Epigenomics

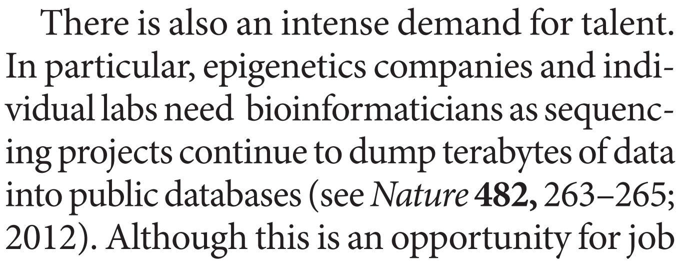
- **Part 1: Intro to epigenomics/technologies**
- Part 2: Computational methods

Mark D. Robinson, Statistical Genomics, IMLS

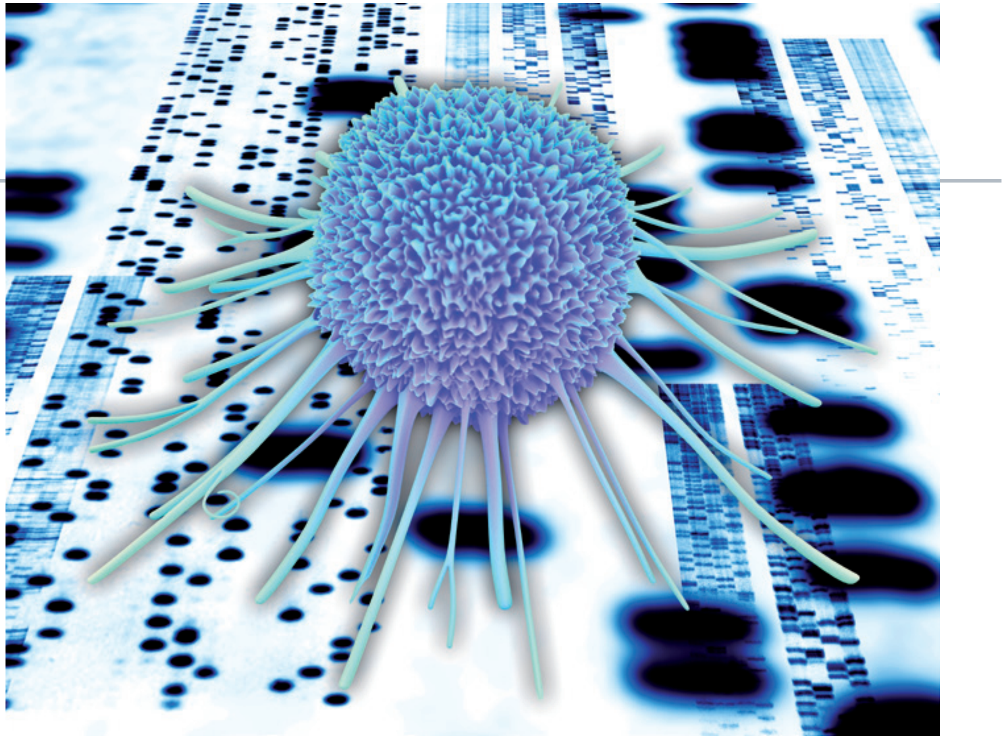
Overview of this lecture

- Some definitions
- Molecular basis: DNA methylation, histone variants and post-translational modifications, RNA
- Some compelling examples: agouti mice, Dutch “hunger winter”, etc.
- GWAS to EWAS
- Epigenetics and disease (cancer, diabetes)
- Epigenetic drugs

A plug for (bioinformatics/ statistics in) epigenomics



There is also an intense demand for talent. In particular, epigenetics companies and individual labs need bioinformaticians as sequencing projects continue to dump terabytes of data into public databases (see *Nature* **482,** 263–265; 2012). Although this is an opportunity for job



PASIEKA/SPL

Computer reconstruction of a cancer cell on a DNA autoradiogram.

EPIGENETICS

Marked for success

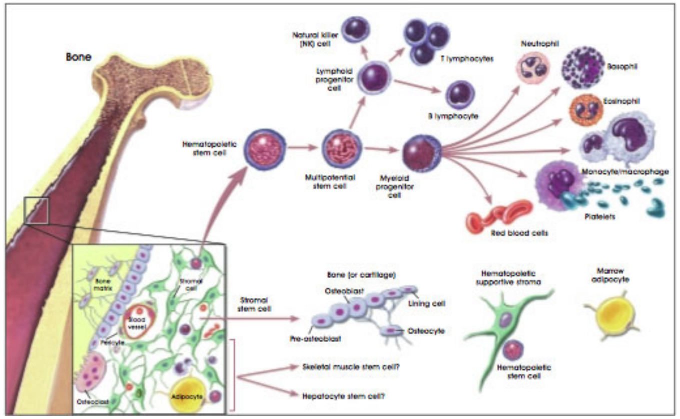
The growing field of cancer epigenetics demands computational expertise and translational research experience. Qualified practitioners are in high demand.

Genetics and Epigenetics

Genetics **can** explain differences between individuals.

Epigenetics can explain difference both ***between*** and ***within*** individuals.

Each cell type has the same DNA sequence, but **very** different epigenetic state.



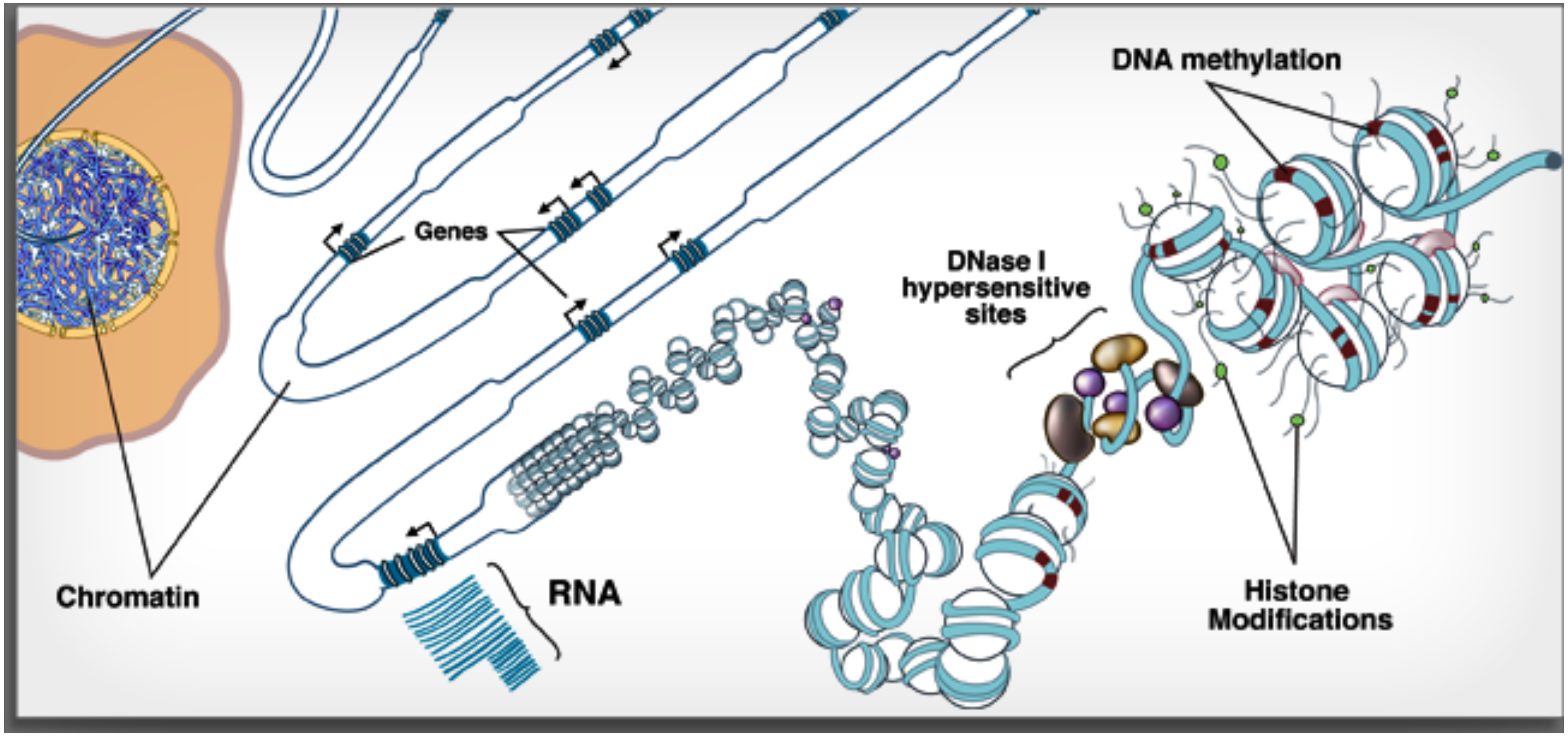
Epigenetics definition

Epi - "on top of" or "in addition to"

“Epigenetics”:

- **heritable alterations in gene expression caused by mechanisms other than changes in DNA sequence.**
- the study of the mechanisms of temporal and spatial control of gene activity during the development of complex organisms
- "epigenetic code" has been used to describe the set of epigenetic features that create different phenotypes in different cells

Molecular basis of epigenetics



Epigenetic analogies

Computer: Two computers have the same specifications and software packages installed ("identical twins"). One user is doing word processing and email, the other is doing email and image processing. That is, the underlying instructions are common, but are being used ("expressed") differently.

Music: Genetics is the music, epigenetics is the musician's interpretation of the notes, rhythm, etc.

Television: You can fine tune the hue, brightness, contrast, etc., but you cannot change the original broadcast.

Recipe: The recipe ("genes") represent the set of instructions for baking something; depending on the person baking, there may be a different result

Script*: The Romeo and Juliet script is a fixed document ("genes"), but the director's interpretation ("epigenetics") can vary drastically (e.g. Baz Luhrmann 1996 Hollywood vs. Shakespeare).

*From The Epigenetics Revolution by Nessa Carey

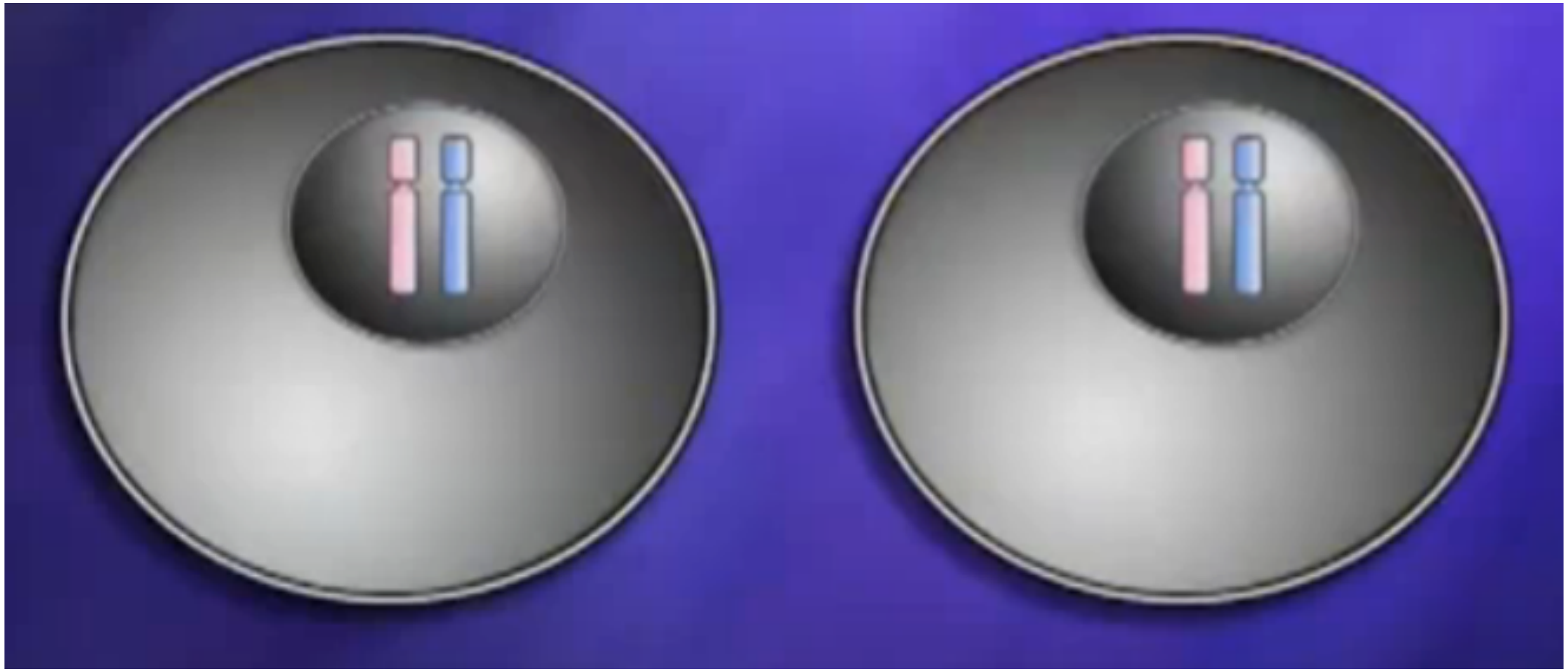
Some compelling examples of epigenetics

1. X-inactivation (coat colour in cats)
2. agouti mice (maternal diet affects coat colour)
3. Dutch "hunger winter" (children conceived/born during famine)

Example 1: X-inactivation

Females have 2 **X**-chromosomes, but one of them is (mostly) silenced. In early embryogenesis, either the maternal or paternal allele is silenced at random, but any subsequent cell divisions will maintain the silenced X. For example, calico coat colour is determined by an X-inactivation outcome (gene is on the X-chromosome).

X-inactivation

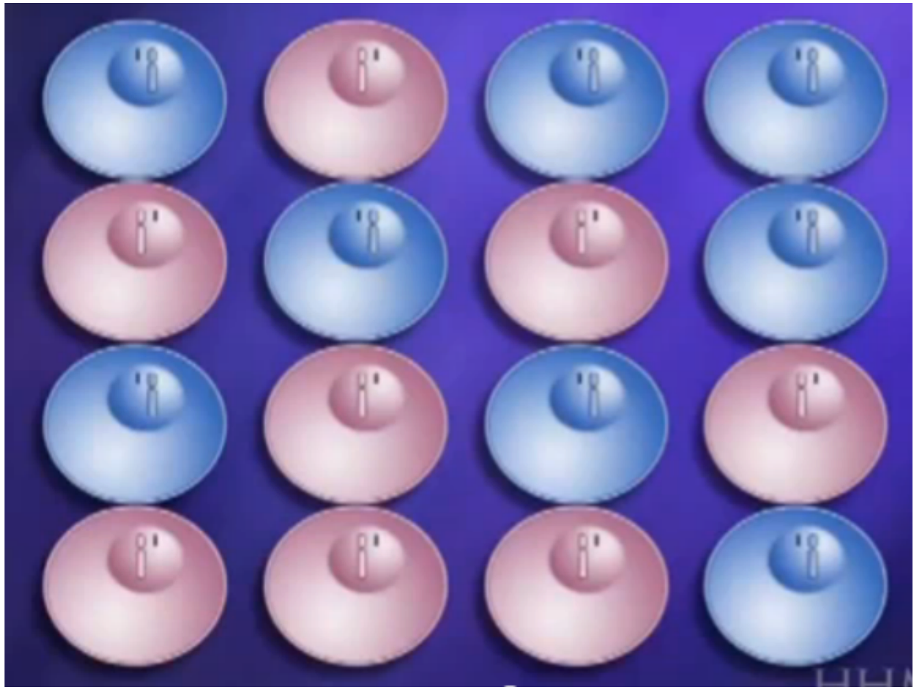


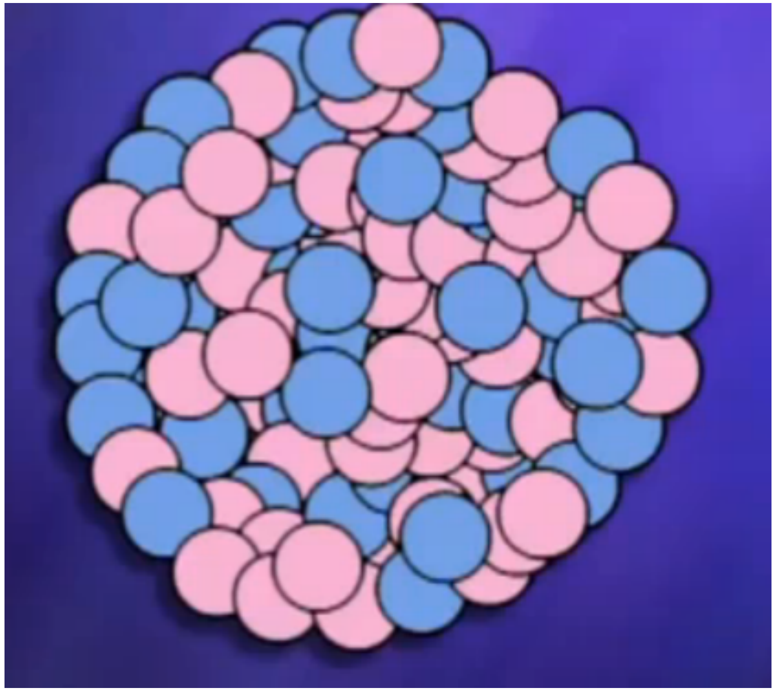
Two cells (from a female), each with 2 X-chromosomes

X-inactivation

One of the X chromosomes is randomly silenced.

