Painters in chromatin: a unified quantitative framework to systematically characterize epigenome regulation and memory

Amith Z. Abdulla[1,2], Cédric Vaillant[2,*], and Daniel Jost[1,*]

[1] Laboratoire de Biologie et Modélisation de la Cellule, École Normale Supérieure de Lyon, CNRS, UMR5239, Inserm U1293, Université Claude Bernard Lyon 1, Lyon, France
[2] École Normale Supérieure de Lyon, CNRS, Laboratoire de Physique, Lyon, France

* Corresponding authors: daniel.jost@ens-lyon.fr and cedric.vaillant@ens-lyon.fr.

Abstract

In eukaryotes, many stable and heritable phenotypes arise from the same DNA sequence, owing to epigenetic regulatory mechanisms relying on the molecular cooperativity of "reader-writer" enzymes. In this work, we focus on the fundamental, generic mechanisms behind the epigenome memory encoded by post-translational modifications of histone tails. Based on experimental knowledge, we introduce a unified modeling framework, the painter model, describing the mechanistic interplay between sequence-specific recruitment of chromatin regulators, chromatin-state-specific reader-writer processes and long-range spreading mechanisms. A systematic analysis of the model building blocks highlights the crucial impact of tridimensional chromatin organization and state-specific recruitment of enzymes on the stability of epigenomic domains and on gene expression. In particular, we show that enhanced 3D compaction of the genome and enzyme limitation facilitate the formation of ultra-stable, confined chromatin domains. The model also captures how chromatin state dynamics impact the intrinsic transcriptional properties of the region, slower kinetics leading to noisier expression. We finally apply our framework to analyze experimental data, from the propagation of $\gamma H2AX$ around DNA breaks in human cells to the maintenance of heterochromatin in fission yeast, illustrating how the painter model can be used to extract quantitative information on epigenomic molecular processes.

1 Introduction

The ability of organisms to precisely regulate gene expression is central to their development. Proper temporal and spatial expressions of genes in eukaryotes require activation of transcription during the appropriate developmental stages. In response to environmental and developmental cues, cells can adopt different gene expression patterns and 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 or significantly weakened. This capacity of translating transient external stimuli into diverse and stable phenotypes without alteration of the genomic sequence is at the heart of the "epigenetic" regulation of gene expression [1]. Epigenetic processes are involved in the control of somatic inheritance and in the maintenance of cellular identity as well as in the transgenerational inheritance of traits by transmission via the germline [2].

A major class of mechanisms driving such epigenetic "memory" relies on chromatin-based processes regulating gene expression by the control of the local biochemical and structural properties of chromatin leading to different chromatin states more or less permissive to transcription [3]. Such control is mediated in part by biochemical modifications of the DNA or of histone tails, the so-called epigenomic marks. The mechanisms of establishment and heritability of these chromatin states, either active or repressive, are governed by similar general rules involving the combined and self-reinforcing action of specific chromatin-binding proteins that add or remove epigenomic marks [4–6].

De novo assembly first proceeds by a nucleation ("forcing") stage via the targeting of specific enzymes at dedicated regulatory sequences by either DNA binding proteins or the RNAi-based pathways [4, 7–11].

Such nucleation elements are often composed by multiple binding sites for different DNA binding proteins that associate to their cognate sequence to stably recruit chromatin regulators. For example, in Drosophila, the Polycomb-based epigenetic repression of developmental genes relies on the targeting of the histone modifying enzymes (HMEs) PRC1 (monoubiquitination of H2AK118) and PRC2 (methylation of H3K27) complexes at specific "silencers" regions, the so-called Polycomb Response Elements, characterized by various combinations of binding sites for adaptor proteins (e.g. Gaga, Zeste, Pho) [12, 13]. Similarly, gene activation relies on the recruitment by Trithorax-group proteins (e.g. MLL2) of acetyltransferases (e.g. p300) and demethylases (e.g. UTX) at gene promoters and enhancers [14, 15].
Once initiated, the state is able to spread to the neighboring sequences and to form a stable chromatin domain [10] that can further propagate through replication and mitosis [16]. The ubiquitous ability of some chromatin regulators to be recruited by (e.g. Clr4) or to have a boosted activity in presence of (e.g. PRC2) the chromatin state they catalyze ("reading"capacity) and to spread this state to the neighbouring sequences ("writing" capacity) introduce an effective positive feedback which is believed to be a key ingredient of epigenetic maintenance [10, 16, 17].

Such "reader-writer" principle of chromatin enzymes and its impact on epigenetic memory has been already investigated using simple mathematical models of chromatin state regulation with a focus on the dynamics of histone marks mediated by HMEs [11, 18–29]. In particular, in their seminal work, Dodd *et al.* [18] suggested that the maintenance of stable, extended active or repressed chromatin domains over generations, even in absence of nucleation signals, is made possible by the reader-writer property of HMEs coupled to their capacity to spread (or write) a mark at long-range along the genome. Indeed, such interplay leads to cooperativity and allows the effective formation of a large reservoir of modified nucleosomes to serve as templates to ensure full recovery after random perturbations such as transcription- [26, 28] or replication-mediated [22, 30] histone turnover. Actually, the long-range spreading property reflects the polymeric nature of the genome that that can bring in close spatial proximity two distant loci. Many experimental and theoretical studies have indeed highlighted the correlation between spatial chromosome organisation and chromatin regulation [31–38], strengthening this hypothesis that the 3D genome impinges on chromatin states assembly and maintenance.

The reader-writer ability of HMEs coupled to long-range spreading however raises concerns about the maintenance of a stable compartmentalization of the genome [24]. Indeed, as observed in many species, the states of chromatin along the genome are linearly organised into consecutive "domains" of finite sizes, leading globally to a 1D compartmentalisation of the genome into active and inactive domains with more or less well defined inter-domain boundaries [33, 40, 41]. Because of the requirement to regulate (activation or silencing) a given part of the genome without affecting its surrounding flanks, chromatin states have thus to be specifically targeted to and stably confined inside specific genomic domains. In standard mathematical models [18, 21, 30], long-term epigenetic memory goes with great difficulties to limit the expansion of the mark, questioning mechanistically how the 1D partitioning of the genome is established and above all maintained. Several hypotheses have been proposed to address this question. Spreading may be slowed down by boundary elements [6, 19, 26, 37], usually termed "insulators", that restrict the local spreading of a mark along the genome. These insulators are often associated with specific DNA-binding proteins (e.g. CTCF, BEAF, CP190) or actively transcribed genes found at the boundaries between antagonistic chromatin states [42–49]. Spreading may also be limited by the formation of spatial compartments [37–39, 50, 60] as genome 3D compartmentalization is strongly correlated with 1D chromatin segmentation [31, 33, 36], thus providing 3D insulation and restricting long-range effects.

Nevertheless, recent experimental studies suggest that, in vivo, the reader-writer mechanism (along with long-range spreading and insulation) might not be strong enough to self-sustain by itself an epigenetic state and that a (compartmentalized) long-term memory may still be dependent on genomic bookmarking by keeping a weak "forcing" activity at regulatory sites [10, 51–53]. Such role of nucleation signals in the maintenance of a stable chromatin state has only been partially addressed [24, 27] and a general mechanistic framework of epigenomic regulation and memory integrating nucleation, reader-writer mechanisms and long-range spreading is still lacking.

In this paper, we provide such a framework and systematically investigate the formation and memory

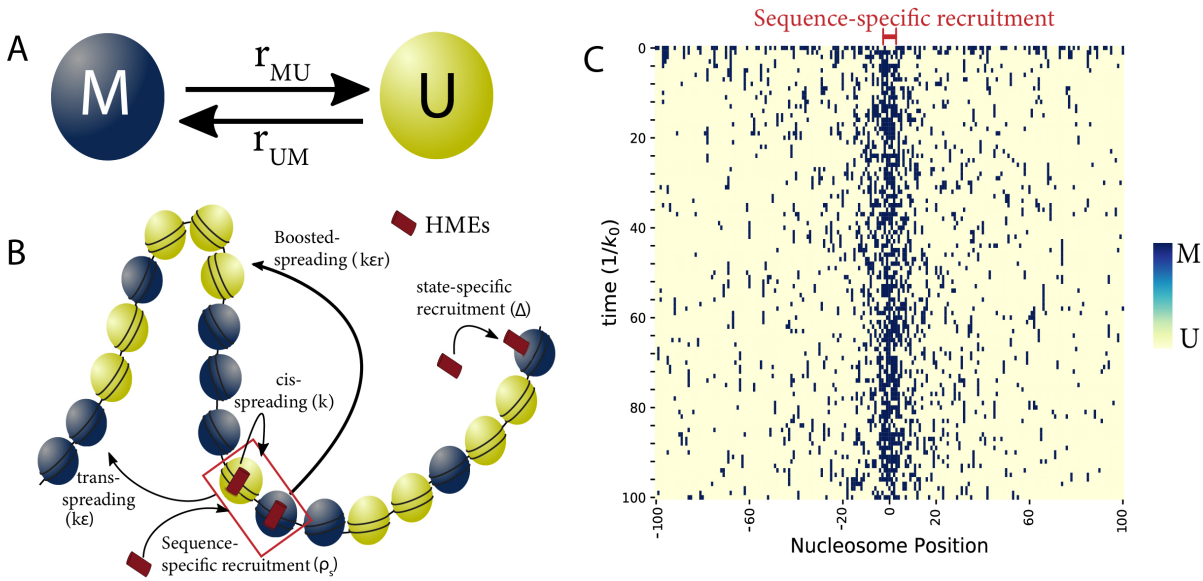

Figure 1. "Painter" Model: (A) Each nucleosome can be in one of the two, unmodified "U" or modified "M", states. The switching between the two states is controlled by transition rates r_{UM}, r_{MU}. (B) Scheme of functionally distinct spreading mechanisms (respective terms are shown as in Eq. 1). The painter region where HMEs are recruited in a sequence-specific manner is framed in red.
(C) An example of the stochastic time evolution of the epigenomic state simulated for a 201 nucleosomes-wide region in the simple painter mode.

of chromatin states. In particular, our modular approach allows to carefully examine the role of sequence-specific nucleation signals, and how they interplay with reader-writer mechanisms and 3D organisation. After a deep analysis of the generic properties of each contributions, we show that such framework allows to rationalize the maintenance of confined chromatin states and to address the role of epigenome regulation in gene expression. Finally, we contextualize our quantitative approach to several concrete experimental systems, from the formation of γH2AX domains around double-stranded breaks and the formation of heterochromatin domains around Transposable Elements (TE) insertion sites in mammals to the heterochromatin memory in fission yeast. The very good agreement between our simple model predictions and experimental data suggests that such multi-modal spreading model may provide a very promising framework to further investigate the link between genomic and epigenomic organisation during normal development, pathologies as well as during evolution.

2 MATERIALS AND METHODS

Our work consists of three main methodological parts: (i) modeling chromatin state dynamics; (ii) modeling its impact on transcription; (iii) comparing model predictions with experimental data.

2.1 Chromatin state dynamics

2.1.1 Model

As in our previous studies [21, 22, 29], chromatin is modeled by a unidimensional array of n nucleosomes. As a proxy for the local chromatin state, we assume that each nucleosome can fluctuate stochastically between a finite number of epigenetic states, each state corresponding to a specific combination of histone marks. To simplify, we consider a generic two-state model between an unmodified/neutral state (U)

and a modified, either active or inactive, state (M) (Fig. 1A). The switching dynamics from M to U (respectively from U to M) is controlled by the transition rate r_{MU} (resp. r_{UM}).

r_{MU} integrates all the molecular processes that promote the removal of the histone modifications. This includes histone turnover [54], enzymatic removal of marks by "erasers" (such as histone deacetylases or demethylases) [55] and histone dilution between sister chromatids at replication forks [56]. In our theoretical study, for simplicity, we lump all these processes into one constant $r_{MU} \equiv k_0$, except in Fig. 7D where we explicitly account for replication by forcing the transition from M to U with a probability 1/2 every generation time $T_{cyc} = 55/k_0$.

r_{UM} accounts for the spreading of the state M by dedicated "writer" enzymes: HMEs are first recruited at some positions and then may "write" the M state to any neighboring U nucleosomes (Fig. 1B).
We assume HMEs can be mobilized to chromatin via two independent, additive, sequence- or state-specific, pathways. The quantity of enzymes bound at position i, $\rho_w(i)$, is then given by $\rho_w(i) = \rho_s(i) + \Delta\delta_{i,M}$ with $\rho_s(i)$ the sequence-specific contribution, $\delta_{i,M} = 1$ if i is in state M, $= 0$ otherwise, and $0 \leq \Delta \leq 1$ is the "reader" recruitment strength accounting for the capacity of some HMEs to co-associate with the same mark they catalyze [5]. For clarity, we only consider logical distributions for ρ_s with $\rho_s(i) = 1$ at given sequence-specific recruitment sites, called the "painter" regions.
Similarly, the activity $k_w(i)$ of a bound enzyme may be sequence- or state-specific: $k_w(i) = k_s(i)(1+r\delta_{i,M})$ with $k_s(i)$ the normal enzymatic activity and $r > 1$ a boost factor accounting for the enhancement of enzymatic activity occurring for some HMEs in the presence of the same mark they catalyze [17]. For simplicity, we assume that $k_s(i) \equiv k$ is homogeneous.
We consider that a bound enzyme at position i may spread the state M not only at nucleosome i (with an on-site (*in cis*) activity $k_w(i)$) but also at longer-range (*in trans*) to any other nucleosome j at a rate $k_{trans}(i \to j) = \epsilon k_w(i) P_c(i,j)$, proportional to the intrinsic enzyme activity $k_w(i)$, to ϵ ($0 \leq \epsilon \leq 1$) a multiplicative factor accounting for a putative modulation of the writing efficiency *in trans* and $P_c(i,j)$ (≤ 1) the probability that i 'communicate' or 'contact' with j to allow spreading (see below).
Altogether, the propensity for a nucleosome i to switch from U to M is thus given by the activities of locally bound HMEs and of HMEs bound to other nucleosomes that may interact with i:

$$\begin{aligned} r_{UM}(i) &= k_w(i)\rho_w(i) + \sum_{j\neq i} \epsilon k_w(j) P_c(i,j) \rho_w(j) \\ &= k \left[(\rho_s(i) + \Delta\delta_{i,M})(1 + r\delta_{i,M}) + \sum_{j\neq i} \epsilon (1 + r\delta_{j,M})(\rho_s(j) + \Delta\delta_{j,M}) P_c(i,j) \right] \end{aligned} \quad (1)$$

2.1.2 Shape of the spreading probability

The capacity for a HME recruited at position i to spread a mark at long-range is encoded into $P_c(i,j)$. P_c may translate various physical mechanisms that allow i to 'communicate' with j.
A natural mechanism is to consider that P_c captures the frequency that i and j are in spatial proximity such that a HME bound to i may catalyze a reaction in j (Fig. 5). In this case, assuming that the 3D chromatin organization equilibrates locally faster than the epigenomic landscape, $P_c(i,j)$ can be approximated by the contact probability between i and j [60] as observed in Hi-C experiments [61]. Indeed, for short genomic distances between i and j (< 100kbp), theoretical and experimental studies suggest that, at this scale, 3D looping rates are $\sim \text{sec}^{-1}$ to min^{-1} [57, 58, 62, 63, 124] while the chromatin modification rates are often $\sim \text{h}^{-1}$ [29, 59]. In this paper, for most of our theoretical results (except in Fig. 7C,D and some cases in Fig. 5, see below), we consider a homogeneous 3D contact probability $P_c(i,j) = 1/|i-j|^\gamma$, with γ reflecting the average compaction level of the chromatin fiber. Experimentally, γ has been shown to vary in the range $[0.5 - 1.5]$ depending on cell cycle stage, organism and/or cell fate [64–66]. In particular, we use the standard, intermediate value $\gamma = 1$. Note that the value of γ does not impact qualitatively the main conclusions of our work but may quantitatively have significant effects on the spreading process (Supplementary Fig. 1) and thus must be carefully adjusted if possible when considering specific experimental systems. In Fig. 7C,D and in Fig. 5, we also consider, for $P_c(i,j)$, contact matrices, representative of chromatin regions with (i) a central strongly-self-interacting domain

of nine nucleosomes; or (2) long-range loops between the painter region (positions $[-2:2]$) and five nucleosomes at positions $[48:52]$. In both cases, $P_c(i,j)$ was computed from polymer simulations using the lattice kinetic Monte-Carlo model developed in [58] (see Supplementary Information for details).
In Fig. 5, we also consider three alternative mechanisms for P_c: (1) HMEs bound at position i may impact only nearest-neighbor (NN) nucleosomes, i.e. $P_c(i,j)=1$ if $|j-i|=1$ and $P_c(i,j)=0$ otherwise; (ii) HMEs recruited at i, while unbinding and diffusing, are more likely to rebind in the 3D vicinity, leading to a concentration gradient around i and to an effective $P_c(i,j)=1/|i-j|^{0.5}$ (see Supplementary Information); (3) HMEs bound at i may spread a mark to a distal nucleosome j only if they are placed in very closed proximity by a loop extruding factor (like cohesin or condensin) [67–70] translocating along the chromatin, leading to $P_c(i,j)=\exp(-|j-i|/s_0)$ (see Supplementary information and Supplementary Fig. 2).

