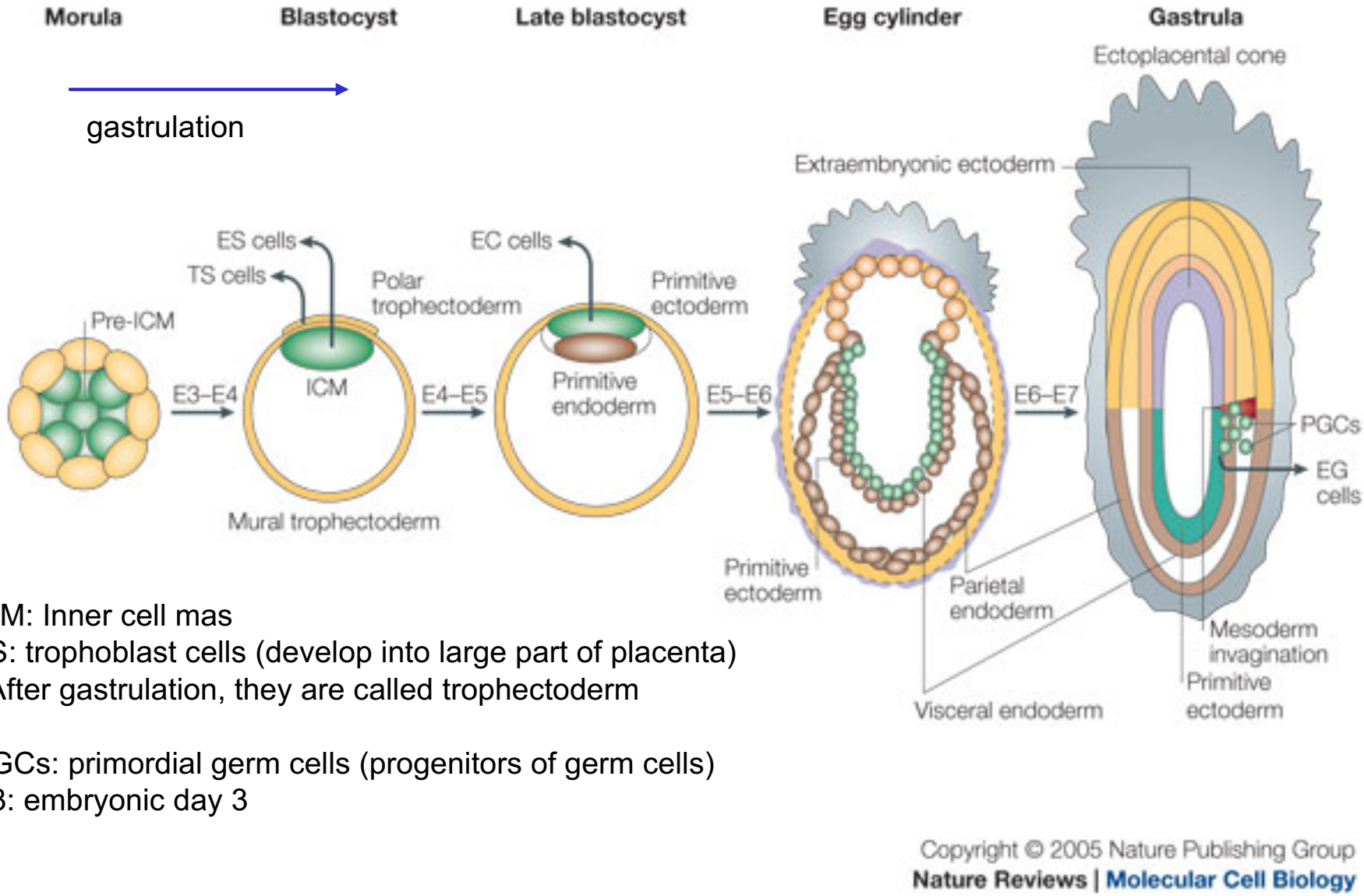
V12: subjects of minitest #3

- Most of lecture V9
- Most of lecture V10
- **No** material from **papers #7, #9 and #11**
- Selected material of papers #8 and #10 (presented in V11)
- Selected material of paper #13 (will be presented today in V12)

(review V9): Embryonic development of mouse

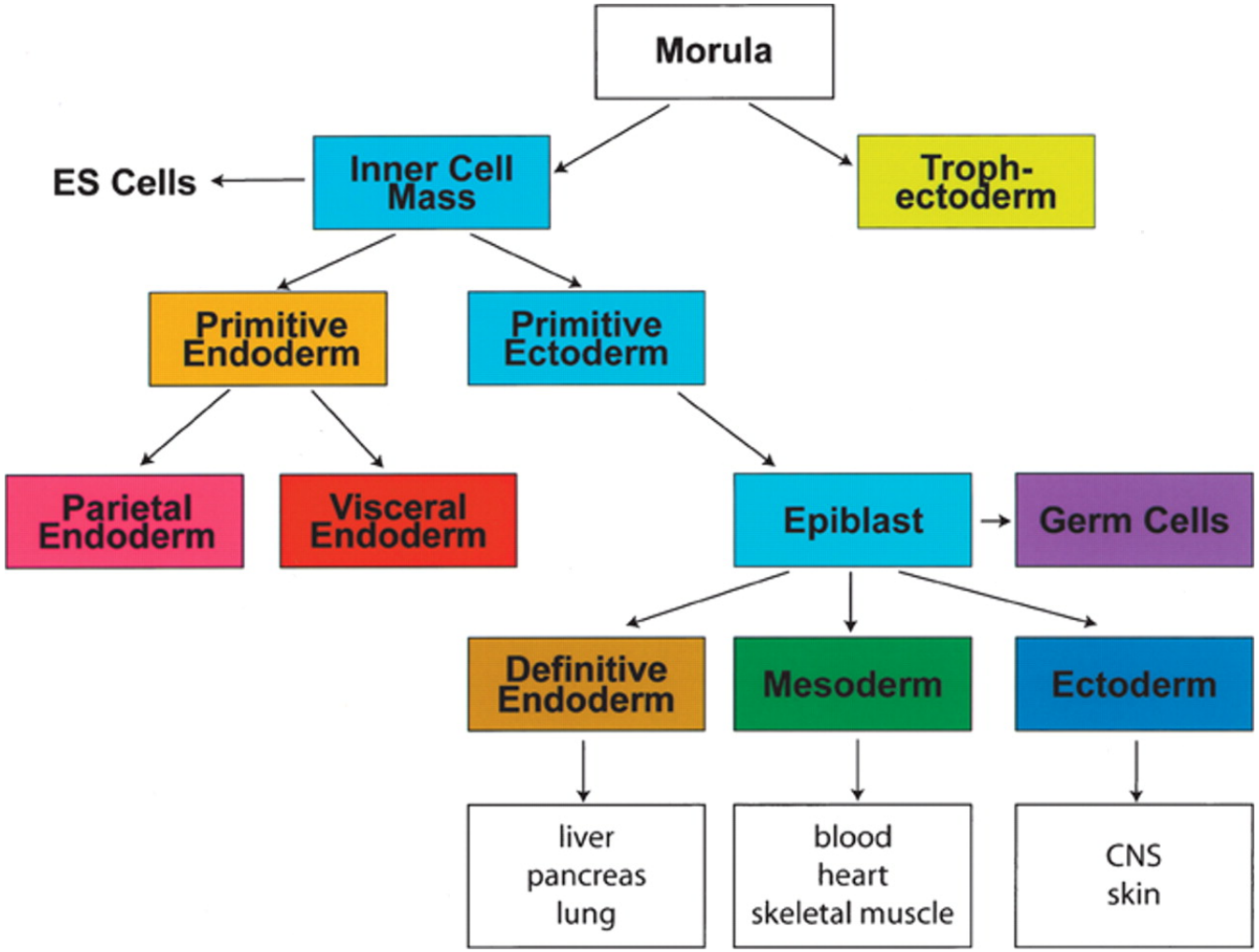


ICM: Inner cell mas
TS: trophoblast cells (develop into large part of placenta)
- After gastrulation, they are called trophectoderm

PGCs: primordial germ cells (progenitors of germ cells)
E3: embryonic day 3

Boiani & Schöler, Nat Rev Mol Cell Biol 6, 872 (2005)

Cell populations in early mouse development



Scheme of early mouse development depicting the relationship of early cell populations to the primary germ layers

Keller, Genes & Dev. (2005) 19: 1129-1155

Types of body cells

3 basic categories of cells make up the mammalian body:

germ cells (oocytes and sperm cells)
somatic cells, and
stem cells.

Each of the approximately 100 trillion (10^{14}) cells in an adult human has its own copy or copies of the genome except certain cell types, such as red blood cells, that lack nuclei in their fully differentiated state.

Most cells are **diploid**; they have two copies of each chromosome.

Cells differentiate to specialize for different functions.

Somatic cells make up most of the human body, such as skin and muscle cells.

