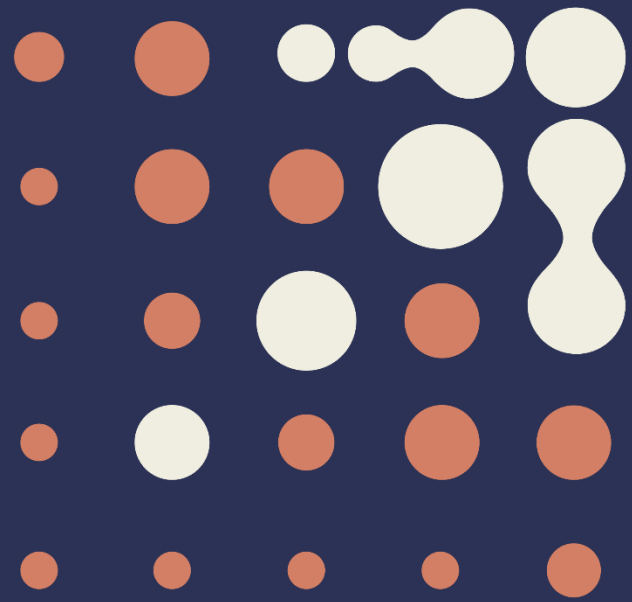
Medical Genomics
Lecture 4: Epigenomics

Alexandra Sexton-Oates, PhD
SextonoatesA@iarc.who.int

23/11/2022

International Agency
for Research on Cancer

World Health
Organization



Outline

Part 1: Introduction to epigenomics

- Concepts: role of epigenetics, refresh gene expression, epigenetic mechanisms
- Techniques: detecting and measuring epigenetic modifications
- Resources: databases of regulatory annotations and epigenetic profiles

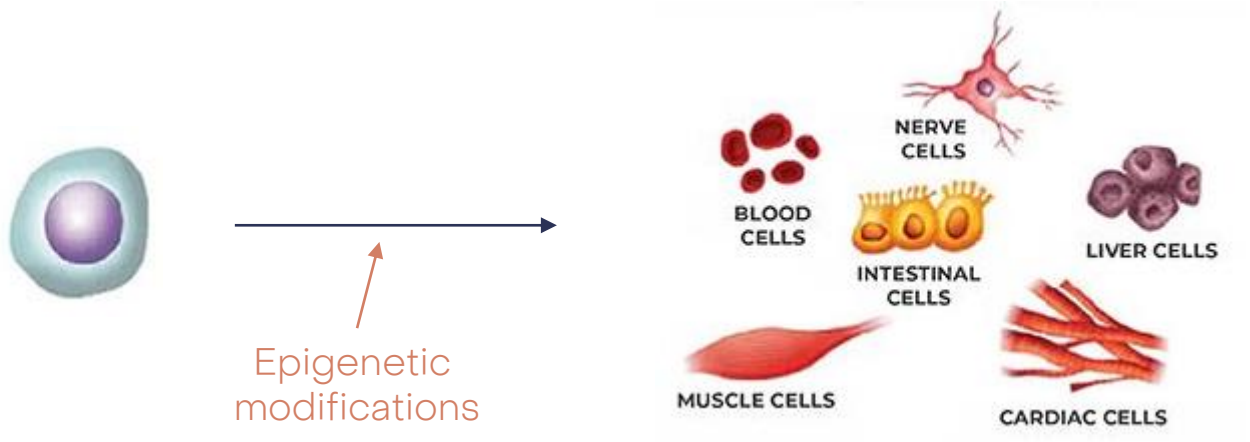
Part 2: Cancer epigenetics

- Concepts: introduction to epigenetics in cancer

Part 3: Application of DNA methylation arrays in cancer

- Methodology: processing DNA methylation arrays, normalisation methods
- Research applications: principal components analysis, global methylation, deconvolution
- Clinical application: DNA methylation arrays for diagnosis and treatment

Part 1: What is epigenetics?



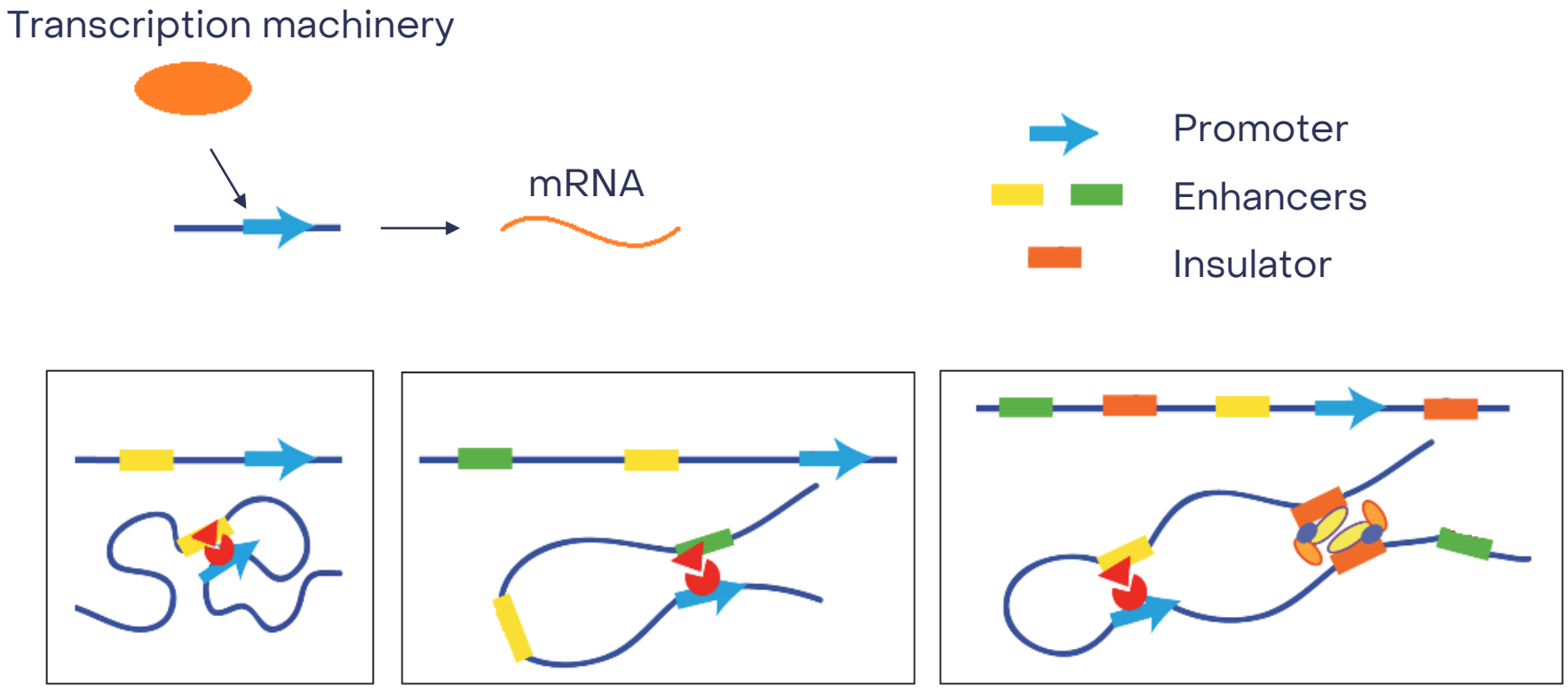
Epigenetics is the study of modifications made to DNA and associated factors that:

- Do not change the DNA sequence itself
- Are maintained during cell division
- Cause stable changes in gene expression

Mechanisms of gene expression: regulatory elements

Promoters
Contain the TSS, RNA polymerase binding site and other transcription factor binding sites

Enhancers
Bind transcription factors to increase the likelihood of, or boost, transcription



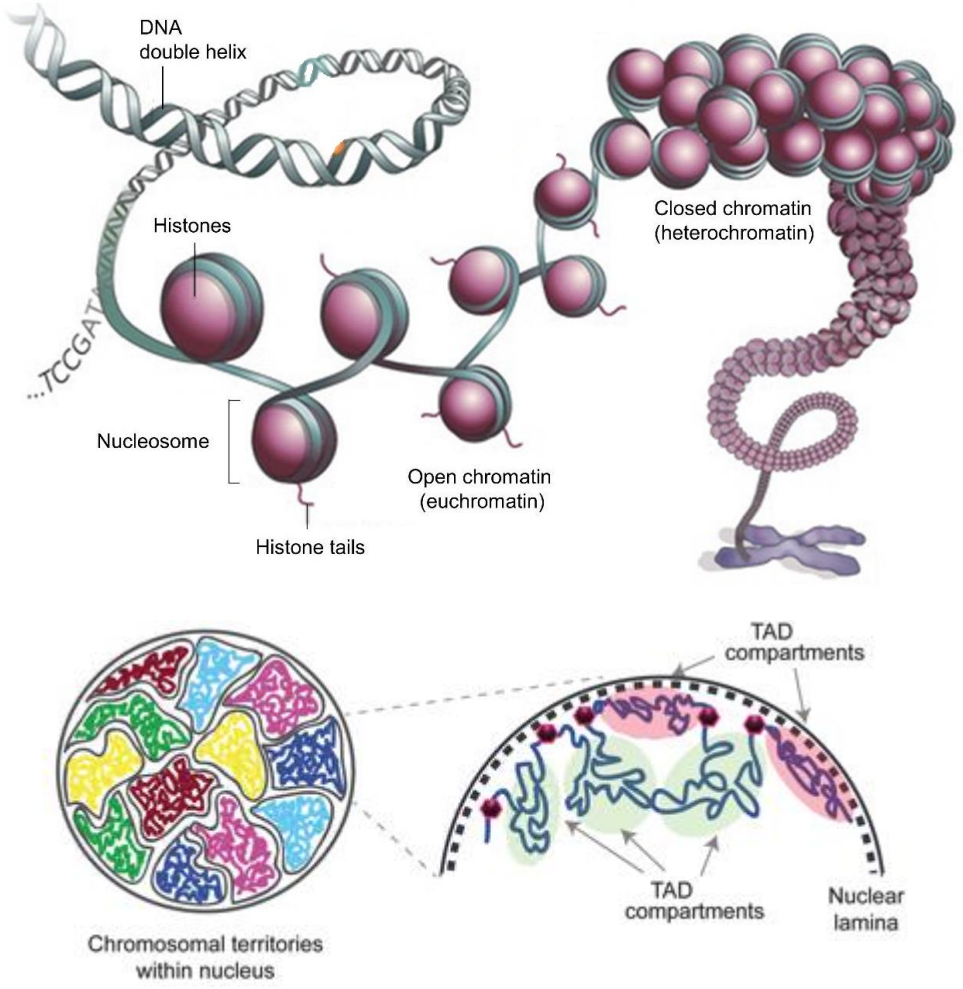
Adapted from: Matharu N & Ahituv N. *PLoS Genetics* 2015; 11(12): e1005640

Mechanisms of gene expression: chromatin structure

Chromatin folding is required to squeeze 2m of DNA into the cell nucleus

Primary structure:
- Double helix winds around the histone octamers to form the nucleosome
- A nucleosome contains 8 histone proteins (H2A, H2B, H3 and H4)
- Euchromatin contains space for transcription machinery

Higher-order structure
- The way in which chromatin loops and folds over itself into topological associated domains and chromosomal territories within the nucleus



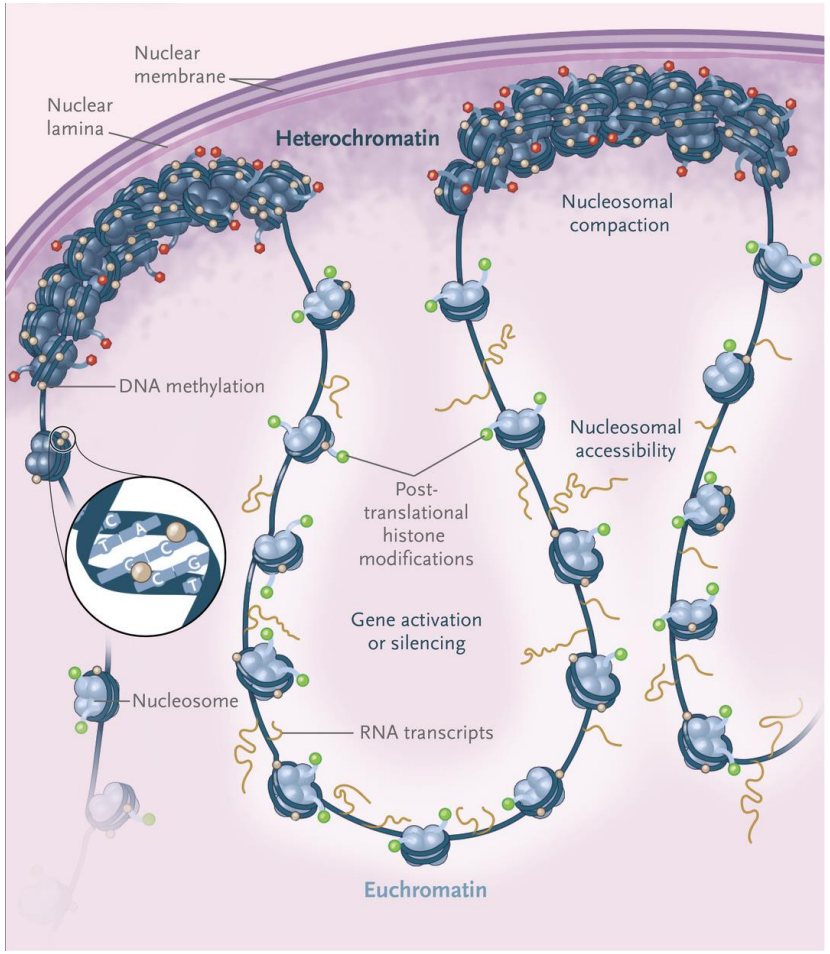
Adapted from: Marx V. *Nature* 2012; 491(17422): 143–147 & Matharu N & Ahituv N. *PLoS Genetics* 2015; 11(12):e1005640

