A Novel Method for Understanding the Effect of Genetic Variation on DNA Methylation

Irene Kaplow

July 16, 2014

STANFORD
COMPUTER SCIENCE

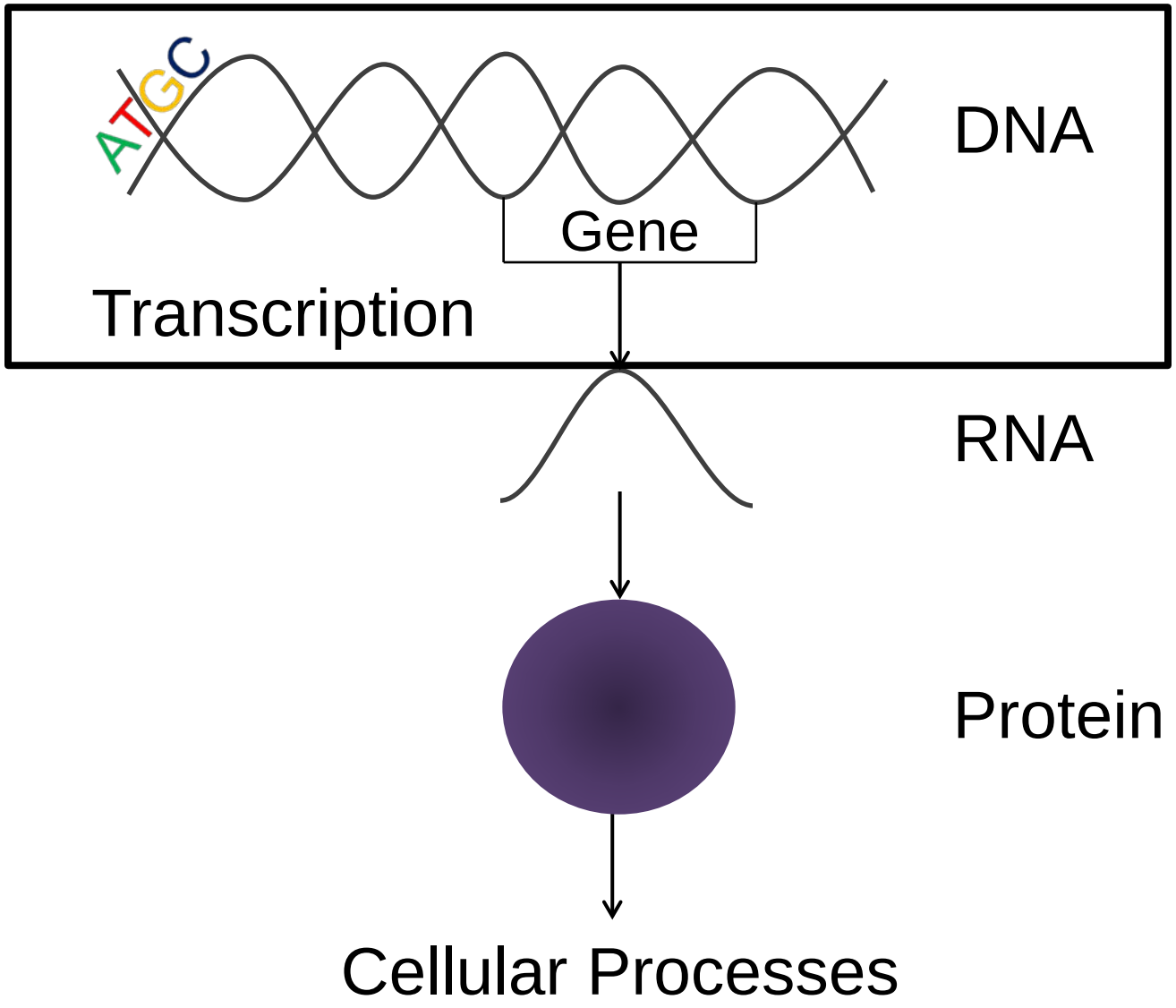
Outline

- Background
- Methods
- Preliminary Results

