PLOS COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Aberrant DNA methylation defines isoform usage in cancer, with functional implications

Yun-Ching Chen, Laura Elnitski*

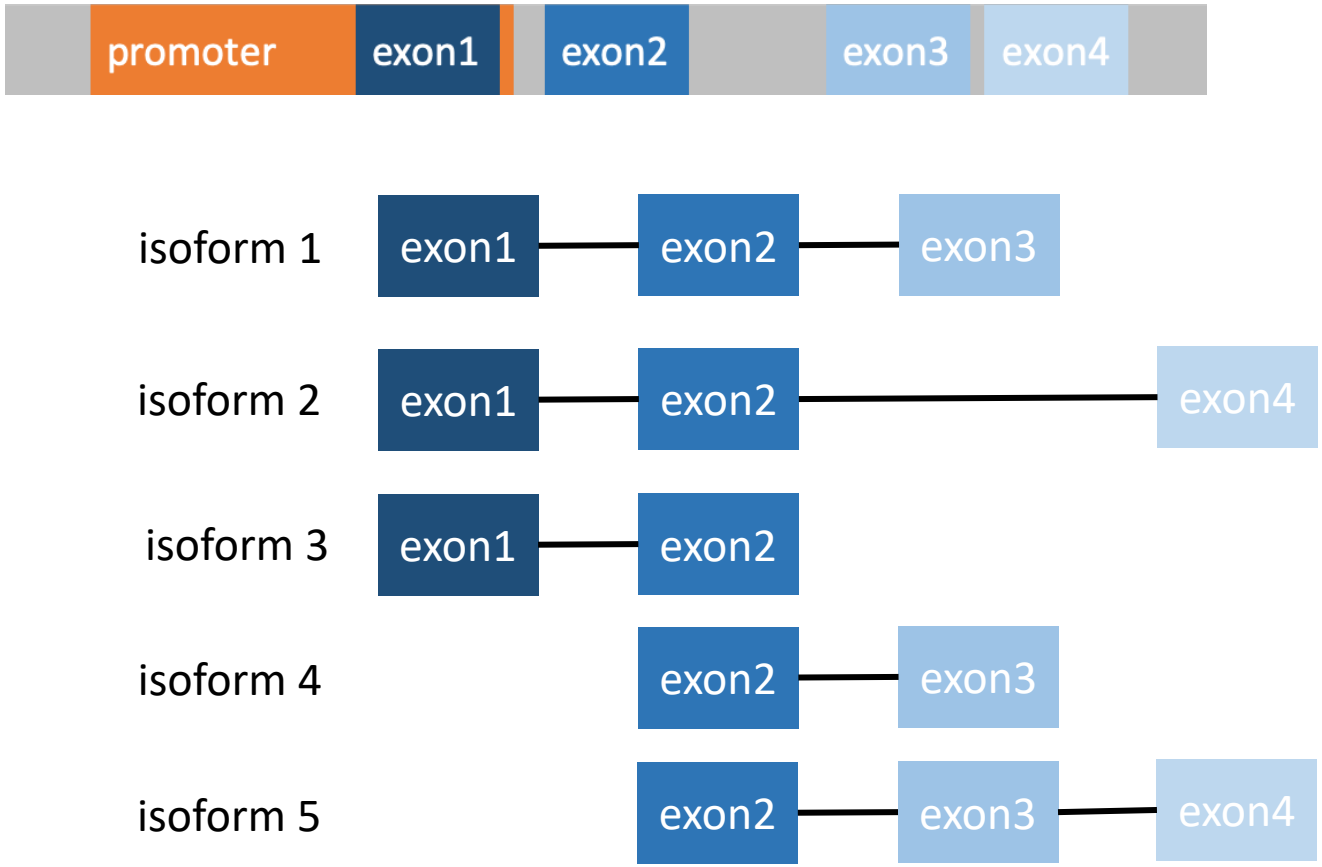
Genomic Functional Analysis Section, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, United States of America

