Modeling the Dynamic Epigenome: from histone modifications towards self-organizing chromatin

Thimo Rohlf[1,4], Lydia Steiner[1,2], Jens Przybilla[1], Sonja Prohaska[2], Hans Binder[1,3] and Jörg Galle[1]*

[1] Interdisciplinary Centre for Bioinformatics of Leipzig University, D-04107 Leipzig, Härtelstr. 16-18, Germany
[2] Computational EvoDevo Group, Institute of Computer Science, University of Leipzig, D-04107 Leipzig, Härtelstr. 16-18, Germany
[3] Leipzig Interdisciplinary Research Cluster of Genetic Factors, Clinical Phenotypes and Environment (LIFE); Universität Leipzig, D-04103 Leipzig, Philipp-Rosenthalstr. 27, Germany
4 Max-Planck-Institute for Mathematics in the Sciences, Inselstr. 22, D-04103 Leipzig, Germany

*Corresponding author: galle@izbi.uni-leipzig.de

Abstract: Epigenetic mechanisms play an important role in regulating and stabilizing functional states of living cells. However, in spite of a fast increasing amount of experimental data, models of transcriptional regulation by epigenetic processes are rather rare. In this review, we focus on epigenetic modes of transcriptional regulation based on histone modifications, and their potential dynamical interplay with DNA methylation and higher-order chromatin structure.

The main purpose of this article is to review recent modeling approaches in this field and to relate them to available experimental data. Thereby, we highlight theoretical predictions that await experimental validation. In particular, we outline the structure of different models for histone modifications based on deterministic, as well as stochastic approaches. We evaluate their assumptions with respect to recruitment of relevant modifiers, establishment and processing of modifications, and compare the emerging stability properties and memory effects. We discuss potential extensions of these models towards multi-scale models of self-organizing chromatin.

In summary, we demonstrate that bottom-up models of epigenetic modes of transcriptional regulation are capable of providing new insights into mechanisms underlying cell differentiation, development and ageing, and represent a basic step towards hypothesis-driven studies in the field.

Keywords: transcriptional regulation, chromatin remodeling, histone modification, heritable cell fates, epigenetic memory, bistability, cooperative binding, multi-state conversion, theoretical model

Introduction

Epigenetic regulation of heritable cell fates involves transcriptional repression and activation of genes. It has been implicated in many biological processes including development and stem cell differentiation as well as ageing and cancer [1, 2]. Epigenetic regulation is based on remodeling of the chromatin structure taking place on different time and length scales [3-5].

The fundamental unit of chromatin is the nucleosome, an octamer of histones which consists of two copies of the core histones H2A, H2B, H3 and H4 and about 150 base pairs of DNA wrapped around it [4]. Histones are subject to large number of covalent posttranslational modifications such as methylation, acetylation, phosphorylation, ubiquitination and sumoylation. These modifications can be added and removed by chromatin modifying enzymes in a reversible manner (for a review see e.g. [6]). While in yeast many of the processes controlling local recruitment of the modifying enzymes are well–studied, in eukaryotes and particularly in mammalian cells they are still far from being understood.

Histones and their modifications have been implicated in the propagation of gene regulatory states from the mother cell to daughter cells along with DNA methylation and small interfering RNAs [7]. This kind of memory not stored in changes of the genome sequence has been called ‘epigenetic memory’. Epigenetic cell memory is particularly important in multi-cellular organisms where cells with identical genomes must maintain distinct functional identities, often in similar or identical environments [8], and at the same time be able to adapt in a flexible manner to changing environments [9].

The question how covalent modifications of histones impact the formation of chromatin structures that are capable of generating an epigenetic memory has been subject to various theoretical investigations. Here, we review the recent development in this research field.

In the following, we briefly address the background of model development. Afterward, we review and compare formal models of histone modification dynamics. Finally, open questions regarding the embedding of these models into a systems biological framework are discussed.

Conceptual models

The basic concept of euchromatin and heterochromatin conformations exists already since decades [10]. While euchromatin is commonly associated with a transcriptional more active state, heterochromatin is considered to repress transcription [11]. Transcription of genes is mainly governed by the action of transcription factors. In order to carry out their regulatory function these factors must gain access to their specific DNA binding sites. Here, they are aided by chromatin remodeling complexes that enable specific sites to become exposed for a fraction of time [4]. It was recognized that these complexes can be selective in the loci they modify, pointing to sequence specific chromatin remodeling and to functional local chromatin conformations. Experimental observations suggest that local chromatin conformations are largely controlled by histone modifications and DNA methylation. Several hypotheses about the underlying organization processes have been developed. For instance, the cooperative interplay between different histone modifications has been suggested to establish a so-called ‘histone code’ which decides about the recruitment of chromatin remodeling complexes [12]. However, such modeling approaches provided a qualitative view rather than a quantitative description of epigenetic regulation of gene activity.

Information-theoretical descriptions of chromatin mechanisms were developed in [13, 14]. In both approaches chromatin is taken as a finite state machine. The possible states are modulated by specific modifications and subsequent rearrangement of chromatin architecture proteins. Such transitions are formally described in terms of rewriting rules of histone

modifications. Accordingly, each architecture protein complex consists of paired selectors and modifiers (to create new modifications), of paired readers and modifiers (to propagate modifications along neighboring histones) and/or of paired readers and effectors (to modulate the transcriptional activity of genes). Implementations of these finite state machines allow to simulate the impact of different rewriting rules onto the dynamics of chromatin modification patterns and thus to address the dynamics of chromatin remodeling.

Quantitative models of the organization, dynamics, stability and inheritance of chromatin, have been developed in the last years [8, 15]. They can be regarded as specific applications of the finite state machines mentioned above. In the next section we focus on such formal modeling approaches.

Formal models

While many different hypotheses have been put forward on how chromatin modification states are established and maintained in living cells, and how combinatorial patterns of these modification states may contribute to transcriptional regulation and cellular memory, very few approaches have been developed so far that rigorously formalize basic dynamical properties of chromatin modification. Such formal models, however, appear as a fundamental step towards general, bottom-up theoretical frameworks of chromatin dynamics that enable to derive quantitative predictions on (epi-)genotype-phenotype maps from general (first) principles. These hypothesis-driven predictions, in turn, can be tested using experimental epigenetic data already available today. In the context of transcription factor networks, such formal approaches ranging from differential equation systems over Boolean networks to Petri networks [16] haven proven tremendously successful and constitute a core part of what nowadays is called ‘systems biology’. The application of these concepts to epigenetics constitutes presently a major challenge for systems biology.

Here we focus on existing formal models of histone modification dynamics (HMD) only. Applications of these models in order to explain genome-wide transcriptional regulation, as well as cross-links of histone modifications with DNA methylation and higher order chromatin structure will be discussed afterwards.

The following – certainly not exhaustive - list of basic problems regarding mathematical formalization of HMD models and their subsequent analysis will guide our overview: 1) representation of chromatin structure in space, 2) mechanisms controlling establishment and maintenance of modifications, 3) propagation of modifications on chromatin, 4) correlations between different types of modifications and 5) predictions derived from the models. Regarding the last point we address i) parameter ranges that ensure switching behavior between different modification states (bi-stability), ii) the directionality of switching (hysteresis), and iii) the inheritance of the epigenetic states (memory).

