

# Motivation: epigenetics & CpG islands

Ben Langmead



JOHNS HOPKINS

WHITING SCHOOL  
*of* ENGINEERING

Department of Computer Science

You are free to use these slides. If you do, please sign the guestbook ([www.langmead-lab.org/teaching-materials](http://www.langmead-lab.org/teaching-materials)), or email me ([ben.langmead@gmail.com](mailto:ben.langmead@gmail.com)) and tell me briefly how you're using them. For original Keynote files, email me.

# Picking up signals

So far, we've focused on how to stitch fragments of evidence into longer units, i.e. genomes

Equipped with genomes, we can ask more questions:

Where are the genes?

Where/what is the *functional* DNA?

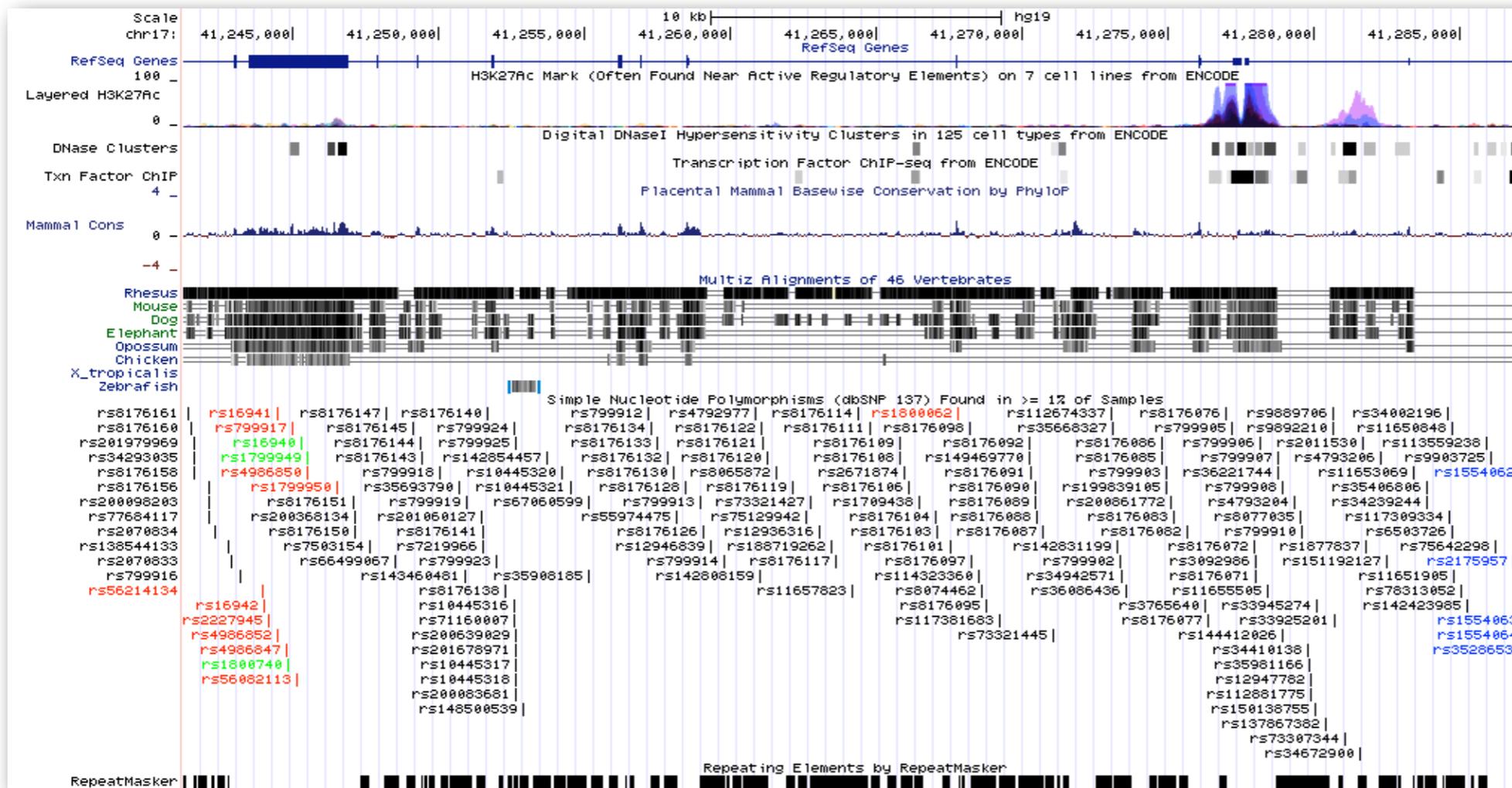
What's different about the DNA in different tissues?

In what abundance do we find various molecules?

What differences exist between individuals?

