The Chromatin Story – Drawing Conclusions about Functional Evolution

Sonja J. Prohaska

Santa Fe Institute for Complex Systems, Santa Fe, NM
Bioinformatics Group, University of Leipzig, Germany
Institute for Theoretical Chemistry, Univeristy of Vienna

February 15, 2009

Regulation – eexcuted by regulators

- *"[Regulation is an] attempt to produce outcomes which might not otherwise occur, produce or prevent outcomes in different places to what might otherwise occur, or produce or prevent outcomes in different timescales than would otherwise occur"* by Wikipedia (english)
- *"Regler beeinflussen selbsttätig in einem meist technischen Prozess eine oder mehrere physikalische Größen auf ein vorgegebenes Niveau unter Reduzierung von Störeinflüssen."* by Wikipedia (german)
- *"In control theory, a controller is a device which monitors and affects the operational conditions of a given dynamical system. The operational conditions are typically referred to as output variables of the system which can be affected by adjusting certain input variables"* by Wikipedia (english)

Regulation in a cell

- monitor a condition (input)
- execute a condition-dependent effect (output)
- single regulator or regulatory cascade
- affect a process (dynamic)
- channel the process
  - restrict the process
  - modify input, output or parameters
  - may induce an outcome that would otherwise not occure

Regulation can't make the impossible possible only the improbable probable.

Under regulation or not?

A gene is regulated if

- its gene expression profile changes with changing conditions
- there is at least one gene whose deletion uncouples the genes expression pattern from a condition

A gene is unregulated if

- its expression pattern does not respond to changes that are sensed or induced by the underlying genome
- there is no gene delition that could alter the response to a changing condition

Concentrate on regulation by a genetic factor/regulator.

What are regulators good for?

- supress selfish DNA and parasites – defense
- reduction of disturbance and fluctuation – robustness
- fine-tuning of cellular processes – economic efficiency
- selective (de)activation of gene sets – adaptation to conditions fluctuating with a high frequency compared to live span of an individual. Good for species with long non-reproductive phases.

Multiple phenotypic states

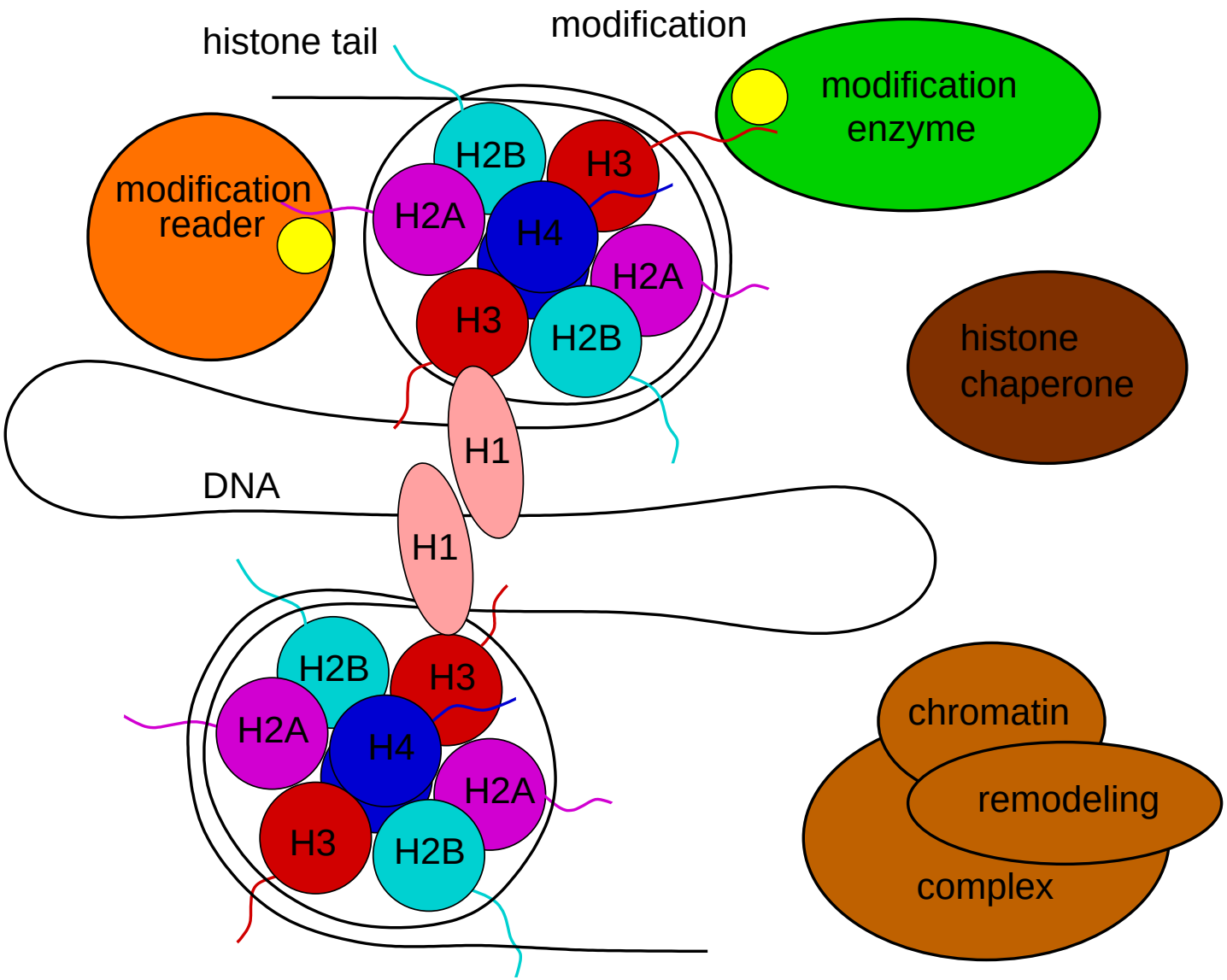
Cells and organisms exist in different phenotypic states: exponetial growth phase, stationary phase, with or without flagellum, spores, etc. The transition between states is necessarily regulated, e.g. differentiation, development.

Gene expression levels from large sets of genes are changed during the transition between states. Resulting in genomic states with certain gene expression profiles.

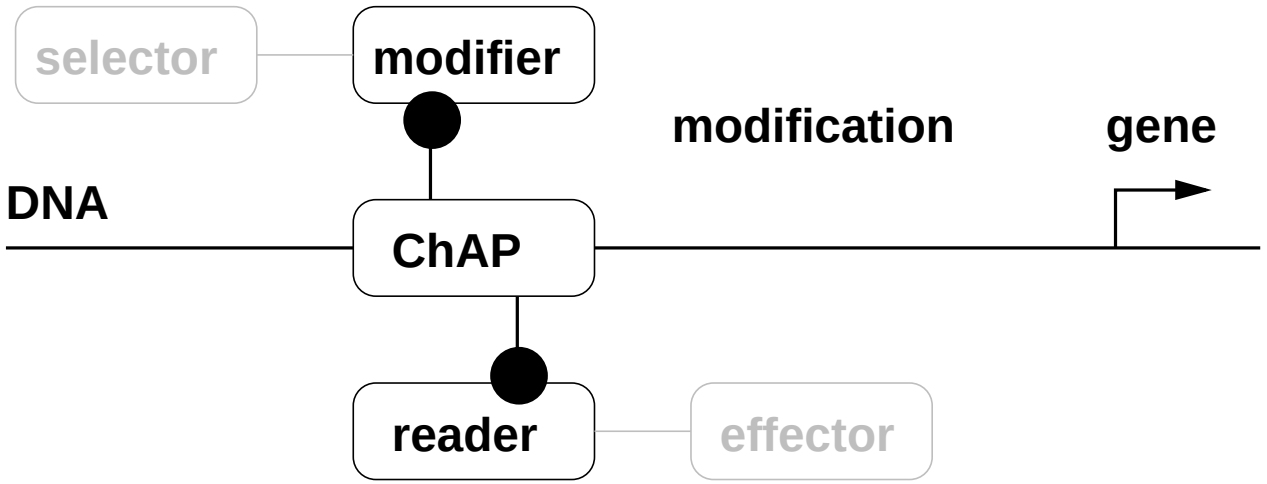
Chromatin regulation

- Chromatin regulation is part of transcriptional regulation. In general, it regulates promoter accessibility, Polymerase recruitment, transcription initiation and elongation and, therefore, the amount of transcripts generated.
- Chromatin regulation is part of epigentic regulation. It can store information about gene expression that can be transmitted to the next generation in an sequence-independent manner.
- Chromatin regulation also denotes the regulation of chromatin structure formation and maintenance.

Chromatin Regulation in Higher Eukaryots



Components of Chromatin Regulation



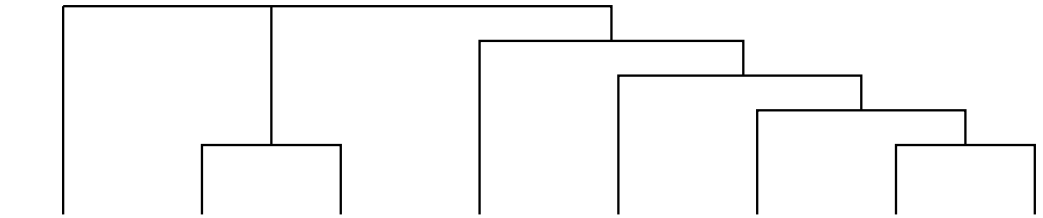
Find Functional Homologs of Chromatin Regulation

- functionally describe the essential components
- search for homologs in all complete genomes
- determine regulatory network homology
- all nodes and edges are homologous $\rightarrow$ full homology
- some nodes and edges are homologous $\rightarrow$ partial homology

Methods

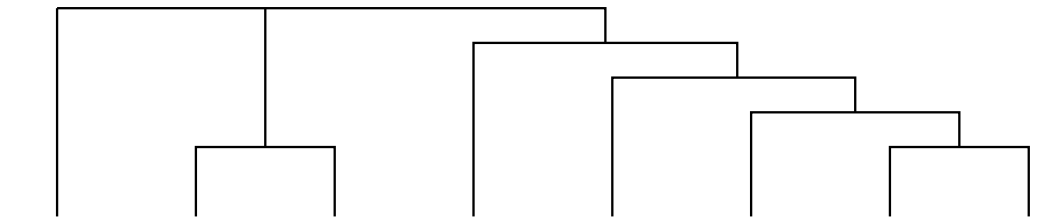
- functional domains from known chromatin regulators
- superfamily database and HMM domain annotation from 27. April 2008
- gene counts for each species and domain
- calculation of the proportion of species per clade with at least one gene
- calculation of average number of genes per species in a clade

Phylogentic Distribution of Modification Enzymes



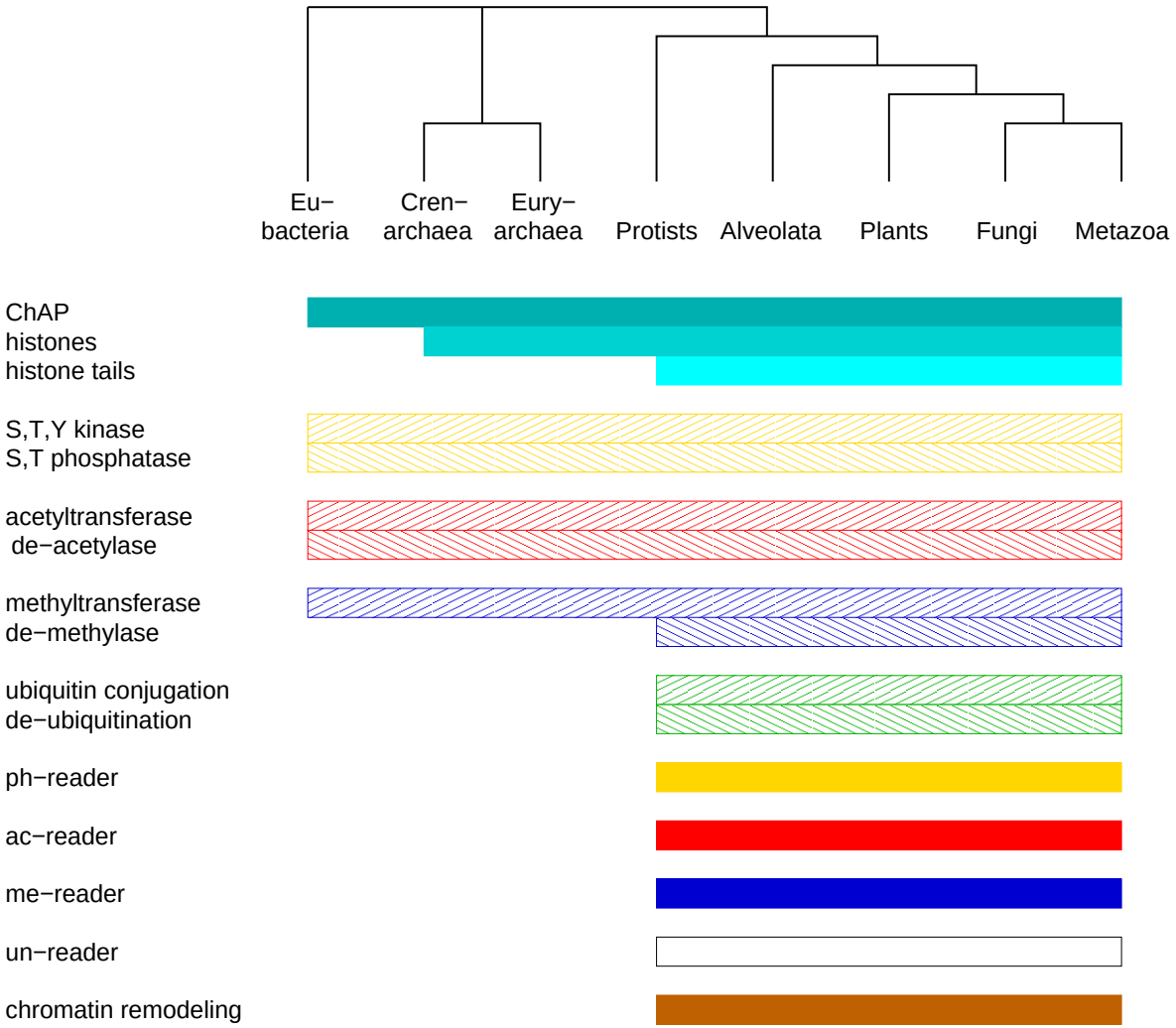
phosphorylation									
kinase	+ph(S,T,Y)	8.4	**4.0**	**2.9**	**373.7**	**352.6**	**825.1**	**147.0**	**606.7**
PP2C-like	-ph(S,T)	2.2	(1.0)	(2.2)	**19.3**	**24.4**	**58.6**	**8.2**	**20.6**
PP1-5	-ph(S,T)	1.9	1.3	(1.4)	**47.9**	**16.8**	**23.2**	**12.3**	**19.2**
acetylation									
NAT	+ac(K)	**22.3**	**7.8**	**12.6**	**27.2**	**13.1**	**45.5**	**35.9**	**38.2**
Sir2	-ac(K)	1.6	–	(1.1)	**6.8**	**2.8**	**3.4**	**5.5**	**7.8**
Rpd3	-ac(K)	1.7	–	–	**4.4**	**4.8**	**12.1**	**4.8**	**13.1**
methylation									
PRMT	+me(R)	2.0	2.0	2.1	**9.2**	**7.0**	**22.1**	**6.6**	**20.8**
SET	+me(K)	–	–	–	**14.6**	**6.8**	**29.1**	**7.9**	**33.6**
Dot1	+me(K)	–	–	–	2.7	–	2.1	1.1	1.8
JmjC	-me(R,K)	–	–	–	6.3	2.8	**14.2**	**4.4**	**19.8**
ubiquitination									
E1	+ub	+	–	–	8.1	**2.9**	**12.9**	**5.3**	**16.4**
E2	+ub	–	–	–	**24.2**	**21.1**	**38.1**	**19.0**	**43.9**
DUB	-ub	–	–	–	**3.6**	**5.7**	**86.6**	**3.5**	**8.8**
DCH	-ub	–	–	–	2.1	2.5	**3.1**	**3.0**	**4.9**
DCH-L3	-ub	–	–	–	**32.3**	**21.2**	**25.1**	**16.1**	**56.9**

