l'autre. Une fois ce choix irréversible fait, il est transmis de cellule-mère en cellule-fille de manière clonale tout au long de la vie de l'individu, bien qu'il ait été observé que divers facteurs génétiques peuvent contribuer à un biais dans ce choix (Chadwick *et al.*, 2006).

Outre ces facteurs génétiques, on peut aussi envisager un choix établi de manière stochastique, non-coordonnée entre les chromosomes, c'est-à-dire où il suffirait qu'un chromosome initie sa phase d'inactivation pour que le processus de comptage fasse en sorte que l'autre chromosome reste actif. De cette manière, un biais pourrait être observé lorsque les deux chromosomes ne ressentent pas ou ne répondent pas de la même manière aux facteurs déclenchant l'initiation de l'inactivation (en l'occurrence la différenciation de la cellule). Ceci pourrait être ce que l'on observe lors de diverses mutations génétiques effectuées sur *Xist* ou son anti-sens *Tsix*. En effet, de manière générale, toute mutation touchant à la régulation de l'inactivation du chromosome X engendre un biais dans le choix de manière plutôt « prévisible ». Par exemple toute mutation causant une perte de fonction de *Xist* engendrera à coup sûr un biais de façon à ce que l'allèle muté devienne le Xa (Penny *et al.*, 1996, Marahrens *et al.*, 1998, Gribnau *et al.*, 2005, Monkhorst *et al.*, 2008) , et de même toute mutation portant atteinte à la fonction de l'anti-sens *Tsix* (dont le rôle principal est de réprimer *Xist*) fera que cet allèle sera automatiquement désigné Xi (Clerc & Avner, 1998 ; Luikenhuis *et al.*, 2001 ; Vigneau *et al.*, 2006 ; Morey *et al.*, 2004). Dans ces deux cas, le choix est biaisé, soit par le fait qu'un allèle ne pourra inactiver, soit car il ne peut s'empêcher

d'inactiver. Cependant, un choix non biaisé dans ces cellules engendrerait une mort cellulaire certaine de par le fait que le comptage ne serait plus respecté.

Une question qui s'applique pour le choix mais aussi pour le comptage est : la cellule inactive-t-elle un nombre variable de Xs, puis, seules les cellules femelles avec un X inactif survivent au cours du développement, ou y a-t-il une véritable communication entre les deux chromosomes X qui leur permet de « sentir » le fait qu'un autre chromosome X est déjà inactivé ou prédisposé à l'être ?

C'est dans cette deuxième optique que les interactions en *trans* (« transvection ») entre les deux chromosomes X, prennent tout leur sens (Marahrens, 1999, Bacher *et al.*, 2006, Xu *et al.*, 2006, Augui *et al.*, 2007). Le rapprochement des deux chromosomes au sein du noyau est observé entre 1,5 et 2 jours après l'induction de la différenciation par acide rétinoïque (Bacher *et al.*, 2006) et entre 2 et 4 jours lorsque initiée par formation de corps embryonnaires (Xu *et al.*, 2006) ce qui correspond approximativement dans les deux cas au moment où l'inactivation est initiée et où l'ARN de *Xist* commence à recouvrir le futur Xi en *cis*. En plus de favoriser une détection du statut d'activité du chromosome X homologue, ce phénomène de transvection possède l'avantage de faciliter la mise en place de la théorie de brisure de symétrie (Nicodemi & Prisco, 2007) engendrée par un facteur bloquant ou bien un activateur de l'inactivation (en fonction du modèle), l'un des chromosomes X étant fixé par toutes les molécules "signales" et le chromosome homologue par aucune.

La transvection des deux chromosomes X à un moment précis dans le développement de la cellule, concordant avec l'initiation de l'inactivation, pourrait jouer un rôle important dans le processus de choix du futur X inactif (Figure 5). Ce phénomène semble être médié par deux régions distinctes : *Xpr* (pour « *X-pairing Region* », situé environ 200kb en amont de *Xist*) ainsi que *Tsix/Xite* (pour « *X-inactivation Intergenic Transcription Elements* ») (*Xite* étant une région activatrice de *Tsix* sur laquelle nous reviendrons dans les chapitres suivants). D'après les résultats obtenus par les équipes d'Edith Heard et de Jeannie Lee, il semblerait que ces deux régions impliquées dans l'association des Xs entre eux agissent de manière différente voir opposée quant à leur effet sur *Xist*. Une zone de contact (*Xpr*) induirait la surexpression modérée de *Xist* sur les allèles en rendant celui-ci compétent pour sa future surexpression (Augui *et al.*, 2007 ; Navarro & Avner, 2010), l'autre zone de contact définirait le futur X actif (*Tsix/Xite*) en agissant potentiellement sur l'expression de *Tsix* en réprimant l'expression de l'antisens monoallèlement, ce qui engendrerait la surexpression de *Xist* sur l'allèle

opposé. En tout cas, la transvection semble nécessaire à l'initiation de l'inactivation, le chromosome X ne s'associant à rien restant actif dans les deux modèles.

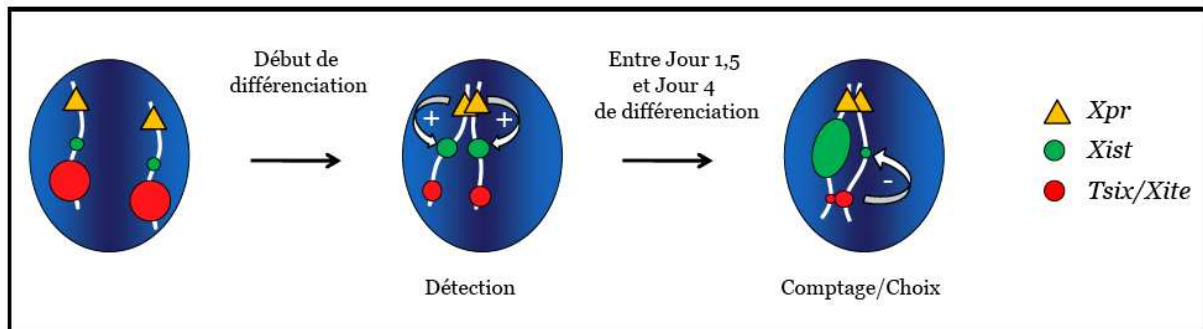


Figure 5 : Le « pairing » au sein de cellules ES femelles. Schématisation de la relation entre les deux chromosomes X en cours de différenciation cellulaire. La transvection des *XICs* est induite par *Xpr* (en jaune) et pourrait représenter un phénomène déclencheur de l'inactivation en surexprimant *Xist* (en vert) au cours d'une étape de détection. Le couplage des *Xpr* pourrait aussi faciliter les liaisons entre les régions *Tsix/Xite* (en rouge), ce qui résulte en une régulation réciproque de *Xist/Tsix*, rendant la surexpression de *Xist* monoallélique. La taille des cercles représentant *Xist* et *Tsix/Xite* correspond à l'intensité de leur activité transcriptionnelle au cours de la différenciation cellulaire.

Lorsque la cellule s'enclenche dans une voie de différenciation, le comptage ainsi que le choix sont effectués et l'inactivation « aléatoire » du chromosome X s'établit, éteignant transcriptionnellement l'un des deux allèles. L'inactivation de cet allèle est maintenue au cours de la vie de la cellule somatique et le choix de l'allèle à inactiver est transmis au cours des générations de manière clonale, donnant lieu à un individu mosaïque quant à la provenance de l'X actif/inactif (Xp/Xm ou *vice-versa*). La dynamique des phénomènes menant à l'état inactif ainsi que les caractéristiques de l'Xi sont décrits ultérieurement dans cette introduction.

Il existe cependant un tissu au sein duquel on observe une réactivation du chromosome X inactif, il s'agit des cellules germinales primordiales.

e. Réactivation du Xi dans la lignée germinale

Les cellules germinales primordiales ou PGC (pour « *Primordial Germ Cell* ») représentent les cellules à l'origine des gamètes (Figure 6). Ces cellules sont spécifiées entre 7,25 et 7,5jpc dans le mésoderme extraembryonnaire (d'origine épiblastique) (Ginsburg *et al.*, 1990, Chuva de Sousa Lopes *et al.*, 2008, Sugimoto & Abe, 2007) avant de migrer dans l'endoderme du

tube digestif postérieur à partir de 8,5jpc, pour finalement coloniser les gouttières génitales entre 10,5 et 11,5jpc (pour revue voir McLaren, 2003).

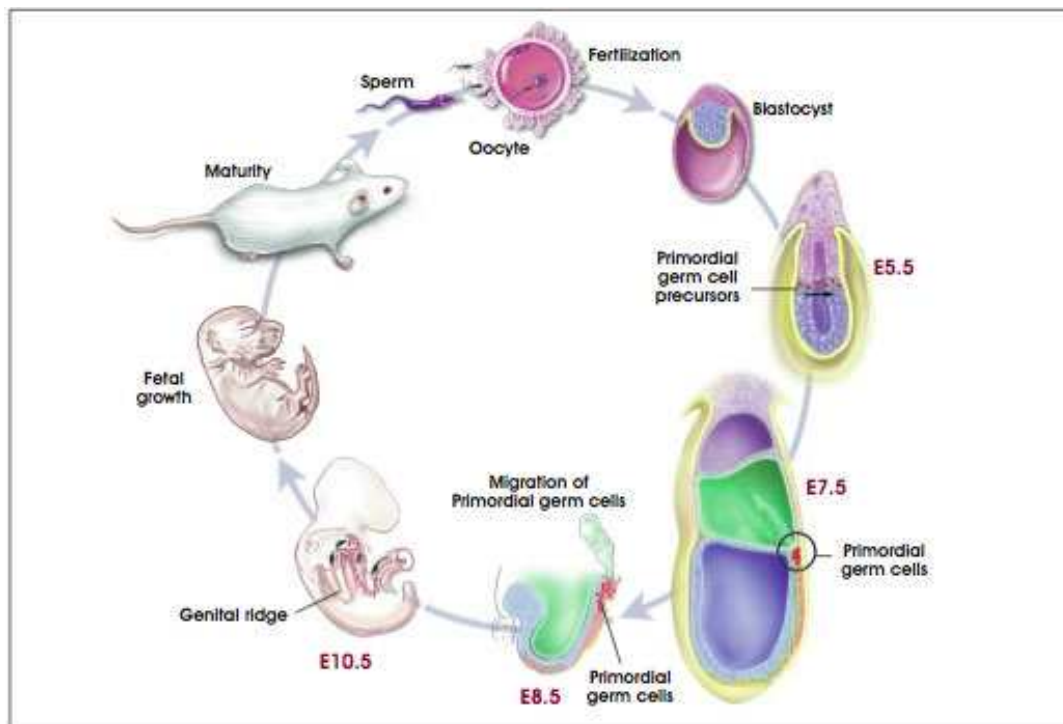


Figure 6 : Développement des cellules germinales primordiales au sein de l'embryon murin. A partir de 5,5jpc on observe la différenciation d'une partie du mésoderme extraembryonnaire en cellules précurseurs de PGCs. Puis, au cours des jours suivants, les PGCs vont acquérir leur véritable identité (un retour à la pluripotence et la réactivation du chromosome X inactif) lors de leur migration le long de la gouttière uro-génitale (période au sein de laquelle il est possible de les dériver en cellules EGs). Chez l'adulte, les PGCs aboutiront à la formation des gamètes. (D'après Winslow & Duckwall, 2001, <http://stemcells.nih.gov/info/scireport/appendixia.asp>)

Les PGCs étant des cellules ayant différencié à partir de l'épiblaste, chez les femelles, leurs chromosomes X sont inactivés de manière aléatoire. Cependant au cours de la migration des PGCs le long de la gouttière uro-génitale, le chromosome X inactif va subir une réactivation concordant avec une déméthylation globale de l'ADN du génome ainsi que l'effacement de l'empreinte parentale (Monk *et al.*, 1987 ; Hajkova *et al.*, 2002 ; Lee *et al.*, 2002), afin d'établir de nouvelles marques sexe-spécifiques durant la maturation des ovocytes.

Cette réactivation du Xi ainsi que la reprogrammation globale du génome sont le fruit d'interactions tissus-spécifiques, le passage de ces cellules le long de la gouttière génitale étant indispensable (Tam *et al.*, 1994, Francavilla & Zamboni, 1985, Chuva de Sousa Lopes *et al.*, 2008).

Curieusement, l'extinction transcriptionnelle de *Xist* s'effectue indépendamment de *Tsix*, une transcription de l'antisens n'étant observée à aucun moment au cours du développement des

PGCs (Nesterova *et al.*, 2002, Sugimoto *et al.*, 2007). Parallèlement à ceci, on observe la ré-expression de facteurs de pluripotence tels que *Oct3/4* (Tsang *et al.*, 2001, de Napoles *et al.*, 2007) et *Nanog* (de Napoles *et al.*, 2007) immédiatement avant la chute d'expression de *Xist*, soulignant une fois de plus les liens étroits entre inactivation du chromosome X et pluripotence cellulaire.

III. Modèles cellulaires

Maintenant que nous avons brièvement vu en quoi consiste l'inactivation du chromosome X, plutôt sous un angle macromoléculaire, l'étape suivante est de décrire ce qui se passe au niveau épigénétique, protéique, et transcriptionnel. Mais avant cela, je pense qu'il est important que je détaille quelque peu les modèles cellulaires majoritairement utilisés dans la description des mécanismes de l'inactivation *ex vivo* chez la souris (Figure 7 et 8).

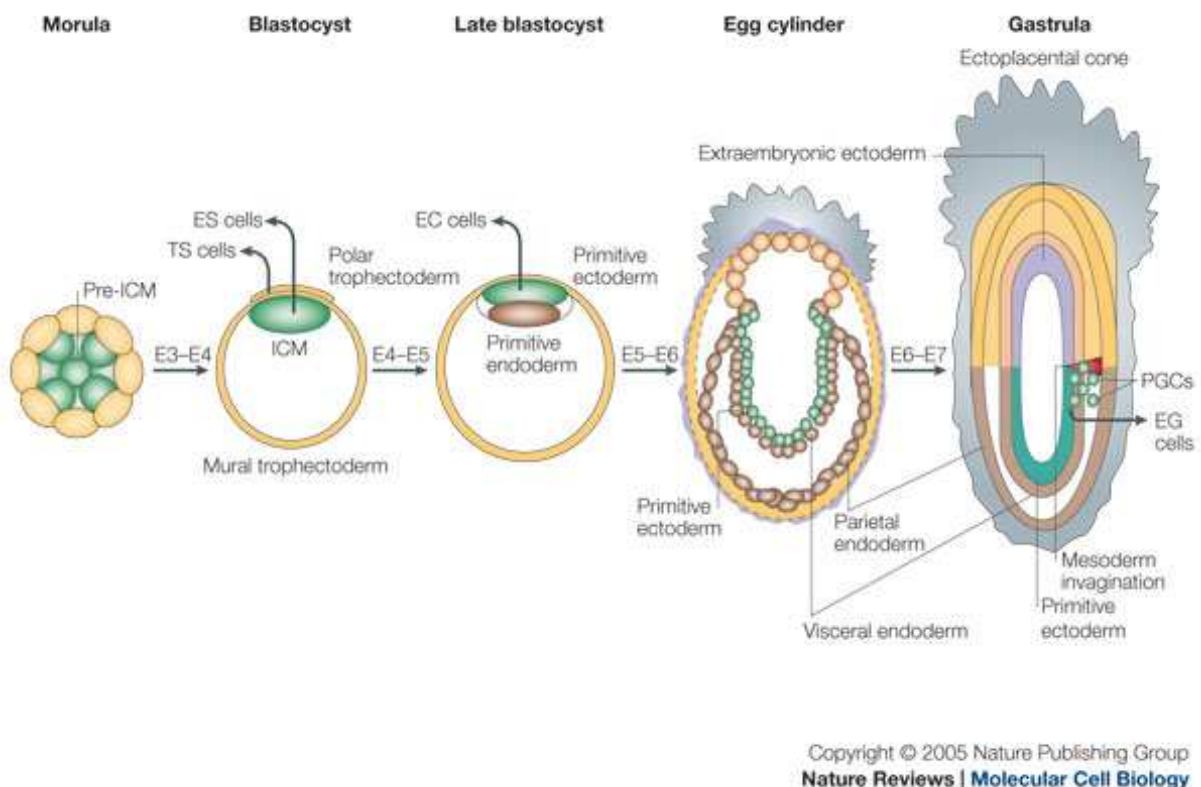


Figure 7 : L'origine des cellules souches dans l'embryon mammifère. Dans cette figure, les cellules pluripotentes de l'embryon sont colorées en vert. ICM : masse cellulaire interne, TS : cellules trophoblastiques, ES : cellules souches embryonnaires, EC : cellules du carcinome embryonnaire, PGCs : cellules germinales primordiales, EG : cellules germinales embryonnaires. (D'après Boiani & Schöler, 2005)

- Les cellules trophoblastiques souches (TS, pour « *Trophoblastic Stem* ») sont dérivées du trophoblaste polaire des blastocystes à 3,5jpc ou de l'ectoderme extraembryonnaire à 6,5jpc (Tanaka *et al.*, 1998); elles maintiennent l'inactivation empreintée acquise plus tôt durant le développement, mais les modalités de ce maintien diffèrent selon qu'elles soient cultivées de manière à préserver leurs caractéristiques de cellules souches ou que l'on induise leur différenciation

- Les cellules souches extraembryonnaires endodermiques (XEN, pour « *eXtra-Embryonic eNdoderm* ») sont dérivées de la masse cellulaire interne de blastocystes à 3,5jpc et expriment à la fois des marqueurs d'endoderme viscéral et d'endoderme pariétal (Kunath *et al.*, 2005); l'inactivation est réalisée de manière empreintée dans ces cellules, de manière quelque peu différente des lignages trophoblastiques

- Les cellules souches embryonnaires (ES, pour « *Embryonic Stem* ») sont dérivées de la masse cellulaire interne de blastocystes à 3,5jpc (Evans et Kaufman, 1981; Martin, 1981) et sont probablement représentatives des cellules de l'épiblaste vers 5,0jpc ; les deux chromosomes X sont actifs lorsqu'elles sont cultivées de manière à préserver leur pluripotence, mais l'un d'eux est inactivé lorsqu'on induit leur différenciation.

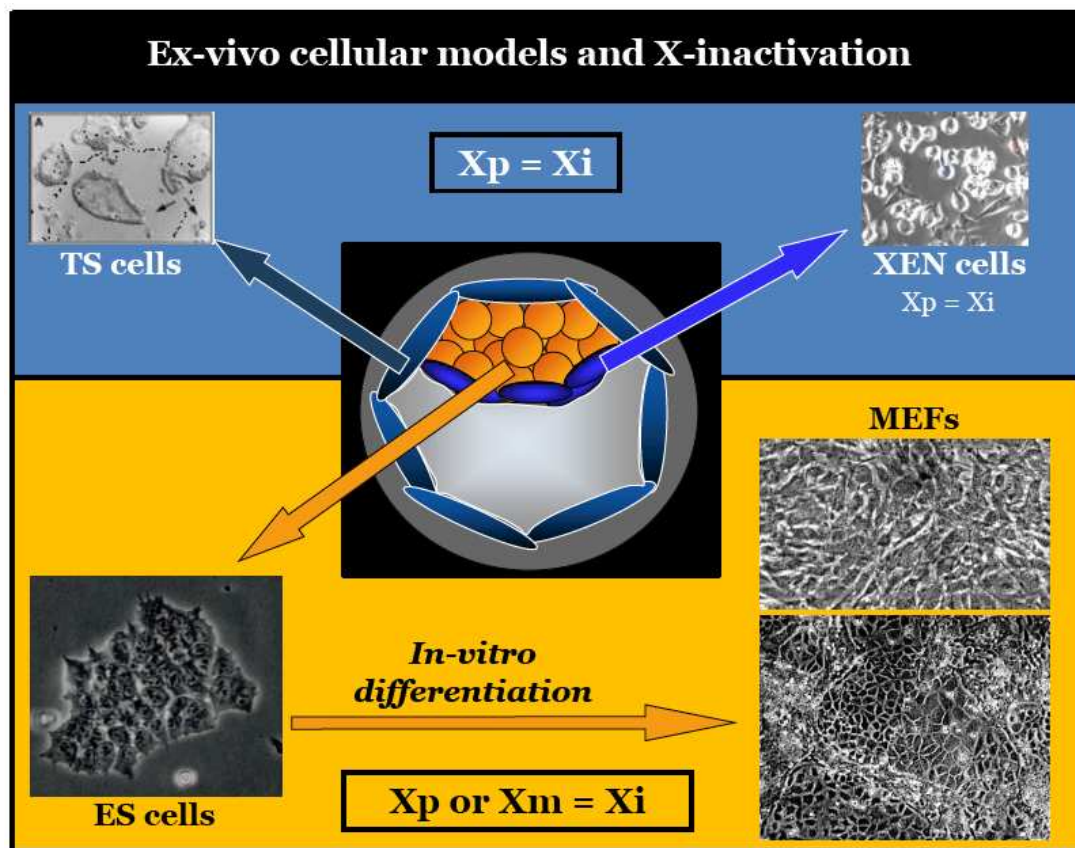


Figure 8 : Modèles cellulaires ex-vivo et inactivation du chromosome X. Représentation des cellules TS provenant du trophoctoderme et des XEN provenant de l'endoderme primitif, dont le chromosome X paternel est toujours celui inactivé suite à l'inactivation empreintée (partie supérieure bleue). Les cellules ES proviennent de la masse cellulaire interne et subissent une inactivation aléatoire suite à une différenciation (ici après 4 jours à l'acide rétinoïque) (partie inférieure jaune). Les MEFs (fibroblastes) étant le modèle principalement utilisé afin d'étudier le statut du chromosome X inactivé aléatoirement depuis de nombreuses divisions cellulaires.

- Les cellules germinales embryonnaires (EGC, pour « *Embryonic Germ Cells* ») sont dérivées par dé-différenciation des PGC prélevées entre 8,5 et 11,5jpc (Durcova-Hills *et al.*, 2006; Matsui *et al.*, 1992; Resnick *et al.*, 1992). Les cellules EG ressemblent en de nombreux points aux cellules ES : leur profil d'expression génique est pratiquement identique à celui des cellules ES (Sharova *et al.*, 2007 ; Mise *et al.*, 2008), leurs deux chromosomes X sont actifs (Stewart *et al.*, 1994) et répliqués précocement et ils expriment *Tsix* (qui n'est jamais observé dans les PGCs) ainsi que *Xist* (mais à son niveau basal) (Nesterova *et al.*, 2002). Et tout cela, quelque soit le stade à partir duquel elles ont été dérivées (8,5jpc à 13,5jpc). Ce qui est le plus frappant est qu'aucun X_i n'est observé dans les EGs, quelque soit le stade des PGCs à partir desquels ils sont dérivés. Cela pourrait laisser supposer que quelque chose de sécrété par les cellules nourricières sur lesquelles elles sont dérivées ou bien un produit dans le milieu de culture (habituellement utilisé pour la culture de cellules ES) induise la réactivation active du

Xi, et de manière plus générale, une convergence développementale vers des cellules ES.

Cependant il existe des différences entre ES et EGs, notamment au niveau de la méthylation de l'ADN des gènes soumis à l'empreinte parentale (Shovlin *et al.*, 2008). Le statut de la méthylation de l'ADN des gènes empreintés semble être hétérogène dans des EGs dérivées précocement (8jpc), indiquant déjà une mise en place de l'effacement de l'empreinte. En effet, dans les EGs dérivées tardivement (12jpc) quasiment toute la méthylation site-spécifique est perdue alors qu'elle est présente dans les ES (Durcova-hills *et al.*, 2001). Toutefois, les domaines de méthylation différentielle (DMD) des gènes Igf2/H19 se trouvent hyperméthylés dans les EGs tardives (Tada *et al.*, 1998, Durcova-Hills *et al.*, 2001, Petkov *et al.*, 2009) mais hypométhylés dans les PGCs correspondantes (Szabo & Mann, 1995), laissant supposer que la reprogrammation épigénétique des PGCs est modifiée par l'environnement cellulaire lorsque dérivées en EGs (soit par une re-méthylation soit par une déméthylation peu efficace), bien que l'étude d'un plus grand nombre de gènes soumis à empreinte serait utile pour véritablement en tirer une conclusion. Une autre différence récemment observée par Ciaudo *et al.* concerne un groupe de microARNs (miRs) exprimés uniquement au cours la différenciation précoce des cellules ES mâles (Ciaudo *et al.*, 2009). Toutefois, ce biais d'expression est retrouvé dans les EGs (dérivées à 8,5 ou 9,5jpc) mais pas dans les cellules embryonnaires gonadales (correspondant donc à un stade plus avancé des PGCs) où il est exprimé par les deux sexes au départ, avant d'être réprimé au cours de la suite de la spécification cellulaire. Bien qu'il soit clair qu'il existe de nombreuses différences entre PGCs et EGs, et de nombreuses similitudes entre EGs et ES, je pense qu'il est trop tôt pour dénigrer l'intérêt des EGs dans un contexte d'étude de cellules germinales. De plus, indépendamment de la comparaison entre les différents types cellulaires, l'étude des EGs peut aussi apporter des informations pertinentes quant à la manière dont elle reprogramme son génome et réactive leur chromosome X inactif.

- Les PGC dérivées *in vitro* à partir de cellules ES (Geijsen *et al.*, 2004) et les cellules germinales (GS, pour « *Germline Stem cells* ») mâles dérivées à partir de testicules néonatales (Kanatsu-Shinohara *et al.*, 2004). Ces cellules GS ont un profil d'expression génique plus proche de celui des PGC *in vivo* (Mise *et al.*, 2008) et pourraient donc être de meilleurs modèles cellulaires pour étudier la réactivation du Xi dans un contexte gonadal.

Comme nous l'avons brièvement évoqué plusieurs fois dans cette introduction, il existe deux préoccupations majeures auxquelles il est difficile de répondre lorsqu'on travaille avec des modèles cellulaires en culture *ex vivo* dans l'étude de processus développementaux :

- À quel point le processus de dérivation de lignées cellulaires (provenant de tissus *in vivo*) ou leur maintien en milieu de culture pendant des périodes prolongées altèrent-ils leurs caractéristiques par rapport au lignage d'origine ?
- Quelle est la capacité de ces modèles cellulaires à récapituler les événements développementaux *in vivo* étudiés *ex vivo* ?

Le premier point peut être adressé en multipliant diverses études comparatives, le facteur limitant étant la quantité de matériel disponible lorsqu'on travaille sur des stades embryonnaires précoces, certaines expériences étant donc plus faciles à réaliser que d'autres. Cependant avec le progrès permanent que l'on observe aussi bien dans les méthodes que dans la machinerie utilisées, de nos jours l'étude d'échantillons cellulaires en faible quantité devient possible.

Le deuxième point concerne surtout, dans notre cas, le type de différenciation cellulaire employé pour récapituler certains phénomènes, comme l'inactivation du chromosome X. Il existe actuellement 3 méthodes pour différencier les cellules souches embryonnaires : l'induction par acide rétinoïque, la formation de corps embryonnaires (ou EBs pour « *Embryoid Bodies* ») ou la différenciation par dilution (Morey *et al.*, 2001), chaque méthode ayant ses atouts et ses inconvénients et privilégie la différenciation des cellules vers un certain type cellulaire ; cependant, au sein de notre laboratoire, nous avons choisi la différenciation par ajout d'acide rétinoïque.

Le fait que différentes méthodes aboutissent à différents résultats, amène certaines équipes à retrouver, après différenciation, une expression ectopique importante de *Xist* chez les mâles mutés pour leur gène *Tsix* (Clerc & Avner, 1998, Morey *et al.*, 2001 *et al.*, 2004, Luikenhuis *et al.*, 2001, Vigneau *et al.*, 2006, Sado *et al.*, 2002, Sado *et al.*, 2001), et d'autres non (Lee *et al.*, 1999 ; Lee, 2005 ; Shibata *et al.*, 2008 ; Sado *et al.*, 2001 ; Ohhata *et al.*, 2006). Dans la grande majorité des cas, une expression ectopique de *Xist* est plus fréquemment observée lorsque les cellules sont dérivées à l'aide d'acide rétinoïque qu'en induisant la formation de corps embryonnaires. Ceci peut s'expliquer par le fait que l'acide rétinoïque réprime la voie de signalisation Wnt/béta-caténine (Verani *et al.*, 2007 ; Zhao & Duester, 2008), la bêta-caténine étant un activateur transcriptionnel du gène *Nanog* (Takao *et al.*, 2006), lui-même un répresseur transcriptionnel direct de *Xist* ainsi qu'un facteur de pluripotence majeur (Navarro

et al., 2008). Ainsi, il serait logique d'assister à une dé-répression bien plus rapide au sein des cellules différenciées à l'acide rétinoïque que par formation de corps embryonnaires. Les techniques d'analyse peuvent aussi être mises en cause de par le fait que des observations sur gel sont bien plus difficiles à comparer que des données provenant de PCR-quantitative.

IV. Les caractéristiques du X inactif

Comme nous avons pu le voir précédemment, la présence d'un chromosome X inactif au sein de toute cellule somatique de mammifère euthérien femelle est indispensable à son bon fonctionnement, ceci étant dû au processus de la compensation de dose qui assure une homéostasie génique au sein de la cellule. Il existe deux instances de mise en place d'un chromosome X inactif : une première fois dans l'ensemble des cellules de l'embryon, entre le stade 2-4 cellules et le stade péri-implantatoire de l'embryon au cours de laquelle le chromosome X d'origine paternelle est toujours celui choisi pour être inactivé, puis une deuxième fois, après réactivation du Xp, au sein des cellules de l'épiblaste et ce de manière aléatoire.

Le processus de l'inactivation aléatoire du chromosome X est un mécanisme qui se met en place au cours de plusieurs divisions cellulaires lors de la différenciation des cellules de la masse interne. La mise en place de cette extinction transcriptionnelle à l'échelle d'un chromosome entier se fait en plusieurs étapes séquentielles à travers la participation de diverses protéines engendrant un ensemble de modifications hétérochromatiques de manière directement liée à l'état de différenciation de la cellule. Deux phases majeures sont distinguables au cours de l'inactivation : l'initiation et le maintien. L'initiation de l'inactivation est engendrée suite à l'établissement du ratio X:A (le comptage) et au choix du chromosome à inactiver. Une fois l'inactivation engagée, il est nécessaire d'engager une répression transcriptionnelle stable afin que cette inactivation puisse être maintenue au cours de la vie de la cellule mais aussi lors des divisions mitotiques. Les deux phases ne font pas forcément appel aux mêmes acteurs et au cours du chapitre suivant je vais tenter de brièvement décrire leur composition ainsi que leur agencement chronologique.

a. L'ARN non-codant *Xist*

En 1949, la présence d'une structure particulièrement dense au sein du noyau a été décrite par Barr grâce à une coloration cytologique dans des cellules de chatte (Barr & Bertram, 1949). Dix ans plus tard, Ohno associe enfin ce corpuscule de Barr à l'un des deux chromosomes X à l'état condensé hétérochromatique (Ohno, 1959 ; Ohno & Hauschka, 1960). Grâce à des études sur le pelage des souris et des chats, et notamment sur des mutations des gènes codant pour la couleur de celui-ci, Mary Lyon a pu, en 1961, émettre l'hypothèse que le corpuscule de Barr correspond à l'un des deux chromosomes X réprimé transcriptionnellement et ce, de manière aléatoire au cours du développement embryonnaire (Lyon, 1961). Il faudra attendre 35 ans pour qu'en 1996, on découvre enfin l'acteur principal de l'inactivation du chromosome X : *Xist* pour « *X Inactive Specific Transcript* » (Penny *et al.*, 1996, Marahrens *et al.*, 1997). La même année, on observe que cet ARN non-codant, déjà décrit à partir de 1991 (Brown *et al.*, 1991, Borsani *et al.*, 1991, Brockdorff *et al.*, 1991), forme un large domaine colocalisant avec le X inactif (Brown *et al.*, 1992, Clemson *et al.*, 1996).

Au cours de ce chapitre nous nous contenterons d'étudier le rôle fonctionnel de *Xist* ; sa structure et sa régulation étant traitées ailleurs (voir le chapitre : « *Xist*, *Tsix* : généralités », p51).

Dans les cellules ES non-différenciées femelles, avant l'inactivation, le gène *Xist* est très faiblement exprimé chez les deux sexes. Lorsque observé par RNA-FISH, à ce stade, il forme un signal ponctiforme (« *pinpoint* ») localisé au site de transcription sur les deux allèles. Dès 24h de différenciation (induit ici à l'acide rétinoïque), on observe une surexpression de transcrits stables de *Xist* qui recouvrent ou "décorent" le futur X inactif ; ces phénomènes sont vraisemblablement les premiers événements de l'initiation de l'inactivation.

Au départ de la différenciation, le domaine formé par *Xist* reste de petite taille, peu intense, et présente un marquage diffus (domaine immature), puis, au fur et à mesure que la différenciation progresse, il va s'accroître et s'intensifier jusqu'à occuper l'intégralité du domaine nucléaire du chromosome à partir duquel il est transcrit (domaine mature). La répression génique commence rapidement après l'accumulation de *Xist*, mais il faut tout de même attendre entre 48 et 72h avant de voir une répression transcriptionnelle des premiers gènes (Sheardown *et al.*, 1997, Heard *et al.*, 2001).

Xist comprend de 4 à 6 régions répétées (en fonction de l'analyse) numérotées de A à F (Brown *et al.*, 1992, Beletskii *et al.*, 2001) (Figure 9). Sur ces 4 à 6 régions l'importance capitale de deux d'entre elles au processus d'inactivation a été démontrée, le *Repeat-A* étant

responsable de la fonction répressive de *Xist* (Wutz *et al.*, 2002) et le *Repeat-C* chargé de sa localisation sur le chromosome X en *cis* (Beletskii *et al.*, 2001), bien que Wutz *et al.* imputent cette action à une fonction redondante mais coopérative de différentes régions répétées de *Xist* (Wutz *et al.*, 2002).

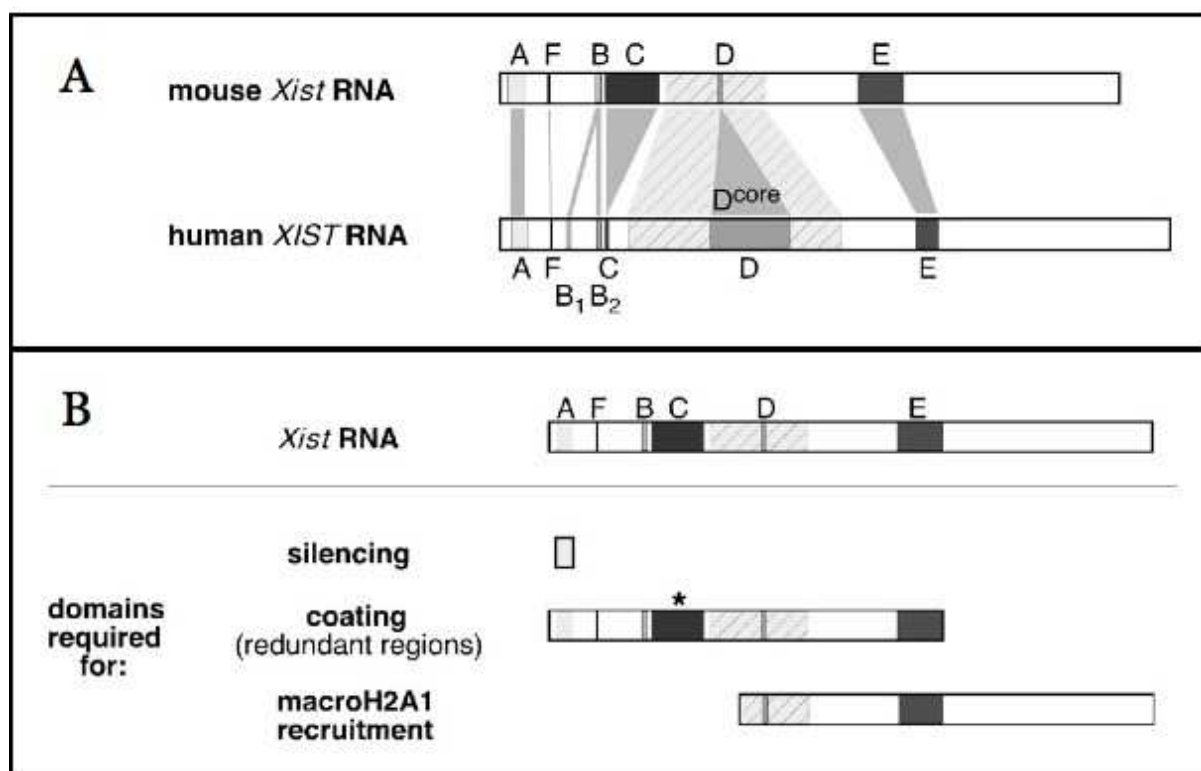


Figure 9 : Les séquences répétées de *Xist*. (A) Conservation de la structure des séquences répétées de *Xist/XIST* entre la souris et l'homme. Les répétitions A, B, C, D et F appartiennent à l'exon 1 tandis que la répétition E se trouve en 5' de l'exon 6 chez l'homme et de l'exon 7 chez la souris. (B) Les domaines fonctionnels de l'ARN *Xist*. L'astérisque représente la zone cible d'un peptide d'acides nucléiques antisense qui lorsque dirigé contre le *Repeat-C* empêche la fixation de *Xist* au chromosome X (Beletskii *et al.*, 2001). (D'après Plath *et al.*, 2002)

i. Propagation de *Xist* et l'extinction transcriptionnelle

La localisation de l'ARN *Xist* est un élément important de sa fonction, et savoir de quelle manière il se lie au chromosome X inactif est du plus grand intérêt. Malheureusement, à ce jour, très peu d'études ont été menées afin d'apporter une réponse précise et fiable à ce sujet. Le fait même de savoir si l'ARN non-codant se lie directement à l'ADN est sujet à controverse (Clemson *et al.*, 1996, Gilbert *et al.*, 2000).

Au cours des premiers instants de la transcription de *Xist*, on observe la présence du transcrit

uniquement au site de transcription, puis au cours de la différenciation, très rapidement, *Xist* va se propager le long de son chromosome X en *cis* (notamment grâce à ses séquences répétées comme vu précédemment), ne s'hybridant jamais avec le chromosome X actif (quelque soit sa proximité) (Wutz communication orale *et al.*, 2009). Sachant que seulement environ 2000 molécules de *Xist* sont présentes dans le noyau de cellules somatiques femelles à un temps donné (Buzin *et al.*, 1994), et ce, afin de réprimer plus de 150Mb, il doit exister une manière efficace d'aider ce transcrit à se propager le long du Xi de manière à recouvrir la quasi-totalité de son domaine chromosomique. Malheureusement, la manière dont se propage *Xist* le long du chromosome X inactif lors de sa surexpression, notamment ses partenaires protéiques et ADN demeure encore mystérieuse. Cependant, depuis maintenant plus d'une dizaine d'années (Lyon, 1998), la théorie que les LINEs (pour « *Long INterspersed Elements* ») de type L1, qui sont des retrotransposons, mesurant jusqu'à 6-7kb, soient impliqués dans ce processus persiste. Ceci est notamment dû au fait qu'ils sont particulièrement enrichies sur le chromosome X, par rapport au reste du génome (Boyle *et al.*, 1990, Bailey *et al.*, 2000, Chureau *et al.*, 2002, Nesterova *et al.*, 2001, Mikkelsen *et al.*, 2007, Waterston *et al.*, 2002), à l'exception des gènes échappant à l'inactivation (Allen *et al.*, 2003, Ross *et al.*, 2005). De plus, leur densité semble corrélée avec l'efficacité de la propagation de l'état répressif le long des autosomes, que ce soit dans des contextes de transgénèse de *Xic* ou bien de translocation X;autosome (Lyon, 1998). Mais cela est aussi dû au fait que les LINEs semblent jouer un rôle capital dans la propagation dans la nucléation d'un compartiment nucléaire répressif autour du Xi (Chow *et al.*, 2010).

Les LINEs anciens (tronqués en 5', transcriptionnellement et rétrotranscriptionnellement incompetents) faciliteraient la mise en place d'un domaine hétérochromatique (et donc peut-être la propagation de *Xist* le long du Xi), ce domaine ayant comme fonction d'inactiver les séquences introniques ainsi que la grande majorité des séquences répétées et d'en exclure l'ARN polymérase II, et ce d'une manière *Repeat-A* indépendante (Chaumeil *et al.*, 2006, Clemson *et al.*, 2006). Puis, au cours d'un processus, cette fois-ci *Repeat-A* dépendant, on assiste au rapatriement des séquences géniques vers le centre du domaine hétérochromatique afin d'induire leur répression, seuls les gènes échappant à l'inactivation restent en bordure du domaine afin de pouvoir continuer à être transcrits. Enfin, par un mécanisme encore inexpliqué, le *Repeat-A* favoriserait la transcription des LINEs récents, ce phénomène

aboutissant à l'extinction transcriptionnelle de certains gènes situés dans des régions plus résistantes à la répression, possiblement *via* un processus d'ARNi (Irvine *et al.*, 2006, Matzke *et al.*, 2007) (Figure 10). De plus, les auteurs décrivent ici, pour la première fois, l'observation de petits foci d'ARN Polymérase II au sein du domaine hétérochromatique formé par *Xist*, ceux-ci se superposant aux L1 récents qui continuent sans doute à contribuer à l'extinction transcriptionnelle à travers leur expression.

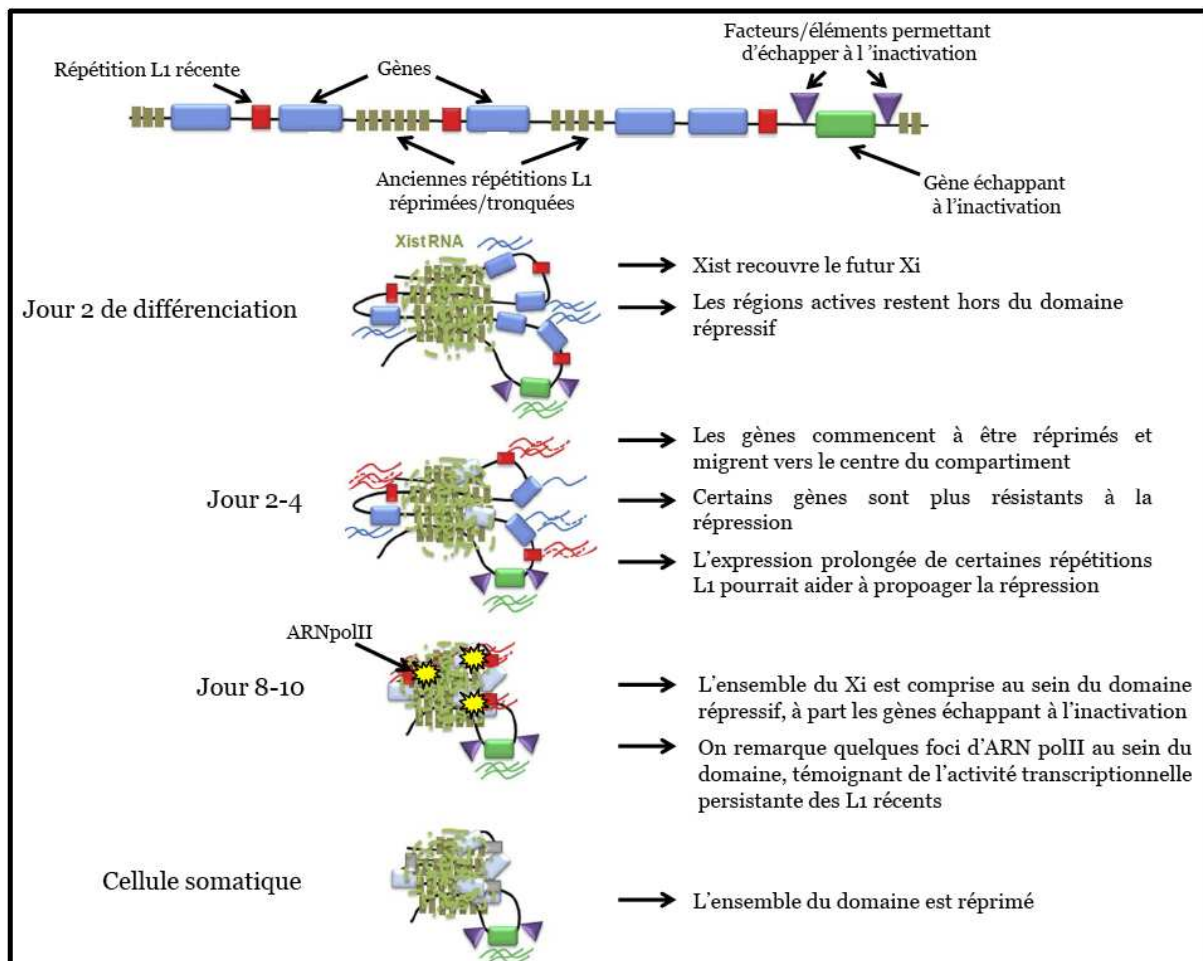


Figure 10 : Modèle du comportement des LINES et de leurs actions au cours de l'inactivation du chromosome X. Les LINES non exprimés pourraient faciliter la formation d'un compartiment nucléaire hétérochromatique de manière précoce au cours de l'inactivation du chromosome X. Les LINES transcriptionnellement actifs pourraient faciliter la propagation de l'extinction transcriptionnelle à certaines régions du chromosome X, particulièrement les zones plus résistantes à l'inactivation. Dans un premier temps ce phénomène aurait lieu autour des domaines échappant à l'inactivation à l'extérieur du compartiment répressif, puis les LINES récents ainsi que les gènes sur lesquels ils influent migrent au centre de ce compartiment, tout en maintenant momentanément l'activité transcriptionnelle des séquences répétées récentes (symbolisé par les foci de d'ARNpolII observable en jaune). (Adaptée de Chow *et al.*, 2010)

Ceci expliquerait donc partiellement qu'en délétant le *Repeat-A*, Anton Wutz n'observe pas

d'extinction génique du Xi malgré le fait que le transcrit arrive à décorer celui-ci en *cis*. De par la nature de cette séquence répétée, qui forme des structures secondaires en tiges-boucles complexes (Maenner *et al.*, 2010), on peut supposer que ce défaut de répression génique est aussi dû à une absence de recrutement de facteurs répressifs, dont PRC2 qui a déjà été démontré comme fixant cette région du transcrit *Xist* (Hendrich *et al.*, 1997, Maenner *et al.*, 2010, Zhao *et al.*, 2008).

ii. *Xist* au fil du temps

Par contre, l'action répressive de *Xist*, attribuée à la région du *Repeat-A* (comme nous venons de le voir) semble n'être fonctionnelle qu'au cours d'une fenêtre de temps limitée (Figure 11). Anton Wutz et Rudolph Jaenisch ont décrit une fenêtre d'opportunité au cours de laquelle *Xist* est capable d'engendrer l'inactivation du chromosome X : les 2 à 3 premiers cycles cellulaires des cellules ES en différenciation (correspondant à de 24h à 48h de différenciation par acide rétinoïque) (Wutz & Jaenisch, 2000). Au cours de cette période l'inactivation est réversible mais une fois la fenêtre dépassée, la cellule perd la capacité d'initier l'inactivation aléatoire (après la fenêtre de temps bien que l'ARN non-codant va recouvrir son chromosome en *cis*, on n'assiste pas à l'induction d'une répression transcriptionnelle). De plus, l'inactivation du chromosome X reste réversible pendant une courte période de 48h post-différenciation lorsque l'on y réprime *Xist*, et induit une sorte de mémoire chromosomique de l'état inactif, puisque si l'on ré-induit l'expression de *Xist* hors de la fenêtre d'opportunité, cette fois-ci, l'inactivation est possible (Wutz *et al.*, 2002) (Figure 11).

D'autres études proposent que *Xist* ne serait pas indispensable au maintien de la répression transcriptionnelle, son expression ne semblant nécessaire qu'au cours des premières divisions cellulaires post-différenciation c'est-à-dire pour l'initiation de l'inactivation (Brown & Willard, 1994 ; Csankovszki *et al.*, 1999 ; Wutz & Jaenisch, 2000), l'Xi ne se réactivant qu'à des taux très faibles dans les cellules au sein desquelles il est délété ou réprimé.

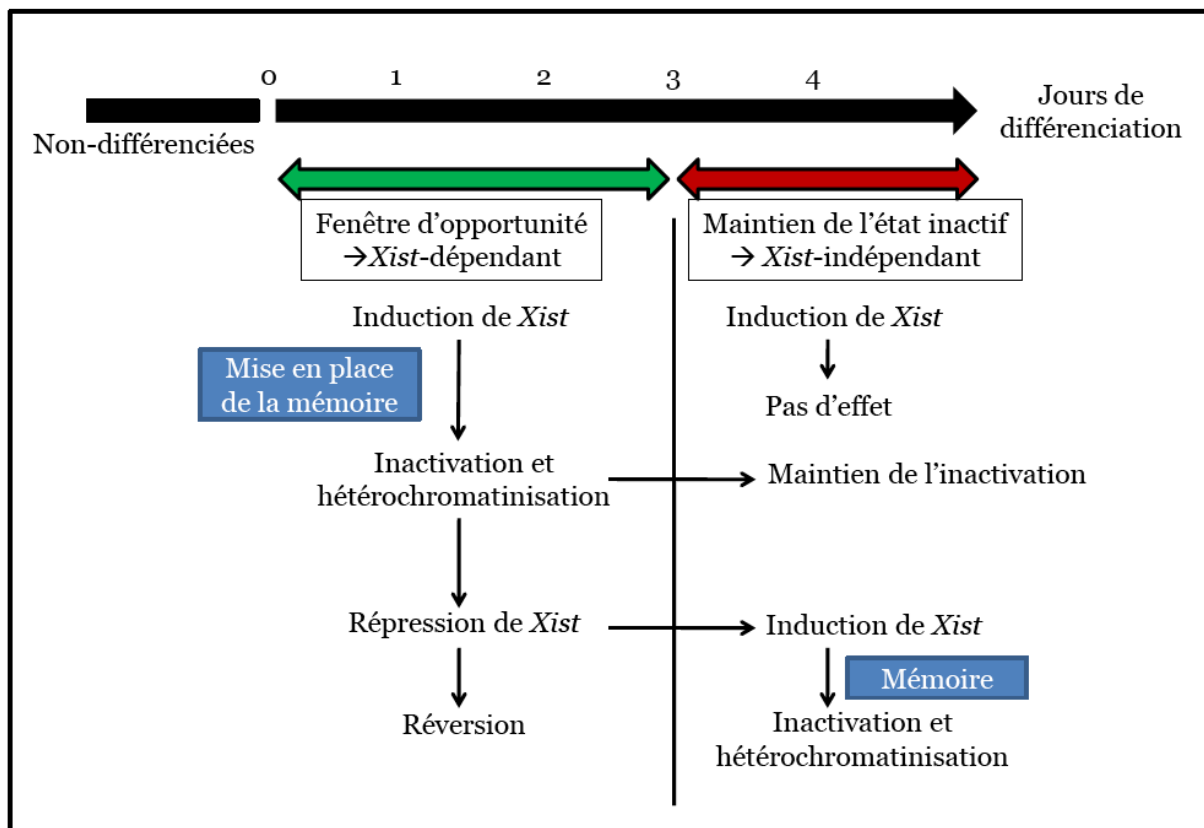


Figure 11 : Mise en évidence de l'existence fenêtré d'opportunité de l'inactivation ainsi que d'une mémoire chromosomique de l'état inactif induite par *Xist*. Représentation schématique de la différenciation des cellules ES. Lorsque *Xist* est induit pendant la fenêtré d'opportunité pour l'inactivation, le chromosome s'inactive en *cis* et s'enrichit en plusieurs marques typiques de l'hétérochromatine facultative. Cependant, l'induction de *Xist* après cette fenêtré d'opportunité n'est plus capable d'induire ni l'inactivation, ni l'hétérochromatinisation du X. Lorsque *Xist* est induit, puis réprimé, au cours de la fenêtré d'opportunité, sa réinduction tardive après la fenêtré d'opportunité est capable d'induire l'inactivation et l'hétérochromatinisation. Ceci indique que le premier événement d'inactivation a induit une mémoire chromosomique de l'état inactif qui a rendu compétent ce chromosome pour s'inactiver même après la fenêtré d'opportunité.

Ceci nous permet tout de même de conclure que *Xist* induit un état de répression épigénétique, dans le sens que la répression génique qu'il établit est ensuite perpétuée par des mécanismes indépendants de *Xist* lui-même, mais que cette initiation doit avoir lieu au cours d'une fenêtré de temps bien spécifique impliquant surement des facteurs développementaux. Passé ce laps de temps, il n'est pas clair s'il s'agit d'une inaptitude du chromosome X à être inactivé ou s'il s'agit d'une incapacité de *Xist* à induire l'inactivation. Cependant, des études récentes menées par le groupe d'Anton Wutz, ont apporté un éclaircissement sur ce dernier point.

Grâce à un travail précédent démontrant la capacité exceptionnelle de *Xist* à induire l'extinction transcriptionnelle au sein de cellules progénitrices du système hématopoïétique des souris

adultes lorsqu'exprimé ectopiquement (Savarese *et al.*, 2006), l'équipe de Wutz a pu identifier, suite à des analyses de transcriptome, *SatB1*, une protéine nucléaire impliquée dans la régulation de la structure chromatinienne et dans la régulation génique des thymocytes, comme protéine candidate pour expliquer cette particularité de l'absence de fenêtre de temps dans ces cellules. À travers diverses expériences ils ont réussi à démontrer que *SatB1/SatB2* (exprimés uniquement au cours des 3 premiers jours de différenciation cellulaire des cellules ES) sont des facteurs d'initiation de l'inactivation et permettent le rétablissement d'un contexte embryonnaire quant à l'extinction génique par *Xist* dans des cellules somatiques (Agrelo *et al.*, 2009). La localisation de *SatB1* en anneau autour du domaine de *Xist* dans les ES présagerait une interaction entre la protéine et l'ARN. De par le rôle connu de *SatB1* dans la jonction entre matrice nucléaire et MARs (pour « *Matrix Attachment Region* », qui sont des séquences servant à organiser la chromatine en domaines structuraux et jouant un rôle important dans l'expression génique) (Dickinson *et al.*, 1992 ; Cunningham *et al.*, 1994 ; Galande *et al.*, 2007 ; Cai *et al.*, 2006), il est possible que cette localisation serve à ancrer la chromatine au sein du compartiment hétérochromatique formé par *Xist* lors de la réorganisation chromatinienne qu'est l'inactivation du chromosome X. L'hypothèse des auteurs quant au fonctionnement de *SatB1* dans l'inactivation est que *Xist*, par l'intermédiaire du *Repeat-A* attire les régions géniques vers la zone de répression grâce aux modifications de la chromatine orchestrées par *SatB1*.

Malgré le rôle évident de *Xist* dans l'hétérochromatinisation du chromosome X inactif au cours de la différenciation cellulaire, ce transcrit pourrait aussi jouer un rôle particulier dans son maintien malgré certaines études déclarant le contraire (Brown & Willard, 1994 ; Csankovszki *et al.*, 1999 ; Wutz & Jaenisch, 2000). En 2007, Zhang *et al.* ont observé que *Xist* est nécessaire au maintien péri-nucléolaire du Xi, cette localisation particulière serait supposément nécessaire à la transmission mitotique de son état hétérochromatique (Zhang *et al.*, 2007) (Figure 12).

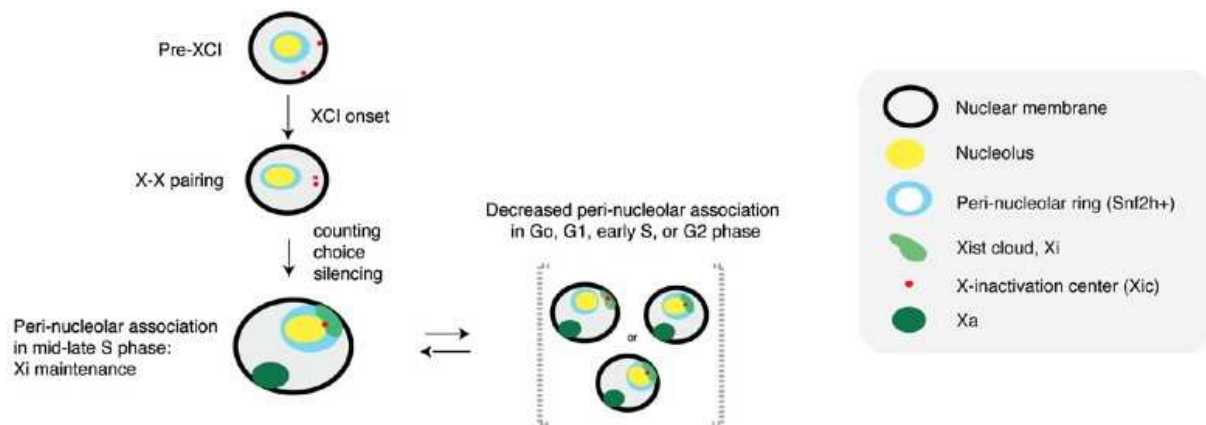


Figure 12 : Modèle proposant que la localisation périnucléolaire est requise pour la maintenance de l'état réprimé du chromosome X inactif. La cascade de l'inactivation du chromosome X est initiée lors de la transvection entre chromosomes X, qui permet un comptage correct ainsi qu'une désignation exclusive mutuelle de l'Xa et de l'Xi. Guidé par *Xist*, le Xi voyage vers la périphérie du nucléole, où il doit séjourner lors de la phase S moyenne à tardive afin de maintenir son état hétérochromatique. Le Xa est aussi associé à l'enveloppe nucléaire. (D'après Zhang *et al.*, 2007)

L'association péri-nucléolaire du Xi avait déjà été décrite de très longue date (Barr & Bertram, 1949 ; Manuelidis & Borden, 1988 ; Bourgeois *et al.*, 1985), bien qu'une association péri-nucléaire ait aussi été observée (Barton *et al.*, 1964 ; Bourgeois *et al.*, 1985 ; Dyer *et al.*, 1989 ; Belmont *et al.*, 1986 ; Eils *et al.*, 1996). L'association du Xi au nucléole dans près de 90% des cellules somatiques femelles est directement médiée par l'expression de *Xist*, les autosomes contenant des transgènes de l'ARN non-codant étant localisés au même endroit. Les auteurs proposent que cette compartementalisation nucléaire jouerait un rôle dans la réplication et la transmission de l'état réprimé du Xi au cours des mitoses, certains types d'hétérochromatine ayant déjà été décrits comme localisant de préférence au même endroit (O'Keefe *et al.*, 1992 ; Huang, 2000). Ils remarquent aussi que la zone péri-nucléolaire est connue pour son enrichissement de la protéine Snf2h, la sous-unité catalytique du complexe ACF1-ISWI qui est nécessaire à la progression de la fourchette de réplication au sein de chromatine condensée (Collins *et al.*, 2002). Ce rôle dans le maintien de l'inactivation au sein de cellules différenciées n'est pas forcément contradictoire avec les études précédentes stipulant que *Xist* n'est pas nécessaire à cette fonction. En effet, ces études ne montrent pas une réactivation brutale de l'Xi mais plutôt un processus lent et passif, nécessitant de multiples passages cellulaires pour apercevoir une réactivation sporadique. Cela nous indique qu'il existe d'autres mécanismes de maintien de l'état inactif au cours de la cellule. De manière intéressante, un autre groupe a démontré une localisation péri-nucléolaire d'un autre domaine chromosomique

hétérochromatique médié par un ARN non-codant exprimé de manière empreintée : l'ARN *Kcnq1ot1* (Mohammad *et al.*, 2008) et suppose aussi que cette localisation est responsable de la transmission de l'état répressif. Ceci conforte donc l'hypothèse de Zhang *et al.*, et laisse supposer qu'il existerait peut-être un mécanisme plus global de localisation péri-nucléolaire de régions transcriptionnellement réprimées par des phénomènes épigénétiques tels que l'empreinte parentale, l'inactivation du chromosome X ou tout simplement par des ARNs non-codants.

b. Suite de l'extinction transcriptionnelle

Lors des premiers instants de l'initiation de l'inactivation du chromosome X dans des cellules ES femelles, suite à la surexpression de *Xist*, on observe l'exclusion de l'ARN polymérase II ainsi que de l'ARN Cot-1 (ces séquences étant composées de séquences hautement répétées et d'éléments non-codants exprimés, tels que les retrotransposons, les séquences répétées centromériques, intergéniques ou introniques) du domaine formé par l'ARN non-codant. Ce processus est orchestré de manière identique aussi bien lors de l'inactivation empreintée qu'aléatoire (Chaumeil *et al.*, 2006, Namekawa *et al.*, 2010, Clemson *et al.*, 2006, Okamoto *et al.*, 2005). Parallèlement à ce remodelage structural du chromosome X inactif, on assiste à une réorganisation épigénétique de celui-ci non moins importante dans l'acquisition d'un chromosome X inactif stable (Figure 13).

c. Les modifications d'histones

i. Généralités

Comme nous avons déjà pu brièvement l'apercevoir au cours de l'introduction, la modification des histones joue un rôle prépondérant dans la régulation épigénétique du génome. Les histones sont des protéines qui vont s'assembler en octamère (généralement l'histone H2A, H2B, H3 et H4, chacun en double exemplaire) constituant ainsi un nucléosome (Luger *et al.*, 1997). Il faut aussi savoir qu'il existe plusieurs variants d'histones qui interviennent en fonction du contexte biologique (répression transcriptionnelle, réparation

de lésions de l'ADN, etc...). La double hélice d'ADN va s'enrouler autour des nucléosomes (stabilisée par

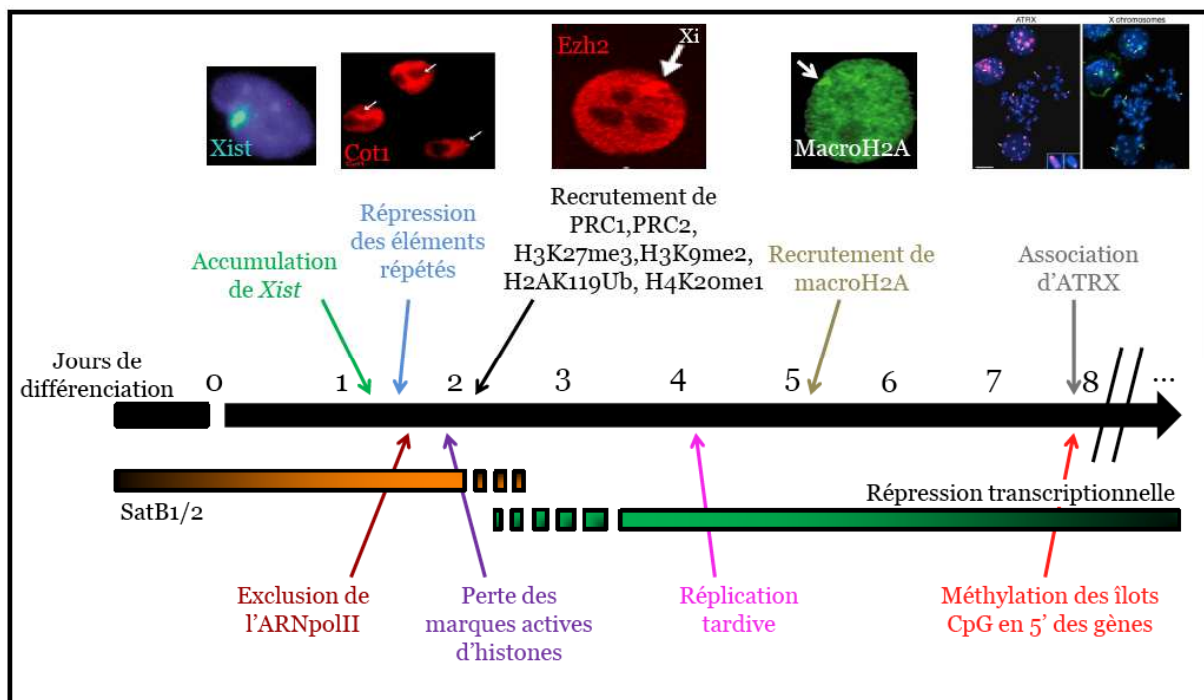


Figure 13 : Cinétique de l'inactivation aléatoire du chromosome X. Divers phénomènes épigénétiques se mettent en place sur le chromosome X étant choisi comme inactif au cours de la différenciation cellulaire. Ces événements se mettent en place à des moments précis, sans doute dû à leur fonction dans la mise en place ou le maintien de l'inactivation chromosomique. En barre orange est représentée la période de temps au cours de laquelle est exprimé SatB1, correspondant à la période pendant laquelle l'expression de Xist a la capacité d'induire la répression transcriptionnelle (Agrelo *et al.*, 2010). La barre verte représente la période au cours de laquelle on observe une extinction transcriptionnelle progressive des gènes du Xi.

l'histone H1 qui joue aussi un rôle dans l'espacement des nucléosomes et leur compaction, Zlatanova & Van holde, 1998) afin d'obtenir une compaction optimale. Cependant, ces histones peuvent être modifiées post-traductionnellement de manière covalente, le plus souvent sur leur queue N-terminale (Fischle *et al.*, 2003 ; Turner, 2002).

Ces modifications sont variées, comprenant des méthylations, des acétylations, des phosphorylations, des ADP-ribosylations, des biotinylation, des sumoylation ou encore des ubiquitination, chacune d'entre-elles ayant sa répercussion différente en *cis*, sur le contexte structural du nucléosome et de l'ADN qui l'entoure, mais aussi et surtout en *trans*, en recrutant des protéines capables notamment d'agir sur la régulation transcriptionnelle des gènes avoisinants (pour revue voir Jenuwein & Allis, 2001 ; Margueron *et al.*, 2005 ; Strahl & Allis, 2000). De cette manière, au cours du développement embryonnaire mais aussi au cours de la vie d'un individu, l'expression d'un gène peut être rapidement modifiée afin de

permettre un comportement approprié face à un environnement changeant. Ces modifications participent aussi à la conservation de l'identité transcriptionnelle entre cellule mère et cellule fille, en conservant la combinaison de gènes exprimés ou réprimés au cours des divisions cellulaires. Le paradigme d'une telle conservation est le chromosome X inactif. L'hétérochromatine est souvent définie comme étant une organisation compacte de la chromatine, moins accessible aux facteurs et machineries de transcription par opposition à l'euchromatine, une structure associée à la portion du génome riche en gènes, généralement transcriptionnellement active et répliquée précocement au cours de la phase S, ces deux situations étant parfaitement reflétées par le Xa et le Xi.

Au vu des multiples combinaisons de modifications d'histones possibles, que ce soit d'une histone à une autre, d'un acide aminé à un autre ou même d'une modification à une autre sur le même acide aminé d'une histone, l'idée d'un "code histone" a été proposée par Strahl et Allis (Strahl & Allis, 2000) comme possibilité pour expliquer les conséquences de ces combinaisons complexes. Ces modifications forment donc un véritable réseau, aussi bien capables de recruter des protéines effectrices que d'influer sur les modifications voisines, certaines étant incompatibles avec d'autres ou bien au contraire, favorisant le dépôt d'autres marques renforçant ainsi son statut épigénétique. Et ce tableau ne s'est pas simplifié lorsqu'en 2006 Bernstein, *et al.* ont décrit l'existence de « domaines bivalents » aux loci de certains gènes développementaux clés dans les cellules ES non-différenciées, comprenant aussi bien des marques hétérochromatiques (H3K27me3) qu'euchromatiques (H3K4me3) (Bernstein *et al.*, 2006 ; Azuara *et al.*, 2006 ; Stock *et al.*, 2007). De plus chez certaines espèces (telle que *S. cerevisiae*), des marques typiquement activatrices (H3K4me) se sont révélées aussi avoir un rôle répresseur (Mueller *et al.*, 2006, Pinskaya *et al.*, 2009). Bien que ce genre de dualité fonctionnelle n'ait pas encore été décrite chez les euthériens, cela complique davantage la possibilité de déchiffrer un éventuel "code" des modifications d'histones.

De manière simplifiée, l'hyperacétylation de résidus sur les histones H3, H4, H2A et H2B, mais également l'hyperméthylation de certaines lysines (en particulier K4 et K79) de H3, sont généralement associées à des gènes transcrits, tandis que l'hyperméthylation de résidus sur les histones H3 et H4 (en particulier K9 et K27 de H3, et K20 de H4 et à quelques exceptions près, dont K4 de H3) correspond en général à des gènes transcriptionnellement inactifs. La phosphorylation des queues d'histones est plutôt impliquée dans la mitose ou la réparation de l'ADN (variant d'histone H2AX), et l'ubiquitination peut être aussi bien associée à

l'activation transcriptionnelle (K120 de H2B) qu'à sa répression (K119 de H2A).

Bien entendu, de nombreuses protéines sont liées au dépôt ou à la reconnaissance de chacune de ces marques et possèdent donc des domaines fonctionnels adéquats. Par exemple les protéines possédant un chromodomaine sont impliquées dans la méthylation, de la même manière que les protéines à bromodomaine s'associent aux acétylations.

Grâce aux avancées réalisées dans le perfectionnement des méthodes utilisées dans l'étude de la distribution de ces marques épigénétiques, telles que l'immunofluorescence, le FISH, ou bien plus récemment le ChIP-on-CHIP (Horak & Snyder, 2002) et le ChIP-Seq (Mardis, 2007), il est devenu très aisé d'étudier celle-ci au niveau de son gène ou chromosome préféré, ou bien du génome entier.

Comme nous venons de le voir, l'acquisition de « marques » (modifications covalentes de leurs queues) par les histones est donc intimement liée à deux notions essentielles : la stabilité de l'information portée par la chromatine et sa transmission au cours des divisions cellulaires d'une part, et le recours à une séquence d'évènements distincts mais coordonnés pour établir une mémoire épigénétique, d'autre part. Au cours du chapitre suivant, nous allons donc voir de quelle manière les modifications d'histones participent à ces deux notions au cours de l'établissement et du maintien de l'inactivation du chromosome X.

ii. Modifications d'histones sur le Xi

Suite au recouvrement du Xi par le transcrit *Xist*, puis par l'exclusion de l'ARN pol II, on assiste tout d'abord à une perte des marques d'histones euchromatiques : H3K9Ac, H3K4me2, H3K4me3 (Chaumeil *et al.*, 2002, Goto *et al.*, 2002, Heard *et al.*, 2001, Okamoto *et al.*, 2004) suivie de près par l'hypoacétylation des histones H4 (Chaumeil *et al.*, 2002, Heard *et al.*, 2001, Keohane *et al.*, 1996). Puis, de manière quasi-simultanée, on observe une apparition de marques hétérochromatiques avec l'enrichissement de H3K27me3 (Okamoto *et al.*, 2004 ; Plath *et al.*, 2003 ; Silva *et al.*, 2003 ; Rougeulle & Avner, 2003), H3K9me2 surtout au niveau des régions promotrices des gènes (Heard *et al.*, 2001 ; Mermoud *et al.*, 2002 ; Okamoto *et al.*, 2004 ; Rougeulle *et al.*, 2003, 2004), H4K20me1 (Kohlmaier *et al.*, 2004) et H2AK119ub1 (Fang *et al.*, 2004 ; Hernandez-Munoz *et al.*, 2005 ; Smith *et al.*, 2004 ; de Napoles *et al.*, 2004) (Figure 14).

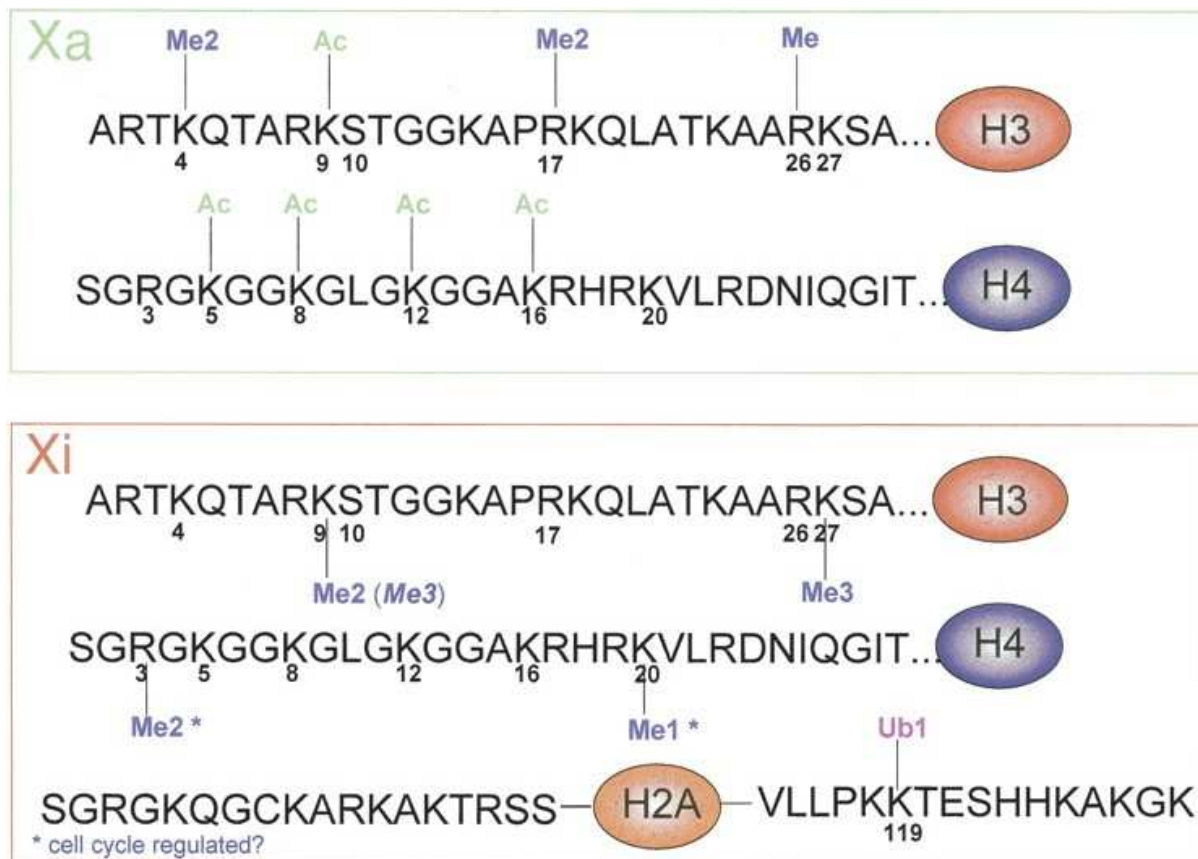


Figure 14 : Marques épigénétiques associées à l'inactivation du chromosome X. Le statut des modifications d'histones globalement associées au chromosome X inactif (Xi) ou de son homologue actif (Xa) est montré sur des cartes des acides aminés composant les queues N- et C-terminales des histones indiquées. (D'après Heard & Distèche, 2006)

De manière significative, il a été montré pour la plupart de ces marques qu'elles restaient associées au chromosome tout au long du cycle cellulaire et également sur les chromosomes métaphasiques, ce qui est cohérent avec le fait qu'elles pourraient participer à la transmission héréditaire de l'état inactif (Mermoud *et al.*, 2002 ; Boggs *et al.*, 2002 ; Chadwick & Willard, 2004 ; Mak *et al.*, 2002 ; Chaumeil *et al.*, 2002).

Malgré le fait qu'une multitude de marques hétérochromatiques soient recrutées sur le Xi lors de l'initiation de l'inactivation, il semblerait que la plupart d'entre-elles ont une activité redondante par rapport aux autres, rendant leur présence individuelle dispensable.

L'exemple typique de cette hypothèse est la marque répressive par excellence H3K27me3. Le dépôt de cette marque, effectué par le complexe PRC2, corrèle de très près avec l'expression de *Xist*, n'apparaissant que lors de sa surexpression et disparaissant lors de son excision (dans les cellules somatiques) (Zhang *et al.*, 2007 ; Plath *et al.*, 2003 ; Silva *et al.*, 2003). Pourtant, l'absence de cette marque engendrée par la mutation d'une sous-unité (*Eed*) de PRC2

n'engendre aucun défaut, ni de l'initiation, ni du maintien de l'inactivation aléatoire (Silva *et al.*, 2003, Kalantry & Magnuson, 2006), bien qu'on assiste à une réactivation partielle du Xi lors de l'inactivation empreintée (sans doute due à un état inactif plus facilement réversible, notamment grâce à l'absence de méthylation de l'ADN) (Wang *et al.*, 2001 ; Mak *et al.*, 2002). La marque H3K27me3 est aussi connue comme signal de recrutement du complexe protéique PRC1, qui est aussi enrichi sur le Xi de manière précoce (de Napoles *et al.*, 2004 ; Fang *et al.*, 2004). PRC1 a non seulement un effet sur la compaction chromatinienne (par sa sous-unité Ring1B) (Francis *et al.*, 2004) mais sa fixation engendre aussi la modification répressive H2AK119ub1, notamment impliquée dans le blocage de l'élongation transcriptionnelle (Stock *et al.*, 2007). Cependant, il semblerait que l'ARN *Xist* soit capable de recruter ces deux complexes au Xi, indépendamment l'un de l'autre ainsi que de la répression génique (Schoeftner *et al.*, 2006, Kohlmaier *et al.*, 2004), et ce, de manière précoce au cours de l'inactivation. Cette capacité de certains ARNs non-codants à recruter les complexes protéiques polycomb est déjà bien documentée (Hekimoglu & Ringrose, 2009 ; Khalil *et al.*, 2009 ; Rinn *et al.*, 2007 ; Yap *et al.*, 2010).

Bien que le recrutement de PRC2 au Xi soit partiellement indépendant du *Repeat-A* (Plath *et al.*, 2003 ; Kohlmaier *et al.*, 2004), il semblerait que cette région répétée soit impliquée dans son recrutement dès le stade de cellule ES non-différenciée (Kanhare *et al.*, 2010 ; Zhao *et al.*, 2008), grâce à un petit transcrit non-codant émanant du *Repeat-A* nommé *RepA*. Ce recrutement par la région répétée *Repeat-A* de PRC2 pourrait être partiellement responsable de la propagation de la marque H3K27me3 le long du Xi lors de l'inactivation (expliquant peut-être la faiblesse du marquage observé en absence de la région répétée, Kohlmaier *et al.*, 2004).

En conclusion, nous pouvons donc supposer que ni PRC1, ni PRC2 ainsi que leurs marques associées ne sont indispensables à la mémoire répressive nécessaire au maintien de l'inactivation aléatoire du chromosome X (Kalantry & Magnuson, 2006 ; Leeb & Wutz, 2007). Pourtant, leur recrutement précoce, corrélé à l'expression et à la localisation de *Xist*, est compatible avec un rôle dans l'établissement d'une structure hétérochromatique qui fonctionne comme une mémoire épigénétique, cette mémoire étant stable et pouvant s'auto-perpétuer de manière indépendante de *Xist*, PRC1 et PRC2. Ces conclusions troublantes pourraient tout simplement être expliquées par l'observation d'un phénomène de redondance entre les complexes protéiques (Leeb *et al.*, 2010), ou bien par la présence de paralogues compensant

pour les absences de certaines protéines lorsque celles-ci sont mutées (de Napoles *et al.*, 2004 ; Leeb & Wutz, 2007). L'accumulation de marques hétérochromatiques sur le Xi pourrait donc être une sorte de "garantie" pour la cellule afin de s'assurer une structure répressive (que ce soit pour les séquences répétées ou géniques) efficace et durable (Rougeulle & Avner, 2003). L'étude des double-mutants PRC1 et PRC2 dans l'inactivation du chromosome X nous fournira des indications précieuses quant aux rôles exacts de ces marques.

De plus il est important de noter que, bien que l'ajout de marques hétérochromatiques aux histones est important à la formation d'une structure chromatinienne répressive, la disparition des marques euchromatiques l'est sûrement tout autant sinon plus, comme le propose Bryan Turner, qui en se servant d'un inhibiteur d'histone déacétylase (la trichostatine A) au cours de la différenciation de cellules ES femelles a réussi à empêcher la formation d'un X inactif (O'Neill *et al.*, 1999). De plus, il se trouve justement que Pasini *et al.* ont démontré l'implication de PRC2 dans la déméthylation de la marque euchromatique H3K4me3 à travers son recrutement de la déméthylase Rbp2 (Jarid1a) (Pasini *et al.*, 2008), la déméthylation en H3K4me3 sur les promoteurs étant en effet corrélée au dépôt de H3K27me3 sur l'ensemble du gène (Marks *et al.*, 2009).

d. Autres modifications épigénétiques liées à l'Xi

i. MacroH2A

Parmi les modifications épigénétiques liées aux histones, outre les marques des queues N-terminales qui peuvent varier en fonction du contexte et du rôle à jouer, le recrutement de variants d'histones peut aussi participer de manière non-négligeable à la modification de l'expression génique.

Le variant d'histone macroH2A est composé d'un domaine N-terminal très similaire à celui de l'histone H2A mais aussi d'un grand domaine C-terminal de fonction inconnue. Lors de l'induction de la différenciation cellulaire, la localisation de macroH2A passe du centrosome au chromosome X inactif (Costanzi & Pehrson, 1998, 2001 ; Chadwick & Willard, 2001 ; Mermoud *et al.*, 1999 ; Rasmussen *et al.*, 2000) et ce, de manière uniforme tout le long du Xi

(Mietton *et al.*, 2009), formant ainsi ce qui est appelé un « *macrochromatin body* » (ou corps macro-chromatinien).

Le recrutement de macroH2A, comme celui de H3K27me3, dépendrait de l'expression du gène *Xist*, mais ne serait pas non plus nécessaire au maintien de l'état hétérochromatique du Xi (Csankovski *et al.*, 1999).

De manière similaire aux complexes protéiques PRC1 ainsi que PRC2 (et leurs marques associées), le variant d'histone macroH2A serait donc un facteur répressif (Doyen *et al.*, 2006, Agelopoulos *et al.*, 2006, Gilbert *et al.*, 2000, Angelov *et al.*, 2003, Nusinow *et al.*, 2007) supplémentaire, dont le rôle est de participer à l'établissement collectif de la répression transcriptionnelle du chromosome X inactif assurée par un système inhibiteur basé sur la redondance fonctionnelle afin de s'assurer de son bon fonctionnement. Toutefois, sa localisation uniforme le long du Xi nous laisse supposer que macroH2A agit plutôt sur la structure chromatinienne qu'en réprimant directement la transcription au niveau des promoteurs (Mietton *et al.*, 2009).

ii. Méthylation de l'ADN

Une autre modification épigénétique du génome est la méthylation de l'ADN. Comme son nom l'indique cette marque n'est plus associée aux histones mais consiste en le fait d'ajouter un groupement méthyle aux cytosines lorsque ces dernières précèdent une guanine (notée CpG). Ces dinucléotides se trouvent répartis sur tout le génome mais sont parfois présents à une très forte densité au sein d'une région précise, ceci constituant ce qui est appelé : un îlot CpG. On retrouve ces îlots CpG, le plus souvent, dans les régions promotrices des gènes, et leur statut de méthylation reflète de manière quasi-invariable leur statut transcriptionnel : la méthylation de l'îlot étant synonyme de répression et *vice-versa*. En revanche, les régions codantes des gènes peuvent être plus ou moins méthylées quelque soit leur niveau d'expression. Les séquences répétées et intergéniques sont, quant à elles, globalement hyperméthylées (sûrement dans le but d'empêcher la relocalisation de certaines séquences répétées ayant conservé leur capacité à se déplacer).

À ce jour, 5 enzymes possédant un domaine ADN méthyltransférase ont été décrites : Dnmt1, Dnmt2 (Goll & Bestor, 2005), Dnmt3a, Dnmt3b, Dnmt3l (Bourc'his *et al.*, 2001, 2004, Hata *et al.*, 2002). Seuls Dnmt1, 3a et 3b semblent posséder un rôle dans l'inactivation. Dnmt1

possède une activité ADN méthyltransférase de maintenance, capable de reconnaître les dinucléotides CpG hémiméthylés issus de la réplication semi-conservative de l'ADN et de convertir les nouvelles cytosines incorporées en 5-méthylcytosines. Sa très faible affinité pour les CpG non-méthylés nous laisse supposer qu'elle joue essentiellement un rôle dans le processus de mémoire cellulaire. Dnmt3a et 3b ont la capacité de transférer les groupements méthyles à des substrats hémi- ou non-méthylés (Okano *et al.*, 1998). Ces protéines sont essentielles au développement murin (Okano *et al.*, 1999). Elles interviennent lors de la vague de méthylation du génome ayant lieu durant les stades post-implantatoires et sont également impliquées dans la méthylation *de novo* des gènes porteurs d'une empreinte maternelle dans les ovocytes (Chen *et al.*, 2003, Okano *et al.*, 1999). De plus, dans les cellules ES, ces gènes jouent un rôle dans l'établissement ainsi que le maintien des profils de méthylation du génome (Chen *et al.*, 2003) et pourraient donc participer à la mise en place de nouvelles empreintes génomiques comme l'inactivation du chromosome X.

Sur le Xi, les îlots CpG situés au niveau des promoteurs des gènes sont hyperméthylés, sauf ceux des gènes échappant à l'inactivation (Distèche *et al.*, 1997), les promoteurs de ces derniers se trouvant donc hypométhylés sur les deux chromosomes X (Mondello *et al.*, 1988). Cette méthylation des promoteurs est une marque plutôt tardive de l'inactivation (Lock *et al.*, 1987), apparaissant aux alentours de 8 jours de différenciation cellulaire (à l'acide rétinoïque), et serait donc vraisemblablement impliquée dans le maintien de la répression transcriptionnelle du Xi plutôt que dans l'initiation. Sur le Xa, c'est plutôt le corps des gènes ainsi que leur régions 3' qui se trouvent méthylés (Bartlett *et al.*, 1991, Lock *et al.*, 1987, Yen *et al.*, 1984).

L'importance de la méthyl-transférase de maintien Dnmt1 dans le maintien de l'inactivation est démontrée dans les travaux orchestrés par le laboratoire d'En Li (Sado *et al.*, 2000), une réactivation partielle du Xi étant observée chez les mutants *Dnmt1*^{-/-}. Confirmant, l'absence de méthylation de l'ADN dans le processus d'inactivation empreintée, les auteurs n'observent aucune différence quant à ce phénomène au sein d'embryons pré-implantatoires.

Paradoxalement, la même équipe n'observe aucun effet d'un double mutant *Dnmt3a*^{-/-} et *Dnmt3b*^{-/-} sur l'inactivation du chromosome X dans son ensemble (Sado *et al.*, 2004). L'absence d'effet de cette double-mutation est plutôt inattendue, car Dnmt1 ne méthyle pas l'ADN *de novo*, contrairement à Dnmt3a/3b, rendant le rôle de Dnmt1 dans l'inactivation à

travers la méthylation de l'ADN contradictoire avec ces derniers résultats. Les auteurs estiment que cette divergence pourrait être due à une sensibilité accrue à l'hypométhylation de la part du transgène (Sado *et al.*, 2000) par rapport aux gènes endogènes à l'X (Sado *et al.*, 2004).

Il existe diverses propositions de mécanismes liant la méthylation de l'ADN à la répression transcriptionnelle. Tout d'abord, la méthylation de l'ADN peut empêcher certains facteurs de transcription (dits "sensibles" à la méthylation) de se fixer à leur site de reconnaissance, inhibant ainsi leur rôle d'activateur transcriptionnel, comme c'est par exemple le cas pour les protéines ETS1, CTCF ou YY1 (à certains de leurs sites de fixation) (Maier *et al.*, 2003 ; Bell & Felsenfeld, 2000 ; Hark *et al.*, 2000 ; Kanduri *et al.*, 2000 ; De la Rosa-Velazquez *et al.*, 2007 ; Rodriguez *et al.*, 2010 ; Satyamoorthy *et al.*, 1993 ; Kim *et al.*, 2003). Ce blocage de la fixation de la machinerie transcriptionnelle ou de leurs recruteurs/activateurs, pourrait aussi être obtenu si la méthylation de l'ADN empêche, par déstabilisation stérique, la formation de structures en boucle établies au sein des îlots CpG, qui faciliteraient l'accès aux séquences promotrices. L'autre possibilité est que l'ADN méthylé ou bien les ADN méthyltransférases interagissent avec d'autres protéines servant à renforcer la structure hétérochromatique telles que MeCP2 (« methylcytosine-binding protein 2 ») ou Mbd2 (« Methyl-CpG-binding domain protein 2 ») (Nan *et al.*, 1997 ; Nan *et al.*, 1998 ; Jones *et al.*, 1998 ; Feng & Zhang, 2001 ; Ng *et al.*, 1999).

Deux autres protéines ont récemment été décrites comme pouvant influencer sur les niveaux de méthylation de l'ADN du chromosome X inactif : ATRX et SmcHd1.

ATRX est été impliquée dans le contrôle de la méthylation de l'ADN (Gibbons *et al.*, 2000 ; McDowell *et al.*, 1999 ; Garrick *et al.*, 2006) mais aussi dans le recrutement de la protéine MeCP2 (Kernohan *et al.*, 2010 ; Nan *et al.*, 2007) et dans l'interaction avec la sous-unité Ezh2 du complexe PRC2 (responsable de H3K27me3) chez l'homme (Cardoso *et al.*, 1998, Kourmouli *et al.*, 2005). Sa localisation sur le Xi est observée de manière tardive après 8 jours de différenciation spontanée (correspondant à un moment situé entre le recrutement de macroH2A et la méthylation de l'ADN du Xi) (Baumann & De La Fuente, 2009) mais aussi au sein des tissus extra-embryonnaires où sa mutation engendre des problèmes d'inactivation

et de développement (Garrick *et al.*, 2006). Son rôle dans l'inactivation aléatoire est inconnu. Smchd1 (pour « *Structural Maintenance of Chromosomes Hinge Domain containing 1* ») est, quant à elle, impliquée dans la maintenance de l'inactivation du chromosome X à travers un rôle dans l'hyperméthylation des îlots CpG associés au Xi (Blewitt *et al.*, 2008). Cette protéine contient un « *SMC Hinge* » généralement impliqué dans la condensation chromosomique, la cohésion, la ségrégation des chromosomes ainsi que la recombinaison et la réparation de l'ADN. Une mutation de cette protéine engendre la réactivation d'un transgène codant pour la GFP à partir du Xp aussi bien dans les tissus extra-embryonnaires que dans l'embryon proprement dit. Ils ont aussi remarqué, dans ces cellules, que l'intégralité des îlots CpG étudiés, parmi ceux qui sont soumis à l'inactivation, étaient hypométhylés au point de retrouver les niveaux présents sur le chromosome X actif mâle, corrélant avec une expression ectopique de la plupart des gènes associés (contrairement aux gènes autosomaux dont l'expression est invariante dans ces mutants). La méthylation de l'ADN du groupe de gènes *Rhox*, connu comme étant non-variant au cours de l'inactivation arrive à maintenir cette caractéristique dans les mutants pour *Smchd1*. Cela nous indique donc que Smchd1 joue uniquement un rôle dans la méthylation des îlots CpG en réponse à l'inactivation du chromosome X dissociant éventuellement un mécanisme de méthylation médié par *Xist* d'une méthylation indépendante, mise en place plus tôt au cours du développement embryonnaire. De plus, grâce à des expériences d'immunofluorescence, ils ont réussi à prouver que Smchd1 colocalisait avec le Xi au sein de cellules fibroblastiques femelles.

Paradoxalement, la protéine Smchd1 semble donc être impliquée non seulement dans la maintenance de l'inactivation aléatoire mais aussi dans l'inactivation empreintée à travers un mécanisme de dépôt de la méthylation de l'ADN ce qui paraît paradoxal sachant que celle-ci n'est pas nécessaire à la maintenance de l'inactivation empreintée (Sado *et al.*, 2000). Ceci laisse supposer un deuxième rôle de cette protéine dans l'inactivation, de manière indépendante de la méthylation de l'ADN (tout du moins dans les tissus extra-embryonnaires).

La découverte du rôle exact de ces deux protéines dans la cellule ainsi que leur fonctionnement seront sûrement d'un grand intérêt pour la compréhension de la manière dont s'établit le maintien de la répression génique sur le Xi. Et il s'avère que la méthylation de l'ADN sur le X inactif semblerait dépendre de beaucoup plus de protéines que Dnmt1. De plus, nous pouvons

aussi noter que ce type de protéine (à domaine SMC) est impliqué dans le phénomène de compensation de dose chez le nématode *C. elegans* (Chuang *et al.*, 1994) nous fournissant ainsi un lien particulièrement excitant entre deux formes de compensation de dose, toutes deux basées sur la répression transcriptionnelle.

iii. Divers

Et enfin, deux autres protéines semblent impliquées dans la maintenance de l'inactivation : ATM (« *Ataxia Telangiectasia Mutated* ») et ATR (« *ATM and Rad3 related* ») (Ouyang *et al.*, 2005). Ces deux protéines kinases ont été initialement caractérisées à travers leurs rôles dans la réparation de l'ADN, la stabilité génomique et la recombinaison. Cependant il a aussi été remarqué qu'elles sont associées à des histones déacétylases en absence de dommage à l'ADN (Schmidt & Schreiber, 1999 ; Wang *et al.*, 2000). En inhibant ATM/ATR (grâce à la 2-aminopurine), Ouyang *et al.* ont aperçu une augmentation de l'acétylation de l'histone H4 sur l'ensemble du Xi ainsi que la réactivation dans certaines cellules d'un transgène GFP inséré dans un des chromosomes X. L'interférence des deux protéines aboutit aussi à une réactivation du transgène dans une fraction de cellules et ceci de manière indépendante du rôle de ces protéines dans la réparation de l'ADN. De manière curieuse, aucune localisation d'ATM sur le Xi n'a pu être observée au sein de cellules humaines, indiquant que leur rôle dans le maintien de l'inactivation s'effectuerait sûrement en amont de la modification de la chromatine.

iv. Réplication Asynchrone

Une dernière marque caractéristique du chromosome X inactif est sa réplication tardive au cours de la phase S (Taylor, 1968 ; Takagi & Sasaki, 1975 ; Morishima *et al.*, 1962), cette caractéristique étant l'un des critères les plus employés pour discriminer le Xi du Xa. En effet, les chromosomes X actif et inactif se répliquent de manière asynchrone et cette différence est très bien conservée chez l'ensemble des mammifères ainsi que les marsupiaux (Sharman, 1971 ; Cooper & Hope, 1993). Cette modification épigénétique se met en place entre 2 et 4 jours après initiation de la différenciation (Chaumeil *et al.*, 2002 ; Okamoto *et al.*, 2005 ; Keohane *et al.*, 1999). Des études par approche locus-spécifique, ont ainsi permis d'observer que les gènes échappant à l'inactivation situés sur le Xi sont répliqués de façon synchrone aux gènes

du X actif en phase S précoce (Boggs & Chinault, 1994 ; Hansen *et al.*, 1996 ; Kawame *et al.*, 1995 ; Torchia *et al.*, 1994). L'exception à ces observations est le gène *Xist*, qui est répliqué tardivement à partir du Xi malgré le fait qu'il est transcriptionnellement actif (Gartler *et al.*, 1999).

Le phénomène de réplication tardive semble être caractéristique des régions non-transcrites (Gilbert, 2002). La réplication tardive du Xi semble parfaitement corrélée avec son état transcriptionnel global. En effet, au sein de mutants conditionnels de *Xist* la réplication tardive est maintenue (Csankovski *et al.*, 1999). Par contre lors de la réactivation partielle du Xi suite à l'utilisation de 5-azacytidine, on retrouve une synchronisation de la réplication des deux chromosomes X (Hansen *et al.*, 1996) (indépendamment de Dnmt3a et 3b, Sado *et al.*, 2004).

Le rôle précis de ce phénomène est encore inconnu, cependant nous pouvons supposer qu'il pourrait constituer un moyen d'isoler temporellement les deux chromosomes X afin de faciliter leur traitement différentiel au sein d'un même noyau, que ce soit de façon à minimiser le contact avec la machinerie de transcription ou encore maximiser l'accessibilité aux enzymes de modification de la chromatine, et contribuer ainsi au maintien de l'état inactif.

Bien que, comme nous venons de le voir, la délétion conditionnelle de *Xist* sur le Xi ne re-synchronise pas la réplication de ce chromosome avec son homologue, un mutant conditionnel pour la délétion de *Xist* sur le Xa induit la réplication retardée de ce dernier, indépendamment de la présence d'un second chromosome X (Diaz-Perez *et al.*, 2005). Une possibilité pour expliquer ceci serait que cette région d'ADN délétée correspondant à *Xist*, soit nécessaire pour la stabilité de l'ADN du chromosome X en *cis*, et qu'en son absence la fragilité de l'ADN engendre une réplication tardive du Xa (Diaz-Perez *et al.*, 2006, Hansen *et al.*, 1993, Hellman *et al.*, 2000, Arlt *et al.*, 2003), une des protéines responsables du maintien de la stabilité de l'ADN aux sites fragiles lors de la réplication étant curieusement ATR (Casper *et al.*, 2002).

v. Conclusion sur les différentes marques épigénétiques

Tout au long de ce chapitre sur les marques caractérisant le chromosome X inactif chez la souris, nous avons pu voir à quel point la cellule emploie un nombre pharamineux de protéines afin d'établir une structure hétérochromatique. Ces protéines ont parfois des fonctions qui leur sont propres et parfois qui sont redondantes, mais dans tous les cas, l'accumulation d'autant

de marques hétérochromatiques doit sans doute avoir une explication du point de vue évolutif. S'agit-il d'un système qui accumule des mécanismes de répression imparfaits ou trop faibles à eux seuls, d'où la nécessité d'en associer autant autour d'un même objectif ? Ou s'agit-il d'un système de prévention, au cas où l'une des marques ne peut être réalisée, les autres pouvant la remplacer ou tout du moins la compenser ? Les deux hypothèses n'étant pas exclusives.

L'idée d'une redondance ou d'une coopération entre les marques hétérochromatiques afin d'initier et de maintenir la répression transcriptionnelle sur le Xi est justement renforcée par les études de Csankovszki et d'Hernandez-Munoz (Csankovszki *et al.*, 1999 *et al.*, 2001, Hernandez-Munoz *et al.*, 2005).

