Overview of the state of epigenetics: from genome to epigenome to drug discovery

Jonathan Pevsner, Ph.D.
Kennedy Krieger Institute
March 10, 2016
ASCPT session:
Using epigenetic markers for drug discovery

Outline

The human genome

Epigenetics and the human epigenome
- The ENCODE Project
- Roadmap Epigenomics Project

Epigenetic markers for drug discovery

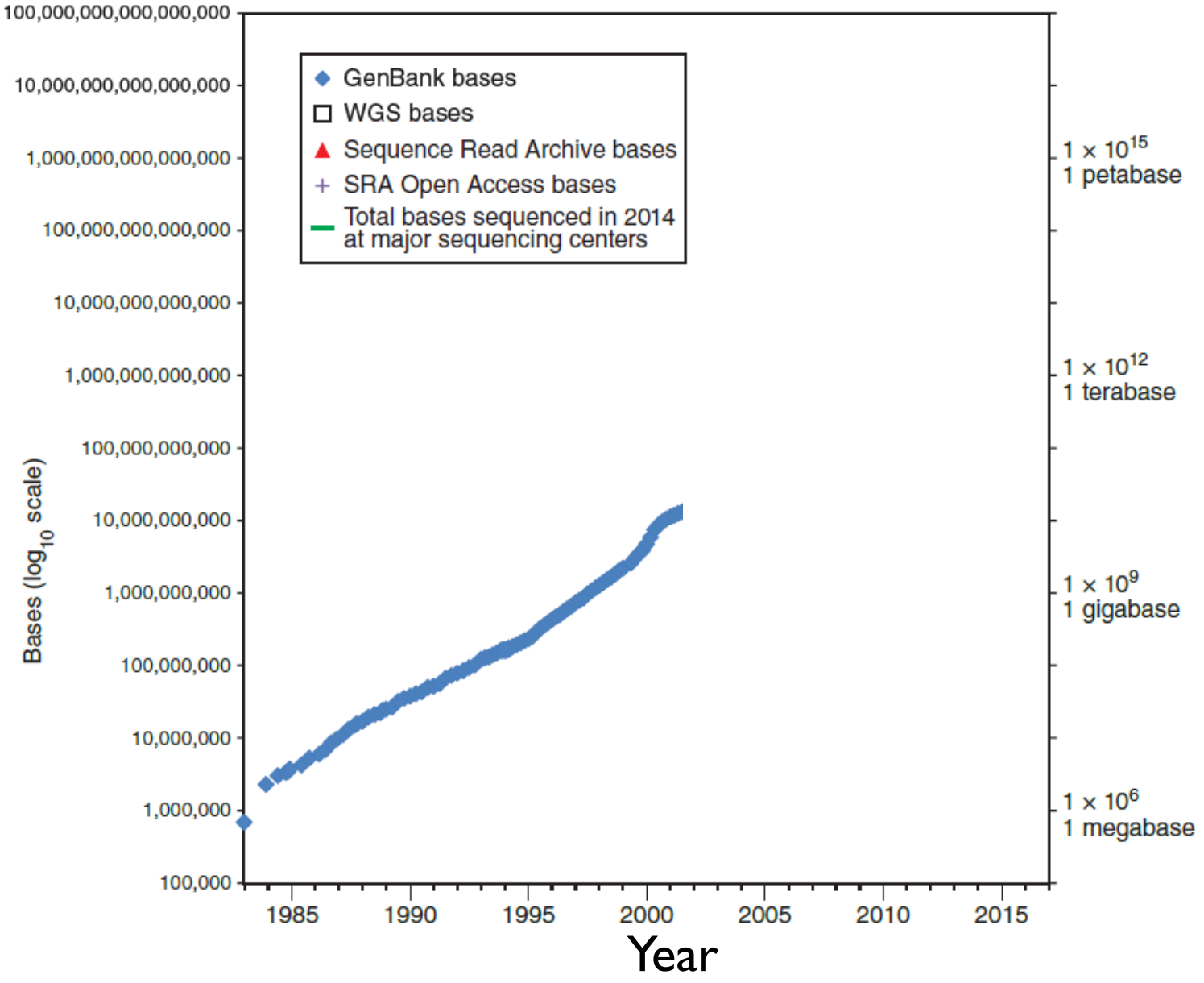
The human genome

2003: Human genome project completed

2007: First individual human genome reported
Emergence of next-generation sequencing

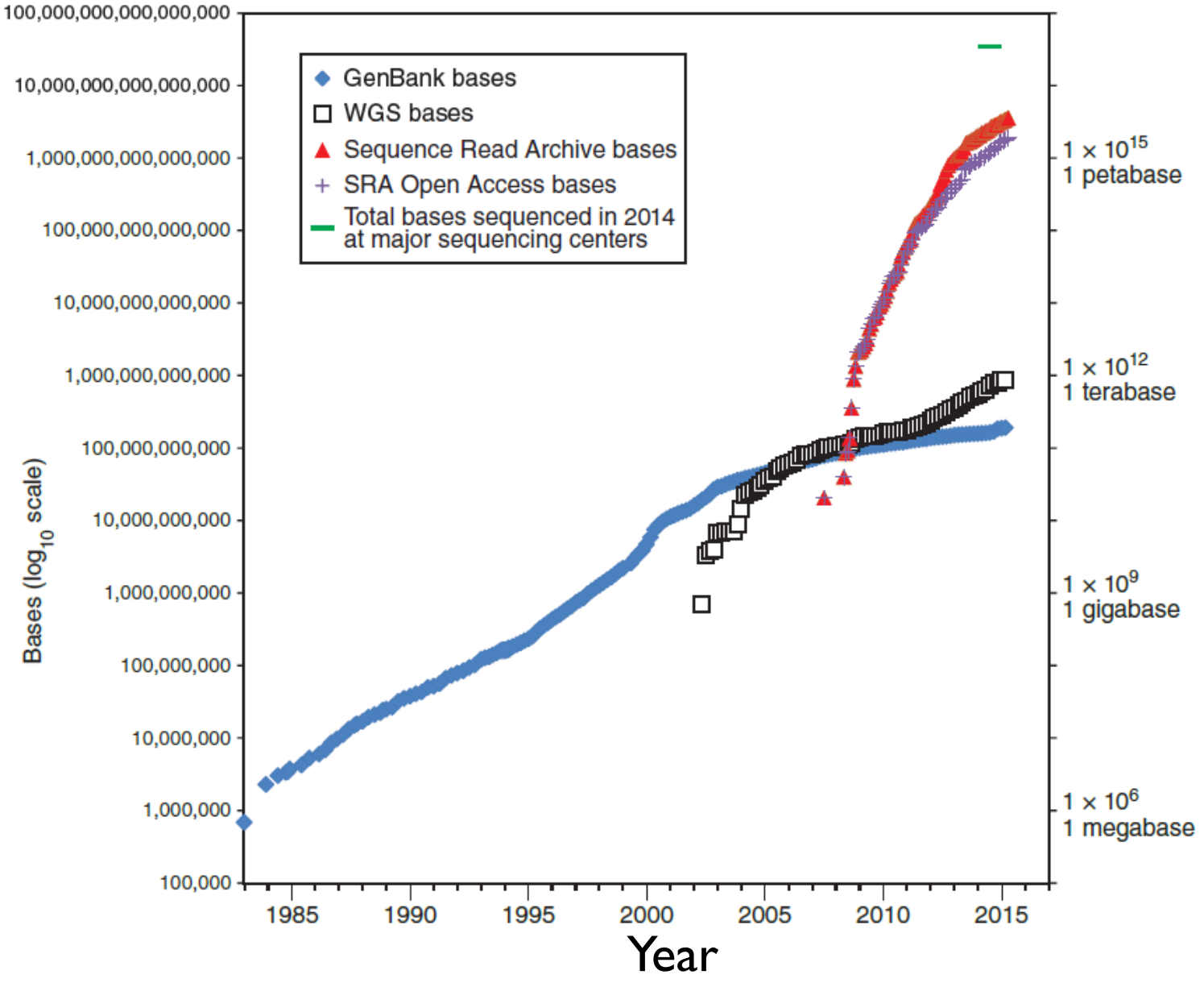
2016: 100,000 genomes, >>200,000 exomes

Growth of DNA sequence in repositories



B&FG 3e
Fig. 2-3
Page 22

Growth of DNA sequence in repositories



B&FG 3e
Fig. 2-3
Page 22

Contents of a human genome

When we sequence a typical genome we find the following:

- 100 billion base pairs sequenced (3 billion x 30)
- 4 million single nucleotide variants
- 600,000 indels and other structural variants
- 20,000 protein-coding genes
- 11,000 nonsynonymous variants
- 100 disease-associated variants
- 0-24 homozygous gene knockouts

