Grundlagen der Systembiologie und der Modellierung epigenetischer Prozesse

Christian Arnold

Bioinformatics Group,
Dept. of Computer Science & Interdisciplinary Center for Bioinformatics,
University of Leipzig

December 6th, 2010

Table of Content

1. Epigenetics and Epigenetic Inheritance

The importance of epigenetic inheritance

Epigenetics: "The study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in the DNA sequence." (Riggs et al. 1996)

- stable propagation of gene expression programs during cell division $\rightarrow$ Cells can remember who they are!
- particularly important in multicellular organisms, where different cells must maintain distinct functional identities

How can particular chromatin states survive through multiple cell divisions?

Epigenetic memory

- DNA provides accurate stability and heritability that is necessary for memory on evolutionary timescales, whereas epigenetic memory provides an inherently reversible and plastic memory that works on shorter timescales
- short-term memory that involves transient signals setting the cell into one of at least two alternative regulatory states
- each cell division is a great challenge for the stability, due to major chromosomal perturbations

Proposed models for the inheritance of histone modifications

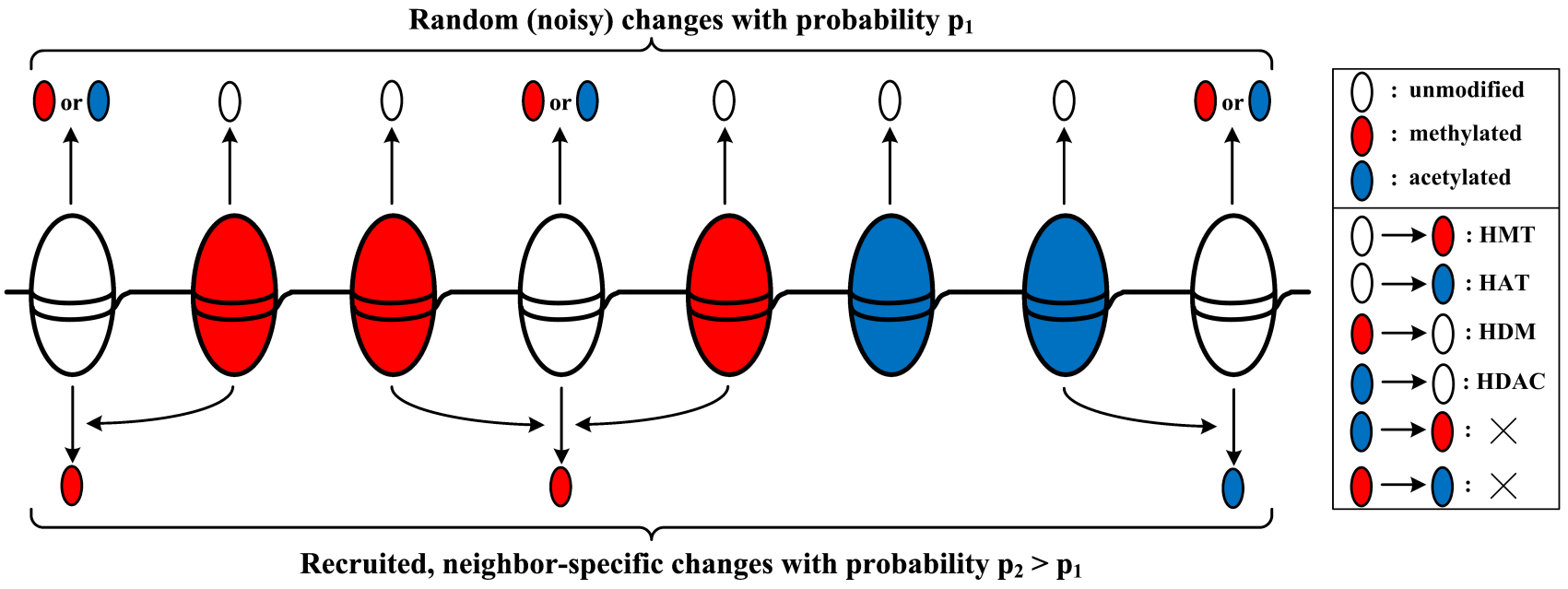
- often based on positive feedback loops in nucleosome modification (e.g., Dodd et al. 2007[1] or David-Rus et al. 2009[2])
- For some modifications, positive feedback loops have indeed been identified (e.g., H3K9me3 and H3K27me3).