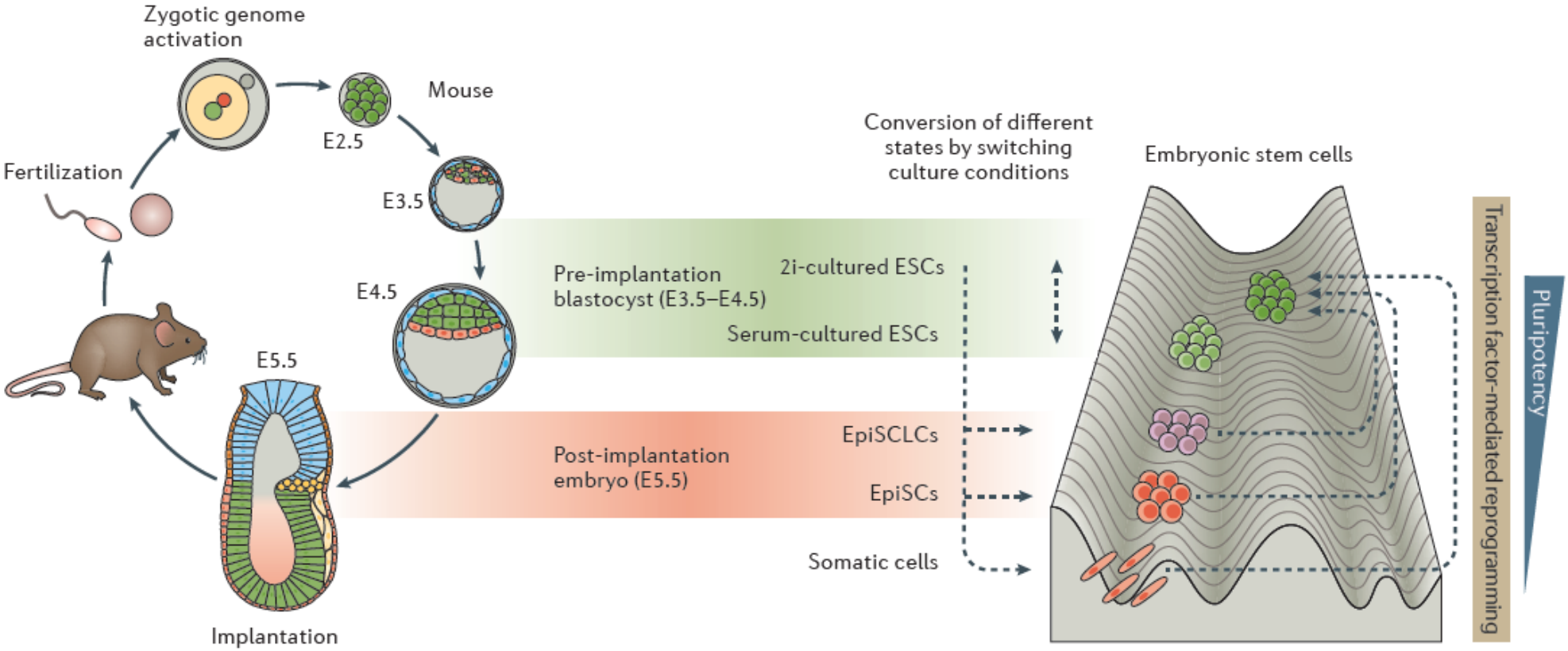
www.wikipedia.org

Different states of pluripotency

E4.5 epiblast cells: represent ground-state pluripotency

Implantation: stage of pregnancy at which the blastocyst adheres to the wall of the **uterus**.

After implantation (E5.5): **epiblast cells** undergo a strong wave of epigenetic reprogramming. They are now „primed“.



Atlasi & Stunnenberg, *Nature Rev Genet* **18**, 643–658 (2017)

Waddington's epigenetic landscape for embryology

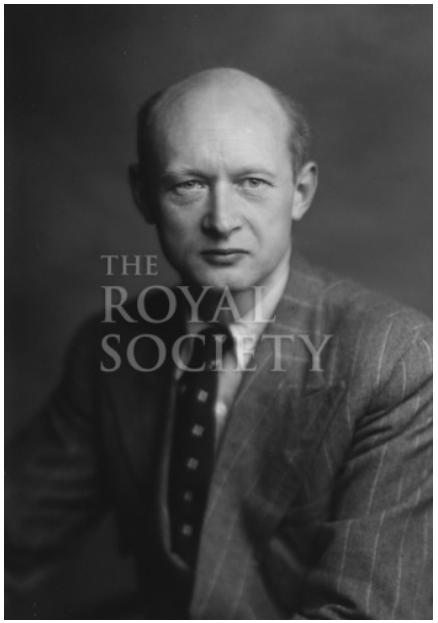
Slack, Nature Rev Genet 3, 889-895 (2002)

Waddington worked in **embryology**

a) is a painting by John Piper that was used as the frontispiece for Waddington's book *Organisers and Genes*.
It represents an epigenetic landscape.

Developmental pathways that could be taken by each cell of the embryo are metaphorically represented by the path taken by water as it flows down the valleys.

b) Later depiction of the epigenetic landscape. The ball represents a cell, and the bifurcating system of valleys represents bundles of trajectories in state space.

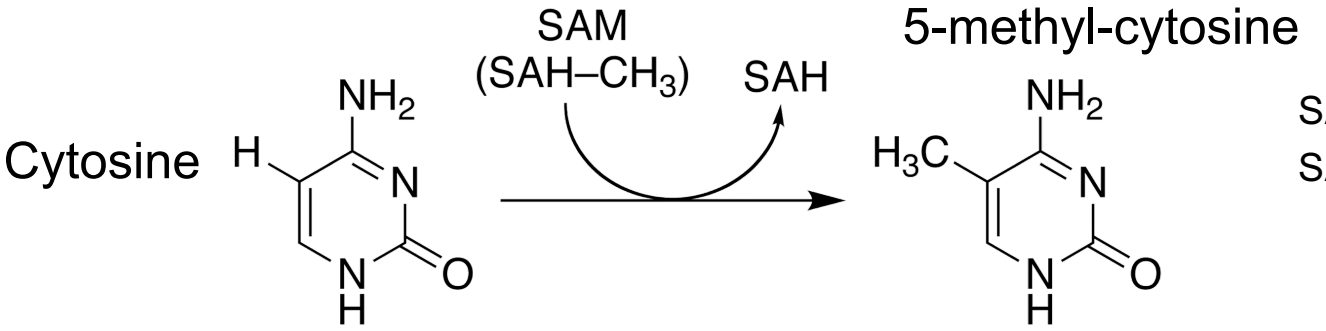


Conrad Hal Waddington (1905 – 1975)
pictures.royalsociety.org

Cytosine methylation

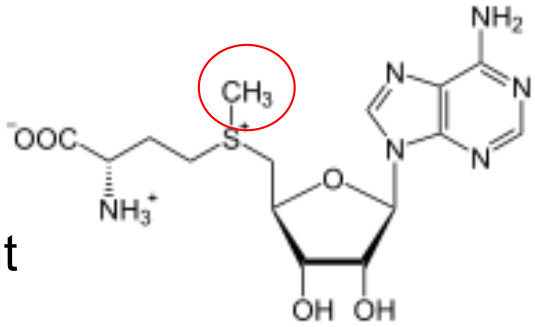
Observation: 3-6 % of all cytosines are methylated in human DNA.

This methylation occurs (almost) exclusively when cytosine is followed by a guanine base -> **CpG dinucleotide**.



SAM: S-adenosyl-methionine

SAH: S-adenosyl-homocysteine



Mammalian genomes contain much fewer (only 20-25 %) of the CpG dinucleotide than is expected by the G+C content (we expect 1/16 ≈ 6% for any random dinucleotide).

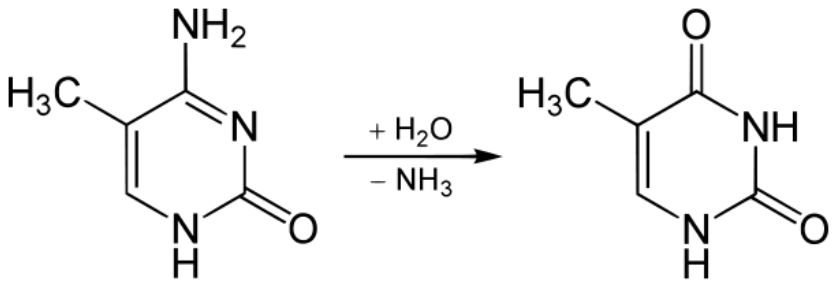
This is typically explained in the following way:

As most CpGs serve as targets of DNA methyltransferases, they are usually methylated …. (see following page)

Esteller, Nat. Rev. Gen. 8, 286 (2007)
www.wikipedia.org

Cytosine methylation

But 5-Methylcytosine can easily **deaminate** to **thymine**.

5-methyl-cytosine 

thymine

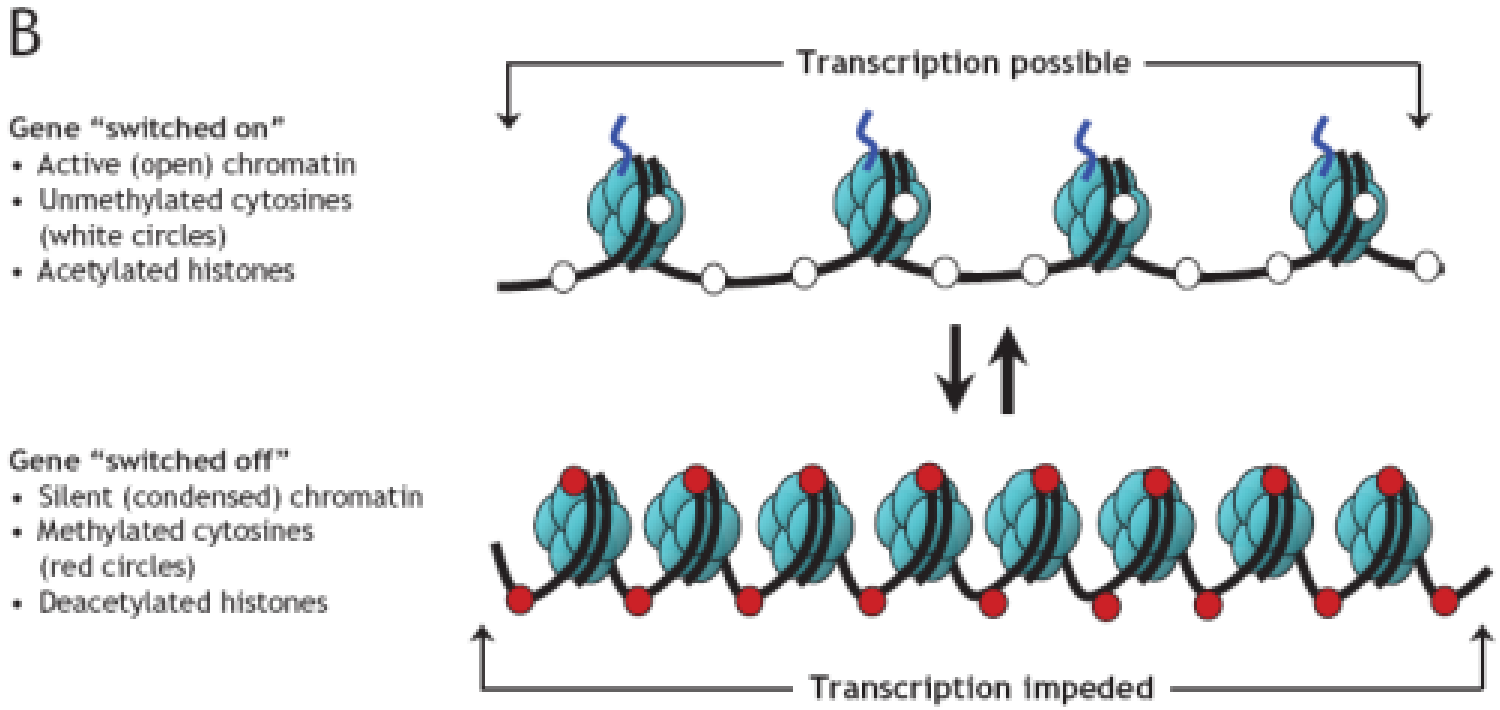
If this mutation is not repaired, the affected CpG is permanently converted to TpG (or CpA if the transition occurs on the reverse DNA strand).