Ces équipes ont entrepris d'étudier les interactions potentiellement synergistiques entre différentes marques répressives grâce à l'utilisation de différents mutants ($\Delta Xist$ et $\Delta Dnmt1$), de différents produits chimiques (la 5-azacytidine qui inhibe les ADN méthyltransférases et la Trichostatine A qui inhibe les déacétylases d'histones) ou en présence de ARNi anti-macroH2A au sein des cellules fibroblastiques murines (MEFs) femelles.

En règle générale, ils observent une quantité croissante de cellules ayant réactivé le transgène lors de l'accumulation d'inhibiteurs de l'hétérochromatinisation (Figure 18).

Taux de réactivation du transgène						
Xist	Recrutement de macroH2A	ARNi macroH2A	Méthylation de l'ADN	Hypoacétylation des histones	Réactivation GFP	Facteur d'augmentation
+	+	+	+	+	0,01%	
+	+	+	+	TSA	0,01%	x1
+	-	ARNi	+	+	0,01%	x1
-	-	+	+	+	0,04%	x4
+	+	+	5-azacytidine	+	0,20%	x20
-	-	+	5-azacytidine	+	0,30%	x30
-	-	+	5-azacytidine	TSA	0,60%	x60
+	+	+	5-azacytidine	TSA	0,3-1%	x30-100
+	-	ARNi	5-azacytidine	TSA	5,60%	x560
+	+	+	Delta-Dnmt1	+	15%	x1500
-	-	+	Delta-Dnmt1	+	25%	x2500

Figure 18 : Tableau récapitulatif du taux de réactivation du Xi lors de diverses mutations et traitements chimiques. Csankovszki *et al.* ainsi qu'Hernandez-Munoz *et al.* ont étudié le taux de réactivation d'un transgène GFP présent sur le chromosome X inactif au sein de mutants de *Xist* ou en utilisant divers inhibiteurs chimiques. Ils observent un effet cumulatif important des différents phénomènes.

Ces résultats sont bien la preuve d'un phénomène accumulatif de marques hétérochromatiques agissant de concert sur le Xi, ainsi que de l'action en cascade des effecteurs de ces marques. Ces marques, en plus de modifier par elles-mêmes le contexte chromatinien, recrutent souvent plusieurs autres facteurs influant sur différentes marques mais toujours de manière répressive, afin d'assurer une extinction transcriptionnelle stable et efficace, bien que parmi ces marques et ces effecteurs aucun ne semble vraiment indispensable.

V. Les principaux acteurs de l'inactivation du chromosome X : *Xist* et son antisens *Tsix*

Au cours de cette introduction, nous avons donc pu voir que l'hétéromorphisme chromosomique entre les sexes nécessite un système de ré-équilibrage transcriptionnel afin qu'une quantité identique de transcrits provienne du simple chromosome X chez le mâle, ainsi que des deux chromosomes X chez la femelle, ce mécanisme étant appelé compensation de dose. Nous avons ensuite pu voir comment se met en place cette compensation, nous intéressant plus particulièrement à l'inactivation aléatoire du chromosome X chez la souris.

Après avoir vu en détail les marques caractéristiques de ce paradigme épigénétique, il est temps de s'intéresser à la "source" de ce phénomène : le gène *Xist*. Dans cette dernière grande partie introductive, nous allons tout d'abord voir ce qu'est le « centre d'inactivation » (la première région étant décrite comme contenant tous les éléments nécessaires à l'inactivation du chromosome X), puis nous verrons les caractéristiques majeures des gènes et des transcrits *Xist* et son antisens *Tsix*, avant de finalement nous intéresser au sujet central de cette thèse : la régulation de l'expression de *Xist*.

a. Le *Xic*

Le centre d'inactivation du chromosome X (*Xic*) peut être défini comme la région minimale capable de réaliser les fonctions de comptage, de choix, et d'initiation de l'inactivation. La fonction de maintien n'est à priori pas assurée par cette région (comme nous l'avons vu dans

le chapitre Conclusion sur les différentes marques épigénétiques, Brown & Willard, 1994).

i. Définition cytogénétique et transgénèse

La définition cytogénétique du *Xic* est venue de l'étude d'un grand nombre d'individus porteurs de translocations équilibrées entre le chromosome X et un autosome. Ces études ont pu démontrer que seuls les translocations contenant une région particulière du chromosome X (le *Xic*) étaient capables d'induire une inactivation ectopique en cis de l'autosome porteur (Rastan, 1983 ; Brown *et al.*, 1991 ; Leppig *et al.*, 1993 ; Cattanaach *et al.*, 1991). Ce type d'analyse a pu permettre l'identification d'une région située, chez l'homme, à la position Xq13.1 mesurant entre 680 et 1200kb (Brown *et al.*, 1991 ; Lafreniere & Willard, 1993 ; Leppig *et al.*, 1993 ; Willard, 1996). La localisation du *Xic* chez la souris a été obtenue en se servant de la synténie des gènes entre les deux espèces. Une région d'environ 700kb a été décrétée comme correspondant au *Xic*, elle se trouve au sein de la partie distale de la bande D du chromosome X (Rastan, 1983, 1985 ; Heard *et al.*, 1994).

Par la suite, de nombreuses études de transgénèse ont été menées à l'aide de YACs (pour « *Yeast Artificial Chromosome* »), de BACs (pour « *Bacterial Artificial Chromosome* ») ou encore de cosmides afin de déterminer la région précise responsable de l'inactivation (Lee *et al.*, 1996, 1999, 2005 ; Herzing *et al.*, 1997 ; Heard *et al.*, 1999) et pouvant récapituler toutes ses caractéristiques, tout en confirmant l'assomption que les *Xics* murin et humain sont bien synténiques. À ce jour, aucun transgène, inséré en simple copie, n'a su récapituler parfaitement les mécanismes attendus lors de l'inactivation aléatoire du chromosome X (Heard *et al.*, 1996 *et al.*, 1997), cependant nous pouvons noter l'absence au sein de ces transgènes de certains gènes maintenant connus comme ayant des rôles clés dans le processus d'inactivation tels que *Rnf12* (rôle dans le comptage) et *Xpct* (contenant la région *Xpr* (« *X-Pairing Region* »), impliqué dans l'association des deux chromosomes X avant l'initiation). Il est toutefois intéressant de noter qu'un transgène inséré en simple copie arrive à récapituler l'inactivation empreintée lorsqu'il provient de la lignée germinale mâle (de la même manière qu'un *Xp*) (Okamoto *et al.*, 2005). Ceci nous prouve donc que l'inactivation aléatoire mais aussi empreintée sont régulées au niveau du *Xic* mais selon des modalités différentes (le choix et le comptage dans le cas de l'inactivation aléatoire étant deux étapes clés dont la cellule ne peut s'affranchir).

ii. Unités fonctionnelles du *Xic*

Le séquençage puis l'annotation des séquences synténiques du *Xic*, chez la souris et chez l'homme, ont permis de générer la première carte des unités de transcription de cette région (Chureau *et al.*, 2002) (Figure 19). Le *Xic* contient plusieurs unités de transcription, la plupart d'entre elles n'ayant pas été reliées fonctionnellement au processus d'inactivation. En particulier, les 7 gènes codants liés au *Xic* murin (*Xpct* (Debrand *et al.*, 1998), *Cnbp2*, *Tsx* (Simmler *et al.*, 1996), *Chic1* (Simmler *et al.*, 1997), *Cdx4* (Gamer & Wright, 1993), *Ppnx*, et *Nap1L2* (Rougeulle & Avner, 1996 ; Rogner *et al.*, 2000 ; Attia *et al.*, 2007) ne semblent pas avoir de relation directe ou indirecte avec l'inactivation du chromosome X. Au cours de cette première annotation, l'identification de 4 unités de transcription produisant des ARNs non-codants de grande taille a été faite : *Xist*, *Tsix*, *Ftx*, et *Jpx*. Comme nous l'avons vu et allons continuer de le voir, le tandem *Xist/Tsix* est au cœur de la mise en place de l'inactivation, tandis que *Ftx* et *Jpx* sont soupçonnés d'y jouer un rôle qui, à l'heure actuelle, demeure encore inconnu. Depuis, quelques autres ARNs non-codants ont été décrits comme émanant du *Xic* et possédant des rôles non-négligeables dans l'inactivation comme le petit transcrit *RepA* (voir le chapitre : « Modification d'histones sur le Xi », p36) (Zhao *et al.*, 2008), les transcrits sens/antisens de la région répétée *DxPas34* (une région impliquée dans la régulation de *Tsix*) (Cohen *et al.*, 2007), ou encore un transcrit intergénique pouvant engendrer une forme longue de *Tsix* émanant d'une région appelée *Xite* (Ogawa & Lee, 2003). Enfin une autre région à transcription sens/antisens non-codante de grande complexité a été identifiée 50kb en amont du gène *Xist* : la « Région B » (Rougeulle & Avner, 2003), celle-ci n'ayant à présent aucune fonction connue.

Chez l'homme, ces gènes, aussi bien codants que non-codants, sont conservés, à l'exception de *Ppnx*, *Tsx* (pseudogénisé) et *Tsix* (sur lequel nous reviendrons ultérieurement).

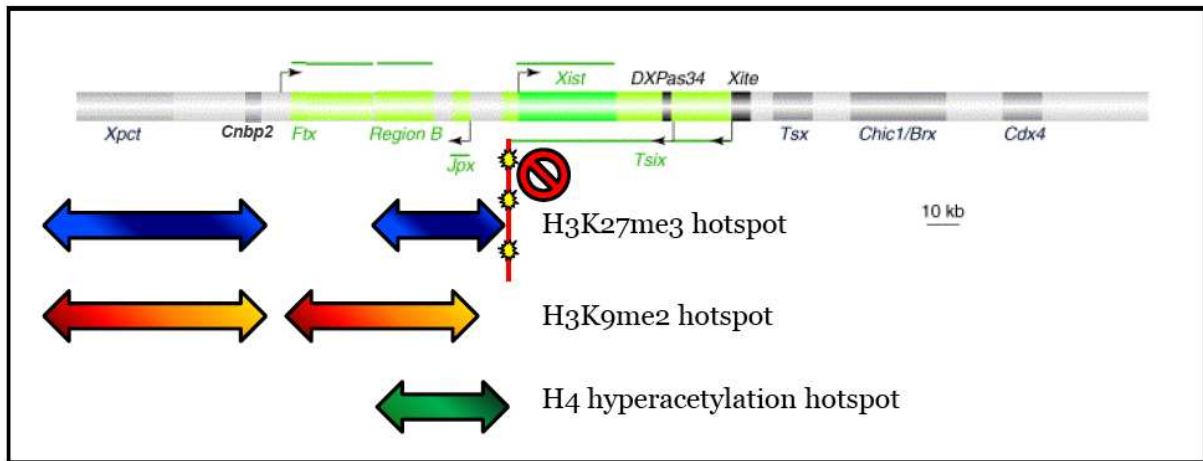


Figure 19 : Éléments fonctionnels au sein du centre d'inactivation du chromosome X. Le XIC, comme défini par des analyses de délétions et de transgénèse, est montré avec ses éléments fonctionnels importants. Les gènes montrés en vert produisent des transcrits non-codants (lignes vertes). Les gènes montrés en gris produisent des transcrits codants pour des protéines et n'ont aucun lien connu avec l'inactivation du chromosome X. Sous la carte, la flèche bleue représente un domaine de H3K27me3 uniquement présente dans les cellules ES, la flèche rouge/orange représente un domaine de H3K9me2 qui est constitutivement présent au sein des cellules ES mâles et femelles aussi bien avant qu'après différenciation. Et la barre verte représente un domaine d'H4Ac, uniquement présent dans les ES femelles. La ligne rouge représente la frontière établie par l'accumulation d'ARNpolII (étoiles jaunes) suite à la fin de la transcription de *Tsix*, cela empêche la propagation du domaine de H3K27me3 en direction de *Xist*. (Inspirée de Heard *et al.*, 2004)

iii. "Point chaud" de modifications épigénétiques en 5' de *Xist*

Des analyses par immunofluorescence et par immunoprécipitation de chromatine (ChIP, voir Annexes 1 et 2, p268-269) ont permis de mettre en évidence une structure chromatiniennne particulière se situant dans la région 5' de *Xist* et étendue sur près de 340kb (Heard *et al.*, 2001, Rougeulle *et al.*, 2004). En effet, au sein de cette région, on observe un fort enrichissement en marques répressives H3K9me2 et H3K27me3, d'où la dénomination commune donnée à cette région : « *hotspot* » de méthylation (Figure 19). L'histone méthyltransférase G9a est responsable du maintien de H3K9me2 au sein du *hotspot* mais pas de sa déposition, fait qui serait antérieur à la dérivation des cellules ES (Rougeulle *et al.*, 2004). De plus, les profils et les cinétiques de fixation de ces marques ne sont pas identiques (le *hotspot* d'H3K9me2 persistant post-différenciation), laissant supposer que leur mise en place ou leur maintien sont effectués de manière indépendante. Cependant, on peut apercevoir une absence de ces marques au niveau des promoteurs de *Jpx* et *Ftx*, celles-ci étant remplacées par une hyperméthylation de H3K4.

Un deuxième enrichissement épigénétique avait déjà été observé en 1999 par O'Neill *et al.*. Il s'agit d'un *hotspot* d'acétylation de l'histone H4, ce marquage ne s'étalant que sur environ la moitié de la longueur du *hotspot* de méthylation, mais présentant la particularité de ne pas être présente dans les cellules ES mâles, ou dans des cellules femelles hétérozygotes pour une délétion du premier exon de *Xist* (O'Neill *et al.*, 1999), ce qui pourrait laisser supposer un rôle pour ce *hotspot* d'H4ac dans le comptage (Figure 19). Le *hotspot* de méthylation a, quant à lui, été proposé comme jouant un rôle de centre de nucléation pour l'inactivation du chromosome X en raison de sa nature hétérochromatique, de la présence de deux ARNs non-codants ainsi que la région B. La région B pourrait représenter une candidate idéale pour le recrutement de la machinerie d'hétérochromatinisation liée à l'interférence ARN (en raison de son transcrit sens/antisens) par des mécanismes similaires à ceux décrits au niveau des centromères de levure dont la chromatine est transcriptionnellement éteinte grâce à de petits ARNs sens et antisens générés à partir de d'ARNs non-codants et par le complexe protéique spécialisé appelé RITS (pour « *RNA-Induced Transcriptional Silencing* ») (Buhler *et al.*, 2006, Volpe *et al.*, 2002, Verdel *et al.*, 2004). Toutefois, à ce jour aucune preuve expérimentale n'est venue étayer cette hypothèse.

b. *Xist* et *Tsix*, généralités

i. *Xist*

La volonté de caractériser le centre d'inactivation du chromosome X, impliqué dans l'établissement de la répression transcriptionnelle de l'intégralité de ce chromosome, a poussé l'équipe de Huntington Willard, en 1991, à étudier le profil d'expression des gènes se trouvant au sein de cette région candidate chez l'humain (Brown *et al.*, 1991). Ces travaux ont permis d'isoler et de caractériser le gène *XIST*, révélé comme étant le seul gène connu étant uniquement exprimé à partir du chromosome X inactif. Peu de temps après, le même genre d'expériences ont été menées chez la souris, parvenant aux mêmes conclusions (Borsani *et al.*, 1991, Brockdorff *et al.*, 1991). Cependant, il faudra attendre 1996 pour que Penny *et al.* démontrent formellement son implication dans le processus d'inactivation en le délétant de manière hétérozygote dans des cellules ES femelles, qui lorsqu'elles sont différenciées

n'inactivent que l'allèle sauvage (Penny *et al.*, 1996).

Le gène *XIST/Xist* s'étend sur 32kb chez l'homme et sur 23kb chez la souris, les transcrits en émanant ne codent aucune protéine (pas de phase ouverte de lecture) malgré le fait qu'ils soient polyadénylés, cappés et épissés. De plus, ces transcrits ne sont pas exportés vers le cytoplasme et restent donc localisés dans le noyau cellulaire (Brown *et al.*, 1992, Brockdorff *et al.*, 1992, Cohen & Panning, 2007). Par la suite, la technique d'hybridation *in situ* fluorescente (RNA FISH) permet d'observer les ARNs en transcription dans le noyau interphasique. Il fut ainsi observé que *Xist*, au lieu d'être représenté par un signal ponctuel au niveau de son site de transcription (comme c'est le cas pour les gènes "classiques"), est visible dans les cellules somatiques sous la forme d'un large domaine, et lors de sa détection simultanée avec le territoire chromosomique du chromosome X (DNA FISH), il a été observé qu'ils colocalisent au sein du noyau (Clemson *et al.*, 1996). L'ARN *XIST/Xist* recouvre donc la quasi-totalité du chromosome X inactif dans les cellules somatiques femelles en *cis*, ne se dissociant que lors de la métaphase chez l'homme, tandis que, chez la souris, il reste localisé au niveau des régions riches en gènes (Duthie *et al.*, 1999) au cours de cette phase.

Des études de comparaison phylogénétiques ont permis de mettre en évidence que *Xist*, au moins en partie, dérive du gène codant *Ln3* et est devenu non-codant suite à la pseudogénéisation d'une grande partie du *Xic*, et ce, après la divergence entre marsupiaux et euthériens (Duret *et al.*, 2006). Huit exons de *Xist* ont été caractérisés, aussi bien chez l'homme (Brockdorff *et al.*, 1992, Brown *et al.*, 1992, Chureau *et al.*, 2002), que chez la vache (Chureau *et al.*, 2002), la souris (Chureau *et al.*, 2002, Sheardown *et al.*, 1997, Simmler *et al.*, 1996) ou le campagnol (Nesterova *et al.*, 2001), cependant certains d'entre eux ne sont pas retrouvés au sein de toutes les espèces (exons 2 et 7 humains, exons 2 et 5 chez la souris et la vache) (Chureau *et al.*, 2002) (Figure 20).

Sachant qu'il existe une protéine (PABP pour « *Poly(A) Binding Protein* ») dont le but est de s'accrocher aux queues poly-adénylées des ARNm afin de les protéger des attaques de nucléases, il est fort possible que ce genre de différence explique les différentes stabilités de l'ARN *Xist* observées entre embryons pré- et post-implantatoire : demi-vie inférieure à 30min et supérieure à 8h respectivement (Panning *et al.*, 1997, Sheardown *et al.*, 1997). Cependant aucun rôle cellulaire ni dans l'inactivation du chromosome X n'a encore été attribué à ces différents isoformes.

Le fait que chez la souris, *Xist* possède deux promoteurs alternatifs P1 et P2, séparés d'environ 1,5kb (Johnston *et al.*, 1998) introduit une source additionnelle de complexité au sein de la multitude de transcrits possibles. Le premier article décrivant l'existence de promoteurs alternatifs tendait à démontrer que le promoteur P2 était le promoteur principal de par le fait qu'environ 75% des transcrits *Xist* présents au sein de cellules somatiques émanent de ce promoteur (Johnston *et al.*, 1998). Cependant, cette hypothèse pose un sérieux problème fonctionnel : le promoteur P2 se trouvant en aval du *Repeat-A*, cela signifie que ses transcrits n'auront pas la capacité de participer à la répression transcriptionnelle du chromosome X inactif lors de la différenciation cellulaire (Wutz *et al.*, 2002). De plus, en 2005, Navarro *et al.* ont démontré que seul le promoteur P1 était développementalement régulé au cours de la différenciation : le recrutement de l'ARN polymérase II et de la sous-unité TFIIB n'augmentant qu'au niveau de P1 sur l'allèle destiné à être inactivé dans les cellules exprimant fortement *Xist* par rapport aux cellules ES indifférenciées (Navarro *et al.*, 2005). Navarro *et al.* proposent que la différence de quantité de transcrits émanants de P1 ou P2 observée par Johnston *et al.* ainsi que par Sado *et al.* (Sado *et al.*, 2005) pourrait être due à une transcription ou une stabilisation plus efficace du transcrit à partir de P2 ou bien à des artéfacts de reverse-transcription liés à l'existence de structures secondaires engendrées par les séquences répétées contenues dans la région d'étude. L'hypothèse d'un plus grand nombre de transcrits provenant de P2 dans les cellules somatiques n'est pas particulièrement choquante (indépendamment du fait que le complexe de pré-initiation de la transcription semble majoritairement se fixer à P1) si l'on considère que *Xist* n'est pas nécessaire au maintien de l'inactivation dans ces cellules. Une étude cinétique de la quantité et de la stabilité de ces deux genres de transcrits à des temps précoces au cours de l'initiation de l'inactivation puis lors du maintien serait sans doute très riche en enseignements quant à l'importance de la présence de

l'un ou l'autre transcrit lors de ces différentes étapes.

En outre, les domaines fonctionnels ne sont pas très bien caractérisés. À part la séquence répétée *Repeat-A*, qui comme nous l'avons vu, est responsable de la répression transcriptionnelle des séquences géniques du chromosome X lors de l'inactivation (Wutz *et al.*, 2002), il existe 5 autres éléments répétés, tous très conservés au cours de l'évolution, à propos desquels nous ne connaissons que peu de choses (voir le chapitre : « L'ARN non-codant *Xist* », p25). La manière dont fonctionne le *Repeat-A* est encore inconnue. Dernièrement, plusieurs articles l'ont mise en cause dans le recrutement du complexe protéique PRC2 notamment grâce à deux grandes boucles formées par quatre répétitions (Zhao *et al.*, 2008, Maenner *et al.*, 2010), une structure particulière qui est elle aussi conservée à travers le groupe des euthériens. Ce genre d'interactions ARN non-codant/PRC2 concorde bien avec des observations faites concernant d'autres ARNnc tels que *HOTAIR* ou encore *Kcnq1ot1* qui sembleraient aussi interagir avec ce complexe protéique (Rinn *et al.*, 2007, Pandey *et al.*, 2008). Mais les avis divergent quant au rôle possible de l'interaction entre PRC2 et *Xist*, certains privilégiant l'hypothèse d'un moyen de propagation de l'hétérochromatinisation lors de l'initiation de l'inactivation, d'autres préférant l'hypothèse d'un rôle dans la régulation de *Xist* (sur laquelle nous reviendrons plus tard).

Le gène *Xist* exprimé à partir de l'allèle inactif dans les cellules somatiques joue donc un rôle primordial dans l'établissement de cette répression génique. La délétion de *Xist* sur l'Xi au sein des cellules différenciées engendre éventuellement une légère dérégulation du Xi mais ne semble pas modifier la réplication tardive du Xi (Csankovszki *et al.*, 1999 *et al.*, 2001). Par contre la délétion du *Xist* non-transcrit sur le Xa (de même que le X unique actif chez les mâles ou les femelles XO) modifie les propriétés réplcatives de celui-ci en *cis*, en retardant sa réplication au cours de la phase S (Diaz-Perez *et al.*, 2005) (voir le chapitre : « Réplication asynchrone », p44). Les mêmes auteurs ont ensuite engendré cette mutation de *Xist* sur les deux allèles du chromosome X dans des cellules somatiques femelles. Cette expérience a démontré le retour de certaines marques euchromatiques le long du Xi (l'acétylation de l'histone H4 ainsi que la méthylation de H3K4) mais aucune réactivation génique n'a été observée (Diaz-Perez *et al.*, 2006). Paradoxalement, la réplication déjà tardive du Xi (signe de l'hétérochromatine) n'a été que plus retardée dans ces cellules mutantes, le même phénotype étant observable pour le Xa. Ces deux phénomènes (marques euchromatiques et réplication

encore plus tardive) seraient apparemment dus aux deux allèles, leur absence engendrant une déstabilisation des chromosomes X et une augmentation notable des points de cassure ainsi que des translocations, symbolisées par le recrutement des modifications épigénétiques typiques de la réparation de l'ADN (γ H2AX, recrutement de p53 phosphorylée). Les mécanismes exacts de ces changements d'état ne sont pas très clairs, cependant ces résultats laissent supposer l'existence d'un élément au sein de ces 21kb de *Xist* jouant un grand rôle de la stabilité de l'ADN du chromosome X, aussi bien en *cis* qu'en *trans*.

Une région promotrice minimale permettant l'expression de *Xist* a été déterminée par l'analyse de différentes constructions, composées d'une séquence en 5' de *Xist* introduite devant un gène rapporteur et transfectées de façon stable ou transitoire dans les cellules ES et (Pillet *et al.*, 1995, Sheardown *et al.*, 1997). Cette région, constituée des 400pb en amont du promoteur P1, a permis de définir plusieurs séquences susceptibles d'être fixées par diverses protéines telles que YY1, Sp1, CBF, CTCF. L'analyse des effets respectifs de ces différents éléments sur l'expression de *Xist* a révélé que cette région est uniquement responsable de l'expression constitutive de ce gène. Les séquences nécessaires à la régulation de *Xist* pourraient être localisées plus en amont ou plus aval. La recherche de sites de fixation de facteurs de transcription par hypersensibilité à la DNase I a d'ailleurs permis l'identification de plusieurs sites aussi bien en amont qu'en aval de P1 (Sheardown *et al.*, 1997, Newall *et al.*, 2001, Tsai *et al.*, 2008).

ii. *Tsix*

En 1999, de manière simultanée, plusieurs groupes ont découvert grâce à des expériences d'ARN-FISH, de Réverse-Transcriptase PCR brin-spécifique et de criblage de banques d'ADNc de cellules ES, un transcrit anti-sens à *Xist* baptisé *Tsix* (Debrand *et al.*, 1999, Lee *et al.*, 1999, Mise *et al.*, 1999). L'unité de transcription majoritaire de ce gène, longue de 40kb et traversant l'intégralité du gène *Xist*, code pour un long ARN non-codant polyadénylé (Debrand *et al.*, 1999, Lee *et al.*, 1999). À l'origine, ce transcrit a été décrit comme émanant d'un îlot CpG 12kb en aval de l'exon 8 de *Xist*, avant qu'un site d'initiation mineur ne soit décrit, positionné encore environ 10kb en aval et associé lui aussi à un petit îlot CpG (Sado *et al.*, 2001). Cette même publication ainsi que les recherches de Shibata et Lee (Shibata & Lee,

2003) ont permis la description et la quantification de formes de transcrits antisens différentiellement épissés et polyadénylés. À l'heure actuelle, 6 exons ont été décrits. Ils génèrent une grande hétérogénéité par épissages alternatifs complexes au niveau de l'extrémité 5' du transcrit initié au promoteur le plus proche de *Xist* (Sado *et al.*, 2001, Shibata & Lee, 2003). De plus, il semblerait que seuls 30 à 60% des transcrits *Tsix* soient épissés, cette proportion restant constante au cours de la différenciation cellulaire. Une autre source de diversité provient des multiples sites d'arrêt transcriptionnels ayant été décrits : il en existe 3 en 5' du promoteur P1 de *Xist* et un quatrième situé juste en amont du mini-satellite *DxPas34*. Les promoteurs alternatifs, les épissages alternatifs et les sites multiples de terminaison, confèrent au transcrit *Tsix* une grande hétérogénéité dont la fonction n'a pas encore été étudiée. Il est toutefois essentiel de noter que seul le dernier exon (appelé exon 4) de *Tsix* est complémentaire à *Xist*, recouvrant les deux promoteurs et donc le *Repeat-A*, quelque soit le site de terminaison utilisé.

Malgré le fait que, comme nous le verrons par la suite, *Tsix* joue un rôle essentiel dans la régulation de *Xist*, paradoxalement, ce transcrit anti-sens ne semble être conservé ni au niveau de sa séquence, ni au niveau de sa fonction au sein des autres espèces eutheriennes (Figure 21). Un transcrit antisens à *XIST* dans le *XIC* humain a bien été identifié, il s'étend sur environ 30kb mais ne semble pas être épissé et surtout, il ne recouvre que les 2 derniers exons de *XIST* (Migeon *et al.*, 2001). De plus, l'analyse du profil d'expression de cet antisens par ARN-FISH a montré sa co-expression avec le transcrit *XIST* au niveau du chromosome X inactif, contrairement à la situation murine (Migeon *et al.*, 2002). Ces données vont clairement à l'encontre de l'existence d'un mécanisme de répression de *XIST* par ce potentiel *TSIX* et suggèrent l'intervention, chez l'homme, d'un autre processus de régulation de *XIST*. Il serait très intéressant de tenter de corrélérer cette perte de fonction de la transcription antisens chez l'homme à l'absence d'une inactivation empreintée du chromosome X.

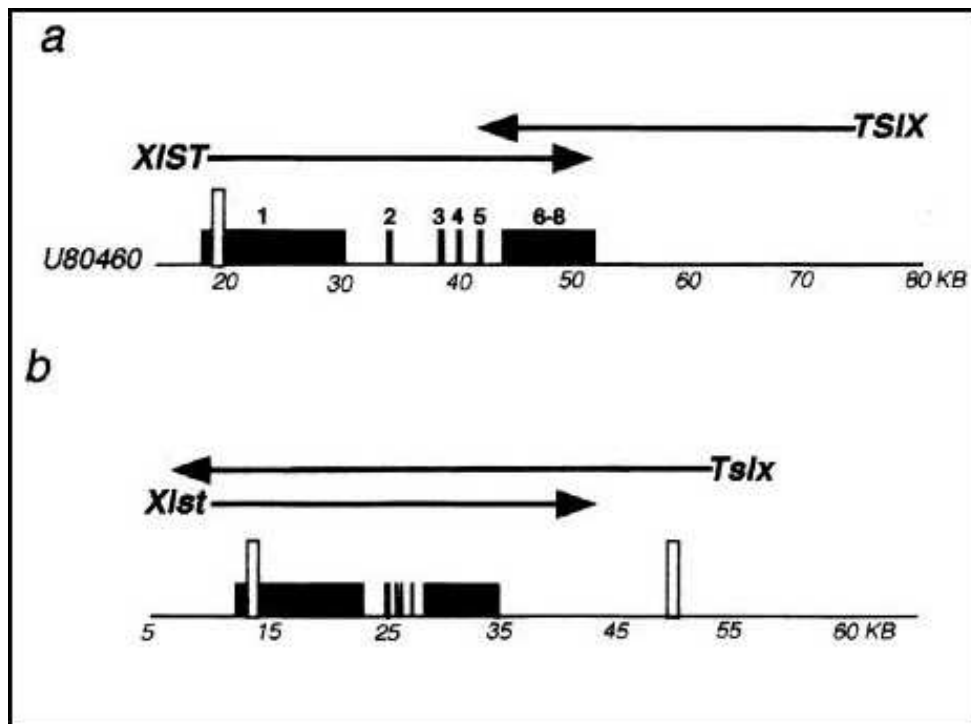


Figure 21 : Comparaison de l'agencement des gènes *XIST/Xist* et *TSIX/Tsix* et de leur îlots CpG entre homme et souris. Les boîtes noires représentent les exons de *XIST/Xist*, les lanches représentent les îlots CpG. Les flèches montrent les transcrits ainsi que leur sens de transcription. (a) Fragment du *XIC* humain. (b) Fragment de *XIC* murin. (D'après Migeon *et al.*, 2002)

Dans les cellules ES indifférenciées murines, *Tsix* est exprimé aussi bien chez les mâles que chez les femelles (Debrand *et al.*, 1999, Lee *et al.*, 1999, Mise *et al.*, 1999, Sado *et al.*, 2001). Dans ce contexte cellulaire, les transcrits *Tsix* sont bien plus nombreux que les transcrits *Xist*, d'environ un facteur 100 au début du transcrit majoritaire de *Tsix* et d'un facteur 10 au niveau de son dernier exon (Shibata & Lee, 2003). Il faut cependant tenir compte du fait que dans les cellules ES, *Xist* est très peu exprimé, à raison de 10 copies par cellule femelle et 4 par cellule mâle (Huynh & Lee, 2003). Il serait intéressant d'apprécier dans quelle mesure le niveau relatif d'expression de *Xist* et de *Tsix* peut varier entre cellules d'une même population. En particulier, on peut se demander si *Xist* est réellement exprimé dans toutes les cellules ou si son niveau moyen d'expression traduit une forte expression dans quelques cellules seulement. D'autre part, nous ne savons pas si, à un instant donné, l'expression de *Xist* et *Tsix* peut avoir lieu à partir du même *Xic*, ou si l'un et l'autre sont exprimés de manière alternative. L'évolution des techniques de marquage des ARN et d'imagerie devrait permettre d'apporter des réponses précises à ces questions.

En cours de différenciation, l'expression de *Tsix* disparaît aussi bien dans les cellules mâles que

les cellules femelles, bien qu'au sein de ces dernières son expression soit perdue de manière plus rapide sur le Xi que sur le Xa (Debrand *et al.*, 1999, Lee *et al.*, 1999).

L'expression de *Tsix* semble donc anticorrélérer celle de *Xist*, justifiant ainsi d'étudier la manière dont est contrôlée l'expression de l'antisens si l'on espère avoir un jour une vue globale de l'ensemble des mécanismes gouvernant l'inactivation du chromosome X.

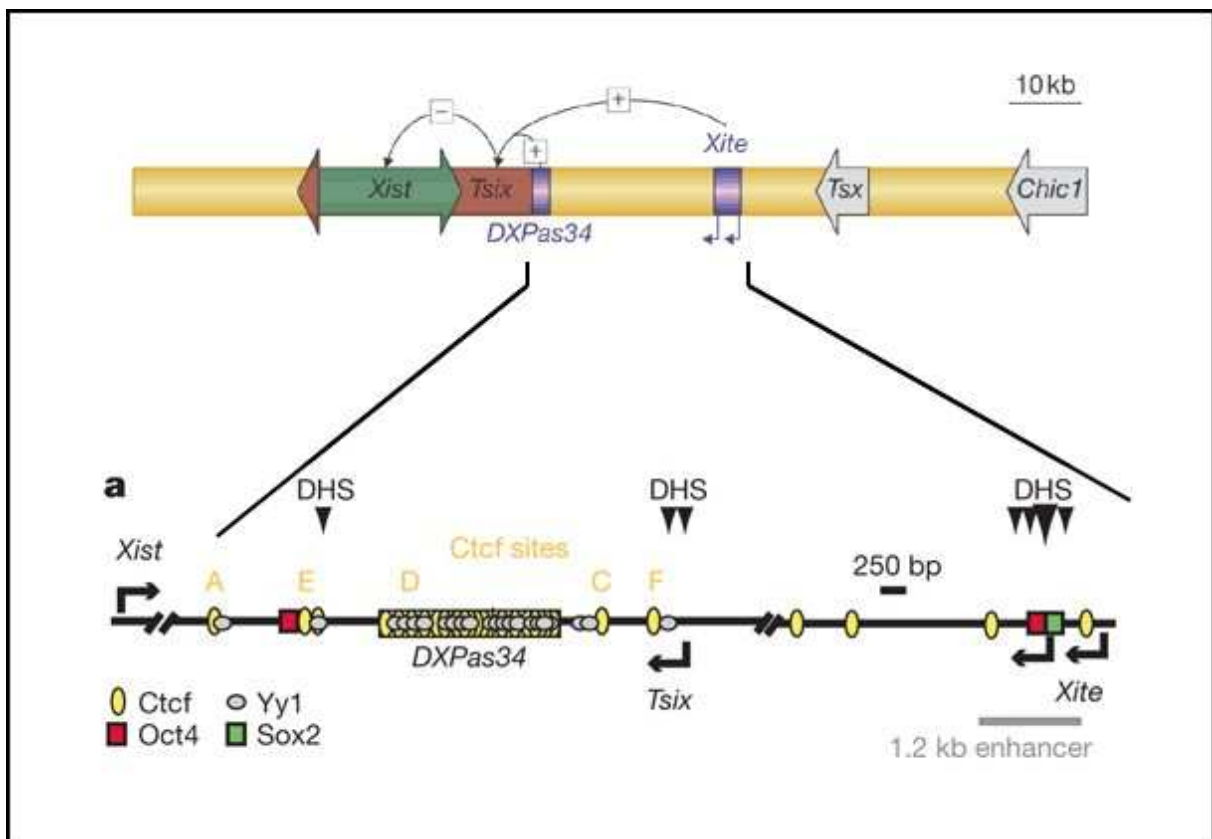


Figure 22 : Schéma descriptif des éléments régulateurs de *Tsix*. Le transcrit *Tsix* (flèche marron), antisens à *Xist* (flèche verte), est principalement initié 16kb en aval de *Xist* et s'étend sur 40kb. L'élément d'ADN régulateur *Xite* (en violet) et le minisatellite *DxPas34* (violet) ont été démontrés comme amplicateurs de l'expression de *Tsix* (+), qui à son tour réprime *Xist* (-). (a) Carte des sites de fixation potentiels de CTCF, YY1, Oct4, Sox2 au sein d'une région de 15kb du *XIC* murin. (Adaptée de Morey & Bickmore, 2006 et Donohoe *et al.*, 2009)

iii. *DxPas34*

Environ 750pb en amont du promoteur de *Tsix* (dans le sens de la transcription de *Xist*) se trouve le minisatellite *DxPas34* qui est constitué d'une séquence d'environ 1,2kb composée de 34 nucléotides répétés en tandem, le nombre de répétitions variant en fonction de la lignée de la cellule (Heard *et al.*, 1993, Courtier *et al.*, 1995, Simmler *et al.*, 1996) (Figure 22). Cette région répétée est le siège d'une méthylation différentielle dans les tissus somatiques : le

minisatellite du Xa se trouvant méthylé mais non pour le Xi (Courtier *et al.*, 1995, Prissette *et al.*, 2001) (contrairement à la situation habituelle où les îlots CpG sont méthyles sur le Xi et non le Xa). Cependant cette méthylation est acquise après l'initiation de l'inactivation aléatoire, ce qui suggère qu'elle n'est pas causalement liée à la régulation de *Xist* ou *Tsix*. Au promoteur de *Tsix*, on observe aussi une association constitutive de la méthylation de H3K4 dans les cellules mâles ainsi qu'une absence de méthylation H3K9 pouvant laisser supposer l'existence d'une structure chromatinienne ouverte de cette région sur le Xa, cependant l'acétylation des histones H3 et H4 est perdue chez les mâles en cours de différenciation (Kimura *et al.*, 2002) correspondant sans doute à l'extinction transcriptionnelle de *Tsix*.

Malgré ce manque de corrélation entre méthylation différentielle et régulation de *Tsix*, *DxPas34* semble tout de même influencer sur l'expression de l'antisens. En effet, en 2005, Stavropoulos *et al.* ont démontré que le promoteur majoritaire de *Tsix* ne pouvait, à lui seul, récapituler la dynamique développementale de l'expression de l'antisens (Stavropoulos *et al.*, 2005). Ils sont ensuite parvenus à prouver qu'une expression supérieure de *Tsix* (par rapport à s'il y a uniquement le promoteur) ainsi qu'une régulation développementale était obtainable lors de l'ajout de certaines séquences flanquantes présentes au locus endogène de *Tsix*. Ces éléments amplificateurs de l'expression (« *enhancers* ») correspondent d'une part au locus *Xite* (voir le chapitre suivant) ainsi qu'un *enhancer* bipartite englobant la région *DxPas34* (Stavropoulos *et al.*, 2005).

DxPas34 semble posséder plusieurs fonctions au niveau de la régulation de *Tsix*. Dans un premier temps, une délétion brute de son locus engendre une forte diminution transcriptionnelle de *Tsix* chez les cellules ES (expression réduite de 70 à 90% selon l'article) (Vigneau *et al.*, 2006, Cohen *et al.*, 2007). Toutefois, ces deux équipes ayant réalisé ces expériences divergent fortement quant aux conclusions tirées de l'étude de leurs cellules mutantes. Vigneau *et al.* observent la disparition transcriptionnelle de *Tsix* au cours de la différenciation (bien qu'elle soit déjà très basse) et parallèlement à ceci, ils observent une surexpression ectopique de *Xist* d'environ un facteur 7 au sein des cellules ES mâles non-différenciées et même une présence de domaines de *Xist* dans environ 15% des cellules mâles différenciées. L'équipe de Jeannie Lee, bien qu'elle observe une diminution de l'expression de *Tsix* dans les cellules ES chez les mutants $\Delta DxPas34$, observe aussi une faible dérégulation de l'antisens au cours des stades tardifs de la différenciation cellulaire sur l'allèle muté. Elle en

conclut donc que *DxPas34* joue aussi bien le rôle d'*enhancer* de *Tsix* dans les cellules ES au cours des stades précoces de l'inactivation, que celui de répresseur dans les cellules somatiques (ceci ne corrélant pas avec leurs données précédemment obtenues par des expériences de luciférase) (Stavropoulos *et al.*, 2005). De plus, de la même manière que Vigneau *et al.*, ils observent une surexpression de *Xist*, mais pas de domaine ectopique dans les ES mâles différenciées. Ces différences phénotypiques pourraient s'expliquer notamment par les techniques de différenciation cellulaires (expliquées au chapitre Modèles cellulaires), par les différences entre lignées cellulaires, ou bien par les faibles différences concernant la séquence délétée.

Dans un second temps, il semblerait que *DxPas34* ait la capacité d'initier une transcription bidirectionnelle (Cohen *et al.*, 2007, Stavropoulos *et al.*, 2005). La transcription antisens pourrait correspondre à certaines formes du transcrit *Tsix* émanant de *DxPas34* (Shibata & Lee, 2003), cependant la fonction de transcrits sens, moins abondants que *Tsix*, mesurant environ 3kb et dont les multiples sites d'initiation semblent chacun corrélér avec une unité de la séquence répétée, est encore inconnue. Curieusement, l'expression de ce transcrit sens (toujours par rapport à *Xist*) émanant de *DxPas34* corrèle avec celle de *Tsix* ainsi qu'avec le profil de méthylation de *DxPas34* dans les cellules mâles (différenciées ou non) (Cohen *et al.*, 2007).

La présence de sites d'hypersensibilité à la DNase, ainsi que de sites pseudo-consensus pour les protéines CTCF et YY1 dans chaque répétition font de cette région une localisation potentielle de facteurs de transcription dont le but serait de réguler *Tsix* (Stavropoulos *et al.*, 2005, Chao *et al.*, 2002, Donohoe *et al.*, 2007) (Figure 22). Et en effet, Donohoe *et al.* ont publié que CTCF aussi bien que YY1 se fixent à *DxPas34* et ses alentours, de manière interdépendante par hétérodimérisation dans les cellules ES. Cependant, l'effet sur la régulation transcriptionnelle de l'antisens par ses protéines reste flou. Le laboratoire de Jeannie Lee a publié en 2007 que lors d'une déplétion de ces protéines par siRNA la quantité de transcrits *Tsix* chute simultanément à une augmentation de transcrits *Xist* dans des cellules ES mâles, corroborant leur hypothèse que CTCF et YY1 sont des activateurs transcriptionnels de *Tsix* (Donohoe *et al.*, 2007), ceci étant aussi confirmé par des expériences de luciférase. Cependant en 2009, la même équipe réalise une fois de plus une expérience de déplétion de CTCF par siRNA mais cette fois ci dans des cellules femelles après 4 jours de différenciation

(en corps embryonnaires) et observe paradoxalement une augmentation de l'expression de *Tsix* (corrélée à une baisse de *Xist*) alors qu'à cette même période elle a considérablement commencé à décroître dans les cellules sauvages (Donohoe *et al.*, 2009). De plus, d'après leurs données, la fixation de CTCF et YY1 autour de *DxPas34* (constitutifs dans les cellules ES mâles et femelles ainsi qu'en cours de différenciation, sauf pour un site CTCF) ne semble pas forcément corrélée avec le profil d'expression de *Tsix*. Cependant en observant le profil de fixation de ces protéines sur le minisatellite même, la corrélation est plus claire, surtout lorsque l'on compare la fixation de CTCF et YY1 dans des cellules ES femelles à celle des MEFs femelles (Navarro et données personnelles).

Dernièrement, le facteur de pluripotence Oct4 a aussi été impliqué dans la régulation de *Tsix*, au vu de la baisse transcriptionnelle de ce dernier lors de l'utilisation de siRNA anti-*Oct4* dans des cellules ES femelles 4 jours post-différenciation en EBs (Donohoe *et al.*, 2009). Ces résultats pourraient s'expliquer par l'observation d'une faible fixation du facteur de pluripotence à un site légèrement (environ 700pb) en amont de *DxPas34* connu pour fixer CTCF et YY1 (le site E). Ils laisseraient aussi supposer l'existence d'un modèle intéressant où Oct4 régulerait négativement l'expression de *Xist* par deux mécanismes distincts : une fixation directe dans l'intron 1 de *Xist* menant à une répression transcriptionnelle directe, ainsi qu'une fixation dans *DxPas34* et *Xite* ayant comme effet de stimuler l'expression de *Tsix* qui réprime *Xist* post-transcriptionnellement (voir les chapitres suivants) (Navarro *et al.*, 2005 *et al.*, 2008, Donohoe *et al.*, 2009). Cependant, cette fixation n'a jamais été observée par Navarro *et al.* (voir l'Article, p186).

La fixation d'Oct4 et CTCF ne semble pas être uniquement impliquée dans la régulation transcriptionnelle de *Tsix* mais pourrait également jouer un rôle important dans le phénomène d'appariement des chromosomes en *trans* médié par la région *Xite/DxPas34* qui a lieu juste avant l'initiation de l'inactivation (Xu *et al.*, 2006 *et al.*, 2007, Donohoe *et al.*, 2009) (voir le chapitre : « Choix », p15). En effet, la déplétion de CTCF ou d'Oct4 semble diminuer la proportion de cellules dans lesquelles on observe du « *pairing* » lors de l'initiation de la différenciation dans les cellules ES femelles (Xu *et al.*, 2007, Donohoe *et al.*, 2009). Cependant de manière surprenante, cette même équipe observe que lors d'une déplétion d'Oct4, non seulement l'appariement en *trans* n'a plus lieu, mais aussi que les deux chromosomes X initient leur inactivation dans environ 20% des cellules (ceci n'étant pas observé dans des

cellules déplétées de CTCF) soulignant un biais sévère du comptage.

Ces observations nous amènent donc à nous interroger sur l'effet exact de CTCF et Oct4 à ce locus. Sachant que CTCF affecte aussi le pairing sans pour autant modifier le comptage, il semblerait qu'OCT4 soit impliqué dans ces deux phénomènes de manière indépendante. L'inconvénient majeur des expériences de déplétion par siRNA est qu'elles déplètent l'ensemble du génome en fixation de la protéine visée, ne permettant pas de savoir si l'effet observé est dû à l'absence de la protéine au site candidat ou s'il est dû à une absence de fixation autre part. Et c'est exactement le cas de la fixation d'Oct4 dans la région *Xite/DxPas34*, pour laquelle nous ne pouvons pas savoir si le biais de comptage des chromosomes X et donc une surexpression de *Xist* de manière biallélique est due à l'absence de répression directe de *Xist* (à travers sa fixation de son intron 1) ou à l'absence d'activation de *Tsix* (à travers sa fixation dans *Xite/DxPas34*), ou encore si les deux n'agissent pas de manière synergétique. Une dernière possibilité pour expliquer ces résultats serait qu'au vu des nombreux sites de fixation pour ces protéines à travers le génome ainsi que leur rôle prépondérant dans la pluripotence (pour Oct4) et la régulation génique (pour Oct4 et CTCF), il est fort possible que de les dépler au niveau de la cellule entière induise des artéfacts dus à des modifications physiologiques/métaboliques de la cellule lors de son entrée en apoptose, par exemple.

iv. *Xite*

Comme nous avons pu le voir au cours des chapitres précédents, un autre locus semble être impliqué dans la régulation transcriptionnelle de *Tsix* ainsi que dans la transvection des deux allèles du chromosome X lors de l'initiation de l'inactivation : le locus *Xite*. Il se trouve à environ 15kb en aval du promoteur majoritaire de *Tsix*, s'étend sur environ 6kb et est situé au niveau de sites initiateurs de transcription antisens intergénique, l'un de ces sites correspondant au promoteur minoritaire de *Tsix* (Ogawa & Lee, 2003) (Figure 22). De manière similaire à *DxPas34*, il est aussi la source de transcription bidirectionnelle mais uniquement dans les cellules ES, la fonction de ce phénomène étant actuellement encore inconnue. Le fait de bloquer la transcription anti-sens émanant de *Xite* (donc potentiellement le transcrit mineur de *Tsix*) n'altère en rien le rôle de *Xite* dans la régulation de *Tsix* (Ogawa & Lee, 2003, Stavropoulos *et al.*, 2005). En corrélation avec sa fonction de régulateur transcriptionnel de

Tsix, l'expression de cette transcription antisens intergénique originant de *Xite* décroît parallèlement à celle de *Tsix* en cours de différenciation des cellules ES mâles et femelles, leurs sites de méthylation différentielle étant d'ailleurs méthylés de la même manière selon les tissus (hypométhylés dans les cellules ES et hyperméthylés dans les EBs) (Ogawa & Lee, 2003 ; Boumil *et al.*, 2006). Cependant ces mêmes sites sont hyperméthylés dans les cellules spermatiques (dont le X est inactivé) mais hypométhylés dans les oocytes (dont le X est actif), ceci semblant logique par rapport à l'expression de *Xist* dans ces cellules et constituant par la même occasion une forme potentielle d'empreinte parentale pour l'inactivation empreintée (Boumil *et al.*, 2006). Toutefois, la délétion de *Xite* semble indiquer que son rôle est de maintenir l'expression de *Tsix* sur le futur Xa lors de l'initiation de l'inactivation, l'expression de *Tsix* étant plus rapidement perdue sur l'allèle muté que sur l'allèle sauvage lors de l'induction de la différenciation (Ogawa & Lee, 2003), aidant ainsi à maintenir la répression de *Xist* sur cet allèle une fois le choix effectué. La mutation induit donc un biais dans le choix de l'allèle qui sera inactivé, favorisant naturellement celui dans lequel *Tsix* sera moins efficacement exprimé, mais ce biais n'est pas total, l'X sauvage étant parfois inactivé. Contrairement aux autres mutations de la région (qui engendrent un biais total) (Penny *et al.*, 1996 ; Marahrens *et al.*, 1997 ; Lee *et al.*, 1999 ; Luikenhuis *et al.*, 2001 ; Clerc & Avner, 1998 ; Morey *et al.*, 2001), la délétion de *Xite* semble plutôt influencer la probabilité, par un mécanisme en *cis*, avec laquelle un chromosome X va être inactivé, ceci laissant entrevoir la possibilité que *Xite* pourrait être un candidat pour *Xce*.

De manière similaire à *DxPas34*, le locus *Xite* contient aussi plusieurs sites d'hypersensibilité à la DNase I régulés développementalement, certains de ces sites étant aussi fixés par CTCF et Oct4, les conclusions obtenues par rapport à *DxPas34* suite à la déplétion de ces protéines étant donc identiques. Par contre, Donohoe *et al.* décrivent aussi une faible fixation de la protéine Sox2 (aussi impliquée dans l'établissement d'un état de pluripotence au sein de la cellule) au niveau de *Xite* (Donohoe *et al.*, 2009). L'influence de cette protéine sur l'expression de l'antisens ne semble cependant pas définie. En effet, lors d'un test d'expression de la luciférase sous le contrôle du promoteur de *Tsix* et de *Xite*, la délétion du site de fixation de Sox2 (tout comme celui d'Oct4) semble engendrer une forte répression de *Tsix* dans des cellules ES mâles indifférenciées, indiquant donc que ces protéines semblent influencer la surexpression de l'antisens. Mais curieusement, lorsque Sox2 est déplétée par interférence

dans des cellules femelles à jour 4 de différenciation (par formation d'EBs), l'expression de *Tsix* et de *Xite* augmente d'un facteur 2. Il semblerait donc que Sox2 influe sur un autre mécanisme ou à partir d'un autre site de fixation, de manière à réprimer l'antisens, ce mécanisme étant apparemment dominant par rapport à l'effet de *Xite*. Il est aussi possible que le fait d'effectuer ces expériences dans des types cellulaires différents influe sur les résultats, le rôle de Sox2 sur l'expression de *Tsix* pouvant varier au cours du développement ou entre les sexes.

Concernant le rôle de la fixation de CTCF et Oct4 au niveau de *Xite* dans le « *pairing* », avec les données actuellement disponibles, nous n'avons aucun moyen de savoir si c'est la fixation de ces protéines à *DxPas34* ou à *Xite* qui influe sur ce phénomène. La région promotrice de *Tsix*, *DxPas34* et *Xite* arrivent donc tous à engendrer un *pairing* ectopique entre chromosome X et autosome (contenant le transgène) dans les ES femelles mais seul le transgène *Xite* à su recréer une interaction en *trans* chez le mâle (mais sans pour autant initier l'inactivation, dans les deux cas) (Lee, 2005, Xu *et al.*, 2006 *et al.*, 2007), indiquant peut-être une plus grande importance de cette région au sein de la transvection. Cependant, en délétant *Xite* dans les cellules sauvages, l'interaction entre les deux allèles n'est que retardée et non abrogée. D'autres parties du *XIC* semblent donc nécessaires pour récapituler l'inactivation dans les cellules transgéniques des deux sexes (voir le chapitre : « Choix », p15). Ces expériences témoignent aussi de l'importance de ce genre d'interactions en *trans* dans l'initiation de l'inactivation, laissant supposer un rôle essentiel de ce phénomène dans le choix et le comptage.

Ces éléments en 3' de *Xist* semblent aussi être impliqués dans une dernière fonction qui pourrait être reliée aux observations évoquées dans les paragraphes précédents. En 2008, Tsai *et al.* ont réalisé la première étude par 3C (pour « *Chromosomal Conformation Capture* ») (une technique permettant de visualiser les interactions inter-loci) dans la région de *Xist/Tsix* (Tsai *et al.*, 2008). Dans cet article, les auteurs décrivent deux domaines de structure chromatinienne distinctes, l'un concernant une interaction entre la région promotrice de *Tsix* et *Xite* et l'autre correspondant à une interaction entre *Jpx* et le promoteur de *Xist*, les deux étant régulées de manière développementale, les interactions semblant toujours plus importantes dans les cellules femelles à jour 0 qu'à jour 10 de différenciation en EBs. Le moment de formation de ces domaines corrélant à peu près avec le moment du *pairing* des deux

chromosomes X lors de l'initiation de l'inactivation pourrait faire croire à une liaison entre les mécanismes, une certaine structure chromatinienne étant peut-être nécessaire à une interaction en *trans* opérationnelle. Par ailleurs, la protéine CTCF semblerait être un très bon candidat pour médier ces interactions en *cis* (et en *trans*) de par son implication dans des phénomènes similaires à d'autres loci (Kurukuti *et al.*, 2006, Ling *et al.*, 2006, Splinter *et al.*, 2006).

En résumé, le transcrit non-codant antisens *Tsix* semble être exprimé de manière opposée à l'expression de *Xist* au cours de l'embryogénèse précoce puis tardive (où *Tsix* n'est plus exprimé). La régulation de l'antisens semble majoritairement être médiée par deux sites, chacun de part et d'autre du promoteur majoritaire, chacun contenant un îlot CpG, des sites d'hypersensibilité à la DNase développementalement et différenciellement orchestrés et une activité transcriptionnelle bi-directionnelle : le minisatellite *DxPas34* et le locus de transcription intergénique *Xite*. Diverses mutations du promoteur majoritaire de *Tsix* et/ou des ses deux « *enhancers* » ont démontré l'implication de ces éléments dans l'expression transcriptionnelle de l'antisens et à travers cela, dans l'initiation de l'inactivation du chromosome X.

Il est donc maintenant temps de se pencher sur le rôle fonctionnel de *Tsix* au sein de l'inactivation et plus précisément de voir comment ce dernier, ainsi que plusieurs autres facteurs, régulent l'expression de *Xist*.

VI. Régulation de *Xist*

Je me permets d'insérer ici une revue que j'ai pu rédiger au cours de mon doctorat pour le magazine Biofutur (n°304, novembre 2009), afin de présenter certains aspects de la régulation de *Xist* sous un aspect quelque peu "vulgarisé". Cependant, par la suite, je reprendrai les différents points abordés au cours de cet article en les détaillant davantage, permettant ainsi au lecteur d'acquérir une vue d'ensemble du processus avant d'en voir les détails.

Revue

X-plications sur le contrôle de l'inactivation : la régulation du gène *Xist*

*Le gène *Xist* est central à l'inactivation du chromosome X. Pour appréhender la régulation de ce processus, il est donc essentiel de comprendre comment l'expression de *Xist* est contrôlée de manière fine et efficace au cours de la mise en place et du maintien de l'inactivation. Si certains acteurs, comme des facteurs développementaux et un transcrit antisens à *Xist* sont déjà connus, d'autres restent à découvrir, en particulier chez des espèces de mammifères encore peu étudiées.*

Andrew Oldfield, Claire Rougeulle

UMR 7216 Epigénétique et destin cellulaire, Université Paris-Diderot Paris 7, CNRS, 35, rue Hélène Brion 75205 Paris Cedex 13
claire.rougeulle@univ-paris-diderot.fr

Dans chaque cellule d'une femelle mammifère adulte, un des deux chromosomes X est aléatoirement inactivé – transcriptionnellement éteint – dans sa quasi-intégralité. L'inactivation du chromosome X est la réponse trouvée par l'évolution pour pallier la divergence gonosomique entre mâle (XY) et femelle (XX). Ce phénomène met ainsi les deux sexes sur un pied d'égalité en limitant la quantité de transcrits provenant des chromosomes X présents dans les cellules femelles.

L'inactivation, un processus dynamique

L'inactivation du chromosome X se met en place très tôt au cours de l'embryogenèse. Chez la souris, organisme modèle pour l'étude de l'inactivation, elle se déroule en plusieurs étapes (figure p.29). La première vague d'inactivation, qui touche systématiquement le chromosome X d'origine paternelle (Xp), débute rapidement après la fécondation et perdure dans les tissus extra-embryonnaires. Au moment de l'implantation, alors que se séparent les lignages embryonnaires et extra-embryonnaires, une transition importante se produit dans la masse cellulaire interne (MCI). En effet, dans ces cellules capables de générer tous les tissus de l'embryon, et donc considérées comme pluripotentes, l'Xp est réactivé et les deux chromosomes X sont actifs. Les cellules souches embryonnaires (CSE), dérivées de la MCI et

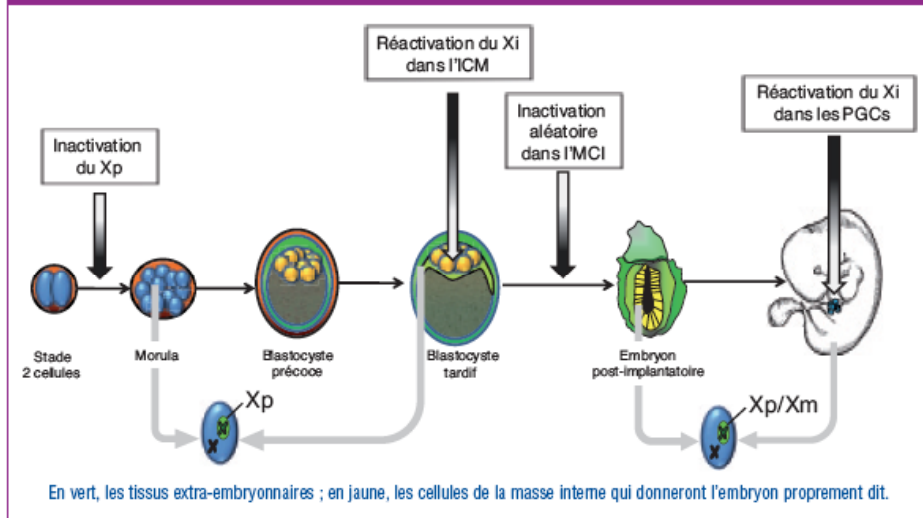
pluripotentes elles aussi, maintiennent également leurs deux chromosomes X actifs. La différenciation de ces cellules pluripotentes in vitro, tout comme celle de la MCI au cours du développement, s'accompagne d'une vague d'inactivation aléatoire qui concerne cette fois soit l'Xp, soit l'X maternel (Xm). Les CSE représentent ainsi un modèle de choix pour l'étude de l'inactivation aléatoire. Une fois mise en place, l'inactivation aléatoire est stable et héritée de manière clonale au cours des divisions cellulaires successives, à l'exception des cellules de la lignée germinale dont l'X se réactive avant l'entrée en méiose.

Le gène *Xist* contrôle l'inactivation

L'acteur principal de l'inactivation est un gène du chromosome X, *Xist* (*X-inactive specific transcript*) (figure p. 30). Celui-ci produit un long ARN non codant (ARNnc) qui reste au sein du noyau et, lorsqu'il est produit en quantité suffisante, recouvre le chromosome X à partir duquel il est transcrit pour l'inactiver. *Xist* est absolument essentiel à l'inactivation : un chromosome X portant une version mutante de *Xist* ne peut jamais être inactivé (1). En revanche, l'expression forcée d'une copie de *Xist* artificiellement placée sur un autre chromosome induit l'inactivation de ce dernier (2). Étant donné son rôle primordial dans l'inactivation, il n'est

(1) Penny GD *et al.* (1996) *Nature* 379, 131-7
(2) Wutz A, Jaenisch R (2000) *Mol Cell* Apr 5, 695-705

Dynamique de l'inactivation du chromosome X au cours de l'embryogenèse précoce



pas surprenant de constater que l'expression de *Xist* varie considérablement au cours de l'embryogenèse précoce. La découverte des facteurs contrôlant son expression est donc une étape essentielle pour comprendre comment l'inactivation du chromosome X s'établit. Le premier facteur à intervenir dans l'initiation de l'inactivation, et donc dans la régulation de *Xist*, est le nombre de chromosomes X présents dans la cellule. En effet, l'unique chromosome X du mâle n'est jamais inactivé et il faut au moins deux X pour que l'inactivation se mette en place. Si l'on ne comprend pas encore très bien comment la cellule arrive à compter, on sait en revanche que l'organisation nucléaire, et plus précisément l'existence d'une interaction physique transitoire entre les deux chromosomes X de la femelle, est une étape importante à la mise en place du processus (voir l'article p. 32).

Outre cet aspect, la régulation de *Xist* implique donc essentiellement deux paramètres :

- le taux d'expression : des niveaux élevés d'ARN *Xist* doivent être produits pour permettre l'inactivation du chromosome en *cis** ;
- l'origine allélique du transcrit : chez les femelles, *Xist* ne doit être exprimé qu'à partir d'un des deux chromosomes X afin que l'autre soit maintenu dans un état actif.

Des facteurs développementaux contrôlent l'inactivation en régulant *Xist*

Bien que l'inactivation du chromosome X soit extrêmement dynamique durant les étapes précoces du développement, les périodes pendant lesquelles les deux chromosomes X de la femelle sont actifs sont très limitées. Seules les cellules pluripotentes, présentes de manière transitoire dans l'embryon, semblent tolérer la présence de deux X actifs. En effet, au cours du développement, l'inactivation n'est réversible de manière spontanée qu'en deux occasions : dans la MCI et dans les cellules germinales primordiales, deux types

cellulaires pluripotents. De plus, dans le cas de cellules femelles, la perte de pluripotence qui se produit lors de la différenciation cellulaire s'accompagne de l'inactivation d'un des chromosomes X. Enfin, la dédifférenciation (reprogrammation artificielle) de cellules somatiques femelles en cellules pluripotentes (cellules pluripotentes induites par exemple) s'accompagne de la réactivation du chromosome X initialement inactif (3).

Toutes ces observations suggèrent donc un lien inverse entre inactivation du X et pluripotence. La nature de ce lien a été mise à jour en 2008 par Pablo Navarro et ses collègues de l'Institut Pasteur et du *Medical Research Council* d'Édimbourg (4) qui ont démontré que trois facteurs de transcription essentiels pour l'état pluripotent, Oct3/4, Nanog et Sox2, sont directement impliqués dans la régulation de *Xist*.

Cette étude a montré que ces trois facteurs interagissent physiquement avec le gène *Xist* dans les cellules pluripotentes, où ils sont simultanément exprimés et où *Xist* est réprimé. L'expression de ces facteurs étant spécifique des cellules indifférenciées, cette fixation est naturellement perdue une fois la différenciation cellulaire enclenchée et l'inactivation initiée. Une abolition rapide et simultanée de la fixation de Oct3/4, Nanog et Sox2 au niveau de *Xist* dans un contexte artificiel induit une expression importante de ce dernier et ce, même dans des cellules mâles, dans lesquelles *Xist* n'est normalement jamais exprimé. Ces résultats démontrent que Nanog, Oct3/4 et Sox2 sont des répresseurs essentiels de *Xist* qui agissent en synergie pour coupler l'inactivation à la pluripotence (figure p. 31).

Si ces données s'intègrent parfaitement dans l'inactivation aléatoire, le rôle (s'il existe) des facteurs de pluripotence au cours de l'inactivation préimplantatoire n'est pas encore défini (5). Toutefois, ces découvertes prodigieuses permettent d'ores et déjà d'établir un premier lien moléculaire entre le développement embryonnaire, la différenciation cellulaire et l'inactivation du chromosome X.

- (3) Maherall N *et al.* (2007) *Cell Stem Cell* 1, 55-70
 (4) Navarro P *et al.* (2008) *Science* 321, 1693-5
 (5) Navarro P, Avner P (2009) *FEBS Lett* 583, 1721-7

* *Xist* n'inactive que le chromosome à partir duquel il est produit.

Un transcrit antisens participe à la régulation de *Xist*

Les facteurs de pluripotence jouent donc un rôle important dans le contrôle du niveau d'expression de *Xist*. Ils ne sont cependant pas les seuls régulateurs connus. Un transcrit antisens à *Xist*, nommé *Tsix*, dont la région transcriptionnelle recouvre intégralement celle de *Xist* (figure p.30), joue également un rôle prépondérant dans la régulation de celui-ci.

Au cours du développement, l'expression de cet antisens est principalement restreinte aux cellules pluripotentes de la masse interne, quand l'Xp est réactivé et les deux chromosomes X actifs. Lorsque ces cellules commencent à se différencier, l'expression de l'antisens est fortement réduite sur le futur X inactif (Xi) de manière concomitante à l'augmentation de *Xist* sur ce même chromosome. Sur l'X qui restera actif (Xa), l'expression de *Tsix* persiste provisoirement puis s'arrête. D'après ces données ainsi que plusieurs études basées sur diverses mutations de *Tsix*, celui-ci semblerait bien être un répresseur supplémentaire de *Xist*, qui agirait cette fois en *cis*. Cependant, il aura fallu attendre 2006 pour connaître avec précision le niveau auquel agit cet ARNnc. À cette date, plusieurs équipes ont démontré que cette répression s'effectue à plusieurs niveaux (6). Une des conséquences de la perte de *Tsix* est une augmentation significative des transcrits *Xist*. Cependant, la transcription de *Tsix* ne semble pas en être affectée, ce qui suggère que *Tsix* joue un rôle au niveau de la stabilité de l'ARN *Xist*, en favorisant par exemple sa dégradation par des mécanismes encore inconnus.

Tsix possède également des activités de remodelage de la chromatine qui touchent l'ensemble de sa région transcrite et plus particulièrement la région promotrice de *Xist* (figure p. 31). Dans la mesure où, au moment de la réversion de l'inactivation du Xp, *Tsix* est produit par les deux chromosomes X actifs dans les cellules de la MCI, il établit un profil chromatinien identique au niveau du locus sur les deux allèles. D'où l'hypothèse que cette activité jouerait un rôle essentiel dans la transition entre les inactivations préimplantatoire spécifique du Xp et post-implantatoire aléatoire, en « effaçant », au niveau de *Xist* des marques chromatiniques asymétriques associées à cette première forme d'inactivation. Suite à cela, la cellule serait donc apte à faire un « choix aléatoire » quant au chromosome X à inactiver.

Par son activité de remodelage de la chromatine au niveau du promoteur de *Xist*, *Tsix* joue aussi un rôle prépondérant dans ce choix (7). Dans cette région, il induit une structure de type répressif. En l'absence de *Tsix*, la région acquiert donc une conformation permissive à une transcription éventuelle. Lors de l'initiation de l'inacti-

vation aléatoire, l'extinction transcriptionnelle de *Tsix* sur un des deux allèles génère une asymétrie chromatinienne au niveau du promoteur de *Xist*, la structure la plus favorable à l'activation de *Xist* se retrouvant sur le chromosome sur lequel *Tsix* n'est plus exprimé (figure ci-contre). La manière dont ces modifications chromatiniques sont engendrées reste inconnue mais deux types de mécanismes peuvent être évoqués. Le premier fait intervenir le transcrit, potentiellement via le processus d'interférence à ARN. Le second suppose que les modifications sont faites de manière concomitante à la transcription grâce à des interactions directes entre la machinerie de transcription et les modificateurs d'histones.

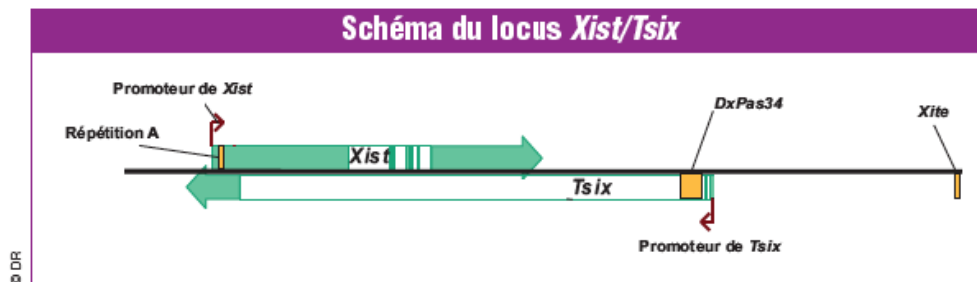
D'autres facteurs interviennent pour activer *Xist*

Xist est donc tout à la fois soumis aux activités de remodelage de *Tsix*, qui exerce ainsi un contrôle allélique sur son expression, et à des facteurs de pluripotence, qui régulent ses niveaux d'expression et assurent le couplage moléculaire entre l'inactivation du chromosome X et le développement embryonnaire. Ces facteurs interviennent principalement au niveau de la répression de *Xist* et, étonnamment, nous n'avons que peu de données concernant l'activation de ce dernier. Seuls deux facteurs candidats, YY1 et CTCF, dont le rôle dans l'activation transcriptionnelle de *Xist* est encore mal défini, ont été décrits (8,9).

Très récemment, le groupe de Takashi Sado, à l'Institut national de génétique du Japon, a apporté un éclaircissement nouveau sur le rôle d'une région de *Xist* : la répétition A (10). Cette région répétée située dans le premier exon de *Xist* (figure ci-dessous) était connue comme la région responsable du pouvoir inactivant de cet ARNnc (11). En l'absence de cette région, l'ARN *Xist* est, en effet, capable de recouvrir le chromosome mais pas d'induire son inactivation. Grâce à une étude *in vivo* chez des souris portant une délétion de la répétition A, Sado et ses collaborateurs ont fait état d'un rôle de cette région dans la régulation de *Xist*. En accord avec le rôle initial de cette région, chez des souris mutantes, le chromosome portant la mutation ne peut être inactivé. Cependant, et de manière surprenante, ce phénomène n'est pas lié en soi à un problème d'inactivation mais à une expression très faible de *Xist* qui pourrait s'expliquer par le fait que, parallèlement, *Tsix* est anormalement exprimé. Ainsi, le rôle joué par la répétition A dans le contrôle de *Xist* serait indirect et passerait par l'antisens *Tsix*. Les mécanismes de cette nouvelle régulation sont pour l'heure inconnus mais ces données soulignent l'importance de connaître la régulation de *Tsix* pour comprendre celle de *Xist*.

- (6) Navarro P *et al.* (2005) *Genes Dev* 19, 1474-84
- (7) Navarro P *et al.* (2006) *Genes Dev* 20, 2787-92
- (8) Kim JD *et al.* (2008) *Genome Res* 16, 901-11
- (9) Pugacheva EM *et al.* (2005) *Hum Mol Genet* 14, 953-65
- (10) Hoki Y *et al.* (2009) *Development* 136, 139-46
- (11) Wutz A *et al.* (2002) *Nat Genet* 30, 167-74

En vert, les exons de *Xist* et *Tsix* ; la flèche indique le sens de la transcription. En orange, certaines des régions impliquées dans la régulation de *Xist* et/ou de son antisens.



Régulation de *Tsix*

Comme dans le cas de *Xist*, la régulation de *Tsix* est complexe et contrôlée par plusieurs éléments. Outre la répétition A, dont la fonction régulatrice de *Tsix* récemment révélée est encore mal connue, deux régions situées respectivement en aval et en amont du promoteur de *Tsix* jouent des rôles activateurs : la région répétée *DXPas34* (figure 2) (12) et la région *Xite* (*X-inactivation intergenic transcription element*, figure 2) (13). On trouve au sein de ces deux régions de nombreux sites de fixation de plusieurs facteurs qui contrôlent ainsi l'expression de *Tsix*. Il est intéressant et frappant de constater que nombre de ces facteurs participent également directement à la régulation de *Xist*. C'est le cas en particulier des protéines YY1 et CTCF. De plus, très récemment, l'équipe de Jeannie Lee, de la *Harvard Medical School*, a suggéré que le facteur de pluripotence Oct3/4 (voire Sox2) jouerait également un rôle d'activateur transcriptionnel de *Tsix* en se fixant au niveau de *DXPas34* et de *Xite* (14).

CTCF intervient donc à la fois dans la régulation de *Xist* et de son antisens *Tsix*. Une des particularités fonctionnelles de cette protéine est qu'elle intervient dans l'organisation tridimensionnelle de la chromatine (15), permettant ainsi de mettre en contact des régions du génome physiquement distantes. La fixation de CTCF à plusieurs endroits stratégiques du locus *Xist/Tsix* (promoteur de *Xist*, *DXPas34*, *Xite*) suggère ainsi qu'une partie de la régulation transcriptionnelle de *Xist/Tsix* pourrait passer par le remodelage de la conformation chromatinienne environnante. Afin d'étudier cette possibilité, une première étude a été menée par l'équipe de Jeannie Lee en se servant de la technique 3C (*chromosomal conformation capture*) (16). Les données obtenues, bien qu'étant très difficiles à interpréter du fait de la proximité des sites observés, nous confortent dans l'idée d'interactions entre les différentes régions jouant un rôle dans la mise en place de l'inactivation du chromosome X. Prometteuses, ces avancées nécessitent néanmoins d'être approfondies de manière à pouvoir établir un modèle dynamique fiable des remaniements chromatinien se déroulant au cours de l'inactivation.

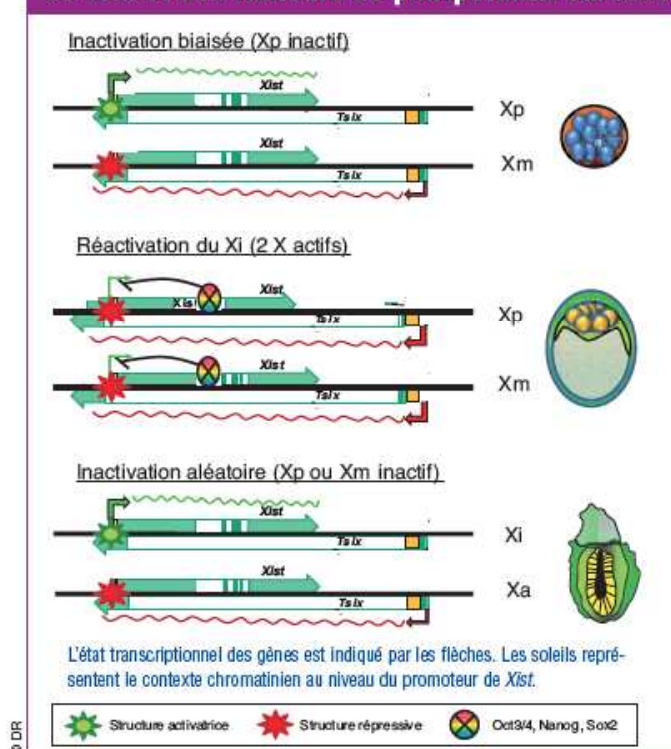
Conservation des mécanismes au cours de l'évolution

L'inactivation du chromosome X est un processus que l'on retrouve chez tous les mammifères, à l'exception des monotrèmes (voir p. 38). On relève cependant des différences phénotypiques et mécanistiques plus ou moins sévères entre les espèces dont la plus frappante concerne les marsupiaux, chez lesquels l'inactivation se produit en l'absence de gène *Xist*.

Bien que *Xist* soit responsable de l'inactivation aléatoire chez tous les mammifères placentaires, les mécanismes contrôlant son expression au sein des organismes autres que la souris, considérée ici comme organisme référent, demeurent encore peu connus.

Chez l'humain, à l'inverse de la souris, l'inactivation semble aléatoire même dans les tissus extraembryonnaires. La présence d'un homologue fonctionnel de *Tsix* est également controversée. Il est probable que ces différences reflètent l'existence d'un réseau de régula-

Modèle dynamique de l'action répressive de *Tsix* et des facteurs de pluripotence sur *Xist*



tion de *Xist* propre à chaque espèce. Toutefois, l'étude des étapes précoces de l'inactivation et de la régulation de ce gène chez l'homme a été jusqu'à présent limitée par des questions éthiques évidentes qui rendent le matériel biologique adéquat, l'embryon, peu accessible.

L'arrivée des CSE humaines comme modèle d'étude est très prometteuse, même si les données préliminaires indiquent une complexité importante. En particulier, et toujours à l'inverse de la souris, l'inactivation chez l'humain semble pouvoir se mettre en place dans des CSE indifférenciés et pluripotents.

D'après nos hypothèses sur la fonction du *Tsix* murin dans la transition entre inactivation préimplantatoire biaisée et inactivation aléatoire, et du fait de l'incertitude concernant l'existence d'un antisens humain, il serait tentant de corrélérer chez l'homme l'absence simultanée de *Tsix* et d'inactivation biaisée.

Le modèle bovin tend à infirmer cette hypothèse, puisque les données récentes indiquent l'existence d'une inactivation biaisée préimplantatoire en l'absence d'un antisens analogue à celui décrit chez la souris. Les données disponibles pour cette espèce sont cependant très préliminaires et des études supplémentaires devront être menées avant de pouvoir conclure définitivement sur ces deux aspects.

Si de nombreuses questions demeurent, tous ces exemples tendent à montrer qu'il existe des divergences importantes entre espèces. Une connaissance approfondie des mécanismes contrôlant *Xist* dans différentes espèces est donc essentielle pour appréhender, d'un point de vue évolutif, leur conservation et, par-là même, la plasticité avec laquelle une finalité, ici l'inactivation d'un chromosome X, peut être atteinte. ●

(12) Vigneau S et al. (2006) *Proc Natl Acad Sci USA* 103, 7390-5

(13) Stavropoulos N et al. (2005) *Mol Cell Biol* 25, 2757-69

(14) Donohoe ME et al. (2009) *Nature* 460, 128-32

(15) Kurukuti S et al. (2006) *Proc Natl Acad Sci USA* 103, 10684-9

(16) Tsai CL et al. (2008) *Dev Biol* 319, 416-25

a. Régulation de *Xist* par *Tsix*

Comme nous avons pu l'entre-apercevoir au cours de cette introduction, l'expression de *Tsix* semble anti-corréler parfaitement celle de *Xist* dans tous les types cellulaires étudiés (à part lors des stades tardifs des PGCs où ni *Xist*, ni *Tsix* ne sont exprimés). De nombreuses expériences ont d'ailleurs montré que toute mutation engendrant une transcription incorrecte de l'antisens permet la surexpression de *Xist* en *cis*, et ce, aussi bien dans les cellules ES mâles que femelles (Clerc & Avner, 1998 ; Morey *et al.*, 2001, 2004 ; Luikenhuis *et al.*, 2001 ; Vigneau *et al.*, 2006 ; Sado *et al.*, 2002 ; Sado *et al.*, 2001 ; Lee *et al.*, 1999 ; Shibata & Lee, 2004 ; Sun *et al.*, 2006).

Afin d'étudier la manière dont *Tsix* établit cette répression, plusieurs équipes ont cherché à caractériser les marques épigénétiques du locus *Xist/Tsix*, en se focalisant sur la région promotrice de *Xist* (ou l'exon 4 de *Tsix*), au sein de cellules ES sauvages et de cellules ES chez qui la transcription de *Tsix* est bloquée (Navarro *et al.*, 2005 *et al.*, 2006, Sun *et al.*, 2006, Sado *et al.*, 2005, Ohhata *et al.*, 2008, Shibata & Yokota, 2008).

i. Régulation post-transcriptionnelle de *Xist* par *Tsix*

L'étude de la lignée ES mâle mutante Ma2L, contenant une séquence d'arrêt transcriptionnel en amont du promoteur de *Tsix*, empêchant ainsi la transcription antisens d'atteindre le gène *Xist*, et de sa lignée contrôle Ma1L (Luikenhuis *et al.*, 2001), a permis à Navarro *et al.* d'établir que l'expression de *Tsix* joue un rôle de répresseur post-transcriptionnel vis-à-vis de *Xist* (Navarro *et al.*, 2005).

Avant la parution de cet article, le promoteur P2 était supposé être le promoteur principal de *Xist* au vu du nombre plus important de transcrits provenant de celui-ci (Johnston *et al.*, 1998), s'ajoutant à ceci, la sur-expression de *Xist* en cours de différenciation était censée être liée à la stabilisation du transcrit pendant ce processus (Panning *et al.*, 1997, Sheardown *et al.*, 1997).

Navarro *et al.* ont cependant démontré, grâce à des expériences d'immunoprécipitation de la chromatine (ChIP), que la machinerie de transcription (dont TFIIB, RNAPolII, TFIIA, TFIIIE, TFIIF, TFIIH) n'est pas recrutée aux promoteurs de *Xist* dans les cellules indifférenciées mutées (Ma2L) par rapport au niveau basal présent dans les cellules contrôles (Ma1L), et ce, malgré une forte augmentation du nombre de transcrits *Xist* (visibles par FISH et RT-PCR)

suite à la mutation de l'antisens. En 2006, Sun *et al.* ont confirmé l'absence d'augmentation de transcription de *Xist* dans des cellules ES mâles mutées pour *Tsix* par des expériences de « run-on » (Sun *et al.*, 2006). Ceci nous indiquant que *Tsix* régule bien *Xist* mais au niveau post-transcriptionnel (Figure 23).

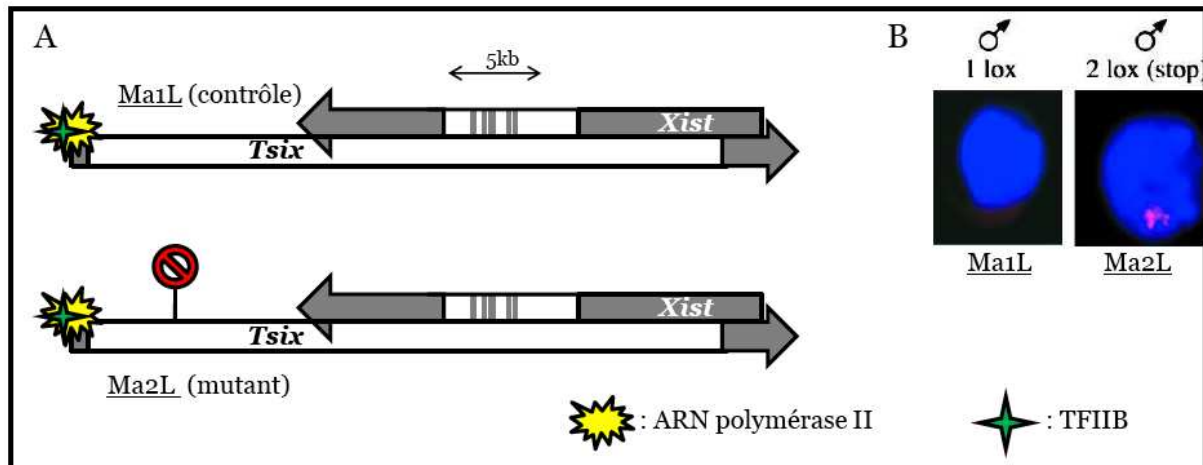


Figure 23 : Schéma de l'analyse transcriptionnelle de cellules ES mâles tronquées ou pas pour *Tsix* en cours de différenciation. Les cellules Ma2L portent un signal d'arrêt transcriptionnel 4kb en aval du promoteur de *Tsix*. Ma1L sont les cellules contrôle correspondantes. (A) Schéma des résultats obtenus par analyse par ChIP de la distribution d'ARN polymérase II et de TFIIB au sein du locus *Xist/Tsix* dans Ma1L et Ma2L. (B) Analyse par FISH de l'expression de *Xist* dans les cellules Ma1L et Ma2L démontrant une surexpression de son transcrit, mais régulé de manière post-transcriptionnelle par l'antisens.

Ces observations supputent malgré tout l'existence d'un mécanisme, agissant indépendamment de *Tsix*, qui réprimerait *Xist* transcriptionnellement, et ce, de manière développementale (agissant uniquement au stade indifférencié), le PIC (pour « Pre-Initiation Complex ») étant recruté à *Xist* seulement lors de la différenciation cellulaire.

La manière dont s'effectue cette régulation post-transcriptionnelle demeure encore mystérieuse. Mais dans le but de discriminer entre un effet de l'action de la transcription de *Tsix* à travers *Xist* ou d'une action du transcrit même, les expériences de Shibata et Lee proposent que le phénotype lié à la délétion de *Tsix* ne peut être réverté lors de l'introduction d'un ADNc exprimant sa version épissée (Shibata & Lee, 2004) indiquant donc une nécessité du passage de la transcription. Cependant, ces expériences ne permettent malheureusement pas de discriminer entre le fait que la transcription de *Tsix* à travers le gène *Xist* pourrait être nécessaire à sa régulation, que le mécanisme de répression serait couplé à la machinerie d'épissage ou encore que le transcrit *Tsix* non-épissé (qui est d'ailleurs la forme majoritaire)

pourrait être nécessaire à ce processus. Certaines de ces lacunes ont été contournées par des expériences de Sado *et al.*, qui ont délété le site accepteur d'épissage de l'exon 4, afin d'analyser la nécessité de l'épissage de l'antisens dans la régulation de l'expression de *Xist* (notamment médiée par le contexte chromatinien de sa région promotrice) (Sado *et al.*, 2006). De manière similaire à Shibata et Lee, Sado *et al.* proposent que l'absence de *Tsix* sous sa forme épissée n'influe aucunement sur la répression de *Xist* médiée par *Tsix* ou sur la configuration chromatinienne caractéristique du promoteur de *Xist* (Sado *et al.*, 2006).

L'hypothèse de la nécessité de la transcription de *Tsix* plutôt que son transcrit a aussi été privilégiée par Ohhata *et al.* (Ohhata *et al.*, 2008) qui n'ont pas observé d'effet répressif de *Tsix* sur *Xist* dans leur lignée mutante pour *Tsix*, malgré le fait que 93% de l'antisens étaient transcrit, bien qu'évidemment les 7% restants correspondant à la région promotrice de *Xist* pourraient naturellement représenter une source de transcrits d'une importance capitale dans l'hypothèse d'interactions entre les transcrits sens et antisens.

Une des possibilités pour expliquer cette répression consisterait en la déstabilisation des transcrits *Xist* par *Tsix*, telle que ce qui avait été proposé par Panning et Sheardown, toutefois ces données ont été remises en cause par le laboratoire de Jeannie Lee, qui n'observe aucune modification de la stabilité de l'ARN *Xist*, que ce soit entre l'état indifférencié-différencié ou bien entre les cellules mutées ou non pour *Tsix* (Sun *et al.*, 2006). Toutefois, cette analyse de stabilité ne porte que sur les transcrits totaux de *Xist* mais ne prend pas en compte ses formes épissées ; or, il a été montré que l'abondance des transcrits *Xist* dans des cellules ES femelles différenciées pouvait être régulée au moment de leur maturation par la voie NMD (pour « *non-sense RNA mediated decay* »). Plus spécifiquement, la voie NMD est requise pour l'obtention de transcrits épissés, mais n'a pas d'effet sur le niveau des transcrits non épissés de *Xist* (Ciaudo *et al.*, 2006). La modification de la stabilité d'ARNs *Xist* épissés est donc envisageable, soit par le système NMD, soit par un mécanisme encore inconnu de *Tsix*.

ii. ARNi

Dû au fait que l'exon 4 de *Tsix* se superpose à la région promotrice de *Xist* et que celui-ci est transcrit en faible quantité dans les cellules ES mâles comme femelles, une des questions les plus débattues dans le domaine de l'inactivation était de savoir si le phénomène de siRNA joue

un rôle dans la régulation de *Xist*, soit en induisant des modifications épigénétiques directement au niveau du promoteur, soit en déclenchant la dégradation de ses transcrits.

A l'heure actuelle, chez les mammifères, il n'y a aucune preuve de l'existence d'un lien direct entre la machinerie à ARNi ou de petits ARNs et l'établissement ou la maintenance d'un état chromatinien hétérochromatique (constitutif ou facultatif), contrairement à chez d'autres eucaryotes telles que les plantes ou la drosophile. Afin de savoir ce qu'il en est vraiment, ces dernières années, plusieurs équipes se sont attelées à prouver l'existence de petits ARNs liés à l'inactivation du chromosome X ainsi que d'étudier le rôle de Dicer (responsable de la formation de petits ARNs à partir d'ARN double-brin chez les mammifères) à travers diverses mutations de celui-ci (Ogawa *et al.*, 2008, Nesterova *et al.*, 2008, Kanellopoulou *et al.*, 2009).

L'article d'Ogawa *et al.* est le seul, à ce jour, à prouver l'existence de petits ARNs issus de duplexes *Xist:Tsix*. En effet, ce papier décrit l'existence de duplexes d'ARN double-brin entre *Xist* et *Tsix*, ceux-ci étant majoritairement présents dans les cellules ES. Lors de l'entrée des cellules en différenciation, la quantité de ces duplexes diminue tandis qu'on observe l'apparition de petits ARNs, baptisés xiRNA, d'une taille de 24 à 42 nucléotides et provenant de diverses localisations de *Xist* et des deux brins, dont le *Repeat-A*, l'exon 7 et la région promotrice. Ces xiRNA deviennent de nouveau non-mesurables dans les cellules somatiques, leur expression correspondant donc à la fenêtre de l'initiation de l'inactivation du chromosome X. Leur présence est aussi corrélée à la transcription de l'antisens, leur nombre chutant lorsque *Tsix* est muté. De plus, bien que dans les cellules ES les duplexes soient produits à partir des deux allèles, il semblerait que seuls ceux provenant du Xa soient ensuite dégradés en xiRNA en cours de différenciation. L'hypothèse des auteurs étant que les duplexes sont initialement formés à partir des deux allèles, puis lorsque le Xa va être choisi en cours de différenciation, la transcription de *Tsix* va se prolonger légèrement plus longtemps sur cet allèle et engendrer la formation de xiRNAs qui seraient responsables d'une répression transcriptionnelle du promoteur de *Xist* en cis, de manière similaire au TGS (pour « *Transcriptional Gene Silencing* ») observé chez la levure (Grewal & Elgin, 2007, Volpe *et al.*, 2002, Kim *et al.*, 2006, Morris *et al.*, 2004). Les auteurs ont ensuite voulu savoir si l'enzyme Dicer avait un rôle dans la fabrication de ces petits ARNs. Et bien que les observations d'Ogawa *et al.* tendent à confirmer que cette enzyme pourrait jouer un rôle dans l'inactivation, les résultats et les conclusions d'autres publications divergent quant à cette hypothèse et laissent planer un

doute de son implication dans ce phénomène. Ogawa *et al.* étudient en effet la fonction de Dicer grâce à des mutants *Dicer*^{-/-} au sein desquels ils expriment tout de même cette protéine de manière faible (environ 5%) afin d'échapper à son phénotype léthal lorsque totalement absent. Dans ces cellules, ils observent une diminution des xiRNAs lorsque la différenciation est induite ainsi qu'un manque de domaine de *Xist* et du marquage d'un des chromosomes X par H3K27me3 caractéristiques du Xi, ceci malgré l'augmentation du nombre de transcrits *Xist* aussi bien dans les cellules ES mâles que femelles corrélant avec l'observation d'une hypométhylation de l'ADN en 5' de *Xist*. Cette augmentation ectopique de *Xist* dans les cellules ES n'est cependant pas maintenue en cours de différenciation, la quantité de transcrits restant inférieure d'un facteur 10 par rapport aux cellules sauvages, ce qui pourrait expliquer l'incapacité à former des domaines et pourrait être dû au fait que la différenciation de ces cellules mutantes se fait excessivement mal (des quantités importantes de Oct4 et Nanog étant encore présentes en fin de différenciation). Les auteurs effectuent ensuite une mutation transcriptionnelle d'un des allèles de *Tsix* dans les cellules *Dicer*^{-/-} et observent la réapparition des domaines de *Xist* dans ces cellules bien que ceux de H3K27me3 demeurent absents. Dicer pourrait cependant aussi avoir un effet sur *Tsix*, la transcription de l'antisens n'étant jamais aussi bien réprimé en cours de différenciation dans les cellules *Dicer*^{-/-} que dans les cellules sauvages, ceci étant même observable, de manière fort paradoxale, dans les cellules *Tsix*^{+/-} (peut-être dû aux problèmes de différenciation). La délétion de l'antisens sur un allèle devrait donc empêcher la formation de duplexes en *cis*, corrélant avec l'hypothèse que l'X muté devenant toujours être le Xi. Cependant, dans un contexte *Dicer*^{-/-} on peut supposer que l'augmentation de transcrits *Xist* soit due au fait que d'une part un des moyens de répression du gène ait été retiré (*Tsix*) et d'autre part, qu'il n'y a plus de transcrits pris dans les duplexes, ils peuvent donc tous participer à la formation d'un domaine.

Les deux autres équipes ayant étudié le rôle de Dicer dans l'inactivation du chromosome X semblent lui accorder bien moins d'importance que ne le fait l'équipe de Jeannie Lee. La première étude, dirigée par Neil Brockdorff, observe que la transcription sens et anti-sens passant par la région promotrice de *Xist* influence sa méthylation de l'ADN (une surexpression en sens induisant une hypométhylation tandis que la transcription de l'antisens est bien connue comme induisant une méthylation de l'ADN au niveau du promoteur de *Xist*, (voir le chapitre suivant) (Nesterova *et al.*, 2008). Cependant, bien que leurs expériences dans

des cellules mutées pour Dicer engendrent une hypométhylation du promoteur de *Xist* ainsi qu'une augmentation du nombre de transcrits dans les ES (de manière similaire à l'étude d'Ogawa *et al.*), ce phénomène est observé à travers de nombreux loci de gènes empreintés (tels que *H19* et *Igf2rAir*) indiquant une perturbation globale de la méthylation de l'ADN, contrairement à l'effet observé par Ogawa qui était restreint à *Xist*. Et se basant sur des articles publiés précédemment stipulant que la mutation de Dicer engendre la réduction des méthyltransférases d'ADN *de novo* (Dnmt3A et 3B), Nesterova *et al.* postulent que l'effet de Dicer sur *Xist* n'est pas direct mais passe plutôt par la régulation de ce dernier sur les Dnmts, engendrant des modifications du profil de méthylation du promoteur de *Xist*. De plus, cette équipe n'observe pas de défaut de fabrication de domaine hétérochromatique par *Xist* dans les mutants pour Dicer.

Et enfin, la dernière équipe à se pencher sur le sujet n'observe tout simplement aucun effet de la délétion de Dicer sur l'inactivation du chromosome X (Kanellopoulou *et al.*, 2009) à part une légère surexpression de *Xist* dans les cellules ES, ceci étant dû à un phénomène post-transcriptionnel, la quantité d'ARN *Xist* non-épissé restant constante entre les cellules contrôles et les mutées *Dicer* *-/-*.

En bref, bien que les trois études convergent sur l'observation d'une surexpression de *Xist* dans les cellules indifférenciées en l'absence de Dicer (peut-être due à son effet sur les ADN méthyltransférases), l'implication directe de cette enzyme dans le processus d'inactivation du chromosome X n'est pas encore clairement démontrée, ceci étant peut-être attribuable à la diversité des méthodes utilisées et des mutants engendrés. Un paradoxe supplémentaire vient du fait que Dicer ait été décrite comme étant une protéine cytoplasmique et qu'aucun fragment de l'ARN de *Xist* ou de *Tsix* n'ait déjà été retrouvé à cet endroit, d'autant plus que Dicer est connu pour engendrer des fragments d'ADN allant de 21 à 25pb, ce qui ne semble pas vraiment correspondre à ce qui est observé par Ogawa *et al.*.

Afin d'y voir plus clair, il est nécessaire de savoir comment sont obtenus les xiRNAs, de découvrir comment ils seraient fabriqués par Dicer au vu de sa localisation cytoplasmique et de savoir ensuite comment ils agissent au niveau du promoteur de *Xist*. Et qui plus est, *Tsix* pourrait aussi être régulé par un mécanisme basé sur les siRNAs de par sa régulation par des régions à transcription bi-directionnelle.

iii. Modification du paysage épigénétique au promoteur de *Xist* par *Tsix*

Bien que nous venons de voir que l'expression de *Tsix* engendre une répression post-transcriptionnelle de *Xist*, dont le processus demeure encore inconnu, il a longtemps été spéculé que la transcription de l'antisens engendre des modifications des caractéristiques épigénétiques de la chromatine le long du locus *Xist/Tsix* (Morey *et al.*, 2004, Vigneau *et al.*, 2006) et plus précisément au niveau du promoteur de *Xist*.

Ces dernières années, plusieurs laboratoires se sont penchés sur ce sujet, et grâce à divers mutants transcriptionnels de l'antisens, ils ont pu observer diverses répercussions engendrées par l'absence de transcription de celui-ci à travers la région promotrice de *Xist* sur le contexte chromatinien caractérisant cette région.