---

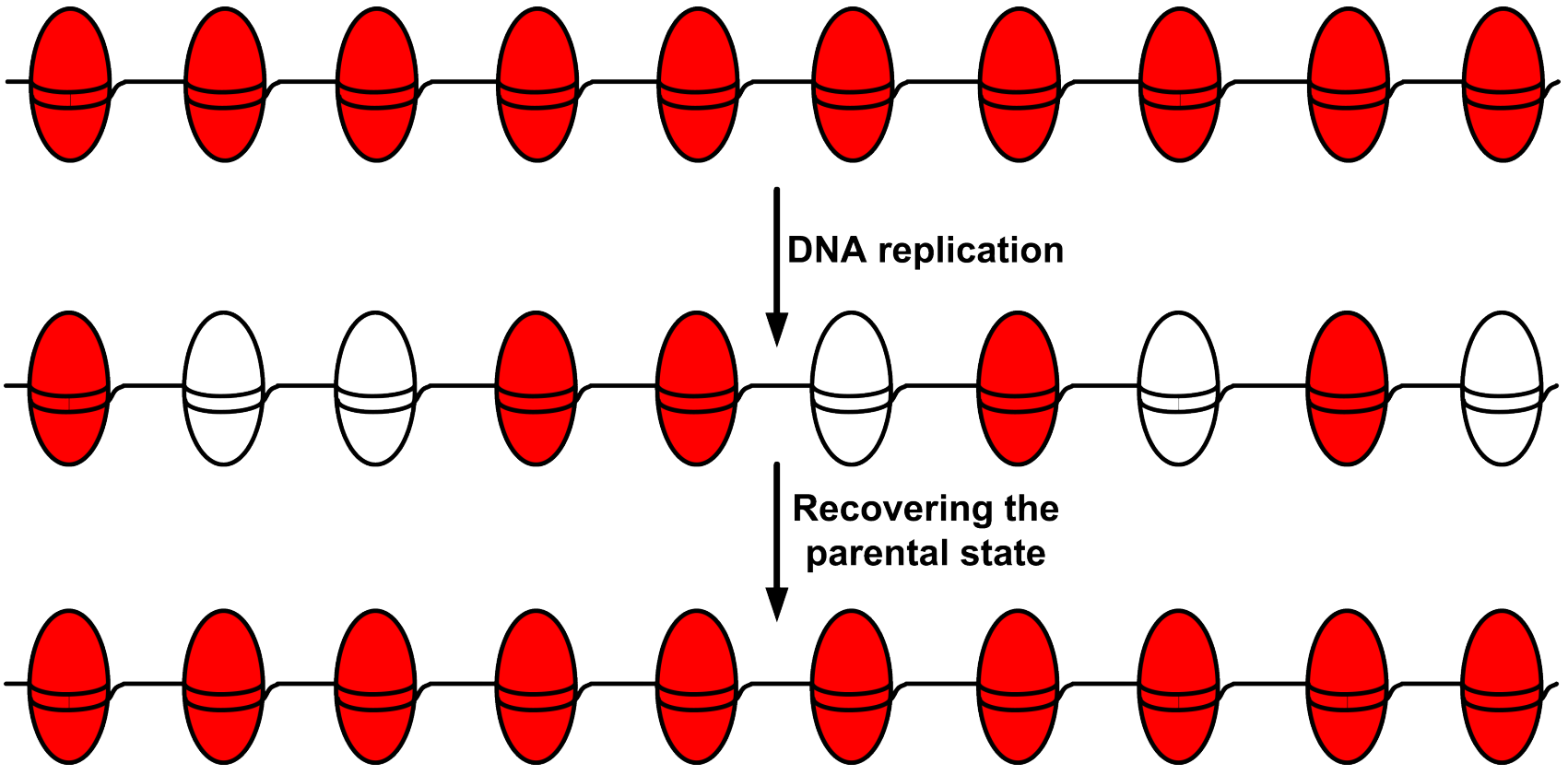
[1] Dodd, I.B, et al. 2007. Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell* **129**:813-822

[2] David-Rus, D., et al. 2009. Inheritance of epigenetic chromatin silencing. *Journal of Theoretical Biology* **258**:112-120

Graphical representation of models that are based on positive feedback loops



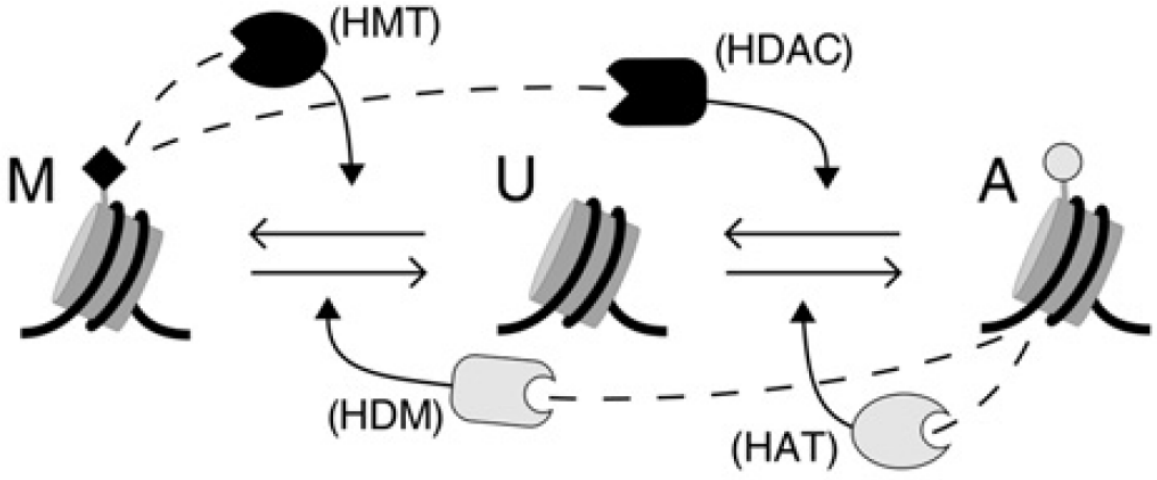
Graphical representation of models that are based on positive feedback loops



2. Modeling the Inheritance of Histone Modifications

The model from Dodd et al. 2007 / David-Rus et al. 2009

1. DNA region with L nucleosomes
2. Each nucleosome at position i can be in one of multiple possible states s_i
3. Nucleosoms can be actively interconverted by histone writers and histone erasers (e.g., HMT,HDM,HAT,HDAC)
4. Changes are either random (noise) or recruited



Starting to model the system: The master equation

Definition: Master equation

The *master equation* is a phenomenological set of first-order differential equations describing the time evolution of the probability of a system to occupy each one of a discrete set of states. In probability theory, this identifies the evolution as a continuous-time Markov process.

$$\frac{d}{dt}P[s_1, ..., s_L; t] = \sum_{i=1}^{L}\sum_{s'}$$
$$R_{is_i s'}[s_1, ..., s_{i-1}, s', s_{i+1}, ..., s_L]P[s_1, ..., s_{i-1}, s', s_{i+1}, ..., s_L; t] -$$
$$R_{is' s_i}[s_1, ..., s_{i-1}, s_i, s_{i+1}, ..., s_L]P[s_1, ..., s_{i-1}, s_i, s_{i+1}, ..., s_L; t]$$

The next step: Including DNA replication in the model

DNA replication

$$P[s_1, ..., s_L; nT+] = \sum_{s'_1,...,s'_L} \prod_{i=1}^{L} \left(\frac{1}{2}\delta_{s_i,s'_i} + \frac{1}{2}p_{s'_i}\right) P[s'_1, ..., s'_L; nT-]$$

- $nT-$: time just before the n-th round of DNA replication
- $nT+$: time just after the n-th round of DNA replication
- p_s : probability of a newly synthesized nucleosome to be in state s
- δ_{s_i,s'_i} : probability of the parental nucleosome state to be retained

Solving the master equation analytically for a long time behavior is generally an impossible task

Definition: Mean-field theory (Molekularfeldtheorie)

Mean-field theory (MFT) is an approximative method that replaces all interactions to other individual particles with an average or effective interaction. This reduces any multi-body problem into an effective one-body problem. This avoids summing over all states, and the ease of solving MFT problems means that some insight into the behavior of the system can be obtained at a relatively low cost.

What does this mean in the context of our model?

- ignore spatial variation of marks and replace them by average concentrations
- focus on regions with the same epigenetic fate
- replace master equation and the equation for DNA replication by its mean-field theory equivalents

Starting simple: The two-state model

- nucleosomal state: presence (A) or absence of a mark (U) for a string of L nucleosomes
- $a(t)$: mean-field probability for acetylation
- Basal rates:
  - $n_{A\to U}$: rate constant for the deacetylation of an acetylated mark
  - $n_{U\to A}$: rate constant for the acetylation of an deacetylated mark
- Recruitment rates:
  - r_{HAT}: rate constant for the recruitment of *HAT*s by acetylated marks
  - r_{HDAC}: rate constant for the recruitment of *HDAC*s by deacetylated marks

Starting simple: The two-state model

Kinetic equation for the concentrations of acetylation

$$\frac{da(t)}{dt} =$$

$$\underbrace{(1 - a(t))(n_{U\to A} + r_{HAT}\ a(t))}_{\text{Acetylation gain}} - \underbrace{a(t)(n_{A\to U} + r_{HDAC}(1 - a(t)))}_{\text{Acetylation loss}}$$

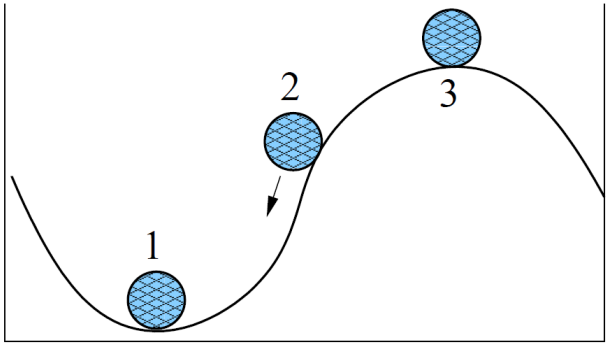
This model has only one steady state $a = 1$ for a vanishing rate of degradation $n_{A\to U}$.

Steady states / Fixed points

Definition: Steady state / Fixed point

In chemistry, a steady state is a situation in which all state variables are constant in spite of ongoing processes that strive to change them. For an entire system to be at steady state, i.e. for all state variables of a system to be constant, there must be a flow through the system.

Steady states can be either *stable* or *unstable*.



The two-state model fails to produce bistability

- no bistability, even in the absence of DNA replication (only one steady state)
- in a two-states model, direct cooperativity of histone modifications is necessary to attain bistability

This system is thus too simple to model epigenetic inheritance.

Cooperativity is necessary to attain bistability

Definition: Cooperativity

Cooperativity is a phenomenon displayed by enzymes or receptors that have multiple binding sites where the affinity of the binding sites for a ligand is increased (positive cooperativity) or decreased (negative cooperativity) upon the binding of a ligand to a binding site.