Representation of space and chromatin structure

Nucleosomes are the backbone of chromatin structure. Hence, assumptions about nucleosome positioning and chromatin folding constitute the basic level for theoretical models of HMD. While complex approaches for nucleosome (re-)positioning and association of histones and DNA have been developed [17], current HMD models typically assume a static chromatin structure represented e.g. by one-dimensional linear chains of nucleosomes [8] or neglect space in a mean field-like manner [18]. Nucleosome position variation – at least on a local scale – is known to be limited [19] and hence may be neglected. In contrast, the linear chain approximation represents a strong restriction, since chromosomes undergo 3-dimensional

folding [20, 21] with large impact on transcription factor binding [22]. Nevertheless, all HDM models discussed in the following rely on the linear chain assumption. Effects of chromosome folding will be discussed later.

Modifying histones

Chemical modifications of histones are under enzymatic control. The respective 'modifiers' must be recruited to the histones before they can start their 'writing' activity. In *S. cerevisiae* and in *D. melanogaster* the recruitment is known to be governed by sequence-dependent interactions with response elements, i.e. specific sequence motifs on the DNA [23, 24]. This fact is often neglected in HMD models [15]. We recently considered sequence specific recruitment of modifiers by assuming a finite but variable number of binding sites per DNA response element (see below). Unmethylated CpGs have been suggested as candidates for such binding sites in mammalian cells [25].

In general, different modification levels can exist, for example, several lysines can be mono-, di- or trimethylated. The coexistence of different histone modification levels in dividing yeast cells has been studied by de Vos *et al.* [26] using rate models. They described the transitions between mono-, di- and tri-methlyation of H3 at lysine 79 (H3K79) by the methylase Dot1 using coupled ordinary differential equations. The authors show that the different methylation levels of H3K79 proceed with different kinetics. This result is consistent with a non-processive mechanism, where each sub-sequent reaction is found to be slower than the one before. In consequence, a steady state modification level is reached for slowly cycling cells only. The model offers a simple approach to link different modification levels with cellular growth dynamics. However, it does not address functions related to explicit epigenetic memory involving feedback and/or cooperativity mechanisms.

Propagation of modifications

Histone modifications can cover extended genomic regions. Adequate models have therefore to address the question how modifications, initially starting from a small nucleation site, propagate along the nucleosome chain.

Early models for nucleosome modification assume a linear stepwise process, where a modified nucleosome stimulates the modification of its nearest neighbors [4]. Yet, doubts on the effectiveness and stability of such a purely local propagation mechanism soon arose. As a consequence, Dodd et al. [8] proposed an alternative propagation scheme: They suggest that, within a confined chromatin region, any other histone potentially affects the modification state of a given histone, thereby allowing a discontinuous, "jump-like" spreading of modifications. This mechanism led for the first time to a description of bistable histone modifications, being prerequisite of an inheritable epigenetic memory. The authors also show that exclusive nearest-neighbor dependent propagation leads to less pronounced bistable behavior.

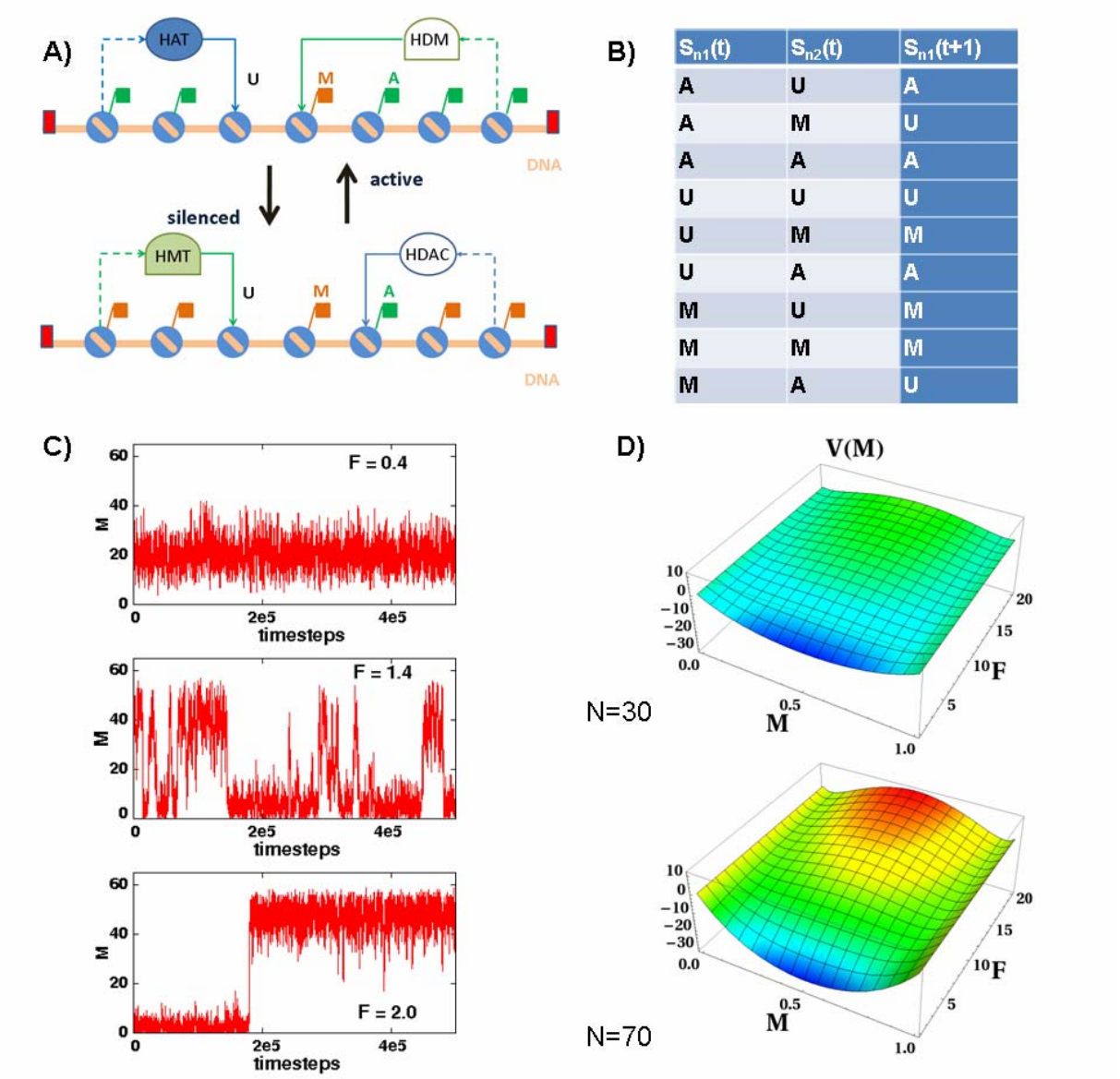
Sedighi and Sengupta [15] showed that also a localized propagation mechanism can lead to a bistable behavior in case that high cooperativity in recruiting (de-)modification complexes is assumed. In a recently developed model, we combined a "non-local" propagation mechanism of modification states with the assumption of cooperative recruitment of modifiers to nucleosomes within a DNA- binding region of definite length. The mean modification state of all nucleosomes within this cooperative region determines their modification state in a self-consistent manner. In the following we will discuss these three HMD models in more detail with the focus on cooperativity and bi-stability.

Cooperativity and bistability

As mentioned above, development and cell differentiation, as well as cellular response to environmental challenges, require the involvement of bi-stable elements which switch gene activity between states of high and low expression levels and vice versa without changes to DNA sequence. Histone modifications provide one option of epigenetic control of such switching elements. However it remains unclear, how this mechanism can induce stable cell fate decisions in the presence of considerable noise at the single nucleosome level due to effects such as the stochasticity of histone modification reactions, or a high turnover of histones themselves [8, 27]. The HMD models discussed in the following rely on two different mechanisms inducing bistable dynamics: i) indirect cooperativity induced by a multi-step conversion between mutually exclusive histone marks [8] and ii) direct cooperativity induced by positive feedback between reading and writing of modifications [15].

Bistability as consequence of multi-step histone state conversion: In their model, Dodd et al. describe histone modification dynamics in a well-defined ~20kb chromatin region of the fission yeast *S. pombe* corresponding to about N=60 subsequent nucleosomes. A sketch of the model is shown in Figure 1A. Only three relevant kinds of nucleosomes are assumed: unmodified (U), methylated (M) and acetylated (A). In a more general interpretation these three states can be considered as "unmodified", "modified" and "anti-modified". Nucleosomes are actively interconverted by modifying and demodifying enzymes (namely, histone methyl transferases HMTs, histone acetyl transferases HATs, histone demethylases HDMs, and histone deacetylases HDAs). Dodd et al. assumed that conversions between A and M always proceed via the 'intermediate' U state. The model was implemented as a stochastic cellular automaton. Accordingly, discrete modification states of single nucleosomes are updated in discrete time steps, with either recruited, i.e. deterministic, or noisy (random) conversions. The update rules for recruited conversions are given in Figure 1B.

To implement these update rules, the algorithm selects at each time step (t) randomly a pair of nucleosomes (n1 and n2) from the ensemble of N nucleosomes considered. Then, in dependence on the actual state of both nucleosomes ($S_{n1}(t)$ and $S_{n2}(t)$) the new state of n1 at t+1 is determined according the rules given in Figure 1B. For example, the combination A (n1) and M (n2) leads to conversion A → U for nucleosome n1. Changes of the type A → M or M → A thus always involve at least two steps. While recruited conversions occur with probability α, noisy conversions occur with probability $1 - \alpha$. In a noisy conversion the state of nucleosome (n_1) is changed towards the possible neighboring states (e.g., A,M →U and vice versa). The degree of bistability exhibited by the system increases with increasing feedback-to-noise ratio $F := \alpha/(1 - \alpha)$.



$S_{n1}(t)$	$S_{n2}(t)$	$S_{n1}(t+1)$
A	U	A
A	M	U
A	A	A
U	U	U
U	M	M
U	A	A
M	U	M
M	M	M
M	A	U

Figure 1: HMD-model of acetylation/methylation balance in *S. pombe* [8]. A) Model sketch. The states (M, U, A) of a randomly selected nucleosome triggers the recruitment of specific enzymes (dotted lines) and a subsequent state transition (solid lines) of a second nucleosome. Involved enzymes are: histone methyl transferases (HMTs), histone acetyl transferases (HATs), histone demethylases (HDMs), histone deacetylases (HDAs). B) Overview over the models' transition rules for recruited (deterministic) conversions. C) Simulation results of the model for different signal to noise ratios F, for N = 60 interacting nucleosomes. With increasing F the system starts exhibiting strong bistability. D) Epigenetic landscapes for the simplified two-state model [28] considering different system sizes (N). The valleys become more pronounced with increasing N.

The cellular automaton defined this way exhibits bi-stable behavior, which is controlled by noise (see Figure 1C). The authors demonstrate that it is the two-step conversion mechanism of the model which inherently implements this cooperative behavior. A conversion from A to M, for example, requires two consecutive recruitments by nearby M nucleosomes, and thus has a rate proportional to M^2. By limiting recruited modification of nucleosome (n1) to stimulation by adjacent nucleosomes (n2) only, the authors show that step-wise, local propagation of modifications produces poor bistability.

In a subsequent publication Dodd and co-workers generalize the proposed mechanisms. Now, bistability is considered to be associated with valleys in a so-called 'epigenetic landscape'

(see Figure 1D, and [28]). For this purpose, a simplified two-state model (states M, A) is applied where cooperativity is taken into account by assuming that two nucleosomes of the same type (e.g. M) are needed to trigger a conversion between the states (A→M). In a formal analysis of this model, they formulated a Langevin equation for the modified state (M) and studied model dynamics in terms of a one-dimensional Fokker-Planck equation. Accordingly, changes of histone modification levels are treated as a diffusion process in an effective potential V(m) that describes the shape of the underlying epigenetic landscape. Interestingly, the valley structure of the epigenetic landscape becomes more pronounced with increasing number N of cooperatively interacting nucleosomes (Fig. 1D), leading to an exponentially increasing stability (average lifetime) of epigenetic states with N. It turns out that bistability requires a minimum number of cooperative histones and that the size of the parameter space allowing for bistability increases with N.

Bistability as consequence of cooperative protein binding to chromatin. A different HMD model was introduced in [15]. Originally it was developed to describe epigenetic silencing in budding yeast *S. cerevisiae*, where the underlying epigenetic mechanisms are relatively well investigated [29]. In this model silencing is considered to result from interplay between histone de-acetylation and cooperative binding of so-called Sir protein complexes. The model is formulated in terms of chemical reaction equations for the local degree of histone acetylation (A) and the local probability of occupation by Sir (S) along a chromatin fiber of undefined length. Independent model parameters are i) the time dependent bulk concentration of ambient Sir complexes and ii) the rate of de-acetylation. The model can be translated to the more general case of two modification states of chromatin where the production of one state depends non-linearly on the other. Accordingly, the authors demonstrate that assumption of either cooperative Sir complex binding or of a 'transcriptional' feedback on the histone acetylation are sufficient to obtain bistable behavior in this model. This features are formally introduced by the functions $f(A)=(1-A)^n$ and $g(S)=(1-S)^m$, modulating the rates of Sir binding and of histone acetylation, respectively (n and m are adjustable parameters). The authors discuss the case of finite supply of Sir proteins as an interesting extension of the model. They considered a decrease of the bulk concentration of Sir assuming that de-acetylation and Sir binding proceed along extended chromatin regions not subjected to partitioning by boundary elements. Sir binding is assumed to deplete its bulk concentration which in consequence decreases the Sir binding rate and thus slows down the propagation of the modification state. This topic has been investigated in more detail in a subsequent publication [30].

A HMD-model comparable to that of Sedighi & Sengupta was recently introduced by Binder *et al.* for epigenetic silencing in eukaryotes [31]. The model is motivated by the structure and function of Polycomb group (PcG) and trithorax group (trxG) complexes in heterochromatin and euchromatin formation [24]. The basic framework of the model is summarized in Figure 2. In this model cooperative binding of a chromatin modifying enzyme to response elements (RE) leads to bistable epigenetic states. Reversible binding of a protein complex to these specific genomic loci is described by a classical binding isotherm (Figure 2B, Equ. 1). The bound complex is capable of writing modifications on the histones associated with the RE. Histone modification further facilitates complex binding. This positive feedback loop gives rise to bistability of the transcriptional activity of the genes associated with the RE.

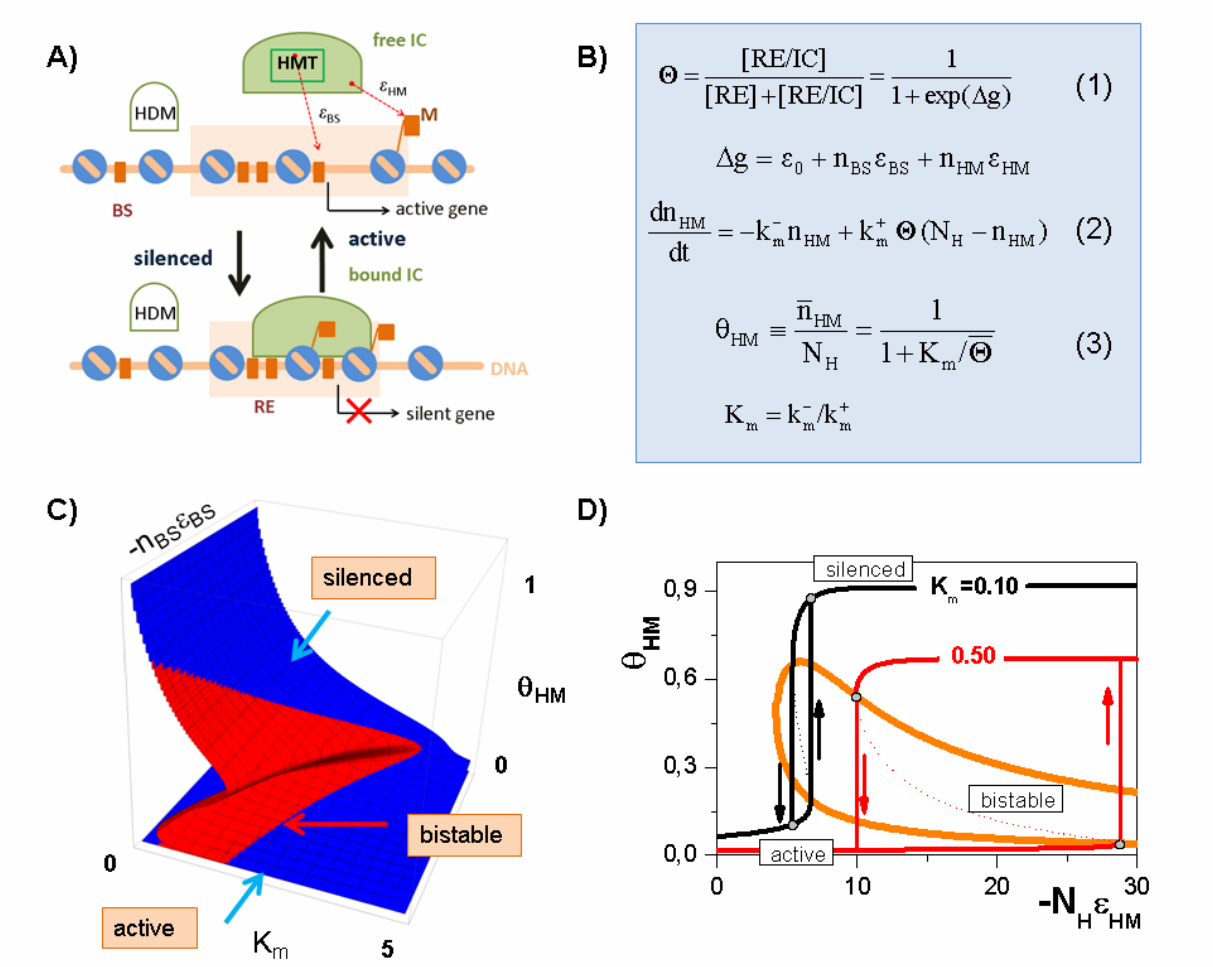


Figure 2: HMD-model of histone methylation in eukaryotic cells [31]. A) Sketch of the model. Free interaction complexes (ICs) interact with binding sites (BS) in the response element (RE) as well as with modified histones (M). Bound ICs trigger additional histone modifications that further improve binding. Assuming a repressive model, binding reversibly silences associated genes that are otherwise active. Unspecific de-modification occurs permanently. B) Main equations as explained in the text: 1): Binding isotherm. The probability of complex binding is given by the free enthalpy Δg of binding. 2): Modification kinetics. 3): Self-consistent steady state solution for the fraction of modified histones. C) Solutions of Equ. 3 in dependence of the sequence specific binding energy $-n_{BS}\varepsilon_{BS}$ and the de-modification constant K_m. D) Hysteretic behavior of the system in dependence of the number of cooperatively acting nucleosomes N_H in units of $-\varepsilon_{HM}$.

Particularly, the model assumes that an RE contains N_H nucleosomes which are under the potential influence of a modification complex. The changes of the number of modified histones n_{HM} per RE are described by Equ. 2 in Figure 2B. Here, k^-_m and $k^+_m\,\Theta$ define the rates of de-modification and modification, respectively. The latter scales with the occupancy of the RE with bound complexes ($0 \le \Theta \le 1$). This implies that histones are modified in the presence of bound complexes only. The RE-occupancy Θ is determined by the free enthalpy Δg of complex binding. Δg depends on three contributions: i) a basal repulsive interaction term $\varepsilon_0 > 0$ which prevents unspecific association of the complex with chromatin and DNA, ii) an attractive interaction term $\varepsilon_{BS} < 0$ describing the specific interaction of the complex with one of the n_{BS} DNA binding sites per RE, and iii) an attractive interaction term $\varepsilon_{HM} < 0$

describing the interaction of the complex with one of the n_H modified histones per RE. Under steady state conditions, the model provides a self-consistent solution which links the RE occupancy with the fraction of modified histones per RE, θ_{HM}, given by Equ.3 in Figure 2B.

The model shows bistable behavior regarding histone modification depending on the specific binding term ($-n_{BS}\varepsilon_{BS}$), and on the ratio $K_m = k^-/k^+$ characterizing the steady state of the modification reaction which is under enzymatic control. Figure 2C shows the respective hyperplane which divides into regions of monostable (blue) and bistable (red) solutions. The monostable region, in turn, is split into two states of high and low modification degree referring to silenced and active gene expression, respectively, in case of a repressive model.

Importantly, the maximum number of histones per RE, N_H, considerably affects systems behavior. It determines the strongest possible attraction that is exerted by the histones on the complex. In consequence, bistability is governed by the length of cooperatively acting chromatin given by N_H. Bistability requires a minimum length of the cooperative unit in agreement with the Dodd-model discussed above (see Figure 2D).

In summary, the three models discussed above demonstrate that bistability in chromatin states, and thereby regulatory switching and hysteretic behavior, can be generated by different mechanisms that, however, share a common fundamental principle, namely cooperative behavior. In the following, we discuss implications of this principle on inheritance of cell fates.

Inheritance of histone marks

In multicellular organisms, epigenetic regulation has evolved towards highly complex organisation. In these systems, epigenetic cell memory is particularly important, because cells with identical genomes first must achieve distinct phenotypes in course of development and differentiation, but later on also be able to stably maintain their identities under cell divisions. This requires stable inheritance of epigenetic states (cell fates) across many cell generations. In this context, DNA replication upon cell division poses a major problem: histone modifications are bound to get diluted during DNA replication, requiring a subsequent reconstitution of the parental state for faithful inheritance. A common model of this process, illustrated in Figure 3A, assumes that the parental nucleosomes are first randomly partitioned between the two daughter strands and then the diluted histone populations are complemented by de novo synthesized and assembled histones [32]. Heritable epigenetic states must be stable against such large-scale perturbations. This problem has been addressed in all three HDM-models discussed above.

In their stochastic model, Dodd et al. [8] investigate the ability of high-M and low-M states to be maintained through DNA replication. In simulation series they replaced each nucleosome at the time of replication by a U (unmodified) nucleosome with a probability of one-half. They found that high stabilities of modification states can be realized at modest feedback-to-noise ratios F. Interestingly, the number of switches becomes independent of generation time τ above a minimal value τ_{min}, and for F larger or equal than 2. This result demonstrates that transitions between modification states are much more likely to occur immediately after replication than at any other point of the cell cycle. This dichotomy can be interpreted as a gradual development in a well-defined epigenetic landscape (compare Fig. 1D) *between* cell divisions, with sudden reshufflings during replication that entirely dominate state transitions [28].

In order to systematically elucidate the conditions for stable inheritance of histone modification states David-Rus et al. [18] formulated a general stochastic model of epigenetic

inheritance. Subsequently, they develop a mean-field theory of the process, ignoring spatial variation of modifications and replacing them by an average value corresponding to the entire region of chromatin considered. They analysed two-and three-state models of modification dynamics. The requirement of stability against perturbations at cell division imposes constraints on rate parameters. For two-state models, cooperative conversion of U ("unmodified") histones into A ("antimodified") ones must occur with a higher rate than vice versa. As a consequence of this asymmetry, fluctuations can flip a low-A state into a high-A states within a few cell cycles.

Three-state models (with the model by Dodd et al. [8] as a particular realisation) are more stable in this regard. The authors suggest, that "the presence of multiple epigenetic marks is a design criterion for epigenetic stability", where the higher dimensionality of the state space gives rise to the increased stability of the system.

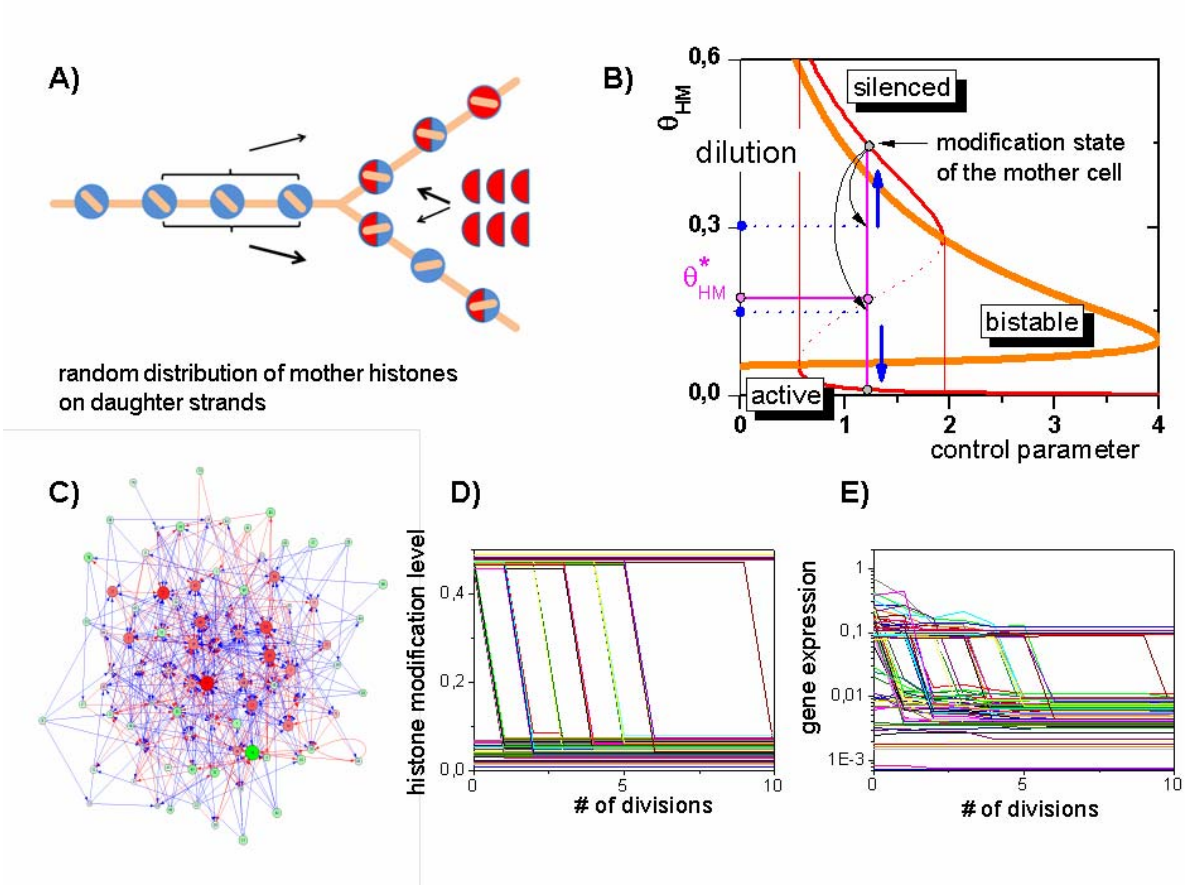


Figure 3: Inheritance of histone marks. A) Dilution of histone marks. During cell division mother histones (blue) are randomly distributed onto the two daughter strands and complemented with de novo synthesized histones (red). B) Consequences of the marker dilution in bistable modified regions. Both daughter strands carry initially a lower marker density then the mother cell, which is assumed to be in a high modification state. Depending on whether the diluted states are located in the attractor of the low modification state (lower blue state) or not (higher blue state) a spontaneous de-modification can appear. C) Transcription factor network generated by a Random Genome (RG) model. Red nodes represent activated genes, green nodes repressed genes. D,E) RG dynamics under the control of an activating histone modification. Partial de-modification of response elements in course of cell division D) leads to decreased gene expression of in the network E).

We recently discussed an additional option to increase this dimensionality (and thereby stability) [31], namely the superposition of different interaction terms governing complex binding and cooperative modification dynamics. In the parameter range of bistability (see Figure 3C), this can lead to an initial cascade of transitions during the first few generations of

cells (reminiscent e.g. of a cell differentiation cascade) that quickly settles onto a stable differentiated stationary state (see Figure 3D,E); an observation that was also made by Micheelsen et al. [28] in the framework of epigenetic landscapes.

These studies demonstrate that epigenetic marks can be stable against perturbations only under very specific conditions. Thus, in cycling cells and to a smaller degree also in quiescent cells permanent epigenetic remodeling has to be considered, even in stable environments. Notably these changes are not purely random. Instead, changes in a particular chromatin region can be more likely than in other regions.

Coupling to transcriptional regulation

Processing of epigenetic states into cellular phenotypes requires transcription, and in the following translation of genes. Effects of chromatin conformations on transcription factor binding are well established [4], however, descriptions of these phenomena often remain qualitative. The HMD models discussed above suggest that this coupling may lead to interesting new phenomena, for example, ultrasensitivity due to highly non-linear amplification effects.

Sneppen *et al.* [33] showed that non-cooperative binding of a TF to a single site can produce a large change in gene expression in response to even a small change in concentration of the TF. This ultrasensitivity is caused by the parallel recruitment of a histone-modifiying enzyme that changes the balance between the assumed feedback loops in histone modification. The resulting asymmetry between modification and anti-modification can be interpreted as a TF-induced deformation of the underlying epigenetic landscape [28], pointing at “dynamic epigenetic landscapes” as a possible, general theoretical framework to model developmental and differentiation processes.

In addition to transcription factors, also signal transduction networks can interfere with epigenetic systems [34], enabling the cell to respond to environmental effects. In this context, the model of Dodd et al. was applied to a Polycomb-based switch with impact for the epigenetic memory in vernalization of Arabidopsis [35], i.e. the acquisition of a plant's ability to flower or germinate in the spring. Experiments show that polycomb PRC2-controlled silencing of the floral repressor FLC is involved in this process: H3K27me3 levels continuously increase within a small nucelation region at the FLC genomic locus during the cold. At more distal regions the increase of H3K27me3 levels depends additionally on the presence of a sub-sequent warm period. These results suggest a possible interference of histone modification with environmental signals.

Sedighi & Sengupta [15] and Mukhopadhyay et al. [30] show that the recruitment of histone modifying enzymes by TFs (or other molecular complexes) can substantially contribute to the formation of cooperative loops for writing and maintenance of modifications. Hence, current theoretical approaches – although far from realistic levels of complexity - already account for the cross-talk between different epigenetic control systems.

In conclusion, future – more general – models will have to address not only the effect of coupling between different types of histone modifications on (epi-)genotype-phenotype maps, but also have to integrate the coupled dynamics between transcription factor networks and chromatin-based regulation, as well as higher order effects e.g. from the three-dimensional structure of chromosomes. Considerations that may guide approaches towards adequate theoretical frameworks will be outlined in the following section.

Towards a systems biology perspective on epigenetics

The genomewide distribution of various histone modifications can be studied using a combination of ChIP with next generation sequencing (ChIP-seq) technology [36]. Large data sets on different types of modifications in different systems have been published, including data on embryonic stem cells, various lineage committed multipotent progenitors and differentiated cells [37-39]. A comprehensive toolbox is available for processing, analysis, and visualisation of these high-dimensional data sets [40]. Although the models discussed here are all inspired by experimental results, direct adaptation of them to genome-wide data has been not provided so far. Angle et al. [35] for the first time fitted model output data to population averaged modification levels obtained by ChIP experiments. Thus, a major challenge posed in systems biology is the direct quantitative analysis of genome-wide modification data using concepts and hypotheses developed in modelling approaches on the dynamics and inheritance of histone modifications.

A quantitative approach of this kind is not only required for genome-wide modification data, but also for related expression data. However, the models discussed here have not been applied in such a context so far although they link the formation of different chromatin structures with switching of genes between active and silenced transcription in ultra-sensitive regulation circuits [33]. An essential step in that direction would be the integration of histone modification models into multi-scale models of transcriptional regulation. Fortunately, the ChIP-seq data sets typically comprise both, modification and expression data and thus they can serve as starting point for such joint quantitative approaches (e.g. Mikkelsen et al. [37]).
Combined simulation of transcriptional regulation by cis-regulatory elements and histone modification will support our understanding of the impact of these different layers of regulation, for example, on development and stem cell differentiation. In this context, integration of other epigenetic modes of transcriptional regulation such as DNA-methylation may be required. Molecular coupling of DNA methylation and histone methylation has been demonstrated recently [41, 42] leading to a coordinated regulation of gene expression involving different time scales [5] and different potentials for stable inheritance of epigenetic information.

Models of the dynamics, stability and inheritance of DNA methylation have been introduced already [43, 44]. They highlight the role of cooperative action of maintenance and de novo DNA methylation for stable inheritance of this epigenetic mark. Moreover, genome-wide high-resolution methylation maps are available that complement histone modification data [45]. However, similar to the histone modification models discussed above, also DNA methylation models await integration into a multi-scale modelling approach to transcriptional regulation.

Artificial genomes may help to solve this problem. They provide a simple framework that allows the straightforward modular integration of different layers of regulation and simple testing of hypotheses about relevant interactions between these layers. Random Genomes (RG) as the simplest type of artificial genomes have been developed a decade ago by Reil [46]. Recently, a specially designed RG model has been applied to analyse global gene expression characteristics [47]. A transcription factor network defined by a particular realisation of a RG is shown in Figure 3C. We have linked this approach to the histone modification model of Binder *et al*. [31]. We assumed that genomic regions defining the genes in the RG are associated with cooperatively acting chromatin regions (separated from each other by, e.g., insulator elements). Moreover, we assumed direct proportionality between

the promoter activity of the genes and the modification level of their regulatory region, as already suggested by Sneppen et al. [33]. In an exemplary study, we simulated how an activating modification corresponding, e.g., to H3K4me3 affects de-modification of chromatin regions in the course of proliferation (see Figure 3D). Figure 3E shows the corresponding progressive decrease of the expression levels of associated genes.

In this example, each gene is regulated individually. Chromatin modifications, however, can induce coordinated expression changes of *groups* of neighbouring genes. A prominent example of this property of chromatin represents regulation of the Drosophila Hox gene cluster. Wit & van Steensel [48] suggested three prototypes of multigene chromatin domains: i) the spreading of chromatin proteins along the DNA, ii) looping of the chromatin fibre, and iii) compartmentalisation of nonadjacent chromatin regions by clustering. The models discussed here refer exclusively to the first option. In this case a simplified one-dimensional view on chromatin appears to be appropriate. Even these simple models can be made more realistic in straight-forward ways, allowing to relate their predictions to available genome-wide data, e.g. on the length distributions of modified regions. We were able to reproduce experimentally observed length distributions of modified H3K4me3 and H3K27me3 regions in murine stem cells by assuming insulator elements, which fragment the chromatin fibre into subdomains [49]. The derived values of model parameters such as modification rates or interaction strengths are critically related to bistability of chromatin modifications and associated gene activity [31].
Explicit consideration of looping and compartmentalisation, however, would require higher dimensional approaches. Note that the models of cooperative binding of histone modifying complexes discussed here already *imply* spatial effects via non-local interaction. Compartmentalization of chromatin into transcription factories providing disjoint local environments has been observed experimentally [50]. Moreover histone modifications and DNA methylation are thought to induce cooperative loopings of the chromatin fibre with implications for the activity of associated genes [50]. The integration of such spatio-temporal aspects of chromatin organisation into multi-scale models of transcriptional regulation may represent another future step towards a comprehensive systems level description of the epigenome.

Conclusions

Epigenetic mechanisms of transcriptional regulation pose new problems in mathematical modeling. Particularly, the description of transcriptional regulation by histone modifications is of high relevance for the understanding of many biological processes, including development and differentiation. As the essential model ingredient required for epigenetic memory and inheritance of epigenetic information, the approaches developed so far always identified positive feedback loops based on cooperative non-local interactions between the histones and the modifying molecules. Moreover, such regulation circuits imply ultrasensitive responses of gene expression. Despite the success achieved in the description of basic principles of epigenetic regulation the comprehensive integration of whole genome transcriptional and epigenetic data into modeling is still missing.