Outline

The human genome

Epigenetics and the human epigenome

- The ENCODE Project
- Roadmap Epigenomics Project

Epigenetic markers for drug discovery

Definition of epigenomics

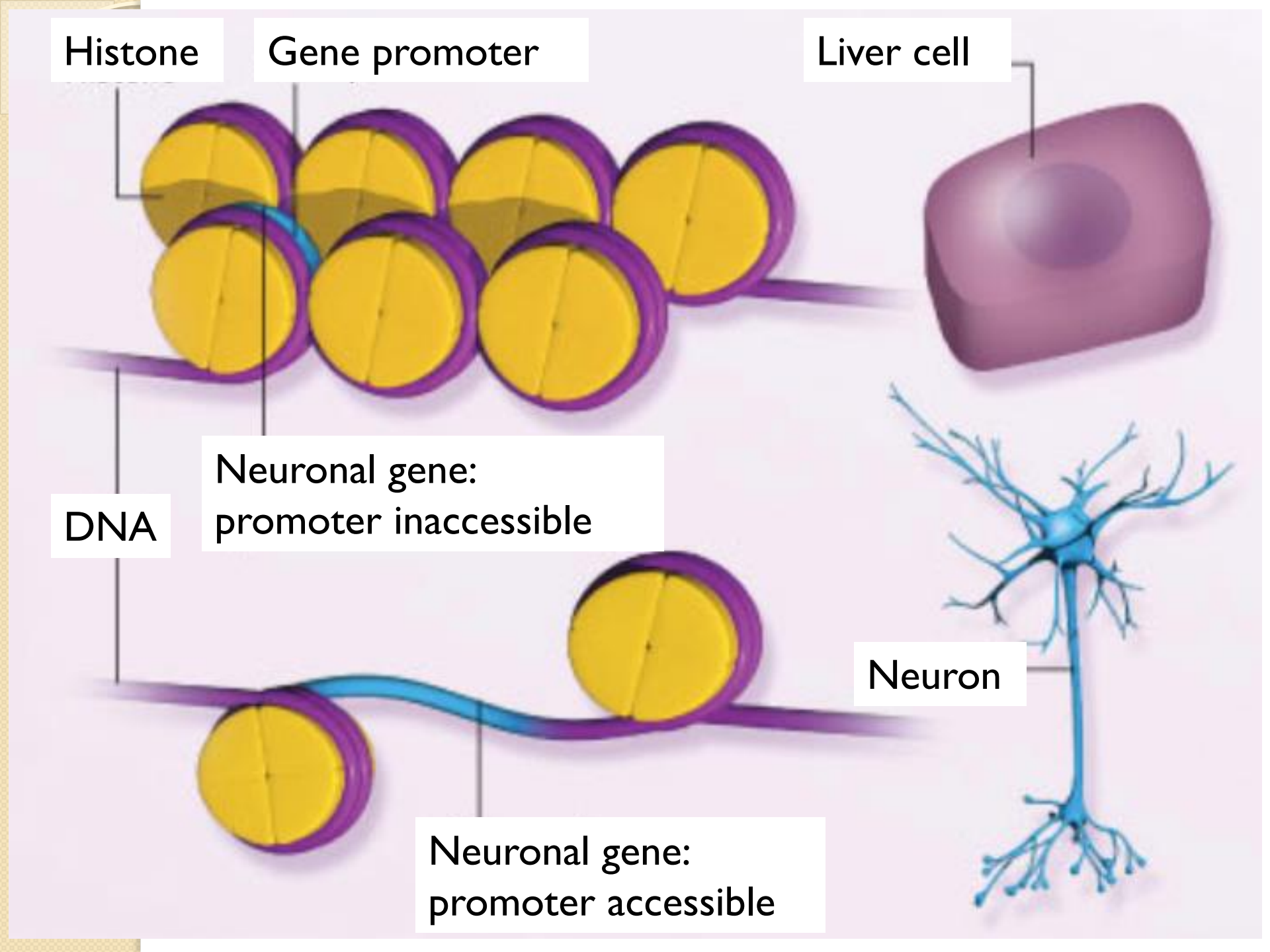
An epigenetic mark is a physical modification of genomic DNA, including methylation of DNA, acetylation of histones, and alterations to chromatin structure. It does not involve changes to DNA sequence itself.

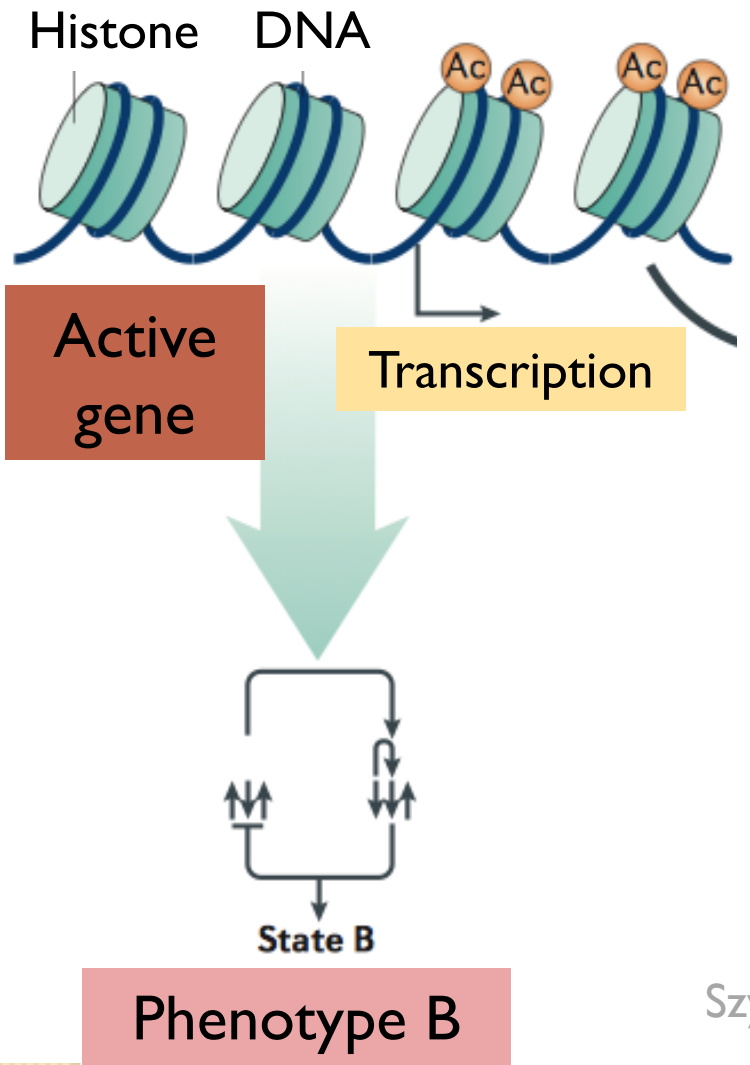
The epigenome is the collection of epigenetic modifications.

Epigenetics concepts

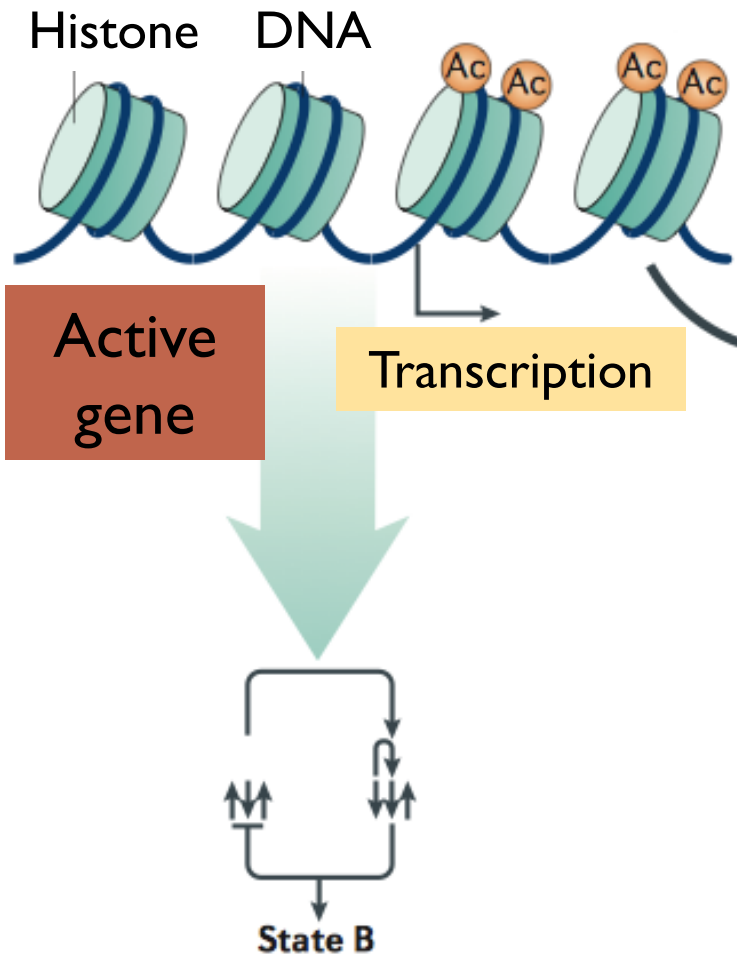
Modifications to DNA (e.g. methylation) and histones (e.g. histone acetylation, phosphorylation, ubiquitination) affect many processes:

- transcription
- repair
- replication
- recombination


Histone
Gene promoter
Liver cell
Neuronal gene:
promoter inaccessible
DNA
Neuron
Neuronal gene:
promoter accessible



Szyf M (2016). Nat. Rev. Drug Discovery 14:461

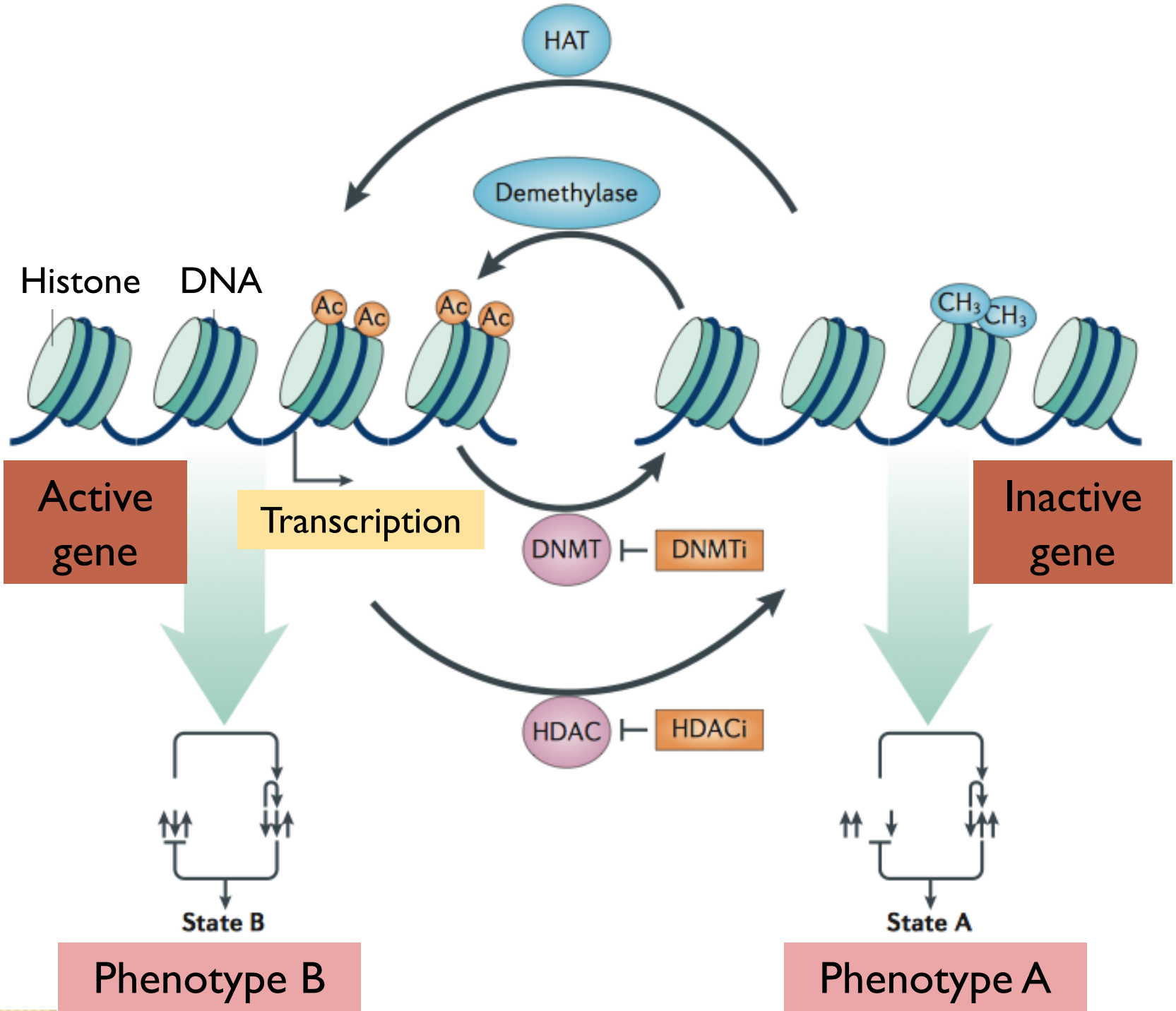


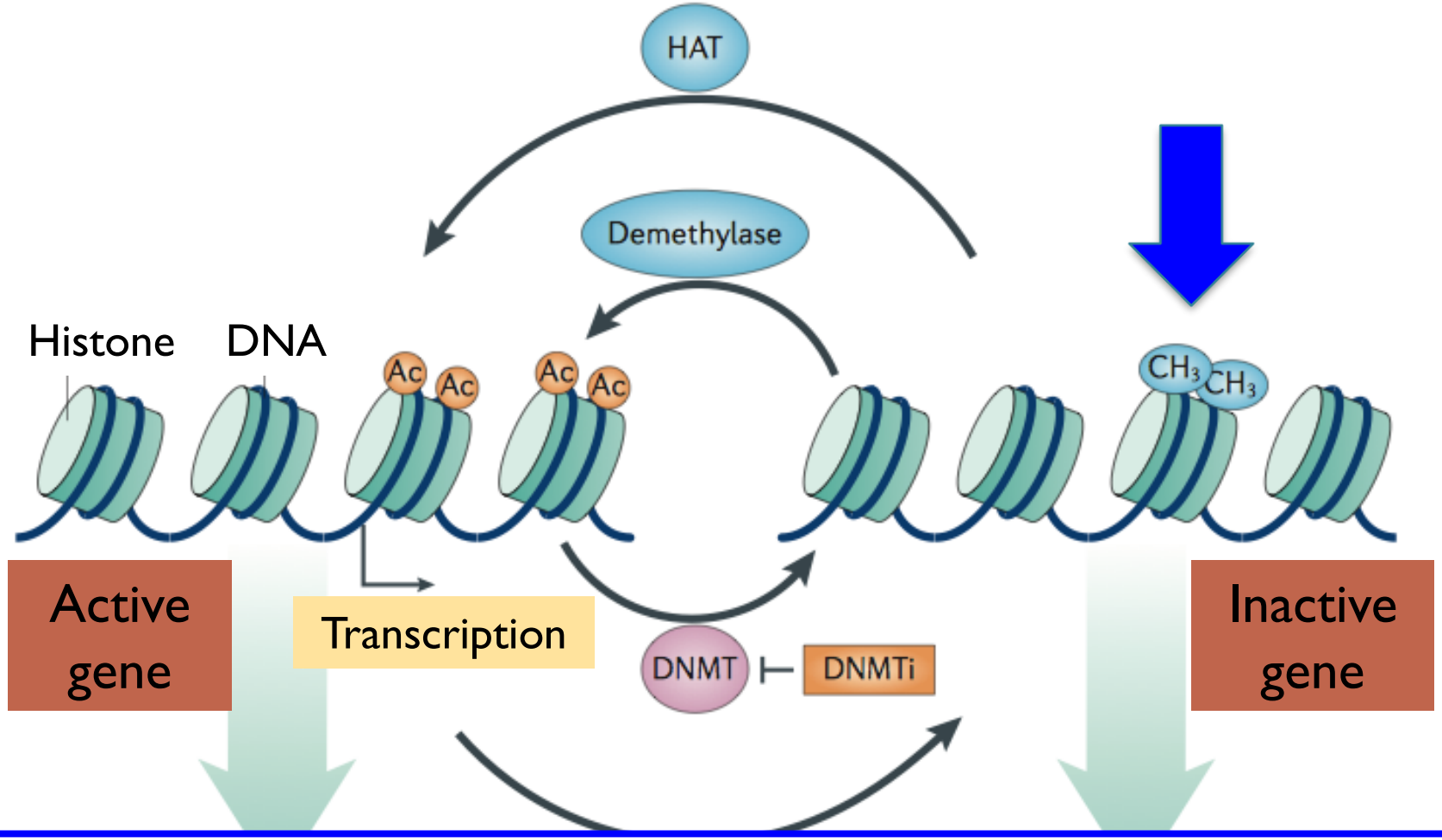
Chromatin:

- Nuclear complex composed of DNA and histones
- Heterochromatin (repressed)
- Euchromatin (open, transcribed)

The nucleosome:

- basic unit of chromatin
- 147 base pairs of DNA wrapped around a histone octamer


HAT
Demethylase
Histone
DNA
Ac
Ac
Ac
Ac
CH3
CH3
Active gene
Transcription
DNMT
DNMTi
Inactive gene
HDAC
HDACi
State B
State A
Phenotype B
Phenotype A



DNA methylation:

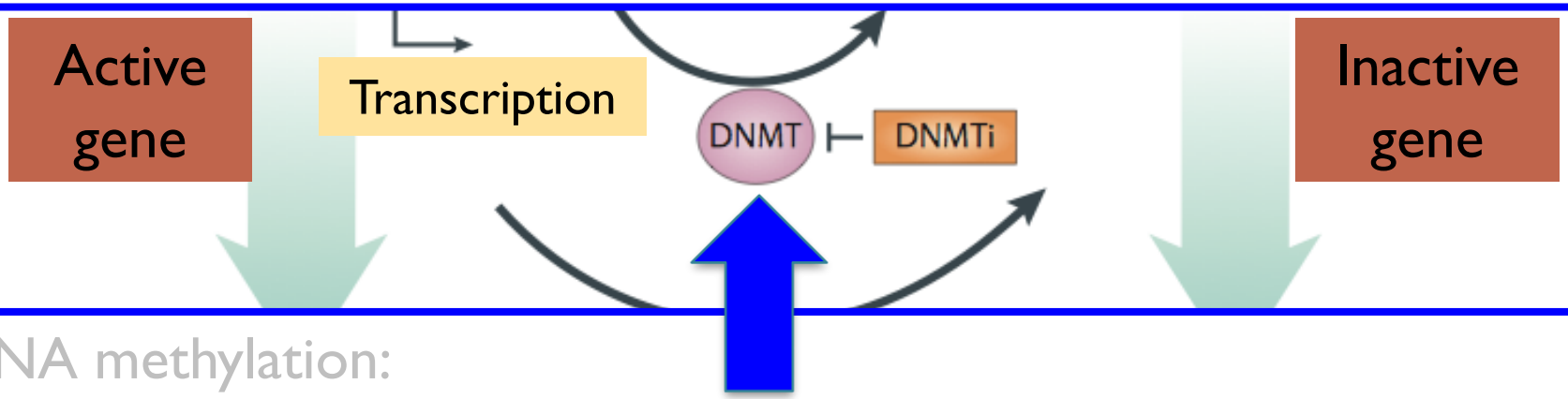
- addition of a methyl group to the 5' carbon of cytosine in CpG dinucleotides
- CpG islands are rich in CpG, present in 70% of gene promoter regions, typically unmethylated in normal cells (changes in cancer)
- CpG islands rich in telomeres, centromeres

DNA methytransferases (DNMTs):

- DNMT1 is a maintenance methyltransferase
- DNMT3A, DNMT3B function in embryogenesis
- Somatic mutations in *DNMT3A* in 25% of acute myeloid leukemia

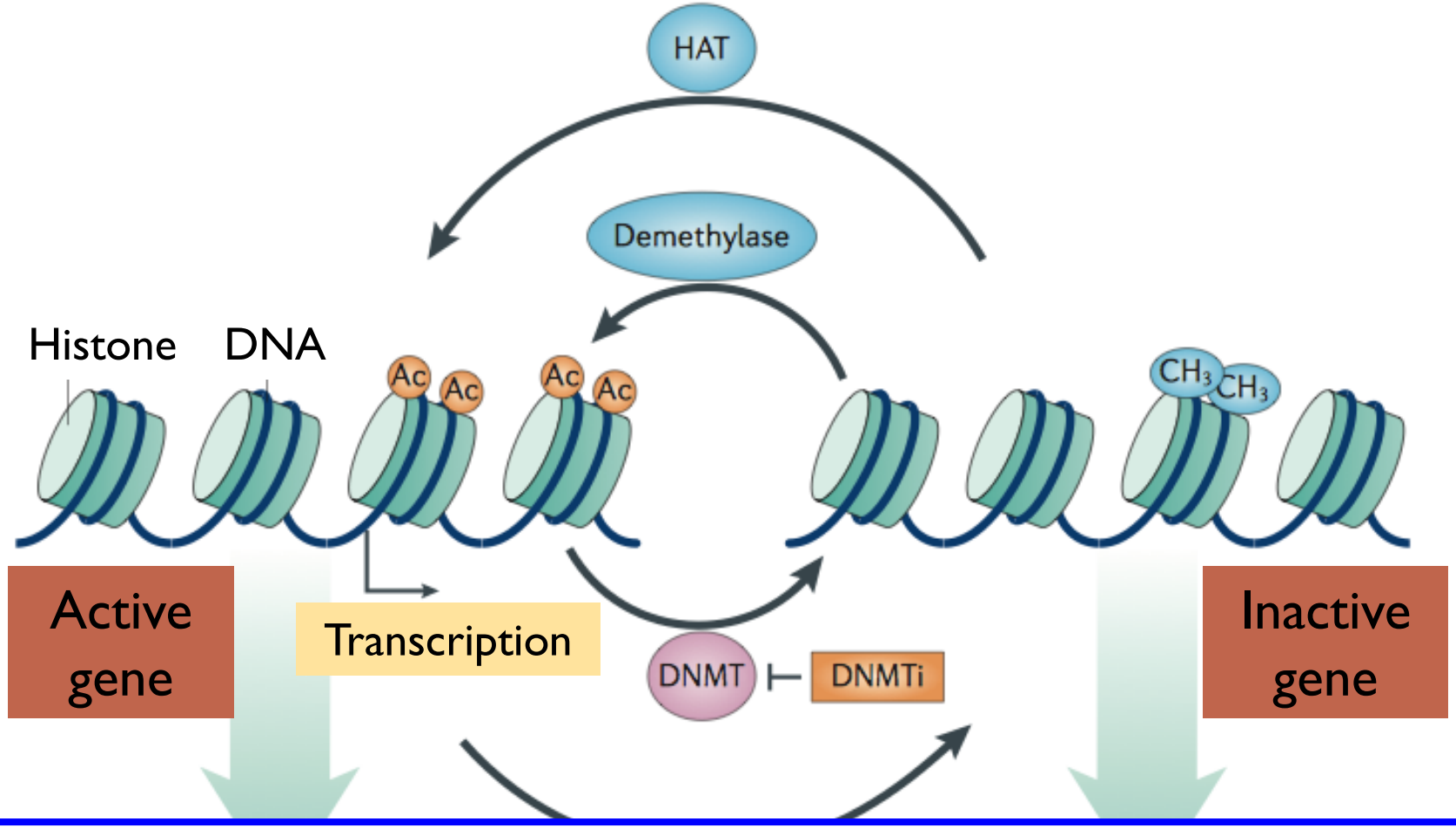
Methyl-binding proteins (MBDs):

- MBD1, MBD2, MBD3, MeCP2
- Mutations in *MECP2* cause Rett syndrome



DNA methylation:

- addition of a methyl group to the 5' carbon of cytosine in CpG dinucleotides
- CpG islands are rich in CpG, present in 70% of gene promoter regions, typically unmethylated in normal cells (change in cancer)
- CpG islands rich in telomeres, centromeres



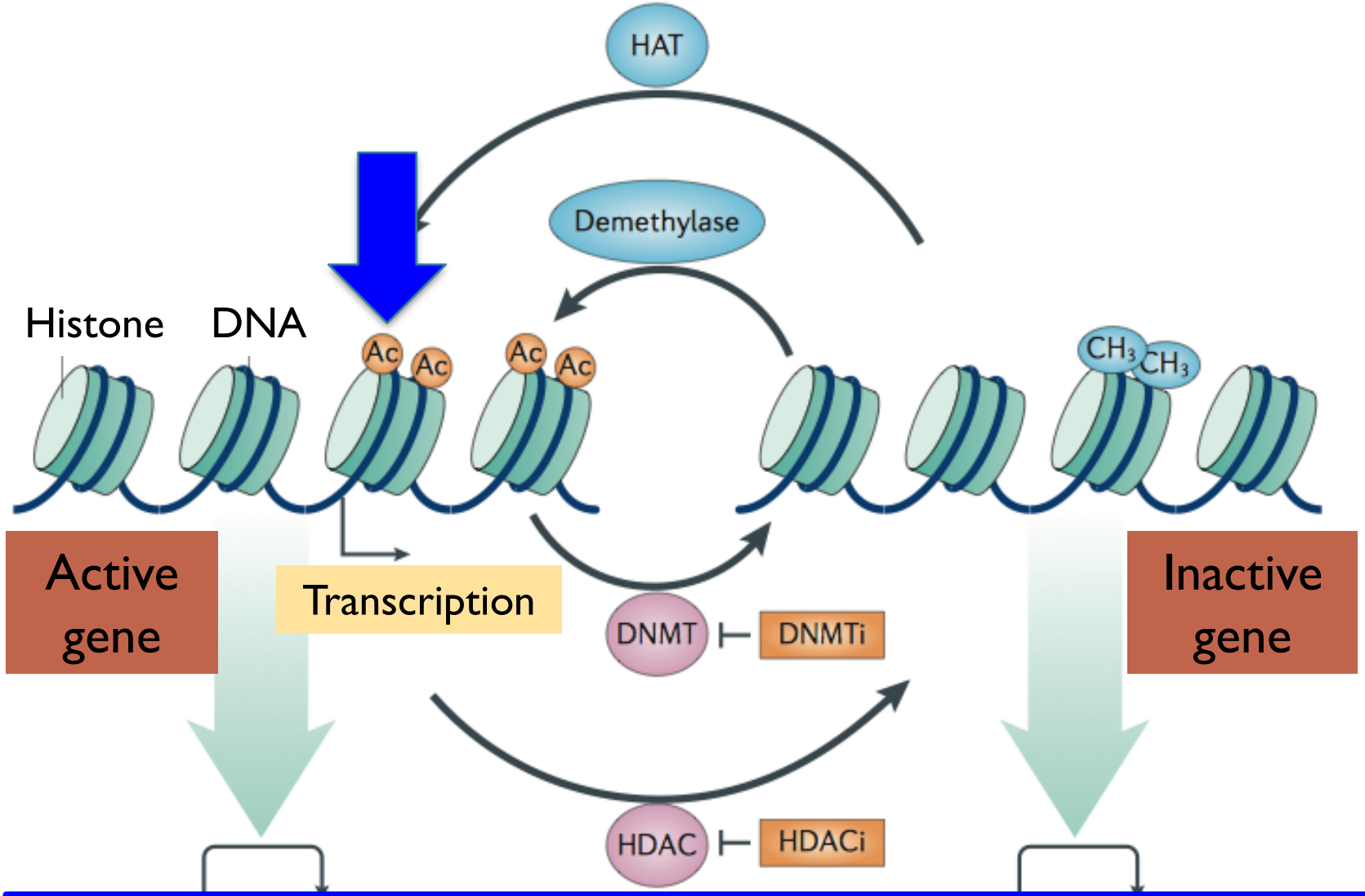
DNA hydroxymethylcytosine (5hmC) also occurs:

- Catalyzed by TET enzymes
- Mutations in *TET1* gene cause cancer (e.g. lymphocytic leukemia)
- Mutations in *TET2* gene cause hematological malignancies

Modifications of histone tails

The main histones are H2A, H2B, H3, and H4. These are highly conserved evolutionarily. Several different enzymes modify histones to alter their function.

- Histone acetyltransferases (HATs)
- Histone deacetylases (HDACs)
- Histone methyltransferases
- Histone demethylases



Histone acetylation

- promotes an open conformation
- often at promoters of actively transcribed genes
- many cancers include mutations in acetyltransferases

Chromatin remodeling complexes

These complexes package chromatin. There are four main types. Each uses ATP hydrolysis to mobilize nucleosomes:

- SWI/SNF
- ISWI
- CHD
- INO80

Many genes encoding protein members of these complexes are mutated in cancer, intellectual disability, autism spectrum disorder, and other conditions.

Outline

The human genome

Epigenetics and the human epigenome

- The ENCODE Project
- Roadmap Epigenomics Project

Epigenetic markers for drug discovery

The ENCODE Project

- The Encyclopedia of DNA Elements (ENCODE) project was designed to catalog functional elements embedded in genomic DNA. Main conclusions include:

  - The genome is pervasively transcribed.
  - >80% of the genome is functional.
  - 56% of the genome is enriched for histone modifications.
  - Of the 80% functional region of the genome, half is RNA and histone elements; the other half includes DNase hypersensitivity sites, transcription factor bindings sites, other regulatory sites.

ENCODE data are available, both raw and processed.

ARTICLE

OPEN

doi:10.1038/nature14248

Integrative analysis of 111 reference human epigenomes

Roadmap Epigenomics Consortium†, Anshul Kundaje[1,2,3]*, Wouter Meuleman[1,2]*, Jason Ernst[1,2,4]*, Misha Bilenky[5]*, Angela Yen[1,2], Alireza Heravi-Moussavi[5], Pouya Kheradpour[1,2], Zhizhuo Zhang[1,2], Jianrong Wang[1,2], Michael J. Ziller[2,6], Viren Amin[7], John W. Whitaker[8], Matthew D. Schultz[9], Lucas D. Ward[1,2], Abhishek Sarkar[1,2], Gerald Quon[1,2], Richard S. Sandstrom[10], Matthew L. Eaton[1,2], Yi-Chieh Wu[1,2], Andreas R. Pfenning[1,2], Xinchen Wang[1,2,11], Melina Claussnitzer[1,2], Yaping Liu[1,2], Cristian Coarfa[7], R. Alan Harris[7], Noam Shoresh[2], Charles B. Epstein[2], Elizabeta Gjoneska[2,12], Danny Leung[8,13], Wei Xie[8,13], R. David Hawkins[8,13], Ryan Lister[9], Chibo Hong[14], Philippe Gascard[15], Andrew J. Mungall[5], Richard Moore[5], Eric Chuah[5], Angela Tam[5], The
Annaick Carles[18], Jesse R. Dixo
Daofeng Li[21], Rebecca Lowdon
Nisha Rajagopal[8,13], Pradipta Ra
Michael Stevens[21,42], Robert E.
Philip L. De Jager[2,23,27], Peggy J
Marco A. Marra[5,32], Michael T.
Wei Wang[8], Robert A. Waterlan
Joseph F. Costello[14]§, Joseph R.
John A. Stamatoyannopoulos[10]

111 reference human epigenomes profiled for:

- histone modification patterns
- DNA accessibility
- DNA methylation
- RNA expression