Cooperativity is one possibility of introducing non-linearity in the feedback loop, which is a requirement for bistability in a system (only positive feedback loops are not sufficient)

Two-state model with explicit cooperativity

Kinetic equations for the concentrations of acetylation with cooperativity

$$\frac{da(t)}{dt} =$$

$$\underbrace{(1-a(t))(n_{U\to A} + r_{HAT}\ a^{\mathbf{n}}(t))}_{\text{Acetylation gain}} - \underbrace{a(t)(n_{A\to U} + r_{HDAC}(1-a(t))^{\mathbf{m}})}_{\text{Acetylation loss}}$$

- n and m: degree of cooperative acetylation and deacetylation, resp.
- n and m introduce non-linearities in the feedback loop

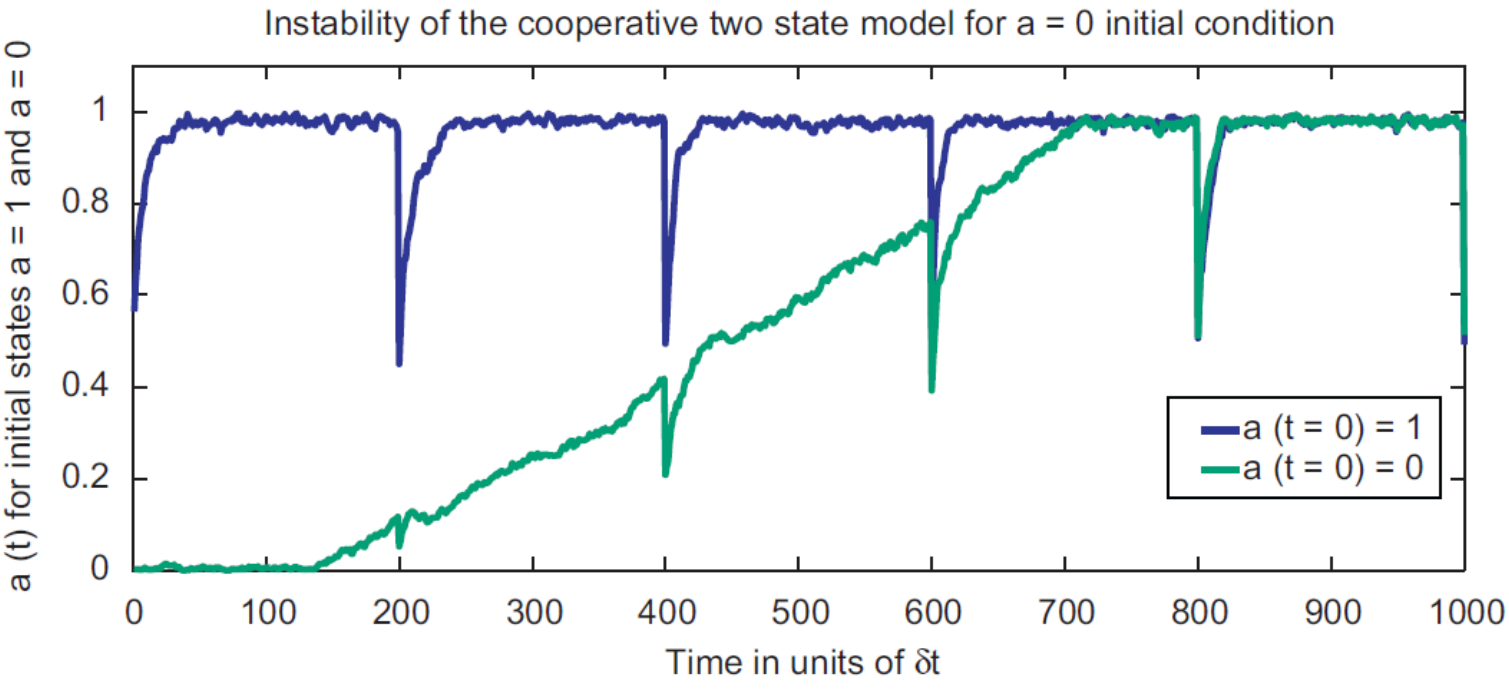
Two-state model with explicit cooperativity

For the simplest case $n = m = 2$ (and negligible basal rates $n_{U \to A}$ and $n_{A \to U}$), the steady states for this system are as follows:

1. $a = 1$
2. $a = 0$
3. $(a = \frac{r_{HDAC}}{r_{HAT} + r_{HDAC}})$

Now, the system has two stable steady states $\rightarrow$ Memory

Chemical noise can quickly change the dynamics of the system



Hypothesis: The presence of multiple epigenetic marks is a design criterion for epigenetic stability.

The three-state model

- $a(t), m(t), u(t)$: mean-field probability for acetylated / methylated / unmodified nucleosomes
- Recruitment rates:
  - r_{HAT}: rate constant for the recruitment of *HAT*s by acetylated marks
  - r_{HDAC}: rate constant for the recruitment of *HDAC*s by deacetylated marks
  - r_{HMT}: rate constant for the recruitment of *HMT*s by methylated marks
  - r_{HDM}: rate constant for the recruitment of *HDM*s by demethylated marks
- Basal rates for conversion and degradation (noise):
  $n_{U\to M}, n_{M\to U}, n_{U\to A}, n_{A\to U}$

The three-state model

Kinetic equations for the concentrations of methylation and acetylation

$$\frac{dm(t)}{dt} = r_{HMT}\ u(t)\ m(t) - \ r_{HDM}\ m(t)\ a(t) + n_{U\to M}u(t) - n_{M\to U}m(t)$$

$$\frac{da(t)}{dt} = r_{HAT}\ u(t)\ \ a(t) - r_{HDAC}\ a(t)\ m(t) + n_{U\to A}u(t) - n_{A\to U}a(t)$$