2.1.3 Stochastic simulations

For a given set of parameters, the stochastic dynamics of the system is simulated using the standard Gillespie algorithm [71] implemented on Python (can be downloaded at https://github.com/physical-biology-of-chromatin/Painter-Model). In all our work, we simulated the dynamics of $n=201$ nucleosomes corresponding to $\sim$ 40kbp-long genomic region. Starting from a random initial macro-state (i.e random choice between U or M states for each nucleosome), the system relaxes to a steady state. Each simulation corresponds to a "single cell" trajectory of the local epigenetic state (Fig. 1C). Unless specified (Fig. 6, 7), characterization of the system was done at steady-state. In each condition, 500 different trajectories were simulated. Times are given in $1/k_0$ unit that characterizes the typical turnover time of histone marks and is $\sim h$ [29, 54].

2.1.4 Analytical solutions

In the case where state-specific terms are negligible ($r=\Delta=0$, painter mode), the steady-state probability $P(M_i)$ of being in modified state M at any position i is simply given by :

$$P(M_i) = \frac{r_{UM}(i)}{r_{UM}(i)+r_{MU}(i)} \tag{2}$$

$$= \frac{(k/k_0)(\rho_s(i)+\varepsilon\sum_{j\neq i}\rho_s(j)*P_c(i,j))}{(k/k_0)(\rho_s(i)+\varepsilon\sum_{j\neq i}\rho_s(j)*P_c(i,j))+1}. \tag{3}$$

In the case where only state-specific recruitment is negligible ($\Delta=0$) and where the sequence-specific recruitment is localized in a finite painter region, by doing a mean-field approximation, the average probability of the M-state inside the painter region $\overline{P}_p$ is given by analytically solving (see Supplementary Information and Supplementary Figure 3):

$$\overline{P}_p = \frac{1+\varepsilon\sum_{j\neq i}P_c(i,j)(1+r\overline{P}_p)}{1+\varepsilon\sum_{j\neq i}P_c(i,j)(1+r\overline{P}_p)+(k_0/k)}. \tag{4}$$

Outside the painter region, $P(M_i)$ is then given by replacing ε by $\varepsilon(1+r\overline{P}_p)$ in Eq. 3.

2.1.5 Enzyme limitation

In the model described above, we assumed that the concentration of HME is large enough to not have to consider the depletion of the pool of freely-diffusing HMEs (that impacts the sequence- and state-specific recruitment strength) by bound HMEs. In Fig. 7, we consider a scenario where the number of enzymes is actually limited. Assuming fast binding-unbinding kinetics for the enzymes, we can show (see Supplementary Information and Supplementary Fig. 4) that such system is equivalent to our simple model (Eq. 1) but by replacing the state-dependent parameter Δ by $P(E)$ the probability that a HME is bound to a modified nucleosome. $P(E)$ explicitly depends on N_e^{tot} the total number of HMEs, k_b and k_u the binding and unbinding rates (see Eq. 7 in Supplementary Information). In the limit of a non-limiting number of enzymes ($N_e^{tot}\gg 1$), $P(E)\approx k_bN_e^{tot}/(k_u+k_bN_e^{tot})\equiv\Delta$. In Fig. 7, we fix $k_b/k_u=1$.

2.1.6 Contact domains

The spreading probability $P_c(i,j)$ represents how frequent pairs of nucleosomes are 'in (effective) contact'. Therefore, the genomic region is not only a unidimensional string of beads with only NN links but a network of interactions. In order to capture at the single-cell level, the propensity of nucleosome in the M-state to be 'connected', we randomly generated for every configuration an undirected graph (M_{net}) for the corresponding set of M-states (Supplementary Fig. 5). More precisely, for all pairs (i,j) of nucleosomes in this set, an edge between them is inserted in M_{net} with a probability $P_c(i,j)$. Then, we infer the disconnected subgraphs of M_{net} that would represent the different clusters of connected M-state nucleosomes. Finally, for each node/nucleosome i, we compute the size of the contact domain it belongs to.

In the reader-writer mode, a M-state nucleosome may influence a U-state nucleosome via P_c. Thus, to behave coherently, a genomic region may not necessarily need to have all nucleosomes in the M-state, but rather that all M-states belong to the same graph and that all U-states can be reached via $P_c(i,j)$ and are connected to such network. To quantify this, for each configuration, we consider the largest M-state subgraph $M_{net}*$ in M_{net}. For each U-state nucleosome i, random links with any M-state nucleosome j of $M_{net}*$ are generated with probability $P_c(i,j)$. i is added to $M_{net}*$ if it has been linked to at least two M-state nodes of $M_{net}*$ (see Supplementary Fig. 5, Fig. 5C). The size of such extended contact domain can thus be considered analogous to a percolation parameter [72], indicative of the extent of M-state spreading in the region.

2.1.7 Nucleosome state correlations

To compute the spatial correlations between nucleosome states, we divide the genomic region into 40 bins of 5 nucleosomes. In one configuration, these bins are thus characterized by a M-state ranging from 0 to 5. Nucleosome state correlation matrix shown in Fig. 2, 3, 4C represents the Pearson correlation between the M-states computed for each pair of nucleosomes over all the configurations extracted from the simulations.

2.2 Transcription dynamics

Assimilating the modified state M to a silencing state, we assume that the instantaneous transcription rate α of a gene localized inside the region of interest is negatively impacted by the current proportion of M-state nucleosomes inside the promoter region. More precisely, we use the cooperative switch model of transcription (Fig. 8A), proposed by Zerihun *et al.* [22].

$$\alpha(\overline{M}) = \alpha_0 \frac{\tanh((\overline{M} - \overline{M}^*)/d) - \tanh((-1 - \overline{M}^*)/d)}{\tanh((1 - \overline{M}^*)/d) - \tanh(-1 - \overline{M}^*)/d)} + \alpha_1 \tag{5}$$

where $\alpha_0 + \alpha_1$ is the maximal transcription rate, α_1 is the minimal, leaky, rate, $\overline{M} = (1/n)\sum \delta_{i,M}$ the current proportion of modified state in a n nucleosome-wide promoter region around, $\overline{M}^*$ the critical value of modification above which the gene is repressed and d defines the sharpness of the transition between activation and repression.

Gene expression dynamics is simulated along with the chromatin state dynamics within the same Gillespie simulations by considering, in addition to $\{r_{UM}(i), r_{MU}(i)\}$, the transcription of mRNA molecules at rate α (Eq. 5), their degradation at rate β, their translation into proteins at rate Γ, the degradation of proteins at rate Ω and the maturation of proteins at rate μ [73]. In Fig. 8, we fix $n = 5$, $\alpha_1 = 0.2\alpha_0$, $d = 0.05$, $\overline{M}^* = 0.15$, $\beta = 0.1\alpha_0$, $\Gamma = \alpha_0$, $\Omega = 0.3\alpha_0$ and $\mu = 0.1\alpha_0$.

2.3 Analysis of experimental data

In Fig. 9A, the average experimental profile $S(i)$ of H2A.X phosphorylation in human cells around double strand breaks is extracted from Sup. Fig. 1k (right) of Arnould *et al.* [70]. We fit the experimental data

by the chromatin state model in the painter mode ($r = \Delta = 0$) with one painter (ATM kinase) at position 0 (see Suppl. Fig. 1k, left of [70]) for a ±1Mbp-region around the break by using Eq. 3 and assuming that $S(i) = AP(M_i) + B$ with A and B two constants. Parameter inference is done by minimizing the L2-distance between predictions and experiments.

Similarly, in Fig. 9B, the average density of H3K9me3 modifications flanking a transposable element in mESC cells is obtained from Robollo *et al.* [74]. Fit to experimental data using Eq. 3 is performed for a 10kbp-region with a centered painter region of size 7 nucleosomes corresponding to 1kbp transposable element. Parameter inference is done by minimizing the L2-distance between predictions and experiments.

In Fig. 9C the distribution of fluorescence of the reporter gene as a function of time and the corresponding fractions of "OFF" cells are extracted from Ragunathan *et al.* [10] for different strains. Modeling of the system is performed using the chromatin state+transcription model (see above) with $k = 0.05$, $\varepsilon = 0.6$, $n = 5$, $\alpha_0 = 1$, $\alpha_1 = 0.2\alpha_0$, $d = 0.05$, $\overline{M}^* = 0.1$, $\beta = 0.1\alpha_0$ (in min^{-1}) and the same parameters for protein dynamics as above. A scaling is applied to convert the number of proteins p into a fluorescence level F ($F = p + C$ with $C = 380$). The fraction of OFF cells is defined by the fraction of cells whose fluorescence is below 530 as defined experimentally [10]. The extracted experimental fluorescence at 6 hours is not aligned with the fluorescence peak observed at other time points. Fit to experimental data is performed by minimizing $\chi^2 = \sum_{t_i}[(OFF_{pred}(t_i) - OFF_{exp}(t_i))/OFF_{exp}(t_i)]^2$, where $OFF_{pred}(t_i)$ is the predicted fraction of OFF cells at time t_i and $OFF_{exp}(t_i)$ the corresponding experimental observation, using a simple grid-search algorithm to scan parameters (see Supplementary Fig. 6).

3 RESULTS

3.1 A generic model of epigenomic regulation

To investigate in detail the main mechanisms driving epigenomic regulation, we develop a simple generic two-state model (Fig. 1A) where the local chromatin state can fluctuate between an unmodified, neutral state U and a modified state M that may correspond to an active (e.g. H3K4me3 or H3K27ac) or repressive (e.g. H3K27me3 or H3K9me2/3) state. We model chromatin as a unidimensional string of nucleosomes, each nucleosome being in state U or M. We consider a ∼40-kbp-long genomic domain (201 nucleosomes), a size that typically encompasses epigenetically-regulated regions like transposable elements [75], gene promoters and enhancers, the Mating type locus or subtelomeres in yeast [76].

The stochastic dynamics of the state is driven by histone-modifying enzymes (HMEs) that deposit ("writers", transition from U to M) or remove ("erasers", from M to U) specific histone modifications [5,10,16]. For example, PRC2 via its methyltransferase subunits EZH1 or EZH2 catalyzes the methylation of H3K27 [2, 77]. Our model integrates the key mechanisms acting on HME recruitment or activity, focusing on "writer" enzymes and lumping all the processes participating in histone mark removal into one effective turnover rate k_0 (see Materials and Methods for details). Briefly (Fig. 1B), we have delineated the chromatin association of HMEs into sequence- and state-specific contributions: (a) writers, in association with DNA-binding proteins, may localize around specific genomic location with probability ρ_s, or (b) HME may be recruited to M-state nucleosomes with efficiency Δ. Recruited HMEs at position i can then write the epigenetic mark on-site with a rate k (action *in cis*) but may also spread it to a distal nucleosome j with a rate $(\varepsilon k)P_c(i, j)$ (action *in trans*) [15], where $P_c(i, j)$ accounts for the capacity of two nucleosomes to interact and may depend on the exact spreading mechanism (see Materials and Methods). In the following, we consider that bound HMEs may catalyze reaction in their 3D neighborhood, as evidenced experimentally for several epigenetic systems including H3K27 methylation [15, 83], H3K9 methylation [78], γH2AX phosphorylation around DNA double strand breaks [79] and $P_c(i, j)$ is thus defined as the average 3D contact frequency between two genomic regions. HME enzymatic activity may be also boosted by allostery by a factor r if bound to a M-state nucleosome [17, 77, 80]. State-specific effects (recruitment and allosteric boost) are the so-called reader-writer mechanisms that are thought to be crucial for the establishment and maintenance of many epigenomic states [16].

In the following sections, we will systematically dissect the role of sequence-specific recruitment and of

these reader-writer processes in epigenetic regulation using extensive simulations of our generic two-state model.

3.2 Spreading by sequence-dependent recruitment of enzymes: the "painter" mode

We first investigate the contribution of a simple mode of spreading where enzymes bind only to specific genomic locations and then may spread epigenetic marks in the 3D vicinity, in absence of reader-writer mechanisms. In this case, HMEs can be considered as "painters" that sit at specific places along the genome and "paint" the surrounding chromatin with a given modification. As sequence-dependent recruitment of enzymes is essential for *de novo* establishment of epigenetic domains [13, 81] and as many writers do not carry a "reader" subunit (that may drive state-dependent effects), such mode of regulation, hereafter called the "painter" mode, is likely to be a major way of epigenetic spreading.

To illustrate this model, we consider a simple case where painters are only recruited to a single 1kbp-long locus located in the middle of the 40kbp-long region (yellow area in Fig. 2A). In this situation, the probability $P(M)$ to have a modified nucleosome at a given position can be derived analytically (see Eq. 3 in Materials and Methods) and only depends on the ratio (k/k_0) between the cis-spreading rate k and the turnover rate k_0 and on the ratio ε between the trans- and cis-activities (Fig. 2A,B). For all positions, we observe that $P(M)$ is a gradually increasing function of both parameters (Fig. 2B, Supplementary Fig. 7). In the limit of very low trans-efficiency ($\varepsilon \approx 0$), modified states are essentially confined within the recruitment region due to the remaining on-site writing activity. For larger ε values, $P(M)$ is peaked at the painter region and decays at large genomic distances. The stronger ε and k/k_0 the wider the peak in $P(M)$ profiles and thus longer the range of spreading of the M-state (Fig. 2A).

Actually, the observed decay of $P(M)$ outside the recruitment zone is translating the decay of contact probability P_c with the painters bound to recruitment zone. More generally, in the limit of low contact probability or low spreading efficiency ($(k/k_0)P_c\varepsilon \ll 1$), $P(M)$ is directly proportional to the spreading probability P_c (see Eq. 3) which, in our case, is taken to be $\propto 1/s^\gamma$ with s the genomic distance to the painter region and $\gamma = 1$ (see also Supplementary Fig. 1). This illustrates the direct relationship between the local 3D organization and the profiles of epigenomic marks around recruitment sites [82].

In the painter mode, the local chromatin state only depends on the position of the bound painters and does not feedback on the recruitment of HMEs or on their activity. This lack of cooperativity leads to the absence of correlation between the nucleosome states at two different positions along the region (Fig. 2C). To quantify the efficiency of the spreading mechanism and its capacity, at the single-cell level, to form more or less expanded, coherent M-state domain, we estimate how nucleosomes in the M-state are effectively 'colocalized' in space. For each configuration and each M-state nucleosome, we compute the number of other M-state nucleosomes present in the same 3D contact domain (see Materials and Methods). Fig. 2D shows that, in average, except around the painter region, M-state nucleosomes are isolated from each other and only part of small domains.

3.3 State-dependent enzymatic activity: the "boosted-painter" mode

In some systems such as Polycomb regulation (associated with the H3K27me3 modified state), it has been observed that the activity of the writers can be "boosted" by the presence of pre-existing modifications [17, 77, 80, 83, 84]. For example, the binding of the Polycomb writer PRC2 to H3K27me3 marked nucleosomes, triggers a boost in methyltransferase activity of its subunit EZH2 by allostery, via its other subunit EED. In our theoretical framework, this effect can be formalized by an increase in the writing rate of HMEs bound to a modified nucleosome via a multiplicative factor $r > 1$ (see Materials and Methods). For example, for mammalian PRC2, in vitro experiments suggested an allosteric boost of $r \lesssim 10$ fold in presence of H3K27me3 peptides [17, 84].

