



**HAL**  
open science

# Architecture chromosomique du locus Xic : implications pour la régulation de l'inactivation du chromosome X

Elphège-Pierre Nora

► **To cite this version:**

Elphège-Pierre Nora. Architecture chromosomique du locus Xic : implications pour la régulation de l'inactivation du chromosome X. Génétique. Université Paris Sud - Paris XI, 2011. Français. NNT : 2011PA112130 . tel-00635540

**HAL Id: tel-00635540**

**<https://theses.hal.science/tel-00635540>**

Submitted on 25 Oct 2011

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

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

## THÈSE DE DOCTORAT

En vue de l'obtention du grade de Docteur de l'Université Paris-Sud XI  
École doctorale Gènes, Génomes, Cellules

Discipline : Génétique

Présentée et soutenue publiquement par

Elphège-Pierre NORA

Le 7 septembre 2011

### **Architecture chromosomique du locus *Xic* : implications pour la régulation de l'inactivation du chromosome X**

Membres du jury

Pr. Pierre CAPY	Président
Dr. Thierry FORNE	Rapporteur
Pr. Reiner VEITIA	Rapporteur
Pr. Wendy BICKMORE	Examineur
Dr. Joost GRIBNAU	Examineur
Dr. Edith HEARD	Directeur de thèse



### ARCHITECTURE CHROMOSOMIQUE DU LOCUS *Xic* :

#### IMPLICATIONS POUR LA RÉGULATION DE L'INACTIVATION DU CHROMOSOME X

---

Le développement embryonnaire précoce des mammifères femelles s'accompagne de l'inactivation transcriptionnelle d'un de leurs deux chromosomes X. Cet évènement est initié suite à l'expression mono-allélique de l'ARN non codant *Xist*, qui est contrôlée par de nombreux éléments *cis*-régulateurs présents dans le *centre d'inactivation du chromosome X (Xic)* – tel son anti-sens répresseur *Tsix*. Mon travail de thèse a consisté à développer des approches permettant d'appréhender le paysage structural dans lequel s'exerce cette régulation. La caractérisation de l'architecture tridimensionnelle du *Xic*, par des techniques basées sur la capture de conformation chromosomique (3C) et l'hybridation *in situ* en fluorescence (FISH), m'a permis de mettre en évidence que les promoteurs respectifs de *Xist* et *Tsix* sont engagés dans des interactions physiques intimes avec des loci distaux, localisés au sein du *Xic*, et de montrer qu'au moins certaines de ces régions exercent un effets régulateurs à longue-distance. Les éléments du *Xic* contactés par les régions promotrices de *Xist* et de *Tsix* sont en outre fondamentalement différents, chacune engageant des associations chromosomiques sur plusieurs centaines de kilobases dans leur direction 5' respective.

Ce travail a également permis de révéler des propriétés insoupçonnées de l'architecture chromosomiques. En effet, le *Xic* apparaît scindé en plusieurs sous-régions, couvrant chacune entre 200kb et 1Mb, à l'intérieur desquelles les interactions chromosomiques sont préférentiellement établies. L'existence de ces domaines d'interaction s'intègre avec d'autres propriétés structurales du génome, tels la composition de la chromatine sous-jacente et l'association à la lamine nucléaire, mais n'apparaît pas en dépendre directement. En étudiant la dynamique de la conformation chromosomique du *Xic* au cours de la différenciation cellulaire, j'ai pu constater la robustesse de cette organisation, sauf sur le chromosome X inactif, qui se distingue par la perte des contacts chromosomiques préférentiels détectables sur son homologue actif.

Enfin, j'ai pu mettre en évidence que la variabilité du repliement général du chromosome X amène à un instant donné chaque allèle de *Tsix* à contacter physiquement des jeux de séquences distales différents, suggérant que l'environnement structural instantané de chacun de ces allèles à l'orée de l'activation mono-allélique de *Xist* est différent.

Ce travail, combinant des approches à l'échelle de la population cellulaire d'une part et de la fibre de chromatine unique d'autre part, apporte une nouvelle vision du paysage structural et régulateur dans lequel s'inscrit le contrôle de l'activité transcriptionnelle de *Xist*, et fourni de nouvelles perspectives concernant les principes fondamentaux de l'organisation topologique des chromosomes chez les mammifères.



---

CHROMOSOMAL ARCHITECTURE OF THE *Xic* LOCUS:

IMPLICATIONS FOR THE REGULATION OF X CHROMOSOME INACTIVATION

---

Early development of female mammals is accompanied by transcriptional inactivation of one of their two X chromosomes. This event is initiated following mono-allelic expression of the *Xist* non-coding RNA – what is achieved by the interplay of numerous *cis*-regulatory elements present within the *X inactivation center* (*Xic*), such as its repressive antisense *Tsix*. Our work aimed at throwing light on the structural landscape that underlies such long-range regulation. Characterization of the three-dimensional architecture of the *Xic*, by the means of Chromosome Conformation Capture (3C)-based techniques and *in situ* fluorescence hybridization (FISH), revealed that the respective promoters of *Xist* and *Tsix* contact many distal genomic elements within the *Xic*, and that at least one of such interacting region exerts long-range *cis*-transcriptional control. Noticeably, *Xist* and *Tsix* promoters associate with different sets of elements in their respective 5' direction that are spread out over several hundreds of kilobases

These experiments also revealed unforeseen properties of chromatin architecture. Indeed, the *Xic* appears to be partitioned in several sub-regions, each spanning between 200kb and 1Mb, inside which chromosomal interactions are preferentially established. The existence of these interaction domains integrates with other structural features of the genome, such as underlying chromatin composition and association with the nuclear lamina, but does not seem to directly depend on them. By analyzing chromosome conformation of the *Xic* during cell differentiation we document the robustness of this organizational principle, with the noticeable exception of the inactive X chromosome that assumes a folding pattern that is more random than its active homolog.

Finally we also bring evidence that variability in the folding pattern of the two X chromosomes in the same cell brings each *Tsix* allele in association with different sets of chromosomal partners at a given moment, suggesting that the instantaneous structural environment of each allele at the onset of mono-allelic *Xist* up-regulation is different.

By combining approaches at the scale of cell populations on the one hand, and at the single chromatin fiber level on the other, this study provides a first vision of the structural landscape in which *Xist* regulation takes place, and brings new insights concerning fundamental properties of chromosome organization in mammals.



# REMERCIEMENTS





Bien que le diplôme porte mon nom, ce travail n'est pas le mien. Il est le fruit de maintes idées plantées au fil des années par les nombreuses personnes que j'ai eu la chance de côtoyer. L'inspiration, scientifique et humaine, que j'ai trouvée parmi mes proches et moins proches est ce qui m'a poussé jusqu'à ce doctorat – et qui me presse d'embrasser la suite.

J'aimerais remercier chaleureusement mon Jury, non seulement pour avoir accepté d'évaluer ce travail, mais également parce ses membres ont été source de cette inspiration, chacun à des façons et des périodes différentes mais néanmoins déterminantes pour le développement du présent document - et de son auteur.

C'est évidemment vers Edith que vont mes sentiments les plus émus à l'idée de terminer ce manuscrit. J'ai considérablement apprécié l'attention avec laquelle tu as guidé ma thèse tout en me laissant suffisamment d'indépendance pour goûter à la fois aux périodes d'errements et à ces instants d'adrénaline intenses - auxquels je crois bien être devenu dépendant. Je te suis particulièrement reconnaissant de ta générosité scientifique et intellectuelle, des occasions incroyables que tu m'as présentées et des personnalités hors du commun que tu m'as permis de rencontrer. Le plaisir éprouvé à travailler au sein de ton équipe a été quotidien, tout simplement.

Sandrine évidemment, qui a m'a encadré au labo pendant mon master puis en photo sur son bureau. Je n'ai jamais compris pourquoi j'ai eu la chance de tomber dans tes bonnes grâces, mais je sais que j'aurais perdu les pédales plus d'une fois si tu n'avais pas été là tout du long, pour me remonter dans les périodes de froid, et me maintenir sur Terre dans les périodes bouillonnantes. Heureusement que tu as été, que tu es, et que tu seras là.

Julie (oui, c'est chronologique), dont les FISH m'ont attiré dans l'équipe : merci d'avoir tout inventé et de rester solidement en contact à travers tous ces kilomètres. Mes hommages à Macha qui m'a montré que si on s'applique ça ne rate pas (reviens !). Iku-senpai, je n'oublierai jamais qu'au début tu me volais ma corbeille et que tu me faisais l'obole des mictions de tes sujets de travail. Le traumatisme passé tu t'es évidemment révélé un collègue patient, intéressé et blagueur – merci pour tes attentions. Mes excuses à Patricia pour toutes les commandes à la dernière minutes ; merci d'être toujours de bonne humeur et de sang-froid. Osamu, que je remercie de m'avoir montré la 4C. Je tiens à témoigner de mon admiration pour Jen, pour m'avoir fait réaliser qu'on pouvait écrire cinq lignes sur un eppendorf et qu'à force d'acharnement on y arrive. La discrétion de Niall et ses passions dans l'autre vie, celle hors des murs du labo, sont un rappel à l'ordre que j'ai trop fait d'ignorer. Ma gratitude va également à Catherine C. pour ses judicieux conseils au début de mes aventures qui m'ont permis de retomber sur mes pieds. Catherine P, pour son don d'ubiquité dont j'aimerais connaître l'astuce.

Christel – que serions-nous sans toi, tout simplement ? Poussé par Constance j'ai appris à violenter ma nonchalance, sortir de mes habitudes, et compris que je ne risque pas d'avoir la réponse si je ne fais pas la manip – ou pas la bonne manip. Un profond merci à Martin pour sa clairvoyance et sa démonstration que la vie est ce qu'on en fait. Simão, merci de m'avoir appriché à parler portuguaich à Neuziñach. Lucie, ton dynamisme ne nous manque que trop. Quant à Anne-Valérie, force tranquille, j'aimerais laisser ta perspicacité, ton bon-sens et ton aplomb m'influencer ; merci pour les clefs de ton appartement. Les continuel encouragements de Katia arrivent toujours à me remonter. Ronan, ton efficacité et ton flegme m'épatent – quel est ton secret ? La franchise et le discernement d'Edda rafraîchissent les têtes enfumées ; j'admire ton jugement et ta présence

d'esprit. Tim, toi qui commence l'aventure, souviens-toi seulement qu'il ne faut pas croire un mot de ce qu'on te raconte, et tout ira bien. Camille, qui fait tant pour le bien commun de l'équipe. Les derniers moments de cette thèse ont été le moment de passer le flambeau. Luca, je te remercie pour réussir mieux que moi les manips que je suis censé t'apprendre, pour analyser les matrices qui me laissent perplexes, et t'enthousiasmer à la perspective de cliquer sur des points pendant plusieurs années ; je suis très honoré de partager un bout de projet avec quelqu'un d'aussi fin et perspicace. Maud enfin, fraîche recrue qui m'est plus familière en plumes et robe de soirée signée BGA qu'en blouse – viens me voir j'ai un super *pet project* pour tes longues journées sans embryons !

En dehors de l'équipe, j'ai eu la chance de pouvoir travailler avec plusieurs personnalités marquantes. Au sein du laboratoire de Job Dekker, qui m'a accueilli plusieurs semaines sous la neige du Massachussetts, Bryan, Nynke, Natasha et Amartya ont acquis mon éternelle gratitude pour m'avoir initié aux 3C et 5C. Nicolas, au sein de l'équipe d'Emmanuel Barillot, qui résout mes problèmes plus vite que je ne les formule. Je suis également reconnaissant à Erik et Wouter de Laat et pour partager leurs découvertes et leur logique implacable. Merci également à Jane Skok et aux Skokettes, notamment Julie (la même !) et Jeanne, pour me donner l'impression de leur être utile. Merci à la Société Française de génétique pour m'avoir permis de me rendre à un symposium au CSHL.

L'aventure prend fin dans l'unité de Biologie du développement, dont je remercie vivement tous les membres pour leur volonté d'échanges scientifiques. Merci à Virginie et Déborah pour leur efficacité. Je tiens particulièrement à témoigner de mon admiration à l'équipe d'imagerie, Olivier R, Olivier L et Tristan qui assurent un service d'une qualité absolument remarquable, et en plus dans la bonne humeur. Promiscuité oblige, je tiens nos voisins Bourchiss Bourrekiss Bourc'his, Rachel, Julian, Mounia, Patricia, Natasha, Sophie et leur chef dans une estime particulière – en vérité surtout grâce à Charlotte qui en plus de m'avoir marrainé à BGA m'a permis de me faire des copains au BDD.

Avant tout cela, il y avait l'UMR218. On était jeunes, alertes, et j'avais un bureau sur des roulettes que les gens déplaçaient en le cognant lorsqu'ils rentraient trop vite dans le labo. Une belle époque. Je tiens à remercier chaleureusement Geneviève Almouzni à la fois pour son soutien et ses conseils, ainsi que pour le dynamisme de son unité avec qui j'ai toujours tant de plaisir de collaborer à la machine à café. Que d'aventures partagées avec Marion et Baptiste, pour qui le mois de septembre s'annonce également chargé... Une pensée à Patricia, qui m'a présenté mon premier DeltaVision – grandes émotions. Au cours de cette thèse, les idées ne me sont jamais apparues plus claires que la veille de mes redoutés tutorats ; merci aux membres de ce comité annuel – Angela Taddei, Catherine Corbel, Edith Heard bien sûr, Geneviève Almouzni, Nathalie Dostatni, Pierre-Antoine Desfossez et Jonathan Weitzman – pour m'avoir ainsi guidé. Un grand merci à tous les membres passés et présents de l'UMR218 pour leur conseils, leur encouragements et leurs moult bières.

Les diverses réunions inter-institut – X.I.D. meeting, B.E.E.R. session... – permettent également de rencontrer les autres, ceux qui font la même chose en différent. Merci à l'équipe Claire Rougeulle, particulièrement à Angélique, Andrew, Mélanie et son membre éponyme pour leur naturel, leur volonté de partager leurs observations, leurs idées et leurs pots de mariage. Merci également à Thomas Boudier et Fabrice Cordelières, sans qui ImageJ serait resté effrayant et avec trop peu de mémoire pour ouvrir mes images.

J'aimerais également profiter de cette occasion pour remercier ceux qui ont joué un rôle déterminant dans mes premiers pas en laboratoire. C'est à Geneviève Barrier et François Jacob que revient d'avoir gratté l'allumette. Fatima Mechta-Grigoriou et Damien Gérald m'ont mis le pied à l'étrier, et permis de rencontrer Chaouki Miled et Jonathan Weitzman chez Moshe Yaniv, personnalités marquantes que je suis heureux d'avoir suivi jusqu'à aujourd'hui. Caroline Parmentier et André Calas ont toute ma reconnaissance pour m'avoir insufflé le plaisir de chercher des explications en biologie. Que Brinda Ravikumar et David Rubinsztein trouvent mon meilleur souvenir pour m'avoir laissé goûter la première fois au goût de l'échec surmonté. Mon séjour chez Rob Martienssen reste un de mes moments les plus intenses de « science de nuit », où Ana Kloc et Mikel Zaratiegui ont été des guides salutaires.

Manifestement la flopée de magistériens cherche sa mention. Les années BGA furent une période profondément exaltante. Merci aux enseignants d'y mettre tellement du vôtre. Merci aux petits qui m'entourent depuis. Françoise, Claire, Jérémy, Kevin, André, bande de sales conspirateurs, ainsi que Soline, Adèle et Marie... Les plus vieux aussi, Aude, Anne, Doruk, Charlotte (encore elle !), Ana-Maria, Michael, Michel, Maud... Des amis qui ne s'endorment pas (toujours) quand on leur parle de chromosomes.

Enfin merci aux amis qui ont un peu moins le nez sur la paillasse, qui m'ont permis de « descendre de mon vélo et de me regarder pédaler » comme dirait Hélène L. Ce fut essentiel pour arriver au bout de ce doctorat. Laétitia, tu n'es pas venue à la soutenance alors j'espère qu'il pleut en Corse. Les alsaciens Maxime, Clémence, Audrey, In(), Damien, Pauline, il faut que je vous appelle pour vous dire que je soutiens enfin et vous expliquer que c'est pour ça qu'on s'est pas vu depuis six mois. Merci à Xavier d'être un lien avec une réalité dont le labo me fait souvent oublier l'existence (*i.e.* la vie). Merci à Yannick d'être au fond celui qui s'inquiètera de mettre tout ça en perspective. Et il a bien raison.

Merci à mon chromosome X de rester actif, du moins jusqu'à la soutenance...



# Table des matières

<b>AVANT-PROPOS : LE PARADIGME DE L'INACTIVATION DU CHROMOSOME X .....</b>	<b>17</b>
La régulation transcriptionnelle de <i>Xist</i> .....	19
La répression transcriptionnelle par l'ARN <i>Xist</i> .....	20
Inactivation du chromosome X et régulation transcriptionnelle à longue distance.....	21
<b>INTRODUCTION : ORGANISATION STRUCTURALE DU GÉNOME ET CONTRÔLE DE L'ACTIVITÉ TRANSCRIPTIONNELLE .....</b>	<b>23</b>
<b>1. TSS et promoteur propre.....</b>	<b>25</b>
<b>2. Les régions régulatrices proximales.....</b>	<b>28</b>
2.1. Généralités.....	28
2.2. Structure et dynamique fonctionnelle.....	29
2.2.1. L'hypersensibilité à la DNaseI.....	29
2.2.2. La fixation de facteurs <i>trans</i> détermine l'activité des régions de contrôle proximales.	30
2.3. Le complexe Mediator.....	31
<b>3. Les régions de contrôle distales.....</b>	<b>32</b>
3.3. Structure et fonction des régions régulatrices distales.....	33
3.4. Activité et mécanismes d'action des régions de contrôle de la transcription.....	46
3.5. Contrôle de l'activation transcriptionnelle par effet isolateur.....	50
<b>4. Voisinage génomique et régulation transcriptionnelle .....</b>	<b>58</b>
4.1. L'effet de position.....	58
4.2. Considérations sur le concept de séquence régulatrice.....	65
4.2.1. Structure n'est pas synonyme de fonction.....	65
4.3. Paysage transcriptionnel et loci co-régulés : l'exemple des <i>Hox</i> .....	67
4.4. Changements transcriptionnels régionaux induits par le repositionnement nucléaire	71

<b>5. Architecture des chromosomes et dynamique de l'organisation nucléaire .....</b>	<b>73</b>
5.3. De la dizaine à la centaine de kb .....	75
5.4. De la centaine de kb au Mb .....	77
5.5. Du Mb à la dizaine de Mb .....	78
5.6. De la dizaine de Mb au chromosome .....	79
5.7. Positionnement relatif des chromosomes et interactions physiques en <i>trans</i> .....	83
5.8. Conclusions et perspectives pour l'étude de la régulation transcriptionnelle .....	84

## **ANNEXES A L'INTRODUCTION ..... 87**

Revue 1 <i>Regulation of X-chromosome inactivation by the X-inactivation centre</i> Augui*, Nora* & Heard 2011 .....	89
----------------------------------------------------------------------------------------------------------------------	----

Revue 2 <i>Chromatin structure and nuclear organization dynamics during X-chromosome inactivation</i> Nora and Heard, 2011 .....	91
----------------------------------------------------------------------------------------------------------------------------------	----

## **CONTEXTE DES PROJETS DÉVELOPPÉS ET RÉSUMÉ DES TRAVAUX.. 93**

Expression mono-allélique de <i>Xist</i> et l'importance du locus <i>Tsix</i> .....	95
Caractérisation de l'organisation chromosomique du locus <i>Tsix</i> et de sa variabilité inter-allélique .....	96
Éléments de contrôle à distance et complexité du locus <i>Xic</i> .....	98
Caractérisation de l'organisation structurale du <i>Xic</i> et des influences régulatrices qu'elle soutend	98

## **RESULTATS ..... 109**

Article 1 <i>Networks of chromosomal interactions provide a structural basis for the regulatory landscape of the X inactivation center</i> Nora et al. – en préparation .....	111
Résultats supplémentaires .....	113

## **ANNEXES AUX RESULTATS ..... 115**

Article 2 <i>Sensing X-chromosomes pairs before X inactivation via a novel X-pairing region of the Xic</i> Augui, Fillion, Huart, Nora et al. 2009 .....	117
----------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Article 3 <i>The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA</i> Splinter, de Wit, Nora et al. 2011 .....	119
------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

**DISCUSSION ET PERSPECTIVES ..... 121**

**ANNEXES..... 137**

Revue 3 *Dynamique nucléaire et inactivation du chromosome X* Nora & Augui 2009..... 139

Revue 4 *Etablissement de l'inactivation transcriptionnelle du chromosome X pendant le développement embryonnaire* Nora & Heard 2009..... 141

Revue 5 *X chromosome inactivation: when dosage counts* Nora & Heard 2009..... 143

Revue 6 *RNF12 – a jack of all trades in X inactivation?* Schulz, Nora and Heard, 2011..... 145

**BIBLIOGRAPHIE ..... 147**





## **AVANT-PROPOS :**

**Le paradigme de l'inactivation du  
chromosome X**



Le développement embryonnaire précoce des mammifères femelles s'accompagne de l'inactivation transcriptionnelle de l'un de leurs deux chromosomes X. Chez la souris, l'inactivation touche aléatoirement l'un des deux chromosomes X et est mise en place dans les cellules de l'embryon. Un locus complexe, le centre d'inactivation du chromosome X (*Xic*), orchestre l'initiation de ce phénomène en contrôlant l'activation transcriptionnelle mono-allélique de *Xist*. L'ARN non-codant *Xist* (*X Inactive Specific Transcript*) est capable de recouvrir en *cis* le chromosome à partir duquel il est produit, et d'induire son hétérochromatinisation. Deux grandes questions se posent alors : comment l'activation transcriptionnelle de *Xist* est-elle contrôlée ? Comment l'expression de l'ARN *Xist* déclenche-t-elle en retour la répression transcriptionnelle ?

Mon travail doctoral a été l'occasion de m'immerger dans la dense littérature qui concerne l'inactivation du chromosome X et de participer à la rédaction de plusieurs revues de synthèse, dont le lecteur trouvera une copie à la fin de ce chapitre introductif ainsi qu'en annexe du manuscrit. Ces articles s'intéressent en particulier à deux volets sur lesquels nos propres travaux ont porté, à savoir la régulation de l'expression de l'ARN *Xist* d'une part, et l'architecture nucléaire des chromosomes X actifs et inactifs d'autre part. Dans Nora et Augui, 2009 et Nora et Heard, 2010 nous expliquons en termes simplifiés des généralités concernant le phénomène d'inactivation transcriptionnelle du chromosome X au cours du développement embryonnaire, et son lien intime avec l'organisation physique de la chromatine à l'intérieur du noyau.

### **La régulation transcriptionnelle de *Xist***

La question complexe de la régulation de l'expression de *Xist* est traitée en détail dans (Augui\*, Nora\* et Heard, 2011), et nous y dégageons l'importance de plusieurs régions génomiques voisines à travers divers mécanismes. Nous y expliquons notamment le rôle répresseur en *cis* de la transcription anti-sens à *Xist* contrôlée par le locus *Tsix*, lui-même influencé par plusieurs éléments régulateurs dont *Xite* et *DXPas34* (*X-inactivation Intergenic Transcription Element* et microsatellite de 34-mer sur la band D du X identifié à l'Institut Pasteur), que nous avons étudié en détail dans les travaux qui seront présentés dans la section de résultats. Nous y discutons également ce qui est sur des mécanismes fondamentalement différents par lequel deux régions distales de *Xist* sont impliquées dans sa *trans* activation. Le locus *Xpr* (*X-Pairing Region*) est engagé dans des associations physiques en *trans*, entre les allèles *Xic*, qui précèdent et accompagnent l'activation de *Xist* ; sa caractérisation initiale a fait l'objet d'une étude à laquelle j'ai participé, et dont le manuscrit est joint à ce volume (Augui et al., 2007). Le locus *Rnf12* (*Really Interesting New Gene Finger [protein] 12*) est quant à lui responsable de la production d'une ubiquitine E3-ligase dont la surexpression active

transcriptionnellement *Xist*. Dans un commentaire de l'étude ayant menée à la reconnaissance initiale du rôle RNF12 nous discutons des implications conceptuelles de ce phénomène sur les différentes étapes de l'initiation de l'inactivation du chromosome X (Nora et Heard, 2009). Les mécanismes par lesquels cette régulation est exercée sont mal compris, d'autant que l'altération de ce locus entraîne des effets différents suivant les expérimentateurs, pour des raisons que nous avons commenté par ailleurs (Schulz et al., 2011). D'autres loci contrôlant la production d'ARN non-codants *Jpx* et *Ftx* (*Juxtaposed to Xist* et *Five-prime to Xist*) ont été impliqués à différentes étapes de la régulation de l'expression de *Xist*. Au vu de l'ensemble de la littérature concernant le contrôle de l'activité transcriptionnelle de *Xist*, nous concluons qu'il existe encore des régions non identifiées dans son voisinage direct ou indirect qui participent directement ou indirectement à sa régulation.

### **La répression transcriptionnelle par l'ARN *Xist***

La capacité de l'ARN *Xist* à induire la répression transcriptionnelle du chromosome X est mal comprise. Comme nous le détaillons dans Nora et Heard, 2011, la chromatine exposée à l'ARN *Xist* subit de profonds changements dans sa composition et son organisation. Des cofacteurs de *Xist* important pour la mise en place de l'état inactif ont été identifiés : SATB1 et HnRNP-U/SAF-A (*Special AT-rich Sequence Binding [protein] 1* et *Heterogeneous Nuclear RiboNucleoProtein U/ Scaffold Attachment Factor A*). De manière intrigante, ce sont des protéines de la matrice nucléaire participant au contrôle de l'architecture du génome. La maintenance de l'état inactif ne requiert pas l'expression continue de *Xist*, mais apparaît impliquer des enzymes responsables de la catalyse de modifications post-traductionnelles des histones, participant au complexe PRC2 ; une réactivation transcriptionnelle n'est cependant observée que dans certains types cellulaires, reflétant probablement la forte redondance des mécanismes répresseurs recrutés par l'ARN *Xist*. La protéine SmcHD1, apparentée aux cohésines, est également nécessaire à la maintenance de l'état inactif, renforçant l'idée d'un lien intime entre la topologie du X inactivé et sa répression transcriptionnelle. L'étude comparative précise de la conformation du X actif et du X inactif, à laquelle nous avons participé en collaboration avec le laboratoire du Dr Wouter de Laat, (Splinter et al., 2011– voir manuscrit joint à ce volume), a révélé que les associations chromosomiques distales entre les différents loci transcrits ou non transcrits du X actif sont perdues sur le chromosome X inactif, où les interactions apparaissent plus nombreuses mais non dirigées vers des partenaires particuliers. La délétion du locus *Xist* sur le X inactif, qui n'entraîne pas la réactivation transcriptionnelle, est cependant suffisante pour remodeler l'architecture chromosomique vers un arrangement topologique qui ressemble à celui du X actif. L'ensemble de ces études suggère donc que dans le

cadre de l'inactivation du chromosome X, il existe un lien plus direct entre conformation chromosomique et composition de la chromatine d'une part qu'entre composition de la chromatine et activité transcriptionnelle d'autre part.

### **Inactivation du chromosome X et régulation transcriptionnelle à longue distance**

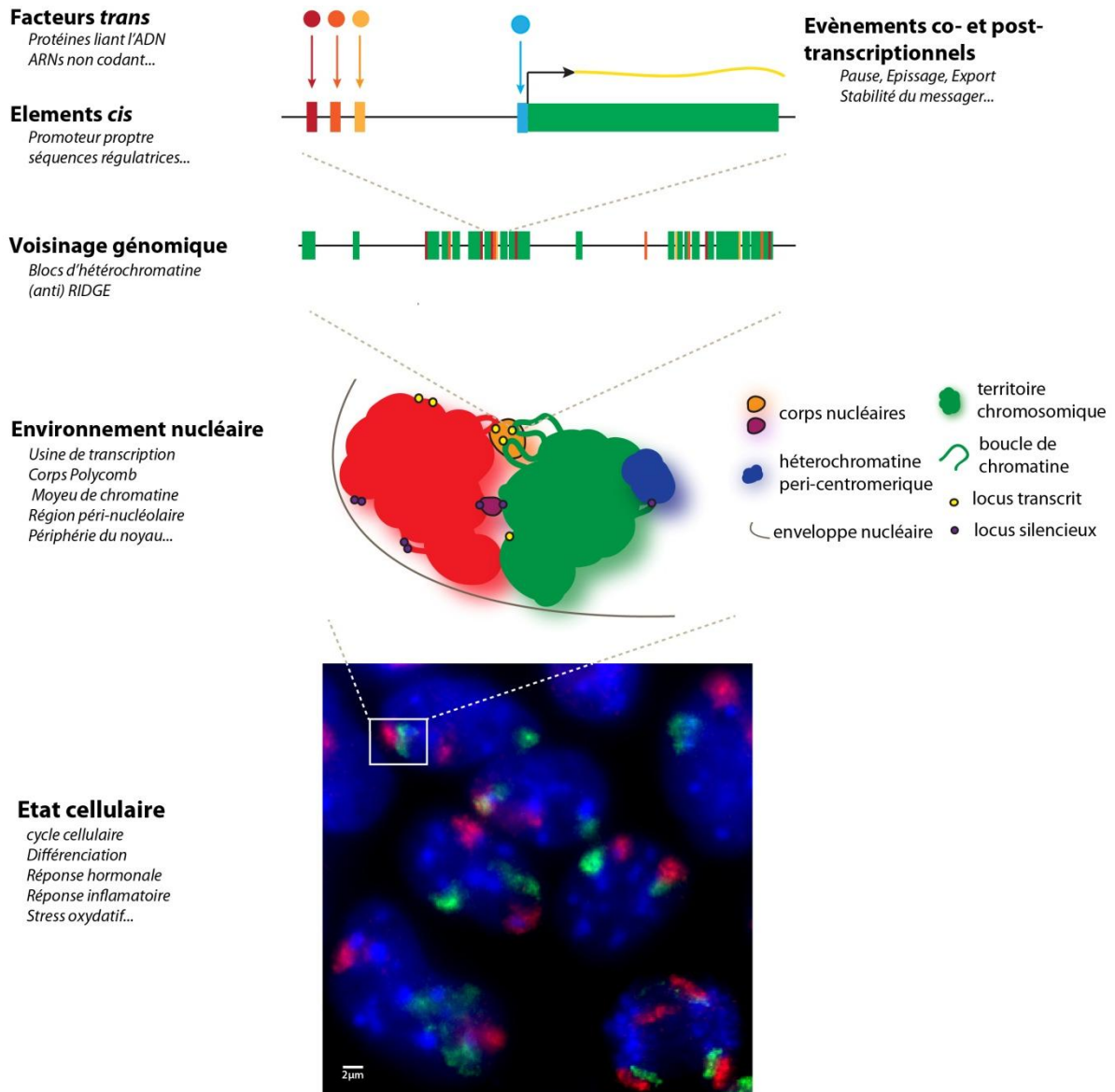
La régulation de l'expression de Xist d'une part, et le remodelage de l'architecture chromatinienne lors de l'inactivation du chromosome X d'autre part sont des illustrations du principe fondamental de contrôle de l'activité transcriptionnelle à longue distance. Afin de replacer nos travaux dans un contexte plus général que celui détaillé dans nos articles de revue, très focalisés autour du modèle de l'inactivation du chromosome X, nous avons donc décidé de consacrer le chapitre introductif qui va suivre aux mécanismes sous-tendant la régulation transcriptionnelle chez les mammifères. Nous nous concentrerons sur les mécanismes contrôlant l'activation de la transcription, en portant une attention particulière sur les relations régulatrices entretenues entre les promoteurs et leur contexte génomique. Nous discuterons ainsi tout d'abord de la régulation locale de la transcription, pour nous intéresser aux communications fonctionnelles entre éléments de plus en plus distaux, pour finir avec des remarques générales sur l'organisation topologique des chromosomes au sein du noyau.



# **INTRODUCTION :**

**Organisation structurale du génome  
et contrôle de l'activité transcriptionnelle**





**Figure 1. Les différentes échelles chromosomiques de la régulation transcriptionnelle chez les mammifères.**

En amont des évènements régulateurs co- et post-transcriptionnels s'exerce le contrôle *via* des couples facteur *trans* / élément *cis*, à plus ou moins grande distance du site d'initiation de la transcription. D'un point de vue plus régional, les chromosomes sont organisés en larges segments où l'activité transcriptionnelle et les effets régulateurs sont particulièrement denses (jungles), intercalés par de longs blocs transcriptionnellement inertes (déserts). Des segments distaux sur l'échelle génomique linéaire peuvent se retrouver physiquement proches dans l'espace nucléaire. L'architecture chromosomique influence, et est en retour influencée, par la présence de nombreux corps nucléaires, qui sont des assemblages multi-protéiques impliqués dans le métabolisme de l'ADN (ex : usine de transcription, de réplication, de réparation) ou des ARN (ex : foci d'épissage [*splicing speckles*]). Des structures formées par des complexes (désoxy)ribonucléoprotéiques (ex : nucléoles, hétérochromatine péri-centromérique) délimitent des environnements nucléaires focaux où l'activité transcriptionnelle est localement contrôlée de façon particulière. Enfin, les stimuli perçus par la cellule peuvent agir plus ou moins directement sur ces différents niveaux d'organisation (ex : récepteurs nucléaires, voies de transduction) et ainsi réguler différents aspects de l'activité transcriptionnelle.

L'expression de Xist au cours de la différenciation repose sur l'activation transcriptionnelle de son promoteur (Sun et al., 2006). Nous allons donc présenter ce qui est su des mécanismes généraux contrôlant le phénomène d'activation transcriptionnelle chez les mammifères. Nous discuterons tout d'abord des événements en jeu au promoteur propre et dans sa région proximale. La régulation de l'expression de Xist étant par ailleurs connue pour faire intervenir de nombreux éléments de contrôle à distance, nous présenterons ensuite ce qui est su du mode d'action de tels éléments, ainsi que de la façon dont leurs effets régulateurs sont ponctués à travers le génome. Enfin, nous évoquerons le contexte structural dans lequel ces mécanismes opèrent en discutant des différentes échelles de l'organisation physique des chromosomes dans le noyau.

### 1. TSS et promoteur propre

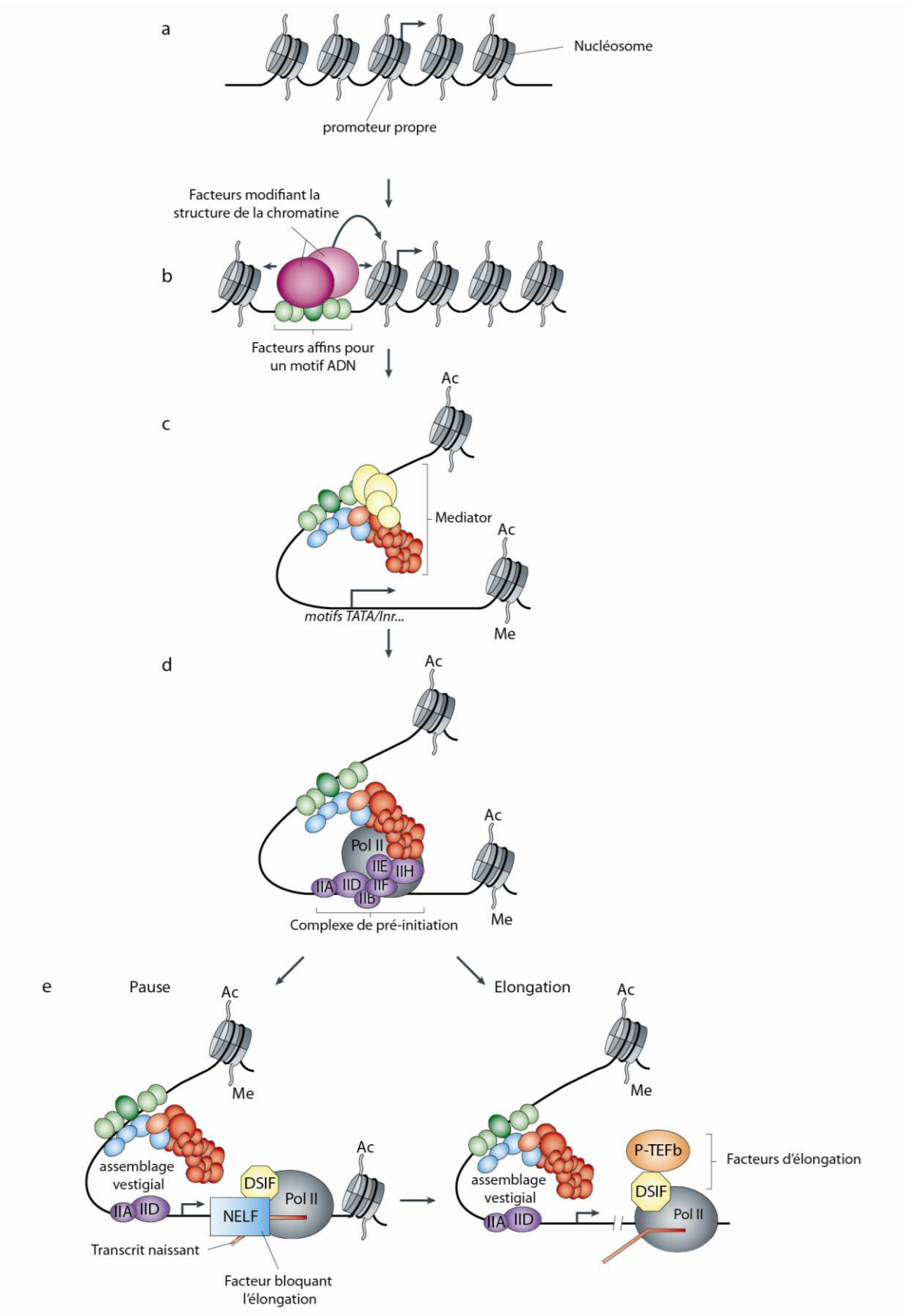
Le site d'initiation de la transcription est déterminé par une combinaison de motifs ADN courts qui forment le promoteur propre. Cet élément génomique est typiquement localisé dans les 80 paires de bases qui entourent le site de départ de l'ARN polymérase II (TSS pour *Transcriptional Start Site* : premier ribonucléotide du futur polymère ARN). La Pol-II est incapable d'initier la transcription en elle-même et requiert la liaison préalable des facteurs dits généraux de transcription (GTF pour *General Transcription Factors* : TFIIA, TFIIB, TFIID, TFIIE, TFIIF et TFIIH, qui sont pour la plupart de gros assemblages multi-protéiques). L'activité promotrice d'une séquence lui est ainsi conférée par sa capacité à recruter ces GTFs, puis la Pol-II – qui ensemble forment le complexe de pré-initiation (ou PIC pour *Pre-Initiation Complex*). Plusieurs motifs promoteurs existent, qui se distinguent par leur affinité relative pour les différents GTFs et leur distance par rapport au TSS. Par exemple, la célèbre TATA box est toujours en amont du TSS, alors que la transcription est initiée dans le motif Inr (*Initiator*) et que les motifs MTE et DPE (*Motif Ten Element* et *Downstream Promoter Element*) sont plus en aval (voir **figure 2**). Quant à leur activité, la boîte TATA peut fonctionner comme séquence promotrice à elle seule ; son activité est cependant modifiée par la présence d'autres séquences dans le promoteur propre, et par exemple elle pourra spécifier le départ de la transcription précisément dans le motif Inr si celui-ci est présent. Les motifs DPE et MTE peuvent également initier la transcription en présence de Inr, mais il est important de garder à l'esprit qu'il n'y a pas une séquence promotrice universelle ; au contraire, chaque promoteur propre est constitué d'un ou plusieurs motifs qui peuvent agir en synergie pour contrôler le recrutement de la Pol-II (pour revue voir Fuda et al., 2009).

Un même promoteur peut entraîner le démarrage de la Pol-II à plusieurs TSS. Là encore il semble que ce soit la combinaison des séquences du promoteur propre et de son voisinage qui détermine si un promoteur fonctionne de manière focale ou dispersée. C'est ainsi que pour les

promoteurs propres à TSS dispersés, qui représentent près des deux tiers des promoteurs chez les vertébrés, les différents sites d'initiation peuvent s'étaler sur une centaine de paires de bases. En terme de structure ils se distinguent par l'absence générale de motif TATA, DPE ou MTE, ainsi que par leur localisation dans les îlots CpG chez les vertébrés (Juven-Gershon et al., 2010 pour revue). Contrairement aux promoteurs à TSS dispersés, qui sont généralement actifs de façon ubiquitaire, l'activité des promoteurs à TSS focalisés est généralement régulée, et ils se distinguent en outre par d'autres propriétés structurales – comme un positionnement des nucléosomes plus contraint, contenant le variant H2A.Z et une sous-unité H3 triméthylée sur le résidu Lysine 4 (H3K4me3) – suggérant des différences importantes dans la façon dont la transcription y est initiée (Rach et al., 2011).

L'empaquetage de l'ADN en chromatine représente une barrière considérable à l'accessibilité des facteurs de transcription aux motifs d'ADN sous-jacents. Les GTFs ne sont pas capables d'altérer eux-mêmes la structure de la chromatine avant de se lier à l'ADN. Leur liaison à une séquence promotrice dépend donc d'évènements de déstabilisation des nucléosomes, qui peuvent soit survenir passivement de par les propriétés locales intrinsèques de la fibre d'ADN, soit être activement catalysés par des facteurs de remodelage de la chromatine. Les propriétés biophysiques de la fibre d'ADN peuvent en effet participer à la création de régions peu fréquemment occupées par un nucléosome (NDR pour *Nucleosome Depleted Region*) : par exemple, un poly(dA:dT) autorise une forte courbure locale et est donc tolérant à l'éviction des nucléosomes (Kaplan et al., 2009). Paradoxalement, des motifs favorables à l'accueil d'un nucléosome déterminent également le positionnement d'une NDR, car l'encombrement stérique dû au positionnement efficace d'un nucléosome en empêchera un autre de se lier dans les cent paires de bases voisines.

**Figure 2. L'initiation de la Transcription.** (A) en absence de fixation de facteurs *trans* les nucléosomes n'occupent généralement pas de position préférentielle, résultant en la couverture des motifs du promoteur propre. (B) La fixation de facteurs *trans* aux régions régulatrices proximales induit le repositionnement des nucléosomes, passivement par effet de phase dû à l'encombrement stérique et parfois activement via le recrutement de complexes de remodelage de la chromatine, aboutissant au démasquage du promoteur propre (C) Le complexe Mediator est recruté par son affinité avec les facteurs *trans* liés aux régions régulatrices. Ce large assemblage macromoléculaire, ainsi que les divers co-facteurs non représentés, entraîne des changements locaux de la structure de la chromatine, comme des modifications post-traductionnelles des histones environnantes telles H3Ac (Ac) et H3K4me3 (Me) par exemple, et stimule l'assemblage du complexe de pré-initiation au promoteur propre (E) L'amorçage de l'ARN polymérase II s'accompagne initialement de la transcription sur quelques dizaines de paires de base seulement (**pause**) ; la progression au-delà du voisinage du promoteur propre n'est initiée qu'après dissociation contrôlée des facteurs bloquant (**élongation**). *Adapté de Malik et Roeder, 2010*



Ainsi, le positionnement d'une NDR peut affecter l'exposition de séquences plus distales, avec une périodicité correspondant à l'espace entre deux nucléosomes (Valouev et al., 2011). Les promoteurs possédant une NDR constitutive déterminée par la composition nucléotidique locale sont qualifiés d'ouverts, et sont généralement actifs de façon ubiquitaire – en tout cas chez la levure. Les promoteurs dont l'activité n'est pas constitutive sont au contraire souvent associés à des nucléosomes dans leur état inactif. Ils sont qualifiés de couverts. Les nucléosomes y sont en compétition avec les autres facteurs capables de lier l'ADN pour l'occupation des séquences spécifiant l'assemblage du PIC. Comment permettre la liaison efficace des GTFs dans un tel contexte ? Éloignons nous du TSS.

## **2. Les régions régulatrices proximales**

### **2.1. Généralités**

La région régulatrice de la transcription, au sens du segment d'ADN contrôlant le recrutement du PIC à un promoteur propre donné, peut s'étendre bien au-delà des 80 paires de bases qui encadrent le(s) TSS. Les motifs qui y sont retrouvés sont complètement différents de ceux du promoteur propre, et ne correspondent plus à des sites de liaison aux GTF mais à des séquences permettant le recrutement de facteurs activateurs ou répresseurs de la transcription. La nature de ces facteurs est aussi variée que les mécanismes par lesquels ils régulent l'activité transcriptionnelle, qui restent généralement mal compris. Il peut s'agir de protéines monomériques ou de grands complexes, faisant parfois même intervenir un ou plusieurs composants ribonucléiques, et ils peuvent moduler différentes étapes de la réaction de transcription – de l'assemblage du PIC à l'efficacité d'élongation – ou même contrôler le métabolisme du transcrit naissant (pour revue voir Saunders et al., 2006). Des facteurs modulant l'activité du PIC via le promoteur proximal, certains exercent leur fonction régulatrice en interagissant physiquement avec les GTF liés au promoteur propre voisin – voire avec la Pol-II elle-même. D'autres agissent en déplaçant localement les nucléosomes ou en modifiant de façon covalente les histones avoisinantes, ce qui modifie en retour la capacité des autres facteurs de transcription à se lier à leur séquence cible, ou crée des sites protéiques de recrutement pour des effecteurs spécialisés.

Certains facteurs de transcription sont directement capables de reconnaître un motif ADN même à la surface du nucléosome. Les termes « facteurs pionniers » sont parfois utilisés pour désigner de tels facteurs car leur rôle est initial dans l'initiation de la transcription – même si leur action peut parfois être très indirecte comme nous le détaillerons plus loin. De manière importante,

les facteurs pionniers peuvent également participer au dégagement actif des nucléosomes sur une large portion d'ADN en recrutant des enzymes de la famille SWI/SNF, capables de déplacer, voire d'évincer, un ou plusieurs nucléosomes (Smale, 2010 pour revue). Certaines protéines de la famille *Forkhead* sont ainsi connues pour pouvoir initier directement le dégagement du promoteur propre ou d'éléments régulateurs (pour revue voir Benayoun et al., 2011).

## **2.2. Structure et dynamique fonctionnelle**

### **2.2.1. L'hypersensibilité à la DNaseI**

Il est connu depuis la fin des années 70 que certaines régions génomiques se distinguent par une structure chromatinienne localement altérée et particulièrement accessible à la DNaseI. Ces régions, appelées sites d'hypersensibilité à la DNaseI (DHS) correspondent à des éléments génomiques très peu fréquemment occupés par des nucléosomes. Plusieurs études séminales dans les années 80, notamment au locus  *$\beta$ -globin* chez les vertébrés et aux loci répondant aux chocs thermiques chez la drosophile, ont mis en évidence que si certains de ces DHS correspondent à des promoteurs actifs, d'autres sont localisés à plusieurs kilobases du plus proche TSS détectable (Wu et al., 1979a, Wu et al., 1979b, Stalder et al., 1980). De manière importante, l'hypersensibilité à la DNaseI de ces régions n'est pas une propriété constitutive ; au contraire, elle n'est observée que dans les types cellulaires où les promoteurs environnants sont actifs, suggérant d'emblée un lien entre la structure chromatinienne de ces éléments distaux et l'activité transcriptionnelle des régions voisines. Le milieu des années 80 est également marqué par le développement de techniques permettant de mesurer la capacité d'une séquence à activer en *cis* la transcription d'un promoteur connexe, hétérologue ou non, en contexte épisomal ou génomique. Historiquement, l'inspection des loci *IgH* et *Ig $\kappa$*  a initialement révélé des similarités – à la fois en termes de séquence nucléotidique et de capacité à activer en *cis* la transcription de promoteurs voisins – entre des DHS introniques et certains éléments activateurs de la transcription présents dans les génomes de virus eucaryotes (Neuberger, 1983 ; Banerji et al., 1983 ; Parslow et Granner, 1983 ; Emorine et al., 1983 ; Picard et Schaffner, 1984). Corroborées par des résultats à de nombreux autres loci, ces observations ont rapidement mené à l'idée que certains DHS distaux correspondent à des éléments régulateurs – activateurs en l'occurrence – de l'activité des promoteurs voisins. Ces tests fonctionnels ont ainsi permis de corrélérer l'état structural d'un élément génomique (hypersensibilité à la DNaseI) à sa capacité à stimuler la transcription par la démonstration que ces deux propriétés sont généralement tissu-spécifiques d'une part, et concomitantes d'autre part. Ces deux observations sont à l'origine de questions encore

d'actualité aujourd'hui, concernant respectivement le rôle des facteurs *trans* dans la spécification de l'activité des séquences *cis*-régulatrices et les mécanismes moléculaires par lesquels celles-ci peuvent stimuler la transcription.

### 2.2.2. La fixation de facteurs *trans* détermine l'activité des régions de contrôle proximales

Le développement des techniques d'immuno-précipitation de la chromatine (ChIP ; Gilmour et Lis, 1984 ; Orlando et Paro, 1993 ; Walter et al., 1994) a permis de tester l'existence d'associations physiques entre un facteur protéique de choix et une séquence ADN donnée *in vivo*. Ainsi une des premières applications de la ChIP chez les mammifères a démontré la capacité d'un élément génomique s'étalant jusqu'à deux cents paires de bases en amont du promoteur du locus *IFN- $\beta$*  à lier de façon coopérative plusieurs facteurs protéiques initialement identifiés comme activateurs de la transcription à partir du promoteur *IFN- $\beta$*  en condition d'infection virale (Wathelet et al., 1998). Ces études ont ainsi d'une part démontré que des séquences distales au promoteur peuvent contrôler l'initiation de la transcription, et ont d'autre part illustré le rôle de l'interaction entre facteurs protéiques *trans* et séquences *cis*-régulatrices dans l'induction transcriptionnelle.

Il a depuis été clairement démontré que la modulation de l'activité transcriptionnelle par la présence de séquences proximales ou distales dépend à la fois de propriétés intrinsèques à ces éléments génomiques, en *cis* du promoteur, et du ou des facteur(s) *trans* s'y liant. Ainsi, un même élément génomique peut sous-tendre différents effets sur l'activité d'un promoteur voisin, suivant qu'il est lié à un facteur donné ou à un autre. Cette propriété est particulièrement importante lors de la différenciation cellulaire, où une même séquence peut servir de plateforme de recrutement pour différents complexes, suivant les facteurs de transcription disponibles dans un contexte cellulaire ou dans un autre (voir **figure 3**). Citons l'exemple du locus *Kit*, dont la transcription est stimulée dans les précurseurs érythroïdes où le facteur GATA-2 est fixé en amont de son promoteur, tandis qu'elle est réprimée une fois que GATA-1 y remplace GATA-2 dans les cellules matures (Jing et al., 2008).



**Figure 3. Dynamique des effets régulateurs.** L'effet régulateur sous-tendu par une séquence distale sur un promoteur dépend de la nature des facteurs qui s'y fixent. Ainsi par exemple, au cours de la différenciation, l'expression tissu-spécifique de certains facteurs *trans* peut conférer un effet séquentiellement activateur puis répresseur de l'activité transcriptionnelle d'un promoteur cible.

### 2.3. Le complexe Mediator

Comment les facteurs protéiques capables de se lier à l'ADN de façon site spécifique activent-ils la transcription ? D'une manière qui peut paraître surprenante, leur action régulatrice est rarement due à une modulation directe de l'assemblage du PIC au promoteur propre. En effet, la fixation de ces facteurs n'est généralement que la première étape dans l'assemblage de complexes macromoléculaires modulaires, qui jouent le rôle d'intermédiaires entre les facteurs directement liés à l'ADN et le PIC (voir **figure 2**). Ces intermédiaires sont regroupés sous le terme général de co-activateurs, et se distinguent des activateurs directs par leur incapacité à se lier à l'ADN de façon site-spécifique. La dénomination plus générale de co-facteur est cependant préférable, car en réalité certains de ces complexes exercent une action répressive sur la transcription.

Le complexe Mediator est l'un des co-facteurs de transcription les plus célèbres, probablement à cause de son rôle essentiel à la majorité des sites de transcription par l'ARN Pol-II et de sa grande conservation structurale et fonctionnelle chez les eucaryotes (pour revue voir Malik et Roeder, 2010). Bien qu'il module l'efficacité d'assemblage de l'ARN Polymérase II au promoteur propre, Mediator n'est généralement pas considéré comme un GTF, d'une part car il ne se fixe pas au promoteur propre mais plusieurs centaines de paires de bases en amont, et d'autre part car il n'est pas toujours nécessaire pour assurer un niveau basal de transcription *in vivo*. Une des particularités les plus notables de Mediator est son extraordinaire complexité : pas moins de 30 sous-unités protéiques y participent, dont certaines de façon tissu-spécifique. Les modules dits de tête et du centre peuvent interagir avec les GTF et l'ARN Polymérase II, tandis que la queue peut s'associer avec



différents facteurs capables de se lier à l'ADN, comme des récepteurs nucléaires (tels les récepteurs à l'oestrogène ER ou aux gluco-corticoïdes GR) ou des effecteurs finaux de voies de transduction (tels le facteur de croissance transformant TGF $\beta$ , la  $\beta$ -caténine ou encore NANOG). L'activation transcriptionnelle est ainsi coordonnée de façon site-spécifique par Mediator, qui est en outre capable de recruter la machinerie de modification de la chromatine via l'histone acétylase p300. Le module à activité kinase s'associe avec le reste du complexe de façon plus labile, et joue un rôle régulateur grâce à sa capacité à masquer le site d'affinité pour l'ARN polymérase II sur les parties de tête et du centre. Mediator est ainsi capable de directement coordonner l'action de la machinerie de modification de la chromatine et l'assemblage du PIC. Une telle situation a par exemple été reportée lorsque le dimère formé du récepteur à l'hormone thyroïdienne et du récepteur à l'acide rétinoïque (TR-RXR $\alpha$ ) se lie à l'ADN, et interagit avec Mediator qui peut alors recruter l'histone acétylase p300 d'une part et TFIID d'autre part (Black et al., 2006).

Si pendant longtemps Mediator a été considéré comme un co-facteur activateur de l'assemblage du PIC, plusieurs études récentes ont étendu son rôle à d'autres étapes – activatrices ou inhibitrices – de la transcription. Ces fonctions non-canoniques sont déployées en fonction de la nature du facteur qui recrute Mediator, qui en retour sert de plateforme d'ancrage pour différents effecteurs. Il a ainsi été rapporté que dans certains contextes cellulaires, Mediator peut être recruté par le facteur REST (Repressor Element 1 Silencing Transcription [factor], affiné pour le motif ADN RE1) et orchestrer l'action de répresseurs transcriptionnels, comme l'histone H3K9 méthyl-transférase G9a (Ding et al., 2008). Dans ce cas, le module kinase décrit précédemment est particulièrement important puisqu'il permet la liaison avec REST et G9a, et inhibe l'engagement avec le PIC.

Le complexe Mediator apparaît ainsi jouer le rôle d'un concentrateur qui intègre de nombreuses afférences régulatrices, de façon locale d'une part – les promoteurs affectés seront ceux avoisinant un élément génomique lié à un facteur affiné pour une de ses sous-unités – et diversifiée d'autre part – son effet sur la transcription sera dépendant de ses co-facteurs : ceux-là même qui lui servent d'ancrage à l'ADN ainsi que ses nombreux autres partenaires diffusibles et souvent dépendants du contexte cellulaire.

### **3. Les régions de contrôle distales**

La littérature distingue de façon arbitraire des éléments génomiques régulateurs proximaux et distaux, en référence à la distance qui les sépare du promoteur propre qu'ils affectent. Ces critères de classification varient cependant grandement suivant les auteurs et l'organisme considéré. On

parle parfois même de promoteur proximal pour la kilobase directement en amont du promoteur propre mais, en dehors de leur faible éloignement du promoteur, ces régions ne sont pas forcément structurellement différentes des éléments régulateurs plus distaux. Il existe en réalité une telle diversité dans la nature et le fonctionnement des séquences régulatrices proximales et distales qu'il est très difficile de les classer sur une autre base que la distance génomique qui les sépare du (ou des) promoteur(s) propre(s) qu'elles affectent.

### 3.3. Structure et fonction des régions régulatrices distales

#### 3.3.1. Diversité fonctionnelle et nomenclature des régions distales régulatrices

Contrairement aux régions régulatrices proches du promoteur propre (<1kb), une région génomique distale peut être impliquée dans le contrôle – parfois opposé – de plusieurs promoteurs propres. L'altération d'un même élément peut ainsi, par exemple, prévenir la transcription à partir d'un promoteur donné alors qu'un autre promoteur se trouvera suractivé. De plus, comme mentionné précédemment, un même élément génomique peut influencer positivement ou négativement sur la transcription des promoteurs environnants suivant la nature des facteurs *trans* qui lui sont liés. Ces propriétés empêchent *de facto* une simple classification sur la base de leur effet sur la transcription. La littérature récente fait pourtant référence à l'existence d'amplificateurs (traduction de l'anglais *enhancer*) et répresseurs (traduction de l'anglais *silencer*) transcriptionnels dès que le taux de transcrits produits par un promoteur jugé cible est respectivement augmenté ou diminué. Une stricte classification de ces éléments en référence à leur effet sur la transcription est en réalité très difficile. *Stricto sensu*, celle-ci ne devrait pas référer aux effets indirects de ces éléments, et le terme *d'enhancer* ou de *silencer* ne devrait donc être attribué qu'une fois le mécanisme précis élucidé. Celui-ci étant rarement pleinement appréhendé, il est prudent d'utiliser ces termes en explicitant le promoteur sur lequel l'effet est mesuré. Pour illustration, dans le cas des amplificateurs transcriptionnels (*enhancer*) de transcrits précurseurs de microARN (Fukao et al., 2007), la délétion d'un élément génomique régulateur interagit génétiquement, négativement, avec la cible de ce microARN mais ne devrait pas pour autant être considéré comme un élément répresseur (au sens de *silencer*) de cette cible.

Pour ces raisons nous n'utiliserons les termes d'éléments amplificateur ou répresseur que dans le contexte où le promoteur sur lequel l'effet transcriptionnel est mesuré est évident ; pour référer à l'élément physique en question nous préférons employer le terme générique d'élément

génomique non-codant, ou région régulatrice quand l'effet sur la transcription est sous-entendu. Nous reviendrons sur ce point au cours de ce chapitre d'introduction.

Par ailleurs, très peu d'études se sont consacrées à la caractérisation systématique des effets de l'altération d'un élément génomique (non promoteur) sur la transcription des loci environnants. Ce fait est à la source de la vision très « promoteur-centrique » que la littérature donne de ces éléments régulateurs, qui tient son origine dans le fait historique que, de façon générale, la région codante d'un locus d'intérêt est d'abord étudiée en détail avant que la recherche d'éléments régulateurs ne soit entreprise. En outre, les connaissances sur les mécanismes précis par lequel un élément distal donné influe sur la transcription d'un ou plusieurs promoteurs sont très limitées, et peuvent parfois provenir de l'observation d'effets très indirects. En effet, certains éléments génomiques peuvent être impliqués dans la *cis*-régulation de la transcription à des étapes ultérieures à l'assemblage du PIC comme l'élongation ou l'épissage – dont l'altération peut influencer très fortement sur le taux de transcrit, comme dans le cas de l'élément *Repeat A* dans la région correspondant au premier exon de *Xist* (Royce-Tolland et al., 2010) par exemple – ou jouer un rôle dans le contrôle d'autres aspects métaboliques et structuraux du génome comme la réplication ou la position dans le noyau (Ahmed et al., 2010).

La diversité des mécanismes de régulation de la transcription par des éléments génomiques distaux aux promoteurs propres sera le sujet de la section suivante, mais nous allons tout d'abord discuter des caractéristiques structurales de ces régions.

### **3.3.2. Identifier des effets régulateurs à distance grâce à la transgénèse**

La prévalence de la contribution de l'environnement génomique dans le contrôle de l'activité transcriptionnelle a été rendue particulièrement évidente de par l'observation que ni le taux de transcrit ni la spécificité tissulaire ne sont généralement correctement déterminés par de petits transgènes composés seulement du promoteur propre et de l'ADN complémentaire du transcrit correspondant. Ces effets sont attribuables à deux facteurs différents, souvent concomitants et parfois liés: l'absence dans le transgène d'éléments de contrôle du locus d'intérêt, introniques ou distaux, qui assurent normalement sa régulation dans le contexte endogène ; la soumission du transgène aux influences régulatrices de son environnement génomique ectopique (dits effets de position, au sens large). L'utilisation de grandes constructions transgéniques, comme des chromosomes artificiels bactériens (BAC), de levure (YAC) ou dérivés du bactériophage P1 (PAC), contenant le locus d'intérêt et plusieurs dizaines – voire même centaines – de kilobases et parfois

nécessaire pour la récapitulation adéquate du patron transcriptionnel en contexte ectopique (pour revue voir Giraldo et Montoliu, 2001).

L'approche de transgénèse a été particulièrement utilisée dans le cas des loci produisant des transcrits dont l'expression est soumise à empreinte parentale (pour revue voir Tunster et al., 2011), où plusieurs équipes ont cherché à identifier les divers centres de contrôle permettant de conférer l'expression mono-allélique, avec les loci *Igf2/H19* et *Igf2r/Airn* initialement étudiés de la sorte (Ainscough et al., 1997 ; Wutz et al., 1997).

Il peut arriver que la récapitulation du patron transcriptionnel ne soit que partielle, en particulier lorsqu'un même promoteur est soumis à plusieurs éléments de contrôle garantissant par exemple son activité dans plusieurs tissus distincts. Le locus *Pax6* en est un exemple bien documenté (pour revue voir Kleinjan et Lettice, 2008). De manière similaire, alors que l'expression de *Xist* au cours du développement murin préimplantatoire est correctement assurée par des transgènes de 200kb (Okamoto et al., 2005), ces constructions ne permettent pas son expression dans les tissus embryonnaires issus de la masse interne (Heard et al., 1996).

La manipulation de grands fragments d'ADN est délicate, et il est fréquent que les YAC ou les BAC recombinent avant ou pendant leur intégration dans le génome. La caractérisation de l'intégrité de ces constructions une fois intégrées est en outre particulièrement difficile de par l'existence de copies homologues endogènes. De plus, la nécessité de générer systématiquement des organismes transgéniques est un frein considérable à ce type d'approche. Le développement de méthodologie basées sur la prédiction d'éléments fonctionnels sur la base de propriétés structurales, que nous allons maintenant détailler, a considérablement simplifié leur identification – mais les expériences de transgénèse additive ou suppressive restent la preuve ultime de l'activité de régions de contrôle ainsi prédites.

### **3.3.3. Identification *in silico* d'éléments génomiques fonctionnels**

#### **3.3.3.1. Utilité et limites prédictives de l'approche évolutive**

Un moyen indirect de postuler qu'un élément génomique sert une fonction biologique est de mesurer s'il a subi une pression de sélection au cours de l'évolution. Cette approche a permis l'identification de nombreuses régions régulant la transcription à des loci impliqués dans l'orchestration de processus développementaux très conservés entre espèces distantes (pour revue voir Elgar et Vavouri, 2008). La force de cette approche est qu'elle repose uniquement sur l'information génomique, et peut donc permettre d'identifier des éléments régulateurs même si

ceux-ci ne sont actifs que dans certains types cellulaires très spécifiques. En revanche, elle se limite par définition aux éléments conservés et n'est donc pas appropriée pour rechercher des éléments génomiques contrôlant un processus développemental évolutivement récent ou divergent, comme le développement cardiaque par exemple (Blow et al., 2010). Ceci peut également être expliqué par la co-évolution des séquences *cis*-régulatrices et des facteurs *trans* qui s'y lient (pour revue voir Simpson et Ayyar, 2008). Dans le cas de l'inactivation du chromosome X, il est connu que des éléments cruciaux à la régulation de l'expression *Xist* chez la souris ne sont pas conservés chez l'humain (comme *Tsix/Xite* – pour revue voir Migeon, 2003). De telles différences pourraient sous-tendre les différences du patron d'expression de *Xist* rapportées entre ces espèces (Okamoto et al., 2011).

### **3.3.3.2. ENCODE et l'annotation structurale du génome**

L'ère du projet ENCODE (ENCyclopedia Of Dna Elements), démarré en 2003, visant à la caractérisation détaillée des éléments fonctionnels du génome humain (The ENCODE Project Consortium, 2011 pour revue), marque un revers conceptuel dans l'appréciation des éléments génomiques régulateurs de la transcription. Cette approche repose sur la description de plusieurs propriétés structurales du génome (composition et structure de la chromatine, fixation de facteurs protéiques, structure et niveau d'expression des ARNs etc...) dans plusieurs types cellulaires, rendue possible par le développement des techniques de puces à ADN et plus récemment de séquençage massif. Ainsi, il est aujourd'hui si aisé de procéder à l'identification des sites de fixation de facteurs protéiques que la caractérisation des propriétés structurales du génome – souvent regroupées sous le chapeau d'épigénome – précède généralement la caractérisation fonctionnelle de ces éléments. L'accomplissement de la phase pilote du projet ENCODE, qui s'est concentrée sur 1% du génome humain, a ainsi illustré que moins de la moitié seulement des DHS correspondent à un site d'initiation de la transcription. En effet, de façon surprenante, la majorité des DHS ont été cartographiés à plus de 2,5kb d'un TSS connu, démontrant l'existence d'un nombre étonnement élevé d'évènements de liaison de facteurs protéiques à des sites distaux aux promoteurs – autant de sites pouvant potentiellement représenter des éléments régulateurs de la transcription.

### **3.3.3.3. La notion de signature épigénomique**

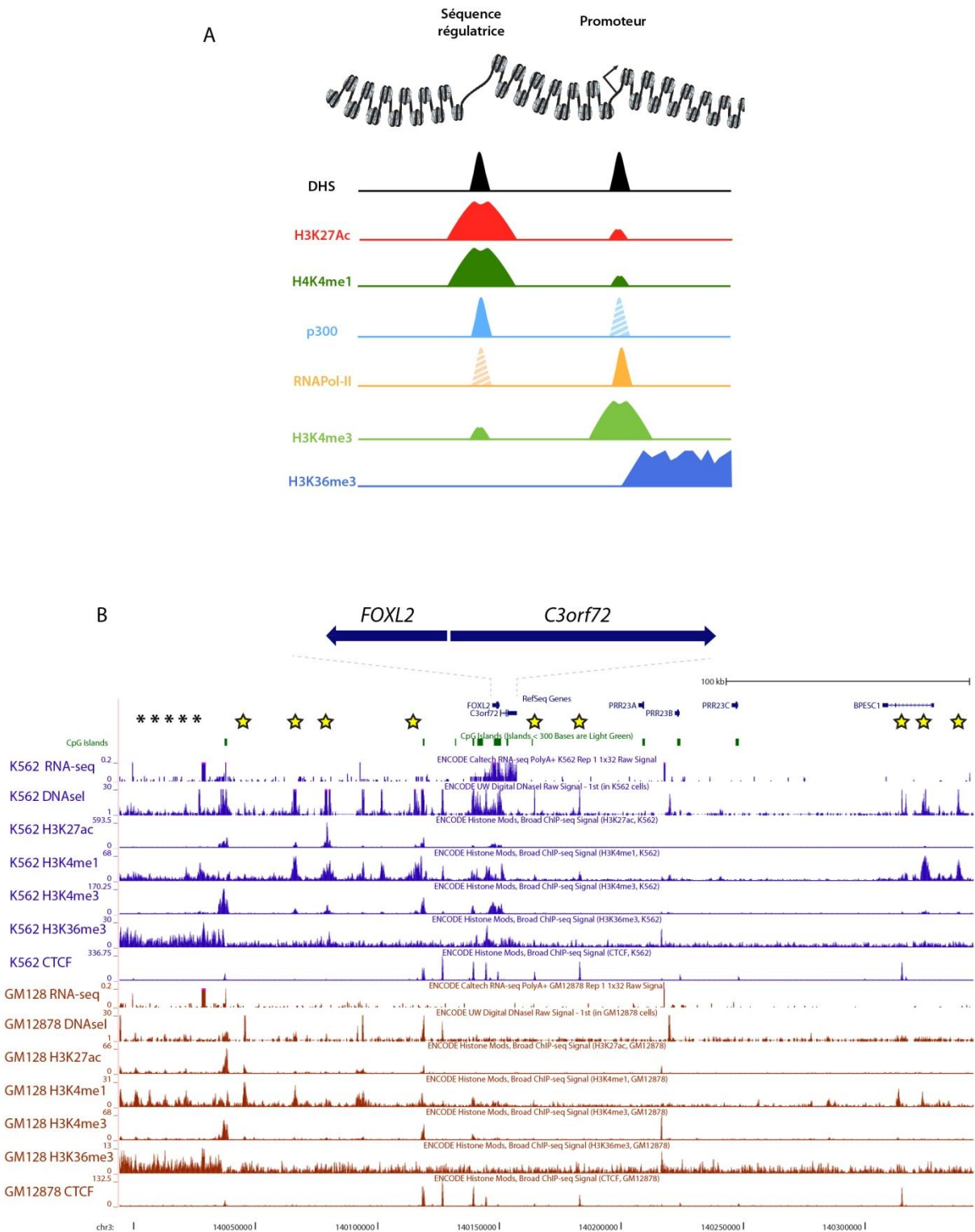
La méthodologie ENCODE est également particulière en ce qu'elle appelle à l'intégration des différentes propriétés ainsi caractérisées. De manière intéressante, il est ainsi apparu que les

nucléosomes autour des DHS correspondant aux promoteurs sont fréquemment porteurs de résidus lysines triméthylés en position 4 de l'histone H3 (H3K4me3) – en particulier pour ceux contenant un îlot CpG (Mikkelsen et al., 2007) – contrairement aux DHS distincts des promoteurs qui eux sont généralement caractérisés par un résidu mono-méthylé (H3K4me1). Certains DHS non promoteurs ayant historiquement été caractérisés comme modulateurs de la transcription, il a été proposé que le potentiel régulateur d'éléments génomiques pourrait être prédit *ab initio* à partir de leur signature épigénomique (Heintzman et al., 2007 ; pour revue voir Schübeler, 2007). Ces études ont également identifié le complexe co-activateur de la transcription acétyl-transférase p300-C/EBP $\beta$  comme marqueur de l'état actif des régions régulatrices amplificatrices de la transcription, dont l'utilité pour leur identification à l'aveugle fut démontré par la suite (Visel et al., 2009 ; Blow et al., 2010 ; Ghisletti et al., 2010). D'autres caractéristiques structurales de ces éléments ont par ailleurs récemment été identifiées en fonction de l'état d'activité de ces séquences, comme la présence de H3K27Ac (Creyghton et al., 2010, Rada-Iglesias et al., 2011) – retrouvée lorsqu'elles sont activement en train de stimuler la transcription – de certains facteurs liant l'ADN de façon site spécifique (Ghisletti et al., 2010) et de l'ARN polymérase II elle-même (De Santa et al., 2010) – qui y est parfois associée à la production de courts transcrits instables et non poly-adenylés (Kim et al., 2010). Ces études illustrent en outre la diversité structurale (dynamique) de ces éléments génomiques ainsi que leur similarité avec les promoteurs, propriété déjà soulignée par des études ciblées à plusieurs loci comme  $\beta$ -*GLOBIN* (Vieira et al., 2004), *Hoxd4* (Rastegar et al., 2004) et *VpreB1* (Szutorisz et al., 2005a - voir Szutorisz et al., 2005b pour revue). Un exemple de carte épigénomique du locus FOXL2 dans deux lignées cellulaires hématopoïétiques illustre comment le principe d'annotation structurale permet l'identification de régions génomiques contenant potentiellement des éléments régulateurs tissu-spécifiques (voir **figure 4**).

#### **3.3.3.4. Prédire la dynamique des effets régulateurs au cours du développement**

Lancé à la suite de la première phase ENCODE, le projet modENCODE (Celniker et al., 2009 et Elsner et Mak, 2011 pour revue) vise à appliquer la méthodologie du projet pilote chez deux organismes multicellulaires modèles : *Drosophila melanogaster* et *Caenorhabditis elegans*. Ces projets ont ainsi développé des techniques d'analyses multivariées permettant de corrélérer la composition de la chromatine, la liaison de facteurs protéiques à l'ADN et plusieurs paramètres relatifs à la réplication ainsi qu'à la production et le métabolisme des ARN à travers plusieurs stades de développement et types cellulaires (The modENCODE Consortium et al., 2010 ; Gerstein et al., 2010). Si l'interprétation de ces informations structurales en termes de régulation transcriptionnelle

est encore un défi, des méthodes basées sur l'analyse combinatoire de la liaison de plusieurs facteurs à l'ADN ont déjà été utilisées avec succès pour prédire la dynamique tissulaire et développementale de l'activité régulatrice d'éléments génomiques (Zinzen et al., 2009).



**Figure 4. Signature épigénomique (A)** Les sites impliqués dans le contrôle de la transcription se distinguent par leur hypersensibilité à la DNaseI (DHS). Les sites régulateurs peuvent se distinguer structurellement des promoteurs par plusieurs caractéristiques, comme H3K27AC, H3K4me1, la présence de p300 et parfois de la RNA-Pol-II ; les promoteurs actifs quant à eux peuvent parfois fixer p300, sont liés à l'ARN Pol-II et sont généralement enrichis en H3K4me3. La séquence transcrite est généralement décorée par H3K36me3 qui peut varier localement, en particulier au niveau des sites d'épissages. **(B)** La construction de cartes épigénomiques multidimensionnelles peut ainsi permettre d'annoter le génome dans plusieurs types cellulaires, comme par exemple ici au locus *FOXL2* dans les cellules érythroïdes humaines K562 (pistes bleues) et lymphoblastoïdes GM12878 (pistes rouges). On remarque que la transcription de *FOXL2* dans la lignée K562, détectable sur la piste de RNA seq, s'accompagne de maintes différences structurales dans son voisinage génomique (marquées par ☆) par rapport à la lignée GM128 où *FOXL2* est silencieux. Ce type de carte épigénomique peut aussi permettre d'identifier des sites de transcription non annotés, comme celui visible à l'extrême gauche (\*\*\*\*). (A) inspirée de Schübeler, 2007 ; (B) visualisation données Chip-seq, RNA-seq et DNaseI-seq ENCODE sur le navigateur de génome UCSC GB.

### 3.3.3.5. Des régions régulatrices distales... jusqu'à quel point ?

La manipulation des organismes modèles permettant une approche de génétique inverse performante, les techniques de génétique directe (traduction de l'anglais *forward genetics*) sont de moins en moins utilisées pour cartographier des séquences *cis*-régulatrices. La dissection fonctionnelle des éléments régulateurs par des méthodes tels que le crible de capture de séquences amplificatrices de la transcription (traduction de l'anglais *enhancer trap*) sont tombées en désuétude au profit d'approches basées sur l'ingénierie génétique dirigée – bien que ces techniques soient toujours utilisées pour la création de banques d'organismes transgéniques permettant de diriger l'expression d'ARN ou de protéines d'intérêt de façon très tissu-spécifique. Les approches de génétique directe sont cependant toujours d'actualité chez l'humain, chez qui le suivi médical permet la détection efficace des pathologies à déterminisme génétique. C'est ainsi que de nombreux syndromes ont été associés à l'altération d'éléments génomiques régulant l'activité de séquences codantes (pour revue voir Kleinjan et Lettice, 2008 ou Maston et al., 2006). Les cas les plus surprenants concernent probablement *SHH* (*SONIC HEDGEHOG*) et *SOX9* (*SEX DETERMINING REGION Y BOX 9*), dont l'activité est altérée spécifiquement dans certains tissus à la suite de micro-délétions distantes de plus d'une mégabase de ces loci, causant respectivement une polydactylie aux mains et les syndromes de la séquence de Pierre Robin (Lettice et al., 2002 ; Benko et al., 2009). Il est aujourd'hui encore inconnu si ces exemples représentent des cas extrêmes de régulation à longue distance, ou si ce sont les illustrations les plus flagrantes de mécanismes fondamentaux de



modulation de l'activité transcriptionnelle. Que des transgènes aussi grands que 460kb ne soient pas capables de spécifier le patron d'expression de *Xist* suggère que de la mise en place de l'inactivation repose sur de tels mécanismes de régulation à très longue distance (Heard et al., 1999).

### **3.3.4. Peut-on vraiment déduire la fonction d'un élément de par sa signature épigénomique ?**

La simple inspection des cartes épigénomiques dressées dans des modèles cellulaires récapitulant la mise en place de l'inactivation du chromosome X, telles les cellules souches embryonnaires murines, pourrait donc fournir un moyen efficace de prédire des éléments régulateurs du locus *Xist/Tsix*. Cependant, il est important de garder à l'esprit que si les propriétés structurales de la chromatine peuvent en effet révéler l'existence de régions fonctionnelles, toutes les régions fonctionnelles connues ne sont pas identifiables par une signature chromatinienne particulière. En outre, bien que la structure et la dynamique des épigénomes puisse maintenant être étudiée rapidement et en grands détails, les analyses fonctionnelles sont quant à elles beaucoup plus lentes car elles nécessitent souvent la transgénèse. De plus, les mécanismes moléculaires à l'origine des patrons chromatiniens utilisés comme signature ne sont pas bien compris, et leur diversité nous échappe encore. De manière plus générale, il apparaît important de garder à l'esprit que si le résultat des études dites fonctionnelles permet de conclure quant au mode d'action des éléments génomiques étudiés dans le contexte expérimental utilisés, il est parfois délicat d'inférer des propriétés générales sur la base de telles observations. De plus, une fonction n'est généralement assignée qu'à la suite d'un test dédié (*i.e.* régulation de la transcription, période de réplication, instabilité génétique etc...) sans être déduite à l'aveugle par une expérience unique. Puisque l'assignation d'une fonction à une région génomique, par exemple en termes de régulation de la transcription, est ainsi effectuée sur la base de phénotypes observés, il convient de garder à l'esprit qu'une expérience différente, ayant mesuré une caractéristique phénotypique différente, aurait pu lui assigner une fonction différente. Un élément génomique donné ne possède donc vraisemblablement pas une fonction unique, et il est donc utile – voir essentiel – de se souvenir du cadre expérimental dans lequel un élément génomique d'intérêt a été initialement caractérisé.

Plus précisément, les expériences dites de caractérisation fonctionnelle consistent généralement à tester qualitativement et parfois quantitativement la *cis*-activation transcriptionnelle d'un rapporteur possédant soit le promoteur endogène cible présumé, soit un promoteur minimal hétérologue. L'introduction ectopique de grands segments contenant différentes versions du locus d'intérêt, portant notamment des mutations aux sites candidats permet également de tester

l'importance fonctionnelle d'un élément génomique dans des conditions où l'intégrité globale du locus d'intérêt est mieux respectée (Kleinjan et al., 2001). La mesure des changements de l'activité transcriptionnelle du locus endogène après délétion d'éléments candidats est également parfois possible. Du fait des difficultés présentées par la création d'organismes transgéniques, la première de ces approches est souvent utilisée au détriment des deux dernières. Bien que rapides et relativement informatives, les méthodes basées sur des rapporteurs ne permettent cependant pas d'exclure un rôle fonctionnel en cas de résultat négatif, en particulier lorsque le mécanisme de régulation en jeu repose sur la synergie de plusieurs éléments dispersés ou si l'élément ne fonctionne qu'en contexte d'une chromatine correctement assemblée. L'approche d'ingénierie de loci ectopiques résout ce problème et rend possible l'analyse combinatoire de différents éléments mais, comme leur insertion dans le génome est rarement contrôlée, la comparaison entre les différents mutants est souvent difficile – notamment à cause des effets de position, des insertions en multicopie et des remaniements aléatoires inhérents à la manipulation de larges fragments d'ADN. Enfin, les effets de la délétion au locus endogène ne sont pas non plus forcément toujours conclusifs car de nombreux éléments régulateurs peuvent agir de façon redondante, et une fonction cryptique n'est parfois révélée qu'en contexte synthétique, ou dans des contextes environnementaux ou fonds génétiques particuliers (comme le cas des régions *shadow enhancers* Frankel et al., 2010).

Ces difficultés expérimentales sont en grande partie responsables du peu de certitudes établies quant aux mécanismes employés par les régions régulatrices pour contrôler les multiples étapes de la transcription. Encore une fois, les analyses fonctionnelles permettent seulement de conclure quant à la régulation du phénotype mesuré. Il est certain que le développement de nouvelles procédures permettant de tester d'autres propriétés élargira considérablement la liste des fonctions attribuées aux éléments non codants du génome.

### **3.1.4. Le rôle des séquences régulatrices dans le contrôle de la dynamique des patrons de transcription**

#### **3.1.4.1. Au cours du développement**

Une observation critique à la compréhension des mécanismes par lesquels des régions génomiques peuvent réguler l'activité transcriptionnelle à distance concerne le fait que l'effet régulateur n'est pas forcément exercé de façon constitutive. Ont ainsi été identifiées de nombreuses séquences génomiques dont l'altération peut n'avoir aucun effet mesurable sur l'activité

transcriptionnelle d'un locus dans certains tissus, alors qu'elle abolit toute transcription dans d'autres. Dans le centre d'inactivation du chromosome X, le locus *Xite* est par exemple connu pour stimuler la transcription de *Tsix* seulement au cours de la différenciation précoce, mais pas à l'état indifférencié (Ogawa et Lee, 2003). Un autre exemple bien documenté est MFCS1, séquence régulatrice du locus *Sonic Hedgehog* (*Shh*) mentionné précédemment, dont la délétion n'affecte l'expression de SHH que dans les bourgeons de membres embryonnaires (Sagai et al., 2005). D'autres séquences régulatrices contrôlent la production de ce morphogène dans d'autres lignages épithéliaux (Sagai et al., 2009). Cette propriété pourrait être expliquée par l'expression tissu-spécifique de facteurs *trans* capables de se fixer sur ces séquences, ne leur conférant ainsi un effet régulateur que dans certains types cellulaires (voir **figure 3**). Selon cette hypothèse, l'expression ectopique de ces facteurs devrait alors être suffisante pour activer la transcription des cibles de ces éléments régulateurs.

Une telle situation a en effet été rapportée pour la protéine PU.1, capable de lier directement l'ADN. L'expression forcée de ce facteur dans des fibroblastes ou des progéniteurs myéloïdes, alors qu'il n'est normalement exprimé que dans la lignée hématopoïétique (et de façon particulièrement élevée dans les macrophages), induit des changements structuraux autour de ses sites de fixation, où apparaît la signature chromatinienne typique d'éléments génomiques régulateurs actifs (déplétion en nucléosome et enrichissement local de H3K4me1 ; Ghisletti et al., 2010 ; Heinz et al., 2010). Ces changements s'accompagnent en outre de l'expression d'ARNm spécifiques de la lignée hématopoïétique, bien qu'à moindre niveau, et de propriétés cellulaires caractéristiques des macrophages (Feng et al., 2008). Le paysage transcriptionnel de ces fibroblastes est cependant loin d'être identique à celui des macrophages, et le fait que de nombreux loci où PU.1 se lie ne se trouvent pas transcriptionnellement activés suggère que d'autres facteurs interviennent. Il a été proposé que l'histoire développementale d'une cellule puisse expliquer ce phénomène.

#### **3.1.4.2. Engrenage épigénétique**

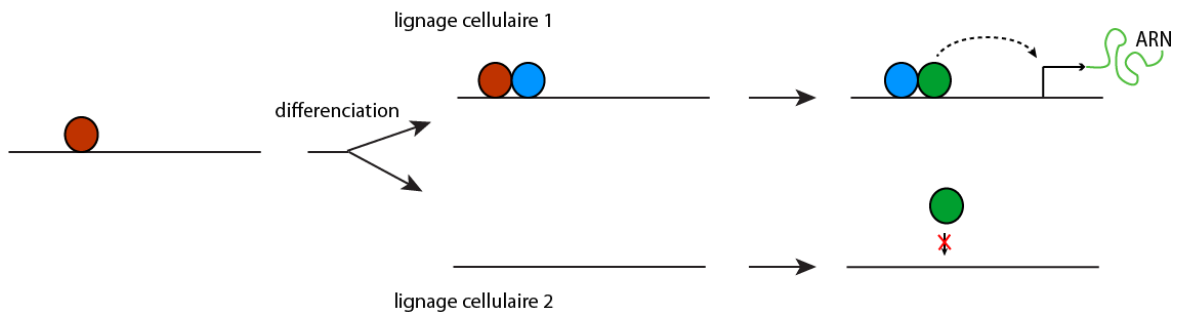
Cette idée s'appuie entre autre sur les expériences de transfert nucléaire et fusion cellulaire qui ont démontré que les facteurs diffusibles présents dans une cellule somatique receveuse sont au mieux capables d'affecter légèrement, mais jamais de remodeler entièrement le patron transcriptionnel du noyau de la cellule somatique donneuse. Ces observations suggèrent que la capacité transcriptionnelle d'un locus peut être contrôlée par des propriétés intrinsèques à la chromatine. Il a ainsi été proposé que l'activation transcriptionnelle tissu-spécifique d'un locus est la conséquence d'interactions protéine-ADN dont le contexte chromatinien a été rendu permissif à la

suite de transitions structurales qui peuvent avoir été amorcées dès les premiers stades de développement. Dans ce modèle, des facteurs pionniers seraient responsables de l'initiation de changements structuraux de la chromatine – tels la déméthylation de l'ADN à leur site de fixation, la modification post-traductionnelle d'histones ou encore l'ancrage de co-facteurs – qui seraient ensuite transmis à travers les générations cellulaires, permettant à des activateurs transcriptionnels tissu-spécifiques de gagner accès au locus dans les types cellulaires où ils sont exprimés. Ces particularités structurales au cours du développement, dans certains tissus par exemple, préviendraient ainsi définitivement toute transcription même si l'activateur transcriptionnel venait à être exprimé (pour revue voir Smale, 2010). Ce phénomène, parfois résumé sous l'appellation d'engrenage épigénétique (traduction de *epigenetic priming*), a été invoqué pour expliquer la transcription tissu-spécifique de plusieurs loci, dont le point commun est d'être sous contrôle d'éléments génomiques distaux marqués par la présence d'une structure chromatinienne généralement associée avec la permissivité transcriptionnelle, à des stades développementaux qui précèdent leur activation transcriptionnelle (Szutorisz et al., 2005a ; Walter et al., 2008 ; Xu et al., 2007 ; Xu et al., 2009 ; Liber et al., 2010). La généralité de ce phénomène reste cependant à explorer. Reste en outre à démontrer formellement que la compétence transcriptionnelle de ces loci dépend réellement de l'établissement préalable des caractéristiques structurales décrites au niveau des éléments génomiques régulateurs incriminés (voir **figure 5**).