Steady states for the three-state model

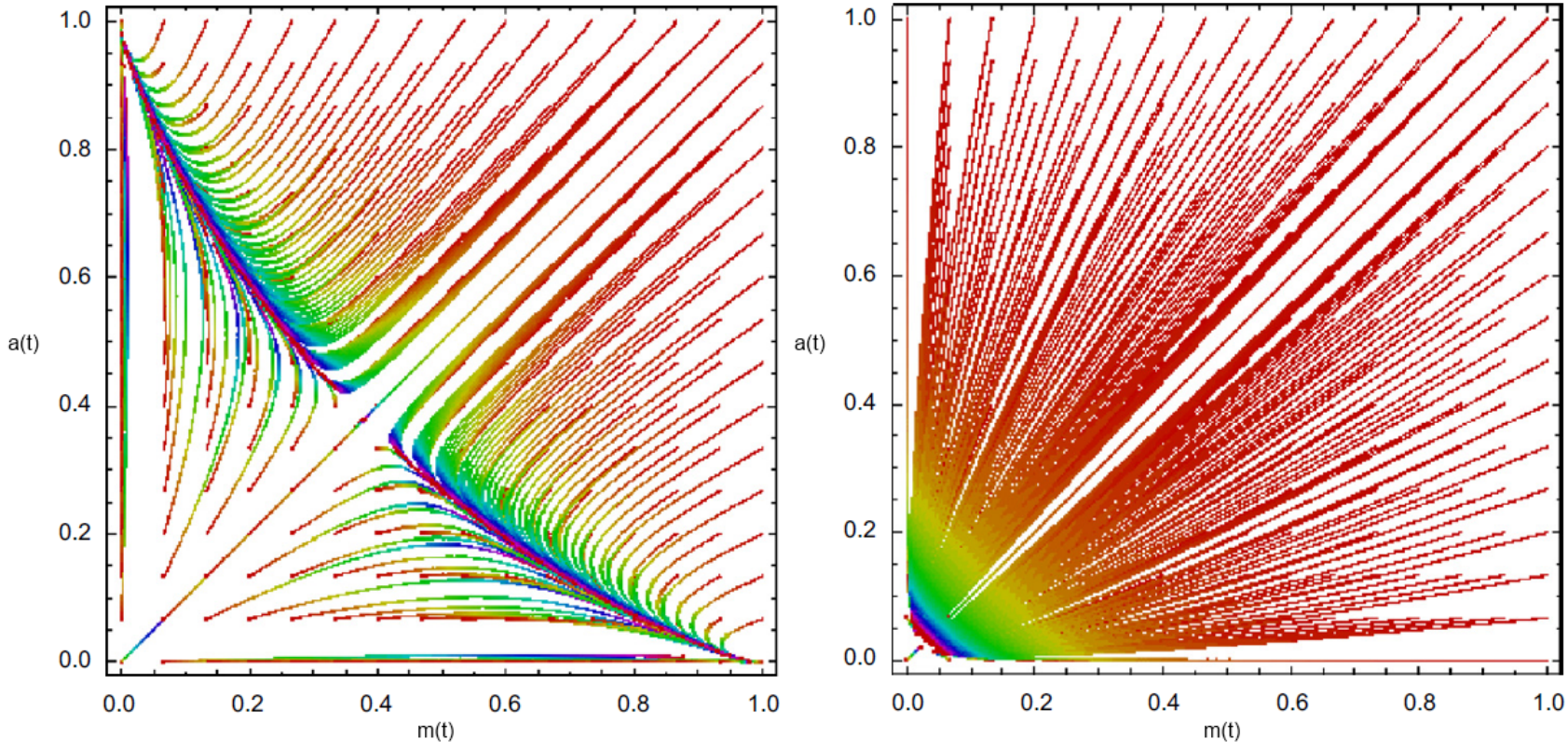
Steady states (if basal rates are small):

1. $a = 1,\ m = 0$
2. $a = 0,\ m = 1$
3. $(a = 0,\ m = 0)$
4. $(a = \frac{r_{HMT}\, r_{HDAC}}{r_{HAT}\, r_{HDM} + r_{HDAC}(r_{HMT} + r_{HDM})},$