Types of epigenetic modifications

Chromatin remodelling

Histone modifications

DNA methylation



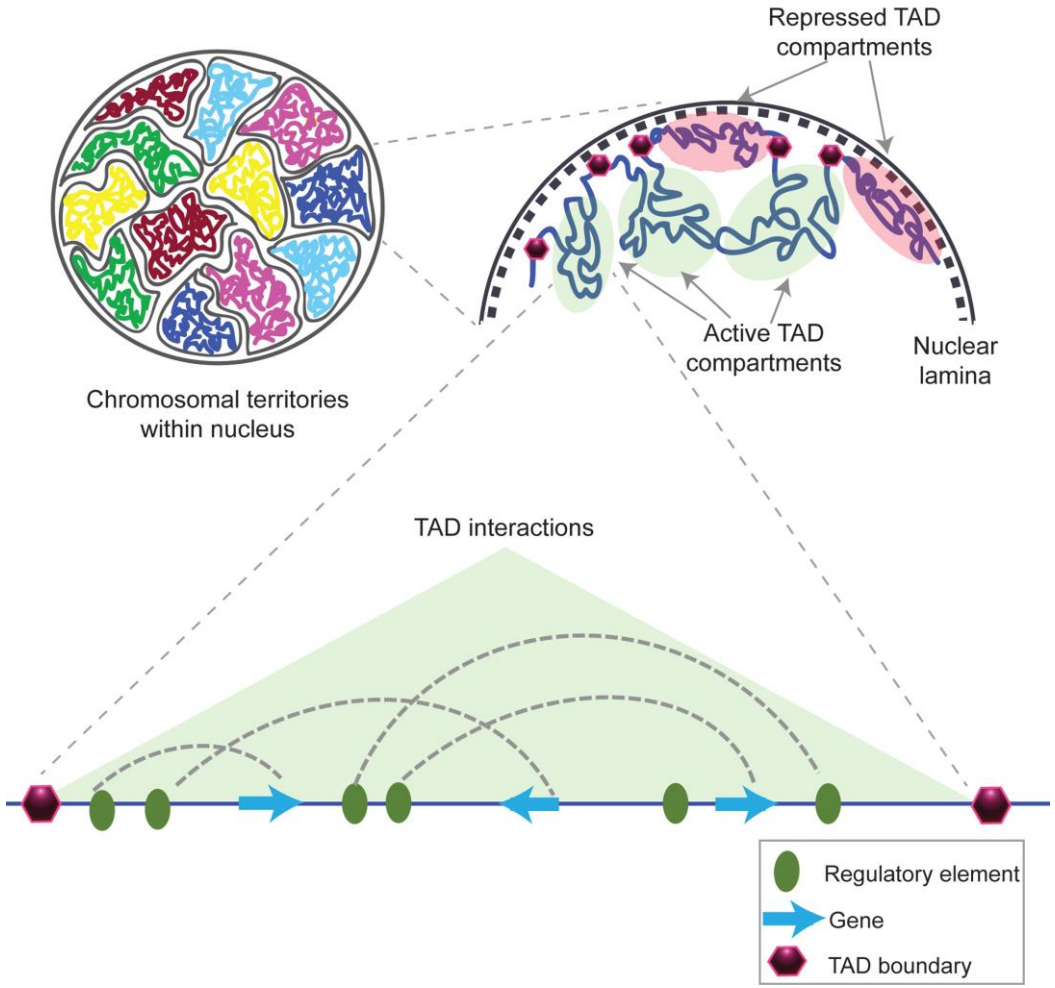
Feinberg A. *New England Journal of Medicine* 2018; 378:1323-1334

Chromatin remodelling

Changes to primary structure
- Heterochromatin ↔ euchromatin

Changes to higher-order structure
- Rearrangements of TADs and chromosomal territories
- Brings distant regions together which require the same active/repressed state

Controlled by:
- Histone modifications
- Chromatin remodelling complexes

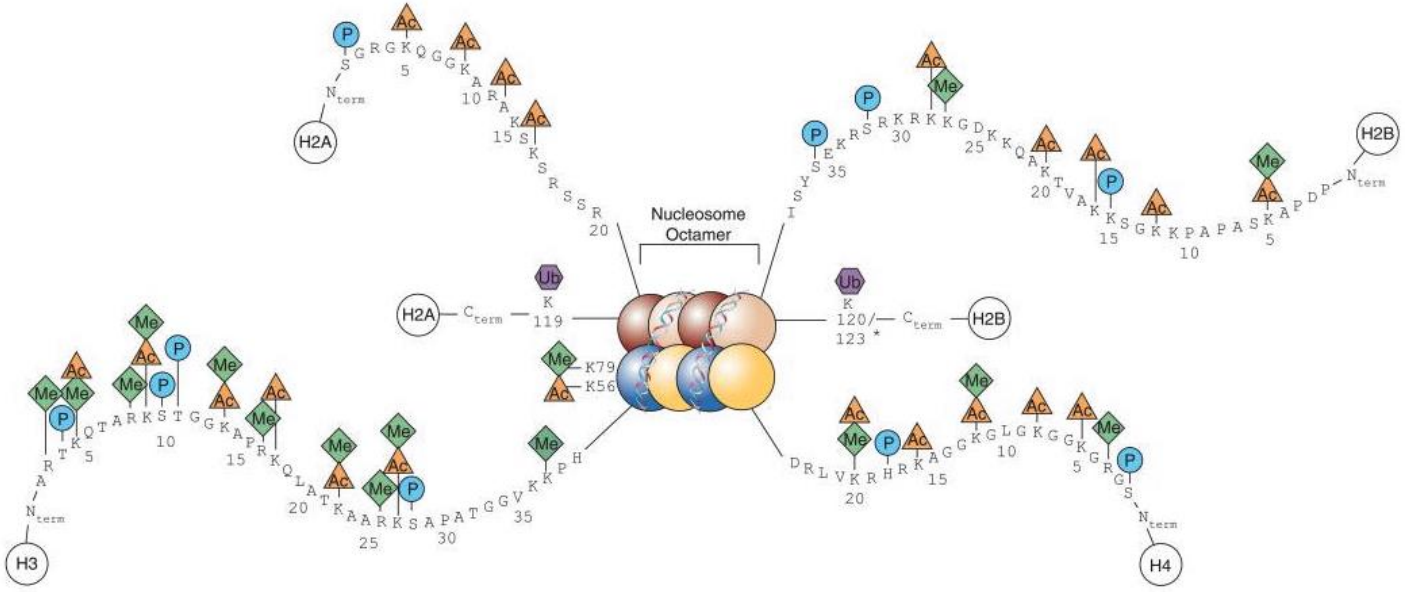


Matharu N & Ahituv N. *PLoS Genetics* 2015; 11(12):e1005640

Histone modifications

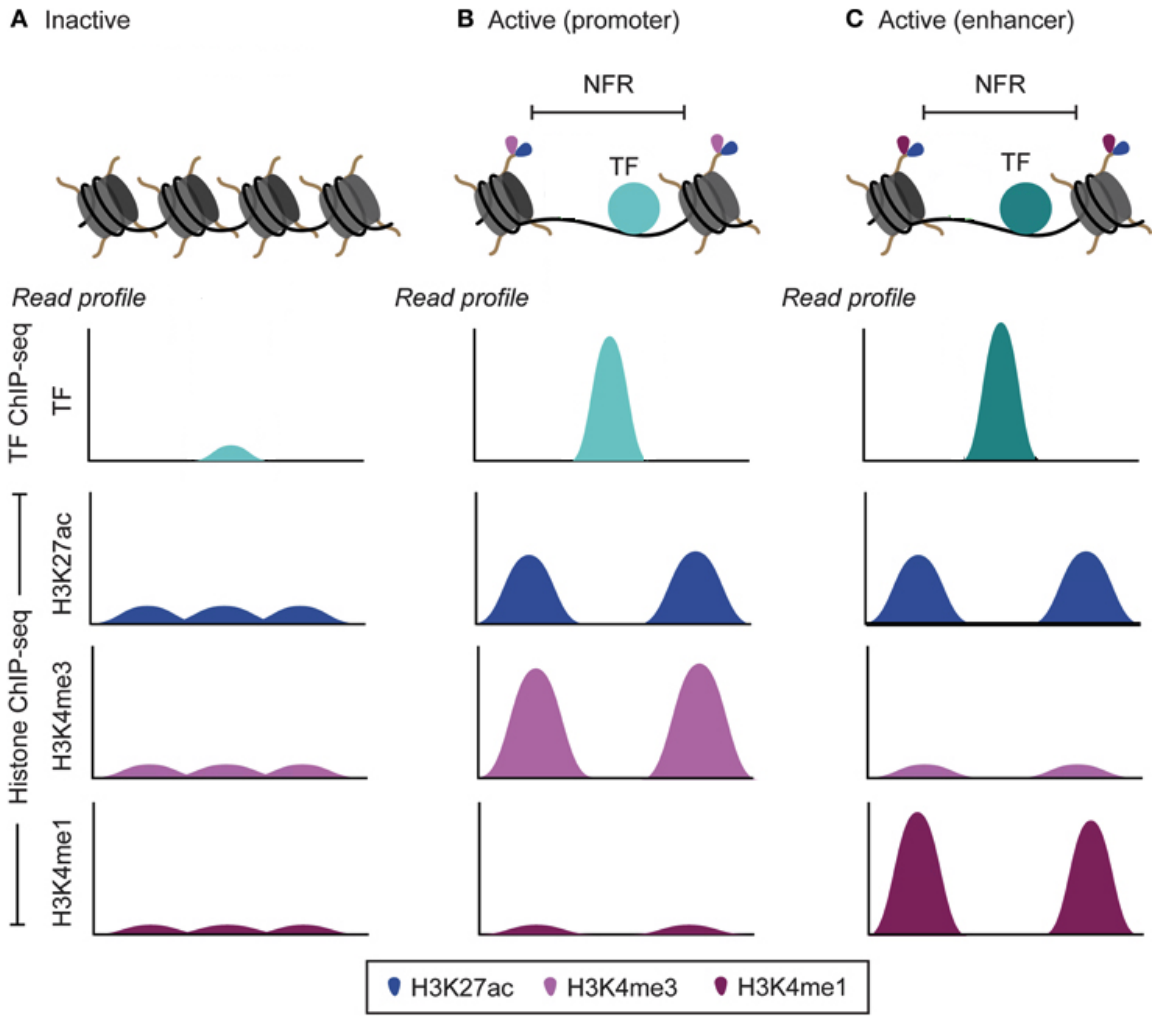
Enzymatic addition and removal of molecules to amino acids on the tails of histones within the histone octamer (nucleosome)

Modifications can:
- Disrupt the contact between nucleosomes
- Disrupt the contact between the histones and the DNA
- Recruit remodelling proteins
- Change how tightly the chromatin is packed and how easily transcriptional machinery can bind to DNA



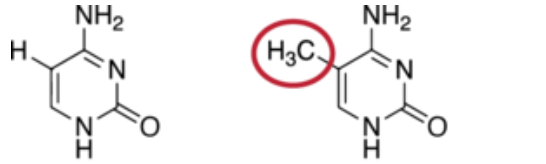
Keppler BR & Archer TK. *Expert Opinion on Therapeutic Targets* 2008; 12(10): 1301-1312.

Histone modifications



Pundhir S, Poirazi P & Gorodkin J. *Frontiers in Genetics* 2015; 6(188).

DNA methylation



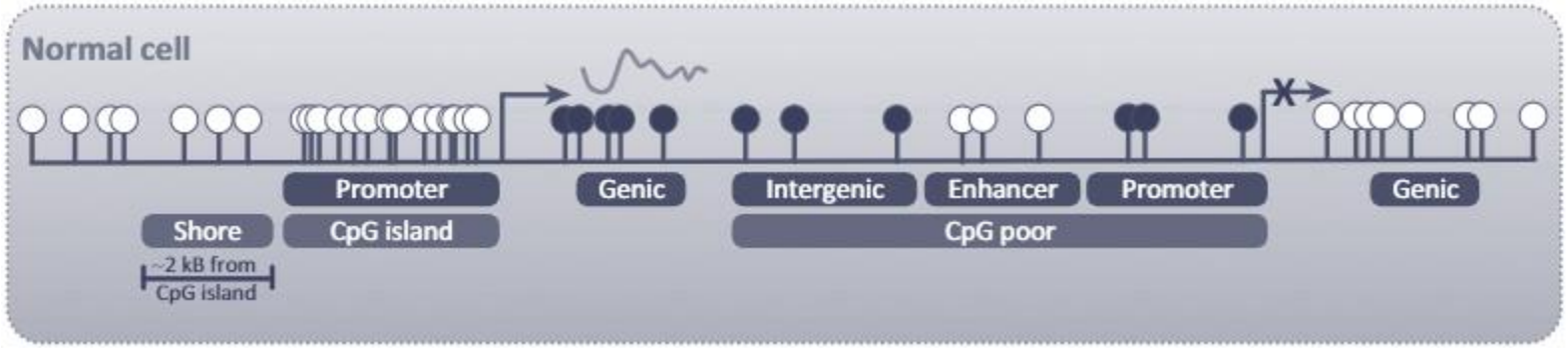
Cytosine methylated Cytosine

Methyl group added to the fifth carbon of a cytosine base, in the context of CG sequences

The effect of DNA methylation on gene expression depends on the location of the methylation

CpG sites in the promoter are typically unmethylated
- Suppression of gene expression is achieved instead through histone modifications
- Promoter hypermethylation does occur in imprinted genes and on the inactive X-chromosome

CpG sites outside the promoter are typically extensively methylated
- Maintains chromosomal stability
- Gene bodies are also methylated – even in transcribed genes



DNA methylation maintained by *DNMT1*, *DNMT3a* and *DNMT3b*
Removed by TET family enzymes with the help of *IDH1/2*

Stirzaker C *et al. Trends in Genetics* 2014; 30(2): 75-84

How do we detect and measure epigenetic modifications?