X-inactivation

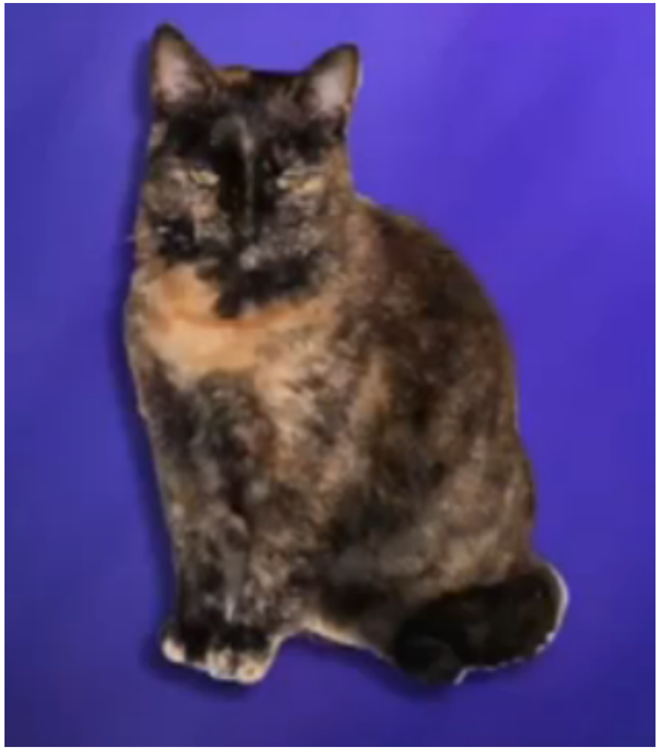




Cells divide, but preserve the inactivated X.

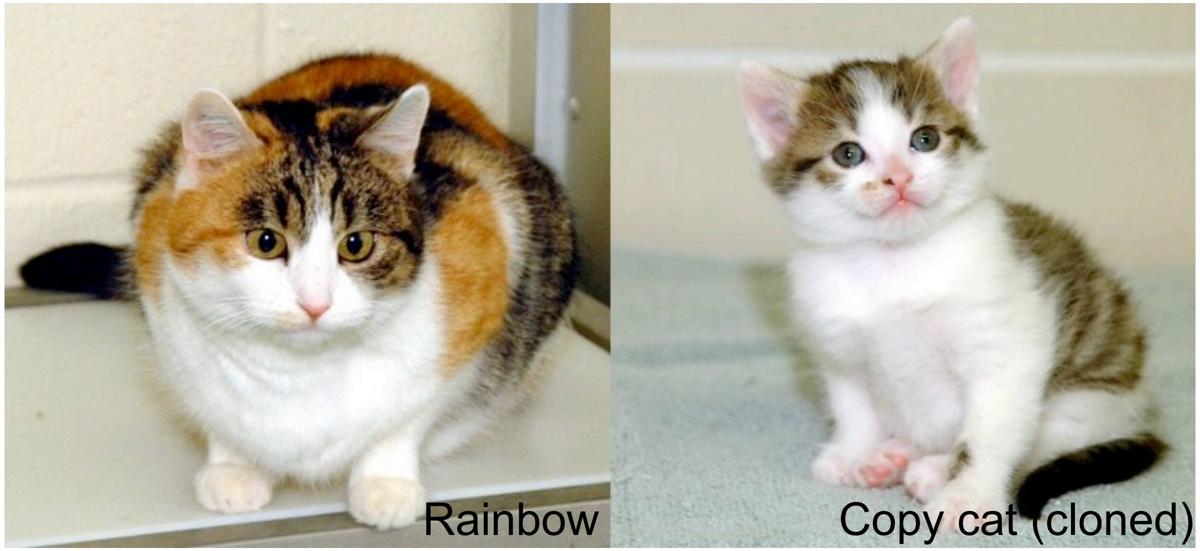
http://www.youtube.com/watch?v=n330FzHpI90

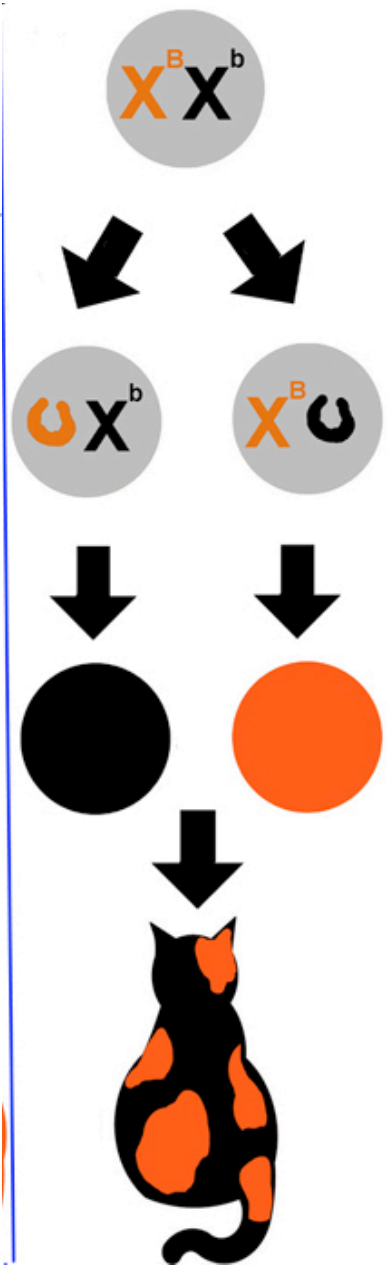
X-inactivation



Result: patchy coat colours in female calico cats.

X-inactivation (randomly initiated)



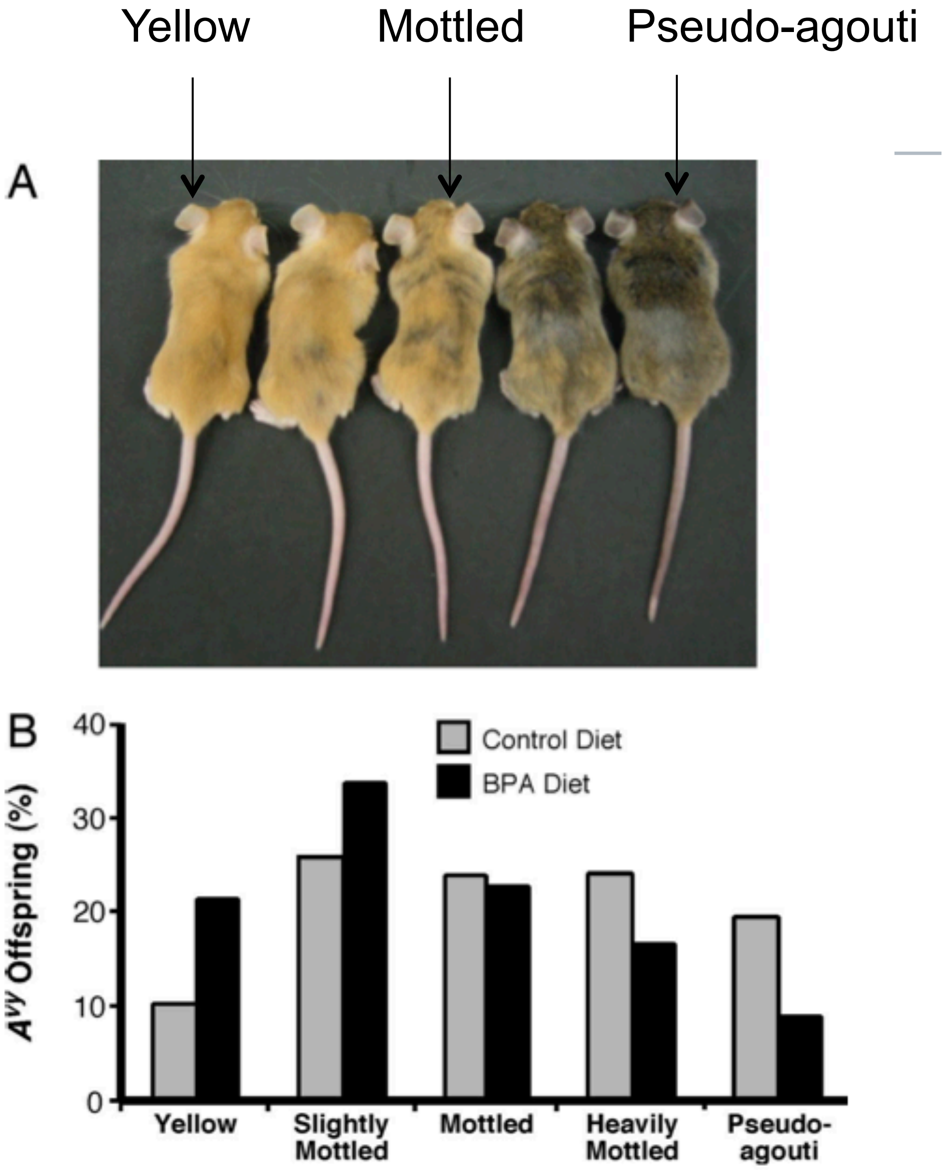


Genetically identically, epigenetically distinct
(Genetic Savings and Clone)

Example 2: Agouti mice

Observation: coat colour in offspring is strongly affected by mother’s diet.

Molecularly, what is driving this?

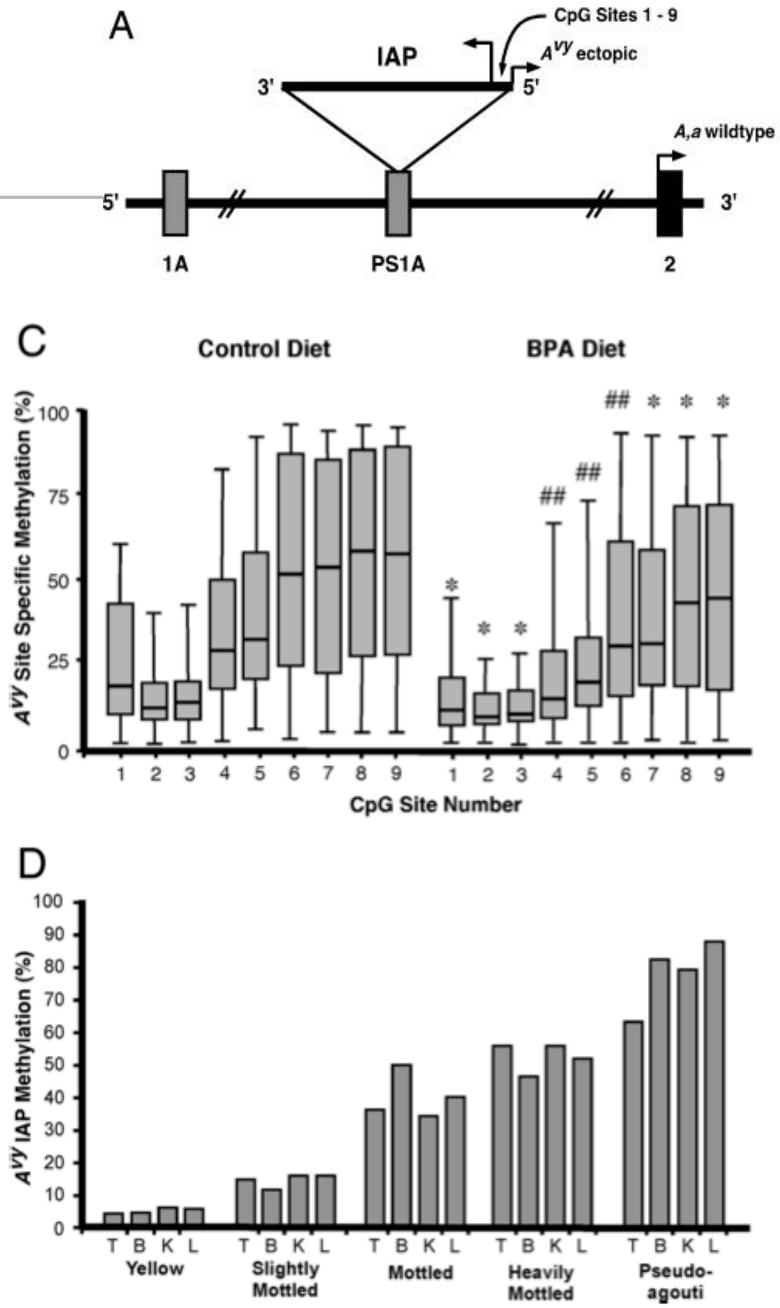


Agouti mice

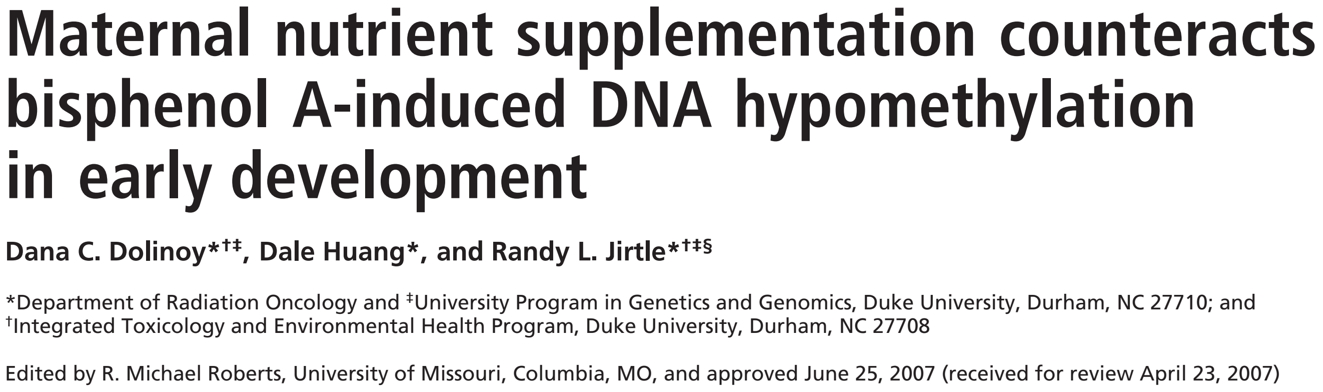
Observations:

1. Methylation level (at promoter upstream of agouti gene) is strongly associated with coat colour.

2. Diet affects methylation level (in several tissues).



Agouti mice



Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development

Dana C. Dolinoy*[†‡], Dale Huang*, and Randy L. Jirtle*[†‡§]

*Department of Radiation Oncology and [‡]University Program in Genetics and Genomics, Duke University, Durham, NC 27710; and [†]Integrated Toxicology and Environmental Health Program, Duke University, Durham, NC 27708

Edited by R. Michael Roberts, University of Missouri, Columbia, MO, and approved June 25, 2007 (received for review April 23, 2007)

Example 3: Dutch “hunger winter”

-- Food shortage in the Netherlands near the end of World War II

“… children of the women who were pregnant during the famine were smaller, as expected. However, surprisingly, when these children grew up and had children those children were also smaller than average.”

http://en.wikipedia.org/wiki/Dutch_famine_of_1944

(Also brought about evidence in the discovery of Coeliac disease)

Example 3: Dutch “hunger winter”

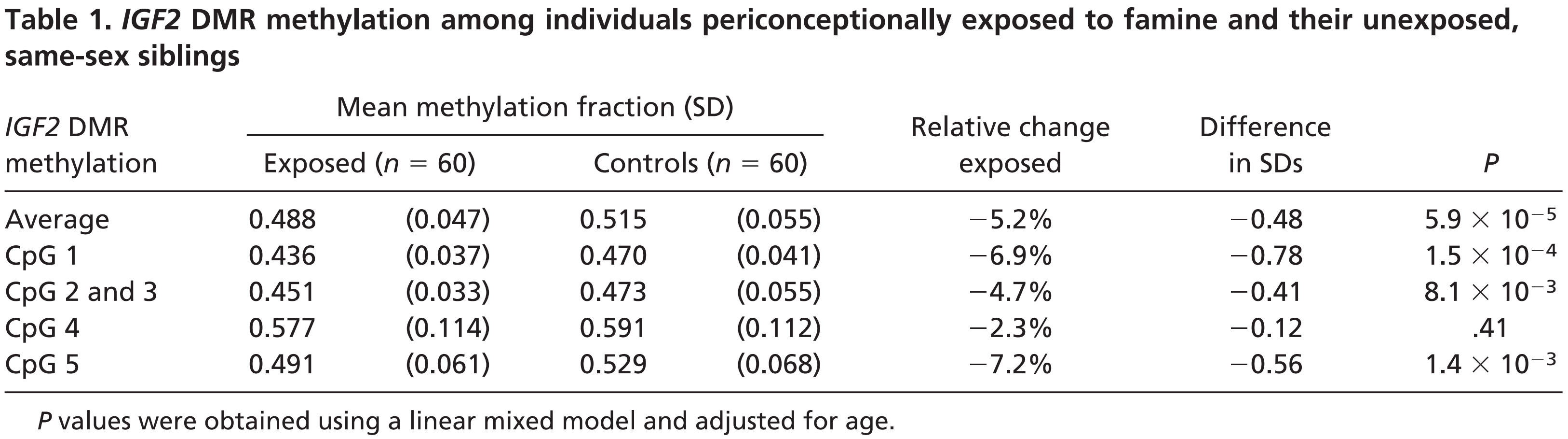


Table 1. *IGF2* DMR methylation among individuals periconceptionally exposed to famine and their unexposed, same-sex siblings

IGF2 DMR methylation	Mean methylation fraction (SD)				Relative change exposed	Difference in SDs	*P*
	Exposed (n = 60)		Controls (n = 60)				
Average	0.488	(0.047)	0.515	(0.055)	−5.2%	−0.48	5.9×10^{-5}
CpG 1	0.436	(0.037)	0.470	(0.041)	−6.9%	−0.78	1.5×10^{-4}
CpG 2 and 3	0.451	(0.033)	0.473	(0.055)	−4.7%	−0.41	8.1×10^{-3}
CpG 4	0.577	(0.114)	0.591	(0.112)	−2.3%	−0.12	.41
CpG 5	0.491	(0.061)	0.529	(0.068)	−7.2%	−0.56	1.4×10^{-3}

P values were obtained using a linear mixed model and adjusted for age.