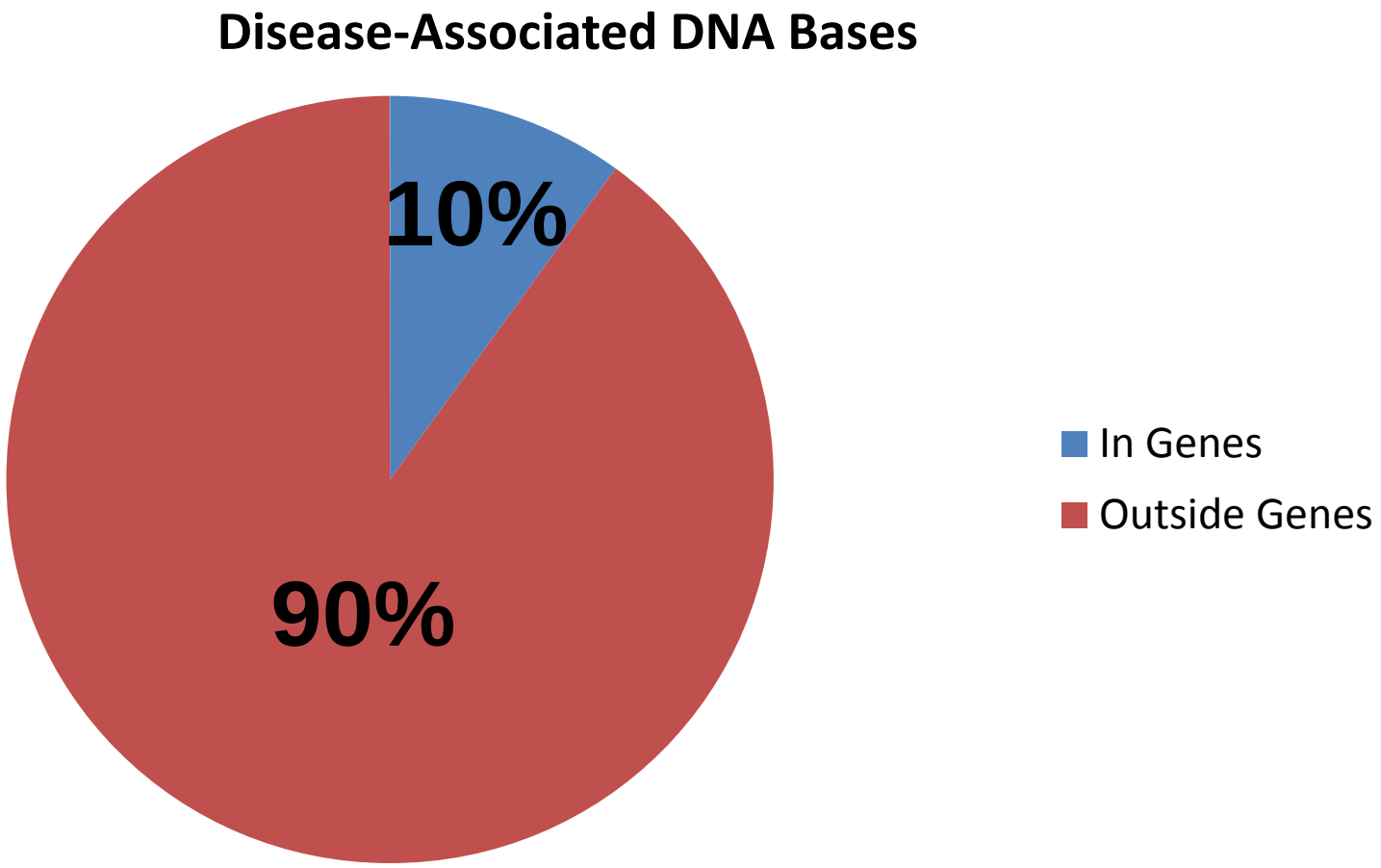
Outline

- **Background**
- Methods
- Preliminary Results

Molecular Biology 101

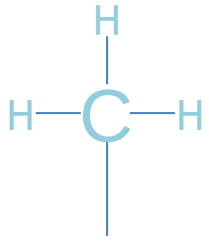


Importance of Transcription



Bases outside genes can affect disease by *regulating transcription*.