En effet, l'absence de transcription à travers la région promotrice de *Xist*, que ce soit dans les cellules mâles Ma2L, dans les mâles Xdc (dont le promoteur de *Tsix* est délété) ou bien les mâles X^{PA} (contenant un arrêt de transcription dans la partie 3' de *Tsix*), semble y engendrer des marques de type euchromatique (Navarro *et al.*, 2005 *et al.*, 2006 *et al.*, 2009, Sado *et al.*, 2005, Ohhata *et al.*, 2008). Au stade ES, Navarro *et al.* observent dans ces mutants une augmentation en *cis* des marques activatrices H3K4me2, H3K4me3, H3K9Ac alors que sont diminuées les marques répressives H3K9me3 et la méthylation des CpG (Navarro *et al.*, 2005 *et al.*, 2006 *et al.*, 2009). L'observation de la même région dans des embryons de 13,5jpc par Sado, Ohhata *et al.* a aussi conduit à l'observation d'une hypométhylation de l'ADN, de l'enrichissement en H3K4me2 et H4Ac et de la diminution en H3K27me3 et en H3K9me2, ces dernières marques étant normalement caractéristiques du Xi dans les cellules femelles à ce stade cellulaire (Sado *et al.*, 2004 *et al.*, 2005). En définitive, nous pouvons donc conclure que *Tsix* semble recruter une série de marques hétérochromatiques au promoteur de *Xist* dans les cellules ES non-différenciées, en accord avec son rôle de répresseur (Figure 24). Ce recrutement pourrait très bien être semblable à ce qui est observé chez la levure *S. cerevisiae*, où l'on observe un lien direct entre méthylation de H3K4 et élongation transcriptionnelle (Bernstein *et al.*, 2002, Krogan *et al.*, 2003, Ng *et al.*, 2003), l'enzyme responsable du dépôt de cette marque (Set1) interagissant directement avec l'ARN polymérase II en élongation.

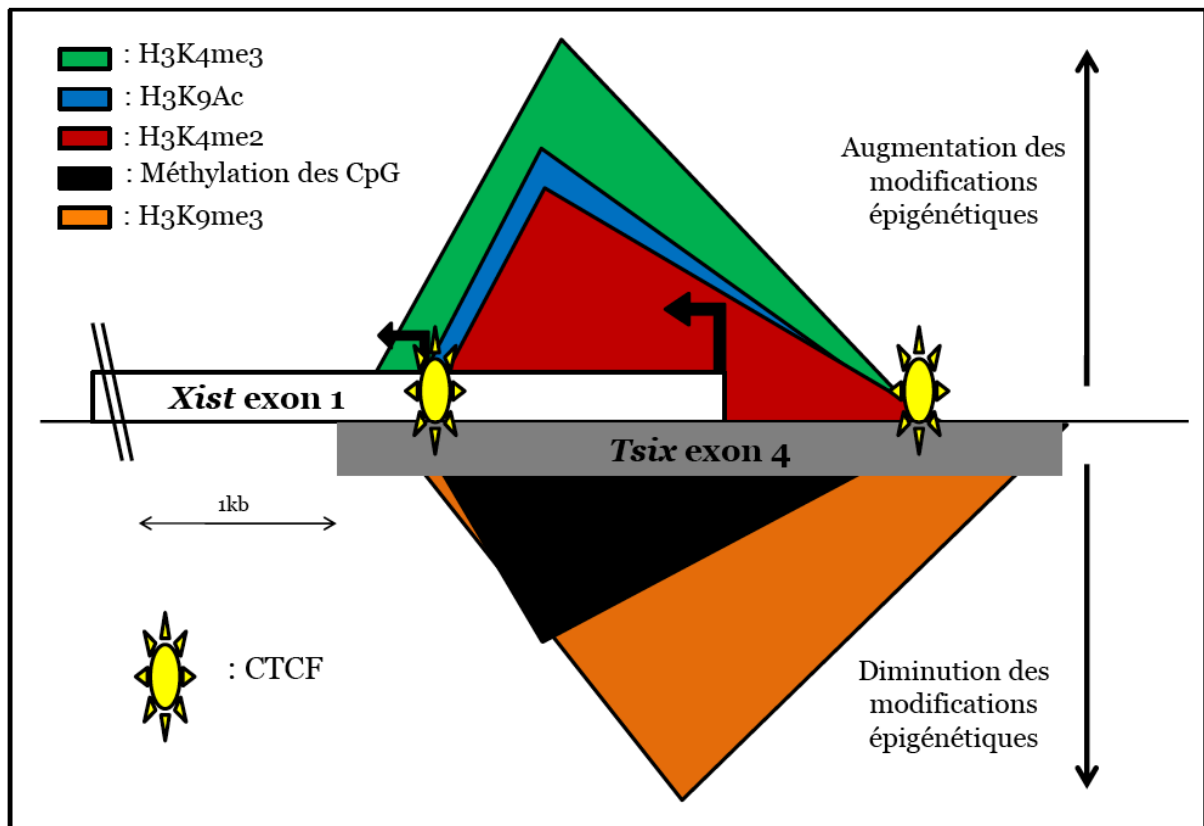


Figure 24 : Remodelage chromatinien de la région promotrice de *Xist*, comprise entre deux sites de fixation de CTCF, par *Tsix*. Le graphique montre le ratio entre Ma2L (cellules ES mâles tronquées pour *Tsix*) et Ma1L (sauvages) obtenu pour différentes modifications épigénétiques. La position des sites de fixation de CTCF sont indiqués par des soleils jaunes. CTCF est présente de part et d'autre du promoteur de *Xist* aussi bien dans les cellules Ma1L qu'au sein des Ma2L au stade non-différencié. Les triangles colorés représentent la modifications des domaines de diverses modifications épigénétiques de la chromatine en présence ou non d'une transcription abortée de l'antisens. (basée sur les résultats de Navarro *et al.*, 2006)

Parallèlement à cette euchromatinisation de la région promotrice, on observe aussi une expression ectopique de *Xist* à partir du chromosome X muté dans ces cellules ES mâles, des domaines de *Xist* étant même observables dus à un recrutement du PIC au promoteur P1 en cours de différenciation (Vigneau *et al.*, 2006) (bien que dans les cellules somatiques, Sado *et al.* confirment les données de Johnston *et al.*, 1998, quant à la majorité des transcrits provenant de P2). À l'inverse, le Xa est normalement caractérisé, dans la région promotrice de *Xist*, par la méthylation de l'ADN et l'enrichissement en H3K27me3, ainsi que l'absence des marques euchromatiques H3K4me2, H3K4me3, H3K9ac et H4ac (McDonald *et al.*, 1998; Navarro *et al.*, 2005, 2006, 2009; Sado *et al.*, 2005; Sun *et al.*, 2006). De manière fort intéressante, Navarro *et al.* observent que l'ensemble de ces modifications chromatinienne, dues à une absence de la transcription de l'antisens en *cis*, se déroulent toutes au sein d'une

région délimitée par deux sites de fixation de la protéine CTCF. Cette protéine est connue pour jouer un rôle d'insulateur afin de définir des domaines chromatinien bien distincts des régions alentour ou encore de bloqueur d'amplificateur (« *enhancer* »), et est notamment impliquée dans l'établissement de domaines géniques échappant à l'inactivation sur le Xi (Filippova *et al.*, 2005). De manière intéressante, CTCF est fixée à environ -1kb et +1,5kb du promoteur P1 aussi bien dans les cellules ES non-différenciées mâles que femelles, et reste présente même dans les cellules mâles au sein desquelles la transcription de *Tsix* est abolie (Navarro *et al.*, 2006). Toutefois, avec ces données, nous ne pouvons conclure sur le rôle de CTCF dans la délimitation de cette région hétérochromatique, étant donné que ce phénomène pourrait aussi être dû au fait qu'elle correspond aussi à l'exon 4 de *Tsix*, autrement dit la seule partie exonique du locus *Xist/Tsix* qui se superpose. Il serait tout de même intéressant de savoir si CTCF reste fixée dans ces cellules en cours de différenciation.

Dans les cellules ES mâles, on peut apercevoir que la région promotrice de *Xist* est aussi bien marquée par la marque euchromatique H3K4me2 que par la marque hétérochromatique H3K9me3 (Navarro *et al.*, 2006). Cette configuration de marques épigénétiques contradictoires au niveau d'un promoteur différentiellement régulé au cours du développement, constitue l'exemple type d'un domaine bivalent tel qu'il a été défini par Azuara, Bernstein *et al.* (Bernstein *et al.*, 2006 ; Azuara *et al.*, 2006 ; Stock *et al.*, 2007). L'existence même d'un tel domaine, correspond tout à fait à l'idée qu'il est nécessaire au promoteur de *Xist* de pouvoir rapidement basculer aussi bien vers une conformation euchromatique qu'hétérochromatique en cours de différenciation en fonction du choix effectué par la cellule quant à l'allèle du chromosome X à inactiver. Dans cette ligne d'idées, l'ADN semble d'ailleurs globalement hypométhylé au niveau de la région promotrice de *Xist* dans les embryons mâles et femelles jusqu'au stade blastocyste inclus (McDonald *et al.*, 1998), ainsi que dans les cellules ES femelles indifférenciées (Norris *et al.*, 1994; Sado *et al.*, 1996; Sun *et al.*, 2006). Mais il semble être tout du moins partiellement méthylé dans les cellules ES mâles, en tout cas pour un petit nombre de positions testées (Norris *et al.*, 1994; Sun *et al.*, 2006). Cette méthylation pourrait dépendre de l'association des ARN *Tsix* avec le méthylase Dnmt3a (Sun *et al.*, 2006, Nesterova *et al.*, 2008).

Une autre étude (Sun *et al.*, 2006) arrive cependant à des conclusions partiellement différentes

de celles de Navarro, Sado, Ohhata, *et al.* concernant la régulation par *Tsix* de la chromatine au promoteur de *Xist*. Dans leur article, les auteurs indiquent n'observer aucune variation de l'état chromatique de la région promotrice de *Xist* concernant les marques H3K4me2 et H4Ac (et une faible augmentation de H3K27me3), dans leurs cellules ES femelles non-différenciées monoalléliquement mutées au niveau du promoteur de *Tsix*. D'autre part, dans des cellules ES mâles $X^{\Delta CpG}Y$, ils ne mesurent pas de diminution de la méthylation de l'ADN à proximité du promoteur de *Xist*, aussi bien lorsque les cellules sont indifférenciées qu'en cours de différenciation. Dans les cellules ES femelles indifférenciées $X^{\Delta CpG}X$, la méthylation de l'ADN au promoteur de *Xist* ne semble pas non plus être altérée sur le chromosome porteur de la délétion, mais après différenciation, elle est observée, de manière attendue, presque exclusivement sur le chromosome X sauvage, qui dans ces lignées n'est jamais inactivé.

Cependant la divergence la plus importante repose sur l'interprétation des données obtenues par le laboratoire de Jeannie Lee. Sun *et al.* proposent que lors de l'initiation de l'inactivation, sur le futur Xi, un état hétérochromatique transitoire s'établit dans la région promotrice de *Xist* permettant sa sur-expression. Cette hypothèse se base sur le fait qu'ils observent une très légère augmentation de la marque H3K27me3 dans les cellules mutées pour l'antisens par rapport aux cellules sauvages, cette augmentation étant d'ailleurs observée de manière beaucoup plus flagrante dans des études menées par d'autres laboratoires (Navarro *et al.*, 2006, 2009 ; Shibata *et al.*, 2008, Shibata & Yokota, 2008). Ils supposent donc que dans un contexte sauvage, la marque H3K27me3 est déposée dans la région promotrice de *Xist* sur le futur Xi parallèlement à l'extinction transcriptionnelle de *Tsix* lors de l'initiation de l'inactivation, ce qui n'a pas été observé par d'autres études beaucoup plus poussées (Navarro *et al.*, 2009, Marks *et al.*, 2009, voir le chapitre suivant). Ils supposent aussi, sur la base de leurs données ne corroborant pas celles déjà publiées (modifications de H3K4me2 et H4Ac), que paradoxalement dans les cellules ES indifférenciées, l'expression de *Tsix* induit une structure euchromatique sur le promoteur de *Xist*, qui serait responsable de sa répression.

En définitive, le laboratoire de Jeannie Lee avance donc l'hypothèse surprenante qu'un état euchromatique empêche la transcription de *Xist* à l'état indifférencié ainsi que lors de l'initiation de la différenciation sur le futur Xa, tandis qu'un état hétérochromatique transitoire (le profil dans les cellules somatiques étant similaires à ceux des autres articles) permet la transcription de *Xist* à partir du futur Xi. Dernièrement ce laboratoire a tenté de corroborer ce

postulat avec l'observation du recrutement du complexe protéique PRC2 (responsable du dépôt de la marque H3K27me3) par un petit transcrit émanant du *Repeat-A* nommé *RepA* et qui serait responsable de ces observations ainsi que de la propagation de cette marque le long du Xi (Zhao *et al.*, 2008) (voir aussi le chapitre : « Modifications d'histones sur le Xi », p36). De manière surprenante, le recrutement de PRC2 par *RepA* n'engendrerait le dépôt de H3K27me3 qu'en cours de différenciation alors que l'interaction entre transcrit et complexe protéique est déjà observable dans les cellules ES non-différenciées (et que cette marque semble déjà présente à jour 0 dans l'article de Sun *et al.*). De plus, l'idée que le *Repeat-A* serve de centre de nucléation de cette marque au promoteur de *Xist* sur le futur Xi semble quelque peu "tirée par les cheveux", sachant qu'une seule position dans la région promotrice de *Xist* a été analysée. De plus H3K27me3 semble se propager à partir du « *hotspot* » dans les cellules mutées par *Tsix*, indiquant plutôt un rôle de l'antisens dans la délimitation de la région marquée par H3K27me3 (Navarro *et al.*, 2006 *et al.*, 2009), ce qui a récemment été confirmé (Navarro *et al.*, 2009, Marks *et al.*, 2009) (voir le chapitre suivant). Par contre, ceci n'empêche pas la possibilité du complexe *RepA*-PRC2 d'agir sur d'autres régions telles que le *hotspot* de méthylation.

Les différences entre les profils présentés par Sun *et al.* avec ceux de Navarro *et al.* appellent une remarque, concernant la résolution des deux études. Ainsi, Sun *et al.* n'ont testé qu'une seule position dans la région promotrice de *Xist*, contrairement à Navarro *et al.*, qui ont réalisé pour chacune des modifications un profil incluant de nombreuses positions et permettant de définir précisément les limites, de part et d'autre des promoteurs P1 et P2, des modifications qu'ils observent. Précisément, les observations de Sun *et al.* ne sont pas en désaccord avec les mesures réalisées par Navarro *et al.* aux mêmes positions, qui sont à la marge de la région euchromatinisée en l'absence de *Tsix* (Navarro *et al.*, 2006).

Pour finir, l'absence de *Tsix*, si elle ne conduit pas à l'activation transcriptionnelle de *Xist* avant différenciation (malgré une augmentation du nombre de transcrits lorsque l'antisens est muté), n'en a pas moins pour conséquence une modification de la chromatine au promoteur de *Xist*, dont il a été proposé qu'elle « prédestinait » l'allèle ainsi ciblé à être surexprimé en cours de différenciation (Navarro *et al.*, 2006), le domaine bivalent basculant soit vers un domaine euchromatique (dans le cas du Xi) soit vers un domaine hétérochromatique (Xa).

D'autre part, l'absence d'activation transcriptionnelle de *Xist* malgré l'euchromatinisation de sa région promotrice, dans les cellules ES indifférenciées mutées pour *Tsix*, suggère l'existence d'un répresseur transcriptionnel de *Xist*, perdu en cours de différenciation (voir le chapitre : « Couplage entre la régulation de l'expression de *Xist* et la pluripotence », p88). On peut également imaginer que, bien que le promoteur de *Xist* soit euchromatinisé, la chromatine associée au reste du locus *Xist* ne soit pas compétente pour permettre l'expression de ce gène à un fort niveau. Des remaniements de nature épigénétique devraient alors avoir lieu en cours de différenciation pour permettre cette expression. Comme nous le verrons à la section suivante, *Tsix* joue aussi un rôle important dans la détermination du profil épigénétique de l'ensemble du locus *Xist/Tsix*.

iv. Modification du paysage épigénétique le long du locus *Xist/Tsix* par *Tsix*

En plus de modifier l'état chromatinien au promoteur de *Xist*, il semblerait que *Tsix* soit aussi impliqué dans le remodelage de la chromatine tout le long du locus *Xist/Tsix* (Figure 25).

Au sein des cellules ES sauvages, le corps du gène *Xist* est associé aux marques H3K4me2 et H3K9me3, ainsi qu'à l'absence des marques H3K4me3, H3K9Ac et H3K27me3 (Navarro *et al.*, 2005 *et al.*, 2006 *et al.*, 2009, Sun *et al.*, 2006, Shibata *et al.*, 2008). Par contre, lorsque *Tsix* est tronqué ou absent, la marque H3K4me2 est pratiquement perdue le long du locus (Morey *et al.*, 2004, Navarro *et al.*, 2005 *et al.*, 2006 *et al.*, 2009, Sun *et al.*, 2006, Shibata *et al.*, 2008). L'absence de *Tsix* conduit aussi à l'apparition de la marque H3K27me3 dans le corps du gène, comme vu au chapitre précédent, mais n'affecte pas les autres marques citées précédemment, à part l'acétylation de l'histone H4 qui n'a été décrite qu'une fois (dans des cellules mutantes pour *Tsix*) (Sun *et al.*, 2006).

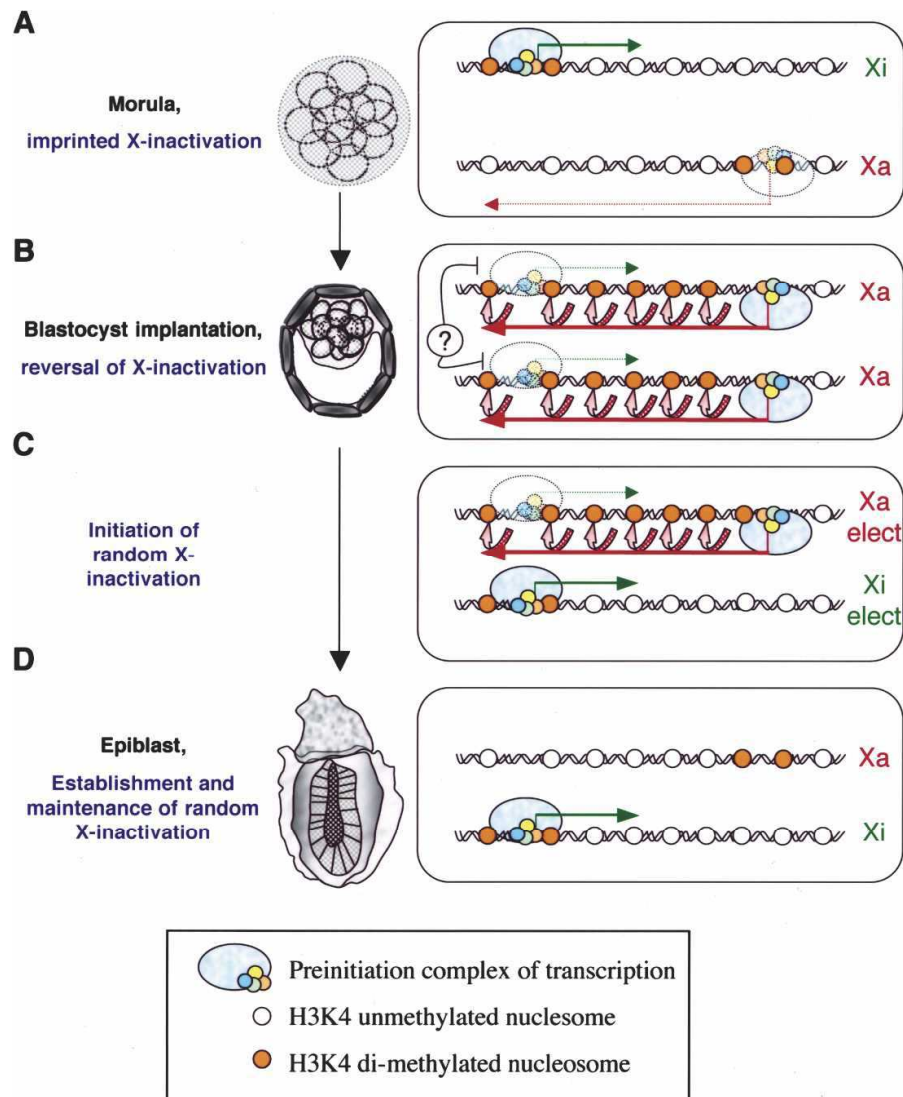


Figure 25 : Modèle de la régulation de la région *Xist/Tsix* au cours du développement embryonnaire précoce murin. Les étapes du développement sont indiquées sur la gauche et le statut de la région *Xist/Tsix* en terme de de transcription et de chromatine à chaque étape est indiqué sur la droite. On observe ici, l'effet de la transcription de *Tsix* à travers le gène *Xist*, qui sert notamment à effacer les marques hétérochromatiques engendrées suite à l'inactivation empreintée du Xp. De cette manière, les deux chromosomes possèdent, en principe, une probabilité équivalente d'être inactivé lors de l'inactivation aléatoire. (D'après Navarro *et al.*, 2005)

Ces modifications épigénétiques lorsque l'antisens n'est plus transcrit indiquent que *Tsix* aurait donc un rôle dans l'euchromatinisation du corps du gène *Xist*, un rôle contraire à ce qui avait été conclu au niveau de la région promotrice de celui-ci. La manière dont ces deux fonctions distinctes de *Tsix* sont isolées l'une de l'autre reste encore à découvrir. Mais à ce titre, il est important de rappeler que la région hétérochromatinisée par *Tsix* (au niveau du promoteur de *Xist*) est délimitée par deux sites de fixation pour CTCF, qui pourraient, ici, jouer le rôle de barrière quant à l'un ou l'autre effet de la transcription de l'antisens. De plus, on ne peut

s'empêcher d'observer que cette région correspond à peu près à l'exon 4 de *Tsix*, ceci laissant entrevoir la possibilité d'une fonction du transcrit *Tsix* ou d'un duplexe *Xist:Tsix* (cette région étant la seule à pouvoir le faire) dans l'établissement d'un état chromatinien particulier.

Il a été proposé que la déposition de H3K4me2 dans le corps de *Xist* des cellules ES non-différenciées par la transcription de *Tsix* pourrait contribuer à effacer les marques d'empreinte ou le reste des marques caractéristiques du Xp inactivé de façon empreintée, et ce, de manière à rendre les deux allèles de *Xist* épigénétiquement équivalents, au moment de la réactivation du Xp dans la masse cellulaire interne des blastocystes (Navarro *et al.*, 2005). Ainsi, la cellule serait apte à choisir quel chromosome X inactiver de manière stochastique lors de l'initiation de l'inactivation aléatoire.

Comme nous l'avons évoqué lors du chapitre précédent, la délétion de *Tsix* ainsi que l'empêchement de sa transcription à passer par la région promotrice de *Xist* semble engendrer un recrutement de la marque H3K27me3 le long de celle-ci (Navarro *et al.*, 2006 *et al.*, 2009, Sun *et al.*, 2006, Shibata *et al.*, 2008). L'hypothèse que cette marque devrait donc être retrouvée sur le futur Xi en cours de différenciation est donc compréhensible (Sun *et al.*, 2006). Cependant dans les mutants *Tsix*, la marque H3K27me3 disparaît totalement du *hotspot* (ainsi que de là où elle s'est propagée chez les mutants) au fur et à mesure de la progression de la différenciation (Navarro *et al.*, 2009, Marks *et al.*, 2009), nous indiquant donc qu'il y a peu de chances qu'elle serve à la surexpression de *Xist*, d'autant plus que dans une lignée mutante pour *Tsix* ainsi que pour *Eed* (protéine faisant partie du complexe PRC2) cette sur-expression se fait mieux en l'absence de *Tsix* et de H3K27me3 que juste en l'absence de l'antisens (Shibata *et al.*, 2008).

De manière surprenante, le dépôt de cette marque dans des cellules ES non-différenciées en l'absence de *Tsix* possède un motif de fixation (« *pattern* ») particulier. La marque semble se propager à partir du *hotspot* de méthylation (voir le chapitre : « "Point chaud" de modifications épigénétiques en 5' de *Xist* », p50) jusqu'à l'arrêt transcriptionnel de *Tsix* où qu'il soit (Navarro *et al.*, 2006 *et al.*, 2009), le *hotspot* démarrant justement à la fin de la région transcriptionnelle de *Tsix* dans les cellules sauvages (Sado *et al.*, 2001, Heard *et al.*, 2001, Rougeulle *et al.*, 2004). Il serait intéressant de savoir jusqu'où le *hotspot* a la capacité de s'étendre dans des cellules mutantes, la mutation la plus distante du *hotspot* utilisée dans

l'étude de Navarro *et al.* (Δ DxPas34) ne comprenant pas le promoteur minoritaire de *Tsix*, d'où l'arrêt de la propagation autour du microsatellite. Théoriquement, quelque soit la mutation, il devrait s'arrêter à la première fin de transcription rencontrée (si ce pouvoir bloquant n'est pas restreint à *Tsix*).

Navarro *et al.* ont ensuite établi que le blocage du *hotspot* de méthylation est dû à l'accumulation d'ARN polymérase II engendrée par la fin de la transcription de l'antisens, de manière similaire avec ce qui est connu concernant l'ARN polymérase III (Simms *et al.*, 2004, Noma *et al.*, 2006). Plusieurs hypothèses existent pour expliquer ce mécanisme (déplacement de nucléosomes, recrutement d'autres facteurs), par contre il est sûr que la fixation de la protéine CTCF ne semble jouer aucun rôle dans la propagation ou le blocage du *hotspot* en 5' de *Xist*, sa fixation étant ubiquitaire à cette position dans les cellules ES non-différenciées mutées pour *Tsix* (chez qui le hotspot se prolonge jusqu'à l'arrêt transcriptionnel de ce dernier) (Navarro *et al.*, 2006). La fonction de *Tsix* en tant qu'obstacle de la marque H3K27me3 protégeant *Xist* semble rejoindre l'idée que les barrières à la propagation hétérochromatinienne pourraient fonctionner en exploitant la fonction d'autres mécanismes de régulation génique, ne nécessitant pas d'appareillage spécifique (Gaszner & Felsenfeld, 2006). Un effet distal des mutants de *Tsix* a pu être observé de manière similaire à la transcription antisense de *Air* et *Kcnq1ot1* qui contrôlent l'expression empreintée en réprimant des gènes que le transcrit traverse ainsi que des gènes situés à de plus grandes distances et ce, en induisant le recrutement de marques hétérochromatiques telles que H3K27me3 et H3K9me2 (Ideraabduallah *et al.*, 2008, Pandey *et al.*, 2008, Terranova *et al.*, 2008). Dans leur article, Navarro *et al.* démontrent une implication de *Tsix* dans l'intensité du marquage en H3K27me3 au sein du hotspot, ainsi que la nécessité de sa transcription complète pour une expression appropriée des gènes *Cnbp2* et *Xpct* (ceux-ci étant partiellement réprimés en corrélation avec une augmentation de H3K27me3 sur l'ensemble du hotspot dans les cellules mutées). L'intensité du *hotspot* augmente effectivement en présence de mutants de l'antisens, les auteurs expliquant cela par la possibilité qu'en absence de *Tsix*, les transcrits *Xist* puissent être fixés par PRC2 (faisant référence à l'article par Zhao *et al.*, 2008) et que ces complexes pourraient être enrichis au niveau du *hotspot* qui seraient le centre de nucléation de l'inactivation du chromosome X. Cependant en cours de différenciation le *hotspot* disparaît dans tous les types cellulaires observés, cet effet serait donc restreint aux cellules ES non-

différenciées. Une autre possibilité serait que ces interactions à longue distance soient médiées par une conformation particulière de la chromatine, surtout au vu des rôles de *Xpct* (*Xpr*) et *Tsix* dans la transvection des deux chromosomes X lors de l'initiation de l'inactivation.

De manière intéressante, il faut finalement noter que la modification des niveaux de H3K27me3 n'a pas affecté la méthylation de H3K4 ou de H3K9 (que ce soit dans le *hotspot* ou au promoteur de *Xist*), indiquant ainsi l'indépendance de la régulation de H3K27me3 vis-à-vis de H3K9me2 (déjà démontré par Rougeulle *et al.*, 2004) et *vice-versa*.

v. Effet de *Xist* sur *Tsix*

Jusqu'à présent, dans cette partie de l'introduction dédiée à la régulation de l'expression de *Xist*, nous avons surtout évoqué les mécanismes mis en place par *Tsix* pour le réprimer. Cependant, il y a peu de temps, le laboratoire de Sado a découvert une mutation de *Xist* qui influence l'expression de *Tsix* (Hoki *et al.*, 2009), une première dans l'histoire de l'inactivation.

La question de savoir si la transcription de *Xist* influence celle de son anti-sens ou bien de savoir si certaines parties du transcrit *Xist* peuvent induire un changement quelconque au niveau de l'expression de *Tsix* n'avait jamais été adressée. Hoki *et al.* ont cherché à obtenir des mutants du *Repeat-A* afin d'étudier son rôle au cours de l'inactivation. Et ils ont bien observé que leur mutant n'arrive pas à éteindre transcriptionnellement son chromosome X au cours de l'inactivation, ce qui était attendu en tenant compte des découvertes de Wutz *et al.* (Wutz *et al.*, 2002). Par contre, curieusement, ils ont aussi découvert que ce phénotype n'est pas dû à une impossibilité de l'ARN à ramener les régions géniques vers le centre hétérochromatique du domaine (Chaumeil *et al.*, 2006) mais à la faible quantité de transcrits *Xist* dans la cellule (surement due à une répression transcriptionnelle, la stabilité des ARNs étant similaire à celle des cellules sauvages). Parallèlement à cela, les auteurs ont aussi observé une expression ectopique de l'antisens provenant du Xp dans ces cellules soumises à l'inactivation empreintée, qui pourrait expliquer l'état réprimé de *Xist*, notamment grâce à une hyperméthylation de l'ilôt CpG au sein de la région promotrice de *Xist* dans les cellules mutées. La raison pour laquelle ces phénotypes n'avaient pas été observés par Wutz *et al.* vient du fait que les systèmes utilisés pour étudier l'effet du *Repeat-A* diffèrent grandement, Wutz utilisant un système transgénique inductible du cDNA, ne contenant donc ni la région

promotrice de *Xist* ni l'antisens. Cela nous laisse supposer que le *Repeat-A* par un processus encore inconnu arrive à réprimer *Tsix*, l'effet de cette répression n'étant visible que lorsque *Xist* est censé être surexprimé. La possibilité existe que *Rep-A* et PRC2 pourraient être impliqués, l'hypothèse de Jeannie Lee que ce petit transcrit joue un rôle dans la surexpression de *Xist* paraît, dans ce contexte, très attrayante. De plus sans savoir si la délétion du *Repeat-A* influence *Xist* ou *Tsix* qui s'influencent réciproquement, on peut imaginer la possibilité que ces interactions se fassent à travers des structures tri-dimensionnelles le long du locus *Xist/Tsix* au vu des distances entre les deux régions promotrices. Cependant cette étude menée par Hoki contraste avec celles de Zhao et Wutz de par le fait qu'elle s'intéresse principalement à l'inactivation empreintée, il sera donc fort intéressant de savoir si les mêmes phénotypes concernant l'expression de *Xist* et *Tsix* sont retrouvés lors de l'inactivation aléatoire, d'autant plus que l'expression de *RepA* n'a jamais été cherchée dans les embryons pré-implantatoires.

b. Couplage entre la régulation de l'expression de *Xist* et la pluripotence

Nous avons vu que *Tsix* réprime *Xist* post-transcriptionnellement de par le fait que lors de la délétion de l'antisens, aucun recrutement de la machinerie transcriptionnelle n'est observé au promoteur majoritaire de *Xist* (Navarro *et al.*, 2005). De plus, bien que *Xist* soit surexprimé dans ces cellules à un stade non-différencié, par un mécanisme encore inconnu, il ne l'est tout de même pas assez pour induire la formation d'un domaine autour du chromosome X et donc d'induire son extinction transcriptionnelle. Or, lors de l'induction de la différenciation cellulaire dans des cellules mâles au sein desquelles la transcription de *Tsix* est bloquée, on observe une expression ectopique de *Xist* suffisamment grande pour effectuer une inactivation chromosomique, indiquant la présence d'un facteur répressif (autre que *Tsix*) dans les cellules ES qui disparaît en cours de différenciation et permet ainsi le recrutement de la machinerie de transcription à son promoteur (Navarro *et al.*, 2005, 2006).

En 2008, la découverte par Navarro *et al.* de ce chaînon manquant à la compréhension des mécanismes répressifs de *Xist* a permis de relier l'initiation de l'inactivation à la pluripotence cellulaire (Navarro *et al.*, 2008). Car de manière somme toute assez logique, ce sont les facteurs de pluripotence Oct4, Nanog et Sox2 qui, en se fixant dans l'intron 1 du gène *Xist*

(préalablement repéré comme contenant un site d'hypersensibilité à la méthylation présent uniquement à l'état indifférencié, Tsai *et al.*, 2008), inhibent le recrutement du PIC à P1.

À l'origine, les auteurs cherchaient un lien entre inactivation et pluripotence, les deux phénomènes étant de toute évidence étroitement liés et *Nanog* étant un candidat idéal de par l'anti-correlation parfaite de son expression vis-à-vis de celle de *Xist* dans tous les types cellulaires, notamment dans l'ICM où *Nanog* se trouve exprimé alors qu'on assiste à la réactivation du Xp suite à l'inactivation empreintée (Mak *et al.*, 2004). En réalisant des expériences de CHIP le long du locus *Xist/Tsix* ils ont trouvé un site de fixation de *Nanog* dans les cellules ES, aussi bien mâles que femelles. Afin de valider le rôle de cette protéine dans la régulation transcriptionnelle de *Xist*, ils ont cherché à regarder comment cette transcription évolue dans des cellules ES délétées pour *Nanog* (3 mutants différents). Il se trouve que dans ces lignées cellulaires, l'expression de *Xist* n'est surexprimée que d'un facteur 4 à 6 par rapport aux lignées contrôles, ceci n'impliquant pas de répression de *Tsix*, ni de différenciation cellulaire étant donné que *Nanog* n'est pas indispensable pour induire ou maintenir la pluripotence cellulaire (Chambers *et al.*, 2007). Dans ces cellules mutantes pour *Nanog*, ainsi que dans les cellules ES différenciées ou bien les fibroblastes, la fixation de *Nanog* n'est plus observée à l'intron 1 de *Xist*. Cependant l'enrichissement plutôt faible de *Xist* par rapport à des mutants de *Tsix* où il est d'environ un facteur 100, a conduit les auteurs à chercher des facteurs additionnels participant à la répression transcriptionnelle de *Xist* en partenariat avec *Nanog*. *Oct4* et *Sox2* se présentent comme des candidats idéaux, au vu de leur rôle dans la maintenance de la pluripotence cellulaire ainsi que le partage de leurs sites de fixation avec *Nanog* à plusieurs sites à travers le génome (Boyer *et al.*, 2005, Loh *et al.*, 2006). La cinétique de fixation de ces protéines est la même que celle de *Nanog* au niveau de l'intron 1 : on ne les retrouve que dans les cellules ES, ces deux protéines restant même attachées dans les cellules mutantes pour *Nanog*. Ce qui nous indique que d'une part *Nanog* n'est pas responsable de ce recrutement et explique d'autre part le faible taux de surexpression de *Xist* observé dans ces cellules. Afin de visualiser quel effet ces nouvelles protéines ont sur l'expression de *Xist*, ces chercheurs se sont servis de cellules ES mâles délétées pour *Oct4* mais ayant reçu un transgène de cette même protéine dont la répression est inductible par la tétracycline. L'ajout de tétracycline au milieu de culture de ces cellules engendre en plus d'une absence d'*Oct4*, une différenciation logique de celles-ci en réponse à laquelle la quantité cellulaire de *Tsix*, *Nanog* et

Sox2 commence à diminuer à partir de 48h. De manière contrastante, *Xist* est systématiquement surexprimé dès 24h de traitement, son niveau d'expression étant comparable avec celui de cellules ES femelles en différenciation. En comparaison, lors de la différenciation de ces cellules induite par ajout d'acide rétinoïque en absence de tétracycline, au même stade (après 24h), Oct4, Nanog et Sox2 sont tous les 3 encore fixés à *Xist*, indiquant que ces 3 facteurs agissent en synergie afin de réprimer *Xist*.

De manière surprenante, seules environ 10% des cellules délétées pour Oct4 acquièrent une accumulation ectopique de *Xist* après 96h de traitement à la tétracycline. Les auteurs expliquent que ceci pourrait être dû au fait que ces 10% de cellules correspondent à des cas où l'expression de *Tsix* aurait été prématurément réprimée (bien que 92% et 36% des cellules *Xist*-positives contiennent encore une expression de *Tsix* à 24 et 96h de traitement, respectivement). Ou alors que cette hétérogénéité de domaines de *Xist* pourrait être due au mosaïcisme de Nanog qui est absent dans environ 20% des cellules ES non-différenciées à tout instant (Chambers *et al.*, 2007) et *Xist* étant partiellement dérprimé dans de telles cellules, ce qui laisserait supposer que celles subissant une forte surexpression de *Xist* seraient aussi celles chez qui Nanog est momentanément réprimé.

L'ensemble de ces données démontre bien l'implication de ces trois facteurs de pluripotence dans la répression transcriptionnelle directe de *Xist* dans les cellules ES, liant par la même occasion l'inactivation du chromosome X au développement embryonnaire. De plus, au vu du rôle de ces trois facteurs dans la régulation de *Xist*, les auteurs proposent qu'il y a de fortes chances pour qu'ils soient aussi impliqués dans la réactivation du Xp lors du stade péri-implantatoire de l'embryon, le même phénomène pouvant se produire lors de la réactivation du Xi dans les PGCs suite à la réexpression de *Nanog* (de Napoles *et al.*, 2007, Sugimoto *et al.*, 2007).

Les mécanismes exacts de la mise en place de cette répression sont encore vagues. Cependant d'après le fait que le Xi ne se réactive qu'une fois *Nanog* réexprimé dans les PGCs alors même qu'*Oct4* et *Sox2* le sont déjà, et qu'il en est de même lors de la fabrication de cellules iPS (Stadtfield *et al.*, 2008) ou de la fusion cellulaire (Do *et al.*, 2008), on peut supposer que Nanog est responsable de l'initiation de cette répression, même si au vu de la faible surexpression de *Xist* dans les cellules *Nanog*^{-/-} on est en droit de supposer qu'il ne joue pas un rôle prépondérant dans la maintenance de la répression. Par contre *Oct4*, dont la mutation induit un

effet beaucoup plus marqué sur la surexpression de *Xist*, semblerait être davantage lié au maintien de cette répression, voire même au recrutement ou au maintien de la fixation des deux autres protéines, leur fixation à l'intron chutant plus rapidement que leur ARNm lors de la répression d'*Oct4* par ajout de tétracycline dans la lignée mutante. L'étude du rôle de Sox2 dans ce système pourra aussi être intéressant, notamment pour tenter d'établir une chronologie de fixation, en plus d'observer son importance au sein de ce mécanisme régulateur, d'autant plus que sa déplétion semble engendrer une surexpression de *Tsix* et de *Xite* dans des cellules femelles à quatre jours de différenciation (Donohoe *et al.*, 2009). Navarro *et al.* constatent qu'il semblerait donc que Nanog soit indispensable soit à la fixation soit à l'activité répressive d'Oct4 et Sox2 dans la régulation de *Xist*, mais qu'après cette étape ces dernières fonctionnent de manière indépendante de Nanog. Pour expliquer cela, ils proposent un modèle où un remodelage chromatinien de l'intron 1 de *Xist* serait orchestré par un mécanisme Nanog-dépendant afin de permettre l'accès d'Oct4 et Sox2 à cette région. Ou encore que Nanog pourrait être responsable de l'effacement d'une empreinte parentale empêchant la fixation de Oct4 et Sox2 sur le Xp lors de l'activation de son expression dans l'épiblaste et de la réactivation du Xp (Navarro & Avner, 2010), ce qui expliquerait la répression de *Xist* sur l'Xm au cours des stades pré-implantatoires en l'absence de Nanog.

Il reste toutefois quelques points d'ombre dans le processus de répression transcriptionnelle de *Xist*, notamment le fait qu'au sein de cellules pluripotentes femelles dérivées à partir de l'épiblaste murin, bien que les trois facteurs de pluripotence soient exprimés, on observe une surexpression de *Xist* qui se traduit par l'inactivation aléatoire d'un des deux chromosomes X (Guo *et al.*, 2009). Ceci implique donc que des niveaux supplémentaires de complexité soient impliqués dans la mise en place de l'inactivation aléatoire au cours d'une période précédant la répression des facteurs de pluripotence, tel que le blocage de leur fixation, l'inhibition de leur fonction, ou encore le maintien de leur fixation uniquement sur le futur Xa. L'étude de ce processus au cours d'une cinétique temporelle très précise ainsi que dans différents tissus (par exemple au cours de la maturation des PGCs) pourra peut-être répondre à certaines questions non-élucidées concernant ce processus captivant.

c. Régulation en *trans* de *Xist* par divers facteurs

Jusqu'à présent nous nous sommes surtout attelés à décrire les effets répressifs de l'antisens et des facteurs de pluripotence sur l'expression de *Xist* qui sont plutôt bien décrits, et leur implication dans le processus de l'inactivation du chromosome X étant sans appel. Cependant, au fil des années plusieurs autres facteurs protéiques ont été décrits comme influençant la régulation transcriptionnelle de *Xist* mais souvent soit de manière controversée, soit de manière pas véritablement approfondie. Au cours du chapitre suivant nous allons donc nous pencher sur d'autres régulateurs, moins évidents, de *Xist*, commençant tout d'abord par des répresseurs avant d'étudier les activateurs potentiels.

i. Les répresseurs

L'implication des méthyl-transférases à ADN dans la régulation de *Xist* est sujette à controverse malgré le nombre d'études sur le sujet. En 2004, l'équipe d'En Li a étudié l'effet de la mutation *Dnmt3a*^{-/-} ainsi que *Dnmt3b*^{-/-} sur l'expression de *Xist*, mais n'a trouvé que peu d'effets de cette délétion sur notre gène d'intérêt (une réactivation partielle très progressive) (Sado *et al.*, 2004, voir le chapitre : « Méthylation de l'ADN », p40). Cette observation est plutôt paradoxale si l'on tient compte de l'effet prononcé d'une mutation de la méthyl-transférase de maintien Dnmt1 sur l'expression de *Xist*. En son absence, la région promotrice de *Xist* étant déméthylée, la cellule n'arrive donc pas à maintenir *Xist* réprimé sur le Xa, engendrant des domaines observables chez les mâles et des cellules à deux domaines de *Xist* chez les femelles post-différenciation, ainsi qu'une dérégulation partielle du Xi dans les cellules somatiques (Beard *et al.*, 1995 ; Panning *et al.*, 1997 ; Sado *et al.*, 2000 ; Csankovszki *et al.*, 2001). Comment le maintien de la méthylation de l'ADN au promoteur de *Xist* joue un rôle essentiel dans sa régulation alors que la méthylation *de novo* n'y jouerait semble-t-il aucun demeure un mystère. Il est possible que l'importance de ces protéines se trouve à un autre niveau que celui de la méthylation de l'ADN, et qu'elles serviraient à recruter d'autres protéines. Dnmt1 et Dnmt3a peuvent par exemple directement recruter des HDACs (HDAC1 mais aussi HDAC2 pour Dnmt1) à leurs sites de fixation (Fuks *et al.*, 2000 *et al.*, 2001, Rountree *et al.*, 2000 ; Robertson *et al.*, 2000), chaque protéine ayant ainsi ses partenaires propres et pouvant donc jouer des rôles dans des processus différents les uns des autres.

Cependant, plusieurs autres articles démontrent une influence subtile de la protéine Dnmt3a dans la régulation de *Xist* : - Nesterova *et al.* imputent l'observation d'une augmentation du nombre des transcrits *Xist* et la présence peu fréquente de domaines ectopiques dans des cellules mâles à la délétion de Dicer qui jouerait un rôle dans la régulation des Dnmt (Nesterova *et al.*, 2008, voir le chapitre : « ARNi », p74)

- Sun *et al.* observent une association de Dnmt3a avec l'ARN de *Tsix* et supposent que cette interaction sert à partiellement réprimer *Xist* sur le Xa lors de l'initiation de l'inactivation, la méthyl-transférase étant déjà recrutée au promoteur de *Xist* dans les cellules ES ou encore les cellules mutantes pour *Tsix* mais, selon eux, sans être active (Sun *et al.*, 2006)

- Do *et al.* voient une surexpression de *Dnmt3a* et de *Tsix* parallèlement à une répression accélérée de *Xist* en se servant de TSA (inhibiteur d'histone acétyl-déacétylase) au cours d'expériences de fusions nucléaires (l'effet de Dnmt3a étant ici difficile à départager de celui de l'antisens) (Do *et al.*, 2008).

Il semblerait donc que, bien que mineur, il existe bien un effet répressif de Dnmt3a sur l'expression de *Xist*, soit de par son rôle prouvé dans la méthylation de l'ilôt CpG soit par son recrutement d'autres protéines jouant des rôles dans l'hétérochromatinisation. Le rôle minime de cette protéine n'est cependant pas surprenant, si l'on considère *Tsix* et les facteurs de pluripotence comme étant les principaux répresseurs de *Xist*, ainsi que la régulation quasi-normale de *Xist* dans les mutants *Dnmt3a/3b*^{-/-}. De plus, la méthylation de l'ADN étant un processus assez long à effacer, il semble normal que l'évolution ait tendance à privilégier des régulateurs à action rapide pour un phénomène aussi subtil que l'inactivation du chromosome X.

En restant dans le domaine de la méthylation de l'ADN, l'équipe d'Adrian Bird a montré que l'expression de *Xist* était faiblement réactivée dans des cellules fibroblastiques mâles déficientes pour la protéine Mbd2 (une protéine ayant la capacité de se fixer sur l'ADN méthylé), cette réactivation étant abolie lors de l'introduction d'une protéine endogène (Barr *et al.*, 2007). Le rôle de Mbd2 dans la répression de *Xist* semblerait agir à travers un recrutement d'histones déacétylases (Feng *et al.*, 1999, Ng *et al.*, 2000), un effet cumulatif étant observé lors de l'ajout de TSA dans le milieu de culture. De plus, cette protéine semble partager une voie d'action avec Dnmt1, ici encore, une expression ectopique cumulative étant observée dans

des cellules *Mbd2*^{-/-} en présence de 5-azacytidine (un inhibiteur de méthylation de l'ADN) ou de siRNA anti-Dnmt1. Cette relation entre Mbd2, la déacétylation des histones et Dnmt1 semble parfaitement corrélée avec les données de Do *et al.*, nous permettant ainsi de progresser dans la compréhension de mécanismes "secondaires" ou de secours de la répression de *Xist*.

ii. Les activateurs

Contrairement aux répresseurs de l'expression de *Xist*, très peu de choses sont connues à propos du mécanisme induisant l'activation transcriptionnelle de ce gène sur le futur Xi lors de l'induction de la différenciation cellulaire. Plusieurs protéines candidates ont été proposées pour cette fonction, mais il existe très peu d'études démontrant véritablement un lien entre ces facteurs et une activation transcriptionnelle.

- PRC2

Bien que le complexe protéique PRC2 soit responsable du dépôt d'une marque généralement associée à l'hétérochromatine (H3K27me3) sur la partie N-terminale de l'histone H3, l'équipe de Jeannie Lee le modélise en tant que facteur essentiel à la surexpression de *Xist* dans le contexte d'un état hétérochromatique transitoire au promoteur sur le futur X inactif lors de l'initiation de l'inactivation (Zhao *et al.*, 2008). Bien que plusieurs équipes ont maintenant démontré la capacité de PRC2 à se lier à un transcrit comprenant le *Repeat-A* (Zhao *et al.*, 2008, Maenner *et al.*, 2010, Kanhere *et al.*, 2010) sûrement grâce à ses multiples structures tige-boucle, son rôle dans l'activation transcriptionnelle de *Xist* fait encore débat, comme cela a déjà été discuté au cours de cette introduction (voir les chapitres : « Modifications d'histones sur le Xi », p36, « Modification du paysage épigénétique au promoteur de *Xist* », p78 et « Effets de *Xist* sur *Tsix* », p87).

- CTCF

La protéine CTCF a été décrite comme jouant un rôle indirect dans la répression transcriptionnelle de *Xist*, à travers son rôle supposé en tant qu'activateur transcriptionnel de *Tsix* et éventuellement de *Xite*. Sa fixation dans cette région serait requise dans l'appariement des chromosomes X au cours de l'initiation de l'inactivation, notamment à la transvection des régions *Tsix/Xite* qui seraient potentiellement responsables du choix du chromosome X actif, à

travers le maintien monoallélique de l'activité transcriptionnelle de l'antisens au cours des premiers instants de la différenciation (Chao *et al.*, 2002, Donohoe *et al.*, 2007 *et al.*, 2009, voir le chapitre : « Le choix », p15). Toutefois, chez l'homme, l'équipe de Victor Lobanenko a décrit une mutation ponctuelle très proche du promoteur P1 de *Xist* qui influence la fixation de CTCF et par la même occasion biaise le choix du chromosome X à inactiver (Pugacheva *et al.*, 2005). En effet, dans le cas d'une mutation C(-43)G, ils ont observé un biais favorisant l'allèle muté à être inactivé tandis que lors d'une mutation C(-43)A l'allèle muté est préférentiellement choisi pour rester actif. Or, cette position se trouve au sein d'un site de fixation de la protéine CTCF et la mutation affecte sa capacité à se lier à cette position. Au vu des corrélations entre fixation et inactivation, ces observations impliquent donc la fixation de CTCF au promoteur P1 dans un rôle d'activateur transcriptionnel de *Xist*. De plus, ce site de fixation semble être conservé chez la souris (d'après des expériences de retard sur gel et de ChIP), et les auteurs ont pu démontrer que la fixation de CTCF à ce site particulier n'est réalisée que sur le Xi dans les cellules somatiques humaines, ainsi qu'uniquement sur l'X paternel inactivé dans les cellules placentaires de souris femelles.

Comme évoqué précédemment, Navarro *et al.* ont décrit deux autres sites de fixation de CTCF, l'un en amont l'autre en aval du promoteur de *Xist* chez la souris (Navarro *et al.*, 2006), cependant cette étude ne fait état de la fixation de CTCF que dans les cellules ES mâles et femelles non-différenciées et n'observe aucune différence de fixation entre les deux sexes (Figure 24). Malgré cela, l'absence de fixation différentielle ne préclut pas la fonction de CTCF à ces deux autres sites de fixation de participer à la régulation transcriptionnelle de *Xist* en cours de différenciation, d'autant plus qu'une étude a démontré la présence de sites d'hypersensibilité à la DNase aux mêmes endroits, tout en observant que le site correspondant à la fixation de CTCF en aval de P1 disparaît en cours de différenciation cellulaire, l'autre site étant constitutif (Tsai *et al.*, 2008). La protéine CTCF est composée de 11 doigts de zinc, ce qui lui permet de se fixer à plusieurs types de séquences nucléotidiques ainsi que d'interagir avec une quantité importante d'autres protéines, lui conférant aussi la possibilité d'avoir de nombreux rôles indépendants au sein de la cellule, en fonction du contexte chromatinien et de ses modifications post-traductionnelles (pour revue voir Phillips & Corces, 2009). Curieusement, malgré la présence d'un site d'hypersensibilité à la DNase au lieu de fixation de CTCF décrit par Pugacheva *et al.* dans les cellules ES mâles non-différenciées, cette fixation

n'est pas retrouvée par Navarro *et al.* lors de son étude de la région par ChIP ; il pourrait s'agir de la fixation d'une autre protéine se fixant à ce site dans les cellules indifférenciées tandis que CTCF n'est peut-être nécessaire que lors de l'inactivation empreintée (et par homologie avec l'homme, dans le maintien de l'activité transcriptionnel de *Xist*) ou alors d'une sensibilité différente des différents anticorps utilisés dans les deux articles.

- YY1

Dès 1997, lors de la caractérisation de la région promotrice de *XIST* chez l'humain, Hendrich *et al.* ont découvert une séquence consensus de fixation de la protéine YY1 chevauchant le promoteur P1 (Hendrich *et al.*, 1997). Cette protéine est connue aussi bien pour activer que pour réprimer la transcription et interagir avec de nombreux partenaires en fonction du contexte (pour une revue voir Thomas & Seto, 1999 ; Shi *et al.*, 1997). Cette fixation a été confirmée *in vitro* par retard sur gel, et une délétion du site lors d'un test luciférase a bien montré une forte réduction du gène associé, l'implication de YY1 dans ce processus étant difficilement vérifiable étant donné la mutation du promoteur par la même occasion.

Il faudra ensuite attendre 2006 pour que l'étude d'une fixation éventuelle de YY1 dans la région promotrice de *Xist* soit reprise. L'équipe de Joomyeong Kim s'est servie d'un logiciel afin de trouver des régions fortement enrichies (« *clusters* ») en sites de fixation de YY1 (Kim *et al.*, 2006). Ceci lui a permis de trouver aussi bien un cluster de 3 sites juste en amont du promoteur P2 de *Xist*, qu'un *cluster* de 25 sites au sein de la répétition *DxPas34* de *Tsix*, ce dernier étant impliqué dans l'activation transcriptionnelle de l'antisens en collaboration avec CTCF (Donohoe *et al.*, 2007) (Figure 22). De plus, l'analyse par ChIP a révélé que YY1 n'était présent que sur l'allèle possédant un îlot CpG non-méthylé au promoteur de *Xist* (les sites de fixation étant compris dans l'îlot). Cela est confirmé par les sites d'hypersensibilité à la DNase à P1 et P2 (qui peuvent aussi correspondre à CTCF) qui disparaissent en cours de différenciation des cellules ES mâles (Tsai *et al.*, 2008), impliquant donc potentiellement cette protéine dans un rôle dans l'activation transcriptionnelle de *Xist* au sein des cellules somatiques et dans les ovocytes (Kim *et al.*, 2006). La suspicion du rôle de YY1 dans la surexpression de *Xist* a été renforcée par la mutation combinatoire des sites de fixation de la région promotrice placée en amont du gène codant pour la luciférase, ces mutations induisant une chute importante de l'expression du gène. Lors de leur publication suivante, la même

équipe a employé la technique d'ARN interférence en ciblant YY1 (Kim & Kim, 2008). Les résultats concernant la régulation de *Xist* ou *Tsix* par YY1 sont peu concluants, la différence d'expression des deux gènes en présence de siRNA anti-YY1 sont faibles et peu reproductibles mais penchent paradoxalement vers une augmentation de la quantité de transcrits dans les deux cas au sein des cellules femelles. Les auteurs ont corrélé cela à une légère hypométhylation du promoteur de *Xist* dans les cellules interférées pour YY1 par rapport aux contrôles. L'année suivante, le même genre d'étude menée cette fois-ci dans les deux sexes, a permis la visualisation d'une surexpression de *Xist* dans les cellules somatiques mâles possédant une quantité réduite de YY1, et parallèlement à cela on observe aussi une diminution de *Tsix* dans les mêmes cellules mâles, mais aucune différence dans les cellules femelles conformément à l'article précédent (Kim & Kim, 2008). Cela tendrait à confirmer la régulation de *Tsix* par YY1, ce qui engendrerait *de facto* une modification de l'expression de *Xist* (YY1 ne se fixant pas au promoteur de *Xist* dans les cellules somatiques mâles). Cependant ces résultats doivent être traités avec précaution, *Tsix* étant censé être transcriptionnellement éteint dans de telles cellules somatiques, nous indiquant qu'ils sont peut être le fruit d'une étude réalisée sur des quantités extrêmement faibles de matériel et ne représentant potentiellement pas la véritable situation biologique.

De plus, il semble exister une relation entre YY1 et méthylation de l'ADN, cette protéine exhibant une relative incapacité à se fixer à des séquences dont l'ADN est méthylé (Kim, 2009), un dinucléotide CpG se trouvant au sein même de sa séquence consensus. Une autre étude a montré qu'en l'absence de YY1, l'ilôt CpG de *Xist* se trouve fortement déméthylé au cours de l'oogénèse (aucune différence n'étant observée au cours de la spermiogénèse) (Kim *et al.*, 2009b) suggérant un rôle de cette protéine dans la méthylation *de novo* de cette région différenciellement méthylée (« DMR » pour *Differentially Methylated Region*). Il est intéressant de noter que CTCF (un partenaire de YY1 d'après Donohoe *et al.*, 2007) joue aussi un rôle dans la méthylation de DMRs, mais en empêchant la méthylation de ceux-ci sur l'allèle maternel durant l'embryogénèse (Fedoriw *et al.*, 2004, Engel *et al.*, 2006) alors que paradoxalement YY1 semblerait être requise pour la méthylation maternelle-spécifique de *Xist* au cours de l'oogénèse. Il est possible que YY1 exerce cette fonction en recrutant des méthyltransférases. Son incapacité à se fixer à l'ADN méthylé pourrait être un moyen de reconnaissance des sites ne nécessitant pas de méthylation *de novo*.

Toutefois, son rôle dans la méthylation de l'ADN semble fortement incompatible avec son rôle supposé en tant qu'activateur transcriptionnel de *Xist*, à moins qu'elle ait des fonctions différentes à ce locus selon l'état développemental et/ou le sexe de la cellule. L'étude de ces phénomènes à d'autres stades embryonnaires ainsi qu'une recherche précise de partenaires sont nécessaires à une véritable caractérisation de son implication dans ces différents processus.

- Rnf12

Rnf12 code pour une E3 ubiquitine-ligase impliquée dans la dégradation de protéines interagissant avec LIM, dans l'activation transcriptionnelle et dans l'homéostasie de la longueur des télomères (Bach *et al.*, 1999 ; Her & Chung, 2009 ; Johnsen *et al.*, 2009).

De manière intéressante, cette protéine surexprimée autour de la période correspondant à la mise en place de l'inactivation (aussi bien chez les mâles que les femelles), induit une proportion anormalement élevée de doubles domaines de *Xist*, lorsque surexprimé au sein de cellules femelles en différenciation (Jonkers *et al.*, 2009). De même lorsque surexprimé dans des cellules mâles en différenciation, les auteurs observent une proportion importante de chromosomes X inactifs. Ils postulent donc avoir trouvé un activateur de l'inactivation du chromosome X qui biaise l'étape de comptage, et par la même occasion un candidat idéal pour le rôle de facteur de compétence (voir le chapitre : « Le comptage », p14) (Figure 26).

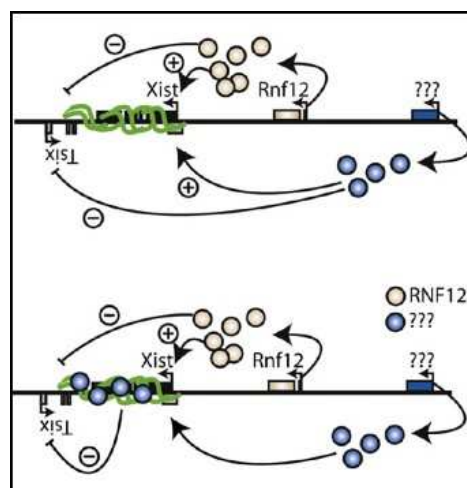


Figure 26 : Modèle d'action de Rnf12 dans l'initiation de l'inactivation du chromosome X. Une activation de l'inactivation médiée par Rnf12 pourrait impliquer une activation de *Xist* ou une répression de *Tsix* à travers des mécanismes directs ou indirects (Rnf12 en beige). L'action d'autres activateurs putatifs de l'inactivation (en bleu) pourrait impliquer un mécanisme similaire à celui imaginé pour Rnf12 (en haut) ou pourrait impliquer la répression de *Tsix* par le recrutement de l'ARN *Xist* (en bas). (D'après Jonkers *et al.*, 2009)

Seule une quantité réduite de *Rnf12* serait nécessaire pour activer *Xist* au sein de cellules mâles mutées pour *Tsix* (Luikenhuis *et al.*, 2001, Vigneau *et al.*, 2006, Sado *et al.*, 2002) ou de cellules contenant des transgènes ectopiques (Herzing *et al.*, 1997, Lee *et al.*, 1999, Migeon *et al.*, 2001, Heard *et al.*, 1999), en raison de l'absence d'un des répresseurs majeurs de *Xist*. Toutefois, les auteurs ignorent encore le mécanisme d'action de *Rnf12* mais proposent qu'il n'agit pas seul de par l'inactivation subsistante (bien que tardive) au sein de cellules femelles *Rnf12*^{+/-} (un seul allèle ne devant pas être suffisant pour induire l'inactivation).

Le fait le plus intrigant est de savoir comment cette protéine arrive à contourner les voies de répression médiées par *Tsix* ainsi que par les facteurs de pluripotence au sein des cellules ES, étant donné l'inactivation ectopique d'un chromosome X au sein de 10 à 40% de cellules femelles non-différenciées chez qui l'on a surexprimé *Rnf12*. Un rôle s'exerçant uniquement sur *Tsix* n'est pas envisageable, la délétion de l'antisens n'engendrant pas une surexpression de *Xist* suffisante pour inactiver le chromosome X de l'allèle ciblé. De plus, cette régulation semblerait s'effectuer de manière indirecte, l'introduction d'un transgène de *Xist/Tsix* en simple copie dans des cellules ES femelles générant très rarement une surexpression de *Xist* à partir de ce transgène (Heard *et al.*, 1999), un facteur intermédiaire paraissant donc requis afin d'établir l'effet observé.

Rnf12 a donc la capacité d'induire l'expression de *Xist* lorsque surexprimé, il est donc capital de le réprimer correctement dans les cellules ES. De manière fort intéressante, plusieurs sites d'Oct4, Sox2 et de Nanog ont été montrés comme se fixant sur le gène *Rnf12* par ChIP-seq dans les cellules ES (Chen *et al.*, 2008 ; Kim *et al.*, 2008). Cette répression éventuelle générée par les facteurs de pluripotence pourrait expliquer la surexpression de *Rnf12* en cours de différenciation.

- Xpr

L'équipe d'Edith Heard a démontré l'importance d'une région nommée *Xpr* (superposée au gène *Slc16a2/Xpct*) dans l'interaction transitoire en *trans* entre les deux chromosomes X avant l'initiation de l'inactivation (autour de 1,5 à 2 jours post-différenciation) (Augui *et al.*, 2007) (Figure 5). Parallèlement à ceci, cette région semblerait être impliquée dans une surexpression ectopique modérée de *Xist*, une accumulation de ses transcrits étant apercevable dans une proportion (au plus 30%) de cellules ES femelles non-différenciées lors de l'intégration d'un

transgène du *Xpr* en simple-copie. Et de manière identique, aucune cellule ES mâle contenant ce transgène n'a pu être obtenue, possiblement à cause d'une surexpression létale de *Xist* au sein de ces cellules (nullisomie fonctionnelle). Tout ceci en fait donc aussi un candidat pour le rôle de facteur de compétence (voir le chapitre : « Le comptage », p14).

Il est tentant de spéculer sur l'interaction de *Rnf12* et de *Xpr* au cours des étapes de comptage et de choix lors des phases précoces de l'inactivation du chromosome X. Leur localisation génétique pourrait même présenter des caractéristiques intéressantes pour un mécanisme en cascade, *Rnf12* étant situé à environ 500kb et *Xpr* (ou *Xpct*) à environ 200kb en amont de *Xist*, les trois gènes étant transcrits dans la même orientation. Le gène *Ftx*, qui semblerait aussi jouer un rôle dans l'activation transcriptionnelle de *Xist* (Chureau *et al.*, soumis), se trouve à 100kb et est aussi transcrit dans la même orientation, offrant l'hypothèse d'une fonction de toute la région 5' de *Xist*, plus particulièrement sur le brin sens, dont les loci présentent tous un pic d'activité lors de l'initiation de l'inactivation (que ce soit au niveau de la transcription ou bien de la localisation nucléaire), dans la régulation de l'ARN non-codant responsable de l'inactivation du chromosome X.

Chapitre II : Projet de Thèse

Depuis maintenant quelques temps, l'équipe de Claire Rougeulle s'intéresse (entre autres) aux différents mécanismes de régulation transcriptionnelle servant à établir la coordination des gènes responsables de l'initiation de l'inactivation du chromosome X. Ces projets s'appuient sur l'utilisation des cellules souches embryonnaires murines comme modèle, celles-ci récapitulant parfaitement le processus de l'inactivation aléatoire lorsque leur différenciation (en l'occurrence, par acide rétinoïque) est induite. Et c'est dans ce contexte, en reprenant certains projets préalablement initiés au laboratoire par Pablo Navarro et en cherchant à les développer, que s'inscrit mon projet de thèse.

Initialement, mon projet de thèse avait pour objectif de confirmer l'implication de certaines protéines candidates dans le rôle d'activateur transcriptionnel du gène *Xist*, car comme nous avons pu le voir au cours de l'introduction, bien que les mécanismes répressifs de ce gène commencent à être bien décrits, les processus permettant sa surexpression sont presque totalement inconnus.

Il y a quatre ans, lors de mon arrivée à l'Institut Pasteur, plusieurs articles venaient d'être publiés mettant principalement en cause deux protéines dans l'activation transcriptionnelle de *Xist* : CTCF et YY1 (Pugacheva *et al.*, 2005, Hendrich *et al.*, 1997, Kim *et al.*, 2006). De plus, Pablo Navarro avait préalablement étudié la fixation de certaines protéines dans toute la région du locus *Xist/Tsix* par immunoprécipitation de chromatine (Collas *et al.*, 2010), ce qui lui a notamment permis de trouver deux sites CTCF encore inconnus autour du promoteur de *Xist* (Navarro *et al.*, 2006). Muni de ces informations et de cette technique, mon objectif était de décrire les schémas de fixation de ces protéines dans la région promotrice de *Xist* au sein de différents types cellulaires et ainsi tenter de relier les résultats obtenus à un rôle dans la régulation de notre gène "fétiche".

Par la suite, afin de valider les résultats obtenus par ChIP, nous avons envisagé d'étudier de plus près l'implication fonctionnelle de la fixation de chacune des protéines à leurs différents sites le long de la région promotrice de *Xist*. Dans ce but, nous avons voulu réaliser des mutations individuelles ou collectives de ces différents sites *in vivo* en utilisant une technique de RMCE modifiée (« *Recombinase-Mediated Cassette Exchange* ») (Bouhassira *et al.*, 1997) afin de voir l'effet de l'absence de fixation de ces protéines sur la régulation de *Xist* ainsi que sur l'inactivation dans son ensemble.

De manière similaire à bon nombre de thèses, ainsi qu'à la recherche en biologie en général, certains aspects de mon projet initial se sont avérés fructueux, et d'autres moins. Mais les découvertes établies par d'autres laboratoires ainsi que les discussions fructueuses avec divers

collègues m'ont permis de quelque peu réorienter certains de mes projets afin de mieux développer certains autres aspects, tout aussi intéressants, du sujet. La caractérisation de nouvelles protéines impliquées dans la régulation transcriptionnelle de *Xist* et de *Tsix* a, notamment, occupé une grande partie de mon temps, me permettant d'acquérir une certaine expérience dans l'intégration d'informations à un modèle précis, en imaginant perpétuellement de nouveaux acteurs et comment ils pourraient participer à la régulation de notre locus d'intérêt.

Au cours de cette prochaine partie mêlant résultats et discussions, je m'attellerai à présenter les résultats que j'ai pu obtenir au cours de ces quatre dernières années passées sous la direction du Dr. Claire Rougeulle.

Les quatre volets que j'ai choisi de vous présenter sont :

- L'étude du réseau transcriptionnel régulant l'expression de *Xist* chez la souris et sa conservation chez l'humain
- La caractérisation fonctionnelle des sites de fixation pour les protéines CTCF et YY1 dans la région promotrice de *Xist* chez la souris
- L'étude du réseau transcriptionnel régulant *Xist* au sein de cellules germinales embryonnaires de souris.
- Le couplage moléculaire de la régulation de *Tsix* et de la pluripotence cellulaire

Chapitre III : Résultats et **Discussions**

I. L'étude du réseau transcriptionnel régulant l'expression de *Xist* chez la souris et sa conservation chez l'humain

Comme nous venons de le voir en introduction, les acteurs gouvernant la répression, aussi bien transcriptionnelle que post-transcriptionnelle, de *Xist* commencent à être connus bien que les mécanismes mêmes de ce processus n'ont pas encore été élucidés. Cependant, dans la cellule ES femelle sauvage, une fois la différenciation initiée et les facteurs répressifs (l'antisens et les facteurs de pluripotence) disparus, des facteurs activant la transcription sont nécessaires à la surexpression de *Xist* menant à l'inactivation du chromosome X en *cis*. En se basant donc sur la mise en avant de quelques protéines candidates à ce rôle (voir le chapitre : « Régulation en *trans* de *Xist* par divers facteurs », p92), nous avons cherché à confirmer certains résultats (Kim *et al.*, 2006, Hendrich *et al.*, 1997, Pugacheva *et al.*, 2005, Navarro *et al.*, 2006) et à les approfondir afin d'essayer d'obtenir une vision plus claire de l'ensemble des mécanismes régulant l'expression de *Xist* et de la manière dont ils se coordonnent les uns par rapport aux autres.

Parmi les protéines candidates, CTCF et YY1 sont toutes deux connues comme jouant de multiples rôles au sein de la cellule, notamment celui d'activateur transcriptionnel. Au cours des dernières années, l'observation de sites de fixation leur appartenant, au sein de la région promotrice de *Xist*, nous a intrigué. En effet, ces protéines ont été décrites se fixant au promoteur P1 de *Xist* dans des cellules somatiques femelles, cette fixation étant femelle-spécifique dans les cellules somatiques (Hendrich *et al.*, 1997, Pugacheva *et al.*, 2005). Puis, plus récemment, plusieurs d'autres sites de fixation de ces protéines ont été décrits par différentes équipes : l'équipe de Joomyeong Kim décrivant trois sites de YY1 plus proches du promoteur P2 de *Xist*, dont l'occupation semble aussi corrélée avec l'expression de ce dernier (Kim *et al.*, 2006), et notre équipe décrivant deux nouveaux sites de CTCF (l'un environ 1kb en amont et l'autre 1,5kb en aval du promoteur P1), ces deux pics semblant ubiquitaires dans les cellules femelles (Navarro *et al.*, 2006).

Avant de commencer à décrire les résultats obtenus au cours de mon doctorat concernant la caractérisation du réseau transcriptionnel de *Xist*, je souhaiterais tout d'abord consacrer quelques mots à la description des protéines CTCF et YY1, la connaissance de leurs rôles et de leurs interactions multiples étant capitale à la compréhension des données présentées par la suite.

a. CTCF

La protéine CTCF est impliquée dans de nombreux mécanismes au sein de la cellule dont l'activation et la répression transcriptionnelle, l'insulation, le blocage de la méthylation de l'ADN et la structure tridimensionnelle de la chromatine. Sa localisation est strictement nucléaire (Klenova *et al.*, 1993), le seul cas de localisation cytoplasmique ayant été décrit dans des cellules tumorales (Rakha *et al.*, 2004, Butcher *et al.*, 2007). Son expression est ubiquitaire (à part au sein des spermatocytes primaires, Loukinov *et al.*, 2002) mais est régulée au cours du cycle cellulaire avec un pic d'expression au cours de la transition entre les phases G2 et S (Klenova *et al.*, 1998). Au sein du noyau, une faible proportion de CTCF est associée la matrice nucléaire (Dunn *et al.*, 2003 ; Ottaviani *et al.*, 2009) ou localise à la périphérie nucléolaire (Yusufzai & Felsenfeld, 2004) (surement grâce à ces interactions avec les lamines et la nucléophosmine, respectivement). Le reste des protéines est réparti le long des chromosomes et plus particulièrement aux centrosomes lors de la mitose (surtout de la métaphase à l'anaphase), potentiellement toujours grâce à son interaction avec la nucléophosmine (Zhang *et al.*, 2004 ; Burke *et al.*, 2005 ; Grisendi *et al.*, 2005). Enfin, CTCF pourrait être impliquée dans la réplication asynchrone de certains gènes exprimés monoallèlement (Ensminger & Chess, 2004) comme c'est le cas à *H19/IGF2* (Bergstrom *et al.*, 2007).

Le fait que CTCF soit impliquée dans autant de processus tient surement en grande partie de sa structure particulière composée de 11 doigts de zinc (Figure 27) grâce auxquels elle peut se lier à diverses séquences nucléotidiques et interagir avec de multiples autres protéines ainsi qu'avec elle-même (Gaszner & Felsenfeld, 2006 ; Pant *et al.*, 2004 ; Yusufzai *et al.*, 2004).

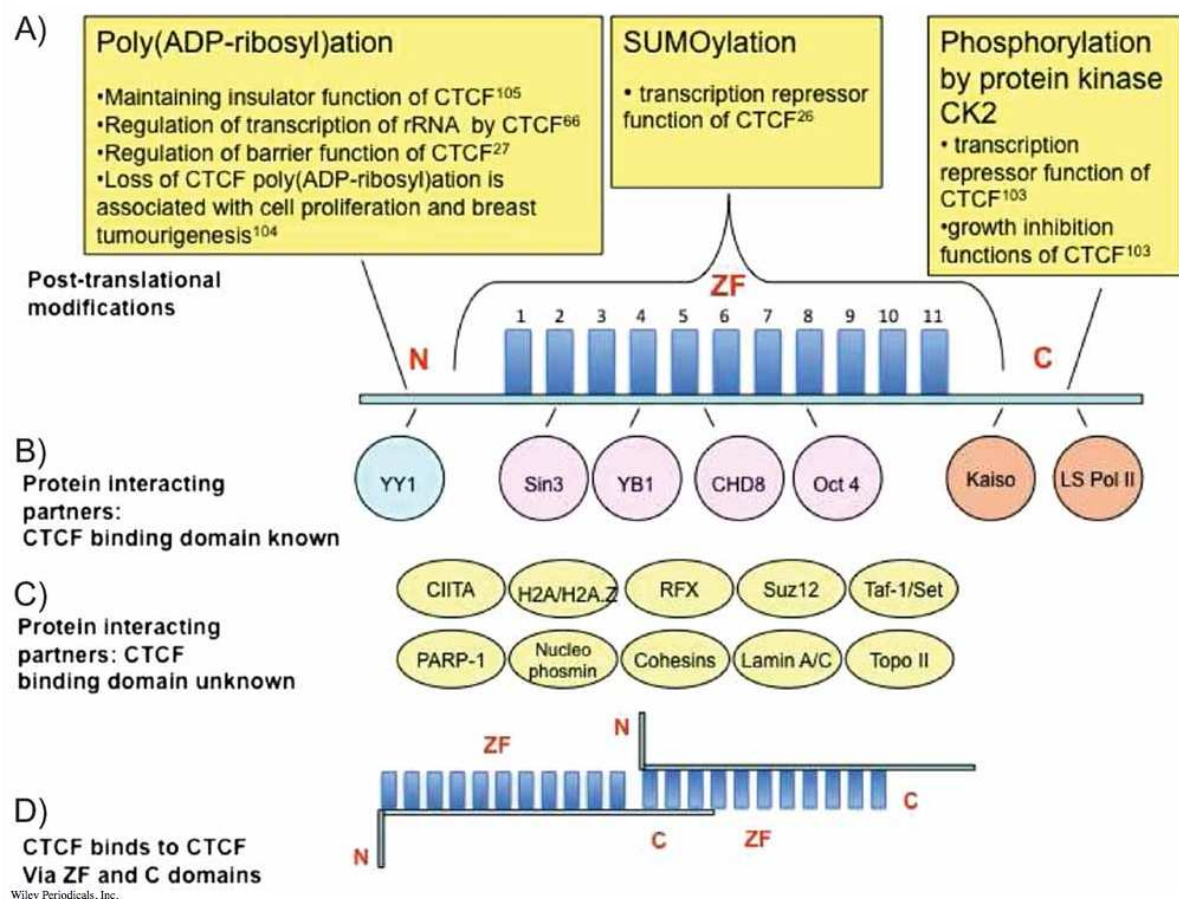


Figure 27 : La structure de CTCF, ainsi que ses modifications et partenaires connus. (A) Représentation schématique de CTCF. Elle est composée de trois domaines distincts : un domaine N-terminal (N), 11 doigts de zinc (ZF) et un domaine C-terminal (C). Les modifications post-traductionnelles de CTCF sont montrées sur la partie supérieure de la figure, avec le domaine modifié. **(B)** Les protéines qui interagissent à travers un domaine connu, avec le domaine d'interaction indiqué. **(C)** Autres protéines interagissant avec CTCF. **(D)** Une interaction entre deux protéines CTCF *via* les ZF et le domaine C-terminal. (D'après Ohlsson *et al.*, 2010)

Bien que cette protéine puisse utiliser différents doigts de zinc pour se lier à l'ADN en fonction du contexte (Filippova *et al.*, 1996, Burcin *et al.*, 1997, Quitschke *et al.*, 2000), grâce à des expériences récentes se servant de l'analyse informatique de données obtenues par ChIP-seq, plusieurs équipes ont pu définir une séquence consensus qui semblerait effectivement corrélérer avec une fixation de CTCF dans 80 à 95% des cas (Jothi *et al.*, 2008, Chen *et al.*, 2008, Cuddapah *et al.*, 2009), cette séquence étant quasiment identique chez l'homme et chez la souris. Ces mêmes expériences ont aussi permis d'identifier près de 40 000 sites de fixation au sein des cellules ES de souris, cette influence monumentale implique cette protéine dans bon nombre de processus cellulaires, expliquant sûrement la létalité pré-implantatoire de l'embryon muté pour CTCF (Heath *et al.*, 2008, Splinter *et al.*, 2006).

Un autre facteur pouvant influencer sur la diversité des fonctions et des interactions de CTCF est le fait que cette protéine peut subir de nombreuses modifications post-traductionnelles telles que la poly-(ADP-ribosyl)ation (Yu *et al.*, 2004, Torrano *et al.*, 2006), la phosphorylation (Klenova *et al.*, 2001 ; El-Kady & Klenova, 2005 ; Delgado *et al.*, 1999) et la sumoylation (MacPherson *et al.*, 2009 ; Kitchen & Schoenherr, 2010) (Figure 27).

i. Fonction activatrice/répressive

CTCF a été décrite comme possédant un rôle dans l'initiation, l'activation, la répression et l'élongation transcriptionnelle au sein du noyau.

Sa fonction d'initiateur pourrait se dérouler principalement à travers son interaction avec la grande sous-unité de l'ARN polymérase II (Rpb1) (Chernukhin *et al.*, 2007), qu'elle soit hypo- ou hyper-phosphorylée (la première forme étant associée à l'initiation, la deuxième à l'élongation de la transcription) (Dahmus, 1996). CTCF semble donc pouvoir jouer le rôle de la TBP (« *TATA-Binding Protein* »), initiant faiblement la transcription d'un gène reporteur sans séquence promotrice en recrutant directement la machinerie de transcription. Cependant, l'affinité de ces deux protéines l'une pour l'autre peut être réduite lors de la phosphorylation de la partie C-terminale de CTCF, partie justement responsable de cette interaction (Chernukhin *et al.*, 2007). Cela paraît plutôt paradoxal sachant que cette modification post-traductionnelle de la protéine est généralement associée à une augmentation de la capacité à activer la transcription (El-Kady & Klenova, 2005 ; Klenova *et al.*, 2001). De plus, CTCF peut aussi engendrer une activation transcriptionnelle simplement en recrutant d'autres protéines jouant aussi ce rôle, telles que YY1 (Donohoe *et al.*, 2007) ou Oct4 (Donohoe *et al.*, 2009).

Son rôle de répresseur, quant à lui, semblerait pouvoir dépendre de plusieurs mécanismes différents. D'une part, bien qu'elle puisse interagir avec l'ARN polymérase II, comme nous venons de le voir, CTCF a aussi la capacité de la bloquer et de la sequestrer en cours d'élongation (Chernukhin *et al.*, 2007, Cho *et al.*, 2005). D'autre part, il a été prouvé que CTCF pouvait interagir avec la protéine Suz12 (Li *et al.*, 2008), membre du complexe protéique PRC2 (responsable du dépôt de la marque H3K27me3) induisant un contexte chromatinien défavorable à la fixation de la machinerie de transcription, mais aussi avec d'autres protéines pouvant réprimer la transcription : Oct4, YB1 (Klenova *et al.*, 2004, Chernukhin *et al.*, 2000), YY1, Sin3 (Lutz *et al.*, 2000).

ii. Fonction d'insulateur

Un insulateur est une séquence d'ADN qui empêche l'activation d'un promoteur par un « *enhancer* » inapproprié et/ou bloque la propagation de chromatine condensée. Chez les mammifères, il a été observé que tous les éléments génétiques qualifiés d'insulateurs sont fixés par CTCF. Cependant, cette protéine semble y jouer d'avantage un rôle dans le blocage d'activation (Bushey *et al.*, 2009 ; Wallace & Felsenfeld, 2007) (Figure 28 haut), que dans l'entrave à la propagation de l'hétérochromatine (Figure 28 bas), cette fonction n'ayant jamais été démontrée expérimentalement de manière définitive (Splinter *et al.*, 2006 ; Recillas-Targa, 2002 ; Ciavatta *et al.*, 2006 ; Gaszner & Felsenfeld, 2006) malgré sa présence à de nombreuses frontières d'états chromatinien (Cho *et al.*, 2005 ; Filippova *et al.*, 2005 ; Barski *et al.*, 2007 ; Cuddapah *et al.*, 2009) et sa capacité à bloquer la propagation de la méthylation d'ADN (Engel *et al.*, 2006 ; Fedoriw *et al.*, 2004 ; Filippova *et al.*, 2005 ; Lewis & Murrell, 2004 ; Pant *et al.*, 2004 ; Schoenherr *et al.*, 2002).

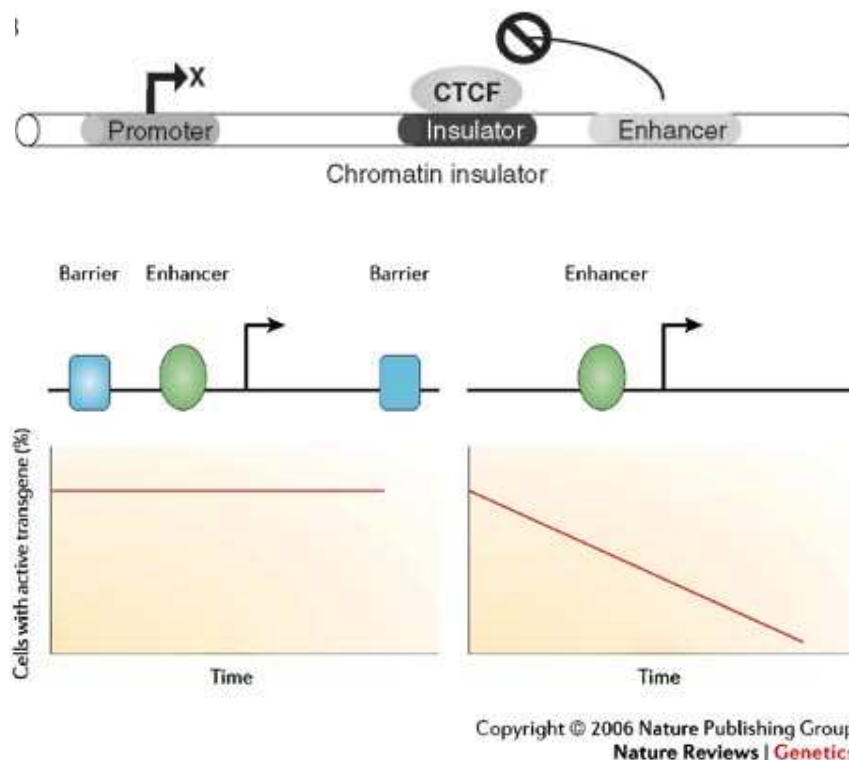


Figure 28 : Les deux rôles possibles pour CTCF en tant qu'insulateur. (Haut) CTCF se fixe à tous les insulateurs connus chez les mammifères et y fonctionne comme une protéine insulatrice de chromatine qui n'affecte pas directement la transcription but empêche une interaction entre un promoteur et des amplificateurs ou des répresseurs à proximité. Ceci empêche l'activation ou la répression inappropriée de gènes en établissant des domaines chromatinien indépendamment régulés. **(Bas)** Un insulateur barrière protège un transgène contre une répression due à l'hétérochromatinisation. Les graphiques montrent l'activité du transgène dans de telles situations. La barrière doit encadrer en 5' et en 3' les extrémités de la construction afin de prévenir de l'hétérochromatinisation pouvant arriver des deux côtés lors de l'insertion du transgène. (D'après Filippova *et al.*, 2005 et Gaszner & Felsenfeld, 2006)

La présence de CTCF aux frontières entre euchromatine et hétérochromatine a notamment déjà été observée en rapport avec l'inactivation du chromosome X. En effet, CTCF pourrait être impliquée dans le fait que certains gènes échappent à l'inactivation sur le Xi, de par sa présence aux frontières des domaines demeurant actifs dans les cellules somatiques femelles (Filippova *et al.*, 2005). De plus, des études récentes par ChIP-seq ont aussi démontré la présence de cette protéine aux frontières entre la marque hétérochromatique H3K27me3 et la marque euchromatique H2AK5Ac (Barski *et al.*, 2007, Cuddapah *et al.*, 2009). Cependant, différentes études prouvent qu'en l'absence de CTCF on n'observe pas de propagation de l'hétérochromatine (Splinter *et al.*, 2006, Recillas-Targa *et al.*, 2002), ceci signifiant que l'entrave à ce phénomène est sans doute sous le contrôle d'une autre protéine (qui interagit potentiellement avec CTCF).

Le mécanisme d'insulation semble passer par un phénomène de formation de structures tridimensionnelles impliquant l'élaboration de structures en boucle favorisant ou défavorisant l'accès aux éléments sur la boucle par d'autres complexes protéiques. Cette implication dans la formation de boucles chromatinienne se reflète d'ailleurs au niveau des protéines interagissant avec CTCF. En effet, la nucléophosmine (Yusufzai & Felsenfeld, 2004) (aussi appelée NPM1 ou B23, est associée aux structures ribonucléoprotéiques nucléolaires), la cohésine (Wendt *et al.*, 2008, Parelho *et al.*, 2008, Stedman *et al.*, 2008) (responsable de l'appariement des chromatides sœurs) et CHD8 (une hélicase, capable de modifier la flexibilité de la chromatine) (Ishihara *et al.*, 2006 ; Wallace & Felsenfeld, 2007) interagissent toutes trois avec CTCF et jouent toute un rôle dans la modification de la structure chromatinienne. Cette interaction avec la nucléophosmine pourrait aussi être responsable de la localisation périnucléolaire des régions insultrices (ottaviani *et al.*, 2009 ; Yusufzai *et al.*, 2004).

D'autres protéines, telles que Kaiso (Defossez *et al.*, 2005), Suz12 (Li *et al.*, 2008) et Parp1 (à travers la poly-(ADP-ribosyl)ation) (Yu *et al.*, 2004) ont été mises en cause par rapport à un rôle dans l'insulation. Par exemple, l'interaction entre Kaiso (un facteur de transcription pouvant aussi bien se lier à l'ADN méthylé que non-méthylé) et CTCF est défavorable à la fonction de bloqueur d'*enhancer* de cette dernière, tandis que sa poly-(ADP-ribosyl)ation semblerait, au contraire, la stimuler (Yu *et al.*, 2004).

iii. Rôle dans la méthylation de l'ADN

CTCF est une protéine dont la fixation est dépendante de l'état de méthylation de son site. En effet, elle accomplit souvent sa tâche d'insulateur à des loci empreintés, ceux-ci étant souvent

différentiellement méthylés (d'où l'appellation de DMR pour « *Differentially Methylated Region* ») afin d'obtenir une expression monoallélique, de par le fait que CTCF ne peut se fixer à ceux déjà méthylés (comme par exemple à *H19/Igf2*, Bell & Felsenfeld, 2000, Hark *et al.*, 2000) (Figure 29A et B). Malgré cela, elle semble aussi avoir la capacité d'empêcher la méthylation *de novo* de ses sites de fixation (Schoenherr *et al.*, 2002) (comme vu précédemment).

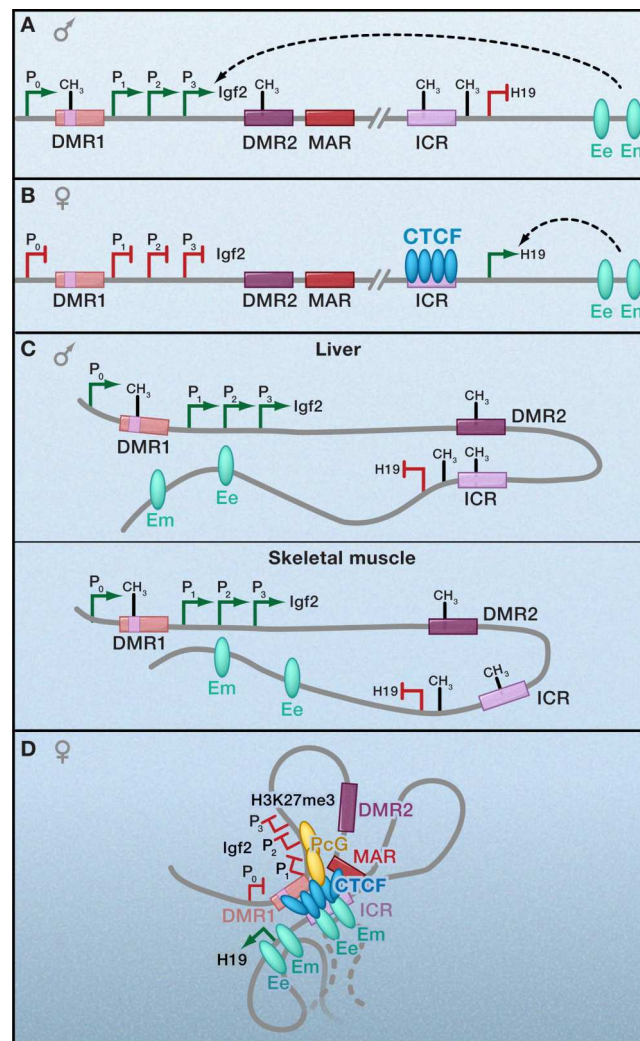


Figure 29 : CTCF organise des contacts chromatinien aux loci empreintés. (A et B) Représentation linéaire du locus *H19/Igf2* murin. Le gène non-codant maternellement exprimé *H19* est situé à environ 90kb du gène codant *Igf2* qui est exprimé uniquement à partir de l'allèle paternel. La région de contrôle de l'empreinte (ICR) environ 2kb en amont d'*H19* contient 4 sites CTCF et est essentiel à la régulation du locus. Les régions de méthylation différentielles (DMRs) agissent en concert afin de réguler réciprocalement l'expression allèle-spécifique d'un groupe partagé d'amplificateurs en aval de *H19*. -CH₃, méthylation de l'ADN. Ovaes verts : amplificateurs. (C et D) Modèles schématisques tridimensionnels, illustrant le schéma allèle-spécifique de la fixation de CTCF, de la méthylation de l'ADN et des boucles chromatinien. Bien que les boucles sont ici représentées *via* une multimerisation de CTCF, il n'est pas clair si ces interactions longue distance peuvent être attribué à la fixation de CTCF à tous les sites suivi d'une dimerisation ou si la détection de CTCF par ChIP est due à des interactions indirectes due aux boucles. Les lignes discontinues indiquent un mouvement des amplificateurs entre l'ICR et le promoteur d'*H19*. (D'après Phillips & Corces, 2009)

Cette capacité pourrait s'expliquer par son interaction avec la protéine Parp1 (Yusufzai *et al.*, 2004, Guastafierro *et al.*, 2008, Caiafa & Zlatanova, 2009). La protéine Parp1 est responsable du dépôt de la marque poly(ADP-ribosyl) sur des protéines chromatinienne, dont elle-même. L'ajout post-transcriptionnel de segments d'ADP-ribose engendre la présence de charges négatives supplémentaires pouvant modifier les interactions entre protéines ou entre protéine et ADN.

Il a été démontré que, bien que Parp1 polyADP-ribosyle CTCF, l'interaction entre ces deux protéines est aussi nécessaire pour engendrer l'automodification de Parp1 (Guastafierro *et al.*, 2008), phénomène responsable de l'inhibition fonctionnelle de l'ADN méthylase de maintien, Dnmt1 (Reale *et al.*, 2005).

Ces interactions entre CTCF, Parp1 et Dnmt1 sont visibles au niveau du génome par son hypométhylation d'environ 70% lorsque CTCF est surexprimée (Guastafierro *et al.*, 2008). Mais elles sont aussi visibles à une échelle plus réduite, la colocalisation de CTCF et de Parp1 étant observée à près de 140 sites (lors d'une expérience de ChIP-CHIP), dont le locus empreinté *H19/Igf2* qui, lors de l'inhibition des protéines de la famille Parp (à l'aide de 3-aminobenzamide), témoigne d'une expression biallélique (Yu *et al.*, 2004). Ceci impliquerait donc cette interaction protéique dans le maintien de l'empreinte parentale, CTCF recrutant Parp1 à certains de ses sites de fixation, cette dernière servant à bloquer la méthylation de l'ADN des régions voisines (Mukhopadhyay *et al.*, 2004).

Il faut aussi noter que la perte de l'hélicase à chromodomaine CHD8 semble être liée à une augmentation de la méthylation (ainsi qu'à une perte d'acétylation de l'histone H3) à certains sites CTCF (Ishihara *et al.*, 2006), bien que l'on ne connaisse pas encore les mécanismes sous-jacents de cette observation malgré un effet de la surexpression de CHD8 sur l'activité de Parp (Nishiyama *et al.*, 2009).

iv. Formation de structures tridimensionnelles

Comme nous l'avons vu, CTCF semble interagir avec plusieurs protéines impliquées dans la formation d'interactions inter- ou intra-chromosomiques (la nucléophosmine, la cohésine, CHD8, Lamine A/C). L'observation que CTCF participe à ce type de phénomène a pu récemment voir le jour grâce aux avancées techniques, fournissant aux chercheurs un moyen de visualiser les différentes interactions à distance existant à travers le génome, notamment grâce aux techniques de 3C, 4C et 5C (Simonis *et al.*, 2007, Fullwood & Ruan, 2009) (voir Annexe 3, p270).

Dernièrement, la découverte d'une interaction entre CTCF et le complexe de la cohésine est venue renforcer le rôle joué par ces structures tri-dimensionnelles dans les mécanismes régulatoires mis en place par CTCF (Wendt *et al.*, 2008, Parelho *et al.*, 2008, Stedman *et al.*, 2008, Rubio *et al.*, 2008).

La cohésine est le complexe responsable de l'appariement des chromatides sœurs au cours des phases S, G2, et en début de mitose. Elle est composée de 4 sous-unités : Smc1, Smc3, Scc1 (ou Mcd1 ou Rad21) et Scc3 (ou SA2 ou Stag2), l'agencement de ces différentes sous-unités permettant de maintenir la proximité de deux brins de chromatine distincts (Figure 30).

Mais d'après ces découvertes récentes, il semblerait que la cohésine puisse aussi jouer un rôle important dans la régulation transcriptionnelle des gènes à travers son interaction avec CTCF (bien qu'une publication récente fasse état d'un tel rôle sans l'intervention de CTCF, Schmidt *et al.*, 2010).

Conformément au fait que CTCF et la cohésine partagent des sites consensus quasi-identiques (voir le chapitre : « Recrutement de Rad21 », p149) (Ohlsson *et al.*, 2009 bioessay), leur co-localisation à travers le génome est extrêmement fréquente : 65 à 90% des sites liant la cohésine lient aussi CTCF, tandis que 55 à 80% des sites CTCF fixent également la cohésine (Parelho *et al.*, 2008, Wendt *et al.*, 2008). Cette interaction se ferait plus particulièrement entre CTCF et la sous-unité Scc3 de la cohésine (Rubio *et al.*, 2008).

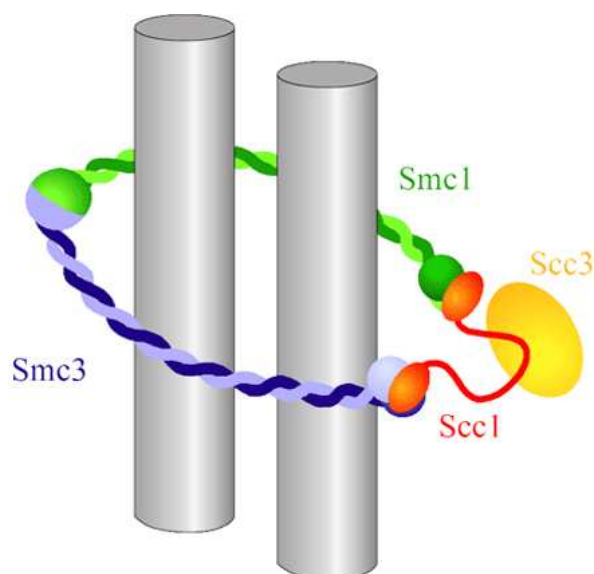


Figure 30 : Schéma modélisant la structure du complexe de la cohésine. D'après Haering *et al.*, 2005, <http://www.sciencemag.org.gate2.inist.fr/feature/data/prizes/ge/2004/haering.dtl>

La mise en place des structures tridimensionnelles, engendrées par l'association de deux sites distincts (distants d'un minimum de 10kb d'après Rippe, 2001), est doute facilitée par la

capacité de CTCF et de la cohésine à homo- et hétérodimériser entre eux (Zhang *et al.*, 2009). L'interdépendance des deux protéines dans le contexte de leur colocalisation est clairement observable lors d'expériences de déplétion. D'après les données obtenues en 3C, CTCF semble recruter la cohésine tandis que la cohésine semble stabiliser la fixation de CTCF (Hou *et al.*, 2010). Cependant, il est fort possible que cette interaction soit cycle cellulaire-dépendante, CTCF pouvant être fixée de manière stable tandis que la cohésine n'interagirait de manière maximale qu'au cours de la phase S, ces interactions étant défaites au cours des phases G2/M comme cela a déjà été démontré (Kang & Lieberman, 2009).

En résumé, la cohésine pourrait soit stabiliser les interactions longue distance mises en place par CTCF, soit être elle-même responsable de ces interactions intra- et inter-chromosomiques à travers le génome. Ces interactions peuvent donc avoir plusieurs fonctions au sein du noyau :

- Être responsable de modifications de régulation par transvection
- Créer des boucles chromatinienne afin de favoriser ou au contraire, d'empêcher l'accès de la machinerie de transcription ou leurs recruteurs à un promoteur précis
- Ancrer des loci à la matrice nucléaire ou à la périphérie du nucléole

Il n'est pas exclu que d'autres protéines participent à ce type de restructuration chromatinienne.

