Project (2017-): Theory of epigenome regulation
Physico-chemical models of chromatin-based control of gene expression

6 janvier 2017

TABLE DES MATIÈRES

Résumé

The general objective of this research project is to build quantitative models of the epigenomic-based mechanisms of gene regulation involved in the short- and long-term cellular response. Using theoretical approaches from statistical and numerical physics and in close collaboration with experimentalists, we propose to study the "1D" assembly of the different eu- and heterochromatin states along the genome (nucleation, spreading, compartmentalization and mitotic/meiotic inheritance), and their "3D" folding and nuclear organization. At the nucleosomal and gene scales, we are currently developing molecular models accounting for the combined action of chromatin regulators and DNA binding proteins to understand the mechanisms by which chromatin states achieve their functions. At larger scale, we use coarse-grained models to derive the structural and dynamical properties of the epigenome : (i) the regulation of epigenomic domains (ii) their folding into spatial compartments and (iii) finally the coupling between 3D organization and 1D assembly. Our main goal is to provide a general framework to understand how the epigenome is regulated and how its affect gene expression : from its establishment during development and differentiation to its deregulation in diseases. Ultimately we propose to understand how the epigenome (via the chromatin

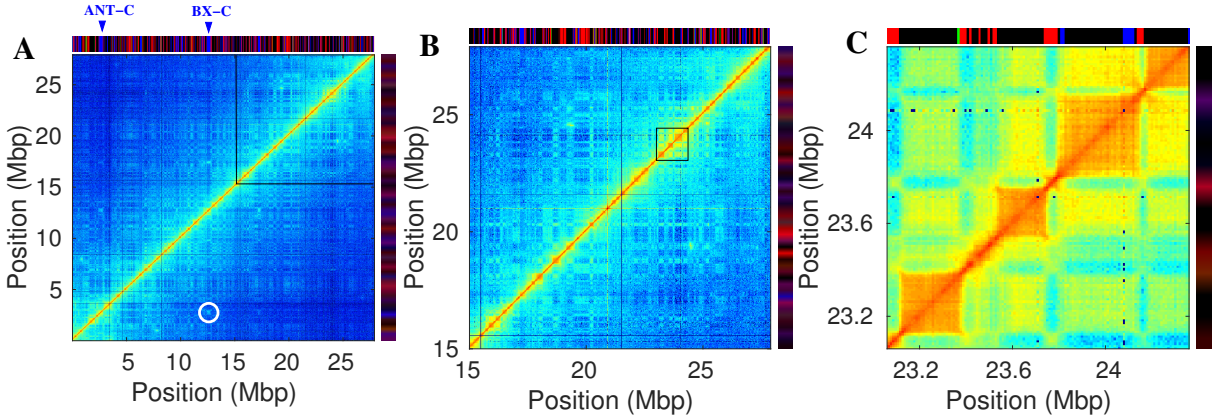


FIGURE 1 – **Compartmentalization of the genome as revealed by Hi-C maps.** (A, B, C) Different magnifications of the Hi-C contact map of chromosome 3R performed in lateëembryonic cellular stages of *D. melanogaster* (1). As clearly shown in (C), the so-called "TADs" correspond to the higher self-interacting genomic domains and strikingly coincide with the active and inactive epigenomic domains obtained by Filion et al (2) (colored-bar segmentation on the top and right in A, B and C) with (red) : active chromatin, (green) : HP1/H3K9me2-3 heterochromatin, (blue) : PcG heterochromatin and (black) : "null" heterochromatin. Additionally, these maps reveal a typical checkerboard pattern corresponding to long-range (up to $\sim$ 15 Mbp) interactions between TADs of the same functional and chromatin state.In (A) the white circle indicates the 10 Mbp long-range contact between the ANT-C and BX-C PolyComb domains.

regulators) "sense" environmental, notably metabolic changes as well as genomic changes and how in some cases, it can translate transient stimuli into stable, inheritable expression pattern changes while in other cases it can act as a "buffer" for these perturbations.

1 INTRODUCTION

The ability of organisms to precisely regulate gene expression is central to their development. Proper temporal and spatial expressions of genes in higher eukaryotes require activation of transcription during the appropriate developmental stages. In response to environmental and developmental cues, cells can adopt different gene expression patterns to differentiate into a variety of cell types ; once established, this pattern is frequently maintained over several cell divisions despite the fact that the initiating signal is no longer present. This ability of translating transient external stimuli into diverse (plasticity) and stable (robustness) phenotypes without alteration of the genomic sequence is at the heart of "epigenetic" regulation of gene expression. Transcription in eukaryotes is regulated by trans-acting factors that associate with the genome in a cell-type and condition-specific manner at specific cis-regulatory elements including proximal promoters, enhancers and repressors. The packaging of eukaryotic DNA into chromatin contribute to this regulation via the modulation of the accessibility and specificity of regulators to their DNA cognate sites. The local chromatin state is characterized by various features like the nucleosome positioning or the covalent modifications of DNA and histones tails. **This pattern of chromatin state along the genome, the so-called "epigenome", is itself regulated by the combined action of different specialized chromatin regulators like chromatin remodellers, modifying enzymes, histone chaperones.**

1.1 *1D segmentation of the epigenome.*

The general picture that emerges from the genome-wide high-resolution profiling of structural and functional chromatin markers obtained in various organisms and cell types (2, 3, 4, 5), is that eukaryotic genomes are linearly organized into distinct epigenomic domains (Fig 1). These domains extend over few kb up to few megabases, are characterized by a specific type of chromatin and are isolated from their neighborhood by boundary elements such as insulators. Euchromatin, less condensed, early replicating and containing most active genes, is generally distinguished from heterochromatin, typically highly condensed, late

replicating and inhibitory to transcriptional machinery. Heterochromatin is classified into two subtypes : constitutive and facultative heterochromatin. Constitutive heterochromatin contains highly repetitive DNA sequences such as those found at centromeres and telomeres, and serves to stably silence transposable elements as well as facilitates chromosome segregation. In contrast, facultative heterochromatin is typically associated with developmentally regulated genes whose chromatin structure may change in response to cellular differentiation signals. In many higher eukaryotes, from plants to mammals, statistical analysis of hundreds of chromatin markers have identified only a small number of main chromatin types (2, 6, 4), typically 4 or 5, covering the well-known constitutive HP-1-like heterochromatin or the facultative Polycomb-like heterochromatin but also a less-characterized ultra-repressive heterochromatin enriched in genes that are expressed in very few tissues, the so-called black chromatin (2, 7).

1.2 ***3D compartmentalization of the epigenome.*** Interestingly, within epigenomic domains, regulatory sequences such as enhancers may be located far from the target genes and multiple elements that are arrayed over large regions may collaborate or compete for the regulation of individual genes or gene clusters. This implies the existence of long-range mecanisms where regulatory elements could act over large genomic distances up to hundreds of Kb or more. A first possibility is the linear spreading of a regulatory signal (e.g. repressive chromatin state) from nucleation sites (e.g. silencers) to target-sites (e.g. promoters). Another -but not exclusive- strategy is to take benefit of the polymeric nature of chromatin and to induce spatial co-localization of regulatory sequences with their target. Recently, 3C-based studies have indeed shown that regulatory elements can act over large genomic distances by chromatin looping (8, 9) forming active or repressive chromatin higher-order structure at particular developmental-regulated genes. These pair-wise 3D interactions are mediated by DNA binding proteins such as insulators or cohesins that would cluster in space and bridge distant cis-regulatory sites. At a genomic scale, the contact maps of drosophila (1, 10) , mouse (11) and human (11, 12) chromosomes have further revealed a remarkable 3D compartmentalization where epigenomic domains fold into independant "spatial domain", the so-called topologically-associating domains (TADs), characterized by (i) high intra-domain contact frequencies ; (ii) three-dimensional insulation between adjacent domains ; (iii) and in many cases, significant contacts between distal domains of the same chromatin type (Fig. 1). This compartmentalization is consistent with the nuclear structure revealed by imaging techniques such as EM and immuno-FISH (13, 14, 15) that clearly shows a phase separation between eu- vs heterochromatin and to some extent between the different heterochromatin types (16).

1.3 ***Epigenome regulation during development, disease and aging*** During development, cell differentiation proceeds by global and concommitent rearrangements of epigenomic profile, chromatin organization and transcriptional activity (17, 18, 19, 20). Developmental cell specification is globally accompanied by a progressive chromatin restriction, starting from open (3D) and permissive (1D) chromatin organization in pluripotent ES to increasingly repressive, compact and segregated state in differentiated cells.

The epigenome of cancer cells exhibit many aberrations when compared with normal cells (21). This includes in particular DNA hypo- and hypermethylation and associated transcriptional derepression, gene silencing and genome instability. Global DNA hypomethylation is thought to cause expression and recombination of repetitive sequences leading to instability of the cancer genome, whereas hypermethylation at CpG islands can contribute to cell transformation by silencing tumour suppressor genes. More recent studies have also linked DNA hypomethylation in cancer cells to formation of repressive chromatin domains and gene silencing. The origin of these aberrations is unknown, but may be linked to perturbations in the DNA modification machinery.

(Text extracted from (22))

The question of whether epigenetic alterations are a driving feature of cellular aging is important, given the longstanding known connections between life span, chromatin enzymes, and metabolic changes. In particular, the discovery that yeast Sir2, a key protein in heterochromatic silencing and promoting longevity, is an NAD^+-dependent HDAC, has suggested that chromatin is involved in aging. Further observations, linking Sirtuin pathways with life span extension by calorie/dietrary restriction, underscored the role of metabolism in regulating longevity. Cellular senescence is a response to environmental stress and cellular damage whereby cells cease to

replicate, but are metabolically active. This process is, in fact, physiologically linked to aging, accelerating the process and involving enormous epigenetic alterations. Thus, as discussed below, numerous findings have been made, establishing the involvement of epigenetic regulation in the aging processes. Loss of repressive chromatin occurs during aging. The level of histone H3K9me3, a heterochromatic modification associated with pericentric and telomeric locations, decreases at these locations and broadly increases at euchromatin genome-wide. This imbalance in repressive chromatin observed with age is mirrored by similar changes in the localization of HP1, a histone H3K9me3-binding protein, as well as changes in transcription across the genome . These observations in the yeast, worm, and fly model systems consistently implicate a reduction of heterochromatin and the relaxation of chromatin-mediated transcriptional control during aging, leading to shortened life span.

Senescence is a stable proliferation arrest (irreversible cell cycle arrest) characterized by profound changes in cellular morphology and metabolism as well as by extensive chromatin reorganization in the nucleus. Cellular senescence occurs in response to stress caused by expression of activated oncogenes or long-term replication, both of which are associated with cellular dysfunction and, thus, the cell cycle arrest is protective against cancer. However, senescence also correlates with aging in primates and, importantly, was recently causally linked to aging in a mouse progerin premature aging model via demonstration that removal of senescent cells ameliorates age-associated tissue pathologies. Hence, senescence is a provocative state of protection against cancerous malignancy, but also correlates with cellular damage associated with aging. Aberrant large-scale chromatin foci, called Senescent-Associated Heterochromatin Foci (SAHF), have been shown to arise during senescence and are enriched in heterochromatin proteins and chromatin changes indicative of DNA damage, such as HMGA (high mobility group AT hook protein) and macroH2A (H2A variant protein with macro domain). These acquired SAHFs repress certain genes required for cell proliferation.

As already noticed, cellular senescence is an important tumour suppressor mechanism. For example, replicative senescence blocks tumour formation by imposing an upper limiton the proliferative capacity of normal cells. To become fully transformed, cancer cells must bypass senescence (by circumventing or inactivating the senescence barrier before or after its imposition, respectively). Interestingly, gains and losses of methylation in replicative senescence are thus qualitatively similar to those in cancer, and this "reprogrammed" methylation landscape is largely retained when cells bypass senescence. Consequently, the DNA methylome of senescent cells might promote malignancy, if these cells escape the proliferative barrier.

1.4 *Metabolo-epigenetics* (Extracted from (24, 22))

To make the appropriate developmental decisions or maintain homeostasis, cells and organisms must coordinate the expression of their genome and metabolic state. However, the molecular mechanisms that relay environmental cues such as nutrient availability to the appropriate gene expression response remain poorly understood. Specialized transcription factors that are activated by metabolic sensors or extracellular signals, such as hormones, direct a transcriptional response to changes in metabolic state. In addition, key components of intermediary metabolism are cofactors or cosubstrates of chromatin-modifying enzymes. Therefore, changes in cofactor availability may affect chromatin structure and gene expression. As such, their concentrations constitute a potential regulatory interface between the metabolic and chromatin states. In addition, there is increasing evidence for a direct involvement of classic metabolic enzymes in gene expression control (Fig. 5). A prominent area in epigenetic research that has emerged in recent years relates to how cellular metabolism regulates various events of chromatin remodeling (25, 22). Cells sense changes in the environment and translate them into specific modulations of the epigenome through a variety of signaling components, several of which are proteins with histone- and DNA-modifying enzymatic activity. There are now a myriad of residues on DNA and histone tails that can undergo modification at a given time. The enzymes that elicit these modifications rely critically on the availability of phosphate, acetyl, and methyl groups, to mention a few. As such, their concentrations constitute a potential regulatory interface between the metabolic and chromatin states. This constitutes an intriguing link between cellular metabolism and epigenetic control that has previously been largely unappreciated. Although the specificity and degree of change in the levels of cellular metabolites may influence the epigenome is not well