Example of reference epigenome data sets

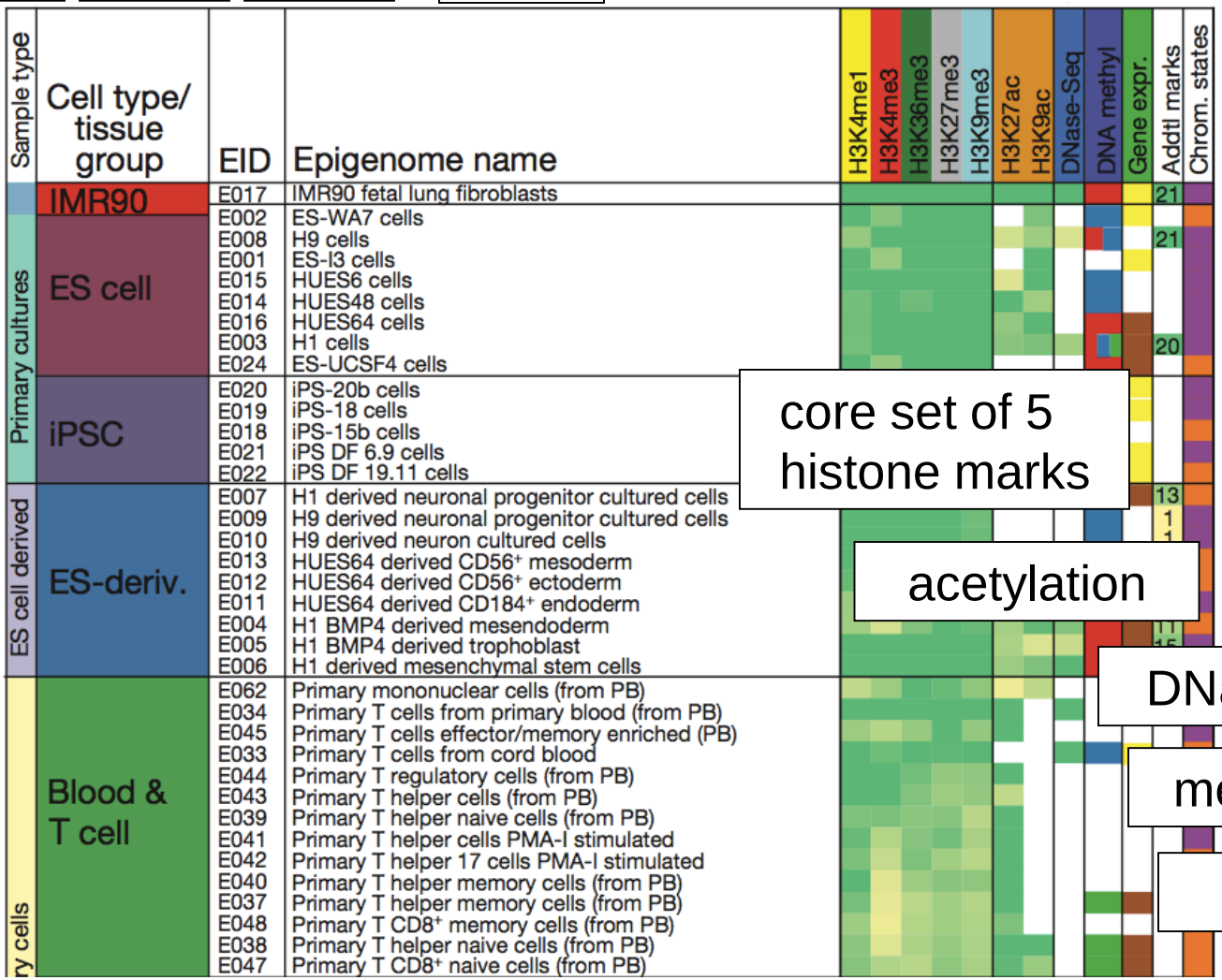
type

loca-tion

EID

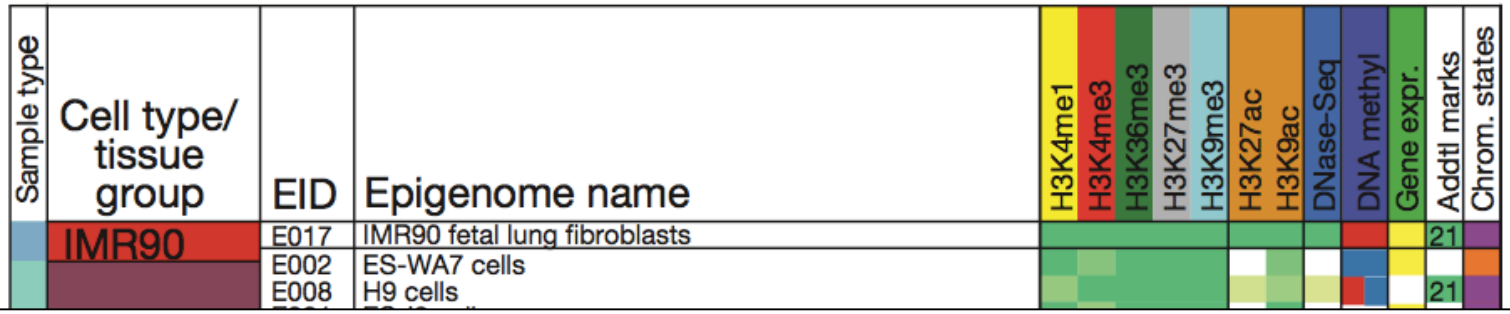
name

normalized strand cross-correlation quality scores (NSC)



Roadmap Epigenomics Consortium *et al. Nature* **518**, 317-330 (2015)

Example of reference epigenome data sets



H3K4me3 (histone H3 lysine 4 trimethylation) promoter regions
H3K4me1 (H3 lysine 4 monomethylation) enhancer regions
H3K36me3 (H3 lysine 36 trimethylation) transcribed regions
H3K27me3 (H3 lysine 27 trimethylation) Polycomb repression
H3K9me3 (H3 lysine 9 trimethylation) heterochromatin

H3K27ac (acetylation) activation of enhancer regions
H3K9ac (acetylation) activation of promoter regions
Dnase hypersensitivity accessible chromatin
DNA methylation active gene expression

Epigenomic information across tissues and marks

3.5 Mb region of chromosome 9

127 reference epigenomes

Promoters (red vertical lines) are constitutive, enhancers (yellow) are highly dynamic

Outline

The human genome

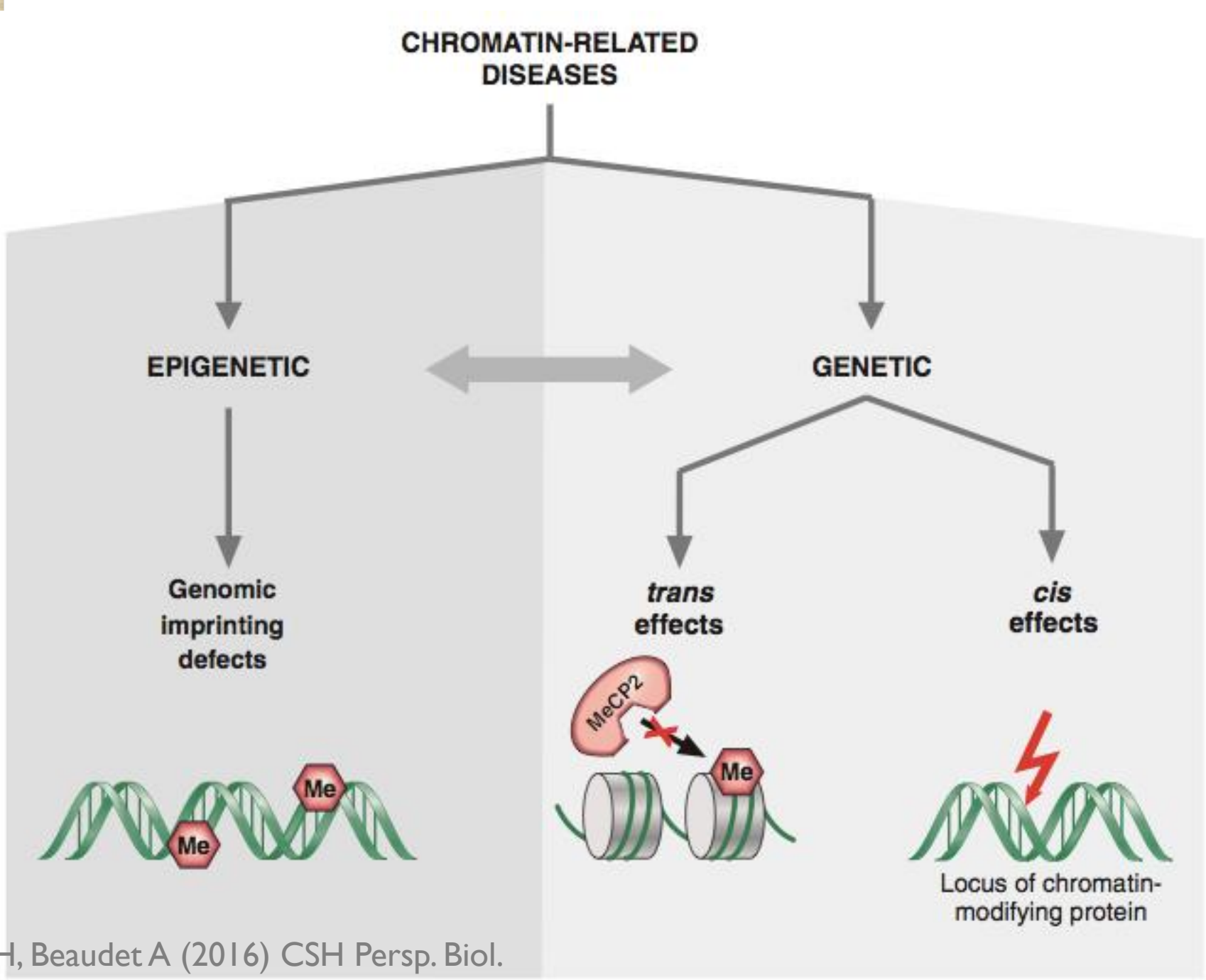
Epigenetics and the human epigenome
- The ENCODE Project
- Roadmap Epigenomics Project

Epigenetic markers for drug discovery

Epigenetics and human disease: two approaches

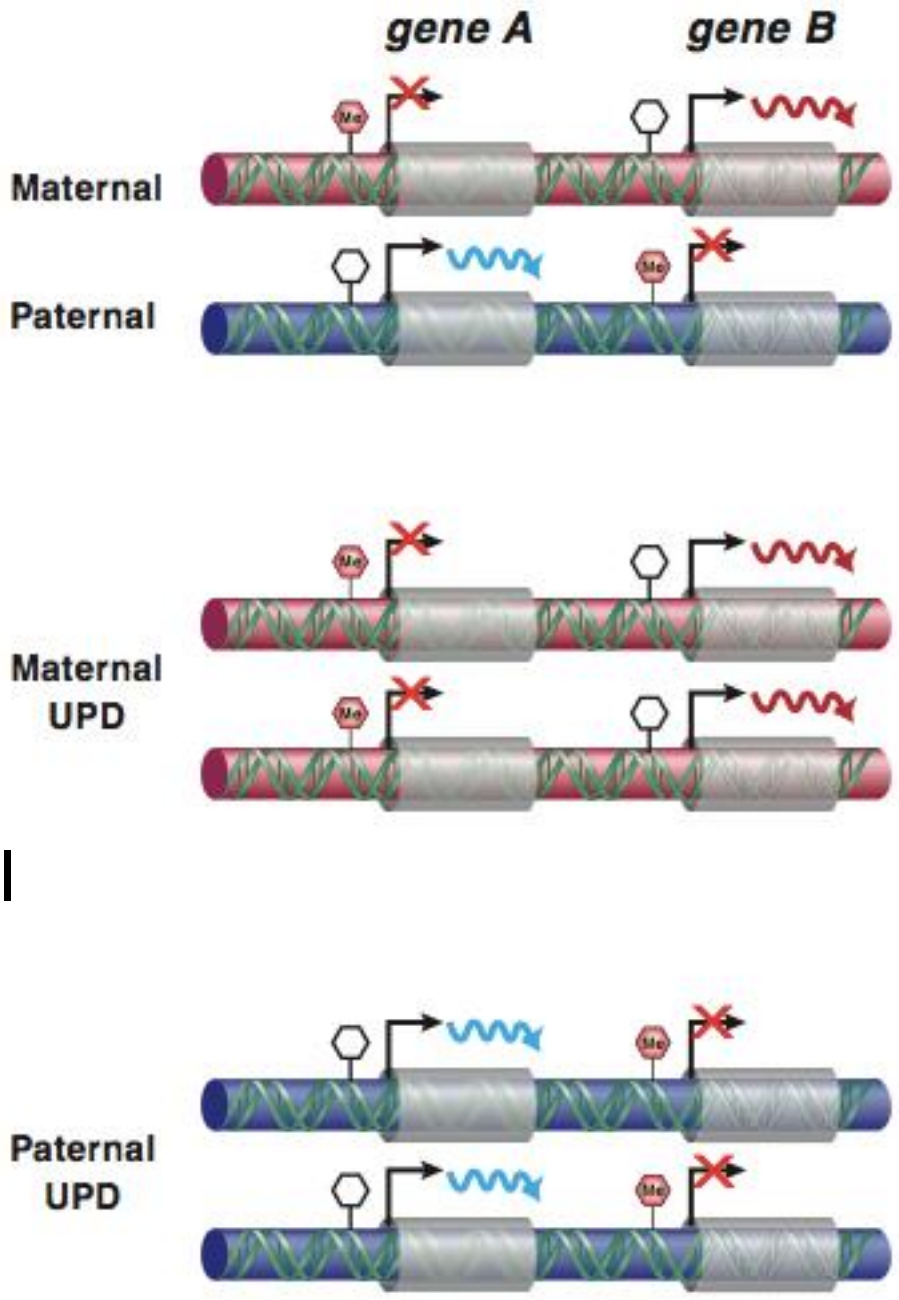
1. Search for changes in the epigenome (e.g. altered methylation patterns) that occur in disease.