Hence, methylCpGs represent **mutational hot spots** in the genome.
If such mutations occur in the germ line, they become heritable.

A constant loss of CpGs over thousands of generations
can explain the low frequency of this
special dinucleotide in the genomes of human and mouse.

Esteller, Nat. Rev. Gen. 8, 286 (2007)
www.wikipedia.org

chromatin organization affects gene expression



Schematic of the reversible changes in chromatin organization that influence gene expression:

genes are expressed (switched on) when the chromatin is **open** (active), and they are inactivated (switched off) when the chromatin is **condensed** (silent).

White circles = unmethylated cytosines;

red circles = methylated cytosines.

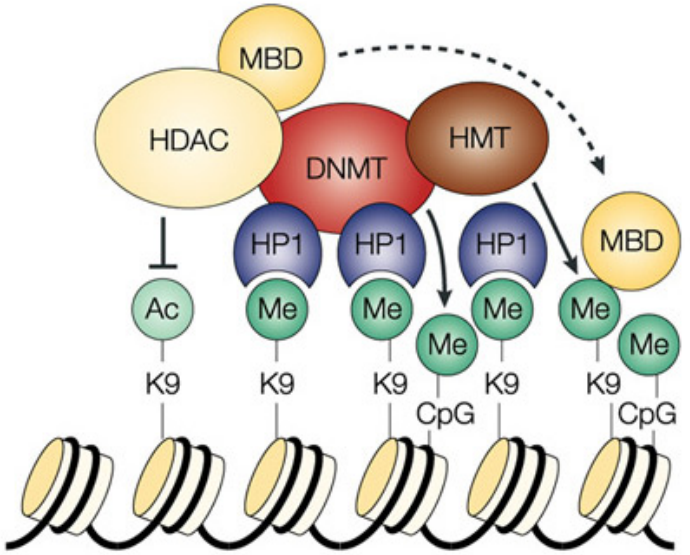
Rodenhiser, Mann, CMAJ 174, 341 (2006)

Enzymes that control DNA methylation and histone modfications

These dynamic chromatin states are controlled by reversible epigenetic patterns of **DNA methylation** and **histone modifications**.

Enzymes involved in this process include
- DNA methyltransferases (DNMTs),
- histone deacetylases (HDACs),
- histone acetylases,
- histone methyltransferases (HMT) and the
- methyl-binding domain protein MECP2

with its methyl-binding domain (MBD) that binds specifically to me-cytosine.



HP1: heterochromatin protein 1

Rodenhiser, Mann, CMAJ 174, 341 (2006)
Feinberg AP & Tycko P (2004) Nature Reviews: 143-153

DNA methylation

Typically, unmethylated clusters of CpG pairs are located in
tissue-specific genes and in essential **housekeeping genes.**

(House-keeping genes are involved in routine maintenance roles and are expressed in most tissues.)

These clusters, or **CpG islands**, are targets for proteins
that bind to unmethylated CpGs and initiate gene transcription.

In contrast, **methylated CpGs** are generally associated with silent DNA,
can block methylation-sensitive proteins and can be easily mutated.

The **loss** of normal DNA methylation patterns is the
best understood epigenetic cause of **disease**.

In animal experiments, the removal of genes that encode DNMTs is lethal;
in humans, overexpression of these enzymes has been linked
to a variety of cancers.

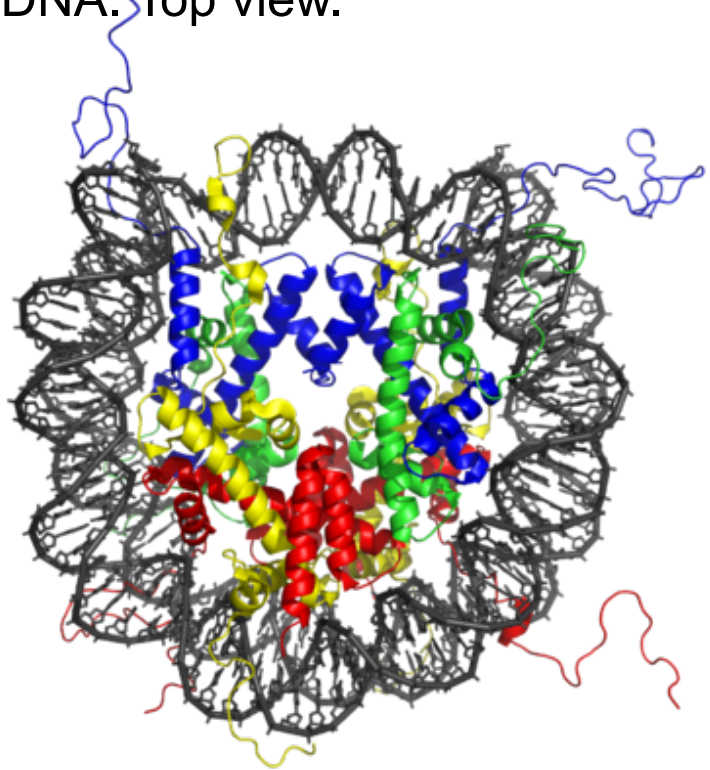
Rodenhiser, Mann, CMAJ 174, 341 (2006)

The histone code

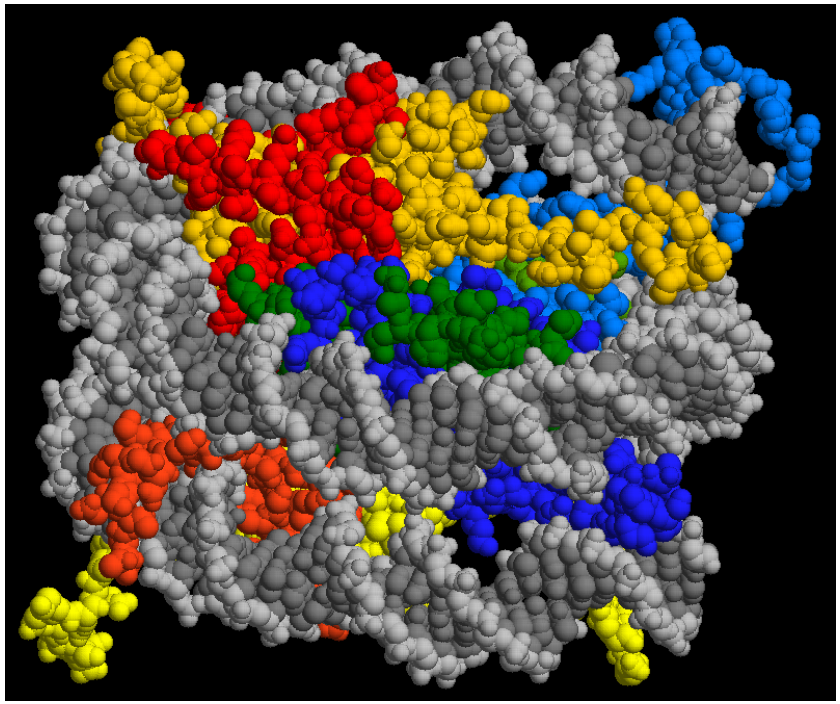
The DNA of eukaryotic organisms is packaged into chromatin, whose basic repeating unit is the **nucleosome**.

A nucleosome is formed by wrapping 147 base pairs of DNA twice around an octamer of four core histones, **H2A** , **H2B** , **H3** and **H4** (2 copies of each one).

X-ray structure of the nucleosome core particle consisting of core histones, and DNA. Top view.



Side view shows two windings of DNA and two histone layers



www.wikipedia.org

Post-translational modifications of histone tails

The disordered histone tails comprise 25-30% of the histone mass.

They extend from the compact histone multimer to provide a platform for various **post-translational modifications (PTMs)**.

These modifications affect the histones' ability to bind DNA and to other histones.