Higher-order chromatin structure
- Chromosome conformation capture with 4C, 5C and Hi-C

Primary chromatin structure
- DNase-seq, ATAC-seq

Histone modifications
- ChIP-seq

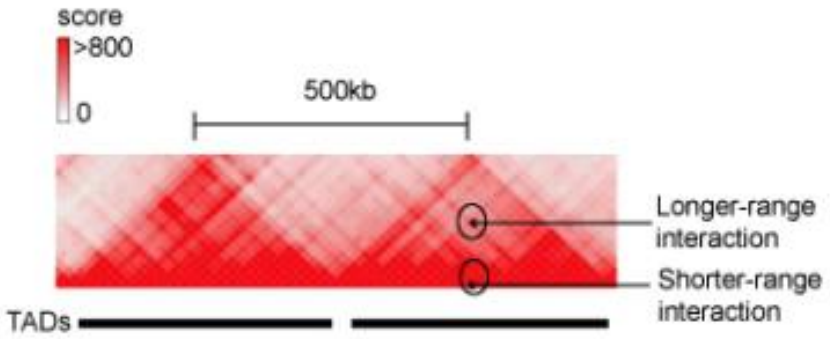
DNA methylation
- Bisulphite sequencing, DNA methylation array

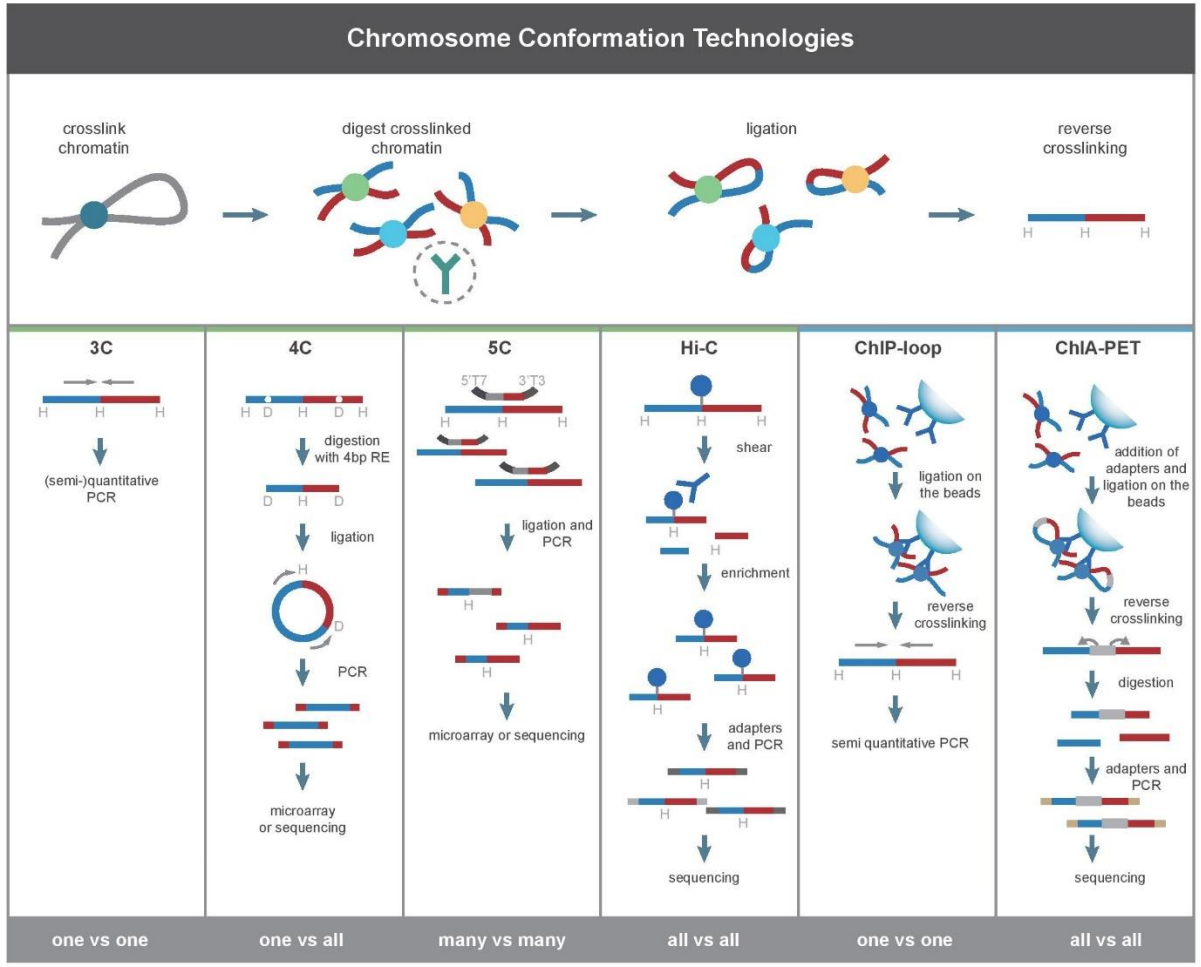
All of these techniques capture a profile for (mostly) one tissue/cell type at one time
- Implications for extrapolating results

Chromosome conformation capture

Chromatin is cross-linked using formaldehyde so that DNA regions within spatial proximity are glued together with protein complexes

DNA is then fragmented, and ends ligated



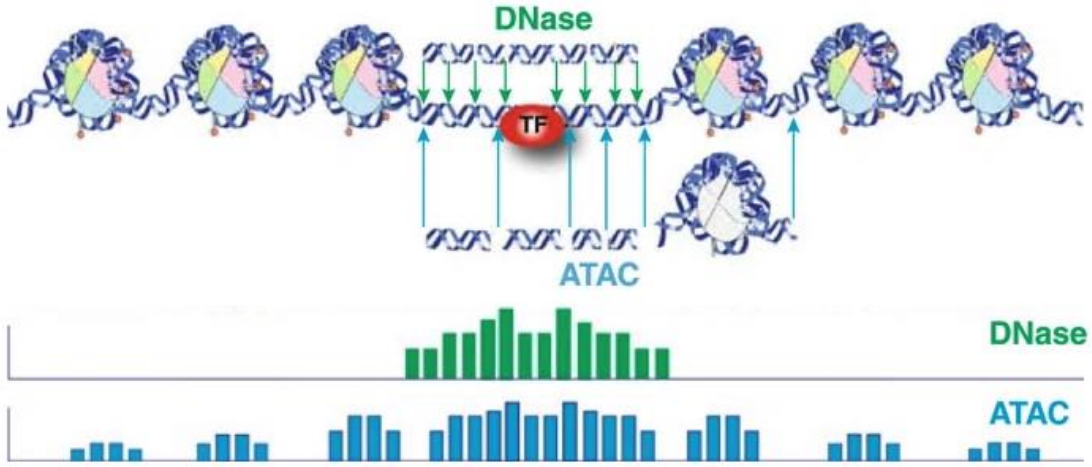


Guoliang L *et al. BMC Genomics* 2014; 15(S11).

Primary chromatin structure

DNase-seq
- Identifies DNase I hypersensitivity sites
- DNase I hypersensitivity sites appear as a result of transcription factor binding and the removal of nucleosomes – in active chromatin

Assay for Transposase Accessible Chromatin (ATAC-seq)
- ‘Tagmentation’ fragmenting and tagging simultaneously with Tn5 transposase
- Identifies sites of nucleosome occupancy and nucleosome-free regulatory regions

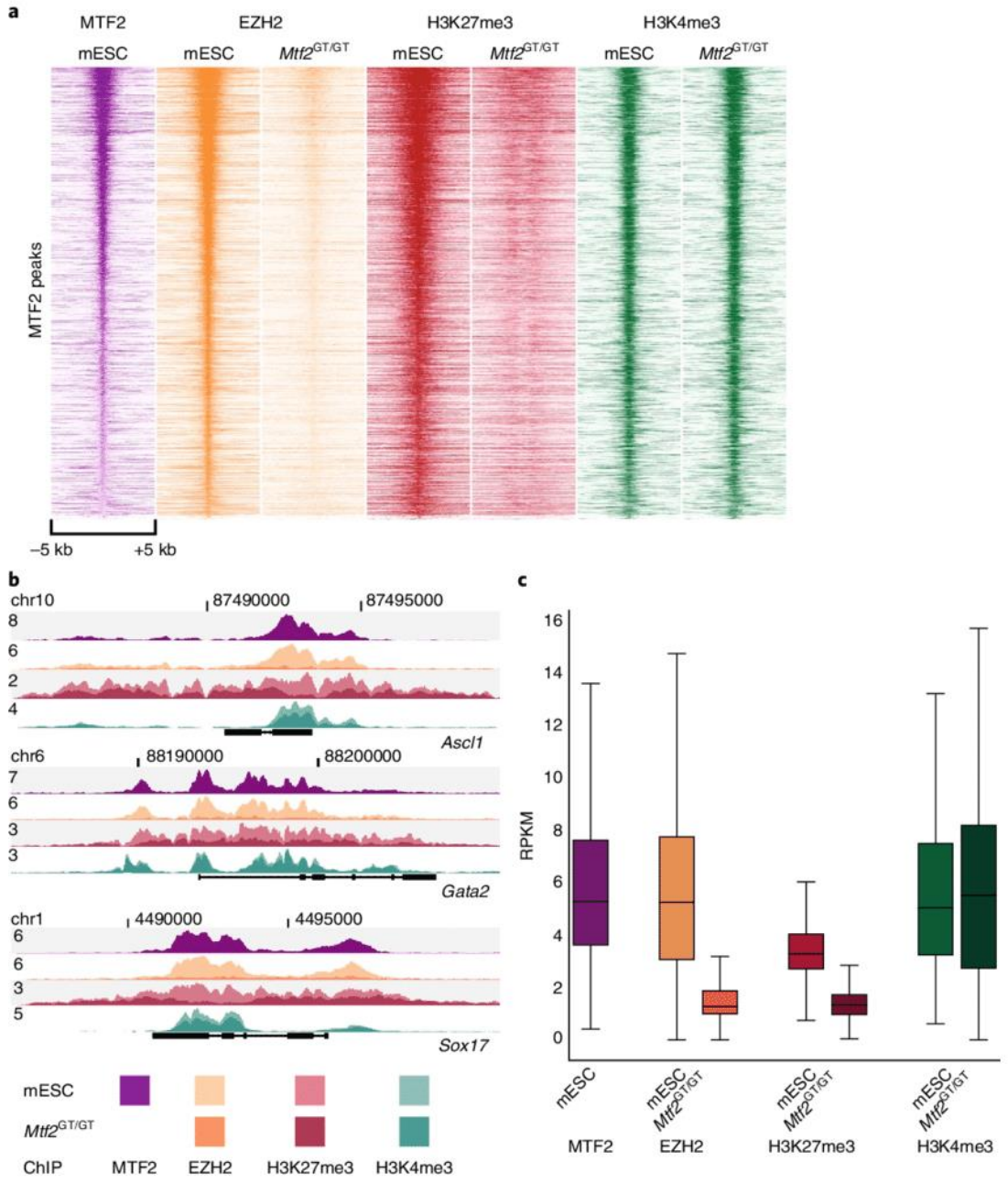


Tsompana M & Buck MJ. *Epigenetics & Chromatin* 2014; 7(33).

Histone modifications

Chromatin immunoprecipitation and sequencing (ChIP-seq)

- Identifies all the binding sites for a particular protein (such as a TF), or histone modification
- Some regions are promoters and enhancers in virtually all cell types, others can be cell-type specific



Perrino M. *et al. Nature Genetics* 2018; 50(7).

