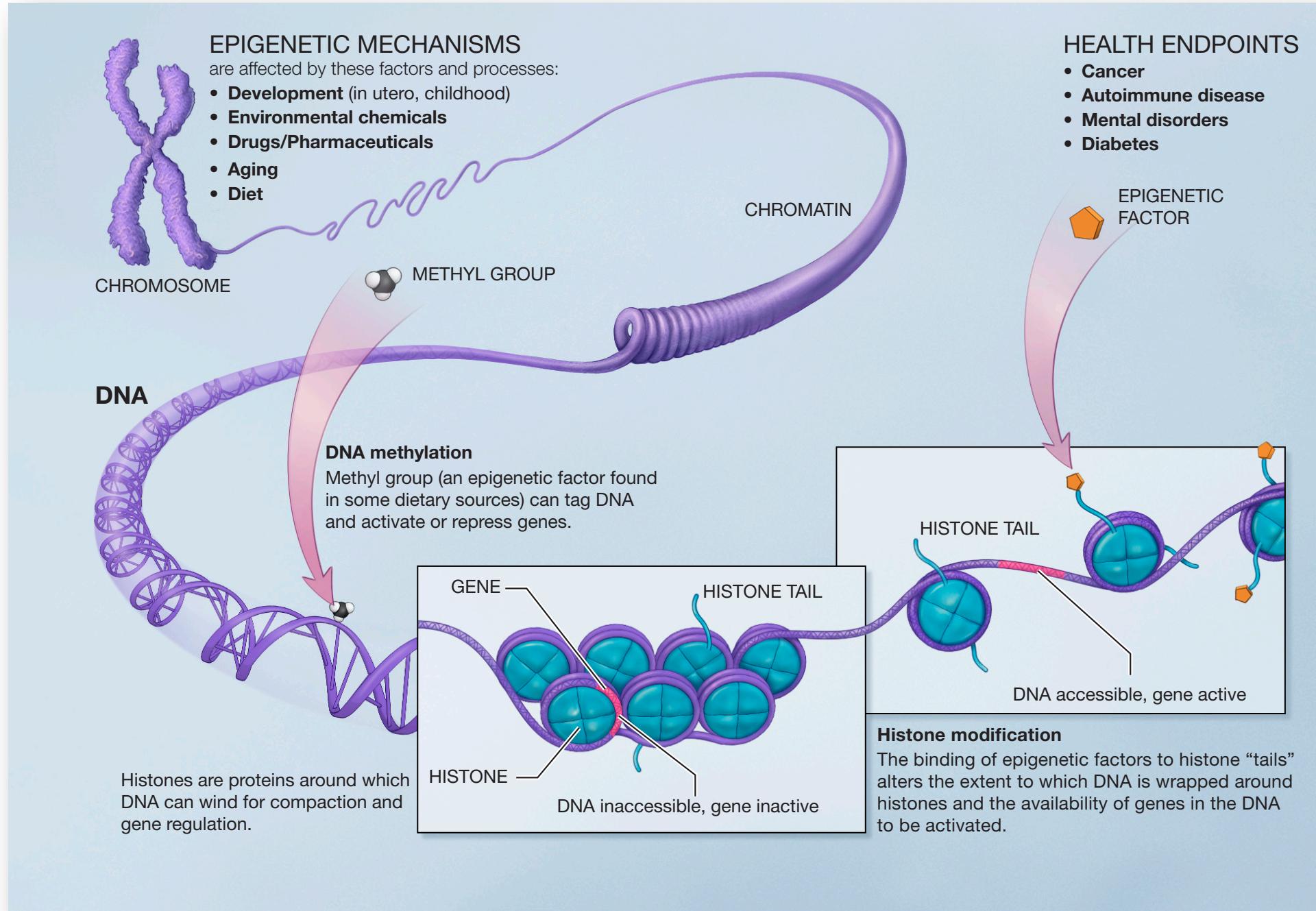


Introduction to Methylation

Epigenomics Workshop 2020