DNA Methylation

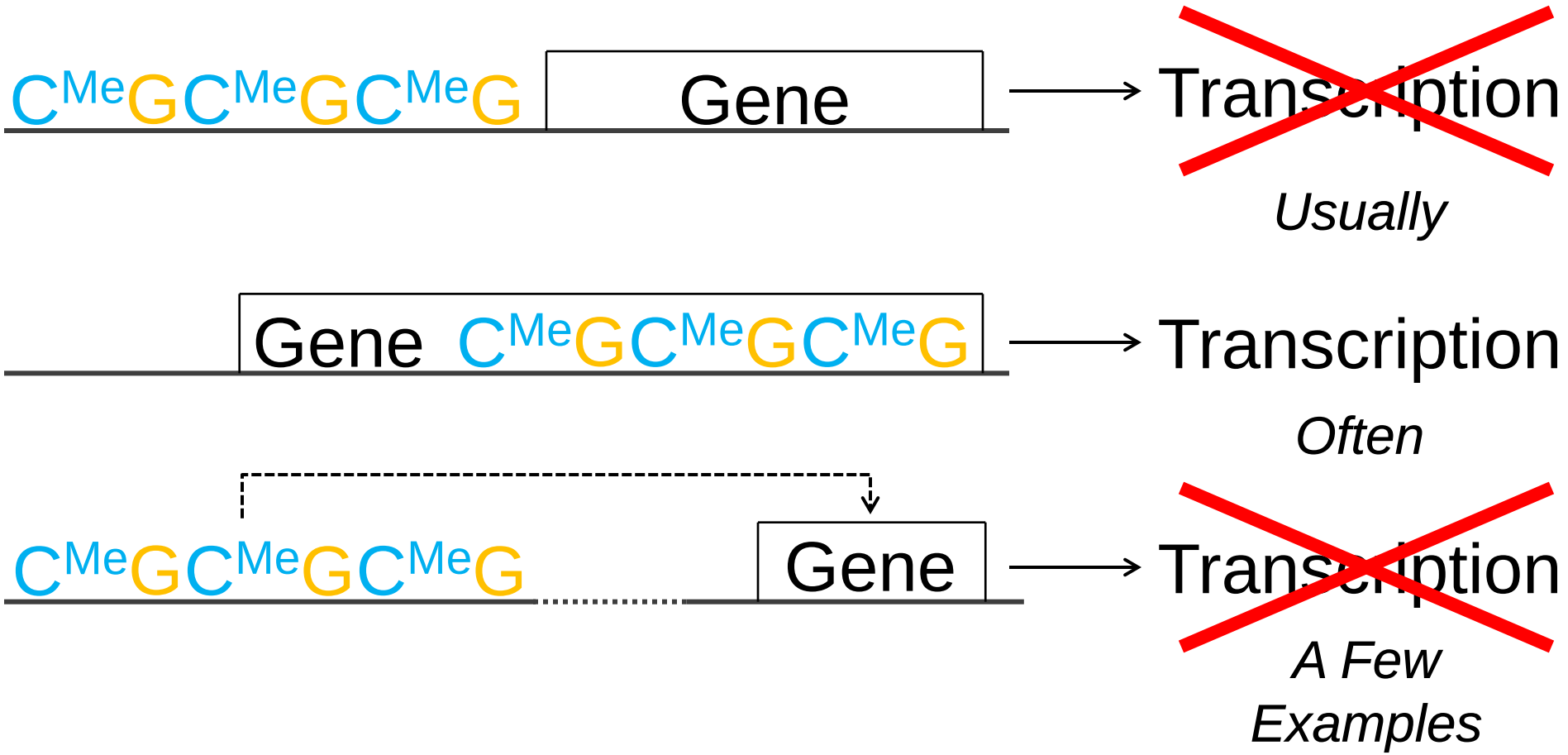


C C = C^{Me}

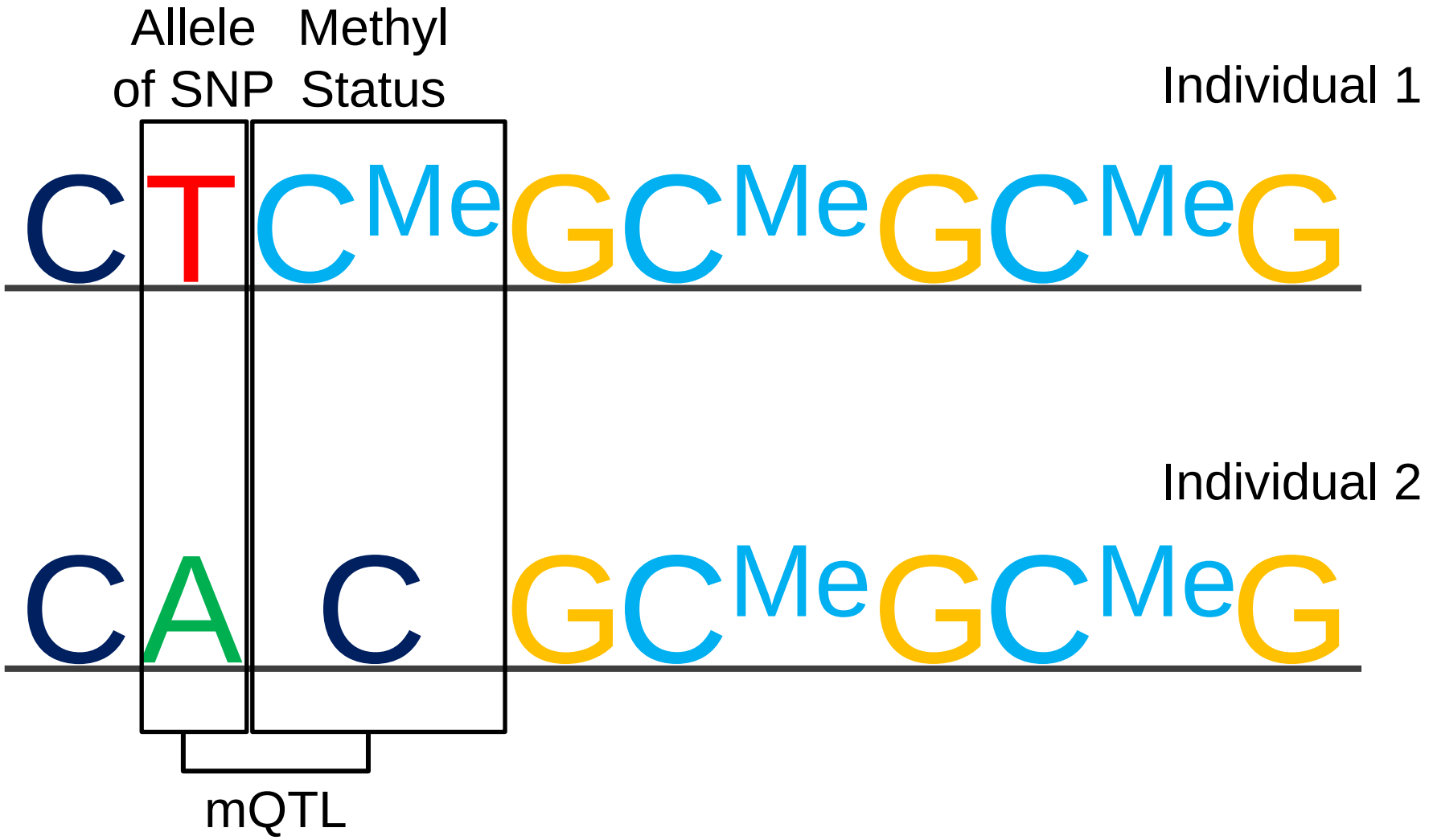
C T C^{Me} G C^{Me} G C^{Me} G

"CpG"