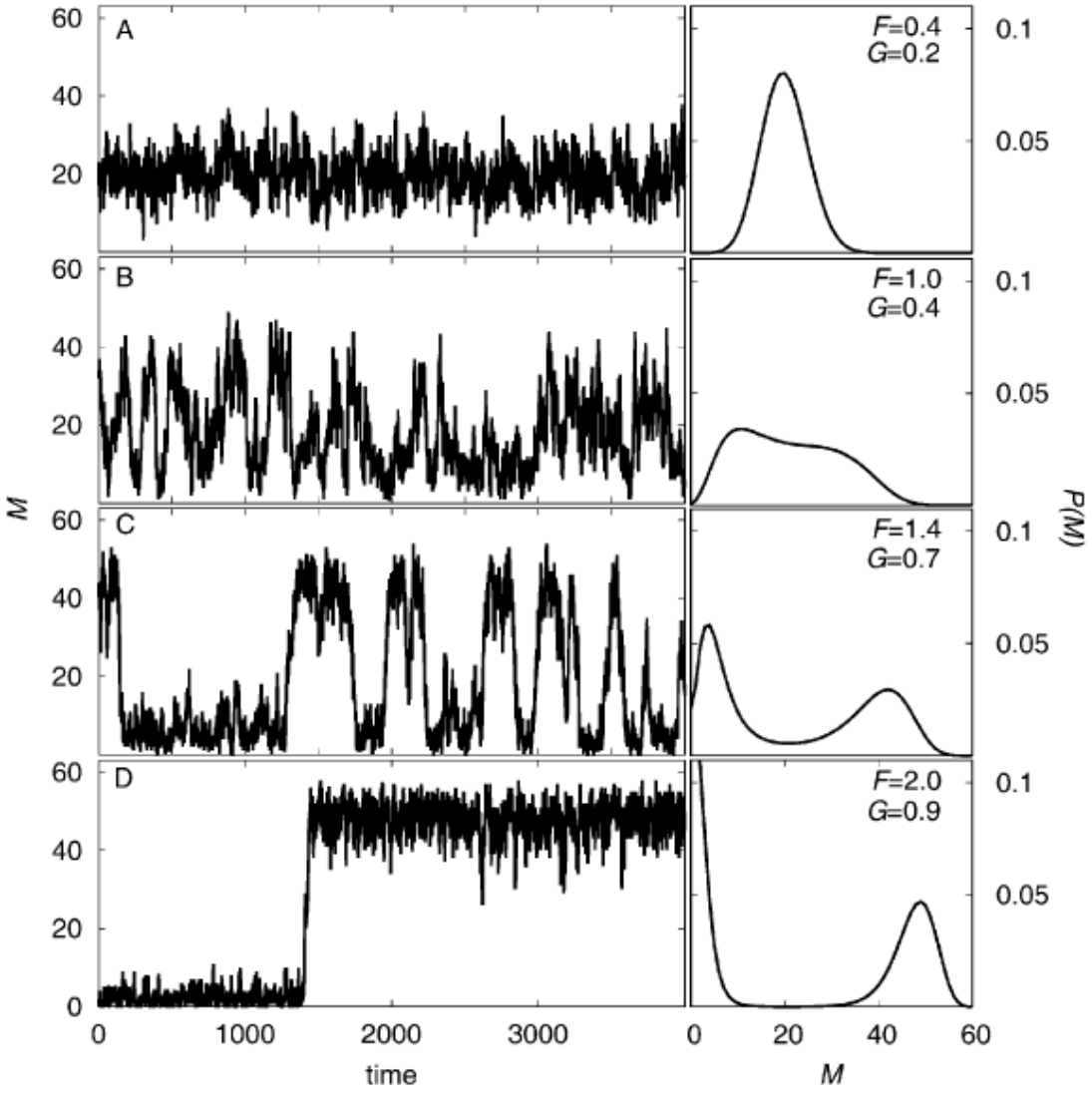
   $m = \frac{r_{HAT}\, r_{HDM}}{r_{HAT}\, r_{HDAC} + r_{HDAC}(r_{HMT} + r_{HDM})})$

The system has also two stable steady states, even though now, cooperativity has not been modeled explicitly!

Phase flow for low (left) and high (right) basal rates

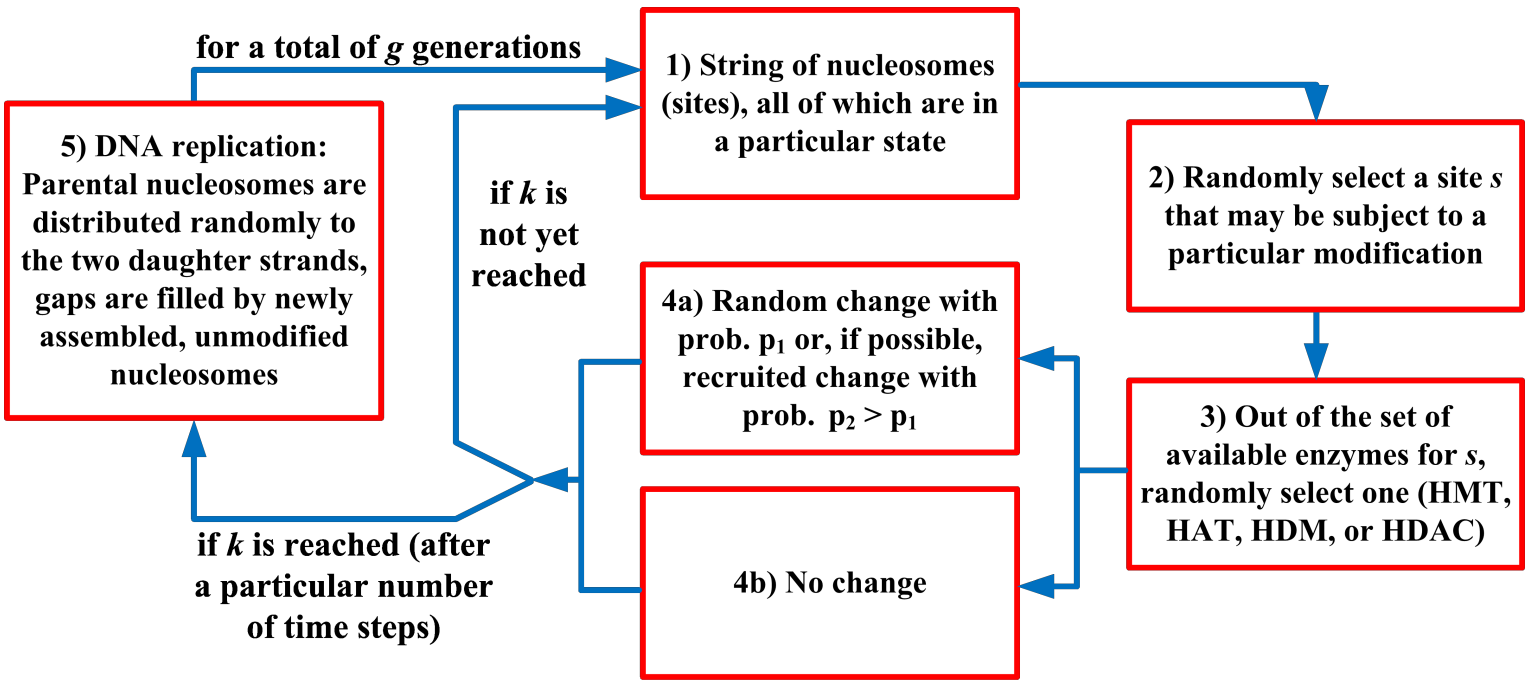


The feedback-to-noise ratio F has a crucial impact on the behavior of the system



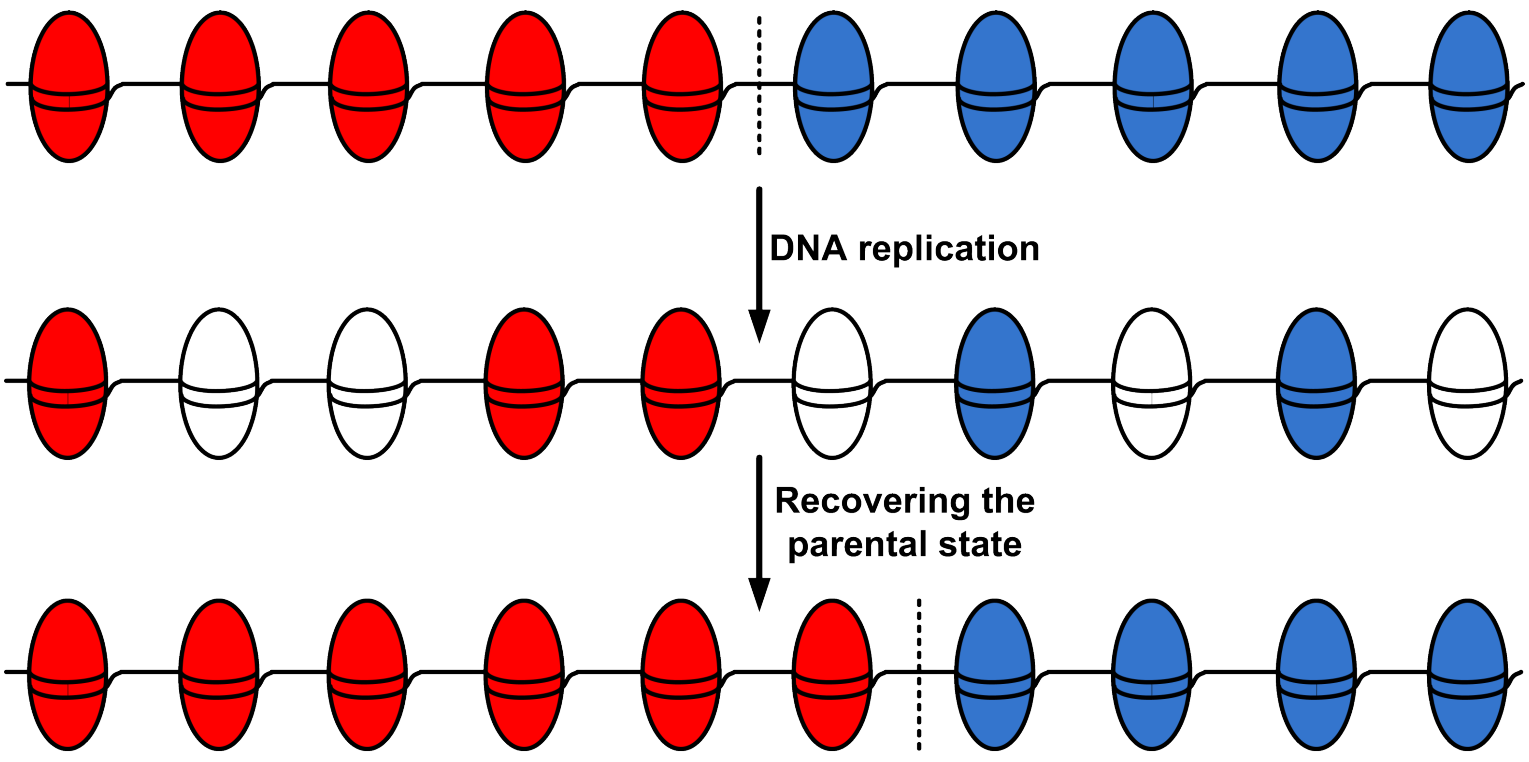
3. Outlook and Conclusions

Research in our lab: Modeling epigenetic inheritance using computational simulations