# Picking up signals

Through many experiments, we know much more about the genome than just its DNA sequence:



Experimentally observed products, e.g. messenger RNAs

Epigenetic marks

Sequence conservation among related species

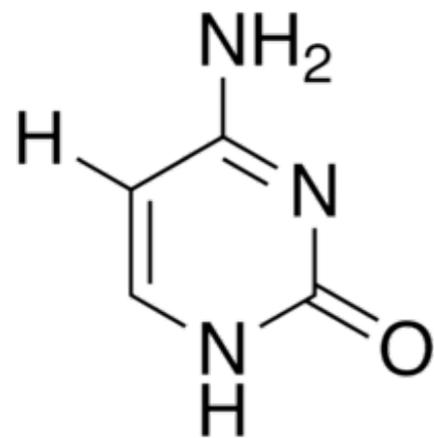
Sites that vary across individuals

40 K nt region of chromosome 17  
<http://genome.ucsc.edu/cgi-bin/hgTracks>

# CpG Islands

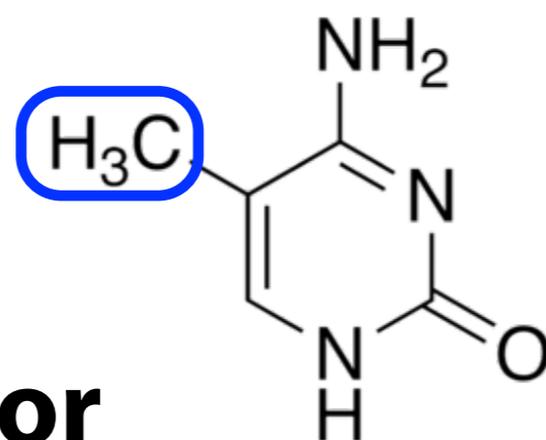
A signal we can discern from genome sequence alone: CpG islands

Dinucleotide "CG" (AKA "CpG") is special because the C can possibly have a *methyl group* attached



Unmethylated

**or**



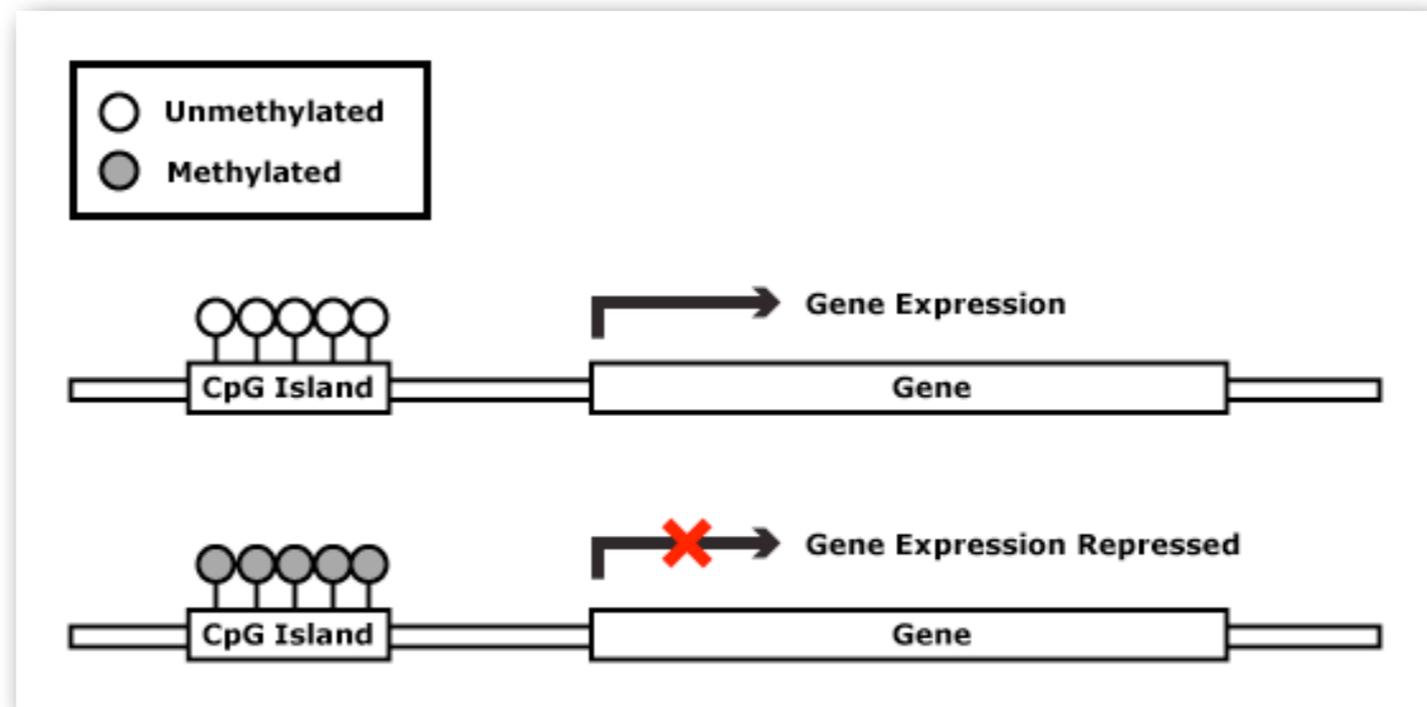
Methylated

Proteins involved in gene expression can be repelled or attracted by the methyl group

# CpG Islands

*CpG island*: part of the genome where CG occurs particularly frequently

CpG islands usually regulate expression of nearby genes

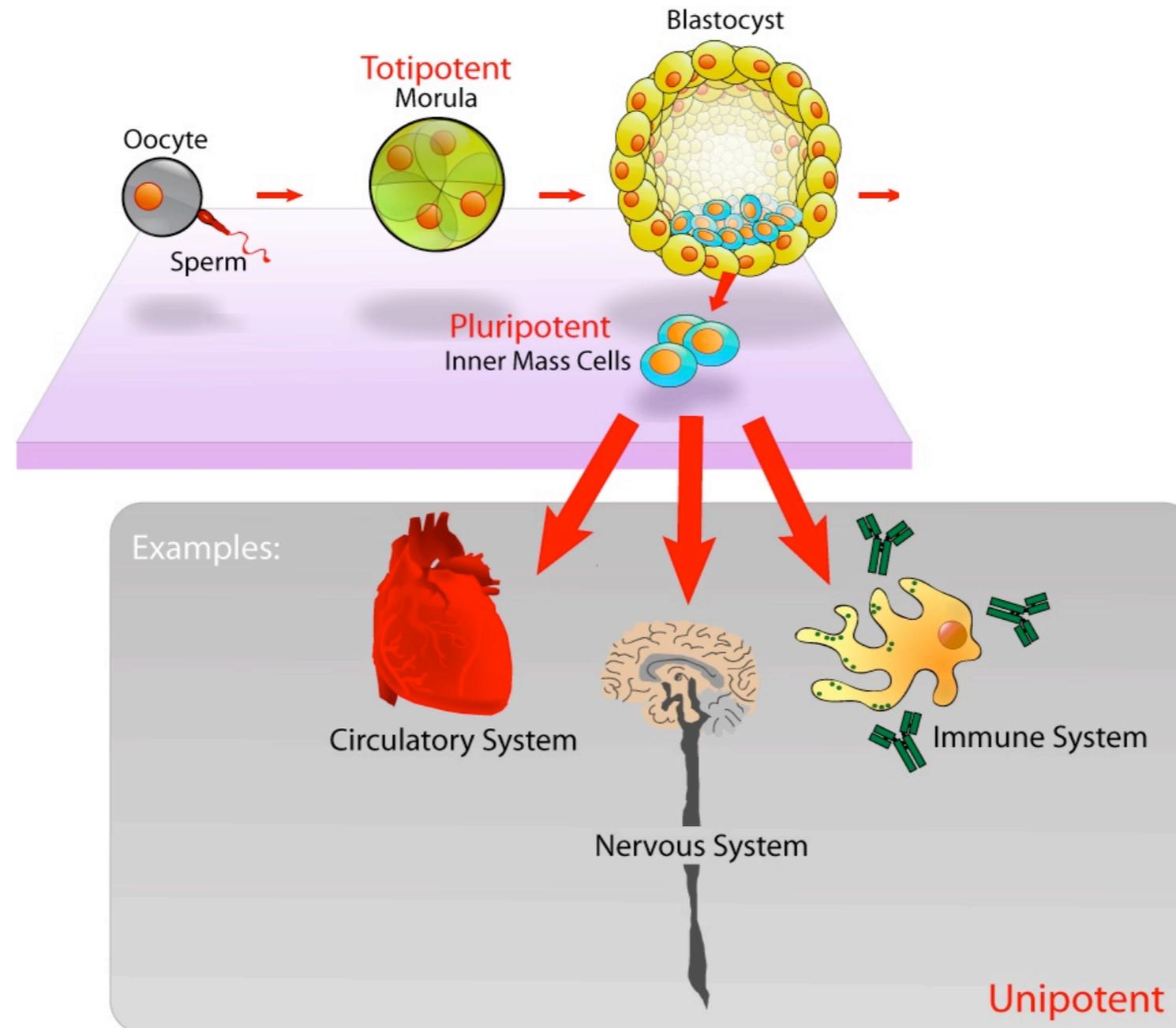


[http://missinglink.ucsf.edu/lm/genes\\_and\\_genomes/methylation.html](http://missinglink.ucsf.edu/lm/genes_and_genomes/methylation.html)

Cells from different tissues have different patterns of CpG methylation, in turn giving them different gene expression profiles

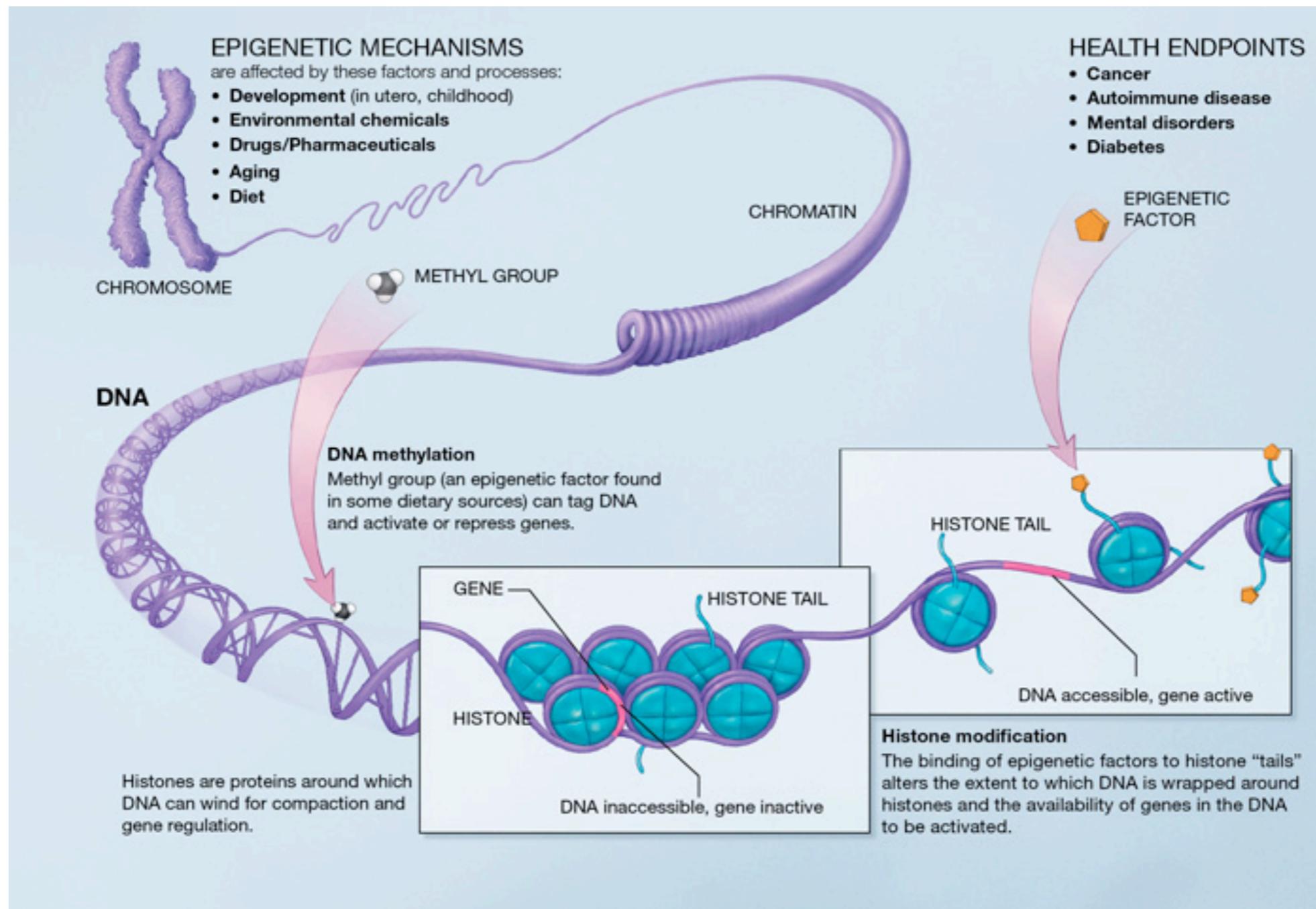
Key *epigenetic* phenomenon

# Background: Epigenetics



[http://en.wikipedia.org/wiki/File:Stem\\_cells\\_diagram.png](http://en.wikipedia.org/wiki/File:Stem_cells_diagram.png)

# Background: Epigenetics



[http://upload.wikimedia.org/wikipedia/commons/d/dd/Epigenetic\\_mechanisms.jpg](http://upload.wikimedia.org/wikipedia/commons/d/dd/Epigenetic_mechanisms.jpg)

# Background: Epigenetics

Study of how characteristics are inherited across generations *without* changes to the DNA sequence itself

How does a heart cell know it's a heart cell?

How does a calico cat get its splotches?

Epigenetic changes are important in various diseases: Fragile X, Rett, and Angelman syndromes, cancer



[http://en.wikipedia.org/wiki/Calico\\_cat](http://en.wikipedia.org/wiki/Calico_cat)