2. Search for mutations in protein-coding genes that function in epigenetics. These may act in *cis* or *trans*.

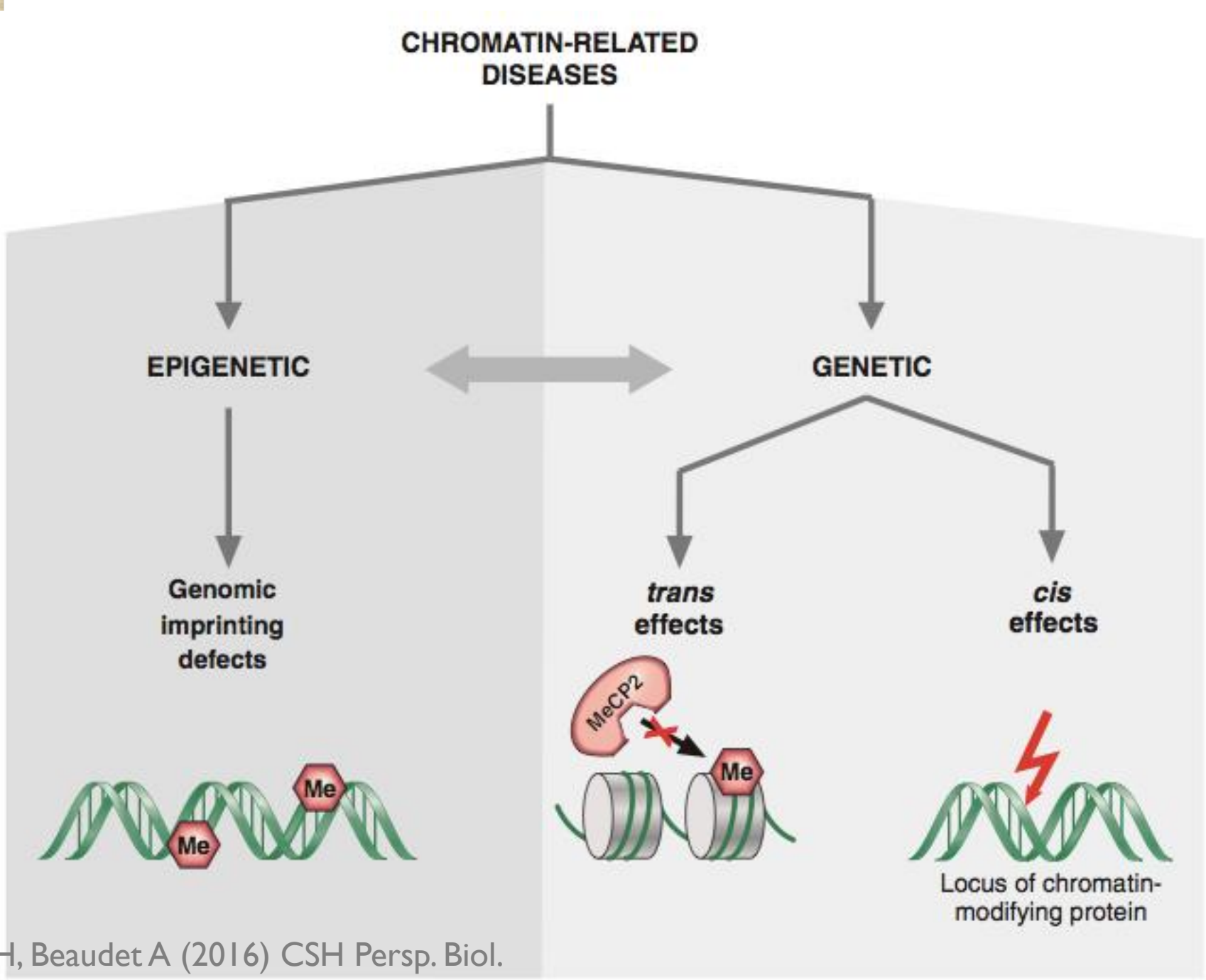


Zoghbi H, Beaudet A (2016) CSH Persp. Biol.

Uniparental disomy



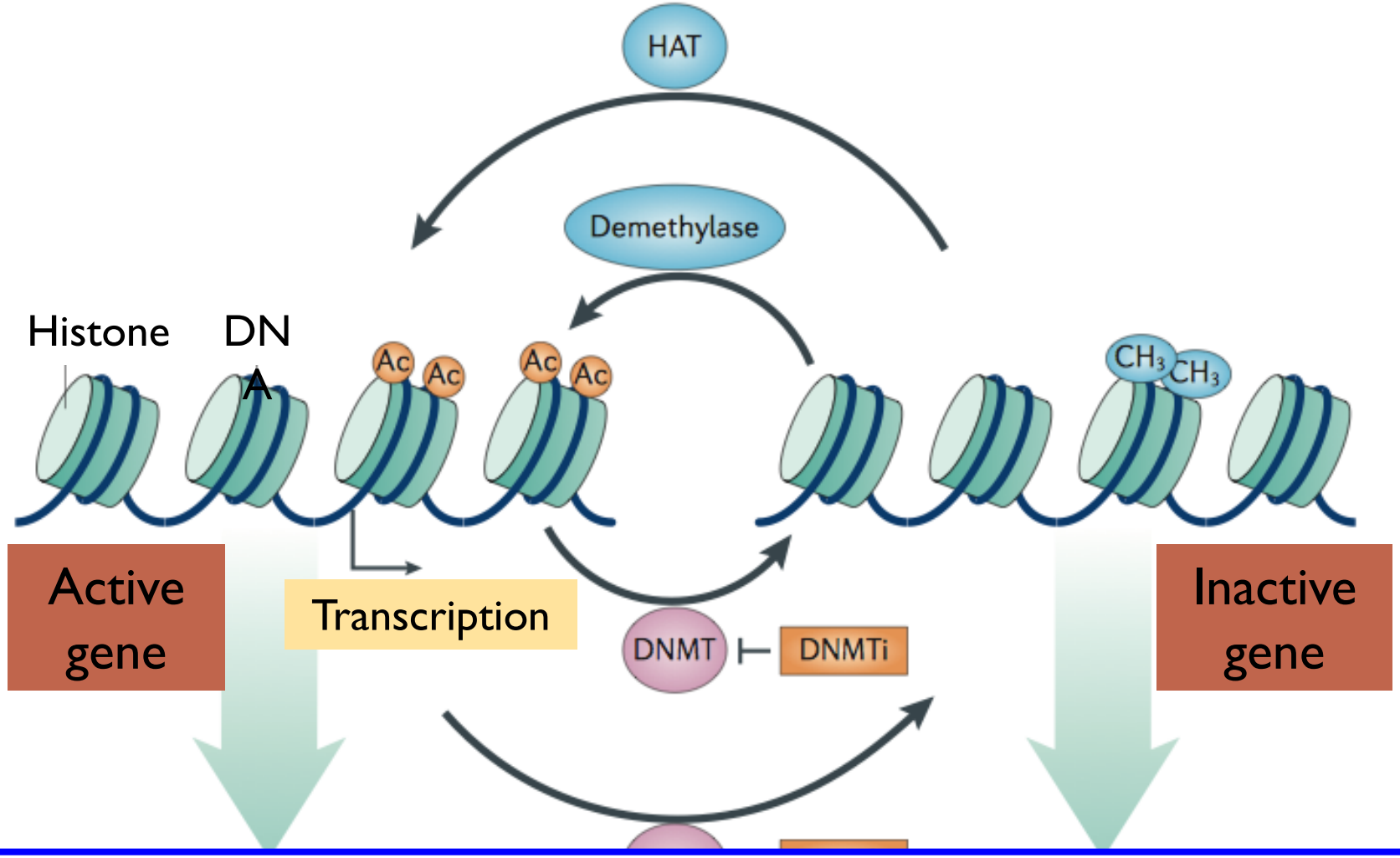
Zoghbi H, Beaudet A (2016) CSH Persp. Biol.



