- Nucleosome structure illustrates the logic of histone modifications
- Nucleosomes are highly dynamic
- Covalent modification of histones define a functional "histone code"
- A battery of histone-modifying enzyme exist
- Several classes of proteins are implicated in the recognition of histone code, to connect "local" modifications (e.g. at promoters or enhancers) with the locus-level chromatin status

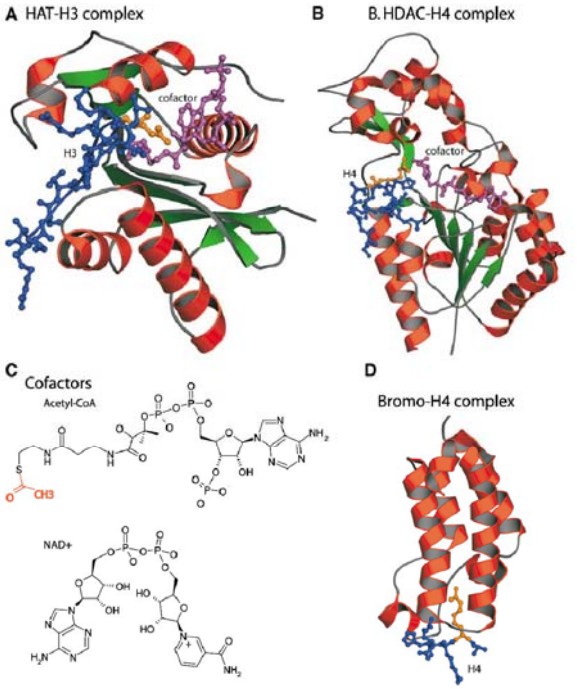


Figure 5. Lysine Acetylation and Deacetylation

(A) A HAT domain is shown in complex with a substrate H4 tail (blue) and its cofactor (purple) (Clements et al., 2003). The substrate lysine is shown in orange.

(B) The structure of yeast Hst2 deacetylase (Sir2 homolog) in complex with H4 substrate (blue) and cofactor (purple) (Zhao et al., 2003). The substrate acetyllysine is shown in orange.

(C) The cofactors of HAT (acetyl-CoA) and HDAC (NAD^+).

(D) The bromodomain of Gcn5 in complex with the H4 tail (blue) with acetyllysine 16 (orange) (Owen et al., 2000).

acetyl-Lys

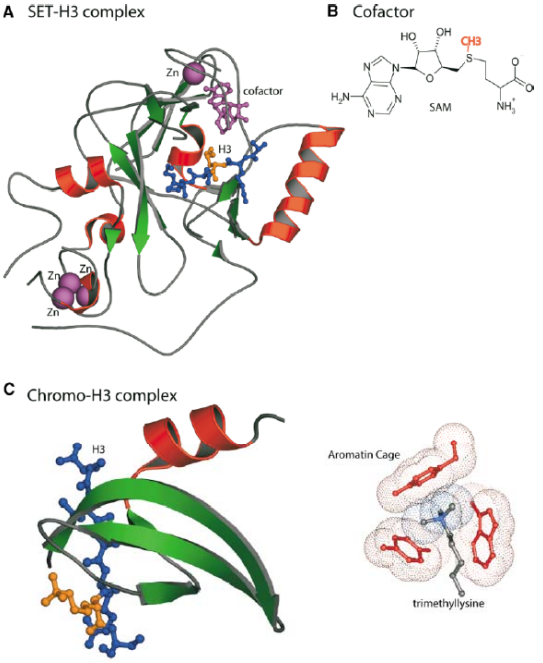


Figure 6. Lysine Methylation

(A) The structure of DIM5, a homolog of SUV39H1, which methylates lysine 9 (orange) in H3, is shown in complex with H3 tail (blue), cofactor (purple), and four zinc ions (pink) (Zhang et al., 2003). The zinc ions stabilize the pre-SET and post-SET regions, both of which are important for catalytic function.
(B) The cofactor S-adenosyl-L-methionine.
(C) HP1 chromodomain in complex with H3 tail (blue) with methyllysine 9 (orange) (Jacobs and Khorasanizadeh, 2002). An aromatic cage recognizes the methyllysine moiety.

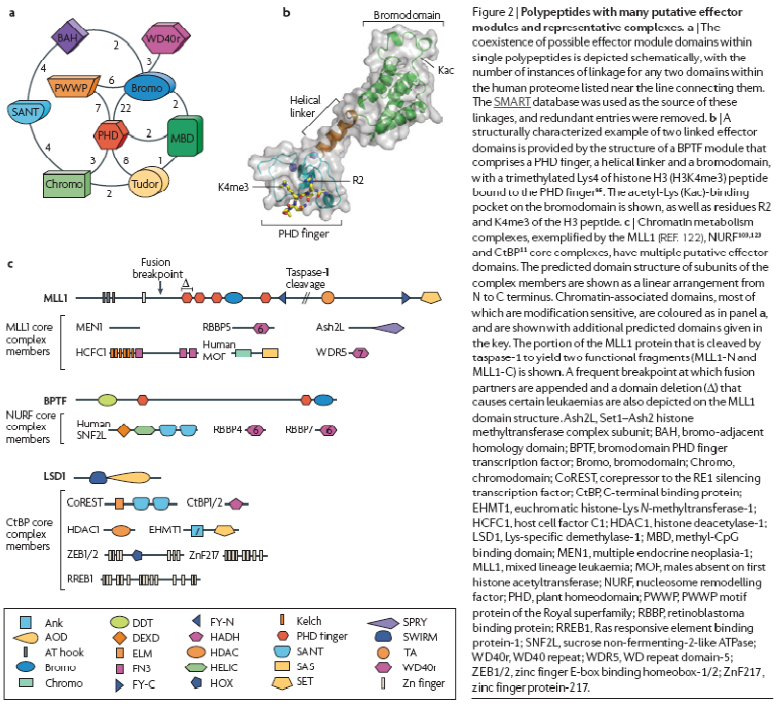
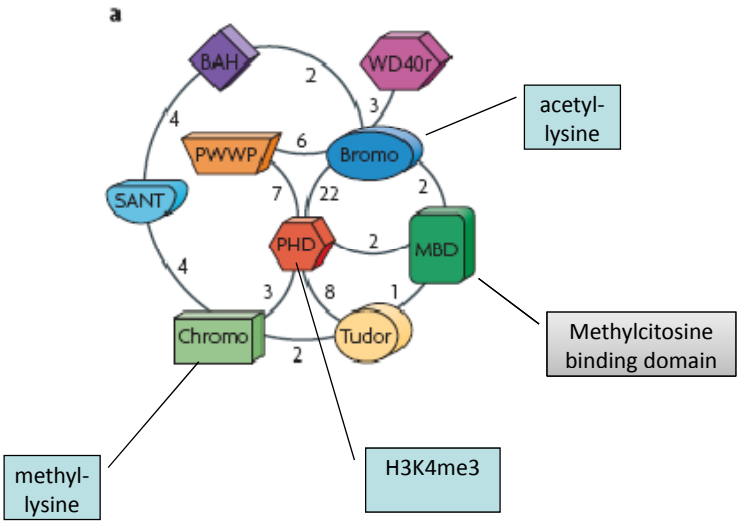