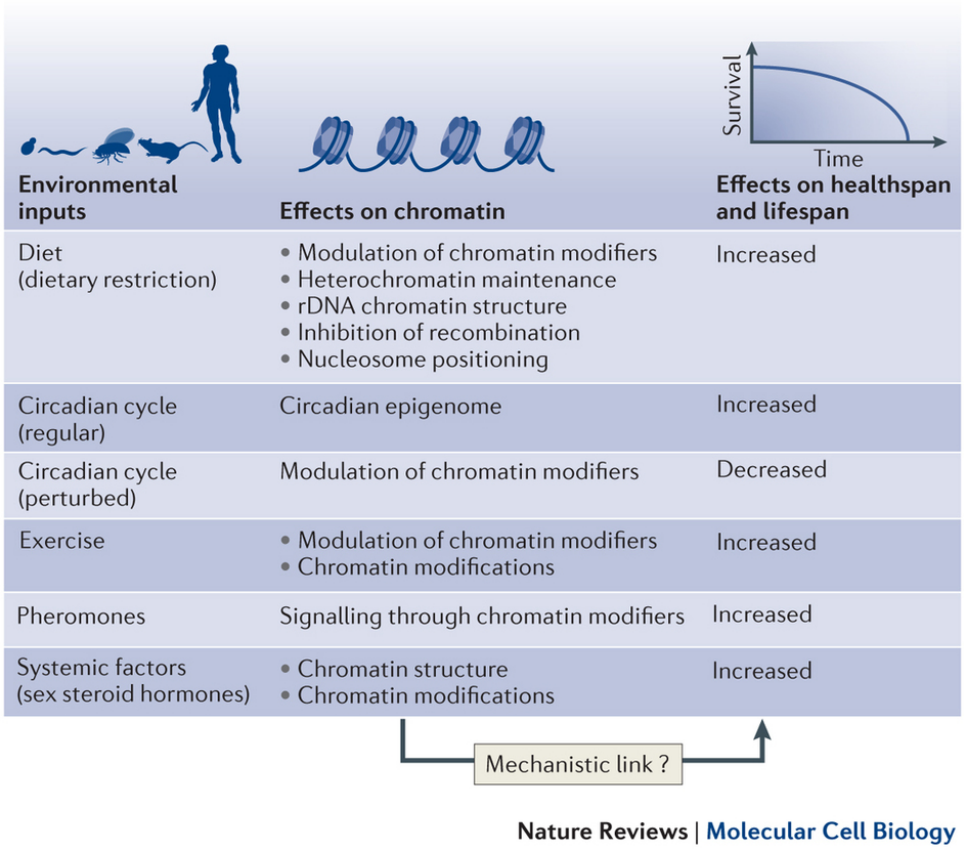


FIGURE 2 – Environmental signals that modulate lifespan might do so by modulating chromatin. Dietary restriction increases lifespan in a range of organisms and has also been linked to changes in the chromatin landscape. These include changes in the expression of chromatin modifiers (for example, increased expression of several sirtuins) or increased maintenance of heterochromatin. Robust circadian light cycles also promote healthspan and lifespan and are linked to circadian epigenomic changes (for example, periodic increases in H3K14ac of circadian promoters by the circadian CLOCK proteins) and modulation of the activity of chromatin modifiers such as NAD-dependent protein deacetylase sirtuin 1 (SIRT1). Physical activity is also beneficial to healthspan and lifespan and has been associated with changes in chromatin modifications (for example, increased H3K36ac in human skeletal muscle) and with the regulation of chromatin modifiers (for example, induced nuclear exclusion of histone deacetylase 4 (HDAC4) and HDAC5 in human skeletal muscle). Recent work has shown that, in *Caenorhabditis elegans*, pheromones may increase lifespan through a mechanism requiring chromatin-modifying enzymes. Finally, in women, the strong decrease in the production of sex steroid hormones (such as oestrogens) with age contributes to age-related diseases, and oestrogens can directly remodel chromatin at target genes through their receptors. Whether there is a linear pathway from the environmental output to the changes in chromatin and, ultimately, to healthspan and lifespan extension, remains untested. From (23)

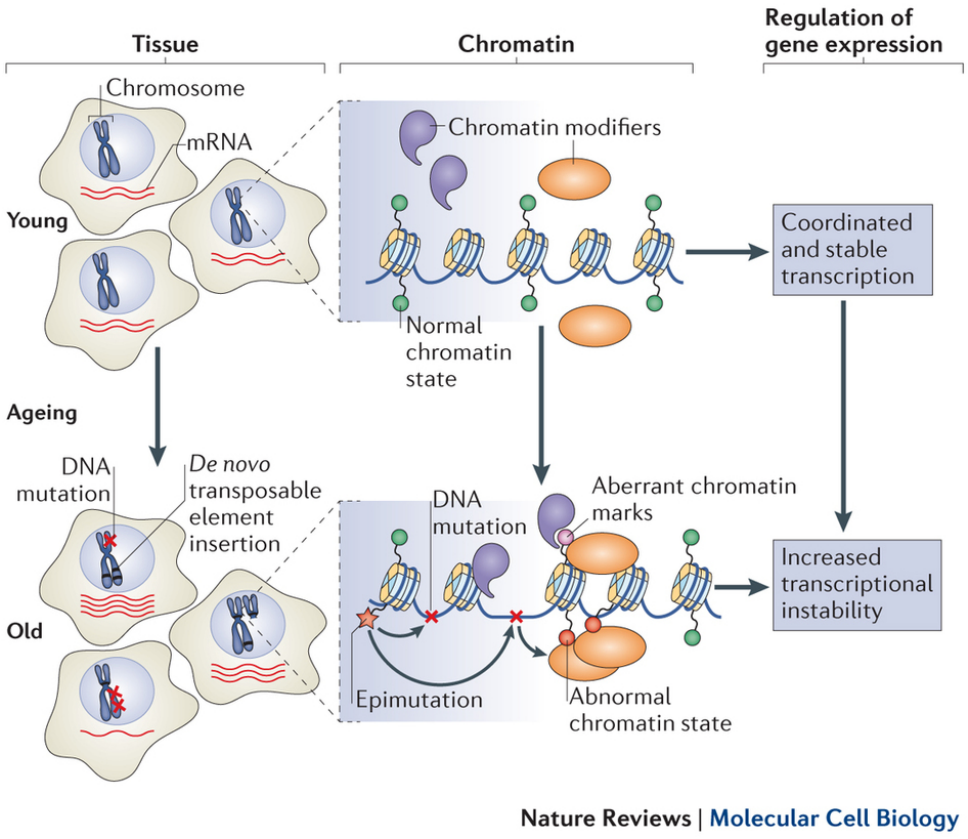


FIGURE 3 – Aging : In cells from young organisms (top left), transcriptional programmes are robustly defined and precise between cells (depicted by consistent mRNA levels). Genomic integrity is maintained (depicted by intact chromosomes), because mutations are rare or correctly repaired. As a result, "normal" chromatin states are found throughout the genome. With increased age (bottom left), transcriptional instability is increased among cells of a tissue (depicted by variable mRNA levels among cells). Genomic instability is also a hallmark of ageing and is increased both at a macro level (for example, aneuploidies, depicted by partial chromosome duplication, or increased transposable element insertions) and more locally by DNA mutations in the form of single-nucleotide mutations or small insertions or deletions. DNA damage can trigger the recruitment of chromatin modifiers and the acquisition of abnormal chromatin states. Thus, genomic instability could modify the epigenetic landscapes of old cells. Reciprocally, aberrant changes in epigenetic marks, known as 'epimutations', can further promote the accumulation of DNA mutations in a feedback-loop mechanism. The epigenetic changes that are acquired during ageing could also decrease the transcriptional precision of neighbouring genes. From (23)

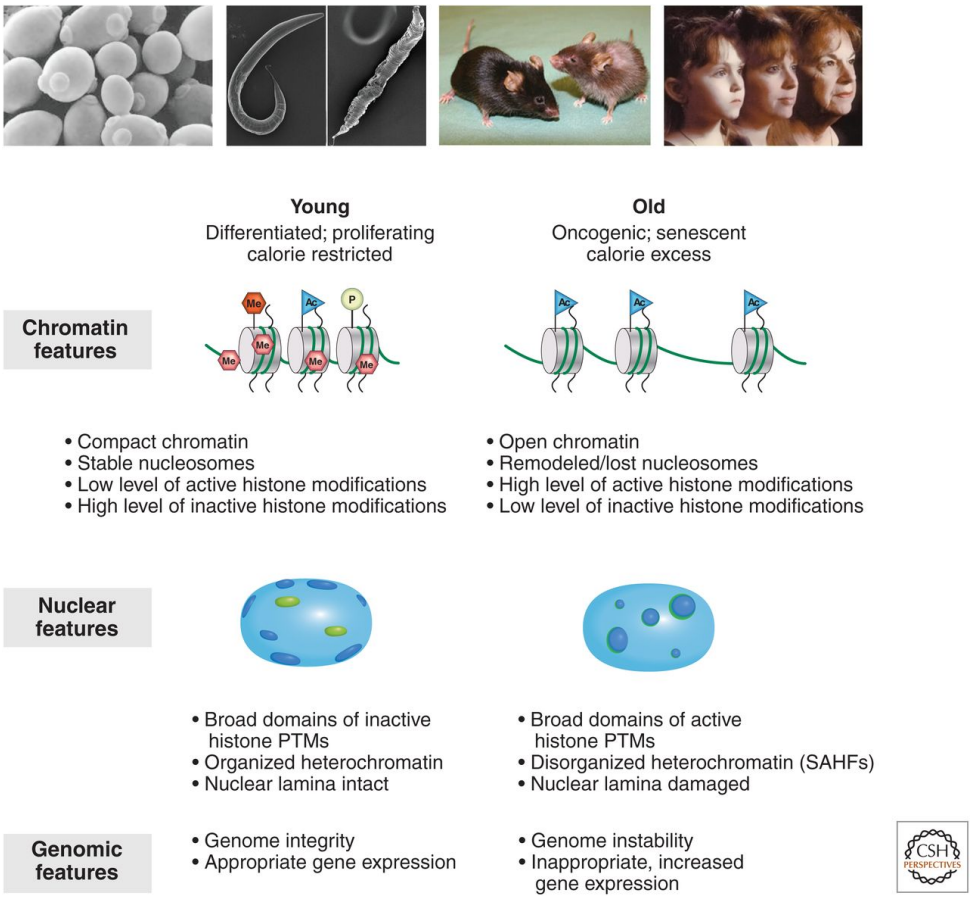


FIGURE 4 – Features of aging and senescent cells. (Top) Aging is a biological inevitability for all eukaryotes, from budding yeast and C. elegans to mice and humans (depicted). The features of senescence and cellular aging are compared between a normal or "young" cell (left) and a diseased or "old" cell (right) at three levels : chromatin, nuclear organization, and integrity of the genome. From (22)

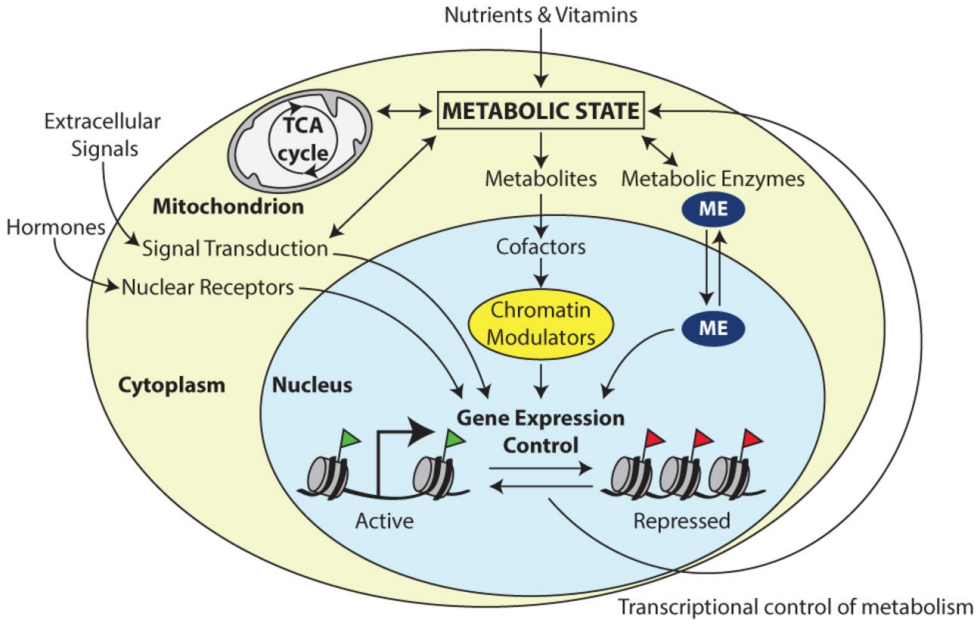


FIGURE 5 – The interface between metabolism and transcription. Specialized transcription factors that are activated by metabolic sensors or extracellular signals, such as hormones, direct a transcriptional response to changes in metabolic state. In addition, key components of intermediary metabolism are cofactors or cosubstrates of chromatin-modifying enzymes. Therefore, changes in cofactor availability may affect chromatin structure and gene expression. Finally, specialized metabolic enzymes (MEs) do double duty as regulators of chromatin and transcription. Frequently, cytoplasmic-nuclear partitioning is used as a regulatory mechanism to link metabolic state to transcriptional outcomes. From (24)

quantified yet, it is becoming more and more clear that the dynamic interplay between cellular metabolism and gene transcription regulation, in normal proliferative cells, in diseases (such as cancer) and aging (including cellular senescence). Both cellular nutrient metabolism and chromatin organization are remodeled in cancer cells, and these alterations play key roles in tumor development and growth.

1.5 *From 3D to 1D : Spatial chromatin compartments and the "nano-reactor" hypothesis*

Functional importance of spatial compartments : increasing the local concentration

As discussed in the "Rapport d'Activité", the spatial organization of chromatin results in part from the clustering of epigenomic chromatin states but a still open question is whether this spatial organization is only a by-product of (epi-)genome activity or is also participating to the regulation of the epigenome assembly and more generally to the regulation of the genome function. The basic concept behind this structural/functional coupling is the increase of local concentration of regulatory proteins due to spatial co-localization. This paradigm has been actually evidenced and formalized for many years in the context of the well-known lac operon system. Molecular crowding and spatial confinement increase the binding affinities of regulators (activators and repressors) to their chromatin/DNA targeted regulatory sequences (ibid). In some sense, the nuclear compartments would correspond to biochemical nano-reactors where a few number of reacting biomolecules are co-localized in space favouring their biochemical (co-)activity on chromatin and in fine on DNA.

In the lac system, the presence of few additional dispersed recruitment sequences (operators) and the ability of the lac-repressor to oligomerize enhance the association of a repressor to the effective "repressing" site. In eukaryotes, similar strategies are acting at the level of enhancer-promoter modules where the action at "distance" of the enhancer sequences are conditioned to their physical contacts with the promoter : as for the lac system, distal enhancer sequences might actually act as secondary recruitment sequences for TFs that, by associating with mediator and other architectural proteins, can promote recruitment and stabilization of the transcriptional machinery at promoters via long-range looping and clustering.