Role of Methylation in Transcription

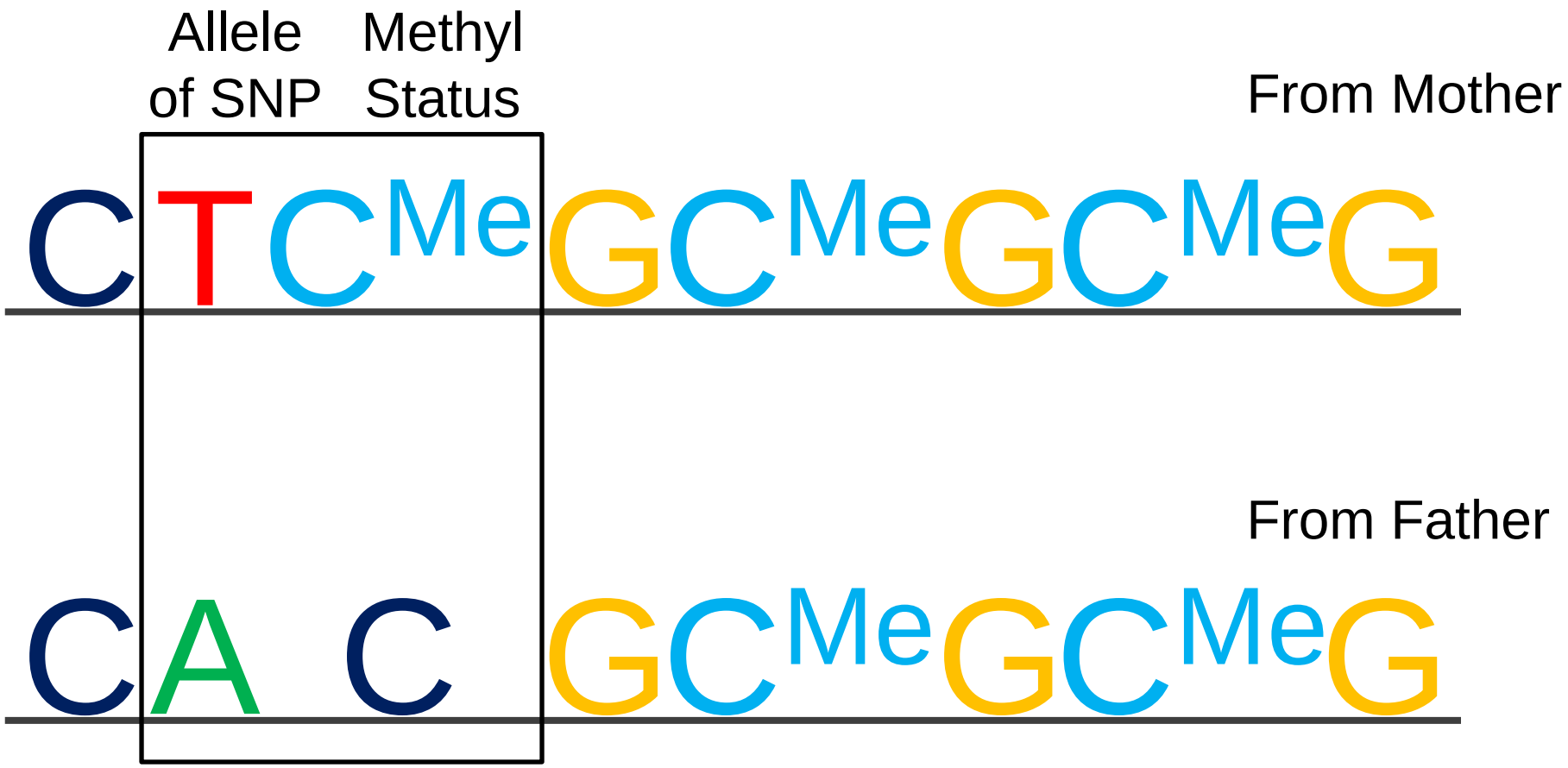


Methylation Quantitative Trait Loci



mQTLs can have dramatic effects on transcription!

Heterozygosity



Goal of Our Study

- Question: ***Which differences in DNA bases are associated with whether near-by Cs are methylated?***
  - In other words, which SNPs are mQTLs?

- Do these associations provide mechanisms that explain how SNPs may affect transcription or disease?

