Chromatin Computation

Christian Arnold

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

February 18th, 2010

Table of Content

Motivation

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)

"The genetic code is the piano,
the epigenetic code the tune."

Epigenetic Phenomenon

- Cell differentiation
- Gene regulation
- Genomic imprinting
- Transgenerational inheritance
- Much more yet to be discovered

The Epigenetic Code

- Genetic code - four-base alphabet, nearly invariant and universal
- Is there a similarly universal epigenetic code?

The Epigenetic Code

- Genetic code - four-base alphabet, nearly invariant and universal
- Is there a similarly universal epigenetic code?
    - most likely not, because the patterns vary too much among organisms and are dynamic in space and time
    - the epigenetic code defines a process involving the recruitment of a myriad of chromatin-remodeling complexes, insulator proteins, histone exchange chaperones, enzymes, coregulators, and effectors

Epigenetics - A BRIEF overview
Chromatin and Cellular Memory
Chromatin Model
Conclusions and Perspectives

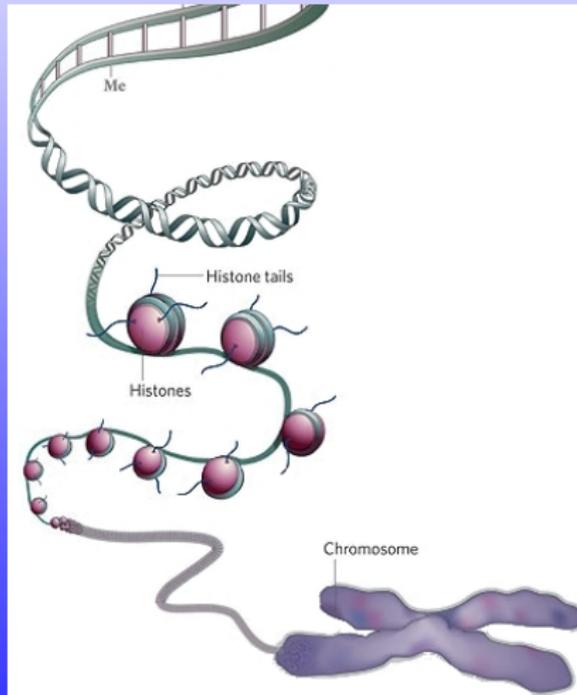Chromatin Structure
Chromatin Regulation

Chromatin

- Chromatin is the complex combination of DNA and proteins that makes up chromosomes
- Functions
  - package DNA into a smaller volume
  - mechanism to control gene regulation and DNA replication
- Heritable changes in chromatin structure through
  - DNA methylation
  - Posttranslational histone modifications
  $\Rightarrow$ Cellular memory

Epigenetics - A BRIEF overview
Chromatin and Cellular Memory
Chromatin Model
Conclusions and Perspectives

Chromatin Structure
Chromatin Regulation

Histones and Nucleosomes

- Histones are the fundamental repeating subunit of chromatin consisting of
    - highly conserved core histones composed of a globular domain and flexible histone tails protruding from the surface
    - linker histones
- Two of each of the 4 core histones assemble to form one octameric nucleosome core particle by wrapping DNA around the protein spool
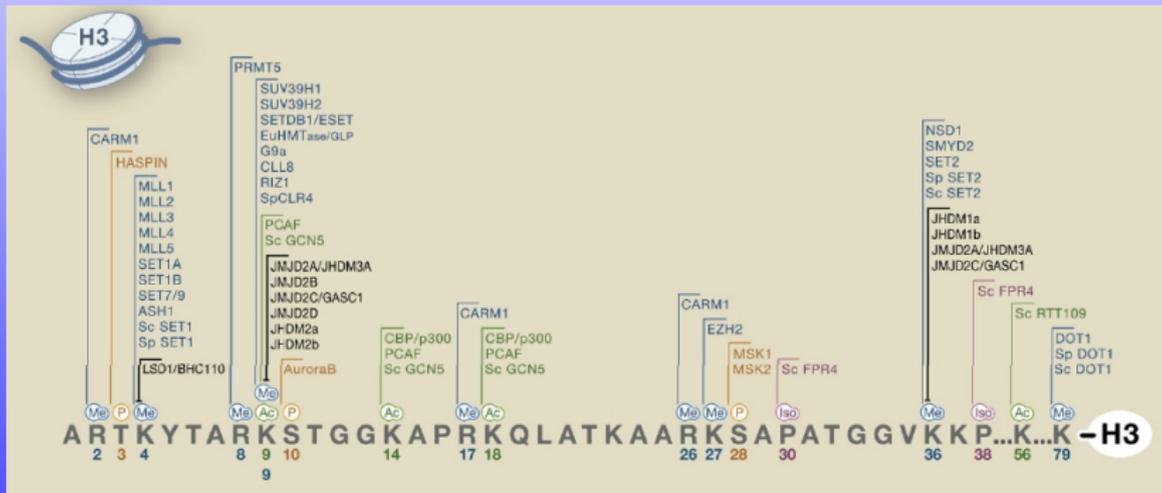
Epigenetics - A BRIEF overview
Chromatin and Cellular Memory
Chromatin Model
Conclusions and Perspectives