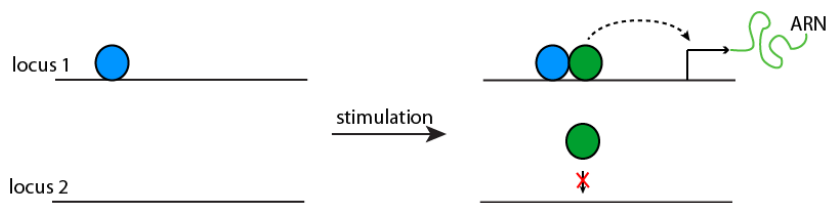
Le modèle d'engrenage épigénétique est séduisant en ce qu'il pourrait expliquer pourquoi l'expression ectopique de facteurs *trans* ne suffit généralement pas à rapidement convertir un type cellulaire en un autre. L'importance biologique de ce phénomène doit cependant être relativisée au vu de la démonstration que la surexpression d'une poignée de facteurs peut, dans certains cas, induire une conversion phénotypique presque totale, même si la fréquence de conversion est extrêmement basse (par exemple moins de 0,02% après près de trois semaines de sélection dans le cas de l'induction de cellules souches pluripotentes murines par la surexpression, *via* transduction, de SOX2, KLF4, C-MYC et OCT3/4, Takahashi et Yamanaka, 2006 ; Huang et al., 2011 ; pour revue voir Masip et al., 2010). De manière intéressante, les cocktails de facteurs utilisés dans les diverses expériences de *trans*-différenciation incluent systématiquement un ou plusieurs facteurs capables d'interagir avec la machinerie de remodelage de la chromatine (par exemple KLF4 interagit avec p300, et C-MYC avec des SWI/SNF), suggérant que la *trans*- ou dé-différenciation est le résultat de la combinaison d'une action inductive (par les facteurs exprimés de façon ectopique et le milieu de culture cellulaire) et sélective (par le temps mais également par les conditions de culture). Ces expériences ne permettent donc pas d'exclure formellement l'existence du phénomène d'engrenage

épigénétique, dont l'importance biologique pourra être évaluée par la caractérisation plus systématique de la transcription tissu-spécifique.

#### A - Engrenage épigénétique au cours du développement



#### B - Engrenage épigénétique lors de la réponse rapide à un stimulus



**Figure 5 L'engrenage épigénétique.** Dans certaines situations, des facteurs *trans* ne peuvent se lier à un élément génomique donné que si celui-ci est déjà lié par un ou plusieurs autres facteurs. Cette propriété est à la base du phénomène d'engrenage épigénétique, où la présence de ces facteurs de permissivité est héritée au travers des générations cellulaires, même en absence du facteur activateur et donc de l'activité du promoteur cible. **(A)** Des exemples ont ainsi été rapportés où, la fixation d'un facteur pionnier (rouge), à des étapes précoces du développement, permet le recrutement subséquent d'autres facteurs (bleu), seulement dans les types cellulaires où il est exprimé. Ces facteurs secondaires peuvent ainsi servir de relai une fois que l'expression du pionnier cesse, si bien qu'après différenciation l'expression pourtant ubiquitaire d'activateurs (vert) n'a d'effet que dans certains lignages, mais pas dans d'autres. **(B)** L'engrenage épigénétique est aussi en jeu lors de l'induction transcriptionnelle rapide, où la présence de facteurs potentialisants (bleu), antérieure au stimulus, détermine la capacité d'un locus à répondre au signal régulateur (vert).

#### 3.1.4.3. L'étude de changements transcriptionnels rapides accrédié l'idée d'engrenage épigénétique à court terme

À côté des processus de différenciation cellulaire au cours du développement existent des cas d'induction transcriptionnelle beaucoup plus rapides. Les voies de signalisation par récepteurs nucléaires offrent par exemple l'opportunité de suivre la dynamique structurale de la chromatine qui

accompagne des changements transcriptionnels importants sur une échelle de temps très courte, de l'ordre de l'heure pour les hormones stéroïdes par exemple. Il a été proposé que la vitesse de l'activation transcriptionnelle repose sur le fait que les récepteurs hormonaux pourraient eux-mêmes agir en tant que facteurs pionniers, en initiant d'eux-mêmes la décompaction de la chromatine nécessaire à l'assemblage du PIC. De manière intéressante, une étude récente supporte au contraire l'existence d'un phénomène global d'engrenage épigénétique. En effet, l'identification par ChIP-seq des sites de fixation du récepteur aux glucocorticoïdes (GR) après une heure d'exposition à la dexaméthasone révèle que ces régions sont, pour la très grande majorité, déjà hypersensibles à la DNaseI avant stimulation (John et al., 2011). De manière surprenante, l'analyse de plusieurs types cellulaires connus pour activer la transcription des jeux de loci différents après stimulation révèle invariablement que l'accessibilité à la DNaseI précède la fixation du récepteur nucléaire. Il apparaît ainsi que les propriétés structurales tissu-spécifique de la chromatine pourraient prédéterminer la réponse transcriptionnelle à la stimulation hormonale.

Cette hypothèse est en outre étayée par une autre étude concernant cette fois la réponse inflammatoire de macrophages exposés au lipopolysaccharide (LPS) pendant seulement quelques heures (Ghisletti et al., 2010). En effet, la fixation de p300 après stimulation dans des régions génomiques distales aux TSS – indicative de l'activation de l'effet régulateur de ces régions – est observée sur plusieurs centaines d'éléments génomiques, dont une fraction importante est déjà porteuse de la signature chromatinienne d'éléments régulateurs (enrichissement en H3K4me1 mais pas en H3K4me3) avant stimulation. La fixation de p300 s'accompagne en outre de la liaison de plusieurs facteurs liant l'ADN seulement après stimulation, comme NF- $\kappa$ B ou des membres de la famille des IRF (dont la capacité à lier l'ADN est régulée par des modifications post-traductionnelles). Cette observation suggère donc que la fixation des activateurs transcriptionnels pourrait être facilitée par une structure de la chromatine préexistante à l'induction. La spécificité de cette réponse pourrait donc être, au moins en partie, due aux facteurs responsables de l'établissement et de la maintenance de ces propriétés structurales. De manière intéressante, dans les macrophages naïfs, la présence de la protéine PU.1 – présente de façon constitutive aux sites où la fixation de p300 est induite après stimulation – apparaît nécessaire à la maintenance de l'enrichissement local en H3K4me1 et de l'état décompacté de la chromatine autour des régions génomiques qu'elle est capable de lier. Par ailleurs, l'expression de PU.1 dans les macrophages est nécessaire pour l'activité amplificatrice de la transcription de certaines de ses cibles d'une part et pour l'expression de marqueurs spécifiques des macrophages d'autre part. Enfin, comme il a été mentionné plus haut, l'expression ectopique de PU.1 dans des fibroblastes est suffisante pour entraîner des modifications de la chromatine à ses loci cibles, ainsi que l'expression de marqueurs spécifiques des macrophages

(Feng et al., 2008). Il a ainsi été proposé que l'expression du facteur PU.1 spécifiquement dans les macrophages pourrait être responsable de la capacité de certaines régions du génome à répondre à la stimulation inflammatoire, en contrôlant des propriétés structurales de la chromatine autour de ses régions cibles.

L'induction transcriptionnelle apparaît donc être la résultante de l'action combinée de facteurs potentialisateurs (parfois tissus-spécifiques comme PU.1) et de facteurs effecteurs (parfois constitutifs mais à activité inductible comme GR, NFκB etc...), à des régions génomiques qui sont souvent distinctes des sites où est *in fine* initiée la transcription. Se pose donc maintenant la question de comprendre par quel(s) mécanisme(s) ces régions distales sont capables de moduler l'activité transcriptionnelle de promoteurs à distance, de qui elles peuvent être séparées par plus d'une mégabase.

### **3.4. Activité et mécanismes d'action des régions de contrôle de la transcription**

#### **3.4.1. Considérations générales**

D'un point de vue phénoménologique, les éléments du promoteur propre ou de son voisinage direct sont impliqués dans la détermination de la fréquence de transcription à partir d'un promoteur (Yunger et al., 2010). En revanche, les éléments distaux amplificateurs de la transcription semblent fonctionner différemment et contrôlent plus généralement la probabilité d'activation d'un promoteur sans pour autant montrer d'effet important sur la quantité de transcrit produit par un promoteur actif (Ko et al., 1990 ; Walters et al., 1995 ; voir Fiering et al., 2000 pour revue). À l'échelle de la population cellulaire, différents éléments activateurs peuvent ainsi donner l'impression que le taux d'un transcrit donné augmente de façon rhéostatique alors qu'en réalité il s'agit du nombre de cellules où le transcrit est produit qui devient plus important. Cette différence est fondamentale car, comme nous le discuterons en détail, elle peut expliquer l'apparition de différences phénotypiques importantes entre les membres d'une population cellulaire soumis à un même stimulus environnemental.

En termes mécanistiques, les régions régulatrices proximales peuvent exercer leur effet directement sur un promoteur propre, par exemple *via* le positionnement d'un nucléosome ou plus indirectement par nucléation d'assemblages protéiques (Herendeen et al., 1992). D'autres mécanismes sont également possibles. Dès la découverte des premières régions régulatrices de la transcription il a été proposé, initialement d'après des résultats obtenus chez les bactéries, que la

structure topologique de la fibre d'ADN pourrait permettre l'association physique de sites distants d'un point de vue génomique, permettant ainsi l'interaction structurale de leurs partenaires protéiques respectifs, et sous-tendant ainsi leur interaction fonctionnelle (pour revue des résultats séminaux voir Ptashne, 1986). Cette proposition a initialement été formulée pour rendre compte d'observations à l'opéron bactérien *araBAD* (arabinose), où l'effet régulateur de ces protéines n'est efficace que si le promoteur et le site opérateur sont distants d'un multiple de la période de la double hélice d'ADN – orientant ainsi ces éléments génomiques du même côté de cette hélice et permettant donc l'interaction des protéines effectrices qui y sont liées, même si leurs points d'attache sur le génome sont séparés de plusieurs centaines de paires de bases (Dunn et al., 1984). De nombreuses autres études, dont les fameux travaux aux opérons gal (Irani et al., 1983) et au prophage lambda (Hochschild et Ptashne, 1986 ; Griffith et al., 1986) pour ne citer qu'eux, ont depuis abouti à la conclusion qu'un contact physique entre éléments génomiques distaux peut exister, non seulement sur de courtes distances génomiques comme chez les procaryotes mais aussi au travers de plus longues distances chez les eucaryotes, et en particulier dans les grands génomes des mammifères. De manière importante, ces travaux ont rapidement souligné la relation à double sens entre les caractéristiques conformationnelles de l'ADN d'une part et les facteurs protéiques s'y liant d'autre part ; ceux-ci sont en effet généralement nécessaires pour ponter les séquences auxquels ils se fixent ainsi que pour le maintien de cette topologie particulière. En retour, les propriétés structurales de ces facteurs sont également influencées par l'augmentation de leur concentration locale qui découle du rapprochement physique de plusieurs de leurs sites de liaison. Cette dernière propriété est particulièrement importante en termes biologiques car les protéines liant l'ADN sont très généralement capables d'homo- ou hétéro- multimérisation. Des régions d'ADN physiquement associées *via* les interactions physiques des facteurs protéiques qui leur sont liés représentent ainsi un site privilégié pour la nucléation d'assemblages macromoléculaires. Par ailleurs leur dynamique apporte l'opportunité d'une modulation des effets coopératifs entre les sous-unités liant directement ou indirectement les séquences impliquées dans la boucle d'ADN. Les propriétés qui peuvent émerger d'une telle situation dans le cas de complexes capables de réguler la transcription sont nombreuses (pour revue voir Vilar et Saiz, 2005), et incluent notamment des situations documentées de modulation du bruit transcriptionnel et de la variabilité phénotypique entre les cellules d'une même population.

Pendant près de vingt-cinq ans le modèle de régulation transcriptionnelle en *cis* par formation de boucles chromatiniennes locales (*looping*, ou modèle de contact) a été étayé par des évidences fonctionnelles indirectes (discutées en détail dans Ptashne, 1986) ou structurales mais *in vitro* (Hochschild et Ptashne, 1988). Ce n'est qu'en 2002 que le développement de la technique de



Capture de Conformation Chromosomique (3C - Dekker et Klecker, 2002, inspiré de la méthodologie de Cullen et al., 1993 et Mukherjee et al., 1988; voir **figure 6** Dekker et Klecker, 2002) et son application à la dissection de l'organisation topologique du locus  *$\beta$ -globin* murin (Tolhuis et al., 2002) a formellement prouvé l'existence de contacts physiques *in vivo* entre éléments génomiques endogènes inter-régulés.

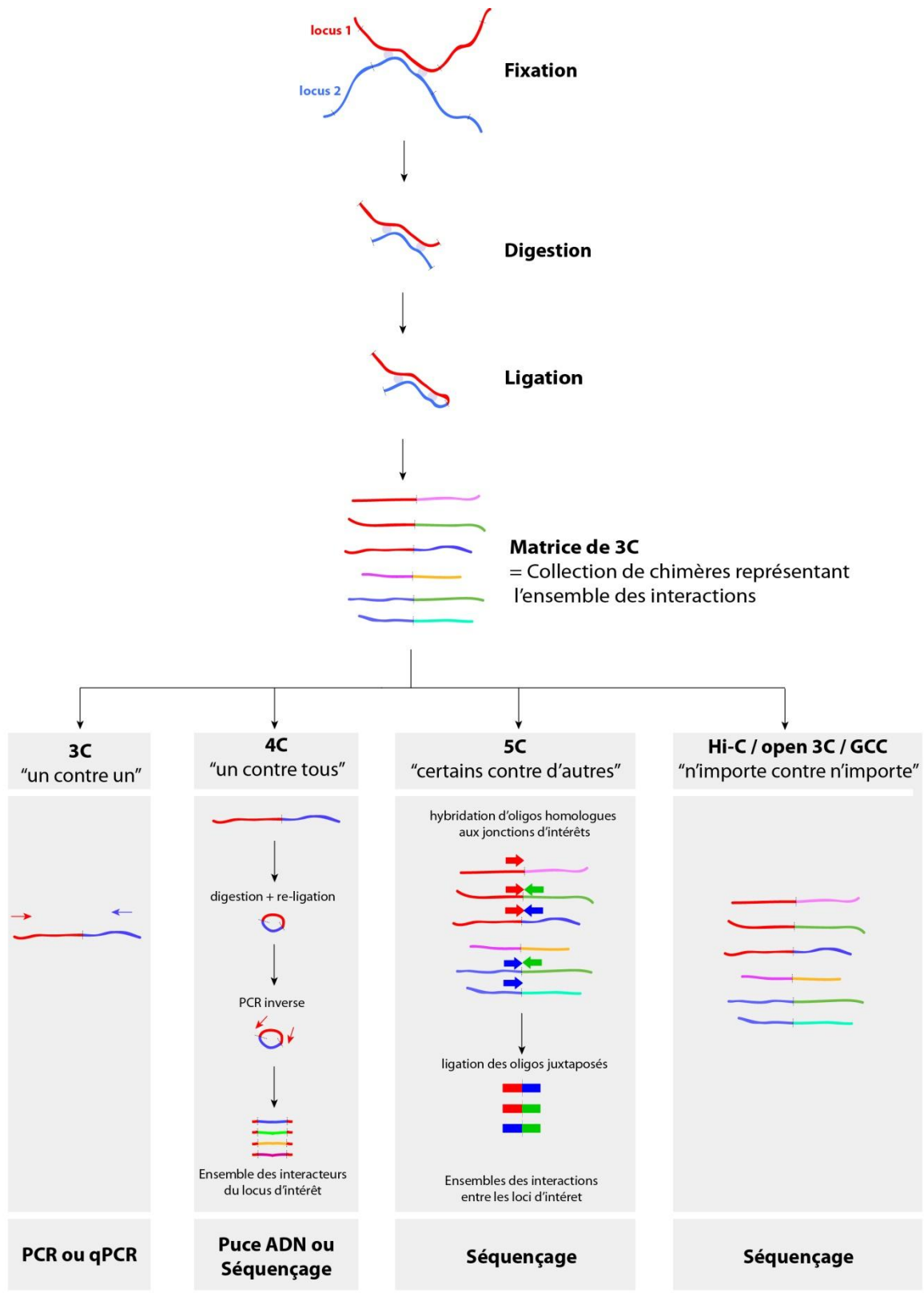
**Figure 6 La 3C.** Le principe de la **Capture de Conformation Chromosomique (3C)** repose sur la fixation de la chromatine dans sa conformation native par incubation de cellules vivantes avec du formaldéhyde. Les loci qui étaient proches physiquement se retrouvent alors pontés par leurs intermédiaires protéiques (histones et autres). La digestion de cette préparation de chromatine par des enzymes de restriction permet le clivage à des sites précis et la libération de petits fragments co-fixés, aux extrémités compatibles. Par soucis de clarté une seule interaction est représentée sur ce schéma, mais il faut imaginer que le pontage au formaldéhyde peut fixer plusieurs fragments chromatiniens ensembles. À ce stade, la dilution extrême des échantillons peut permettre de favoriser la ligation intra-moléculaire, après ajout de ligase T4, des fragments chromatiniens qui ont été fixés ensemble. La déprotéinisation permet ensuite de libérer ces molécules d'ADN chimériques dont l'existence reflète l'existence de leur interaction physique initiale, et l'abondance la fréquence de cette interaction. Lors de la procédure de 3C, tout le génome est fixé, digéré et ligué, si bien que la matrice obtenue reflète l'ensemble des interactions possibles – si suffisamment de cellules sont analysées. Il est important de remarquer qu'au maximum seulement deux interacteurs peuvent être capturés par locus donné par cellule (diploïde). Il convient donc d'analyser un nombre suffisamment élevé de cellules, typiquement plusieurs millions, afin d'être sûr de ne pas sous-échantillonner l'ensemble des interactions possibles.

La quantification des fréquences d'interaction peut ensuite se faire de plusieurs manières, à plus ou moins grande échelle, suivant le but poursuivi. Dans le cas d'un nombre restreint d'interactions à tester, la procédure classique emploie la PCR ou la qPCR avec des amorces situées sur les fragments d'ADN dont on veut quantifier l'interaction.

Lorsque l'on désire quantifier toutes les interactions impliquant un locus d'intérêt, on peut employer la **3C Circulaire (4C)**, qui repose sur la re-circularisation de l'échantillon et la PCR inverse, cette fois-ci employant deux amorces divergentes dans le fragment de restriction d'intérêt.

Si l'on cherche à quantifier de nombreuses interactions en parallèle, la **Copie Conforme de 3C (5C)** peut être utilisée. On hybride alors une collection d'oligonucléotides s'hybridant aux jonctions des fragments d'intérêts. Leur ligation par Taq ligase, seulement possible si ces oligonucléotides se sont hybridés face à face, permet alors de générer des doublets d'oligonucléotides, correspondant à une représentation réduite de la matrice de 3C initiale, et ne contenant que les interactions d'intérêt qui peuvent donc alors être directement analysées par séquençage massif.

Enfin, il est possible de procéder au séquençage de la matrice de 3C, purifiée ou non, suivant plusieurs protocoles dits de **Capture de Conformation Génomique (GCC)**, **Hi-C** ou encore **open 3C**. Cette approche ne bénéficie pas de la même résolution que les précédentes, (il faudrait séquencer extrêmement profondément pour retrouver plusieurs fois la même interaction par chance) mais permet de dresser des cartes topologiques à l'échelle du génome à résolution modérée (pour les grands génomes).



Dekker et Kleckner 2002; Tolhuis et al., 2002  
qPCR: Hagège et al., 2007; Court et al., 2011

Simonis et al., 2006; Zhao et al., 2006;  
Würtele et Chartrand 2006;  
Lomvardas et al., 2006

Dostie et al., 2006

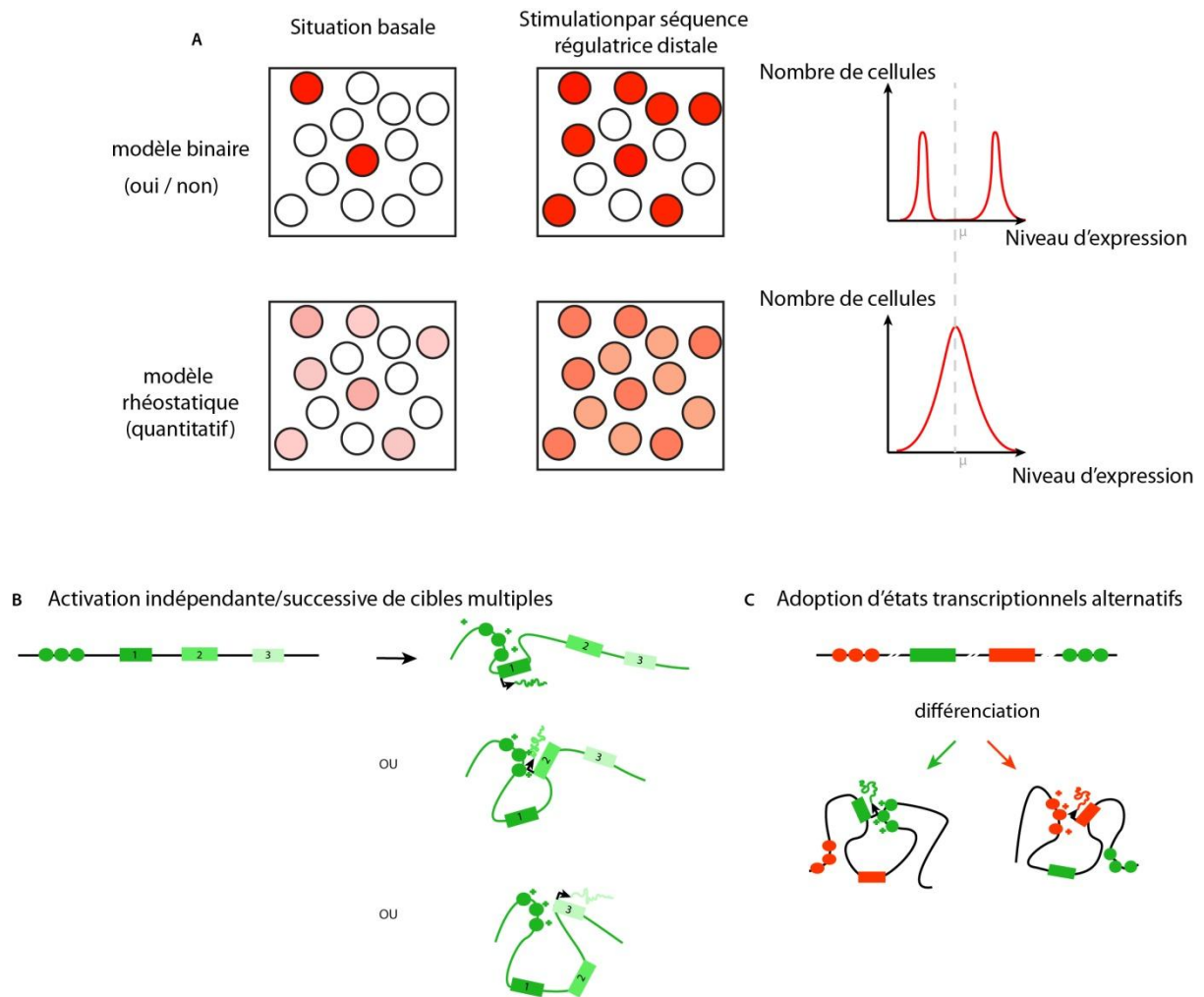
Rodley et al., 2009; Lieberman-Aiden et al., 2009  
Tanizawa et al., 2010; Duan et al., 2010

### 3.4.2. Prédications du modèle d'activation par contact

Plusieurs prédictions sont envisageables quant au mécanisme d'action des éléments génomiques nécessitant la formation de boucles chromatinienne pour exercer leur régulation. D'une part, si une région du génome n'est soumise à l'effet régulateur d'une autre région que si elles co-localisent de façon suffisamment étroite pour permettre l'interaction physique des facteurs qui leurs sont liés, le système est à même d'adopter un comportement binaire. En effet, soit l'interaction physique a lieu, et l'effet régulateur peut être exercé, soit les éléments génomiques sont trop éloignés pour interagir. Cette première propriété présente des parallèles avec la capacité des amplificateurs transcriptionnels à moduler la probabilité de la transcription, plutôt que son taux, discutée précédemment (voir **figure 7**). Une deuxième prédiction est que les facteurs qui contrôlent la conformation de la chromatine devraient jouer un rôle important dans la régulation de la transcription. Ceci a maintenant été observé dans maints systèmes pour plusieurs facteurs, tels CTCF, le complexe cohésine ou SATB1 par exemple, dont nous discuterons en détail ultérieurement (voir Gause et al., 2008 pour revue). Une troisième prédiction est que lorsque les interactions sont établies par paires, un élément régulateur ne peut moduler l'activité que d'une seule cible à la fois ; ainsi plusieurs promoteurs pourraient entrer en compétition avec un élément régulateur. Cette propriété pourrait expliquer que plusieurs promoteurs soumis à l'influence d'un même élément régulateur initient généralement la transcription de façon exclusive, au moins temporellement, comme aux loci  $\beta$ - et  $\alpha$ -globin (Gribnau et al., 1998 ; Trimborn et al., 1999 – voir **figure 7B**). Qui plus est, l'adoption de configurations chromosomiques locales alternatives pourrait également expliquer l'émergence ou le maintien de patrons transcriptionnels distincts lors de la différenciation cellulaire (voir **figure 7C** ; pour revue voir West et Fraser, 2005). L'ensemble des observations expérimentales ne sont cependant pas uniquement compatibles avec le modèle d'activation par contact, et il est important de garder à l'esprit que plusieurs processus différents peuvent sous-tendre la communication régulatrice à distance, suivant le type d'éléments génomiques en jeu et les facteurs *trans* impliqués (comme les phénomènes de nucléation locale linéaire [dits de *tracking*] ou tridimensionnelle [dits de *oozing*], pour discussion voir Bulger et Groudine, 2011).

### 3.5. Contrôle de l'activation transcriptionnelle par effet isolateur

Une question se pose alors concernant le mécanisme d'action des éléments génomiques capables de réguler la transcription à longue distance : qu'est ce qui détermine le rayon d'activité de ces séquences ? Plusieurs expériences ont démontré que la proximité physique de certains



**Figure 7 Modèles binaires et rhéostatiques ; prédictions du modèle de régulation par boucles chromosomiques.**

(A) Dans le régime binaire strict, la transcription est soit active soit inactive. Il en résulte que la population cellulaire se partage entre cellules qui transcrivent le locus en question, et d'autres qui ne le transcrivent pas. Les éléments régulateurs qui stimulent la transcription de façon binaire n'augmentent pas le taux de transcription mais la probabilité d'activation de la transcription. La régulation rhéostatique stricte quant à elle permet de moduler le taux de transcrit par cellule, qui se distribue donc de façon continue à travers la population cellulaire. Notons que la valeur moyenne du taux de transcrit dans la population ( $\mu$ , ligne hachurée verticale) est la même dans les deux cas, et que seules des expériences à l'échelle de la cellule unique permettent de distinguer les deux régimes. La littérature rapporte que les éléments génomiques sous-tendant des effets d'amplification transcriptionnelle obéissent généralement au mode binaire. L'activation transcriptionnelle par formation de contacts chromosomiques s'accommode aisément de l'invocation d'un régime binaire, puisque l'interaction a lieu ou non – sans état continu intermédiaire possible. Une composante rhéostatique pourrait cependant être introduite secondairement en fonction mécanisme précis par lequel l'interaction stimule la transcription. Plusieurs situations développementales sont explicables par la régulation binaire de la transcription via contacts chromosomiques, tels les cas d'activation séquentielle et non-simultanée de plusieurs promoteurs cibles (B), comme au locus  $\beta$ -globin, ou l'adoption d'états transcriptionnels alternatifs par configurations chromosomiques incompatibles entre elles (C). *Inspiré de Blackwood et Kadonaga 1998 ; West et Fraser 2005 ; Lidstrom et Konopka, 2010*

amplificateurs transcriptionnels actif et d'un promoteur peut être suffisante pour stimuler la transcription, même si ces éléments ne sont pas présents en *cis* sur la même molécule d'ADN. Par exemple des éléments génomiques de SV40 ou du cytomégalo virus peuvent augmenter l'activité transcriptionnelle *in vitro* d'un promoteur  $\beta$ -globin lorsque ceux-ci sont attachés de façon non-covalente par un pont avidine-streptavidine (Müller et al., 1989). Cela suggère donc que les facteurs qui contrôlent la topologie de la chromatine et des chromosomes pourraient jouer un rôle clef dans la régulation transcriptionnelle, sans exclure que l'activité de l'élément ou la transcription elle-même puisse intervenir en retour dans le contrôle de la conformation chromatinienne. Nous allons maintenant détailler les caractéristiques fonctionnelles et structurales des éléments génomiques capables d'isoler un locus des effets régulateurs exercés par son environnement épigénomique, après des remarques générales concernant l'étude de l'effet d'isolation régulatrice.

### 3.5.1. Considérations générales

La recherche de caractéristiques structurales générales permettant d'expliquer l'effet d'isolation souffre de la grande diversité fonctionnelle des éléments capables de sous-tendre un tel effet. Qui plus est, l'assignation d'un effet isolateur à un élément génomique se fait suite à des tests fonctionnels impliquant non seulement cette séquence, mais aussi la région régulatrice dont l'effet est bloqué, et le promoteur sur lequel l'effet est mesuré. L'effet observé résulte donc de la convolution des relations qu'entretiennent ces trois éléments, dont les propriétés peuvent en outre varier d'un type cellulaire à l'autre, suivant les facteurs *trans* qui se lient à chacun d'entre eux. L'effet isolateur n'apparaît donc pas comme une simple propriété constitutive et intrinsèque à un élément génomique donné. Nous tenons donc à généraliser ici une remarque mentionnée plus haut concernant l'amplification transcriptionnelle : faire référence à un élément génomique en tant qu'« isolateur » (traduction de l'anglais *insulator*) n'a de sens que dans le contexte où cet effet prend place. Il est ainsi difficile d'utiliser cette dénomination pour référer de façon générale à un élément génomique capable d'isolation puisque cet effet ne s'exerce que dans certains contextes. La compréhension de l'effet d'isolation nécessite donc la dissection des relations tripartites entre l'élément isolant, les régions environnantes dont l'influence régulatrice est bloquée, et le promoteur qui se trouve isolé.

### 3.5.2. Les origines de la notion d'isolation transcriptionnelle

Les premiers effets d'isolation ont été découverts au locus *87A7* chez *D.melanogaster*, sur la base de leur hypersensibilité à la DNaseI et de leur localisation aux deux extrémités d'un locus

contenant deux promoteurs inductibles par la chaleur (chacun associé avec une séquence codant la protéine HSP70 ; Udvardy et al., 1985). De manière intéressante, l'omission de ces séquences dans un transgène 87A7 entraîne la sensibilité de promoteurs rapporteurs internes aux influences régulatrices, stimulatrices ou inhibitrices, exercées par les séquences génomiques environnantes (Kellum et Schedl, 1991). Ces séquences n'agissent pas en tant qu'amplificateur transcriptionnel dominant car elles sont incapables d'activer directement en *cis* ces promoteurs. L'effet régulateur d'une région activatrice sur le promoteur *hsp70* dans des constructions ectopiques est bloqué lorsqu'une de ces séquences est présente, mais seulement si elle est insérée entre ces deux éléments (Kellum et Schedl, 1992). Cette propriété, partagée par d'autres éléments génomiques régulateurs, est appelée effet bloqueur d'amplificateur transcriptionnel (traduction de l'anglais *enhancer-blocking* – **Figure 8**), et nous allons maintenant détailler les propriétés de ce phénomène et les explications mécanistiques avancées pour en rendre compte.

### **3.5.3. Interruption de l'effet régulateur d'éléments génomiques distaux ponctuels**

Tout comme pour les activateurs transcriptionnels, l'effet régulateur d'éléments isolateurs est sous-tendu par les protéines capables de s'y fixer, de par leur affinité pour des motifs ADN particuliers, ainsi qu'à leurs (nombreux) cofacteurs secondaires qu'elles recrutent. Plusieurs des protéines critiques pour déterminer l'effet isolateur de ces éléments chez *D. melanogaster* n'étant pas conservées chez les vertébrés, nous nous concentrerons sur les études concernant ce dernier taxon, bien que celles-ci fussent historiquement largement inspirées des travaux chez la drosophile. Le premier élément à activité isolatrice identifié chez les mammifères fut l'homologue de *cHS4*, une région du locus  *$\beta$ -globin* aviaire hypersensible à la DNaseI dans les érythrocytes (Chung et al., 1993 ; Li et Stamatoyannopoulos, 1994), dont la capacité à bloquer en *cis* l'effet d'un amplificateur de la transcription dépend de sa liaison avec CTCF (Bell et al., 1999)

#### **3.5.3.1. L'isolation transcriptionnelle à *Igf2/H19***

La recherche de sites de liaison de CTCF dans le génome mammifère a rapidement abouti à l'identification d'une autre séquence aux propriétés isolatrices, bloqueuses d'amplification transcriptionnelle en l'occurrence, au locus *Igf2/H19* murin (Kanduri et al., 2000 ; Bell et Felsenfeld, 2000 ; Hark et al., 2000 ; Szabó et al., 2000 ; Kaffer et al., 2000). Ce locus est particulièrement complexe de par la présence de plusieurs amplificateurs transcriptionnels agissant de façon tissu-spécifique et allèle-spécifique, dont l'orchestration aboutit à l'expression sous empreinte dans

certaines organes d'Igf2 (codant pour le facteur de croissance ressemblant à l'insuline n°2 INSULIN-LIKE GROWTH FACTOR 2), H19 et d'autres ARN non codants récemment caractérisés (Bertheaux et al., 2008 ; Court et al., 2011a). La région différentiellement méthylée dans la lignée germinale (gDMR pour *Differentially Methylated Regions [between male and female gametes]*), correspondant à la région déterminant l'expression sous empreinte de Igf2 et H19 (ICR pour *Imprinting Control Region*), est située entre le promoteur *H19* et les promoteurs *Igf2* et comporte plusieurs sites de liaison à CTCF. CTCF n'est pas affiné pour l'ADN méthylé et n'est retrouvé que sur l'ICR maternelle, où il est nécessaire pour bloquer l'influence régulatrice d'amplificateurs transcriptionnels situés au-delà de H19. La transcription d'*Igf2* à partir de l'allèle paternel apparaît ainsi être la résultante de l'absence cet effet d'isolation. En effet, l'altération génétique des sites de liaison à CTCF dans l'ICR maternelle induit l'expression d'*Igf2* en *cis* dans certains tissus. D'élégantes études de CHIP et de 3C permettant de différencier l'allèle paternel de l'allèle maternel ont démontré que la topologie de la chromatine à ce locus est différente suivant son origine parentale (Murrell et al., 2004 ; Hagège et al., 2006 ; Kurukuti et al., 2006; Yoon et al., 2007 Qiu et al., 2008 ; Engel et al., 2008 ; Court et al., 2011a). De manière générale, les amplificateurs transcriptionnels actifs en fonction du tissu considéré apparaissent en contact physique avec le promoteur qu'ils régulent, et l'isolation d'*Igf2* des effets régulateurs environnants s'accompagne de l'association physique de ses extrémités 5' et 3' avec l'ICR seulement sur l'allèle maternel. De manière importante, ces caractéristiques sont en outre affectées par la délétion de l'élément génomique porteur des sites de liaison à CTCF. Il est cependant difficile de conclure quant à un rôle architectural direct et systématique de CTCF sur la simple base de ces observations, étant donné que cette délétion s'accompagne de nombreux autres changements structuraux au locus *Igf2/H19* qui pourraient *in fine* être à l'origine la perte de ces associations chromosomiques. L'étude de nombreuses altérations génomiques du locus montrent en effet que plusieurs régions *cis*-régulatrices différentes peuvent intervenir en fonction du type cellulaire (pour revue voir Arney, 2003) et que les facteurs liés à l'ICR maternelle sont vraisemblablement responsables de plusieurs caractéristiques structurales établies de façon hiérarchique au cours du développement (Lopes et al., 2003).

### 3.5.3.2. L'organisation structurale du locus *$\beta$ -globin*

L'étude du locus  *$\beta$ -globin* murin a également impliqué CTCF dans la formation de contacts chromosomiques entre de multiples éléments génomiques. La liaison de ce facteur à différents DHS du locus a permis d'identifier plusieurs éléments génomiques capables de bloquer – de façon modérée toutefois – l'influence du site m5'HS2 sur un promoteur hétérologue dans des

constructions reportrices (Farrell et al., 2002 ; Bulger et al., 2003). L'hypersensitivité à la DNaseI et la présence de CTCF à ces sites varie en fonction du type cellulaire et corrèle avec l'échafaudage d'une topologie faisant coalescer physiquement ces différentes régions. Cette configuration existe dans les cellules érythroïdes immatures – à l'exception des promoteurs *β-globin*, qui ne sont pas encore actifs à cette étape de différenciation – suggérant qu'elle n'est pas en elle-même capable de déterminer l'état d'activité des promoteurs du locus (Palstra et al., 2003). L'expression de CTCF est nécessaire au maintien de ces associations chromosomiques, et la mutation de son site de liaison dans un des DHS (*m3'HS1*) diminue fortement la fréquence d'interaction de cet élément avec les autres DHS du locus (Splinter, 2006). De manière intéressante, cette mutation ne s'accompagne ni d'une diminution du taux d'ARN  $\beta$ -Globin ni de l'activation ectopique des promoteurs voisins ( $\epsilon\gamma$  et  $\beta H1$  encodent la chaîne B de l'hémoglobine embryonnaire et, lorsqu'ils sont transcrits, sont sous contrôle de la même région de contrôle que les loci encodant les chaînes B chez l'adulte ; les *mOlf* encodent des récepteurs olfactifs exprimés seulement dans certains neurones). Qui plus est, la collection d'éléments régulateurs présente dans la région de contrôle (LCR pour *Locus Control Region*) du locus *β-globin* contacte toujours physiquement les promoteurs *β-major* dans ces cellules érythroïdes matures mutantes. Ces observations (pour revue voir Palstra et al., 2008a) ont mené à la proposition que si CTCF peut bien être dans certains cas nécessaire à l'organisation structurale de la chromatine des régions auxquelles il se fixe, sa présence n'est pas forcément associée à un effet isolateur – ni même régulateur – de la transcription. Nous tenons cependant à faire remarquer que ces conclusions ont été émises au vu des effets de l'altération du site *m3'HS1*, dont la conséquence sur la régulation du locus sont minimales ; il sera intéressant d'explorer la relation entre la structure chromosomique sous-tendue par les DHS liés à CTCF plus internes au locus *β-globin* et dont le rôle régulateur est avéré.

### 3.5.3.3. CTCF et les effets d'isolation transcriptionnelle à travers le génome

L'intérêt général récent pour CTCF tient probablement du fait que la présence de cette protéine semble être un point commun à tous les éléments génomiques capables d'isoler un promoteur des influences régulatrices exercées par les régions voisines. La cartographie des sites CTCF à travers le génome permet d'identifier plus de 10.000 sites de liaison, largement invariants entre les types cellulaires (Kim et al., 2007 ; Chen et al., 2008). Tous ces sites représentent-ils vraiment des instances d'isolation régulatrice ? De manière intrigante, il a été rapporté que le taux de transcrits provenant de loci séparés par une région génomique conservée comprenant une séquence consensus CTCF est moins corrélé que ceux qui sont juxtaposés sans consensus CTCF



intervenant (Xie et al., 2007). Cette observation a mené à la suggestion que les sites CTCF pourraient fonctionner généralement en tant qu'isolateurs en partitionnant le génome en blocs à l'activité transcriptionnelle indépendante (Cuddapah et al., 2009). Cependant, il n'est pas clair si ces séquences consensus correspondent aux régions effectivement liées par CTCF identifiables par ChIP-seq, ni dans quelle proportion CTCF serait un élément déterminant de ce phénomène ou seulement une de ses signatures. La protéine CTCF est en effet capable d'interagir avec de nombreux cofacteurs dont la nature module et spécialise l'effet régulateur imposé par CTCF (Zlatanova et Caiafa, 2009).

Afin de discriminer ces éventualités, les effets transcriptionnels découlant de l'inactivation de CTCF ont récemment été étudiés de façon plus systématique. De manière surprenante, l'inactivation conditionnelle de CTCF dans les bourgeons de membres embryonnaires n'entraîne qu'un effet modéré sur le paysage transcriptionnel, avec la dérégulation – qui plus est modeste – de moins de 5% des transcrits analysés (Soshnikova et al., 2010). Les loci correspondants ne se distinguent pas par une position génomique particulière, et les promoteurs proches d'amplificateurs transcriptionnels actifs établis ne semblent pas être activés ectopiquement. L'effet de la déplétion de CTCF a cependant un effet drastique sur la survie cellulaire et les bourgeons de membres mutants dégénèrent rapidement – observation qui supporte l'idée d'un rôle essentiel de CTCF mais via le contrôle d'un nombre limité de cibles directes (Gomes et Espinosa, 2010 ; Wan et al., 2008).

Au vu de toutes ces données, nous pensons qu'il est prématuré de vouloir conclure à un rôle général et direct de CTCF dans la spécification du rayon d'action des éléments capables d'amplification transcriptionnelle. Aucune étude n'a à ce jour utilisé de concert la prédiction d'éléments amplificateurs d'une part et isolateurs d'autre part afin de pouvoir évaluer la perturbation de leurs relations fonctionnelles, en terme de régulation transcriptionnelle, après inactivation de CTCF. L'interprétation de telles analyses est cependant rendue difficile de par la méconnaissance générale des mécanismes avec lesquels les amplificateurs stimulent la transcription. Le développement d'approches globales, reposant sur l'intégration des observations obtenues par les récentes techniques de run-on, ChIP et 3C à l'échelle du génome entier, représente une voie prometteuse pour avancer vers l'élucidation de ce problème.