Outline

- Background
- **Methods**
- Preliminary Results

Identifying Cytosine Methyl Status

Methylation arrays:

Also identifies alleles for only subset of SNPs

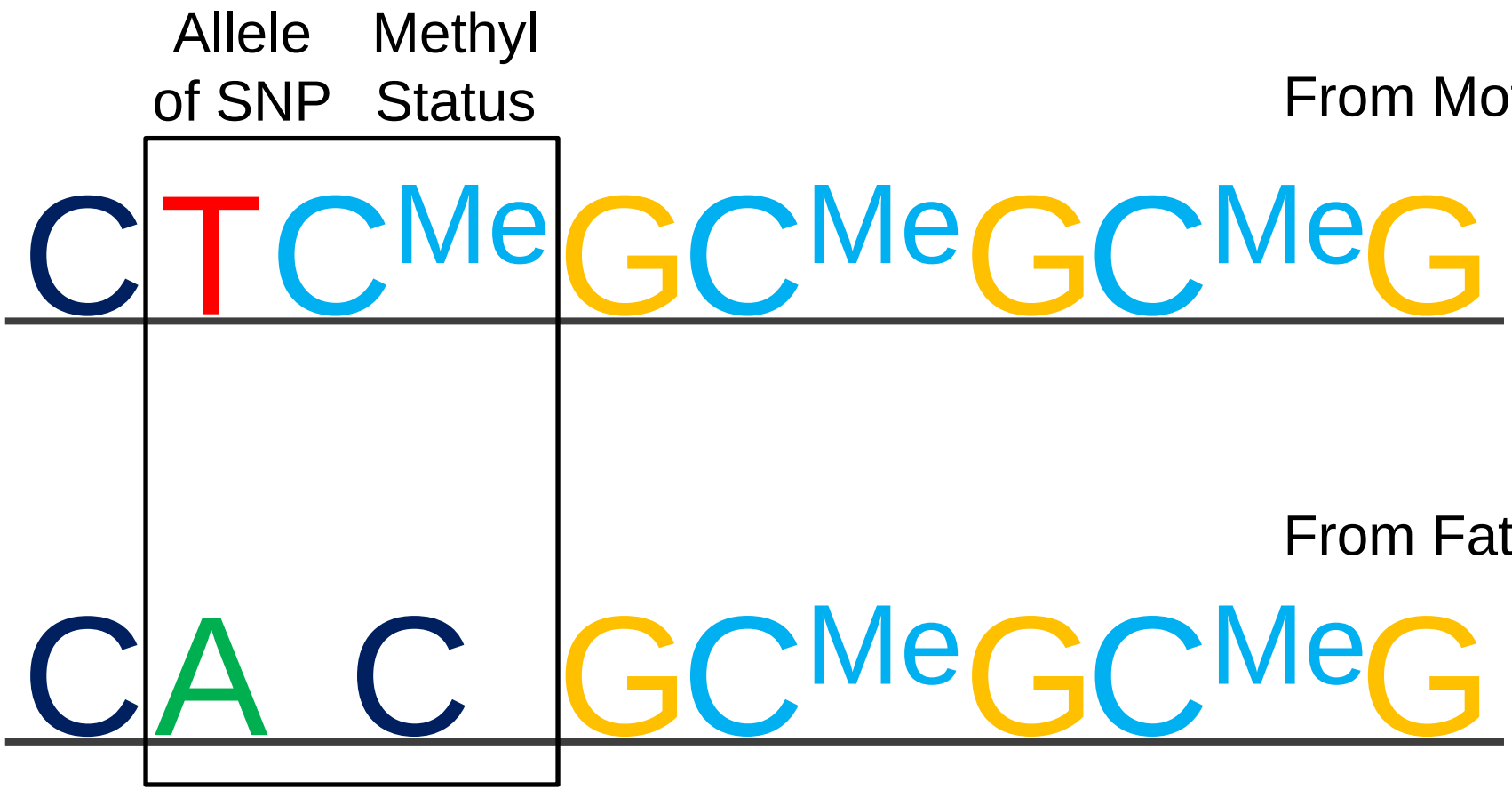
Whole-genome bisulfite sequencing:

From Mother

From Father

Also identifies alleles for all SNPs for DNA from each parent

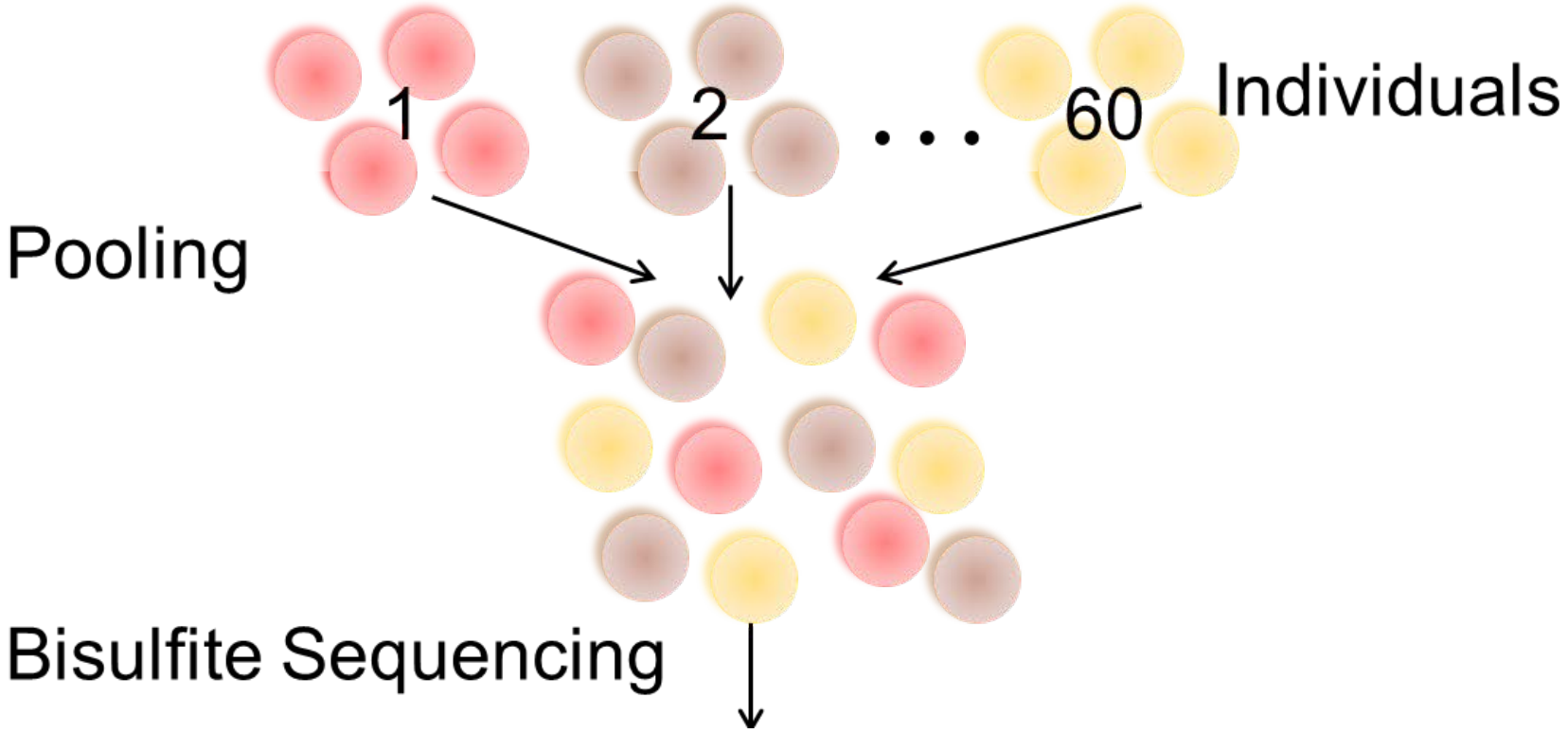
Methyl Arrays vs Bisulfite Sequencing



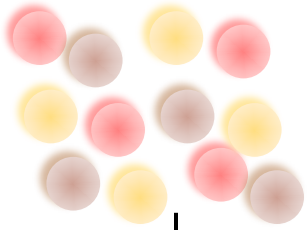
Allele of SNP	Methyl Status		
0.5	0.5	Methylation Array	***Averaging***
1	1	Bisulfite Sequencing	**Expensive!**
0	0		