Pour conclure sur l'ensemble des rôles exercés par la protéine CTCF, il semblerait que celle-ci ait typiquement un rôle de protéine d'échafaudage, c'est-à-dire dont la fonction s'accomplit surtout à travers de nombreuses interactions protéiques ayant chacune un rôle bien particulier, permettant ainsi à CTCF de jouer un rôle majeur dans bon nombre de processus transcriptionnels et épigénétiques. Ces interactions étant déterminées par la séquence nucléotidique ainsi que le contexte chromatinien du site de fixation, cela permet aux différents domaines de CTCF (notamment les doigts de zinc) de bénéficier d'un éventail d'interacteurs.

b. YY1

De manière similaire à CTCF, YY1 possède une fonction indispensable dans le développement embryonnaire, des mutants homozygotes démontrant une mortalité péri-

implantatoire, et les hétérozygotes étant retardés développementalement et présentant des défauts de neurulation (Donohoe *et al.*, 1999).

YY1 possède 4 doigts de zinc, localisés à l'extrémité C-terminale, qui lui permettent de lier l'ADN ainsi que diverses protéines (Figure 33). Mais elle est aussi composée d'un domaine N-terminal impliqué dans l'activation transcriptionnelle ainsi qu'un domaine regroupant le milieu de la protéine et les doigts de zinc fonctionnant comme répresseur (Figure 31).

Comme semblerait l'indiquer sa structure, YY1 est impliquée dans l'activation et la répression transcriptionnelle, mais aussi dans l'initiation de la transcription (Galvin & Shi, 1997 ; Shi *et al.*, 1997 ; Thomas & Seto, 1999). De plus, cette protéine possède une séquence de liaison à l'ARN au sein de ses doigts de zinc (Ficzycz & Ovsenek, 2002), lui permettant ainsi de participer à l'assemblage de mRNPs (pour « Messenger RiboNucleoprotein Particles ») dans les oocytes (Belak & Ovsenek, 2007).

Comme CTCF, YY1 est aussi sujette à plusieurs types de modifications post-traductionnelles incluant la glycosylation (Hiromura *et al.*, 2003, Ozcan *et al.*, 2010), la phosphorylation (Austen *et al.*, 1997), l'acétylation (Yao *et al.*, 2001) et la polyADP-ribosylation (Oei & Shi, 2001).

Le rôle de la glycosylation est de dissocier le complexe YY1-protéine du rétinoblastome (qui empêche sa fixation) d'une manière cycle-cellulaire dépendante, rendant ainsi YY1 libre d'influer sur la transcription (Hiromura *et al.*, 2003).

La phosphorylation sur cette protéine n'a, à ce jour, été impliqué que dans l'entrave de la fixation de YY1 au promoteur du gène talin2 (Zheng *et al.*, 2010).

L'acétylation de la région centrale de la protéine (riche en glycines et lysines), ajoutée par p300 et PCAF, est nécessaire à une activité répressive maximale, alors que celle ayant lieu au niveau des doigts de zinc (et qui n'est pas déacétylable), sert d'appât aux HDACs résultant en une activité d'histone déacétylase, cela malgré une baisse de l'affinité pour ses sites de fixation engendrée par l'encombrement de ses doigts de zinc (Yao *et al.*, 2001).

Et enfin, la polyADPriboseylation de la protéine semble avoir lieu majoritairement lors de lésions de l'ADN qui activent Parp1, cette activation ayant pour conséquence de modifier post-traductionnellement YY1 qui accuse une baisse de son affinité à ses sites de fixation habituels (Oei & Shi, 2001).

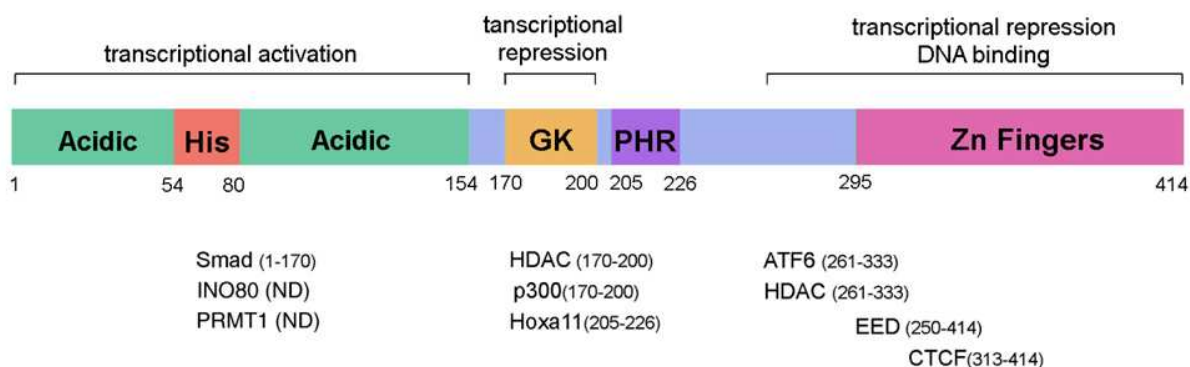


Figure 31 : Diagramme schématisique de YY1 et de ses partenaires. Les domaines d'activation et de répression transcriptionnelle sont indiqués, ainsi que le domaine de liaison à l'ADN. His : domaine riche en histidine, GK : domaine riche en glycine-lysine, PHR : région d'homologie à PHO. Les protéines/complexes associés avec YY1 sont indiquées sous chaque région d'interaction. Entre parenthèses sont les résidus impliqués dans l'interaction. ND : non déterminé. (D'après He & Casaccia-Bonnel, 2008)

Comme décrit dans le tableau ci-dessous, YY1 interagit avec une quantité importante de protéines. Le fait que YY1 soit impliquée dans autant de processus dont le contrôle du cycle cellulaire, l'embryogénèse, les infections virales, l'apoptose ou encore l'oncogénèse, pourrait expliquer qu'une équipe de chercheurs observe que ce facteur est impliqué dans la régulation de près de 10% des gènes chez les mammifères (Schug *et al.*, 2005).

Protéine	Rôle	Référence
c-Myc	Abolit les fonctions activatrices et répressives de YY1	Shrivastava <i>et al.</i> 1993
ARNpolII	Initiation de transcription	Usheva & Shenk 1994
ATF6	Active l'élément de stress du réticulum endoplasmique	Li <i>et al.</i> 2000
CBP	Fonction activatrice réprimée par YY1 au promoteur de AP-1	O'Connor <i>et al.</i> 1996
CTCF	Activateur transcriptionnel de Tsix	Donohoe <i>et al.</i> 2007
E1A	Protéine virale, nécessaire à l'activation transcriptionnelle médiée par YY1 du gène viral AAV P5	Shi <i>et al.</i> 1991
Eed	Induction de tissu neural	Satijn <i>et al.</i> 2001
GATA1	Amplifie l'activation transcriptionnelle de YY1 à l'alpha-globine	Rincon-Arango <i>et al.</i> 2005
Gcn5	Décalage de nucléosomes par hyperacétylation	Klar & Bode, 2005
HDAC1	Répression par déacétylation des histones	Yang <i>et al.</i> 1996, Yang <i>et al.</i> 1997
HDAC2	Répression par déacétylation des histones	Yang <i>et al.</i> 1996, Yang <i>et al.</i> 1997
HDAC3	Répression par déacétylation des histones	Yang <i>et al.</i> 1996, Yang <i>et al.</i> 1997
HoxA11	Réprime la transcription	Luke <i>et al.</i> 2006
Ino80	Permet la fixation de YY1 et aide à	Cai <i>et al.</i> 2007, Wu <i>et</i>

	l'activation transcriptionnelle, essentiel à la réparation de l'ADN basée sur la recombinaison homologue	<i>al.</i> 2007
p300	Répresseur transcriptionnel lorsqu'interagissant avec YY1	Lee <i>et al.</i> 1992
Prmt1	Induit H4R3me aux promoteurs activés par YY1	Rezai-Zadeh <i>et al.</i> 2003
Retinoblastoma protein	Inhibe la fonction d'initiation de transcription de YY1 en bloquant sa fixation à l'ADN	Petkova <i>et al.</i> 2001
Smad	YY1 réprime la fonction d'activateur transcriptionnel de Smad	Kurisaki <i>et al.</i> 2003
Sox2	Activateur transcriptionnel de Xite	Donohoe <i>et al.</i> 2009
Sp1	Amplifie la capacité de YY1 à initier la transcription	Seto <i>et al.</i> 1993, Lee <i>et al.</i> 1994
Suz12	Recruté par YY1 pour réprimer CEBPD	Ko <i>et al.</i> 2008
TFIIB	Initiation de transcription	Usheva & Shenk 1994
YAF2	Sert d'intermédiaire entre YY1 et les protéines du groupe PcG responsables de répression transcriptionnelle	Wilkinson <i>et al.</i> 2010
YY1AP	Amplifie l'activation transcriptionnelle de YY1	Wang <i>et al.</i> 2004

Figure 32 : Tableau récapitulatif des différents interacteurs de YY1 et le but de leur coopération.

i. Protéines PcG

Les protéines du groupe polycomb (PcG) sont responsables du maintien de la répression transcriptionnelle de gènes importants pour le développement. Cependant, les mécanismes de leur recrutement à des sites spécifiques de la chromatine ne sont pas clairs. YY1 est une des rares protéines PcG à posséder un domaine de liaison à l'ADN (ses doigts de zinc). De plus, YY1 possède un domaine baptisé REPO (pour « REcruitment of POLycomb ») correspondant au domaine PHR (pour « PHO Homology Region ») visible sur le schéma ci-dessus (Figure 31), qui permet justement de recruter les protéines notamment responsables du dépôt de la marque H3K27me3 : le complexe PRC2. Ce recrutement s'effectue par l'intermédiaire d'une autre protéine PcG : la protéine Yaf2 (Wilkinson *et al.*, 2010).

Ces informations sont particulièrement intéressantes par rapport aux données de Zhao *et al.* qui observe le recrutement de PRC2 sur un petit transcrit émanant de la région du *Repeat-A* de *Xist* et qui servirait à marquer d'abord le promoteur de celui-ci avant de se propager le long de l'ensemble du chromosome X inactif (Zhao *et al.*, 2008). Dans cette situation, YY1 pourrait jouer le même rôle de recruteur que pour l'ADN, mais ici, sur l'ARN grâce à son domaine de liaison connu pour engendrer la formation de mRNPs dans les oocytes (Belak & Ovsenek, 2007). Il serait donc d'un intérêt certain de réaliser des expériences de RNA-ChIP avec des

anticorps anti-YY1 et anti-Yaf2. Cependant, le complexe PRC2 a aussi été observé se liant à l'ARN non-codant (ncARN) *HOTAIR*. Les auteurs n'observent cependant pas d'implication de YY1 dans la liaison entre les deux protagonistes et proposent plutôt une fixation directe de PRC2 au ncARN par l'intermédiaire du doigt de zinc unique de Suz12 (Rinn *et al.*, 2007).

ii. Formation de structures tridimensionnelles

De manière similaire à CTCF, YY1 semble aussi être impliquée dans les structures tridimensionnelles de l'ADN. En effet, en 1995, l'équipe de Gary Stein s'est rendue compte que la protéine NMP-1 (pour « *Nuclear Matrix Protein 1* ») n'était autre que YY1 (Guo *et al.*, 1995). NMP-1 avait préalablement été décrite comme un facteur de liaison entre l'ADN et la matrice nucléaire (Dworetzky *et al.*, 1992). De plus, des études d'immunofluorescence ont montré qu'une fraction de YY1 est localisée à la matrice nucléaire, principalement mais pas exclusivement associée aux nucléoles. Le fait que YY1 puisse s'attacher aux MARs (pour « *Matrix Attachment Region* » qui sont des séquences de chromosomes qui s'attachent à la matrice nucléaire) nous fait soupçonner un rôle de cette protéine dans la médiation d'interactions gène-matrice, la matrice nucléaire pouvant compartimenter fonctionnellement le noyau eucaryote afin d'assister la régulation de l'expression génique.

Une autre équipe propose que la fixation de complexes protéiques au niveau des promoteurs, des *enhancers* et des « *silencers* » engendre des perturbations du cheminement de la double-hélice d'ADN et suppose que YY1 pourrait participer à un tel phénomène (Natesan & Gilman, 1993). Pour corroborer leur théorie, ils démontrent que la fixation de YY1 au promoteur de *c-fos* induit une courbure de l'ADN de près de 78°, pouvant favoriser ou défavoriser certaines interactions de ce dernier avec des protéines avoisinantes. L'orientation même du site de fixation de YY1 pourrait influencer sur un tel phénomène, sa fonction passant de répressive à activatrice lors de l'inversion de son site de fixation, toujours au promoteur de *c-fos*. L'importance capitale du sens du site de fixation est confirmée par Kim *et al.* concernant des sites de liaison de YY1 à son propre promoteur (Kim *et al.*, 2009).

Ce phénomène de courbure de l'ADN ainsi que les nombreuses interactions de YY1 avec des protéines modificatrices d'histones pourraient aussi expliquer son importance dans la contraction des loci des immunoglobulines dans les cellules progénitrices des lymphocytes B. Lors de la déplétion de la cellule en YY1, on assiste à une absence de contraction chromatinienne du locus IgH, phénomène qui est responsable de la recombinaison des gènes Ig afin de créer de nouveaux anticorps (Liu *et al.*, 2007)

iii. Rôle dans la méthylation de l'ADN

Il est aussi nécessaire de rappeler le rôle probable de YY1 dans la méthylation des DMRs au sein des cellules germinales, notamment son importance dans la méthylation de *Xist* au cours de l'oogénèse mais aussi dans la protection contre cette méthylation à la même DMR durant la spermatogénèse (voir le chapitre : « Régulation en *trans* de *Xist* par différents facteurs », p92).

Comme nous venons de le voir, les deux protéines candidates proposées pour le rôle d'activateurs transcriptionnels de *Xist* ont chacune un grand nombre de fonctions, celles-ci étant médiées grâce à une multitude d'interacteurs. Il s'avère donc que ces protéines fonctionnent rarement seules et sont essentielles à bon nombre de processus cellulaires. Au cours de mon doctorat, j'ai donc tenté de caractériser le rôle de CTCF et YY1 dans la régulation du gène *Xist*, notamment en confirmant que leur fixation corrèle bien avec une fonction d'activateur, tout en tentant d'impliquer des partenaires de ces deux protéines dans le même processus à travers la recherche de leur fixation le long du locus *Xist/Tsix*.

c. Résultats

i. Fixation de CTCF le long du locus *Xist/Tsix*

Suite aux données publiées par Hendrich, Chao, Pugacheva, Kim et Navarro *et al.* (Hendrich *et al.*, 1997, Chao *et al.*, 2002, Pugacheva *et al.*, 2005, Kim *et al.*, 2006, Navarro *et al.*, 2006, voir le chapitre : « Régulation en *trans* de *Xist* par différents facteurs », p92 et la Figure 33), nous avons décidé, à l'aide de la technique d'immunoprécipitation de chromatine suivie d'une analyse par PCR-quantitative (voir les Annexes 1 et 2, p268-269), caractériser le profil de fixation de CTCF le long du locus *Xist/Tsix* dans différents types cellulaires qui reflètent les différentes étapes de l'inactivation du chromosome X.

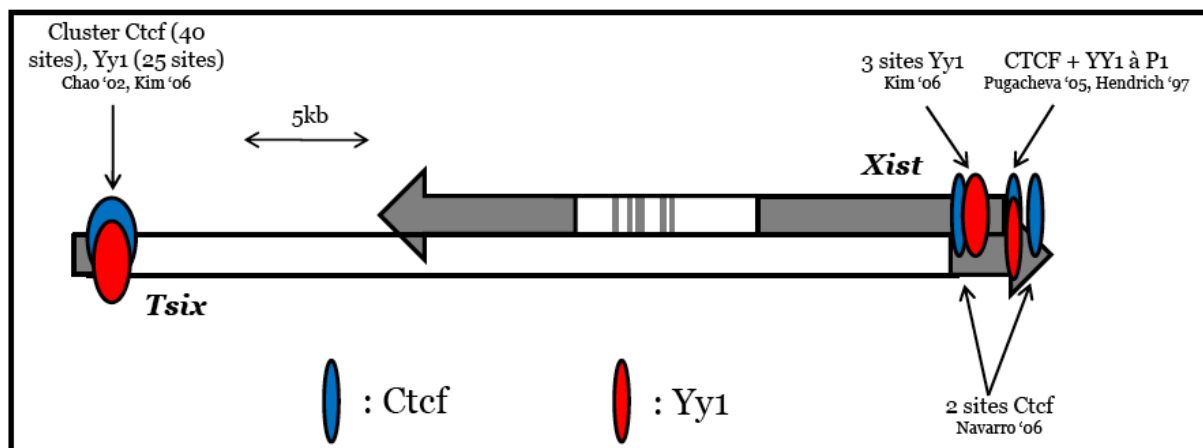


Figure 33 : Récapitulatif des sites de fixation de CTCF et de YY1 le long du locus *Xist/Tsix* déjà publiés. La largeur des ovales représentant les protéines est indicatif du nombre de sites de fixation.

Dans un premier temps, nous avons regardé la fixation de CTCF au sein de cellules souches embryonnaires murines femelles (cellules LF2, de la lignée *129*) ainsi qu'au sein de cellules fibroblastiques femelles (aussi de la lignée *129*) ayant été isolées à partir d'embryons à 13,5jpc, récapitulant ainsi les deux statuts "extrêmes" de l'inactivation (l'activité transcriptionnelle des deux X dans les ES, et l'inactivation aléatoire d'un des deux X maintenue depuis plusieurs divisions cellulaires au sein des MEFs).

Au sein des cellules ES (Figure 34B), nous retrouvons deux pics de part et d'autre du promoteur P1 de *Xist* décrit par Navarro *et al.* (Navarro *et al.*, 2006). Cependant, nous observons aussi plusieurs pics au niveau du locus *Xist/Tsix* dont certains n'ont jamais été décrits et semblent être placés à des positions "stratégiques" par rapport à la séquence de *Xist* et de son antisens : recouvrant d'une part les petits exons 2,3, 4 et 5, et la fin du transcrit de *Xist*, et d'autre part, une région en amont du promoteur de *Tsix*. CTCF se fixe aussi plus faiblement au sein de l'exon 1. Toutefois, nous ne voyons aucune fixation de CTCF sur le promoteur même comme décrit par Pugacheva *et al.*, ce qui pourrait être dû au fait que ces observations avaient été faites dans des cellules somatiques et non pas ES ou bien à la différence entre les anticorps utilisés. Curieusement, le pic de CTCF attendu à *DxPas34* (Chao *et al.*, 2002, Donohoe *et al.*, 2007) est très faible ; il est toutefois possible que l'enrichissement au sein du minisatellite soit plus important que ce qui est représenté, sachant que de par la nature répétitive de la région, il est impossible de dessiner des amorces en son centre, nous limitant à l'étude de ses extrémités.

De manière intéressante, dans les cellules fibroblastiques femelles (Figure 34C), chez qui un chromosome X est inactivé, on assiste à une diminution voire une disparition de 6 des 8 pics

de fixation. Les seuls pics persistants sont ceux qui entourent le promoteur P1, et l'on observe même une apparition de pic au sein de l'exon 1 (fixation maximale décalée par rapport aux cellules ES). Il est intéressant de se demander si les modifications de fixation de CTCF entre cellules indifférenciées et différenciées à ce locus sont fonctionnellement reliées à la disparition de l'expression biallélique de *Tsix* ou bien à la surexpression monoallélique de *Xist*. Par contre, contrairement aux résultats de Pugacheva *et al.*, et confirmant ce qui a été observé dans les ES (Navarro *et al.*, 2006), nous n'observons toujours pas de fixation au niveau de P1, bien qu'il y ait un léger décrochage (propre aux MEFs) du profil à cet endroit.

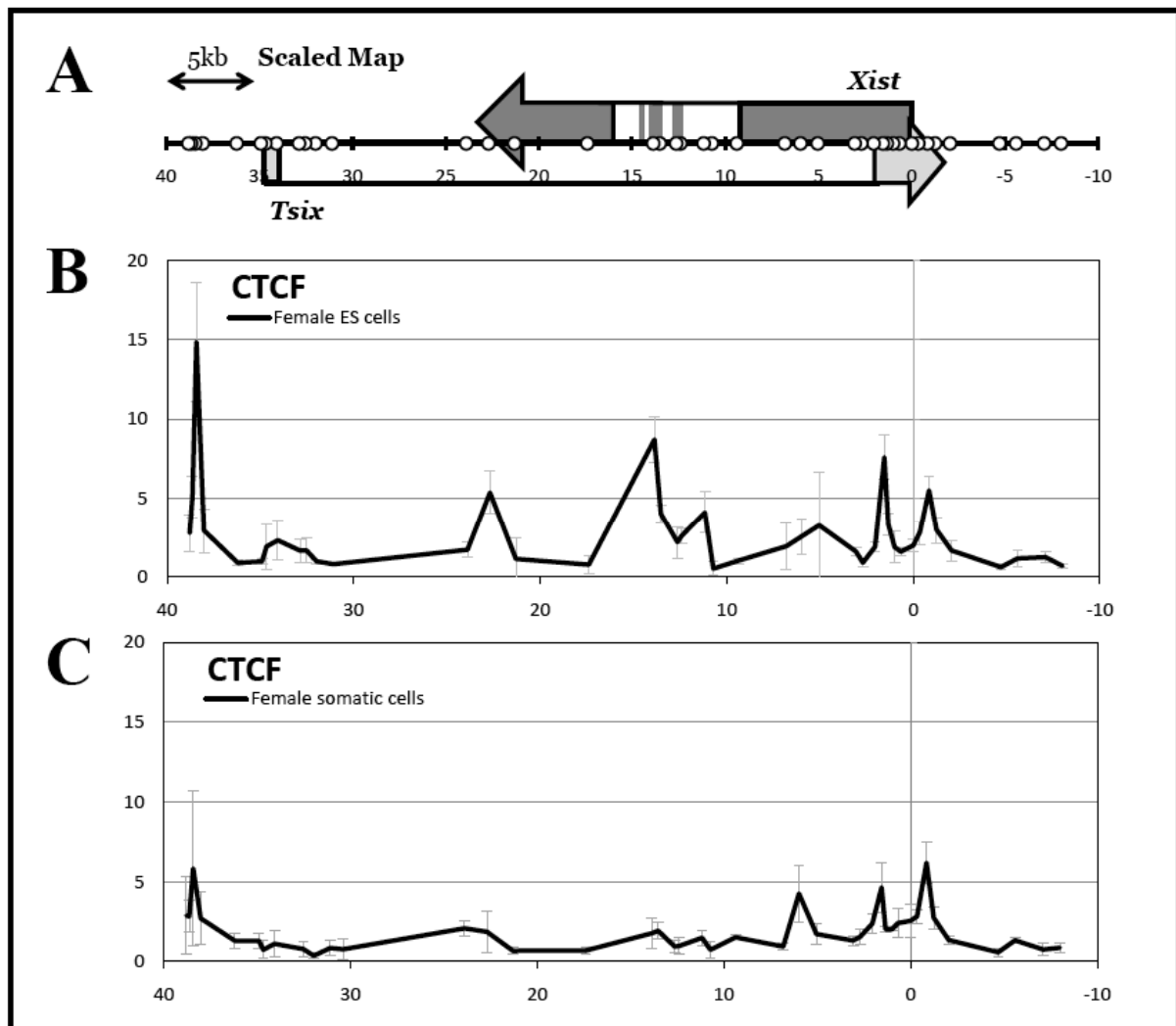


Figure 34 : Distribution de CTCF dans une région de 50kb du XIC murin au sein de cellules ES et fibroblastiques femelles. (A) Diagramme schématique des 50kb analysés par des expériences de ChIP ($n > 3$) montrant les différentes unités de transcription du locus *Xist/Tsix*. Les flèches indiquent le sens transcriptionnel. En gris foncé : les exons de *Xist*, en gris clair : les exons de *Tsix*. À chaque expérience de ChIP, nous avons utilisé 42 paires d'amorces éparpillées le long de la région. (B) Analyse de ChIP de CTCF dans des cellules ES femelles indifférenciées et (C) dans des cellules femelles fibroblastiques. Les deux graphiques montrent le facteur d'enrichissement des immunoprécipitations (après normalisation par rapport à l'input) par rapport au bruit de fond déterminé en moyennant la valeur de 10 positions, au sein de la région étudiée, montrant une absence de fixation. L'origine de l'axe des ordonnées correspondant au +1 de transcription du gène *Xist*.

Afin de savoir si le profil de fixation de CTCF est identique sur les deux Xs actifs dans les cellules indifférenciées et si les modifications de profil observées dans les cellules somatiques sont spécifiques au Xa ou au Xi, nous avons entrepris de réaliser la même expérience mais dans des cellules mâles, ES (CK35) et fibroblastes (lignée 129).

Le profil de fixation de CTCF et le niveau d'enrichissement dans les cellules ES mâles (Figure 35A) sont identiques à ceux qui sont observés chez les femelles, nous indiquant qu'*a priori* chez la femelle les deux allèles possèdent le même schéma de fixation. En effet dans le cas contraire nous aurions observé des différences d'intensité de la fixation de la protéine par rapport aux mâles (sachant que les données de ChIP sont toujours normalisées par rapport à l'ADN génomique non immunoprécipité). La seule différence est une meilleure définition du pic à *DxPas34* mais qui demeure toujours aussi faible. Les similitudes observées entre les deux sexes pourraient refléter la nécessité des deux allèles (aussi bien le futur X actif qu'inactif) d'être sur un pied d'égalité afin que le choix du chromosome à inactiver ne soit pas biaisé lors de l'induction de la différenciation.

Dans les cellules fibroblastiques mâles la quasi-totalité du profil est identique à celui des femelles. Il existe cependant une différence : la disparition de la fixation de CTCF près du promoteur P2 (Figure 35B). Cette différence pourrait être la cause ou la conséquence de l'expression de *Xist* à partir du chromosome X inactif chez les femelles, sachant que CTCF ne se fixe pas à ce site sur le X actif.

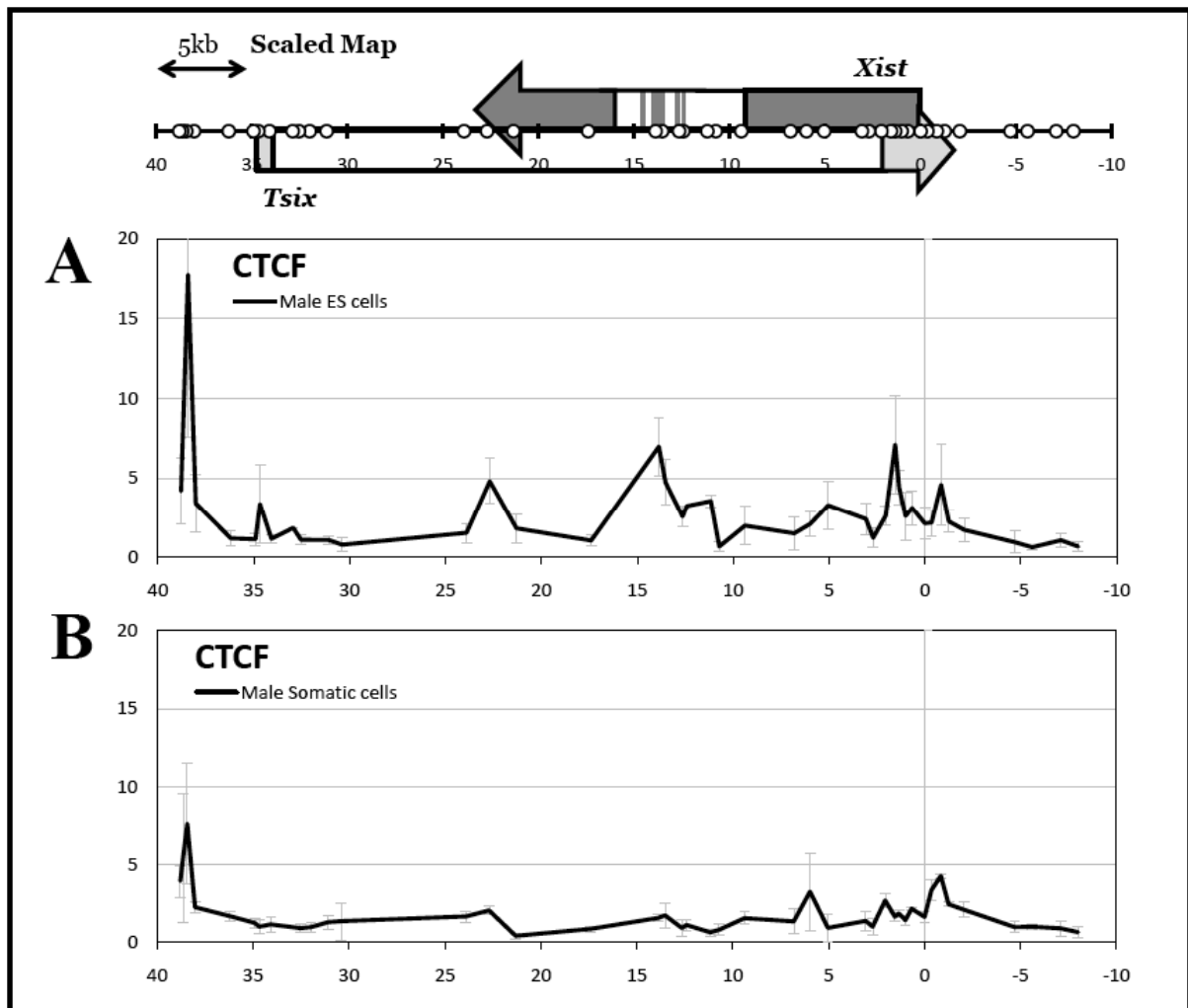


Figure 35 : Étude de la fixation de CTCF le long du locus *Xist*/*Tsix* dans des cellules ES et fibroblastiques mâles. (A) La fixation de CTCF le long du locus *Xist*/*Tsix* dans les cellules ES mâles indifférenciées est très similaire à ce qui est observé pour les cellules femelles du même stade. **(B)** Au sein des cellules mâles somatiques, on remarque principalement la disparition du premier pic en aval du promoteur de *Xist*.

Le profil de fixation de CTCF au promoteur P2 corrélant avec un rôle hypothétique dans l'activation transcriptionnelle de *Xist* (bien qu'il soit déjà fixé à ce site dans les ES, sans qu'il y ait expression de *Xist*), nous avons voulu savoir ce qu'il en était pour YY1, facteur transcriptionnel ayant été décrit comme se fixant aussi autour de P1 et de P2 afin d'activer l'expression de *Xist* (Hendrich *et al.*, 1997, Kim *et al.*, 2006) (Figure 33).

ii. Fixation de YY1 en 5' de *Xist*

Nous observons une fixation de YY1 entre 1 et 1,5kb en aval du promoteur P1 de *Xist* dans les cellules ES femelles (Figure 36B), des résultats similaires ont depuis été publiés par Kim *et al.* (Kim *et al.*, 2006). Malheureusement, de par la proximité importante des trois sites de fixation potentiels (observés par Kim *et al.*), nous avons été dans l'impossibilité de dessiner

des amorces pouvant nous indiquer lesquels sont sollicité par la protéine. On observe également un décalage du pic de fixation de YY1 entre les cellules ES et les cellules somatiques femelles, cela étant peut-être le témoin d'une fixation différentielle à ces sites en fonction du contexte cellulaire et/ou de son rôle à ce locus. Ces trois sites se trouvent à une distance comprise entre 20 et 230pb du site CTCF, cette proximité pouvant indiquer une interaction entre les deux protéines (comme cela a été décrit à *DxPas34*, Donohoe *et al.*, 2007).

Ce pic de YY1 est également présent dans les cellules ES mâles (Figure 36C), reflétant une fois de plus une similitude de la fixation des protéines étudiées entre les deux sexes et potentiellement les deux allèles. Et de manière identique au pic de CTCF situé à P2, la fixation de YY1 est abolie au sein des cellules mâles somatiques, cette cinétique de fixation corrélant une fois de plus avec l'expression de *Xist*. De plus, ici encore, nous n'observons pas la fixation de YY1 au promoteur de P1 décrit par Hendrich *et al.* (Hendrich *et al.*, 1997), ceci pouvant dû au fait que ces observations avaient été effectuées chez l'homme, malgré la conservation d'un site consensus (n'étant plus actuel) chez plusieurs espèces, ou bien au fait que leurs expériences ont été réalisées *in vitro*.

Nous ne montrons ici que la partie 5' de *Xist*, le reste du locus *Xist/Tsix* (tel qu'il est montré dans les figures précédentes) n'étant pas occupé par cette protéine, à part à *DxPas34* où il se fixe sur le minisatellite de manière similaire à CTCF, c'est-à-dire plutôt faiblement (données non montrées).

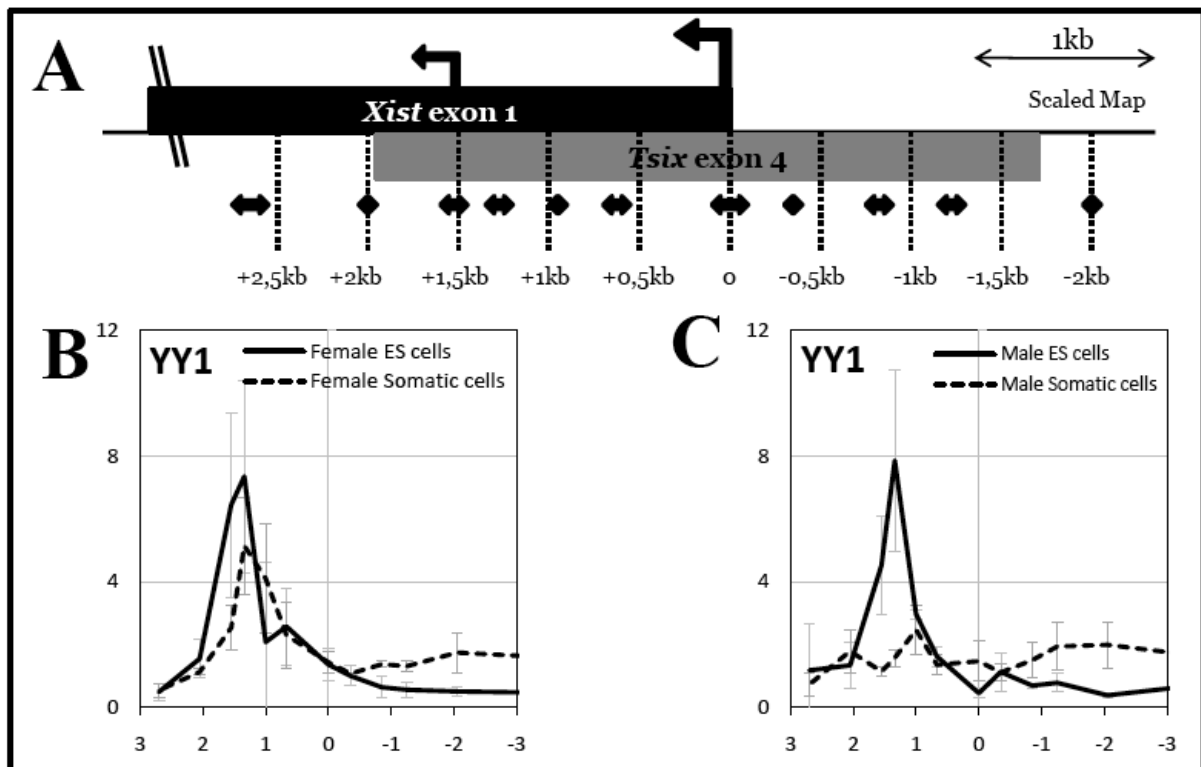


Figure 36 : Étude de la fixation de YY1 dans la région promotrice de *Xist* dans des cellules ES et somatiques mâles et femelles. (A) Représentation schématique de la région promotrice de *Xist*. Le rectangle noir représente l'exon 1 de *Xist*, le gris, l'exon 4 de *Tsix*. Les deux flèches unidirectionnelles représentent les deux promoteurs P1 et P2 de *Xist*. Les flèches bidirectionnelles représentent les régions amplifiées lors de la PCR-Quantitave. Le point 0 est toujours le +1 transcriptionnel de *Xist*. (B) En trait plein est représentée la fixation de YY1 dans les cellules ES femelles indifférenciées et en pointillés dans des cellules somatiques femelles. (C) Idem à (B) mais chez le mâle.

iii. Corrélation entre fixation de CTCF et YY1 et surexpression de *Xist* en cours de différenciation

Afin de savoir si les différences de fixation de ces protéines au promoteur P2 de *Xist* sont bien le reflet de l'activation transcriptionnelle de ce dernier, nous avons entrepris de regarder la fixation de CTCF et YY1 au niveau du X actif et du X inactif, en utilisant notamment différentes lignées de cellules ES mâles mutantes, notamment les lignées Ma2L, Δ DxPas34 et Δ 65 (Luikenhuis *et al.*, 2001, Vigneau *et al.*, 2006, Morey *et al.*, 2004) (Figure 37).

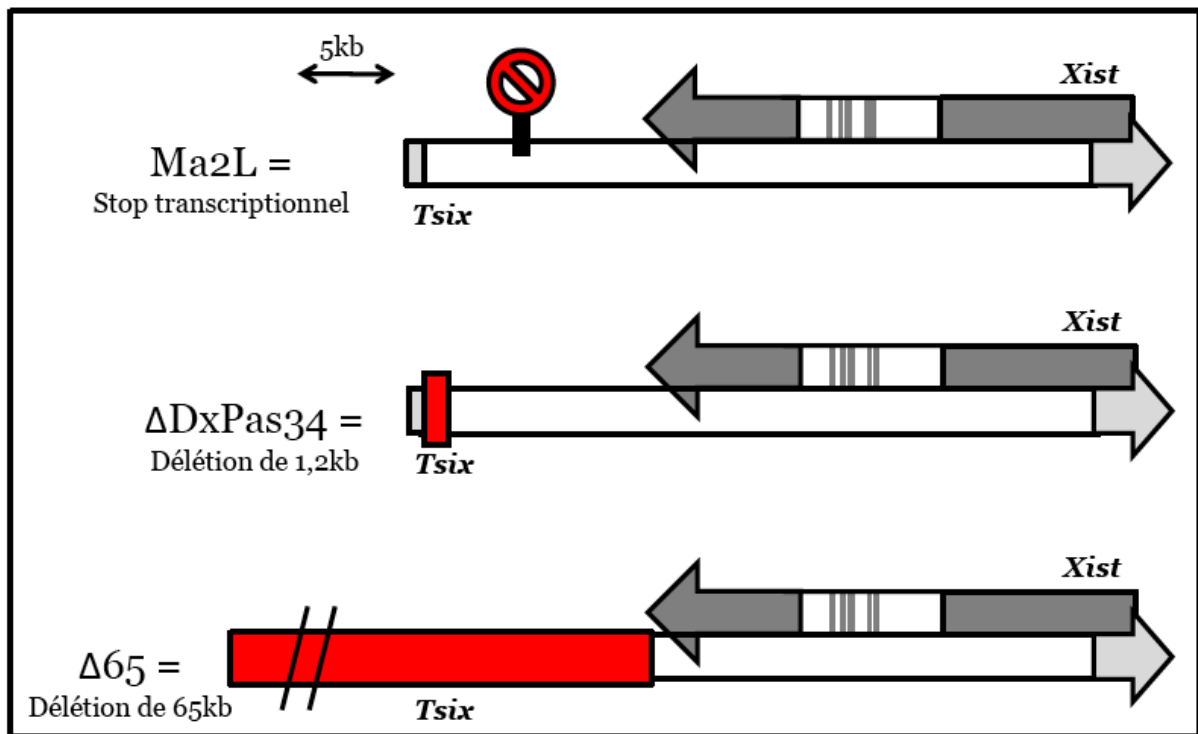


Figure 37 : Schéma Récapitulatif des différentes mutations génétiques présentes au sein de cellules mâles qui ont servi à l'étude. Dans Ma2L, Luikenhuis *et al.* ont inséré une séquence d'arrêt transcriptionnel à environ 4kb en aval du promoteur de *Tsix* (Luikenhuis *et al.*, 2001), empêchant ainsi la machinerie de transcription de traverser la région promotrice de *Xist*. Δ DxPas34 possède une délétion d'environ 1,2kb du microsatellite jouant le rôle d'amplificateur de *Tsix*, ceci réduit l'activité transcriptionnelle de l'antisens d'environ 90% (Vigneau *et al.*, 2006). Δ 65 représente une lignée de cellules ES mâles dans lesquels il manque un fragment de 65kb à partir de la fin du transcrit de *Xist* et dans la direction de *Tsix*, abolissant toute transcription à partir de *Tsix* ou de *Xite* (entre autres) (Morey *et al.*, 2005). En rouge sont représentées les séquences altérées.

Ces lignées ont toutes la particularité d'affecter la transcription de *Tsix*, soit en empêchant sa transcription à travers la région codante de *Xist* (Ma2L), soit en réduisant sa transcription (d'environ 90%) par délétion d'une région stimulatrice (Δ DxPas34) ou encore en supprimant ses promoteurs (majeur et mineur) (Δ 65). La particularité de ces cellules est qu'elles sont toutes capables d'induire l'inactivation ectopique de l'unique chromosome X lorsque la différenciation est induite (Figure 38), nous permettant ainsi d'observer quelles sont les caractéristiques de fixation de nos protéines d'intérêt exclusivement sur le X inactif. Les cellules ES males sauvages, quant à elles, en particulier en cours de différenciation, permettent d'étudier la fixation de CTCF sur le X actif (ou sur le promoteur inactif de *Xist*).

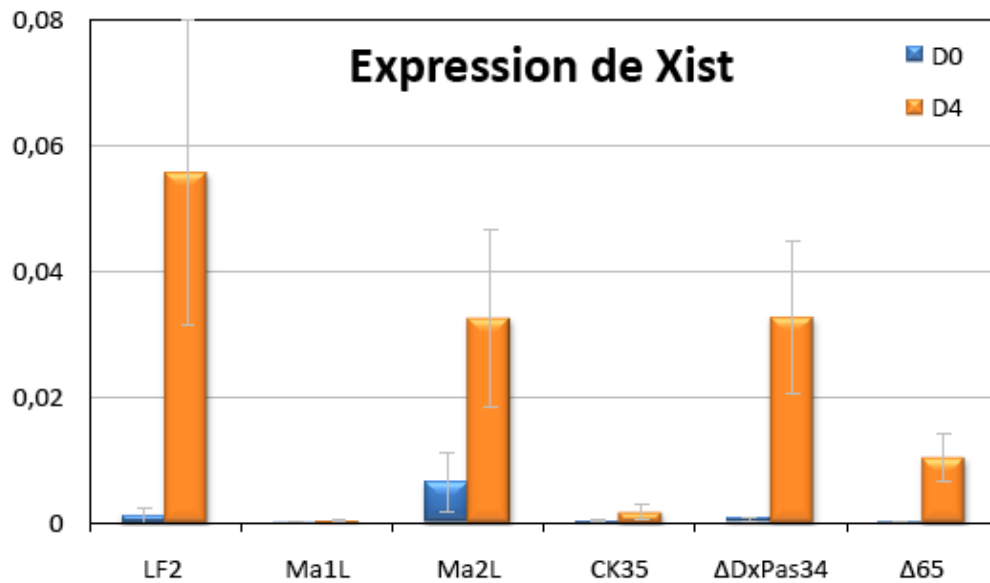


Figure 38 : Expression de *Xist* au sein des différentes lignées cellulaires étudiées. Ces résultats ont été générés suite à la reverse-transcription des ARNs suivie d’une analyse par PCR-quantitative. Tous les résultats ont été normalisés par rapport à l’expression du gène *Arpo*. En bleu les cellules ES indifférenciées et en orange, après 4 jours de différenciation. LF2 : cellules ES femelles sauvages, CK35 : cellules ES mâles sauvages.

La fixation de CTCF et YY1 dans les cellules ES mâles contrôles Ma1L (contenant simplement un site *Lox*, correspondant au site d’excision du « *stop* » transcriptionnel) est identique à celle des cellules ES mâles CK35 (Figure 39A et B). De plus, nous pouvons remarquer que cette fixation est abolie dès le quatrième jour de différenciation. Ainsi, le détachement définitif des protéines a lieu dans une fenêtre de temps qui correspond à celle de l’initiation de l’inactivation chez les femelles. À l’inverse, chez les cellules mutantes (ici Ma2L) en cours de différenciation, on observe un maintien, voire une augmentation, de la fixation des deux protéines au site en aval de P1 (Figure 39C et D). Ces résultats nous indiquent que la fixation de CTCF et YY1 en aval de P1 se fait de manière allélique en cours de différenciation cellulaire, et que cette fixation allélique corrèle à une surexpression de *Xist*, le Xa perdant la fixation des protéines, tandis que le Xi qui surexprime *Xist*, la conserve. L’expression de *Tsix* ne semble pas avoir d’effet sur la fixation de ces protéines puisqu’à l’état indifférencié, CTCF et YY1 sont fixés autour du promoteur P2 indépendamment de l’expression de *Tsix* (et quelque soit le mutant observé). Il est important de noter que nous observons, ici encore, à un décalage du pic de YY1 sur le (futur) Xi en cours de différenciation (Figure 39D), nous renforçant dans l’idée d’une utilisation différentielle des 3 sites de fixation en cours de différenciation et peut-être lors de l’activation de *Xist*.

Une fois de plus, la fixation des deux protéines sur le reste du locus *Xist/Tsix* n'est pas montrée ici, aucune des 3 mutations ne modifiant quoique ce soit au reste du profil.

Cependant, j'aimerais signaler la particularité que seules les cellules Ma2L à jour 0 subissent une augmentation notable de *Xist* par rapport aux cellules contrôles (Figure 38), bien que les trois mutations aient pour conséquence de diminuer l'expression de *Tsix*, censé être un répresseur post-traductionnel de *Xist* (Navarro *et al.*, 2006). Cela nous permet de suggérer que la région *DxPas34* (délétée dans $\Delta DxPas34$ et $\Delta 65$ mais pas dans Ma2L) pourrait avoir un rôle dans la surexpression de *Xist* (d'une manière post-traductionnelle étant donné qu'il n'y a pas de recrutement du PIC à P1 chez ses mutants) dans les cellules ES indifférenciées lorsque la transcription de *Tsix* est inhibée. Ce rôle semble disparaître en cours de différenciation puisque les taux d'expression de *Xist* sont très similaires entre Ma2L et $\Delta DxPas34$ à jour 4 d'induction par ajout d'acide rétinolique. Une autre possibilité est que l'insertion de la séquence d'arrêt transcriptionnel entourée de deux sites Lox pourrait contribuer à la perturbation d'un élément impliqué dans la répression de *Xist*, cette hypothèse paraissant bien moins probable de par la petite taille de l'insertion ainsi que l'inexistence d'éléments régulateurs de quelque espèce que ce soit autour de la région dans laquelle elle se trouve.

Une fois une corrélation indéniable établie entre la fixation des protéines CTCF et YY1 autour du promoteur P2 de *Xist* et la surexpression de celui-ci, nous avons souhaité nous intéresser aux autres pics de CTCF le long du locus *Xist/Tsix*, notamment en cherchant à leur attribuer une fonction, leur fixation se rapportant plutôt au développement cellulaire qu'à l'expression de *Xist* ou de son anti-sense.

iv. Un rôle pour les autres pics de CTCF ?

Étant donné la distribution des pics de CTCF le long du locus *Xist/Tsix* et les nombreux articles impliquant cette protéine dans la structuration tridimensionnelle de la chromatine, aussi bien lors d'interactions intra- (Li *et al.*, 2008) qu'interchromosomiques (Ling *et al.*, 2006), il semblait fort concevable qu'elle ait un rôle similaire à cet endroit. De plus, il a dernièrement été prouvé que CTCF et la cohésine interagissent à de nombreux sites à travers le génome, ce complexe protéique étant présent dans 55 à 80% des cas aux sites CTCF (Parelho *et al.*, 2008, Wendt *et al.*, 2008).

Afin de tester cette hypothèse, nous avons examiné la distribution de la sous-unité Rad21/Sccl de la cohésine au niveau du locus *Xist/Tsix* dans nos différents modèles cellulaires.

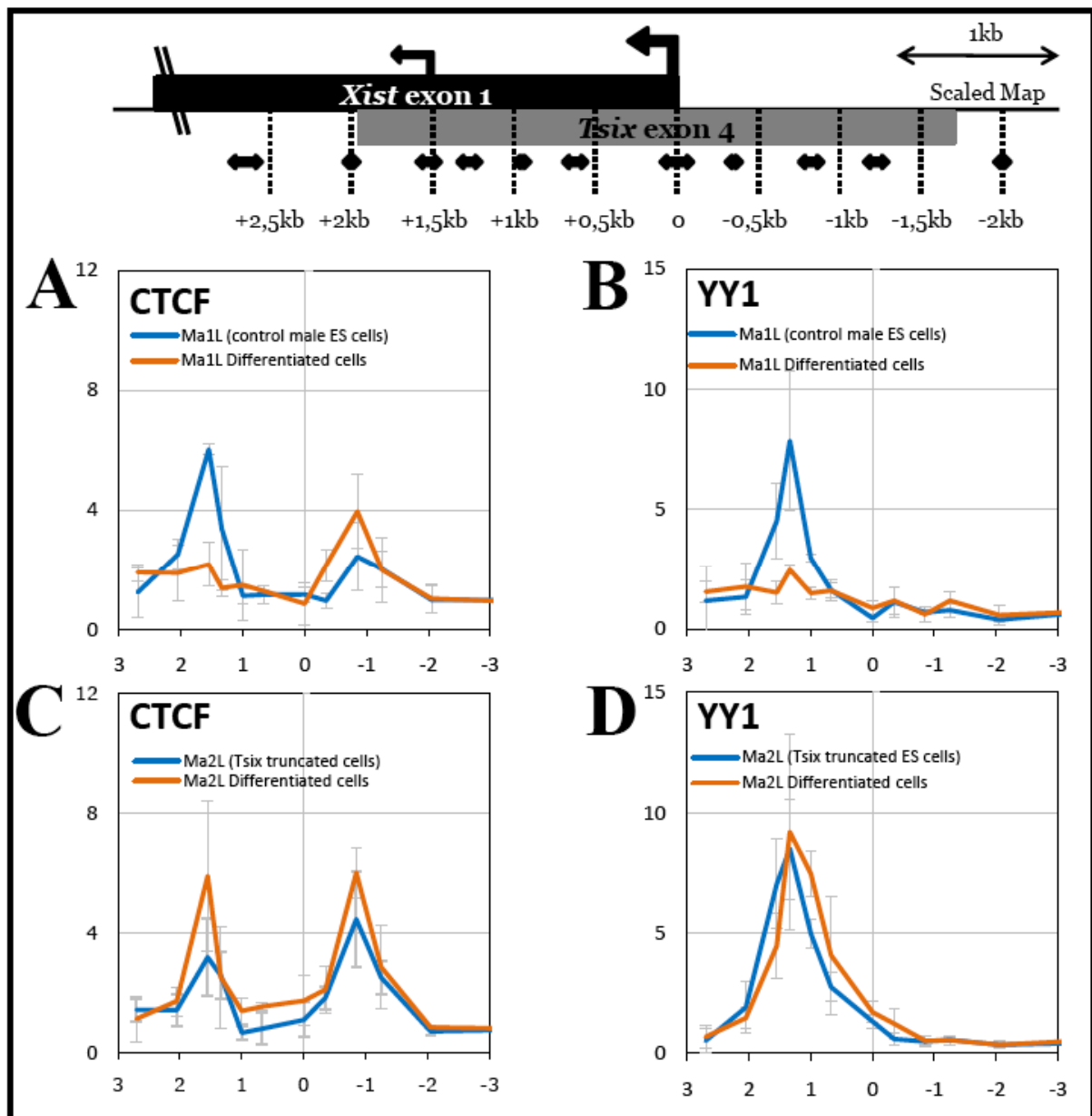


Figure 39 : Fixation de CTCF et de YY1 à la région promotrice sur le Xa ou le Xi en cours de différenciation. Fixation de (A) CTCF et (B) YY1 dans les cellules ES mâles sauvages Ma1L, au stade indifférencié (bleu) ou après 4 jours de différenciation à l'acide rétinoïque (orange). Étude identique de (C) CTCF et (D) YY1 dans les cellules ES mâles Ma2L tronquées pour *Tsix*. En cours de différenciation le chromosome X est inactivé chez Ma2L, tandis qu'il reste actif chez Ma1L.

Deux pics majeurs de Rad21 sont observables sur l'ensemble du locus *Xist/Tsix* dans les cellules ES femelles (Figure 41A) qui corrélient tous deux avec des pics CTCF : un premier en amont de P1 et un second en amont du promoteur majoritaire de *Tsix*. Ce résultat nous suggère que ces deux pics de CTCF/Rad21 pourraient être impliqués dans un mécanisme visant à modifier la structure tridimensionnelle du locus. Un pic de très faible intensité est aussi identifiable, sa localisation correspondant justement à celle du pic de CTCF à P2.

Le profil de fixation de Rad21 dans des MEFs femelles rejoint clairement celui de CTCF, le pic en amont de *Xist* étant ubiquitaire, le pic en amont de *Tsix* disparaissant totalement une fois la différenciation cellulaire terminée comme le fait également le petit pic à P2 (Figure 40B). Cependant, de manière étrange, on assiste à la naissance d'autres petits pics qui étaient majoritairement inexistants dans les ES. On ignore comment s'opère la fixation de ces petits pics, CTCF ayant totalement abandonné ces régions après différenciation.

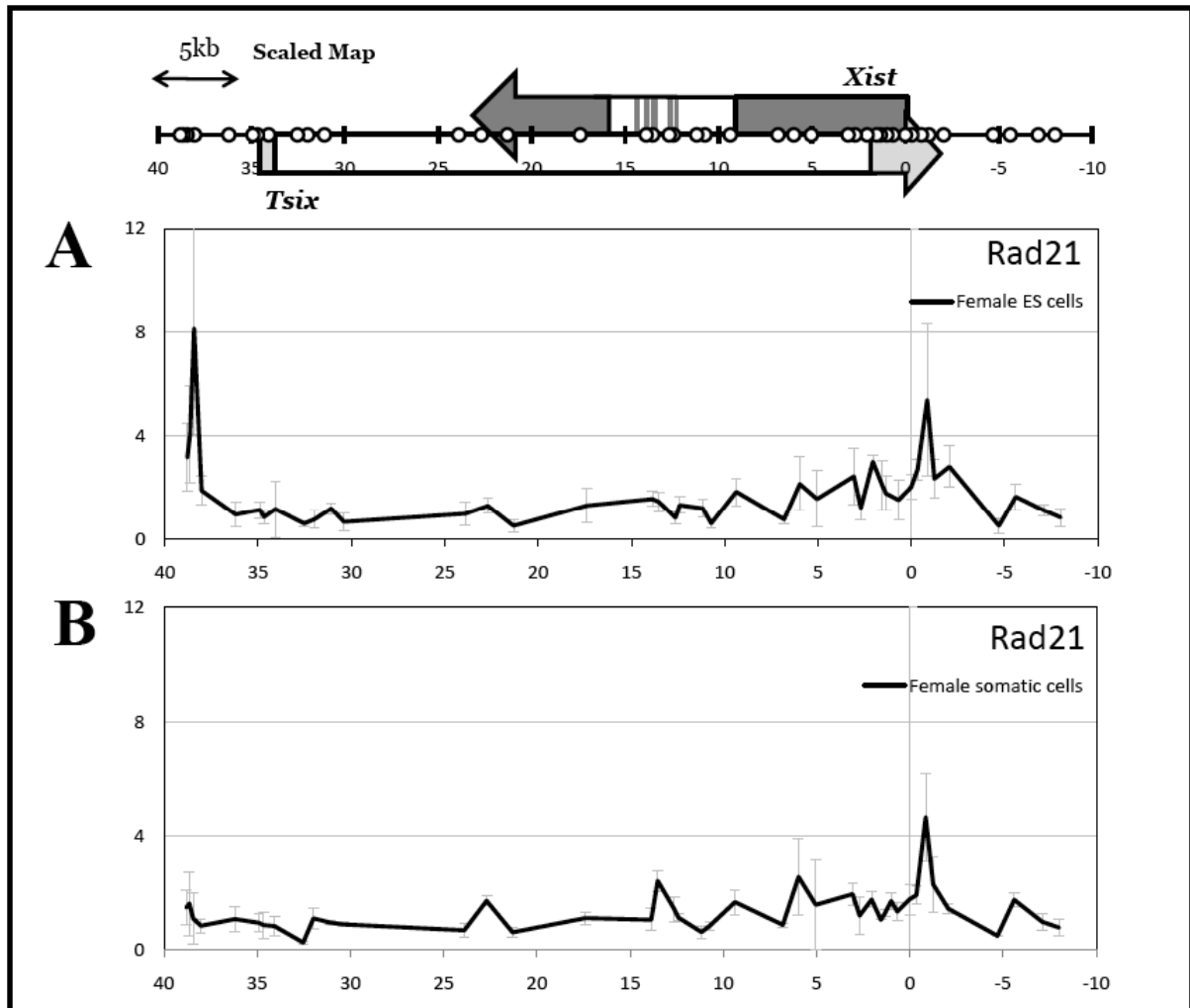


Figure 40 : Étude de la fixation de Rad21 le long du locus *Xist/Tsix* dans des cellules ES et fibroblastiques femelles. (A) Fixation de Rad21 le long du locus *Xist/Tsix* dans les cellules ES femelles indifférenciées. (B) Fixation de Rad21 le long du locus *Xist/Tsix* dans les cellules fibroblastiques femelles.

Au sein des cellules mâles, le profil est sensiblement similaire à celui des femelles (Figure 41). Dans les cellules ES, les différences majeures se situent autour du promoteur de *Xist*. En effet, le pic en amont de P1 est plutôt faible comparé à celui chez la femme, tandis que le petit pic à P2 semble mieux défini et plus grand (Figure 41A). Il est possible que ces changements

reflètent une différence allélique de marquage par Rad21, ceci n'ayant pas été observé pour CTCF. Tandis que dans les MEFs mâles, le même schéma se met en place que chez les cellules somatiques femelles : maintien du pic en 5' de *Xist* et disparition de celui en 5' de *Tsix* (Figure 41fB). Cependant, les petits pics n'apparaissent plus ici, indiquant soit leur origine artéfactuelle, soit le fait qu'ils se trouvent surement enrichis en l'absence de fixation de CTCF et uniquement sur le Xi, ce qui paraît peu probable. Le fait que la fixation de CTCF et Rad21 n'ait plus lieu en 5' de *Tsix* dans les cellules somatiques nous renseigne sur sa fenêtre temporelle d'action et nous indique qu'elle n'est pas nécessaire à une expression maintenue de *Xist* au cours de la vie cellulaire.

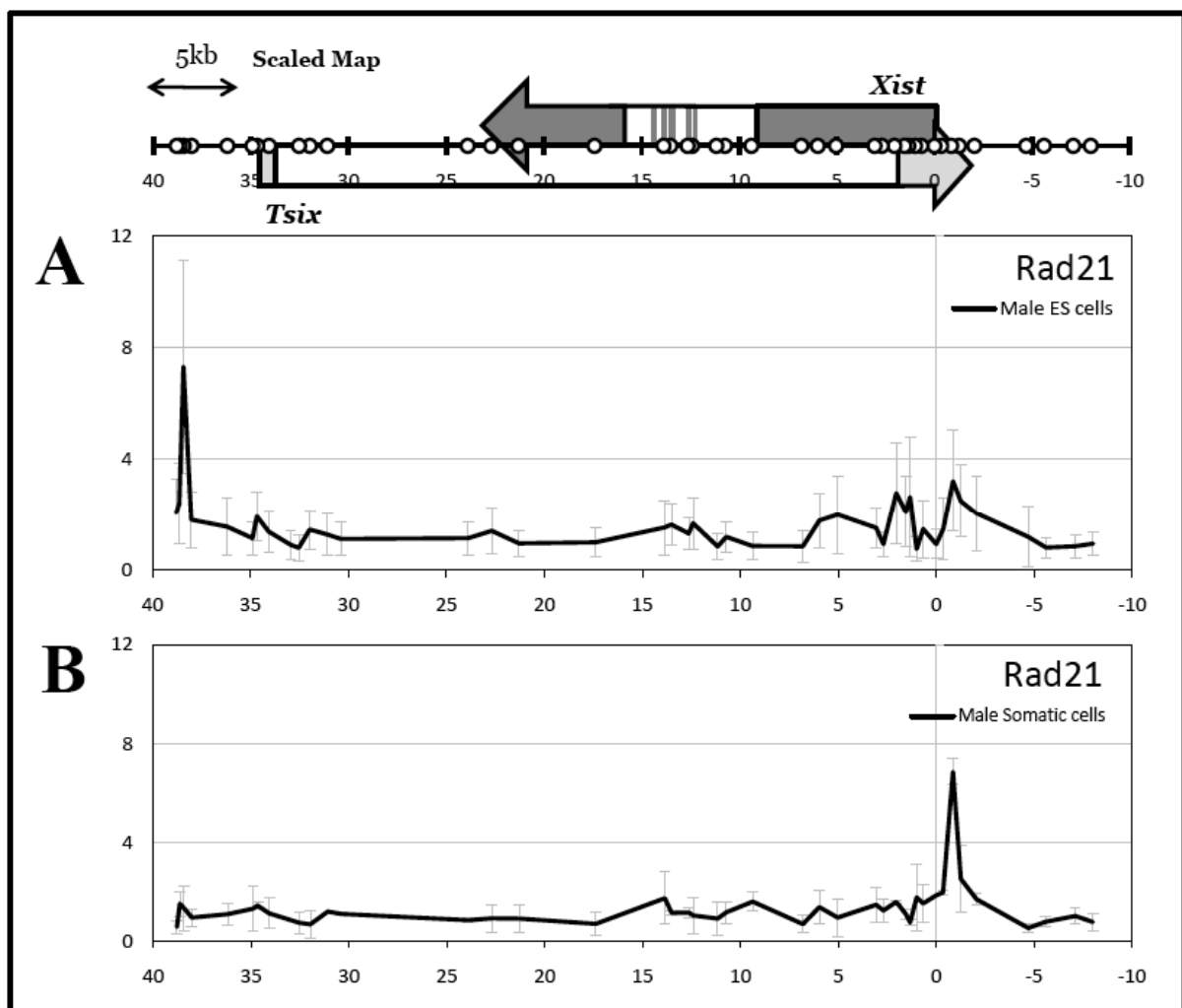


Figure 41 : Étude de la fixation de Rad21 le long du locus *Xist/Tsix* dans des cellules ES et fibroblastiques mâles. (A) Fixation de Rad21 le long du locus *Xist/Tsix* dans les cellules ES mâles indifférenciées. On peut remarquer une baisse de la fixation en amont du promoteur P1 de *Xist* par rapport à chez les femelles. (B) Fixation de Rad21 le long du locus *Xist/Tsix* dans les cellules fibroblastiques mâles. Comme dans les cellules fibroblastiques femelles, la fixation en amont de *Tsix* disparaît dans les cellules somatiques, parallèlement à une augmentation importante du pic en amont de *Xist*.

Nous avons ensuite étudié la fixation de Rad21 dans des cellules portant des mutations de *Tsix* afin d'observer si l'absence de transcription de l'antisens ou bien la délétion du site de fixation de CTCF/Rad21 en 5' de *Tsix* (dans les cellules $\Delta 65$), nous apporteraient des informations sur une nécessité quelconque de la cohésine dans l'inactivation ou bien sur une éventuelle interaction entre les deux pics.

Les profils observés entre Ma1L et Ma2L sont très similaires à ce qui avait été observé chez les ES (Figures 42A et B). L'accentuation de l'intensité du pic de Rad21 en amont de P1 chez Ma2L par rapport à Ma1L nous fournit un début d'explication quant à la différence observée entre les ES mâles et femelles. La même chose est observée pour les cellules $\Delta 65$ et $\Delta Dxpas34$ par rapport aux cellules contrôles CK35 (données non montrées).

Cette similitude nous permet donc une précision concernant le rôle du pic en amont de *Tsix* : cette fixation se fait de manière identique sur les deux allèles quelque soit le type cellulaire étudié. Cependant la fixation de CTCF et Rad21 à ce site disparaît totalement au sein des cellules somatiques, ce qui pourrait correspondre avec un effet éventuel de l'expression de *Tsix*, bien qu'à jour 4 l'expression de celui-ci a déjà considérablement diminué alors que nous n'observons aucun changement de la fixation des protéines au site le plus proche de son promoteur.

Quelque soit la lignée cellulaire étudiée, il semblerait que la fixation de Rad21 dans les ES se fasse de manière plus importante sur le futur allèle inactif que sur le futur actif. Et en regardant les graphiques de la Figure 42A et B, on ne peut s'empêcher de se demander si les différences d'intensités de fixation observées entre Ma1L et Ma2L sont significatives, reflétant peut-être une activité de structuration tridimensionnelle plus importante sur le futur X inactif, bien qu'on ne les observe pas en étudiant la fixation de CTCF (données non montrées).

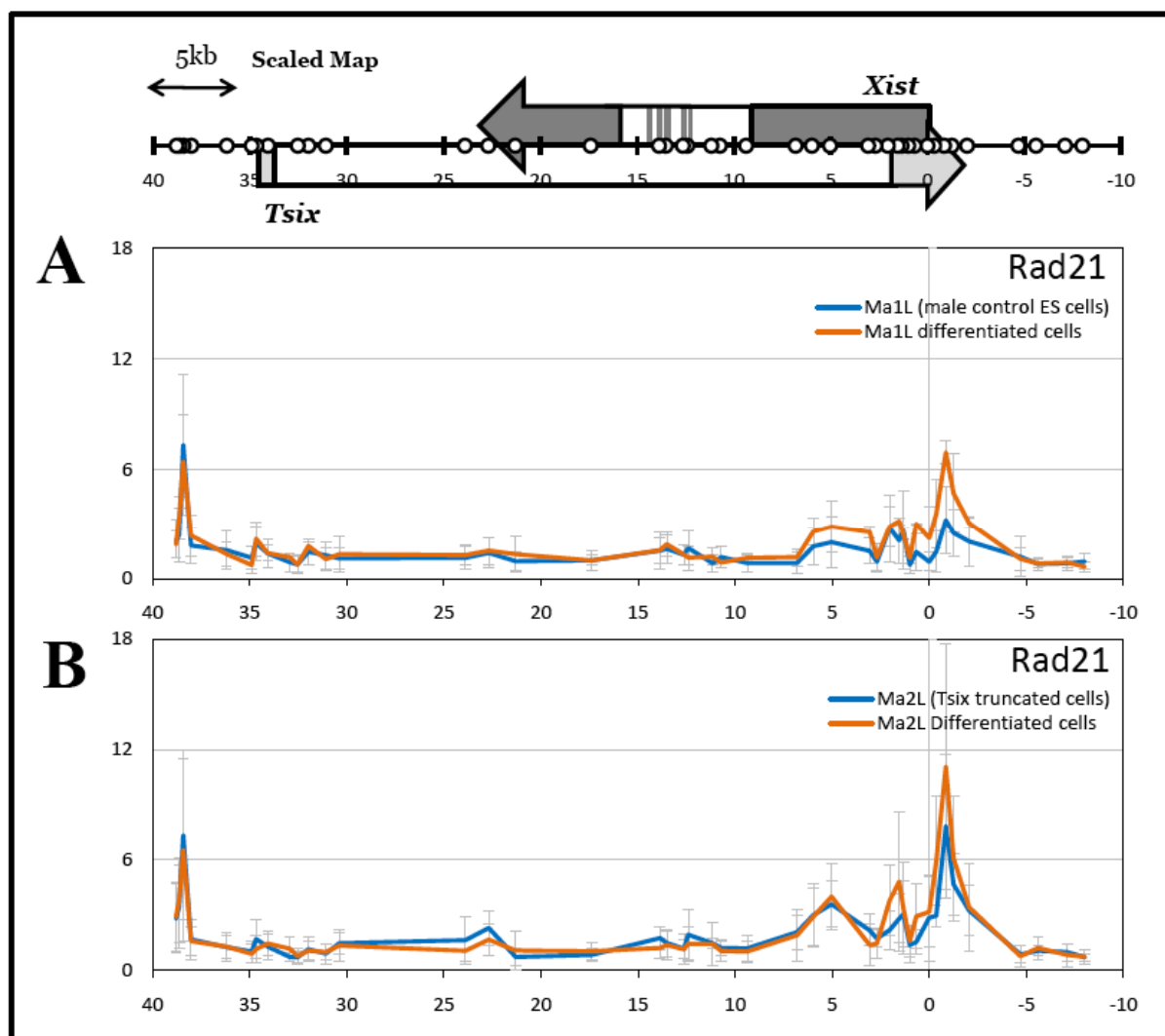


Figure 42 : **Fixation de Rad21 le long du locus *Xist/Tsix* sur le Xa ou le Xi en cours de différenciation.** (A) Fixation de Rad21 dans les cellules ES mâles sauvages Ma1L, au stade indifférencié (bleu) ou après 4 jours de différenciation à l'acide rétinoïque (orange). (B) Étude identique de Rad21 dans les cellules ES mâles Ma2L tronquées pour *Tsix*. En cours de différenciation le chromosome X est inactivé chez Ma2L, tandis qu'il reste actif chez Ma1L.

À travers ce profilage des sites de fixation de CTCF, YY1 et Rad21 à travers le locus *Xist/Tsix*, nous sommes parvenus à montrer une colocalisation de CTCF et YY1 dans la région promotrice de *Xist*, leur présence en cours de différenciation étant corrélée avec une surexpression de celui-ci. Nous avons aussi décrit deux nouveaux sites de fixation pour la cohésine, qui colocalise avec CTCF de manière ubiquitaire en amont de *Xist* (surtout sur le futur Xi), mais seulement dans les cellules non différenciées ou en cours de différenciation en amont de *Tsix*. Ces résultats renforcent l'idée grandissante que CTCF peut jouer des rôles différents en fonction des sites concernés et des protéines qu'elle y recrute. Ceci serait médié par la séquence nucléotidique autour du site de fixation ainsi que par le contexte

chromatinien, qui lui permettraient d'engendrer de nombreuses interactions, orchestrées avec précision, grâce à ses divers domaines, notamment les 11 doigts de zinc.

v. Conservation des fixations de CTCF, YY1 et Rad21 chez l'homme

Un des moyens d'aborder le rôle de CTCF, YY1 et Rad21 dans la régulation de *Xist* d'un point de vue fonctionnelle, est d'étudier la conservation de la fixation de ces protéines dans la région promotrice de *Xist* dans d'autres espèces.

Les protéines CTCF (Filippova *et al.*, 1996, Klenova *et al.*, 1998, Burke *et al.*, 2002, Pugacheva *et al.*, 2006) (Figure 43), YY1 (Inouye & Seto, 1994, Kim *et al.*, 2006) et Rad21 sont fortement conservées à travers l'évolution, surtout entre la souris et l'homme (CTCF : 98%, YY1 : 99%, Rad21 : 97%).

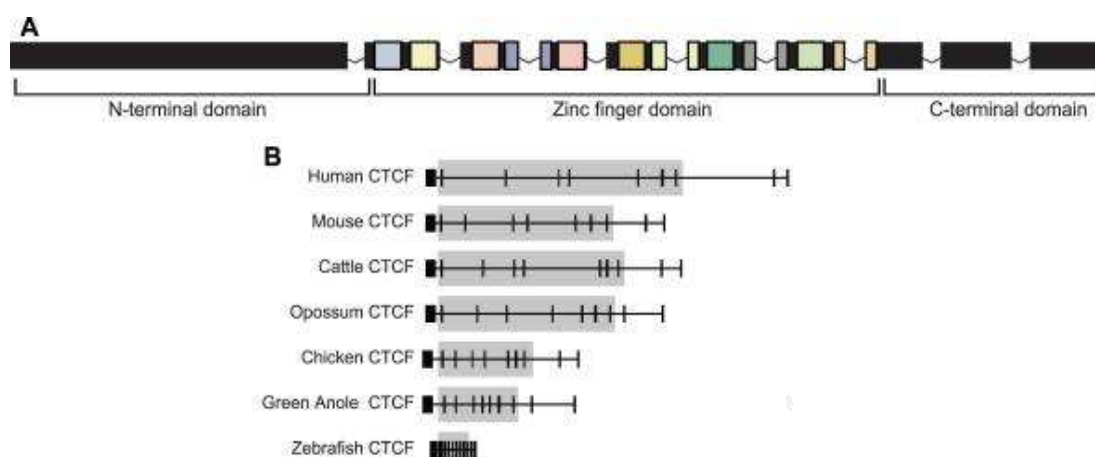


Figure 43 : Structure et conservation du gène CTCF. (A) Sont représentés les 3 domaines structuraux de CTCF : les 11 doigts de zinc, entourés d'un domaine N- et C-terminal. (B) Tous les orthologues de CTCF chez les vertébrés possèdent 10 exons. Les frontières intron-exons sont conservées au sein de la zone grisée entre tous les orthologues de CTCF. (D'après Hore *et al.*, 2008)

Grâce à ses expériences menées au laboratoire, et aux nombreuses données générées par ChIP-seq dans différentes lignées cellulaires humaines, la doctorante Mélanie Makhoulouf, spécialiste du laboratoire en culture de cellules ES humaines, a pu établir un profil de fixation pour nos protéines d'intérêt dans plusieurs lignées. Par soucis de brièveté, du fait que ma participation à ce projet se cantonne à l'analyse des données de ChIP-seq et que ces résultats sont plutôt préliminaires, je me contenterai de ne montrer que les résultats obtenus par analyse de cellules fibroblastiques : MRC-5 pour les mâles, et IMR-90 pour les femelles.

Ainsi, nous avons pu comparer le profil de fixation des protéines CTCF, RAD21 et YY1 dans la région promotrice de *XIST* dans les cellules humaines à celui obtenu chez la souris.

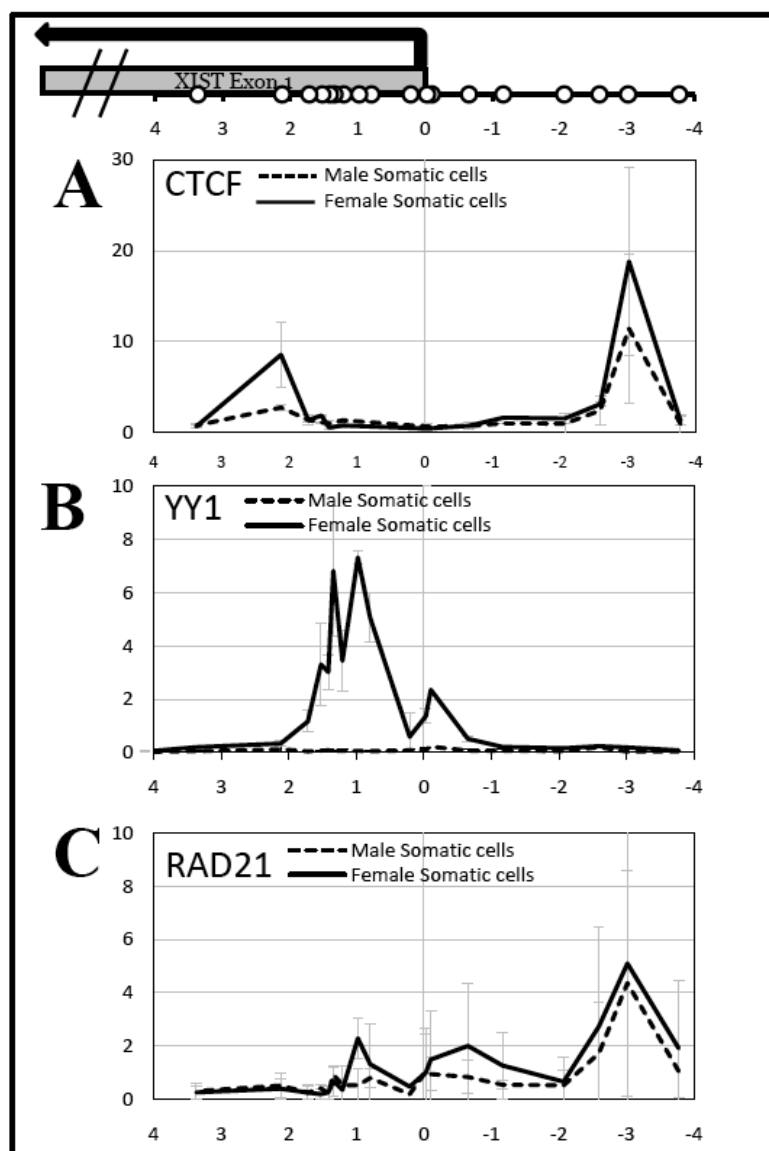


Figure 44 : Fixation de CTCF, YY1 et Rad21 au sein de la région promotrice du *XIST* humain dans des cellules fibroblastiques mâles et femelles. Expériences de ChIP (n=3) mesurées par PCR-quantitative à 21 positions différentes le long de la région promotrice du *XIST* humain. L'enrichissement est présenté en tant que ratio par rapport au bruit de fond, une fois normalisé par l'input (Figure 35A). Fixation de (A) CTCF, (B) YY1 et (C) Rad21 dans des cellules fibroblastiques mâles (pointillés) et femelles (trait plein).

De manière spectaculaire, les 3 protéines se fixent de façon quasi-identique dans la région promotrice du *XIST* humain que ce qui est observé chez la souris (Figures 44A, B, et C ; 34C ; 35B ; 36B et C ; 40B et 41B). En effet, CTCF se fixe aussi bien en amont qu'en aval du promoteur dans les cellules somatiques femelles, mais uniquement en amont dans les cellules mâles (Figure 44A). Les distances de fixation des protéines par rapport aux promoteurs sont différentes de chez la souris, mais cela est attendu compte tenu de l'évolution des séquences (le gène *Xist* étant plus long chez l'homme que chez la souris, voir le chapitre : « *Xist* », p51).

YY1 se fixe aussi en aval du promoteur et, similairement à CTCF, uniquement chez les femelles (Figure 44B). Ce profil de fixation à cet endroit semble indiquer que les protéines étudiées ne se fixent que sur l'allèle inactif du chromosome X à ce stade du développement. Cependant, concernant YY1, on peut observer quelques différences par rapport à la situation murine : d'une part le pic majeur est très large, faisant donc état de plusieurs sites occupés par cette protéine (sachant qu'il existe 8 séquences consensus de fixation à cette région, Kim *et al.*, 2006) ; et il sera intéressant, dans l'avenir, de comparer ce profil à celui existant dans les cellules ES humaines pour voir si, comme chez la souris, il serait possible que différents sites soient utilisés lors de différentes situations. D'autre part, nous pouvons remarquer un pic mineur, juste en amont du promoteur, qui pourrait correspondre à l'observation faite par Hendrich *et al.* qu'il existe une fixation différentielle de YY1 à ce site corrélant avec l'expression de *XIST* dans les cellules somatiques (Hendrich *et al.*, 1997).

Enfin, RAD21 colocalise parfaitement avec CTCF au site en 5' de *XIST*, étant présent aussi bien chez les mâles que chez les femelles (Figure 44C). Nous pouvons aussi nous apercevoir qu'il semblerait que la fixation de RAD21 à la région promotrice de *XIST* soit légèrement plus intense chez la femelle que chez le mâle, phénomène aussi observé chez les cellules somatiques murines femelles ou chez le mutant Ma2L en cours de différenciation. Cette divergence pourrait être le symbole d'une activité d'interaction tridimensionnelle plus importante au sein de la région promotrice de *XIST/Xist* sur le X inactif.

Il semblerait donc que non seulement le réseau d'activation transcriptionnel géré par CTCF et YY1 soit conservé chez l'homme, mais aussi que la fixation de CTCF et RAD21 en amont du promoteur de *XIST* semble avoir une véritable importance de par sa conservation, bien que nous ne connaissions pas encore sa fonction.

Afin de voir l'étendue de cette conservation, j'ai entrepris d'effectuer une recherche *in silico* de sites de fixation potentiels de CTCF, YY1 et Rad21 dans la région promotrice de *Xist* de 8 espèces euthériennes grâce aux sites consensus préalablement obtenus par d'autres laboratoires (Kim *et al.*, 2007, Kim *et al.*, 2009, Parelho *et al.*, 2009) (Figure 45).

Ceci nous permet de visualiser l'importance indéniable de la conservation des sites de fixation de ces protéines à la région promotrice de *Xist*. Bien entendu, cette étude basée uniquement sur des séquences consensus, ne prouve en rien la fixation réelle de ces protéines. Cependant, une séquence consensus n'étant valable, pour révéler un véritable site de fixation, que dans une certaine fraction des cas, il est assez remarquable d'observer une telle conservation de ces

séquences, d'autant plus qu'elles fixent ces protéines chez la souris et l'homme, nous laissant croire qu'il y ait de fortes chances pour qu'elles fassent de même chez les autres espèces.

Cependant, nous pouvons remarquer que seule la souris possède un site consensus pour CTCF en aval du promoteur de *Xist* et que par la même occasion, on y observe aussi un site pour Rad21, malgré une très faible fixation *in vivo*. De plus, à part le chat et en partie le cheval, le rat est la seule espèce à ne pas présenter de site consensus pour CTCF et Rad21 en amont de *Xist*, et ce malgré la proximité évolutive importante par rapport à la souris.

Mais le fait le plus marquant, est que des sites consensus de YY1 (oscillant entre 3 et 8 selon l'espèce) en aval du promoteur sont présents chez toutes les espèces (à part l'éléphant et éventuellement le chat). Cette donnée suggère fortement que YY1 soit le véritable activateur de *Xist* en cours de différenciation, CTCF à ce site, a peut-être une fonction différente ou redondante par rapport à YY1 ou un rôle différent selon les espèces, la régulation de *Xist* pouvant posséder des particularités propres à chacune d'entre elles.

Le fait que nous n'observions pas de fixation de YY1 chez le chat est sans doute dû à un manque de recouvrement du séquençage, un segment conséquent de la région 5' de *Xist* n'étant pas représenté, de même que pour le site CTCF/Rad21 chez l'éléphant. Par contre, concernant le pic chez le rat, le chat et le cheval, il se pourrait que le site ne soit pas conservé ou bien que la séquence de fixation ait légèrement varié avec l'évolution, celle-ci divergeant donc trop de la séquence consensus pour être reconnue par le logiciel.

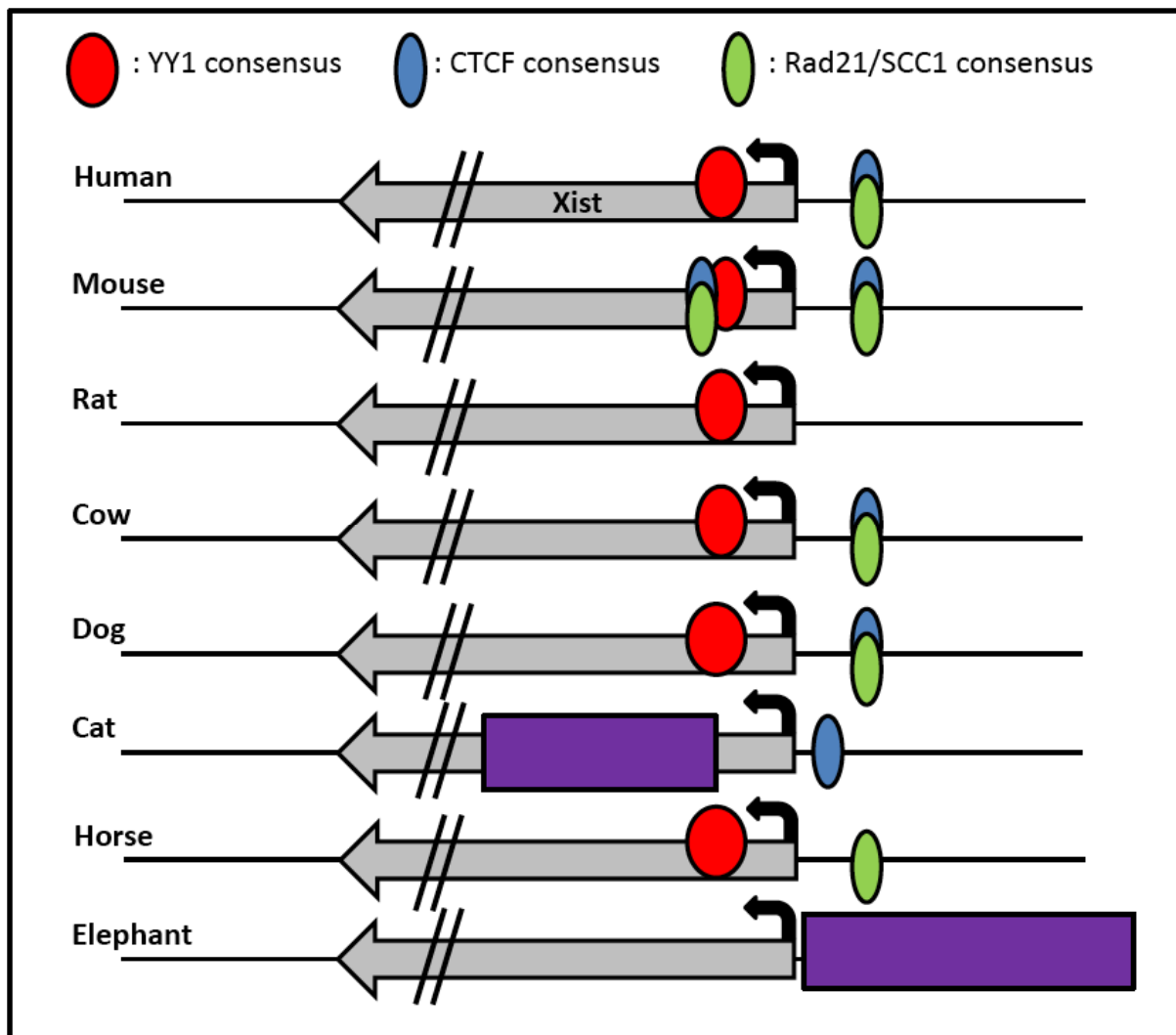


Figure 45 : Conservation des sites consensus de fixation pour CTCF, YY1 et Rad21 dans la région promotrice de *Xist* chez les euthériens. Représentation schématique des sites consensus de CTCF (en bleu), YY1 (en rouge) et Rad21 (en vert) chez 8 espèces de mammifères placentaires. Les analyses ont été réalisées grâce au logiciel VectorNTI sur les séquences fournies par le site UCSC (<http://genome.ucsc.edu/cgi-bin/hgGateway>). Les consensus utilisés sont les suivants, CTCF : CCDSYAGRKG DHR, YY1 : CGCCATNTT, et Rad21 : CCWSYDGRK GKM. Les blocs violets représentent des zones dont la séquence n'est pas disponible.

d. Discussion

Dans le but de caractériser le réseau transcriptionnel de *Xist*, notamment de ses activateurs, au cours de l'inactivation du chromosome X, nous avons entrepris une étude à moyenne échelle (sur 50kb) du profil de fixations de CTCF et YY1, deux candidates préalablement soupçonnés d'agir positivement sur l'expression de *Xist*. À travers cette étude, nous avons pu voir que la protéine à doigts de zinc CTCF se lie à de multiples sites le long du locus *Xist/Tsix*. Au niveau du promoteur P2 de *Xist*, la fixation de CTCF corrèle parfaitement avec la surexpression du

gène non-codant au cours de la différenciation : présent sur le Xi, absent sur le Xa. Il se trouve que YY1 se fixe également à un (ou plusieurs) site(s) situé(s) très près de P2 et aussi de manière corrélative à l'expression de *Xist* lors de l'initiation de l'inactivation. Cela, nous permet de soumettre l'hypothèse que CTCF et YY1 pourraient bien être les activateurs transcriptionnels de *Xist* tant recherchés.

Grâce à la découverte par d'autres laboratoires d'une interaction fréquente de CTCF, à travers tout le génome, avec le complexe de la cohésine, nous avons pu lier deux autres sites de fixation CTCF à une interaction avec la sous-unité Rad21 de ce complexe. De manière intéressante, ces deux sites se trouvent chacun de part et d'autre du locus *Xist/Tsix*, nous amenant à supposer qu'ils pourraient avoir un rôle dans sa régulation. Cependant, au vu des différentes modifications de leur profils de fixation au cours de la différenciation cellulaire, il semblerait que la fonction de chaque pic se fasse indépendamment l'un de l'autre. Cette découverte nous permet d'envisager un nouveau type de régulation du locus *Xist/Tsix* qui s'établirait à travers des structures tridimensionnelles grâce à des interactions entre CTCF et la cohésine.

i. Nos données sont corrélées par la littérature

- Les sites d'hypersensibilité à la DNase

Par le passé, plusieurs équipes de recherche ont cherché des sites d'hypersensibilité à la DNase (ou DHS pour « *DNase Hypersensitive Site* ») à travers le locus *Xist/Tsix* (Figure 46) (Sheardown *et al.*, 1997, Sado *et al.*, 2005, Tsai *et al.*, 2008). Au cours de ces études, plusieurs sites se sont avérés corrélés avec l'expression de *Xist* mais sans que l'on connaisse l'identité de la ou des protéines responsables de cette hypersensibilité.

Ainsi, en comparant nos données à ces DHS, on remarque que les sites 3/HS3/2a et HS5/2c (Figure 46) correspondent à notre pic ubiquitaire de CTCF en amont de *Xist* et à notre site CTCF/YY1 au promoteur P2 de *Xist*, respectivement. Cette corrélation est aussi bien valable au niveau de la localisation que de la régulation en cours de différenciation de ces sites.

De manière intéressante, les DHS observés par Sado *et al.* au sein de leurs cellules mutantes X^{dc} (des cellules somatiques mâles, invalidées pour *Tsix*) dans le même article récapitulent le schéma de fixation de CTCF et YY1 que l'on observe dans Ma1L et Ma2L, c'est-à-dire leur présence ectopique post-différenciation chez les mâles mutés.

De plus, Tsai *et al.*, dont l'étude se focalise sur une région plus étendue, trouvent aussi des DHS dans l'exon 1 de *Xist* (ou nous avons un pic CTCF, surtout après différenciation), dans l'intron 1 de *Xist* (correspondant au site de fixation pour Oct4/Nanog et Sox2) mais aussi à l'exon 8 de *Xist* (pic de CTCF) et en amont de *Tsix* (pic de CTCF et Rad21), corroborant ainsi nos résultats sur le reste du locus. Par contre cette analyse ne semble pas trouver de DHS au niveau des petits exons 2, 3, 4 et 5 où nous observons un fort attachement de CTCF (surtout dans les ES).

De plus, les trois études observent un (ou deux dans le cas de Sheardown *et al.*) sites d'hypersensibilité à la DNase, au niveau du promoteur P1 de *Xist*. Ce site (1/HS1/2b) (Figure 46) est ubiquitaire dans les cellules femelles mais est perdu lors de la différenciation chez les mâles. Il pourrait correspondre à la fixation basale de la machinerie de transcription, qui serait recrutée de manière bien plus importante en cours de différenciation, mais uniquement sur l'allèle inactif chez les femelles.

Le pic 2 de Sheardown *et al.* est quelque peu mystérieux, dans le sens où il n'a pas été identifié par les autres équipes (peut-être dû à un manque de définition de l'analyse en raison des plus grandes régions étudiées) et de par sa présence restreinte aux cellules ES. Cette caractéristique de fixation pourrait nous laisser croire à l'existence d'un facteur répressif autre qu'Oct4/Nanog/Sox2 qui se fixerait en amont de P1 afin de réprimer *Xist* transcriptionnellement. Ce facteur serait sans doute aussi régulé de manière développementale de par son absence dans les cellules somatiques. Sheardown *et al.* confirment bien la fixation d'une protéine d'identité inconnue à ce site par des expériences de retard sur gel (Sheardown *et al.*, 1997).

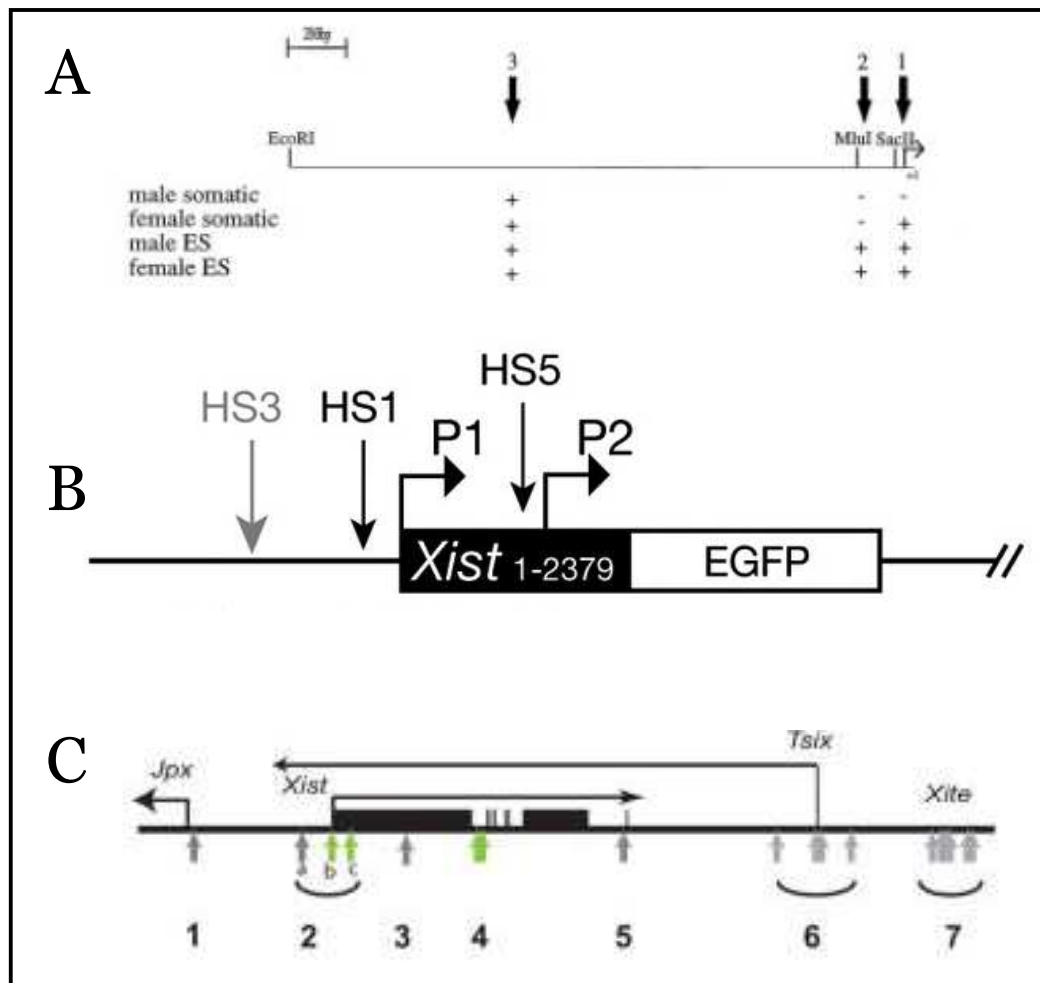


Figure 46 : Les sites d'hypersensibilité à la DNase le long du locus *Xist/Tsix*. Représentation schématique des différents sites d'hypersensibilité à la DNase tel que décrits par (A) Sheardown *et al.*, 1997, dans les cellules ES et fibroblastiques mâles et femelles ; (B) Sado *et al.*, 2005, avec en noir les sites spécifiques à l'X inactif et en gris le site ubiquitaire et (C) Tsai *et al.*, 2008, avec en gris les sites constitutifs et en vert les sites spécifiques des cellules ES.

- Tests luciférase

Différents tests à la luciférase semblent corroborer notre hypothèse que la fixation de CTCF et YY1 à P2 est responsable de la surexpression de *Xist* au cours de l'initiation de l'inactivation. Sun *et al.* ainsi que Kim *et al.* ont réalisé des constructions comprenant des régions promotrices de *Xist* différentes suivies d'un gène codant pour la luciférase (Sun *et al.*, 2006, Kim *et al.*, 2006). Leurs deux constructions minimales (-540/+9 et -1143/+956pb par rapport à P1) ne contiennent pas les sites de fixation de CTCF et YY1 autour de P2, l'expression de la luciférase y étant très faible, même plus faible que l'expression de la luciférase en absence de promoteur (Figure 47). Bien que dans l'expérience de Sun *et al.* cette expression soit augmentée en cours de différenciation, elle est loin d'atteindre les valeurs attendues dans ce

genre de contexte. Il semblerait donc qu'il manque quelque chose à la capacité du promoteur à être surexprimé. Et c'est en étudiant une autre construction contenant cette fois-ci 1814pb en aval de P1 (et donc les sites de fixation pour YY1 et CTCF ainsi que P2) que Kim *et al.* obtiennent une surexpression de la luciférase (le fait que par la suite ils retrouvent une expression faible du gène rapporteur suite à des mutations individuelles des sites YY1 semble indiquer que ce n'est pas le promoteur P2, en tant que tel, qui est responsable de la surexpression) (Figure 47). Cette expérience semble bien confirmer la fonction de YY1 dans la surexpression de *Xist* dans un contexte dépourvu des répresseurs principaux (*Tsix* et Oct4/Nanog/Sox2).