#### **3.5.3.4. Mécanistique des effets de blocage**

Comment des contraintes topologiques exercées sur la chromatine pourraient-elles interférer avec la communication entre éléments génomiques ? Dans le cas d'amplificateurs transcriptionnels nécessitant un contact physique direct avec leur promoteur cible pour moduler la transcription, il est

aisé d’imaginer comment le partitionnement des deux éléments dans deux boucles chromatiniennes différentes pourrait sévèrement diminuer leur fréquence de contact, et donc abroger leur interaction fonctionnelle. Quant aux amplificateurs transcriptionnels exerçant leur influence par nucléation, ils pourraient trouver la propagation de leur effet stoppée par des complexes macromoléculaires stablement assemblés sur la chromatine. Bien qu’attractifs, ces modèles trouvent peu de support univoque dans la littérature – probablement parce que les mécanismes par lesquels les amplificateurs transcriptionnels stimulent la transcription restent encore largement à explorer. Par ailleurs, les résultats de 3C reflètent un comportement statistique moyenné sur plusieurs millions de cellules, faisant ainsi parfois se superposer des configurations incompatibles mais co-existantes à l’échelle de la population cellulaire. Il est tentant mais souvent incorrect – comme le montre l’imagerie (Sandhu et al., 2009 ; Bantignies et al., 2011) ou la modélisation (Baù et al., 2010) – de les interpréter comme reflétant simplement un comportement uniforme de la fibre de chromatine.

La compréhension de l’effet bloqueur suppose donc en plus de posséder au préalable des informations mécanistiques quant aux influences régulatrices auxquelles le promoteur d’intérêt est directement soumis. Pour cette raison il est difficile même aujourd’hui d’inférer des généralités concernant les mécanismes sous-tendant le blocage d’activité régulatrice par des éléments génomiques ponctuels. Les récents développements permettant des analyses de 3C à haut débit, que nous détaillerons plus loin, seront sans nul doute instrumentaux pour tester la généralité du modèle de ségrégation des boucles chromatiniennes. Enfin, il est clair que la caractérisation du contenu protéique conférant à un élément génomique la capacité de l’isoler des influences régulatrices environnantes sera un prérequis indispensable à toute valeur explicative de ces modèles mécanistiques.

Les mécanismes de contrôle présentés dans cette section apparaissent ainsi capables de confiner les influences régulatrices sur l’activité transcriptionnelle dans des domaines génomiques discrets. Ces observations ont ainsi fait émerger l’idée que le génome pourrait être partitionné en segments contigus plus ou moins étendus, dont les frontières pourraient éventuellement varier en fonction du type cellulaire, et où l’activité transcriptionnelle résulterait de l’intégration de multiples afférences régulatrices. Il apparaît donc important de considérer l’environnement génomique d’un locus afin de comprendre pleinement sa régulation transcriptionnelle. Appréhender le contrôle de l’expression de Xist nécessite donc de décrypter les relations structurales et fonctionnelle que ce locus entretient avec le reste du centre d’inactivation du chromosome X. Nous allons maintenant commenter ce qui est su à propos des effets régulateurs régionaux, et discuter de la généralisation de la notion de voisinage génomique en termes de régulation transcriptionnelle.

## 4. Voisinage génomique et régulation transcriptionnelle

Les études de l'effet de position, principalement à travers ses manifestations bigarrées (traduction de l'anglais *variegation* ; PEV pour *Position Effect Variegation*) chez la drosophile (Muller, 1930), ont mené à l'élaboration du modèle archétypal de répression transcriptionnelle par propagation de structures hétérochromatiques (pour revue générale voir Girton et Johansen, 2008). Ce concept a inspiré tellement d'idées et de travaux concernant la régulation transcriptionnelle à distance qu'il nous semble important de les discuter ici, bien que sa description originale ne soit pas basée sur des observations faites chez les vertébrés. Au lieu d'en faire un historique détaillé nous avons choisi de les commenter à la lumière de plusieurs observations récentes qui incitent à la reconsidération de plusieurs aspects de ce modèle. Nous commenterons ensuite un des modèles les plus paradigmatiques de la régulation transcriptionnelle à distance chez les vertébrés, les clusters *Hox*, qui illustrent l'intrication extrême des multiples afférences régulatrices qui parcourent le génome des mammifères.

### 4.1. L'effet de position

Les deux idées principales à la base du modèle d'effet de position par propagation de structures hétérochromatiques répressives sont les suivantes. D'une part, lorsque des réarrangements chromosomiques sont à l'origine des effets bigarrés, ceux-ci impliquent systématiquement des régions euchromatiques d'un côté et des régions d'hétérochromatine constitutive de l'autre – notamment péri-centromériques. Qui plus est, la fréquence d'inactivation transcriptionnelle apparaît inversement corrélée avec la distance du point de cassure. D'autre part, la plupart des mutations augmentant ou diminuant l'effet bigarré affectent des loci contrôlant la production d'enzymes du métabolisme de la chromatine, suggérant que l'instruction de la répression dépend de sa composition ou de son organisation (Csink et Henikoff, 1996).

#### 4.1.1. Reconsidération du mécanisme de propagation linéaire

L'existence d'une structure stable et répressive qui se propagerait en *cis* à partir de régions péri-centromériques (Locke et al., 1988) apparaît cependant difficilement compatible avec la mobilité des protéines de la chromatine, dont l'association avec l'hétérochromatine cytologique est hautement dynamique, même en interphase (Cheutin et al., 2003 ; Cheutin et al., 2004). Ces idées, même avant leur démonstration expérimentale, ont mené à la proposition que l'effet de position pourrait résulter de la propagation d'un effet répressif non pas de façon strictement polaire mais discontinue (Henikoff, 1996). En effet, certains réarrangements chromosomiques sont accompagnés

de la répression de marqueurs distaux sans que certains marqueurs proximaux soient affectés (Talbert et Henikoff, 2000). Les idées séduisantes d'effet de position par propagation linéaire de structures chromatinienne répressives sont si communément admises qu'il est surprenant que les patrons de redistribution de ces structures n'aient été analysés de façon systématique que récemment (Vogel et al., 2009). Ces études ont rapporté plusieurs résultats surprenants. Si la cartographie de répartition de HP1 sur le chromosome réarrangé révèle, de façon attendue, un enrichissement sur la partie proximale au point de cassure péri-centromérique, ces changements sont répartis de façon très hétérogène sur les quelques centaines de kilobases concernés, et ne touchent qu'une vingtaine de séquences codantes. De manière intrigante l'enrichissement en HP1 et H3K9me2 sont les plus forts au locus *white* (Vogel et al., 2009 ; Rudolph et al., 2007), le rapporteur utilisé dans la très grande majorité des analyses de PEV. L'étude transcriptomique par puce révèle de façon surprenante qu'en réalité, seule l'expression de *white* est réprimée de façon mesurable par l'invasion de HP1 et H3K9me2 (Vogel et al., 2009) dans cette région. Ces résultats suggèrent que dans ce contexte le locus *white* a non seulement la propriété particulière de retenir HP1 et H3K9me2, mais également celle d'être particulièrement sensible à la répression transcriptionnelle par effet de position. Ces observations démontrent ainsi que les effets de régulateurs induits par de tels remaniements chromosomiques ne peuvent être simplement expliqués par la nucléation linéaire d'effets répresseurs, tel qu'il a originalement été proposé. La propension à l'établissement – ou à la rétention – de modifications structurales la chromatine, et la sensibilité de l'activité transcriptionnelle à de tels changements apparaissent ainsi extrêmement dépendantes du contexte génomique local dans lequel ils s'exercent. Cette propriété illustre l'importance de l'influence d'éléments *cis*-régulateurs dans la manifestation de l'effet de position, et qui, dans certains cas extrêmes ont même été proposés comme agissant en tant qu'obstacle à son exercice.

### **1.1.1. Isolation de l'effet de position : les effets barrière**

#### **1.1.1.1. Définitions initiales et généralités**

Les études séminales sur l'effet d'isolation que nous avons détaillées précédemment ont également révélé la capacité de certains éléments à annuler l'effet de position, en limitant considérablement la variabilité intercellulaire dans l'expression de rapporteurs physiquement liés à ces séquences. Ainsi, l'interposition des séquences distales du locus *87A7* abolit non seulement la sensibilité d'un promoteur aux influences d'amplificateurs transcriptionnels (Kellum et Schedl, 1992), mais empêche également sa soumission aux effets répresseurs qui accompagnent la juxtaposition de

séquences hétérochromatiques (Kellum et Schedl, 1991). Cet effet, attribué à l'existence d'une activité barrière, est mécaniquement distinct du bloc de l'effet régulateur de séquences *cis*-activatrices, quant à lui généralement dénommé effet bloqueur. En effet, un élément génomique capable de l'un n'est pas forcément également capable de l'autre, suggérant que les acteurs moléculaires peuvent être différents (pour revue voir Gaszner et Felsenfeld, 2006). Là où un effet bloqueur peut être mis en évidence par la capacité d'une séquence à isoler un promoteur d'un amplificateur transcriptionnel dans un contexte épisomal, l'effet barrière ne peut s'exercer que dans le contexte où un promoteur d'intérêt et ses régions environnantes sont assemblés en chromatine. Pour cette raison le test d'effet barrière requiert souvent une transgénèse plus lourde que le test d'effet bloqueur.

Il n'est pas forcément trivial de distinguer les deux effets dans le cas d'expériences utilisant des intégrations stables de constructions ectopiques ou caractérisant l'altération d'un locus endogène. En outre les deux effets peuvent parfois aller de pair, découler l'un de l'autre, ou même témoigner du même phénomène d'isolation. Comme nous l'avons expliqué, l'effet bloqueur est défini par l'altération de l'influence d'amplificateurs transcriptionnels sur l'activité transcriptionnelle d'un promoteur par la présence d'un élément génomique. Il dépend donc des relations qu'entretiennent chacun de ces trois éléments entre eux. L'effet barrière quant à lui dépend des relations régulatrices entre un promoteur, un élément sous-tendant l'activité isolatrice et leur contexte génomique.

#### **1.1.1.2. Effets barrières et chromatine**

Il est très important de garder à l'esprit que l'effet barrière a été historiquement défini sur la base de sa manifestation sur la modulation de l'activité transcriptionnelle de rapporteurs soumis à un effet de position, lui-même dû à la proximité ectopique de chromatine péri-centromérique (Kellum et Schedl, 1991). Le terme barrière ne réfère ainsi ni à l'existence de facteurs physiques ni à la rupture des patrons structuraux de la chromatine, qui peuvent par ailleurs caractériser certains éléments génomiques capables d'affranchir un promoteur des effets de position du à la juxtaposition de fragments d'hétérochromatine. Il ne fait non plus pas référence à la nature de l'effet qui sous-tend la répression transcriptionnelle imposée par la proximité inhabituelle d'hétérochromatine péri-centromérique. On trouve pourtant souvent dans la littérature le terme d'effet barrière pour référer à des éléments génomiques capables de telles propriétés – cette fois-ci indépendamment de leur effet transcriptionnel sur d'éventuels promoteurs connexes (West et al., 2004 ; Gaszner et Felsenfeld, 2006). Les termes de barrière de chromatine ou de barrière chromosomique sont

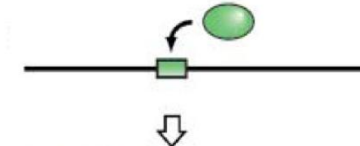
également parfois utilisés de manière plus ou moins interchangeable. Ces terminologies découlent des modèles qui expliquent l'effet barrière en invoquant la capacité de certaines séquences à empêcher la propagation de structures chromatiniennes répressives (voir pour exemple Capelson et Corces, 2004 ; **figure 8**). Comme nous allons le discuter, l'effet barrière ne peut cependant pas simplement s'expliquer par de tels mécanismes.

Plusieurs couples de séquence *cis* - facteur *trans* ont été identifiés pour déstabiliser très localement la structure de la chromatine (ex : chez la drosophile *su(hw)*, BEAF32 etc...) et parfois la transcription de rapporteurs ectopiques liés (sans inférer un effet causal entre ces deux phénomènes), mais sans que l'effet barrière en relation avec l'effet de position n'ait nécessairement été cherché à être expliqué. Cependant, comme nous l'avons commenté dans la section précédente, il apparaît en réalité que non seulement ces structures répressives ne sont pas simplement nucléées linéairement sur de grandes distances le long de la fibre chromatinienne, mais en outre que leur effet répresseur sur la transcription ne suit pas une relation déterministe explicite – même dans ces exemples prototypiques d'effets de position chez la drosophile. Il est en effet très difficile d'identifier la nature initiale des évènements menant *in fine* à la répression transcriptionnelle (inactivation ? défaut général d'activation ? etc..) des mécanismes impliqués dans la maintenance de l'état inactif, et le rôle exact des différentes caractéristiques de la chromatine y est largement débattu.

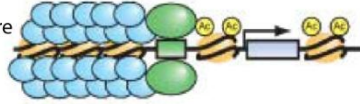
Ainsi, contrairement à ce qui est parfois discuté dans la littérature, il ne semble pas possible d'expliquer généralement l'effet barrière, en termes de modulation de l'activité transcriptionnelle, par l'existence d'éléments génomiques discrets qui seraient capables de partitionner les états structuraux de larges segments de chromatine. Au contraire, même chez les vertébrés, les effets barrières apparaissent s'exercer de façon locale, les plus grandes distances reportées étant de l'ordre de la dizaine de kilobases seulement (**figure 8** ; Lunyak et al., 2007 ; pour revue des effets barrières chez les mammifères voir Amouyal, 2010).

**A.**

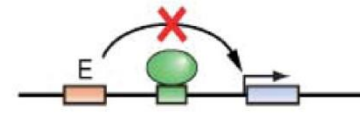
recrutement d'effecteurs



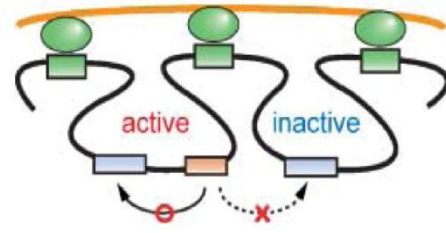
édification d'une rupture locale de la structure de la chromatine



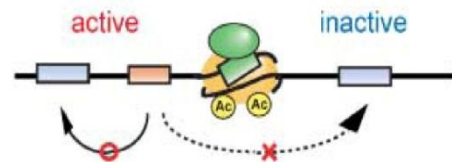
effet de bloc des influences transcriptionnelles



Architecture nucléaire



Partitionnement des influences régulatrices

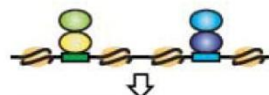


**B.**

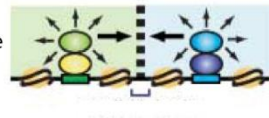
1. liaison d'un facteur *trans* pionnier sur l'ADN



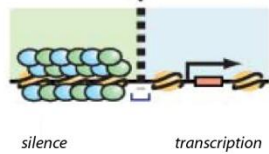
2. Recrutement de co-facteurs protéiques



3. Modification de la structure chromatinienne locale (ex: histones)

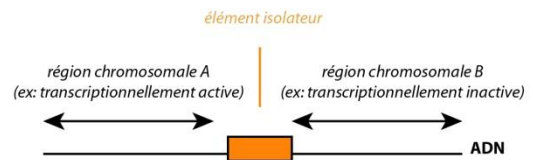


4. Modulation de l'activité transcriptionnelle locale

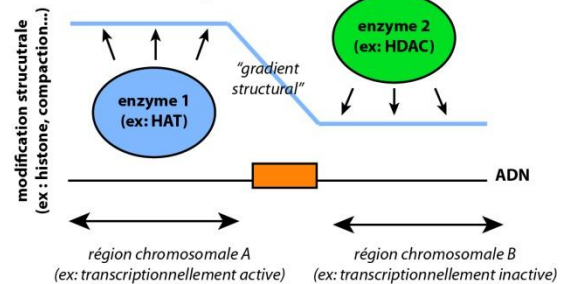


**C.**

**(i) Frontière Fixe**



**(ii) Frontière négociable**



**Figure 8 Principes du partitionnement structural et fonctionnel des chromosomes (A)** Plusieurs mécanismes ont été avancés pour expliquer les effets d'isolation transcriptionnelle, qui reposent tous sur l'introduction d'une rupture locale de la structure de la chromatine (au sens large : composition en histone, organisation topologique etc...) **(B)** Mécanismes théoriques d'édification d'une frontière structurale **(C)** La position de telles frontières peut être fixe (i), ou bien négociable (ii) suivant le mécanisme en jeu (l'exemple montré ici fait intervenir la diffusion d'histones acétylases, HAT, ou déacétylase, HDAC ; la position exacte de la frontière n'est alors pas déterminée mais résulte de la collision de ces deux activités enzymatiques opposées et peut ainsi varier de cellule à cellule ou au cours du temps). *Tiré de Kimura et Horikoshi, 2004*

#### 4.1.2.3. Barrières, frontières et bordures

Le partitionnement de la fibre chromatinienne en larges domaines structuraux discrets (>50kb) a longuement été invoqué pour inférer l'existence d'éléments génomiques ponctuels agissant comme barrières à l'envahissement d'un domaine par ses voisins. Chez les vertébrés, où ces domaines sont particulièrement abondants, certaines séquences ont été identifiées comme pouvant recruter des facteurs *trans* capables de maintenir localement des propriétés structurales particulières de la chromatine, de façon dominante sur les effets environnants. Cependant, pour les quelques cas où la délétion de ces éléments de leur contexte endogène a été étudiée, il apparaît que la structure chromatinienne du domaine correspondant n'est pas drastiquement affectée (citons les exemples mentionnés précédemment des loci *B-GLOBIN* et *Igf2*; Chan et al., 2008, Singh et Srivastava, 2008). Ces observations suggèrent que la ponctuation des domaines structuraux de la chromatine n'est pas forcément définie par des éléments génomiques précis, mais pourrait émerger de la rencontre de deux activités antithétiques. Au lieu de refléter une limite de résolution expérimentale, l'effet de gradient local souvent observé à de telles frontières pourrait ainsi être le résultat de la moyenne des configurations structurales au travers de la population cellulaire étudiée (**figure 8cii**). Ces idées ont notamment été étayées par l'observation des transitions du contenu chromatinien en histones hyper- et hypo- acétylées aux régions sub-téломériques chez *S. cerevisiae* (Kimura et al., 2002), mais n'ont pas encore été directement appliquées aux observations faites chez d'autres organismes.

#### 4.1.2. Mosaïcisme de l'effet de position

##### 4.1.2.1. Des barrières locales négociables ?

Il est intéressant de revenir sur la manifestation bigarrée de l'effet de position à la lumière de ces idées. Si nous avons discuté des influences régulatrices à longue distance en jeu dans ce phénomène, nous n'avons pas encore proposé d'explication pour la variation inter-cellulaire à l'origine de la bigarrure. Lors des études séminales chez la drosophile impliquant le locus *white*, l'inspection détaillée de la répartition des ommatidies blanches et rouges révèle leur organisation en groupes cellulaires, suggérant une héritabilité clonale des états transcriptionnels (pour revue, voir Girton et Johansen, 2008). De manière intéressante, de nombreuses mutations augmentent ou diminuent le mouchetage, sans pour autant affecter l'intensité de pigmentation dans les ommatidies rouges ni faire apparaître d'ommatidie blanches dans les groupes d'ommatidies rouges. L'effet bigarré apparaît donc résulter de variabilité dans l'activation transcriptionnelle du locus, plutôt que dans l'héritabilité mitotique de cet état. Il a ainsi été proposé que cette variation inter-cellulaire est liée à des différences structurales de la chromatine, dont les états seraient influencés par



l'hétérochromatine péri-centromérique de façon méta-stable pendant les phases mitotiques du développement (Harmon et Sedat, 2005). La position de la bordure entre ces états chromatiniens dans les cellules post-mitotiques composant l'ommatidie serait ainsi différente, déterminant l'aspect moucheté de l'œil suivant que le locus *white* est affecté ou non. L'effet de position mosaïque serait ainsi un exemple de bordure négociable, au cours du développement.

#### 4.1.2.2. Effets doses

Comment la probabilité d'activation transcriptionnelle – ou inversement de répression – est-elle contrôlée ? Comme mentionné plus haut, les mutations modifiant l'étendue de la bigarrure affectent pour la plupart l'expression de protéines impliquées dans le métabolisme de la chromatine, suggérant que la modification de sa composition ou de son organisation sont à même d'affecter la probabilité d'activation transcriptionnelle de *white*, plutôt que le taux d'expression du transcrit et de la protéine correspondante. Nombre de ces loci contrôlent la bigarrure de *white* de façon dose-dépendante : *su(var)2-5* et *su(var)3-9*, codant respectivement un homologue de HP1 et une histone H3K9 méthyl-transférase, sont par exemple connus pour leur effet haplo-suppresseur triplo-aggravateur (les drosophiles porteuses d'un seul allèle *su(var)2-5* ou *su(var)3-9* sauvage dans le contexte de la translocation de *white* ont les yeux rouges, alors que deux allèles sauvages y confèrent des yeux mouchetés et trois des yeux blancs (Eissenberg et al., 1990). L'analyse récente de la distribution de HP1 dans de telles mouches hétérozygotes pour *su(var)3-9* a montré que là encore, seuls quelques loci dans le voisinage de la translocation menant à l'effet de position sont en réalité affectés (Vogel et al., 2009). De manière intéressante, ces loci sont les mêmes que ceux qui se trouvent enrichis en HP1 lors de la translocation des séquences péri-centromériques responsables de l'effet de position sur *white*. Qui plus est, *white* est le seul locus transcriptionnellement activé lors de la réduction du dosage de *su(var)3-9*. Il ne faut donc pas voir le modèle de bordure négociable comme une variation moins robuste du modèle postulant l'existence de séquences barrant la nucléation linéaires de propriétés structurales de la chromatine, mais comme un phénomène pouvant s'exercer ponctuellement et très localement, à l'instar de l'effet barrière, mais de façon encore plus dépendante du contexte génomique puisqu'il est la manifestation de la collision des influences structurales qui s'y exercent (voir **figure 8**).

Quels sont les facteurs pouvant moduler localement la structure de la chromatine et l'activité transcriptionnelle ? Cette question revient à se demander quelle est l'origine de la capacité de certains éléments génomiques à imposer un effet régulateur localement, s'additionnant ou parfois court-circuitant l'influence du contexte épigénomique. Une réponse définitive requerrait l'annotation

systematique des éléments régulateurs et la dissection des mécanismes moléculaires sous-tendant ces effets, ce qui a été engagé par les projets modENCODE, mais est encore hors de portée d'une généralisation chez les vertébrés. De manière intéressante, les séquences dérivant d'éléments transposables apparaissent contribuer de façon extrêmement importante aux régions de contrôle de la structure et de l'activité transcriptionnelle des génomes eucaryotes (pour revue voir Slotkin et Martienssen, 2007).

## **4.2. Considérations sur le concept de séquence régulatrice**

### **4.2.1. Structure n'est pas synonyme de fonction**

Les stratégies moléculaires sous-tendant l'effet barrière sont nombreuses, et dépendent de la nature des contextes chromatinien qu'elles partitionnent. Il est important de garder à l'esprit que toute transition dans la structure ou l'organisation de la chromatine n'implique pas forcément l'existence d'un élément à effet barrière. En effet, une frontière entre deux domaines structuraux homogènes peut émerger de la collision d'influences régulatrices antagonistes – et être par conséquent plus variables en termes de positionnement. Dans les cas d'éléments génomiques conférant activement un effet barrière, l'identification des facteurs *trans* recrutés par ces éléments peut permettre d'obtenir des informations importantes quant au mécanisme en jeu au locus considéré. S'il est encore prématuré de tracer des conclusions générales chez les mammifères, les mécanismes conférant respectivement les effets amplificateurs, isolateurs, et parfois répresseurs, apparaissent très divers, au point de sembler différent dans presque chaque situation. En revanche, certains acteurs moléculaires sont clairement partagés entre les éléments conférant ces différents effets, ce qui nous amène à suggérer que la discrimination des éléments génomiques sur la base de leur effet régulateur reflète plus probablement les différences expérimentales à l'origine de leur caractérisation qu'une réelle différence structurale. Par exemple, les amplificateurs transcriptionnels peuvent initier la transcription (Kim et al., 2010 ; De Santa et al., 2010), des promoteurs peuvent conférer un effet isolateur (pour revue voir Core et Lis, 2009), tout comme un effet bloqueur d'amplificateur transcriptionnel dominant. Chacun de ces effets n'est donc pas sous-tendu par un jeu de facteurs spécifiques dont la simple présence pourrait servir à la classification fonctionnelle des éléments génomiques qui les sous-tendent. Les effets régulateurs observés apparaissent au contraire émerger de l'interaction de différents éléments. Il n'est donc au final pas forcément surprenant que leur considération séparée, dans le but d'identifier leurs propriétés intrinsèques, se heurte à l'impossibilité de tracer des relations d'équivalences entre la structure de chacun de ces éléments et la fonction, en termes d'activité transcriptionnelle, de leur ensemble.

#### **4.2.2 La classification des éléments structuraux du génome dans des catégories définies fonctionnellement est délicate**

Nous pensons ainsi qu'il est incorrect et trompeur de raisonner en termes d'éléments transcriptionnels (*enhancers*), ou isolateurs transcriptionnels (*insulators*), et voulons insister sur ce point important. Les sites génomiques en question, et leurs facteurs associés, sont des entités physiques préhensibles tandis que les effets régulateurs sont des caractéristiques subjectives déterminées par un observateur. Les caractéristiques fonctionnelles d'un élément génomique étant décrites dans un contexte expérimental, utiliser l'appellation des effets mesurés par l'expérimentateur pour référer de manière générale à cet élément mène à l'idée qu'une de leur propriété intrinsèque est, dans ce cas par exemple, de stimuler la transcription ou de bloquer les effets régulateurs environnants. Devons-nous nous étonner alors que des éléments génomiques désignés sous le même nom (*enhancers, insulators* etc...) puissent profondément différer en termes structuraux ? Dans le contexte de l'étude de la régulation transcriptionnelle, il est important de garder à l'esprit que les mots comme « activateur », « répresseur », « amplificateur », « isolateur » ne sont pas des noms mais des adjectifs, et qu'il se réfèrent à des effets mesurés expérimentalement – et non pas aux caractéristiques structurales des entités physiques qui sous-tendent ces effets. Cette remarque est d'ailleurs largement généralisable à de nombreuses autres notions, jusqu'aux termes « épigénétique » et « génétique », dont la dérive sémantique entraîne l'apparition de paradoxes conceptuels dont beaucoup peuvent cependant se résoudre d'eux-mêmes par l'emploi d'une terminologie appropriée (Fox Keller et Harel, 2007 ; Gerstein et al., 2007).

#### **4.2.3. Nécessité d'approcher le rôle des éléments génomiques dans leur contexte chromosomique**

L'effet régulateur sur la transcription d'un promoteur des régions constituant l'environnement épigénomique (les séquences et les facteurs qui leur sont liés) est qualifié d'activateur, répresseur ou isolateur par l'expérimentateur qui cherche généralement à les isoler de leur contexte natif pour leur assigner une fonction intrinsèque. Si cette démarche a été productive dans l'identification de grands types de relations régulatrices entre certains éléments du génome, elle l'a également été dans la démonstration de la nécessité de développer des approches permettant leur étude dans leur contexte endogène, où s'exerce l'entièreté des interactions fonctionnelles – en termes de contrôle transcriptionnel. Ces considérations doivent rester à l'esprit de celui qui cherche à généraliser les mécanismes à l'origine des effets régulateurs observés à un nombre restreint d'éléments

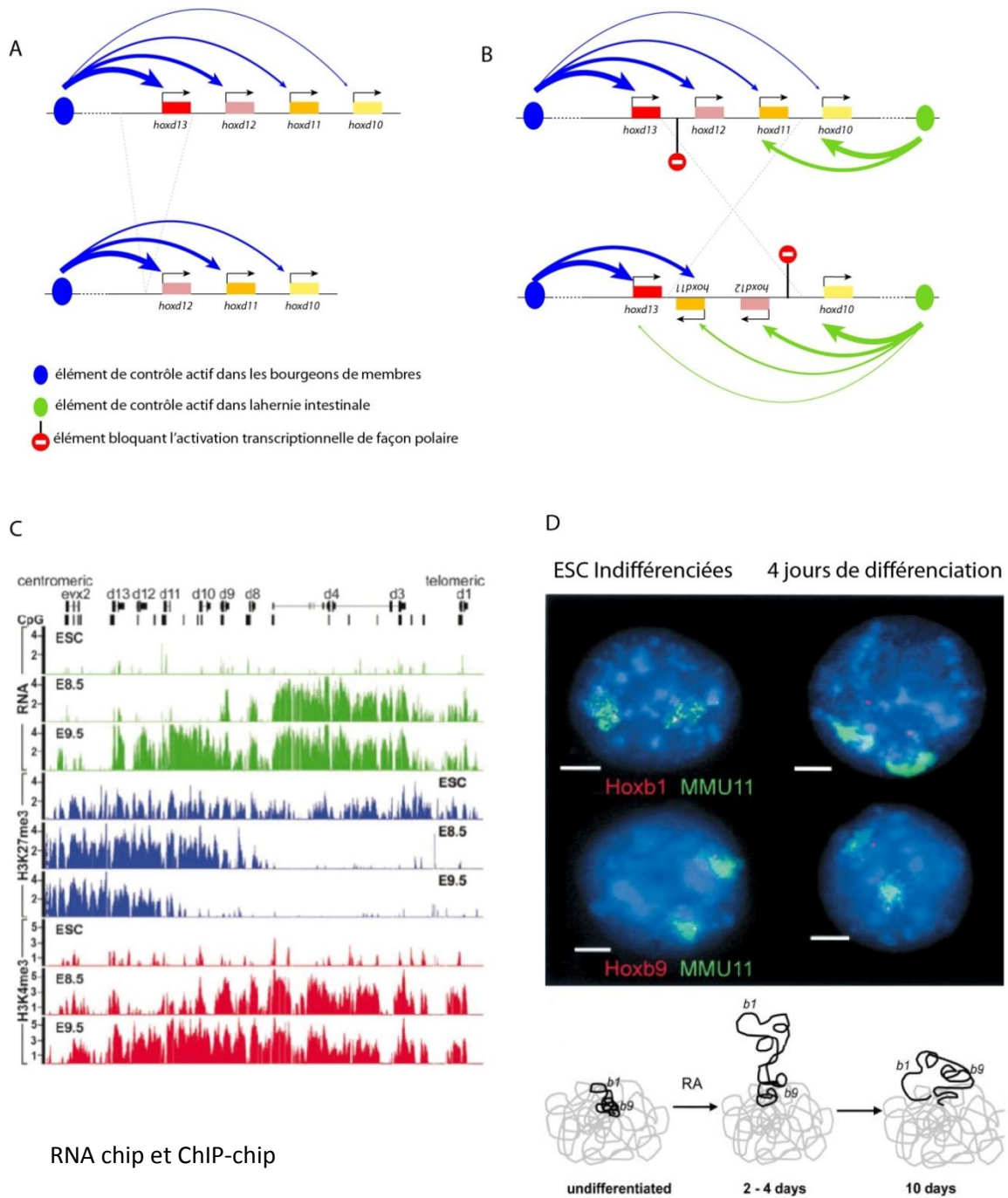
génomiques, ou inversement à généraliser les effets régulateurs associés à un nombre restreint de caractéristiques structurales (lors de l'interprétation des cartes épigénomiques par exemple). Nous allons maintenant discuter des études qui se sont attachées à caractériser les effets régionaux, qui émergent de la convolution de multiples afférences régulatrices en un point du génome, et qui reflètent le contexte dans lequel la dynamique transcriptionnelle s'inscrit *in vivo*.

#### **4.3. Paysage transcriptionnel et loci co-régulés : l'exemple des *Hox***

L'exemple des clusters *Hox* est d'un intérêt particulier car il illustre l'interconnexion de multiples mécanismes impliqués dans la régulation de domaines génomiques au cours du développement (pour revue détaillée, voir Spitz et Duboule, 2008). De nombreuses expériences de génétique murine ont permis d'y démontrer l'existence d'influences régulatrices hautement dynamiques, à la fois au cours du développement et en fonction du type cellulaire, et pouvant s'exercer sur un ou plusieurs membres des différents complexes *Hox*. Nous allons maintenant discuter de quelques observations qui illustrent l'idée que l'activité transcriptionnelle à une région donnée est le résultat de l'intégration de multiples afférences régulatrices, pouvant s'exerçant à la fois au niveau local et régional.

Tout d'abord, l'étude du cluster *HoxB* a montré qu'un même élément génomique peut exercer un effet régulateur partagé par plusieurs promoteurs dans le même tissu – assurant dans ce cas la co-expression de *Hoxb3* et *Hoxb4* dans une partie précise du rhombocéphale murin en développement (Gould et al., 1997). Ces observations suggèrent d'entrée que d'autres mécanismes que l'activation par contact via boucles chromosomiques peuvent sous-tendre la régulation transcriptionnelle à distance.

L'utilisation de transgènes comprenant toutes les séquences codantes du complexe *HoxD* a démontré que les différents promoteurs de ce locus sont soumis à l'influence de plusieurs éléments de contrôle distaux, distincts en fonction du type cellulaire à un même stade de développement (Spitz et al., 2001). En effet, ces larges transgènes récapitulent l'expression tissu-spécifique des différents loci *HoxD* le long de l'axe antéro-postérieur dans le tronc, mais ne sont pas transcrits en dehors de ces tissus – et restent en particulier silencieux dans les bourgeons de membres où *Hoxd11*, *Hoxd12* et *Hoxd13* sont normalement particulièrement actifs.



**Figure 9. Influences régulatrices et organisation structurale des loci *Hox* murins (A)** Un même élément régulateur peut contrôler plusieurs promoteurs différents. Dans le cas présenté, la spécificité d'action de l'élément de contrôle des bourgeons de membres est sous-tendue par la distance génomique qui le sépare de ses promoteurs cibles, comme l'illustre l'activation ectopique des membres *HoxD* proximaux en cas de délétion de la partie distale du locus. **(B)** L'effet bloqueur d'influence transcriptionnel peut être exercé de façon polaire, comme l'illustre l'analyse d'une inversion partielle du locus *HoxD* **(C)** La différenciation cellulaire entre le stade cellule souche (ESC), E8.5 et E9.5 s'accompagne de l'activation des membres proximaux dans le bourgeon caudal, visible sur les pistes de RNA-chip (vert). Cette activation transcriptionnelle est concomitante avec la régression polaire de la décoration des parties proximales du locus par H3K27me3 (bleu), témoignant de la présence des protéines du groupe Polycomb, et du gain de H3K4me3 (rouge), témoignant de l'activité locale des protéines MLL/Trithorax. **(D)** La différenciation *in vitro* des ESC entraîne la décompaction du locus *Hoxb9* et la projection des membres proximaux, premiers à être activés transcriptionnellement, en dehors du territoire du chromosome 11 (MMU11). (A) Inspiré de Kmita et al., 2002 et Zeller et Deschamps, 2002 (B)

Ces mêmes observations ont également suggéré l'existence d'éléments de contrôle tissu-spécifiques agissant sur les promoteurs du complexe *HoxD* mais situés en dehors du cluster homéotique. En effet, une construction *Hoxd11/LacZ* insérée au point de délétion du complexe est transcrite dans les bourgeons de membres – mais pas dans le tronc (Spitz et al., 2001).

Les complexes *Hox* sont en outre célèbres de par le fameux phénomène de colinéarité qui s'y exerce. Ainsi, un promoteur du complexe *HoxD* sera toujours transcrit plus efficacement dans les bourgeons de membres s'il est (naturellement ou artificiellement) placé en position proximale (nous emploierons ici « proximal » et « distal » en référence au numéro du locus *Hox* en question, et non à l'orientation centro-télomérique ou à l'orientation de la transcription). De plus, des délétions internes raccourcissant le complexe entraînent l'activation transcriptionnelle ectopique des parties distales du cluster, dans les tissus où les loci évincés sont normalement transcrits (Kmita et al., 2002). Ces observations suggèrent que la régulation transcriptionnelle en jeu s'exerce de façon polaire et non spécifique des promoteurs impliqués (**figure 9A**).

Des effets barrières sont également en jeu de façon tissu-spécifique. En effet, dans la hernie intestinale *Hoxd12* et *Hoxd13* ne sont pas transcrits alors que les membres plus proximaux du complexe sont stimulés par un élément activateur en 5' du locus. De manière surprenante, l'inversion de *Hoxd11-Hoxd12-Hoxd13* entraîne la transcription ectopique dans la hernie de toute la partie distale, suggérant que cet effet barrière agit lui aussi de façon polaire (Kmita et al., 2000). Les mécanismes sous-tendant ces phénomènes ne sont pas encore élucidés (**figure 9B**).

Les effets de régulation régionale visibles aux clusters *Hox* reflètent-elles un contrôle plus global du paysage transcriptionnel ? L'analyse du patron d'activité des promoteurs bordant *HoxD* (*Evx2* 8kb en 5' et *Lnp* 90kb en 3') a révélé que ces loci sont transcriptionnellement actifs dans de nombreux tissus, à l'exception notable du tronc (Spitz et al., 2003). Ces observations ont mené à l'idée qu'une région de contrôle globale exerce son influence sur une large portion chromosomique, qui se manifeste en terme de modulation de l'activité transcriptionnelle des promoteurs de cette région génomique en fonction des autres influences exercées plus localement de façon tissu- et stade-spécifique.

Les mécanismes moléculaires sous-tendant la colinéarité temporelle d'activation transcriptionnelle des *Hox* restent encore largement incompris, mais l'hypothèse selon laquelle une transition progressive des états chromatinien le long du locus au cours du temps (Dolle et al., 1989) a trouvé support à plusieurs niveaux, à la fois en termes de la composition et de l'organisation spatiale de la chromatine. En effet, il est connu que l'invalidation des complexes polycomb PRC2 ou PRC1 induit l'expression prématurée des membres les plus distaux des complexes *Hox*. Le fort

enrichissement en H3K27me3 et H2AK119Ub à ces loci dans les cellules souches embryonnaires – avant leur activation transcriptionnelle – y témoigne en outre de la présence de PRC2 et PRC1 respectivement. L'étude du développement du bourgeon caudal a dévoilé une corrélation frappante entre l'activation transcriptionnelle et les changements de la composition de la chromatine des membres les plus distaux du locus *HoxD* (Soshnikova et Duboule, 2009), révélant la progression polaire de l'enrichissement en H3K4me3 et H3Ac à partir de l'extrémité 5' d'une part et la rétraction correspondante de H3K27me3 dans les partie 3' (**figure 9C**). De manière importante, une inversion du locus intervenant entre *Hoxd10* et *Hoxd11* a en outre permis de démontrer que la colinéarité génomique du cluster est nécessaire pour l'établissement précis des patrons chromatiniens initiaux aux deux extrémités du point d'inversion.

En terme d'organisation topologique, l'activation transcriptionnelle de *HoxD* s'accompagne également d'une décompaction globale du locus mesurable par DNA FISH (Morey et al., 2007 – tout comme celle de *HoxB*; Chambeyron et Bickmore, 2004, Chambeyron et al., 2005 – et de la réorganisation du positionnement du locus par rapport à son territoire chromosomique, également visible par DNA FISH (**figure 9D**). L'inactivation de PRC2 ou PRC1 entraîne la décompaction prématurée de *HoxD* et *HoxB* dans les cellules souches embryonnaires, suggérant que ces complexes pourraient jouer un rôle direct dans le maintien de l'architecture chromosomique locale (Eskeland et al., 2010). La décondensation et la réorganisation nucléaire du territoire chromosomique ne sont pas deux facettes du même phénomène, puisqu'à l'échelle allélique l'un peut être visible sans l'autre – suggérant que bien que concomitants à l'échelle de la population ces évènements ne sont pas mécaniquement interdépendants (Morey et al., 2007). Les mécanismes qui entraînent la protrusion de *HoxD* et *HoxB* en dehors du territoire de leur chromosome au moment leur activation transcriptionnelle ne sont pas compris. Plusieurs observations ont cependant démontré que cette réorganisation n'est pas dépendante de l'activité transcriptionnelle. En effet, l'analyse d'un transgène *Hoxb1* inséré à l'extrémité distale de *HoxD* a montré que dans certains tissus ce transgène est capable d'être extrudé du territoire chromosomique et d'induire la décompaction partielle de *HoxD* bien qu'il n'y soit pas transcrit (Morey et al., 2008). De plus, l'éloignement de *HoxB* et *HoxD* de leurs territoires chromosomiques correspondants pendant la différenciation des cellules souches embryonnaires ne s'accompagne pas de changement majeur dans l'activité transcriptionnelle des loci voisins, qui participent pourtant également à cette réorganisation de l'architecture chromosomique (**figure 9D**). Les facteurs responsables de la réorganisation chromosomique des loci *Hox* ne sont pas connus. En particulier, il reste à élucider à quel point cette propriété est impliquée dans leur régulation transcriptionnelle ou si elle est simplement symptomatique de la dynamique structurale de leur chromatine pendant le développement.

Une dernière caractéristique de la régulation transcriptionnelle des *Hox* que nous voulons mentionner est le phénomène de contrôle en *trans* entre les différents clusters. En effet, l'activité régulatrice de certains éléments génomiques des loci *Hox* est modulée par la liaison de certaines protéines HOX elles-mêmes (Maconochie et al., 1997). L'association du phénomène de partage de régions de contrôle et de boucles auto- ou rétro-régulatrices pourrait ainsi permettre la maintenance robuste de patrons transcriptionnels. Dans certains cas, les produits d'activités des loci *Hox* responsables de leur inter-régulation en *trans* n'est pas de nature protéique, mais ribonucléique. Par exemple, chez l'humain, un ARN non codant (HOTAIR) produit par le locus *HOXC* a été impliqué dans la régulation de la structure de la chromatine de *HOXD* (Rinn et al., 2007), bien que ce rôle ne semble pas être conservé chez la souris (Schorderet et Duboule, 2011). De nombreux ARN non codants sont produits par les différents clusters *Hox* murins, et leur implication dans la régulation en *cis* ou en *trans* de ces loci reste une question totalement en suspens.

En conclusion les clusters *Hox* semblent particuliers de par la complexité de leur contrôle transcriptionnel. En outre, ils sont spécialement extraordinaires de par la finesse des études génétiques qui y ont été réalisées, et il est probable que la difficulté de modéliser simplement les influences *cis*-régulatrices s'y exerçant provient de la quantité d'information qui a été produite les concernant. À quel point la situation aux *Hox* reflète-t-elle ce qui se passe aux autres régions génomiques ? Difficile de répondre aujourd'hui, mais il est clair que les loci pour lesquels l'ingénierie génétique a été intensive chez les mammifères ont presque systématiquement révélé des niveaux de complexité inattendus en terme de régulation transcriptionnelle.

#### **4.4. Changements transcriptionnels régionaux induits par le repositionnement nucléaire**

Nous avons évoqué plusieurs fois la périphérie du noyau dans les paragraphes précédents, notamment pour la faible activité transcriptionnelle des segments chromosomiques qui ont tendance à y résider. La région corticale du noyau est très hétérogène, et accueille des loci aux patrons transcriptionnels variés, en fonctions des zones membranaires et des propriétés des loci considérés (pour revue, voir Akhtar et Gasser, 2007). Il a longtemps été difficile de directement tester si la localisation en périphérie du noyau a un impact direct sur l'activité transcriptionnelle chez les mammifères, et en particulier à quel point le positionnement nucléaire d'un locus peut affecter le comportement de son voisinage chromosomique. Le développement de stratégies moléculaires permettant d'amarrer une région génomique à la périphérie du noyau dans des cellules humaines, via l'intégration de séquences *LacO* et l'expression de protéines membranaires ou laminaires fusionnées au LacI (pour revue voir Ruault et al., 2008), a permis d'apporter des éléments de réponse



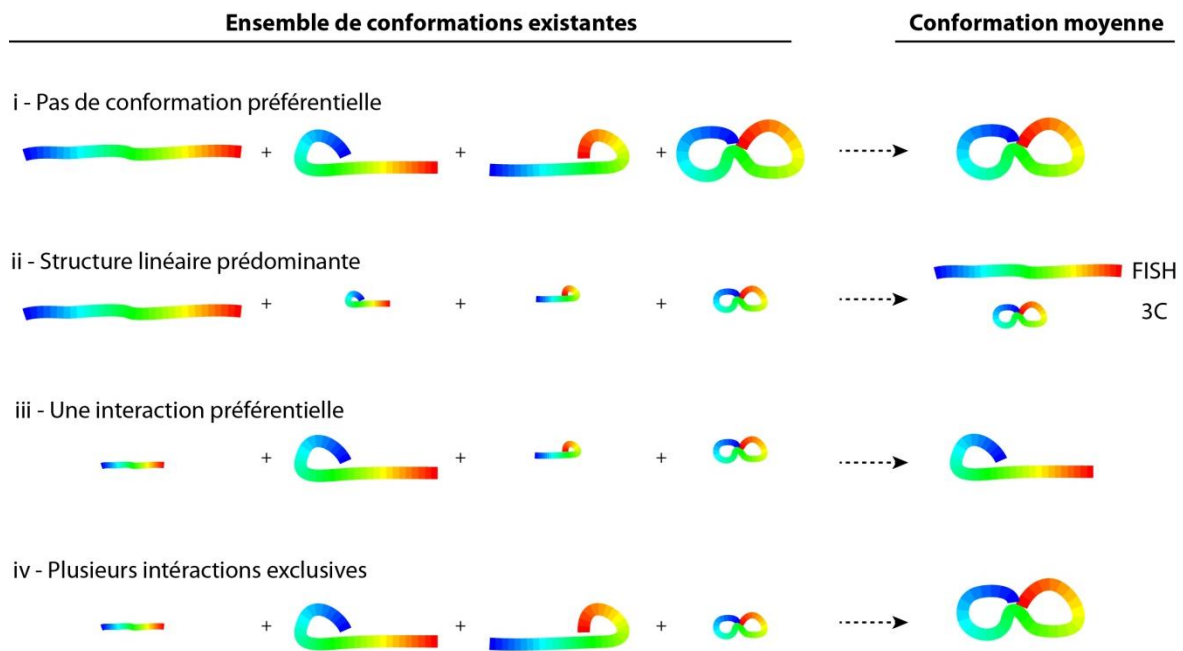
à ces questions. La baisse modeste du taux de transcrits provenant de rapporteurs juxtaposés aux sites *LacO* en situation d'ancrage à la membrane (ou à la lamine nucléaire) laisse suggérer que le positionnement en périphérie du noyau a un effet négatif sur l'activité transcriptionnelle, sans pour autant lui être complètement réfractaire (Kumaran, 2008 ; Reddy et al., 2008 ; Finlan et al., 2008). De manière surprenante, le locus ancré à la membrane (ou à la lamine) ne se désolidarise pas du territoire du chromosome auquel il est physiquement lié : le chromosome transgénique entier tend à être plus fréquemment en périphérie (Finlan et al., 2008). Des changements transcriptomiques sont en outre détectés (diminution pour la majorité), principalement dans les 10Mb encadrant l'insertion *LacO*, révélant un effet régional qui ne concerne pas tout le chromosome repositionné. Plus précisément, il apparaît que le taux de transcrits venant de certains loci seulement est affecté, suggérant que des propriétés locales modulent probablement la sensibilité de certaines régions chromosomiques à l'effet répresseur du repositionnement en périphérie du noyau. Si la nature de ces propriétés n'est pas connue, il a été montré que des inhibiteurs d'histone déacétylases (HDAC) contrecarrent la baisse du taux de transcrits induite par l'ancrage à la membrane nucléaire – où ces enzymes sont normalement particulièrement enrichies. Il a ainsi été proposé que le repositionnement en périphérie du noyau pourrait être un mécanisme de modulation de l'activité transcriptionnelle, probablement sous-tendue par la forte concentration locale en HDAC, et dont les effets peuvent d'une part s'exercer à longue distance des éléments qui sous-tendent l'association au cortex nucléaire, et d'autre part se révéler qualitativement et quantitativement différents en fonction des propriétés locales des différents loci concernés.

Ces observations intéressantes suggèrent que la dynamique nucléaire d'un locus peut avoir des conséquences à longue distance sur son voisinage chromosomique. L'ancrage artificiel à la membrane nucléaire impose probablement des contraintes de diffusion inhabituelles, notamment de par la stabilité atypique de la liaison LacI-LacO (Dubarry et al., 2011). Cependant, de nombreuses observations suggèrent que des changements du positionnement nucléaire et de l'architecture chromosomique participent pleinement à la modulation de l'activité transcriptionnelle (pour revue voir Meister et al., 2011). L'étude de translocations balancées dans des cellules humaines a par exemple révélé que l'altération de l'organisation chromosomique s'accompagne de changements transcriptionnels globaux ainsi que de la modification du positionnement nucléaire des chromosomes concernés par la translocation, bien que la ploïdie soit conservée (Harewood et al., 2010). Il apparaît donc important d'appréhender l'architecture chromosomique dans le contexte de la dynamique de l'organisation nucléaire afin de comprendre dans quel contexte structural s'inscrivent les nombreux mécanismes de régulation de l'activité transcriptionnelle décrits jusqu'à présent. Ce sujet est l'objet du chapitre suivant.

## 5. Architecture des chromosomes et dynamique de l'organisation nucléaire

La description des patrons d'organisation de la chromatine dans son contexte nucléaire pose des défis techniques importants. Si la microscopie photonique peut sembler l'approche la plus directe pour appréhender la conformation des chromosomes, le nombre restreint de fluorophores disponibles pour le marquage des sites génomiques d'intérêts d'une part, et les résolutions spatiale et temporelle limitées d'autre part, rendent son utilisation difficile pour la construction de modèles détaillés de l'architecture des chromosomes. D'un autre côté, le développement récent des techniques comme la capture de conformation chromosomique (3C), l'immuno-précipitation de chromatine (ChIP), l'identification de cibles d'adénine méthyl-transférases recombinantes (DamID) et leurs possibles combinaisons et variantes permettent de produire des cartes d'interactions chromosomiques et protéiques statistiques à l'échelle du génome (pour revue voir van Steensel et Dekker, 2010). Ces techniques génomiques reflètent un comportement statistique dans le sens où les tendances qu'elles mettent en évidence émergent de la moyenne de millions de cellules, et ne donne donc pas directement accès à la variabilité intercellulaire ni ne permet pas d'appréhender la co-occurrence des différents évènements étudiés à l'échelle de la cellule (ou du chromosome) unique.

En effet, comme il n'est pas possible de suivre la dynamique de l'ensemble des repères génomiques au court du temps, il est capital de garder à l'esprit que l'interprétation de données – aussi bien microscopiques que génomiques – obtenues sur plusieurs chromosomes différents est automatiquement associée à plusieurs postulats. D'un point de vue d'échantillonnage spatial d'une part, toute représentation moyenne qui émerge de telles observations sera donc statistique et ne reflétera pas la diversité des réalisations topologiques uniques. Certaines de ces représentations feront parfois même coexister des conformations non compatibles entre elles (voir **figure 10**). Par exemple, si un locus A est parfois trouvé physiquement associé à un locus B, et que le locus B est parfois retrouvé avec le locus C, le comportement moyen suggérera que les loci A,B et C peuvent être co-localisés – même si on sait par ailleurs que le locus B ne peut avoir qu'un seul partenaire à la fois. La microscopie en fluorescence permettant des mesures multi-variées, elle peut apporter des solutions à ce problème dans certains cas. Ensuite, d'un point de vue d'échantillonnage temporel, on suppose d'entrée de jeu que le système est ergodique, c'est-à-dire que l'analyse de nombreuses conformations simultanées permet d'approcher la variabilité de l'une d'entre elle dans le temps. Ce principe est probablement satisfait par la plupart des approches génomiques, qui analysent des millions de cellules à la fois, mais il n'est pas toujours possible d'en présager, surtout en microscopie où l'analyse se base sur un nombre bien plus limité de cellules. L'imagerie de cellules vivantes peut permettre d'apporter de précieuses informations quant à la dynamique de l'organisation nucléaire, mais est difficile en pratique, en particulier chez les mammifères.



**Figure 10. Limites des représentations moyennes de la structure des chromosomes.** Il existe généralement plusieurs configurations chromosomiques possibles pour un même segment chromatinien. Le calcul (par exemple en moyennant des distances mesurées par FISH) ou l'expérience (par exemple lors de la 3C où des millions de cellules sont analysées simultanément), peut mener à la perception d'une moyenne de ces configurations qui ne reflète pas la diversité des conformations réelles. Dans certains cas (iv) la moyenne de plusieurs conformations exclusives peut même mener à une représentation moyenne qui ne reflète aucune des réalisations topologiques réelles. Dès qu'une conformation moyenne est considérée, il est donc utile de se demander jusqu'à quel point celle-ci reflète le comportement d'une fibre de chromatine unique.

En somme, lorsque les échelles étudiées le permettent, il est donc important d'essayer d'intégrer les données microscopiques et génomiques. Ces approches peuvent être aidées par la modélisation sur des principes de physique statistique, qui permet d'analyser un grand nombre de réalisations topologiques compatibles avec les mesures expérimentales (pour revue voir Baù et Marti-Renom, 2010). En termes pratiques il est impossible d'inspecter manuellement chacune de ces réalisations, mais il est tout à fait possible de leur appliquer les analyses classiques de formation de cluster. Leur regroupement permet notamment de quantifier la variabilité et le degré de ressemblance entre les différentes conformations, et d'apprécier s'il existe des types de configurations dominantes. Les hypothèses nécessaires pour bâtir de tels modèles sont parfois

nombreuses et arbitraires, mais cette démarche peut néanmoins être particulièrement utile pour guider les expériences qui permettront de tester leur exactitude.

Après ces considérations, nous allons maintenant détailler quelques propriétés qui nous semblent représentatives de l'organisation topologique de la chromatine à différentes échelles. Comme précédemment, nous concentrerons principalement nos commentaires sur des travaux effectués chez les mammifères, en mentionnant quelques travaux effectués chez la drosophile qui nous apparaissent particulièrement relevant pour cet exposé.

### **5.3. De la dizaine à la centaine de kb**

#### **5.3.1. Généralités**

L'organisation *in vivo* de la fibre chromatinienne chez les mammifères reste très peu étudiée en deçà de la dizaine de kilobases (ce qui pourrait être assemblé en une cinquantaine de nucléosomes), à cause des limites de résolution des techniques mentionnées précédemment. Quant aux connaissances de l'architecture des chromosomes entre la dizaine et la centaine de kilobases, elles reposent presque exclusivement sur les études biochimiques – les sondes utilisées pour l'hybridation *in situ* en fluorescence étant trop grandes pour être utilisées dans ce contexte (elles-mêmes mesurant entre 10 et 100kb). Plusieurs études employant la méthode capture de conformation chromosomique (3C) de façon systématique, ou sa variante multiplex dite de copie-conforme de 3C (5C - voir **figure 6**), ont permis de décrire le repliement statistique local d'un nombre limité de loci dans plusieurs types cellulaires humains: *FMR1*, *HOX*, A-GLOBIN, et un désert riche en A/T du chromosome 16 (Gheldof et al., 2006 ; Fraser et al., 2009 ; Ferraiuolo et al., 2010 ; Baù et al., 2010 ; Dostie et al., 2006 ). Le point commun de ces études est d'avoir rapporté que, quel que soit le locus étudié, les régions contigües le long de la fibre chromatinienne ont tendance à interagir physiquement beaucoup plus fréquemment que des régions éloignées. En revanche, la vitesse avec laquelle la probabilité d'interaction décroît avec la distance génomique est variable en fonction des loci étudiés, et parfois même au sein du même voisinage chromosomique. L'absence d'interaction spécifique à longue distance dans le désert riche en A/T étudié suggère que la région n'adopte pas de conformation préférentielle (Dostie et al., 2006 ), pouvant être expliqué par une configuration linéaire décompactée ou au contraire un arrangement confiné mais variable entre chaque chromosome. En contraste, lorsque la topologie de la chromatine autour de promoteurs ou de leurs éléments de contrôle est analysée, de nombreuses interactions ponctuelles et avec des régions parfois très distantes (>100kb) sont généralement observées – confirmant les observations rapportés dans les expériences de 3C dédiées (Tolhuis et al., 2002 et bien d'autres).

### 5.3.2. Dynamique développementale des interactions chromosomiques en *cis* ponctuelles

Les associations focales engageant des éléments génomiques discrets corrélient généralement avec l'état d'activité transcriptionnelle de ces loci, mais peuvent être dus aux facteurs *trans* qui lient ces éléments et stimulent la transcription sans être directement une conséquence de l'acte transcriptionnel. Par exemple, les promoteurs *Il4*, *Il5* et *Il13*, dispersés sur 100kb chez la souris interagissent physiquement quelque-soit le type cellulaire regardé (Spilianakis et Flavell, 2004); dans les lymphocytes T, un élément de contrôle distal participe à ce réseau d'interaction, aussi bien dans les cellules T naïves que dans le lignage Th1, où ces loci ne sont pas transcrits. Ce n'est que dans le lignage Th2 où la fixation de facteurs tissu-spécifique tels GATA3 est associée à leur activation transcriptionnelle. Qui plus est, l'expression ectopique de GATA3 dans les fibroblastes stimulés entraîne l'apparition de la configuration chromosomique typique des lymphocytes, sans pour autant entraîner la production des interleukines 4, 5 ou 13. L'organisation topologique de la chromatine en deçà de 100kb peut donc être contrôlée par des éléments *cis* qui réagissent à l'influence de facteurs *trans* ubiquitaires ou tissu-spécifiques et qui peuvent en outre, de façon parallèle et concurrente, stimuler la transcription. Les mécanismes permettant le contrôle coordonné mais parallèle de l'architecture locale de la chromatine et de l'activité transcriptionnel est illustré par l'association physique des complexes cohésine et Mediator, récemment proposé pour sous-tendre localement la communication structurale et fonctionnelle de dizaines milliers d'éléments à travers le génome (Kagey et al., 2010).

### 5.3.3. Rôle des facteurs *trans* dans l'architecture locale de la chromatine

Il ne faut cependant pas imaginer que l'architecture de la chromatine et la transcription sont coordonnées par un mécanisme unique. Certains facteurs *trans* sont en effet capables d'agir en amont à la fois de l'activation transcriptionnelle et du contrôle de la conformation chromosomique. Par exemple, l'étude simultanée des événements de liaison du récepteur  $\alpha$  de l'œstrogène (ER $\alpha$ ) et des interactions chromosomiques faisant intervenir cette protéine à l'échelle du génome humain entier, par analyse des interactions de la chromatine *via* séquençage d'étiquettes en extrémités appareillées (*Chromatin Interaction Analysis by Paired-End Tag sequencing*, ChIA-PET), a révélé que près de la moitié des sites de fixation de ce récepteur hormonal sont physiquement associés (Fullwood et al., 2009). Ces interactions concernent presque exclusivement des éléments génomiques distants de moins de 1Mb, la grande majorité étant en fait séparée de moins de 100kb. De manière intéressante, certaines de ces interactions n'apparaissent qu'après traitement hormonal, suggérant que l'ER $\alpha$  pourrait être responsable du remodelage local de la conformation

chromosomique – bien qu’il n’ait pas été clairement établi quelle est la proportion de ces interactions qui préexistent à la fixation du récepteur nucléaire à la chromatine. Il est important de noter que d’autres facteurs que l’ER $\alpha$  sont clairement nécessaires pour déterminer l’interactivité de ses sites de fixation, puisque la moitié de ses évènements de liaison à la chromatine ne fait pas intervenir de telles interactions physiques. En revanche, près dans près d’un tiers des cas plusieurs sites sont associés entre eux d’une façon qui pourrait s’expliquer par la formation d’une série de boucles concentriques ancrées entre elles sur un court segment. Le suivi de la cinétique d’activation transcriptionnelle suivant le traitement hormonal a révélé une corrélation surprenante dans la rapidité de réponse des loci localisés dans différentes régions de ces structures. En effet, les loci situés dans la partie coalescente de ces boucles putatives sont plus rapidement activés transcriptionnellement au cours du traitement hormonal que ceux situés dans les parties projetées, eux-mêmes tout de mêmes plus rapides que les loci qui ne sont pas engagés dans de telles structures topologiques. La cartographie exhaustive des interactions impliquant l’ER $\alpha$  a ainsi non seulement révélé que la topologie locale de la chromatine peut être contrôlée par la fixation d’un récepteur nucléaire, mais aussi que cette architecture est directement en relation avec l’efficacité d’activation transcriptionnelle des loci sous-jacents. A quel point de telles boucles concentriques reflètent le moyennage de multiples configurations possibles, et probablement dynamique au fur et à mesure de la progression de la réponse hormonale, reste une question totalement ouverte.

#### **5.4. De la centaine de kb au Mb**

En sus de telles interactions spécifiques entre éléments génomiques discrets existe un niveau basal de collision, détectable entre n’importe quelle paire de régions chromosomiques. Le perfectionnement récent des méthodes de quantification de 3C chez les mammifères a permis de mettre en évidence une périodicité de 90kb dans la fréquence d’établissement de contacts physiques le long de la fibre de chromatine, qui peut être expliquée par une configuration statistique hélicoïdale à l’échelle supra-nucléosomale (Court et al., 2011b). A quel point cette observation correspond à un arrangement rigide de la fibre chromatinienne dans le noyau ou reflète une tendance commune d’une forte proportion des configurations possibles reste à être exploré (**figure 10**). En effet, les oscillations étant rapidement amorties (l’effet devenant difficilement mesurable après trois périodes), il est probable qu’une variabilité importante existe entre les différentes configurations possibles à un même locus, laissant la question de l’existence *in vivo* d’une périodicité structurale de la fibre chromatinienne à l’échelle supra-nucléosomale encore en suspens. Ces observations ont mené à la proposition intéressante que cette périodicité structurale pourrait

participer à la modulation de l'exercice d'effets régulateurs entre loci distants. De manière surprenante, cette organisation cyclique semble réservée aux régions riches en G/C (et donc généralement riches en promoteurs), mais n'apparaît pas dépendante de l'activité transcriptionnelle des loci sous-jacents. De plus, à distance génomique égale, la fréquence de rencontre entre deux sites donnés varie suivant les régions étudiées, suggérant que les régions riches en A/T, dont le potentiel transcriptionnel est généralement faible, sont localement moins interactives que les régions riches en G/C. L'augmentation du nombre de loci étudiés permettra de tester l'idée que les isochores génomiques corrént avec des modulations régionales de l'interactivité chromosomique locale. L'organisation précise de la fibre chromatinienne entre la centaine de kb et le mégabase reste peu documentée. Comme il sera détaillé dans la section suivante, mon travail s'est attaché à avancer dans la description de ce niveau d'organisation en développant des méthodes basées sur le 3C à grande échelle et l'imagerie à haute résolution (voir partie « Résultats »).

### 5.5. Du Mb à la dizaine de Mb

A des échelles supérieures au mégabase, les approches de DNA FISH deviennent réalisables en routine. Les premières études corrént la distance génomique, en kb, à la distance physique mesurée entre les signaux de FISH, en nm, a suggéré que la chromatine adopte globalement un comportement de marche aléatoire (van den Engh et al., 1992). Le coefficient de corrélation entre la distance en kb et la distance physique en interphase varie cependant entre les différentes régions du génome suggérant qu'en interphase certains segments chromosomiques sont plus compactés que d'autres dans un même type cellulaire (Yokota et al., 1997). L'étude de plusieurs régions chromosomiques, comme les clusters *Hox* ou des loci contrôlant la production d'immunoglobulines, ont montré que la compaction d'un même segment génomique peut également varier en fonction du type cellulaire (Chambeyron et Bickmore, 2004, Roldan et al., 2005). Cette évolution est souvent liée à l'état transcriptionnel du locus, mais n'en est ni une conséquence ni un déterminant, et sont parfois liées à d'autres activités métaboliques du génome que la transcription.

Par exemple, le locus *IgH* est particulièrement compacté dans les cellules pro-B murines, avant que les parties proximales et les parties distales ne s'engagent dans les évènements de recombinaison qui participent à la diversité antigénique (Kosak et al., 2002). La dissection fine de ce locus complexe a en outre révélé que, de manière surprenante, la distance physique moyenne séparant la partie  $D_H$  des différentes répétitions  $V_H$  dans les cellules pro-B est équivalente quelque-soit le segment  $V_H$  considéré, bien que ceux-ci se répartissent sur près de deux Mb (Jhunjhunwala et al., 2008). L'analyse par imagerie a permis de suggérer que cette situation est due à l'adoption d'un

ensemble variable de configurations, et que cette variabilité est en outre plus importante dans les pro-B que dans leurs progéniteurs. Cette observation est particulièrement intéressante au vu du fait que chaque segment  $V_H$  a une probabilité quasi-équivalente de participer à la recombinaison antigénique avec les parties proximales du locus *IgH* lors de la recombinaison antigénique. Cette topologie est donc incompatible avec un comportement de simple marche aléatoire et suggère que l'organisation architecturale de la fibre chromatinienne peut être contrainte de façon régulée à l'échelle de quelques mégabases. La modélisation des données expérimentale par divers arrangements topologiques simulés d'une fibre chromatinienne de 30-nm a mené à la proposition de plusieurs types de configurations possibles, notamment celles dites de marche aléatoire et boucles géantes d'une part – où de larges segments de plusieurs Mb forment des boucles attachées à une matrice nucléaire (Yokota et al., 1995 ; Sachs et al., 1995) – et celles dites de sous-compartiment multilobés d'autre part – impliquant la formation de rosettes de petites boucles reliées par des joints chromatiniens de taille variable (Münkel et al., 1999). Le cas du locus *Igh* dans les pro-B est compatible avec une organisation en rosettes reliées par des segments de jointure (Jhunjhunwala et al., 2008). Que ce cas reflète le comportement général de la fibre chromatinienne ou les spécificités architecturales de ce locus dans le lignage lymphocytaire est une question ouverte, dont la réponse nécessitera l'examen précis d'un nombre plus important de régions chromosomiques. En particulier, ces modélisations postulent l'arrangement topologique d'une fibre de 30-nm. Il est important de garder à l'esprit que l'existence d'une telle structure n'a été reportée que dans des tampons *in vitro*, et n'a pas été observée *in vivo* de façon reproductible (pour revue voir Fussner et al., 2011). Comme nous le discuterons dans nos résultats, il sera intéressant de dériver des modèles se passant d'un arrangement en fibre de 30-nm.

## **5.6. De la dizaine de Mb au chromosome**

### **5.6.1. Variabilité régionale de la condensation de la chromatine**

Ce niveau d'organisation est mieux compris, et peut être décrit à la fois par des techniques d'imagerie classique et de variantes de 3C. Comme mentionné précédemment, les chromosomes ne sont pas des unités structurales homogènes, et la compaction de la chromatine en interphase est hautement variable en fonction des régions étudiées. Les domaines décondensés correspondent aux régions particulièrement sensibles à la MNase, une DNase clivant entre les nucléosomes dans des extraits chromatiniens, suggérant que l'accessibilité locale de la chromatine est liée à des propriétés architecturales de la fibre à des échelles d'organisation supérieure (Gilbert et al., 2004).



Les loci particulièrement sensibles à la MNase sont en outre fréquemment retrouvés en dehors du cœur du territoire du chromosome qui les porte.

### **5.6.2. L'état transcriptionnel et le contenu C/G régional corrélient avec l'établissement d'interactions chromosomiques préférentielles**

Le développement des techniques de Capture Circulaire de Conformation Chromosomique (4C), permettant de quantifier la fréquence d'interaction d'un locus d'intérêt avec le reste du génome, a permis de grandement raffiner les connaissances de l'organisation statistique interne des chromosomes (Zhao et al., 2006 ; Simonis et al., 2006 ; Würtele et Chartrand, 2006 ; Lomvardas et al., 2006 ; voir **figure 6**). Ainsi, si la dizaine de mégabase encadrant un locus donné contient la majorité de ses partenaires physiques préférentiels, de nombreuses interactions reproductibles sont généralement détectables tout le long du chromosome auquel il est physiquement lié (Simonis et al., 2006 ; Würtele et Chartrand, 2006). Les nombreux loci étudiés par cette technique ont permis de dégager la tendance générale des régions transcrites à contacter préférentiellement d'autres régions transcrites, et inversement. Toutes les régions transcrites ne sont cependant pas contactées avec égale fréquence, suggérant l'existence d'affinités électives pour certains partenaires plutôt que d'autres. En particulier, les régions transcriptionnellement actives dans les domaines G/C riches du génome, où le potentiel de transcription est particulièrement dense, ont plutôt tendance à contacter d'autres régions riches en G/C que des régions transcriptionnellement actives mais riches en A/T. L'étude de loci dont la transcription est tissu-spécifique a en outre montré que cette préférence d'association en fonction du contenu en G/C est valable quel que soit l'état d'activité transcriptionnelle : les partenaires d'un locus silencieux dans un environnement riche en G/C seront d'autres loci silencieux dans un environnement riche en G/C. Les expériences de capture de conformation chromosomique globale (Hi-C) suggèrent que cette ségrégation spatiale des domaines génomique est un principe fondamental de l'organisation tridimensionnelle de la chromatine dans les cellules mammifères (Lieberman-Aiden et al., 2009).

Quels sont les facteurs qui amènent les loci transcriptionnellement actifs à se retrouver fréquemment dans l'espace tridimensionnel du noyau ? L'observation que la distribution de la machinerie de transcription dans le noyau est hétérogène (Jackson et al., 1993), et qu'elle se concentre en foci dont le nombre est bien moindre que celui des loci transcriptionnellement actifs (« usines de transcription »), a mené à la suggestion que l'ensemble des coordonnées nucléaires possibles pour un locus transcrit est en réalité limité, ce qui expliquerait *de facto* la forte probabilité de rencontre physique de deux loci transcrits (Osborne et al., 2004). Il n'est pas complètement clair

comment expliquer que des régions transcrites dans des domaines riches en G/C ne rencontre pas fréquemment des régions transcrites riches en A/T, mais il semblerait qu'un élément de réponse puisse tenir dans des différences dans les caractéristiques cinétiques de la transcription, vraisemblablement moins fréquente dans les régions riches en A/T. Comme il a été rapporté dans une étude à laquelle nous avons participé, parmi les partenaires préférentiels des régions A/T riches transcrites sont parfois retrouvées des régions transcriptionnellement inactives, reflétant probablement la variabilité intercellulaire de l'activité transcriptionnelle de ce type de loci (Splinter et al., 2011).

### 5.6.3. L'activité transcriptionnelle n'est pas déterminante dans la topologie chromosomique

L'acte transcriptionnel est-il le facteur déterminant l'association physique des régions transcrites du génome ? La réponse à cette question nécessite de garder à l'esprit que la transcription est un processus dynamique, qu'un même allèle n'est pas transcrit de façon continue au cours du temps et qu'à un instant donné le même locus peut être transcrits dans certaines cellules et pas dans d'autres (Yunger et al., 2010). Le traitement pendant plusieurs heures par des drogues inhibant l'initiation ou l'élongation de la transcription n'ayant pas d'effet mesurable sur les interactions chromosomiques mesurées par 4C, il a été suggéré que l'acte transcriptionnel n'est pas ce qui détermine l'association physique des loci transcrits entre eux (Palstra et al., 2008b). Il a ainsi été proposé que d'autres paramètres sont impliqués dans le maintien des interactions chromosomiques de ces loci, comme par exemple la présence de facteurs *trans* ou l'organisation structurale des nucléosomes, qui reste inchangées par les traitements d'inhibition de transcription. Plus généralement, ces études pointent vers l'idée que les caractéristiques de l'organisation architecturale des chromosomes pourraient être spécifiées indépendamment, ou même en amont, de l'activité métabolique régionale de ses différents loci. Cette idée est supportée par l'étude de loci transcriptionnellement activé en une heure après traitement hormonal au récepteur des glucocorticoïdes (GR), dont les partenaires chromosomiques distaux (>10Mb), contenant d'autres loci responsifs au GR, sont identiques avant ou après stimulation (Hakim et al., 2011). Mises en parallèle avec l'étude de ChIA-PET de ER $\alpha$  mentionné précédemment, ces observations suggèrent donc que la fixation de facteurs *trans* pourrait induire la transcription d'une part et moduler la topologie chromatinienne locale (<100kb) d'autre part, dans le contexte d'une architecture chromosomique globalement fixe. Une dynamique surprenante de l'architecture chromosomique a cependant été rapportée après traitement à l'ER $\alpha$  (Hu et al., 2008) mais la reproductibilité de ce phénomène reste à démontrer (Kocanova et al., 2010) avant d'en tirer des conclusions sur les

mécanismes généraux gouvernant la coordination de l'organisation de la chromatine et de la transcription (pour discussion voir Belmont, 2010).

#### 5.6.4. Dynamique développementale de l'architecture chromosomique

Nous avons mentionné précédemment la décompaction locale des clusters *HoxB* et *HoxD* qui accompagne leur activation transcriptionnelle pendant le développement précoce, et leur repositionnement cytotologique concomitant à l'extérieur du territoire du chromosome auquel ils sont lié (Chambeyron et Bickmore, 2004 ; Chambeyron et al., 2005 ; Morey et al., 2007). Si la réorganisation dramatique des *Hox* est probablement une caractéristique particulière de ces loci atypiques, l'architecture interne des chromosomes apparaît hautement dynamique au cours de la différenciation cellulaire. Le repositionnement de loci précis peut avoir lieu par rapport à des sous-régions du territoire : par exemple, au cours du développement lymphocytaire murin, les loci *Cd4* et *Cd8* restent dans le territoire du chromosome 6 auquel ils sont physiquement liés, mais des sondes spécifiques de leur voisinage chromosomique permettent de mettre en évidence qu'ils se désolidarise de leur environnement génomique direct respectif (Ktistaki et al., 2010). Une étude à laquelle j'ai participé a par ailleurs démontré que ces deux loci, distants de plus de 50 mégabases s'associent physiquement en *cis* à plusieurs phases précises de la maturation des lymphocytes T (Collins et al., 2011). La cinétique de ces évènement dépend de l'expression de plusieurs protéines spécifiques de chacun de ces stades, et qui sont par ailleurs capables de se lier à des régions non codantes de *Cd4* et *Cd8* et de contrôler leur transcription. De manière intéressante, l'expression concomitante de CD4 et CD8 dans les lymphocytes T précoces devient exclusive dans les lignages qui émergent de leur maturation. Ainsi, le locus *Cd4* est plus interne par rapport à son territoire sous-chromosomique local dans les cellules où il est silencieux (Ktistaki et al., 2010), mais en reste tout de même désolidarisé, ce qui correspond au maintien de son association avec *Cd8* (Collins et al., 2011). En revanche dans le lignage où *Cd4* est transcrit, *Cd8* est silencieux et toujours physiquement associé avec son voisinage génomique (Ktistaki et al., 2010), et les deux loci sont tenus à distance (Collins et al., 2011). Si les mécanismes précis permettant ces interactions chromosomiques ne sont pas connus, leur dynamisme est contrôlé positivement et négativement par plusieurs facteurs *trans*, liant *Cd4* et *Cd8*, et exprimés à différents stades de développement (Collins et al., 2011). Encore une fois, la transcription n'apparaît pas jouer un rôle directif puisque *Cd4* et *Cd8* restent associés lorsque seul *Cd8* est transcrit. Cet exemple illustre comment l'architecture chromosomique peut être dynamiquement remodelée au cours du développement, menant à l'appariement ou à la

dissociation physique de loci dont l'activation transcriptionnelle est coordonnée par des facteurs *trans* tissus-spécifiques.

### 5.7. Positionnement relatif des chromosomes et interactions physiques en *trans*

La technique de 4C, permettant d'interroger les partenaires physiques d'un locus d'intérêt dans tout le génome (**Figure 6**), a également révélé la capacité de plusieurs loci à s'associer avec certaines régions en *trans*. Cette propriété peut paraître surprenante car, si l'on comprend bien que l'existence des territoires chromosomiques favorise les rencontres entre deux loci physiquement liés, tel n'est pas le cas pour des régions n'appartenant pas au même chromosome. Certaines régions particulières du génome sont pourtant bien connues pour avoir la capacité à coalescer dans l'espace nucléaire, comme les segments péri-centromériques et les loci *rDNA* contrôlant la production des ARN ribosomiaux. De plus, l'organisation relative des territoires chromosomiques n'est cependant pas aléatoire, et il est connu que les petits chromosomes riches en A/T ont tendance à occuper des positions plus périphériques dans le noyau (Boyle et al., 2001), et que l'organisation radiale et relative des chromosomes est dynamique au cours du développement (Mayer et al., 2005). Un locus appartenant à un territoire chromosomique donné aura donc plus d'opportunité d'interagir en *trans* avec d'autres loci sur les voisins préférentiels de ce territoire. De manière intéressante, la capacité à établir des contacts en *trans* varie profondément entre les différents loci qui ont été étudiés par 4C (Roby et al., 2011). De façon étonnante, certains loci semblent même établir préférentiellement des interactions en *trans*, avec un jeu de partenaires limité, plutôt qu'avec autres régions du chromosome auquel il appartient (Lomvardas et al., 2006). Dans le seul cas où les profils 4C et les territoires chromosomiques ont été étudiés conjointement, la capacité à établir des interactions en *trans* a été corrélée avec une protrusion du locus concerné en dehors du territoire du chromosome auquel il est lié (Sandhu et al., 2009).

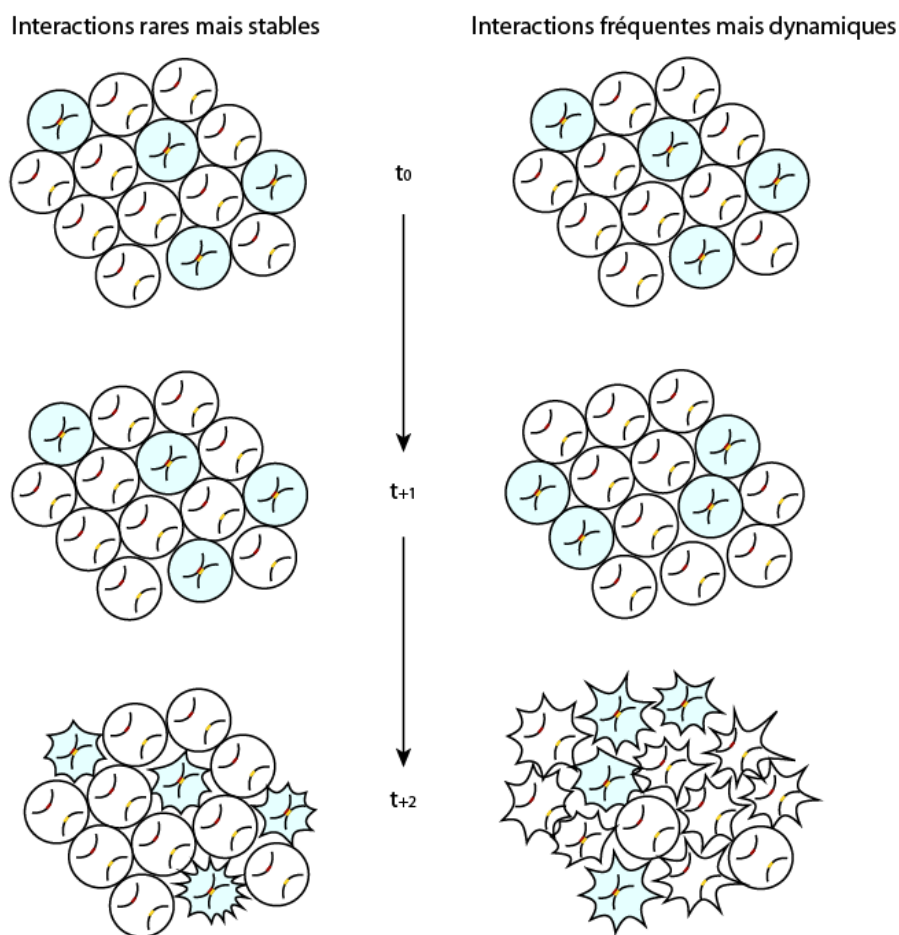
Ces interactions en *trans* sont-elles symptomatiques de propriétés communes des loci impliqués ? Pourraient-elles sous-tendre une co-régulation transcriptionnelle entre des éléments génomiques situés sur des chromosomes différents ? Plusieurs études ont rapporté que des loci dont la transcription est co-régulée peuvent s'associer physiquement dans le noyau à des fréquences importantes, souvent modulées au cours du développement (Spilianakis et al., 2005 ; Brown et al., 2006). Ces interactions peuvent en outre concerner des loci impliqués dans des voies biologiques similaires, comme l'a montré l'identification des partenaires physiques des loci  $\alpha$ -globin et  $\beta$ -globin activement transcrits, qui contactent préférentiellement des loci codant des protéines impliqués dans l'érythropoïèse (Schoenfelder et al., 2010). De manière intrigante, l'allèle maternel d'*Igf2/H19*

contacte de nombreux autres régions génomiques sous empreinte parentale, et une mutation de la région de contrôle de l'empreinte d'*Igf2/H19* altère leur réplication asynchrone (Zhao et al., 2006 ; Sandhu et al., 2009). Un autre exemple est la capacité du locus codant pour l'interferon- $\beta$  humain à engager, lors de son activation transcriptionnelle, des interactions chromosomiques avec plusieurs régions génomiques où le facteur Nf $\kappa$ B se lie constitutivement, qui a mené à la proposition que ces éléments joueraient un rôle important dans la *trans*-activation du locus par ce facteur (Apostolou et Thanos, 2008). Enfin, certains loci ont également été décrits comme pouvant établir des interactions homologues en *trans*, entre les deux allèles (Bacher et al., 2006 ; Xu et al., 2006 ; Brown et al., 2008 ; Hewitt et al., 2009). Ces associations sont également régulées de façon dynamique au cours du développement, et ont parfois pu être associées à des situations où une asymétrie apparaît entre les deux allèles, comme dans le cas de la recombinaison mono-allélique de *IgH* ou de l'activation transcriptionnelle mono-allélique de *Xist*, qui a été un de nos centre d'attention et sera discuté dans la section suivante.

## 5.8. Conclusions et perspectives pour l'étude de la régulation transcriptionnelle

En somme, il apparaît que l'étude des partenaires chromosomiques d'un locus est hautement informative pour comprendre les particularités de sa régulation transcriptionnelle. Si des méthodes comme le 4C permettent de cribler le génome pour les interacteurs préférentiels d'un locus d'intérêt, la quantification des fréquences d'interaction par FISH révèle que ces événements d'association ne concernent généralement qu'une fraction mineure de la population cellulaire à un instant donné (Simonis et al., 2006 ; Zhao et al., 2006). Une telle situation peut être expliquée soit par l'établissement d'interactions stables mais occasionnelles, qui ne concerneraient qu'un petit nombre de cellules, ou à l'inverse par une forte dynamique et un caractère transitoire des contacts, qui concerneraient une forte proportion de cellules au cours du temps mais un petit nombre de façon simultanée (Williams et al., 2010). De manière importante, ces deux scénarios prédisent des conséquences très différentes, en particulier lorsque ces interactions sont liées à la régulation transcriptionnelle des loci impliqués (voir **figure 11**). Le suivi en temps réel de la dynamique nucléaire de loci endogènes est un challenge technique chez les mammifères, car le double étiquetage nécessite deux événements de recombinaisons homologues, pour cibler chacun des partenaires de l'interaction d'intérêt. Le seul exemple rapporté concerne la visualisation de l'association physique inter-allélique des centres d'inactivations (*Xic*) du chromosome X qui précède l'activation transcriptionnelle mono-allélique de *Xist* (Masui et al., 2011). Cette étude a suggéré que la position relative des *Xic* en *trans* est hautement dynamique, bien que leur mobilité relative soit plus faible dans les cellules où ils sont à proximité – reflétant peut-être l'ancrage à une structure nucléaire.

Cependant, chaque locus est probablement particulier dans sa régulation et dans sa relation avec les autres régions de son chromosome, voire des autres chromosomes. Le développement de techniques d'étiquetage systématique permettra de mieux comprendre la dynamique de l'architecture du noyau. Central à ces questions est de comprendre comment le contexte structural et le statut transcriptionnel de régions chromosomique affecte et est affecté par la plasticité de l'organisation nucléaire. En outre, seule la comparaison de méthodes permettant d'affecter directement et sélectivement l'organisation génomique, la structure de la chromatine, l'activité transcriptionnelle ou la position nucléaire permettra de disséquer l'interdépendance de ces phénomènes, dont il est tentant d'anticiper qu'ils entretiennent des relations différentes en fonction du locus et du type cellulaire.



**Figure 11 Dynamique des interactions chromosomiques régulatrices.** Considérons l'existence d'interactions chromosomiques détectées dans une faible proportion des cellules à un instant donné, et qui corrèle avec l'activation transcriptionnelle d'un promoteur cible. Dans l'éventualité où cette conformation est stable, la population cellulaire évoluera de façon mosaïque, l'activation transcriptionnelle de la cible ne concernant que les quelques cellules au sein desquelles l'interaction a lieu (colonne de gauche). En revanche, si cette conformation est dynamique, le nombre de cellules dans lesquelles l'interaction est retrouvée reste le même au court du temps, mais l'identité des cellules impliquées change, si bien que la population évolue de façon concertée – la majorité des cellules ayant au final subit



# **ANNEXES A L'INTRODUCTION**





# Revue 1

*Regulation of X-chromosome inactivation by the X-inactivation centre*

Augui\*, Nora\* & Heard 2011

Cette revue illustre les complexités de la littérature concernant la régulation de l'expression de Xist, et est le fruit d'un long travail que nous avons effectué main dans la main avec Sandrine Augui, ancienne collègue et présente amie.



## Revue 2

### *Chromatin structure and nuclear organization dynamics during X-chromosome inactivation*

Nora and Heard, 2011

Cet article est une revue que nous avons eu l'occasion d'écrire à la suite de notre participation au soixante-quinzième symposium de biologie quantitative tenu au laboratoire de Cold Spring Harbor, en 2010, sur l'organisation du noyau et ses fonctions. Il fait partie du volume correspondant édité par CSHL press, et résume la cascade d'évènement déclenchée par l'expression de Xist, et aboutissant à l'inactivation transcriptionnelle du chromosome X.