Future Perspective

Epigenetic memory allows cells with identical DNA to maintain distinct functional identities. Patterns of epigenetic modifications have been demonstrated to diverge in monozygotic twins as they become older [51]. Such differences have been early related either to environmental

factors or to reduced inheritance of the epigenetic information during ageing [52]. Regardless of the huge amount of experimental data that has been emerged since then, the mechanisms of epigenetic remodeling are still poorly understood. Reconstructing epigenetic networks ('modification webs') is still a largely unsolved problem [53] that provokes even more detailed and comprehensive measurements. However, there is increasing evidence that generally not all possible combinations of modifications can be observed and that large but specific patterns of modifications can characterize a single functional state [38]. This questions the design of many experimental studies.

Mathematical models of the dynamics, stability and inheritance of epigenetic marks allow to generate well-founded hypotheses regarding the mechanisms on work and to design effective protocols for their experimental validation. Thus, they will support cost-efficient research approaches. However, prerequisite of a Renaissance of such hypotheses-driven research approaches will be an explanatory understanding of measured absolute chromatin modifications levels.

An emerging field in epigenetic research is ageing. Epigenetic changes have been linked to a decline in stem cells function [54, 55] and epigenetic 'reprogramming' is considered in future therapeutic applications [56, 57]. Recent findings demonstrate that age-associated hypermethylation occurs in bivalent modified chromatin domains pointing to a close link between the different layers of regulation also in this process [58]. Aberrant epigenetic changes have also been recognized in cancer development and 'cancer epigenetics' has reached mainstream oncology [59]. Only recently it has been shown that age-dependent DNA methylation at genes that are suppressed in stem cells is a hallmark of cancer explaining age as major risk factor in cancer [60].

In both, ageing and cancer development the question arises why clones of cells which carrying a particular modified epigenome become dominate or vanish over time. Thus, in order to understand epigenetic phenomena in ageing and cancer clonal competition in stem cell niches has to be considered. This will require simulation on the cell level describing growing populations and regenerating tissues. Individual cell-based models of such systems have been established [61-63]. We envision an integration of complex models of transcriptional regulation with these approaches into a comprehensive model framework.

Executive Summary

Conceptual models

- Cooperative interplay between different histone modifications has been suggested to establish a ‘histone code’
- Finite state machines constitute a general, information-theoretical framework to study the impact of different rewriting rules on chromatin modification patterns

Formal models

- Models of histone modification dynamics use simplified spatial structures (e.g. linear genome chain) and predict that long-range interactions between nucleosomes are essential for effective modification propagation
- Different types of cooperative interactions can lead to bistability (switching between different modification states), effective parameter range increases with system size
- Replication is introduced as a global stochastic fluctuation; the models dicussed can reproduce stable inheritance of modification marks, as well as differentiation

- Coupling to transcription is possible and predicts novel effects (e.g. ultra-sensitive gene regulation)

Towards a systems biology perspective on epigenetics

- Quantitative, multi-scale approaches are needed that link predictive, formal models of the modification dynamics of different epigenetic marks and transcriptional regulation to genome-wide experimental data
- Artificial genomes represent a first step towards such a comprehensive modeling framework in this direction

Future perspective

- Mathematical models of the dynamics, stability and inheritance of epigenetic marks will lead to novel hypotheses guiding future design of experimental protocols
- Particularly promising application fields are ageing and cancer development

Bibliography

1. Reik W: Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 447, 425-431 (2007).
2. Oakley EJ, Zant GV: Unraveling the complex regulation of stem cells: implications for aging and cancer. *Leukemia* 21(4), 612--621 (2007).
3. Bernstein BE, Meissner A, Lander ES: The mammalian epigenome. *Cell* 128(4), 669--681 (2007).
4. Felsenfeld G, Groudine M: Controlling the double helix. *Nature* 421, 448-453 (2003).
5. Kuzawa CW: Timescales of human adaptation: the role of epigenetic processes. *Epigenomics* 3(2), 221-234 (2011).
6. Latham JA, Dent SYR: Cross-regulation of histone modifications. *Nat Struct Mol Biol* 14(11), 1017--1024 (2007).
7. Probst AV, Dunleavy E, Almouzni G: Epigenetic inheritance during the cell cycle. *Nat Rev Mol Cell Biol* 10(3), 192--206 (2009).
8. Dodd IB, Micheelsen MA, Sneppen K, Thon G: Theoretical Analysis of Epigenetic Cell Memory by Nucleosome Modification. *Cell* 129, 813-822 (2007).

** In this paper, the first mathematical model of histone modification dynamics is introduced that leads to bistable behavior from cooperative multi-state conversions.

9. Szyf M: Implicatiions of a life-long dynamic epigenome. *Epigenomics* 1(1), 9-12 (2009).
10. Sutton E: The Structure of Salivary Gland Chromosomes of Drosophila Melanogaster in Exchanges between Euchromatin and Heterochromatin. *Genetics* 25(5), 534–540. (1940).
11. Huisinga KL, Brower-Toland B, Elgin SC: The contradictory definitions of heterochromatin: transcription and silencing. *Chromosoma* 115(2), 110-122 (2006).
12. Jenuwein T, Allis CD: Translating the histone code. *Science* 293(5532), 1074--1080 (2001).
13. Benecke A, Group SE: Chromatin code, local non-equilibrium dynamics, and the emergence of transcription regulatory programs. *Eur Phys J E Soft Matter* 19(3), 353--366 (2006).
14. Prohaska SJ, Stadler PF, Krakauer DC: Innovation in Gene Regulation: The Case of Chromatin Computation. *J. Theor. Biol.* 265, 27-44 (2010).

* In addition to a detailed evolutionary study of histone modification readers and -writers, a finite state machine approach to chromatin computation is proposed.

15. Sedighi M, Sengupta AM: Epigenetic chromatin silencing: bistability and front propagation. *Physical Biology* 4(4), 246 (2007).

** In this paper, a complementary mathematical approach to Dodd et al. for histone modification dynamics is studied. Explicit cooperativity of histone de-acetylations and Sir complex binding leads to bistability and stable inheritance of modifications.

16. Karlebach G, Shamir R: Modelling and analysis of gene regulatory networks. *Nat Rev Mol Cell Biol* 9(10), 770--780 (2008).
17. Segal E, Widom J: What controls nucleosome positions? *Trends Genet* 25(8), 335--343 (2009).
18. David-Rus D, Mukhopadhyay S, Lebowitz JL, Sengupta AM: Inheritance of epigenetic chromatin silencing. *J Theor Biol* 258(1), 112--120 (2009).
19. Arya G, Maitra A, Grigoryev SA: A structural perspective on the where, how, why, and what of nucleosome positioning. *J Biomol Struct Dyn* 27(6), 803--820 (2010).
20. Misteli T: Beyond the sequence: cellular organization of genome function. *Cell* 128(4), 787--800 (2007).