In Fig. 3, we characterize the impact of this boost on the simple painter mode described above (stably bound writers at a specific 1kbp-long region), still neglecting state-specific recruitment of HMEs ($\Delta = 0$). As shown in Fig. 3A, a weak ε value that essentially confines the M-state to the painter region in the simple painter mode ($r = 0$, green line), can be compensated by a strong boost term ($r = 20$, black line)

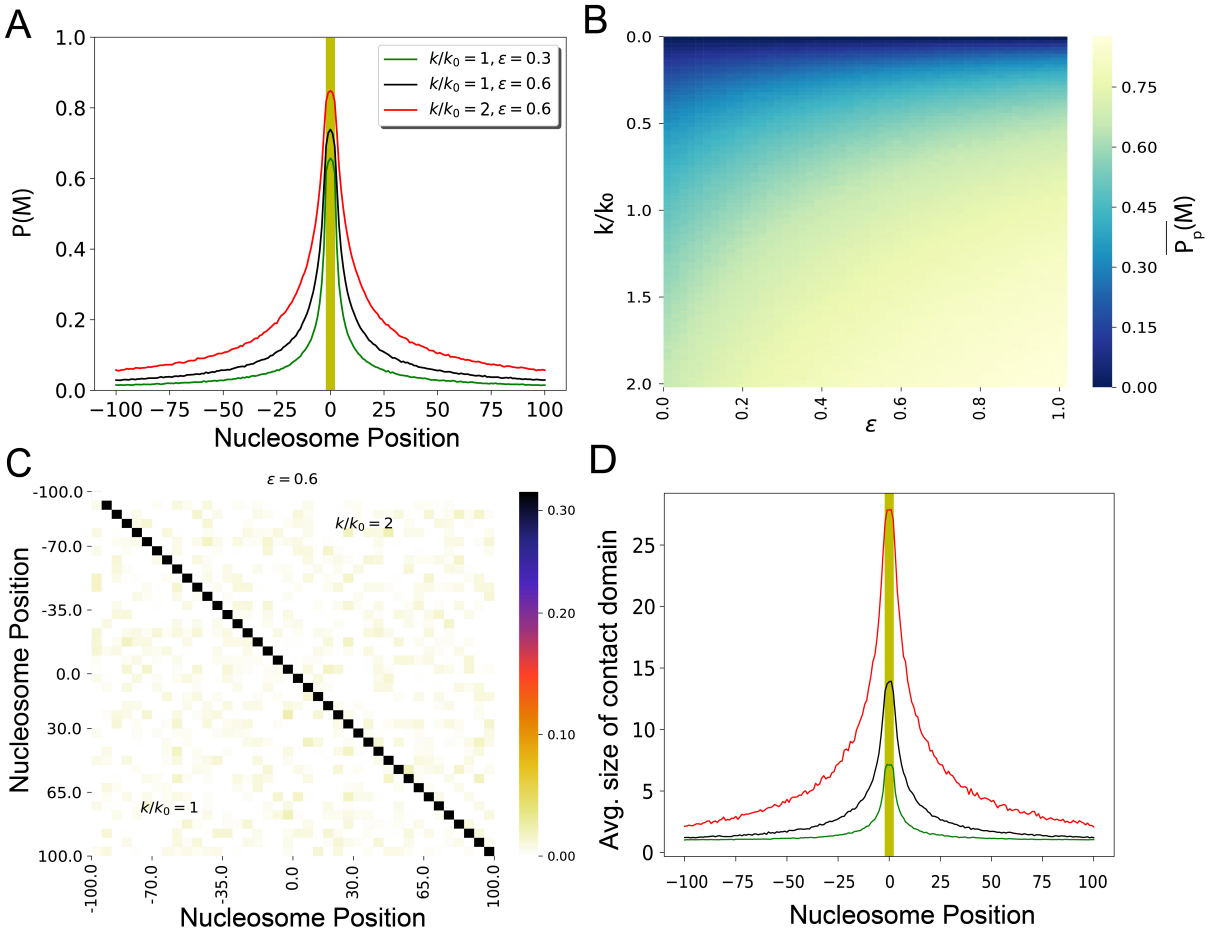

Figure 2. Sequence-specific recruitment - Painter mode : (A) Virtual Chip-seq profiles of the M-state. Probability $P(M)$ to be modified as a function of the position around the 5 nucleosomes-long region where HMEs are recruited (yellow region) for different spreading efficiency ε and painter activity k/k_0. (B) Average probability $\overline{P}_p(M)$ of the M-state within the recruitment region as a function of ε and k/k_0. (C) Spatial correlation matrix between the nucleosome state at two positions (see Materials and Methods) for $\varepsilon = 0.6$ and $k/k_0 = 1$ (lower part) or $= 2$ (upper part). (D) Average M-state contact domain size as a function of the nucleosome position for different ε and k/k_0 values.

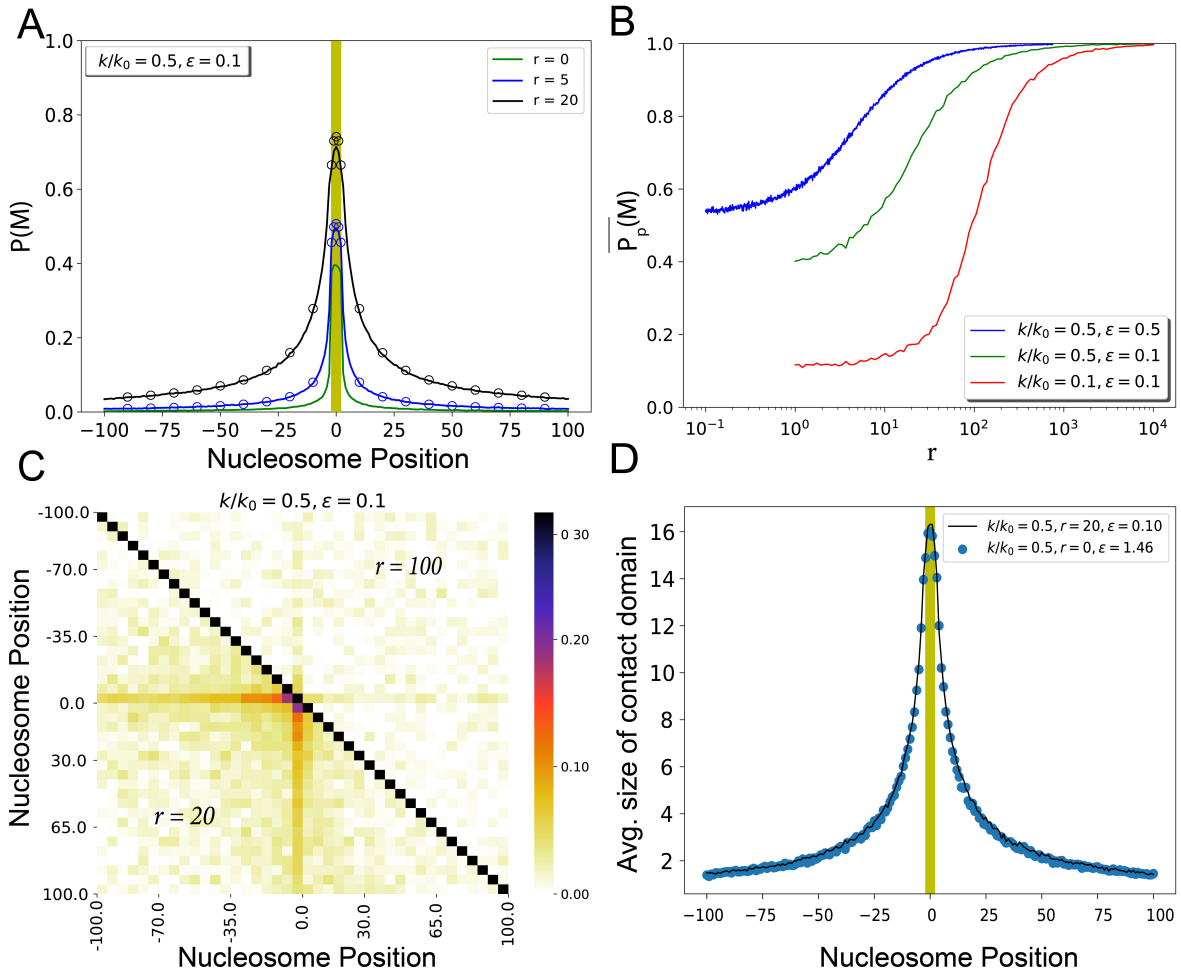

Figure 3. State-specific enzyme activity - boosted painter mode : (A) Virtual Chip-seq profiles of the M-state. $P(M)$ as a function of the position around the 5 nucleosomes-long region where HMEs are recruited (yellow region) for different values of the boost factor r. The scatter points are the analytical solution of the model. (B) $\overline{P}_p(M)$ within the recruitment region as a function of r. (C) Spatial correlation matrix between the nucleosome state at two positions for $\varepsilon = 0.1$, $k/k_0 = 0.5$ and $r = 20$ (lower part) or $= 100$ (upper part). (D) Average M-state contact domain size as a function of the nucleosome position with boost $r = 20$ (full line). Dots correspond to a simple painter mode with an adjusted ε value to reproduce the $P(M)$ profile.

with a significant overall enhancement of $P(M)$: the presence of modified nucleosomes in the painter region due to the "on site", cis-activity of painters (controlled by k/k_0) boosts globally the cis- and trans-spreading capacity of bound writers (via r) and favors further spreading inside and outside the painter region.

To get a better understanding of this "boosted-painter" mode, we first focus on the painter region by computing $\overline{P}_p(M)$, the average value of $P(M)$ inside this region (Fig. 3B). As expected, $\overline{P}_p(M)$, is an increasing function of k/k_0 and r. However, contrary to the simple painter mode (Supplementary Fig. 7B), the transition from low to high M-state is a much sharper sigmoid function. Such switch-like behavior suggests a phase transition [27] inside the painter region that reflects the cooperative dynamics between writer-bound nucleosomes due to the state-dependent boost. The sharpness and position of the transition depends on k/k_0 and ε. High values lead (i) to smoother transition as the simple painter mode is significant enough to buffer the boost effect and (ii) to lower critical r-values as less boost is required to get high enzymatic activity.

In this spreading mode, the chromatin states of nucleosomes localized outside the painter region are now dynamically coupled to the ones inside, leading to positive spatial correlations that are maximal for r-values around the critical boost (Fig. 3C). Such coupling makes an exact analytical treatment of $P(M)$ intractable. However, using a mean-field approximation inside the painter region, we can derive an expression for $\overline{P}_p(M)$ (see Materials and Methods and Supplementary Information). Outside the painter region, $P(M)$ is thus just as in the simple painter mode but for a boosted trans-activity $\varepsilon(1 + r\overline{P}_p(M))$ (blue and black circles in Fig. 3A). Hence, regarding the profile of epigenomic mark and of coherent 3D contact domain size (Fig. 3D), the boosted painter mode is indistinguishable from a simple painter mode with a greater "on site" and "off site" spreading rates (Blue dots, Fig. 3D). Only the presence of spatial correlations of nucleosome states with the recruitment region (Fig. 3C) might be a clear signature of this mode of propagation.

3.4 State-specific recruitment of enzymes: the "reader-writer" mode

Some writer enzymes have been shown to be recruited in a state-dependent manner to chromatin, having the ability to "read" (i.e. to be recruited at) a particular histone tail modification and "write" the same on another nucleosome in the 3D vicinity [2, 3, 5, 10]. For example, the methyltransferase Clr4 (associated with heterochromatin formation and H3K9me2/3 modifications in fission yeast) contains a chromodomain that may trigger its recruitment by H3K9me2/3 [10]. In our framework, we introduce this "reader-writer" mode by assuming that the probability of finding a HME bound at a M-state nucleosome is enhanced by a factor Δ (see Materials and Methods). In the following part of this section, we focus on this effect coupled to the simple painter mode neglecting possible state-specific boost ($r = 0$).

In that case, thanks to the long-range action of M-state bound writers, the state dynamics of every nucleosome is coupled to the states of all the other nucleosomes which is well illustrated in Fig. 4C by the global increase in spatial correlations. Such reader-writer mode introduces a positive feedback in the global M-state dynamics which has been shown to promote the formation and inheritance of extended, stable M state domains [18, 32]. As shown in Fig. 4A, when combined with the simple painter mode, state-specific recruitment strongly modifies the spreading pattern by notably increasing M-state occurrence away from painter region. In particular, we observe heavy tails for $P(M)$ (black and blue lines in Fig. 4A) that qualitatively differs from the typical $1/s^\gamma$ contact decay observed for the sole painter mode (red dotted curve in Fig. 4A). Such signature (heavy tails, deviation from simple painter model) in experimental profiles of epigenomic marks may thus be suggestive of a dominant reader-writer mechanism. In this regime, we also observe that the reader-writer process facilitates the formation of expanded 3D M-state contact domains away from the sequence-dependent recruitment region (fig. 4D) unlike the simple and boosted painter modes.

To quantify the resulting global increase in $P(M)$, we systematically compute $\overline{P(M)}$, the average value of $P(M)$ inside the whole genomic region (Fig. 4B). For a given trans-spreading activity (ε), we observe a sharp transition when state-specific recruitment efficiency (Δ) augments: from a pure simple painter profile localized around the painter region ($\overline{P(M)} \ll 0.5$) to a globally modified state ($\overline{P(M)} \gg 0.5$). As in the boosted-painter mode, this also suggests a phase transition but here that reflects the cooperative

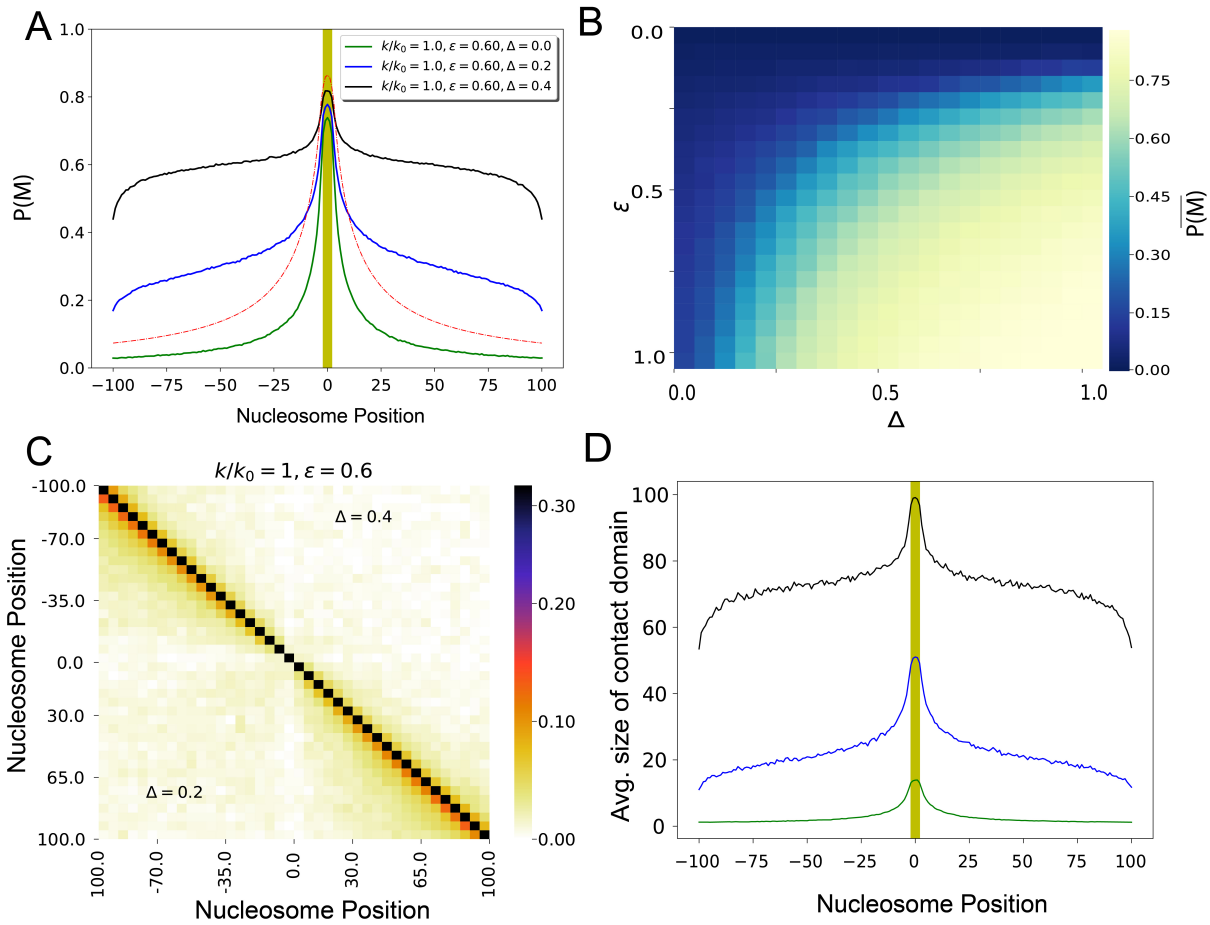

Figure 4. State-specific recruitment of enzyme - reader-writer mode: (A) Virtual Chip-seq profiles of the M-state. $P(M)$ as a function of the position around the 5 nucleosomes-long region where HMEs are recruited (sequence specific) for different values of reader recruitment strength Δ (at the critical point $\Delta = 0.2$ and above $= 0.4$). (B) Average value $\overline{P(M)}$ of $P(M)$ inside the whole genomic region as a function of ε and Δ ($k/k_0 = 1$). (C) Spatial correlation matrix between the nucleosome state at two positions for $\varepsilon = 0.6$, $k/k_0 = 1$ and $\Delta = 0.2$ (lower part) or $= 0.4$ (upper part). (D) Average M-state contact domain size as a function of the nucleosome position for different Δ values. Colors as in (A).

dynamics between *all* nucleosomes of the region (see also next section). Around the critical Δ-value (e.g. $\Delta_c \approx 0.2$ for $k/k_0 = 1$ and $\varepsilon = 0.6$), fluctuations in $P(M)$ are maximal and lead to high inter-nucleosome correlations (Fig. 4C and Supplementary Fig.8).