Figure 2 | **Polypeptides with many putative effector modules and representative complexes. a** | The coexistence of possible effector module domains within single polypeptides is depicted schematically, with the number of instances of linkage for any two domains within the human proteome listed near the line connecting them. The SMART database was used as the source of these linkages, and redundant entries were removed. **b** | A structurally characterized example of two linked effector domains is provided by the structure of a BPTF module that comprises a PHD finger, a helical linker and a bromodomain, with a trimethylated Lys4 of histone H3 (H3K4me3) peptide bound to the PHD finger[66]. The acetyl-Lys (Kac)-binding pocket on the bromodomain is shown, as well as residues R2 and K4me3 of the H3 peptide. **c** | Chromatin metabolism complexes, exemplified by the MLL1 (REF. 122), NURF[109,123] and CtBP[111] core complexes, have multiple putative effector domains. The predicted domain structure of subunits of the complex members are shown as a linear arrangement from N to C terminus. Chromatin-associated domains, most of which are modification sensitive, are coloured as in panel a, and are shown with additional predicted domains given in the key. The portion of the MLL1 protein that is cleaved by taspase-1 to yield two functional fragments (MLL1-N and MLL1-C) is shown. A frequent breakpoint at which fusion partners are appended and a domain deletion (Δ) that causes certain leukaemias are also depicted on the MLL1 domain structure. Ash2L, Set1–Ash2 histone methyltransferase complex subunit; BAH, bromo-adjacent homology domain; BPTF, bromodomain PHD finger transcription factor; Bromo, bromodomain; Chromo, chromodomain; CoREST, corepressor to the RE1 silencing transcription factor; CtBP, C-terminal binding protein; EHMT1, euchromatic histone-Lys *N*-methyltransferase-1; HCFC1, host cell factor C1; HDAC1, histone deacetylase-1; LSD1, Lys-specific demethylase-1; MBD, methyl-CpG binding domain; MEN1, multiple endocrine neoplasia-1; MLL1, mixed lineage leukaemia; MOF, males absent on first histone acetyltransferase; NURF, nucleosome remodelling factor; PHD, plant homeodomain; PWWP, PWWP motif protein of the Royal superfamily; RBBP, retinoblastoma binding protein; RREB1, Ras responsive element binding protein-1; SNF2L, sucrose non-fermenting-2-like ATPase; WD40r, WD40 repeat; WDR5, WD repeat domain-5; ZEB1/2, zinc finger E-box binding homeobox-1/2; ZnF217, zinc finger protein-217.



Principal characteristics of chromatin-interacting proteins

1. "Modules" or "domains" that recognize specific histone position / PMT are present in a plethora of chromatin metabolizing enzymes and proteins

2. These modular proteins are always present in megadalton, heteromeric and multifunctional complexes

The figures shown in preceding slides are from:

Ruthenburg et al., 2007, Nat. Rev. Mol. Cell Biol. 8: 983.

REVIEW 1

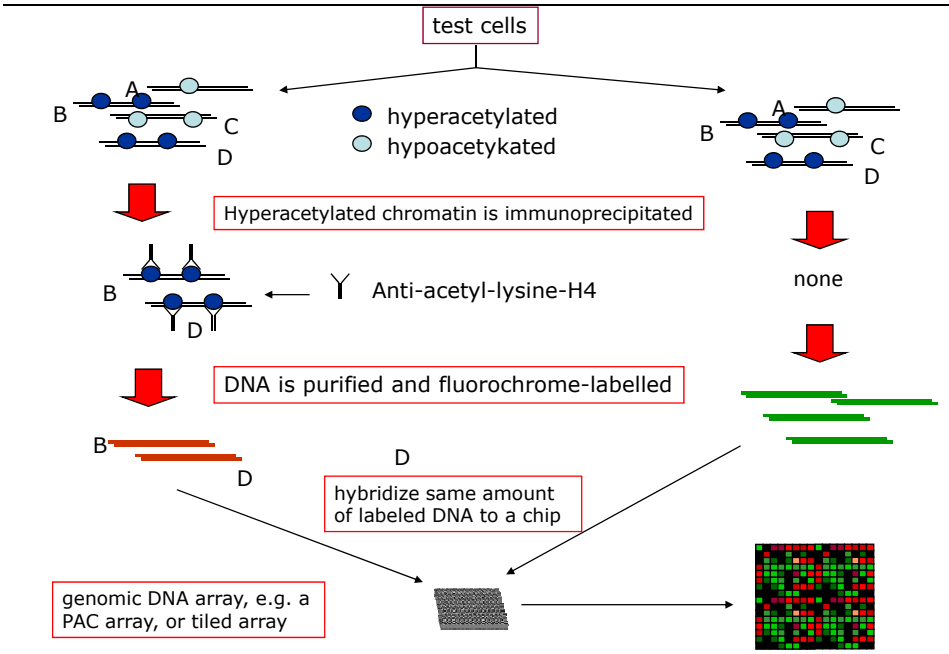
Methods and technologies used to study chromatin status at the genome-wide level are reviewed here

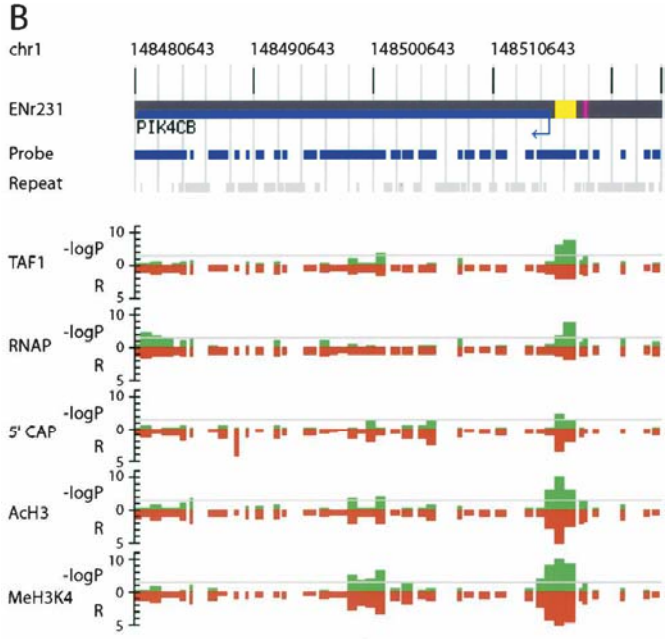
Microarray technology: beyond transcript profiling and genotype analysis