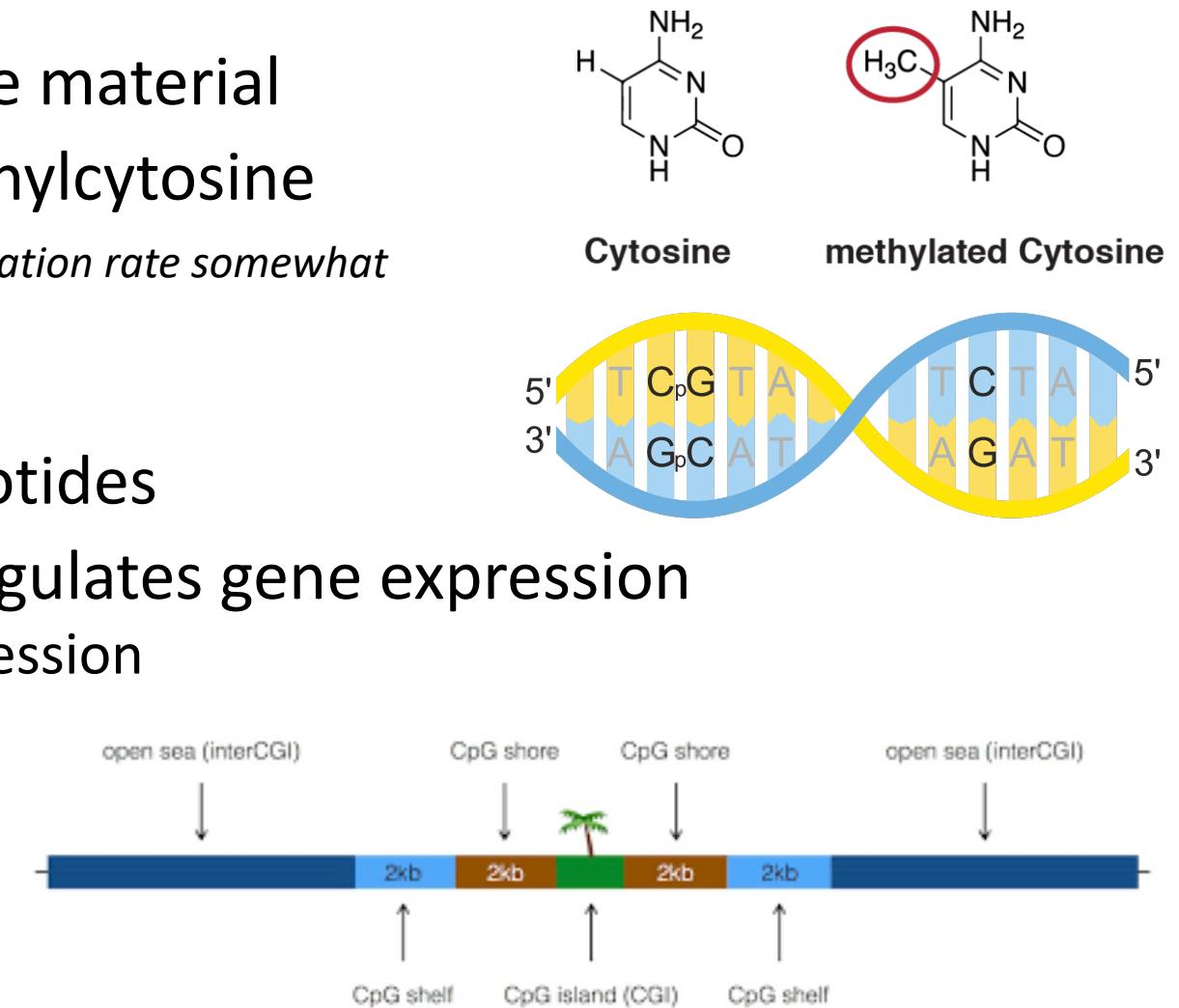


A brief history of DNA methylation

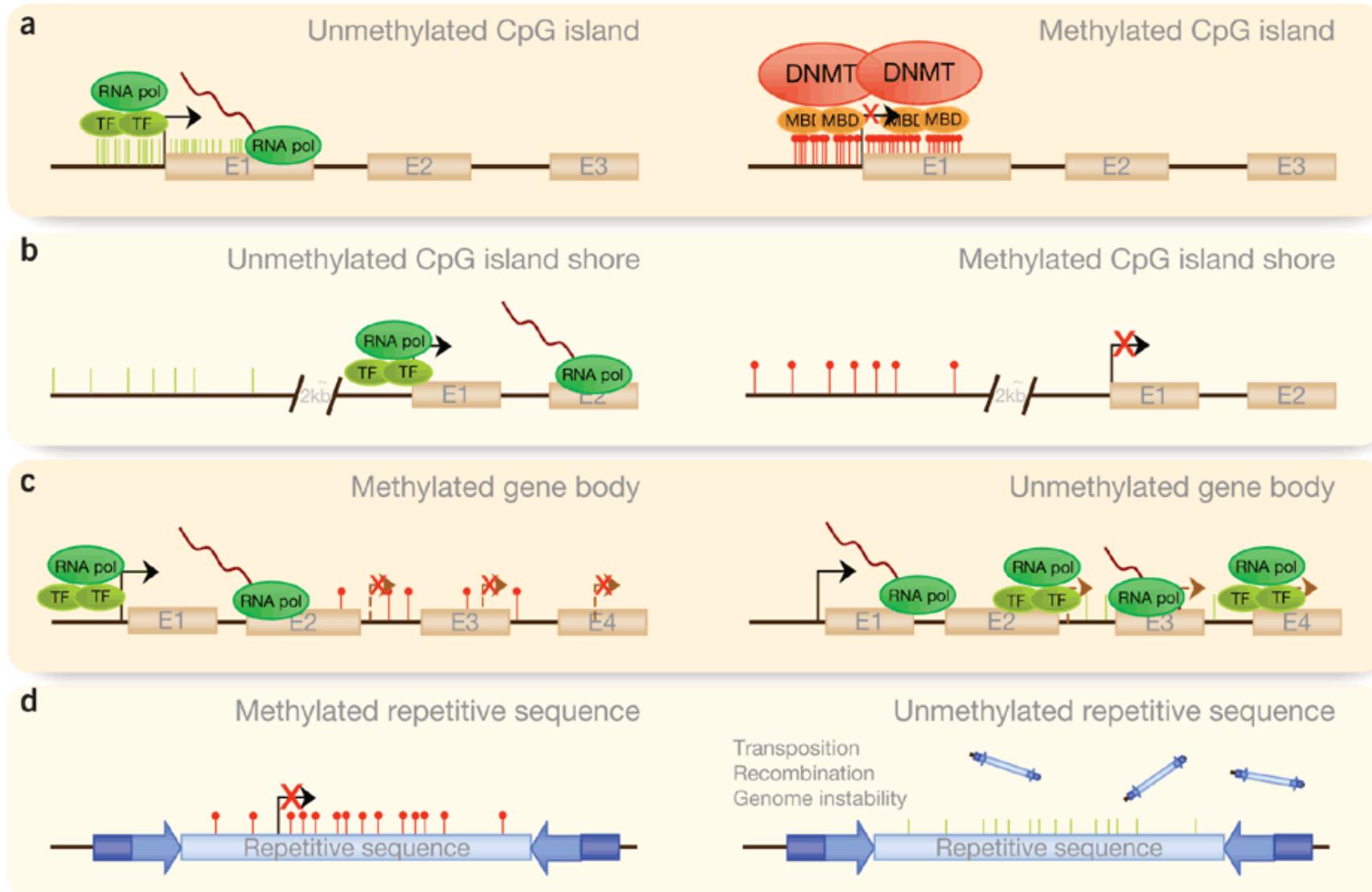
- 1944: Avery identifies DNA as gene material
- 1948: Hotchkiss discovered 5-methylcytosine