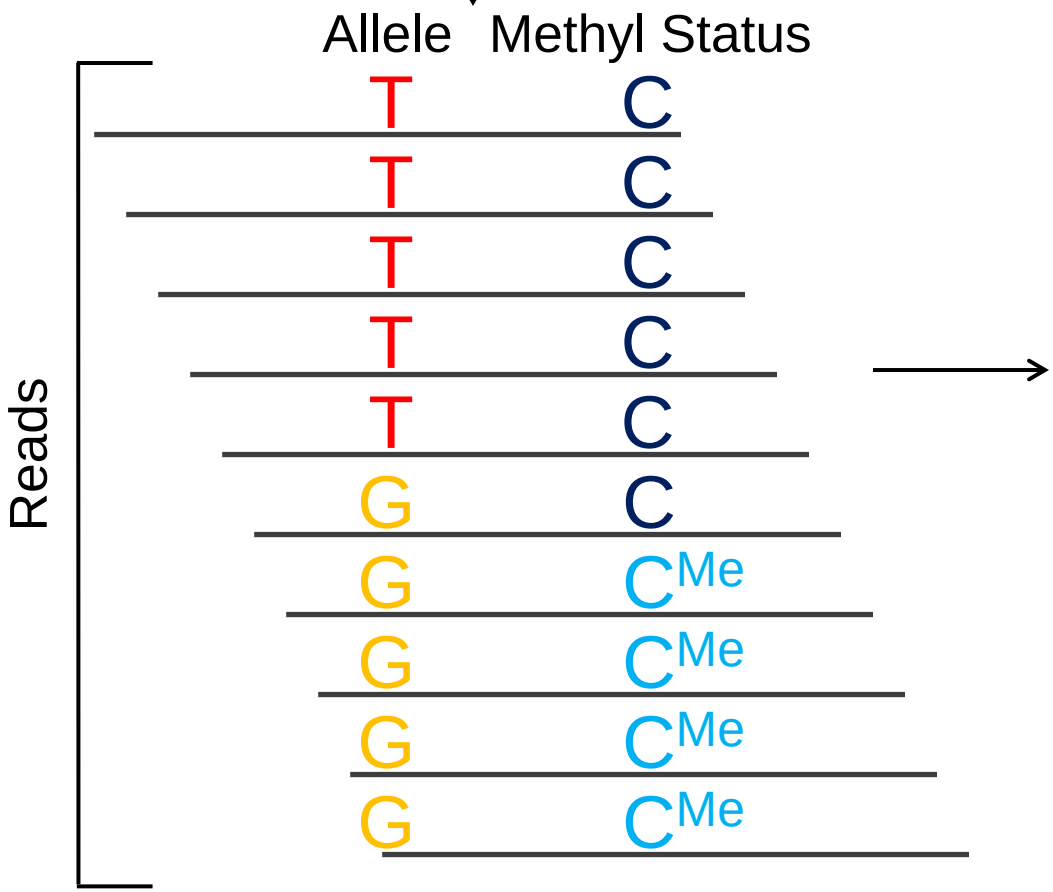
Bisulfite Sequencing on Pool

- Cell type: Lymphoblastoid cell lines (LCLs)
- Individuals: 60 Yorubans used in other studies



Identifying mQTLs from Pooled Data





	Unmethyl	Methyl
T Allele	5	0
G Allele	1	4

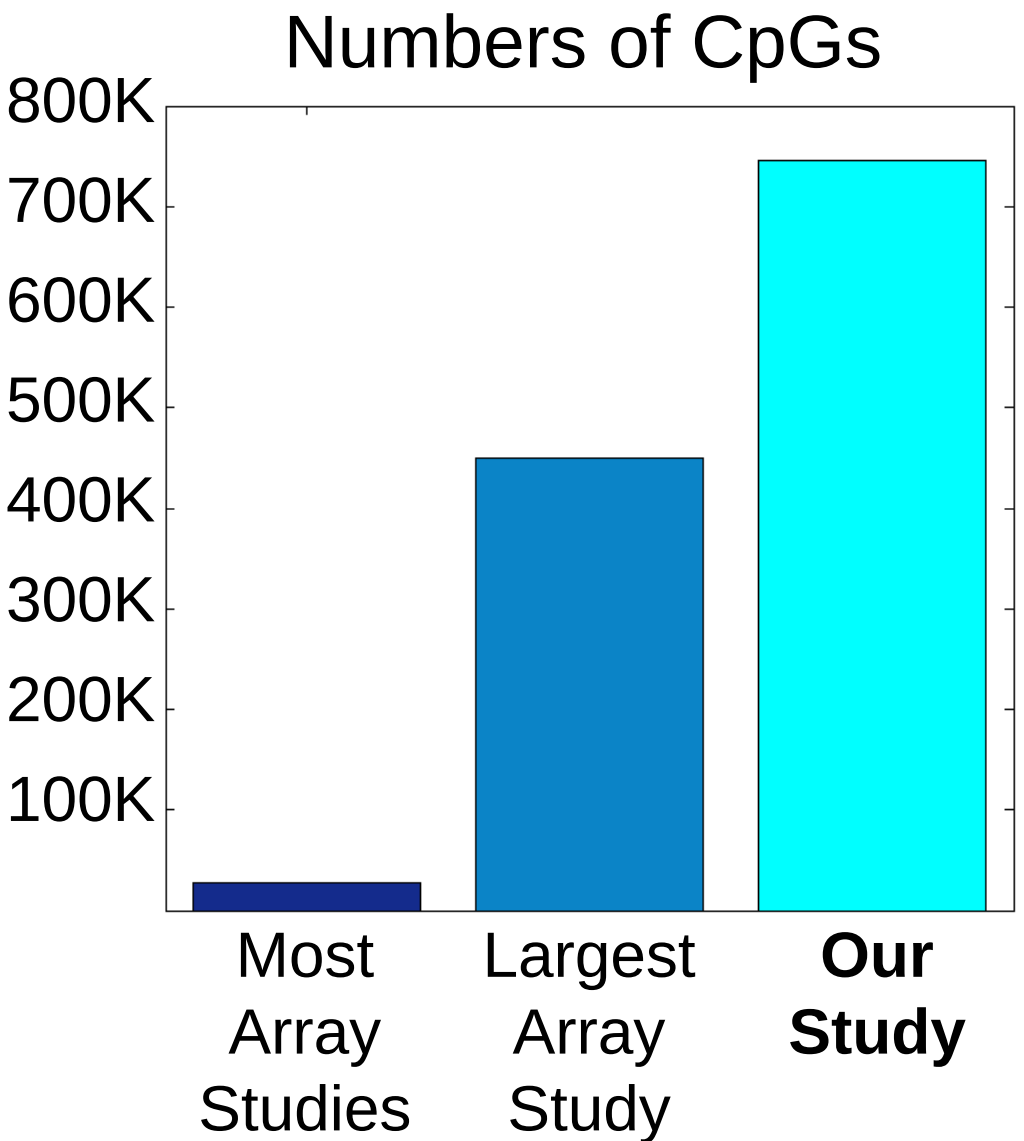
Fisher's Exact Test

Outline

- Background
- Methods
- **Preliminary Results**

Preliminary Results Summary

- Evaluated 746,210 CpGs
- 2,276 mQTLs
- Some overlap with:
  - SNPs associated with transcription
  - SNPs associated with physical traits



