

Illumina 450k: A microarray for the study of DNA methylation

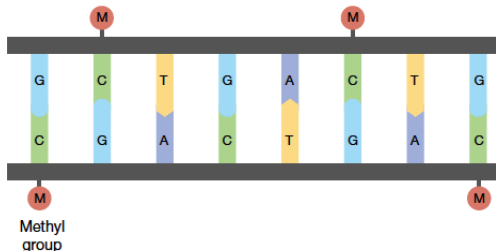
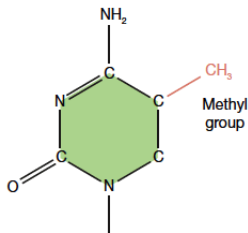
Jean-Philippe Fortin

May 15, 2013

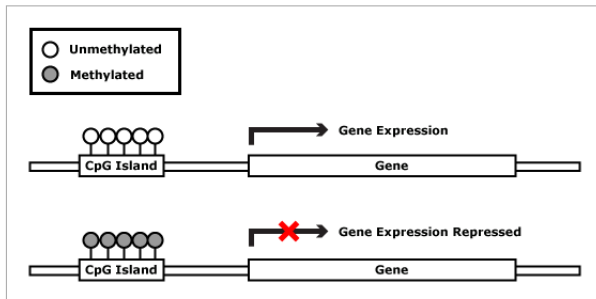
- DNA Methylation
- Illumina Infinium 450k Human Methylation Assay
- Probes Design
- Statistical Challenges

DNA methylation

- **DNA methylation:** 5-methylcytosine residues found in CpG dinucleotides. One of the most studied epigenetics marks.
- **Epigenetics:** study of gene expression variants caused by mechanisms that do not involve a change in the nucleotide sequence



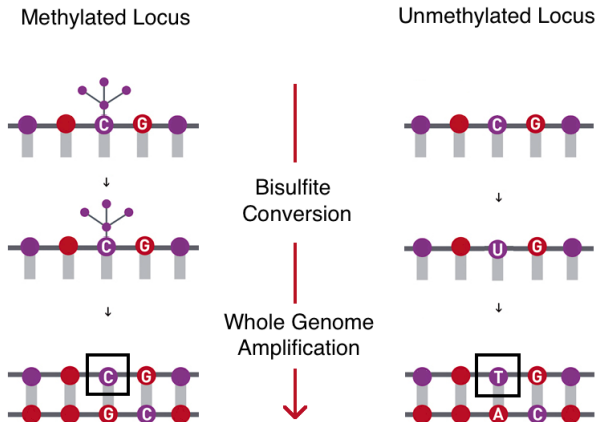
DNA Methylation



2

² <http://missinglink.ucsf.edu/lm/genes-and-genomes/methylation.html>

Bisulfite Conversion

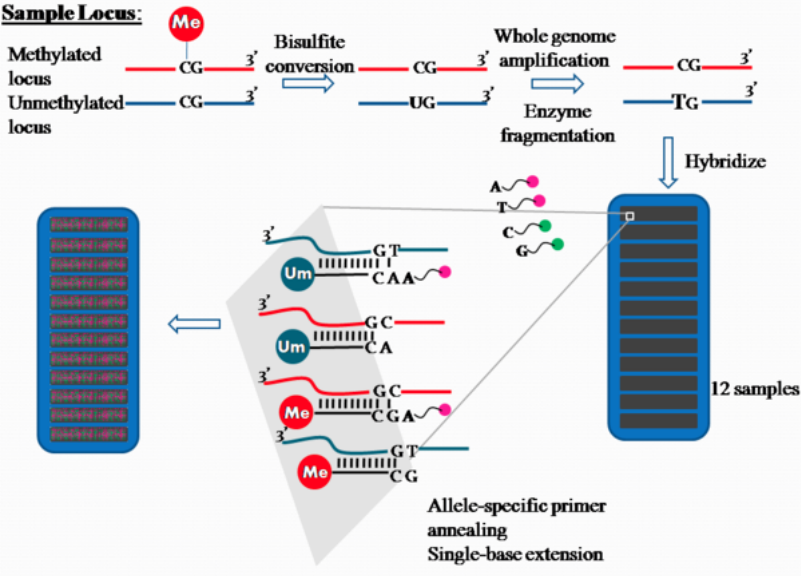


After treating the DNA with bisulfite conversion, the analysis of DNA methylation is reduced to an analysis of single nucleotide polymorphisms (SNPs) for T's and C's:

- C is found: original cytosine was methylated
 - T is found: original cytosine was non-methylated
- Can use **Next Generation Sequencing** or **Microarrays**

- A recent popular methylation array that allows the interrogation of more than 485,000 methylation sites per sample
- **Contains two chemistry technologies:**
 - 135,000 probes from Infinium I array
 - 350,000 probes from Infinium II.

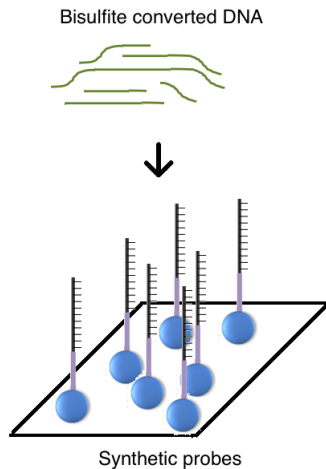
Sample Locus:



4

4 Figure from <http://en.wikipedia.org/wiki/Illumina-Methylation-Assay>

Probe Design



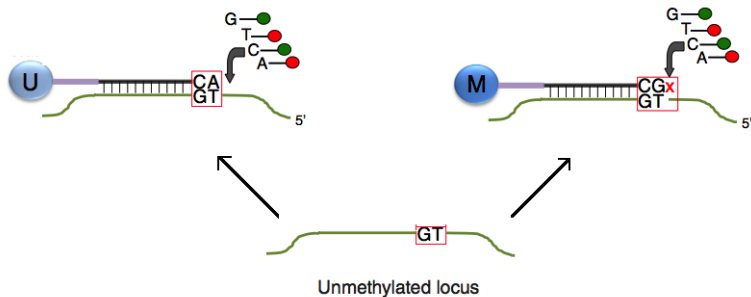
Infinium I Probes:

- One probe for the methylated locus
- One probe for the non-methylated locus



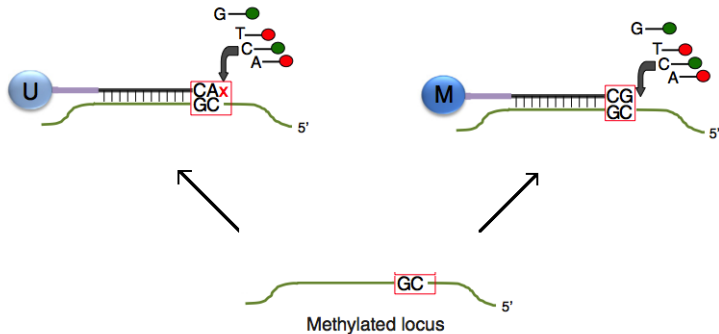
Infinium I Probes:

- Hybridization of the **unmethylated locus**:

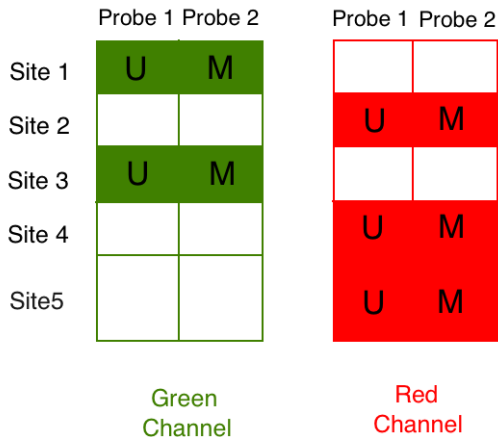


Infinium I Probes:

- Hybridization of the **methylated locus**:



Infinium I Design



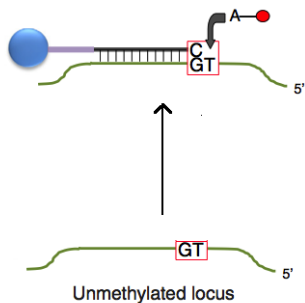
Infinium II Probes:

- Only one probe to interrogate both loci
- The interrogated CpG site is at the end of the probe
- The last base corresponding to the G is not included



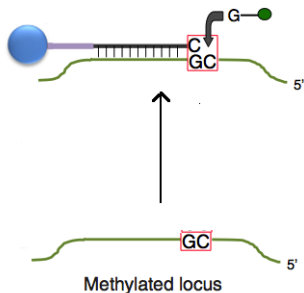
Infinium II Probes:

- Hybridization of the **unmethylated** locus = RED

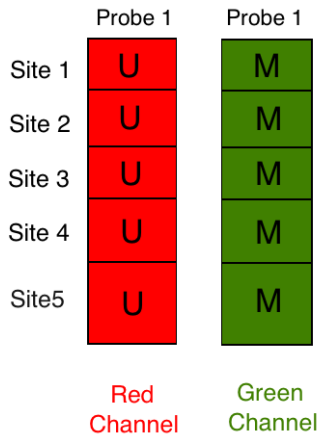


Infinium II Probes:

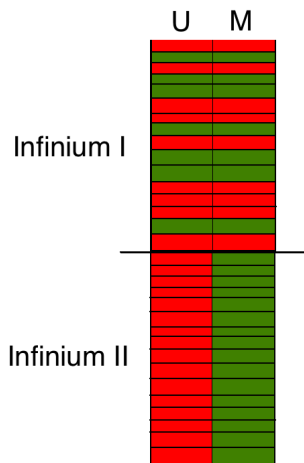
- Hybridization of the **methylated** locus = GREEN



Infinium II Design



Infinium I + Infinium II



Beta-value:

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

$\beta = 0$: All cells are non-methylated

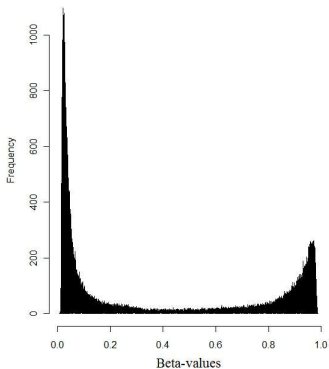
$\beta = 1$: All cells are methylated

M value:

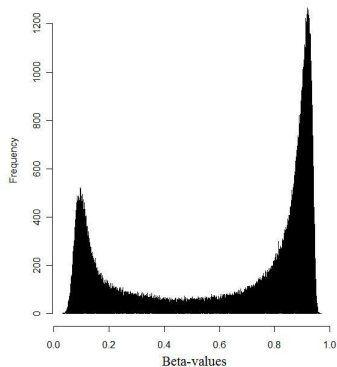
$$Mvalue = \log\left(\frac{M}{U}\right)$$

The Beta-value distributions are not the same for the two designs:

Infinium I

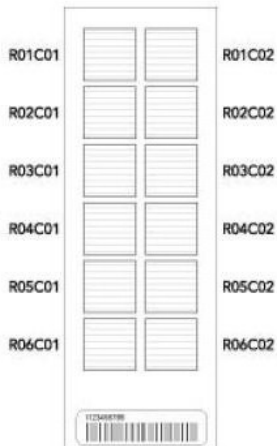


Infinium II

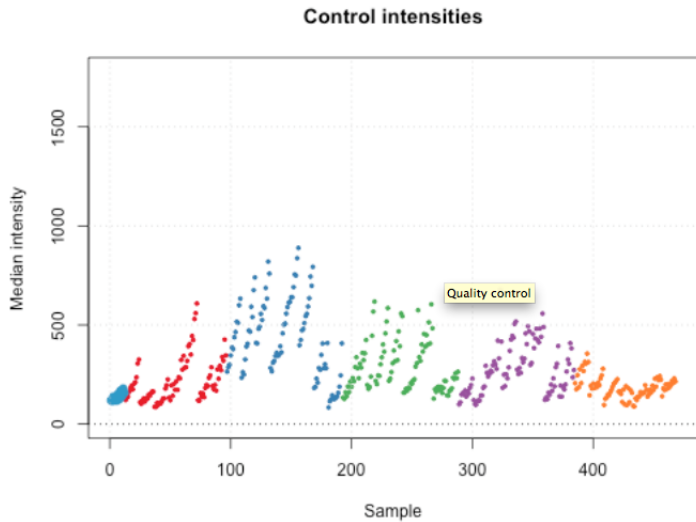


Chip Configuration

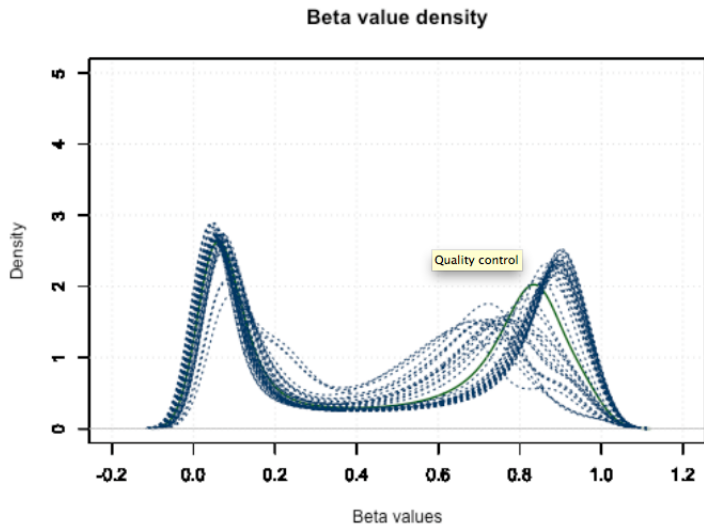
Array Sample Section Nomenclature



Background intensities



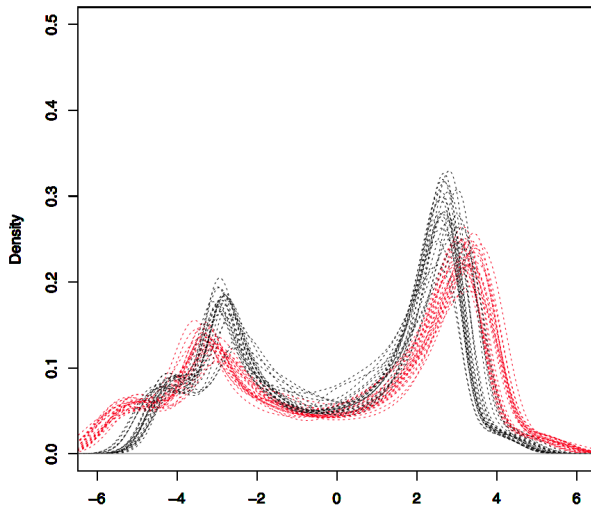
Beta-value densities



- How to reduce technical variation? Background, dye bias and spatial effects
- Presence of two types of probes
- Statistical model for differential analysis
- Blood mixture

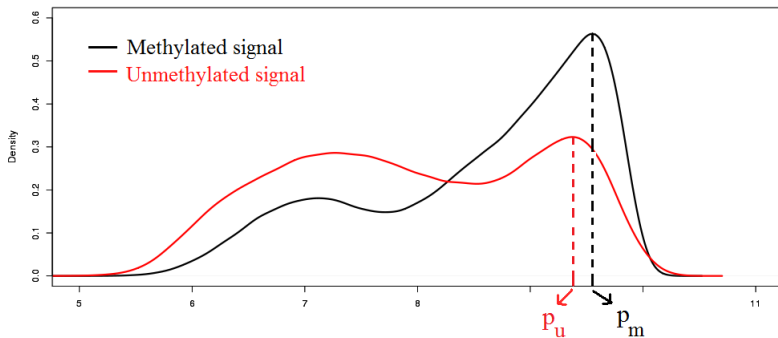
Background Effects

Raw data



- For Type I probes, always same color.
- Not a problem if we normalize green and red probes separately.
- For Type II probes
 - GREEN signal = methylation
 - RED signal = no methylation
- Dye bias can be a concern if inconsistent across samples

Type II Probes: Separate distributions



Relation between control dye bias and difference in peaks

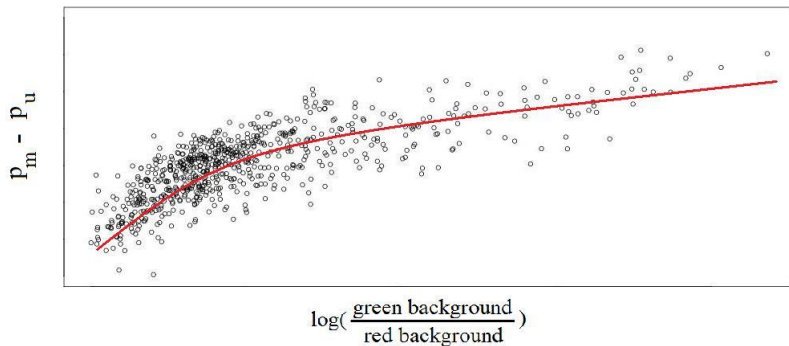


Figure 1: <http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Sequencing/Epigenetic-Sequencing/Methylation-Analysis.html>

Figure 2: <http://missinglink.ucsf.edu/lm/genes-and-genomes/methylation.html>

Figure 3: <http://www.diagenode.com/en/applications/bisulfite-conversion.php>

Figure 4: <http://en.wikipedia.org/wiki/Illumina-Methylation-Assay>