Example 3: Dutch “hunger winter”

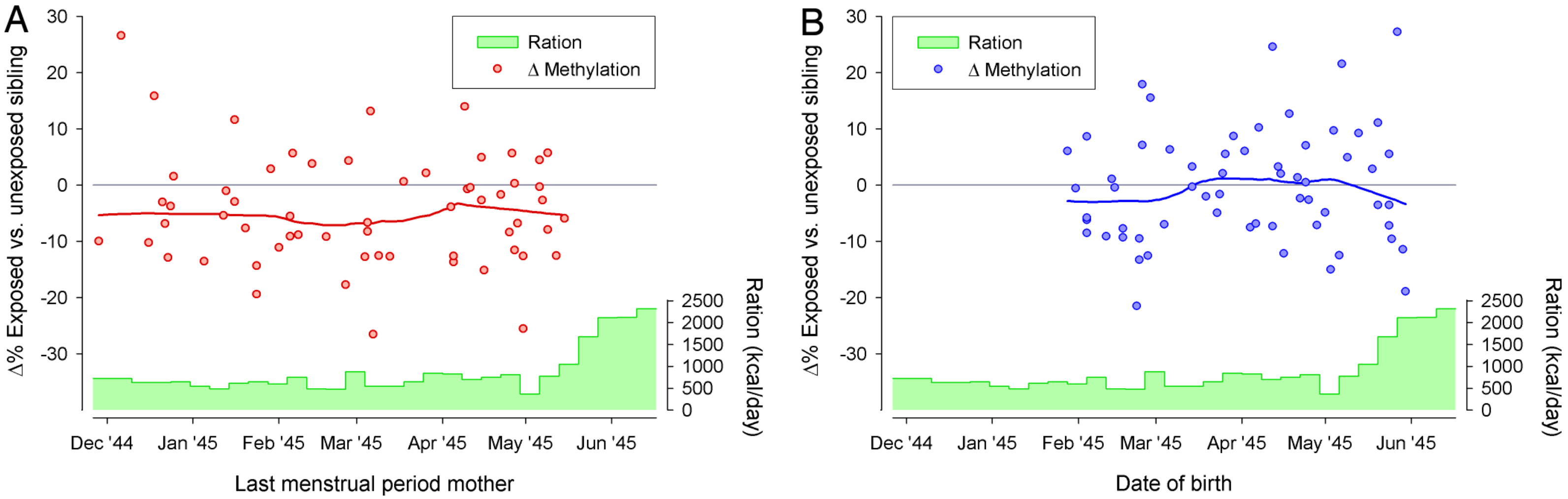
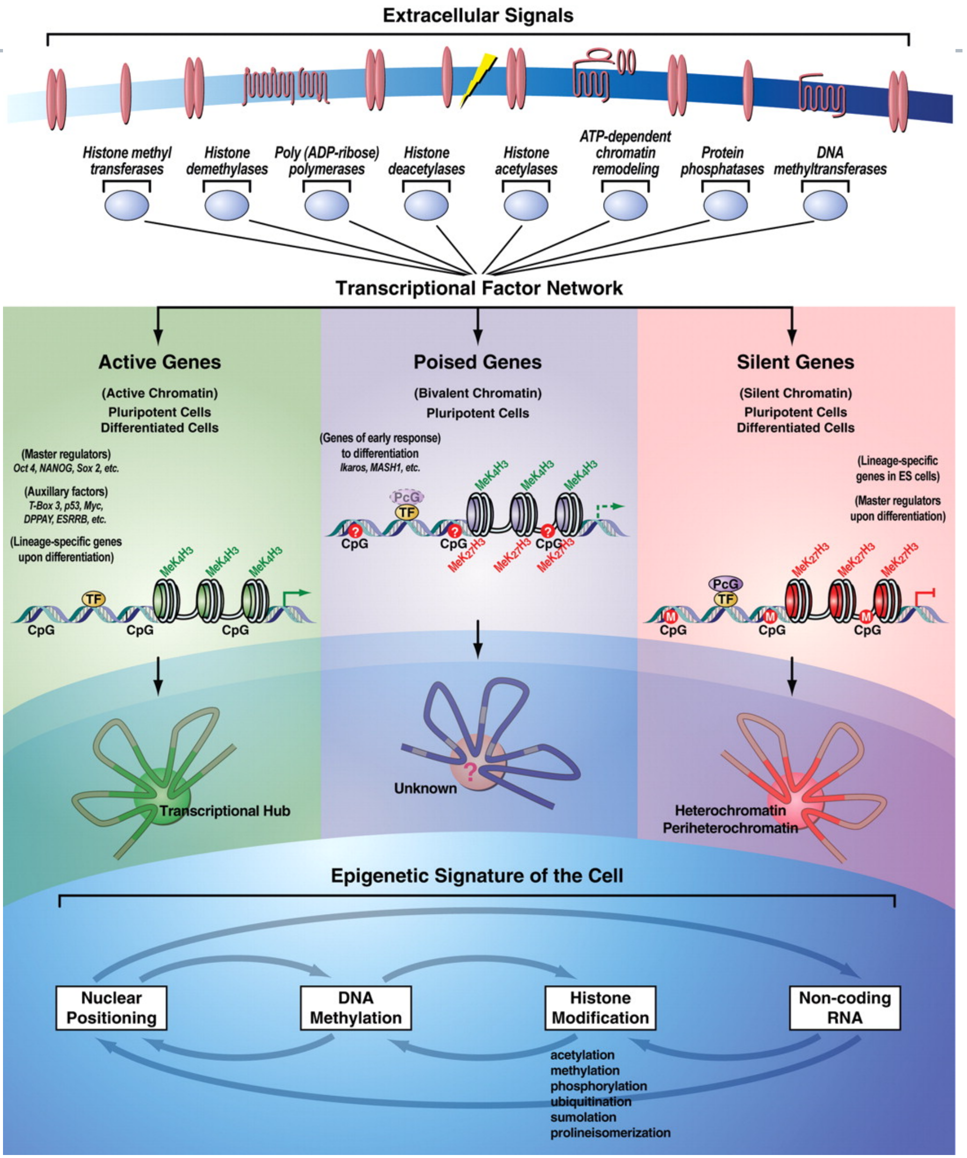


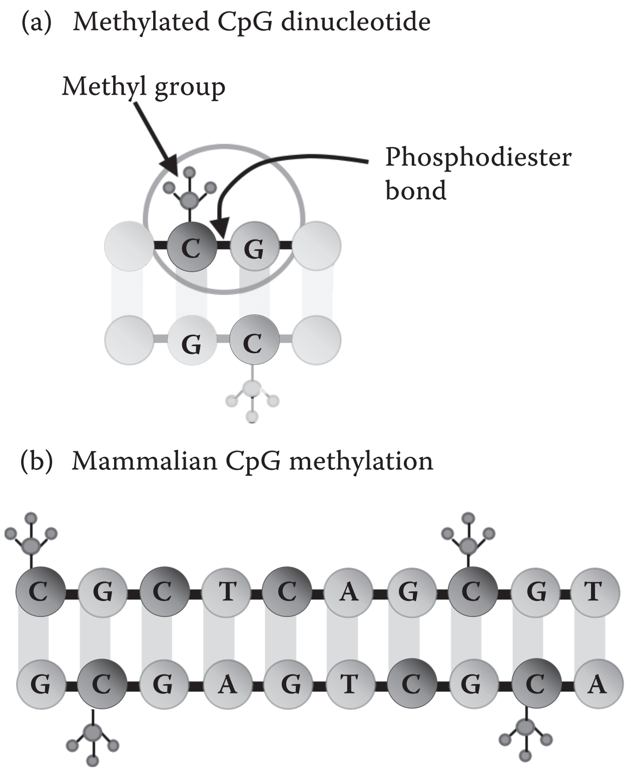
Fig. 1. Difference in *IGF2* DMR methylation between individuals prenatally exposed to famine and their same-sex sibling. (*A*) Periconceptional exposure: Difference in methylation according to the mother’s last menstrual period (a common estimate of conception) before conception of the famine-exposed individual. (*B*) Exposure late in gestation: Difference in methylation according to the date of birth of the famine-exposed individual. To describe the difference in methylation according to estimated conception and birth dates, a lowess curve (red or blue) is drawn. The average distributed rations (in kcal/day) between December 1944 and June 1945 are depicted in green.

Epigenetics in concert with TF network

"… **suggests** that epigenetic players such as histone modifications, DNA methylation, the alteration of chromatin structure due to chromatin remodeling, and non-coding RNAs represent another crucial mechanism, besides the transcriptional factor network, which is designed by nature for the regulation of gene expression and cellular differentiation. Elucidating epigenetic mechanisms **promise** to have important implications for advances in stem cell research and nuclear reprogramming and **may offer** novel targets to combat human disease **potentially** leading to new diagnostic and therapeutic avenues in medicine."



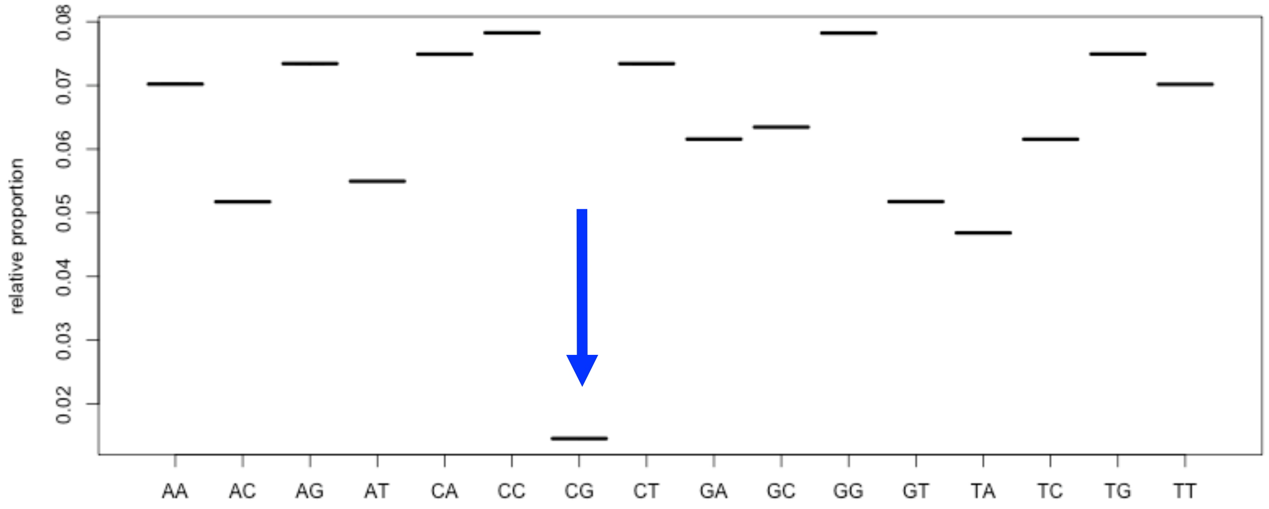
DNA methylation



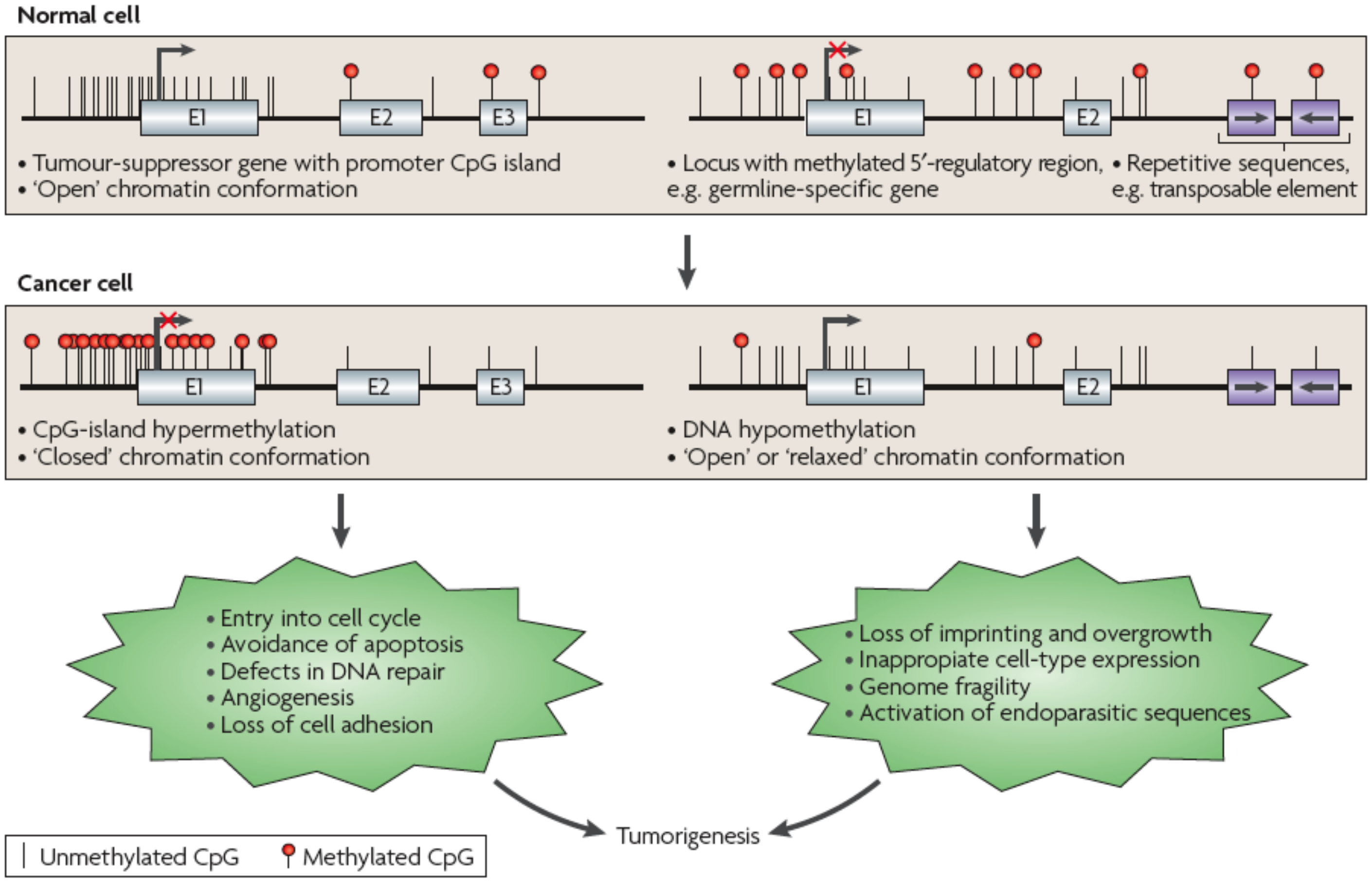
Covalent addition of methyl group (CH_3) to cytosine (almost exclusively at CpG sites in mammals); **binary status** at individual sites