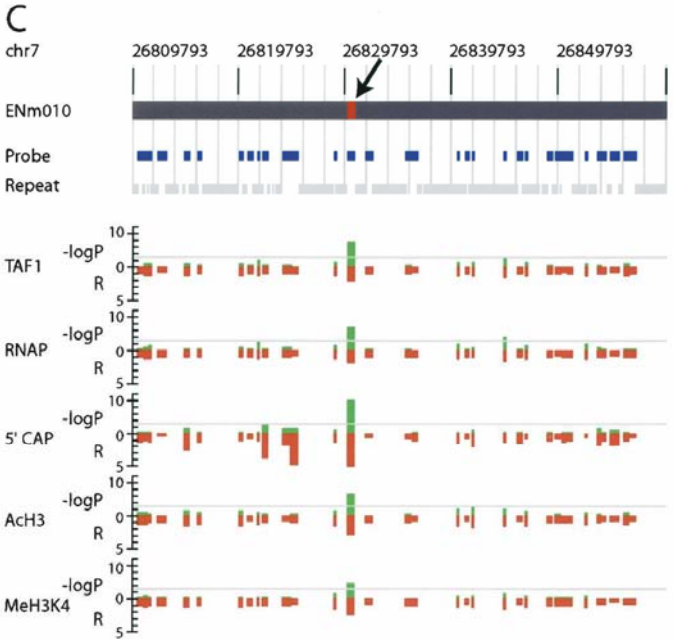
Jörg D. Hoheisel

Abstract | Understanding complex functional mechanisms requires the global and parallel analysis of different cellular processes. DNA microarrays have become synonymous with this kind of study and, in many cases, are the obvious platform to achieve this aim. They have already made important contributions, most notably to gene-expression studies, although the true potential of this technology is far greater. Whereas some assays, such as transcript profiling and genotyping, are becoming routine, others are still in the early phases of development, and new areas of application, such as genome-wide epigenetic analysis and on-chip synthesis, continue to emerge.

Identification of hyperacetylated nucleosomes by genome-wide chromatin immunoprecipitation



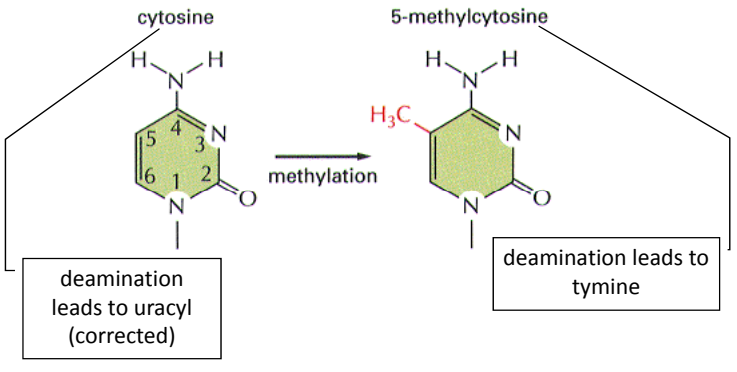

B
chr1
148480643
148490643
148500643
148510643
ENr231
PIK4CB
Probe
Repeat
TAF1
RNAP
5' CAP
AcH3
MeH3K4
-logP
R
10
0
5


C
chr7
26809793
26819793
26829793
26839793
26849793
ENm010
Probe
Repeat
TAF1
RNAP
5' CAP
AcH3
MeH3K4
-logP
R
10
0
5

Heterochromatic	Euchromatic
more packed	less packed
silenced	active or poised
rich in HP1α	poor in HP1α
Histones hypoacetylated	Histones hyperacetylated
H3K9me H3K27me other H metylations	H3K4me
CpG islands methylated	CpG islands un-methylated

CpG methylation

Analysis of heterochromatic regions and of silenced gene promoters has shown that a modification of 5'-CpG-3' sequences with methylation of the C-5 of cytosine is very frequent



CpG methylation is the only epigenetic modification concerning the DNA

Among experimental evidences obtained:

a. Methylation of CpG is observed at regulatory regions of silenced genes

b. Hypermethylation is observed through the inactive X chromosome

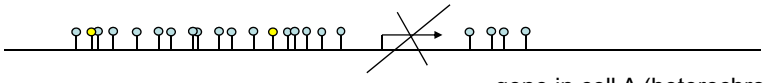
c. Methylated DNA can not usually be expressed as a transgene

d. Housekeeping genes (constitutively expressed) show unmethylated CpG islands

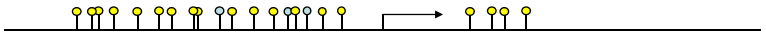
e. CpG methylation is observed at imprinted genes

f. CpG methylation profiles can be reproduced with fidelity during DNA replication and cell division

Important: CpG methylation is common in Mammals and Plants, but is not used in the same way in *S. cerevisiae* and *C. elegans*



gene in cell A (heterochromatic)



gene in cell B (euchromatic)

unmethylated CpG

methylated CpG

Why are methylated genes silenced ?

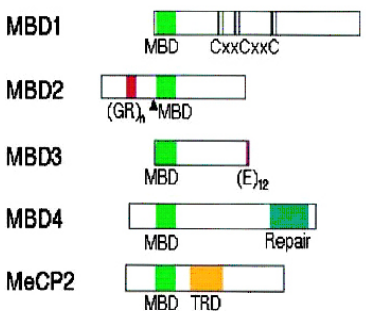
Usually CpG methylation does not limit access of transcription factors to DNA (there are few exceptions).

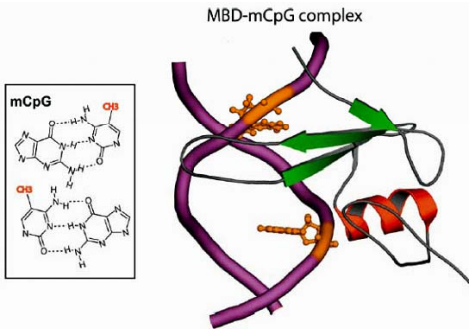
Is there any link between DNA CpG methylation and histone modifications typical of heterochromatin ?

The answer is: YES !

There are several links between DNA CpG methylation and histone de-acetylation and H3-K9 methylation

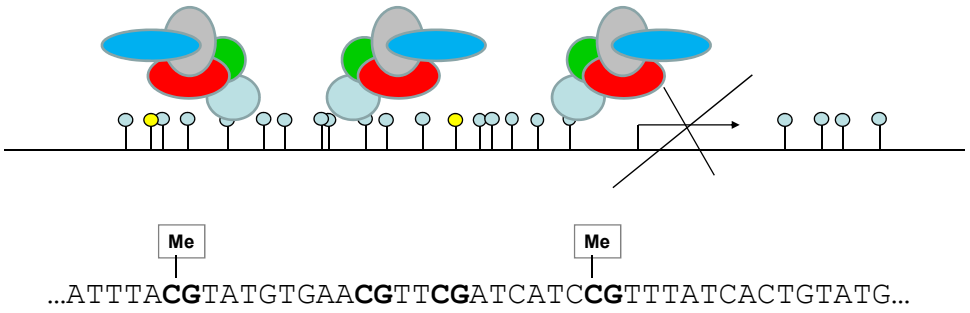
There are several Methyl CpG binding proteins in eukaryotes, which deserve different functions. The protein involved in heterochromatinization is most likely MeCP2 that possesses both a Methyl CpG Binding Domain (MBD) and a transcriptional repression domain (TRD).