Coriell Journal Club 2019-08-20

Kelsey Keith

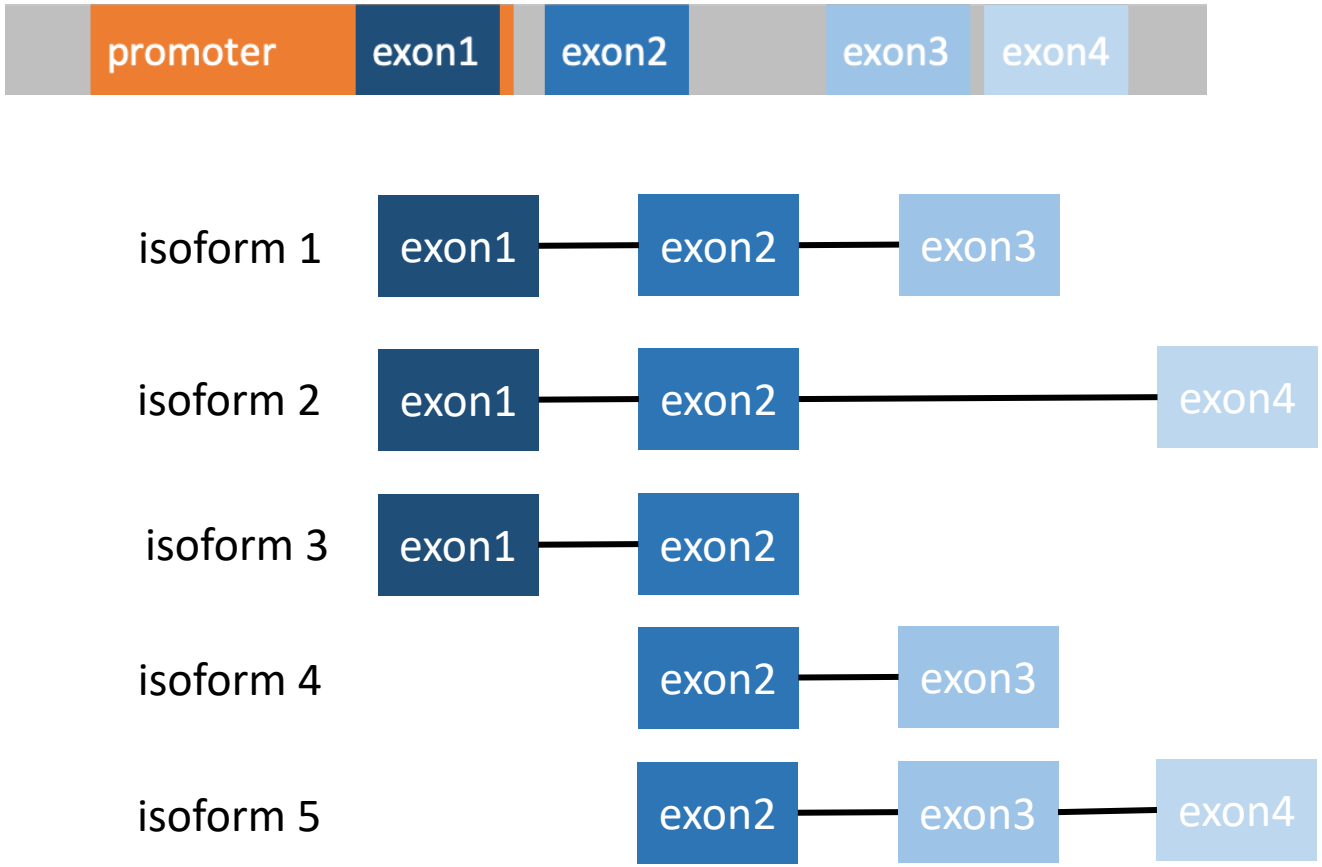
Background

Isoform Switching



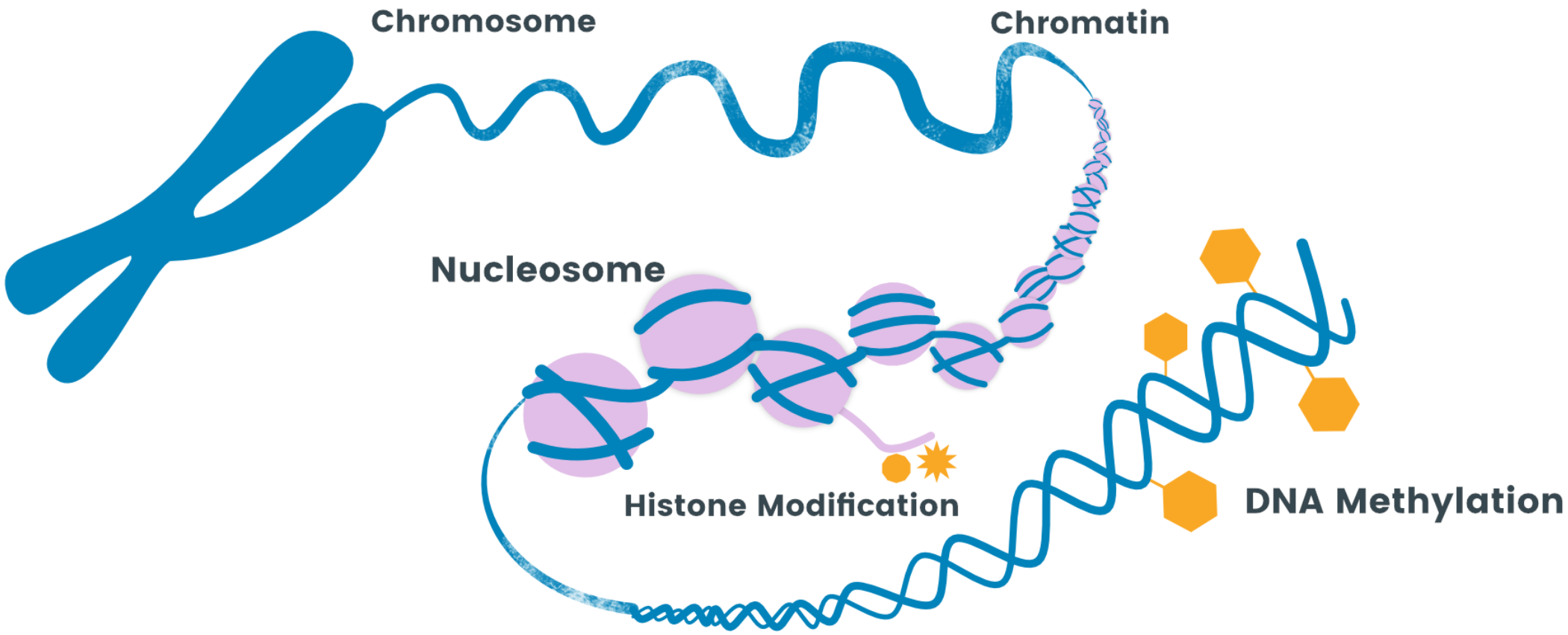
- 90% of genes
- Strategies
  - Alternative start or end
  - Alternative splice sites
- Happens in normal cells, but is co-opted in cancer
  - Loss of function
  - Some gain of function

Isoform Switching

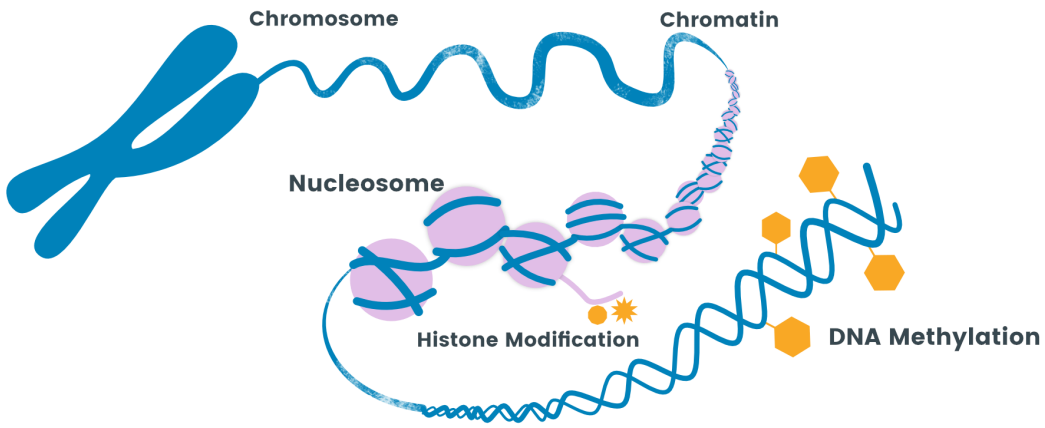


1. Modify protein-coding potential of mRNA transcript
2. Swapping functional domains
3. Gain or loss of introns
4. Inducing nonsense mediated decay of mRNA transcript
5. Changing the length of open reading frames
6. Toggling signal peptide inclusion

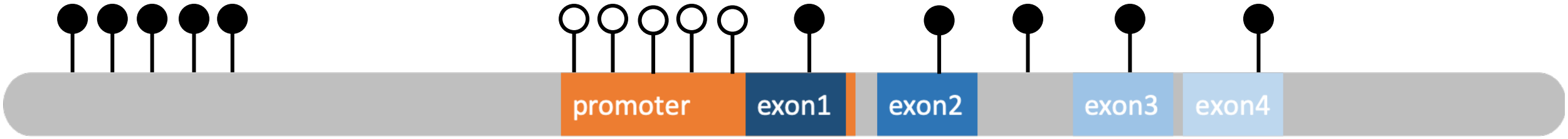
Methylation



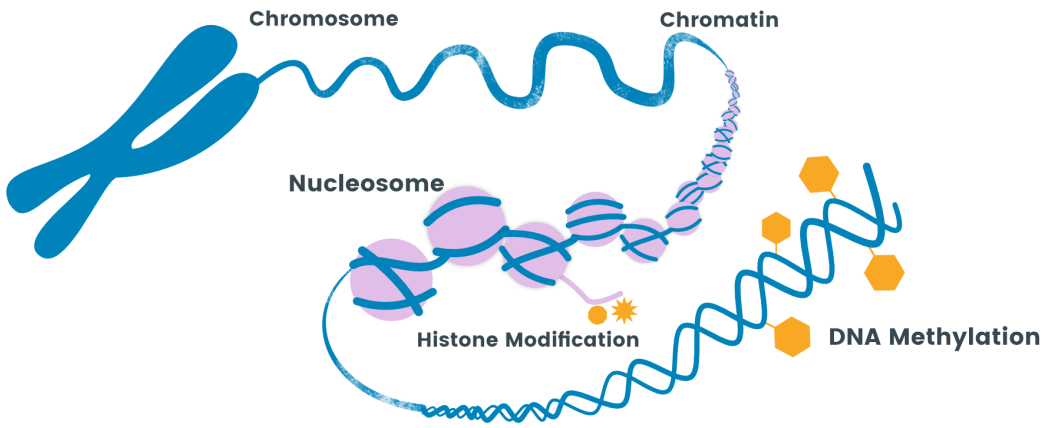
Methylation



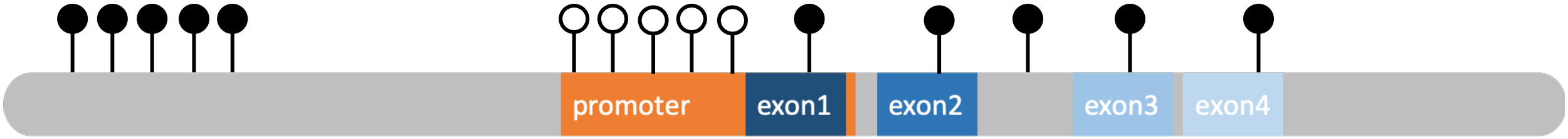
normal



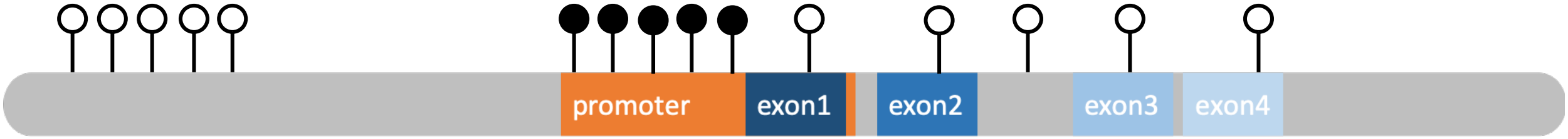
Methylation



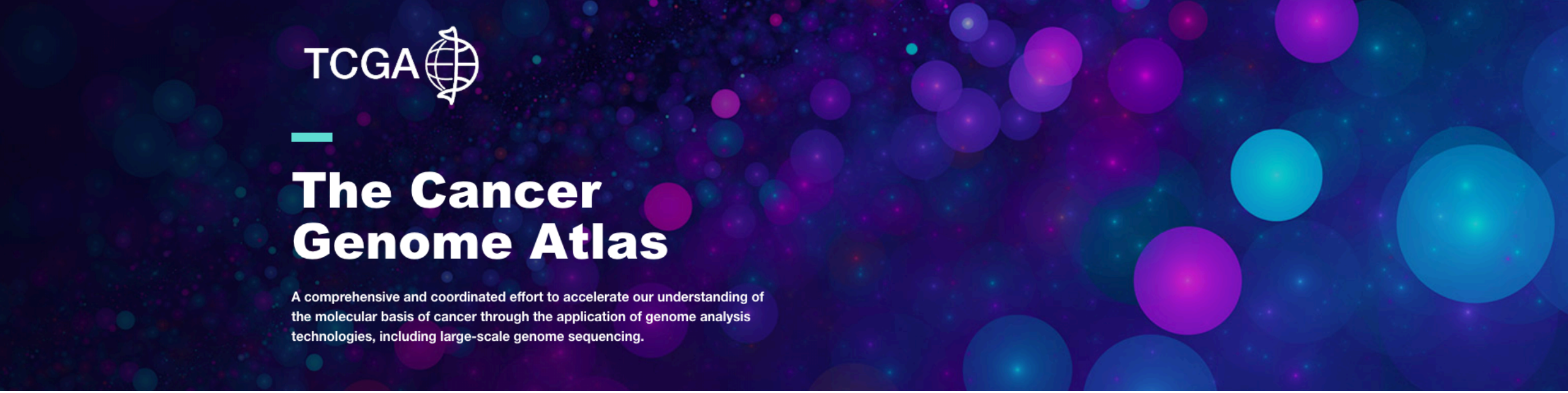
normal



cancer

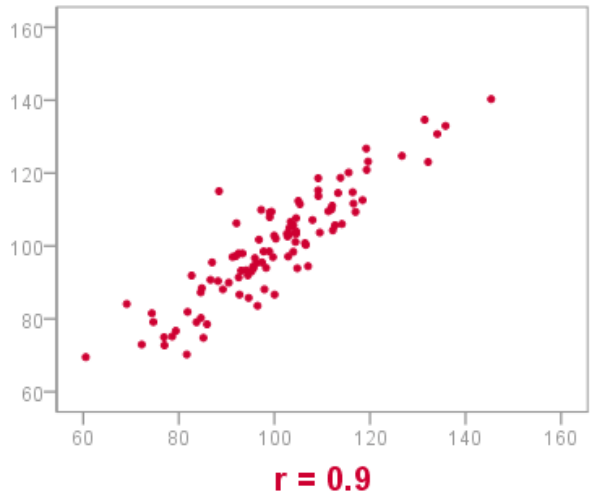


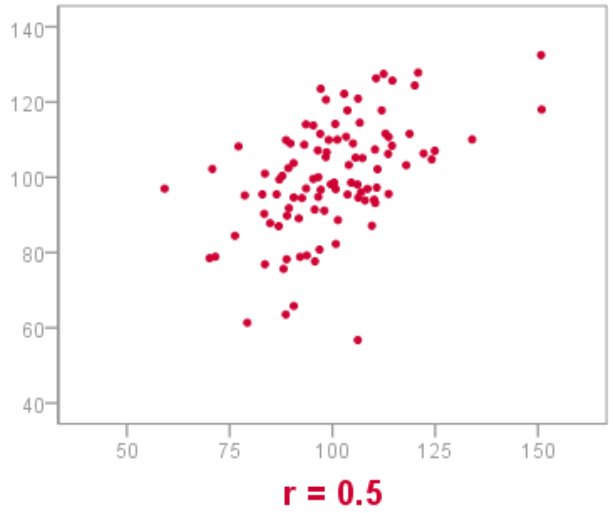
Methods

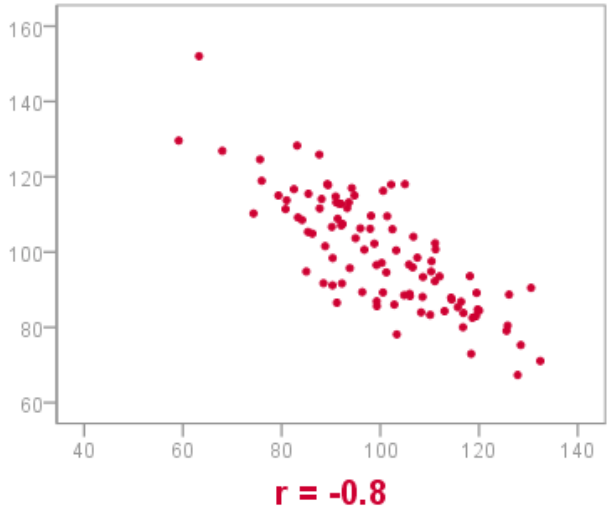


- Joint project between the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI)
- 2006 - 2018
- 33 cancer types
- 20,000 primary cancer and matched normal samples
- PanCancer Atlas
- Illumina Infinium HumanMethylation450 BeadChip array methylation data
- RNA-seq expression data

Pearson Correlation

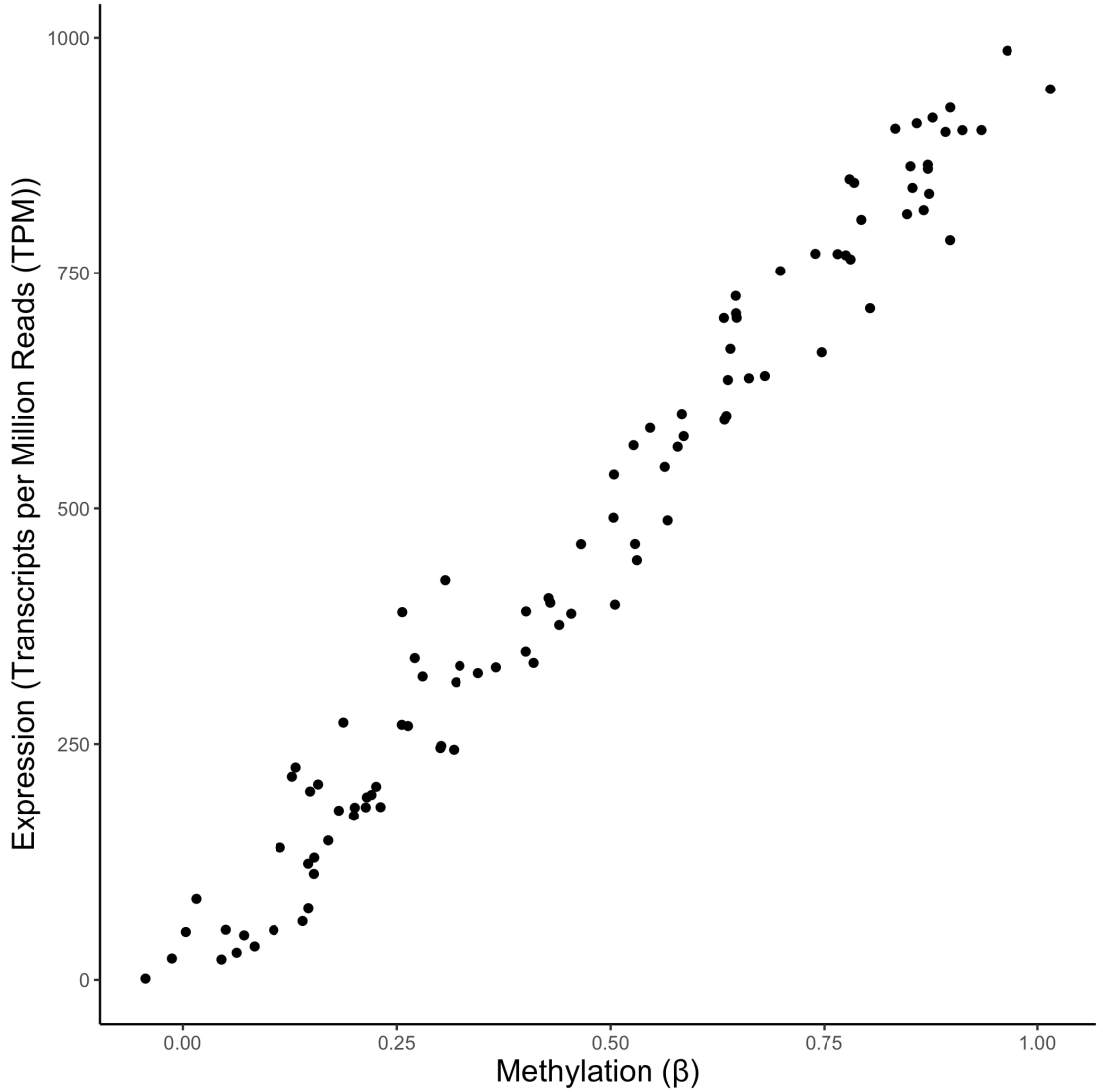






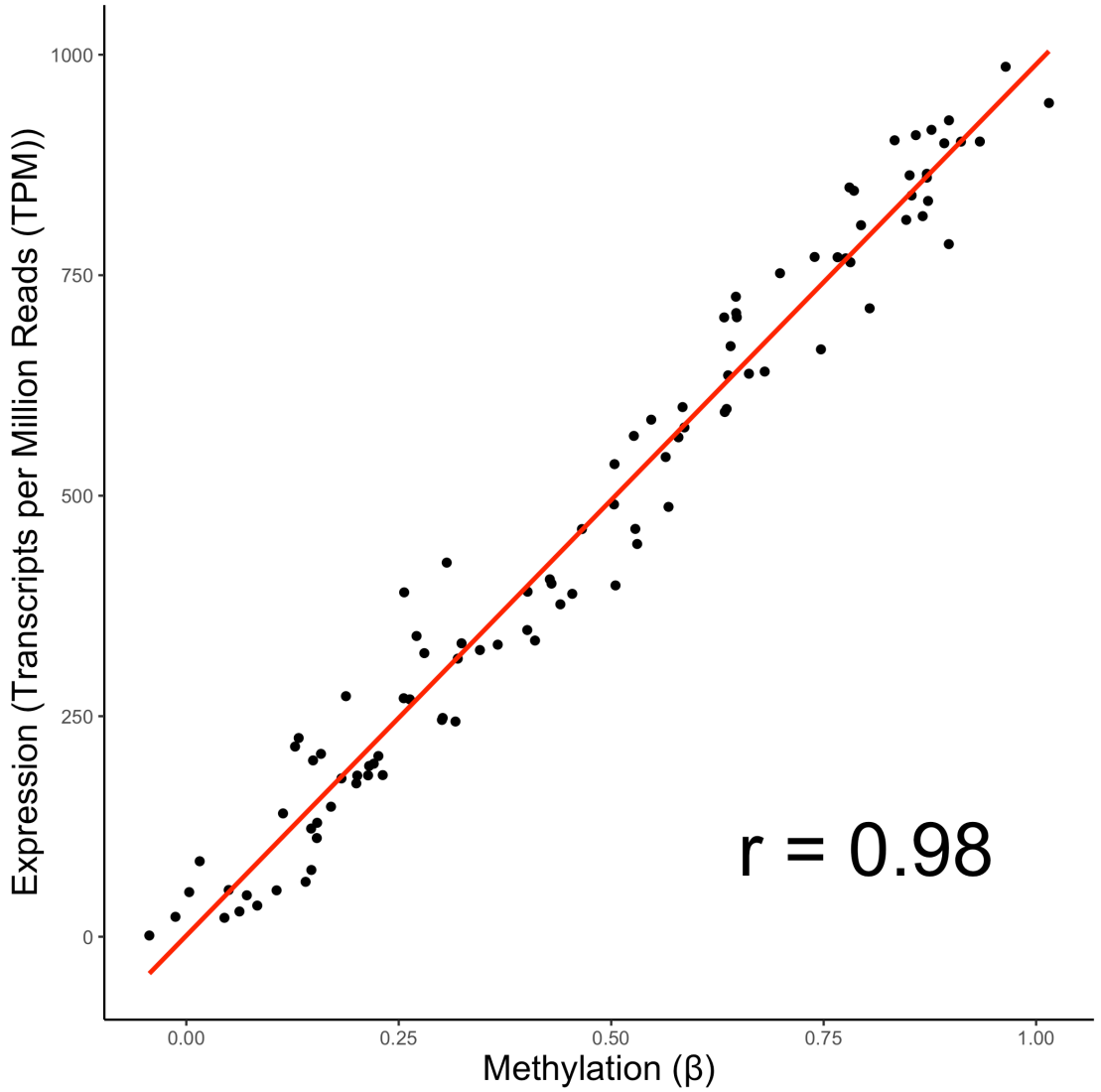
- Correlation coefficient describes how closely points lie along a line
- Between -1 and 1
- Sign indicates direction
- Closer to 0, the worse the correlation is

Pearson Correlation



- Correlation coefficient describes how closely points lie along a line
- Between -1 and 1
- Closer to 0, the worse the correlation is
- Sign indicates direction

Pearson Correlation



- Correlation coefficient describes how closely points lie along a line
- Between -1 and 1
- Closer to 0, the worse the correlation is
- Sign indicates direction

Pearson Correlation b/t Methylation and Expression

- Looked at 11 cancer types in the TCGA

Pearson Correlation b/t Methylation and Expression

- Looked at 11 cancer types in the TCGA
- Correlation between isoform and methylation
  - Take 2 kb around gene
  - Calculate correlation between methylation and isoform expression for each probe

Pearson Correlation b/t Methylation and Expression

- Looked at 11 cancer types in the TCGA
- Correlation between isoform and methylation
  - Take 2 kb around gene
  - Calculate correlation between methylation and isoform expression for each probe
- Positive correlation means isoform use increased, while a negative correlation mean isoform use decreased

Pearson Correlation b/t Methylation and Expression

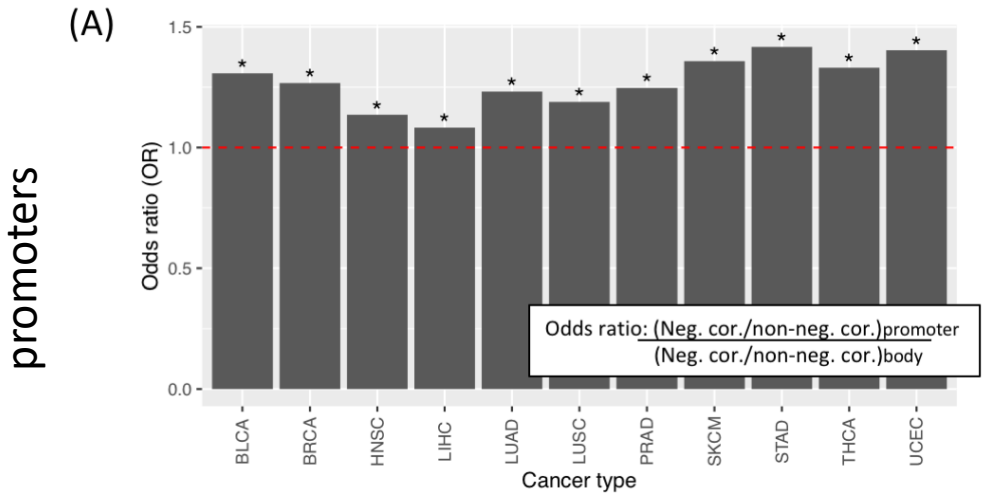
- Looked at 11 cancer types in the TCGA
- Correlation between isoform and methylation
  - Take 2 kb around gene
  - Calculate correlation between methylation and isoform expression for each probe
- Positive correlation means isoform use increased, while a negative correlation mean isoform use decreased
- Tests were corrected using empirical FDR

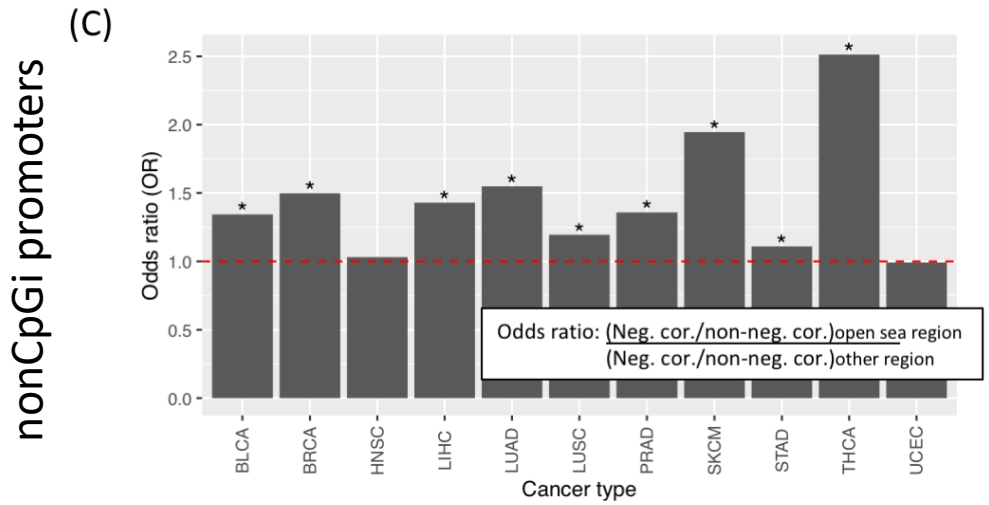
Pearson Correlation b/t Methylation and Expression

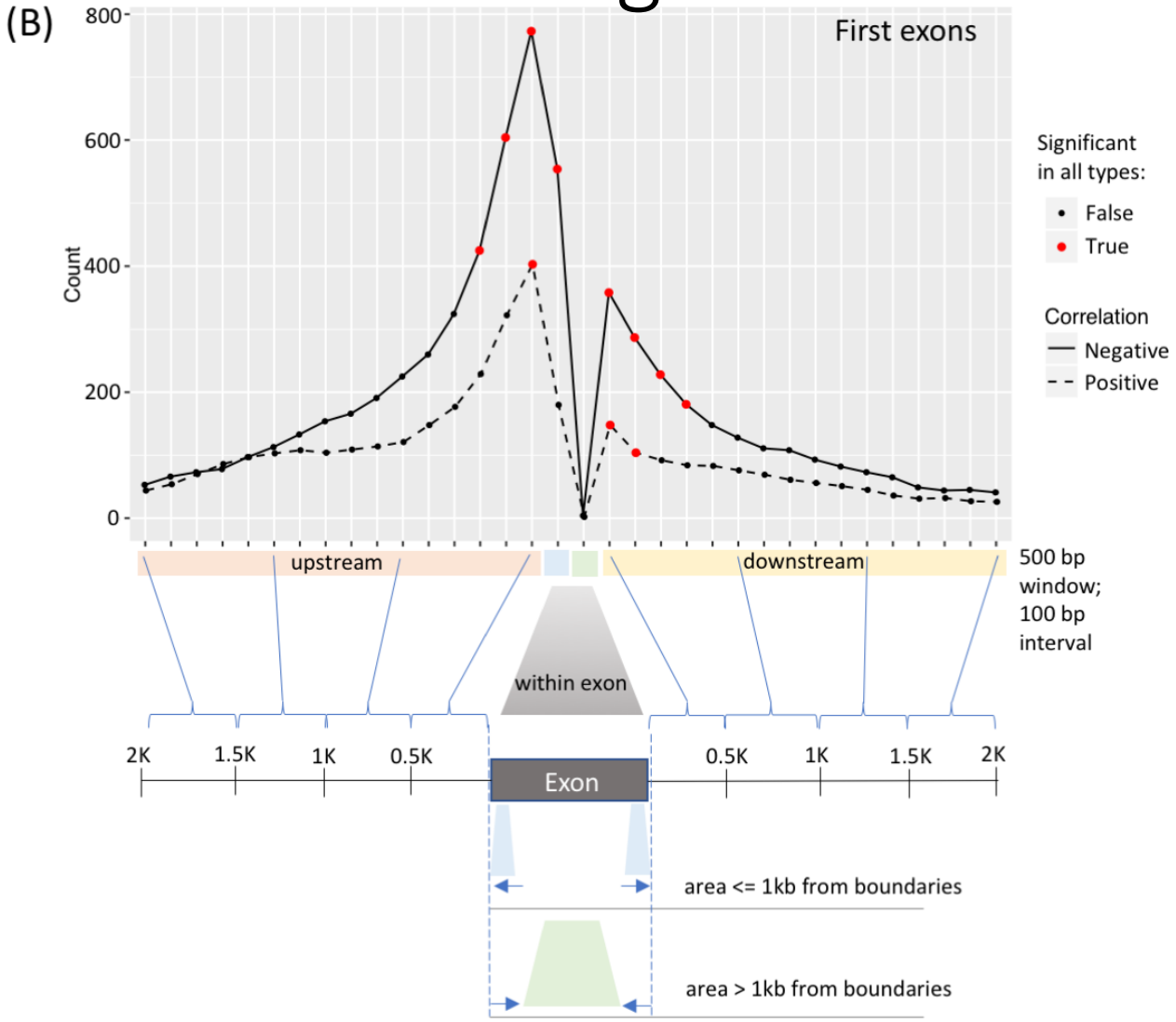
- Looked at 11 cancer types in the TCGA
- Correlation between isoform and methylation
  - Take 2 kb around gene
  - Calculate correlation between methylation and isoform expression for each probe
- Positive correlation means isoform use increased, while a negative correlation mean isoform use decreased
- Tests were corrected using empirical FDR
- Filtered results:
  - Took methylation probes that were highly variable (standard deviation in the top 25%) in tumors
  - Correlation greater than 0.3

Results

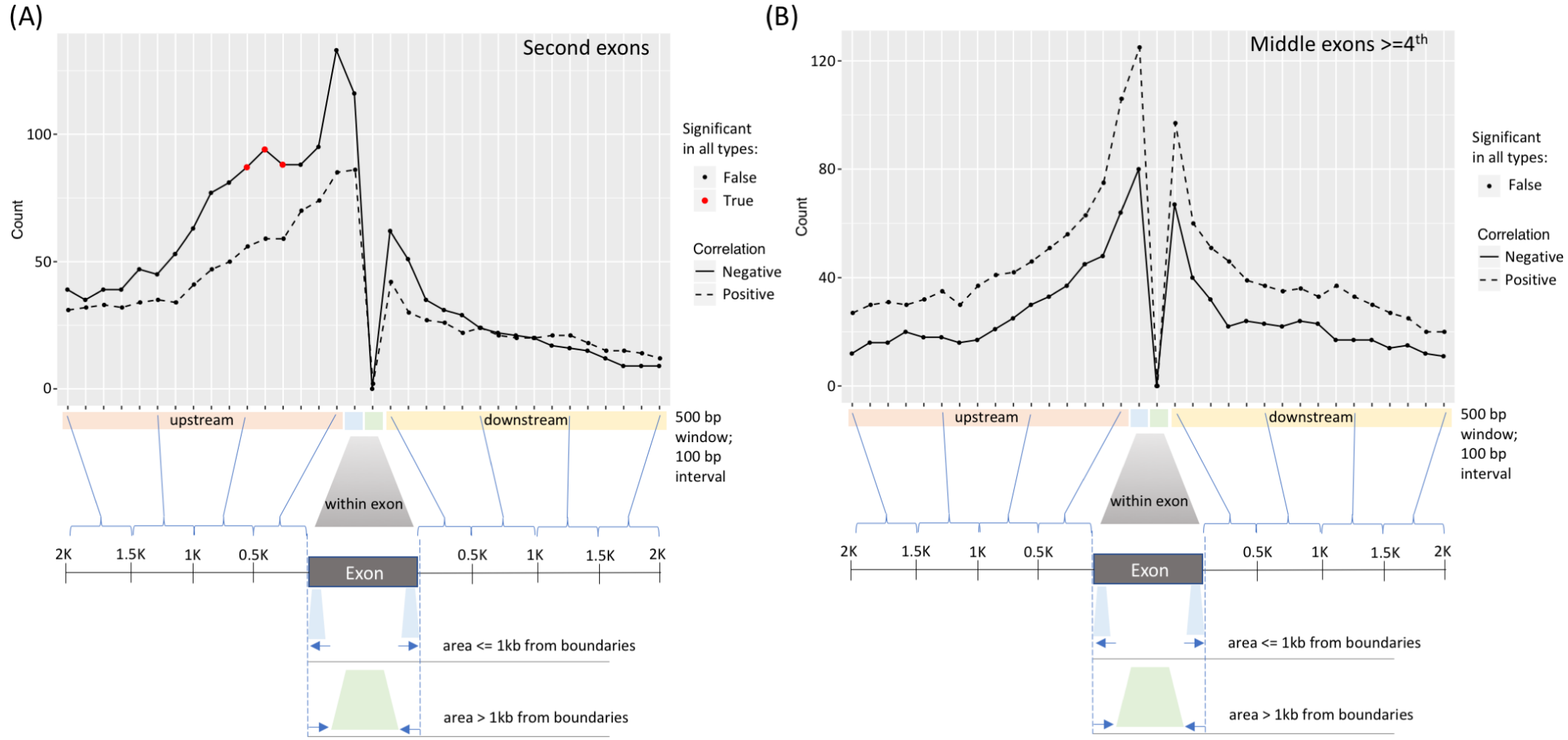
DNAm around the transcription start site is correlated with decreased isoform usage





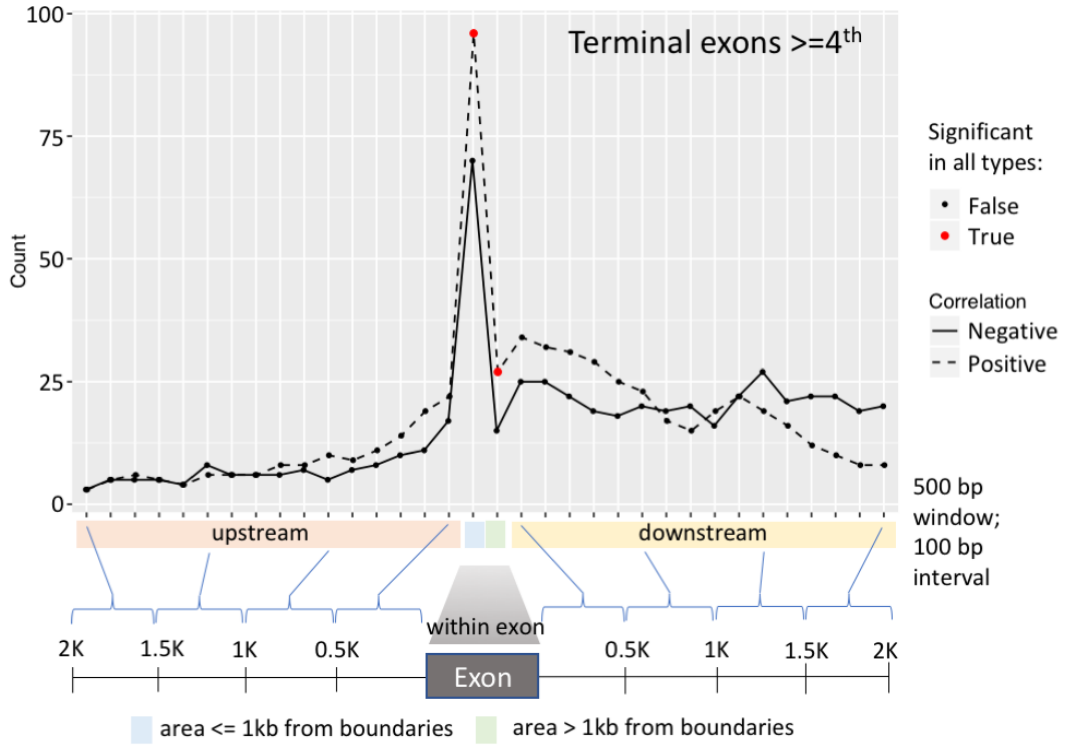


DNAm in downstream isoform positions is correlated with increased isoform usage

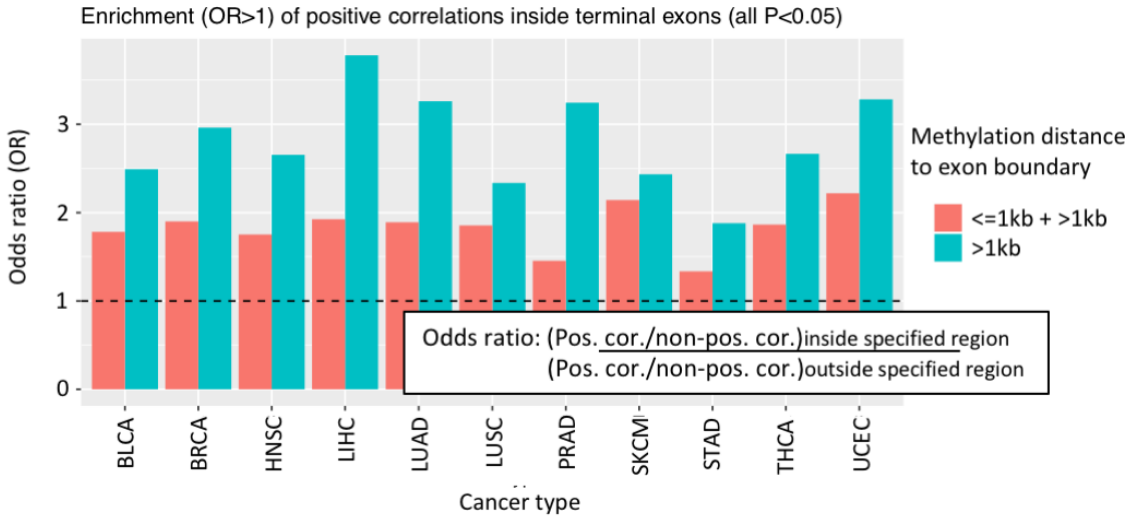


DNAm near the TTS may define the 3’ isoform boundary

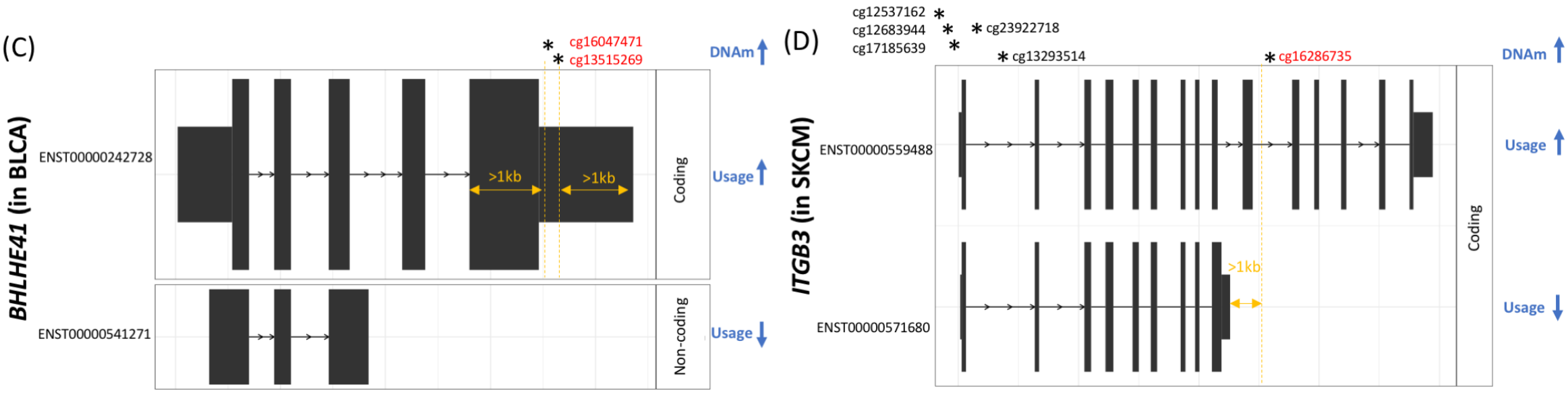
(A)