Phylogentic Distribution of Modification Readers



phosphorylated									
14-3-3	ph	–	–	–	**5.3**	**4.2**	**8.6**	**2.2**	**8.3**
2×BRCT	ph	–	–	–	3.6	(2.2)	**4.0**	**2.0**	**9.2**
acetylated									
bromo	ac	–	–	–	**19.2**	**9.5**	18.1	**8.5**	**43.1**
methylated									
WD40 repeat	me	(2.8)	(1.0)	(2.5)	**200.7**	**164.8**	**223.3**	**122.1**	**313.4**
PHD	me	–	–	(1.0)	**4.6**	**10.5**	**40.4**	**11.6**	**65.5**
chromo	me	–	–	–	9.0	**6.1**	**8.5**	**14.8**	**22.1**
Tudor	me	–	–	–	1.0	1.5	**7.9**	1.5	**20.9**
MBT	me	–	–	–	–	–	–	–	**12.2**
un-modified									
Myb/SANT	–	–	–	–	**59.7**	**27.6**	**108.6**	**11.8**	**25.?**

The elements of chromatin regulation did not come into existance all at once.



Chromatin unites a architectural and regulatory role.

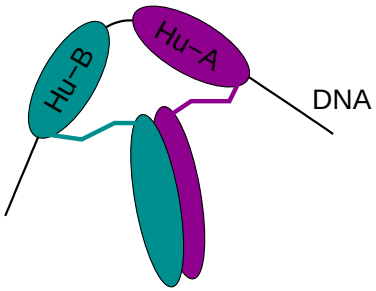
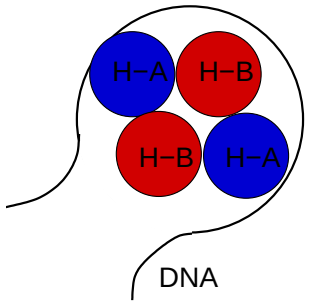
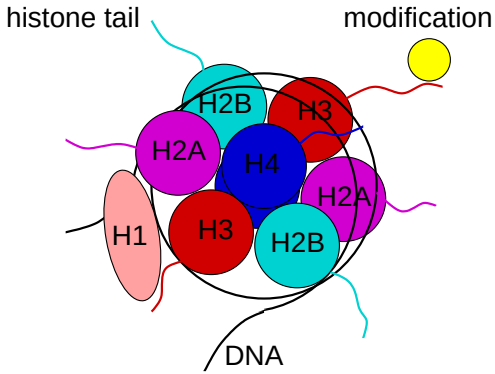
architectural role

- protect DNA from tangling and aggregation
- protect DNA from extrem conditions and damage
- organize the DNA into domains and structures
- reduce the contour length (compactation/packaging)

regulatory role

- change local DNA tension and supercoiling
- define states of gene expression on a genomewide scale
- set long-term non-expression
- reduce transcriptional noise and internal initiation

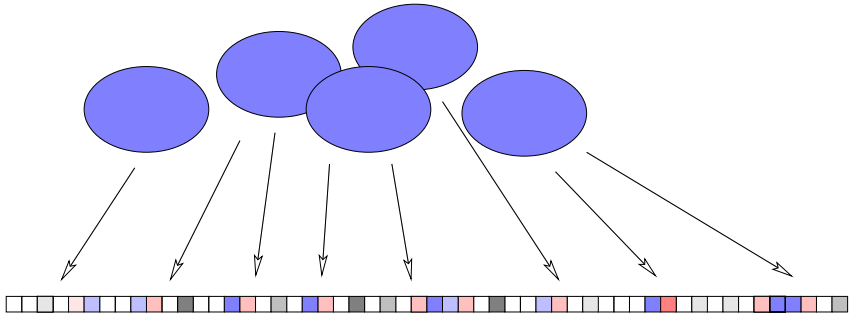
Bacteria, Archaea and Eukarya do have Chromatin.

euBacteria	**Archaea**	**Eukarya**
HU	HU, Alba, histone, a.o.	histone
	no histone tails	histone tails
HU-α, HU-β	H-α, H-β	H3, H4, H2A, H2B
dimers	tetramers	octamers
homo- (and hetero-)	(homo- and) hetero-	heterodimers
		