CpG islands

- CG dinucleotides are under-represented in the genome, but often occur in "clusters" called CpG islands (CGIs); various CGI definitions

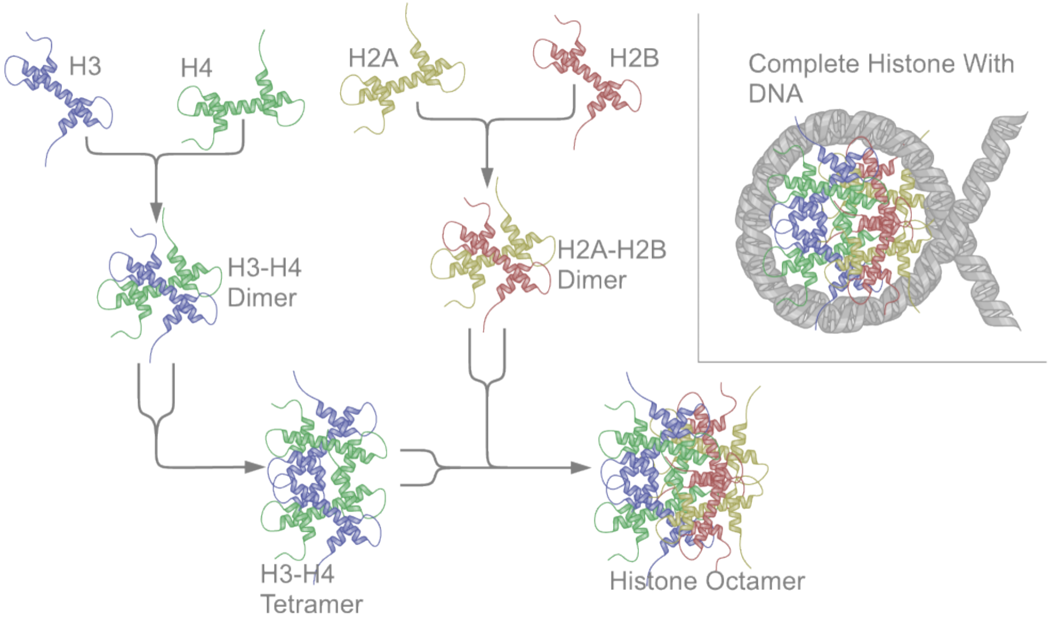


Dogma: CpG methylation and transcription


Normal cell
E1
E2
E3
• Tumour-suppressor gene with promoter CpG island
• 'Open' chromatin conformation
E1
E2
• Locus with methylated 5′-regulatory region, e.g. germline-specific gene
• Repetitive sequences, e.g. transposable element
Cancer cell
E1
E2
E3
• CpG-island hypermethylation
• 'Closed' chromatin conformation
E1
E2
• DNA hypomethylation
• 'Open' or 'relaxed' chromatin conformation
• Entry into cell cycle
• Avoidance of apoptosis
• Defects in DNA repair
• Angiogenesis
• Loss of cell adhesion
• Loss of imprinting and overgrowth
• Inappropiate cell-type expression
• Genome fragility
• Activation of endoparasitic sequences
Tumorigenesis
| Unmethylated CpG
Methylated CpG

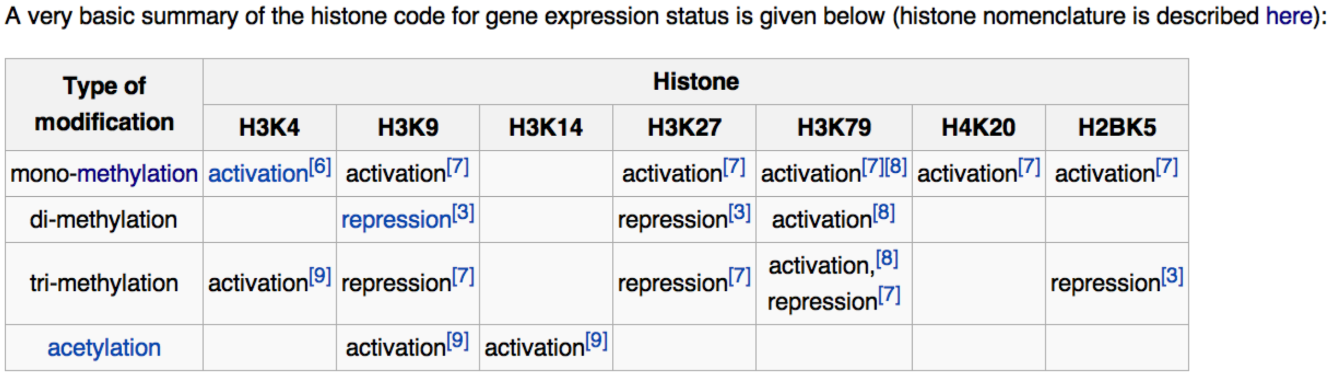
Esteller, Nature Reviews Genetics 2005

Histone variants and post-translation modifications



Two of each of H2A, H2B, H3 and H4 form a “nucleosome”, which 147bp of DNA can wrap around

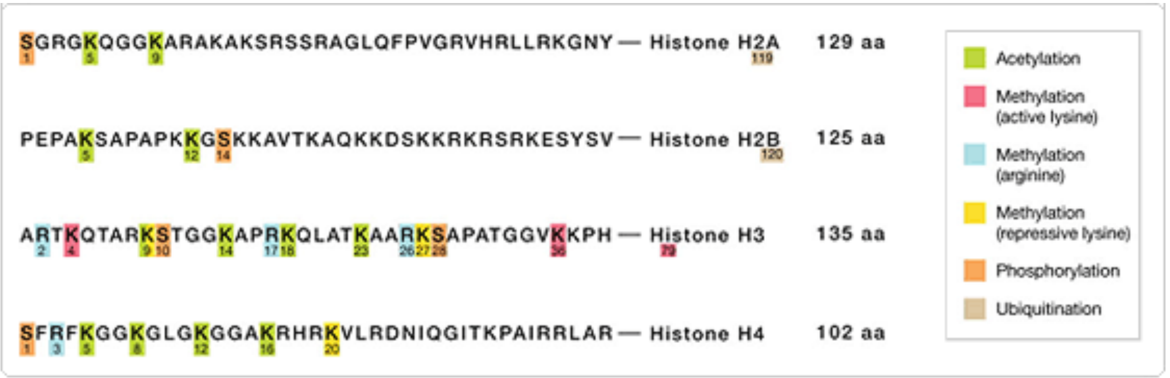
Histone variants and post-translation modifications



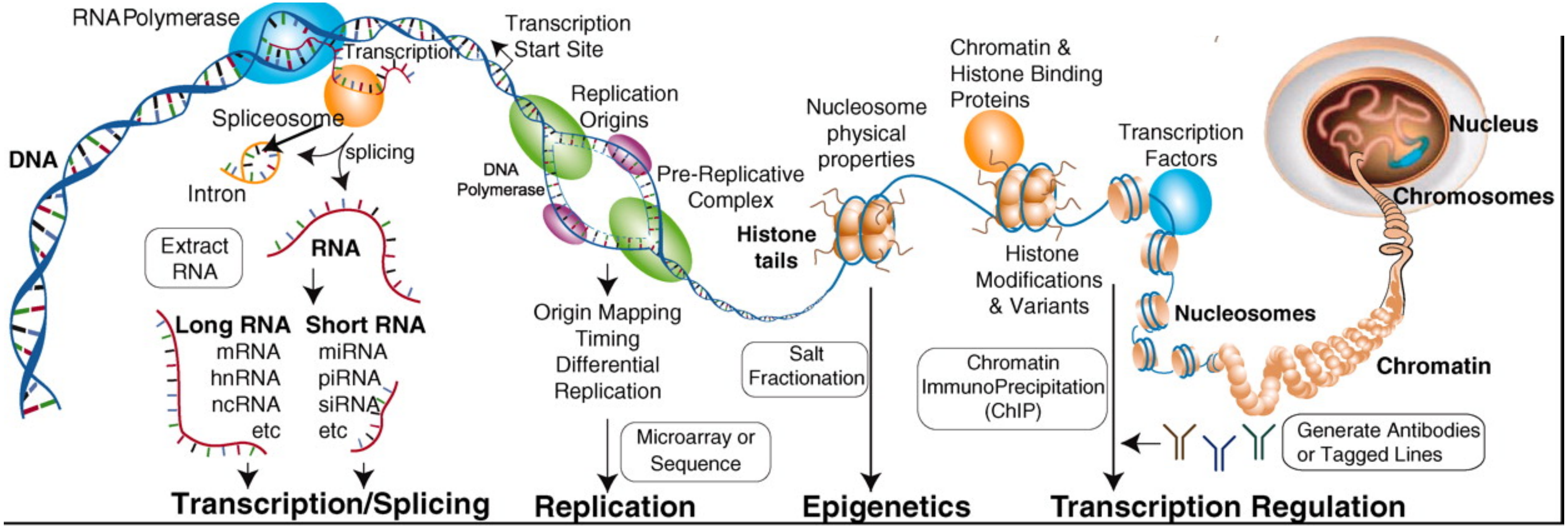
A very basic summary of the histone code for gene expression status is given below (histone nomenclature is described here):

Type of modification	Histone						
	H3K4	H3K9	H3K14	H3K27	H3K79	H4K20	H2BK5
mono-methylation	activation[6]	activation[7]		activation[7]	activation[7][8]	activation[7]	activation[7]
di-methylation		repression[3]		repression[3]	activation[8]		
tri-methylation	activation[9]	repression[7]		repression[7]	activation,[8] repression[7]		repression[3]
acetylation		activation[9]	activation[9]				

- H3K4me3 is found in actively transcribed promoters, particularly just after the transcription start site.
- H3K9me3 is found in constitutively repressed genes.
- H3K27me is found in facultatively repressed genes.[7]
- H3K36me3 is found in actively transcribed gene bodies.
- H3K9ac is found in actively transcribed promoters.
- H3K14ac is found in actively transcribed promoters.


SGRGKQGGKARAKAKSRSSRAGLQFPVGRVHRLLRKGNY — Histone H2A
1
5
9
119
129 aa
PEPAKSAPAPKKGSKKAVTKAQKKDSKKRKRSRKESYSV — Histone H2B
5
12
14
120
125 aa
ARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPH — Histone H3
2
4
9 10
14
17 18
23
26 27 28
56
79
135 aa
SFRFKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLAR — Histone H4
1
3
5
8
12
16
20
102 aa
Acetylation
Methylation (active lysine)
Methylation (arginine)
Methylation (repressive lysine)
Phosphorylation
Ubiquitination

Various other epigenetic (and regulator) factors



Roy et al. *Science* 2010

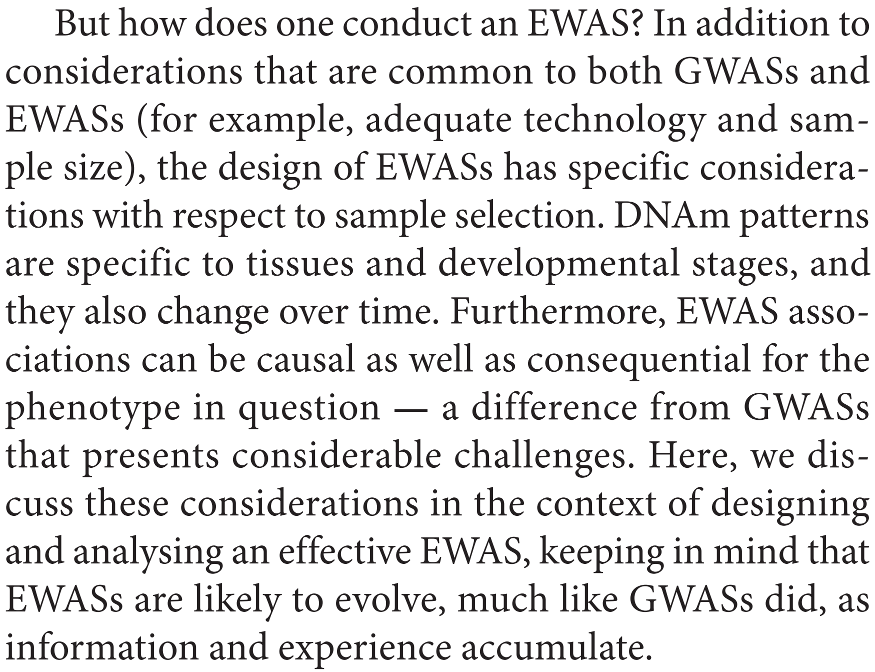
Genome/Epigenome Wide Association Studies (GWAS/EWAS)

GWAS – associating genotype to phenotype

EWAS – association "epitype" to phenotype

Genetics does not explain a high amount of causality in common diseases

Challenge is far greater – there is 1 genome, but 1000s of epigenomes (100s of cell types, 10s-100s of epigenome dimensions)



But how does one conduct an EWAS? In addition to considerations that are common to both GWASs and EWASs (for example, adequate technology and sample size), the design of EWASs has specific considerations with respect to sample selection. DNAm patterns are specific to tissues and developmental stages, and they also change over time. Furthermore, EWAS associations can be causal as well as consequential for the phenotype in question — a difference from GWASs that presents considerable challenges. Here, we discuss these considerations in the context of designing and analysing an effective EWAS, keeping in mind that EWASs are likely to evolve, much like GWASs did, as information and experience accumulate.

Rakyan et al. 2011, Nature Reviews Genetics

Epigenetics and cancer

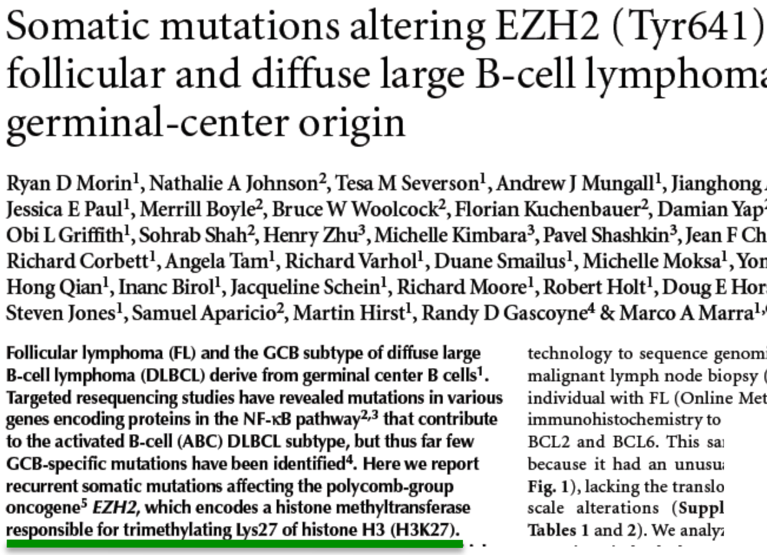
Most is known about DNA methylation. Cancers typically exhibit (of varying degrees associated with severity):

- Global DNA hypomethylation
- Region-specific hypermethylation, typically at CpG-island-associated promoters

Recent evidence highlights interruptions of epigenetic machinery from genetic mutations in cancer

Cancer Genome Sequencing -> Epigenetics

Morin et al. (2010) Nature Genetics.