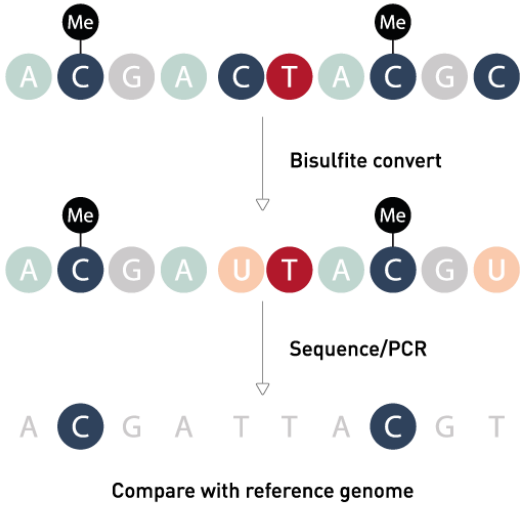
DNA methylation

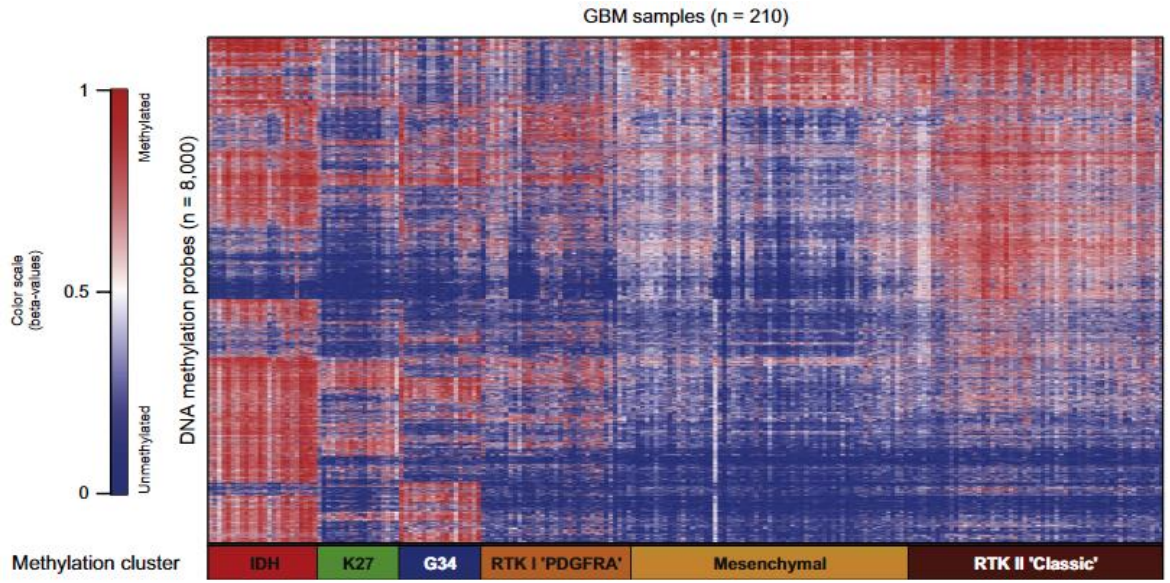
Bisulphite conversion techniques

- Whole genome bisulphite sequencing (~95%)
- Reduced representation bisulphite sequencing (~4%)
- Methylation microarrays (~2%)

Non-BC techniques

- MBDCap-seq (~18%)



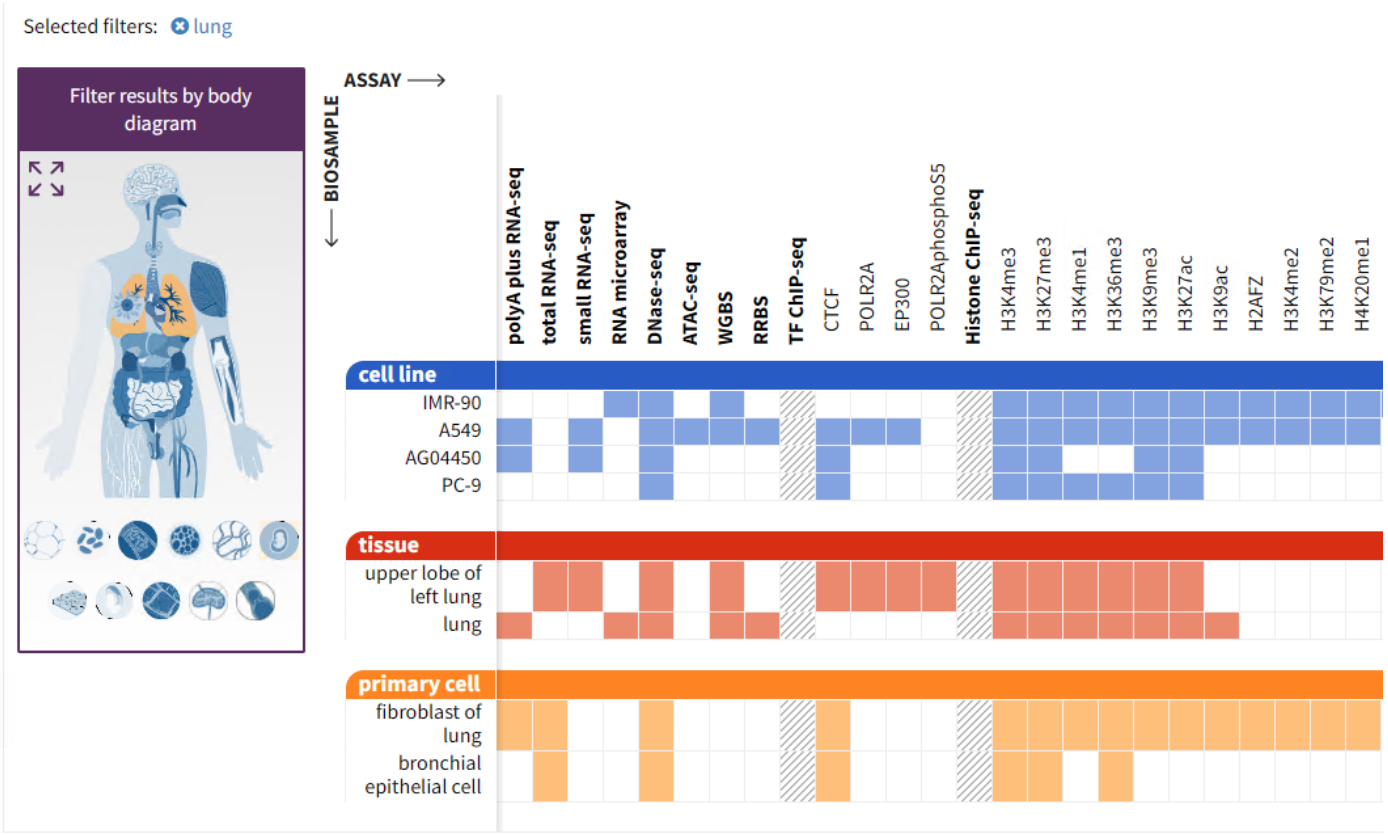


Sturm D *et al. Cancer Cell* 2012; 22(4): 417–418.

Epigenetic profile resources

International Human Epigenome Consortium

– Blueprint, ENCODE, Roadmap



Part 1: Summary

- Epigenetic mechanisms alter gene expression without changing the underlying DNA sequence
- They are dynamic, responding to the environment, and ensure that a cell behaves the way it's supposed to
- Can be measured with sequencing and/or microarrays
- Require level of computational expertise

Outline

Part 1: Introduction to epigenomics

- Concepts: role of epigenetics, refresh gene expression, epigenetic mechanisms
- Techniques: detecting and measuring epigenetic modifications
- Resources: databases of regulatory annotations and epigenetic profiles

Part 2: Cancer epigenetics

- Concepts: introduction to epigenetics in cancer

Part 3: Application of DNA methylation arrays in cancer

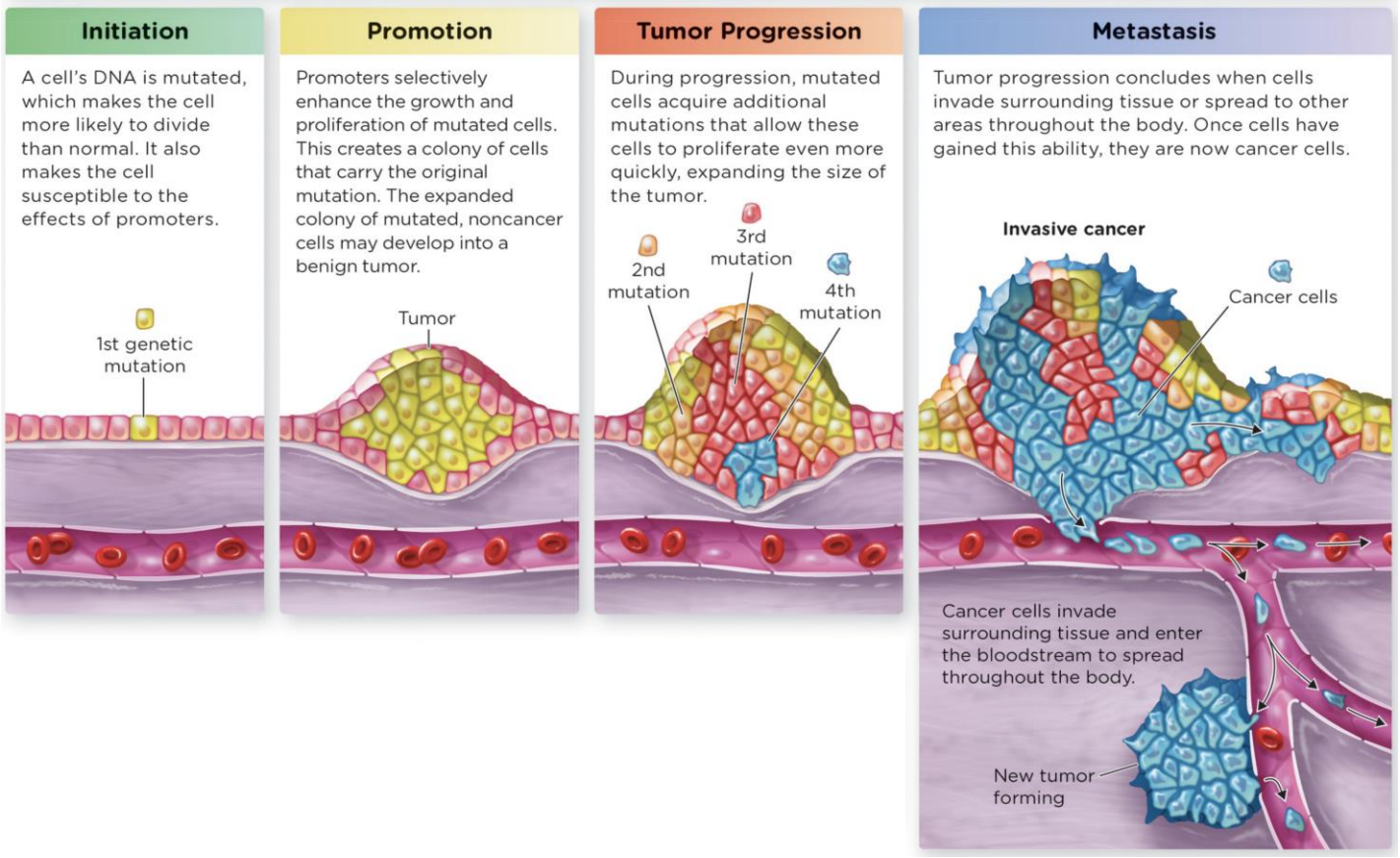
- Methodology: processing DNA methylation arrays, normalisation methods
- Research applications: principal components analysis, global methylation, deconvolution
- Clinical application: DNA methylation arrays for diagnosis and treatment

Part 2: Epigenetics in cancer

Classical model of cancer development

- Cancer arises from single cells that accumulate genetic changes during subsequent divisions to form a sub-population of cells with selective advantage over non-tumour cells
- Genetic changes allow for dysregulation of the cell's proliferative mechanisms, bypassing of checkpoints that regulate cell growth and division, and evasion of immune surveillance = tumour growth and malignancy

- Genetic changes aren't the only way a cell can acquire disrupted gene function – epigenetic mechanisms play a key role in the tumourogenic process



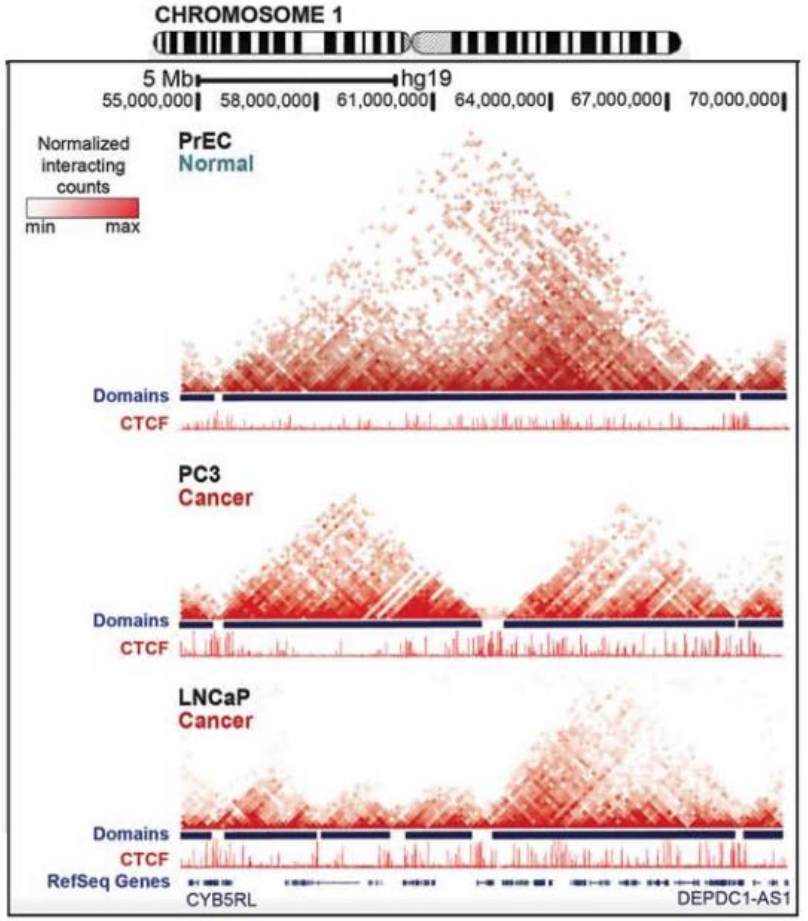
Chromatin structure in cancer

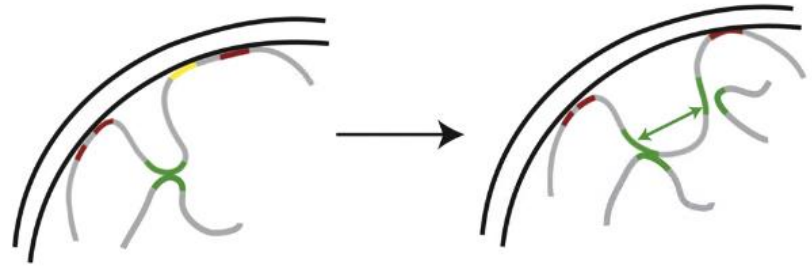
A notable feature of cancer cells is abnormal chromatin and nuclear architecture

Cancer cells retain the ability to segment their genomes into topological associated domains

- In cancer these TADs are more numerous and smaller
- Often associated with changes in copy number
- Produce novel cancer-specific chromatin interactions to alter gene expression
- Alterations to TAD structure increased genomic instability

- Mechanism for change to TADs in cancer is unknown





Taberlay PC *et al*. *Genome Research* 2016; 26: 719–731; Reddy KL & Feinberg AP. *Seminars in Cancer Biology* 2013; 23: 109–1135.