$-$ supercoiling	$+$ or $-$ supercoiling	$-$ supercoiling
contacts 9bp	contacts 60bp	contacts 146bp
not wrapped	$\sim$ 1 turn	$\sim$ 2 turns

ChAPs Have Similarities with General Transcription Factors.

specific regulator

general regulator



	specific	**general**
binding specificity	high	low
binding affinity	high	low
protein abundance	low	high
DNA coverage	very low	moderate/high
interaction	certain location	everywhere
interaction	certain time	all the time
regulated genes	few	many

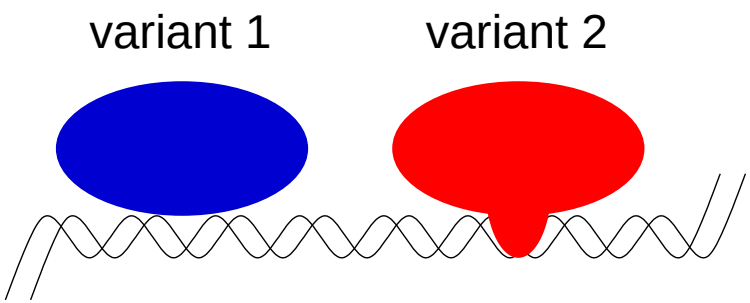
ChAP variants enable transition between genomic states.

- **Eukarya**
  H3: deposition at replication forks only
  H3.3: deposition throughout the cell cycle at euchromatin
  CENP-A: deposition at centromers
  MacroH2A: deposition at the inactive X chromosome
  H2A.Z: less stable, inhibits spreading of silent chromatin
- **Archaea**
  H-β causes greater compactation
  (H-α,H-α) homodimers: in exponential growth phase
  (H-α,H-β) heterodimers: in stationary phase
- **euBacteria**
  different cellular forms have different nucleoid structures
  H-NS is is a general regulator for enviromentally responsive genes and a ChAP

This is due to different biochemical properties, cooperative effects and minor differences in sequence preference.

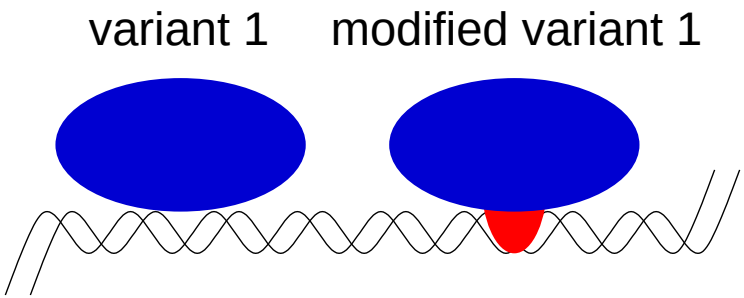
Variants from gene expression or modification are functionally equivalent.

variants



- different biochemical properties
- different conformation
- different interaction interfaces
- active and inactiv variants

modifications



- alter biochemical properties
- induce conformational changes
- alter interaction interfaces
- activate/inactivate proteins

Modifications

Modification enzymes in bacteria likely don't modify ChAPs.

- modify harmful substances (antibiotics)
- modify cellular targets of antibiotics
- only phosphorylation is extensively used for signalling
- phosphorylate His rather than Ser/Thr/Tyr
- no ubiquitin
- no modifications on ChAPs observed so far

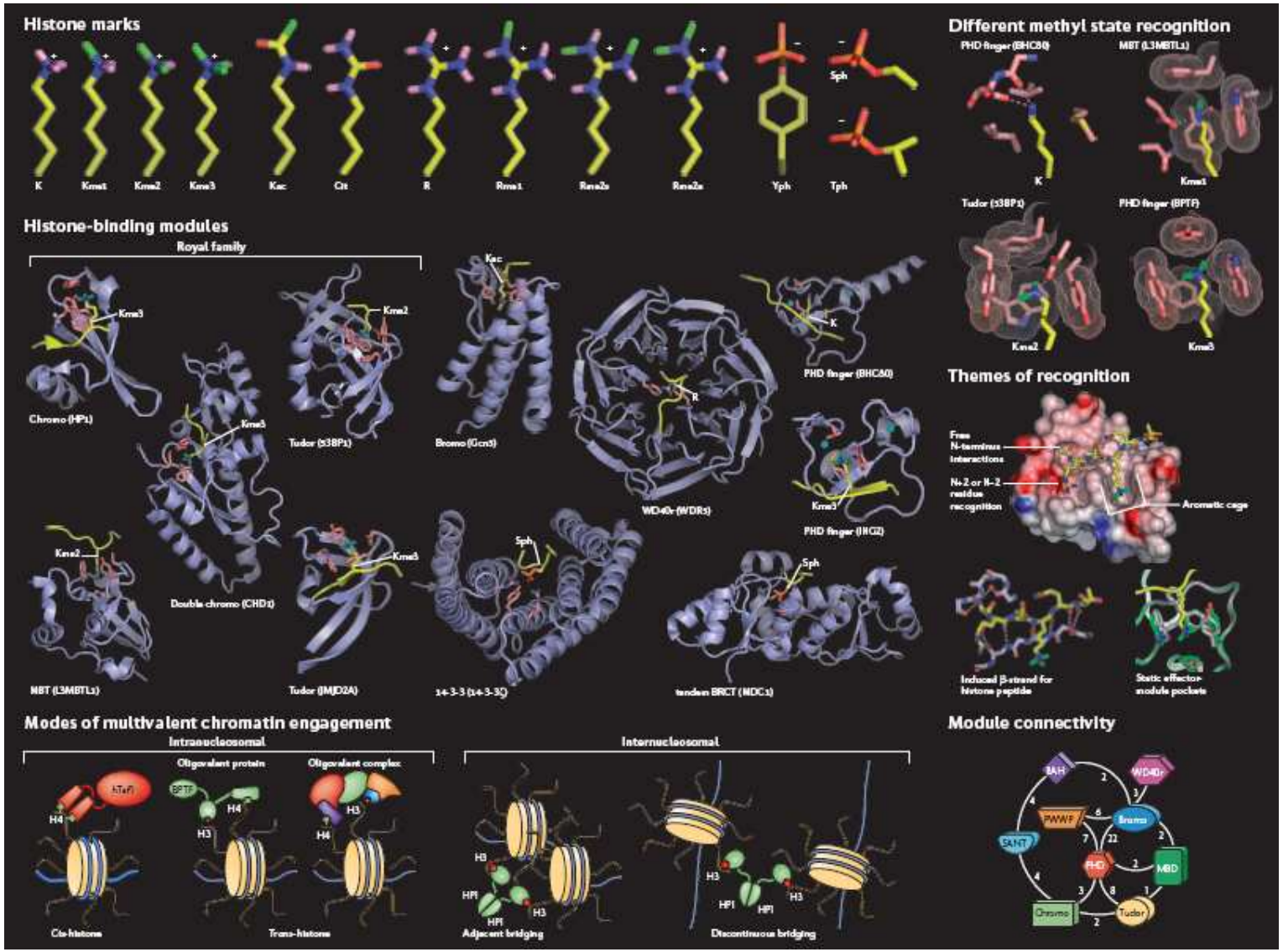
Archaea do modify their ChAPs but not the histones.

- archaeal histones do not have tails, the eucaryotic site of modification
- Alba (a ChAP) is acetylated and de-acetylated
- acetylated Alba has a reduced binding affinity
- de-acetylation causes transcriptional repression
- (de-)acetylation is coupled to the cellular NAD level

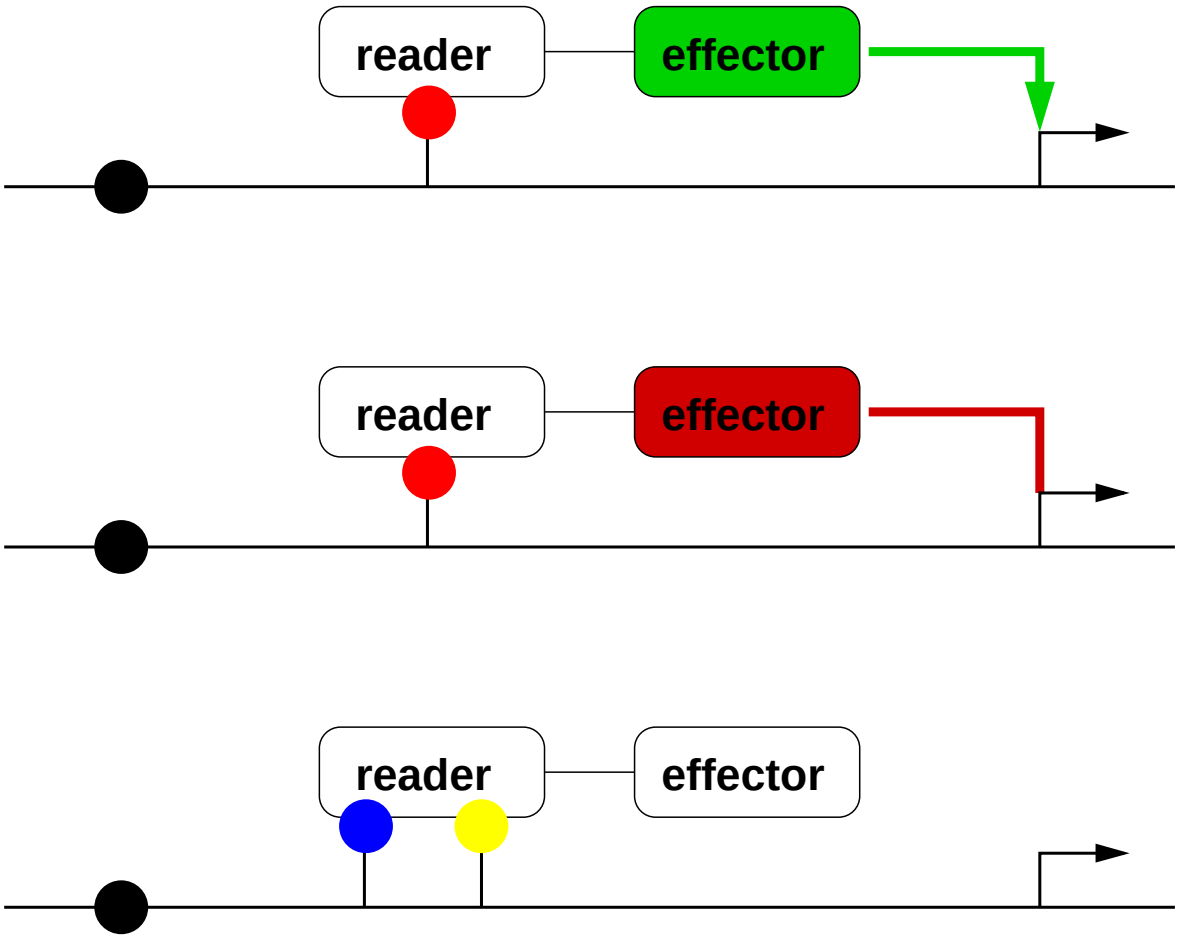
Eukaryotic chromatin is modified and stores and transmits information.