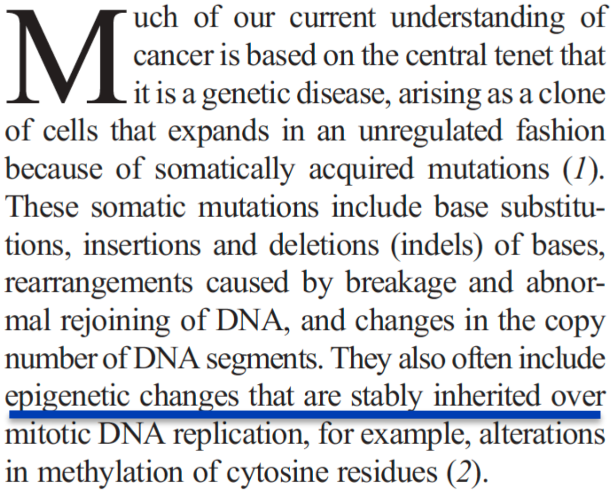


Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin

Ryan D Morin[1], Nathalie A Johnson[2], Tesa M Severson[1], Andrew J Mungall[1], Jianghong An[1], Rc
Jessica E Paul[1], Merrill Boyle[2], Bruce W Woolcock[2], Florian Kuchenbauer[2], Damian Yap[2], R Kei
Obi L Griffith[1], Sohrab Shah[2], Henry Zhu[3], Michelle Kimbara[3], Pavel Shashkin[3], Jean F Charlot[3],
Richard Corbett[1], Angela Tam[1], Richard Varhol[1], Duane Smailus[1], Michelle Moksa[1], Yongjun Z
Hong Qian[1], Inanc Birol[1], Jacqueline Schein[1], Richard Moore[1], Robert Holt[1], Doug E Horsman[4],
Steven Jones[1], Samuel Aparicio[2], Martin Hirst[1], Randy D Gascoyne[4] & Marco A Marra[1,6]

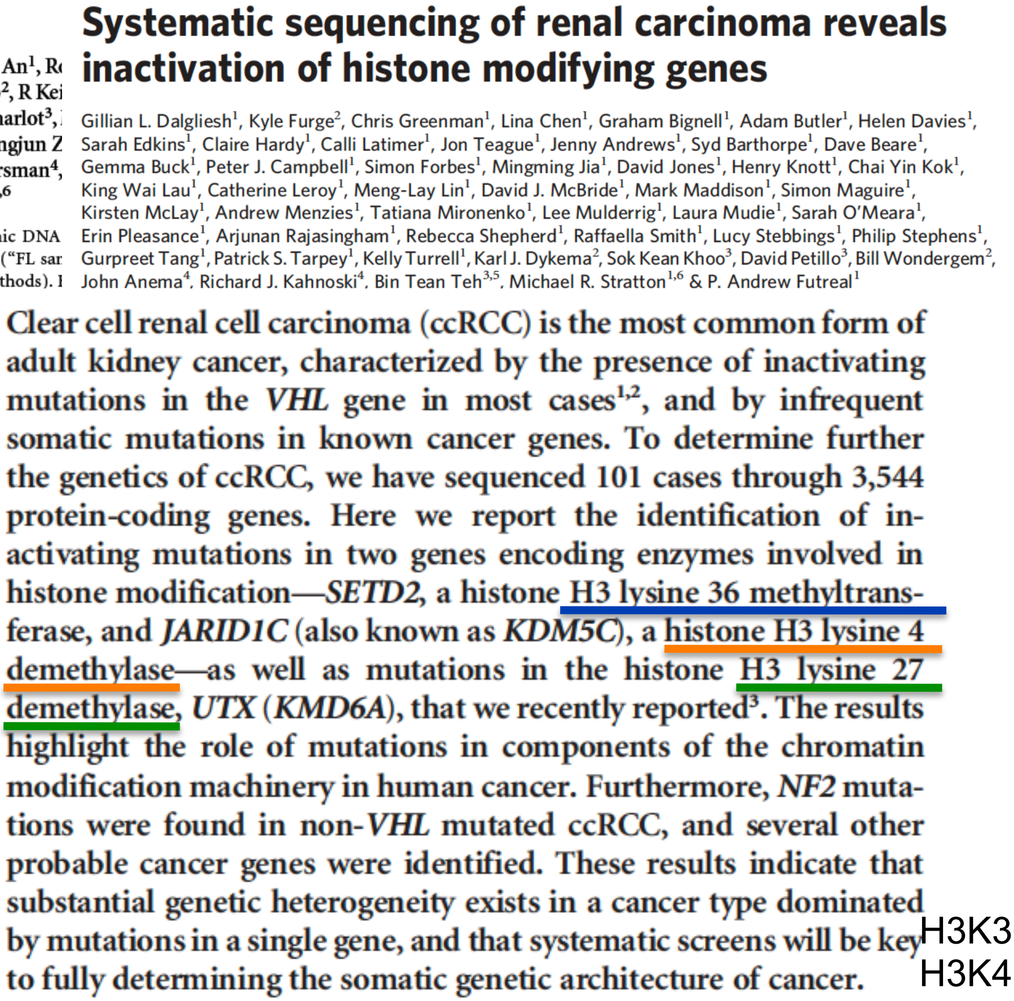
Follicular lymphoma (FL) and the GCB subtype of diffuse large B-cell lymphoma (DLBCL) derive from germinal center B cells[1]. Targeted resequencing studies have revealed mutations in various genes encoding proteins in the NF-κB pathway[2,3] that contribute to the activated B-cell (ABC) DLBCL subtype, but thus far few GCB-specific mutations have been identified[4]. Here we report recurrent somatic mutations affecting the polycomb-group oncogene[5] *EZH2*, which encodes a histone methyltransferase responsible for trimethylating Lys27 of histone H3 (H3K27).

technology to sequence genomic DNA
malignant lymph node biopsy ("FL sar
individual with FL (Online Methods). I
immunohistochemistry to
BCL2 and BCL6. This sa
because it had an unusu
Fig. 1), lacking the translo
scale alterations (Suppl
Tables 1 and 2). We analyz



Much of our current understanding of cancer is based on the central tenet that it is a genetic disease, arising as a clone of cells that expands in an unregulated fashion because of somatically acquired mutations (*1*). These somatic mutations include base substitutions, insertions and deletions (indels) of bases, rearrangements caused by breakage and abnormal rejoining of DNA, and changes in the copy number of DNA segments. They also often include epigenetic changes that are stably inherited over mitotic DNA replication, for example, alterations in methylation of cytosine residues (*2*).

Stratton (2011) Science.



Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes

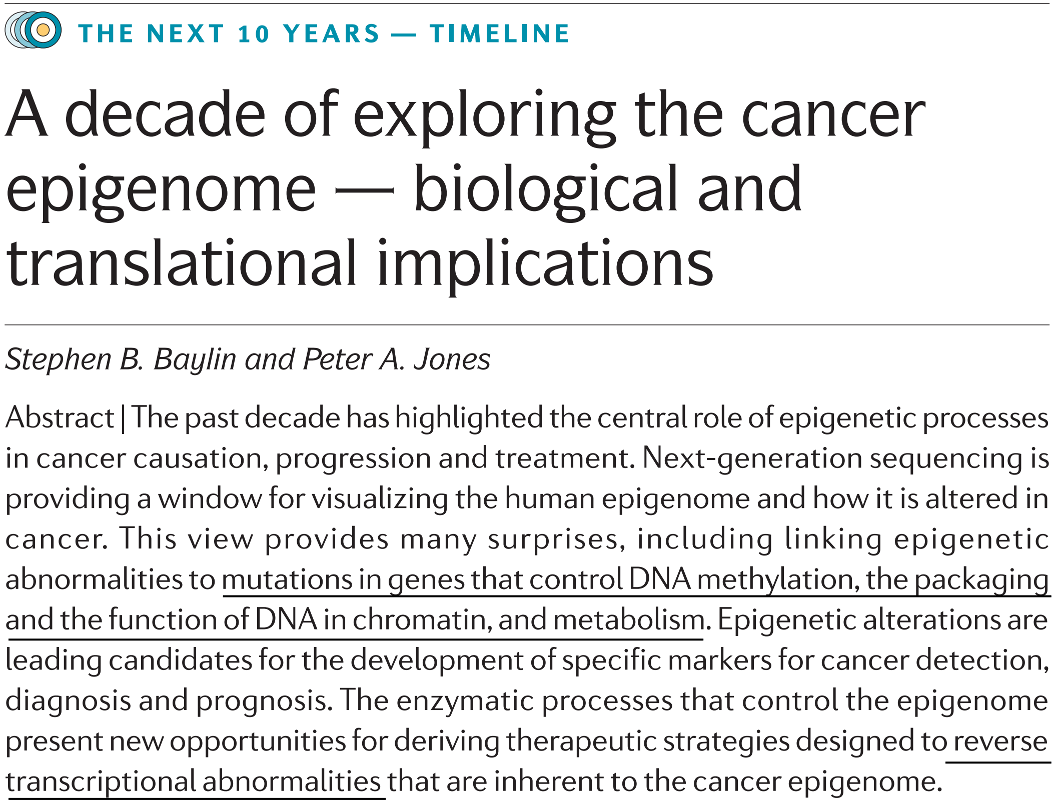
Gillian L. Dalgliesh[1], Kyle Furge[2], Chris Greenman[1], Lina Chen[1], Graham Bignell[1], Adam Butler[1], Helen Davies[1], Sarah Edkins[1], Claire Hardy[1], Calli Latimer[1], Jon Teague[1], Jenny Andrews[1], Syd Barthorpe[1], Dave Beare[1], Gemma Buck[1], Peter J. Campbell[1], Simon Forbes[1], Mingming Jia[1], David Jones[1], Henry Knott[1], Chai Yin Kok[1], King Wai Lau[1], Catherine Leroy[1], Meng-Lay Lin[1], David J. McBride[1], Mark Maddison[1], Simon Maguire[1], Kirsten McLay[1], Andrew Menzies[1], Tatiana Mironenko[1], Lee Mulderrig[1], Laura Mudie[1], Sarah O'Meara[1], Erin Pleasance[1], Arjunan Rajasingham[1], Rebecca Shepherd[1], Raffaella Smith[1], Lucy Stebbings[1], Philip Stephens[1], Gurpreet Tang[1], Patrick S. Tarpey[1], Kelly Turrell[1], Karl J. Dykema[2], Sok Kean Khoo[3], David Petillo[3], Bill Wondergem[2], John Anema[4], Richard J. Kahnoski[4], Bin Tean Teh[3,5], Michael R. Stratton[1,6] & P. Andrew Futreal[1]

Clear cell renal cell carcinoma (ccRCC) is the most common form of adult kidney cancer, characterized by the presence of inactivating mutations in the *VHL* gene in most cases[1,2], and by infrequent somatic mutations in known cancer genes. To determine further the genetics of ccRCC, we have sequenced 101 cases through 3,544 protein-coding genes. Here we report the identification of inactivating mutations in two genes encoding enzymes involved in histone modification—*SETD2*, a histone H3 lysine 36 methyltransferase, and *JARID1C* (also known as *KDM5C*), a histone H3 lysine 4 demethylase—as well as mutations in the histone H3 lysine 27 demethylase, *UTX* (*KMD6A*), that we recently reported[3]. The results highlight the role of mutations in components of the chromatin modification machinery in human cancer. Furthermore, *NF2* mutations were found in non-*VHL* mutated ccRCC, and several other probable cancer genes were identified. These results indicate that substantial genetic heterogeneity exists in a cancer type dominated by mutations in a single gene, and that systematic screens will be key to fully determining the somatic genetic architecture of cancer.

Dalgliesh et al. (2010) Nature.

H3K36me3,
H3K4me3,
H3K27me3

Epigenetic drugs



THE NEXT 10 YEARS — TIMELINE

A decade of exploring the cancer epigenome — biological and translational implications

Stephen B. Baylin and Peter A. Jones

Abstract | The past decade has highlighted the central role of epigenetic processes in cancer causation, progression and treatment. Next-generation sequencing is providing a window for visualizing the human epigenome and how it is altered in cancer. This view provides many surprises, including linking epigenetic abnormalities to mutations in genes that control DNA methylation, the packaging and the function of DNA in chromatin, and metabolism. Epigenetic alterations are leading candidates for the development of specific markers for cancer detection, diagnosis and prognosis. The enzymatic processes that control the epigenome present new opportunities for deriving therapeutic strategies designed to reverse transcriptional abnormalities that are inherent to the cancer epigenome.

Translational advances:

Biomarkers (e.g. GSTP1 in prostate cancer)

Therapeutics (e.g. azacitidine and decitabine have FDA approval for myelodisplastic syndrome, which can lead to leukemia)

FDA approval of vorinostat and romidepsin for cutaneous
T cell lymphoma

HDAC inhibitors in clinical trials.

….

DNA methylation

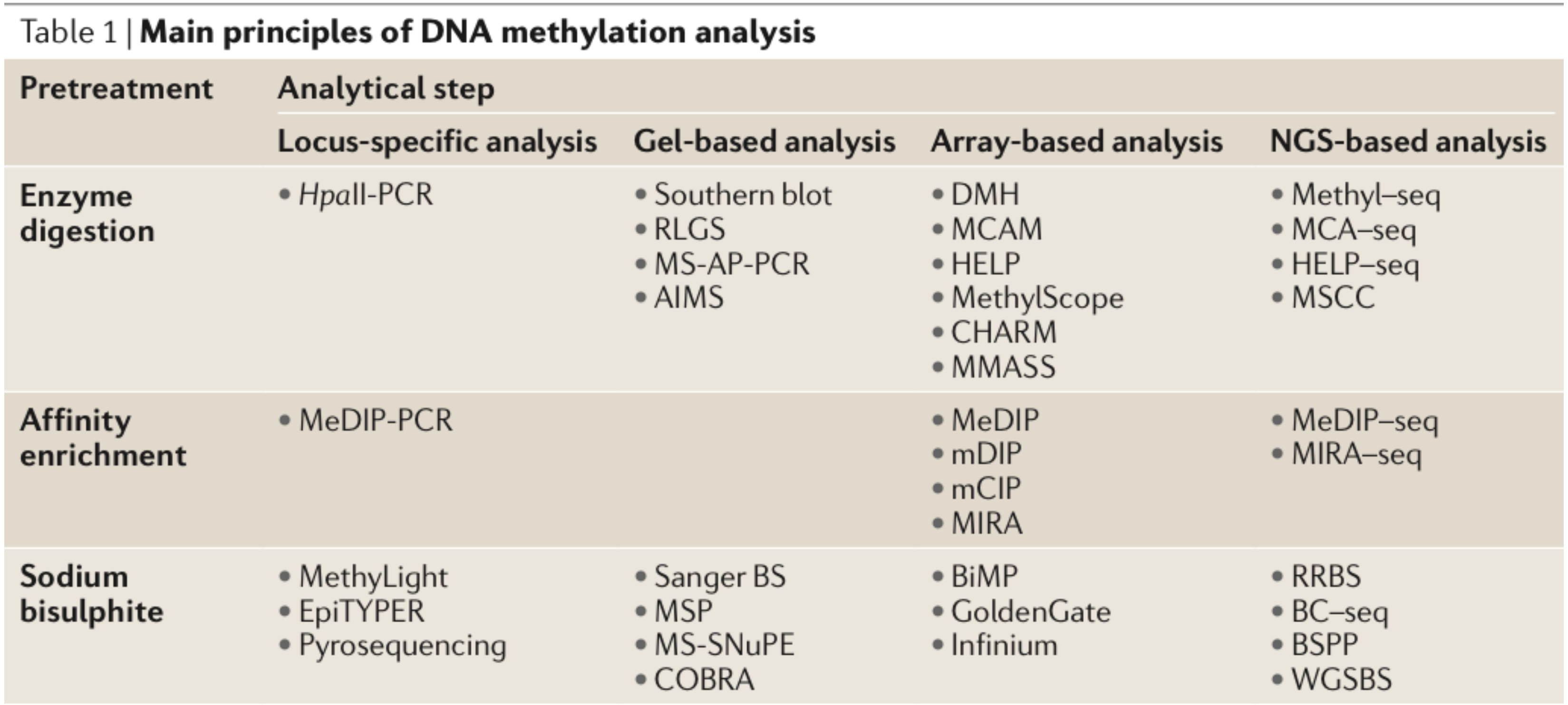


Table 1 | **Main principles of DNA methylation analysis**

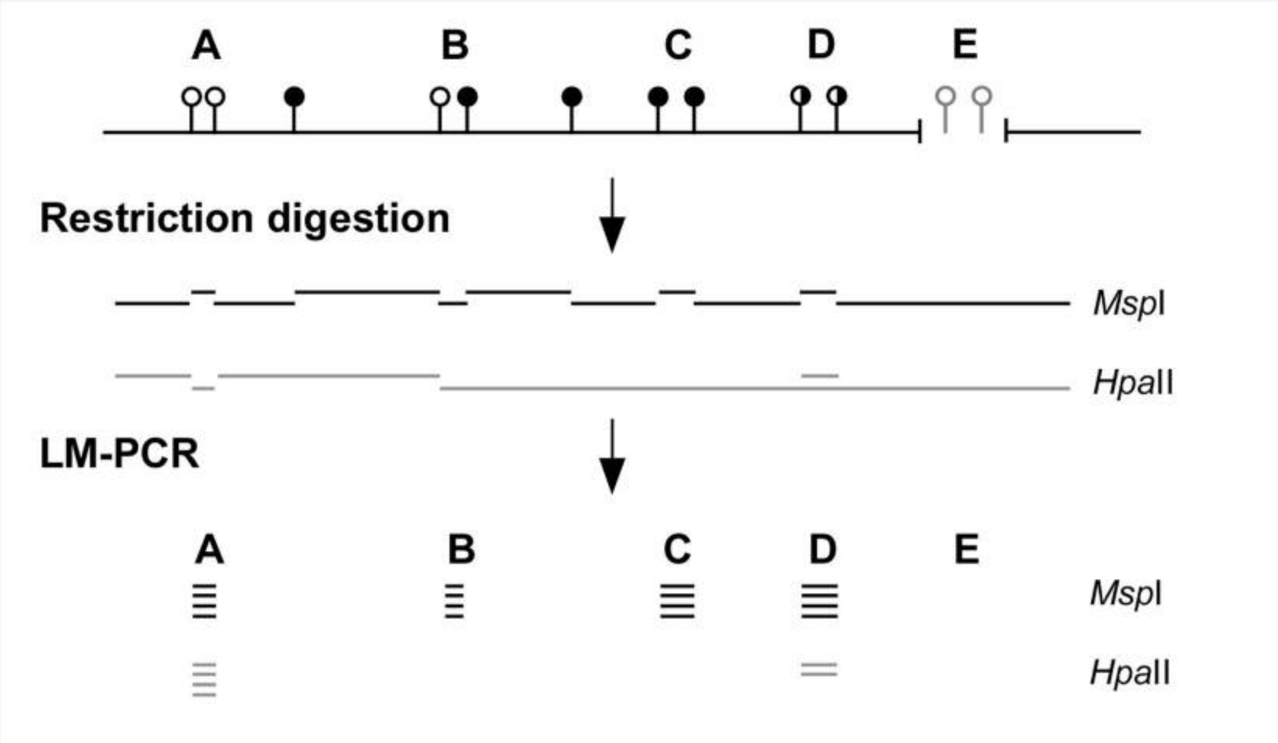
Pretreatment	Analytical step			
	Locus-specific analysis	Gel-based analysis	Array-based analysis	NGS-based analysis
Enzyme digestion	• *Hpa*II-PCR	• Southern blot • RLGS • MS-AP-PCR • AIMS	• DMH • MCAM • HELP • MethylScope • CHARM • MMASS	• Methyl–seq • MCA–seq • HELP–seq • MSCC
Affinity enrichment	• MeDIP-PCR		• MeDIP • mDIP • mCIP • MIRA	• MeDIP–seq • MIRA–seq
Sodium bisulphite	• MethyLight • EpiTYPER • Pyrosequencing	• Sanger BS • MSP • MS-SNuPE • COBRA	• BiMP • GoldenGate • Infinium	• RRBS • BC–seq • BSPP • WGSBS

Direct sequencing

Oxford Nanopore
Pacific Biosciences
etc.