3.5 The spreading probability drives the percolation of the chromatin landscape

A key element of the different writing modes investigated above is the capacity of recruited HMEs to spread an epigenetic signal [2, 3, 5, 10]. This property depends on the spreading probability $P_c(i, j)$ that captures the ability of two nucleosomes to interact. In the previous sections, we assumed that $P_c(i, j) = 1/|j - i|^\gamma$ with $\gamma = 1$, a generic scaling law accounting for a mechanism of spreading via 3D contacts as already observed for some HMEs like Polycomb group proteins (H3K27 methylation) [15], Clr4 methyltransferase (H3K9) [10], ATM kinase (γH2AX) [70] and characteristic of the polymeric nature of chromatin [92]. In this section, we explore the role of the shape of $P_c(i, j)$ in epigenomic regulation.

This shape may depend on the specific spreading mechanism and on the experimental system or genomic region under study. Here, we consider five alternative forms for $P_c(i, j)$ (see Materials and Methods). (1) Two that still correspond to a 3D contact spreading process but with contextualized, heterogeneous $P_c(i, j)$ (Supplementary Fig. 9A): one for a small 3D compact domain localized around the recruitment zone and another for a region with a strong loop between the painter area and a locus 10kbp away, mimicking, for example, a repressed MAT locus in yeasts and a promoter-enhancer loop in mammals, respectively. (2) A shape accounting for a spreading to only nearest-neighbor (NN) nucleosomes that might be relevant in scenarios where an epigenomic state propagates unidimensionally along the chromatin [24, 91, 115]. (3) A form compatible with an effective contact spreading mechanism where HMEs recruited at a given position may diffuse in 3D to nearby positions and thus impact the state of distal nucleosomes ($P_c(i, j) = 1/|j - i|^{0.5}$). (4) A scenario where two nucleosomes may influence each other only if they have been placed in very close proximity by loop extruding factors (eg, cohesin or condensin) that are molecular motors translocating along the chromatin and implicated in TAD formation in mammals [67–70] ($P_c(i, j) = e^{-|j-i|/s_0}$).

In the simple painter mode (Fig. 5A, $\Delta = 0$), spreading from the central recruitment zone is limited by the shape of $P_c(i, j)$: long-range effective interactions like in the diffusion scenario lead to more extended $P(M)$ profiles compared to localized spreading as the NN case. In the 3D loop case, the M-state is able to stably propagate distally thanks to enriched 3D contacts with the painter area of sequence-specific recruitment.

In the reader-writer mode (Fig. 5A, $\Delta > 0$), as observed in the previous section, state-specific recruitment allows a facilitated spreading and modifies the shape of $P(M)$. For a given reader-writer strength Δ, scenarios with longer-range $P_c(i, j)$ are more impacted. However, for all scenarios, it exists a critical Δ value (Δ_c) above which the M-state spreads over the whole region (Fig. 5B), more local $P_c(i, j)$ shapes (eg, NN or compact domain cases) exhibiting larger Δ_c values. Such transition driven by the reader-writer capacity of HMEs resembles actually to a generic percolation transition [27] where the system starts to be fully connected (or percolated) [72]. In our context, it corresponds to a situation where all the M-state nucleosomes form a unique expanded contact domain from which all the remaining U-state nucleosomes can be reached via $P_c(i, j)$ to allow spreading (see Materials and Methods, Supplementary Fig.5,Fig.10). To quantify this, for each configuration, we look how the most expanded M-state contact domain is connected to the unmodified regions. Fig. 5C shows the evolution of the average size of such extended generalized contact domains (containing U- and M-state nucleosomes) for three different scenarios. The global epigenomic state becomes percolated when this size approaches 200, meaning that the extended domain covers the whole genomic region (Fig. 5D). Interestingly, a full overall M-state ($\overline{P(M)} \sim 1$) is not required to reach percolation, in particular for long-range spreading mechanism. For example, in the $P_c(i, j) = 1/|j - i|$ case, percolation transition occurs at $\Delta_c \approx 0.2$ where $\overline{P(M)} \approx 0.3$ (Supplementary Fig. 10).

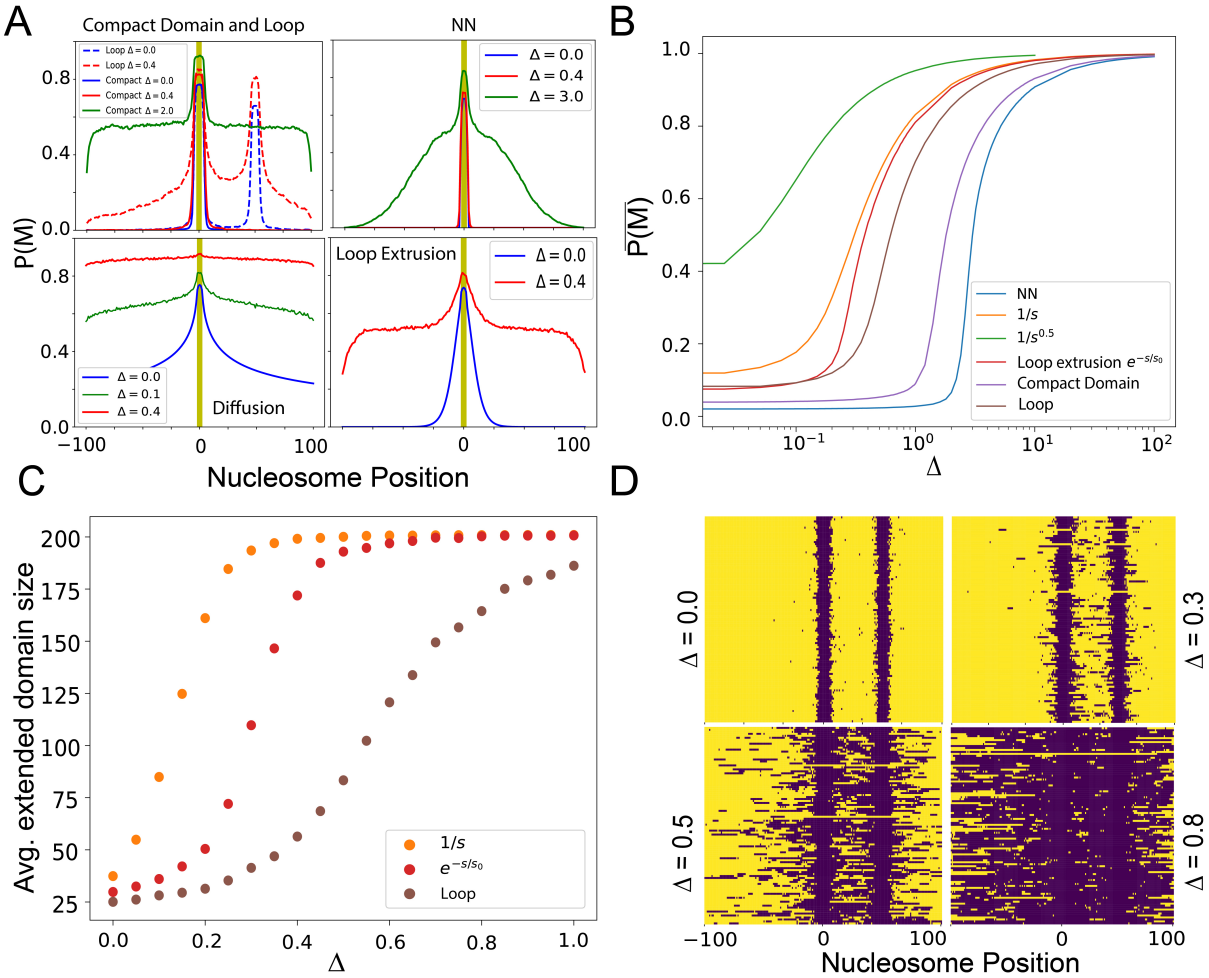

Figure 5. Spreading probability and percolation. (A) Virtual Chip-seq profiles of the M-state. $P(M)$ as a function of the position around the 5 nucleosomes-long region where HMEs are recruited (yellow) for different $P_c(i,j)$ (see Materials and Methods) and for $k/k_0 = 1$, $\varepsilon = 0.6$ with $\Delta = 0$ (painter mode, Blue) and $\Delta > 0$ (reader-writer mode, Red and Green). (B) Average value $\overline{P(M)}$ of $P(M)$ inside the whole genomic region in the reader-writer mode as a function of Δ ($k/k_0 = 1$, $\varepsilon = 0.6$) for different $P_c(i,j)$. (C) Average size of the most extended generalized domain as a function of Δ for three situations described in (B). (D) Illustration of the percolation of the system. For four different Δ values in the loop case (brown curves in (B) and (C)), most extended generalized domains observed in different configurations (100 examples) are stacked together to show cell to cell variability (all nucleosomes belonging to a domain are colored in blue).

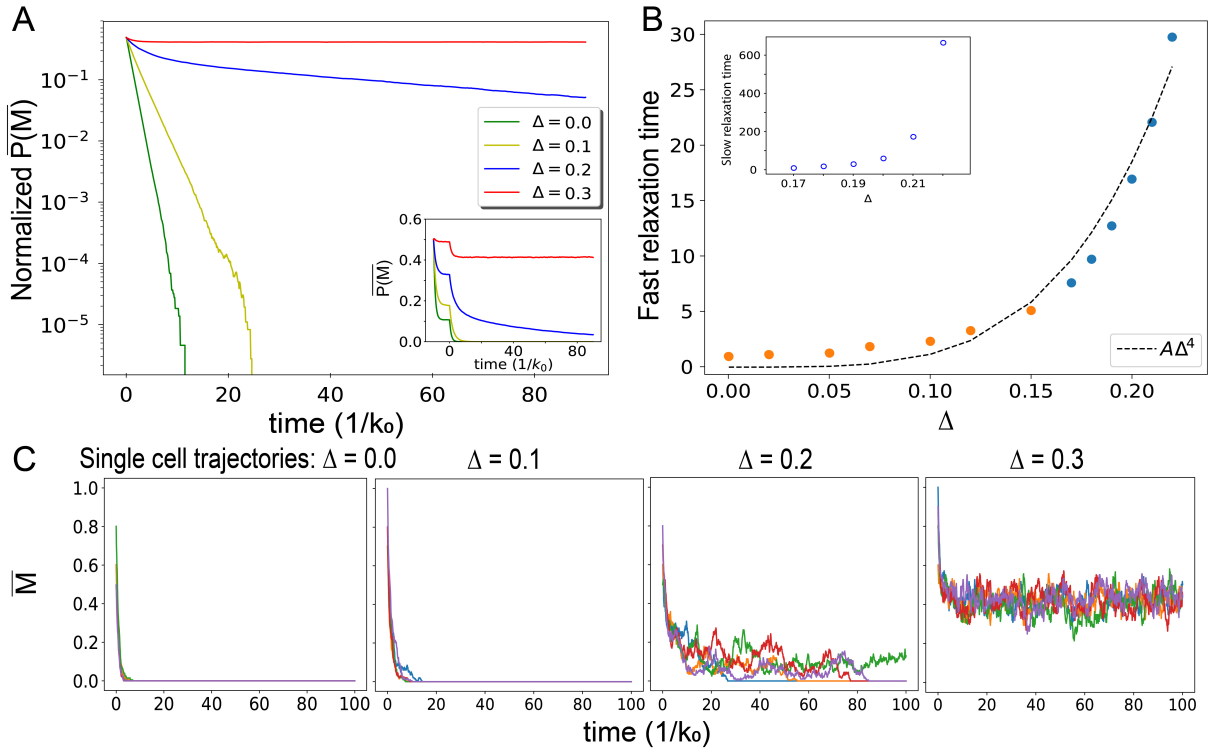

Figure 6. Epigenetic memory: (A) Normalized time-evolution of $\overline{P(M)}$ after the unbinding of sequence recruited painters at $t = 0$ for different values of Δ and with $k/k_0 = 1, \varepsilon = 0.6$. Inset shows the non-normalized evolution. (B) Fast (main) and slow (inset) relaxation times as a function of Δ obtained by fitting the curve in (A) by one or two time-scale exponential decay (see text). The fast time scales as $\approx \Delta^4$. (C) Examples of "single cell" time-evolution of the M state decay ($\sum \delta_{i,M}/201$ vs t) for different values of Δ.

3.6 The reader-writer mode may lead to epigenetic memory

Having characterized how the various writing modes establish typical epigenomic profiles from the painter region, we now investigate how these different modes impact the "epigenetic memory" of these states. By definition, epigenetic memory stands for the ability of maintaining a given transcriptional or chromatin state in the absence of the initial, sequence-dependent stimulus (e.g. transcription factors or HMEs).

For such purpose, focusing on a 3D spreading mechanism ($P_c(i,j) = 1/|j-i|$), we follow the time evolution of the mean genomic M-state $\overline{M} \equiv 1/n \sum \delta(i, M)$ once the sequence-specific recruitment of painters have been released (for examples of individual trajectories, see Fig. 6C). In Fig. 6A, we plot the ensemble average-value $\overline{P(M)}$ of $\overline{M}$ for different situations. In the simple painter mode ($\Delta = 0$, green line), the M-state relaxes very quickly to the U state. Indeed, in absence of recruited writers in the painter area, only transitions from M to U are possible and the global state decays exponentially with a characteristic time $1/k_0$. Similar behaviors are observed in the booster-painter mode (Supplementary Fig. 11).

In the reader-writer mode, for low Δ values, the decay remains fast and exponential ($\Delta = 0.1$ in Fig. 6A,C) with a relaxation time that increases with Δ (orange dots in Fig. 6B). Above a critical value, state-dependent recruitment is strong enough to stably maintain a M-state after HMEs unbind from the painter region ($\Delta = 0.3$ in Fig. 6A,C) and to keep the memory of an initial (even small) M-state enrichment (Supplementary Fig. 12). The initial profile along the genome is lost and gives rise to an uniform spreading of the M-state (Supplementary Fig. 13).

Interestingly, for Δ-values just below this critical point, the relaxation dynamics is better described by a two time-scale exponential (Supplementary Fig. 14) kinetics ($\Delta = 0.2$ in Fig. 6A,C): a fast initial decay following the trend of the low-Δ regime (blue dots in Fig. 6B) and a slow decay at larger time (inset in Fig. 6B). In this regime, random $M \to U$ conversions dominate but the state-dependent recruitment allows self-maintenance of a coherent M-state for long time periods before reaching a global, absorbing U-state ($\overline{M} = 0$) from which the system cannot escape as also observed by [27]. Such maintenance time-period is very stochastic ($\Delta = 0.2$ in Fig. 6C) while, in the low and high Δ-regimes, convergence to steady-state behaves quite uniformly.

3.7 Maintenance of a confined chromatin state requires chromatin compaction and enzymatic titration

In the previous section, we showed that epigenetic memory was possible for strong-enough state-specific recruitment. However, in this regime of parameter, the spreading of the M-state along the genome is not constrained [39], leading to unconfined memory [24] (Fig. 7A). Of course, experimentally, epigenomic domains are confined to specific regions in the genome and do not spread ubiquitously [10, 16]. In this section, we investigate what could be the minimal changes to our model to support confined memory and the maintenance of a localized M-state even in absence of sequence-specific recruitment.

Our simple epigenomic model assume implicitly that the number of HMEs is not limiting in the system. Recent quantitative proteomics experiments however suggest that such assumption might not be satisfied for some epigenomic marks like the Polycomb/H3K27me3 system during fly embryogenesis [85]. Therefore, a mechanism that could possibly restrict the spreading of M-state might be enzyme limitation. In the model, under the approximation of fast binding-unbinding enzyme kinetics at M-state nucleosomes, it is possible to account for enzyme titration via an effective Δ parameter that depends explicitly on the total number of HMEs N_e^{tot} and on the current fraction of bound HMEs (see Materials and Methods), high fractions being associated with low effective state-dependent recruitment. Limiting the number of HMEs thus reduces the spreading efficiency (Supplementary Fig. 4), however in ranges of parameters that support epigenetic memory, the long-range-spreading activity of bound HMEs still lead to unconfined epigenomic domains (Fig. 7B).

As long-range spreading via the trans-activity of HMEs seems to promote unconfined memory, another possible mechanism to restrict spreading might thus be to modulate the 3D communication between loci [37]. Some epigenomic marks are associated with architectural proteins that may indeed impact chromatin organization. For example, PRC1 and HP1 that can bind H3K27me3- and H3K9me2/3-marked chromatin respectively, are known to promote chromatin compaction in vitro [86, 87] and in vivo [88–90]. To explore how 3D compaction of domains would influence epigenetic memory, we introduce a nucleosome-nucleosome contact probability $P_c(i, j)$ that is consistent with the formation of a compact domain around the painter region (see Materials and Methods) with P_c inside the domain being stronger than outside (Supplementary figure 9). In the limit of high N_e^{tot} (Fig. 7C), such 3D organization clearly facilitates the maintenance of a stable M-state inside the compacted region even in absence of sequence dependent stimulus, but exhibits "flooding" of the M-state outside this region with time due to the residual spreading between nearest-neighbor sites. However, when we couple 3D compaction and enzyme limitation, the initial domain remains very stable with limited flooding even under strong perturbation like replication (Fig. 7D), suggesting that both ingredients can lead to confined epigenetic memory.

3.8 Chromatin state dynamics regulates transcriptional noise

Histone modifications are likely to play a role in the regulation of genome accessibility thereby impacting transcriptional activity of genes [3]. To understand how gene expression may be affected by the dynamics of chromatin states, we consider a simple toy model where the transcriptional state of a gene depends on the current mean M-state proportion $\overline{M} = \frac{1}{n}\sum \delta_{i,M}$ over a 1kbp-wide ($n = 5$) region that would represent the gene promoter. We consider in this section that the state M is a repressive state (e.g. constitutive heterochromatin via H3K9me2/3) and that the transcription rate is inhibited by the presence

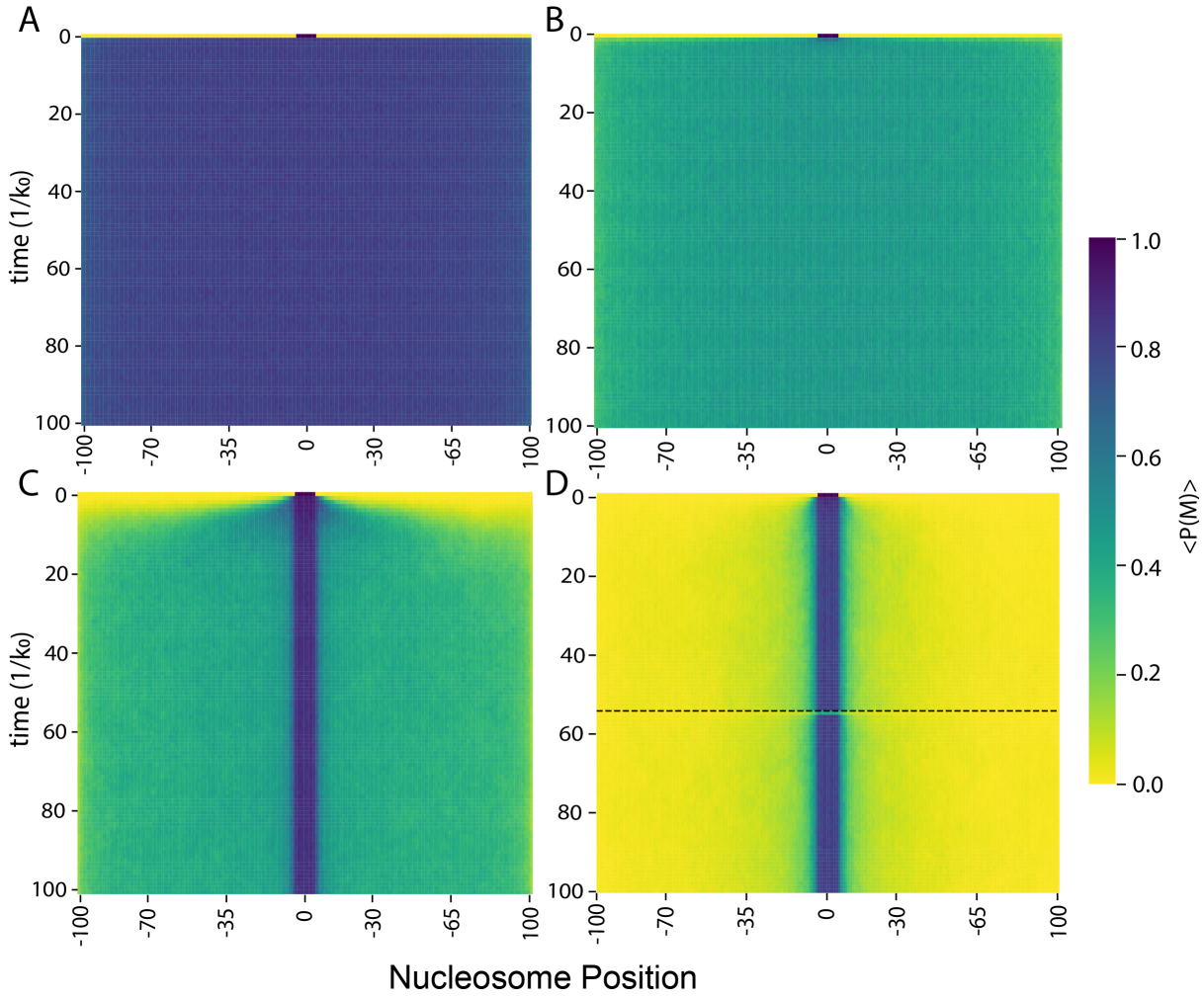

Figure 7. Confined epigenetic memory. (A) Evolution of $P(M)$ in the enzyme non-limiting regime $N^{tot} = 45$. The initial configuration is a domain of modified states around nucleosome position 0 and there is no sequence-specific recruitment for $t > 0$. $k/k_0 = 2, \epsilon = 0.9$. (B) As in (A) but with enzyme limitation, $N_e^{tot} = 9$. (C) As in (A) but with 3D compaction of the domain (Supplementary Fig.9A shows the corresponding contact probability matrix), in large $N_e^{tot} = 45$ limit. (D) As in (A) but with enzyme limitation ($N_e^{tot} = 9$) and 3D compaction. Modified state domain remains confined even under replicative dilution ($T_{cyc} = 55/k_0$, dashed black line).

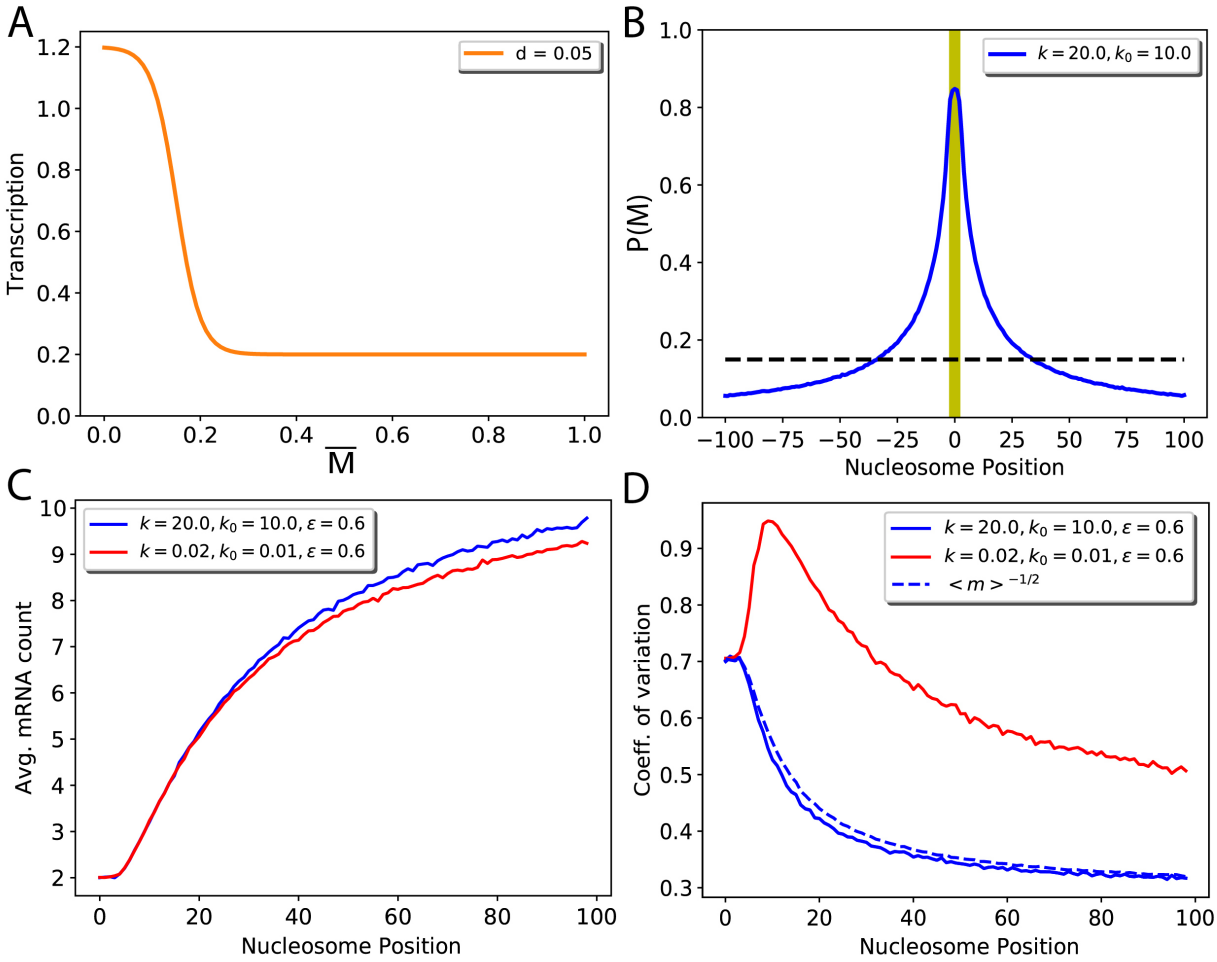

Figure 8. Chromatin state dynamics and transcription: (A) Cooperative switch model: Transcription rate as a function of M state proportion ($\overline{M} = \frac{1}{n}\sum \delta_{i,M}$) with steepness $d = 0.05$. (B) $P(M)$ profile for $k/k_0 = 2, \varepsilon = 0.6$. Black dotted line indicates the value of $P(M)$ at the transcriptional switch ($P(M) = M^*$,see (A)). (C) Average mRNA count for the two different epigenomic dynamics as a function of the distance of the gene from the painter region. (D) The coefficient of variation of the mRNA count for the same parameters as in (C). The dashed line shows $1/\sqrt{\langle mRNA \rangle}$ with $\langle mRNA \rangle$ taken from panel (C).

of modified nucleosomes at the promoter ($\overline{M} \geq \overline{M}^*$) in a switch-like manner [22](Fig. 8A, see Materials and Methods).

To only focus on the epigenetics-transcription relation, we consider a simple painter mode (as in Fig. 2) with $k/k_0 = 2$ and investigate the transcriptional properties of a gene as a function of the proximity of its promoter to the painter region (Fig. 8). We observe that the steady-state average number of mRNA per cell $\langle mRNA \rangle$ for this gene is larger for more distant promoters (Fig. 8C), qualitatively mirroring the decrease in $P(M)$ (Fig. 8B). Interestingly, for the same *relative* k/k_0 value (and thus the same average M-state profile, see Eq. 3, Fig. 8B), the average transcription level depends on the *absolute* values of k and k_0 that control the kinetics of M state dynamics, comparatively to the transcription dynamics that is itself driven by the transcription and mRNA degradation rates (see Materials and Methods). We detect that slow epigenetic dynamics (red line in Fig. 8C) tend to favor a slightly more repressed state compared to fast dynamics (blue line in Fig. 8C).

Such kinetic effect is more striking when considering the intrinsic stochastic fluctuations of gene expression by estimating the coefficient of variation (CV) (Fig. 8D), defined as the ratio between the standard deviation of the corresponding steady-state distribution of mRNA number per cell and its average value, a high CV meaning a noisy, highly fluctuating gene. For fast dynamics (blue line in Fig. 8D), relative fluctuations decrease with the distance from the painter region. CV follows $\sim 1/\sqrt{\langle mRNA \rangle}$ (dashed blue line in Fig. 8D), characteristic of the intrinsic transcriptional noise found in elementary gene expression models [93]: lowly expressed gene being relatively more noisy. For slow dynamics (red line in Fig. 8D), we instead observe a sharp increase up to a distance s^* where fluctuations become maximal, followed by a gradual decrease at larger distances; fluctuations remaining always larger than in the fast dynamics case. Indeed, in our transcription rate model (Fig. 8A), close to the transition point $\overline{M}^*$, small variations of $\overline{M}$ result in large deviations for the transcription rate. If the epigenetic dynamics is faster than mRNA production and degradation rates, the transcription level cannot adjust to the rapid fluctuations of $\overline{M}(t)$ that are filtered by the slow mRNA dynamics [94]. However, if the chromatin dynamics is much slower, mRNA level adapts to the current, slowly fluctuating M-state. Hence expression will stochastically switch between a repressed ($\overline{M}(t) > \overline{M}^*$) and expressed state ($\overline{M}(t) < \overline{M}^*$) leading to large fluctuations of expression. The peak observed in CV in this case thus translates the interplay between such ultra-sensitivity to epigenetic fluctuations with the more standard dependence of the noise to $\langle mRNA \rangle$ (fast chromatin dynamics case). Biologically, this strong propagation of epigenetic fluctations to the gene expression level may be related to the well known variegation phenomenon where genes inserted near heterochromatin domains may have variable transcriptional pattern, depending on the distance of their insertion sites to the heterochromatin domains [95].

3.9 Applications to diverse biological contexts

Above we studied systematically the behavior of our generic model of epigenomic regulation and its consequences on epigenetic memory and transcription. In the next, we will describe three applications of this framework to specific biological situations.

3.9.1 Phosphorylation of $H2AX$ around double-stranded breaks.

During our analysis of the simple painter mode (Fig. 2), we showed that there is a direct relationship between the distribution of marks around the painter binding sites and the local 3D chromatin organization. This 3D↔1D relationship has been evidenced experimentally in the context of DNA double-stranded breaks (DSB) in humans [70]. After a break had occured, variant histones H2AX present in the chromatin are phosphorylated (the so-called γH2AX mark) by the ATM kinase recruited at the DSB site. The correlative analysis of the 4C and γH2AX Chip-seq signals around DSBs clearly reveal high similarities (e.g. see Fig. 1A in [70]), suggesting that the spreading by painter mode is at work in that system. In Fig. 9A, we extracted the average γH2AX profile observed around DSBs (red dots) from [70] that extends over Mbps. Using the analytical formula for a simple painter model (Eq. 3) with a painter region localized at the DSB and $P_c(s) \sim 1/s$ the typical average contact probability found in human cells [61], we were able to fit very well the experimental data (blue line in Fig. 9A, L2-distance $= 1.4 * 10^{-4}$)(see Materials and

Methods). This leads to one identifiable parameter $(k/k_0)\varepsilon = 1213$. In the original article, Arnould *et al.* suggested that the loop extrusion mechanism might be directly implicated in the spreading mechanism [70]. To test this alternative hypothesis, we perform a similar inference using the loop extrusion-like spreading probability ($P_c(s) \sim \exp[-s/s_0]$) which leads to a less precise fit even if the model has an additional parameter (processivity $s_0 \sim 150 - 200$kbp) (green line in Fig. 9A, L2-distance= $5.1 * 10^{-4}$). Note that the diffusion-like spreading mechanism ($P_c(s) \sim 1/s^{0.5}$) is also inconsistent with the experimental data (orange line in Fig. 9A, L2-distance= $9.0 * 10^{-4}$). These results differ from a similar analysis performed by Li et al. [91] suggesting, using a two-state mathematical model neglecting turnover rate of the mark, that Tel1, the yeast homolog of ATM, spreads phosphorylation around DSB via a loop extrusion-like mechanism while within our framework the best prediction (minimum L2-distance) is given by the spreading-by-3D-contact mechanism ($P_c(i,j) = 1/|j-i|$)).

3.9.2 Heterochromatin formation around retrotransposons.

An entirely different system where the simple painter mode may also be operative is the spreading of H3k9me2/3 modifications over few kbps around retrotransposons in mouse ES cells by Setdb1 [74, 96], a lysine methyltransferase recruited by the KRAB–Zinc Finger Protein KAP1 [97]. As in the previous example, we can fit the experimental Chip-seq profiles (red dots in Fig. 9B) by a simple model with a painter zone corresponding to the typical size range of a transposable element $\sim 1kbp$ [75] (blue line in Fig. 9B). In this case, we found $(k/k_0)\varepsilon = 0.78$. Note that these predictions were done without invoking the reader-writer mechanism which might be relevant in other heterochromatin contexts (see below).

Compared to γH2AX, the spreading of H3K9me2/3 is much less efficient leading to smaller domains (kbp- vs Mbp-wide). This may translate fundamental differences in the spreading and turnover rates as histone methylation maintenance in mammalian cell lines could be slow ($\sim$ 10-20 hours) [29, 59] while establishment of γH2AX foci is much faster ($\sim$ 1-30 min) [98, 99].

