Advanced Cellular Biology

Week 2- Epigenetic Modifications
Histone Code Hypothesis
Epigenetic Memory

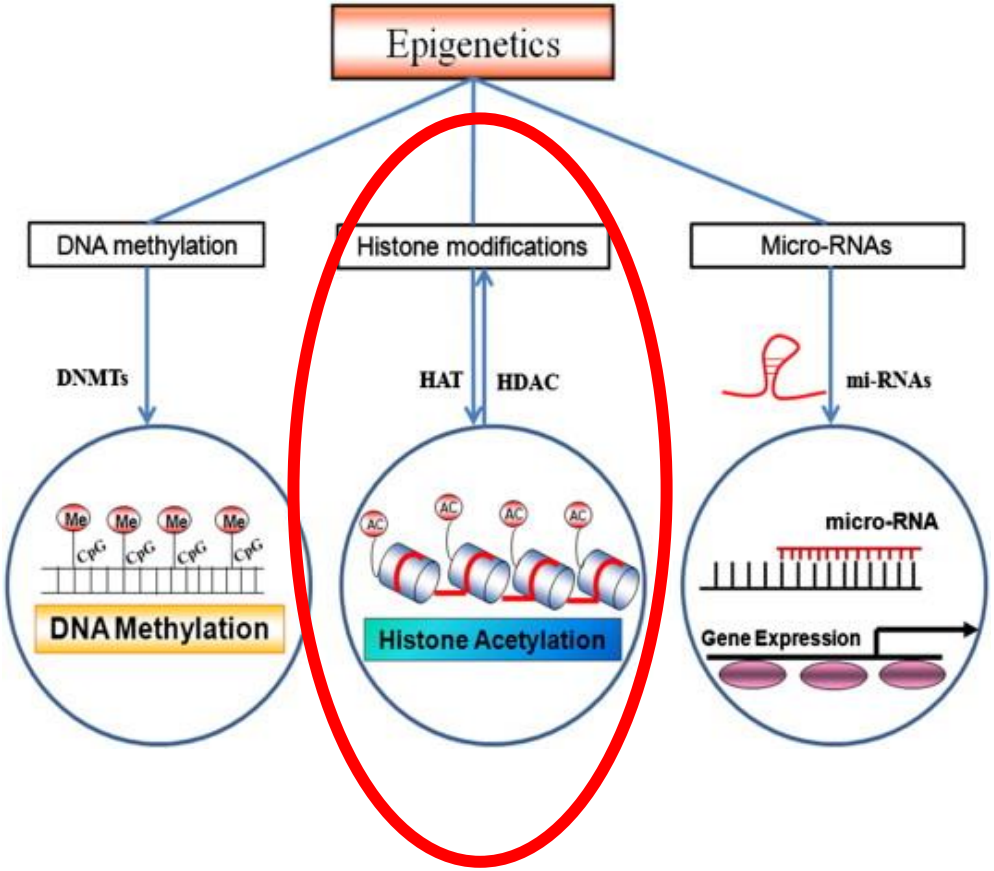
Epigenetic mechanisms

- The term, "epigenetics," was first used to refer to the complex interactions between the genome and the environment that are involved in development and differentiation in higher organisms.

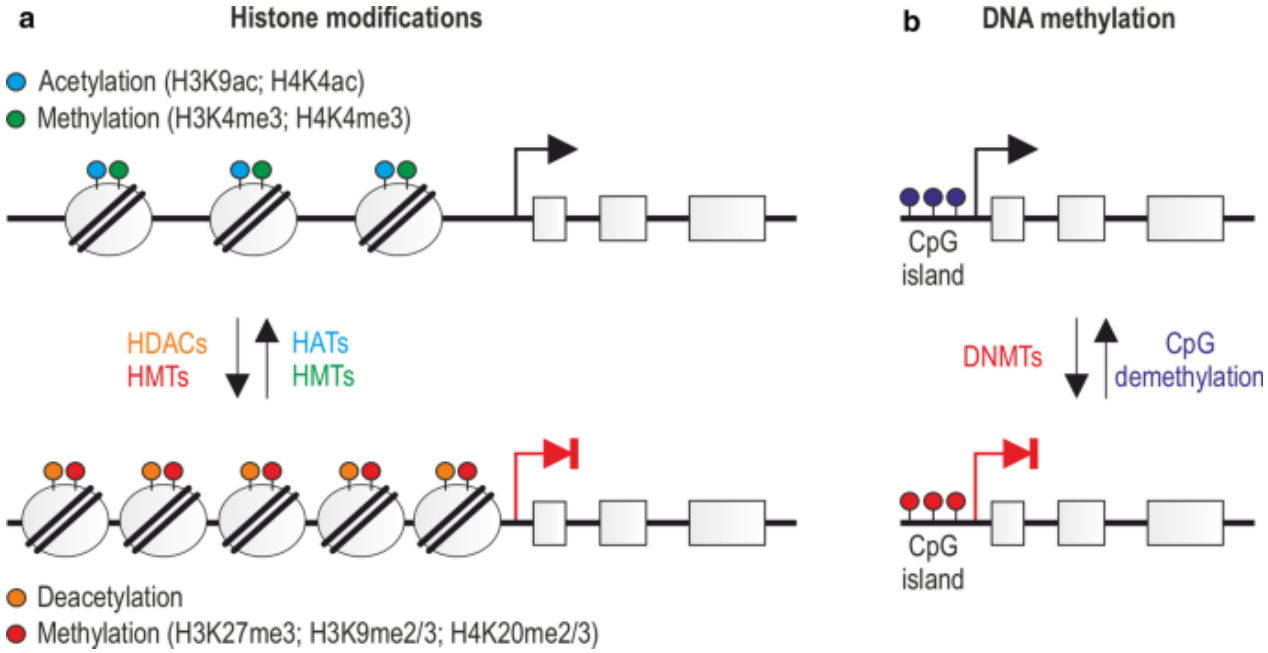
- Today, this term is used to refer to heritable alterations that are not due to changes in DNA sequence. Rather, epigenetic modifications, or "tags," such as DNA methylation and histone modification, alter DNA accessibility and chromatin structure, thereby regulating patterns of gene expression.

- These processes are crucial to normal development and differentiation of distinct cell lineages in the adult organism. They can be modified by exogenous influences, and, as such, can contribute to or be the result of environmental alterations of phenotype. Importantly, epigenetic programming has a crucial role in the regulation of pluripotency genes, which become inactivated during differentiation.

Epigenetic modifications



Epigenetic modifications which are defined by DNA methylation, histone modifications and microRNA mediated gene regulation, have been found to be associated with stem cell differentiation, especially at developmental, tissue regeneration and cancer stage



Covalent modification of core histone tails: Known modification of the four histone core proteins. K=lysine, S=serine.

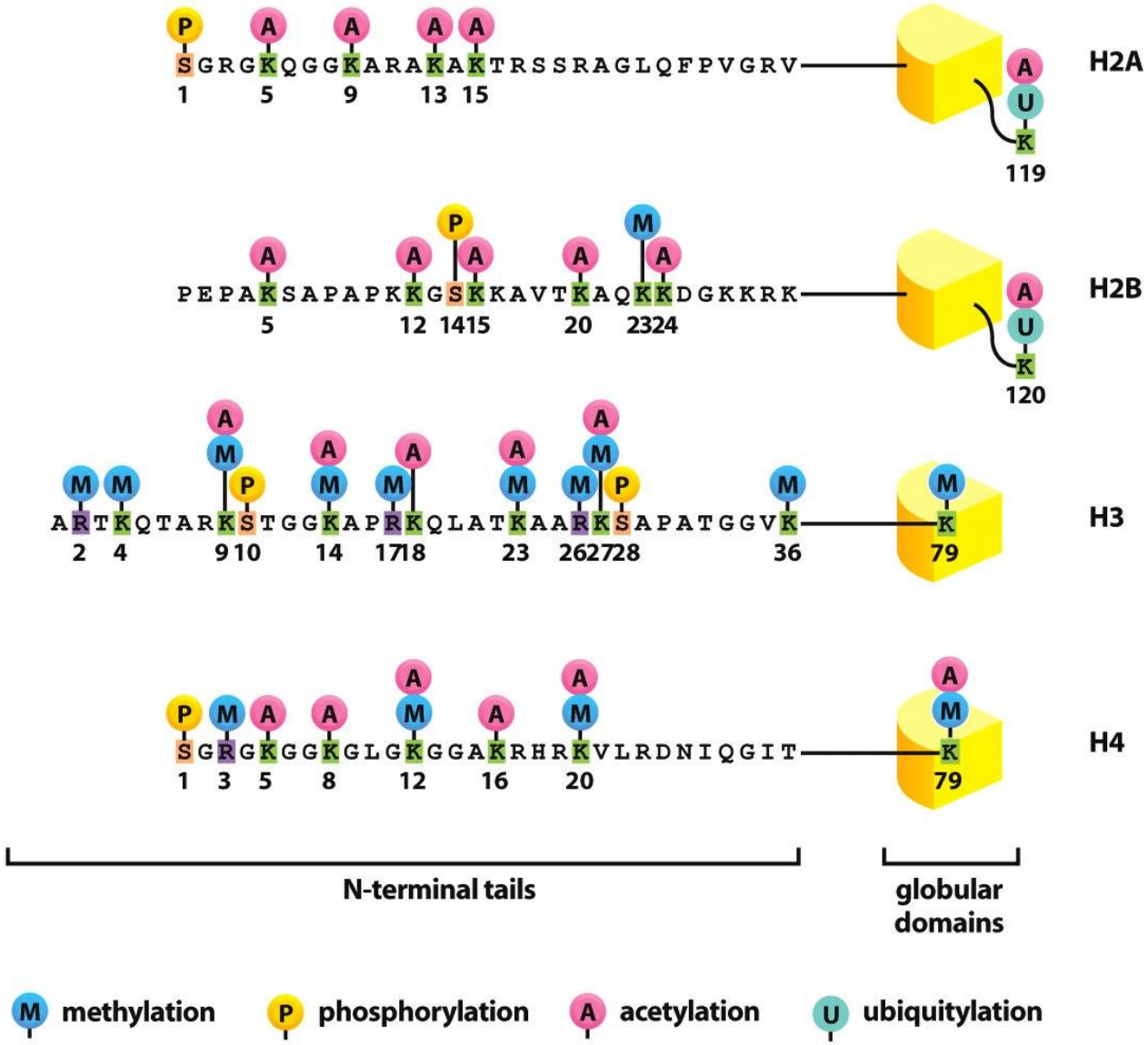


Figure 4-39b Molecular Biology of the Cell 5/e (© Garland Science 2008)

types of covalent amino acid-chain modifications found on nucleosomal histones

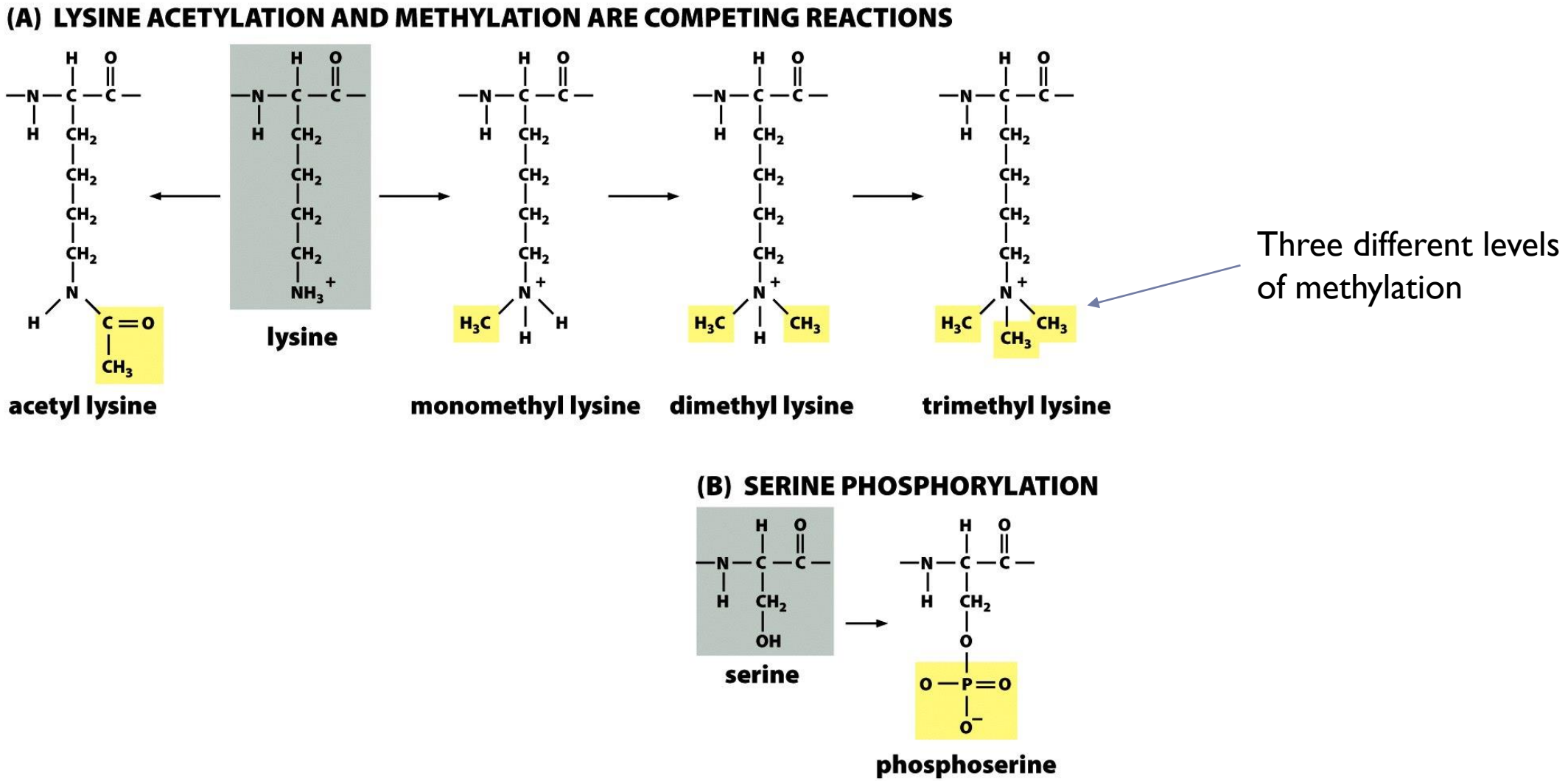


Figure 4-38 Molecular Biology of the Cell 5/e (© Garland Science 2008)

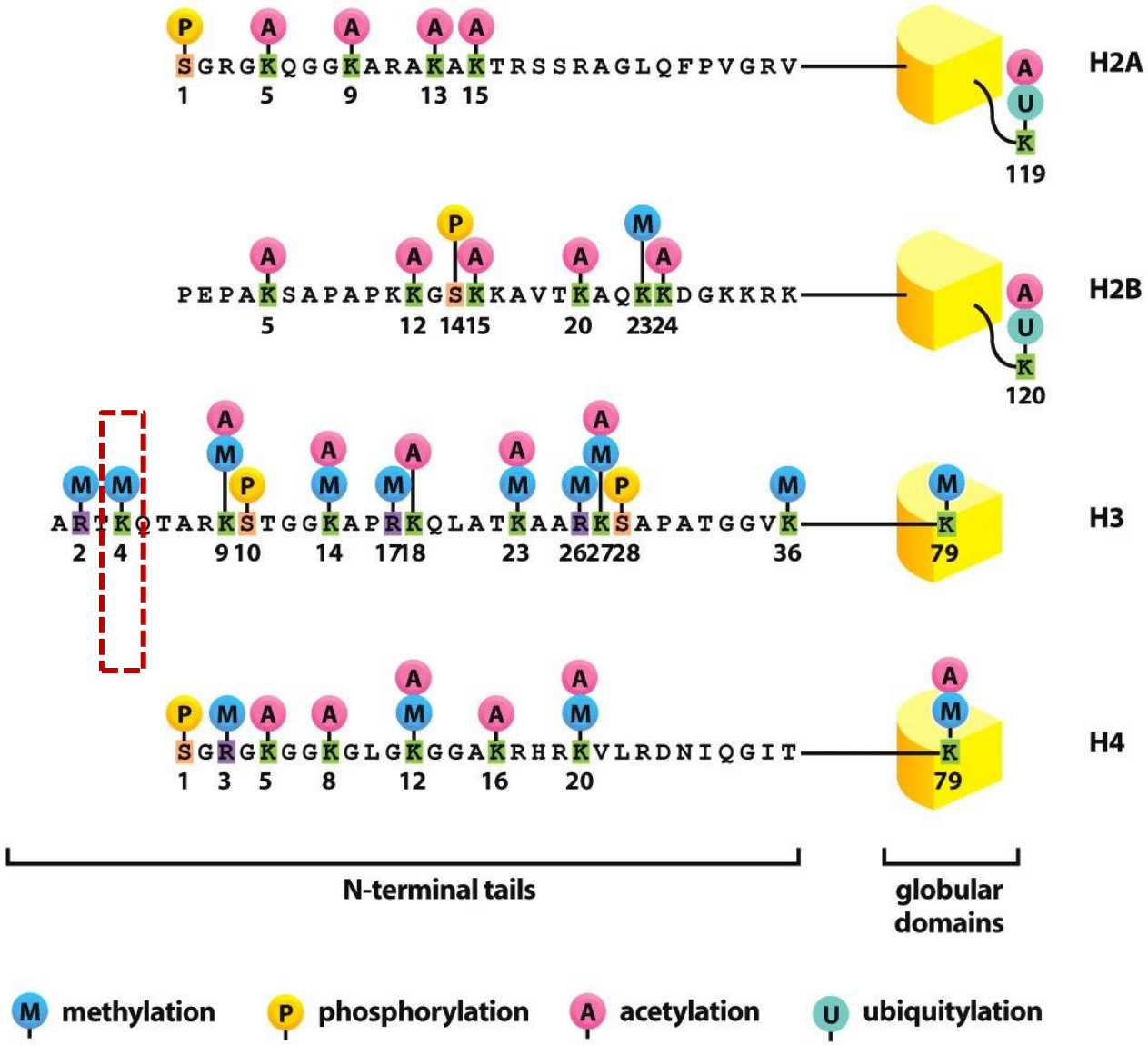


Figure 4-39b Molecular Biology of the Cell 5/e (© Garland Science 2008)