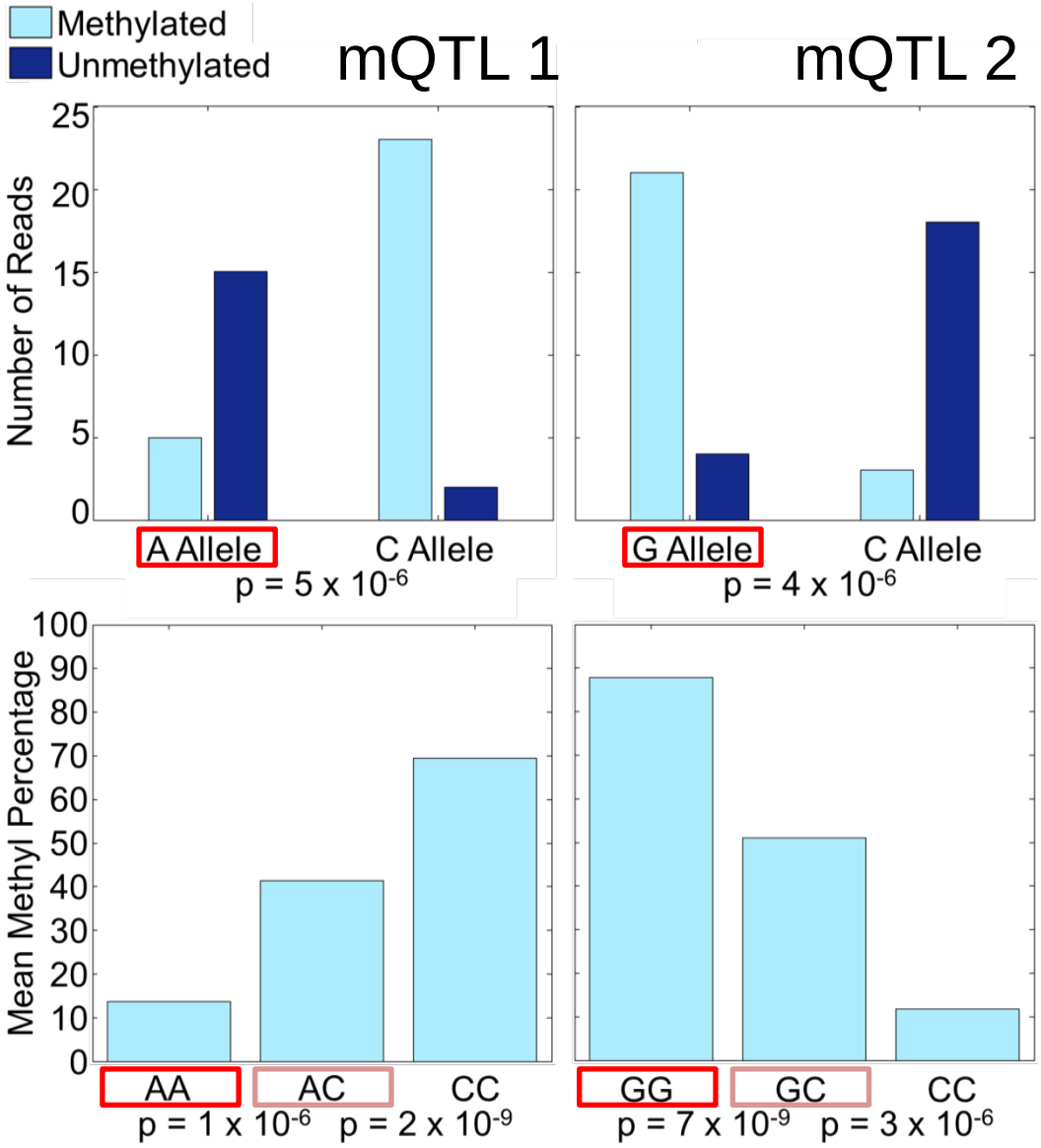
mQTL Overlaps with Disease SNPs

- 4 exact overlaps, including:
  1. SNP associated with age-related macular degeneration in multiple studies
  2. SNP associated with ratio of visceral adipose tissue to subcutaneous adipose tissue
     - Measure of obesity
- Not found in previous mQTL studies
- Validated these 2 overlaps using pyrosequencing

mQTLs and Their Validation

Bisulfite Sequencing

Pyrosequencing Validation



Contributions

- Provided novel pooling and association methods for identifying mQTLs
- Used these methods to identify over 2,000 mQTLs, most of which have not been found in previous studies
- Found mQTLs that are also associated with disease and transcription, suggesting a mechanism for influence of SNPs

Acknowledgements

- Adviser: Hunter Fraser
- Co-adviser: Anshul Kundaje
- Experimentalists:
  - Michael Kobor (UBC)
  - Sarah Mah (UBC)
  - Julie MacIsaac (UBC)
  - Yiqi Zhou
- Labs:
  - Fraser Lab
  - Koller Lab
  - Kundaje Lab
- Funding: