

A Brief Overview of Epigenomics

BIOTRAC Lecture

Sean, Soonweng, Cho, PhD

Kennedy Krieger Institute

and

Johns Hopkins University School of Medicine



sean.cho@jhmi.edu



github.com/sean-cho

Outline

What is epigenomics?

Why epigenomics?

How

- DNA methylation
- Histone modifications
- Chromatin features
 - Accessibility
 - HiC
- Integrative epigenomics
 - Imputation
 - Chromatin state
- Integrative analyses
- Study Design
 - Limitations and considerations
 - Balancing and Blocking

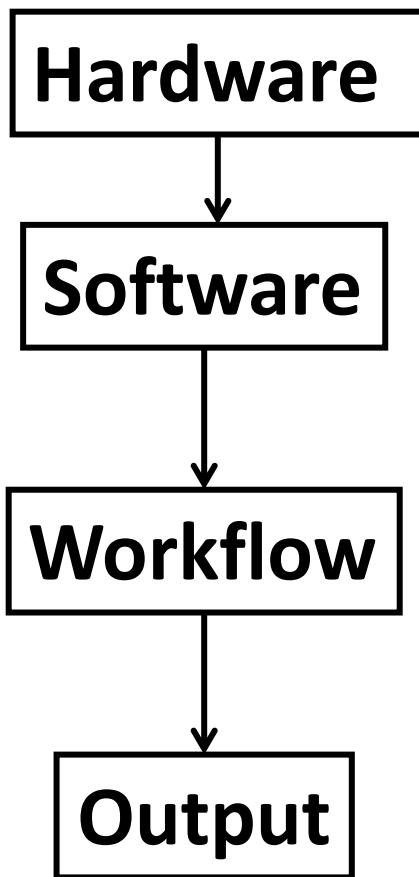
The future of epigenomics

- Nanopore/MinION sequencing
- Single cell epigenomics

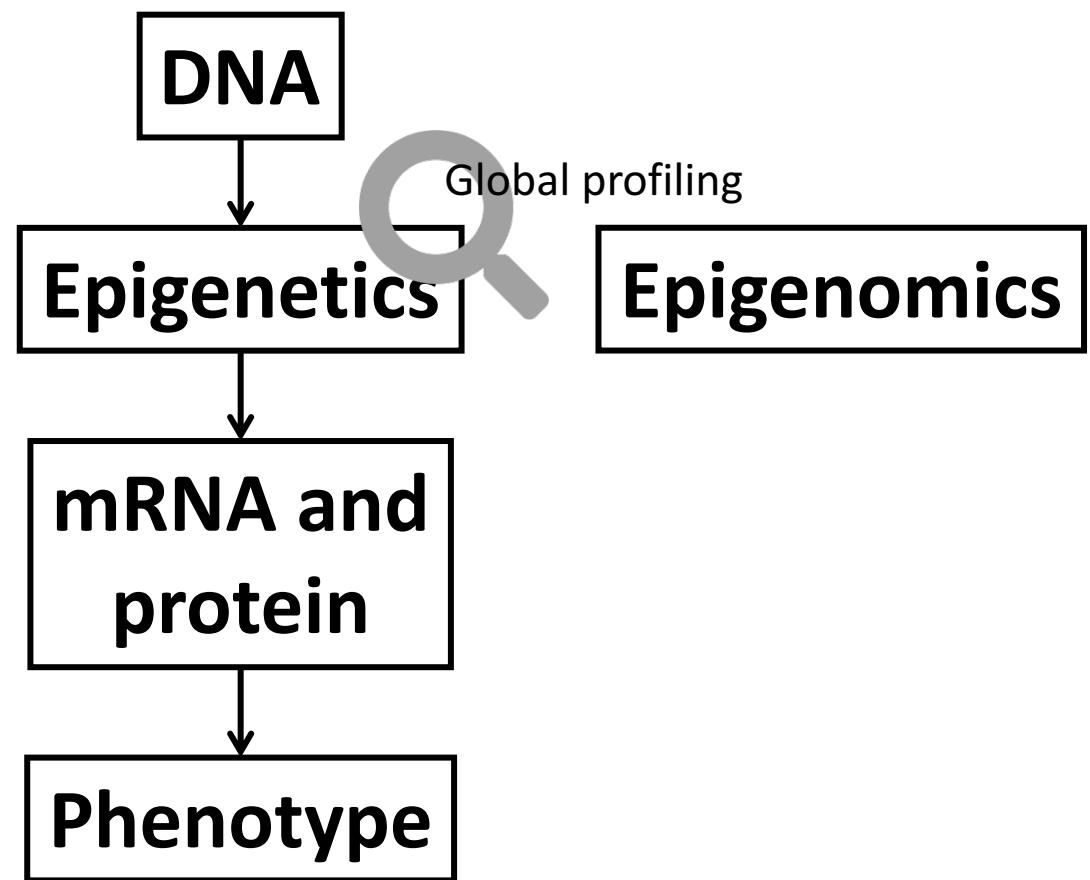
What is epigenetics?

Epigenetics

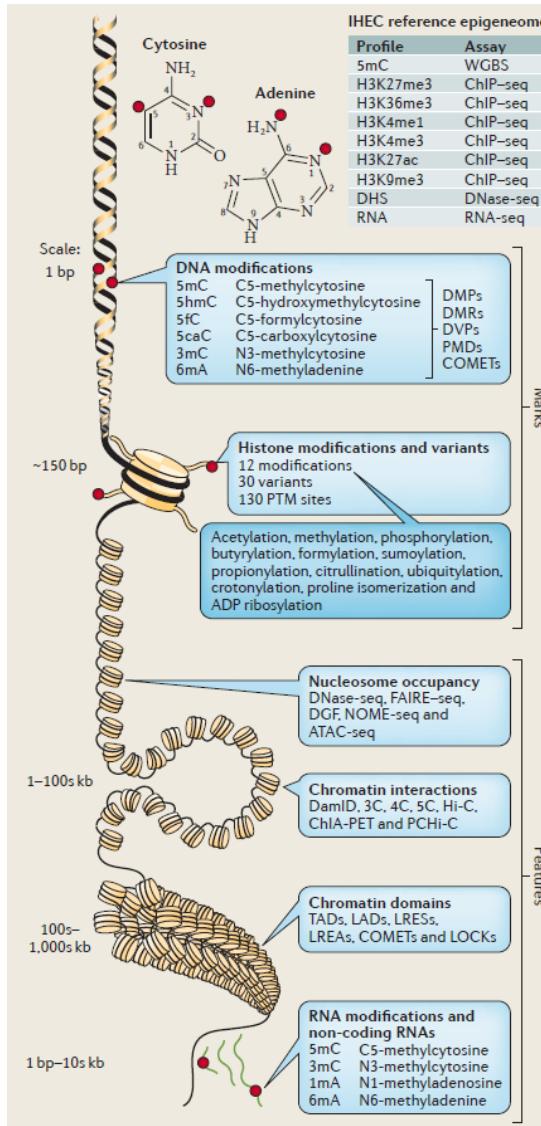
Computer
Science



Biology

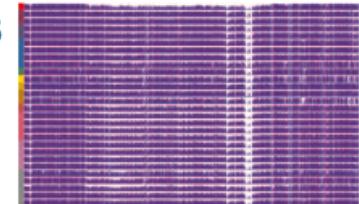


Epigenetics → epigenomics



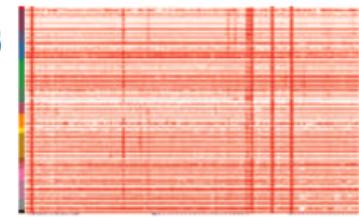
DNA methylation

WGBS



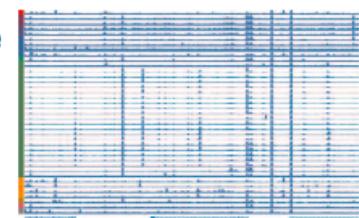
Histone modifications

H3K4me3



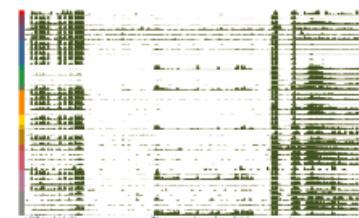
Chromatin features

DNase

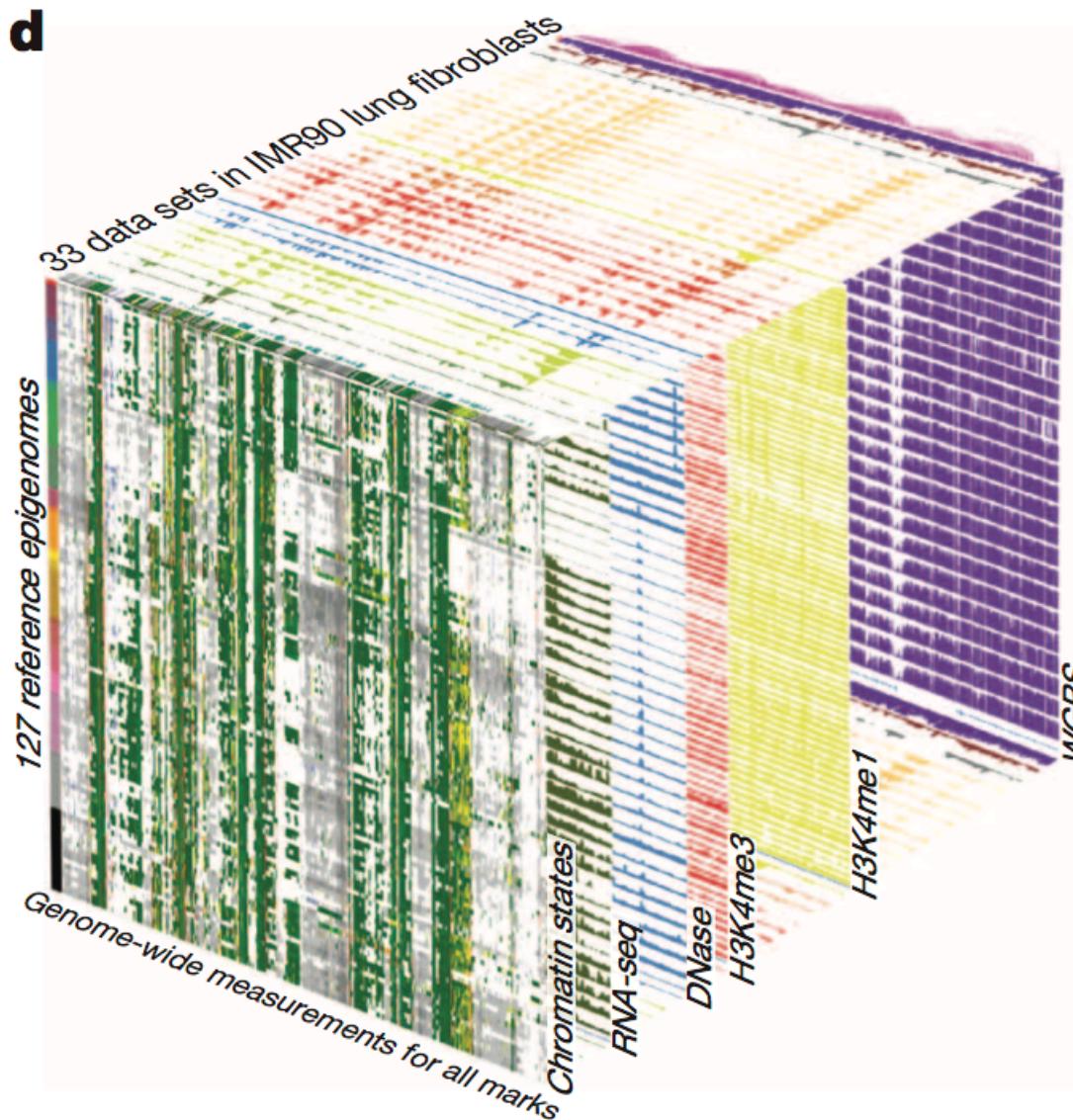


Non-coding RNA

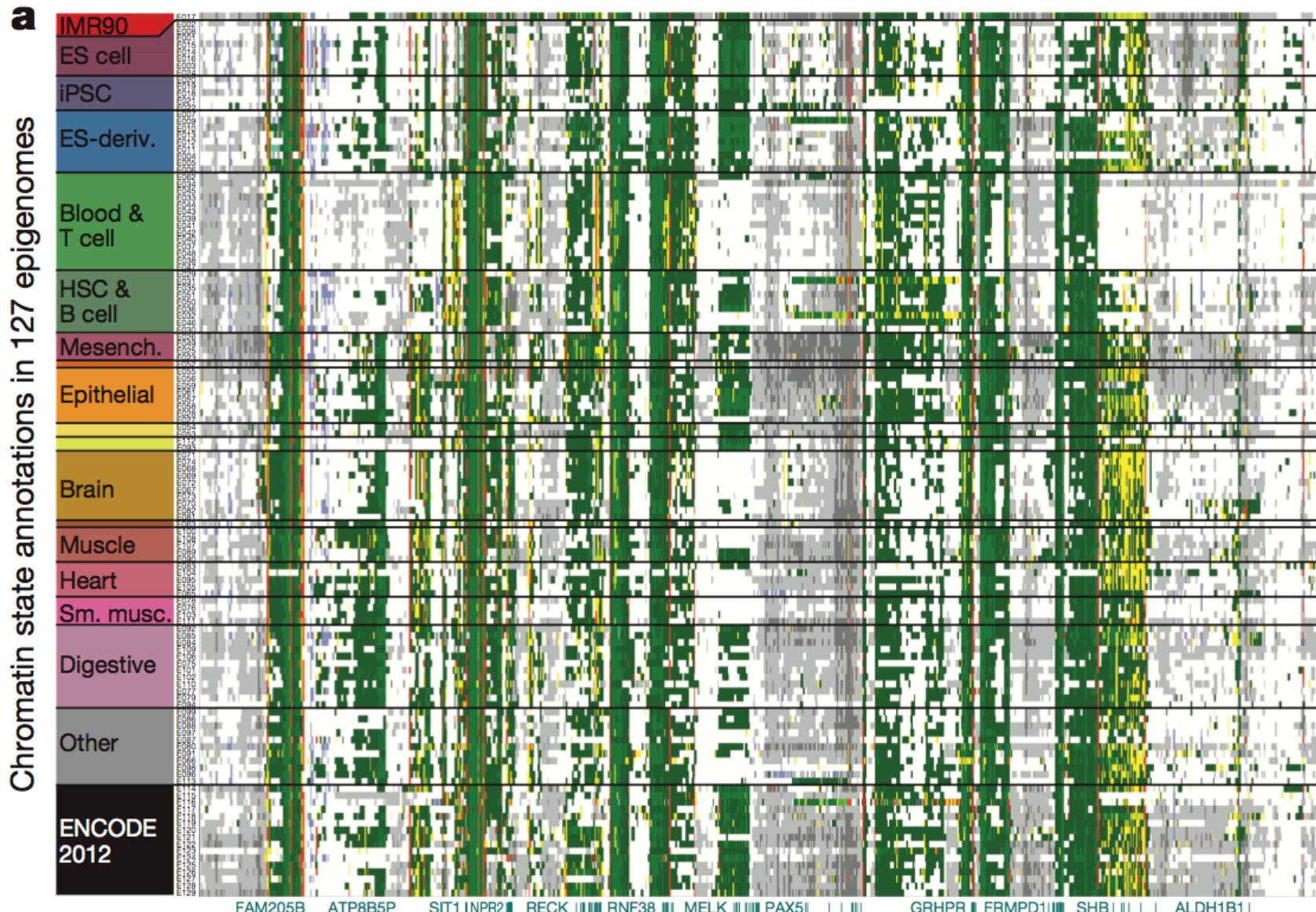
RNA-seq



Integrated into cell specific epigenomes
that govern the function, state, and fate of cells



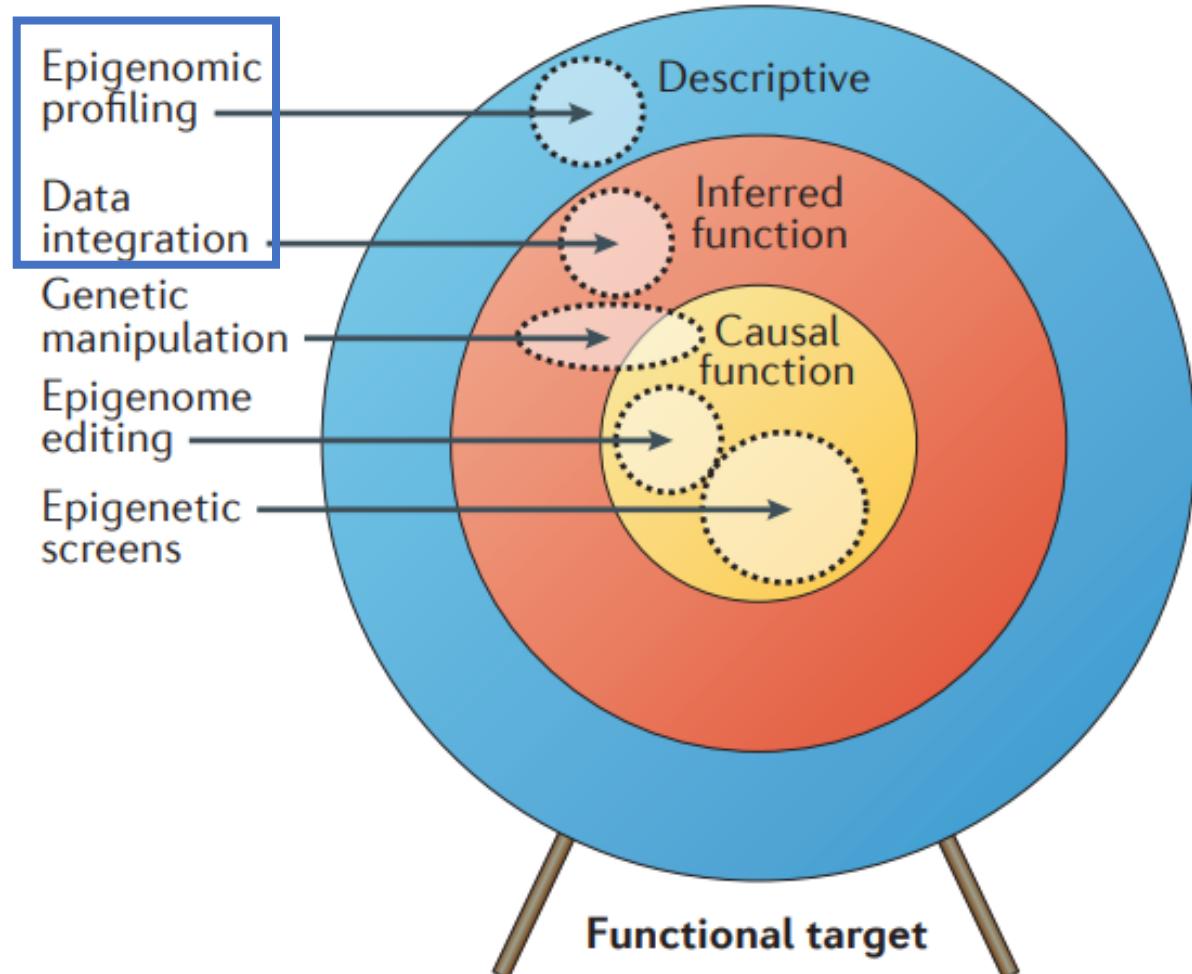
Roadmap Epigenomics Consortium



Why epigenomics?

Why epigenomics?

Hypothesis generating

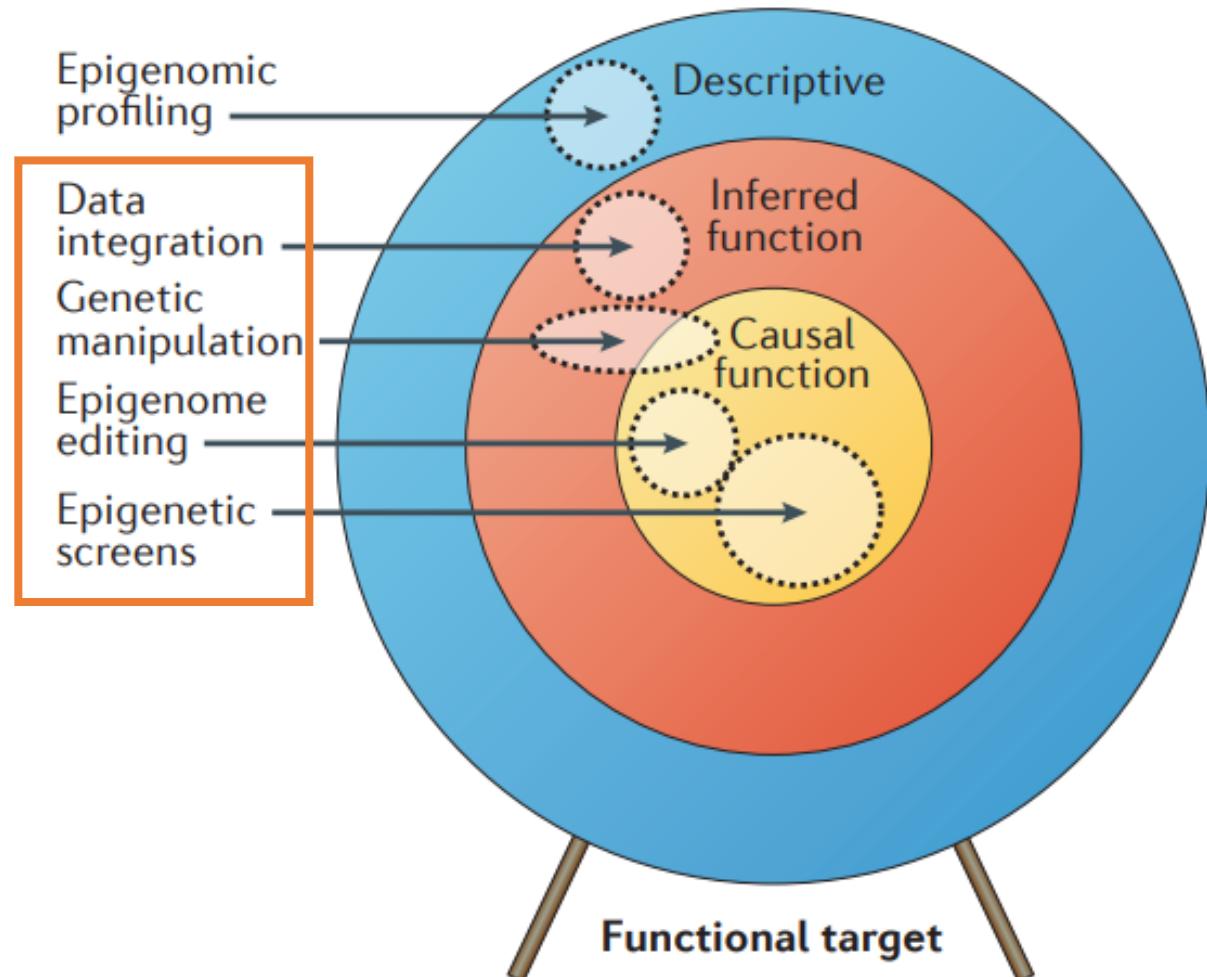


Hypothesis generating experiments

- Identify loci of interest with differential epigenetic features
 - Disease vs. normal
 - Cell type specific epigenetic marks
- Biomarker discovery
- Identify DNA-binding proteins or other factors for further study
 - Within topologically associated domains (TADs), nucleosome occupancy sites, etc.
 - Motifs
- Infer functional relationships through integrative analyses
 - Chromatin state in relation with DNA methylation
 - DNA methylation and effect on the transcriptome
 - Nucleosome occupancy interplay with DNA variation

Why epigenomics?

Hypothesis testing



Hypothesis testing experiments

EXPERIMENT:

Coupled with experimentation or manipulation of sample of interest with epigenomics methods

- Gene overexpression/knockdown
- Treatment with drug
- Introduce mutation

MEASURE:

- Global epigenetic changes
 - Global changes in DNA methylation
 - Alteration in nucleosome occupancy profile
- Site specific epigenetic changes
 - Promoters or enhancers
 - Motifs of interest

Case 01: Identification of differentially methylated loci and biologically relevant clusters in breast cancer

EMBO
Molecular Medicine

Research Article
Epigenetic portraits of human breast cancers

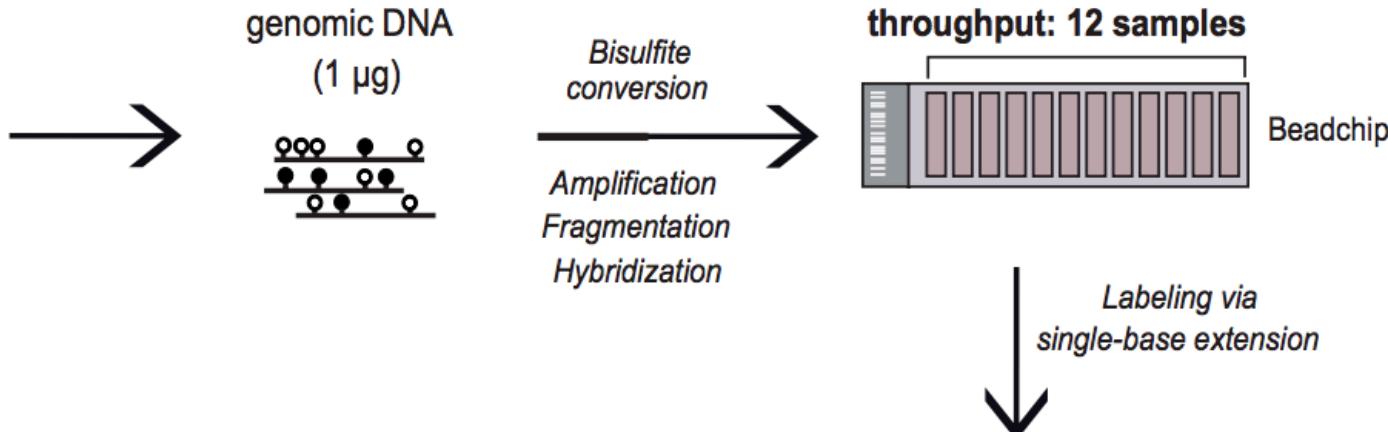
DNA methylation profiling reveals
a predominant immune component
in breast cancers

Sarah Dedeurwaerder^{1†}, Christine Desmedt^{2†}, Emilie Calonne¹, Sandeep K. Singhal²,
Benjamin Haibe-Kains^{2,3}, Matthieu Defrance¹, Stefan Michiels², Michael Volkmar¹, Rachel Deplus¹,
Judith Luciani¹, Françoise Lallemand², Denis Larsimont⁴, Jérôme Toussaint², Sandy Haussy²,
Françoise Rothé², Ghizlane Rouas², Otto Metzger², Samira Majjaj², Kamal Saini², Pascale Putmans¹,
Gérald Hames⁵, Nicolas van Baren⁶, Pierre G. Coulie⁵, Martine Piccart⁷,
Christos Sotiriou^{2***,†}, François Fuks^{1*,†}

Overview

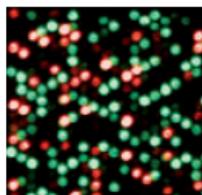
MAIN SET OF PATIENTS:

123 breast tissues

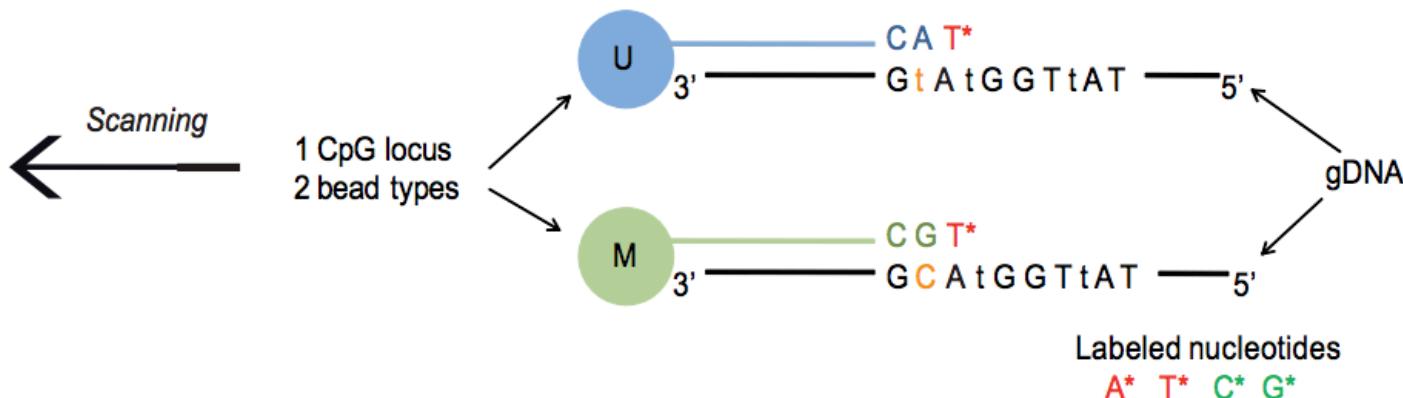


VALIDATION SET OF PATIENTS:

125 breast tissues



14,475 genes
(27,578 CpGs)
investigated per sample

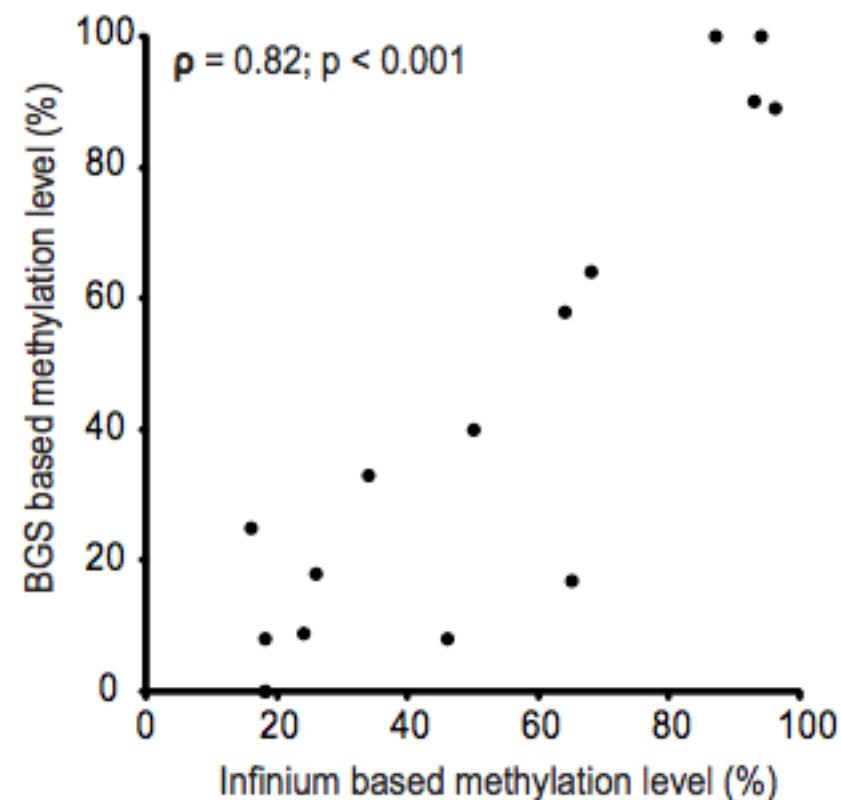


Identification of differentially methylated loci

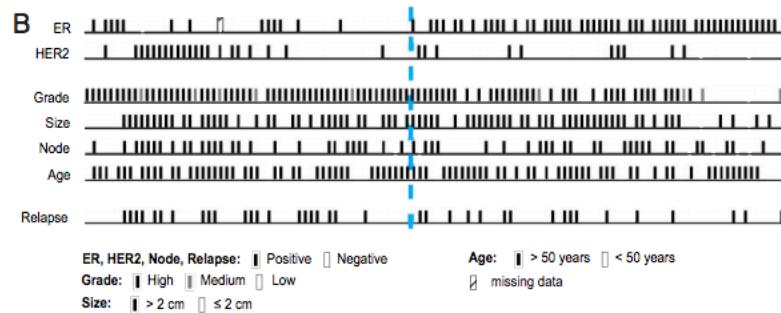
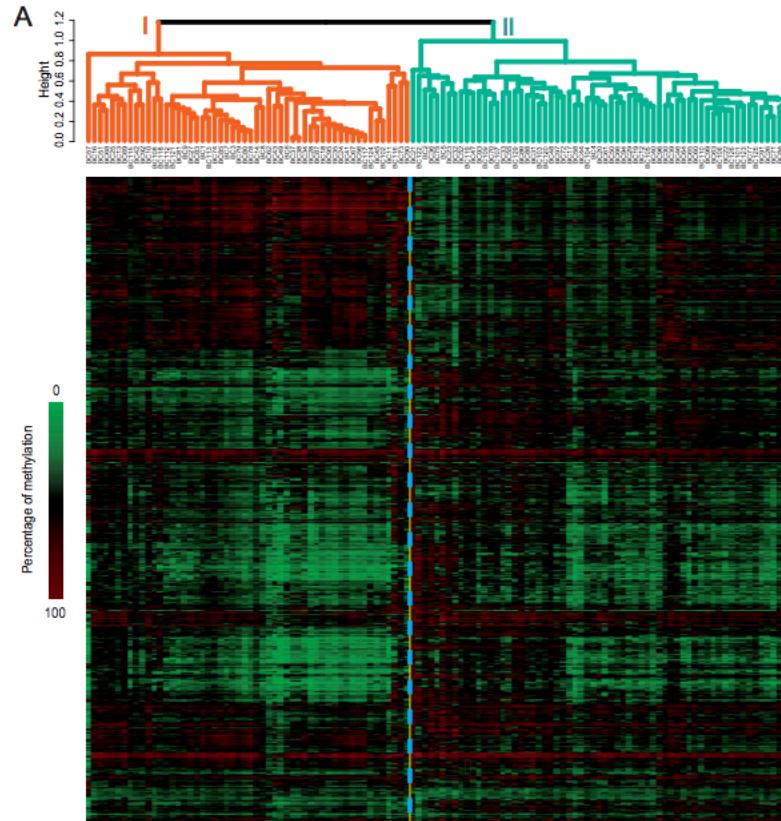
Gene	Illumina ID	Infinium methylation (frequency, %)	Reported methylation data (frequency, %; technique)*	Correlation Infinium vs. reported methylation data
<i>RASSF1A</i>	cg00777121	71	70; MSP / 56; MSP / 56; MPS	++
	cg08047457	72	65; MSP	++
	cg21554552	70	65; MSP	++
<i>CCND2</i>	cg25425078	9	46; MSP / 28; MSP / 55; MSP	+
<i>APC</i>	cg16970232	39	45; MSP / 28; MSP / 39; MSP / 49; MSP	++
	cg20311501	35	45; MSP / 28; MSP / 39; MSP / 49; MSP	++
<i>RARβ2</i>	cg27486427	12	17; PS / 0; PS	++
	cg26124016	4	23; MSP	+
<i>CDH13</i>	cg08747377	17	33; MSP	++
<i>SDHB</i>	cg24305835	0	0; MS-HRM	++
	cg03861428	0	0; MS-HRM	++
<i>FH</i>	cg06806184	0	0; MS-HRM	++

Identification of differentially methylated loci

Gene	Illumina ID	Infinium methylation (frequency, %)	Reported methylation data (frequency, %; technique)*	Correlation Infinium vs. reported methylation data
<i>RASSF1A</i>	cg00777121	71	70; MS	
	cg08047457	72	65; MS	
	cg21554552	70	65; MS	
<i>CCND2</i>	cg25425078	9	46; MS	
<i>APC</i>	cg16970232	39	45; MS	
	cg20311501	35	45; MS	
<i>RARβ2</i>	cg27486427	12	17; PS	
	cg26124016	4	23; MS	
<i>CDH13</i>	cg08747377	17	33; MS	
<i>SDHB</i>	cg24305835	0	0; MS-	
	cg03861428	0	0; MS-	
<i>FH</i>	cg06806184	0	0; MS-	



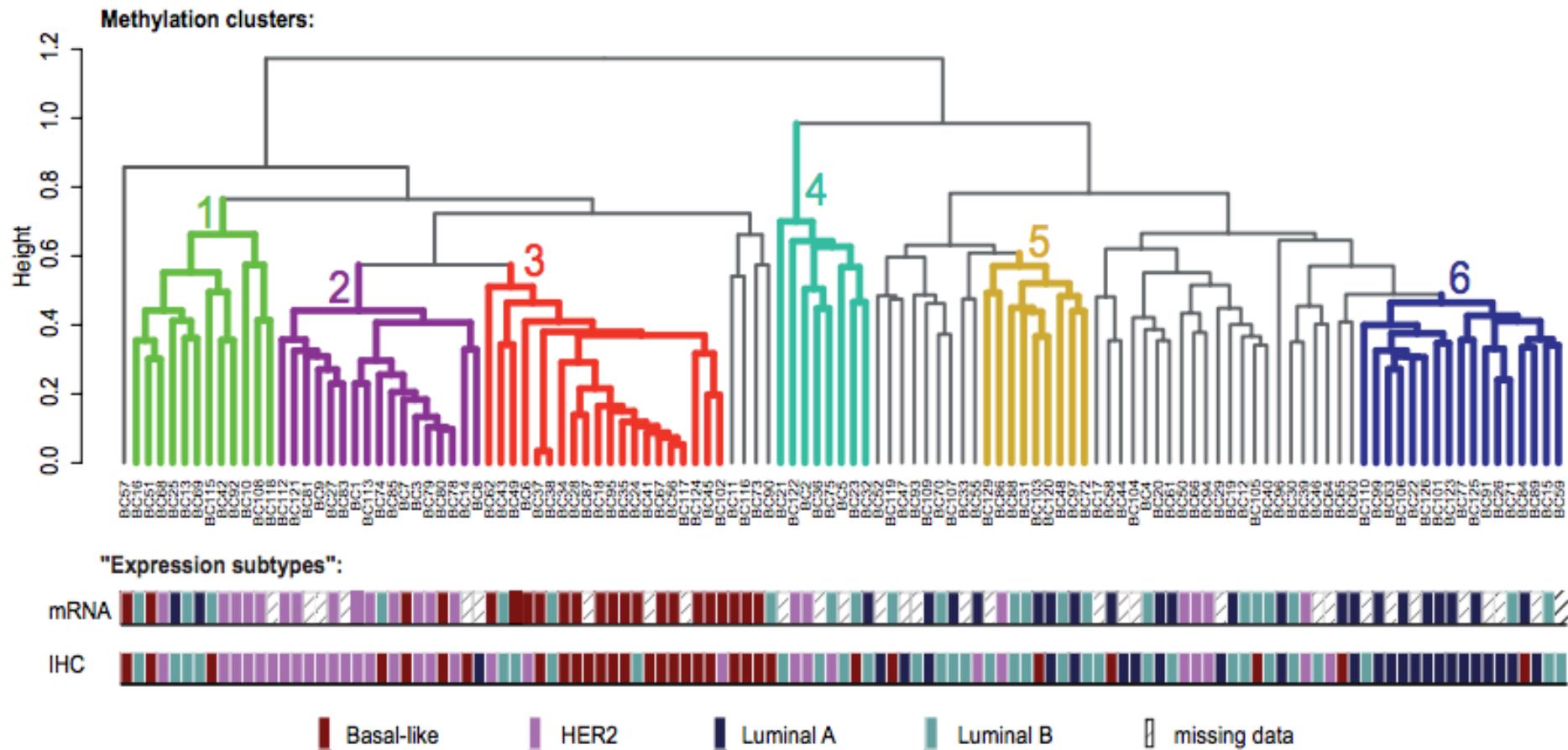
Unsupervised clustering identifies primary clusters by ER status



Six stable, biologically relevant clusters

A

UNSUPERVISED CLUSTERING ON THE MAIN SET

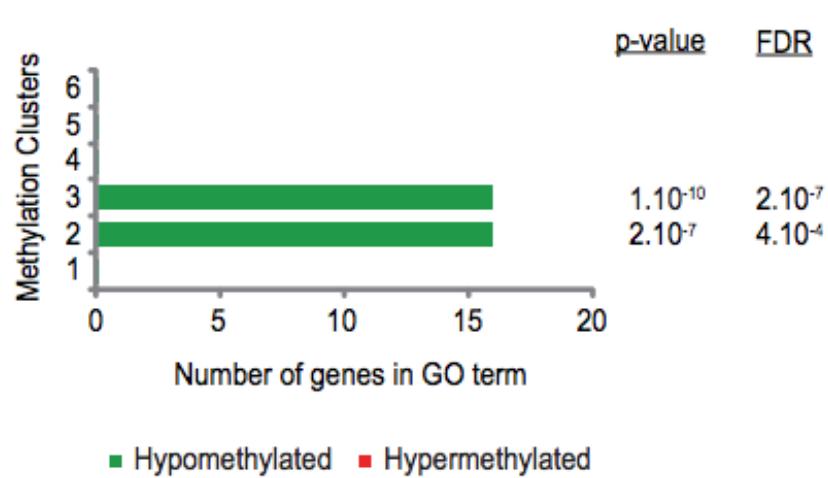


Validated the presence of these clusters in an independent data set

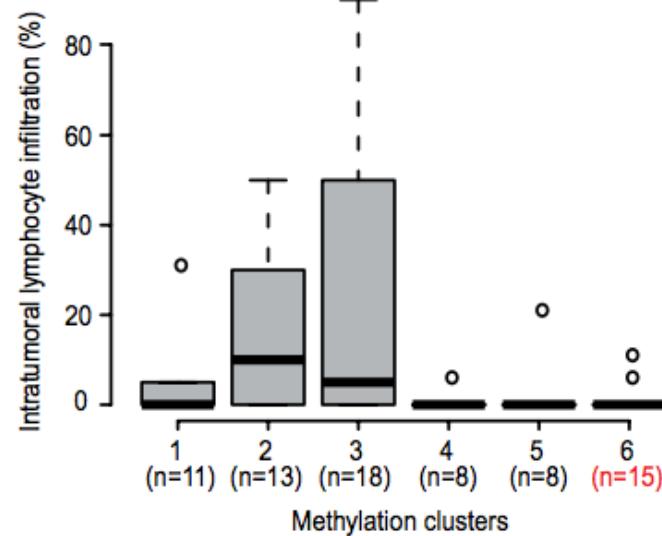
Identification of immune-related clusters

B

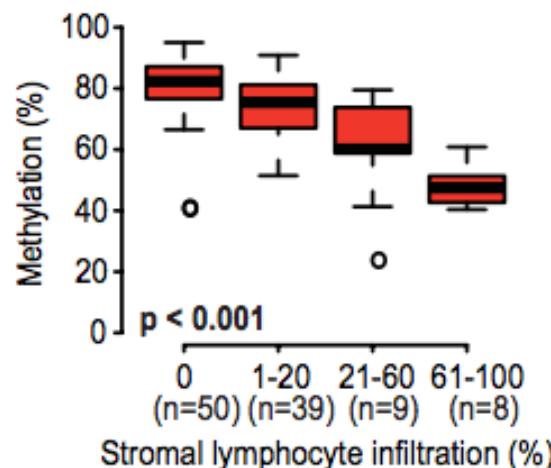
Immune system process



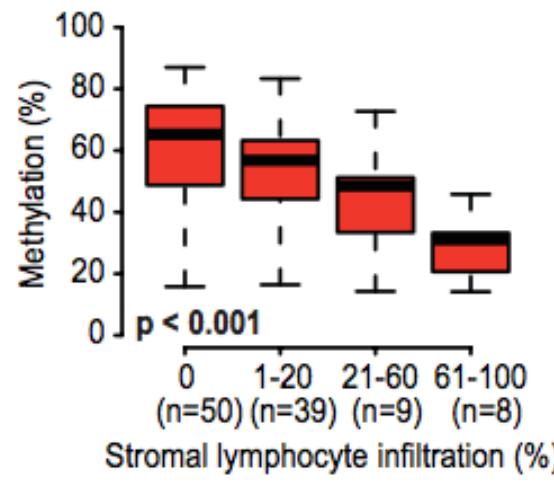
Intratumoral lymphocyte infiltration



LCK



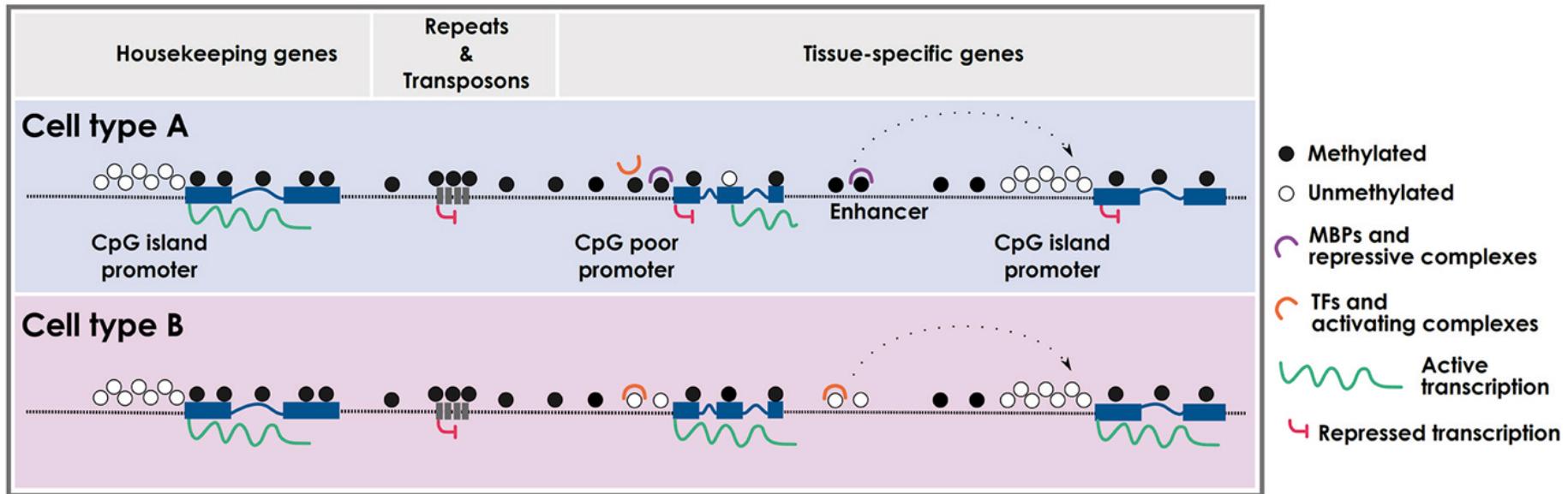
ITGAL



Epigenomic methods

DNA Methylation

“Typical” 5mC DNA methylation in mammalian genome



- CpG island promoter and enhancer methylation is inversely proportional to gene expression
- Gene body methylation is correlated with gene expression of highly transcribed genes
- CpG methylation is typical in repeats and transposons
- Changes in these patterns, both globally and loci-specific, have been observed in disease progression

a Assays for DNA methylation mapping

Bisulphite sequencing

DNA treatment with bisulphite specifically introduces mutations at unmethylated Cs. These mutations are mapped by next-generation sequencing

Bisulphite microarrays

DNA-methylation-specific mutations are introduced by bisulphite treatment. These mutations are mapped using a genotyping microarray that covers a selection of Cs

Enrichment-based methods

Methylated (alternatively, unmethylated) DNA fragments are enriched in a DNA library. The library composition is quantified by next-generation sequencing

Unprocessed DNA sequencing or microarray data (assay-specific)

b Data processing and quality control

Processing bisulphite-sequencing data

- Bisulphite sequence alignment
- Quantification of absolute DNA methylation at single-base resolution
- Quality control

Processing bisulphite microarray data

- Data normalization
- Quantification of absolute DNA methylation at single-base resolution
- Quality control

Processing enrichment-based data

- DNA sequence alignment
- Quantification of relative enrichment
- Statistical inference of absolute DNA methylation corrected for CpG density
- Quality control

Table with DNA methylation levels for each CpG in each sample (assay-independent)

c Data visualization and statistical analysis

Visualizing DNA methylation data

- Visual inspection of selected regions in a genome browser
- Global visualization of the distribution of DNA methylation
- Clustering-based assessment of global similarity and differences in a set of samples

Identifying differentially methylated regions

- Statistical testing for differential DNA methylation at single CpGs and/or larger genomic regions
- Statistical correction for multiple hypothesis testing
- Ranking based on statistical significance and effect size

List of DMRs that are statistically significant

d Validation and interpretation

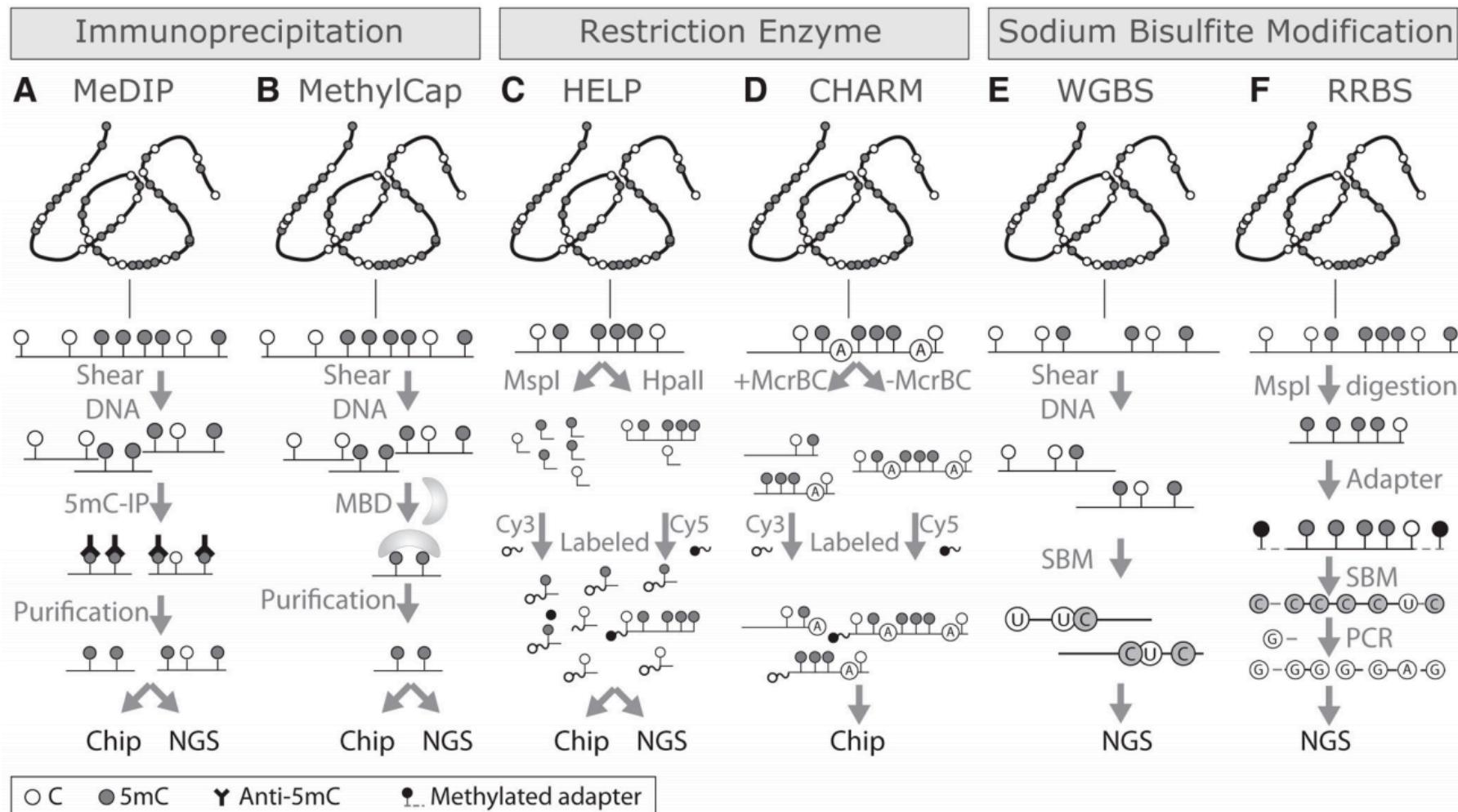
Verifying and validating differences in DNA methylation

- Global analysis of DMR list: volcano plots, Q-Q plots, Manhattan plots
- Manual or computational ranking and selection of promising DMRs for experimental verification and/or validation
- Computational design of high-throughput assays for confirming the sensitivity and specificity of DMR identification in large sample cohorts

Interpreting differences in DNA methylation

- Integrative analysis in the context of other genomic data sets
- Search for significant enrichment of gene functions and regulatory elements among the DMRs
- Statistical assessment of confounding factors to assess whether it would be plausible to hypothesize causal effects

Epigenomics methods for studying 5-mC DNA methylation



a Assays for DNA methylation mapping

Bisulphite sequencing

DNA treatment with bisulphite specifically introduces mutations at unmethylated Cs. These mutations are mapped by next-generation sequencing

Bisulphite microarrays

DNA-methylation-specific mutations are introduced by bisulphite treatment. These mutations are mapped using a genotyping microarray that covers a selection of Cs

Enrichment-based methods

Methylated (alternatively, unmethylated) DNA fragments are enriched in a DNA library. The library composition is quantified by next-generation sequencing

↓
Unprocessed DNA sequencing or
microarray data (assay-specific)

b Data processing and quality control

Processing bisulphite-sequencing data

- Bisulphite sequence alignment
- Quantification of absolute DNA methylation at single-base resolution
- Quality control

Processing bisulphite microarray data

- Data normalization
- Quantification of absolute DNA methylation at single-base resolution
- Quality control

Processing enrichment-based data

- DNA sequence alignment
- Quantification of relative enrichment
- Statistical inference of absolute DNA methylation corrected for CpG density
- Quality control

↓
Table with DNA methylation levels for each
CpG in each sample (assay-independent)

c Data visualization and statistical analysis

Visualizing DNA methylation data

- Visual inspection of selected regions in a genome browser
- Global visualization of the distribution of DNA methylation
- Clustering-based assessment of global similarity and differences in a set of samples

Identifying differentially methylated regions

- Statistical testing for differential DNA methylation at single CpGs and/or larger genomic regions
- Statistical correction for multiple hypothesis testing
- Ranking based on statistical significance and effect size

↓
List of DMRs that are statistically significant

d Validation and interpretation

Verifying and validating differences in DNA methylation

- Global analysis of DMR list: volcano plots, Q-Q plots, Manhattan plots
- Manual or computational ranking and selection of promising DMRs for experimental verification and/or validation
- Computational design of high-throughput assays for confirming the sensitivity and specificity of DMR identification in large sample cohorts

Interpreting differences in DNA methylation

- Integrative analysis in the context of other genomic data sets
- Search for significant enrichment of gene functions and regulatory elements among the DMRs
- Statistical assessment of confounding factors to assess whether it would be plausible to hypothesize causal effects

Goal: Identify differentially methylated probe/site

Linear model

Observed signal for probe p

Coefficients for probe p

$$E(Y_p) = X\beta_p$$

Design matrix

$$\sim \beta_D + \beta_1 + \cdots + \beta_n + \epsilon$$

Identifier	Disease	Gender	...	HRT
Case01	1	1	...	1
Case02	1	0	...	0
Ctrl01	0	1	...	0
Ctrl02	0	0	...	1

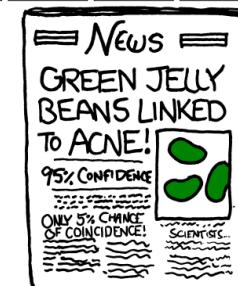
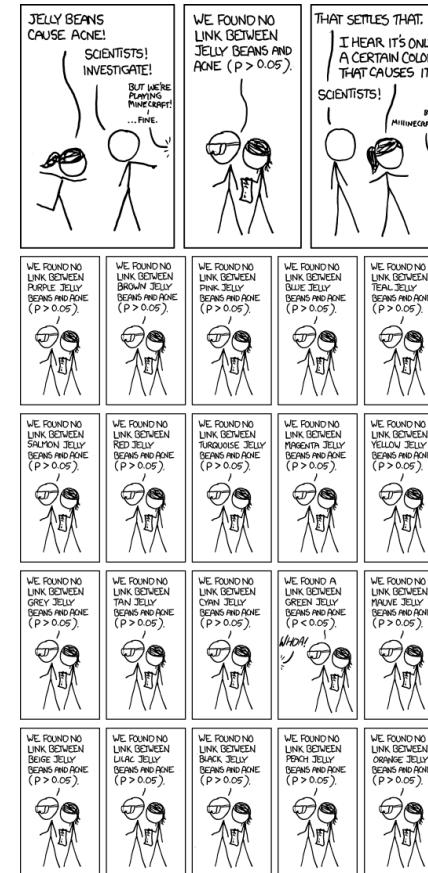
Multiple hypothesis correction

$$E(Y_1) = X\beta_1$$

...

$$E(Y_n) = X\beta_n$$

n = 450,000



```
library(ggplot2)

## Seed for reproducibility
set.seed(0)

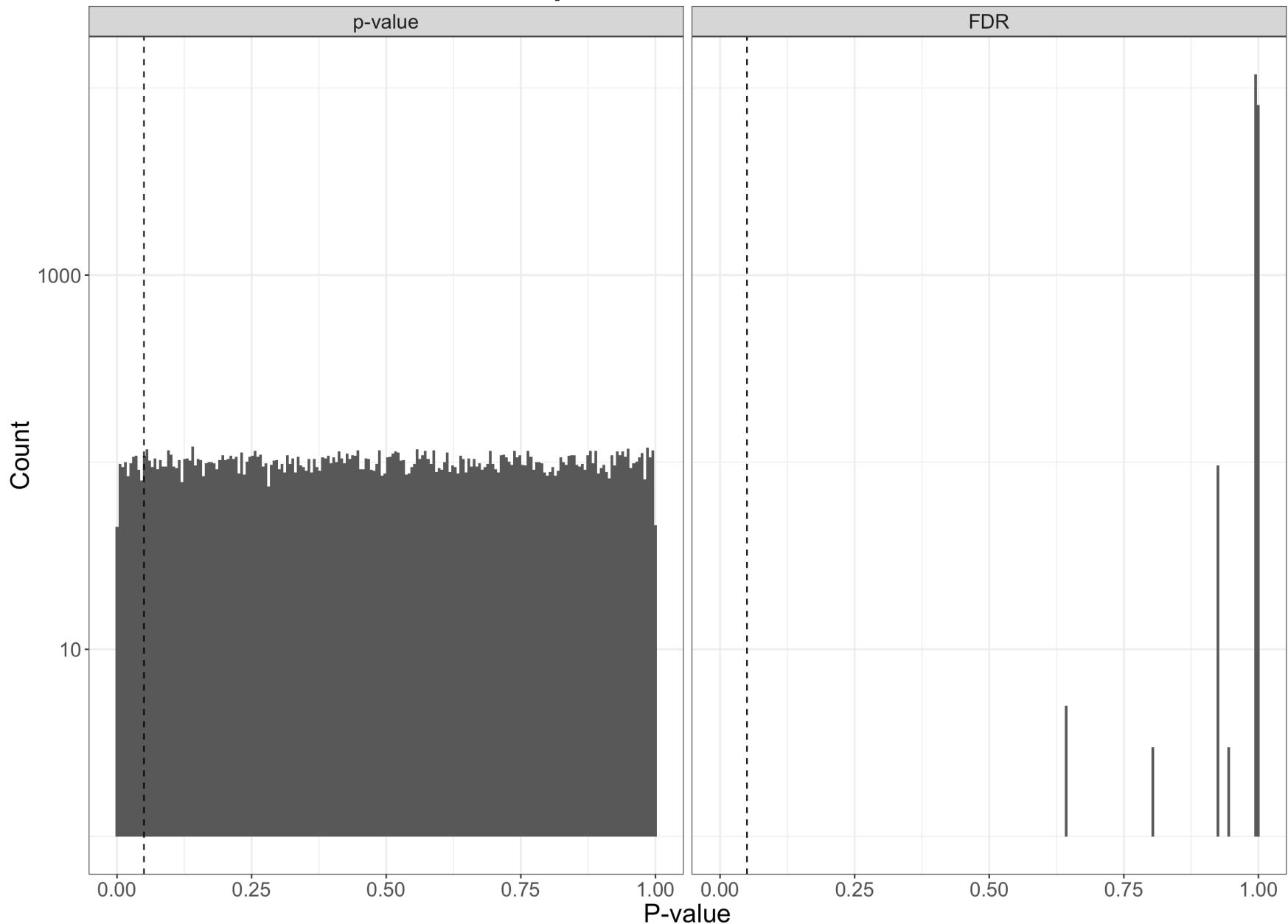
## Permutation function
permute <- function(x){
  ## 100 samples each
  a <- rnorm(100)
  b <- rnorm(100)
  ans <- t.test(a,b)
  return(ans$p.value)
}

## Run 20,000 permutations
p <- apply(X = 1:20000, permute)
mean(p <= 0.05)
#> [1] 0.0477

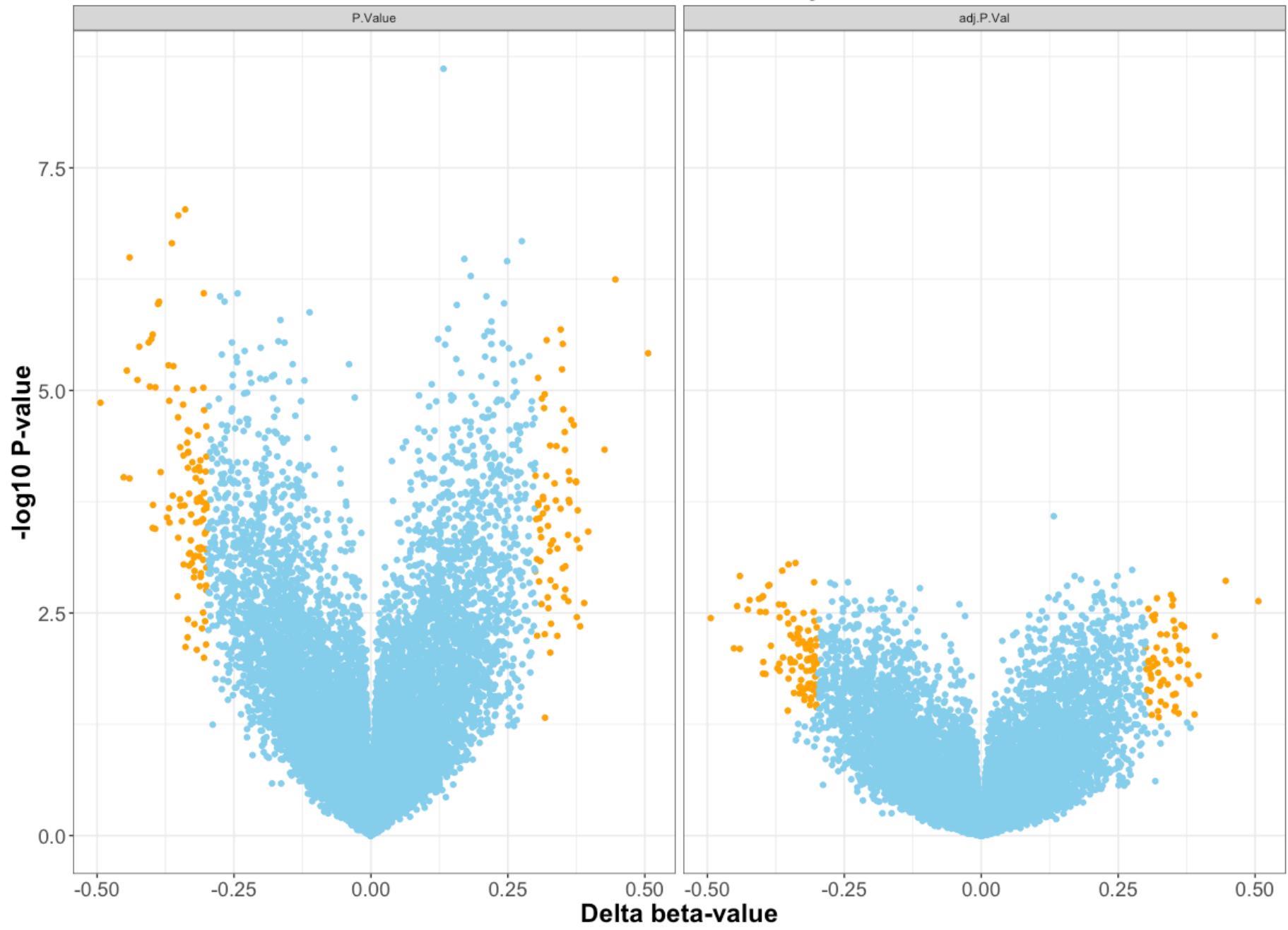
## P-value adjustment
q <- p.adjust(p, method = 'fdr')
mean(q <= 0.05)
#> [1] 0

pq <- rbind(data.frame(p = 'p-value', val = as.numeric(p)),
             data.frame(p = 'FDR', val= as.numeric(q)))
ggplot(pq, aes(x = val)) +
  geom_histogram(bins = 200) +
  facet_grid(. ~ p) +
  scale_y_log10() +
  geom_vline(xintercept = 0.05, lty = 2) +
  ylab('Count') + xlab('P-value') +
  ggtitle('20,000 permutation of H0 = True') +
  theme_bw() +
  theme(plot.title = element_text(hjust = 0.5,
                                  face = 2, size = 22),
        strip.text = element_text(size = 14),
        axis.title = element_text(size = 18),
        axis.text = element_text(size = 14))
#> Warning: Transformation introduced infinite values in continuous y-axis
#> Warning: Removed 192 rows containing missing values (geom_bar).
```

20,000 permutation of H0 = True



P-values before and after adjustment

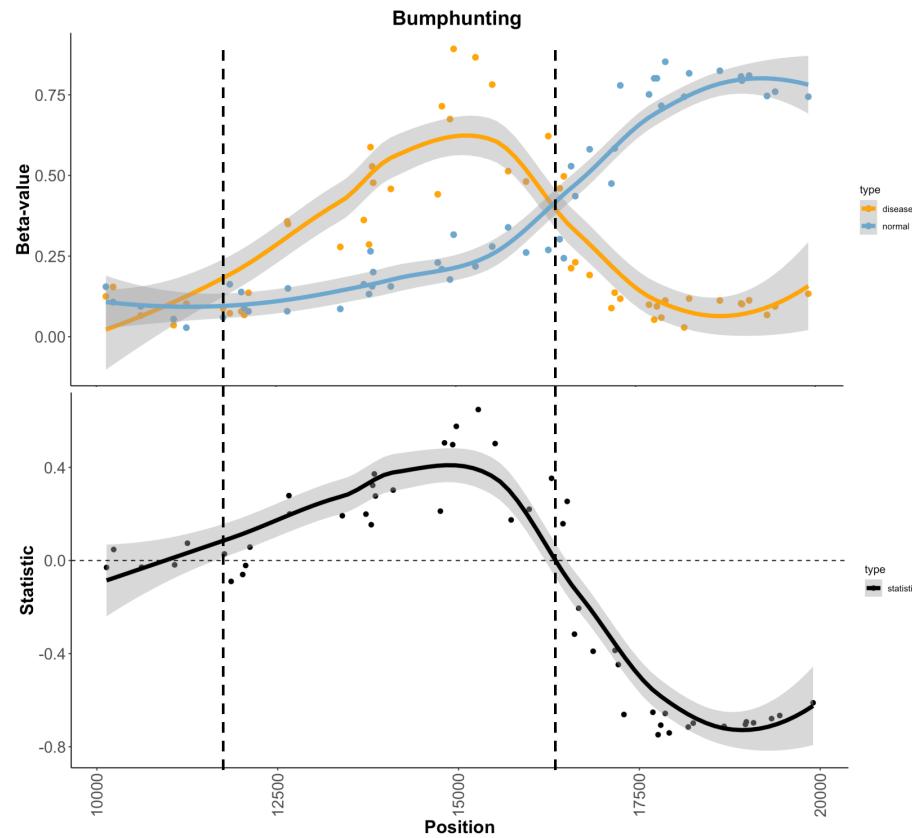


Differentially methylated regions (DMRs)

- Biologically more relevant
- More methylation status independence between DMRs compared to individual CpGs
- Generate significance against permuted null

Software:

- minfi::bumphunter
- DMRcate
- CHAMP::probe lasso
- comb-p



a Assays for DNA methylation mapping

Bisulphite sequencing

DNA treatment with bisulphite specifically introduces mutations at unmethylated Cs. These mutations are mapped by next-generation sequencing

Bisulphite microarrays

DNA-methylation-specific mutations are introduced by bisulphite treatment. These mutations are mapped using a genotyping microarray that covers a selection of Cs

Enrichment-based methods

Methylated (alternatively, unmethylated) DNA fragments are enriched in a DNA library. The library composition is quantified by next-generation sequencing

↓
Unprocessed DNA sequencing or
microarray data (assay-specific)

b Data processing and quality control

Processing bisulphite-sequencing data

- Bisulphite sequence alignment
- Quantification of absolute DNA methylation at single-base resolution
- Quality control

Processing bisulphite microarray data

- Data normalization
- Quantification of absolute DNA methylation at single-base resolution
- Quality control

Processing enrichment-based data

- DNA sequence alignment
- Quantification of relative enrichment
- Statistical inference of absolute DNA methylation corrected for CpG density
- Quality control

↓
Table with DNA methylation levels for each
CpG in each sample (assay-independent)

c Data visualization and statistical analysis

Visualizing DNA methylation data

- Visual inspection of selected regions in a genome browser
- Global visualization of the distribution of DNA methylation
- Clustering-based assessment of global similarity and differences in a set of samples

Identifying differentially methylated regions

- Statistical testing for differential DNA methylation at single CpGs and/or larger genomic regions
- Statistical correction for multiple hypothesis testing
- Ranking based on statistical significance and effect size

↓
List of DMRs that are statistically significant

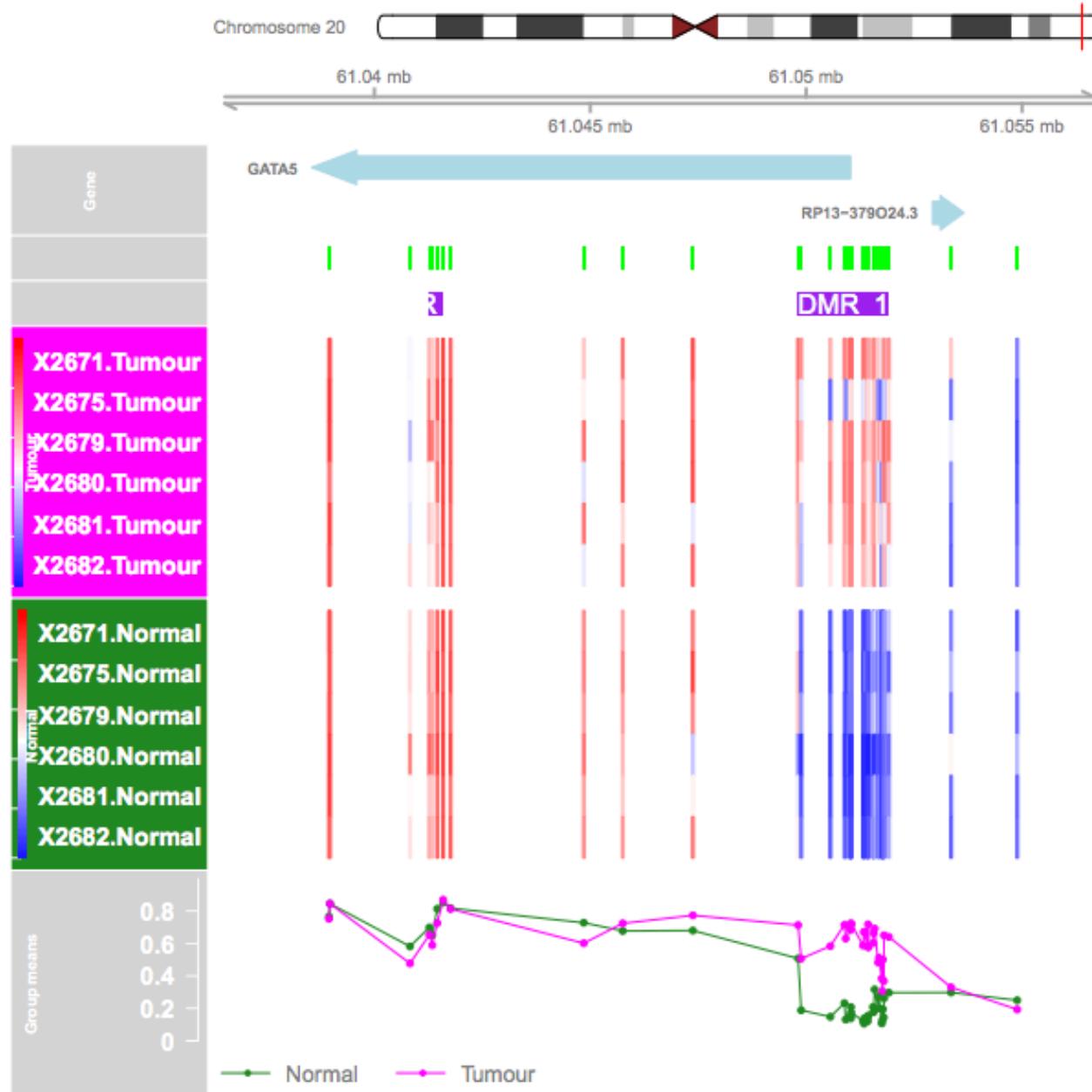
d Validation and interpretation

Verifying and validating differences in DNA methylation

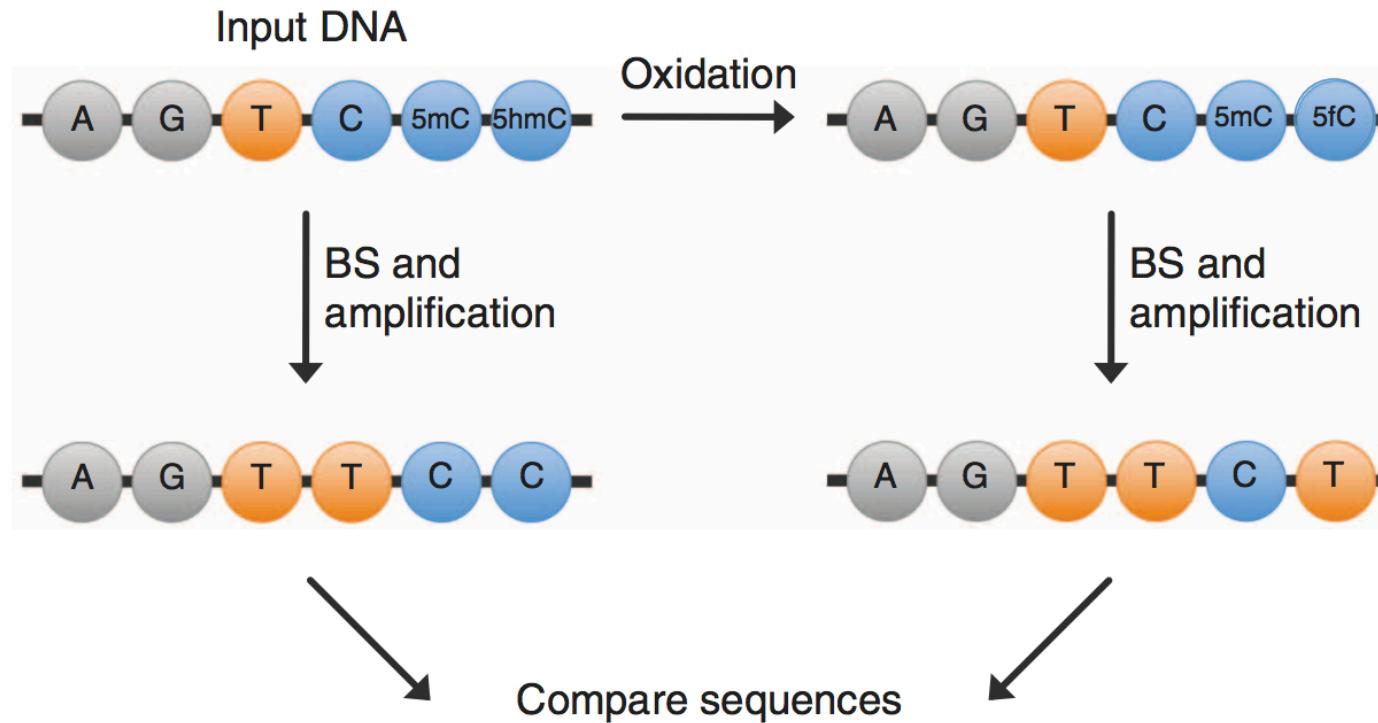
- Global analysis of DMR list: volcano plots, Q-Q plots, Manhattan plots
- Manual or computational ranking and selection of promising DMRs for experimental verification and/or validation
- Computational design of high-throughput assays for confirming the sensitivity and specificity of DMR identification in large sample cohorts

Interpreting differences in DNA methylation

- Integrative analysis in the context of other genomic data sets
- Search for significant enrichment of gene functions and regulatory elements among the DMRs
- Statistical assessment of confounding factors to assess whether it would be plausible to hypothesize causal effects

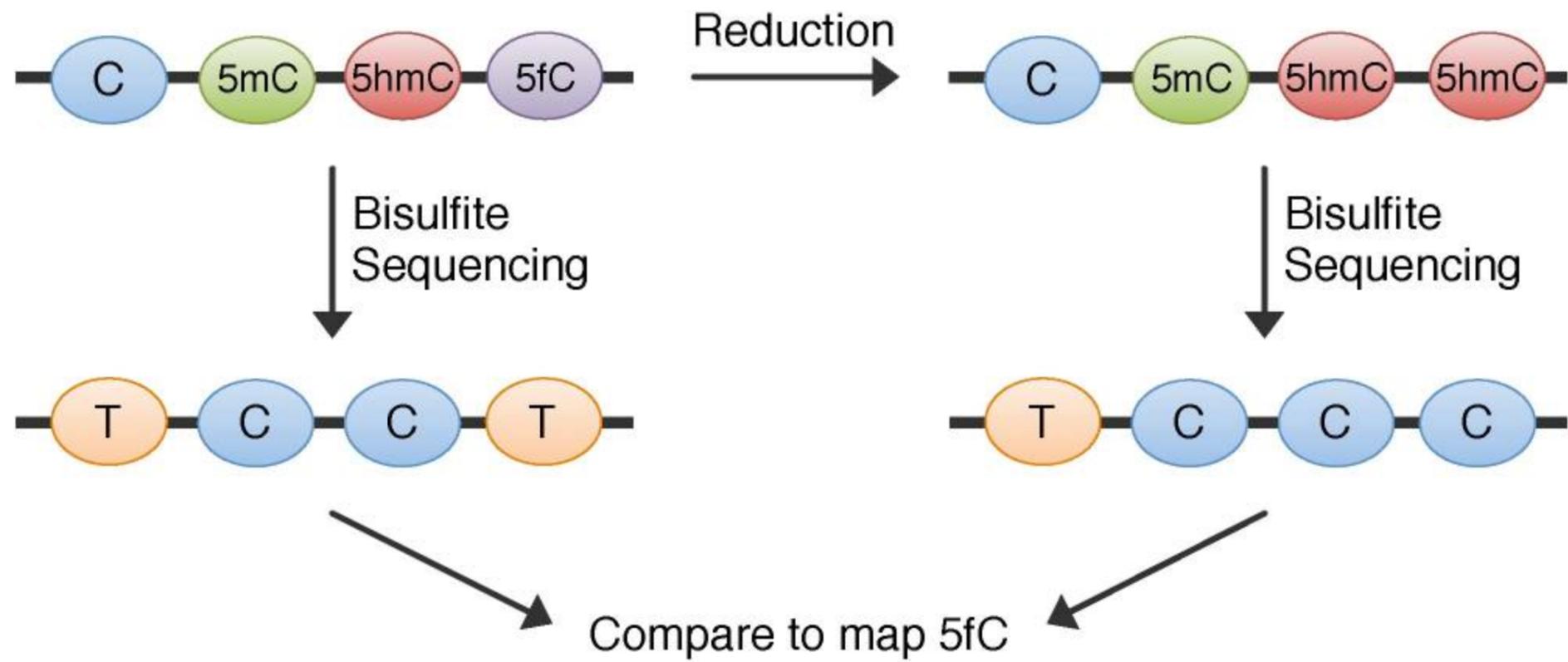


Detecting 5hmC: Oxidative BS-seq



Base	Sequence	BS Sequence	oxBS Sequence
C	C	T	T
5mC	C	C	C
5hmC	C	C	T

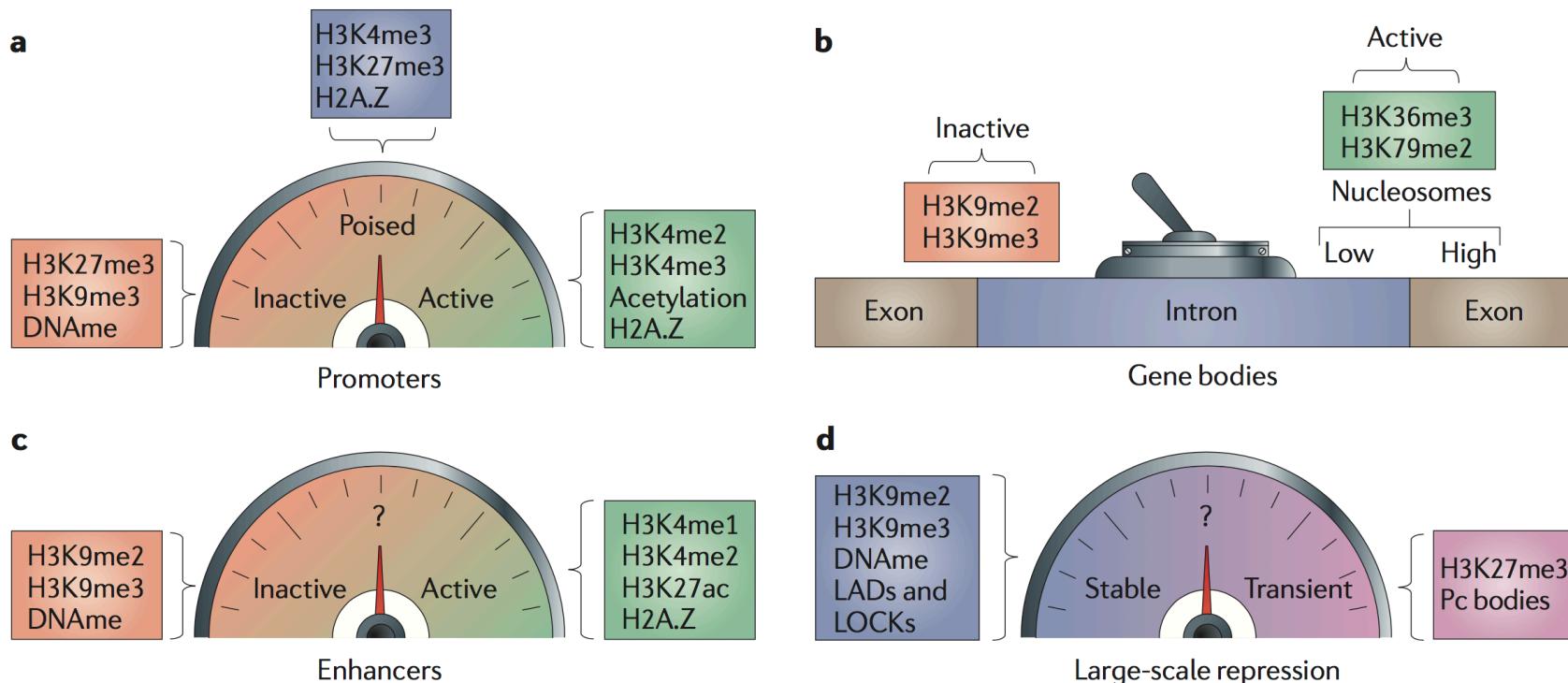
Detecting 5fC: Reduced BS-seq



Chromatin modifications

Chromatin marks: The histone code hypothesis

- Combinatorial histone post-translational modifications (PTMs) encrypt recruitment of different histone effectors
- These combinations define specific chromatin states
- These states are heritable

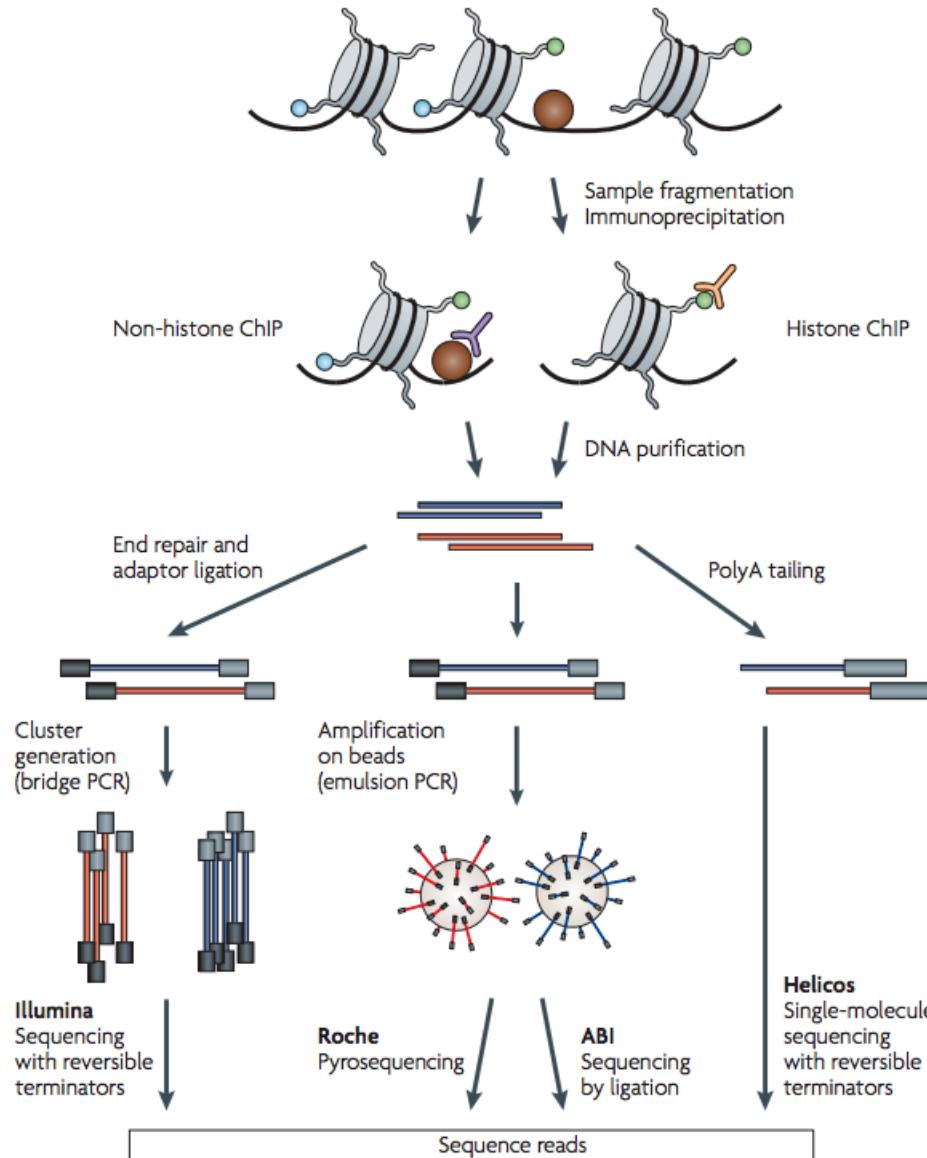


Strahl, B. D., & Allis, C. D. (2000). The language of covalent histone modifications. *Nature*, 403(6765), 41

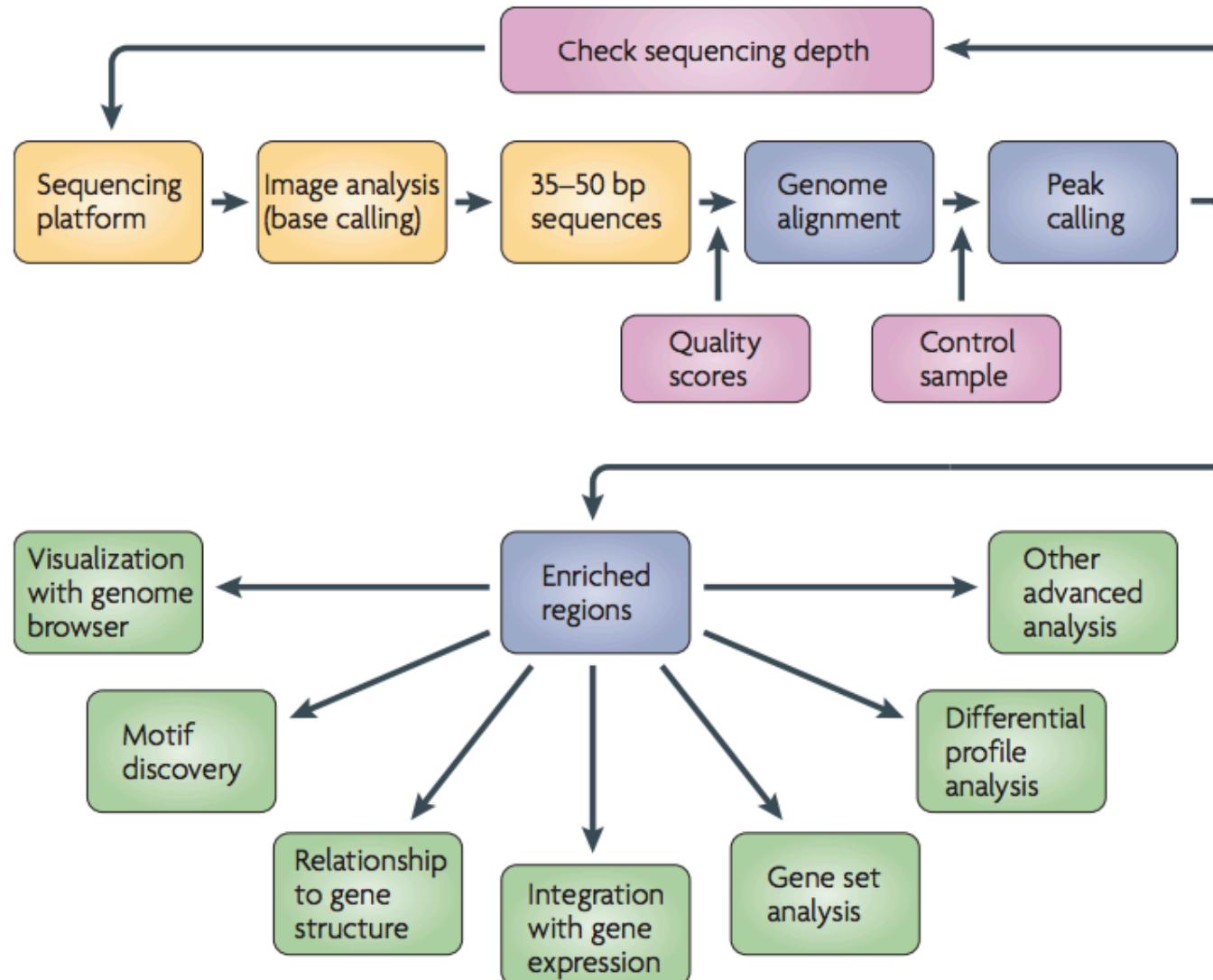
Turner, B. M. (2000). Histone acetylation and an epigenetic code. *Bioessays*, 22(9), 836-845

Zhou, V. W., Goren, A., & Bernstein, B. E. (2011). Charting histone modifications and the functional organization of mammalian genomes. *Nature Reviews Genetics*, 12(1), 7