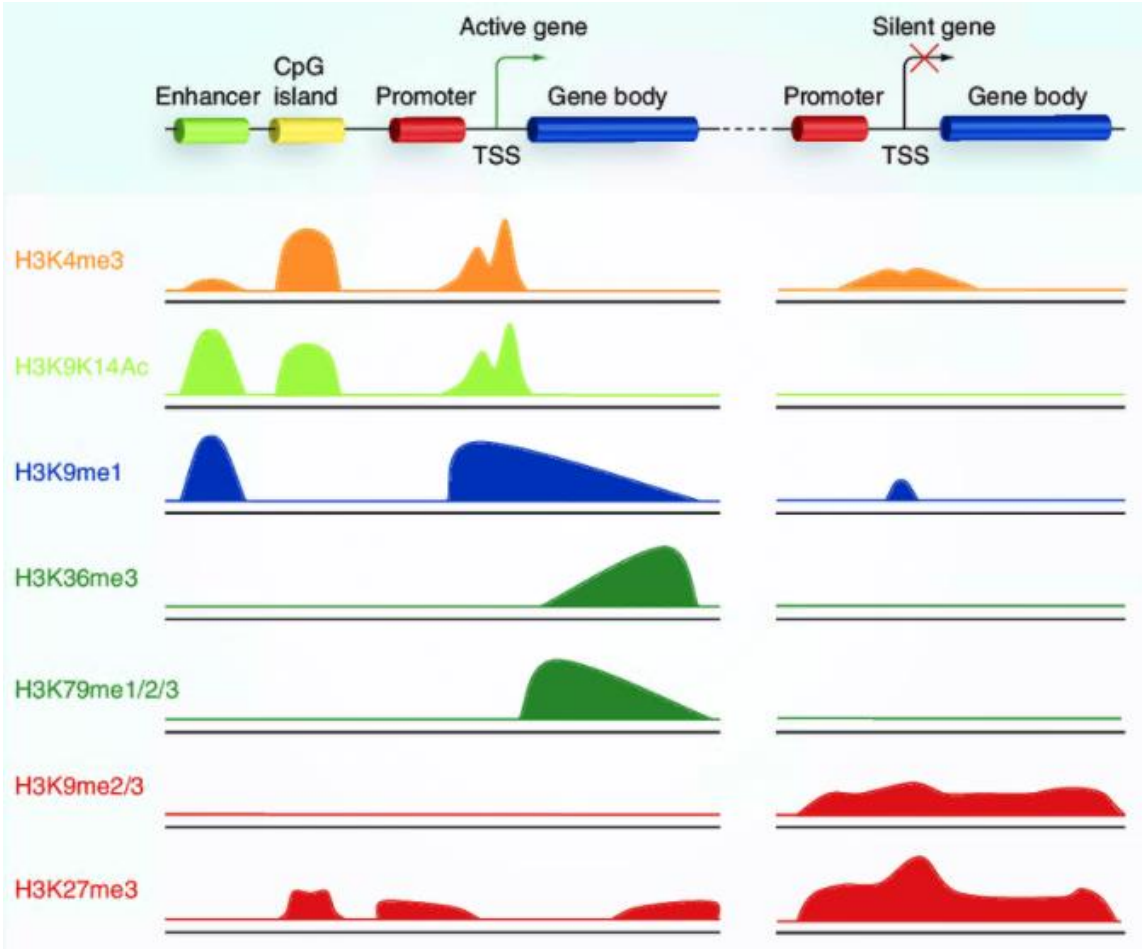
Histone modifications

Histone modifications facilitate or block transcription

Aberrant patterns of histone modifications in cancer result in activation and silencing of genes to promote tumourogenesis

Typically, patterns are altered by acquisition of genetic changes to histone modifying enzymes

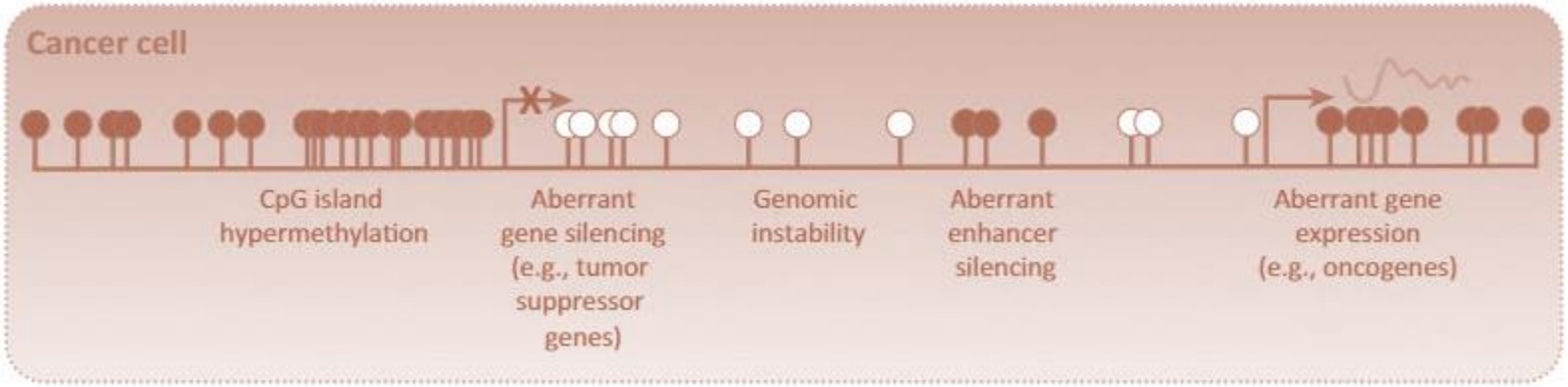
AML1-ETO fusion protein recruits HDAC1 → transcriptional repression



Siew Lim P, Shannon MF & Hardy K. *Epigenomics* 2010; 2(6): 775–795.

DNA methylation in cancer

Nearly all human neoplasms exhibit aberrant DNA methylation patterns
- Genome-wide hypomethylation
- Promoter-specific hypermethylation



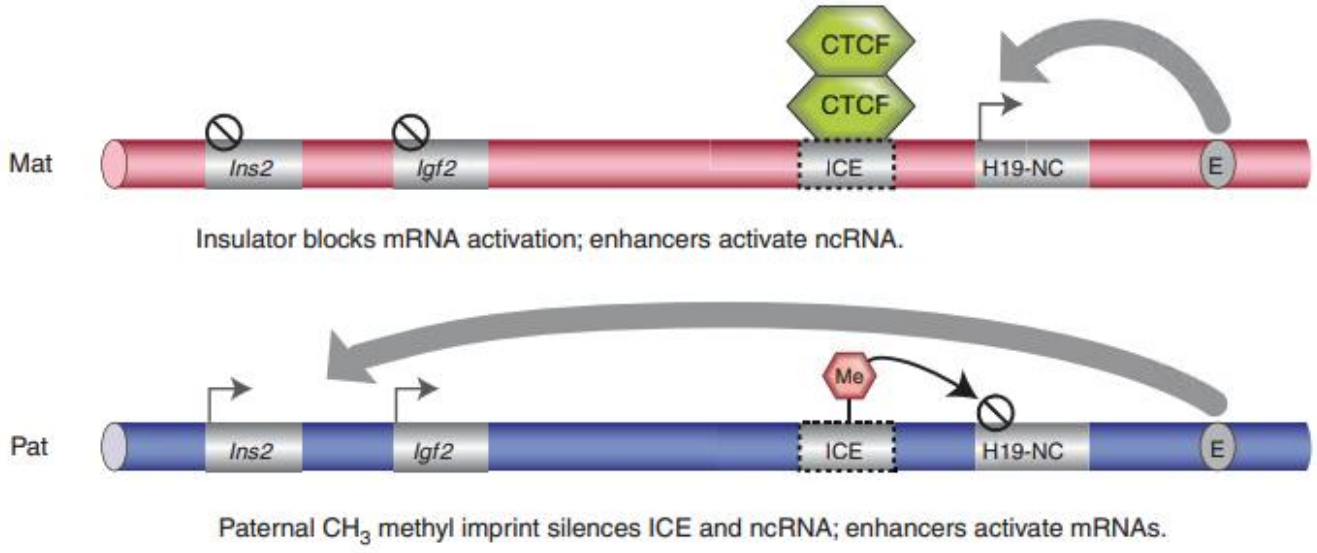
Genome-wide hypomethylation
- Reactivation of transposable elements, chromosomal instability, aneuploidy, translocations, loss of imprinting

Promoter hypermethylation
- Silencing of tumour suppressor genes

Stirzaker C *et al. Trends in Genetics* 2014; 30(2): 75-84

DNA methylation in cancer

Methylation of other regulatory regions



Barlow DP & Bartolomei MS. *Cold Spring Harbor Perspectives in Biology* 2014; 6.

Part 2 Summary

What comes first? Epigenetic or genetic changes…

Benign tumours and pre-malignant lesions display genome-wide hypomethylation
- May be triggered by chronic inflammation

Methylation-induced silencing of TSGs can occur prior to mutations in known driver genes
- May be triggered by environmental factors such as synthetic hormones, heavy metals, and smoking

However, genetic changes can alter expression of epigenetic modifiers leading to widespread epigenetic changes
- IDH1/2 mutations in brain cancer → no α-ketoglutarate → no demethylation → hypermethylation of TSGs

Epigenetic mechanisms are disrupted in cancer, they both lead to, and are a product of, cancer development

Outline

Part 1: Introduction to epigenomics

- Concepts: role of epigenetics, refresh gene expression, epigenetic mechanisms
- Techniques: detecting and measuring epigenetic modifications
- Resources: databases of regulatory annotations and epigenetic profiles

Part 2: Cancer epigenetics

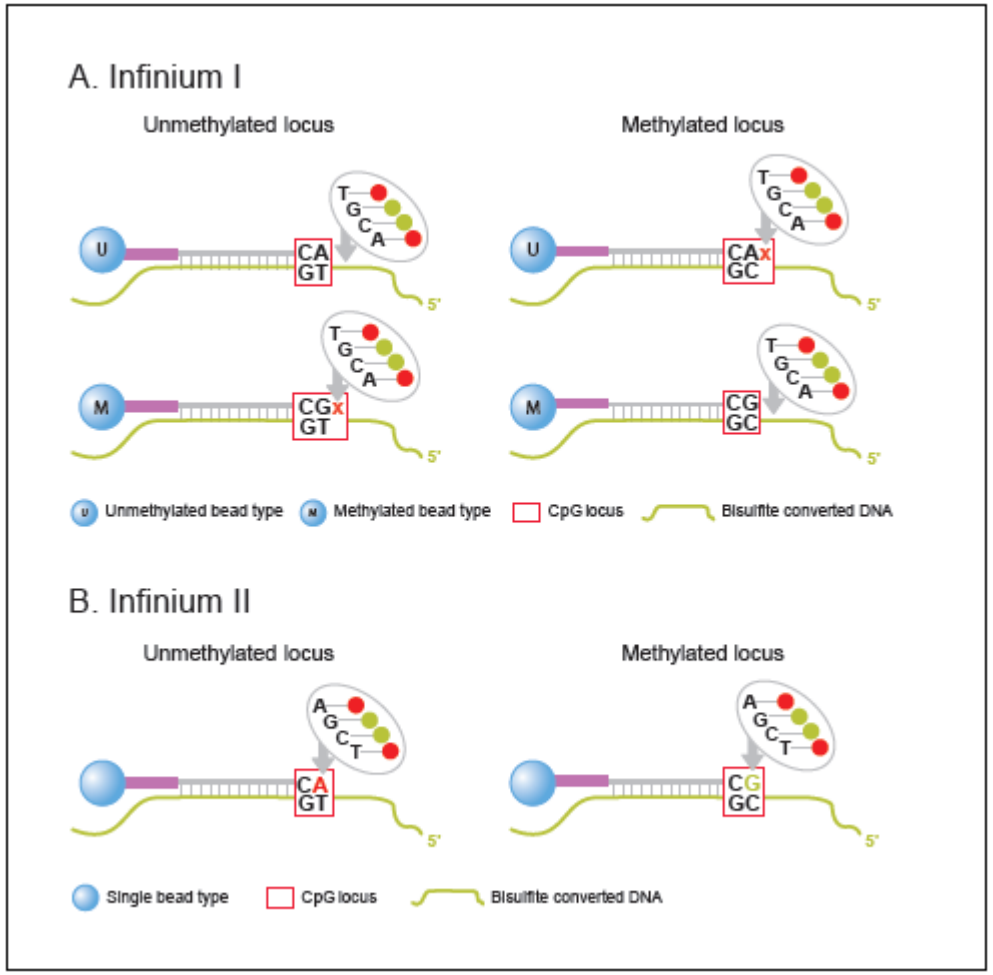
- Concepts: introduction to epigenetics in cancer