Chromatin Structure
Chromatin Regulation

Chromatin Structure - A Hierarchy of Organisation

Epigenetics - A BRIEF overview
Chromatin and Cellular Memory
Chromatin Model
Conclusions and Perspectives

Chromatin Structure
Chromatin Regulation

Posttranslational Histone Modifications

- chemical modifications of histones that affect transcription
- especially the residues on the histone tails are subject to those modifications
  - Phosphorylation
  - Acetylation
  - Methylation
  - Ubiquitylation and sumoylation
  - ADP-ribosylation, biotinylation, proline isomeration, and others
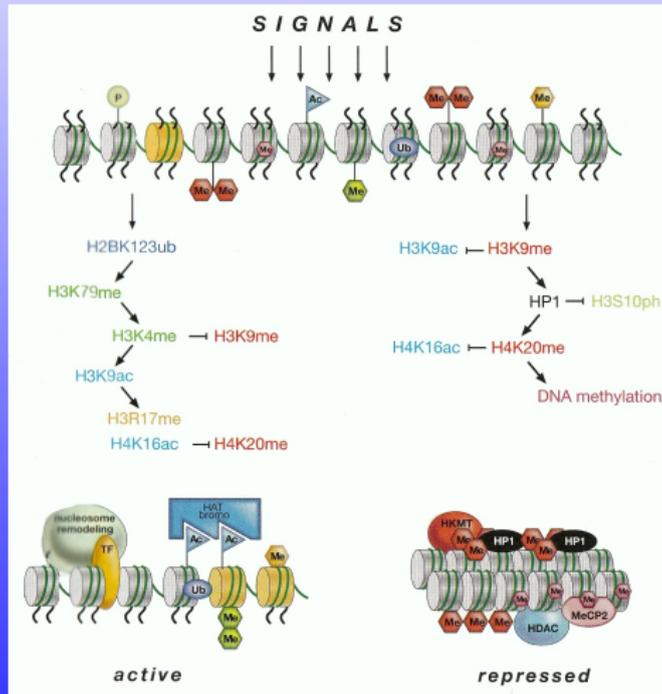
Epigenetics - A BRIEF overview
Chromatin and Cellular Memory
Chromatin Model
Conclusions and Perspectives

Chromatin Structure
Chromatin Regulation

Example: Histone H3 modifications

Epigenetics - A BRIEF overview
Chromatin and Cellular Memory
Chromatin Model
Conclusions and Perspectives

Chromatin Structure
Chromatin Regulation

Chromatin Regulation

- Regulation of chromatin structure formation and maintenance
- Part of both transcriptional regulation and epigenetic regulation
- Modifications can be read, written, or deleted by various proteins
- Information can be stored and propagated along the chromosome
  $\Rightarrow$ Chromatin is a cellular memory device

Epigenetics - A BRIEF overview
Chromatin and Cellular Memory
Chromatin Model
Conclusions and Perspectives

Chromatin Structure
Chromatin Regulation

Example

Epigenetics - A BRIEF overview
Chromatin and Cellular Memory
Chromatin Model
Conclusions and Perspectives

Goals
Model Details
Challenges
Current Progress

Questions

- How efficient, reliable, and redundant is the information propagation process?

- How is chromatin computation conceptually different from transcription factor based regulatory networks?

- Under which conditions can one get distinct, genome-wide chromatin states? How many states are possible?

- What are potential models for stable replication of epigenomic information?

- Given the set and interactions of readers, writers and erasers, how much information can be stored epigenetically?

- What is the role of sites that link the genomic and epigenomic information?

Epigenetics - A BRIEF overview
Chromatin and Cellular Memory
Chromatin Model
Conclusions and Perspectives

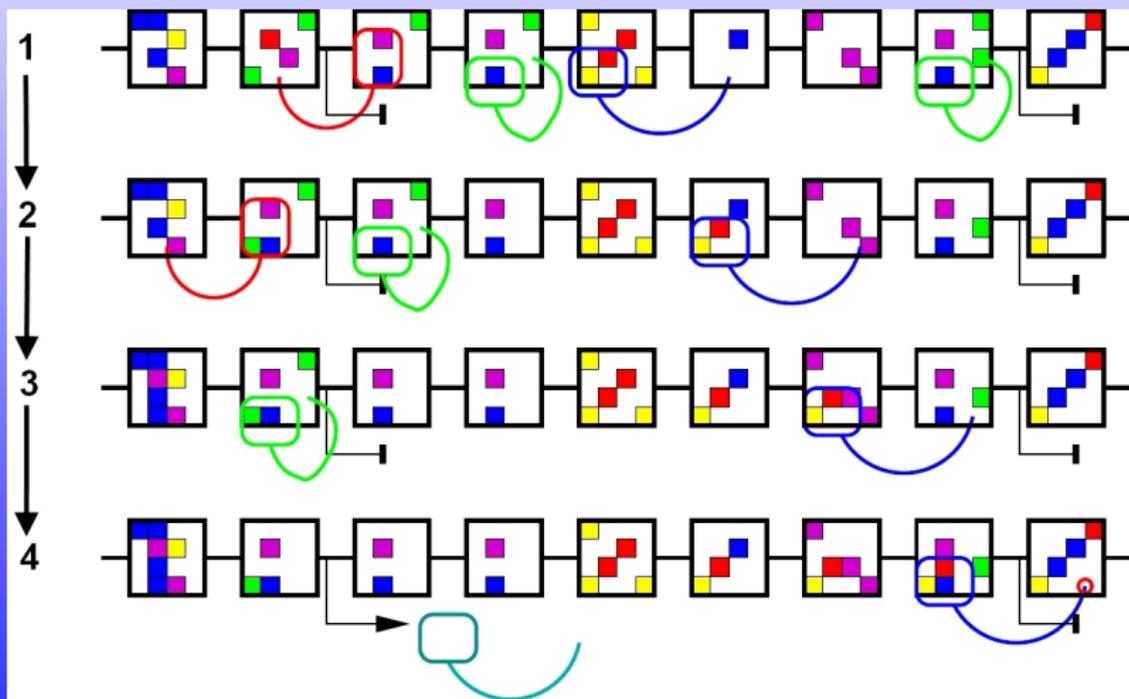Goals
Model Details
Challenges
Current Progress

Model Details

- Model independent of underlying DNA for now
- Memory is organized in a linear sequence of pages (here: eukaryotic nucleosomes)
  - each nucleosome has a fixed number of possible modification sites from all eight histones
  - at most one modification at each modification site (phosphorylation, acetylation, methylation, ubiquitylation/sumoylation)

Epigenetics - A BRIEF overview
Chromatin and Cellular Memory
Chromatin Model
Conclusions and Perspectives

Goals
Model Details
Challenges
Current Progress

Model Details

- State transitions are determined by rules that can read the information stored on a particular position and change its content of the same or an adjacent memory page
- Information propagation
  - to the nucleosome itself
  - to one of the two neighbor nucleosomes
  - to a non-adjacent nucleosome
- DNA methylation not yet included

Epigenetics - A BRIEF overview
Chromatin and Cellular Memory
Chromatin Model
Conclusions and Perspectives

Goals
Model Details
Challenges
Current Progress

Graphical Representation

Epigenetics - A BRIEF overview
Chromatin and Cellular Memory
Chromatin Model
Conclusions and Perspectives

Goals
Model Details
Challenges
Current Progress

Challenges

Overall complexity far from being completely understood

Epigenetics - A BRIEF overview
Chromatin and Cellular Memory
Chromatin Model
Conclusions and Perspectives

Goals
Model Details
Challenges
Current Progress

Challenges

Overall complexity far from being completely understood

- Histone code is not simply binary
- Histone variants, histone replacement, histone repositioning, different subtypes of a particular modification
- Available enzymes and rules often specific to a particular lineage, species, or tissue
- RNA-based mechanisms of regulation
    - mRNA-like ncRNAs(mlonRNAs or mlncRNAs) such as HOTAIR, AIR, XIST or TSIX
    - RNA interference
- ...

Epigenetics - A BRIEF overview
Chromatin and Cellular Memory
Chromatin Model
Conclusions and Perspectives

Goals
Model Details
Challenges
Current Progress

Current Progress

- Gathering information from literature
  - state transition rules and conditions
  - universal vs. lineage/species-specific rules
  - meaningful parameter values
  - ...
- Programming a flexible and expandable model

Conclusions and Perspectives

- We are aware of the complexity of these epigenetic mechanisms, but we are far away from making sense of that complexity

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

- Dynamics of the system only poorly understood

Thanks!

Thanks to Sonja and Peter!

Thank you for your attention!
Any questions?