



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *Université Toulouse 3 - Paul Sabatier*

Discipline ou spécialité : *Biologie moléculaire et cellulaire*

Présentée et soutenue par *Magali Hennion*
Le 3 décembre 2010

Titre : *Mécanismes de régulation de la transcription par les insulateurs
impliquant leur rôle dans l'organisation de la chromatine*

Ecole doctorale : *École Doctorale Biologie, Santé, Biotechnologies*
Unité de recherche : *Laboratoire de Biologie Moléculaire Eucaryote*
Directeur(s) de Thèse : *Dr. Emmanuel Käs et Dr. Olivier Cuvier*

Rapporteurs :

M. Christophe Escudé MNHN, Paris
M. Olivier Hyrien ENS, Paris
M. Stefan Dimitrov IAB, Grenoble

Autre(s) membre(s) du jury :

M. Emmanuel Käs	LBME, Toulouse	Directeur de thèse
M. Olivier Cuvier	LBME, Toulouse	Directeur de thèse
M. Pierre-Emmanuel Gleizes	LBME, Toulouse	Président du jury
M. Frédéric Bantignies	IGH, Montpellier	Examinateur

Remerciements

Tout d'abord, je voudrais remercier chaleureusement l'ensemble des membres de mon jury. En premier lieu, les rapporteurs de mon manuscrit : Stefan Dimitrov qui n'a malheureusement pas pu assister à ma soutenance et que je regrette de n'avoir pas encore rencontré, j'espère que l'occasion se présentera dans l'avenir ; Olivier Hyrien et Christophe Escudé, pour la qualité de vos questions lors de la soutenance et pour vos encouragements. Je remercie également Frédéric Bantignies pour sa participation en tant qu'examinateur et Pierre-Emmanuel Gleizes qui a accepté de présider ce jury. L'ensemble de vos interventions ont fait de la discussion un moment très agréable et riche scientifiquement.

Je remercie l'ensemble de l'équipe Käs. Un grand merci à Emmanuel, qui m'a encadrée et soutenue pendant tout le début de ma thèse malgré des circonstances extérieures difficiles. Prends soin de toi, maintenant que tu m'as menée à bon port, tu n'as plus à te faire de soucis pour moi ! Merci à Roxane et Caroline pour m'avoir tout appris sur le travail de paillasse. Sans vous je ne m'en serais pas sortie ! Merci aussi aux petites mains qui m'ont bien aidée pendant leur stage : Sana et Quitterie.

Je remercie aussi l'équipe Cuvier en croissance exponentielle. Merci à Olivier, d'abord pour la collaboration fructueuse du début de ma thèse puis pour ton encadrement ces derniers 18 mois. J'ai beaucoup appris à ton contact et je pars bien formée pour mon post-doc. Bon courage pour la suite ! Merci à Sophie qui est venue me tenir compagnie avant l'arrivée du chef, merci pour toute l'aide que tu m'as apportée au labo, ainsi que pour les conseils en pâtisserie ! Merci à Priscillia, notre petite princesse, pour ta fraîcheur et ta bonne humeur, pour les coups de mains que tu m'as donnés et pour ta gestion exemplaire du stock de thés et tisanes ! Merci à Marine qui m'a bien aidée pendant son stage productif. Bonne chance pour la suite ! Un grand merci à Adrien pour ta patience avec ces biologistes qui ont tant de mal avec les biostats ! Merci aux renforts : Pauline, pour ton soutien sans faille en particulier pendant ces dernières semaines de rush, et pour m'avoir évité de dormir dans ma voiture... Je vous laisse à toi et Sophie, un gros fardeau de trucs à finir, désolée. Jun, à qui je laisse aussi pas mal de boulot, bon courage ! Laurent, merci beaucoup pour ton aide durant ces dernières semaines, en particulier pendant la préparation de ma soutenance. Merci aussi à Manu pour ton soutien moral et tes conseils avisés pour la préparation de mon oral. Je regrette de partir quand vous arrivez tous les deux ! Enfin, merci beaucoup à Eldon, notre collaborateur outre-atlantique qui a abattu un travail de biostats considérable. Ça a été un plaisir de t'avoir parmi nous quelques mois, j'aurais aimé avoir plus de temps pour discuter avec toi et profiter de ta science.

Je voudrais aussi remercier l'ensemble du LBME : en premier lieu, l'équipe Bystricky qui m'a adoptée pendant ma période solitaire. Merci à tous les Bystrickets et Bystrickettes, en particulier Imen, Isa, Mathieu, Luca, Silvia et les anciennes Steph et Magali pour tous les bons moments passés ! Merci à l'ensemble du rez-de-chaussée, les Grigoriev et les Gadaux pour les échanges d'idées et de bons procédés. Merci aux gens du second, aux Ferrer-Henry, en particulier Michou pour sa direction du labo et ses encouragements, aux Cavaillés, aux Kiss et aux Gleizes. C'est un grand plaisir de travailler dans ce laboratoire où tout le monde s'entraide volontiers dans la bonne humeur. Merci aux collègues du LMGM qui partagent aussi leur matériel et leurs connaissances généreusement. Enfin, merci aux non-biologistes qui font tourner le labo : l'accueil, le service info, l'équipe Technik, la laverie, et bien sûr nos gestionnaires préférées Safi, Cathy et Isabelle.

Merci à mes copains Toulousains : Jeannot, Coco, Vaness, Alan et Sophie, Ambre et Alex, Tif et Isa, Anna et Yann, Ju et Noëla, pour toutes ces bons moments passées ensemble autour d'un repas ou d'un jeu. Vous allez me manquez! Merci aussi aux amies d'« enfance » de CP, Sigo et Sandra pour les heures passées au téléphone! Enfin, un grand merci à ma petite Nan, pour les grands moments de complicité à chaque retour sur Gre. Merci à Da aussi, c'est beau de continuer à se taper des fous rires délirants après toutes ces années, vous êtes mes jumelles préférées!

Je voudrais bien sûr remercier Thomas, mon amour. Professionnellement d'abord pour le développement du logiciel pyQPCR intégrant tous mes caprices et qui m'a fait gagner énormément de temps. Et surtout personnellement pour la joie toujours renouvelée de vivre ensemble. C'est bon d'avoir quelqu'un sur qui s'appuyer en toute circonstance. Je suis très heureuse de te rejoindre bientôt dans ce pays froid pour écrire une nouvelle page de notre histoire.

Pour finir, je voudrais remercier ma famille adorée. Ma mamie Jeanne pour la correspondance abondante (parfois à sens unique, pardonne-moi) que nous entretenons. Bon Pa qui est si fier d'avoir un docteur de plus dans la famille, merci beaucoup pour tes encouragements! Mes frangins que j'aimerais voir plus souvent : Alex et toutes ses adorables femmes, Benji, matheux mais sociable, indulgent avec sa soeur qui n'y comprend rien, et Roméo, notre poète qui dessine des monstres en attendant qu'on le paye pour ça, je crois en vous deux les petits! Et mes parents bien sûr, merci de votre soutien quelque soit la direction que je prends, c'est fou cette confiance que vous avez en moi! Bonne chance pour votre nouvelle vie dans le pays pluvieux où vous vous êtes exilés. On viendra vous voir, ça nous changera de la neige!



Sommaire

Sommaire	5
Liste des abréviations utilisées	7
Introduction	11
1 Contexte	13
2 Démarche scientifique et résultats	59
3 Discussion et perspectives	125
Conclusion	137
A Nomenclature	139
B Matériel et méthodes	141
Liste des figures	147
Liste des tables	148
Bibliographie	149
Table des matières	173



Liste des notations utilisées

3C Chromosome Conformation Capture (Capture de Conformation Chromosomique)
abd-A Abdominal A
Abd-B Abdominal B
ACF ATP-utilizing Chromatin assembly and remodelling Factor
ADN Acide Désoxyribonucléique
APC Anaphase Promoting Complex
ARN Acide Ribonucléique
ARNdb ARN double-brin
ARNt ARN de transfert
Ash1 Absent, small or homeotic discs 1
ATP Adénosine Triphosphate
BE Boundary Element
BE28 Boundary Element 28
BEAF Boundary Element Associated Factor
BESS BEAF/Su(var)3-7/Stonewall
BID BEAF self-Interaction Domain
BLAST Basic Local Alignment Search Tool
BRE TFIIB Responsive Element
BTB Bric-a-brac, Tramtrack, Broad-complex
CAF1 Chromatin assembly factor 1
Cdk7 Cyclin-dependante kinase 7
Cdk9 Cyclin-dependante Kinase 9
cDNA ADN complémentaire
CEMOVIS Cryo-Electron Microscopy of Vitreous Sections
CENP- A Centromere Protein A
CHD Chromodomain, Helicase, DNA binding
ChIP-chip Immunoprécipitation de Chromatine suivie d'une hybridation sur puce
ChIP-seq Immunoprécipitation de Chromatine suivie du séquençage des ADN
CHRAC Chromatin Accessibility Complex
COMPASS Complex of Proteins Associated with Set1
CP190 Centrosomal Protein 190 Kd
CREB/ATF cAMP Response Element Binding protein/Activating Transcription Factor
CTCF CCCTC-binding Factor
CTD Carboxy-Terminal Domain
DNase I Désoxyribonucléase I

Dnmt2 DNA methyltransferase 2
DPE Downstream Promoter Element
DRB 5,6-dichloro-1- β -D-ribofuranosylbenzimidazole
DREF DNA Replication-related Element Factor
DSIF DRB Sensitivity-Inducing Factor
ELL Eleven nineteen Lysine-rich Leukemia
E(Var) Activateur (Enhancer) de Variégation
FACS Fluorescence-activated cell sorting
FACT Facilitates Chromatin Transcription
Fcp1 TFIIF interacting CTD phosphatase
FISH Fluorescent *In Situ* Hybridization
GAF GAGA factor
HAT Histone Acétyl-Transférase
HDAC Histone Désacétylase
HIRA Histone Regulating Protein A
HMT Histone Methyl Transférase
HP1 Heterochromatin Protein 1
HP2 Heterochromatin Protein 2
HS site hypersensible
HSA helicase-SANT
HSF Heat Shock Factor
ICR Imprinted Control Region
INO80 Inositol requiring 80
Inr élément initiateur
ISWI Imitation Switch
LCR Locus Control Region
MAP-kinase Mitogen-Activated Protein kinase
Mi-2/NuRD Nucleosome Remodeling Histone Deacetylase
MNase Micrococcal nuclease
MNase-seq MNase suivie du séquençage des ADN
mod(mdg4) modifier of mdg4
MOF Males absent On the First
Nap1 Nucleosome assembly protein 1
NASP Nuclear Autoantigenic Sperm Protein
NELF Negative Elongation Factor
NFR Nucleosome Free Region
NURF Nucleosome Remodeling Factor
Paf1 RNA polymerase II associated factor 1
pb paire de base
PCR Polymerase Chain Reaction
Pcf11 Protein 1 of Cleavage and polyadenylation Factor 1
Pc-G Polycomb Group

PEV Position-Effect Variegation
PHD Plant Homeo Domain
PIC Preinitiation complex (Complexe de préinitiation)
piRNA Piwi-interacting RNA
pol II ARN Polymérase II
PRC1 Polycomb Repressive Complex 1
PRC2 Polycomb Repressive Complex 2
PRE PcG Response Element
P-TEFb Positive Transcription Elongation Factor b
rDNA locus de transcription des ARN ribosomiques
RDRC RNA-Directed RNA polymerase Complex
RITS RNA-induced initiation of transcriptional gene silencing
RNAi interférence à l'ARN
RNA-seq Séquençage de l'ensemble des transcrits
rNTP ribonucléotide Triphosphate
rRNA ARN ribosomique
RSC Chromatin Structure Remodeling
RT-qPCR Reverse Transcription quantitative PCR
SANT Switching-defective protein 3, Adaptor 2, Nuclear receptor co-repressor, Transcription factor IIIB
SAT III séquences satellites de type III
scs et scs' Specialized chromatin structures
SETDB1 SET Domain Bifurcated 1
siRNA small interfering RNA (petit ARN interférent)
SLIDE SANT-like ISWI
SOM Self-Organizing Map
SRB Suppressors of RNA polymerase B mutations
Su(f) Suppressor or forked
Su(Hw) Suppressor of Hairy wing
Su(Var) Suppresseur de Variégation
SWI/SNF Switching defective/Sucrose Non Fermenting
SWR1 SWI/SNF-related 1
TAF TBP-Associated Factor
TBP TATA-Binding Protein
TF Transcription Factor
Topors Topoisomerase-I-interacting protein
TRE TrxG Response Element
Trl Trithorax-like
trx trithorax
TSS Transcription Start Site (site de démarrage de la transcription)
Ubx Ultrabithorax
wm white-mottled
Zw5 Zeste-white 5

Nota : Pour la nomenclature des nucléotides, voir code IUPAC en annexe [A.1](#). Pour la nomenclature des acides aminés, voir [A.2](#).

Introduction

Ce manuscrit comporte trois chapitres principaux ainsi que diverses annexes. Le premier chapitre est une présentation du contexte de mon travail de thèse (§1). Il comprend l'ensemble des informations importantes pour comprendre la suite du manuscrit. Mon travail s'intéresse au rôle des insulateurs liés par BEAF (Boundary Element Associated Factor) dans l'organisation de la chromatine et la régulation transcriptionnelle. La première partie de ce chapitre définit la chromatine et détaille sa structure et sa dynamique (§1.1). La seconde expose les mécanismes de la transcription par l'ARN polymérase II qui est responsable de la synthèse des ARN messagers (§1.2), ce qui constitue la première étape de régulation de l'expression des gènes. Vient ensuite une partie présentant l'influence de la chromatine sur le niveau de transcription (§1.3). Enfin, la dernière partie introduit les éléments insulateurs ainsi que les protéines impliquées dans la fonction insulatrice (§1.4).

Le second chapitre de ce manuscrit décrit la démarche expérimentale suivie au cours de ma thèse ainsi que les résultats obtenus (§2). Les génomes eucaryotes s'organisent avec différents niveaux de compaction de la fibre chromatinienne. L'hétérochromatine est généralement associée à la mise en silence de vastes domaines, tandis que l'euchromatine est une structure plus ouverte qui comprends la majorité des gènes exprimés. Les zones de transitions entre ces deux types de structures sont mal caractérisées à l'heure actuelle. Je me suis tout d'abord intéressée à l'organisation de la région de transition entre l'euchromatine et l'hétérochromatine sur le chromosome X de la drosophile. Il a été proposé que les insulateurs pourraient jouer un rôle dans la compartimentation de la chromatine. Nous avons étudié un insulateur situé dans cette région de transition et dont la fonction insulatrice est assurée par la protéine BEAF (§2.2). Il a été montré que les insulateurs liés par BEAF sont également présents dans les régions euchromatiniennes. Nous nous sommes donc ensuite intéressés à ces insulateurs et nous avons pu mettre en évidence leur rôle dans la régulation de la transcription, à la fois en activant directement certains gènes et en entrant en compétition avec l'activateur DREF en se liant sur d'autres promoteurs (§2.3). Dans une troisième partie, nous avons cherché à comprendre les mécanismes impliqués dans cette régulation. Nous nous sommes intéressés au rôle de BEAF dans le positionnement des nucléosomes qui apparaît comme important dans la régulation de l'étape de pause de l'ARN polymérase II (§2.4).

Le troisième chapitre de ma thèse discute mes résultats ainsi que les différents modèles proposés pour expliquer le rôle *in vivo* des insulateurs et de la protéine BEAF (§3). Une comparaison avec les autres insulateurs présents chez la drosophile est proposée. Les perspectives de ce travail sont également présentées. Elles concernent notamment l'étude des mécanismes de régulation de la fonction des insulateurs et de la formation de boucles de chromatine dans laquelle ils sont impliqués.



Contexte



Sommaire

1.1	La chromatine	14
1.1.1	Historique des découvertes majeures sur la chromatine	14
1.1.2	Structure de la chromatine	19
1.1.3	Domaines chromatinien	22
1.1.4	Modifications de la chromatine	24
1.1.5	Dynamique de la chromatine	27
1.2	La transcription des ARN messagers	35
1.2.1	Structure de la polymérase II	35
1.2.2	Formation du complexe de préinitiation	36
1.2.3	Initiation et pause	38
1.2.4	Reprise de l'élongation	40
1.2.5	Terminaison	41
1.2.6	Régulations de la transcription basale	42
1.3	Chromatine et transcription	43
1.3.1	Positionnement des nucléosomes	43
1.3.2	Remodelage de la chromatine aux promoteurs	47
1.3.3	Chromatine et élongation	47
1.4	Les insulateurs	49
1.4.1	Définition classique d'un insulateur	49
1.4.2	BEAF et Zw5	50
1.4.3	L'insulateur <i>gypsy</i>	54
1.4.4	Les insulateurs CTCF	54
1.4.5	Le facteur GAGA	55
1.4.6	CP190	55
1.4.7	les différents modèles d'insulation	56
1.4.8	Rôle des insulateurs <i>in vivo</i>	57

Ce chapitre détaille le contexte de mon travail de thèse. Après une présentation de la chromatine, de sa structure et sa dynamique, je résumerai les mécanismes de la transcription par l'ARN polymérase II ainsi que le rôle de la structure de la chromatine dans la régulation de cette activité. Enfin, je présenterai les insulateurs qui sont des éléments d'ADN impliqués dans l'organisation de la chromatine.

1.1 La chromatine

1.1.1 Historique des découvertes majeures sur la chromatine

On peut considérer que l'histoire de la chromatine (pour revue : [Olins & Olins 2003](#)) commence avec Flemming qui a inventé le terme « chromatine » (Gr. chroma, couleur) dans les années 1880 pour identifier une substance dans le noyau qui avait tendance à fixer les colorants. Il étudiait alors la division cellulaire. À cette époque se développaient différentes techniques de préparation, de coloration et de fixation des échantillons, et le microscope s'améliorait. Des études biochimiques émergeaient également, et en 1871 Miescher, qui avait mis au point une technique pour isoler des noyaux de leucocytes à partir de pus récupéré sur des bandages chirurgicaux, décrit pour la première fois un acide riche en phosphore et en azote qu'il appela « nucléine » ([Miescher 1871](#), et pour revue : [Dahm & Miescher \(2008\)](#)). Plus tard, en étudiant des spermatozoïdes de saumon, il montra que la nucléine était associée à une molécule basique qu'il nomma la « protamine ».

Ces études furent continuées par Kossel qui caractérisa le composant basique associé à la nucléine et qu'il appela en 1884 « histone » ([Kossel 1911](#), les protamines de Miescher sont des protéines différentes qui remplacent les histones uniquement dans les spermatozoïdes). Il précisa également la nature de la nucléine (qu'il appela acide nucléique en pensant avoir découvert une autre substance) en montrant la présence de bases puriques et pyrimidiques, d'un sucre, et d'acide phosphorique. Il reçut le prix Nobel en 1910 pour ces études.

À cette époque, le lien avec la transmission héréditaire des caractères mise en évidence en 1865 par Mendel n'a pas encore été établi. Mendel était moine et botaniste et avait installé un jardin expérimental dans son monastère pour étudier la formation des hybrides chez le pois (Fig. 1.1). Après dix années de travaux minutieux, il présente les résultats de ces études en 1865 et les publie l'année suivante ([Mendel 1866](#)). Il en ressort trois lois fondamentales de la génétique (appelées plus tard « lois de Mendel » par Correns) :

Première loi : Uniformité des hybrides de première génération : aucune forme intermédiaire n'apparaît quand les parents sont de souches pures. Le concept de l'hérédité par mélange est réfuté.

















Deuxième loi : Ségrégation indépendante de plusieurs couples de caractères en seconde génération. Cela correspond à la loi de pureté des gamètes : chaque gamète ne contient toujours qu'un seul facteur héréditaire pour un caractère donné (une seule version de chaque gène). Cela s'explique par l'haploïdie des gamètes qui résulte de la méiose cellulaire spécifique à la gamétogénèse.





Troisième loi : Ségrégation indépendante des caractères en seconde génération. Les différentes versions d'un caractère se séparent et se réassortissent indépendamment de celle d'un autre caractère.

En conclusion, Mendel propose que les caractéristiques héréditaires des êtres vivants sont gouvernées chacune par une double commande (une paire d'allèles) et que seule une sur deux est transmise au descendant par chaque parent. Il pose les bases théoriques de la génétique et de l'hérédité modernes.

Les travaux de Mendel n'ont pas un fort impact à l'époque. Quasiment oubliés, ils seront redécouverts en 1900 par de Vries qui les généralise à l'ensemble des plantes ([De Vries 1900](#)), et par Cuénot qui décrit les mêmes lois chez les animaux ([Cuénot 1908](#)).

En 1902, Boveri et Sutton ont développé indépendamment la théorie chromosomique, le premier en étudiant le développement embryonnaire chez l'oursin ([Boveri 1902](#)) et le second en travaillant sur la méiose chez la sauterelle ([Sutton 1902](#)). Leur théorie prédisant que les chromosomes sont les facteurs de l'hérédité mendélienne fut controversée jusqu'à sa démonstration formelle dans les années 1910 par les travaux de Morgan chez la mouche *Drosophila melanogaster* (travaux détaillés dans son livre [Morgan et al. 1915](#)). La première mutation qu'il observe est une mouche mâle présentant des yeux blancs alors qu'ils sont

		♂ gametes			
		RY $\frac{1}{4}$	Ry $\frac{1}{4}$	ry $\frac{1}{4}$	rY $\frac{1}{4}$
♀ gametes	RY $\frac{1}{4}$	$RR YY$ $\frac{1}{16}$ 	$RR Yy$ $\frac{1}{16}$ 	$Rr Yy$ $\frac{1}{16}$ 	$Rr YY$ $\frac{1}{16}$ 
	Ry $\frac{1}{4}$	$RR Yy$ $\frac{1}{16}$ 	$RR yy$ $\frac{1}{16}$ 	$Rr yy$ $\frac{1}{16}$ 	$Rr Yy$ $\frac{1}{16}$ 
	ry $\frac{1}{4}$	$Rr Yy$ $\frac{1}{16}$ 	$Rr yy$ $\frac{1}{16}$ 	$rr yy$ $\frac{1}{16}$ 	$rr Yy$ $\frac{1}{16}$ 
	rY $\frac{1}{4}$	$Rr YY$ $\frac{1}{16}$ 	$Rr Yy$ $\frac{1}{16}$ 	$rr Yy$ $\frac{1}{16}$ 	$rr YY$ $\frac{1}{16}$ 

9  : 3  : 3  : 1 





 Round, yellow  Winkled, yellow
 Round, green  Winkled, green

Figure 1.1 – Schéma d'hybridation des pois par Mendel (1866).

rouges chez l'animal sauvage (Fig. 1.2). Il remarque que ce caractère n'est présent que chez les mâles, ce qui le conduit à proposer que le facteur déterminant ce caractère est porté par un chromosome sexuel. Ainsi, un facteur mendélien est pour la première fois expérimentalement assigné à un chromosome défini.

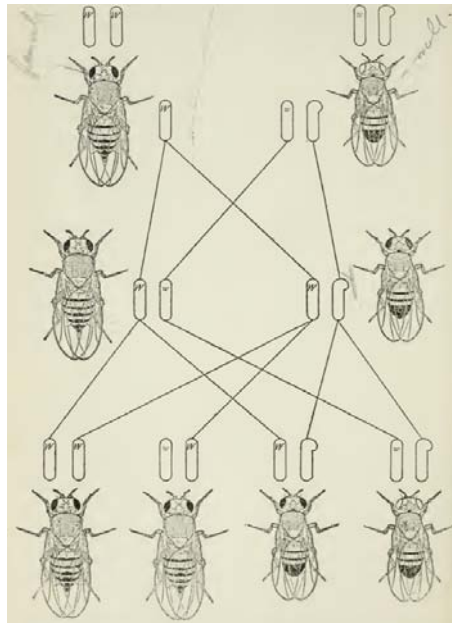


Figure 1.2 – La mutation *white* et son hérédité liée au sexe (Morgan 1919). Morgan observa que seules les mouches mâles présentent le phénotype d'oeil blanc, ce qui implique que ce caractère est défini par un chromosome sexuel.

Ainsi, ce début de XX^e siècle voit de grandes avancées en génétique, mais la connaissance de la structure de la chromatine avance peu. L'observation de chromosomes mitotiques de plantes conduit Heitz à définir deux types de chromatine (Heitz 1928). L'hétérochromatine, littéralement « l'autre chromatine », se distingue de l'euchromatine, la « vraie chromatine », par une coloration au carmin acétique plus forte (Fig. 1.3) et par le fait que cette chromatine reste condensée après la mitose (voir §1.1.3). Heitz émet alors la théorie

cette fonction (Avery *et al.* 1944). Un certain « principe transformant » avait été décrit en 1928 par Griffith (Griffith 1928). Il avait constaté qu'en injectant à une souris des pneumocoques de souche R vivants (« rough » et non virulents) ainsi qu'une petite quantité de pneumocoques S (« smooth », encapsulés et virulents) tués, la souris meurt et que son sang contient des pneumocoques des deux souches (Fig. 1.6). Ainsi, l'information contenue dans les bactéries mortes est passée dans les vivantes par ce qu'on appelle encore aujourd'hui « transformation ». Avery, MacLeod et McCarty ont montré que les extraits bactériens restaient capables de « transformer » après un traitement avec des protéases (qui dégradent les protéines) ou avec une ribonucléase (qui dégrade l'ARN) mais pas après un traitement avec une « dépolymérase à ADN » (une préparation brute capable de dégrader l'ADN, on parle aujourd'hui de « nucléase »), prouvant ainsi que c'est bien l'ADN qui contient l'information génétique.

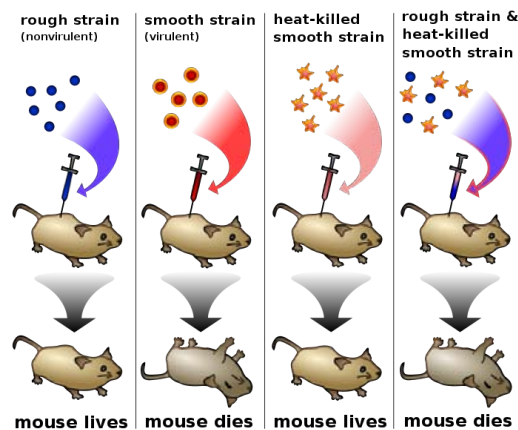


Figure 1.6 – Expérience de Griffith mettant en évidence l'existence d'un « principe transformant » (voir texte, image Wikimedia Commons).

En 1953, Watson et Crick publient la structure de l'ADN (Watson & Crick 1953) en utilisant la diffraction aux rayons X (images de Franklin & Gosling 1953). Il s'agit d'une double hélice formée de deux brins composés de groupements phosphates et de sucres (Fig. 1.7). Des bases azotées (adénine, thymine, guanine et cytosine) sont liées sur les sucres des deux brins, une cytosine s'appariant à une guanine, et une adénine s'appariant à une thymine (Fig. 1.8). Cette structure est en accord avec les résultats de Chargaff qui avait observé qu'il y avait autant de cytosines que de guanines et autant d'adénines que de thymines dans l'ADN purifié (Chargaff 1950).



Figure 1.7 – Représentation schématique de la double hélice d'ADN (Watson & Crick 1953).

Après le succès de la résolution de la structure de l'ADN par la diffraction aux rayons X, cette technique est utilisée dans l'espoir de résoudre la structure de la nucléohistone que l'on pense être une macromolécule fibreuse où l'ADN est recouvert par une gaine de protéines (les histones). En 1970, Pardon et Richards proposent une structure en super-

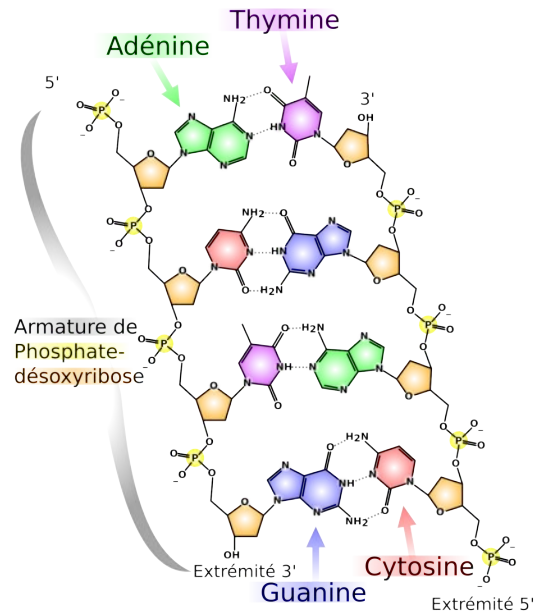


Figure 1.8 – Structure chimique de l’ADN (image Wikimedia Commons). Le squelette est formé d’une succession de groupements phosphates et de sucres sur lesquels sont liées les bases azotées. Les tirets figurent les liaisons hydrogènes impliquées dans l’appariement des bases.

hélice de 65 Å de rayon et de 120 Å de pas mais considèrent que leurs données ne sont pas suffisantes pour résoudre précisément la structure (Pardon *et al.* 1974). Bram propose quant à lui, un rayon de 50 Å et un pas nettement plus petit de 45 Å (DuPraw 1974). Ces variations importantes viennent du fait que les fibres de chromatines ne donnent pas des spots assez fins par diffraction, contrairement à ce qui a été observé sur l’ADN.

Les différentes données de microscopie électronique obtenues dans ces années là ne correspondaient pas toujours avec ces modèles de super-hélice mais étaient assez peu considérées. Swift écrit même en 1974 que l’observation de chromosomes entiers sous un microscope électronique ressemble au mieux à un mauvais jour dans une usine de macaronis... (Swift 1974).

Il avait été observé que certaines régions de la chromatine étaient accessibles à la dégradation par une nucléase et par conséquent non recouvertes totalement par des protéines (Clark & Felsenfeld 1971), mais personne à l’époque n’imaginait l’organisation en nucléosomes. C’est en 1973 que Olins et Olins observent pour la première fois la fameuse structure en « collier de perles » qu’ils appellent alors les « ν bodies » (Olins & Olins 1974). En se basant sur ses dimensions, ils proposent un modèle où un « ν body » contiendrait deux histones de chaque type (H2A, H2B, H3 et H4) qui entoureraient l’ADN. Ils prédisent aussi que ces sphéroïdes pourraient s’associer entre eux pour former un repliement ou une hélice compacte. En 1975, le terme « nucléosome » apparaît (Oudet *et al.* 1975). Des études indépendantes mesurent la force des interactions entre les histones et suggèrent plutôt un cœur protéique compact autour duquel l’ADN serait enroulé (Kornberg 1974).

La découverte du nucléosome révolutionne la perception de la chromatine, l’ADN n’est plus recouvert par des histones mais enroulé autour d’un cœur globulaire formé par les histones et donc accessible pour interagir avec d’autres protéines nucléaires.

La résolution de la structure cristalline du nucléosome devient un enjeu majeur. En 1984, la structure est publiée avec une résolution de 7 Å (Richmond *et al.* 1984) puis affinée en 1997 à 2,8 Å (Fig. 1.9; Luger *et al.* 1997). En 2002, la structure de l’ADN nucléosomique a été précisée à 1,9 Å (Davey *et al.* 2002; Richmond & Davey 2003).

Ainsi, le premier niveau d’organisation de la chromatine a été caractérisé. La compaction de l’ADN au sein des nucléosomes diminue l’accessibilité de la chromatine qui

peut cependant être modulée. En effet, il avait été montré dès 1964 que des modifications post-traductionnelles (acétylations et méthylations) des histones pouvaient réguler la transcription (Allfrey *et al.* 1964). Près de quarante ans plus tard, la grande diversité des modifications post-traductionnelles des extrémités N-terminales des histones (voir Fig. 1.13) et leur association à des processus biologiques spécifiques ont conduit à proposer l'hypothèse d'un « langage » baptisé le « code des histones » en 2001 (Jenuwein & Allis 2001). Ce code serait « lu » par d'autres protéines ou complexes protéiques, capables de reconnaître et d'interpréter des profils de modification spécifiques. La notion d'épigénétique regroupe ces modifications qui ne sont pas codées directement dans la séquence d'ADN mais peuvent cependant être transmises pendant la division cellulaire.

Récemment, le développement des études à l'échelle du génome a apporté des précisions importantes sur l'organisation globale de la chromatine. Un nombre important de modifications des histones a été cartographié, ainsi que les sites de liaison de nombreux facteurs ayant un rôle structural ou fonctionnel dans l'organisation de la chromatine. L'ensemble de ces recherches a permis une grande avancée dans la connaissance de la chromatine. L'étude de son implication dans la régulation de l'expression génique est un domaine de recherche actuellement en grand essor.

1.1.2 Structure de la chromatine

Le matériel génétique ayant une taille considérable chez les eucaryotes pluri-cellulaires, une compaction très élevée de la chromatine est nécessaire pour la contenir de façon ordonnée dans un noyau. L'ADN d'une cellule de drosophile mesure linéairement ~ 6 cm (~ 2 m pour l'humain) et doit tenir dans un noyau de quelques microns, soit un facteur de compaction (volume de l'ADN libre/volume du noyau) de l'ordre de 10 000 (Holmes & Cozzarelli 2000).

Une partie des mécanismes et des structures mises en jeu est connue mais certaines étapes restent débattues (voir Fig. 1.10).

1.1.2.1 Le nucléosome

L'unité fondamentale de la chromatine eucaryote est le nucléosome (Olins & Olins 1974). Il se compose d'une particule de cœur composée de l'ADN enroulé autour d'une particule protéique, l'octamère d'histones, et d'un ADN de liaison qui relie entre eux les particules de cœur. L'ADN de la particule de cœur a une taille d'environ 147 paires de bases (pb) et s'enroule autour de l'octamère d'histones en faisant 1,65 tour orienté à gauche (pour revue : Felsenfeld & Groudine 2003). L'ADN de liaison a une taille variable suivant les espèces et le type cellulaire (de l'ordre de 50 pb en moyenne).



Figure 1.9 – Structure cristallographique du nucléosome (Luger *et al.* 1997). L'hélice d'ADN est représentée en marron et vert. Les histones sont représentées en bleu (H3), vert (H4), rose (H2B) et orange (H2A).

L'octamère d'histones est composé d'un tétramère H3-H4 associé à deux dimères H2A-H2B. Cette structure octamérique particulière a pu être caractérisée de manière très précise par cristallographie aux rayons X (Luger *et al.* 1997) (voir Fig. 1.9). Les histones sont de petites protéines extrêmement conservées au cours de l'évolution. Chaque histone comporte un domaine central globulaire et des extrémités moins structurées qui sortent du nucléosome et sont le siège de diverses modifications post-traductionnelles (voir §1.1.4).

Ainsi, le premier niveau de compaction de la chromatine est le nucléofilament de 11 nm, une structure dite en « collier de perles » constituée d'un enchaînement régulier de nucléosomes.

In vitro, les nucléosomes peuvent s'auto-assembler à partir d'histones natives ou recombinantes. Une technique classique consiste à incuber une molécule d'ADN nu avec des octamères d'histones en présence de fortes concentrations de sel et à réduire peu à peu la concentration saline par dialyse (Spadafora *et al.* 1978). Pour obtenir un espacement des nucléosomes plus proche de ce qui est observé dans la chromatine native, il est également possible d'assembler des nucléosomes en ajoutant des chaperons d'histones et des facteurs de remodelage de la chromatine (voir ci-dessous) recombinants à l'ADN et aux histones, en présence d'ATP (Adénosine Triphosphate; Ito *et al.* 1999). *In vivo*, l'assemblage des nucléosomes sur l'ADN se fait notamment pendant la réplication juste après le passage des fourches. Les chaperons d'histones sont des protéines qui lient les histones et participent à leur dynamique intracellulaire, leur stockage et leur déposition sur l'ADN (pour revue : Park & Luger 2006). Le complexe CAF1 (Chromatin Assembly Factor 1) dépose un nouveau tétramère H3/H4. Deux dimères H2A/H2B sont ensuite ajoutés sur le tétramère H3/H4. Il n'est pas clair si cette étape est effectuée par FACT (Facilitates Chromatin Transcription, facteur également impliqué dans l'élongation de la transcription, voir §1.3.3) et/ou Nap1 (Nucleosome assembly protein 1, facteur impliqué dans l'import de ces histones dans le noyau après leur synthèse; Haushalter & Kadonaga 2003) qui peuvent tous deux chaperonner le dimère H2A/H2B. Pendant cette étape, les histones (qui sont acétylées après leur synthèse) sont désacétylées après leur incorporation (Annunziato & Seale 1983). La régularité dans l'espacement des nucléosomes est obtenue suite à une étape maturation. Les complexes de remodelage de la chromatine CHRAC (Chromatin Accessibility Complex) et ACF (ATP-dependent Chromatin assembly and remodelling Factor) utilisent l'énergie libérée par l'hydrolyse de l'ATP pour déplacer les nucléosomes et les espacer régulièrement (Fyodorov *et al.* 2004, voir §1.1.5.3). L'assemblage des nucléosomes peut aussi avoir lieu de façon indépendante de la réplication. Par exemple, le chaperon d'histones HIRA (Histone Regulating Protein A) permet l'assemblage des nucléosomes *in vitro* en l'absence de réplication (Green *et al.* 2005) et facilite la mise en silence de certains promoteurs *in vivo* en favorisant la formation d'une chromatine répressive (Anderson *et al.* 2009).

1.1.2.2 La fibre de 30 nm

L'association d'une histone de liaison H1 à 20 pb de l'ADN de liaison fixe deux tours d'ADN en interagissant avec la surface de la particule de cœur. La particule obtenue constituée d'environ 167 pb d'ADN, d'un octamère d'histones et d'une histone H1 est appelée le chromatosome (Simpson 1978b). L'histone H1 est ajoutée sur le nucléosome par le chaperon NASP (Nuclear Autoantigenic Sperm Protein) (Finn *et al.* 2008).

Contrairement aux histones de cœur, l'inactivation de l'histone de liaison n'est pas létale (Shen *et al.* 1995), mais induit une altération globale de la structure de la chromatine, suggérant un important rôle de H1 dans le maintien de cette structure (pour revue : Khochbin 2001).

Le second niveau de compaction observé *in vitro* est la fibre de 30 nm qui se forme grâce aux interactions entre les nucléosomes via l'histone H1 qui interagit à la fois avec

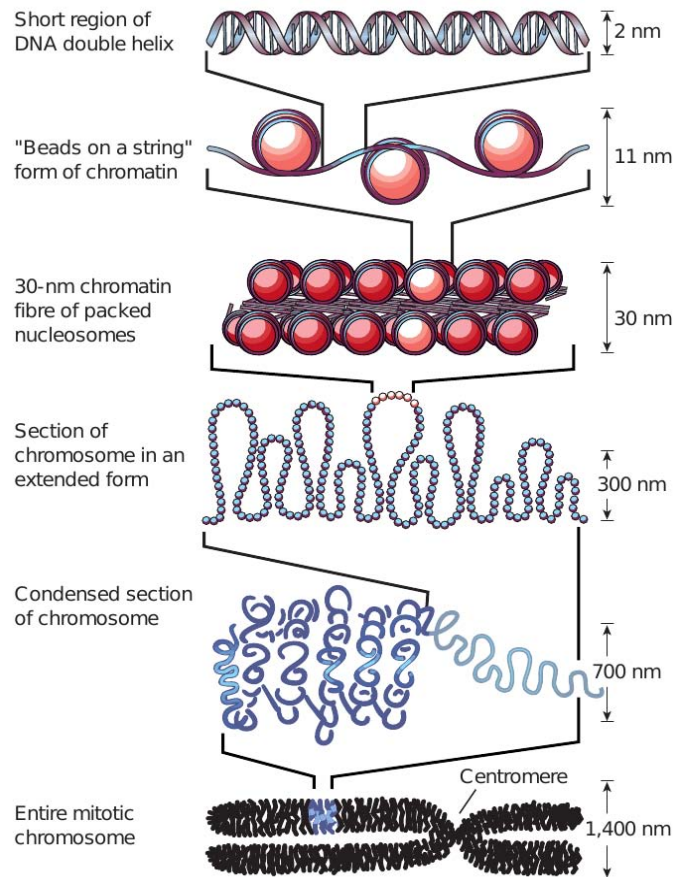


Figure 1.10 – Les différents niveaux de compaction de la chromatine (Felsenfeld & Groudine 2003).

l'ADN de liaison et avec le cœur des nucléosomes. Il y a plus de trente ans, Finch et Klug ont proposé que la chaîne de nucléosomes se repliait pour former une fibre de 30 nm en mettant en jeu les histones de liaison H1 et des ions Mg^{2+} (Finch & Klug 1976). En effet, la chromatine isolée observée au microscope électronique à transmission présente une structure de fibre de 30 nm de diamètre. Ils ont proposé une structure en solénoïde où deux nucléosomes consécutifs se retrouvent côte à côte dans une hélice simple. Plus tard, un second modèle d'hélice en zig-zag à deux départs a été suggéré, dans lequel un nucléosome est en contact avec ses seconds voisins et non avec le nucléosome consécutif (Woodcock *et al.* 1984). Diverses variations autour de ce modèle ont également été proposées.

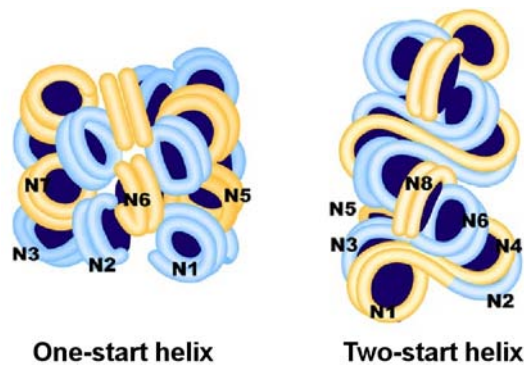


Figure 1.11 – Différents modèles de la fibre de 30 nm (Maeshima *et al.* 2010).

Diverses études, notamment de cryo-microscopie, ont apporté des arguments en faveur de l'un ou l'autre modèle (Dorigo *et al.* 2004; Robinson *et al.* 2006). Pour expliquer ces résultats contradictoires, le groupe de Rhodes a proposé que le choix du mode de compaction

dépendait de la longueur de l'ADN de liaison (Routh *et al.* 2008). Enfin, il a été montré récemment que les deux structures pourraient coexister au sein de la fibre (Grigoryev *et al.* 2009). Cependant, l'arrangement des nucléosomes *in vivo* pourrait ne pas être si régulier, avec des régions ordonnées en zig-zag alternant avec des régions localement désordonnées (voir Discussion §3).

1.1.2.3 Structures plus condensées

Finalement, il est proposé que plusieurs repliements successifs conduisent à des niveaux d'organisation supérieurs en domaines spécifiques dans le noyau. La fibre de 30 nm pourrait s'organiser en boucles pour atteindre un niveau de compaction maximum dans le chromosome métaphasique (pour revue : Woodcock & Dimitrov 2001; Felsenfeld & Grudin 2003). Cependant les preuves expérimentales de ce type d'organisation font toujours défaut à l'heure actuelle (voir Discussion §3). Après la réplication, les chromatides soeurs sont maintenues ensemble par le complexe cohesin jusqu'à la ségrégation des chromosomes en anaphase. À ce moment là, la kleisin (une sous-unité du complexe cohesin) est protéolysée par l'enzyme separase et les deux chromatides peuvent être séparées. Le complexe condensin permet la compaction des chromosomes au début de la mitose et la résolution des caténations vraisemblablement en recrutant la topoisomérase II (Hudson *et al.* 2003), ce qui permet la ségrégation (pour revue : Wood *et al.* 2010).

1.1.3 Domaines chromatiniens

Les génomes eucaryotes sont organisés en domaines comprenant des gènes individuels ou des groupes de gènes ayant des profils d'expression différents. Ces domaines sont définis de différentes façons. Cela peut-être leur statut d'expression (actif ou inactif), leur sensibilité à la DNase I (Désoxyribonucléase I), la présence de certaines marques sur les histones, la présence de variants d'histones, ou encore la distribution de certaines protéines non-histones.

La coloration des chromosomes en métaphase génère, le long de ces structures, une alternance de bandes transversales de colorations distinctes. Les bandes ainsi obtenues prennent le nom du type de traitement qui les génère (pour revue : Craig & Bickmore 1993).

Les principales bandes sont

- les bandes G et C qui sont répliquées tardivement en phase S (Synthèse ; phase de réplication de l'ADN) et correspondent à une chromatine condensée ou hétérochromatine.
- les bandes R qui sont répliquées précocement en phase S et correspondent à une chromatine plus ouverte ou euchromatine.

1.1.3.1 L'euchromatine

Les domaines de chromatine active ou euchromatine ont des caractéristiques propres. Riches en gènes actifs, ils sont relativement plus accessibles à diverses enzymes comme la DNase I, et leur profil de sédimentation suggère une structure plus ouverte avec un arrangement de nucléosomes peu régulier. Ces domaines peuvent être déplétés en histone H1. Les histones sont acétylées et riches en méthylation sur la lysine 4 de l'histone H3 (voir le §1.1.4 qui présente ces modifications ; Strahl & Allis 2000). L'euchromatine est enrichie en histones acétylées qui sont conservées entre l'interphase et la mitose, suggérant un rôle de marqueur fonctionnel de cette modification pouvant servir de mémoire de l'organisation de ces domaines au cours du cycle cellulaire.

1.1.3.2 L'hétérochromatine

L'hétérochromatine est la chromatine mise en silence par répression épigénétique. L'hétérochromatine est riche en séquences répétées et relativement pauvre en gènes. Cette chromatine silencieuse est stable et héréditaire et peut être visualisée par coloration de l'ADN. Elle reste en apparence condensée tout au long du cycle cellulaire (Fig. 1.3; Heitz 1928). Les histones de l'hétérochromatine sont hypoacétylées, pauvres en méthylation de la lysine 4 de l'histone H3 et enrichies en méthylation de la lysine 9 de l'histone H3 (voir §1.1.4; Strahl & Allis 2000). Des facteurs spécifiques lient la chromatine et entraînent sa condensation et l'inhibition de la transcription. Dans la plupart des cellules somatiques de la drosophile, 30 à 35% du génome est sous forme d'hétérochromatine (le chromosome Y, 40% du X, 25 % des chromosomes 2 et 3 et plus de la moitié du 4; Girton & Johansen 2008). L'ADN de l'hétérochromatine est organisée en un enchaînement régulier de nucléosomes et est relativement résistant à la digestion par les nucléases (Sun *et al.* 2001).

L'hétérochromatine constitutive est définie comme la chromatine qui reste condensée dans tous les types cellulaires. Elle est composée de séquences répétées et est située au niveau des centromères et des télomères des chromosomes. Il existe également quelques blocs isolés d'hétérochromatine constitutive dispersés dans le génome. Elle est relativement pauvre en gènes, bien qu'un certain nombre de gènes essentiels y soit localisés.

L'hétérochromatine facultative est définie comme de l'euchromatine qui est mise en silence à certains stades du développement ou dans certains types cellulaires (Grigoryev *et al.* 2006).

1.1.3.3 Organisation nucléaire de la chromatine

Pendant l'interphase, la chromatine s'organise dans le noyau. Les régions d'hétérochromatine se regroupent à la périphérie nucléaire, créant ainsi des foyers riches en protéines répressives, ce qui permet une grande stabilité de la mise en silence des gènes. L'euchromatine réside quant à elle dans la partie interne du nucléoplasme (Dillon 2004). Il a de plus été montré que les chromosomes n'occupent pas aléatoirement l'espace nucléaire, mais sont organisés en territoires distincts (Cremer *et al.* 1982; Cremer & Cremer 2010, Fig. 1.12).

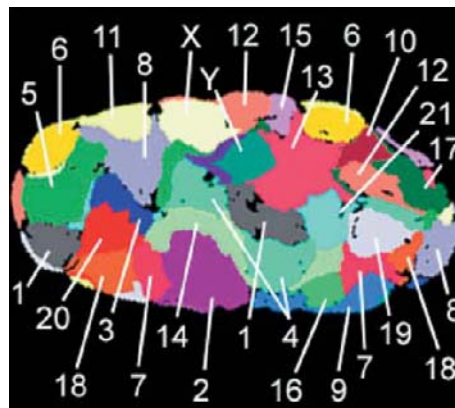


Figure 1.12 – Représentation en fausses couleurs des territoires chromosomiques d'un fibroblaste humain (Bolzer *et al.* 2005). Chaque région colorée représente un chromosome. Ce type de carte est obtenue par FISH (Fluorescent *in situ* hybridization) en 3D avec des sondes spécifiques de chaque chromosome (whole-chromosome painting).

Cette organisation nucléaire semble importante pour la régulation des gènes. En effet, quelques exemples montrent que la mise en silence ou l'activation d'un gène est accompagnée d'une relocalisation de ce gène dans le noyau (Dundr *et al.* 2007; Chuang *et al.* 2006). De plus, en déplaçant artificiellement un gène vers une autre région du noyau (comme la périphérie nucléaire) on observe un changement dans son expression (Finlan & Bickmore

2008; Reddy *et al.* 2008), ce qui suggère que la position d'un gène dans le noyau participe à sa régulation et n'est pas seulement une conséquence du niveau d'expression de ce gène.

1.1.4 Modifications de la chromatine

Dès les premières étapes de l'assemblage de la chromatine, la particule élémentaire peut être soumise à des variations au niveau de l'ADN (notamment par méthylation), ou au niveau des histones qui peuvent soit présenter différentes modifications post-traductionnelles, soit exister sous forme de variants (pour revue : Campos & Reinberg 2009).

Ces variations peuvent avoir des effets intrinsèques, extrinsèques ou dépendants d'un facteur effecteur. On parle d'un effet intrinsèque quand la modification altère la structure même du nucléosome. C'est notamment le cas des variants d'histones (voir §1.1.4.2) qui modifient la structure du coeur du nucléosome. Un effet extrinsèque décrit un effet direct sur les contacts inter-nucléosomiques et la structure de la chromatine. Enfin, les effets dépendants d'un effecteur sont les variations dépendantes de la liaison de facteurs spécifiques non-histones modifiant la chromatine. Ce sont ces effets qui sont à la base de l'hypothèse du code des histones.

1.1.4.1 Les modifications des queues des histones

Les queues terminales des histones sont assez courtes (moins de 40 acides aminés) et riches en résidus basiques. Elles sortent nettement de la structure du nucléosome et sont sujettes à de nombreuses modifications post-traductionnelles (Fig. 1.13). La liaison des effecteurs se fait de façon spécifique via des domaines particuliers comme le chromodomaine et le domaine PHD (Plant Homeo Domain) qui reconnaissent des résidus méthylés ou le bromodomaine qui reconnaît les résidus acétylés. Quelques exemples de modifications sont présentés ci-dessous.

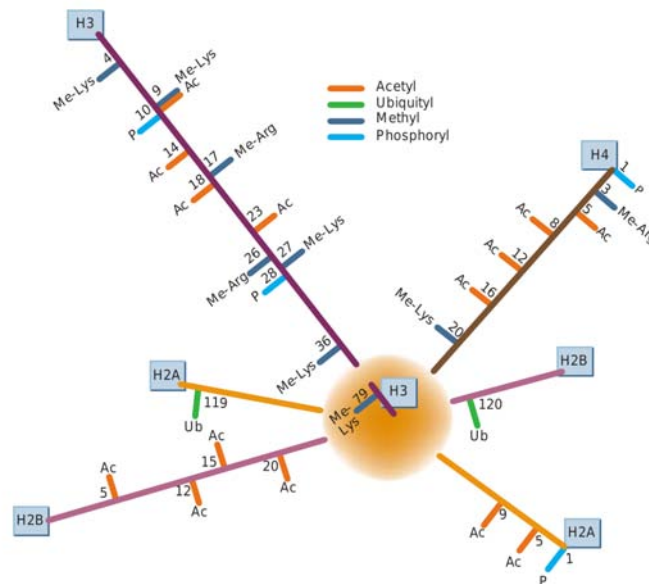


Figure 1.13 – Les principales modifications post-traductionnelles de la queue des histones (Felsenfeld & Groudine 2003).

La modification la mieux étudiée jusqu'à présent est l'acétylation des résidus lysines. Cette modification neutralise la charge positive des lysines et est suspectée de perturber les interactions histone-ADN. Bien que l'effet sur la structure même du nucléosome soit faible (Ausio & van Holde 1986), l'acétylation des histones induit un net changement dans la compaction de la chromatine. En effet, la chromatine acétylée est plus sensible à

la digestion par une endonucléase (Simpson 1978a), ce qui suggère une structure moins compacte. L'état d'acétylation résulte d'un équilibre entre deux activités antagonistes : l'activité histone-acétyltransférase (HAT) et l'activité histone-désacétylase (HDAC) (Suka *et al.* 2002). De nombreuses protéines jouant un rôle dans la régulation de la transcription ont une activité histone-acétyltransférase intrinsèque. De même, des histone-désacétylases ont été décrites comme faisant partie de complexes multiprotéiques de remodelage associés à la compaction de la chromatine. Un exemple bien étudié est l'acétylation de la lysine K16 de l'histone H4 qui modifie directement les contacts inter-nucléosomiques ce qui perturbe la structure chromatiniennne (nomenclature des acides aminés en annexe A.2). En effet, les résidus 16 à 20 de l'histone H4 interagissent avec deux zones acides des hélices C-terminales de l'histone H2A du nucléosome adjacent (Zhou *et al.* 2007). L'acétylation de H4K16 de 20 nucléosomes dans une fibre de chromatine reconstituée *in vitro* de 60 nucléosomes suffit à empêcher la compaction normalement observée en modifiant la salinité de la solution (Robinson *et al.* 2008). Cette modification est effectuée par le complexe MOF (Males absent On the First) dont la déficience entraîne la létalité chez l'embryon de souris (Thomas *et al.* 2008) et des défauts du cycle cellulaire et de l'apoptose dans des cellules en culture. Ces observations montrent l'importance de cette marque dans le maintien de la structure chromatiniennne. Chez la drosophile, MOF est indispensable à l'ouverture de la chromatine du chromosome X nécessaire à la compensation de dose chez le mâle (Hilfiker *et al.* 1997). En effet, l'unique chromosome X du mâle est surexprimé d'un facteur deux pour compenser le fait qu'il y a deux chromosomes X chez la femelle (Hamada *et al.* 2005). Cette surexpression met en jeu une ouverture globale de la chromatine qui permet la surexpression des gènes. En plus de modifier directement la structure de la chromatine, l'acétylation repousse certains facteurs comme le complexe ACF qui déplace les nucléosomes pour les espacer régulièrement à condition que les histones soient désacétylées (Shogren-Knaak *et al.* 2006).

La méthylation des histones est une marque très étudiée qui peut être présente sous trois formes : mono, bi ou tri-méthylation (pour revue : Campos & Reinberg 2009). Ces marques peuvent être associées à de la chromatine silencieuse (H3K9me3, H4K20me3 ou H3K27me3) ou active (H3K27me, H3K9me, H4K20me, H3K4me1/2/3 ou H3K36me3 ; Barski *et al.* 2007). La chromatine inactive présente les marques H3K9me3 et H3K20me3, la première étant reconnue par la protéine HP1 (Heterochromatin Protein 1) qui recrute les effecteurs de ces modifications et permet la propagation de la structure (voir §1.1.5.2). La mise en silence de certains gènes par les protéines Polycomb implique la déposition de la marque H3K27me3 (voir §1.1.5.4). Les régions transcrites présentent généralement des groupements méthyles sur H3K4 ainsi que la marque H3K36me3. Cependant, l'étude de la répartition de ces marques à l'échelle du génome montre qu'il existe de nombreuses exceptions (par exemple, H3K9me3 peut être associé à des gènes actifs ou H3K4me3 à des gènes mis en silence ; Barski *et al.* 2007). Ainsi, le code des histones est beaucoup plus complexe qu'initialement suggéré et la combinaison de toutes ces marques semble plus importante qu'une marque isolée. De plus, l'organisation des ces marques est différente au niveau des promoteurs et à l'intérieur des gènes. Ainsi, dans les gènes actifs, la marque H3K4me3 est présente uniquement au niveau des promoteurs alors que la marque H3K36me3 y est absente. Celle-ci est par contre présente à l'intérieur des gènes et sa densité augmente graduellement plus on avance le long de la région codante (Campos & Reinberg 2009).

L'ubiquitination de la lysine 119 de l'histone H2A est un exemple d'effet intrinsèque sur la structure du chromatosome. En effet, la désubiquitination de K119 favorise la phosphorylation de H1 et son décrochage de la chromatine (Zhu *et al.* 2007), ce qui facilite l'ouverture de celle-ci. Cette marque, déposée par PRC1 (Polycomb Repressive Complex 1) chez la drosophile (de Napoles *et al.* 2004), est associée majoritairement à de la chromatine silencieuse et la désubiquitination active l'initiation de la transcription (Nakagawa *et al.* 2008). De plus, cette marque empêche le désassemblage des histones par le complexe

FACT nécessaire à l'activité d'élongation de l'ARN polymérase II (pol II; [Zhou et al. 2008](#)). À l'inverse, l'ubiquitination de la lysine 120 de H2B est associée aux promoteurs actifs. Cette marque est reconnue par le complexe COMPASS (Complex of Proteins Associated with Set1) qui méthyle à son tour H3K4 et H3K79 ([Lee et al. 2007](#); [McGinty et al. 2008](#)) et facilite la transcription.

La phosphorylation est également une modification importante qui peut se trouver sur les sérines et les thréonines. Par exemple, la phosphorylation de la sérine 10 de l'histone H3 a été impliquée dans la condensation des chromosomes ainsi que la régulation transcriptionnelle d'un nombre important de gènes ([Nowak & Corces 2004](#)). Cette marque intervient notamment pour activer rapidement certains gènes après activation de la voie MAP-kinase (Mitogen-Activated Protein kinase) en réponse à des signaux extérieurs à la cellule.

1.1.4.2 Les variants d'histones

Contrairement aux histones H2B et H4, H2A et H3 possèdent de nombreux variants. L'incorporation de variants d'histones est importante pour des domaines spécifiques du génome et est associée à des fonctions très diverses. Ainsi CENP-A (Centromere Protein A), variant de l'histone H3, est associé aux régions centromériques. Le variant H2A.X intervient dans la formation de foyers contenant des facteurs de réparation dans des régions où des cassures d'ADN double brin ont eu lieu. Quant à H2A.Z, ce variant pourrait jouer un rôle en modifiant la structure de la chromatine afin de réguler la transcription.

La structure cristalline du nucléosome contenant H2A.Z montre une déstabilisation de l'interaction H3-H2A.Z et la formation d'une poche de liaison à un ion qui pourrait affecter la liaison de H1 ([Suto et al. 2000](#)). Cela implique un effet intrinsèque sur la structure du chromatosome et sur le repliement de la chromatine. Le rôle de H2A.Z reste cependant mal compris. Les analyses à l'échelle du génomes ont montré que H2A.Z est enrichi sur les nucléosomes -2, +1 et +2, mais déplété sur le -1 des gènes actifs et faciliterait l'initiation de la transcription ([Schones et al. 2008](#)).

Le variant H3.3 représente environ un quart des histones H3 chez la drosophile et est associé aux régions transcrites ([McKittrick et al. 2004](#)). Pour justifier cet enrichissement, il a été proposé que le variant H3.3 est intégré dans ces régions de façon indépendante de la réplication, mais dépendante de la transcription. Ainsi après le passage de l'ARN polymérase II, les nucléosomes sont réassemblés et le variant H3.3 est préféré à H3. L'invalidation du variant H3.3 n'a cependant pas de conséquences importantes sur la transcription, il est alors remplacé par l'histone H3 canonique ([Sakai et al. 2009](#)). Ces auteurs montrent cependant que ce variant est indispensable à une ségrégation méiotique normale.

1.1.4.3 La méthylation de l'ADN

Chez de nombreux organismes, une marque supplémentaire existe, à savoir la méthylation de l'ADN. Il s'agit de l'addition d'un groupement méthyle en position 5 des cytosines ou en position 6 des adénines. Chez les vertébrés, la majorité des sites méthylés sont des cytosines localisées au niveau des îlots CpG sur les promoteurs ou les premiers exons des gènes. La méthylation des CpG des promoteurs inhibe l'initiation de la transcription et permet la mise en silence des gènes notamment sur le chromosome X inactif ou sur des gènes soumis à l'empreinte parentale (pour revues : [Jones & Takai 2001](#); [Lan et al. 2010](#)).

Pendant longtemps, les chercheurs ont rapporté l'absence de cytosines méthylés chez la drosophile ([Urieli-Shoval et al. 1982](#)). Cette vision a été révisée depuis que des études ont mis en évidence, non seulement des résidus méthylés ([Gowher et al. 2000](#); [Lyko et al. 2000](#)), mais aussi l'existence d'une enzyme méthyltransférase d'ADN (Dnmt2 (DNA methyltransferase 2), [Kunert et al. 2003](#)). Cette marque est présente en assez grande quantité

dans les embryons mais rare dans les adultes où quelques séquences méthylées ont cependant été identifiées (Salzberg *et al.* 2004). Alors que chez les mammifères 2 à 10% des cytosines sont méthylées, seulement 0,1 à 0,6% de ces résidus sont méthylés chez la drosophile (Lyko *et al.* 2000). Contrairement aux vertébrés où ce sont les îlots CpG qui sont préférentiellement méthylés (Bird 2002), la machinerie de méthylation de l'ADN de la drosophile préfère les cytosines des motifs CpA et CpT (Kunert *et al.* 2003).

1.1.5 Dynamique de la chromatine

L'organisation de la chromatine doit répondre à deux obligations difficilement conciliables : le polymère doit être suffisamment compact pour être contenu dans le noyau tout en permettant que l'ADN soit accessible aux différentes machineries ayant besoin de lire le code génétique, telles que les machineries de transcription, de réparation ou de réplication. Ces obligations opposées peuvent s'accommoder si le niveau de compaction de la chromatine varie en fonction des besoins. Ainsi, les génomes eucaryotes sont organisés par des domaines plus ou moins compacts et donc plus ou moins accessibles qui doivent être régénérés pendant chaque cycle de réplication et dynamiquement modulés pendant les phases de transcription. Différents aspects de cette dynamique chromatiniennne sont présentés ci-dessous.

1.1.5.1 La PEV (Position-Effect Variegation) ou mosaïcisme par effet de position

Un effet de position désigne une dérégulation de l'expression d'un gène après son déplacement dans une autre région du génome où il est soumis à des éléments régulateurs différents de ceux de son emplacement d'origine. On parle de PEV (Position-Effect Variegation) quand le phénotype de cette dérégulation est bigarré, c'est à dire qu'il n'est pas homogène parmi toutes les cellules et l'effet ressemble à une mosaïque. Une inactivation a lieu quand des gènes de l'euchromatine, comme par exemple le gène *white* chez la drosophile, sont relocalisés à proximité de l'hétérochromatine. La PEV peut également être observée lorsque des gènes normalement situés dans ou près de l'hétérochromatine sont relocalisés à proximité de l'euchromatine, comme dans le cas du gène *cubitus interruptus* (Dubidin & Sidorov 1934; Locke & Tartof 1994; Demakova *et al.* 1998). Des études menées sur ces réarrangements chromosomiques suggèrent que c'est une propagation de la structure chromatiniennne adjacente qui induit les changements d'expression observés (pour revue : Girton & Johansen 2008).

Un des exemples le plus célèbre de PEV est le mutant *w^{m4h}* décrit pour la première fois par Muller en 1930 (Muller 1930). Il provient d'un réarrangement chromosomique induit par des traitements aux rayons X. Le gène *white* code pour un transporteur de pigment nécessaire à la couleur rouge normale des yeux de drosophile. L'inversion *w^{m4h}* déplace le gène *white* à proximité de l'hétérochromatine péricentromérique du chromosome X, alors que ce gène est normalement situé à l'extrémité distale de ce même chromosome (Schultz 1936). Cette relocalisation est associée à une mise en silence du gène *white* dans certaines cellules. Cette répression est héritée de façon clonale, elle traduit une régulation négative stochastique du gène *white* lors du développement de l'œil. Le phénotype alors observé est une coloration bigarrée de l'œil de la mouche avec des facettes (ou ommatidies) pâles quand le gène est partiellement éteint et des ommatidies rouges quand il est exprimé.

Le rôle de nombreuses protéines dans la formation ou le maintien de la structure hétérochromatiniennne a été mis en évidence grâce à leur capacité à modifier le phénotype de PEV notamment dans les lignées *white-mottled* (*w^m*) de drosophile (pour revues : Schotta *et al.* 2003; Girton & Johansen 2008; Schulze & Wallrath 2007)

Des mutations générées dans cette lignée ont permis de définir deux classes :

- les gènes dont la mutation réduit le phénotype de PEV : gènes suppresseurs de « variégation » (anglicisme couramment utilisé pour « mosaïcisme ») ou *Su(var)*,
- les gènes dont la mutation l'accroît : gènes activateurs de « variégation » ou *E(var)*.

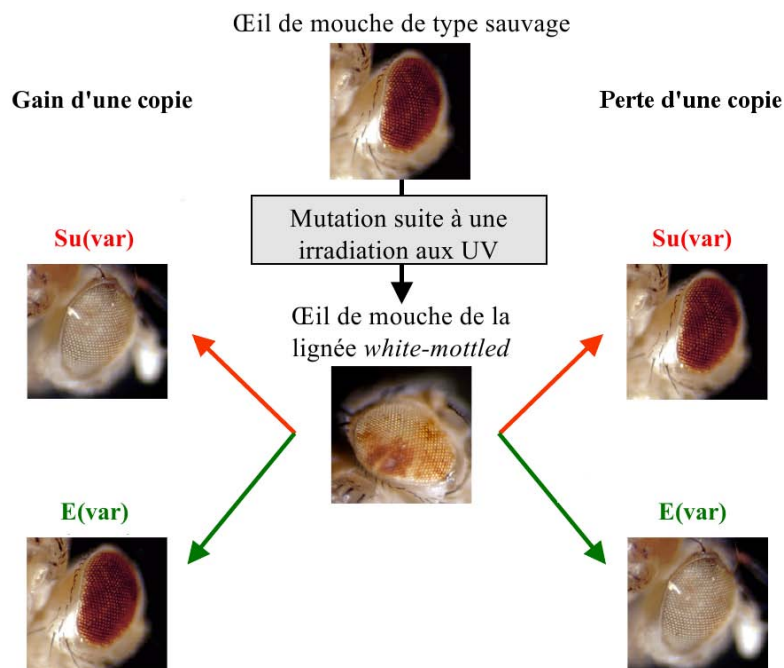


Figure 1.14 – Phénotypes des activateurs et des répresseurs de « variégation » dans la lignée *wm*^{4h}.

Certains gènes codant les protéines impliquées dans la formation de l'hétérochromatine sont des *Su(var)* qui présentent un phénotype d'haplo-suppresseur/triplo-amplificateur de PEV (Wustmann *et al.* 1989). Celui-ci correspond à une diminution de la PEV lorsque une copie du gène est éteinte, et à une augmentation de la PEV lorsqu'une copie supplémentaire est insérée (Fig. 1.14). Cette sensibilité au dosage montre l'importance de la balance stoechiométrique entre les différents facteurs de régulation de la structure chromatinienne. Parmi les gènes ainsi identifiés on peut citer par exemple *su(var)3-9* qui s'est avéré coder pour une méthyltransférase spécifique de la lysine 9 de l'histone H3 qui est conservée chez tous les eucaryotes (Rea *et al.* 2000). Un autre exemple ayant considérablement fait progresser la compréhension des mécanismes de formation de l'hétérochromatine est l'étude du gène *su(var)2-5* qui code pour la protéine HP1 -également conservée- qui reconnaît la méthylation de H3K9 et permet la propagation de la structure (Hall *et al.* 2002).

Parmi les gènes *E(var)*, on peut citer le gène *Trl* (Trithorax-like) codant pour la protéine GAF (GAGA Factor) dont la liaison aux promoteurs permet l'ouverture de la chromatine (voir §1.4.5 ; Farkas *et al.* 1994). Le gène *mod(mdg4)* a également été initialement décrit comme un *E(var)* (Weiler & Wakimoto 2002). Il code pour une protéine associée à la fonction insulatrice des éléments *gypsy* (voir §1.4.3). À l'inverse des gènes *Su(var)*, les gènes *E(var)* sont impliqués dans l'ouverture de la chromatine pour permettre l'expression génique (Fig. 1.15).

1.1.5.2 Formation de l'hétérochromatine

Les modèles classiques permettant d'expliquer les mécanismes moléculaires de la formation de l'hétérochromatine postulent l'existence de sites d'initiation et de terminaison (Locke *et al.* 1988).

La découverte d'un lien entre la voie d'interférence à l'ARN (RNAi) et la formation de l'hétérochromatine chez la levure *Schizosaccharomyces pombe* a permis de postuler la nature des sites d'initiation. En effet, il a été montré que des ARN double-brins (ARNdb)

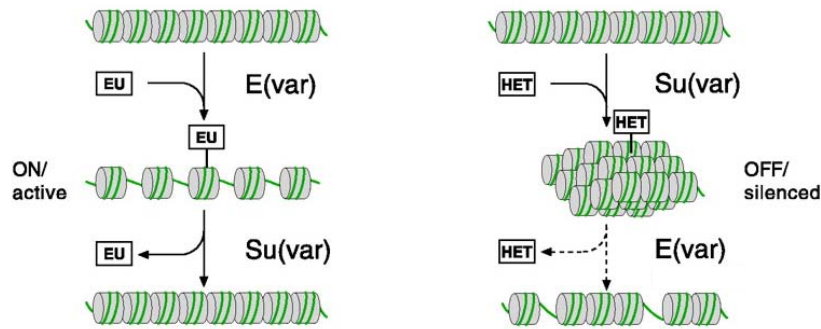


Figure 1.15 – Lien entre modificateurs de variégation et structure de la chromatine (adapté de [Jenuwein & Allis 2001](#)). Un *E(var)* a tendance à ouvrir la chromatine, ce qui facilite l'activation des gènes. Un *Su(var)* a tendance à compacter la chromatine, ce qui met en silence les gènes. EU : facteur euchromatinien ; HET : facteur hétérochromatinien.

sont générés par transcription bidirectionnelle de séquences péri-centromériques répétées. Ils sont alors maturés en petits ARN interférents (siRNA) par l'endonucléase Dicer1 ([Volpe et al. 2002](#)). Ces siRNA sont alors pris en charge par le complexe RITS (RNA-induced Initiation of Transcriptional gene Silencing; [Verdel et al. 2004](#)). Le complexe RITS est composé des protéines Chp1, Ago1 et Tas3. Ago1 lie les siRNA formés par Dicer et permet le ciblage en retour des répétitions centromériques, grâce à l'homologie de séquence des transcrits naissants. L'ARN polymérase II synthétise un transcrite qui sert de plateforme de recrutement de RITS, RDRC (RNA-Directed RNA polymerase Complex) et peut-être d'autres complexes nécessaires à l'assemblage de l'hétérochromatine (pour revue : [Grewal 2010](#)). Le complexe RDR permet l'amplification de la production des ARNdb au site d'initiation et participe ainsi à une boucle de rétro-contrôle positive. Le recrutement de RITS sur le site de transcription nécessite la méthyltransférase Clr4 (orthologue de *Su(var)3-9*) qui méthyle H3K9 ([Noma et al. 2004](#)). Cette marque sert également de site de liaison de Swi6 (orthologue de HP1) et permet la propagation de l'hétérochromatine.

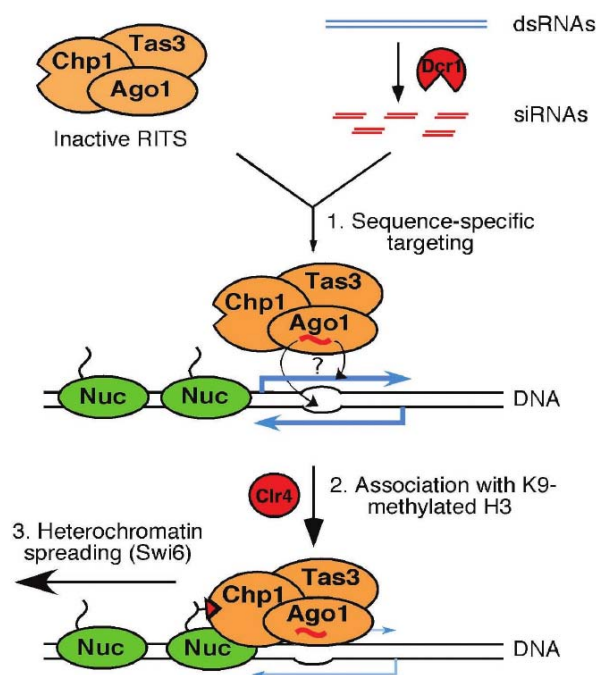


Figure 1.16 – Modèle d'initiation de la formation de l'hétérochromatine ([Verdel et al. 2004](#), voir texte).

Il est important de noter que la nucléation de l'hétérochromatine chez *S. pombe* peut

aussi avoir lieu de façon indépendante de la machinerie RNAi. Par exemple, en l'absence de l'interférence à l'ARN, les protéines ATF/CREB (cAMP Response Element Binding protein / Activating Transcription Factor) peuvent être activées par un stress et permettre la formation de l'hétérochromatine au locus de choix du type sexuel (Jia *et al.* 2004). De même, la protéine Taz1 peut recruter Swi6 aux télomères de façon indépendante de la voie RNAi (Kanoh *et al.* 2005).

La conservation de ce mécanisme d'initiation impliquant la machinerie RNAi chez les eucaryotes pluricellulaires est sujette à débat. Tout d'abord, les mouches et les vertébrés ne semblent pas avoir de gènes codant pour des protéines orthologues des RDR, ce qui indique que cette étape d'amplification n'existe probablement pas chez ces organismes. Chez la drosophile, il a été montré que des protéines de la voie RNAi étaient nécessaires à la mise en silence de répétitions de transgènes *white* insérés dans l'hétérochromatine péricentromérique (Pal-Bhadra *et al.* 2004). En effet des mutations dans les gènes *piwi*, *aubergine* et *homeless (spindle E)* induisent une réexpression de ces transgènes. Ces mutations s'accompagnent d'une réduction de la marque H3K9me3 et une délocalisation des protéines HP1 et HP2 (Heterochromatin Protein 2, un facteur interagissant avec HP1 qui est également nécessaire au maintien de la structure hétérochromatinienne; Shaffer *et al.* 2006). Ces travaux sont contredits par une autre étude qui montre au contraire que les protéines argonautes Ago2, *piwi* et *aubergine* ne sont pas nécessaires à la mise en silence des regroupements des gènes codant les piRNA (Piwi-interacting RNA) qui sont hétérochromatiniens et liés par HP1 (Moshkovich & Lei 2010). À l'inverse, les mutations de ces protéines entraînent une augmentation de la liaison de HP1 et de la mise en silence de ce locus. Pour réconcilier ces études, on peut suggérer qu'il existe différents mécanismes de nucléation de l'hétérochromatine suivant les loci considérés.