H3 K 4 me

H3K4me

Some specific meaning of histone code

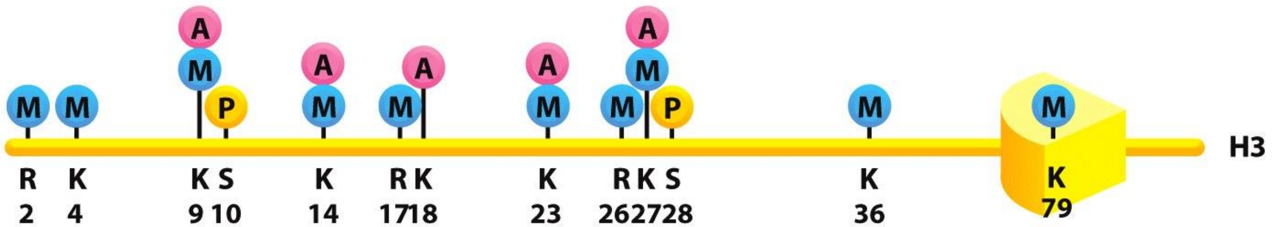


Figure 4-44a Molecular Biology of the Cell 5/e (© Garland Science 2008)

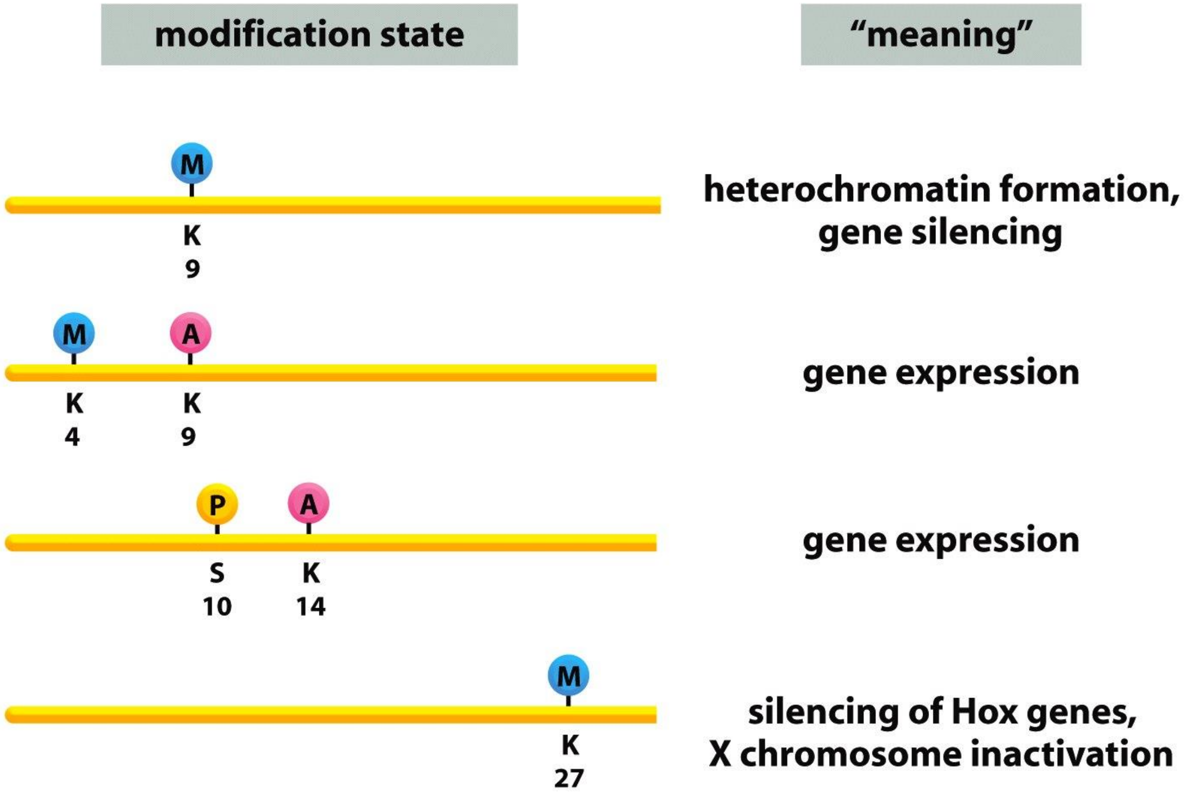
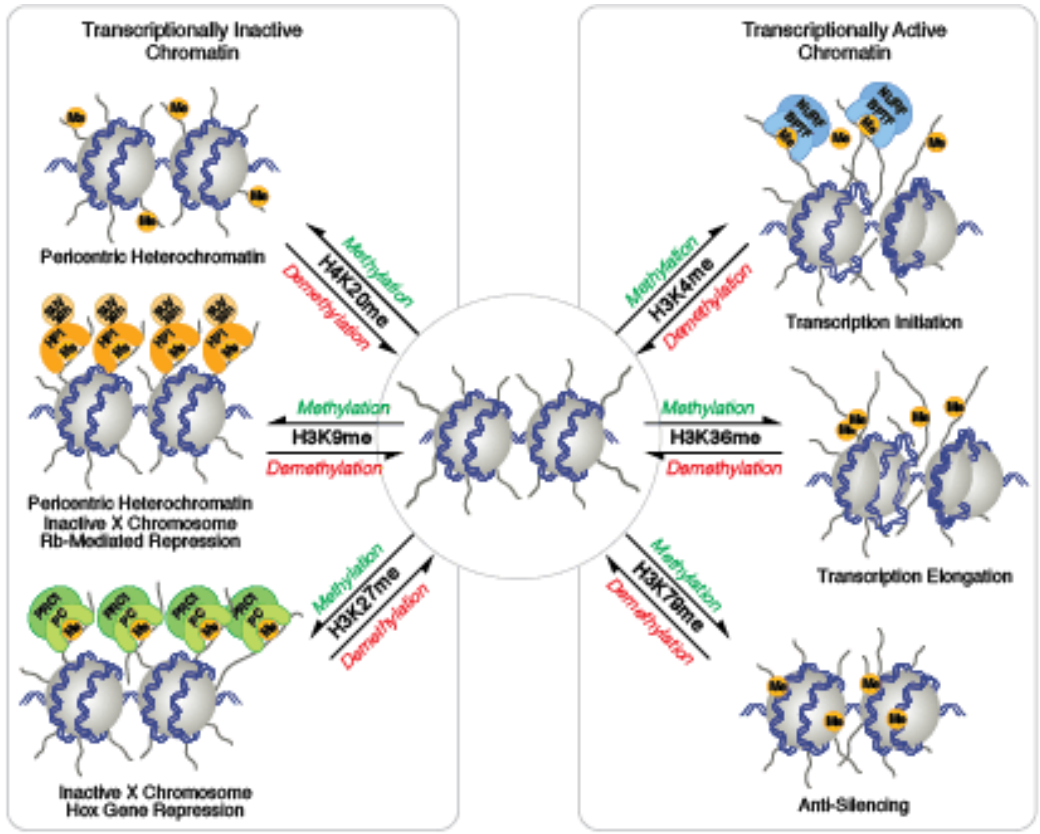


Figure 4-44b Molecular Biology of the Cell 5/e (© Garland Science 2008)

The H3 tail can be marked by different combinations that convey a specific meaning to stretch of chromatin where this combination occurs



Methylation Degree: + Mono † Di ‡ Tri

H4K20me

Methylases:

SET8/KMT5A (+)
NSD1/KMT3B (+,†)
ASH1L/KMT2H (+,†,‡)
SUV420H1/KMT5B (†,‡)
SUV420H2/KMT5C (†,‡)
NSD2/KMT3G (+,†)

Demethylases:

JHDM1D/KDM7A (+,†)
PHF8/KDM7B (+,†)

H3K9me

Methylases:

PRDM3/KMT8E (+)
PPRDM16/KMT8F (+)
G9a/EHMT2/KMT1C (+,†)
EHMT1/KMT1D (+,†)
ASH1L/KMT2H (+,†,‡)
PRDM2/KMT8A (+,†,‡)
PRDM8/KMT8D (†)
SUV39H1/KMT1A (†,‡)
SUV39H2/KMT1B (†,‡)
ESET/KMT1E (†,‡)
CLLD8/KMT1F (†,‡)

Demethylases:

AOF1/KDM1B (+,‡)
JMJD1A/KDM3A (+,†)
JMJD1B/KDM3B (+,†)
JMJD1C/KDM3C (+,‡)
JHDM1D/KDM7A (+,†)
PHF8/KDM7B (+,†)
JMJD2A/KDM4B (†,‡)
JMJD2B/KDM4B (†,‡)
JMJD2C/KDM4C (†,‡)
JMJD2D/KDM4D (†,‡)
KDM4E/KDM4DL (†,‡)

H3K27me

Methylases:

Ezh2/KMT6 (+,†,‡)
NSD2/KMT3G (+,†,‡)
NSD3/KMT3F (†,‡)

Demethylases:

JHDM1D/KDM7A (+,†)
PHF8/KDM7B (+,†)
UTX/KDM6A (†,‡)
JMJD3/KDM6B (†,‡)

H3K4me

Methylases:

MLL1/KMT2A (+,†,‡)
MLL2/KMT2B (+,†,‡)
MLL3/KMT2C (+)
MLL4/KMT2D (+)
SET1A/KMT2F (+,†,‡)
SET1B/KMT2G (+,†,‡)
NSD3/KMT3F (+,†)
NSD2/KMT3G (+,†)
SET7/KMT7 (†)
SMYD3/KMT3E (†,‡)
ASH1L/KMT2H (+,†,‡)
PRDM9/KMT8B (‡)

Demethylases:

LSD1/KDM1A (+,†)
AOF1/KDM1B (+,‡)
JARID1A/KDM5A (†,‡)
JARID1B/KDM5B (†,‡)
JARID1C/KDM5C (†,‡)
JARID1D/KDM5D (†,‡)

H3K36me

Methylases:

NSD1/KMT3B (+,†)
SMYD2/KMT3C (+,†)
NSD2/KMT3G (+,†)
SET2/KMT3A (‡)

Demethylases:

JMJD1A/KDM2A (+,†)
JMJD1B/KDM2B (+,†)
JMJD2A/KDM4A (†,‡)
JMJD2B/KDM4B (†,‡)
JMJD2C/KDM4C (†,‡)
JMJD2D/KDM4D (†,‡)
KDM4DL/KDM4E (†,‡)

H3K79me

Methylases:

DOT1L/KMT4 (+,†,‡)

Demethylases:

???

How each mark on a nucleosome is read:

- The structure of a protein module that specifically recognizes histone H3 trimethylated on lysine 4 (H3K4me3) is shown. These modules are thought to act in concert with other modules as a part of a *code-reader complex*, attract other protein complexes that execute an appropriate biological function at the right time.

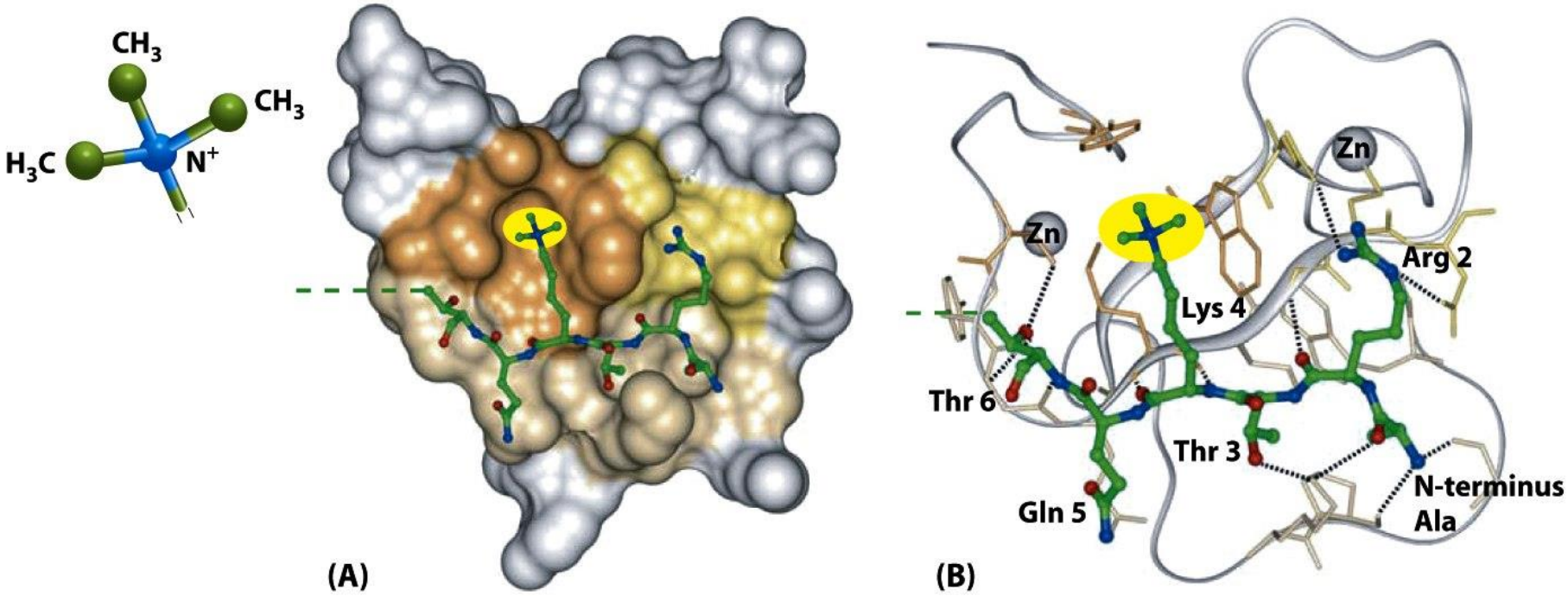
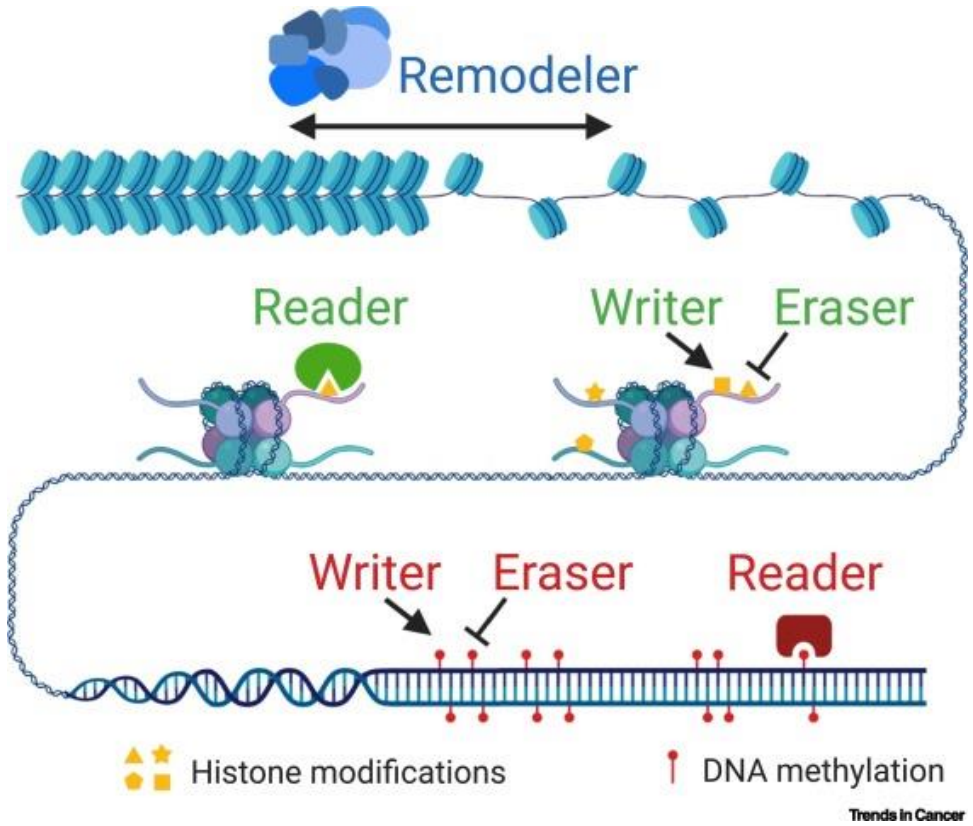


Figure 4-42 Molecular Biology of the Cell 5/e (© Garland Science 2008)

According to the histone code hypothesis, distinct combinations of covalent post-translational modifications of histones influence chromatin structure and lead to varied transcriptional outputs



epigenetic regulators: 'writers', 'readers', 'erasers', and 'remodelers'

Lysine methylations mark various sites on the tail and globular domains of histones and their levels are precisely balanced by the action of methyltransferases ('writers') and demethylases ('erasers').

In addition, distinct effector proteins ('readers') recognize specific methyl-lysines in a manner that depends on the neighboring amino-acid sequence and methylation state.

Misregulation of histone lysine methylation has been implicated in several cancers and developmental defects.

How the histone code could be read by a code-reader complex:

A large protein complex that contains a series of protein modules, each recognizes a specific histone mark (green). This code-reader complex will bind tightly only to a region of chromatin that contains different histone marks that it recognizes. Therefore, only a specific combination of marks will cause the complex to bind to chromatin and attract additional protein complexes (purple) that catalyzes a biological function.

Reading the histone code generally involves the joint recognition of the marks at other sites on the nucleosome along with the indicated tail recognition.

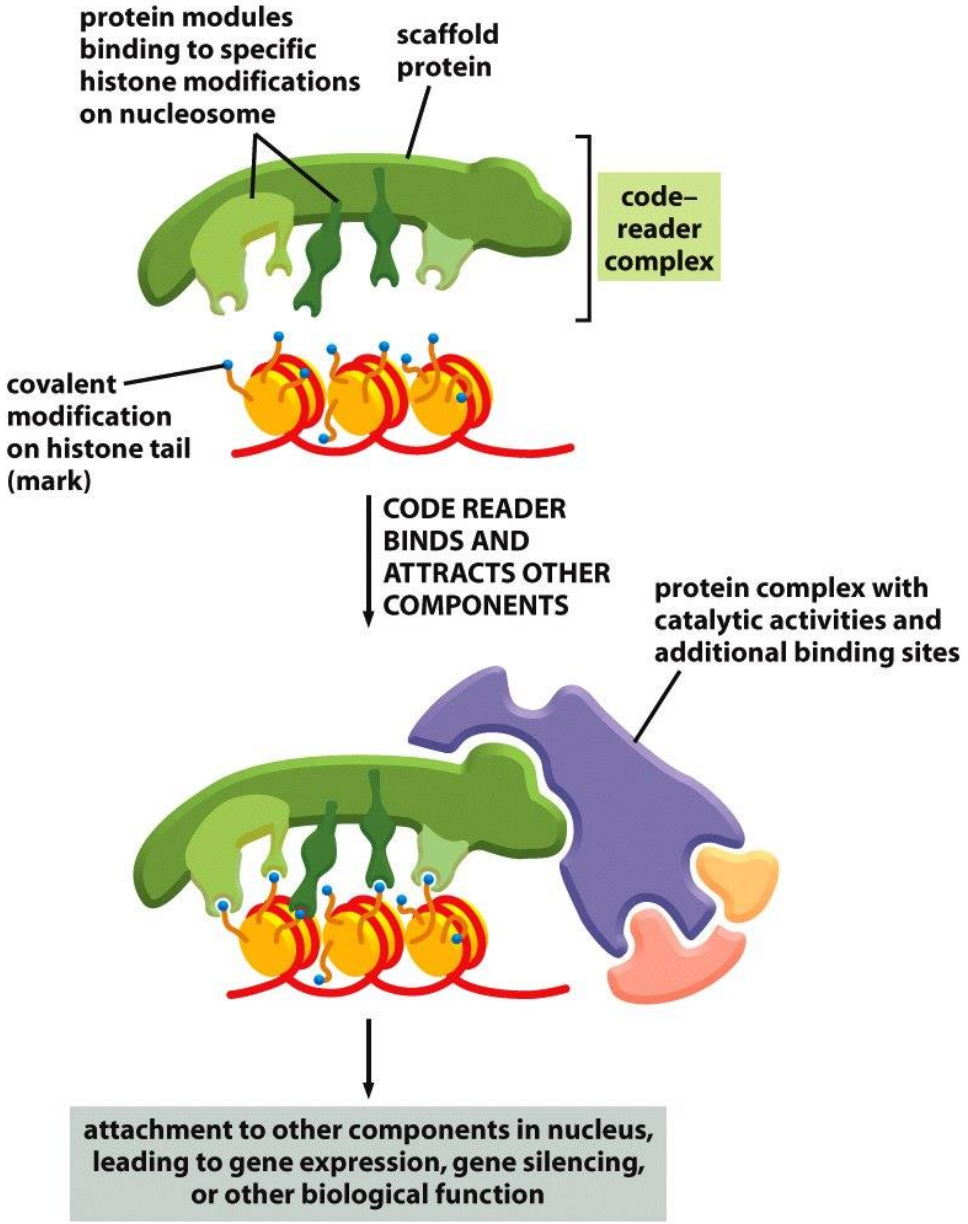


Figure 4-43 Molecular Biology of the Cell 5/e (© Garland Science 2008)

How code-reader and code-writer complexes can spread chromatin modifications along a chromosome:

The *code-writer* is an enzyme that creates a specific modification on one or more of the four nucleosome histones.

After its recruitment to a specific site on a nucleosome by a *gene regulatory protein*, the writer *collaborates with a code-reader* protein **to spread its** mark from nucleosome to nucleosome by means of indicating **reader-writer complex**.

To work, the reader must recognize the same histone modification mark that the writer produces.

ATP-dependent remodelling complexes are type of reader-writer complex.

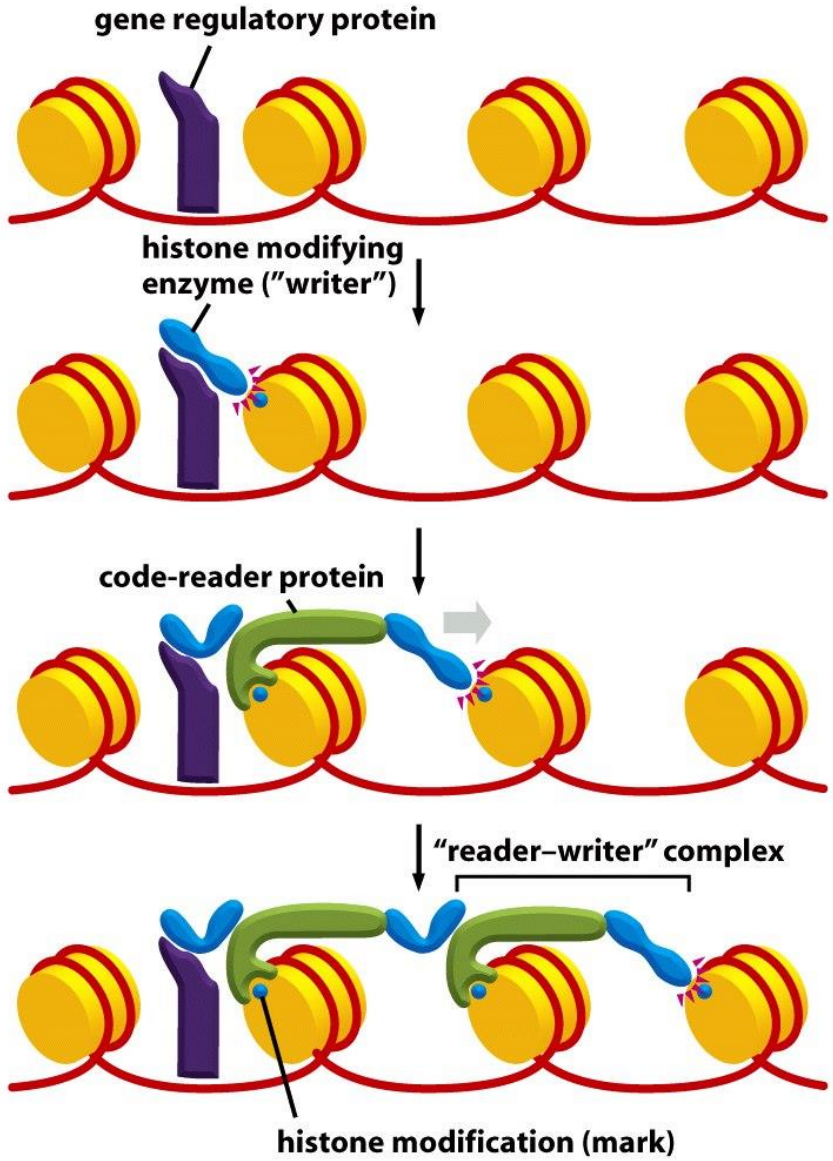


Figure 4-45 Molecular Biology of the Cell 5/e (© Garland Science 2008)

How a complex containing reader-writer and ATP-dependent chromatin remodeling proteins can spread chromatin changes along the chromosome:

A spreading wave of chromatin condensation. The reader-writer complex collaborates with an ATP-dependent chromatin remodeling protein to reposition nucleosomes and pack them into highly condensed arrays.

The heterochromatin specific **protein HP1 plays a major role in this process**. HP1 binds to trimethyl lysine 9 on histone H3 (H3K9me3), and it remains associated with condensed chromatin.

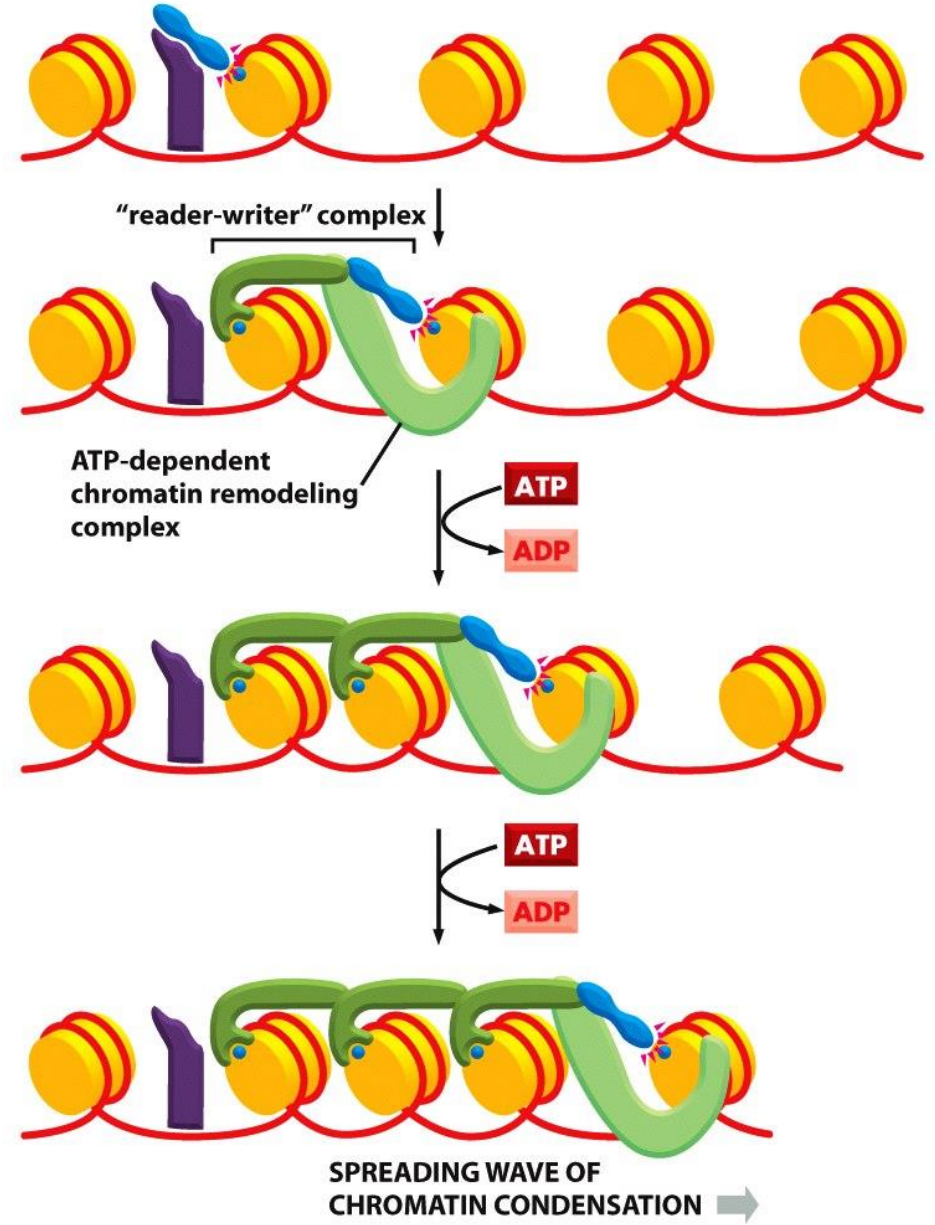


Figure 4-46 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Barrier DNA sequences block the spread of Reader-Writer complexes and thereby separate neighboring chromatin domains: *Models*

a.The tethering of a region of chromatin to a large fixed site, such as nuclear pore complex, can form a barrier that stops the spreading of heterochromatin.

b.The tight binding of barrier proteins to a group of nucleosomes can compete with heterochromatin spreading.

c.by recruiting a group of highly active histone-modifying enzymes, barriers can erase the histone marks that are required for heterochromatin to spread.

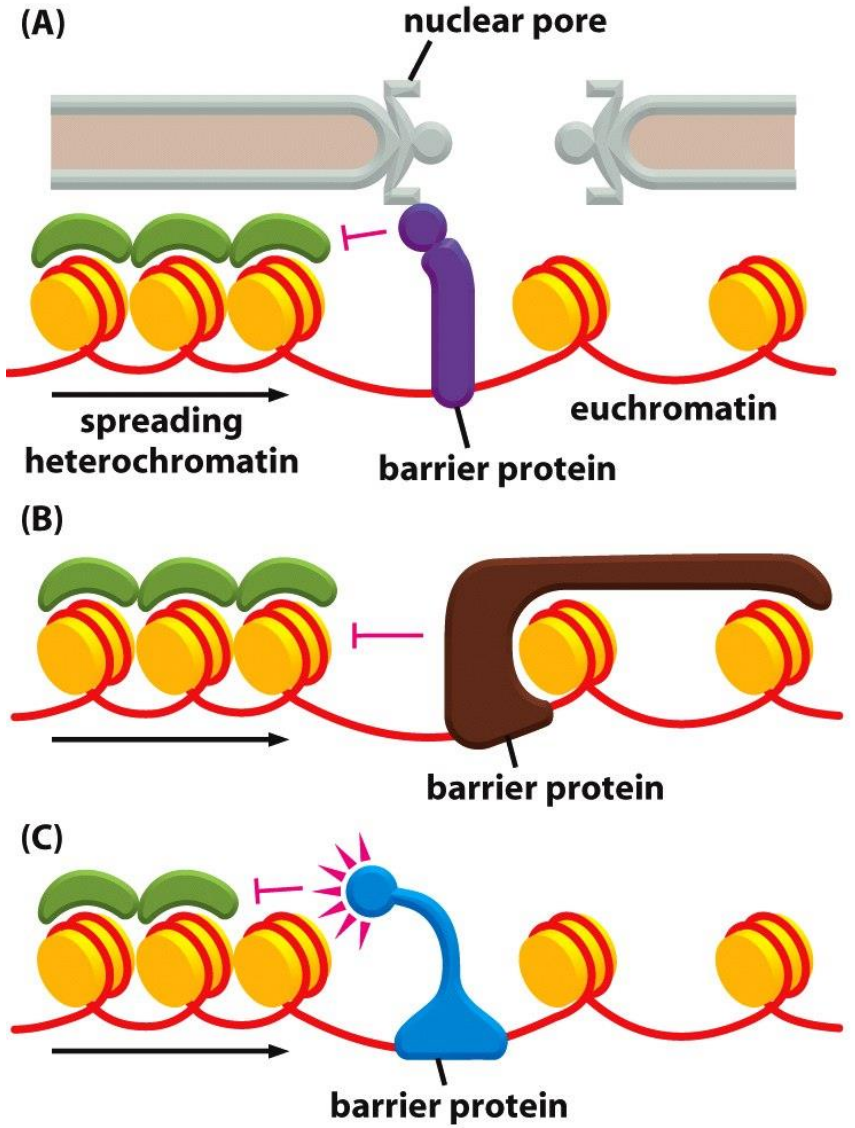
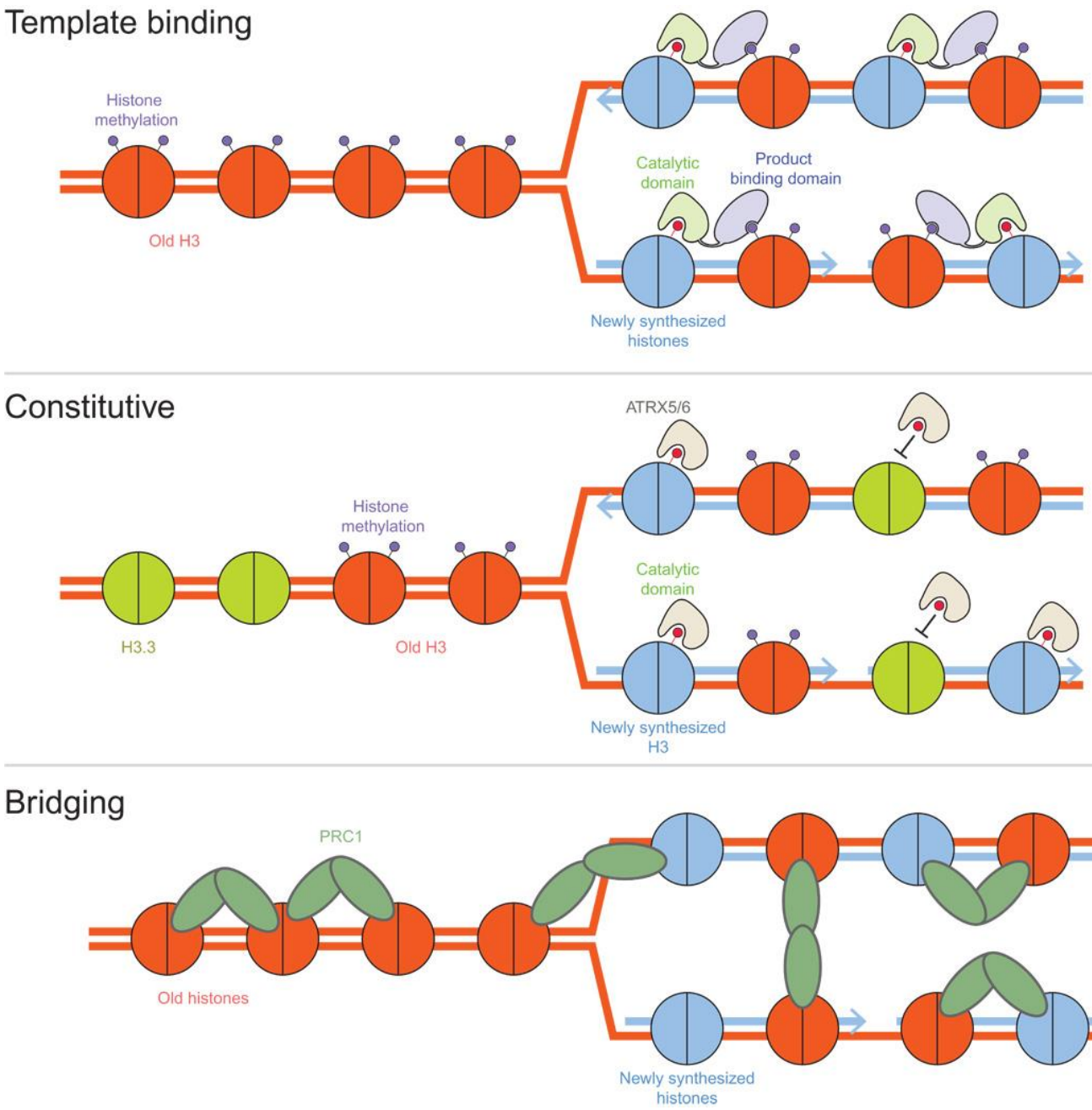


Figure 4-47 Molecular Biology of the Cell 5/e (© Garland Science 2008)



Three models of propagating histone modifications through replication.