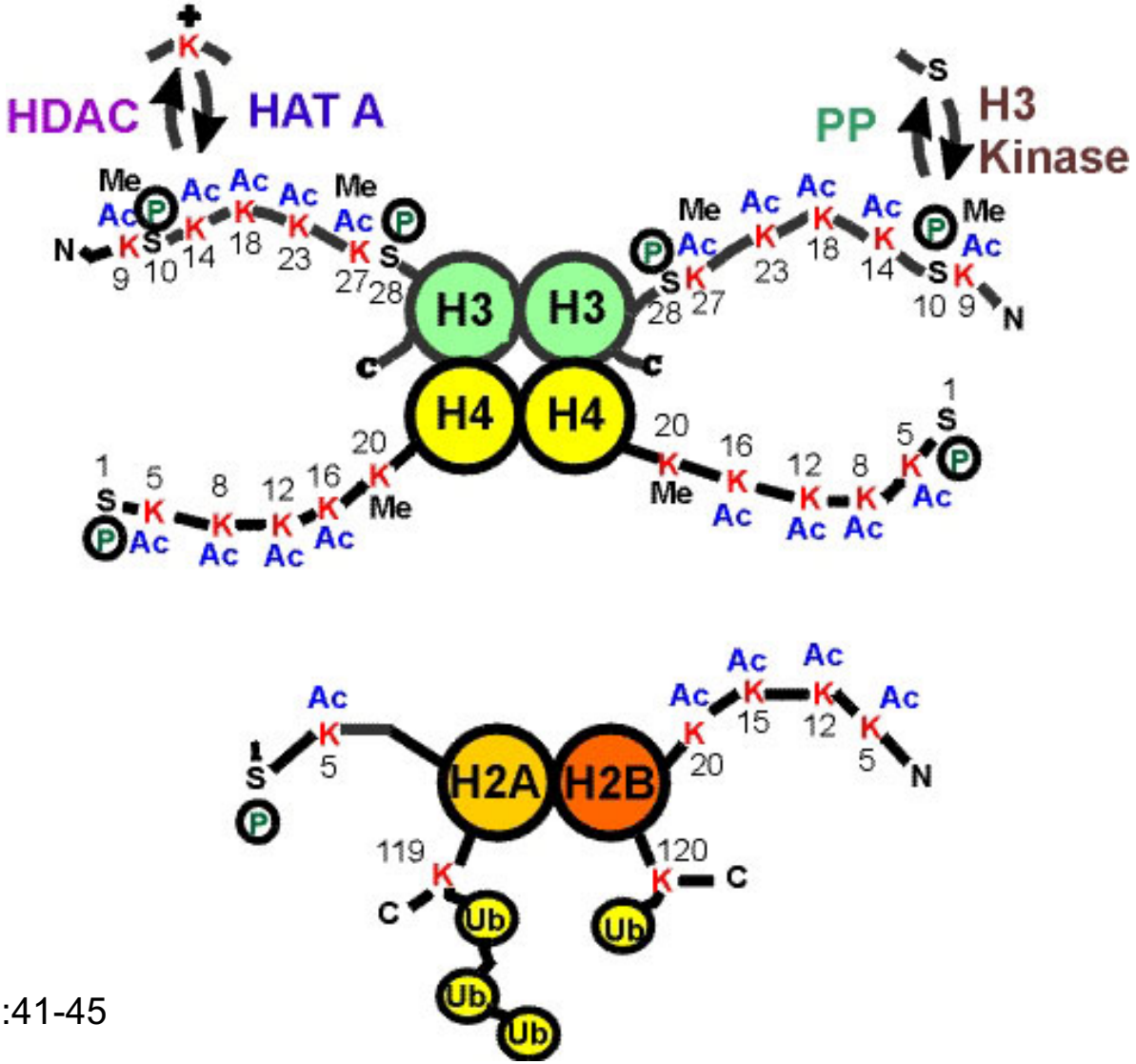
This, in turn, affects **gene expression.**

*ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS**

By V. G. Allfrey, R. Faulkner, and A. E. Mirsky

The Rockefeller Institute

PNAS 1964;51:786
First report on PTMs of histones



Strahl BD and Allis CD, 2000. Nature 403:41-45

Mode of action of histone PTMs

Histone PTMs exert their effects via two main mechanisms.

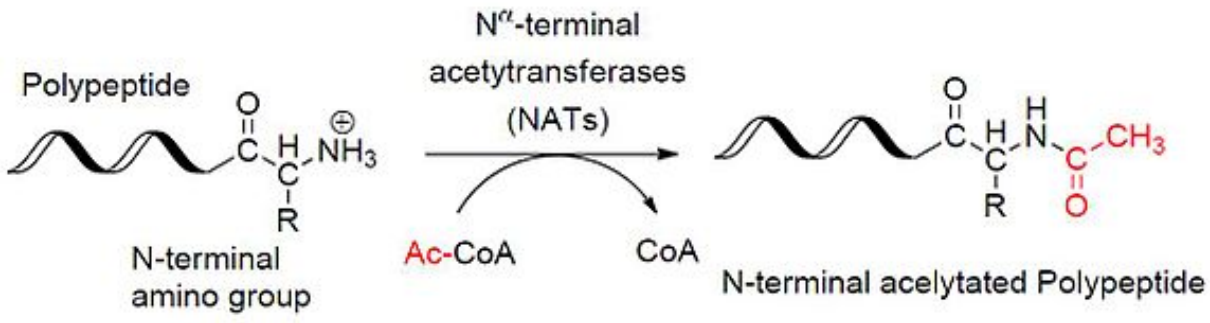
(1) PTMs directly influence the overall structure of chromatin, either over short or long distances.

(2) PTMs regulate (either positively or negatively) the binding of effector molecules.

Bannister, Kouzarides, Cell Res. (2011) 21: 381–395.

PTMs of histone tails

Histone **acetylation** and **phosphorylation** effectively reduce the positive charge of histones.



This potentially disrupts electrostatic interactions between histones and DNA.

This presumably leads to a less compact chromatin structure, thereby facilitating DNA access by protein machineries such as those involved in transcription.

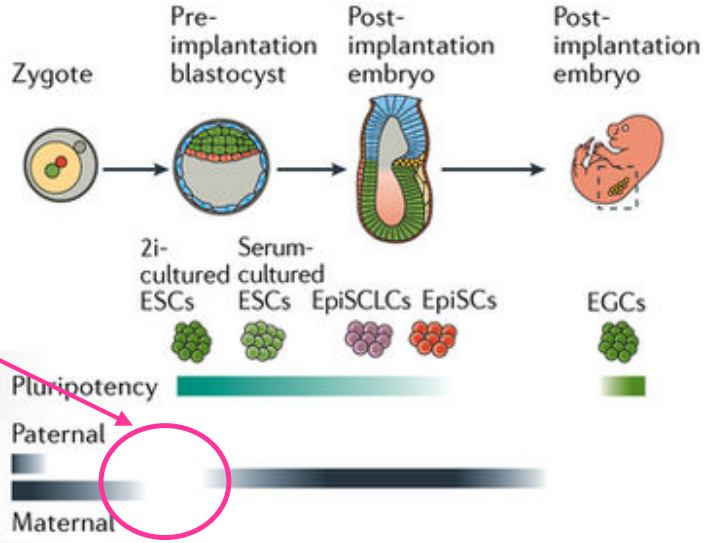
Histone **methylation** mainly occurs on the side chains of lysines and arginines.

Unlike acetylation and phosphorylation, however, histone methylation does not alter the charge of the histone protein.

Bannister, Kouzarides, Cell Res. (2011) 21: 381–395.
By Ybs.Umich - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=31240656

Dynamics of epigenetic modifications

DNA methylation is erased in the paternal and maternal genomes after fertilization and is put back on at later developmental stages.



Chromatin modification	Writers	Erasers	Location	Function
DNA methylation	DNMT1, DNMT3A and DNMT3B	TET1, TET2 and TET3	CpG dinucleotides	Silencing and others
H3K27me3	PRC2	• UTX1 • JMJD3	CpG-rich promoters and intergenic regions	Silencing
H3K9me2	G9A and GLP	• JMJD2A, JMJD2B, JMJD2C and JMJD2D • JMJD1A, JMJD1B and JMJD1C	Gene bodies, intergenic regions and enhancers	Silencing
H3K4me3	COMPASS-like proteins (SET1, MLL1–MLL2)	• JARID1A, JARID1B, JARID1C and JARID1D • KDM2B	Mainly promoters	Possibly activating
H3K27ac	HATs (including CBP/p300, GNATs and MYSTs)	HDACs and sirtuins	Promoters and enhancers	Activating
H3K4me1	COMPASS-like proteins (MLL3–MLL4)	LSD1 and LSD2	Promoters, enhancers and intergenic regions	Priming and/or activating

Atlasi & Stunnenberg, *Nature Rev Genet* **18**, 643–658 (2017)

(review V10): Epigenetics of stem cells

During development, epigenetic information is acquired in a progressive manner.

These changes regulate the transcriptional programme during **lineage commitment**.

Dynamic regulation of the epigenome underlies **cellular plasticity** and provides a **heritable** response to environmental and developmental cues.

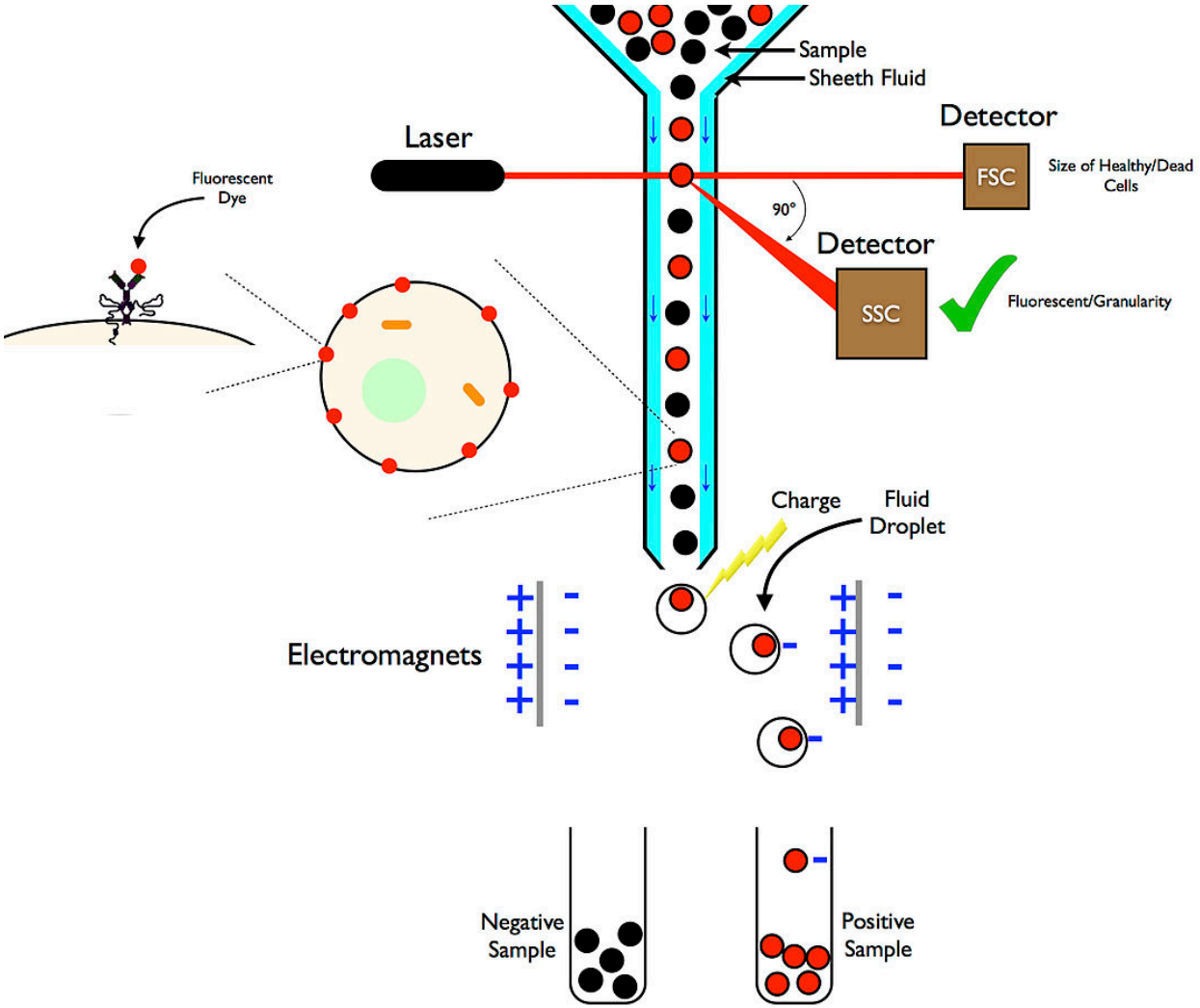
The different layers of epigenetic information are closely **interconnected**.