Les étapes suivantes sont bien décrites et conservées de la levure à l'humain. La lysine 9 de l'histone H3 est d'abord désacétylée par une HDAC (Histone désacétylase). Il s'agit de HDAC1 chez la drosophile. Une fois désacétylée, la lysine 9 est méthylée par Su(var)3-9 (Czermin *et al.* 2001). Cette histone modifiée est alors reconnue par le chromodomaine de HP1 (Bannister *et al.* 2001; Lachner *et al.* 2001) qui s'oligomérisent. HP1 recrute à son tour Su(var)3-9 qui méthyle l'histone du nucléosome voisin, permettant ainsi la propagation de la structure. HP1 recrute également la méthyltransférase Su(var)4-20 qui triméthyle la lysine 20 de H4 (Kourmouli *et al.* 2004; Schotta *et al.* 2004).

Chez les vertébrés, les protéines Su(Var)3-9 recrutent également des méthyltransférases à ADN qui méthylent les CpG, assurant un niveau supplémentaire de répression. Chez la drosophile, quand la méthylation de l'ADN a été découverte, elle n'a d'abord pas été associée à la mise en silence des loci, contrairement à ce qui est observé chez les mammifères. Au contraire, il a été montré que cette marque est présente dans des régions codantes transcrites (Pour revue : Mandrioli & Borsatti 2006). Il a été proposé que cette marque puisse être importante pour éviter l'initiation de la transcription à des sites aberrants dans les régions codantes (Mandrioli 2004). Il a cependant été montré récemment que cette marque peut être déposée via SETDB1 (SET Domain Bifurcated 1) et participe à la mise en silence de gènes et de rétrotransposons (Gou *et al.* 2010). SETDB1 tri-méthyle H3-K9 et reconnaît les motifs CpA méthylés. La marque H3K9 tri-méthylée est alors reconnue par Dnmt2 et HP1, ce qui permet la propagation de la structure. Les motifs CpA méthylés sont également reconnus par Su(var)4-20 qui triméthyle H4K20 (Phalke *et al.* 2009). Ainsi la méthylation de l'ADN semble avoir un rôle conservé entre la drosophile et l'humain, bien que ce mécanisme soit nettement moins largement utilisé chez la première.

Cet enchaînement d'étapes permet une propagation linéaire de la structure hétérochromatinienne. Une question importante est celle des mécanismes permettant l'arrêt de cette propagation pour protéger les régions qui doivent rester euchromatiniennes. Les insulateurs sont des éléments d'ADN qui semblent de bons candidats car ils ont la propriété d'arrêter cette propagation dans des expériences faisant intervenir des transgènes (voir

§1.4). En effet, lorsqu'un transgène est inséré artificiellement dans l'hétérochromatine il est mis en silence. Par contre si un insulateur est ajouté de chaque côté de ce transgène, celui-ci devient insensible à son environnement et est à nouveau exprimé (Kellum & Schedl 1991). Ce rôle des insulateurs dans un contexte sauvage n'est cependant pas démontré et d'autres modèles permettant d'expliquer l'arrêt de la propagation sont également proposés dans la littérature et seront discutés plus loin (voir §3).

L'étude de la PEV est un bon modèle pour adresser ces questions de propagation des structures chromatiniennes. En effet, après relocalisation d'un domaine euchromatinien près d'une région hétérochromatinienne, les gènes proches de la cassure sont moins exprimés. Cela suggère que l'hétérochromatine s'est étendue et a mis en silence ces gènes. Il a été montré que les gènes les plus proches de la cassure sont les plus affectés et que l'effet diminue en s'éloignant (Lewis 1950). L'étude de la redistribution de la marque H3K9me2 suggère la formation d'un gradient d'hétérochromatine en accord avec ce modèle de propagation linéaire (Rudolph *et al.* 2007). Ces données sont cependant controversées. En effet, il a été montré qu'un gène peut être affecté par ce type d'inversion sans qu'un gène plus proche de la cassure ne le soit. C'est notamment le cas du gène *roughest* (*rst*) qui est affecté plus fréquemment que le gène *white* (*w*) dans la lignée w^{mMc} (Talbert & Henikoff 2000, Fig. 1.17). Dans cette lignée, les yeux peuvent en effet présenter le phénotype d'oeil rugueux (*rst*) alors que la plupart des cellules ont une couleur rouge normale (w^+). Pour expliquer cette observation, il a été proposé qu'il se forme des boucles qui mettent en contact des répétitions hétérochromatiniennes avec certains loci euchromatiniens mais pas avec d'autres qui auraient moins d'affinité pour ce type de chromatine.

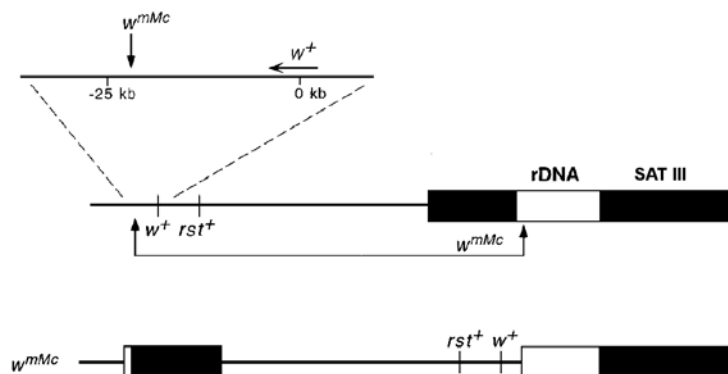


Figure 1.17 – Schéma de l'inversion w^{mMc} (Talbert & Henikoff 2000). Dans cette lignée, bien que le gène *white* soit plus proche de l'hétérochromatine que le gène *roughest*, il arrive fréquemment que ce dernier soit affecté (aspect rugueux de l'oeil) sans que le premier soit touché (ommatidie de couleur rouge normale). rDNA : région formée de répétitions codants les ARN ribosomiques ; SAT III : séquences satellites de type III.

Une limite de cette étude est soulignée par les auteurs eux-mêmes. En effet, on ne peut pas prouver que les effets observés sur l'expression de *rst* et *w* reflètent ce qu'il se passe dans une cellule donnée. Le phénotype de la dérégulation de *rst* est une déformation de l'oeil par inclusion de cellules surnuméraires et il n'est pas nécessaire que toutes les cellules soient *rst* pour observer ce phénotype. Ainsi un mélange de cellules $w^+ rst^+$ et de cellules $w rst$ pourrait produire le même phénotype. Il n'y a donc pas forcément de cellules où le gène le plus proche de la cassure (*w*) n'est pas affecté alors que le gène plus éloigné (*rst*) l'est. Seule une étude mesurant l'expression dans la même cellule pourrait lever cette ambiguïté. Ce type de mise en silence par formation de boucles n'est cependant pas exclu

et sera discuté dans ce manuscrit (voir §3).

1.1.5.3 Les complexes de remodelage de la chromatine

Les différentes machineries cellulaires impliquées dans la transcription, la réplication ou la réparation ont besoin de pouvoir accéder à l'information codée dans l'ADN. Comme la plus grande partie du génome des métazoaires est sous forme condensée et inaccessible, la formation et le maintien d'un état ouvert de la chromatine est un processus actif indispensable. Celui-ci requiert des éléments d'ADN spécifiques qui recrutent des complexes de remodelage et de modification de la chromatine.

L'étude des facteurs qui favorisent un espacement régulier des nucléosomes pendant l'assemblage de la chromatine a permis d'identifier plusieurs complexes multiprotéiques capables au moins *in vitro* de faire glisser des nucléosomes sur l'ADN (pour revue : [Clapier & Cairns 2009](#)). Les facteurs de remodelage sont impliqués dans le déplacement ou la restructuration des nucléosomes (Fig. 1.18). Il existe 4 familles de complexes de remodelage de la chromatine qui sont les familles SWI/SNF (Switching defective / Sucrose Non Fermenting), ISWI (Imitation Switch), CHD (Chromo-Helicase/ATPase DNA binding) et INO80 (Inositol requiring). Ces complexes utilisent l'énergie libérée par l'hydrolyse de l'ATP pour modifier l'interaction entre les histones et l'ADN. Elles ont en commun un domaine ATPase coupé en deux parties, le domaine DExx et le domaine HELICc (voir Fig. 1.19).

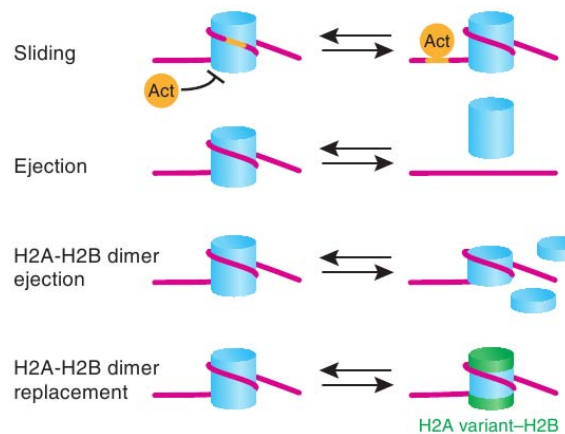


Figure 1.18 – Actions des remodeleurs de la chromatine ([Cairns 2007](#)). Les remodeleurs peuvent faire glisser un nucléosome, l'éjecter entièrement ou partiellement (dimère H2A/H2B), ou échanger un variant d'histone par un autre.

La famille SWI/SNF est caractérisée par la présence d'un domaine HSA (helicase-SANT) et d'un bromodomaine (pour revue : [Mohrmann & Verrijzer 2005](#)). Ces complexes peuvent faire glisser ou éjecter les nucléosomes dans divers processus mais a priori pas pendant l'assemblage de la chromatine. Parmi eux se trouvent les complexes SWI/SNF chez l'humain et Brahma chez la drosophile.

La famille ISWI est caractérisée par les domaines SANT et SLIDE (SANT-like ISWI) en C-terminal (pour revue : [Corona & Tamkun 2004](#)). Elle contient notamment les complexes CHRAC, ACF et NURF (Nucleosome Remodeling Factor). Les complexes ACF et CHRAC ont pour rôle d'optimiser l'espacement des nucléosomes pour favoriser l'assemblage de la chromatine et réprimer la transcription. Il a été montré chez la drosophile que l'ATPase ISWI est nécessaire au recrutement de l'histone de liaison H1 et que son invalidation entraîne une décondensation de la chromatine ([Corona et al. 2007](#)). Le complexe NURF au contraire peut désorganiser l'espacement des nucléosomes et favoriser la transcription par la pol II.

La famille CHD est caractérisée par 2 chromodomains organisés en tandem à l'extrémité N-terminale de la sous-unité catalytique (pour revue : [Marfella & Imbalzano 2007](#)). Certains remodeleurs CHD peuvent faire glisser ou éjecter les nucléosomes et favoriser la transcription. D'autres comme le complexe Mi-2/NuRD (Nucleosome Remodeling Histone Deacetylase) ont au contraire un rôle répressif vis à vis de la transcription.

La famille INO80 a la particularité de présenter une longue insertion entre les deux parties du domaine ATPase (pour revue : [Bao & Shen 2007](#)). Les complexes INO80 ont diverses fonctions, ils peuvent notamment favoriser la transcription et ont un rôle dans la réparation de l'ADN. Le complexe SWR1 (SWI/SNF-related 1) a la capacité d'échanger un dimère H2A-H2B par le variant H2A.Z-H2B.

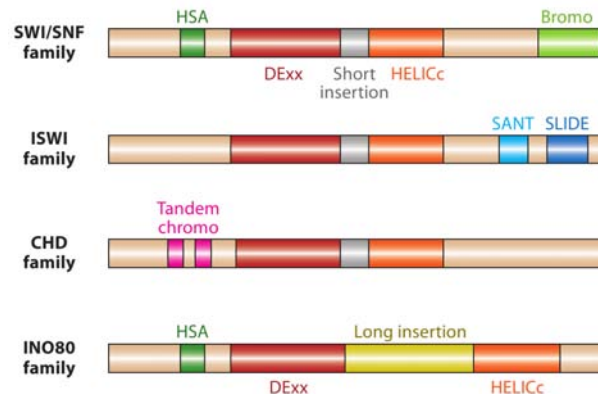


Figure 1.19 – Structure des différentes familles de remodeleurs de la chromatine ([Clapier & Cairns 2009](#)).

Les complexes de remodelage de la chromatine et de modifications des histones fonctionnent en liaison étroite. La compensation de dose chez la drosophile est un exemple d'interaction fonctionnelle entre ces deux types de complexes. Elle nécessite une augmentation de deux fois du niveau de transcription du chromosome X du mâle. Cette augmentation se fait grâce à une décompaction globale du chromosome notamment par l'hyperacétylation de la lysine 16 de l'histone H4 par MOF ([Rea et al. 2007](#)). Non seulement cette marque inhibe la formation d'une fibre compacte, mais elle diminue également l'activité de ISWI qui participe au recrutement de H1. Il y a cependant un équilibre entre l'ouverture de la chromatine par H4K16Ac et sa fermeture par ISWI. En effet, une invalidation du complexe ISWI entraîne une décondensation encore plus forte du chromosome X du mâle ([Corona et al. 2007](#)).

Ces complexes de remodelage de la chromatine interviennent dans tous les processus nécessitant un accès à l'ADN tels que la réparation et la réplication. Ils ont également un important rôle de régulation de la transcription des gènes (voir §1.3).

1.1.5.4 Polycomb et Trithorax

Les protéines Polycomb et Trithorax sont des régulateurs clés du développement de la drosophile (pour revues : [Kerppola 2009](#); [Schuettengruber et al. 2007](#)). Ces protéines répriment ou activent, respectivement, un nombre important de gènes en modifiant la chromatine. Elles ne sont pas responsables de l'initiation de l'état transcriptionnel de ces gènes, mais du maintien de cet état à travers les divisions cellulaires. Ces protéines reconnaissent des éléments d'ADN spécifiques, les séquences PRE (PcG Response Element) et TRE (TrxG Response Element). Le locus le plus étudié est le complexe *bithorax* qui contient trois gènes clés du développement spatio-temporel de la drosophile : *Abdominal A* et *B* (*abd-A* et *Abd-B*) et *Ultrabithorax* (*Ubx*) dont l'expression est finement régulée. Une fois qu'un profil d'expression de ces gènes est établi dans un type cellulaire, les protéines Polycomb et Trithorax assurent la transmission de ce profil au cours des divisions

cellulaires.

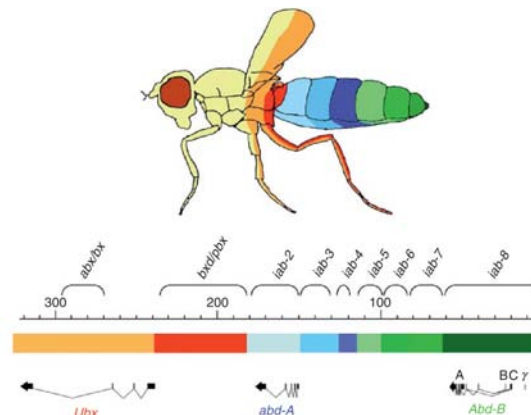


Figure 1.20 – Organisation du locus *bithorax* (Maeda & Karch 2009). Ce locus comprend trois gènes clés du développement de la drosophile. Les régions contrôlées par les différents éléments régulateurs sont représentées en couleur.

1.1.5.5 Activation par les protéines Trithorax

Les protéines Trithorax appartiennent à différentes classes. La première comprend des protéines possédant un domaine SET dont la protéine Trx qui lie l'élément *Fab-7* et maintient l'activation d'*Ubx* au sein du complexe *bithorax* (Dejardin & Cavalli 2005). Les gènes *Ash1* (*absent, small or homeotic discs 1*) et *trx* (*trithorax*) encodent des méthyltransférases qui triméthylent H3K4, ce qui est une marque de chromatine transcriptionnellement active (Byrd & Shearn 2003). Cette modification est importante car elle empêche la mise en silence de ces régions par les protéines Polycomb (Klymenko & Muller 2004).

La seconde classe inclut des composants des complexes de remodelage de la chromatine SWI/SNF et NURF (§1.1.5.3) dont le recrutement est facilité en présence de la marque H3K4me3 (Wysocka *et al.* 2006). Ces complexes facilitent le remodelage des nucléosomes aux promoteurs et le long des gènes pour permettre la liaison de l'ARN polymérase II ainsi que l'élongation (voir §1.3). Ainsi, dans ce groupe de protéines, des modificateurs d'histones et des complexes de remodelage collaborent pour maintenir l'activation des gènes.

1.1.5.6 Répression par les protéines Polycomb

Sur les chromosomes polytènes, les protéines Polycomb sont localisées sur des centaines de sites dans l'euchromatine qui ne se superposent pas avec les sites de HP1 (Schulze & Wallrath 2007), ce qui montre que la répression par ces protéines fonctionne indépendamment de la mise en silence par HP1 et H3K9me3 présentée précédemment (voir §1.1.5.2). Les régions mises en silence par les Polycomb sont un exemple typique d'hétérochromatine facultative.

Les protéines PcG appartiennent à 3 complexes. Le complexe PRC2 est responsable de la di- et tri-méthylation de la lysine 27 de l'histone H3 qui est une marque reconnue par le complexe PRC1 dont la liaison favorise la compaction de la chromatine par un mécanisme qui reste à déterminer (Cao & Zhang 2004). De plus, PRC1 est responsable de la monoubiquitination de la lysine 119 de H2A qui corrèle, comme H3K27me3, avec les régions transcriptionnellement réprimées. Le troisième complexe est PhoRC qui possède un domaine de liaison à l'ADN et semble important pour cibler le recrutement de PRC2 sur les PRE (Wang *et al.* 2004).

Ces protéines collaborent pour maintenir un état de la chromatine répressif à la transcription qui est hérité pendant de nombreux cycles cellulaires.

1.2 La transcription des ARN messagers

La transcription est la première étape de l'expression des gènes. Sa régulation est finement contrôlée et met en œuvre de nombreux facteurs. Chez les eucaryotes, la transcription nucléaire est assurée par les ARN polymérases I, II et III (principalement). L'ARN polymérase I est responsable de la transcription du précurseur des ARN ribosomiques (ARNr). L'ARN polymérase II transcrit majoritairement les ARN messagers (ARNm), ainsi que certains ARN non-codants. Le nombre d'ARN non-codants connus s'est considérablement accru ces dernières années, notamment grâce à l'émergence des techniques de séquençage à haut débit. Enfin, l'ARN polymérase III synthétise les ARN de transfert (ARNt), l'ARN ribosomique 5S ainsi que certains petits ARN non-codants.

Le cycle de transcription par les ARN polymérases peut être divisé en trois étapes principales : l'initiation, au cours de laquelle l'ARN polymérase et ses facteurs d'initiation sont recrutés et permettent la formation de la première liaison phosphodiester, l'élongation pendant laquelle la polymérase synthétise la molécule d'ARN, et enfin la terminaison au cours de laquelle la polymérase se décroche et l'ARN est libéré.

1.2.1 Structure de la polymérase II

L'ARN polymérase II (pol II) est l'unité catalytique de la synthèse d'ARN. Elle est composée de 12 sous-unités nommées Rpb1 (RNA polymerase B 1, la plus grande) à Rpb12 (la plus petite) et conservées de la levure à l'homme. Les sous-unités Rpb1, Rpb2 et Rpb3 sont homologues à des sous-unités des ARN polymérases I et III, mais également aux sous-unités β' , β et α respectivement, de l'ARN polymérase d'*E. coli*.

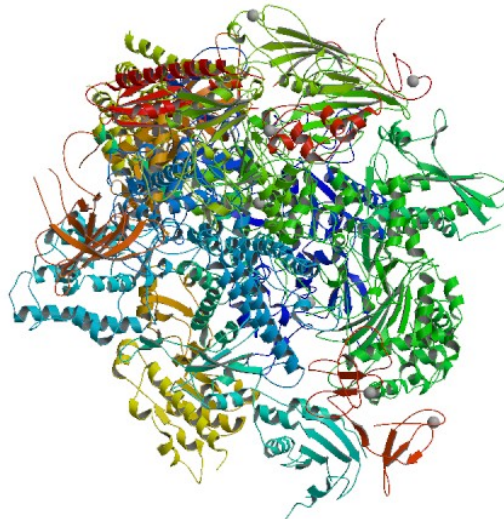


Figure 1.21 – Structure cristallographique de la polymérase II (<http://dx.doi.org/10.2210/pdb3fki/pdb>; Meyer *et al.* 2009). Les douze sous-unités sont représentées par des couleurs différentes.

La plus grosse sous-unité de pol II eucaryote, Rpb1, possède un domaine remarquable à son extrémité C-terminale (CTD, pour carboxy-terminal domain). Ce domaine consiste en une répétition de l'heptapeptide Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7 (Corden *et al.* 1985). La séquence de ce peptide a été extrêmement conservée au cours de l'évolution, mais le nombre de répétitions varie d'une espèce à l'autre (26 fois chez la levure *Saccharomyces cerevisiae* (*S. cerevisiae*), 44 fois chez la drosophile et 52 fois chez les mammifères). Le CTD est essentiel pour l'activité de l'ARN pol II *in vivo* et sa délétion est létale chez la levure, la drosophile et la souris (Allison *et al.* 1988). L'heptapeptide répété du CTD comporte 5

cibles possibles de phosphorylation et l'état de phosphorylation module l'activité de l'ARN pol II pendant les différentes étapes de la transcription.

1.2.2 Formation du complexe de préinitiation

La première étape est la formation du complexe de préinitiation (PIC) au niveau du promoteur. Le PIC est le complexe minimal nécessaire pour l'initiation de la transcription *in vitro*. Il est composé de l'ARN polymérase II ainsi que des facteurs généraux de transcription TFIIB, TFIID, TFIIE, TFIIF et TFIIH (Reese 2003). TFIIA est un sixième facteur non indispensable *in vitro* qui augmente l'efficacité de transcription. Les facteurs généraux de transcription interviennent dans la reconnaissance du promoteur, le recrutement de l'ARN polymérase II et l'ouverture de la bulle de transcription (Fig. 1.22).

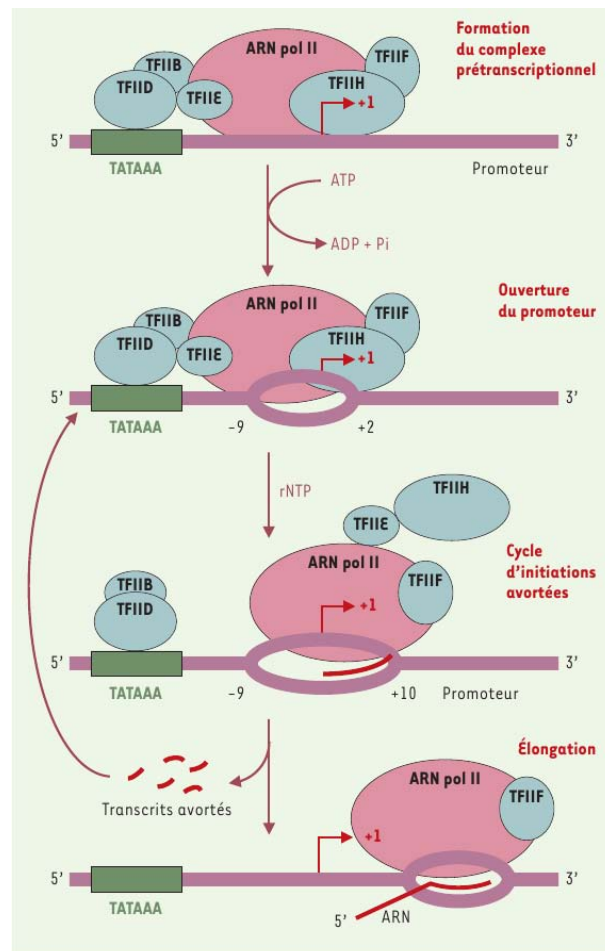


Figure 1.22 – Initiation de la transcription et facteurs généraux de transcription (Langelier *et al.* 2002).

Le promoteur correspond à une région non transcrite de l'ADN en amont du début de la région transcrite. Chez les eucaryotes, le promoteur contient un ou plusieurs motifs caractéristiques, dont le plus connu est la boîte TATA (découverte en 1979, aussi appelée la boîte Goldberg-Hogness du nom de ses découvreurs ; Breathnach & Chambon 1981). Il s'agit d'une séquence TATAWAWR (avec des variations possibles), située entre 25 et 30 paires de bases en amont du site de démarrage de la transcription (TSS). La boîte TATA et la protéine TBP qui la reconnaît sont conservées des archées aux humains. De nombreux promoteurs sont cependant dépourvus de boîte TATA (Yang *et al.* 2007). D'autres éléments tels que le DPE (Downstream Promoter Element), les éléments de reconnaissance par TFIIB (BRE^u et BRE^d) ou l'élément initiateur (Inr) ont été identifiés (voir Fig. 1.23 ; pour revue : Juven-Gershon & Kadonaga 2010).

L'élément Inr est situé autour du TSS et est constitué d'une séquence riche en pyrimidines moins conservée que la boîte TATA (YYANWYY chez l'humain et TCAKTY chez la drosophile). L'Inr est probablement le motif le plus fréquent dans les promoteurs. Il est reconnu par TFIID.

La séquence BRE (TFIIB Responsive Element), riche en guanines a d'abord été identifiée comme un élément reconnu spécifiquement par le facteur général de transcription TFIIB, situé en amont de la boîte TATA dans une partie des promoteurs en contenant (Lagrange *et al.* 1998). Un second site de liaison de TFIIB a été découvert juste en aval de la boîte TATA (Deng & Roberts 2006) et a été nommé BRE^d, le BRE original étant renommé BRE^u. Ces séquences fonctionnent avec la boîte TATA et peuvent soit augmenter soit diminuer le niveau de transcription basale (Lagrange *et al.* 1998; Evans *et al.* 2001; Deng & Roberts 2006).

La séquence DPE est une séquence faiblement conservée et située à environ trente paires de bases à l'intérieur des gènes (Burke & Kadonaga 1996). Elle est reconnue par TFIID qui lie de façon coopérative les motifs Inr et DPE. La distance entre ces deux éléments est cruciale pour l'activité transcriptionnelle des promoteurs dépendant d'un DPE (Kutach & Kadonaga 2000). Les promoteurs DPE ne contiennent généralement pas de boîte TATA, bien que des exceptions existent.

Le MTE (Motif Ten Element) a été décrit en 2002 suite à une étude visant à chercher de nouveaux motifs sur-représentés dans les régions promotrices chez la drosophile (Ohler *et al.* 2002). Le MTE est situé juste en amont du DPE et fonctionne avec l'Inr. De plus, il a un effet synergique avec le motif DPE mais aussi avec la boîte TATA (Lim *et al.* 2004). Le MTE est vraisemblablement reconnu par les sous-unités TAF6 (TBP-Associated Factor 6) et TAF9 de TFIID (Theisen *et al.* 2010).

Enfin, le DCE (Downstream Core Element) a été décrit assez récemment. Il est composé de 3 sous-éléments qui sont situés en aval du TSS (Lee *et al.* 2005). Il est reconnu par la sous-unité TAF1 de TFIID.

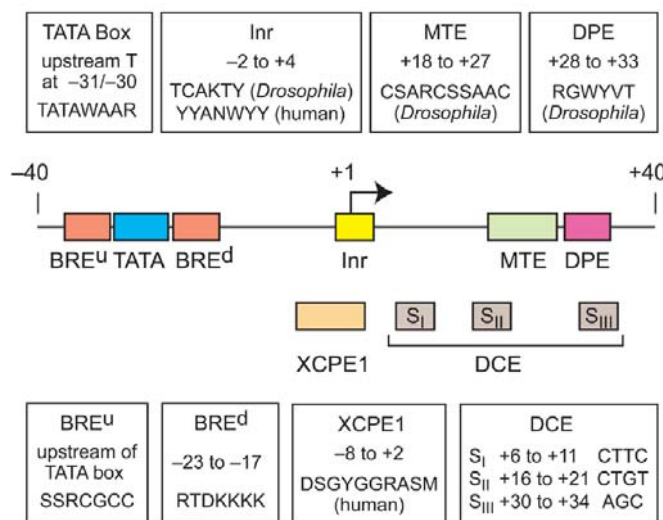


Figure 1.23 – Localisation et séquences consensus des boîtes caractéristiques des promoteurs (Juven-Gershon & Kadonaga 2010). Abréviations : voir code IUPAC en annexe A.1.

Diverses études à des échelles variables ont été conduites pour déterminer la proportion de gènes possédant tel ou tel motif. Les résultats de ces études sont assez différents et dépendent de la méthode bioinformatique utilisée. Il est encore aujourd'hui difficile de se faire une idée précise de la fréquence de ces boîtes. Chez la drosophile, une des premières études conduite sur 200 gènes révélait qu'environ la moitié des gènes contiennent une boîte TATA, et 40% des gènes contiennent un DPE dont 14% contenant les 2 motifs (Kutach &

Kadonaga 2000). Deux ans plus tard, une étude sur 2000 gènes arrivait à 19% contenant une boîte TATA, 16% un DPE et 26% une séquence Inr (Ohler *et al.* 2002). En 2006 une étude sur ~3400 promoteurs décrit que 16% d'entre eux contiennent une boîte TATA, 22% un DPE, 66% un élément Inr et 10% un MTE (Gershenson *et al.* 2006) alors qu'une seconde étude sur 10 000 promoteurs conclut que seuls 5% contiennent une boîte TATA, 14% un Inr et 2% un DPE (FitzGerald *et al.* 2006).

L'existence d'une grande variété de types de promoteurs suggère que le rôle de ces motifs n'est pas seulement de spécifier le site de démarrage de la transcription. D'une part, le niveau basal d'expression peut-être modulé par la présence de ces différents motifs et d'autre part, il a été suggéré que ces motifs peuvent contribuer à la spécificité de l'interaction entre un activateur et un promoteur. En effet certains activateurs préfèrent agir sur des promoteurs contenant une boîte TATA plutôt que des promoteurs avec un DPE et pas de boîte TATA (Butler & Kadonaga 2001). Les facteurs interagissant avec les activateurs et intervenant dans le choix d'un type de promoteur par rapport à un autre restent à identifier.

Des études réalisées *in vitro* ont permis de proposer un assemblage séquentiel des facteurs généraux de la transcription (pour revue : Hampsey 1998). Le promoteur est reconnu par le complexe TFIID qui modifie la courbure de l'ADN et agit comme une plate-forme pour l'arrivée des autres facteurs généraux. La sous-unité TBP (TATA-Binding Protein) de TFIID interagit avec la boîte TATA (Parker & Topol 1984; Buratowski *et al.* 1988). D'autres sous-unités (TAFs) de TFIID sont impliquées dans la reconnaissance des autres boîtes DCE, DPE et Inr des différents types de promoteurs (pour une revue détaillée voir Thomas & Chiang 2006). TFIIB vient ensuite stabiliser TFIID en interagissant avec lui et les séquences proches de la boîte TATA. TFIIB recrute alors un complexe formé par TFIIF et l'ARN polymérase II sous sa forme hypophosphorylée (pol IIA), mais la transcription ne peut s'initier que lorsque TFIIE et TFIIH sont incorporés au PIC. L'activité hélicase de TFIIH catalyse l'ouverture de la bulle de transcription au niveau du TSS. La conservation de la distance entre la boîte TATA et le TSS permet que la liaison de TFIID sur la boîte TATA induise la localisation du site actif de la polymérase précisément au niveau du TSS.

Un mécanisme alternatif de formation du PIC où les différents facteurs sont préassemblés avant de lier l'ADN a été proposé après observation de plusieurs complexes de la polymérase avec des SRB (Suppressors of RNA polymerase B mutations) et une partie des facteurs généraux de transcription (Kim *et al.* 1994; Koleske & Young 1994). La composition de ces complexes, appelés holoenzymes, peut varier en fonction de la méthode de purification utilisée, mais elle ne comprend pas TFIID (Wu & Chiang 1998), ce qui suggère que c'est ce facteur qui se lie à la région promotrice puis recrute le complexe holoenzyme.

La prédominance de l'un ou l'autre de ces mécanismes *in vivo* reste débattue et les deux modèles sont toujours d'actualité.

Ce PIC minimal suffit pour la transcription *in vitro* mais non *in vivo* où la polymérase doit pouvoir interagir avec les activateurs et initier la transcription sur une matrice chromatinienne. Ces étapes nécessitent l'intervention de différents coactivateurs tels que les complexes de remodelage (voir §1.1.5.3) de la chromatine ou le Médiateur (voir §1.2.6.2).

1.2.3 Initiation et pause

L'initiation de la transcription corrèle avec la phosphorylation de la sérine 5 du CTD de la polymérase (Komarnitsky *et al.* 2000). Cette marque est déposée par la kinase Cdk7 (Cyclin-dépendante kinase 7) appartenant au complexe TFIIH (Shiekhattar *et al.* 1995). Il est proposé que la kinase Cdk9 (Cyclin-dépendante Kinase 9) du complexe P-TEFb (Positive Transcription Elongation Factor b) soit également impliquée dans la phosphorylation de ce résidu (Eissenberg *et al.* 2007). Cette modification est essentielle pour le recrutement des facteurs nécessaires à la mise en place de la coiffe des ARN messagers (Fabrega *et al.*

2003).

Avant l'élongation, plusieurs cycles d'initiation abortive ont lieu, pendant lesquels l'ARN polymérase II synthétise de courts ARN de quelques nucléotides qui sont relargués (Carpousis & Gralla 1980, pour revue : Hsu (2009)). Des progrès dans la détermination de la structure du PIC ont permis de comprendre le mécanisme de production de ces petits ARN. En effet, il a été montré que le domaine N-terminal du facteur TFIIB forme une boucle à l'intérieur du site actif de la polymérase II à l'endroit normalement occupé par l'hybride ARN-ADN. Tant que l'ARN ne dépasse pas 9 bases, cette boucle ne gêne pas (elle pourrait même faciliter l'incorporation des premières bases). Par contre, quand l'ARN dépasse 9 bases, il y a compétition entre la boucle et l'ARN qui amène soit la victoire de l'ARN, l'éjection de TFIIB et le passage en élongation productive (« promoter escape »), soit la victoire de TFIIB et l'avortement de l'initiation qui conduit aux petits ARN observés (Boeger *et al.* 2005; Bushnell *et al.* 2004).

La fuite du promoteur (« promoter escape ») décrit le moment où l'ARN polymérase II rompt ses contacts avec les séquences promotrices et une partie des facteurs liés au promoteur et entre en phase d'élongation (Saunders *et al.* 2006).

Pour certains gènes il a été montré que la polymérase s'arrête peu de temps après son départ du promoteur. Cette pause a été d'abord observée sur le gène inductible par choc thermique *hsp70* chez *Drosophila melanogaster* (Gilmour & Lis 1986; Rougvie & Lis 1988). Avant le choc thermique, l'ARN polymérase II est en pause après avoir transcrit 20 à 40 nucléotides. Elle ne peut alors progresser efficacement vers une élongation productive qu'en réponse à une stimulation par des signaux extérieurs comme le choc thermique. D'abord considérée comme un phénomène marginal, la pause à proximité du promoteur semble au contraire être un phénomène très répandu qui régule la transcription de nombreux gènes chez les eucaryotes. Elle a été particulièrement bien étudiée dans le cas de la drosophile (pour revues : Saunders *et al.* 2006; Gilmour 2009). Cependant, ce type de régulation a également été proposé chez les mammifères où de nombreux gènes inactifs possèdent les marques de l'initiation (présence du PIC (Kim *et al.* 2005), présence des marques H3K4me3, H3K9Ac et K14Ac (Guenther *et al.* 2007, voir §1.1.4,)).

L'induction du locus *hsp70* nécessite la liaison des protéines GAF et HSF (Heat Shock Factor) en amont du promoteur. GAF est présent avant le choc thermique et est nécessaire pour la pause de la polymérase II et pour l'établissement d'une structure permettant la liaison de HSF (Shopland *et al.* 1995). Un choc thermique induit le recrutement de HSF qui favorise alors celui de P-TEFb qui permet la reprise de l'élongation (Boehm *et al.* 2003). Une étude récente à l'échelle du génome de la drosophile montre que ce mécanisme n'est pas restreint à ce gène mais intervient dans la régulation transcriptionnelle de centaines d'autres (Lee *et al.* 2008). Ces auteurs ont effectué une immunoprécipitation de la chromatine suivie d'une hybridation sur puce (ChIP-chip) de l'ADN associé à NELF et ont montré que ce facteur est localisé légèrement en aval du promoteur de plus de 2000 gènes où la pol II était généralement en pause. Ils ont également montré que GAF était associé à 40% de ces gènes.

Les mécanismes de régulation de cette pause impliquent les complexes DSIF (DRB Sensitivity-Inducing Factor), NELF (Negative Elongation Factor) et P-TEFb. Les facteurs DSIF et NELF agissent de façon coopérative pour bloquer l'élongation (pour revue : Gilmour 2009). Bien que les modèles antérieurs suggéraient que c'était la sous-unité NELF-E qui liait l'ARN naissant et bloquait son extrusion (Yamaguchi *et al.* 2002; Gilmour 2009), une étude récente chez la drosophile montre que c'est la sous-unité Spt5 de DSIF qui interagit avec l'ARN (seulement quand il atteint une taille supérieure à 18 nucléotides et émerge de la surface de la pol II) et qui est responsable du recrutement de NELF (Missra & Gilmour 2010). Ce dernier stabilise fortement l'interaction DSIF/pol II. Les deux complexes NELF et DSIF reconstitués *in vitro* ne suffisent pas à arrêter complètement la pol II mais la ralentissent seulement. Cela suggère que d'autres facteurs, dont probablement

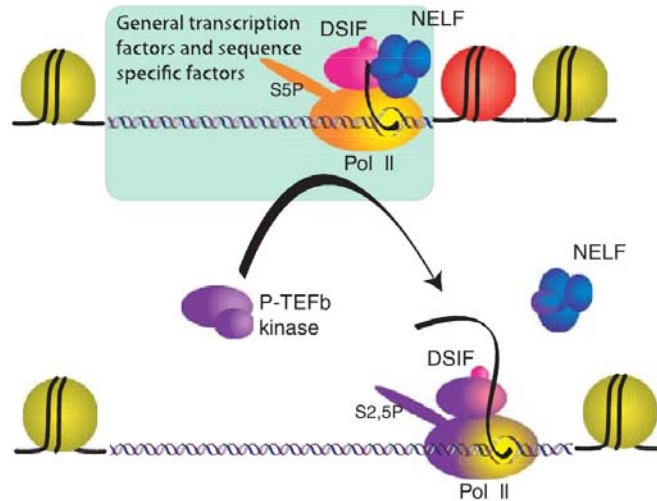


Figure 1.24 – Modèle de la pause à proximité du promoteur (Gilmour 2009). La pause est maintenue par les facteurs NELF et DSIF. L'intervention de la kinase de P-TEFb permet le décrochage de NELF et la reprise de l'élongation.

les nucléosomes, sont nécessaires à un arrêt complet (Missra & Gilmour 2010).

Cette étape est un point important de régulation de l'élongation de la transcription par l'ARN polymérase II *in vivo*. La pause pourrait à la fois servir de point de contrôle pour la vérification de la préparation de l'ARN polymérase II à son entrée en élongation productive, et permettre une induction rapide de l'expression des gènes. En effet, beaucoup des gènes régulés par la pause appartiennent à la catégorie des gènes répondant à un stimulus environnemental ou développemental (Muse *et al.* 2007). Cependant, la déplétion de NELF supprime la pause de la pol II mais ne permet pas d'atteindre les niveaux d'expression observés après induction, ce qui suggère la nécessité d'un activateur supplémentaire dont la fonction n'est pas seulement de faire repartir la polymérase arrêtée (Gilchrist *et al.* 2008). Il est également possible que la déplétion de NELF diminue la quantité de pol II prête à entrer en élongation. Une telle diminution globale du recrutement de la pol II a été observée et permet d'expliquer le comportement de centaines de gènes dont l'expression est diminuée après déplétion de NELF (Gilchrist *et al.* 2008). Cette étude montre une augmentation de la densité des nucléosomes près du promoteur en l'absence de NELF. Ainsi, le rôle de la polymérase en pause pourrait être de maintenir la chromatine dans un état ouvert permettant un haut niveau de transcription. En effet, beaucoup de gènes parmi les plus hautement transcrits présentent un enrichissement de NELF et de la pol II en pause (Lee *et al.* 2008).

1.2.4 Reprise de l'élongation

L'entrée en élongation productive coïncide avec le recrutement du complexe P-TEFb, dont la kinase Cdk9 est responsable de la phosphorylation de DSIF, NELF et de la sérine 2 du CTD de l'ARN polymérase II (Saunders *et al.* 2006). La phosphorylation de la sérine 2 de la pol II par P-TEFb semble être l'étape clé du relargage de la pause (Boehm *et al.* 2003). En effet, c'est la forme IIO de la polymérase où le CTD est phosphorylé sur les sérines 2 et 5 qui est impliquée dans l'élongation. Le CTD subit plusieurs cycles de phosphorylation/déphosphorylation au cours du processus de transcription. Cependant, l'utilisation du flavopiridol, un inhibiteur spécifique de P-TEFb (Chao *et al.* 2000; Chao & Price 2001) soulève de nouvelles questions. En effet, après traitement avec cette drogue, la polymérase reprend effectivement l'élongation mais s'arrête de nouveau après avoir transcrit environ 100 nucléotides (Ni *et al.* 2008). Le facteur ELL (Eleven nineteen Lysine-rich

Leukemia) est aussi impliqué dans la reprise de l'élongation. En effet, il a été montré que sa déplétion entraîne une réduction de la polymérase phosphorylée sur la sérine 2 sur le gène *hsp70* activé par choc thermique, ainsi qu'une diminution de la transcription (Smith *et al.* 2008). Quand la pol II reprend l'élongation, NELF se dissocie (Wu *et al.* 2003) alors que DSIF reste associé et devient un activateur de l'élongation (Yamada *et al.* 2006). Une étude (Ivaldi *et al.* 2007) controversée (Cai *et al.* 2008) propose l'intervention de la kinase Jil1 qui phosphoryle la sérine 10 de l'histone H3, ce qui permettrait le recrutement de P-TEFb. Tout récemment, une autre étude a proposé que la régulation de la pause par P-TEFb faisait aussi intervenir son rôle dans la phosphorylation de H1. Cette phosphorylation permet la dissociation de H1, rendant potentiellement le nucléosome plus facile à remodeler ou éjecter, ce qui favoriserait l'entrée de la polymérase en élongation productive (O'Brien *et al.* 2010).

Lors de la transition entre l'initiation et l'élongation, les facteurs généraux de transcription, à l'exception de TFIIF, sont progressivement relâchés et la polymérase peut alors s'associer aux facteurs d'élongation. Ces derniers préviennent les pauses et les arrêts de la polymérase durant la synthèse de l'ARNm (Conaway *et al.* 2000). Au niveau du complexe d'élongation, un hybride ADN-ARN de 8 à 9 nucléotides se forme au sein de la bulle qui reste ouverte entre -9 et +2 (le site +1 correspondant à la position du dernier rNTP (ribonucléotide Triphosphate) ajouté dans le complexe en élongation). Les sous-unités Rpb1 et Rpb2 forment le centre catalytique de l'enzyme (Coulombe & Burton 1999).

La vérification de la fidélité de la transcription est indispensable pour une expression correcte des gènes. Celle-ci se fait tout au long de l'élongation. Une boucle de Rpb1 située près du site actif interagit avec le nucléotide entrant (Wang *et al.* 2006). Si celui-ci est complémentaire avec la base de l'ADN qui se trouve dans le site actif, la conformation obtenue favorise l'attaque nucléophile et la formation de la liaison phosphodiester entre l'ARN déjà formé et le nouveau nucléotide. Si le nucléotide entrant n'est pas complémentaire, il se forme un mésappariement dans le site actif et la vitesse de formation de la liaison est nettement réduite. À ce moment, soit le nucléotide ressort du site actif, ce qui permet d'en tester un autre, soit la liaison se forme malgré tout. La conformation anormale de l'hybride ADN/ARN ralentit le processus d'ajout du nucléotide suivant et le transcrit devient sensible à l'activité endonucléase intrinsèque de la polymérase qui peut revenir en arrière sur quelques nucléotides et couper le transcrit puis recommencer à ajouter des nucléotides (Sydow & Cramer 2009). Cette activité endonucléase importante pour la fidélité de la transcription est stimulée par TFIIS (Thomas *et al.* 1998).

L'élongation est couplée à la biogenèse de l'ARNm qui comporte la mise en place de la coiffe, l'épissage, le clivage du transcrit et sa polyadénylation. La transcription affecte, et est affectée par, les autres processus métaboliques de l'ADN que sont la réparation, la recombinaison et la réplication.

1.2.5 Terminaison

Au fur et à mesure de la transcription, le CTD de la polymérase est phosphorylé sur la sérine 2 d'un nombre croissant de répétitions de l'heptapeptide Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7. Cette marque est importante pour le recrutement de la machinerie de maturation des ARN en 3' des gènes. Pcf11 (Protein 1 of Cleavage and polyadenylation Factor 1) est un facteur important dans cet étape. Il interagit avec le CTD et l'ARN et permet la dissociation du complexe d'élongation (Zhang & Gilmour 2006). Comme la phosphorylation, la déphosphorylation du CTD joue un rôle important dans le processus transcriptionnel en permettant à l'ARN pol II de revenir à sa forme hypophosphorylée (la seule capable d'intégrer le complexe de pré-initiation) et d'entamer un nouveau cycle de transcription. La protéine phosphatase Fcp1 (TFIIF interacting CTD phosphatase) est responsable de la déphosphorylation du CTD (Cho *et al.* 1999). Fcp1 pourrait jouer un

rôle dans le recyclage de la polymérase bien qu'il ne soit pas connu si la déphosphorylation du CTD est une cause ou une conséquence de l'arrêt de la transcription.

1.2.6 Régulations de la transcription basale

1.2.6.1 Les activateurs

Les activateurs (*enhancers* en anglais) sont des séquences spécifiques qui activent la transcription à distance de façon indépendante de l'orientation. Ils sont reconnus par des facteurs de transcription qui peuvent être exprimés de façon ubiquitaire ou spécifique de certains tissus. Ces facteurs activent la transcription par différents mécanismes. Ils augmentent la probabilité ou la vitesse de l'initiation soit en ouvrant la chromatine, soit en recrutant la machinerie de transcription, soit en facilitant son passage en élongation après la pause. Deux modèles permettent d'expliquer comment un activateur peut communiquer avec sa cible :

- Le premier modèle propose que la protéine activatrice initialement recrutée au niveau de l'activateur se déplace le long de l'ADN jusqu'à trouver sa cible.
- Le second modèle propose la formation d'une boucle de chromatine qui permet un regroupement spatial de l'activateur et du promoteur. Cette boucle peut se former soit par collisions aléatoires jusqu'à ce que l'interaction spécifique stabilise cette structure, soit par un balayage linéaire de l'activateur le long de l'ADN jusqu'à trouver sa cible.

Ces deux modèles sont probablement applicables suivant le type d'activateur concerné. Les activateurs peuvent être localisés à une très grande distance de leurs promoteurs cibles (plusieurs milliers de paires de bases) et même être localisés sur un autre chromosome (pour revue : Williams *et al.* 2010). Dans de nombreux cas, la formation d'une boucle de chromatine permettant de les rapprocher physiquement a été démontrée par capture de conformation chromosomique (3C ; Dekker *et al.* 2002; Lieberman-Aiden *et al.* 2009). Un exemple bien caractérisé de ce type d'organisation est le locus β -globin chez l'humain (Tolhuis *et al.* 2002, Fig. 1.25). La formation d'une boucle en *cis* rapproche physiquement la région LCR (Locus Control Region) de gènes situés à ~ 50 Kpb de distance, formant une région de chromatine transcriptionnellement active (active chromatin hub) qui permet l'expression des gènes. Dans d'autres types cellulaires, cette boucle ne se forme pas et les gènes ne sont pas exprimés.

Les mécanismes permettant l'interaction entre un activateur et un promoteur sont encore mal connus. Plus de quinze facteurs parmi lesquels des facteurs de transcription, de remodelage de la chromatine, des insulateurs et des protéines ayant un rôle architectural comme la cohesin ont été impliqués dans la formation ou la stabilisation de ces boucles (pour revue : Sexton *et al.* 2009). Cette interaction est transitoire et sa force module la stabilité de la liaison et la durée de l'activation, et donc le niveau de transcription. Elle est généralement assez stable pour perdurer pendant plusieurs cycles de transcription, et quand elle est perdue, l'activateur peut sélectionner un nouveau partenaire (Geyer & Clark 2002). Cette régulation à très longue distance nécessite des mécanismes rigoureux pour maintenir la spécificité de l'activation. Une partie de cette spécificité est assurée par les insulateurs qui limitent l'activité des activateurs et des répresseurs (voir §1.4).

1.2.6.2 Le complexe Médiateur

Médiateur est un grand complexe multiprotéique qui est recruté pendant l'assemblage du PIC (pour revues : Casamassimi & Napoli 2007; Sikorski & Buratowski 2009). Ce complexe active la transcription en servant d'adaptateur entre les protéines activatrices liées à une séquence régulatrice en *cis* et la pol II sur le promoteur. Le mécanisme d'activation par le Médiateur n'est pas précisément connu. Il interagit avec le CTD de la pol II non

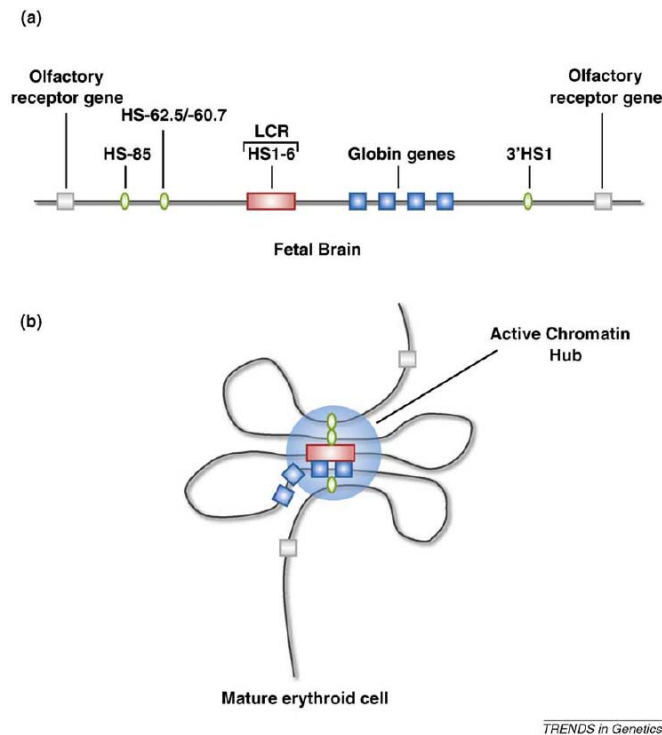


Figure 1.25 – Modèle d'organisation du locus β -globin (Williams *et al.* 2010). (a) Dans le cerveau foetal, le locus adopte une conformation linéaire et les gènes β -globin ne sont pas exprimés. (b) Dans les cellules érythroïdes, le locus forme des boucles de chromatine qui rapprochent la LCR ainsi que des sites hypersensibles (HS) distants des gènes β -globin. Cette conformation crée une région transcriptionnellement active et les gènes sont exprimés.

phosphorylé et pourrait faciliter le recrutement de certains éléments du PIC (Thompson *et al.* 1993). Il régule également l'assemblage du PIC en recrutant l'acétyltransférase p300. Ce recrutement bloque l'assemblage jusqu'à ce que p300 acétyle ces cibles (les histones et lui-même) ce qui provoque son détachement (Black *et al.* 2006). Ainsi le Médiateur permet de contrôler l'assemblage du PIC pour que les modifications des histones nécessaires soient bien effectuées. La phosphorylation du CTD de la polymérase pendant l'initiation décroche le complexe Médiateur (Max *et al.* 2007) qui pourrait avoir un rôle dans la ré-initiation rapide de la transcription tant que le signal activateur perdure.

1.3 Chromatine et transcription

Bien que moins compacte que l'hétérochromatine, l'euchromatine est tout de même très dense en nucléosomes qui doivent être déplacés par les complexes de remodelage pour permettre un accès complet à l'information génétique (Workman 2006). La capacité des nucléosomes à réprimer la transcription a été démontrée *in vitro* (Laybourn & Kadonaga 1991) et *in vivo* (Han & Grunstein 1988). La chromatine est à la fois un obstacle à l'initiation de la transcription et un frein à l'élongation. Ainsi, la connaissance des mécanismes contrôlant la localisation des nucléosomes devrait permettre une avancée majeure dans la compréhension de la régulation la transcription.

1.3.1 Positionnement des nucléosomes

La notion de positionnement des nucléosomes est différente de la notion d'occupation. En effet, l'occupation définit la densité globale des nucléosomes alors que le positionnement définit s'ils sont précisément positionnés ou si leur localisation est floue, c'est à dire qu'elle

varie fortement d'une cellule à l'autre au sein d'une population (Fig. 1.26). Ces deux propriétés sont importantes pour appréhender l'effet des nucléosomes sur la transcription.

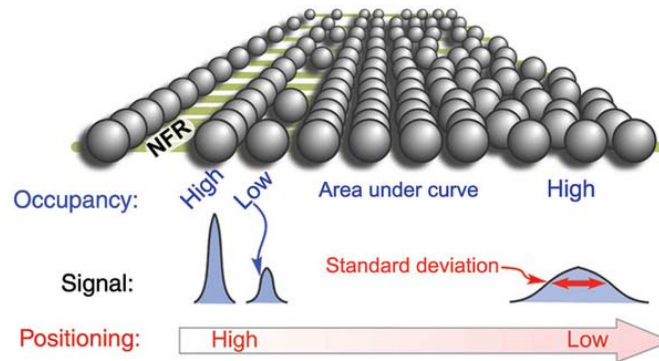


Figure 1.26 – Occupation et positionnement (Pugh 2010). L'occupation désigne la densité globale des nucléosomes alors que le positionnement exprime le fait que leur position est globalement conservée au sein d'une population ou si elle varie. Ces deux notions sont disjointes et toutes les combinaisons de ces deux propriétés existent.

In vivo, trois mécanismes sont impliqués dans la réduction de l'occupation des nucléosomes à certains loci :

- Certains complexes de remodelage de la chromatine recrutés par différentes protéines activatrices peuvent générer des régions pauvres en nucléosomes indépendamment de la transcription (Schwabish & Struhl 2007).
- L'élongation par la polymérase II nécessite des cycles de désassemblage/réassemblage des nucléosomes dans les régions codantes, ce qui génère pour les gènes hautement transcrits une diminution de la densité nucléosomique (Schwabish & Struhl 2004).
- Enfin, certaines séquences comme les séries de poly(dA :dT) (séries de ~ 10 (ou plus) nucléotides désoxyadénosines consécutifs sur le même brin) défavorisent intrinsèquement l'enroulement de l'ADN autour des histones *in vitro* et rendent la chromatine accessible *in vivo*. Ces séquences sont nettement enrichies dans les régions NFR (Nucleosome Free Region ; Segal & Widom 2009a).

Seulement 75 à 90% de l'ADN est contenu dans un nucléosome, laissant entre 20 et 50 pb d'ADN de liaison entre chacun d'entre eux. En ce qui concerne le positionnement des nucléosomes, a priori n'importe quelle séquence d'ADN peut être enroulée autour d'un octamère d'histones, mais il apparaît que les nucléosomes sont précisément positionnés dans certaines régions du génome. Les propriétés de l'ADN engagé dans un nucléosome et de l'ADN de liaison sont très différentes. En effet, dans le nucléosome l'ADN est fortement courbé et il est stériquement encombré par les histones et par le second tour d'ADN (Richmond & Davey 2003). Ces différences structurelles modifient l'affinité de liaison des protéines sur l'ADN. Ainsi, la position génomique des nucléosomes influence la liaison de multiples régulateurs, ce qui fait des nucléosomes des répresseurs ou des activateurs de l'expression génique suivant le contexte (Wyrick *et al.* 1999). La compréhension des mécanismes mis en jeu dans ce positionnement et son rôle dans la régulation de l'expression constitue actuellement un enjeu majeur (pour revue : Segal & Widom 2009b).

L'étude du positionnement des nucléosomes a récemment profité de l'émergence des technologies d'hybridation sur puce puis de séquençage à haut débit. Ainsi, pour connaître le positionnement précis des nucléosomes, des études à l'échelle du génome ont été conduites chez différentes espèces : *S. cerevisiae* (Yuan *et al.* 2005), *H. sapiens* (Schones *et al.* 2008), *C. elegans* (Valouev *et al.* 2008), *D. melanogaster* (Mavrich *et al.* 2008b) et *S. pombe* (Lantermann *et al.* 2010). Ces études montrent une grande conservation du positionnement des nucléosomes vis à vis du TSS des gènes. Les nucléosomes les mieux positionnés sont ceux en aval des TSS des gènes transcrits (nommés +1, +2, ... à partir du TSS, voir Fig.1.27).

Juste en amont du TSS il y a une région déplétée en nucléosomes (NFR) qui correspond à la région promotrice où s'assemble le complexe de préinitiation.

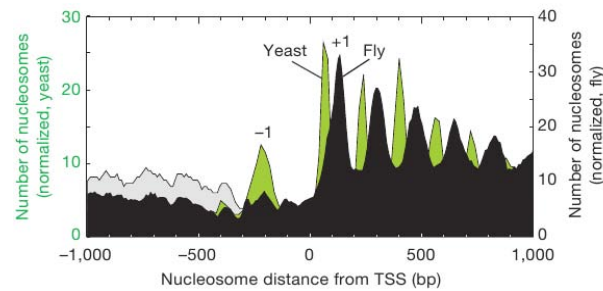


Figure 1.27 – Positionnement des nucléosomes par rapport au TSS chez la levure (vert) et la drosophile (noir ; Mavrich *et al.* 2008b).

Les mécanismes permettant un tel positionnement sont à l'heure actuelle mal connus, et plusieurs modèles complémentaires sont proposés.

1.3.1.1 Influence de la séquence

L'influence de la séquence de l'ADN est le facteur le plus étudié pour expliquer le positionnement des nucléosomes. Il existe une très grande différence d'affinité (jusqu'à 1000 fois) pour l'octamère d'histones suivant la séquence de l'ADN (Thastrom *et al.* 2004). Il n'y a pas de contacts dépendant spécifiquement d'une base entre l'ADN et les histones, par contre, l'enchaînement des bases modifie nettement la flexibilité de l'ADN et son affinité pour l'octamère d'histones (Widom 2001). Différents modèles ont été mis au point pour prédire le positionnement des nucléosomes à partir de la séquence (Segal *et al.* 2006). Ces modèles rendent assez bien compte de l'emplacement des nucléosomes sur de l'ADN *in vitro* mais ne suffisent pas à expliquer le positionnement *in vivo* (Zhang *et al.* 2009). Bien que la séquence influence le positionnement des nucléosomes *in vivo*, elle ne suffit pas à prédire ce positionnement.

1.3.1.2 Barrière d'énergie

Pour expliquer les données obtenues *in vivo* chez la levure, un modèle de positionnement statistique comprenant une barrière d'énergie près de laquelle les nucléosomes sont bien positionnés a été proposé (Mavrich *et al.* 2008a; Milani *et al.* 2009). Ce modèle propose que la région déplétée en nucléosomes en amont du TSS impose la localisation des nucléosomes adjacents. Il reproduit assez bien les données expérimentales mais il reste controversé notamment par une étude récente qui montre qu'une région pauvre en nucléosomes en amont du TSS n'est pas suffisante pour définir le nucléosome +1 (Zhang *et al.* 2009).

1.3.1.3 Transcription

L'article cité précédemment propose que c'est le processus d'initiation de la transcription qui positionne le nucléosome +1 (Zhang *et al.* 2009). Ainsi, le fait que la région en amont du TSS soit intrinsèquement pauvre en nucléosomes facilite l'assemblage du complexe de préinitiation mais ce sont des facteurs appartenant à (ou interagissant transitoirement avec) ce complexe qui positionnent le nucléosome +1. Ce modèle permet d'expliquer pourquoi ce sont les gènes les plus fortement transcrits (souvent ré-initiés) qui présentent un profil de positionnement précis.

1.3.1.4 Motifs spécifiques en phase ou en opposition de phase

L'analyse de motifs d'ADN conservés parmi différentes espèces de drosophiles (Stark *et al.* 2007) montre que certains motifs s'organisent vis à vis des nucléosomes. Ainsi, il y a des motifs nucléosomiques en phase avec les nucléosomes, des motifs anti-nucléosomiques en opposition de phase et des motifs n'ayant pas d'organisation particulière (Mavrich *et al.* 2008b).

Ces motifs peuvent être liés par des facteurs spécifiques qui semblent avoir un rôle dans le positionnement des nucléosomes. En effet, la déviation entre les données de positionnement *in vivo* et *in vitro* sur certains sites chez la levure peut s'expliquer par la liaison de facteurs spécifiques (Kaplan *et al.* 2009). Un exemple d'un tel facteur est la protéine insulatrice CTCF (CCCTC-binding factor) humaine (voir §1.4.4). En effet, l'étude de la chromatine autour des sites occupés par CTCF montre une organisation très précise avec une dizaine de nucléosomes bien positionnés de chaque côté du site de liaison de CTCF qui lui est nettement déplété en nucléosomes (Fu *et al.* 2008). Il n'est cependant pas clair dans quelle mesure CTCF pourrait avoir un rôle direct dans ce positionnement.

1.3.1.5 Lien entre positionnement des nucléosomes et transcription

Il y a encore peu de données permettant d'appréhender comment la machinerie de transcription s'assemble et fonctionne dans le contexte de la chromatine. Il est probable que certains (ou peut-être même la plupart) des facteurs impliqués se lient à un ou plusieurs nucléosomes et non à l'ADN nu. Les expériences d'immunoprécipitation de la chromatine ne permettent pas de distinguer ces deux types d'interaction. Pour interagir avec un nucléosome, une protéine peut soit interagir avec des histones, soit avec l'ADN nucléosomique, soit encore entrer en collision avec lui sans avoir d'affinité de liaison. En ajoutant une étape de digestion MNase (Micrococcal nuclease) au protocole de ChIP, l'équipe de Pugh a étudié les interactions entre certains facteurs et les nucléosomes (Koerber *et al.* 2009). Ils ont ainsi pu montrer que des protéines s'associent à des nucléosomes précis (par exemple Bdf1 qui est associé aux nucléosomes +1 et +2 simultanément) alors que d'autres s'associent aux nucléosomes de façon non spécifique (exemple de Smr1). Ils ont aussi observé que les nucléosomes associés à la pol II n'étaient pas précisément localisés ce qui est en accord avec le modèle selon lequel la pol II entre en collision avec les nucléosomes, les déplaçant aléatoirement avant leur éjection ou désassemblage.

1.3.1.6 Lien entre positionnement des nucléosomes et pause à proximité du promoteur

La localisation de l'histone H2A.Z a été précisément déterminée par séquençage à haut débit après immunoprécipitation (ChIP-seq) (Mavrich *et al.* 2008b). Cette marque se retrouve sur les gènes qui sont ou viennent d'être transcrits. Cette étude montre que le 1^{er} nucléosome après le TSS est centré sur la position +135 dans le cas général, et +145 dans le cas des gènes ayant une pol II en pause. Cela suggère que le 1^{er} nucléosome est en contact avec la polymérase en pause, ce qui le déplace légèrement. Reste à savoir si le nucléosome a un rôle actif dans la localisation de la pause de la pol II ou si au contraire c'est la pol II qui impose la localisation du nucléosome.

De façon intéressante, la déplétion du facteur de stabilisation de la pause NELF induit une nette perturbation de la structure chromatiniennne avec une augmentation de l'occupation des gènes par les nucléosomes corrélée à une diminution de l'expression (Gilchrist *et al.* 2008). Cette donnée suggère que la polymérase en pause pourrait imposer une certaine organisation des nucléosomes en aval du TSS.

Après un choc thermique, les gènes du locus *Hsp70* sont rapidement et fortement exprimés. L'étude de l'occupation en nucléosomes de ces gènes montre une éviction d'une

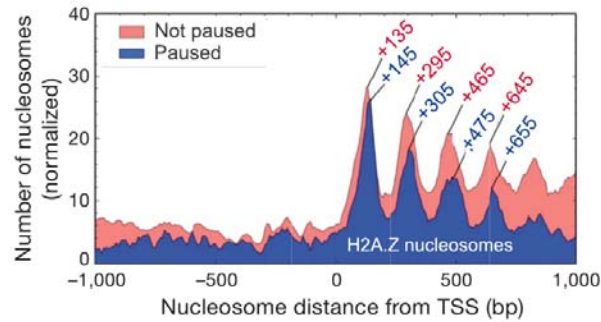


Figure 1.28 – Positionnement des nucléosomes dans les gènes en pause et dans les gènes sans pause (Mavrich *et al.* 2008b).

grande partie des nucléosomes après le choc thermique (Petesch & Lis 2008). De façon intéressante, ce remodelage se fait de façon indépendante de la transcription. En effet, il a lieu nettement avant la reprise de l'élongation et même lorsque la polymérase est bloquée artificiellement sur le site de pause. Ce résultat montre que, au moins dans ce cas, ce n'est pas la polymérase en pause qui impose le positionnement des nucléosomes et suggère que d'autres facteurs inconnus sont impliqués dans ce positionnement.

1.3.2 Remodelage de la chromatine aux promoteurs

Afin d'initier la transcription dans un contexte chromatinien, des facteurs de transcription et des complexes de remodelage de la chromatine sont recrutés sur le promoteur. Ces complexes remplacent notamment les histones canoniques par des variants, modifient les histones, et repoussent ou éjectent les nucléosomes pour créer une structure adaptée au recrutement ou au départ de la machinerie de transcription (pour revues : Cairns 2009; Clapier & Cairns 2009).

Les promoteurs des gènes sont généralement pauvres en nucléosomes (Lee *et al.* 2004), mais leur activation diminue encore nettement leur densité (Schwabish & Struhl 2004). Les complexes SWI/SNF sont importants pour ce remodelage à certains promoteurs. En effet, chez la levure, une mutation de ce complexe ralentit considérablement ce remodelage au niveau du promoteur PHO8 (Gregory *et al.* 1999). Il a également été montré chez la drosophile que Brahma est nécessaire à l'activation de la plupart des gènes (Armstrong *et al.* 2002).

Il y a une forte interdépendance entre les modifications des histones et le remodelage de la chromatine. L'expression d'un gène peut nécessiter la déstabilisation des nucléosomes sur le promoteur par un complexe de remodelage pour permettre l'accès à une histone acétyltransférase (Cosma *et al.* 1999), alors qu'un autre nécessitera à l'inverse d'abord l'acétylation des histones et la liaison de la pol II pour permettre le recrutement d'un complexe de remodelage (Agalioti *et al.* 2000). Il ne semble donc pas exister de mécanisme général avec conservation de l'ordre d'intervention des différents facteurs.

1.3.3 Chromatine et élongation

L'élongation ayant lieu dans le contexte de la chromatine, la polymérase doit passer la barrière constituée par l'enchaînement des nucléosomes. Plusieurs mécanismes sont proposés pour permettre ce passage (pour revue : Selth *et al.* 2010). Une hypothèse est que le nucléosome soit désassemblé devant la pol II et réassemblé après son passage. Il a en effet été montré chez la levure que la densité de nucléosomes sur un gène diminue après son activation (Schwabish & Struhl 2004). Chez la drosophile, il a été observé que la densité d'histones est réduite dans les gènes actifs et que le variant H3.3 s'y incorpore plus rapidement que dans les régions inactives, ce qui suggère un réassemblage rapide après passage

de la polymérase (Mito *et al.* 2005). Il n'est cependant pas certain que le désassemblage des nucléosomes soit nécessaire au passage de la pol II. En effet, des expériences *in vitro* ont suggéré que la polymérase peut passer à travers un nucléosome isolé sans le déplacer. La machinerie de transcription semble pouvoir tourner autour du nucléosome déplaçant ainsi la bulle de transcription autour de lui (Studitsky *et al.* 1997). Plusieurs données suggèrent également qu'il peut y avoir un désassemblage partiel des nucléosomes avec l'éviction d'un dimère H2A/H2B (Belotserkovskaya *et al.* 2003, Fig. 1.29).

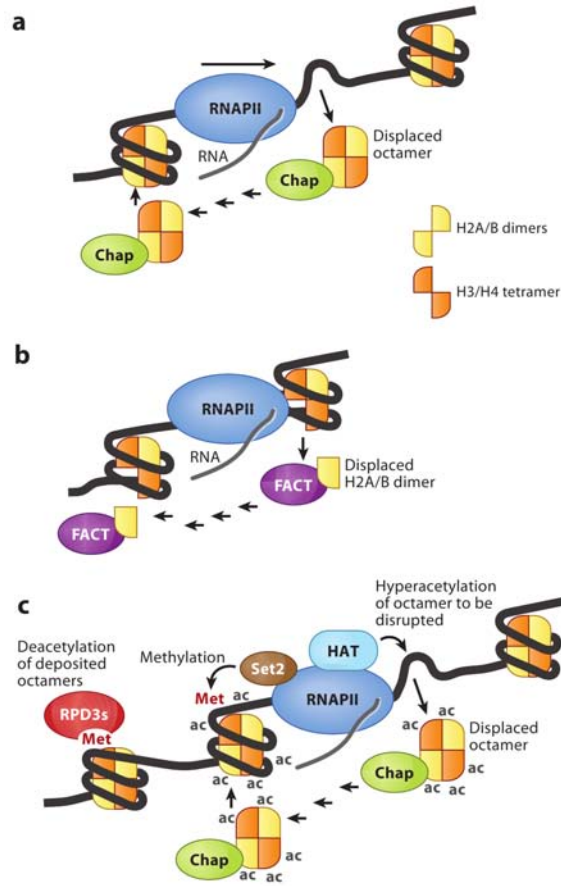


Figure 1.29 – Transcription à travers les nucléosomes (Selth *et al.* 2010). a. Un chaperon (Chap) facilite le décrochage de l'octamère d'histones (ou le prend en charge après son décrochage par la polymérase) puis son réassemblage après le passage de la polymérase. b. FACT déstabilise le nucléosome en chaperonnant un dimère H2A/H2B, ce qui pourrait suffire à permettre le passage de la pol II, puis participe au réassemblage. c. Les histones sont hyperacétylées devant la polymérase pour faciliter leur prise en charge par des chaperons. Les histones peuvent alors être méthylées, ce qui permet le recrutement d'une désacétylase qui remet les histones dans leur état d'origine.

Il existe des facteurs qui facilitent ce passage. Parmi eux se trouvent des facteurs de remodelage de la chromatine, ainsi que des chaperons d'histones. Des enzymes de modification des histones interviennent également dans ces processus.

Plusieurs facteurs de remodelage de la chromatine (voir §1.1.5.3) sont impliqués dans l'élongation de la transcription. Par exemple, RSC (Chromatin Structure Remodeling), de la famille SWI-SNF, peut faciliter l'élongation à travers un mono-nucléosome *in vitro* (Carey *et al.* 2006). Cette activité est augmentée lorsque les histones sont acétylées car l'affinité de RSC pour le nucléosome s'en trouve augmentée. En accord avec ces résultats, des travaux récents montrent que RSC est recruté dans les régions codantes de gènes activés par un stress et qu'il est nécessaire à la progression de la pol II (Mas *et al.* 2009). *In vitro*,

il a été montré que RSC se lie à un nucléosome et déstabilise l'interaction entre l'ADN et les histones, ce qui rend le nucléosome plus mobile, avant de le déplacer à l'extrémité du substrat d'ADN (Shukla *et al.* 2010). Cette étape de déstabilisation de l'interaction entre l'ADN et les histones pourrait faciliter l'éviction du nucléosome devant la pol II. Le complexe SWI/SNF intervient également dans cette étape chez la levure (Schwabish & Struhl 2004) ainsi que dans la reprise de l'élongation sur le locus *hsp70* humain (Brown *et al.* 1996). Le facteur de remodelage Chd1 est également impliqué dans l'élongation. Chez la levure *S. cerevisiae*, il est présent sur des régions activement transcrites et interagit avec les facteurs d'élongation DSIF, FACT et Paf1 (RNA polymerase II associated factor 1; Simic *et al.* 2003)

Plusieurs chaperons d'histones régulent également l'élongation. La purification de la polymérase en élongation a permis l'identification du facteur FACT (Facilitates Chromatin Transcription) qui permet à une ARN pol II arrêtée au niveau d'un nucléosome de reprendre la transcription indépendamment de l'hydrolyse d'ATP et des facteurs TFIIF ou TFIIS (Orphanides *et al.* 1998). FACT agit en déstabilisant le nucléosome et en décrochant un dimère H2A-H2B (Belotserkovskaya *et al.* 2003; Reinberg & Sims 2006). Le chaperon de H3/H4 Spt6 est nécessaire à la restauration des nucléosomes après le passage de la pol II. Des défauts de ce chaperon et de plusieurs autres entraînent le même phénotype de déplétion des histones dans les régions codantes et l'activation de sites d'initiation cryptiques (Kaplan *et al.* 2003), ce qui suggère que les chaperons en général sont impliqués à la fois dans le décrochage des nucléosomes au moment du passage de la pol II et dans leur réassemblage après cette étape.

Les modifications des histones interviennent également pendant l'élongation, soit en modifiant directement les interactions internucléosomales, soit en permettant le recrutement de complexes facilitant l'élongation. En plus de l'hyperacétylation des histones au niveau des promoteurs actifs, cette marque intervient dans les régions codantes. Le niveau d'acétylation est plus modeste mais la présence d'acétyltransférases et de désacétylases dans ces régions suggère la possibilité d'un important renouvellement de l'acétylation. Ces deux types d'enzymes étant nécessaires à l'élongation, il a été proposé que les histones sont acétylées devant la pol II pour être plus facilement prises en charge par les chaperons; les nucléosomes sont réassemblés après le passage de la pol II et les histones sont alors désacétylées (Svejstrup 2003, Fig. 1.29).

1.4 Les insulateurs

Les chromosomes eucaryotes sont constitués de domaines de chromatine plus ou moins compacts et accessibles. La régulation des gènes a lieu dans ce contexte et fait intervenir des éléments activateurs ou répresseurs. Le mécanisme de transcription est le même pour l'ensemble des gènes mais ils sont régulés indépendamment les uns des autres. La spécificité de la régulation nécessite une compartimentation des gènes dans des domaines transcriptionnellement indépendants. Parmi les éléments impliqués dans cette compartimentation se trouvent les insulateurs qui ont pour rôle d'isoler des gènes d'un domaine chromatinien donné de la régulation anormale par des activateurs ou des répresseurs de domaines voisins. Plusieurs exemples de translocations chromosomiques qui invalident des insulateurs ont pour conséquences un profil d'expression génique perturbé qui induit notamment des défauts développementaux (Feinberg 2007).

1.4.1 Définition classique d'un insulateur

Un insulateur est un élément d'ADN classiquement défini par deux propriétés qui sont vérifiées à l'aide de tests fonctionnels (Fig. 1.30; pour revues : Gaszner & Felsenfeld 2006; Valenzuela & Kamakaka 2006; Gurudatta & Corces 2009).

- Tout d’abord lorsqu’un transgène encadré par deux insulateurs est inséré dans un locus, il n’est pas affecté par la structure de la chromatine environnante (rôle de barrière). Ces insulateurs empêchent notamment la progression de la structure répressive quand le transgène est inséré dans l’hétérochromatine et permettent ainsi l’expression du gène. Ces insulateurs ont été identifiés à partir d’expériences de sensibilité à la DNase I. En effet, les régions de transition entre la chromatine ouverte et fermée sont hypersensibles (Udvardy *et al.* 1985).
- D’autre part, lorsqu’un insulateur est inséré artificiellement entre un activateur et un promoteur il bloque la communication entre ces éléments et empêche l’activation du gène (rôle de bloqueur d’activateur). Un insulateur bloqueur d’activateur interfère avec cette communication quand il est placé entre l’activateur et le promoteur, mais n’a aucun effet quand il est situé d’un côté ou de l’autre du couple activateur-promoteur (Gaszner & Felsenfeld 2006).