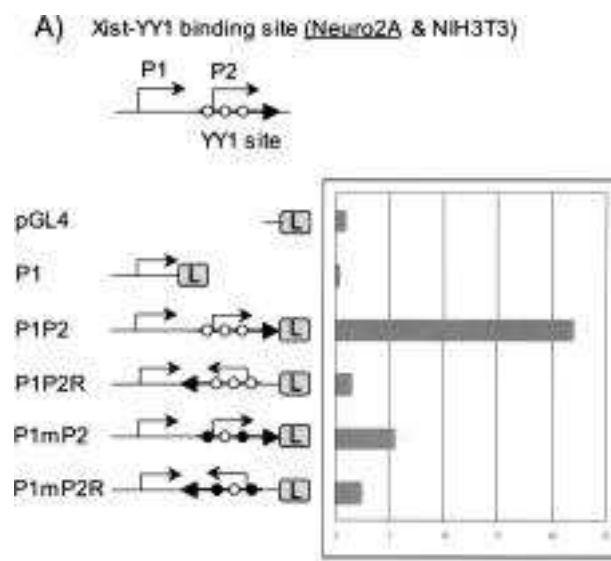


Figure 47 : Activité transcriptionnelle d'un groupe de sites YY1 au promoteur P2 de *Xist*. L'implication transcriptionnelle du groupe de sites YY1 a été testé en insérant la région promotrice endogène de *Xist* devant le gène sans promoteur de la luciférase. Cinq constructions et un témoin négatif (pGL4) ont été analysés. Chaque construction a été analysée plus de trois dans le type cellulaire entre parenthèses (cellules somatiques). La mutation des sites YY1 est : CATTATNTT au lieu de CGCCATNTT. Les séquences des fragments de promoteurs utilisés correspondent à : construction P1 (AJ421479; 105154–107233), construction P1P2 (AJ421479; 105154–108089). (D'après Kim *et al.*, 2006)

ii. Situation dans les cellules ES non-différenciées

Bien que l'association de CTCF et de YY1 au promoteur de *Xist* dans des cellules en différenciation semble garantir une surexpression de celui-ci, paradoxalement, leur présence aux mêmes sites dans les cellules ES indifférenciées ne semble pas fournir la même certitude. En effet, nous avons observé la présence de CTCF et YY1 au promoteur P2 dans tous les

types cellulaires sauvages pluripotents étudiés, alors que l'expression de *Xist* dans ces cellules est basale.

L'explication la plus simple à ce paradoxe est que les facteurs responsables de la répression de *Xist*, à savoir Oct4/Nanog/Sox2 et *Tsix*, empêchent CTCF et YY1 de jouer leur rôle, que ce soit par modification de conformation tridimensionnelle de la région (bloquant ainsi des interactions nécessaires à leur bon fonctionnement) ou tout simplement en exerçant une répression plus intense que ne l'est la surexpression exercée par nos activateurs.

Pour s'en assurer, on pourrait effectuer un test de luciférase dans des cellules ES. En effet, en ne mettant que la région promotrice de *Xist* (comprenant, bien sûr, les sites de YY1 et CTCF) devant le gène de la luciférase, le promoteur ne serait plus sous le contrôle des deux systèmes répresseurs démontrés à ce jour. L'expression du gène rapporteur dans un contexte au sein duquel son promoteur est généralement réprimé, montrerait bien que l'absence d'effet de CTCF et YY1 dans les cellules ES ne serait due qu'à un effet dominant exercé par les mécanismes répresseurs. Le fait d'effectuer cette expérience dans des cellules pluripotentes nous permettrait aussi de vérifier deux autres possibilités qui pourraient expliquer l'absence d'expression de *Xist* malgré la présence des ses activateurs dans un contexte sauvage :

- les activateurs ne sont peut-être pas fonctionnels dans les cellules pluripotentes de par la présence ou l'absence d'une modification post-traductionnelle inhibant leur fonction à ce stade cellulaire
- la fonction des activateurs passe potentiellement par le recrutement d'autres protéines effectrices, celles-ci n'étant peut-être pas exprimées à ce stade.

Une autre possibilité venant s'ajouter à un effet dominant des facteurs de pluripotence et de l'antisens, pourrait être l'absence d'une quantité suffisante de Rnf12, qui semble aussi être réprimée au sein des cellules ES par les facteurs de pluripotence (Figure 48 et le chapitre : « Régulation en *trans* par divers facteurs », p92). Cette protéine, qui lorsqu'elle est surexprimée dans les cellules ES mâles, arrive à engendrer l'initiation de l'inactivation du chromosome X, pourrait être responsable d'une capacité de *Xist* à être surexprimé en cours de différenciation (Jonkers *et al.*, 2009). La manière dont Rnf12 arrive à contourner les différents types de répression dans les cellules ES est encore inconnue, mais une analyse par ChIP de Rnf12 au sein du locus *Xist/Tsix* serait sans doute très bénéfique à une meilleure compréhension du phénomène. Le fait qu'une surexpression de Rnf12 dans les ES mâles suffise à surexprimer *Xist* a conduit les auteurs de l'article à stipuler que : « En outre, l'activation de l'inactivation du chromosome X dans des cellules mâles transgéniques pour *Rnf12* ne peut pas être expliqué par un effet de Rnf12 sur la différenciation. Il est donc plus

probable que les cellules ES transgéniques pour *Rnf12* restent indifférenciées, ce qui implique que tous les autres facteurs nécessaires à l'initiation de l'inactivation sont déjà présents dans les cellules ES indifférenciées, bien que le niveau de ces autres facteurs peut augmenter lors de la différenciation. Nos résultats suggèrent également que la répression de *Xist* peut être surmonté par une expression accrue de RNF12, même dans les cellules ES indifférenciées. », corroborant la possibilité que CTCF et YY1 puissent être actifs au sein de cellules ES non-différenciées lorsque l'on leur en donne l'occasion.

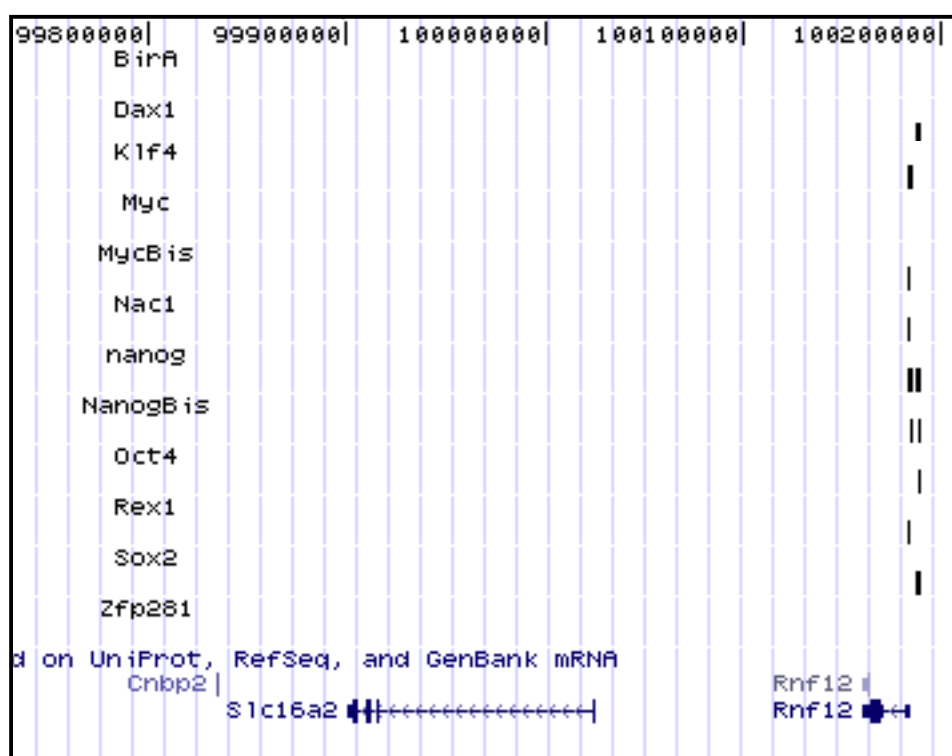


Figure 48 : Fixation de multiples facteurs de transcription au promoteur de *Rnf12* dans les cellules ES de souris. Données de ChIP-CHIP obtenues par Kim *et al.*, 2008, analysées grâce au site UCSC Genome Browser. Les résultats sont obtenus lors de l'utilisation des données ChIP-seq de Chen *et al.*, 2008.

iii. Différents mutants de *Tsix* ont différents effets sur l'expression de *Xist*

Je pense qu'il est intéressant de revenir sur la différence d'expression de *Xist* observée entre les différents mutants de *Tsix* (Figure 38) préalablement évoquée au cours de l'analyse de ces résultats.

Tandis que *Xist* est surexprimé d'environ un facteur 100 entre Ma1L et Ma2L ou entre CK35 et ΔAV dans les cellules ES mâles indifférenciées, ce facteur n'est plus que d'environ 3 ou 4 entre les CK35 et les Δ65/ΔDxPas34 ou encore entre les E3.4 et les D102/c5.1 (Figure 38 ; Morey *et al.*, 2001, 2004 ; Vigneau *et al.*, 2006). Il semblerait donc il y ait autre chose en jeu

que simplement la répression post-transcriptionnelle décrite par Navarro *et al.* (Navarro *et al.*, 2005), qui ne saurait expliquer les différences conséquentes observées, sachant que l'expression de *Tsix* est abolie dans l'ensemble des mutants (à part $\Delta DXPas34$ chez qui il reste de 10 à 15% d'expression).

En considérant précisément la nature des différentes mutations, on remarque que l'on peut distinguer deux catégories différentes : la première étant constituée de mutations ne touchant uniquement à *Tsix* (Ma2L et ΔAv) et la seconde regroupant les mutations de plusieurs dizaines de kb (D102, c5.1, $\Delta 65$, D102 étant aussi une délétion de 65kb mais engendrée dans une lignée femelle), $\Delta DXPas34$ étant un cas à part, sa faible surexpression de *Xist* étant sans doute due à la transcription résiduelle de l'antisens.

Tout ceci me permet de suggérer qu'il existe un élément au sein d'une région de 50kb en aval de *Tsix* (par rapport au sens de la transcription de *Xist*) qui aurait un effet activateur sur l'expression de *Xist*. Cependant, nous ne pouvons pas déterminer, dans l'état actuel de nos connaissances, s'il s'agit d'un effet médié par un contact à travers une interaction à longue distance nécessitant une structuration tridimensionnelle particulière ou s'il est effectué grâce à un élément codant pour une protéine, un microARN (bien qu'aucun microARN n'ait été détecté dans cette région) ou autre. Cet effet est d'ailleurs encore plus remarquable à jour 4 de différenciation à l'acide rétinoïque où l'expression de *Xist* est 3 fois plus élevée dans les Ma2L/ ΔAv (et même $\Delta DXPas34$) que dans les $\Delta 65$ (Figure 38, Morey *et al.*, 2004, Vigneau *et al.*, 2006).

Indépendamment, nous pouvons aussi remarquer que l'expression de *Xist* chez les mutants mâles n'arrive jamais à atteindre le niveau que l'on mesure chez les femelles, tout du moins après quatre jours de différenciation. Il est possible que cela soit dû au manque de Rnf12 chez les mâles, les femelles bénéficiant d'une expression biallélique dont la production devrait, au final, se fixer sur un seul allèle d'après la théorie de brisure de symétrie (Nicodemi & Prisco, 2007).

iv. Interaction entre CTCF & YY1 ?

D'après les résultats que nous avons obtenus, il semblerait que CTCF et YY1 soient impliquées dans la surexpression de *Xist*. Bien que CTCF se fixe en plusieurs endroits le long du gène *Xist*, seule sa fixation près du promoteur de P2 semble requise à sa fonction potentielle d'activateur. YY1, quant à lui, ne se fixe à *Xist* qu'en amont de P2 mais possède 3 sites à cet endroit, ceux-ci étant séparés au maximum d'environ 220pb. Les distances entre le site CTCF et les sites YY1 à P2 sont comprises entre 36 et 245pb. Ces faibles distances, ainsi

que la description d'une hétérodimerisation possible des deux protéines (Donohoe *et al.*, 2007) nous amène à nous demander si, à ce locus, CTCF et YY1 fonctionnent en collaboration ou indépendamment.

Curieusement, l'article qui propose cette interaction observe des cinétiques de recrutement différentes pour ces protéines à plusieurs sites autour du microsatellite *DxPas34*, remettant en cause un lien direct entre-elles, tout du moins concernant leur recrutement à ces sites là. De plus, d'après des données de ChIP-seq, il semblerait que ces deux protéines ne colocalisent pas particulièrement au niveau du génome entier (communication personnelle avec J.J. Miranda *et al.*, 2010).

Cependant, d'après Donohoe *et al.*, CTCF se lierait à YY1 grâce à son domaine N-terminal, celui-ci ayant aussi été reconnu comme essentiel à la capacité de CTCF à activer transcriptionnellement l'expression du gène *APP* (pour « *Amyloid Precursor Protein* ») dans les cellules HeLa (Vostrov *et al.*, 2002). Il serait donc intéressant de regarder si YY1 se fixe aussi à la région contrôlant l'activité de ce gène, d'autant plus qu'à ce jour, YY1 est la seule protéine décrite comme se fixant à ce domaine de CTCF. La fonction activatrice de CTCF à ce locus passerait donc éventuellement par le recrutement de YY1 et non pas un effet direct.

En définitive, la seule manière de savoir si une des deux protéines recrute l'autre sera de déléter au locus endogène l'un de leur sites et de voir ce que cela engendre vis-à-vis de la fixation de l'autre protéine au site voisin.

v. De quelle manière l'activation de *Xist* est-elle obtenue ?

Nous n'avons pas encore abordé la manière dont CTCF et YY1 effectuent la surexpression de *Xist*. Bien que les deux protéines soient capables de recruter directement l'ARN polymérase II afin d'initier une activation transcriptionnelle, plusieurs autres mécanismes sont envisageables afin d'obtenir une conclusion similaire.

Le mécanisme le plus simple à envisager, à part le recrutement de l'ARN polII, est la génération d'une structure chromatinienne euchromatique au promoteur de *Xist* par interaction avec des modificateurs d'histones (comme CIITA pour CTCF, ou p300 et Gcn5 pour YY1 par exemple) (voir les chapitres : « CTCF », p106 et « YY1 », p114), cette structure euchromatique permettant ainsi la fixation par le PIC une fois les mécanismes répressifs éteints.

Ou bien, on pourrait tout simplement envisager l'existence d'un recrutement indirect du PIC au promoteur de *Xist*. Des protéines interactrices avec CTCF ou YY1 seraient, elles, responsables de recruter la machinerie de transcription (par exemple Sp1 pour YY1) (voir le

chapitre : « YY1 », p114). Mais le seul moyen de savoir si une certaine protéine est impliquée dans la régulation effectuée par CTCF ou YY1 à ce locus est de tester sa présence par ChIP puis d'empêcher sa fixation afin de voir son effet sur le mécanisme.

Dans notre cas, la fixation de CTCF et de YY1 à P2 n'engendre ni le recrutement de l'ARNpolIII, ni l'euchromatinisation de la région au sein des cellules ES non-différenciées mais ces rôles sont toutefois envisageables lorsque la différenciation s'initie.

vi. CTCF et YY1 impliquées dans la méthylation de l'ADN du promoteur de *Xist* ?

Les protéines CTCF et YY1 sont toutes deux connues comme ayant une fixation sensible à la méthylation, c'est-à-dire qu'elles sont incapables de se fixer à de l'ADN méthylé. Cependant, des équipes ont récemment mis en évidence leur capacité à influencer sur cette même méthylation à travers le génome. Tandis que YY1 serait en partie responsable de la méthylation de l'îlot CpG de *Xist* au cours de l'oogénèse et de la maintenance d'une absence de cette dernière durant la spermiogénèse (Kim *et al.*, 2009, voir le chapitre : « Régulation en *trans* de *Xist* par divers facteurs », p92), CTCF serait capable de recruter et d'activer Parp-1 qui à son tour est capable d'inhiber l'action de l'ADN méthyltransférase de maintenance Dnmt1 (Guastafierro *et al.*, 2008, Reale *et al.*, 2005, Engel *et al.*, 2006).

Il s'avère justement que les sites de fixation de CTCF et YY1 liés à la surexpression de *Xist* sont localisés en bordure de l'îlot CpG de son promoteur (Figure 49).

Le promoteur de *Xist* est mosaïquement méthylé sur les deux allèles dans les cellules ES femelles indifférenciées (tandis que légèrement hyperméthylé chez les mâles) (Norris *et al.*, 1994, Sun *et al.*, 2006, Sado *et al.*, 2006, Rougeulle non publié). Par contre, en cours de différenciation cet îlot CpG va subir une méthylation intense de son ADN sur le X actif tandis que le promoteur du X inactif restera déméthylé afin de pouvoir exprimer *Xist* convenablement (Norris *et al.*, 1994, Sado *et al.*, 2004).

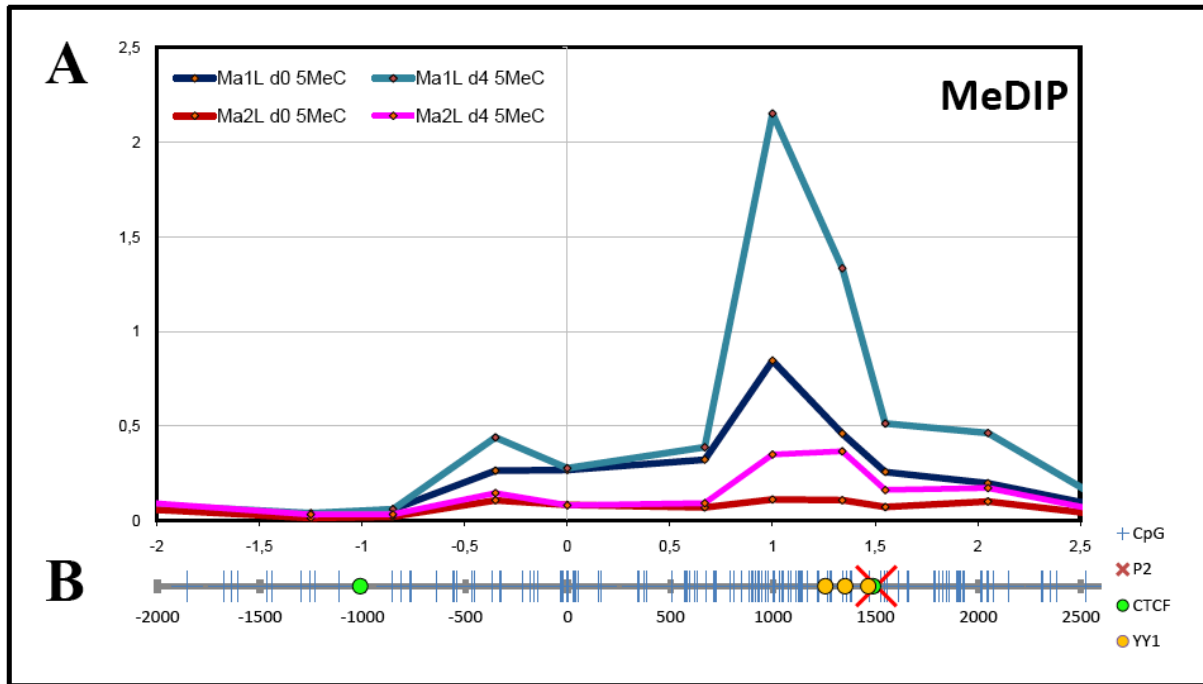


Figure 49 : Analyse de la méthylation de l'ADN au sein de la région promotrice dans Ma1L et Ma2L en cours de différenciation. (A) Expérience de Me-DIP (pour « *Methyl-DNA ImmunoPrecipitation* ») menée par Claire Rougeulle dans la région promotrice de *Xist* aux mêmes positions que le Figure 36A. On peut observer une différence dès le stade ES indifférencié entre futur Xa (Bleu marine) et futur Xi (rouge) qui est accentuée en cours de différenciation. (B) Représentation de l'ensemble des dinucléotides CpG (traits bleus) dans la région promotrice de *Xist*. Sont aussi représentés les différents sites de fixation de CTCF (rond vert) et YY1 (rond jaune) ainsi que le promoteur P2 (croix rouge), le promoteur P1 étant placé à l'abscisse 0.

Des données récentes, publiées par le laboratoire d'Adrian Bird, confirment une fois de plus l'observation d'une méthylation différentielle entre les deux allèles dans des cellules somatiques. De plus ils ont pu identifier des protéines se fixant à cet îlot CpG en fonction de son état de méthylation : MeCp2 et Cfp1 (Thomson *et al.*, 2010, Skene *et al.*, 2010), MeCp2 se fixant sur l'allèle méthylé et Cfp1 sur le déméthylé. Le rôle de Cfp1 n'est pas encore clair mais il semblerait capable de recruter la méthyltransférase H3K4 Setd1, favorisant ainsi la mise en place d'une structure euchromatinienne. À l'inverse, MeCp2 recruterait des histones déacétylases de manière à organiser un domaine répressif autour de son site de fixation.

Au vu de l'interaction de CTCF avec Parp1 et leur influence sur la répression de Dnmt1 (voir le chapitre : « CTCF », p108), il semblerait fort crédible que CTCF serve à la maintenance de l'état déméthylé du promoteur de *Xist* sur l'allèle inactif du chromosome X (allèle actif de *Xist*) dans les cellules somatiques. Ce rôle pourrait s'effectuer soit en parallèle de celui d'activateur transcriptionnel, soit dans un rôle exclusivement lié au blocage la méthylation de l'ADN (permettant ainsi la fixation de YY1). Cependant la question dans cette situation est de

savoir si l'absence de CTCF est la cause ou la conséquence de la méthylation de cet îlot CpG. Une fois de plus, le seul moyen d'obtenir une réponse serait de muter le site de fixation de CTCF mais dans les cellules Ma2L par exemple. Ainsi si l'on observe une absence de fixation de YY1 à P2 en cours de différenciation, on devrait sûrement aussi observer une absence de surexpression de *Xist* ainsi qu'une forte méthylation de l'îlot CpG, tendant à prouver que c'est bien CTCF ou YY1 qui est responsable du profil de méthylation (reste à découvrir la raison pour laquelle elle ne se trouve plus fixée à l'allèle actif dans les cellules femelles différenciées). Dans le cas contraire, cela signifierait que c'est la méthylation qui est responsable de l'éviction de CTCF et de YY1 à P2 du Xa ; reste à savoir comment cette méthylation se retrouve empêchée sur le Xi. Cependant, ceci paraît peu probable, de par l'observation de Sado *et al.* qu'en absence de méthylation de la l'ADN (double-mutant *Dnmt3a* et *3b*^{-/-}), la surexpression de *Xist* s'effectue tout de même de manière monoallélique (Sado *et al.*, 2004), nous laissant supposer que cette absence de méthylation à son promoteur ne permet pas le maintien biallélique de YY1 et de CTCF (ce qui reste à être vérifié).

vii. Recrutement de Rad21

CTCF possède un pic de fixation, de part et d'autre du locus *Xist/Tsix*. Il se trouve que Rad21 colocalise avec ces derniers au sein des cellules ES (bien que plutôt faiblement chez les mâles). Par contre après l'induction de la différenciation, celui se trouvant en amont de *Xist* se maintient tandis que celui fixé en amont de *Tsix* disparaît.

À travers des études se servant de CHIP-seq afin de cartographier la fixation de ces protéines à l'échelle du génome entier (Parelho *et al.*, 2008, Wendt *et al.*, 2008, Stedman *et al.*, 2008), différents groupes ont remarqué que CTCF et la sous-unité de la cohésine Rad21 colocalisent en de très nombreuses circonstances (voir le chapitre : « CTCF », p108). De plus, ils se sont aperçus que leur présence mutuelle à un même site de fixation renforçait leur stabilité, la fixation de CTCF étant considérablement réduite dans des cellules déplétées pour Rad21 (Parelho *et al.*, 2008, Wendt *et al.*, 2008, Hou *et al.*, 2010), tandis que la présence de CTCF est, la plupart du temps, nécessaire à celle de Rad21 (Rubio *et al.*, 2008, Wendt *et al.*, 2008, Parelho *et al.*, 2008). On ignore encore comment s'effectue la fixation de la cohésine à l'ADN lors d'une interaction avec CTCF : il pourrait aussi bien s'agir d'une interaction protéine-protéine où CTCF est seule attachée à l'ADN, ou alors, cette dernière pourrait moduler l'ADN afin que la cohésine puisse aussi s'y accrocher.

Par contre, curieusement, CTCF et Rad21 ont une séquence consensus quasiment identique, nous indiquant que les deux protéines se fixent exactement au même endroit sur l'ADN. On

pourrait ainsi se demander si le nombre de sites auxquels se fixe la cohésine indépendamment de CTCF (comme lors de sa fonction telle qu'initialement décrite : dans l'appariement des chromatides sœurs) est suffisamment élevé pour que ce consensus soit vraiment le reflet de sa véritable séquence de fixation ou si le fait que CTCF et Rad21 colocalisent à autant de sites à travers le génome, fasse que le consensus de Rad21 soit en fait celui de la protéine qui la recrute.

viii. Interactions 3D ?

La manière dont sont placés les deux pics de CTCF liant aussi la cohésine dans les cellules ES, nous laisse, à première vue, supposer une interaction directe entre les deux loci, la distance minimale de 10kb pour réunir deux régions étant largement atteinte (Gondor & Ohlsson, 2009). Cependant, le pic en amont de *Xist* est à peu près ubiquitaire (sa fixation étant faible dans les cellules ES mâles sauvages) tandis que celui en 5' de *Tsix* n'est pas maintenu dans les cellules somatiques, indiquant une forte probabilité qu'ils ne soient pas associés fonctionnellement. De plus, le ChIP dans les cellules $\Delta 65$, au sein desquelles le pic en amont de *Tsix* est délété, ne témoigne d'aucune modification de la fixation de CTCF ou Rad21 en 5' de *Xist* (données non montrées), ce qui en soi n'est pas étonnant, mais le cas contraire aurait pu être indicatif.

La présence de sites de fixation multiples de CTCF tout le long du locus *Xist/Tsix* demeure encore mystérieuse. Ces pics situés à des positions clés de la séquence de *Xist*, ne sont encore associés à rien de connu mais pourraient très bien être associés entre eux afin de former une structure tri-dimensionnelle nécessaire à la régulation de *Xist* et *Tsix* dans les ES.

La technique vraisemblablement la plus utile pour étudier d'éventuelles interactions tridimensionnelles entre diverses séquences est le 3C (pour « *Chromatin Conformation Capture* ») (Dekker *et al.*, 2002). Le laboratoire de Jeannie Lee ayant effectué des analyses en 3C sur toute la région, nous avons maintenant une première idée d'interactions possibles de nos sites *in vivo* (Tsai *et al.*, 2008). Même s'il manque plusieurs contrôles afin de s'assurer de l'exactitude des interactions observées et que d'après les spécialistes, il est difficile d'être sûr d'une interaction si elle est composée de deux points distancés de moins de 30kb, il semblerait que ces données puissent être majoritairement reproduites en utilisant des conditions bien plus sûres (Elphège Nora, communication orale *et al.*, 2010).

Tsai *et al.* semblent observer deux domaines d'interactions : un premier qui comprend du promoteur de *Jpx* à l'exon 8 de *Xist*, tandis que le deuxième irait de l'exon huit de *Xist* à *Xite*,

et ce de manière similaire qu'il s'agisse de cellules ES mâles ou femelles, les cellules différenciées mâles montrant par la suite de plus faibles interactions que les femelles. Cependant, en se servant de la région promotrice de *Xist* (comprenant le pic CTCF/Rad21) comme amorce, on peut apercevoir une interaction avec la région correspondant au 2^{ème} site CTCF/Rad21, cette interaction n'étant pas visible dans les segments de *Tsix* avoisinants (Figure 50). Il est donc possible que malgré l'existence supposée de deux blocs interagissants de manière maximale dans les ES, il pourrait y avoir une interaction entre les 2 sites fixant CTCF et Rad21.

Le fait que cet article propose la formation de deux domaines d'interactions tridimensionnelles distincts, séparés au niveau de l'exon huit, pourrait impliquer fonctionnellement la fixation de CTCF à ce site dans ce phénomène.

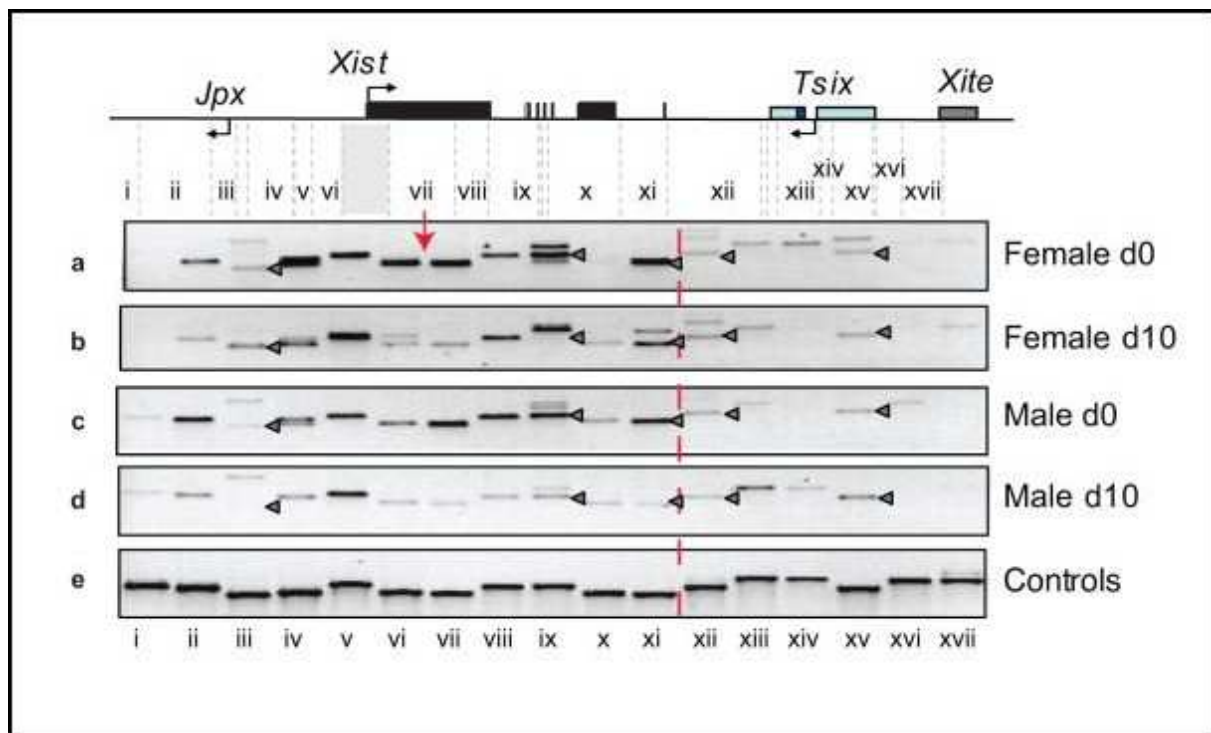


Figure 50 : Interactions tridimensionnelles d'un fragment du XIC avec la région promotrice de *Xist*. Analyse par 3C (Annexe 3, 270) des interactions d'une partie du XIC dans des cellules ES et différenciées en corps embryonnaires mâles (J1) et femelles (EL16.7) en utilisant les enzymes de restriction *Bgl*II et *Bam*HI et un site d'ancrage contenant le promoteur de *Xist* (région grisée et flèche rouge). La fin du domaine *Xist* est indiqué par les traits rouges. Les triangles gris représentent des produits PCR spécifiques. La présence de bandes multiples dans certaines colonnes est due à une digestion incomplète de l'ADN. (D'après Tsai *et al.*, 2008)

Ce genre d'expérience est très utile pour mettre en évidence des interactions tridimensionnelles agissantes en *cis* à relativement courte distance, que l'on appelle aussi intrachromosomiques. Par contre, l'étude à grande échelle pour observer notamment des

interactions en *trans* ou interchromosomiques, nécessite d'autres techniques telles que le 4C (pour « *circular 3C* ») ou le 5C (pour « *3C carbon copy* ») (Simonis *et al.*, 2007, Fullwood & Ruan, 2009). Et au vu de la capacité de CTCF à orchestrer ce genre d'interaction (Ling *et al.*, 2006), il sera essentiel d'effectuer ces expériences afin de connaître les partenaires et éventuellement la fonction de ces pics.

S'il semble donc qu'il existe bien une structure tridimensionnelle complexe au locus *Xist/Tsix*, nous manquons aujourd'hui de preuves pour affirmer que ce réseau interactif est bien l'œuvre (tout du moins en partie) des multiples sites de fixation de CTCF et de la cohésine que nous avons découverts. La répétition de ces expériences dans des cellules déplétées pour CTCF ou pour la cohésine pourrait être très intéressante, de même que dans des cellules mutées au niveaux des sites de fixation.

Le fait que la fixation de Rad21 en amont du promoteur du P1 de *Xist* soit plus faible sur le futur X actif que sur le futur X inactif (Figure 42) pourrait être représentatif d'un rôle plus important ou d'un état de compaction différentiel du locus au sein des cellules indifférenciées, le Xi possédant donc potentiellement une compaction plus importante que le futur Xa au promoteur de *Xist*. Une hypothèse intéressante consisterait à se demander si ce phénomène pourrait être relié à l'observation récente par Mlynarczyk-Evans *et al.* de l'existence du chromosome X dans les cellules ES femelles sous deux formes de compaction différentes (Mlynarczyk-Evans *et al.*, 2006).

En étudiant des cellules ES femelles par FISH avec une sonde dirigée le *XIC*, cette équipe a observé que chaque chromosome X pouvait émettre un signal simple ou en doublon. Puis en analysant des cellules hétérozygotes pour *Xist* ou *Tsix*, ils se sont aperçus que le type de signal dépend du devenir du chromosome : le futur Xa montrant un doublet, le futur Xi un point unique. Dans les cellules sauvages, ces simplets/doublets peuvent aussi représenter le devenir transcriptionnel du X. Cependant ces états distincts fluctuent d'un chromosome à l'autre, ne se fixant définitivement que lorsque la différenciation est engagée.

Ce phénomène pourrait s'expliquer par des forces différentes de cohésion entre les chromatides sœurs après la réplication de l'ADN, mais il tendrait surtout à prouver que le devenir de chaque X est fixé avant l'initiation de l'inactivation, bien qu'il fluctue entre les deux Xs au cours de la participation des cellules à l'épiblaste. Et il est aussi possible qu'au sein des cellules hétérozygotes pour *Xist/Tsix*, les différences de transcription de ces gènes (les deux sont exprimées dans les cellules ES, même si *Xist* l'est très faiblement, Lee *et al.*,

1999) engendrées par les mutations soient suffisantes pour indiquer la probabilité d'un état futur.

Ceci étant, cette théorie n'explique pas pourquoi cette disparité d'intensité des pics de fixation disparaît en cours de différenciation, ni pourquoi cette fixation est maintenue suite à l'inactivation si elle est responsable du choix du chromosome X à inactiver. De plus, l'intensité de fixation de CTCF au même site ne variant que très peu entre les différentes lignées cellulaires nous amène à être très prudents par rapports aux conclusions tirées.

ix. Un rôle dans le *kissing* ?

Au cours des premiers jours (entre 1,5 et 2 jours) de la différenciation de cellules ES femelles, il est possible d'observer un phénomène de transvection des deux chromosomes X (Bacher *et al.*, 2006, Xu *et al.*, 2006) (voir le chapitre : « Le choix », p15). Il semblerait que ce phénomène soit orchestré par deux régions distinctes, l'une étant *Xpr* (située environ 200kb en amont de *Xist*) (Augui *et al.*, 2007), l'autre étant un segment d'environ 15kb de la région *Tsix/Xite* mais dont de petits sous-fragments de 1 à 2kb peuvent récapituler le phénomène de transvection (Xu *et al.*, 2006 *et al.*, 2007).

Plusieurs études ont maintenant mis en cause CTCF dans ce phénomène (Xu *et al.*, 2007, Donohoe *et al.*, 2009), la déplétion de cette protéine inhibant le rapprochement en cours de différenciation chez la femelle. Il semblerait que chaque sous-fragment, malgré une absence de similarité de séquence entre eux, soit capable d'induire le « *kissing* » lorsque introduit ectopiquement dans des cellules mâles. D'après les auteurs cela serait dû au fait qu'il contiendrait un site de fixation pour CTCF. Cependant même un fragment ne contenant pas de site CTCF y parvient aussi.

La possibilité que CTCF joue un rôle dans la transvection médiée par *Xpr* est aussi envisageable, le BAC 5 utilisé par Augui *et al.* recouvrant cette région contient 3 sites consensus pour CTCF (2 en amont du promoteur de *Xpct* et un au sein de son premier intron). De plus, un autre BAC (le BAC 8) utilisé par cette équipe au cours de cette étude, contient une large région dont font partie *Xist* et *Tsix*, et il apparaît que, corrélant en partie les données de l'équipe Lee, on observe une quantité importante de cellules dont les chromosomes X sont à proximité, aussi bien à jour 2 que, étonnamment, à jour 4.

D'après nos résultats, CTCF et la cohésine colocalisent de part et d'autre de *Xist* et de *Tsix* tout au long de cette période, et leur rôle avéré dans les interactions interchromosomiques font de ces sites des candidats intéressants pour expliquer une sorte de dialogue entre les deux

chromosomes avant l'initiation de l'inactivation afin d'établir le choix et peut-être le comptage.

De plus, Augui *et al.* démontrent que cette proximité des deux *Xics* a surtout lieu au cours de la phase S du cycle cellulaire, phase qui justement, se trouve être celle où l'on observe le maximum d'interactions entre CTCF et la cohésine (Figure 51).

Etudier la transvection des chromosomes X dans des cellules déplétées pour la cohésine pourrait fournir une première réponse à la participation potentielle de nos pics CTCF/Rad21 à ce phénomène, l'autre solution étant, encore une fois, de muter les sites de fixation endogènes sur un des deux chromosomes et d'observer s'il y a rapprochement des deux X en cours de différenciation, ainsi que d'analyser le choix et le comptage dans ces cellules afin de s'assurer que les mutations n'induisent pas de biais.

CTCF a par ailleurs été proposée comme responsable de la transvection périnucléolaire des allèles du gène *CCND1* chez l'homme dans des cellules de lymphome de manteau (Liu *et al.*, 2008). Les acteurs de cette découverte proposent d'ailleurs que CTCF remplisse ce rôle en collaboration avec la nucléophosmine, celle-ci ayant été reconnue comme une interactrice majeure de CTCF. Elle fonctionnerait en joignant des promoteurs à des séquences régulatrices, à l'origine très éloignées, à la périphérie nucléolaire (Yusufzai & Felsenfeld, 2004). Ce genre de phénomène favoriserait non pas une interaction interchromosomique directe entre les sites CTCF mais un rapprochement nucléaire considérable en fixant les chromosomes de manière proche au nucléole ou bien à la matrice nucléaire.

x. Un rôle pour le pic CTCF/Cohésine situé en amont de *Tsix*

Comme nous l'avons évoqué au début de la partie Résultats et comme nous venons de le revoir dans la partie précédente, CTCF a la capacité d'interagir avec la nucléophosmine et les lamines afin d'attacher des séquences régulatrices ou insulatrices à des S/MARs (pour « *Scaffold/Matrix-Attachment-Region* ») (Kurukuti *et al.*, 2006, Yusufzai & Felsenfeld, 2004 mol cell) (Figure 52).

Par chance, Tsai *et al.* avaient fait une analyse bioinformatique du locus *Xist/Tsix* à la recherche de séquences ayant des caractéristiques de MARs (Tsai *et al.*, 2008).

En comparant nos données obtenues en ChIP pour CTCF, YY1 et Rad21 et en réalisant nous même l'analyse bioinformatique grâce au programme MAR-Wiz, nous avons pu nous apercevoir que le pic CTCF/Rad21 situé en amont de *Tsix*, colocalise parfaitement avec une

région soupçonnée d'être associée à la matrice nucléaire et qui, d'après Tsai *et al.*, serait conservée chez les humains (Figure 53).

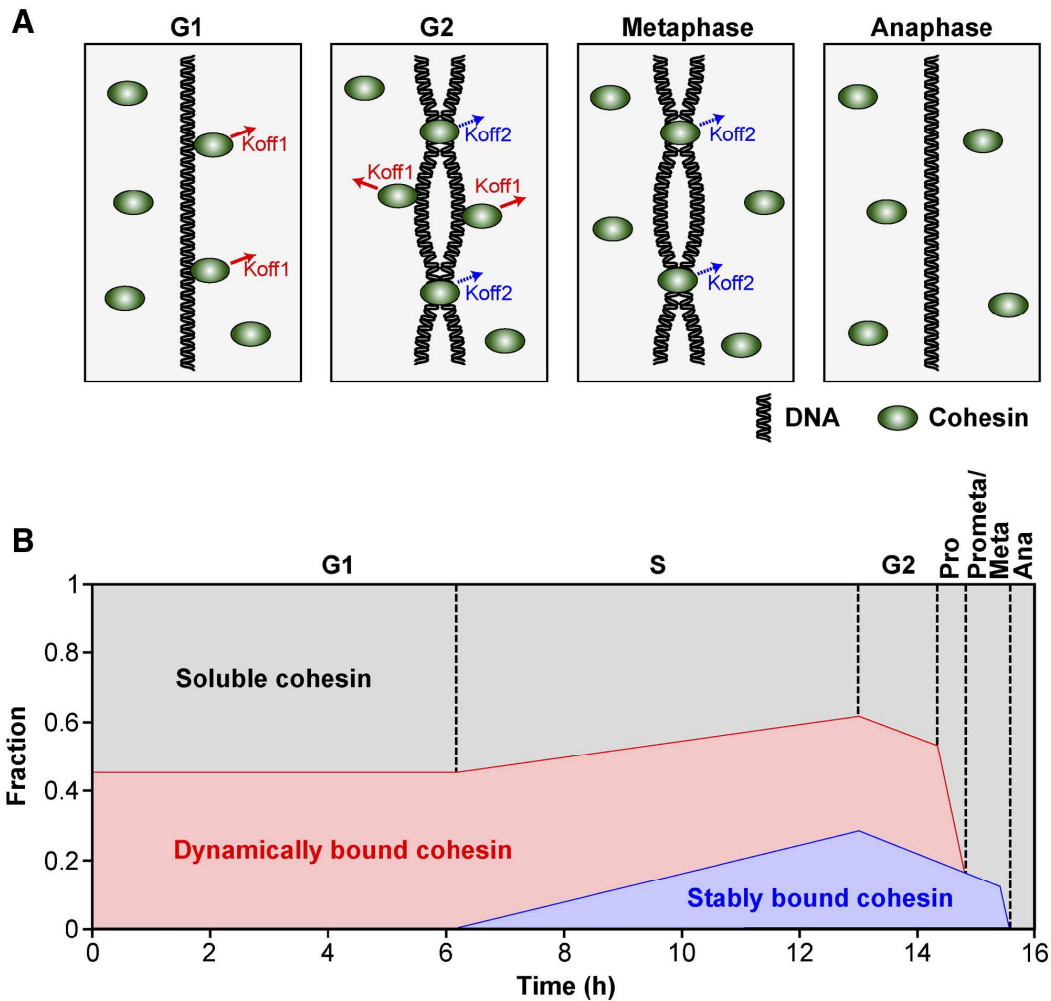


Figure 51 : Modèle de la fixation cycle-cellulaire dépendante de la cohésine à la chromatine. (A) Différents modes d'interaction entre la cohésine et la chromatine. En G1, une fraction de la cohésine se fixe dynamiquement à la chromatine avec un temps de résidence réduit, se dissociant au taux élevé de k_{off1} . En G2, une fraction additionnelle de cohésine se fixe à la chromatine avec persistance, se dissociant avec un taux faible de k_{off2} . De par le fait que les chromosomes métaphasiques contiennent exclusivement une cohésine stablement fixée, les auteurs spéculent que la fraction de complexes stablement fixés sont responsables du maintien de la liaison des chromatides sœurs jusqu'à rupture protéolytique au début de la ségrégation. (B) Quantités relatives de cohésine nucléaire durant le cycle cellulaire. La cohésine fixée de manière dynamique est présente sur la chromatine tout au long de l'interphase mais sur les chromosomes mitotiques. La cohésine attachée de manière stable apparaît post-réplication et persiste sur la chromatine mitotique jusqu'à l'anaphase. La cohésine est dégradée immédiatement avant l'amorçage de la ségrégation et est absent des chromosomes en anaphase. (D'après Gerlich *et al.*, 2006)

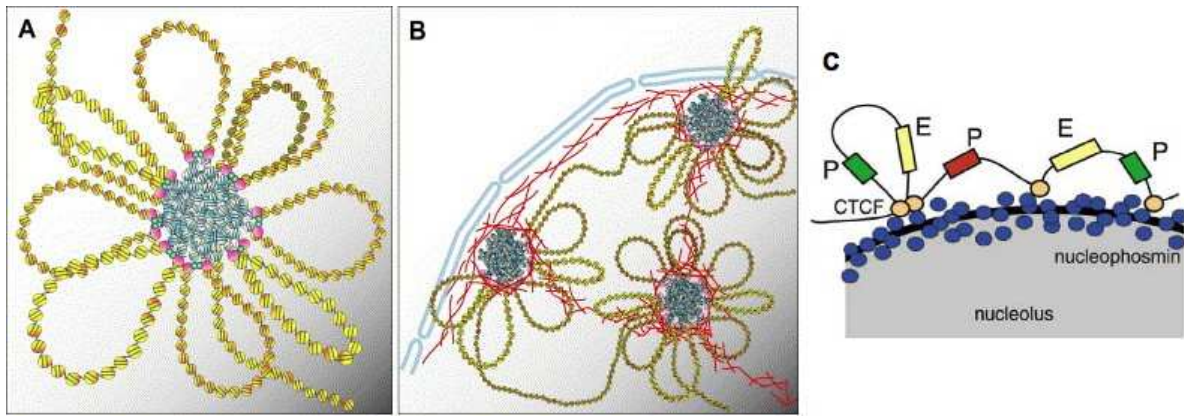


Figure 52 : Les insulateurs comme organisateurs de domaines de chromatine indépendants au sein du noyau. (A) et (B) Les interactions entre les éléments frontières (roses et vertes) ou leur interaction avec un substrat nucléaire fixé (rouge), forment les insulateurs qui peuvent participer dans la répartition structurale de la chromatine en domaines topologiques indépendants. L'indépendance fonctionnelle des domaines chromatiniens créés peut venir de la séparation physique d'éléments régulateurs localisés dans des domaines individuels, de l'association d'insulateurs avec des signaux activateurs ou répressifs au sein du noyau, ou du blocage de la propagation de la tension topologique. (C) Modèle de CTCF attachant une région insulatrice à la périphérie nucléolaire. CTCF s'attache à la partie externe du nucléole et fixe l'insulateur séparant ainsi l'amplificateur E des promoteurs (P). Seuls les promoteurs verts peuvent interagir avec l'amplificateur. (D'après Capelson & Corces 2004, et Yusufzai *et al.*, 2004)

De manière surprenante, il semblerait que le site de fixation pour Oct4/Nanog/Sox2 dans l'intron 1 soit aussi associé à la matrice nucléaire (et lui aussi conservé chez l'homme). Il serait intéressant de savoir si en cours de différenciation, ce site est encore associé à la matrice, sachant qu'une fois les facteurs de pluripotence détachés, il n'aurait à priori, plus aucune raison de se trouver ainsi localisé.

Enfin, il existerait un 3^{ème} MAR situé environ 4,5kb en amont de *Xist*, pour lequel nous ne possédons aucune indication quant aux protéines qui pourraient s'y fixer.

Ces 3 MARs espacés le long de *Xist/Tsix* pourraient donc être responsables de la formation de boucles chromatiniennes afin de rapprocher certaines séquences géniques ou régulatrices entre elles. Curieusement, seul le MAR situé en 5' de *Tsix* semble être lié par CTCF et Rad21, lui conférant peut-être des caractéristiques particulières par rapport aux deux autres.

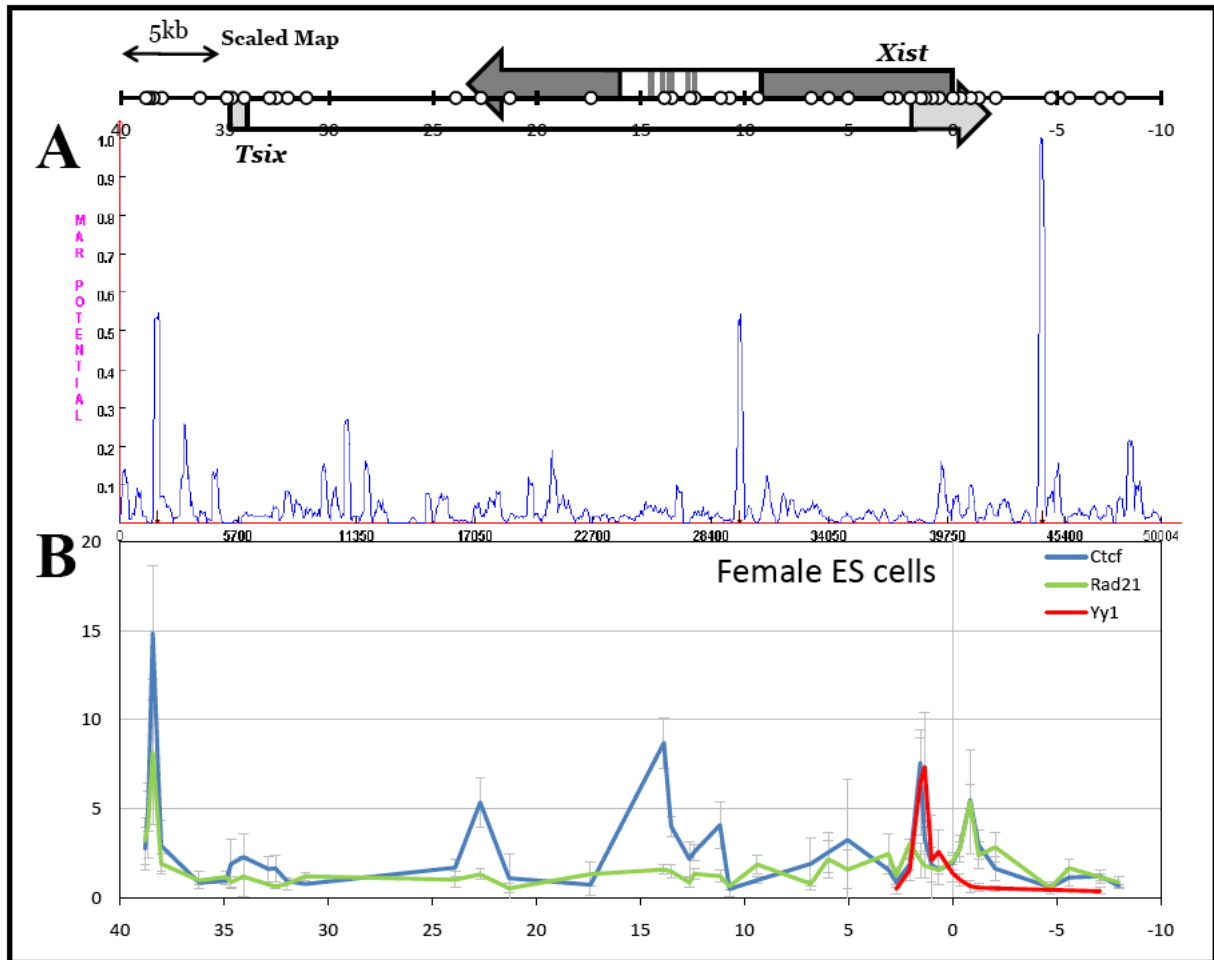


Figure 53 : Corrélation entre la présence de séquences potentielles de MARs et la fixation de CTCF au locus *Xist/Tsix*. (A) Prédiction de MARs le long de *Xist/Tsix* d'après le logiciel MAR-Wiz, qui tient compte, entre autre, des origines de réplication, la proportion de dinucléotides AT et TG, la courbure de l'ADN et la reconnaissance de la topoisomérase II. Largeur de fenêtre = 300, slide distance = 50, run length = 3. (B) Récapitulation des données obtenues concernant la fixation de CTCF, YY1 et Rad21 le long du locus *Xist/Tsix* dans des cellules ES femelles indifférenciées.

Une des protéines remarquée pour son rôle important dans la régulation des gènes au niveau des MARs est SatB1 (Alvarez *et al.*, 2000, Banan *et al.*, 1997, Kohwi-Shigematsu *et al.*, 1997). Cette protéine serait impliquée dans l'attachement de la partie basale de la boucle chromatinienne engendrée par la fixation des MARs à la matrice (Figure 54). Comme vu au cours de l'introduction (voir le chapitre : « l'ARN non-codant *Xist* », p25), SatB1 et son paralogue SatB2 sont aussi impliqués dans l'inactivation du chromosome X. Cette implication est visible lors de leur expression les 3 premiers jours après l'induction de la différenciation cellulaire, en tant que facteurs permettant l'initiation de l'inactivation et permettant le rétablissement d'un contexte embryonnaire quant à l'extinction génique par *Xist* dans les cellules somatiques (Agrelo & Wutz *et al.*, 2009). Il serait donc remarquable d'arriver à relier

l'expression de *SatB1*, permettant l'acquisition d'une fenêtre de temps afin que *Xist* initie l'inactivation, à un rôle dans l'attachement du locus *Xist/Tsix* à la matrice nucléaire dans le but de réguler ses gènes (CTCF ayant déjà été impliquée comme nécessaire dans la régulation génique de *SatB1* pour certains gènes, Ribeiro De almeida *et al.*, 2009). Ou bien que ce type d'interactions permette un traitement particulier de l'ARN *Xist* lui fournissant la capacité à inactiver transcriptionnellement les gènes du chromosome X.

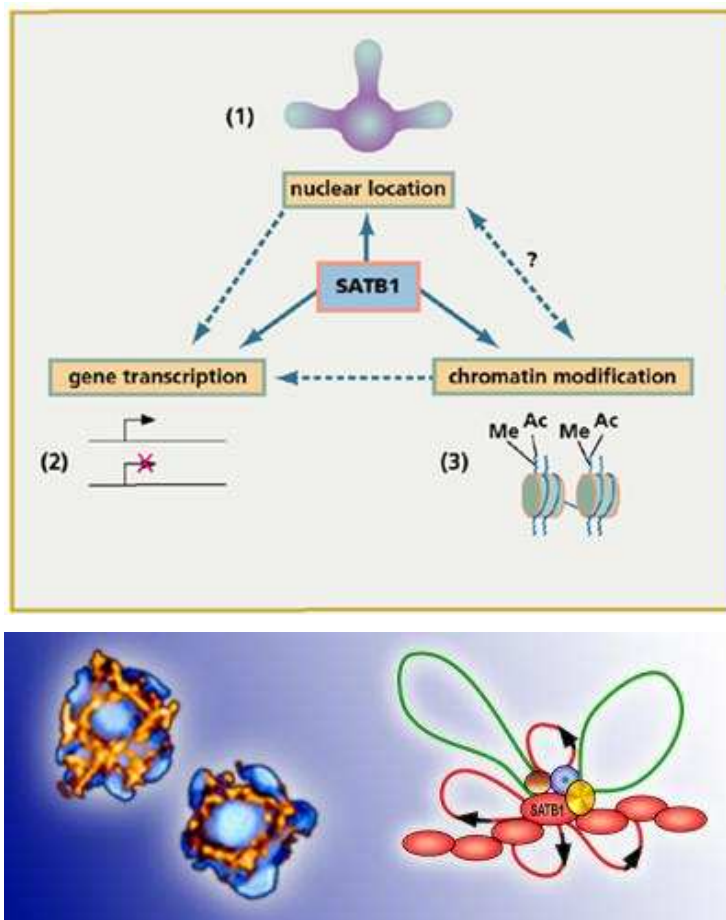


Figure 54 : L'implication de SatB1 la régulation épigénétique des gènes. (Haut) La protéine architecte du noyau SatB1 est clé dans plusieurs aspects de l'organisation nucléaire ainsi que dans plusieurs autres fonctions. SatB1 est localisé à des régions spécifiques du noyau en dehors de l'hétérochromatine et à la base des boucles formant des domaines chromatinien. (1) et (2) Elle est aussi nécessaire pour l'ancrage de gènes à un endroit spécifique. La localisation de ces gènes corrèle avec leur régulation correcte. (3) SatB1 affecte aussi l'état des modifications d'histones à leur site de fixation et aux gènes alentour. Ces modifications peuvent aussi bien avoir un effet activateur que répresseur. Cependant, il n'a pas encore été déterminé si les modifications chromatinien sont dues à SatB1-même ou à la localisation des régions à la base des boucles chromatinien. **(Bas)** SatB1 forme une cage tridimensionnelle au sein du noyau (à gauche) qui se fixe à des régions d'ADN spécifiques au sein de gènes puis réorganise la chromatine et recrute des enzymes régulatrices de la transcription (droite). (D'après Shannon, 2003 et <http://www.sciencedaily.com>)

xi. Conservation

Il apparaît clairement que, d'après les données générées par Mélanie Makhlouf, la fixation de CTCF, YY1 et Rad21 est conservée chez l'homme. Cependant, nous ne savons pas si le reste des pics CTCF fixés le long du gène *Xist*, mais surtout le pic CTCF/Rad21 en amont de *Tsix*, sont conservés. En effet, la séquence, ainsi semble-t-il que la fonction de l'antisens, sont très mal conservées entre la souris et l'homme (Migeon *et al.*, 2001). Qui plus est, les données de Chip-seq disponibles, concernant CTCF et RAD21 ne proviennent que de cellules somatiques, dans lesquelles nous ne nous attendons pas à voir ce pic s'il y a bien conservation. Cribler tout le locus de *XIST/TSIX* manuellement serait trop couteux en temps et en main d'oeuvre. Toutefois, d'après nos observations effectuées chez la souris, il ne semble pas que cette fixation soit reliée à la régulation de *Tsix*. De plus, Tsai *et al.* observent la conservation chez l'homme du MAR colocalisant avec le pic CTCF/Rad21 en 5' de *Tsix* chez la souris, nous fournissant ainsi une indication pour rechercher ce site, s'il est bel et bien conservé.

Dans tous les cas, la conservation des sites de fixation de YY1 et du pic CTCF/Rad21 en 5' de *Xist* chez autant d'espèces différentes, nous fournit des raisons de privilégier YY1 comme activateur transcriptionnel majeur de *Xist* et nous indique que quel qu'il soit, le rôle de cette fixation de CTCF/Rad21 est d'une importance capitale.

xii. Conclusion

Toutes les informations obtenues concernant le réseau régulateur du gène *Xist* au cours de ces 4 dernières années nous ont permis de proposer un modèle (Figure 55) au sein duquel de nombreux éléments restent à confirmer ou à découvrir, mais qui fournit les bases pour une caractérisation des mécanismes permettant la régulation de *Xist* (plus particulièrement sa surexpression) au cours de différentes étapes du développement embryonnaire post-implantatoire.

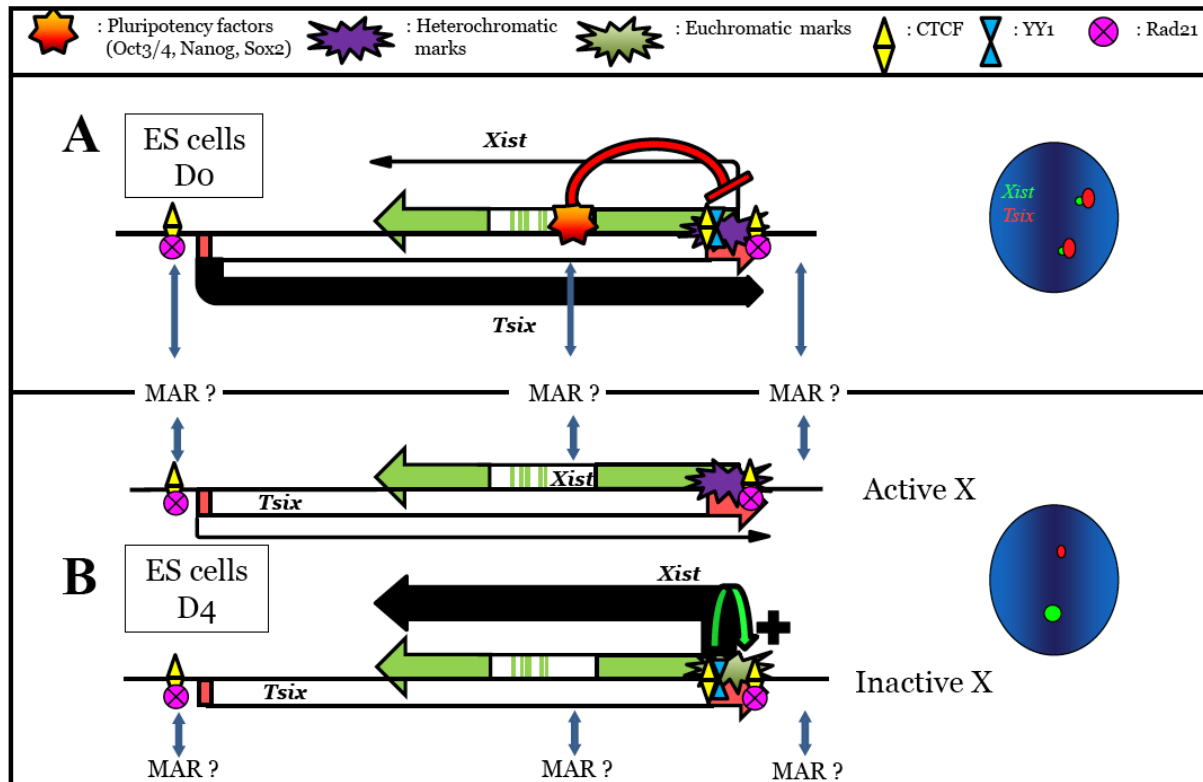


Figure 55 : Modèle schématique représentant les différents phénomènes au sein du locus *Xist/Tsix* influant sur la régulation de *Xist*. (A) Au sein des cellules ES indifférenciées, les deux chromosomes sont identiques quant aux caractéristiques représentés ici (à part un recrutement moins important de Rad21 sur le futur Xi), d'où la représentation d'un chromosome X unique. Dans ces cellules, *Xist* est réprimé par une alliance entre *Tsix* (post-transcriptionnellement et en médiant un domaine hétérochromatique à son promoteur, en violet) et les facteurs de pluripotence (répression transcriptionnelle, étoile orange). On observe la fixation de CTCF (diamant jaune) et YY1 (sablier bleu) (activateurs transcriptionnels supposés) en aval du promoteur de *Xist* de manière anticipée, prêts à jouer leur rôle dès l'initiation de la différenciation et la disparition des mécanismes répressifs. Le site CTCF et Rad21 (rond rose) en amont de *Tsix* correspond potentiellement à un MAR et pourrait donc potentiellement être impliqué dans la localisation nucléaire du locus ainsi que dans sa régulation (éventuellement en collaboration avec deux autres MARs : un à l'intron 1, l'autre en amont de *Xist*). Pour l'instant nous ignorons le rôle du pic en amont de *Xist*. (B) Après quatre jours de différenciation cellulaire des cellules ES femelles à l'acide rétinoïque, on assiste à une disparition de l'expression de *Tsix* (légèrement maintenue sur le Xa) ainsi que de la fixation des facteurs de pluripotence de manière biallélique (bien que la cinétique exacte des phénomènes à chaque allèle pourrait jouer un rôle déterminant dans le choix du chromosome X à inactiver). Ceci engendre l'apparition d'un domaine euchromatique (vert foncé) au promoteur de *Xist*, ainsi que sa surexpression médiée par CTCF et surtout YY1, ces phénomènes n'étant observables que sur le chromosome ayant été choisi comme inactif. Le chromosome actif maintient sa structure hétérochromatique au promoteur de *Xist* et n'est plus fixé par CTCF et YY1 dû à un mécanisme encore inconnu. Après quatre jours de différenciation, les pics CTCF/Rad21 sont toujours présents de part et d'autre du locus *Xist/Tsix*. Cependant au sein des cellules somatiques, on observe la disparition de celui se trouvant en amont de *Tsix*, nous indiquant que son rôle est cantonné à une fenêtre de temps parallèle à celle de la mise en place de l'inactivation du chromosome X.

À travers la recherche de rôles pour ces différentes protéines au sein du locus *Xist/Tsix*, il m'apparaît comme de plus en plus évident qu'il est très difficile de cantonner des protéines telles que CTCF ou YY1 à un rôle précis, tel qu'insulateur ou répresseur transcriptionnel,

respectivement. La multitude d'interactions dont elles bénéficient en fonction de la séquence nucléotidique, du contexte cellulaire et épigénétique, ainsi que de leurs propres modifications post-traductionnelles, font de ces protéines de véritables couteaux suisses (à lames de zinc). Et c'est dans cet état d'esprit que je trouve la proposition de l'existence d'un « code CTCF » par Ohlsson, Lobanenko et Klenova particulièrement judicieuse (Ohlsson *et al.*, 2010).

II. La caractérisation fonctionnelle des sites de fixation pour les protéines CTCF et YY1 dans la région promotrice de *Xist* chez la souris

a. Introduction

Il semblerait donc que le maintien en cours de différenciation de la fixation de CTCF et YY1 à un site situé environ 1,5kb en aval du promoteur P1 de *Xist* soit nécessaire à l'activation transcriptionnelle de ce dernier. Curieusement, ces deux protéines ont déjà été impliquées dans un rôle similaire dans la régulation de l'expression de l'antisens *Tsix* (Chao *et al.*, 2002, Donohoe *et al.*, 2007). En effet, comme nous l'avons vu au cours de l'Introduction (voir le chapitre : « *DxPas34* », p59), CTCF et YY1 pourraient agir en hétérodimère lors de leur recrutement à la région répétée *DxPas34*. Ceci semble en effet plutôt paradoxal considérant que cette région répétée a été décrite comme étant un « enhancer » de *Tsix*, signifiant qu'elle est partiellement responsable de l'activation transcriptionnelle de ce gène (Vigneau *et al.*, 2006, Cohen *et al.*, 2007). CTCF et YY1 possèderaient donc une fonction activatrice de *Tsix* dans les cellules ES (ainsi qu'en début de différenciation sur le X actif) mais aussi sur *Xist*, qui est censé être réprimé par son anti-sens (Navarro *et al.*, 2005).

Lors d'expériences d'ARNi (un type d'expérience classiquement utilisé pour observer l'effet de l'absence de fixation d'une protéine à un certain locus) la diminution intracellulaire des protéines visées influe sur leur fixation à tous leurs sites. Dans notre cas, cela signifie qu'en déplaçant la cellule de CTCF ou de YY1, on empêche leur fixation aussi bien dans la région promotrice de *Xist* qu'à celle de *Tsix*, rendant donc impossible le fait de départager leur fonction à ces deux sites. Théoriquement, en déplaçant ces protéines dans des cellules ES on devrait assister à une répression de *Tsix* engendrant une surexpression de *Xist* (similaire à celle observée dans les lignées cellulaires Ma2L, $\Delta Dxpas34$, ou $\Delta 65$, Vigneau *et al.*, 2006, Morey *et al.*, 2004). Cependant, la déplétion de CTCF ou de YY1 engendre aussi leur absence au promoteur de *Xist*, l'incapacitant potentiellement à être surexprimé. Et c'est ce qui est observé expérimentalement par le laboratoire de Jeannie Lee, que ce soit en utilisant des siRNA anti-YY1 ou anti-CTCF on encore en générant des souris *YY1*^{+/-} (Donohoe *et al.*, 2007).

Lors des expériences de siRNA au sein de cellules ES mâles, Donohoe *et al.* observent effectivement une diminution de l'expression de *Tsix* (d'un facteur 2 à 3 par interférence mais de 12,5 par mutation) ainsi qu'une augmentation minime de l'expression de *Xist* (d'un facteur

3 à 7). Cette faible surexpression de *Xist* pourrait être due à deux phénomènes non exclusifs : soit la quantité de transcription antisense est encore trop élevée, soit l'absence de CTCF et de YY1 au promoteur de *Xist* ne lui permet pas d'être surexprimé (indépendamment des facteurs de pluripotence qui sont aussi présent dans Ma2L par exemple, mais où l'on observe une surexpression d'un facteur 100).

Les expériences ayant été menées par Vigneau *et al.* peuvent éventuellement nous apporter un éclaircissement (Vigneau *et al.*, 2006). L'effet des mutants $\Delta Dxpas34$ de Vigneau *et al.* sur l'expression de *Xist* (augmentation d'un facteur 7) et de *Tsix* (diminution d'un facteur 10) dans les cellules mâles indifférenciées est comparable à celui observé dans les mutants *YY1*^{+/-}, bien qu'au sein de ces cellules CTCF et YY1 sont encore fixés au promoteur de *Xist* (Figure 56). Ceci tendrait à favoriser l'hypothèse d'une transcription résiduelle encore trop élevée dans les expériences de Donohoe *et al.* pour expliquer la faible surexpression de *Xist* et nous indique donc que les 10% de transcrits résiduels sont suffisants pour empêcher le passage de la surexpression de *Xist* d'un facteur 7 à un facteur 100.

Par contre en cours de différenciation, bien que les systèmes censés réprimer *Xist* disparaissent graduellement dans tous les types de mutants, Vigneau *et al.* observent une surexpression de *Xist* permettant même d'obtenir des domaines de transcrits ectopiques (dans une faible proportion de cellules), tandis que Donohoe *et al.* observent une extinction totale de l'expression de *Xist* dans leurs mutants *YY1*^{+/-} malgré l'absence de l'antisens et des facteurs de pluripotence, soulignant ainsi l'importance de YY1 dans la surexpression de *Xist* en cours de différenciation.

Ces résultats ajoutés aux nôtres (présentés au chapitre précédent) et à ceux obtenus par Kim *et al.* lors d'un test à la luciférase (Kim *et al.*, 2006, voir le chapitre : « Les activateurs », p94) (Figure 47), nous permettent donc d'envisager avec une grande confiance l'importance capitale de YY1 dans la régulation transcriptionnelle de *Xist*. Toutefois, afin de s'en assurer définitivement, nous avons choisi d'entreprendre la seule étude permettant de véritablement démontrer ce lien ainsi que de nous fournir une idée des rôles joués par CTCF et la cohésine au sein de la région promotrice de *Xist* : la mutation ponctuelle de leurs sites endogènes.

Pour réaliser ce projet, nous avons, notamment grâce à l'expertise du Dr. Philippe Clerc dans le domaine de la biologie moléculaire, entrepris d'utiliser une méthode appelée R.M.C.E. (pour « *Recombinase-Mediated Cassette Exchange* ») (Bouhassira *et al.*, 1997) (Figure 57).

Le principe de cette technique consiste à échanger une région d'intérêt par une version mutée grâce à l'utilisation du phénomène de recombinaison *Lox-Cre* (pour « *Causes REcombination* »)

(Sauer, 1987). En l'occurrence, nous avons choisi comme région d'étude l'exon 4 de *Tsix* (correspondant à la région promotrice de *Xist*).

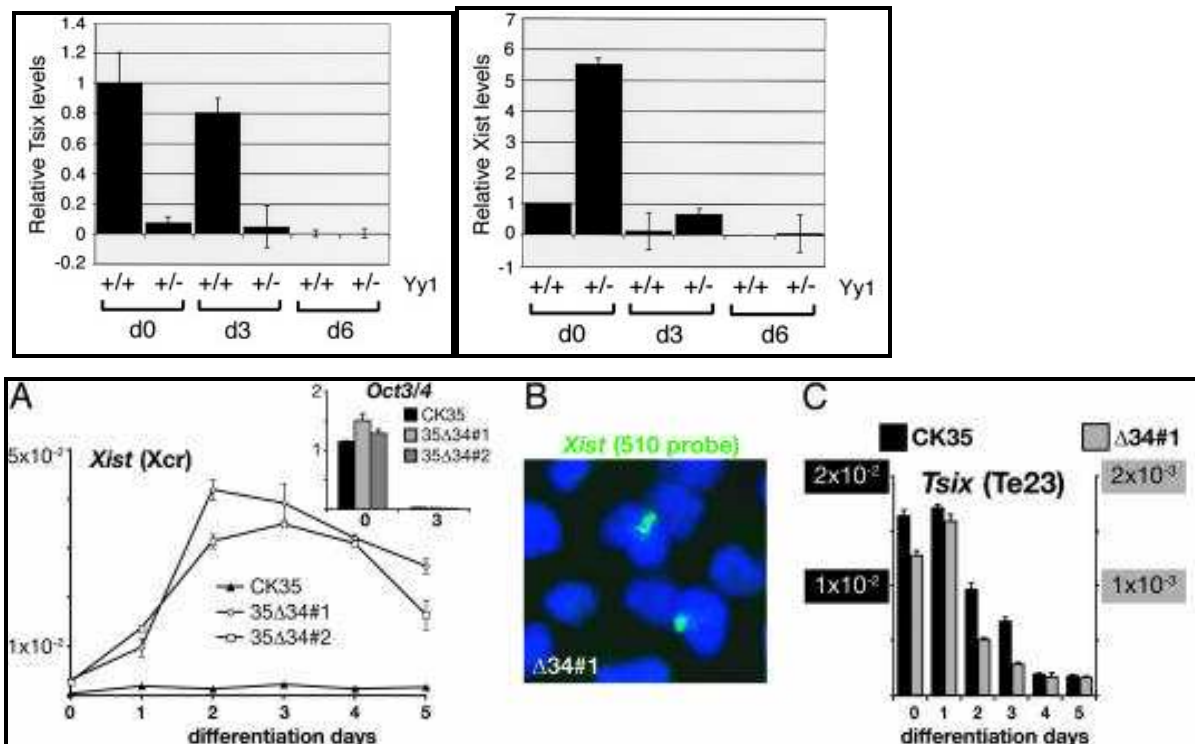


Figure 56 : Effet d'une réduction partielle de l'expression de *Tsix* sur celle de *Xist*. (Haut) Analyse de RT-PCR quantitative de l'expression de *Xist* et *Tsix* dans des clones ES mâles *YY1*^{+/+} ou *YY1*^{+/-} par amplification en temps réel (Biorad). Les cellules ES ont été différenciées entre 0 et 6 jours, tel qu'indiqué. (Bas) La délétion de *Dxpas34* engendre une surexpression ectopique de *Xist* dans les cellules mâles. (A) Analyse par qRT-PCR de l'expression de *Xist* normalisée à celle d'*Arpo* dans des cellules différenciées à l'acide rétinoïque. L'expression d'*Oct4* a aussi été analysée de manière similaire pour contrôler la différenciation cellulaire. (B) ARN-FISH de *Xist* (en vert) sur des noyaux colorés au DAPI (en bleu) après 3 jours de différenciation cellulaire. Des domaines ectopiques de *Xist* sont présents dans 15% des cellules délétées pour *Dxpas34* mais n'ont jamais été observés dans les cellules sauvages. (C) Analyse par qRT-PCR des transcrits *Tsix* épiciés normalisés par rapport à *Arpo* au sein de cellules ES différenciées à l'acide rétinoïque. (D'après Donohoe *et al.*, 2007 et Vigneau *et al.*, 2006)

Dans un premier temps, il est nécessaire d'intégrer une cassette de sélection entourée de sites *Lox* au locus que l'on souhaite étudier. Ceci est réalisé par la transfection d'un vecteur plasmidique contenant des bras d'homologie et des sites *Lox* de part et d'autre de la cassette de sélection. Ensuite, une fois la recombinaison homologue validée, il est aisé d'échanger la cassette de sélection par une version modifiée de la région promotrice. Cela permet d'observer l'effet des modifications apportées à son locus endogène, dans un contexte cellulaire, grâce aux sites *Lox*. Lors de la construction de notre plasmide, nous avons choisi d'insérer ces sites *Lox* en sens inverse l'un par rapport à l'autre afin de favoriser les interactions entre les sites insérés au génome et ceux du plasmide, plutôt que de les laisser

dans le même sens, ce qui aurait favorisé l'excision de la séquence contenue entre les deux sites insérés dans le génome.

Notre avons aussi choisi d'employer une cassette de sélection contenant des gènes qui nous permettrons d'effectuer une sélection négative (Thymidine Kinase, rendant les cellules sensibles au Gancyclovir) mais aussi une sélection positive (par la Puromycine). La sélection positive est nécessaire au cours de l'étape de sélection de clones recombinants après la recombinaison homologue initiale. La sélection négative servira lors de la réinsertion d'une séquence modifiée, pour s'assurer que l'on a bien remplacé la cassette de sélection par la région promotrice de *Xist* mutée. De plus, des sites *FRT* (pour « *Flippase Recognition Target* ») ont aussi été ajoutés de part et d'autre de la cassette de sélection afin de pouvoir l'exciser et ainsi étudier l'effet de la délétion de la séquence promotrice de *Xist* (ou l'exon 4 de *Tsix*).

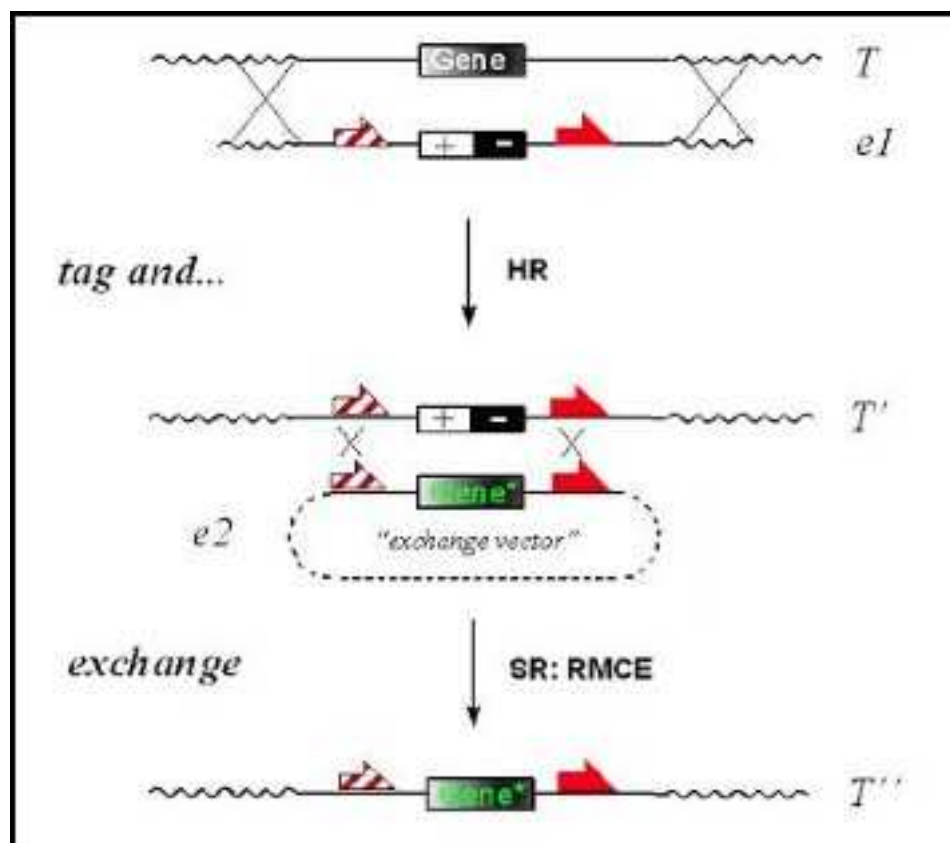


Figure 57 : Stratégie générale de *Recombinase-Mediated Cassette Exchange*. Cette stratégie permet le remplacement d'un gène par sa copie modifiée en 2 étapes : une première étape de recombinaison homologue afin d'insérer les sites (généralement *Lox* ou *FRT*) qui serviront aux recombinaisons suivantes ainsi qu'une cassette de sélection. Et une recombinaison "facilitée" médiée par les sites préalablement insérés permettant l'insertion de la version modifiée de son gène. (D'après Oumard *et al.*, 2006)

Une fois la recombinaison homologue effectuée, il est donc aisé d'échanger la région promotrice de *Xist* par une autre version de celle-ci contenant des mutations des sites de fixation des protéines, et ceci, simplement en transfectant cette nouvelle version du promoteur ainsi qu'en induisant l'expression de la recombinaise : la *Cre*. Comme vu lors de la définition du projet de thèse, l'objectif de cette expérience est de muter individuellement ou collectivement les sites des protéines CTCF et YY1 dans la région promotrice de *Xist* et ainsi pouvoir visualiser leur véritable effet fonctionnel sur son expression mais aussi sur le rôle de CTCF et la cohésine dans la structure chromatinienne tridimensionnelle du locus *Xist/Tsix*.

Afin de pouvoir étudier un effet de ces mutations sur l'inactivation tout en observant les variations de fixation des protéines de manière allélique, il a été décidé d'effectuer la recombinaison homologue du vecteur plasmidique dans des cellules ES femelles de la lignée *PGK1*. Ces cellules sont originellement de la lignée *129* mais ont la particularité de posséder un *XIC* d'origine *PGK1a* (l'autre allèle étant donc *129*) permettant ainsi une étude allélique de la fixation des protéines ainsi que de l'expression des gènes grâce à de nombreux polymorphismes existant entre les deux fonds génétiques. Puis, cette lignée cellulaire étant particulièrement résistante à la recombinaison homologue, nous avons tenté d'induire ce phénomène dans des cellules *EL16.7* généreusement fournies par Jeannie Lee, contenant un chromosome X d'origine *129* et un d'origine *Castaneus*, cette lignée étant censée être plus facile à recombiner.

b. Le vecteur

La partie centrale du vecteur est composée des deux gènes de sélection, ceux-ci étant entourés des sites *Lox* en sens inversés et *FRT* dans le même sens l'un par rapport à l'autre (Figure 58). Cependant, je tiens à signaler ici, que bien que j'ai intégralement réalisé la conception de cette construction moléculaire (avec l'aide du Dr. Philippe Clerc), Sylvain Maenner (du laboratoire de Maturation des ARNs de la Faculté des Sciences et Techniques de Nancy) a entamé la construction de cette partie du plasmide en rajoutant les sites *Lox* et *FRT* de part et d'autre des cassettes, elles-mêmes obtenues à partir d'un autre plasmide qui nous a aimablement été fourni par Sharaghim Tajbakhsh de l'Institut Pasteur. Une fois la cassette de double-sélection entourée des sites de recombinaison grâce à une amplification par PCR, cette construction fut clônée au sein du plasmide pCR2.1 de chez invitrogen. J'ai ensuite entrepris d'insérer les bras d'homologie par digestion/ligation, le bras gauche mesurant approximativement 4,7kb et le bras droit 8kb, ceux-ci ayant été amplifiés par PCR à partir du BAC 399K20 (Chureau *et al.*, 2002) provenant d'un fond génétique *129*.

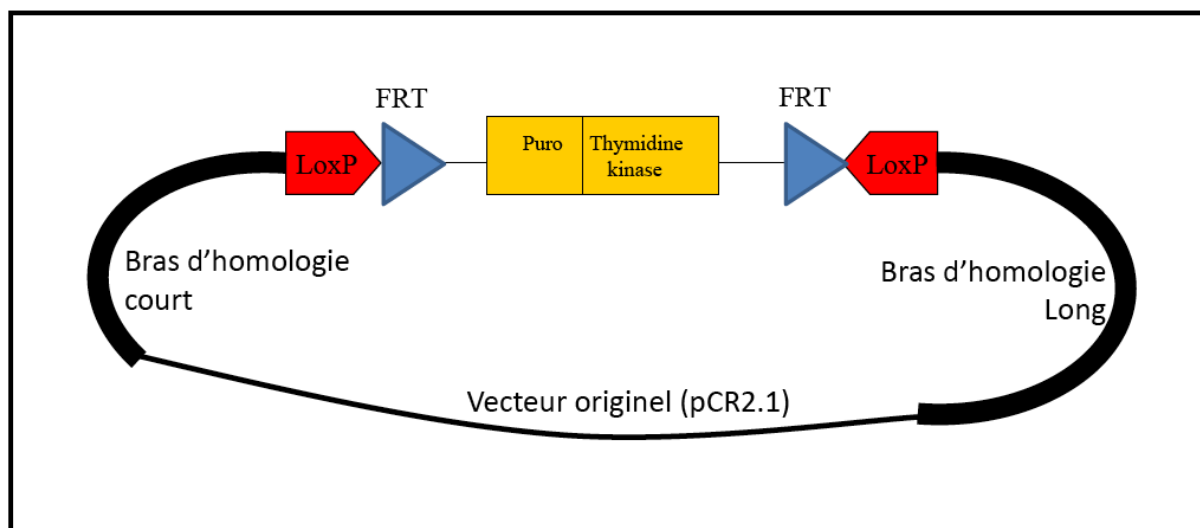


Figure 58 : Schéma du vecteur servant à introduire les sites de recombinaison au locus endogène. Les séquences LoxP servant à la réinsertion du gène muté sont représentés en rouge. Les sites FRT servant à l'étude d'un délétant de la région promotrice de Xist sont en bleu. En jaune, la double-cassette de sélection. En noir épais les bras d'homologie et en noir fin la séquence originelle du plasmide.

Ce clonage, réalisé en plusieurs étapes, a nécessité le criblage d'un nombre important de clones à chacune d'entre elles sans doute à cause de la taille importante des fragments à insérer. En effet, le plasmide de base mesurait 3,9kb et en définitive avoisine les 18,5kb, ce qui approche de la limite maximale pour la taille d'un plasmide que l'on peut aisément manipuler. Cependant au cours de l'analyse par profils de digestions, il s'est avéré que tous les clones ayant correctement inséré les deux bras d'homologie contenaient une délétion d'environ 1 à 1,5kb correspondant à une partie du *Repeat-C* (présent dans le bras droit). Toutefois, nous avons construit en parallèle un autre plasmide identique au précédent (mis à part le bras gauche qui mesurait 400pb de plus), nous permettant ainsi de mettre au point les conditions de criblage des clones recombinés par PCR. Et il s'est avéré que non seulement ces clones ne contenaient pas de délétion dans le bras droit mais de plus, il existait des sites de digestion capables d'engendrer un bras gauche plus court (d'environ 1kb donc mesurant environ 4kb au final) (Figure 59) qui pourrait servir à la recombinaison homologue. Et c'est donc avec ce vecteur là que nous avons entamé une série d'électroporations en espérant obtenir des cellules ES ayant incorporé la cassette à l'endroit désiré.

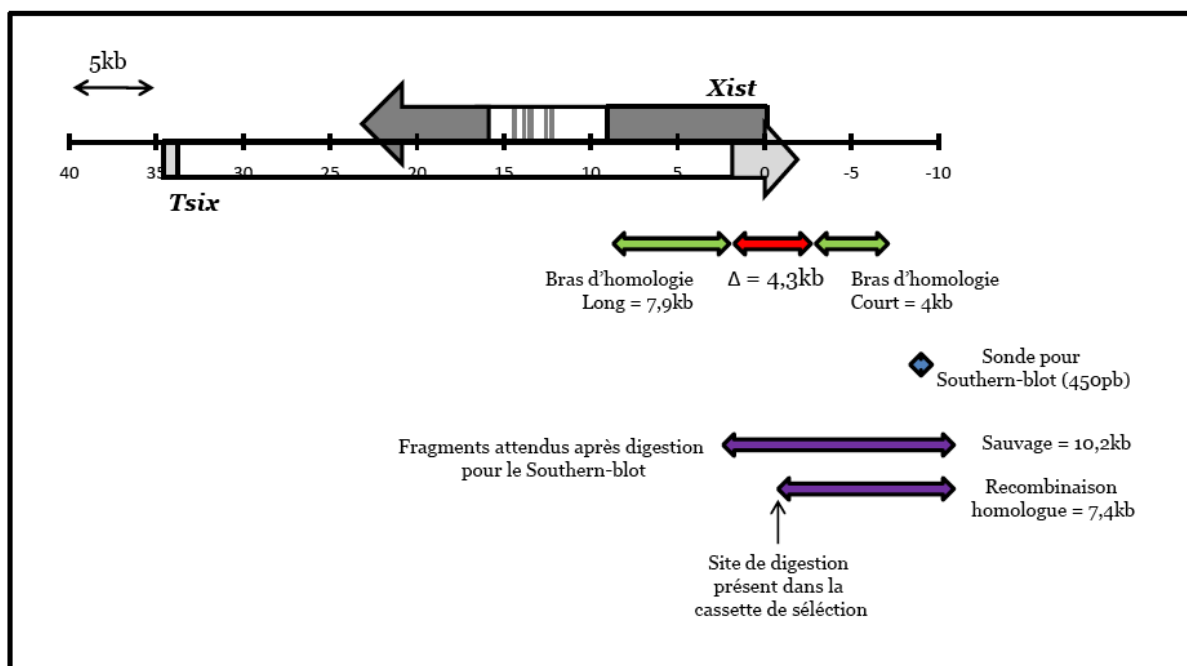


Figure 59 : Schéma des régions du locus *Xist/Tsix* impliquées dans la construction. En vert sont représentés les bras d'homologie servant à la recombinaison homologue. En rouge, on distingue la région délétée, remplacée dans un premier temps par la cassette de sélection, puis par sa version modifiée. En bleu, la séquence correspondant à la sonde utilisée pour le criblage par Southern-blot. En violet, sont représentés les fragments d'ADN pouvant être hybridés par la sonde suite à une digestion de l'ADN génomique par l'enzyme de restriction *SphI*.

c. L'électroporation

Dans le dessein d'obtenir un clone correctement recombiné plusieurs méthodes d'électroporation ont été utilisées (Figure 60).

Méthodes d'électroporation	Protocole 1	Protocole 2	Protocole 3	Protocole 4
Nombre de cellules	1.10^7	$1,5.10^7$	2.10^7	$1,25.10^7$
Quantité de plasmide linéarisé ($\mu\text{g}/\mu\text{l}$)	15	20	20	15
Appareil utilisé	Biorad	Biorad	BTX	BTX
Voltage (V)	200	240	700	700
Durée (ms)	37	12	0,125	0,125
Capacitance (μF)	1500	800	-	-
Nombre de colonies obtenues	+	+++	+++	++

Figure 60 : Tableau récapitulant les différents protocoles d'électroporation utilisés.

Les protocoles 1 et 2 m'ont été conseillés par des collègues de l'Institut Pasteur (D. Rocancourt et P. Clerc, respectivement), le protocole 3 correspond à un protocole utilisé par l'Institut Sanger à Cambridge modifié et le protocole 4 représente le protocole exact de l'Institut Sanger (l'électroporation doit aussi être faite dans du PBS et non de l'eau).

d. Le criblage

Suite à l'électroporation, il est nécessaire d'effectuer une sélection grâce à l'ajout d'un antibiotique (ici, la puromycine, utilisée à 1µg/ml), nous permettant de discriminer entre les cellules ayant ou n'ayant pas intégré les cassettes de sélection entourées des sites de recombinaison au sein de leur génome. Cependant, nous n'avons rajouté la puromycine qu'entre 30 et 40h après la transfection afin de laisser les cellules adhérer correctement et commencer à croître dans un environnement non stressant.

Après 8 jours de culture en présence de milieu de sélection, les clones sont d'une taille suffisamment grande pour être repiqués.

Notre approche de criblage a été basée sur un test PCR, en employant des amorces qui se fixent d'une part en amont du bras d'homologie gauche et d'autre part au sein de la cassette de sélection, nous fournissant ainsi une première idée de si la recombinaison homologue a eu lieu au bon locus. Pour ce faire, dans un premier temps, nous avons entrepris d'utiliser une méthode de regroupement (« *pooling* ») des clones : prenant un aliquot des cellules et le mélangeant à 7 autres clones, tout en les repiquant individuellement dans des plaques 96 puits. Une PCR de criblage est ensuite réalisée sur le *pool* de cellules directement issues du clone. Cette méthode a l'avantage de réduire le nombre d'échantillons à tester (grâce à la technique de regroupement) et permet également d'obtenir une première réponse quant à la validité des clones avant même de les expandre ou de les congeler. Cependant l'inconvénient majeur de cette approche est que l'on ignore si on teste véritablement tous les clones, l'utilisation d'un témoin positif de PCR (l'amplification d'une région éloignée de la zone supposée de recombinaison) nous indiquant seulement si parmi les 8 clones présents dans le *pool*, au moins l'un d'entre eux est amplifiable. Afin d'améliorer notre sensibilité de détection et donc de s'assurer que nous ne rations pas de clones positifs parmi les *pools*, nous avons ensuite transféré les résultats de PCR sur une membrane de nitrocellulose afin de l'hybrider à l'aide d'une sonde radioactive (de 450kb, située entre l'amorce 5' de criblage et le début du bras d'homologie gauche) censée s'accrocher aux amplicons PCR correspondant à un clone présentant une recombinaison correcte.

L'ensemble des clones du *pool* montrant un marquage par hybridation d'un produit correspondant approximativement à la taille attendue est ensuite cultivé de manière à obtenir suffisamment de cellules pour effectuer de nouveau les mêmes tests (criblage PCR suivi d'une hybridation radioactive) sur les clones individuels, nous permettant d'identifier le clone positif parmi tous ceux du *pool* et de confirmer sa validité. Suite à cette deuxième étape, les clones suspectés comme étant bons sont cultivés en grande quantité, rendant ainsi possible

l'extraction d'une quantité suffisante d'ADN pour effectuer un troisième type de criblage : le « *Southern-blot* ». L'enzyme employée pour digérer l'ADN génomique des clones suspectés comme étant positifs est SphI. La révélation de l'hybridation de la sonde radioactive, suite au transfert de l'ADN génomique digéré sur une membrane de nitrocellulose, devrait montrer un fragment à une distance de migration correspondant à 7,4kb en cas de clone positif et à 10,2kb en cas de clone négatif (Figure 59).

Toutefois, au fur et à mesure des électroporations et des criblages, nous avons opté pour une méthode nous assurant du criblage de chaque clone individuel en passant par une étape d'expansion, puis de congélation pour tous les clones. Pendant que les cellules étaient congelées nous avons donc le temps de tester chaque clone par PCR, à l'aide des amorces de criblage mais aussi avec des amorces constituant un témoin positif afin de s'assurer de bien pouvoir amplifier l'ADN extrait à partir de chacun d'entre eux. Cette méthode nécessite beaucoup plus de temps et de travail, mais en contrepartie est bien plus fiable.

e. Le bilan

Finalement, après 6 électroporations utilisant diverses méthodes et près de 2700 clones criblés, nous n'avons jamais été en mesure de valider un clone positif.

Ce résultat peut paraître surprenant si l'on considère que le taux de réussite d'une recombinaison homologue chez la souris doit se situer, en moyenne, entre 0,5 et 5%.

Cependant je pense qu'il existe plusieurs facteurs pouvant expliquer ce manque de réussite :

- Tout d'abord, le criblage par regroupement de clones ne me semble pas être une méthode très fiable. Bien que nous ayons réalisé des tests de « dilutions-limites » d'ADN matrice et comparé les résultats obtenus à la proportion de plasmide intégré par rapport à la quantité d'ADN génomique totale contenue dans une cellule afin de vérifier la robustesse des amorces utilisées, il est impossible de s'assurer de la capacité d'amplification de l'ADN de tous les clones mis en commun au sein d'un même *pool*. Par contre, en testant individuellement chaque clone, on est sûr de savoir s'il est positif ou négatif. Malheureusement, le test individuel des clones n'a été mis en place que pour la deuxième moitié des clones criblés.
- Ensuite, le choix du programme d'électroporation constitue un point plus controversé. Comme je l'ai présenté précédemment (Figure 60), nous avons utilisé 3 programmes différents, tout en variant aussi le nombre de cellules et la quantité d'ADN plasmidique linéarisé. L'inconvénient de cette méthode est que chaque personne l'ayant déjà employée, préconise l'utilisation de son programme "fétiche",

celui-ci ayant donné des résultats probants lors de son emploi dans des conditions bien particulières, mais en définitive, il n'existe aucune étude scientifique (à ma connaissance) portant sur les conséquences de différentes modifications d'un programme d'électroporation sur la capacité à induire une recombinaison homologue correcte. Une question intéressante se pose à ce sujet : la présence d'un faible nombre de clones suite à l'électroporation est-elle un signe positif de la réussite de la recombinaison ? Car on pourrait s'imaginer que plus la mortalité cellulaire est importante et plus la recombinaison a été difficile à réaliser, nous donnant donc à penser que parmi les clones ayant survécu, on a une forte chance d'en trouver des positifs. D'autant plus que l'on favorise la possibilité d'une recombinaison homologue au locus souhaité à travers la présence de longs bras d'homologie. Mais l'inverse peut aussi être vrai, si l'on estime que plus le nombre de clones ayant subi une intégration de la cassette de sélection à son génome est important, plus on a de chances pour que parmi ces clones, quelques-uns l'aient intégré au bon endroit.

- Et enfin, une dernière source de problème pourrait se situer au niveau du choix de la lignée cellulaire utilisée. En effet, pour des raisons que l'on ignore, il semblerait que certaines lignées soient plus résistantes à la recombinaison homologue que d'autres. La lignée majoritairement utilisée au cours des expériences a été la lignée PGK1. Cette même lignée a déjà été utilisée par le laboratoire de Neil Brockdorff lors de la construction du délétant $\Delta Xist$ (Penny *et al.*, 1996). La différence majeure par rapport à notre expérience est la taille des bras d'homologie, qui sont bien plus petits (1,8 et 2,7kb). Cependant le programme d'électroporation ainsi que les quantités de cellules et d'ADN sont similaires ($1,5 \cdot 10^7$ cellules, 15µg d'ADN, 500V, Biorad). Et d'après l'article dans lequel ces données ont été publiées, un seul clone recombinant au bon locus sur 2588 criblés aurait été obtenu, indiquant donc la résistance particulière de ces cellules à la recombinaison homologue. Et c'est dans l'objectif d'améliorer ce rendement que nous avons ensuite changé de lignée cellulaire (pour les 600 derniers clones) en utilisant la lignée de cellules ES femelles *EL16.7* (associées au protocole 2, Figure 60), composées d'un allèle du chromosome X de la lignée 129 et d'un allèle *Castaneus* (de *Mus castaneus*), nous permettant ainsi de conserver l'approche d'analyse à l'aide de polymorphismes (Xu *et al.*, 2006 ; Stavropoulos *et al.*, 2005 ; Lee, 2005 ; Shibata & Lee, 2004 ; Ogawa & Lee, 2003 ; Lee *et al.*, 1999). Bien que dans ses articles, Jeannie Lee n'indique que très rarement le rendement des recombinaisons homologues, dans l'un d'entre eux (Lee *et al.*, 1999), elle affirme

avoir obtenu 3 clones positifs sur 150 criblés lors d'une recombinaison homologue cherchant à déléter la région 5' de *Tsix* ($Tsix^{\Delta CpG}$). Et ce, en utilisant un protocole très similaire à notre protocole n°2 : un programme d'électroporation similaire (240V, 500 μ F, Biorad), une solution proportionnellement proche (3.10^7 cellules, 40 μ g d'ADN), et des bras d'homologie de tailles semblables (2 et 7,7kb).

