

# Rapport d'Activité

4 janvier 2015

## TABLE DES MATIÈRES

<b>1</b>	<b>Parcours : Formation et Recherche</b>	<b>2</b>
<b>2</b>	<b>Activités de Recherche : Objectifs généraux</b>	<b>4</b>
<b>3</b>	<b>Chromatine et organisation du génome</b>	<b>4</b>
3.1	Effet de séquence sur l'élasticité de chaînes ADN : . . . . .	5
3.2	Positionnement des nucléosomes : . . . . .	5
<b>4</b>	<b>Origines de réplication et nucléosomes</b>	<b>11</b>
<b>5</b>	<b>Chromatine et insertion virale</b>	<b>11</b>
<b>6</b>	<b>Positionnement des Nucleosomes : vers le <i>in vivo</i></b>	<b>12</b>
6.1	Modélisation des effets "extrinsèques" . . . . .	12
6.1.1	Facteurs de remodelage : modèles effectifs Forces & Fluidification . . . . .	13
6.1.2	Modèle de compétition nucléosome vs. facteurs de transcriptions . . . . .	13
6.2	Les mic MAC de la Chromatine . . . . .	13
6.3	Mécano-transduction et chromatine chez les plantes . . . . .	14
<b>7</b>	<b>Repliement de l'épigénome</b>	<b>14</b>
7.1	Objectifs . . . . .	14
7.2	Application (I) : Drosophile . . . . .	15
7.3	Application (II) : <i>C. elegans</i> . . . . .	18
<b>8</b>	<b>Motilité bactérienne</b>	<b>19</b>

## 1 PARCOURS : FORMATION ET RECHERCHE

Cédric Vaillant  
43 ans  
Vie Maritale, 3 Enfants (9, 7 et 3 ans)

Chargé de Recherche (CR1)  
Docteur en Sc. Physiques  
Ingénieur Supélec

cedric.vaillant@ens-lyon.fr

### ANNEES DE FORMATION PRE-DOCTORALES

1996-1998 : Coopération scientifique (CSNE) au Centre de Sismologie de Mongolie (Oulan Bator). Responsable local de la maintenance technique et de l'extension du réseau d'analyse sismique installé par le CEA en Mongolie.

1995-1996 : DEA "Champs, Particules, Matières" (Université Paris VI). Mai-Juin 96 : Stage de DEA au Laboratoire de l'Accélérateur Linéaire (Orsay) sur le projet "Virgo" de détection d'ondes gravitationnelles. Etudes relatives à l'asservissement des miroirs des cavités interférométriques.

1993-1996 : Diplôme d'Ingénieur Supélec (Ecole Supérieure d'Electricité).

1990-1993 : Classes préparatoires scientifiques (Math-sup, Math-spé P', Lycée Pothiers, Orléans).

### ANNEES DE RECHERCHE DOCTORALES et Post-DOCTORALES

2004-2006 : Post-doctorat au Laboratoire Statistique et Génome. Projet de recherche : Organisation et dynamique de la chromatine.

2002-2004 : Années de Post-doctorat à l'Ecole Polytechnique Fédérale de Lausanne (Lausanne, Suisse) dans le Groupe de John Maddocks ("Laboratory for Computation and Visualization in Mathematics and Mechanics", Faculté des Sciences de Base, Institut Bernoulli). Sujet : Influence de la séquence sur les propriétés de conformation et d'organisation des chaînes ADN.

1998-2001 : Thèse au Centre de Recherche Paul Pascal (CNRS, Bordeaux), sous la direction de Alain Arnéodo et soutenue le 14 Décembre 2001 pour l'obtention du grade de Docteur en Sciences Physiques. Sujet : Influence de la séquence sur les propriétés élastiques des chaînes ADN. Jury de thèse : Martine Ben Amar (Présidente), Vincent Croquette, Marc Mézard (rapporteurs), Alain Arnéodo (directeur de thèse), Didier Chatenay, Richard Lavery, Claude Thermes.

### ANNEES DE RECHERCHE CR CNRS

2012- : Membre du laboratoire de Physique de l'ENS de Lyon et du Centre Blaise Pascal. Co-direction avec Daniel Jost (CR2 CNRS TIMC-IMAG, Grenoble) du Groupe "Physics of Epigenome Régulation" (<http://perso.ens-lyon.fr/cedric.vaillant>). Directeur de thèse de Noelle Haddad (Sept. 2013-). Membre associé du réseau Epigenesys (<http://www.epigenesys.eu/en/>). Membre du comité de pilotage de Semovi (<http://www.biosyl.org/resources/semovi>).

2011 : Soutenance de mon Habilitation à Diriger des Recherches : "Chapelet Nucléosomal et Organisation du Génome" (<http://perso.ens-lyon.fr/cedric.vaillant/doku.php?id=divers>)

2006-2011 : CR dans l'équipe d'Alain Arnéodo "Chromatine et Génome" au sein du Laboratoire Joliot-Curie (<http://www.ens-lyon.fr/Joliot-Curie/>).

**Publications : 33**

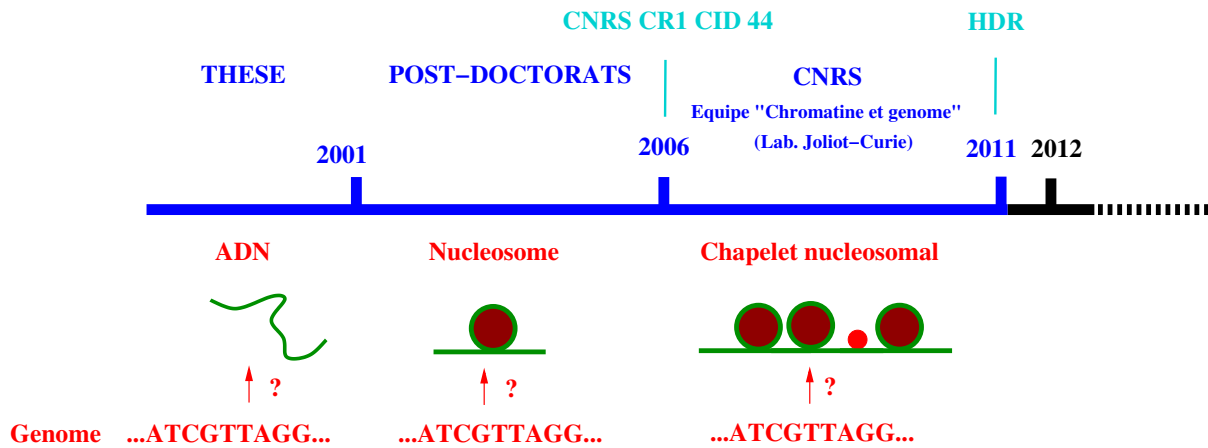


FIGURE 1: Parcours de recherche I

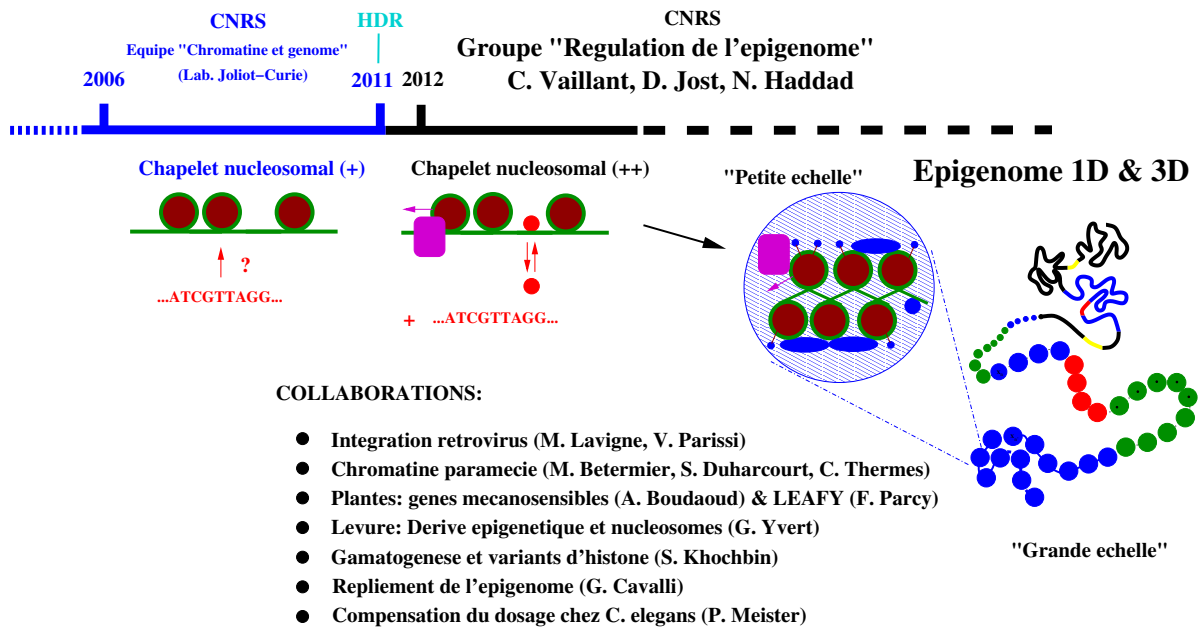


FIGURE 2: Parcours de recherche II

Remarque :

On peut trouver ma liste de publication à l'adresse :  
<http://perso.ens-lyon.fr/cedric.vaillant/doku.php?id=cedpublications>),  
 les 5 publications que j'ai selectionnées, ce Rapport d'activite, le Projet de Recherche à l'adresse :  
<http://perso.ens-lyon.fr/cedric.vaillant/doku.php?id=divers>).

## 2 ACTIVITÉS DE RECHERCHE : OBJECTIFS GÉNÉRAUX

L'ADN chromosomique des cellules eucaryotes est fortement condensé au sein d'un complexe nucléo-protéique, la chromatine. Le premier niveau de compaction, le nucléosome, correspond à un enroulement de 146 paires de bases autour d'un octamère d'histone. L'arrangement linéaire de ces nucléosomes le long de la chaîne ADN forme le chapelet nucléosomal ou "fibre de 10 nm". Cette fibre, notamment grâce à la fixation d'histones de liaison ou autre protéines architecturales pourrait se compacter en une fibre plus compacte, dite "fibre de 30 nm" et/ou s'organiser, à plus grande échelle, soit dans des phases denses de type "fondu de polymère" soit dans des structures plus ouvertes de type "blobs" ou "boucles". Par ailleurs que ce soit au niveau de l'ADN, avec la méthylation, ou au niveau des histones, avec les modifications covalentes des queues ou l'insertion de variants, la chromatine se caractérise localement par une signature biochimique. Or, ces marqueurs biochimiques sont impliqués soit directement dans la structuration de la fibre (par exemple en modulant la stabilité des nucléosomes, ou l'interaction entre nucléosomes..) soit dans le recrutement de facteurs auxiliaires régulateurs de la chromatine comme les facteurs de remodelage. Il apparaît ainsi, qu'aussi bien l'organisation spatiale que la composition biochimique de la chromatine, en modulant l'accessibilité des différents complexes enzymatiques à leurs sites nucléiques joue un rôle fondamental dans la régulation du programme transcriptionnel (quels genes actifs et quand?) des cellules : à temps court, dans le cas de la réponse au stress et à temps long, dans le cas de la spécification et maintenance (hérédité) d'un type cellulaire au cours du développement ou lors de maladie comme le cancer.

Mon objectif est d'identifier et de modéliser par des approches de physique statistique, de l'équilibre et du hors-équilibre, les mécanismes de régulation "chromatinienne" de la transcription : comment la structure et la dynamique de la chromatine à toute échelle est-elle impliquée dans la dynamique d'activation/répression des gènes? Quels sont les mécanismes épigénétiques chromatinien à l'origine de la plasticité cellulaire, c'est à dire les mécanismes qui puisse à partir d'un même génotype conduire à différents phénotypes stables et héréditaires? Dans quelle mesure la séquence génomique et l'organisation du génome contribue à cette régulation "épigénétique"?

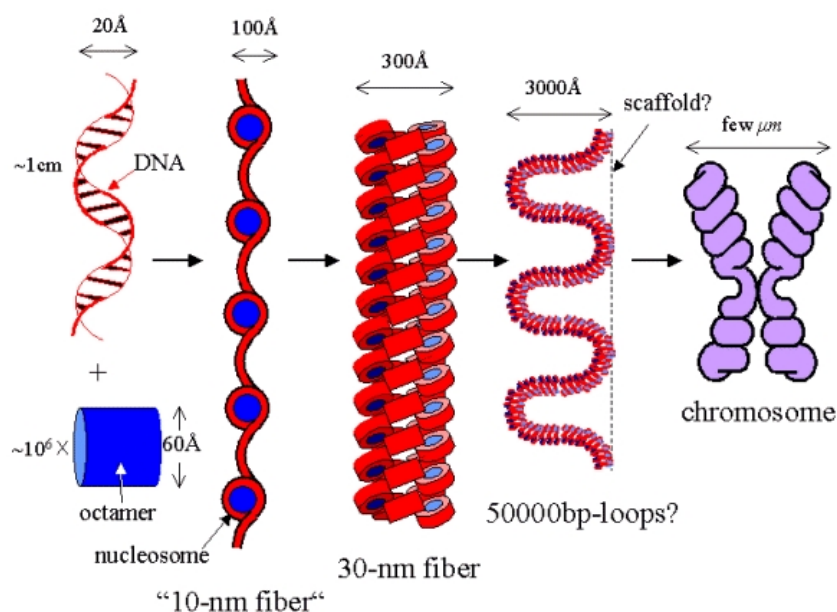


FIGURE 3: **Chromatin organization.** Hierarchical multi-scale organization of chromatin in eukaryotic nuclei, from DNA to metaphasic chromosome.