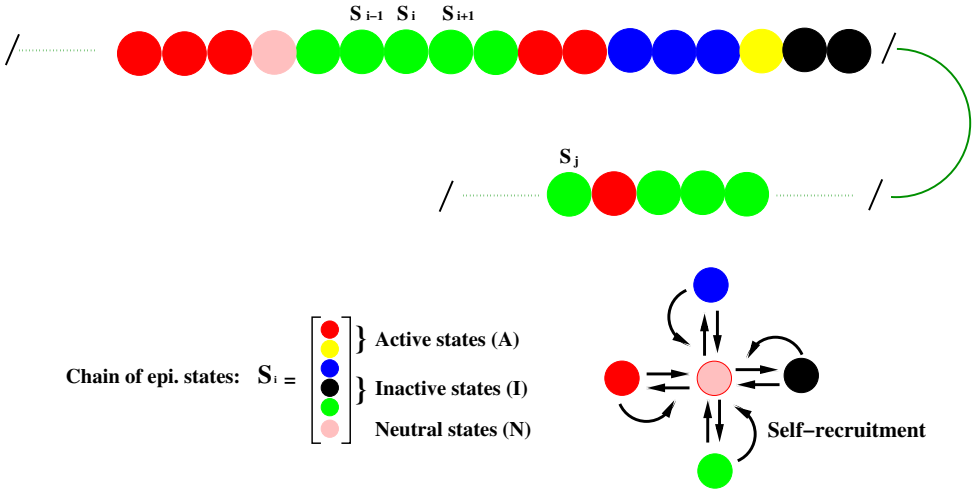


FIGURE 6 – Effective spin-like model of epigenome assembly

Along the same line, in drosophila, PcG-mediated gene repression involves the spatial colocalisation of the silencer elements PREs into 3D compartments, the so-called PcG bodies, mediated by the Polycomb proteins. Such clustering operates "in cis" ie within an epigenomic domains but also "in trans" between distant domains as for example between the ANT-C and BX-C domains (white circle in Fig. 1A) where the level of repression has been directly correlated to the level of clustering between PcG domains. Same colocalisation mechanisms of PcG repressed genes are also observed in mammals. In budding yeast, repression by the SIR system, has been linked to spatial clustering and perinuclear anchoring of Sir-bound telomeres at the nuclear membrane. And a similar coupling between clustering and repression at the nuclear enveloppe also operates in higher eukaryotes such as worms and mammals, wich has been recently remarkably evidenced in the chromosome-wide inactivation process of the X chromosome.

The mechanisms that drive this nano-reactor formation has been discussed before : the polymeric nature of chromatin induce a "natural" confinement since dispersed sequences on the same chain have a greater probability to colocalize due to chain looping. Every process that promotes this looping probability also enhances local confinement. In particular the multimerization of regulatory DNA binding proteins can promote physical bridging between enhancer and promoter, and between silencers (Fig. 6A). Additionally insulator proteins such as dCTCF associated with cohesins, may contribute to the structural but also selective confinement of active/repressive modules by forming "insulated neighborhood". At larger scale, TADs that can be either constitutively or facultatively formed during development contribute also to the confinement of the "sub-TADs" modules, providing a "basal" (large scale) level of confinement and of selectivity that are then finer-tuned at lower scale within sub-TADs modules. Implication of TADs in regulating transcription has been also recently proposed in the process of mammalian X inactivation : consistently with this nano-reactor hypothesis, the expression of the Tsix transcript was positively correlated with the compaction level of its embedding TAD. And more generally, recent studies have shown that perturbing TADs integrity may indeed lead to transcriptional deregulation and diseases.

2 GENERAL OBJECTIVES

Altogether, these experimental data reveal a clear link between the epigenome, the 3D architecture and the gene transcription pattern. However, it is still unclear what are the mecanisms behind one-dimensional epigenome assembly and its three-dimensional folding, and how these 1D and 3D organizations precisely contribute to epigenetic gene regulation. The general objective of this long term project is, in continuity to my previous and current research activities, by using theoretical approaches from equilibrium and non equilibrium statistical physics, to identify and model some chromatin-based mechanisms of gene regulation involved in cellular development : from first principles and with a minimum (ie tractable) parametrization, can we account for

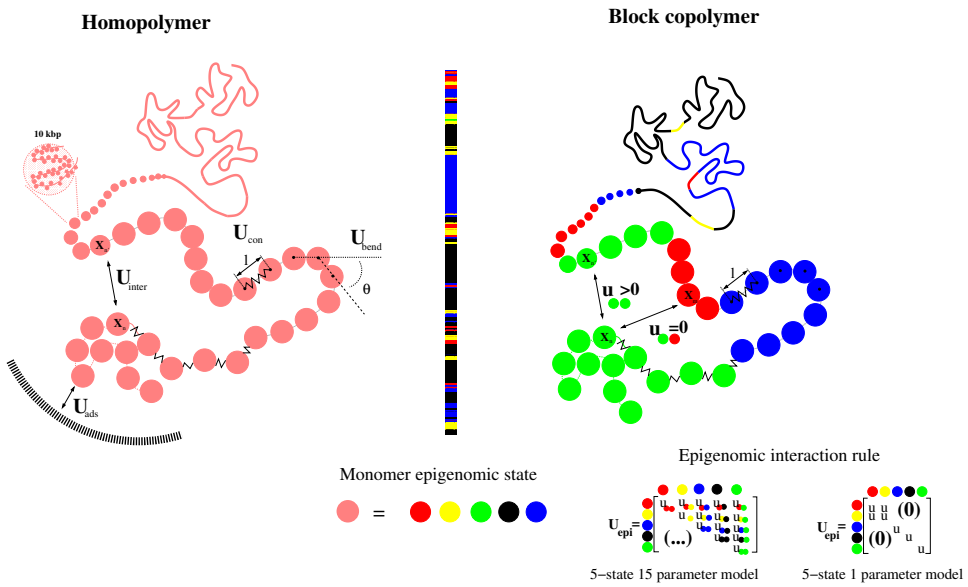


FIGURE 7 – Effective model of chromatin folding.

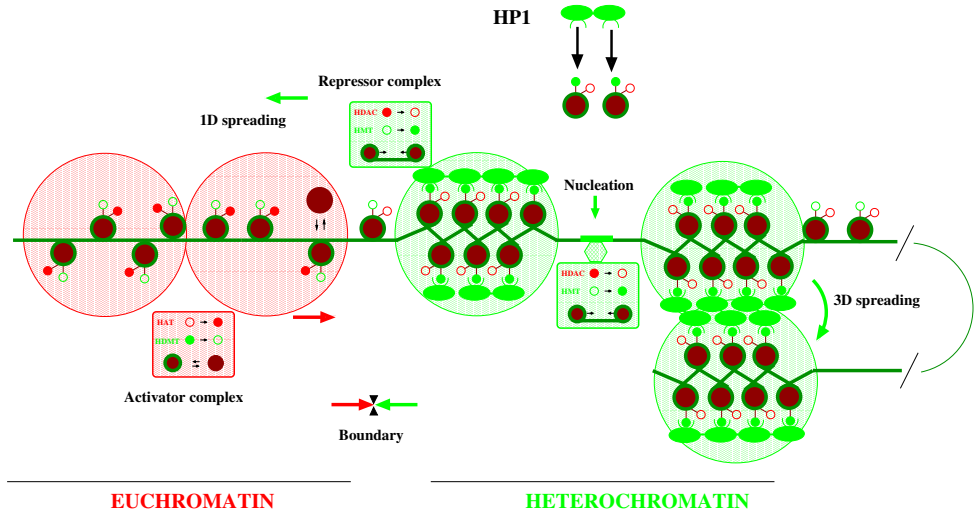


FIGURE 8 – Molecular model of epigenome assembly (and folding)

the different chromatin states, their spatial folding and nuclear organization ? For their dynamics during cell cycle , differentaition ? Can we understand how these 1D & 3D organizations of the genome contribute to stably activate or repress genome activity ? How chromatin regulatory pathway "sense" environmental changes, in particular metabolic changes, and, in some case, translate transient stimuli into stable, inheritable expression pattern changes while in other cases act as a "buffer" of external as well as genetic/genomic perturbations ?

There are two main axes that are in continuity with my previous and current research activity and a third one, more exploratory :

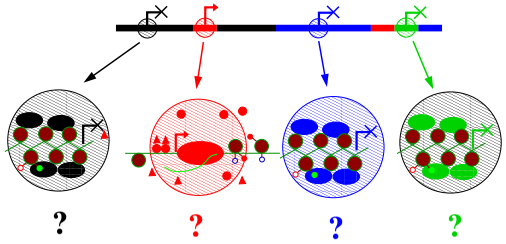
- **Epigenome assembly and folding : "living chromatin" and the nano-reactor hypothesis (Fig. 9)**

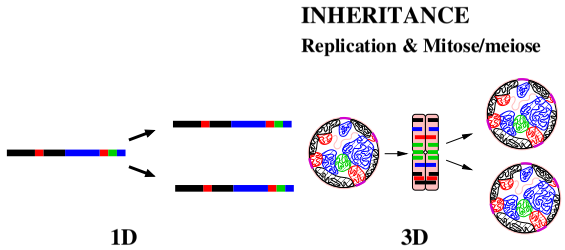
  Question : Can we account for the differents epigenomic domains and of their spatial organisation observed in the various cell types and organisms ? What controls boundaries between 1D & 3D domains ?

  Based upon current *in vivo* and *in vitro*-derived experimental knowledges, we will introduce local molecular models (Fig. 8) to investigate the macromolecular structure and dynamics associated to the different types of chromatin. In a thermodynamical framework, as a function of relevant control parameters (i.e. chromatin regulators activity, architectural proteins availability, binding strength, interactions strength, effective temperature) we will compute the equilibrium distribution of chromatin states along genomes. For a given genomic domain, we will explore the paramater space to show conditions at which thermodynamically distinct chromatin states (heterochromatin/euchromatine/intermediate) can occur at equilibrium and thus delineate regions of monostability (i.e. where there is a single stable chromatin state) and regions of multistability (ie where different stable chromatin states can occur).
  Additionnaly, we will further investigate the coupling between epigenome assembly and the 3D organisation of chromatin. We will refine and extend the simple block copolymer model (see below) that we already developped ; in particular, the idea is to introduce finer scale epigenomic variability as well as other constraints such as the ability of anchoring at the nuclear membrane and of long-range bridging (site-specific looping) between regulatory sites. Topological constraints (excluded volume) that are likely to control large-scale organisation of chromosomes will be treated more accurately. Using analytico-numerical mean-field approaches and efficient numerical simulations we will compute equilibrium phase diagrams. However this model states that the epigenome is "quenched". The main task of this project will be to relax this hypothesis and to develop a theoretical framework that couples the dynamics of epigenome assembly and its 3D organisation : the "living chromatin" framework. The principle behind this, is that there is a self-reinforcing positive feedback loop between 1D and 3D organization which mainly relies, for example in the case of heterochromatin, on the oligomerization and self-recruitement ability of repressive complexes (26, 27, 28, 29) : the 3D compartimentalization of chromatin might be a way of coordinating and reinforcing the functionnal output (repression, activation) of genomic regions by colocalization and mutualization of the same specific regulators (Fig 11), ie by the self-assmebly of a "nano-reactor" (30, 31). Initiation (nucleation) for de novo activation/repression may be performed via the primary targeting of activating/silencing complexes to specific genomic sites mediated by DNA binding proteins (or by non-coding RNAs) which is then followed by the coupled self-assembly of the chromatin and structural state that further enhances activation/repression and perpetuates the active/repressive environment throughout cell division. This may correspond to a conversion from (i) a "hard-wired" targeting of regulators to few discrete recruitment and bridging genomic sites at the induction stage (in response to developmental or environmental cues), to (ii) a "soft-wired" targeting of regulators with weaker interactions but associated to a larger number of spatially concentrated secondary sites at the maintenance stage (no more or lower cues). In that context, the structural/functional nuclear compartments would correspond to the self-assembly of a robust nano-reactor where the biochemical reactions (DNA/chromatin binding of regula-

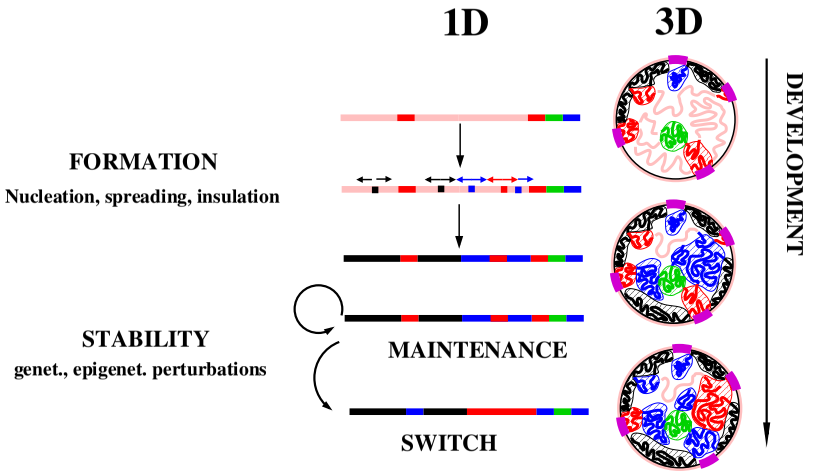
QUESTIONS:

1– STRUCTURE AND MODE OF REPRESSION & ACTIVATION ?





2– REGULATION OF EPIGENOMIC DOMAINS ?



1D/3D COUPLING ?

FIGURE 9 – General objectives

tors, multimerization of regulators, spreading of marks, transcription, replication...) would both depend on and induce (reinforce) their spatial confinement. The "living chromatin" model might be a first attempt to formalize this nano-reactor hypothesis in the context of epigenome regulation. Along the same line, it has been recently suggested by Katada *et al.* (**?**) that chromatin compartments ("microdomaines") may indeed constitute functional "niches" (either active or inactive nano-reactor) by increasing the spatial colocalisation of metabolites with their co-enzymes.

- **Epigenome dynamics : Inheritance (maintenance) or Switch (reprogramming) (Fig. 9) ?**

  Question : Can we account for the regulatory pathways driving chromatin state and folding transitions during development ?

  Having identified the "equilibrium" configurations, we will investigate their stability to various perturbations such as molecular noise, genetic polymorphism, or environmental perturbation. In our thermodynamical framework, molecular noise account for the effective temperature : for given control parameters values (chromatin regulators activity), it controls the "intrinsic" fluctuations of the chromatin states ; genetic polymorphisms may account for genome modifications (insertion of parasitic DNA, mutations in regulatory sites) or mutations in the regulators-encoding genes : this might induce modification of the phase diagram. Environmental perturbation would mainly account for chromatin regulators activity modulation and thus for changes in the control parameter space. Formally, this mean that we will study the relaxation dynamics of a given equilibrium state following a change in the control parameter : what changes (of what amplitude and duration ?) are needed to induce a transition between two equilibrium states ? What is the corresponding kinetic and conformational path during the transition ? What is the role of spontaneous fluctuation (noise) in driving transitions ? Of particular interest dor differentiation process will be the study of multistability which indeed accounts for plasticity but also for "memory" effects ie irreversible switch as well as for glassy dynamics (cf our work on epigenome folding).
  As a main task here, we will investigate the influence of replication on chromatin state stability and maintenance, since replication is a periodic strong transient perturbation of the epigenome. The understanding of how chromatin states are redistributed and mainteianed(or not) after replication is obviously at the heart of epigenetic inheritance. This theoretical investigation is this of outstanding interest and importance. Provided a molecular model of chromatin state "redistribution"(32), we will study the non equilibrium dynamic of chromatins states when periodically submitted to replication-induced rearrangement. Depending on the "transmision" model (equirepartition of chromatin states between the two replicated strands, biased transmisssion due to "spatial" effects....) and on the duration time of cell cycle, some equilibrium states might be driven through transition or reversely, consolidated.

  The objective is to perform this theoretical study in the "Living Chromatin" framework, ie when both 3D and 1D dynamics are coupled.

- **Cellular Signaling : Metabolic signaling to chromatin**

  Question : How epigenome "sense" environnment ?

  Having formalized the generic principles driving the regulation of the epigenome a final question would be : what regulates epigenome regulators ? How indeed environmental signals such as temperature, mechanical and nutrient availability changes can be translated into stable modification of the epigenome and consequently modifications of transcriptional program ?
  As already pointed out, chromatin regulation involves enzymes that use cofactors for the reactions that modify DNA or histones. These enzymes either attach small chemical units

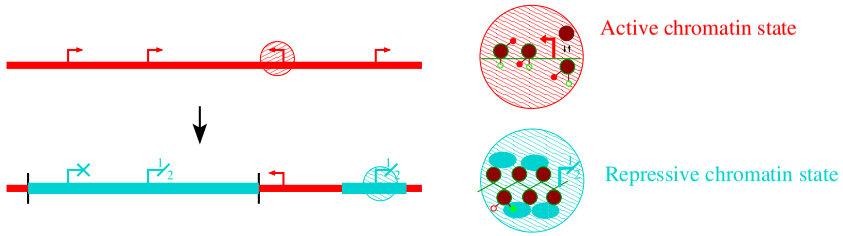
(i.e., posttranslational modifications or PTMs) or alter nucleosome positioning or composition (i.e., of histone variants). It is assumed that this control depends partly on the variable levels of cellular metabolites acting as enzyme cofactors. For example, acetyltransferases use acetyl-coenzyme A (acetyl-CoA), methyltransferases use S-adenosyl methionine, and kinases use ATP as donors of acetyl, methyl, or phospho groups, respectively ; deacetylases can use nicotinamide adenine dinucleotide (NAD), and demethylases can use flavin adenine dinucleotide (FAD) or α-ketoglutarate as coenzymes. In addition, another relevant example relates to remodeler complexes that use ATP for moving, ejecting, or restructuring nucleosomes. As such, epigenome alteration might somehow reflect metabolism alterations and conversely, metabolisms alteration might result from misregulation of epigenome or/and mutations at some metabolic genes. In addition, such mutations can also be linked to altered epigenome. All this strongly support an intricate interplay between epigenome/metabolome/genome dynamics. Our objective will be to formlize this coupling using effective modelling approaches. In particular we will investigate the effect of metabolite restriction (including ATP), and chromatin enzymes defects in epigenome assembly and maintenance.

3 MODELS AND THEORETICAL APPROACHES

3.1 "1D" : Epigenome assembly Current model of epigenome assembly mainly consider a spin-like chain models (Fig 6) where nucleosomes can adopt 2 or 3 "internal" (or more) epigenomic states (Active/Represssive/Unmarked) (33, 34, 35). Then dynamics of conversion (from Active to Repressive....) is controlled by the interplay between a random interconversion (kind of effective temperature in spin analogy) and guided conversion (coupling factors in spin analogy) between the states of the nucleosomes ; it is assumed that nucleosomes are fixed along a given genomic domain. Large-scale (at the domain scale) equilibrium and dynamical properties are then computed via stochastic simulation or in a mean-field approximation to extract phase and bifurcation diagrams as a function of coupling factors strength. Recently we have extended the work of Daniel, by investigating the effect of replication on epigenetic memory and on gene expression : in particular, due to the fork replication progression and the subsequent disruption of repressive state, there is a lag time needed for the epigenome to reassemble when bursting of gene expression may occur (35). This might offer a "window of opportunity" for spurious transcription, for exemple for repeated sequences and mobile DNA elements. The effect of replication will be further tested in a context of "finite ressources" (titration) for the the availability of the different chromatin states : indeed, depending on the timing of their replication, some part of the genome might maintain their chromatin state more easily by consuming the first the available ressource.

The next idea is to now refine these approaches by introducing local molecular models of the different chromatin states and of their dynamics of formation : nucleosome positioning and turn-over, binding of proteins such as HP1, modification of the histone-tails .... (Fig. 8) will be included to define local molecular and epigenetic states. The work concerning the improvement of the nucleosome positioning model presented in the "Rapport d'Activité" will be performed in that context of molecular modeling of epigenome. This, of course in order to better account for the complexity of local epigenome regulation. We assume indeed here that a minimal additional layer of complexity can be introduced (in comparison to the "spin-like" models) in a still easy tractable way (ie without not so much computational effort while keeping some genericity). Moreover, high-resoluted biochemical mapping allow now to build such molecular models.

3.2 "3D" : The copolymer framework. *Folding of an homopolymer.* A generic and minimal model to investigate the large-scale 3D organisation and dynamics of chromosomes is to consider chromatin as a semi-flexible self-avoiding homopolymer. The conformation of such polymer is controlled by the interplay between thermal motion, steric repulsion and effective monomer-monomer interactions that may account for steric confinement, true attractive or repulsive interactions between monomers or interactions mediated by the solvent. For example, in a "good solvent" condition, i.e. when steric repulsion dominates, the chain adopts a swollen coil state with a typical size that scales with the number of monomers as $R_g \sim N^{3/5}$ and the contact frequency is given by $P(s) \propto s^{-9/5}$ with s the linear distance between monomers (36) ; in a "poor solvent"

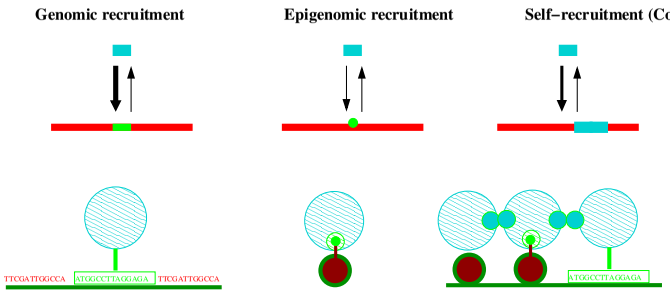


Questions: **1) Mecanismes of activation/repression ?**

2) Mecanisms and dynamics of assembly ?

Epigenome assembly:

Key Principle: Site−specific recruitment + Self−recruitment of repressif chromatin state



3 PHASES:

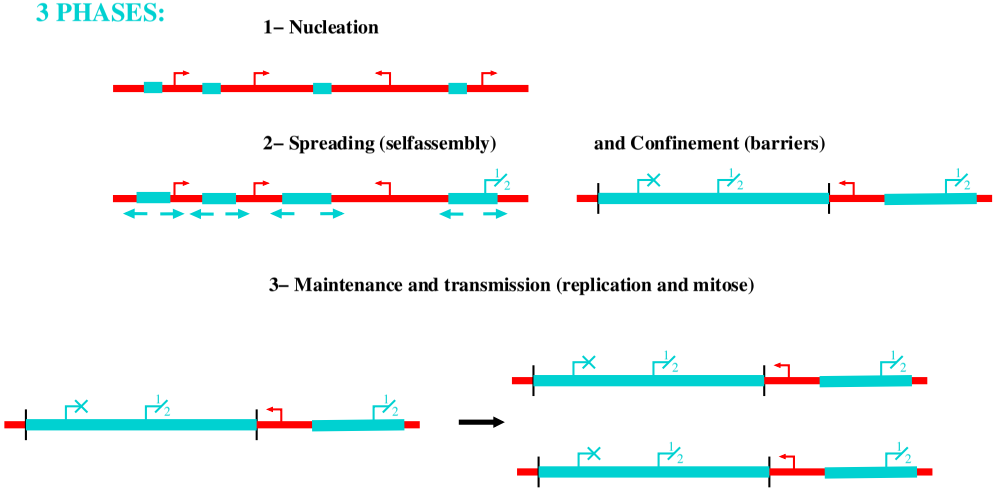


FIGURE 10 – The "1D" problem

condition, i.e. when effective attraction prevails, the polymer chain becomes compressed and equilibrates into a collapsed globular phase characterized by a high monomer density, a typical size $R_g \sim N^{1/3}$ and a contact frequency that is almost constant at large separation distance ; in between, i.e. at the transition point (the so-called θ temperature) where repulsion and attraction counterbalance, the chain adopts a gaussian coil state with $R_g \sim N^{1/2}$ and $P(s) \sim s^{-3/2}$.

Such homopolymeric model with appropriate geometric constraints provide a fairly good description of large-scale conformational properties from yeast to human and in particular a good fit of different scaling behaviours like the dependence of the contact frequency or of the mean distance on the linear genomic distance. In yeast, static and dynamic studies suggest a brush-like equilibrated organisation (37, 38) of chromosomes with scaling properties compatible with a weakly collapsed state ; in fly and human, where chromosomes are much longer, large-scale organisation is characterized by a dense state and a contact probability $P(s) \propto s^{-1}$ which has been primarily associated to a non-equilibrium dense phase composed of spatially segregated long-lived globular domains (39) but which can also be mapped to a semidilute solution of non concatenated rings (40) at equilibrium. A recent paper by Barbieri et al. (41) has shown that this fractal-like state could also be described by introducing effective attraction via binding molecules in an homopolymer model.

Folding of block copolymer. However most of these models consider an homopolymer that cannot obviously account for the regional variability of the chromatin organization and in particular for the spatial compartmentalization of the epigenome. One has to introduce a genomic and/or epigenomic specificity in the folding model. In a recent collaboration with G. Cavalli (42) we posit that chromatin folding is driven by effective epigenomic-dependent interactions between chromatin loci. To test this hypothesis, we propose a generic and new theoretical approach by treating chromatin as a block copolymer, where each block corresponds to an epigenomic domain and where each monomer interacts preferentially with other monomers of the same chromatin type. This is largely motivated by the observations of self-interactions between chromatin type (1) and is also supported by increasing evidence that some architectural proteins might promote physical bridging (29, 28). As already mentionned, we adapt a Gaussian self-consistent approach (43, 44) to derive equilibrium phase diagram and contact maps for any block copolymers, as a function of two control parameters, the specific and non specific monomer-monomer interaction. When considering block copolymers build from the epigenomic landscape of Drosophila (2), we have shown that this simple physical model accounts very well for the folding patterns in TADs observed in Hi-C experiments (1). As a main and very original outcome, this model provides a physical basis for the existence of multistability in chromosome organization. We show indeed how some experimental patterns are fully consistent with multistable conformations where topologically-associated domains of the same epigenomic state interact transiently or long-lastly with each other.

Further extension and improvement of the block copolymer model

In our recent work (see "Rapport de Recherche") , we were interested by identifying some general principles governing chromatin folding. Therefore, for simplicity, we limit our approach to the simplest version of a block copolymer that we can build from the compartmentalization of the epigenome. It assumes that all specific interactions have the same strength whatever the chromatin type and whatever the location along the genome. However such two-parameter model needs to be refined in order to gain in predictability. In particular, the current model considers each domain as an homopolymer with uniform monomer-monomer interactions. This can only provide a coarse-grained understanding of epigenome folding and cannot account for variation in contact frequency and in particular for preferential pairwise (long-range) contact between discrete genomic loci. Recent studies (9, 10) have proposed that these site- and lineage-specific contacts mediated by architectural proteins (insulators, cohesin, mediators) might indeed play a key role in the folding of chromosomes at the sub-Mbp scale. Along the same line, anchoring at the membrane of particular sequence or/and epigenomic domains (via their association with lamina or nuclear pores) has been shown to be crucial for regulating both folding, expression and stability of the genome(45, 46, 15, 47). In addition to the global, non site-specific, interactions investigated by our current model, focal large-scale looping and anchoring might indeed contribute to spatial

compartmentalization of domains (1, 10, 48). Therefore, further improvements of the model will require to augment the number of parameters by allowing for variability of interaction at the monomer scale, and to infer specific interaction strengths that predict at best the observed contact maps (see below).

Moreover, we made the assumptions that the polymer chains are equilibrated. Detailed simulations of confined self-avoiding homopolymers (40) have suggested that strong topological constraints may slow down the chain dynamics such that equilibration of very long confined chains can be extremely long (40), well above the cell cycle length. However, recent theoretical studies (Rosa & Everaers, personal communications) have estimated that below few Mbps, the typical length scale investigated in our preliminary, topological confinement is negligible and polymer chains may safely be considered as equilibrated. This is no more the case if considering the full genome (except for small genomes like yeasts).

However, it is clear (even at short length scale) that the self-consistent approach can only give an approximate solution of chromatin folding due to the rough treatment of excluded volume. The next step will be to use efficient full mumerical simulation such as the one developped by P. Carrivain who was post-doc in Cavalli's lab and is currently in our group under the supervision of R. Everaers and myself. Pascal is working on simulating the nuclear organisation of the complete diploid genome of Drosophila : starting from an initial condition (the so-called Rabl configuration) that corresponds to the chromosome organisation at the entry of the cell cycle (after mitosis) in this specie, he is currently simultaing the relaxation dynamics toward equilibrium (Fig. 13) for a simple generic model that only takes into account geometrical and topological constraints ; the comparison with recent Hi-C maps obtained in the Cavalli's lab at different stages of embryonic developement as well as for mutants (see ANR project below) will allow us to decipher to what extend we may need to introduce additional specific epigenomic-driven interaction between monomers or between monomer and the nuclear membrane. As shown in 13, we have already very good results that we are aiming to publish very soon.

Development of a fast numerical method for simulating nuclear organisation of chromosomes. Using polymer physics and numerical simulations, we recently showed that attractive interactions between loci of the same chromatin state might be the driving forces of the folding of chromatin inside the nucleus in the fly *D. melanogaster* (42, 49). We also showed how interactions of chromosomes with the nuclear membrane may explain the coupling between global gene regulation and 3D compaction of X chromosomes in the worm *C. elegans* (50). However, in these preliminary studies, we restricted our modeling to relatively small systems (subparts of chromosomes or chromosomes at a very coarse-grained resolution). With the recent advances in experimental methods that can now produce high-resoluted structural and epigenomic genome-wide datas, there is a clear need to perform simulations at the full genomic scales in all species ranging from yeast (12 Mbp) to human (3 Gbp). Improving the actual numerical tools in order to significantly reduce computation time is thus a crucial prerequisite that will allow to refine our understanding of the complexity of "spatial" regulation of the genomes.