Epigenetic deregulation is directly linked to a wide spectrum of **diseases** ranging from developmental disorders associated with aberrant genetic imprinting to various cancers that have defects in protein complexes involved in histone or DNA modifications.

The fact that epigenetic modifications are, in principle, **reversible** renders epigenetic regulation amenable to **pharmacological intervention**.

Atlasi & Stunnenberg, *Nature Rev Genet* **18**, 643–658 (2017)

FACS

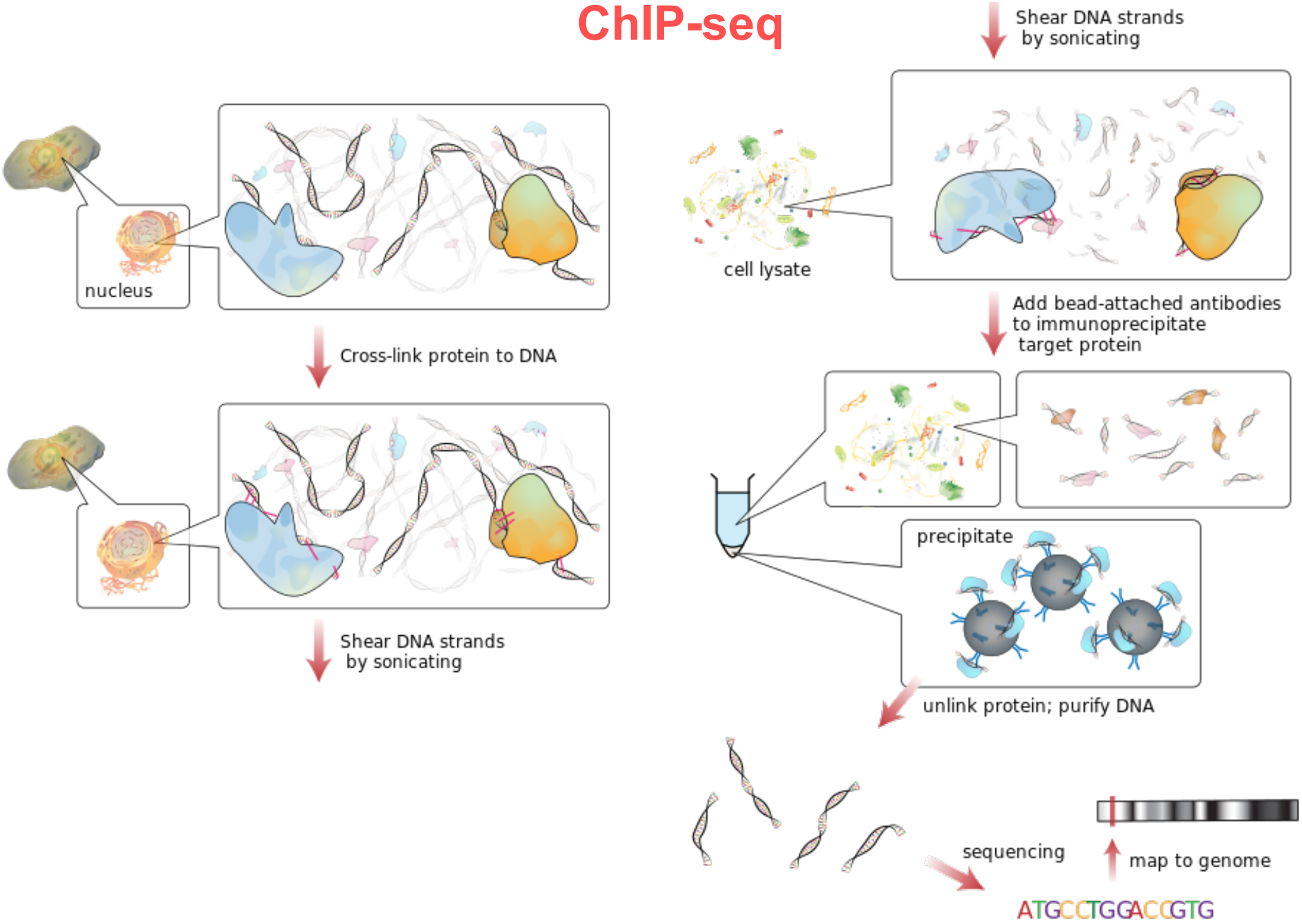


Fluorescence-activated cell sorting (FACS) is a specialized type of flow cytometry. It provides a method for sorting a heterogeneous mixture of biological cells into two or more containers, one cell at a time, based upon the specific light scattering and fluorescent characteristics of each cell.[

www.wikipedia.org

By SariSabban - Sabban, Sari (2011) https://commons.wikimedia.org/w/index.php?curid=18139883

ChIP-seq



www.wikipedia.org

TFs in Core Pluripotency Network

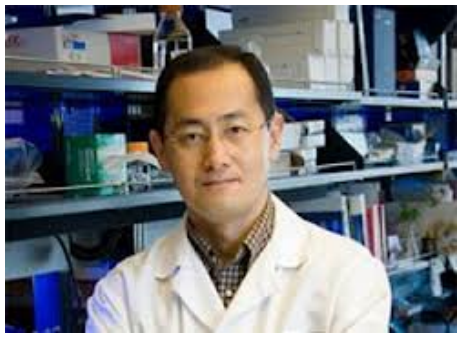
Oct4, encoded by ***Pou5f1***, is a POU domain-containing TF
that is essential to ES cells and early embryonic development.

Oct4 binds to **Sox2**, another TF.

Genome-wide mapping of OCT4 and SOX2 sites
in human ES cells shows that they **co-target** multiple genes.

Oct4 and Sox2, along with **c-Myc** and **Klf4**, appear to be sufficient for reprogramming fibroblasts to **induced pluripotent stem cells (iPS),** which are functionally similar to ES cells (→ **Yamanaka factors**).

Shinya Yamanaka
noble price for medicine 2012



Chen et al., Cell 133, 1106-1117 (2008)

Chromatin states

Analyze previously identified **informative chromatin states**

- H3K4me3+H3K27me3 (bivalent/poised promoter);
  „Poised“ genes: RNA-Polymerase II is located at their promoters in the absence of ongoing transcription, the genes are loaded to be transcribed soon
- H3K4me3+H3K27ac (active promoter); gene is actively transcribed
- H3K4me3 (initiating promoter);
- H3K27me3+H3K4me1 (poised developmental enhancer);
- H3K4me1 (poised enhancer);
- H3K27ac+H3K4me1 (active enhancer); and
- H3K27me3 (Polycomb repressed); and
- H3K9me3 (heterochromatin).

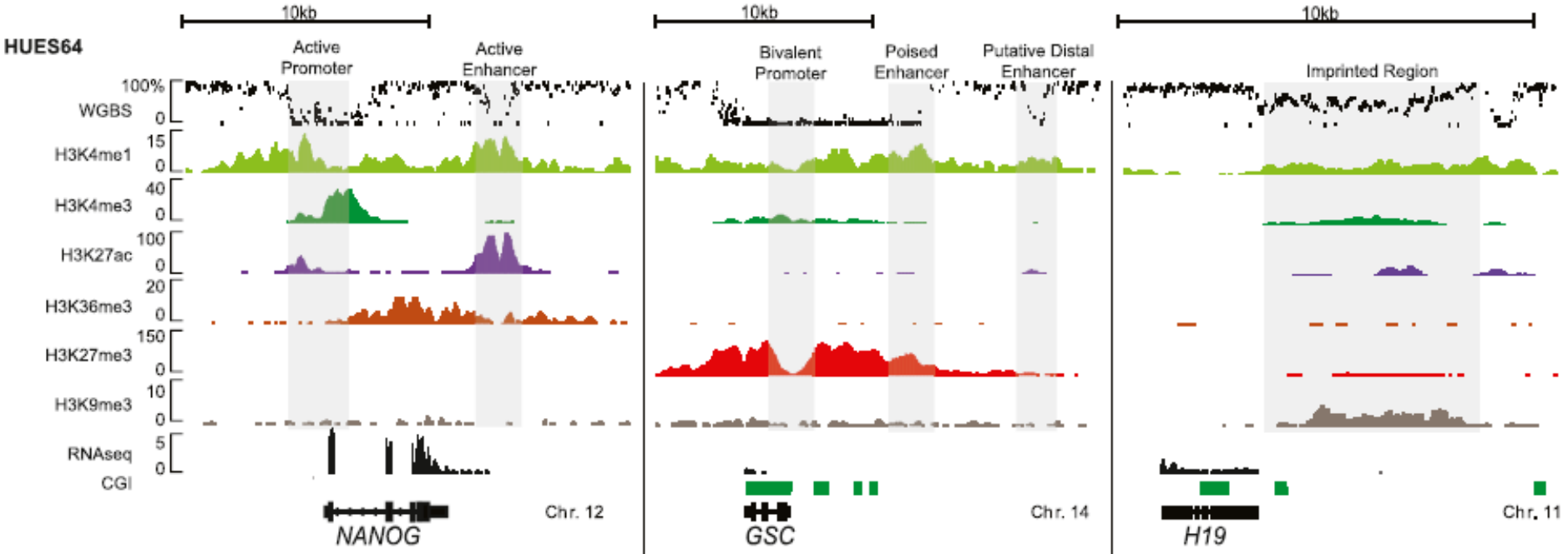
The WGBS data was segmented into three levels of DNA methylation:

- highly methylated regions (HMRs: > 60%),
- intermediately methylated regions (IMRs: 11%– 60%), and
- unmethylated regions (UMRs: 0%–10%).

Gifford et al., Cell 153, 1149-1163 (2013)

Epigenetic Data for hESC

One allele fully methylated,
other allele unmethylated
-> gene appears half methylated



Data for the undifferentiated hESC line HUES64 at 3 loci: NANOG, GSC, and H19 WholeGenomeBisulfiteSequencing (% methylation), ChIP-seq (read count normalized to 10 million reads), and RNA-seq (FPKM = fragments per kilobase of exon per million fragments mapped). **CpG islands** are indicated in **green**.

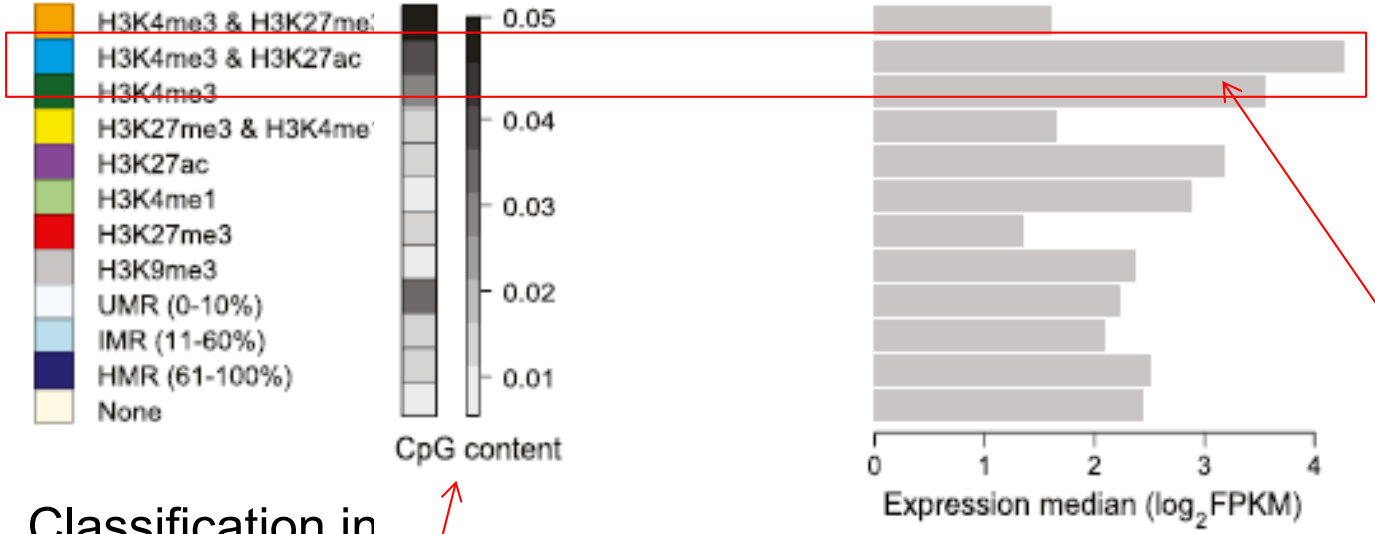
Same data was also collected for dEC, dME, and dEN cells (ca. 12 million cells each)

Bivalent promoter: carries activating (e.g. H3K4me3) and repressive (e.g. H3K27me3) histone marks

Poised enhancer: closed enhancer having H3K4me1 along with H3K27me3 and devoid of H3K27ac marks

Gifford et al., Cell 153, 1149-1163 (2013)

35% of epigenetic marks are linked to expression levels



Classification in distinct epige-netic states:

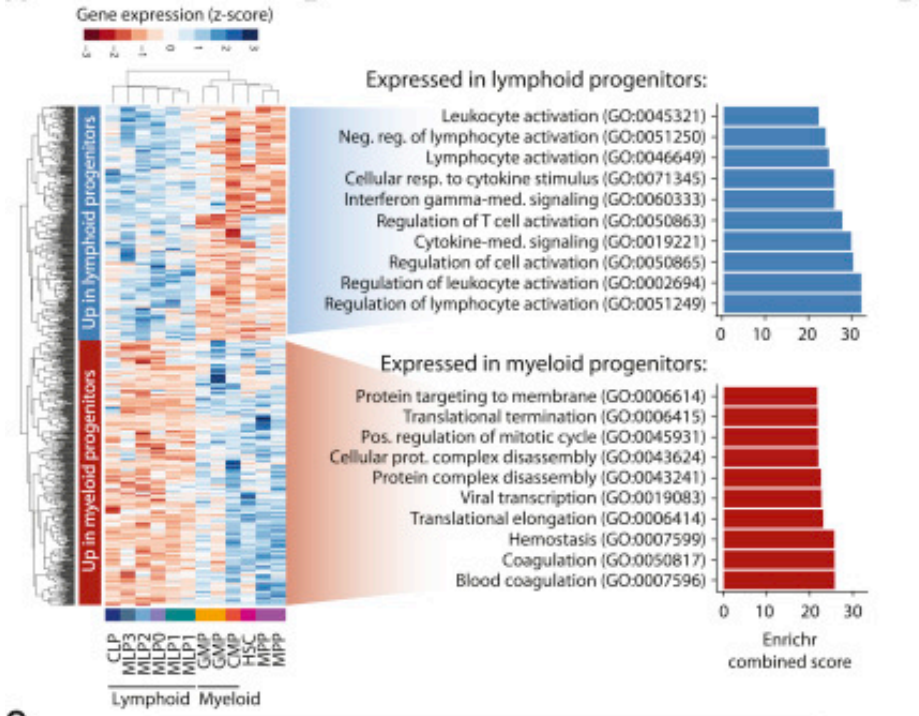
The combination of **H3K4me3** and **H3K27me3** exhibits the **highest CpG content**.

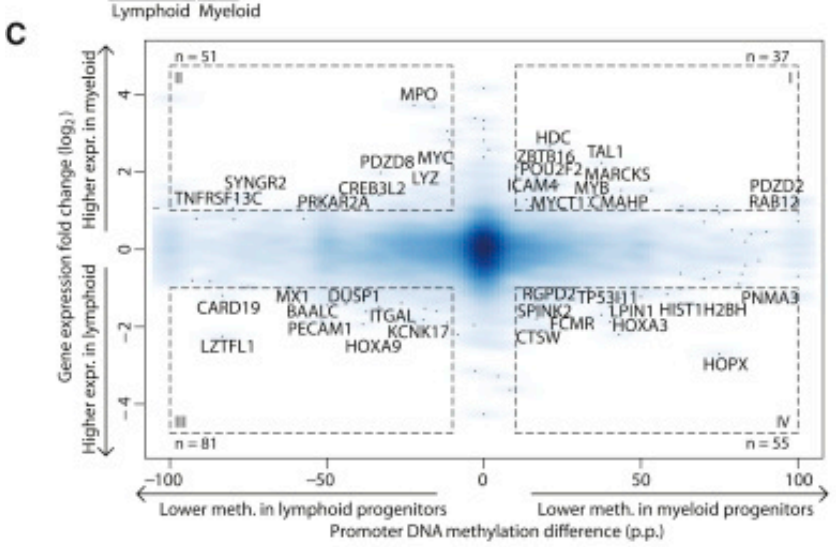
Right: Median expression level of epigenetic states based on assignment of each region to the nearest RefSeq gene.
Regions of open chromatin (active promoter; **H3K4me3 & H3K27ac**) have highest expression.

Note that many (ca. 65%) epigenetic remodeling events are not directly linked to transcriptional changes based on the expression of the nearest gene.

Gifford et al., Cell 153, 1149-1163 (2013)

Cell-type specific expression levels





656 genes were differentially expressed between myeloid and lymphoid progenitors.

Only few genes (left, bottom) showed concordant methylation and expression changes.

Farlik M et al. Cell Stem Cell (2016) 19:808-822

Tissue signature enrichment levels

DNA methyltransferase (DNMT) inhibitors and histone deacetylase (HDAC) inhibitors are in clinical trials.

A few molecules have already been approved as drugs.

Paper #8 (Fawaz, Salem, Hera): Moignard et al.

Decoding the regulatory network of early blood development from single-cell gene expression measurements

Nature Biotechnology 33, 269–276 (2015)

doi:10.1038/nbt.3154

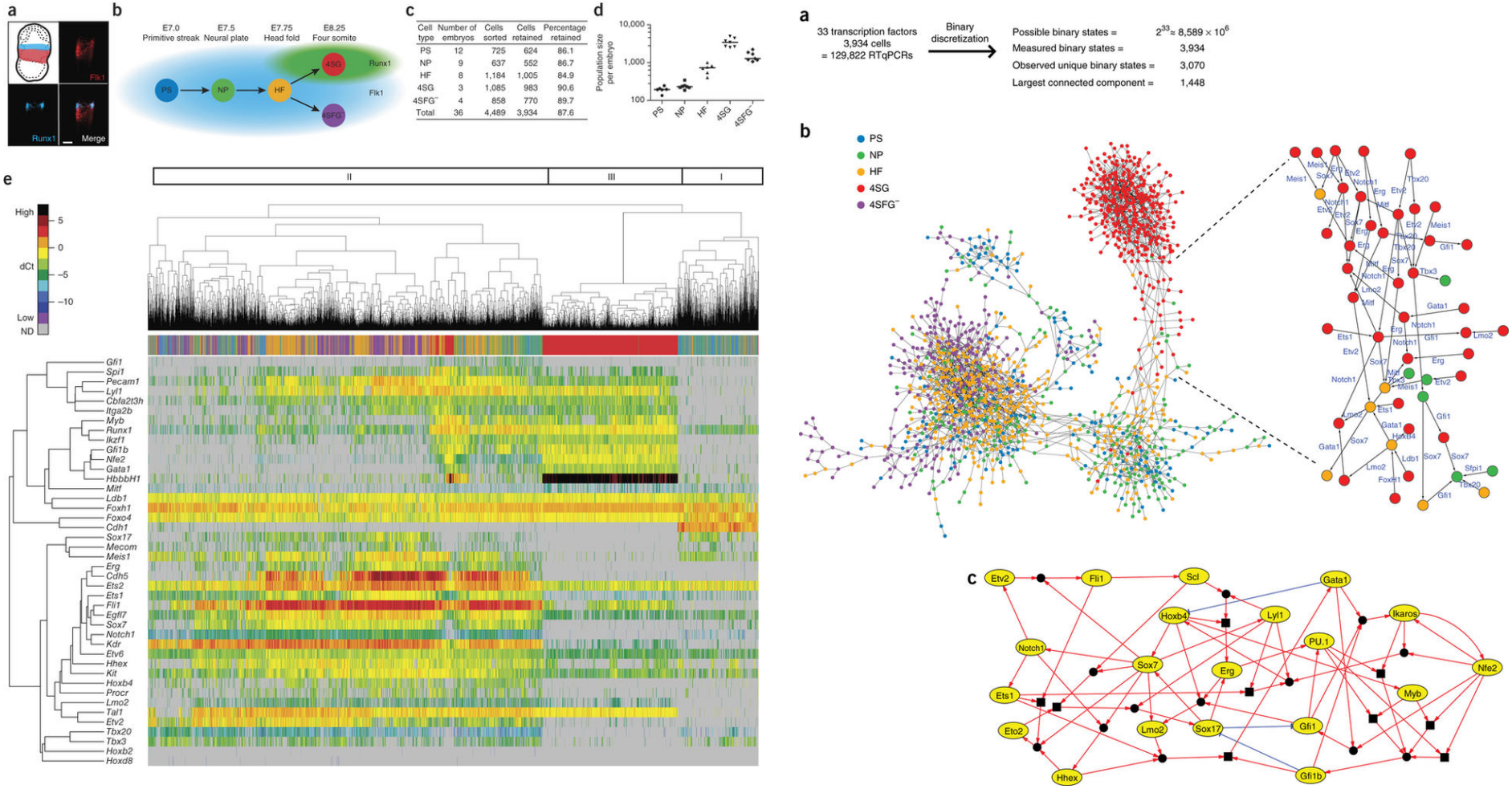
Paper #10 (Samira, Aryan, Jeenu): Göke J, et al.