Zoghbi H, Beaudet A (2016) CSH Persp. Biol.

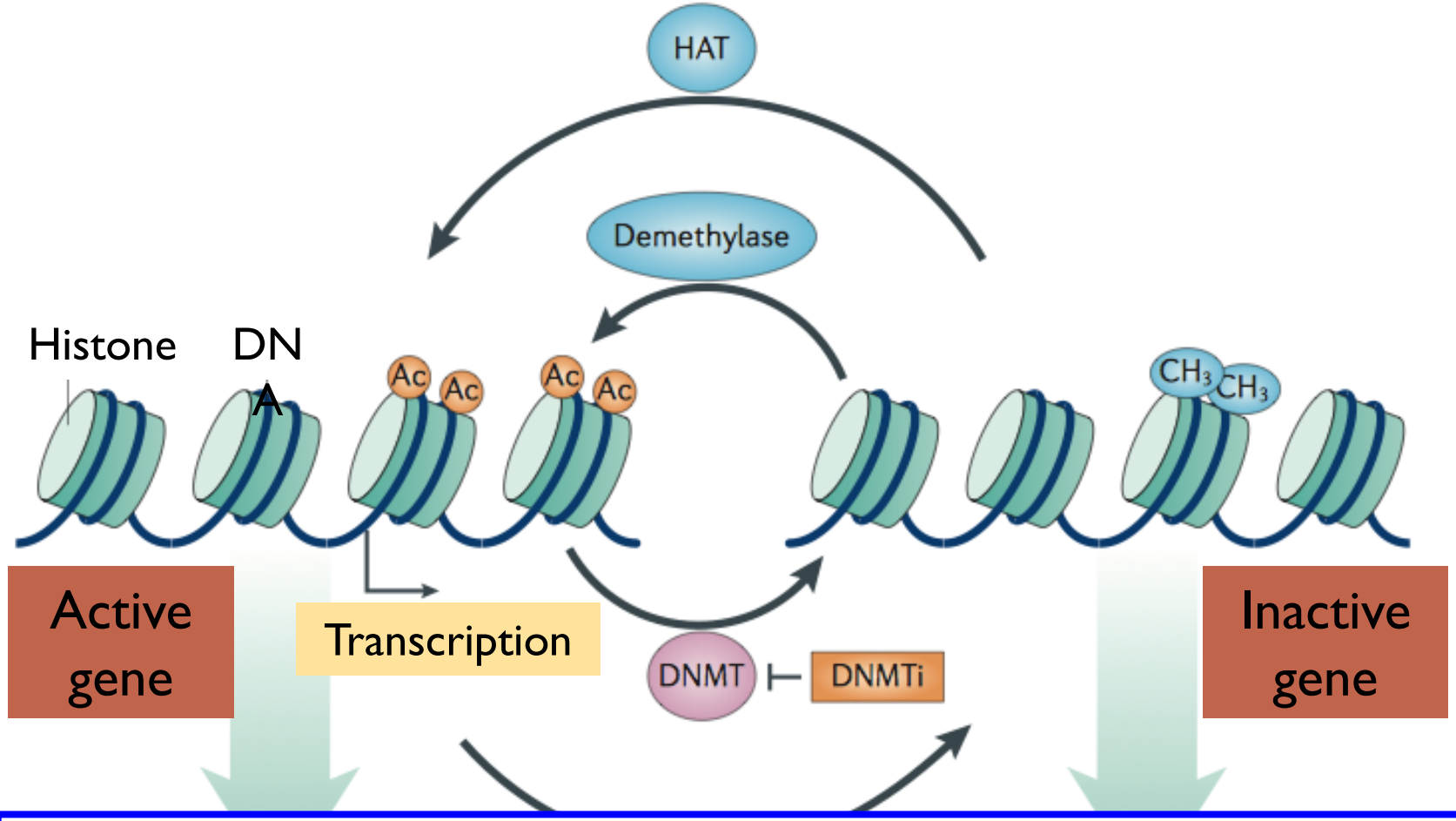
Examples of epigenetic drug targets

- The DNA repair gene *MGMT* is hypermethylated in gliomas and glioblastomas. Extent of hypermethylation predicts clinical response to treatment with alkylating agents (e.g. temozolomide).
- DNMT inhibitors (vidaza, decitabine) are FDA approved to treat myelodysplastic syndrome.
- HDAC inhibitors (orinostat, romidepsin), are approved to treat a rare cutaneous T-cell lymphoma.



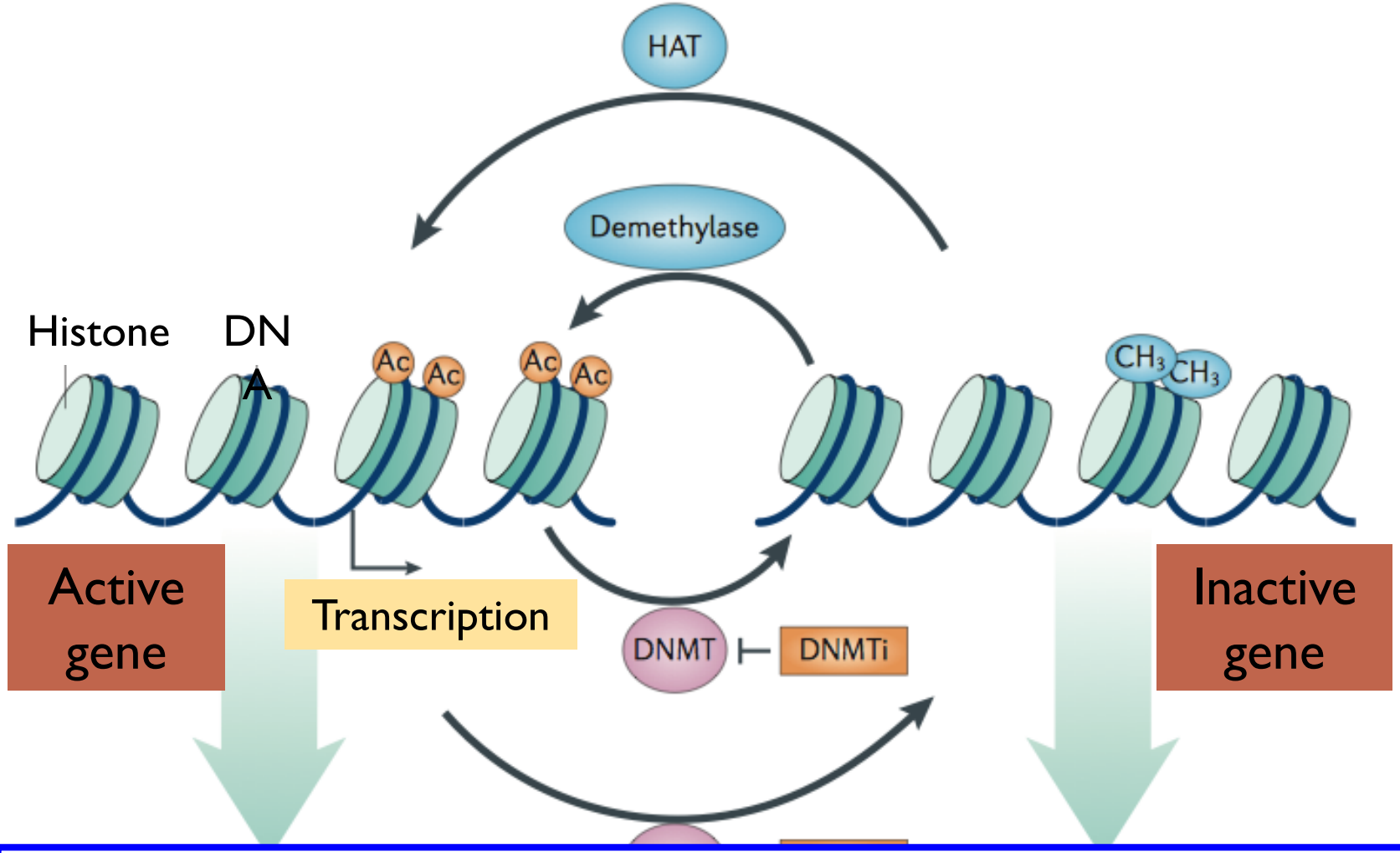
Thomas Paul (Pfizer) will discuss validated chemical probes for understanding responders to epigenetic drugs.

- Imbalances in the SWI/SNF subunits render cells tumorigenic.
- Role of lysine-specific demethylase (LSD1) in small lung cell cancer.



Robert Georgantas (Abbvie) will discuss the use of DNA methylation as biomarkers for drug effects.

- DNA methylation of the *IL1b* gene predicts response of Hepatitis C patients to peg-interferon.
- DNA methylation patterns reflect pharmacologic action of drugs.



Elizabeth Thomas (Scripps Research Institute) will discuss histone acetylation and HDAC inhibitors. She will focus on epigenetic effects in Huntington's disease and Friedreich ataxia.