- many modification encymes in eukarya
  (e.g. 150-600 kinases, 15-45 NATs, 15-50 PRMTs)
- significantly fewer de-modification enzymes
  (e.g. 20-70 phosphatases, 10-20 DACs, 0-20 DMEs)
- specific modification, unspecific de-modification?
- attempted irreversibility?
  Vorwärts immer, rückwärts nimmer!
- local storage of information (history and fate)
- transmission of information to the next generation

Eukaryots have a diverse set of reader domains.

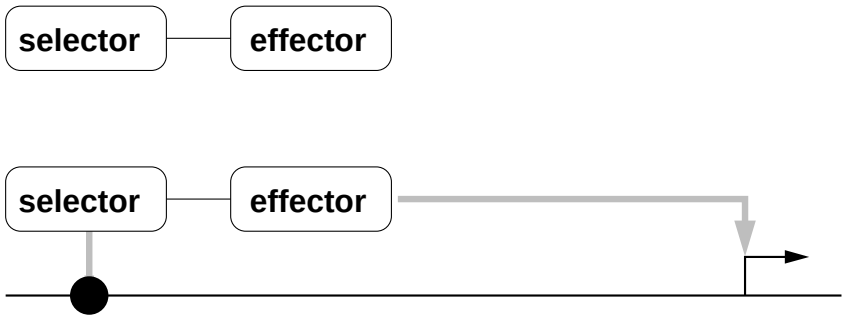


Modification readers allow for complex signal processing.



Hypotetical Evolution of Chromatin Regulation

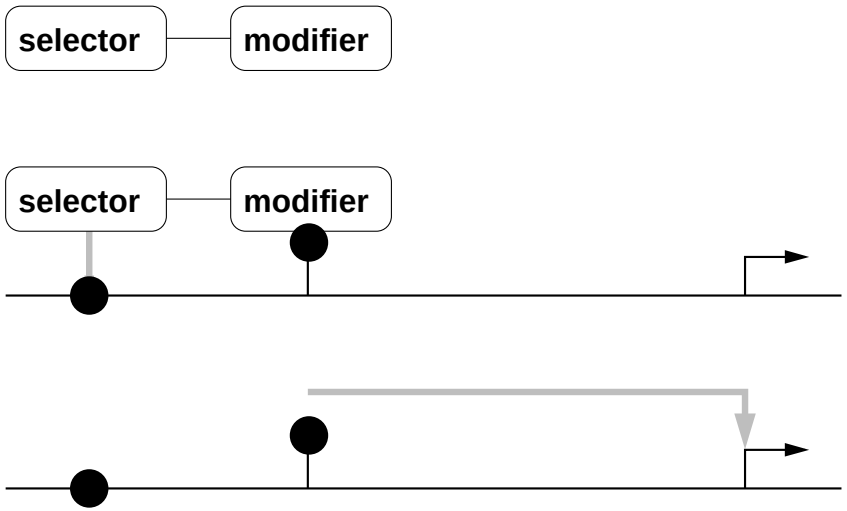
euBacteria: the if statement



- regulator is present when effect is carried out
- e.g. transcription factors, ChAP variants

Hypotetical Evolution of Chromatin Regulation

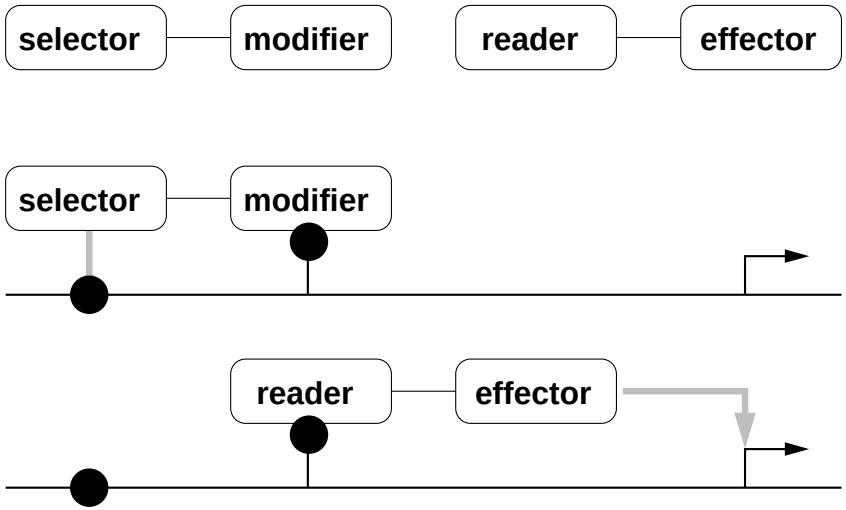
Archaea: the while statement



- regulator only initiates an effect
- the modification itself is responsible for the effect
- termination requires a second regulator
- e.g. acetylation of histones and Alba

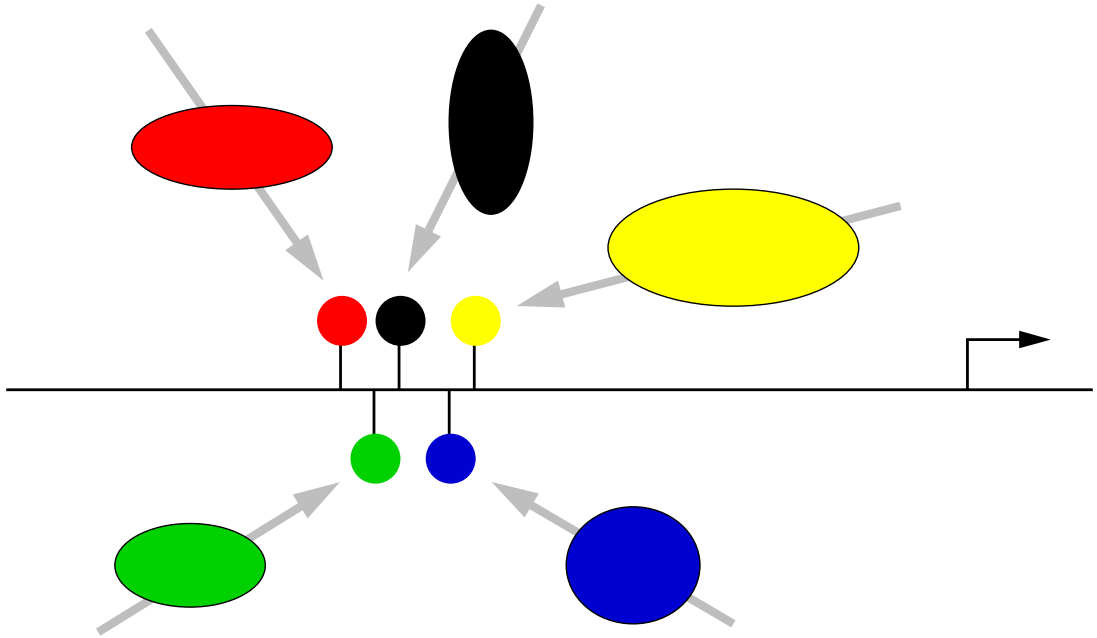
Hypotetical Evolution of Chromatin Regulation

Eukarya: ?



- regulator sets a signal
- any effect is coupled to the signal interpretation
- e.g. most modifications in eukarya

Setting and Interpreting Signals

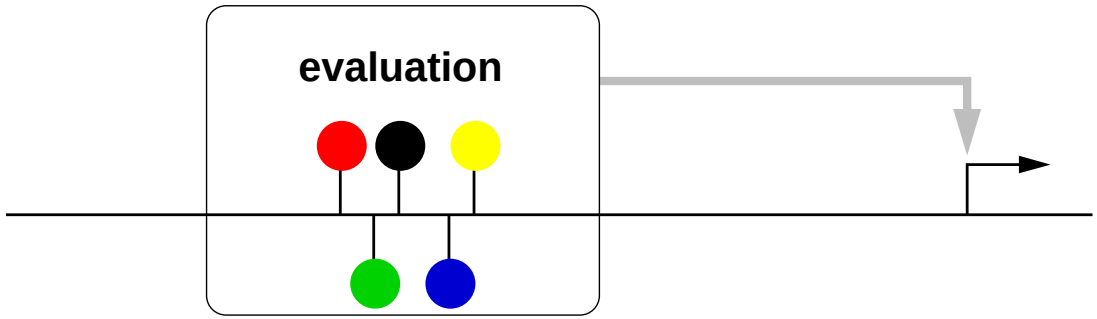


Chromatin acts as an interface to collect and store the input from multiple cellular processes commenting on the expression status of a certain locus.

A local memory holds the information recieved in the past for future use.

The information can be accessed and evaluated when a decision is due.

Setting and Interpreting Signals



Chromatin acts as an interface to collect and store the input from multiple cellular processes commenting on the expression status of a certain locus.
A local memory holds the information recieved in the past for future use.
The information can be accessed and evaluated when a decision is due.

Summary - A Hypothesis

- chromatin could have evolved from general regulators
- which evolved to switch between genomic/phenotypic states
- the major innovation in chromatin regulation is the use of modifications as signals (i.e. information)
- modifications as signals emerge with the appearance of modification readers
- and give rise to local memory and complex evaluation of manifold inputs

Contributors

Sonja J. Prohaska

David C. Krakauer

Douglas Erwin

Peter F. Stadler