Laird, Nature Reviews Genetics, March 2010

Enzyme digestion example

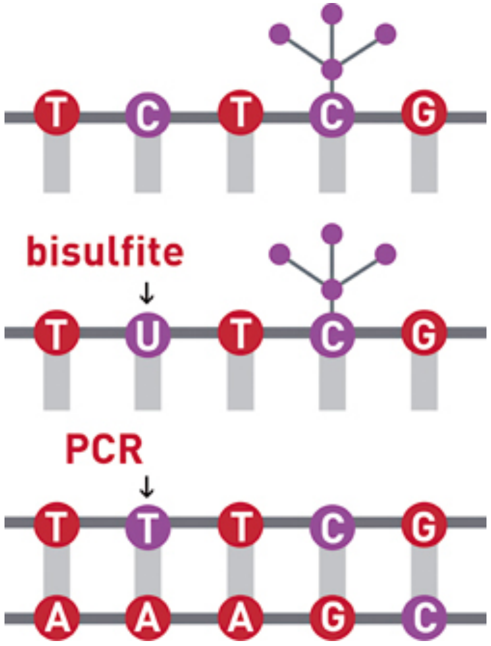


MspI – cuts at CCGG or CCGG sites

HpaII – cuts only at CCGG

CG - unmethylated
CG - methylated

Bisulphite sequencing



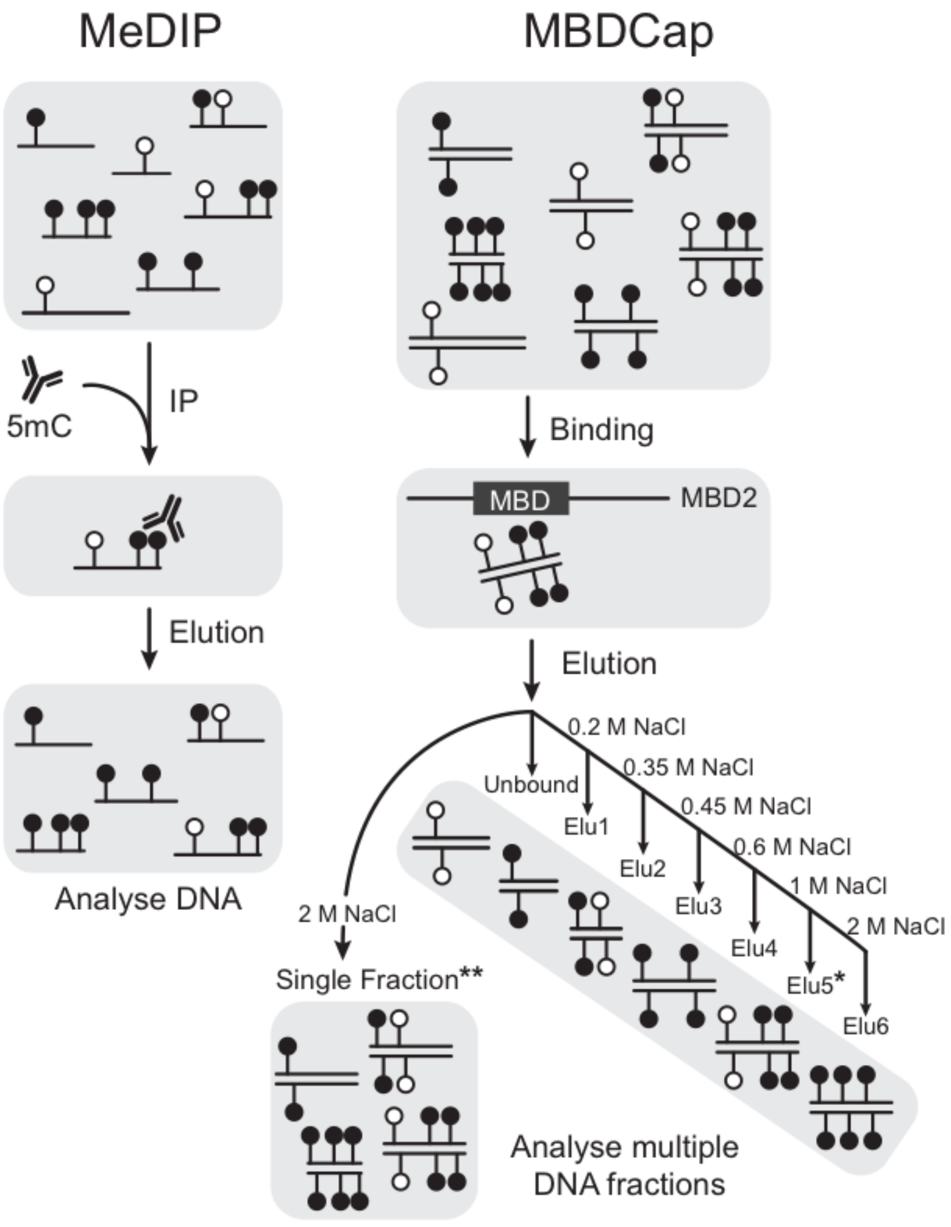
Sodium bisulphite converts methylated **C**ytosine into **U**racil, which can be read as **T**hymine after PCR

In combination with sequencing (Sanger or NGS), can achieve methylation mapping at single base resolution

Can be nicely combined with genotyping arrays (e.g. Illumina HumanMethylation 450k)

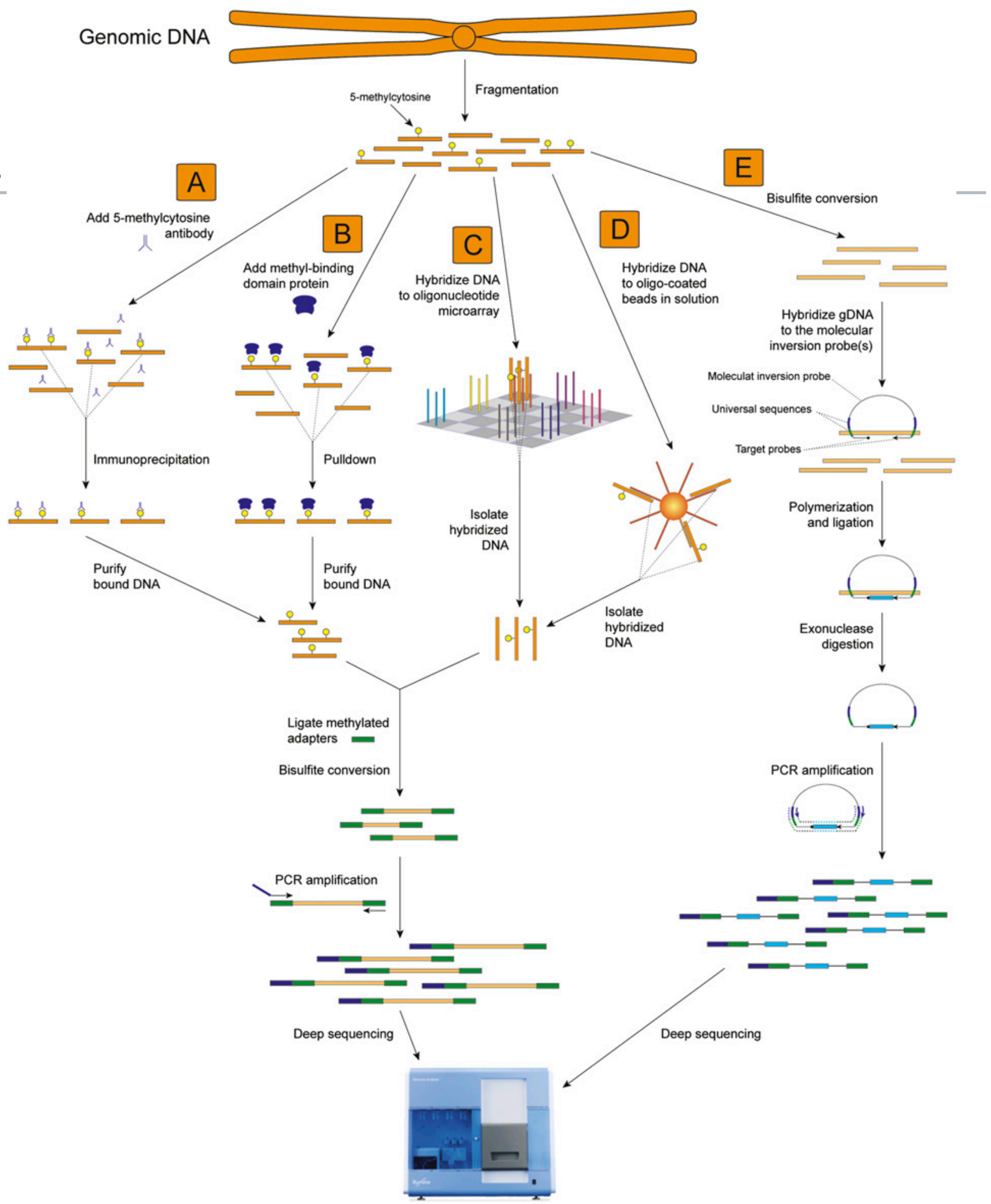
http://www.diagenode.com/en/applications/bisulfite-conversion.php

Affinity capture of methylated DNA



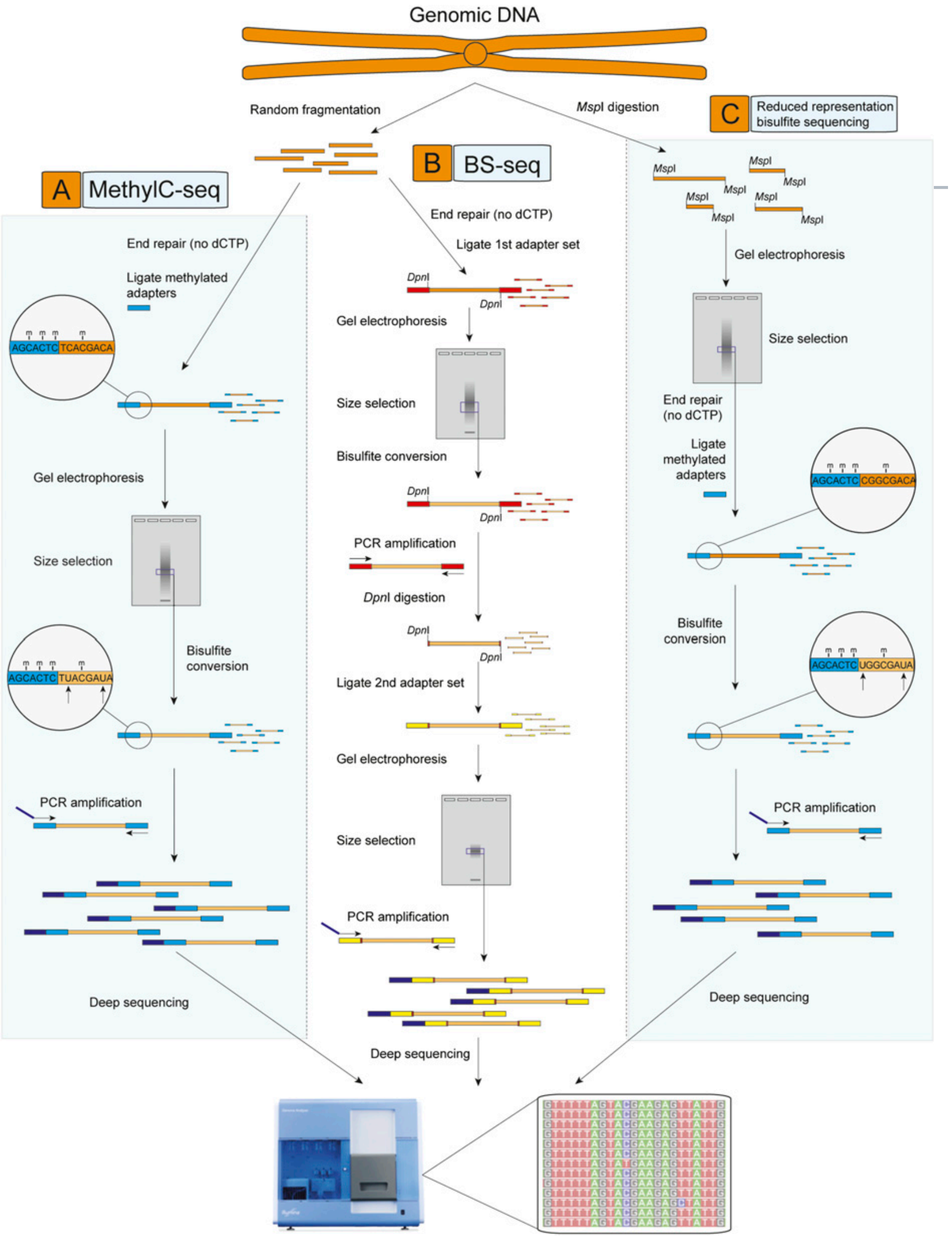
Robinson et al. 2010

Methods for DNA methylation that use “capture” with NGS

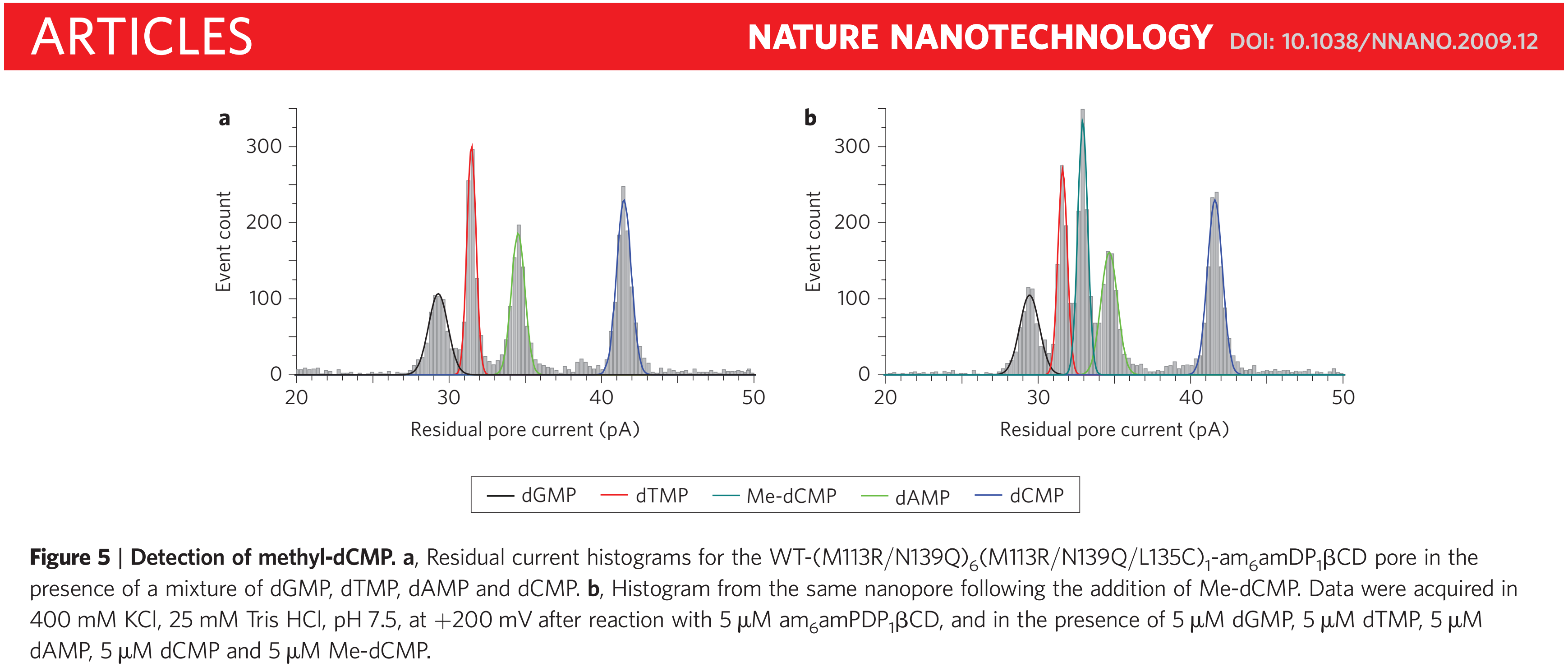


Lister and Ecker, Genome Research (review) 2009

DNAme methods that use bisulphite conversion with NGS



DNA methylation by direct sequencing (Oxford Nanopore)