## 3 CHROMATINE ET ORGANISATION DU GÉNOME

Je suis entré au CNRS en Septembre 2006, au sein du Laboratoire Joliot Curie (Directeur Philippe Bouvet) et du Laboratoire de Physique (Directeur Jean-François Pinton) dans le groupe "Chromatine et génome" (Responsable d'équipe Françoise Argoul). Conformément au projet de recherche proposé lors de ma candidature, j'ai depuis approfondi et étendu mes travaux initiés en thèse et poursuivi durant mes années de recherche post-doctorale (à l'EPFL puis à Evry) concernant le rôle de la séquence génomique sur les propriétés de conformation et d'organisation de l'ADN au sein des cellules eucaryotes.

L'ensemble de ces travaux (depuis ma thèse) a fait l'objet de mon manuscrit de HDR qu'on peut trouver à l'adresse suivante : <http://perso.ens-lyon.fr/cedric.vaillant/doku.php?id=divers>.

**3.1 Effet de séquence sur l'élasticité de chaînes ADN :** J'ai participé étroitement aux travaux théoriques menés par Julien Moukhtar dans le cadre de sa thèse dirigée par Alain Arneodo sur l'effet de la séquence génomique sur les propriétés élastiques de chaînes ADN nues, confinées à 2D, travail prolongeant mes recherches de thèse (1). Notre collaboration avec des expérimentateurs au sein du LJC (Groupe de Françoise Argoul) a notamment permis de confirmer nos prédictions théoriques quant au rôle de la séquence et notamment des corrélations à longue portée dans les propriétés de conformations de longues chaînes ADN (2).

**3.2 Positionnement des nucléosomes :** Mes recherches principales se sont portées sur l'étude du positionnement des nucléosomes le long des génomes, travail que j'avais entamé durant mes post-doctorats successifs. Par la mise en place d'un modèle énergétique basé sur les propriétés élastiques de l'ADN, j'ai pu montrer théoriquement (3), en collaboration avec V. Miele (LBBE, UCB Lyon I), Thierry Grange (Institut Jacques Monod, CNRS, Paris 7) et C. Thermes (CGM, CNRS, Gif/Yvette) que les promoteurs des gènes de la levure et de la drosophile correspondaient à des régions énergétiquement défavorables aux nucléosomes et ce en accord avec les récentes données expérimentales mesurant la densité locale en nucléosome le long du chromosome III de la levure (Yuan *et al.*, *Science*, **309**, 626-630 (2005)). Ce modèle montrait que la séquence en modulant l'énergie de formation des nucléosomes participait de façon significative un peu partout au positionnement des nucléosomes.

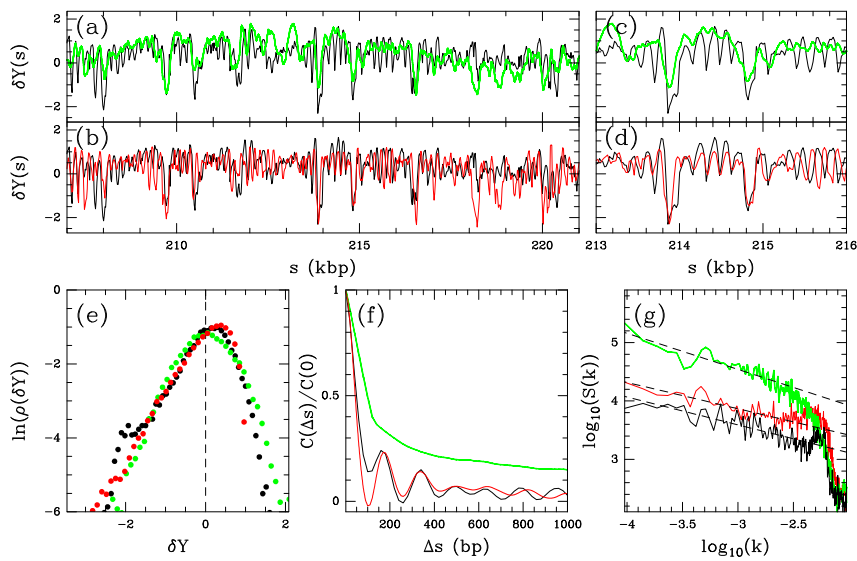


FIGURE 4: Statistical analysis and modelling of Yuan *et al.* nucléosome positioning data (Yuan *et al.*, *Science*, **309**, 626-630 (2005)). (a-d) The experimental Occupancy Landscape :  $\delta Y(s) = Y(s) - \bar{Y}$  (black), the theoretical negative Energy Landscape :  $-\beta\delta F_{Pnuc}(s, 125)$  (green) for a wrapping length  $l = 125$  bp, and the numerical Monte Carlo Occupancy Landscape :  $\delta \ln \rho(s)$  (red) obtained from Monte Carlo simulations (see text). (e) Pdfs, (f) two-point correlation function  $C(\Delta s)$ , (g) power spectrum of  $\delta Y(s)$  (black),  $-\beta\delta F_{Pnuc}(s, 125)$  (green) and  $\delta \ln \rho(s)$  (red). In (g) the dashed lines correspond to power-law scaling exponents  $\nu = 0.60, 0.48$  and  $0.45$  corresponding to  $H = 0.80, 0.74$  and  $0.73$  Long Range Correlation properties respectively. From (4).

Cependant, lorsqu'on étudie les données expérimentales on s'aperçoit qu'il existe des régions le long du génome qui présentent une organisation périodique des nucléosomes, avec une période d'environ  $167$ pb (Figure 4(c)), et ce souvent au voisinage d'une région de déplétion. Or cette organisation périodique n'est pas prédite par ce modèle énergétique. En suivant l'idée de Kornberget Stryer (*Nucleic Acids Res.*, **16**, 6677-6690) qui avait proposé que ces nucléosomes périodiques n'étaient pas en fait positionnés par des séquences favorables ("puits énergétiques") disposées périodiquement le long du génome mais par un mécanisme d'ordonnement statistique dû à la présence de zones d'exclusion. Pour tester cette hypothèse nous avons introduit un modèle thermodynamique simple en considérant la chromatine comme un fluide de nucléosomes pouvant s'adsorber sur et se desorber de l'ADN ; en considérant (i) le fluide en équilibre avec un bain thermique, et un réservoir de "nucléosome" (à un potentiel chimique donné), (ii) un potentiel d'adsorption non uniforme dépendant de la séquence (l'énergie de formation du nucléosome) et (iii) une interaction de type "sphère dure" entre nucléosomes, il m'a été possible grâce à des simulations de Monte Carlo de calculer un profil de positionnement des nucléosomes le long du chromosome III de la levure que l'on a pu directement comparer au profil expérimental obtenu par Yuan *et al.* (*Science*, **309**, 626-630 (2005)) (4). La très bonne prédiction théorique obtenue (Figure 4) nous a donc permis de montrer que (1) la chromatine pouvait effectivement être modélisée par un modèle simple de fluide non uniforme à l'équilibre, (2) que l'inhomogénéité observée dans le profil expérimental à petite échelle pouvait être en

partie due à un effet de séquence : certaines zones de déplétion seraient effectivement dues à la présence de barrières énergétiques induites par des séquences défavorables et les périodicités observées notamment aux abords de ces régions déplétées en nucléosome seraient bien la conséquence d'une "cristallisation" entropique due au confinement induit par ces barrières énergétiques (5). Ces résultats ont aussi permis de montrer que les corrélations à longue portée observées expérimentalement à grande échelle dans le profil de positionnement était bel et bien le reflet des corrélations à longue portée présente dans la séquence, nous permettant par ailleurs de confirmer notre conjecture (faite durant ma thèse) quant au lien entre l'existence de corrélations à longue portée dans la séquence génomique et la présence de nucléosomes.

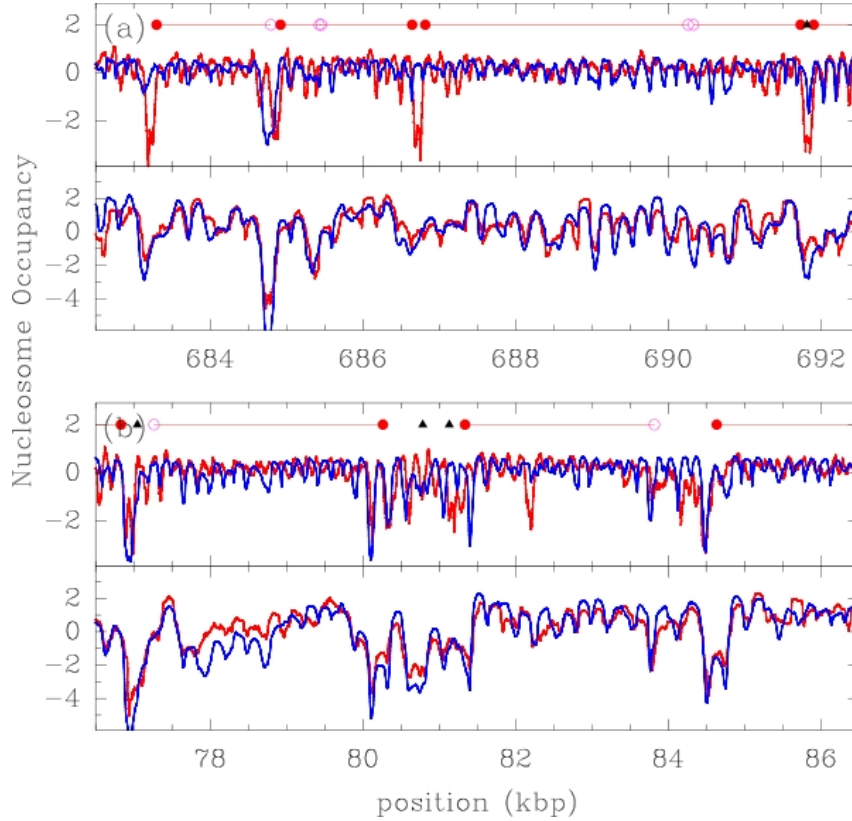


FIGURE 5: (a) Nucléosome occupancy ( $\log_2$ ) profile along the yeast chromosome II genome : *in vivo* (Lee *et al.*, *Nature Genetics*, **39**, 1235-1244 (2007)) (top red curve), *in vitro* (Kaplan *et al.*, *Nature*, **458**, 362-366 (2009)) (bottom red curve), theoretical profiles at high density (1 *nucl./200 bp*) (top blue curve) and low density (1 *nucl./500 bp*) (bottom blue curve). (b) Same as in (a) for chromosome XII. In (a) and (b) are indicated the positions of Transcription Start Site (red dots), Transcription Termination Site (red circles) and of Transcription Factor binding sites (black triangles). From (6).

Suite à cette étude, j'ai eu l'opportunité de diriger la thèse de Guillaume Chevereau (7), dont l'objectif était d'étendre et approfondir nos résultats. Nous avons alors mis en place une modélisation basée sur la résolution analytico-numérique (Vanderlick *et al.*, *Phys. Rev. A*, **34**, 5130-5131 (1986)) de la relation exacte obtenue par Percus (Percus, *J. Stat. Phys.*, **15**, 505-511 (1976)) reliant la densité locale en nucléosomes au potentiel inhomogène d'adsorption en fonction du potentiel chimique et de la température. Cela nous a permis de pouvoir comparer nos résultats aux données expérimentales qui venaient d'être produites sur tout l'ensemble du génome de la levure (Lee *et al.*, *Nature Genetics*, **39**, 1235-1244 (2007)) mais également de l'homme et de *C. elegans*, et ainsi de procéder à une étude beaucoup plus exhaustive de la relation entre séquence et positionnement des nucléosomes pour un coup computationnel très faible. Sur le fond, ces nouvelles données ont permis de confirmer les résultats obtenus précédemment. La corrélation entre les données et nos prédictions sont effectivement très bonnes (Corrélation de Pearson  $r \sim 0.3$  sur tout l'ensemble du génome) (Figure 5) (5). Cependant on voit bien qu'une majorité du positionnement est encore assez mal prédite par notre modèle (Figure 5). Ainsi, si expérimentalement, on observe une grande majorité de promoteurs des gènes qui sont déplétés en nucléosome, seule une fraction, à savoir 30%, sont prédits par notre modèle; à la fin des gènes, on observe également une déplétion prédite pour 50% des cas par notre modèle. La question qui se posait était donc de savoir si notre modèle, notamment le modèle énergétique était suffisamment prédictif ou non, ou si cette différence était due à l'action extrinsèque de facteurs biologiques tels que les facteurs de transcription, enzymes de remodelages ou polymérases connus pour perturber le positionnement des nucléosomes. Or nos prédictions se trouvent en parfait accord avec les données récentes de positionnement obtenues sur des chromatines reconstituées

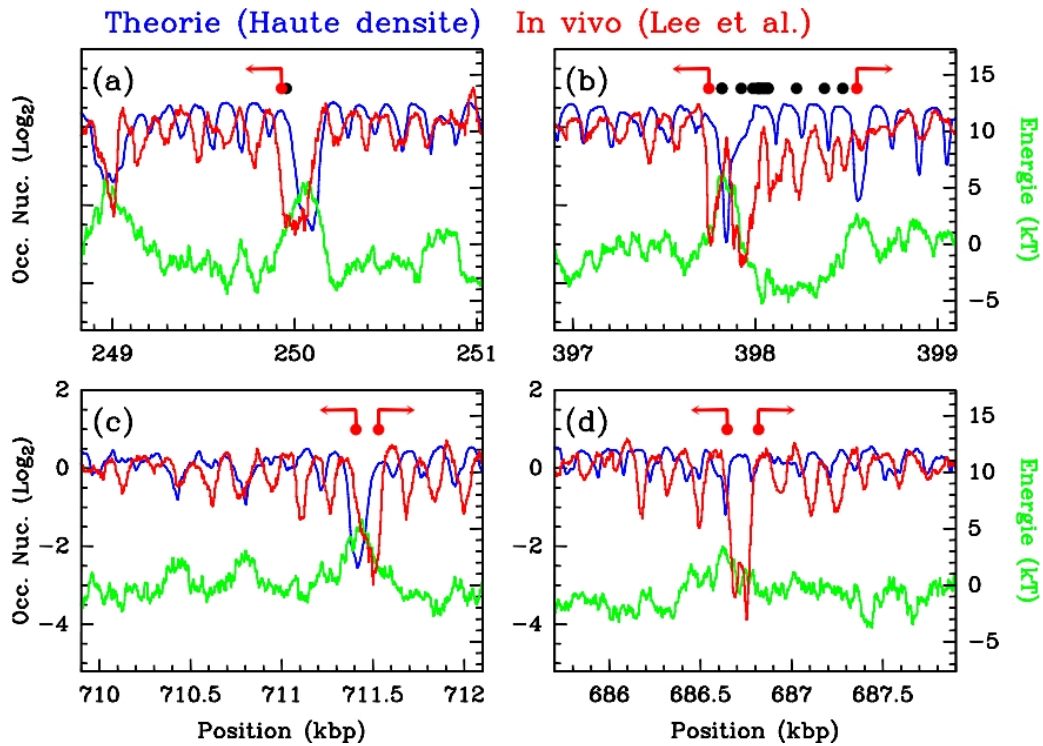


FIGURE 6: Des exemples de profils d'occupation théoriques (en bleu) similaires aux profils expérimentaux (8) (en rouge) à un remodelage globale près. Les gènes sont indiqués par les points (TSS) et flèches rouges. Les sites de fixation des facteurs de transcription sont marqués par les points noirs.

*in vitro* (6, 5) (dans ce cas le positionnement n'est dicté que par l'interaction ADN-histones et non pas par des facteurs extrinsèques), indiquant que notre modèle énergétique "intrinsèque" capte de façon extrêmement satisfaisante l'influence de la séquence génomique (Figure 5). Pour mieux rendre compte des profils *in vivo* il est donc indispensable de considérer l'action des facteurs extrinsèques.

Comparer nos modélisations à haute densité et les données *in vivo* peut nous renseigner sur les mécanismes extrinsèques qui agissent et modèlent *in fine* le chapelet nucléosomal. Si les profils expérimentaux *in vivo* ne concordent pas tout le temps avec la prédiction, c'est que d'autres facteurs que la séquence interviennent. Toutefois, il est très fréquent que les profils nucléosomaux générés par la séquence soient similaires au profil expérimental à un léger décalage près. La figure 6 présente quelques exemples, situés à proximité des gènes de la levure. Il suffit parfois de décaler légèrement ces profils pour obtenir un "fit" correct des données (figure 6 (a) et (c)). Dans d'autre cas, il faut parfois invoquer la suppression d'un nucléosome et un remodelage plus complexe (figure 6 (b) et (d)). Enfin, certains cas révèlent très bien la compétition entre les facteurs de transcription et les nucléosomes comme l'illustre la figure 6(b).

La mise en place de modèles de positionnement des nucléosomes avec "régulateurs extrinsèques" a été débuté (cf Projet en cours) et constitue un des objectifs majeurs de mon Projet de Recherche.

Un autre résultat important que nous avons obtenu concerne l'organisation de la chromatine à l'intérieur des gènes de la levure et la relation entre cette organisation et l'activité transcriptionnelle des gènes. Expérimentalement on observe entre le début et la fin des gènes une organisation très cristalline pour les petits gènes ( $< 1200 bp$ ) (Figure 7(b)). Selon la taille du gène on observe ainsi successivement deux, trois ... jusqu'à huit nucléosomes très bien positionnés. Pour les grands gènes cette cristallisation s'atténue lorsqu'on s'éloigne des deux extrémités du gène pour finalement disparaître au milieu où le positionnement devient "flou" (Figure 7(a)). Grâce à un modèle simple (Figure 6(a-c)) où on ajoute à notre potentiel dépendant de la séquence des barrières énergétique au début et en fin pour rendre compte de l'effet dominant de facteurs extrinsèque, on a réussi à très bien modéliser cette organisation particulière de la chromatine intra-génique (Figure 7) : en particulier, la forte cristallisation pour les petits gènes provient du fort confinement induit par les deux barrières aux deux extrémités conformément au principe de "positionnement statistique" (Figure 8(a) et (c)) (9, 10, 5). Notre modèle rend également compte du profil apparemment non périodique ("flou") observé sur certains petits gènes (Figure 7) : ces gènes, qu'on a appelés "bistables", ont une taille qui autorise la cristallisation de  $n$  ou de  $n + 1$  nucléosomes. Ainsi le profil observé résulte de la superposition de deux états cristallins, l'état dense à  $n + 1$  nucléosomes et l'état dilué à  $n$  nucléosome; le profil apparait donc comme étant "flou" (Figure 8(b)). La mesure expérimentale étant faite sur un grand échantillon de cellules, cela signifie que certaines cellules ont ce gène dans l'état cristallin dilué et d'autres dans l'état cristallin dense. Enfin, nous avons montré que le taux de transcription des petits gènes cristallins augmentait avec la période nucléosomale (Figures 8(f) et 9A). Conformément à ce principe, nous avons montré de plus que les gènes bistables présentaient bel et bien

une plus grande variabilité transcriptionnelle, ceux-ci pouvant facilement transiter d'un état plus dense et plus actif à un état plus dilué et moins actif (Figures 8(f) et 9C). C'est un résultat très original ouvrant de nombreuses perspectives de recherches et de collaborations futures. Globalement tous ces résultats sont le fruit d'une collaboration étroite au sein de l'équipe d'Alain Arnéodo (Guillaume Chevereau, Benjamin Audit et Leonor Palmeira) et aussi avec l'équipe de Claude Thermes (CGM, Gif/Yvette).

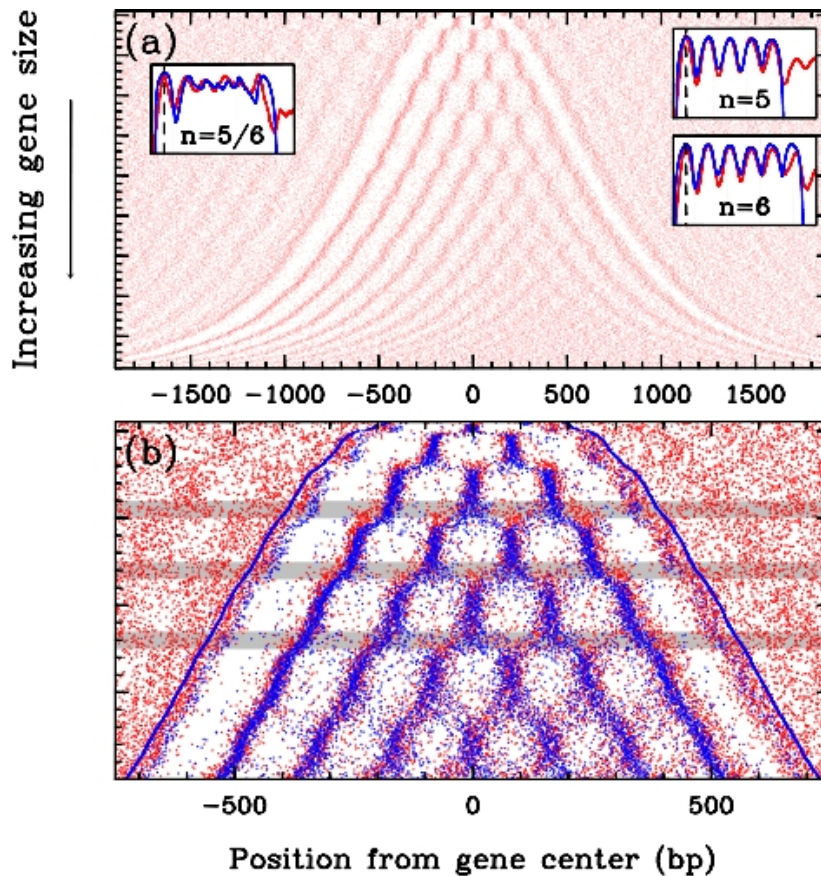


FIGURE 7: (a) 2D map of local minima (red) of the experimental *in vivo* nucleosome occupancy profile at yeast genes (Lee *et al.*, *Nature Genetics*, **39**, 1235-1244 (2007)); genes are ordered vertically by the distance  $L$  between the +1 and -1 nucleosomes. Insets : mean experimental (red) and one individual theoretical (blue) nucleosome occupancy profiles for “crystal” genes harboring 5 nucleosomes (right, top), 6 nucleosomes (right, bottom) and the bi-stable genes with 5/6 nucleosomes. (b) Zoom on the first 2000 genes in (a); on the top of the experimental data (red) are superimposed the predictions of our physical modeling (blue); horizontal grey-shaded bands correspond to some “bi-stable”  $L$ -domains. From (11)

Récemment, nous avons de plus confirmé, grâce à des mesures par microscopie AFM du positionnement de nucléosomes reconstitués sur des petites séquences ADN, que le positionnement des nucléosomes pouvait effectivement résulter du confinement induit par l'exclusion de séquence défavorables (Figure 10). C'est une preuve expérimentale de plus du rôle joué par la séquence et du principe d'ordonnement statistique, et qui a impliqué une grande partie du Laboratoire Joliot-Curie, pour la théorie (groupe d'Alain Arnéodo), la biophysique (groupe de Françoise Argoul) et la biologie (groupe Philippe Bouvet).



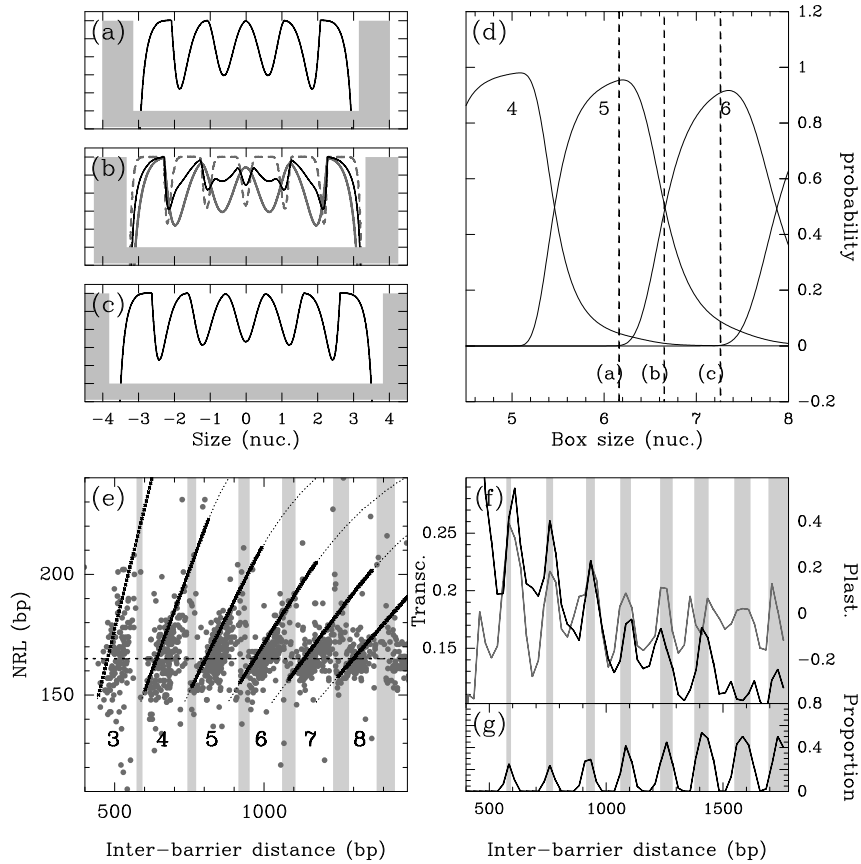


FIGURE 8: Theoretical probability of nucléosome occupancy at each point of a box bordered by two infinite walls mimicking excluding barriers at gene extremities : (a) box large enough to shelter  $n = 5$  nucléosomes (black); (b) larger box where the two  $n = 5$  and  $6$  configurations (grey) are possible; the weighted average of the  $5$  and  $6$  nucléosome crystal-like profiles yields a *fuzzy*-looking profile (black); (c) larger box where  $6$  nucléosomes can be inserted (black). (d) Probability of crystal configurations with a fixed number  $n$  of nucléosomes *vs.* the box size. Vertical dashed lines correspond to the inter-barrier distances used in (a), (b) and (c) respectively. (e) Nucléosome Repeat Length (NRL) dependency on the box size : thin black dotted lines correspond to a fixed number  $n$  of nucléosomes and the thick black dotted lines to the NRL at a fixed nucléosome density ( $\sim 85\%$ ); grey dots correspond to individual crystal gene values. Vertical grey-shaded bands correspond to the experimental “bi-stable”  $\mathcal{L}$ -domains. (f) Average transcription rate (black) and transcriptional plasticity (grey) defined as the square  $\log_2$  expression ratio from a large number of microarray experiments over a  $50 \text{ bp}$  sliding window. (g) Proportion of bistable genes. From (9).

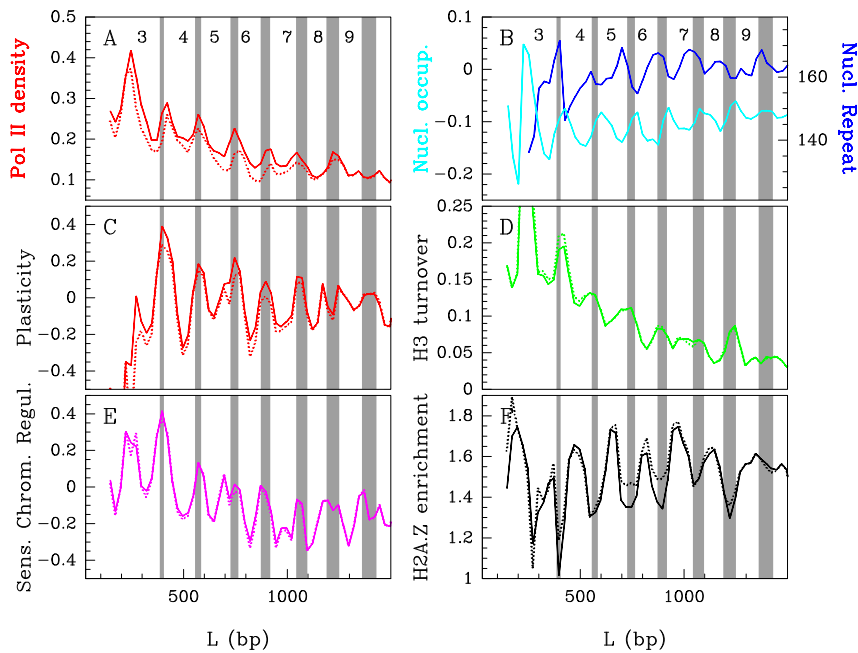


FIGURE 9: Bi-stable nucleosome organization controls gene expression. Sliding window (50 bp) analysis of average (A) transcription rate estimated by Pol II density, (B) nucleosome repeat length (dark blue) and intra-genic mean nucleosome occupancy (light blue), (C) transcriptional plasticity, (D) H3 turnover rate, (E) sensitivity to chromatin regulators disruption and (F) H2A.Z occupancy as a function of the distance  $L$ . The vertical grey bands define bi-stability domains. The dotted curves correspond to the results obtained when excluding the 175 (/4554) ribosomal protein genes from the analysis; no significant changes are observed. From (11).

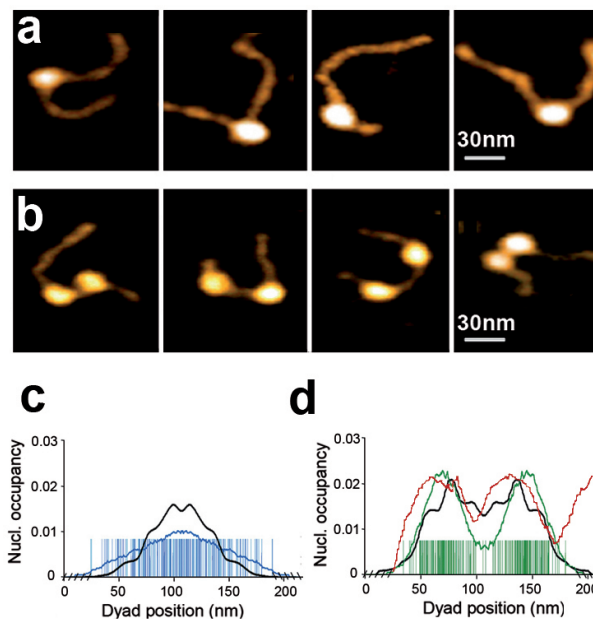


FIGURE 10: **Excluding energy barriers condition the positioning of nucleosome consistently with equilibrium statistical ordering principles (12).** AFM imaging in liquid of mono-nucleosomes (a) and di-nucleosomes (b) along a yeast (chr. 7) DNA fragment ( $L = 595$  bp) containing the gene YRG105W. (c) Statistical analysis of mono-nucleosome positioning ( $N = 113$  molecules); blue bars correspond to experimentally detected dyad locations and to their symmetrical position with respect to the center of the fragment. The experimental nucleosome occupancy profile (blue curve) is compared to the theoretical predictions of our physical modeling after symmetrization (black curve). (d) Statistical analysis of di-nucleosome positioning ( $N = 62$ ); for each image, green bars were drawn at both the 2 dyad positions and their symmetric locations with respect to the center of the fragment. The experimental nucleosome occupancy probability profile (green curve) is compared to the symmetrized theoretical profile (black curve) and to the *in vivo* occupancy profile obtained by Lee et al. (*Nature Genetics*, **39**, 1235-1244 (2007)) (red curve).

## 4 ORIGINES DE RÉPLICATION ET NUCLÉOSOMES

Par ailleurs, nous avons pu montrer qu’avec ces prédictions *in silico* de la chromatine *in vitro* il était possible d’obtenir des résultats très intéressants sur le lien entre l’organisation chromatinienne et la régulation de la réplication. Ainsi nous avons montré (13) qu’autour des origines de réplication prédites par les groupes d’Alain Arnéodo (Benjamin Audit et Lamia Zaghoul) et Claude Thermes (CGM, Gif/Yvette), la chromatine présentait un fort enrichissement en régions de déplétion en nucléosome reflétant une structure chromatinienne ouverte favorable à l’activation de ces origines (Figure 11). Notons que notre modèle de positionnement de nucléosome reproduit de façon très satisfaisante les données expérimentales (5).

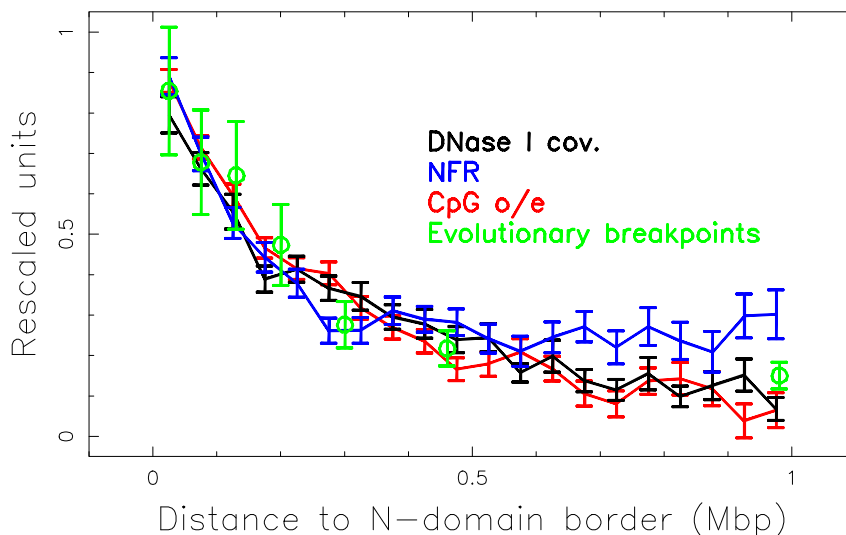


FIGURE 11: Mean profiles of DNase I HS sites coverage (black), NFR density (GC content < 41%, blue) and CpG observed/expected (red) as a function of the distance to the closest N-domain border after masking CGIs and genes extended by 2 kb at both extremities.

Cette étude et celle reportée dans (14), en accord avec d’autres études récentes, ont permis de préciser les “déterminants” chromatinien des origines de réplifications, des levures *S. cerevisiae* & *S. Pombe* aux mammifères en passant par la mouche *D. melanogaster* (15, 16, 17) : les origines de réplication sont des régions qui sont plutôt accessibles, enrichies en “trous” de nucléosomes, en promoteurs de gènes actifs (gènes de ménages), en insulateurs de type CTCF et correspondent en générale aux frontières des domaines topologiques et épigénétiques identifiée par les récentes expériences de “Capture de conformation” (18, 14, 19). La réplication semble donc fortement couplée à l’organisation spatiale (3D) et épigénétique (1D) de la chromatine et une question encore ouverte est de savoir dans quelle mesure celle-ci conditionne cette organisation 1D et 3D (17) ? Un objectif de mon projet de recherche sera notamment d’étudier comment la réplication participe au contrôle épigénétiques des domaines hétérochromatiniens.

## 5 CHROMATINE ET INSERTION VIRALE

Dans une première collaboration avec Marc Lavigne (Institut Pasteur) et Vincent Parissi (Université de Bordeaux 2) (20) nous avons étudié dans quelle mesure le chapelet nucléosomal *in vitro* pouvait influencer l’activité de l’intégrase du virus HIV. Nous avons montré, expériences *in vitro* et modélisation à l’appui, que l’intégration est favorisée dans des chapelets plus labiles, avec des nucléosomes plus faciles à déplacer (ceci n’excluant pas le fait que l’intégration puisse *in fine* être favorisée au niveau d’un nucléosome). Comme le montre la figure 12, nos prédictions du chapelet nucléosomal indiquent que les sites d’intégration de HIV dans des cellules humaines (21) sont effectivement plutôt localisées dans un chapelet nucléosomal mais globalement moins dense qu’en moyenne. Plus récemment le groupe de V. Parissi a étendu cette étude d’autres intégrases retrovirales et a montré, toujours par un système d’intégration sur une chromatine reconstituée que l’effet du nucléosome dépendait de l’intégrase : les intégrases de PLV (et MLV) ont plutôt tendance à intégrer dans des nucléosomes stables tandis que les intégrases d’HIV et ASV intègrent mieux leur ADN dans des régions moins denses. L’interprétation avancée est que c’est la structure du complexe Intégrase+ADN viral qui contribue à cette sélectivité. Pour compléter cette étude, une comparaison avec les données d’intégration *in vivo* indique en effet une différence entre les profils des nucléosomes autour des sites d’intégration de HIV et de MLV (cf Fig 13) (occupation globale plus faible chez HIV que MLV).

Ces travaux font l'objet d'un article en cours de révision :

MS Benleulmi, J. Matysiak, D.Rodrigo Henriquez, C. Vaillant, P. Lesbats, C. Calmels, M. Naughtin, O. Leon, A. Marie Skalka, M. Ruff, M. Lavigne, ML Andreola and V. Parissi. Intasome architecture and chromatin density modulate retroviral integration into nucleosome. *Retrovirology* In revision.

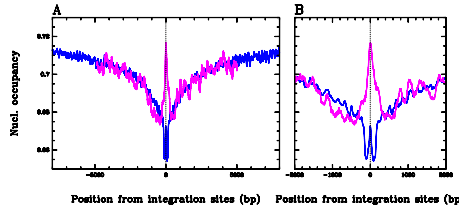


FIGURE 12: Profils “moyens” théoriques (bleu) et expérimentaux (violet) d’occupation en nucléosome (22) autour de sites d’intégration de HIV dans des cellules humaines (23)

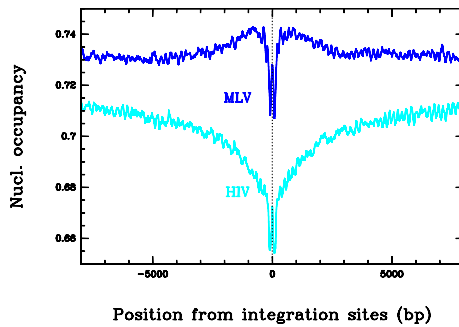


FIGURE 13: Profils théoriques “moyens” d’occupation en nucléosome autour de sites d’intégration de HIV et MLV (23)

En parallèle, l'équipe de Marc Lavigne (alors au LJC a l'ENS d Lyon et actuellement a l'Institut Pasteur à Paris) a procédé à des expériences d'intégration sur 4 séquences “naturelles” où un site d'intégration avait été observé par (21) et pour trois conditions différentes : intégration sur l'ADN nu et intégration sur une chromatine reconstituée *in vitro* à faible densité et à plus haute densité. La position des nucléosomes ainsi que les positions des sites d'intégration ont été cartographié par séquençage haut-débit. Ce qui apparait clairement c'est que les sites d'intégration sont modifiés par la présence de nucléosomes. Une analyse spectrale nous révèle par ailleurs une périodicité de  $10bp$  dans les sites d'intégration uniquement lorsque celle-ci se fait sur l'ADN chromatinisé, ce qui est clairement le signe d'une intégration dans un nucléosome (le grand sillon n'étant “accessible” que toutes les  $10 pb$ ).

Ces résultats ont fait l'objet d'un article prochainement soumis à *Retrovirology* :

M. Naughtin, Z. Haftek-Terreau, J. Xavier, S. Meyer, M. Silvain, Y. Jaszczyszyn, N. Levy, V. Miele, MS. Benleulmi, M. Ruff, V. Parissi, C. Vaillant and M. Lavigne. “HIV-1 integrase selectivity depends on nucleosome positions and physical features of the target DNA”.

Dans ces travaux ma contribution est l'élaboration des prédictions théoriques de l'occupation en nucléosome le long des séquences. Jusqu'à présent nous avons utilisé le modèle simple qui ne dépend que de la séquence génomique, ce qui est justifié par le fait que les expériences sont faites sur des chromatines reconstituées. Pour aller vers l'intégration *in vivo*, il faudra recourir à des modèles étendus comme ceux présentés ci-après :

## 6 POSITIONNEMENT DES NUCLEOSOMES : VERS LE *IN VIVO*

**6.1 Modélisation des effets “extrinsèques”** Une première étape fondamentale vers la modélisation du chapelet nucleosomal “in vivo” est d'intégrer l'action des facteurs de transcription/insulateurs et des facteurs de remodelages (24, 25). Ce sont deux composantes essentielles de la régulation chromatinienne de la transcription que ce soit pour reprimer ou activer ; leurs rôles est essentiel . En effet, c'est souvent par le biais d'un facteur de transcription qui se lie à un site donné qu'un facteur de remodelage spécifique est recruté de façon stable à un locus donné. Leur modes d'action est divers : soit éjecter (ex RSC, Hartley (26, 27, 28)), soit repositionner (ex Iswi2 (29), NoRC (30)) soit équirepartir (31) ou randomiser (24).

**6.1.1 Facteurs de remodelage : modèles effectifs Forces & Fluidification** Une première approche (dite “effective”) est de considérer que les facteurs de remodelage modifient le positionnement “intrinsèque” du nucleosome pour aboutir à une distribution “stationnaire”. Donc on peut toujours considérer le même modèle thermodynamique que précédemment mais avec un profil énergétique effectif qui rendrait compte de l’effet des facteurs de remodelage. Pour bien modéliser l’organisation intra-génique (11) chez *S. cerevisiae* nous avons déjà dû introduire l’effet des facteurs de remodelage : (1) les barrières énergétiques en début et fin de gènes ont été considérées non plus “verticales” mais “en rampe” c’est à dire résultants d’une force localisée au niveau du début et de la fin des gènes (2) à l’intérieur des gènes l’effet de la séquence a été atténué ce qui correspond à une fluidification induite par les remodeleurs. Cette idée de fluidification est intéressante car elle souligne le fait que les facteurs de remodelages agissent en partie pour compenser l’effet du positionnement intrinsèque : il y aurait une activité “ciblée” à certains sites pour exécuter une tâche spécifique (e.g. appliquer une force) et une activité “dispersée” pour “élever” la température effective du chapelet nucleosomal. Je suis en train de mener une étude théorique de ce mécanisme de fluidisation induit par les remodeleurs chez *S. cerevisiae* et *S. Pombe*. L’étude plus fine et détaillée de l’effet des facteurs de remodelage sera entreprise dans le cadre des modèles hétérochomatiniens, présentés dans mon Projet de Recherche.

**6.1.2 Modèle de compétition nucléosome vs. facteurs de transcriptions** L’objectif initial ici était (i) de modéliser la compétition entre les nucléosomes et les facteurs de transcription et (ii) de pouvoir par ailleurs introduire des interactions entre nucléosomes autres que le volume-exclu, interactions qui pourraient rendre de manière effective de l’effet de certains facteurs de remodelages (ceux qui équirepartissent). Il s’agit donc toujours de modéliser un fluide 1D mais multi-composants et avec des interactions à courtes portées quelconques. Au cours de son stage de M2 au printemps 2011, Philippe Robert a mis au point une méthode algorithmique qui permet de modéliser ce type de fluide très efficacement. Philippe a par ailleurs développé une application logicielle graphique d’utilisation simple pour calculer et observer les profils de densités le long de séquences génomiques (avec deux types de composants, nucléosome et facteur de transcription, par exemple)<sup>14</sup>. Récemment, nous avons encore amélioré le modèle du chapelet nucleosomal en intégrant la possibilité de fluctuation de la longueur de complexation de l’ADN autour de l’octamère d’histones (auparavant on la fixait à  $l = 146$  bases, qui n’est en fait qu’une valeur moyenne). Cette possibilité de “respiration” de l’ADN sur le nucléosome le rend plus accessible aux facteurs de transcription et donc peut jouer un rôle important dans la régulation transcriptionnelle. Cette extension de notre modèle a été mise en place via l’algorithme de P. Robert et intégrée dans son application logicielle. Nous sommes en train de rédiger deux articles à ce sujet, une note “technique” et un article de recherche. Une application Web et un service de téléchargement du logiciel est disponible à : <http://joliot-curie.ens-lyon.fr/partjoliot-dev/>

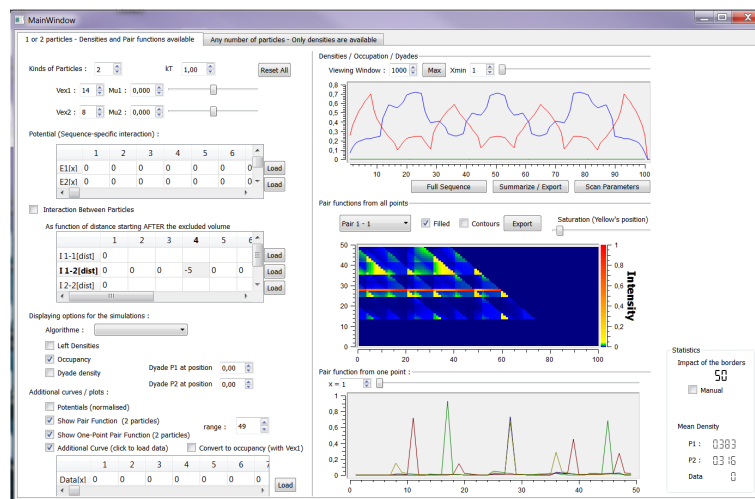


FIGURE 14: Capture d’écran du logiciel de calcul de la densité d’un fluide 1D multi-composant développé par P. Robert.

**6.2 Les mic MAC de la Chromatine** Je travaille aussi dans le cadre d’un projet mené par M. Betermier (CGM, Gif sur Yvette, <http://www.cgm.cnrs-gif.fr/spip.php?article119>) et S. Duharcourt (Institut Jacques-Monot, <http://www.ijm.fr/recherche/equipes/regulation-epigenetique/>) en collaboration avec Claude Thermes et Chunlong Chen (CGM, Gif) et qui vise à comprendre dans quelle mesure le chapelet nucleosomal conditionne le processus d’excision de séquences ADN lors de la création du noyau somatique MAC à partir du noyau germinale mic, dans l’organisme unicellulaire modèle *Paramecium tetraurelia*. La comparaison entre nos prédictions théoriques et les profils expérimentaux (obtenus par expériences de “MNase-seq” sur le noyau Mac) donne des résultats très corrects (Fig. 15). Cependant,

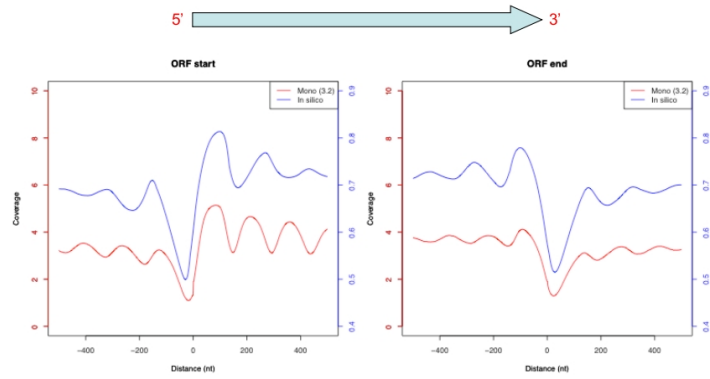


FIGURE 15: Position des nucléosomes en début et fin de gènes : Comparaison entre prédiction (blue) et expériences (rouge) chez la Paramécie (Macro-noyau)

et comme pour la levure et les autres organismes, une meilleure prédiction nécessite de tenir compte des facteurs extrinsèques et notamment l'action de remodelleurs. La période nucléosomale mesurée est en effet très courte ( $150pb$ ) alors que celle prédite est de  $\sim 175pb$ . L'application de notre "modèle étendu" à cet organisme est en cours.

**6.3 Mécano-transduction et chromatine chez les plantes** Je participe au projet "Approche multi-échelle de la dynamique de régulation des gènes mécano-sensibles chez les cellules souches végétales", d'abord dans le cadre d'un financement par la "Mission pour l'interdisciplinarité CNRS reporting PEPS Bio math Info 2012 2013", en collaboration avec A.Boudaoud, O. Hamant (du Laboratoire de Recherche en Développement de Plantes et Laboratoire Joliot-Curie) et Benjamin Audit (Laboratoire de Physique) puis dans le cadre de la boure ERC d'O. Hamant. Le projet en est actuellement au stade de la "transcriptomique" et "épigénomique" : des expériences de perturbation mécanique dans des souches sauvages et variantes d'*Arabidopsis thaliana* ont été menées et l'analyse par RNA-seq et Chip-seq est en cours. Elle devra nous révéler les jeux de gènes sur ou sous exprimés et le lien avec les éventuelles variations épigénomiques. La localisation de ces gènes dans les domaines épigénétiques (comme ceux obtenus par Roudier *et al.* (32)) nous permettra de préciser le mode épigénétique de répression/activation de ces gènes (Hétérochromatine Polycomb ?). Ce projet s'inscrit dans mon projet de recherche, puisque ma contribution ici sera la modélisation des mécanismes épigénétiques "chromatiniens" de répression/dérépression stable de ces gènes (cf Projet de Recherche) au cours du développement. Ce projet interroge aussi la manière dont des stimuli extérieurs (en l'occurrence ici mécaniques) transitoires peuvent être in fine traduits en modifications transcriptionnelles stables et héréditaires via notamment des modifications épigénomiques.

## 7 REPLIEMENT DE L'ÉPIGÉNOME

Mon activité de recherche pour l'année 2014 a tourné essentiellement autour de la modélisation du repliement à grande échelle de la fibre de chromatine. C'est un travail théorique conduit avec Daniel Jost depuis 2013. Nous co-encadrons la thèse de Noelle Haddad sur le sujet depuis Septembre 2013.

**7.1 Objectifs** Les chromosomes eucaryotes sont composés de deux types de domaines épigénétiques structurels et fonctionnels chromatiniens : l'euchromatine, ouverte et généralement accessible, où l'on retrouve la plupart des gènes actifs et l'hétérochromatine, plus fortement condensée, et repressive. L'étude des profils de distributions de protéines régulatrices et des marqueurs épigénétiques obtenus par les récentes méthodes haut-débit ont permis de préciser un peu plus cette compartimentation 1D des génomes (35) : de la drosophile (33), Figure 16) à l'homme (36), en passant par les plantes (32) et *C. elegans* (37) on peut distinguer quatre types de domaines chromatiniens : l'euchromatine qui contient les gènes actifs (gènes toujours actifs, chromatine "jaune"/gènes spécifiques à certains tissus, "rouge"), l'hétérochromatine constitutive de type "HP1/H3K9me" plutôt enrichie en éléments transposables et en séquences répétées, une hétérochromatine facultative (dite "Polycomb", bleue) enrichie en gènes impliqués dans la régulation de la différenciation et du développement et une hétérochromatine "ultra-repressive" (noire) enrichie en gènes qui ne sont exprimés que dans très peu de tissus. L'hétérochromatine constitutive est présente de façon permanente, tandis que l'hétérochromatine facultative permute entre un état hétérochromatinien et un état euchromatinien selon le contexte biologique. D'un point de vue fonctionnel, l'hétérochromatine contrôle plusieurs aspects fondamentaux du fonctionnement nucléaire : (i) assemblage du kinetochore (ii) cohésion des chromatides sœurs assurant ainsi la bonne ségrégation durant la division cellulaire (iii) recombinaison : inhibition de toute recombinaison inopinée au niveau des séquences répétées garantissant ainsi une stabilité génomique (iv) expression des gènes : répression de la transcription des séquences sous-jacentes et voisines (v) activation/repression de certaines interactions à longue distance

### Compartmentation 1D

### Compartmentation 3D

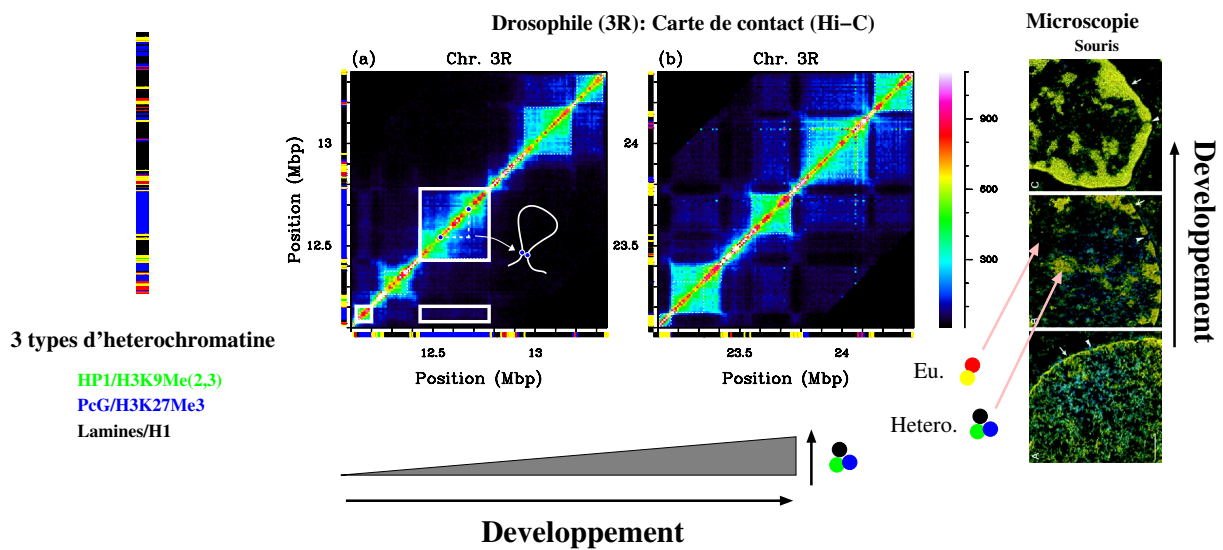


FIGURE 16: (gauche) Schématisation de la partition du génome en domaines épigénétiques chromatiniens. Chaque type de domaine dénoté par une couleur, est caractérisé par une composition spécifique en protéines architecturale, marques épigénétiques et organisation du chapelet de nucléosome. Les domaines verts, bleus et noirs sont des domaines heterochromatiniens, plutôt répressifs transcriptionnellement (avec des modes operatoires de repression differents selon le type de domain); les domaines rouges et jaunes sont enrichie en gènes actifs (Tiré de (33)). (centre) Carte de contact d'une portion de 2 Mega Bases du chromosome 3R de la drosophile : Fréquence de contact entre sites génomiques mesurée par la méthode de "Capture de conformation" haut-débit dite "Hi-C" (18) (droite) microscopie electronique du noyau de cellules de souris à différents stades du développement (34) (de cellules pluripotentes à des cellules différenciées); la coloration jaune indique la présence de marqueurs hétérochromatiniens.

impliquées notamment dans la régulation du développement. L'hétérochromatine est en effet associée à la différenciation cellulaire, même dans les organismes unicellulaires où elle contrôle le type cellulaire et la reproduction sexuée. Dans les organismes multicellulaires l'hétérochromatine est impliquée dans la maintenance de l'identité cellulaire au cours du développement. D'un point de vue moléculaire, l'hétérochromatine se caractérise par des signatures biochimiques et structurales particulières (38) : l'hypo-acétylation des histones, la méthylation spécifique H3K9me (resp. H3K27me), l'association avec des protéines structurales de la famille HP1 (resp. PcG) (ou H1/lamine) et par une distribution des nucléosomes très périodique. Cette organisation épigénétique du génome est dynamique : en effet, le développement des cellules germinales et la différenciation des cellules souches en cellules somatiques sont associés à une reprogrammation de ces domaines chromatiniques (39) ; la reprogrammation pathologique de ces domaines contribue souvent à la cancérogénèse ou à d'autres pathologies (40). La régulation de ces domaines est confrontée à un double challenge : à la fois permettre une plasticité chromatinienne pour garantir une plasticité phénotypique au cours du développement et assurer une robustesse de ces états chromatiniques pour maintenir l'identité phénotypique des cellules dans un environnement fluctuant (41). De façon assez remarquable, les récentes expériences de "Capture de conformation de la chromatine" (42, 18) et de microscopie (43, 44) ont permis de montrer que cette partition 1D du génome se retrouve également dans l'organisation spatiale de la chromatine (45, 35) : les domaines épigénétiques (1D) s'organisent en effet en domaines topologiques (micro-phases 3D ou "TAD" pour "Topological Associated Domains") caractérisés par des interactions spatiales essentiellement "intra-domaine" : les domaines adjacents apparaissent ainsi "isolés" les uns par rapports aux autres. Les domaines topologiques de même "couleur" épigénétique ont de plus tendance à interagir entre eux (18) suggérant ainsi un mécanisme moléculaire d'interaction spécifique entre mêmes états épigénétiques locaux. L'organisation de ces domaines n'est par ailleurs pas aléatoire au sein du noyau (34, 45, 46, 43) : l'hétérochromatine a tendance à se former plutôt à la périphérie du noyau ainsi qu'autour du nucléole tandis que l'euchromatine est localisée plutôt à l'intérieur (47). Cette organisation spatiale évolue au cours du développement, en passant d'une faible organisation globale dans les cellules pluripotentes à une ségrégation et compartimentalisation fortes dans les cellules différenciées (48, 49, 34). Cette réorganisation spatiale est dans le cas de l'embryogénèse fortement corrélée à la dynamique des domaines épigénétiques (39) mais au cours d'autres processus développementaux une réorganisation spatiale peut intervenir indépendamment de la dynamique des marques épigénétiques (43).

L'objectif a donc été de comprendre à travers un modèle de polymère comment la compartimentalisation en domaines épigénomiques pouvait rendre compte de l'organisation spatiale de la chromatine et notamment de la compartimentation 3D observée. Nous avons introduit un modèle de copolymère par blocs où chaque bloc correspond à un domaine épigénomique (Fig. 17). Les interactions de volume entre monomères sont de deux types : interactions non-spécifiques (pour rendre compte du confinement global) et interactions spécifiques entre monomères de même "couleur" épigénétique. Cette spécificité est motivée par de nombreuses données expérimentales révélant une interaction effective entre fragment de même "état" chromatinien (couleur). En effet il a été montré que les complexes protéiques du groupe Polycomb (50, 51, 52) et HP1 (53) peuvent créer (par oligomérisation) des pontages physiques entre sites distants. De plus, la mutualisation des polII (usines de transcription), la formation de boucles entre sites de régulation par des protéines insulatrices du type CTCF, ou l'ancrage à des pores nucléaires peuvent conduire aussi à une interaction effective entre sites actifs (54, 55). Enfin, la chromatine dite "noire" est souvent associée aux "lamines" (33) suggérant une interaction effective via l'ancrage à la membrane.

**7.2 Application (I) : Drosophile** Comme première approximation, nous avons en effet considéré que les interactions spécifiques avaient la même valeur quelle que soit le type épigénomique en question (i.e, par exemple deux monomères "Polycomb" interagissent avec la même intensité que deux monomères "HP1"...). Nous avons modélisé la dynamique de la chaîne polymérique par une équation de Langevin non-linéaire qu'on a résolue par une méthode d'approximation de champ gaussien auto-consistant. Nous avons aussi mis en place des simulations de dynamique moléculaire. La méthode approchée nous permet de façon très efficace d'obtenir des diagrammes de phases, comme celui reporté à la figure 18 dans l'exemple pédagogique du copolymère ( $A_{10}B_{10}$ ). Dans ce cas simple on voit déjà apparaître la complexité du diagramme de phase avec notamment une zone de paramètres correspondant à de la multistabilité.

Appliqué à la drosophile, à savoir en considérant en entrée la séquence 1D des domaines épigénomiques obtenus par (33), nous avons montré que notre modèle de copolymère "minimal" pouvait déjà très bien rendre compte de l'organisation 3D telle qu'elle a été mesurée par les cartes Hi-C de (18) (Fig 19). Nos résultats suggèrent par ailleurs que les données expérimentales sont compatibles avec des configurations multistables à savoir des configurations caractérisées par les domaines épigénomiques repliés sur eux-mêmes pour former les domaines topologiques (les "TADs") qui peuvent ensuite interagir entre eux à longue portée (si de même type épigénomique) mais de façon dynamique (cf exemple pédagogique 18). L'existence de cette multistabilité suggère par ailleurs que la relaxation au cours du cycle cellulaire peut être fortement ralentie par cette multistabilité de sorte que l'équilibre pourrait finalement ne jamais être



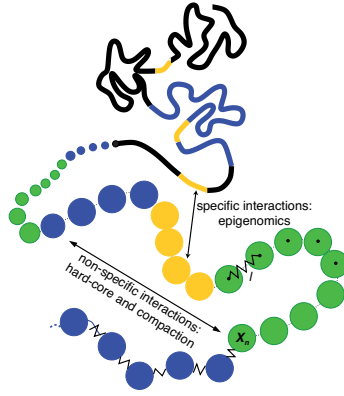


FIGURE 17: Block copolymer model : the chromatin is modeled as a self-avoiding bead-and-spring chain (56) where each monomer represents a portion of DNA (10 kbp) and is characterized by its epigenetic state : yellow (active), green (HP1-like heterochromatin), blue (Polycomb-like heterochromatin), black (repressive chromatin) (33). The model integrates non-specific and specific short-range interactions to account respectively for the effective compaction of the chain and for epigenomically-related affinities between monomers.

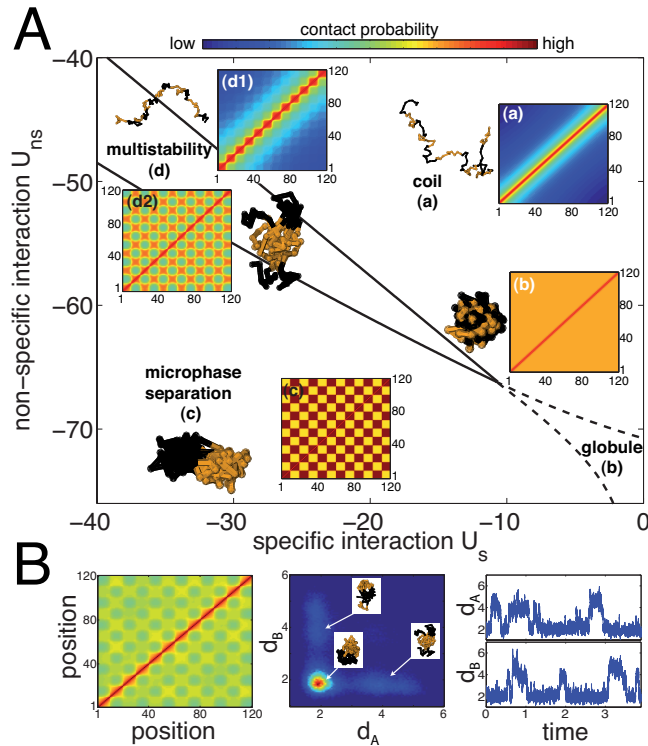


FIGURE 18: (A) Phase diagram of the copolymer  $(A_{10}B_{10})_6$  as a function of the strength of specific and non-specific interactions (in  $k_B T$ -unit). Insets represent typical heat maps of the probability of contacts between two monomers (in log-unit) for the different phases : coil (a), globule (b), multiphase separation (MPS) (c) and multistability (d). Snapshots results from full numerical simulations of the system. (B) Contact map (Left), joint-probability distribution function for the root mean squared distance (r.m.s.d)  $d_A$  between  $A$ -monomers and the r.m.s.d  $d_B$  between  $B$ -monomers (Center), and typical time-evolution of  $d_A$  and  $d_B$  along one simulated trajectory (Right), obtained from full numerical simulations for a parameter set inside the multistability region (d2 in A).

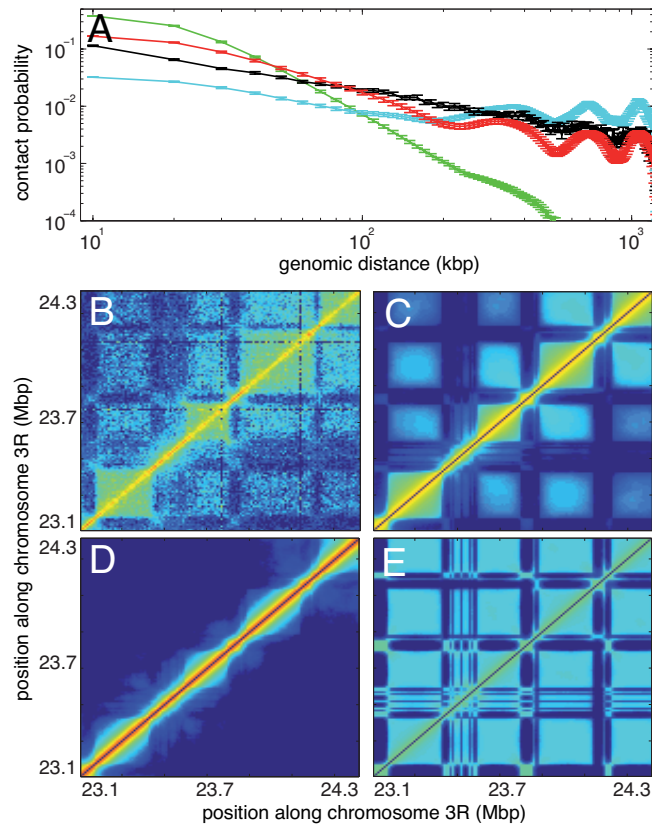


FIGURE 19: (A) Average contact probability between two loci of the chromatin region located between 23.05 and 24.36 Mbp of chromosome 3R, as a function of their genomic distance. (B,C,D,E) Steady-state contact maps. Taken from experimental data (black, B) or from full numerical simulations for parameter sets at the multistability/coil boundary (green, D), at the multistability/MPS boundary (cyan, E) or inside the multistability region (red, C).

atteint dans des cellules qui cyclent et donc périodiquement “réinitialisées” dans l’état mitotique.

Ce travail est le fruit d’une collaboration étroite avec G. Cavalli de l’IGH (Montpellier) et de Pascal Carrivain, physicien, actuellement post-doc conjoint chez G. Cavalli. Il s’est fait dans le cadre de deux financements ( $2 \times 5000$  euros) obtenus par le biais de l’appel à projet “PEPS Interdisciplinaires PTI Physique théorique et ses interfaces”, 2013 et 2014. Le projet est disponible à l’adresse <http://perso.ens-lyon.fr/cedric.vaillant/doku.php?id=divers>

Un article sur les résultats obtenus a été publié dans *Nucleic Acids Research* en 2014 :

D. Jost, P. Carrivain, G. Cavalli C. Vaillant. “Modeling epigenome folding : formation and dynamics of topologically associated chromatin domains.” *Nucleic Acids Res* **42**, 9553–9561 (2014).

Téléchargeable à l’adresse : <http://perso.ens-lyon.fr/cedric.vaillant/doku.php?id=divers>

Ce travail a été présenté dans différentes conférences à travers un poster aux conférence “Nuclear Structure and Dynamics” (L’Isle sur la Sorgue, Octobre 2013), “Advanced Workshop on Interdisciplinary Views on Chromosome Structure & Function” (Trieste, Septembr 2014), “Lyon SysBio 2014” (Lyon, Novembre 2014), “Physique théorique et ses Interfaces” (Paris, Novembre 2014).

Ces résultats ont permis de pérenniser la collaboration avec G. Cavalli qui s’est concrétisée par le dépôt d’un projet ANR en 2014 (non retenu) et de nouveau d’un pre-projet à l’automne 2014, toujours avec la participation de Ralf Everaers, professeur à l’ENS, collègue du Laboratoire de Physique et du Centre Blaise Pascal.

**7.3 Application (II) : *C. elegans*** Dans le même esprit, avec Daniel Jost, nous avons travaillé avec le groupe de Peter Meister (<http://www.izb.unibe.ch/content/groups/meister/>) sur la modélisation de l’organisation spatiale des chromosomes X chez le ver *C. elegans*. L’objectif est ici de comprendre les mécanismes moléculaires impliqués dans le processus dite de “compensation du dosage” des chromosomes X. Ce processus permet d’adapter le niveau d’expression des genes des chromosomes X selon le nombre de ces chromosomes présent dans les cellules. Ainsi chez les mammifères, l’un des deux chromosomes X chez les femelles est totalement réprimé transcriptionnellement au cours de l’embryogenèse précoce des lignées somatiques et cet état cellulaire persiste tout au long de la vie de l’individu : moléculairement, ou encore “épigénomiquement” parlant, ce chromosome inactivé forme un seul domaine épigénomique heterochromatinien avec un mode de repression de type “Polycomb”. Chez *C. elegans*, ce sont les deux chromosomes X de l’individu hermaphrodite qui sont exprimés de moitié par rapport au chromosome X mâle (57). Une des signatures “épigénomique” de ce processus de compensation est que les deux chromosomes X hermaphrodites, contrairement au X mâle, constituent chacun d’eux un seul domaine chromatinien caractérisé par la présence tout le long du complexe protéique dit “DCC” et de marques d’histones répressives H4K20me1. Il apparait par ailleurs que cette différence épigénomique entre le X mâle et les X hermaphrodites soit associée à des organisation spatiales bien distinctes : des expériences de microscopie couplée à des expériences de mesures biochimiques d’interaction avec la membrane menées récemment dans le groupe de P. Meister ont permis de montrer que (i) les chromosomes X sont davantage condensés que les autosomes, (ii) que le X mâle est davantage localisé à la périphérie et interagit davantage avec les composantes des pores nucléaires (que les X herma. et autosomes) et (iii) que les X hermaphrodites sont plus à l’intérieur du noyau et légèrement plus compacts que les X mâles. Pour interpréter ces résultats nous avons mis en place des dynamiques moléculaires en ne considérant cette fois-ci qu’un seul domaine épigénomique couvrant tout le chromosome (donc modèle d’homopolymère) ; l’effet du DCC dans le cas compensé a été introduit par une interaction attractive entre monomères (le DCC pouvant, comme HP1, PcG... pointer deux fragments de chromatine) ; des interactions avec la membrane ont été introduites, soit de façon “continue” (interaction de type chromatine-lamines) soit de façon discrète (type chromatine-pores nucléaires) (Fig. 20). Dans le cadre de ce modèle, la configuration du X lorsque que l’interaction effective monomères-monomères est plus forte que celle monomères-pore nucléaires (du fait de la présence du complexe DCC) est celle adoptée par les X hermaphrodites (Fig. 20). Inversement lorsqu’il n’y a pas de DCC, l’interaction avec les pores induit une colocalisation à la membrane qui dans le cas de non équilibration, peut induire une compaction assez importante. Dans le cas d’une interaction continue ou alors discrète mais équilibrée cette compaction serait moins importante. Donc, le X mâle serait “figé” par les pore dans une configuration encore proche de la configuration mitotique et resterait assez compact malgré l’absence de complexes de condensation (type DCC).

Ce travail a fait l’objet d’une publication en 2014 :

R. Sharma, D. Jost, J. Kind, G. Gómez-Saldivar, B. van Steensel, P. Askjaer, C. Vaillant P. Meister. Differential spatial and structural organization of the X chromosome underlies dosage compensation in *C. elegans*. *Genes Dev.* **28**, 2591–2596.

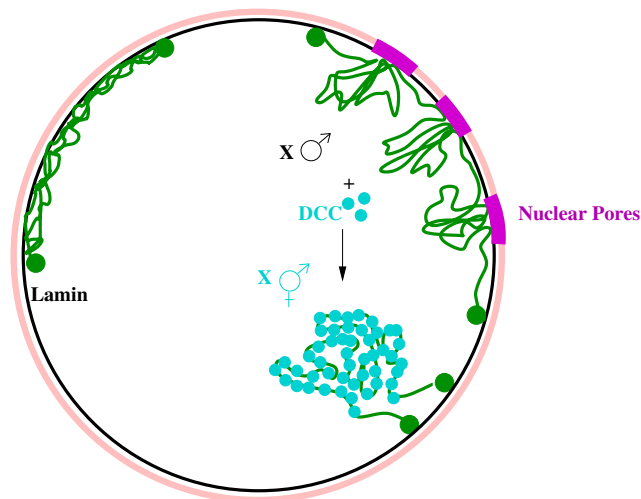


FIGURE 20: Mécanisme associé à la compensation du dosage des chromosomes X : en absence de DCC les chromosomes X chez le mâle sont localisés à la membrane par des interactions discrètes aux pores nucléaires à des séquences spécifiques (dites “REX”). Chez les hermaphrodites l’association des complexes DCC à la chromatine tout d’abord à ces sites REX puis ensuite au reste du chromosome, induit une relocalisation vers l’intérieur du noyau et à une augmentation de la compaction.

Téléchargeable à :

[http://perso.ens-lyon.fr/cedric.vaillant/lib/exe/fetch.php?media=genes\\_dev.-2014-sharma-2591-6.pdf](http://perso.ens-lyon.fr/cedric.vaillant/lib/exe/fetch.php?media=genes_dev.-2014-sharma-2591-6.pdf)

Un projet ANR (PRCI avec la Suisse) avec le groupe de P. Meister sera présenté début 2015. (cf Projet de Recherche) (Un projet HFSP avait été présenté en 2014 mais il n’avait pas été retenu).

Il faut mentionner ici dans le cadre de ces travaux que je viens d’être sollicité pour faire partie du réseau européen de recherche Epigenesys (<http://www.epigenesys.eu/en/>) en tant que membre associé.

## 8 MOTILITÉ BACTÉRIENNE

Je collabore avec Laurence Lemelle, Christophe Place et JF Palierne sur des travaux concernant la motilité de bactéries flagellées près de surfaces (<http://www.ens-lyon.fr/Joliot-Curie/spip.php?rubrique65>). Par des méthodes de microscopie à très haute résolution et haute cadence, elle a pu obtenir un grand nombre de séries temporelles caractérisant les cinématiques du corps des bactéries (trajectoire du centre de masse et rotation du corps). Ma contribution consiste en l’analyse du signal de ces données ainsi qu’à leur modélisation. Cette analyse a permis notamment de caractériser très finement les étapes associées aux événements de “tumbling”, à savoir de réorientation aléatoire des trajectoires bactériennes. Un modèle simple permet de suggérer que les tumbling observés au cours des nages interfaciales peuvent être en grande partie expliqués par un modèle de tumbling en 3D (modélisé lui-même par une diffusion rotationnelle isotrope) mais cette fois-ci confiné à 2D et biaisé par la composante rotationnelle déterministe caractéristique de la nage près des surfaces. Cela fait l’objet d’un article prochainement soumis à PNAS :

L. Lemelle, T. Cajgfinger, C.C. Nguyen, A. Dominjon, C. Place, R. Barbier, J-F Palierne and C. Vaillant. “Tumbling kinematics in swimming *Escherichia coli* near a solid surface”.

Un pré-projet ANR (“PINBALL : Primo INteraction de la BActérie *Pseudomonas aeruginosa* avec des ceLLules hôtes”) sur l’étude de la motilité près de surfaces cellulaires lors de phases de pré-infection par *Pseudomonas aeruginosa* a été déposée à l’automne 2014 (Coordinateur : L. Lemelle).

## PUBLICATIONS

- [1] Moukhtar, J., Vaillant, C., Audit, B., and Arneodo, A. (2009) Generalized wormlike chain model for long-range correlated heteropolymers. *Europhys. Lett.*, **86**, 48001.
- [2] Moukhtar, J., Favier-Moskalenko, C., Milani, P., Audit, B., Vaillant, C., Fontaine, E., Mongelard, F., Lavorel, G., St-Jean, P., Bouvet, P., Argoul, F., and Arneodo, A. (2010) Effect of genomic long-

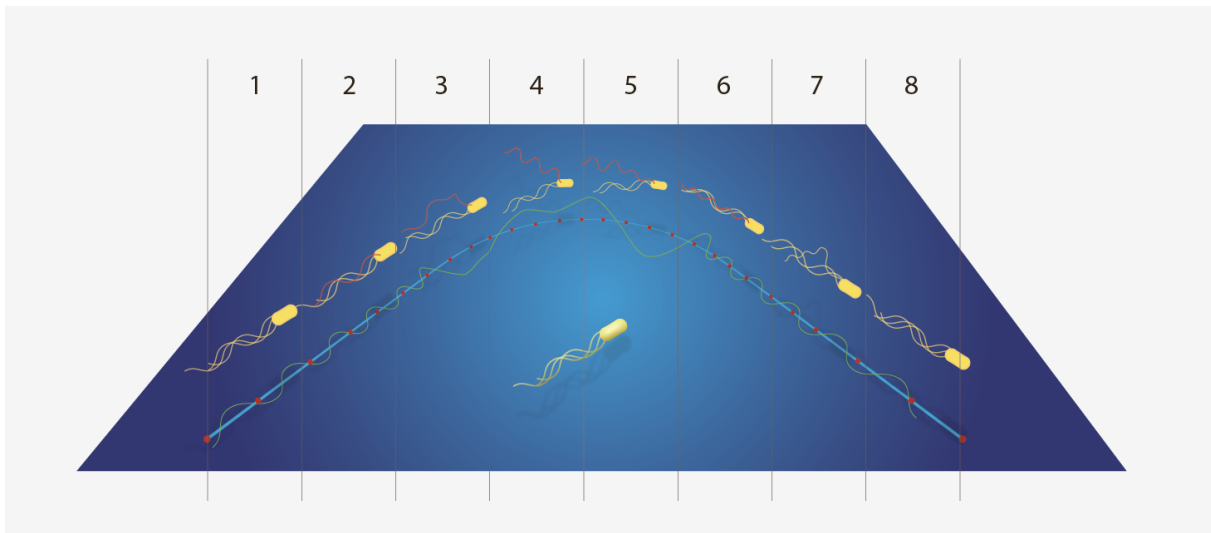


FIGURE 21: **Model sequence of the body motion in a tumble occurring on a solid surface.** This figure was constructed based on that of Darnton et al. (Darnton N C, Turner L, Rojevsky S and Berg H C (2007) On Torque and Tumbling in Swimming Escherichia coli. *J Bacteriol*, 189 :1756-1764.) that sums up the observations of the polymorphic transformations of the flagella and their interpretation in terms of reorientation of the motion within a canonical tumble caused by the reversal of one single motor in bulk solution. The different body motions deduced from the speed,  $V$ , the direction of the motion,  $\phi$ , and the body orientation relative to the trajectory,  $\theta_c$  analyses are superimposed on this schema. The speed is rendered graphically by the length between consecutive red dots representing the centroids of the image of the body projected on the surface that are sampled at a constant frequency. The motion direction is indicated by the blue line passing through the red dots. For the sake of simplicity, the CW circular trajectory in the run phases was represented with a very low curvature value. The wobbling of the body and its fluctuation are represented by the green line oscillating around the line of propagation.

- range correlations on DNA persistence length : From theory to single molecules experiments. *J. Phys. Chem. B*, **114**, 5125–5143.
- [3] Miele, V., Vaillant, C., d'Aubenton-Carafa, Y., Thermes, C., and Grange, T. (Jun, 2008) DNA physical properties determine nucleosome occupancy from yeast to fly. *Nucleic Acids Res.*, **36**(11), 3746–3756.
  - [4] Vaillant, C., Audit, B., and Arneodo, A. (Nov, 2007) Experiments confirm the influence of genome long-range correlations on nucleosome positioning. *Phys. Rev. Lett.*, **99**(21), 218103.
  - [5] Chevereau, G., Arneodo, A., and Vaillant, C. (2011) Influence of the genomic sequence on the primary structure of chromatin. *Frontiers in Life Science*, **5**, 29–68.
  - [6] Chevereau, G., Palmeira, L., Thermes, C., Arneodo, A., and Vaillant, C. (Oct, 2009) Thermodynamics of intragenic nucleosome ordering. *Phys. Rev. Lett.*, **103**(18), 188103.
  - [7] Chevereau, G. Thermodynamique du positionnement des nucléosomes PhD thesis Université de Lyon-Ecole Normale Supérieure, Lyon, France. (2010).
  - [8] Lee, W., Tillo, D., Bray, N., Morse, R. H., Davis, R. W., Hughes, T. R., and Nislow, C. (Oct, 2007) A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Genet.*, **39**(10), 1235–1244.
  - [9] Chevereau, G., Palmeira, L., Thermes, C., Arneodo, A., and Vaillant, C. (2009) Thermodynamics of intra-genic nucleosome ordering. *Phys. Rev. Lett.*, **103**, 188103.
  - [10] Möbius, W. and Gerland, U. (2010) Quantitative test of the barrier nucleosome model for statistical positioning of nucleosomes up- and downstream of transcription start sites. *PLoS Comput. Biol.*, **6**(8), e1000891.
  - [11] Vaillant, C., L.Palmeira, Chevereau, G., Audit, B., d'Aubenton-Carafa, Y., Thermes, C., and Arneodo, A. (2010) A novel strategy of transcription regulation by intra-genic nucleosome ordering. *Genome Res.*, **20**, 59–67.
  - [12] Milani, P., Chevereau, G., Vaillant, C., Audit, B., Haftek-Terreau, Z., Marilley, M., Bouvet, P., Argoul, F., and Arneodo, A. (Dec, 2009) Nucleosome positioning by genomic excluding-energy barriers. *Proc. Natl. Acad. Sci. U S A*, **106**(52), 22257–22262.

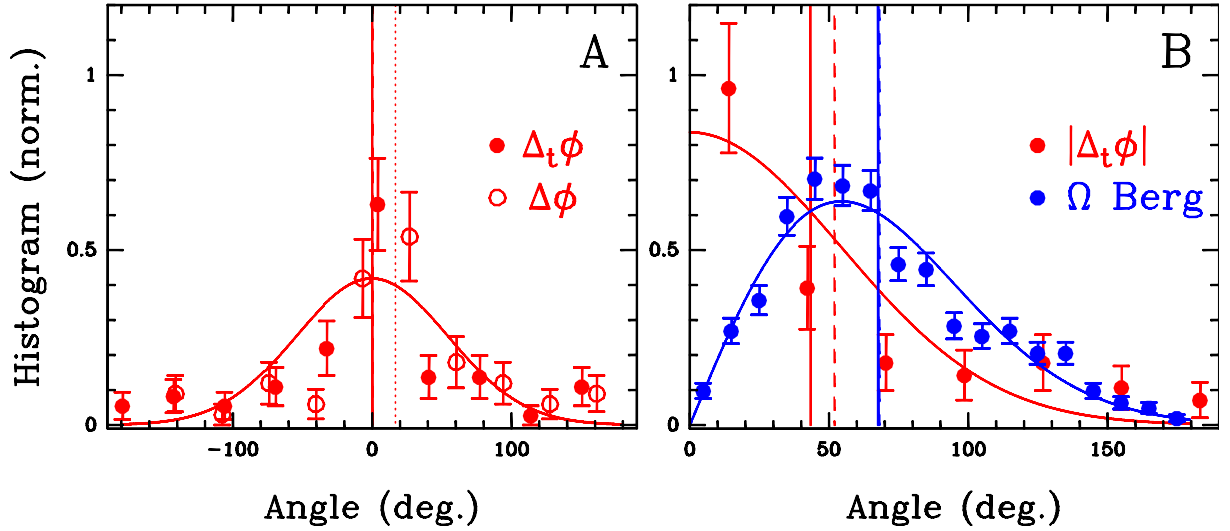


FIGURE 22: **Distributions of “tumbling” reorientation angles** (A) Normalized distribution of the reorientation angles  $\Delta\phi$  extracted from the 52 “low-speed” events (bin size  $18^\circ$ ) (red open circle). The forward motion is oriented along the  $0^\circ$ , the CW motion towards the positive values of angles. Mean value (red vertical dotted line) is  $17^\circ \pm 70^\circ$ . Corresponding normalized distribution of the centered reorientation angle  $\Delta_t\phi = \Delta\phi - \Delta\phi_p$  (red dots) and the theoretical distribution (red curve) using a diffusion constant  $d_t = 2.6 \text{rad}^2 \cdot \text{s}^{-1}$  (see (B)) Mean of experimental (red vertical line) and theoretical (red vertical dashed line) distributions are  $0.5 \pm 77^\circ$  and  $0 \pm 54^\circ$ . (B) Bulk distribution of the reorientation angle  $\Omega$  first measured by Berg and Brown (blue dots) (1) and the corresponding theoretical distribution (blue curve) when considering a diffusion constant  $d^t = 2.6 \text{rad}^2 \cdot \text{s}^{-1}$  (see Mat.& Meth). Mean of experimental (blue vertical line) and theoretical (blue vertical dashed line) distributions are  $68 \pm 37^\circ$  and  $68 \pm 36^\circ$ . Normalized distributions of the near surface “unsigned” reorientation angle ( $\Omega = |\Delta_t\phi|$ ) obtained experimentally (red dots) (same data as in (A)) and theoretically (red curve) when using the same diffusion constant  $d_t = 2.6 \text{rad}^2 \cdot \text{s}^{-1}$  as for bulk. Mean of experimental (red vertical line) and theoretical (red vertical dashed line) distributions are  $52 \pm 37^\circ$  and  $43 \pm 33^\circ$ .

- [13] Audit, B., Zaghloul, L., Vaillant, C., Chevereau, G., d'Aubenton-Carafa, Y., Thermes, C., and Arneodo, A. (Oct, 2009) Open chromatin encoded in DNA sequence is the signature of “master” replication origins in human cells. *Nucleic Acids Res.*, **37**(18), 6064–6075.
- [14] Baker, A., Audit, B., Chen, C.-L., Moindrot, B., Leleu, A., Guilbaud, G., Rappailles, A., Vaillant, C., Goldar, A., Mongelard, F., d'Aubenton-Carafa, Y., Hyrien, O., Thermes, C., and Arneodo, A. (2012) Replication fork polarity gradients revealed by megabase-sized u-shaped replication timing domains in human cell lines. *PLoS Comput Biol*, **8**(4), e1002443.
- [15] Eaton, M. L., Galani, K., Kang, S., Bell, S. P., and MacAlpine, D. M. (Apr, 2010) Conserved nucleosome positioning defines replication origins. *Genes Dev*, **24**(8), 748–753.
- [16] Eaton, M. L., Prinz, J. A., MacAlpine, H. K., Tretyakov, G., Kharchenko, P. V., and MacAlpine, D. M. (Feb, 2011) Chromatin signatures of the drosophila replication program. *Genome Res*, **21**(2), 164–174.
- [17] Ding, Q. and MacAlpine, D. M. (Apr, 2011) Defining the replication program through the chromatin landscape. *Crit Rev Biochem Mol Biol*, **46**(2), 165–179.
- [18] Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (Feb, 2012) Three-dimensional folding and functional organization principles of the drosophila genome. *Cell*, **148**(3), 458–472.
- [19] Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., and Ren, B. (May, 2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, **485**(7398), 376–380.
- [20] Lesbats, P., Botbol, Y., Chevereau, G., Vaillant, C., Calmels, C., Arneodo, A., Andreola, M.-L., Lavigne, M., and Parissi, V. (2011) Functional coupling between hiv-1 integrase and the swi/snf chromatin remodeling complex for efficient in vitro integration into stable nucleosomes. *PLoS Pathog*, **7**(2), e1001280.
- [21] Wang, G. P., Ciuffi, A., Leipzig, J., Berry, C. C., and Bushman, F. D. (Aug, 2007) Hiv integration site selection : analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Res*, **17**(8), 1186–1194.
- [22] Schones, D. E., Cui, K., Cuddapah, S., Roh, T.-Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (Mar, 2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell*, **132**(5), 887–898.
- [23] Cattoglio, C., Pellin, D., Rizzi, E., Maruggi, G., Corti, G., Miselli, F., Sartori, D., Guffanti, A., Serio, C. D., Ambrosi, A., Bellis, G. D., and Mavilio, F. (Dec, 2010) High-definition mapping of retroviral integration sites identifies active regulatory elements in human multipotent hematopoietic progenitors. *Blood*, **116**(25), 5507–5517.
- [24] Clapier, C. R. and Cairns, B. R. (2009) The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.*, **78**, 273–304.
- [25] Wang, X., Bai, L., Bryant, G. O., and Ptashne, M. (Dec, 2011) Nucleosomes and the accessibility problem. *Trends Genet.*, **27**(12), 487–492.
- [26] Hartley, P. D. and Madhani, H. D. (May, 2009) Mechanisms that specify promoter nucleosome location and identity. *Cell*, **137**(3), 445–458.
- [27] Garcia, J. F., Dumesic, P. A., Hartley, P. D., El-Samad, H., and Madhani, H. D. (Aug, 2010) Combinatorial, site-specific requirement for heterochromatic silencing factors in the elimination of nucleosome-free regions. *Genes Dev.*, **24**(16), 1758–1771.
- [28] Floer, M., Wang, X., Prabhu, V., Berrozpe, G., Narayan, S., Spagna, D., Alvarez, D., Kendall, J., Krasnitz, A., Stepansky, A., Hicks, J., Bryant, G. O., and Ptashne, M. (Apr, 2010) A rsc/nucleosome complex determines chromatin architecture and facilitates activator binding. *Cell*, **141**(3), 407–418.
- [29] Whitehouse, I. and Tsukiyama, T. (Jul, 2006) Antagonistic forces that position nucleosomes *in vivo*. *Nat. Struct. Mol. Biol.*, **13**(7), 633–640.
- [30] Németh, A., Strohner, R., Grummt, I., and Längst, G. (2004) The chromatin remodeling complex NoRC and TTF-I cooperate in the regulation of the mammalian rRNA genes in vivo. *Nucleic Acids Res*, **32**(14), 4091–4099.

- [31] Pointner, J., Persson, J., Prasad, P., Norman-Axelsson, U., StrÄlfors, A., Khorosjutina, O., Krietenstein, N., Svensson, J. P., Ekwall, K., and Korber, P. (Nov, 2012) Chd1 remodelers regulate nucleosome spacing in vitro and align nucleosomal arrays over gene coding regions in *s. pombe*.. *EMBO J*, **31**(23), 4388–4403.
- [32] Roudier, F. and other (May, 2011) Integrative epigenomic mapping defines four main chromatin states in arabidopsis.. *EMBO J*, **30**(10), 1928–1938.
- [33] Filion, G. J., vanBemmel, J. G., Braunschweig, U., Talhout, W., Kind, J., Ward, L. D., Brugman, W., deCastro, I. J., Kerkhoven, R. M., Bussemaker, H. J., et al. (2010) Systematic protein location mapping reveals five principal chromatin types in drosophila cells. *Cell*, **143**(2), 212–224.
- [34] Ahmed, K., Deghani, H., Rugg-Gunn, P., Fussner, E., Rossant, J., and Bazett-Jones, D. P. (2010) Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo.. *PLoS One*, **5**(5), e10531.
- [35] vanSteensel, B. (May, 2011) Chromatin : constructing the big picture.. *EMBO J*, **30**(10), 1885–1895.
- [36] Ernst, J., Kheradpour, P., Mikkelson, T. S., Shores, N., Ward, L. D., Epstein, C. B., Zhang, X., Wang, L., Issner, R., Coyne, M., Ku, M., Durham, T., Kellis, M., and Bernstein, B. E. (May, 2011) Mapping and analysis of chromatin state dynamics in nine human cell types.. *Nature*, **473**(7345), 43–49.
- [37] Gerstein, M. B. and other (Dec, 2010) Integrative analysis of the caenorhabditis elegans genome by the modencode project.. *Science*, **330**(6012), 1775–1787.
- [38] Beisel, C. and Paro, R. (Feb, 2011) Silencing chromatin : comparing modes and mechanisms.. *Nat Rev Genet*, **12**(2), 123–135.
- [39] Hawkins, R. D. and other (May, 2010) Distinct epigenomic landscapes of pluripotent and lineage-committed human cells.. *Cell Stem Cell*, **6**(5), 479–491.
- [40] Feinberg, A. P. (May, 2007) Phenotypic plasticity and the epigenetics of human disease.. *Nature*, **447**(7143), 433–440.
- [41] Pujadas, E. and Feinberg, A. P. (Mar, 2012) Regulated noise in the epigenetic landscape of development and disease.. *Cell*, **148**(6), 1123–1131.
- [42] Bantignies, F. and Cavalli, G. (Nov, 2011) Polycomb group proteins : repression in 3d.. *Trends Genet*, **27**(11), 454–464.
- [43] Chandra, T. and other (Jul, 2012) Independence of repressive histone marks and chromatin compaction during senescent heterochromatic layer formation.. *Mol Cell*, **47**(2), 203–214.
- [44] Cheutin, T. and Cavalli, G. (Jan, 2012) Progressive polycomb assembly on h3k27me3 compartments generates polycomb bodies with developmentally regulated motion.. *PLoS Genet*, **8**(1), e1002465.
- [45] Naumova, N. and Dekker, J. (Jun, 2010) Integrating one-dimensional and three-dimensional maps of genomes.. *J Cell Sci*, **123**(Pt 12), 1979–1988.
- [46] Rapkin, L. M., Anchel, D. R. P., Li, R., and Bazett-Jones, D. P. (Feb, 2012) A view of the chromatin landscape.. *Micron*, **43**(2-3), 150–158.
- [47] Zullo, J. M., Demarco, I. A., PiquÄl-Regi, R., Gaffney, D. J., Epstein, C. B., Spooner, C. J., Luperchio, T. R., Bernstein, B. E., Pritchard, J. K., Reddy, K. L., and Singh, H. (Jun, 2012) Dna sequence-dependent compartmentalization and silencing of chromatin at the nuclear lamina.. *Cell*, **149**(7), 1474–1487.
- [48] Meister, P., Towbin, B. D., Pike, B. L., Ponti, A., and Gasser, S. M. (Apr, 2010) The spatial dynamics of tissue-specific promoters during *c. elegans* development.. *Genes Dev*, **24**(8), 766–782.
- [49] Meister, P., Mango, S. E., and Gasser, S. M. (Apr, 2011) Locking the genome : nuclear organization and cell fate.. *Curr Opin Genet Dev*, **21**(2), 167–174.
- [50] Francis, N. J., Kingston, R. E., and Woodcock, C. L. (Nov, 2004) Chromatin compaction by a polycomb group protein complex.. *Science*, **306**(5701), 1574–1577.
- [51] Lo, S. M., Follmer, N. E., Lengsfeld, B. M., Madamba, E. V., Seong, S., Grau, D. J., and Francis, N. J. (Jun, 2012) A bridging model for persistence of a polycomb group protein complex through dna replication in vitro.. *Mol Cell*, **46**(6), 784–796.



- [52] Isono, K., Endo, T. A., Ku, M., Yamada, D., Suzuki, R., Sharif, J., Ishikura, T., Toyoda, T., Bernstein, B. E., and Koseki, H. (Sep, 2013) Sam domain polymerization links subnuclear clustering of prc1 to gene silencing.. *Dev Cell*, **26**(6), 565–577.
- [53] Canzio, D., Liao, M., Naber, N., Pate, E., Larson, A., Wu, S., Marina, D. B., Garcia, J. F., Madhani, H. D., Cooke, R., Schuck, P., Cheng, Y., and Narlikar, G. J. (Apr, 2013) A conformational switch in hp1 releases auto-inhibition to drive heterochromatin assembly.. *Nature*, **496**(7445), 377–381.
- [54] Phillips-Cremins, J. E., Sauria, M. E. G., Sanyal, A., Gerasimova, T. I., Lajoie, B. R., Bell, J. S. K., Ong, C.-T., Hookway, T. A., Guo, C., Sun, Y., Bland, M. J., Wagstaff, W., Dalton, S., McDevitt, T. C., Sen, R., Dekker, J., Taylor, J., and Corces, V. G. (Jun, 2013) Architectural protein subclasses shape 3d organization of genomes during lineage commitment.. *Cell*, **153**(6), 1281–1295.
- [55] Hou, C., Li, L., Qin, Z. S., and Corces, V. G. (Nov, 2012) Gene density, transcription, and insulators contribute to the partition of the drosophila genome into physical domains.. *Mol Cell*, **48**(3), 471–484.
- [56] Grest and Kremer (May, 1986) Molecular dynamics simulation for polymers in the presence of a heat bath.. *Phys Rev A*, **33**(5), 3628–3631.
- [57] Ferrari, F., Alekseyenko, A. A., Park, P. J., and Kuroda, M. I. (Feb, 2014) Transcriptional control of a whole chromosome : emerging models for dosage compensation.. *Nat Struct Mol Biol*, **21**(2), 118–125.