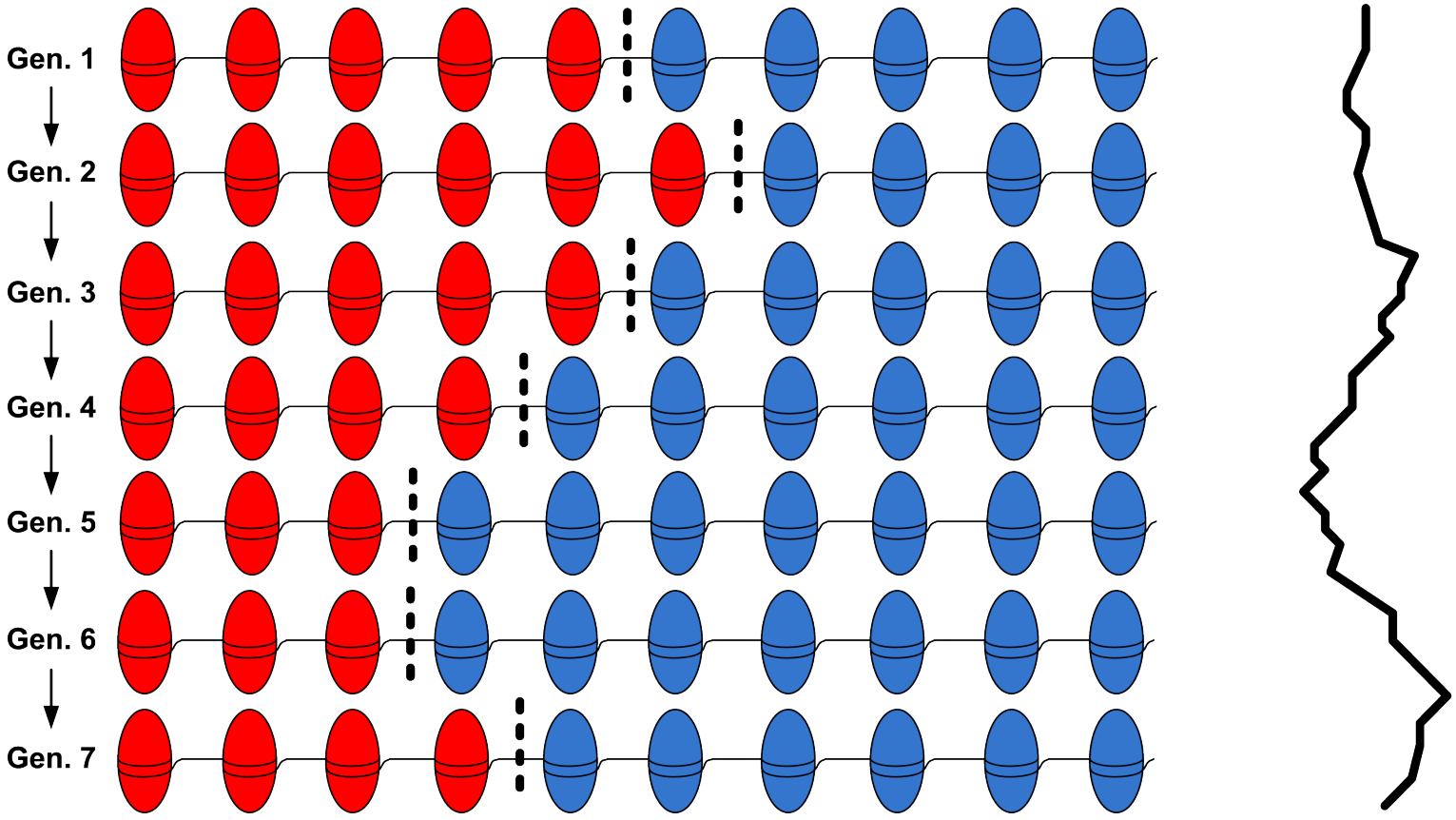


k = average number of modifications per site

Research in our lab: Modeling epigenetic inheritance using computational simulations



Research in our lab: Modeling epigenetic inheritance using computational simulations



Overall complexity of epigenetic inheritance far from being completely understood

- histone variants, histone repositioning, different variants of a particular modification
- DNA dependence (sequence specificity and methylation status)
- nucleosome positioning and turnover, chromosomal rearrangements, ncRNA-related mechanisms
- histone code not simply binary
- available enzymes and rules often specific to a particular cell type, developmental stage, species, or even genomic loci

Overall complexity of epigenetic inheritance far from being completely understood

- abundance of the histone modifying enzymes and their individual reaction rates are unknown
- Increasing evidence suggests that histone modifications cannot be stably inherited in absence of the initial trigger $\rightarrow$ Other factors are mandatory to maintain an epigenetic memory (e.g., DNA methylation).
- ...

How well do the current models reflect biology?

Three take-home messages

1. Chromatin can be seen as a powerful computational device capable of storing, processing and propagating information in a context-sensitive and massively parallel manner.

Three take-home messages

1. Chromatin can be seen as a powerful computational device capable of storing, processing and propagating information in a context-sensitive and massively parallel manner.
2. Models of epigenetic inheritance predict that some kind of cooperativity is necessary to achieve bistability. Bistable models can even withstand the major perturbations through DNA replication. The parental state can be recomputed using a set of predefined rules.

Three take-home messages

1. Chromatin can be seen as a powerful computational device capable of storing, processing and propagating information in a context-sensitive and massively parallel manner.
2. Models of epigenetic inheritance predict that some kind of cooperativity is necessary to achieve bistability. Bistable models can even withstand the major perturbations through DNA replication. The parental state can be recomputed using a set of predefined rules.
3. Simple models of epigenetic inheritance may be analyzed analytically, and they give insights to some basic properties that epigenetic inheritance systems have to exhibit. More complex models can often only be solved numerically.

Fate?

Thank you for your attention!