"minor constituent designated epicytosine [with] a migration rate somewhat greater than that of cytosine"

- 1953: DNA structure resolved
- 1962: Methylation in CpG dinucleotides
- Around 1980: DNA methylation regulates gene expression
 - McGhee & Ginder: Beta-globin expression
 - Jones & Taylor: Cytidine analogs
- Early 1980s: CpG Island discovery



Methylation in health and disease

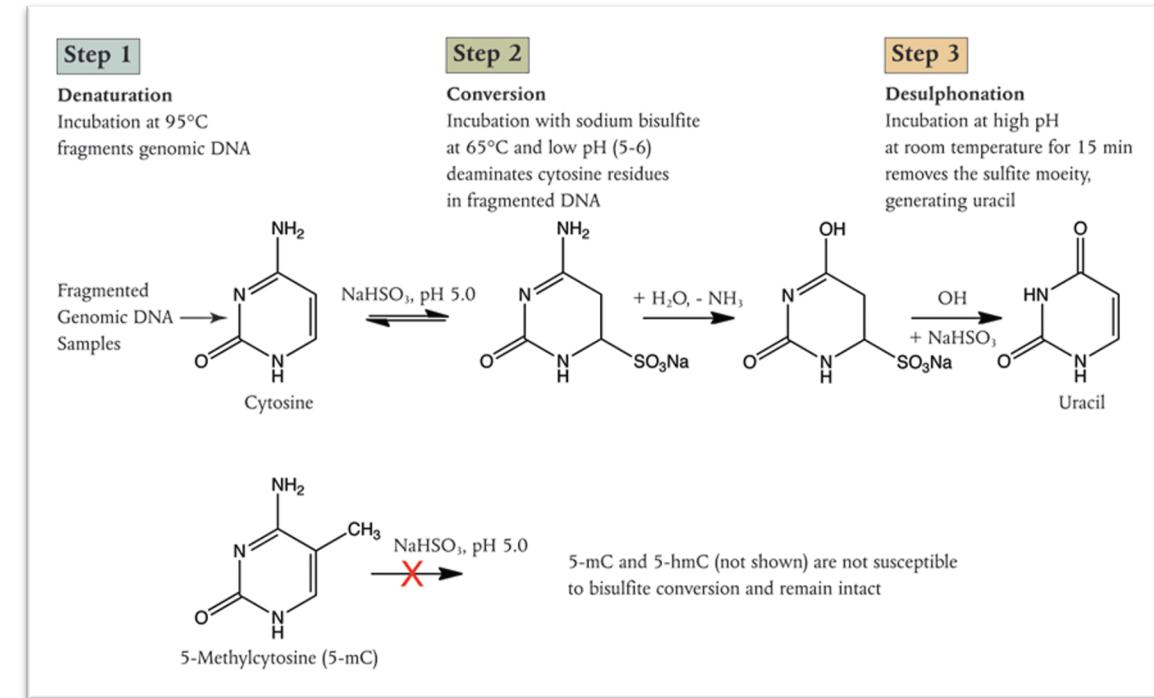
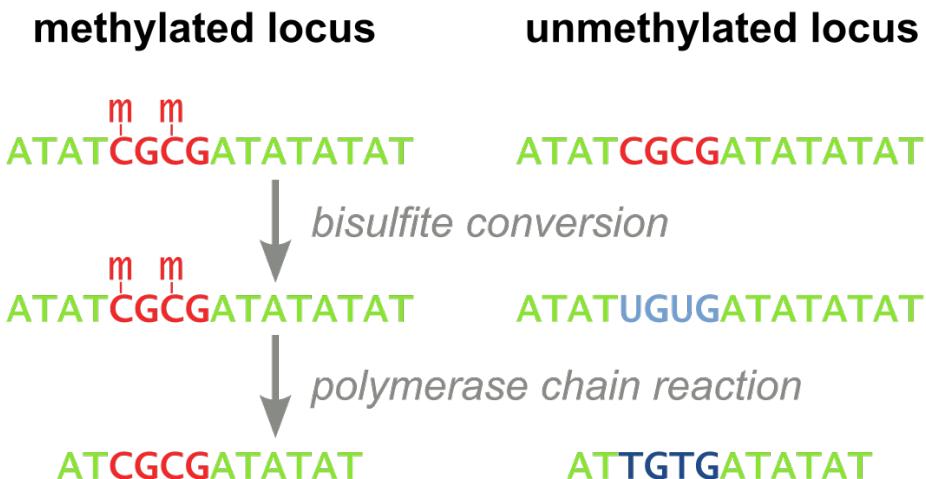