In the template-binding model , adjacent nucleosomes are modified by a histone-modifying enzyme that binds the modified residue on a nearby tail.

In the constitutive model , H3K27 methylation is restored by recognition of by ATXR5/6, such that only replication-coupled (H3) nucleosomes, not replication-independent (H3.3) nucleosomes, are methylated on H3K27.

In the bridging model, PRC1 bridges nucleosomes across daughter chromatids.

Bivalent Histone Modifications

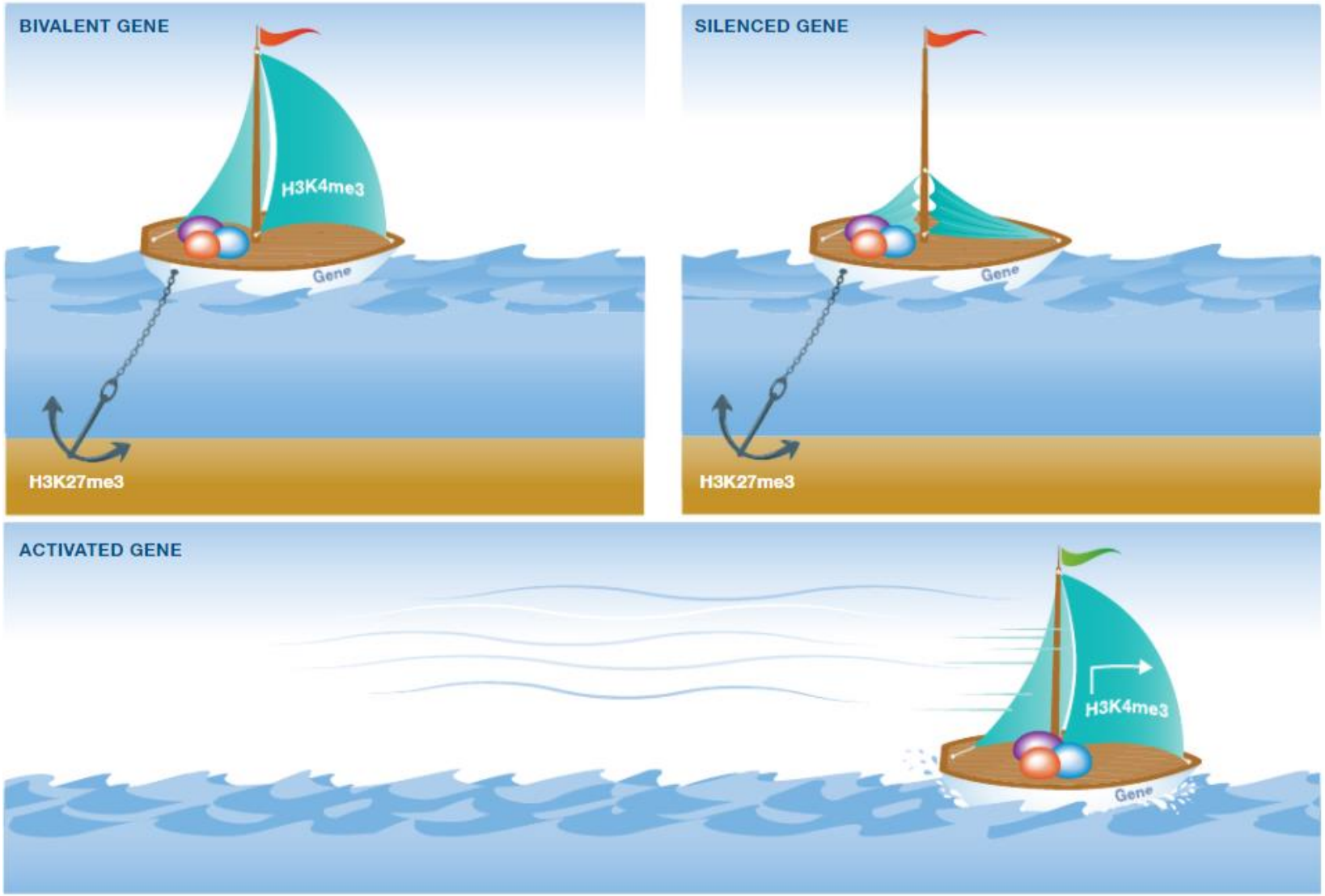
- It is believed that *histone methylation* potentially takes charge of cell fate determination and differentiation. The synchronous existence of functionally opposite histone marks at transcript start sequence (TSS) is defined as "Bivalency", which mainly mark development related genes.

- H3K4me3 and H3K27me3, the prominent histone methylations of bivalency, are implicated in transcriptional activation and transcriptional repression respectively.

- i.e: Several thousands of the H3K4me3-enriched promoters in pluripotent cells also contain H3K27me3 mark.

- bivalent promoters are not unique to pluripotent cells, but they are relatively enriched in these cell types, largely marking developmental and lineage-specific genes, primordial germ cells, and male germ cells which are silent but poised for immediate and rapid activation upon cell differentiation

- The delicate balance between H3K4me3 and H3K27me3 produces diverse chromatin architectures, resulting in different transcription states of downstream genes: "poised", "activated" or "repressed".



A bivalent gene, depicted as a boat (top left), is ready to go (sail up: H3K4me3) but **is held in check** (anchor: H3K27me 3). Once the sail is down (top right), the gene is **stably silenced** (only H3K 27me3), but if instead the anchor is lifted (bottom), the gene is promptly **activated** (only H3K4me3).

Chromatin remodeling and bivalent histone modifications in embryonic stem cells. EMBO Reports. DOI 10.15252/embr.201541011 | Received 13 July 2015

Review > Curr Stem Cell Res Ther. 2018;13(2):83-90. doi: 10.2174/1574888X12666170123144743.

Bivalent Histone Modifications and Development

Feifei Li [1 2], Mian Wan [1 2], Binpeng Zhang [1 2], Yiran Peng [1 2], Yachuan Zhou [1 2], Caixia Pi [1 2], Xin Xu [1 2], Ling Ye [1 2], Xuedong Zhou [1 2], Liwei Zheng [1 2]

Bivalent histone modifications are considered to set up genes for activation during lineage commitment by H3K4me3 and repress lineage control genes to maintain pluripotency by H3K27me3.

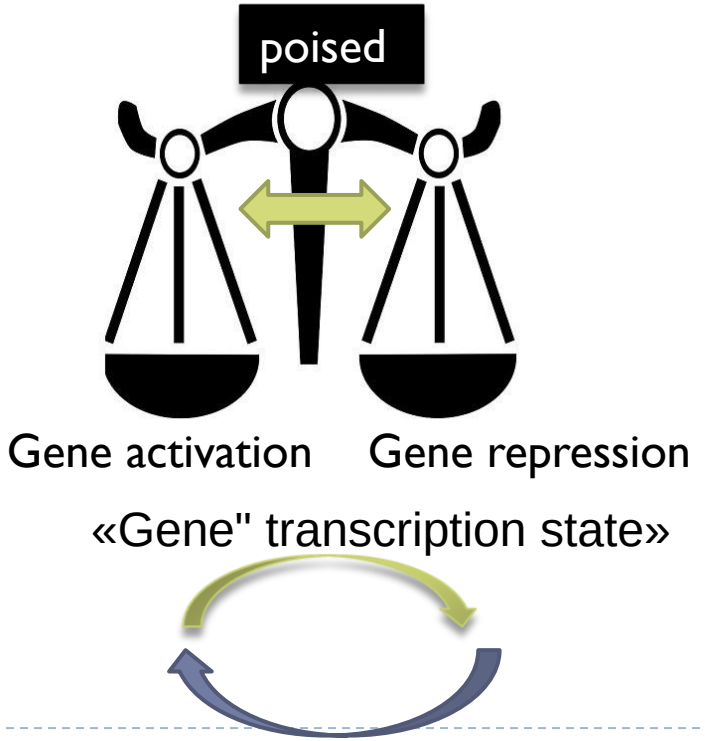
received developmental signals

↓

switch on or switch off

↓

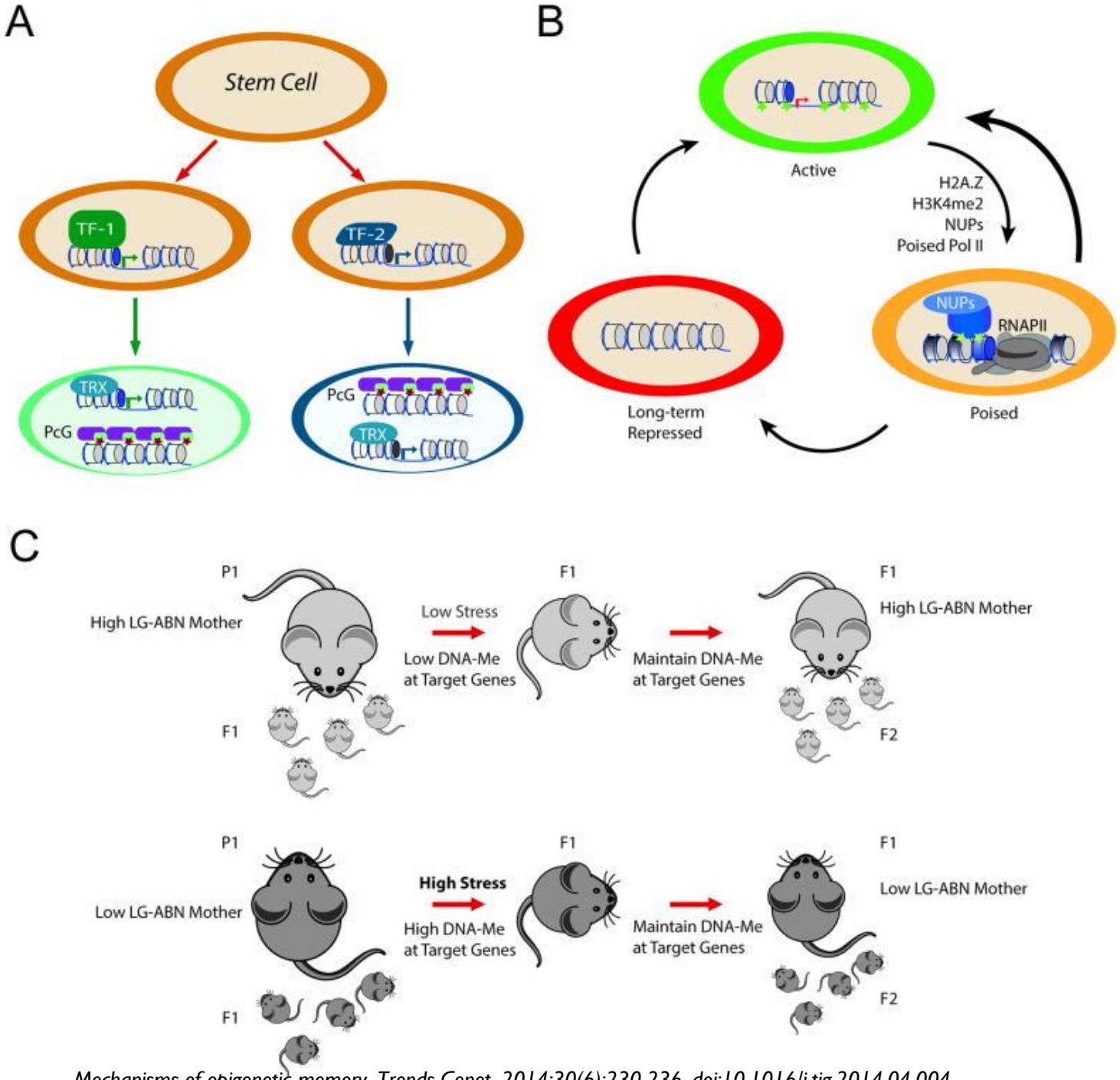
irreversible and spontaneous differentiation procedures.



Summarily, bivalency in stem cells keeps stemness via poising differentiation relevant genes.

what is epigenetic memory?

The epigenetic memory of a cell defines the set of modifications to the cell's deoxyribonucleic acid (DNA) that do not alter the DNA sequence, and have been inherited from the cell from which it descends.



Mechanisms of epigenetic memory. Trends Genet. 2014;30(6):230-236. doi:10.1016/j.tig.2014.04.004

types of epigenetic memory

1) cellular memory, mitotically heritable transcriptional states established during development in response to developmental cues,

2) 2) transcriptional memory, mitotically heritable changes in the responsiveness of organisms to environmental stimuli due to previous experiences

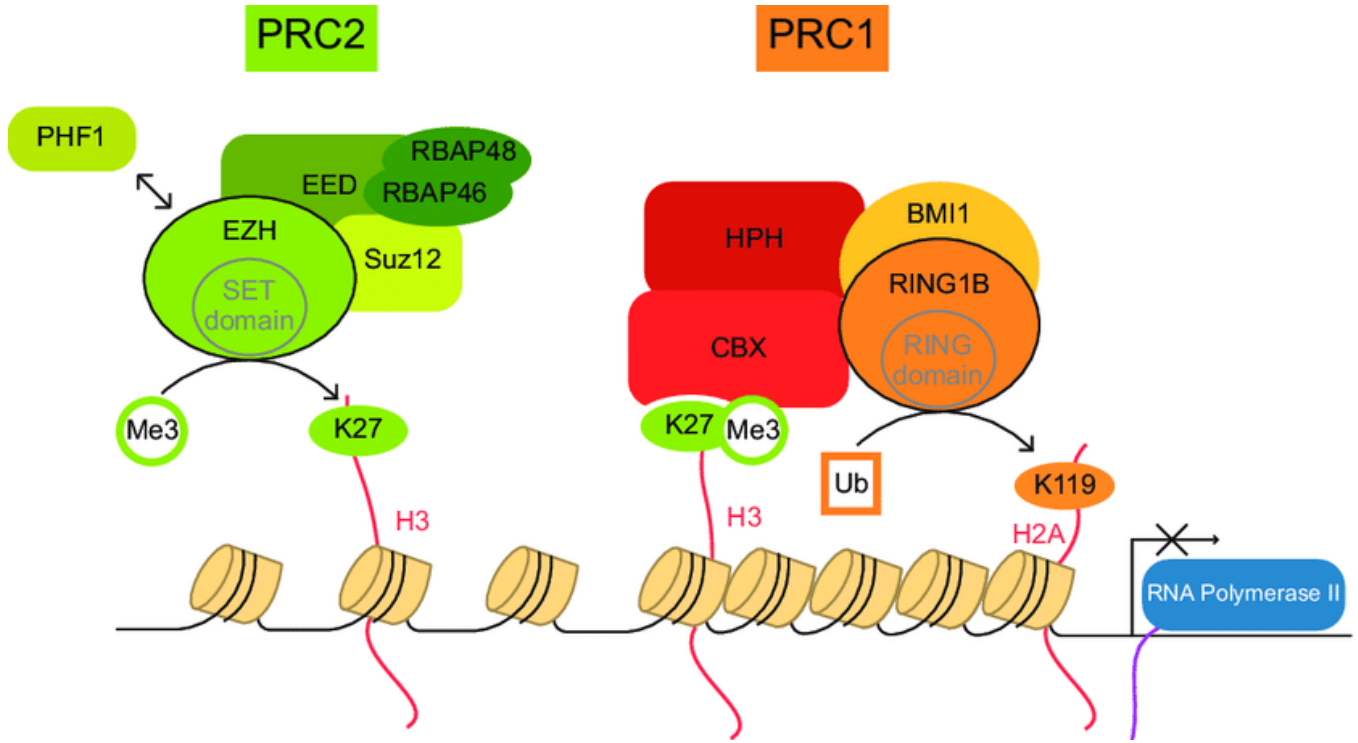
3) 3) transgenerational memory, meiotically heritable changes in the gene expression and physiology of organisms in response to experiences in the previous generations

- This cellular memory requires the 2 antagonistic protein groups: **Trithorax and Polycomb group (PcG) proteins**
- The Trithorax complex mediates methylation of histone H3 on lysine 4 (H3K4me) and is required to maintain genes in an active state
- The PcG proteins have two biochemically characterized repressive complexes: PRC1 and PRC2. PRC2 methylates histone 3 lysine 27 (H3K27me) at genes targeted for silencing. PRC1 binds H3K27me and induces spreading of structural changes in the chromatin .This histone mark has been proposed to act as a repressive "bookmark" during mitosis where it is maintained through cell division and transmitted through DNA replication in the absence of the initial stimuli