Combinatorial Binding in Human and Mouse Embryonic Stem Cells Identifies Conserved Enhancers Active in Early Embryonic Development.

PLoS Comput Biol 7(12): e1002304 (2011)

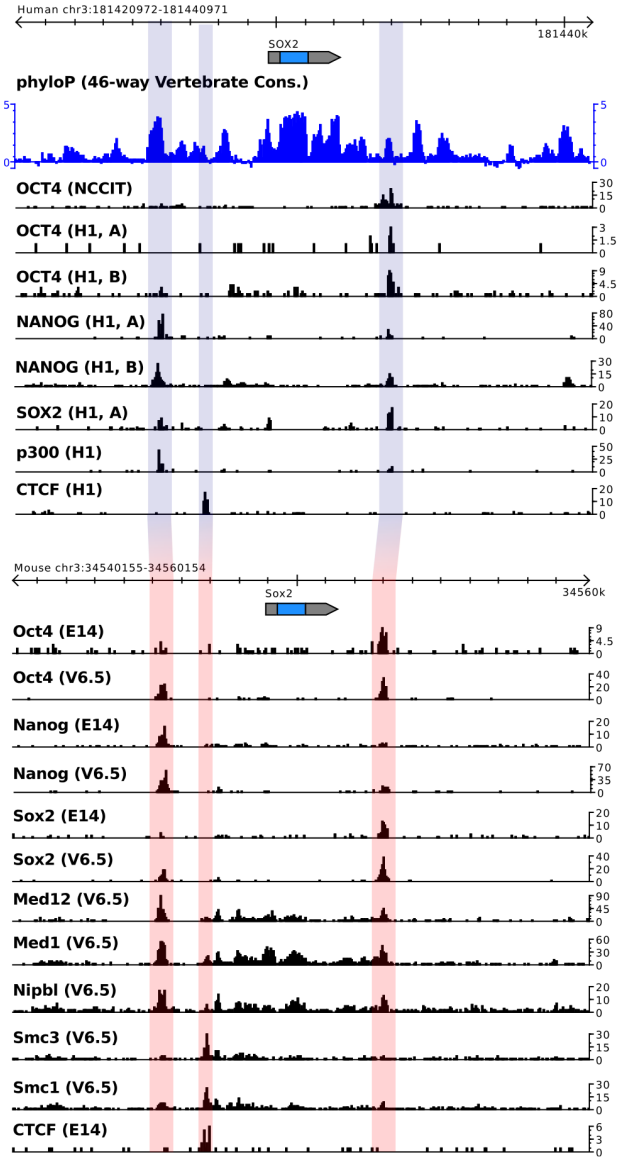
https://doi.org/10.1371/journal.pcbi.1002304

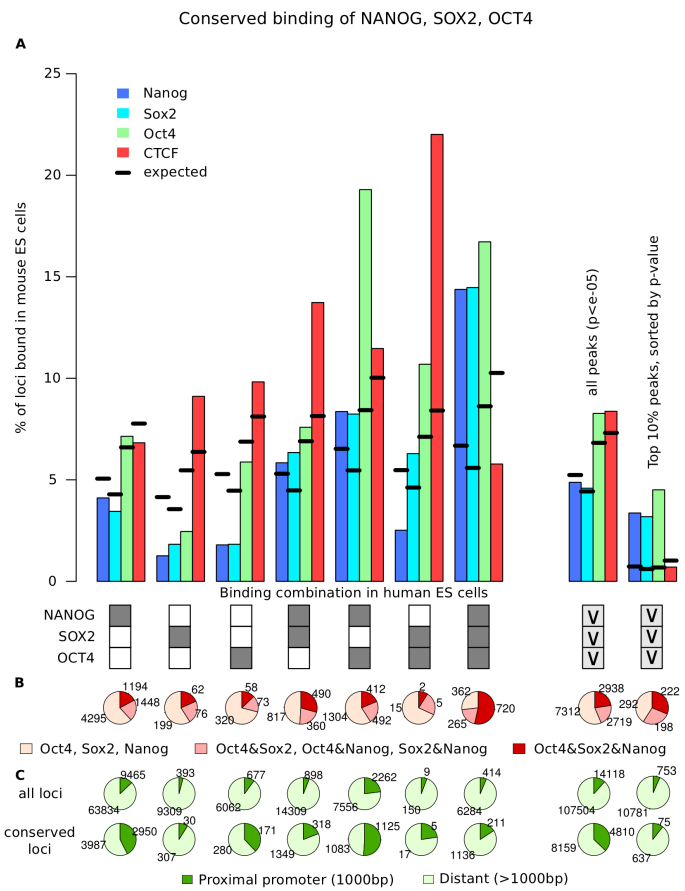
Material from Paper #8 relevant for minitest 3

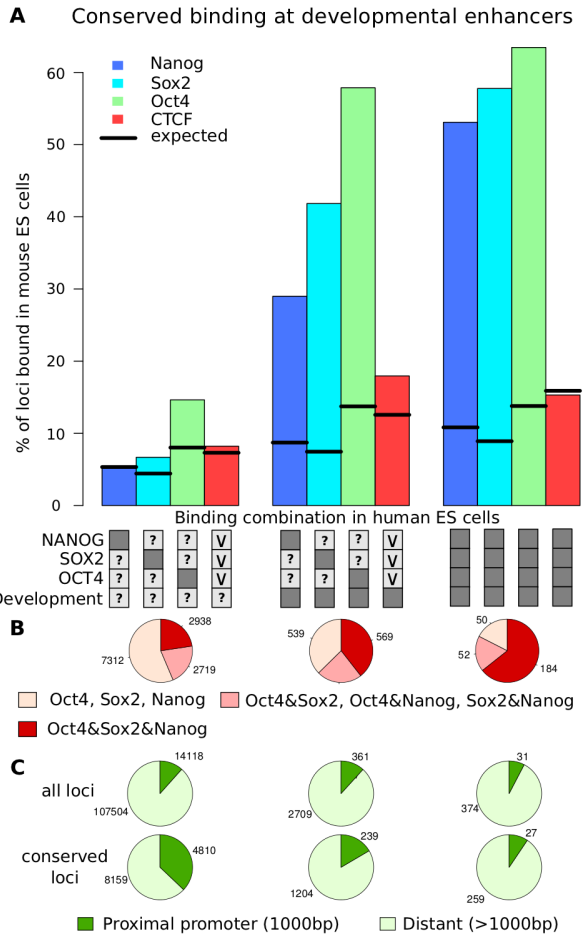


Göke et al., PLoS Comput Biol 7, e1002304 (2011)

Material from Paper #10 relevant for minitest 3







Göke et al., PLoS Comput Biol 7, e1002304 (2011)

Explanations for paper #10
Compare Pluripotency Networks of Mouse vs. Human

Given: the core of the regulatory network that maintains the pluripotent state is a set of **TFs**. Among these, **OCT4** seems to play a key role.

OCT4 co-occupies many regulatory sites together with SOX2 and NANOG.

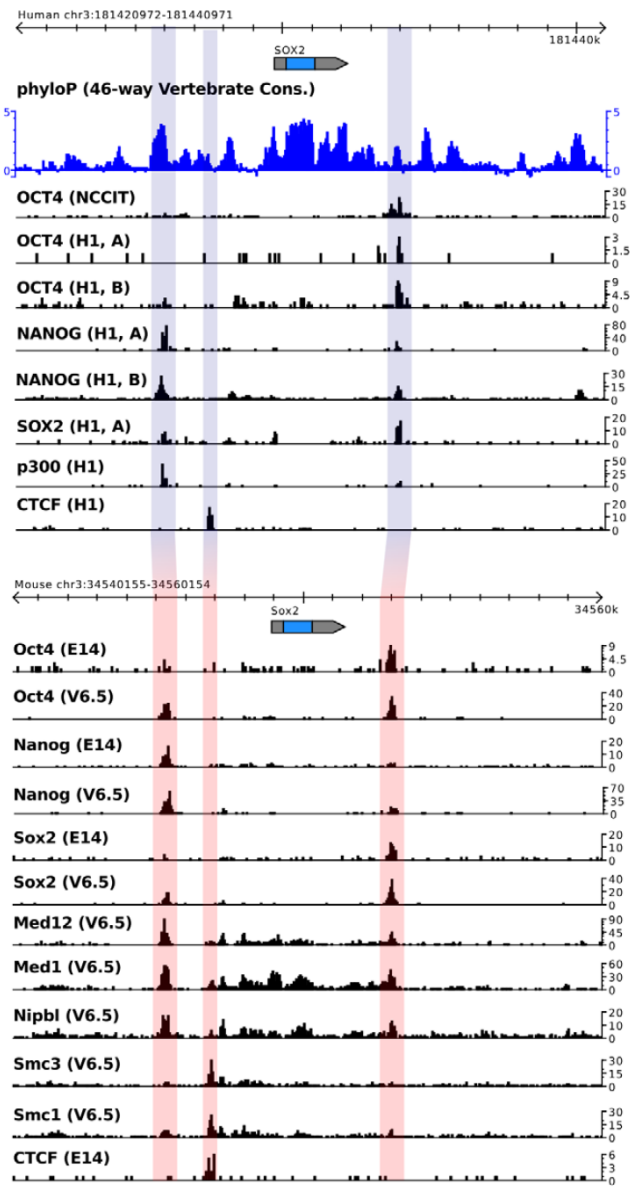
Many genes which are important for early embryogenesis have a conserved function in mouse and human.

BUT: only about 5% of binding events of the key pluripotency factors OCT4 and NANOG are conserved at orthologous genomic locations in mouse and human ES cells.

This is also true for other TFs. E.g. the liver TFs CEBP and HNF4 only showed 7% conserved binding events between human and mouse.

Göke et al., PLoS Comput Biol 7, e1002304 (2011)

Pluripotency Network in mouse and human ES cells



Overview of genome-wide CHiP-seq binding data in human H1 ES cells and various mouse embryonic stem cells and embryonal carcinoma cells.

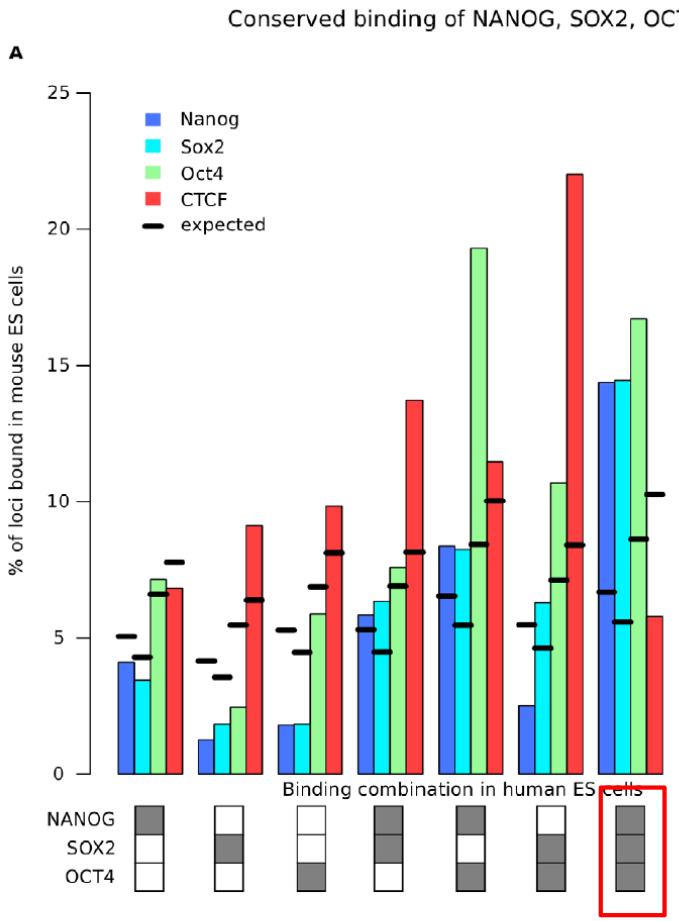
Note: human genes use CAPITAL letters (OCT4), mouse genes small letters (Oct4).

Shown is the locus of the SOX2 gene in the human genome (top), along with mapped reads for OCT4, SOX2, NANOG and p300. Individual experiments are shown separately.

The orthologous locus in the mouse genome is aligned at the bottom along with mapped reads from individual experiments.

Göke et al., PLoS Comput Biol 7, e1002304 (2011)

Combined binding of Oct4, Sox2 and Nanog



The combination of OCT4, SOX2 and NANOG influences conservation of binding events.

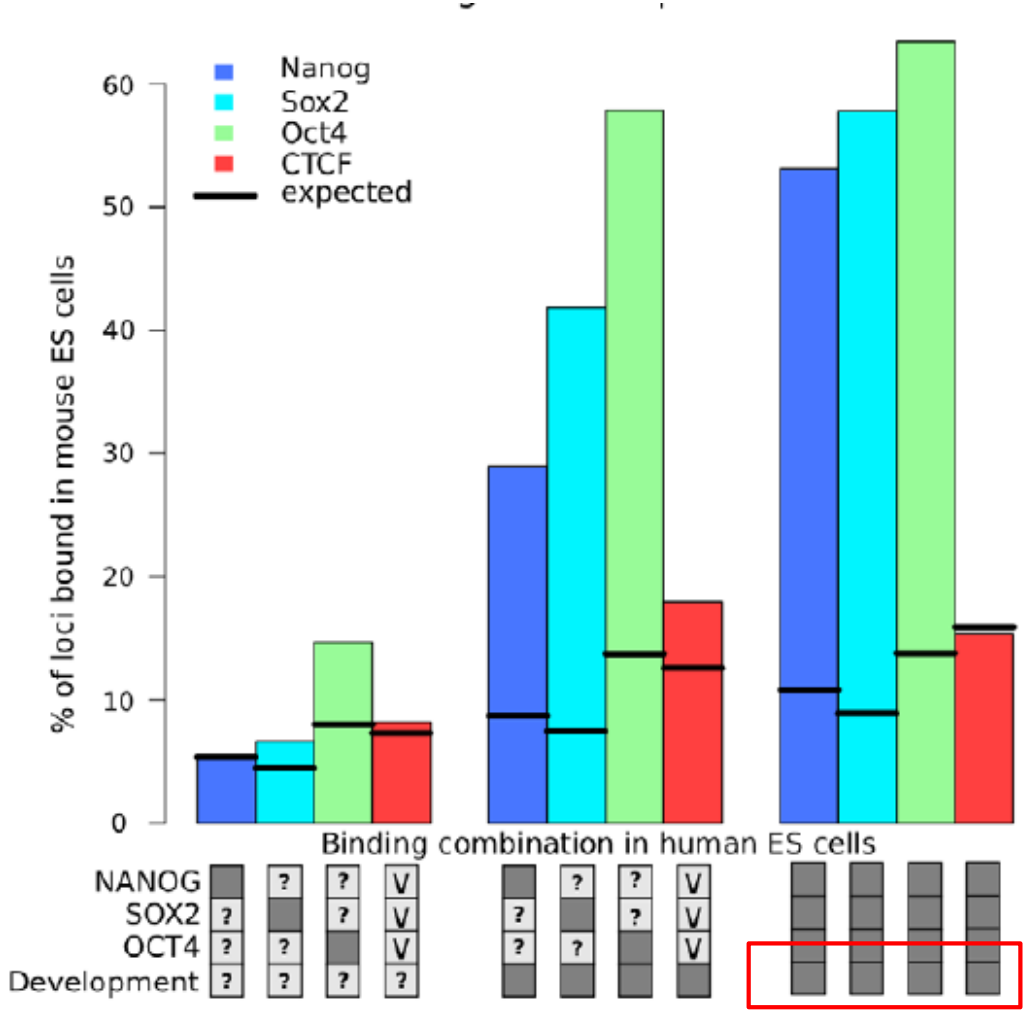
(A) Bars indicate the fraction of loci where binding of Nanog, Sox2, Oct4 or CTCF can be observed at the orthologous locus in mouse ES cells for all combinations of OCT4, SOX2 and NANOG in human ES cells as indicated by the boxes below.

Dark boxes indicate binding, white boxes indicate no binding (“AND” relation).

Combinatorial binding of OCT4, SOX2 and NANOG shows the largest fraction of conserved binding for Oct4, Sox2 and Nanog in mouse.

Göke et al., PLoS Comput Biol 7, e1002304 (2011)

Increased Binding conservation in ES cells at developmental enhancers



Fraction of loci where binding of Nanog, Sox2, Oct4 and CTCF can be observed at the orthologous locus in mouse ESC.

Combinations of OCT4, SOX2 and NANOG in human ES cells are discriminated by developmental activity as indicated by the boxes below.
Dark boxes : "AND" relation,
light grey boxes with "v" : "OR" relation, '
"?" : no restriction.

Combinatorial binding events at developmentally active enhancers show the highest levels of binding conservation between mouse and human ES cells.

Göke et al., PLoS Comput Biol 7, e1002304 (2011)

Summary

26% of combinatorially bound loci which are conserved between mouse and human ES cells are **developmental enhancers** in the mouse.
This suggests that many enhancers bound by OCT4, SOX2 and NANOG are also developmental enhancers in human.

The very same regulatory elements bound by key pluripotency factors in ES cells frequently act as enhancers during early development.

$\rightarrow$ the gene regulatory networks of ES cells and early development are linked at the level of transcriptional regulation.

The finding that binding at these developmental enhancers are highly conserved in mouse and human ES cells suggests that these elements are crucial for the maintenance of the pluripotent state.

Göke et al., PLoS Comput Biol 7, e1002304 (2011)

Paper presentations in V12 and V13

V12 Paper #11 (Girmay, Schowing, Reis): Chao Lu et al.
IDH mutation impairs histone demethylation and results in a block to cell differentiation
Nature 483, 474–478 (2012)
doi:10.1038/nature10860

Paper #13 (Grandke, Thedinga, Rubab): Athanasiadis EI, et al.
Single-cell RNA-sequencing uncovers transcriptional states and fate decisions in haematopoiesis
Nature Communications 8, 2045 (2017)
doi: 10.1038/s41467-017-02305-6.

V13 (second half) Paper #12 (Dash, Prudhan, Shreshta): Monika E. Hegi, et al.
David Brocks et al.
Intratumor DNA Methylation Heterogeneity Reflects Clonal Evolution in Aggressive Prostate Cancer
Cell Reports 8, 798-806 (2014)
https://doi.org/10.1016/j.celrep.2014.06.053

Paper #14 (Flohr, Bauer, Lüssem): Shaffer SM, et al.
Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance
Nature 546, 431–435 (2017)
doi:10.1038/nature22794

Material from paper #13 relevant for minitest 3

Paper #13 (Grandke, Thedinga, Rubab):

Athanasiadis EI, et al.

Single-cell RNA-sequencing uncovers transcriptional states and fate decisions in haematopoiesis

Nature Communications 8, 2045 (2017)

doi: 10.1038/s41467-017-02305-6.