ARTICLES

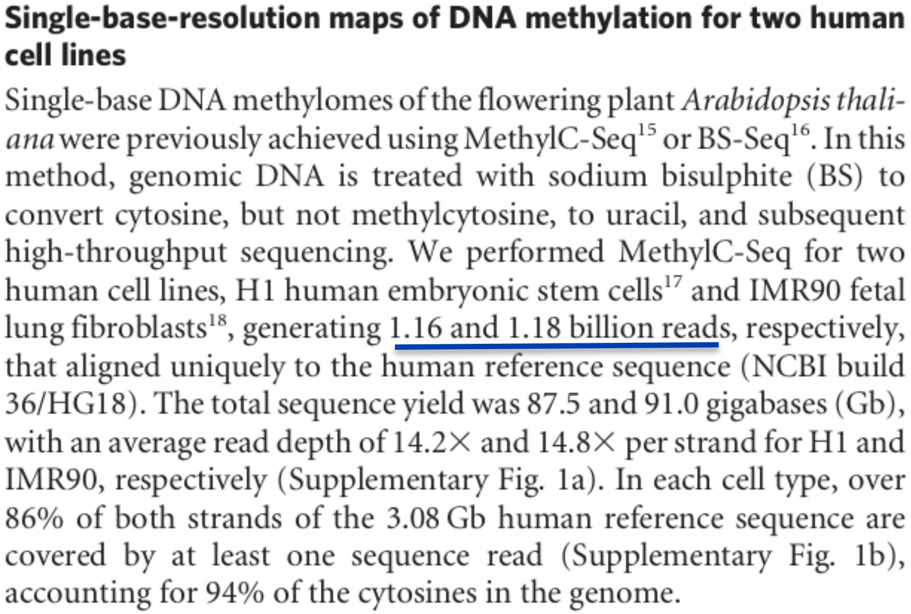
NATURE NANOTECHNOLOGY DOI: 10.1038/NNANO.2009.12

Figure 5 | Detection of methyl-dCMP. a, Residual current histograms for the WT-(M113R/N139Q)$_6$(M113R/N139Q/L135C)$_1$-am$_6$amDP$_1$βCD pore in the presence of a mixture of dGMP, dTMP, dAMP and dCMP. **b**, Histogram from the same nanopore following the addition of Me-dCMP. Data were acquired in 400 mM KCl, 25 mM Tris HCl, pH 7.5, at +200 mV after reaction with 5 μM am$_6$amPDP$_1$βCD, and in the presence of 5 μM dGMP, 5 μM dTMP, 5 μM dAMP, 5 μM dCMP and 5 μM Me-dCMP.

Other remarks into DNA methylation data

- Whole genome bisulphite sequencing is the most accurate, but expensive and somewhat inefficient
- Performance of affinity capture can vary drastically according to exact specifications of the protocol
- Difficult to compare methods since platforms have different coverage, different resolution

Whole genome BS sequencing can be inefficient



Single-base-resolution maps of DNA methylation for two human cell lines

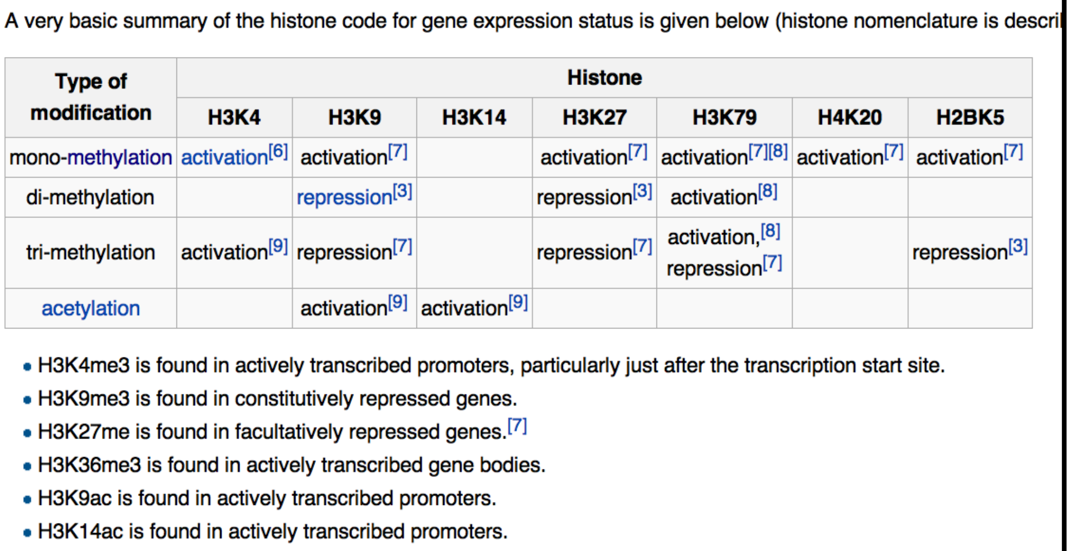
Single-base DNA methylomes of the flowering plant *Arabidopsis thaliana* were previously achieved using MethylC-Seq[15] or BS-Seq[16]. In this method, genomic DNA is treated with sodium bisulphite (BS) to convert cytosine, but not methylcytosine, to uracil, and subsequent high-throughput sequencing. We performed MethylC-Seq for two human cell lines, H1 human embryonic stem cells[17] and IMR90 fetal lung fibroblasts[18], generating 1.16 and 1.18 billion reads, respectively, that aligned uniquely to the human reference sequence (NCBI build 36/HG18). The total sequence yield was 87.5 and 91.0 gigabases (Gb), with an average read depth of 14.2× and 14.8× per strand for H1 and IMR90, respectively (Supplementary Fig. 1a). In each cell type, over 86% of both strands of the 3.08 Gb human reference sequence are covered by at least one sequence read (Supplementary Fig. 1b), accounting for 94% of the cytosines in the genome.

Lister et al. 2009, Nature

Notes re: WGSBS:

1. Mapping is done on BS-converted reads/genome (i.e.3 bases), requires mapping separately to each strand – need longer (paired) reads and high coverage
2. Of the 1.18B reads, approximately 670M (56%) do NOT overlap a CpG site
3. There may be a fair amount of regions that are completely unmethylated

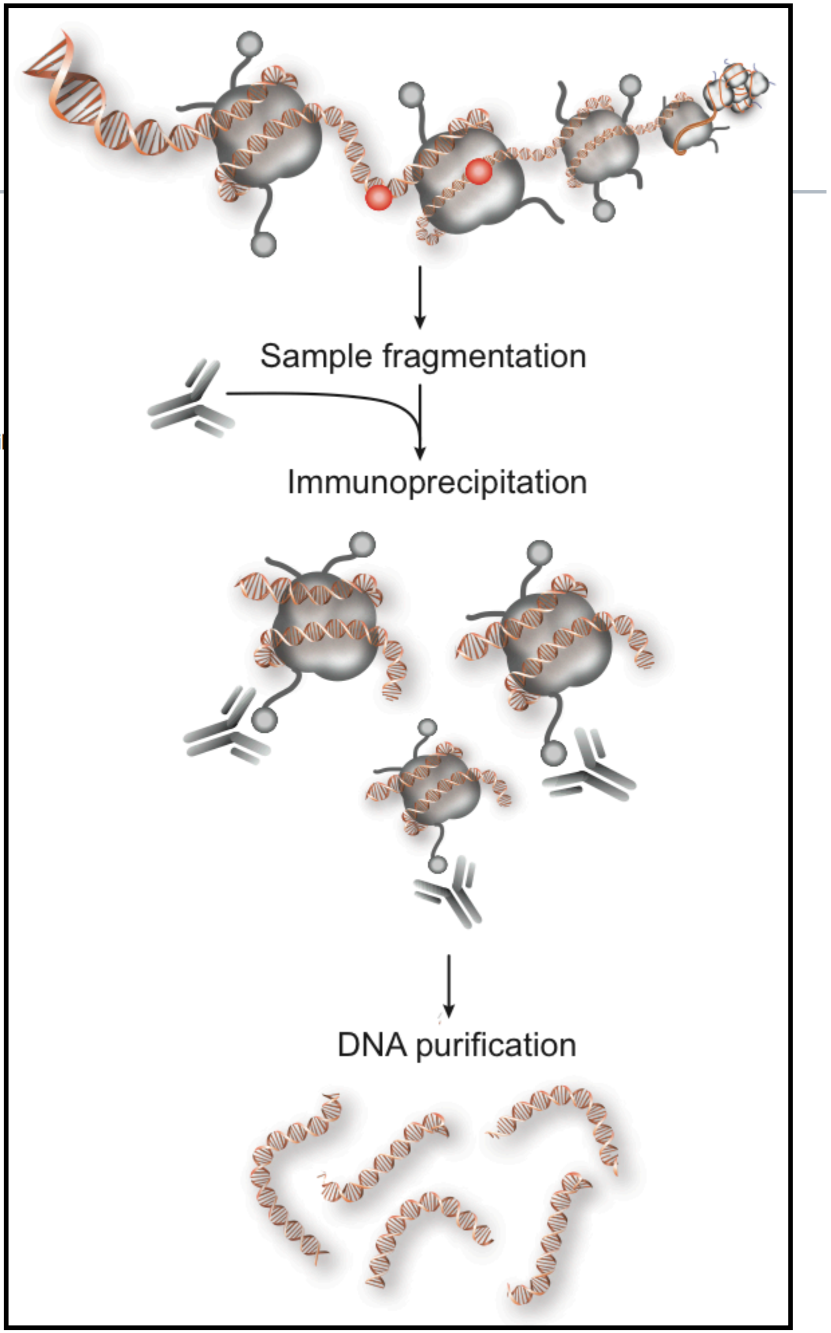
Chromatin immunoprecipitation for protein-DNA interactions



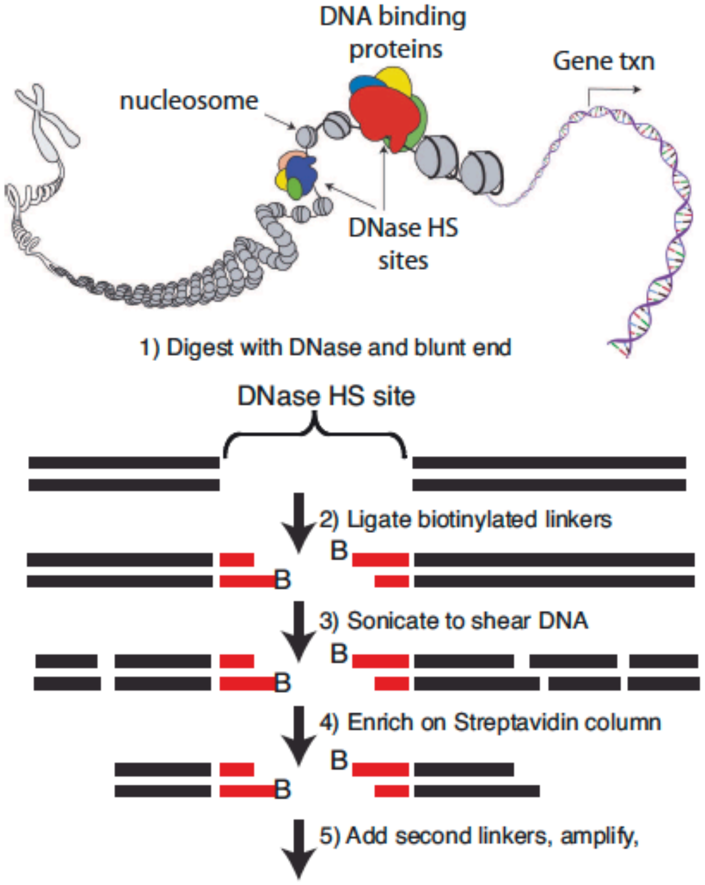
A very basic summary of the histone code for gene expression status is given below (histone nomenclature is descri

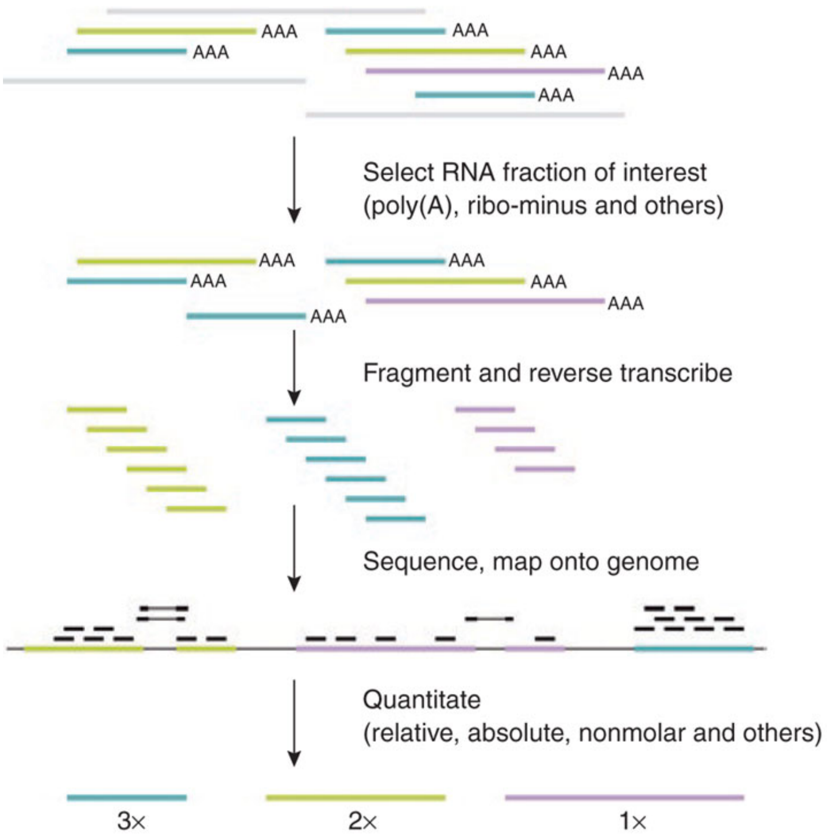
Type of modification	Histone						
	H3K4	H3K9	H3K14	H3K27	H3K79	H4K20	H2BK5
mono-methylation	activation[6]	activation[7]		activation[7]	activation[7][8]	activation[7]	activation[7]
di-methylation		repression[3]		repression[3]	activation[8]		
tri-methylation	activation[9]	repression[7]		repression[7]	activation,[8] repression[7]		repression[3]
acetylation		activation[9]	activation[9]				

- H3K4me3 is found in actively transcribed promoters, particularly just after the transcription start site.
- H3K9me3 is found in constitutively repressed genes.
- H3K27me is found in facultatively repressed genes.[7]
- H3K36me3 is found in actively transcribed gene bodies.
- H3K9ac is found in actively transcribed promoters.
- H3K14ac is found in actively transcribed promoters.