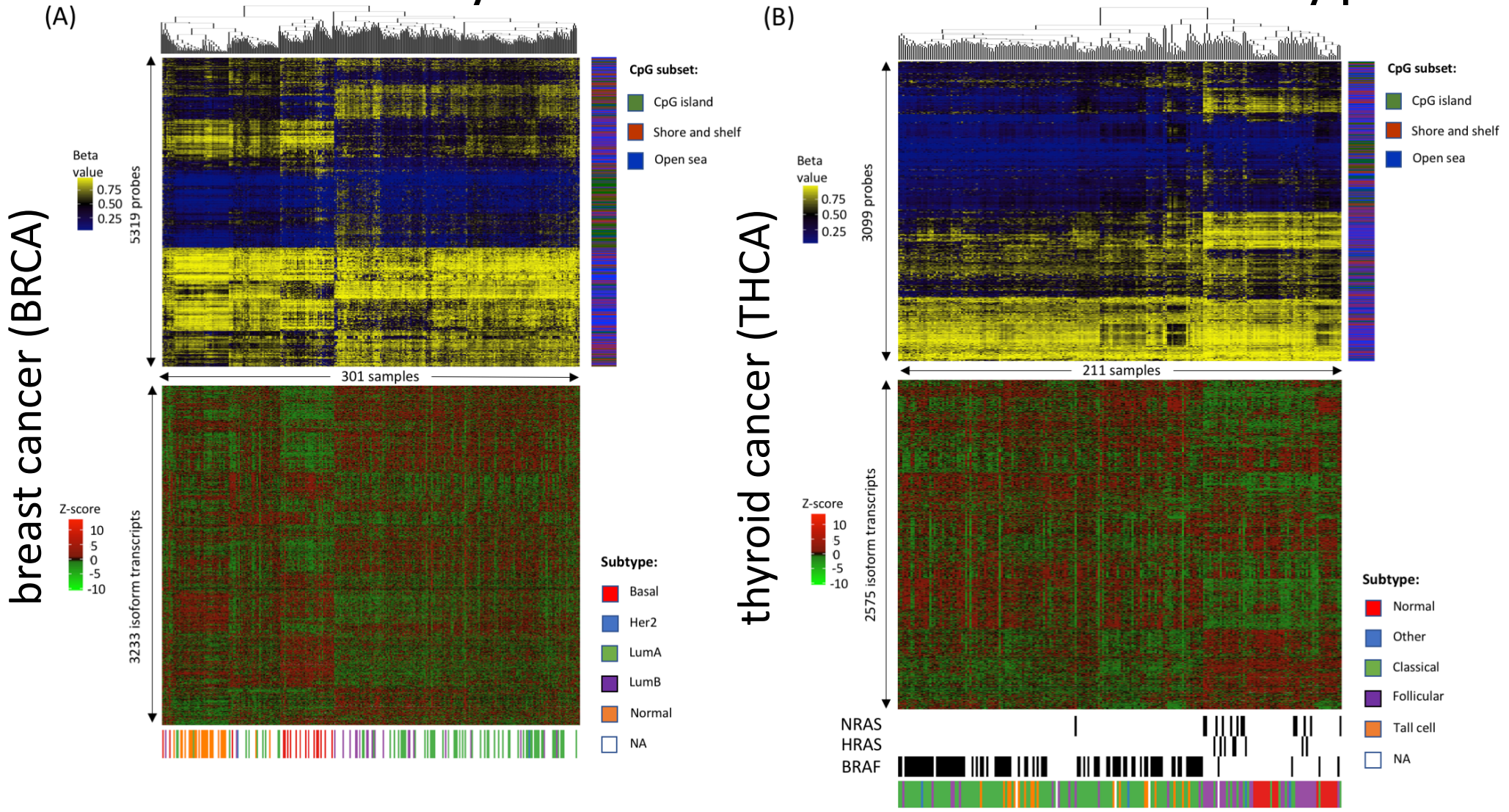
(B)



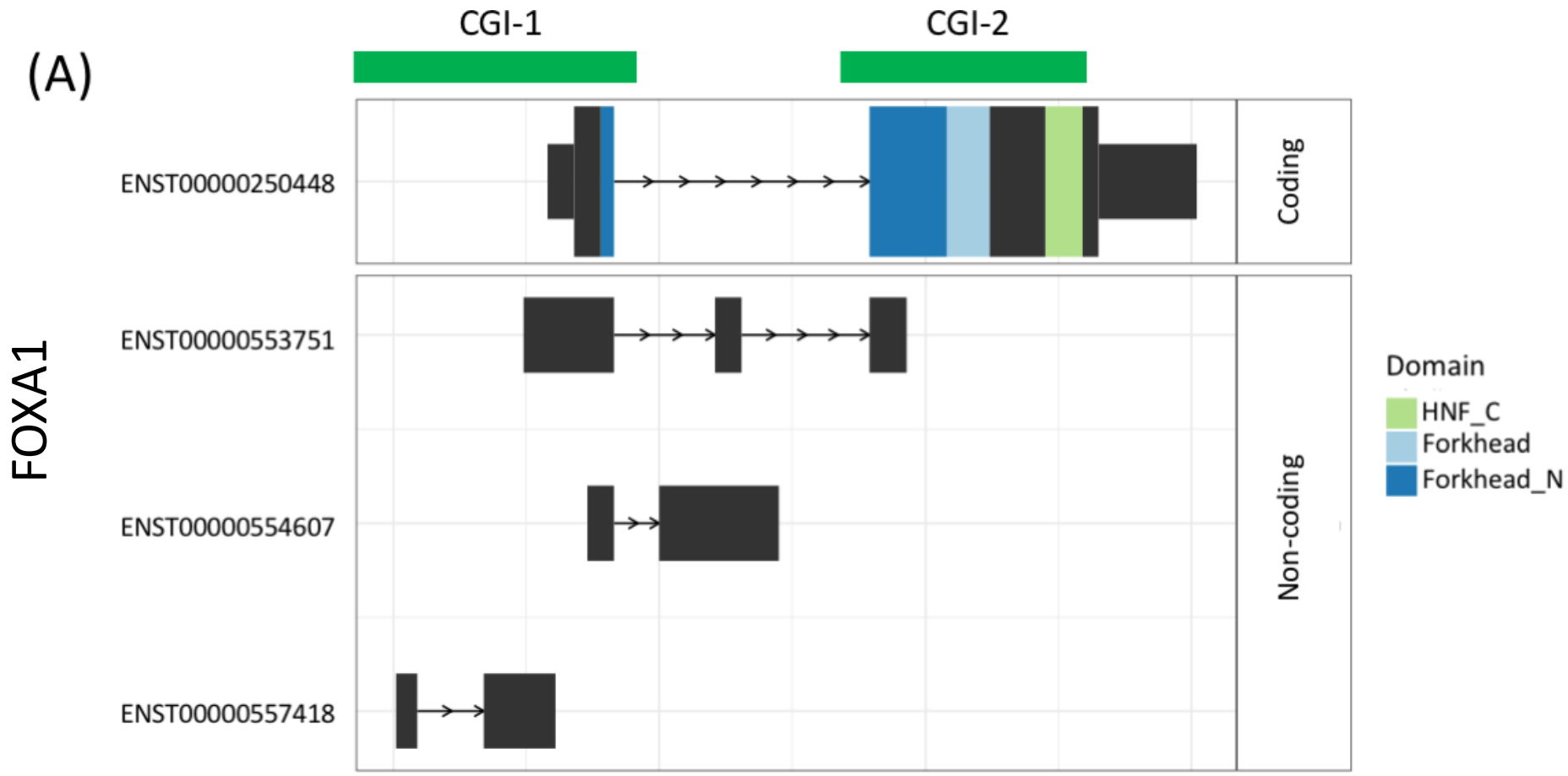
DNAm near the TTS may define the 3’ isoform boundary



Correlated DNAm and isoform usage can be used to classify tumors into know subtypes