# **Contexte des projets développés et résumé des travaux**



Mon travail doctoral s'est intéressé au contexte chromosomique dans lequel la régulation de la transcription prend place. L'étude de l'inactivation du chromosome X nous a fourni un modèle de choix pour les raisons que nous allons brièvement détailler.

Les mécanismes contrôlant l'initiation de l'inactivation chez la souris convergent vers le contrôle transcriptionnel de *Xist*, situé dans le centre d'inactivation du chromosome X (*Xic*; voir **Figure 12**). Les études génétiques visant à déterminer la région génomique minimale capable de récapituler le patron d'expression développemental de *Xist* ont mis en évidence que l'activité transcriptionnelle du locus est contrôlée en *cis* et en *trans* par plusieurs éléments génomiques appartenant à l'intervalle *Xic* (voir Augui et al., 2011, ci-joint, pour revue). Ce locus représente donc un modèle de choix pour comprendre comment l'intrication de multiples afférences régulatrices permet d'orchestrer la dynamique développementale de l'activité transcriptionnelle.

### **Expression mono-allélique de *Xist* et l'importance du locus *Tsix***

La région de contrôle la mieux caractérisée est responsable de la production du transcrit antisens à *Xist*, *Tsix*, responsable de la répression de *Xist* à la fois en *cis* directement et en *trans* indirectement (voir Augui et al., 2011 en annexe de l'introduction pour revue). La transcription à *Tsix* est elle-même modulée par plusieurs éléments, notamment *Xite* et *DXPas34*, qui contiennent des sites de fixation pour plusieurs facteurs *trans* dont l'altération perturbe l'expression de *Xist*. L'ablation hétérozygote de *Tsix* entraîne non seulement l'activation en *cis* de *Xist* mais aussi sa répression en *trans*, menant à l'élection systématique du chromosome mutant comme X inactif (voir Augui et al., 2011 pour discussion de ce phénomène de rétrocontrôle). La délétion homozygote de *Tsix* diminue l'efficacité globale de l'activation transcriptionnelle de chaque allèle de *Xist* d'une part, et dans une proportion importante des cas celle-ci se fait de façon concomitante sur les deux allèles – révélant ainsi une sorte d'effet coopératif en *trans*. Une telle coordination en *trans* des patrons d'expression de *Xist* apparaît faire intervenir des événements à la fois en amont et en aval de l'activation transcriptionnelle de *Xist*. En effet, avant que l'expression de *Xist* ne soit détectée, les deux *Xic* ont tendance à être fréquemment observés à proximité dans le noyau - comportement perdu dans les mutants *Tsix*; le suivi en temps réel de la dynamique nucléaire de ces loci a en outre montré que suite à la rencontre physique transitoire des *Xics* l'expression de *Tsix* est altérée de façon mono-allélique, créant une fenêtre d'opportunité pour l'activation mono-allélique de *Xist* (Masui et al., 2011). L'expression de l'ARN *Xist* entraîne l'inactivation transcriptionnelle et des changements de la structure chromatinienne en *cis*, résultant en la diminution de l'expression de facteurs *trans*-régulateurs de *Xist* liés au X, comme RNF12; ce type de boucle de contrôle permettrait ainsi de diminuer la probabilité d'activation transcriptionnelle des allèles de *Xist* une fois que l'un a été activé, assurant que l'inactivation du chromosome X soit mono-allélique (voir Nora & Heard 2009 en

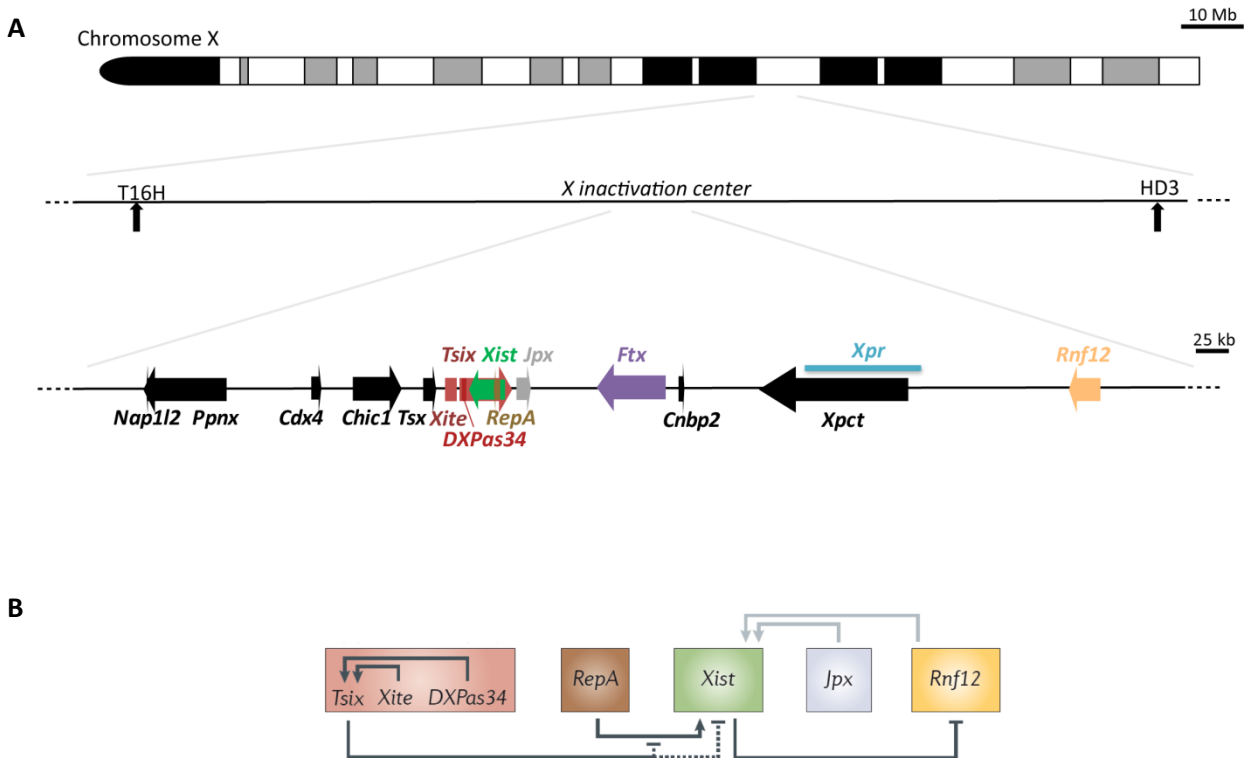


annexe pour plus de détails). Si des mécanismes contrôlant l'activation transcriptionnelle de *Xist* ont commencé à être identifiés, les bases moléculaires sous-tendant son expression mono-allélique sont largement incomprises. *Tsix* joue clairement un rôle central dans ce processus, mais les mécanismes à l'origine du traitement différentiel des deux allèles de *Tsix* après leur appariement transitoire restent à appréhender.

- **Quels sont les partenaires physiques contactés par *Tsix*, à part son homologue en *trans* ?**
- **À quel titre ces contacts chromosomiques participent-ils à ou témoignent-ils de l'activité transcriptionnelle du locus *Tsix* à l'échelle allélique ?**

### **Caractérisation de l'organisation chromosomique du locus *Tsix* et de sa variabilité inter-allélique**

Guidé par les études génétiques qui ont démontré l'importance de la région *Tsix* dans le contrôle de l'activation transcriptionnelle mono-allélique de *Xist*, nous avons tout d'abord voulu avancer dans la description de l'organisation structurale de ce locus, et en particulier de sa variabilité inter-allélique. Nous avons ainsi développé une approche permettant de caractériser l'organisation topologique du locus *Tsix*, dont nous rapportons les résultats dans la section suivante (« Résultats supplémentaires »). L'identification des partenaires chromosomiques préférentiels de *Tsix*, par capture circulaire de conformation chromosomique (4C), a révélé la prévalence – si ce n'est l'exclusivité – des contacts en *cis* sur le chromosome X. Ce comportement corrèle avec le positionnement de ce locus à l'intérieur du territoire du chromosome X. L'étude de la signature épigénomique des interacteurs préférentiels de *Tsix* a révélé que ces associations chromosomiques sont probablement sous-tendues par divers mécanismes. En particulier, les protéines impliquées dans l'appariement homologue en *trans* des allèles *Tsix* ne sont pas systématiquement retrouvées à ses partenaires chromosomiques contactés en *cis*. Enfin, les techniques de DNA FISH nous ont permis de montrer que les interacteurs préférentiels de *Tsix* sont contactés à faible fréquence à travers la population cellulaire, si bien que l'environnement génomique des deux allèles *Tsix* dans le même noyau à un instant donné est différent. Nos observations soulignent le fait que deux loci identiques au point de vue de leur séquence ADN ne sont pas nécessairement structurellement équivalents à tout instant. Nous discutons enfin des implications de telles observations dans le cadre du phénomène d'activation transcriptionnelle aléatoire de l'un ou l'autre des deux allèles *Xist*, et du déclenchement mono-allélique de l'inactivation du chromosome X.



**Figure 12 Le centre d'inactivation du chromosome X (*Xic*)** est une région de plusieurs mégabases, identifiée par l'étude de translocations et délétions chromosomiques sur la base de sa capacité à orchestrer l'initiation de l'inactivation du chromosome X. Elle est notamment responsable de la production de l'ARN *Xist*, dont l'expression entraîne la répression transcriptionnelle en *cis*. Le *Xic* contient en outre les éléments *cis*-régulateurs nécessaires à la définition du patron transcriptionnel de *Xist*, dont certains ont été identifiés, comme *Tsix*, *Xite*, *DXPas34* et *RepA*. Le *Xic* contient également des loci contrôlant en *trans* l'expression de *Xist* via divers mécanismes, comme *Jpx* et *Ftx* (production d'un ARN non codant), *Xpr* (appariements chromosomiques) et *Rnf12* (codant une ubiquitine E3 ligase). **(A)** carte physique du *Xic* ; les éléments impliqués dans la régulation de l'activité transcriptionnelle de *Xist* sont figurés en couleur **(B)** réseau régulateur au locus *Xist*. Noir : régulation en cis ; gris : régulation en trans. Adapté de Augui\*, Nora\* et Heard, 2011

## Éléments de contrôle à distance et complexité du locus *Xic*

Si *Tsix* joue clairement un rôle crucial dans la régulation de l'expression de *Xist*, d'autres régions du *Xic* sont impliquées dans le contrôle de l'initiation de l'inactivation. En effet, de larges transgènes contenant *Xist* et *Tsix* sont incapables d'exprimer correctement *Xist*. Nous avons donc souhaité développer une approche permettant d'appréhender le paysage *cis*-régulateur du *Xic* dans son ensemble.

- **Quels sont les éléments du *Xic* qui participent à la régulation de l'expression de *Xist* ?**
- **Quelle est l'étendue fonctionnelle du *Xic* ?**

## Caractérisation de l'organisation structurale du *Xic* et des influences régulatrices qu'elle sous-tend

Nous avons ainsi entrepris de cribler le *Xic* pour des éléments régulateurs de *Xist*, et nous avons pour ce faire choisi de cartographier à grande échelle les interactions physiques établies entre les différentes parties du *Xic*. Qui plus est, cette stratégie nous a permis d'étudier l'architecture chromosomique à l'échelle du mégabase et en deçà, niveaux d'organisation jusqu'alors mal caractérisés de façon générale. Nous avons ainsi appliqué la récente technique de copie conforme de capture de conformation chromosomique (5C), qui nous a permis de dresser une carte topologique des quatre mégabases qui encadrent *Xist*. Cette région englobe tous ses éléments régulateurs connus ainsi que les intervalles dans lesquels la présence d'éléments régulateurs additionnels a été génétiquement prédite. Les résultats de ces travaux sont réunis dans le premier manuscrit accompagnant la section suivante.

Nous avons mis en évidence que la topologie chromatiniennne ségrégue physiquement le promoteur de *Tsix* d'une part et de promoteur de *Xist* d'autre part. Ainsi, le promoteur de *Xist* contacte fréquemment les loci *Jpx*, *Ftx*, *Xpr* et *Rnf12* alors que les partenaires physiques locaux du promoteur de *Tsix* impliquent *Xite* et jusqu'à 200kb en amont. En comparant les patrons d'expression de *Tsix* à partir de transgènes comprenant ou non ces régions nous montrons qu'elles abritent un ou des éléments de contrôle en *cis* de *Tsix*, démontrant ainsi que les informations structurales obtenues par les cartes topologiques de la chromatine peuvent être utilisées afin de prédire des relations régulatrices des loci sous-jacents. L'exploration des cartes transcriptomiques et épigénomiques de la région nous a permis d'identifier un ARN nucléaire jusqu'alors non caractérisé, que nous avons baptisé *Linx*, produit par une des régions interagissant avec *Tsix/Xite* et dont la co-expression en *cis* avec *Tsix* pourrait être compatible avec une fonction régulatrice.

Par ailleurs, en rapportant que le même jeu d'éléments distaux contacte *Xist* et *Tsix* dans les cellules mâles et les cellules femelles, nous suggérons que les facteurs *trans* spécifiques aux femelles ne sont pas capables de remodeler le paysage structural du *Xic*, mais exercent leur action dans un contexte topologique déterminé indépendamment de la transcription. L'utilisation de cellules mâles sur-exprimant un facteur *trans* activant *Xist*, dans lesquelles la conformation topologique est identique aux cellules sauvages, confirme que la topologie du locus est déterminée de façon indépendante de l'activation transcriptionnelle de *Xist*. Ces résultats suggèrent ainsi que l'activation transcriptionnelle de *Xist* pourrait non seulement nécessiter un jeu de facteur *trans* adéquat, mais également une conformation chromosomique en *cis* appropriée.

Enfin, nous avons mis en évidence qu'à l'échelle de la mégabase et en deçà, le repliement de la chromatine ne suis absolument pas un modèle de marche aléatoire. Au contraire, le *Xic* est partitionnée en blocs contigus à l'intérieur desquels les interactions chromosomiques sont préférentiellement établies. En appuyant le dessin de sondes de DNA-FISH sur nos cartes de 5C, nous avons pu observer ce comportement par microscopie à haute résolution. La position de ces domaines d'interactions chromosomiques correspond avec précision avec le patron de répartition de plusieurs composants intrinsèques de la chromatine, notamment H3K27me3 et H3K9me2. Toutefois, l'organisation topologique du *Xic* ne semble pas affectée en leur absence. Un certain dynamisme des patrons conformationnels est cependant observé au cours de la différenciation cellulaire. En comparant la topologie du *Xic* entre les chromosomes X actifs et inactifs, nous apportons la preuve que l'hétérochromatinisation s'accompagne de changements architecturaux à l'échelle du mégabase, se manifestant notamment par la perte des interactions préférentielles à longue distance.

Les perspectives ouvertes par nos travaux seront ensuite discutées, à la fois en termes de la compréhension de la régulation de *Xist* et de l'initiation de l'inactivation du chromosome X, mais aussi plus généralement en termes des mécanismes fondamentaux reliant les propriétés structurales du génome et le contrôle de l'activité transcriptionnelle.



# **MATERIELS ET METHODES**



## **Culture cellulaire**

### **1. Cellules souches embryonnaires indifférenciées (mESCs)**

Toutes les mESCs utilisées ont été maintenues en atmosphère humidifiée à 37°C et 8% de CO<sub>2</sub>. Les mESCs XY de la lignée E14 (Hooper et al., 1987) ont été obtenue par voie commerciale et cultivées d'après les conseils du fournisseur sans fibroblaste nourricier sur flacons gélatinisés (Milieu Minimum Glasgow [Gibco], 15% sérum de veau foetal [invitrogen], 10<sup>-4</sup> M de β-mercaptoethanol [Sigma], 2mM de L-Glutamine, 0,1mM d'acides aminés non-essentiels et 1mM de pyruvate de sodium, complémenté par 1000U/mL de Leuhemia inhibitory factor [LIF, Chemicon]). Les mESCs XX stables de la lignée PGK12.1 (Penny et al., 1996) sont un don du Dr Neil Brockdorff (Université d'Oxford) ; les cellules mutantes EED (allèle I7Rn5-3354SB, Morin-Kensicki et al., 2001) sont un don du Dr Terry Magnuson ; elles ont été cultivées dans les conditions correspondant à leur dérivation, sans fibroblaste nourricier sur flacons gélatinisés (Milieu Eagle modifié de Dulbecco [DMEM, Gibco], 15% sérum de veau foetal [invitrogen], 10<sup>-4</sup> mM de β-mercaptoethanol [Sigma] complémenté par 1000U/mL de Leuhemia inhibitory factor [LIF, Chemicon]). Le même milieu de culture a été utilisé pour les lignées TT2 et TT2-G9a (clone 22.10 ;Tachibana et al., 2002) sont un don des Drs Joseph Wagstaff (Université de Virginie) et Yoichi Shinkai (Université de Kyoto), et ont été cultivées sur fibroblastes nourriciers.

### **2. Différenciation des cellules souches embryonnaires**

Pour l'étude des stades précoces (jour deux), la différenciation des mESCs est induite par deux lavages au PBS suivis de l'utilisation d'un milieu de culture particulier (Milieu Eagle modifié de Dulbecco [DMEM, Gibco], 10% sérum de veau foetal [invitrogen], 10<sup>-4</sup> mM de β-mercaptoethanol [Sigma], 100nM de *trans*-acide rétinolique).

Pour la dérivation de progéniteurs neuronaux (NPC) à partir des cellules E14, un protocole adapté de Conti et al., 2005 et décrit dans Splinter et al., 2011, joint à ce manuscrit, a été utilisé. Les cellules sontensemencées à raison de 15.000 cellules par cm<sup>2</sup> sur flacon gélatinisé et cultivées en monocouche adhérente en milieu N2B27 (StemCell Sciences) en 5% de CO<sub>2</sub> pendant sept jours. Elles sont ensuite décollées par traitement à la trypsine (Gibco), re-suspendues dans du N2B27 contenant 10ng/mL des facteurs de croissance EGF et FGF (Preprotech) et cultivées en suspension pendant trois jours. Les agrégats ont ensuite été déposés sur flacon gélatinisés d'où les NPC irradiant progressivement, en N2B27 complémenté en EGF et FGF.

### **3. Fibroblastes embryonnaires murins primaires**

Les fibroblastes embryonnaires murins (MEF) ont été produits par la dissociation via traitement à la trypsine d'embryons au stade E13.5 de fond génétique C57BL/6 et cultivés pendant trois passages (8% de CO<sub>2</sub> , milieu Eagle modifié de Dulbecco [DMEM, Gibco], 10% sérum de veau foetal [invitrogen], 10<sup>-4</sup> mM de β-mercaptoethanol [Sigma], 100nM de *trans*-acide rétinolique)



## **Copie Conforme de Capture de Conformation Chromosomique (5C)**

La 5C a été effectuée d'après le protocole détaillé dans Dostie et Dekker, 2007, avec certaines modifications.

### **1. Fixation des cellules**

Les cellules adhérentes sont fixées dans les boîtes de Pétri où elles ont été fixées à température ambiante pendant 10 minutes (dans du milieu Eagle modifié de Dulbecco [DMEM, Gibco], 10% sérum de veau fœtal [invitrogen], 10-4 mM de b-mercaptoethanol [Sigma] contenant 2% de para-formaldéhyde [Electron microscopy]). La fixation est arrêtée par l'ajout de Glycine en solution à concentration finale de 0.125M, suivi de 5 minutes d'incubation à température ambiante et 15 minutes sur glace. Deux lavages en Glycine 0,125M dans du PBS1X permettent d'éliminer les débris, cellulaires, et les cellules sont ensuite décrochées par grattage manuel, centrifugées à 950rpm à 4°C pendant 10 minutes et regroupées par lots de 5 à  $10 \times 10^7$  cellules.

### **2. 3C**

La procédure et les décrits sont décrits dans Dostie et Dekker, 2007, à l'exception des modifications suivantes. Points 45 et 47 : une centrifugation à 1000rpm permet d'empêcher la formation d'agrégats difficiles à re-suspendre. Point 53 : cette étape est remplacée par une incubation à 37°C pendant 1h à 800rpm dans un Thermomixer de paillasse (Eppendorf). Une incubation similaire est effectuée après l'étape 54. Point 55 : 1000U de l'enzyme HindIII fortement concentrée (NEB) ont été utilisés. L'étape 57 est remplacée par une incubation à 37°C pendant 1h à 800rpm dans un Thermomixer de paillasse (Eppendorf). Etape 58 : l'ATP et la ligase, à raison de 6.000U totales, ne sont pas ajoutés dans la réaction seulement après 30 minutes d'incubation à 37°C dans un agitateur de cultures bactériennes à 200rpm. Etape 59 : 4h d'incubation. L'étape 70 est précédée d'un lavage à l'éthanol 70°C qui permet de diminuer considérablement la précipitation des sels aux étapes ultérieures.

### **3. 5C**

1µg d'ADN produit par l'étape de 3C est utilisé par réaction de 5C. Pour chaque banque de 5C 16 réactions de 5C ont été effectuées en parallèle (étape 90) avec l'ensemble des oligonucléotides décrits dans l'annexe, chacun à concentration finale de 10 fM lors de l'étape d'annealing (point 94). L'ensemble des 17 réactions est utilisé pour 96 réactions de PCR (étape 97) où l'ampliTaq Gold et le tampon associé remplacent les réactifs de l'étape 97, pour 25 cycles lors de la PCR (étape 98). Le produit de l'ensemble de ces réactions est concentré par la purification sur deux colonnes d'affinités du kit PCR clean-up (Qiagen).

### **4. Préparation des échantillons de 5C pour le séquençage massif**

3µg de la banque de 5C sont utilisés pour une réaction d'A-tailing classique suivie de la ligation des amorces commerciales Illumina dédiées au séquençage massif par extrémités appareillées (Paired-end), dans les conditions recommandées par le fournisseur. Le produit de réaction correspondant à la banque de 5C liguée à deux amorces est purifié par électrophorèse en gel d'agarose grâce au kit Qiagen correspondant. Six réactions de PCR en 25µL contenant chacune environ 7,5ng du produit de purification sont ensuite effectuées avec les réactifs du kit Phusion PCR (Finnzymes/ThermoScientific)

en parallèle, avec les amorces pour séquençage en extrémités appareillés Illumina. Le produit PCR est à nouveau purifié sur gel et séquençé sur un appareil GenomeAnalyzerII (Illumina) au centre de séquençage NEMO du campus de UMASS à Worcester.

### **5. Visualisation et analyse de la 5C**

Les résultats du séquençage sont alignés au génome murin (assemblage mm9) grâce au logiciel Novoalign (<http://www.novocraft.com>). Entre 5 et 20 millions d'interactions sont typiquement alignées par expérience, chacune effectuée deux fois sur des échantillons biologiques différents. La visualisation des résultats est rendue possible par l'interface my5C (<http://my5C.umassmed.edu> ; Lajoie et al., 2009), une suite de programmes dédiés développés sous R (<http://r-project.org> ; N. Servant, Emmanuel Barillot, manuscrit en préparation), et le navigateur de génome de l'université de Santa Cruz (Kent et al., 2002 ; <http://genome.ucsc.edu>)

### **Capture de Conformation Chromosomique Circulaire (4C)**

La 4C a été effectuée d'après le protocole décrit dans Simonis et al., 2006. L'enzyme HindIII a été utilisé pour la première digestion, et NlaIII pour la seconde. Une PCR nichée a été utilisée sur les échantillons de 4C finaux afin d'amplifier spécifiquement les partenaires physiques du promoteur de Tsix :

Première PCR :

NH1F-XhoI AACTCGAGTCTCAAAGGCAAATGGGCTC  
HN1R-Not1 AAAAGCGCCGCACTGAGCAATCTGCCTAGGT

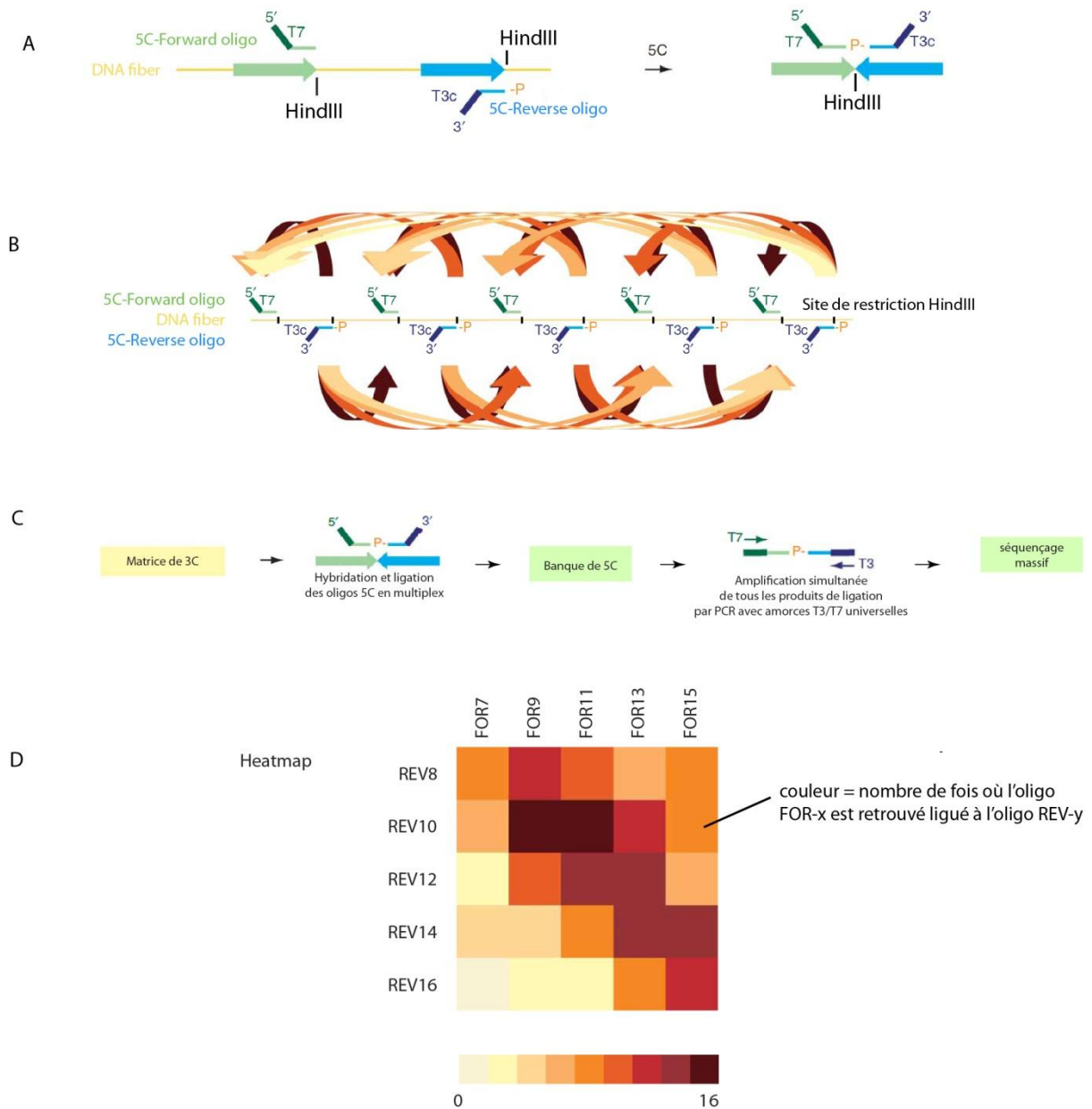
Seconde PCR :

HN2-F GAAATAGCCTCAGATCTGGTGA  
NH2-R ATTCTTTTCTGTGACTCCAC

Les sondes sur la puce ADN utilisée (NimbleGen) sont des oligonucléotides de 60bp homologues aux fragments HindIII-HindIII ou HindIII-NlaIII des chromosomes X, 3, 4, 7, 8 et 17 entiers, ainsi que plusieurs autres régions autosomales.

### **FISH et microscopie**

Les cellules sontensemencées sur des lamelles de verre et cultivées ainsi 24h. Pour les mESCs différenciées deux jours, le milieu de différenciation est ajouté après 24h de culture sur lamelle, et la culture est prolongée de 48h. La procédure de préparation des échantillons pour l'hybridation *in situ* en fluorescence de l'ADN ou de l'ARN est décrite en détail dans Chaumeil et al., 2008. L'imagerie a été effectuée sur un système de microscopie à champ large DeltaVision (Applied Precision) permettant la déconvolution grâce au programme propriétaire Softworx. Les sondes utilisées sont décrites dans Augui et al., 2007, joint à ce manuscrit. L'analyse positionnelle en trois-dimension a été effectuée grâce à un algorithme que nous avons développé sous ImageJ (<http://rsb.info.nih.gov/ij/>), faisant usage du plugin 3D object counter (Bolte et Cordelières, 2006), et permettant la détection semi-automatique des centres de masse des signaux de FISH et le calcul de la distance euclidienne qui les sépare.



**Figure 13 Principe détaillé de la Copie Conforme de Capture de Conformation Chromosomique (3C)** (A) Une collection d'oligo-nucléotides est créée, dont la particularité est de pouvoir s'hybrider à l'extrémité des sites de restrictions utilisés dans la procédure de de 3C (détaillée figure 6), qui entraîne la ligation des fragments de restrictions qui sont physiquement proches dans le noyau. Lorsqu'une telle interaction physique existe, les oligos de 5C vont alors s'hybrider nez-à-nez sur la matrice de 3C, et vont ainsi pouvoir être ligués ensemble. (B) Le dessin de nombreux oligos de 5C sur de grandes régions génomiques rend possible la mesure de la fréquence d'interaction entre tous les fragments de restriction Forward et Reverse, en parallèle. (C) le multiplexage est rendu possible par l'ajout d'extrémités T3 ou T7 sur chaque oligo Forward ou Reverse respectivement. La fréquence de ligation entre un oligo 5C Forward X et un oligo 5C Reverse Y est mesurée en séquençant directement la banque de 5C. (D) Cette valeur est visualisée sous forme de matrice bi-dimensionnelle où la fréquence de ligation est représentée par la couleur du pixel de coordonnée (x ,y). Adapté de Dostie & Dekker, 2007





# RESULTATS



# Article 1

**A structural basis for the regulatory landscape of the *X inactivation center***

Nora et al. – en préparation





## **Résultats supplémentaires**



# **ANNEXES AUX RESULTATS**



## Article 2

***Sensing X-chromosomes pairs before X inactivation via a novel X-pairing region of the Xic***

Augui, Fillion, Huart, Nora et al. 2009

Cet article décrit l'identification et la caractérisation initiale de la région *Xpr*, et décrit en particulier sa capacité à sous-tendre des événements d'appariements chromosomiques homologues avant l'activation transcriptionnelle de *Xist*. Il y est proposé que ces interactions chromosomiques participent à l'activation de *Xist* spécifiquement chez les femelles. Ma contribution a consisté à caractériser l'organisation nucléaire des chromosomes lors des événements d'appariements de *Xpr* par peinture chromosomique.



## Article 3

***The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA***

Splinter, de Wit, Nora et al. 2011

Cette collaboration avec le laboratoire de Wouter de Laat a permis de démontrer que l'architecture du chromosome X inactif se distingue de celle de son homologue actif par l'absence de conformation préférentielle. Des loci qui échappent à l'inactivation dérogent cependant à cette règle, et l'identification de leurs partenaires physiques permet ainsi de d'identifier de nouveaux évadés de l'inactivation. Enfin, l'ARN Xist est requis pour le maintien de cet état architectural particulier, et son absence entraîne une reconfiguration de cette organisation indépendamment de tout changement transcriptionnel. Ma contribution à cette étude a principalement consisté à effectuer des expériences de 3D-DNA FISH et d'appliquer des techniques d'analyse d'image que nous avons développées.





## **DISCUSSION ET PERSPECTIVES**



Les différentes études auxquelles j'ai participé se sont attachées à la caractérisation de l'architecture tridimensionnelle du *centre d'inactivation du X (Xic)*, de sa conformation locale (article 1) jusqu'à ses interactions avec le reste du chromosome X en *cis* (résultats supplémentaires et article 3) ainsi qu'en *trans* avec son homologue (article 2).

### **Cheminement de la démarche expérimentale**

Lorsque j'ai commencé mon travail de thèse, le seul élément *cis*-régulateur connu de l'expression de *Xist* était le locus *Tsix*. L'activité transcriptionnelle de *Tsix* jouant un rôle majeur dans le contrôle de l'initiation de l'inactivation du chromosome X, nous avons initialement cherché à caractériser de nouvelles propriétés structurales de ce locus. Peu avant mon arrivée au laboratoire, l'équipe avait montré que la région *Tsix/Xite* est capable d'engager des interactions chromosomiques en *trans* avec son homologue. Nous avons donc voulu savoir si ce locus était également capable d'établir d'autres types d'interactions chromosomiques, en *cis* avec le chromosome X ou en *trans* avec des autosomes. Pour cela, nous avons utilisé la méthode de 4C (Zhao et al., 2006) sur le promoteur de *Tsix*, en utilisant des puces ADN couvrant plusieurs chromosomes entiers, technique qui à l'époque n'avait été utilisée que pour l'étude des loci  *$\beta$ -globin* et *Rad23* (Simonis et al., 2006). Ce travail a été réalisé avec l'aide du laboratoire du Dr Rolf Ohlsson (Karolinska Institute).

Les premières expériences ont semblé révéler plusieurs interactions entre *Tsix* et des régions autosomales. Néanmoins, ne disposant pas du recul que les nombreuses études basées sur le 4C ont depuis apporté, nous avons passé beaucoup de temps à essayer de valider ces quelques interactions en *trans*, pour au final réaliser, après plusieurs réplicats biologiques, qu'elles étaient en réalité peu reproductibles. En l'occurrence, dans les cellules souches embryonnaires le locus *Tsix* s'est avéré se trouver la plupart du temps au cœur du chromosome X, comme l'ont montré par la suite des expériences de peinture chromosomiques par DNA FISH (résultats supplémentaires). *Tsix* apparaît donc préférentiellement contacter des régions sur son propre chromosomes.

Toutefois, *Tsix* ne s'associe pas avec n'importe quelle séquence sur le chromosome X. En dehors de son voisinage direct, où le signal 4C est toujours extrêmement fort, et ce quel que soit le locus étudié, plusieurs régions distantes de plusieurs dizaines de mégabases sont préférentiellement contactées. La fréquence avec laquelle *Tsix* rencontre ces régions, quantifiée par DNA FISH, s'est pourtant avérée assez faible. Nous nous sommes donc demandés si ce que nous pensions initialement refléter du signal aspécifique, au voisinage génomique direct de *Tsix*, ne renfermait pas en réalité ce que nous recherchions.

Afin de cribler à haute résolution le voisinage de *Tsix*, nous avons besoin d'une technologie plus résolutive que le 4C sur puce. Nous avons alors décidé d'entreprendre la cartographie des interactions chromosomiques internes au *Xic* par 5C, méthode qui nous permettait également de nous intéresser à d'autres loci en parallèle (Dostie et Dekker, 2007). Une approche permettant d'appréhender la conformation globale du *Xic*, y compris *Tsix* et *Xist*, nous a donc semblé la plus appropriée. Au vu de l'ensemble des résultats de 3C publiés à l'époque sur des échelles de la centaine de kilobases, nous nous attendions à n'identifier que quelques interactions à longue distance entre loci et séquences régulatrices.

La mise au point du 5C a été longue et délicate mais a aboutie grâce à l'aide des membres du laboratoire du Dr Job Dekker (University of Massachusetts Medical School). La première surprise que nous avons eue au vu des cartes topologiques que nous avons dressées dans les cellules souches embryonnaires (ESC) est l'ampleur du réseau structural définissant la conformation du *Xic* (article 1). Certes, quelques interactions très spécifiques entre deux fragments étaient détectées dans quelques cas – dont *Tsix* et *Xite* – mais après amélioration du protocole de 5C nous étions à même de détecter de façon reproductible des interactions sur près d'une mégabase dans certain cas. La deuxième surprise que nous avons rencontrée est la simplicité de ce réseau d'interaction : dans les cellules ES il est principalement partitionné en une série de longs blocs de taille variable (0.2 à 1Mb), à l'intérieur desquels peuvent exister d'autres sous-blocs où les interactions sont encore plus fréquentes.

### **Le partitionnement physique du génome**

Peu après nos premiers résultats, le laboratoire du Dr Job Dekker a décrit l'existence d'une organisation similaire à l'échelle du génome, grâce au développement de la méthode de Hi-C (Lieberman-Aiden et al., 2009). Cette étude suggère en outre que la fréquence de contact de deux régions à travers le génome est liée au contenu en G/C et à la richesse en séquences transcrites, faisant apparaître l'agrégation des interactions chromosomiques en blocs de plusieurs mégabases de long. Ces blocs semblent pouvoir interagir entre eux et se séparent en deux groupes, à l'intérieur desquels les contacts sont plus fréquents que les contacts inter-groupes.

En accord avec ces données, nous avons montré que *Tsix* interagit fréquemment avec des régions du chromosome X riches en G/C (résultats supplémentaires). En revanche, l'architecture interne du *Xic* que nous avons mise en évidence diffère du principe d'organisation avancé par Lieberman-Aiden et al. En effet, le 5C nous permet de regarder à des échelles inférieures à celle du mégabase, correspondant à la limite d'application du modèle de globule fractal déduit des données

de Hi-C (Lieberman-Aiden et al., 2009). Ainsi, nous avons montré que les blocs conformationnels que nous observons n'interagissent pas entre eux, même pour les plus distaux.

### **Conformation topologique et composition de la chromatine**

Au moment où nous avons obtenu nos premières cartes de 5C (article 1), le laboratoire venait d'achever une collaboration avec l'équipe du Dr Henk Stunnenberg en cartographiant l'enrichissement pour divers histones modifiées et facteurs de transcription généraux sur le chromosome X au cours de la différenciation des cellules souches embryonnaires (Marks et al., 2009). Ceci nous a donné l'opportunité d'intégrer la structure topologique du *Xic* à la composition biochimique de sa chromatine. Avec grande surprise à l'époque, nous avons vu que les domaines d'interactions dans nos cartes de 5C s'alignent avec les domaines chromatiniens, notamment en H3K9me2 et H3K27me3.

Peu après cette observation, plusieurs études ont indépendamment rapporté que de tels blocs d'enrichissement, faible mais reproductible, pour ces histones modifiées existent à travers le génome murin. En effet, plusieurs formes d'histones, en particulier H3K27me3 (Pauler et al., 2009) et H3K9me2 (Wen et al., 2009 ; Lienert et al., 2011), en plus d'être parfois enrichies à certains endroits ponctuels (généralement des promoteurs transcriptionnellement inactifs pour H3K27me3), sont retrouvées recouvrant de larges segments chromosomiques pouvant aller jusqu'à plusieurs centaines de kb contigus. Il n'est pas compris pourquoi l'enrichissement moyen dans ces domaines est globalement moindre que sur les pics ponctuels, en particulier pour H3K27me3 (Pauler et al., 2009). La variabilité intercellulaire, l'absence de positionnement préférentiel et une dynamique importante (par exemple au cours du cycle cellulaire) sont des explications envisageables. De manière intéressante, ces domaines H3K27me3 et H3K9me2 sont généralement distincts – sauf pour ceux dans le *Xic*, étrangement (Rougeulle et al., 2004 ; Marks et al., 2009) – reflétant la diversité des mécanismes par lesquels les histones méthyl-transférases correspondantes sont recrutées, et suggérant que plusieurs états structuraux régionaux existent. Il nous apparaît probable que ces blocs de H3K27me3 et H3K9me3 reflètent l'existence de blocs conformationnels d'une nature similaire à ceux que nous rapportons aux *Xic* à travers le génome. Cependant tous les blocs que nous voyons ne sont pas systématiquement définis par ces marques chromatiniennes, en particulier pour les plus petits où les interactions sont les plus fréquentes (article 1). Nous nous sommes donc intéressés aux autres propriétés structurales de la chromatine se manifestant sur des échelles similaires aux domaines d'interactions que nous avons mis en évidence.

## Les domaines associés à la lamine

La présence en continu sur de longs domaines génomiques n'est pas une propriété réservée à certains types d'histones. En effet, la cartographie des sites de liaison à la protéine LaminB1, intégrale au cortex nucléoplasmique, a révélé que les chromosomes peuvent être subdivisés en larges segments, où les événements de liaison à la lamine sont soit particulièrement fréquents (et sont alors appelés LAD pour *Lamina Associated Domains*), soit relativement occasionnels. De telles observations ont été rapportées à la fois chez l'humain, la souris et la drosophile, suggérant qu'elles reflètent des propriétés fondamentales de l'organisation de la chromatine chez les métazoaires (Guelen et al., 2008 ; Peric-Hupkes et al., 2010 ; Pickersgill et al., 2006). L'analyse du positionnement de ces régions dans le noyau par FISH confirme la tendance des LAD à être proches de la membrane. Elle suggère en outre que les domaines d'enrichissement émergent probablement du moyennage statistique des conformations différentes provenant des millions de cellules utilisées lors de la Dam-ID, plutôt que d'une organisation rigide et invariante de certaines portions chromosomiques à la lamine (Peric-Hupkes et al., 2010 ; Bas van Steensel et Jop Kind, communication personnelle). Les LAD correspondent généralement à des régions peu transcrites, pauvres en promoteurs, enrichies en H3K27me3 et en moindre mesure H3K9me2, sont généralement bordés par des îlots CpG, des promoteurs ou des sites de fixation à CTCF, et s'étendent de 100kb à plusieurs Mb chez la souris et l'humain (Guelen et al., 2008 ; Peric-Hupkes et al., 2010). L'analyse de la différenciation *in vitro* des cellules souches embryonnaires murines a montré que ces différents domaines épigénomiques (au moins pour H3K9me2 et les LAD) ne changent pas du tout au tout (Lienert et al., 2011; Peric-Hupkes et al., 2010). Si l'architecture basale de ces domaines est conservée au cours de la différenciation, leur organisation locale à plusieurs centaines de sites est cependant altérée, corrélant la plupart du temps avec des changements de l'activité transcriptionnelle des loci sous-jacents.

## Intégration des différentes propriétés structurales du génome

Là encore, la position des LADs s'aligne avec celle des domaines d'interaction que nous détectons par 5C sur le *Xic*, non seulement dans les cellules souches embryonnaires mais également dans les progéniteurs neuronaux issus de leur différenciation *in vitro* ainsi que dans les fibroblastes embryonnaires issus d'embryons murins (Peric-Hupkes et al., 2010). Un seul LAD est présent dans le *Xic*, alors que plusieurs blocs d'interaction chromosomiques existent ; ce LAD est entièrement compris dans un bloc conformationnel, et ses frontières sont proches de celles du cluster d'interaction. Au cours de la différenciation en progéniteurs neuronaux, le domaine d'interaction voisin du LAD présent dans les cellules ES devient lui aussi associé à la lamine, sans que ses

propriétés topologiques en terme d'associations chromosomiques semblent altérées. De manière intéressante, ces deux LAD voisins restent des blocs conformationnels distincts, suggérant que leur localisation à la membrane nucléaire n'augmente pas leur chance de se contacter physiquement. Dans les fibroblastes embryonnaires l'un de ces LADs est interrompu par une région de plusieurs dizaines de kilobases qui n'est pas associée à la lamine. L'analyse par 5C montre que les contacts physiques entre cette région et le reste du domaine d'interaction sont fortement diminuées, mais restent plus fréquents que ceux avec des régions en dehors du bloc conformationnel.

Ces observations suggèrent que le principe à l'origine du partitionnement du génome en domaines d'interaction exerce également son influence au niveau de plusieurs autres caractéristiques structurales, comme l'incorporation de modifications post-traductionnelles des histones composant la chromatine et son association à la lamine nucléaire. Les tests génétiques que nous avons effectués dans l'article 1 ont montré que la perte en H3K9me2 et H3K27me3 n'affecte pas le partitionnement en domaines conformationnels. Cela suggère donc que le phénomène à l'origine des clusters d'interaction chromosomiques agit en amont ou de façon indépendante aux enzymes catalysant ces modifications chromatiniennes. La précision de l'alignement entre l'agencement en domaine des interactions chromosomiques et de l'enrichissement en H3K27me3 ou H3K9me2 nous amène cependant à penser que ces phénomènes sont liés, et donc à privilégier la première de ces hypothèses. Il sera donc intéressant d'essayer de tester si l'organisation topologique de la chromatine peut influencer sa structure biochimique, en analysant par exemple des remaniements génomiques intervenant dans les domaines d'interaction et en s'intéressant à la perturbation éventuelle de la répartition de H3K27me3, H3K9me2 et de l'association à la lamine.

### **Les régions de réplication précoce ou tardive**

Le programme de réplication (traduction de l'anglais *replication timing*) est une autre propriété structurale du génome qui se manifeste de façon régionale, et qui est liée à la période de la phase S à laquelle des groupes d'origines de réplication s'activent de façon coordonnée (pour revue voir Pope et al., 2010). La cartographie des régions génomiques précoces ou tardives à la réplication chez la souris, l'humain et la drosophile a en effet montré une concentration en larges domaines (jusqu'à 100kb chez la mouche, entre 200kb et 2Mb chez les mammifères), dont la composition génomique et épigénomique est distincte (Hiratani et al., 2008 ; Ryba et al., 2010 ; Eaton et al., 2011). Il a été proposé à partir d'expériences de DNA FISH que les régions devenant tardivement répliquées sont relocalisées dans le noyau en position plus périphérique, et inversement pour celles devenant répliquées précocement. Les LAD sont en effet des régions qui tendent à être répliquées



tardivement (Pickersgill et al., 2006), bien que ne s'alignant pas parfaitement avec les domaines à réplication tardive, notamment au niveau de leurs transitions (Peric-Hupkes et al., 2010). De manière intéressante l'élimination de l'histone méthyl-transférase G9a, responsable de la présence de H3K9me2 au niveau de certaines régions du génome préférentiellement positionnées à la périphérie du noyau des cellules ES murines (Wu et al., 2005), entraîne l'activation transcriptionnelle ponctuelle de certains loci dans des domaines à réplication tardive, mais sans provoquer de changement dans leur programme de réplication ni de repositionnement nucléaire (Yokochi et al., 2009). Ces observations suggèrent ainsi que si la structure de la chromatine, la position dans le noyau et le programme de réplication sont corrélés, ces propriétés ne sont pas forcément interdépendantes.

La taille des domaines de réplication est bien supérieure à celle des domaines topologiques que nous rapportons dans l'article 1 (1-2 à 2-3 mégabases en fonction du type cellulaire; Hiratani et al., 2008). Le programme de réplication a en effet déjà été corrélé avec un niveau d'organisation topologique supérieur du génome, au-dessus du mégabase (Ryba et al., 2010). Cependant, il a été rapporté que la différenciation cellulaire s'accompagne de changements locaux du programme de réplication, concernant principalement l'homogénéisation des profils par l'adoption par les petits domaines du comportement de ses voisins plus grands (Hiratani et al., 2008 ; Hiratani et al., 2010). La taille des domaines de réplication qui passent de tôt à tardif, ou inversement, correspond à celle des domaines conformationnels que nous identifions dans l'article 1 (Hiratani et al., 2008). Le programme de réplication étant stable à travers les différents types cellulaires testés dans le *Xic* (sauf sur le chromosome X inactif, globalement répliqué de façon tardive), nous n'avons pu formellement tester l'hypothèse que les domaines d'interaction chromosomiques pourraient sous-tendre le comportement modulaire de certaines régions du génome à changer de programme de réplication. Une analyse par 5C telle que celle que nous avons utilisée pour le *Xic*, mais appliquée à des loci au profil de réplication plus dynamique au cours de la différenciation cellulaire, permettrait de répondre directement à cette question.

### **L'architecture chromosomique comme nouvelle signature des états chromatiniens ?**

Les analyses systématiques des propriétés structurales de la chromatine, en particulier de la constitution et de l'organisation des nucléosomes ainsi que des sites de fixation préférentiels de facteurs *trans* se sont considérablement développées au cours de la dernière décennie. De manière intéressante, la considération combinatoire des nombreuses annotations épigénomiques maintenant disponibles chez de nombreux organismes fait émerger un nombre limité de patrons structuraux, dont l'intégration avec les données transcriptomiques permet l'indexation des différentes régions du

génomique de façon tissu-spécifique (Birney et al., 2007; Ernst et Kellis, 2010; Ernst et al., 2011 chez l'humain. Fillion et al., 2010; Ernst et Kellis, 2010; Nègre et al., 2011; The modENCODE Consortium et al., 2010; Xie et al., 2007 chez *D. melanogaster*. Gerstein et al., 2010 chez *C. elegans* ; Roudier et al., 2011 chez *Arabidopsis thaliana*). La comparaison de ces états structuraux entre différents types cellulaires et stades de développement permet par exemple d'identifier des segments génomiques où la dynamique est corrélée, et d'inférer ainsi des relations régulatrices.

Les résultats de ce type d'analyse intégrative ne sont pas encore disponibles pour le génome murin, et nous attendons avec impatience de pouvoir les confronter aux données topologiques produites dans les articles 1, 2 et 4. Plus généralement, nous pensons que les caractéristiques conformationnelles de la fibre d'ADN pourraient faire partie des propriétés structurales utilisées par la cartographie intégrative des états chromatinien (Ernst et Kellis, 2010). Ce défi computationnel est l'un des objectifs de la nouvelle phase du projet ENCODE (The ENCODE Project Consortium, 2011).

### **Comment relier structure et fonction du génome à grande échelle ?**

Dans l'article 1, nous étayons l'idée que l'étude de la structure chromatinienne au sens large (topologie, modifications des histones et liaison de facteurs *trans*) peut être utilisée pour mettre en évidence des relations régulatrices entre des éléments génomiques distants. En émettant l'hypothèse, raisonnable, que le partitionnement en blocs conformationnel du *Xic* reflète une propriété générale de l'organisation topologique du génome, il devient intéressant de se demander, d'une part, à quel point la régulation fonctionnelle à longue distance en termes transcriptionnels est un phénomène courant et, d'autre part, si les influences régulatrices sont également partitionnées en domaines génomiques.

### **Activité transcriptionnelle et voisinages génomiques**

L'analyse des premières données transcriptomiques a révélé des similitudes inattendues quant aux patrons transcriptionnels de loci adjacents dans le génome. En effet, les ARN qu'ils peuvent produire ont tendance à être co-exprimés dans les mêmes types cellulaires ou aux mêmes stades développementaux. Ce phénomène a été initialement observé lors de l'examen des premières banques d'EST (témoins de séquences transcrites pour la traduction de l'anglais *Expressed Sequence Tags*) chez la drosophile (Spellman et Rubin, 2002 ; Boutanaev et al., 2002). Chez les mammifères, des analyses initiales par SAGE (analyse en série de l'activité transcriptionnelle pour la traduction de

l'anglais *Serial Analysis of Gene Expression*) et banques d'EST ont révélé que si la tendance majeure à la co-expression concerne des groupes de loci contigus transcrits de façon constitutive (Lercher et al., 2002), ils impliquent également des régions où plusieurs unités de transcription sont actives de concert mais de façon tissu-spécifique.

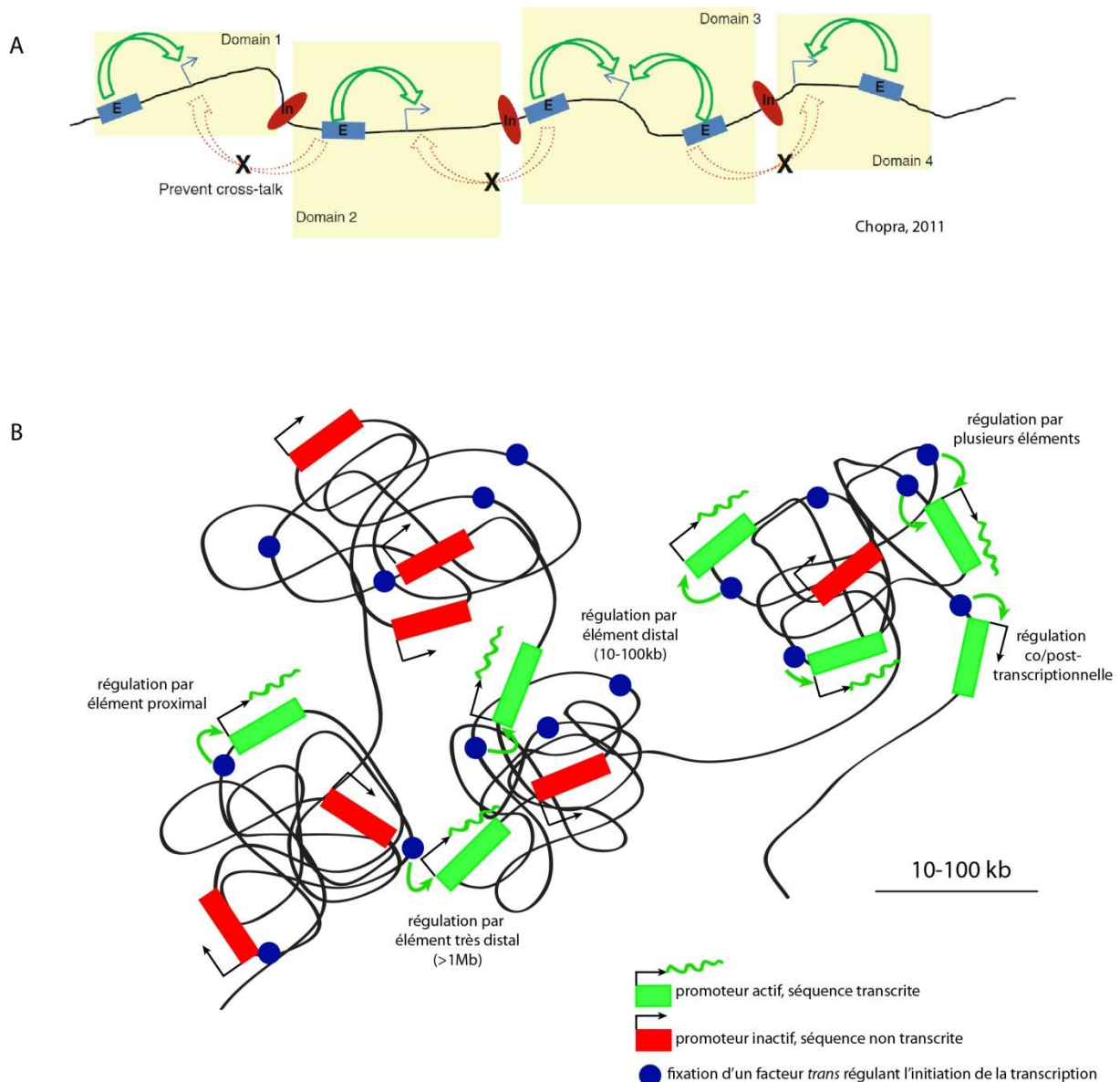
Une partie de ces loci adjacents co-régulés concernent des régions issues de duplications en tandem. Les éléments génétiques  $\gamma$  sont structurellement semblables, car évolutivement apparentées, et sont généralement impliqués dans des mécanismes biologiques similaires – suivant leur état de dérivation. Dans certains cas, le maintien de leur organisation en groupe témoigne d'une pression de sélection importante, imputable aux mécanismes de contrôle de leur activité transcriptionnelle. Les loci  $\alpha$ -globin,  $\beta$ -globin et *Hox* déjà évoqués en sont des exemples bien documentés, même si leurs patrons transcriptionnels sont contrôlés par des éléments régulateurs de nature complètement différente (pour revue détaillée de la complexité de ces loci paradigmatiques, voir Higgs et al., 2008, Palstra et al., 2008a et Spitz et Duboule, 2008).

Ce type de région dupliquée en tandem représente cependant une fraction mineure du génome, la majorité des paires de séquences codantes adjacentes n'étant pas évolutivement apparentées. La tendance à la co-expression de leurs produits transcriptionnels peut d'autant plus surprendre à la lumière de l'observation que des loci codant des protéines impliquées dans les mêmes voies biologiques sont fréquemment voisins dans les génomes eucaryotes – sans pour autant que ces arrangements soient nécessairement conservés entre les taxons éloignés (Lee et Sonnhammer, 2003). Il est ainsi tentant de spéculer que du fait de leur proximité génomique, plusieurs loci pourraient être soumis à des influences régulatrices similaires – ce qui favoriserait ainsi leur co-expression et rendrait possible l'interaction physique et fonctionnelle de leurs produits ribonucléiques ou protéiques. De manière intéressante, un tel scénario pourrait expliquer la fixation évolutive de tels arrangements génomiques une fois de telles relations fonctionnelles tissées, puisque les évènements de recombinaison séparant ces loci aboliraient leur capacité à être co-exprimé et seraient contre-sélectionnés. De nombreuses études à travers les phyla eucaryotes ont maintenant confirmé la généralité de cette tendance à la transcription concertée des loci juxtaposés (pour revue voir Michalak, 2008). Si cette corrélation est statistiquement très significative, il est important de garder à l'esprit qu'elle est généralement de faible amplitude. Si l'impact de la régulation à l'échelle de domaine chromosomique n'apparaît donc pas le déterminant majeur des patrons transcriptionnels, son effet n'en est pas moins suffisant pour contrôler des phénotypes soumis en toute vraisemblance à sélection évolutive.

Ces idées sont étayées par plusieurs évidences expérimentales que nous allons maintenant détailler, et qui ont révélé que des promoteurs insérés de façon ectopique dans diverses régions du génome des mammifères ont un niveau d'activité qui dépend de leur région d'insertion.

### **Cadastre des influences régulatrices et voisinages transcriptionnels**

S'il est connu depuis longtemps que l'efficacité d'expression de rapporteurs transgéniques dépend de leur site d'insertion dans le génome (Gossen et Bujard, 1992), la première étude ayant rapporté des observations systématiques chez les mammifères est relativement récente (Gierman et al., 2007). La comparaison d'une centaine de lignées cellulaires HEK293 (dérivées de tissu rénal embryonnaire humain), chacune porteuse d'une intégration différente d'un rapporteur (*green fluorescent protein - gfp*) sous promoteur minimal (*phospho-glycerate kinase - pgk*), a ainsi montré que l'activité transcriptionnelle varie en fonction du contexte génomique. De manière intéressante, le taux d'expression des transcrits issus des rapporteurs corrèle avec celui des régions environnantes, et peut différer grandement en fonction des sites d'intégration. Cette corrélation est plus marquée lorsque l'on regarde la moyenne du taux d'expression des vingt à quarante unités de transcription voisines plutôt que celle des régions qui flanquent directement l'intégration. Cette observation surprenante suggère que la dépendance du contexte génomique sur l'expression de ces rapporteurs reflète d'avantage l'existence de variations importantes dans les influences régulatrices régionales du génome que leur soumission directe aux éléments de contrôle locaux. Si le mécanisme à l'origine de ces influences transcriptionnelles à grande échelle n'est pas élucidé, il a été remarqué que les insertions les plus actives transcriptionnellement résident dans des domaines génomiques où la chromatine est plus décondensée et occupe une position plus centrale dans le noyau (d'après des mesures d'hybridation in situ de l'ADN en trois dimensions). Le génome apparaît ainsi partitionné en larges blocs à l'intérieur desquels s'exercent différentes influences régulatrices sur l'activité transcriptionnelle. Est-il possible d'imaginer cartographier ces influences régulatrices à l'échelle du génome, et leur éventuelle dynamique développementale ?



**Figure 14 Modèle de travail (A)** Représentation du phénomène de paysage régulateur tiré de Chopra, 2011. Les éléments génomiques (boîtes bleues E) où des facteurs *trans* se fixent peuvent moduler l'activité transcriptionnelle des promoteurs environnant (flèches fines bleues), sauf si un élément bloquant leur effet intervient (ovales rouges Ins), ce qui aboutit au partitionnement du génome en domaines de régulation distincts. **(B)** À la suite de nos travaux, nous proposons une évolution de ce modèle qui rend compte du partitionnement physique des interactions chromosomiques aux échelles inférieures au mégabase. Le repliement globulaire de la chromatine qui peut rendre compte de ce phénomène explique facilement le partitionnement conséquent des influences régulatrices. Ce modèle fait plusieurs prédictions testables, telles l'existence de globules chromatiniens, la diversité mécanistique des effets d'isolation (changement de globule, organisation interne du globule etc...) ou encore l'émergence des paysages régulateurs sur la base de l'existence de globules chromatiniens.

## Identification et dissection des cadastres régulateurs par le système GROMIT

Les cribles de capture de séquences amplificatrices de la transcription (traduction de l'anglais *enhancer-trap*) permettent d'avancer vers ce but, mais ils ne peuvent prétendre à l'exhaustivité puisqu'il est impossible de les faire arriver à saturation, particulièrement chez les mammifères, étant donnée la grande taille de leurs génomes. Un élégant système de dissection des influences régulatrices locales dans le génome murin a cependant été récemment mis au point, permettant de combiner leur détection à l'induction de remaniements génomiques (Ruf et al., 2011). Cette méthode, nommée GROMIT (cartographie de l'organisation régulatrice du génome par intégration de transposons pour l'anglais Genome Regulatory Organization Mapping with Integrated Transposons) est basée sur l'insertion aléatoire d'un élément mobile particulier, dérivé de *Sleeping Beauty* et porteur d'un rapporteur à promoteur minimum ( *$\beta$ -globin*) et d'un site Lox. La capacité de cet élément à transposer préférentiellement en *cis* permet de générer des collections d'insertions denses autour du point d'intégration initial et, en combinaison avec à l'expression inductible de la recombinase CRE, rend possible d'entraîner des réarrangements génomiques locaux dans ces mêmes régions. De manière surprenante, l'analyse de la centaine d'insertions initiales – distribuées aléatoirement dans le génome – révèle une activation tissu-spécifique transcriptionnelle du rapporteur dans la majorité des cas, même lorsque l'insertion est située à plusieurs centaines de kilobases de toute séquence codante. Le génome apparaît donc parcouru d'influences régulatrices, y compris dans ses portions où elles passent habituellement inaperçues de par l'absence de promoteur activable, mais qui deviennent évidente par l'utilisation d'un système rapporteur approprié.

Cet effet est-il simplement explicable par l'activation tissu-spécifique d'amplificateurs transcriptionnels agissant à distance ? L'effet intrinsèque de ces éléments, mesurable par des rapporteurs en contexte ectopique, révèle que la majorité des insertions localisées près d'éléments de contrôle connus est en effet soumise à leur contrôle transcriptionnel (Ruf et al., 2011). Pour tous les éléments analysés, des différences subtiles mais reproductibles existent cependant, suggérant que la précision des patrons transcriptionnels résultent en réalité non pas de l'effet d'un élément de contrôle unique, mais de la convolution de multiples afférences régulatrices, activatrices et inhibitrices. La remobilisation des transposons rapporteurs révèle une grande variabilité quant à la dynamique avec laquelle ces influences régulatrices changent le long du génome : certains rapporteurs situés à plus d'un mégabase sont parfois exprimés dans des tissus très similaires, alors que d'autres séparées de seulement quelques dizaines de kilobases ont parfois des patrons d'expression complètement différents. Les transitions semblent moins abruptes dans les régions peu denses en promoteurs endogènes - probablement reflétant que l'hétérogénéité de répartition des séquences *cis*-régulatrices (à effet activateur ou isolateur) suit celle des promoteurs. Le croisement

d'animaux porteurs de plusieurs insertions du système GROMIT permettant d'induire des remaniements chromosomiques, *via* induction de la recombinaise CRE fournit enfin ainsi l'opportunité d'étudier l'effet de l'altération de l'environnement génomique des rapporteurs. La disponibilité de plusieurs centaines d'insertions remobilisables fait de l'approche GROMIT un des systèmes particulièrement prometteur pour avancer dans la compréhension de l'architecture des influences régulatrices du génome murin, et de leur dynamique développementale.

L'induction de remaniements chromosomiques entre les sites d'insertion des constructions GROMIT pourrait fournir le cadre expérimental dans lequel tester la robustesse des principes gouvernant la conformation chromosomique. Il serait particulièrement intéressant d'induire par exemple la délétion d'une frontière entre deux domaines conformationnels, pour savoir si leur partitionnement repose sur des contraintes internes à ces domaines ou sur un effet d'isolation topologique sous-tendu par des éléments génomiques à leurs bordures. De plus grands remaniements tels des inversions chromosomiques ou même des translocations réciproques sont également envisageables avec un tel système. Il deviendrait alors intéressant de tester à quel point un élément génomique donné continue à contacter le même jeu de partenaires chromosomiques. Des études récentes chez la drosophile ont suggéré que de tels remaniements influencent fortement la topologie chromosomique (Tolhuis et al., 2011). Qu'en serait-il pour des loci dont la capacité à établir des interactions chromosomiques est conservée en contexte ectopique, tel le *Xic* (Bacher et al., 2006 ; Xu et al., 2006 ; Augui et al., 2007)?

## Conclusion

En somme, il existe de nombreux aspects structuraux de la chromatine opérant au niveau régional et qui pourraient sous-tendre les tendances à la régulation transcriptionnelle collective. Ce type de contrôle n'est bien sûr pas incompatible avec une régulation individuelle des différents membres de tels groupes, les comportements collectifs relevant souvent plus de la tendance que des particularités ponctuelles. La caractérisation de l'organisation topologique du centre d'inactivation du chromosome X, qui a été le but principal de ma thèse, a permis de mettre en évidence de nouveaux régulateurs potentiels de l'expression de *Xist* et *Tsix*, et de fournir une base structurale au paysage régulateur au locus *Xic*. Il apparaît probable que l'étendue des régions, respectivement capables de contacter physiquement *Xist* ou *Tsix*, reflète l'existence de multiples éléments de contrôle distaux. Une question qui m'intrigue particulièrement maintenant est de savoir si ces éléments contactent et agissent de concert sur leur cible, ou si chaque interaction possible reflète des états structuraux – et régulateurs – différents. Cet aspect m'apparaît d'autant plus important

que le *Xic* gère l'activation mono-allélique de *Xist* ; le fait l'activation transcriptionnelle n'est pas concomitante entre les deux allèles signifie qu'ils doivent différer structuralement à un moment donné. Si la nature de cette différence structurale reste évasive, il semble intéressant de penser que l'architecture chromosomique locale, et sa variabilité inter-allélique, puisse jouer un rôle direct. Les études de microscopie à haute résolution que nous avons entamées, permettant la détection combinée de l'activité transcriptionnelle de *Xist* (par RNA FISH) et de la conformation interne du *Xic* (par DNA FISH), sont une façon d'appréhender directement question. Mon sentiment est que la variabilité inter-cellulaire et inter-allélique de l'organisation de la chromatine au sein du noyau pourrait être à même d'expliquer la variabilité des patrons transcriptionnels, *via* l'établissement de contacts chromosomiques régulateurs alternatifs. Le rapport récent que la *trans*-induction mosaïque du locus  *$\beta$ -globin* murin par un élément de contrôle transgénique est, au moins en partie, soutenue par l'établissement ponctuel d'associations chromosomiques supporte cette idée (Noordermeer et al., 2011). Comprendre enfin comment un tel phénomène s'intègre à la dynamique de diffusion de la fibre chromatinienne en quatre dimensions représente un défi technique qui ne manquera pas de nous occuper dans les prochaines années.





# ANNEXES



# Revue 3

## *Dynamique nucléaire et inactivation du chromosome X*

Nora & Augui 2009

Cet article de vulgarisation fait partie d'un numéro spécial de Biofutur sur l'inactivation du chromosome X édité par Edith Heard et explique en termes accessibles des concepts fondamentaux concernant l'architecture des chromosomes et sa dynamique pendant la répression transcriptionnelle du chromosome X.



# Revue 4

*Etablissement de l'inactivation transcriptionnelle du chromosome X pendant le développement embryonnaire*

Nora & Heard 2009

Nous avons eu l'occasion d'écrire cet article destiné à un public scientifique à la suite de ma participation à une séance spéciale de la Société Française de Biologie consacrée à l'organisation nucléaire, organisée par Eve Devinoy.



# Revue 5

*X chromosome inactivation: when dosage counts*

Nora & Heard 2009

Commentaire accompagnant la publication de l'article de Jonkers et al., 2009, rapportant la capacité de la surdose de la protéine RNF12 à activer l'expression de Xist.





# Revue 6

## ***RNF12 – a jack of all trades in X inactivation?***

Schulz, Nora and Heard, 2011

Commentaire accompagnant la publication de l'article de Barakat et al., 2011, et Shin et al., 2010 rapportant différentes conséquences de l'inactivation homozygote de *Rnf12*. En anglais, le proverbe « *A jack of all trade, master of none* » désigne quelqu'un qui se distingue par la diversité de ses compétences mais n'excelle dans aucune. Il est souvent terminé par la mention « *though oftentimes better than master of one* », ce qui résume bien les raisons de l'intérêt général pour la protéine RNF12...



# **BIBLIOGRAPHIE**



- Ahmed, S., Brickner, D. G., Light, W. H., Cajigas, I., McDonough, M., Froysheter, A. B., Volpe, T., et Brickner, J. H. (2010). DNA zip codes control an ancient mechanism for gene targeting to the nuclear periphery. *Nat. Cell Biol* *12*, 111-118.
- Ainscough, J. F., Koide, T., Tada, M., Barton, S., et Surani, M. A. (1997). Imprinting of *Igf2* and *H19* from a 130 kb YAC transgene. *Development* *124*, 3621-3632.
- Akhtar, A., et Gasser, S. M. (2007). The nuclear envelope and transcriptional control. *Nat Rev Genet* *8*, 507-517.
- Amouyal, M. (2010). Gene insulation. Part II: natural strategies in vertebrates. *Biochemistry and Cell Biology* *88*, 885-898.
- Apostolou, E., et Thanos, D. (2008). Virus Infection Induces NF-kappaB-dependent interchromosomal associations mediating monoallelic IFN-beta gene expression. *Cell* *134*, 85-96.
- Arney, K. L. (2003). *H19* and *Igf2*-enhancing the confusion? *Trends in Genetics* *19*, 17-23.
- Augui, S., Filion, G. J., Huart, S., Nora, E., Guggiari, M., Maresca, M., Stewart, A. F., et Heard, E. (2007). Sensing X Chromosome Pairs Before X Inactivation via a Novel X-Pairing Region of the *Xic*. *Science* *318*, 1632-1636.
- Augui, S., Nora, E. P., et Heard, E. (2011). Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat. Rev. Genet* *12*, 429-442.
- Bacher, C. P., Guggiari, M., Brors, B., Augui, S., Clerc, P., Avner, P., Eils, R., et Heard, E. (2006). Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. *Nat Cell Biol* *8*, 293-299.
- Banerji, J., Olson, L., et Schaffner, W. (1983). A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* *33*, 729-740.
- Bantignies, F., Roure, V., Comet, I., Leblanc, B., Schuettengruber, B., Bonnet, J., Tixier, V., Mas, A., et Cavalli, G. (2011). Polycomb-Dependent Regulatory Contacts between Distant Hox Loci in *Drosophila*. *Cell* *144*, 214-226.
- Barakat, T. S., Gunhanlar, N., Gontan Pardo, C., Achame, E. M., Ghazvini, M., Boers, R., Kenter, A., Rentmeester, E., Grootegoed, J. A., et Gribnau, J. (2011). RNF12 Activates Xist and Is Essential for X Chromosome Inactivation. *PLoS Genet* *7*, e1002001.
- Baù, D., et Marti-Renom, M. A. (2010). Structure determination of genomic domains by satisfaction of spatial restraints. *Chromosome Res* *19*, 25-35.
- Baù, D., Sanyal, A., Lajoie, B. R., Capriotti, E., Byron, M., Lawrence, J. B., Dekker, J., et Marti-Renom, M. A. (2010). The three-dimensional folding of the  $\alpha$ -globin gene domain reveals formation of chromatin globules. *Nat Struct Mol Biol* *18*, 107-114.
- Bell, A. C., et Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* *405*, 482-485.
- Bell, A. C., West, A. G., et Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* *98*, 387-396.
- Belmont, A. S. (2010). Estrogen fueled, nuclear kiss: Did it move for you? *Nucleus* *1*, 440-443.
- Benayoun, B. A., Caburet, S., et Veitia, R. A. (2011). Forkhead transcription factors: key players in health and disease. *Trends Genet*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21507500> [Consulté mai 5, 2011].
- Benko, S., Fantes, J. A., Amiel, J., Kleinjan, D.-J., Thomas, S., Ramsay, J., Jamshidi, N., Essafi, A., Heaney, S., Gordon, C. T., et al. (2009). Highly conserved non-coding elements on either side of *SOX9* associated with Pierre Robin sequence. *Nat Genet* *41*, 359-364.
- Berteaux, N., Aptel, N., Cathala, G., Genton, C., Coll, J., Daccache, A., Spruyt, N., Hondermarck, H., Dugimont, T., Cury, J.-J., et al. (2008). A novel *H19* antisense RNA overexpressed in breast

- cancer contributes to paternal IGF2 expression. *Mol. Cell. Biol* 28, 6731-6745.
- Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigó, R., Gingeras, T. R., Margulies, E. H., Weng, Z., Snyder, M., Dermitzakis, E. T., Stamatoyannopoulos, J. A., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447, 799-816.
- Black, J. C., Choi, J. E., Lombardo, S. R., et Carey, M. (2006). A mechanism for coordinating chromatin modification and preinitiation complex assembly. *Mol. Cell* 23, 809-818.
- Blow, M. J., McCulley, D. J., Li, Z., Zhang, T., Akiyama, J. A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2010). ChIP-Seq identification of weakly conserved heart enhancers. *Nat Genet* 42, 806-810.
- Bolte, S., et Cordelières, F. P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. *J Microsc* 224, 213-232.
- Boutanaev, A. M., Kalmykova, A. I., Shevelyov, Y. Y., et Nurminsky, D. I. (2002). Large clusters of co-expressed genes in the Drosophila genome. *Nature* 420, 666-669.
- Boyle, S., Gilchrist, S., Bridger, J. M., Mahy, N. L., Ellis, J. A., et Bickmore, W. A. (2001). The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum. Mol. Genet* 10, 211-219.
- Brown, J. M., Green, J., das Neves, R. P., Wallace, H. A. C., Smith, A. J. H., Hughes, J., Gray, N., Taylor, S., Wood, W. G., Higgs, D. R., et al. (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *The Journal of Cell Biology* 182, 1083.
- Brown, J. M., Leach, J., Reittie, J. E., Atzberger, A., Lee-Prudhoe, J., Wood, W. G., Higgs, D. R., Iborra, F. J., et Buckle, V. J. (2006). Coregulated human globin genes are frequently in spatial proximity when active. *The Journal of cell biology* 172, 177.
- Bulger, M., et Groudine, M. (2011). Functional and Mechanistic Diversity of Distal Transcription Enhancers. *Cell* 144, 327-339.
- Bulger, M., Schübeler, D., Bender, M. A., Hamilton, J., Farrell, C. M., Hardison, R. C., et Groudine, M. (2003). A complex chromatin landscape revealed by patterns of nuclease sensitivity and histone modification within the mouse beta-globin locus. *Mol. Cell. Biol* 23, 5234-5244.
- Capelson, M., et Corces, V. G. (2004). Boundary elements and nuclear organization. *Biology of the Cell* 96, 617-629.
- Celniker, S. E., Dillon, L. A. L., Gerstein, M. B., Gunsalus, K. C., Henikoff, S., Karpen, G. H., Kellis, M., Lai, E. C., Lieb, J. D., MacAlpine, D. M., et al. (2009). Unlocking the secrets of the genome. *Nature* 459, 927-930.
- Chambeyron, S., et Bickmore, W. A. (2004). Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev* 18, 1119-1130.
- Chambeyron, S., Da Silva, N. R., Lawson, K. A., et Bickmore, W. A. (2005). Nuclear re-organisation of the Hoxb complex during mouse embryonic development. *Development* 132, 2215-2223.
- Chan, P. K., Wai, A., Philipsen, S., et Tan-Un, K.-C. (2008). 5'HS5 of the human beta-globin locus control region is dispensable for the formation of the beta-globin active chromatin hub. *PLoS ONE* 3, e2134.
- Chaumeil, J., Augui, S., Chow, J. C., et Heard, E. (2008). Combined immunofluorescence, RNA fluorescent in situ hybridization, and DNA fluorescent in situ hybridization to study chromatin changes, transcriptional activity, nuclear organization, and X-chromosome inactivation. *Methods Mol. Biol* 463, 297-308.
- Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V. B., Wong, E., Orlov, Y. L., Zhang, W., Jiang, J., et al. (2008). Integration of External Signaling Pathways with the Core

- Transcriptional Network in Embryonic Stem Cells. *Cell* 133, 1106-1117.
- Cheutin, T., Gorski, S. A., May, K. M., Singh, P. B., et Misteli, T. (2004). In vivo dynamics of Swi6 in yeast: evidence for a stochastic model of heterochromatin. *Mol. Cell Biol* 24, 3157-3167.
- Cheutin, T., McNairn, A. J., Jenuwein, T., Gilbert, D. M., Singh, P. B., et Misteli, T. (2003). Maintenance of Stable Heterochromatin Domains by Dynamic HP1 Binding. *Science* 299, 721-725.
- Chung, J. H., Whiteley, M., et Felsenfeld, G. (1993). A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* 74, 505-514.
- Collins, A., Hewitt, S. L., Chaumeil, J., Sellars, M., Micsinai, M., Allinne, J., Parisi, F., Nora, E. P., Bolland, D. J., Corcoran, A. E., et al. (2011). RUNX Transcription Factor-Mediated Association of Cd4 and Cd8 Enables Coordinate Gene Regulation. *Immunity* 34, 303-314.
- Conti, L., Pollard, S. M., Gorba, T., Reitano, E., Toselli, M., Biella, G., Sun, Y., Sanzone, S., Ying, Q.-L., Cattaneo, E., et al. (2005). Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. *PLoS Biol* 3, e283.
- Core, L. J., et Lis, J. T. (2009). Paused Pol II captures enhancer activity and acts as a potent insulator. *Genes Dev* 23, 1606-1612.
- Court, F., Baniol, M., Hagege, H., Petit, J. S., Lelay-Taha, M.-N., Carbonell, F., Weber, M., Cathala, G., et Forne, T. (2011a). Long-range chromatin interactions at the mouse *Igf2/H19* locus reveal a novel paternally expressed long non-coding RNA. *Nucleic Acids Res*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21478171> [Consulté mai 8, 2011].
- Court, F., Miro, J., Braem, C., Lelay-Taha, M.-N., Brisebarre, A., Atger, F., Gostan, T., Weber, M., Cathala, G., et Forne, T. (2011b). Modulated contact frequencies at gene-rich loci support a statistical helix model for mammalian chromatin organization. *Genome Biol* 12, R42.
- Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences* 107, 21931-21936.
- Csink, A. K., et Henikoff, S. (1996). Genetic modification of heterochromatic association and nuclear organization in *Drosophila*. *Nature* 381, 529-531.
- Cuddapah, S., Jothi, R., Schones, D. E., Roh, T.-Y., Cui, K., et Zhao, K. (2009). Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res* 19, 24-32.
- Cullen, K. E., Ron, M. P., et Seyfred, M. A. (1993). Interaction Between Transcription Regulatory Regions of Prolactin Chromatin. *Science* 261, 203.
- Dekker, J., et Klecker, N. (2002). Capturing Chromosome Conformation. *Science* 295, 1306-1311.
- Ding, N., Zhou, H., Esteve, P.-O., Chin, H. G., Kim, S., Xu, X., Joseph, S. M., Friez, M. J., Schwartz, C. E., Pradhan, S., et al. (2008). Mediator links epigenetic silencing of neuronal gene expression with x-linked mental retardation. *Mol. Cell* 31, 347-359.
- Dolle, P., Izpisua-Belmonte, J.-C., Falkenstein, H., Renucci, A., et Duboule, D. (1989). Coordinate expression of the murine *Hox-5* complex homoeobox-containing genes during limb pattern formation. *Nature* 342, 767-772.
- Dostie, J., Richmond, T. A., Arnaout, R. A., Selzer, R. R., Lee, W. L., Honan, T. A., Rubio, E. D., Krumm, A., Lamb, J., Nusbaum, C., et al. (2006). Chromosome Conformation Capture Carbon Copy (5C): A massively parallel solution for mapping interactions between genomic elements. *Genome Research* 16, 1299-1309.
- Dostie, J., et Dekker, J. (2007). Mapping networks of physical interactions between genomic elements



- using 5C technology. *Nat Protoc* 2, 988-1002.
- Dunn, T. M., Hahn, S., Ogden, S., et Schleif, R. F. (1984). An operator at-280 base pairs that is required for repression of araBAD operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. *Proceedings of the National Academy of Sciences of the United States of America* 81, 5017.
- Eaton, M. L., Prinz, J. A., MacAlpine, H. K., Tretyakov, G., Kharchenko, P. V., et MacAlpine, D. M. (2011). Chromatin signatures of the *Drosophila* replication program. *Genome Res* 21, 164-174.
- Eissenberg, J. C., James, T. C., Foster-Hartnett, D. M., Hartnett, T., Ngan, V., et Elgin, S. C. (1990). Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A* 87, 9923-9927.
- Elgar, G., et Vavouri, T. (2008). Tuning in to the signals: noncoding sequence conservation in vertebrate genomes. *Trends Genet* 24, 344-352.
- Elsner, M., et Mak, H. C. (2011). A modENCODE snapshot. *Nature Biotechnology* 29, 238-240.
- Emorine, L., Kuehl, M., Weir, L., Leder, P., et Max, E. E. (1983). A conserved sequence in the immunoglobulin J kappa-C kappa intron: possible enhancer element. *Nature* 304, 447-449.
- Engel, N., Raval, A. K., Thorvaldsen, J. L., et Bartolomei, S. M. (2008). Three-dimensional conformation at the H19/Igf2 locus supports a model of enhancer tracking. *Human Molecular Genetics* 17, 3021 -3029.
- van den Engh, G., Sachs, R., et Trask, B. J. (1992). Estimating genomic distance from DNA sequence location in cell nuclei by a random walk model. *Science* 257, 1410.
- Ernst, J., Kheradpour, P., Mikkelsen, T. S., Shores, N., Ward, L. D., Epstein, C. B., Zhang, X., Wang, L., Issner, R., Coyne, M., et al. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 473, 43-49.
- Ernst, J., et Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat Biotechnol* 28, 817-825.
- Eskeland, R., Leeb, M., Grimes, G. R., Kress, C., Boyle, S., Sproul, D., Gilbert, N., Fan, Y., Skoultschi, A. I., et Wutz, A. (2010). Ring1B Compacts Chromatin Structure and Represses Gene Expression Independent of Histone Ubiquitination. *Molecular Cell* 38, 452-464.
- Farrell, C. M., West, A. G., et Felsenfeld, G. (2002). Conserved CTCF insulator elements flank the mouse and human beta-globin loci. *Mol. Cell. Biol* 22, 3820-3831.
- Feng, R., Desbordes, S. C., Xie, H., Tillo, E. S., Pixley, F., Stanley, E. R., et Graf, T. (2008). PU. 1 and C/EBP alpha/beta convert fibroblasts into macrophage-like cells. *Proceedings of the National Academy of Sciences* 105, 6057.
- Ferraiuolo, M. A., Rousseau, M., Miyamoto, C., Shenker, S., Wang, X. Q. D., Nadler, M., Blanchette, M., et Dostie, J. (2010). The three-dimensional architecture of Hox cluster silencing. *Nucleic Acids Research* 38, 7472.
- Fiering, S., Whitelaw, E., et Martin, D. I. K. (2000). To be or not to be active: the stochastic nature of enhancer action. *Bioessays* 22, 381-387.
- Filion, G. J., van Bommel, J. G., Braunschweig, U., Talhout, W., Kind, J., Ward, L. D., Brugman, W., de Castro, I. J., Kerkhoven, R. M., Bussemaker, H. J., et al. (2010). Systematic Protein Location Mapping Reveals Five Principal Chromatin Types in *Drosophila* Cells. *Cell* 143, 212-224.
- Finlan, L. E., Sproul, D., Thomson, I., Boyle, S., Kerr, E., Perry, P., Ylstra, B., Chubb, J. R., et Bickmore, W. A. (2008). Recruitment to the Nuclear Periphery Can Alter Expression of Genes in Human Cells. *PLoS Genetics* 4, e1000039.

- Fox Keller, E., et Harel, D. (2007). Beyond the Gene. *PLoS ONE* 2, e1231.
- Frankel, N., Davis, G. K., Vargas, D., Wang, S., Payre, F., et Stern, D. L. (2010). Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* 466, 490-493.
- Fraser, J., Rousseau, M., Shenker, S., Ferraiuolo, M. A., Hayashizaki, Y., Blanchette, M., et Dostie, J. (2009). Chromatin conformation signatures of cellular differentiation. *Genome biology* 10, R37.
- Fuda, N. J., Ardehali, M. B., et Lis, J. T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* 461, 186–192.
- Fukao, T., Fukuda, Y., Kiga, K., Sharif, J., Hino, K., Enomoto, Y., Kawamura, A., Nakamura, K., Takeuchi, T., et Tanabe, M. (2007). An Evolutionarily Conserved Mechanism for MicroRNA-223 Expression Revealed by MicroRNA Gene Profiling. *Cell* 129, 617-631.
- Fullwood, M. J., Liu, M. H., Pan, Y. F., Liu, J., Xu, H., Mohamed, Y. B., Orlov, Y. L., Velkov, S., Ho, A., Mei, P. H., et al. (2009). An oestrogen-receptor- $\alpha$ -bound human chromatin interactome. *Nature* 462, 58-64.
- Fussner, E., Ching, R. W., et Bazett-Jones, D. P. (2011). Living without 30nm chromatin fibers. *Trends Biochem. Sci* 36, 1-6.
- Gaszner, M., et Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. *Nature Reviews Genetics* 7, 703–713.
- Gause, M., Schaaf, C. A., et Dorsett, D. (2008). Cohesin and CTCF: cooperating to control chromosome conformation? *Bioessays* 30, 715-718.
- Gerstein, M. B., Bruce, C., Rozowsky, J. S., Zheng, D., Du, J., Korbil, J. O., Emanuelsson, O., Zhang, Z. D., Weissman, S., et Snyder, M. (2007). What is a gene, post-ENCODE? History and updated definition. *Genome Research* 17, 669-681.
- Gerstein, M. B., Lu, Z. J., Van Nostrand, E. L., Cheng, C., Arshinoff, B. I., Liu, T., Yip, K. Y., Robilotto, R., Rechtsteiner, A., Ikegami, K., et al. (2010). Integrative Analysis of the *Caenorhabditis elegans* Genome by the modENCODE Project. *Science* 330, 1775-1787.
- Gheldof, N., Tabuchi, T. M., et Dekker, J. (2006). The active FMR1 promoter is associated with a large domain of altered chromatin conformation with embedded local histone modifications. *Proceedings of the National Academy of Sciences* 103, 12463.
- Ghisletti, S., Barozzi, I., Mietton, F., Polletti, S., De Santa, F., Venturini, E., Gregory, L., Lonie, L., Chew, A., Wei, C.-L., et al. (2010). Identification and Characterization of Enhancers Controlling the Inflammatory Gene Expression Program in Macrophages. *Immunity* 32, 317-328.
- Gierman, H. J., Indemans, M. H. G., Koster, J., Goetze, S., Seppen, J., Geerts, D., van Driel, R., et Versteeg, R. (2007). Domain-wide regulation of gene expression in the human genome. *Genome Research* 17, 1286-1295.
- Gilbert, N., Boyle, S., Fiegler, H., Woodfine, K., Carter, N. P., et Bickmore, W. A. (2004). Chromatin Architecture of the Human Genome:: Gene-Rich Domains Are Enriched in Open Chromatin Fibers. *Cell* 118, 555–566.
- Gilmour, D. S., et Lis, J. T. (1984). Detecting protein-DNA interactions in vivo: distribution of RNA polymerase on specific bacterial genes. *Proc Natl Acad Sci U S A* 81, 4275-4279.
- Giraldo, P., et Montoliu, L. (2001). Size matters: use of YACs, BACs and PACs in transgenic animals. *Transgenic research* 10, 83–103.
- Girton, J. R., et Johansen, K. M. (2008). Chromatin structure and the regulation of gene expression: the lessons of PEV in *Drosophila*. *Advances in genetics* 61, 1–43.
- Gomes, N. P., et Espinosa, J. M. (2010). Gene-specific repression of the p53 target gene PUMA via

- intragenic CTCF-Cohesin binding. *Genes Dev* 24, 1022-1034.
- Gossen, M., et Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U.S.A* 89, 5547-5551.
- Gould, A., Morrison, A., Sproat, G., White, R. A., et Krumlauf, R. (1997). Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping Hox expression patterns. *Genes Dev* 11, 900-913.
- Gribnau, J., De Boer, E., Trimborn, T., Wijgerde, M., Milot, E., Grosveld, F., et Fraser, P. (1998). Chromatin interaction mechanism of transcriptional control in vivo. *The EMBO Journal* 17, 6020-6027.
- Griffith, J., Hochschild, A., et Ptashne, M. (1986). DNA loops induced by cooperative binding of lambda repressor. *Nature* 322, 750-752.
- Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M. B., Talhout, W., Eussen, B. H., de Klein, A., Wessels, L., de Laat, W., et al. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453, 948-951.
- Hagège, H., Nasser, R., Weber, M., Milligan, L., Aptel, N., Jacquet, C., Drewell, R. A., Dandolo, L., Surani, M. A., Cathala, G., et al. (2006). The 3' portion of the mouse H19 Imprinting-Control Region is required for proper tissue-specific expression of the Igf2 gene. *Cytogenet. Genome Res* 113, 230-237.
- Hakim, O., Sung, M.-H., Voss, T. C., Splinter, E., John, S., Sabo, P. J., Thurman, R. E., Stamatoyannopoulos, J. A., de Laat, W., et Hager, G. L. (2011). Diverse gene reprogramming events occur in the same spatial clusters of distal regulatory elements. *Genome Research* 21, 697-706.
- Harewood, L., Schütz, F., Boyle, S., Perry, P., Delorenzi, M., Bickmore, W. A., et Reymond, A. (2010). The effect of translocation-induced nuclear reorganization on gene expression. *Genome Res* 20, 554-564.
- Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M., et Tilghman, S. M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 405, 486-489.
- Harmon, B., et Sedat, J. (2005). Cell-by-Cell Dissection of Gene Expression and Chromosomal Interactions Reveals Consequences of Nuclear Reorganization. *PLoS Biol* 3, e67.
- Heard, E., Kress, C., Mongelard, F., Courtier, B., Rougeulle, C., Ashworth, A., Vourc'h, C., Babinet, C., et Avner, P. (1996). Transgenic mice carrying an Xist-containing YAC. *Human molecular genetics* 5, 441.
- Heard, E., Mongelard, F., Arnaud, D., et Avner, P. (1999). Xist yeast artificial chromosome transgenes function as X-inactivation centers only in multicopy arrays and not as single copies. *Molecular and cellular biology* 19, 3156.
- Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., Hawkins, R. D., Barrera, L. O., Van Calcar, S., Qu, C., Ching, K. A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* 39, 311-318.
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., et Glass, C. K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Molecular Cell* 38, 576-589.
- Henikoff, S. (1996). Dosage-dependent modification of position-effect variegation in *Drosophila*. *BioEssays* 18, 401-409.
- Herendeen, D., Kassavetis, G., et Geiduschek, E. (1992). A transcriptional enhancer whose function imposes a requirement that proteins track along DNA. *Science* 256, 1298-1303.

- Hewitt, S. L., Yin, B., Ji, Y., Chaumeil, J., Marszalek, K., Tenthorey, J., Salvagiotto, G., Steinel, N., Ramsey, L. B., Ghysdael, J., et al. (2009). RAG-1 and ATM coordinate monoallelic recombination and nuclear positioning of immunoglobulin loci. *Nat Immunol* *10*, 655-664.
- Higgs, D. R., Vernimmen, D., et Wood, B. (2008). Chapter 5 Long-Range Regulation of alpha-Globin Gene Expression. Dans *Long-Range Control of Gene Expression* Academic Press, p. 143-173. Available at: <http://www.sciencedirect.com/science/article/B7CT7-4RVDGXJ-7/2/94b1068d9fbb959fa11bfb12900993f9> [Consulté mai 21, 2011].
- Hiratani, I., Ryba, T., Itoh, M., Rathjen, J., Kulik, M., Papp, B., Fussner, E., Bazett-Jones, D. P., Plath, K., Dalton, S., et al. (2010). Genome-wide dynamics of replication timing revealed by in vitro models of mouse embryogenesis. *Genome Res* *20*, 155-169.
- Hiratani, I., Ryba, T., Itoh, M., Yokochi, T., Schwaiger, M., Chang, C.-W., Lyou, Y., Townes, T. M., Schübeler, D., et Gilbert, D. M. (2008). Global Reorganization of Replication Domains During Embryonic Stem Cell Differentiation. *Plos Biol* *6*, e245.
- Hochschild, A., et Ptashne, M. (1986). Cooperative binding of lambda repressors to sites separated by integral turns of the DNA helix. *Cell* *44*, 681-687.
- Hochschild, A., et Ptashne, M. (1988). Interaction at a distance between [lambda]repressors disrupts gene activation. *Nature* *336*, 353-357.
- Hooper, M., Hardy, K., Handyside, A., Hunter, S., et Monk, M. (1987). HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature* *326*, 292-295.
- Hu, Q., Kwon, Y.-S., Nunez, E., Cardamone, M. D., Hutt, K. R., Ohgi, K. A., Garcia-Bassets, I., Rose, D. W., Glass, C. K., Rosenfeld, M. G., et al. (2008). Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules. *Proceedings of the National Academy of Sciences* *105*, 19199 -19204.
- Huang, P., He, Z., Ji, S., Sun, H., Xiang, D., Liu, C., Hu, Y., Wang, X., et Hui, L. (2011). Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature advance online publication*. Available at: <http://dx.doi.org/10.1038/nature10116> [Consulté mai 11, 2011].
- Irani, M. H., Orosz, L., et Adhya, S. (1983). A control element within a structural gene: The gal operon of Escherichia coli. *Cell* *32*, 783-788.
- Jackson, D. A., Hassan, A. B., Errington, R. J., et Cook, P. R. (1993). Visualization of focal sites of transcription within human nuclei. *EMBO J* *12*, 1059-1065.
- Jhunjunwala, S., van Zelm, M. C., Peak, M. M., Cutchin, S., Riblet, R., van Dongen, J. J. M., Grosveld, F. G., Knoch, T. A., et Murre, C. (2008). The 3D Structure of the Immunoglobulin Heavy-Chain Locus: Implications for Long-Range Genomic Interactions. *Cell* *133*, 265-279.
- John, S., Sabo, P. J., Thurman, R. E., Sung, M.-H., Biddie, S. C., Johnson, T. A., Hager, G. L., et Stamatoyannopoulos, J. A. (2011). Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nat Genet* *43*, 264-268.
- Jonkers, I., Barakat, T. S., Achame, E. M., Monkhorst, K., Kenter, A., Rentmeester, E., Grosveld, F., Grootegoed, J. A., et Gribnau, J. (2009). RNF12 Is an X-Encoded Dose-Dependent Activator of X Chromosome Inactivation. *Cell* *139*, 999-1011.
- Kaffer, C. R., Srivastava, M., Park, K. Y., Ives, E., Hsieh, S., Batlle, J., Grinberg, A., Huang, S. P., et Pfeifer, K. (2000). A transcriptional insulator at the imprinted H19/Igf2 locus. *Genes Dev* *14*, 1908-1919.
- Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., van Berkum, N. L., Ebmeier, C. C., Goossens, J., Rahl, P. B., Levine, S. S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature* *467*, 430-435.

- Kanduri, C., Holmgren, C., Pilartz, M., Franklin, G., Kanduri, M., Liu, L., Ginjala, V., Ullerås, E., Mattsson, R., et Ohlsson, R. (2000). The 5' flank of mouse H19 in an unusual chromatin conformation unidirectionally blocks enhancer-promoter communication. *Curr. Biol* 10, 449-457.
- Kellum, R., et Schedl, P. (1991). A position-effect assay for boundaries of higher order chromosomal domains. *Cell* 64, 941-950.
- Kellum, R., et Schedl, P. (1992). A group of scs elements function as domain boundaries in an enhancer-blocking assay. *Molecular and cellular biology* 12, 2424.
- Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., et Haussler, and D. (2002). The Human Genome Browser at UCSC. *Genome Research* 12, 996 -1006.
- Kim, T. H., Abdullaev, Z. K., Smith, A. D., Ching, K. A., Loukinov, D. I., Green, R. D., Zhang, M. Q., Lobanenko, V. V., et Ren, B. (2007). Analysis of the Vertebrate Insulator Protein CTCF-Binding Sites in the Human Genome. *Cell* 128, 1231-1245.
- Kim, T.-K., Hemberg, M., Gray, J. M., Costa, A. M., Bear, D. M., Wu, J., Harmin, D. A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182-187.
- Kimura, A., Umehara, T., et Horikoshi, M. (2002). Chromosomal gradient of histone acetylation established by Sas2p and Sir2p functions as a shield against gene silencing. *Nat Genet* 32, 370-377.
- Kleinjan, D. A., Seawright, A., Schedl, A., Quinlan, R. A., Danes, S., et van Heyningen, V. (2001). Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6. *Hum. Mol. Genet* 10, 2049-2059.
- Kleinjan, D. A., et Lettice, L. A. (2008). Long-range gene control and genetic disease. *Advances in Genetics* 61, 339-388.
- Kmita, M., Kondo, T., et Duboule, D. (2000). Targeted inversion of a polar silencer within the HoxD complex re-allocates domains of enhancer sharing. *Nature genetics* 26, 451-454.
- Kmita, M., Fraudeau, N., Herault, Y., et Duboule, D. (2002). Serial deletions and duplications suggest a mechanism for the collinearity of Hoxd genes in limbs. *Nature* 420, 145-150.
- Ko, M. S., Nakauchi, H., et Takahashi, N. (1990). The dose dependence of glucocorticoid-inducible gene expression results from changes in the number of transcriptionally active templates. *EMBO J* 9, 2835-2842.
- Kocanova, S., Kerr, E. A., Rafique, S., Boyle, S., Katz, E., Caze-Subra, S., Bickmore, W. A., et Bystrycky, K. (2010). Activation of estrogen-responsive genes does not require their nuclear co-localization. *PLoS Genet* 6, e1000922.
- Kosak, S. T., Skok, J. A., Medina, K. L., Riblet, R., Le Beau, M. M., Fisher, A. G., et Singh, H. (2002). Subnuclear Compartmentalization of Immunoglobulin Loci During Lymphocyte Development. *Science* 296, 158 -162.
- Ktistaki, E., Garefalaki, A., Williams, A., Andrews, S. R., Bell, D. M., Foster, K. E., Spilianakis, C. G., Flavell, R. A., Kosyakova, N., Trifonov, V., et al. (2010). CD8 locus nuclear dynamics during thymocyte development. *J. Immunol* 184, 5686-5695.
- Kumaran, R. (2008). Chromatin Dynamics and Gene Positioning. *Cell* 132, 929-934.
- Kurukuti, S., Tiwari, V. K., Tavoosidana, G., Pugacheva, E., Murrell, A., Zhao, Z., Lobanenko, V., Reik, W., et Ohlsson, R. (2006). CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proceedings of the National Academy of Sciences* 103, 10684.
- Lajoie, B. R., van Berkum, N. L., Sanyal, A., et Dekker, J. (2009). My5C: web tools for chromosome conformation capture studies. *Nat Meth* 6, 690-691.

- Lee, J. M., et Sonnhammer, E. L. L. (2003). Genomic Gene Clustering Analysis of Pathways in Eukaryotes. *Genome Research* 13, 875-882.
- Lercher, M. J., Urrutia, A. O., et Hurst, L. D. (2002). Clustering of housekeeping genes provides a unified model of gene order in the human genome. *Nat. Genet.* 31, 180-183.
- Lettice, L. A., Horikoshi, T., Heaney, S. J. H., van Baren, M. J., van der Linde, H. C., Breedveld, G. J., Joosse, M., Akarsu, N., Oostra, B. A., Endo, N., et al. (2002). Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly. *Proc. Natl. Acad. Sci. U.S.A* 99, 7548-7553.
- Li, Q., et Stamatoyannopoulos, G. (1994). Hypersensitive site 5 of the human beta locus control region functions as a chromatin insulator. *Blood* 84, 1399-1401.
- Liber, D., Domasch, R., Holmqvist, P.-H., Mazzarella, L., Georgiou, A., Leleu, M., Fisher, A. G., Labosky, P. A., et Dillon, N. (2010). Epigenetic Priming of a Pre-B Cell-Specific Enhancer through Binding of Sox2 and Foxd3 at the ESC Stage. *Cell Stem Cell* 7, 114-126.
- Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., et al. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* 326, 289-293.
- Lienert, F., Mohn, F., Tiwari, V. K., Baubec, T., Roloff, T. C., Gaidatzis, D., Stadler, M. B., et Schübeler, D. (2011). Genomic Prevalence of Heterochromatic H3K9me2 and Transcription Do Not Discriminate Pluripotent from Terminally Differentiated Cells. *PLoS Genet* 7, e1002090.
- Locke, J., Kotarski, M. A., et Tartof, K. D. (1988). Dosage-dependent modifiers of position effect variegation in *Drosophila* and a mass action model that explains their effect. *Genetics* 120, 181-198.
- Lomvardas, S., Barnea, G., Pisapia, D. J., Mendelsohn, M., Kirkland, J., et Axel, R. (2006). Interchromosomal Interactions and Olfactory Receptor Choice. *Cell* 126, 403-413.
- Lopes, S., Lewis, A., Hajkova, P., Dean, W., Oswald, J., Forné, T., Murrell, A., Constância, M., Bartolomei, M., Walter, J., et al. (2003). Epigenetic modifications in an imprinting cluster are controlled by a hierarchy of DMRs suggesting long-range chromatin interactions. *Hum. Mol. Genet* 12, 295-305.
- Lunyak, V. V., Prefontaine, G. G., Nunez, E., Cramer, T., Ju, B.-G., Ohgi, K. A., Hutt, K., Roy, R., Garcia-Diaz, A., Zhu, X., et al. (2007). Developmentally Regulated Activation of a SINE B2 Repeat as a Domain Boundary in Organogenesis. *Science* 317, 248-251.
- Maconochie, M. K., Nonchev, S., Studer, M., Chan, S. K., Pöpperl, H., Sham, M. H., Mann, R. S., et Krumlauf, R. (1997). Cross-regulation in the mouse HoxB complex: the expression of Hoxb2 in rhombomere 4 is regulated by Hoxb1. *Genes Dev* 11, 1885-1895.
- Malik, S., et Roeder, R. G. (2010). The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nature Reviews Genetics* 11, 761-772.
- Marks, H., Chow, J. C., Denissov, S., Francoijs, K.-J., Brockdorff, N., Heard, E., et Stunnenberg, H. G. (2009). High-resolution analysis of epigenetic changes associated with X inactivation. *Genome Research* 19, 1361-1373.
- Masip, M., Veiga, A., Izpisua Belmonte, J. C., et Simón, C. (2010). Reprogramming with defined factors: from induced pluripotency to induced transdifferentiation. *Mol. Hum. Reprod* 16, 856-868.
- Maston, G. A., Evans, S. K., et Green, M. R. (2006). Transcriptional regulatory elements in the human genome. *Annu. Rev. Genomics Hum. Genet.* 7, 29-59.
- Masui, O., Bonnet, I., Le Baccon, P., Brito, I., Pollex, T., Murphy, N., Hupé, P., Barillot, E., Belmont,

- A. S., et Heard, E. (2011). Live-Cell Chromosome Dynamics and Outcome of X Chromosome Pairing Events during ES Cell Differentiation. *Cell* 145, 447-458.
- Mayer, R., Brero, A., Von Hase, J., Schroeder, T., Cremer, T., et Dietzel, S. (2005). Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. *BMC cell biology* 6, 44.
- Meister, P., Mango, S. E., et Gasser, S. M. (2011). Locking the genome: nuclear organization and cell fate. *Curr. Opin. Genet. Dev* 21, 167-174.
- Michalak, P. (2008). Coexpression, coregulation, and cofunctionality of neighboring genes in eukaryotic genomes. *Genomics* 91, 243–248.
- Migeon, B. R. (2003). Is Tsix repression of Xist specific to mouse? *Nat. Genet* 33, 337; author reply 337-338.
- Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R. P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553-560.
- Morey, C., Da Silva, N. R., Perry, P., et Bickmore, W. A. (2007). Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. *Development* 134, 909-919.
- Morey, C., Da Silva, N. R., Kmita, M., Duboule, D., et Bickmore, W. A. (2008). Ectopic nuclear reorganisation driven by a Hoxb1 transgene transposed into Hoxd. *Journal of Cell Science* 121, 571 -577.
- Morin-Kensicki, E. M., Faust, C., LaMantia, C., et Magnuson, T. (2001). Cell and tissue requirements for the gene *eed* during mouse gastrulation and organogenesis. *Genesis* 31, 142-146.
- Mukherjee, S., Erickson, H., et Bastia, D. (1988). Enhancer-origin interaction in plasmid R6K involves a DNA loop mediated by initiator protein. *Cell* 52, 375-383.
- Muller, H. J. (1930). Types of visible variations induced by X-rays in *Drosophila*. *Journ. of Gen.* 22, 299-334.
- Müller, H.-P., Sogo, J., et Schaffner, W. (1989). An enhancer stimulates transcription in *Trans* when attached to the promoter via a protein bridge. *Cell* 58, 767-777.
- Münkel, C., Eils, R., Dietzel, S., Zink, D., Mehring, C., Wedemann, G., Cremer, T., et Langowski, J. (1999). Compartmentalization of Interphase Chromosomes Observed in Simulation and Experiment. *Journal of Molecular Biology* 285, 1053-1065.
- Murrell, A., Heeson, S., et Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nat Genet* 36, 889-893.
- Nègre, N., Brown, C. D., Ma, L., Bristow, C. A., Miller, S. W., Wagner, U., Kheradpour, P., Eaton, M. L., Loriaux, P., Sealfon, R., et al. (2011). A cis-regulatory map of the *Drosophila* genome. *Nature* 471, 527–531.
- Neuberger, M. S. (1983). Expression and regulation of immunoglobulin heavy chain gene transfected into lymphoid cells. *EMBO J* 2, 1373-1378.
- Noordermeer, D., de Wit, E., Klous, P., van de Werken, H., Simonis, M., Lopez-Jones, M., Eussen, B., de Klein, A., Singer, R. H., et de Laat, W. (2011). Variegated gene expression caused by cell-specific long-range DNA interactions. *Nat Cell Biol.* Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21706023> [Consulté juillet 3, 2011].
- Nora, E. P., et Heard, E. (2011). Chromatin Structure and Nuclear Organization Dynamics during X-Chromosome Inactivation. *Cold Spring Harbor Symposia on Quantitative Biology.* Available at: <http://www.symposium.cshlp.org/cgi/doi/10.1101/sqb.2010.75.032> [Consulté mai 8, 2011].

- Nora, E. P., et Heard, E. (2010). Établissement de l'inactivation transcriptionnelle du chromosome X pendant le développement embryonnaire. *Biologie Aujourd'hui* 204, 6.
- Nora, E. P., et Heard, E. (2009). X Chromosome Inactivation: When Dosage Counts. *Cell* 139, 865-867.
- Nora, E., et Augui, S. (2009). Dynamique nucléaire et inactivation du chromosome X. *Biofutur*.
- Ogawa, Y., et Lee, J. T. (2003). Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. *Molecular cell* 11, 731-743.
- Okamoto, I., Arnaud, D., Le Baccon, P., Otte, A. P., Disteché, C. M., Avner, P., et Heard, E. (2005). Evidence for de novo imprinted X-chromosome inactivation independent of meiotic inactivation in mice. *Nature* 438, 369-373.
- Okamoto, I., Patrat, C., Thépot, D., Peynot, N., Fauque, P., Daniel, N., Diabangouaya, P., Wolf, J.-P., Renard, J.-P., Duranthon, V., et al. (2011). Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. *Nature* 472, 370-374.
- Orlando, V., et Paro, R. (1993). Mapping Polycomb-repressed domains in the bithorax complex using in vivo formaldehyde cross-linked chromatin. *Cell* 75, 1187-1198.
- Osborne, C. S., Chakalova, L., Brown, K. E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J. A., Lopes, S., Reik, W., et al. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat. Genet* 36, 1065-1071.
- Palstra, R.-J., de Laat, W., et Grosveld, F. (2008a). Beta-globin regulation and long-range interactions. *Adv. Genet* 61, 107-142.
- Palstra, R.-J., Simonis, M., Klous, P., Brasset, E., Eijkelkamp, B., et de Laat, W. (2008b). Maintenance of long-range DNA interactions after inhibition of ongoing RNA polymerase II transcription. *PLoS ONE* 3, e1661.
- Palstra, R.-J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., et de Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet* 35, 190-194.
- Parslow, T. G., et Granner, D. K. (1983). Structure of a nuclease-sensitive region inside the immunoglobulin kappa gene: evidence for a role in gene regulation. *Nucleic Acids Res* 11, 4775-4792.
- Pauler, F. M., Sloane, M. A., Huang, R., Regha, K., Koerner, M. V., Tamir, I., Sommer, A., Aszodi, A., Jenuwein, T., et Barlow, D. P. (2009). H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies a histone banding pattern on a mouse autosomal chromosome. *Genome Research* 19, 221-233.
- Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S., et Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. *Nature* 379, 131-137.
- Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S. W. M., Solovei, I., Brugman, W., Gräf, S., Flicek, P., Kerkhoven, R. M., van Lohuizen, M., et al. (2010). Molecular Maps of the Reorganization of Genome-Nuclear Lamina Interactions during Differentiation. *Molecular Cell* 38, 603-613.
- Picard, D., et Schaffner, W. (1984). A lymphocyte-specific enhancer in the mouse immunoglobulin kappa gene. *Nature* 307, 80-82.
- Pickersgill, H., Kalverda, B., de Wit, E., Talhout, W., Fornerod, M., et van Steensel, B. (2006). Characterization of the *Drosophila melanogaster* genome at the nuclear lamina. *Nat. Genet* 38, 1005-1014.
- Pope, B. D., Hiratani, I., et Gilbert, D. M. (2010). Domain-wide regulation of DNA replication timing during mammalian development. *Chromosome Res* 18, 127-136.



- Ptashne, M. (1986). Gene regulation by proteins acting nearby and at a distance. *Nature* 322, 697-701.
- Qiu, X., Vu, T. H., Lu, Q., Ling, J. Q., Li, T., Hou, A., Wang, S. K., Chen, H. L., Hu, J. F., et Hoffman, A. R. (2008). A complex deoxyribonucleic acid looping configuration associated with the silencing of the maternal *Igf2* allele. *Mol. Endocrinol* 22, 1476-1488.
- Rach, E. A., Winter, D. R., Benjamin, A. M., Corcoran, D. L., Ni, T., Zhu, J., et Ohler, U. (2011). Transcription Initiation Patterns Indicate Divergent Strategies for Gene Regulation at the Chromatin Level. *PLoS Genet* 7, e1001274.
- Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S. A., Flynn, R. A., et Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 470, 279-283.
- Rastegar, M., Kobrossy, L., Kovacs, E. N., Rambaldi, I., et Featherstone, M. (2004). Sequential histone modifications at *Hoxd4* regulatory regions distinguish anterior from posterior embryonic compartments. *Mol. Cell. Biol* 24, 8090-8103.
- Reddy, K. L., Zullo, J. M., Bertolino, E., et Singh, H. (2008). Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* 452, 243-247.
- Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., Goodnough, L. H., Helms, J. A., Farnham, P. J., Segal, E., et al. (2007). Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell* 129, 1311-1323.
- Robyr, D., Friedli, M., Gehrig, C., Arcangeli, M., Marin, M., Guipponi, M., Farinelli, L., Barde, I., Verp, S., Trono, D., et al. (2011). Chromosome conformation capture uncovers potential genome-wide interactions between human conserved non-coding sequences. *PLoS ONE* 6, e17634.
- Roldan, E., Fuxa, M., Chong, W., Martinez, D., Novatchkova, M., Busslinger, M., et Skok, J. A. (2005). Locus « decontraction » and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. *Nat Immunol* 6, 31-41.
- Roudier, F., Ahmed, I., Bérard, C., Sarazin, A., Mary-Huard, T., Cortijo, S., Bouyer, D., Caillieux, E., Duvernois-Berthet, E., Al-Shikhley, L., et al. (2011). Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *EMBO J* 30, 1928-1938.
- Rougeulle, C., Chaumeil, J., Sarma, K., Allis, C. D., Reinberg, D., Avner, P., et Heard, E. (2004). Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Molecular and cellular biology* 24, 5475.
- Royce-Tolland, M. E., Andersen, A. A., Koyfman, H. R., Talbot, D. J., Wutz, A., Tonks, I. D., Kay, G. F., et Panning, B. (2010). The A-repeat links ASF/SF2-dependent Xist RNA processing with random choice during X inactivation. *Nat Struct Mol Biol* 17, 948-954.
- Ruault, M., Dubarry, M., et Taddei, A. (2008). Re-positioning genes to the nuclear envelope in mammalian cells: impact on transcription. *Trends Genet* 24, 574-581.
- Rudolph, T., Yonezawa, M., Lein, S., Heidrich, K., Kubicek, S., Schäfer, C., Phalke, S., Walther, M., Schmidt, A., Jenuwein, T., et al. (2007). Heterochromatin formation in *Drosophila* is initiated through active removal of H3K4 methylation by the LSD1 homolog SU(VAR)3-3. *Mol. Cell* 26, 103-115.
- Ruf, S., Symmons, O., Uslu, V. V., Dolle, D., Hot, C., Eттwiller, L., et Spitz, F. (2011). Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor. *Nat Genet* 43, 379-386.
- Ryba, T., Hiratani, I., Lu, J., Itoh, M., Kulik, M., Zhang, J., Schulz, T. C., Robins, A. J., Dalton, S., et Gilbert, D. M. (2010). Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Research* 20, 761-770.

- Sachs, R. K., van den Engh, G., Trask, B., Yokota, H., et Hearst, J. E. (1995). A random-walk/giant-loop model for interphase chromosomes. *Proceedings of the National Academy of Sciences* 92, 2710-2714.
- Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M., et Shiroishi, T. (2005). Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development* 132, 797.
- Sagai, T., Amano, T., Tamura, M., Mizushina, Y., Sumiyama, K., et Shiroishi, T. (2009). A cluster of three long-range enhancers directs regional Shh expression in the epithelial linings. *Development* 136, 1665-1674.
- Sandhu, K. S., Shi, C., Sjölander, M., Zhao, Z., Gündör, A., Liu, L., Tiwari, V. K., Guibert, S., Emilsson, L., Imreh, M. P., et al. (2009). Nonallelic transvection of multiple imprinted loci is organized by the H19 imprinting control region during germline development. *Genes & development* 23, 2598.
- De Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B. K., Muller, H., Ragoussis, J., Wei, C.-L., et Natoli, G. (2010). A Large Fraction of Extragenic RNA Pol II Transcription Sites Overlap Enhancers. *PLoS Biol* 8, e1000384.
- Saunders, A., Core, L. J., et Lis, J. T. (2006). Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol* 7, 557-567.
- Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N. F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J. A., Umlauf, D., Dimitrova, D. S., et al. (2010). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat. Genet* 42, 53-61.
- Schorderet, P., et Duboule, D. (2011). Structural and functional differences in the long non-coding RNA hotair in mouse and human. *PLoS Genet* 7, e1002071.
- Schübeler, D. (2007). Enhancing genome annotation with chromatin. *Nature Genetics* 39, 283-284.
- Schulz, E. G., Nora, E. P., et Heard, E. (2011). Rnf12—A Jack of All Trades in X Inactivation? *PLoS Genetics* 7.
- Shin, J., Bossenz, M., Chung, Y., Ma, H., Byron, M., Taniguchi-Ishigaki, N., Zhu, X., Jiao, B., Hall, L. L., Green, M. R., et al. (2010). Maternal Rnf12/RLIM is required for imprinted X-chromosome inactivation in mice. *Nature* 467, 977-981.
- Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., et de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture—on-chip (4C). *Nat Genet* 38, 1348-1354.
- Simpson, P., et Ayyar, S. (2008). Chapter 3 Evolution of Cis[hyphen (true graphic)]Regulatory Sequences in Drosophila. Dans *Long-Range Control of Gene Expression Academic Press*, p. 67-106. Available at: <http://www.sciencedirect.com/science/article/pii/S006526600700003X> [Consulté juin 29, 2011].
- Singh, V., et Srivastava, M. (2008). Enhancer blocking activity of the insulator at H19-ICR is independent of chromatin barrier establishment. *Mol. Cell. Biol* 28, 3767-3775.
- Slotkin, R. K., et Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* 8, 272-285.
- Smale, S. T. (2010). Pioneer factors in embryonic stem cells and differentiation. *Current Opinion in Genetics & Development* 20, 519-526.
- Soshnikova, N., et Duboule, D. (2009). Epigenetic Temporal Control of Mouse Hox Genes in Vivo. *Science* 324, 1320-1323.
- Soshnikova, N., Montavon, T., Leleu, M., Galjart, N., et Duboule, D. (2010). Functional Analysis of CTCF During Mammalian Limb Development. *Developmental Cell* 19, 819-830.

- Spellman, P. T., et Rubin, G. M. (2002). Evidence for large domains of similarly expressed genes in the *Drosophila* genome. *Journal of Biology* 1, 5.
- Spilianakis, C. G., et Flavell, R. A. (2004). Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat Immunol* 5, 1017-1027.
- Spilianakis, C. G., Lalioti, M. D., Town, T., Lee, G. R., et Flavell, R. A. (2005). Interchromosomal associations between alternatively expressed loci. *Nature* 435, 637-645.
- Spitz, F., Gonzalez, F., Peichel, C., Vogt, T. F., Duboule, D., et Zákány, J. (2001). Large scale transgenic and cluster deletion analysis of the *HoxD* complex separate an ancestral regulatory module from evolutionary innovations. *Genes Dev* 15, 2209-2214.
- Spitz, F., et Duboule, D. (2008). Global control regions and regulatory landscapes in vertebrate development and evolution. *Advances in Genetics* 61, 175–205.
- Spitz, F., Gonzalez, F., et Duboule, D. (2003). A global control region defines a chromosomal regulatory landscape containing the *HoxD* cluster. *Cell* 113, 405–417.
- Splinter, E. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes & Development* 20, 2349-2354.
- Splinter, E., de Wit, E., Nora, E. P., Klous, P., van de Werken, H. J. G., Zhu, Y., Kaaij, L. J. T., van IJcken, W., Gribnau, J., Heard, E., et al. (2011). The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. *Genes & Development*. Available at: <http://genesdev.cshlp.org/content/early/2011/06/16/gad.633311.abstract> [Consulté juin 23, 2011].
- Stalder, J., Larsen, A., Engel, J. D., Dolan, M., Groudine, M., et Weintraub, H. (1980). Tissue-specific DNA cleavages in the globin chromatin domain introduced by DNAase I. *Cell* 20, 451-460.
- van Steensel, B., et Dekker, J. (2010). Genomics tools for unraveling chromosome architecture. *Nat. Biotechnol* 28, 1089-1095.
- Sun, B. K., Deaton, A. M., et Lee, J. T. (2006). A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization. *Molecular cell* 21, 617–628.
- Szabó, P., Tang, S. H., Rentsendorj, A., Pfeifer, G. P., et Mann, J. R. (2000). Maternal-specific footprints at putative CTCF sites in the H19 imprinting control region give evidence for insulator function. *Curr. Biol* 10, 607-610.
- Szutorisz, H., Canzonetta, C., Georgiou, A., Chow, C. M., Tora, L., et Dillon, N. (2005a). Formation of an active tissue-specific chromatin domain initiated by epigenetic marking at the embryonic stem cell stage. *Molecular and cellular biology* 25, 1804.
- Szutorisz, H., Dillon, N., et Tora, L. (2005b). The role of enhancers as centres for general transcription factor recruitment. *Trends in Biochemical Sciences* 30, 593-599.
- Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., et al. (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes & Development* 16, 1779 -1791.
- Takahashi, K., et Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* 126, 663-676.
- Talbert, P. B., et Henikoff, S. (2000). A reexamination of spreading of position-effect variegation in the white-rough region of *Drosophila melanogaster*. *Genetics* 154, 259-272.
- The ENCODE Project Consortium (2011). A User's Guide to the Encyclopedia of DNA Elements (ENCODE). *PLoS Biol* 9, e1001046.
- The modENCODE Consortium, Roy, S., Ernst, J., Kharchenko, P. V., Kheradpour, P., Negre, N., Eaton, M. L., Landolin, J. M., Bristow, C. A., Ma, L., et al. (2010). Identification of

- Functional Elements and Regulatory Circuits by *Drosophila* modENCODE. *Science* 330, 1787-1797.
- Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F., et de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active [beta]-globin locus. *Molecular cell* 10, 1453–1465.
- Tolhuis, B., Blom, M., Kerkhoven, R. M., Pagie, L., Teunissen, H., Nieuwland, M., Simonis, M., de Laat, W., van Lohuizen, M., et van Steensel, B. (2011). Interactions among Polycomb domains are guided by chromosome architecture. *PLoS Genet* 7, e1001343.
- Trimborn, T., Gribnau, J., Grosveld, F., et Fraser, P. (1999). Mechanisms of developmental control of transcription in the murine alpha- and beta-globin loci. *Genes Dev* 13, 112-124.
- Tunster, S. J., Van De Pette, M., et John, R. M. (2011). BACs as Tools for the Study of Genomic Imprinting. *J Biomed Biotechnol* 2011.
- Udvardy, A., Maine, E., et Schedl, P. (1985). The 87A7 chromomere : Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. *Journal of Molecular Biology* 185, 341-358.
- Valouev, A., Johnson, S. M., Boyd, S. D., Smith, C. L., Fire, A. Z., et Sidow, A. (2011). Determinants of nucleosome organization in primary human cells. *Nature advance online publication*. Available at: <http://dx.doi.org/10.1038/nature10002> [Consulté juin 10, 2011].
- Vieira, K. F., Levings, P. P., Hill, M. A., Crusselle, V. J., Kang, S.-H. L., Engel, J. D., et Bungert, J. (2004). Recruitment of transcription complexes to the beta-globin gene locus in vivo and in vitro. *J. Biol. Chem* 279, 50350-50357.
- Vilar, J. M., et Saiz, L. (2005). DNA looping in gene regulation: from the assembly of macromolecular complexes to the control of transcriptional noise. *Current Opinion in Genetics & Development* 15, 136-144.
- Visel, A., Blow, M. J., Li, Z., Zhang, T., Akiyama, J. A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457, 854-858.
- Vogel, M. J., Pagie, L., Talhout, W., Nieuwland, M., Kerkhoven, R. M., et Van Steensel, B. (2009). High-resolution mapping of heterochromatin redistribution in a *Drosophila* position-effect variegation model. *Epigenetics & Chromatin* 2, 1.
- Walter, J., Dever, C. A., et Biggin, M. D. (1994). Two homeo domain proteins bind with similar specificity to a wide range of DNA sites in *Drosophila* embryos. *Genes Dev* 8, 1678-1692.
- Walter, K., Bonifer, C., et Tagoh, H. (2008). Stem cell-specific epigenetic priming and B cell-specific transcriptional activation at the mouse Cd19 locus. *Blood* 112, 1673.
- Walters, M. C., Fiering, S., Eidemiller, J., Magis, W., Groudine, M., et Martin, D. I. (1995). Enhancers increase the probability but not the level of gene expression. *Proceedings of the National Academy of Sciences of the United States of America* 92, 7125.
- Wan, L.-B., Pan, H., Hannenhalli, S., Cheng, Y., Ma, J., Fedoriw, A., Lobanenko, V., Latham, K. E., Schultz, R. M., et Bartolomei, M. S. (2008). Maternal depletion of CTCF reveals multiple functions during oocyte and preimplantation embryo development. *Development* 135, 2729-2738.
- Wathelet, M. G., Lin, C. H., Parekh, B. S., Ronco, L. V., Howley, P. M., et Maniatis, T. (1998). Virus infection induces the assembly of coordinately activated transcription factors on the IFN-beta enhancer in vivo. *Mol. Cell* 1, 507-518.
- Wen, B., Wu, H., Shinkai, Y., Irizarry, R. A., et Feinberg, A. P. (2009). Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat. Genet* 41, 246-250.
- West, A. G., et Fraser, P. (2005). Remote control of gene transcription. *Human molecular genetics* 14,

R101.

- West, A. G., Huang, S., Gaszner, M., Litt, M. D., et Felsenfeld, G. (2004). Recruitment of Histone Modifications by USF Proteins at a Vertebrate Barrier Element. *Molecular Cell* *16*, 453-463.
- Williams, A., Spilianakis, C. G., et Flavell, R. A. (2010). Interchromosomal association and gene regulation in trans. *Trends in Genetics* *26*, 188-197.
- Wu, C., Bingham, P. M., Livak, K. J., Holmgren, R., et Elgin, S. C. (1979a). The chromatin structure of specific genes: I. Evidence for higher order domains of defined DNA sequence. *Cell* *16*, 797-806.
- Wu, C., Wong, Y. C., et Elgin, S. C. (1979b). The chromatin structure of specific genes: II. Disruption of chromatin structure during gene activity. *Cell* *16*, 807-814.
- Wu, R., Terry, A. V., Singh, P. B., et Gilbert, D. M. (2005). Differential subnuclear localization and replication timing of histone H3 lysine 9 methylation states. *Mol. Biol. Cell* *16*, 2872-2881.
- Würtele, H., et Chartrand, P. (2006). Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. *Chromosome Res* *14*, 477-495.
- Wutz, A., Smrzka, O. W., Schweifer, N., Schellander, K., Wagner, E. F., et Barlow, D. P. (1997). Imprinted expression of the Igf2r gene depends on an intronic CpG island. *Nature* *389*, 745-749.
- Xie, X., Mikkelsen, T. S., Gnirke, A., Lindblad-Toh, K., Kellis, M., et Lander, E. S. (2007). Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. *Proc. Natl. Acad. Sci. U.S.A* *104*, 7145-7150.
- Xu, J., Pope, S. D., Jazirehi, A. R., Attema, J. L., Papathanasiou, P., Watts, J. A., Zaret, K. S., Weissman, I. L., et Smale, S. T. (2007). Pioneer factor interactions and unmethylated CpG dinucleotides mark silent tissue-specific enhancers in embryonic stem cells. *Proceedings of the National Academy of Sciences* *104*, 12377.
- Xu, J., Watts, J. A., Pope, S. D., Gadue, P., Kamps, M., Plath, K., Zaret, K. S., et Smale, S. T. (2009). Transcriptional competence and the active marking of tissue-specific enhancers by defined transcription factors in embryonic and induced pluripotent stem cells. *Genes & Development* *23*, 2824-2838.
- Xu, N., Tsai, C. L., et Lee, J. T. (2006). Transient homologous chromosome pairing marks the onset of X inactivation. *Science* *311*, 1149.
- Yokochi, T., Poduch, K., Ryba, T., Lu, J., Hiratani, I., Tachibana, M., Shinkai, Y., et Gilbert, D. M. (2009). G9a selectively represses a class of late-replicating genes at the nuclear periphery. *Proceedings of the National Academy of Sciences* *106*, 19363.
- Yokota, H., van den Engh, G., Hearst, J. E., Sachs, R. K., et Trask, B. J. (1995). Evidence for the organization of chromatin in megabase pair-sized loops arranged along a random walk path in the human G0/G1 interphase nucleus. *J. Cell Biol* *130*, 1239-1249.
- Yokota, H., Singer, M. J., Van den Engh, G. J., et Trask, B. J. (1997). Regional differences in the compaction of chromatin in human G 0/G 1 interphase nuclei. *Chromosome Research* *5*, 157-166.
- Yoon, Y. S., Jeong, S., Rong, Q., Park, K.-Y., Chung, J. H., et Pfeifer, K. (2007). Analysis of the H19ICR insulator. *Mol. Cell. Biol* *27*, 3499-3510.
- Yunger, S., Rosenfeld, L., Garini, Y., et Shav-Tal, Y. (2010). Single-allele analysis of transcription kinetics in living mammalian cells. *Nat Meth* *7*, 631-633.
- Zhao, Z., Tavoosidana, G., Sjölander, M., Göndör, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Singh Sandhu, K., Singh, U., et al. (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal

interactions. *Nat Genet* 38, 1341-1347.

Zinzen, R. P., Girardot, C., Gagneur, J., Braun, M., et Furlong, E. E. M. (2009). Combinatorial binding predicts spatio-temporal cis-regulatory activity. *Nature* 462, 65-70.

Zlatanova, J., et Caiafa, P. (2009). CTCF and its protein partners: divide and rule? *Journal of Cell Science* 122, 1275.