Methods commonly used for detection of DNA methylation

CpG Methylation					
A Site-Specific Methylation			B Global Methylation		
Method	Refs	Outcome	Method	Refs	Outcome
* Methylation-specific PCR	Herman <i>et al.</i> , 1996 Karouzakis <i>et al.</i> , 2009	Detection of methylation of a specific gene or a region	* Methylation-specific PCR of repetitive sequences	Yang <i>et al.</i> , 2004	Methylation status of repetitive sequences of the genome
* Quantitative PCR (i.e. High Melting Resolution analysis)	Candiloro <i>et al.</i> , 2011 Newman <i>et al.</i> , 2012 Kristensen <i>et al.</i> , 2013	Methylation level of a specific gene/regions of the genome	* HPLC	Kuo <i>et al.</i> , 1980 Ehrlich <i>et al.</i> , 1982	
* COBRA (combined bisulfite and restriction analysis)	Xiong and Laird, 1997 Lahtz <i>et al.</i> , 2013	Quantification of methylation frequencies at individual consecutive CpG sites	* HPCE	Li <i>et al.</i> , 2009	
* Pyrosequencing	Candiloro <i>et al.</i> , 2011 Kristensen <i>et al.</i> , 2013		* Mass spectrometry (Flow cytometry, microscopy etc.)	Annan <i>et al.</i> , 1989 Coolen <i>et al.</i> , 2007	Total amount of methylated cytosines in the genome
			* Anti-5meC immunological methods	Habib <i>et al.</i> , 1999 Piyathilake <i>et al.</i> , 2004 Brown <i>et al.</i> , 2008 Karouzakis <i>et al.</i> , 2009 Schneider and Fagagna, 2012	
			* Microarray	Weber <i>et al.</i> , 2005 Bar-Nur <i>et al.</i> , 2011 Bocke <i>et al.</i> , 2011 Walker <i>et al.</i> , 2011	Genome-mapping (methylation status of large DNA fragments)
C Global Methylation Detection Using Proxy Markers			* Next-generation sequencing (i.e. Illumina platform)	Bibikova <i>et al.</i> , 2009 Russnes <i>et al.</i> , 2011 Zong <i>et al.</i> , 2012 Glossop <i>et al.</i> , 2013 Renner <i>et al.</i> , 2013	(1) Methylation status of individual CpG dinucleotides, (2) Methylation status of gene regions with sites in the promoter region, 5'UTR, first exon, gene body, 3'UTR, and (3) Methylation status of CpG islands, shore and shelf regions (distance from the CpG islands), and non-CpG islands of the genome
Marker	Example Methods		Interpretation	Refs	
	* MBD domain of MBD1 attaches to a luciferase sensor (luminometer)			Badran <i>et al.</i> , 2011	
MBD1	* Dot blot analysis of MBD1 protein		Global DNA methylation	Zhang <i>et al.</i> , 2012	
	* Illumina sequencing of methylated DNA enriched by the MBD domain of MBD1			Morita <i>et al.</i> , 2012	

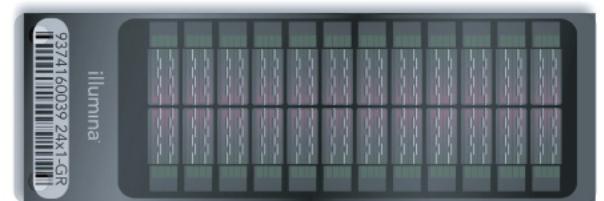
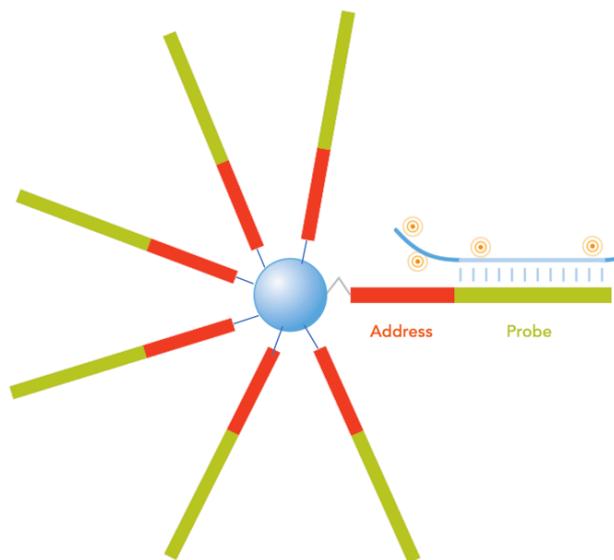
Bisulfite Conversion

- Bisulfite treatment crucial for both arrays and NGS
- C → U (-> T)
- mC → mC (-> C)
- methylation-specific PCR, high resolution melting curve analysis, microarray-based approaches, and next-generation sequencing



Methylation Arrays

- Infinium Methylation BeadChip arrays from Illumina: 27K, 450K and 850K (or EPIC)
- >480,000 CpG loci, covers 99% of RefSeq genes
- Distributed over various functional elements; covers 96% of CGI
- 50bp single-stranded DNA oligos (“probes”) attached to silica beads, 2 detection channels (red and green)
- Hybrid of 2 different probe designs!