For that purpose we have already developed a lattice polymer model with kinetic Monte-Carlo simulations (49). The objective will be to improve this method by implementing a numerical scheme that has been developed few years ago by Ostrovsky et al. (Parallel Computing **27** : 613-641 (2001)) ; using this algorithm we expect, in a serial implementation, a significant > 10 fold decrease of equilibration times for a yeast genome (from 3 mn to 20 s) and a $> 1O00$ fold decrease for the human genome (from few days to few minutes) ! Full parallelization will be implemented and choice between CPU vs GPU architecture will be also tested. By achieving such optimizations we will make a clear breakthrough in that booming field of genome 3D organisation.

Inference of structure and potentials from Hi-C maps. The copolymer framework associated with the self-consistent Gaussian approximation may represent an efficient formalism to extract from the available experimental data the effective genomic and epigenomic interactions between chromatin loci. As a promising outcome of such inference process, would be a powerful tool to predict the chromatin organization in various conditions, allowing to investigate *in silico* changes in TAD formations and long-range contacts when altering the epigenome. Hence our model may provide a very interesting framework for understanding how epigenome regulation (resp. dysregulation) during development (resp. disease) could lead to cell phenotypic variations

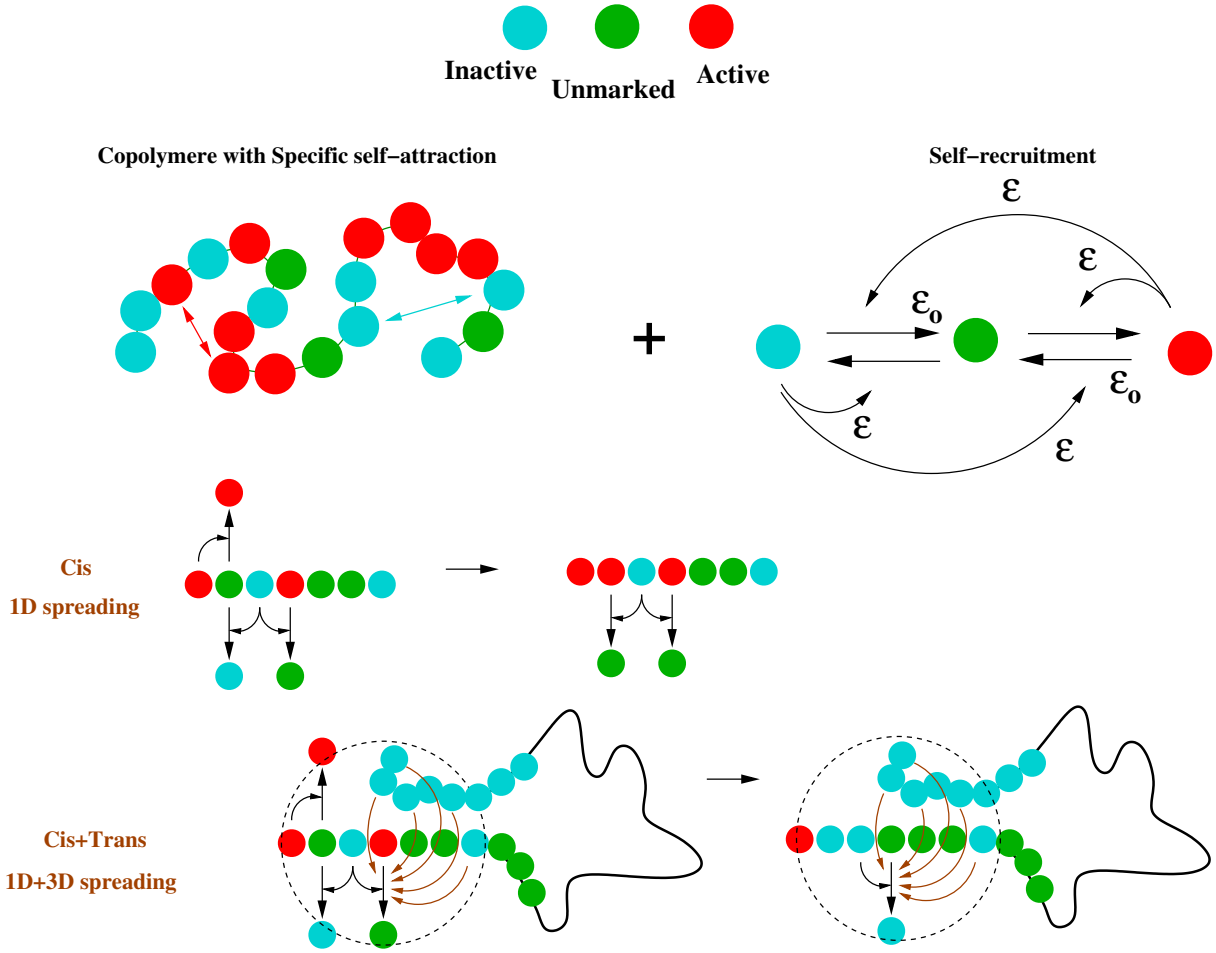


FIGURE 11 – The "Living Chromatin" model : (Top) The living chromatin model is a combination of the copolymer model (42) and of epigenome regulation model (33, 42). Each monomer can be in one of the 3 states : A, U and I ; the inter-conversion dynamics between this states result either from "noisy" conversions (with rate ϵ_o) or "recruited" conversion (with rate ϵ). The chain is modelled by a bead-spring model with specific attraction between monomer of the same active or inactive state. (Bottom) Recruited conversion is achieved either by recruitment *in cis* (1D spreading) or by recruitment *in trans* (3D spreading)

via large-scale chromatin reorganization. This inference (based on bayesian inference) was a project carried out by Noelle Haddad during her PhD thesis (51, 31). As shown in the "Rapport d'Activité", promising results are already obtained. However, the input of this inference approach was the contact maps and we are now building aa approach that will start with the epigenomic domains instead : indeed, the working hypothesis is that pair-wise self-attraction mainly depend on the epigenomic states, such that we aim at providing the minimal model of epigenome-driven attarction that acounts at best for the experimental observations. Then, we have to make the connection between these GSC dervide parameters and those of the more realistic models that we are using for thewhoel genome simulations. We will be able then make prediction of the 3D folding of altered epigenome of for other drosophila species.

3.3 Coupling 1D and 3D : the "Living Chromatin" approach and the "nano-reactor" hypothesis

Local epigenome assembly relies on the combined action of various enzymes ("chromatin regulators") that are primarily targeted at given DNA sequences to produce either a permissive (by activators) or a repressive (by repressors) transcriptional state around this "nucleation" sequence. The ability of these regulators to be secondarily targeted at chromatin fragments carrying the epigenomic state that they catalyze increase the cooperativity of their action leading to the 1D spreading (in "cis") and to an enhanced robustness (stability) of the corresponding epigenomic states (that can be maintained even in the absence of the initial nucleation process). In a sense

this molecular cooperativity balance the small absolute number of regulators that are effectively acting (or even produced) by an increase of the effective local concentration of enzyme. However chromatin is not a 1D object and fragments that are genomically distants can be colocalized spatially leading to a 3D spreading ("in trans") of the epigenomic states : the enzymes recruited at a given chromatin fragment can catalyze the states of all chromatin fragments that are in the spatial vicinity (and not anly genomic vicinity). It would correspond to a 1D system (like spin chains) with long-range interactions leading to an even higher cooperativity (again because of an enhanced local concentration of enzymes). The abovementioned 1D models (Dodd et al. (33), Jost (34, 35)) include long-range interaction in a mean-field way and our idea is to couple our "structural" polymeric approach (used in our work concerning epigenome folding) (Fig. 7) and their dynamical spin chain model (Fig. 6). Full numerical simulations will be developped by implementing kinetic Monte-Carlo algorithms for the polymer and the chromatin states. Within such model, the "spin coupling" constant will depend on the spatial proximity and as before, specific volume interactions between monomers will depend on their epigenomic state (Fig. 11). Thus, the very interesting and original feature that we will be able to study is the "living" feature of chromatin : due to the ability of chromatin fragments of same epigenomic states to cluster in space, the spatial organization will be dynamically coupled to the 1D assembly of the epigenomic states. This correspond to a mechano-chemical coupling that will further increase the cooperativity of the enzymatic action of regulators and thus the robustness of the epigenomic domains. As discussed in (30, 31),

We have already obtained significant results and we are currently writing an article on that topic. We have developepd a kinetic Monte Carlo algorithm to simulate the combined dynamic of the local epigenomic states and of the polymeric chain on a FCC lattice (see above). As a main outcome, we have shown how chain compaction can induce bistability (phase transition from a monostable state to a bistable state) (Fig. 12) and how compaction may help to maintain epigenome compartmentalization. It is notably at the heart of the project with Peter Meister concerning dosage compensation of X chromosomes in *C. elegans* (see below). But the scope of application of this model is much wider and it is likely a very well suited framework for studying the Hox gene co-regulation systems among others (see below).

4 EXPERIMENTAL SYSTEMS AND COLLABORATIONS

4.1 Discovery and modeling of epigenetically regulated genomic domains in lung cancer

Granted INSERM project (Research projects in the field of multidisciplinary approaches in modeling complex biological processes applied to cancer (Systems biology)) 2015-2018 : Coord. D. Jost, Coll. : S. Khochbin and E. Brambillat (Institute Albert Bonniot, Grenoble)

Scientific background : Recent investigations of the cancer cell epigenome are now highlighting a central role played by epigenetics in malignant transformation. In particular, aberrant gene activity of tumor-suppressor or tissue-specific genes can be directly associated to modification of the epigenome. While increasing amount of genome-wide data are accumulating, a main challenge in cancer systems biology remains to understand and propose a unifying quantitative model unraveling the fundamental rules that link the epigenetic deregulations to gene expression and hence to specific cancer phenotype. Description of the project : We propose here to developing an interdisciplinary strategy to investigate the interplay between genomic, epigenomic and transcriptomic alterations in the context of the lung cancer. Combining statistical analysis and quantitative mathematical modeling with molecular biology experiments on tumors and on specific cell lines , we aim at discovering epigenetically regulated genomic domains in lung cancer, as well as at characterizing and modeling these epigenetic "hot" domains and their association with tumor progression and aggressivness. Expected results : Development of a challenging mathematical model of epigenetic regulations, trained on experimental data, will be used to infer failures in the epigenetic maintenance mechanisms that lead to deregulation of the "hot" domains. Results obtained for lung tumors will be correlated with clinical data (prognosis, lung cancer origin, treatment response, etc.) allowing for the discovery of new oncogenic mechanisms, epigenetic biomarkers and would open the door for the establishment of new anti-cancer therapeutic strategies.

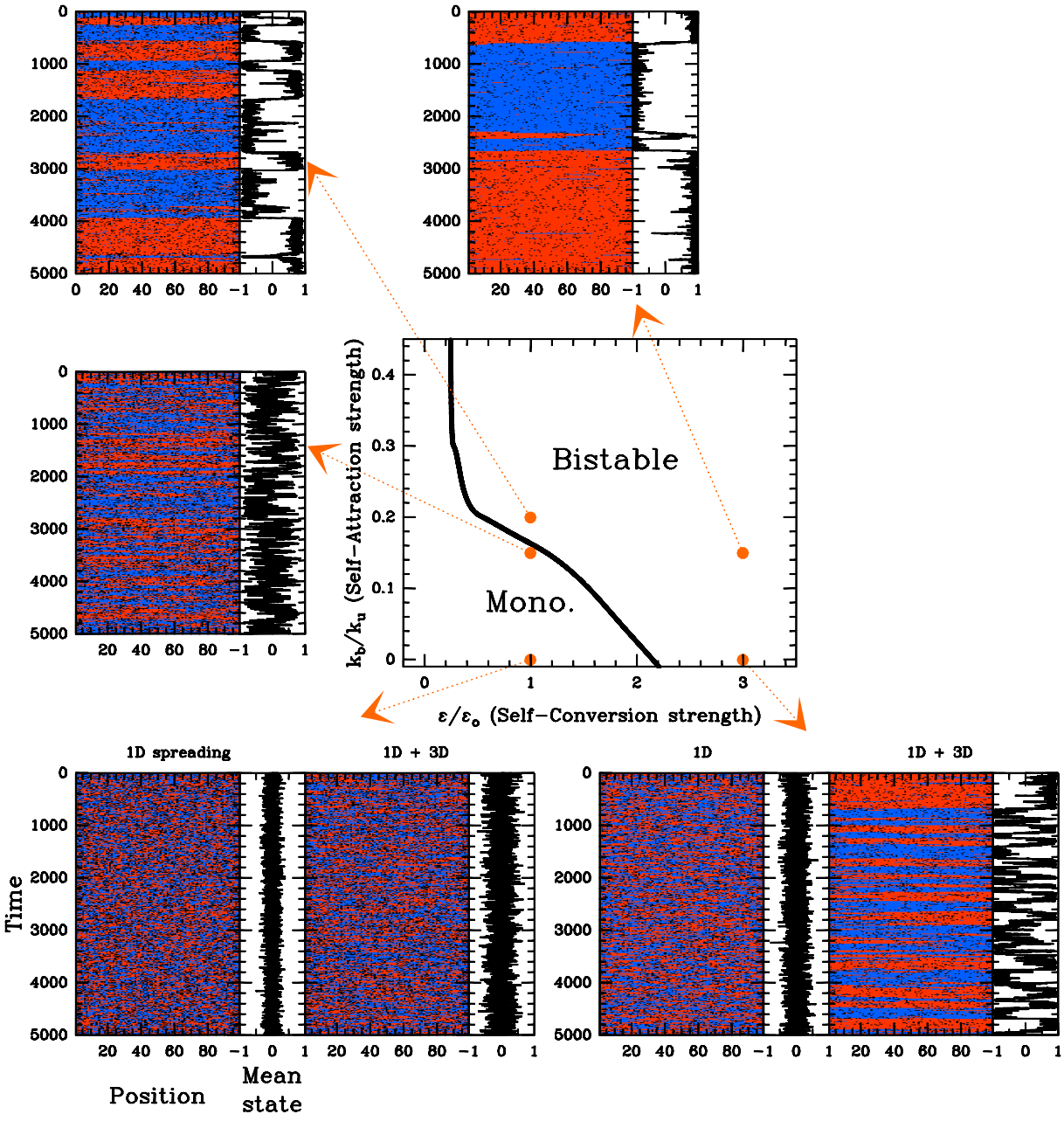


FIGURE 12 – **Epigenomic Phase Diagrams**. Epigenomic Phase diagrams in the $(k_b/k_u, \epsilon/\epsilon_o)$ plane for the 1D+3D models with $\epsilon_o = 0.001$. Some typical time evolution of the monomer chromatin states (active in blue, inactive in red, unmarked in black) and of the corresponding mean epigenomic state $\bar{e}$ are reported for the 1D and 1D+3D models at different spreading and attraction strength

4.2 Epigenome reprogramming during spermatogenesis *Collaboration with Saadi Khochbin, Institut A. Bonniot,Genoble*

An essential feature of spermatogenesis is the generation of a transportable genome placed in the nucleus of highly specialized cells, spermatozoa, capable of leaving and surviving the parent organism. After meiosis, young spermatids (known as round spermatids) inherit a chromatin-containing transcriptionally active genome that, during subsequent stages in elongating spermatids, undergoes a genome-wide histone hyperacetylation followed by histone removal and the assembly of transition proteins (TPs) and protamines in condensing spermatids. These chromatin transitions constitute a unique feature among eukaryotes, since the universal nucleosome-based organization of the genome undergoes a metamorphosis into new and unique genome-packaging structures based on nonhistone proteins. This transition (nucleoosme-TP-protamines) has been shown to occur through the complete replacement of conventional histones H2B by one of its variant TH2B (52). Although essential to the life cycle, the molecular basis of these dramatic changes remains one of the most obscure issues in modern biology.

The project here is to provide a quantitative modelling of the different stages of these global transition. The analyis of all the epigenomic datas (nucleosome positioning, histone variant replacement, histone acetylation...) obtained, in particular in the group of S. Khochbin at the different stages of spermatogenesis will allow us to build dynamical models of this global epigenome reprogramming. The main objective of this project is to provide fundamental insights concerning transgenerational epigenetic ineritance via the epigenome regulation of male gametes.

This project will involve the "1D" theoretical modeling of epigenome (nucleomes position) dynamics ; however, the "3D" spatial compartmentalization may also contribute in such large scale replacement of histone since this replacement is catalyzed by the primary modification (acetylation) of histones that would be facilitated by spatial colocalization of enzymes (and metobolic cosubstrate). As such teh "living chromatin" approach may alo be well suited for that project.

4.3 Fly development : Epigenetic regulation of development : towards a predictive mathematical modeling of three-dimensional genome folding and cellular memory *Granted ANR project ("EpiDevoMath" 2015-2018) : Coord. G. Cavalli (IGH, Montpellier), Coll. : D. Jost, R. Everaers*

The Cavalli lab has adapted the original Hi-C method to map chromosome contacts in order to improve its resolution and simplify the protocol. This allowed obtaining high-resolution contact maps of the genome at late Drosophila embryogenesis, which showed the existence of domains of 100 Kb average size. Many contacts are formed within each domain, whereas much fewer contacts are formed between domains, such that one may define these as physical or topological domains. The analysis showed that physical domains are strongly correlated with the chromatin states of underlying genes, i.e. most genes within the same physical domain have similar chromatin states. They showed the existence of four different types of physical domains, with predominant i) active chromatin, ii) heterochromatin, iii) Polycomb chromatin and iv) silent chromatin without notable marks. Furthermore, we identified interdomain contacts that associate preferentially physical domains of the same kind. In order to improve this analysis we plan to increase the sequencing depth by over one order of magnitude (from 300 million to approximately 5 billion reads). Moreover, we will produce a developmental time course in order to deduce when physical domains are formed and compare this to epigenomic domains. We will focus on early stages of embryogenesis, starting from stages preceding the onset of transcription (syncytial nuclear cleavage 8), then stages between cleavage 8 and cellularization (nuclear cleavage 14), and finally cellular blastoderm embryos. We can obtain at least 106 cells for each of the stages and this is sufficient to perform Hi-C. If we find that the earliest point still resembles late embryogenesis, we will establish Hi-C maps of earlier stages, up to the oocyte. We have already schemes that allow to obtain thousands of unfertilized eggs, and may establish single-cell Hi-C to analyze hundreds of different eggs if necessary. In order to complement the developmental maps, we will also analyze cultured cells derived from embryos in G1, S, G2 and M states, obtained by FACS sorting. This will allow deriving the stability of chromatin interactions during the cell cycle. For comparison, we will also produce Hi-C maps for different Drosophila species, such as *Drosophila simulans* and *Drosophila pseudoobscura*.

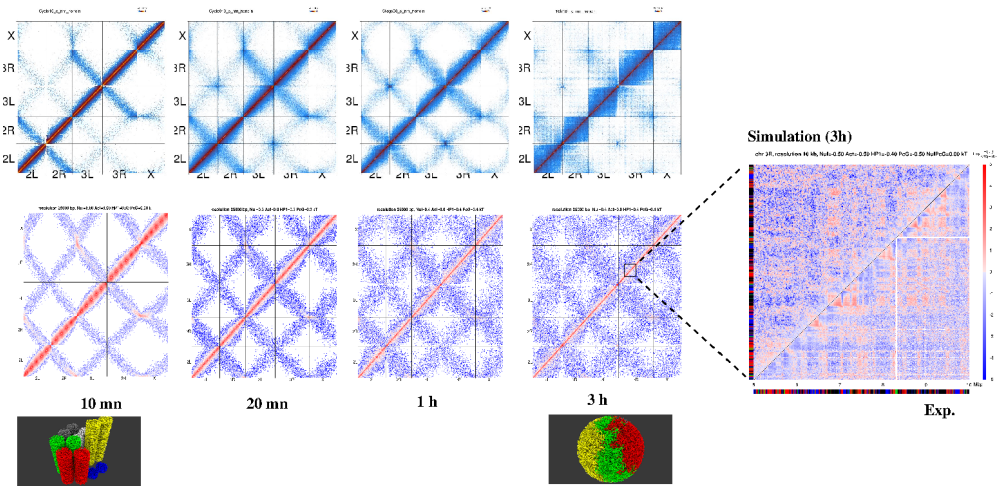


FIGURE 13 – Nuclear organization of drosophila chromosomes at different embryonic stages (Top) Experimental Hi-C maps obtained in the Cavalli's group (Bottom) Prédiction when starting with a Rab-like initial condition of chromosomes. From left to right : Mixed nuclei from stages 1-2 (cell cycle of $\sim 10mn$) , stages 3-4 (cell cycles of $\sim 20mn$), stages 5-8 (cell cycle $\sim 1h$) and late stage 16 ($> 10h$ after last mitose). (Bottom left and right images) The initial and a relaxed configurations of the chromosomes

In the course of this project, we will use additional experimental data to refine the estimates of interaction forces. In particular, we will test if the mathematical models make quantitative predictions in the form of virtual Hi-C maps for the effects of erasing topological insulation effects, for the consequence of chromosomal inversion, or for the folding of chromosomes in different Drosophila species. To this end, we will produce Hi-C maps in the In(3LR)sep inversion as well as for mutants of border binding proteins such as CTCF and Cp190. We will also test predicted distance distributions between selected loci, as predicted in our molecular dynamics simulations of chromosome folding, by Fluorescent in situ DNA hybridization (DNA FISH), using both confocal and superresolution microscopy with an OMX apparatus . Furthermore, we will attempt to improve the quality of the biological Hi-C maps by deeper sequencing (depending on cost reduction in the coming years). These gradual refinements of the predictions followed by experimental testing will lead to improvements that should ultimately lead to a very good match between predictions and data for the fly genome and represent the first step of a revolution in our understanding the relation between chromatin components and their role in setting up chromosome architecture.

Extending predictive 3D epigenomics to mouse and human cells. Recent work has shown that topological domains exist in mouse and human cells, with the notable difference in size, which approaches 1Mb in these organisms . However, the existing maps are restricted to very low sequence coverage when compared to those of the Drosophila genome. We calculated that at least 100 fold increase in sequencing depth would be required to match the resolution achieved in flies, and the Cavalli lab secured funds to perform Hi-C of 1011 paired-end reads in total, that will be used to map mouse ES cells as well as their differentiated counterparts that give rise to neuronal progenitors or terminally differentiated neurons. We will thus produce predictive Hi-C maps for these cell types and match them to the experimental maps that we will obtain by the end of 2015. In the context of the present project, we will then predict the effect of reducing Polycomb proteins and insulators and produce knockdown and knockout cells for these chromatin components and compare predictions with Hi-C data in a similar way as in flies.

4.4 The physicochemistry of nuclear reorganisation during cellular senescence

Granted FRM project (Fond pour la Recherche Médicale. Innovative physicochemical studies for biology and medicine) 2015-2018 : Coord. G. Cavalli (IGH, Montpellier), Coll. : D. Jost

As a second mammalian model for nuclear organization change, we will tackle cellular senescence. Cell senescence is the chief cause of aging, while also being involved in early development,

and constitutes a crucial genome defense mechanism against oncogenesis. The onset of senescence is marked by profound changes in 3D genome organization. Senescence is characterized by an irreversible cell cycle arrest in response to various forms of stress, including activation of oncogenes, shortened telomeres (replicative senescence), DNA damage, oxidative stress and mitochondrial dysfunction. Senescent cells acquire an alteration of cell morphology and metabolism, epigenetic changes, increase in senescence-associated β-galactosidase activity and a major change in nuclear architecture with the formation of senescence-associated heterochromatic foci, so called SAHFs. A recent analysis suggests that SAHFs do not arise as a consequence of major changes in the epigenome marks along the chromosome, but rather on a change in their 3D organization . We will revisit this notion by investigating a series of epigenome marks as well as by performing Hi-C experiments in normal versus senescent cells. For this, we dispose of a experimental "model" where we can induce either replicative senescence or oncogene-dependent senescence. The two systems will be compared and virtual Hi-C maps will also be produced for these cells in order to compare them with experimental data. This project aims to study the role of physicochemical mechanisms on the 3D folding of the genome during senescence. Our previous observation that subchromosomal spatial domains can emerge from epigenetically driven chromatin interactions has led us to the hypothesis that the altered 3D chromosomal organization in senescent cells might result from a microphase separation process among different chromatin types as senescence sets in (Fig. 14). We shall challenge this hypothesis by combining polymer modeling with mapping of chromatin interactions and their effect on global transcription regulation during the onset of senescence. The outcomes of this project will impact interventions in pathologies such as progeria and cancer.

4.5 Functional Folding of chromosomes : Dosage Compensation in *C. elegans* as a model to understand the formation of functional subnuclear domains

Collaboration with Peter Meister's group in Bern (`http://www.izb.unibe.ch/content/groups/meister/`*) and with D. Jost*

Although very different in their shape, physiology and developmental history, all cells of multicellular organisms share the same genetic material. Cell fate is therefore not a function of genome sequence but of genome expression. Gene expression is regulated at many levels, from local DNA packaging and transcription factor binding to higher order 3D chromatin functional domains. There is increasing evidence that the formation and maintenance of these subnuclear domains are strongly correlated with cell fate, suggesting a role in the epigenetic determination of gene expression (53, 54). Therefore, models of subnuclear domains are important to uncover the general principles of chromatin 3D folding and their influence on gene regulation. Dosage compensation (DC) is a promising system to understand how a nuclear domain is formed and maintained over the entire life of an organism. This essential function widely observed in metazoans, ensures adjusted expression of X-linked genes to the X to autosome ratio. The X chromosome forms a specific subnuclear domain, restricted to a single chromosome. In C. elegans hermaphrodites, DC is performed by downregulation of transcription from the two X chromosomes (55). Transcription is not completely shut down but only half downregulated, a situation much closer to general gene regulation than the complete heterochromatinization seen in mammalian female cells. Transcriptional downregulation is achieved via association of the Dosage Compensation Complex (DCC, a condensin-like complex). DCC loading and spreading along the X chromosome leads to the restriction of the methylation levels of a specific histone tail residue, a decrease in active chromatin marks and partial eviction of RNA polymerase II from the X chromosome [Ferrari]. The functional link between DCC loading and the downregulation of transcription remains elusive, as there is no direct correlation between DCC presence along the chromosome and transcriptional silencing. Together with the structure of the DCC ; this suggested a role for tridimensional nuclear organization in the creation of a specific "half-repressive" subnuclear domain containing the X chromosome. In previous work of this consortium, we showed that the X chromosome undergoes major structural rearrangements inside the nuclear space in relation with DC (50). In males, the X chromosome is located at the nuclear periphery and interacts with nuclear pore components, which had previously been shown to increase transcriptional output in a variety of biological systems. In hermaphrodite, the X chromosome is located in the middle of the nucleus and its localization depends on DCC loading, which may impair pore interaction and transcrip-

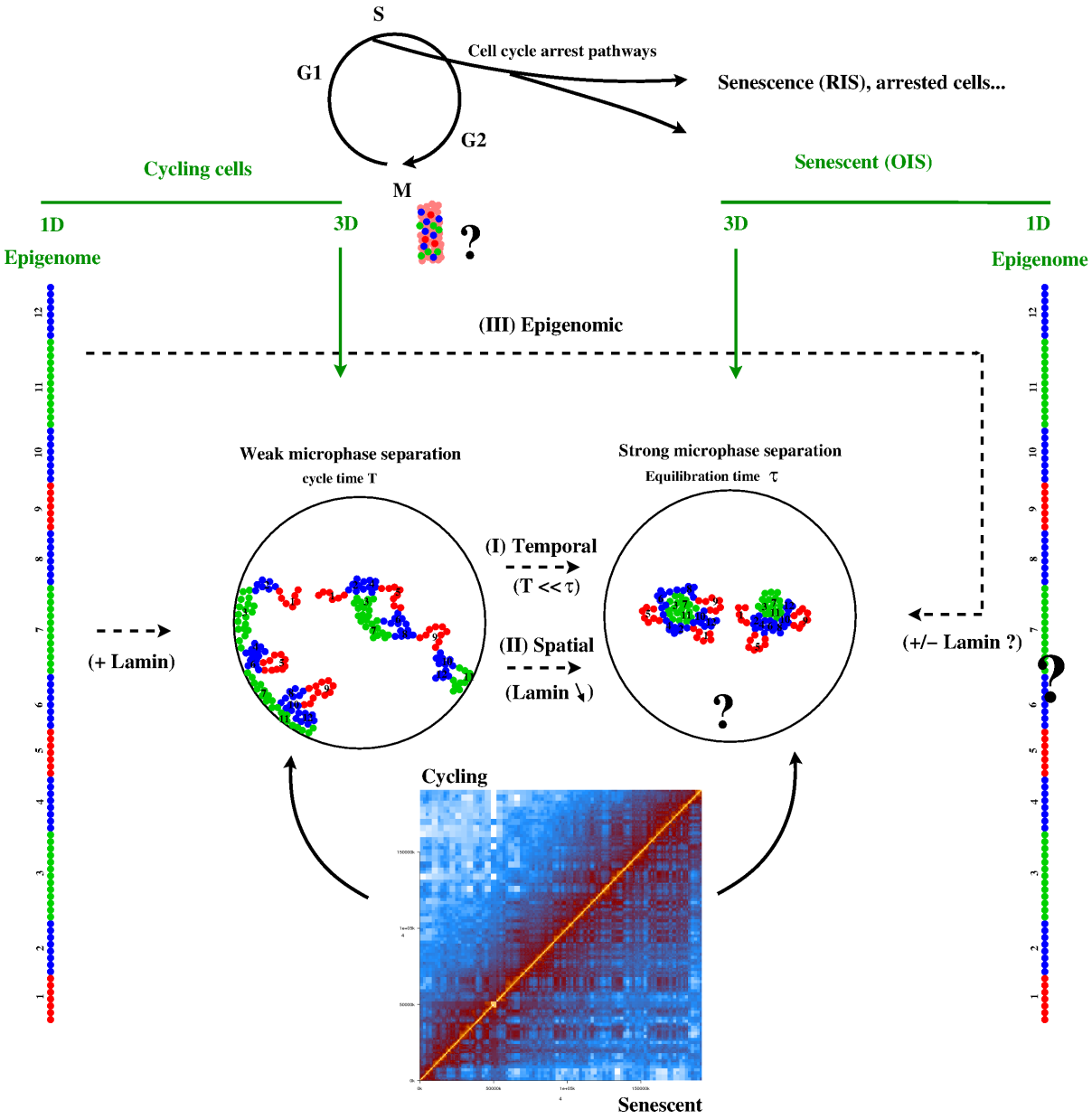


FIGURE 14 – The working hypothesis : Nuclear organisation as a micro-phase separation of chromatin states. In this interdisciplinary project we aim to understand the physico-chemical and biological mecanisms that drive spatial organisation of chromosomes. A very well suited system to test this working hypothesis is cellular senescence, a natural cell cycle arrest pathway that is accompanied by a strong microphase separation as seen in imunofluorescence and preliminary HiC maps. We propose three main non-exclusive scenarios that are consistent with our theoretical co-polymer folding model of the epigenome : (I) Temporal scenario : escape from cycling constraints may allow relaxation of the out-of-equilibrium weakly micro-phase separated state of cycling cells toward the equilibrated fully micro-phase separated state. (II) Spatial scenario : a downregulation of the lamin-heterochromatin interaction that maintains a weak micro-phase separation in normal cycling cells may induce relocalization and bulk long-range association of heterochromatin in senescent cells (III) Epigenomic scenario :changes of the epigenome, in particular transcriptional changes, may induce changes in physico-chemical properties of chromatin states (self-interaction, mobility...) leading to a different folding pattern. Within our theoretical co-polymer framework, we aim to infer interaction parameters from experiments and make quantitative predictions that we will test with dedicated experiments.

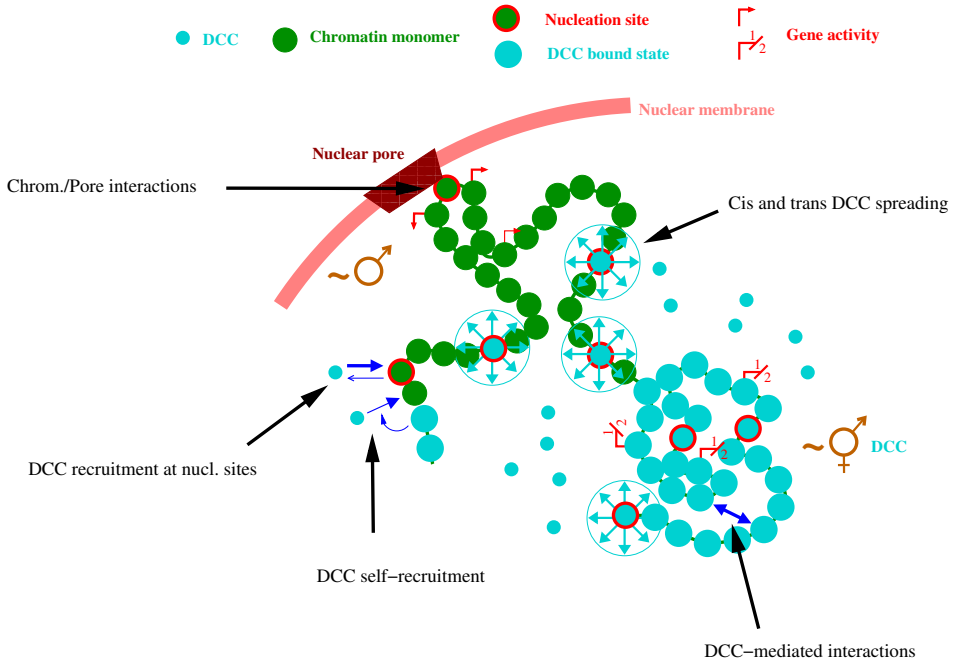


FIGURE 15 – **The "Living Chromatin model" applied to DC** : Chromatin is modeled as a polymer where monomers interact with each other or with nuclear landmarks, depending on their chromatin state. In males, interactions with nuclear pores induce segregation of the X chromosomes at the nuclear periphery. In hermaphrodite, DCC is recruited at specific sites and spread in cis and in trans by subsequent self-recruitment that is enhanced by 3D long-range interactions.

tional activation. Together, this strongly suggests a model of DC in which the X chromosome is creating two different subnuclear domains in males (non compensated) and hermaphrodites (compensated) and these domains impinge on X-linked gene transcription. The dosage compensated X is a perfect model to understand the function of nuclear domains in relation to physical properties of the chromatin fiber : DC occurs in all cells of the organism, the loading sites of the DCC and interaction sites at the periphery are known, interactions with nuclear landmarks have been characterized, the X domain is different between sexes and the transcriptional output is clear and phenotypically screenable. With this project, we aim to understand how the physical properties of the chromatin folding coupled to the loading and spreading of DCC in one sex impact on the embryonic creation and life-long maintenance of functional X chromosome domains. In particular, we want to elucidate the physical and molecular mechanisms of domain wide gene regulation. To this aim, our approach is cross-disciplinary, combining computation modeling of chromatin behavior and a highly sensitive genome-wide mapping technique (DamID-seq). The iterative process of model implementation, confrontation with the data, formulation of predictions and experimental verification will provide new insights into the mechanism of DC.

The project is subdivided into 4 specific aims :

1. Detailed experimental description of the establishment of the compensated X domain
2. Development of a modular computational model of functional subnuclear domains (Fig. 15) : In parallel to the characterization of X chromosome conformation in vivo, we will develop an original physical model ("the Living Chromatin" model) that couples the dynamics of the chromatin folding - modeled as a semi-flexible heterogeneous polymer chain - to the stochastic loading and spreading of architectural proteins and epigenetic marks along chromosomal domains. While most of the published computational chromatin models consider chromosomes as homogenous polymers, our recent work led to the development of a heterogeneous polymer model which accounts for the local variability of epigenomic properties (42). This allowed interrogating the role of chromatin landscape along the genome in the formation of subnuclear domains and their impact on gene transcription. In particular, we showed that formation and dynamics of topologically-associated domains may be

associated with the folding of chromatin domains due to protein-mediated interactions. To describe the establishment and maintenance of DC, we aim to generalize this approach by modeling the dynamics of chromatin regulators and its crosstalk with 3D chromatin folding and gene expression. To date, models that explicitly assess for this dynamic coupling remains to be developed. Here, we propose to complete the polymer modeling developed in (42) by a stochastic description of the local chromatin state (34). In particular, X chromosomes will be modeled as a confined long semi-flexible polymer where each monomer represents a portion of DNA. The chromatin state of each monomer would dynamically fluctuate between several flavors depicting the presence or not of DCC or other factors. DCC-mediated interactions between monomers will impact on the local and higher-order chromatin organization, while spatial localization will affect DCC spreading and recruitment. Ultimately, we aim to understand the functional impact of chromatin state, folding and nuclear localization in the non-compensated vs compensated X chromosome : for that purpose and as first attempt, local gene activity will be modeled as depending on local compaction/chromatin state and on proximity to nuclear pores. Outcomes of the model will be generated using advanced numerical kinetic Monte-Carlo methods. For a given set of parameters and molecular mechanisms, the living chromatin model would be able to make quantitative predictions on the nuclear 3D chromatin organization of the X chromosomes as well as on X gene expression during the life span of the animal. It will allow relating experimental observations to the critical underlying physical and molecular mechanisms involved in DC establishment and maintenance. Moreover, it is likely that the model will formulate new predictions that will strengthen the description of DC processes and that can be tested in vivo.

3. Role and function of DCC in hermaphrodite X domain maintenance
4. Interpretation of experimental results by the model and experimental tests of the model predictions.

4.6 Modelling Hox gene regulation In the frame of G. Cavalli's collaboration, I propose to investigate the fascinating phenomenon of Hox genes regulation in both Drosophila and vertebrate systems (56). During embryogenesis, Hox genes control the specification of the anterior to posterior axis ; in vertebrate they also organize structures along other body axis such as the appendicular (57, 17). All these genes are organized into genomic clusters and, remarkably, the relative order of Hox genes within this cluster corresponds to the relative position of the structures they instruct along the anterior-to-posterior body axis (Fig. 16 A,B). Recent studies have shown that this "colinear" activation is associated with a a "wave" of concommittent spatial and epigenomic reorganisation of chromatin along the clusters (Fig. 16A,B). As proposed by D. Duboule (56), the emerging picture is that such "collinearity" is based on five main principles . Hox gene regulation and the question of colinearity might thus be a great modelling task to investigate within our "living polymer" framework.

5 CONCLUSION

This project rely on the strong collaboration with different groups of biologists G. Cavalli, P. Meister, S. Khochbin..., international leaders in their field ; my objective is, in a joint work with Daniel Jost (TIMC-IMAG Grenoble), Ralf Everaers (Lab. Physique ENSL), A. Rosa in Trieste (`https://sites.google.com/site/angelosissa/`), P. Carrivain and the future post-docs and PhD students to combine our theoretical and numerical skills in order to provide original multi-scale theoretical modelling of the regulation of epigenome assembly and folding. I propose to carry out this project in the Laboratory of Physics of "ENS de Lyon". My objective is first to work on developing our "non-localized"computational biology group with D. Jost ("Physics of epigenome regulation") to explore topics related to epigenetic gene regulation in various systems in both multicellular and unicellular organisms. At the ENS de Lyon, I am currently collaborating with Gael Yvert from the LBMC on topics related to nucleosome positioning and evolutionnary drifts in Yeast ; future collaborations with François Roudier from the RDP concerning evolutionnary dynamics and the PcG repression system in plants are also planned. I am currently coordinating the set up of a new structure hosted at the Centre Blaise Pascal et ENS de Lyon, called "Atelier

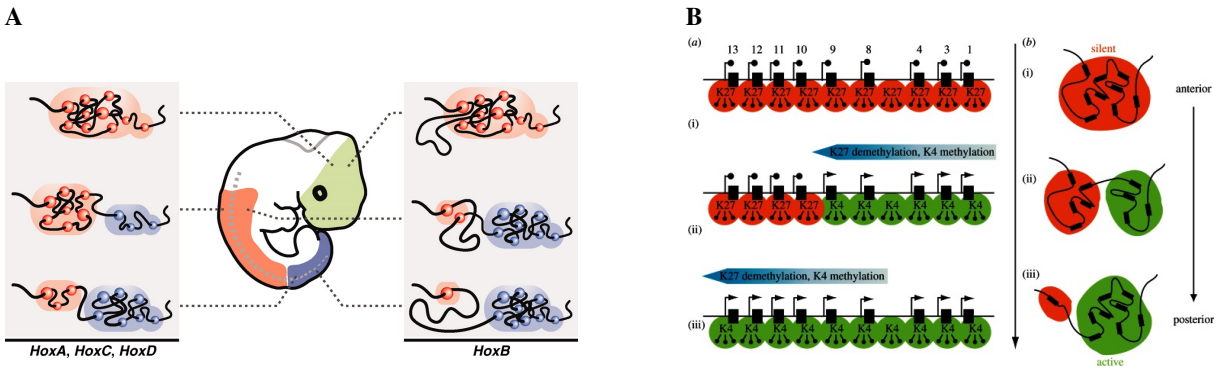


FIGURE 16 – A) Model of the 3D organization of Hox gene clusters, at various stages of colinear gene activation. Transcriptionally inactive genes are depicted in red and active genes in blue. Gene activation is paralleled by a transition from one 3D domain, matching the presence of H3K27me3, to another domain of active transcription (marked with H3K4me3). Although the same dynamics are observed for the HoxA, HoxC, and HoxD clusters (left), the HoxB cluster (right) shows a slight variation with a large piece of intergenic DNA looping out from these two domains. B) Collinearity during trunk extension and chromatin dynamics at Hox clusters. Expression of Hox gene along the anterior-to-posterior (AP) embryonic axis is collinear with gene order within the cluster. (a) During axial extension, the sequential onset of Hox gene transcriptional activation is accompanied by a transition in histone modifications over the gene cluster. In ES cells (i), the whole cluster is labelled with H3K27me3 (orange), a mark associated with Polycomb-mediated silencing. In the developing embryo, this mark is progressively erased and replaced by H3K4me3 (green), concomitantly with gene activation. (b) Active and silent Hox loci segregate into distinct spatial compartments along the AP axis. In embryonic tissues where the whole cluster is repressed, such as the forebrain (i), Hox clusters form a compact three-dimensional structure. In regions where subsets of Hox genes are expressed (anterior trunk, ii), active and silent genes segregate in distinct compartments, labelled with either H3K27me3 (silent compartment) or H3K4me3 (active compartment). In posterior embryonic regions (iii), most genes are transcribed and participate in the active compartment.

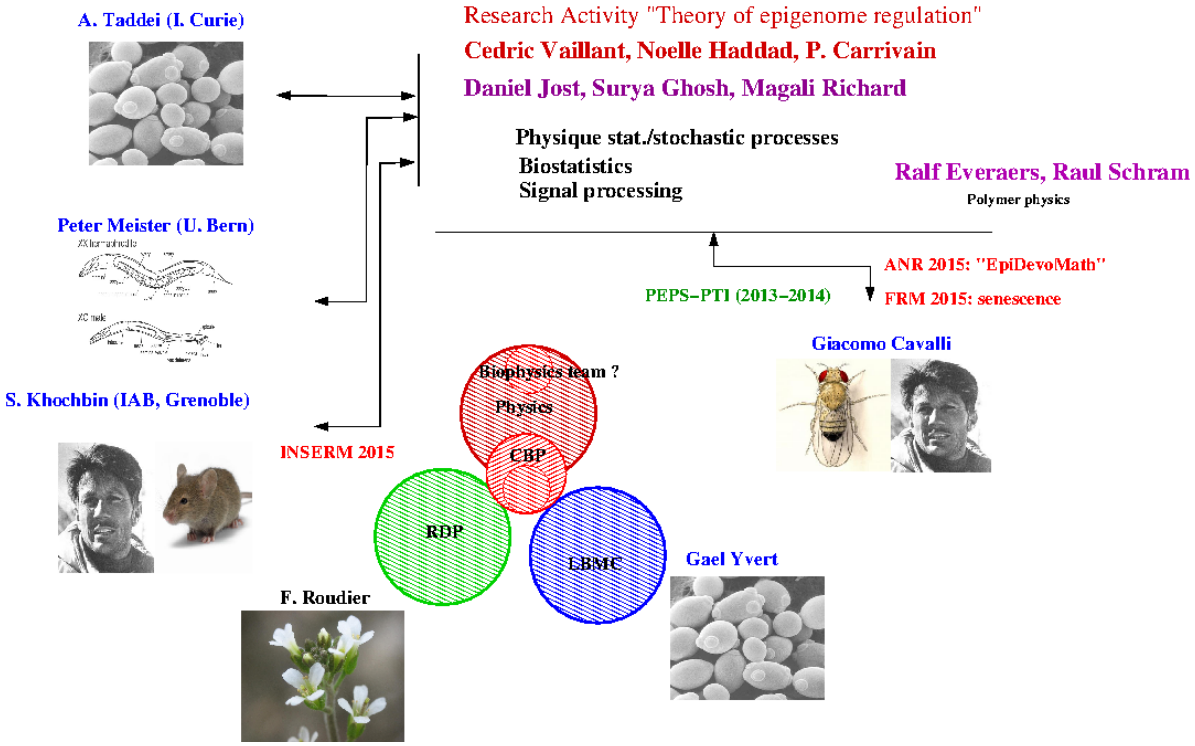


FIGURE 17 – Who ? With whom ? Where ?

de Biologie Numérique" whose main objective will be to mutualize and organize the numerical ressources provided by the CBP/PSMN (`http://www.cbp.ens-lyon.fr/doku.php`) as well as to share knowledges and skills in computational biology.

[1] Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. Three-dimensional folding and functional organization principles of the drosophila genome. *Cell* **148**(3), 458–472, Feb (2012).

[2] Filion, G. J., van Bemmel, J. G., Braunschweig, U., Talhout, W., Kind, J., Ward, L. D., Brugman, W., de Castro, I. J., Kerkhoven, R. M., Bussemaker, H. J., et al. Systematic protein location mapping reveals five principal chromatin types in drosophila cells. *Cell* **143**(2), 212–224 (2010).

[3] Ernst, J., Kheradpour, P., Mikkelsen, T. S., Shoresh, 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. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**(7345), 43–49, May (2011).

[4] Roudier, F. and other. Integrative epigenomic mapping defines four main chromatin states in arabidopsis. *EMBO J* **30**(10), 1928–1938, May (2011).

[5] Gerstein, M. B. and other. Integrative analysis of the caenorhabditis elegans genome by the modencode project. *Science* **330**(6012), 1775–1787, Dec (2010).

[6] Kharchenko, P. V., Alekseyenko, A. A., Schwartz, Y. B., Minoda, A., Riddle, N. C., Ernst, J., Sabo, P. J., Larschan, E., Gorchakov, A. A., Gu, T., Linder-Basso, D., Plachetka, A., Shanower, G., Tolstorukov, M. Y., Luquette, L. J., Xi, R., Jung, Y. L., Park, R. W., Bishop, E. P., Canfield, T. K., Sandstrom, R., Thurman, R. E., MacAlpine, D. M., Stamatoyannopoulos, J. A., Kellis, M., Elgin, S. C. R., Kuroda, M. I., Pirrotta, V., Karpen, G. H., and Park, P. J. Comprehensive analysis of the chromatin landscape in drosophila melanogaster. *Nature* **471**(7339), 480–485, Mar (2011).

[7] van Steensel, B. Chromatin : constructing the big picture. *EMBO J* **30**(10), 1885–1895, May (2011).

[8] Ong, C.-T. and Corces, V. G. Enhancer function : new insights into the regulation of tissue-specific gene expression. *Nat Rev Genet* **12**(4), 283–293, Apr (2011).

[9] 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. Architectural protein subclasses shape 3d organization of genomes during lineage commitment. *Cell* **153**(6), 1281–1295, Jun (2013).

[10] Hou, C., Li, L., Qin, Z. S., and Corces, V. G. Gene density, transcription, and insulators contribute to the partition of the drosophila genome into physical domains. *Mol Cell* **48**(3), 471–484, Nov (2012).

[11] Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., and Ren, B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**(7398), 376–380, May (2012).

[12] Nagano, T., Lubling, Y., Stevens, T. J., Schoenfelder, S., Yaffe, E., Dean, W., Laue, E. D., Tanay, A., and Fraser, P. Single-cell hi-c reveals cell-to-cell variability in chromosome structure. *Nature* **502**(7469), 59–64, Oct (2013).

[13] Ahmed, K., Dehghani, H., Rugg-Gunn, P., Fussner, E., Rossant, J., and Bazett-Jones, D. P. Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. *PLoS One* **5**(5), e10531 (2010).

[14] Rapkin, L. M., Anchel, D. R. P., Li, R., and Bazett-Jones, D. P. A view of the chromatin landscape. *Micron* **43**(2-3), 150–158, Feb (2012).

[15] Solovei, I., Wang, A. S., Thanisch, K., Schmidt, C. S., Krebs, S., Zwerger, M., Cohen, T. V., Devys, D., Foisner, R., Peichl, L., Herrmann, H., Blum, H., Engelkamp, D., Stewart, C. L., Leonhardt, H., and Joffe, B. Lbr and lamin a/c sequentially tether peripheral heterochromatin and inversely regulate differentiation. *Cell* **152**(3), 584–598, Jan (2013).

[16] Chandra, T. and other. Independence of repressive histone marks and chromatin compaction during senescent heterochromatic layer formation. *Mol Cell* **47**(2), 203–214, Jul (2012).

[17] Noordermeer, D., Leleu, M., Splinter, E., Rougemont, J., Laat, W. D., and Duboule, D. The dynamic architecture of hox gene clusters. *Science* **334**(6053), 222–225, Oct (2011).

[18] Noordermeer, D. and Duboule, D. Chromatin architectures and hox gene collinearity. *Curr Top Dev Biol* **104**, 113–148 (2013).

[19] Hawkins, R. D. and other. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* **6**(5), 479–491, May (2010).

[20] Zhu, J., Adli, M., Zou, J. Y., Verstappen, G., Coyne, M., Zhang, X., Durham, T., Miri, M., Deshpande, V., Jager, P. L. D., Bennett, D. A., Houmard, J. A., Muoio, D. M., Onder, T. T., Camahort, R., Cowan, C. A., Meissner, A., Epstein, C. B., Shoresh, N., and Bernstein, B. E. Genome-wide chromatin state transitions associated with developmental and environmental cues. *Cell* **152**(3), 642–654, Jan (2013).

[21] Jones, P. A. and Baylin, S. B. The epigenomics of cancer. *Cell* **128**(4), 683–692, Feb (2007).

[22] Berger, S. and Sassone-Corsi, P. Metabolic signaling to chromatin. *Cold Spring Harbor Perspectives in Biology* **8**, a019463 (2016).

[23] Benayoun, B., Pollina, E., and A., B. Epigenetic regulation of ageing : linking environmental inputs to genomic stability. *Nat Rev Mol Cell Biol* **16**, 593–610 (2015).

[24] van der Knaap, J. and Verrijzer, C. Undercover : gene control by metabolites and metabolic enzymes. *Genes & Dev* **30**, 2345âĂŞ2369 (2016).

[25] Katada, S., Imhof, A., and Sassone-Corsi, P. Connecting threads : epigenetics and metabolism. *Cell* **148**(1-2), 24–28, Jan (2012).

[26] Francis, N. J., Kingston, R. E., and Woodcock, C. L. Chromatin compaction by a polycomb group protein complex. *Science* **306**(5701), 1574–1577, Nov (2004).

[27] Lo, S. M., Follmer, N. E., Lengsfeld, B. M., Madamba, E. V., Seong, S., Grau, D. J., and Francis, N. J. A bridging model for persistence of a polycomb group protein complex through dna replication in vitro. *Mol Cell* **46**(6), 784–796, Jun (2012).

[28] Isono, K., Endo, T. A., Ku, M., Yamada, D., Suzuki, R., Sharif, J., Ishikura, T., Toyoda, T., Bernstein, B. E., and Koseki, H. Sam domain polymerization links subnuclear clustering of prc1 to gene silencing. *Dev Cell* **26**(6), 565–577, Sep (2013).

[29] 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. A conformational switch in hp1 releases auto-inhibition to drive heterochromatin assembly. *Nature* **496**(7445), 377–381, Apr (2013).

[30] Jost, D., Vailant, C., and Meister, P. Coupling 1D modifications and 3D nuclear organization : data, models and function. *Curr Opin Cell Biol* **44**, 20–27 (2017).

[31] Haddad, N., Jost, D., and Vaillant, C. Perspectives : Using polymer modeling to understand the formation and function of nuclear compartments. *Chromosome Res* **-**, In press (2017).

[32] Ramachandran, S. and Henikoff, S. Replicating nucleosomes. *Sci Adv* **1**(7), Aug (2015).

[33] Dodd, I. B., Micheelsen, M. A., Sneppen, K., and Thon, G. Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell* **129**(4), 813–822, May (2007).

[34] Jost, D. Bifurcation in epigenetics : implications in development, proliferation, and diseases. *Phys Rev E Stat Nonlin Soft Matter Phys* **89**(1), 010701, Jan (2014).

[35] Zerihun, M. B., Vaillant, C., and Jost, D. Effect of replication on epigenetic memory and consequences on gene transcription. *Phys Biol* **12**(2), 026007, Apr (2015).

[36] Grosberg, A. Y. and Khokhlov, A. R. *Statistical physics of macromolecules*. AIP Press, (1994).

[37] Wong, H., Marie-Nelly, H., Herbert, S., Carrivain, P., Blanc, H., Koszul, R., Fabre, E., and Zimmer, C. A predictive computational model of the dynamic 3d interphase yeast nucleus. *Curr Biol* **22**(20), 1881–1890, Oct (2012).

[38] Hajjoul, H., Mathon, J., Ranchon, H., Goiffon, I., Mozziconacci, J., Albert, B., Carrivain, P., Victor, J.-M., Gadal, O., Bystricky, K., and Bancaud, A. High-throughput chromatin motion tracking in living yeast reveals the flexibility of the fiber throughout the genome. *Genome Res* **23**(11), 1829–1838, Nov (2013).

[39] 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., and Dekker, J. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**(5950), 289–293, Oct (2009).

[40] Rosa, A. and Everaers, R. Structure and dynamics of interphase chromosomes. *PLoS Comput. Biol.* **4**(8), e1000153 (2008).

[41] Barbieri, M., Chotalia, M., Fraser, J., Lavitas, L.-M., Dostie, J., Pombo, A., and Nicodemi, M. Complexity of chromatin folding is captured by the strings and binders switch model. *Proc Natl Acad Sci U S A* **109**(40), 16173–16178, Oct (2012).

[42] Jost, D., Carrivain, P., Cavalli, G., and Vaillant, C. Modeling epigenome folding : formation and dynamics of topologically associated chromatin domains. *Nucleic Acids Res* **42**(15), 9553–9561 (2014).

[43] Timoshenko, E., Kuznetsov, Y. A., and Dawson, K. Conformational transitions of heteropolymers in dilute solutions. *Phys. Rev. E* **57**(6), 6801–6814 (1998).

[44] Ramalho, T., Selig, M., Gerland, U., and Enßlin, T. A. Simulation of stochastic network dynamics via entropic matching. *Phys. Rev. E* **87**(2), 022719 (2013).

[45] Meuleman, W., Peric-Hupkes, D., Kind, J., Beaudry, J.-B., Pagie, L., Kellis, M., Reinders, M., Wessels, L., and van Steensel, B. Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res* **0**, doi : 10.1101/gr.141028.112, Dec (2012).

[46] Vaquerizas, J. M., Suyama, R., Kind, J., Miura, K., Luscombe, N. M., and Akhtar, A. Nuclear pore proteins nup153 and megator define transcriptionally active regions in the drosophila genome. *PLoS Genet* **6**(2), e1000846, Feb (2010).

[47] Mekhail, K. and Moazed, D. The nuclear envelope in genome organization, expression and stability. *Nat Rev Mol Cell Biol* **11**(5), 317–328, May (2010).

[48] Mukhopadhyay, S., Schedl, P., Studitsky, V. M., and Sengupta, A. M. Theoretical analysis of the role of chromatin interactions in long-range action of enhancers and insulators. *Proc Natl Acad Sci U S A* **108**(50), 19919–19924, Dec (2011).

[49] Olarte, J., Haddad, N., Vaillant, C., and Jost, D. The folding landscape of the epigenome. *Phys Biol* **13**, 026001 (2016).

[50] Sharma, R., Jost, D., Kind, J., Gómez-Saldivar, G., van Steensel, B., Askjaer, P., Vaillant, C., and Meister, P. Differential spatial and structural organization of the x chromosome underlies dosage compensation in c. elegans. *Genes Dev* **28**(23), 2591–2596, Dec (2014).

[51] Haddad, N. *Analyse et modélisation du repliement spatial de l'épigénome.* PhD thesis, Ecole Normale Supérieure de Lyon, (2016).

[52] Montellier, E., Boussouar, F., Rousseaux, S., Zhang, K., Buchou, T., Fenaille, F., Shiota, H., Debernardi, A., Héry, P., Curtet, S., Jamshidikia, M., Barral, S., Holota, H., Bergon, A., Lopez, F., Guardiola, P., Pernet, K., Imbert, J., Petosa, C., Tan, M., Zhao, Y., Gérard, M., and Khochbin, S. Chromatin-to-nucleoprotamine transition is controlled by the histone h2b variant th2b. *Genes Dev* **27**(15), 1680–1692, Aug (2013).

[53] Meister, P., Mango, S. E., and Gasser, S. M. Locking the genome : nuclear organization and cell fate. *Curr Opin Genet Dev* **21**(2), 167–174, Apr (2011).

[54] Cavalli, G. and Misteli, T. Functional implications of genome topology. *Nat Struct Mol Biol* **20**(3), 290–299, Mar (2013).

[55] Ferrari, F., Alekseyenko, A. A., Park, P. J., and Kuroda, M. I. Transcriptional control of a whole chromosome : emerging models for dosage compensation. *Nat Struct Mol Biol* **21**(2), 118–125, Feb (2014).

[56] Noordermeer, D. and Duboule, D. Chromatin looping and organization at developmentally regulated gene loci. *Wiley Interdiscip Rev Dev Biol* **2**(5), 615–630 (2013).

[57] Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., Trono, D., Spitz, F., and Duboule, D. A switch between topological domains underlies hoxd genes collinearity in mouse limbs. *Science* **340**(6137), 1234167, Jun (2013).