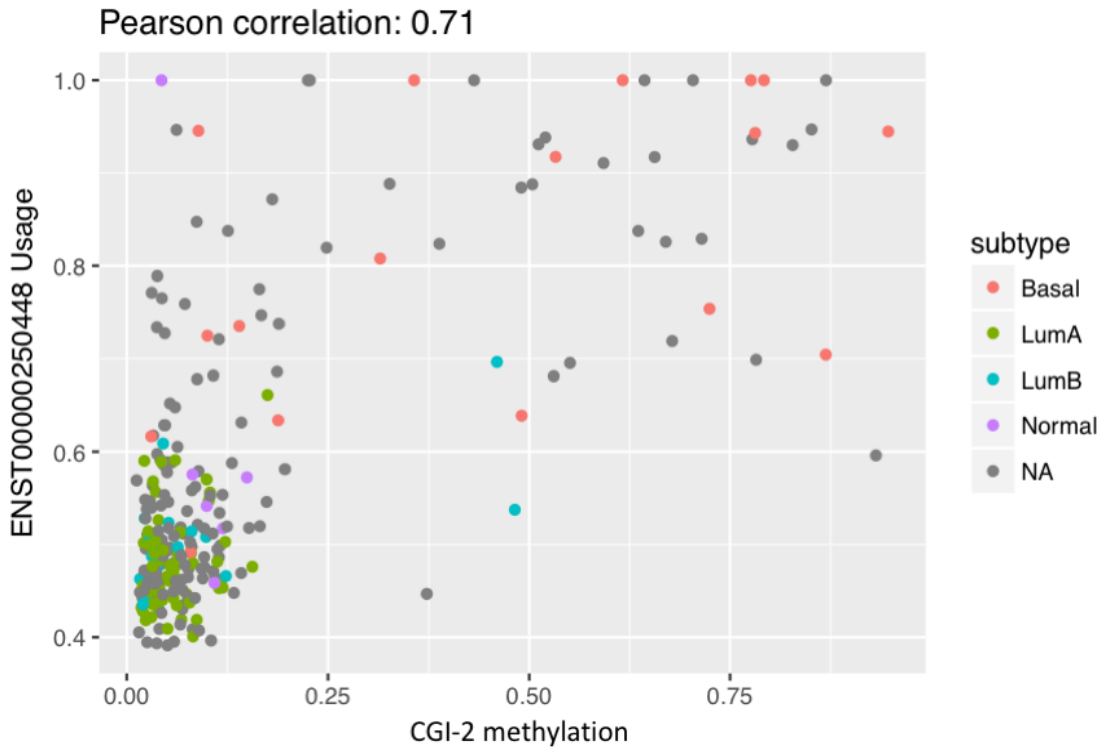


Functional implications of DNAm-correlated isoform switching in cancer

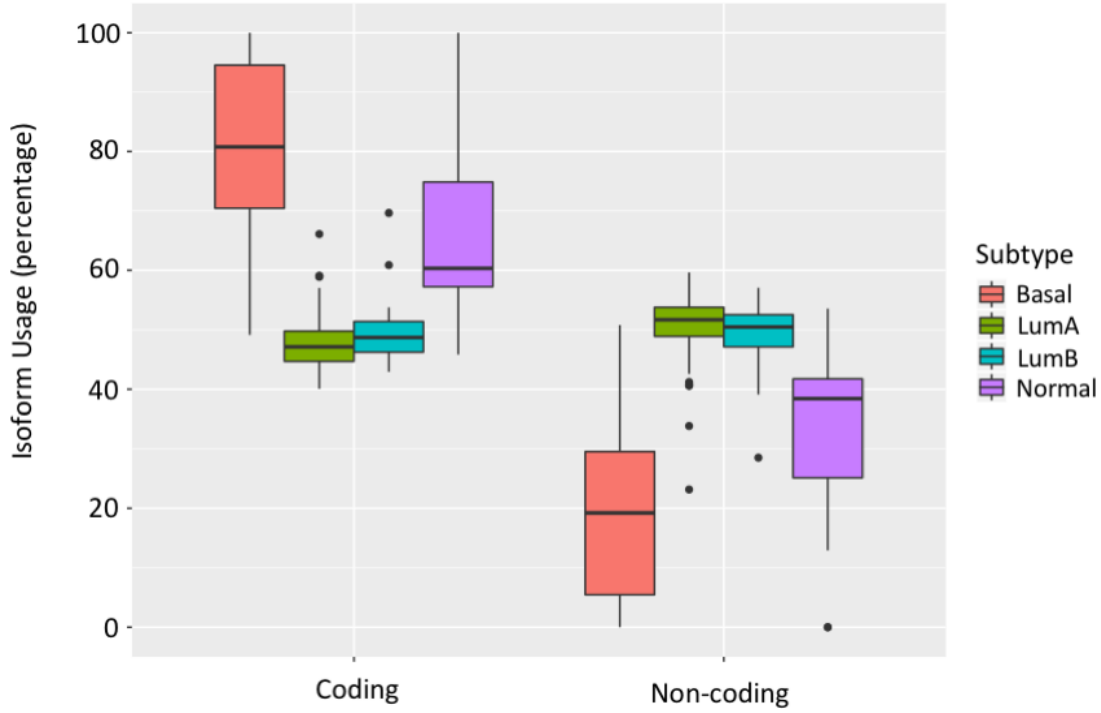


Functional implications of DNAm-correlated isoform switching in cancer

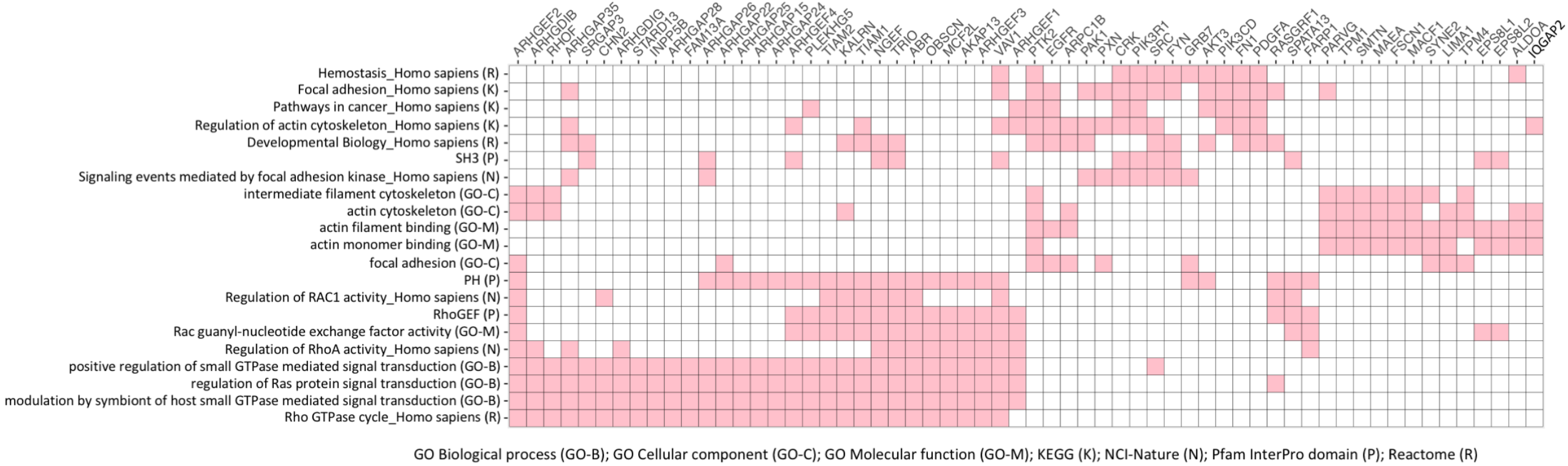
(B)



(C)



DNAm-isoform correlations denote genes enriched in cancer and other biological pathways



GO Biological process (GO-B); GO Cellular component (GO-C); GO Molecular function (GO-M); KEGG (K); NCI-Nature (N); Pfam InterPro domain (P); Reactome (R)

Conclusions

Conclusions

- Aberrant DNA methylation in cancer is correlated with isoform usage, over different cancer types
- Most of the trends reported have been identified in normal cells, but this paper demonstrates them in a cancer context.
- There are novel observations, for example, the observation that methylation in alternative terminal exons is correlated with their inclusion, leading to the proposal that the methylation signals for inclusion, that would benefit from further validation

Weakness

- Didn’t consider age-related methylation at all as far as I could tell
- Didn’t really look at the relationship between methylation and alternative splicing
- Had to use array data for methylation
  - Limits the number of CpGs for analysis
  - Arrays have known accuracy problems
- No code available
- No validation of work in other data, wet lab

Thank you!