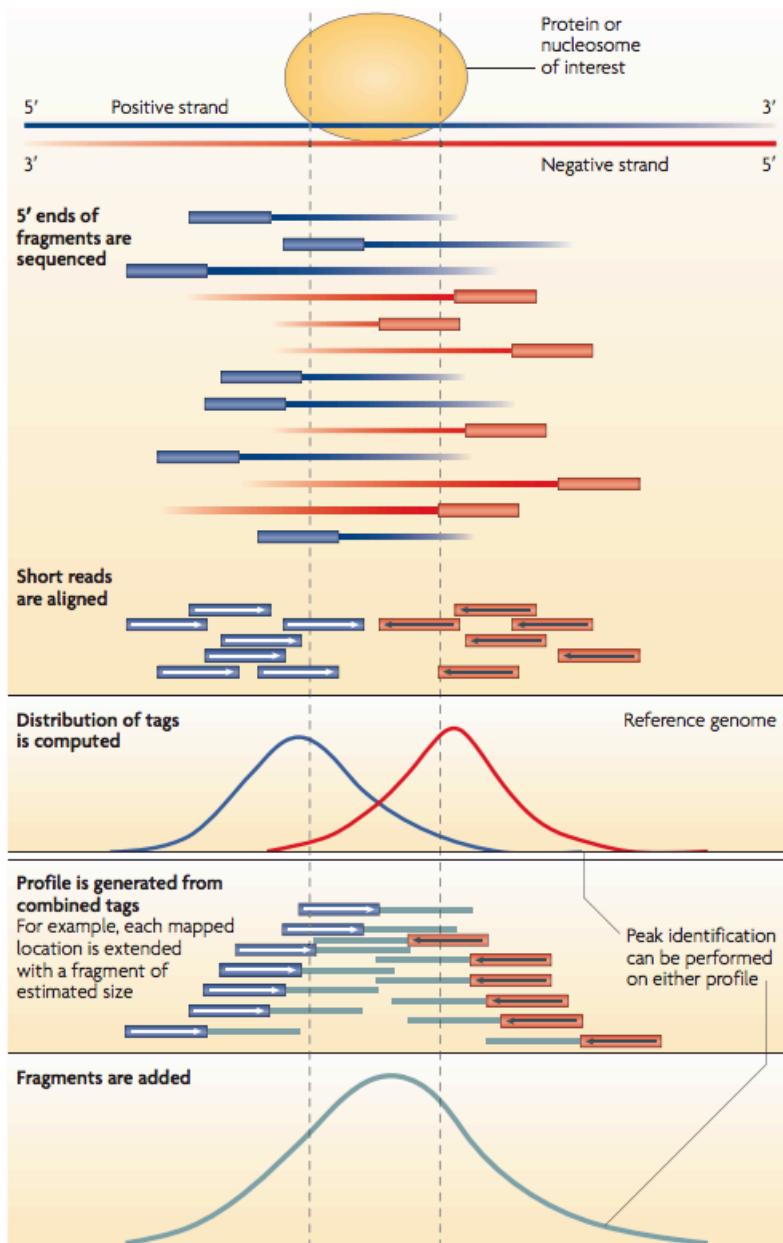
ChIP-Seq: Assay



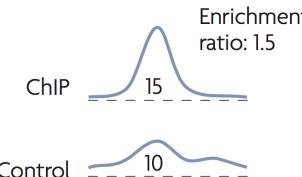
ChIP-Seq: Analysis pipeline



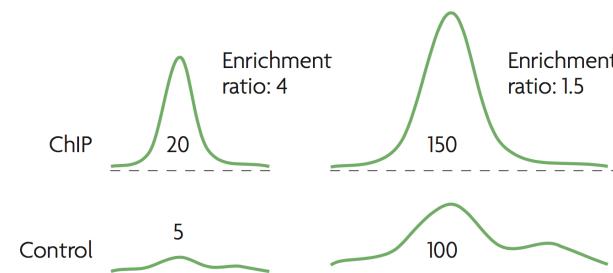
Goal: Identify enrichment peaks



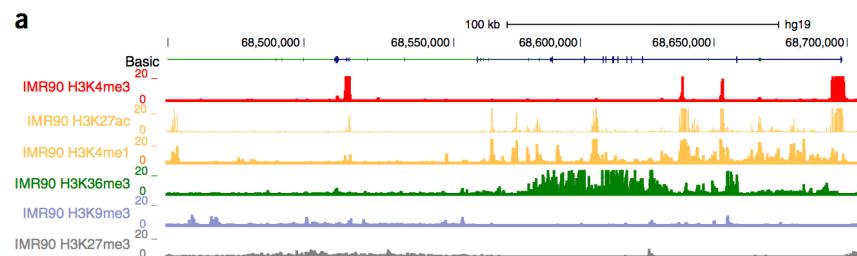
Ba Not statistically significant



Bb Statistically significant



a

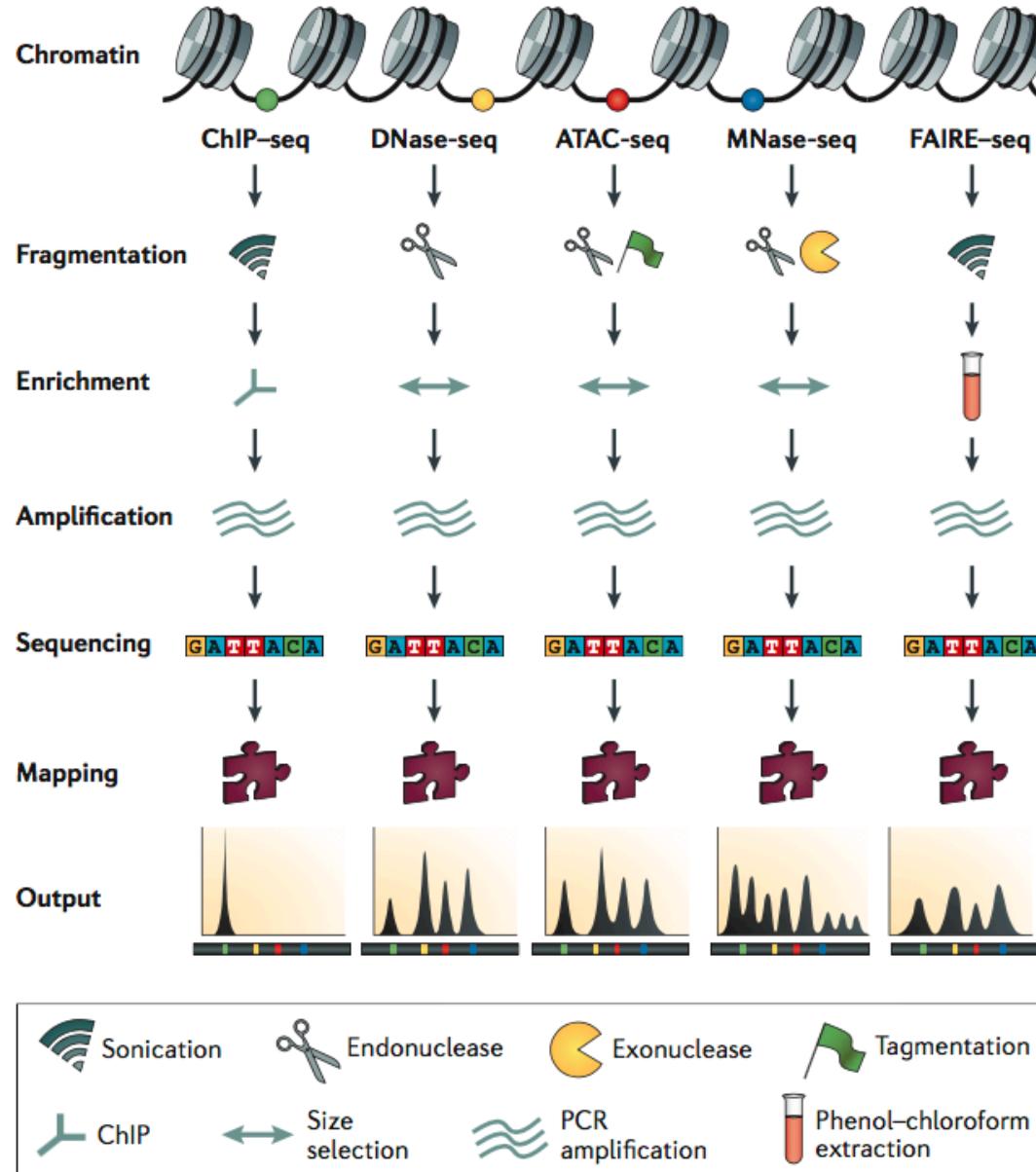


Park, P. J. (2009). ChIP-seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics*, 10(10), 669

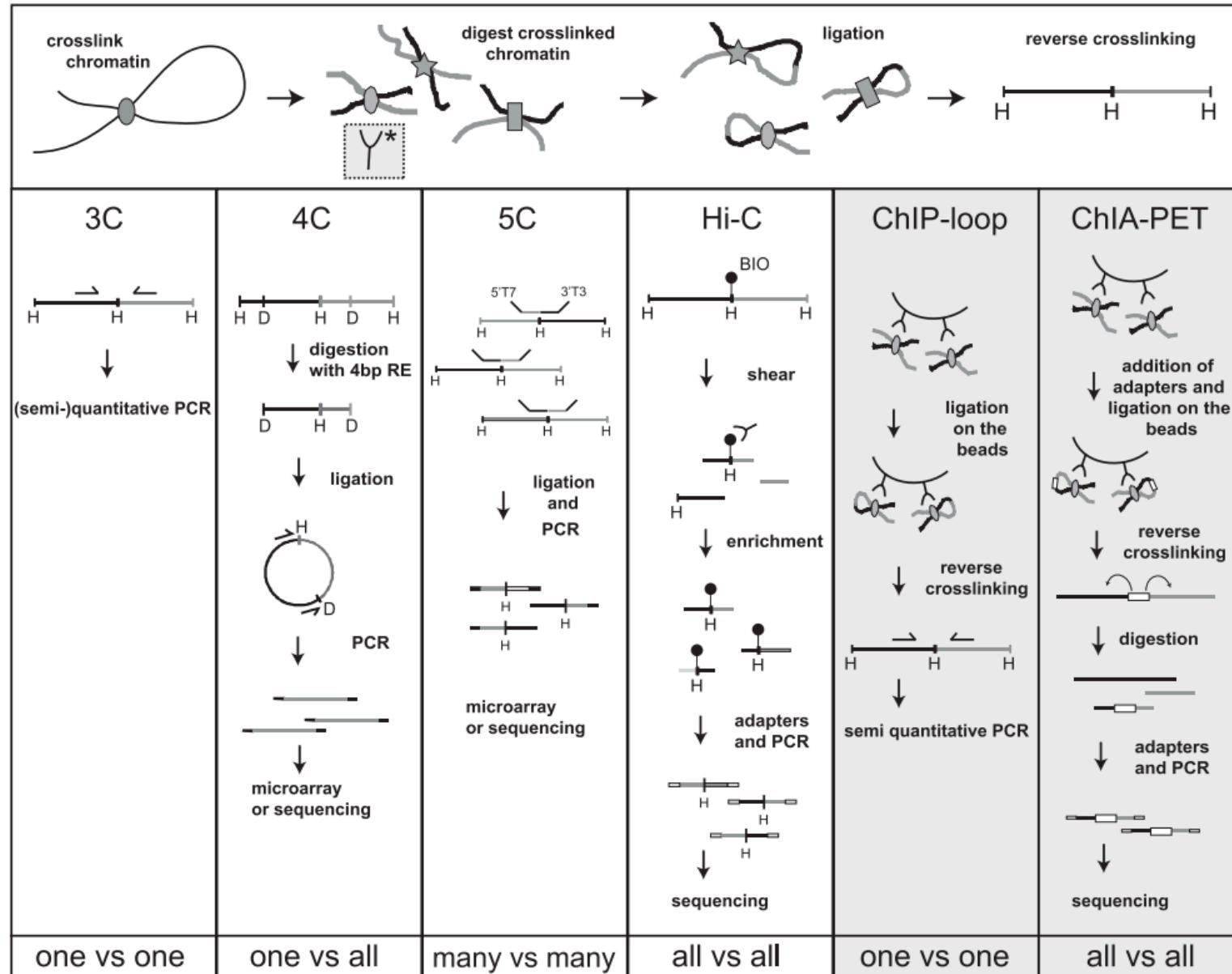
Ernst, J., & Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nature biotechnology*, 28(8), 817

Chromatin features

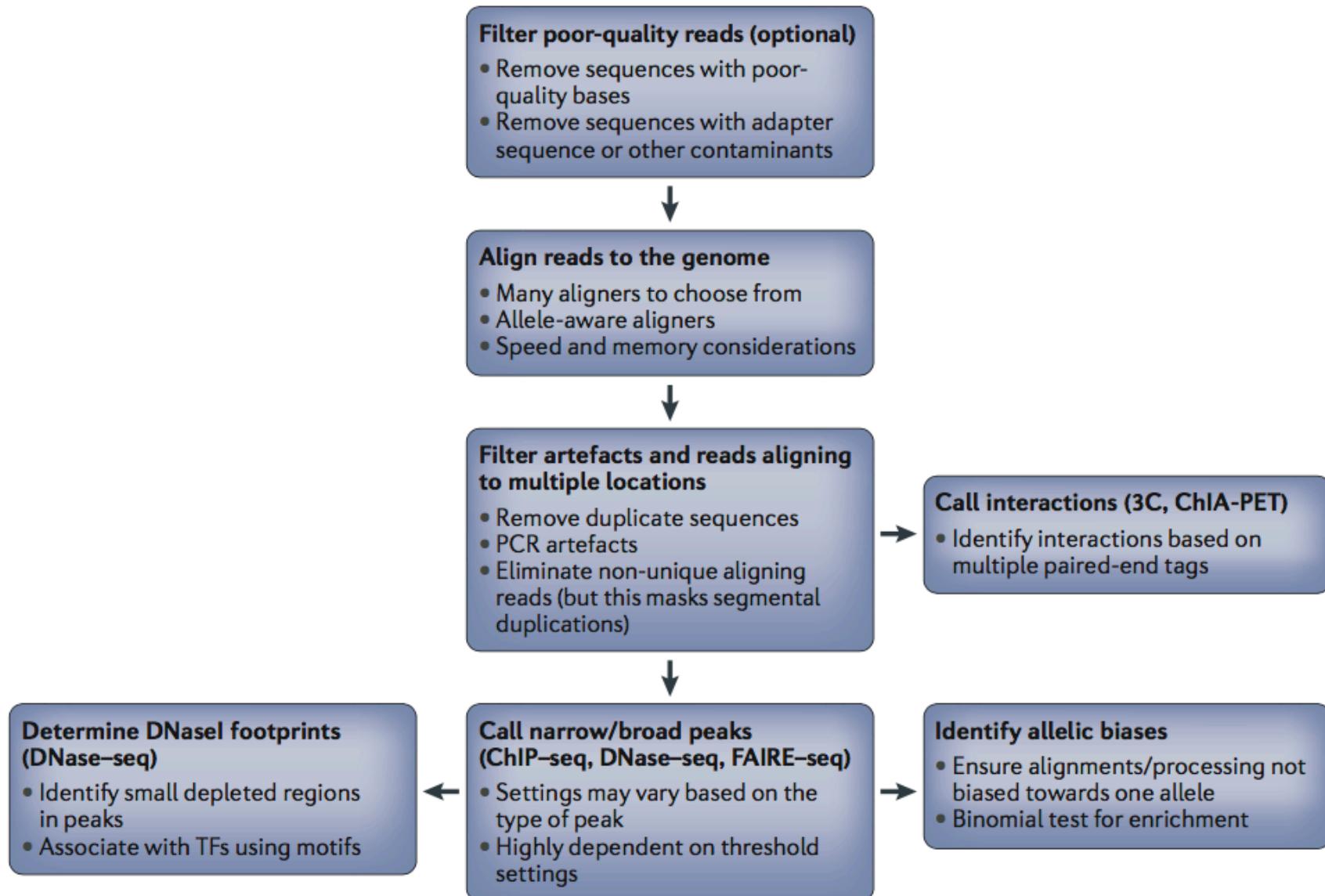
Chromatin features: accessibility



Chromatin features: long range interactions



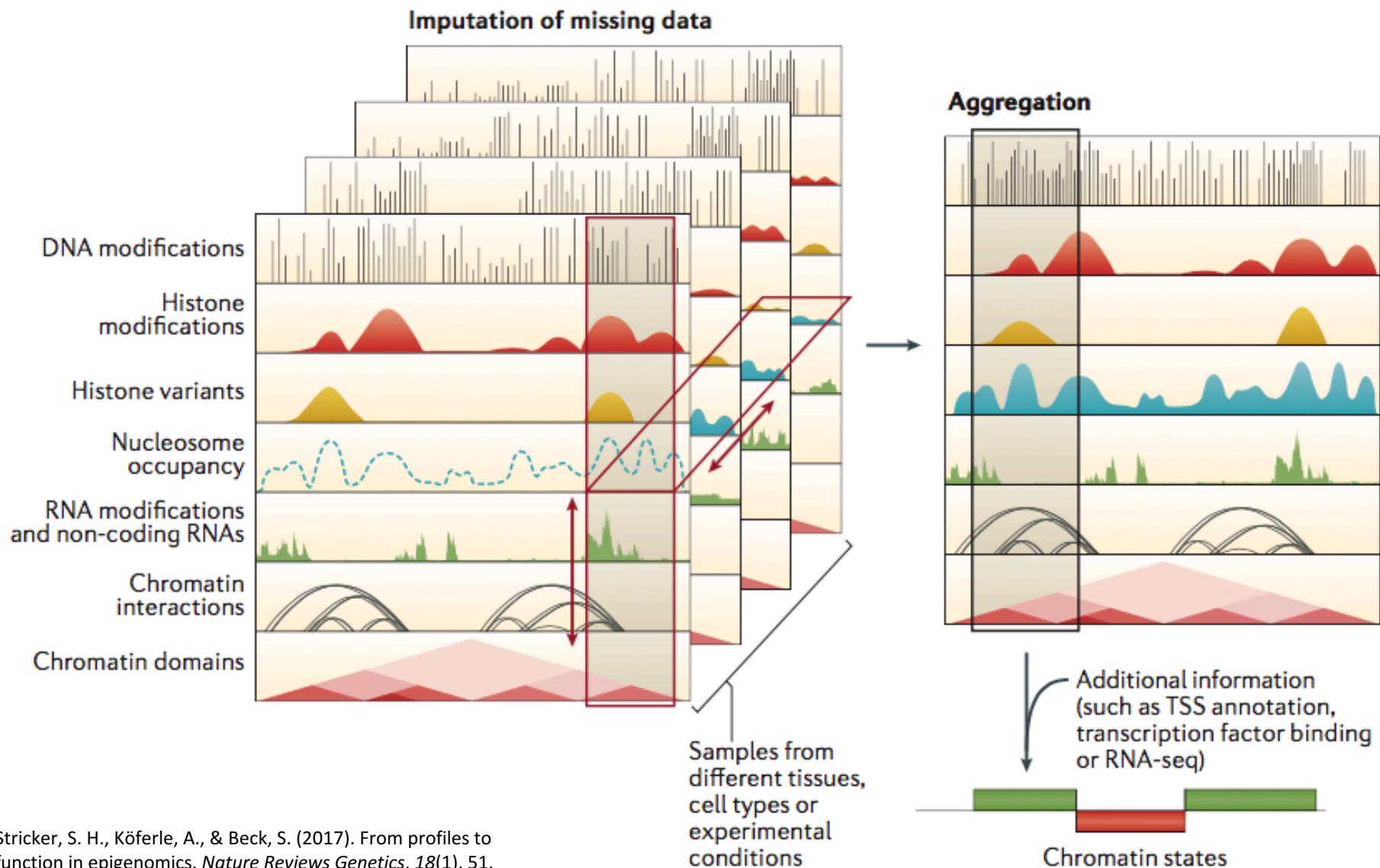
Post-sequencing bioinformatics pipeline



Integration

Putting it all together

Imputation and integration to infer functional states

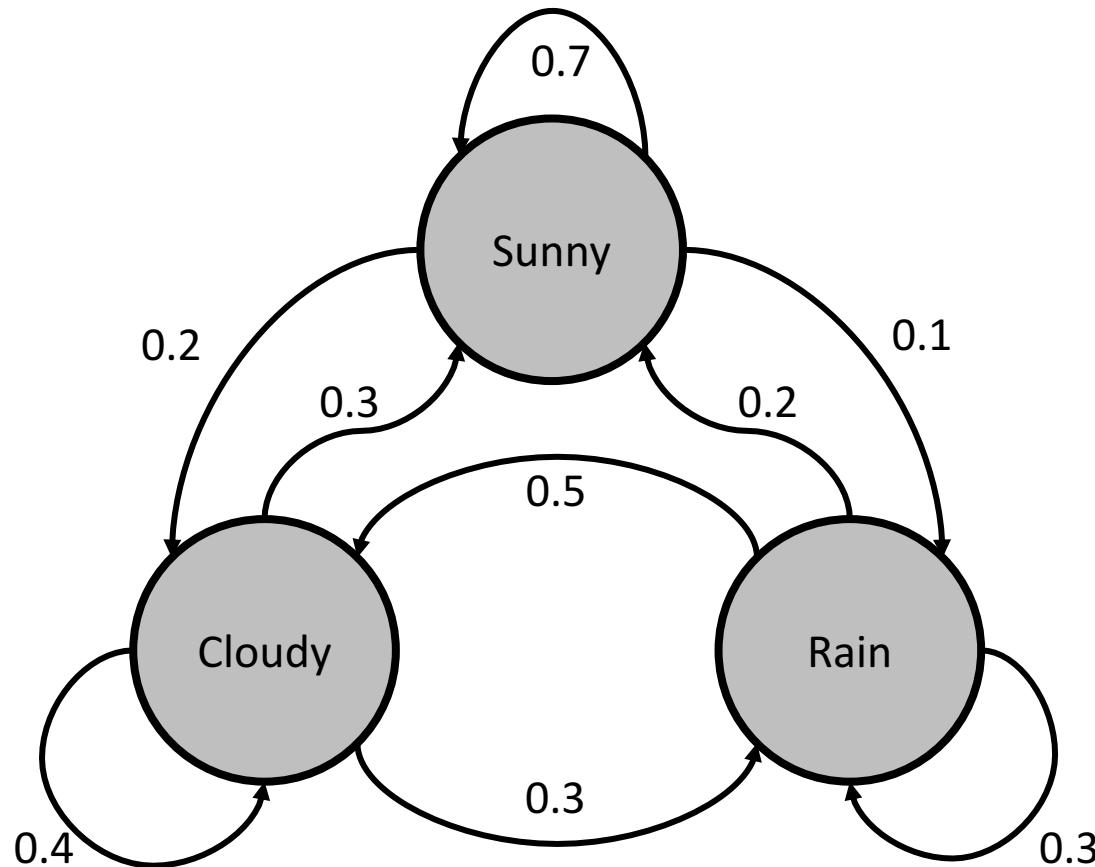


ChromHMM: Inferring chromatin states

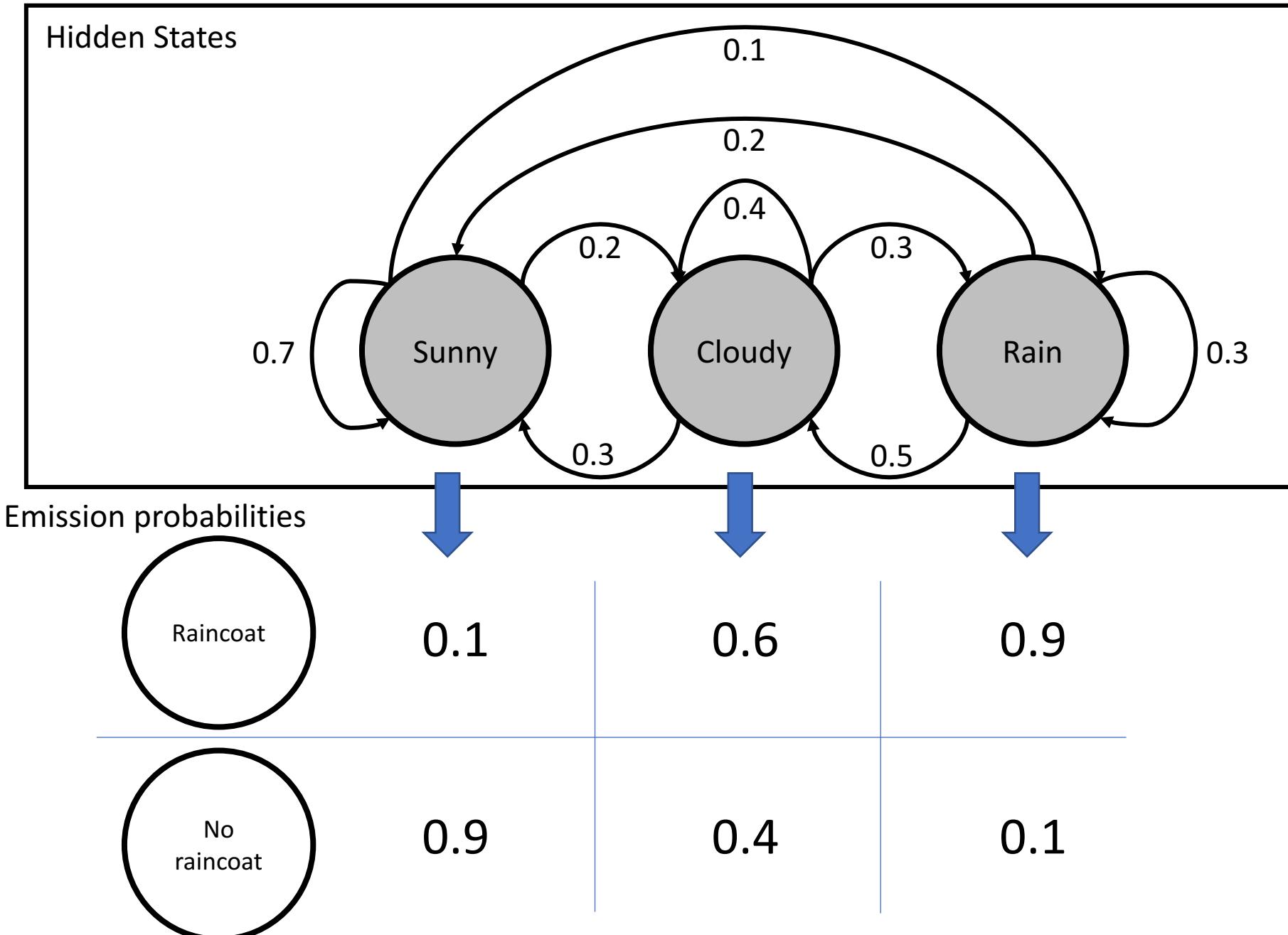
- Purpose: identify segments of the genome into biologically relevant functional elements
- ChromHMM uses a multivariate hidden Markov model (HMM) to infer chromatin states
- Chromatin states are derived from combinations of epigenetics marks (histone modifications, chromatin features)
- Chromatin states include:
 - Active/Weak
 - Promoter, enhancer, insulator
 - Polycomb

Simple Hidden Markov Model

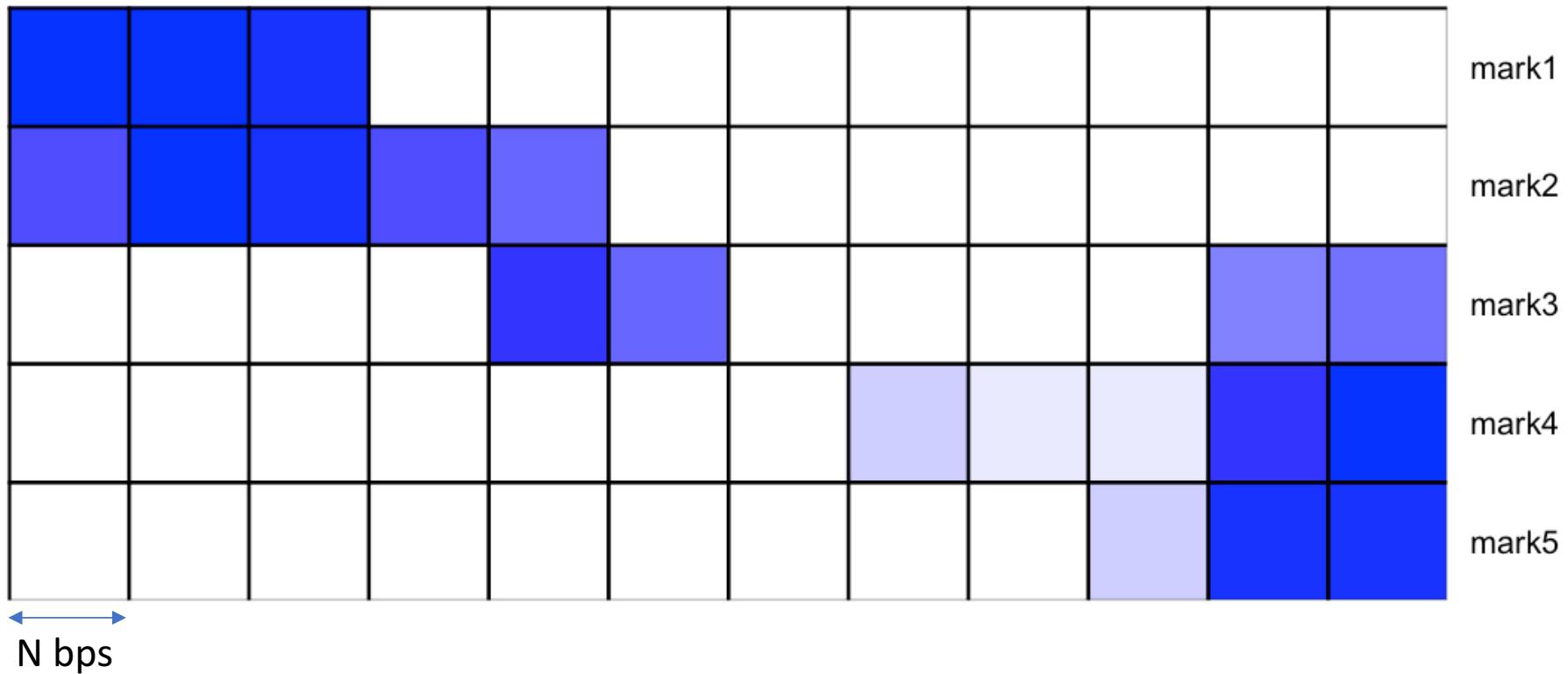
Markov Chain for weather



Simple Hidden Markov Model



Observation

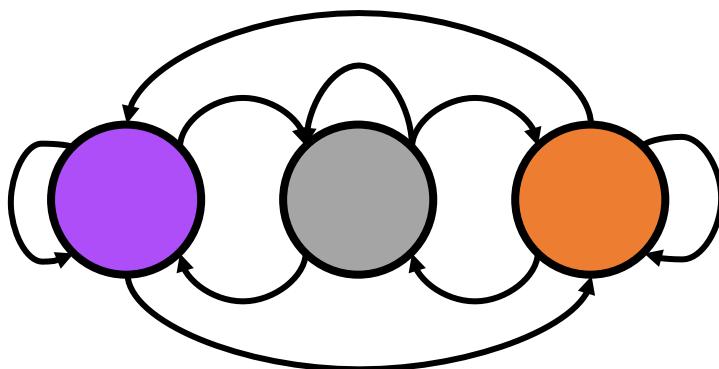


Observation

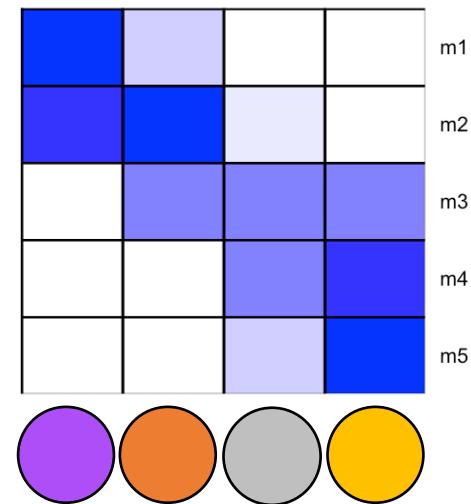
mark1													
mark2													
mark3													
mark4													
mark5													

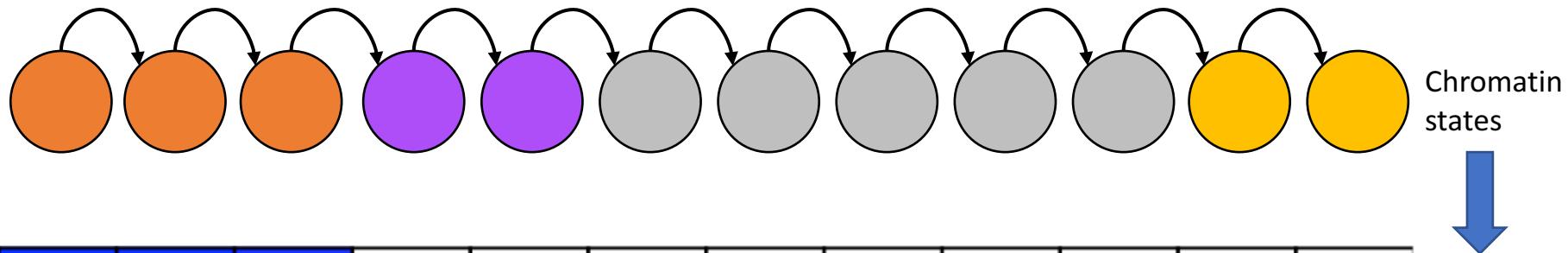
Learn from the data

1. Transition probabilities

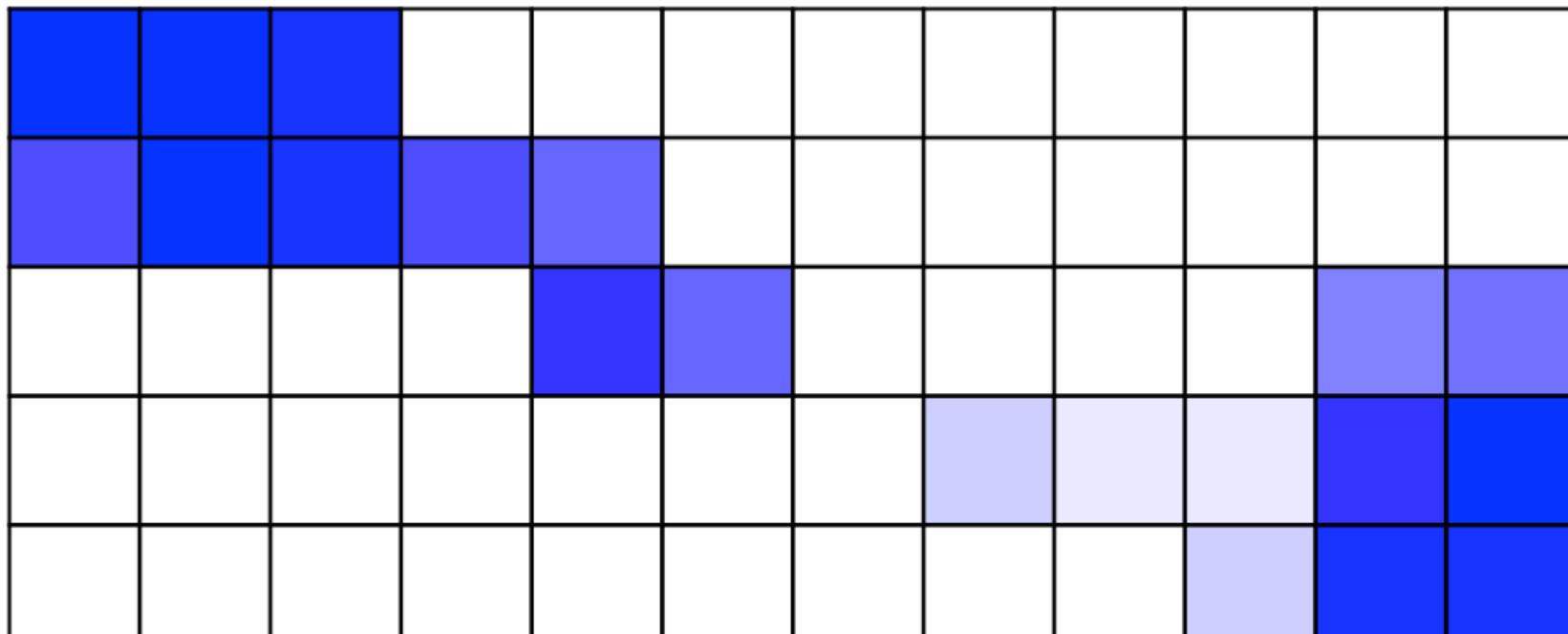


2. Emission probabilities





Chromatin
states



Genomic Annotations

TSS

Gene body

Enhancer

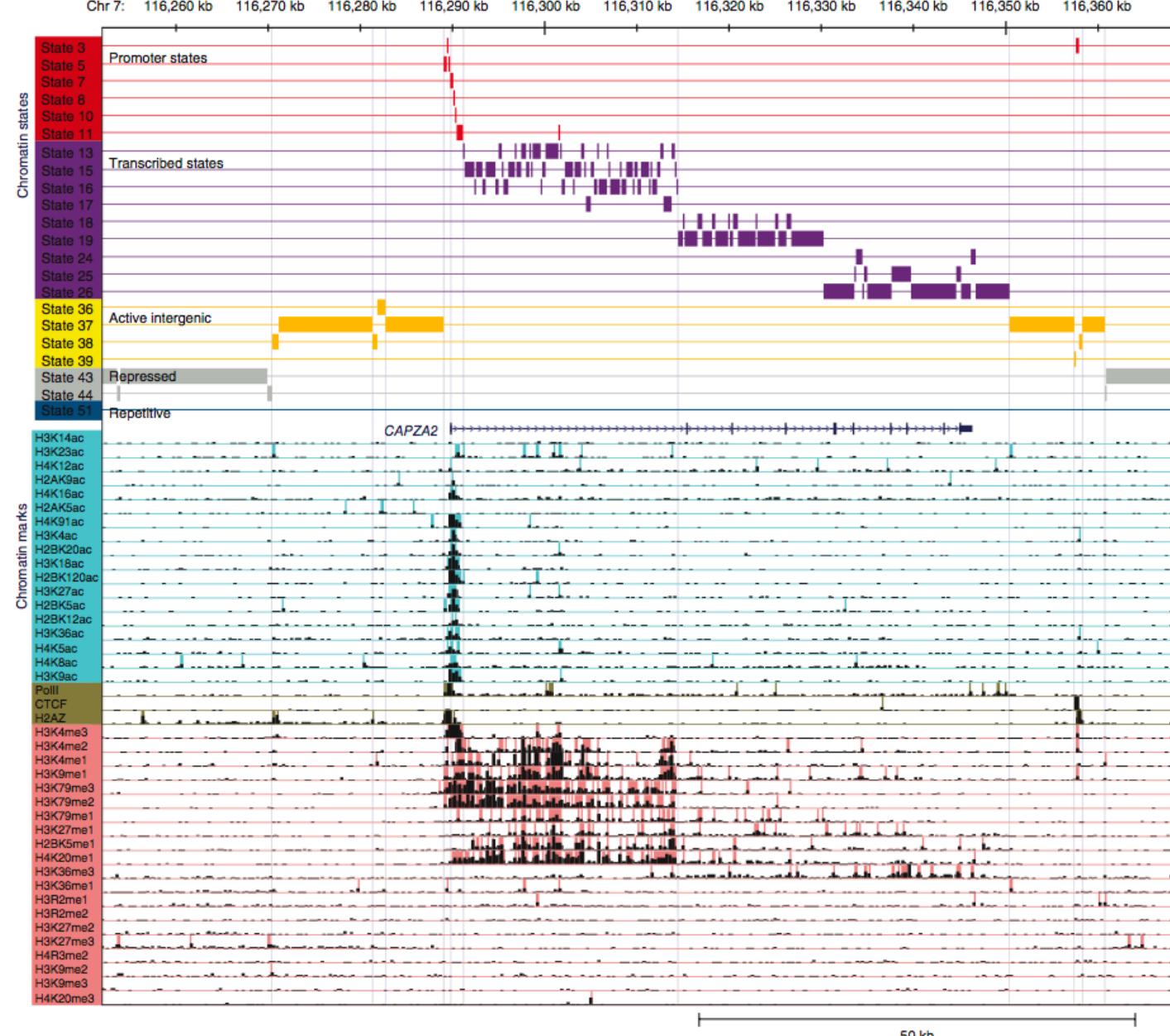
Chromatin states

Promoters

Transcribed states

Active intergenic

Repressed



Integrative genomics

Case 02:

Integrating genomics, epigenomics, and experimental biology to identify causal allele and regulatory mechanism in obesity



ESTABLISHED IN 1812

SEPTEMBER 3, 2015

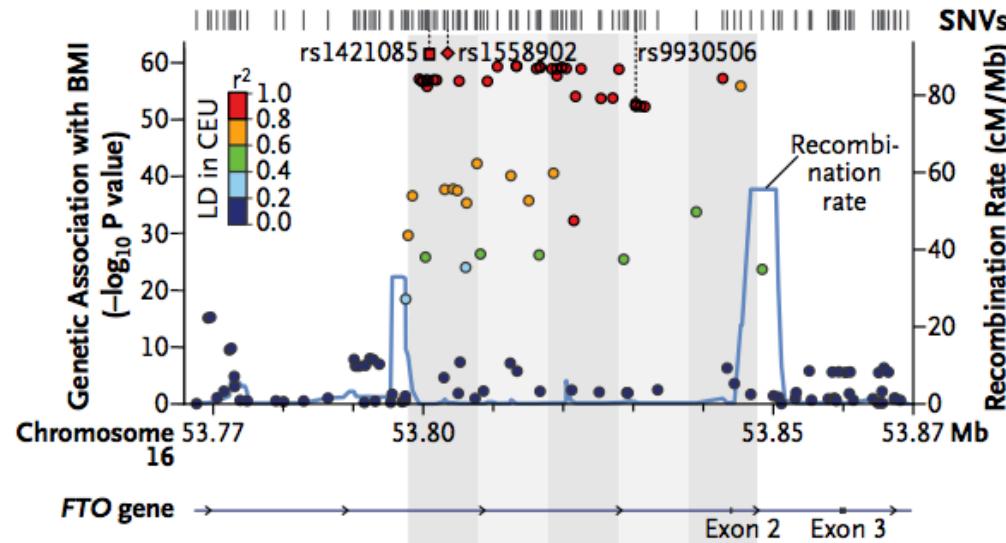
VOL. 373 NO. 10

FTO Obesity Variant Circuitry and Adipocyte Browning in Humans

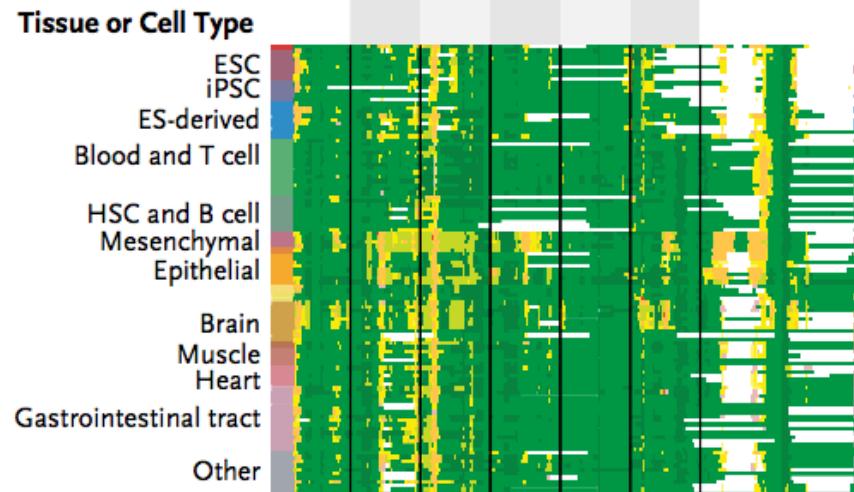
Melina Claussnitzer, Ph.D., Simon N. Dankel, Ph.D., Kyoung-Han Kim, Ph.D., Gerald Quon, Ph.D.,
Wouter Meuleman, Ph.D., Christine Haugen, M.Sc., Viktoria Glunk, M.Sc., Isabel S. Sousa, M.Sc.,
Jacqueline L. Beaudry, Ph.D., Vijitha Puvilindran, B.Sc., Nezar A. Abdennur, M.Sc., Jannel Liu, B.Sc.,
Per-Arne Svensson, Ph.D., Yi-Hsiang Hsu, Ph.D., Daniel J. Drucker, M.D., Gunnar Mellgren, M.D., Ph.D.,
Chi-Chung Hui, Ph.D., Hans Hauner, M.D., and Manolis Kellis, Ph.D.

GWAS locus prioritization using epigenome map of human cells

A

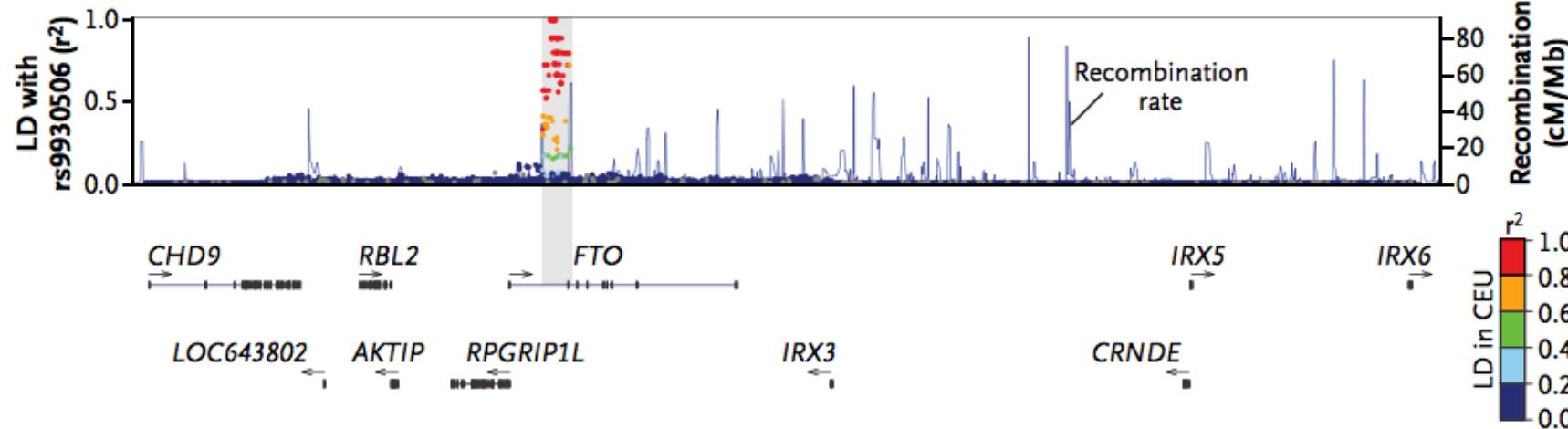


B

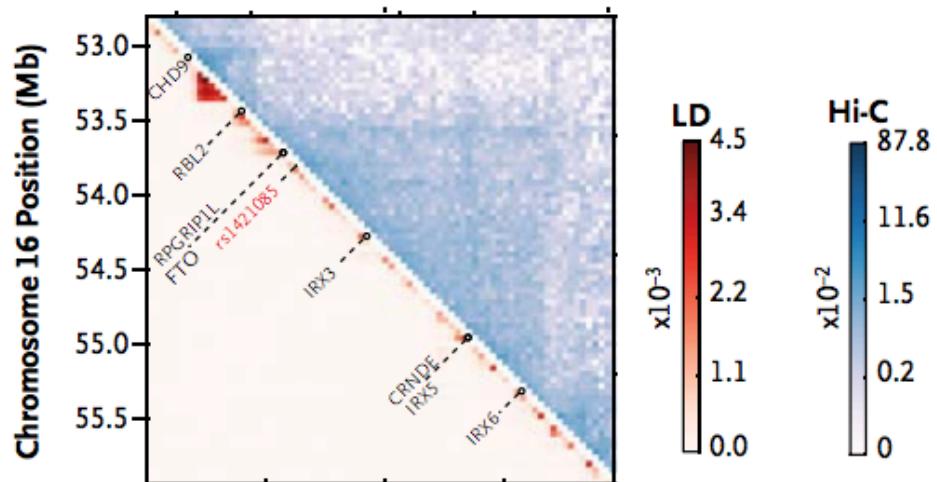


Hi-C data reveals 2-Mb topologically associated domain (TAD)

A

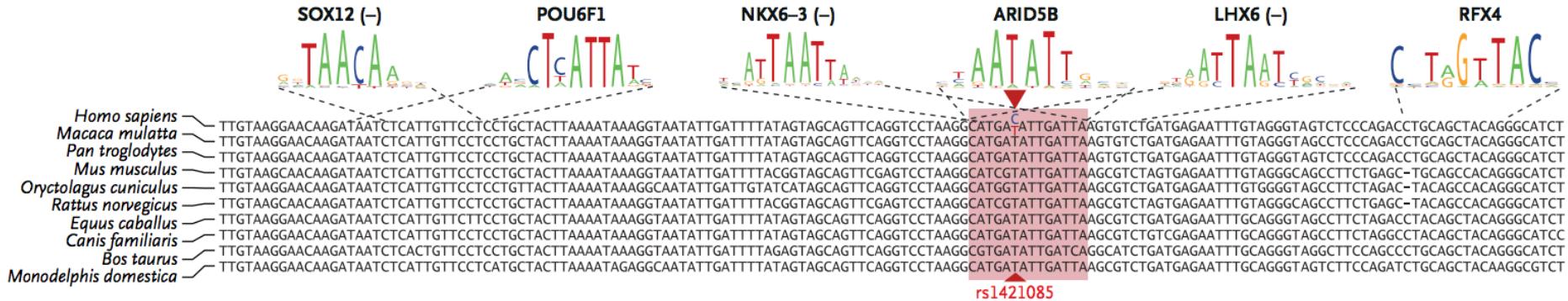


B



Identification of an SNV in linkage disequilibrium with candidate risk allele which disrupts an ARID5B repressor motif

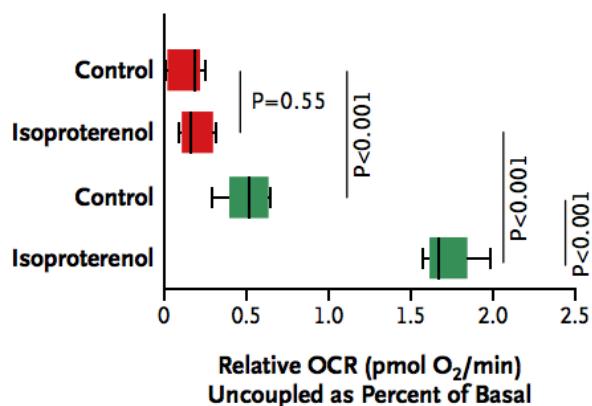
A



CRISPR-CAS9 rescue of homozygous risk allele restores oxygen consumption in adipocytes from carriers

rs1421085 (risk background)

- CC risk allele
- CC→TT rescue (CRISPR–Cas9 editing)



Limitations and considerations

Enrichment methods

Bisulfite-based methods

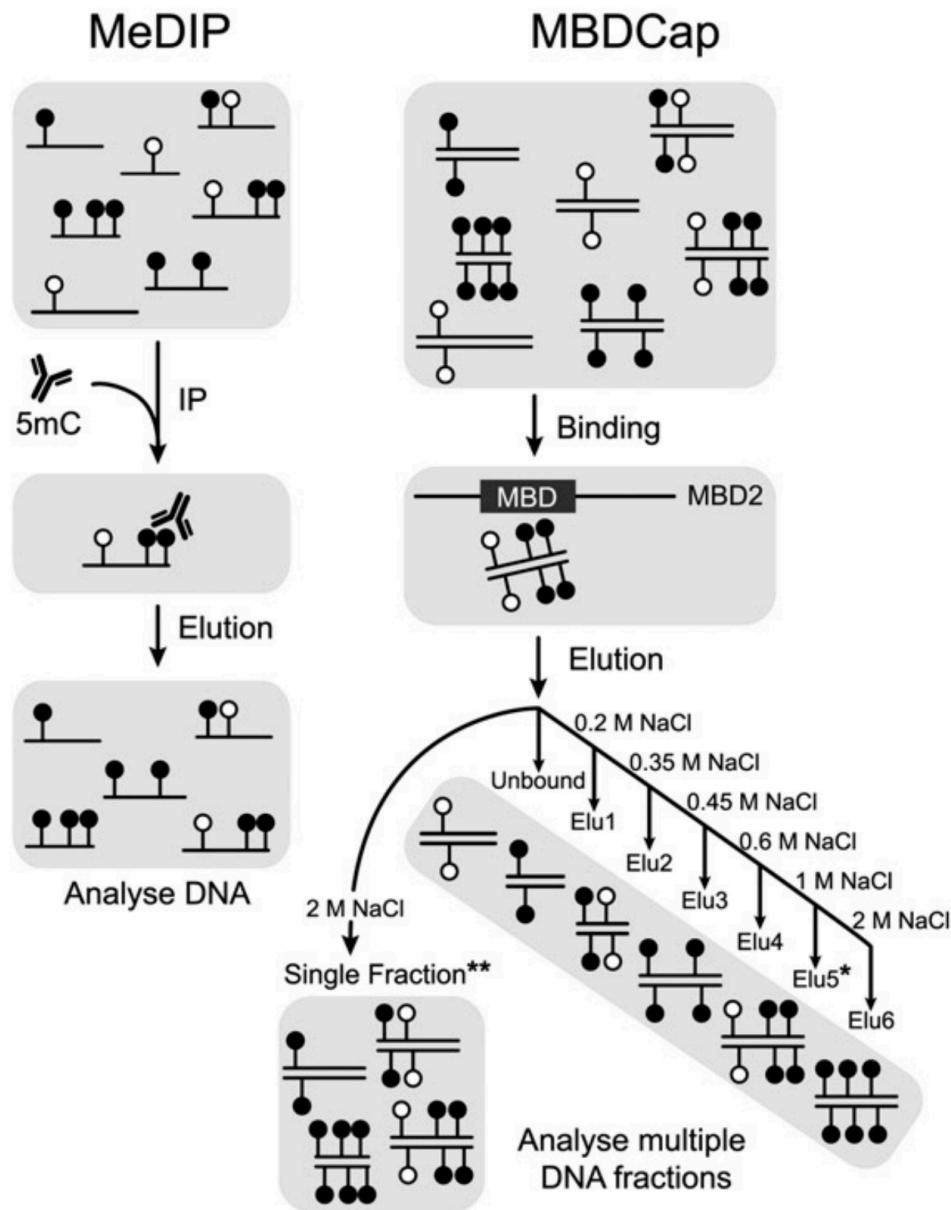
- Incomplete bisulfite conversion
 - Leads to increased methylation signal
 - Identification
 - Control probes in microarray
 - Methylation of non-CpG sites
 - Methylation-specific PCR of non-CpG sites

Affinity-based methods

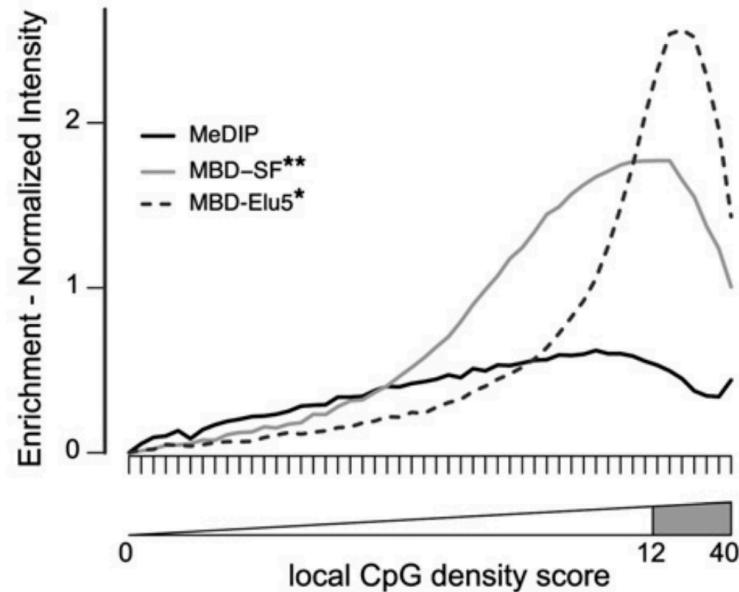
- Highly dependent on antibody used
 - Include input control and other appropriate controls

Differences between MeDIP and MBDCap

A



B



Robinson, M. D., Stirzaker, ... & Clark, S. J. (2010). Evaluation of affinity-based genome-wide DNA methylation data: effects of CpG density, amplification bias, and copy number variation. *Genome research*

Read-out

Array

- Sequence specific bias within reference genome
- Neighboring SNPs may interfere with binding

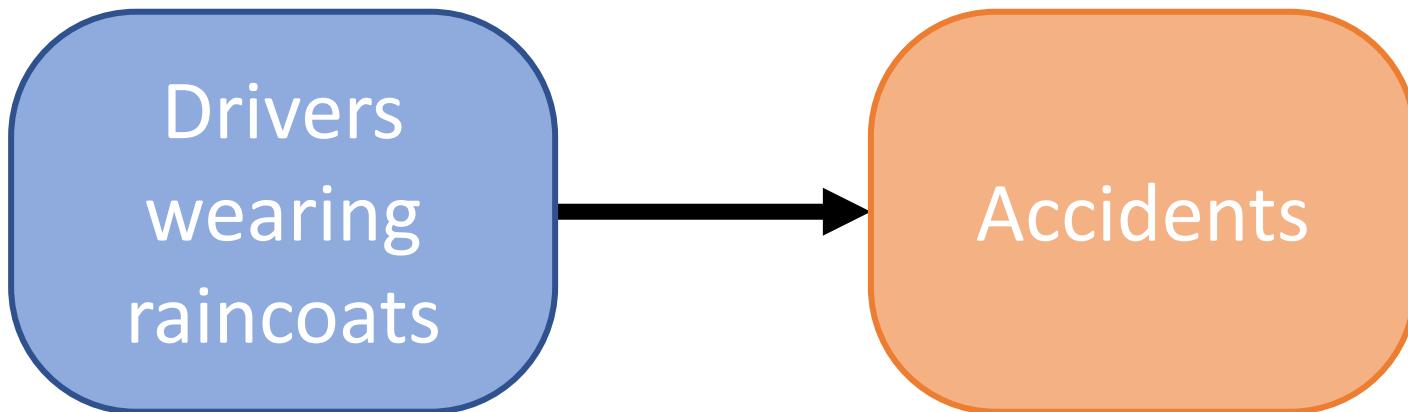
Sequencing

- Alignment
 - Repetitive regions are difficult to align
 - WGBS aligners have different biases depending on reference genome
- Uneven depth across genome, may bias peak calling
 - Include input DNA for capture-based methods

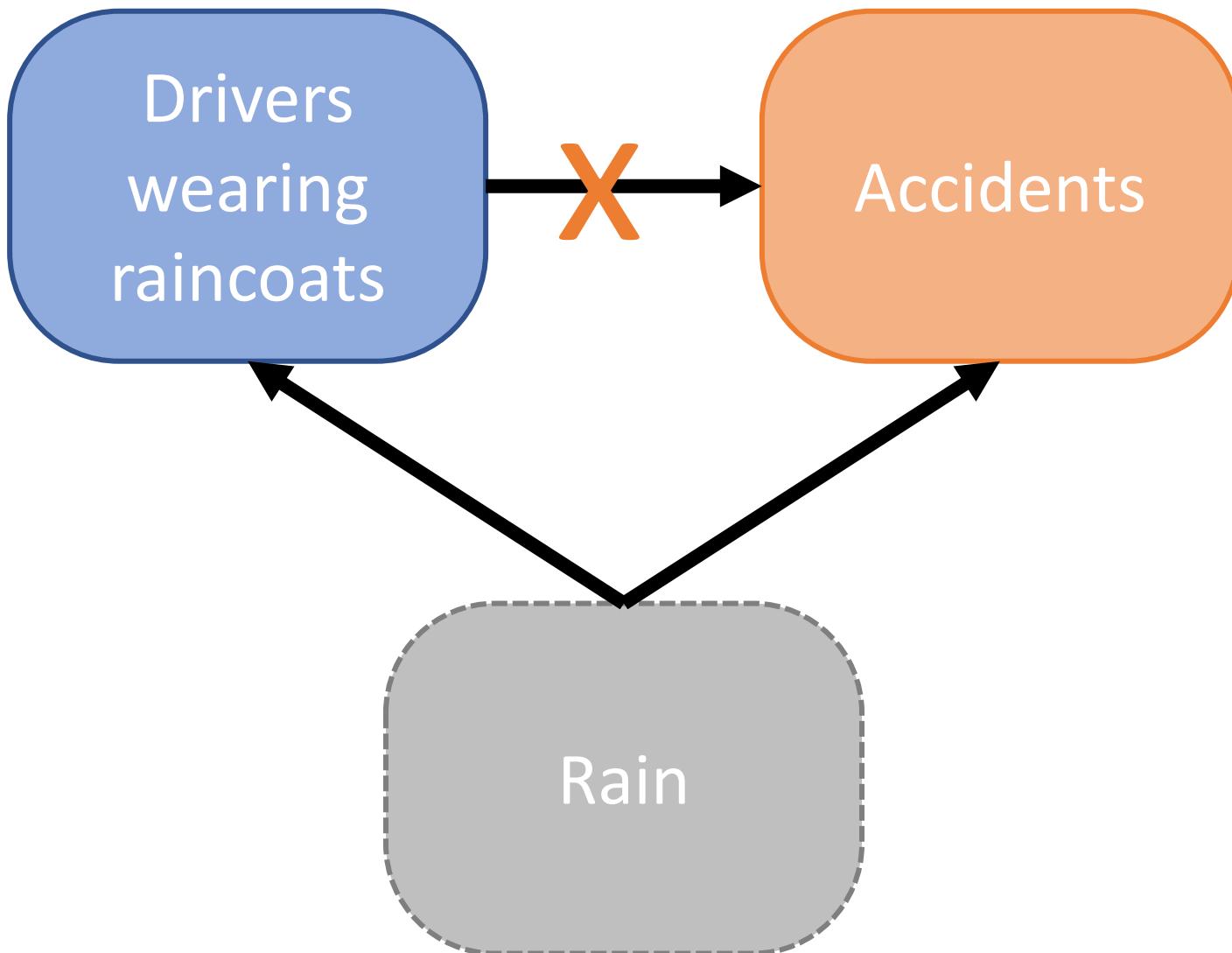
General statistical & design considerations

- Correlation does not mean causation
 - Experiments are required to establish causality
 - Inactive CRISPR-Cas9 linked to epigenetic effectors for targeted epigenetic alteration
- Balanced study
- Replication using an independent cohort is important
 - Especially for biomarker or clustering studies
 - Training and validation cohorts should be representative of each other
- Orthogonal validation of sequencing or microarray results
- Accounting for potential sources of variation

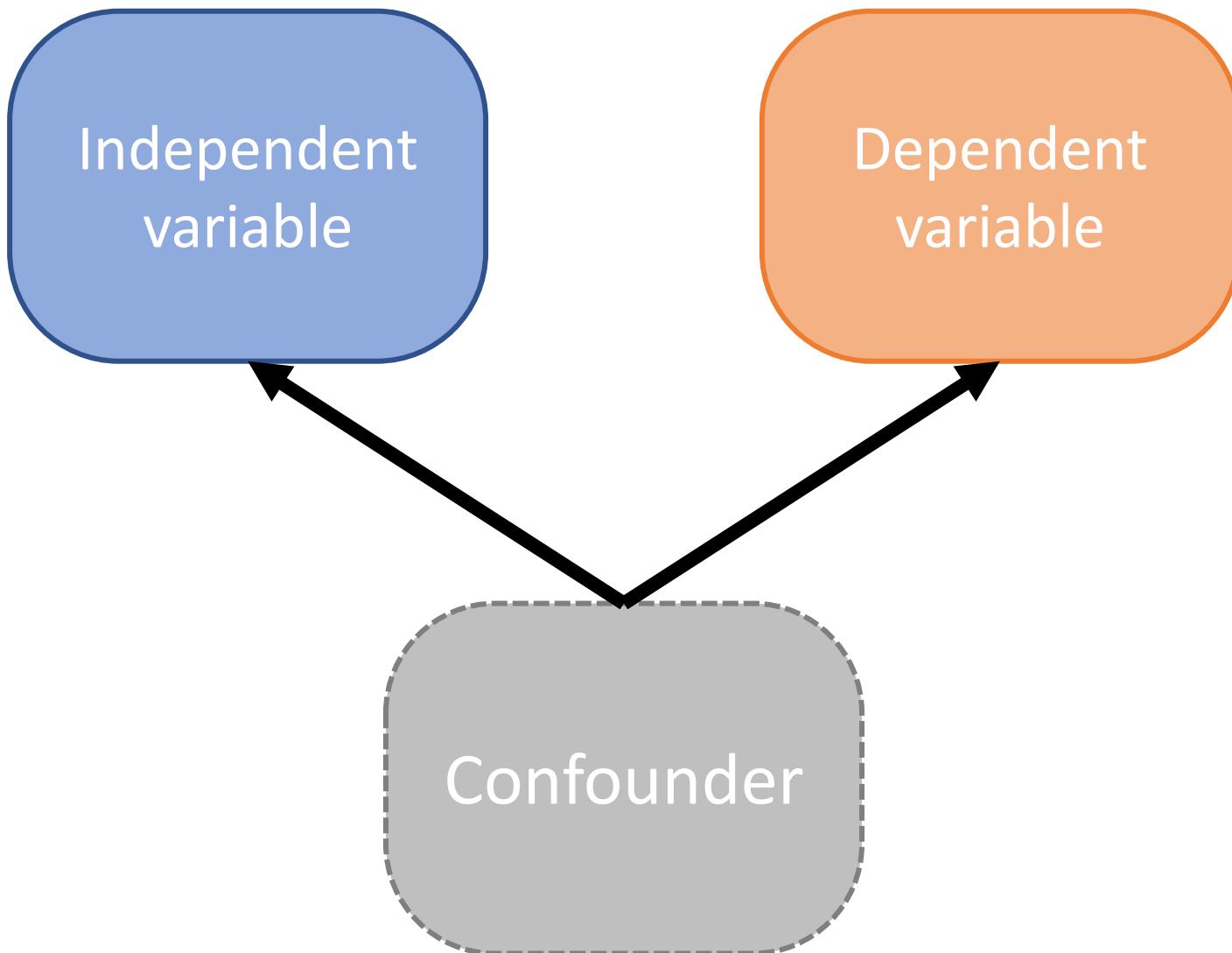
Confounding



Confounding

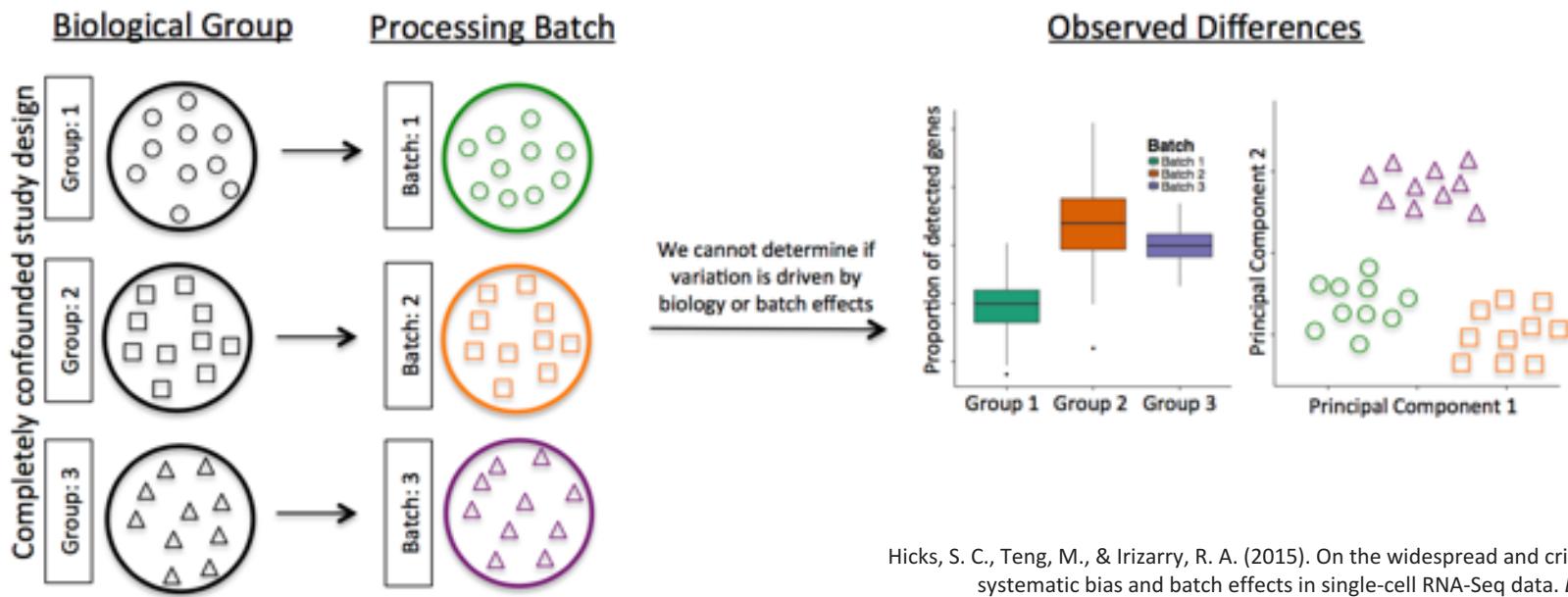


Confounding



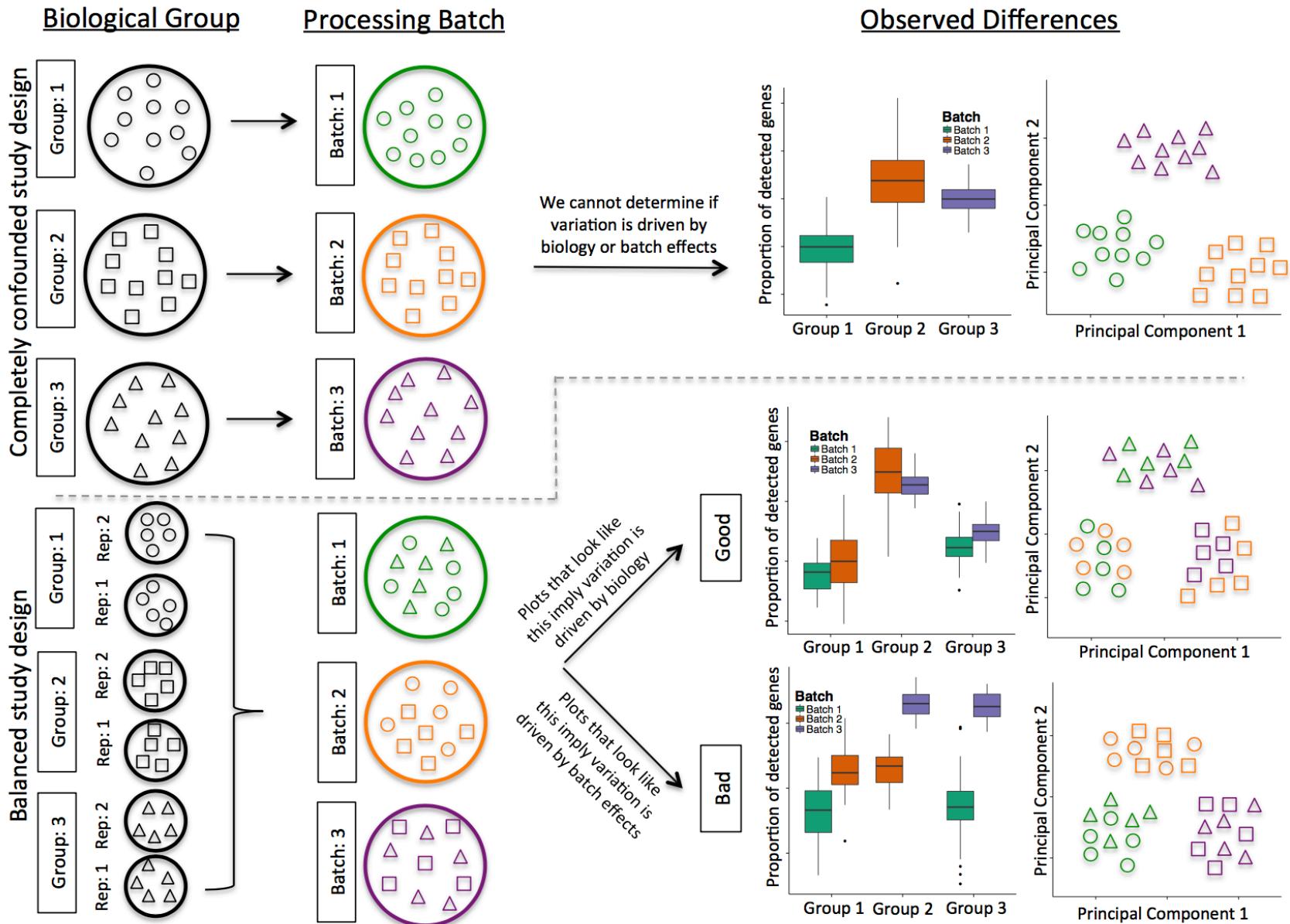
Batch as a confounder

Disease Date
Treatment Operator
Handedness Plate

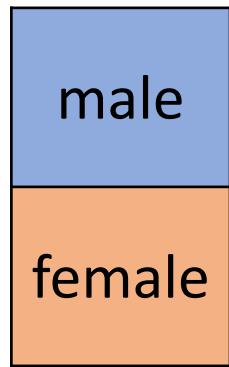


- Study design: randomization or blocked design
- Batch correction: estimate signals contributed from known batches and account in model

Batch as a confounder



Perfectly confounded

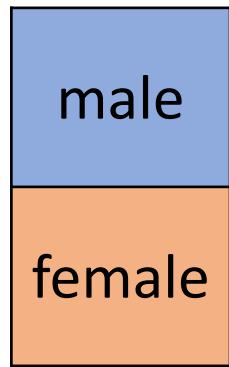


D = disease
C = control

D	D
D	D
D	D
D	D
D	D

C	C
C	C
C	C
C	C
C	C

Still confounded

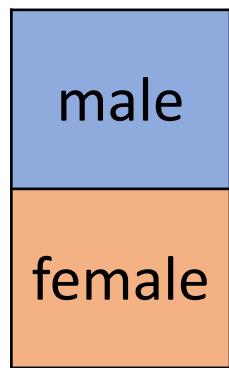


D = disease
C = control

D	D
D	D
D	D
D	D
D	D
D	D

C	C
C	C
C	C
C	C
C	C
C	C

Better

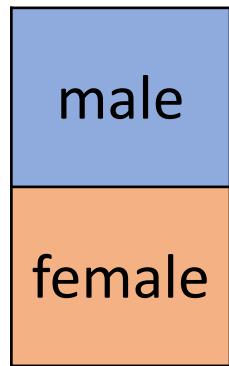


D = disease
C = control

D	D
D	D
D	D
C	C
C	C
C	C

D	D
D	D
D	D
C	C
C	C
C	C

Randomized



D = disease

C = control

D	D
C	C
C	C
D	D
C	C
D	D
D	C
C	D

C	C
D	D
D	D
C	C
D	C
C	D

General batch correction model

Observed signal (beta-value)

Uncaptured error/noise

$$\beta = \alpha + \beta_G + \beta_1 + \cdots + \beta_n + \delta + \epsilon$$

Average signal

Biological covariates

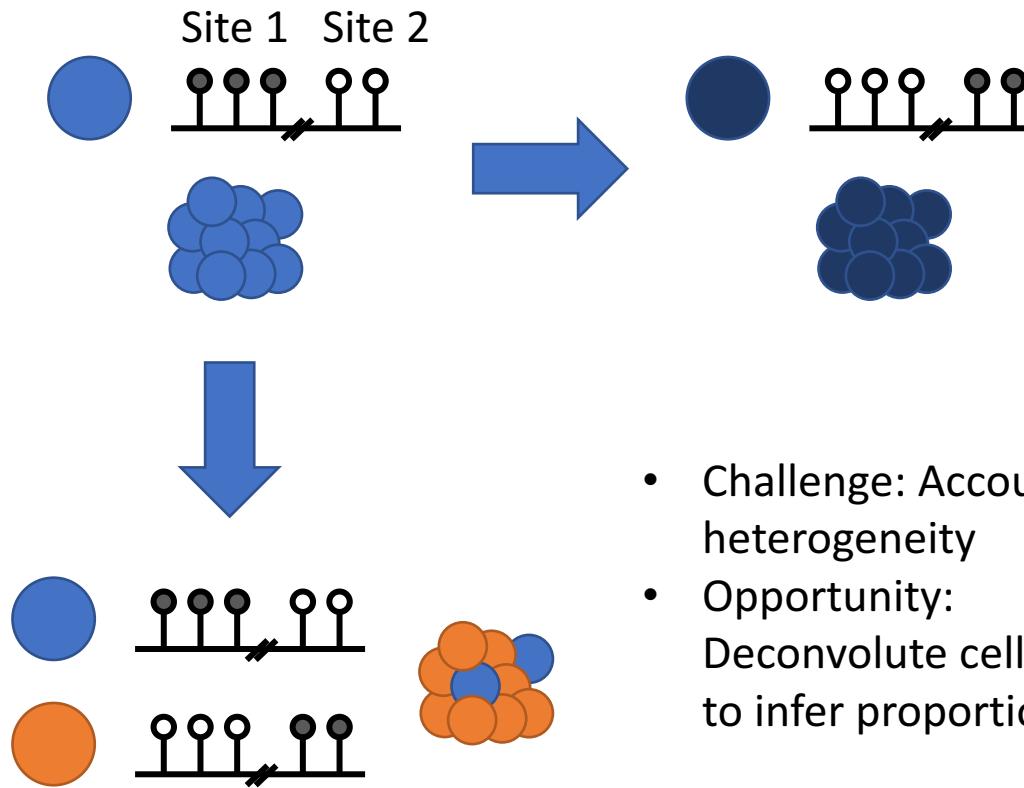
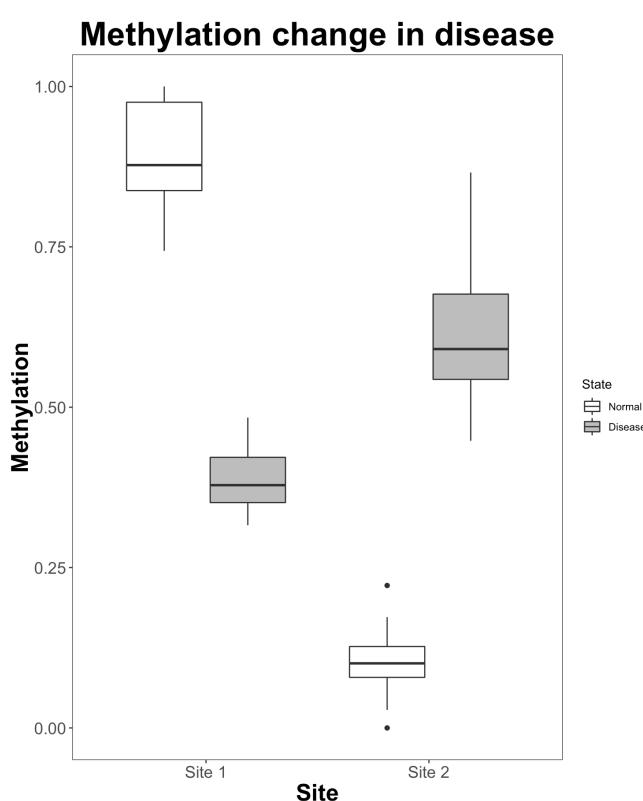
Independent
variable of interest
(e.g. disease state)

Batch

If the batch perfectly segregates with the independent biological variable of interest, the model cannot distinguish between β_G and δ , leading to removal of biological signal.

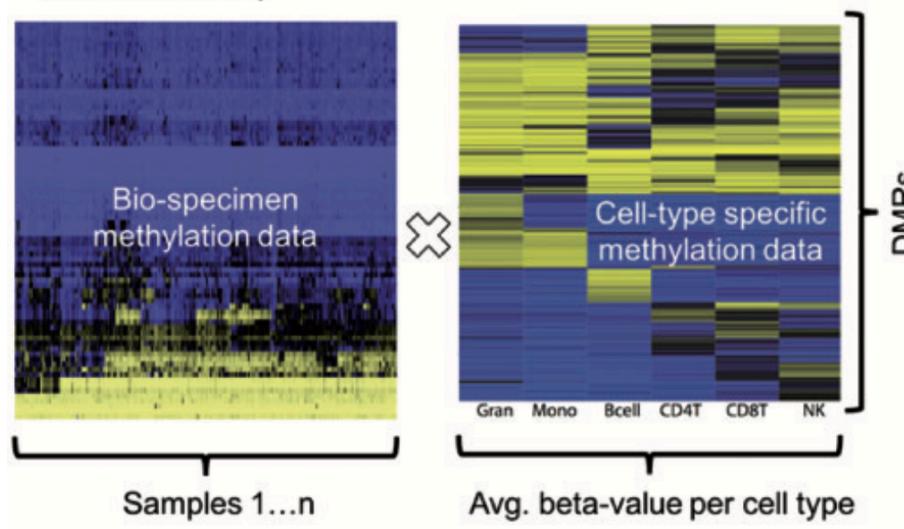
Heterogeneity

- Cell methylation exists in 3 states:
 - Unmethylated
 - Hemi-methylated
 - Methylated
- NOTE: Measurement includes a 4th state; missing data
- How are we measuring methylation as a continuous value from 0 – 100%?
- We are measuring a heterogeneous admixture of cells



- Challenge: Account for heterogeneity
- Opportunity: Deconvolute cell types to infer proportion

1. Reference-based cell type deconvolution using immune cell DMRs (e.g. Houseman method)



Result: a matrix of samples with estimated immune cell type proportions

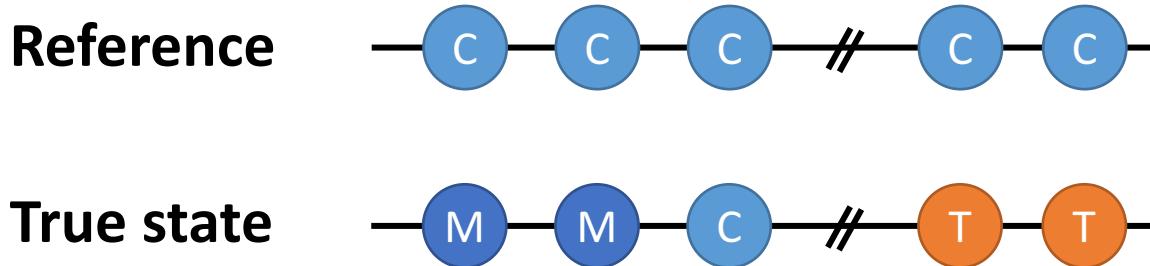
The resulting matrix is a table where rows represent 'Samples 1...n' and columns represent 'Immune proportion estimates for samples 1...n'. The columns are labeled: Gran, Mono, B-cell, CD4T, CD8T, and NK.

Samples 1...n		Gran	Mono	B-cell	CD4T	CD8T	NK
Gran ₁	Mono ₁	B-cell ₁	CD4T ₁	CD8T ₁	NK ₁		
Gran ₂	Mono ₂	B-cell ₂	CD4T ₂	CD8T ₂	NK ₂		
...
Gran _n	Mono _n	B-cell _n	CD4T _n	CD8T _n	NK _n		

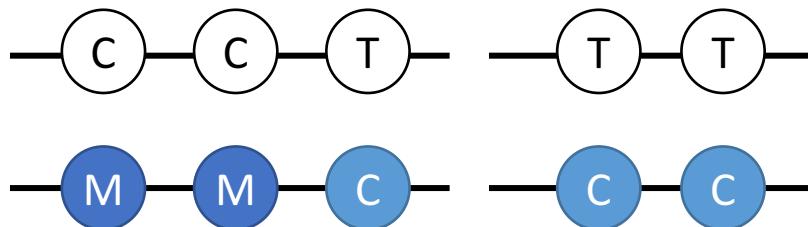
Emerging technologies in epigenomics.

Nanopore sequencing

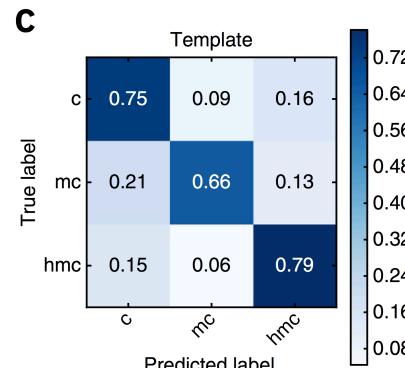
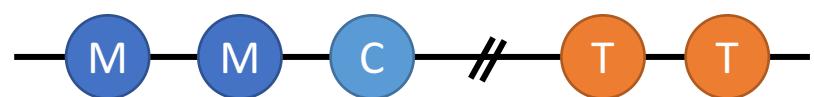
Allows for the detection of 5mC and 5hmC without bisulfite treatment and distinguishes T's from unmethylated C's.



Bisulfite sequencing

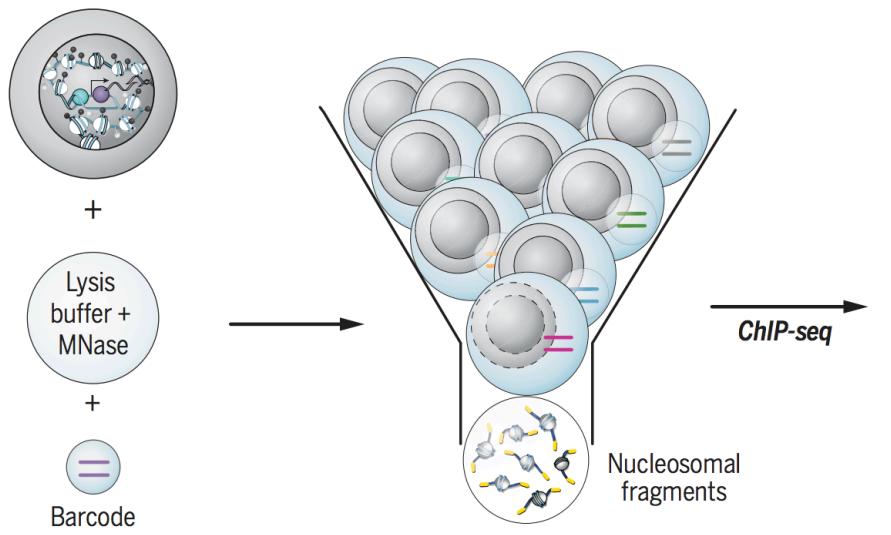


Nanopore sequencing

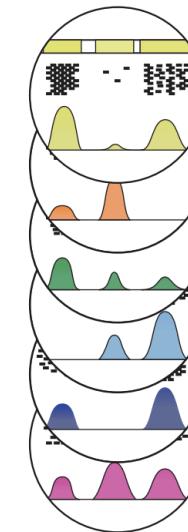
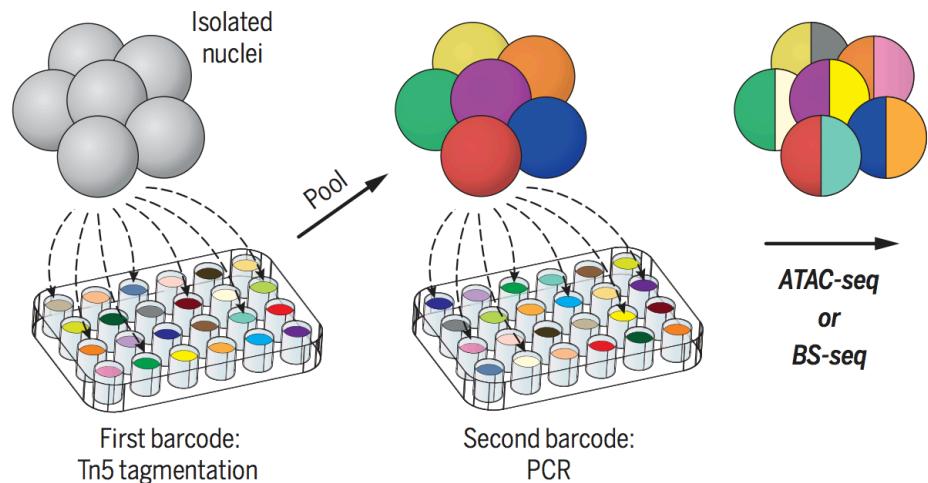


Single cell epigenomics

Droplet barcoding

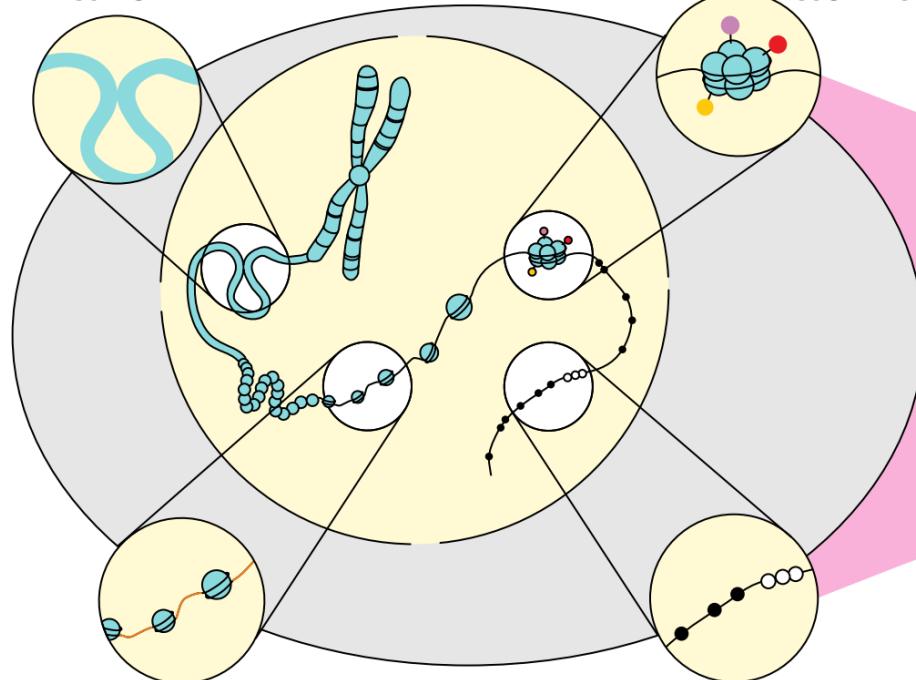


Combinatorial barcoding



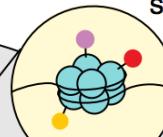
Chromosome conformation

scHiC



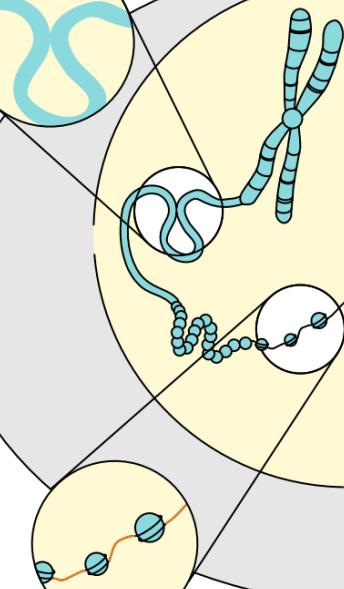
Histone modifications

scChIP-seq



DNA accessibility

scATAC-seq, scDNase-seq



DNA modifications

scRRBS, scBS-seq

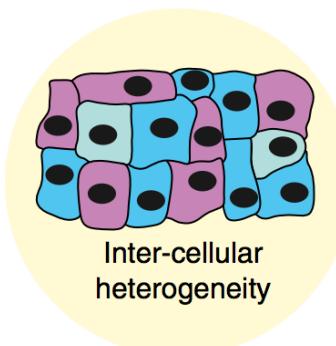
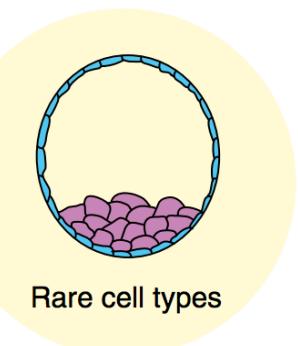
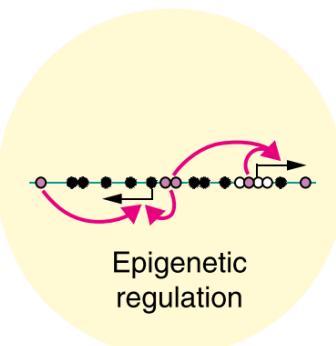


Transcriptomics

Epigenomics

Genomics

Integrated multi-Omics



Thank you.

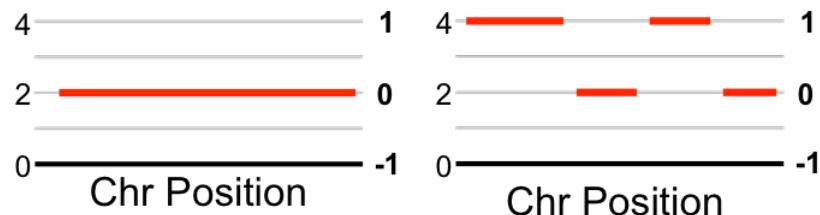
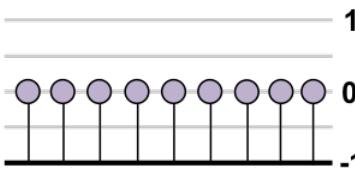
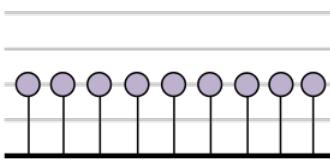
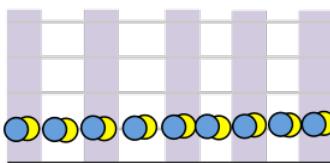
QUESTIONS?

APPENDIX

SNP Array

Normal Sample

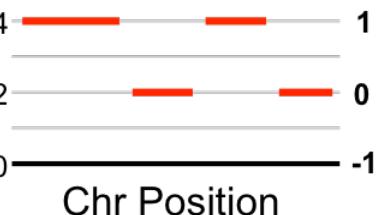
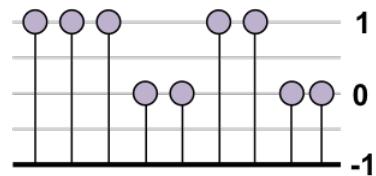
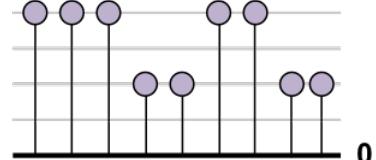
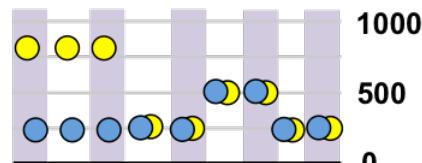
- Allele A = A_i
- Allele B = B_i



SNP Array

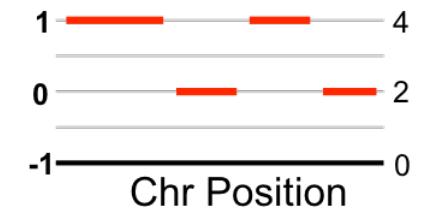
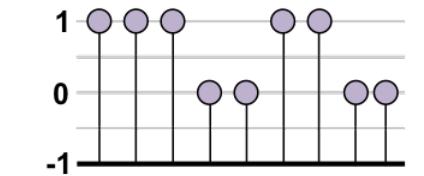
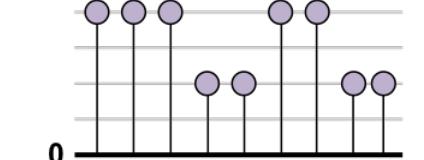
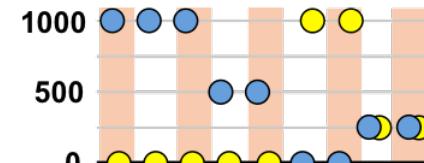
Tumor Sample

- Allele A = A_i
- Allele B = B_i



Methylation Array

- Unmethylated, T = A_i
- Methylated, C = B_i



Raw signal intensity:
 A_i, B_i

↓
 Observed copy number:
 $R_i = A_i + B_i$

↓
 Log R Ratio:
 $LRR = \log(R_i/R')$

↓
 Segmented LRR