On considère généralement qu’un insulateur possède à la fois ces deux fonctions mais, par extension, une séquence possédant l’une ou l’autre de ces fonctions est également appelée insulateur.

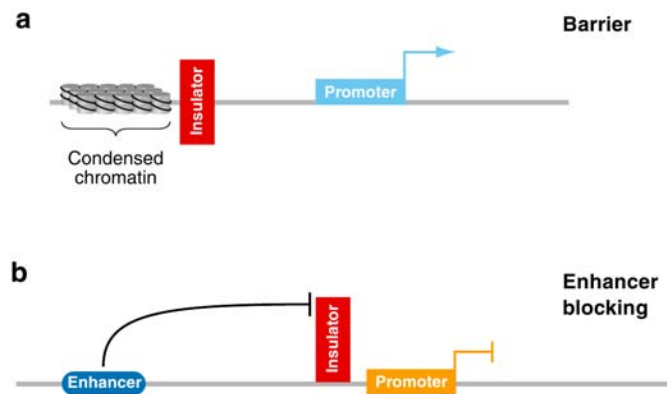


Figure 1.30 – Propriétés définissant les insulateurs (Valenzuela & Kamakaka 2006). Les insulateurs sont définis par leur capacité à bloquer la propagation de la structure chromatinienne environnante et à empêcher la communication entre un activateur et un promoteur.

Chez la drosophile, différentes séquences insulatriques ont été caractérisées. On les classe en fonction des protéines qui les lient et qui confèrent à ces séquences leurs propriétés d’isolation.

1.4.2 BEAF et Zw5

1.4.2.1 Exemple du locus 87A7 de la drosophile

Parmi les premières études sur les insulateurs se trouvent celles menées sur le locus *87A7* chez *Drosophila melanogaster*. Ce locus contient quatre gènes dont les deux gènes *hsp70* transcrits de façon divergente et activés lors d’un choc thermique (Fig. 1.32). L’observation des chromosomes polytènes après un choc thermique montre qu’il se forme une structure ouverte appelée « puff » en anglais et dont la propagation est limitée dans l’espace (Semeshin *et al.* 1982, et Fig. 1.31).

Cette région est relativement ouverte et est flanquée de deux zones ayant une structure chromatinienne inhabituelle appelées *scs* et *scs'* (specialized chromatin structures) séparées par environ 15 Kpb (Udvardy *et al.* 1985). La caractérisation de ces deux séquences ont participer à la définition des deux propriétés classiques des insulateurs (Kellum & Schedl 1991, 1992). Ces propriétés ont amené l’hypothèse selon laquelle ces séquences seraient responsables de l’arrêt de la propagation du « puff » (Udvardy & Schedl 1993).

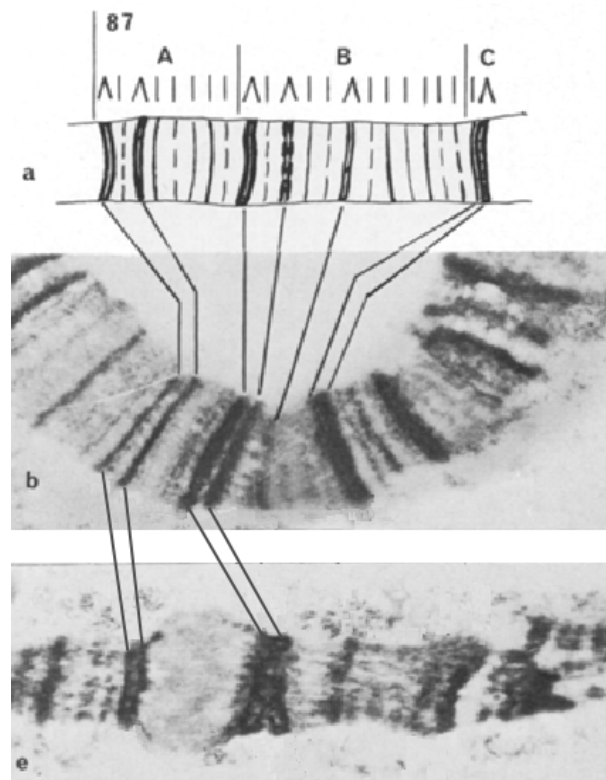


Figure 1.31 – Visualisation sur chromosome polytène de l'effet d'un choc thermique au locus 87A7 (Semeshin *et al.* 1982). Le choc thermique induit la formation d'une structure ouverte ou « puff » (en bas).

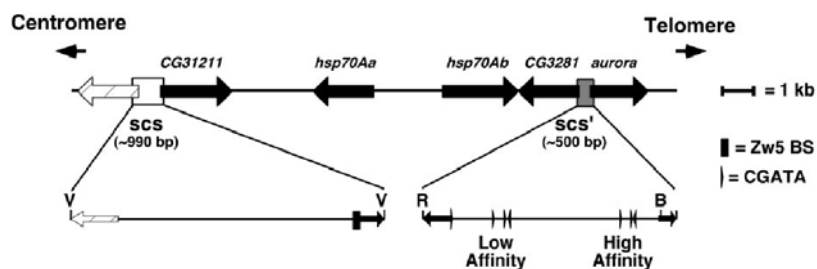


Figure 1.32 – Structure du locus 87A7 de la drosophile (Kuhn *et al.* 2004). Ce locus contient cinq gènes indiqués par les flèches noires : une paire de gènes *hsp70* transcrits de façon divergente, *CG31211* (anciennement *CG14732*), *CG3281* et *aurora*. Les séquences *scs* et *scs'* sont indiquées. Le site de liaison (BS) de Zw5 est représenté par un rectangle noir. Les sites de liaison de BEAF (un de faible affinité (low) et un de forte affinité (high)) ainsi que la position des motifs CGATA sont indiqués (Zhao *et al.* 1995).

L'insulateur *scs* se trouve à proximité du promoteur du gène *CG31211* et est reconnu par la protéine Zw5 (Zeste-white) (Gaszner *et al.* 1999) tandis que l'insulateur *scs'* est proche (<100 pb) des promoteurs de deux gènes transcripts de façon divergente *CG3281* et *aurora* et est reconnu par la protéine BEAF (Boundary Element-Associated Factor) (Zhao *et al.* 1995).

1.4.2.2 Caractérisation de BEAF

La protéine de drosophile BEAF a tout d'abord été caractérisée comme interagissant avec l'élément frontière *scs'* en reconnaissant des motifs palindromiques de type CGATA / TATCG (Zhao *et al.* 1995). Deux isoformes de la protéine BEAF existent : les protéines BEAF32A et BEAF32B (Hart *et al.* 1997). Elles proviennent du même gène et diffèrent par environ 80 acides aminés dans la région N-terminale. Cette région possède un motif atypique de liaison à l'ADN en doigt de zinc appelé « BED finger » (Aravind 2000) qui est essentiel à la fonction insulatrice des protéines (Cuvier *et al.* 1998). La protéine BEAF32B est essentielle à la viabilité de la drosophile contrairement à la protéine BEAF32A dont les mutations ne provoquent pas de défauts phénotypiques majeurs (Roy *et al.* 2007). Le domaine C-terminal de BEAF est impliqué dans la trimérisation de la protéine (Hart *et al.* 1997). Il contient un domaine BESS (BEAF/Su(var)3-7/Stonewall) (Delattre *et al.* 2002) présent dans 19 autres protéines qui pourrait être à l'origine de la trimérisation observée (Bhaskar & Courey 2002).

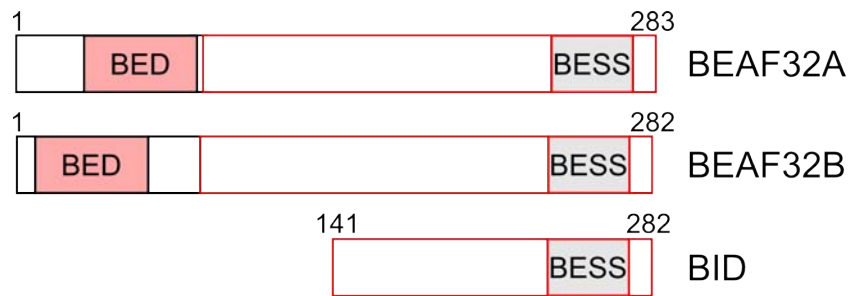


Figure 1.33 – Structure des protéines BEAF. Les domaines « BED finger » (BED) et BESS sont indiqués. En rouge est représentée la partie identique entre ces deux protéines. Le mutant BID (BEAF self-Interaction Domain Gilbert *et al.* 2006) qui ne comprend que les 141 acides aminés C-terminaux est également représenté.

En plus de l'insulateur *scs'*, il a été montré sur des chromosomes polytènes que BEAF se lie à des centaines d'autres sites dont certains ont aussi les propriétés d'un insulateur (Zhao *et al.* 1995; Cuvier *et al.* 1998). Cela suggère que les insulateurs utilisant la protéine BEAF sont largement répandus dans le génome de la drosophile.

La création de BID, un mutant dominant négatif de BEAF, a permis de montrer que BEAF est nécessaire au maintien de la structure de la chromatine dans les chromosomes polytènes (Gilbert *et al.* 2006). Ce mutant est délété de la partie N-terminale de la protéine qui est nécessaire à la liaison à l'ADN (Fig. 1.33). BID séquestre la protéine BEAF endogène, ce qui entraîne une désorganisation complète de la structure en bandes normalement observée. Cette observation suggère un rôle majeur de BEAF dans l'organisation chromatiniennne.

L'élément frontière ou insulateur BE28 (Boundary Element 28) correspond à une séquence de 269 pb qui se retrouve environ 150 fois dans le génome de la drosophile. Sur le chromosome X, elle est incluse dans une région répétée de 1,2 kb localisée entre l'euchromatine et l'hétérochromatine (Cuvier *et al.* 2002). La séquence BE28 a la propriété de bloquer la communication entre un activateur et un promoteur. Cette séquence contient un site de liaison de BEAF et est proche d'une séquence riche en nucléotides A et T qui est liée par la protéine D1 qui pourrait interagir physiquement avec BEAF. D1 est une pro-

téine essentielle exprimée de façon ubiquitaire au cours du développement de la drosophile (Aulner *et al.* 2002). D1 contient 10 motifs *AT-hook* qui lient l'ADN et permet l'association de D1 avec les séquences satellites SAT I et SAT III. D1 pourrait être impliqué dans le recrutement sur ces séquences des protéines HP1 et topoisomérase II et participer la formation de l'hétérochromatine (Blattes *et al.* 2006). Il a été proposé que les répétitions BE28 pourraient définir une frontière entre l'hétérochromatine et l'euchromatine et que les protéines BEAF et D1 pourraient être nécessaires au maintien de cette frontière (Cuvier *et al.* 2002).

Une protéine importante pour la régulation du cycle cellulaire, DREF (DNA Replication-related Element Factor), se lie à la séquence DRE (TATCGATA) qui contient le motif consensus de BEAF. Ces protéines entrent en compétition pour la liaison à la séquence DRE qui est présente dans plusieurs Boundary Elements (BE) liés par BEAF (Hart *et al.* 1999). Il a de plus été montré qu'une surexpression de BEAF supprime le phénotype de développement anormal des ailes induit par la surexpression de DREF, ce qui est en accord avec un mécanisme de compétition (Yoshida *et al.* 2001). DREF active des gènes impliqués dans la prolifération cellulaire (pour revue : Matsukage *et al.* 2008) et la compétition avec BEAF pourrait être importante pour limiter l'activité de DREF.

1.4.2.3 Caractérisation de Zw5

La protéine Zw5, précédemment dénommée SBP, est le produit du gène *zeste-white 5* (Gaszner *et al.* 1999). Elle joue un rôle dans l'activité isolatrice de l'élément *scs* en permettant la formation d'une boucle de chromatine, réunissant les éléments frontières *scs* et *scs'*. Cette boucle, qui formerait un mini-domaine contenant les deux gènes *hsp70*, se constitue grâce à l'interaction entre les protéines Zw5 et BEAF32 (Blanton *et al.* 2003, Fig. 1.34).

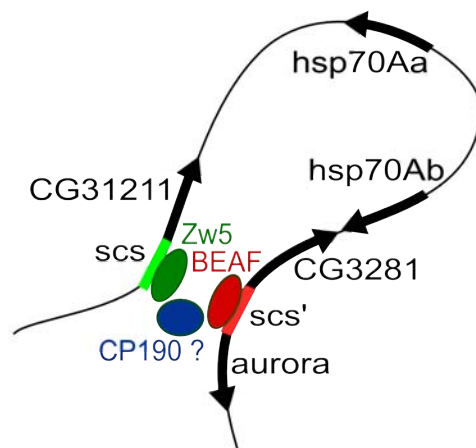


Figure 1.34 – Boucle formée entre les éléments *scs* et *scs'*. Le rapprochement de ces séquences peut être effectué via une interaction directe entre BEAF et Zw5 (qui interagissent effectivement *in vitro*) ou par l'intermédiaire d'un autre facteur comme CP190 (voir §1.4.6).

Les études visant à caractériser Zw5 sont encore peu nombreuses. Une mutation de *zeste-white 5* est létale, ce qui prouve l'importance de cette protéine au delà de la régulation du locus 87A7 (Gaszner *et al.* 1999). Il est important de noter que la séquence *scs* a les propriétés d'un insulateur mais que la séquence de liaison de Zw5 prise isolément n'a pas cette activité isolatrice (Gaszner *et al.* 1999). Cela suggère que d'autres protéines non encore caractérisées s'associent à Zw5 sur ce locus pour lui conférer les propriétés observées.

1.4.3 L'insulateur *gypsy*

La plupart des travaux initiaux de caractérisation de l'activité bloqueur d'activateur des insulateurs a été menée en utilisant l'insulateur *gypsy* de *D. melanogaster* (voir §3; Geyer *et al.* 1986). Celui-ci est situé en 5' du rétrotransposon *gypsy*. Trois protéines ont été décrites dans le complexe associé à cet insulateur : Su(Hw) (Supressor of Hairy wing), Mod(mdg4) et CP190 (Centrosomal Protein 190 Kd).

1.4.3.1 Su(Hw)

Su(Hw) est un facteur de transcription qui possède 12 sites de liaison sur la séquence non traduite en 5' du rétrotransposon. Le site de liaison de Su(Hw) est suffisant pour le blocage de la communication entre un activateur et un promoteur (Geyer & Corces 1992), et confère une protection contre les effets de position dans les tests transgéniques (Roseman *et al.* 1993). Une mutation invalidant Su(Hw) n'est pas létale mais conduit à la stérilité des femelles (Harrison *et al.* 1993).

1.4.3.2 Mod(mdg4)

Le gène *mod(mdg4)* code pour plus de 20 isoformes différentes (Buchner *et al.* 2000) dont l'isoforme Mod(mdg4)2.2 qui est nécessaire à la fonction insulatrice de Su(Hw) avec lequel il interagit (Ghosh *et al.* 2001). Une mutation invalidant le gène *mod(mdg4)* est létale mais une mutation n'affectant que l'isoforme 2.2 est viable et montre des défauts de la fonction insulatrice de *gypsy*. En effet, en l'absence de Mod(mdg4), su(Hw) perd sa fonction insulatrice et acquiert une fonction de répresseur bidirectionnel (Gerasimova *et al.* 1995).

1.4.3.3 CP190

CP190 interagit avec Su(Hw) et avec Mod(mdg4) et est également indispensable à la fonction insulatrice de *gypsy* (Pai *et al.* 2004). Cette protéine contient un domaine BTB (Bric-a-brac, Tramtrack, Broad-complex) et trois domaines en doigt de zinc. Une mutation l'invalidant est létale. Bien que CP190 puisse lier l'ADN avec une faible affinité, Su(Hw) est nécessaire au recrutement de CP190 sur le rétrotransposon. De façon intéressante, CP190 est recruté de façon indépendante de Su(Hw) sur des centaines d'autres loci chez la drosophile (Pai *et al.* 2004).

1.4.3.4 Topors

Ce complexe interagit également avec la protéine Topors (Topoisomerase-I-interacting protein) qui est liée à la lamine nucléaire (Capelson & Corces 2005). Ces interactions induisent le regroupement d'éléments *gypsy* situés à une grande distance chromosomique dans des « foyers d'insulateurs » localisés à la périphérie nucléaire (Gerasimova *et al.* 2000). Cette localisation périphérique n'est cependant pas indispensable à la fonction insulatrice de cet insulateur (Xu *et al.* 2004).

1.4.4 Les insulateurs CTCF

Le premier insulateur étudié chez les vertébrés est l'élément *cHS4* localisé dans la LCR du locus *β -globin* du poulet (Chung *et al.* 1993). L'activité bloqueur d'activateur de cet élément dépend de la liaison de la protéine CTCF (Bell *et al.* 1999), une protéine à doigt de zinc conservée de la drosophile à l'humain (Moon *et al.* 2005). Cette protéine a initialement été purifiée lors d'une étude recherchant les protéines se liant à la région flanquant en 5' le gène *c-myc* chez le poulet (Lobanenkov *et al.* 1990). Il a été montré que

l'orthologue humain de CTCF se lie de la même façon à proximité du promoteur de *c-myc* dont il réprime l'expression (Filippova *et al.* 1996).

Chez les mammifères, le rôle de CTCF a été en particulier étudié sur le locus soumis à l’empreinte parentale *Igf2* qui est exprimé uniquement depuis l’allèle paternel. L’expression est contrôlée par la région ICR (Imprinted Control Region) qui possède des sites de liaison de CTCF. Dans l’allèle maternel, l’ICR est lié par CTCF qui bloque l’activation et empêche l’expression du locus. Dans l’allèle paternel, les îlots CpG de l’ICR sont méthylés, ce qui empêche la liaison de CTCF et permet l’activation de la transcription (Bell & Felsenfeld 2000; Hark *et al.* 2000). Cet exemple est important car il est l’un des premiers où l’activité de bloqueur d’activateur d’un insulateur a été montrée sur un locus endogène et non par l’insertion de transgènes (voir Fig. 3.6).

Le complexe *bithorax* de la drosophile contient tout un ensemble de séquences régulatrices qui organisent l’expression spatio-temporelle de trois gènes essentiels au développement. Les interactions entre ces séquences sont régulées par plusieurs éléments insulateurs (*Mcp*, *Fab-6*, *Fab-7* et *Fab-8*). Parmi eux, *Fab-8* interagit avec l’orthologue de la protéine CTCF des vertébrés. Les mutations de cette protéine sont létales et provoquent d’importants défauts homéotiques (Mohan *et al.* 2007). CTCF est aussi recruté sur les insulateurs *Mcp* et *Fab-6*. *Fab-7* est quant à lui reconnu par GAF qui est nécessaire à la fonction insulatrice de cet élément (Schweinsberg *et al.* 2004) (voir ci-dessous).

1.4.5 Le facteur GAGA

La protéine GAF (pour revue : Adkins *et al.* 2006) est le produit de l’expression du gène *Trl* qui est nécessaire à l’expression normale de plusieurs gènes homéotiques (Biggin & Tjian 1988). L’étude de la localisation de GAF sur les chromosomes polytènes montre que cette protéine se lie à de nombreux sites dans l’euchromatine (Granok *et al.* 1995). Parmi eux se trouve l’insulateur *Fab-7* qui nécessite la liaison de GAF pour présenter une propriété insulatrice (Schweinsberg *et al.* 2004). La délétion du site de liaison de GAF en amont du gène *hsp26* induit une perturbation de la structure chromatinienne et une perte de l’expression inductible du gène (Lu *et al.* 1993). Ce résultat suggère un important rôle de GAF dans le maintien d’une structure ouverte permissive à la transcription. De plus, la liaison de GAF est nécessaire à la pause de la polymérase sur le gène *hsp70* et à la transcription de ce gène (Lee *et al.* 1992). Cette activité semble venir du fait que GAF permet l’ouverture de la chromatine nécessaire au recrutement de la polymérase II. En effet, GAF a la capacité d’ouvrir la chromatine reconstituée *in vitro* (Tsukiyama *et al.* 1994) et modifie la PEV (rôle *E(var)*, Farkas *et al.* 1994). Plus récemment, il a été montré que GAF est nécessaire à l’éviction massive des nucléosomes du locus *hsp70* observée après un choc thermique. L’ensemble de ces données montre un important rôle de GAF dans l’ouverture de la chromatine aux promoteurs. Très récemment, l’homologue de GAF chez les vertébrés, c-Krox/Th-POK a été décrit (Matharu *et al.* 2010).

1.4.6 CP190

La protéine CP190 a d’abord été trouvée dans le complexe *gypsy* (voir ci-dessus) mais elle a par la suite été associée à d’autres insulateurs. Il a notamment été montré qu’elle interagit également avec CTCF (Mohan *et al.* 2007; Gerasimova *et al.* 2007). Des études de ChIP à l’échelle du génome montre que ces protéines colocalisent sur des centaines de sites. C’est également le cas de la protéine BEAF qui se retrouve très fréquemment sur des sites occupés par CP190 (Bartkuhn *et al.* 2009; Bushey *et al.* 2009). Ces données suggèrent un important rôle pour CP190 dans la régulation de la fonction insulatrice et l’existence d’un mécanisme commun d’insulation pour Su(Hw), BEAF, et CTCF. GAF n’interagit pas avec CP190 mais interagit avec Su(Hw) (Melnikova *et al.* 2004) qui elle

interagit avec CP190, et pourrait donc avoir un mécanisme d'action similaire aux autres protéines insulatrices.

1.4.7 les différents modèles d'insulation

Pour expliquer les propriétés des insulateurs, plusieurs modèles ont été proposés (pour revue : [Bushey *et al.* 2008](#), et Fig. 1.35).

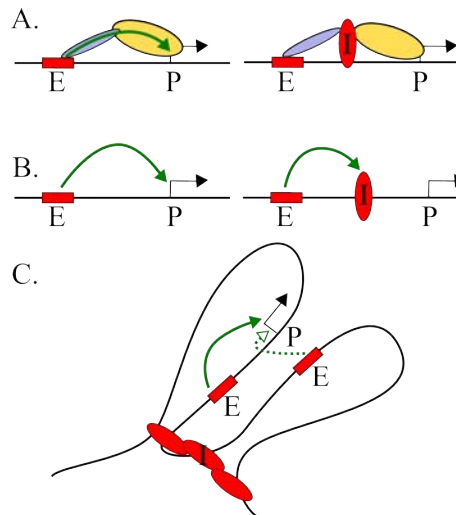


Figure 1.35 – Les différents modèles d'insulation. A. La barrière physique où l'insulateur (I) empêche les protéines impliquées dans la communication entre l'activateur (E) et le promoteur (P) d'interagir. B. Le leurre de promoteur où l'insulateur empêche l'activateur d'atteindre sa cible en le séquestrant. C. Les boucles de chromatine où un activateur ne peut interagir avec sa cible que s'il est situé dans la même boucle.

1.4.7.1 Barrière physique

La première idée est que les séquences insulatrices agissent comme une barrière physique bloquant la communication entre un activateur et un promoteur ou la propagation de la structure chromatinienne. Dans le premier cas, un signal moléculaire comme une polymérase en élongation part de l'activateur et rencontre l'insulateur qui le bloque et l'empêche d'atteindre le promoteur ([Kong *et al.* 1997](#)). En effet, l'insertion d'un élément terminateur de la transcription entre un activateur et un promoteur diminue l'efficacité de l'activateur ([Ling *et al.* 2004](#)). De plus, quand l'insulateur HS4 (Hypersensitive Site 4) est placé entre un activateur et un promoteur, la polymérase s'accumule au niveau de l'insulateur et non plus au niveau du promoteur ([Zhao & Dean 2004](#)). Il a cependant été montré que certains insulateurs sont situés dans un intron et peuvent inactiver un activateur situé en aval tout en permettant la transcription normale du gène par la polymérase qui peut donc passer à travers ces éléments ([Geyer & Corces 1992](#)). Ce modèle de barrière physique est généralement proposé pour expliquer le blocage de la propagation de la chromatine environnante et la compartimentation du génome en domaines indépendants ([Valenzuela & Kamakaka 2006](#)).

1.4.7.2 Leurre de promoteur

Les insulateurs pourraient fonctionner en interagissant avec un activateur et en le séquestrant, empêchant ainsi toute interaction avec un promoteur cible. C'est le modèle du « Promoter decoy » où l'insulateur est décrit comme un leurre imitant un promoteur et attirant à lui différents composants de la machinerie de transcription ainsi que l'activateur

(Geyer 1997). Ce modèle est soutenu par le fait que des séquences activatrices et insulatrices colocalisent (Yoon *et al.* 2007). Ce modèle ne permet cependant pas d'expliquer la directionnalité de l'effet bloqueur d'activateur qui n'a lieu que lorsqu'un insulateur est placé entre un activateur et un promoteur. En effet, dans ce modèle, un insulateur pourrait très bien séquestrer un activateur en étant placé de l'autre côté de l'activateur, et non entre ces deux éléments.

1.4.7.3 Boucles de chromatine

Le troisième modèle propose l'intervention de boucles de chromatine indépendantes. Les insulateurs interagiraient entre eux ou avec des éléments de la structure nucléaire pour former ces boucles. Cette hypothèse est confortée par de nombreuses données et tend à dominer les autres dans la communauté scientifique. Par exemple, des expériences de FISH ont donné des résultats qui concordent avec ce modèle en démontrant la formation d'une boucle entre deux insulateurs *gypsy*. De plus, l'ajout d'un troisième insulateur entre les deux autres engendre la formation de deux boucles plus petites (Byrd & Corces 2003). Dans le même ordre d'idée, des expériences de 3C (Chromosome Conformation Capture) ont montré que les séquences *scs* et *scs'* sont proches spatialement, et que la séquence entre ces loci forme une boucle (Blanton *et al.* 2003). Le rapprochement de ces séquences peut être effectué via une interaction directe entre BEAF et Zw5 (qui interagissent effectivement *in vitro*) ou par l'intermédiaire de CP190 (Fig. 1.34). Le modèle propose que les boucles ainsi formées se comportent comme des domaines indépendants. Ainsi, quand un activateur se trouve dans la même boucle qu'un promoteur, la communication entre ces éléments est possible et le gène est activé. Par contre, quand un insulateur est inséré entre ces éléments, alors ils se retrouvent dans des boucles différentes et la communication est impossible.

1.4.8 Rôle des insulateurs *in vivo*

Les insulateurs sont définis par des propriétés spécifiques testées par des expériences transgéniques. Bien que quelques exemples d'insulateurs ont des propriétés similaires à leur emplacement endogène, ces propriétés ne reflètent pas nécessairement le rôle réel de tous les insulateurs *in vivo*. Mon travail de thèse s'est intéressé à la compréhension du rôle de la protéine insulatrice BEAF dans l'organisation du génome de la drosophile en domaine chromatiniens indépendants, ainsi qu'à son implication dans la régulation transcriptionnelle.



Sommaire

2.1	Problématique	60
2.2	Frontière hétérochromatine-euchromatine et gènes hétérochromatiniens	60
2.2.1	Présentation de la région modèle	60
2.2.2	Déplétion par interférence à l'ARN	63
2.2.3	Effet sur la croissance des cellules	63
2.2.4	Effets sur l'expression des gènes euchromatiniens	63
2.2.5	Effets sur l'expression des gènes hétérochromatiniens	65
2.2.6	Effets sur l'expression du rDNA	65
2.2.7	Discussion	67
2.3	Prédiction des sites de liaison de BEAF et régulation des gènes euchromatiniens	68
2.3.1	Introduction	68
2.3.2	Article paru dans <i>PLoS Biology</i>	69
2.3.3	Rappel des principaux résultats	85
2.3.4	Retour sur le motif <i>dual-core</i>	85
2.3.5	Conclusion	86
2.4	Mécanismes de régulation transcriptionnelle	87
2.4.1	Article soumis à <i>Genes and Development</i>	87
2.4.2	Rappel des principaux résultats	119
2.4.3	Quelques données supplémentaires	120
2.4.4	Conclusion	123

CE chapitre présente la démarche suivie et les résultats obtenus pendant la durée de ma thèse. Je me suis tout d'abord intéressée au rôle potentiel de BEAF dans l'établissement d'une frontière entre l'hétérochromatine et l'euchromatine du chromosome X de la drosophile. J'ai ensuite étudié l'implication de BEAF dans la régulation transcriptionnelle de l'expression des gènes puis je me suis intéressée aux mécanismes mis en jeu dans cette régulation.

2.1 Problématique

Les insulateurs sont définis par deux propriétés : la première consiste à bloquer la propagation d'une structure chromatinienne pour préserver un environnement favorable à l'expression des gènes. Cette propriété suggère que les insulateurs pourraient avoir ce rôle de barrière au niveau des régions de transition entre l'hétérochromatine constitutive et l'euchromatine *in vivo* (Kellum & Schedl 1991). Dans la première partie de ma thèse je me suis intéressée au rôle potentiel de la protéine BEAF dans l'organisation de ces régions. Pour cela, j'ai utilisé comme modèle le chromosome X de la drosophile où il a été proposé que des répétitions de sites de liaison de BEAF (le BE28) pourraient être impliqués dans la séparation de l'hétérochromatine et de l'euchromatine (Cuvier *et al.* 2002). J'ai étudié l'influence de cet insulateur sur l'expression de gènes situés de part et d'autre, ainsi que sur la transcription du locus des ARN ribosomiques.

La seconde propriété des insulateurs est qu'ils peuvent bloquer la communication entre un activateur et un promoteur quand ils sont intercalés entre eux. Cette propriété suggère un rôle potentiel dans la régulation des gènes euchromatiniens. De plus, la localisation de BEAF sur les chromosomes polytènes n'est pas restreinte aux frontières hétéro/euchromatiniennes mais comprend aussi de nombreux loci euchromatiniens où le rôle de BEAF n'est pas déterminé (Cuvier *et al.* 1998). La seconde partie de ma thèse s'est intéressée à la caractérisation de ces sites et du rôle de BEAF dans ce contexte.

2.2 Frontière hétérochromatine-euchromatine et gènes hétérochromatiniens

Durant le début de ma thèse, je me suis intéressée à l'organisation des zones frontières entre euchromatine et hétérochromatine. Pour cela, j'ai utilisé comme modèle d'étude le chromosome X de *Drosophila melanogaster*.

2.2.1 Présentation de la région modèle

Le chromosome X est composé d'environ 22 Mpb d'euchromatine et de 15 Mpb d'hétérochromatine dont les 4 Mpb de séquences codant pour les ARN ribosomiques (rDNA) et 11 Mpb de séquences satellites dont les répétitions SAT III (Fig. 2.1). Les satellites de type III sont formés par des répétitions (30 000 fois) d'une séquence de 359 pb (Hsieh & Brutlag 1979). Ces séquences sont reconnues par les protéines D1, HP1 et par la topoisomérase II (Blattes *et al.* 2006).

Seule l'euchromatine est séquencée car le rDNA et les séquences satellites sont constitués de multiples répétitions qui ne peuvent pas être ordonnées aussi facilement que les séquences uniques. Une très petite partie de l'hétérochromatine du chromosome X a été séquencée (~200 Kpb sur plus de 15 Mpb) et est référencée sur Flybase¹ comme *XHet*. Cette région a pu être séquencée car elle est assez riche en gènes et pauvre en répétitions. Elle n'a pas encore été raccordée avec la séquence principale qui par ailleurs comporte un grand nombre de trous (on les repère aisément en regardant le niveau du pourcentage en G/C qui est nul dans ces régions ou en regardant la séquence où il y a de longues séries de « n »).

Une étude antérieure propose que des répétitions de l'élément BE28 situées dans la région entre l'hétérochromatine et l'euchromatine du chromosome X pourraient constituer une frontière physique permettant la compartimentation de cette région (Cuvier *et al.* 2002). Cet élément est lié par la protéine BEAF et possède les propriétés d'un insulateur. De façon intéressante, dans le mutant *um^{4h}*, le BE28 est relocalisé au bout du bras

1. <http://flybase.org/>; release 5.30

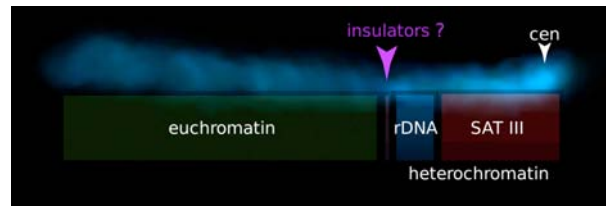


Figure 2.1 – Organisation du chromosome X. Le chromosome X est composé d'une région d'euchromatine et d'une large région d'hétérochromatine comprenant les répétitions codant pour les ARN ribosomiques. Entre ces deux régions, des insulateurs pourraient être impliqués dans le maintien de la compartimentation. L'indication « cen » localise le centromère, rDNA, la région codant les ARN ribosomiques et SAT III, les séquences satellites de type III.

euchromatinien (données non publiées, voir Fig. 2.2). Cette inversion induit une chute d'expression des gènes euchromatiniens relocalisés près de l'hétérochromatine (effet connu depuis longtemps pour *white* (Muller 1930) et qui s'étend jusqu'aux gènes *kirre* et *notch*), ce qui suggère un changement de la structure chromatinienne sur au moins 200 Kpb. Ce changement pourrait être en partie dû à la perte d'éléments frontières comme le BE28, s'il s'avère que cet élément est effectivement important pour maintenir la compartimentation dans cette région du chromosome X.

Pour vérifier si les répétitions BE28 ont un rôle dans la compartimentation du génome en domaines indépendants, je me suis intéressée à la régulation de l'expression des gènes présents de part et d'autre. Une recherche BLAST (Basic Local Alignment Search Tool) de la séquence du BE28 montre que cet élément se retrouve à environ 600 Kpb de l'extrémité de la partie séquencée du chromosome X. Cet élément est situé dans une région vide de gènes d'au moins 350 Kpb. Une partie de cette région (au moins 70 Kpb) n'est pas séquencée, probablement à cause de la présence de séquences répétées qui pourraient contenir d'autres sites de liaison de BEAF (un autre élément proche du BE28, le BE51, a déjà été caractérisé par Cuvier *et al.* (1998)). Entre cette région et la fin de la partie séquencée du chromosome X, une vingtaine de gènes sont répertoriés (Fig. 2.3, à droite). Le *XHet* non encore rattaché à la séquence principale mais situé entre cette région et le rDNA, contient un dizaine de gènes supplémentaires dont le gène *Su(f)* codant pour la protéine essentielle Suppressor of forked.

Du côté centromère-distal du BE28 se trouve un regroupement de huit gènes dont six sont en réalité trois répétitions en tandem de deux gènes. Sur la figure 2.3 (à gauche) et dans la suite, ils seront représentés par C2 (correspondant aux gènes *CG32500*, *CG32857* et *CG33502*) et C3 (*CG32819*, *CG32820* et *CG17450*). Ce groupe est séparé d'environ 80 Kpb des prochains gènes centromère-distaux.

En analysant cette région, je me suis proposée de vérifier le modèle suivant : les régions euchromatiniennes et hétérochromatiniennes sont séparées par des séquences spécifiques qui recrutent des protéines insultrices pour empêcher la propagation d'un type de structure dans l'autre. Au niveau du chromosome X, les répétitions BE28 liées par la protéine BEAF semblent de bons candidats pour cette fonction. Si ce modèle linéaire est exact, alors les gènes du côté du rDNA sont hétérochromatiniens alors que ceux de l'autre côté sont euchromatiniens. La déplétion de la protéine BEAF devrait perturber l'expression de ces gènes en supprimant la compartimentation hypothétique qui sous-tend le modèle classique de séparation entre eu- et hétérochromatine. On pourrait avoir soit une propagation de la structure hétérochromatinienne comme proposé par les modèles expliquant la PEV (voir §1.1.5.1), et une diminution globale de l'expression des gènes euchromatiniens, soit à l'inverse la propagation de la structure euchromatinienne et un changement d'expression des gènes hétérochromatiniens, soit une combinaison de ces deux effets.

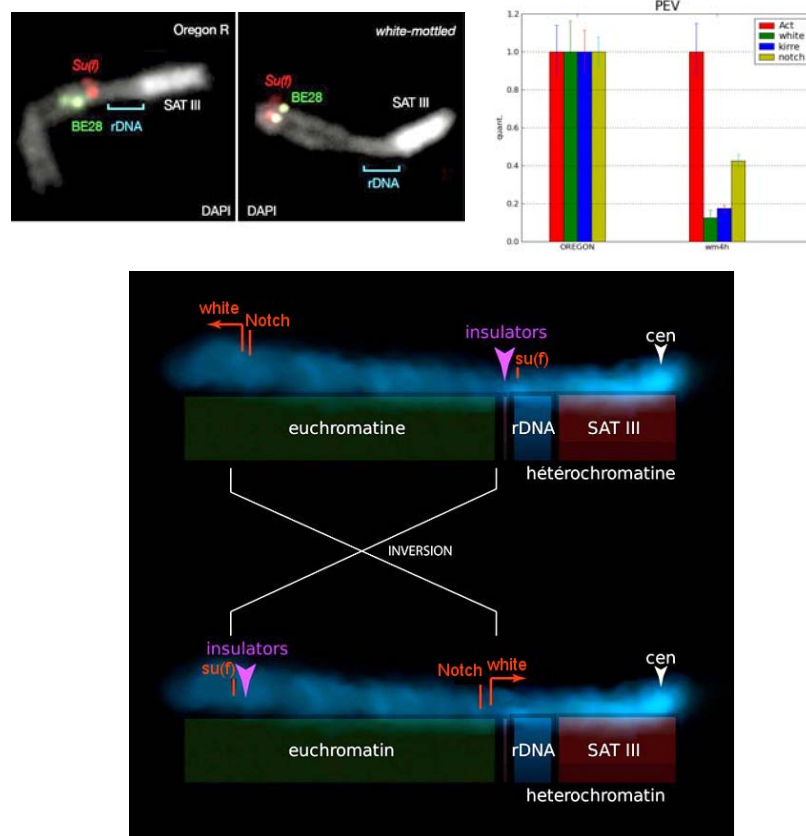


Figure 2.2 – Relocalisation du BE28 dans les mutants wm^{4h} et expression des gènes déplacés. L'expérience de FISH présentée montre le changement de localisation du BE28 dans la lignée wm^{4h} (expérience de Roxane Blattes). Une analyse par RT-qPCR montre que les gènes relocalisés à proximité de l'hétérochromatine voient leur expression affectée. Les barres d'erreur correspondent à un intervalle de confiance de 90%. Le schéma en bas permet de localiser les différents éléments déplacés par l'inversion.

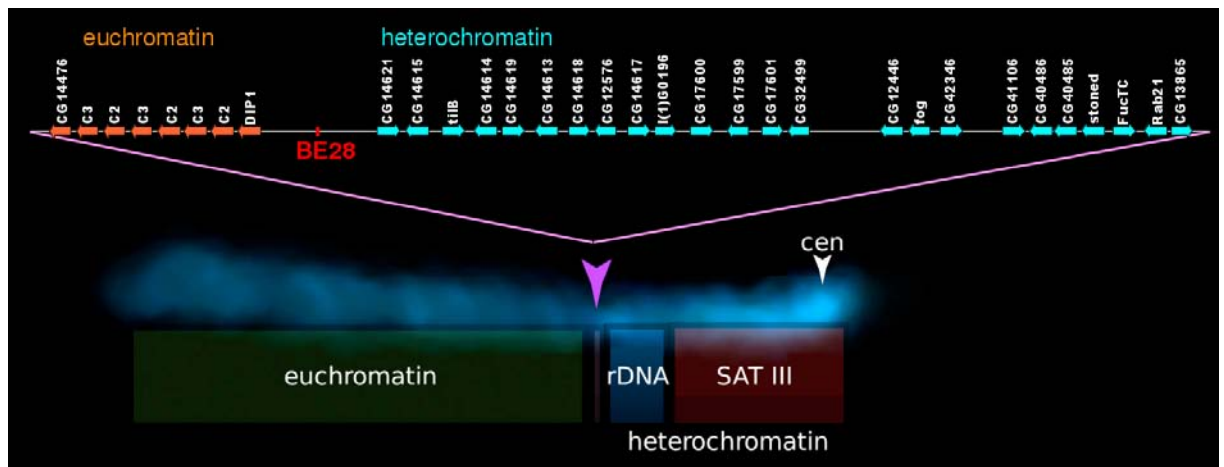


Figure 2.3 – Représentation des gènes proches des répétitions BE28. Le BE28 se situe près de la fin de la partie séquencée du chromosome X. Mon étude s'est focalisée sur les gènes représentés schématiquement de part et d'autre de cet élément.

2.2.2 Déplétion par interférence à l'ARN

Pour étudier le rôle de la protéine BEAF dans l'expression des gènes du locus modèle, j'ai utilisé une procédure de déplétion par interférence à l'ARN. Pour cela, des ARN double-brin (ARNdb) sont générés à partir d'ADN complémentaire (cDNA) à l'ARN messager de BEAF. L'ADN complémentaire est sélectionné et amplifié par PCR (Polymerase Chain Reaction) sur une banque de cDNA obtenue par rétrotranscription d'ARN totaux extraits de cellules en culture. Les amorces utilisées sont formées de la région complémentaire à l'ARN messager de BEAF (ciblant les deux formes A et B, voir §1.4.2.2) et de la séquence promotrice T7 qui est ainsi ajoutée en 5' et en 3' du cDNA (voir Matériel et Méthodes). Les cDNA ainsi amplifiés sont alors transcrits en ARNdb par la polymérase T7. Les cellules de drosophile (Kc ou S2) sont traitées en phase exponentielle de croissance. Elles sont incubées dans du milieu sans serum en présence des ARNdb, la déprivation étant un stress suffisant à l'entrée des ARN dans la cellule. Les cellules ainsi traitées sont remises en condition de croissance pendant 4 à 5 jours pour obtenir une bonne déplétion. L'efficacité du RNAi est mesurée par RT-qPCR (Reverse Transcription quantitative PCR) sur les ARN totaux extraits des cellules, ainsi que par *western blot* sur des extraits protéiques nucléaires (voir Fig. 2.4).

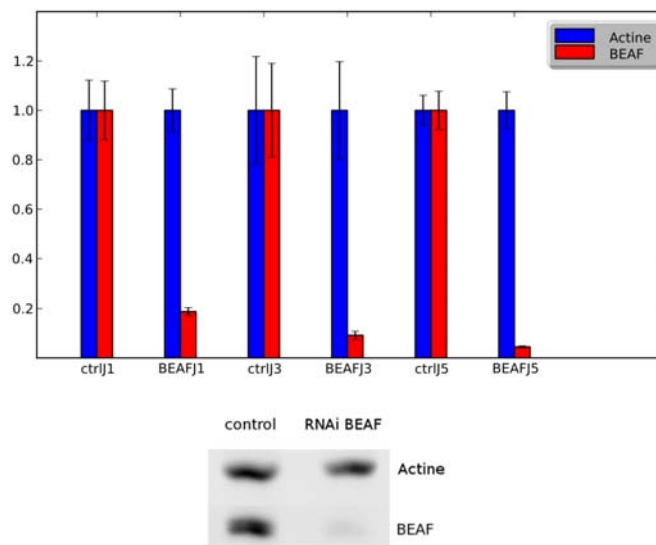


Figure 2.4 – Déplétion de la protéine BEAF par interférence à l'ARN. Le graphique du haut représente la quantité d'ARN messager de BEAF avant et après déplétion à des temps différents. Les barres d'erreur correspondent à un intervalle de confiance de 90%. Le *western blot* présenté en bas montre la quantité de protéine restant après 5 jours de traitement. La quantité d'actine est utilisée comme contrôle.

2.2.3 Effet sur la croissance des cellules

Un comptage quotidien des cellules a permis de montrer que la déplétion de la protéine BEAF entraîne un défaut de croissance des cellules après environ trois jours de traitement (Fig. 2.5). Ce défaut de croissance sera en partie expliqué grâce aux résultats obtenus dans l'étude suivante (voir 2.3).

2.2.4 Effets sur l'expression des gènes euchromatiniens

Des analyses par RT-qPCR sur les gènes du groupe euchromatinien ont été conduites afin de mesurer si le niveau d'expression de ces gènes étaient affecté par l'absence de BEAF. Si le modèle de propagation linéaire de la structure compacte est correct, on

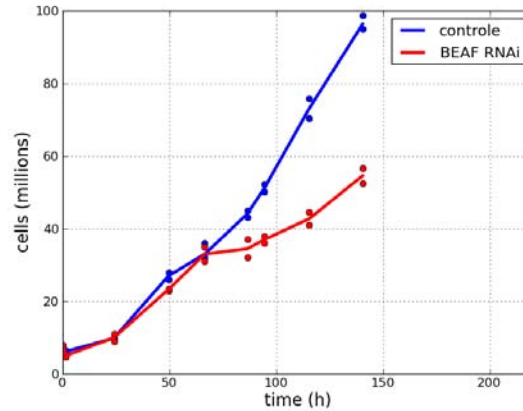


Figure 2.5 – Croissance des cellules après déplétion de la protéine BEAF. Les cellules sont comptées quotidiennement. Deux expériences indépendantes sont présentées.

devrait observer une diminution globale de l'expression de l'ensemble de ces gènes. Ces analyses ont donné des résultats inattendus qui prouvent que le modèle classique de simple propagation régulée par la présence de barrières est vraisemblablement trop réducteur. En effet, suivant les expériences de déplétion, une très grande variabilité de l'expression de ces gènes a été observée. Les effets de la déplétion sont toujours importants mais leur sens varie considérablement (voir les trois exemples de la Fig. 2.6). Ainsi un gène qui est surexprimé après la déplétion de BEAF dans une expérience peut être sous-exprimé dans une seconde expérience indépendante mais conduite dans les mêmes conditions. L'analyse de 10 expériences n'a fait ressortir aucune tendance si ce n'est cette grande variabilité (Fig. 2.7).

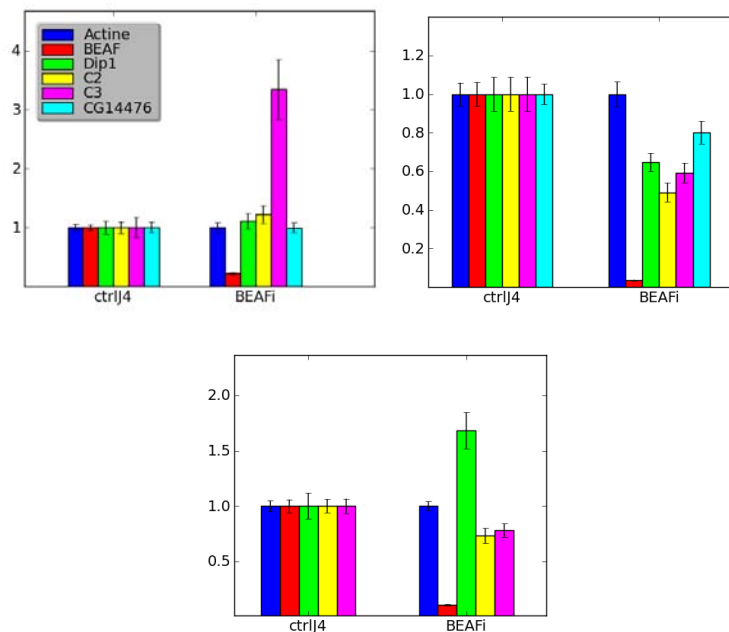


Figure 2.6 – Effet de la déplétion de BEAF sur l'expression des gènes du groupe euchromatinien. Trois exemples d'expériences indépendantes sont présentées. La quantification relative par rapport à l'échantillon contrôle est représentée. Les barres d'erreur correspondent à un intervalle de confiance de 90%.

Il est difficile de conclure sur ces observations. Toutefois, il semble clair qu'il n'y a pas, après déplétion de BEAF, de propagation de structure compacte répressive comme

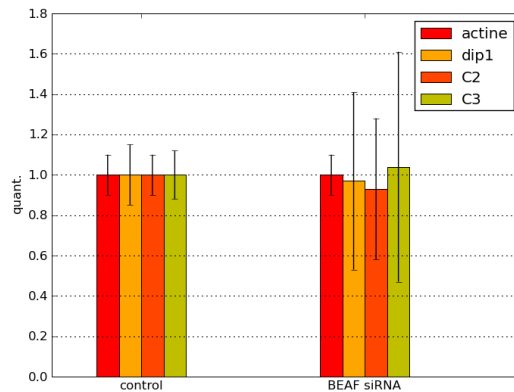


Figure 2.7 – Effet de la déplétion de BEAF sur l’expression des gènes du groupe euchromatinien. Les variations moyennes de dix expériences sont représentées ainsi que l’écart type correspondant.

dans le modèle classique expliquant la PEV. Ce modèle ne semble pas applicable dans le cas de ce locus. En revanche, il y a clairement des effets qui peuvent être interprétés comme une perturbation importante de la structure chromatinienne fortement dépendante de la quantité restante de protéine BEAF. Nous avons cherché un lien entre l’efficacité de la déplétion et les profils d’expression obtenus mais aucune tendance n’est ressortie. Un autre élément important à prendre en compte pour expliquer ces résultats est le fait que les gènes C2 et C3 sont répétés en tandem. Ainsi, lorsqu’on mesure la quantité d’ARNm, on mesure la somme de l’expression des trois gènes. Si ces gènes ne sont pas affectés de la même façon alors il n’est pas étonnant d’observer une grande variabilité. Cette observation n’est cependant pas valable pour le gène Dip1 qui n’est pas répété. Une possibilité intéressante serait que ces variations importantes reflètent simplement la variabilité naturelle de l’expression de ces gènes en l’absence du régulateur qu’est BEAF.

2.2.5 Effets sur l’expression des gènes hétérochromatiniens

La même approche de PCR quantitative a été utilisée pour étudier l’effet de la déplétion de BEAF sur les gènes hétérochromatiniens. La plupart des gènes étudiés ne montrent pas de changement de leur expression en l’absence de BEAF. C’est notamment le cas du gène *CG14621* qui est pourtant le plus proche du BE28 et qui devrait être le plus affecté si le modèle classique de propagation linéaire était correct. Cependant certains gènes pourtant plus éloignés voient leur expression diminuer de façon reproductible après ce traitement (Fig. 2.8).

Ainsi, il n’y a pas de changement global de l’expression de ces gènes. Cela va à l’encontre d’un modèle d’organisation linéaire avec propagation des structures actives ou répressives en l’absence de la barrière. Au contraire, certains gènes bien spécifiques sont affectés. Les données obtenues dans les études suivantes permettent de comprendre une partie de ces observations (voir §2.4.3.3).

2.2.6 Effets sur l’expression du rDNA

De façon assez paradoxale, les gènes les plus transcrits du génome sont localisés dans un contexte hétérochromatinien. En effet, le locus de transcription des ARN ribosomiques (rDNA) est entièrement hétérochromatinien (voir Fig. 2.1). En partant de l’idée que l’élément BE28 pourrait être impliqué dans la compartimentation de la chromatine du chromosome X, j’ai étudié l’effet de la déplétion de BEAF sur l’expression du rDNA. Pour cela, j’ai analysé par *northern blot* les transcrits issus du rDNA dans des conditions sauvages

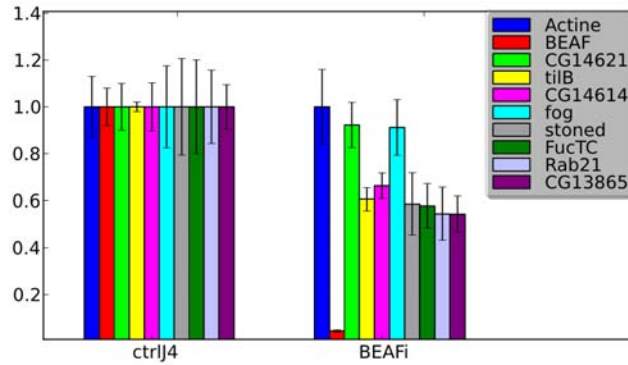


Figure 2.8 – Effet de la déplétion de BEAF sur l’expression des gènes du côté hétérochromatinien. La plupart des gènes ne sont pas affectés. Cependant certains gènes comme *CG14614*, *tilB*, *FucTC* ou *stoned* voit leur expression décroître après la déplétion de BEAF. Les quantifications relatives par rapport à l’échantillon contrôle sont représentées. Les barres d’erreur correspondent à un intervalle de confiance de 90%.

ou en l’absence de BEAF.

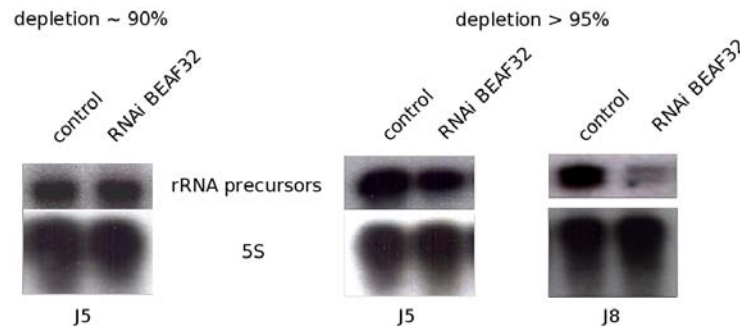


Figure 2.9 – Effet de la déplétion de BEAF sur l’expression du rDNA. Ce *northern blot* mesure la quantité de transcrite primaire 38S dans des cellules contrôles ou déplétées de BEAF. L’ARN 5S est produit par la polymérase III à partir d’un autre locus situé sur le bras droit euchromatinien du chromosome 2, indépendamment de la production des autres ARN ribosomiques (pol I). Il est ici utilisé comme contrôle. A gauche est présentée une expérience où la déplétion est moyenne. A droite se trouve une expérience avec une meilleure déplétion où la quantité de transcrits est mesurée après cinq ou huit jours de traitement.

Après cinq jours de traitement, le niveau de protéine BEAF est bas, mais peut varier d’une expérience à l’autre. En analysant les transcrits primaires issus du rDNA (38S), j’ai constaté qu’ils étaient parfois affectés mais pas à chaque expérience. Lorsque l’efficacité de déplétion est moyenne (diminution de ~90% de l’ARNm) il n’y a pas d’effet sur la transcription du rDNA. Cependant, lorsqu’elle est plus importante, alors on observe une petite diminution de la transcription du rDNA. De plus, quand on pousse la déplétion plus longtemps (huit jours) alors l’effet devient très net. Ces observations semblent indiquer que BEAF pourrait être impliqué dans la régulation transcriptionnelle du rDNA. Cependant vu le temps de réponse du système, on ne peut exclure que cette observation soit une conséquence indirecte de la dérégulation de gènes euchromatiniens impliqués dans le cycle cellulaire ou l’apoptose (voir §2.3). Le défaut de croissance observé en l’absence de BEAF peut être soit une cause soit une conséquence du défaut de production des ARN ribosomiques.

2.2.7 Discussion

Cette première étude visait à caractériser le rôle possible des répétitions BE28 dans la compartimentation du chromosome X. Il est apparu que ces éléments ne constituent pas une barrière physique bloquant la propagation des structures chromatinienne. En effet, les changements d'expression des gènes de part et d'autre de ces éléments en l'absence de la protéine isolatrice BEAF ne concordent pas avec un modèle linéaire de propagation simple de structures chromatinienne corrélées à l'expression génique.

L'expression des gènes du groupe centromère-distal est perturbée par l'absence de BEAF mais d'une façon bien plus complexe que par la simple propagation d'une structure répressive. Une étude récente de la relocalisation de HP1 dans des lignées *white-mottled* est en accord avec nos observations. Il a été montré que la localisation de HP1 s'étend effectivement sur ~ 200 Kpb dans la région normalement euchromatinienne près du point de cassure de l'inversion (Vogel *et al.* 2009). Par contre, cette extension ne se fait pas de façon homogène, contrairement à ce que suggère le modèle de propagation linéaire qui formerait un gradient d'hétérochromatine comme proposé précédemment (Rudolph *et al.* 2007). A l'inverse, le niveau de liaison de HP1 varie de façon importante et certaines régions sont nettement plus favorables à la liaison de HP1. C'est notamment le cas du locus du gène *white*, ce qui explique le phénotype des yeux de ces drosophiles.

Un modèle proposant la formation de boucles de chromatine semble plus approprié pour expliquer les effets observés au niveau de ces régions frontière. Il a d'ailleurs été proposé que le rDNA forme une boucle qui rapproche les séquences SAT III de l'euchromatine en interphase (Blattes *et al.* 2006, Fig. 2.10).

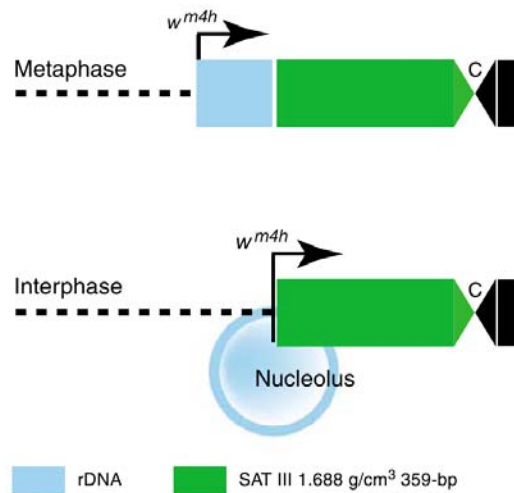


Figure 2.10 – Modèle proposant la formation d'une boucle du rDNA (Blattes *et al.* 2006). En interphase, le rDNA s'organise en boucle pour former le nucléole, ce qui rapproche les séquences SAT III de l'euchromatine.

Des contacts entre certaines répétitions hétérochromatiniennes et des loci euchromatiniens rapprochés artificiellement pourraient induire la mise en silence observée. La « variéation » pourrait alors s'expliquer par une grande dynamique de ces contacts. Cette idée peut être rapprochée de ce qui est observé dans le cas de l'inactivation en *trans* qui a été montrée pour certains loci comme dans le cas du mutant *brown Dominant*. Dans ce mutant, de l'ADN de séquences satellites hétérochromatiniennes a été inséré dans un des deux homologues du gène *brown*. Il en résulte une inactivation de l'autre allèle qui est relocalisé près de l'hétérochromatine centromérique (Henikoff & Dreesen 1989; Nisha *et al.* 2008). Cette observation prouve que la mise en silence d'un locus peut se faire par des contacts avec une région hétérochromatinienne sans propagation le long d'un chromosome (pour revue Talbert & Henikoff 2006). La formation de boucles de chromatine

pourrait reproduire ce type d'interaction en *trans* et expliquer l'absence d'un gradient d'hétérochromatine aux régions de transition.

Ces boucles pourraient être stabilisées par BEAF comme cela a été montré dans le cas de la boucle formée entre les éléments *scs* et *scs'* (Blanton *et al.* 2003). Dans ce cas, la déplétion de BEAF perturberait l'organisation tri-dimensionnelle de la région étudiée. Cela pourrait expliquer la grande variabilité de l'expression des gènes du groupe centromère-distal qui résulterait d'associations anormales, dynamiques et non contrôlées entre différentes régions chromatinienne. Nous nous sommes focalisés dans ce travail sur l'insulateur BE28 qui est précisément localisé sur la séquence du chromosome X. Il n'est cependant pas exclu que d'autres insulateurs (liés par BEAF ou d'autres protéines insulatrices, voir §3) puissent être impliqués dans l'organisation de cette région et fonctionner soit de façon autonome, soit en association étroite, ce qui pourrait participer au mécanisme de formation des boucles de chromatine en accord avec le modèle proposé ci-dessus. En effet, d'autres sites de liaison de BEAF sont présents dans la région centromère-proximale séquencée (voir §2.4.3.3). Ils pourraient interagir ensemble et/ou avec le BE28. Pour tester cette hypothèse, il serait intéressant de mettre en oeuvre des techniques de 3C qui permettraient de visualiser les interactions entre les régions chromatinienne au niveau de ces zones de transition. La comparaison de ces résultats dans des cellules sauvages et dans des cellules déplétées de BEAF pourrait apporter une meilleure compréhension des effets observés dans notre étude et, à terme, du rôle de BEAF dans l'organisation des régions frontières entre euchromatine et hétérochromatine. En plus des sites de liaison de BEAF présents dans la région centromère-proximale séquencée, il est possible que d'autres insulateurs soient présents dans les trous non séquencés de cette région de transition et soient impliqués dans des interactions à plus grande distance. Le manque de données de séquences reste un obstacle important pour tester cette hypothèse.

Certains gènes du côté centromère-proximal du BE28 sont également affectés et une partie de ces données sera expliquée par les études qui suivent. En effet, cette région du génome possède d'autres sites de liaison de BEAF bien localisés (voir Fig. 2.14). Ainsi la régulation de ces gènes ne semble pas liée à la présence de l'élément BE28 mais plutôt à celle de sites de liaison de BEAF plus proches des gènes et qui régulent localement leur expression (voir §2.4.3.3). Ainsi, la protéine BEAF a à la fois des effets localisés en régulant précisément un gène et des effets à plus grande échelle observés notamment dans le cas du groupe de gènes centromère-distal.

L'effet sur le rDNA est difficile à appréhender compte tenu de l'absence de données sur la séquence de ces régions qui est un obstacle majeur à l'étude des loci hétérochromatiniens. En effet la connaissance des séquences entre les répétitions du rDNA pourrait permettre de vérifier s'il y a ou non d'autres sites de liaison de BEAF, ou d'autres insulateurs, plus proches, qui pourraient participer à la régulation de ce locus.

2.3 Prédiction des sites de liaison de BEAF et régulation des gènes euchromatiniens

2.3.1 Introduction

2.3.1.1 Contexte

L'étude de la localisation de BEAF sur des chromosomes polytènes montre que ce facteur se lie sur des centaines de sites dans le génome (Cuvier *et al.* 1998). Ces sites ne sont pour la plupart pas caractérisés et la fonction cellulaire de BEAF reste inconnue. L'étude du mutant BID (présenté au §1.4.2.2) a permis de montrer que BEAF est essentiel au maintien d'une structure chromatinienne ordonnée dans les chromosomes polytènes (Gilbert *et al.* 2006). Les propriétés insulatrices des quelques sites de liaisons connus

suggèrent que BEAF pourrait avoir un rôle dans la compartimentation du génome en domaines actifs et inactifs. La connaissance de l'ensemble des sites de liaison de ce facteur devrait permettre d'avoir une vision globale de l'organisation du génome et de comprendre le rôle des insulateurs. L'article présenté ci-après combine une approche informatique pour prédire les sites de liaison de BEAF avec une approche expérimentale qui permet de proposer un mécanisme par lequel BEAF régule l'expression de gènes essentiels.

2.3.1.2 Un travail en collaboration

Nos compétences concernant le traitement des cellules pour obtenir des déplétions efficaces par RNAi nous a amené à collaborer avec Olivier Cuvier avant son changement de laboratoire. Ainsi, j'ai mis en oeuvre l'ensemble des méthodes concernant le traitement des cellules en culture du travail présenté ci-après : la production d'ARNdb ciblant BEAF et DREF, le traitement par RNAi et la vérification par RT-qPCR de l'efficacité de la déplétion. J'ai également transfecté les cellules avec les différentes constructions pour l'étude du mécanisme de compétition entre BEAF et DREF, puis extrait les ARN pour les analyses subséquentes. Enfin, j'ai préparé les échantillons de chromatine utilisés pour les différentes expériences de ChIP.

2.3.2 Article paru dans *PLoS Biology*

Formal Correction: This article has been **formally corrected** to address the following errors.

1. The following information was missing from the Funding section: MH was supported by a fellowship from the Ligue Nationale Contre le Cancer (LNCC).

BEAF Regulates Cell-Cycle Genes through the Controlled Deposition of H3K9 Methylation Marks into Its Conserved *Dual-Core* Binding Sites

Eldon Emberly¹✉*, Roxane Blattes²†, Bernd Schuettengruber³†, Magali Hennion^{2,4}, Nan Jiang³, Craig M. Hart³, Emmanuel Käs², Olivier Cuvier⁴✉*

1 Physics Department, Simon Fraser University, Burnaby, British Columbia, Canada, **2** CNRS, Laboratoire de Biologie Moléculaire Eucaryote, Université de Toulouse, UPS, France, **3** Department of Biological Sciences, Louisiana State University, Baton Rouge, Louisiana, United States of America, **4** Institut de Genetique Humaine, Department of Genome Dynamics, CNRS, Montpellier, France

Chromatin insulators/boundary elements share the ability to insulate a transgene from its chromosomal context by blocking promiscuous enhancer–promoter interactions and heterochromatin spreading. Several insulating factors target different DNA consensus sequences, defining distinct subfamilies of insulators. Whether each of these families and factors might possess unique cellular functions is of particular interest. Here, we combined chromatin immunoprecipitations and computational approaches to break down the binding signature of the *Drosophila* boundary element–associated factor (BEAF) subfamily. We identify a dual-core BEAF binding signature at 1,720 sites genome-wide, defined by five to six BEAF binding motifs bracketing 200 bp AT-rich nuclease-resistant spacers. Dual-cores are tightly linked to hundreds of genes highly enriched in cell-cycle and chromosome organization/segregation annotations. siRNA depletion of BEAF from cells leads to cell-cycle and chromosome segregation defects. Quantitative RT-PCR analyses in BEAF-depleted cells show that BEAF controls the expression of dual core–associated genes, including key cell-cycle and chromosome segregation regulators. *beaf* mutants that impair its insulating function by preventing proper interactions of BEAF complexes with the dual-cores produce similar effects in embryos. Chromatin immunoprecipitations show that BEAF regulates transcriptional activity by restricting the deposition of methylated histone H3K9 marks in dual-cores. Our results reveal a novel role for BEAF chromatin dual-cores in regulating a distinct set of genes involved in chromosome organization/segregation and the cell cycle.

Citation: Emberly E, Blattes R, Schuettengruber B, Hennion M, Jiang N, et al. (2008) BEAF regulates cell-cycle genes through the controlled deposition of H3K9 methylation marks into its conserved *dual-core* binding sites. *PLoS Biol* 6(12): e327. doi:10.1371/journal.pbio.0060327

Introduction

Chromatin insulators/boundary elements (BEs) [1,2] are defined as sequences able to insulate a transgene from its chromosomal context and to block promiscuous enhancer–promoter interactions or heterochromatin spreading [1,3–5]. These elements are thought to subdivide the genome into functional chromosome domains, through their ability to cluster DNA loops [1,2] and to control the deposition of histone epigenetic marks [6–8] to regulate chromatin accessibility for gene expression [9–13].

No common signature and/or mechanism of action has been identified among characterized insulators/boundary elements [2]. Rather, several factors confer insulating activity by targeting different DNA consensus sequences in the known insulators. In *Drosophila*, insulating factors include dCTCF [14,15], Zw5 [16], boundary element–associated factor (BEAF) [17], and the well-characterized suppressor of Hairy-wing (Su(Hw)) [1,18,19], which targets hundreds of distinct, largely uncharacterized genomic sites [20–22]. Whether each of these factors and subfamily of insulators might possess distinct cellular functions is of particular interest.

BEAF blocks both enhancer–promoter communication [17,23–25] and repression by heterochromatin, as shown using reporter transgenes [5,25]. This insulating activity of BEAF was also evidenced by a genetic screen in *yeast* [4], confirming that, unlike de-silencing activity, BEAF binding

sites must bracket a transgene for insulation. The hundreds of BEAF binding sites have not been characterized in situ, however, and the cellular function of BEAF remains to be elucidated in vivo.

Here we have combined computational and experimental approaches to address the function of BEAF binding sites in vivo. We have identified $\approx 1,720$ BEAF *dual-core* elements genome-wide that share an unusual organization conserved over 600 bp. The dual-core signature consists of five to six BEAF binding motifs bracketing 200 bp AT-rich nuclease-resistant spacers. BEAF dual-cores juxtapose to hundreds of

Academic Editor: Tom Misteli, National Cancer Institute, United States of America

Received: August 25, 2008; **Accepted:** November 11, 2008; **Published:** December 23, 2008

Copyright: © 2008 Emberly et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abbreviations: AA, anacardic acid; BE, boundary element; BEAF, boundary element–associated factor; cdk, cyclin-dependent kinases; GOs, gene-class ontologies; H3K9me3, histone H3 tri-methylated on lysine 9; HAT, histone acetyltransferase; NPS, nucleosome-positioning sequence; Su(Hw), suppressor of Hairy-wing; TSS, transcriptional start site

* To whom correspondence should be addressed. E-mail: eemberly@sfu.ca (EE); cuvier@igh.cnrs.fr (OC)

✉ These authors contributed equally to this work.

† These authors also contributed equally to this work.

Author Summary

The genome of eukaryotes is packaged in chromatin, which consists of DNA, histones, and accessory proteins. This leads to a general repression of genes, particularly for those exposed to mostly condensed, heterochromatin regions. DNA sequences called chromatin insulators/boundary elements are able to insulate a gene from its chromosomal context by blocking promiscuous heterochromatin spreading. No common feature has been identified among the insulators/boundary elements known so far. Rather, distinct subfamilies of insulators harbor different DNA consensus sequences targeted by different DNA-binding factors, which confer their insulating activity. Determining whether distinct subfamilies possess distinct cellular functions is important for understanding genome regulation. Here, using *Drosophila*, we have combined computational and experimental approaches to address the function of the boundary element-associated factor (BEAF) subfamily of insulators. We identify hundreds of BEAF dual-cores that are defined by a particular arrangement of DNA sequence motifs bracketing nucleosome binding sequences, and that mark the genomic BEAF binding sites. BEAF dual-cores are close to hundreds of genes that regulate chromosome organization/segregation and the cell cycle. Since BEAF acts by restricting the deposition of repressing epigenetic histone marks, which affects the accessibility of chromatin, its depletion affects the expression of cell-cycle genes. Our data reveal a new role for BEAF in regulating the cell cycle through its binding to highly conserved chromatin dual-cores.

genes highly enriched in gene annotations regulating chromosome organization/segregation and the cell cycle. Accordingly, BEAF depletion leads to cell-cycle and chromosome segregation defects. Quantitative RT-PCR analyses further show that dual-cores regulate the expression of key cell-cycle genes including *cdk7* and *mei-S332*. These results are also reproduced in embryos expressing truncated *beaf* mutants, which abolish the proper targeting of BEAF to dual-cores and its insulating activity. Chromatin immunoprecipitation analyses show that BEAF acts by restricting the deposition of methylated H3K9 marks in dual-cores. Our data reveal a new role for BEAF in regulating chromosome organization/segregation and the cell cycle through its binding to highly conserved chromatin dual-cores.

Results

Breaking Down the Binding Code of BEAF to Dual-Cores

The DNA-binding activity of BEAF has been well-characterized in vitro [17,20,23,24]. Each subunit of the BEAF complex targets one CGATA motif. Point mutations within this consensus abolish both its binding and insulating activities. Clusters of three to four CGATA motifs can create high-affinity ($K_d \sim 10\text{--}25$ pM) BEAF in vitro binding sites, which we call single elements. A computational scan of the *Drosophila* genome revealed thousands of single elements, yet immunostaining analysis demonstrated that they were not good predictors for BEAF binding in vivo. For example, Chromosome 4 was found to contain hundreds of single elements, yet immunodetection analysis showed only three major BEAF signals on this chromosome (Figure 1A). Interestingly, statistical analysis showed that single elements were often organized in a pair-wise configuration. Genome-wide, 988 single elements form 494 so-called “dual-cores,” which harbor two separate clusters of three CGATAs, a

statistically significant result (p -value $\sim 1e-9$). Moreover, 1,226 additional “dual-core-like” elements have a second cluster of two (instead of three) CGATAs. These elements include all characterized BEAF insulators whose activity involves a second, lower-affinity CGATA cluster ($K_d \sim 400\text{--}600$ pM) where BEAF binding is abolished when the first high-affinity cluster is mutated [20,23].

Detailed analysis by alignment of all 1,720 dual-core and dual-core-like elements showed a highly organized distribution of their 12,058 CGATAs, which preferentially segregate into two clusters separated by spacers of approximately 200 bp (Figure 1B). For *scs'* and other characterized BEAF insulators, these spacers were found to be relatively AT-rich [20,24,26]. Scanning the 1,720 dual-cores for A+T content showed that they all harbor significant AT-rich (>70%) sequences in their spacers (Figure 1C, Figure S1). The remarkably conserved organization of dual-cores indicates that they likely correspond to a highly specific BEAF-binding signature.

We tested this possibility by assaying BEAF binding to dual-cores by chromatin immunoprecipitation (ChIP) and ChIP-on-chip (Figures 1D and 2). Based on the signals obtained with anti-BEAF antibodies, dual-cores are expected to be precipitated much more efficiently than single elements (Figure 1A). Indeed, ChIP analysis confirmed that single elements were not bound by BEAF (Figure 1D). In contrast, dual-cores from the 7C locus of the X chromosome were efficiently bound by BEAF (Figure 1D, probes 4 and 5), while nearby control sequences or single elements were not (probes 1, 2, 3, and 6). Altogether, 25 out of 25 dual-cores and dual-core-like elements assayed by ChIP were found to be efficiently bound by BEAF in vivo (Figures 1D and 2; unpublished data). The *actin* promoter region, which contains six unclustered CGATA motifs, was not bound by BEAF (Figure 1D; last row), indicating that the distribution of CGATAs in dual-cores, rather than the number of CGATAs *per se*, is important for BEAF binding. Furthermore, ChIP-on-chip analysis over 350 Kbp of the X chromosome strengthens our conclusions, as all major peaks corresponding to regions where BEAF binds in vivo fit into a dual-core or a dual-core-like element (hereafter called “dual-cores”, see black rectangles in Figure 2; see our database at <http://www.sfu.ca/~eemberly/insulator/> for a complete listing). We note that computer analysis occasionally retrieved minor peaks present in the shoulder of the major BEAF peaks (enrichment <2 ; red bars in Figure 2) that may be attributed to the cooperative binding of BEAF to additional CGATA motifs present in single elements juxtaposed to dual-cores (Figure 2, see black bars for “single”). However, no peaks were present in regions corresponding to dispersed single elements (Figure 2; see our database), confirming that they are not sufficient for BEAF binding. These results establish that BEAF elements organized into dual-cores indeed define a characteristic in vivo BEAF-binding signature (Figure 1E).

BEAF Dual-Cores Are Tightly Linked to a Discrete Set of Gene Ontologies

Analysis of the positioning of dual-cores relative to genes showed that they are preferentially associated with gene-dense regions. 545 dual-cores reside within 500 bp of promoter/transcriptional start sites (TSSs) (p -value = $6.7e-119$) (Figure 3A), and more than 850 are within 2,000 bp. As

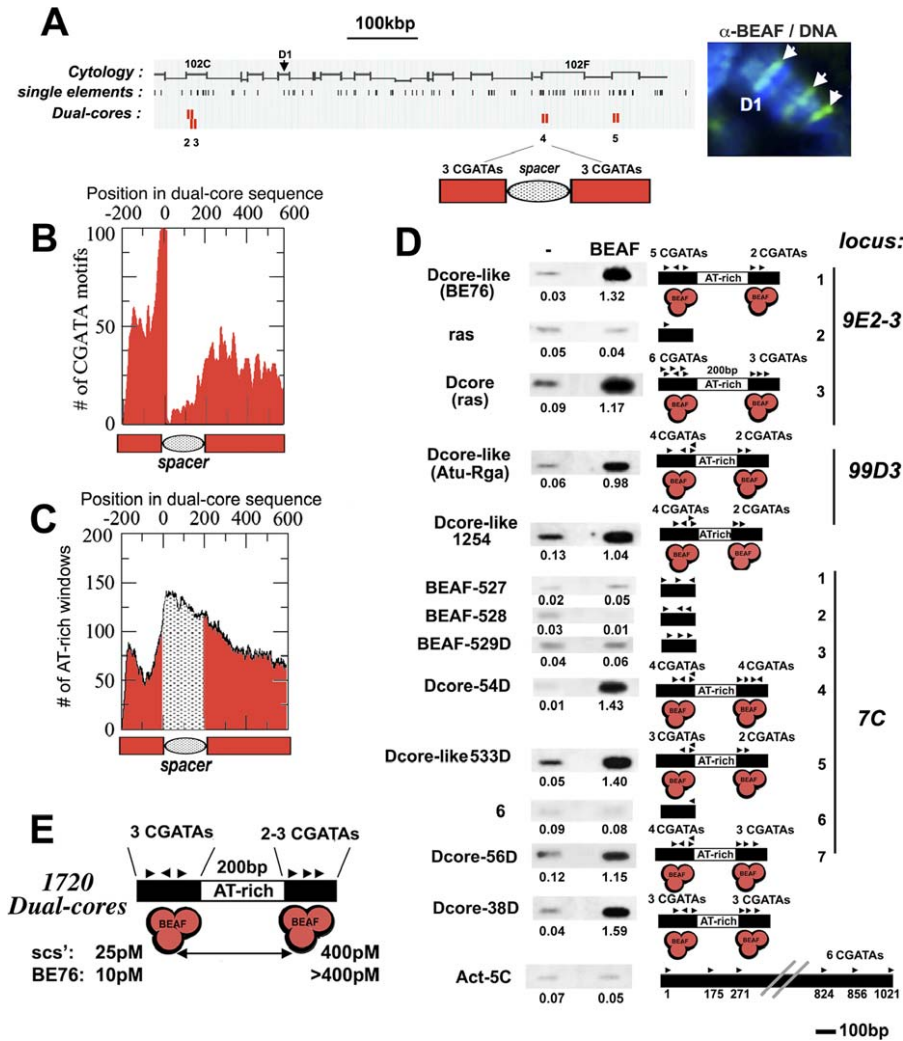


Figure 1. BEAF Dual-Cores Define a New Class/Family of Chromatin Elements

(A) BEAF single-element and dual-core predictions are shown in parallel with immunostaining with anti-BEAF antibodies for Chromosome 4 (1.2 Mbp; (A); D1 = recognizable band of Chromosome 4). Each single element contains ≥ 3 BEAF CGATA consensus motifs within 200 bp, and each dual-core corresponds to two juxtaposed single BEAF elements (see text).

(B,C) Statistical analysis of dual-cores. (B) Shows the distribution of 12,058 CGATA motifs of dual-cores into two clusters (3CGATAs \times 2) separated by a CGATA-free spacer. (C) Shows the localization of AT-rich 200-bp windows ($>70\%$ A+T) in the spacer. Position 0 is the location of the right-most CGATA motif in the first cluster of dual-cores. This analysis includes dual core-like elements, which contain two (instead of three) CGATAs within 100 bp in the second cluster.

(D) ChIP analysis with anti-BEAF or control IgG antibodies on DNA sequences corresponding to the indicated dual-cores or control regions. CGATAs are represented by arrowheads. Numbers below each blot represent the percentage of immunoprecipitated DNA over input genomic DNA as standard. (E) The BEAF dual-core signature.

doi:10.1371/journal.pbio.0060327.g001

dual-cores are preferentially distributed in pairs separated by approximately 5–15 kbp (p -value = $1.01e-33$; Figure 3B), the remaining elements might be found at the 3' borders of genes. However, we could not find any specific enrichment for dual-cores in the 3' UTR of genes (unpublished data), indicating that the clustering of dual-cores can be attributed to the clustering of genes/TSS rather than the bracketing of genes by dual-cores per se. These features (see our genome-wide database) raise the possibility that dual-cores might exert a function distinct from that of Su(Hw) binding sites, which rarely juxtapose the TSS of genes [21,22,27].

Strikingly, genes containing a dual-core near their promoter were statistically enriched in gene-class ontology (GO) groups that include the cell cycle, chromosome organization/

segregation, apoptosis, and sexual reproduction (p -value $< 1e-6$; Figure 3C). These essential cellular processes require constitutive regulation, whereas genes associated with non-constitutive processes such as sensing and behavior were not enriched for BEAF dual-cores (Table S1). Inspection of Table S1 also shows that other cell functions enriched in BEAF dual-cores include GOs that can be linked to phenotypes observed in *beaf* mutants [25,28,29], such as chromosome architecture, germ-cell and imaginal-disc development, and eye morphogenesis defects. We asked whether BEAF might be involved in regulating the cell-cycle and/or chromosome organization by siRNA-mediated depletion of BEAF from cells. Reduction of BEAF levels to background occurred from day 3–4 (Figure 3D), when defects in cell growth are first

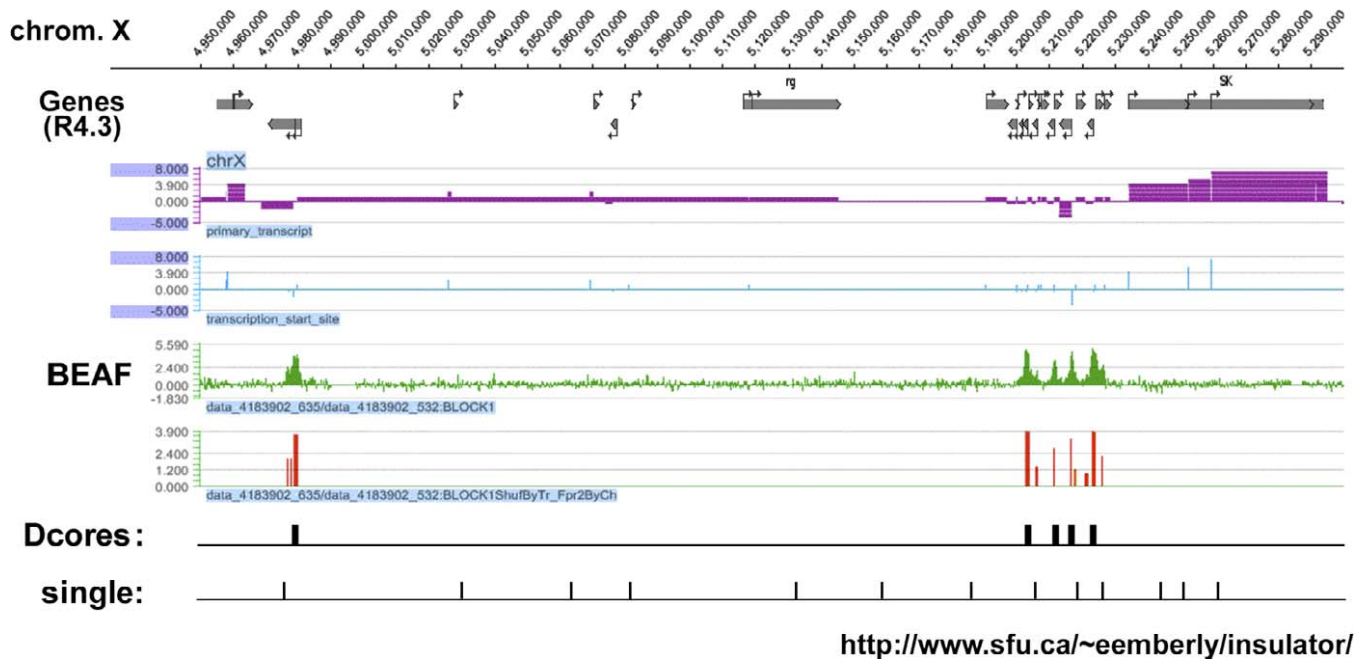


Figure 2. ChIP-on Chip Analysis Shows That BEAF Peaks Corresponding to *Dual-Cores*

The panel shows an alignment of ChIP-on-chip analysis using anti-BEAF antibodies (see the graph in green with red bars marking the position of significant peaks), with our dual-core database (Dcores, black rectangles) (<http://www.sfu.ca/~eemberly/insulator/>), over a region of the X chromosome (nucleotide positions 4,950,000 to 5,300,000). Single elements not forming dual-cores are also shown (single, black bars). Note that all five peaks fit into dual-cores. The second peak from the left is X-Dcore38_D, which juxtaposes the *cdk7* gene. Transcription start sites (blue bars) and primary transcripts (purple lines) are shown on top.
doi:10.1371/journal.pbio.0060327.g002

observed (Figure 3E). In addition, FACS and microscope analyses showed that BEAF depletion led to a significant and reproducible increase ($>3\times$) in the proportion of cells with 4N DNA content and with phenotypes typical of chromosome segregation defects (Figure 3F and 3G). These observations support our conclusion that the selective association of the corresponding GOs with closely linked dual-cores likely reflects a biologically significant localization.

BEAF Dual-Cores Control the Expression of Key Cell-Cycle Regulators

We next asked whether the phenotypes observed upon BEAF-depletion can be attributed to the loss of activity of BEAF dual-cores associated with 160 genes that control cell-cycle chromosome dynamics. These include *mei-S332* and *cdk7*, two major chromosome-segregation and cell-cycle regulators [30–32] whose promoter regions are bound by BEAF in vivo (Dual-cores 38/56, Figure 1D). Remarkably, further DNA-motif searches showed that the dual-cores associated with *cdk7* and *mei-S332*, and more generally with all genes belonging to the cell-cycle and/or chromosome dynamic GOs, also contain the TATCGATA consensus sequence recognized by DREF (*p*-value $\sim 2.4e-6$; Figure 4A). DREF activates hundreds of cell-cycle regulatory genes [33] and, importantly, might compete with BEAF for binding to the overlapping consensus [34]. Hence, DREF-regulated dual-cores may define a distinct regulatory subclass (Figure 4A, right).

To test how BEAF might affect the expression of genes associated with dual-cores that do or do not contain a DREF consensus site, we performed quantitative RT-PCR expres-

sion analysis from BEAF-depleted or control cells (Figure 4). BEAF depletion did not affect the expression of control genes (see Figure 4A, left), including those located near a single element (Figure 4B; *actin*, CG9745) where BEAF does not bind (Figure 1). The expression of all genes associated with a dual core lacking a DREF element was consistently found to be positively regulated by BEAF by approximately 4-fold to 5-fold (*CG1430*, *CG10946*, *CG1444*, *snf*, *ras*, *janus*; Figure 4B). These data are in complete agreement with previous work showing that BEAF has a positive effect on gene expression by de-repressing a transgene from surrounding chromatin [17,20,23,24]. In stark contrast, the expression of all genes associated with a dual-core harboring a DREF consensus, including *cdk7* and *mei-S332*, specifically increased by approximately 4- to 6-fold upon depletion of BEAF (Figure 4B; *CG32676*, *mei-S332*, *cdk7*, *CG10944*, *ser*). Accordingly, Western blot analysis showed that Cdk7 and Mei-S332 protein levels increased under these conditions (Figure S2). Therefore, two categories of dual-cores may be found. In those lacking a DREF consensus, BEAF positively regulates gene expression; in those that contain a DREF consensus, BEAF may prevent binding of DREF to its overlapping consensus, thereby controlling the activation of the associated cell-cycle and chromosome organization/segregation GOs.

Mutagenesis of the DREF Site from DREF Binding Dual-Cores Reveals the Positive Effect of BEAF

Quantitative RT-PCR analysis showed that DREF depletion resulted in a more than 10-fold down-regulation of *cdk7* (Figure 5), confirming the role of DREF as a transcriptional activator of this gene. To further characterize the respective

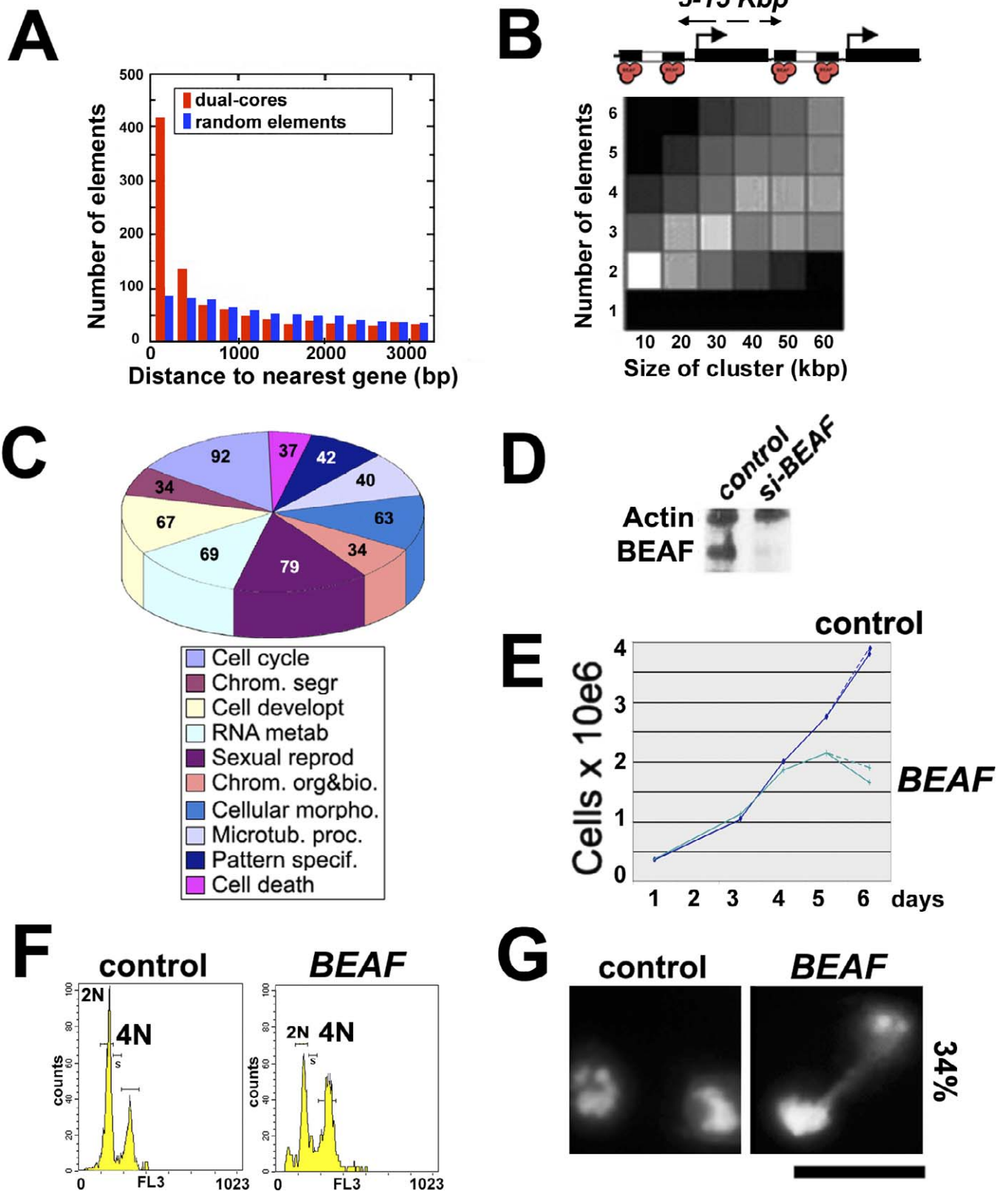


Figure 3. Genomic Features Associated with BEAF Dual-Cores

(A) Distribution of the distances between TSSs and dual-cores (red) compared to a theoretical distribution for randomly placed elements (blue) (see Materials and Methods). *y*-Axis, number of dual-cores; *x*-axis, distance to the nearest gene (in bp).

(B) Statistical analyses of the distribution of dual-cores. Dual-cores are preferentially organized in pairs separated by 5–15 Kbp as depicted on top (white square: *p*-value = 1.01e-33) most likely due to the organization of genes in clusters (see text). *y*-Axis, number of dual-cores; *x*-axis, distance between two dual-cores. Lighter (/darker) boxes represent the most (/less) significant *p*-values (see Materials and Methods).

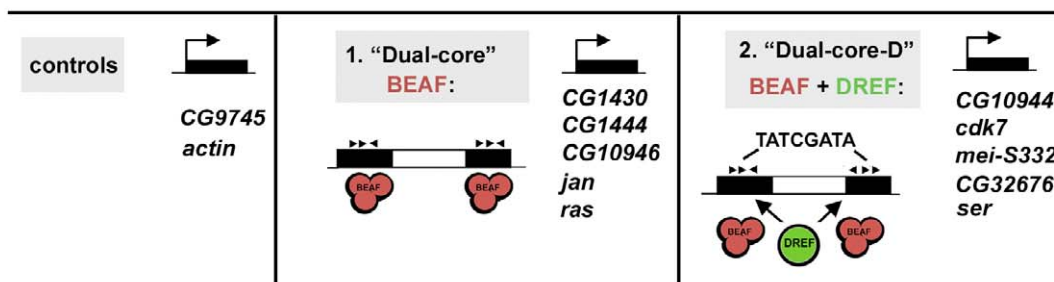
(C) Pie chart showing the most statistically significant (p -value $< 10e-6$) GOs for dual-core-associated genes (<500 bp of TSS) (see Materials and Methods).

(D) Western blotting of control or *beaf* siRNA-treated SL2 cells using anti-actin (dual-core-free gene) (see Figure 1D) and anti-BEAF antibodies. (E–G) Cell growth (E), FACS (F), and microscopy (G) analyses performed in parallel from the same BEAF-depleted or control cell samples shown in (D). In (E), the two curves for BEAF-depleted cells show the standard variation from two independent experiments. In (F), the y -axis indicates the number of cells (counts) and the x -axis the FL3 channel used to measure the staining of nuclei with propidium iodide. 4N indicates the number of G2/M cells with 4N DNA content. In (G), 34% indicates the increase in the percentage of BEAF-depleted cells with apparent chromosome segregation defects compared to control cells. DNA was stained with Hoechst. Bar, 10 μ m.
doi:10.1371/journal.pbio.0060327.g003

roles of BEAF and DREF in regulating cell-cycle regulatory genes by binding to dual-cores, we eliminated the DREF consensus from the dual-core associated with *cdk7* (*dre* mutant, Figure 5A) and transfected this construct or its wild-type version into cells depleted of BEAF or of DREF by

siRNA (Figure 5B). Because the *dre* mutant does not modify the CGATA BEAF consensus and still harbors the dual-core signature (2 \times 3CGATAs separated by the spacer; Figure S7B), this construct may be used to reveal the effect of the BEAF dual-core on the expression of *cdk7* independently of DREF.

A



B

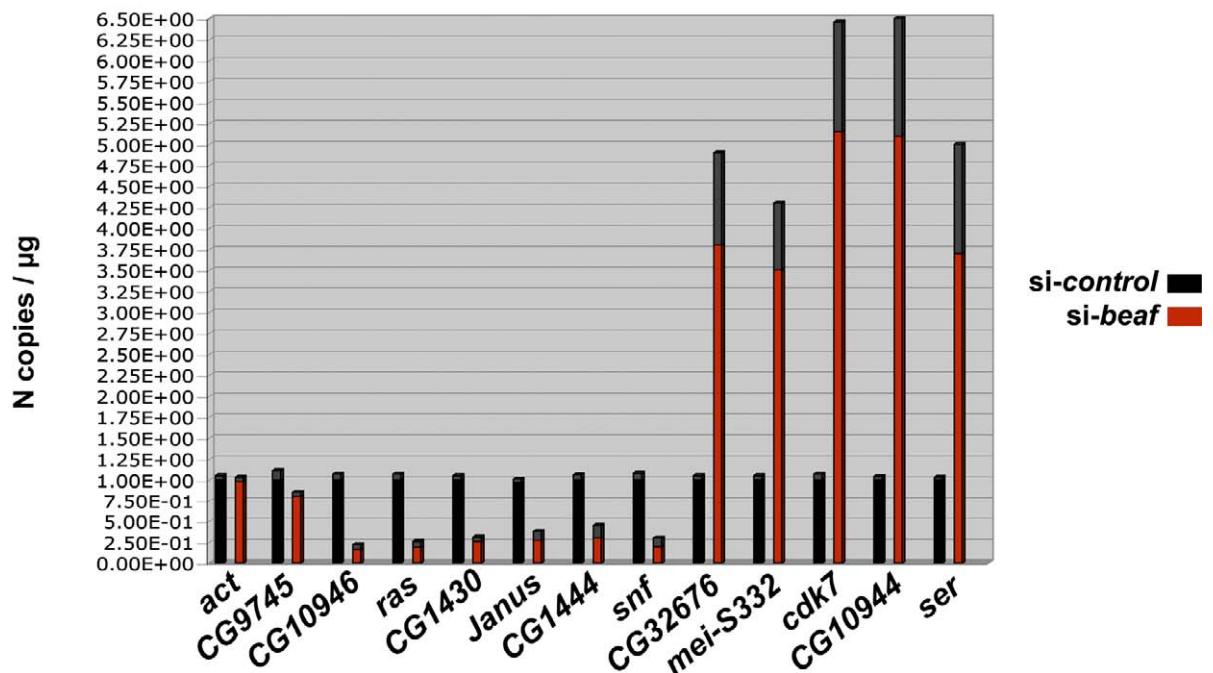
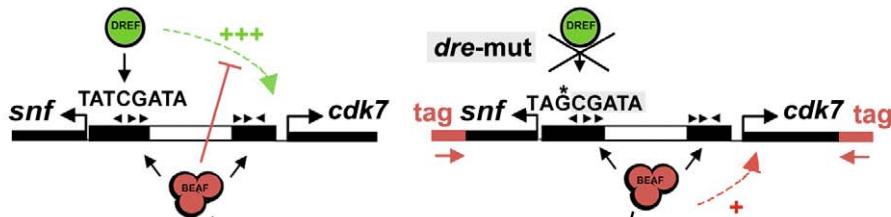


Figure 4. DREF Modulates the Activity of BEAF Dual-Cores Associated with Key Cell-Cycle Genes

(A) Schematic representation of the two distinct classes of BEAF dual-cores depending on the presence of an additional consensus (TATCGATA) for the transcription factor DREF (denoted “Dual-core_D”; see our database and text).

(B) Quantitative RT-PCR analysis of the expression of a control gene (left; *actin*, CG9745), of dual core-associated genes (without a TATCGATA DREF element, middle; CG1430, CG1444, CG10946, *janus* (*jan*), *rasberry* (*ras*)) or of dual core_D-associated genes (with a DREF element, right; CG10944, *cdk7*, *mei-S332*, CG32676, *serendipity* (*ser*)) in BEAF-depleted (red bars) or control cells (black bars). The y -axis shows the number of copies of amplification products per μ g of RNA normalized for each gene in control cells (see Materials and Methods), where $N=1$ corresponds to 8,450 copies of *actin*, 14,930 copies of CG9745, 29,300 copies of CG10946, 24,800 copies of *ras*, 23,020 copies of CG1430, 43,020 copies of *janus*, 29,370 copies of CG1444, 80,390 copies of *snf*, 48,393 copies of CG32676, 56,300 copies of *mei-S332*, 114,500 copies of *cdk7*, 32,830 copies of CG10944, and 49,620 copies of *serendipity*. The copy number was estimated using a standard curve obtained from more than three different quantities of genomic DNA. Experimental error from three independent experiments is denoted by the differentially colored (gray) portion at the top of each bar.
doi:10.1371/journal.pbio.0060327.g004

A



B

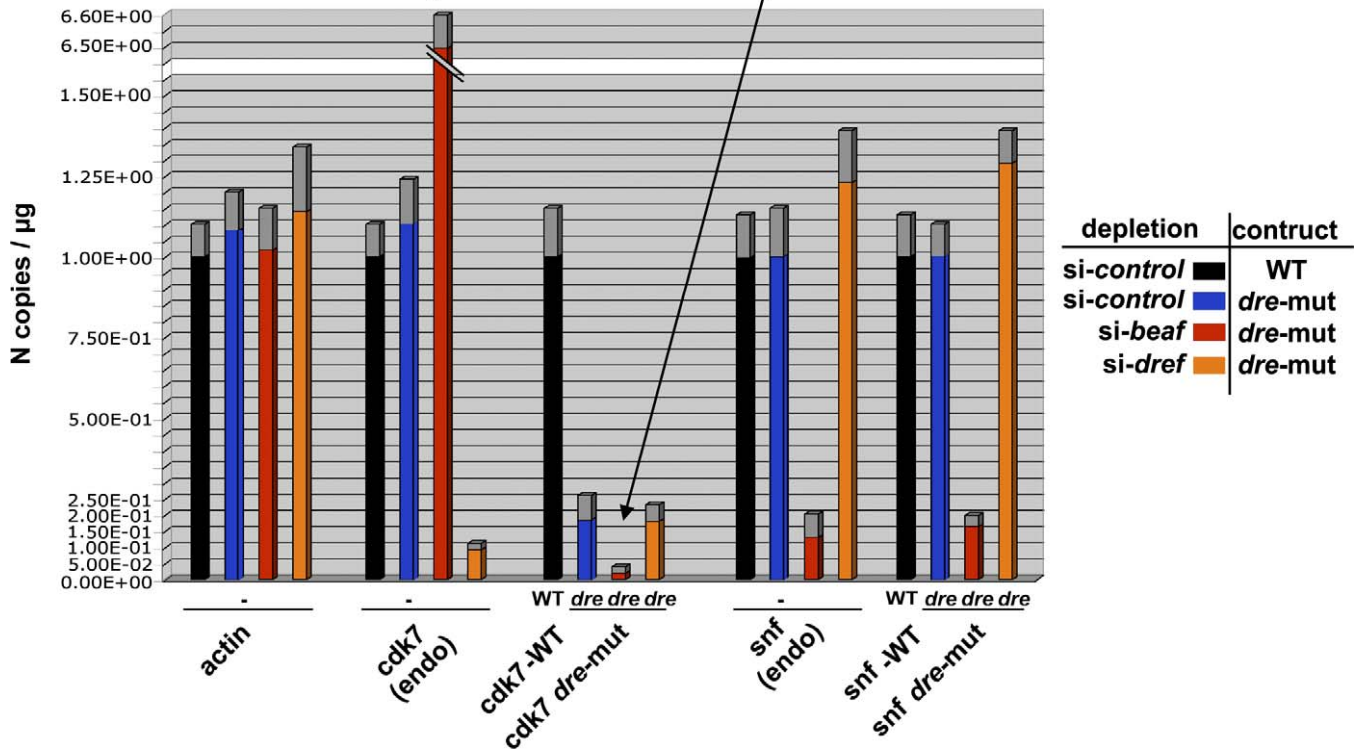


Figure 5. BEAF Dual-Cores Have a Positive Effect on Gene Expression Independent of DREF

(A) Schematic representation of the transfected constructs potentially regulated by BEAF–DREF competition for binding to the consensus TATCGATA present in the dual-core_D associated with *cdk7* and *snf* (X_Dcore 38D). Both the wild-type (left) and its DREF-response element (*dre*) mutant form (right) are shown. The transfected Dual-core harbors an additional sequence (tag) for specific amplification by Q-PCR. The +++ sign indicates the activation of *cdk7* by DREF, and the + sign represents the positive effect of BEAF obtained in the absence of the DREF binding site (B).

(B) Quantitative RT-PCR analysis in control (black and blue bars), BEAF-depleted (red bars), or DREF-depleted cells (orange bars) of the expression of endogenous genes (*actin*; *cdk7-endo*; *snf-endo*) and of transfected genes associated with the wild-type (WT) or the mutated version (*dre*; without DREF binding site) of Dcore_38D, as indicated. The y-axis shows the number of copies of amplification products per μg of RNA normalized for each gene in control cells, where the copy number $N=1$ corresponds to 9,680 copies of *actin*, 93,690 copies of *snf*, 138,900 copies of *cdk7*. Note that the expression of the transgenes (WT or *dre*) and of endogenous genes was measured from the same batch of cells. For transfected constructs, gene expression was normalized to the DNA copy-number in the input (for details, see Materials and Methods). Experimental error is denoted by the differentially colored (gray) portion at the top of each bar.
doi:10.1371/journal.pbio.0060327.g005

Importantly, mutating the DREF consensus site led to a down-regulation of *cdk7* (Figure 5B, *cdk7-mut*, blue bar), similar to what is found by depleting DREF. Strikingly, BEAF depletion further impaired the expression of *cdk7* by approximately 5-fold (Figure 5B, *cdk7-mut*, red bar) compared to the expression of the identical construct in control cells (Figure 5B, *cdk7-mut*, blue bar). We conclude that, although BEAF regulates DREF-mediated activation, it additionally positively regulates the expression of *cdk7*, as found for other genes associated with a dual-core lacking a DREF consensus. In support of this

conclusion, we obtained a similar result for *snf*, which is transcribed in opposite orientation relative to *cdk7* (Figure 5A). *Snf* is under the influence of the same dual-core as *cdk7*, yet its expression is not regulated by DREF (Figure 5B). However, BEAF depletion reproducibly impaired *snf* expression by approximately 6-fold, similar to what we obtained for *cdk7* in the absence of DREF. These results show that BEAF has a positive role on the expression of genes associated with dual-cores, in addition to its role in controlling activation by DREF.

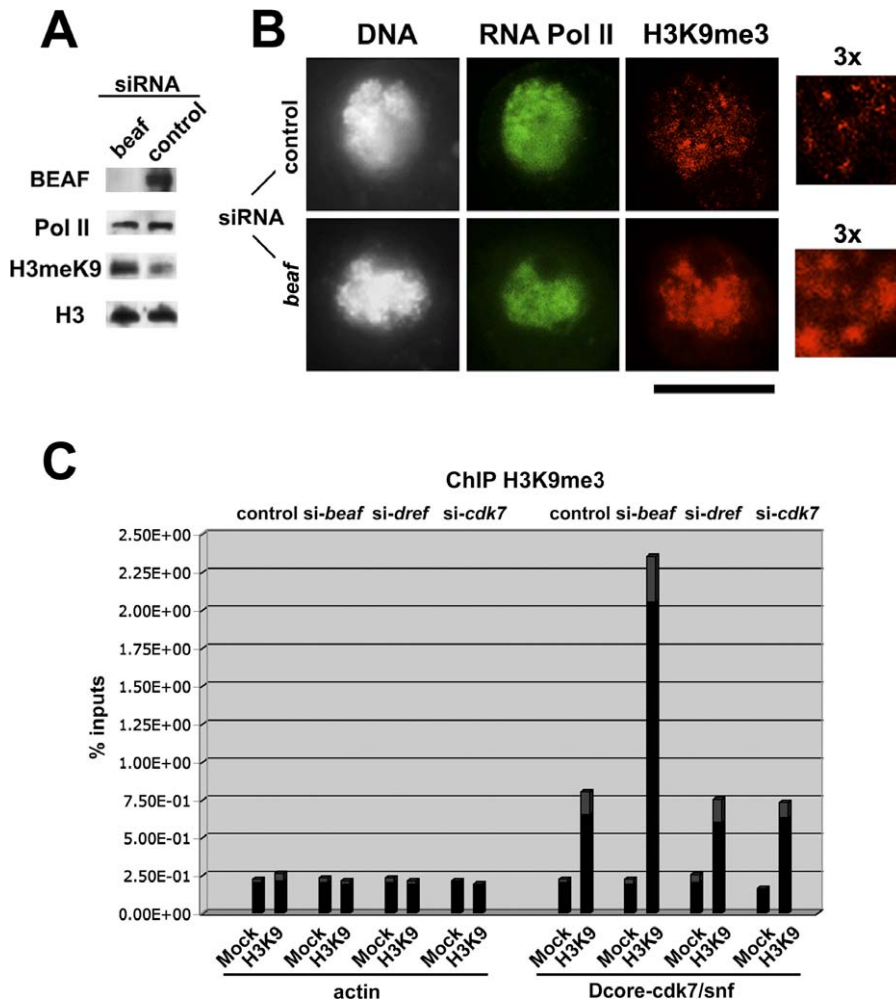


Figure 6. BEAF Restricts the Deposition of H3K9me3 Marks in Dual-Cores

(A,B) Western blot (A) and immunostaining (B) analysis with the indicated antibodies of chromatin-associated proteins in BEAF-depleted or control cells. Enlargements of confocal images are also shown (3 \times). Bar, 10 μ m.

(C) Quantitative PCR analysis following ChIP with mock control or with anti-H3K9me3 ('H3K9') antibodies in control (left), BEAF-depleted (second from left), DREF-depleted (second from right), or CDK7-depleted (right) cells. The graph shows the analysis of the *actin* control region and of dual-core_{38D} associated with *cdk7-snf*. Note that similar results were obtained for the dual-cores associated with *mei-S332* (unpublished data). The y-axis shows the percentage of material precipitated from inputs (see Materials and Methods for details). Experimental error is denoted by the differentially colored (gray) portion at the top of each bar.

doi:10.1371/journal.pbio.0060327.g006

BEAF Restricts the Deposition of H3K9me3 in Dual-Cores

BEAF insulating activity can protect a transgene from repression by chromatin [5,25]. The expression of genes positively regulated by dual-cores might implicate mechanisms similar to those required for insulation, and we asked whether BEAF might control the deposition of epigenetic marks, as shown for other types of insulators [7,35,36]. We tested this possibility by measuring the levels of histone H3 methylated on lysine 9 (H3K9me3), a characteristic mark of heterochromatin, as a function of BEAF depletion. The deposition of H3K9me3 was strongly increased upon BEAF depletion (Figure 6A). Double immunostaining analysis showed that this increase was specific, as RNA polymerase II, actin, or unmodified histone H3 levels were unchanged (Figure 6A and 6B, Figure S3A and S3B). Numerous and broader H3K9me3 foci not restricted to heterochromatin regions appeared in BEAF-depleted cells (Figure 6B, 3 \times panels; [37]), strengthening the view that H3K9me3 also acts

to influence gene expression in euchromatin [8,38,39]. ChIP-on-chip analysis confirmed that discrete H3K9me3 peaks are found in many promoter regions, including those associated with a dual-core (Figure S3C). However, these H3K9me3 peaks appear to be distinct from the broader and larger H3K9me3 peaks found in regions where genes are known to be repressed (e.g., *eye*, Figure S3C) and where the methylK27 mark is also present (not shown; B. Schuettengruber unpublished data).

We further tested if BEAF affects the deposition of H3K9me3 marks into dual-cores by performing ChIP analysis using anti-H3K9me3 antibodies on BEAF-depleted, DREF-depleted, or control cells (Figure 6C). BEAF-depletion led to a significant and reproducible increase of approximately 8-fold in H3K9me3 levels for the dual-cores linked to *snf-cdk7*, similar to that obtained for *mei-S332* and *CG1430*, and in stark contrast to the stable levels found for the *actin* control (Figure 6C; unpublished data). In contrast, no variation in H3K9me3

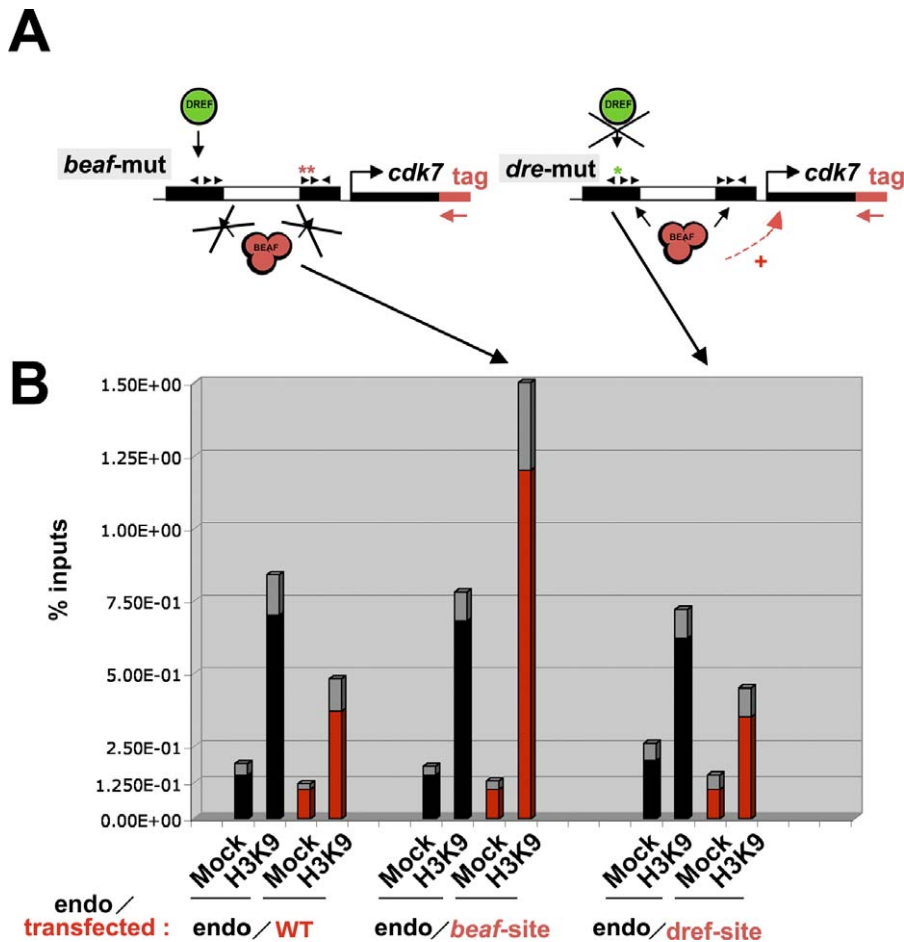


Figure 7. BEAF—But Not DREF—Binding Sites Restrict the Deposition of H3K9me3 Marks in Dual-Cores

(A) Schematic representation of Dcore_38D mutated for its DREF (*dre-mut*; see Figure 5) or BEAF (*beaf-mut*) binding sites. The mutated CGATAs (arrowheads) are represented by the *.

(B) Quantitative PCR analysis following ChIP with mock control or with anti-H3K9me3 antibodies of both transfected wild-type (WT) or mutated constructs (A) and endogenous ("endo") Dcore_38D from the same batch of cells. The y-axis shows the percentage of material precipitated from the inputs. For transfected constructs, the percentage was normalized to the DNA copy-number in the input (for details, see Materials and Methods). Experimental error is denoted by the differentially colored (gray) portion at the top of each bar.

doi:10.1371/journal.pbio.0060327.g007

levels could be found upon DREF depletion (Figure 6C), showing that this increase is specific to BEAF depletion. This result also rules out that the changes we observe overall could be due to off-target effects. Moreover, CDK7 depletion, which severely impaired cell-cycle progression (unpublished data), did not affect the levels of H3K9me3 (Figure 6C), indicating that their increase is not due to an indirect perturbation of the cell cycle upon BEAF depletion. Finally, H3K9me3 levels did not vary in control regions located a few kbp away from the dual-core, suggesting that BEAF controls the deposition of this mark locally (unpublished data). These results show that BEAF dual-cores are involved in blocking the deposition of H3K9me3 marks, fully consistent with their ability to positively regulate the expression of dual core-associated genes.

BEAF Positively Regulates Gene Expression by Restricting the Deposition of H3K9me3

To confirm that the observed increase in H3K9me3 levels is directly linked to the activity of BEAF, we introduced mutations in two of the CGATA motifs of the dual-core

associated with *cdk7* ("*beaf-mut*", Figure 7A) and transfected this construct or constructs harboring a wild-type dual-core or a dual-core mutated in the DREF site (*dre* mutant, Figure 7A) into cells. Quantitative PCR analysis of chromatin immunoprecipitated with anti-H3K9me3 antibodies showed that mutation of the BEAF site led to an increase in H3K9me3 levels of approximately 3.8-fold compared to wild-type or *dre* mutant constructs (Figure 7B), establishing that BEAF directly controls the deposition of H3K9me3. This did not affect the levels of H3K9me3 in the endogenous *cdk7* dual-core, as measured from the same batch of transfected cells, showing that the observed increase is indeed specific for the mutated dual-core. We conclude that BEAF serves to restrict the deposition of H3K9me3 marks into dual-cores.

The deposition of epigenetic marks is critical for regulating gene activity at the level of chromatin accessibility [9,12,13], which may account for the positive effect of BEAF on gene expression. We sought to determine whether BEAF-regulated deposition of H3K9me3 marks affects the expression of cell-cycle genes. BEAF-depleted or control cells were treated with anacardic acid (AA), a histone acetyltransferase (HAT)

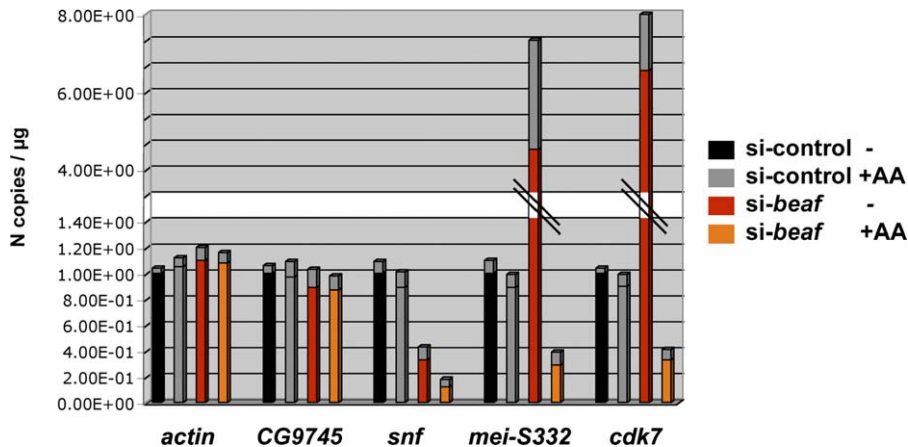


Figure 8. BEAF Dual-Cores Protect Genes from Chromatin Repression

Quantitative RT-PCR analysis of the expression of—from left to right—*actin*, *CG9745*, *snf*, *cdk7*, and *mei-S332* in control cells (black and grey bars, in the absence or presence of 5 μ M AA, respectively) or BEAF-depleted cells (red and orange bars, in the absence or presence of 5 μ M AA, respectively). The y-axis shows the number of copies of amplification products per μ g of RNA which was normalized for each gene in control cells, where N=1 corresponds to 11,200 copies of *actin*, 12,900 copies of *CG9745*, 75,430 copies of *snf*, 64,500 copies of *mei-S332*, and 129,800 copies of *cdk7*. Experimental error is denoted by the differentially colored portion (gray) at the top of each bar. doi:10.1371/journal.pbio.0060327.g008

inhibitor that globally affects gene expression by altering the accessibility of chromatin [40]. AA treatment did not affect the expression of either control genes or dual core-associated genes (compare grey and black bars in Figure 8). In contrast, AA severely compromised the activation of *snf*, *cdk7*, or *mei-S332* upon BEAF depletion compared to untreated BEAF-depleted cells (Figure 8; unpublished data). This result strongly supports a model whereby BEAF restricts the deposition of methylated H3K9 marks, thereby protecting the expression of dual core-associated genes from repression by chromatin (see Discussion).

Regulation of Gene Expression Involves the Cooperative Binding of BEAF to Dual-Cores

Are these variations in gene expression related to the cooperative binding of BEAF to the two clusters of CGATAs present in dual-cores? We sought to answer this question by using transgenic fly lines expressing the C-terminal BEAF self-interaction domain (BID in Figure 9A) under the control of a GAL4 activator. BID lacks the BEAF DNA-binding domain, impairing the insulating activity of BEAF [25] by preventing its cooperative binding to two nearby CGATA clusters (Figure 9B). Importantly, defects in expression of *cdk7*, *snf*, and/or *mei-S332* were highly similar in embryos expressing BID to that observed in BEAF-depleted cells (compare Figures 9C and 4B). This result supports our conclusion that BEAF binding is required to regulate these genes in vivo. It also suggests that the cooperative binding of BEAF to conserved *dual-cores*, which is abolished by BID, may be important for the regulation of gene expression by BEAF. Accordingly, cell functions enriched in BEAF dual-cores include GOs (Table S1) that correspond to phenotypes observed following expression of *beaf* mutants, which are lethal to flies [25], or to GOs found to genetically interact with these mutants [28].

Discussion

Results of our in silico analysis reveal \sim 1,720 BEAF dual-cores in the *Drosophila melanogaster* genome that share a

striking organization (Figure 1E). Genome-wide ChIP-on-chip analysis detects approximately 1,800 significant BEAF binding sites (C. M. Hart, unpublished observations), suggesting that our dual-core database encompasses most in vivo BEAF binding sites. The few (<100) additional peaks not included in our database but detected by ChIP-on-chip analysis may correspond to elements initially scored as single elements but whose organization is close to that of dual-cores. These rare exceptions are in part due to the computer stringency of the dual-core signature. For example, BEAF-1255 can be bound by BEAF in vivo (Figure S4), yet this element could not be scored as a dual-core because one out of five of its clustered CGATA motifs lies 3 bp outside the defined 100-bp window ('out' in Figure S4). Furthermore, approximately 10% of the minor BEAF sites are found in regions lacking any CGATA motifs, including the *scs* insulator (unpublished data) [16]. Since this region is not directly bound by BEAF, it is thus possible that some of the minor BEAF peaks are due to indirect interactions between BEAF and other insulator proteins, as previously suggested for the *scs'*-*scs* pair of insulators [16]. Other protein-protein interactions that regulate BEAF binding could also involve the splicing variant of the *beaf* gene itself, called BEAF-32A [23], which does not harbor the BEAF DNA-binding domain that recognizes clustered CGATA motifs. ChIP-on-chip analysis using antibodies that also recognize this isoform showed no additional major peaks (Figure S5, compare '-32A' with '+32A'), indicating that dual-cores constitute the main binding sites for both BEAF isoforms. Finally, we note that the BEAF-32A isoform is unlikely to play a major role in the activities described here, as its binding is dispensable for the insulating function of BEAF [20], and its expression is not essential for the development of embryos into adult flies [29]. Taken together, our results show that the BEAF dual-core signature is a bona fide mark that identifies a *cis*-regulatory element that regulates the expression of nearby genes.

Results of our experiments using both BEAF depletion in tissue culture cells and BID expression in vivo provide clear evidence for specific functions of the BEAF dual-cores,

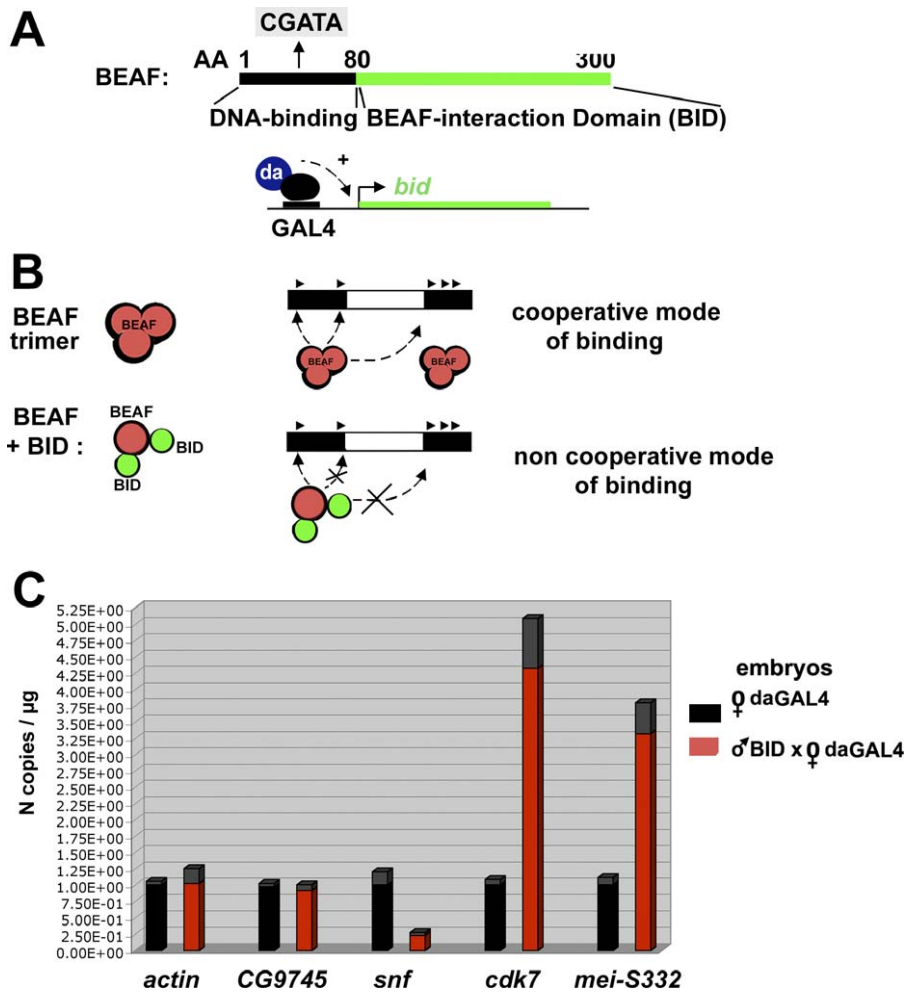


Figure 9. Regulated Gene Expression Requires Cooperative Binding of BEAF to Dual-Cores and Is Impaired by BID

(A) Schematic representation of BEAF. Its N-terminus is the DNA binding domain (AA 1–80) which targets one CGATA motif [17,23–25]. Its C-terminus, required for assembly of BEAF complexes, is called the BEAF self-interaction domain (BID; AA 80–280). As a result, expression of BID under the control of the GAL4-daughterless (da) activator in transgenic flies dominantly competes with the cooperative binding of BEAF [25], as represented in (B).

(B) Schematic representation showing how the cooperative binding of BEAF complexes to clustered CGATAs of dual-cores is compromised by expression of BID upon its assembly with BEAF subunits.

(C) Quantitative RT-PCR analysis in 4–8-hour embryos expressing a GAL4-driven BID transgene. The results are expressed as copy number of amplification products measured for each gene in embryos expressing BID or in embryos expressing the da-GAL4 driver alone. The y-axis shows the number of copies of amplification products per μ g of RNA normalized for each gene in control cells, where $N=1$ corresponds to 9,900 copies of *actin*, 11,200 copies of *CG9745*, 77,300 copies of *snf*, 59,400 copies of *mei-S332*, and 116,200 copies of *cdk7*. Experimental error is denoted by the differentially colored (gray) portion at the top of each bar. Note that BID expression leads to lethality at later stages of development [25].

doi:10.1371/journal.pbio.0060327.g009

reflected by a selective association with genes that control cell-cycle and/or chromosome organization/segregation. The competition between DREF and BEAF for binding to nested consensus sequences is also supported by ChIP analyses showing that DREF targets' identical sites [34] clearly enriched nearby genes associated with the cell cycle and chromosome dynamic GOs (Figure S6; unpublished data). Thus, while DREF levels increase at the G1/S transition to activate *mei-S332* and *cdk7* within the appropriate window for cell-cycle progression [30–32], BEAF may further facilitate this activation by restricting the deposition of H3K9me marks. Indeed, over-expressing BEAF was shown to reduce the phenotypes related to cell-cycle progression in flies that over-express DREF [33], supporting a role for BEAF in controlling the cell cycle. Such a model is also supported by our observation that AA treatment strongly represses these

genes in BEAF-depleted cells and that mutation of the BEAF-binding site in a dual-core results in a local increase in H3K9m3 levels. In addition, computer analysis of micro-array expression data for *Drosophila* embryos during early development shows that the 545 genes associated with dual-cores are positively correlated with *beaf* expression (Figure S7A), in contrast to genes unlinked to these elements (p -value $\sim 3e-17$ according to the Kolmogorov-Smirnov test). This strict correlation further indicates that BEAF has a global positive role on gene expression genome-wide, and similar analyses did not reveal any significant correlation change between genes whose TSS is closely juxtaposed (<100 bp) to dual-cores, including *snf* or *cdk7* (Figure S7B), compared to genes whose TSS is more distant (500 bp). Accordingly, the cell-cycle and chromosome dynamics GOs that include *cdk7* and *mei-S332* are enriched for positively correlated genes (see our

database for a detailed list). Taken together, our results show that BEAF could play an important role in chromosome organization during the cell cycle through a regulated switch involving the BEAF–DREF competition: According to such a mechanism, BEAF would restrict the deposition of H3K9me₃, allowing dual-core-associated genes to remain in a potentially active state, while controlling the time of activation of cell-cycle GOs by DREF. Accordingly, BEAF depletion leads to down-regulation of genes associated with a dual-core lacking a DREF element (CG10946, *ras*, CG1430, *Janus*, CG1444), but to increased expression of CG32676, *mei-S332*, *cdk7*, CG10944, and *ser*, which are under the control of DREF-associated dual-cores (Figure 4). In the latter case, the apparent contradiction between the positive—restriction of H3K9me₃ deposition—and negative effects of BEAF can be reconciled by our results showing that BEAF controls the activation of these genes by DREF. BEAF depletion relieves the competition for binding by DREF, leading to the increased expression of *cdk7* or *mei-S332* in spite of an increased deposition of H3K9me₃ marks under these conditions. Mutating the DREF or BEAF binding sites of DREF-associated dual-cores (Figures 5 and 7) allows for distinguishing between these different effects on the expression of linked genes.

It is intriguing that the spacers of dual-cores are well-conserved. One possibility is that they may be preferentially bound by a nucleosome, as recently shown for CTCF insulators [41]. Supporting this idea, the known dual core-spacers correspond to nuclease-resistant “cores”, between two nuclease-hypersensitive sites (BE76, *scs'*) [20,24,26] (Figure S8), where a nucleosome may be present (C. M. Hart, unpublished observations). Indeed, we found that dual core-spacers fall within predicted nucleosome-positioning sequence (NPS) databases [42–44], as indicated by NPS/dual-core sequence alignments (Figure S8; not shown), possibly accounting for the conserved organization of dual-cores. Our results further suggest that the cooperative binding of BEAF across these AT-rich spacers may be important for BEAF function. Indeed, expression of BID, which prevents its cooperative binding across the spacers, mimics the effect of BEAF depletion on the expression of dual-core-associated genes, as also found by mutagenesis of two CGATA motifs from one dual-core cluster. However, BEAF still efficiently binds *in vivo* to the few dual-cores that harbor a shorter spacer (<150 bp; e.g., see Dual-core 1,254, Figure 1; unpublished data), indicating that the conserved dual-core-spacer is dispensable for BEAF binding. Recent reports have shown that gene expression is differentially regulated through nucleosome positioning in several species [12,13,42,43]. Positioned nucleosomes may restrict promoter accessibility in yeast, and pausing of RNA polymerase II facing the +1 nucleosome may be regulated through nucleosome positioning in *Drosophila* [44]. Similarly, dual-cores are also closely associated with TSSs, and a potential link to nucleosome positioning strengthens the view that BEAF may regulate chromatin accessibility for gene expression through a restriction of the deposition of methylated H3K9 marks into dual-cores.

Our model whereby dual-cores regulate the deposition of specific epigenetic marks is in agreement with the activity of other known insulators [6,7,9–11]. Variations in H3K9me₃ levels might affect the interplay between the deposition of

H3K9me₃ and acetylated histone H4 (H4Ac) marks [45]. However, no variation in the deposition of H4Ac could be found in dual-cores compared to control regions after BEAF depletion (unpublished data). This is not surprising, as BEAF has no de-silencing activity on its own [5,25]. Computer analysis failed to reveal any enrichment of dual-cores near the 3'UTR of genes, and the activity of dual-cores may thus essentially play a role in regulating chromatin accessibility near promoter regions, but not within the 3' border of genes. Furthermore, the insulating activity of BEAF was demonstrated in the context of two dual-cores bracketing a transgene [5,25], and most likely also involved higher-level chromatin organization [2]. Although not enriched near the 3'UTR of genes, dual-cores still bracket/separate groups of genes clustered within 5–15 Kbp, a genomic context that may further require insulating activity to block promiscuous enhancer–promoter interactions and involve DNA looping between distant insulators [2]. It has recently been shown for a Su(Hw) insulator that the regulation of gene expression may further depend on its genomic environment [46]. Also, other dual-cores are often found in the vicinity of genes exposed to repression by heterochromatin (see our genome-wide database), and the function of BEAF may be particularly important in this context [17,20,23,24]. We propose that the BEAF dual-cores closely linked to a restricted array of several hundred genes define a family of insulators that provide a link between chromatin organization and the cell cycle.

Materials and Methods

Bioinformatic analysis, availability of predictions, dual-core sequences. All genome-wide predictions and analyses are available on our Web site: <http://www.sfu.ca/~eemberly/insulator/>. Additional information, including DNA sequences of single elements, dual-cores or dual-core-like elements, and their position relative to genes or other genomic features (GOs) can be directly retrieved from our Web site.

Dual-core-like prediction, distribution of CGATA sites in dual-cores. Each single BEAF element that was not a part of a dual-core element was analyzed for the presence of a “dual-core-like” signature. We define single elements as consisting of three CGATAs within 200 bp, and a dual-core-like element as a single BEAF element (three CGATAs) associated with a second nearby (<800 bp) cluster of two CGATA sites within a 100-bp window. 1,226 BEAF elements fit into this classification, including all previously identified dual-cores (BE76, BE28, BE51, Jan/Ser(BE83)). The position of each CGATA site within a dual-core sequence was analyzed relative to the position of the rightmost site of the first BEAF single element. In Figure S1, the position of each CGATA motif was measured from the average position (taken as position 0 on the x-axis) of all the CGATA locations in the first BEAF single element of the dual-core. This removes any ambiguity in defining the starting position of the sequence, allowing more precise mapping of dual-cores with respect to gene promoters.

Statistical significance of dual-cores. We predicted dual-cores by pairing together the genome-wide set of 7,045 single BEAF elements that were separated by a spacer <L bp. The statistical significance of the number of predicted dual-cores as a function of spacer length L was assessed by comparing it to the expected number for randomly spaced elements. The *p*-value was found to reach a flat minima for 600 bp < L < 3,000 bp. For larger L values, the predictions decreased in significance, eventually becoming no more significant than chance. There are 1,720 dual-cores, L = 800 bp with a *p*-value of 1e-9, in the sequenced *Drosophila melanogaster* genome.

Statistical significance of promoter distances to dual-cores. The statistical significance of the number of dual-cores within ±d bp of a promoter was assessed by comparing it to the number expected for randomly placed elements. Out of 1,720 dual-core elements, 545 fall within ±500 bp of a promoter. Beyond this distance, the *p*-value was found to decrease in statistical significance, yet 850 dual-cores reside within 2,000 bp of a promoter. Additional dual-cores are found close to genes or groups of genes (see our database).

Statistical significance of the distribution of BEAF dual-cores. In order to analyze the distribution of dual-cores, we calculated the statistical significance for a minimum number of dual-cores, 2, 3, ... x dual-cores (DC) to be found along 5, 10, ... 100 kbp of DNA (W). For a given W and DC, we predicted $N(W, DC)$, the frequency of dual-cores for a certain DNA length. To assess the significance of $N(W, DC)$, we compared it to the number of randomly distributed elements for the same DNA length. If the probability of a random dual-core element to occur within a window of size W is p, then the probability that there are $\geq DC$ elements in W is $P(W, DC) = B(x > DC, W, p)$, where B is the binomial distribution. The expected number of domain predictions for these random elements is then $E(W, DC) = N_{win}(W) * P(W, DC)$, where $N_{win}(W)$ is the number of non-overlapping windows of size W in the entire genome. The p-value for $N(W, DC)$ can then be evaluated using the expected number $E(W, DC)$ as a function of W and DC. We find $W = 10$ kb and $DC = 2$ to yield the statistically most significant BEAF dual-core distribution in pairs (p-value $\sim 1.01e-33$).

GO analysis. The statistical significance of a GO class was assessed using the binomial distribution, p -value = $B(x, N, p)$, where x is the number of genes within the given GO class in a set of N predicted genes, and p is the probability of that GO class in the entire annotation. See our database for a complete listing of all GO analyses of positively correlated genes with or without BEAF dual-cores or DREF elements in their promoters.

Genomic expression analysis and microarray data Genome-wide *Drosophila* gene expression data (Figure S7) covering the first 12 hours of embryonic development are available from the Berkeley *Drosophila* Genome Project. Twelve time points were collected, each with three replicates. Each gene g in the genome has an expression profile containing 12 data points ($g_i = (x_1, x_2, \dots, x_{12})$). For a given pair of genes, we calculated the Pearson correlation coefficient between their respective expression profiles. We then calculated the correlation coefficient between a given set of genes and a given reference gene. To test whether two sets of genes had statistically different correlation coefficient profiles, we used the Kolmogorov-Smirnov test, which assigns a p-value to the likelihood that two samples of a continuous random variable come from the same parent distribution.

Chromatin immunoprecipitation of BEAF, H3K9me3. Chromatin immunoprecipitation (ChIP) was done according to the Upstate protocol using control or *beaf* siRNA-treated cells. Equivalent amounts of chromatin samples were sonicated using a Diagenode Bioruptor and immunoprecipitated with 4 μ l of anti-H3K9me3 (Abcam). Precipitated DNA was analyzed by real-time PCR in parallel with genomic DNA using a Roche Light Cycler and a Light Cycler FastStart DNA Master SYBR green kit. The amplified DNA fragments (<250 bp) cover regions corresponding to the indicated elements (Figures 6 and 7). ChIP with anti-BEAF was performed as previously described [34] with 10 μ l affinity-purified anti-BEAF antibodies that recognize (Figure S5) or not (Figure 2) the BEAF-32A isoform or IgG. The immunoprecipitated DNA was analyzed in parallel with input genomic DNA as a standard. For ChIP-on-chip assays using H3K9 antibodies, precipitated DNA was amplified by ligation-mediated PCR (LM-PCR). 4 μ g of each amplified sample was used to hybridize on 3×385 K tiling microarrays representing the euchromatic, non-repetitive regions of the *Drosophila melanogaster* genome sequence (Flybase release 4.3) from Nimblegen Systems (GEO accessions: GPL3352, GPL3353, GPL3354). To calculate whether the levels of enrichment are statistically significant for each array, a normal distribution was calculated, with the assumption that the mode and median absolute deviation of the normalized \log_2 ratios are the average and the standard deviation of the normal distribution, respectively. Assuming that the normal distribution covers the entire background noise (non-significant signals), a p-value was calculated for each oligonucleotide signal. For the two replicate samples of each profile, each pair of probe p-values were then combined using a Chi Square law with 4 degrees of freedom. Finally, correction for multiple testing [47] was applied to the combined p-values. Only oligonucleotides with final p-values (for combined replicates) $< 1E-04$ were considered to be significantly enriched for the signal.

Expression analyses, siRNA treatments, transfections, expression of *beaf* mutants in embryos. For siRNA treatments, exponentially growing *Drosophila* Schneider SL2 cells were maintained between 1 and 4×10^6 cells/ml in Schneider's *Drosophila* medium (SDM, GIBCO, Invitrogen) supplemented with 10% Fetal Bovine Serum (FBS, Sigma) and 1% penicillin/streptomycin (GIBCO, Invitrogen). Cells were diluted to a final concentration of 1×10^6 cells/ml in SDM without FBS, and 400 μ l of 2 μ M *beaf32*, *dref* or *cdk7* double-stranded RNAs (dsRNA) were added directly to 10 ml of cells which were then plated on 75-cm² T-flasks (Sarstedt), immediately followed by vigorous

agitation. dsRNAs were synthesized using full-length cDNAs of the above genes as templates. Primers consisted of a complementary template portion, a floating end with a T7 promoter and an EcoRI site located at the other end. 5 μ g of DNA template were transcribed for 2 hours at 37 °C in the presence of 0.5 mM rNTPs, 10 mM DTT, 120 units RNase inhibitor, 60 units T7 polymerase in its 1 \times buffer in a 100 μ l final volume. cDNA degradation was performed for 30 to 40 minutes at 37 °C in the presence of 4 units RQDNase in a 400 μ l final volume of the recommended buffer. dsDNAs were then extracted with phenol/chloroform, ethanol-precipitated, and solubilized in 20 μ l TE, pH 7.5. The resulting sequences were checked for potential off-target effects by performing searches with dsCheck [48] (<http://dsCheck.RNAi.jp/>). Treated cells were incubated for 2 hours at 25°C, followed by addition of 20 ml of SDM containing FBS, and cells were incubated for an additional 5 days. Depletion of *beaf32* mRNA was assayed by RT-PCR at 1, 3, or 5 days after treatment. Cells were grown for 5–6 days, and samples were recovered for total RNA, immunostaining, or immunoblotting analysis. FACS analyses were performed after resuspending control or BEAF-depleted cells and staining their DNA with propidium iodide. Analysis of gene expression was performed by quantitative RT-PCR on cDNAs prepared by RT-PCR from BEAF-depleted or control cells (+5–6 days), untreated or treated with AA (5 μ M) for 24 hours. Each measurement was reproduced three times and in two independent RNA extraction experiments. For gene expression analysis, cDNAs prepared from control or BEAF-depleted cells were quantified in parallel with genomic DNA by RT-PCR using a Qiagen Light Cycler. Transfections of plasmids were performed using Lipofectamine (Invitrogen) for 2 hours according to the manufacturer's instructions, 48 hours before RNA purification. Measurements of gene expression for the transfected (wild-type or mutant) constructs were performed using primers that specifically amplify cDNAs from the tags introduced at the 5' and 3' borders (see Figure 5) and that were unable to amplify cDNAs from untransfected cells (unpublished data). Expression was normalized to the copy number of transfected constructs estimated by quantitative PCR of input genomic DNA. For endogenous genes, the primer sequences were selected from the coding regions ($\approx 1,000$ bp 3' from promoter start) of each gene. For endogenous *cdk7/snf*, the selected primers lie outside (15 bp 5' or 3') of the tags. For other analyses, two primer pairs were used alone or in combination to confirm the specific increase/decrease in gene expression, using actin as a control. For quantitative RT-PCR analysis in embryos, males with the BID transgene on Chromosome 2 (CyO/Sp1; BID2B) were crossed with virgin females harboring an embryonic da-GAL4 driver (*daughterless*) on Chromosome 3. The corresponding measurements were compared to those from embryos expressing da-GAL4 alone or from BID2B embryos without a da-GAL4 driver.

Mutagenesis of dual-cores. For mutagenesis of Dcore38_D, a genomic DNA fragment harboring the first exons of *cdk7* and *snf* was cloned, and PCR-mediated mutagenesis was performed using primers that contain mismatches as followed: the *dre* (DREF site) mutant sequence is TAGCGATA and disrupts DREF binding but preserves the CGATA consensus of BEAF. The BEAF site mutant was produced by mutagenesis of two of the CGATA consensus in one cluster of the dual-core, using the *t*ATA mismatches critical for BEAF binding [17,23–25].

Polytene chromosomes, immunostaining analyses, Western blotting, and mapping of nuclease-sensitive sites. Immunostaining analyses were performed using affinity-purified mouse or rabbit anti-BEAF-32B (1:100) as previously described [34,49], using the indicated affinity-purified antibodies or commercially available anti-acetyl-Histone H4, anti-H3K9me3, anti-H3, anti-RNA polymerase II (Upstate), or anti-actin antibodies (Sigma). Double immunostaining of siRNA-treated cells was performed in duplicates and in parallel for control or BEAF-depleted cells treated for 1, 3, or 5 days. Each experiment was repeated three times. DNA was stained with 500 ng/ml DAPI or 1 μ g/ml Hoechst, and coverslips were mounted with 4 μ l of antifading mix and sealed with nail polish. Slides with siRNA control or BEAF-depleted cells were analyzed using the same acquisition parameters using a Leica DMRA2 microscope. Mapping of BEAF dual-cores and immunolocalization of anti-BEAF signals was performed over >10 Mbp for Chromosome 2 and X chromosome, showing striking correspondence (analysis available upon request). For mapping of nucleases-sensitive sites (Figure S8), freshly isolated nuclei from approximately 10^8 cells were digested with very low concentrations of either micrococcal nucleases or DNAase I essentially as previously described [17,20,23,24], and the purified DNA was further digested with *PvuII* and run onto a 1.2% agarose gel for Southern blotting. Naked DNA controls were similarly digested. A *PvuII-EcoRI* end-labeled DNA fragment was used to probe specifically

the region containing the dual-core region. Western blotting was performed using anti-actin or anti-BEAF antibodies. As a control, genomic DNA was first purified and then digested with MNase and Pvu II (+/- EcoRI to mark the 5' border of the dual-core) before analysis by Southern blotting. Western blotting was performed as previously described [17,24] using anti-actin, anti-H3K9me3, anti-mei-S332, or anti-BEAF antibodies.

Supporting Information

Figure S1. Statistical Analysis of Dual-Cores

(A,B) Plots showing the distribution of all 12,058 CGATA motifs from dual-cores (A) and the locations of their AT-rich spacers (B) as in Figure 1C and 1D, except that positions were calculated according to average positions of the three CGATAs in the first (left) cluster to define position zero.

(C) CGATA motifs in the second cluster are enriched near the border of the spacer (+200–300 bp), while fewer localize at larger distances.

Found at doi:10.1371/journal.pbio.0060327.sg001 (89 KB PDF).

Figure S2. Depletion of BEAF Impairs Protein Levels of Key Cell-Cycle Factors

Immunoblotting experiment showing the protein levels of BEAF, MEI-S332, and CDK7 compared to loading controls (ACTIN, DSP1), after siRNA-mediated depletion of BEAF or control treatment. 1.0, 3.0: standard, or 3-fold excess protein loading, respectively.

Found at doi:10.1371/journal.pbio.0060327.sg002 (43 KB PDF).

Figure S3. BEAF Controls the Levels of H3K9me3 Marks (A,B) Immunostaining analysis using (A) anti-histone H3 (green) and anti-actin (red) antibodies or (B) anti-H3K9me3 antibodies, in SL2 control ("control") or BEAF-depleted ("beaf") cells. Enlargements of confocal images after staining with anti-H3K9me3 antibodies are also shown (3X). DNA was counterstained with Hoechst. Bar, 10µm.

(C) Profile of H3K9me3 and position of BEAF Dcores on the X chromosome corresponding to the Xdcore_38D region (first dual-core from right) or to the *eye* locus from ChIP-on-chip data. Note that promoter regions often fit into discrete H3K9me3 peaks distinct from the major H3K9me3 peaks of repressed loci (e.g., *eye*) that are also enriched for the H3 methylK27 mark (see text).

Found at doi:10.1371/journal.pbio.0060327.sg003 (156 KB PDF).

Figure S4. BEAF Elements Resembling Dual-Cores Are also Bound by BEAF In Vivo

The figure shows one of the exceptions for a region where some BEAF binding is detected (graph in green) by genome-wide ChIP-on-chip analysis (approximately 1,800 peaks total) yet which is not included in our database of dual-cores (1,720 dual-cores). This region was not scored in the dual-core database because the second CGATA in the first cluster is 103 bp away ("out") instead of the defined window of 100 bp. TSSs and primary transcript are depicted on the top graphs (see purple bars and blue line, respectively).

Found at doi:10.1371/journal.pbio.0060327.sg004 (474 KB TIF).

Figure S5. The BEAF-32A Splicing Variant also Binds to Dual-Cores

The panel shows an alignment of ChIP-on-chip analysis (graphs in green) using anti-BEAF antibodies that recognize the BEAF-32A splicing variant (+32A) or not (-32A). The red bars mark the position of significant peaks over the same region of the X chromosome (nucleotide positions 4,950,000 to 5,300,000) as shown in Figure 2. TSS (blue bars) and primary transcripts (purple lines) are shown on top.

Found at doi:10.1371/journal.pbio.0060327.sg005 (1.4 MB TIF).

Figure S6. Respective Enrichment of BEAF Dual-Cores and DREF Elements for Several Gene-Class Ontologies

p-Values for gene annotations (GOs) of BEAF dual-cores-only ("dual-cores-only") versus dual-cores containing additional TATCGATA consensus sites for DREF ("dual-cores-DREF") [50]. The ratio of *p*-values is shown for each independent GO category and highlights a greater enrichment for BEAF dual-cores-only sites in chromosome organization (left) and for dual-cores-DREF sites in cell-cycle and apoptosis (right). DREF competes with BEAF for binding to a nested consensus sequence [34] present in dual-cores marked by a "_D" sign (see our Web site). These are significantly enriched in common GOs, including cell-cycle, in agreement with genetic interactions between *beaf* and *dref* [33,50] (see text for details).

Found at doi:10.1371/journal.pbio.0060327.sg006 (124 KB PDF).

Figure S7. Genome-Wide Analysis of the Impact of BEAF Dual Cores on Transcription

(A) BEAF dual-cores have a global positive impact on transcription. Distribution of correlation coefficients between the expression profile of genes with (red) or without (black) BEAF dual-cores (i) in their promoters (see Materials and Methods). "+" and "-" signs indicate statistical enrichment for co-regulated and anti-correlated gene expression profiles, respectively. As a positive control, the target genes for DREF [50] are enriched, as expected, for a minor subpopulation highly co-expressed with DREF (ii), but less significantly (*p*-value of 0.004 according to the Kolmogorov-Smirnov test) than BEAF, which has a more global positive effect on gene expression (*p*-value ~ 3e-17 according to the Kolmogorov-Smirnov test).

(B) Distribution of the BEAF (CGATA, green boxes) and DREF (TATCGATA, red) motifs in the Dual-core 38_D with respect to *snf* and *cdk7* (TSS corresponds to the first colored bp).

Found at doi:10.1371/journal.pbio.0060327.sg007 (309 KB PDF).

Figure S8. The Spacers of BEAF Dual-Cores Fit into Nucleosome-Positioning Sequences

(A,B) Relative positioning of CGATA BEAF consensus binding motifs and the position of putative NPSs predicted by submitting dual-core sequences to available databases [42,43] in the *cdk7* and *mei-S332* promoter regions (A) as well as in >20 cell-cycle regulatory genes (B) (see our Web site for a list). Predicted NPSs are indicated by purple boxes below dual-cores (A) or as an overlay of predicted NPSs (B). The relative position of nucleosome-resistant cores is indicated (N; according to experiments as shown in (C)). These predictions fit with the positions of AT-rich dual-core spacers (see Figure 1D).

(C) Mapping of the accessibility of naked DNA control (top photograph) and of chromatin by nuclease digestion of nuclei (MNase, "M"; or DNAase I, "D"; see Materials and Methods). To map nuclease-resistant/sensitive regions with respect to CGATA clusters of dual-cores, purified genomic DNA was further digested with a second enzyme (*PvuII* + *NotI* or *EcoRI*) which cuts into the first CGATA cluster or 50 bp 3' of the second CGATA cluster, respectively (see dotted lines below the autoradiogram). The dual-core spacer fits into a nuclease-resistant core region bracketed by hypersensitive sites. Note that these features are not found in the naked DNA control, where genomic DNA was first purified before MNase digestion.

Found at doi:10.1371/journal.pbio.0060327.sg008 (239 KB PDF).

Table S1. Gene-Class Ontologies Associated with BEAF Dual-Cores

GO terms for 1,720 BEAF dual-core target genes, which contain a dual-core within +/- 1,000 bp of their promoter. 671 dual-core elements hit one promoter in the genome. The second column gives the number of annotated genes in that GO class, the third column gives the number of genes in dual-core/promoter sets in that GO class, the fourth column shows the expected number of genes in the predicted set given the observed class frequency. The corresponding *p*-value is given in the fifth column. GO terms have been binned into larger categories. Low-scoring GO classes underrepresented in the set of dual-core target genes are shown at the bottom. See our database for a complete listing and additional GO analysis.

Found at doi:10.1371/journal.pbio.0060327.st001 (25 KB DOC).

Acknowledgments

We thank Uli Laemmli and Terry Orr-Weaver for anti-BEAF and anti-MeiS332 antibodies, G. Cavalli for sharing unpublished ChIP-on-chip data for H3K9me3, C. Bez for initial siRNA experiments, F. Maschat for da-GAL4 lines, B. Leblanc for his help with analysis of histone marks, and K. Ishii and M. Méchali for their support and comments.

Author contributions. OC conceived and designed the experiments. EE, RB, BS, MH, NJ, and OC performed the experiments. EE, RB, BS, MH, NJ, CMH, EK, and OC analyzed the data. EE, BS, MH, NJ, CMH, EK, and OC contributed reagents/materials/analysis tools. OC wrote the paper.

Funding. EE would like to acknowledge the support of NSERC and CIAR. Work in EK's lab was supported by ARC, ACI Cancéropôle GSO, and ANR grants. RB was supported by the French Research Ministry and ARC fellowships. BS was supported by a fellowship of

the Fondation de la Recherche Médicale (FRM) and by the Austrian FWF (Fonds zur Förderung der wissenschaftlichen Forschung). CMH was supported by grants from Louisiana Board of Regents LA BOR, LSU Faculty Research Grant LSU FRG, and the NSF. OC would like to

acknowledge the support of the Human Frontier Science Program, CNRS, and INSERM.

Competing interests. The authors have declared that no competing interests exist.

References

- Capelson M, Corces VG (2004) Boundary elements and nuclear organization. *Biol Cell* 96: 617–629.
- Maeda RK, Karch F (2007) Making connections: Boundaries and insulators in *Drosophila*. *Curr Opin Genet Dev* 17: 394–399.
- Gerasimova TI, Corces VG (1999) Domains and boundaries in chromosomes. *Results Probl Cell Differ* 25: 229–249.
- Ishii K, Arib G, Lin C, Van Houwe G, Laemmli UK (2002) Chromatin boundaries in budding yeast: The nuclear pore connection. *Cell* 109: 551–562.
- Ishii K, Laemmli UK (2003) Structural and dynamic functions establish chromatin domains. *Mol Cell* 11: 237–248.
- West AG, Gaszner M, Felsenfeld G (2002) Insulators: Many functions, many mechanisms. *Genes Dev* 16: 271–288.
- Gaszner M, Felsenfeld G (2006) Insulators: Exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet* 7: 703–713.
- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, et al. (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129: 823–837.
- Jenuwein T, Allis CD (2001) Translating the histone code. *Science* 293: 1074–1080.
- Bannister AJ, Schneider R, Kouzarides T (2002) Histone methylation: Dynamic or static? *Cell* 109: 801–806.
- Grewal SI, Moazed D (2003) Heterochromatin and epigenetic control of gene expression. *Science* 301: 798–802.
- Albert I, Mavrich TN, Tomsho LP, Qi J, Zanton SJ, et al. (2007) Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature* 446: 572–576.
- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, et al. (2007) Response: Mapping nucleosome positions using ChIP-Seq data. *Cell* 131: 832–833.
- Bell AC, West AG, Felsenfeld G (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98: 387–396.
- Moon H, Filippova G, Loukinov D, Pugacheva E, Chen Q, et al. (2005) CTCF is conserved from *Drosophila* to humans and confers enhancer blocking of the Fab-8 insulator. *EMBO Rep* 6: 165–170.
- Blanton J, Gaszner M, Schedl P (2003) Protein:protein interactions and the pairing of boundary elements in vivo. *Genes Dev* 17: 664–675.
- Zhao K, Hart CM, Laemmli UK (1995) Visualization of chromosomal domains with boundary element-associated factor BEAF-32. *Cell* 81: 879–889.
- Corces VG (1995) Chromatin insulators. Keeping enhancers under control. *Nature* 376: 462–463.
- Gerasimova TI, Corces VG (1996) Boundary and insulator elements in chromosomes. *Curr Opin Genet Dev* 6: 185–192.
- Cuvier O, Hart CM, Laemmli UK (1998) Identification of a class of chromatin boundary elements. *Mol Cell Biol* 18: 7478–7486.
- Ramos E, Ghosh D, Baxter E, Corces VG (2006) Genomic organization of gypsy chromatin insulators in *Drosophila melanogaster*. *Genetics* 172: 2337–2349.
- Parnell TJ, Kuhn EJ, Gilmore BL, Helou C, Wold MS, et al. (2006) Identification of genomic sites that bind the *Drosophila* suppressor of Hairy-wing insulator protein. *Mol Cell Biol* 26: 5983–5993.
- Hart CM, Zhao K, Laemmli UK (1997) The scs' boundary element: Characterization of boundary element-associated factors. *Mol Cell Biol* 17: 999–1009.
- Cuvier O, Hart CM, Kas E, Laemmli UK (2002) Identification of a multicopy chromatin boundary element at the borders of silenced chromosomal domains. *Chromosoma* 110: 519–531.
- Gilbert MK, Tan YY, Hart CM (2006) The *Drosophila* boundary element-associated factors BEAF-32A and 32B affect chromatin structure. *Genetics* 173: 1365–1365.
- Udvardy A, Maine E, 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. *J Mol Biol* 185: 341–358.
- Adryan B, Woerfel G, Birch-Machin I, Gao S, Quick M, et al. (2007) Genomic mapping of suppressor of Hairy-wing binding sites in *Drosophila*. *Genome Biol* 8: R167.
- Roy S, Tan YY, Hart CM (2007) A genetic screen supports a broad role for the *Drosophila* insulator proteins BEAF-32A and BEAF-32B in maintaining patterns of gene expression. *Mol Genet Genomics* 277: 273–286.
- Roy S, Gilbert MK, Hart CM (2007) Characterization of BEAF mutations isolated by homologous recombination in *Drosophila*. *Genetics* 176: 801–813.
- Kerrebrock AW, Moore DP, Wu JS, Orr-Weaver TL (1995) Mei-S332, a *Drosophila* protein required for sister-chromatid cohesion, can localize to meiotic centromere regions. *Cell* 83: 247–256.
- Larochelle S, Pandur J, Fisher RP, Salz HK, Suter B (1998) Cdk7 is essential for mitosis and for in vivo Cdk-activating kinase activity. *Genes Dev* 12: 370–381.
- Chen J, Larochelle S, Li X, Suter B (2003) Xpd/Erec2 regulates CAK activity and mitotic progression. *Nature* 424: 228–232.
- Yoshida H, Inoue YH, Hirose F, Sakaguchi K, Matsukage A, et al. (2001) Over-expression of DREF in the *Drosophila* wing imaginal disc induces apoptosis and a notching wing phenotype. *Genes Cells* 6: 877–886.
- Hart CM, Cuvier O, Laemmli UK (1999) Evidence for an antagonistic relationship between the boundary element-associated factor BEAF and the transcription factor DREF. *Chromosoma* 108: 375–383.
- Garrett FE, Emelyanov AV, Sepulveda MA, Flanagan P, Volpi S, et al. (2005) Chromatin architecture near a potential 3' end of the igh locus involves modular regulation of histone modifications during B-Cell development and in vivo occupancy at CTCF sites. *Mol Cell Biol* 25: 1511–1525.
- Splinter E, Heath H, Kooren J, Palstra RJ, Klous P, et al. (2006) CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev* 20: 2349–2354.
- Blattes R, Monod C, Susbielle G, Cuvier O, Wu JH, et al. (2006) Displacement of D1, HP1 and topoisomerase II from satellite heterochromatin by a specific polyamide. *EMBO J* 25: 2397–2408.
- Huisinga KL, Brower-Toland B, Elgin SC (2006) The contradictory definitions of heterochromatin: Transcription and silencing. *Chromosoma* 115: 110–122.
- Hediger F, Gasser SM (2006) Heterochromatin protein 1: Don't judge the book by its cover! *Curr Opin Genet Dev* 16: 143–150.
- Mantelingu K, Reddy BA, Swaminathan V, Kishore AH, Siddappa NB, et al. (2007) Specific inhibition of p300-HAT alters global gene expression and represses HIV replication. *Chem Biol* 14: 645–657.
- Fu Y, Sinha M, Peterson CL, Weng Z (2008) The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet* 4: e1000138. doi/10.1371/journal.pgen.1000138
- Ioshikhes IP, Albert I, Zanton SJ, Pugh BF (2006) Nucleosome positions predicted through comparative genomics. *Nat Genet* 38: 1210–1215.
- Segal E, Fondufe-Mittendorf Y, Chen L, Thastrom A, Field Y, et al. (2006) A genomic code for nucleosome positioning. *Nature* 442: 772–778.
- Mavrich TN, Jiang C, Ioshikhes IP, Li X, Venters BJ, et al. (2008) Nucleosome organization in the *Drosophila* genome. *Nature* 453: 358–362.
- Taddei A, Maison C, Roche D, Almouzni G (2001) Reversible disruption of pericentric heterochromatin and centromere function by inhibiting deacetylases. *Nat Cell Biol* 3: 114–120.
- Soshnev AA, Li X, Wehling MD, Geyer PK (2008) Context differences reveal insulator and activator functions of a Su(Hw) binding region. *PLoS Genet* 4: e1000159. doi/10.1371/journal.pgen.1000159
- Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I (2001) Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* 125: 279–284.
- Naito Y, Yamada T, Matsumiya T, Ui-Tei K, Saigo K, et al. (2005) dsCheck: Highly sensitive off-target search software for double-stranded RNA-mediated RNA interference. *Nucleic Acids Res* 33: W589–W591.
- Cuvier O, Lutzmann M, Mechali M (2006) ORC is necessary at the interphase to mitosis transition to recruit cdc2 kinase and disassemble RPA foci. *Curr Biol* 16: 516–523.
- Hirose F, Ohshima N, Shiraki M, Inoue YH, Taguchi O, et al. (2001) Ectopic expression of DREF induces DNA synthesis, apoptosis, and unusual morphogenesis in the *Drosophila* eye imaginal disc: Possible interaction with Polycomb and trithorax group proteins. *Mol Cell Biol* 21: 7231–7242.

2.3.3 Rappel des principaux résultats

Par des analyses statistiques des motifs CGATA (reconnus par BEAF) présents dans le génome de la drosophile, un motif de liaison de BEAF *in vivo* a pu être proposé. Ce motif appelé « *dual-core* » est composé de deux groupes de 2 ou 3 motifs CGATA séparés par une séquence d'environ 200 pb en moyenne, riche en nucléotides A et T. Des analyses d'immunoprécipitation de la chromatine ont permis de confirmer sur un certain nombre d'exemples que BEAF se lie effectivement à ce motif *in vivo*.

L'analyse de la position des *dual-cores* dans le génome a montré que plus de 500 d'entre eux sont associés à la région promotrice de gènes. Le classement de ces gènes par fonction montre un enrichissement de gènes impliqués dans la régulation du cycle cellulaire, dans la ségrégation des chromosomes ou dans l'apoptose. La déplétion par ARN interférence de BEAF conduit à un ralentissement de la phase G2/M corrélé à un défaut de croissance, ce qui fait partie des phénotypes attendus après dérégulation de gènes importants pour la progression dans le cycle cellulaire.

Ces phénotypes pourraient être des conséquences de la dérégulation de gènes activés par DREF dont la séquence consensus de liaison contient le motif de liaison de BEAF et qui pourrait entrer en compétition avec BEAF pour la liaison sur ces séquences. Des expériences de RT-qPCR ont permis de confirmer que BEAF régule effectivement les gènes associés à un *dual-core*. Il est alors apparu deux classes de gènes régulés par BEAF. La première contient des gènes sans motifs de liaison de DREF et sont régulés positivement par BEAF. La seconde contient quant à elle des gènes dont les promoteurs arborent un motif de liaison de DREF. Cette seconde classe est régulée négativement par BEAF et nous avons montré que c'est en entrant en compétition avec l'activateur DREF que BEAF contrôle l'expression de ces gènes.

Nous avons ensuite montré par ChIP que l'activation des gènes par BEAF se faisait en empêchant la déposition de la marque répressive H3K9me3 dans les *dual-cores*. De façon intéressante, ce mécanisme n'est pas impliqué dans l'activation des gènes par DREF.

Enfin, l'utilisation de mouches transgéniques exprimant BID qui est un dominant négatif capable de séquestrer la protéine BEAF endogène, a permis de confirmer les résultats obtenus dans des cellules en culture.

2.3.4 Retour sur le motif *dual-core*

Certains sites où BEAF est détecté ne présentent pas le motif proposé dans ce travail. En effet, certains sites n'ont pas une structure en *dual-core*, ni même de motifs CGATA, alors que d'autres présentent un *dual-core* avec une séquence entre les deux groupes de CGATA nettement plus courte. Pour expliquer la première observation, on peut supposer qu'on détecte une interaction indirecte de BEAF via une interaction avec d'autres protéines insultrices. C'est notamment le cas de l'insulateur *scs* lié par Zw5 qui interagit avec BEAF pour regrouper les éléments *scs* et *scs'* (Blanton *et al.* 2003, voir§1.4.2.3). Ainsi, malgré l'absence de motif CGATA, on détecte sur *scs* un petit signal de BEAF qui est dû à l'interaction avec Zw5. Le second cas montre que la séquence de l'espaceur entre les motifs CGATA n'est pas importante pour la liaison de BEAF. Des études effectuées depuis 2008 ont montré que les *dual-cores* ne sont pas les seuls motifs de liaison de BEAF. L'analyse de l'ensemble de ces sites par immunoprécipitation de chromatine d'embryons puis hybridation sur puce montre que de nombreux sites ne présentent pas cette structure, mais également que de nombreux *dual-cores* ne sont pas liés par BEAF (Jiang *et al.* 2009). En effet, d'après cette étude, sur les 1720 *dual-cores* identifiés par notre approche informatique, seulement 25% présentent un pic BEAF. Par contre, si on se limite aux *dual-cores* proches du site de démarrage de transcription (TSS) d'un gène (<600pb, 581 *dual-cores*), alors près de 70% sont effectivement liés par BEAF. Ainsi, d'autres facteurs sont probablement important pour stabiliser la liaison de BEAF au niveau des promoteurs

et non sur les autres *dual-cores*. En effet, les sites de liaison effectifs de BEAF se retrouvent très proches des TSS (85% des sites sont situés à moins de 300 pb d'un TSS), en particulier de gènes actifs. Cette observation suggère que le rôle principal de BEAF est la régulation transcriptionnelle. Cependant, l'étude présentée ci-dessous cartographie les sites de liaison de BEAF par ChIP-seq (~5400 sites) et montre que plus de 80% des *dual-cores* sont effectivement liés par BEAF, ce qui prouve que les résultats de l'étude de Jiang *et al.* (2009) sont la conséquence d'un manque de sensibilité (1820 sites) et non d'une mauvaise prédiction du motif de liaison de BEAF.

Pour expliquer le motif en *dual-core*, nous pensons à présent que c'est un motif particulier retrouvé lorsque BEAF se lie entre deux gènes divergents (Fig. 2.11), ce qui est très fréquent. Cette configuration est en effet retrouvée dans environ 50% des sites liés par BEAF (Jiang *et al.* 2009).

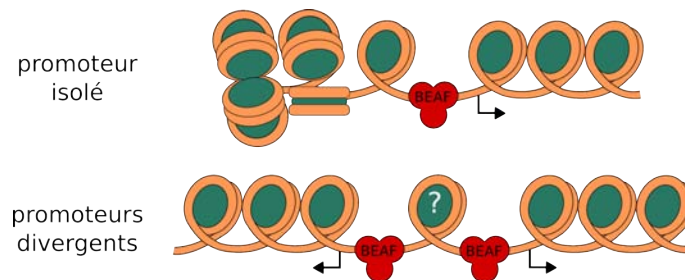


Figure 2.11 – Modèle pour expliquer le motif *dual-core*. En haut se trouve la représentation d'un gène isolé où il n'y a qu'un seul groupe de motifs CGATA qui est lié par un trimère de BEAF. En bas est représenté un locus avec deux gènes divergents. Il y a dans ce cas là deux groupes de motifs CGATA liés par BEAF, ce qui forme le motif *dual-core* proposé dans l'article.

Ainsi les gènes isolés ne présentent pas de motifs en *dual-core* mais seulement un seul groupe de CGATA qui suffit à la liaison de BEAF et à la régulation de l'expression du gène. Cependant, tous les groupes de motifs CGATA du génome ne sont pas liés par BEAF et il serait intéressant de comprendre ce qui restreint la liaison de BEAF, à la fois sur ces groupes isolés et sur les 20% de *dual-cores* non liés.

Le motif en *dual-core* s'est trouvé enrichi dans notre approche informatique à cause d'un biais dans l'organisation des gènes régulés par BEAF. La localisation préférentielle de BEAF entre les gènes divergents pourrait avoir une raison fonctionnelle qu'il serait intéressant de comprendre. Ce facteur pourrait être important dans la régulation de ce type de gènes, soit en permettant une expression indépendante de ces gènes, soit en facilitant leur corégulation.

2.3.5 Conclusion

Cette étude propose un motif particulier de liaison de BEAF et montre que BEAF régule les gènes associés en empêchant l'accumulation de la marque répressive H3K9me3. Parmi les gènes ainsi régulés se trouvent de nombreux régulateurs clés du cycle cellulaire. Certains sont associés avec l'activateur DREF. La compétition entre ces deux facteurs apparaît comme un mécanisme important de régulation du cycle. Ainsi BEAF empêche généralement l'activation par DREF et nous proposons qu'à un certain moment du cycle, un échange entre ces deux facteurs permettent l'induction par DREF et la progression dans le cycle cellulaire. Ce travail apporte une avancée importante dans la compréhension de la fonction de BEAF. En effet, cette étude est la première qui propose que BEAF participe directement à la régulation transcriptionnelle des gènes en se liant dans la région promotrice. Ces observations changent considérablement la vision que l'on avait de cette protéine. Les propriétés d'insulateurs des sites de liaison de BEAF suggéraient plutôt un rôle dans la compartimentation du génome et il apparaît à présent que sa fonction première

pourrait être la régulation directe de l'expression des gènes. L'étude qui suit dissèque les mécanismes de cette régulation.

2.4 Mécanismes de régulation transcriptionnelle

Après l'arrivée d'Olivier Cuvier au laboratoire, mon travail s'est orienté vers la compréhension des mécanismes de régulation de la transcription par les insulateurs BEAF. En effet, nous avons précédemment montré que la plupart des gènes associés à un *dual-core* sont régulés par BEAF (voir §2.3.2), mais les mécanismes moléculaires mis en jeu restaient encore mal connus. Nous avons montré que la marque H3K9me3 était impliquée dans cette régulation, ce qui montre que BEAF intervient dans la régulation de la structure chromatinienne. Pour comprendre précisément le rôle de BEAF dans l'organisation de la chromatine nous avons étudié, à l'échelle du génome entier, le positionnement des nucléosomes en présence ou en absence de BEAF et établi un lien avec la régulation de centaines de gènes.

2.4.1 Article soumis à *Genes and Development*

**Insulators Specify Discrete +1/+2 Nucleosome-Positions to Regulate
Gene Expression through Transcriptional Pausing**

Magali Hennion^{**}, Suresh Cuppadra^{^*}, Adrien Gamot^{**}, Dany Severac[~], Christelle
Dantec[~], Keji Zhao^{^@}, Eldon Emberly^{#@}, and Olivier Cuvier^{+@}

⁺ Laboratoire de Biologie Moléculaire Eucaryote (LBME); CNRS; Université de
Toulouse (UPS); F-31000 Toulouse; France

[^] Laboratory of Molecular Immunology (LMI), Division of Intramural Research (DIR),
National Heart, Lung, and Blood Institute (NHLBI), National Institutes of Health (NIH),
Bethesda, USA

[~]Institut de Génomique Fonctionnelle, MGX-Montpellier GenomiX & Mass-
spectrometry facility Montpellier, France

[#] Physics Department, Simon Fraser University (SFU), Burnaby, BC, Canada

* these authors contributed equally to the work

[@]co-corresponding authors:

E-mail: zhaok@nhlbi.nih.gov; eemberly@sfu.ca ; cuvier@ibcg.biotoul.fr

Abbreviations: NELF; Negative Elongation Factor; SOM, Self-organization map; Pol II, RNA polymerase II; BEAF, Boundary Element-Associated Factor; GAF, GAGA Factor; CTCF, CCCTC-binding protein;

Key words: Chromatin organization ; Nucleosome boundary ; Transcriptional stalling / pausing

Running title: Genome-wide analysis of Insulators

ABSTRACT

Insulators maintain proper expression of eukaryotic genomes involving both local chromatin organization and long-range interactions. It remains unclear how insulator proteins bound close to the transcription start sites exert these functions. Here we report that the binding of BEAF insulator protein specify the position of +1 and +2 nucleosomes that are coupled to NELF-regulated RNA polymerase II pausing/stalling. Insulator-depletion impairs these nucleosome-positioning patterns associated with Pol II-stalling resulting in decreased expression of hundreds of genes including oncogenes. Insulators may thus function through nucleosome-positioning to regulate transcriptional pausing that has important implications in understanding how they restrict promiscuous long-range interactions to participate in the functional organization of chromatin.

INTRODUCTION

Chromatin folding facilitates interactions among distant regulatory elements controlling gene expression (Lieberman-Aiden et al., 2009; Groudine et al., 2009), which may be restricted by chromatin insulators (Reviewed in Gaszner and Felsenfeld, 2006; Maeda and Karch, 2007; Bushey et al., 2009). Insulators block enhancer-promoter communication when interposed (Cai and Levine, 1995), as they can trap enhancers (Geyer, 1997) through mechanisms that remain not fully understood. Recent data show that transcription pausing recapitulates insulating activity (Chopra et al., 2009), as a paused polymerase, or open chromatin, may also trap upstream enhancer(s). Whether this model is applicable to the recently identified binding regions of CTCF and BEAF insulator proteins (Emberly et al., 2008; Bushey et al., 2009) that juxtapose to thousands of transcription start sites (TSS) where nucleosomes are lost (Bartkuhn et al., 2009; Nègre et al., 2010) remains to be determined.

By integrating three independent genome-wide approaches including ChIP-Seq, MNase-Seq and RNA-Seq, we mapped the nucleosome positions and expression profiles in insulator-depleted and control cells. Our results highlight that insulators play key functions in maintaining discrete +1 and +2 nucleosome positions tightly associated with transcriptional pausing. Insulator protein depletion specifically alters these positioned nucleosomes leading to the decreased expression of hundreds of genes. Our data reveal a key role of insulators at the pausing stage of transcription (Chopra et al., 2009) that has

important implications in understanding how they may participate in the functional organization of chromatin.

RESULTS AND DISCUSSION

The thousands of genomic BEAF32-insulator binding sites specify the +1 nucleosome profile associated with transcriptional pausing

CTCF, GAF and BEAF32 insulators juxtapose to hundreds of genes (Emberly et al., 2008; Jiang et al., 2009; Bushey et al., 2009; Jin et al., 2009; Nègre et al., 2010; reviewed in Raab and Kamakaka, 2010) and their binding has been associated with the loss of nucleosome occupancy resulting in a Nucleosome Free Region (NFR) in promoters (Bartkuhn et al., 2009). To test whether insulators might regulate nucleosome-positioning specifically, we mapped the genome-wide binding sites of BEAF32 using ChIP-seq and determined nucleosome positions using MNase-Seq in *Drosophila* S2 cells (Barski et al., 2007; Schones et al., 2008). Peaks of BEAF32 binding were mapped precisely to nucleosome free regions (Figure 1). Mapping the top and bottom strands separately on the genome provided high resolution BEAF sites (Supplementary Figure S1A), showing that they juxtaposed (<250bp) to the TSS of as many as 5,407 genes (Figure 1A) including 97.2% of the binding sites identified in embryos (Jiang et al., 2009) and ~3,000 new sites. A motif search in the BEAF32-bound regions revealed CGATA, the consensus motif essential for BEAF32 binding and activity (Cuvier et al., 1998; 2002), as the top motif. The other significant motifs include the E-box and DRE that are recognized by key cell cycle regulator (Supplementary Fig S1B; e-value~1e-60), in agreement with data showing that BEAF/CTCF insulators juxtapose cell-cycle genes (Emberly et al., 2008; Bushey et al., 2009).

High-resolution mapping of BEAF32 binding sites showed that they fell into the most prominent NFR genome-wide (Mavrich et al., 2008), where nucleosome-occupancy was particularly low (Figure 1B). BEAF32 binds to motifs whose distribution is specifically ‘anti-phased’ with respect to the position of the juxtaposed nucleosomes, a common feature shared with numerous sequence specific DNA-binding factors (Kaplan et al., 2009) that may provide a high-energy barrier which may organize the juxtaposed nucleosomes (Mavrich et al., 2008b; Ioshikhes et al., 2006; Milani et al., 2009; Zhang et al., 2009). Our MNase-Seq assays in wild-type cells accurately mapped the positions of nucleosomes (Figure 1C-D), showing that BEAF32 sites juxtaposed the most prominent +1 and +2 nucleosomes throughout the genome (Figure 1B) both in regions harboring two divergently transcribed genes or a single gene (‘lone’ genes; Figure 1D; Supplementary Figure S3). In the later case, the juxtaposed nucleosomes downstream of TSSs were highly positioned (Figure 1D), raising the possibility that the function of insulators may involve nucleosome-positioning at target genes.

Nucleosome-positioning has been associated with the stalling/pausing of the RNA polymerase II (Pol II), a stage that is specifically regulated by NELF (Muse et al., 2007; Zeitlinger et al., 2007; Lee et al., 2008; Mavrich et al., 2008; Core and Lis, 2008; Gilchrist et al., 2008; Adelman et al., 2009) and that controls the cell-cycle (Rahl et al., 2010). We sought to test whether it might be transcriptional stalling/pausing rather than insulator protein binding that drives the positioning of the juxtaposed nucleosomes. Comparison of the nucleosome profiles at gene TSSs grouped according to their association with and without BEAF32/NELF (Adelman et al., 2009) showed that the presence of BEAF32 binding correlated with higher signals of positioned nucleosomes

(Figure 1D) compared to the nucleosome profile of NELF-only bound genes (Figure 1D; data not shown). The nucleosome-positioning signal of all genes bound by Pol II was markedly reduced in the absence of BEAF32 (Supplementary Figure S2). BEAF32 may also contribute to the NELF-mediated pausing/stalling (Core and Lis, 2008) via specifying nucleosome-positioning in as much as 81% of NELF-bound promoters that are also associated with BEAF32 (Figure 1E).

To further test the potential role of BEAF32 in regulating nucleosome-positioning we determined the positions of nucleosomes in cells where *beaf* expression was efficiently knocked-down (~95%; Figure 2A and 2B; BEAF-KD), in parallel to mock-treated control cells (WT). Aligning the nucleosome profiles of genes bound by BEAF32 showed that a marked loss in the +1 positioned nucleosome occurred in BEAF-KD cells compared to control cells in various genomic contexts (Figure 2C-D; Supplementary Figure S3). These results support a role of BEAF32 in maintaining the positioning of the juxtaposed nucleosomes.

BEAF32 regulates expression of hundreds of genes through its role in nucleosome-positioning

We next asked whether the regulation of gene expression by insulators involves their potent function in nucleosome-positioning. Profiling of RNAs using RNA-Seq (see Material and Methods) revealed that the expression of 1,901 genes was significantly reduced in BEAF-KD compared to wild-type cells (FDR<0.01). Intersecting the RNA-seq data with BEAF32 binding data showed that a large majority (up to ~80%) of the

genes that were down-regulated genes upon BEAF32-KD harbored a juxtaposed (<250bp) BEAF32 binding site (p-value~1e-174; Figure 3A). The proportion of the down-regulated genes that were recognized by BEAF32 decreased proportionally (from ~80% down to ~20%; Figure 3B), suggesting that BEAF32 binding may directly contribute to the expression of thousands of genes. Gene ontology analysis indicated that the down-regulated genes were enriched in processes including cell cycle and cytoskeleton organization (Supplementary Table S1), in agreement with recent data showing that pausing regulates cell-cycle genes (Rahl et al., 2010). The 1,901 genes covered > 85% of the down-regulated genes identified through our microarray analysis (p-value ~1e-28; Supplementary Figure S4; Supplementary Table S2), confirming the reliability of our RNA-seq data.

Alignment of the nucleosome profiles of the 1,901 BEAF32-regulated genes highlighted a common pattern of positioned nucleosomes (Figure 3C, red curve), in contrast to control genes whose expression was not affected upon knockdown of BEAF32 (Figure 3C, black curve). This observation indicated that nucleosomes were not randomly positioned among the BEAF32-regulated genes, suggesting that BEAF32 might regulate their expression through a common mechanism involving nucleosome-positioning.

Clusters highlight distinct functional patterns of nucleosome positions at +1 and +2 specifically associated with insulators

The genomic context of BEAF32 sites might also impact how its depletion affects nucleosome-positioning. We thus clustered the nucleosome profiles using self-organization mapping (SOM) and k-means (Hartigan and Wong, 1979) to identify more

systematically all possible changes in nucleosome-positioning that might be specified by insulators. This allowed us to identify significant groups/clusters of genes sharing a particular pattern of positioned nucleosome (Figure 4 and 5). SOM analysis showed that BEAF32-associated promoters were highly enriched for several nucleosome patterns, mostly centered near, +1 and +2 positions (Figure 4A, right panel) that was not found for control promoters (Figure 4A, left panel). Similar results were obtained using the 5' or 3' reads, separately, of MNase-Seq data that delimit the boundaries of the nucleosomes (Figure 5A; data not shown) confirming the significance of the distinct +1 and +2 nucleosome patterns.

The nucleosomes centered near +1 contained two highly significant clusters corresponding to +1 and to +1-shifted ($\sim+25$ bp) nucleosomes (Figure 4B and C). The shifted +1 nucleosome pattern was the most significant pattern associated with hundreds of BEAF32-bound promoters (p -value $\sim 4e-29$; Supplementary Figure S7A). Importantly, both SOM and k-means showed that this cluster was specifically lost in BEAF-KD cells compared to WT cells (Figure 4B and C; data not shown), confirming that insulators regulate this discrete shift of nucleosome-positioning. Therefore, although the pattern of nucleosome-positioning varies according to genes and/or genomic contexts (see our genome-browser dedicated to nucleosome-positioning genome-wide), clustering approaches showed that highly significant positioned nucleosomes near TSSs of hundreds of promoters may be functionally specified by insulators.

Insulator-coupled +1/+2 nucleosome positioning regulates gene expression and transcriptional pausing

Further SOM analysis on the 1,901 genes that are differentially expressed (Figure 4D; WT>BEAF-KD) showed that the loss of +1-shifted and the +2 nucleosome were encountered in 304 (p-value~1e-18) and 388 (p-value~1e-24) down-regulated genes, respectively, corresponding to the two most significantly altered nucleosome patterns associated with changes in gene expression upon BEAF-KD (Figure 4D; Figure5). The +1 pattern was tightly associated with transcriptional pausing (Figure 1E; Adelman et al., 2009; Rahl et al., 2010), strengthening our data showing that insulators may regulate gene expression through this stage. These results also indicate that the +2 positioned nucleosome is not always preceded by a +1 positioned nucleosome, unlike shown in current models based on the mean values of nucleosome-positioning by aligning promoters (Figure 1B). The functional significance of a distinct +2 positioned nucleosome was confirmed by intersecting the different clusters of promoters sharing a similar nucleosome-positioning pattern with ChIP-seq data for BEAF32, Pol II or NELF binding(Supplementary Figure S7B; Adelman et al., 2009). These analyses showed that the +2 pattern was specifically associated with BEAF insulators and NELF-regulated pausing of transcription (p-value~1e-4), where higher levels of Pol II binding were detected over the pausing sites near TSS (Adelman et al., 2009; Rahl et al., 2010) as compared to the gene body. In contrast, this signature was not detected for inactive or active genes where its levels of binding are similar over the whole gene body (Adelman et al., 2009; Rahl et al., 2010) strongly suggesting that +2 positioned nucleosomes may serve to accommodate for the pausing of RNA polymerase II complexes. Further

examination of Pol II binding using ChIP in BEAF32-KD/WT control cells (Figure 5B) showed that it was specifically decreased by ~2 fold upon BEAF-depletion in the proximal sites (+ 50bp) corresponding to the Pol II pausing sites, compared to almost no change in the gene body (+1000 bp) of the same genes (Figure 5B). In addition no significant changes were found for control genes where no binding of BEAF32 or paused Pol II was detected (Figure 5B; data not shown). These results confirmed that insulators may impact transcription at the stage of Pol II pausing as confirmed by our RNA-seq data showing that the differentially expressed genes (WT>KD) were significantly enriched in genes harboring a paused polymerase at genome-wide levels (Figure 5C). Taken together, our data show that insulators function by regulating the progression of Pol II through pausing involving newly identified discrete +1-shifted and +2 nucleosome-positioning patterns.

Our results show that the insulator-mediated regulation of gene expression through nucleosome-positioning may not only involve the NFR (Jin et al., 2009), but also includes novel discrete positioning of nucleosomes over the +1/+2 positions that are conserved through evolution including humans (data not shown). This finding highlights novel chromatin patterns that mark 'paused' genes and all types of insulator sites including GAF/CTCF that also flank NELF/paused sites (Supplementary Figure S8). Insulators could thus serve to guide the +1 followed by a +2-phased nucleosomes to stabilize the binding/pausing of Pol II until activation is triggered by oncogenes (see model in Supplementary Figure S14), thereby synchronizing cell-cycle genes through pausing (Rahl et al., 2010). This activity may be particularly important to regulate

pausing at oncogene-promoters that are highly enriched for BEAF/CTCF binding sites (Emberly et al., 2008; Bushey et al., 2009) as these gene ontologies are enriched in our gene expression analysis. Such disruption of nucleosome phasing may also involve the passage of a transcribing polymerase (Koerber et al., 2009), histone variants (Jin et al., 2009) or remodeling complexes that lead to discrete nucleosome shift/loss upon transcriptional activation through their intrinsic nucleosome disassembly activity and/or the frequent burst of initiation (Boeger et al., 2003; Boeger et al., 2008; Dechassa et al., 2010). Supporting this idea, the insulator co-factor CP190 may recruit remodeling complexes (Bartkuhn et al., 2009) to all types of insulator protein binding sites where it binds (Bushey et al., 2009; Nègre et al., 2010), close to paused genes (Supplementary Figure S8; data not shown). The +2 nucleosomes are more enriched at promoters where no CP190 is present (data not shown) further suggesting that CP190 and additional co-factors will likely contribute to regulate pausing through discrete steps of nucleosome positioning.

The insulator-mediated pausing of Pol II may serve to favor the interactions between promoters and distant elements over alternative/promiscuous promoters as Pol II pausing restricts long-range contacts with bona-fide regulatory elements (see model in Supplementary Figure S14) as shown (Chopra et al., 2009). The thousands of distinct insulators that mark specific nucleosome patterns could thus play a key role in the functional partitioning of genomes (Gaszner and Felsenfeld, 2006; Maeda and Karch, 2007; Bartkuhn and Renkawitz, 2008; Bushey et al., 2009), providing a unique link between the functional organization of chromatin at the level of nucleosome-positioning with higher-order chromatin organization.

MATERIAL AND METHODS

RNAi-mediated depletions, micro-array analysis, Deep-sequencing / RNA-sequencing (Digital Gene Expression), quantitative-RT-PCR

Drosophila Schneider SL2 cells were treated with specific interfering RNAs to knock down BEAF32 or control RNA as a mock control as described previously (Emberly et al., 2008) and mRNA levels were quantified in parallel with cDNA standard curves (Supplementary Figure S12; see Supplementary Information for details). BEAF-depleted (~95%) or control samples obtained from three experiments were analysed by the MGX genomic facility (<http://mgx.transcriptome.eu/>) using both micro-array analysis (Affymetrix Genechip) and RNA-sequencing (RNA-seq; Digital Gene Expression) with the Genome Analyzer II (Illumina)(see Supplementary Information for details).

ChIP-Seq, MNase-Seq, and Statistical Analyses

For ChIP experiments using BEAF32 and Pol II antibodies, chromatin extracts were prepared by sonication of the cells cross-linked with formaldehyde as described previously (Barski et al., 2007). For BEAF32 ChIP-Seq analysis, the ChIP samples were processed to construct the sequencing libraries and sequenced using Illumina Genome Analyzer II (GA II; Barski et al., 2007). To map the genome-wide nucleosome positions using MNase-Seq, BEAF32-depleted and control cells were cross-linked with formaldehyde and digested with micrococcal nuclease. Mononucleosomes were isolated, ligated to Solexa sequencing adaptors, and sequenced using GA II as described

previously (Schones et al., 2008). To generate genome-wide profiles from sequencing reads, first reads were shifted by an average fragment size that was determined from the spatial correlation function between the (+) and (-) reads. Genome-wide profiles were then generated from the read density using a Gaussian smoothing procedure (see Supplementary Information for details). The Pol II ChIP DNA was analysed by qPCR using the oligos listed (Supplementary Table S3) and the pyQPCR software developed in the laboratory available at <http://pyqpcr.sourceforge.net/>.

Self Organization Map (SOM)

Smoothed nucleosome profiles from MNase-Seq in BEAF-bound promoters (Figure 2) were clustered using hierarchical clustering (UPGMA; see Supplementary Information). To generate nucleosome profiles by self-organizing-map (SOM) or k-means, we combined WT and BEAF-KD smoothed nucleosome profiles using a 9x9 SOM and a training on 100,000 samplings using a Gaussian spatial profile with a sigmoidially decreasing radius function ($R/2$) occurring at 30,000 iterations (see Supplementary Information). The result is a trained nucleosome profile at each node in the SOM. To test for enrichment of each nucleosome profile for a particular protein's binding site (Supplementary Figure S7B), a Fisher's exact test was applied to each cluster, using lone promoters/genes (any TSS locating >1.5 kbp away from the next/closest TSS). The significantly enriched nucleosome profiles were further analyzed in each WT or BEAF-KD condition, with Proportional Venn diagram as described (<http://bioinforx.com/>).

ACKNOWLEDGMENTS

We thank D. Gilmour and K. Adelman for discussions and for suggestions, S. Gadat as well as members of O.C.'s group including S. Queille and P. Morillon for suggestions regarding statistical analysis or for technical help, Artem Barski and Kairong Cui for performing the GA II sequencing. E.E's lab was supported by NSERC and the Canadian Institute For Advanced Research (CIFAR). K.Z's lab was supported by the Division of Intramural Research Program of the National Heart, Lung and Blood Institute, NIH. M.H. was supported by a fellowship from La Ligue Nationale Contre le Cancer. O.C.'s lab was supported by grants from the Midi-Pyrénées Region and from the ATIP-AVENIR and joint ARC Program of the CNRS and Inserm.

REFERENCES

Adelman K, Kennedy MA, Nechaev S, Gilchrist DA, Muse GW, Chinenov Y, Rogatsky I. 2009. Immediate mediators of the inflammatory response are poised for gene activation through RNA polymerase II stalling. *Proc Natl Acad Sci USA* **106**: 18207-18212.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* **129**: 823-837.

Bartkuhn M, Renkawitz R. 2008. Long range chromatin interactions involved in gene regulation. *Biochim Biophys Acta* **1783**: 2161-2166.

Bartkuhn M, Straub T, Herold M, Herrmann M, Rathke C, Saumweber H, Gilfillan GD, Becker PB, Renkawitz R. 2009. Active promoters and insulators are marked by the centrosomal protein 190. *EMBO J* **28**: 877-888.

Boeger H, Griesenbeck J, Strattan JS, Kornberg RD. 2003. Nucleosomes unfold completely at a transcriptionally active promoter. *Mol Cell* **11**: 1587-1598.

Boeger H, Griesenbeck J, Kornberg RD. 2008. Nucleosome retention and the stochastic nature of promoter chromatin remodeling for transcription. *Cell* **133**: 716-726.

Bushey AM, Ramos E, Corces VG. 2009. Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. *Genes Dev* **23**: 1338-1350.

Cai H, Levine M. 1995. Modulation of enhancer-promoter interactions by insulators in the *Drosophila* embryo. *Nature* **376**: 533-536.

Chopra VS, Cande J, Hong JW, Levine M. 2009. Stalled Hox promoters as chromosomal boundaries. *Genes Dev* **23**: 1505-1509.

Core LJ, Lis JT. 2008. Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* **319**: 1791-1792.

Cuvier O, Hart CM, Laemmli UK. 1998. Identification of a class of chromatin boundary elements. *Mol Cell Biol* **18**: 7478-7486.

Cuvier O, Hart CM, Käs E, Laemmli UK. 2002. Identification of a multicopy chromatin boundary element at the borders of silenced chromosomal domains. *Chromosoma* **110**: 519-31.

Dechassa ML, Sabri A, Pondugula S, Kassabov SR, Chatterjee N, Klädde MP, Bartholomew B. 2010. SWI/SNF has intrinsic nucleosome disassembly activity that is dependent on adjacent nucleosomes. *Mol Cell* **38**: 590-602.

Emberly E, Blattes R, Schuettengruber B, Hennion M, Jiang N, Hart CM, Käs E, Cuvier O. 2008. BEAF regulates cell-cycle genes through the controlled deposition of H3K9 methylation marks into its conserved dual-core binding sites. *PLoS Biol* **6**: 2896-2910.

Gaszner M, Felsenfeld G. 2006. Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet* **7**: 703-713.

Geyer PK. 1997. The role of insulator elements in defining domains of gene expression. *Curr Opin Genet Dev* **7**: 242-248.

Gilchrist DA, Nechaev S, Lee C, Ghosh SK, Collins JB, Li L, Gilmour DS, Adelman K. 2008. NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev* **22**: 1921-1933.

Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**: 289-293.

Hartigan JA, Wong MA. 1979. A k-means clustering algorithm. *Appl Stat* **28**: 100-108.

Hartley PD, Madhani HD. 2009. Mechanisms that specify promoter nucleosome location and identity. *Cell* **137**: 445-458.

Ioshikhes, I. P., Albert, I., Zanton, S. J. & Pugh, B. F. 2006. Nucleosome positions predicted through comparative genomics. *Nat. Genet.* **38**, 1210-1215.

Jiang N, Emberly E, Cuvier O, Hart CM. 2009. Genome-wide mapping of boundary element-associated factor (BEAF) binding sites in *Drosophila melanogaster* links BEAF to transcription. *Mol Cell Biol* **29**: 3556-3568.

Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, Felsenfeld G. 2009. H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat Genet* **41**:941-945.

Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, LeProust EM, Hughes TR, Lieb JD, Widom J, Segal E. 2009. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**: 362-366.

Koerber RT, Rhee HS, Jiang C, Pugh BF. 2009. Interaction of transcriptional regulators with specific nucleosomes across the *Saccharomyces* genome. *Mol Cell* **35**: 889-902.

Lee C, Li X, Hechmer A, Eisen M, Biggin MD, Venters BJ, Jiang C, Li J, Pugh BF, Gilmour DS. 2008. NELF and GAGA factor are linked to promoter-proximal pausing at many genes in *Drosophila*. *Mol Cell Biol* **28**: 3290-3300.

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Maeda RK, Karch F. 2007. Making connections: boundaries and insulators in *Drosophila*. *Curr Opin Genet Dev* **17**: 394-399.

Mavrich TN, Jiang C, Ioshikhes IP, Li X, Venters BJ, Zanton SJ, Tomsho LP, Qi J, Glaser RL, Schuster SC, Gilmour DS, Albert I, Pugh BF. 2008. Nucleosome organization in the *Drosophila* genome. *Nature* **453**: 358-362.

Mavrich TN, Ioshikhes IP, Venters BJ, Jiang C, Tomsho LP, Qi J, Schuster SC, Albert I, Pugh BF. 2008b. A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res* **18**: 1073-1083.

Milani P, Chevereau G, Vaillant C, Audit B, Haftek-Terreau Z, Marilley M, Bouvet P, Argoul F, Arneodo A. 2009. Nucleosome positioning by genomic excluding-energy barriers. *Proc Natl Acad Sci USA* **106**: 22257-22262.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. 2007. RNA polymerase is poised for activation across the genome. *Nat. Genet.* **39**: 1507–1511.

Nègre N, Brown CD, Shah PK, Kheradpour P, Morrison CA, Henikoff JG, Feng X, Ahmad K, Russell S, White RA, Stein L, Henikoff S, Kellis M, White KP. 2010. A

comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genet* **6**: e1000814.

Raab JR, Kamakaka RT. 2010. Insulators and promoters: closer than we think. *Nat Rev Genet* **11**:439-446.

Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA, Young RA. 2010. c-Myc regulates transcriptional pause release. *Cell* **141**: 432-445.

Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G, Zhao K. 2008. Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132**: 887-898.

Zeitlinger, J., Stark, A., Kellis, M., Hong, J.W., Nechaev, S., Adelman, K., Levine, M., and Young, R.A. 2007. RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat. Genet.* **39**: 1512–1516.

Zhang Y, Moqtaderi Z, Rattner BP, Euskirchen G, Snyder M, Kadonaga JT, Liu XS, Struhl K. 2009. Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. *Nat Struct Mol Biol* **16**: 847-852.

FIGURE LEGENDS

Figure 1. BEAF insulators specify the +1 nucleosome-positioning associated with NELF-binding sites genome-wide

A. The top scheme summarizes our ChIP-Seq data for the genome-wide mapping of BEAF32 insulators juxtaposing (<250bp) to 5,400 promoters (mean distances ~51 bp from TSS), with respect to the positions of nearby nucleosomes (bottom). BEAF32 sites occasionally juxtapose to two genes transcribing in opposite orientation (genes in grey and black).

B. Plot representing the averaged nucleosome-occupancy (y-axis; ChIP-seq data from Mavrich et al., 2008) with all 5,400 aligned BEAF sites, along their promoter regions (x-axis; TSS=0). The red-dotted line highlights how the mean positions of BEAF sites fall within 'gap' regions, where nucleosome-occupancy is low. The black dotted-lines mark the juxtaposed -1, +1 and +2 nucleosomes from 5' to 3'.

C. Top: scheme representing the general pattern of nucleosome-positioning associated with the TSS of 'lone' genes (see text for a definition).

D. Average nucleosome-positioning (mean nucleosome-occupancy; y-axis) measured by MNase-Seq in wild-type cells (this study) and on distinct groups of promoters bound by BEAF32 only (red curve), BEAF32 and NELF (black curve), NELF only (blue curve) or no BEAF32 neither NELF (green curve). The nucleosome profiles are shown for promoter regions of lone genes (x-axis; TSS=0). A similar number of promoters was taken to measure nucleosome-positioning for each distinct group of promoters.

E. Venn diagram showing the overlap between BEAF32 (this study) and NELF (Adelman et al., 2009) binding sites from genome-wide mapping in *Drosophila* S2 cells.

Figure 2. BEAF32 depletion may stabilize nucleosome-positioning

A. Graph representing the results from 3 independent RT-qPCR analyses to measure the mRNA levels of *beaf* (red) and *actin* (blue) in BEAF32 knock-down (BEAF-KD) or WT cells. The error bar corresponds to the standard deviation of the measure.

B. Western blotting of nuclear extracts prepared from control WT- or BEAF-KD cells using the anti-BEAF32 or anti-actin antibodies as indicated on the right.

C. Clustering analysis of MNase-Seq nucleosome signals surrounding TSSs in WT (left) and BEAF-KD (right) cells for hundreds of aligned promoters (y-axis). Each line (x-axis; TSS=0) shows the nucleosome-positioning for the same promoter in WT and BEAF-KD. Nucleosome occupancy is indicated by a color gradient, where blue<green<red, as previously done (Hartley and Madhani, 2009). The scheme on top represents the general pattern of nucleosome-positioning according to our MNase-Seq data (Figure 1C). Note that the BEAF binding sites were similarly distributed with respect to the hundreds of juxtaposed TSSs of down-regulated or unaffected genes (Supplementary Figure S6), showing that the contribution of BEAF32 binding to gene expression does not strictly correlate with a given distance to TSS. This rules out the possibility that BEAF32 affects gene expression only if bound at a given distance from the juxtaposed TSS, as this would indirectly lead the down-regulated genes to align and, consequently, to exhibit a similar nucleosome pattern.

D. The positioning of +1 nucleosome was weakened by knocking down BEAF32. The graph represents the mean value of nucleosome signals obtained from 175 distinct promoter regions.

Figure 3. BEAF32 insulators regulate gene expression involving their function in nucleosome-positioning

A. Proportional Venn diagram showing the intersection analysis between differentially expressed genes (1,901 in euchromatin with WT>BEAF-KD; FDR<0.01) and BEAF32-bound genes.

B. The graph shows the percentage of genes whose promoter is bound by BEAF32 (y-axis; BEAF32 site <250bp from their TSS) as a function of their differential expression in WT compared to BEAF-KD cells (red line = FDR<0.01). ‘Diff genes’ indicate the 1,901 differentially expressed genes (WT>BEAF-KD; FDR<0.01) and ‘control genes’ the same number of control genes (WT ~ BEAF-KD). The x-axis shows the total number of genes (2000, 4000 etc.) included in each differential expression level calculated by a t-test (see Material and Methods). RNASeq data were verified by micro-array and RT-qPCR analysis of gene expression showing a tight correlation (Pearson’s correlation 0.83; p-value~ 3.2e-5) for all genes tested in distinct cell types (Supplementary Figure S4 and S5).

C. Average nucleosome profiles (y=axis) of all differentially expressed genes (WT>BEAF-KD; FDR<0.01) compared to the same number of control genes (WT ~BEAF-KD; see panel B). The nucleosome profiles are shown for promoter regions of lone genes (x-axis; TSS=0).

Figure 4. Clustering approaches identify discrete patterns of insulator-regulated nucleosome-positioning accounting for their role in gene expression

A. Graphs representing the significant nucleosome profiles found in control promoters (left graph) or in BEAF32-bound promoters (right graph; FDR<0.01), as obtained by running self organization map (SOM; see Material and Methods). The '+1' and '+2' signs highlight a distinct feature of BEAF32-bound promoters where the numerous nucleosome profiles are centered near two major positions, +1 and +2. y-axis: nucleosome occupancy; x-axis: position (bp) surrounding the TSS (0).

B. Graphs representing trained grids to identify the most significant/recurrent normalized profiles of nucleosomes obtained by running SOM (see Material and Methods) for control promoter regions (left) or BEAF32-bound promoters in WT (middle) or in BEAF-KD (right) cells. Each square/node corresponds to a learned nucleosome profile (see example in C) and is colored according to its significance, i.e. to the number of promoters sharing a particular profile (brown>red>yellow>green>blue). This identifies three significant nodes associated with BEAF32-bound promoters in WT cells (WT; +1, +1-shifted, +2) and one in BEAF-KD (+1).

C. Two highly significant and distinct +1 nucleosome-positioning profiles (nodes; see B) among WT cells mark BEAF32 insulator sites. Y-axis: nucleosome profiles; x-axis: promoter regions of lone genes (TSS=0).

D. Results from SOM analysis on the 1,901 differentially expressed genes (WT>BEAF-KD; see Figure 3), showing the preferred nucleosome profiles (y-axis) along promoter regions of lone genes (x-axis; TSS=0) independently of BEAF32 binding. SOM was run

on combined WT+BEAF-KD nucleosome-profiles, before testing (Fisher exact test) if a given profile was significant for each condition (WT or BEAF-KD).

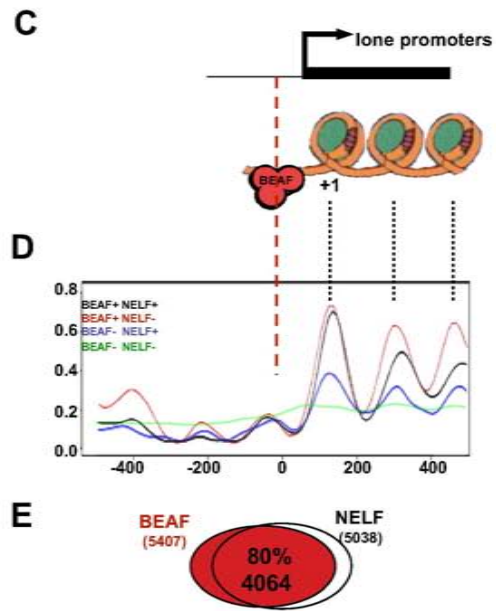
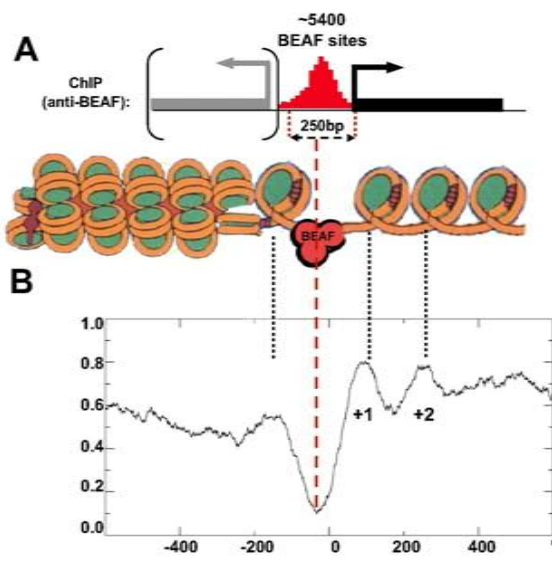
Figure 5. Insulators function to regulate transcriptional pausing

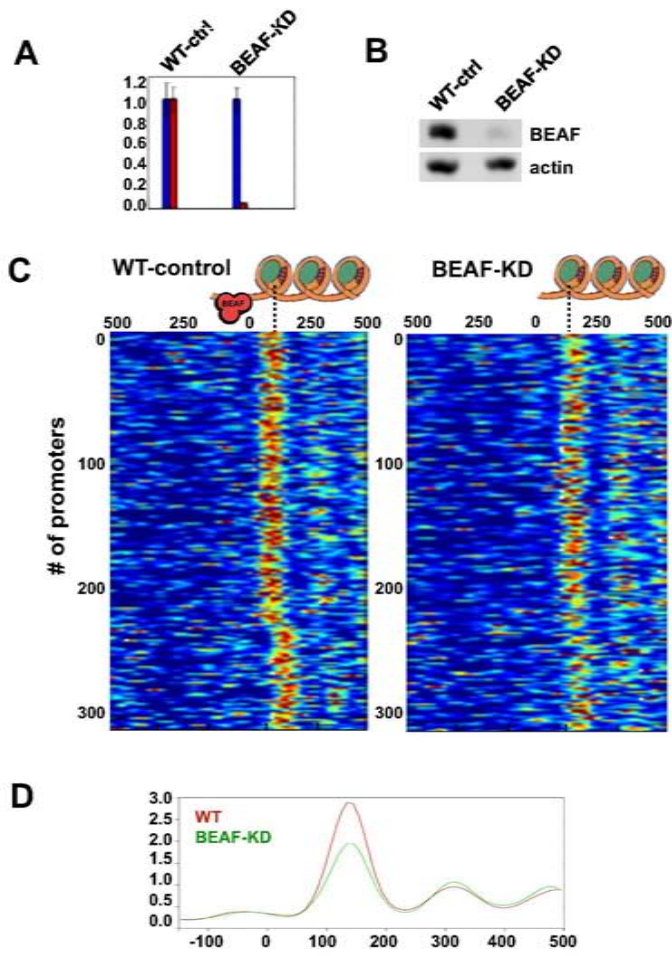
A. Graphs showing the most significant +1/+2 nucleosome profiles (y-axis) along promoter regions of lone genes (x-axis; TSS=0), identified by running k-means independently on positive (+) or on negative (-) sequence reads, for lone-promoters. The nucleosome depicted on top shows where the positions of the +/- reads are located, -73 or +73 bp respectively, in respect to the center of the nucleosome.

B. Graph representing the relative enrichment of Pol II at the sites of pausing (~+50) or gene body (~+1000) obtained by qPCR analysis of ChIP samples from BEAF-KD and WT control cells with anti-Pol II antibodies over IgG controls and plotted as a percentage of input materials. This analysis was done on control 'active' (or inactive) genes and 'paused' genes. Errors bars correspond to a confidence interval of 90%.

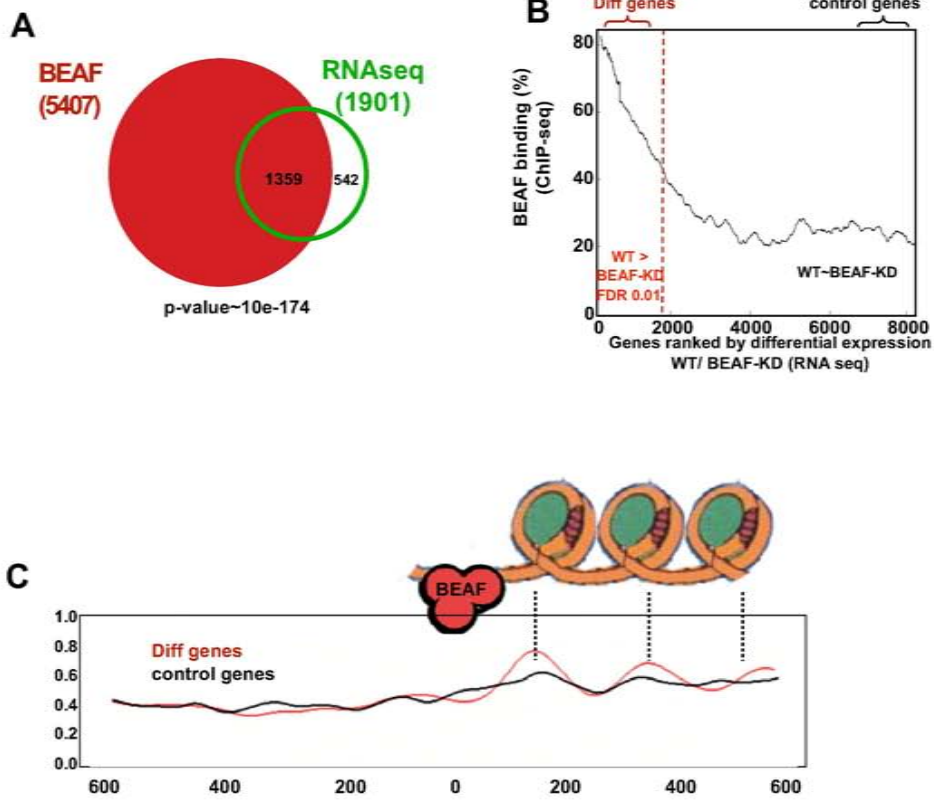
C. Graph showing the intersection analysis of differentially expressed genes ranked according to their t-test metric (x-axis; see Material and Methods) as a function of the percentage of 'paused' over total genes (y-axis). The dotted-line mark the FDR<0.01 limit, for the 1,901 differentially regulated genes (WT>KD; left).

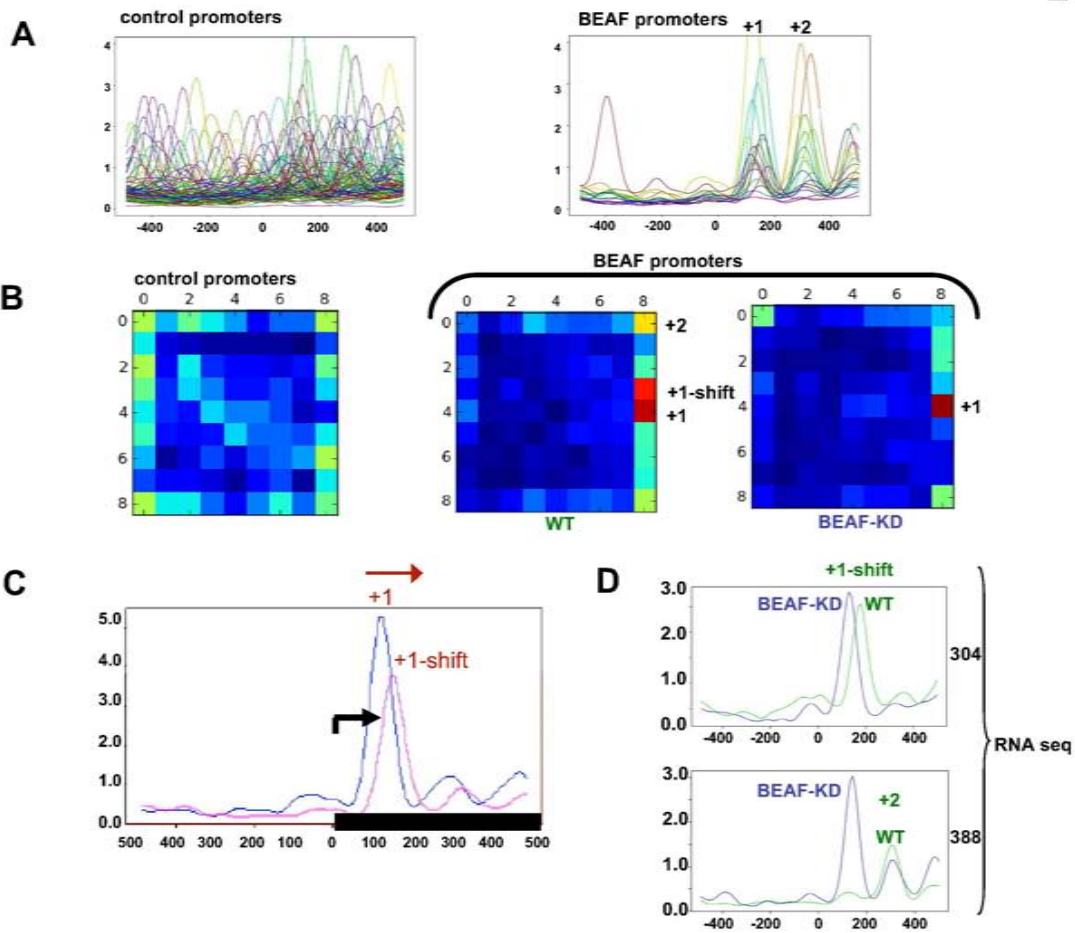
Hennion_Fig1

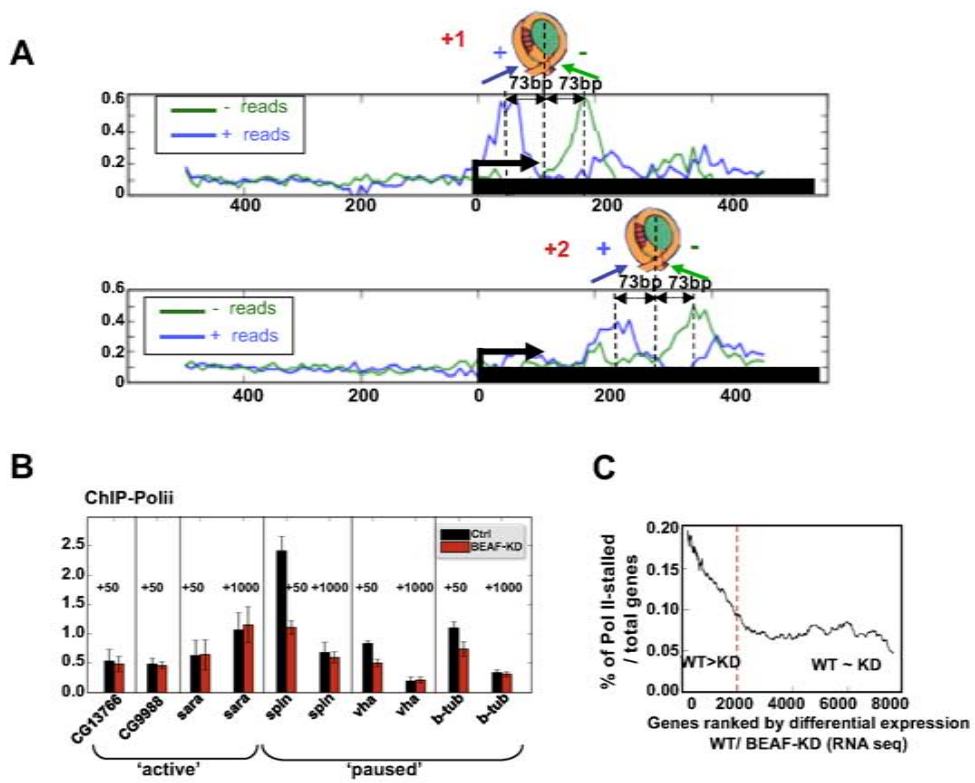




Hennion_Fig3







2.4.2 Rappel des principaux résultats

La première étape de cette étude a consisté à cartographier précisément l'ensemble des sites de liaison de BEAF. Pour cela nous avons fait produire un anticorps spécifique de BEAF efficace pour des immunoprécipitations. Nous avons alors préparé de la chromatine à partir de cellules sauvages que nous avons immunoprécipitée grâce à cet anticorps. L'ADN lié à BEAF a alors été purifié puis séquencé. Plus de 5400 sites de liaison ont ainsi pu être déterminés. De façon intéressante, la plupart de ces sites sont localisés juste à côté du TSS d'un gène, précisément dans la région déplétée en nucléosomes (NFR).

Nous avons également cartographié la position des nucléosomes par une digestion à la MNase, l'isolation de mono-nucléosomes et le séquençage des séquences d'ADN purifié (MNase-seq). Cela a permis de montrer que les sites de liaison de BEAF sont très proches de nucléosomes précisément positionnés (+1 et +2). L'étude du positionnement des nucléosomes par rapport à l'ensemble des gènes a fait ressortir une observation importante, à savoir que les gènes associés à BEAF ont des nucléosomes beaucoup plus précisément localisés que les gènes contrôles n'ayant pas de site de liaison de BEAF. Comme l'étape de pause de la polymérase avait été associée à un positionnement précis des nucléosomes, nous avons vérifié si la pause en tant que telle pouvait positionner ces nucléosomes. De façon intéressante, 80% des gènes associés à BEAF sont également liés par le facteur NELF qui régule cette étape (voir §1.2.3). Cependant, ni la présence de NELF ni celle de la pol II ne suffisent à positionner fortement les nucléosomes. Ces observations nous ont conduit à l'hypothèse selon laquelle BEAF pourrait positionner les nucléosomes et ainsi favoriser l'étape de pause de la pol II.

Pour vérifier l'importance de BEAF dans le positionnement des nucléosomes, nous avons utilisé la même technique de MNase-seq pour cartographier les nucléosomes dans les cellules déplétées de cette protéine. Il en résulte une nette diminution du positionnement des nucléosomes, notamment du +1, ce qui confirme le rôle de BEAF dans le maintien de cette organisation.

Comme nous avons montré précédemment que BEAF régule de nombreux gènes (§2.3), nous avons vérifié si son rôle dans le positionnement des nucléosomes était impliqué dans cette régulation. Pour cela, nous avons séquencé l'ensemble des transcrits de cellules contrôles ou déplétées de BEAF (RNA-seq). Cette étude globale a permis d'identifier ~1900 gènes dont l'expression diminue après la déplétion de BEAF et plus de 500 dont l'expression est augmentée. L'analyse des profils de nucléosomes de ces deux types de gènes montre que les premiers ont un profil très précis, ce qui n'est pas le cas des seconds, ni des gènes non affectés par ce traitement. Pour comprendre le rôle du positionnement des nucléosomes par BEAF dans la régulation transcriptionnelle, nous nous sommes donc focalisés sur la première classe de gènes. La majorité de ces gènes sont effectivement liés par BEAF, ce qui suggère un effet direct de cette protéine.

Pour progresser dans la compréhension des mécanismes mis en jeu, nous avons adopté une technique de regroupement des gènes suivant le profil de nucléosomes qu'ils arborent. Pour cela, une analyse par SOM (Self-Organizing Map) conduite par notre collaborateur Eldon Emberly a fait ressortir les profils les plus significatifs. Il ressort de ces données que les gènes ayant un site de liaison de BEAF sont enrichis en trois profils : un profil ayant un nucléosome +1 bien positionné, un profil assez proche mais où ce nucléosome est décalé d'un dizaine de paires de bases (+1-shifted) et un profil où c'est le second nucléosome qui est le mieux positionné (+2). De façon intéressante, ces deux derniers profils sont perdus lorsque les cellules sont déplétées pour BEAF, ce qui suggère un rôle de BEAF dans le maintien de ces positions.

Pour vérifier si ces changements de profils pouvaient expliquer les changements d'expression observés, nous avons utilisé le même type d'approche (SOM), mais cette fois-ci sur les gènes exprimés différemment après la déplétion de BEAF. Les deux mêmes profils

sont ressortis de cette étude. La disparition du profil +1-shifted permet d'expliquer la régulation de 304 gènes et la disparition du profil +2 celle de 388 autres gènes.

Comme ces profils sont enrichis en gènes en pause, nous avons finalement vérifié l'effet de la déplétion sur la pause de la polymérase. Pour cela, j'ai fait des immunoprécipitations de chromatine avec un anticorps pol II qui montrent que la pause est effectivement stabilisée par BEAF car sa déplétion entraîne une diminution de la quantité de polymérase sur le site de pause, mais pas plus en aval dans le gène.

2.4.3 Quelques données supplémentaires

2.4.3.1 BEAF et NELF

Pour confirmer le rôle de BEAF dans la stabilisation de la pause, nous avons regardé si ce facteur pouvait interagir avec NELF. En effet, BEAF et NELF ont $\sim 80\%$ de sites de liaison en commun, ce qui suggère qu'une interaction est possible. Pour vérifier cela nous avons fait une immuno-précipitation de BEAF sur des extraits nucléaires et regardé si NELF était co-immunoprécipité. On voit sur le *western blot* 2.12 qu'une petite partie de NELF est effectivement liée à BEAF. Ce résultat demande cependant à être confirmé, notamment par l'expérience inverse en immunoprécipitant NELF. Cette interaction pourrait expliquer d'un point de vue plus fonctionnel l'effet de BEAF sur la stabilisation de la pause de la pol II.

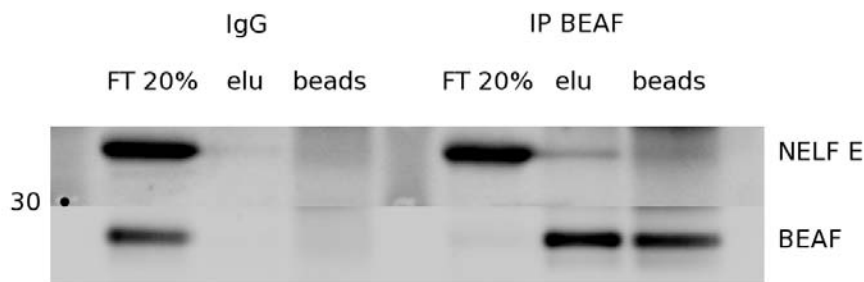


Figure 2.12 – Co-immunoprécipitation de BEAF et NELF. Ce *western blot* montre la liaison de BEAF sur les billes et la fraction éluée (ligne « elu »), ainsi que le fait qu'une fraction de la protéine NELF est effectivement co-précipitée avec BEAF. En contrôle est présentée la même expérience conduite avec des IgG purifiés à la place de l'anticorps anti-BEAF. FT : flow-through (ce qui n'est pas retenu sur les billes); elu : fraction éluée; beads : fraction retenue sur les billes après l'éluion.

2.4.3.2 Vérification de la reproductibilité

Pour conforter nos données de MNase-seq et vérifier la reproductibilité des résultats, j'ai digéré avec de la MNase de la chromatine issue de cellules contrôles ou déplétées de la protéine BEAF (expérience indépendante de celles présentées dans l'article). J'ai alors effectué un *Southern blot* avec une sonde spécifique d'un gène associé à BEAF. Cette expérience montre qu'il y a effectivement un changement dans l'accessibilité de la chromatine dû à la déplétion de BEAF (Fig. 2.13). En effet, après traitement des cellules, la chromatine est moins digérée par la MNase, ce qui suggère une structure plus compacte, en accord avec la présence de la marque répressive H3K9me3. En plus de ce changement dans l'accessibilité de la chromatine, le positionnement des nucléosomes est altéré par la déplétion de BEAF. En comparant ce gel avec nos données de MNase-seq, on observe une bonne adéquation entre ces deux techniques bien qu'elles aient été utilisées sur des expériences de déplétion de BEAF indépendantes. Malgré la résolution bien plus basse du *Southern blot*, on retrouve un profil similaire, avec notamment le changement de localisation du premier nucléosome en aval du TSS. Cela suggère que la cartographie

des nucléosomes est assez reproductible. L'analyse d'un seul locus n'est cependant pas suffisante et ne donne qu'une indication sur cette reproductibilité. En particulier, il faudrait également analyser un locus contrôle où BEAF n'a pas d'effet sur le positionnement des nucléosomes.

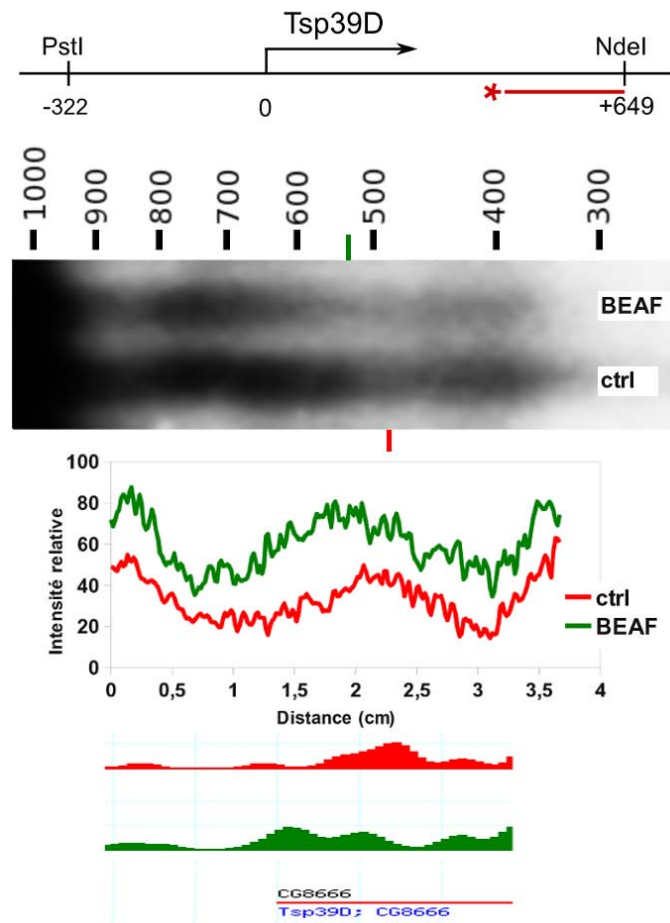


Figure 2.13 – Adéquation du profil de nucléosomes obtenu par *Southern blot* et par MNase-seq. La photo illustre un *Southern blot* avec une sonde spécifique de la région 5' du gène *Tsp39D*. Le signal a été quantifié grâce au logiciel ImageJ. En dessous sont reproduites les données de MNase-seq qui correspondent assez bien avec les régions moins digérées sur le gel.

Il serait intéressant de confirmer ce changement d'accessibilité de la chromatine qui suggère un changement dans l'occupation des nucléosomes en plus de l'effet sur le positionnement. Cette question de l'occupation n'a pas été étudiée dans cet article et mérite de l'être dans le futur.

2.4.3.3 Retour sur la région frontière eu-/hétérochromatine

Ces données permettent d'expliquer en partie les résultats présentés dans la première partie (§2.2.5). Nous avons montré que l'expression de gènes associés à des sites prédits de liaison de BEAF est affectée lorsque ce facteur est déplété. La plupart des gènes du côté hétérochromatinien du BE28 se comportent de la même façon que les gènes euchromatiniens étudiés dans cet article. Ceux qui sont associés à un site de liaison de BEAF (voir Fig. 2.14) sont activés par BEAF de la même façon que les gènes associés à un *dual-core* sans site de liaison de DREF. Cela peut s'expliquer par l'hypothèse selon laquelle au moins une partie de ces gènes se comportent comme des îlots d'euchromatine au sein d'un contexte hétérochromatinien, ce type de structure ayant été observé dans le chromosome 4 (Sun *et al.* 2004; Riddle *et al.* 2008).

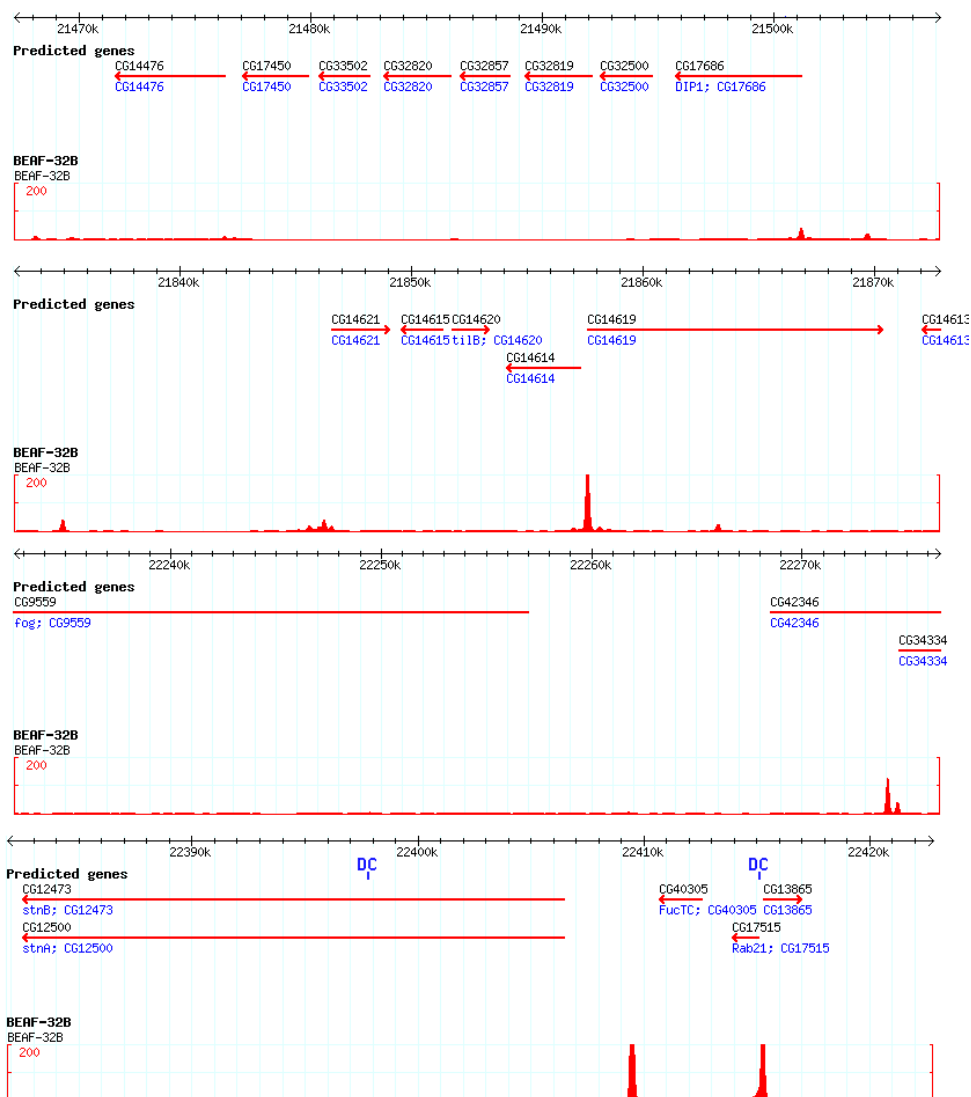


Figure 2.14 – Localisation des sites de liaison de BEAF sur l'extrémité de la partie séquencée du chromosome X. Les sites réels de liaison de BEAF sont représentés (pics rouges) par rapport aux gènes d'intérêt (voir §2.2.1 et Fig. 2.3). Les *dual-cores* sont indiqués en bleu (DC).

Pour vérifier cela, j'ai traité les cellules avec des ARNdb permettant de dépléter la protéine HP1 qui est nécessaire à l'expression de gènes hétérochromatiniens (Lu *et al.* 2000). Cette analyse montre que la plupart des gènes sensibles à la déplétion de BEAF et associés à un site de liaison de cette protéine ne sont pas affectés par l'absence de HP1 (Fig. 2.15, *Rab21*, *CG13865* et *CG14624*). Cela suggère qu'ils puissent avoir une structure euchromatinienne. Le gène *stoned* constitue cependant une exception notable. En effet, son expression chute considérablement après déplétion de HP1. Ce gène est assez typique des gènes hétérochromatiniens dans la mesure où il est relativement long (~23 Kpb) et possède de longs introns composés d'éléments transposables (Dimitri *et al.* 2003). Sa sensibilité à la déplétion de HP1 va aussi en ce sens. Cela prouve qu'au moins une partie des gènes de cette région ont besoin d'un contexte hétérochromatinien pour être exprimés normalement. L'effet de BEAF sur ce gène pourrait être dû à un mécanisme différent de ce qui est observé pour *Rab21*, *CG13865* et *CG14624*. En effet, la localisation du site de BEAF est nettement plus éloignée du TSS, ce qui suggère que BEAF pourrait affecter ce gène par un effet indirect. Si, comme dans les gènes euchromatiniens (voir §2.3.2), la marque H3K9me3 augmente lorsque BEAF est déplétée, il se pourrait que HP1 soit relocalisé sur cette marque et soit donc moins disponible pour faciliter l'expression de *stoned*. Il serait intéressant de tester cette hypothèse par des expériences de ChIP anti-HP1 et anti-H3K9me3. Ces observations sont en accord avec l'idée qu'il existe aux régions frontières des sous-domaines où des régions euchromatiniennes alternent avec des régions hétérochromatiniennes (voir 3).

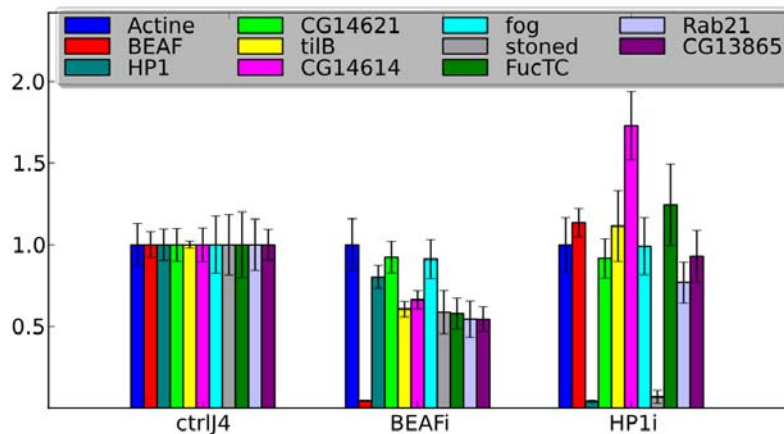


Figure 2.15 – Effet de la déplétion de HP1 sur les gènes centromère-proximaux du BE28 (voir Fig. 2.3). Les quantifications relatives par rapport à l'échantillon contrôle sont représentées. Les barres d'erreur correspondent à un intervalle de confiance de 90%.

Les gènes *tilB* et *FucTC* sortent également du schéma présenté dans l'étude précédente. En effet, ils ne présentent pas de site de liaison de BEAF mais leur expression est systématiquement réduite après sa déplétion. Ils ne sont par contre pas affectés par la diminution de HP1. Dans ce cas, il semble clair que ce n'est pas un effet direct de BEAF. Cela peut provenir de la dérégulation d'un autre facteur qui serait important pour l'expression de ces gènes. Alternativement, cela pourrait aussi être une conséquence de changements de l'organisation du domaine dus à la perte de BEAF sur le promoteur d'un gène proche (*CG14614* pour *tilB* ou *Rab21* pour *FucTC*) ou plus éloigné.

2.4.4 Conclusion

Cette étude conduit à des avancées importantes dans la compréhension des mécanismes de régulation transcriptionnelle par les insulateurs. Au moins dans le cas de BEAF, celle-

ci fait intervenir le positionnement précis des nucléosomes +1 et +2. De plus, un lien intéressant avec l'étape de pause de la polymérase a été découvert. En effet, BEAF s'associe à NELF pour stabiliser cette étape à de nombreux loci et éventuellement réguler la progression de la pol II (Fig. 2.16).

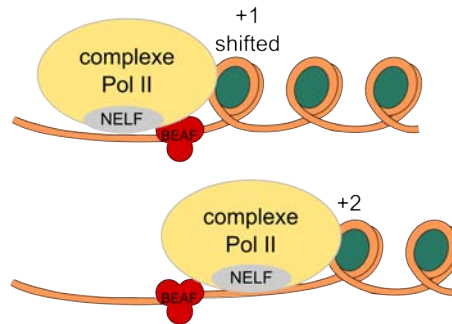


Figure 2.16 – Représentation de la régulation de la pause de polymérase par BEAF via la positionnement des nucléosomes. BEAF interagit avec NELF et positionne les nucléosomes (+1 et +2) pour stabiliser la pause.

Notre modèle propose que la pause de la pol II stabilisée par les insulateurs pourrait favoriser l'interaction d'un élément régulateur distant avec le « bon » promoteur et éviter l'activation anormale d'autres promoteurs (Fig. 2.17). Nous proposons que cet état de pause perdure jusqu'à ce qu'un signal d'activation (la liaison d'un facteur oncogène par exemple) permette la reprise de l'élongation et une expression rapide du gène. Ce mécanisme pourrait permettre de synchroniser l'activation de gènes impliqués dans la réponse à un même stimulus (voir §3).

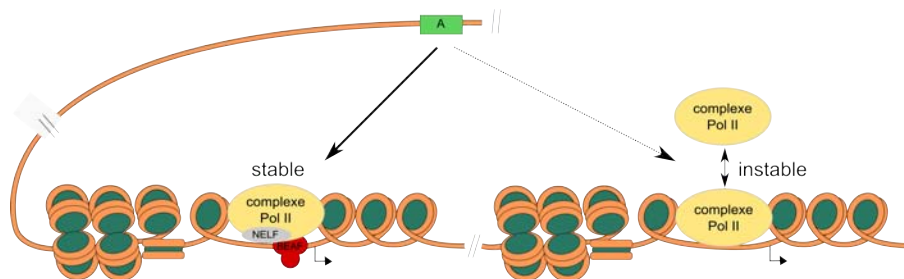


Figure 2.17 – Modèle montrant que la pause de la pol II pourrait favoriser l'interaction avec un activateur. Au moment de l'arrivée d'un signal activateur, l'activateur (A) pourrait agir plus facilement avec une polymérase en pause, par exemple en ayant plus de temps pour interagir avec elle, qu'avec une polymérase en liaison instable avec un promoteur.



Discussion et perspectives

3

Sommaire

3.1	Nature des régions frontières	126
3.2	Fonctions de BEAF <i>in vivo</i>	127
3.2.1	Les éléments <i>scs</i> et <i>scs'</i>	127
3.2.2	Activation transcriptionnelle	127
3.2.3	Répression transcriptionnelle	129
3.3	Fonctions des autres insulateurs	129
3.3.1	Points communs	129
3.3.2	Spécialisation des insulateurs	130
3.3.3	Rôles dans d'autres processus cellulaires que la transcription	131
3.4	Formation de boucles de chromatine	132
3.4.1	Contraintes mécaniques	132
3.4.2	Implication du complexe cohesin	133
3.5	Régulation de la fonction insulatrice	133
3.5.1	Régulation du recrutement des facteurs insulateurs par d'autres protéines	133
3.5.2	Modifications post-traductionnelles des protéines insulatrices	134
3.5.3	Implication d'ARN	135
3.6	Les insulateurs chez les vertébrés	135

Ce chapitre discute les résultats obtenus et propose des perspectives à ce travail. Je discuterai tout d'abord la nature des régions de transition entre l'hétérochromatine et l'euchromatine. Je m'intéresserai ensuite aux différents rôles de BEAF *in vivo* ainsi qu'à ceux des autres protéines insulatrices de la drosophile. Je soulèverai alors la question des mécanismes de formation des boucles de chromatine, puis celle de la régulation de la fonction insulatrice. Enfin, j'étendrai les perspectives de ce travail aux insulateurs présents chez les vertébrés.

3.1 Nature des régions frontières

Nous avons montré que les répétitions d'éléments BE28 liés par BEAF ne constituent pas une barrière empêchant la propagation des structures chromatinienne. Les changements d'expression observés dans la lignée *w^{m4}* ne sont donc pas seulement la conséquence de la perte de la fonction insulatrice des BE28. En effet, la déplétion de la protéine insulatrice BEAF dans la lignée sauvage n'induit pas une diminution de l'expression des gènes du côté euchromatinien. Nous nous sommes focalisés sur l'élément BE28 mais il est possible que la frontière entre euchromatine et hétérochromatine, si elle existe, soit constituée par d'autres éléments. Une séquence plus proche du rDNA a d'ailleurs été proposée pour remplir cette fonction. Il s'agit d'une région constituée d'au moins quatre répétitions en tandem de 5.9 kb qui a été séquencée par une marche chromosomique depuis le gène *su(f)* (O'Hare *et al.* 2002), mais aucune caractérisation fonctionnelle de cette région n'a été effectuée depuis.

Il est cependant probable qu'il n'y ait pas une unique région définissant une frontière nette entre l'euchromatine et l'hétérochromatine. Ainsi, cette région de transition semble plutôt constituée de séquences hétérochromatiniennes (dont probablement les différents trous dans la séquence dus aux répétitions) contenant des îlots euchromatiniens. C'est ce que nous observons dans la région de transition du chromosome X où certains gènes sont activés par HP1 alors que d'autres ne sont pas affectés. Ce type d'organisation est d'ailleurs retrouvé sur le chromosome 4. En effet, ce petit chromosome (4,2 Mpb) contient une région de 1,2 Mpb observable dans le chromosome polytène (donc endo-répliquée contrairement à l'hétérochromatine centromérique) qui est constituée d'un mélange des deux types de structures avec ~35 % de séquences répétées (contre 6% dans les régions euchromatiniennes des autres chromosomes) et une densité génique élevée mais sans qu'il y ait de grands blocs de l'une ou l'autre structure (pour revue : Riddle *et al.* 2009).

S'il n'y a pas de frontière physique entre l'hétérochromatine et l'euchromatine, d'autres mécanismes doivent être impliqués dans le maintien de ces structures. Une possibilité repose simplement sur la loi d'action des masses. En effet, si la quantité de facteurs maintenant une structure (par exemple HP1) est limitée, alors il se forme naturellement un gradient entre les régions ayant une forte affinité pour ces facteurs et les autres. À ce moment là, la quantité de protéines constituant un type de chromatine devrait influencer la taille des domaines (Locke *et al.* 1988). C'est ce qui se produit quand la synthèse de constituants essentiels de l'hétérochromatine est bloquée : on observe alors une réduction de la formation de l'hétérochromatine et de la mise en silence associée. Au contraire, quand la synthèse de protéines euchromatiniennes est bloquée, les domaines d'hétérochromatine s'étendent. Ainsi les frontières des domaines chromatinien pourraient ne pas être déterminés par des éléments d'ADN, mais plutôt résulter d'une compétition entre deux processus dépendants de la concentration en activateurs et en répresseurs (Kimura & Horikoshi 2004, Fig. 3.1).

Il existe des mécanismes actifs permettant de préserver les domaines euchromatiniens de l'invasion par l'hétérochromatine. Par exemple, des enzymes spécifiques peuvent être recrutées. C'est notamment le cas de la kinase Jil1 qui arrête la propagation de l'hétérochromatine en phosphorylant H3S10, ce qui empêche la méthylation d'H3K9 et le recrutement de HP1 (Zhang *et al.* 2006).

Ce type de propagation ne suffit cependant pas à expliquer la localisation de HP1 observée dans une lignée *white-mottled*. En effet, celle-ci ne forme pas un gradient au niveau de la cassure mais des pics discontinus (Vogel *et al.* 2009). Pour expliquer cette observation, un modèle en boucles a été proposé (voir §2.2.7). La connaissance de la structure tri-dimensionnelle des régions de transitions pourrait apporter d'importantes informations sur les mécanismes de régulation des gènes qui s'y trouvent.

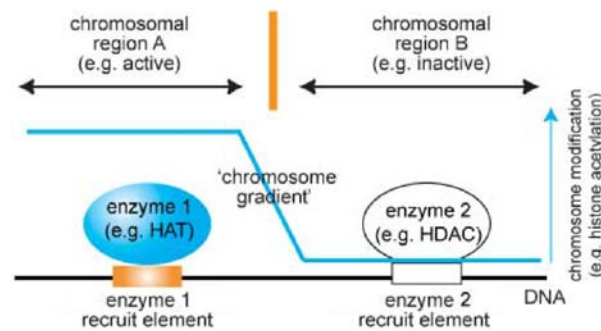


Figure 3.1 – Modèle représentant une barrière non fixe résultant de la compétition entre deux types de structure (adapté de Kimura & Horikoshi 2004). Des éléments d'ADN recrutent des effecteurs qui propagent des marques épigénétiques. Lorsque deux structures mutuellement exclusives entrent en collision, un gradient de ces marques se forme par compétition.

3.2 Fonctions de BEAF *in vivo*

3.2.1 Les éléments *scs* et *scs'*

La démonstration de la fonction insulatrice des éléments *scs* et *scs'* a amené l'hypothèse que ces éléments limitaient la propagation du « puff » de chromatine ouverte observé lors d'un choc thermique (Udvardy & Schedl 1993). Ce modèle de frontière a cependant été controversé par une étude cartographiant précisément ces éléments sur des chromosomes polytènes. Ce travail montre que ces éléments ne sont pas à la frontière entre le « puff » engendré par choc thermique et l'hétérochromatine mais bien à l'intérieur du « puff » (Kuhn *et al.* 2004). Ainsi, *scs* et *scs'* ne forment pas une barrière et la question de savoir ce qui limite réellement la propagation du « puff » reste entière. Il a été suggéré que *scs* et *scs'* pourraient plutôt avoir un rôle dans la régulation transcriptionnelle des promoteurs associés (Kuhn *et al.* 2004; Gaszner & Felsenfeld 2006). Cependant, l'étude des insulateurs sur les chromosomes polytènes pose un problème. En effet, de plus en plus de données suggèrent que les insulateurs agissent en formant des boucles de chromatine et de telles boucles sont difficilement observables sur les polytènes qui présentent généralement une structure linéaire. Une étude récente sur la réorganisation des nucléosomes de cette région après un choc thermique a montré que les nucléosomes sont éjectés indépendamment de la transcription dans toute la région entre les séquences *scs* et *scs'* (Petesch & Lis 2008). Par contre, les nucléosomes à l'extérieur de cette région ne sont pas affectés par le choc thermique. Ces données montrent que les éléments *scs* et *scs'* ont effectivement la capacité de partitionner ce locus en domaines de chromatine indépendants. Il est à noter que le gène *aurora* qui est situé dans le « puff » mais à l'extérieur du domaine défini par les séquences *scs* et *scs'* n'est pas induit au moment du choc thermique, ce qui corrèle avec l'observation qu'il n'y a pas de changement de densité de nucléosomes dans ce gène. Ainsi, l'insulateur *scs'* et BEAF pourraient être importants pour maintenir la répression de *aurora* au moment du choc thermique.

3.2.2 Activation transcriptionnelle

Nous avons établi un lien important entre la protéine BEAF et la régulation de la transcription. En effet, de nombreux sites de liaison de BEAF sont localisés à proximité d'un site de démarrage de la transcription d'un gène actif. La déplétion de BEAF induit une dérégulation positive ou négative du gène suivant les cas. Les gènes dont le promoteur ne présente pas de site de liaison de l'activateur DREF sont généralement activés par BEAF (voir §2.3). Le mécanisme de cette activation met en jeu l'inhibition de la déposition de la marque répressive H3K9me3 ainsi qu'un contrôle du positionnement des nucléosomes en

aval du TSS.

Nous avons montré que BEAF avait un rôle dans le positionnement des nucléosomes qui est impliqué dans l'activation transcriptionnelle de centaines de gènes (voir §2.4). En effet, les gènes dont l'expression baisse après déplétion de BEAF présentent un positionnement très précis qui est perdu lors de ce traitement. Des profils particuliers de nucléosomes (+1-shifted et +2) sont perdus et ces changements sont associés à la dérégulation de gènes précis.

Le mécanisme de ce repositionnement des nucléosomes par BEAF reste à élucider. En effet, il pourrait s'agir soit d'un effet direct de la liaison BEAF, soit du recrutement par BEAF de facteurs de remodelage de la chromatine. L'étude des facteurs associés à BEAF est en cours au laboratoire et devrait permettre de tester cette hypothèse.

Nous avons montré que ce changement de positionnement corrèle avec la déstabilisation de la pause de la polymérase II, ce qui explique le phénotype observé. Ainsi, un nouveau rôle important et inattendu de BEAF a été mis en évidence, à savoir la stabilisation de l'étape de pause qui fait vraisemblablement intervenir une interaction de BEAF avec le facteur NELF.

Il serait intéressant de déterminer s'il existe un lien entre la régulation du positionnement des nucléosomes et la formation de boucles de chromatine. En effet, il n'est pas évident de comprendre en quoi la formation de boucles pourrait influencer la localisation précise des nucléosomes, ou à l'inverse comment un positionnement particulier pourrait stabiliser ce type de structure en boucles. Il est possible que les deux mécanismes soient disjoints. BEAF pourrait positionner précisément les nucléosomes et interagir avec d'autres facteurs comme CP190 pour regrouper des gènes en pause et permettre leur corégulation en réponse à des stimuli extérieurs. Pour étudier l'organisation tri-dimensionnelle des gènes en pause régulés par BEAF, nous voulons mettre en place des expériences de 3C.

Une étude récente suggère une interprétation intéressante de nos résultats. En effet, il a été montré que le promoteur de gènes en pause avait lui-même une fonction d'insulateur (Chopra *et al.* 2009). Pour expliquer cette observation, il a été proposé que les promoteurs en pause pourraient interagir avec des insulateurs et former des boucles reproduisant la fonction de blocage d'un activateur (Fig. 3.2).

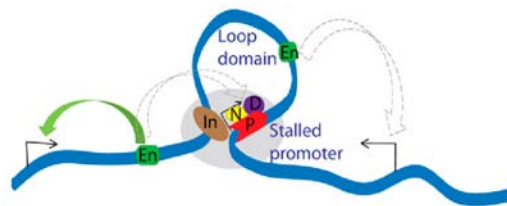


Figure 3.2 – Modèle expliquant la fonction insulatrice d'un promoteur en pause (*stalled*; Chopra *et al.* 2009). Un promoteur en pause pourrait interagir avec un insulateur (In) et former une boucle qui contient et isole l'activateur (En) des promoteurs (représentés par des flèches noires) situés à l'extérieur de la boucle. P : polymérase II; D : DSIF; N : NELF.

Nous avons à présent une autre explication étant donné que BEAF stabilise la pause en se liant directement au promoteur. Ainsi, la propriété d'insulateur des séquences liées par BEAF pourrait être une conséquence de sa capacité à stabiliser la polymérase en pause. L'insertion d'un insulateur entre un activateur et promoteur pourrait engendrer un recrutement ectopique de la polymérase II sur l'insulateur qui interagirait alors avec l'activateur, l'empêchant d'activer sa cible. Cette hypothèse correspond au modèle de leurre de promoteur. En accord avec cette idée, lorsque l'insulateur HS4 est placé entre un activateur et un promoteur, la polymérase s'accumule au niveau de l'insulateur et non plus au niveau du promoteur (Zhao & Dean 2004). La question de la directionnalité de l'activité insulatrice n'est cependant pas expliquée par ce modèle. Il est intéressant de constater que

le test fonctionnel permettant de définir un insulateur ne reflète pas le cas général *in vivo*. En effet, au moins dans le cas des insulateurs liés par BEAF, ces éléments ne sont quasiment jamais situés entre un activateur et un promoteur comme dans ce test, mais au niveau des promoteurs eux-mêmes. Le modèle de boucles indépendantes où un activateur ne pourrait agir avec un promoteur que s'il est dans la même boucle ne fonctionne pas dans le cas où l'insulateur est localisé au niveau du promoteur. Des études supplémentaires sont nécessaires pour imaginer puis tester pourquoi un insulateur possède cette fonction de bloqueur de communication uniquement quand il est inséré entre un activateur et un promoteur.

3.2.3 Répression transcriptionnelle

Nous avons montré qu'environ 500 gènes voient leur expression augmenter lorsque BEAF est déplété. Il pourrait s'agir d'un effet indirect dû à la dérégulation de l'expression d'autres loci. Cependant, les promoteurs de ces gènes sont souvent associés à un site de liaison de BEAF, ce qui suggère un effet direct pour au moins une fraction de ces gènes. Une partie d'entre eux présente un site de liaison de DREF qui chevauche le site de BEAF. Nous avons montré qu'un mécanisme de compétition entre ces deux facteurs explique la fonction répressive de BEAF sur ces gènes (voir §2.3). En effet, la présence de BEAF limite la liaison de DREF et ainsi empêche l'activation par ce second facteur. En l'absence de BEAF, DREF se lie et le gène est surexprimé. Cependant, il n'est pas certain que tous les gènes surexprimés quand BEAF est déplété présentent un site de liaison de DREF. D'autres mécanismes pourraient être impliqués.

On peut faire un parallèle avec ce qui a été observé lors de la déplétion de NELF (Gilchrist *et al.* 2008). En effet, là aussi la réponse dépend des gènes, certains sont activés alors que d'autres sont réprimés. Pour expliquer cela, il a été proposé que la réponse dépend de l'étape limitante entre le recrutement de la pol II sur le site de pause, et la reprise de l'élongation. Ainsi, pour les gènes qui sont activés en l'absence de NELF, ce facteur stabilise la pause et son absence accélère la reprise de l'élongation. À l'inverse, pour les gènes qui sont réprimés en l'absence de NELF, ce facteur apparaît comme important pour le recrutement de la pol II et l'étape de pause n'est pas limitante. Nous avons mis au point des expériences de déplétion de NELF et nous allons vérifier par RNA-seq si les gènes affectés par l'absence de BEAF ou NELF sont globalement les mêmes. Étant donné le large recouvrement entre les sites de liaison de ces deux facteurs (~80 %) il est probable qu'au moins une partie des gènes affectés soit commune. Il n'est cependant pas certain que le sens (activation ou répression) d'action de ces facteurs soit le même. Cette étude devrait nous aider à comprendre le mécanisme de répression de certains gènes par BEAF.

Parmi ces gènes activés en l'absence de BEAF se trouvent les gènes *hsp70A* localisés entre les insulateurs *scs* et *scs'*. Comme il a été montré que BEAF intervient dans la formation d'une boucle entre ces insulateurs, on peut supposer que la déplétion de BEAF ouvre la boucle, ce qui pourrait avoir pour conséquence l'activation anormale des gènes qui sont normalement en pause (rapidement activables par choc thermique). Ce cas est différent de ce qui a été présenté précédemment car BEAF n'est pas présent sur les promoteurs des gènes *hsp70A* et n'a donc pas un rôle direct de stabilisation de la pause (Fig. 3.3).

3.3 Fonctions des autres insulateurs

3.3.1 Points communs

Quelle que soit la protéine insulatrice qui les lie, une grande partie des insulateurs de la drosophile colocalisent avec CP190 (Bartkuhn *et al.* 2009). De plus, CTCF interagit avec CP190 (Mohan *et al.* 2007; Gerasimova *et al.* 2007) et Su(Hw) nécessite ce facteur pour être recruté sur les insulateurs *gypsy* (Pai *et al.* 2004). Des données préliminaires de notre



Figure 3.3 – Localisation de BEAF dans la région des gènes *hsp70A* (ChIP-seq). On voit nettement le pic de BEAF sur l'élément *scs'*. Un petit pic dû à la formation de la boucle est visible sur *scs*. Par contre, il n'y a pas de signal sur les promoteurs des gènes *hsp70A* qui sont pourtant activés quand BEAF est déplété.

groupe indiquent que BEAF pourrait également interagir avec CP190. Ce facteur semble donc être important pour le fonctionnement de tous les insulateurs de la drosophile. Il pourrait participer à la formation de boucles de chromatine comme cela a été proposé pour les insulateurs *gypsy* (Pai *et al.* 2004).

Un autre point commun entre BEAF et les autres protéines insulatrices concerne leur rôle dans le positionnement des nucléosomes. Ainsi, les sites de liaison de CTCF sont associés à un fort positionnement d'un dizaine de nucléosomes de chaque côté du site (Fu *et al.* 2008). Ce résultat suggère que CTCF pourrait avoir le même rôle que ce que nous avons montré pour BEAF qui est aussi associé à des sites où les nucléosomes sont très précisément positionnés. Un rôle important de GAF est sa capacité à recruter le complexe de remodelage de la chromatine NURF qui remodelé les nucléosomes pour créer une configuration de chromatine ouverte (Adkins *et al.* 2006). GAF est également nécessaire à l'éviction des nucléosomes des gènes *hsp70A* après un choc thermique (Petesch & Lis 2008). Cependant son rôle dans le positionnement des nucléosomes n'a pas été adressé à l'heure actuelle.

Il serait intéressant d'étudier plus précisément le rôle des autres protéines insulatrices de la drosophile dans le positionnement des nucléosomes, notamment de celles qui sont conservées chez d'autres espèces animales. Cela pourrait permettre de prédire si ces mécanismes de régulation pourraient exister chez ces espèces. C'est assez probable car le positionnement des nucléosomes au niveau des TSS est conservé, y compris chez l'humain où CTCF et probablement GAF pourraient jouer le même rôle que ce que nous avons montré pour BEAF chez la drosophile.

3.3.2 Spécialisation des insulateurs

L'étude de la localisation des différents types d'insulateurs de la drosophile par rapport aux gènes montre des différences qui suggèrent qu'ils ont des fonctions spécifiques. En effet, alors que Su(Hw) est localisé plutôt loin des gènes, CTCF et BEAF sont retrouvés en majorité juste en amont des TSS de gènes fortement exprimés (Emberly *et al.* 2008; Bushey *et al.* 2009; Jiang *et al.* 2009).

Les différentes protéines insulatrices de la drosophile peuvent favoriser des interactions inter- ou intra-chromosomiques en interagissant entre elles via leur domaine BTB. Ainsi, il a été montré par immunofluorescence que les protéines Su(Hw), Mod(mdg4)2.2 (Gerasimova *et al.* 2000), CP190 (Pai *et al.* 2004) et CTCF (Gerasimova *et al.* 2007) ont une distribution en foyers d'un nombre limité dans le noyau des cellules interphasiques. Ces foyers d'insulateurs étant d'un nombre bien inférieur au nombre de sites de liaison de ces protéines, il a été proposé qu'il s'agisse de grands agrégats de protéines insulatrices interagissant entre elles, et regroupant ainsi les séquences cibles (Fig. 3.4). Cela pourrait permettre de coréguler des gènes en augmentant localement la concentration en protéines régulatrices.

La localisation de Su(Hw) dans des régions intergéniques suggère que cette protéine



Figure 3.4 – Modèle représentant un foyer d’insulateurs (Capelson & Corces 2004). En jaune et bleu sont représentés les nucléosomes de la fibre chromatinienne. En rose et vert sont représentées des protéines insulatriques qui regroupent des gènes pour qu’ils soient corégulés et/ou qui définissent des domaines de chromatine indépendants (boucles).

pourrait permettre la formation de boucles de chromatine pour délimiter des domaines fonctionnels. Les foyers formés par les insulateurs *gypsy* sont associés à la lamine et se trouvent à la périphérie nucléaire (Gerasimova *et al.* 2000). Ces insulateurs ne semblent pas avoir de rôle direct dans la régulation de la transcription des gènes. Une partie des sites liés par CTCF sont également retrouvés dans des régions intergéniques et pourraient fonctionner comme les sites liés par Su(Hw).

La localisation des sites de liaison de BEAF et du reste de ceux de CTCF près du TSS de gènes hautement transcrits peut signifier que, par des mécanismes analogues, ils servent à localiser ces régions dans des compartiments spécialisés du noyau comme les usines de transcription (Gurudatta & Corces 2009). En effet, la colocalisation des insulateurs CTCF pourrait refléter leur regroupement au sein d’un même foyer de transcription. Des expériences d’immunofluorescence sont actuellement en cours au sein de notre groupe pour étudier la localisation de BEAF par rapport foyers de la pol II.

3.3.3 Rôles dans d’autres processus cellulaires que la transcription

Il existe quelques données qui suggèrent que les insulateurs pourraient participer à la régulation de la réplication. Il a été montré chez la drosophile que les loci des groupes de gènes codant les protéines du chorion sont normalement amplifiés environ 80 fois pendant l’ovogenèse, mais que cette amplification décroît considérablement si un de ces loci est inséré à un autre site génomique. En ajoutant des insulateurs *gypsy* de part et d’autre de ce locus, celui-ci devient insensible à cet effet de position (Lu & Tower 1997). De plus, l’insertion d’un insulateur entre l’élément régulateur ACE3 et l’origine ori- β supprime l’amplification (Lu *et al.* 2001). Cette rupture de communication est comparable au cas d’un insulateur inséré entre un activateur transcriptionnel et un promoteur. Ces données suggèrent que les insulateurs pourraient interagir avec les mécanismes de contrôle de la réplication de la même façon qu’avec les mécanismes de contrôle de la transcription. Il a été proposé que les insulateurs liés par CTCF dans le locus β -globin du poulet pourraient limiter l’action des éléments réplicateurs contenus dans cette région (Prioleau *et al.* 2003) mais aucune étude n’a été conduite à ce jour pour vérifier cette hypothèse. Les insulateurs étant impliqués dans la formation de boucles de chromatine et le regroupement de séquences au sein de foyers, il serait intéressant d’étudier si ces éléments pourraient être impliqués dans la formation des foyers de réplication qui sont observés (pour revue : Natsume & Tanaka 2010).

Il apparaît que BEAF reste lié à la chromatine pendant la mitose (Hart *et al.* 1999) et il a été proposé que ce facteur pourrait avoir un rôle dans la maintenance de l'information épigénétique (Gilbert *et al.* 2006). Il est en effet possible que l'organisation de la chromatine par les insulateurs participe à la mémoire cellulaire. Cette possibilité a été étudiée plus précisément chez les vertébrés (pour revue : Phillips & Corces 2009). CTCF reste lié sur environ la moitié de ces sites pendant la mitose (Burke *et al.* 2005). Cette liaison est indispensable pour la maintenance de l'état hypométhylé de certains îlots CpG comme, par exemple, ceux de l'allèle maternelle du locus *Igf2* (Fedoriw *et al.* 2004, voir §1.4.4 et Fig. 3.6). Comme il n'y a pas de données indiquant si la méthylation de l'ADN chez la drosophile pourrait avoir un lien avec les insulateurs, ces observations ne sont pas transposables. Cependant, de façon intéressante, la boucle de chromatine formée entre l'ICR et l'élément DMR1 du locus *Igf2* reste intacte pendant la mitose (Burke *et al.* 2005). Ainsi, le maintien de boucles de chromatine par les insulateurs pourrait participer à la mémoire cellulaire. Cette possibilité mériterait d'être étudiée chez les différents organismes.

3.4 Formation de boucles de chromatine

3.4.1 Contraintes mécaniques

Quelles que soient les discussions à propos de la structure de la fibre de 30 nm, son existence *in vivo* dans les structures de compaction d'ordres supérieurs (en interphase ou dans les chromosomes mitotiques) n'est généralement pas mise en doute. Il y a pourtant des études plus ou moins récentes qui suggèrent qu'il n'y a pas de fibre de 30 nm dans les chromosomes mitotiques et que cette structure est rare en interphase (pour revue : Maeshima *et al.* 2010). En 1988, Dubochet *et al.* avaient déjà suggéré, d'après des observations de sections de cellules mitotiques par microscopie CEMOVIS (Cryo-Electron Microscopy of Vitreous Sections), que le chromosome était constitué d'un agrégat compact liquide de fibres de 11 nm (Dubochet *et al.* 1988) et non de fibres de 30 nm. Pendant longtemps, l'impossibilité d'observer la fibre de 30 nm a été attribuée à un artefact dû à la technique utilisée. Récemment la technique a été améliorée et le groupe de Dubochet a confirmé ces anciennes observations (Eltsov *et al.* 2008). Le spectre obtenu par transformée de Fourier spatiale (pour faire ressortir les tailles caractéristiques des motifs réguliers) des images de chromosomes montre bien un pic à 11 nm correspondant à la structure en « collier de perles » mais pas de pic à 30 nm. Ces auteurs proposent qu'à partir d'une certaine densité de nucléosomes, la fibre de 30 nm, observée *in vitro* à de faibles concentrations, fond pour former un agrégat liquide où les interactions inter-fibres dominent les interactions entre nucléosomes proches sur une même fibre.

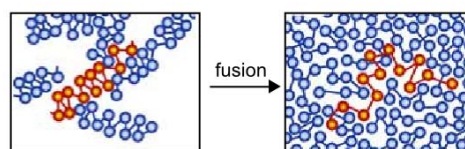


Figure 3.5 – Modèle de fusion de la fibre de 30 nm (adapté de Maeshima *et al.* 2010).

En ce qui concerne l'interphase, la fibre de 30 nm est également difficilement observable et ne semble pas être la structure dominante (à l'exception de certains types cellulaires comme les érythrocytes chez le poulet (Woodcock 1994)). Il n'est cependant pas exclu qu'une partie du génome soit sous forme de fibre de 30 nm et qu'il y ait des transitions dynamiques entre domaines organisés et domaines fondus.

Le modèle d'un polymère en fusion implique une grande dynamique des fibres de chromatine. Comme le repliement des fibres détermine l'accessibilité à l'ADN, ces mouvements rapides permanents pourraient permettre l'accès des facteurs de transcription et la formation de boucles transitoires facilitant l'interaction entre un promoteur et les séquences activatrices. La multiplication des observations de ce type de boucles semble en faveur de ce modèle où la chromatine est nettement plus dynamique que dans une fibre structurée de 30 nm.

3.4.2 Implication du complexe cohesin

La grande dynamique des fibres chromatiniennes permet à un insulateur ou un activateur de trouver ses cibles vraisemblablement par collisions aléatoires jusqu'à ce qu'une interaction spécifique ait lieu. Ces interactions peuvent être de forces variables et doivent être stabilisées par des co-facteurs pour perdurer. La recherche de ces co-facteurs est importante pour comprendre la dynamique de l'organisation tri-dimensionnelle de la chromatine.

Récemment, un lien important a été découvert entre les insulateurs et le complexe cohesin (pour revue : [Wood *et al.* 2010](#)). La cohesin a été initialement étudiée dans le cadre de la mitose où elle maintient les chromatides soeurs ensemble. Plusieurs études à l'échelle du génome ont montré que l'insulateur CTCF et le complexe cohesin ont un grand nombre de sites de liaison communs chez l'humain. En effet, 65% à 90% des sites de liaison de la cohesin contiennent CTCF et à l'inverse 55% à 80% des sites de CTCF lient aussi la cohesin ([Parelho *et al.* 2008](#); [Wendt *et al.* 2008](#)). Le recrutement de la cohesin est dépendant de CTCF car sa déplétion entraîne la disparition de la cohesin sur les sites partagés. De plus, par des expériences de 3C, plusieurs groupes ont montré que le complexe cohesin est impliqué dans la formation des boucles de chromatine par CTCF à différents loci ([Hadjur *et al.* 2009](#); [Hou *et al.* 2010](#); [Mishiro *et al.* 2009](#); [Nativio *et al.* 2009](#)). Ainsi la cohesin semble stabiliser des interactions entre différents sites chromatiniens en plus de son rôle dans la cohésion des chromatides soeurs. Sur le locus *Igf2* décrit précédemment (voir §1.4.4), il a été montré que la cohesin lie la région ICR comme CTCF et que sa déplétion déstabilise la structure en boucle nécessaire à la répression du locus maternel qui devient exprimé dans ces conditions ([Nativio *et al.* 2009](#), Fig. 3.6).

Chez la drosophile, la cohesin se retrouve plutôt dans les régions transcriptionnellement actives et colocalise avec la polymérase II ([Misulovin *et al.* 2008](#)). Il n'y a pas de nette colocalisation de CTCF avec la cohesin. En revanche, 80% des sites de liaison de la stromalin (une sous-unité du complexe cohesin) se superposent avec CP190 ([Bartkuhn *et al.* 2009](#)), ce qui suggère tout de même que la cohesin pourrait avoir un rôle dans la fonction insulatrice. L'étude d'une éventuelle activité de la cohesin dans la formation ou la stabilisation de boucles formées par BEAF serait intéressante pour mieux comprendre comment ces structures sont régulées.

3.5 Régulation de la fonction insulatrice

Les insulateurs sont importants pour maintenir le profil d'expression des gènes. Il est probable qu'ils aient un rôle dans la définition des types cellulaires. Il apparaît donc important que la fonction insulatrice puisse être régulée.

3.5.1 Régulation du recrutement des facteurs insulateurs par d'autres protéines

Bien que la majorité des sites de liaison des protéines insulatrices soit conservée entre les différents types cellulaires étudiés jusqu'à présent, des différences de localisation ont

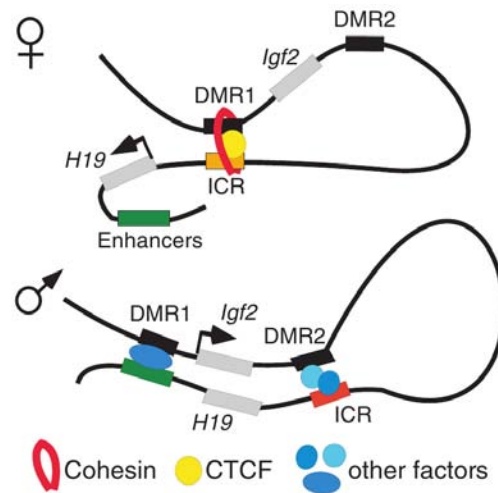


Figure 3.6 – Modèle simplifié représentant le locus soumis à l’empreinte parentale *Igf2-H19* (Ong & Corces 2009). L’expression monoallélique de ce locus est régulée par la liaison de CTCF et de la cohesin sur la région ICR. Sur l’allèle maternel, CTCF permet l’interaction entre l’ICR et la région DMR1 (DNA methylated region 1) qui nécessite aussi le complexe cohesin. La formation de cette structure isole le gène *Igf2* de l’influence des activateurs. Sur l’allèle paternel, l’ICR est méthylé, ce qui empêche la liaison de CTCF et de la cohesin, et permet aux activateurs d’agir sur *Igf2* qui est alors exprimé.

été observées. Une étude menée sur des cellules Kc et Mbn2 montre que par exemple que 18% des sites de Su(Hw) sont spécifiques aux cellules Kc et 5% aux cellules Mbn2 (Bushey *et al.* 2009). Il en est de même pour CTCF (18% et 37% respectivement), BEAF (11% dans les deux lignées) et CP190 (17% et 14% respectivement). Cela suggère que le recrutement des protéines insulatrices sur l’ADN, ainsi que celui de CP190, sont contrôlés et permettent la régulation de l’activité des éléments d’ADN insulateurs.

Le facteur DREF présenté au §1.4.2.2 est un exemple de protéine impliquée dans la régulation de la fonction de BEAF. En effet, les motifs consensus de liaison de BEAF (CGATA) et de DREF (TATCGATA) se chevauchent et nous avons montré que ces facteurs peuvent entrer en compétition pour la liaison de certains promoteurs et réguler différemment les gènes associés (voir §2.3).

Dans le cas des protéines CTCF et GAF, il a été montré que des facteurs de remodelage de la chromatine pouvaient moduler les propriétés insulatrices des séquences associées. En effet, la déplétion de protéines du complexe NURF supprime la fonction de bloqueur d’activation de *Fab7* et *SF1* qui sont reconnus par GAF, et augmente cette fonction pour l’insulateur *Fab8* qui lui est reconnu par CTCF (Li *et al.* 2010). Ces réponses opposées sont intéressantes car elles suggèrent des mécanismes d’action différents des insulateurs lié par GAF et CTCF.

3.5.2 Modifications post-traductionnelles des protéines insulatrices

La régulation des insulateurs peut également se faire par des modifications post-traductionnelles des protéines insulatrices. Par exemple, la protéine ubiquitine-ligase E3 Topors est nécessaire à la fonction insulatrice de *gypsy* (Capelson & Corces 2005). Il n’est cependant pas déterminé si la fonction ubiquitine-ligase est réellement impliquée dans l’insulation par *gypsy*. Par contre, Topors empêche la sumoylation de CP190 et mod(mdg4)2.2 qui, elle, inhibe la fonction insulatrice de *gypsy* (Capelson & Corces 2006). Ainsi, des modifications post-traductionnelles peuvent effectivement moduler l’activité des insulateurs. En ce qui concerne BEAF, plusieurs modifications ont été observées (phosphorylation Hart *et al.* (1997) ; O-glycosylation Pathak *et al.* (2007)) mais leur implication dans la fonction

de la protéine reste à élucider.

3.5.3 Implication d'ARN

En plus de diverses protéines, des ARN semblent également impliqués dans la régulation de la fonction insulatrice. En effet, des mutations des composants *piwi* et *aubergine* entraînent une diminution de l'activité insulatrice de *gypsy* (Lei & Corces 2006). Cet effet suggère l'intervention de petits ARN dans le mécanisme responsable de l'insulation. De plus, l'hélicase à ARN Rm62 interagit avec CP190 de façon dépendante d'un ou plusieurs ARN (perte de l'interaction après traitement à la RNase) et son invalidation augmente l'efficacité de l'insulateur. Ainsi des ARN participent à la fonction insulatrice de *gypsy* et Rm62 régule négativement leur effet. Comme CP190 interagit avec d'autres protéines insulatrices, il est possible que ce mécanisme puisse également réguler les autres types d'insulateurs. Cette hypothèse n'a cependant pas encore été étudiée.

3.6 Les insulateurs chez les vertébrés

De façon assez paradoxale, au moins quatre types d'insulateurs ont été décrits chez la drosophile, alors que les vertébrés semblent n'utiliser qu'un seul type d'insulateur, à savoir CTCF. Cette absence de corrélation entre la taille du génome et la complexité des insulateurs est une donnée étonnante. Une possibilité est qu'il reste des insulateurs à découvrir chez les vertébrés. Une étude recherchant les orthologues des protéines insulatrices de la drosophile chez différentes espèces à partir de la séquence de leur génome montre qu'il ne semble pas en exister en dehors des insectes (Schoborg & Labrador 2010). Cependant, une étude très récente décrit la découverte d'un orthologue de GAF, c-Krox/Th-POK (Matharu *et al.* 2010). c-Krox/Th-POK est encore très peu caractérisé, mais il a été montré que ce facteur lie également les séquences riches en motifs GAGA à la fois *in vitro* et sur quelques sites testés *in vivo*. Cette découverte contredit l'étude précédente (Schoborg & Labrador 2010) et conforte l'idée que des orthologues de BEAF ou de Su(Hw) pourraient exister. Il pourrait s'agir d'orthologues structuraux codés par des séquences trop différentes pour être identifiées facilement.

Les études sur la protéine CTCF des mammifères sont nombreuses et montrent que ce facteur régule l'expression d'oncogènes et de gènes suppresseurs de tumeurs (dont *MYC*, *Rb*, *ARF*, *hTERT* et *Igf2* ; pour revue : Filippova 2008). Plusieurs mutations de CTCF sont d'ailleurs retrouvées dans différents types de tumeurs (Filippova *et al.* 2002). Ces mutations altèrent la spécificité de liaison de CTCF et certains de ces gènes cibles suppresseurs de tumeurs ne sont plus activés. De la même façon, nous avons montré que BEAF est impliqué dans la régulation de multiples gènes associés au contrôle de la prolifération cellulaire. De plus, BEAF entre en compétition avec l'activateur DREF qui régule plusieurs gènes suppresseurs de tumeurs dont *p53*, *myb* ou *raf* (pour revue : Matsukage *et al.* 2008). Ainsi, les protéines insulatrices semblent importantes pour le contrôle de la prolifération cellulaire et des dysfonctionnements de ces protéines pourraient être impliqués dans le développement de cancers.



Conclusion

Ce travail de thèse s'est focalisé sur la protéine BEAF, sur son rôle dans la compartimentation du génome ainsi que dans la régulation transcriptionnelle.

Nous avons pu montrer que la liaison de BEAF au niveau de l'élément BE28 situé dans la région de transition entre l'euchromatine et l'hétérochromatine du chromosome X ne suffit pas à former une frontière entre ces deux types de chromatine. Nous avons vu que la déplétion de BEAF entraîne des dérégulations variables des gènes centromères distaux du BE28 et nous proposons que plusieurs insulateurs pourraient interagir entre eux pour organiser tri-dimensionnellement cette région. Des études supplémentaires sont nécessaires pour comprendre comment sont organisées ces régions de transition entre euchromatine et hétérochromatine et en quoi les insulateurs sont impliqués dans cette organisation.

Nous avons également abordé la question du rôle de BEAF dans la régulation des gènes euchromatiniens. Nous avons montré que BEAF se lie sur les promoteurs de centaines de gènes, et qu'il régule leur expression positivement ou négativement suivant les loci. Cette régulation fait intervenir, pour les gènes activés par BEAF, une restriction de la déposition de la marque H3K9me3, et pour les gènes réprimés par BEAF, une compétition avec DREF.

Enfin, nous avons étudié le rôle de BEAF dans le positionnement des nucléosomes et nous avons montré que BEAF positionne précisément les nucléosomes +1 et +2 en aval du TSS des gènes. Ce positionnement est important pour la régulation de la transcription de ces gènes et est corrélé avec une stabilisation de l'étape de pause de l'ARN polymérase II.

Ces différentes études amènent une meilleure compréhension de la fonction de BEAF *in vivo* et suggèrent que ce facteur est un régulateur clé impliqué dans le contrôle de la prolifération cellulaire. Cette propriété est conservée dans le cas de l'insulateur CTCF des vertébrés et il serait intéressant de voir dans quelle mesure les mécanismes impliqués dans la régulation de la transcription, notamment le positionnement des nucléosomes et la stabilisation de l'étape de pause, que nous avons observés chez la drosophile pourraient être conservés chez les vertébrés.



Nomenclature



Sommaire

A.1 Code IUPAC	139
A.2 Nomenclature des acides aminés	140

A.1 Code IUPAC

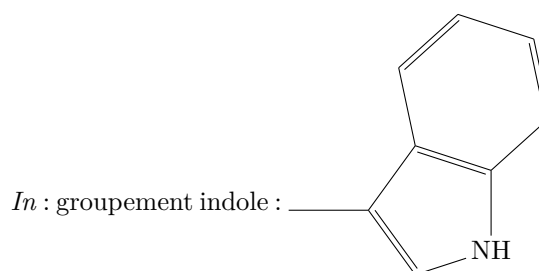
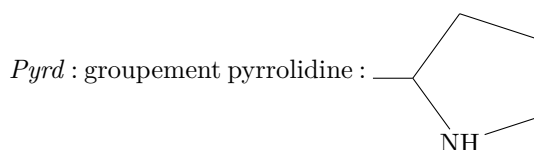
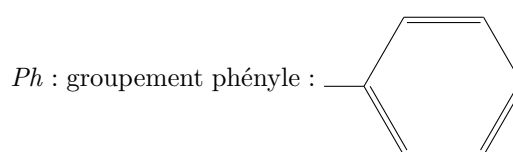
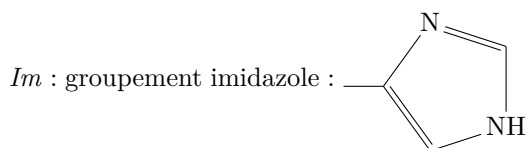
Symboles des nucléotides notation IUPAC		
A	A	A denine
C	C	C ytosine
G	G	G uanine
T	T	T hymine
U	U	U racile
R	A ou G	pu R ine
Y	C ou T (U)	p Y rimidine
M	A ou C	a M ino
K	G ou T (U)	K eto
S	C ou G	S trong (3 liaisons H)
W	A ou T (U)	W weak (2 liaisons H)
B	C ou G ou T (U)	pas A
D	A ou G ou T (U)	pas C
H	A ou C ou T (U)	pas G
V	A ou C ou G	pas T (U)
N	A ou C ou G ou T (U)	a N y

Table A.1 – Code IUPAC de nomenclature des nucléotides.

A.2 Nomenclature des acides aminés

Acide aminé	code 3 lettres	code 1 lettre	MW (Da)	Structure
Alanine	Ala	A	89,1	$\text{CH}_3 - \text{CH}(\text{NH}_2) - \text{COOH}$
Arginine	Arg	R	174,2	$\text{HN}=\text{C}(\text{NH}_2) - \text{NH} - (\text{CH}_2)_3 - \text{CH}(\text{NH}_2) - \text{COOH}$
Asparagine	Asn	N	132,1	$\text{H}_2\text{N} - \text{CO} - \text{CH}_2 - \text{CH}(\text{NH}_2) - \text{COOH}$
Acide aspartique	Asp	D	133,1	$\text{HOOC} - \text{CH}_2 - \text{CH}(\text{NH}_2) - \text{COOH}$
Cystéine	Cys	C	121,2	$\text{HS} - \text{CH}_2 - \text{CH}(\text{NH}_2) - \text{COOH}$
Acide glutamique	Glu	E	147,1	$\text{HOOC} - (\text{CH}_2)_2 - \text{CH}(\text{NH}_2) - \text{COOH}$
Glutamine	Gln	Q	146,1	$\text{H}_2\text{N} - \text{CO} - (\text{CH}_2)_2 - \text{CH}(\text{NH}_2) - \text{COOH}$
Glycine	Gly	G	75,1	$\text{H}_2\text{N} - \text{CH}_2 - \text{COOH}$
Histidine	His	H	155,2	<i>Im</i> - $\text{CH}_2 - \text{CH}(\text{NH}_2) - \text{COOH}$
Isoleucine	Ile	I	131,2	$\text{CH}_3 - \text{CH}_2 - \text{CH}(\text{CH}_3) - \text{CH}(\text{NH}_2) - \text{COOH}$
Leucine	Leu	L	131,2	$(\text{CH}_3)_2\text{CH} - \text{CH}_2 - \text{CH}(\text{NH}_2) - \text{COOH}$
Lysine	Lys	K	146,2	$\text{H}_2\text{N} - (\text{CH}_2)_4 - \text{CH}(\text{NH}_2) - \text{COOH}$
Méthionine	Met	M	149,2	$\text{CH}_3 - \text{S} - (\text{CH}_2)_2 - \text{CH}(\text{NH}_2) - \text{COOH}$
Phénylalanine	Phe	F	165,2	<i>Ph</i> - $\text{CH}_2 - \text{CH}(\text{NH}_2) - \text{COOH}$
Proline	Pro	P	115,1	<i>Pyrd</i> - COOH
Sérine	Ser	S	105,1	$\text{HO} - \text{CH}_2 - \text{CH}(\text{NH}_2) - \text{COOH}$
Thréonine	Thr	T	119,1	$\text{CH}_3 - \text{CH}(\text{OH}) - \text{CH}(\text{NH}_2) - \text{COOH}$
Tryptophane	Trp	W	204,2	<i>In</i> - $\text{CH}_2 - \text{CH}(\text{NH}_2) - \text{COOH}$
Tyrosine	Tyr	Y	181,2	<i>pPhOH</i> - $\text{CH}_2 - \text{CH}(\text{NH}_2) - \text{COOH}$
Valine	Val	V	117,1	$(\text{CH}_3)_2\text{CH} - \text{CH}(\text{NH}_2) - \text{COOH}$

Table A.2 – Nomenclature des acides aminés.



Matériel et méthodes

B

Sommaire

B.1	Traitements des cellules	141
B.1.1	Culture cellulaire	141
B.1.2	Interférence ARN	141
B.1.3	FACS	142
B.2	Analyses de l'expression	142
B.2.1	PCR quantitative et RNA-seq	142
B.2.2	<i>Northern blot</i>	142
B.3	Préparation de noyaux, digestion à la MNase et analyses	143
B.3.1	Préparation de noyaux	143
B.3.2	Digestion à la MNase	143
B.3.3	<i>Southern blot</i>	143
B.4	Analyses des protéines	143
B.4.1	Extraits protéiques et <i>western blot</i>	143
B.4.2	Extraits nucléaires et immunoprécipitation	144
B.5	Immunoprécipitation de Chromatine (ChIP)	144
B.6	Séquences des oligonucléotides	146

B.1 Traitements des cellules

B.1.1 Culture cellulaire

Les cellules de drosophile S2 ou Kc sont cultivées en suspension entre 1×10^6 et 5×10^6 cellules/ml dans du milieu Schneider (GIBCO, Invitrogen) supplémenté avec 10% de sérum de veau fœtal (SVF, Sigma) et de 50 U/mL de pénicilline et 50 μ g/mL de streptomycine (GIBCO, Invitrogen).

B.1.2 Interférence ARN

Les ARNdb sont synthétisés à partir de régions d'environ 1Kb des ADNc des gènes cibles. L'amplification des régions d'intérêt est effectuée en utilisant des amorces permettant d'ajouter le promoteur de l'ARN polymérase T7 (voir séquences au §B.6). Les produits de PCR sont transcrits avec le kit TranscriptAid T7 High Yield (Fermentas). Les ARNdb sont ensuite purifiés par extraction au phénol/chloroforme puis précipitation à l'éthanol 100%. Ils sont ensuite solubilisés dans de l'eau à une concentration finale de 3 μ g/ μ L environ. Pour favoriser l'appariement des brins d'ARN complémentaires, les échantillons sont chauffés à 70°C pendant 15 min puis refroidis lentement en 30 min jusqu'à la température ambiante.

Pour le traitement par interférence à l'ARN, les cellules sont centrifugées à $300 \times g$ pendant 5 min, le milieu de culture est enlevé et elles sont reprises à une concentration finale de $3 \cdot 10^6$ cellules/ml dans du milieu sans sérum. 300 μ g d'ARNdb sont ajoutés dans 10 ml de cellules. Les cellules sont incubées 2h à 25°C. L'absence de sérum pendant cette incubation induit un stress suffisant pour permettre la pénétration des ARNdb. 20 ml de milieu avec sérum sont ajoutés pour

obtenir une concentration finale de 1.10^6 cellules/ml puis les cellules sont remises en condition de croissance pendant 4 à 5 jours à 25°C. Elles sont comptées tous les jours pour estimer leur croissance (Z1 Coulter, Beckman) et la déplétion des ARNm est vérifiée par RT-qPCR (voir §B.2.1) et par *western blot* (§B.4.1).

B.1.3 FACS

500 000 cellules sont collectées par 5 min de centrifugation à $400 \times g$. Elles sont rincées dans un grand volume de PBS (Sambrook & Russell 2001) puis reprises dans 400 μ L de PBS. Tout en vortexant doucement les cellules, 3 volumes d'éthanol absolu sont ajoutés goutte à goutte. Les cellules ainsi fixées peuvent être conservées à -20°C. Le jour de l'analyse, les cellules sont rincées deux fois au PBS puis incubées dans 500 μ L de PBS additionné de 1 mg/mL de RNase A (Sigma Aldrich) pendant 45 min à 37°C. 10 μ g/ μ L de iodure de propidium (Sigma Aldrich) sont ajoutés. Après une incubation d'au moins 30 min, les échantillons sont analysés dans un cytomètre FACSCalibur (BD).

B.2 Analyses de l'expression

B.2.1 PCR quantitative et RNA-seq

Pour l'analyse de l'expression des gènes, les ARN totaux de 5 millions de cellules sont extraits sur une colonne RNeasy (Qiagen) sur laquelle un traitement à la DNase (RNase-Free DNase Set) est effectué. Ils sont alors dosés par mesure de l'absorbance à l'aide d'un spectrophotomètre Nanovue (GE Healthcare, utilisé pour l'ensemble des dosages des acides nucléiques). Une partie de ces ARN a été analysée par séquençage à haut débit (plateforme MGX, Montpellier). Pour les analyses par PCR quantitative, 1 μ g d'ARN est rétro-transcrit en ADNc (Invitrogen SuperScript VILO cDNA synthesis) dans un volume de 20 μ L. Les échantillons d'ADNc sont traités à la RNase H (5U) pendant 30 min à 37°C. L'analyse est alors effectuée par PCR quantitative sur ces ADNc qui sont dosés par mesure de l'absorbance. Dans un volume total de 15 μ L, 30 ng d'ADNc, 200 nM de chaque amorce et 7,5 μ L de mix Biorad iQ SYBR Green 2X sont déposés dans chaque puits (plaque de 96 puits). Chaque mesure est effectuée en triplicata. La PCR quantitative est réalisée dans un thermocycleur Realplex d'Eppendorf. Le cycle de PCR est le suivant : 2 min à 95°C puis 40 cycles à deux étapes (10 s à 95°C puis 30 s à 60°C). Les amorces sont validées à l'aide d'une courbe standard obtenue à partir de dilutions d'ADNc qui donne l'efficacité de la PCR, et d'une courbe de fusion (de 60°C à 95°C en 20 minutes) qui permet de vérifier la pureté du produit de PCR. L'analyse des données brutes est effectuée à l'aide du logiciel pyQPCR¹

B.2.2 Northern blot

5 μ g d'ARN totaux extraits des cellules (voir ci-dessus) sont préalablement dilués dans 3 volumes de tampon de *northern blot* (MOPS 1X (Sambrook & Russell 2001) ; 10% formaldéhyde ; 50% formamide et 0,5 μ g/mL de BET) puis chauffés à 65°C 5 min. Ils sont déposés sur un gel contenant 1% d'agarose et 10% de formaldéhyde dans du MOPS 1X et mis à migrer sur la nuit (~ 20 Vxh/cm). Le gel est lavé deux fois pendant 15 minutes dans de l'eau pour enlever le formaldéhyde, puis incubé dans une solution de 10 mM NaOH et 3 M NaCl pendant 20 min pour faciliter le transfert des transcrits de grande taille. Les ARN sont alors transférés sur une membrane (GeneScreen Plus) par électro-transfert (tampon : 12 mM Tris ; 6 mM NaAc ; 0,3 mM EDTA). Les ARN sont fixés sur la membrane par exposition aux UV (Stratalinker UV Crosslinker de Stratagene, mode Autocrosslink). Pour détecter les précurseurs des ARN ribosomiques, 10 pmol de sonde oligonucléotidique (spécifique de l'ARN 5S ou de l'ITS2b du grand précurseur 38S des autres ARN ribosomiques) est radio-marquée en l'incubant pendant 45 min à 37°C en présence de 25 μ Ci de γ^{32} P-ATP et de 10U de Polynucléotide Kinase T4 dans le tampon fourni (Promega). La réaction est stoppée par l'ajout d'EDTA à la concentration finale de 30 mM puis l'échantillon est passé à travers une colonne MicroBiospin 6 (Biorad) pour éliminer le γ^{32} P-ATP non incorporé. La sonde est ensuite ajoutée sur la membrane préalablement pré-hybridée 1h (tampon : 6X SSPE ; 5X Denhart (Sambrook & Russell 2001) ; 0,5% SDS ; 50 μ g/mL d'ADN de sperme de saumon (soniqué et dénaturé, Roche)). L'hybridation est conduite sur la nuit à 40°C. Après rinçages de la membrane

1. logiciel libre disponible à l'adresse suivante : <http://pyqpcr.sourceforge.net/>

(4 fois 15 min à 40°C dans du tampon SSC 1X, SDS 0,1%), le signal est détecté grâce à des films Biomax MS (Kodak).

B.3 Préparation de noyaux, digestion à la MNase et analyses

B.3.1 Préparation de noyaux

Les cellules sont collectées par centrifugation (5 min à $400 \times g$) puis rincées 2 fois dans 15 mL de tampon d'isolation (3,75 mM Tris-HCl pH 7,4; 0,05 mM spermine (Sigma Aldrich); 0,125 mM spermidine (Sigma Aldrich); 0,5 mM EDTA/KOH pH 7,4; 20 mM KCl; 0,5% (v/v) thiodiglycol; 0,1% Trasylol (Bayer) et 0,1 mM de PMSF (Sigma Aldrich)) à température ambiante (centrifugation 5 min à $1800 \times g$). Le culot est resuspendu dans 15 mL de tampon d'isolation glacé supplémenté de 0,1% de digitonine (Sigma Aldrich) et de 1% de Trasylol puis les cellules sont lysées dans un homogénéisateur de type Dounce par 15 à 20 aller-retours avec un piston B (vérification de la lyse sous un microscope). Les noyaux sont rincés par au moins trois lavages successifs dans le même tampon (centrifugation 5 min à $1250 \times g$). Ils sont dosés par absorbance à 260 nm après dilution d'un aliquot dans du SDS 1% et brève sonication.

B.3.2 Digestion à la MNase

Les noyaux sont repris dans du tampon MNase (10 mM Tris-Cl pH 8; 50 mM NaCl; 0,01 mg/mL BSA, 3 mM CaCl_2 ; 1 mM MgCl_2 ; 0,1% digitonine et 0,1 mM PMSF) à une concentration de 10 unités A_{260} . Chaque digestion est conduite dans 350 μL . Des concentrations croissantes de MNase sont ajoutées (de 0,01 à 0,5 U/mL; Sigma Aldrich). La digestion est effectuée pendant 10 min à température ambiante puis arrêtée par ajout d'un volume de solution STOP contenant de la RNase A (50 $\mu\text{g}/\text{mL}$) pour dégrader les ARN (100 mM Tris-Cl pH 7,4; 0,5% SDS; 62,5 mM EDTA; 12,5 mM EGTA). Les échantillons sont incubés 1h à 37°C puis de la protéinase K est ajoutée à 0,4 mg/mL et l'incubation est poursuivie sur la nuit. L'ADN est alors purifié par extraction au phénol/chloroforme/alcool isoamylique (25/24/1) puis précipitation à l'éthanol. Il est ensuite solubilisé dans de l'eau. Pour l'analyse MNase-seq, les fragments d'ADN correspondants aux mononucléosomes ont été purifiés sur gel d'agarose puis analysés par séquençage à haut débit (collaboration avec Keji Zhao, NIH).

B.3.3 Southern blot

2 μg d'ADN sont digérés par 20 U d'enzymes de restriction PstI et NdeI (New England Biolabs) pendant 2h à 37°C. L'ADN est à nouveau purifié et solubilisé dans de l'eau puis séparé sur un gel de 2% d'agarose, TBE 0,5X (Sambrook & Russell 2001). L'ADN est alors transféré sur une membrane de nylon (GeneScreen Plus, Perkin Elmer) par électro-transfert (tampon : 12 mM Tris; 6 mM NaAc; 0,3 mM EDTA). La sonde de ~ 200 pb est obtenue par PCR sur de l'ADN génomique puis purification sur gel d'agarose (Qiaquick gel extraction kit, Qiagen). 25 ng de sonde sont marquées avec le kit High Prime DNA labeling (Roche) et 50 μCi de $\alpha^{32}\text{P}$ dCTP (Perkin Elmer). L'ADN est ponté sur la membrane par exposition aux UV (Stratalinker UV Crosslinker de Stratagene, mode Autocrosslink) puis celle-ci est préhybridée dans du tampon de composition 6X SSPE, 5X Denhart, 0,5% SDS et 1 $\mu\text{g}/\text{mL}$ d'ADN de sperme de saumon (soniqué et dénaturé, Roche) pendant 1 heure à 60°C. La sonde est dénaturée 5 min à 100°C puis ajoutée sur la membrane. L'hybridation est conduite à 60°C sur la nuit. Après rinçages de la membrane (2 fois 15 min à 60°C dans du tampon SSPE 2X, SDS 0,1% puis 2 fois 15 min à 60°C dans du tampon SSPE 1X, SDS 0,1%), le signal est détecté grâce à des films Biomax MS (Kodak). Après photographie du film, un densitogramme est effectué grâce au logiciel ImageJ en sélectionnant les différentes pistes du gel.

B.4 Analyses des protéines

B.4.1 Extraits protéiques et western blot

Les noyaux (l'équivalent de 10 μL de noyaux à 20 unités A_{260} par piste; voir ci-dessus) sont centrifugés 5 min à $600 \times g$ puis repris dans 20 μL de tampon de digestion (15 mM Tris-Cl pH

7,4; 0,05 mM spermine; 0,125 mM spermidine; 0,5 mM EDTA/KOH pH 7,4; 20 mM KCl; 10 mM MgCl₂; 70 mM NaCl; 1% Trasylol; 0,1% digitonine et 0,1 mM PMSF). 1 µL de DNase I est ajouté et les échantillons sont incubés 30 min à 37°C.

Après ajout du tampon de charge (NuPage 4X, Invitrogen) et de DTT à la concentration finale de 100 mM, les échantillons sont soniqués, chauffés à 95°C 5 min puis chargés sur gel SDS-PAGE 4-12% d'acrylamide (Invitrogen) pour une migration à 90V. Les protéines du gel sont ensuite transférées passivement par diffusion sur 2 membranes de nitrocellulose Hybond-C Extra (Amersham) placées de part et d'autre du gel sur la nuit à 55°C dans du tampon de transfert (50 mM NaCl; 1 mM EDTA; 10mM Tris-Cl pH 8,0; 0,1 mM DTT; Cuvier *et al.* 2002). Les membranes sont bloquées 1h dans du PBS additionné de 0,05% de Tween-20 et de 5% de lait écrémé en poudre puis incubées sur la nuit à 4°C avec l'anticorps primaire (anti-BEAF (produit à façon par Eurogentec; 1/10 000), anti-NELF (don de D. S. Gilmour; 1/1000) et anti-actine (Chemicon; 1/20 000)) dilué dans du PBS-0,05% Tween-20. Après 5 rinçages de 5 min dans du PBS-0,05% Tween-20, les membranes sont incubées 1h dans une solution d'anticorps secondaire anti-lapin ou souris (suivant l'anticorps primaire) couplé à la horseradish-peroxydase dilué au 1/10 000^{ème} (GE Healthcare). Les membranes sont à nouveau rincées 5 fois 5 min dans du PBS-0,05% Tween-20 et les signaux sont révélés avec le kit de détection ECL Plus (GE Healthcare). La chemiluminescence est mesurée par un analyseur Luminescent Image LAS-400 de FUJIFILM.

B.4.2 Extraits nucléaires et immunoprécipitation

Des noyaux préparés comme décrit précédemment à partir de 250 millions de cellules sont repris dans 650 µL de NEB20 (20 mM KCl; 10 mM Hepes/KOH pH 7,6; 3 mM MgCl₂; 0,1 mM EDTA; 1 mM DTT; 0,1% Trasylol; 0,2 mM PMSF; 10,0% glycerol; cocktail d'inhibiteurs de protéases (Complete, Roche)). Un volume de NEB700 est ajouté (idem sauf 700 mM de KCl) puis les échantillons sont incubés 30 min à 4°C. Ils sont alors centrifugés à 125 000 × *g* 1h à 4°C. Le surnageant est récupéré et 0,4 volume de NEB20 y est ajouté pour réduire la concentration en KCl à 250 mM. Les extraits sont alors aliquotés et congelés par immersion dans de l'azote liquide, puis conservés à -80°C. La qualité des extraits est vérifiée après migration sur gel d'acrylamide et coloration au bleu de Coomassie. Leur concentration est estimée par comparaison avec des dilutions d'extraits commerciaux (Active motif) de concentration connue.

Pour les immunoprécipitations, 20 µL de billes d'agarose couplées à la protéine A (Invitrogen) sont rincées puis couplées à 25 µg d'anticorps BEAF (Eurogentec) ou d'IgG (Tebu-Bio) dans 10 volumes de PBS additionné de 0,05% de NP40 pendant 2h à température ambiante. Les billes sont rincées deux fois par 500 µL de borate de sodium 0,2 M (pH 9, adapté par ajout de NaOH). Les anticorps sont crosslinkés sur les billes par une incubation de 30 min dans 500 µL de borate de sodium 0,2 M pH 9.0 additionné de 20 mM de diméthylpipéridimidate (Sigma Aldrich). Les billes sont rincées une fois par 500 µL d'éthanolamine 0.2M (pH 8.0; adapté par ajout de HCl) puis incubées 2h dans cette solution. Elles sont alors rincées cinq fois dans du PBS/0,05% NP40 et deux fois dans du tampon d'IP (PBS; 0,1% NP40; cocktail d'inhibiteurs de protéases). ~50 µg d'extraits nucléaires sont dilués dans 10 volumes de tampon d'IP et incubés sur les billes couplées aux anticorps pendant 2h à 4°C. 20% du surnageant est conservé (FT). Les billes sont lavées 4 fois dans le tampon d'IP additionné de 10% de glycérol puis deux fois dans ce tampon additionné de 0,2 M de KCl et à nouveau deux fois dans le tampon sans KCl. Les protéines sont éluées dans 40 µL de tampon d'IP additionné de 0,5% de SDS par 10 min d'incubation à température ambiante. Elles sont alors analysées par *western blot* (voir ci-dessus).

B.5 Immunoprécipitation de Chromatine (ChIP)

Le pontage des protéines et de l'ADN est effectué sur cellules fraîches (rincées 2 fois dans du PBS froid avec inhibiteurs de protéases) dans du PBS additionné de 0,08% de formaldéhyde pendant 10 min, puis la réaction est arrêtée en ajoutant de la glycine (125mM final) et en incubant 5 min à température ambiante. Les cellules sont rincées 2 fois dans du PBS. Des noyaux sont alors préparés (voir protocole de préparation de noyaux). Les noyaux sont perméabilisés pendant 20 min à température ambiante en présence de 0,2% de Triton X-100. Ils sont alors rincés dans du tampon de lyse (KCl 60 mM; NaCl 15 mM; MgCl₂ 4 mM; HEPES pH 7,6 15 mM; Triton 0,5%; DTT 0,5 mM; inhibiteurs de protéases) puis incubés dans 100 µl de tampon de lyse supplémenté de 1% de SDS et de 0,5% de lauroylsarcosyl pendant 30 min à 4°C. La chromatine est soniquée par 20 cycles à haute puissance de 30 secondes séparés par 30 s de repos dans un Biorupteur (Diagenode). La

sonication est vérifiée sur l'équivalent de 200 000 cellules. Pour cela, les échantillons sont digérés à la RNase A (1µg/µL) puis à la protéinase K (1µg/µL) 3h à 55°C. L'ADN est alors purifié par extraction des protéines au phénol/chloroforme/alcool isoamylique et analysé sur gel d'agarose 1%. Si la taille des ADN soniqués est encore trop grande, quelques cycles supplémentaire de sonication sont effectués pour obtenir des fragments d'une taille moyenne de 300 à 400 pb. Les billes d'agarose couplées à la protéine G (Genscript) sont bloquées durant 2h à 4°C sur roue dans le tampon de ChIP (tampon de lyse supplémenté de 0,1% de SDS et de 0,5% de lauroylsarcosyl) additionné de 0,1 mg/ml de BSA, puis rincées deux fois dans du tampon de ChIP. La chromatine est diluée 10 fois puis incubée sur la nuit à 4°C avec des billes bloquées à la BSA pour supprimer les complexes ADN-protéines ayant une forte affinité pour les billes. D'autres billes (25 µl) sont saturées en anticorps par une incubation sur la nuit à 4°C avec 20 µg d'anticorps (pol II, Covance; IgG, Tebu-Bio) dans du tampon de ChIP (volume total : 200 µl). Elles sont ensuite rincées 4 fois dans 500 µL de tampon de ChIP. L'équivalent en chromatine (pré-incubée sur les billes) de 2 millions de cellules est ajouté à 10 µL de billes couplées aux anticorps. 10% de la chromatine est prélevée pour constituer l'« input ». La chromatine est incubée sur les billes 4 à 6h sur roue à 4°C. Ensuite, les billes sont rincées 4 fois avec du tampon de ChIP, puis deux fois avec du TE (Tris-HCl 10 mM pH 8; EDTA 0,1 mM) sans inhibiteur de protéases. La chromatine est alors éluée par deux incubations de 20 min à 70°C sous agitation, l'une dans 100 µl de Tris-HCl 50 mM pH 8/EDTA 10 mM/SDS 1%, la suivante dans 150 µl de TE additionné de 0,67% de SDS. Les 2 éluats sont mélangés. La réaction de réversion des pontages est réalisée sur la nuit à 65°C. 250 µL de solution de protéinase K (140 µg/mL de glycogène et 400 µg/mL de protéinase K dans du TE) sont ajoutés pour une digestion de 2h à 50°C. Enfin, l'ADN est purifié en présence de 0,4 M de LiCl par extraction au phénol/chloroforme/alcool isoamylique. L'ADN est précipité par ajout de 1,3 ml d'éthanol absolu froid et incubation pendant au moins deux heures à -20°C. Après 30 min de centrifugation à 16 000 × *g*, le culot est rincé à l'éthanol 70%, puis séché avant d'être repris dans 15 µL d'eau. L'ADN est ensuite dilué pour être quantifié par qPCR (séquences des amorces en B.6). L'échantillon d'« input » est utilisé pour réaliser une gamme standard par dilution et permettre une quantification précise de l'ADN précipité par la ChIP grâce au logiciel pyQPCR.

B.6 Séquences des oligonucléotides

	Forward	Reverse
Expression		
Beaf32	GATGAGTAGGCGGTCCATTA	GATTACGACACGCTGATTTGC
Actine 5C	TTCACCACGACCGCTGAGCG	GGCCATCTCCTGGTCAAAGTC
CG12842	GTCCCTGAGTGAGGCCAATA	CACATTCAAAGGCGTCGTTTC
tsp39D	CTGTGGCAGTCATTTGCTTC	CGGTATGCTTACATAGCCA
CG5568	AGATTAAGCGTGGCTTGGTT	AGGTTGACCACTCACTTGAC
CG32115	CGAATCGTTAGGGACTGGGA	TCCGGTCAAAGGTTTATGGTT
CG14561	TGCCCTACACGAGATAAACG	CTTGGGCGTAAATCGGGAAT
Ror	GACGGAAGAGAACCAAGTCC	GTGCTGAACACATCCTCGTA
CG4250	GCACACATCGGTCCGAATAA	TGGAGCTGAGGATAGACCTG
Dip2-CG9771	TTTCCCAAACAACCGCAAAG	GGAGCACATAAAAAAGCTGGC
CG4210	ACTGGAATGGAATCCTGCAC	AGTAGTGCCAGCCCTCTTTA
CG12880	GCCCAAGTCACCCACATATT	TTCTGGTGCTCCTTGATGAC
su(var)205 (HP1)	CGAAAGCGAAAAAGAACGAT	TCCGATGCCTTAAGAGTTGG
Dref	GGCTCAAGATGGGACACATT	TGAATCGCTTGTGAAGTCC
CG13865	GTCCGGACGAAGCATAACAT	CACCCGAATCGATATTCCAC
white	TCCAGCAGGATGACCTCTTT	GCCGATAGGTGATGATCCGT
CG32820-C3	CCTGGAGATACCACTGCACA	CACGATGTCATGCAACCTTCT
DIP1-CG17686	CGGTGAGGATGAGGCCGATGCACCG	TCCGCGAACAGCTTGTAGATG
notch	CGGCCATCTCGGACTAGACCTGCCG	CCACATCGTCTCGGACAGAATC
kirre	CGTACATGCCGAGGCAACGCTGTACG	CTTAACGCGACCAACCACTCC
CG14476	CGGACACGACTCTAAGCCCTACGGTCCG	CTGTCCGGCGTGTCTGGAATAC
CG32857-C2	CGCATGTGGTCACTCAGTGGGATGCC	GATGAACATGCTCCTCCTGCAC
CG14621	CTCCGTCACTCGCACATCTC	CGGTGCTGGCATTGTGGCTTTGACCG
CG14614	CGCTCTGTTTGCCTCGTGGGAGCG	ATGGTTCGAGTGTCCAAGTGC
fog	CGGTTCATGGGTGGCGTCTGGACCG	GCAACAACAGCGACTGATCCTC
stoned	ATCAGCATAATGGACGGATTTGG	CGGCCACCTCCCAAGCATCGCCG
FucTC	CGCAACCTACAGCTCCCGTTTTGCG	TGGTGTAGAACTGTCTTTGTG
ti1B	TCCAGGTAATTGAACACATCCAGA	CGCGGTGCAAGTTCTCCAGGCGCG
Rab21	CGGTCTGTGCTGCGCTACATGGAAGACCG	CATCCTCCAGGGACATCTTGC
Chip		
betaTub56D-50	ACTGTCCGCTTACACACTTGG	CCACACTGCGGCCATCGTAT
betaTub56D-1000	GAGAAGCCGCGGAAGTAGCG	GTGTTTCGTAGGCTGCATTAGCTGT
spin-50	TCCGCCGACGGTCCACACTTA	TGCATTTGCAGTGTCTCTCGCT
spin-1000	GCTTTGCACCGAAACTCGGGA	ACCGCCGTGCTCACTCACTC
vha100-2-50	CCGCGAGTGGGAGCGATAAAT	GACGTGTGCTGCGCTCCATT
vha100-2-1000	TTGCCGACCACTTCCCGCAT	TAAAGCCAAAGCCCGGCCAC
CG9988-50	ATCGTTGGCTGCTAAACAGT	TGGGTTTTGATGCCGAGAAT
CG13766-50	AAAGTTTCTGCAATCGCCAC	GCCGATCAAAGTGACAGTTC
sara-50	TGCAACACTGTTTCCATGGCT	TGAGGAAAACAAGACGTATTCCG
sara-1000	CACTGCGTCCGGCACAGAAT	GGGCGACGATTCTCCTCCGAGA
T7		
BEAF32	GGAATTCTAATACGACTCACTATAGGGAGA-	GGAATTCTAATACGACTCACTATAGGGAGA-
DREF	GACAAAACCGTGCAGGATTACC	TGCGAACGGTTCCTACTCATC
HP1	CGGGTCAAACCTGGAGTTTCAT	CACGGTGGCATAACAGCATAAC
CDK7	CGCAAGGATGAGGAGAAGTC	TTTTTCGCTTTTCGATGATCC
	GACACGGTGACCAACCAAAAT	TCCCAGTGTGGAAAAGATCC
Sonde Southern blot		
tsp39D	GGATGGTGTCCAGGTGCTGGG	AACATATGAATCAATGTCCGCC
Sonde Northern blot		
ITS2b-38S	GGTTATGTTATTATTCTTCGTTGGTTCCG	
5S	CTAACCCGCGCCGACGCTGCTTAATTTCCG	

Table B.1 – Séquences des oligonucléotides utilisés dans cette étude.

Liste des figures

1.1	Schéma d'hybridation des pois par Mendel	15
1.2	La mutation <i>white</i> et son hérédité liée au sexe	15
1.3	Représentation des chromosomes de différentes espèces de drosophile	16
1.4	Schéma original du chromosome X par Painter	16
1.5	Photo de chromosomes polytènes	16
1.6	Expérience de Griffith mettant en évidence l'existence d'un « principe transformant »	17
1.7	Représentation schématique de la double hélice d'ADN	17
1.8	Structure chimique de l'ADN	18
1.9	Structure cristallographique du nucléosome	19
1.10	Les différents niveaux de compaction de la chromatine	21
1.11	Différents modèles de la fibre de 30 nm	21
1.12	Représentation des territoires chromosomiques d'un fibroblaste humain	23
1.13	Les principales modifications post-traductionnelles de la queue des histones	24
1.14	Phénotypes des activateurs et des répresseurs de « variégation » dans la lignée <i>w^{m4h}</i>	28
1.15	Lien entre modificateurs de variégation et structure de la chromatine	29
1.16	Modèle d'initiation de la formation de l'hétérochromatine	29
1.17	Schéma de l'inversion <i>w^{mMc}</i>	31
1.18	Actions des remodeleurs de la chromatine	32
1.19	Structure des différentes familles de remodeleurs de la chromatine	33
1.20	Organisation du locus <i>bithorax</i>	34
1.21	Structure cristallographique de la polymérase II	35
1.22	Initiation de la transcription et facteurs généraux de transcription	36
1.23	Localisation et séquences consensus des boîtes caractéristiques des promoteurs	37
1.24	Modèle de la pause à proximité du promoteur	40
1.25	Modèle d'organisation du locus β -globin	43
1.26	Occupation et positionnement	44
1.27	Positionnement des nucléosomes par rapport au TSS chez la levure et la drosophile	45
1.28	Positionnement des nucléosomes dans les gènes en pause et dans les gènes sans pause	47
1.29	Transcription à travers les nucléosomes	48
1.30	Propriétés définissant les insulateurs	50
1.31	Visualisation sur chromosome polytène de l'effet d'un choc thermique au locus 87A7	51
1.32	Structure du locus 87A7 de la drosophile	51
1.33	Structure des protéines BEAF	52
1.34	Boucle formée entre les éléments <i>scs</i> et <i>scs'</i>	53
1.35	Les différents modèles d'insulation	56
2.1	Organisation du chromosome X	61
2.2	Relocalisation du BE28 dans les mutants <i>w^{m4h}</i> et expression des gènes déplacés	62
2.3	Représentation des gènes proches des répétitions BE28	62
2.4	Déplétion de la protéine BEAF par interférence à l'ARN	63
2.5	Croissance des cellules après déplétion de la protéine BEAF	64
2.6	Effet de la déplétion de BEAF sur l'expression des gènes du groupe euchromatinien	64
2.7	Effet de la déplétion de BEAF sur l'expression des gènes du groupe euchromatinien	65
2.8	Effet de la déplétion de BEAF sur l'expression des gènes du côté hétérochromatinien	66
2.9	Effet de la déplétion de BEAF sur l'expression du rDNA	66
2.10	Modèle proposant la formation d'une boucle du rDNA	67
2.11	Modèle pour expliquer le motif <i>dual-core</i>	86
2.12	Co-immunoprécipitation de BEAF et NELF	120

2.13	Adéquation du profil de nucléosomes obtenu par <i>Southern blot</i> et par MNase-seq	121
2.14	Localisation des sites de liaison de BEAF	122
2.15	Effet de la déplétion de HP1 sur les gènes centromère-proximaux du BE28	123
2.16	Représentation de la régulation de la pause de polymérase par BEAF via la positionnement des nucléosomes	124
2.17	Modèle montrant que la pause de la pol II pourrait favoriser l'interaction avec un activateur	124
3.1	Modèle représentant une barrière non fixe résultant de la compétition entre deux types de structure	127
3.2	Modèle expliquant la fonction insulatrice d'un promoteur en pause	128
3.3	Localisation de BEAF dans la région des gènes <i>hsp70A</i>	130
3.4	Modèle représentant un foyer d'insulateurs	131
3.5	Modèle de fusion de la fibre de 30 nm	132
3.6	Modèle simplifié représentant le locus soumis à l'empreinte parentale <i>Igf2-H19</i>	134

Liste des tables

A.1	Code IUPAC de nomenclature des nucléotides.	139
A.2	Nomenclature des acides aminés	140
B.1	Séquences des oligonucléotides utilisés dans cette étude.	146



Bibliographie

- ADKINS, N. L., HAGERMAN, T. A. & GEORGEL, P. 2006 GAGA protein : a multi-faceted transcription factor. *Biochem. Cell Biol.* **84**, 559–567, [DOI :[10.1139/o06-062](https://doi.org/10.1139/o06-062)] [PubMed :[16936828](https://pubmed.ncbi.nlm.nih.gov/16936828/)].
- AGALIOTI, T., LOMVARDAS, S., PAREKH, B., YIE, J., MANIATIS, T. & THANOS, D. 2000 Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. *Cell* **103**, 667–678, [PubMed :[11106736](https://pubmed.ncbi.nlm.nih.gov/11106736/)].
- ALLFREY, V. G., FAULKNER, R. & MIRSKY, A. E. 1964 Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **51**, 786–794, [PubMed :[14172992](https://pubmed.ncbi.nlm.nih.gov/14172992/)].
- ALLISON, L. A., WONG, J. K., FITZPATRICK, V. D., MOYLE, M. & INGLES, C. J. 1988 The C-terminal domain of the largest subunit of RNA polymerase II of *Saccharomyces cerevisiae*, *Drosophila melanogaster*, and mammals : a conserved structure with an essential function. *Mol. Cell. Biol.* **8**, 321–329, [PubMed :[3122024](https://pubmed.ncbi.nlm.nih.gov/3122024/)].
- ANDERSON, H. E., WARDLE, J., KORKUT, S. V., MURTON, H. E., LOPEZ-MAURY, L., BAHLER, J. & WHITEHALL, S. K. 2009 The fission yeast HIRA histone chaperone is required for promoter silencing and the suppression of cryptic antisense transcripts. *Mol. Cell. Biol.* **29**, 5158–5167, [DOI :[10.1128/MCB.00698-09](https://doi.org/10.1128/MCB.00698-09)] [PubMed :[19620282](https://pubmed.ncbi.nlm.nih.gov/19620282/)].
- ANNUNZIATO, A. T. & SEALE, R. L. 1983 Histone deacetylation is required for the maturation of newly replicated chromatin. *J. Biol. Chem.* **258**, 12675–12684, [PubMed :[6226660](https://pubmed.ncbi.nlm.nih.gov/6226660/)].
- ARAVIND, L. 2000 The BED finger, a novel DNA-binding domain in chromatin-boundary-element-binding proteins and transposases. *Trends Biochem. Sci.* **25**, 421–423, [PubMed :[10973053](https://pubmed.ncbi.nlm.nih.gov/10973053/)].
- ARMSTRONG, J. A., PAPOULAS, O., DAUBRESSE, G., SPERLING, A. S., LIS, J. T., SCOTT, M. P. & TAMKUN, J. W. 2002 The *Drosophila* BRM complex facilitates global transcription by RNA polymerase II. *EMBO J.* **21**, 5245–5254, [PubMed :[12356740](https://pubmed.ncbi.nlm.nih.gov/12356740/)].
- AULNER, N., MONOD, C., MANDICOURT, G., JULLIEN, D., CUVIER, O., SALL, A., JANSSEN, S., LAEMMLI, U. K. & KAS, E. 2002 The AT-hook protein D1 is essential for *Drosophila melanogaster* development and is implicated in position-effect variegation. *Mol. Cell. Biol.* **22**, 1218–1232, [PubMed :[11809812](https://pubmed.ncbi.nlm.nih.gov/11809812/)].
- AUSIO, J. & VAN HOLDE, K. E. 1986 Histone hyperacetylation : its effects on nucleosome conformation and stability. *Biochemistry* **25**, 1421–1428, [PubMed :[3964683](https://pubmed.ncbi.nlm.nih.gov/3964683/)].
- AVERY, O. T., MACLEOD, C. M. & MCCARTY, M. 1944 Studies on the chemical nature of the substance inducing transformation of pneumococcal types : Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *J. Exp. Med.* **79**, 137–158, [PubMed :[19871359](https://pubmed.ncbi.nlm.nih.gov/19871359/)].
- BALBIANI, EG. 1881 Sur la structure du noyau des cellules salivaires chez les larves de *Chironomus*. *Zool. Anz* **4**, 637–641.
- BANNISTER, A. J., ZEGERMAN, P., PARTRIDGE, J. F., MISKA, E. A., THOMAS, J. O., ALLSHIRE, R. C. & KOUZARIDES, T. 2001 Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124, [DOI :[10.1038/35065138](https://doi.org/10.1038/35065138)] [PubMed :[11242054](https://pubmed.ncbi.nlm.nih.gov/11242054/)].

- BAO, Y. & SHEN, X. 2007 INO80 subfamily of chromatin remodeling complexes. *Mutat. Res.* **618**, 18–29, [DOI :[10.1016/j.mrfmmm.2006.10.006](https://doi.org/10.1016/j.mrfmmm.2006.10.006)] [PubMed :[17316710](https://pubmed.ncbi.nlm.nih.gov/17316710/)].
- BARSKI, A., CUDDAPAH, S., CUI, K., ROH, T. Y., SCHONES, D. E., WANG, Z., WEI, G., CHEPELEV, I. & ZHAO, K. 2007 High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837, [DOI :[10.1016/j.cell.2007.05.009](https://doi.org/10.1016/j.cell.2007.05.009)] [PubMed :[17512414](https://pubmed.ncbi.nlm.nih.gov/17512414/)].
- BARTKUHN, M., STRAUB, T., HEROLD, M., HERRMANN, M., RATHKE, C., SAUMWEBER, H., GILFILLAN, G. D., BECKER, P. B. & RENKAWITZ, R. 2009 Active promoters and insulators are marked by the centrosomal protein 190. *EMBO J.* **28**, 877–888, [DOI :[10.1038/emboj.2009.34](https://doi.org/10.1038/emboj.2009.34)] [PubMed :[19229299](https://pubmed.ncbi.nlm.nih.gov/19229299/)].
- BELL, A. C. & FELSENFELD, G. 2000 Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* **405**, 482–485, [DOI :[10.1038/35013100](https://doi.org/10.1038/35013100)] [PubMed :[10839546](https://pubmed.ncbi.nlm.nih.gov/10839546/)].
- BELL, A. C., WEST, A. G. & FELSENFELD, G. 1999 The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**, 387–396, [PubMed :[10458613](https://pubmed.ncbi.nlm.nih.gov/10458613/)].
- BELOTSEKOVSKAYA, R., OH, S., BONDARENKO, V. A., ORPHANIDES, G., STUDITSKY, V. M. & REINBERG, D. 2003 FACT facilitates transcription-dependent nucleosome alteration. *Science* **301**, 1090–1093, [DOI :[10.1126/science.1085703](https://doi.org/10.1126/science.1085703)] [PubMed :[12934006](https://pubmed.ncbi.nlm.nih.gov/12934006/)].
- BHASKAR, V. & COUREY, A. J. 2002 The MADF-BESS domain factor Dip3 potentiates synergistic activation by Dorsal and Twist. *Gene* **299**, 173–184, [PubMed :[12459265](https://pubmed.ncbi.nlm.nih.gov/12459265/)].
- BIGGIN, M. D. & TJIAN, R. 1988 Transcription factors that activate the Ultrabithorax promoter in developmentally staged extracts. *Cell* **53**, 699–711, [PubMed :[2897243](https://pubmed.ncbi.nlm.nih.gov/2897243/)].
- BIRD, A. 2002 DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21, [DOI :[10.1101/gad.947102](https://doi.org/10.1101/gad.947102)] [PubMed :[11782440](https://pubmed.ncbi.nlm.nih.gov/11782440/)].
- BLACK, J. C., CHOI, J. E., LOMBARDO, S. R. & CAREY, M. 2006 A mechanism for coordinating chromatin modification and preinitiation complex assembly. *Mol. Cell* **23**, 809–818, [DOI :[10.1016/j.molcel.2006.07.018](https://doi.org/10.1016/j.molcel.2006.07.018)] [PubMed :[16973433](https://pubmed.ncbi.nlm.nih.gov/16973433/)].
- BLANTON, J., GASZNER, M. & SCHEDL, P. 2003 Protein :protein interactions and the pairing of boundary elements in vivo. *Genes Dev.* **17**, 664–675, [DOI :[10.1101/gad.1052003](https://doi.org/10.1101/gad.1052003)] [PubMed :[12629048](https://pubmed.ncbi.nlm.nih.gov/12629048/)].
- BLATTES, R., MONOD, C., SUSBIELLE, G., CUVIER, O., WU, J. H., HSIEH, T. S., LAEMMLI, U. K. & KAS, E. 2006 Displacement of D1, HP1 and topoisomerase II from satellite heterochromatin by a specific polyamide. *EMBO J.* **25**, 2397–2408, [DOI :[10.1038/sj.emboj.7601125](https://doi.org/10.1038/sj.emboj.7601125)] [PubMed :[16675949](https://pubmed.ncbi.nlm.nih.gov/16675949/)].
- BOEGER, H., BUSHNELL, D. A., DAVIS, R., GRIESENBECK, J., LORCH, Y., STRATTAN, J. S., WESTOVER, K. D. & KORNBERG, R. D. 2005 Structural basis of eukaryotic gene transcription. *FEBS Lett.* **579**, 899–903, [DOI :[10.1016/j.febslet.2004.11.027](https://doi.org/10.1016/j.febslet.2004.11.027)] [PubMed :[15680971](https://pubmed.ncbi.nlm.nih.gov/15680971/)].
- BOEHM, A. K., SAUNDERS, A., WERNER, J. & LIS, J. T. 2003 Transcription factor and polymerase recruitment, modification, and movement on *dhsp70* in vivo in the minutes following heat shock. *Mol. Cell. Biol.* **23**, 7628–7637, [PubMed :[14560008](https://pubmed.ncbi.nlm.nih.gov/14560008/)].
- BOLZER, A., KRETH, G., SOLOVEI, I., KOEHLER, D., SARACOGLU, K., FAUTH, C., MULLER, S., EILS, R., CREMER, C., SPEICHER, M. R. & CREMER, T. 2005 Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol.* **3**, e157, [DOI :[10.1371/journal.pbio.0030157](https://doi.org/10.1371/journal.pbio.0030157)] [PubMed :[15839726](https://pubmed.ncbi.nlm.nih.gov/15839726/)].
- BOVERI, T. 1902 Über mehrpolige Mitosen als Mittel zur Analyse des Zellkerns. *Verh. d. phys.-med. Ges. Würzburg N.F.* **35**, 67–90.
- BREATHNACH, R. & CHAMBON, P. 1981 Organization and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.* **50**, 349–383, [DOI :[10.1146/annurev.bi.50.070181.002025](https://doi.org/10.1146/annurev.bi.50.070181.002025)] [PubMed :[6791577](https://pubmed.ncbi.nlm.nih.gov/6791577/)].

- BROWN, S. A., IMBALZANO, A. N. & KINGSTON, R. E. 1996 Activator-dependent regulation of transcriptional pausing on nucleosomal templates. *Genes Dev.* **10**, 1479–1490, [PubMed :[8666232](#)].
- BUCHNER, K., ROTH, P., SCHOTTA, G., KRAUSS, V., SAUMWEBER, H., REUTER, G. & DORN, R. 2000 Genetic and molecular complexity of the position effect variegation modifier *mod(mdg4)* in *Drosophila*. *Genetics* **155**, 141–157, [PubMed :[10790390](#)].
- BURATOWSKI, S., HAHN, S., SHARP, P. A. & GUARENTE, L. 1988 Function of a yeast TATA element-binding protein in a mammalian transcription system. *Nature* **334**, 37–42, [DOI :[10.1038/334037a0](#)] [PubMed :[3290687](#)].
- BURKE, L. J., ZHANG, R., BARTKUHN, M., TIWARI, V. K., TAVOOSIDANA, G., KURUKUTI, S., WETH, C., LEERS, J., GALJART, N., OHLSSON, R. & RENKAWITZ, R. 2005 CTCF binding and higher order chromatin structure of the H19 locus are maintained in mitotic chromatin. *EMBO J.* **24**, 3291–3300, [DOI :[10.1038/sj.emboj.7600793](#)] [PubMed :[16107875](#)].
- BURKE, T. W. & KADONAGA, J. T. 1996 *Drosophila* TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes Dev.* **10**, 711–724, [PubMed :[8598298](#)].
- BUSHEY, A. M., DORMAN, E. R. & CORCES, V. G. 2008 Chromatin insulators : regulatory mechanisms and epigenetic inheritance. *Mol. Cell* **32**, 1–9, [DOI :[10.1016/j.molcel.2008.08.017](#)] [PubMed :[18851828](#)].
- BUSHEY, A. M., RAMOS, E. & CORCES, V. G. 2009 Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. *Genes Dev.* **23**, 1338–1350, [DOI :[10.1101/gad.1798209](#)] [PubMed :[19443682](#)].
- BUSHNELL, D. A., WESTOVER, K. D., DAVIS, R. E. & KORNBERG, R. D. 2004 Structural basis of transcription : an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. *Science* **303**, 983–988, [DOI :[10.1126/science.1090838](#)] [PubMed :[14963322](#)].
- BUTLER, J. E. & KADONAGA, J. T. 2001 Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs. *Genes Dev.* **15**, 2515–2519, [DOI :[10.1101/gad.924301](#)] [PubMed :[11581157](#)].
- BYRD, K. & CORCES, V. G. 2003 Visualization of chromatin domains created by the gypsy insulator of *Drosophila*. *J. Cell Biol.* **162**, 565–574, [DOI :[10.1083/jcb.200305013](#)] [PubMed :[12925706](#)].
- BYRD, K. N. & SHEARN, A. 2003 ASH1, a *Drosophila* trithorax group protein, is required for methylation of lysine 4 residues on histone H3. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 11535–11540, [DOI :[10.1073/pnas.1933593100](#)] [PubMed :[13679578](#)].
- CAI, W., BAO, X., DENG, H., JIN, Y., GIRTON, J., JOHANSEN, J. & JOHANSEN, K. M. 2008 RNA polymerase II-mediated transcription at active loci does not require histone H3S10 phosphorylation in *Drosophila*. *Development* **135**, 2917–2925, [DOI :[10.1242/dev.024927](#)] [PubMed :[18667461](#)].
- CAIRNS, B. R. 2007 Chromatin remodeling : insights and intrigue from single-molecule studies. *Nat. Struct. Mol. Biol.* **14**, 989–996, [DOI :[10.1038/nsmb1333](#)] [PubMed :[17984961](#)].
- CAIRNS, B. R. 2009 The logic of chromatin architecture and remodelling at promoters. *Nature* **461**, 193–198, [DOI :[10.1038/nature08450](#)] [PubMed :[19741699](#)].
- CAMPOS, E. I. & REINBERG, D. 2009 Histones : annotating chromatin. *Annu. Rev. Genet.* **43**, 559–599, [DOI :[10.1146/annurev.genet.032608.103928](#)] [PubMed :[19886812](#)].
- CAO, R. & ZHANG, Y. 2004 The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Curr. Opin. Genet. Dev.* **14**, 155–164, [DOI :[10.1016/j.gde.2004.02.001](#)] [PubMed :[15196462](#)].
- CAPELSON, M. & CORCES, V. G. 2004 Boundary elements and nuclear organization. *Biol. Cell* **96**, 617–629, [DOI :[10.1016/j.biolcel.2004.06.004](#)] [PubMed :[15519696](#)].

- CAPELSON, M. & CORCES, V. G. 2005 The ubiquitin ligase dTopors directs the nuclear organization of a chromatin insulator. *Mol. Cell* **20**, 105–116, [DOI :[10.1016/j.molcel.2005.08.031](https://doi.org/10.1016/j.molcel.2005.08.031)] [PubMed :[16209949](https://pubmed.ncbi.nlm.nih.gov/16209949/)].
- CAPELSON, M. & CORCES, V. G. 2006 SUMO conjugation attenuates the activity of the gypsy chromatin insulator. *EMBO J.* **25**, 1906–1914, [DOI :[10.1038/sj.emboj.7601068](https://doi.org/10.1038/sj.emboj.7601068)] [PubMed :[16628226](https://pubmed.ncbi.nlm.nih.gov/16628226/)].
- CAREY, M., LI, B. & WORKMAN, J. L. 2006 RSC exploits histone acetylation to abrogate the nucleosomal block to RNA polymerase II elongation. *Mol. Cell* **24**, 481–487, [DOI :[10.1016/j.molcel.2006.09.012](https://doi.org/10.1016/j.molcel.2006.09.012)] [PubMed :[17081996](https://pubmed.ncbi.nlm.nih.gov/17081996/)].
- CARPOUSIS, A. J. & GRALLA, J. D. 1980 Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation in vitro at the lac UV5 promoter. *Biochemistry* **19**, 3245–3253, [PubMed :[6996702](https://pubmed.ncbi.nlm.nih.gov/6996702/)].
- CASAMASSIMI, A. & NAPOLI, C. 2007 Mediator complexes and eukaryotic transcription regulation : an overview. *Biochimie* **89**, 1439–1446, [DOI :[10.1016/j.biochi.2007.08.002](https://doi.org/10.1016/j.biochi.2007.08.002)] [PubMed :[17870225](https://pubmed.ncbi.nlm.nih.gov/17870225/)].
- CHAO, S. H., FUJINAGA, K., MARION, J. E., TAUBE, R., SAUSVILLE, E. A., SENDEROWICZ, A. M., PETERLIN, B. M. & PRICE, D. H. 2000 Flavopiridol inhibits P-TEFb and blocks HIV-1 replication. *J. Biol. Chem.* **275**, 28345–28348, [DOI :[10.1074/jbc.C000446200](https://doi.org/10.1074/jbc.C000446200)] [PubMed :[10906320](https://pubmed.ncbi.nlm.nih.gov/10906320/)].
- CHAO, S. H. & PRICE, D. H. 2001 Flavopiridol inactivates P-TEFb and blocks most RNA polymerase II transcription in vivo. *J. Biol. Chem.* **276**, 31793–31799, [DOI :[10.1074/jbc.M102306200](https://doi.org/10.1074/jbc.M102306200)] [PubMed :[11431468](https://pubmed.ncbi.nlm.nih.gov/11431468/)].
- CHARGAFF, E. 1950 Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* **6**, 201–209, [PubMed :[PubMed](https://pubmed.ncbi.nlm.nih.gov/)].
- CHO, H., KIM, T. K., MANCEBO, H., LANE, W. S., FLORES, O. & REINBERG, D. 1999 A protein phosphatase functions to recycle RNA polymerase II. *Genes Dev.* **13**, 1540–1552, [PubMed :[10385623](https://pubmed.ncbi.nlm.nih.gov/10385623/)].
- CHOPRA, V. S., CANDE, J., HONG, J. W. & LEVINE, M. 2009 Stalled Hox promoters as chromosomal boundaries. *Genes Dev.* **23**, 1505–1509, [DOI :[10.1101/gad.1807309](https://doi.org/10.1101/gad.1807309)] [PubMed :[19515973](https://pubmed.ncbi.nlm.nih.gov/19515973/)].
- CHUANG, C. H., CARPENTER, A. E., FUCHSOVA, B., JOHNSON, T., DE LANEROLLE, P. & BELMONT, A. S. 2006 Long-range directional movement of an interphase chromosome site. *Curr. Biol.* **16**, 825–831, [DOI :[10.1016/j.cub.2006.03.059](https://doi.org/10.1016/j.cub.2006.03.059)] [PubMed :[16631592](https://pubmed.ncbi.nlm.nih.gov/16631592/)].
- CHUNG, J. H., WHITELEY, M. & 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, [PubMed :[8348617](https://pubmed.ncbi.nlm.nih.gov/8348617/)].
- CLAPIER, C. R. & CAIRNS, B. R. 2009 The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* **78**, 273–304, [DOI : [10.1146/annurev.biochem.77.062706.153223](https://doi.org/10.1146/annurev.biochem.77.062706.153223)] [PubMed :[19355820](https://pubmed.ncbi.nlm.nih.gov/19355820/)].
- CLARK, R. J. & FELSENFELD, G. 1971 Structure of chromatin. *Nature New Biol.* **229**, 101–106, [PubMed :[5283616](https://pubmed.ncbi.nlm.nih.gov/5283616/)].
- CONAWAY, J. W., SHILATIFARD, A., DVIR, A. & CONAWAY, R. C. 2000 Control of elongation by RNA polymerase II. *Trends Biochem. Sci.* **25**, 375–380, [PubMed :[10916156](https://pubmed.ncbi.nlm.nih.gov/10916156/)].
- CORDEN, J. L., CADENA, D. L., AHEARN, J. M. & DAHMUS, M. E. 1985 A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7934–7938, [PubMed :[2999785](https://pubmed.ncbi.nlm.nih.gov/2999785/)].
- CORONA, D. F., SIRIACO, G., ARMSTRONG, J. A., SNARSKAYA, N., MCCLYMONT, S. A., SCOTT, M. P. & TAMKUN, J. W. 2007 ISWI regulates higher-order chromatin structure and histone H1 assembly in vivo. *PLoS Biol.* **5**, e232, [DOI :[10.1371/journal.pbio.0050232](https://doi.org/10.1371/journal.pbio.0050232)] [PubMed :[17760505](https://pubmed.ncbi.nlm.nih.gov/17760505/)].

- CORONA, D. F. & TAMKUN, J. W. 2004 Multiple roles for ISWI in transcription, chromosome organization and DNA replication. *Biochim. Biophys. Acta* **1677**, 113–119, [DOI :[10.1016/j.bbaexp.2003.09.018](https://doi.org/10.1016/j.bbaexp.2003.09.018)] [PubMed :[15020052](https://pubmed.ncbi.nlm.nih.gov/15020052/)].
- COSMA, M. P., TANAKA, T. & NASMYTH, K. 1999 Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell* **97**, 299–311, [PubMed :[10319811](https://pubmed.ncbi.nlm.nih.gov/10319811/)].
- COULOMBE, B. & BURTON, Z. F. 1999 DNA bending and wrapping around RNA polymerase : a "revolutionary" model describing transcriptional mechanisms. *Microbiol. Mol. Biol. Rev.* **63**, 457–478, [PubMed :[10357858](https://pubmed.ncbi.nlm.nih.gov/10357858/)].
- CRAIG, J. M. & BICKMORE, W. A. 1993 Chromosome bands—flavours to savour. *Bioessays* **15**, 349–354, [DOI :[10.1002/bies.950150510](https://doi.org/10.1002/bies.950150510)] [PubMed :[8343145](https://pubmed.ncbi.nlm.nih.gov/8343145/)].
- CREMER, T., CREMER, C., SCHNEIDER, T., BAUMANN, H., HENS, L. & KIRSCH-VOLDERS, M. 1982 Analysis of chromosome positions in the interphase nucleus of Chinese hamster cells by laser-UV-microirradiation experiments. *Hum. Genet.* **62**, 201–209, [PubMed :[7169211](https://pubmed.ncbi.nlm.nih.gov/7169211/)].
- CREMER, T. & CREMER, M. 2010 Chromosome territories. *Cold Spring Harb Perspect Biol* **2**, a003889, [DOI :[10.1101/cshperspect.a003889](https://doi.org/10.1101/cshperspect.a003889)] [PubMed :[20300217](https://pubmed.ncbi.nlm.nih.gov/20300217/)].
- CUVIER, O., HART, C. M., KAS, E. & LAEMMLI, U. K. 2002 Identification of a multicopy chromatin boundary element at the borders of silenced chromosomal domains. *Chromosoma* **110**, 519–531, [PubMed :[12068969](https://pubmed.ncbi.nlm.nih.gov/12068969/)].
- CUVIER, O., HART, C. M. & LAEMMLI, U. K. 1998 Identification of a class of chromatin boundary elements. *Mol. Cell. Biol.* **18**, 7478–7486, [PubMed :[9819433](https://pubmed.ncbi.nlm.nih.gov/9819433/)].
- CUÉNOT, L. 1908 Les idées nouvelles sur l'origine des espèces par mutation. *Rev. gén. sc. pures appl.* **19**, 860–871.
- CZERMIN, B., SCHOTTA, G., HULSMANN, B. B., BREHM, A., BECKER, P. B., REUTER, G. & IMHOF, A. 2001 Physical and functional association of SU(VAR)3-9 and HDAC1 in Drosophila. *EMBO Rep.* **2**, 915–919, [DOI :[10.1093/embo-reports/kve210](https://doi.org/10.1093/embo-reports/kve210)] [PubMed :[11571273](https://pubmed.ncbi.nlm.nih.gov/11571273/)].
- DAHM, R. & MIESCHER, F. 2008 Discovering DNA : Friedrich Miescher and the early years of nucleic acid research. *Hum. Genet.* **122**, 565–581, [DOI :[10.1007/s00439-007-0433-0](https://doi.org/10.1007/s00439-007-0433-0)] [PubMed :[17901982](https://pubmed.ncbi.nlm.nih.gov/17901982/)].
- DAVEY, C. A., SARGENT, D. F., LUGER, K., MAEDER, A. W. & RICHMOND, T. J. 2002 Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **319**, 1097–1113, [DOI :[10.1016/S0022-2836\(02\)00386-8](https://doi.org/10.1016/S0022-2836(02)00386-8)] [PubMed :[12079350](https://pubmed.ncbi.nlm.nih.gov/12079350/)].
- DE VRIES, H. 1900 Sur la loi de disjonction des Hybrides. *Comptes Rendus de l'Académie des Sciences (Paris)* **130**, 845–847.
- DEJARDIN, J. & CAVALLI, G. 2005 Epigenetic inheritance of chromatin states mediated by Polycomb and trithorax group proteins in Drosophila. *Prog. Mol. Subcell. Biol.* **38**, 31–63, [PubMed :[15881890](https://pubmed.ncbi.nlm.nih.gov/15881890/)].
- DEKKER, J., RIPPE, K., DEKKER, M. & KLECKNER, N. 2002 Capturing chromosome conformation. *Science* **295**, 1306–1311, [DOI :[10.1126/science.1067799](https://doi.org/10.1126/science.1067799)] [PubMed :[11847345](https://pubmed.ncbi.nlm.nih.gov/11847345/)].
- DELATTRE, M., SPIERER, A., HULO, N. & SPIERER, P. 2002 A new gene in Drosophila melanogaster, Ravus, the phantom of the modifier of position-effect variegation Su(var)3-7. *Int. J. Dev. Biol.* **46**, 167–171, [PubMed :[11902679](https://pubmed.ncbi.nlm.nih.gov/11902679/)].
- DEMAKOVA, O. V., BELIAEVA, E. S. & ZHIMULEV, I. F. 1998 [Somatic pairing of homologs of the fourth chromosome as a reason for suppression of the Dubinin effect in Drosophila melanogaster]. *Genetika* **34**, 629–634, [PubMed :[9719912](https://pubmed.ncbi.nlm.nih.gov/9719912/)].
- DENG, W. & ROBERTS, S. G. 2006 Core promoter elements recognized by transcription factor IIB. *Biochem. Soc. Trans.* **34**, 1051–1053, [DOI :[10.1042/BST0341051](https://doi.org/10.1042/BST0341051)] [PubMed :[17073748](https://pubmed.ncbi.nlm.nih.gov/17073748/)].

- DILLON, N. 2004 Heterochromatin structure and function. *Biol. Cell* **96**, 631–637, [DOI :[10.1016/j.biocel.2004.06.003](https://doi.org/10.1016/j.biocel.2004.06.003)] [PubMed :[15519697](https://pubmed.ncbi.nlm.nih.gov/15519697/)].
- DIMITRI, P., JUNAKOVIC, N. & ARCA, B. 2003 Colonization of heterochromatic genes by transposable elements in *Drosophila*. *Mol. Biol. Evol.* **20**, 503–512, [DOI :[10.1093/molbev/msg048](https://doi.org/10.1093/molbev/msg048)] [PubMed :[12654931](https://pubmed.ncbi.nlm.nih.gov/12654931/)].
- DORIGO, B., SCHALCH, T., KULANGARA, A., DUDA, S., SCHROEDER, R. R. & RICHMOND, T. J. 2004 Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* **306**, 1571–1573, [DOI :[10.1126/science.1103124](https://doi.org/10.1126/science.1103124)] [PubMed :[15567867](https://pubmed.ncbi.nlm.nih.gov/15567867/)].
- DUBIDIN, N. P. & SIDOROV, V. N. 1934 Relation between the effect of a gene and its position in the system. *Am. Naturalist* **48**, 377–381, [JSTOR :[2456938](https://www.jstor.org/stable/2456938)].
- DUBOCHET, J., ADRIAN, M., CHANG, J. J., HOMO, J. C., LEPAULT, J., MCDOWALL, A. W. & SCHULTZ, P. 1988 Cryo-electron microscopy of vitrified specimens. *Q. Rev. Biophys.* **21**, 129–228, [PubMed :[3043536](https://pubmed.ncbi.nlm.nih.gov/3043536/)].
- DUNDR, M., OSPINA, J. K., SUNG, M. H., JOHN, S., UPENDER, M., RIED, T., HAGER, G. L. & MATERA, A. G. 2007 Actin-dependent intranuclear repositioning of an active gene locus in vivo. *J. Cell Biol.* **179**, 1095–1103, [DOI :[10.1083/jcb.200710058](https://doi.org/10.1083/jcb.200710058)] [PubMed :[18070915](https://pubmed.ncbi.nlm.nih.gov/18070915/)].
- DUPRAW, E. J. 1974 Quantitative constraints in the arrangement of human DNA. *Cold Spring Harb. Symp. Quant. Biol.* **38**, 87–98, [PubMed :[4524792](https://pubmed.ncbi.nlm.nih.gov/4524792/)].
- EISSENBERG, J. C., SHILATIFARD, A., DOROKHOV, N. & MICHENER, D. E. 2007 Cdk9 is an essential kinase in *Drosophila* that is required for heat shock gene expression, histone methylation and elongation factor recruitment. *Mol. Genet. Genomics* **277**, 101–114, [DOI :[10.1007/s00438-006-0164-2](https://doi.org/10.1007/s00438-006-0164-2)] [PubMed :[17001490](https://pubmed.ncbi.nlm.nih.gov/17001490/)].
- ELTSOV, M., MACLELLAN, K. M., MAESHIMA, K., FRANGAKIS, A. S. & DUBOCHET, J. 2008 Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 19732–19737, [DOI :[10.1073/pnas.0810057105](https://doi.org/10.1073/pnas.0810057105)] [PubMed :[19064912](https://pubmed.ncbi.nlm.nih.gov/19064912/)].
- EMBERLY, E., BLATTES, R., SCHUETTENGROBER, B., HENNION, M., JIANG, N., HART, C. M., KAS, E. & CUVIER, O. 2008 BEAF regulates cell-cycle genes through the controlled deposition of H3K9 methylation marks into its conserved dual-core binding sites. *PLoS Biol.* **6**, 2896–2910, [DOI :[10.1371/journal.pbio.0060327](https://doi.org/10.1371/journal.pbio.0060327)] [PubMed :[19108610](https://pubmed.ncbi.nlm.nih.gov/19108610/)].
- EVANS, R., FAIRLEY, J. A. & ROBERTS, S. G. 2001 Activator-mediated disruption of sequence-specific DNA contacts by the general transcription factor TFIIB. *Genes Dev.* **15**, 2945–2949, [DOI :[10.1101/gad.206901](https://doi.org/10.1101/gad.206901)] [PubMed :[11711430](https://pubmed.ncbi.nlm.nih.gov/11711430/)].
- FABREGA, C., SHEN, V., SHUMAN, S. & LIMA, C. D. 2003 Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Mol. Cell* **11**, 1549–1561, [PubMed :[12820968](https://pubmed.ncbi.nlm.nih.gov/12820968/)].
- FARKAS, G., GAUSZ, J., GALLONI, M., REUTER, G., GYURKOVICS, H. & KARCH, F. 1994 The *Trithorax*-like gene encodes the *Drosophila* GAGA factor. *Nature* **371**, 806–808, [DOI :[10.1038/371806a0](https://doi.org/10.1038/371806a0)] [PubMed :[7935842](https://pubmed.ncbi.nlm.nih.gov/7935842/)].
- FEDORIW, A. M., STEIN, P., SVOBODA, P., SCHULTZ, R. M. & BARTOLOMEI, M. S. 2004 Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. *Science* **303**, 238–240, [DOI :[10.1126/science.1090934](https://doi.org/10.1126/science.1090934)] [PubMed :[14716017](https://pubmed.ncbi.nlm.nih.gov/14716017/)].
- FEINBERG, A. P. 2007 Phenotypic plasticity and the epigenetics of human disease. *Nature* **447**, 433–440, [DOI :[10.1038/nature05919](https://doi.org/10.1038/nature05919)] [PubMed :[17522677](https://pubmed.ncbi.nlm.nih.gov/17522677/)].
- FELSENFELD, G. & GROUDINE, M. 2003 Controlling the double helix. *Nature* **421**, 448–453, [DOI :[10.1038/nature01411](https://doi.org/10.1038/nature01411)] [PubMed :[12540921](https://pubmed.ncbi.nlm.nih.gov/12540921/)].
- FILIPPOVA, G. N. 2008 Genetics and epigenetics of the multifunctional protein CTCF. *Curr. Top. Dev. Biol.* **80**, 337–360, [DOI :[10.1016/S0070-2153\(07\)80009-3](https://doi.org/10.1016/S0070-2153(07)80009-3)] [PubMed :[17950379](https://pubmed.ncbi.nlm.nih.gov/17950379/)].

- FILIPPOVA, G. N., FAGERLIE, S., KLENOVA, E. M., MYERS, C., DEHNER, Y., GOODWIN, G., NEIMAN, P. E., COLLINS, S. J. & LOBANENKOV, V. V. 1996 An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Mol. Cell. Biol.* **16**, 2802–2813, [PubMed :[8649389](#)].
- FILIPPOVA, G. N., QI, C. F., ULMER, J. E., MOORE, J. M., WARD, M. D., HU, Y. J., LOUKINOV, D. I., PUGACHEVA, E. M., KLENOVA, E. M., GRUNDY, P. E., FEINBERG, A. P., CLETON-JANSEN, A. M., MOERLAND, E. W., CORNELISSE, C. J., SUZUKI, H., KOMIYA, A., LINDBLOM, A., DORION-BONNET, F., NEIMAN, P. E., MORSE, H. C., COLLINS, S. J. & LOBANENKOV, V. V. 2002 Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter its DNA-binding specificity. *Cancer Res.* **62**, 48–52, [PubMed :[11782357](#)].
- FINCH, J. T. & KLUG, A. 1976 Solenoidal model for superstructure in chromatin. *Proc. Natl. Acad. Sci. U.S.A.* **73**, 1897–1901, [PubMed :[1064861](#)].
- FINLAN, L. E. & BICKMORE, W. A. 2008 Porin new light onto chromatin and nuclear organization. *Genome Biol.* **9**, 222, [DOI :[10.1186/gb-2008-9-5-222](#)] [PubMed :[18495049](#)].
- FINN, R. M., BROWNE, K., HODGSON, K. C. & AUSIO, J. 2008 sNASP, a histone H1-specific eukaryotic chaperone dimer that facilitates chromatin assembly. *Biophys. J.* **95**, 1314–1325, [DOI :[10.1529/biophysj.108.130021](#)] [PubMed :[18456819](#)].
- FITZGERALD, P. C., STURGILL, D., SHYAKHTENKO, A., OLIVER, B. & VINSON, C. 2006 Comparative genomics of Drosophila and human core promoters. *Genome Biol.* **7**, R53, [DOI :[10.1186/gb-2006-7-7-r53](#)] [PubMed :[16827941](#)].
- FRANKLIN, R. E. & GOSLING, R. G. 1953 Molecular configuration in sodium thymonucleate. *Nature* **171**, 740–741, [PubMed :[13054694](#)].
- FU, Y., SINHA, M., PETERSON, C. L. & WENG, Z. 2008 The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet.* **4**, e1000138, [DOI :[10.1371/journal.pgen.1000138](#)] [PubMed :[18654629](#)].
- FYODOROV, D. V., BLOWER, M. D., KARPEN, G. H. & KADONAGA, J. T. 2004 Acf1 confers unique activities to ACF/CHRAC and promotes the formation rather than disruption of chromatin in vivo. *Genes Dev.* **18**, 170–183, [DOI :[10.1101/gad.1139604](#)] [PubMed :[14752009](#)].
- GASZNER, M. & FELSENFELD, G. 2006 Insulators : exploiting transcriptional and epigenetic mechanisms. *Nat. Rev. Genet.* **7**, 703–713, [DOI :[10.1038/nrg1925](#)] [PubMed :[16909129](#)].
- GASZNER, M., VAZQUEZ, J. & SCHEDL, P. 1999 The Zw5 protein, a component of the scs chromatin domain boundary, is able to block enhancer-promoter interaction. *Genes Dev.* **13**, 2098–2107, [PubMed Central :[PMC316952](#)] [PubMed :[10465787](#)].
- GERASIMOVA, T. I., BYRD, K. & CORCES, V. G. 2000 A chromatin insulator determines the nuclear localization of DNA. *Mol. Cell* **6**, 1025–1035, [PubMed :[11106742](#)].
- GERASIMOVA, T. I., GDULA, D. A., GERASIMOV, D. V., SIMONOVA, O. & CORCES, V. G. 1995 A Drosophila protein that imparts directionality on a chromatin insulator is an enhancer of position-effect variegation. *Cell* **82**, 587–597, [PubMed :[7664338](#)].
- GERASIMOVA, T. I., LEI, E. P., BUSHEY, A. M. & CORCES, V. G. 2007 Coordinated control of dCTCF and gypsy chromatin insulators in Drosophila. *Mol. Cell* **28**, 761–772, [DOI :[10.1016/j.molcel.2007.09.024](#)] [PubMed :[18082602](#)].
- GERSHENZON, N. I., TRIFONOV, E. N. & IOSHIKHES, I. P. 2006 The features of Drosophila core promoters revealed by statistical analysis. *BMC Genomics* **7**, 161, [DOI :[10.1186/1471-2164-7-161](#)] [PubMed :[16790048](#)].
- GEYER, P. K. 1997 The role of insulator elements in defining domains of gene expression. *Curr. Opin. Genet. Dev.* **7**, 242–248, [PubMed :[9115431](#)].

- GEYER, P. K. & CLARK, I. 2002 Protecting against promiscuity : the regulatory role of insulators. *Cell. Mol. Life Sci.* **59**, 2112–2127, [PubMed :[12568337](#)].
- GEYER, P. K. & CORCES, V. G. 1992 DNA position-specific repression of transcription by a *Drosophila* zinc finger protein. *Genes Dev.* **6**, 1865–1873, [PubMed :[1327958](#)].
- GEYER, P. K., SPANA, C. & CORCES, V. G. 1986 On the molecular mechanism of gypsy-induced mutations at the yellow locus of *Drosophila melanogaster*. *EMBO J.* **5**, 2657–2662, [PubMed :[3096713](#)].
- GHOSH, D., GERASIMOVA, T. I. & CORCES, V. G. 2001 Interactions between the Su(Hw) and Mod(mdg4) proteins required for gypsy insulator function. *EMBO J.* **20**, 2518–2527, [DOI :[10.1093/emboj/20.10.2518](#)] [PubMed :[11350941](#)].
- GILBERT, M. K., TAN, Y. Y. & HART, C. M. 2006 The *Drosophila* boundary element-associated factors BEAF-32A and BEAF-32B affect chromatin structure. *Genetics* **173**, 1365–1375, [DOI :[10.1534/genetics.106.056002](#)] [PubMed :[16648647](#)].
- GILCHRIST, D. A., NECHAEV, S., LEE, C., GHOSH, S. K., COLLINS, J. B., LI, L., GILMOUR, D. S. & ADELMAN, K. 2008 NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev.* **22**, 1921–1933, [DOI :[10.1101/gad.1643208](#)] [PubMed :[18628398](#)].
- GILMOUR, D. S. 2009 Promoter proximal pausing on genes in metazoans. *Chromosoma* **118**, 1–10, [DOI :[10.1007/s00412-008-0182-4](#)] [PubMed :[18830703](#)].
- GILMOUR, D. S. & LIS, J. T. 1986 RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in *Drosophila melanogaster* cells. *Mol. Cell. Biol.* **6**, 3984–3989, [PubMed :[3099167](#)].
- GIRTON, J. R. & JOHANSEN, K. M. 2008 Chromatin structure and the regulation of gene expression : the lessons of PEV in *Drosophila*. *Adv. Genet.* **61**, 1–43, [DOI :[10.1016/S0065-2660\(07\)00001-6](#)] [PubMed :[18282501](#)].
- GOU, D., RUBALCAVA, M., SAUER, S., MORA-BERMUDEZ, F., ERDJUMENT-BROMAGE, H., TEMPST, P., KREMMER, E. & SAUER, F. 2010 SETDB1 is involved in postembryonic DNA methylation and gene silencing in *Drosophila*. *PLoS ONE* **5**, e10581, [DOI :[10.1371/journal.pone.0010581](#)] [PubMed :[20498723](#)].
- GOWHER, H., LEISMANN, O. & JELTSCH, A. 2000 DNA of *Drosophila melanogaster* contains 5-methylcytosine. *EMBO J.* **19**, 6918–6923, [DOI :[10.1093/emboj/19.24.6918](#)] [PubMed :[11118227](#)].
- GRANOK, H., LEIBOVITCH, B. A., SHAFFER, C. D. & ELGIN, S. C. 1995 Chromatin. Ga-ga over GAGA factor. *Curr. Biol.* **5**, 238–241, [PubMed :[7780729](#)].
- GREEN, E. M., ANTCZAK, A. J., BAILEY, A. O., FRANCO, A. A., WU, K. J., YATES, J. R. & KAUFMAN, P. D. 2005 Replication-independent histone deposition by the HIR complex and Asf1. *Curr. Biol.* **15**, 2044–2049, [DOI :[10.1016/j.cub.2005.10.053](#)] [PubMed :[16303565](#)].
- GREGORY, P. D., SCHMID, A., ZAVARI, M., MUNSTERKOTTER, M. & HORZ, W. 1999 Chromatin remodelling at the PHO8 promoter requires SWI-SNF and SAGA at a step subsequent to activator binding. *EMBO J.* **18**, 6407–6414, [DOI :[10.1093/emboj/18.22.6407](#)] [PubMed :[10562552](#)].
- GREWAL, S. I. 2010 RNAi-dependent formation of heterochromatin and its diverse functions. *Curr. Opin. Genet. Dev.* **20**, 134–141, [DOI : [10.1016/j.gde.2010.02.003](#)] [PubMed :[20207534](#)].
- GRIFFITH, F. 1928 The Significance of Pneumococcal Types. *J Hyg (Lond)* **27**, 113–159, [PubMed :[20474956](#)].
- GRIGORYEV, S. A., ARYA, G., CORRELL, S., WOODCOCK, C. L. & SCHLICK, T. 2009 Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13317–13322, [DOI :[10.1073/pnas.0903280106](#)] [PubMed :[19651606](#)].

- GRIGORYEV, S. A., BULYNKO, Y. A. & POPOVA, E. Y. 2006 The end adjusts the means : heterochromatin remodelling during terminal cell differentiation. *Chromosome Res.* **14**, 53–69, [DOI :[10.1007/s10577-005-1021-6](https://doi.org/10.1007/s10577-005-1021-6)] [PubMed :[16506096](https://pubmed.ncbi.nlm.nih.gov/16506096/)].
- GUENTHER, M. G., LEVINE, S. S., BOYER, L. A., JAENISCH, R. & YOUNG, R. A. 2007 A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**, 77–88, [DOI :[10.1016/j.cell.2007.05.042](https://doi.org/10.1016/j.cell.2007.05.042)] [PubMed :[17632057](https://pubmed.ncbi.nlm.nih.gov/17632057/)].
- GURUDATTA, B. V. & CORCES, V. G. 2009 Chromatin insulators : lessons from the fly. *Brief Funct Genomic Proteomic* **8**, 276–282, [DOI :[10.1093/bfgp/elp032](https://doi.org/10.1093/bfgp/elp032)] [PubMed :[19752045](https://pubmed.ncbi.nlm.nih.gov/19752045/)].
- HADJUR, S., WILLIAMS, L. M., RYAN, N. K., COBB, B. S., SEXTON, T., FRASER, P., FISHER, A. G. & MERKENSCHLAGER, M. 2009 Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* **460**, 410–413, [DOI :[10.1038/nature08079](https://doi.org/10.1038/nature08079)] [PubMed :[19458616](https://pubmed.ncbi.nlm.nih.gov/19458616/)].
- HALL, I. M., SHANKARANARAYANA, G. D., NOMA, K., AYOUB, N., COHEN, A. & GREWAL, S. I. 2002 Establishment and maintenance of a heterochromatin domain. *Science* **297**, 2232–2237, [DOI : [10.1126/science.1076466](https://doi.org/10.1126/science.1076466)] [PubMed :[12215653](https://pubmed.ncbi.nlm.nih.gov/12215653/)].
- HAMADA, F. N., PARK, P. J., GORDADZE, P. R. & KURODA, M. I. 2005 Global regulation of X chromosomal genes by the MSL complex in *Drosophila melanogaster*. *Genes Dev.* **19**, 2289–2294, [DOI :[10.1101/gad.1343705](https://doi.org/10.1101/gad.1343705)] [PubMed :[16204180](https://pubmed.ncbi.nlm.nih.gov/16204180/)].
- HAMPSEY, M. 1998 Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol. Mol. Biol. Rev.* **62**, 465–503, [PubMed :[9618449](https://pubmed.ncbi.nlm.nih.gov/9618449/)].
- HAN, M. & GRUNSTEIN, M. 1988 Nucleosome loss activates yeast downstream promoters in vivo. *Cell* **55**, 1137–1145, [PubMed :[2849508](https://pubmed.ncbi.nlm.nih.gov/2849508/)].
- HARK, A. T., SCHOENHERR, C. J., KATZ, D. J., INGRAM, R. S., LEVORSE, J. M. & TILGHMAN, S. M. 2000 CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405**, 486–489, [DOI :[10.1038/35013106](https://doi.org/10.1038/35013106)] [PubMed :[10839547](https://pubmed.ncbi.nlm.nih.gov/10839547/)].
- HARRISON, D. A., GDULA, D. A., COYNE, R. S. & CORCES, V. G. 1993 A leucine zipper domain of the suppressor of Hairy-wing protein mediates its repressive effect on enhancer function. *Genes Dev.* **7**, 1966–1978, [PubMed :[7916729](https://pubmed.ncbi.nlm.nih.gov/7916729/)].
- HART, C. M., CUVIER, O. & LAEMMLI, U. K. 1999 Evidence for an antagonistic relationship between the boundary element-associated factor BEAF and the transcription factor DREF. *Chromosoma* **108**, 375–383, [PubMed :[10591997](https://pubmed.ncbi.nlm.nih.gov/10591997/)].
- HART, C. M., ZHAO, K. & LAEMMLI, U. K. 1997 The scs' boundary element : characterization of boundary element-associated factors. *Mol. Cell. Biol.* **17**, 999–1009, [PubMed Central :[PMC231825](https://pubmed.ncbi.nlm.nih.gov/PMC231825/)] [PubMed :[9001253](https://pubmed.ncbi.nlm.nih.gov/9001253/)].
- HAUSHALTER, K. A. & KADONAGA, J. T. 2003 Chromatin assembly by DNA-translocating motors. *Nat. Rev. Mol. Cell Biol.* **4**, 613–620, [DOI :[10.1038/nrm1177](https://doi.org/10.1038/nrm1177)] [PubMed :[12923523](https://pubmed.ncbi.nlm.nih.gov/12923523/)].
- HEITZ, E. 1928 Das Heterochromatin der Moose. *Jahrb Wiss Botanik* **69**, 762—818.
- HEITZ, E. 1933 Die somatische Heteropyknose bei *Drosophila melanogaster* und ihre genetische Bedeutung. *Cell Tissue Res* **20**, 237–287, [SpringerLink :[m041027563526145](https://www.springerlink.com/m041027563526145/)].
- HEITZ, E. & BAUER, H. 1933 Beweise für die Chromosomennatur der Kernschleifen in den Knäuelkernen von *Bibio hortulanus* L. *Z. Zellforsch. mikrosk. Anat.* **17**, 67–82.
- HENIKOFF, S. & DREESEN, T. D. 1989 Trans-inactivation of the *Drosophila* brown gene : evidence for transcriptional repression and somatic pairing dependence. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6704–6708, [PubMed :[2505257](https://pubmed.ncbi.nlm.nih.gov/2505257/)].
- HILFIKER, A., HILFIKER-KLEINER, D., PANNUTI, A. & LUCCHESI, J. C. 1997 mof, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in *Drosophila*. *EMBO J.* **16**, 2054–2060, [DOI :[10.1093/emboj/16.8.2054](https://doi.org/10.1093/emboj/16.8.2054)] [PubMed :[9155031](https://pubmed.ncbi.nlm.nih.gov/9155031/)].

- HOLMES, V. F. & COZZARELLI, N. R. 2000 Closing the ring : links between SMC proteins and chromosome partitioning, condensation, and supercoiling. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1322–1324, [DOI :[10.1073/pnas.040576797](https://doi.org/10.1073/pnas.040576797)] [PubMed :[10677457](https://pubmed.ncbi.nlm.nih.gov/10677457/)].
- HOU, C., DALE, R. & DEAN, A. 2010 Cell type specificity of chromatin organization mediated by CTCF and cohesin. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 3651–3656, [DOI :[10.1073/pnas.0912087107](https://doi.org/10.1073/pnas.0912087107)] [PubMed :[20133600](https://pubmed.ncbi.nlm.nih.gov/20133600/)].
- HSIEH, T. & BRUTLAG, D. 1979 Sequence and sequence variation within the 1.688 g/cm³ satellite DNA of *Drosophila melanogaster*. *J. Mol. Biol.* **135**, 465–481, [PubMed :[231676](https://pubmed.ncbi.nlm.nih.gov/231676/)].
- HSU, L. M. 2009 Monitoring abortive initiation. *Methods* **47**, 25–36, [DOI :[10.1016/j.ymeth.2008.10.010](https://doi.org/10.1016/j.ymeth.2008.10.010)] [PubMed :[18948204](https://pubmed.ncbi.nlm.nih.gov/18948204/)].
- HUDSON, D. F., VAGNARELLI, P., GASSMANN, R. & EARNSHAW, W. C. 2003 Condensin is required for nonhistone protein assembly and structural integrity of vertebrate mitotic chromosomes. *Dev. Cell* **5**, 323–336, [PubMed :[12919682](https://pubmed.ncbi.nlm.nih.gov/12919682/)].
- ITO, T., LEVENSTEIN, M. E., FYODOROV, D. V., KUTACH, A. K., KOBAYASHI, R. & KADONAGA, J. T. 1999 ACF consists of two subunits, Acf1 and ISWI, that function cooperatively in the ATP-dependent catalysis of chromatin assembly. *Genes Dev.* **13**, 1529–1539, [PubMed :[10385622](https://pubmed.ncbi.nlm.nih.gov/10385622/)].
- IVALDI, M. S., KARAM, C. S. & CORCES, V. G. 2007 Phosphorylation of histone H3 at Ser10 facilitates RNA polymerase II release from promoter-proximal pausing in *Drosophila*. *Genes Dev.* **21**, 2818–2831, [DOI :[10.1101/gad.1604007](https://doi.org/10.1101/gad.1604007)] [PubMed :[17942706](https://pubmed.ncbi.nlm.nih.gov/17942706/)].
- JENUWEIN, T. & ALLIS, C. D. 2001 Translating the histone code. *Science* **293**, 1074–1080, [DOI :[10.1126/science.1063127](https://doi.org/10.1126/science.1063127)] [PubMed :[11498575](https://pubmed.ncbi.nlm.nih.gov/11498575/)].
- JIA, S., NOMA, K. & GREWAL, S. I. 2004 RNAi-independent heterochromatin nucleation by the stress-activated ATF/CREB family proteins. *Science* **304**, 1971–1976, [DOI :[10.1126/science.1099035](https://doi.org/10.1126/science.1099035)] [PubMed :[15218150](https://pubmed.ncbi.nlm.nih.gov/15218150/)].
- JIANG, N., EMBERLY, E., CUVIER, O. & HART, C. M. 2009 Genome-wide mapping of boundary element-associated factor (BEAF) binding sites in *Drosophila melanogaster* links BEAF to transcription. *Mol. Cell. Biol.* **29**, 3556–3568, [DOI :[10.1128/MCB.01748-08](https://doi.org/10.1128/MCB.01748-08)] [PubMed :[19380483](https://pubmed.ncbi.nlm.nih.gov/19380483/)].
- JONES, P. A. & TAKAI, D. 2001 The role of DNA methylation in mammalian epigenetics. *Science* **293**, 1068–1070, [DOI :[10.1126/science.1063852](https://doi.org/10.1126/science.1063852)] [PubMed :[11498573](https://pubmed.ncbi.nlm.nih.gov/11498573/)].
- JUVEN-GERSON, T. & KADONAGA, J. T. 2010 Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev. Biol.* **339**, 225–229, [DOI :[10.1016/j.ydbio.2009.08.009](https://doi.org/10.1016/j.ydbio.2009.08.009)] [PubMed :[19682982](https://pubmed.ncbi.nlm.nih.gov/19682982/)].
- KANO, J., SADAIE, M., URANO, T. & ISHIKAWA, F. 2005 Telomere binding protein Taz1 establishes Swi6 heterochromatin independently of RNAi at telomeres. *Curr. Biol.* **15**, 1808–1819, [DOI :[10.1016/j.cub.2005.09.041](https://doi.org/10.1016/j.cub.2005.09.041)] [PubMed :[16243027](https://pubmed.ncbi.nlm.nih.gov/16243027/)].
- KAPLAN, C. D., LAPRADE, L. & WINSTON, F. 2003 Transcription elongation factors repress transcription initiation from cryptic sites. *Science* **301**, 1096–1099, [DOI :[10.1126/science.1087374](https://doi.org/10.1126/science.1087374)] [PubMed :[12934008](https://pubmed.ncbi.nlm.nih.gov/12934008/)].
- KAPLAN, N., MOORE, I. K., FONDUFE-MITTENDORF, Y., GOSSETT, A. J., TILLO, D., FIELD, Y., LEPROUST, E. M., HUGHES, T. R., LIEB, J. D., WIDOM, J. & SEGAL, E. 2009 The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**, 362–366, [DOI :[10.1038/nature07667](https://doi.org/10.1038/nature07667)] [PubMed :[19092803](https://pubmed.ncbi.nlm.nih.gov/19092803/)].
- KELLUM, R. & SCHEDL, P. 1991 A position-effect assay for boundaries of higher order chromosomal domains. *Cell* **64**, 941–950, [PubMed :[1848159](https://pubmed.ncbi.nlm.nih.gov/1848159/)].
- KELLUM, R. & SCHEDL, P. 1992 A group of scs elements function as domain boundaries in an enhancer-blocking assay. *Mol. Cell. Biol.* **12**, 2424–2431, [PubMed Central :[PMC364415](https://pubmed.ncbi.nlm.nih.gov/PMC364415/)] [PubMed :[1569958](https://pubmed.ncbi.nlm.nih.gov/1569958/)].

- KERPPOLA, T. K. 2009 Polycomb group complexes—many combinations, many functions. *Trends Cell Biol.* **19**, 692–704, [DOI :[10.1016/j.tcb.2009.10.001](https://doi.org/10.1016/j.tcb.2009.10.001)] [PubMed :[19889541](https://pubmed.ncbi.nlm.nih.gov/19889541/)].
- KHOCHBIN, S. 2001 Histone H1 diversity : bridging regulatory signals to linker histone function. *Gene* **271**, 1–12, [PubMed :[11410360](https://pubmed.ncbi.nlm.nih.gov/11410360/)].
- KIM, T. H., BARRERA, L. O., ZHENG, M., QU, C., SINGER, M. A., RICHMOND, T. A., WU, Y., GREEN, R. D. & REN, B. 2005 A high-resolution map of active promoters in the human genome. *Nature* **436**, 876–880, [DOI :[10.1038/nature03877](https://doi.org/10.1038/nature03877)] [PubMed :[15988478](https://pubmed.ncbi.nlm.nih.gov/15988478/)].
- KIM, Y. J., BJORKLUND, S., LI, Y., SAYRE, M. H. & KORNBERG, R. D. 1994 A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77**, 599–608, [PubMed :[8187178](https://pubmed.ncbi.nlm.nih.gov/8187178/)].
- KIMURA, A. & HORIKOSHI, M. 2004 Partition of distinct chromosomal regions : negotiable border and fixed border. *Genes Cells* **9**, 499–508, [DOI :[10.1111/j.1356-9597.2004.00740.x](https://doi.org/10.1111/j.1356-9597.2004.00740.x)] [PubMed :[15189444](https://pubmed.ncbi.nlm.nih.gov/15189444/)].
- KLYMENKO, T. & MULLER, J. 2004 The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep.* **5**, 373–377, [DOI :[10.1038/sj.embor.7400111](https://doi.org/10.1038/sj.embor.7400111)] [PubMed :[15031712](https://pubmed.ncbi.nlm.nih.gov/15031712/)].
- KOERBER, R. T., RHEE, H. S., JIANG, C. & PUGH, B. F. 2009 Interaction of transcriptional regulators with specific nucleosomes across the *Saccharomyces* genome. *Mol. Cell* **35**, 889–902, [DOI :[10.1016/j.molcel.2009.09.011](https://doi.org/10.1016/j.molcel.2009.09.011)] [PubMed :[19782036](https://pubmed.ncbi.nlm.nih.gov/19782036/)].
- KOLESKE, A. J. & YOUNG, R. A. 1994 An RNA polymerase II holoenzyme responsive to activators. *Nature* **368**, 466–469, [DOI :[10.1038/368466a0](https://doi.org/10.1038/368466a0)] [PubMed :[8133894](https://pubmed.ncbi.nlm.nih.gov/8133894/)].
- KOMARNITSKY, P., CHO, E. J. & BURATOWSKI, S. 2000 Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* **14**, 2452–2460, [PubMed :[11018013](https://pubmed.ncbi.nlm.nih.gov/11018013/)].
- KONG, S., BOHL, D., LI, C. & TUAN, D. 1997 Transcription of the HS2 enhancer toward a cis-linked gene is independent of the orientation, position, and distance of the enhancer relative to the gene. *Mol. Cell. Biol.* **17**, 3955–3965, [PubMed :[9199330](https://pubmed.ncbi.nlm.nih.gov/9199330/)].
- KORNBERG, R. D. 1974 Chromatin structure : a repeating unit of histones and DNA. *Science* **184**, 868–871, [PubMed :[4825889](https://pubmed.ncbi.nlm.nih.gov/4825889/)].
- KOSSEL, A. 1911 Ueber die chemische Beschaffenheit des Zellkerns. *Munchen Med. Wochenschrift* **58**, 65–69.
- KOURMOULI, N., JEPPESEN, P., MAHADEVHAIHAH, S., BURGOYNE, P., WU, R., GILBERT, D. M., BONGIORNI, S., PRANTERA, G., FANTI, L., PIMPINELLI, S., SHI, W., FUNDELE, R. & SINGH, P. B. 2004 Heterochromatin and tri-methylated lysine 20 of histone H4 in animals. *J. Cell. Sci.* **117**, 2491–2501, [DOI :[10.1242/jcs.01238](https://doi.org/10.1242/jcs.01238)] [PubMed :[15128874](https://pubmed.ncbi.nlm.nih.gov/15128874/)].
- KUHN, E. J., HART, C. M. & GEYER, P. K. 2004 Studies of the role of the *Drosophila* scs and scs' insulators in defining boundaries of a chromosome puff. *Mol. Cell. Biol.* **24**, 1470–1480, [PubMed :[14749365](https://pubmed.ncbi.nlm.nih.gov/14749365/)].
- KUNERT, N., MARHOLD, J., STANKE, J., STACH, D. & LYKO, F. 2003 A Dnmt2-like protein mediates DNA methylation in *Drosophila*. *Development* **130**, 5083–5090, [DOI :[10.1242/dev.00716](https://doi.org/10.1242/dev.00716)] [PubMed :[12944428](https://pubmed.ncbi.nlm.nih.gov/12944428/)].
- KUTACH, A. K. & KADONAGA, J. T. 2000 The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters. *Mol. Cell. Biol.* **20**, 4754–4764, [PubMed :[10848601](https://pubmed.ncbi.nlm.nih.gov/10848601/)].
- LACHNER, M., O'CARROLL, D., REA, S., MECHTLER, K. & JENUWEIN, T. 2001 Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116–120, [DOI :[10.1038/35065132](https://doi.org/10.1038/35065132)] [PubMed :[11242053](https://pubmed.ncbi.nlm.nih.gov/11242053/)].

- LAGRANGE, T., KAPANIDIS, A. N., TANG, H., REINBERG, D. & EBRIGHT, R. H. 1998 New core promoter element in RNA polymerase II-dependent transcription : sequence-specific DNA binding by transcription factor IIB. *Genes Dev.* **12**, 34–44, [PubMed :9420329].
- LAN, J., HUA, S., HE, X. & ZHANG, Y. 2010 DNA methyltransferases and methyl-binding proteins of mammals. *Acta Biochim. Biophys. Sin. (Shanghai)* **42**, 243–252, [PubMed :20383462].
- LANGELIER, M. F., TRINH, V. & COULOMBE, B. 2002 Gros plan sur l'ARN polymérase II. *Medecine Sciences* **18**, 210–216, [EDK :EDK].
- LANTERMANN, A. B., STRAUB, T., STRALFORS, A., YUAN, G. C., EKWALL, K. & KORBER, P. 2010 Schizosaccharomyces pombe genome-wide nucleosome mapping reveals positioning mechanisms distinct from those of Saccharomyces cerevisiae. *Nat. Struct. Mol. Biol.* **17**, 251–257, [DOI :10.1038/nsmb.1741] [PubMed :20118936].
- LAYBOURN, P. J. & KADONAGA, J. T. 1991 Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II. *Science* **254**, 238–245, [PubMed :1718039].
- LEE, C., LI, X., HECHMER, A., EISEN, M., BIGGIN, M. D., VENTERS, B. J., JIANG, C., LI, J., PUGH, B. F. & GILMOUR, D. S. 2008 NELF and GAGA factor are linked to promoter-proximal pausing at many genes in Drosophila. *Mol. Cell. Biol.* **28**, 3290–3300, [DOI :10.1128/MCB.02224-07] [PubMed :18332113].
- LEE, C. K., SHIBATA, Y., RAO, B., STRAHL, B. D. & LIEB, J. D. 2004 Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat. Genet.* **36**, 900–905, [DOI :10.1038/ng1400] [PubMed :15247917].
- LEE, D. H., GERSHENZON, N., GUPTA, M., IOSHIKHES, I. P., REINBERG, D. & LEWIS, B. A. 2005 Functional characterization of core promoter elements : the downstream core element is recognized by TAF1. *Mol. Cell. Biol.* **25**, 9674–9686, [DOI :10.1128/MCB.25.21.9674-9686.2005] [PubMed :16227614].
- LEE, H., KRAUS, K. W., WOLFNER, M. F. & LIS, J. T. 1992 DNA sequence requirements for generating paused polymerase at the start of hsp70. *Genes Dev.* **6**, 284–295, [PubMed :1737619].
- LEE, J. S., SHUKLA, A., SCHNEIDER, J., SWANSON, S. K., WASHBURN, M. P., FLORENS, L., BHAUMIK, S. R. & SHILATIFARD, A. 2007 Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. *Cell* **131**, 1084–1096, [DOI :10.1016/j.cell.2007.09.046] [PubMed :18083099].
- LEI, E. P. & CORCES, V. G. 2006 RNA interference machinery influences the nuclear organization of a chromatin insulator. *Nat. Genet.* **38**, 936–941, [DOI :10.1038/ng1850] [PubMed :16862159].
- LEWIS, E. B. 1950 The phenomenon of position effect. *Adv. Genet.* **3**, 73–115, [PubMed :15425389].
- LI, M., BELOZEROV, V. E. & CAI, H. N. 2010 Modulation of chromatin boundary activities by nucleosome-remodeling activities in Drosophila melanogaster. *Mol. Cell. Biol.* **30**, 1067–1076, [DOI :10.1128/MCB.00183-09] [PubMed :19995906].
- 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., SANDSTROM, R., BERNSTEIN, B., BENDER, M. A., GROUDINE, M., GNIRKE, A., STAMATOYANNOPOULOS, J., MIRNY, L. A., LANDER, E. S. & DEKKER, J. 2009 Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293, [DOI :10.1126/science.1181369] [PubMed :19815776].
- LIM, C. Y., SANTOSO, B., BOULAY, T., DONG, E., OHLER, U. & KADONAGA, J. T. 2004 The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes Dev.* **18**, 1606–1617, [DOI :10.1101/gad.1193404] [PubMed :15231738].
- LING, J., AINOL, L., ZHANG, L., YU, X., PI, W. & TUAN, D. 2004 HS2 enhancer function is blocked by a transcriptional terminator inserted between the enhancer and the promoter. *J. Biol. Chem.* **279**, 51704–51713, [DOI :10.1074/jbc.M404039200] [PubMed :15465832].

- LOBANENKOV, V. V., NICOLAS, R. H., ADLER, V. V., PATERSON, H., KLENOVA, E. M., POLOTSKAJA, A. V. & GOODWIN, G. H. 1990 A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene* **5**, 1743–1753, [PubMed :2284094].
- LOCKE, J., KOTARSKI, M. A. & 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, [PubMed Central :PMC1203489] [PubMed :3146523].
- LOCKE, J. & TARTOF, K. D. 1994 Molecular analysis of cubitus interruptus (ci) mutations suggests an explanation for the unusual ci position effects. *Mol. Gen. Genet.* **243**, 234–243, [PubMed :8177220].
- LU, B. Y., EMTAGE, P. C., DUYF, B. J., HILLIKER, A. J. & EISSENBERG, J. C. 2000 Heterochromatin protein 1 is required for the normal expression of two heterochromatin genes in *Drosophila*. *Genetics* **155**, 699–708, [PubMed :10835392].
- LU, L. & TOWER, J. 1997 A transcriptional insulator element, the su(Hw) binding site, protects a chromosomal DNA replication origin from position effects. *Mol. Cell. Biol.* **17**, 2202–2206, [PubMed :9121470].
- LU, L., ZHANG, H. & TOWER, J. 2001 Functionally distinct, sequence-specific replicator and origin elements are required for *Drosophila* chorion gene amplification. *Genes Dev.* **15**, 134–146, [PubMed :11157771].
- LU, Q., WALLRATH, L. L., GRANOK, H. & ELGIN, S. C. 1993 (CT)_n (GA)_n repeats and heat shock elements have distinct roles in chromatin structure and transcriptional activation of the *Drosophila* hsp26 gene. *Mol. Cell. Biol.* **13**, 2802–2814, [PubMed :8474442].
- LUGER, K., MADER, A. W., RICHMOND, R. K., SARGENT, D. F. & RICHMOND, T. J. 1997 Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260, [DOI :10.1038/38444] [PubMed :9305837].
- LYKO, F., RAMSAHOYE, B. H. & JAENISCH, R. 2000 DNA methylation in *Drosophila melanogaster*. *Nature* **408**, 538–540, [DOI :10.1038/35046205] [PubMed :11117732].
- MAEDA, R. K. & KARCH, F. 2009 The bithorax complex of *Drosophila* an exceptional Hox cluster. *Curr. Top. Dev. Biol.* **88**, 1–33, [DOI :10.1016/S0070-2153(09)88001-0] [PubMed :19651300].
- MAESHIMA, K., HIHARA, S. & ELTSOV, M. 2010 Chromatin structure : does the 30-nm fibre exist in vivo? *Curr Opin Cell Biol* [DOI :10.1016/j.ceb.2010.03.001] [PubMed :20346642].
- MANDRIOLI, M. 2004 Epigenetic tinkering and evolution : is there any continuity in the role of cytosine methylation from invertebrates to vertebrates? *Cell. Mol. Life Sci.* **61**, 2425–2427, [DOI :10.1007/s00018-004-4184-y] [PubMed :15526149].
- MANDRIOLI, M. & BORSATTI, F. 2006 DNA methylation of fly genes and transposons. *Cell. Mol. Life Sci.* **63**, 1933–1936, [DOI :10.1007/s00018-006-6039-1] [PubMed :16847580].
- MARFELLA, C. G. & IMBALZANO, A. N. 2007 The Chd family of chromatin remodelers. *Mutat. Res.* **618**, 30–40, [DOI :10.1016/j.mrfmmm.2006.07.012] [PubMed :17350655].
- MAS, G., DE NADAL, E., DECHANT, R., RODRIGUEZ DE LA CONCEPCION, M. L., LOGIE, C., JIMENO-GONZALEZ, S., CHAVEZ, S., AMMERER, G. & POSAS, F. 2009 Recruitment of a chromatin remodelling complex by the Hog1 MAP kinase to stress genes. *EMBO J.* **28**, 326–336, [DOI :10.1038/emboj.2008.299] [PubMed :19153600].
- MATHARU, N. K., HUSSAIN, T., SANKARANARAYANAN, R. & MISHRA, R. K. 2010 Vertebrate homologue of *Drosophila* GAGA factor. *J. Mol. Biol.* **400**, 434–447, [DOI :10.1016/j.jmb.2010.05.010] [PubMed :20471984].
- MATSUKAGE, A., HIROSE, F., YOO, M. A. & YAMAGUCHI, M. 2008 The DRE/DREF transcriptional regulatory system : a master key for cell proliferation. *Biochim. Biophys. Acta* **1779**, 81–89, [DOI :10.1016/j.bbagr.2007.11.011] [PubMed :18155677].

- MAVRICH, T. N., IOSHIKHES, I. P., VENTERS, B. J., JIANG, C., TOMSHO, L. P., QI, J., SCHUSTER, S. C., ALBERT, I. & PUGH, B. F. 2008a A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.* **18**, 1073–1083, [DOI :[10.1101/gr.078261.108](https://doi.org/10.1101/gr.078261.108)] [PubMed :[18550805](https://pubmed.ncbi.nlm.nih.gov/18550805/)].
- MAVRICH, T. N., JIANG, C., IOSHIKHES, I. P., LI, X., VENTERS, B. J., ZANTON, S. J., TOMSHO, L. P., QI, J., GLASER, R. L., SCHUSTER, S. C., GILMOUR, D. S., ALBERT, I. & PUGH, B. F. 2008b Nucleosome organization in the *Drosophila* genome. *Nature* **453**, 358–362, [DOI :[10.1038/nature06929](https://doi.org/10.1038/nature06929)] [PubMed :[18408708](https://pubmed.ncbi.nlm.nih.gov/18408708/)].
- MAX, T., SOGAARD, M. & SVEJSTRUP, J. Q. 2007 Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator. *J. Biol. Chem.* **282**, 14113–14120, [DOI :[10.1074/jbc.M701345200](https://doi.org/10.1074/jbc.M701345200)] [PubMed :[17376774](https://pubmed.ncbi.nlm.nih.gov/17376774/)].
- MCGINTY, R. K., KIM, J., CHATTERJEE, C., ROEDER, R. G. & MUIR, T. W. 2008 Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* **453**, 812–816, [DOI :[10.1038/nature06906](https://doi.org/10.1038/nature06906)] [PubMed :[18449190](https://pubmed.ncbi.nlm.nih.gov/18449190/)].
- MCKITTRICK, E., GAFKEN, P. R., AHMAD, K. & HENIKOFF, S. 2004 Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1525–1530, [DOI :[10.1073/pnas.0308092100](https://doi.org/10.1073/pnas.0308092100)] [PubMed :[14732680](https://pubmed.ncbi.nlm.nih.gov/14732680/)].
- MELNIKOVA, L., JUGE, F., GRUZDEVA, N., MAZUR, A., CAVALLI, G. & GEORGIEV, P. 2004 Interaction between the GAGA factor and Mod(mdg4) proteins promotes insulator bypass in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14806–14811, [DOI :[10.1073/pnas.0403959101](https://doi.org/10.1073/pnas.0403959101)] [PubMed :[15465920](https://pubmed.ncbi.nlm.nih.gov/15465920/)].
- MENDEL, G. 1866 Versuche uber Pflanzen-Hybriden. *Verhandlungen des naturforschenden Vereines in Brunn (Abhandlungen)* **4**, 3–47.
- MEYER, P. A., YE, P., SUH, M. H., ZHANG, M. & FU, J. 2009 Structure of the 12-subunit RNA polymerase II refined with the aid of anomalous diffraction data. *J. Biol. Chem.* **284**, 12933–12939, [DOI :[10.1074/jbc.M809199200](https://doi.org/10.1074/jbc.M809199200)] [PubMed :[19289466](https://pubmed.ncbi.nlm.nih.gov/19289466/)].
- MIESCHER, F. 1871 Ueber die chemische Zusammensetzung der Eiterzellen. *Hoppe-Seyler, med. chem. Unters.* **4**, 441–460.
- MILANI, P., CHEVEREAU, G., VAILLANT, C., AUDIT, B., HAFTEK-TERREAU, Z., MARRILLEY, M., BOUVET, P., ARGOUL, F. & ARNEODO, A. 2009 Nucleosome positioning by genomic excluding-energy barriers. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 22257–22262, [DOI :[10.1073/pnas.0909511106](https://doi.org/10.1073/pnas.0909511106)] [PubMed :[20018700](https://pubmed.ncbi.nlm.nih.gov/20018700/)].
- MISHIRO, T., ISHIHARA, K., HINO, S., TSUTSUMI, S., ABURATANI, H., SHIRAHIGE, K., KINOSHITA, Y. & NAKAO, M. 2009 Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster. *EMBO J.* **28**, 1234–1245, [DOI :[10.1038/emboj.2009.81](https://doi.org/10.1038/emboj.2009.81)] [PubMed :[19322193](https://pubmed.ncbi.nlm.nih.gov/19322193/)].
- MISSRA, A. & GILMOUR, D. S. 2010 Interactions between DSIF (DRB sensitivity inducing factor), NELF (negative elongation factor), and the *Drosophila* RNA polymerase II transcription elongation complex. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11301–11306, [DOI :[10.1073/pnas.1000681107](https://doi.org/10.1073/pnas.1000681107)] [PubMed :[20534440](https://pubmed.ncbi.nlm.nih.gov/20534440/)].
- MISULOVIN, Z., SCHWARTZ, Y. B., LI, X. Y., KAHN, T. G., GAUSE, M., MACARTHUR, S., FAY, J. C., EISEN, M. B., PIRROTTA, V., BIGGIN, M. D. & DORSETT, D. 2008 Association of cohesin and Nipped-B with transcriptionally active regions of the *Drosophila melanogaster* genome. *Chromosoma* **117**, 89–102, [DOI :[10.1007/s00412-007-0129-1](https://doi.org/10.1007/s00412-007-0129-1)] [PubMed :[17965872](https://pubmed.ncbi.nlm.nih.gov/17965872/)].
- MITO, Y., HENIKOFF, J. G. & HENIKOFF, S. 2005 Genome-scale profiling of histone H3.3 replacement patterns. *Nat. Genet.* **37**, 1090–1097, [DOI :[10.1038/ng1637](https://doi.org/10.1038/ng1637)] [PubMed :[16155569](https://pubmed.ncbi.nlm.nih.gov/16155569/)].
- MOHAN, M., BARTKUHN, M., HEROLD, M., PHILIPPEN, A., HEINL, N., BARDENHAGEN, I., LEERS, J., WHITE, R. A., RENKAWITZ-POHL, R., SAUMWEBER, H. & RENKAWITZ, R. 2007 The *Drosophila* insulator proteins CTCF and CP190 link enhancer blocking to body patterning. *EMBO J.* **26**, 4203–4214, [DOI :[10.1038/sj.emboj.7601851](https://doi.org/10.1038/sj.emboj.7601851)] [PubMed :[17805343](https://pubmed.ncbi.nlm.nih.gov/17805343/)].

- MOHRMANN, L. & VERRIJZER, C. P. 2005 Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. *Biochim. Biophys. Acta* **1681**, 59–73, [DOI :[10.1016/j.bbaexp.2004.10.005](https://doi.org/10.1016/j.bbaexp.2004.10.005)] [PubMed :[15627498](https://pubmed.ncbi.nlm.nih.gov/15627498/)].
- MOON, H., FILIPPOVA, G., LOUKINOV, D., PUGACHEVA, E., CHEN, Q., SMITH, S. T., MUNHALL, A., GREWE, B., BARTKUHN, M., ARNOLD, R., BURKE, L. J., RENKAWITZ-POHL, R., OHLSSON, R., ZHOU, J., RENKAWITZ, R. & LOBANENKOV, V. 2005 CTCF is conserved from *Drosophila* to humans and confers enhancer blocking of the Fab-8 insulator. *EMBO Rep.* **6**, 165–170, [DOI :[10.1038/sj.embor.7400334](https://doi.org/10.1038/sj.embor.7400334)] [PubMed :[15678159](https://pubmed.ncbi.nlm.nih.gov/15678159/)].
- MORGAN, T. H. 1919 The Physical basis of heredity. *J.B. Lippincott Company* [OpenLibrary :[OpenLibrary](https://openlibrary.org/subjects/biology)].
- MORGAN, T. H., STURTEVANT, A. H., MULLER, H. J. & BRIDGES, C. B. 1915 The Mechanism of Mendelian heredity. *H. Holt and company*.
- MOSHKOVICH, N. & LEI, E. P. 2010 HP1 recruitment in the absence of argonaute proteins in *Drosophila*. *PLoS Genet.* **6**, e1000880, [DOI :[10.1371/journal.pgen.1000880](https://doi.org/10.1371/journal.pgen.1000880)] [PubMed :[20300658](https://pubmed.ncbi.nlm.nih.gov/20300658/)].
- MULLER, H. J. 1930 Types of visible variations induced by X-Rays in *Drosophila*. *J Genet* **22**, 299–334, [SpringerLink :[817gt77381utg356](https://www.springerlink.com/doi/10.1007/978-1-4939-9736-1_10)].
- MUSE, G. W., GILCHRIST, D. A., NECHAEV, S., SHAH, R., PARKER, J. S., GRISSOM, S. F., ZEITLINGER, J. & ADELMAN, K. 2007 RNA polymerase is poised for activation across the genome. *Nat. Genet.* **39**, 1507–1511, [DOI :[10.1038/ng.2007.21](https://doi.org/10.1038/ng.2007.21)] [PubMed :[17994021](https://pubmed.ncbi.nlm.nih.gov/17994021/)].
- NAKAGAWA, T., KAJITANI, T., TOGO, S., MASUKO, N., OHDAN, H., HISHIKAWA, Y., KOJI, T., MATSUYAMA, T., IKURA, T., MURAMATSU, M. & ITO, T. 2008 Deubiquitylation of histone H2A activates transcriptional initiation via trans-histone cross-talk with H3K4 di- and trimethylation. *Genes Dev.* **22**, 37–49, [DOI :[10.1101/gad.1609708](https://doi.org/10.1101/gad.1609708)] [PubMed :[18172164](https://pubmed.ncbi.nlm.nih.gov/18172164/)].
- DE NAPOLES, M., MERMOUD, J. E., WAKAO, R., TANG, Y. A., ENDOH, M., APPANAH, R., NESTEROVA, T. B., SILVA, J., OTTE, A. P., VIDAL, M., KOSEKI, H. & BROCKDORFF, N. 2004 Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* **7**, 663–676, [DOI :[10.1016/j.devcel.2004.10.005](https://doi.org/10.1016/j.devcel.2004.10.005)] [PubMed :[15525528](https://pubmed.ncbi.nlm.nih.gov/15525528/)].
- NATIVIO, R., WENDT, K. S., ITO, Y., HUDDLESTON, J. E., URIBE-LEWIS, S., WOODFINE, K., KRUEGER, C., REIK, W., PETERS, J. M. & MURRELL, A. 2009 Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS Genet.* **5**, e1000739, [DOI :[10.1371/journal.pgen.1000739](https://doi.org/10.1371/journal.pgen.1000739)] [PubMed :[19956766](https://pubmed.ncbi.nlm.nih.gov/19956766/)].
- NATSUME, T. & TANAKA, T. U. 2010 Spatial regulation and organization of DNA replication within the nucleus. *Chromosome Res.* **18**, 7–17, [DOI :[10.1007/s10577-009-9088-0](https://doi.org/10.1007/s10577-009-9088-0)] [PubMed :[19856119](https://pubmed.ncbi.nlm.nih.gov/19856119/)].
- NI, Z., SAUNDERS, A., FUDA, N. J., YAO, J., SUAREZ, J. R., WEBB, W. W. & LIS, J. T. 2008 P-TEFb is critical for the maturation of RNA polymerase II into productive elongation in vivo. *Mol. Cell. Biol.* **28**, 1161–1170, [DOI :[10.1128/MCB.01859-07](https://doi.org/10.1128/MCB.01859-07)] [PubMed :[18070927](https://pubmed.ncbi.nlm.nih.gov/18070927/)].
- NISHA, P., PLANK, J. L. & CSINK, A. K. 2008 Analysis of chromatin structure of genes silenced by heterochromatin in trans. *Genetics* **179**, 359–373, [DOI :[10.1534/genetics.107.084004](https://doi.org/10.1534/genetics.107.084004)] [PubMed :[18493059](https://pubmed.ncbi.nlm.nih.gov/18493059/)].
- NOMA, K., SUGIYAMA, T., CAM, H., VERDEL, A., ZOFALL, M., JIA, S., MOAZED, D. & GREWAL, S. I. 2004 RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. *Nat. Genet.* **36**, 1174–1180, [DOI :[10.1038/ng1452](https://doi.org/10.1038/ng1452)] [PubMed :[15475954](https://pubmed.ncbi.nlm.nih.gov/15475954/)].
- NOWAK, S. J. & CORCES, V. G. 2004 Phosphorylation of histone H3 : a balancing act between chromosome condensation and transcriptional activation. *Trends Genet.* **20**, 214–220, [DOI :[10.1016/j.tig.2004.02.007](https://doi.org/10.1016/j.tig.2004.02.007)] [PubMed :[15041176](https://pubmed.ncbi.nlm.nih.gov/15041176/)].

- O'BRIEN, S. K., CAO, H., NATHANS, R., ALI, A. & RANA, T. M. 2010 P-TEFB kinase complex phosphorylates histone H1 to regulate expression of cellular and HIV-1 genes. *J Biol Chem* [DOI :[10.1074/jbc.M110.125997](https://doi.org/10.1074/jbc.M110.125997)] [PubMed :[20551309](https://pubmed.ncbi.nlm.nih.gov/20551309/)].
- O'HARE, K., CHADWICK, B. P., CONSTANTINOU, A., DAVIS, A. J., MITCHELSON, A. & TUDOR, M. 2002 A 5.9-kb tandem repeat at the euchromatin-heterochromatin boundary of the X chromosome of *Drosophila melanogaster*. *Mol. Genet. Genomics* **267**, 647–655, [DOI :[10.1007/s00438-002-0698-x](https://doi.org/10.1007/s00438-002-0698-x)] [PubMed :[12172804](https://pubmed.ncbi.nlm.nih.gov/12172804/)].
- OHLER, U., LIAO, G. C., NIEMANN, H. & RUBIN, G. M. 2002 Computational analysis of core promoters in the *Drosophila* genome. *Genome Biol.* **3**, RESEARCH0087, [PubMed :[12537576](https://pubmed.ncbi.nlm.nih.gov/12537576/)].
- OLINS, A. L. & OLINS, D. E. 1974 Spheroid chromatin units (v bodies). *Science* **183**, 330–332, [PubMed :[4128918](https://pubmed.ncbi.nlm.nih.gov/4128918/)].
- OLINS, D. E. & OLINS, A. L. 2003 Chromatin history : our view from the bridge. *Nat. Rev. Mol. Cell Biol.* **4**, 809–814, [DOI :[10.1038/nrm1225](https://doi.org/10.1038/nrm1225)] [PubMed :[14570061](https://pubmed.ncbi.nlm.nih.gov/14570061/)].
- ONG, C. T. & CORCES, V. G. 2009 Insulators as mediators of intra- and inter-chromosomal interactions : a common evolutionary theme. *J. Biol.* **8**, 73, [DOI :[10.1186/jbiol165](https://doi.org/10.1186/jbiol165)] [PubMed :[19725934](https://pubmed.ncbi.nlm.nih.gov/19725934/)].
- ORPHANIDES, G., LEROY, G., CHANG, C. H., LUSE, D. S. & REINBERG, D. 1998 FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* **92**, 105–116, [PubMed :[9489704](https://pubmed.ncbi.nlm.nih.gov/9489704/)].
- OUDET, P., GROSS-BELLARD, M. & CHAMBON, P. 1975 Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* **4**, 281–300, [PubMed :[1122558](https://pubmed.ncbi.nlm.nih.gov/1122558/)].
- PAI, C. Y., LEI, E. P., GHOSH, D. & CORCES, V. G. 2004 The centrosomal protein CP190 is a component of the gypsy chromatin insulator. *Mol. Cell* **16**, 737–748, [DOI :[10.1016/j.molcel.2004.11.004](https://doi.org/10.1016/j.molcel.2004.11.004)] [PubMed :[15574329](https://pubmed.ncbi.nlm.nih.gov/15574329/)].
- PAINTER, T. S. 1933 A new method for the study of chromosome rearrangements and the plotting of chromosome maps. *Science* **78**, 585–586, [DOI :[10.1126/science.78.2034.585](https://doi.org/10.1126/science.78.2034.585)] [PubMed :[17801695](https://pubmed.ncbi.nlm.nih.gov/17801695/)].
- PAL-BHADRA, M., LEBOVITCH, B. A., GANDHI, S. G., RAO, M., BHADRA, U., BIRCHLER, J. A. & ELGIN, S. C. 2004 Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. *Science* **303**, 669–672, [DOI :[10.1126/science.1092653](https://doi.org/10.1126/science.1092653)] [PubMed :[14752161](https://pubmed.ncbi.nlm.nih.gov/14752161/)].
- PARDON, J. F., RICHARDS, B. M. & COTTER, R. I. 1974 X-ray diffraction studies on oriented nucleohistone gels. *Cold Spring Harb. Symp. Quant. Biol.* **38**, 75–81, [PubMed :[4524787](https://pubmed.ncbi.nlm.nih.gov/4524787/)].
- PARELHO, V., HADJUR, S., SPIVAKOV, M., LELEU, M., SAUER, S., GREGSON, H. C., JARMUZ, A., CANZONETTA, C., WEBSTER, Z., NESTEROVA, T., COBB, B. S., YOKOMORI, K., DILLON, N., ARAGON, L., FISHER, A. G. & MERKENSCHLAGER, M. 2008 Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**, 422–433, [DOI :[10.1016/j.cell.2008.01.011](https://doi.org/10.1016/j.cell.2008.01.011)] [PubMed :[18237772](https://pubmed.ncbi.nlm.nih.gov/18237772/)].
- PARK, Y. J. & LUGER, K. 2006 Structure and function of nucleosome assembly proteins. *Biochem. Cell Biol.* **84**, 549–558, [DOI :[10.1139/o06-088](https://doi.org/10.1139/o06-088)] [PubMed :[16936827](https://pubmed.ncbi.nlm.nih.gov/16936827/)].
- PARKER, C. S. & TOPOL, J. 1984 A *Drosophila* RNA polymerase II transcription factor contains a promoter-region-specific DNA-binding activity. *Cell* **36**, 357–369, [PubMed :[6537904](https://pubmed.ncbi.nlm.nih.gov/6537904/)].
- PATHAK, R. U., RANGARAJ, N., KALLAPPAGAUDAR, S., MISHRA, K. & MISHRA, R. K. 2007 Boundary element-associated factor 32B connects chromatin domains to the nuclear matrix. *Mol. Cell. Biol.* **27**, 4796–4806, [DOI :[10.1128/MCB.00305-07](https://doi.org/10.1128/MCB.00305-07)] [PubMed :[17485444](https://pubmed.ncbi.nlm.nih.gov/17485444/)].
- PETESCH, S. J. & LIS, J. T. 2008 Rapid, transcription-independent loss of nucleosomes over a large chromatin domain at Hsp70 loci. *Cell* **134**, 74–84, [DOI :[10.1016/j.cell.2008.05.029](https://doi.org/10.1016/j.cell.2008.05.029)] [PubMed :[18614012](https://pubmed.ncbi.nlm.nih.gov/18614012/)].

- PHALKE, S., NICKEL, O., WALLUSCHECK, D., HORTIG, F., ONORATI, M. C. & REUTER, G. 2009 Retrotransposon silencing and telomere integrity in somatic cells of *Drosophila* depends on the cytosine-5 methyltransferase DNMT2. *Nat. Genet.* **41**, 696–702, [DOI :[10.1038/ng.360](https://doi.org/10.1038/ng.360)] [PubMed :[19412177](https://pubmed.ncbi.nlm.nih.gov/19412177/)].
- PHILLIPS, J. E. & CORCES, V. G. 2009 CTCF : master weaver of the genome. *Cell* **137**, 1194–1211, [DOI :[10.1016/j.cell.2009.06.001](https://doi.org/10.1016/j.cell.2009.06.001)] [PubMed :[19563753](https://pubmed.ncbi.nlm.nih.gov/19563753/)].
- PRIOLEAU, M. N., GENDRON, M. C. & HYRIEN, O. 2003 Replication of the chicken beta-globin locus : early-firing origins at the 5' HS4 insulator and the rho- and betaA-globin genes show opposite epigenetic modifications. *Mol. Cell. Biol.* **23**, 3536–3549, [PubMed :[12724412](https://pubmed.ncbi.nlm.nih.gov/12724412/)].
- PUGH, B. F. 2010 A preoccupied position on nucleosomes. *Nat. Struct. Mol. Biol.* **17**, 923, [DOI :[10.1038/nsmb0810-923](https://doi.org/10.1038/nsmb0810-923)] [PubMed :[20683475](https://pubmed.ncbi.nlm.nih.gov/20683475/)].
- REA, S., EISENHABER, F., O'CARROLL, D., STRAHL, B. D., SUN, Z. W., SCHMID, M., OPRAVIL, S., MECHTLER, K., PONTING, C. P., ALLIS, C. D. & JENUWEIN, T. 2000 Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593–599, [DOI :[10.1038/35020506](https://doi.org/10.1038/35020506)] [PubMed :[10949293](https://pubmed.ncbi.nlm.nih.gov/10949293/)].
- REA, S., XOURI, G. & AKHTAR, A. 2007 Males absent on the first (MOF) : from flies to humans. *Oncogene* **26**, 5385–5394, [DOI :[10.1038/sj.onc.1210607](https://doi.org/10.1038/sj.onc.1210607)] [PubMed :[17694080](https://pubmed.ncbi.nlm.nih.gov/17694080/)].
- REDDY, K. L., ZULLO, J. M., BERTOLINO, E. & SINGH, H. 2008 Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* **452**, 243–247, [DOI :[10.1038/nature06727](https://doi.org/10.1038/nature06727)] [PubMed :[18272965](https://pubmed.ncbi.nlm.nih.gov/18272965/)].
- REESE, J. C. 2003 Basal transcription factors. *Curr. Opin. Genet. Dev.* **13**, 114–118, [PubMed :[12672487](https://pubmed.ncbi.nlm.nih.gov/12672487/)].
- REINBERG, D. & SIMS, R. J. 2006 de FACTo nucleosome dynamics. *J. Biol. Chem.* **281**, 23297–23301, [DOI :[10.1074/jbc.R600007200](https://doi.org/10.1074/jbc.R600007200)] [PubMed :[16766522](https://pubmed.ncbi.nlm.nih.gov/16766522/)].
- RICHMOND, T. J. & DAVEY, C. A. 2003 The structure of DNA in the nucleosome core. *Nature* **423**, 145–150, [DOI :[10.1038/nature01595](https://doi.org/10.1038/nature01595)] [PubMed :[12736678](https://pubmed.ncbi.nlm.nih.gov/12736678/)].
- RICHMOND, T. J., FINCH, J. T., RUSHTON, B., RHODES, D. & KLUG, A. 1984 Structure of the nucleosome core particle at 7 Å resolution. *Nature* **311**, 532–537, [PubMed :[6482966](https://pubmed.ncbi.nlm.nih.gov/6482966/)].
- RIDDLE, N. C., LEUNG, W., HAYNES, K. A., GRANOK, H., WULLER, J. & ELGIN, S. C. 2008 An investigation of heterochromatin domains on the fourth chromosome of *Drosophila melanogaster*. *Genetics* **178**, 1177–1191, [DOI :[10.1534/genetics.107.081828](https://doi.org/10.1534/genetics.107.081828)] [PubMed :[18245350](https://pubmed.ncbi.nlm.nih.gov/18245350/)].
- RIDDLE, N. C., SHAFFER, C. D. & ELGIN, S. C. 2009 A lot about a little dot - lessons learned from *Drosophila melanogaster* chromosome 4. *Biochem. Cell Biol.* **87**, 229–241, [DOI :[10.1139/o08-119](https://doi.org/10.1139/o08-119)] [PubMed :[19234537](https://pubmed.ncbi.nlm.nih.gov/19234537/)].
- ROBINSON, P. J., AN, W., ROUTH, A., MARTINO, F., CHAPMAN, L., ROEDER, R. G. & RHODES, D. 2008 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. *J. Mol. Biol.* **381**, 816–825, [DOI :[10.1016/j.jmb.2008.04.050](https://doi.org/10.1016/j.jmb.2008.04.050)] [PubMed :[18653199](https://pubmed.ncbi.nlm.nih.gov/18653199/)].
- ROBINSON, P. J., FAIRALL, L., HUYNH, V. A. & RHODES, D. 2006 EM measurements define the dimensions of the "30-nm" chromatin fiber : evidence for a compact, interdigitated structure. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6506–6511, [DOI :[10.1073/pnas.0601212103](https://doi.org/10.1073/pnas.0601212103)] [PubMed :[16617109](https://pubmed.ncbi.nlm.nih.gov/16617109/)].
- ROSEMAN, R. R., PIRROTTA, V. & GEYER, P. K. 1993 The su(Hw) protein insulates expression of the *Drosophila melanogaster* white gene from chromosomal position-effects. *EMBO J.* **12**, 435–442, [PubMed Central :[PMC413226](https://pubmed.ncbi.nlm.nih.gov/PMC413226/)] [PubMed :[8382607](https://pubmed.ncbi.nlm.nih.gov/8382607/)].
- ROUGVIE, A. E. & LIS, J. T. 1988 The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged. *Cell* **54**, 795–804, [PubMed :[3136931](https://pubmed.ncbi.nlm.nih.gov/3136931/)].

- ROUTH, A., SANDIN, S. & RHODES, D. 2008 Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 8872–8877, [DOI :[10.1073/pnas.0802336105](https://doi.org/10.1073/pnas.0802336105)] [PubMed :[18583476](https://pubmed.ncbi.nlm.nih.gov/18583476/)].
- ROY, S., GILBERT, M. K. & HART, C. M. 2007 Characterization of BEAF mutations isolated by homologous recombination in *Drosophila*. *Genetics* **176**, 801–813, [DOI :[10.1534/genetics.106.068056](https://doi.org/10.1534/genetics.106.068056)] [PubMed :[17435231](https://pubmed.ncbi.nlm.nih.gov/17435231/)].
- RUDOLPH, T., YONEZAWA, M., LEIN, S., HEIDRICH, K., KUBICEK, S., SCHAFFER, C., PHALKE, S., WALTHER, M., SCHMIDT, A., JENUWEIN, T. & REUTER, G. 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, [DOI :[10.1016/j.molcel.2007.02.025](https://doi.org/10.1016/j.molcel.2007.02.025)] [PubMed :[17434130](https://pubmed.ncbi.nlm.nih.gov/17434130/)].
- SAKAI, A., SCHWARTZ, B. E., GOLDSTEIN, S. & AHMAD, K. 2009 Transcriptional and developmental functions of the H3.3 histone variant in *Drosophila*. *Curr. Biol.* **19**, 1816–1820, [DOI :[10.1016/j.cub.2009.09.021](https://doi.org/10.1016/j.cub.2009.09.021)] [PubMed :[19781938](https://pubmed.ncbi.nlm.nih.gov/19781938/)].
- SALZBERG, A., FISHER, O., SIMAN-TOV, R. & ANKRI, S. 2004 Identification of methylated sequences in genomic DNA of adult *Drosophila melanogaster*. *Biochem. Biophys. Res. Commun.* **322**, 465–469, [DOI :[10.1016/j.bbrc.2004.07.134](https://doi.org/10.1016/j.bbrc.2004.07.134)] [PubMed :[15325253](https://pubmed.ncbi.nlm.nih.gov/15325253/)].
- SAMBROOK, J. & RUSSELL, D. 2001 Molecular Cloning : A Laboratory Manual. *Cold Spring Harbor Laboratory Press* .
- SAUNDERS, A., CORE, L. J. & LIS, J. T. 2006 Breaking barriers to transcription elongation. *Nat. Rev. Mol. Cell Biol.* **7**, 557–567, [DOI :[10.1038/nrm1981](https://doi.org/10.1038/nrm1981)] [PubMed :[16936696](https://pubmed.ncbi.nlm.nih.gov/16936696/)].
- SCHOBORG, T. A. & LABRADOR, M. 2010 The phylogenetic distribution of non-CTCF insulator proteins is limited to insects and reveals that BEAF-32 is *Drosophila* lineage specific. *J. Mol. Evol.* **70**, 74–84, [DOI :[10.1007/s00239-009-9310-x](https://doi.org/10.1007/s00239-009-9310-x)] [PubMed :[20024537](https://pubmed.ncbi.nlm.nih.gov/20024537/)].
- SCHONES, D. E., CUI, K., CUDDAPAH, S., ROH, T. Y., BARSKI, A., WANG, Z., WEI, G. & ZHAO, K. 2008 Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132**, 887–898, [DOI :[10.1016/j.cell.2008.02.022](https://doi.org/10.1016/j.cell.2008.02.022)] [PubMed :[18329373](https://pubmed.ncbi.nlm.nih.gov/18329373/)].
- SCHOTTA, G., EBERT, A., DORN, R. & REUTER, G. 2003 Position-effect variegation and the genetic dissection of chromatin regulation in *Drosophila*. *Semin. Cell Dev. Biol.* **14**, 67–75, [PubMed :[12524009](https://pubmed.ncbi.nlm.nih.gov/12524009/)].
- SCHOTTA, G., LACHNER, M., SARMA, K., EBERT, A., SENGUPTA, R., REUTER, G., REINBERG, D. & JENUWEIN, T. 2004 A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* **18**, 1251–1262, [DOI :[10.1101/gad.300704](https://doi.org/10.1101/gad.300704)] [PubMed :[15145825](https://pubmed.ncbi.nlm.nih.gov/15145825/)].
- SCHUETTENGROBER, B., CHOURROUT, D., VERVOORT, M., LEBLANC, B. & CAVALLI, G. 2007 Genome regulation by polycomb and trithorax proteins. *Cell* **128**, 735–745, [DOI :[10.1016/j.cell.2007.02.009](https://doi.org/10.1016/j.cell.2007.02.009)] [PubMed :[17320510](https://pubmed.ncbi.nlm.nih.gov/17320510/)].
- SCHULTZ, J. 1936 Variegation in *Drosophila* and the Inert Chromosome Regions. *Proc. Natl. Acad. Sci. U.S.A.* **22**, 27–33, [PubMed :[16588038](https://pubmed.ncbi.nlm.nih.gov/16588038/)].
- SCHULZE, S. R. & WALLRATH, L. L. 2007 Gene regulation by chromatin structure : paradigms established in *Drosophila melanogaster*. *Annu. Rev. Entomol.* **52**, 171–192, [DOI :[10.1146/annurev.ento.51.110104.151007](https://doi.org/10.1146/annurev.ento.51.110104.151007)] [PubMed :[16881818](https://pubmed.ncbi.nlm.nih.gov/16881818/)].
- SCHWABISH, M. A. & STRUHL, K. 2004 Evidence for eviction and rapid deposition of histones upon transcriptional elongation by RNA polymerase II. *Mol. Cell. Biol.* **24**, 10111–10117, [DOI :[10.1128/MCB.24.23.10111-10117.2004](https://doi.org/10.1128/MCB.24.23.10111-10117.2004)] [PubMed :[15542822](https://pubmed.ncbi.nlm.nih.gov/15542822/)].
- SCHWABISH, M. A. & STRUHL, K. 2007 The Swi/Snf complex is important for histone eviction during transcriptional activation and RNA polymerase II elongation in vivo. *Mol. Cell. Biol.* **27**, 6987–6995, [DOI :[10.1128/MCB.00717-07](https://doi.org/10.1128/MCB.00717-07)] [PubMed :[17709398](https://pubmed.ncbi.nlm.nih.gov/17709398/)].

- SCHWEINSBERG, S., HAGSTROM, K., GOHL, D., SCHEDL, P., KUMAR, R. P., MISHRA, R. & KARCH, F. 2004 The enhancer-blocking activity of the Fab-7 boundary from the *Drosophila bithorax* complex requires GAGA-factor-binding sites. *Genetics* **168**, 1371–1384, [DOI :[10.1534/genetics.104.029561](https://doi.org/10.1534/genetics.104.029561)] [PubMed :[15579691](https://pubmed.ncbi.nlm.nih.gov/15579691/)].
- SEGAL, E., FONDUFE-MITTENDORF, Y., CHEN, L., THASTROM, A., FIELD, Y., MOORE, I. K., WANG, J. P. & WIDOM, J. 2006 A genomic code for nucleosome positioning. *Nature* **442**, 772–778, [DOI :[10.1038/nature04979](https://doi.org/10.1038/nature04979)] [PubMed :[16862119](https://pubmed.ncbi.nlm.nih.gov/16862119/)].
- SEGAL, E. & WIDOM, J. 2009a Poly(dA :dT) tracts : major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.* **19**, 65–71, [DOI :[10.1016/j.sbi.2009.01.004](https://doi.org/10.1016/j.sbi.2009.01.004)] [PubMed :[19208466](https://pubmed.ncbi.nlm.nih.gov/19208466/)].
- SEGAL, E. & WIDOM, J. 2009b What controls nucleosome positions? *Trends Genet.* **25**, 335–343, [DOI :[10.1016/j.tig.2009.06.002](https://doi.org/10.1016/j.tig.2009.06.002)] [PubMed :[19596482](https://pubmed.ncbi.nlm.nih.gov/19596482/)].
- SELTH, L. A., SIGURDSSON, S. & SVEJSTRUP, J. Q. 2010 Transcript Elongation by RNA Polymerase II. *Annu. Rev. Biochem.* **79**, 271–293, [DOI :[10.1146/annurev.biochem.78.062807.091425](https://doi.org/10.1146/annurev.biochem.78.062807.091425)] [PubMed :[20367031](https://pubmed.ncbi.nlm.nih.gov/20367031/)].
- SEMESHIN, V. F., BARICHEVA, E. M., BELYAEVA, E. S. & ZHIMULEV, I. F. 1982 Electron microscopical analysis of *Drosophila* polytene chromosomes. I. Mapping of the 87A and 87C heat shock puffs in development. *Chromosoma* **87**, 229–237.
- SEXTON, T., BANTIGNIES, F. & CAVALLI, G. 2009 Genomic interactions : chromatin loops and gene meeting points in transcriptional regulation. *Semin. Cell Dev. Biol.* **20**, 849–855, [DOI :[10.1016/j.semcdb.2009.06.004](https://doi.org/10.1016/j.semcdb.2009.06.004)] [PubMed :[19559093](https://pubmed.ncbi.nlm.nih.gov/19559093/)].
- SHAFFER, C. D., CENCI, G., THOMPSON, B., STEPHENS, G. E., SLAWSON, E. E., ADU-WUSU, K., GATTI, M. & ELGIN, S. C. 2006 The large isoform of *Drosophila melanogaster* heterochromatin protein 2 plays a critical role in gene silencing and chromosome structure. *Genetics* **174**, 1189–1204, [DOI :[10.1534/genetics.106.057604](https://doi.org/10.1534/genetics.106.057604)] [PubMed :[16980400](https://pubmed.ncbi.nlm.nih.gov/16980400/)].
- SHEN, X., YU, L., WEIR, J. W. & GOROVSKY, M. A. 1995 Linker histones are not essential and affect chromatin condensation in vivo. *Cell* **82**, 47–56, [PubMed :[7606784](https://pubmed.ncbi.nlm.nih.gov/7606784/)].
- SHIEKHATTAR, R., MERMELSTEIN, F., FISHER, R. P., DRAPKIN, R., DYNLACHT, B., WESSLING, H. C., MORGAN, D. O. & REINBERG, D. 1995 Cdk-activating kinase complex is a component of human transcription factor TFIIH. *Nature* **374**, 283–287, [DOI :[10.1038/374283a0](https://doi.org/10.1038/374283a0)] [PubMed :[7533895](https://pubmed.ncbi.nlm.nih.gov/7533895/)].
- SHOGREN-KNAAK, M., ISHII, H., SUN, J. M., PAZIN, M. J., DAVIE, J. R. & PETERSON, C. L. 2006 Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**, 844–847, [DOI :[10.1126/science.1124000](https://doi.org/10.1126/science.1124000)] [PubMed :[16469925](https://pubmed.ncbi.nlm.nih.gov/16469925/)].
- SHOPLAND, L. S., HIRAYOSHI, K., FERNANDES, M. & LIS, J. T. 1995 HSF access to heat shock elements in vivo depends critically on promoter architecture defined by GAGA factor, TFIIID, and RNA polymerase II binding sites. *Genes Dev.* **9**, 2756–2769, [PubMed :[7590251](https://pubmed.ncbi.nlm.nih.gov/7590251/)].
- SHUKLA, M. S., SYED, S. H., MONTEL, F., FAIVRE-MOSKALENKO, C., BEDNAR, J., TRAVERS, A., ANGELOV, D. & DIMITROV, S. 2010 Remosomes : RSC generated non-mobilized particles with approximately 180 bp DNA loosely associated with the histone octamer. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 1936–1941, [DOI :[10.1073/pnas.0904497107](https://doi.org/10.1073/pnas.0904497107)] [PubMed :[20080697](https://pubmed.ncbi.nlm.nih.gov/20080697/)].
- SIKORSKI, T. W. & BURATOWSKI, S. 2009 The basal initiation machinery : beyond the general transcription factors. *Curr. Opin. Cell Biol.* **21**, 344–351, [DOI :[10.1016/j.ceb.2009.03.006](https://doi.org/10.1016/j.ceb.2009.03.006)] [PubMed :[19411170](https://pubmed.ncbi.nlm.nih.gov/19411170/)].
- SIMIC, R., LINDSTROM, D. L., TRAN, H. G., ROINICK, K. L., COSTA, P. J., JOHNSON, A. D., HARTZOG, G. A. & ARNDT, K. M. 2003 Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes. *EMBO J.* **22**, 1846–1856, [DOI :[10.1093/emboj/cdg179](https://doi.org/10.1093/emboj/cdg179)] [PubMed :[12682017](https://pubmed.ncbi.nlm.nih.gov/12682017/)].
- SIMPSON, R. T. 1978a Structure of chromatin containing extensively acetylated H3 and H4. *Cell* **13**, 691–699, [PubMed :[657272](https://pubmed.ncbi.nlm.nih.gov/657272/)].

- SIMPSON, R. T. 1978b Structure of the chromosome, a chromatin particle containing 160 base pairs of DNA and all the histones. *Biochemistry* **17**, 5524–5531, [PubMed :728412].
- SMITH, E. R., WINTER, B., EISSENBERG, J. C. & SHILATIFARD, A. 2008 Regulation of the transcriptional activity of poised RNA polymerase II by the elongation factor ELL. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 8575–8579, [DOI :10.1073/pnas.0804379105] [PubMed :18562276].
- SPADAFORA, C., OUDET, P. & CHAMBON, P. 1978 The same amount of DNA is organized in in vitro-assembled nucleosomes irrespective of the origin of the histones. *Nucleic Acids Res.* **5**, 3479–3489, [PubMed :214759].
- STARK, A., LIN, M. F., KHERADPOUR, P., PEDERSEN, J. S., PARTS, L., CARLSON, J. W., CROSBY, M. A., RASMUSSEN, M. D., ROY, S., DEORAS, A. N., RUBY, J. G., BRENNECKE, J., HODGES, E., HINRICHS, A. S., CASPI, A., PATEN, B., PARK, S. W., HAN, M. V., MAEDER, M. L., POLANSKY, B. J., ROBSON, B. E., AERTS, S., VAN HELDEN, J., HASSAN, B., GILBERT, D. G., EASTMAN, D. A., RICE, M., WEIR, M., HAHN, M. W., PARK, Y., DEWEY, C. N., PACHTER, L., KENT, W. J., HAUSSLER, D., LAI, E. C., BARTEL, D. P., HANNON, G. J., KAUFMAN, T. C., EISEN, M. B., CLARK, A. G., SMITH, D., CELNIKER, S. E., GELBART, W. M. & KELLIS, M. 2007 Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures. *Nature* **450**, 219–232, [DOI :10.1038/nature06340] [PubMed :17994088].
- STRAHL, B. D. & ALLIS, C. D. 2000 The language of covalent histone modifications. *Nature* **403**, 41–45, [DOI :10.1038/47412] [PubMed :10638745].
- STUDITSKY, V. M., KASSAVETIS, G. A., GEIDUSCHEK, E. P. & FELSENFELD, G. 1997 Mechanism of transcription through the nucleosome by eukaryotic RNA polymerase. *Science* **278**, 1960–1963, [PubMed :9395401].
- SUKA, N., LUO, K. & GRUNSTEIN, M. 2002 Sir2p and Sas2p opposingly regulate acetylation of yeast histone H4 lysine16 and spreading of heterochromatin. *Nat. Genet.* **32**, 378–383, [DOI :10.1038/ng1017] [PubMed :12379856].
- SUN, F. L., CUAYCONG, M. H. & ELGIN, S. C. 2001 Long-range nucleosome ordering is associated with gene silencing in Drosophila melanogaster pericentric heterochromatin. *Mol. Cell. Biol.* **21**, 2867–2879, [DOI :10.1128/MCB.21.8.2867-2879.2001] [PubMed :11283265].
- SUN, F. L., HAYNES, K., SIMPSON, C. L., LEE, S. D., COLLINS, L., WULLER, J., EISSENBERG, J. C. & ELGIN, S. C. 2004 cis-Acting determinants of heterochromatin formation on Drosophila melanogaster chromosome four. *Mol. Cell. Biol.* **24**, 8210–8220, [DOI :10.1128/MCB.24.18.8210-8220.2004] [PubMed :15340080].
- SUTO, R. K., CLARKSON, M. J., TREMETHICK, D. J. & LUGER, K. 2000 Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat. Struct. Biol.* **7**, 1121–1124, [DOI :10.1038/81971] [PubMed :11101893].
- SUTTON, W. S. 1902 On the morphology of the chromosome group in *Brachystola Magna*. *Biol. Bull.* **4**, 24–39, [BioBull :24].
- SVEJSTRUP, J. Q. 2003 Transcription. Histones face the FACT. *Science* **301**, 1053–1055, [DOI :10.1126/science.1088901] [PubMed :12933997].
- SWIFT, H. 1974 The organization of genetic material in eukaryotes : progress and prospects. *Cold Spring Harb. Symp. Quant. Biol.* **38**, 963–979, [PubMed :4524797].
- SYDOW, J. F. & CRAMER, P. 2009 RNA polymerase fidelity and transcriptional proofreading. *Curr. Opin. Struct. Biol.* **19**, 732–739, [DOI :10.1016/j.sbi.2009.10.009] [PubMed :19914059].
- TALBERT, P. B. & HENIKOFF, S. 2000 A reexamination of spreading of position-effect variation in the white-rough region of Drosophila melanogaster. *Genetics* **154**, 259–272, [PubMed :10628986].
- TALBERT, P. B. & HENIKOFF, S. 2006 Spreading of silent chromatin : inaction at a distance. *Nat. Rev. Genet.* **7**, 793–803, [DOI :10.1038/nrg1920] [PubMed :16983375].

- THASTROM, A., BINGHAM, L. M. & WIDOM, J. 2004 Nucleosomal locations of dominant DNA sequence motifs for histone-DNA interactions and nucleosome positioning. *J. Mol. Biol.* **338**, 695–709, [DOI :[10.1016/j.jmb.2004.03.032](https://doi.org/10.1016/j.jmb.2004.03.032)] [PubMed :[15099738](https://pubmed.ncbi.nlm.nih.gov/15099738/)].
- THEISEN, J. W., LIM, C. Y. & KADONAGA, J. T. 2010 Three key subregions contribute to the function of the downstream RNA polymerase II core promoter. *Mol. Cell. Biol.* **30**, 3471–3479, [DOI :[10.1128/MCB.00053-10](https://doi.org/10.1128/MCB.00053-10)] [PubMed :[20457814](https://pubmed.ncbi.nlm.nih.gov/20457814/)].
- THOMAS, M. C. & CHIANG, C. M. 2006 The general transcription machinery and general co-factors. *Crit. Rev. Biochem. Mol. Biol.* **41**, 105–178, [DOI :[10.1080/10409230600648736](https://doi.org/10.1080/10409230600648736)] [PubMed :[16858867](https://pubmed.ncbi.nlm.nih.gov/16858867/)].
- THOMAS, M. J., PLATAS, A. A. & HAWLEY, D. K. 1998 Transcriptional fidelity and proofreading by RNA polymerase II. *Cell* **93**, 627–637, [PubMed :[9604937](https://pubmed.ncbi.nlm.nih.gov/9604937/)].
- THOMAS, T., DIXON, M. P., KUEH, A. J. & VOSS, A. K. 2008 Mof (MYST1 or KAT8) is essential for progression of embryonic development past the blastocyst stage and required for normal chromatin architecture. *Mol. Cell. Biol.* **28**, 5093–5105, [DOI :[10.1128/MCB.02202-07](https://doi.org/10.1128/MCB.02202-07)] [PubMed :[18541669](https://pubmed.ncbi.nlm.nih.gov/18541669/)].
- THOMPSON, C. M., KOLESKE, A. J., CHAO, D. M. & YOUNG, R. A. 1993 A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell* **73**, 1361–1375, [PubMed :[8324825](https://pubmed.ncbi.nlm.nih.gov/8324825/)].
- TOLHUIS, B., PALSTRA, R. J., SPLINTER, E., GROSVELD, F. & DE LAAT, W. 2002 Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell* **10**, 1453–1465, [PubMed :[12504019](https://pubmed.ncbi.nlm.nih.gov/12504019/)].
- TSUKIYAMA, T., BECKER, P. B. & WU, C. 1994 ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. *Nature* **367**, 525–532, [DOI :[10.1038/367525a0](https://doi.org/10.1038/367525a0)] [PubMed :[8107823](https://pubmed.ncbi.nlm.nih.gov/8107823/)].
- UDVARDY, A., MAINE, E. & 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. *J. Mol. Biol.* **185**, 341–358, [PubMed :[2997449](https://pubmed.ncbi.nlm.nih.gov/2997449/)].
- UDVARDY, A. & SCHEDL, P. 1993 The dynamics of chromatin condensation : redistribution of topoisomerase II in the 87A7 heat shock locus during induction and recovery. *Mol. Cell. Biol.* **13**, 7522–7530, [PubMed :[8246970](https://pubmed.ncbi.nlm.nih.gov/8246970/)].
- URIELI-SHOVAL, S., GRUENBAUM, Y., SEDAT, J. & RAZIN, A. 1982 The absence of detectable methylated bases in *Drosophila melanogaster* DNA. *FEBS Lett.* **146**, 148–152, [PubMed :[6814955](https://pubmed.ncbi.nlm.nih.gov/6814955/)].
- VALENZUELA, L. & KAMAKAKA, R. T. 2006 Chromatin insulators. *Annu. Rev. Genet.* **40**, 107–138, [DOI :[10.1146/annurev.genet.39.073003.113546](https://doi.org/10.1146/annurev.genet.39.073003.113546)] [PubMed :[16953792](https://pubmed.ncbi.nlm.nih.gov/16953792/)].
- VALOUEV, A., ICHIKAWA, J., TONTHAT, T., STUART, J., RANADE, S., PECKHAM, H., ZENG, K., MALEK, J. A., COSTA, G., MCKERNAN, K., SIDOW, A., FIRE, A. & JOHNSON, S. M. 2008 A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res.* **18**, 1051–1063, [DOI :[10.1101/gr.076463.108](https://doi.org/10.1101/gr.076463.108)] [PubMed :[18477713](https://pubmed.ncbi.nlm.nih.gov/18477713/)].
- VERDEL, A., JIA, S., GERBER, S., SUGIYAMA, T., GYGI, S., GREWAL, S. I. & MOAZED, D. 2004 RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**, 672–676, [DOI :[10.1126/science.1093686](https://doi.org/10.1126/science.1093686)] [PubMed :[14704433](https://pubmed.ncbi.nlm.nih.gov/14704433/)].
- VOGEL, M. J., PAGIE, L., TALHOUT, W., NIEUWLAND, M., KERKHOVEN, R. M. & VAN STEENSEL, B. 2009 High-resolution mapping of heterochromatin redistribution in a *Drosophila* position-effect variegation model. *Epigenetics Chromatin* **2**, 1, [DOI :[10.1186/1756-8935-2-1](https://doi.org/10.1186/1756-8935-2-1)] [PubMed :[19178722](https://pubmed.ncbi.nlm.nih.gov/19178722/)].
- VOLPE, T. A., KIDNER, C., HALL, I. M., TENG, G., GREWAL, S. I. & MARTIENSSSEN, R. A. 2002 Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**, 1833–1837, [DOI : [10.1126/science.1074973](https://doi.org/10.1126/science.1074973)] [PubMed :[12193640](https://pubmed.ncbi.nlm.nih.gov/12193640/)].

- WANG, D., BUSHNELL, D. A., WESTOVER, K. D., KAPLAN, C. D. & KORNBERG, R. D. 2006 Structural basis of transcription : role of the trigger loop in substrate specificity and catalysis. *Cell* **127**, 941–954, [DOI :[10.1016/j.cell.2006.11.023](https://doi.org/10.1016/j.cell.2006.11.023)] [PubMed :[17129781](https://pubmed.ncbi.nlm.nih.gov/17129781/)].
- WANG, L., BROWN, J. L., CAO, R., ZHANG, Y., KASSIS, J. A. & JONES, R. S. 2004 Hierarchical recruitment of polycomb group silencing complexes. *Mol. Cell* **14**, 637–646, [DOI :[10.1016/j.molcel.2004.05.009](https://doi.org/10.1016/j.molcel.2004.05.009)] [PubMed :[15175158](https://pubmed.ncbi.nlm.nih.gov/15175158/)].
- WATSON, J. D. & CRICK, F. H. 1953 The structure of DNA. *Cold Spring Harb. Symp. Quant. Biol.* **18**, 123–131, [PubMed :[13168976](https://pubmed.ncbi.nlm.nih.gov/13168976/)].
- WEILER, K. S. & WAKIMOTO, B. T. 2002 Suppression of heterochromatic gene variegation can be used to distinguish and characterize E(var) genes potentially important for chromosome structure in *Drosophila melanogaster*. *Mol. Genet. Genomics* **266**, 922–932, [DOI :[10.1007/s00438-001-0633-6](https://doi.org/10.1007/s00438-001-0633-6)] [PubMed :[11862486](https://pubmed.ncbi.nlm.nih.gov/11862486/)].
- WENDT, K. S., YOSHIDA, K., ITOH, T., BANDO, M., KOCH, B., SCHIRGHUBER, E., TSUTSUMI, S., NAGAE, G., ISHIHARA, K., MISHIRO, T., YAHATA, K., IMAMOTO, F., ABURATANI, H., NAKAO, M., IMAMOTO, N., MAESHIMA, K., SHIRAHIGE, K. & PETERS, J. M. 2008 Cohesin mediates transcriptional insulation by CCTC-binding factor. *Nature* **451**, 796–801, [DOI :[10.1038/nature06634](https://doi.org/10.1038/nature06634)] [PubMed :[18235444](https://pubmed.ncbi.nlm.nih.gov/18235444/)].
- WIDOM, J. 2001 Role of DNA sequence in nucleosome stability and dynamics. *Q. Rev. Biophys.* **34**, 269–324, [PubMed :[11838235](https://pubmed.ncbi.nlm.nih.gov/11838235/)].
- WILLIAMS, A., SPILIANAKIS, C. G. & FLAVELL, R. A. 2010 Interchromosomal association and gene regulation in trans. *Trends Genet.* **26**, 188–197, [DOI :[10.1016/j.tig.2010.01.007](https://doi.org/10.1016/j.tig.2010.01.007)] [PubMed :[20236724](https://pubmed.ncbi.nlm.nih.gov/20236724/)].
- WOOD, A. J., SEVERSON, A. F. & MEYER, B. J. 2010 Condensin and cohesin complexity : the expanding repertoire of functions. *Nat. Rev. Genet.* **11**, 391–404, [DOI :[10.1038/nrg2794](https://doi.org/10.1038/nrg2794)] [PubMed :[20442714](https://pubmed.ncbi.nlm.nih.gov/20442714/)].
- WOODCOCK, C. L. 1994 Chromatin fibers observed in situ in frozen hydrated sections. Native fiber diameter is not correlated with nucleosome repeat length. *J. Cell Biol.* **125**, 11–19, [PubMed :[8138565](https://pubmed.ncbi.nlm.nih.gov/8138565/)].
- WOODCOCK, C. L. & DIMITROV, S. 2001 Higher-order structure of chromatin and chromosomes. *Curr. Opin. Genet. Dev.* **11**, 130–135, [PubMed :[11250134](https://pubmed.ncbi.nlm.nih.gov/11250134/)].
- WOODCOCK, C. L., FRADO, L. L. & RATTNER, J. B. 1984 The higher-order structure of chromatin : evidence for a helical ribbon arrangement. *J. Cell Biol.* **99**, 42–52, [PubMed :[6736132](https://pubmed.ncbi.nlm.nih.gov/6736132/)].
- WORKMAN, J. L. 2006 Nucleosome displacement in transcription. *Genes Dev.* **20**, 2009–2017, [DOI :[10.1101/gad.1435706](https://doi.org/10.1101/gad.1435706)] [PubMed :[16882978](https://pubmed.ncbi.nlm.nih.gov/16882978/)].
- WU, C. H., YAMAGUCHI, Y., BENJAMIN, L. R., HORVAT-GORDON, M., WASHINSKY, J., ENERLY, E., LARSSON, J., LAMBERTSSON, A., HANDA, H. & GILMOUR, D. 2003 NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in *Drosophila*. *Genes Dev.* **17**, 1402–1414, [DOI :[10.1101/gad.1091403](https://doi.org/10.1101/gad.1091403)] [PubMed :[12782658](https://pubmed.ncbi.nlm.nih.gov/12782658/)].
- WU, S. Y. & CHIANG, C. M. 1998 Properties of PC4 and an RNA polymerase II complex in directing activated and basal transcription in vitro. *J. Biol. Chem.* **273**, 12492–12498, [PubMed :[9575207](https://pubmed.ncbi.nlm.nih.gov/9575207/)].
- WUSTMANN, G., SZIDONYA, J., TAUBERT, H. & REUTER, G. 1989 The genetics of position-effect variegation modifying loci in *Drosophila melanogaster*. *Mol. Gen. Genet.* **217**, 520–527, [PubMed :[2505058](https://pubmed.ncbi.nlm.nih.gov/2505058/)].
- WYRICK, J. J., HOLSTEGE, F. C., JENNINGS, E. G., CAUSTON, H. C., SHORE, D., GRUNSTEIN, M., LANDER, E. S. & YOUNG, R. A. 1999 Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast. *Nature* **402**, 418–421, [DOI :[10.1038/46567](https://doi.org/10.1038/46567)] [PubMed :[10586882](https://pubmed.ncbi.nlm.nih.gov/10586882/)].

- WYSOCKA, J., SWIGUT, T., XIAO, H., MILNE, T. A., KWON, S. Y., LANDRY, J., KAUER, M., TACKETT, A. J., CHAIT, B. T., BADENHORST, P., WU, C. & ALLIS, C. D. 2006 A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**, 86–90, [DOI :[10.1038/nature04815](https://doi.org/10.1038/nature04815)] [PubMed :[16728976](https://pubmed.ncbi.nlm.nih.gov/16728976/)].
- XU, Q., LI, M., ADAMS, J. & CAI, H. N. 2004 Nuclear location of a chromatin insulator in *Drosophila melanogaster*. *J. Cell. Sci.* **117**, 1025–1032, [DOI :[10.1242/jcs.00964](https://doi.org/10.1242/jcs.00964)] [PubMed :[14996934](https://pubmed.ncbi.nlm.nih.gov/14996934/)].
- YAMADA, T., YAMAGUCHI, Y., INUKAI, N., OKAMOTO, S., MURA, T. & HANDA, H. 2006 P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Mol. Cell* **21**, 227–237, [DOI :[10.1016/j.molcel.2005.11.024](https://doi.org/10.1016/j.molcel.2005.11.024)] [PubMed :[16427012](https://pubmed.ncbi.nlm.nih.gov/16427012/)].
- YAMAGUCHI, Y., INUKAI, N., NARITA, T., WADA, T. & HANDA, H. 2002 Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase II complex and RNA. *Mol. Cell. Biol.* **22**, 2918–2927, [PubMed :[11940650](https://pubmed.ncbi.nlm.nih.gov/11940650/)].
- YANG, C., BOLOTIN, E., JIANG, T., SLADEK, F. M. & MARTINEZ, E. 2007 Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene* **389**, 52–65, [DOI :[10.1016/j.gene.2006.09.029](https://doi.org/10.1016/j.gene.2006.09.029)] [PubMed :[17123746](https://pubmed.ncbi.nlm.nih.gov/17123746/)].
- YOON, Y. S., JEONG, S., RONG, Q., PARK, K. Y., CHUNG, J. H. & PFEIFER, K. 2007 Analysis of the H19ICR insulator. *Mol. Cell. Biol.* **27**, 3499–3510, [DOI :[10.1128/MCB.02170-06](https://doi.org/10.1128/MCB.02170-06)] [PubMed :[17339341](https://pubmed.ncbi.nlm.nih.gov/17339341/)].
- YOSHIDA, H., INOUE, Y. H., HIROSE, F., SAKAGUCHI, K., MATSUKAGE, A. & YAMAGUCHI, M. 2001 Over-expression of DREF in the *Drosophila* wing imaginal disc induces apoptosis and a notching wing phenotype. *Genes Cells* **6**, 877–886, [PubMed :[11683916](https://pubmed.ncbi.nlm.nih.gov/11683916/)].
- YUAN, G. C., LIU, Y. J., DION, M. F., SLACK, M. D., WU, L. F., ALTSCHULER, S. J. & RANDO, O. J. 2005 Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* **309**, 626–630, [DOI :[10.1126/science.1112178](https://doi.org/10.1126/science.1112178)] [PubMed :[15961632](https://pubmed.ncbi.nlm.nih.gov/15961632/)].
- ZHANG, W., DENG, H., BAO, X., LERACH, S., GIRTON, J., JOHANSEN, J. & JOHANSEN, K. M. 2006 The JIL-1 histone H3S10 kinase regulates dimethyl H3K9 modifications and heterochromatic spreading in *Drosophila*. *Development* **133**, 229–235, [DOI :[10.1242/dev.02199](https://doi.org/10.1242/dev.02199)] [PubMed :[16339185](https://pubmed.ncbi.nlm.nih.gov/16339185/)].
- ZHANG, Y., MOQTADERI, Z., RATTNER, B. P., EUSKIRCHEN, G., SNYDER, M., KADONAGA, J. T., LIU, X. S. & STRUHL, K. 2009 Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. *Nat. Struct. Mol. Biol.* **16**, 847–852, [DOI :[10.1038/nsmb.1636](https://doi.org/10.1038/nsmb.1636)] [PubMed :[19620965](https://pubmed.ncbi.nlm.nih.gov/19620965/)].
- ZHANG, Z. & GILMOUR, D. S. 2006 Pcf11 is a termination factor in *Drosophila* that dismantles the elongation complex by bridging the CTD of RNA polymerase II to the nascent transcript. *Mol. Cell* **21**, 65–74, [DOI :[10.1016/j.molcel.2005.11.002](https://doi.org/10.1016/j.molcel.2005.11.002)] [PubMed :[16387654](https://pubmed.ncbi.nlm.nih.gov/16387654/)].
- ZHAO, H. & DEAN, A. 2004 An insulator blocks spreading of histone acetylation and interferes with RNA polymerase II transfer between an enhancer and gene. *Nucleic Acids Res.* **32**, 4903–4919, [DOI :[10.1093/nar/gkh832](https://doi.org/10.1093/nar/gkh832)] [PubMed :[15371553](https://pubmed.ncbi.nlm.nih.gov/15371553/)].
- ZHAO, K., HART, C. M. & LAEMMLI, U. K. 1995 Visualization of chromosomal domains with boundary element-associated factor BEAF-32. *Cell* **81**, 879–889, [PubMed :[7781065](https://pubmed.ncbi.nlm.nih.gov/7781065/)].
- ZHOU, J., FAN, J. Y., RANGASAMY, D. & TREMETHICK, D. J. 2007 The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. *Nat. Struct. Mol. Biol.* **14**, 1070–1076, [DOI :[10.1038/nsmb1323](https://doi.org/10.1038/nsmb1323)] [PubMed :[17965724](https://pubmed.ncbi.nlm.nih.gov/17965724/)].
- ZHOU, W., ZHU, P., WANG, J., PASCUAL, G., OHGI, K. A., LOZACH, J., GLASS, C. K. & ROSENFELD, M. G. 2008 Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. *Mol. Cell* **29**, 69–80, [DOI :[10.1016/j.molcel.2007.11.002](https://doi.org/10.1016/j.molcel.2007.11.002)] [PubMed :[18206970](https://pubmed.ncbi.nlm.nih.gov/18206970/)].

- ZHU, P., ZHOU, W., WANG, J., PUC, J., OHGI, K. A., ERDJUMENT-BROMAGE, H., TEMPST, P., GLASS, C. K. & ROSENFELD, M. G. 2007 A histone H2A deubiquitinase complex coordinating histone acetylation and H1 dissociation in transcriptional regulation. *Mol. Cell* **27**, 609–621, [DOI :[10.1016/j.molcel.2007.07.024](https://doi.org/10.1016/j.molcel.2007.07.024)] [PubMed :[17707232](https://pubmed.ncbi.nlm.nih.gov/17707232/)].



Table des matières

Sommaire	5
Liste des abréviations utilisées	7
Introduction	11
1 Contexte	13
1.1 La chromatine	14
1.1.1 Historique des découvertes majeures sur la chromatine	14
1.1.2 Structure de la chromatine	19
1.1.3 Domaines chromatinien	22
1.1.4 Modifications de la chromatine	24
1.1.5 Dynamique de la chromatine	27
1.2 La transcription des ARN messagers	35
1.2.1 Structure de la polymérase II	35
1.2.2 Formation du complexe de préinitiation	36
1.2.3 Initiation et pause	38
1.2.4 Reprise de l'élongation	40
1.2.5 Terminaison	41
1.2.6 Régulations de la transcription basale	42
1.3 Chromatine et transcription	43
1.3.1 Positionnement des nucléosomes	43
1.3.2 Remodelage de la chromatine aux promoteurs	47
1.3.3 Chromatine et élongation	47
1.4 Les insulateurs	49
1.4.1 Définition classique d'un insulateur	49
1.4.2 BEAF et Zw5	50
1.4.3 L'insulateur <i>gypsy</i>	54
1.4.4 Les insulateurs CTCF	54
1.4.5 Le facteur GAGA	55
1.4.6 CP190	55
1.4.7 les différents modèles d'insulation	56
1.4.8 Rôle des insulateurs <i>in vivo</i>	57
2 Démarche scientifique et résultats	59
2.1 Problématique	60
2.2 Frontière hétérochromatine-euchromatine et gènes hétérochromatiniens	60
2.2.1 Présentation de la région modèle	60
2.2.2 Déplétion par interférence à l'ARN	63
2.2.3 Effet sur la croissance des cellules	63
2.2.4 Effets sur l'expression des gènes euchromatiniens	63
2.2.5 Effets sur l'expression des gènes hétérochromatiniens	65
2.2.6 Effets sur l'expression du rDNA	65
2.2.7 Discussion	67
2.3 Prédiction des sites de liaison de BEAF et régulation des gènes euchromatiniens	68
2.3.1 Introduction	68
2.3.2 Article paru dans <i>PLoS Biology</i>	69
2.3.3 Rappel des principaux résultats	85

2.3.4	Retour sur le motif <i>dual-core</i>	85
2.3.5	Conclusion	86
2.4	Mécanismes de régulation transcriptionnelle	87
2.4.1	Article soumis à <i>Genes and Development</i>	87
2.4.2	Rappel des principaux résultats	119
2.4.3	Quelques données supplémentaires	120
2.4.4	Conclusion	123
3	Discussion et perspectives	125
3.1	Nature des régions frontières	126
3.2	Fonctions de BEAF <i>in vivo</i>	127
3.2.1	Les éléments <i>scs</i> et <i>scs'</i>	127
3.2.2	Activation transcriptionnelle	127
3.2.3	Répression transcriptionnelle	129
3.3	Fonctions des autres insulateurs	129
3.3.1	Points communs	129
3.3.2	Spécialisation des insulateurs	130
3.3.3	Rôles dans d'autres processus cellulaires que la transcription	131
3.4	Formation de boucles de chromatine	132
3.4.1	Contraintes mécaniques	132
3.4.2	Implication du complexe cohesin	133
3.5	Régulation de la fonction insulatrice	133
3.5.1	Régulation du recrutement des facteurs insulateurs par d'autres protéines	133
3.5.2	Modifications post-traductionnelles des protéines insulatrices	134
3.5.3	Implication d'ARN	135
3.6	Les insulateurs chez les vertébrés	135
	Conclusion	137
A	Nomenclature	139
A.1	Code IUPAC	139
A.2	Nomenclature des acides aminés	140
B	Matériel et méthodes	141
B.1	Traitements des cellules	141
B.1.1	Culture cellulaire	141
B.1.2	Interférence ARN	141
B.1.3	FACS	142
B.2	Analyses de l'expression	142
B.2.1	PCR quantitative et RNA-seq	142
B.2.2	<i>Northern blot</i>	142
B.3	Préparation de noyaux, digestion à la MNase et analyses	143
B.3.1	Préparation de noyaux	143
B.3.2	Digestion à la MNase	143
B.3.3	<i>Southern blot</i>	143
B.4	Analyses des protéines	143
B.4.1	Extraits protéiques et <i>western blot</i>	143
B.4.2	Extraits nucléaires et immunoprécipitation	144
B.5	Immunoprécipitation de Chromatine (ChIP)	144
B.6	Séquences des oligonucléotides	146
	Liste des figures	147
	Liste des tables	148
	Bibliographie	149
	Table des matières	173



Mechanisms of transcriptional regulation by insulators involve their role in chromatin organization

Magali HENNION

Ph.D supervisors : Emmanuel KÄS and Olivier CUVIER

Abstract

Chromosomes organization is essential for many processes including genes activity, faithful genetic transmission during cell division and cell proliferation. Insulators are specific DNA sequences that specify chromatin borders, dividing the genome into functionally independent sub-domains. Insulators have been identified in most eukaryotes but the mechanisms that enable them to define borders and/or regulate nearby genes are still unclear. During my PhD studies, I focused on specific insulators whose function is carried out by the BEAF protein in *Drosophila melanogaster*.

By depleting BEAF by RNA interference in cultured cells, I've studied the effect of this protein on chromatin organization at a putative border between euchromatin and heterochromatin, as well as on the expression of surrounding genes. I then extended these studies to the mechanisms of transcriptional regulation that involve BEAF. Localized near many promoters, BEAF insulators regulate the transcription of genes whose expression is essential for cells. Combining several genome-wide approaches, we have shown that BEAF is involved in the precise positioning of nucleosomes at transcription start sites. This positioning enables the regulation of an early step of polymerase II transcription, namely the pause near promoters which appears to be essential for the normal expression of many genes.

Keywords : Chromatin - Insulators - BEAF - Transcription - Nucleosome Positioning

Mécanismes de régulation de la transcription par les insulateurs impliquant leur rôle dans l'organisation de la chromatine

Magali HENNION

Directeurs de thèse : Emmanuel KÄS et Olivier CUVIER

Thèse (spécialité : **Biologie moléculaire et cellulaire**) soutenue le **3 décembre 2010** à
Toulouse

Résumé

L'organisation des chromosomes est essentielle à la régulation de nombreux processus incluant l'activité des gènes, leur transmission fidèle au cours de la division cellulaire et la prolifération des cellules. Les insulateurs sont des séquences d'ADN particulières qui marquent des zones frontières dans la chromatine, subdivisant ainsi le génome en sous-domaines fonctionnellement indépendants. Bien que des insulateurs aient été identifiés chez la plupart des eucaryotes, les mécanismes qui leur permettent de définir des frontières et/ou de réguler les gènes voisins restent à préciser. Mon travail de thèse a porté sur l'étude des insulateurs dont la fonction est assurée par la protéine BEAF chez *Drosophila melanogaster*.

Par une approche de déplétion de BEAF par interférence à l'ARN dans des cellules en culture, j'ai étudié l'influence de cette protéine sur l'organisation d'une frontière putative entre euchromatine et hétérochromatine, ainsi que sur l'expression des gènes environnants. Je me suis ensuite intéressée aux mécanismes de régulation transcriptionnelle impliquant BEAF. Proches de nombreux promoteurs, les insulateurs BEAF régulent la transcription de gènes dont l'expression est essentielle pour la cellule. En combinant plusieurs approches à l'échelle du génome, nous avons montré que BEAF est impliqué dans le positionnement précis des nucléosomes au niveau du site de démarrage de la transcription. Ce positionnement permet la régulation d'une étape précoce de la transcription par la polymérase II, à savoir l'étape de pause à proximité du promoteur qui apparaît comme primordiale pour l'expression normale d'un grand nombre de gènes.

Mots-clés : Chromatine - Insulateurs - BEAF - Transcription - Positionnement des nucléosomes

Laboratoire de Biologie Moléculaire Eucaryote (UMR 5099)
Bâtiment IBCG
118, route de Narbonne
31062 Toulouse cedex 9
FRANCE