Part 3: Application of DNA methylation arrays in cancer

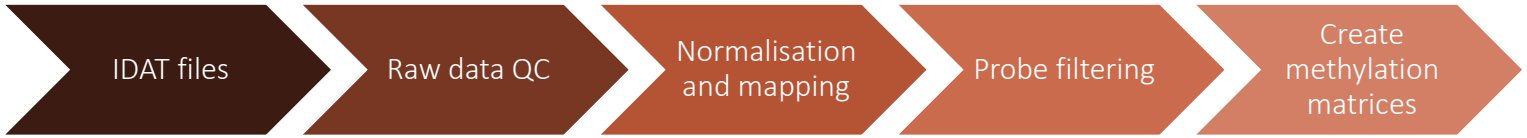
- Methodology: processing DNA methylation arrays, normalisation methods
- Research applications: principal components analysis, global methylation, deconvolution
- Clinical application: DNA methylation arrays for diagnosis and treatment

Illumina InfiniumEPIC BeadChip Array

1. Bisulphite converted DNA is hybridised to oligonucleotide coated beads
   - Two probe types: Infinium I and Infinium II
2. Array is fluorescently stained
3. Intensity of methylated and unmethylated probe signals captured as IDAT files



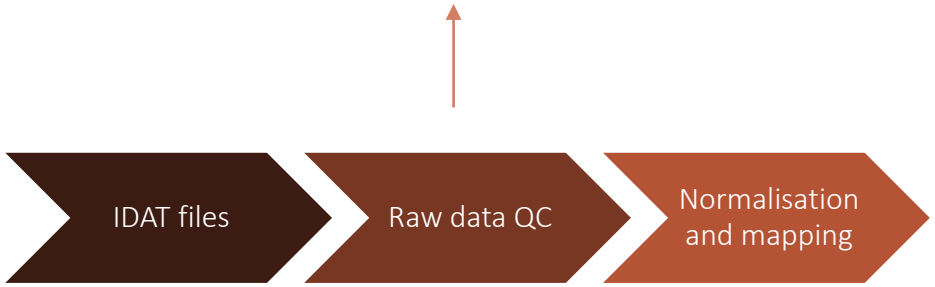
Processing IDAT files to methylation matrices



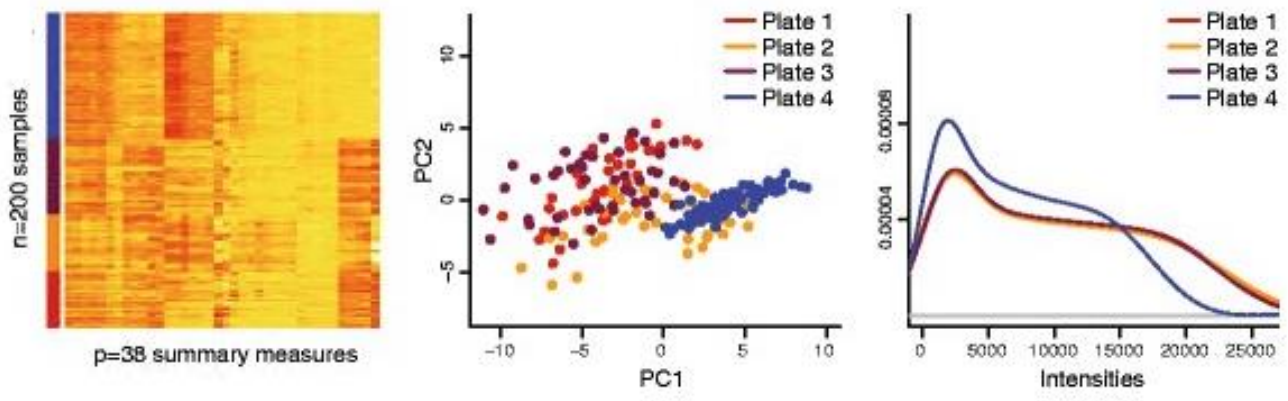
plotCtrl: examine efficiency of technical steps

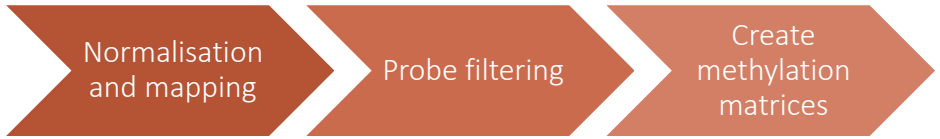
getQC: examine signal intensity

detection: examine probe detection level



Reduce within and between array technical effects





Removal of:

- Poor performing probes
- Cross-reactive probes
- SNP-associated probes
- Sex chromosome probes

Beta values

$$\beta = \frac{M}{M+U+100}$$

M values

$$\mathrm{M} = \log_2(\frac{M}{U})$$

Replacing –Inf/+Inf values

Applications of DNA methylation array data for research

What can we learn from running tumour samples on DNA methylation arrays?

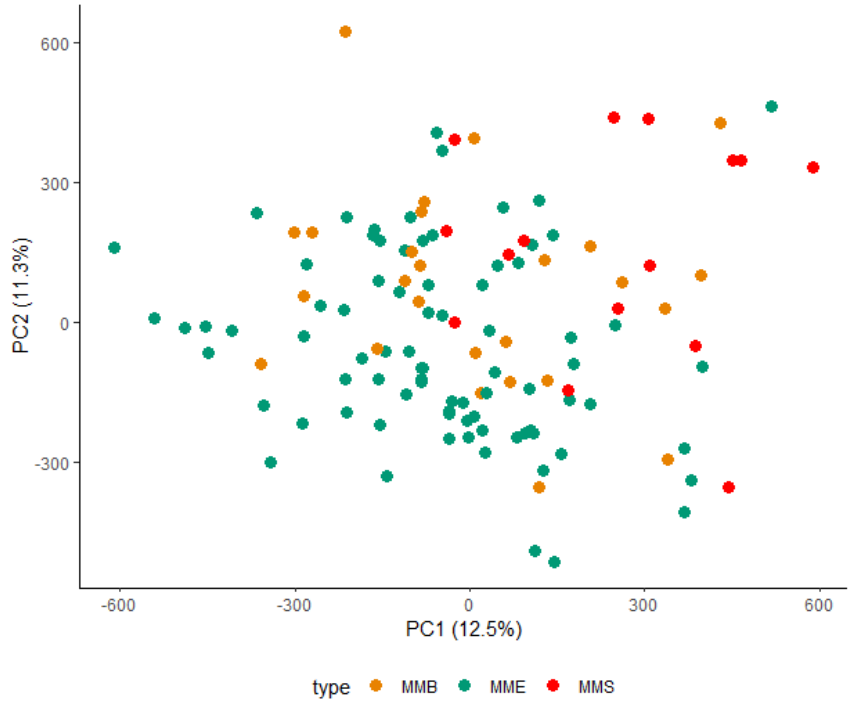
Methods

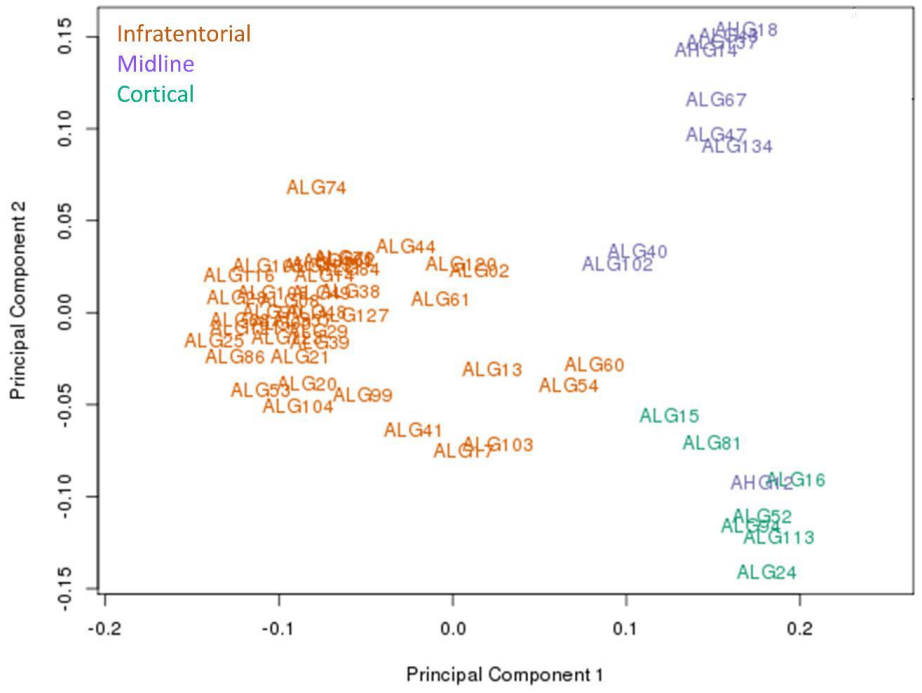
- Principal components analysis and beta density
- Identifying differentially methylated sites and regions
- Deconvolution of cell types
- Estimating tumour purity
- Predicting smoking status

Principal components analysis

PCA allows us to

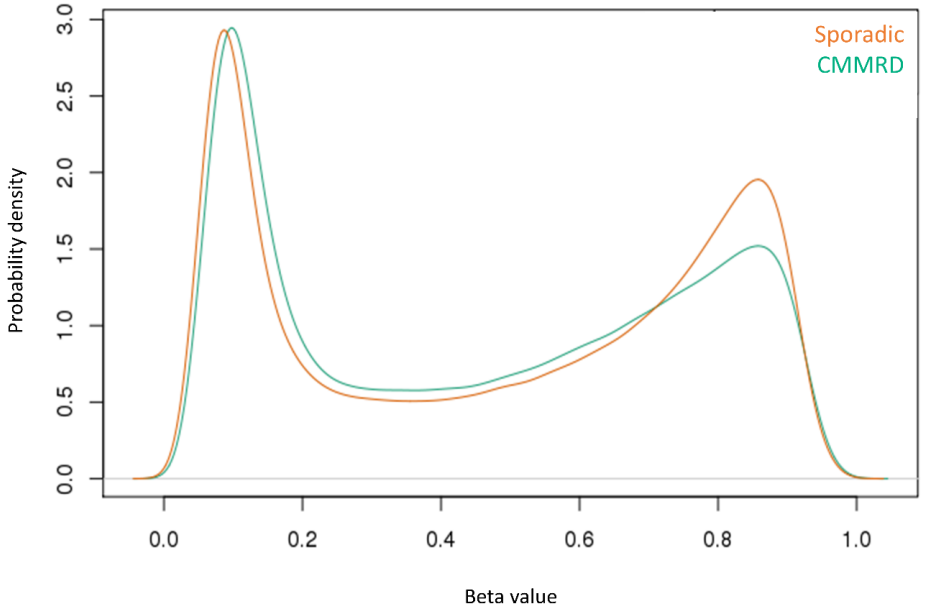
- Examine clinical and technical sources of variation within the data
- Identify interesting drivers of DNA methylation profile

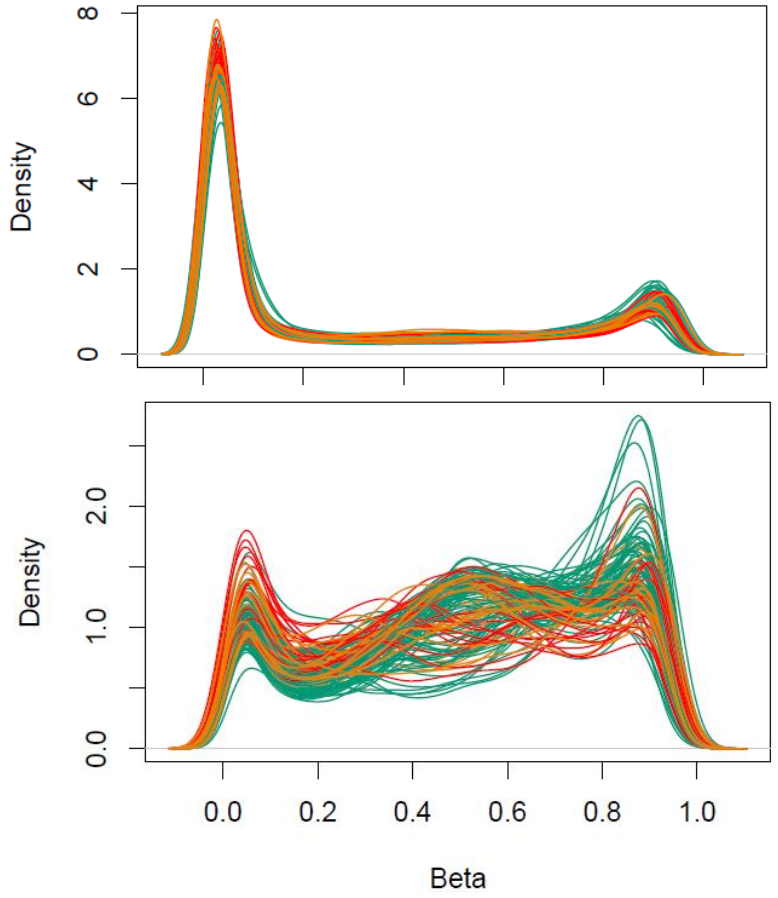




Exploring beta density

Beta density plotting allows us to compare profiles between groups and regions

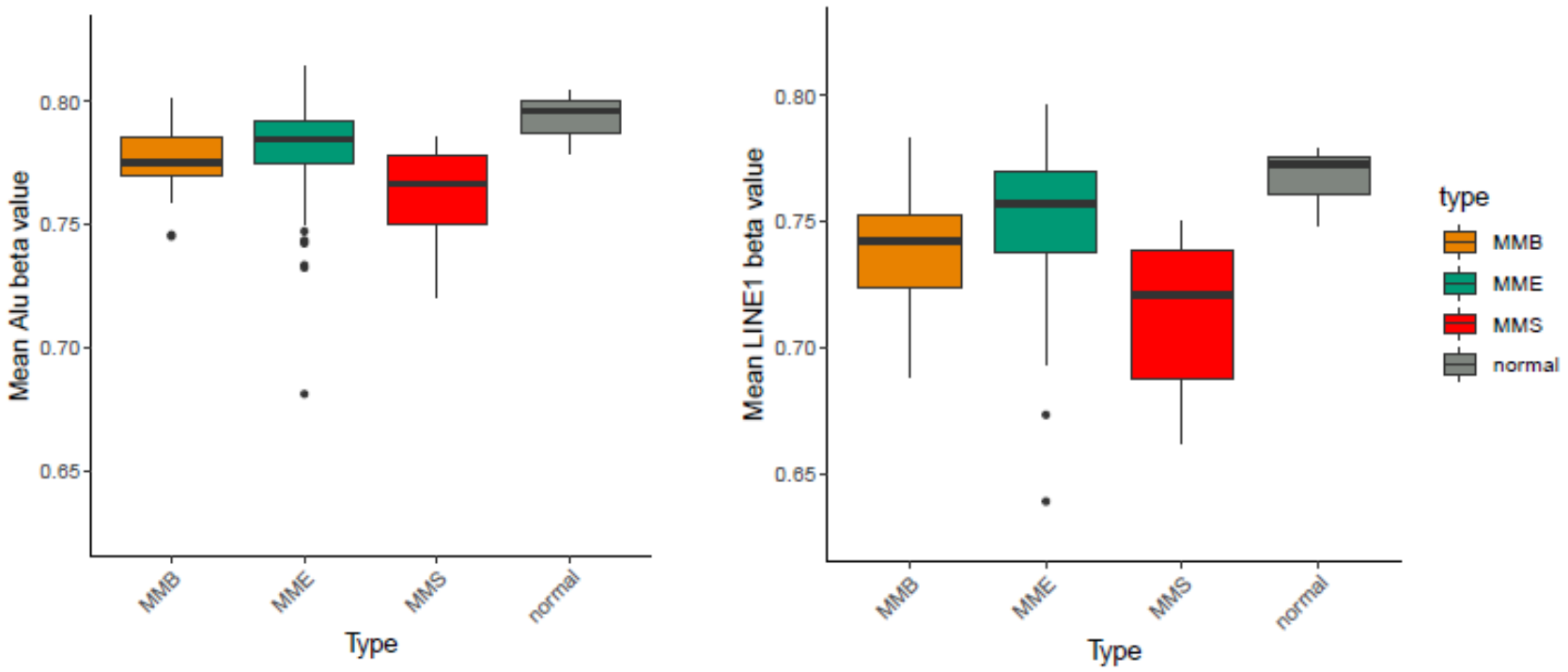




Estimating global methylation

Using REMP (R) to identify CpG sites within Alu and LINE1 elements

Average methylation at LINE1 or Alu can be used as a proxy for global methylation



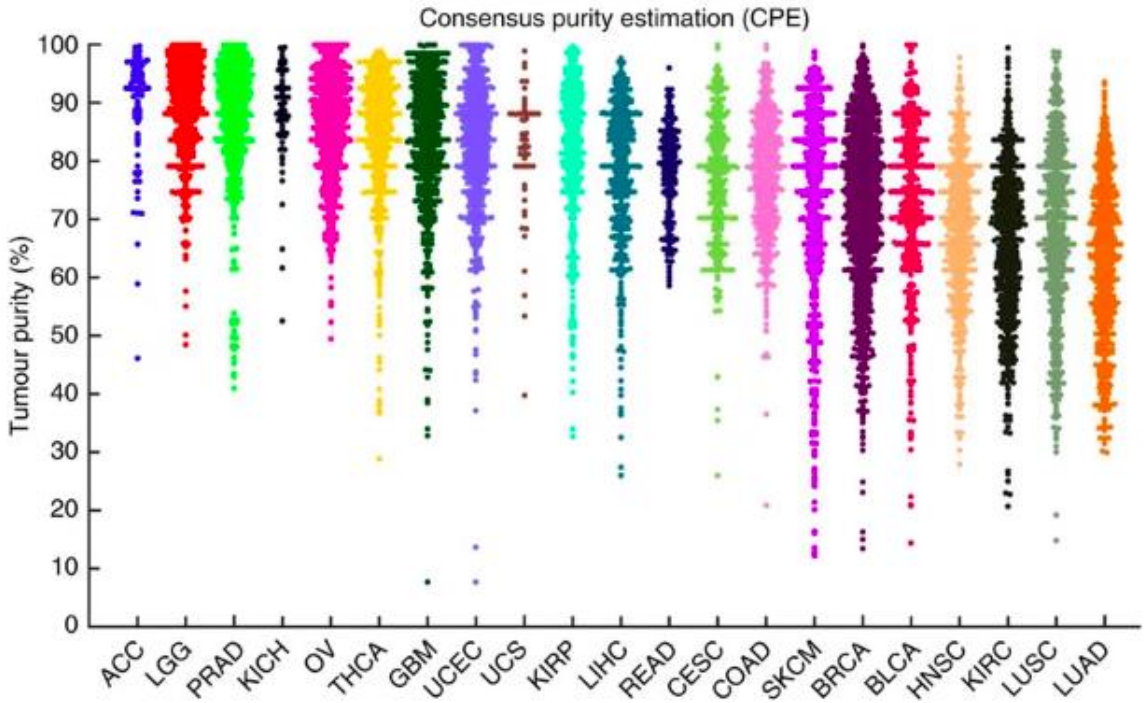
Estimating tumour purity

Tumour purity varies for intrinsic and extrinsic reasons

- Immune and stromal infiltration within a tumour
- Amount of normal tissue surrounding tumour biopsy

LUMP – leukocyte unmethylation to infer tumour purity

- 44 CpG sites as the junction between 30,106 sites consistently unmethylated in leukocytes (<5%) and 174,696 consistently methylated sites (>30%) in 21 cancer types
- Purity estimate = mean(LUMP)/0.85

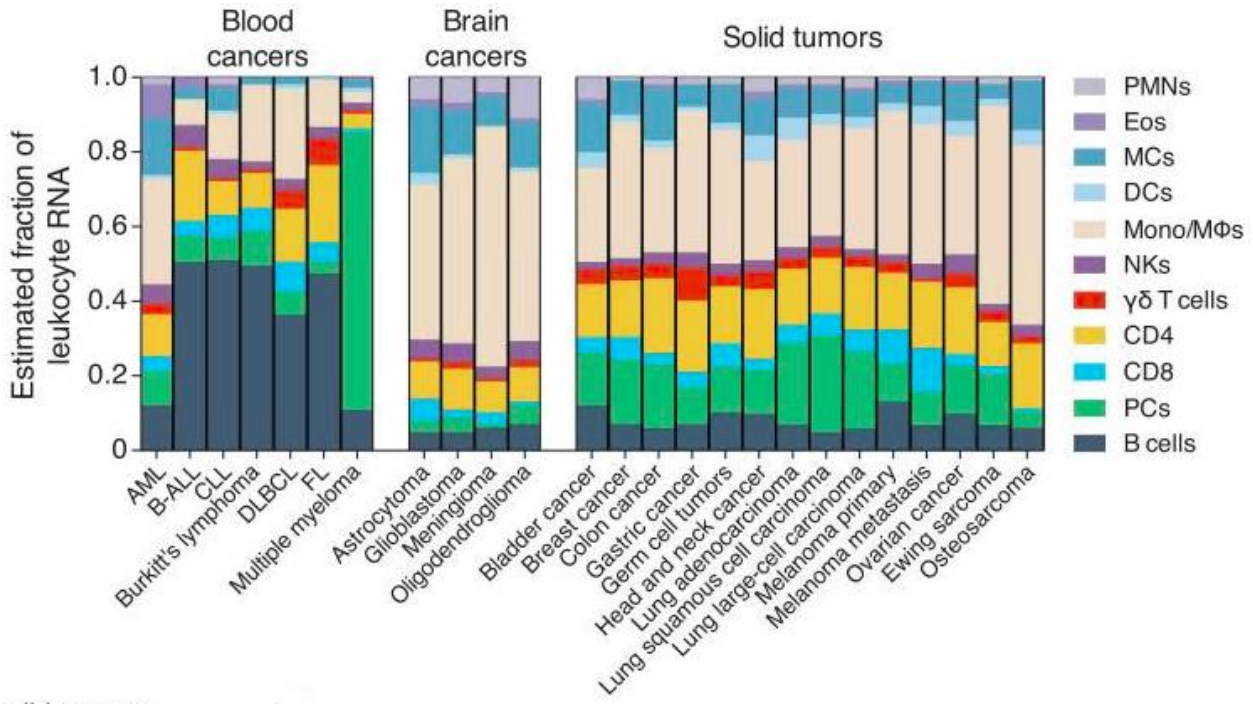


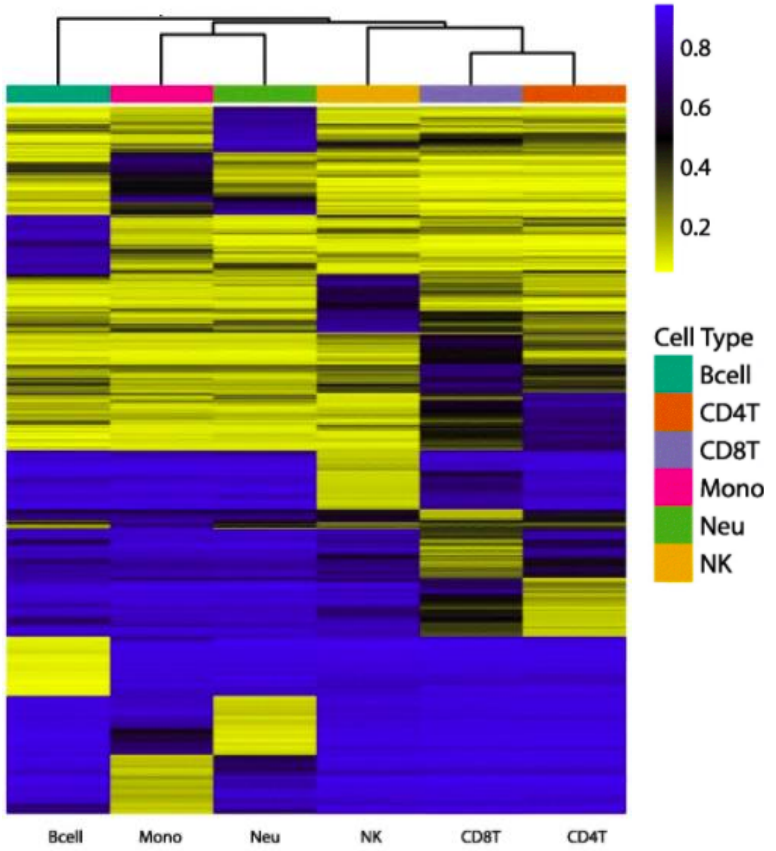
Aran D, Sirota M & Butte A. *Nature Communications* 2015. 6(8971).

Deconvolution of cell types

Using deconvolution methods, we can identify proportions of:

- Whole blood cell types
- Umbilical cord blood cell types
- Frontal cortex cell types



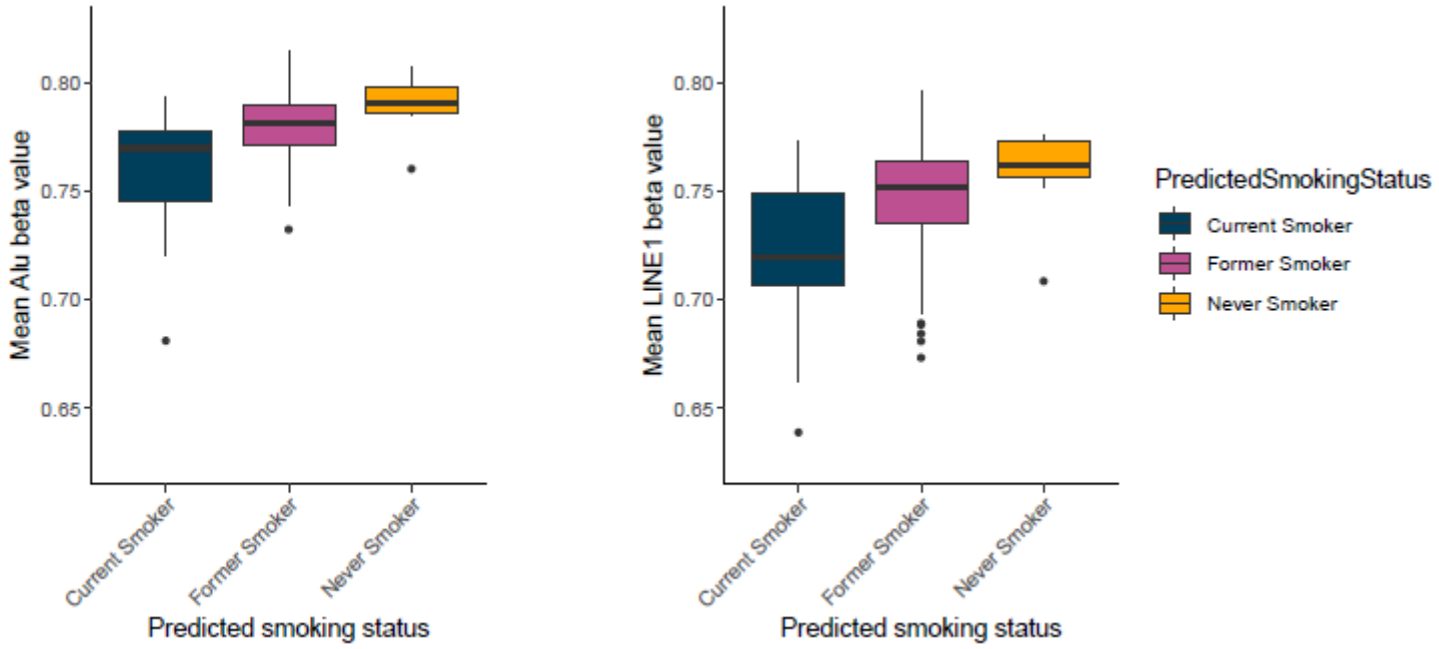


Gentles AJ *et al. Nature Medicine* 2015; 21(8): 938–945; Salas LA *et al. Genome Biology* 2018. 19(64).

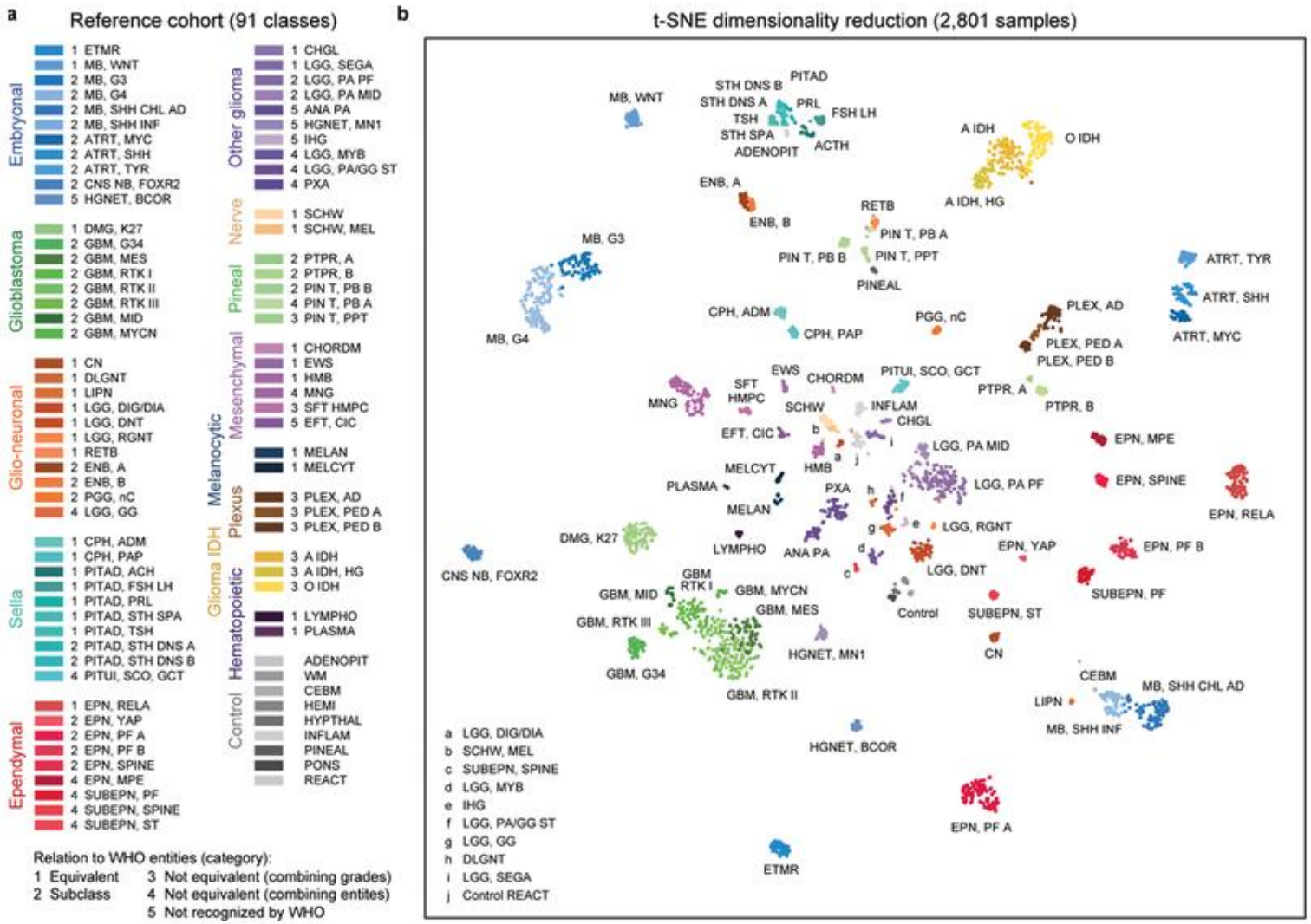
Predicting smoking status

Smoking has a large impact on DNA methylation level in blood and lung tissue

EpiSmokeR (R) can be used to predict current, former, and never smoking status

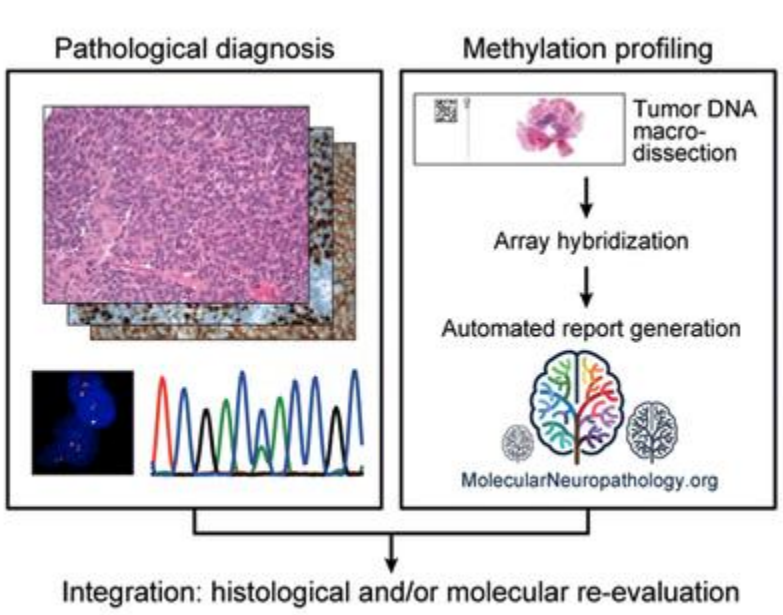


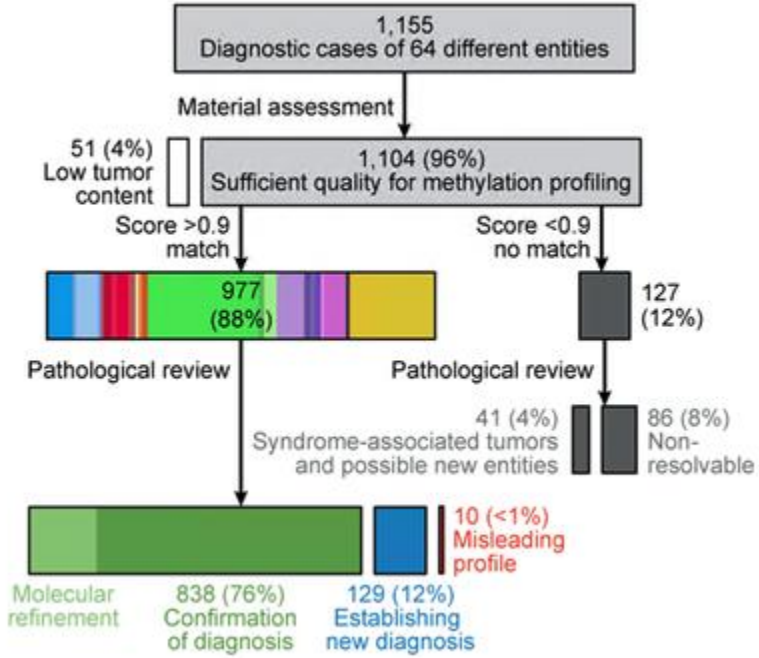
Applications of DNA methylation array data for clinical oncology



Capper D *et al. Nature* 2018. 55(7679): 469–474.

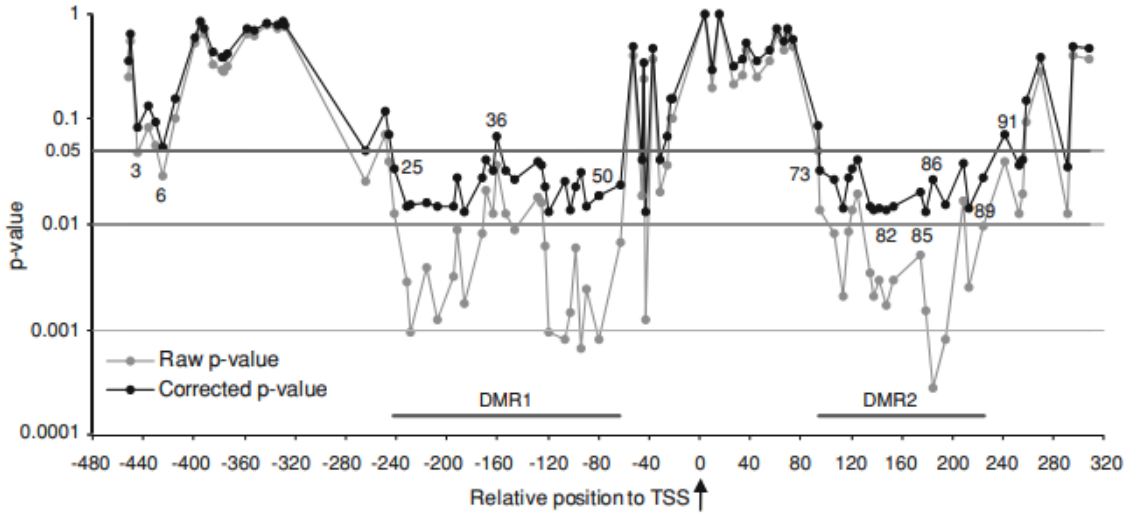
Applications of DNA methylation array data for clinical oncology

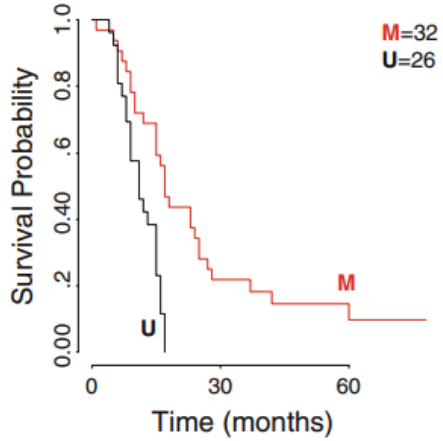


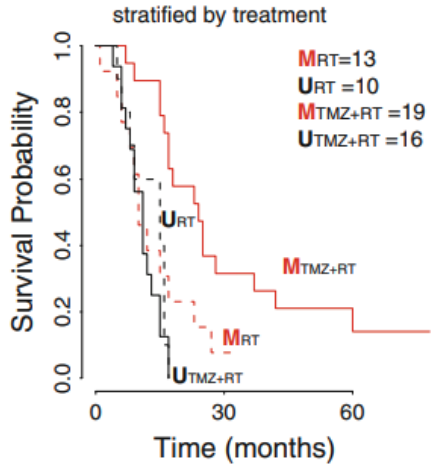


Capper D *et al. Nature* 2018. 55(7679): 469–474.

Differentially methylated sites and regions







↑ DMR methylation

↓

↓ MGMT expression

Temozolomide induces DNA alkylation

↓

No repair of DNA damage

↓

Tumour cell death

Malley DS *et al. Acta Neuropathologica* 2011. 121: 651–661; Bady P *et al. Acta Neuropathologica* 2012. 124: 547–560.

Summary Part 3

DNA methylation arrays are a useful tool in cancer research and clinical oncology

Explore global and site-specific profiles, tumour bulk composition, improve classification

Can be combined with other omics techniques to get a more complete picture of the tumour epigenome

Take home messages

Epigenetic modifications influence gene expression and therefore normal cellular functioning and disease development

- Do so without changing underlying DNA sequence
- Dynamic changes which respond to environmental stimuli

Several levels of modifications from higher-order chromosome structure to single nucleotide alterations

Patterns can be detected with a range of technologies with utility in research and clinical practice

Any questions: SextonoatesA@iarc.who.int