Ces trois arguments sont à mon avis les raisons principales du non-aboutissement de ce projet. Ceci étant, il est fort possible que la complexité de la région promotrice de *Xist*, contenant de nombreuses régions répétées de grande taille (celles-ci étant pour certaines comprises dans les bras d'homologie, notamment le *Repeat-C*) contribue aussi à ce manque de réussite.

f. Perspectives

Malheureusement, faute de temps, je n'ai pu prolonger le criblage de clones outre mesure. Toutefois, Damien Neuillet, assistant-ingénieur au laboratoire Épigenétique du Chromosome X de l'UMR7216, a repris le projet à son compte. En se servant toujours du vecteur initial, qui a entre-temps, été intégralement validé par séquençage, il a effectué une électroporation des cellules *EL16.7* suivant le protocole 4 (Figure 60), et a criblé individuellement près de 600 clones par PCR. Sur l'ensemble des colonies criblées, 6 d'entre elles présentent un produit PCR de la taille attendue suite à une recombinaison homologue au bon locus. Il est actuellement dans le processus d'effectuer un criblage par *southern-blot* pour ces candidats.

Une fois un (ou plusieurs) clone positif identifié, l'étape suivante sera d'y réinsérer les versions mutées de la région promotrice de *Xist*. Ces mutations seront engendrées sous la forme de micro-délétions d'environ 5pb au cœur du site consensus de fixation des protéines CTCF et YY1 (Kim *et al.*, 2006 ; Kim *et al.*, 2007) grâce au procédé « *QuikChange Lightning Site-Directed Mutagenesis Kit* » produit par Agilent (Figure 61). Elles seront tout d'abord réalisées individuellement avant d'ensuite être étudiées collectivement. Les sites sur lesquels nous nous focaliserons sont :

- les sites YY1 en amont du promoteur P2 de *Xist*, en délétant d'abord les 3 sites d'un coup, suivi des 3 individuellement afin de voir si la fixation de cette protéine se déplace en fonction du contexte cellulaire et donc si cela influe sur son rôle au niveau de la régulation transcriptionnelle de *Xist*
- le site de fixation de CTCF au niveau du promoteur P2 de *Xist*, lui aussi potentiellement impliqué dans la surexpression de *Xist*

- et enfin le site de fixation de CTCF et Rad21 en amont du promoteur P1 de *Xist*, dont la délétion se révélera sûrement très utile à l'identification de sa fonction.

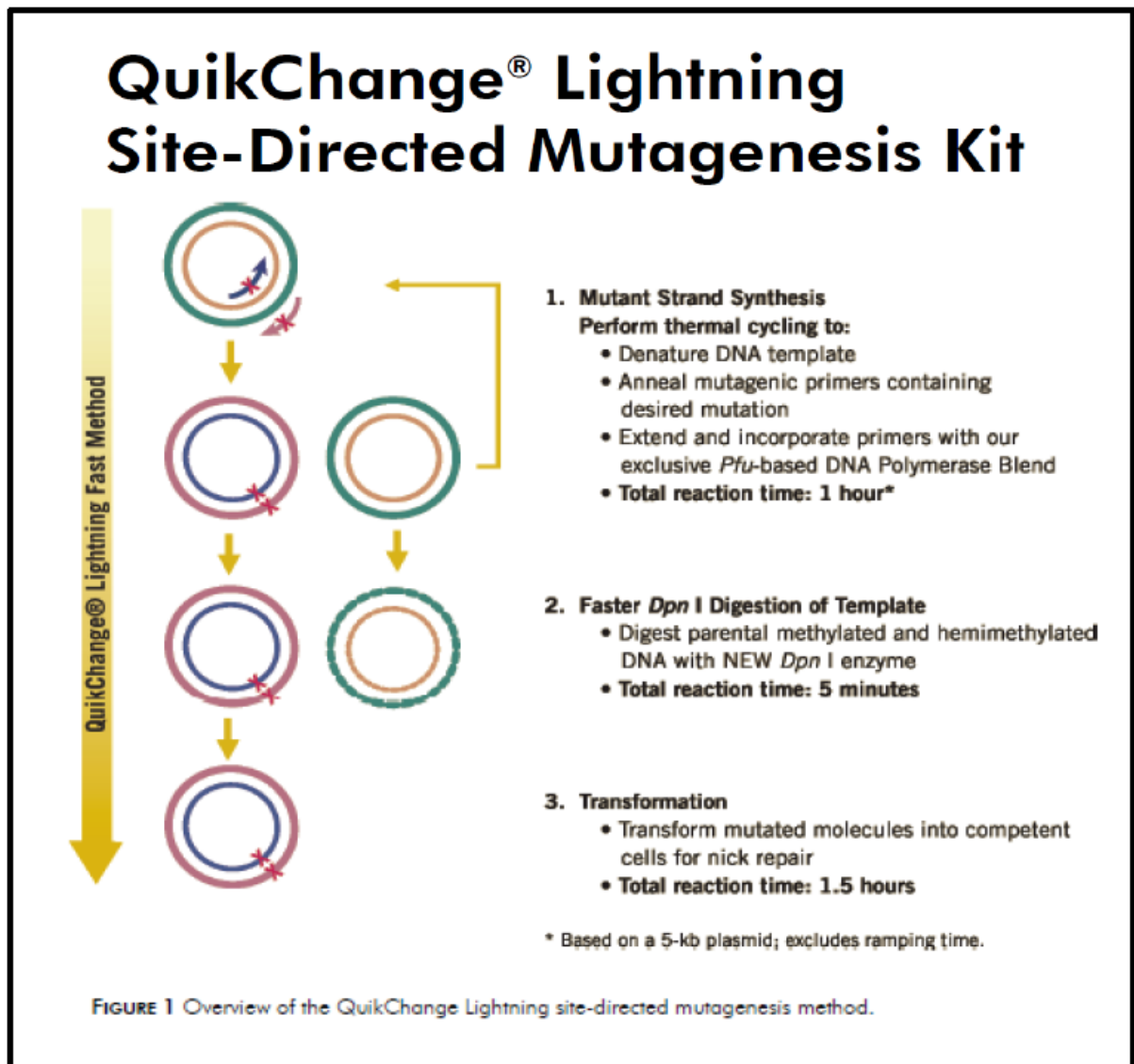


Figure 61 : Schéma du protocole utilisé pour générer des mutations spécifiques au sein des sites de fixation des protéines d'intérêt. (Agilent)

Non-seulement ces mutations joueront un rôle essentiel dans la caractérisation de la fonction de CTCF, YY1 et Rad21 à ce locus par rapport à l'expression de *Xist*, mais elles permettront aussi d'observer quelles sont les conditions requises à leur fixation vis-à-vis les unes des autres. C'est-à-dire que cette étude devrait nous permettre d'observer les recrutements ou les coopérations entre les protéines, particulièrement pour savoir si la fixation de CTCF est bien nécessaire au recrutement de Rad21 en amont de P1 et surtout si CTCF et YY1 agissent en hétérodimère à P2, Donohoe *et al.* n'ayant pas regardé la fixation de l'une lors de

l'interférence de l'autre afin de valider leur interaction au sein de *DxPas34* (Donohoe *et al.*, 2007). Une collaboration avec le laboratoire d'Edith Heard pourra aussi être envisagée dans le but d'étudier l'effet de ces mutations sur la structure tridimensionnelle du *Xic* mais aussi éventuellement sur le *pairing*.

En attendant d'obtenir un (ou plusieurs) clone(s) recombinant(s) correcte(s), nous avons entrepris d'étudier l'utilité des différents sites de fixation de CTCF et YY1 au promoteur de *Xist* dans la surexpression transcriptionnelle de ce dernier, sur l'expression de la luciférase. Cette étude, qui vient d'être initiée par Mélanie Makhoul (avec l'aide d'Olivier Leroy), se sert de la même méthode de mutagenèse de sites prévue pour le RMCE mais au sein d'une étude de l'expression de la luciférase, lorsque placée sous le contrôle du promoteur de *Xist*, ou plus exactement de la région correspondant à l'exon 4 de *Tsix*. Cette expérience devrait nous apporter une première indication quant à la nécessité de la présence de ces protéines dans la surexpression de *Xist*, d'autant plus que les séquences responsables de sa répression (l'antisens, l'intron 1 de *Xist*) ne font (*a priori*, si ce sont bien les seuls répresseurs) par partie de la région étudiée, ce qui devrait nous permettre d'étudier ce processus aussi bien dans les cellules ES que les MEFs (mâles ou femelles).

III. L'étude du réseau transcriptionnel régulant *Xist* au sein de cellules germinales embryonnaires de souris

a. Introduction

Comme nous venons de le voir, et comme nous l'avons vu lors de l'Introduction (voir le chapitre : « Modèles cellulaires », p19), il semblerait que les mécanismes régulant l'inactivation du chromosome X divergent entre les cellules germinales primordiales et les cellules germinales embryonnaires. Bien que dans les deux cas on assiste à une réactivation du chromosome X simultanément à un retour à l'état pluripotent de la cellule, d'une part la dynamique de la réactivation est différente, celle-ci se faisant progressivement au cours de la migration le long de la gouttière uro-génitale pour les PGCs (Sugimoto *et al.*, 2007), tandis qu'aucun chromosome X inactif n'est observé à partir du moment où l'on commence à dériver les PGCs en EGs (peu importe le stade de départ). D'autre part les profils transcriptionnels entre les deux types cellulaires sont différents, celui des PGCs étant bien particulier tandis que celui des EGs se rapproche beaucoup de celui des cellules ES ou des iPS (Mise *et al.*, 2008, Sharova *et al.*, 2007).

Au vu des similitudes entre ES, iPS et EGs, nous serions en droit d'imaginer qu'il existe une forte similitude entre les mécanismes de réactivation du chromosome X inactif au sein de ces trois types cellulaires. D'autant plus qu'elle peut s'effectuer de manière très rapide, peut-être même en une division cellulaire (tout du moins dans les ES), contrairement aux PGCs chez qui la réactivation se fait très progressivement sur plusieurs jours (Sugimoto *et al.*, 2008, Abe *et al.*, 2007, de Napoles *et al.*, 2007, Chuva de Sousa Lopes *et al.*, 2008). Cependant, on ne peut exclure la possibilité de l'existence de divergences entre ces trois types cellulaires, notamment dues à la nature de l'inactivation du chromosome X avant la réactivation. En effet, il existe plusieurs différences entre l'X inactivé de manière empreintée et l'X inactivé de manière aléatoire, la dissimilitude la plus flagrante étant l'absence de méthylation de l'ADN aux promoteurs des gènes (Kratzer *et al.*, 1983, Sado *et al.*, 2000), sur le X inactivé de manière empreintée. Il est donc possible que la réactivation du chromosome X inactif dans les cellules PGCs, EGs et iPS engage des mécanismes plus drastiques afin de pallier à une "complexité" ou à une robustesse supérieure de l'hétérochromatine de l'X inactivé aléatoirement par rapport au X inactivé de manière empreintée.

L'étude du réseau transcriptionnel régulant l'expression de *Xist* dans les EGs va donc d'abord nous permettre de découvrir la manière dont *Xist* est régulé au sein de ce type cellulaire

particulier, puis de comparer ce réseau à celui des cellules ES, nous permettant ainsi d'observer si les deux sortes de réactivation du chromosome X engendrent des différences quant au maintien de l'activité transcriptionnelle sur les deux X dans des cellules pluripotentes.

Dans la suite de ce chapitre, je tiens à préciser que tous les résultats sont préliminaires et nécessitent donc d'être reproduits plusieurs fois avant que l'on puisse définitivement conclure sur leur signification.

b. Résultats

i. Expression de *Xist*, *Tsix*, *Oct4* et *Nanog* dans les EGs

Grâce à la générosité du Pr. Stéphane Viville (Laboratoire d'ontogénèse des cellules germinales et primordiales de l'IGBMC de Strasbourg) nous avons pu acquérir plusieurs lignées de cellules EGs, aussi bien mâles que femelles, dérivées à partir de PGCs de stades différents (8,5 ou 9,5jpc). Dans un premier temps, nous n'avons entrepris la caractérisation du réseau transcriptionnel de *Xist* que dans deux lignées : Nero7 (cellules femelles, dérivées à 9,5jpc) et Nero8 (cellules mâles, dérivées à 9,5jpc) (Figure 62A).

Toutefois, nous avons d'abord voulu observer l'expression de plusieurs gènes d'intérêt afin de s'assurer de l'état pluripotent des cellules (Figure 62B, C et D). Comme moyen de comparaison nous avons rajouté les données obtenues pour des cellules ES femelles (lignée LF2) avant et en cours de différenciation.

Comme nous pouvons le voir sur la Figure 62B, les niveaux de transcrits *Oct4* sont très similaires à ceux de cellules ES non différenciées. Par contre, de manière surprenante, les niveaux de *Nanog* sont particulièrement bas dans les EGs mâles et femelles, correspondant davantage à un niveau habituellement propre à des cellules ES après 4 jours de différenciation cellulaire par acide rétinoïque. Ceci est surprenant, car malgré le fait que *Nanog* n'est pas nécessaire au cours des premières étapes de l'établissement de la pluripotence, il joue un rôle primordial dans son maintien et dans la formation des cellules germinales (Chambers *et al.*, 2007). Sa expression est d'ailleurs détectée à partir de 8jpc dans les PGCs (correspondant au début de migration le long de la gouttière uro-génitale) (Yamaguchi *et al.*, 2005) et est décrite comme active dans EGs (Sharova *et al.*, 2007 ; Chambers *et al.*, 2003).

Cependant, malgré une réactivation plutôt faible de *Nanog*, une répression importante de *Xist* a bien lieu suite à la dérivation des cellules EGs (Figure 62D), l'expression de *Xist* étant encore observée jusqu'à 13,5jpc dans les PGCs (Nestervova *et al.*, 2002). Cette répression semblerait se faire indépendamment de l'antisens, car bien que celui-ci soit ré-exprimé par rapport à son absence dans les PGCs, son taux d'expression se rapproche d'avantage de celui des cellules ES femelles en cours de différenciation (Figure 62D) que des non-différenciées (Figure 62C).

En comparant l'expression de ces quatre gènes entre EGs mâles et femelles, nous n'observons que très peu de différences. La divergence la plus flagrante semble être la quantité très faible de transcrits *Xist* chez les mâles par rapport à chez la femelle, cette différence inter-sexe étant aussi observée dans les cellules ES, et est sans doute due à l'expression basale de *Xist* provenant d'un ou de deux chromosomes X en fonction du sexe.

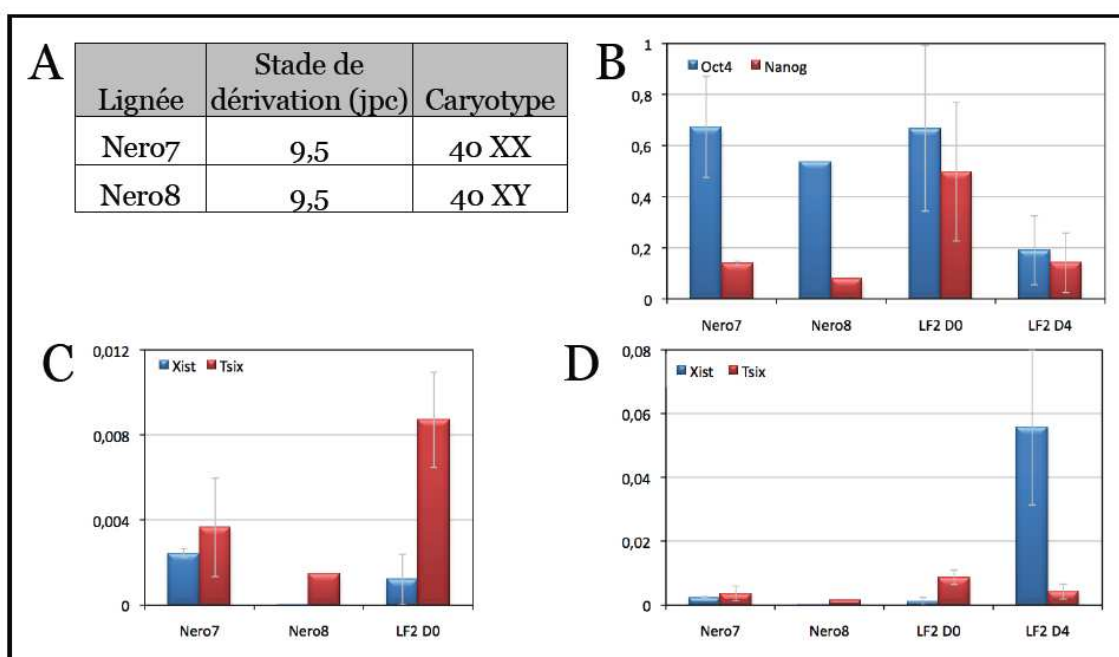


Figure 62 : Statut transcriptionnel de plusieurs gènes clés à notre sujet d'étude au sein de cellules EGs. (A) Description sommaire des deux lignées étudiées. (B) Analyse de l'expression d'*Oct4* et de *Nanog* par qRT-PCR suite à une normalisation par *Arpo* au sein des cellules EGs et comparaison avec les résultats obtenus dans des cellules ES femelles indifférenciées. (C) et (D) Même type d'analyse pour *Xist* et *Tsix*. (C) étant un agrandissement de (D) ne concernant que Nero7, Nero8 et LF2D0.

ii. Mécanisme de répression de *Xist* dans les EGs

Xist étant correctement réprimé au sein des cellules EGs, et ce, malgré une expression faible de *Nanog* et de *Tsix*, nous avons cherché à comprendre comment s'était effectuée cette inhibition. Pour cela nous avons entrepris de cartographier la fixation des facteurs de

pluripotence Oct4, Nanog et Sox2, connus comme étant responsables de la répression transcriptionnelle de *Xist* au sein des cellules ES (Navarro *et al.*, 2008).

Dans les cellules EGs femelles Oct4, Nanog et Sox2 se fixent bien à l'intron 1 de *Xist* de manière identique à ce qui est observé dans les ES (Figure 63B), nous indiquant donc qu'il existe une très forte probabilité que leur rôle soit le même à ce locus dans les deux types cellulaires. Cette fixation est d'autant plus surprenante tenant compte de la faible expression de *Nanog* (plus proche des cellules en cours de différenciation que pluripotentes), mais qui se fixe de manière parfaitement claire à l'intron 1 de *Xist*.

Par la même occasion, sachant que la fixation des facteurs de pluripotence est identique à l'intron 1 de *Xist* entre les EGs et les ES, nous en avons déduit qu'il devrait en être de même pour leur fixation en 5' de *Tsix*, et avons donc cherché à reproduire les résultats du laboratoire de Jeannie Lee concernant la localisation d'Oct4 et éventuellement Sox2 autour de *DxPas34* (Donohoe *et al.*, 2009) (Figure 63A). Paradoxalement, malgré le fait qu'il semblerait que nous ayons réussi à reproduire ces résultats (si l'on se base sur un ratio : valeur du pic dans l'intron 1 par rapport à la valeur du pic à *DxPas34*), il s'avère tout de même qu'il ne semble pas y avoir de fixation de ces facteurs autour du minisatellite, aussi bien dans les EGs que dans les ES (voir l'Article : Figure supplémentaire 1B,C, p209).

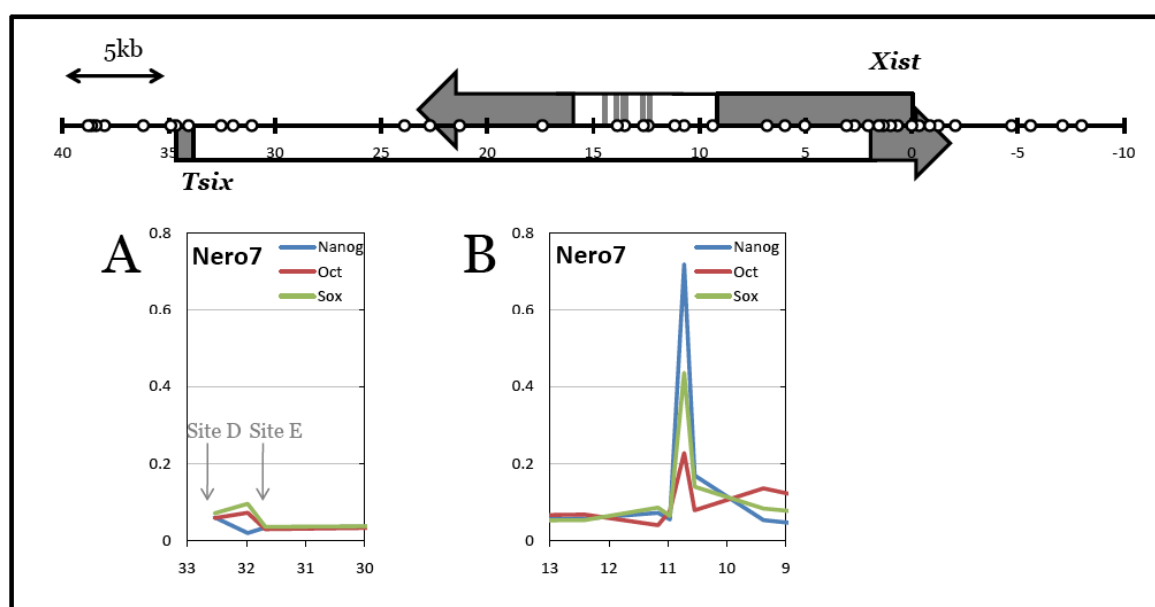


Figure 63 : Fixation des facteurs de pluripotence Oct4, Nanog et Sox2 au sein de l'intron 1 et du minisatellite DxPas34 dans les EGs femelles. (A) Étude de la fixation des facteurs de pluripotence au sein du minisatellite *DxPas34* dans les EGs femelles. Sont indiqués par des flèches les sites où Donohoe *et al.* (Donohoe *et al.*, 2009) prétendent observer une fixation d'Oct4 et éventuellement de Sox2. (B) Étude de ces mêmes facteurs au sein de l'intron 1 de *Xist*, site décrit par Navarro *et al.* (Navarro *et al.*, 2009).

iii. Fixation de CTCF et YY1 le long de *Xist/Tsix*

Au vu des similitudes, aussi bien transcriptionnelles que mécanistiques, entre EGs et ES, nous avons subséquemment cherché à établir un profil de fixation des protéines CTCF et YY1 le long du locus *Xist/Tsix* afin de voir si la cellule conserve aussi ses mécanismes préparant à la surexpression de *Xist* même si les cellules à partir desquelles elles sont dérivées (les PGCs) n'ont aucune utilité pour ce genre de mécanisme (développé dans la discussion).

Que ce soit dans les cellules EGs mâles ou femelles, nous avons observé un profil de fixation de CTCF et YY1 quasi-identique à celui des cellules ES le long de *Xist* et son antisens, que ce soit au niveau de l'emplacement des pics ou de leur intensité par rapport au bruit de fond (pour CTCF, les données pour YY1 n'ayant pas été normalisées faute d'un nombre suffisant de points de bruit de fond) (Figure 64). L'aspect mal défini du pic de CTCF en amont de *Xist* étant sans doute dû à un manque de reproduction de l'expérience, ainsi qu'à de légers problèmes de contamination d'ADN présents au laboratoire lors de la réalisation de cette expérience.

c. Discussion

Conformément aux données publiées par une autre équipe (Nesterova *et al.*, 2002), nous avons pu confirmer que *Xist* est bien réprimé transcriptionnellement au sein des cellules EGs aussi bien mâles que femelles. Et bien que l'état transcriptionnel de *Xist* ne reflète pas celui du chromosome X entier à partir duquel il est exprimé, l'observation que *Xist* se trouve sous sa forme réprimée nous rassure par rapport à la réactivation de l'Xi, sachant que ces cellules ne sont pas mutées et n'auraient donc aucune raison de réprimer *Xist* si elles n'avaient pas subi une réactivation du chromosome X inactif suite au retour à la pluripotence.

Une différence majeure entre EGs et PGCs est le fait qu'en examinant les chromosomes X des EGs, l'on n'assiste jamais à une réactivation progressive du X inactif corrélant avec le stade des PGCs à partir desquels ils ont été dérivés, tandis qu'en observant les PGCs on observe bien une réactivation graduelle, sûrement passive, au cours de leur migration le long de la gouttière uro-génitale. Cependant, une étude récente a démontré que cette réactivation progressive pouvait être observée dans les EGs mais au cours des premiers jours de leur dérivation (Chuva de Sousa Lopes *et al.*, 2008).

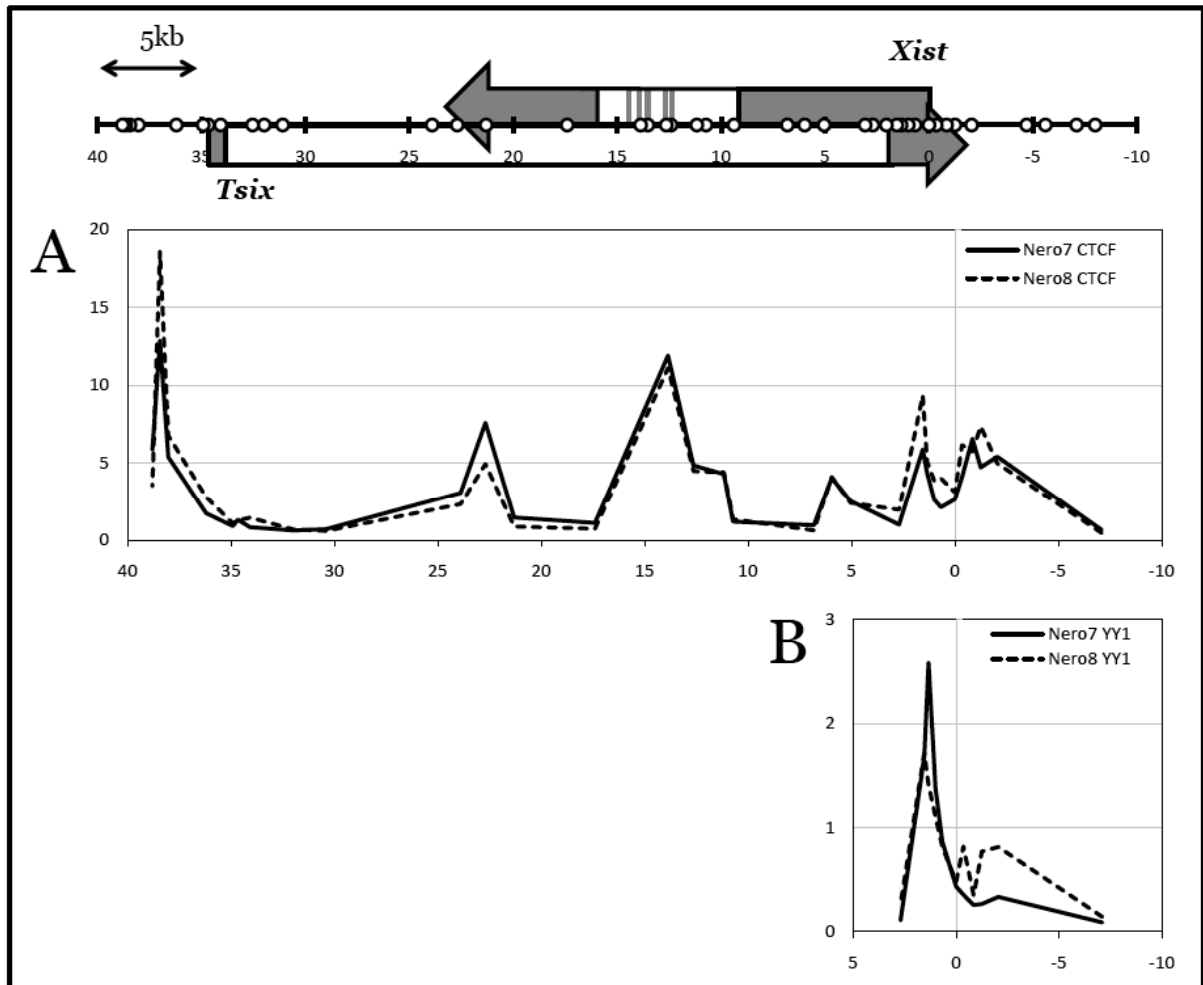


Figure 64 : Fixation de CTCF et YY1 le long du locus *Xist/Tsix* au sein de cellules EGs. (A) Analyse de la fixation de CTCF par ChIP suivie de PCR-quantitative le long du locus *Xist/Tsix* au sein de cellules EGs mâles (Nero8, pointillés) et d'EGs femelles (Nero7, trait plein), analysée de la même manière que décrit à la Figure 34. (B) *Idem*, mais concernant la protéine YY1 ainsi que ces données n'ont pas été normalisées par rapport au bruit de fond (faute d'un nombre suffisant de sites à fixation négatives) mais elles ont tout de même été normalisées par rapport à l'input.

Ceci laisse entrevoir la possibilité que les mécanismes de réactivation du chromosome X ne sont pas si différents entre PGCs et EGs, seulement pour le savoir, il faudra pouvoir étudier ce phénomène au cours de la dérivation et non pas une fois celle-ci terminée. Et cela pourra être particulièrement utile dans la détermination du rôle de *Tsix* dans la réactivation du X inactif, sachant que son expression n'est pas nécessaire à la répression de *Xist* comme en témoigne ce processus dans les PGCs. En effet, contrairement aux PGCs, les EGs expriment *Tsix*, potentiellement de manière plutôt faible, bien que visible par immunofluorescence (Nesterova *et al.*, 2002). Il sera donc intéressant de savoir à quel moment *Tsix* est exprimé lors de la dérivation des EGs et de comparer cette ré-expression à l'extinction transcriptionnelle de *Xist* afin de savoir si ces processus de réactivation diffèrent entre les deux types cellulaires.

Le fait que l'expression de *Tsix* soit aussi basse dans les EGs paraît surprenant. Si ces données sont confirmées, il sera particulièrement intéressant de regarder comment se comportent les facteurs de pluripotence impliqués dans sa régulation (Rex1, Klf4, c-Myc, Sp1, CTCF, YY1). Cependant, une étude menée par Sharova *et al.* démontre que leur expression paraît être très similaire dans les deux types cellulaires, cette similitude étant aussi observable pour *Xist* et *Tsix* (Sharova *et al.*, 2007) (Figure 65), remettant en cause nos résultats (concernant l'expression de *Tsix*).

	ES	EG	Ratio ES/EG
Klf4	4,34155	4,341183333	1,000084462
Rnf12	4,3433	4,3251	1,004207995
Tsix	3,171633333	2,997983333	1,05792227
Xist	2,190883333	2,41755	0,906241167
Myc	3,95235	3,867983333	1,021811538
Sp1	3,991433333	3,95845	1,008332386
Zfp42	4,337183333	4,159616667	1,042688228
Nanog	4,041416667	4,071666667	0,99257061
CTCF	4,23895	4,1885	1,012044885
YY1	3,2294	3,291133333	0,98124253

Figure 65 : Moyennes des variances de l'expression log-transformée de gènes d'intérêt à nos travaux dans différentes lignées de cellules EGs et ES. Ce tableau démontre la non-variabilité de l'expression des gènes majeurs impliqués dans la pluripotence et/ou dans la régulation du locus *Xist/Tsix* entre cellules germinales embryonnaires et cellules souches embryonnaires. (Basée sur les résultats de Sharova *et al.*, 2007)

Ces données pourraient aussi nous faire pencher vers une théorie de l'artéfact concernant la faible expression de *Nanog* observée au cours de nos expériences. De plus, le fait que *Nanog* se lie à l'intron 1 de *Xist* semblerait signifier que même si son expression est faible, cela n'influe pas sur sa capacité à jouer son rôle de répresseur transcriptionnel de *Xist*, bien que la majeure partie de cette fonction semble être générée par Oct4 (Navarro *et al.*, 2008).

Le fait que le profil de fixation de CTCF et YY1 soit identique entre cellules ES et EG engendre aussi de nombreuses questions. Nous pouvons en effet nous demander si la présence de ces protéines le long de *Xist/Tsix* correspond vraiment à une réalité biologique reflétant la situation au sein des PGCs. D'après les divergences entre PGCs et EGs il semblerait plus vraisemblable que ce ne soit pas le cas. Cependant, nos données n'excluent en aucune manière la présence d'un profil de fixation similaire dans les PGCs. Toutefois, si notre hypothèse, que CTCF et YY1 sont présentes au niveau du promoteur P2 de *Xist* dans le but d'être prêtes à le surexprimer dès que les facteurs répressifs sont partis, est correcte, il

semblerait curieux qu'elles aient le même rôle dans les PGCs. En effet, sachant que dans le cas de cellules germinales mâles le X est inactivé en cours de méiose par le phénomène de MSCI (voir le chapitre : « Inactivation méiotique des chromosomes sexuels », p8) indépendamment de *Xist* (Turner, 2002, McCarrey *et al.*, 2002) et que dans le cas de cellules germinales femelles leur X n'est pas réactivé avant l'inactivation aléatoire dans le futur embryon, il semblerait curieux que *Xist* ait besoin d'être prédisposé à la surexpression. Par contre, même si *Xist* n'est pas nécessaire à l'inactivation du chromosome X dans les cellules germinales mâles post-méiotiques (MSCI ou PMSC) celui-ci est tout de même exprimé (McCarrey & Dilworth, 1992 ; Salido & Shapiro *et al.*, 1992). Il serait donc très intéressant de voir si CTCF et YY1 sont fixées au promoteur P2 de *Xist* dans de telles cellules et de comparer ces résultats avec ce qui est observé au cours de l'oogénèse, les cellules germinales femelles n'exprimant jamais *Xist* et ne devant donc pas montrer un tel profil de fixation. Avec les avancées considérables constatées dans la perfection de la méthode de ChIP, celle-ci nécessitant de moins en moins de chromatine pour fournir des résultats fiables, il devrait être possible, dans un avenir proche, d'étudier ce genre de phénomène dans des tissus ne fournissant qu'une quantité très limitée de cellules (telles que les cellules germinales).

Il pourrait aussi être très intéressant de regarder l'état de la méthylation du promoteur de *Xist* dans les EGs car Hajkova *et al.* montrent que celui-ci est totalement déméthylé dans les cellules PGC mâles à partir de 12,5jpc lors de la déméthylation globale du génome et l'effacement de l'empreinte parentale (Reik *et al.*, 2001 ; Surani, 2007), mais la répression de *Xist* est tout de même maintenue (Hajkova *et al.*, 2002). Et au vu des différences dans la méthylation des DMDs entre ES, EGs et PGCs, l'étude des profils de méthylation dans ces cellules pourraient fournir des indications sur les mécanismes menant à, ou maintenant, cette répression. Parallèlement à cela, Navarro et Avner supposent que les sites de fixation d'Oct4 et Sox2 à l'intron 1 de *Xist* sont différenciellement méthylés au cours de la gamétogenèse, empêchant ainsi leur fixation dans les cellules germinales mâles mais pas femelles expliquant ainsi l'expression sexe-spécifique de *Xist* à ce stade du développement (Nanog étant, dans ce modèle, responsable de l'effacement de l'empreinte empêchant la fixation d'Oct4 et Sox2 sur l'allèle maternel au stade morule).

En définitive, nous pourrions supposer que les similitudes observées entre EGs et ES, ainsi que les divergences entre EGs et PGCs, sont somme toute logiques, si l'on considère le fait que lors de la dérivation des cellules EG, on oriente les cellules vers un devenir de pluripotence et d'auto-renouvellement perpétuel pouvant ensuite donner n'importe quel type

cellulaire, à la différence des PGCs qui, suite à leur migration, rentrent en méiose pour en fin de compte ne pouvoir produire uniquement des gamètes. Il semblerait donc, que l'environnement cellulaire des PGCs, nécessaire à leur développement en cellules germinales, produit des signaux différents de ceux fournis par les cellules nourricières (ou « *feeders* ») ou présents dans le milieu de culture des EGs lors de leur dérivation. Et c'est cette différence de contexte de développement cellulaire qui pourrait pousser les cellules EGs à se rapprocher transcriptionnellement et épigénétiquement (à part pour la méthylation des gènes soumis à l'empreinte parentale) des cellules souches embryonnaires, dérivées des cellules de la masse interne dans des conditions similaires. Par la suite, il sera donc particulièrement intéressant de dériver des EGs à l'aide d'acide rétinoïque, afin de voir si ces correspondances par rapport à la régulation de *Xist* entre ES et EGs sont maintenues lors de l'initiation de l'inactivation. De plus, l'étude de la méthylation de l'îlot CpG au promoteur de *Xist* dans ces différents contextes cellulaires pourra peut-être nous éclairer par rapport à ce lien de conservation, semble-t-il unique, existant entre EGs et PGCs.

IV. Le couplage moléculaire de la régulation de *Tsix* et de la pluripotence cellulaire

a. Introduction

Depuis la découverte en 2005 par l'équipe de Claire Rougeulle du fait que *Tsix* ne régule pas transcriptionnellement *Xist* (Navarro *et al.*, 2005) (ultérieurement confirmé par Sun *et al.*, 2006), une des questions majeures concernant l'initiation de l'inactivation du chromosome X était : « comment *Xist* est-il transcriptionnellement régulé en cours de différenciation ? ». Il aura fallu attendre 2008, pour que cette même équipe découvre le mécanisme réprimant transcriptionnellement *Xist* dans les cellules ES non différenciées tout en établissant un lien direct avec la pluripotence cellulaire (Navarro *et al.*, 2008). En effet, les facteurs de pluripotence Oct4, Nanog et Sox2 se fixent au sein de l'intron de 1 de *Xist* dans ce type de cellules (voir le chapitre : « Couplage de la régulation de l'expression de *Xist* et la pluripotence », p88), cette découverte majeure ayant permis d'expliquer le lien étroit observé depuis bon nombre d'années entre l'initiation de l'inactivation et l'état du développement de la cellule.

Cependant, il semblerait que *Tsix* ait aussi un lien particulier avec la différenciation cellulaire, son expression diminuant fortement lors de l'induction de celle-ci. De plus, théoriquement, la répression de l'expression de *Tsix* précède la surexpression de *Xist* sur le X inactif lors de l'initiation de l'inactivation afin d'établir un biais entre les deux allèles, permettant ainsi l'exécution du choix du chromosome X à inactiver. Et sachant que *Xist* est transcriptionnellement régulé par des facteurs de pluripotence, il semblerait logique que l'antisens le soit aussi afin d'obtenir la régulation fine, précise et synchrone, nécessaire à l'équilibre transcriptionnel entre les deux gènes.

Mais jusqu'à présent, la régulation de l'expression de *Tsix* a été décrite comme étant principalement dirigée par deux régions : le minisatellite *DxPas34* et la région de transcription intergénique *Xite* (Stavropoulos *et al.*, 2005). Au cours des années, le laboratoire de Jeannie Lee a publié des données concernant des protéines recrutées à ces deux régions et la manière dont ils influent sur la transcription de l'antisens. C'est ainsi que CTCF, YY1, Oct4 et Sox2 ont chacune été impliquée dans la régulation (surtout la surexpression) transcriptionnelle de *Tsix* (Donohoe *et al.*, 2007, Chao *et al.*, 2002, Donohoe *et al.*, 2009) (voir les chapitres : « *Tsix* », p56 ; « *DxPas34* », p59 ; et « *Xite* », p63). Cependant, il faut signaler

que les expériences de déplétion de ces protéines (notamment CTCF et Sox2) ne semblent pas corrélées avec l'ensemble de leurs études, la fixation même de CTCF et YY1 autour de *DxPas34* ne semblant pas non plus clairement refléter l'activité transcriptionnelle de *Tsix* lors de l'initiation de l'inactivation (Donohoe *et al.*, 2007).

C'est donc dans ce contexte de recherche de liens entre l'initiation de l'inactivation et la pluripotence, que Pablo Navarro, avec mon assistance, a entrepris la recherche de facteurs de pluripotence dans la régulation de l'antisens, certaines études ayant déjà démontré des corrélations transcriptionnelles entre des marqueurs de la pluripotence (notamment Rex1) et *Tsix* (Masui *et al.*, 2008, Silva *et al.*, 2008).

À l'origine de mon implication dans ce projet se trouvait mon désir d'associer mes découvertes de CTCF et YY1 comme activateurs potentiels de *Xist* (ou de *Tsix*, d'après Donohoe *et al.*, 2007) et les résultats de Pablo Navarro, concernant la répression de ce gène par les facteurs de pluripotence Oct4, Nanog et Sox2. Il s'est avéré que le laboratoire de Stuart Orkin avait récemment publié un article sur les réseaux d'interactions protéiques responsables de la pluripotence au sein des cellules souches embryonnaires (Wang *et al.*, 2006). Au sein de cette étude Wang *et al.* ont observé des interactions protéine-protéine entre YY1 et Rex1 ainsi qu'entre Rex1 et Oct4, nous indiquant que Rex1 pouvait être un lien entre ces deux groupes de facteurs transcriptionnels. Comme démontré au sein de la Figure 1 de l'article (voir Figure 1d et e, p205), j'ai pu mettre en évidence une fixation de Rex1 au microsatellite *DxPas34* dans les cellules ES mâles et femelles, engendrant par la suite l'étude de son implication, ainsi que celle d'autres facteurs de pluripotence (Klf4 et c-Myc) dans la régulation de l'antisens *Tsix*.

Dans un second temps, j'ai aussi pu vérifier une hypothèse quant au moyen de mobilisation de Klf4 à *DxPas34* en démontrant le recrutement (avec un profil de fixation quasiment identique) de la protéine Sp1 à ce même locus (voir Figure Supplémentaire 5c, p213). Sp1 servirait donc d'intermédiaire entre soit YY1 et Klf4 et/ou entre c-Myc et Klf4 étant donné ses interactions prouvées avec ces 3 autres protéines (Lee *et al.*, 1993 ; Gartel *et al.*, 2001 ; Zhang *et al.*, 1998).

Article
(Accepté dans Nature)

Molecular coupling of *Tsix* regulation and pluripotency.

Authors:

Pablo Navarro^{1, 2, *}, Andrew Oldfield^{1, 3}, Julie Legoupi¹, Nicola Festucia², Agnès Dubois¹,
Mikael Attia¹, Claire Rougeulle^{1, 3}, Ian Chambers², and Philip Avner^{1*}.

Authors affiliation:

1- Unité de Génétique Moléculaire Murine, URA 2578, Institut Pasteur 75724, Paris Cedex 15, France.

2- Medical Research Council (MRC). Centre Development in Stem Cell Biology, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, MRC. EH9 3JQ, Edinburgh, UK.

3- UMR 7216 Epigénétique et Destin Cellulaire, CNRS/Université Paris Diderot, Case 7042, 75205 Paris Cedex 13, France.

*- Co-corresponding authors. To whom correspondence should be addressed.

E-mail: philip.avner@pasteur.fr

E-mail: pablo.navarro@ed.ac.uk

Main

The reprogramming of X-chromosome inactivation during the acquisition of pluripotency *in-vivo* and *in-vitro*¹, is accompanied by the repression of *Xist*², the trigger of X-inactivation³, and the upregulation of its antisense counterpart *Tsix*⁴. We have shown that key factors supporting pluripotency —Nanog, Oct4, and Sox2— bind within *Xist* intron 1 in undifferentiated embryonic stem (ES) cells to repress *Xist* transcription⁵.

However, the relationship between transcription factors of the pluripotency network and *Tsix* regulation has remained unclear^{5, 6}. Here, we show that *Tsix* upregulation in ES cells depends on the recruitment of the pluripotent marker Rex1, and of the reprogramming-associated factors Klf4 and c-Myc, by the *DXPas34* minisatellite associated with the *Tsix* promoter. Upon deletion of *DXPas34*, binding of the three factors is abrogated and the transcriptional machinery no longer efficiently recruited to the *Tsix* promoter. Additional analyses including knock-down experiments further demonstrate that Rex1 is critically important for efficient transcription elongation of *Tsix*. Hence, distinct ES cell-specific complexes couple X-inactivation reprogramming and pluripotency, with Nanog, Oct4 and Sox2 repressing *Xist* to facilitate the reactivation of the inactive X, and Klf4, c-Myc and Rex1 activating *Tsix* to remodel *Xist* chromatin⁷⁻¹⁰ and ensure random X-inactivation upon differentiation¹. The holistic scenario of *Xist/Tsix* regulation by pluripotent factors that we have identified suggests a general direct governance of complex epigenetic processes by the machinery dedicated to pluripotency.

X-inactivation reprogramming in female mice is a paradigm for the epigenetic processes underlying the acquisition of pluripotency¹. The inactivation of the paternal X-chromosome that characterises the earliest cleavage-stages of development is followed by X-chromosome reactivation in the pluripotent inner cell mass (ICM) of the blastocyst^{11, 12}.

Following differentiation, X-inactivation is established randomly on either the paternal or the maternal X-chromosome¹. The developmental plasticity of X-inactivation during early embryogenesis is paralleled by female induced Pluripotent Stem (iPS) cells¹³: the inactive X is reactivated in iPS cells and random X-inactivation *de novo* established upon loss of pluripotency¹⁴. Hence, two distinct processes affect X-inactivation during the generation of pluripotency: the reactivation of the inactive X *per se*, likely initiated by the repression of *Xist* by Nanog, Oct4, and Sox2^{5, 15}, and the acquisition by both X-chromosomes of an equal competence for future random X-inactivation, a process directly controlled by *Tsix*. Indeed, invalidation of *Tsix* in ES cells⁸ and embryos¹⁰ leads to drastic, stable remodelling of *Xist* chromatin in *cis*⁷⁻¹⁰, associated with the systematic upregulation of *Xist* from the *Tsix*-null allele upon differentiation. Thus, maximal *Tsix* activity is required in pluripotent cells to erase inherited *Xist* chromatin modifications and provide each X-chromosome with equal probabilities of *Xist* upregulation during differentiation¹⁶.

It has been proposed that the high levels of *Tsix* transcription characterising ES cells depend on binding of Oct4 and Sox2⁶, but not Nanog, at *DXPas34*¹⁷ and at *Xite*¹⁸, two enhancers of *Tsix*¹⁹. Oct4 and Sox2, by simultaneously controlling *Xist* and *Tsix* could be acting at the top of the X-inactivation regulatory hierarchy in pluripotent cells. Uncertainties however surround this hypothesis since (i) the reported binding of Oct4 and Sox2 to *DXPas34*, the strong ES-specific enhancer of *Tsix*, is not reproducible (Supplementary Fig. 1), (ii) binding levels at *Xite* are very low (Supplementary Fig. 1), and (iii) *Xite* is a weak enhancer of *Tsix*¹⁸. It appears likely that Oct4 and Sox2 play a minor role, if any, in the establishment of *Tsix* transcription in undifferentiated ES cells. In agreement, *Tsix* remains unaffected after 24h of *Oct4* knock-down, whilst *Xist* upregulation is already established (Supplementary Fig. 1). Thus, additional factors are implicated in controlling *Tsix* transcription in undifferentiated ES cells. Three pluripotency factors, Klf4, c-Myc and Rex1, attracted our attention.

Since Oct4 and Sox2 directly mediate repression of *Xist* in ES cells, it appeared of interest to extend our analysis to Klf4 and c-Myc, the two other factors that are commonly used to generate iPS cells¹³. Rex1 is a marker of pluripotency whose deletion is associated with decreased *Tsix* expression as revealed by microarray analysis²⁰. During iPS cells generation, *Tsix* re-expression temporally correlates with the induction of *Rex1*²¹.

We initially determined whether binding activity of Rex1, Klf4 and c-Myc could be detected at the *Xist/Tsix* region (Fig. 1a). Binding to the *Tsix* 5' region was found for all three factors, in both female (Fig. 1 d, g, j) and male (Fig. 1 e, h, k) ES cells. Only Rex1 displays binding at both ends of *DXPas34*, suggesting that Rex1 is recruited within *DXPas34* itself whilst Klf4 and c-Myc are bound between *DXPas34* and the *Tsix* promoter. As expected given its ChIP profile, Rex1 binding is lost upon the targeted deletion of *DXPas34* in male ES cells (Δ Pas34 cell line¹⁷, Fig. 1c, f). Strikingly, Klf4 and c-Myc binding is similarly affected in Δ Pas34 (Fig. 1i, l), suggesting that *DXPas34* influences Klf4 and c-Myc recruitment. This drastic perturbation in transcription factor binding correlates with a 90% reduction in *Tsix* RNA levels¹⁷, mediated by a strong reduction of RNAPII recruitment at the *Tsix* promoter (Supplementary Fig. 2). Thus, *DXPas34* orchestrates the recruitment of Rex1, Klf4 and c-Myc, to activate *Tsix* transcription in ES cells.

We next silenced *Oct4* in ZhbTc4.1 ES cells to induce the loss of pluripotency²² and the consequent downregulation of Rex1 and Klf4 (Fig. 2a). *Tsix* was downregulated (Fig. 2a) by transcriptional mechanisms (Fig. 2b, c), and binding of Rex1, Klf4 and c-Myc was reduced (Fig. 2d, e, f). Similar observations were made in terminally differentiated, *Tsix*-silenced⁹, mouse embryonic fibroblasts (Fig. 2g, h, i). Rex1, Klf4 and c-Myc are therefore critically important to couple *Tsix* regulation to pluripotency. We also analysed Trophectoderm Stem (TS) cells, which display reduced *Tsix* transcription⁹, and found that only Rex1 is absent from the *Tsix* 5' region (Fig. 2g, h, i). Although the existence of differentiation-dependent repressors or additional ES-specific activators cannot be excluded, this suggests that Rex1 plays an

important role in *Tsix* transcription in ES cells. In agreement with this, transient *Rex1* interference demonstrates that *Tsix* is a rapid *Rex1*-responsive gene in ES cells (Supplementary Fig. 3).

Next, we generated stable *Rex1* knock-down lines and confirmed that *Rex1* downregulation is accompanied by *Tsix* downregulation (Supplementary Fig. 3). The stably-interfered clone displaying the highest level of *Rex1* knock-down, in which *Tsix* expression and *Rex1* binding are reduced by half (Fig. 3a, b), was analysed in detail. We observed that neither *Klf4* nor *c-Myc* are down-regulated (Fig. 3a), nor affected in their binding to the *Tsix* 5' region (Fig. 3b). No drastic effect on RNAPII recruitment at the *Tsix* promoter (Fig. 3c), or on TFIIB, H3K4Me3, and H3K9Ac levels (Supplementary Fig. 3), was observed. Thus, *Rex1* is not an essential factor for recruiting the transcriptional machinery at the *Tsix* promoter, or for triggering the accumulation of chromatin marks characteristic of transcription initiation²³. Analysis of RNAPII (Fig. 3c) and H3K36Me3 (Fig. 3d, Supplementary Fig. 3) across the *Tsix* transcription unit did however reveal reduced levels in *Rex1* knock-down cells, in particular at the *Tsix* 3' end. Since the amount of 3' end-associated RNAPII at the *Tsix* locus was previously proven to be a good readout of *Tsix* transcription⁷, and H3K36Me3 levels correlate well with transcriptional activity²⁴, we conclude that *Rex1* is required for efficient elongation of *Tsix*. Statistical analysis of the ChIP profiles provides further support to the reduction of RNAPII and H3K36Me3 levels (Supplementary Fig. 3).

c-Myc has been shown to be a global regulator of transcription elongation in ES cells²⁵, and affects in particular *Tsix* (Supplementary Fig. 4). This suggests that *Rex1* and *c-Myc* functionally interact to bring about efficient *Tsix* elongation. The fact that *c-Myc* remains associated to the *Tsix* 5' region in *Rex1* knock-down cells indicates that *Rex1* acts downstream of *c-Myc*. The differential segregation of *Rex1*/*c-Myc* from *Nanog*/*Oct4*/*Sox2* targets²⁶ further indicates that *Rex1* might act as a global regulator of transcription elongation in ES cells by providing developmental specificity to *c-Myc* function.

How these factors biochemically interact at the *Tsix* 5' region remains unknown. Given that (i) Yy1 (a factor evolutionary related to Rex1²⁷) and Ctfp show positive binding at both extremities of *DXPas34*, (ii) Sp1 binds at the *Tsix* 5' region, and (iii) interactions between these factors have been previously reported (Supplementary Fig. 5), we propose that *DXPas34* acts as a DNA platform directly recruiting Rex1, Ctfp, and Yy1, which in turn facilitate the recruitment of Sp1, Klf4 and c-Myc. If Rex1 and c-Myc regulate *Tsix* elongation, the parallel recruitment of the RNAPII and Klf4 at the *Tsix* 5' region observed in our experimental conditions leads us to speculate that Klf4 might be critically important to load the transcriptional machinery at the *Tsix* promoter.

Persistent binding of Nanog, Oct4 and Sox2 to *Xist* intron 1 in Δ Pas34 ES cells (Supplementary Fig. 6) maintains transcriptional *Xist* silencing in the absence of *Tsix* transcription (Supplementary Fig. 2). Conversely, *Nanog* and *Oct4* inducible mutant ES cells retain normal regulation of *Tsix*⁵. Moreover, *Xist* upregulation is observed from both wildtype and *Tsix*-deleted alleles upon *Oct4* knock-down (Supplementary Fig. 1). Based on these results we propose that two distinct pluripotency-related complexes act independently but in parallel to specify the reactivation of the inactive X and the resetting of the epigenetic conditions required for *de novo* random X-inactivation, through their direct regulation of *Xist* and *Tsix*, respectively (Fig. 3e). Whilst Oct4, Sox2 and Klf4 show only low level binding at *Xite* (Fig. 3e and Supplementary Fig. 5), inspection of available ChIP-Seq data sets²⁸ indicates that other genes located within the X-inactivation centre, including *Rnf12*²⁹, are targets of the pluripotency-associated machinery. Further complexities of the molecular system coupling pluripotency and X-inactivation are therefore to be expected. Interestingly, the connection of pluripotency regulators with *Xist* and *Tsix* may also apply to other epigenetic phenomena. Examination of ChIP-Seq data sets shows abundant binding of pluripotent factors at known imprinting centres including that of the *Dlk1-Dio3* cluster (Supplementary Fig. 7), which is inappropriately regulated during iPS cell generation³⁰. Hence, over and above their importance

for pluripotency and self-renewal, the pluripotent factors may be key components of other, more specific epigenetic processes occurring in pluripotent cells, notably in the germ line. Cohorts of pluripotency regulators could be involved, as we previously hypothesized for *Xist*^{1, 5}, in either the erasure and/or the establishment of epigenetic imprints, both at *Tsix* and at other imprinted loci.

Methods Summary.

Cell lines: female ES cells: LF2; male ES cells: CK35; Δ Pas34 ES cells: #BH9 and #BD7; male TS cells: F2; male MEFs: derived from E13,5 embryos.

ChIP and RT-PCR were performed as previously described⁵. Chromatin and RNA preparations were isolated in parallel from the same culture batches.

Transient transfections of *Oct4* siRNA (Dharmacon) and the *Rex1* shRNA expressing vectors (a kind gift of Jon Schoorlemmer) were performed using a nucleofector (Lonza) and the manufacturer's protocol (program A30). Results were normalised to the RNA levels of cells nucleofected with non-targeting siRNA (Dharmacon), or with a vector expressing a shRNA against *geminin*. Sequences targeting *Rex1* and *geminin* are available upon request from Jon Schoorlemmer.

Stable integration of shRNA vectors was performed after electroporation of LF2 ES cells, and selection in hygromycin B for two weeks. Resistant clones were individually expanded. Clones carrying the shRNA vectors against *Rex1* or *geminin* were generated and analysed in parallel.

For primer sequences and antibody information see Supplementary Figure 8.

In the figure legends, 'n' indicates the number of independent cell cultures analysed to control for biological variation. When indicated, 'm' shows the number of experiments performed with the 'n' independent extracts. All the results are expressed as means $\pm$ sem.

References.

- ¹ Navarro, P. and Avner, P., When X-inactivation meets pluripotency: an intimate rendezvous. *FEBS Lett.* **583**, 1721-1727 (2009).
- ² Borsani, G. et al., Characterization of a murine gene expressed from the inactive X chromosome. *Nature* **351**, 325-329 (1991).
- ³ Penny, G. D. et al., Requirement for Xist in X chromosome inactivation. *Nature* **379**, 131-137 (1996).
- ⁴ Lee, J. T., Davidow, L. S., and Warshawsky, D., Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat. Genet.* **21**, 400-404 (1999).
- ⁵ Navarro, P. et al., Molecular coupling of Xist regulation and pluripotency. *Science* **321**, 1693-1695 (2008).
- ⁶ Donohoe, M. E. et al., The pluripotency factor Oct4 interacts with Ctf and also controls X chromosome pairing and counting. *Nature* **460**, 128-132 (2009).
- ⁷ Navarro, P. et al., A role for non-coding Tsix transcription in partitioning chromatin domains within the mouse X-inactivation centre. *Epigenetics Chromatin* **2**, 8 (2009).
- ⁸ Navarro, P., Page, D. R., Avner, P., and Rougeulle, C., Tsix-mediated epigenetic switch of a CTCF-flanked region of the Xist promoter determines the Xist transcription program. *Genes Dev.* **20**, 2787-2792 (2006)
- ⁹ Navarro, P. et al., Tsix transcription across the Xist gene alters chromatin conformation without affecting Xist transcription: implications for X-chromosome inactivation. *Genes Dev.* **19**, 1474-1484 (2005).
- ¹⁰ Sado, T., Hoki, Y., and Sasaki, H., Tsix silences Xist through modification of chromatin structure. *Dev. Cell* **9**, 159-165 (2005).
- ¹¹ Mak, W. et al., Reactivation of the paternal X chromosome in early mouse embryos. *Science* **303**, 666-669 (2004).

- ¹² Okamoto I., Otte AP., Allis CD., Reinberg D., and Heard E., Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science* **303**, 664-666 (2004).
- ¹³ Takahashi, K. and Yamanaka, S., Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676 (2006).
- ¹⁴ Maherali, N. et al., Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* **1**, 55-70 (2007).
- ¹⁵ Chambers, I. and Tomlinson, S. R., The transcriptional foundation of pluripotency. *Development* **136**, 2311-2322 (2009).
- ¹⁶ Navarro P., and Avner P., An embryonic story: analysis of the gene regulative network controlling Xist expression in mouse embryonic stem cells. *Bioessays* **32**, 581-588 (2010).
- ¹⁷ Vigneau, S. et al., An essential role for the DXPas34 tandem repeat and Tsix transcription in the counting process of X chromosome inactivation. *Proc. Natl. Acad. Sci. U S A* **103**, 7390-7395 (2006).
- ¹⁸ Ogawa, Y. and Lee, J. T., Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. *Mol. Cell* **11**, 731-743 (2003).
- ¹⁹ Stavropoulos, N., Rowntree, R. K., and Lee, J. T., Identification of developmentally specific enhancers for Tsix in the regulation of X chromosome inactivation. *Mol. Cell. Biol.* **25**, 2757-2769 (2005).
- ²⁰ Masui, S. et al., Rex1/Zfp42 is dispensable for pluripotency in mouse ES cells. *BMC Dev. Biol.* **8**, 45 (2008).
- ²¹ Silva, J. et al., Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol.* **6**, e253 (2008).
- ²² Niwa, H., Miyazaki, J., and Smith, A. G., Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* **24**, 372-376(2000).

- ²³ Kouzarides, T., Chromatin modifications and their function. *Cell* **128**, 693-705 (2007).
- ²⁴ Guttman, M. et al., Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223-227 (2009).
- ²⁵ Rahl, P. B. et al., c-Myc regulates transcriptional pause release. *Cell* **141**, 432-445 (2010).
- ²⁶ Kim, J. et al., An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **132**, 1049-1061 (2008).
- ²⁷ Kim JD., Faulk C., and Kim J., Retroposition and evolution of the DNA-binding motifs of YY1, YY2 and REX1. *Nucleic Acids Res.* **35**, 3442-3452 (2007).
- ²⁸ Chen, X. et al., Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**, 1106-1117 (2008).
- ²⁹ Jonkers, I. et al., RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. *Cell* **139**, 999-1011 (2009).
- ³⁰ Stadtfeld, M. et al., Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* **465**, 175-181 (2010).

End Notes.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank Philippe Clerc for critical discussions. We are very grateful to Jon Schoorlemmer for the kind gift of Rex1 interfering vectors and for communicating information prior to publication. P.N was initially supported by recurrent funding from the Institut Pasteur and then by the UK Royal Society (Newton International Fellowship), P.A was supported by recurrent funding from the CNRS and the Institut Pasteur, contracts 05-JCJC-0166-01 and 07-BLAN-0047-01 from the ANR, and funding from the EU Epigenome Network of Excellence, C.R was supported by the INSERM “Avenir” program and by the European Research Council « starting grant » program. Research in I.C laboratory was supported by The Wellcome Trust and by an MRC studentship (N.F).

Author Contributions P.N conceived the study, designed, carried out and analysed the experiments, and co-wrote the manuscript. A.O, J.L, N.F, M.A, and A.D provided technical help. C.R, I.C, and P.A provided financial and conceptual support. P.A conceived of the study and co-wrote the manuscript. All authors have read and approved the final manuscript.

Authors information Reprints and permissions information is available at www.nature.com/nature. The authors declare that they have no competing financial interests. Correspondence and requests for materials should be addressed to P.A (pavner@pasteur.fr) or P.N (pablo.navarro@ed.ac.uk).

Figure Legends.

Figure 1: *DXPas34* orchestrates Rex1, Klf4 and c-Myc recruitment to the *Tsix* 5' region in pluripotent ES cells. **a, b** and **c**, Schematic representation of the sub-regions of the *Xist/Tsix* locus analysed by Chromatin immunoprecipitation (ChIP) in female ES cells (**a**), male ES cells (**b**), and male Δ Pas34 ES cells (**c**). *Xist* exons are in green, the arrows indicate the direction of transcription of *Xist* (top) and *Tsix* (bottom). *DXPas34* is shown in blue. The location of each primer pair is indicated by a black circle. In (**b**), the blue circles show the primer pairs flanking *DXPas34* that are absent in Δ Pas34 ES cells, whilst in (**c**) the red triangle show the location of the loxP site remaining after *DXPas34* deletion, and the new primer pair designed at the position shown in red. ChIP analysis of Rex1 (**d, e, f**), Klf4 (**g, h, i**), and c-Myc (**j, k, l**) in the indicated cell lines. The X-axis shows the genomic coordinates (Kb) relative to *Xist* transcription start site. The vertical black bars represent the *Xist* and *Tsix* transcription start sites. The Y-axis shows the average percentage of immunoprecipitation. The number of biological replicates used was: d: n=2, e: n=3, f: n=6; g: n=3, h: n=3, i: n=6; j: n=3, k: n=3, and l: n=6. All the results are expressed as means $\pm$ sem.

Figure 2: Developmentally-induced loss of Rex1, Klf4 and c-Myc binding correlates with *Tsix* repression. **a**, Relative RNA levels of the indicated genes in undifferentiated (-Tc, black, set to one) and differentiating (96h Tc treatment, +Tc, red) ZhbTc4.1 cells (n=4). **b, c, d, e, f**, ChIP analysis of RNAPII (**b**), TFIIB (**c**), Rex1 (**d**), Klf4 (**e**) and c-Myc (**f**) across the *Tsix* 5' region in the same cellular conditions (n=2). **g, h, i**, Analysis of Rex1, Klf4 and c-Myc binding at the *Tsix* 5' region in ES cells (ES, CK35, black, n=3), TS cells (TS, F2, red, n=3), and MEFs (green, n=3). The X-axis shows the genomic coordinates (Kb) relative to *Xist* transcription start site. All the results are expressed as means $\pm$ sem.

Figure 3: Rex1 is required for efficient elongation of *Tsix* transcription. **a**, Relative gene expression of a stable clone (Clone #1a in Supplementary Fig. 3) expressing an shRNA against *Rex1* (n=4). **b**, ChIP analysis of Rex1, Klf4 and c-Myc binding at the positions identified as providing maximal binding in Fig. 1 (labelled *Tsix*) and at a negative control position (labelled *ctl* and corresponding to *Tpg*, see Supplementary Fig.8) in control (black) and *Rex1*-interfered cells (red, n=3). **c**, ChIP analysis of the RNAPII across the *Tsix* 3' and 5' regions in control and *Rex1*-interfered cells (n=3, m=5). **d**, ChIP analysis of H3K36Me3 across the *Tsix* 3' and 5' regions in control and *Rex1*-interfered cells (n=2, m=3). The X-axis shows the genomic coordinates (Kb) relative to *Xist* transcription start site. All the results are expressed as means $\pm$ sem. **e**, Transcriptional network coupling pluripotency regulators to X-inactivation. In a previous study we showed that Nanog, Oct4 and Sox2 bind *Xist* intron 1 to repress *Xist* transcription in ES cells. Here we have shown that the pluripotency-associated Rex1 protein, in conjunction with the reprogramming factors Klf4 and c-Myc, bind to the *DXPas34/Tsix* 5' region to confer *Tsix* maximal activity in ES cells. Moreover, Oct4, Sox2 and Klf4 bind to *Xite*, a weak enhancer of *Tsix*. Whilst Nanog, Oct4 and Sox2 suppress *Xist* transcription, facilitating the reactivation of the inactive X, Rex1, Klf4 and c-Myc transactivate *Tsix*, which, in turn, modifies the *Xist* chromatin structure to render all *Xist* alleles epigenetically indistinguishable and permitting random *Xist* transcription upon differentiation. *Tsix* may additionally block *Xist* RNA accumulation at the post-transcriptional level of regulation. We conclude that the road to pluripotency and the path of X-inactivation regulation during both development and *in-vitro* reprogramming experiments are directly coupled through the stringent control of the two main non-coding actors by distinct pluripotency-associated regulatory complexes.

Methods.

RNA Extraction and Random Reverse Transcription

One to 5 million cells were lysed in TRIZOL (Invitrogen) and RNA was then chlorophorm extracted and isopropanol precipitated. After DNase treatment (Quiagen), RNA was phenol/chlorophorm re-extracted, ethanol precipitated, resuspended in water and quantified.

One to 4µg of RNA were used per RT reaction. RNA was denatured in the presence of 1µg of random hexamers (Roche) for 5' at 90°C, and reverse transcribed in a final volume of 20µl with 100U of SuperScriptII (Invitrogen) at 42°C for 60' followed by heat inactivation at 70°C for 15'. Synthesised cDNAs were diluted in 280µl of water and stored at -20°C until used.

Chromatin Extraction and Chromatin ImmunoPrecipitation

Twenty million cells were resuspended in 3ml of pre-warmed DMEM-FCS10% and crosslinked with 1% formaldehyde (Sigma) for 10' at room temperature. The reaction was quenched with 0.125mM glycine for 5' at room temperature. Cells were spun down for 3' at 1300rpm at 4°C, and washed twice with cold PBS1X (Invitrogen). Cell pellets were then vigorously resuspended in 300µl of Sweeling buffer (5mM Pipes pH8, 85mM KCl) freshly supplemented with 1X protease inhibitor cocktail (Roche) and 0.5% NP-40. The suspension was incubated for 20' on ice with occasional gently shaking. One microliter of suspension was used to check for the completeness of total nuclei extraction under the microscope. Nuclei were spun down in 15ml conical tubes for 10' at 1500rpm at 4°C and resuspended in 1,5ml of TSE150 (0.1% SDS, 1% Triton, 2mM EDTA, 20mM Tris-HCl pH8, 150mM NaCl) buffer, freshly supplemented with 1X protease inhibitor cocktail. Samples were sonicated at 4°C in

15ml conical tubes using a Bioruptor (Diagenode) for 5 cycles of 10' divided into 30''ON-30''OFF subcycles at maximum power. Chromatin was then transferred into 1,5ml tubes and centrifuged for 30' at 14000rpm at 4°C. Soluble chromatin was aliquoted and stored at -80°C until use. Twenty microlitres were used for quantity and quality controls of the DNA.

Twenty micrograms of DNA were used for each ChIP. For each experiment, the required amount of chromatin was defrosted (generally between 40 to 100µg -1 to 5 ChIPs per sample) and pre-cleared for 1h30 with rotation at 4°C in 1ml of TSE150 with 50µl of pA/pG sepharose beads (Sigma) 50% slurry, previously blocked with 500µg/ml of molecular grade BSA (Roche) and 1µg/ml of yeast tRNA (Invitrogen). Pre-cleared chromatin was transferred into fresh tubes after 1' centrifugation at 3000rpm and aliquoted accordingly. Twenty micrograms of diluted chromatin was aliquoted for input DNA extraction and precipitation. Immunoprecipitation with specific antibodies (1 to 5µg each, see Supplementary Figure 8) was performed overnight with rotation at 4°C, in a final volume of 500µl. Immunocomplexes were recovered with 50µl of blocked pA/pG sepharose beads 50% slurry for 1h30 with rotation at 4°C. Beads were recovered by 1' centrifugation at 3000rpm and washed at room temperature in 1ml of TSE150, TSE500 (0.1% SDS, 1% Triton, 2mM EDTA, 20mM Tris-HCl pH8, 500mM NaCl), Washing Buffer (10mM Tris-HCl pH8, 0.25M LiCl, 0.5% NP40, 0.5% Na-Deoxycholate, 1mM EDTA), TE (10mM Tris-HCl pH8, 1mM EDTA) and TE. Each wash was performed for 5' with rotation at room temperature. After the last wash, elution was performed in 100µl of Elution Buffer (1% SDS, 10mM EDTA, 50mM Tris-HCl pH8) for 15' at 65°C after vigorous vortexing. Eluates were collected after 1' centrifugation at 14000rpm, and the beads rinsed in 150µl of TE-SDS1%. After 1' centrifugation at 14000rpm, the supernatant was pooled with the corresponding first eluate. Crosslinking of ChIP and input fractions was reversed overnight at 65°C, followed by proteinase K treatment (Invitrogen), phenol/chlorophorm extraction and ethanol precipitation.

DNA pellets corresponding to the input fractions were resuspended in 300µl of water, whilst those corresponding to the ChIP fraction were resuspended in 100µl in the case of transcription factors, or in 300µl in the case of histone modifications.

Transient siRNA knock-downs

The medium of subconfluent ES cell cultures was changed 6h before nucleofection. After cell harvesting, 5 million cells were pelleted in individual tubes, washed with PBS1X (Invitrogen), and resuspended in 90µl of completed nucleofection solution (Amaxa). This cellular suspension was mixed with 10µl of siRNA (150nM) or shRNA expressing vectors (4µg) and transferred into nucleofection cuvettes (Amaxa) that were placed in the nucleofection device. Program A30, which was used for all experiments, consistently gave more than 80% efficiency as evaluated by the nucleofection of a GFP expressing vector, with around 50% immediate mortality. Nucleofected cells were collected using the pipettes provided by Amaxa into 500µl of prewarmed ES cell medium, and transferred into 25cm² gelatinised flasks containing 10ml of prewarmed ES cell medium. Twenty four hours later, cells were collected for analysis and RNA extraction.

Real-Time PCR analysis

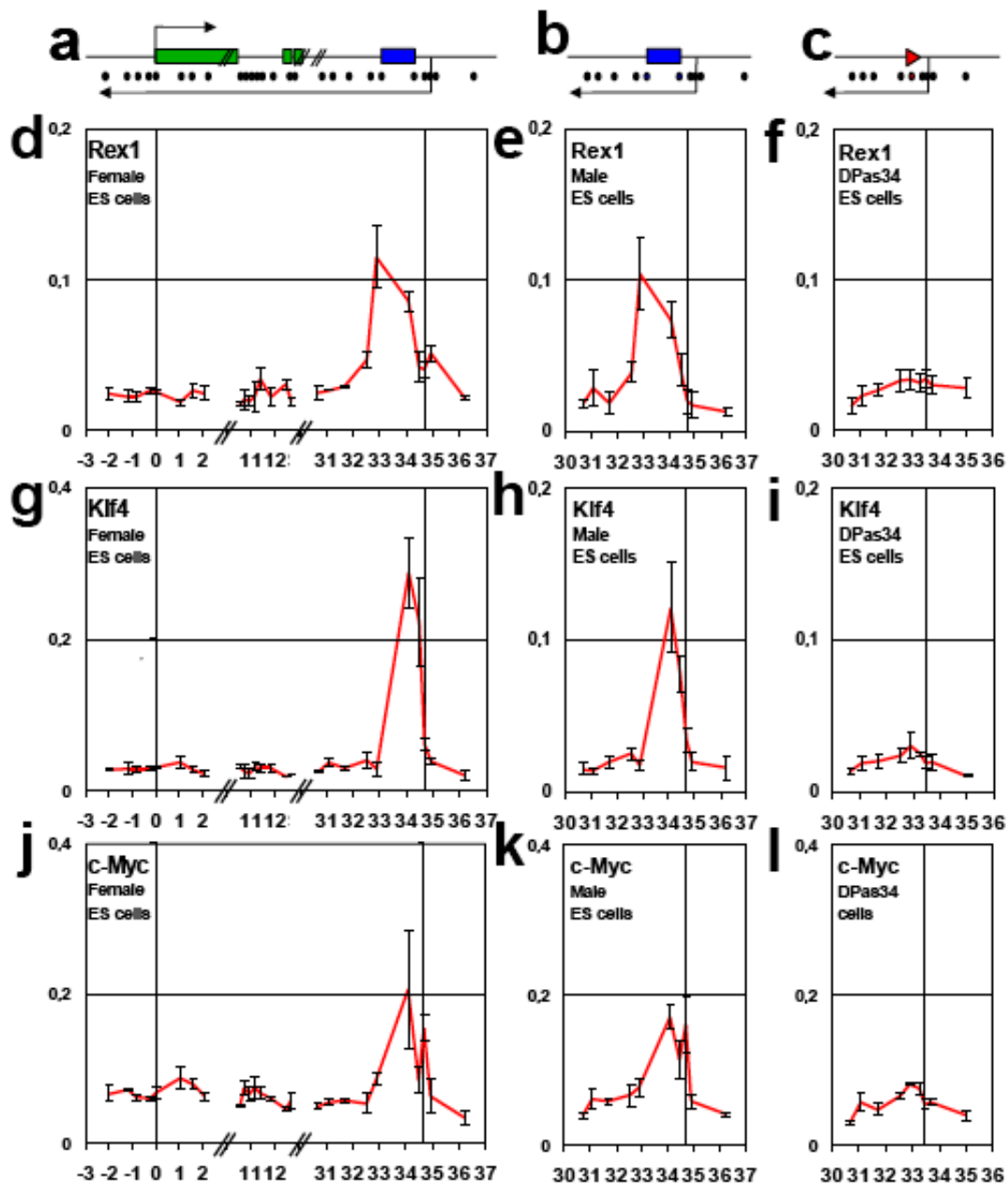
Two systems of PCR analysis were exploited. All analyses except those corresponding to the *Rex1* knock-downs were analysed in 96-well plates using a StepOnePlus PCR machine (Applied Biosystems) and the Power Sybr Green PCR Master Mix (Applied Biosystems). *Rex1* knock-down experiments were analysed in 384-wells plates with a 480 LightCycler (Roche) using LightCycler 480 SYBR Green I Master (Roche). All reactions were performed in duplicate. Five microliters of DNA were used per reaction. Standard curves of all primers were performed to check for efficient amplification (above 90%). Melting curves were also performed to verify production of single DNA species

with each primer pair. All primer sequences are available in Supplementary Figure 8.

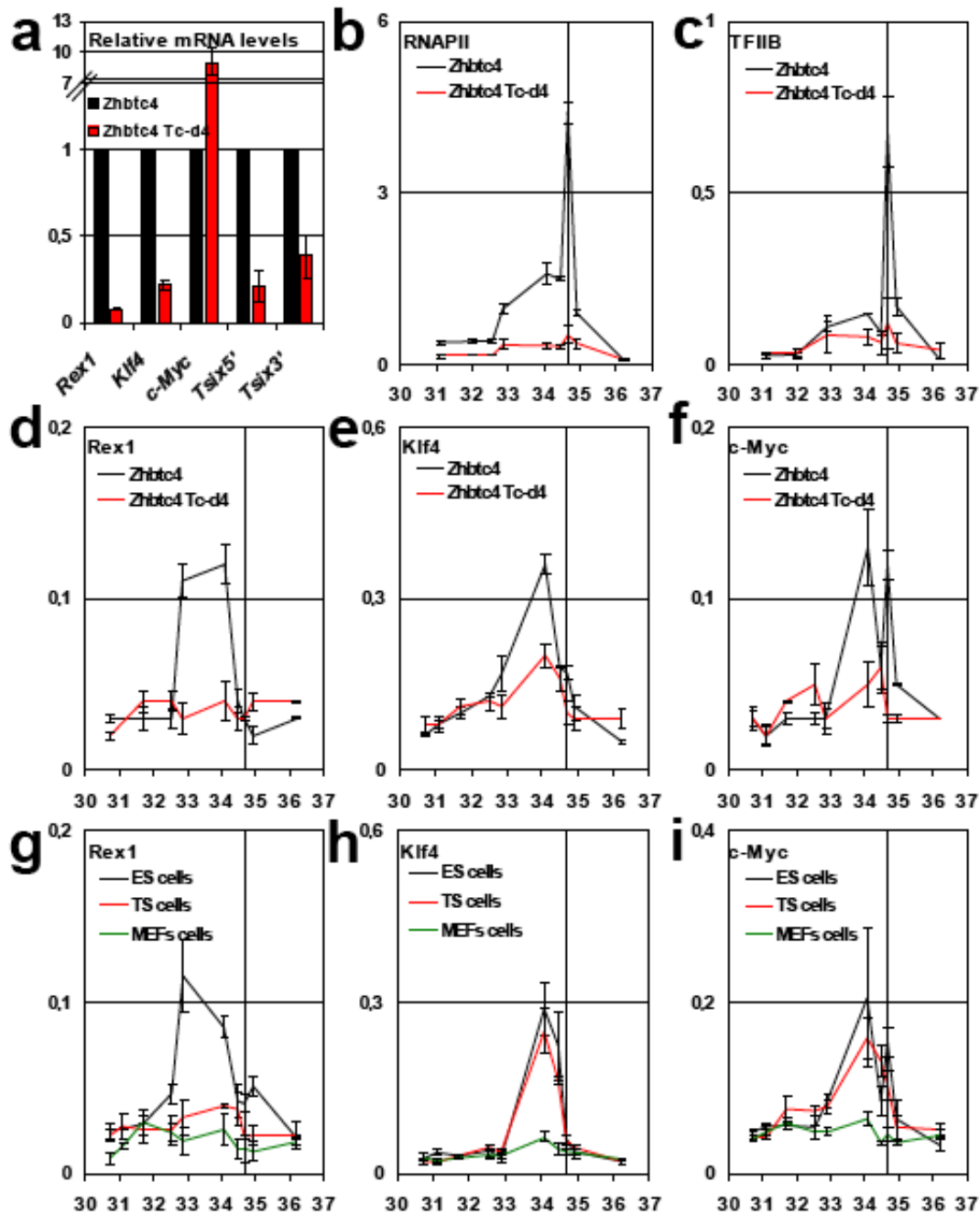
Relative levels of expression in each assay were obtained through the $\Delta\Delta C_t$ method, using (i) *ArpoP0* mRNA levels as a reporter in all experiments except in *Rex1* knockdowns in which *Tbp* was used, and (ii) the appropriate control cell line or cellular state as the reference sample.

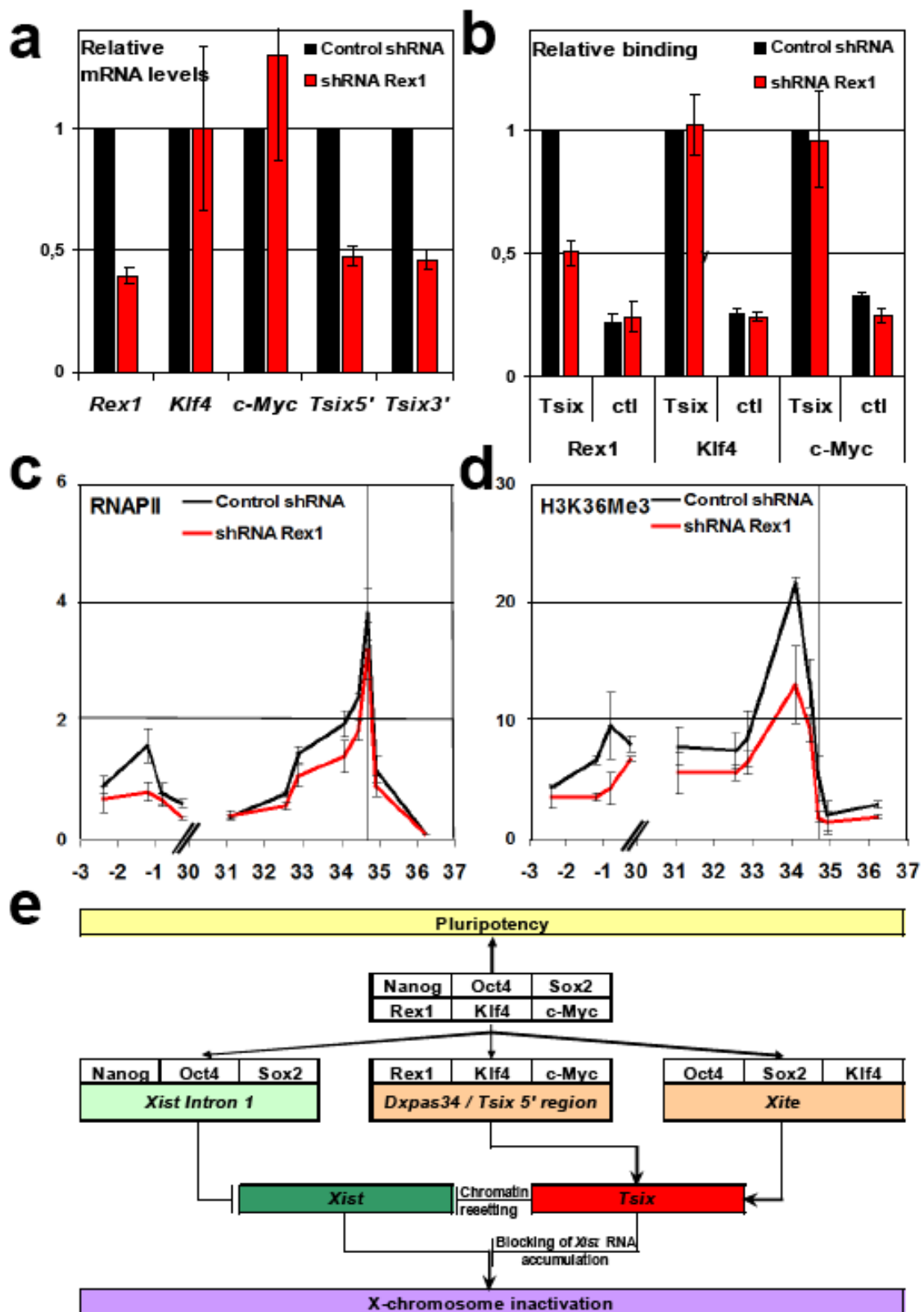
Enrichment levels in ChIP assays are expressed as a percentage of immunoprecipitation relative to the input. Essentially, the ΔC_t method was used to calculate a ChIP over input ratio that was corrected by the appropriate dilution factor of each analysed fraction, and multiplied by 100 to get the percentage of immunoprecipitation.

Navarro et al. Figure 1



Navarro et al. Figure 2





Supplementary Information – Navarro et al.

Molecular coupling of *Tsix* regulation and pluripotency.

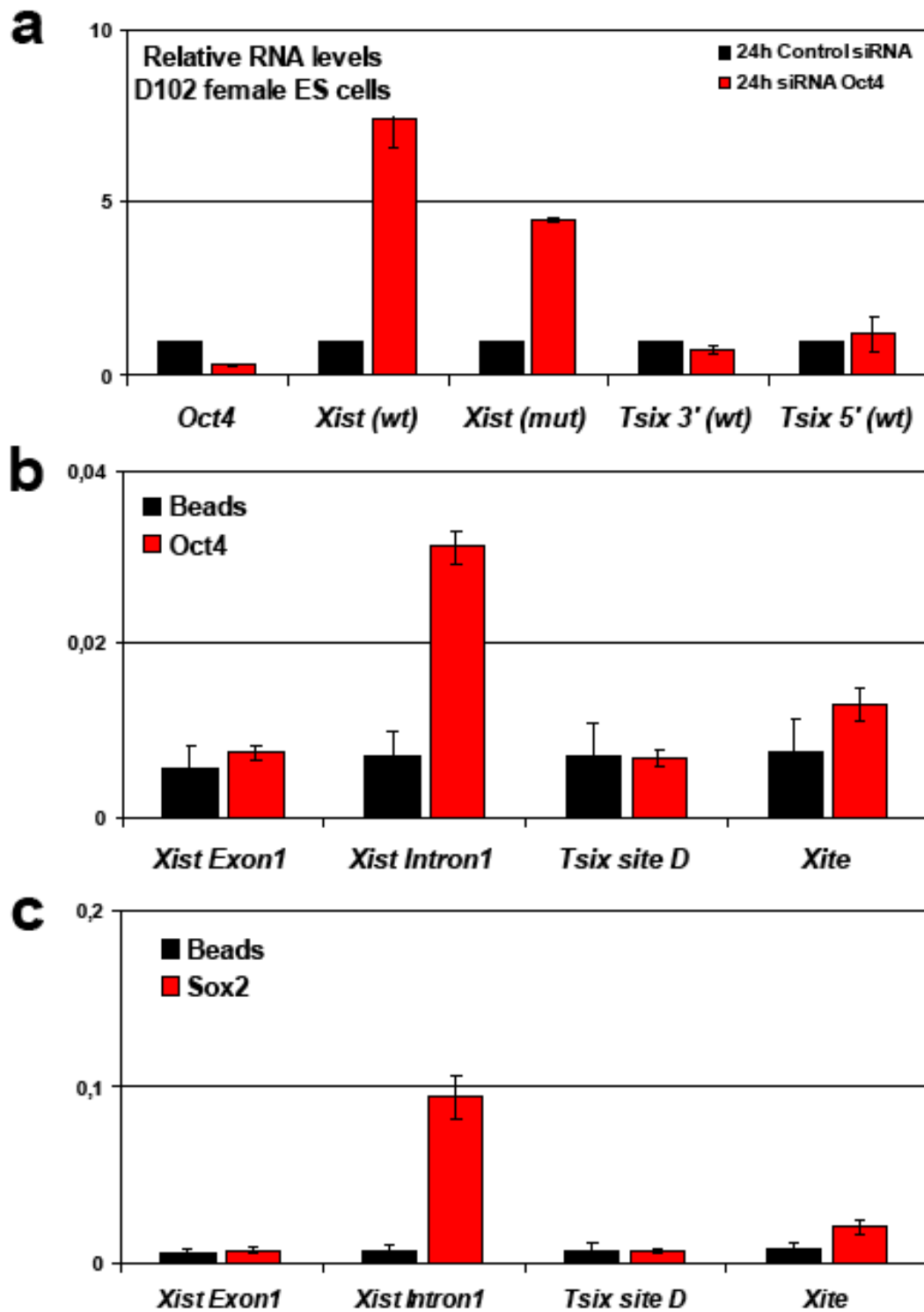
This file contains:

1/ Supplementary Figures 1 to 8: pages 2 to 9

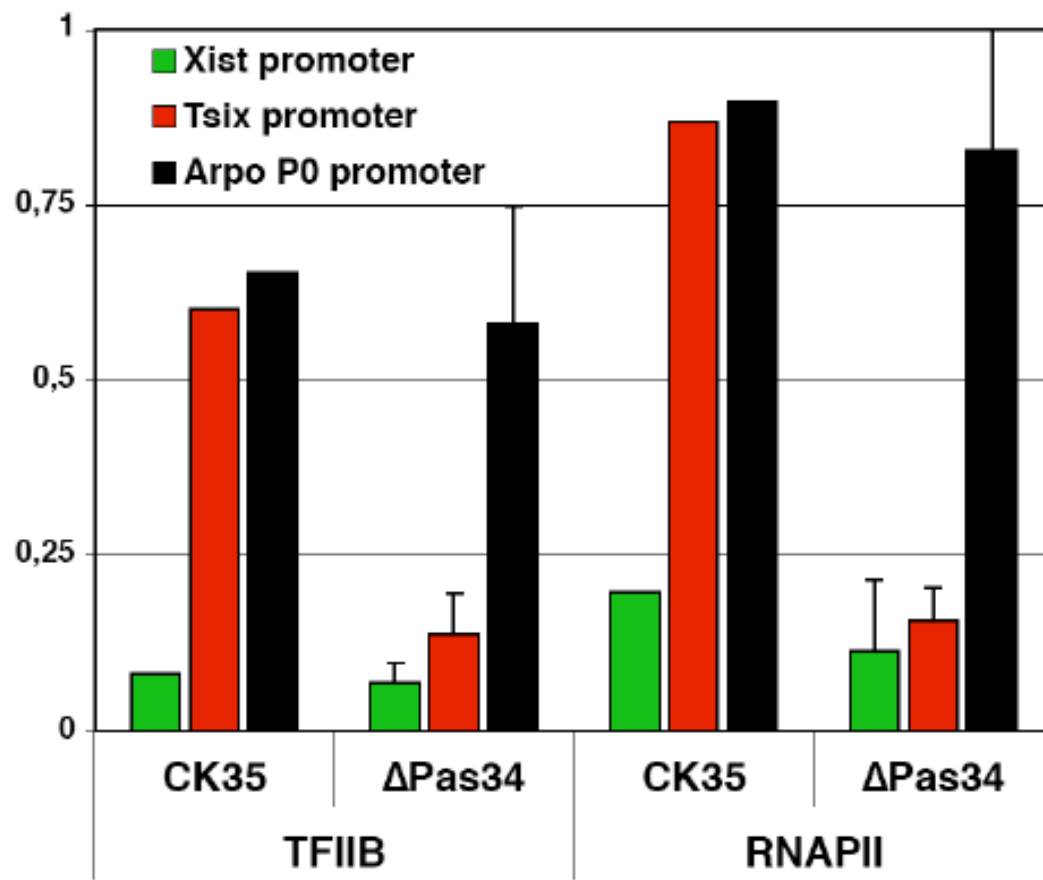
2/ Legends to Supplementary Figures 1 to 8: pages 10 to 13

3/ Supplementary References: pages 14 and 15

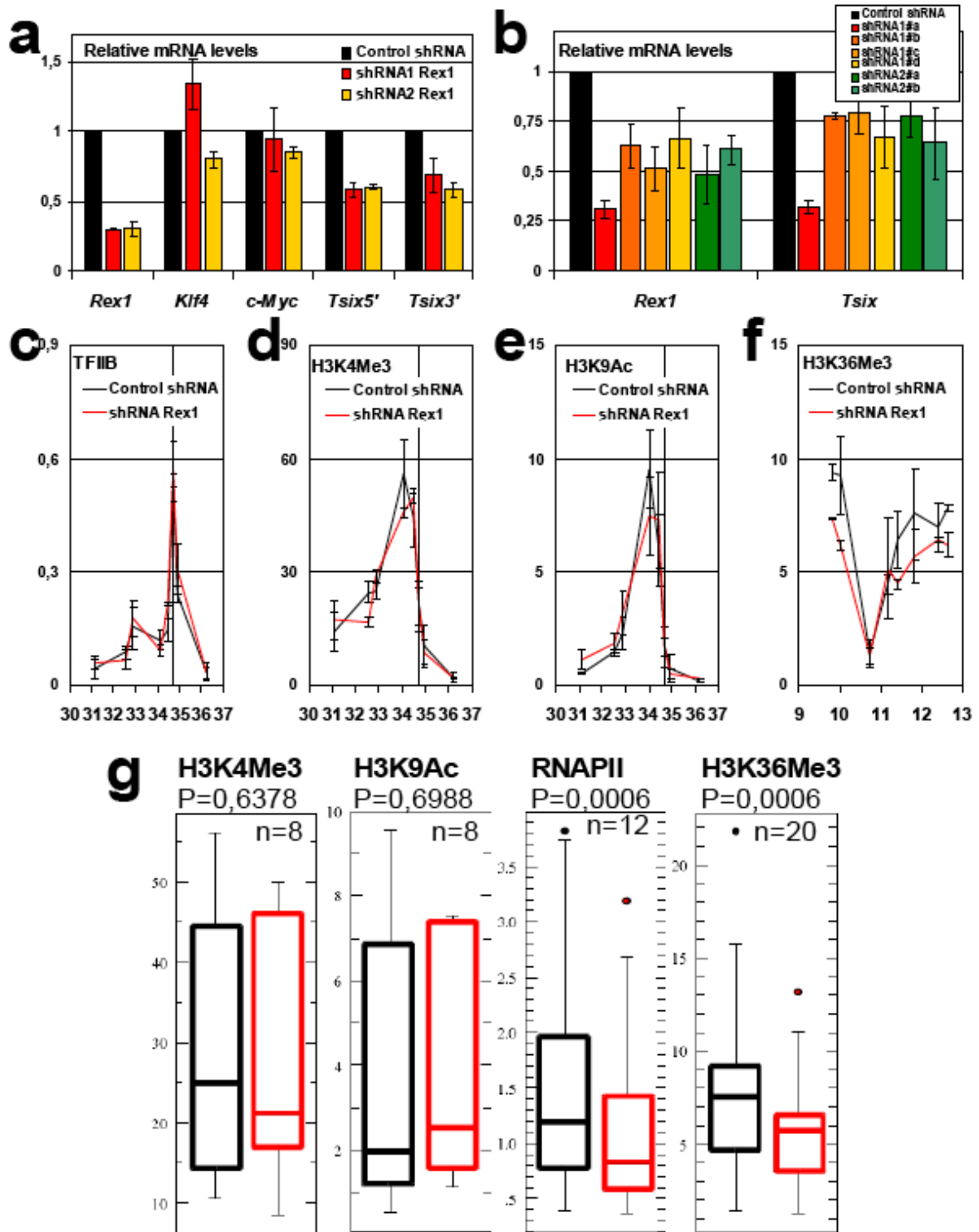
Navarro et al. Supplementary Figure 1



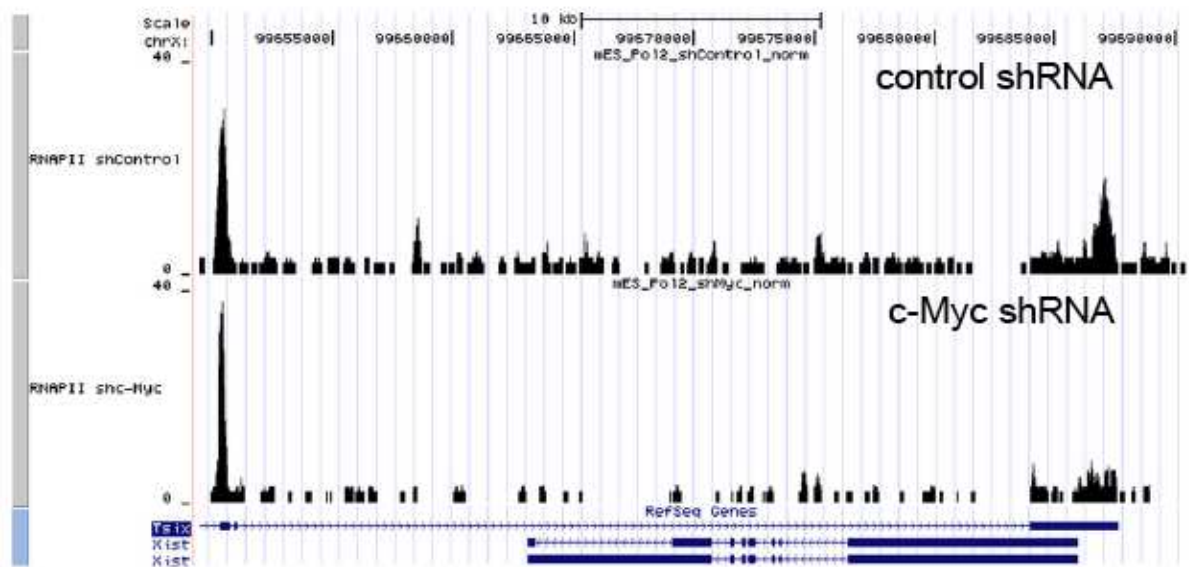
Navarro et al. Supplementary Figure 2



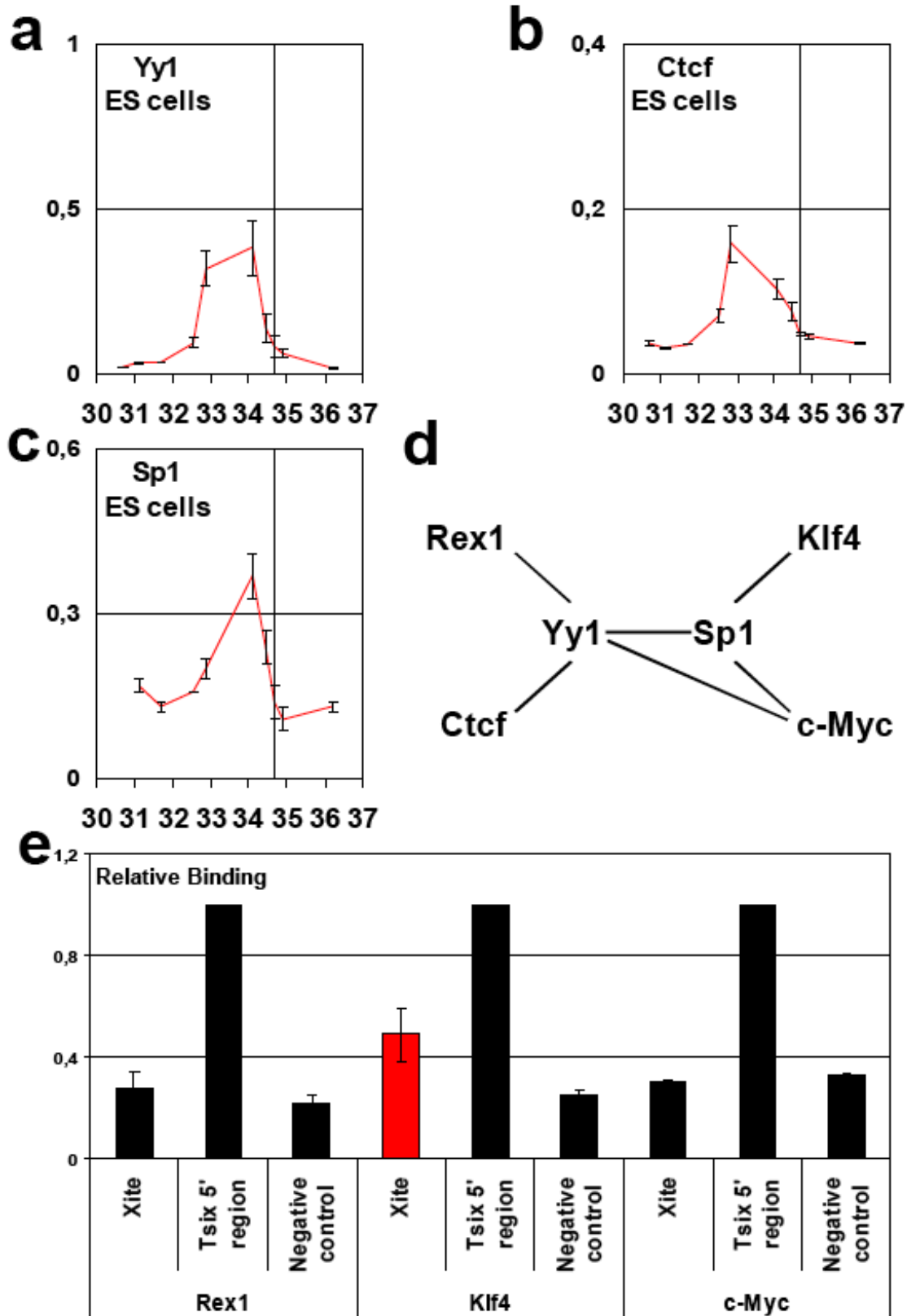
Navarro et al. Supplementary Figure 3



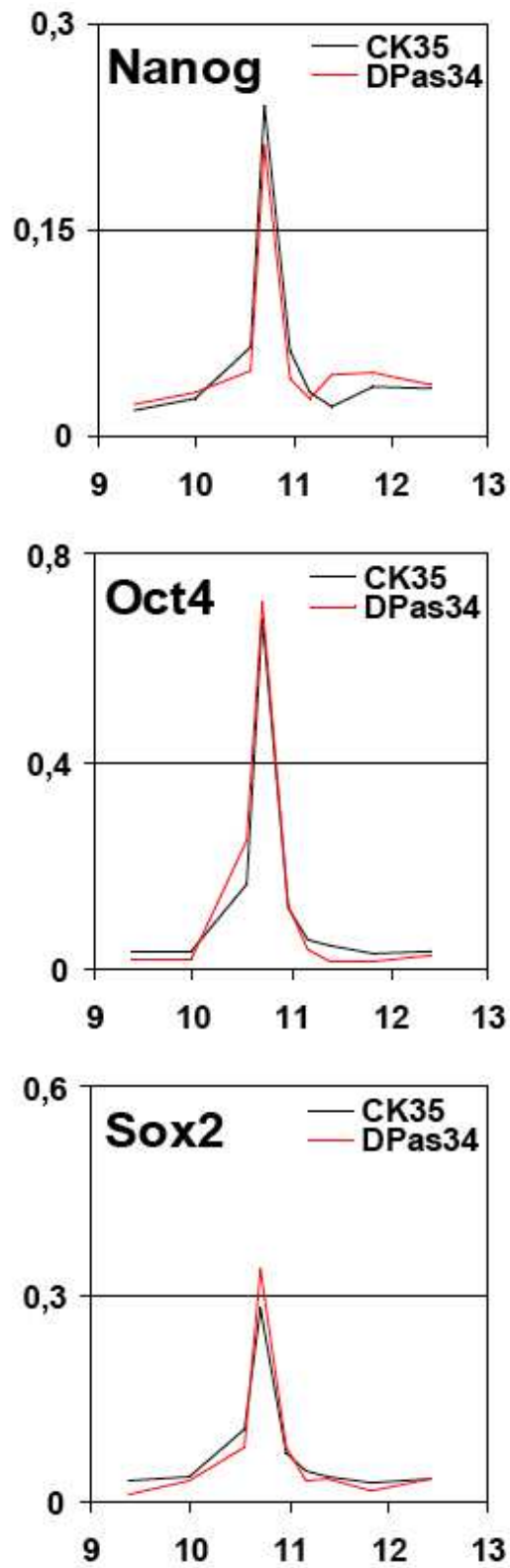
Navarro et al. Supplementary Figure 4



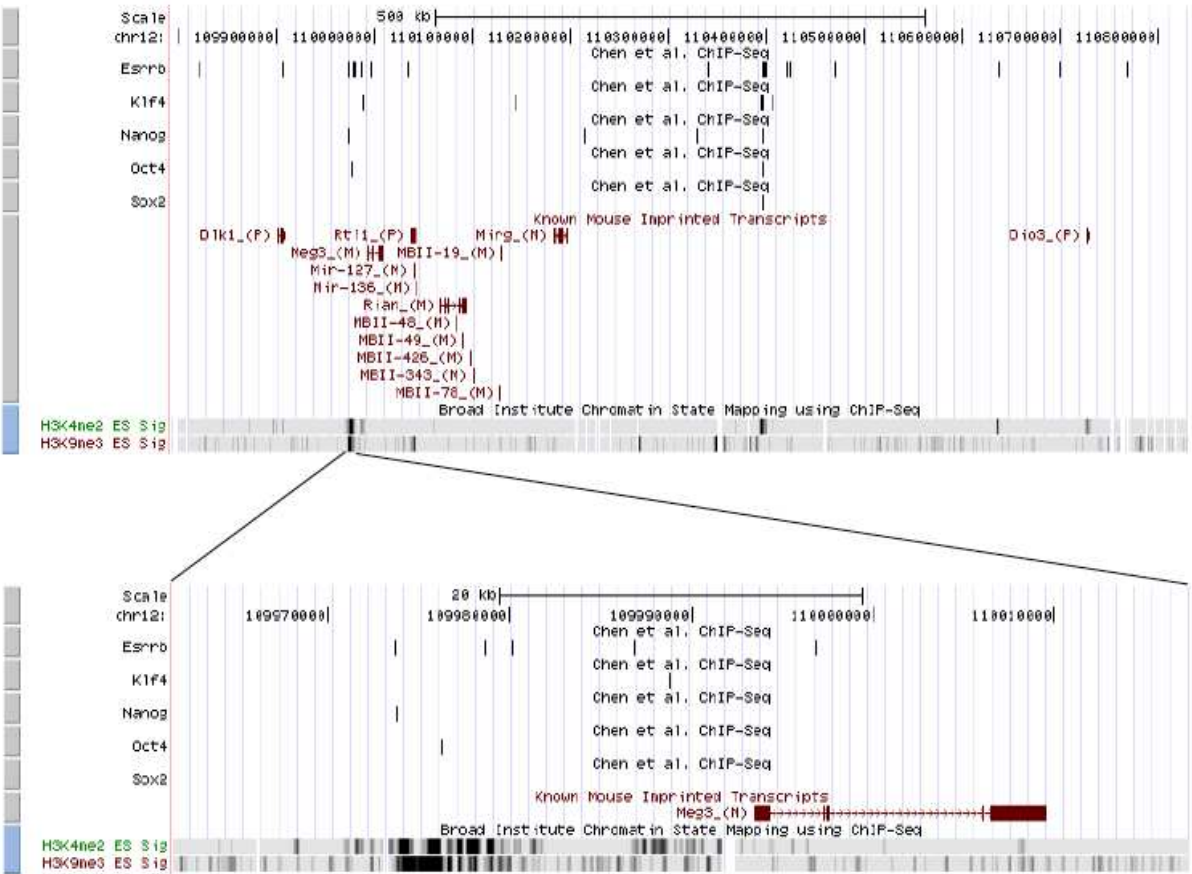
Navarro et al. Supplementary Figure 5



Navarro et al. Supplementary Figure 6



Navarro et al. Supplementary Figure 7



Navarro et al. Supplementary Figure 8

Antibodies		
Nanog	CosmoBio	REC-RCAB0002PF
Oct3/4	Santa Cruz	sc-8628
Sox2	Santa Cruz	sc-17320
Rex1	Santa Cruz	sc-50668X
Klf4	Santa Cruz	sc-20691
c-Myc	Abcam	ab56
Yy1	Santa Cruz	sc-1703X
Ctcf	Santa Cruz	sc-15914X
Sp1	Santa Cruz	sc-59X
TFIIB	Santa Cruz	sc-225
RNAPII	Euromedex	PB-7C2
H3K4Me3	Diagenode	pAb-003-050
H3K9Ac	Diagenode	pAb-004-044
H3K36Me3	Diagenode	pAb-058-050

ChIP primers			
coordinate	name	forward	reverse
-2,4	u3	cttgggagtggttttcttcaaccat	gcacccagttctctccagggaat
-1,2	x9	tgaccagtgacctgcaagttc	ctaagagcacctggctccac
-0,8	u2	ccctacctgaaccacctcaatagt	agttcccttttaggcgtccat
-0,25	u1	ggaaccaaggagccattttgt	cttctgcattagttggcgacc
0	X1	taaaggtccaataagatgtcagaa	ggagagaaaccacgggaagaa
1,05	XP1-2	acggctattctcagagccagtt	attggttgcttttatccagt
1,57	X5	tcattgtgacctgacctctagt	cacctaccataatgcacca
2,5	XP2-3	ttgttgatgctatattgactaatggtttt	cagaaggttgaaatgttaggccttaatt
9,8	XEx1b	catcaggtctggcagcaa	ttcatcagcaatgtcatatcaaaaca
10	XIn1b	ccttggttggtccagacgaactatt	ctttccatgcttccaagaggtag
10,72	Xin1	aacccttttaagtcactgtaaattcc	tagagagccagacaatgtctaagcc
11,17	Xin1c	tcaagttgtgtgagcaggcag	gccgtggaggtcatgttagga
11,4	Xin1d	gtcctgaccagccctca	aggcagcctgaagatggtga
11,82	Xin1e	ccagaccaggtctttgtatgca	cactggactgggagagagggg
12,41	XEx2	gtgctcctgctgctcaagaagaa	gcactcttccctcctctaaatccag
12,64	XEx3	caatccctatgtgagctcaggact	cagagtgcgaggccttgcaagaga
31,09	cta	ggagccctaaacctgtctgtc	gtgtgtcctagctcaaggagg
32,54	tpc	taatccaaacgggtacaaacagaa	cccgccatgtgattatgctaa
32,86	tpd	acgttttgcatctctccctgtaaca	ctgggcagagcagaggtga
34,08	ctc	taaccacctgtgaaggacag	aatgcttgccagctatgcgg
34,46	tpe	aacaagggaacttcgggcaact	tgagatgcgggctaaggaga
34,68	x11	gcgcttgccaggtacttttg	aagagccttaggtcccgcc
34,91	tpf	gaccttggtgggagtcacagaaaa	aggcaaatgggctcagggt
36,21	tpg	tggagctctttctatgtttctcctt	atgaatgggcttcttgaatttctact
	Tsix siteE	CGTAGAGATGGTTGGCTAAAATGTTAC	CCCTGATGTGAGTGTGTGGATTTC
	Xite	CAAGGTTGGGAACAAGGTATATCAGG	GGACAAGGGACAGAAAGTCTTATTTTAC

RT-PCR primers		
Nanog	aggatgaagtgcagcggtg	tgttgagcccttctgaatcag
Oct4	ccccaatgccgtgaagttg	tcagcagcttggcaaaactgtt
Sox2	cacagatgcaaccgatgca	ggtgccctgtcgcagta
Rex1	CAGCTCCTGCACACAGAAAG	ACTGATCCGCAAAACCTG
Klf4	CAAGTCCCCTCTCTCCATTATCAAGAG	CCACTACGTGGGATTTAAAAGTGCCTC
cMyc	tgagccctagtgtgcatt	tcttgcctctcttcagagtcgct
Tbp	ggggagctgtgatgtgaagt	ccaggaaataattctggctca

Supplementary Figure Legends.

Supplementary Figure 1: Oct4/Sox2 are not the main regulators of *Tsix* in ES cells. a,

In a previous report, Lee and colleagues showed that Oct4 and Sox2 bind to *DXPas34* and *Xite* to control *Tsix* transcription in ES cells¹. We have shown that Oct4 and Sox2 bind at *Xist* intron 1 to silence *Xist*². Given that *Tsix* negatively influences *Xist* expression, the finding by Lee et al. suggests that the upregulation of *Xist* we had observed upon depletion of Oct4 could be dependent on the regulation of *Tsix* by Oct4. To test this, we knocked-down *Oct4* mRNA levels in female D102 cells³. In these cells, one X-chromosome carries a 65 kb-long deletion that encompasses the *Tsix* promoter, *DXPas34* and *Xite*. We reasoned that, if the effect of *Oct4* knock-down on *Xist* were dependent on a downregulation of *Tsix*, then *Xist* should not be upregulated from the mutant X-chromosome. After 24h post nucleofection, *Oct4* mRNA levels are greatly reduced and *Xist* upregulation established from both the wt and mutant X chromosomes (previously described allelic real-time PCR was used for this assay³). This demonstrates that *Tsix* is not required for *Xist* upregulation upon *Oct4* knock-down. Importantly, we could not detect, after 24h of interference, any significant modification of *Tsix* expression, which in this cell line is produced by the wt allele only. Thus, we conclude that Oct4 mediates *Xist* silencing independently of *Tsix*. Further, the absence of modification of *Tsix* expression by *Oct4* knock-down indicates that *Tsix* is not a primary target of Oct4. **b, c,** Since *Tsix* expression does not show any major variation upon *Oct4* knock-down, and we were not able in our previous report² to detect Oct4 binding across the *Tsix* 5' region, we wished to try to reproduce the ChIP experiments published by Lee and colleagues. To this end we used the same primer pairs amplifying *DXPas34* and *Xite* that had been shown by Lee

and colleagues to be positive in their Oct4 and Sox2 ChIP experiments. Oct4 (**b**) and Sox2 (**c**) showed very low levels of binding to *Xite* but not to *DXPas34* (Tsix site D). Since binding to *DXPas34* has similarly not been reported in published ChIP-Seq experiments⁴, we conclude that binding at *DXPas34*, if any, is unlikely to be of biological significance. Low levels of binding are detected at *Xite*, which is not, however, a strong enhancer of *Tsix*. If our results argue against a major role of Oct4 and Sox2 on the regulation of *Tsix* in undifferentiated ES cells, the crucial function that Lee and colleagues attributed to them during the early stages of differentiation¹, in particular concerning X-chromosome pairing, is not addressed by our work.

Supplementary Figure 2: Binding of TFIIB and RNAPII at *Xist* and *Tsix* promoters in wild-type and Δ Pas34 ES cells. ChIP analysis of TFIIB and RNAPII binding to the *Xist*, *Tsix* and *Arpo P0* promoters in wild-type and Δ Pas34 ES cells, demonstrating that in the absence of *DXPas34* the *Tsix* promoter is repressed (note that the *Xist* promoter remains repressed in Δ Pas34 ES cells).

Supplementary Figure 3: Additional analyses of *Rex1* knock-down cells. **a**, Relative gene expression 24 hours after *Rex1* knock-down using the independent shRNA1Rex1 and shRNA2Rex1 constructs (n=2). **b**, Relative *Rex1* and *Tsix* expression levels in 4 independent clones stably carrying a vector expressing shRNA1Rex1 and 2 independent clones stably expressing shRNA2Rex1, compared to a clone stably expressing a control shRNA against *geminin* (set here arbitrarily to one). Clone #1a was retained for further analysis.

c, d, e, Analysis of TFIIB (**c**, n=2), H3K4Me3 (**d**, n=2), and H3K9Ac (**e**, n=2) across the *Tsix* 5' region, in control and *Rex1*-interfered cells (Clone #1a). **f**, ChIP analysis of H3K36Me3 across *Xist* intron 1, showing reduced levels of enrichment in *Rex1* knock-down cells (Clone #1a). The position which shows very low levels of enrichment in both control and knockdown cells corresponds to the region of Nanog, Oct4 and Sox2 binding, where a DnaseI hypersensitive site has also been shown to be located. **g**, Boxplots of ChIP data corresponding to the factor indicated (in this case, 'n' refers to the number of positions analysed by ChIP and used to generate the boxplots). P-values were obtained with a paired t-test in which the average %IP obtained at each location was compared between control and *Rex1* knock-down cells.

Supplementary Figure 4: c-Myc also controls *Tsix* transcription elongation. ChIP-seq results of RNAPII binding in control and c-Myc knock-down ES cells⁵ were uploaded to the UCSC browser (<http://genome.ucsc.edu/cgi-bin/hgGateway>) as custom tracks. The top panel shows RNAPII distribution across the *Tsix* locus. Note that ChIP-Seq identifies the two regions of accumulation (corresponding to the 5' and 3' ends) that we have identified here and elsewhere. The bottom panel shows the distribution of RNAPII in *c-Myc* knock-down cells: the density of signal across *Tsix*, in particular at the 3' end, is reduced as compared to the control. In contrast, the *Tsix* promoter remains competent to recruit the RNAPII at high levels. These results are highly similar to that we report here upon *Rex1* knock-down (Figure 3, and Supplementary Fig. 3). Note that the orientations of *Xist/Tsix* are the opposite to that of our figures.

Supplementary Figure 5: Additional binding activities at *DXPas34* suggest a protein interaction network. **a, b**, Validation of previously reported binding of Yy1 (**a**) and Ctfc (**b**) to *DXPas34*. Yy1 has been shown to interact with Rex1⁶, Ctfc⁷, c-Myc⁸ and Sp1⁹. Sp1 also interacts with Klf4¹⁰ and c-Myc¹¹. The binding of Sp1 to the *Tsix* 5' region that we report here (**c**) supports a protein interaction network in which Yy1 and Sp1 play a fundamental structural role (**d**). **e**, Klf4 binds at *Xite* (red bar), although at levels lower than to the *Tsix* 5' region. Data corresponding to this panel was generated in parallel to that of Fig. 3b.

Supplementary Figure 6: Binding of Nanog, Oct4 and Sox2 to *Xist* intron 1 in wild type and Δ Pas34 ES cells. ChIP analysis of Nanog (**a**), Oct4 (**b**), and Sox2 (**c**) binding across *Xist* intron 1 in CK35 and Δ Pas34 cells.

Supplementary Figure 7: Abundant binding across the imprinted locus on chromosome 12qF1. ChIP-seq⁴ identified binding sites for Nanog, Oct4, Sox2, Klf4 and Esrrb were uploaded to the UCSC browser (<http://genome.ucsc.edu/cgi-bin/hgGateway>) as custom tracks. Known imprinted genes were similarly uploaded. Binding profiles were correlated to enrichment for H3K4Me2 and H3K9Me3 that mark imprinting centres¹². Nanog, Oct4 and Esrrb bind around the upstream region of the non-coding *Meg3* gene, where the imprinting centre of the cluster is located.

Supplementary Figure 8: Primer and antibody information.

Supplementary references.

- ¹ Donohoe, M. E. et al., The pluripotency factor Oct4 interacts with Ctf and also controls X-chromosome pairing and counting. *Nature* **460** (7251), 128 (2009).
- ² Navarro, P. et al., Molecular coupling of Xist regulation and pluripotency. *Science* **321** (5896), 1693 (2008).
- ³ Clerc, P. and Avner, P., Role of the region 3' to Xist exon 6 in the counting process of X-chromosome inactivation. *Nat Genet* **19** (3), 249 (1998).
- ⁴ Chen, X. et al., Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133** (6), 1106 (2008).
- ⁵ Rahl, P. B. et al., c-Myc regulates transcriptional pause release. *Cell* **141** (3), 432 (2010).
- ⁶ Wang, J. et al., A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444** (7117), 364 (2006).
- ⁷ Donohoe, M. E. et al., Identification of a Ctf cofactor, Yy1, for the X chromosome binary switch. *Mol Cell* **25** (1), 43 (2007).
- ⁸ Stojanova, A. et al., Repression of the human immunodeficiency virus type-1 long terminal repeat by the c-Myc oncoprotein. *J Cell Biochem* **92** (2), 400 (2004).
- ⁹ Lee, J. S., Galvin, K. M., and Shi, Y., Evidence for physical interaction between the zinc-finger transcription factors YY1 and Sp1. *Proc Natl Acad Sci U S A* **90** (13), 6145 (1993).

¹⁰ Zhang, W. et al., The gut-enriched Kruppel-like factor suppresses the activity of the CYP1A1 promoter in an Sp1-dependent fashion. *J Biol Chem* **273** (28), 17917 (1998).

¹¹ Gartel, A. L. et al., Myc represses the p21(WAF1/CIP1) promoter and interacts with Sp1/Sp3. *Proc Natl Acad Sci U S A* **98** (8), 4510 (2001).

¹² Mikkelsen, T. S. et al., Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448** (7153), 553 (2007).

b. Discussion

Le fait de pouvoir relier la pluripotence cellulaire à l'expression de *Tsix* souligne bien son importance capitale dans l'établissement d'un état inhibant l'initiation de l'inactivation (à travers son action répressive sur *Xist*), tout en assurant un choix non-biaisé du futur allèle inactif au sein des cellules ES indifférenciées (grâce à l'expression de *Tsix* qui établit une structure chromatinienne identique sur les deux allèles).

L'expression de *Tsix* semble en effet jouer un double rôle par rapport à *Xist* (voir le chapitre : « Régulation de *Xist* par *Tsix* », p72) : sa transcription à travers le gène *Xist* induisant des modifications hétérochromatiques au niveau du promoteur ainsi qu'une répression post-transcriptionnelle de ce dernier par des mécanismes encore inconnus (Navarro *et al.*, 2005), tandis qu'elle établit plutôt une structure euchromatinienne le long du locus *Xist/Tsix* dans le but d'effacer les différences chromatinienes induites par une inactivation passée du chromosomes X (Navarro *et al.*, 2006) (l'inactivation empreintée).

À travers cet article, nous avons démontré le rôle de plusieurs facteurs de pluripotence dans la surexpression de *Tsix* au sein des cellules non-différenciées. Il semblerait en effet que dans ce contexte, Rex1 et c-Myc agissent de manière synergique afin d'obtenir une élongation particulièrement efficace de l'antisens. Comme stipulé dans l'article, il est possible que Rex1 fournisse une spécificité fonctionnelle en fonction du développement cellulaire à c-Myc, étant donné qu'ils sembleraient tous deux être impliqués dans la même fonction à ce locus et que l'expression de c-Myc n'est pas réprimée en cours de différenciation (Figure 2a, p206). Ou alors, ces deux protéines pourraient influencer indépendamment sur l'élongation, leur fixation cumulative étant nécessaire pour une expression optimale de *Tsix*, Rex1 étant ensuite décrochée en cours de différenciation suite à son extinction transcriptionnelle tandis que la fixation de c-Myc serait abolie due à une absence de recrutement par d'autres protéines (comme YY1 ou Sp1) (Stojanova *et al.*, 2004, Gartel *et al.*, 2001).

Il est aussi très intéressant de noter que Rex1 semble différentiellement exprimée en fonction de la lignée cellulaire étudiée. En effet, un facteur 10 du nombre de transcrits *Rex1* est observable en comparant son expression au sein de cellules ES 129 et C57BL/6 (Sharova *et al.*, 2007), ce phénomène étant observable pour plusieurs centaines de gènes. Il serait néanmoins intéressant d'étudier le profil de fixation de Rex1 à *DxPas34* dans ces cellules C57BL/6 afin de voir si cette carence en transcrits (qui a un plus grand effet que les déplétions que nous avons généré par ARNi) influence la transcription de *Tsix* et donc, potentiellement, l'initiation de l'inactivation du chromosome X. Dans la même ligne d'idées, Rex1 ne semble

pas être exprimée au sein des EpiSCs (pour « *EPIblast Stem Cells* ») (Bao *et al.*, 2009), un stade cellulaire légèrement plus différencié que les ES, au sein desquelles l'inactivation aléatoire a eu lieu malgré le fait que les facteurs de pluripotence (Oct4/Nanog/Sox2) sont encore exprimés (Guo *et al.*, 2009). Il serait donc intéressant d'étudier l'élongation de la transcription de *Tsix* dans de telles cellules, un mécanisme de répression de *Tsix* pourrait en partie expliquer l'observation d'*X* inactifs dans ces lignées cellulaires.

Cependant, ni les mécanismes de la régulation transcriptionnelle de l'antisens, ni le rôle de la fixation des divers autres facteurs de transcription CTCF, YY1, Sp1, Klf4, Oct4 et Sox2 aux alentours de la région promotrice de *Tsix*, ne sont connus. Étant donné que la délétion de *DxPas34* empêche la fixation de l'ensemble de ces protéines (Oct4 et Sox2 non vérifiées) ainsi que le recrutement de l'ARN polymérase II et TFIIB (Sup. Figure 2, p210 ; Vigneau *et al.*, 2006), il est tentant d'émettre l'hypothèse qu'au moins l'une de ces protéines pourrait être responsable du recrutement du complexe de pré-initiation de transcription au promoteur majoritaire de *Tsix* dans les cellules ES.

Comme évoqué et démontré dans cet article, la fixation des facteurs Oct4 et Sox2 à *DxPas34* est controversée de par le fait que nous n'avons pas réussi à la reproduire et que leur fixation à cette position dans le papier original est déjà minime (Donohoe *et al.*, 2009), notamment par rapport à leur fixation dans l'intron 1 de *Xist* (Navarro *et al.*, 2008). De plus, la fixation des facteurs de CTCF et YY1 ne semble pas corrélée à l'expression de *Tsix* lorsque la différenciation est induite (Donohoe *et al.*, 2007, données personnelles). Par soustraction, parmi les facteurs de transcription dont la fixation autour du promoteur de *Tsix* est avérée, il ne nous reste que Klf4 et Sp1 comme protéines candidates au recrutement de la machinerie de transcription. De manière intéressante, la fixation de Klf4 entre *DxPas34* et le promoteur de *Tsix*, corrèle parfaitement avec l'expression de l'antisens dans tous les tissus étudiés (ES, TS si elle se fixe majoritairement sur le Xa, MEFs, mutants). Il est connu que Klf4 et Sp1 interagissent et possèdent des séquences de fixation très similaires (Zhang & Bieker, 1998), agissant même conjointement au promoteur du gène *LAMC1* chez le rat dans des rôles d'activateurs transcriptionnels (Higaki *et al.*, 2002).

Klf4 est connue pour aussi bien pouvoir jouer le rôle d'activateur que de répresseur transcriptionnel en fonction du contexte chromatinien grâce à deux domaines distincts (Yet *et al.*, 1998). Ce rôle d'activateur pourrait en partie être médié par son interaction avec les histone acétyl-transférases p300 et CBP (pour « *CREB Binding Protein* »), qui en acétylant, entre autres H3K9Ac (Tjeertes *et al.*, 2009, Hargreaves *et al.*, 2009), au promoteur du gène cible de Klf4, engendrent l'ouverture locale de la chromatine permettant ainsi à la machinerie

de transcription de se fixer (Evans *et al.*, 2007). Cette acétylation est d'ailleurs visible dans la Figure supplémentaire 3e, p211, et dont le profil corrèle aussi parfaitement avec celui de Sp1 et Klf4 (de même que celui de H3K4me3, Figure supplémentaire 3d, p211). D'autres mécanismes de régulation transcriptionnelle impliquant des partenaires de Klf4 sont envisageables, notamment à travers des interactions directes avec des sous-unités de la machinerie de transcription (résumé dans la Figure 66).

Nous pourrions imaginer que la fixation et les interactions de YY1, Sp1 et Klf4 autour de *DxPas34*, engendreraient la formation d'une structure euchromatinienne (grâce à Klf4 et peut-être p300 et CBP) qui permettrait à son tour le recrutement de la machinerie de transcription : d'abord TFIID par Sp1, suivi de TFIIB par YY1 et éventuellement par p300 et/ou CBP, engendrant ainsi une activité transcriptionnelle importante une fois l'ARN polymérase II fixée et activée. Dans l'avenir, il sera donc particulièrement intéressant de chercher une localisation de p300 et CBP autour de *DxPas34* et de *Xite* (Klf4 s'y fixant aussi plus faiblement, ce qui pourrait aussi influencer sur l'expression de *Tsix*), ainsi que de réaliser des déplétions de Klf4 et de Sp1 afin de voir si l'on affecte le recrutement des autres protéines à *DxPas34* ou celle du PIC au promoteur de *Tsix*.

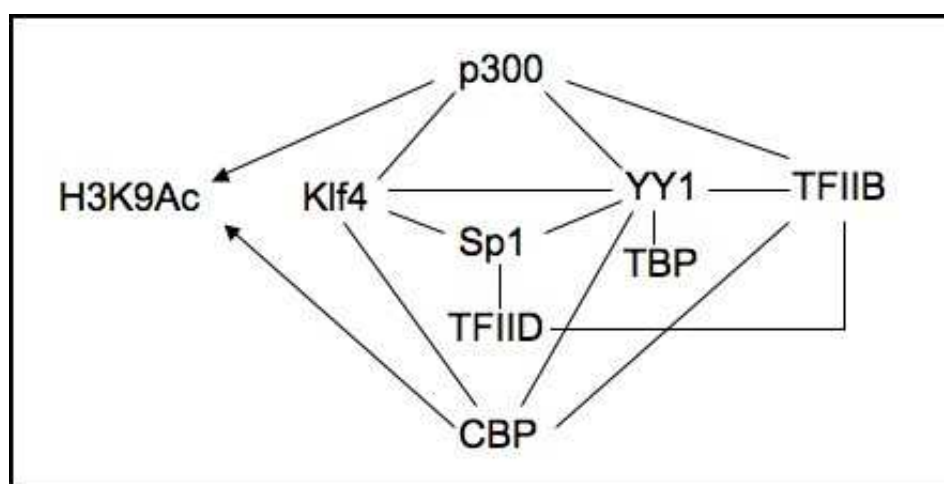


Figure 66 : Réseau d'interactions entre différents facteurs de transcription étant, ou pouvant être, impliqués dans la régulation de l'expression de *Tsix* au sein des cellules ES indifférenciées. Ce réseau schématise les différentes possibilités de recrutement de TFIIB au promoteur de *Tsix* afin d'initier la transcription. En règle générale, cette protéine est recrutée aux promoteurs par TFIID mais il arrive qu'elle soit directement recrutée par une autre protéine faisant office d'initiatrice de la transcription. Références des interactions : YY1, p300, CBP et TFIIB (Kwok *et al.*, 1994 ; Lee *et al.*, 1996 ; Usheva *et al.*, 1994). YY1 et TBP (pour « *TATA-Binding Protein* ») (Lee *et al.*, 1992). TFIIB, TFIID et Sp1 (Liu *et al.*, 2009 ; Smale *et al.*, 1990). Sp1, YY1 et Klf4 (Lee *et al.*, 1993, Zhang & Bieker, 1998). CBP, p300 et H3K9Ac (Tjeertes *et al.*, 2009, Hargreaves *et al.*, 2009).

L'expression de *Tsix* semble donc bien être liée de manière étroite à la pluripotence cellulaire, corrélant ainsi avec l'hypothèse que son expression transcriptionnelle, détectable au sein de cellules dont le chromosome X inactif vient d'être réactivé (ES, iPS, cellules de l'épiblaste) engendre des modifications chromatinienne le long du locus *Xist/Tsix* dans le but d'établir une probabilité égale des deux chromosomes X à être inactivé lors d'une éventuelle différenciation cellulaire, tout en réprimant post-transcriptionnellement *Xist*. Cependant, il existe une lignée cellulaire pluripotente où l'on n'observe pas d'expression de l'antisens : les cellules germinales primordiales (Sugimoto *et al.*, 2007). De manière fascinante, aussi bien Rex1 (aussi appelée Zfp42) que Klf4 sont très faiblement exprimées, voir réprimées, dans les PGCs par rapport aux cellules ES et EGs où l'on observe une expression de *Tsix* (Mise *et al.*, 2008). Cette constatation introduit un lien supplémentaire entre ces facteurs de transcription et l'expression de l'antisens. On pourrait imaginer que la cellule germinale primordiale peut se passer de la transcription de *Tsix* par rapport à l'expression de *Xist*, les facteurs de pluripotence Oct4, Nanog et Sox2 étant réactivés au sein de ces cellules, le réprimant donc transcriptionnellement. Concernant le rôle de *Tsix* dans la structure chromatinienne de *Xist/Tsix* impliquée dans le choix du chromosome à inactiver, il est possible que l'effacement des marques restantes de l'ancien Xi ne soit pas nécessaire dans les PGCs. Ce phénomène pourrait être effectué lors de la reprogrammation du génome suite à la déméthylation globale de l'ADN ainsi qu'à la suppression de l'empreinte parentale (Monk *et al.*, 1987, Hajkova *et al.*, 2002, Lee *et al.*, 2002), qui sera suivie par l'établissement de nouvelles marques sexe-spécifiques impliquées dans l'inactivation empreintée du chromosome X au cours de la maturation des gamètes.

En l'espace de deux ans Navarro *et al.* ont donc établi une corrélation directe entre l'état cellulaire de pluripotence et la régulation des deux acteurs principaux de l'inactivation du chromosome X : *Xist* et *Tsix*.

En étudiant les banques de données de ChIP-CHIP ou de ChIP-seq, on peut s'apercevoir que les facteurs de pluripotence mis en cause dans la régulation de l'inactivation semblent aussi être liés à la régulation d'autres gènes potentiellement liés à ce phénomène, plusieurs d'entre eux (Klf4, c-Myc, Nanog, Oct4, Rex1, Sox2) se trouvant liés autour du promoteur de Rnf12 (Kim *et al.*, 2008 ; Chen *et al.*, 2008), sûrement dans le but d'empêcher sa surexpression avant l'initiation de la différenciation. La découverte de liens entre pluripotence et inactivation ne semble donc, *a priori*, pas terminée. Cependant ce phénomène de coopération de facteurs dans la régulation génique ne semble pas être restreint aux gènes impliqués dans

l'inactivation, ce type de sites cumulatifs autour de promoteurs étant visibles dans l'ensemble du génome, ceci permettant sans doute les multiples différences transcriptionnelles entre cellules pluripotentes et différenciées.

Chapitre IV : Épilogue

Au cours de ce manuscrit, nous avons donc pu nous apercevoir de la multitude de phénomènes pouvant rentrer en jeu dans la régulation transcriptionnelle d'un gène. En effet, que ce soit l'implication de divers facteurs de transcription, les modifications épigénétiques de la chromatine, les interactions tridimensionnelles entre loci ou encore la localisation nucléaire du gène étudié, chacun de ces facteurs joue un rôle prépondérant dans la manière dont celui-ci est régulé. L'exemple de la régulation du tandem de gènes *Xist/Tsix* est particulièrement intéressant à ce point de vue de par le fait qu'il allie l'ensemble de ces phénomènes.

Au locus *Xist/Tsix* on a donc pu observer l'implication primordiale des facteurs de pluripotence, dont la fixation (dans l'intron 1 de *Xist* et dans *DxPas34*) semble expliquer en grande partie la relation étroite entre pluripotence cellulaire et réactivation du X inactif : Oct4, Nanog et Sox2 réprimant *Xist*, tandis que Rex1, Klf4 et c-Myc activent *Tsix* au sein des cellules souches embryonnaires. Cependant dans les deux cas, des questions demeurent quant à l'implication exacte de chaque facteur dans la régulation du gène associé. Et dans l'avenir, il sera donc primordial d'étudier le statut de fixation et la répercussion de celle-ci sur l'expression de *Xist/Tsix* au sein de cellules permettant une extinction individuelle inductible (soit par mutagenèse, soit par interférence stable) de l'ensemble des facteurs de transcription impliqués (la délétion des sites de fixation de ces facteurs semblant peu recommandable, Oct4, Nanog et Sox2 se fixant tous dans une même région très restreinte, et *DxPas34* étant une région hautement répétée donc qui sera particulièrement difficile à manipuler par des approches de biologie moléculaire).

De plus, il serait intéressant de savoir quelle est la cinétique d'action (ou plutôt de fixation) des différents facteurs de pluripotence lors de l'initiation de l'inactivation. Car on pourrait supposer que les protéines régulant l'expression de *Tsix* devraient disparaître en premier (par rapport à celles régulant *Xist*) mais de manière momentanément monoallélique, afin de permettre une surexpression de *Xist* de la même façon. Mais la suite de ce projet devra surtout passer par l'identification des protéines responsables de l'activation transcriptionnelle de *Tsix*, celles-ci se fixant (directement ou pas) à *DxPas34*, le recrutement du PIC ne se faisant pas en son absence.

La fixation de CTCF à de nombreux sites le long du locus *Xist/Tsix* semble être potentiellement une grande source de diversité mécanistique concernant la régulation de ces deux gènes, ceci étant principalement dû aux nombreux rôles exercés par cette protéine à travers diverses interactions protéiques. Bien que nous ayons émis plusieurs hypothèses

concernant le rôle de certains sites de fixation de CTCF dans la région contrôlant la mise en place de l'inactivation nous sommes encore loin d'avoir découvert l'intégralité de la panoplie fonctionnelle de ces fixations dont pourraient faire partie :

- la régulation transcriptionnelle de *Xist* à P2 en association avec YY1
- l'association à la matrice nucléaire en amont de *Tsix*
- le rôle dans la structure tridimensionnelle du locus en amont de *Xist* et de *Tsix*, en association avec la cohésine.

La mutation des sites endogènes de fixation (aussi bien pour CTCF que pour YY1 et la cohésine) dans la région promotrice de *Xist* permettra dans un premier temps de valider certaines hypothèses (notamment le rôle de CTCF et YY1 dans la surexpression de *Xist*). Par la suite, la banalisation des études de la fixation de diverses protéines sur l'ensemble du génome grâce à la méthode de Chip-seq devrait permettre d'apporter des indices précieux quant au rôle exact de chaque pic CTCF dès lors où l'on sait quelle région l'on tient à analyser (ce qui, dans notre cas, pourra notamment servir dans l'élucidation du rôle de CTCF au sein de l'exon 1, 5 et 8 de *Xist*, sites pour lesquels nous n'avons aucune autre information). Cependant, une question fondamentale à la compréhension des différents mécanismes d'interactions engendrées par cette protéine à 11 doigts de zinc demeure de savoir ce qui différencie chaque site de fixation et chaque interaction par rapport aux autres. Il sera en effet d'un intérêt indéniable de découvrir les mécanismes, surement structuraux (variation des doigts de zinc utilisés) et post-traductionnels (poly-(ADP-ribosyl), la phosphorylation et sumoylation), gouvernant l'action précise de CTCF à chacun de ses sites de fixation. Ce travail impliquera sans doute la mutation des différents doigts de zinc ainsi que l'utilisation d'inhibiteurs des modifications post-traductionnels tout en corrélant ces modifications aux différentes interactions ou fixations de la protéine.

Dans l'avenir immédiat, l'obtention de clones recombinants permettant la délétion des sites de fixation des protéines CTCF, YY1 et Rad21 dans la région promotrice de *Xist*, va nous fournir un outil précieux qui devrait être suffisant pour vérifier certaines des hypothèses émises tout au long de ce manuscrit ainsi que de répondre aux questions que l'on se pose à leur sujet, notamment :

- YY1 et CTCF sont-elles les deux activateurs transcriptionnels majeurs de *Xist* ?
- Comment fonctionnent-elles ?
- Bien que YY1 semble véritablement être impliquée dans l'activation de *Xist* (Kim *et al.*, 2006, Figure 47), qu'en est-il de CTCF ? Est-elle fixée à ce site uniquement pour permettre la

fixation ou le recrutement de YY1 ou a-t-elle un rôle direct dans le recrutement du PIC en cours de différenciation ?

- La fixation de ces protéines au promoteur P2 de *Xist* dans les cellules ES indifférenciées joue-t-elle un rôle ?

- Quel est le rôle de la fixation de CTCF et de la cohésine en amont de *Xist* ?

Et en attendant l'obtention de ces lignées cellulaires mutantes, l'étude de la mutation des sites de fixation de ces protéines au sein de la région promotrice de *Xist* placée en amont du gène de la luciférase devrait nous permettre d'au moins répondre dans un premier temps à l'hypothèse que CTCF et YY1 sont bien les activateurs transcriptionnels majeurs de *Xist*.

Toutefois, comme mentionné précédemment, nous ignorons encore tout de la fonction des sites CTCF répartis le long de *Xist* : en 3' de l'exon 1, sur les petits exons 2, 3, 4, 5 ainsi qu'à l'exon 8 de *Xist*. Ces sites, dont la présence semble être reliée au développement, pourraient être liés à l'établissement d'une structure tridimensionnelle particulière au sein du locus *Xist/Tsix*, étants donc peut-être liés à un système de répression de *Xist* dans les cellules ES. Néanmoins, concernant cette hypothèse, on peut se demander quel serait le mécanisme de cette structure tridimensionnelle ? Et plus, globalement, CTCF a-t-elle la capacité d'établir de telles structures en l'absence d'une coopération avec la cohésine ? Pour répondre à cela, il faudra envisager des expériences de 3C ou de 4C, expériences que nous ne sommes pas encore en mesure de réaliser au laboratoire.

Selon toute vraisemblance, CTCF recrute sans doute d'autres protéines à ces positions afin d'y jouer un rôle précis, sa fixation seule devenant une véritable rareté dans la bibliographie actuelle. Il est aussi possible qu'au vu de la localisation de ces sites, CTCF fonctionne dans la régulation de l'épissage le long du gène *Xist*, d'une manière cependant différente d'ASF/SF2 qui semble jouer son rôle par l'intermédiaire de sa fixation au *Repeat-A* (Royce-Tolland *et al.*, 2010).

L'étude des cellules germinales embryonnaires a sûrement aussi encore beaucoup de choses à nous apprendre, notamment en ce qui concerne les mécanismes permettant la réactivation de l'X inactif. Le fait que l'on observe la conservation de la fixation de toutes les protéines que nous avons étudié, de manière identique à ce qui a été vu au sein des cellules ES, nous apporte une preuve supplémentaire de la ressemblance des deux lignées. Cependant, des données récentes ont montré que les EGs maintenaient certaines similitudes avec les PGCs (à partir desquelles elles sont dérivées), particulièrement concernant la méthylation de l'ADN (dont

l'état varie en fonction du stade auquel les cellules ont été dérivées), et éventuellement dans l'expression de certains groupes de microARNs (Tada *et al.*, 1998 ; Durcova-Hills *et al.*, 2001 ; Petkov *et al.*, 2009 ; Ciaudo *et al.*, 2010). Il sera donc particulièrement intéressant d'étudier les mécanismes de méthylation de l'ADN au niveau du promoteur de *Xist* dans ces cellules (notamment dans des cellules dérivées à différents stades : 7.5, 8.5 et 9.5 jpc), dans le dessein d'y observer des variations développementalement liées. Et il pourrait aussi être intéressant de différencier des cellules EGs afin de vérifier si elles ont bien conservé leur capacité à inactiver un X de manière aléatoire.

Au final, nous pouvons donc constater que la régulation de l'expression de *Xist* (mais aussi de son antisens *Tsix*) est un phénomène extrêmement complexe qui semble dépendre de multiples phénomènes plus ou moins indépendants, influants sur plusieurs paramètres dans le seul but d'obtenir une expression importante et monoallélique de *Xist* en cours de différenciation de cellules femelles. Bien qu'au cours de mon doctorat, je n'ai pu qu'effleurer le potentiel de ce sujet, l'abondance de directions dans lesquelles on pourrait se lancer afin d'approfondir nos connaissances sur le sujet présage encore de nombreuses découvertes passionnantes quant à la mise en place de l'inactivation du chromosome X à travers la surexpression de son acteur principal : *Xist*.

À la conclusion de ces quatre années de thèse, et je pense que c'est le cas pour bon nombre de doctorants, je ne peux avoir qu'un sentiment mitigé vis-à-vis du travail effectué. L'impression dominante d'abandonner un projet extrêmement prometteur et loin d'être fini pour devoir en commencer un autre est assez troublante et laisse une désagréable sensation d'inachevé. Cependant, je ne puis m'empêcher de ressentir également une certaine fierté lorsque je regarde le travail accompli, aussi bien au niveau des données produites et des perspectives prometteuses développées qu'au niveau personnel où, au cours de la rédaction du manuscrit, l'on se rend véritablement compte du progrès en termes de connaissances mais aussi de réflexion qu'on a acquis au cours de ces années de doctorat. De plus, l'idée d'entamer un nouveau sujet, dans un autre laboratoire, sûrement à l'étranger, est aussi très excitante, même si au fond de moi, je nourris l'espoir de pouvoir retravailler sur l'inactivation du chromosome X un jour.

Bibliographie :

- Agelopoulos, M. and D. Thanos (2006). "Epigenetic determination of a cell-specific gene expression program by ATF-2 and the histone variant macroH2A." EMBO J 25(20): 4843-4853.
- Agrelo, R., A. Souabni, M. Novatchkova, C. Haslinger, M. Leeb, V. Komnenovic, H. Kishimoto, L. Gresh, T. Kohwi-Shigematsu, L. Kenner and A. Wutz (2009). "SATB1 defines the developmental context for gene silencing by Xist in lymphoma and embryonic cells." Dev Cell 16(4): 507-516.
- Akbadian, S., R. Z. Chen, J. Gribnau, T. P. Rasmussen, H. Fong, R. Jaenisch and E. G. Jones (2001). "Expression pattern of the Rett syndrome gene MeCP2 in primate prefrontal cortex." Neurobiol Dis 8(5): 784-791.
- Allen, E., S. Horvath, F. Tong, P. Kraft, E. Spiteri, A. D. Riggs and Y. Marahrens (2003). "High concentrations of long interspersed nuclear element sequence distinguish monoallelically expressed genes." Proc Natl Acad Sci U S A 100(17): 9940-9945.
- Alvarez, J. D., D. H. Yasui, H. Niida, T. Joh, D. Y. Loh and T. Kohwi-Shigematsu (2000). "The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development." Genes Dev 14(5): 521-535.
- Angelov, D., A. Molla, P. Y. Perche, F. Hans, J. Cote, S. Khochbin, P. Bouvet and S. Dimitrov (2003). "The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling." Mol Cell 11(4): 1033-1041.
- Arlt, M. F., A. M. Casper and T. W. Glover (2003). "Common fragile sites." Cytogenet Genome Res 100(1-4): 92-100.
- Attia, M., C. Rachez, A. De Pauw, P. Avner and U. C. Rogner (2007). "Nap112 promotes histone acetylation activity during neuronal differentiation." Mol Cell Biol 27(17): 6093-6102.
- Augui, S., G. J. Fillion, S. Huart, E. Nora, M. Guggiari, M. Maresca, A. F. Stewart and E. Heard (2007). "Sensing X chromosome pairs before X inactivation via a novel X-pairing region of the Xic." Science 318(5856): 1632-1636.
- Austen, M., B. Luscher and J. M. Luscher-Firzlaff (1997). "Characterization of the transcriptional regulator YY1. The bipartite transactivation domain is independent of interaction with the TATA box-binding protein, transcription factor IIB, TAFII55, or cAMP-responsive element-binding protein (CPB)-binding protein." J Biol Chem 272(3): 1709-1717.
- Ayoub, N., C. Richler and J. Wahrman (1997). "Xist RNA is associated with the transcriptionally inactive XY body in mammalian male meiosis." Chromosoma 106(1): 1-10.
- Azuara, V., P. Perry, S. Sauer, M. Spivakov, H. F. Jorgensen, R. M. John, M. Gouti, M. Casanova, G. Warnes, M. Merkenschlager and A. G. Fisher (2006). "Chromatin signatures of pluripotent cell lines." Nat Cell Biol 8(5): 532-538.
- Babu, M. M., N. M. Luscombe, L. Aravind, M. Gerstein and S. A. Teichmann (2004). "Structure and evolution of transcriptional regulatory networks." Curr Opin Struct Biol 14(3): 283-291.
- Bach, I., C. Rodriguez-Esteban, C. Carriere, A. Bhushan, A. Krones, D. W. Rose, C. K. Glass, B. Andersen, J. C. Izpisua Belmonte and M. G. Rosenfeld (1999). "RLIM inhibits functional activity of LIM homeodomain transcription factors via recruitment of the histone deacetylase complex." Nat Genet 22(4): 394-399.

- Bacher, C. P., M. Guggiari, B. Brors, S. Augui, P. Clerc, P. Avner, R. Eils and E. Heard (2006). "Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation." *Nat Cell Biol* 8(3): 293-299.
- Bailey, J. A., L. Carrel, A. Chakravarti and E. E. Eichler (2000). "Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis." *Proc Natl Acad Sci U S A* 97(12): 6634-6639.
- Banan, M., I. C. Rojas, W. H. Lee, H. L. King, J. V. Harriss, R. Kobayashi, C. F. Webb and P. D. Gottlieb (1997). "Interaction of the nuclear matrix-associated region (MAR)-binding proteins, SATB1 and CDP/Cux, with a MAR element (L2a) in an upstream regulatory region of the mouse CD8a gene." *J Biol Chem* 272(29): 18440-18452.
- Bao, S., F. Tang, X. Li, K. Hayashi, A. Gillich, K. Lao and M. A. Surani (2009). "Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells." *Nature* 461(7268): 1292-1295.
- Barr, H., A. Hermann, J. Berger, H. H. Tsai, K. Adie, A. Prokhortchouk, B. Hendrich and A. Bird (2007). "Mbd2 contributes to DNA methylation-directed repression of the Xist gene." *Mol Cell Biol* 27(10): 3750-3757.
- Barr, M. L. and E. G. Bertram (1949). "A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis." *Nature* 163(4148): 676.
- Barski, A., S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev and K. Zhao (2007). "High-resolution profiling of histone methylations in the human genome." *Cell* 129(4): 823-837.
- Bartlett, M. H., C. N. Adra, J. Park, V. M. Chapman and M. W. McBurney (1991). "DNA methylation of two X chromosome genes in female somatic and embryonal carcinoma cells." *Somat Cell Mol Genet* 17(1): 35-47.
- Barton, D. E., F. N. David and M. Merrington (1964). "The Positions of the Sex Chromosomes in the Human Cell in Mitosis." *Ann Hum Genet* 28: 123-128.
- Baumann, C. and R. De La Fuente (2009). "ATR-X marks the inactive X chromosome (Xi) in somatic cells and during imprinted X chromosome inactivation in trophoblast stem cells." *Chromosoma* 118(2): 209-222.
- Beard, C., E. Li and R. Jaenisch (1995). "Loss of methylation activates Xist in somatic but not in embryonic cells." *Genes Dev* 9(19): 2325-2334.
- Belak, Z. R. and N. Ovsenek (2007). "Assembly of the Yin Yang 1 transcription factor into messenger ribonucleoprotein particles requires direct RNA binding activity." *J Biol Chem* 282(52): 37913-37920.
- Bell, A. C. and G. Felsenfeld (2000). "Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene." *Nature* 405(6785): 482-485.
- Belmont, A. S., F. Bignone and P. O. Ts'o (1986). "The relative intranuclear positions of Barr bodies in XXX non-transformed human fibroblasts." *Exp Cell Res* 165(1): 165-179.
- Bergstrom, R., J. Whitehead, S. Kurukuti and R. Ohlsson (2007). "CTCF regulates asynchronous replication of the imprinted H19/Igf2 domain." *Cell Cycle* 6(4): 450-454.
- Bernstein, B. E., E. L. Humphrey, R. L. Erlich, R. Schneider, P. Bouman, J. S. Liu, T. Kouzarides and S. L. Schreiber (2002). "Methylation of histone H3 Lys 4 in coding regions of active genes." *Proc Natl Acad Sci U S A* 99(13): 8695-8700.
- Bernstein, B. E., T. S. Mikkelsen, X. Xie, M. Kamal, D. J. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath, R. Jaenisch, A. Wagschal, R. Feil, S. L. Schreiber and E. S. Lander (2006). "A bivalent chromatin structure marks key developmental genes in embryonic stem cells." *Cell* 125(2): 315-326.
- Blewitt, M. E., A. V. Gendrel, Z. Pang, D. B. Sparrow, N. Whitelaw, J. M. Craig, A.

- Apedaile, D. J. Hilton, S. L. Dunwoodie, N. Brockdorff, G. F. Kay and E. Whitelaw (2008). "SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation." Nat Genet 40(5): 663-669.
- Boggs, B. A., P. Cheung, E. Heard, D. L. Spector, A. C. Chinault and C. D. Allis (2002). "Differentially methylated forms of histone H3 show unique association patterns with inactive human X chromosomes." Nat Genet 30(1): 73-76.
- Boggs, B. A. and A. C. Chinault (1994). "Analysis of replication timing properties of human X-chromosomal loci by fluorescence in situ hybridization." Proc Natl Acad Sci U S A 91(13): 6083-6087.
- Boiani, M. and H. R. Scholer (2005). "Regulatory networks in embryo-derived pluripotent stem cells." Nat Rev Mol Cell Biol 6(11): 872-884.
- Borsani, G., R. Tonlorenzi, M. C. Simmler, L. Dandolo, D. Arnaud, V. Capra, M. Grompe, A. Pizzuti, D. Muzny, C. Lawrence and *et al.* (1991). "Characterization of a murine gene expressed from the inactive X chromosome." Nature 351(6324): 325-329.
- Bouhassira, E. E., K. Westerman and P. Leboulch (1997). "Transcriptional behavior of LCR enhancer elements integrated at the same chromosomal locus by recombinase-mediated cassette exchange." Blood 90(9): 3332-3344.
- Boumil, R. M., Y. Ogawa, B. K. Sun, K. D. Huynh and J. T. Lee (2006). "Differential methylation of Xite and CTCF sites in Tsix mirrors the pattern of X-inactivation choice in mice." Mol Cell Biol 26(6): 2109-2117.
- Bourc'his, D. and T. H. Bestor (2004). "Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L." Nature 431(7004): 96-99.
- Bourc'his, D., G. L. Xu, C. S. Lin, B. Bollman and T. H. Bestor (2001). "Dnmt3L and the establishment of maternal genomic imprints." Science 294(5551): 2536-2539.
- Bourgeois, C. A., F. Laquerriere, D. Hemon, J. Hubert and M. Bouteille (1985). "New data on the in-situ position of the inactive X chromosome in the interphase nucleus of human fibroblasts." Hum Genet 69(2): 122-129.
- Boyer, L. A., T. I. Lee, M. F. Cole, S. E. Johnstone, S. S. Levine, J. P. Zucker, M. G. Guenther, R. M. Kumar, H. L. Murray, R. G. Jenner, D. K. Gifford, D. A. Melton, R. Jaenisch and R. A. Young (2005). "Core transcriptional regulatory circuitry in human embryonic stem cells." Cell 122(6): 947-956.
- Boyle, A. L., S. G. Ballard and D. C. Ward (1990). "Differential distribution of long and short interspersed element sequences in the mouse genome: chromosome karyotyping by fluorescence in situ hybridization." Proc Natl Acad Sci U S A 87(19): 7757-7761.
- Brockdorff, N., A. Ashworth, G. F. Kay, P. Cooper, S. Smith, V. M. McCabe, D. P. Norris, G. D. Penny, D. Patel and S. Rastan (1991). "Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome." Nature 351(6324): 329-331.
- Brockdorff, N., A. Ashworth, G. F. Kay, V. M. McCabe, D. P. Norris, P. J. Cooper, S. Swift and S. Rastan (1992). "The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus." Cell 71(3): 515-526.
- Brown, C. J., A. Ballabio, J. L. Rupert, R. G. Lafreniere, M. Grompe, R. Tonlorenzi and H. F. Willard (1991). "A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome." Nature 349(6304): 38-44.
- Brown, C. J., B. D. Hendrich, J. L. Rupert, R. G. Lafreniere, Y. Xing, J. Lawrence and H. F. Willard (1992). "The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus." Cell 71(3): 527-542.
- Brown, C. J. and H. F. Willard (1994). "The human X-inactivation centre is not required for

- maintenance of X-chromosome inactivation." Nature 368(6467): 154-156.
- Buhler, M., A. Verdel and D. Moazed (2006). "Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing." Cell 125(5): 873-886.
- Burcin, M., R. Arnold, M. Lutz, B. Kaiser, D. Runge, F. Lottspeich, G. N. Filippova, V. V. Lobanenko and R. Renkawitz (1997). "Negative protein 1, which is required for function of the chicken lysozyme gene silencer in conjunction with hormone receptors, is identical to the multivalent zinc finger repressor CTCF." Mol Cell Biol 17(3): 1281-1288.
- Burke, L. J., T. Hollemann, T. Pieler and R. Renkawitz (2002). "Molecular cloning and expression of the chromatin insulator protein CTCF in *Xenopus laevis*." Mech Dev 113(1): 95-98.
- Burke, L. J., R. Zhang, M. Bartkuhn, V. K. Tiwari, G. Tavoosidana, S. Kurukuti, C. Weth, J. Leers, N. Galjart, R. Ohlsson and R. Renkawitz (2005). "CTCF binding and higher order chromatin structure of the H19 locus are maintained in mitotic chromatin." EMBO J 24(18): 3291-3300.
- Bushey, A. M., E. Ramos and V. G. Corces (2009). "Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions." Genes Dev 23(11): 1338-1350.
- Butcher, D. T. and D. I. Rodenhiser (2007). "Epigenetic inactivation of BRCA1 is associated with aberrant expression of CTCF and DNA methyltransferase (DNMT3B) in some sporadic breast tumours." Eur J Cancer 43(1): 210-219.
- Buzin, C. H., J. R. Mann and J. Singer-Sam (1994). "Quantitative RT-PCR assays show Xist RNA levels are low in mouse female adult tissue, embryos and embryoid bodies." Development 120(12): 3529-3536.
- Cai, S., C. C. Lee and T. Kohwi-Shigematsu (2006). "SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes." Nat Genet 38(11): 1278-1288.
- Cai, Y., J. Jin, T. Yao, A. J. Gottschalk, S. K. Swanson, S. Wu, Y. Shi, M. P. Washburn, L. Florens, R. C. Conaway and J. W. Conaway (2007). "YY1 functions with INO80 to activate transcription." Nat Struct Mol Biol 14(9): 872-874.
- Caiafa, P. and J. Zlatanova (2009). "CCCTC-binding factor meets poly(ADP-ribose) polymerase-1." J Cell Physiol 219(2): 265-270.
- Capelson, M. and V. G. Corces (2004). "Boundary elements and nuclear organization." Biol Cell 96(8): 617-629.
- Cardoso, C., S. Timsit, L. Villard, M. Khrestchatisky, M. Fontes and L. Colleaux (1998). "Specific interaction between the XNP/ATR-X gene product and the SET domain of the human EZH2 protein." Hum Mol Genet 7(4): 679-684.
- Carrel, L. and H. F. Willard (2005). "X-inactivation profile reveals extensive variability in X-linked gene expression in females." Nature 434(7031): 400-404.
- Carroll, J. S., C. A. Meyer, J. Song, W. Li, T. R. Geistlinger, J. Eeckhoute, A. S. Brodsky, E. K. Keeton, K. C. Fertuck, G. F. Hall, Q. Wang, S. Bekiranov, V. Sementchenko, E. A. Fox, P. A. Silver, T. R. Gingeras, X. S. Liu and M. Brown (2006). "Genome-wide analysis of estrogen receptor binding sites." Nat Genet 38(11): 1289-1297.
- Casper, A. M., P. Nghiem, M. F. Arlt and T. W. Glover (2002). "ATR regulates fragile site stability." Cell 111(6): 779-789.
- Cattanach, B. M., C. Rasberry, E. P. Evans, L. Dandolo, M. C. Simmler and P. Avner (1991). "Genetic and molecular evidence of an X-chromosome deletion spanning the tabby (Ta) and testicular feminization (Tfm) loci in the mouse." Cytogenet Cell Genet 56(3-4): 137-143.
- Chadwick, B. P. and H. F. Willard (2001). "Histone H2A variants and the inactive X

- chromosome: identification of a second macroH2A variant." Hum Mol Genet 10(10): 1101-1113.
- Chadwick, B. P. and H. F. Willard (2004). "Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome." Proc Natl Acad Sci U S A 101(50): 17450-17455.
- Chadwick, L. H., L. M. Pertz, K. W. Broman, M. S. Bartolomei and H. F. Willard (2006). "Genetic control of X chromosome inactivation in mice: definition of the Xce candidate interval." Genetics 173(4): 2103-2110.
- Chambers, I., D. Colby, M. Robertson, J. Nichols, S. Lee, S. Tweedie and A. Smith (2003). "Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells." Cell 113(5): 643-655.
- Chambers, I., J. Silva, D. Colby, J. Nichols, B. Nijmeijer, M. Robertson, J. Vrana, K. Jones, L. Grotewold and A. Smith (2007). "Nanog safeguards pluripotency and mediates germline development." Nature 450(7173): 1230-1234.
- Chao, W., K. D. Huynh, R. J. Spencer, L. S. Davidow and J. T. Lee (2002). "CTCF, a candidate trans-acting factor for X-inactivation choice." Science 295(5553): 345-347.
- Chaumeil, J., P. Le Baccon, A. Wutz and E. Heard (2006). "A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced." Genes Dev 20(16): 2223-2237.
- Chaumeil, J., I. Okamoto, M. Guggiari and E. Heard (2002). "Integrated kinetics of X chromosome inactivation in differentiating embryonic stem cells." Cytogenet Genome Res 99(1-4): 75-84.
- Chen, T., Y. Ueda, J. E. Dodge, Z. Wang and E. Li (2003). "Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b." Mol Cell Biol 23(16): 5594-5605.
- Chen, X., H. Xu, P. Yuan, F. Fang, M. Huss, V. B. Vega, E. Wong, Y. L. Orlov, W. Zhang, J. Jiang, Y. H. Loh, H. C. Yeo, Z. X. Yeo, V. Narang, K. R. Govindarajan, B. Leong, A. Shahab, Y. Ruan, G. Bourque, W. K. Sung, N. D. Clarke, C. L. Wei and H. H. Ng (2008). "Integration of external signaling pathways with the core transcriptional network in embryonic stem cells." Cell 133(6): 1106-1117.
- Chernukhin, I., S. Shamsuddin, S. Y. Kang, R. Bergstrom, Y. W. Kwon, W. Yu, J. Whitehead, R. Mukhopadhyay, F. Docquier, D. Farrar, I. Morrison, M. Vigneron, S. Y. Wu, C. M. Chiang, D. Loukinov, V. Lobanenko, R. Ohlsson and E. Klenova (2007). "CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide." Mol Cell Biol 27(5): 1631-1648.
- Chernukhin, I. V., S. Shamsuddin, A. F. Robinson, A. F. Carne, A. Paul, A. I. El-Kady, V. V. Lobanenko and E. M. Klenova (2000). "Physical and functional interaction between two pluripotent proteins, the Y-box DNA/RNA-binding factor, YB-1, and the multivalent zinc finger factor, CTCF." J Biol Chem 275(38): 29915-29921.
- Cho, D. H., C. P. Thienes, S. E. Mahoney, E. Analau, G. N. Filippova and S. J. Tapscott (2005). "Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF." Mol Cell 20(3): 483-489.
- Chow, J. and E. Heard (2009). "X inactivation and the complexities of silencing a sex chromosome." Curr Opin Cell Biol 21(3): 359-366.
- Chow, J. C., C. Ciaudo, M. J. Fazzari, N. Mise, N. Servant, J. L. Glass, M. Attreed, P. Avner, A. Wutz, E. Barillot, J. M. Greally, O. Voinnet and E. Heard (2010). "LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation." Cell 141(6): 956-969.
- Chuang, P. T., D. G. Albertson and B. J. Meyer (1994). "DPY-27: a chromosome condensation protein homolog that regulates *C. elegans* dosage compensation through

- association with the X chromosome." *Cell* 79(3): 459-474.
- Chureau, C., M. Prissette, A. Bourdet, V. Barbe, L. Cattolico, L. Jones, A. Eggen, P. Avner and L. Duret (2002). "Comparative sequence analysis of the X-inactivation center region in mouse, human, and bovine." *Genome Res* 12(6): 894-908.
- Chuva de Sousa Lopes, S. M., K. Hayashi, T. C. Shovlin, W. Mifsud, M. A. Surani and A. McLaren (2008). "X chromosome activity in mouse XX primordial germ cells." *PLoS Genet* 4(2): e30.
- Ciaudo, C., N. Servant, V. Cognat, A. Sarazin, E. Kieffer, S. Viville, V. Colot, E. Barillot, E. Heard and O. Voinnet (2009). "Highly dynamic and sex-specific expression of microRNAs during early ES cell differentiation." *PLoS Genet* 5(8): e1000620.
- Ciavatta, D., S. Kalantry, T. Magnuson and O. Smithies (2006). "A DNA insulator prevents repression of a targeted X-linked transgene but not its random or imprinted X inactivation." *Proc Natl Acad Sci U S A* 103(26): 9958-9963.
- Clemson, C. M., L. L. Hall, M. Byron, J. McNeil and J. B. Lawrence (2006). "The X chromosome is organized into a gene-rich outer rim and an internal core containing silenced nongenic sequences." *Proc Natl Acad Sci U S A* 103(20): 7688-7693.
- Clemson, C. M., J. A. McNeil, H. F. Willard and J. B. Lawrence (1996). "XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure." *J Cell Biol* 132(3): 259-275.
- Clerc, P. and P. Avner (1998). "Role of the region 3' to Xist exon 6 in the counting process of X-chromosome inactivation." *Nat Genet* 19(3): 249-253.
- Cohen, D. E., L. S. Davidow, J. A. Erwin, N. Xu, D. Warshawsky and J. T. Lee (2007). "The DXPas34 repeat regulates random and imprinted X inactivation." *Dev Cell* 12(1): 57-71.
- Cohen, H. R. and B. Panning (2007). "XIST RNA exhibits nuclear retention and exhibits reduced association with the export factor TAP/NXF1." *Chromosoma* 116(4): 373-383.
- Collas, P. (2010). "The current state of chromatin immunoprecipitation." *Mol Biotechnol* 45(1): 87-100.
- Collins, N., R. A. Poot, I. Kukimoto, C. Garcia-Jimenez, G. Dellaire and P. D. Varga-Weisz (2002). "An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin." *Nat Genet* 32(4): 627-632.
- Cooper, S. J. and R. M. Hope (1993). "Evolution and expression of a beta-like globin gene of the Australian marsupial *Sminthopsis crassicaudata*." *Proc Natl Acad Sci U S A* 90(24): 11777-11781.
- Costanzi, C. and J. R. Pehrson (1998). "Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals." *Nature* 393(6685): 599-601.
- Costanzi, C. and J. R. Pehrson (2001). "MACROH2A2, a new member of the MARCOH2A core histone family." *J Biol Chem* 276(24): 21776-21784.
- Courtier, B., E. Heard and P. Avner (1995). "Xce haplotypes show modified methylation in a region of the active X chromosome lying 3' to Xist." *Proc Natl Acad Sci U S A* 92(8): 3531-3535.
- Csankovszki, G., A. Nagy and R. Jaenisch (2001). "Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation." *J Cell Biol* 153(4): 773-784.
- Csankovszki, G., B. Panning, B. Bates, J. R. Pehrson and R. Jaenisch (1999). "Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation." *Nat Genet* 22(4): 323-324.
- Cuddapah, S., R. Jothi, D. E. Schones, T. Y. Roh, K. Cui and K. Zhao (2009). "Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals

- demarcation of active and repressive domains." *Genome Res* 19(1): 24-32.
- Cunningham, J. M., M. E. Purucker, S. M. Jane, B. Safer, E. F. Vanin, P. A. Ney, C. H. Lowrey and A. W. Nienhuis (1994). "The regulatory element 3' to the A gamma-globin gene binds to the nuclear matrix and interacts with special A-T-rich binding protein 1 (SATB1), an SAR/MAR-associating region DNA binding protein." *Blood* 84(4): 1298-1308.
- Dahmus, M. E. (1996). "Reversible phosphorylation of the C-terminal domain of RNA polymerase II." *J Biol Chem* 271(32): 19009-19012.
- De La Rosa-Velazquez, I. A., H. Rincon-Arango, L. Benitez-Bribiesca and F. Recillas-Targa (2007). "Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF." *Cancer Res* 67(6): 2577-2585.
- de Napoles, M., J. E. Mermoud, R. Wakao, Y. A. Tang, M. Endoh, R. Appanah, T. B. Nesterova, J. Silva, A. P. Otte, M. Vidal, H. Koseki and N. Brockdorff (2004). "Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation." *Dev Cell* 7(5): 663-676.
- de Napoles, M., T. Nesterova and N. Brockdorff (2007). "Early loss of Xist RNA expression and inactive X chromosome associated chromatin modification in developing primordial germ cells." *PLoS One* 2(9): e860.
- Debrand, E., C. Chureau, D. Arnaud, P. Avner and E. Heard (1999). "Functional analysis of the DXPas34 locus, a 3' regulator of Xist expression." *Mol Cell Biol* 19(12): 8513-8525.
- Debrand, E., E. Heard and P. Avner (1998). "Cloning and localization of the murine Xpet gene: evidence for complex rearrangements during the evolution of the region around the Xist gene." *Genomics* 48(3): 296-303.
- Defossez, P. A., K. F. Kelly, G. J. Filion, R. Perez-Torrado, F. Magdinier, H. Menoni, C. L. Nordgaard, J. M. Daniel and E. Gilson (2005). "The human enhancer blocker CTC-binding factor interacts with the transcription factor Kaiso." *J Biol Chem* 280(52): 43017-43023.
- Dekker, J., K. Rippe, M. Dekker and N. Kleckner (2002). "Capturing chromosome conformation." *Science* 295(5558): 1306-1311.
- Delgado, M. D., I. V. Chernukhin, A. Bigas, E. M. Klenova and J. Leon (1999). "Differential expression and phosphorylation of CTCF, a c-myc transcriptional regulator, during differentiation of human myeloid cells." *FEBS Lett* 444(1): 5-10.
- Diaz-Perez, S., Y. Ouyang, V. Perez, R. Cisneros, M. Regelson and Y. Marahrens (2005). "The element(s) at the nontranscribed Xist locus of the active X chromosome controls chromosomal replication timing in the mouse." *Genetics* 171(2): 663-672.
- Diaz-Perez, S. V., D. O. Ferguson, C. Wang, G. Csankovszki, S. C. Tsai, D. Dutta, V. Perez, S. Kim, C. D. Eller, J. Salstrom, Y. Ouyang, M. A. Teitell, B. Kaltenboeck, A. Chess, S. Huang and Y. Marahrens (2006). "A deletion at the mouse Xist gene exposes trans-effects that alter the heterochromatin of the inactive X chromosome and the replication time and DNA stability of both X chromosomes." *Genetics* 174(3): 1115-1133.
- Dickinson, L. A., T. Joh, Y. Kohwi and T. Kohwi-Shigematsu (1992). "A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition." *Cell* 70(4): 631-645.
- Disteche, C. M., G. N. Filippova and K. D. Tsuchiya (2002). "Escape from X inactivation." *Cytogenet Genome Res* 99(1-4): 36-43.
- Do, J. T., D. W. Han, L. Gentile, I. Sobek-Klocke, M. Stehling and H. R. Scholer (2008). "Enhanced reprogramming of Xist by induced upregulation of Tsix and Dnmt3a." *Stem Cells* 26(11): 2821-2831.
- Donohoe, M. E., S. S. Silva, S. F. Pinter, N. Xu and J. T. Lee (2009). "The pluripotency factor

- Oct4 interacts with Ctfc and also controls X-chromosome pairing and counting." *Nature* 460(7251): 128-132.
- Donohoe, M. E., L. F. Zhang, N. Xu, Y. Shi and J. T. Lee (2007). "Identification of a Ctfc cofactor, Yy1, for the X chromosome binary switch." *Mol Cell* 25(1): 43-56.
- Donohoe, M. E., X. Zhang, L. McGinnis, J. Biggers, E. Li and Y. Shi (1999). "Targeted disruption of mouse Yin Yang 1 transcription factor results in peri-implantation lethality." *Mol Cell Biol* 19(10): 7237-7244.
- Doyen, C. M., W. An, D. Angelov, V. Bondarenko, F. Miettinen, V. M. Studitsky, A. Hamiche, R. G. Roeder, P. Bouvet and S. Dimitrov (2006). "Mechanism of polymerase II transcription repression by the histone variant macroH2A." *Mol Cell Biol* 26(3): 1156-1164.
- Dunn, K. L., H. Zhao and J. R. Davie (2003). "The insulator binding protein CTCF associates with the nuclear matrix." *Exp Cell Res* 288(1): 218-223.
- Durcova-Hills, G., J. Ainscough and A. McLaren (2001). "Pluripotential stem cells derived from migrating primordial germ cells." *Differentiation* 68(4-5): 220-226.
- Durcova-Hills, G., P. Hajkova, S. Sullivan, S. Barton, M. A. Surani and A. McLaren (2006). "Influence of sex chromosome constitution on the genomic imprinting of germ cells." *Proc Natl Acad Sci U S A* 103(30): 11184-11188.
- Duret, L., C. Chureau, S. Samain, J. Weissenbach and P. Avner (2006). "The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene." *Science* 312(5780): 1653-1655.
- Duthie, S. M., T. B. Nesterova, E. J. Formstone, A. M. Keohane, B. M. Turner, S. M. Zakian and N. Brockdorff (1999). "Xist RNA exhibits a banded localization on the inactive X chromosome and is excluded from autosomal material in cis." *Hum Mol Genet* 8(2): 195-204.
- Dworetzky, S. I., K. L. Wright, E. G. Fey, S. Penman, J. B. Lian, J. L. Stein and G. S. Stein (1992). "Sequence-specific DNA-binding proteins are components of a nuclear matrix-attachment site." *Proc Natl Acad Sci U S A* 89(9): 4178-4182.
- Dyer, K. A., T. K. Canfield and S. M. Gartler (1989). "Molecular cytological differentiation of active from inactive X domains in interphase: implications for X chromosome inactivation." *Cytogenet Cell Genet* 50(2-3): 116-120.
- Eils, R., S. Dietzel, E. Bertin, E. Schrock, M. R. Speicher, T. Ried, M. Robert-Nicoud, C. Cremer and T. Cremer (1996). "Three-dimensional reconstruction of painted human interphase chromosomes: active and inactive X chromosome territories have similar volumes but differ in shape and surface structure." *J Cell Biol* 135(6 Pt 1): 1427-1440.
- El-Kady, A. and E. Klenova (2005). "Regulation of the transcription factor, CTCF, by phosphorylation with protein kinase CK2." *FEBS Lett* 579(6): 1424-1434.
- Engel, N., J. L. Thorvaldsen and M. S. Bartolomei (2006). "CTCF binding sites promote transcription initiation and prevent DNA methylation on the maternal allele at the imprinted H19/Igf2 locus." *Hum Mol Genet* 15(19): 2945-2954.
- Ensminger, A. W. and A. Chess (2004). "Coordinated replication timing of monoallelically expressed genes along human autosomes." *Hum Mol Genet* 13(6): 651-658.
- Evans, M. J. and M. H. Kaufman (1981). "Establishment in culture of pluripotential cells from mouse embryos." *Nature* 292(5819): 154-156.
- Evans, P. M., W. Zhang, X. Chen, J. Yang, K. K. Bhakat and C. Liu (2007). "Kruppel-like factor 4 is acetylated by p300 and regulates gene transcription via modulation of histone acetylation." *J Biol Chem* 282(47): 33994-34002.
- Fang, J., T. Chen, B. Chadwick, E. Li and Y. Zhang (2004). "Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation." *J Biol Chem* 279(51): 52812-52815.

- Farrar, D., S. Rai, I. Chernukhin, M. Jagodic, Y. Ito, S. Yammine, R. Ohlsson, A. Murrell and E. Klenova (2010). "Mutational analysis of the poly(ADP-ribosylation) sites of the transcription factor CTCF provides an insight into the mechanism of its regulation by poly(ADP-ribosylation)." Mol Cell Biol 30(5): 1199-1216.
- Fedoriw, A. M., P. Stein, P. Svoboda, R. M. Schultz and M. S. Bartolomei (2004). "Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting." Science 303(5655): 238-240.
- Feng, Q. and Y. Zhang (2001). "The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes." Genes Dev 15(7): 827-832.
- Feng, Y. Q., R. Alami and E. E. Bouhassira (1999). "Enhancer-dependent transcriptional oscillations in mouse erythroleukemia cells." Mol Cell Biol 19(7): 4907-4917.
- Ficzycz, A. and N. Ovsenek (2002). "The Yin Yang 1 transcription factor associates with ribonucleoprotein (mRNP) complexes in the cytoplasm of *Xenopus* oocytes." J Biol Chem 277(10): 8382-8387.
- Filippova, G. N., M. K. Cheng, J. M. Moore, J. P. Truong, Y. J. Hu, D. K. Nguyen, K. D. Tsuchiya and C. M. Distèche (2005). "Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development." Dev Cell 8(1): 31-42.
- Filippova, G. N., S. Fagerlie, E. M. Klenova, C. Myers, Y. Dehner, G. Goodwin, P. E. Neiman, S. J. Collins and V. V. Lobanenko (1996). "An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes." Mol Cell Biol 16(6): 2802-2813.
- Fischle, W., Y. Wang and C. D. Allis (2003). "Binary switches and modification cassettes in histone biology and beyond." Nature 425(6957): 475-479.
- Francavilla, S. and L. Zamboni (1985). "Differentiation of mouse ectopic germinal cells in intra- and perigonadal locations." J Exp Zool 233(1): 101-109.
- Francis, N. J., R. E. Kingston and C. L. Woodcock (2004). "Chromatin compaction by a polycomb group protein complex." Science 306(5701): 1574-1577.
- Fuks, F., W. A. Burgers, A. Brehm, L. Hughes-Davies and T. Kouzarides (2000). "DNA methyltransferase Dnmt1 associates with histone deacetylase activity." Nat Genet 24(1): 88-91.
- Fuks, F., W. A. Burgers, N. Godin, M. Kasai and T. Kouzarides (2001). "Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription." EMBO J 20(10): 2536-2544.
- Fullwood, M. J. and Y. Ruan (2009). "ChIP-based methods for the identification of long-range chromatin interactions." J Cell Biochem 107(1): 30-39.
- Galante, S., P. K. Purbey, D. Notani and P. P. Kumar (2007). "The third dimension of gene regulation: organization of dynamic chromatin loopscape by SATB1." Curr Opin Genet Dev 17(5): 408-414.
- Galvin, K. M. and Y. Shi (1997). "Multiple mechanisms of transcriptional repression by YY1." Mol Cell Biol 17(7): 3723-3732.
- Gamer, L. W. and C. V. Wright (1993). "Murine Cdx-4 bears striking similarities to the *Drosophila* caudal gene in its homeodomain sequence and early expression pattern." Mech Dev 43(1): 71-81.
- Garrick, D., J. A. Sharpe, R. Arkell, L. Dobbie, A. J. Smith, W. G. Wood, D. R. Higgs and R. J. Gibbons (2006). "Loss of Atrx affects trophoblast development and the pattern of X-inactivation in extraembryonic tissues." PLoS Genet 2(4): e58.
- Gartel, A. L., X. Ye, E. Goufman, P. Shianov, N. Hay, F. Najmabadi and A. L. Tyner (2001).

- "Myc represses the p21(WAF1/CIP1) promoter and interacts with Sp1/Sp3." Proc Natl Acad Sci U S A 98(8): 4510-4515.
- Gartler, S. M., L. Goldstein, S. E. Tyler-Freer and R. S. Hansen (1999). "The timing of XIST replication: dominance of the domain." Hum Mol Genet 8(6): 1085-1089.
- Gaszner, M. and G. Felsenfeld (2006). "Insulators: exploiting transcriptional and epigenetic mechanisms." Nat Rev Genet 7(9): 703-713.
- Geijsen, N., M. Horoschak, K. Kim, J. Gribnau, K. Eggan and G. Q. Daley (2004). "Derivation of embryonic germ cells and male gametes from embryonic stem cells." Nature 427(6970): 148-154.
- Gelbart, M. E. and M. I. Kuroda (2009). "Drosophila dosage compensation: a complex voyage to the X chromosome." Development 136(9): 1399-1410.
- Gerlich, D., B. Koch, F. Dupeux, J. M. Peters and J. Ellenberg (2006). "Live-cell imaging reveals a stable cohesin-chromatin interaction after but not before DNA replication." Curr Biol 16(15): 1571-1578.
- Giannelli, F. (1963). "The pattern of X-chromosome deoxy-ribonucleic acid synthesis in two women with abnormal sex-chromosome complements." Lancet 1(7286): 863-865.
- Gibbons, R. J., T. L. McDowell, S. Raman, D. M. O'Rourke, D. Garrick, H. Ayyub and D. R. Higgs (2000). "Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation." Nat Genet 24(4): 368-371.
- Gilbert, D. M. (2002). "Replication timing and transcriptional control: beyond cause and effect." Curr Opin Cell Biol 14(3): 377-383.
- Gilbert, S. L., J. R. Pehrson and P. A. Sharp (2000). "XIST RNA associates with specific regions of the inactive X chromatin." J Biol Chem 275(47): 36491-36494.
- Ginsburg, M., M. H. Snow and A. McLaren (1990). "Primordial germ cells in the mouse embryo during gastrulation." Development 110(2): 521-528.
- Goll, M. G. and T. H. Bestor (2005). "Eukaryotic cytosine methyltransferases." Annu Rev Biochem 74: 481-514.
- Gondor, A. and R. Ohlsson (2009). "Chromosome crosstalk in three dimensions." Nature 461(7261): 212-217.
- Goto, Y., M. Gomez, N. Brockdorff and R. Feil (2002). "Differential patterns of histone methylation and acetylation distinguish active and repressed alleles at X-linked genes." Cytogenet Genome Res 99(1-4): 66-74.
- Greaves, I. K., D. Rangasamy, M. Devoy, J. A. Marshall Graves and D. J. Tremethick (2006). "The X and Y chromosomes assemble into H2A.Z-containing [corrected] facultative heterochromatin [corrected] following meiosis." Mol Cell Biol 26(14): 5394-5405.
- Grewal, S. I. and S. C. Elgin (2007). "Transcription and RNA interference in the formation of heterochromatin." Nature 447(7143): 399-406.
- Gribnau, J., S. Luikenhuis, K. Hochedlinger, K. Monkhorst and R. Jaenisch (2005). "X chromosome choice occurs independently of asynchronous replication timing." J Cell Biol 168(3): 365-373.
- Grisendi, S., R. Bernardi, M. Rossi, K. Cheng, L. Khandker, K. Manova and P. P. Pandolfi (2005). "Role of nucleophosmin in embryonic development and tumorigenesis." Nature 437(7055): 147-153.
- Grumbach, M. M., A. Morishima and J. H. Taylor (1963). "Human Sex Chromosome Abnormalities in Relation to DNA Replication and Heterochromatinization." Proc Natl Acad Sci U S A 49(5): 581-589.
- Guastafierro, T., B. Cecchinelli, M. Zampieri, A. Reale, G. Riggio, O. Sthandier, G. Zupi, L. Calabrese and P. Caiafa (2008). "CCCTC-binding factor activates PARP-1 affecting DNA methylation machinery." J Biol Chem 283(32): 21873-21880.
- Guo, G., J. Yang, J. Nichols, J. S. Hall, I. Eyres, W. Mansfield and A. Smith (2009). "Klf4

- reverts developmentally programmed restriction of ground state pluripotency." Development 136(7): 1063-1069.
- Gupta, V., M. Parisi, D. Sturgill, R. Nuttall, M. Doctolero, O. K. Dudko, J. D. Malley, P. S. Eastman and B. Oliver (2006). "Global analysis of X-chromosome dosage compensation." J Biol 5(1): 3.
- Hajkova, P., S. Erhardt, N. Lane, T. Haaf, O. El-Maarri, W. Reik, J. Walter and M. A. Surani (2002). "Epigenetic reprogramming in mouse primordial germ cells." Mech Dev 117(1-2): 15-23.
- Handel, M. A. (2004). "The XY body: a specialized meiotic chromatin domain." Exp Cell Res 296(1): 57-63.
- Hansen, R. S. (2003). "X inactivation-specific methylation of LINE-1 elements by DNMT3B: implications for the Lyon repeat hypothesis." Hum Mol Genet 12(19): 2559-2567.
- Hansen, R. S., T. K. Canfield, A. D. Fjeld and S. M. Gartler (1996). "Role of late replication timing in the silencing of X-linked genes." Hum Mol Genet 5(9): 1345-1353.
- Hargreaves, D. C., T. Horng and R. Medzhitov (2009). "Control of inducible gene expression by signal-dependent transcriptional elongation." Cell 138(1): 129-145.
- Hark, A. T., C. J. Schoenherr, D. J. Katz, R. S. Ingram, J. M. Levorse and S. M. Tilghman (2000). "CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus." Nature 405(6785): 486-489.
- Hata, K., M. Okano, H. Lei and E. Li (2002). "Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice." Development 129(8): 1983-1993.
- He, Y. and P. Casaccia-Bonnel (2008). "The Yin and Yang of YY1 in the nervous system." J Neurochem 106(4): 1493-1502.
- Heard, E. (2005). "Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome." Curr Opin Genet Dev 15(5): 482-489.
- Heard, E., P. Avner and R. Rothstein (1994). "Creation of a deletion series of mouse YACs covering a 500 kb region around Xist." Nucleic Acids Res 22(10): 1830-1837.
- Heard, E., P. Clerc and P. Avner (1997). "X-chromosome inactivation in mammals." Annu Rev Genet 31: 571-610.
- Heard, E. and C. M. Disteche (2006). "Dosage compensation in mammals: fine-tuning the expression of the X chromosome." Genes Dev 20(14): 1848-1867.
- Heard, E., C. Kress, F. Mongelard, B. Courtier, C. Rougeulle, A. Ashworth, C. Vourc'h, C. Babinet and P. Avner (1996). "Transgenic mice carrying an Xist-containing YAC." Hum Mol Genet 5(4): 441-450.
- Heard, E., F. Mongelard, D. Arnaud and P. Avner (1999). "Xist yeast artificial chromosome transgenes function as X-inactivation centers only in multicopy arrays and not as single copies." Mol Cell Biol 19(4): 3156-3166.
- Heard, E., F. Mongelard, D. Arnaud, C. Chureau, C. Vourc'h and P. Avner (1999). "Human XIST yeast artificial chromosome transgenes show partial X inactivation center function in mouse embryonic stem cells." Proc Natl Acad Sci U S A 96(12): 6841-6846.
- Heard, E., C. Rougeulle, D. Arnaud, P. Avner, C. D. Allis and D. L. Spector (2001). "Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation." Cell 107(6): 727-738.
- Heard, E., M. C. Simmler, Z. Larin, C. Rougeulle, B. Courtier, H. Lehrach and P. Avner (1993). "Physical mapping and YAC contig analysis of the region surrounding Xist on the mouse X chromosome." Genomics 15(3): 559-569.
- Heath, H., C. Ribeiro de Almeida, F. Sleutels, G. Dingjan, S. van de Nobelen, I. Jonkers, K. W. Ling, J. Gribnau, R. Renkawitz, F. Grosveld, R. W. Hendriks and N. Galjart

- (2008). "CTCF regulates cell cycle progression of alphabeta T cells in the thymus." EMBO J 27(21): 2839-2850.
- Hekimoglu, B. and L. Ringrose (2009). "Non-coding RNAs in polycomb/trithorax regulation." RNA Biol 6(2): 129-137.
- Hellman, A., A. Rahat, S. W. Scherer, A. Darvasi, L. C. Tsui and B. Kerem (2000). "Replication delay along FRA7H, a common fragile site on human chromosome 7, leads to chromosomal instability." Mol Cell Biol 20(12): 4420-4427.
- Hendrich, B. D., R. M. Plenge and H. F. Willard (1997). "Identification and characterization of the human XIST gene promoter: implications for models of X chromosome inactivation." Nucleic Acids Res 25(13): 2661-2671.
- Her, Y. R. and I. K. Chung (2009). "Ubiquitin Ligase RLIM Modulates Telomere Length Homeostasis through a Proteolysis of TRF1." J Biol Chem 284(13): 8557-8566.
- Hernandez-Munoz, I., A. H. Lund, P. van der Stoep, E. Boutsma, I. Muijers, E. Verhoeven, D. A. Nusinow, B. Panning, Y. Marahrens and M. van Lohuizen (2005). "Stable X chromosome inactivation involves the PRC1 Polycomb complex and requires histone MACROH2A1 and the CULLIN3/SPOP ubiquitin E3 ligase." Proc Natl Acad Sci U S A 102(21): 7635-7640.
- Hernandez-Munoz, P., A. Kanavouras, R. Villalobos and A. Chiralt (2004). "Characterization of biodegradable films obtained from cysteine-mediated polymerized gliadins." J Agric Food Chem 52(26): 7897-7904.
- Hernandez-Munoz, P., J. M. Lagaron, A. Lopez-Rubio and R. Gavara (2004). "Gliadins polymerized with cysteine: effects on the physical and water barrier properties of derived films." Biomacromolecules 5(4): 1503-1510.
- Hernandez-Munoz, P., A. Lopez-Rubio, V. Del-Valle, E. Almenar and R. Gavara (2004). "Mechanical and water barrier properties of glutenin films influenced by storage time." J Agric Food Chem 52(1): 79-83.
- Hernandez-Munoz, P., A. Lopez-Rubio, J. M. Lagaron and R. Gavara (2004). "Formaldehyde cross-linking of gliadin films: effects on mechanical and water barrier properties." Biomacromolecules 5(2): 415-421.
- Herzing, L. B., J. T. Romer, J. M. Horn and A. Ashworth (1997). "Xist has properties of the X-chromosome inactivation centre." Nature 386(6622): 272-275.
- Higaki, Y., D. Schullery, Y. Kawata, M. Shnyreva, C. Abrass and K. Bomsztyk (2002). "Synergistic activation of the rat laminin gamma1 chain promoter by the gut-enriched Kruppel-like factor (GKLF/KLF4) and Sp1." Nucleic Acids Res 30(11): 2270-2279.
- Hiromura, M., C. H. Choi, N. A. Sabourin, H. Jones, D. Bachvarov and A. Usheva (2003). "YY1 is regulated by O-linked N-acetylglucosaminylation (O-glcNAcylation)." J Biol Chem 278(16): 14046-14052.
- Hoki, Y., N. Kimura, M. Kanbayashi, Y. Amakawa, T. Ohhata, H. Sasaki and T. Sado (2009). "A proximal conserved repeat in the Xist gene is essential as a genomic element for X-inactivation in mouse." Development 136(1): 139-146.
- Hong, Y. K., S. D. Ontiveros, C. Chen and W. M. Strauss (1999). "A new structure for the murine Xist gene and its relationship to chromosome choice/counting during X-chromosome inactivation." Proc Natl Acad Sci U S A 96(12): 6829-6834.
- Hong, Y. K., S. D. Ontiveros and W. M. Strauss (2000). "A revision of the human XIST gene organization and structural comparison with mouse Xist." Mamm Genome 11(3): 220-224.
- Horak, C. E. and M. Snyder (2002). "ChIP-chip: a genomic approach for identifying transcription factor binding sites." Methods Enzymol 350: 469-483.
- Hou, C., R. Dale and A. Dean (2010). "Cell type specificity of chromatin organization mediated by CTCF and cohesin." Proc Natl Acad Sci U S A 107(8): 3651-3656.

- Huang, S. (2000). "Review: perinucleolar structures." *J Struct Biol* 129(2-3): 233-240.
- Huynh, K. D. and J. T. Lee (2003). "Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos." *Nature* 426(6968): 857-862.
- Huynh, K. D. and J. T. Lee (2005). "X-chromosome inactivation: a hypothesis linking ontogeny and phylogeny." *Nat Rev Genet* 6(5): 410-418.
- Ideraabdullah, F. Y., S. Vigneau and M. S. Bartolomei (2008). "Genomic imprinting mechanisms in mammals." *Mutat Res* 647(1-2): 77-85.
- Ilik, I. and A. Akhtar (2009). "roX RNAs: non-coding regulators of the male X chromosome in flies." *RNA Biol* 6(2): 113-121.
- Inouye, C. J. and E. Seto (1994). "Relief of YY1-induced transcriptional repression by protein-protein interaction with the nucleolar phosphoprotein B23." *J Biol Chem* 269(9): 6506-6510.
- Irvine, D. V., M. Zaratiegui, N. H. Tolia, D. B. Goto, D. H. Chitwood, M. W. Vaughn, L. Joshua-Tor and R. A. Martienssen (2006). "Argonaute slicing is required for heterochromatic silencing and spreading." *Science* 313(5790): 1134-1137.
- Ishihara, K., M. Oshimura and M. Nakao (2006). "CTCF-dependent chromatin insulator is linked to epigenetic remodeling." *Mol Cell* 23(5): 733-742.
- Jacob, F. and J. Monod (1961). "Genetic regulatory mechanisms in the synthesis of proteins." *J Mol Biol* 3: 318-356.
- Jenuwein, T. and C. D. Allis (2001). "Translating the histone code." *Science* 293(5532): 1074-1080.
- Johnsen, S. A., C. Gungor, T. Prenzel, S. Riethdorf, L. Riethdorf, N. Taniguchi-Ishigaki, T. Rau, B. Tursun, J. D. Furlow, G. Sauter, M. Scheffner, K. Pantel, F. Gannon and I. Bach (2009). "Regulation of estrogen-dependent transcription by the LIM cofactors CLIM and RLIM in breast cancer." *Cancer Res* 69(1): 128-136.
- Johnston, C. M., T. B. Nesterova, E. J. Formstone, A. E. Newall, S. M. Duthie, S. A. Sheardown and N. Brockdorff (1998). "Developmentally regulated Xist promoter switch mediates initiation of X inactivation." *Cell* 94(6): 809-817.
- Jones, P. L., G. J. Veenstra, P. A. Wade, D. Vermaak, S. U. Kass, N. Landsberger, J. Strouboulis and A. P. Wolffe (1998). "Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription." *Nat Genet* 19(2): 187-191.
- Jonkers, I., T. S. Barakat, E. M. Achame, K. Monkhorst, A. Kenter, E. Rentmeester, F. Grosveld, J. A. Grootegeed and J. Gribnau (2009). "RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation." *Cell* 139(5): 999-1011.
- Jothi, R., S. Cuddapah, A. Barski, K. Cui and K. Zhao (2008). "Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data." *Nucleic Acids Res* 36(16): 5221-5231.
- Kalantry, S. and T. Magnuson (2006). "The Polycomb group protein EED is dispensable for the initiation of random X-chromosome inactivation." *PLoS Genet* 2(5): e66.
- Kalantry, S., S. Purushothaman, R. B. Bowen, J. Starmer and T. Magnuson (2009). "Evidence of Xist RNA-independent initiation of mouse imprinted X-chromosome inactivation." *Nature* 460(7255): 647-651.
- Kanatsu-Shinohara, M., K. Inoue, J. Lee, M. Yoshimoto, N. Ogonuki, H. Miki, S. Baba, T. Kato, Y. Kazuki, S. Toyokuni, M. Toyoshima, O. Niwa, M. Oshimura, T. Heike, T. Nakahata, F. Ishino, A. Ogura and T. Shinohara (2004). "Generation of pluripotent stem cells from neonatal mouse testis." *Cell* 119(7): 1001-1012.
- Kanduri, C., V. Pant, D. Loukinov, E. Pugacheva, C. F. Qi, A. Wolffe, R. Ohlsson and V. V. Lobanenko (2000). "Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive." *Curr Biol* 10(14): 853-856.

- Kanellopoulou, C., S. A. Muljo, S. D. Dimitrov, X. Chen, C. Colin, K. Plath and D. M. Livingston (2009). "X chromosome inactivation in the absence of Dicer." Proc Natl Acad Sci U S A 106(4): 1122-1127.
- Kang, H. and P. M. Lieberman (2009). "Cell cycle control of Kaposi's sarcoma-associated herpesvirus latency transcription by CTCF-cohesin interactions." J Virol 83(12): 6199-6210.
- Kanhere, A., K. Viiri, C. C. Araujo, J. Rasaiyaah, R. D. Bouwman, W. A. Whyte, C. F. Pereira, E. Brookes, K. Walker, G. W. Bell, A. Pombo, A. G. Fisher, R. A. Young and R. G. Jenner (2010). "Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2." Mol Cell 38(5): 675-688.
- Kawame, H., S. M. Gartler and R. S. Hansen (1995). "Allele-specific replication timing in imprinted domains: absence of asynchrony at several loci." Hum Mol Genet 4(12): 2287-2293.
- Keohane, A. M., A. L. Barlow, J. Waters, D. Bourn and B. M. Turner (1999). "H4 acetylation, XIST RNA and replication timing are coincident and define x;autosome boundaries in two abnormal X chromosomes." Hum Mol Genet 8(2): 377-383.
- Keohane, A. M., P. O'Neill L, N. D. Belyaev, J. S. Lavender and B. M. Turner (1996). "X-Inactivation and histone H4 acetylation in embryonic stem cells." Dev Biol 180(2): 618-630.
- Kernohan, K. D., Y. Jiang, D. C. Tremblay, A. C. Bonvissuto, J. H. Eubanks, M. R. Mann and N. G. Berube (2010). "ATRX partners with cohesin and MeCP2 and contributes to developmental silencing of imprinted genes in the brain." Dev Cell 18(2): 191-202.
- Khalil, A. M., F. Z. Boyar and D. J. Driscoll (2004). "Dynamic histone modifications mark sex chromosome inactivation and reactivation during mammalian spermatogenesis." Proc Natl Acad Sci U S A 101(47): 16583-16587.
- Khalil, A. M., M. Guttman, M. Huarte, M. Garber, A. Raj, D. Rivea Morales, K. Thomas, A. Presser, B. E. Bernstein, A. van Oudenaarden, A. Regev, E. S. Lander and J. L. Rinn (2009). "Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression." Proc Natl Acad Sci U S A 106(28): 11667-11672.
- Kim, D. H., L. M. Villeneuve, K. V. Morris and J. J. Rossi (2006). "Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells." Nat Struct Mol Biol 13(9): 793-797.
- Kim, J. (2009). "YY1's longer DNA-binding motifs." Genomics 93(2): 152-158.
- Kim, J., J. Chu, X. Shen, J. Wang and S. H. Orkin (2008). "An extended transcriptional network for pluripotency of embryonic stem cells." Cell 132(6): 1049-1061.
- Kim, J. and J. D. Kim (2008). "In vivo YY1 knockdown effects on genomic imprinting." Hum Mol Genet 17(3): 391-401.
- Kim, J., A. Kollhoff, A. Bergmann and L. Stubbs (2003). "Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3." Hum Mol Genet 12(3): 233-245.
- Kim, J. D., C. Faulk and J. Kim (2007). "Retroposition and evolution of the DNA-binding motifs of YY1, YY2 and REX1." Nucleic Acids Res 35(10): 3442-3452.
- Kim, J. D., A. K. Hinz, A. Bergmann, J. M. Huang, I. Ovcharenko, L. Stubbs and J. Kim (2006). "Identification of clustered YY1 binding sites in imprinting control regions." Genome Res 16(7): 901-911.
- Kim, J. D., A. K. Hinz, J. H. Choo, L. Stubbs and J. Kim (2007). "YY1 as a controlling factor for the Peg3 and Gnas imprinted domains." Genomics 89(2): 262-269.
- Kim, J. D., K. Kang and J. Kim (2009). "YY1's role in DNA methylation of Peg3 and Xist." Nucleic Acids Res 37(17): 5656-5664.

- Kim, J. D., S. Yu and J. Kim (2009). "YY1 is autoregulated through its own DNA-binding sites." BMC Mol Biol 10: 85.
- Kim, T. H., Z. K. Abdullaev, A. D. Smith, K. A. Ching, D. I. Loukinov, R. D. Green, M. Q. Zhang, V. V. Lobanenko and B. Ren (2007). "Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome." Cell 128(6): 1231-1245.
- Kimura, H., M. Tada, S. Hatano, M. Yamazaki, N. Nakatsuji and T. Tada (2002). "Chromatin reprogramming of male somatic cell-derived XIST and TSIX in ES hybrid cells." Cytogenet Genome Res 99(1-4): 106-114.
- Kitchen, N. S. and C. J. Schoenherr (2010). "Sumoylation modulates a domain in CTCF that activates transcription and decondenses chromatin." J Cell Biochem.
- Klar, M. and J. Bode (2005). "Enhanceosome formation over the beta interferon promoter underlies a remote-control mechanism mediated by YY1 and YY2." Mol Cell Biol 25(22): 10159-10170.
- Klenova, E., A. C. Scott, J. Roberts, S. Shamsuddin, E. A. Lovejoy, S. Bergmann, V. J. Bubh, H. D. Royer and J. P. Quinn (2004). "YB-1 and CTCF differentially regulate the 5-HTT polymorphic intron 2 enhancer which predisposes to a variety of neurological disorders." J Neurosci 24(26): 5966-5973.
- Klenova, E. M., I. V. Chernukhin, A. El-Kady, R. E. Lee, E. M. Pugacheva, D. I. Loukinov, G. H. Goodwin, D. Delgado, G. N. Filippova, J. Leon, H. C. Morse, 3rd, P. E. Neiman and V. V. Lobanenko (2001). "Functional phosphorylation sites in the C-terminal region of the multivalent multifunctional transcriptional factor CTCF." Mol Cell Biol 21(6): 2221-2234.
- Klenova, E. M., S. Fagerlie, G. N. Filippova, L. Kretzner, G. H. Goodwin, G. Loring, P. E. Neiman and V. V. Lobanenko (1998). "Characterization of the chicken CTCF genomic locus, and initial study of the cell cycle-regulated promoter of the gene." J Biol Chem 273(41): 26571-26579.
- Klenova, E. M., R. H. Nicolas, H. F. Paterson, A. F. Carne, C. M. Heath, G. H. Goodwin, P. E. Neiman and V. V. Lobanenko (1993). "CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms." Mol Cell Biol 13(12): 7612-7624.
- Ko, C. Y., H. C. Hsu, M. R. Shen, W. C. Chang and J. M. Wang (2008). "Epigenetic silencing of CCAAT/enhancer-binding protein delta activity by YY1/polycomb group/DNA methyltransferase complex." J Biol Chem 283(45): 30919-30932.
- Kohlmaier, A., F. Savarese, M. Lachner, J. Martens, T. Jenuwein and A. Wutz (2004). "A chromosomal memory triggered by Xist regulates histone methylation in X inactivation." PLoS Biol 2(7): E171.
- Kohwi-Shigematsu, T., K. Maass and J. Bode (1997). "A thymocyte factor SATB1 suppresses transcription of stably integrated matrix-attachment region-linked reporter genes." Biochemistry 36(40): 12005-12010.
- Kourmouli, N., Y. M. Sun, S. van der Sar, P. B. Singh and J. P. Brown (2005). "Epigenetic regulation of mammalian pericentric heterochromatin in vivo by HP1." Biochem Biophys Res Commun 337(3): 901-907.
- Kratzer, P. G., V. M. Chapman, H. Lambert, R. E. Evans and R. M. Liskay (1983). "Differences in the DNA of the inactive X chromosomes of fetal and extraembryonic tissues of mice." Cell 33(1): 37-42.
- Krogan, N. J., M. Kim, A. Tong, A. Golshani, G. Cagney, V. Canadien, D. P. Richards, B. K. Beattie, A. Emili, C. Boone, A. Shilatifard, S. Buratowski and J. Greenblatt (2003). "Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II." Mol Cell Biol 23(12): 4207-4218.
- Kunath, T., D. Arnaud, G. D. Uy, I. Okamoto, C. Chureau, Y. Yamanaka, E. Heard, R. L.

- Gardner, P. Avner and J. Rossant (2005). "Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts." Development 132(7): 1649-1661.
- Kurisaki, K., A. Kurisaki, U. Valcourt, A. A. Terentiev, K. Pardali, P. Ten Dijke, C. H. Heldin, J. Ericsson and A. Moustakas (2003). "Nuclear factor YY1 inhibits transforming growth factor beta- and bone morphogenetic protein-induced cell differentiation." Mol Cell Biol 23(13): 4494-4510.
- Kurukuti, S., V. K. Tiwari, G. Tavoosidana, E. Pugacheva, A. Murrell, Z. Zhao, V. Lobanenko, W. Reik and R. Ohlsson (2006). "CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2." Proc Natl Acad Sci U S A 103(28): 10684-10689.
- Kwok, R. P., J. R. Lundblad, J. C. Chrivia, J. P. Richards, H. P. Bachinger, R. G. Brennan, S. G. Roberts, M. R. Green and R. H. Goodman (1994). "Nuclear protein CBP is a coactivator for the transcription factor CREB." Nature 370(6486): 223-226.
- Lafreniere, R. G. and H. F. Willard (1993). "Pulsed-field map of Xq13 in the region of the human X inactivation center." Genomics 17(2): 502-506.
- Lee, J., K. Inoue, R. Ono, N. Ogonuki, T. Kohda, T. Kaneko-Ishino, A. Ogura and F. Ishino (2002). "Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells." Development 129(8): 1807-1817.
- Lee, J. S., K. M. Galvin and Y. Shi (1993). "Evidence for physical interaction between the zinc-finger transcription factors YY1 and Sp1." Proc Natl Acad Sci U S A 90(13): 6145-6149.
- Lee, J. S., X. Zhang and Y. Shi (1996). "Differential interactions of the CREB/ATF family of transcription factors with p300 and adenovirus E1A." J Biol Chem 271(30): 17666-17674.
- Lee, J. T. (2000). "Disruption of imprinted X inactivation by parent-of-origin effects at Tsix." Cell 103(1): 17-27.
- Lee, J. T. (2005). "Regulation of X-chromosome counting by Tsix and Xite sequences." Science 309(5735): 768-771.
- Lee, J. T., L. S. Davidow and D. Warshawsky (1999). "Tsix, a gene antisense to Xist at the X-inactivation centre." Nat Genet 21(4): 400-404.
- Lee, J. T., N. Lu and Y. Han (1999). "Genetic analysis of the mouse X inactivation center defines an 80-kb multifunction domain." Proc Natl Acad Sci U S A 96(7): 3836-3841.
- Lee, J. T., W. M. Strauss, J. A. Dausman and R. Jaenisch (1996). "A 450 kb transgene displays properties of the mammalian X-inactivation center." Cell 86(1): 83-94.
- Lee, T. C., Y. Shi and R. J. Schwartz (1992). "Displacement of BrdUrd-induced YY1 by serum response factor activates skeletal alpha-actin transcription in embryonic myoblasts." Proc Natl Acad Sci U S A 89(20): 9814-9818.
- Lee, T. C., Y. Zhang and R. J. Schwartz (1994). "Bifunctional transcriptional properties of YY1 in regulating muscle actin and c-myc gene expression during myogenesis." Oncogene 9(4): 1047-1052.
- Leeb, M., D. Pasini, M. Novatchkova, M. Jaritz, K. Helin and A. Wutz (2010). "Polycomb complexes act redundantly to repress genomic repeats and genes." Genes Dev 24(3): 265-276.
- Leeb, M. and A. Wutz (2007). "Ring1B is crucial for the regulation of developmental control genes and PRC1 proteins but not X inactivation in embryonic cells." J Cell Biol 178(2): 219-229.
- Leppig, K. A., C. J. Brown, S. L. Bressler, K. Gustashaw, R. A. Pagon, H. F. Willard and C. M. Disteche (1993). "Mapping of the distal boundary of the X-inactivation center in a rearranged X chromosome from a female expressing XIST." Hum Mol Genet 2(7):

883-887.

- Lewis, A. and A. Murrell (2004). "Genomic imprinting: CTCF protects the boundaries." Curr Biol 14(7): R284-286.
- Li, M., P. Baumeister, B. Roy, T. Phan, D. Foti, S. Luo and A. S. Lee (2000). "ATF6 as a transcription activator of the endoplasmic reticulum stress element: thapsigargin stress-induced changes and synergistic interactions with NF-Y and YY1." Mol Cell Biol 20(14): 5096-5106.
- Li, T., J. F. Hu, X. Qiu, J. Ling, H. Chen, S. Wang, A. Hou, T. H. Vu and A. R. Hoffman (2008). "CTCF regulates allelic expression of Igf2 by orchestrating a promoter-polycomb repressive complex 2 intrachromosomal loop." Mol Cell Biol 28(20): 6473-6482.
- Lifschytz, E. and D. L. Lindsley (1972). "The role of X-chromosome inactivation during spermatogenesis (Drosophila-allocyly-chromosome evolution-male sterility-dosage compensation)." Proc Natl Acad Sci U S A 69(1): 182-186.
- Lindsley, D. L., L. Sandler, B. S. Baker, A. T. Carpenter, R. E. Denell, J. C. Hall, P. A. Jacobs, G. L. Miklos, B. K. Davis, R. C. Gethmann, R. W. Hardy, A. H. Steven, M. Miller, H. Nozawa, D. M. Parry and M. Gould-Somero (1972). "Segmental aneuploidy and the genetic gross structure of the Drosophila genome." Genetics 71(1): 157-184.
- Ling, J. Q., T. Li, J. F. Hu, T. H. Vu, H. L. Chen, X. W. Qiu, A. M. Cherry and A. R. Hoffman (2006). "CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1." Science 312(5771): 269-272.
- Liu, H., J. Huang, J. Wang, S. Jiang, A. S. Bailey, D. C. Goldman, M. Welcker, V. Bedell, M. L. Slovak, B. Clurman, M. Thayer, W. H. Fleming and E. Epner (2008). "Transvection mediated by the translocated cyclin D1 locus in mantle cell lymphoma." J Exp Med 205(8): 1843-1858.
- Liu, H., M. Schmidt-Suppran, Y. Shi, E. Hobeika, N. Barteneva, H. Jumaa, R. Pelanda, M. Reth, J. Skok and K. Rajewsky (2007). "Yin Yang 1 is a critical regulator of B-cell development." Genes Dev 21(10): 1179-1189.
- Liu, W. L., R. A. Coleman, E. Ma, P. Grob, J. L. Yang, Y. Zhang, G. Dailey, E. Nogales and R. Tjian (2009). "Structures of three distinct activator-TFIID complexes." Genes Dev 23(13): 1510-1521.
- Lock, L. F., N. Takagi and G. R. Martin (1987). "Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation." Cell 48(1): 39-46.
- Loh, Y. H., Q. Wu, J. L. Chew, V. B. Vega, W. Zhang, X. Chen, G. Bourque, J. George, B. Leong, J. Liu, K. Y. Wong, K. W. Sung, C. W. Lee, X. D. Zhao, K. P. Chiu, L. Lipovich, V. A. Kuznetsov, P. Robson, L. W. Stanton, C. L. Wei, Y. Ruan, B. Lim and H. H. Ng (2006). "The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells." Nat Genet 38(4): 431-440.
- Loukinov, D. I., E. Pugacheva, S. Vatolin, S. D. Pack, H. Moon, I. Chernukhin, P. Mannan, E. Larsson, C. Kanduri, A. A. Vostrov, H. Cui, E. L. Niemitz, J. E. Rasko, F. M. Docquier, M. Kistler, J. J. Breen, Z. Zhuang, W. W. Quitschke, R. Renkawitz, E. M. Klenova, A. P. Feinberg, R. Ohlsson, H. C. Morse, 3rd and V. V. Lobanenkov (2002). "BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma." Proc Natl Acad Sci U S A 99(10): 6806-6811.
- Luger, K., T. J. Rechsteiner, A. J. Flaus, M. M. Wayne and T. J. Richmond (1997). "Characterization of nucleosome core particles containing histone proteins made in bacteria." J Mol Biol 272(3): 301-311.

- Luikenhuis, S., A. Wutz and R. Jaenisch (2001). "Antisense transcription through the Xist locus mediates Tsix function in embryonic stem cells." Mol Cell Biol 21(24): 8512-8520.
- Luke, M. P., G. Sui, H. Liu and Y. Shi (2006). "Yin Yang 1 physically interacts with Hoxa11 and represses Hoxa11-dependent transcription." J Biol Chem 281(44): 33226-33232.
- Lutz, M., L. J. Burke, G. Barreto, F. Goeman, H. Greb, R. Arnold, H. Schultheiss, A. Brehm, T. Kouzarides, V. Lobanenko and R. Renkawitz (2000). "Transcriptional repression by the insulator protein CTCF involves histone deacetylases." Nucleic Acids Res 28(8): 1707-1713.
- Lyon, M. F. (1961). "Gene action in the X-chromosome of the mouse (*Mus musculus* L.)." Nature 190: 372-373.
- Lyon, M. F. (1962). "Sex chromatin and gene action in the mammalian X-chromosome." Am J Hum Genet 14: 135-148.
- Lyon, M. F. (1998). "X-chromosome inactivation: a repeat hypothesis." Cytogenet Cell Genet 80(1-4): 133-137.
- Lyon, M. F. and S. Rastan (1984). "Parental source of chromosome imprinting and its relevance for X chromosome inactivation." Differentiation 26(1): 63-67.
- Ma, M. and W. M. Strauss (2005). "Analysis of the Xist RNA isoforms suggests two distinctly different forms of regulation." Mamm Genome 16(6): 391-404.
- MacPherson, M. J., L. G. Beatty, W. Zhou, M. Du and P. D. Sadowski (2009). "The CTCF insulator protein is posttranslationally modified by SUMO." Mol Cell Biol 29(3): 714-725.
- Maenner, S., M. Blaud, L. Fouillen, A. Savoye, V. Marchand, A. Dubois, S. Sanglier-Cianferani, A. Van Dorsselaer, P. Clerc, P. Avner, A. Visvikis and C. Branlant (2010). "2-D structure of the A region of Xist RNA and its implication for PRC2 association." PLoS Biol 8(1): e1000276.
- Maier, H., J. Colbert, D. Fitzsimmons, D. R. Clark and J. Hagman (2003). "Activation of the early B-cell-specific mb-1 (Ig-alpha) gene by Pax-5 is dependent on an unmethylated Ets binding site." Mol Cell Biol 23(6): 1946-1960.
- Mak, W., J. Baxter, J. Silva, A. E. Newall, A. P. Otte and N. Brockdorff (2002). "Mitotically stable association of polycomb group proteins eed and enx1 with the inactive x chromosome in trophoblast stem cells." Curr Biol 12(12): 1016-1020.
- Mak, W., T. B. Nesterova, M. de Napoles, R. Appanah, S. Yamanaka, A. P. Otte and N. Brockdorff (2004). "Reactivation of the paternal X chromosome in early mouse embryos." Science 303(5658): 666-669.
- Manuelidis, L. and J. Borden (1988). "Reproducible compartmentalization of individual chromosome domains in human CNS cells revealed by in situ hybridization and three-dimensional reconstruction." Chromosoma 96(6): 397-410.
- Marahrens, Y. (1999). "X-inactivation by chromosomal pairing events." Genes Dev 13(20): 2624-2632.
- Marahrens, Y., J. Loring and R. Jaenisch (1998). "Role of the Xist gene in X chromosome choosing." Cell 92(5): 657-664.
- Marahrens, Y., B. Panning, J. Dausman, W. Strauss and R. Jaenisch (1997). "Xist-deficient mice are defective in dosage compensation but not spermatogenesis." Genes Dev 11(2): 156-166.
- Mardis, E. R. (2007). "ChIP-seq: welcome to the new frontier." Nat Methods 4(8): 613-614.
- Margueron, R., P. Trojer and D. Reinberg (2005). "The key to development: interpreting the histone code?" Curr Opin Genet Dev 15(2): 163-176.
- Marks, H., J. C. Chow, S. Denissov, K. J. Francoijs, N. Brockdorff, E. Heard and H. G. Stunnenberg (2009). "High-resolution analysis of epigenetic changes associated with

- X inactivation." Genome Res 19(8): 1361-1373.
- Martin, G. R. (1981). "Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells." Proc Natl Acad Sci U S A 78(12): 7634-7638.
- Martin, G. R., C. J. Epstein, B. Travis, G. Tucker, S. Yatziv, D. W. Martin, Jr., S. Clift and S. Cohen (1978). "X-chromosome inactivation during differentiation of female teratocarcinoma stem cells in vitro." Nature 271(5643): 329-333.
- Massie, C. E. and I. G. Mills (2008). "ChIPping away at gene regulation." EMBO Rep 9(4): 337-343.
- Matsui, Y., K. Zsebo and B. L. Hogan (1992). "Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture." Cell 70(5): 841-847.
- Matzke, M., T. Kanno, B. Huettel, L. Daxinger and A. J. Matzke (2007). "Targets of RNA-directed DNA methylation." Curr Opin Plant Biol 10(5): 512-519.
- Maya-Mendoza, A., R. Hernandez-Munoz, P. Gariglio and A. Aranda-Anzaldo (2004). "Gene positional changes relative to the nuclear substructure during carbon tetrachloride-induced hepatic fibrosis in rats." J Cell Biochem 93(6): 1084-1098.
- McCarrey, J. R. and D. D. Dilworth (1992). "Expression of Xist in mouse germ cells correlates with X-chromosome inactivation." Nat Genet 2(3): 200-203.
- McCarrey, J. R., C. Watson, J. Atencio, G. C. Ostermeier, Y. Marahrens, R. Jaenisch and S. A. Krawetz (2002). "X-chromosome inactivation during spermatogenesis is regulated by an Xist/Tsix-independent mechanism in the mouse." Genesis 34(4): 257-266.
- McDowell, T. L., R. J. Gibbons, H. Sutherland, D. M. O'Rourke, W. A. Bickmore, A. Pombo, H. Turley, K. Gatter, D. J. Picketts, V. J. Buckle, L. Chapman, D. Rhodes and D. R. Higgs (1999). "Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes." Proc Natl Acad Sci U S A 96(24): 13983-13988.
- McLaren, A. (2003). "Primordial germ cells in the mouse." Dev Biol 262(1): 1-15.
- Memili, E., Y. K. Hong, D. H. Kim, S. D. Ontiveros and W. M. Strauss (2001). "Murine Xist RNA isoforms are different at their 3' ends: a role for differential polyadenylation." Gene 266(1-2): 131-137.
- Mermoud, J. E., C. Costanzi, J. R. Pehrson and N. Brockdorff (1999). "Histone macroH2A1.2 relocates to the inactive X chromosome after initiation and propagation of X-inactivation." J Cell Biol 147(7): 1399-1408.
- Mermoud, J. E., B. Popova, A. H. Peters, T. Jenuwein and N. Brockdorff (2002). "Histone H3 lysine 9 methylation occurs rapidly at the onset of random X chromosome inactivation." Curr Biol 12(3): 247-251.
- Messina, D. N., J. Glasscock, W. Gish and M. Lovett (2004). "An ORFeome-based analysis of human transcription factor genes and the construction of a microarray to interrogate their expression." Genome Res 14(10B): 2041-2047.
- Meyer, B. J. (2005). "X-Chromosome dosage compensation." WormBook: 1-14.
- Mietton, F., A. K. Sengupta, A. Molla, G. Picchi, S. Barral, L. Heliot, T. Grange, A. Wutz and S. Dimitrov (2009). "Weak but uniform enrichment of the histone variant macroH2A1 along the inactive X chromosome." Mol Cell Biol 29(1): 150-156.
- Migeon, B. R., A. K. Chowdhury, J. A. Dunston and I. McIntosh (2001). "Identification of TSIX, encoding an RNA antisense to human XIST, reveals differences from its murine counterpart: implications for X inactivation." Am J Hum Genet 69(5): 951-960.
- Migeon, B. R., C. H. Lee, A. K. Chowdhury and H. Carpenter (2002). "Species differences in TSIX/Tsix reveal the roles of these genes in X-chromosome inactivation." Am J Hum Genet 71(2): 286-293.

- Mikkelsen, T. S., M. J. Wakefield, B. Aken, C. T. Amemiya, J. L. Chang, S. Duke, M. Garber, A. J. Gentles, L. Goodstadt, A. Heger, J. Jurka, M. Kamal, E. Mauceli, S. M. Searle, T. Sharpe, M. L. Baker, M. A. Batzer, P. V. Benos, K. Belov, M. Clamp, A. Cook, J. Cuff, R. Das, L. Davidow, J. E. Deakin, M. J. Fazzari, J. L. Glass, M. Grabherr, J. M. Greally, W. Gu, T. A. Hore, G. A. Huttley, M. Kleber, R. L. Jirtle, E. Koina, J. T. Lee, S. Mahony, M. A. Marra, R. D. Miller, R. D. Nicholls, M. Oda, A. T. Papenfuss, Z. E. Parra, D. D. Pollock, D. A. Ray, J. E. Schein, T. P. Speed, K. Thompson, J. L. VandeBerg, C. M. Wade, J. A. Walker, P. D. Waters, C. Webber, J. R. Weidman, X. Xie, M. C. Zody, J. A. Graves, C. P. Ponting, M. Breen, P. B. Samollow, E. S. Lander and K. Lindblad-Toh (2007). "Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences." Nature 447(7141): 167-177.
- Mise, N., T. Fuchikami, M. Sugimoto, S. Kobayakawa, F. Ike, T. Ogawa, T. Tada, S. Kanaya, T. Noce and K. Abe (2008). "Differences and similarities in the developmental status of embryo-derived stem cells and primordial germ cells revealed by global expression profiling." Genes Cells 13(8): 863-877.
- Mise, N., Y. Goto, N. Nakajima and N. Takagi (1999). "Molecular cloning of antisense transcripts of the mouse *Xist* gene." Biochem Biophys Res Commun 258(3): 537-541.
- Mlynarczyk-Evans, S., M. Royce-Tolland, M. K. Alexander, A. A. Andersen, S. Kalantry, J. Gribnau and B. Panning (2006). "X chromosomes alternate between two states prior to random X-inactivation." PLoS Biol 4(6): e159.
- Mohammad, F., R. R. Pandey, T. Nagano, L. Chakalova, T. Mondal, P. Fraser and C. Kanduri (2008). "Kcnq1ot1/Lit1 noncoding RNA mediates transcriptional silencing by targeting to the perinucleolar region." Mol Cell Biol 28(11): 3713-3728.
- Mondello, C., P. J. Goodfellow and P. N. Goodfellow (1988). "Analysis of methylation of a human X located gene which escapes X inactivation." Nucleic Acids Res 16(14B): 6813-6824.
- Monesi, V. (1965). "Differential rate of ribonucleic acid synthesis in the autosomes and sex chromosomes during male meiosis in the mouse." Chromosoma 17(1): 11-21.
- Monk, M., M. Boubelik and S. Lehnert (1987). "Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development." Development 99(3): 371-382.
- Monk, M. and M. I. Harper (1979). "Sequential X chromosome inactivation coupled with cellular differentiation in early mouse embryos." Nature 281(5729): 311-313.
- Monkhorst, K., I. Jonkers, E. Rentmeester, F. Grosveld and J. Gribnau (2008). "X inactivation counting and choice is a stochastic process: evidence for involvement of an X-linked activator." Cell 132(3): 410-421.
- Moreira de Mello, J. C., E. S. de Araujo, R. Stabellini, A. M. Fraga, J. E. de Souza, D. R. Sumita, A. A. Camargo and L. V. Pereira (2010). "Random X inactivation and extensive mosaicism in human placenta revealed by analysis of allele-specific gene expression along the X chromosome." PLoS One 5(6): e10947.
- Morey, C., D. Arnaud, P. Avner and P. Clerc (2001). "Tsix-mediated repression of *Xist* accumulation is not sufficient for normal random X inactivation." Hum Mol Genet 10(13): 1403-1411.
- Morey, C. and W. Bickmore (2006). "Sealed with a X." Nat Cell Biol 8(3): 207-209.
- Morey, C., P. Navarro, E. Debrand, P. Avner, C. Rougeulle and P. Clerc (2004). "The region 3' to *Xist* mediates X chromosome counting and H3 Lys-4 dimethylation within the *Xist* gene." EMBO J 23(3): 594-604.
- Morishima, A., M. M. Grumbach and J. H. Taylor (1962). "Asynchronous duplication of human chromosomes and the origin of sex chromatin." Proc Natl Acad Sci U S A 48:

- 756-763.
- Morris, K. V., S. W. Chan, S. E. Jacobsen and D. J. Looney (2004). "Small interfering RNA-induced transcriptional gene silencing in human cells." Science 305(5688): 1289-1292.
- Mueller, J. E., M. Canze and M. Bryk (2006). "The requirements for COMPASS and Paf1 in transcriptional silencing and methylation of histone H3 in *Saccharomyces cerevisiae*." Genetics 173(2): 557-567.
- Mukhopadhyay, R., W. Yu, J. Whitehead, J. Xu, M. Lezcano, S. Pack, C. Kanduri, M. Kanduri, V. Ginjala, A. Vostrov, W. Quitschke, I. Chernukhin, E. Klenova, V. Lobanenko and R. Ohlsson (2004). "The binding sites for the chromatin insulator protein CTCF map to DNA methylation-free domains genome-wide." Genome Res 14(8): 1594-1602.
- Muyrers-Chen, I., I. Hernandez-Munoz, A. H. Lund, M. E. Valk-Lingbeek, P. van der Stoop, E. Boutsma, B. Tolhuis, S. W. Bruggeman, P. Taghavi, E. Verhoeven, D. Hulsman, S. Noback, E. Tanger, H. Theunissen and M. van Lohuizen (2004). "Emerging roles of Polycomb silencing in X-inactivation and stem cell maintenance." Cold Spring Harb Symp Quant Biol 69: 319-326.
- Namekawa, S. H., P. J. Park, L. F. Zhang, J. E. Shima, J. R. McCarrey, M. D. Griswold and J. T. Lee (2006). "Postmeiotic sex chromatin in the male germline of mice." Curr Biol 16(7): 660-667.
- Namekawa, S. H., B. Payer, K. D. Huynh, R. Jaenisch and J. T. Lee (2010). "Two-step imprinted X inactivation: repeat versus genic silencing in the mouse." Mol Cell Biol 30(13): 3187-3205.
- Nan, X., F. J. Campoy and A. Bird (1997). "MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin." Cell 88(4): 471-481.
- Nan, X., J. Hou, A. Maclean, J. Nasir, M. J. Lafuente, X. Shu, S. Kriaucionis and A. Bird (2007). "Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation." Proc Natl Acad Sci U S A 104(8): 2709-2714.
- Nan, X., H. H. Ng, C. A. Johnson, C. D. Laherty, B. M. Turner, R. N. Eisenman and A. Bird (1998). "Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex." Nature 393(6683): 386-389.
- Natesan, S. and M. Z. Gilman (1993). "DNA bending and orientation-dependent function of YY1 in the c-fos promoter." Genes Dev 7(12B): 2497-2509.
- Navarro, P. and P. Avner (2010). "An embryonic story: analysis of the gene regulative network controlling Xist expression in mouse embryonic stem cells." Bioessays 32(7): 581-588.
- Navarro, P., I. Chambers, V. Karwacki-Neisius, C. Chureau, C. Morey, C. Rougeulle and P. Avner (2008). "Molecular coupling of Xist regulation and pluripotency." Science 321(5896): 1693-1695.
- Navarro, P., S. Chantalat, M. Foglio, C. Chureau, S. Vigneau, P. Clerc, P. Avner and C. Rougeulle (2009). "A role for non-coding Tsix transcription in partitioning chromatin domains within the mouse X-inactivation centre." Epigenetics Chromatin 2(1): 8.
- Navarro, P., D. R. Page, P. Avner and C. Rougeulle (2006). "Tsix-mediated epigenetic switch of a CTCF-flanked region of the Xist promoter determines the Xist transcription program." Genes Dev 20(20): 2787-2792.
- Navarro, P., S. Pichard, C. Ciaudo, P. Avner and C. Rougeulle (2005). "Tsix transcription across the Xist gene alters chromatin conformation without affecting Xist transcription: implications for X-chromosome inactivation." Genes Dev 19(12): 1474-1484.

- Nesterova, T. B., J. E. Mermoud, K. Hilton, J. Pehrson, M. A. Surani, A. McLaren and N. Brockdorff (2002). "Xist expression and macroH2A1.2 localisation in mouse primordial and pluripotent embryonic germ cells." Differentiation 69(4-5): 216-225.
- Nesterova, T. B., B. C. Popova, B. S. Cobb, S. Norton, C. E. Senner, Y. A. Tang, T. Spruce, T. A. Rodriguez, T. Sado, M. Merckenschlager and N. Brockdorff (2008). "Dicer regulates Xist promoter methylation in ES cells indirectly through transcriptional control of Dnmt3a." Epigenetics Chromatin 1(1): 2.
- Nesterova, T. B., S. Y. Slobodyanyuk, E. A. Elisaphenko, A. I. Shevchenko, C. Johnston, M. E. Pavlova, I. B. Rogozin, N. N. Kolesnikov, N. Brockdorff and S. M. Zakian (2001). "Characterization of the genomic Xist locus in rodents reveals conservation of overall gene structure and tandem repeats but rapid evolution of unique sequence." Genome Res 11(5): 833-849.
- Newall, A. E., S. Duthie, E. Formstone, T. Nesterova, M. Alexiou, C. Johnston, M. L. Caparros and N. Brockdorff (2001). "Primary non-random X inactivation associated with disruption of Xist promoter regulation." Hum Mol Genet 10(6): 581-589.
- Ng, H. H., P. Jeppesen and A. Bird (2000). "Active repression of methylated genes by the chromosomal protein MBD1." Mol Cell Biol 20(4): 1394-1406.
- Ng, H. H., F. Robert, R. A. Young and K. Struhl (2003). "Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity." Mol Cell 11(3): 709-719.
- Ng, H. H., Y. Zhang, B. Hendrich, C. A. Johnson, B. M. Turner, H. Erdjument-Bromage, P. Tempst, D. Reinberg and A. Bird (1999). "MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex." Nat Genet 23(1): 58-61.
- Nguyen, D. K. and C. M. Disteche (2006). "Dosage compensation of the active X chromosome in mammals." Nat Genet 38(1): 47-53.
- Nicodemi, M. and A. Prisco (2007). "Symmetry-breaking model for X-chromosome inactivation." Phys Rev Lett 98(10): 108104.
- Nishiyama, M., K. Oshikawa, Y. Tsukada, T. Nakagawa, S. Iemura, T. Natsume, Y. Fan, A. Kikuchi, A. I. Skoultschi and K. I. Nakayama (2009). "CHD8 suppresses p53-mediated apoptosis through histone H1 recruitment during early embryogenesis." Nat Cell Biol 11(2): 172-182.
- Noma, K., H. P. Cam, R. J. Maraia and S. I. Grewal (2006). "A role for TFIIC transcription factor complex in genome organization." Cell 125(5): 859-872.
- Nusinow, D. A., I. Hernandez-Munoz, T. G. Fazio, G. M. Shah, W. L. Kraus and B. Panning (2007). "Poly(ADP-ribose) polymerase 1 is inhibited by a histone H2A variant, MacroH2A, and contributes to silencing of the inactive X chromosome." J Biol Chem 282(17): 12851-12859.
- O'Connor, M. J., S. H. Tan, C. H. Tan and H. U. Bernard (1996). "YY1 represses human papillomavirus type 16 transcription by quenching AP-1 activity." J Virol 70(10): 6529-6539.
- Oei, S. L. and Y. Shi (2001). "Poly(ADP-ribosylation) of transcription factor Yin Yang 1 under conditions of DNA damage." Biochem Biophys Res Commun 285(1): 27-31.
- Ogawa, Y. and J. T. Lee (2003). "Xite, X-inactivation intergenic transcription elements that regulate the probability of choice." Mol Cell 11(3): 731-743.
- Ogawa, Y., B. K. Sun and J. T. Lee (2008). "Intersection of the RNA interference and X-inactivation pathways." Science 320(5881): 1336-1341.
- Ohhata, T., Y. Hoki, H. Sasaki and T. Sado (2006). "Tsix-deficient X chromosome does not undergo inactivation in the embryonic lineage in males: implications for Tsix-independent silencing of Xist." Cytogenet Genome Res 113(1-4): 345-349.
- Ohhata, T., Y. Hoki, H. Sasaki and T. Sado (2008). "Crucial role of antisense transcription

- across the Xist promoter in Tsix-mediated Xist chromatin modification." Development 135(2): 227-235.
- Ohlsson, R., V. Lobanenko and E. Klenova (2010). "Does CTCF mediate between nuclear organization and gene expression?" Bioessays 32(1): 37-50.
- Ohno, S., W. Becak and M. L. Becak (1964). "X-Autosome Ratio and the Behavior Pattern of Individual X-Chromosomes in Placental Mammals." Chromosoma 15: 14-30.
- Ohno, S. and T. S. Hauschka (1960). "Allocycly of the X-chromosome in tumors and normal tissues." Cancer Res 20: 541-545.
- Ohno, S., W. D. Kaplan and R. Kinosita (1959). "Formation of the sex chromatin by a single X-chromosome in liver cells of Rattus norvegicus." Exp Cell Res 18: 415-418.
- Okamoto, I., D. Arnaud, P. Le Baccon, A. P. Otte, C. M. Disteche, P. Avner and E. Heard (2005). "Evidence for de novo imprinted X-chromosome inactivation independent of meiotic inactivation in mice." Nature 438(7066): 369-373.
- Okamoto, I. and E. Heard (2006). "The dynamics of imprinted X inactivation during preimplantation development in mice." Cytogenet Genome Res 113(1-4): 318-324.
- Okamoto, I., A. P. Otte, C. D. Allis, D. Reinberg and E. Heard (2004). "Epigenetic dynamics of imprinted X inactivation during early mouse development." Science 303(5658): 644-649.
- Okamoto, I., S. Tan and N. Takagi (2000). "X-chromosome inactivation in XX androgenetic mouse embryos surviving implantation." Development 127(19): 4137-4145.
- Okano, M., D. W. Bell, D. A. Haber and E. Li (1999). "DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development." Cell 99(3): 247-257.
- Okano, M., S. Xie and E. Li (1998). "Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells." Nucleic Acids Res 26(11): 2536-2540.
- O'Keefe, R. T., S. C. Henderson and D. L. Spector (1992). "Dynamic organization of DNA replication in mammalian cell nuclei: spatially and temporally defined replication of chromosome-specific alpha-satellite DNA sequences." J Cell Biol 116(5): 1095-1110.
- O'Neill, L. P., A. M. Keohane, J. S. Lavender, V. McCabe, E. Heard, P. Avner, N. Brockdorff and B. M. Turner (1999). "A developmental switch in H4 acetylation upstream of Xist plays a role in X chromosome inactivation." EMBO J 18(10): 2897-2907.
- Ottaviani, A., S. Rival-Gervier, A. Boussouar, A. M. Foerster, D. Rondier, S. Sacconi, C. Desnuelle, E. Gilson and F. Magdinier (2009). "The D4Z4 macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy." PLoS Genet 5(2): e1000394.
- Oumard, A., J. Qiao, T. Jostock, J. Li and J. Bode (2006). "Recommended Method for Chromosome Exploitation: RMCE-based Cassette-exchange Systems in Animal Cell Biotechnology." Cytotechnology 50(1-3): 93-108.
- Ouyang, Y., J. Salstrom, S. Diaz-Perez, S. Nahas, Y. Matsuno, D. Dawson, M. A. Teitell, S. Horvath, A. D. Riggs, R. A. Gatti and Y. Marahrens (2005). "Inhibition of Atm and/or Atr disrupts gene silencing on the inactive X chromosome." Biochem Biophys Res Commun 337(3): 875-880.
- Ozcan, S., S. S. Andrali and J. E. Cantrell (2010). "Modulation of transcription factor function by O-GlcNAc modification." Biochim Biophys Acta 1799(5-6): 353-364.
- Pandey, R. R., T. Mondal, F. Mohammad, S. Enroth, L. Redrup, J. Komorowski, T. Nagano, D. Mancini-Dinardo and C. Kanduri (2008). "Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation." Mol Cell 32(2): 232-246.
- Panning, B., J. Dausman and R. Jaenisch (1997). "X chromosome inactivation is mediated by

- Xist RNA stabilization." *Cell* 90(5): 907-916.
- Pant, V., S. Kurukuti, E. Pugacheva, S. Shamsuddin, P. Mariano, R. Renkawitz, E. Klenova, V. Lobanenko and R. Ohlsson (2004). "Mutation of a single CTCF target site within the H19 imprinting control region leads to loss of Igf2 imprinting and complex patterns of de novo methylation upon maternal inheritance." *Mol Cell Biol* 24(8): 3497-3504.
- Parelho, V., S. Hadjur, M. Spivakov, M. Leleu, S. Sauer, H. C. Gregson, A. Jarmuz, C. Canzonetta, Z. Webster, T. Nesterova, B. S. Cobb, K. Yokomori, N. Dillon, L. Aragon, A. G. Fisher and M. Merkenschlager (2008). "Cohesins functionally associate with CTCF on mammalian chromosome arms." *Cell* 132(3): 422-433.
- Pasini, D., K. H. Hansen, J. Christensen, K. Agger, P. A. Cloos and K. Helin (2008). "Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2." *Genes Dev* 22(10): 1345-1355.
- Patrat, C., I. Okamoto, P. Diabangouaya, V. Vialon, P. Le Baccon, J. Chow and E. Heard (2009). "Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice." *Proc Natl Acad Sci U S A* 106(13): 5198-5203.
- Payer, B. and J. T. Lee (2008). "X chromosome dosage compensation: how mammals keep the balance." *Annu Rev Genet* 42: 733-772.
- Penny, G. D., G. F. Kay, S. A. Sheardown, S. Rastan and N. Brockdorff (1996). "Requirement for Xist in X chromosome inactivation." *Nature* 379(6561): 131-137.
- Petkov, S. G., W. A. Reh and G. B. Anderson (2009). "Methylation changes in porcine primordial germ cells." *Mol Reprod Dev* 76(1): 22-30.
- Petkova, V., M. J. Romanowski, I. Suljoadikusumo, D. Rohne, P. Kang, T. Shenk and A. Usheva (2001). "Interaction between YY1 and the retinoblastoma protein. Regulation of cell cycle progression in differentiated cells." *J Biol Chem* 276(11): 7932-7936.
- Phillips, J. E. and V. G. Corces (2009). "CTCF: master weaver of the genome." *Cell* 137(7): 1194-1211.
- Pillet, N., C. Bonny and D. F. Schorderet (1995). "Characterization of the promoter region of the mouse Xist gene." *Proc Natl Acad Sci U S A* 92(26): 12515-12519.
- Pinskaya, M., S. Gourvennec and A. Morillon (2009). "H3 lysine 4 di- and tri-methylation deposited by cryptic transcription attenuates promoter activation." *EMBO J* 28(12): 1697-1707.
- Plath, K., J. Fang, S. K. Mlynarczyk-Evans, R. Cao, K. A. Worringer, H. Wang, C. C. de la Cruz, A. P. Otte, B. Panning and Y. Zhang (2003). "Role of histone H3 lysine 27 methylation in X inactivation." *Science* 300(5616): 131-135.
- Plath, K., S. Mlynarczyk-Evans, D. A. Nusinow and B. Panning (2002). "Xist RNA and the mechanism of X chromosome inactivation." *Annu Rev Genet* 36: 233-278.
- Prissette, M., O. El-Maarri, D. Arnaud, J. Walter and P. Avner (2001). "Methylation profiles of DXPas34 during the onset of X-inactivation." *Hum Mol Genet* 10(1): 31-38.
- Prothero, K. E., J. M. Stahl and L. Carrel (2009). "Dosage compensation and gene expression on the mammalian X chromosome: one plus one does not always equal two." *Chromosome Res* 17(5): 637-648.
- Pugacheva, E. M., Y. W. Kwon, N. A. Hukriede, S. Pack, P. T. Flanagan, J. C. Ahn, J. A. Park, K. S. Choi, K. W. Kim, D. Loukinov, I. B. Dawid and V. V. Lobanenko (2006). "Cloning and characterization of zebrafish CTCF: Developmental expression patterns, regulation of the promoter region, and evolutionary aspects of gene organization." *Gene* 375: 26-36.
- Pugacheva, E. M., V. K. Tiwari, Z. Abdullaev, A. A. Vostrov, P. T. Flanagan, W. W. Quitschke, D. I. Loukinov, R. Ohlsson and V. V. Lobanenko (2005). "Familial cases of point mutations in the XIST promoter reveal a correlation between CTCF binding

- and pre-emptive choices of X chromosome inactivation." Hum Mol Genet 14(7): 953-965.
- Quitschke, W. W., M. J. Taheny, L. J. Fochtmann and A. A. Vostrov (2000). "Differential effect of zinc finger deletions on the binding of CTCF to the promoter of the amyloid precursor protein gene." Nucleic Acids Res 28(17): 3370-3378.
- Rakha, E. A., S. E. Pinder, C. E. Paish and I. O. Ellis (2004). "Expression of the transcription factor CTCF in invasive breast cancer: a candidate gene located at 16q22.1." Br J Cancer 91(8): 1591-1596.
- Rasmussen, T. P., M. A. Mastrangelo, A. Eden, J. R. Pehrson and R. Jaenisch (2000). "Dynamic relocation of histone MacroH2A1 from centrosomes to inactive X chromosomes during X inactivation." J Cell Biol 150(5): 1189-1198.
- Rastan, S. (1983). "Non-random X-chromosome inactivation in mouse X-autosome translocation embryos--location of the inactivation centre." J Embryol Exp Morphol 78: 1-22.
- Rastan, S. and E. J. Robertson (1985). "X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation." J Embryol Exp Morphol 90: 379-388.
- Reale, A., G. D. Matteis, G. Galleazzi, M. Zampieri and P. Caiafa (2005). "Modulation of DNMT1 activity by ADP-ribose polymers." Oncogene 24(1): 13-19.
- Recillas-Targa, F. (2002). "DNA methylation, chromatin boundaries, and mechanisms of genomic imprinting." Arch Med Res 33(5): 428-438.
- Reik, W., W. Dean and J. Walter (2001). "Epigenetic reprogramming in mammalian development." Science 293(5532): 1089-1093.
- Resnick, J. L., L. S. Bixler, L. Cheng and P. J. Donovan (1992). "Long-term proliferation of mouse primordial germ cells in culture." Nature 359(6395): 550-551.
- Rezai-Zadeh, N., X. Zhang, F. Namour, G. Fejer, Y. D. Wen, Y. L. Yao, I. Gyory, K. Wright and E. Seto (2003). "Targeted recruitment of a histone H4-specific methyltransferase by the transcription factor YY1." Genes Dev 17(8): 1019-1029.
- Ribeiro de Almeida, C., H. Heath, S. Krpic, G. M. Dingjan, J. P. van Hamburg, I. Bergen, S. van de Nobelen, F. Sleutels, F. Grosveld, N. Galjart and R. W. Hendriks (2009). "Critical role for the transcription regulator CCCTC-binding factor in the control of Th2 cytokine expression." J Immunol 182(2): 999-1010.
- Riggs, A. D., G. Reiness and G. Zubay (1971). "Purification and DNA-binding properties of the catabolite gene activator protein." Proc Natl Acad Sci U S A 68(6): 1222-1225.
- Rincon-Arano, H., V. Valadez-Graham, G. Guerrero, M. Escamilla-Del-Arenal and F. Recillas-Targa (2005). "YY1 and GATA-1 interaction modulate the chicken 3'-side alpha-globin enhancer activity." J Mol Biol 349(5): 961-975.
- Rinn, J. L., M. Kertesz, J. K. Wang, S. L. Squazzo, X. Xu, S. A. Brugmann, L. H. Goodnough, J. A. Helms, P. J. Farnham, E. Segal and H. Y. Chang (2007). "Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs." Cell 129(7): 1311-1323.
- Rippe, K. (2001). "Making contacts on a nucleic acid polymer." Trends Biochem Sci 26(12): 733-740.
- Robertson, K. D., S. Ait-Si-Ali, T. Yokochi, P. A. Wade, P. L. Jones and A. P. Wolffe (2000). "DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters." Nat Genet 25(3): 338-342.
- Rodriguez, C., J. Borgel, F. Court, G. Cathala, T. Forne and J. Piette (2010). "CTCF is a DNA methylation-sensitive positive regulator of the INK/ARF locus." Biochem Biophys Res Commun 392(2): 129-134.
- Rogner, U. C., D. D. Spyropoulos, N. Le Novere, J. P. Changeux and P. Avner (2000).

"Control of neurulation by the nucleosome assembly protein-1-like 2." Nat Genet 25(4): 431-435.

Ross, M. T., D. V. Grafham, A. J. Coffey, S. Scherer, K. McLay, D. Muzny, M. Platzer, G. R. Howell, C. Burrows, C. P. Bird, A. Frankish, F. L. Lovell, K. L. Howe, J. L. Ashurst, R. S. Fulton, R. Sudbrak, G. Wen, M. C. Jones, M. E. Hurles, T. D. Andrews, C. E. Scott, S. Searle, J. Ramser, A. Whittaker, R. Deadman, N. P. Carter, S. E. Hunt, R. Chen, A. Cree, P. Gunaratne, P. Havlak, A. Hodgson, M. L. Metzker, S. Richards, G. Scott, D. Steffen, E. Sodergren, D. A. Wheeler, K. C. Worley, R. Ainscough, K. D. Ambrose, M. A. Ansari-Lari, S. Aradhya, R. I. Ashwell, A. K. Babbage, C. L. Bagguley, A. Ballabio, R. Banerjee, G. E. Barker, K. F. Barlow, I. P. Barrett, K. N. Bates, D. M. Beare, H. Beasley, O. Beasley, A. Beck, G. Bethel, K. Blechschmidt, N. Brady, S. Bray-Allen, A. M. Bridgeman, A. J. Brown, M. J. Brown, D. Bonnin, E. A. Bruford, C. Buhay, P. Burch, D. Burford, J. Burgess, W. Burrill, J. Burton, J. M. Bye, C. Carder, L. Carrel, J. Chako, J. C. Chapman, D. Chavez, E. Chen, G. Chen, Y. Chen, Z. Chen, C. Chinault, A. Ciccodicola, S. Y. Clark, G. Clarke, C. M. Clee, S. Clegg, K. Clerc-Blankenburg, K. Clifford, V. Cobley, C. G. Cole, J. S. Conquer, N. Corby, R. E. Connor, R. David, J. Davies, C. Davis, J. Davis, O. Delgado, D. Deshazo, P. Dhami, Y. Ding, H. Dinh, S. Dodsworth, H. Draper, S. Dugan-Rocha, A. Dunham, M. Dunn, K. J. Durbin, I. Dutta, T. Eades, M. Ellwood, A. Emery-Cohen, H. Errington, K. L. Evans, L. Faulkner, F. Francis, J. Frankland, A. E. Fraser, P. Galgoczy, J. Gilbert, R. Gill, G. Glockner, S. G. Gregory, S. Gribble, C. Griffiths, R. Grocock, Y. Gu, R. Gwilliam, C. Hamilton, E. A. Hart, A. Hawes, P. D. Heath, K. Heitmann, S. Hennig, J. Hernandez, B. Hinzmann, S. Ho, M. Hoffs, P. J. Howden, E. J. Huckle, J. Hume, P. J. Hunt, A. R. Hunt, J. Isherwood, L. Jacob, D. Johnson, S. Jones, P. J. de Jong, S. S. Joseph, S. Keenan, S. Kelly, J. K. Kershaw, Z. Khan, P. Kioschis, S. Klages, A. J. Knights, A. Kosiura, C. Kovar-Smith, G. K. Laird, C. Langford, S. Lawlor, M. Leversha, L. Lewis, W. Liu, C. Lloyd, D. M. Lloyd, H. Louseged, J. E. Loveland, J. D. Lovell, R. Lozado, J. Lu, R. Lyne, J. Ma, M. Maheshwari, L. H. Matthews, J. McDowall, S. McLaren, A. McMurray, P. Meidl, T. Meitinger, S. Milne, G. Miner, S. L. Mistry, M. Morgan, S. Morris, I. Muller, J. C. Mullikin, N. Nguyen, G. Nordsiek, G. Nyakatura, C. N. O'Dell, G. Okwuonu, S. Palmer, R. Pandian, D. Parker, J. Parrish, S. Pasternak, D. Patel, A. V. Pearce, D. M. Pearson, S. E. Pelan, L. Perez, K. M. Porter, Y. Ramsey, K. Reichwald, S. Rhodes, K. A. Ridler, D. Schlessinger, M. G. Schueler, H. K. Sehra, C. Shaw-Smith, H. Shen, E. M. Sheridan, R. Shownkeen, C. D. Skuce, M. L. Smith, E. C. Sotheran, H. E. Steingruber, C. A. Steward, R. Storey, R. M. Swann, D. Swarbreck, P. E. Tabor, S. Taudien, T. Taylor, B. Teague, K. Thomas, A. Thorpe, K. Timms, A. Tracey, S. Trevanion, A. C. Tromans, M. d'Urso, D. Verduzco, D. Villasana, L. Waldron, M. Wall, Q. Wang, J. Warren, G. L. Warry, X. Wei, A. West, S. L. Whitehead, M. N. Whiteley, J. E. Wilkinson, D. L. Willey, G. Williams, L. Williams, A. Williamson, H. Williamson, L. Wilming, R. L. Woodmansey, P. W. Wray, J. Yen, J. Zhang, J. Zhou, H. Zoghbi, S. Zorilla, D. Buck, R. Reinhardt, A. Poustka, A. Rosenthal, H. Lehrach, A. Meindl, P. J. Minx, L. W. Hillier, H. F. Willard, R. K. Wilson, R. H. Waterston, C. M. Rice, M. Vaudin, A. Coulson, D. L. Nelson, G. Weinstock, J. E. Sulston, R. Durbin, T. Hubbard, R. A. Gibbs, S. Beck, J. Rogers and D. R. Bentley (2005). "The DNA sequence of the human X chromosome." Nature 434(7031): 325-337.

Rougeulle, C. and P. Avner (1996). "Cloning and characterization of a murine brain specific gene Bpx and its human homologue lying within the Xic candidate region." Hum Mol Genet 5(1): 41-49.

Rougeulle, C. and P. Avner (2003). "Controlling X-inactivation in mammals: what does the

- centre hold?" Semin Cell Dev Biol 14(6): 331-340.
- Rougeulle, C., J. Chaumeil, K. Sarma, C. D. Allis, D. Reinberg, P. Avner and E. Heard (2004). "Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome." Mol Cell Biol 24(12): 5475-5484.
- Rougeulle, C., P. Navarro and P. Avner (2003). "Promoter-restricted H3 Lys 4 di-methylation is an epigenetic mark for monoallelic expression." Hum Mol Genet 12(24): 3343-3348.
- Rountree, M. R., K. E. Bachman and S. B. Baylin (2000). "DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci." Nat Genet 25(3): 269-277.
- Royce-Tolland, M. E., A. A. Andersen, H. R. Koyfman, D. J. Talbot, A. Wutz, I. D. Tonks, G. F. Kay and B. Panning (2010). "The A-repeat links ASF/SF2-dependent Xist RNA processing with random choice during X inactivation." Nat Struct Mol Biol.
- Rubio, E. D., D. J. Reiss, P. L. Welch, C. M. Disteche, G. N. Filippova, N. S. Baliga, R. Aebbersold, J. A. Ranish and A. Krumm (2008). "CTCF physically links cohesin to chromatin." Proc Natl Acad Sci U S A 105(24): 8309-8314.
- Sado, T., M. H. Fenner, S. S. Tan, P. Tam, T. Shioda and E. Li (2000). "X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation." Dev Biol 225(2): 294-303.
- Sado, T., Y. Hoki and H. Sasaki (2005). "Tsix silences Xist through modification of chromatin structure." Dev Cell 9(1): 159-165.
- Sado, T., Y. Hoki and H. Sasaki (2006). "Tsix defective in splicing is competent to establish Xist silencing." Development 133(24): 4925-4931.
- Sado, T., E. Li and H. Sasaki (2002). "Effect of TSIX disruption on XIST expression in male ES cells." Cytogenet Genome Res 99(1-4): 115-118.
- Sado, T., M. Okano, E. Li and H. Sasaki (2004). "De novo DNA methylation is dispensable for the initiation and propagation of X chromosome inactivation." Development 131(5): 975-982.
- Sado, T., Z. Wang, H. Sasaki and E. Li (2001). "Regulation of imprinted X-chromosome inactivation in mice by Tsix." Development 128(8): 1275-1286.
- Salido, E. C., P. H. Yen, T. K. Mohandas and L. J. Shapiro (1992). "Expression of the X-inactivation-associated gene XIST during spermatogenesis." Nat Genet 2(3): 196-199.
- Sandhu, K. S., C. Shi, M. Sjolinder, Z. Zhao, A. Gondor, L. Liu, V. K. Tiwari, S. Guibert, L. Emilsson, M. P. Imreh and R. Ohlsson (2009). "Nonallelic transvection of multiple imprinted loci is organized by the H19 imprinting control region during germline development." Genes Dev 23(22): 2598-2603.
- Satijn, D. P., K. M. Hamer, J. den Blaauwen and A. P. Otte (2001). "The polycomb group protein EED interacts with YY1, and both proteins induce neural tissue in Xenopus embryos." Mol Cell Biol 21(4): 1360-1369.
- Satyamoorthy, K., K. Park, M. L. Atchison and C. C. Howe (1993). "The intracisternal A-particle upstream element interacts with transcription factor YY1 to activate transcription: pleiotropic effects of YY1 on distinct DNA promoter elements." Mol Cell Biol 13(11): 6621-6628.
- Sauer, B. (1987). "Functional expression of the cre-lox site-specific recombination system in the yeast *Saccharomyces cerevisiae*." Mol Cell Biol 7(6): 2087-2096.
- Savarese, F., K. Flahndorfer, R. Jaenisch, M. Busslinger and A. Wutz (2006). "Hematopoietic precursor cells transiently reestablish permissiveness for X inactivation." Mol Cell Biol 26(19): 7167-7177.
- Schmidt, D., P. C. Schwalie, C. S. Ross-Innes, A. Hurtado, G. D. Brown, J. S. Carroll, P. Flicek and D. T. Odom (2010). "A CTCF-independent role for cohesin in tissue-

- specific transcription." *Genome Res* 20(5): 578-588.
- Schmidt, D. R. and S. L. Schreiber (1999). "Molecular association between ATR and two components of the nucleosome remodeling and deacetylating complex, HDAC2 and CHD4." *Biochemistry* 38(44): 14711-14717.
- Schoeftner, S., A. K. Sengupta, S. Kubicek, K. Mechtler, L. Spahn, H. Koseki, T. Jenuwein and A. Wutz (2006). "Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing." *EMBO J* 25(13): 3110-3122.
- Schoenherr, C. J., J. M. Levorso and S. M. Tilghman (2003). "CTCF maintains differential methylation at the Igf2/H19 locus." *Nat Genet* 33(1): 66-69.
- Schug, J., W. P. Schuller, C. Kappen, J. M. Salbaum, M. Bucan and C. J. Stoeckert, Jr. (2005). "Promoter features related to tissue specificity as measured by Shannon entropy." *Genome Biol* 6(4): R33.
- Seto, E., B. Lewis and T. Shenk (1993). "Interaction between transcription factors Sp1 and YY1." *Nature* 365(6445): 462-464.
- Shannon, M. F. (2003). "A nuclear address with influence." *Nat Genet* 34(1): 4-6.
- Shao, C. and N. Takagi (1990). "An extra maternally derived X chromosome is deleterious to early mouse development." *Development* 110(3): 969-975.
- Sharman, G. B. (1971). "Late DNA replication in the paternally derived X chromosome of female kangaroos." *Nature* 230(5291): 231-232.
- Sharova, L. V., A. A. Sharov, Y. Piao, N. Shaik, T. Sullivan, C. L. Stewart, B. L. Hogan and M. S. Ko (2007). "Global gene expression profiling reveals similarities and differences among mouse pluripotent stem cells of different origins and strains." *Dev Biol* 307(2): 446-459.
- Sheardown, S. A., S. M. Duthie, C. M. Johnston, A. E. Newall, E. J. Formstone, R. M. Arkell, T. B. Nesterova, G. C. Alghisi, S. Rastan and N. Brockdorff (1997). "Stabilization of Xist RNA mediates initiation of X chromosome inactivation." *Cell* 91(1): 99-107.
- Sheardown, S. A., A. E. Newall, D. P. Norris, S. Rastan and N. Brockdorff (1997). "Regulatory elements in the minimal promoter region of the mouse Xist gene." *Gene* 203(2): 159-168.
- Shi, Y., J. S. Lee and K. M. Galvin (1997). "Everything you have ever wanted to know about Yin Yang 1." *Biochim Biophys Acta* 1332(2): F49-66.
- Shi, Y., E. Seto, L. S. Chang and T. Shenk (1991). "Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by adenovirus E1A protein." *Cell* 67(2): 377-388.
- Shibata, S. and J. T. Lee (2003). "Characterization and quantitation of differential Tsix transcripts: implications for Tsix function." *Hum Mol Genet* 12(2): 125-136.
- Shibata, S. and J. T. Lee (2004). "Tsix transcription- versus RNA-based mechanisms in Xist repression and epigenetic choice." *Curr Biol* 14(19): 1747-1754.
- Shibata, S. and T. Yokota (2008). "Alteration of histone tail modifications in the Xist locus in wild-type and Tsix-mutant male embryonic stem cells during differentiation." *Exp Anim* 57(2): 153-157.
- Shibata, S., T. Yokota and A. Wutz (2008). "Synergy of Eed and Tsix in the repression of Xist gene and X-chromosome inactivation." *EMBO J* 27(13): 1816-1826.
- Shovlin, T. C., G. Durcova-Hills, A. Surani and A. McLaren (2008). "Heterogeneity in imprinted methylation patterns of pluripotent embryonic germ cells derived from pre-migratory mouse germ cells." *Dev Biol* 313(2): 674-681.
- Shrivastava, A., S. Saleque, G. V. Kalpana, S. Artandi, S. P. Goff and K. Calame (1993). "Inhibition of transcriptional regulator Yin-Yang-1 by association with c-Myc." *Science* 262(5141): 1889-1892.
- Silva, J., O. Barrandon, J. Nichols, J. Kawaguchi, T. W. Theunissen and A. Smith (2008).

- "Promotion of reprogramming to ground state pluripotency by signal inhibition." PLoS Biol 6(10): e253.
- Silva, J., W. Mak, I. Zvetkova, R. Appanah, T. B. Nesterova, Z. Webster, A. H. Peters, T. Jenuwein, A. P. Otte and N. Brockdorff (2003). "Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes." Dev Cell 4(4): 481-495.
- Simmler, M. C., D. B. Cunningham, P. Clerc, T. Verinat, B. Caudron, C. Cruaud, A. Pawlak, C. Szpirer, J. Weissenbach, J. M. Claverie and P. Avner (1996). "A 94 kb genomic sequence 3' to the murine Xist gene reveals an AT rich region containing a new testis specific gene Tsx." Hum Mol Genet 5(11): 1713-1726.
- Simmler, M. C., E. Heard, C. Rougeulle, C. Cruaud, J. Weissenbach and P. Avner (1997). "Localization and expression analysis of a novel conserved brain expressed transcript, Brx/BRX, lying within the Xic/XIC candidate region." Mamm Genome 8(10): 760-766.
- Simms, T. A., E. C. Miller, N. P. Buisson, N. Jambunathan and D. Donze (2004). "The *Saccharomyces cerevisiae* TRT2 tRNAThr gene upstream of STE6 is a barrier to repression in MATalpha cells and exerts a potential tRNA position effect in MATa cells." Nucleic Acids Res 32(17): 5206-5213.
- Simonis, M., J. Kooren and W. de Laat (2007). "An evaluation of 3C-based methods to capture DNA interactions." Nat Methods 4(11): 895-901.
- Sirbu, I. O., X. Zhao and G. Duester (2008). "Retinoic acid controls heart anteroposterior patterning by down-regulating *Isl1* through the *Fgf8* pathway." Dev Dyn 237(6): 1627-1635.
- Skene, P. J., R. S. Illingworth, S. Webb, A. R. Kerr, K. D. James, D. J. Turner, R. Andrews and A. P. Bird (2010). "Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state." Mol Cell 37(4): 457-468.
- Smale, S. T., M. C. Schmidt, A. J. Berk and D. Baltimore (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(12): 4509-4513.
- Smith, E. R., C. D. Allis and J. C. Lucchesi (2001). "Linking global histone acetylation to the transcription enhancement of X-chromosomal genes in *Drosophila* males." J Biol Chem 276(34): 31483-31486.
- Smith, K. P., M. Byron, C. M. Clemson and J. B. Lawrence (2004). "Ubiquitinated proteins including uH2A on the human and mouse inactive X chromosome: enrichment in gene rich bands." Chromosoma 113(6): 324-335.
- Splinter, E., H. Heath, J. Kooren, R. J. Palstra, P. Klous, F. Grosveld, N. Galjart and W. de Laat (2006). "CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus." Genes Dev 20(17): 2349-2354.
- Stadtfield, M., N. Maherali, D. T. Breault and K. Hochedlinger (2008). "Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse." Cell Stem Cell 2(3): 230-240.
- Starmer, J. and T. Magnuson (2009). "A new model for random X chromosome inactivation." Development 136(1): 1-10.
- Stavropoulos, N., R. K. Rowntree and J. T. Lee (2005). "Identification of developmentally specific enhancers for *Tsix* in the regulation of X chromosome inactivation." Mol Cell Biol 25(7): 2757-2769.
- Stedman, W., H. Kang, S. Lin, J. L. Kissil, M. S. Bartolomei and P. M. Lieberman (2008). "Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators." EMBO J 27(4): 654-666.
- Stewart, C. L., I. Gadi and H. Bhatt (1994). "Stem cells from primordial germ cells can

- reenter the germ line." Dev Biol 161(2): 626-628.
- Stock, J. K., S. Giadrossi, M. Casanova, E. Brookes, M. Vidal, H. Koseki, N. Brockdorff, A. G. Fisher and A. Pombo (2007). "Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells." Nat Cell Biol 9(12): 1428-1435.
- Stojanova, A., C. Caro, R. J. Jarjour, S. K. Oster, L. Z. Penn and R. J. Germinario (2004). "Repression of the human immunodeficiency virus type-1 long terminal repeat by the c-Myc oncoprotein." J Cell Biochem 92(2): 400-413.
- Strahl, B. D. and C. D. Allis (2000). "The language of covalent histone modifications." Nature 403(6765): 41-45.
- Sugimoto, M. and K. Abe (2007). "X chromosome reactivation initiates in nascent primordial germ cells in mice." PLoS Genet 3(7): e116.
- Sun, B. K., A. M. Deaton and J. T. Lee (2006). "A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization." Mol Cell 21(5): 617-628.
- Surani, M. A. (2007). "Germ cells: the eternal link between generations." C R Biol 330(6-7): 474-478.
- Szabo, P. E. and J. R. Mann (1995). "Biallelic expression of imprinted genes in the mouse germ line: implications for erasure, establishment, and mechanisms of genomic imprinting." Genes Dev 9(15): 1857-1868.
- Tada, T., Y. Obata, M. Tada, Y. Goto, N. Nakatsuji, S. Tan, T. Kono and N. Takagi (2000). "Imprint switching for non-random X-chromosome inactivation during mouse oocyte growth." Development 127(14): 3101-3105.
- Tada, T., M. Tada, K. Hilton, S. C. Barton, T. Sado, N. Takagi and M. A. Surani (1998). "Epigenotype switching of imprintable loci in embryonic germ cells." Dev Genes Evol 207(8): 551-561.
- Takagi, N. and M. Sasaki (1974). "A phylogenetic study of bird karyotypes." Chromosoma 46(1): 91-120.
- Takagi, N. and M. Sasaki (1975). "Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse." Nature 256(5519): 640-642.
- Takagi, N., M. A. Yoshida, O. Sugawara and M. Sasaki (1983). "Reversal of X-inactivation in female mouse somatic cells hybridized with murine teratocarcinoma stem cells in vitro." Cell 34(3): 1053-1062.
- Takao, Y., T. Yokota and H. Koide (2007). "Beta-catenin up-regulates Nanog expression through interaction with Oct-3/4 in embryonic stem cells." Biochem Biophys Res Commun 353(3): 699-705.
- Tam, P. P., S. X. Zhou and S. S. Tan (1994). "X-chromosome activity of the mouse primordial germ cells revealed by the expression of an X-linked lacZ transgene." Development 120(10): 2925-2932.
- Tan, S. S., E. A. Williams and P. P. Tam (1993). "X-chromosome inactivation occurs at different times in different tissues of the post-implantation mouse embryo." Nat Genet 3(2): 170-174.
- Tanaka, M., Y. Shigetani, S. Sugiyama, K. Tamura, H. Nakamura and H. Ide (1998). "Apical ectodermal ridge induction by the transplantation of En-1-overexpressing ectoderm in chick limb bud." Dev Growth Differ 40(4): 423-429.
- Taylor, J. H. (1968). "Rates of chain growth and units of replication in DNA of mammalian chromosomes." J Mol Biol 31(3): 579-594.
- Taylor, J. H. and P. Miner (1968). "Units of DNA replication in mammalian chromosomes." Cancer Res 28(9): 1810-1814.
- Terranova, R., S. Yokobayashi, M. B. Stadler, A. P. Otte, M. van Lohuizen, S. H. Orkin and

- A. H. Peters (2008). "Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos." Dev Cell 15(5): 668-679.
- Thomas, M. J. and E. Seto (1999). "Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key?" Gene 236(2): 197-208.
- Thomson, J. P., P. J. Skene, J. Selfridge, T. Clouaire, J. Guy, S. Webb, A. R. Kerr, A. Deaton, R. Andrews, K. D. James, D. J. Turner, R. Illingworth and A. Bird (2010). "CpG islands influence chromatin structure via the CpG-binding protein Cfp1." Nature 464(7291): 1082-1086.
- Tjeertes, J. V., K. M. Miller and S. P. Jackson (2009). "Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells." EMBO J 28(13): 1878-1889.
- Torchia, B. S., L. M. Call and B. R. Migeon (1994). "DNA replication analysis of FMR1, XIST, and factor 8C loci by FISH shows nontranscribed X-linked genes replicate late." Am J Hum Genet 55(1): 96-104.
- Torrano, V., J. Navascues, F. Docquier, R. Zhang, L. J. Burke, I. Chernukhin, D. Farrar, J. Leon, M. T. Berciano, R. Renkawitz, E. Klenova, M. Lafarga and M. D. Delgado (2006). "Targeting of CTCF to the nucleolus inhibits nucleolar transcription through a poly(ADP-ribosyl)ation-dependent mechanism." J Cell Sci 119(Pt 9): 1746-1759.
- Tsai, C. L., R. K. Rowntree, D. E. Cohen and J. T. Lee (2008). "Higher order chromatin structure at the X-inactivation center via looping DNA." Dev Biol 319(2): 416-425.
- Tsang, T. E., P. L. Khoo, R. V. Jamieson, S. X. Zhou, S. L. Ang, R. Behringer and P. P. Tam (2001). "The allocation and differentiation of mouse primordial germ cells." Int J Dev Biol 45(3): 549-555.
- Turner, B. (2007). "Diagnosis and treatment of patients with prostate cancer: the nurse's role." Nurs Stand 21(39): 48-56; quiz 58.
- Turner, B. J., J. D. Atkin, M. A. Farg, D. W. Zang, A. Rembach, E. C. Lopes, J. D. Patch, A. F. Hill and S. S. Cheema (2005). "Impaired extracellular secretion of mutant superoxide dismutase 1 associates with neurotoxicity in familial amyotrophic lateral sclerosis." J Neurosci 25(1): 108-117.
- Turner, B. M. (2002). "Cellular memory and the histone code." Cell 111(3): 285-291.
- Turner, B. M. (2005). "Reading signals on the nucleosome with a new nomenclature for modified histones." Nat Struct Mol Biol 12(2): 110-112.
- Turner, B. M. (2007). "Defining an epigenetic code." Nat Cell Biol 9(1): 2-6.
- Turner, J. M., S. K. Mahadevaiah, D. J. Elliott, H. J. Garchon, J. R. Pehrson, R. Jaenisch and P. S. Burgoyne (2002). "Meiotic sex chromosome inactivation in male mice with targeted disruptions of Xist." J Cell Sci 115(Pt 21): 4097-4105.
- Usheva, A. and T. Shenk (1994). "TATA-binding protein-independent initiation: YY1, TFIIB, and RNA polymerase II direct basal transcription on supercoiled template DNA." Cell 76(6): 1115-1121.
- Verani, R., I. Cappuccio, P. Spinsanti, R. Gradini, A. Caruso, M. C. Magnotti, M. Motolese, F. Nicoletti and D. Melchiorri (2007). "Expression of the Wnt inhibitor Dickkopf-1 is required for the induction of neural markers in mouse embryonic stem cells differentiating in response to retinoic acid." J Neurochem 100(1): 242-250.
- Verdel, A., S. Jia, S. Gerber, T. Sugiyama, S. Gygi, S. I. Grewal and D. Moazed (2004). "RNAi-mediated targeting of heterochromatin by the RITS complex." Science 303(5658): 672-676.
- Vigneau, S., S. Augui, P. Navarro, P. Avner and P. Clerc (2006). "An essential role for the DXPas34 tandem repeat and Tsix transcription in the counting process of X chromosome inactivation." Proc Natl Acad Sci U S A 103(19): 7390-7395.

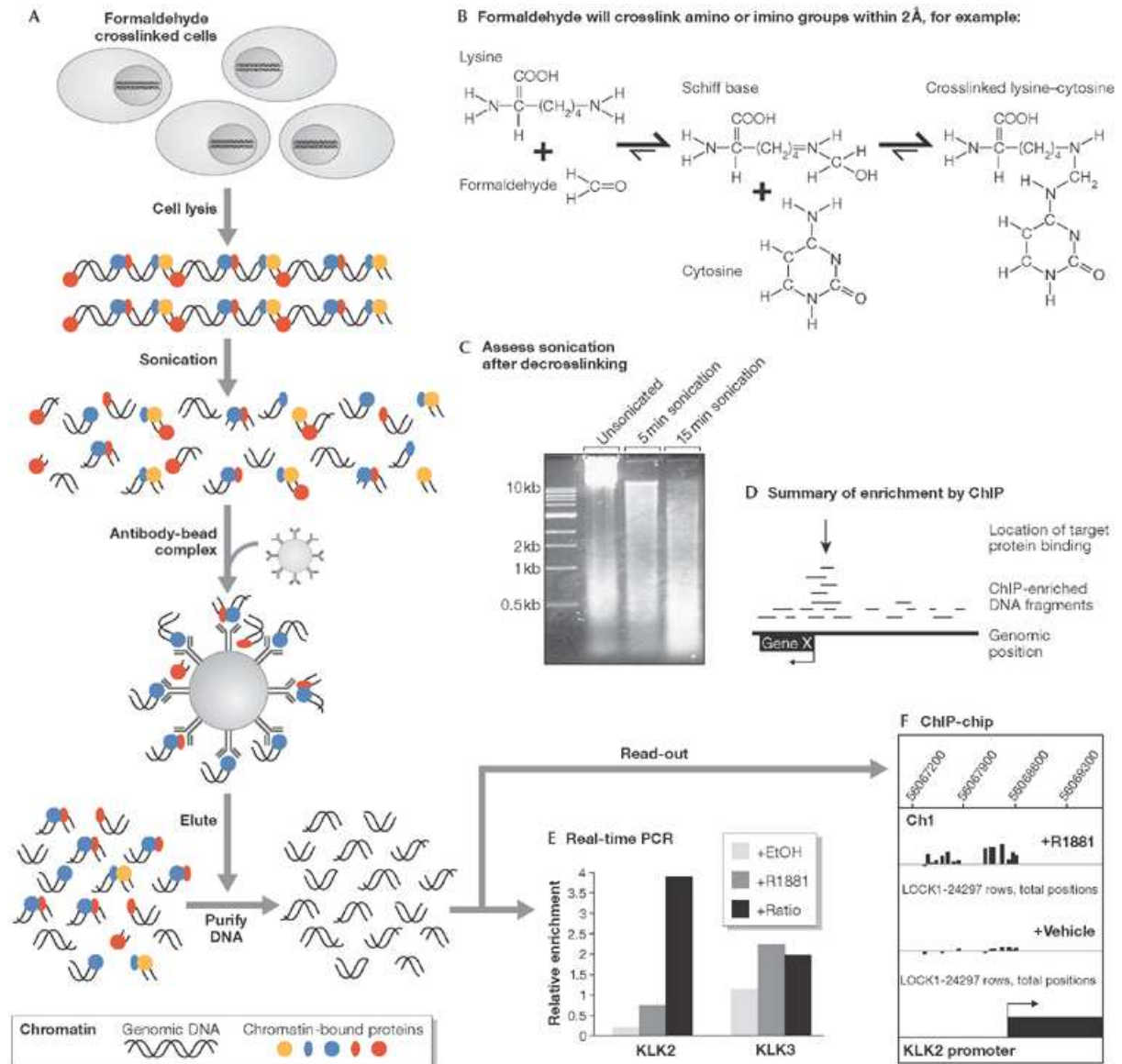
- Volpe, T. A., C. Kidner, I. M. Hall, G. Teng, S. I. Grewal and R. A. Martienssen (2002). "Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi." *Science* 297(5588): 1833-1837.
- Vostrov, A. A., M. J. Taheny and W. W. Quitschke (2002). "A region to the N-terminal side of the CTCF zinc finger domain is essential for activating transcription from the amyloid precursor protein promoter." *J Biol Chem* 277(2): 1619-1627.
- Wallace, J. A. and G. Felsenfeld (2007). "We gather together: insulators and genome organization." *Curr Opin Genet Dev* 17(5): 400-407.
- Wang, C. Y., Y. J. Liang, Y. S. Lin, H. M. Shih, Y. S. Jou and W. C. Yu (2004). "YY1AP, a novel co-activator of YY1." *J Biol Chem* 279(17): 17750-17755.
- Wang, J., J. Mager, Y. Chen, E. Schneider, J. C. Cross, A. Nagy and T. Magnuson (2001). "Imprinted X inactivation maintained by a mouse Polycomb group gene." *Nat Genet* 28(4): 371-375.
- Wang, P. J., D. C. Page and J. R. McCarrey (2005). "Differential expression of sex-linked and autosomal germ-cell-specific genes during spermatogenesis in the mouse." *Hum Mol Genet* 14(19): 2911-2918.
- Wang, Y., D. Cortez, P. Yazdi, N. Neff, S. J. Elledge and J. Qin (2000). "BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures." *Genes Dev* 14(8): 927-939.
- Wang, J., S. Rao, J. Chu, X. Shen, D. N. Levasseur, T. W. Theunissen and S. H. Orkin (2006). "A protein interaction network for pluripotency of embryonic stem cells." *Nature* 444(7117) : 364-8.
- Waterston, R. H., K. Lindblad-Toh, E. Birney, J. Rogers, J. F. Abril, P. Agarwal, R. Agarwala, R. Ainscough, M. Alexandersson, P. An, S. E. Antonarakis, J. Attwood, R. Baertsch, J. Bailey, K. Barlow, S. Beck, E. Berry, B. Birren, T. Bloom, P. Bork, M. Botcherby, N. Bray, M. R. Brent, D. G. Brown, S. D. Brown, C. Bult, J. Burton, J. Butler, R. D. Campbell, P. Carninci, S. Cawley, F. Chiaromonte, A. T. Chinwalla, D. M. Church, M. Clamp, C. Clee, F. S. Collins, L. L. Cook, R. R. Copley, A. Coulson, O. Couronne, J. Cuff, V. Curwen, T. Cutts, M. Daly, R. David, J. Davies, K. D. Delehaunty, J. Deri, E. T. Dermitzakis, C. Dewey, N. J. Dickens, M. Diekhans, S. Dodge, I. Dubchak, D. M. Dunn, S. R. Eddy, L. Elnitski, R. D. Emes, P. Eswara, E. Eyra, A. Felsenfeld, G. A. Fewell, P. Flicek, K. Foley, W. N. Frankel, L. A. Fulton, R. S. Fulton, T. S. Furey, D. Gage, R. A. Gibbs, G. Glusman, S. Gnerre, N. Goldman, L. Goodstadt, D. Grafham, T. A. Graves, E. D. Green, S. Gregory, R. Guigo, M. Guyer, R. C. Hardison, D. Haussler, Y. Hayashizaki, L. W. Hillier, A. Hinrichs, W. Hlavina, T. Holzer, F. Hsu, A. Hua, T. Hubbard, A. Hunt, I. Jackson, D. B. Jaffe, L. S. Johnson, M. Jones, T. A. Jones, A. Joy, M. Kamal, E. K. Karlsson, D. Karolchik, A. Kasprzyk, J. Kawai, E. Keibler, C. Kells, W. J. Kent, A. Kirby, D. L. Kolbe, I. Korf, R. S. Kucherlapati, E. J. Kulbokas, D. Kulp, T. Landers, J. P. Leger, S. Leonard, I. Letunic, R. Levine, J. Li, M. Li, C. Lloyd, S. Lucas, B. Ma, D. R. Maglott, E. R. Mardis, L. Matthews, E. Mauceli, J. H. Mayer, M. McCarthy, W. R. McCombie, S. McLaren, K. McLay, J. D. McPherson, J. Meldrim, B. Meredith, J. P. Mesirov, W. Miller, T. L. Miner, E. Mongin, K. T. Montgomery, M. Morgan, R. Mott, J. C. Mullikin, D. M. Muzny, W. E. Nash, J. O. Nelson, M. N. Nhan, R. Nicol, Z. Ning, C. Nusbaum, M. J. O'Connor, Y. Okazaki, K. Oliver, E. Overton-Larty, L. Pachter, G. Parra, K. H. Pepin, J. Peterson, P. Pevzner, R. Plumb, C. S. Pohl, A. Poliakov, T. C. Ponce, C. P. Ponting, S. Potter, M. Quail, A. Reymond, B. A. Roe, K. M. Roskin, E. M. Rubin, A. G. Rust, R. Santos, V. Sapojnikov, B. Schultz, J. Schultz, M. S. Schwartz, S. Schwartz, C. Scott, S. Seaman, S. Searle, T. Sharpe, A. Sheridan, R. Shownkeen, S. Sims, J. B. Singer, G. Slater, A. Smit, D. R. Smith, B. Spencer, A.

- Stabenau, N. Stange-Thomann, C. Sugnet, M. Suyama, G. Tesler, J. Thompson, D. Torrents, E. Trevaskis, J. Tromp, C. Ucla, A. Ureta-Vidal, J. P. Vinson, A. C. Von Niederhausern, C. M. Wade, M. Wall, R. J. Weber, R. B. Weiss, M. C. Wendl, A. P. West, K. Wetterstrand, R. Wheeler, S. Whelan, J. Wierzbowski, D. Willey, S. Williams, R. K. Wilson, E. Winter, K. C. Worley, D. Wyman, S. Yang, S. P. Yang, E. M. Zdobnov, M. C. Zody and E. S. Lander (2002). "Initial sequencing and comparative analysis of the mouse genome." *Nature* 420(6915): 520-562.
- Wendt, K. S., K. Yoshida, T. Itoh, M. Bando, B. Koch, E. Schirghuber, S. Tsutsumi, G. Nagae, K. Ishihara, T. Mishiro, K. Yahata, F. Imamoto, H. Aburatani, M. Nakao, N. Imamoto, K. Maeshima, K. Shirahige and J. M. Peters (2008). "Cohesin mediates transcriptional insulation by CCCTC-binding factor." *Nature* 451(7180): 796-801.
- Wilkinson, F., H. Pratt and M. L. Atchison (2010). "PcG recruitment by the YY1 REPO domain can be mediated by Yaf2." *J Cell Biochem* 109(3): 478-486.
- Willard, H. F. (1996). "X chromosome inactivation, XIST, and pursuit of the X-inactivation center." *Cell* 86(1): 5-7.
- Wu, S., Y. Shi, P. Mulligan, F. Gay, J. Landry, H. Liu, J. Lu, H. H. Qi, W. Wang, J. A. Nickoloff and C. Wu (2007). "A YY1-INO80 complex regulates genomic stability through homologous recombination-based repair." *Nat Struct Mol Biol* 14(12): 1165-1172.
- Wutz, A. and R. Jaenisch (2000). "A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation." *Mol Cell* 5(4): 695-705.
- Wutz, A., T. P. Rasmussen and R. Jaenisch (2002). "Chromosomal silencing and localization are mediated by different domains of Xist RNA." *Nat Genet* 30(2): 167-174.
- Xu, N., M. E. Donohoe, S. S. Silva and J. T. Lee (2007). "Evidence that homologous X-chromosome pairing requires transcription and Ctf protein." *Nat Genet* 39(11): 1390-1396.
- Xu, N., C. L. Tsai and J. T. Lee (2006). "Transient homologous chromosome pairing marks the onset of X inactivation." *Science* 311(5764): 1149-1152.
- Yamaguchi, S., H. Kimura, M. Tada, N. Nakatsuji and T. Tada (2005). "Nanog expression in mouse germ cell development." *Gene Expr Patterns* 5(5): 639-646.
- Yang, F., T. Babak, J. Shendure and C. M. Disteche (2010). "Global survey of escape from X inactivation by RNA-sequencing in mouse." *Genome Res* 20(5): 614-622.
- Yang, W. M., C. Inouye, Y. Zeng, D. Bearss and E. Seto (1996). "Transcriptional repression by YY1 is mediated by interaction with a mammalian homolog of the yeast global regulator RPD3." *Proc Natl Acad Sci U S A* 93(23): 12845-12850.
- Yang, W. M., Y. L. Yao, J. M. Sun, J. R. Davie and E. Seto (1997). "Isolation and characterization of cDNAs corresponding to an additional member of the human histone deacetylase gene family." *J Biol Chem* 272(44): 28001-28007.
- Yao, Y. L., W. M. Yang and E. Seto (2001). "Regulation of transcription factor YY1 by acetylation and deacetylation." *Mol Cell Biol* 21(17): 5979-5991.
- Yap, K. L., S. Li, A. M. Munoz-Cabello, S. Raguz, L. Zeng, S. Mujtaba, J. Gil, M. J. Walsh and M. M. Zhou (2010). "Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a." *Mol Cell* 38(5): 662-674.
- Yen, P. H., P. Patel, A. C. Chinault, T. Mohandas and L. J. Shapiro (1984). "Differential methylation of hypoxanthine phosphoribosyltransferase genes on active and inactive human X chromosomes." *Proc Natl Acad Sci U S A* 81(6): 1759-1763.
- Yet, S. F., M. M. McA'Nulty, S. C. Folta, H. W. Yen, M. Yoshizumi, C. M. Hsieh, M. D. Layne, M. T. Chin, H. Wang, M. A. Perrella, M. K. Jain and M. E. Lee (1998). "Human EZF, a Kruppel-like zinc finger protein, is expressed in vascular endothelial

- cells and contains transcriptional activation and repression domains." J Biol Chem 273(2): 1026-1031.
- Yu, W., V. Ginjala, V. Pant, I. Chernukhin, J. Whitehead, F. Docquier, D. Farrar, G. Tavoosidana, R. Mukhopadhyay, C. Kanduri, M. Oshimura, A. P. Feinberg, V. Lobanenko, E. Klenova and R. Ohlsson (2004). "Poly(ADP-ribosylation) regulates CTCF-dependent chromatin insulation." Nat Genet 36(10): 1105-1110.
- Yusufzai, T. M. and G. Felsenfeld (2004). "The 5'-HS4 chicken beta-globin insulator is a CTCF-dependent nuclear matrix-associated element." Proc Natl Acad Sci U S A 101(23): 8620-8624.
- Yusufzai, T. M., H. Tagami, Y. Nakatani and G. Felsenfeld (2004). "CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species." Mol Cell 13(2): 291-298.
- Zhang, L. F., K. D. Huynh and J. T. Lee (2007). "Perinucleolar targeting of the inactive X during S phase: evidence for a role in the maintenance of silencing." Cell 129(4): 693-706.
- Zhang, R., L. J. Burke, J. E. Rasko, V. Lobanenko and R. Renkawitz (2004). "Dynamic association of the mammalian insulator protein CTCF with centrosomes and the midbody." Exp Cell Res 294(1): 86-93.
- Zhang, W. and J. J. Bieker (1998). "Acetylation and modulation of erythroid Kruppel-like factor (EKLF) activity by interaction with histone acetyltransferases." Proc Natl Acad Sci U S A 95(17): 9855-9860.
- Zhang, W., J. M. Shields, K. Sogawa, Y. Fujii-Kuriyama and V. W. Yang (1998). "The gut-enriched Kruppel-like factor suppresses the activity of the CYP1A1 promoter in an Sp1-dependent fashion." J Biol Chem 273(28): 17917-17925.
- Zhao, J., B. K. Sun, J. A. Erwin, J. J. Song and J. T. Lee (2008). "Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome." Science 322(5902): 750-756.
- Zheng, H., J. Chu, Y. Zeng, H. H. Loh and P. Y. Law (2010). "Yin Yang 1 phosphorylation contributes to the differential effects of mu-opioid receptor agonists on microRNA-190 expression." J Biol Chem 285(29): 21994-22002.
- Zlatanova, J. and K. van Holde (1998). "Linker histones versus HMG1/2: a struggle for dominance?" Bioessays 20(7): 584-588.

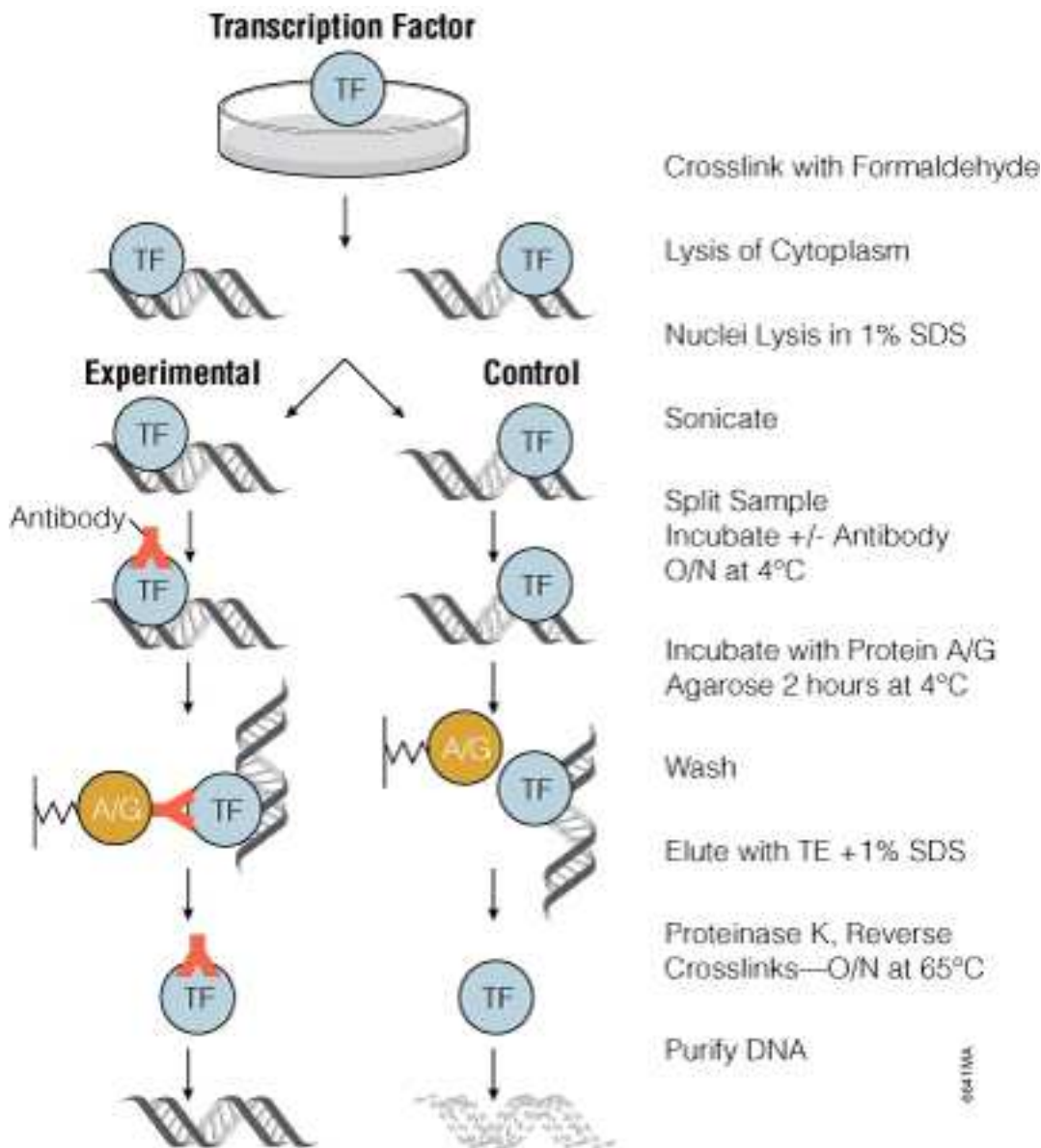
Annexes

Annexe 1 :



Vue d'ensemble de la stratégie d'immunoprécipitation de chromatine. (A) Résumé de la méthodologie d'immunoprécipitation de chromatine. (B) Formation de la liaison protéine-ADN par la foraldéhyde. (C) Exemple d'analyse par gel d'électrophorèse de la chromatine avant et après sonication. (D) Résumé de l'enrichissement de fragments d'ADN par ChIP, montrant les différences d'abondance entre les fragments spécifiques, enrichis par ChIP et les non-spécifiques. (E) Exemple d'analyse de données par PCR-quantitative (F) ou par CHIP-CHIP ou par ChIP-seq. (D'après Massie & Mills, 2008)

Annexe 2 :



Vue d'ensemble de l'immunoprécipitation de chromatine. Les cellules sont cultivées dans des conditions idéales à l'interaction entre le facteur de transcription (FT) d'intérêt et l'ADN. Afin de préserver l'association ADN-protéine au cours de la lyse cellulaire, de la formaldéhyde est ajoutée, engendrant des liaisons entre l'ADN et la protéine. Un extrait total de la cellule est préparé, et l'ADN lié est fragmenté par sonication pour réduire la taille moyenne des fragments d'ADN. Un anticorps qui reconnaît la protéine est ajouté et incubé sur la nuit. Des billes d'agarose protéine A ou protéine G sont ajoutées pour capturer le complexe, puis lavées. L'anticorps doit être spécifique et fortement fixé à la protéine en tenant compte des conditions de lavage utilisées. Enfin, l'inversement de la liaison par chauffage permet l'obtention et l'analyse de l'ADN immunoprécipité. En parallèle, il est nécessaire d'effectuer la même expérience mais sans l'utilisation d'anticorps, ce qu'on appelle Input. L'ADN récupéré de l'input servira à normaliser les données. (<http://www.promega.com/paguide/chap11.htm>)

Annexe 3 :

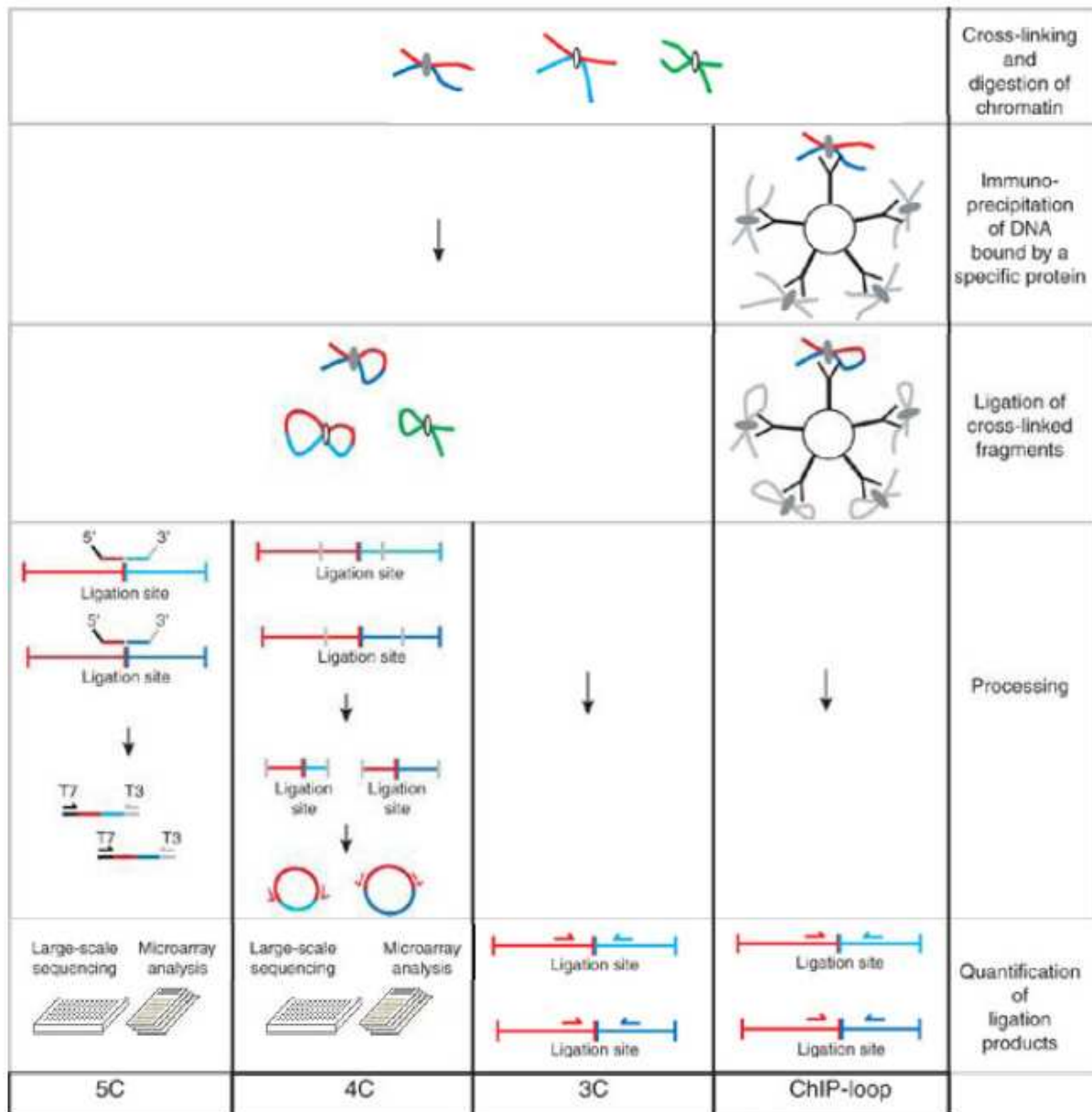


Schéma du protocole de 3C, 4C, 5C. (Simonis *et al.*, 2007)

Résumé :

L'inactivation du chromosome X est la réponse trouvée par l'évolution pour pallier à la divergence gonosomique entre mâle (XY) et femelle (XX). Ce phénomène sert donc à mettre les deux sexes sur un pied d'égalité en limitant la quantité de transcrits provenant des chromosomes X présents dans les cellules femelles.

Au cours de mon doctorat, j'ai tenté de contribuer à l'étude des mécanismes de régulation transcriptionnelle, notamment l'activation, des deux acteurs principaux de l'inactivation: *Xist* (pour « *X-inactive specific transcript* ») et *Tsix*, son transcrit antisens. Ces deux gènes produisent des transcrits non-codants essentiels à une initiation correcte de ce phénomène épigénétique. L'ARN *Xist* "enveloppe" un des deux chromosomes X dans les cellules femelles pour induire son inactivation, et ce de manière extrêmement liée au développement de l'embryon précoce.

Pendant ces 4 années, j'ai entrepris de cartographier à moyenne échelle le profil de fixation de plusieurs protéines le long du locus *Xist/Tsix*, dans le but d'obtenir une meilleure compréhension des mécanismes permettant une surexpression de *Xist* lors de la disparition de ses facteurs répressifs en cours de différenciation. Cette cartographie a été effectuée dans plusieurs lignées sauvages ou mutantes de cellules souches embryonnaires de souris (qui ont comme caractéristique de récapituler l'inactivation du chromosome X *ex-vivo*).

J'ai donc pu établir un modèle de régulation transcriptionnelle de l'ARN non-codant *Xist*, impliquant plusieurs protéines connues pour leur rôle dans la régulation transcriptionnelle (CTCF et YY1) aussi bien que dans la formation de structures tridimensionnelles (la cohésine). La pertinence de ce modèle est renforcée par nos études montrant que de nombreux aspects de ce modèle sont conservés à travers l'évolution (notamment chez l'homme).

J'ai également pu contribuer à la découverte de nouveaux activateurs de *Tsix*, certains facteurs de pluripotence se fixant au minisatellite *DxPas34* afin de réguler l'élongation de la transcription de l'antisens.

Ces résultats apportent donc d'importantes informations concernant les mécanismes régulant la mise en place du phénomène d'inactivation du chromosome X au cours du développement précoce de l'embryon.

Mots clés : Inactivation du chromosome X, épigénétique, régulation transcriptionnelle, structure tridimensionnelle, ARN non-codant.