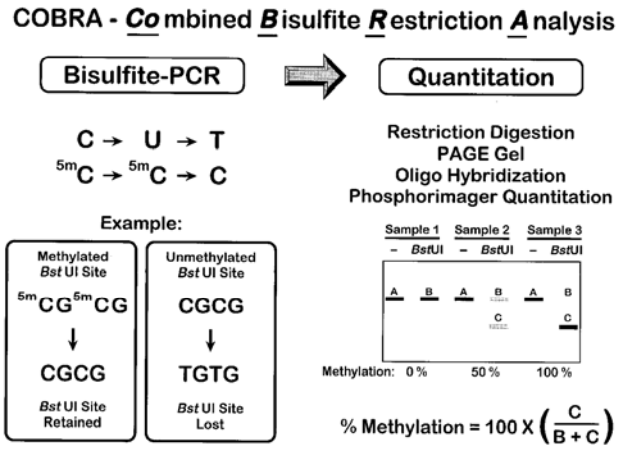


The structure of Methyl Binding Domain (MBD) in complex with methyl CpG (orange)

CpG methylation is physically linked to campaction of chromatin through the MeCP2 protein and other MBD proteins that bind meCpG and make part of large protein complexes containing Histone Deacetylases (HDAC), Histone methyltransferases (HMT) etc... to the site.



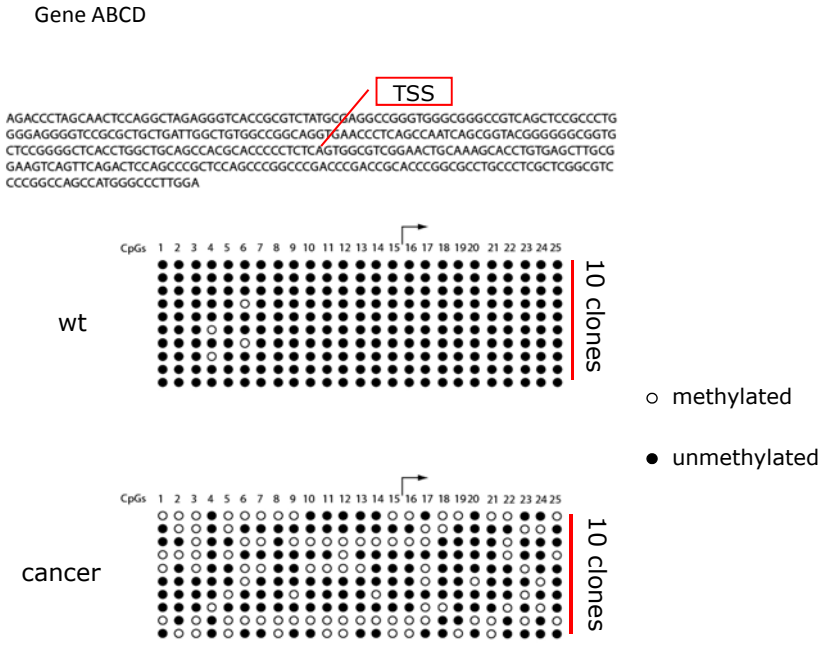
Most common methods to detect CpC methylation are based on bisulfite, that converts C (but not methyl-C) to Uridine. Uridine in DNA is then replaced by T in the following PCR.
Conservation of C (in the case of methyl-C) or change in T are subsequently detected (simplest) by restriction site analysis, as in the **example** following:



In a promoter sequence, CpG “profile” varies from cell to cell

Therefore, to determine the methylation profile:

1) extract DNA
2) bisulfite treatment
3) PCR the fragment
4) clone individual fragment in a vector
5) sequence a representative number of clones



During development, a progressive selection of genes to be expressed and genes to be silenced occurs.

Each cell type establishes its own profile of heterochromatic and euchromatic partition of the genome.

We see this quite clearly by following development using DNA microarray analysis of gene expression

The question now is:

How is the heterochromatic or euchromatic status transmitted to daughter cells during DNA replication and cell division ?

Epigenetic determinants

The chromatic status is **metastable**

How is silencing / expression (i.e. heterochromatic / euchromatic status) inherited by daughter cells after cell division ?

How is heterochromatin established and maintained ?

Is there any kind of epigenetic inheritance in germinal cells ?

Chromatin in pluripotent embryonic stem cells and differentiation

Eran Meshorer and Tom Misteli

Abstract | Embryonic stem (ES) cells are unique in that they are pluripotent and have the ability to self-renew. The molecular mechanisms that underlie these two fundamental properties are largely unknown. We discuss how unique properties of chromatin in ES cells contribute to the maintenance of pluripotency and the determination of differentiation properties.

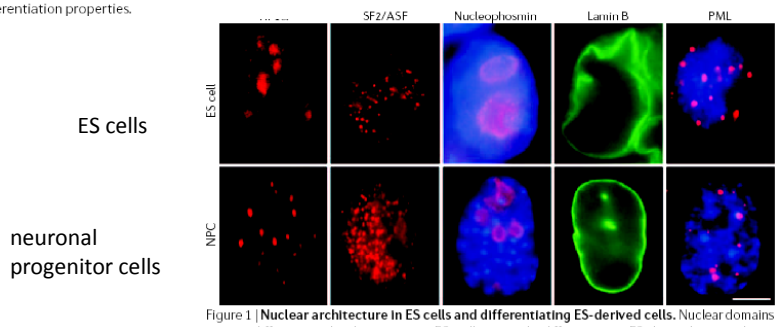


Figure 1 | **Nuclear architecture in ES cells and differentiating ES-derived cells.** Nuclear domains in an undifferentiated embryonic stem (ES) cell (top) and a differentiating ES-derived neuronal progenitor cell (NPC, bottom). From left to right: heterochromatin, as detected with an anti-HP1α antibody, is confined to fewer and larger foci in ES cells compared with NPCs; nuclear speckles, as detected with an anti-SF2/ASF antibody, appear as small, dispersed foci in ES cells and become more conspicuous in NPCs; nucleoli, as identified with an anti-nucleophosmin antibody, appear larger in ES cells compared with NPCs; the ill-defined nuclear lamina in ES cells, stained with an anti-lamin B antibody, becomes round and distinct in NPCs; promyelocytic leukaemia (PML) bodies labelled with an anti-PML antibody show similar patterns in ES cells and NPCs. DAPI, blue. Scale bar, 5 μm.

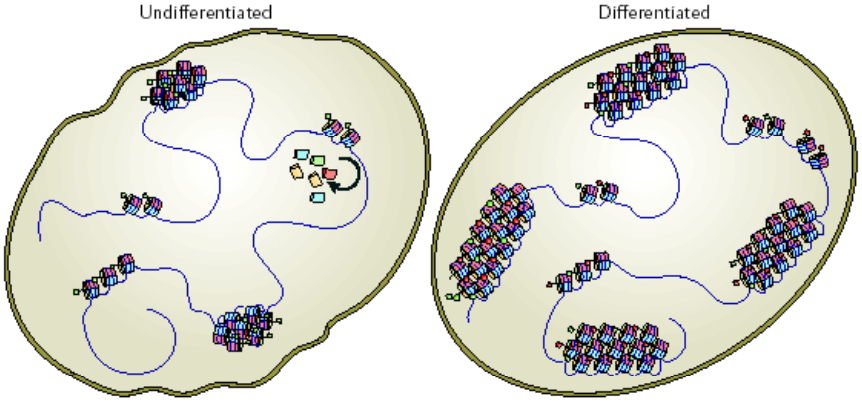


Figure 2 | **Chromatin during ES-cell differentiation.** In pluripotent embryonic stem (ES) cells (left), chromatin is globally decondensed, enriched in active histone marks (green circular tags), and contains a fraction of loosely bound architectural chromatin proteins. As cells differentiate (right), regions of condensed heterochromatin form, silencing histone marks (red circular tags) accumulate, and structural chromatin proteins become more stably associated with chromatin.

NATURE|Vol 447|24 May 2007|doi:10.1038/nature05918 INSIGHT REVIEW

Review 1

Stability and flexibility of epigenetic gene regulation in mammalian development

Wolf Reik[1]

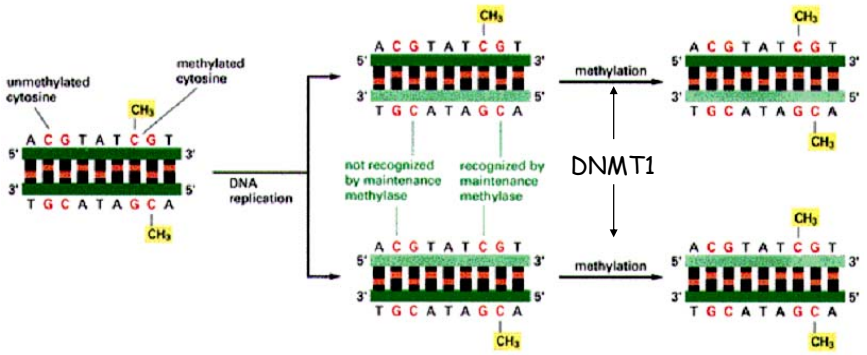
During development, cells start in a pluripotent state, from which they can differentiate into many cell types, and progressively develop a narrower potential. Their gene-expression programmes become more defined, restricted and, potentially, 'locked in'. Pluripotent stem cells express genes that encode a set of core transcription factors, while genes that are required later in development are repressed by histone marks, which confer short-term, and therefore flexible, epigenetic silencing. By contrast, the methylation of DNA confers long-term epigenetic silencing of particular sequences — transposons, imprinted genes and pluripotency-associated genes — in somatic cells. Long-term silencing can be reprogrammed by demethylation of DNA, and this process might involve DNA repair. It is not known whether any of the epigenetic marks has a primary role in determining cell and lineage commitment during development.

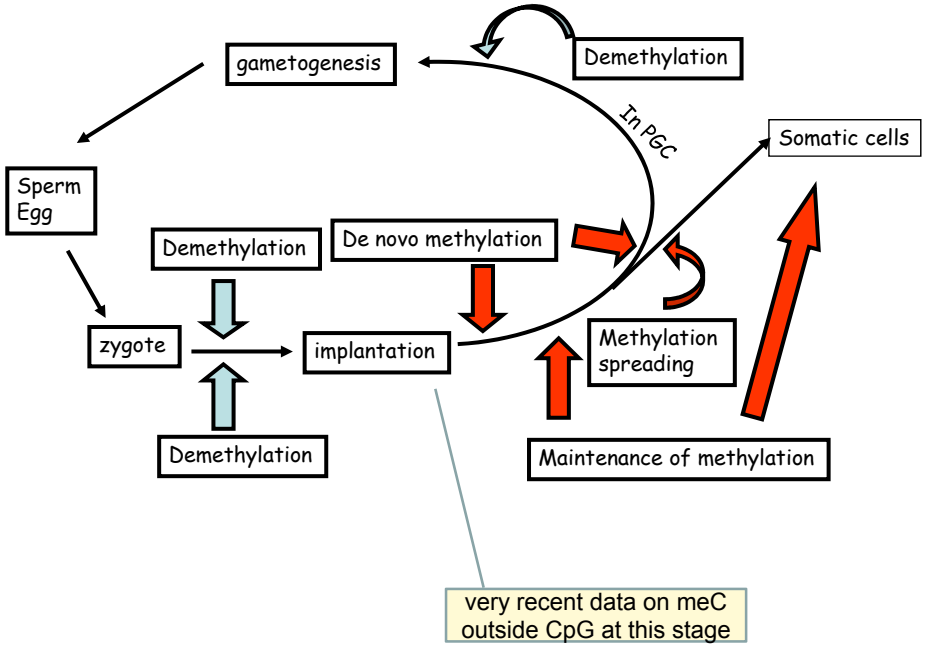
Nature 2007, 447: 425-432.

Determinants of epigenetic transmission:

1. histone marks
2. CpG methylation

The DNMT1 methyl transferase act to maintain the methylation profiles when cells divide, as follows:





Vol 462|19 November 2009|doi:10.1038/nature08514 nature

Human DNA methylomes at base resolution show widespread epigenomic differences

Ryan Lister[1]*, Mattia Pelizzola[1]*, Robert H. Dowen[1], R. David Hawkins[2], Gary Hon[2], Julian Tonti-Filippini[4], Joseph R. Nery[1], Leonard Lee[2], Zhen Ye[2], Que-Minh Ngo[2], Lee Edsall[2], Jessica Antosiewicz-Bourget[5,6], Ron Stewart[5,6], Victor Ruotti[5,6], A. Harvey Millar[4], James A. Thomson[5,6,7,8], Bing Ren[2,3] & Joseph R. Ecker[1]

DNA cytosine methylation is a central epigenetic modification that has essential roles in cellular processes including genome regulation, development and disease. Here we present the first genome-wide, single-base-resolution maps of methylated cytosines in a mammalian genome, from both human embryonic stem cells and fetal fibroblasts, along with comparative analysis of messenger RNA and small RNA components of the transcriptome, several histone modifications, and sites of DNA–protein interaction for several key regulatory factors. Widespread differences were identified in the composition and patterning of cytosine methylation between the two genomes. Nearly one-quarter of all methylation identified in embryonic stem cells was in a non-CG context, suggesting that embryonic stem cells may use different methylation mechanisms to affect gene regulation. Methylation in non-CG contexts showed enrichment in gene bodies and depletion in protein binding sites and enhancers. Non-CG methylation disappeared upon induced differentiation of the embryonic stem cells, and was restored in induced pluripotent stem cells. We identified hundreds of differentially methylated regions proximal to genes involved in pluripotency and differentiation, and widespread reduced methylation levels in fibroblasts associated with lower transcriptional activity. These reference epigenomes provide a foundation for future studies exploring this key epigenetic modification in human disease and development.

EPIGENOMICS

Methylation matters

Dirk Schübeler

Genome-wide maps of methylated cytosine bases at single-base-pair resolution in human cells reveal distinct differences between cell types. These maps provide a starting point to decode the function of this enigmatic mark.

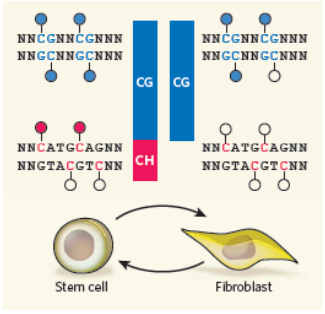


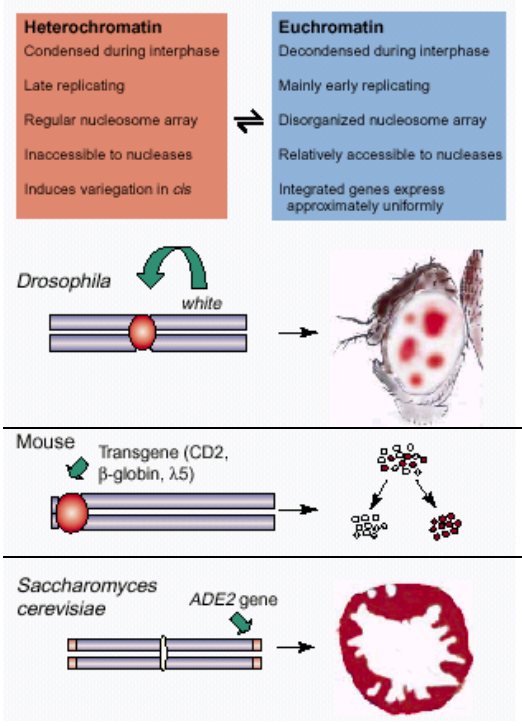
Figure 1 | DNA methylation patterns differ between stem cells and differentiated cells[1]. In stem cells, regions of DNA with CpG methylation (blue) are mostly uniformly methylated, whereas this modification is more heterogeneous in fibroblasts. Non-CpG methylation (red), which occurs primarily at CA nucleotides, is detected only in stem cells, yet is asymmetric and more scarce and patchy than CpG methylation. If fibroblasts are converted to induced pluripotent stem cells they regain non-CpG methylation. Filled circles, methylated cytosines; unfilled circles, unmethylated cytosines. H stands for A, C or T; N stands for any nucleotide.

CpG methylation mechanisms of establishment and maintenance explain some properties of epigenetics

However, some organisms do not use CpG methylation as epigenetic marks...

Could chromatin marks (i.e. histone modifications) also participate in the epigenetic transmission of chromatic status ?

We need to examine some further aspects of heterochromatin (HC) ...

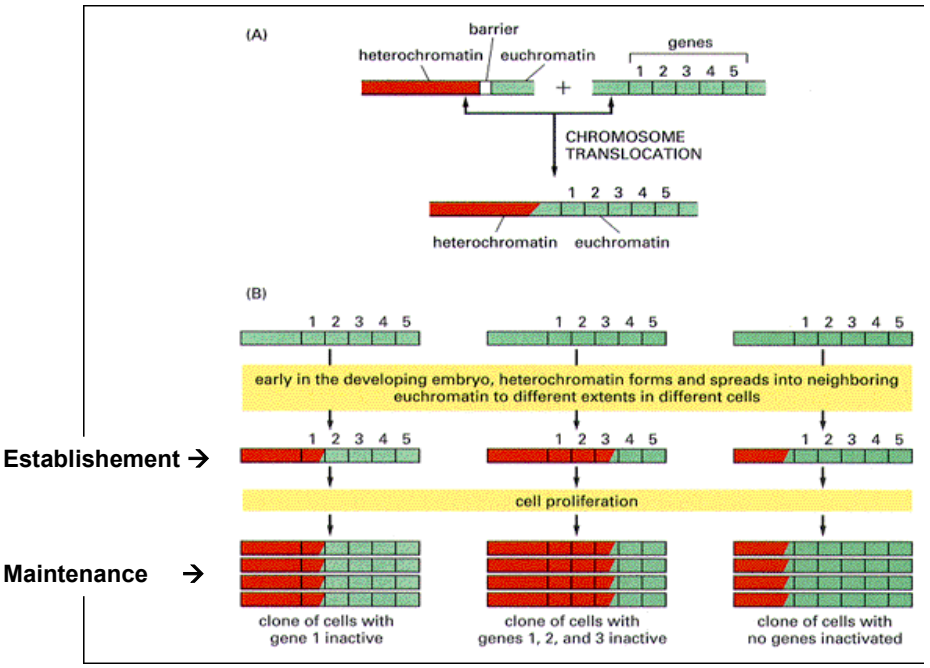


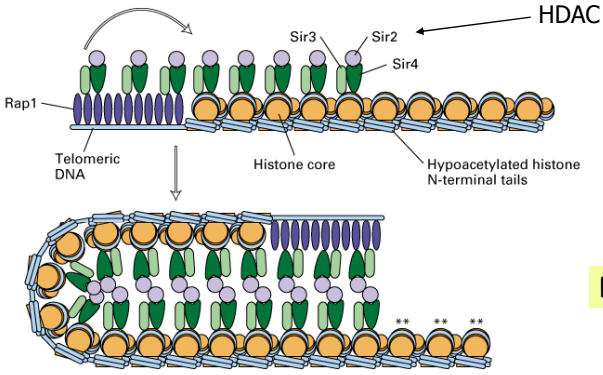
PEV = Positional Effect Variegation
Silencing effects of heterochromatin in different organisms.

Translocation of the Drosophila *white* gene close to the centromere results in mosaic (variegated) expression in the eye.

Insertion of transgenes into peri-centromeric regions in mice gives expression in a proportion of cells from expressing tissues. Expression is clonally stable through multiple cell divisions but also undergoes stochastic switching between the active and inactive states.

Insertion of the Ade2 gene into the telomeric region of Saccharomyces cerevisiae causes metastable silencing of the gene resulting in sectored colonies containing white (expressing) and red (nonexpressing) cells.





replicazione dei cromosomi lineari
struttura dei telomeri
telomerasi

budding yeast

Model for the formation of telomeric heterochromatin. Black lines wrapped around nucleosomes represent DNA.

(a) Core telomeric heterochromatin in wild-type cells containing only a single genomic copy of SIR3. It is proposed that the RAP1-containing telosome folds back onto subtelomeric regions. In this manner, RAP1–SIR–histone interactions are all required for stability of the complex.

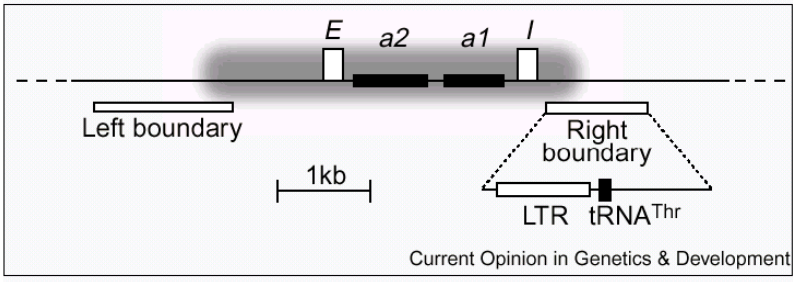
(b) Upon SIR3 overexpression, telomere position effect and the presence of SIR3 is extended up to some 16–20 kb from the telomere. SIR3 overexpression causes loss of some SIR4 and most SIR2 from the complex. Due to the interdependence of RAP1–SIR3–H4 interactions, and because all three SIR proteins are required for extension of heterochromatin by SIR3, it is proposed that the complex necessary for the initiation of heterochromatin formation requires RAP1, the SIR proteins and H4.

Elements that promote heterochromatin are called "silencers"

Elements that stop heterochromatin spreading are called "insulators" or "boundaries"

Boundaries of heterochromatin are cis-elements regulated by trans-factors

Chromosomal boundaries in S. cerevisiae
Xin Bi* and James R Broach †
Current Opinion in Genetics & Development 2001, 11:199–204

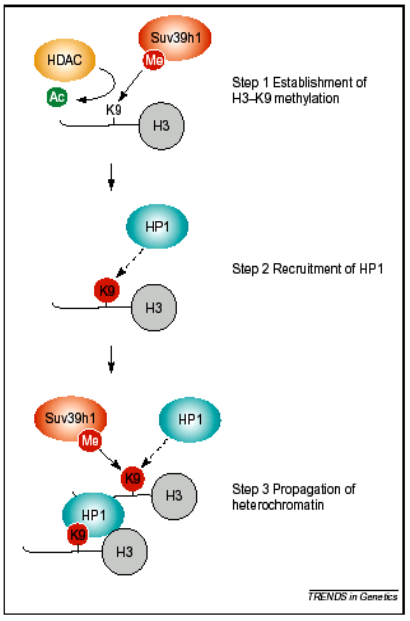


Organization of heterochromatin barriers surrounding *HMR*. The *HMR* locus is diagrammed, showing the location of the mating type genes *a1* and *a2*, the *E* and *I* heterochromatin organizing centers (silencers) and the left and right heterochromatin barriers as defined in [1••]. Background shading indicates the extent of the repressed domain. An expansion of the right barrier shows the location of a Ty1 LTR and the gene for $tRNA^{Thr}$. Earlier evidence suggested that both these elements contribute to barrier activity [1••] but more recent data indicate that the $tRNA^{Thr}$ gene is necessary and sufficient for full barrier activity

How does heterochromain “spread” ?

The simples model is that all the enzymes that co-operate to make heterochromatin make part of complexes that also contain “modules” for all the possible HC markers on DNA or histones.

The original, simple model of spreading



Active genes have HATs locally that acetylate many hstone positions, especially H3-K9.

Acetylation of H3-K9 prevents H3-K9 methylation; thus, **step one** involves the deacetylation of H3-K9 by specific HDACs, and the subsequent methylation by a histone lysine acetyltransferase (HKMT), such as Suv39H1.

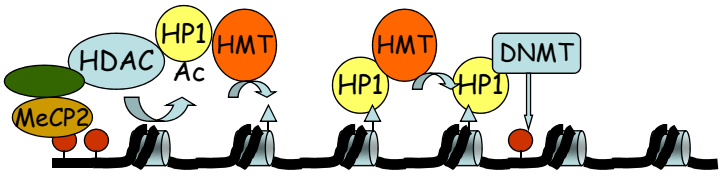
In **step two**, HP1 selectively recognizes methylated H3-K9 through its **chromodomain**

Step three involves propagation of heterochromatin through HP1 recruitment of Suv39H1 via protein association. Suv38H1 is one methyltransferase specific for H3-K9

Once nearby H3-K9 sites are methylated, additional heterochromatin-associated protein 1 (HP1) molecules recognize this mark through its chromo-domain. HP1 then recruits Suv39H1 through protein interactions recruiting further H3-K9 methylation activity.

.

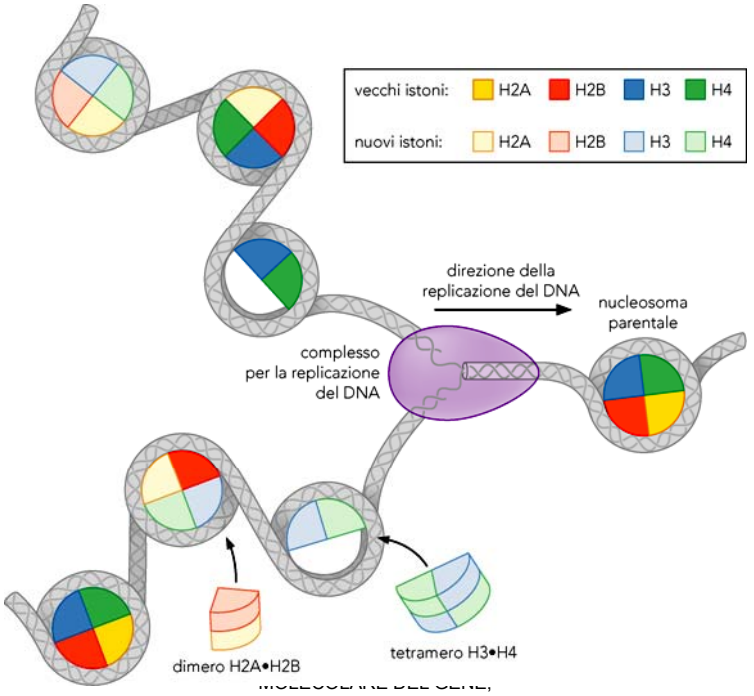
A model for heterochromatin establishment and propagation

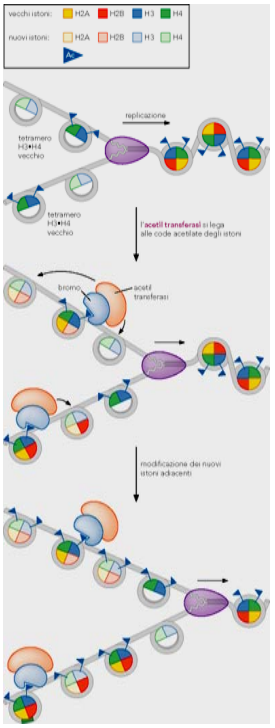


Ac – acetyl groups

- methyl lysine 9 (H3)

- methyl CpG





imprinting

Some genes show a very interesting properties:

.