21. Duan Z, Andronescu M, Schutz K *et al.*: A three-dimensional model of the yeast genome. *Nature* 465(7296), 363--367 (2010).
22. Vilar JMG, Leibler S: DNA Looping and Physical Constraints on Transcription Regulation. *J. Mol. Biol.* 331, 981–989 (2003).
* A mathematical model for the impact of DNA looping on transcription factor binding is studied. Similar approaches may be applied to develop 3D-models of epigenetic dynamics.
23. Pirrotta V, Gross DS: Epigenetic silencing mechanisms in budding yeast and fruit fly: different paths, same destinations. *Mol Cell* 18(4), 395--398 (2005).
24. Simon JA, Kingston RE: Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat Rev Mol Cell Biol* 10(10), 697-708 (2009).
25. Thomson JP, Skene PJ, Selfridge J *et al.*: CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature* 464(7291), 1082-1086 (2010).
26. Vos DD, Frederiks F, Terweij M *et al.*: Progressive methylation of ageing histones by Dot1 functions as a timer. *EMBO Rep* 12(9), 956--962 (2011).
27. Dion MF, Kaplan T, Kim M, Buratowski S, Friedman N, Rando OJ: Dynamics of replication-independent histone turnover in budding yeast. *Science* 315(5817), 1405--1408 (2007).
28. Micheelsen MA, Mitarai N, Sneppen K, Dodd IB: Theory for the stability and regulation of epigenetic landscapes. *Physical Biology* 7(2), 026010 (2010).
29. Rusche LN, Kirchmaier AL, Rine J: Ordered nucleation and spreading of silenced chromatin in Saccharomyces cerevisiae. *Mol Biol Cell* 13(7), 2207-2222 (2002).
30. Mukhopadhyay S, Nagaraj VH, Sengupta AM: Locus dependence in epigenetic chromatin silencing. *Biosystems* 102(1), 49--54 (2010).
31. Binder H, Steiner L, Wirth H, Rohlf T, Prohaska SJ, Galle J: Transcriptional memory emerges from cooperative histone modifications. *BMC System Biology*, (submitted 2011).
**The model presented here, besides cooperative histone modifications, also considers DNA-specific binding of modification complexes, and is fitted to genome-wide length distributions of H3K4me3 and H3K27me3 modified regions.
32. Annunziato AT: Split Decision: What Happens to Nucleosomes during DNA Replication? *J. Biol. Chem.* 280(13), 12065-12068 (2005).
33. Sneppen K, Micheelsen MA, Dodd IB: Ultrasensitive gene regulation by positive feedback loops in nucleosome modification. *Mol. Sys. Biol.* 4, 182 (2008).
34. Mohammad HP, Baylin SB: Linking cell signaling and the epigenetic machinery. *Nat Biotechnol* 28(10), 1033-1038 (2010).
35. Angel A, Song J, Dean C, Howard M: A Polycomb-based switch underlying quantitative epigenetic memory. *Nature* 476, 105–108 (2011).
36. Park PJ: ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet* 10(10), 669--680 (2009).
37. Mikkelsen TS, Ku M, Jaffe DB *et al.*: Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448(7153), 553-560 (2007).
38. Wang Z, Zang C, Rosenfeld JA *et al.*: Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* 40(7), 897--903 (2008).
*Genome-wide analysis of histone acetylations and methylations in the human genome shows that certain specific, large-scale modification patterns are statistically overrepresented.
39. Cui K, Zang C, Roh T-Y *et al.*: Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. *Cell Stem Cell* 4(1), 80--93 (2009).
40. Bock C, Lengauer T: Computational epigenetics. *Bioinformatics* 24(1), 1--10 (2008).

41. Ooi SKT, Qiu C, Bernstein E *et al.*: DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* 448(7154), 714--717 (2007).
42. Hashimoto H, Vertino PM, Cheng X: Molecular coupling of DNA methylation and histone methylation. *Epigenomics* 2(5), 657--669 (2010).
43. Sontag LB, Lorincz MC, Georg Luebeck E: Dynamics, stability and inheritance of somatic DNA methylation imprints. *J Theor Biol* 242(4), 890-899 (2006).
* The first formal model that describes the dynamics, stability and inheritance of somatic DNA methylation imprints. An interesting approach for integration in comprehensive artifical genome models.
44. Genereux DP, Miner BE, Bergstrom CT, Laird CD: A population-epigenetic model to infer site-specific methylation rates from double-stranded DNA methylation patterns. *Proceedings of the National Academy of Sciences of the United States of America* 102(16), 5802-5807 (2005).
45. Meissner A, Mikkelsen TS, Gu H *et al.*: Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454(7205), 766-770 (2008).
46. Reil T: Dynamics of gene expression in an artificial genome - Implications for biological and artificial ontogeny. *Lect Notes Artif Int* 1674, 457-466 (1999).
* The first proposal for in-silico construction of transcription factor networks from first principles based on sequence binding.
47. Binder H, Wirth H, Galle J: Gene expression density profiles characterize modes of genomic regulation: theory and experiment. *J Biotechnol* 149(3), 98-114 (2010).
48. De Wit E, Van Steensel B: Chromatin domains in higher eukaryotes: insights from genome-wide mapping studies. *Chromosoma* 118(1), 25-36 (2009).
49. Bushey AM, Dorman ER, Corces VG: Chromatin insulators: regulatory mechanisms and epigenetic inheritance. *Mol Cell* 32(1), 1--9 (2008).
50. Razin SV, Gavrilov AA, Pichugin A, Lipinski M, Iarovaia OV, Vassetzky YS: Transcription factories in the context of the nuclear and genome organization. *Nucleic Acids Res*, (2011).
51. Fraga MF, Ballestar E, Paz MF *et al.*: Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* 102(30), 10604--10609 (2005).
52. Cooney CA: Are somatic cells inherently deficient in methylation metabolism? A proposed mechanism for DNA methylation loss, senescence and aging. *Growth Dev Aging* 57(4), 261--273 (1993).
53. Hayashi Y, Senda T, Sano N, Horikoshi M: Theoretical framework for the histone modification network: modifications in the unstructured histone tails form a robust scale-free network. *Genes to Cells* 14(7), 789--806 (2009).
54. Liu L, Rando TA: Manifestations and mechanisms of stem cell aging. *J Cell Biol* 193(2), 257--266 (2011).
55. Pollina EA, Brunet A: Epigenetic regulation of aging stem cells. *Oncogene* 30(28), 3105--3126 (2011).
56. Ordovas JM, Smith CE: Epigenetics and cardiovascular disease. *Nature Reviews Cardiology* 7(9), 510-519 (2010).
57. Penner MR, Roth TL, Chawla MK *et al.*: Age-related changes in Arc transcription and DNA methylation within the hippocampus. *Neurobiol Aging*, (2010).
58. Rakyan VK, Down TA, Maslau S *et al.*: Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Research* 20(4), 434-439 (2010).
59. Rodríguez-Paredes M, Esteller M: Cancer epigenetics reaches mainstream oncology. *Nat Med* 17(3), 330--339 (2011).

60. Teschendorff AE, Menon U, Gentry-Maharaj A *et al.*: Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Res* 20(4), 440-446 (2010).
61. Galle J, Hoffmann M, Aust G: From single cells to tissue architecture-a bottom-up approach to modelling the spatio-temporal organisation of complex multi-cellular systems. *J Math Biol* 58(1-2), 261-283 (2009).
62. Galle J, Aust G, Schaller G, Beyer T, Drasdo D: Individual cell-based models of the spatial-temporal organization of multicellular systems--achievements and limitations. *Cytometry A* 69(7), 704-710 (2006).
63. Thorne BC, Bailey AM, Desimone DW, Peirce SM: Agent-based modeling of multicell morphogenic processes during development. *Birth Defects Res C Embryo Today* 81(4), 344--353 (2007).