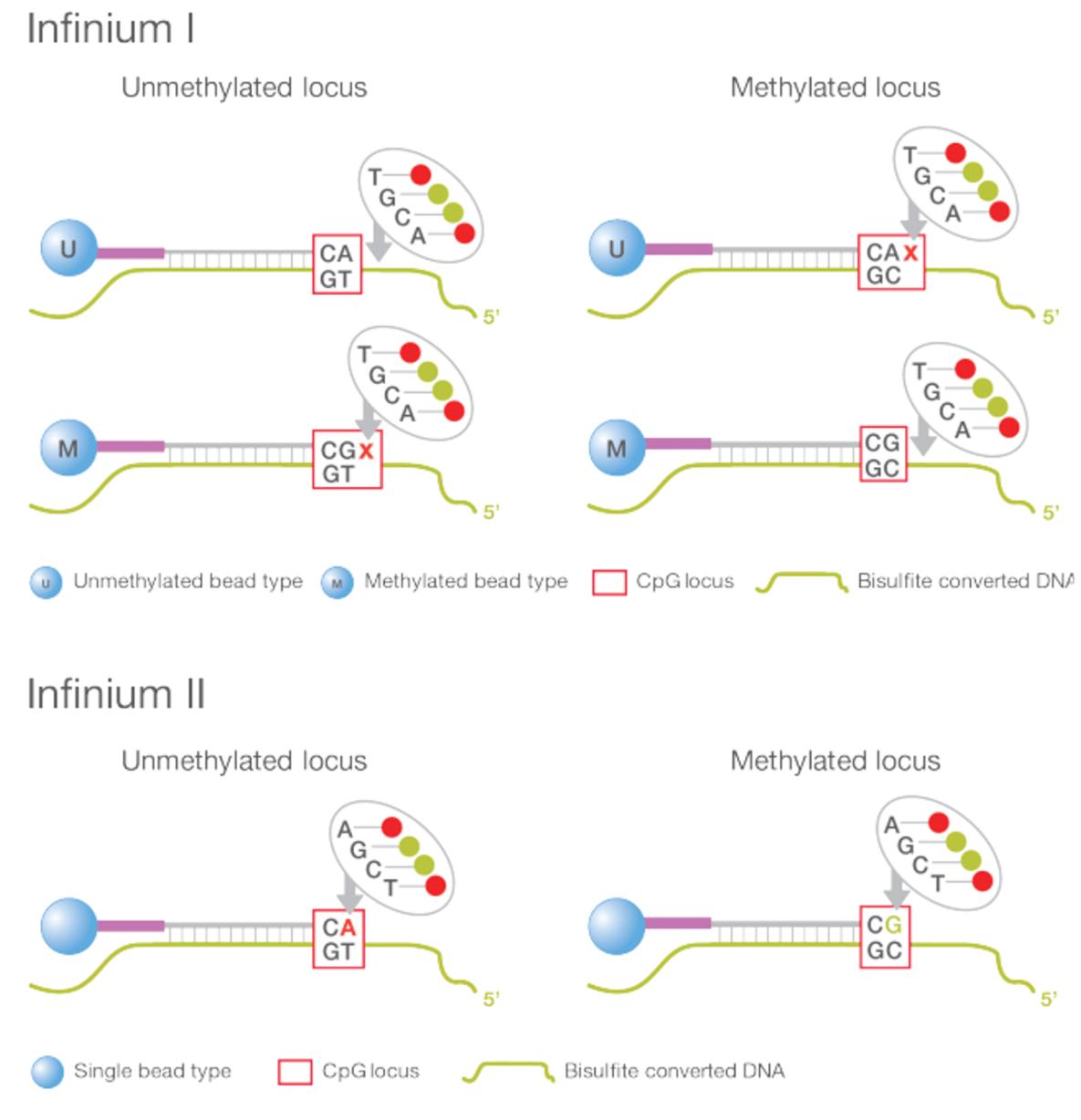
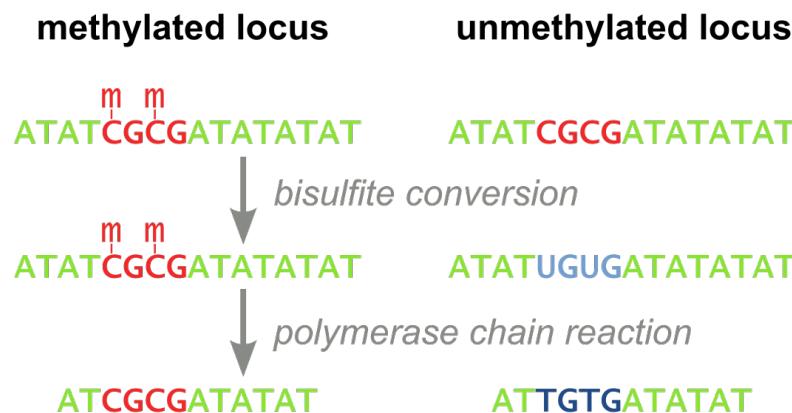


BeadChip Array

Infinium: Type I vs II

- Type I: single color detection, two beads
- Type II: two color detection, single bead

Type I	Type II
Same chemistry as 27K	New for 450K
2 beads/CpG	1 bead/CpG (fit more)
Better for CpG dense regions	Better for less CpG dense regions
More stable and reproducible	Lower dynamic range



From colors to methylation

- Intensities are used to estimate Beta values; for both probe designs

$$\beta = M / (M + U + 100)$$

- Beta value between 0 and 1 (represents the fraction methylation)
- Easily interpretable, but related M value has better statistical properties

$$Mvalue = \log_2(M/U)$$

This step and the next will be part of the tutorial for downstream analysis, so time to get your setup ready!