1. They are expressed only from 1 allele, and their expression depends on whether they are inherited from either the mother or the father.
2. One allele is DNA CpG hypermethylated in a specific control region
3. The methylation is made in germinal cells

This is called germinal epigenetic inheritance (to distingiush from somatic epigenetic inheritance discussed above).

We say these genes are **imprinted**

and the phenomenon is known **as genomic imprinting**

Genomic imprinting was discovered twenty years ago, as an outcome of nuclear transplantation experiments in the mouse.

About 80-90 genes are known to be imprinted in mice and humans, and imprinting is conserved in ruminant species as well.

Many imprinted genes are involved in foetal development and growth, and some influence behaviour.

What is the biological role of imprinting ?

It represents possibly a form of sexual predominance, where a character from one of the partners is programmed to prevail upon the one from the other partner

Usually grouped into Large Chromosomal Domains (LCD)

Transgenic studies have identified sequence elements in these domains that are essential for the imprinted gene expression:

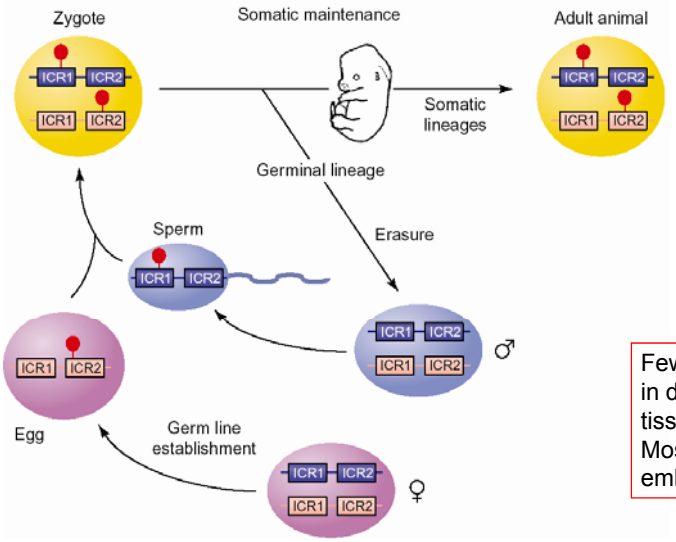
the ICRs - imprinting control regions

- regulated by epigenetic modifications
- up to several kilobases in length
- rich in CpG dinucleotides (many correspond to CpG islands)
- CpG DNA methylation on one of the two parental alleles (DMR=differentially methylated region).

At most ICRs, the allelic methylation originates from the egg.
At only a few, it is established during spermatogenesis.

Following fertilisation, allelic methylation marks are maintained throughout development and they mediate imprinted expression.

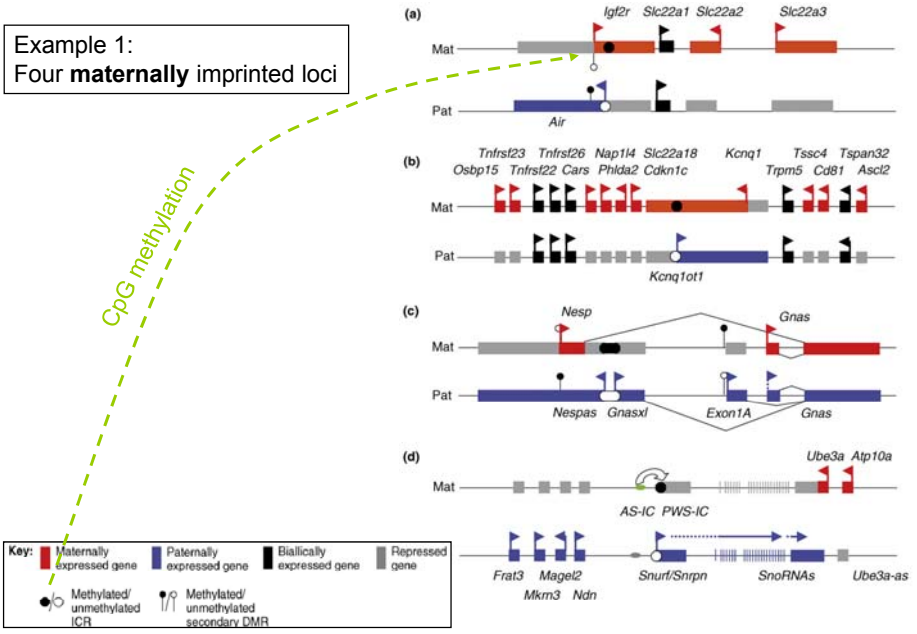
Parental imprints are established during oogenesis, or spermatogenesis, at sequence elements that control the imprinted expression (ICRs). After fertilisation of the egg by the sperm, these imprints are maintained throughout development. DNA methylation (lollypop) is the most consistent hallmark of imprints. Two examples of ICRs are depicted: ICR with paternally-derived (ICR1); ICR with maternally-derived DNA methylation (ICR2).



Few genes expressed in developing and adult tissues
Most of them in extra-embryonic tissues

Loci showing imprinting are often very complex

Example 1:
Four **maternally** imprinted loci



Therefore

(at least in the yeast *S. pombe*)

The RNA interference machinery triggered by endogenous siRNA can lead to both:

post-transcriptional silencing (through the RISC pathway)

and

transcriptional silencing (heterochromatin formation through RITS)

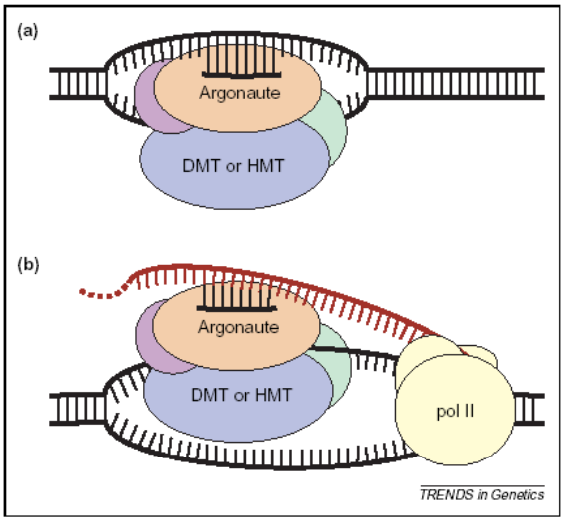


Figure 1. Models for siRNA-directed chromatin modification. DNA (DMT) or histone (HMT) methyltransferase activity is thought to be recruited to target loci by a RITS-like complex that includes an argonaute protein and an siRNA. Two potential mechanisms for target recognition are: **(a)** siRNA binding to target DNA; or **(b)** siRNA binding to nascent transcripts produced from target DNA by RNA polymerase II (pol II).

Two different hypothetical modes RITS can use ss-RNAi to target the gene it derives from

In the first model, a DNA-RNA hybrid is depicted and transcription is **not** required.

In the second model, transcription initiation is required, since the siRNA is thought to form a RNA-RNA hybrid with the nascent transcript.

Indeed, RNA Pol activity seems necessary to target the RITS complex to the locus, although a complete demontration is still lacking.