3.9.3 Memory of heterochromatin in fission yeast.

Finally, we apply our chromatin state dynamical model to quantitatively characterize the stability of epigenetic memory for fission yeast heterochromatin. In *Schizosaccharomyces pombe*, establishment of H3K9me2/3 domains is in part regulated by the balance between the methyltransferase activity of Clr4, its state-dependent recruitment by H3K9me histones and the demethylase-like action of Epe1 [10, 100]. In [10], a heterochromatin domain is established via the sequence-specific recruitment of Clr4 at an ectopic locus containing a fluorescent reporter gene, leading to the repression of the gene. Epigenetic memory of this domain and involved mechanisms are then characterized by releasing sequence tethering and by tracking the progressive re-activation of the gene as a function of time (full lines in Fig. 9C). In particular, they considered four different strains: (i) $TetR - Clr4 - I, epe1+$ with a mutated Clr4 ($Clr4 - I$) with no reader-writer property and a normal demethylase-like activity ($epe1+$), the model analogue being $\Delta = 0, k_0 = x$; (ii) $TetR - Clr4 - I, epe1\Delta$ same as (i) but without demethylase-like activity ($epe1\Delta$) (model equivalent - $\Delta = 0, k_0 < x$); (iii) $TetR - Clr4 - I, Clr4+, epe1+$ with wt-like Clr4 and Epe1 activities ($\Delta > 0, k_0 = y$); and (iv) $TetR - Clr4 - I, Clr4+, epe1\Delta$ as in (iii) but in $epe1\Delta$ background ($\Delta > 0, k_0 < y$). They observed that memory of a repressed state was enhanced by the reader-writer module of Clr4 and the absence of Epe1 (dots in Fig. 9D).

To rationalize these experiments with our quantitative framework, we simulate the experimental memory assay like in Fig. 6 but for various values of Δ and k_0 and by tracking the transcriptional activity of the region using the model described in Fig. 8 (see Materials and Methods). More precisely, we follow the experimental protocol and monitor the time-evolution of the distribution of fluorescence of the reporter gene within the cell population (Fig. 9C), from which we can define the proportion of cells that are still repressed ("OFF") by heterochromatin (see Materials and Methods) (Fig. 9D). The predicted time-evolution of this proportion for each scanned (Δ,k_0) value is then quantitatively compared to the experimental data obtained for each of the four strains (Supplementary Fig. 6, see Materials and Methods). For strain (i), gene activation is fast (Supplementary Fig. 15, black dots in Fig. 9D) and data are compatible with a simple painter mode with high turnover rate ($k_0 > 0.007\ min^{-1}$, $\Delta = 0$)

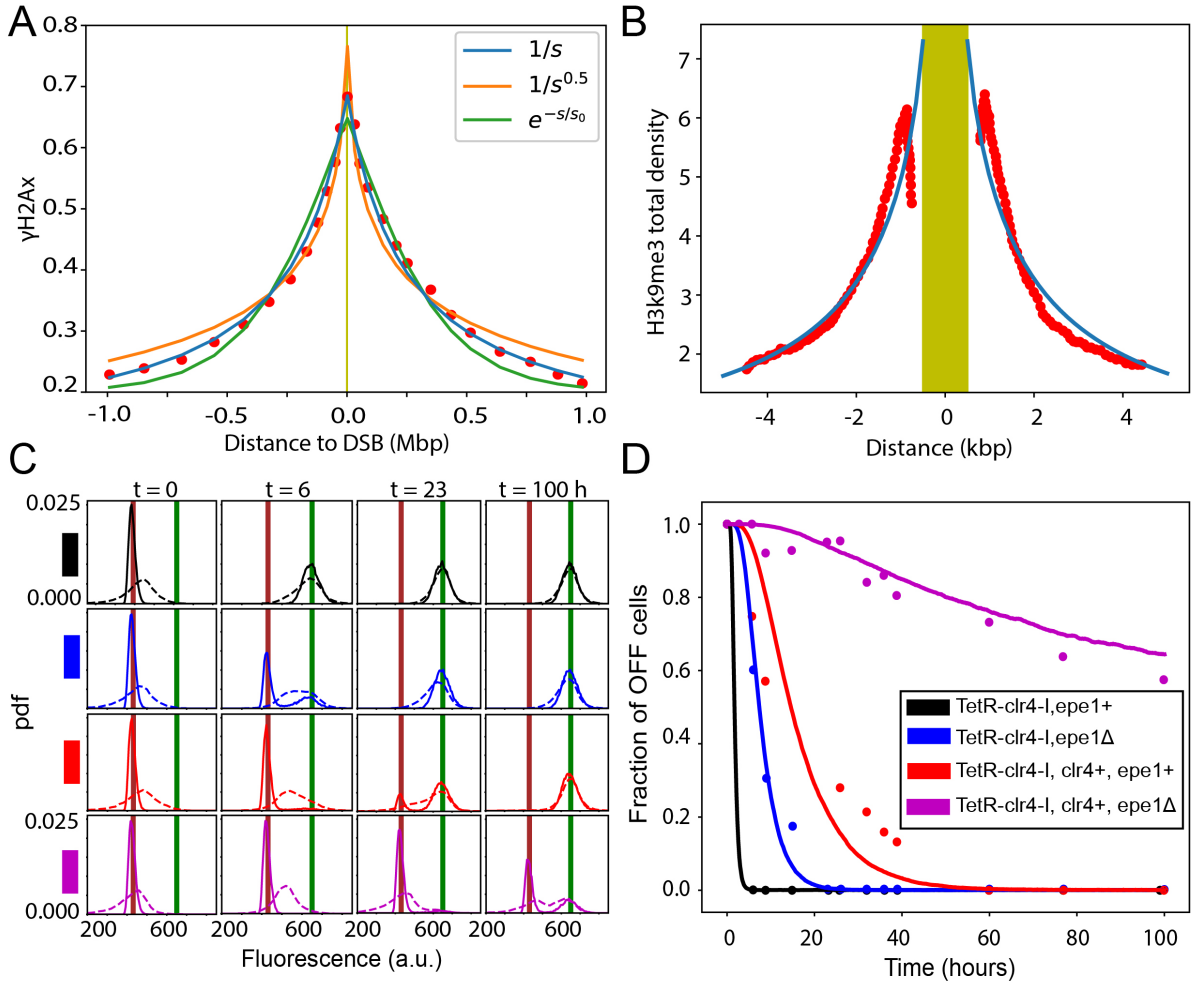

Figure 9. (A) Experimental phosphorylation profile (dots) around a DNA double strand break [70] and the corresponding model prediction (i) $S(i) = 0.58P(M_i) + 0.11$, $(k/k_0)\varepsilon = 1213$ $P_c(i,j) = 1/|i-j|$ (blue) (ii) $S(i) = 0.85P(M_i)$, $(k/k_0)\varepsilon = 29.6$, $P_c(i,j) = 1/|i-j|^{0.5}$ (orange) (iii) $S(i) = 0.9P(M_i) + 0.2$, $(k/k_0)\varepsilon = 1$, $P_c(i,j) = e^{-|i-j|/s_0}$, $s_0 = 1050$ (green). (B) Experimental H3K9me3 density (dots) flanking a transposable element in mouse ES cells [74] and the corresponding model prediction $S(i) = 10.75P(M_i) + 0.1$ (full line). (C) Heterochromatin maintenance in fission yeast. Distributions of fluorescence expressed from a reporter gene at specific time points after painter release at $t = 0$ for different strains (experiments: dashed lines [10], model predictions: full lines). Brown and green vertical lines mark the mean fluorescences in the OFF and ON states respectively. (D) Experimentally-measured (dots) and predicted (full lines) time-evolution of the proportion of OFF cells in the population for the various strains.

(Supplementary Fig. 6A, and black line in Fig. 9D), consistent with the presence of Epe1 and the absence of reader-writer recruitment which, we demonstrated in Fig. 6, is essential for epigenetic memory. For strain (ii), the gene de-repression is slower than in (i) (blue dots in Fig. 9D) but can still be well captured by a simple painter mode ($\Delta = 0$) (blue line in Fig. 9D) associated with decrease of k_0 by at least 6 fold (Supplementary Fig. 6B), consistent with absence of Epe1 in this strain and thus with less turnover. In both backgrounds (iii)(*epe*1+) and (iv)(*epe*1Δ), the reader-writer capacity of Clr4 strongly enhances memory and slows down gene activation (red and purple dots in Fig. 9D). This behavior can only be reproduced (red and purple lines in Fig. 9D) by a reader-writer mode with $\Delta = 0.004$ (Supplementary Fig. 6C,D). Interestingly, for strain (iv) where memory is maximal, the model predicts that the biological system is in a parameter range just below the critical regime of the reader-writer mode (like the blue line in Fig. 6A). In this regime, the epigenetic state is long-lastly maintained until rapid removal of the marks ($\Delta = 0.2$ in Fig. 6C), leading to a stochastic switch in transcriptional activity and bimodal distributions for protein levels (Fig. 9C, purple at t = 100h), as observed experimentally [10].

4 DISCUSSION AND CONCLUSION

In this work, we developed a unified mathematical two-state model of epigenomic regulation, integrating key mechanisms like reader-writer processes. Our simple framework states that the establishment and maintenance of an epigenetic state result from the recruitment and spreading activity of histone modifying enzymes (HMEs). Recruitment of HMEs can be mediated by specific genomic sequences or by the local epigenomic state. Spreading encompasses an on-site action of the HMEs and a long-range, trans activity modulated by the local chromatin state.

In particular, we systematically studied three generic modes of regulation that recapitulate most of the experimentally-known epigenetic systems: (i) a simple painter mode (Fig. 2) in which HMEs are targeted to specific genomic regions and spread via their trans-activity around these binding sites; this mode may be representative of the regulation of small, local epigenomic domains like acetylation marks (e.g. H3K27ac or H3K9ac) around promoters and enhancers; (ii) a boosted-painter mode (Fig. 3) in which recruited HMEs have an enhanced activity if bound to specifically-modified regions; this mode may account for the allosteric boost observed for PRC2 in presence of H3K27me3 [17]; (iii) a reader-writer mode (Fig. 4) in which HMEs may also be recruited by specific epigenetic signals; this mode may capture the regulation of extended chromatin domains like heterochromatic regions (e.g. H3K9me2/3).

In the simple and boosted-painter modes, we found that there is a direct, simple relationship between the binding profiles of HMEs, the chromatin organization and the profile of epigenomic state (Eq. 3). Verification of this relation using experimental measurements of these three information (e.g. bulk Chip-seq and Hi-C data) in wild-type-like conditions may be a strong evidence for one of these two spreading modes [29]. A perfect illustration of this is our application to γH2AX around DSBs in which the experimental profile is well fitted by the model (Fig. 9A) suggesting a simple or boosted-painter mode with a spreading-by-3D-contact scenario rather than a more complex cohesin-mediated mechanism [70].

Distinguishing the boosted- from the simple-painter mode would then require additional information like for example to estimate the spatial correlations between chromatin state that have very different signatures between the two modes (Fig. 2C *vs* 3C). However, this information is currently very difficult to access experimentally as it would require a single-cell assay to estimate covariations of chromatin states; but recent progresses in single-cell Chip-seq experiments [105, 106] may open new venues. Another promising application of Eq. 3 in the simple/boosted-painter modes is that it can be employed to address the inverse problem of inferring the HME binding sites knowing the 3D chromosome organization (e.g. via Hi-C) and the Chip-seq profile of epigenetic states, or of estimating the chromatin folding properties from HME and histone mark Chip-seq profiles. Actually, the latter strategy was recently used by Redolfi *et al.* [82], in the so-called Dam-C technique, to extract 4C-like contact probability information based on the spreading of Dam-mediated DNA methylation from a painter region.

In the booster-painter and reader-writer modes that both involve a "reader" capacity of HME, we observed phase-transitions and critical behaviors via respectively the strengths of boosting activity (r) and of state-dependent recruitment (Δ) (Figs. 3B, 4B). These behaviors are driven by the effective positive

feedback loops and cooperative effects [18] that emerge from the enhanced spreading efficiency of some HMEs (more activity or more recruitment) if the histone modification they catalyze is already present locally. These effects are more important in the reader-writer mode and lead to qualitatively different epigenomic profiles, with extended domains around the binding peaks of HME (Fig. 4A), and distal spatial correlations (Fig. 4C). In particular, we found that these profiles are very sensitive to the underlying spreading capacity of HMEs [101], longer-range mechanisms facilitating the percolation of the system into large contact domains at low state-dependent recruitment (Fig. 5C) [27]. Eventually, these domains may be maintained in the absence of genomic bookmarking (Fig. 6A) for strong enough state-dependent recruitment. However, our analysis of heterochromatin memory in fission yeast (Fig. 9C,D) suggests that such self-sustainable memory may not be occurring for H3K9me2/3 chromatin domains in wild-type conditions and that (weak) nucleation, sequence-dependent signaling might still be needed for maintaining a stable epigenetic landscape [37, 39]; even if a reduction in the effective histone turnover rate (e.g. by a lower demethylase activity as in the *epe*1Δ mutants in [10] or by longer cell-cycles [22, 30]) might trigger the system near to the critical point and may lead to long-term memory [27], but potentially also to high sensitivity to external cues [32]. As shown by previous modeling works [24, 27], we also confirmed that purely local spreading mechanisms (as the NN scenario) would require too strong feedback and state-dependent recruitment to stabilize extended, coherent epigenomic domains during a long time period, even if being able to maintain them over few cell cycles [103].

One theoretical issue however with standard reader-writer models with long-range spreading is the difficulty to stabilize finite-size epigenomic domain without genomic bookmarking and to avoid unconfined spreading at long-time scale, as already pointed out by Erdel and Greene [24] and Dodd and Sneppen [19]. While strong insulators or barriers preventing local spreading [19, 37] or/and the formation of compact 3D domains [37] may slow down but not prevent uncontrollable spreading, we showed here that the combination of 3D compartmentalization with enzyme titration may lead to ultra-stable confined memory of small domains even under strong perturbations like replication (Fig. 7D), as also previously observed by Sandholtz *et al.* using an explicit polymer model of epigenetic regulation with controlled HMEs concentrations [102]. This may be particularly relevant for the Polycomb regulatory system as it involves HMEs (PRC2) in limited numbers, as measured recently during fly embryogenesis [85] and evidenced also in mammals [107], and is associated with PRC1, a protein complex promoting the compaction of H3K27me3-tagged regions [31, 86]. Similarly, maintenance of confined silenced domains in yeast may rely on the combination of their spatial compartmentalisation via Sir3-mediated self-attraction [108] and Sir4-driven tethering to the nuclear envelop [109]) and of the titration of the Sir2 deacetylase [110, 111]. Note that, in our model, 3D compartmentalization was accounted via an attractive interaction between M states, mimicking PRC1, HP1 or Sir3 modus operandi. However, additional architectural mechanisms such as the spatial co-association of boundary elements are likely to be at work: in mammals, the stability of Polycomb domains may be reinforced by contacts between CTCF insulator sites [112] possibly via the cohesin-SA2-mediated loop-extrusion process [113]; in yeasts, colocalization of TFIIIC-binding boundary elements participate to chromatin domain integrity [43, 48]. Additionally, the presence of antagonistic chromatin states domains may act as competitive barriers to epigenomic spreading [6] but also may favor 3D compartmentalisation by strengthening phase separation [114]. Actually, all kind of processes that reinforce *intra-* vs *inter-* contact domains is expected to promote confined memory by limiting long-range "contamination": in a condition of limited number of HMEs, increasing compartmentalization will further increase their local concentration (by sequestration), leading to a stronger stability and weaker pervasive long-range spreading (Fig. 7D), in agreement with recent experimental observations in fission yeast suggesting that a "critical density" of H3K9me3 is required for the stable inheritance of confined heterochromatin [104].

To go beyond the mere description of epigenomic regulation, we also proposed a simple mathematical model describing the impact of chromatin state dynamics on gene expression [22]. Using this generic and modular framework, we quantified the role of epigenetic fluctuations in transcriptional noise. It allowed us to investigate positional variegation as well as epigenomic memory in terms of gene expression in addition to histone marks. In particular, this makes possible to analyze and interpret the transcriptional outputs of epigenetic-related experiments. For example, using this framework, we

quantitatively described heterochromatin assays by fitting gene reporter expression distribution versus time and strains. Introducing a feedback of transcription on the epigenomic dynamics via the recruitment of HMEs [26, 28] or changes in turnover rates [115] may allow a finer description of the interplay between epigenomics and transcription.

Previous mathematical models have also addressed various aspects of the establishment and maintenance of epigenomic information. Several works [18, 20, 26, 32, 118] have focused on bivalent regions, regulated by reader-writer and long-range spreading mechanisms, that can alternate between different bistable, coherent states (active or inactive). While our framework can be easily generalized to account for such multistability, the scope of our study was to specifically investigate the interplay between reader-writer processes, spreading capacity and nucleation mechanisms. In that context, our approach generalizes or complements previous two-state descriptions of epigenome regulation, like the 3D looping models of Erdel and Greene [24] and Katava et al. [39] and the stochastic model by Ancona et al. [27]. 3D looping models with sequence-dependent recruitment and a long-range spreading mechanism based on 3D contacts are by essence very similar to our simple painter and reader-writer modes. In [24], authors discussed the role of looping (vs NN scenario) in epigenomic profiles and of reader-writer feedback in memory and proposed that non-homogeneous state-specific recruitment may lead to stable confined domains. In [39], they investigated the interplay between spreading dynamics and chromatin motion, highlighting the role of nucleation in the maintenance of stable domains in a regime of fast epigenetic spreading. In their stochastic model, Ancona et al. [27] studied a reader-writer-mode like model in absence of nucleation and showed that local mark turnover coupled to long-range spreading mechanisms may lead to percolation transition and bistability.

In our work, we focused on a single painter site, but to go further the painter model may provide a very interesting framework to specifically address the pivotal role of the recruitment painter sites in globally shaping the epigenome. In particular to investigate how their sizes, strengths or genomic distributions influence epigenomic regulation [101]. Interestingly, within some epigenomic domains, some sequences can recruit HMEs while being unable to nucleate chromatin states on their own [7]. These elements may actually act as secondary recruitment sites after the initial nucleation events at the strong, autonomous sites [116] and, in the presence of numerous weak sites, a stable domain can be formed even in the absence of strong sites [117]. Based on our results, an appealing hypothesis would be that, the numerous dispersed sites along the genomic domain might compensate their "individual" weakness by maintaining "collectively and cooperatively" a high local concentration of HMEs via long-range 3D "communication".

One strong hypothesis that we made in order to focus on the painter and reader-writer processes and their consequences on epigenetic memory, was to neglect the interplay between several chromatin modifications [118, 119]. Indeed, the regulation of one silencing epigenomic mark often results from the competition with antagonistic activation marks. Such tug-of-war is driven by reader-writer processes complemented with reader-eraser mechanisms that actively remove opposite marks [118, 120]. This may lead to bistability between active and inactive chromatin states and participate in reinforcing epigenetic memory [18, 26, 121]. As for writers, recruitment of eraser enzymes may be sequence- or state-specific. For example, the UTX/JMJD3 demethylase is recruited by Trithorax-group proteins at active developmental genes to remove the repressive Polycomb H3K27me3 mark [14, 122]. In addition, the regulation of histone turnover may also participate in the erasing mechanisms. Indeed, genomic regions tagged with inactive modifications like H3K9me2/3 are know to display low histone turnover rate, while transcribed regions tends to exhibit higher rates [54, 115, 123]. All this may be responsible of complementary feedback loops in the epigenomic regulatory dynamics that finely tune the different levels of epigenomic modifications. For example, transcription may mediate such a loop not only by depending on the current epigenomic state, as we tentatively modeled in Fig. 8, but also, in return, by impacting epigenomic stability via an increase in histone turnover rate [26, 28].

Another simplification made in our work was to assume that the 3D chromatin organization (that mediates the long-range spreading) was fast compared to epigenomic-associated rates and was independent of the current chromatin state. While the former hypothesis may be satisfied for short genomic distances, for large, Mbp-wide, chromatin domains like pericentromeric, H3K9me2/3-tagged regions in higher eucaryotes,

the dynamics of contact may be slow [58,124] leading to less efficient spreading [37,60]. As discussed above, heterochromatin domains (e.g. H3K9me3, H3K27me3) are often associated with architectural proteins (e.g. HP1, PRC1) that may impact on their compaction. Therefore, it introduces again a dynamical feedback loop between long-range epigenomic spreading and 3D chromatin organization: nucleation of a chromatin state drives its compaction that in turn facilitate spreading and maintenance [35,87]. Recently, we and other groups investigated more carefully this coupling by developing a model that explicitly account for both 3D and 1D dynamics of chromatin [37–39, 102], highlighting the key role played by genome folding on epigenomic regulation. Combining the formalism developed here with such more detailed framework would allow to better characterize the structure-function relationship of chromatin including the formation and maintenance of confined chromatin domain.

To conclude the formalism developed here is generic and modular, as it provides a simple description of epigenomic regulation in terms of HMEs recruitment and enzymatic activity. It can be easily upgraded to include more chromatin modifications and cross-talks and feedback loops and can thus be contextualized to a variety of specific chromatin regulatory systems [29].

A very promising application of our painter model might be to explore how epigenomic (and corresponding transcriptional) dynamics might influence genome evolutionary processes. Of particular interest is the challenging question of the role of epigenome in the control of the integration and selection of transposable elements (TEs). TEs are mobile genomic elements that can be inserted at new genomic positions or can modify their location either via a "copy and paste" or in a "cut and paste" mechanism [75]. They comprise as much as $40 - 60\%$ of mammalian genomes. For active retrotransposons, a specific class of TEs, host cells have developed defense strategies in order to limit their invasion and expansion and to maintain genome integrity. Most of these elements are targeted by epigenomic silencing mechanisms (e.g. H3K9me2/3) [125–127] that limit their expression and thus, in *fine*, restrict their capacity of "parasitic" transposition [128]. This silencing is usually achieved by site specific recruitment of HMEs such as Setdb1 and Su(var)3-9s that nucleate and further spread the H3K9me2/3 marks [125,129,130] within TE elements and beyond into the flanking region, as shown in Fig. 9B with the simple painter mode in mESC. Very similar spreading patterns have also been observed around several other TE families in two Drosophila species, *D. melanogaster* and *D. simulans* [131]. As expected and confirmed by our theoretical approach that couples epigenomic and transcriptional dynamics (Fig. 8C), such long-range spreading of repressive chromatin state from TE elements has been shown to mediate long-range silencing of flanking genes [132]. This is actually reminiscent of the well known phenomenon of position effect variegation (PEV) observed in various organisms, where the incidental repositioning of a normally expressed/repressed genes, by translocation or transposition, next to a hetero/euchromatin fuzzy domain boundary induces its stochastic repression/expression depending on the genomic distance between the insertion and the boundary [95, 133, 134]. On one hand, there is a generic selection against TE transposition into euchromatin due to their deleterious effects for host cells that might be not only related to potential genetic alterations but also to epigenomic alteration via PEV. As shown in [131], the species-specific spreading ability of TEs is likely to be a driving force of their counter-selection. On the other hand, TEs are main contributors of genome evolution, by promoting genomic diversity and allowing regulatory innovations. Co-opted TEs play a crucial regulatory role in various nuclear processes, in particular in gene regulation during development [135, 136] as well as in 3D genome organisation [137]. TEs are constitutively silenced in most cell types and this silencing is required not only to limit their transposition but also to maintain proper cell-type 3D organisation and gene expression pattern. Hence, deregulation of TE has been associated to pathologies and is a clear hallmark of cancer [138, 139]. Overall this clearly indicates the need to develop a quantitative model based on the painter framework that would describe the epigenomic control of TEs and how they affect in turn the epigenome dynamics of flanking genomic regions. This will pave the way for a better understanding of the role of epigenome dynamics in TE-based genome plasticity during evolution as well as in pathologies.

5 ACKNOWLEDGEMENTS

We are grateful to Geneviève Fourel, Kapil Newar, Gaëlle Legube and Raphaël Mourad for fruitful discussions. We acknowledge Agence Nationale de la Recherche (DJ: ANR-18-CE12-0006-03; AZB, DJ: ANR-18-CE45-0022-01; DJ, CV: ANR-21-CE45-0011-01) for funding. We thank PSMN (Pôle Scientifique de Modélisation Numérique) and CBP (Centre Blaise Pascal) of the ENS de Lyon for computing resources.

Conflict of interest statement.

None declared.

References

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

2. Gaydos, L. J., Wang, W. and Strome, S. (2014) Gene repression H3K27me and PRC2 transmit a memory of repression across generations and during development. *Science* **345**, 1515–1518.

3. Allis, C., Jenuwein, T. and Reinberg, D. (2007) *Epigenetics.* Cold Spring Harbor Laboratory Press, NY.

4. Beisel, C. and Paro, R. (2011) Silencing chromatin: comparing modes and mechanisms. *Nat Rev Genet* **12**, 123–135.

5. Zhang, T., Cooper, S. and Brockdorff, N. (2015) The interplay of histone modifications - writers that read. *EMBO Rep* **16**, 1467–1481.

6. Fourel, G., Magdinier, F. and Gilson, E. (2004) Insulator dynamics and the setting of chromatin domains. *Bioessays* **26**, 523–532.

7. Fourel, G., Lebrun, E. and Gilson, E. (2002) Protosilencers as building blocks for heterochromatin. *Bioessays* **24**, 828–835.

8. Reyes-Turcu, F. E. and Grewal, S. I. (2012) Different means, same end-heterochromatin formation by RNAi and RNAi-independent RNA processing factors in fission yeast. *Curr Opin Genet Dev* **22**, 156–163.

9. Hathaway, N. A., Bell, O., Hodges, C., Miller, E. L., Neel, D. S. and Crabtree, G. R. (2012) Dynamics and memory of heterochromatin in living cells. *Cell* **149**, 1447–1460.

10. Ragunathan, K., Jih, G. and Moazed, D. (2015) Epigenetics. epigenetic inheritance uncoupled from sequence-specific recruitment. *Science* **348**, 1258699.

11. Angel, A., Song, J., Dean, C. and Howard, M. (2011) A polycomb-based switch underlying quantitative epigenetic memory. *Nature* **476**, 105–108.

12. Kassis, J. A. and Brown, J. L. (2013) Polycomb group response elements in Drosophila and vertebrates. *Adv Genet* **81**, 83-118.

13. Dorafshan, E., Kahn, T. G. and Schwartz, Y. B. (2017) Hierarchical recruitment of Polycomb complexes revisited. *Nucleus* **8**, 496-505.

14. Wang, S. P., Tang, Z., Chen, C. W., Shimada, M., Koche, R. P., Wang, L. H. *et al.* (2017) A UTX-MLL4-p300 transcriptional regulatory network coordinately shapes active enhancer landscapes for eliciting transcription. *Mol Cell* **67**, 308–21.e6.

15. Schuettengrube, B., Bourbon, H., Di Croce, L. and Cavalli, G. (2017) Genome regulation by Polycomb and Trithorax: 70 years and counting. *Cell* **171**, 34-57.

16. Probst, A. V., Dunleavy, E. and Almouzni, G. (2009) Epigenetic inheritance during the cell cycle. *Nat Rev Mol Cell Biol***10**, 192-206.

17. Margueron, R. *et al.* (2009) Role of the polycomb protein Eed in the propagation of repressive histone marks. *Nature* **461**, 762-767.

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

19. Dodd, I. B. and Sneppen, K. (2011) Barriers and silencers: A theoretical toolkit for control and containment of nucleosome-based epigenetic states. *J Mol Biol* **414**, 624-637.

20. Dayarian, A. and Sengupta, A. (2013) Titration and hysteresis in epigenetic chromatin silencing. *Phys Biol* **10**, 036005.

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

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

23. Erdel, F., Muller-Ott, K. and Rippe, K. (2013) Establishing epigenetic domains via chromatin-bound histone modifiers. *Ann N Y Acad Sci* **1305**, 29–43.

24. Erdel, F. and Greene, E.C. (2016) Generalized nucleation and looping model for epigenetic memory of histone modifications. *Proc Natl Acad Sci U S A***113**, E4180–E4189.

25. Cortini, R., Barbi, M., Caré, B., Lavelle, C., Lesne, A., Mozziconacci, J. and Victor, J.-M. (2016) The physics of epigenetics. *Rev Mod Phys* **88**, 025002.

26. Berry, S., Dean, C. and Howard, M. (2017) Slow chromatin dynamics allow polycomb target genes to filter fluctuations in transcription factor activity. *Cell Systems* **4**, 1–13.

27. Ancona, M., Michieletto, D. and Marenduzzo, D. (2020) Competition between local erasure and long-range spreading of a single biochemical mark leads to epigenetic bistability. *Phys Rev E* **101**, 042408.

28. Holoch, D., Wassef, M., Lövkvist, C. *et al.* (2021) A cis-acting mechanism mediates transcriptional memory at Polycomb target genes in mammals. *Nat Genet* **53**, 1686–1697.

29. Newar, K., Fanchon, E. and Jost, D. (2021) Dynamical modeling of the H3K27 epigenetic landscape in mouse embryonic stem cells. *bioRxiv* 2021.09.30.462529.

30. Reinig, J., Ruge, F., Howard, M. *et al.* (2020) A theoretical model of Polycomb/Trithorax action unites stable epigenetic memory and dynamic regulation. *Nat Commun* **11**, 4782.

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

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

33. Ho, J. W. K., Jung, Y. L., Liu, T., Alver, B. H., Lee, S., Ikegami, K., Sohn, K.-A., Minoda, A., Tolstorukov, M. Y., Appert, A., Parker, S. C. J., Gu, T., Kundaje, A., Riddle, N. C., Bishop, E., Egelhofer, T. A., Hu, S. S., Alekseyenko, A. A., Rechtsteiner, A., Asker, D., Belsky, J. A., Bowman, S. K., Chen, Q. B., Chen, R. A.-J., Day, D. S., Dong, Y., Dose, A. C., Duan, X., Epstein, C. B., Ercan, S., Feingold, E. A., Ferrari, F., Garrigues, J. M., Gehlenborg, N., Good, P. J., Haseley, P., He, D., Herrmann, M., Hoffman, M. M., Jeffers, T. E., Kharchenko, P. V., Kolasinska-Zwierz, P., Kotwaliwale, C. V., Kumar, N., Langley, S. A., Larschan, E. N., Latorre, I., Libbrecht, M. W., Lin, X., Park, R., Pazin, M. J., Pham, H. N., Plachetka, A., Qin, B., Schwartz, Y. B., Shoresh, N., Stempor, P., Vielle, A., Wang, C., Whittle, C. M., Xue, H., Kingston, R. E., Kim, J. H., Bernstein, B. E., Dernburg, A. F., Pirrotta, V., Kuroda, M. I., Noble, W. S., Tullius, T. D., Kellis, M., MacAlpine, D. M., Strome, S., Elgin, S. C. R., Liu, X. S., Lieb, J. D., Ahringer, J., Karpen, G. H. and Park, P. J. (2014) Comparative analysis of metazoan chromatin organization. *Nature* **512**, 449–52.

34. Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S. and Aiden, E. L. (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–80.

35. Haddad, N., Jost, D., and Vaillant, C. (2017) Perspectives: using polymer modeling to understand the formation and function of nuclear compartments. *Chromosome Research* **25**, 35.

36. Haddad, N., Vaillant, C. and Jost, D. (2017) Ic-finder: inferring robustly the hierarchical organization of chromatin folding. *Nucleic Acids Res* **45**, e81.

37. Jost, D. and Vaillant, C. (2018) Epigenomics in 3D: importance of long-range spreading and specific interactions in epigenomic maintenance. *Nucleic Acids res* **46**, 2252–2264.

38. Michieletto, D., Orlandini, E. and Marenduzzo, D. (2016) Polymer model with epigenetic recoloring reveals a pathway for the de novo establishment and 3D organization of chromatin domains. *Phys Rev X* **6**, 041047.

39. Katava, M., Shi, G. and Thirumalai, D. (2021) Chromatin dynamics controls epigenetic domain formation. *bioRxiv* 2021.01.18.427115v2.

40. 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.* (2010) Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell* **143**, 212–224.

41. Roudier, F., Ahmed, I., Bérard, C., Sarazin, A., Mary-Huard, T., Cortijo, S., Bouyer, D., Caillieux, E., Duvernois-Berther, E., Al-Shikhley, L. *et al.* (2011) Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *EMBO J.* **30**, 1928–1938.

42. Kueng, S., Oppikofer, M. and Gasser, S. M. (2013) SIR Proteins and the Assembly of Silent Chromatin in Budding Yeast *Annu Rev Genet* **47**, 293-324.

43. Thurtle, D. M. and Rine, J. (2014) The molecular topography of silenced chromatin in *Saccharomyces cerevisiae* *Genes Dev* **28**, 245-258.

44. De, S. *et al.* (2020) Defining the Boundaries of Polycomb Domains in Drosophila *Genetics* **216**, 689-700.

45. Kirkland, J. G., Raab, J. R. and Kamakaka, R.T. (2013) TFIIIC bound DNA elements in nuclear organization and insulation *Biochim Biophys Acta* **1829**, 418-424.

46. Bonev, B., Mendelson, C. N., Szabo, Q., Fritsch, L., Papadopoulos, G. L., Lubling, Y., Xu, X., Lv, X., Hugnot, J. P., Tanay, A. and Cavalli, G. (2017) Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* **171**, 557-572.e24.

47. Hou, C., Li, L., Qin, Z. S. and Corces, V. G. (2012) Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains. *Mol Cell* **48**, 471-84.

48. Charlton, S. J. *et al.* (2020) Integrity of a heterochromatic domain ensured by its boundary elements *Proc Natl Acad Sci U S A* **117**, 21504-511.

49. Donze, D., Kamakaka, R. T. (2001) RNA polymerase III and RNA polymerase II promoter complexes are heterochromatin barriers in Saccharomyces cerevisiae. *EMBO J.* **20**, 520–31.

50. Michieletto, D., Chiang, M., Colì, D., Papantonis, A., Orlandini, E., Cook, P. R. and Marenduzzo, D. (2018) Shaping epigenetic memory via genomic bookmarking *Nucleic Acids res* **46**, 83–93.

51. Cheng, T. H. and Gartenberg, M. R. (2000) Yeast heterochromatin is a dynamic structure that requires silencers continuously. *Genes Dev* **14**, 452-463.

52. Laprell, F., Finkl, K. and Muller, J. (2017) Propagation of Polycomb-repressed chromatin requires sequence-specific recruitment to DNA. *Science* **356**, 85–88.

53. Coleman, R. and Struhl, G. (2017) Causal role for inheritance of H3K27me3 in maintaining the OFF state of a *Drosophila* HOX gene. *Science* **356**, eaai8236.

54. Deal, R. B., Henikoff, J. G. and Henikoff, S. (2010) Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science* **328**, 1161-1164.

55. Jenuwein, T. and Allis, C.D. (2001) Translating the histone code. *Science* **293**, 1074-1080.

56. Margueron, R. and Reinberg, D. (2010) Chromatin structure and the inheritance of epigenetic information. *Nat Rev Genet* **11**, 285-296.

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

58. Ghosh, S. K. and Jost, D. (2018) How epigenome drives chromatin folding and dynamics, insights from efficient coarse-grained models of chromosomes. *PLoS Comput Biol* **14**, e1006159.

59. Alabert, C., Barth, T. K., Reveròn-Gòmez, N., Sidoli, S., Schmidt, A., Jensen, O. N., Imhof, A. and Grot, A. (2015) Two distinct modes for propagation of histone PTMs across the cell cycle. *Genes Dev* **29**, 585–590.

60. Sandholtz, S. H., et al. (2019) Physical modeling of the spreading of epigenetic modifications through transient DNA looping. *J Phys A: Math Theor* **52**, 434001.

61. Lieberman-Aiden, E., Berkum, N. L. V., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R, Sabo, P. J, Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., Dekker, J. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289-93.

62. Kumari, K., Prakash, R. and Padinhateeri, R. (2022) Heterogeneous interactions and polymer entropy decide organization and dynamics of chromatin domains. *bioRxiv* 2021.02.17.431616.

63. Gabrielle, M. et al. (2022) Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging. *Science* **376**, 496-501.

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

65. Cattoni, D.I., Cardozo Gizzi, A.M. and Georgieva, M. *et al.* (2017) Single-cell absolute contact probability detection reveals chromosomes are organized by multiple low-frequency yet specific interactions. *Nat Commun* **8**, 1753.

66. Gibcus, J. H. *et al.* (2018) A pathway for mitotic chromosome formation. *Science* **359**, eaao6135.

67. Alipour, E. and Marko, J. F. (2012) Self-organization of domain structures by DNA-loop-extruding enzymes. *Nucleic Acids Res.* **40**, 11202–11212.

68. Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N. and Mirny, L. A. (2016) Formation of chromosomal domains by loop extrusion. *Cell Rep.*, **15**, 2038–2049.

69. Ghosh, S. K. and Jost, D. (2020) Genome organization via loop extrusion, insights from polymer physics models. *Brief Funct Genomics.* **23**, 119-127.

70. Arnould, C. *et al.* (2021) Loop extrusion as a mechanism for formation of DNA damage repair foci. *Nature* **590**, 660-665.

71. Gillespie, D.T. (1977) Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* **81**, 25.

72. Hinrichsen, H. and Howard, M. (1999) A model for anomalous directed percolation. *Euro Phys B* **7**, 635–643.

73. Koern, M. *et al.* (2005) Stochasticity in gene expression: from theories to phenotypes. *Nat Rev Genet* **6**, 451-464.

74. Robello, R. *et al.* (2011) Retrotransposon-induced heterochromatin spreading in the mouse revealed by insertional polymorphisms. *PLoS Genet* **7**, 9.

75. Wells, J. N. and Feschotte, C. (2020) A field guide to eukaryotic transposable elements. *Annu Rev Genet* **54**, 539-561.

76. Grunstein, M. and Gasser, S.M. (2013) Epigenetics in *Saccharomyces cerevisiae. Cold Spring Harb Perspect Biol* **5**, a017491.

77. Sanulli, S. *et al.* (2015) Jarid2 methylation via the PRC2 complex regulates H3K27me3 deposition during cell differentiation. *Mol. Cell*, **57**, 769-783.

78. Zhang, K., Mosch, K., Fischle, W. *et al.* (2008) Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nat Struct Mol Biol* **15**, 381–388.

79. Caron, P. *et al.* (2015) Non-redundant functions of ATM and DNA-PKcs in response to DNA double-strand breaks. *Cell Rep* **13**, 1598–1609.

80. Simon, J. A. and Kingston, R. E. (2013) Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. *Mol Cell*, **49**, 808-824.

81. Brothers, M. and Rine, J. (2022) Distinguishing between recruitment and spread of silent chromatin structures in *Saccharomyces cerevisiae. eLife* **11**, e75653.

82. Redolfi, J., Zhan, Y., Valdes-Quezada, C., Kryzhanovska, M., Guerreiro, I., Iesmantavicius, V., Pollex, T., Grand, R. S., Mulugeta, E., Kind, J., Tiana, G., Smallwood, S. A., de Laat, W. and Giorgetti, L. (2019) DamC reveals principles of chromatin folding in vivo without crosslinking and ligation. *Nat Struct Mol Biol* **26**, 471-480.

83. Oksuz, O. *et al.* (2018) Capturing the onset of PRC2-mediated repressive domain formation. *Mol Cell* **70**, 1149–1162.

84. Lee, C. H. *et al.* (2018) Distinct stimulatory mechanisms regulate the catalytic activity of polycomb repressive complex 2. *Mol Cell* **70**, 435–48.e5.

85. Bonnet, J. *et al.* (2019) Quantification of proteins and histone marks in Drosophila embryos reveals reveals stoichiometric relationships impacting chromatin regulation. *Dev Cell* **51**, 632–644.

86. Isono, K., Endo, T. A., Ku, M., Yamada, D., Suzuki, R., Sharif, J., Ishikura, T., Toyoda, T., Bernstein, B. E. and Koseki, H. (2013) SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. *Dev Cell* **26**, 565–577.

87. Larson, A. G., Elnatan, D., Keenen, M. M., Trnka, M. J., Johnston, J. B., Burlingame, A. L., Agard, D. A., Redding, S. and Narlikar, G. J. (2017) Liquid droplet formation by $HP1\alpha$ suggests a role for phase separation in heterochromatin. *Nature* **547**, 236–240.

88. Boettiger, A., Bintu, B., Moffitt, J. *et al.* (2016) Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418–422.

89. Szabo, Q., Jost, D., Chang, J. M., Cattoni, D. I., Papadopoulos, G. L., Bonev, B., Sexton, T., Gurgo, J., Jacquier, C., Nollmann, M., Bantignies, F. and Cavalli, G. (2018) TADs are 3D structural units of higher-order chromosome organization in Drosophila. *Sci Adv* **4**, eaar8082.

90. Strom, A. R., Emelyanov, A. V., Mir, M., Fyodorov, D. V., Darzacq, X. and Karpen, G. H. (2017) Phase separation drives heterochromatin domain formation. *Nature* **547**,241–245.

91. Li, K., Bronk, G., Kondev, J. and Haber, J. E (2020) Yeast ATM and ATR kinases use different mechanisms to spread histone H2A phosphorylation around DNA double-strand break. *Proc Natl Acad Sci U S A* **117**, 21354-21363.

92. Mirny, L. A. (2011) The fractal globule as a model of chromatin architecture in the cell. *Chrom. Res.* **19**(1), 37-51.

93. Walczak, A. M., Mugler, A. and Wiggins, C. H. (2012) Analytic methods for modeling stochastic regulatory networks. In: Liu X., Betterton M. (eds) Computational modeling of signaling networks. Methods in molecular biology (Methods and Protocols), vol 880. Humana Press, Totowa, NJ.

94. Simpson, M. L., Cox, C. D. and Sayler, G. S. (2004) Frequency domain chemical Langevin analysis of stochasticity in gene transcriptional regulation. *Journal of Theoretical Biology* **229**, 383-394.

95. Elgin, S. C. R. and Reuter, G. (2013) Position-effect variegaion, heterochromatin fomration and gene silencing in Drosophila. *Cold Spring Harb Perspect Biol* **5**, a017780.

96. Montavon, T., Shukeir, N., Erikson, G. *et al.* (2021) Complete loss of H3K9 methylation dissolves mouse heterochromatin organization. *Nat Commun* **12**, 4359.

97. Wolf, D. and Goff, S. P. (2009) Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature* **458**, 1201–1204.

98. Liu, Y., et al. (2020) Very fast CRISPR on demand. *Science* **368**(6496):1265–1269.

99. Markova, E., Schultz, N. and Belyaev, I. Y. (2007) Kinetics and dose-response of residual 53BP1/gamma-H2AX foci: co-localization, relationship with DSB repair and clonogenic survival. *Int J Radiat Biol.* **83**(5), 319-29.

100. Audergon, P. N. C. B., Catania, S., Kagansky, A., Tong, P., Shukla, M., Pidoux, A. L. and Allshire, R. C. (2015) Restricted epigenetic inheritance of H3K9 methylation. *Science* **348**, 132-135.

101. Wakim, J. G., Sandholtz, S. H and Spakowitz, A. J (2021) Impact of chromosomal organization on epigenetic drift and domain stability revealed by physics-based simulations. *Biophys J* **120**, 4932–4943.

102. Sandholtz, S. H, MacPherson, Q. and Spakowitz, A. J (2020) Physical modeling of the heritability and maintenance of epigenetic modifications. *Proc Natl Acad Sci U S A* **117**, 20423-20429.

103. Ramakrishnan, N., Pillai, S. R. B and Padinhateeri, R. (2022) High fidelity epigenetic inheritance: Information theoretic model predicts threshold filling of histone modifications post replication. *PLoS Comput Biol* **18**(2), e1009861.

104. Cutter DiPiazza, A. R., Taneja, N., Dhakshnamoorthy, J., Wheeler, D., Holla, S. and Grewal, S. (2021) Spreading and epigenetic inheritance of heterochromatin require a critical density of histone H3 lysine 9 tri-methylation. *Proc Natl Acad Sci U S A* **118**, e2100699118.

105. Grosselin, K., Durand, A., Marsolier, J. *et al.* (2019) High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. *Nat Genet* **51**, 1060–1066.

106. Prompsy, P., Kirchmeier, P., Marsolier, J. *et al.* (2020) Interactive analysis of single-cell epigenomic landscapes with ChromSCape. *Nat Commun* **11**, 5702.

107. Weber, C. M., Hafner, A., Kirkland, J. G. *et al.* (2021) mSWI/SNF promotes Polycomb repression both directly and through genome-wide redistribution. *Nat Struct Mol Biol* **28**, 501–511.

108. Gartenberg, M. R. and Smith, J. S. (2016) The nuts and bolts of transcriptionally silent chromatin in saccharomyces cerevisiae. *Genetics* **203**, 1563–1599.

109. Meister, P. and Taddei, A. (2013) Building silent compartments at the nuclear periphery: a recurrent theme. *Curr Opin Genet Dev* **23**, 96–103.

110. Smith, J. S., Brachmann, C. B., Pillus, L. and Boeke, J. D. (1998). Distribution of a limited Sir2 protein pool regulates the strength of yeast rDNA silencing and is modulated by Sir4p. *Genetics* **149**, 1205-1219.

111. Michel, A. H., Kornmann, B., Dubrana, K. and Shore, D. (2005). Spontaneous rDNA copy number variation modulates Sir2 levels and epigenetic gene silencing. *Genes Dev* **19**, 1199-1210.

112. Dowen, J. M., Fan, Z. P., Hnisz, D., Ren, G., Abraham, B. J., Zhang, L. N., Weintraub, A. S., Tong Ihn Lee, J. S. and Zhao, K. (2014) Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**, 374–387.

113. Cuadrado, A., *et al.* (2019) *Cell Rep* **12**, 3500-3510.

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

115. Chory, E. J., Calarco, J. P., Hathaway, N. A., Bell, O., Neel, D. S. and Crabtree, G. R. (2019) Nucleosome Turnover Regulates Histone Methylation Patterns over the Genome. *Mol Cell* **73**, 61–72.

116. Radman-Livaja, M. *et al.* (2011) Dynamics of Sir3 spreading in budding yeast. *EMBO J* **30**, 1012–1026.

117. De, S., Mitra, A., Cheng, Y., Pfeifer, K. and Kassis, J. A. (2016). Formation of a polycomb-domain in the absence of strong polycomb response elements. *PLoS Genet* **12**, e1006200.

118. Sneppen, K. and Ringrose, L. (2019) Theoretical analysis of Polycomb-Trithorax systems predicts that poised chromatin is bistable and not bivalent. *Nat Commun* **10**, 2133.

119. Alabert, C., Loos, C., Voelker-Albert, M., Graziano, S., Forné, I., Reveron-Gomez, N., Schuh, L., Hasenauer, J., Marr, C., Imhof, A. and Groth, A. (2020) Domain Model Explains Propagation Dynamics and Stability of Histone H3K27 and H3K36 Methylation Landscapes. *Cell Rep* **30**, 1223-1234.e8.

120. Xu, Y., Zhang, S., Lin, S., Guo, Y., Deng, W., Zhang, Y. and Xue, Y. (2017) WERAM: a database of writers, erasers and readers of histone acetylation and methylation in eukaryotes. *Nucleic Acids Res* **45**, D264–D270.

121. Micheelsen, M. A. *et al.* (2010) Theory for the stability and regulation of epigenetic landscapes. *Phys Biol* **7**, 026010.

122. Agger, K., Cloos, P. A. C, Christensen, J., Pasini, D., Rose, S., Rappsilber, J. *et al.* (2007) UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449**, 731–4.

123. Torné, J., Ray-Gallet, D., Boyarchuk, E., Garnier, M., Le Baccon, P., Coulon, A. *et al.* (2020) Two HIRA-dependent pathways mediate H3.3 de novo deposition and recycling during transcription. *Nat Struct Mol Biol* **27**, 1057–68.

124. Salari, H., Di Stefano, M. and Jost, D. (2022) Spatial organization of chromosomes leads to heterogeneous chromatin motion and drives the liquid- or gel-like dynamical behavior of chromatin. *Genome Res* **32**(1), 28-43.

125. Groh, S. and Schotta, G. (2017) Silencing of endogenous retroviruses by heterochromatin. *Cell Mol Life Sci* **74**, 2055–2065.

126. Varshney, D. *et al.* (2015) Selective repression of SINE transcription by RNA polymerase III. *Mobile Genetic Elements* **5:6**, 86-91.

127. Choi, J. *et al.* (2020) DNA methylation and histone H1 jointly repress transposable elements and aberrant intragenic transcripts *Molecular Cell* **77**, 1-14.

128. Slotkin, R. K. and Martienssen,R. (2007) Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genetics* **8**, 272–285.

129. Coluccio, A., Ecco, G., Duc, J. *et al.* (2018) Individual retrotransposon integrants are differentially controlled by KZFP/KAP1-dependent histone methylation, DNA methylation and TET-mediated hydroxymethylation in naïve embryonic stem cells. *Epigenetics & Chromatin* **11**, 7.

130. Rowe, H. M. *et al.* (2010) KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **463**, 237-240.

131. Lee, Y. C. G and Karpen, G. H. (2017) Pervasive epigenetic effects of Drosophila euchromatic transposable elements impact their evolution. *eLife* **6**, e25762.

132. Groner, A. C. *et al.* (2010) KRAB–Zinc Finger Proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS Genet* **6**, e1000869.

133. Abramov, Y. A. *et al.* (2016) The differences between *cis*- and *trans*-gene inactivation caused by heterochromatin in Drosophila. *Genetics* **202**, 93–106.

134. Kitada, T., Kuryan, B. G., Tran, N. N., Song, C., Xue, Y., Carey, M. and Grunstein, M. (2012) Mechanism for epigenetic variegation of gene expression at yeast telomeric heterochromatin. *Genes Dev* **26**, 2443-55.

135. Chuong, E. B., Elde, N. E. and Feschotte, C. (2017) Regulatory activities of transposable elements: from conflicts to benefits. *Nat Rev Genet* **18**, 71-86

136. Erwin, J. A., Marchetto, M. C. and Gage, F. H. (2014) Mobile DNA elements in the generation of diversity and complexity in the brain. *Nat Reviews Neuro* **15**, 497-506.

137. Lu, J. Y. *et al.* (2021) Homotypic clustering of L1 and B1/Alu repeats compartmentalizes the 3D genome. *Cell research* **31**, 613-630.

138. Tubio, J. M. C. *et al.* (2014) Mobile DNA in cancer. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. *Science* **345**, 1251343.

139. Burns, K. H. (2017) Transposable elements in cancer. *Nat Reviews Cancer* **17**, 415-424.