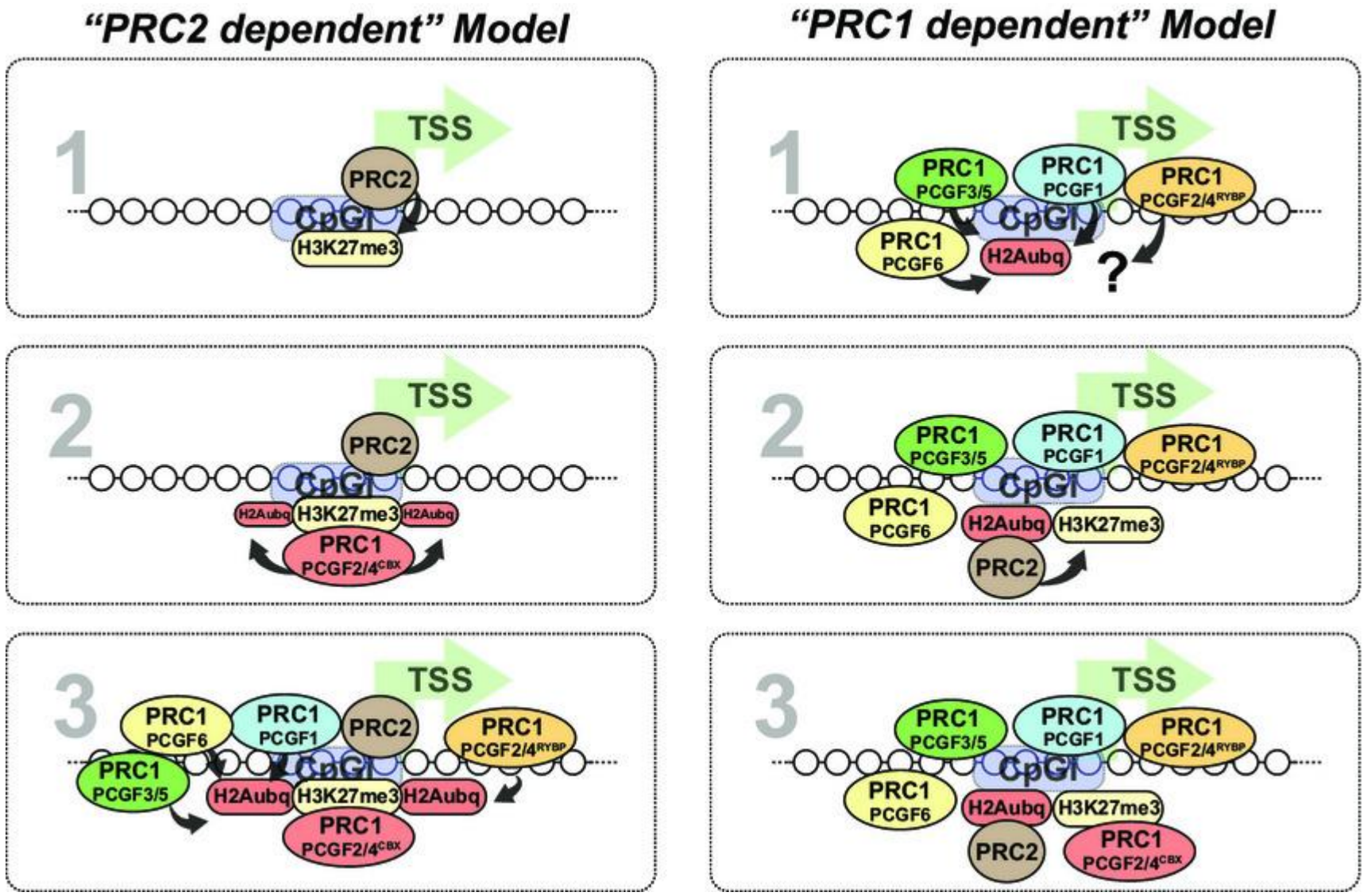
PcG and trxG are organized into large multimeric complexes which act on their target genes by modulating chromatin structure.

PcG : PRC1 / PRC2

- Consecuitive steps in repression: transcriptionally silent state
- Methylates H3K27 and H3K9
- E(Z) component: functional part !
- Following PRC2-catalyzed modification, PRC1 binds via the chromodomain to stabilize silencing.
- PRC1 includes chromodomain part. This complex seems to be recruited to chromatin through the ability of the CBX proteins to bind the H3K27me3 deposited by PRC2 in order to maintainance of silent state.
- The Chromodomain is found to bind to methyl moieties at H3K27 and H3K9
- RING1B is associated with ubiquitylated H2A (i.e on the inactive X chromosome) and the maintenance of this histone mark is dependent on RING1 proteins



- Highlighted that PcG enrichment and H3K27me3 deposition are not restricted to PREs and can be frequently found also at gene promoters
- In mammals, PcGs preferentially associate with CpG-rich promoters
- Indeed, **no clear-cut data** have been published about the role of the direct recognition of these DNA elements by PRC1 and/or PRC2 complexes and their recruitment to target promoters.

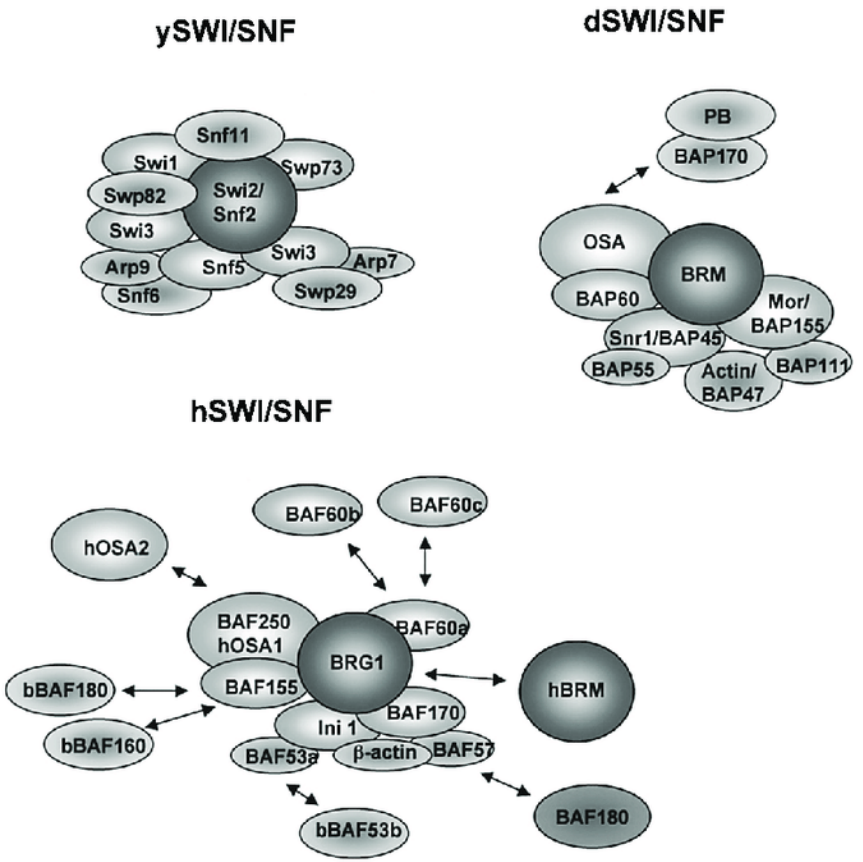


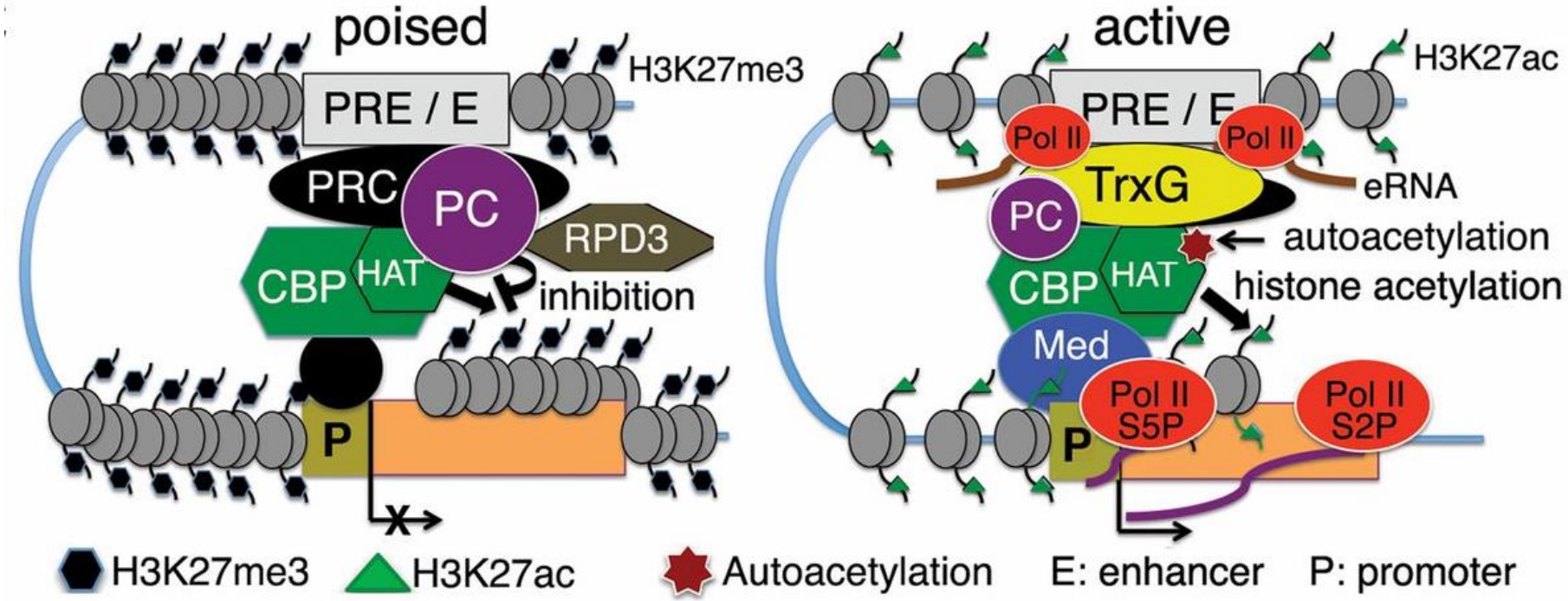
Recent reports have further challenged the previous model showing that the PRC2 complex can be recruited to chromatin by non-canonical PRC1 sub-complexes. While PRC1-PCGF1^{RYBP} and PRC1-PCGF3/5^{RYBP} are able to deposit H2Aubq and induce the recruitment of PRC2 activity, the forced recruitment of PRC1-PCGF2/4^{CBX} to the same genomic loci failed to deposit H2Aubq and to recruit PRC2 activity . Furthermore, PRC2 was recently shown to bind *in vivo* and *in vitro* H2Aubq . Together, these observations suggest a novel mechanism in which non-canonical PRC1 complexes are first recruited to establish H2Aubq domains, which mediate PRC2 association and H3K27me3.

PcG and trxG are organized into large multimeric complexes which act on their target genes by modulating chromatin structure.

trxG group

- Required for the maintenance of ON state throughout the lifetime of an organism.
- By analogy, trxG members covalently modify histone tails via methyltransferase and acetyltransferase activities (TRX and ASH)
- Directly regulates the targetting of activities of ATP-dependent remodelling complexes.
- In addition to ability to directly acvitave transcription, trxG protein have ability to block function of PcG repressors.
- ATP remodelling complexes such as the one that contains trxG proteins :BRM and MOR, which are functional core of SWI/SNF complexes, incrase the ability of DNA-binding proteins to bind to chromatin.
- SWI/SNF contain bromodomain: a protein motif associated with the binding of certain acetylated histones
- Several trxG are able to covalenty modify histone tails:
- MLL1, ASH1 = H3K4 methylation

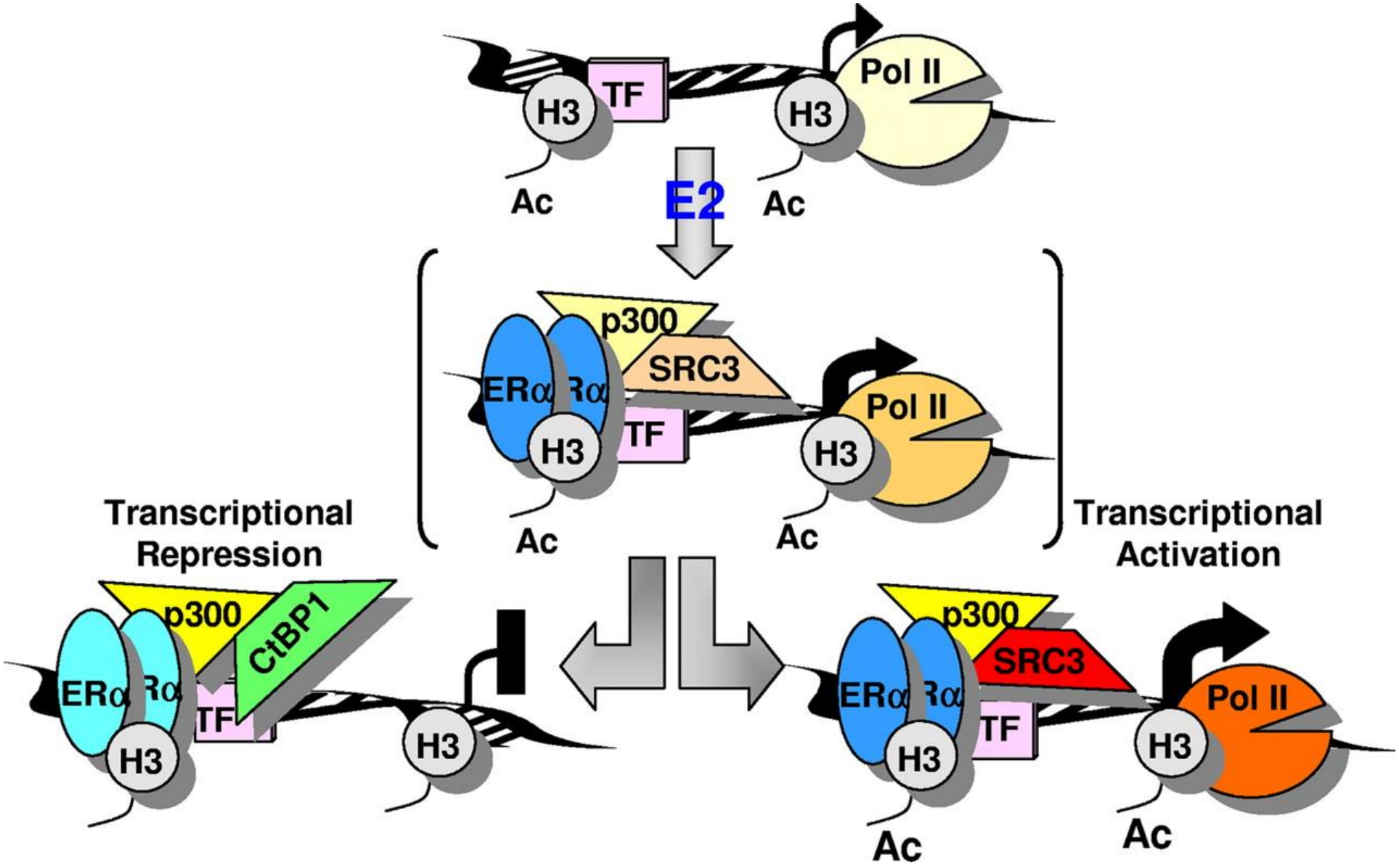




In the repressive state, chromatin is condensed. CBP, PC, and PRC complexes are present at poised enhancer and PRE (marked by H3K27me3 and low-level H3K4me1). PC is associated with unacetylated CBP. CBP HAT activity is likely inhibited by PC. PC- and CBP-associated RPD3 may maintain H3K27 in a deacetylated state.

In the active state, chromatin is relaxed. H3K27 acetylation occurs, but its level is inversely proportional to PC level. Autoacetylated CBP, Pol II, and TrxG proteins, and a relatively lower level of PC are present at active enhancers (marked by H3K27ac and high-level H3K4me1) that are transcribed into enhancer RNA (eRNA)

Polycomb inhibits histone acetylation by CBP by binding directly to its catalytic domain. *PNAS February 9, 2016 113 (6) E744-E753;*



Estrogen receptor alpha represses transcription of early target genes via p300 and CtBP1. Mol Cell Biol. 2009 Apr;29(7):1749-59. doi: 10.1128/MCB.01476-08.

A second major chromatin regulating system is that…

- HP1a, which is normally critical for the formation of constitutive heterochromatin, also affects the generation of the epigenetic marks of the Polycomb/trithorax groups of proteins, chromatin modifiers which are key to maintaining gene expression in euchromatin.

- The small non-histone protein Heterochromatin protein 1a (HP1a) plays a vital role in packaging chromatin, most notably in forming constitutive heterochromatin at the centromeres and telomeres.

Published in final edited form as:
Nature. 2010 September 16; 467(7313): 285–290. doi:10.1038/nature09342.

Epigenetic memory in induced pluripotent stem cells

K Kim[1], **A Doi**[2], **B Wen**[2], **K Ng**[1], **R Zhao**[1], **P Cahan**[1], **J Kim**[3], **MJ Aryee**[2,4], **H Ji**[2], **L Ehrlich**[5], **A Yabuuchi**[1], **A Takeuchi**[1], **KC Cunniff**[1], **H Hongguang**[1], **S Mckinney-Freeman**[1], **O Naveiras**[1], **TJ Yoon**[6,#], **RA Irizarry**[2], **N Jung**[2], **J Seita**[5], **J Hanna**[7], **P Murakami**[2], **R Jaenisch**[7], **R Weissleder**[6], **SH Orkin**[3], **IL Weissman**[5], **AP Feinberg**[2,*], and **GQ Daley**[1,*]

[1]Stem Cell Transplantation Program, Division of Pediatric Hematology/Oncology, Manton Center for Orphan Disease Research, Howard Hughes Medical Institute, Children's Hospital Boston and Dana Farber Cancer Institute; Division of Hematology, Brigham and Women's Hospital; Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School; Harvard Stem Cell Institute; Boston, MA 02115, USA

[2]Center for Epigenetics and Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

[3]Department of Pediatric Oncology, Howard Hughes Medical Institute, Children's Hospital Boston and Dana Farber Cancer Institute; Boston, MA 02115, USA

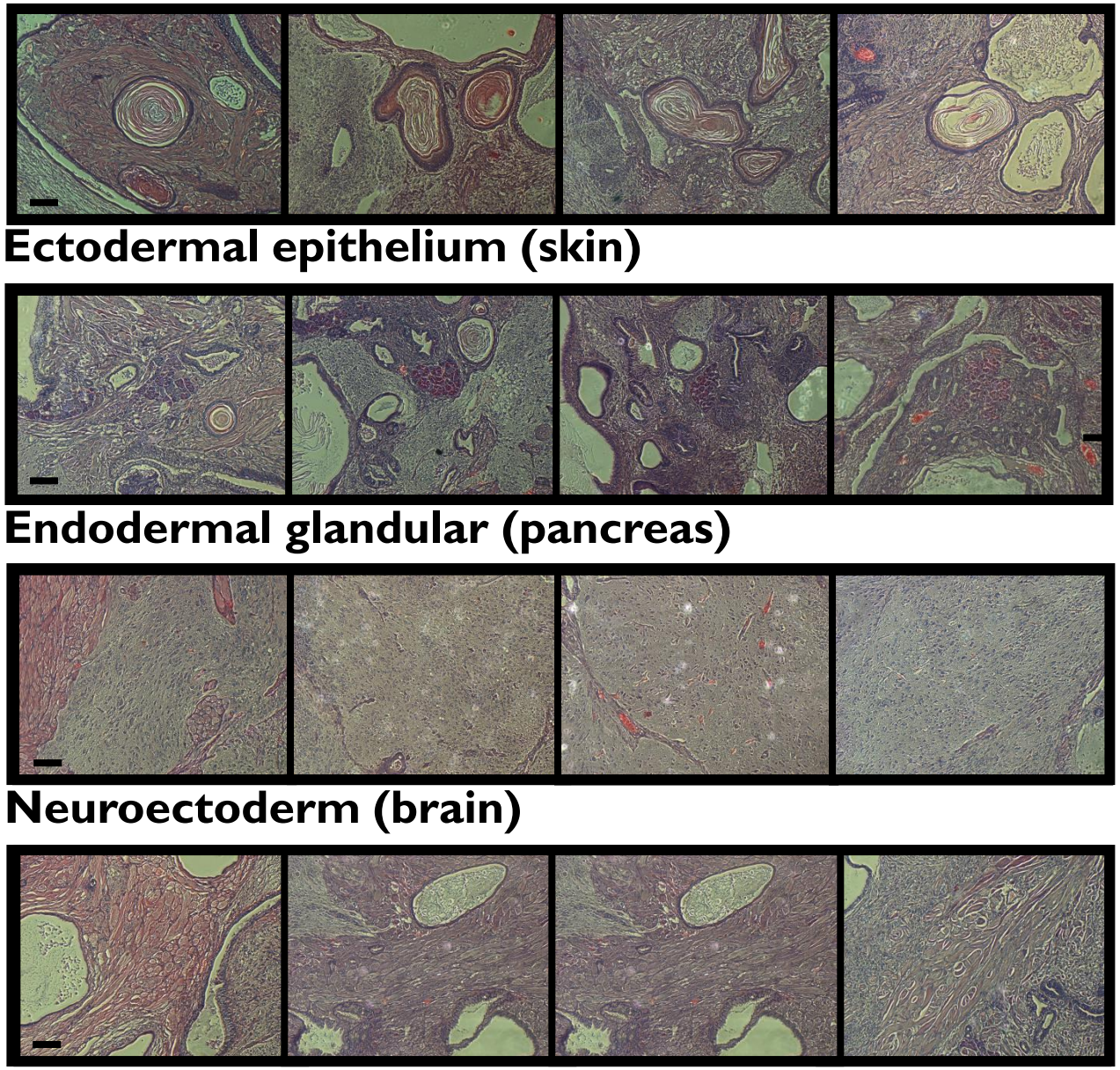
[4]Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University, Baltimore, Maryland, USA

[5]Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, California 94305, USA

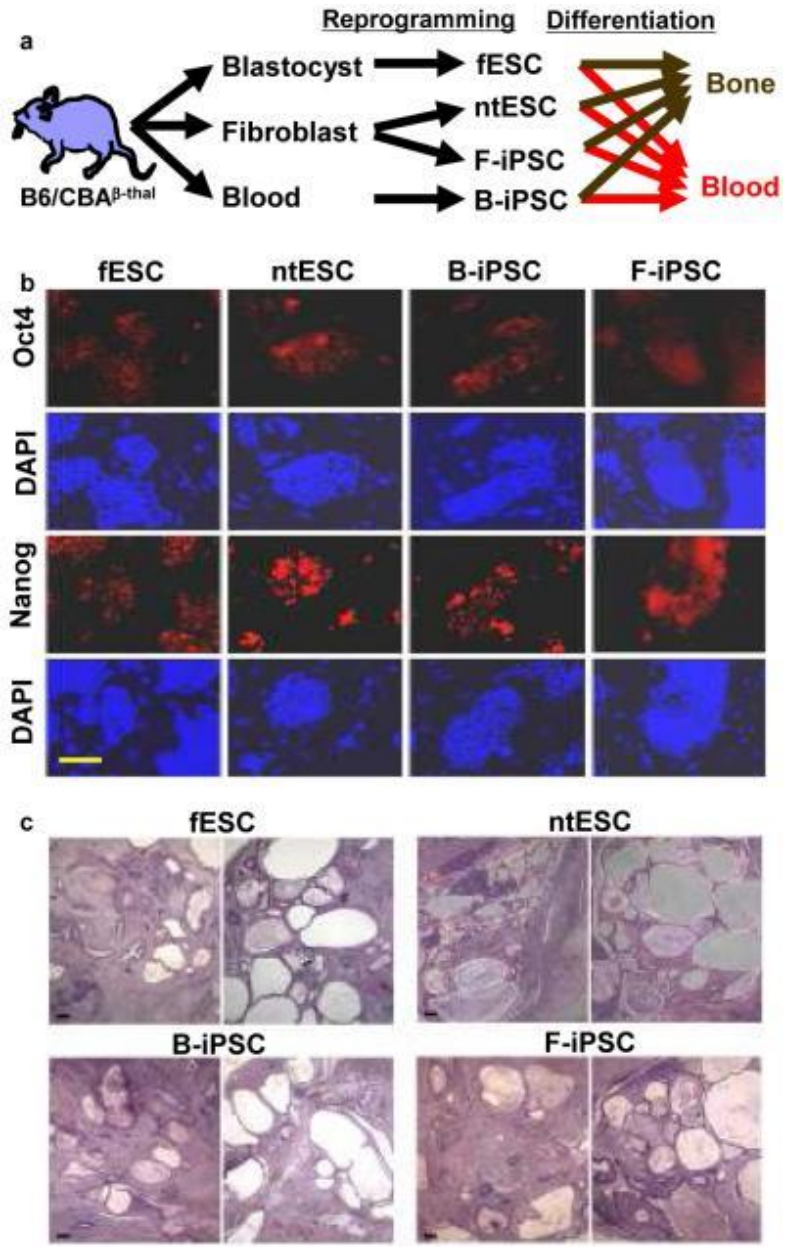
[6]Center for Systems Biology, Massachusetts General Hospital / Harvard Medical School, 185 Cambridge Street, CPZN 5206, Boston, MA 02114, USA

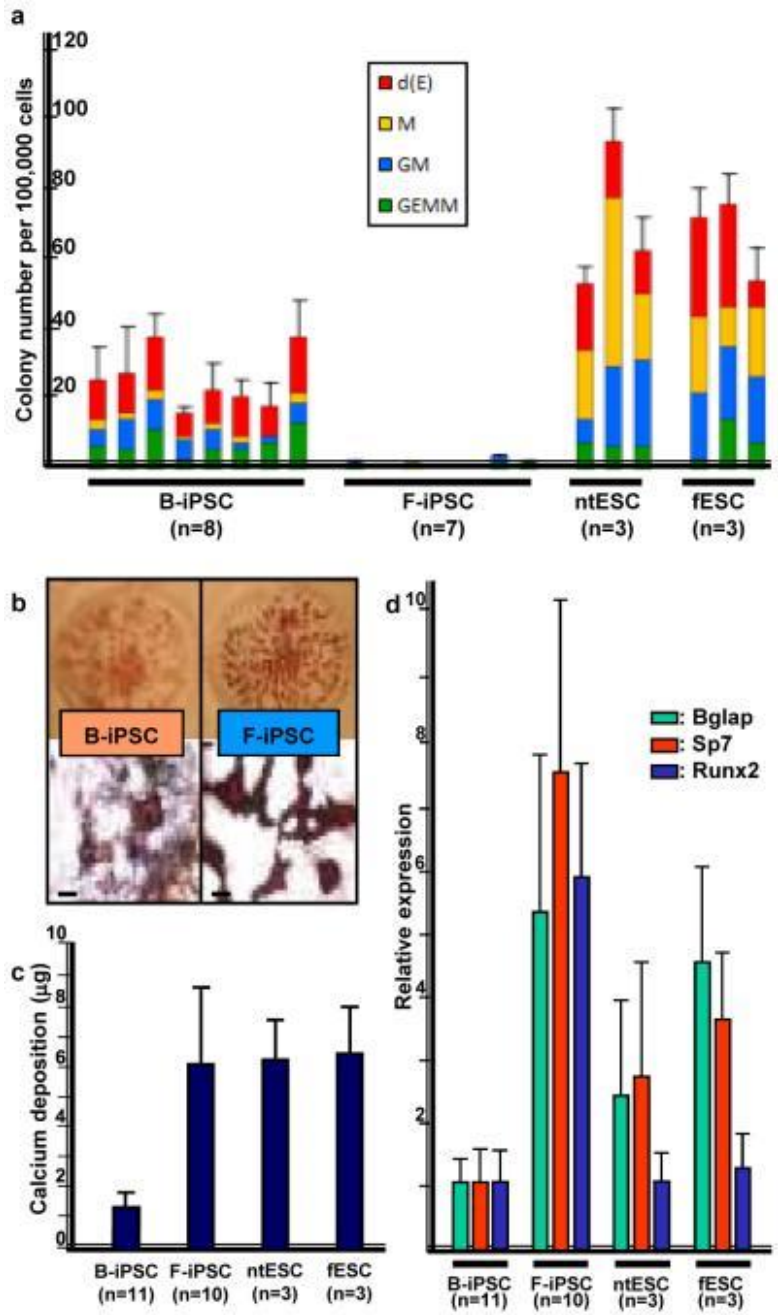
[7]Whitehead Institute for Biomedical Research, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

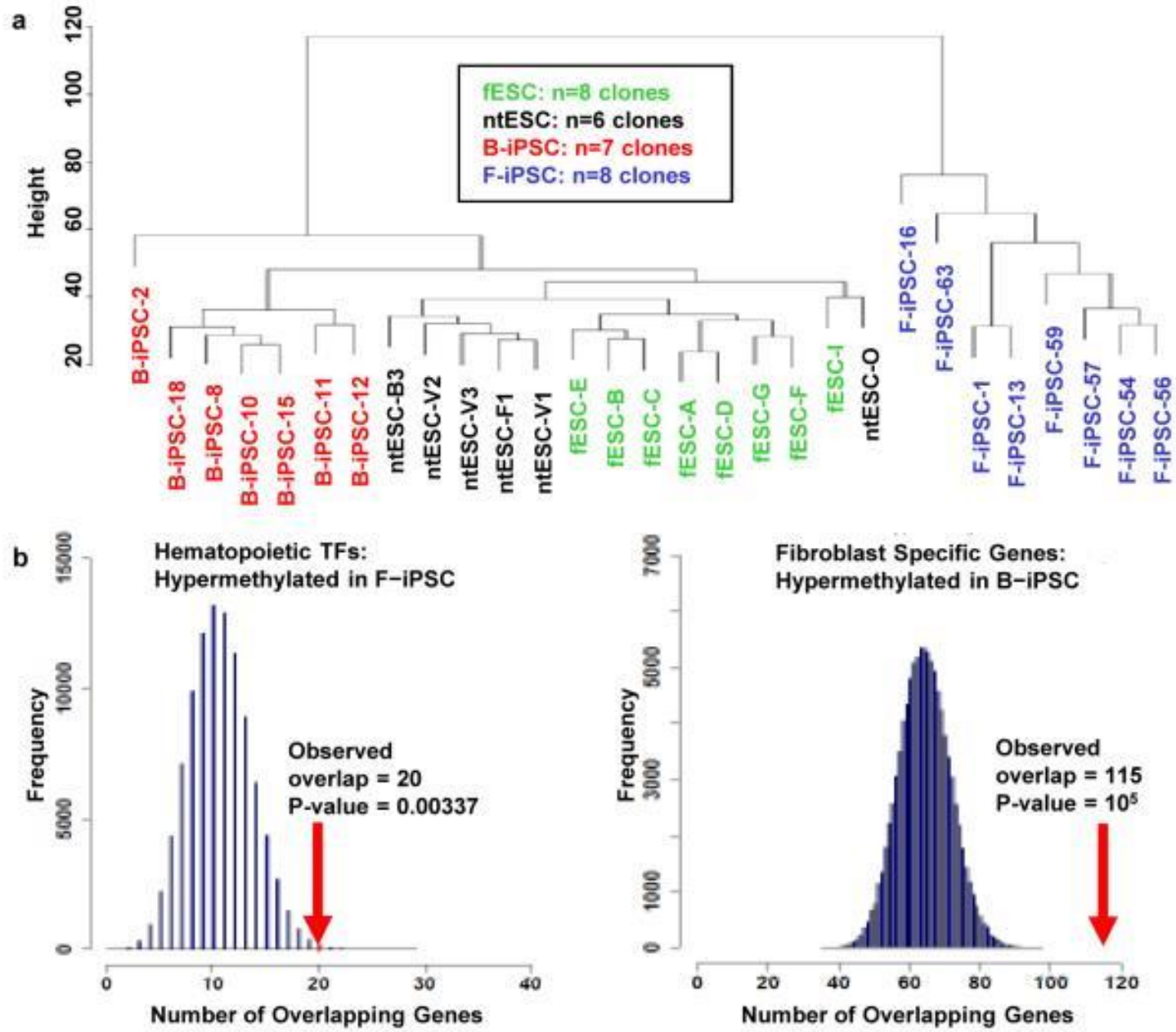
Supplementary Figures and Legends



Supplementary Figure 1 | Teratoma analysis of the fESC, ntESC, B-iPSC, and F-iPSC. Ectodermal epithelium (skin), Endodermal glandular (pancreas), Neuroectoderm (brain), Mesodermal stratified myoepithelium (muscle), and Ectodermal respiratory epithelium (Bronchi). Scale bar, 500μm.

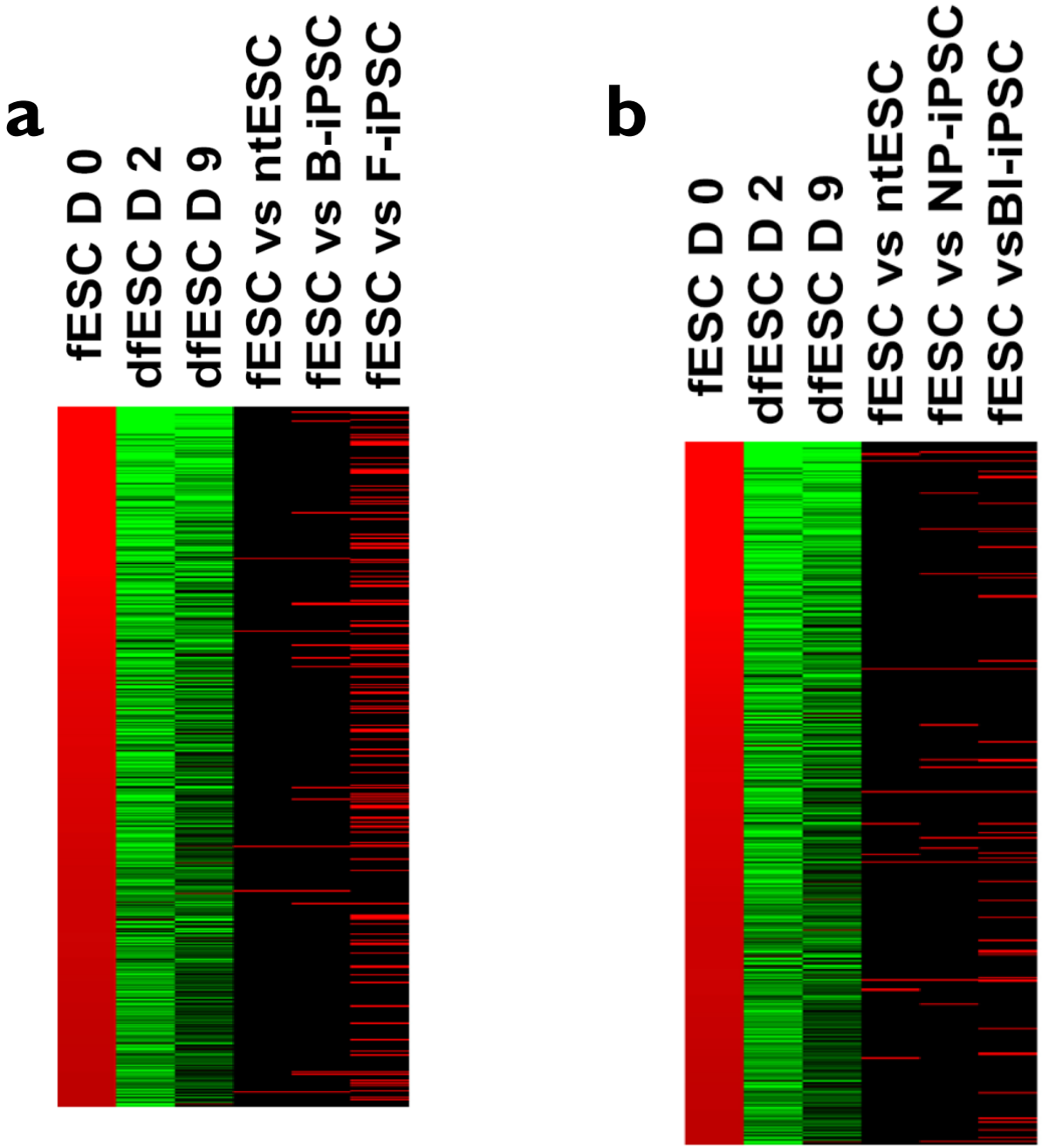

a
Reprogramming
Differentiation
B6/CBA β-thal
Blastocyst
Fibroblast
Blood
fESC
ntESC
F-iPSC
B-iPSC
Bone
Blood
b
fESC
ntESC
B-iPSC
F-iPSC
Oct4
DAPI
Nanog
DAPI
c
fESC
ntESC
B-iPSC
F-iPSC


a
Colony number per 100,000 cells
120
100
80
60
40
20
d(E)
M
GM
GEMM
B-iPSC
(n=8)
F-iPSC
(n=7)
ntESC
(n=3)
fESC
(n=3)
b
B-iPSC
F-iPSC
c
Calcium deposition (μg)
10
8
6
4
2
0
B-iPSC
(n=11)
F-iPSC
(n=10)
ntESC
(n=3)
fESC
(n=3)
d
Relative expression
10
8
6
4
2
: Bglap
: Sp7
: Runx2
B-iPSC
(n=11)
F-iPSC
(n=10)
ntESC
(n=3)
fESC
(n=3)

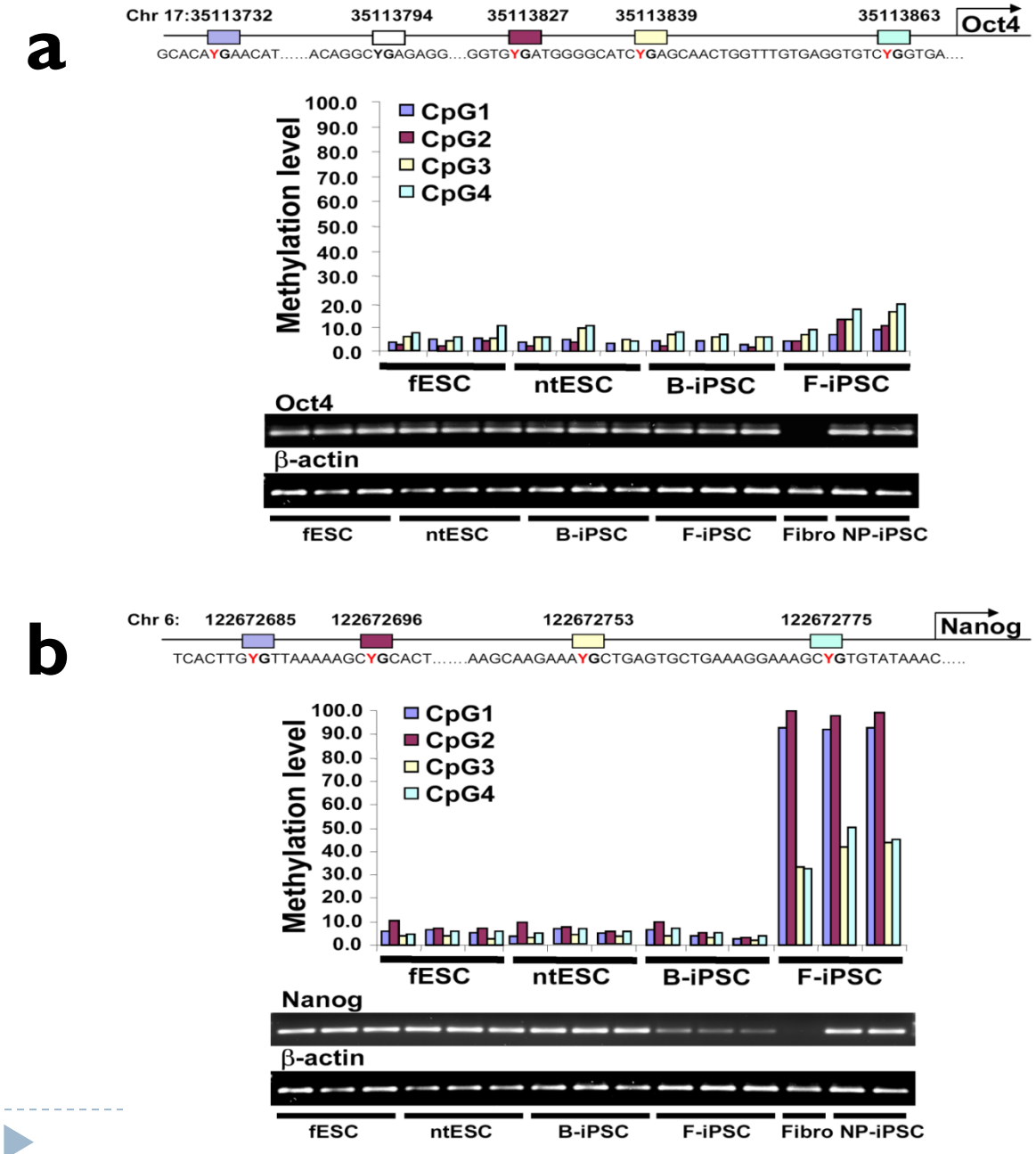

a
Height
120
100
80
60
40
20
fESC: n=8 clones
ntESC: n=6 clones
B-iPSC: n=7 clones
F-iPSC: n=8 clones
B-iPSC-2
B-iPSC-18
B-iPSC-8
B-iPSC-10
B-iPSC-15
B-iPSC-11
B-iPSC-12
ntESC-B3
ntESC-V2
ntESC-V3
ntESC-F1
ntESC-V1
fESC-E
fESC-B
fESC-C
fESC-A
fESC-D
fESC-G
fESC-F
fESC-I
ntESC-O
F-iPSC-16
F-iPSC-63
F-iPSC-1
F-iPSC-13
F-iPSC-59
F-iPSC-57
F-iPSC-54
F-iPSC-56
b
Hematopoietic TFs:
Hypermethylated in F−iPSC
Frequency
15000
10000
5000
0
Observed
overlap = 20
P-value = 0.00337
0
10
20
30
40
Number of Overlapping Genes
Fibroblast Specific Genes:
Hypermethylated in B−iPSC
Frequency
7000
5000
3000
1000
0
Observed
overlap = 115
P-value = 10^5
0
20
40
60
80
100
120
Number of Overlapping Genes

Rank	Gene	Function	Reference
1	Kcnrg		
2	Mast1		
3	Mab21L1	Osteogenetic differentiation	4
4	Atbf1	Myb mediated hematopoietic growth regulation	5
5	Hand1	Cardiac development	6
6	Zfp423	Enhance Hematopoietic activity	7
7	Pcdhga10	protocadherin	8
8	Dlx1	Hematopoietic development with BMP4	9
9	Pim1	Hematopoietic proliferation	10
10	Efnb2	Developmental events, especially in the nervous system and in erythropoiesis	11
11	Asns	Hematopoietic proliferation	12
12	Tradd	programmed cell death	13
13	Ebf2	Osteogenetic differentiation	14
14	Slc13a4	Sodium/sulfate cotransporter	15
15	Osr2	Osteogenetic development	16
16	Igsf4c	Immunoglobulin superfamily	17
17	Meis1	Definitive hematopoiesis	18
18	CD37	T cell B cell interaction and proliferation	19 20
19	Slc38a4		
20	Pcdhga12	Protocadherin	8
21	Pcdhga7	Protocadherin	8
22	Map2k7	Hematopoietic growth	21
23	Sall4	Bmi-1 mediated hematpoietic self-renewal	22
24	Pcdhgb6	Protocadherin	8

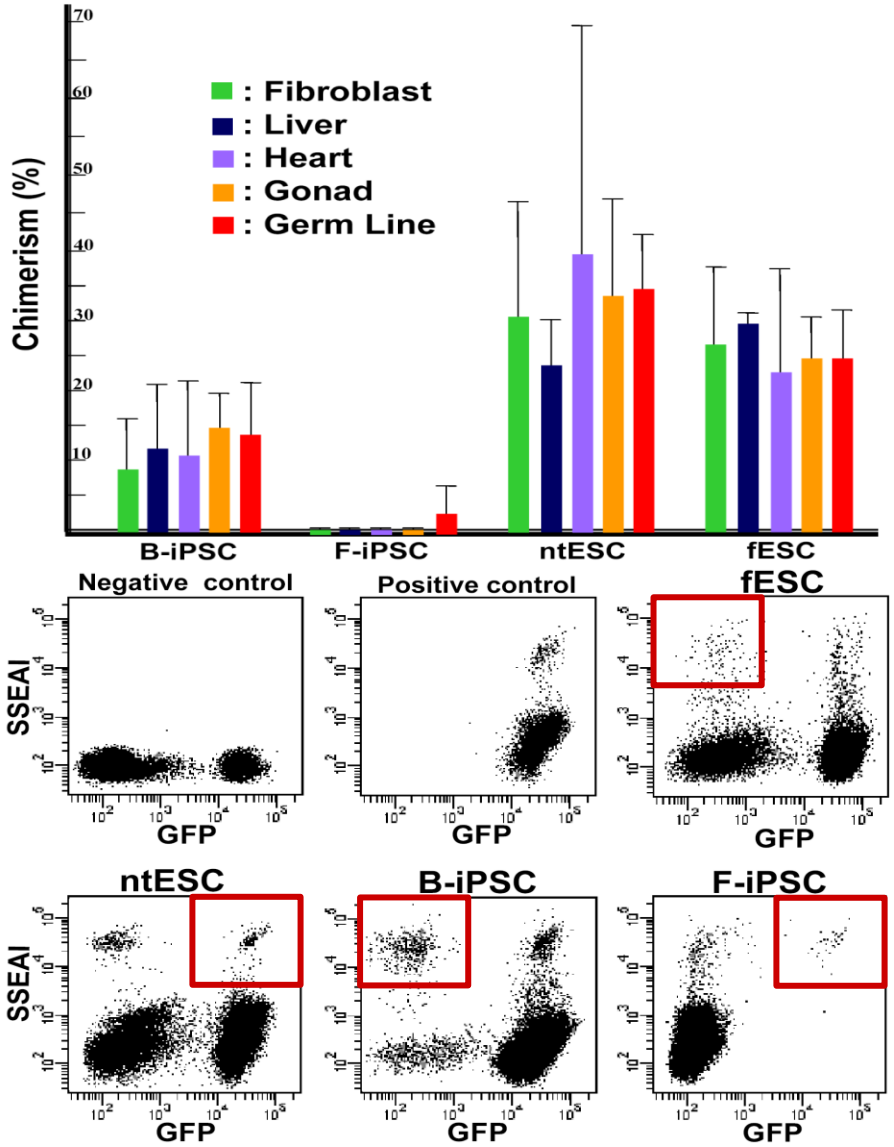
Supplementary Table 2 | Top 24 Differentially Methylated Regions (DMRs) between B-iPSC and F-iPSC. Blood-related genes are shaded in red; bone-related genes are shaded in light brown.



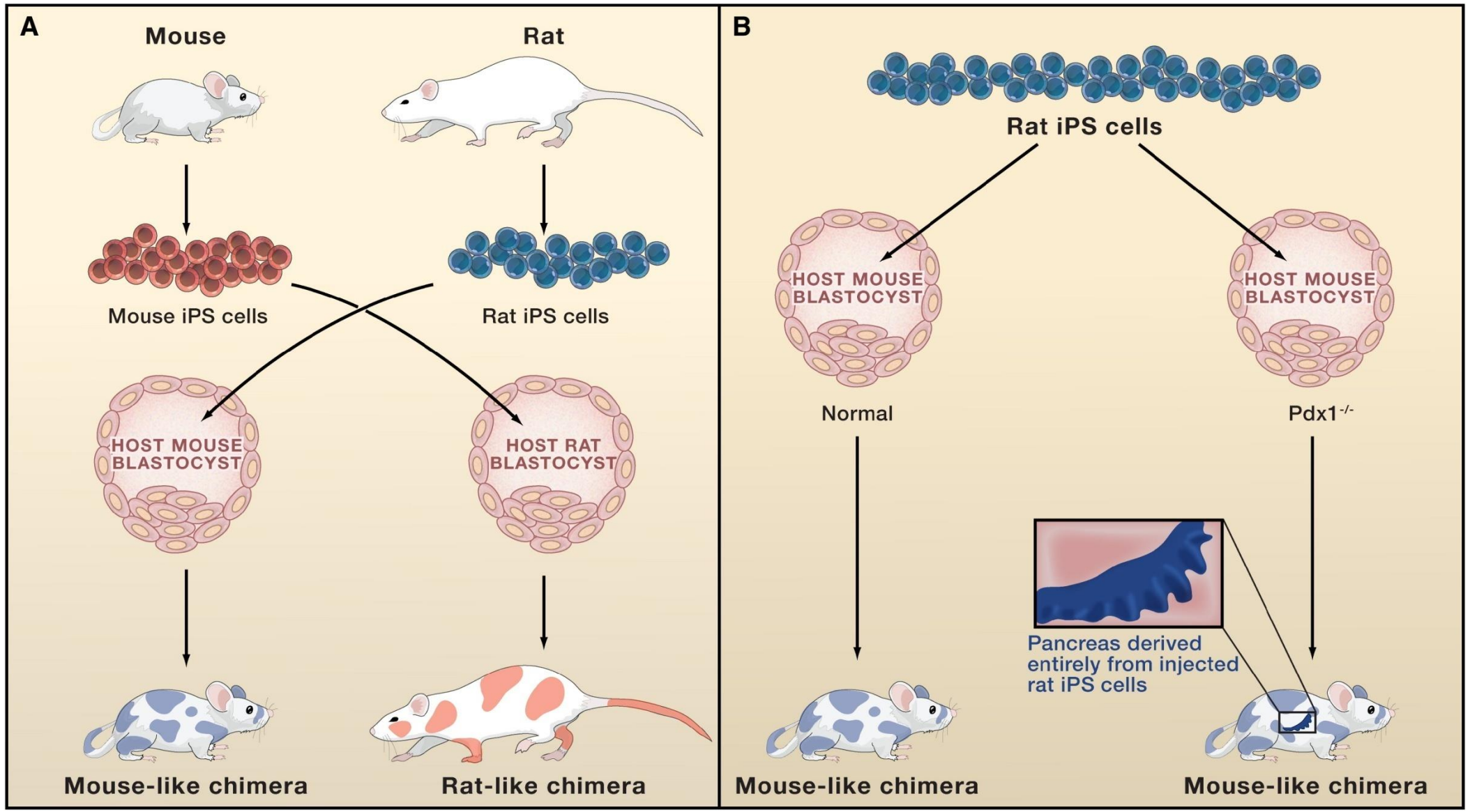
Supplementary Figure 5 | Overlap of DMRs with loci of genes showing fESC-specific gene expression (determined from compiled microarray data[2]**. Heat maps reflect expression values of fESC-specific genes in undifferentiated Stat**e (fESC D0; top 5% highly expressed genes**; 554 genes)** and after differentiation for 2 and 9 days (differentiated fESC day 2; dfESC D2 and day 9; dfESC D9). **a,** Red bars in the right three lanes indicate number of fESC-specific genes that overlap with DMRs (ntESC, n=5; B-iPSC, n=18; F-iPSC, n=114). **b,** Red bars in the right three lanes indicate number of fESC-specific genes that overlap with DMRs (ntESC, n=12; NP-iPSC, n=16; BI-iPSC, n=45).

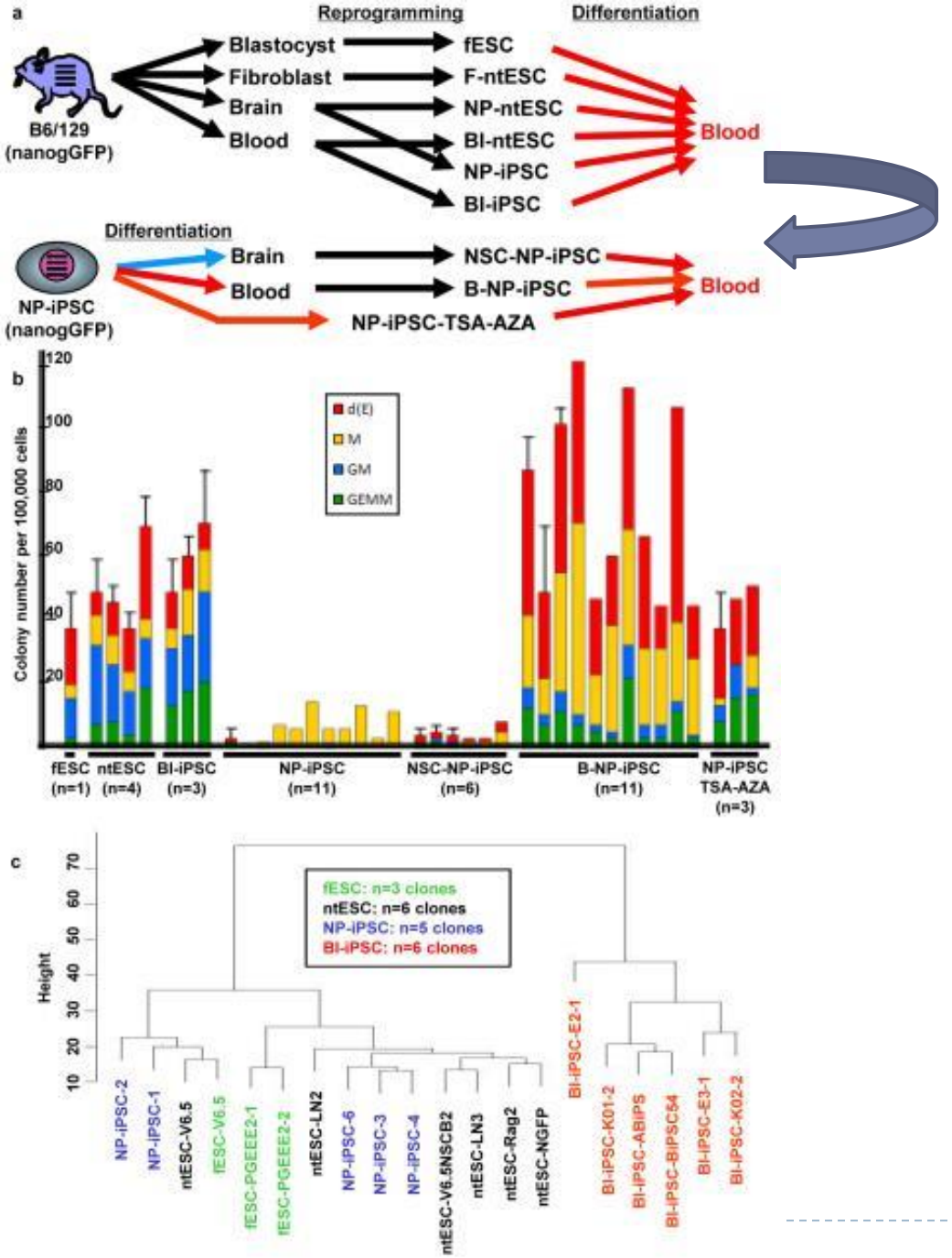


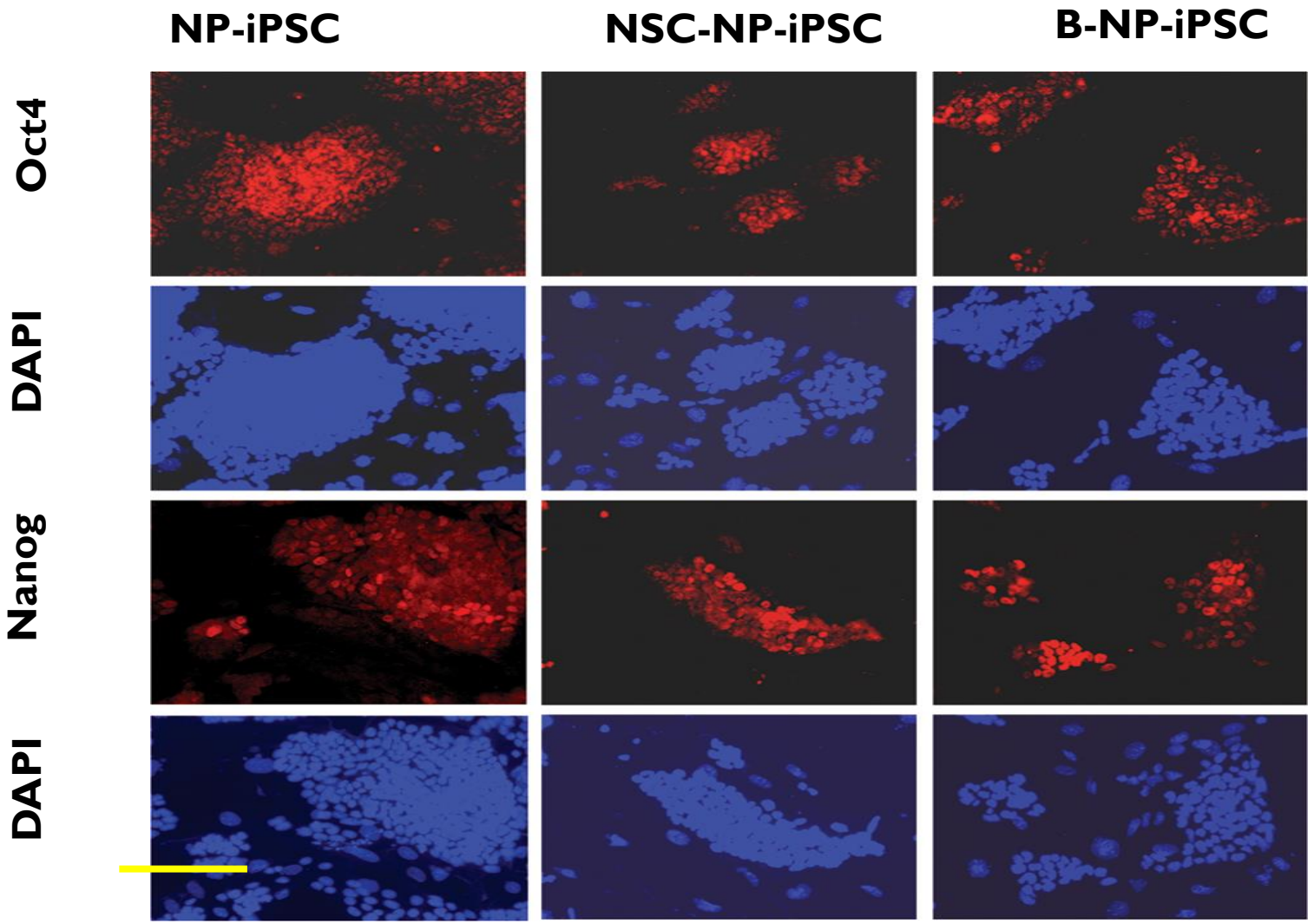
Supplementary Figure 6 | DNA demethylation of promoters and gene expression on the selected pluripotent gene loci. a, Oct4 **b,** Nanog. Schematic structure of the promoters are shown on top, and methylation status of the CpG sites measured by bisulfite pyrosequencing with three independent samples of fESC, ntESC, B-iPSC, and F-iPSC are shown in middle graphs. Detection of Oct4 and Nanog gene expression by RT-PCR with three independent samples of fESC, ntESC, B-iPSC, and F-iPSC are shown below each panel.



Supplementary Figure 7 | Chimera analysis of the fESC, ntESC, B-iPSC, and F-iPSC (refer to Fig. 1a). a, Organ chimerism. B6CBA-derived cells were injected into blastocysts and transferred to pseudopregnant mice (N=3 clones of each stem cell type). Organs from E12.5 embryo (B-iPSC, n=14; F-iPSC, n=8; ntESC, n=15; fESC, n=13) were analyzed by flow cytometry to determine % GFP+ cells. Positive control: SSEA1 staining of gonad cells from GFP+ transgenic mouse.


A
Mouse
Rat
Mouse iPS cells
Rat iPS cells
HOST MOUSE BLASTOCYST
HOST RAT BLASTOCYST
Mouse-like chimera
Rat-like chimera
B
Rat iPS cells
HOST MOUSE BLASTOCYST
HOST MOUSE BLASTOCYST
Normal
Pdx1-/-
Pancreas derived entirely from injected rat iPS cells
Mouse-like chimera
Mouse-like chimera


a
Reprogramming
Differentiation
B6/129
(nanogGFP)
Blastocyst
Fibroblast
Brain
Blood
fESC
F-ntESC
NP-ntESC
Bl-ntESC
NP-iPSC
Bl-iPSC
Blood
Differentiation
NP-iPSC
(nanogGFP)
Brain
Blood
NSC-NP-iPSC
B-NP-iPSC
NP-iPSC-TSA-AZA
Blood
b
Colony number per 100,000 cells
d(E)
M
GM
GEMM
fESC (n=1)
ntESC (n=4)
Bl-iPSC (n=3)
NP-iPSC (n=11)
NSC-NP-iPSC (n=6)
B-NP-iPSC (n=11)
NP-iPSC TSA-AZA (n=3)
c
Height
fESC: n=3 clones
ntESC: n=6 clones
NP-iPSC: n=5 clones
Bl-iPSC: n=6 clones
NP-iPSC-2
NP-iPSC-1
ntESC-V6.5
fESC-V6.5
fESC-PGEEE2-1
fESC-PGEEE2-2
ntESC-LN2
NP-iPSC-5
NP-iPSC-3
NP-iPSC-4
ntESC-V6.5NSCB2
ntESC-LN3
ntESC-Rag2
ntESC-NGFP
Bl-iPSC-E2-1
Bl-iPSC-K01-2
Bl-iPSC-ABIP5
Bl-iPSC-BIPSC54
Bl-iPSC-E3-1
Bl-iPSC-K02-2



Supplementary Figure 8 | Immunohistochemistry of NP-iPSC, NSC-NP-iPSC, and B-NP-iPSC for OCT4 and NANOG expression, as indicated. 4,6-Diamidino-2-phenylindole (DAPI) staining for total cell content. Fibroblasts surrounding pluripotent colonies serve as negative controls for immunohistochemistry staining. Scale bar, 200μm.

a

Comparison	Number of DMRs	Methylation	Number of DMRs
fESC *vs.* F-iPSC	5304	fESC>F-iPSC	1955
		fESC<F-iPSC	3349
fESC *vs.* B-iPSC	694	fESC>B-iPSC	178
		fESC<B-iPSC	516
fESC *vs.* ntESC	229	fESC>ntESC	173
		fESC<ntESC	56
F-iPSC *vs.* B-iPSC	5202	F-iPSC>B-iPSC	2850
		F-iPSC<B-iPSC	2352
F-iPSC *vs.* ntESC	6255	F-iPSC>ntESC	4077
		F-iPSC<ntESC	2178
B-iPSC *vs.* ntESC	995	B-iPSC>ntESC	897
		B-iPSC<ntESC	98

*area cutoff of 2.0.

b

Comparison	Number of DMRs	Methylation	Number of DMRs
BI-iPSC *vs.* fESC	1485	BI-iPSC>fESC	1423
		BI-iPSC<fESC	62
BI-iPSC *vs.* NP-iPSC	2344	BI-iPSC>NP-iPSC	2326
		BI-iPSC<NP-iPSC	18
BI-iPSC *vs.* ntESC	3053	BI-iPSC>ntESC	3000
		BI-iPSC<ntESC	53
fESC *vs.* NP-iPSC	553	fESC>NP-iPSC	136
		fESC<NP-iPSC	417
fESC *vs.* ntESC	679	fESC>ntESC	399
		fESC<ntESC	280
NP-Ipsc *vs.* ntESC	571	NP-iPSC>ntESC	469
		NP-iPSC<ntESC	102

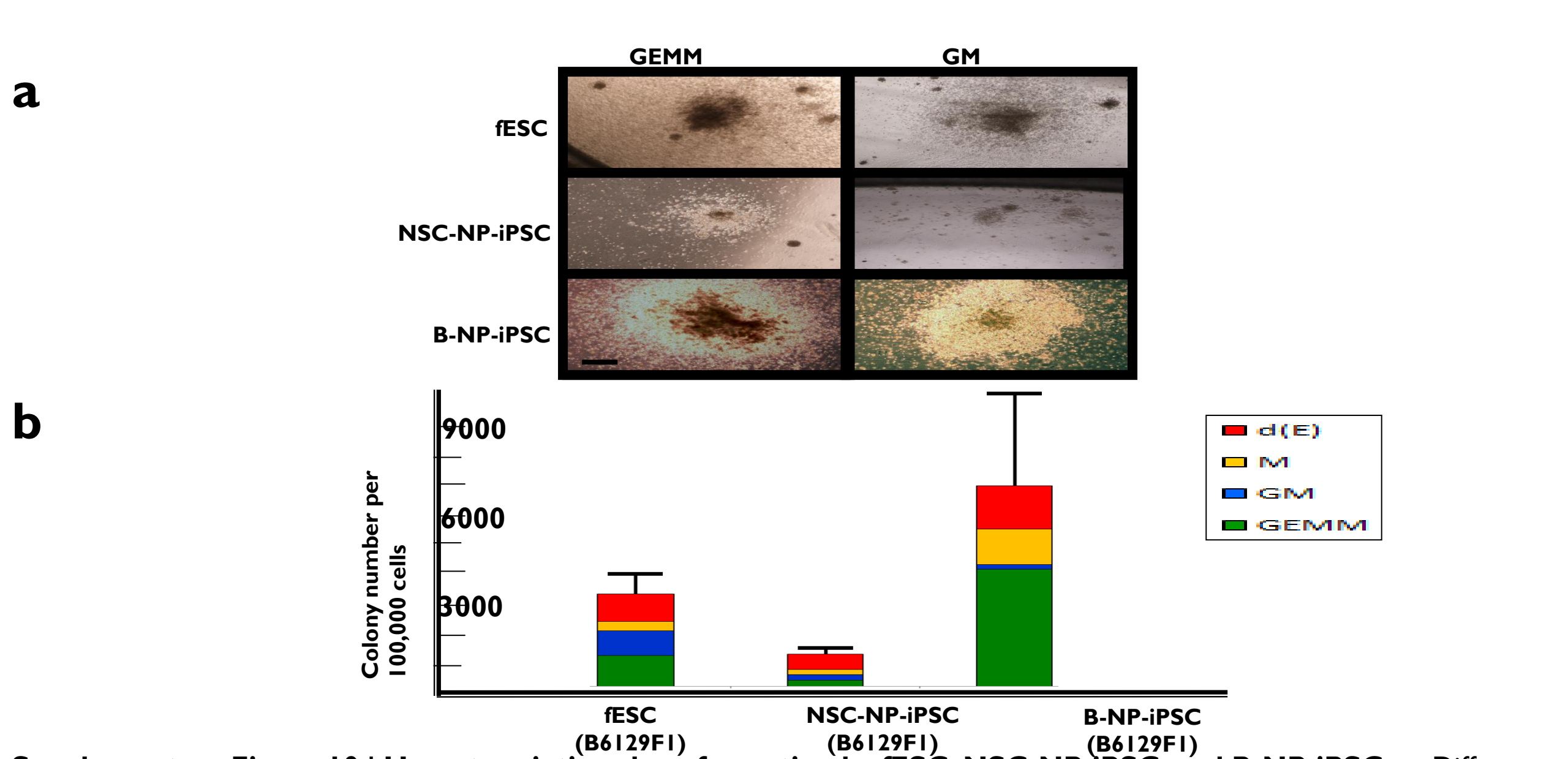
*area cutoff of 2.0.

c

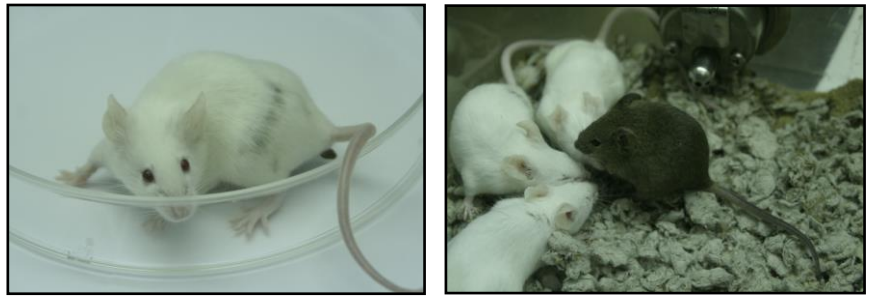
Comparison	Number of DMRs	Methylation	Number of DMRs
NP-iPSC *vs.* NSC-NP-iPSC	107	NP-iPSC>NSC-NP-iPSC	46
		NSC-NP-iPSC>NP-iPSC	61
NP-iPSC *vs.* B-NP-iPSC	803	NP-iPSC>B-NP-iPSC	593
		B-NP-iPSC>NP-iPSC	210
NP-iPSC *vs.* NP-iPSC-TSA-AZA	938	NP-iPSC >NP-iPSC-TSA-AZA	626
		NP-iPSC-TSA-AZA >NP-iPSC	312
NSC-NP-iPSC vs. B-NP-iPSC	688	NSC-NP-iPSC >B-NP-iPSC	632
		B-NP-iPSC >NSC-NP-iPSC	56

* area cutoff of 2.0.

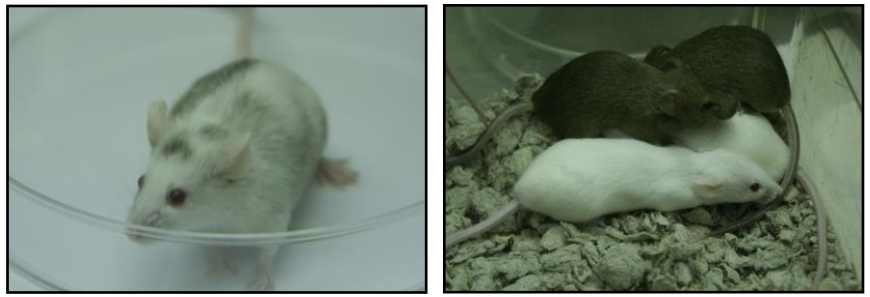
Supplementary Table 1 | . DMRs by CHARM analysis. a, fESC, ntESC, B-iPSC, and F-iPSC (refer to Fig. 1a), **b,** fESC, ntESC, NP-iPSC, and BI-iPSC (refer to Fig. 4a upper schema), **c,** NP-iPSC, NSC-NP-iPSC, NP-iPSC-TSA-AZA, and B-NP-iPSC (refer to Fig. 4a lower schema).



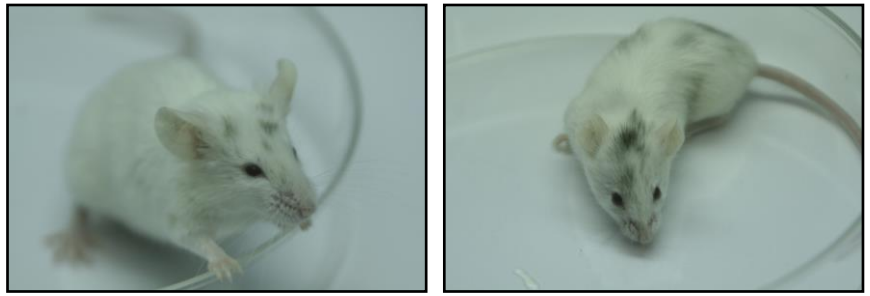
Supplementary Figure 10 | Hematopoietic colony formation by fESC, NSC-NP-iPSC, and B-NP-iPSC. a, Different sizes of GEMM and GM colonies in methylcellulose cultures of fESC, NSC-NP-iPSC, and B-NP-iPSC. (40X magnification). Scale bar, 500μm. **b,** Average cell number per colony among 20 randomly picked colonies from fESC, NSC-NP-iPSC, and B-NP-iPSC. Error bars = s.d.



fESC Chimera and germ line transmission



ntESC chimera and germ line transmission



B-NP-iPSC chimera NSC-NP-iPSC chimera

Supplementary Figure 9 | Mouse chimerism and germ line transmission of the fESC, ntESC, B-NP-iPSC, and NSC-NP-iPSC (refer to Fig. 4a).