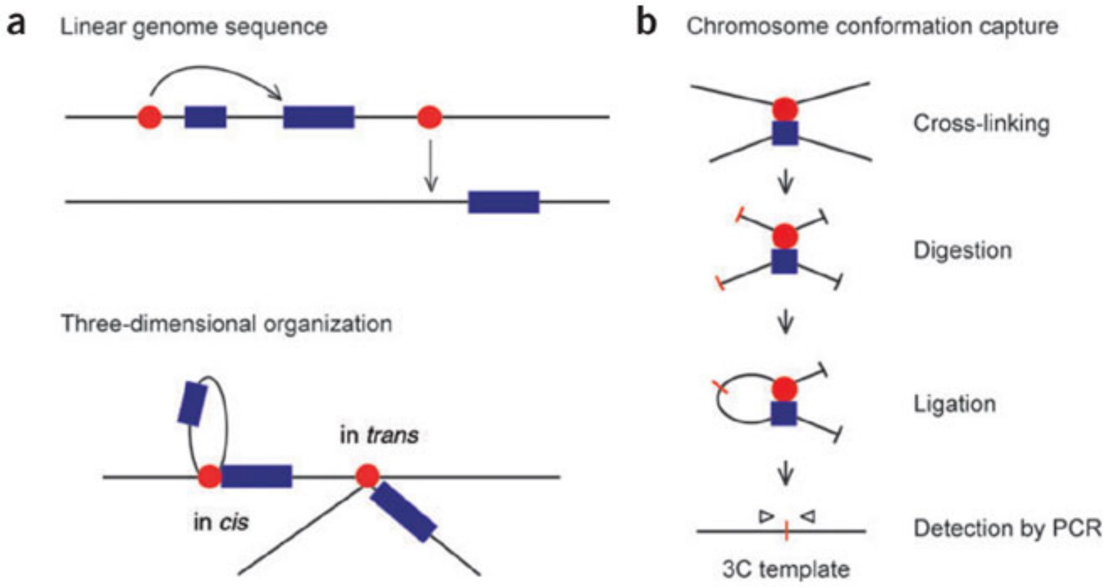


Techniques: DNaseI, RNA-seq


DNA binding proteins
Gene txn
nucleosome
DNase HS sites
1) Digest with DNase and blunt end
DNase HS site
2) Ligate biotinylated linkers
B
B
3) Sonicate to shear DNA
B
B
4) Enrich on Streptavidin column
B
B
5) Add second linkers, amplify,


AAA
AAA
AAA
AAA
AAA
AAA
Select RNA fraction of interest (poly(A), ribo-minus and others)
AAA
AAA
AAA
AAA
AAA
AAA
Fragment and reverse transcribe
Sequence, map onto genome
Quantitate (relative, absolute, nonmolar and others)
3×
2×
1×

Higher-order chromatin structure

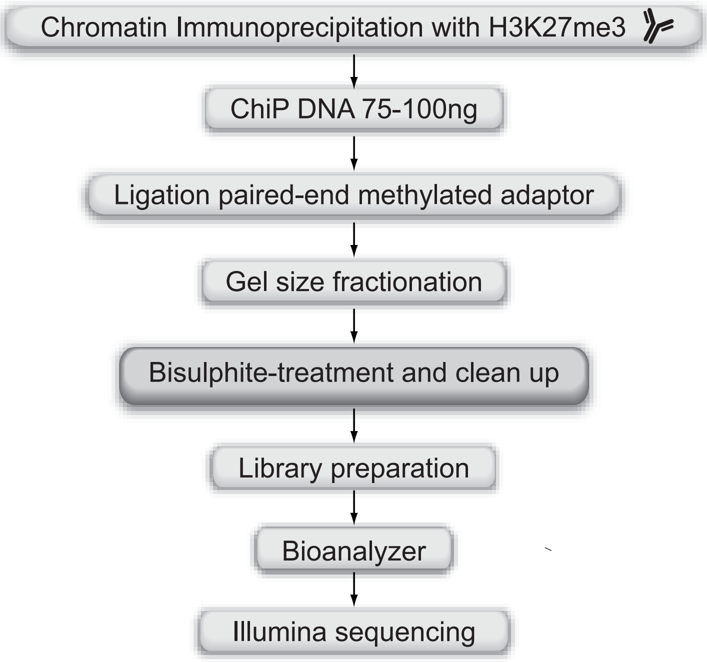


Assaying combinations of epigenetic factors

- Chromatin immunoprecipitation + bisulphite treatment == ChIP-BS-seq
- Nucleosome Occupancy + Methylation == NOME-seq
- Variations on RNA

ChIP-BS-seq

A few tricks on the technical side to facilitate this.



Chromatin Immunoprecipitation with H3K27me3

ChiP DNA 75-100ng

Ligation paired-end methylated adaptor

Gel size fractionation

Bisulphite-treatment and clean up

Library preparation

Bioanalyzer

Illumina sequencing

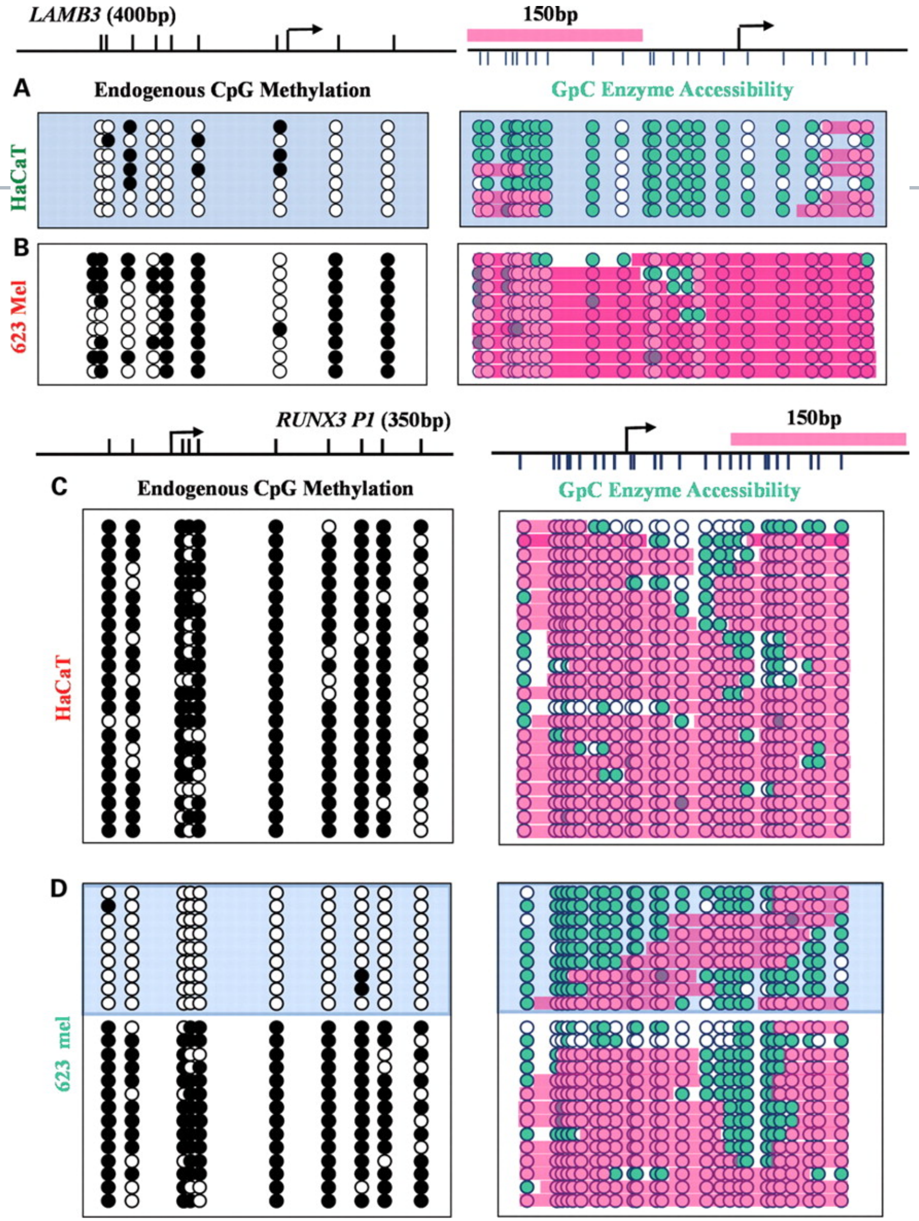
Statham*, Robinson* et al. (2012), Genome Research

NOME-seq

M.CviPI enzyme is used to methylate GpC sites **not bound by nucleosomes**

Both GpC methylation and CpG methylation can be readout (on the same clone) after bisulphite treatment

Pink: nucleosome-bound (not methylated by M.CviPI)
Green: accessible



Remarks: Allele-specific epigenetics, cell populations

- A couple key points to recognize:
  - Typically, MBD-seq/ChIP-seq/etc. are analyzing populations of cells (e.g. patient tumours that may contain normal cell types as well) – so we are really studying the population average!
  - In some instances, we may be able to combine the information we get from genome sequencing (e.g. SNPs) to partition transcription and epigenetic factors by allele

Technical limitation in the amount of DNA need to create library and sequence

- We often want to know about several factors on a single population of cells – requires a lot of DNA/RNA
- New technologies are trying to address this
- Patient (e.g. tumour sample) cell population purity?

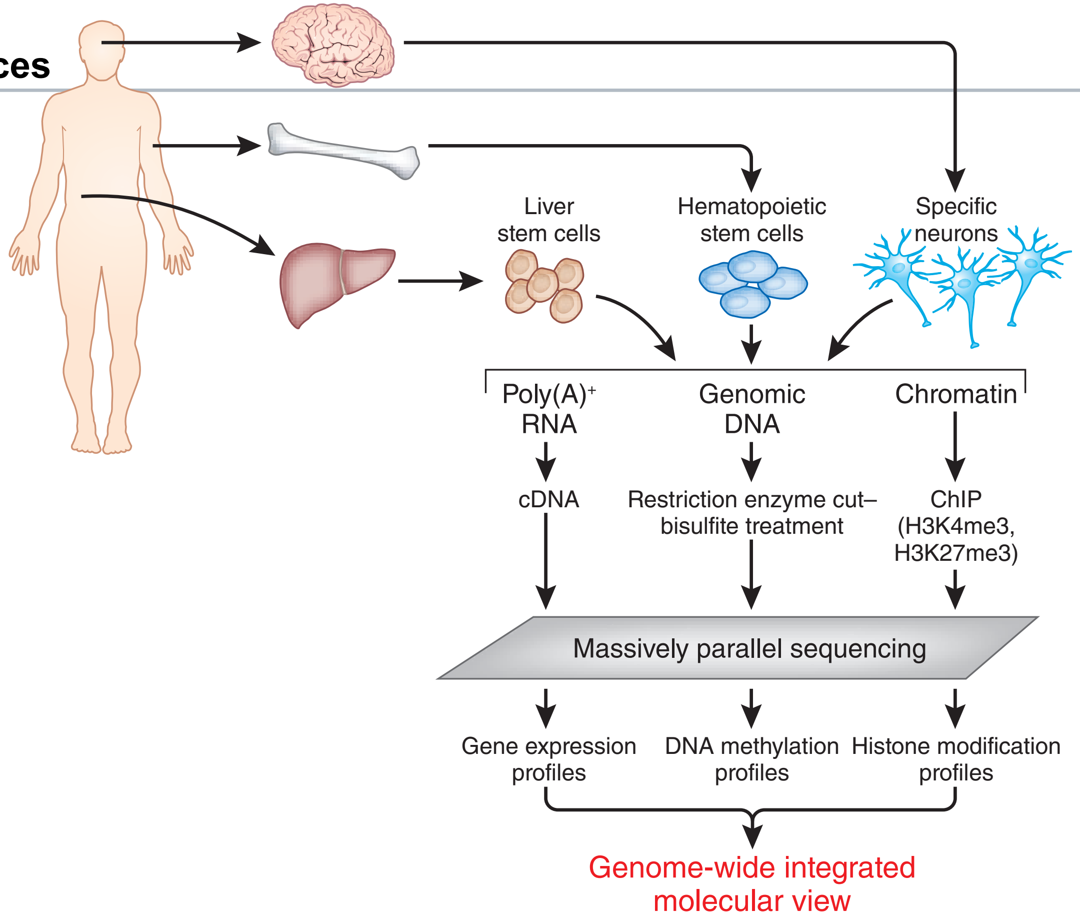
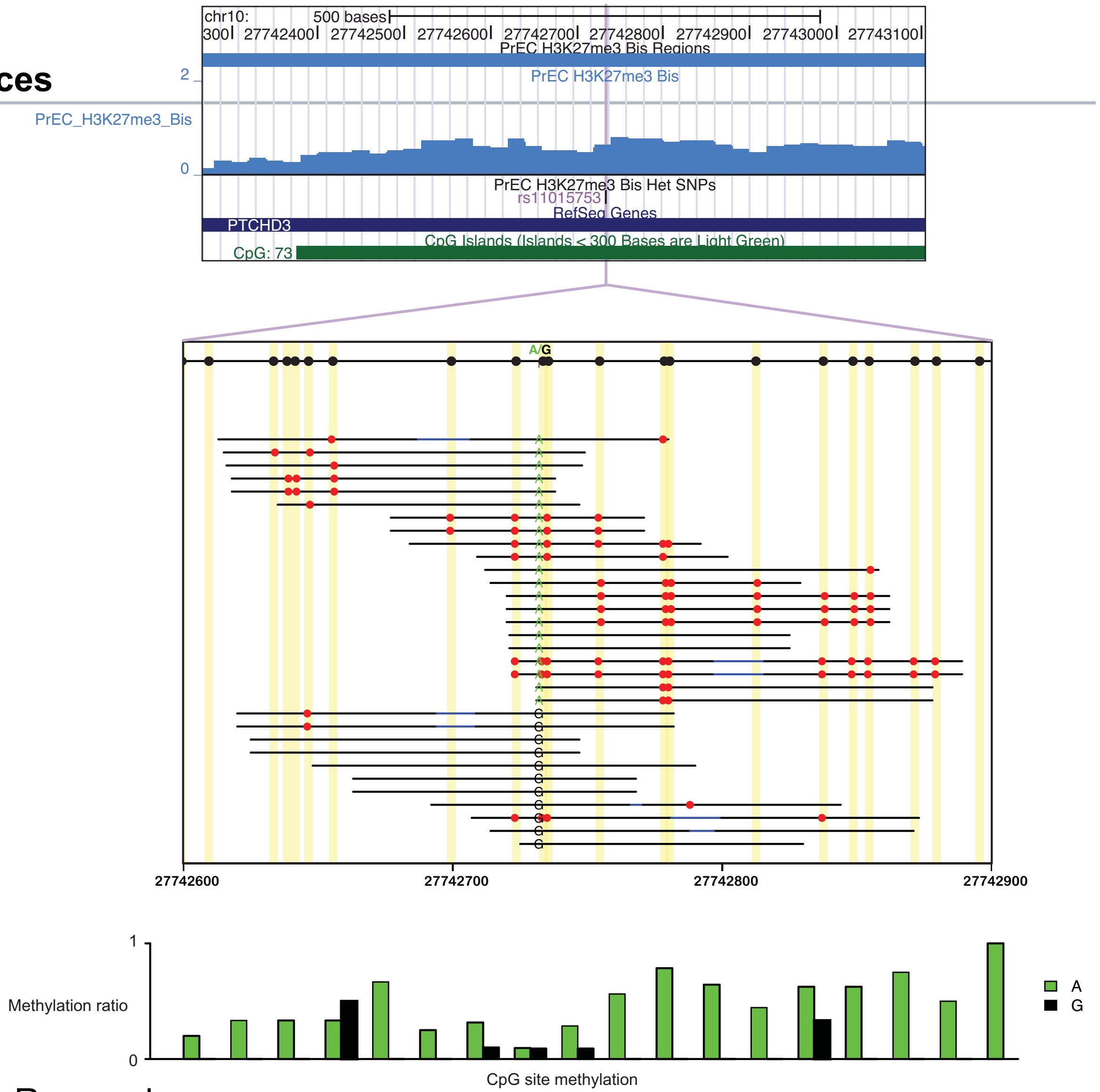


Figure 1 | Schematic flow chart of experimental design. Rare cell types are isolated from specific organs and used for RNA and DNA preparation, and ChIP. Combining gene expression, DNA methylation and histone modification profiles gives an integrated view of the epigenome.

Allele-specific methylation

- Biologically, what affect does this have?
- How prominent is this?



Statham*, Robinson* et al. (2012), Genome Research

Summary

Many approaches for DNA methylation

Chromatin immunoprecipitations for protein-DNA

Higher order structures