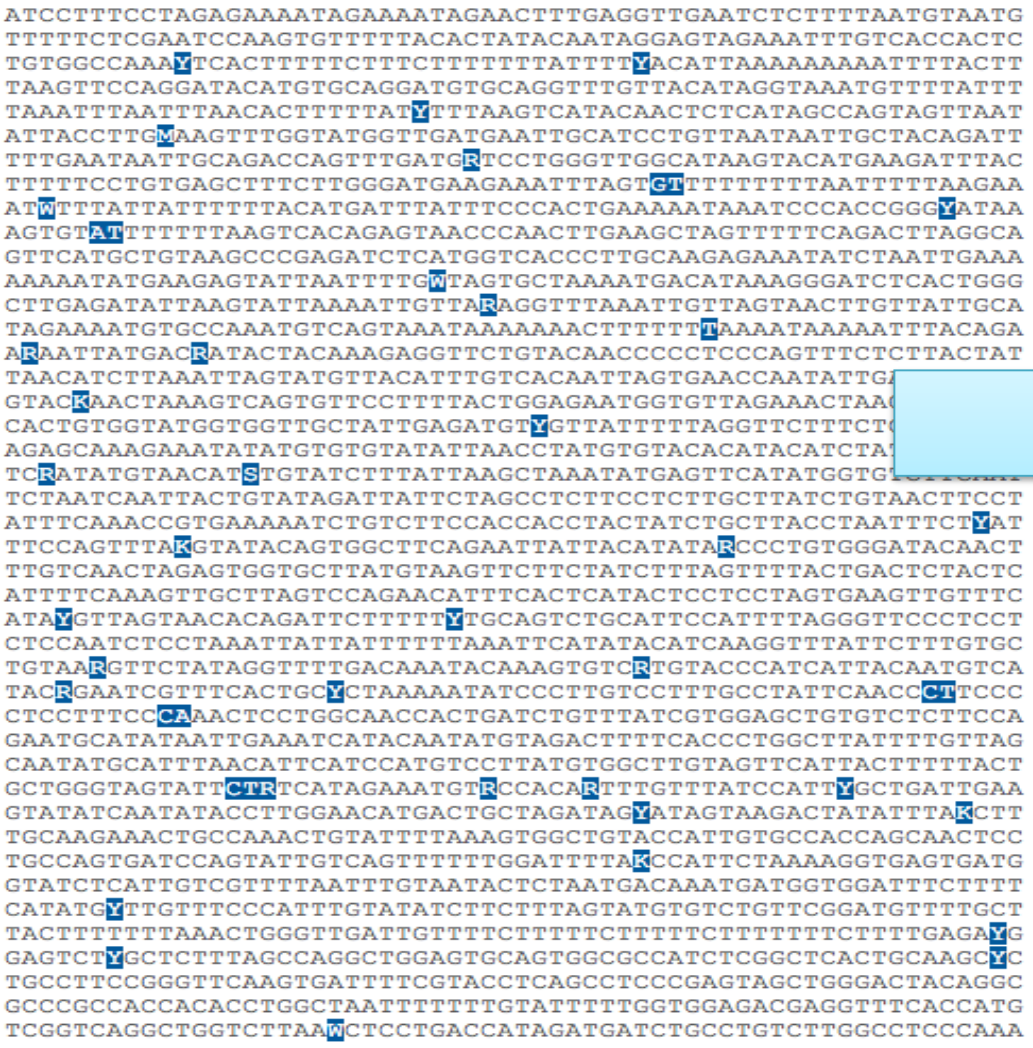
Genetic Variation and Transcriptional Regulation

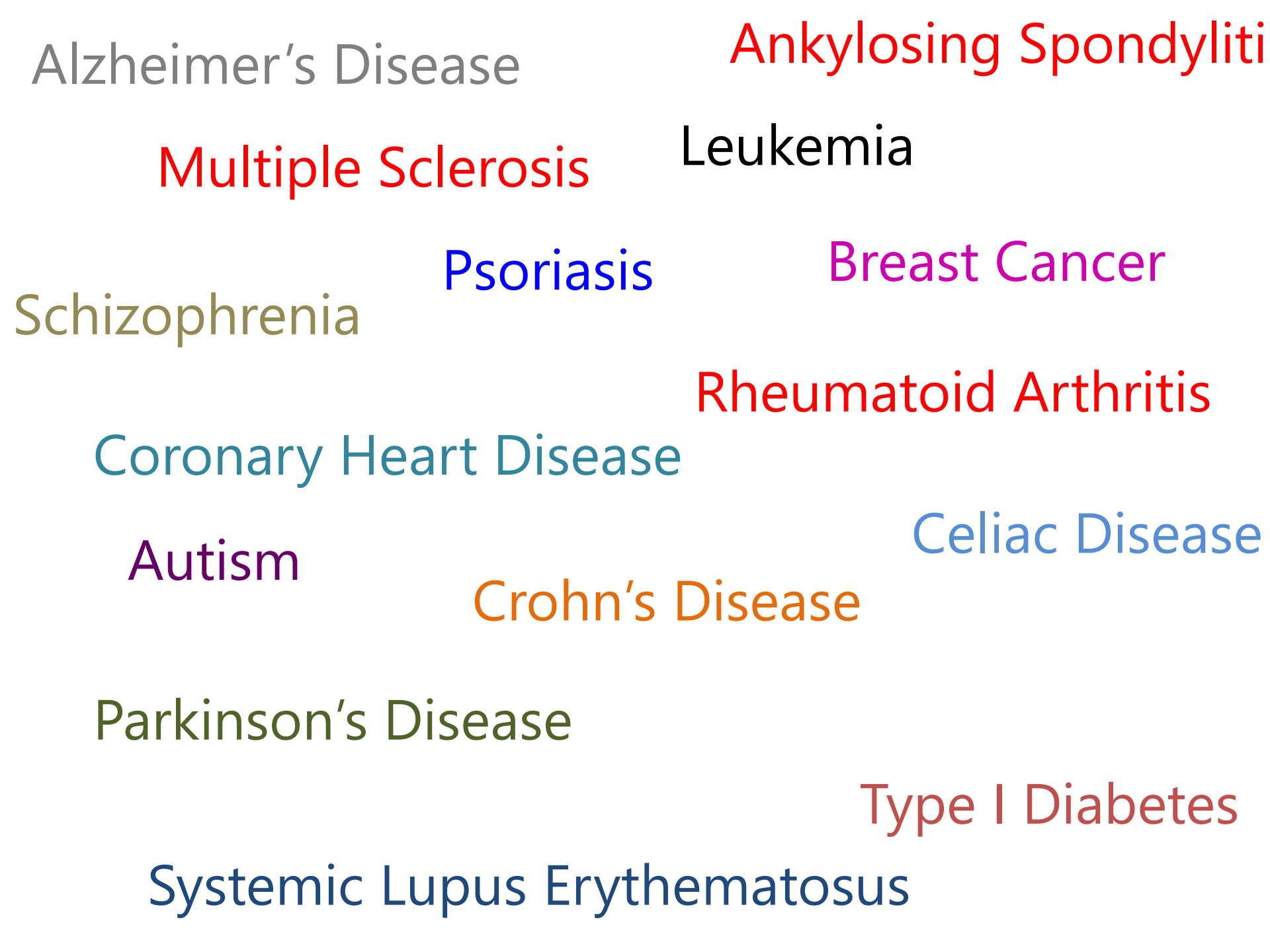
BIOM262 - **Quantitative Methods in Genetics and Genomics**

Graham McVicker – Salk Institute

Which genetic variants affect human traits?



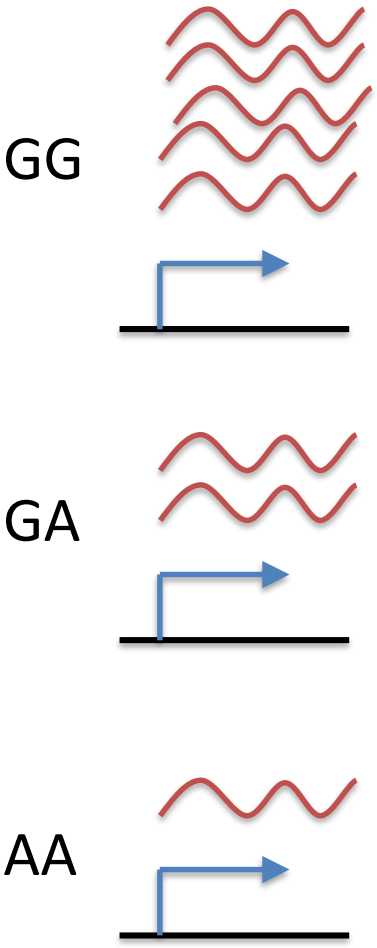
?

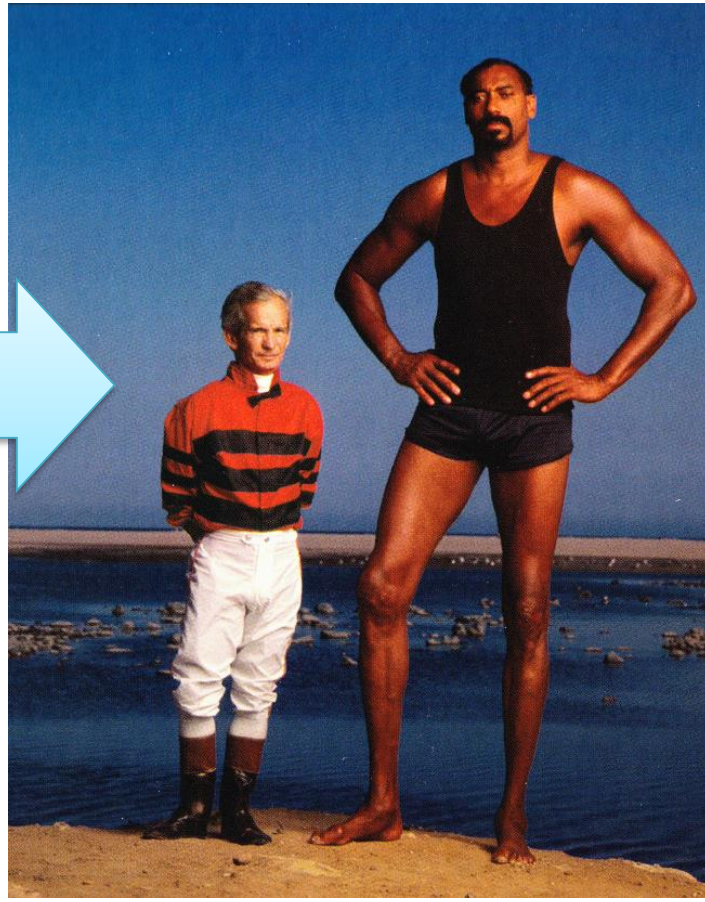

Alzheimer's Disease
Ankylosing Spondyliti
Leukemia
Multiple Sclerosis
Psoriasis
Breast Cancer
Schizophrenia
Rheumatoid Arthritis
Coronary Heart Disease
Celiac Disease
Autism
Crohn's Disease
Parkinson's Disease
Type I Diabetes
Systemic Lupus Erythematosus

Topics

- Genome-wide association studies
- Molecular quantitative trait loci (QTLs)
- Gene expression QTLs
- Chromatin QTLs
- DNA methylation QTLs
- Intersection of molecular QTLs and GWAS

Molecular traits as an intermediate phenotype





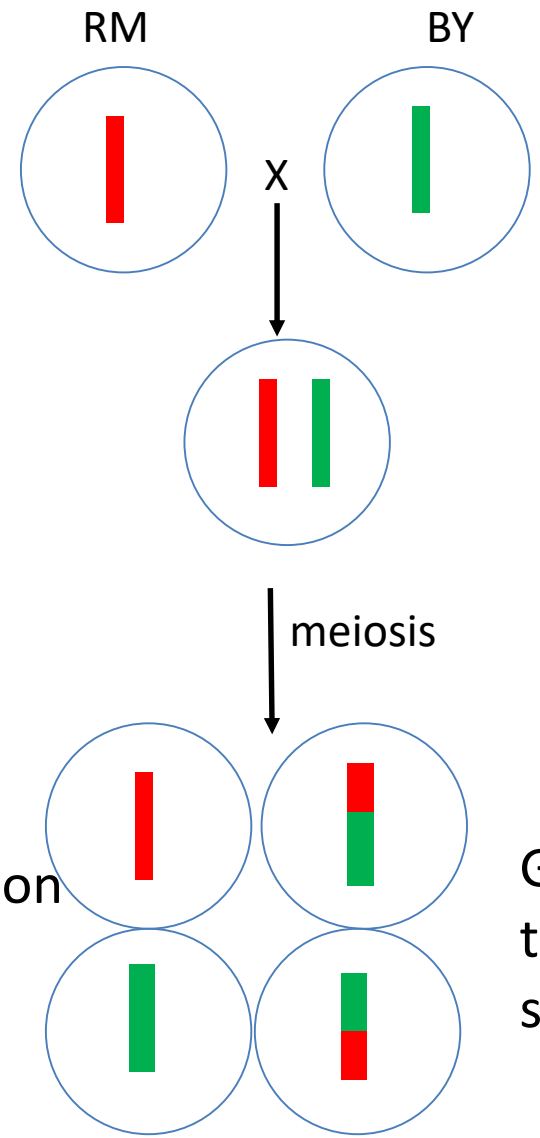
Why use molecular traits?

- Close to underlying genetics
  - affected by small number of polymorphisms
  - require smaller sample sizes
- Can measure 1000s of traits in single experiment (e.g. RNA-seq)
- Reveal molecular basis of organismal traits
  - implicate specific cell types

Discussion

- What else could be used as molecular or cellular traits?
- How would you measure these?

The first gene expression QTL mapping study



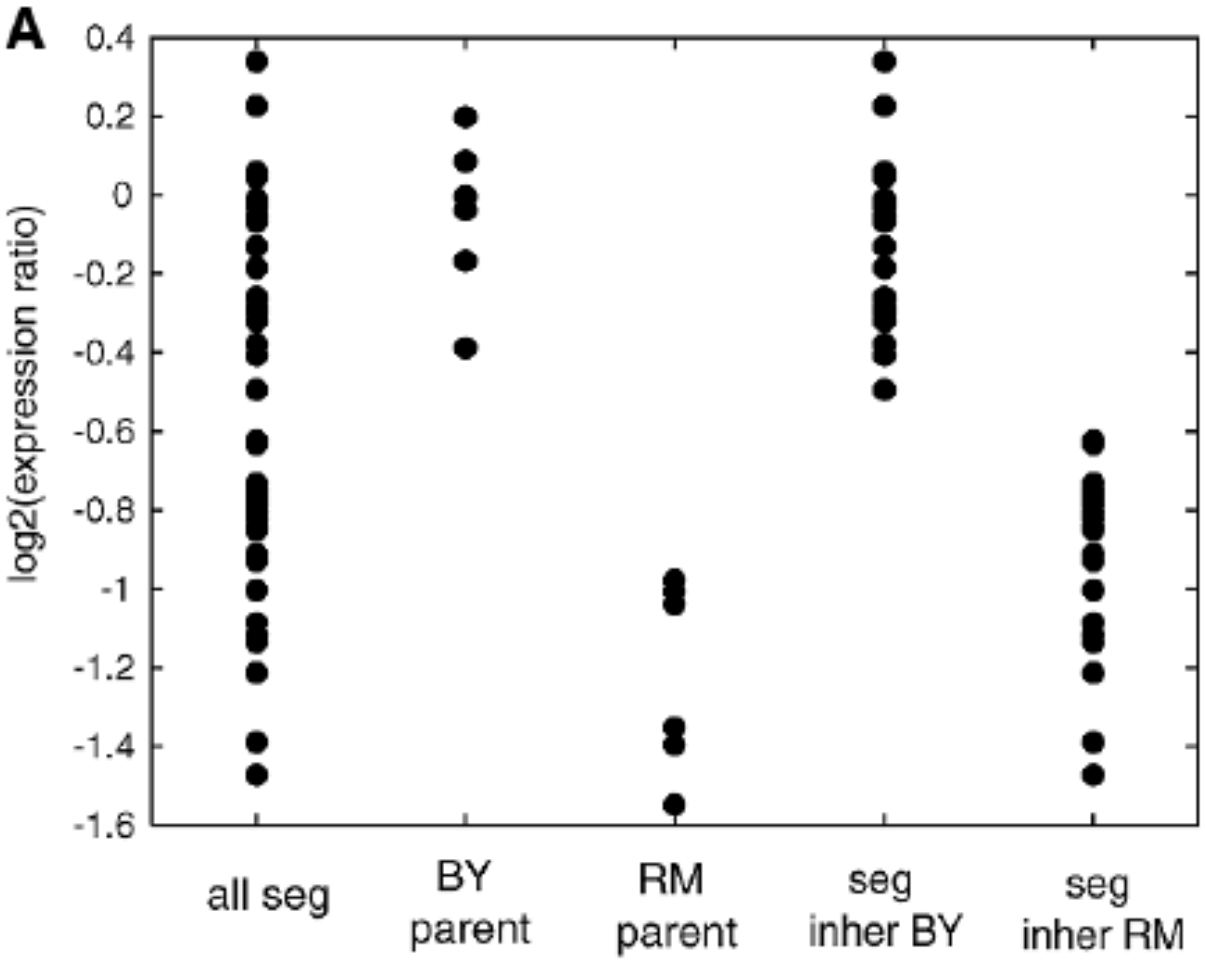
Cross two diverged yeast strains “RM” and “BY”

Measure Gene Expression with microarrays

Genotype “segregants” to identify parental genome segments

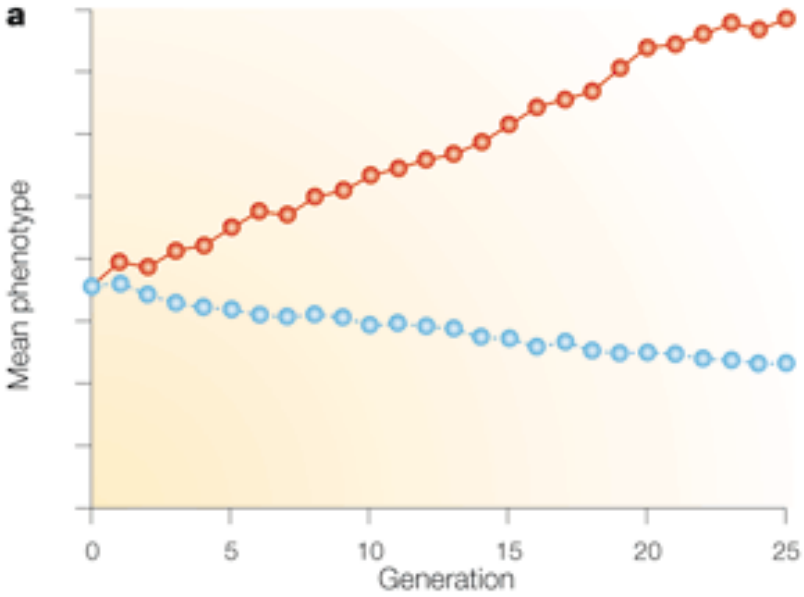
Brem et al. 2002

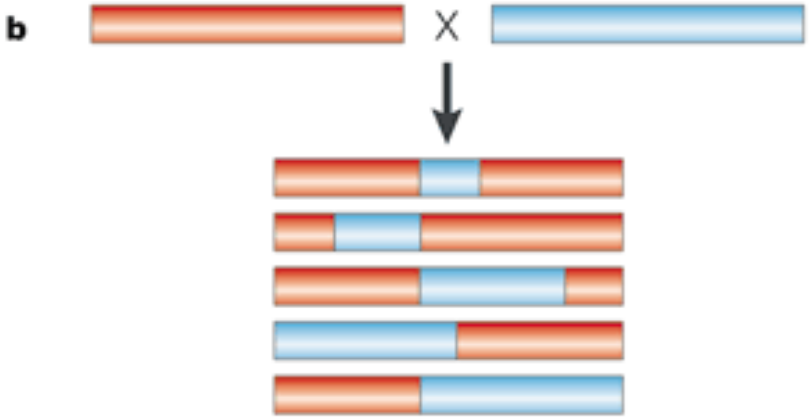
A yeast eQTL

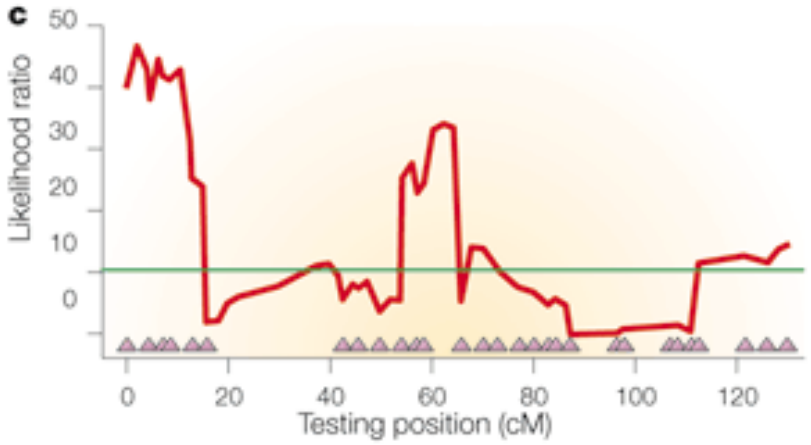


Brem et al. 2002

Quantitative Trait Locus Mapping with recombinant inbred lines

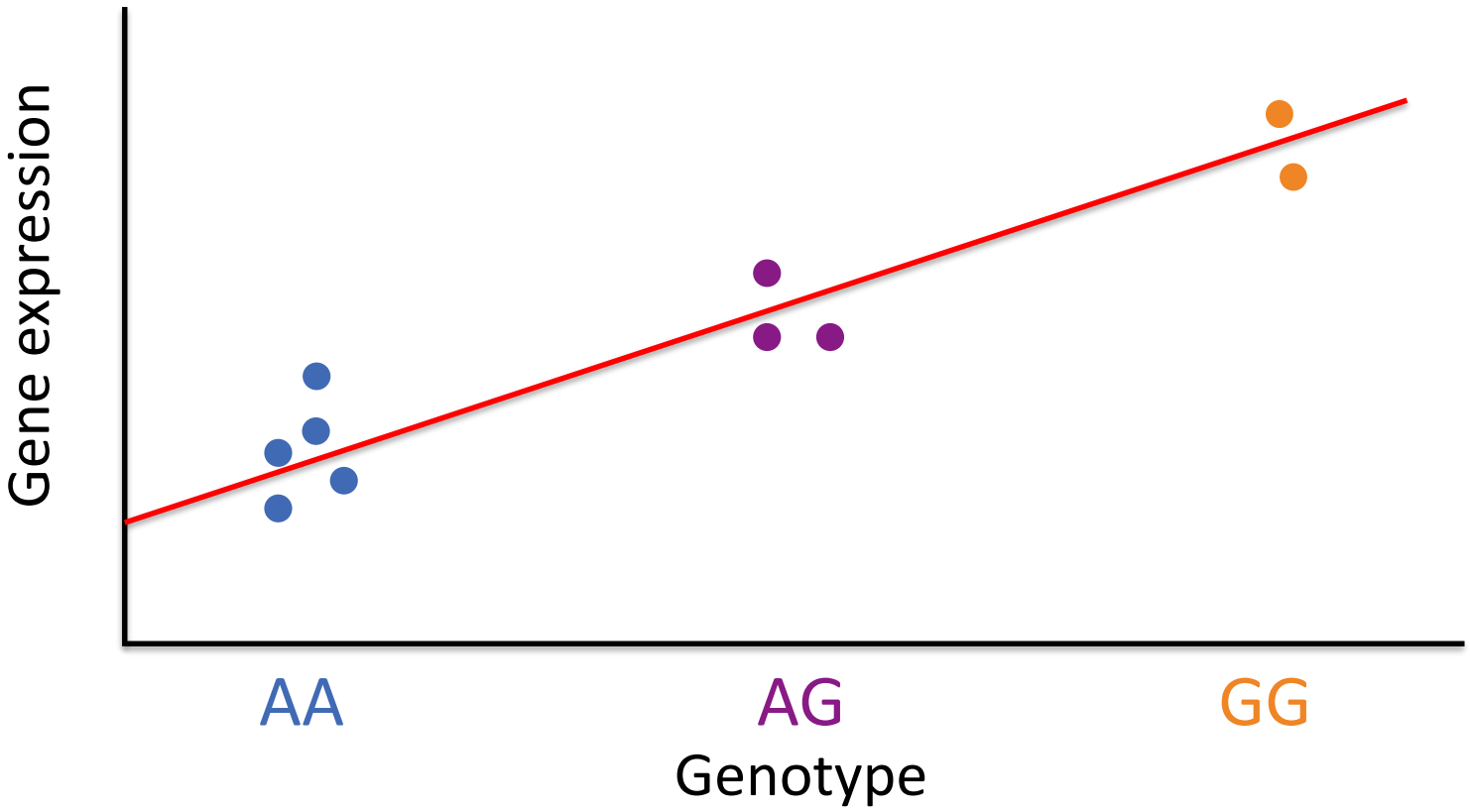




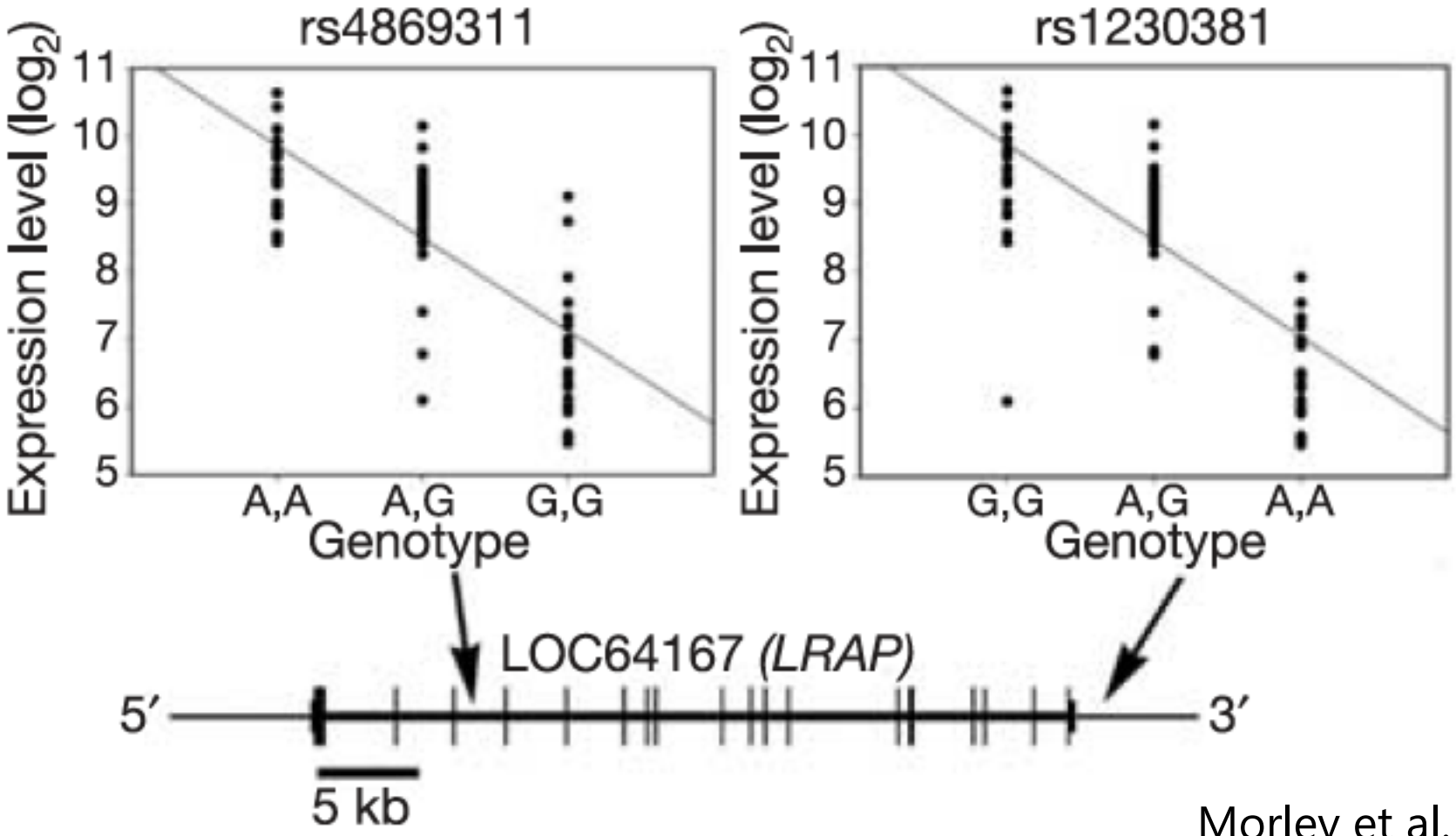


Lango-Allen et al. 2010

eQTL mapping in humans

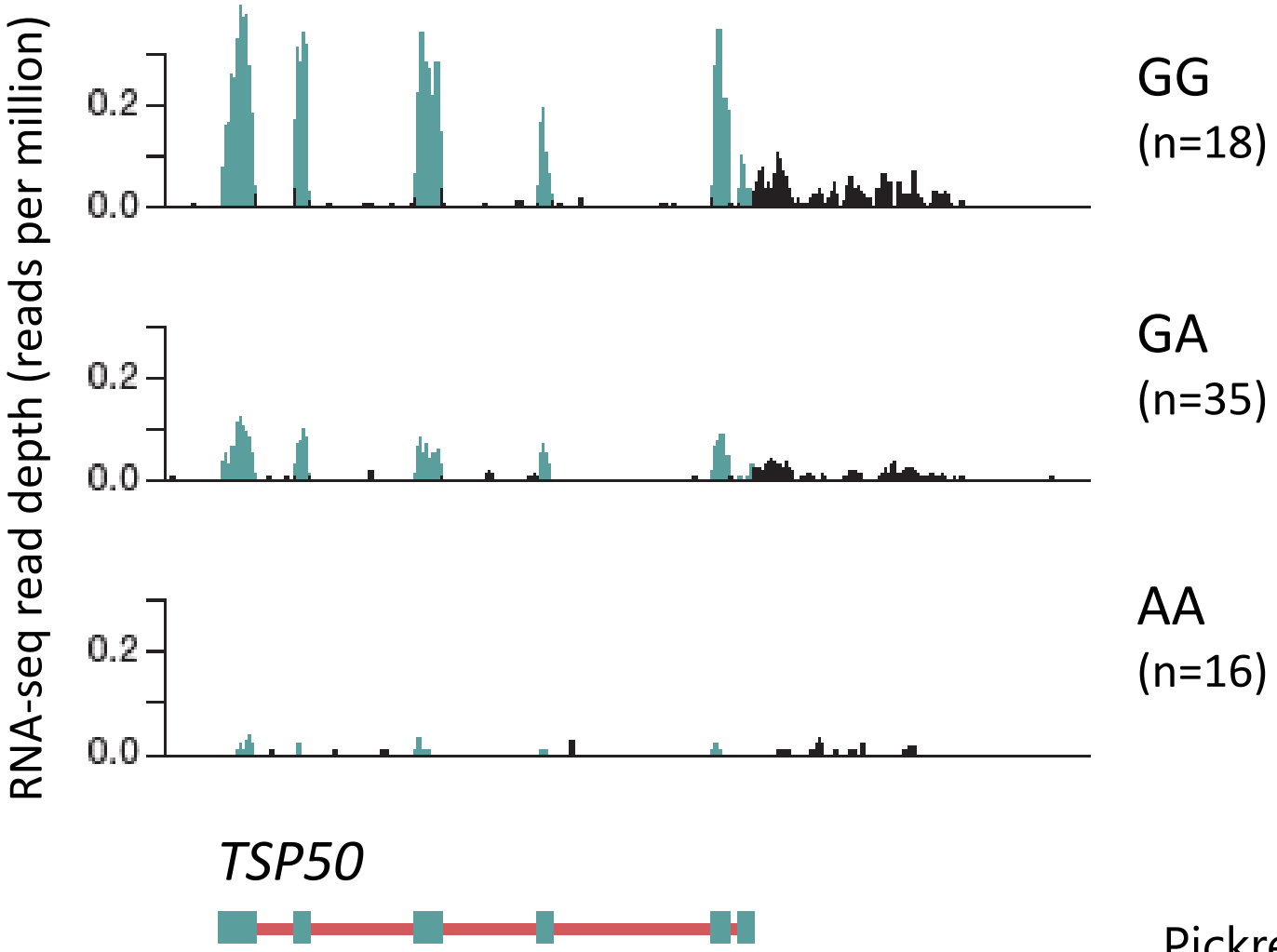


Pedigree-based mapping of eQTLs in human B cell lines



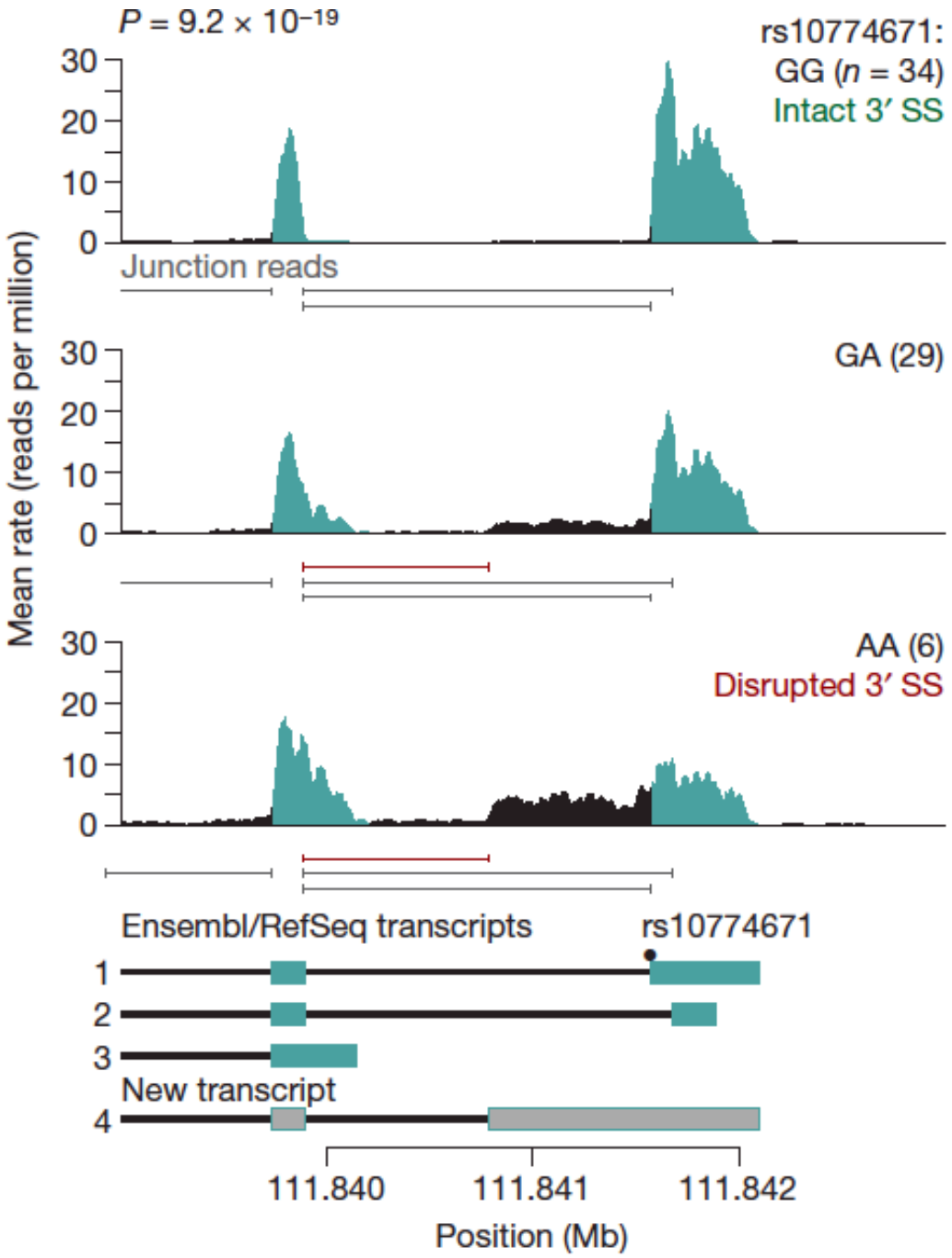
Morley et al. 2004

eQTL mapping with RNA-seq



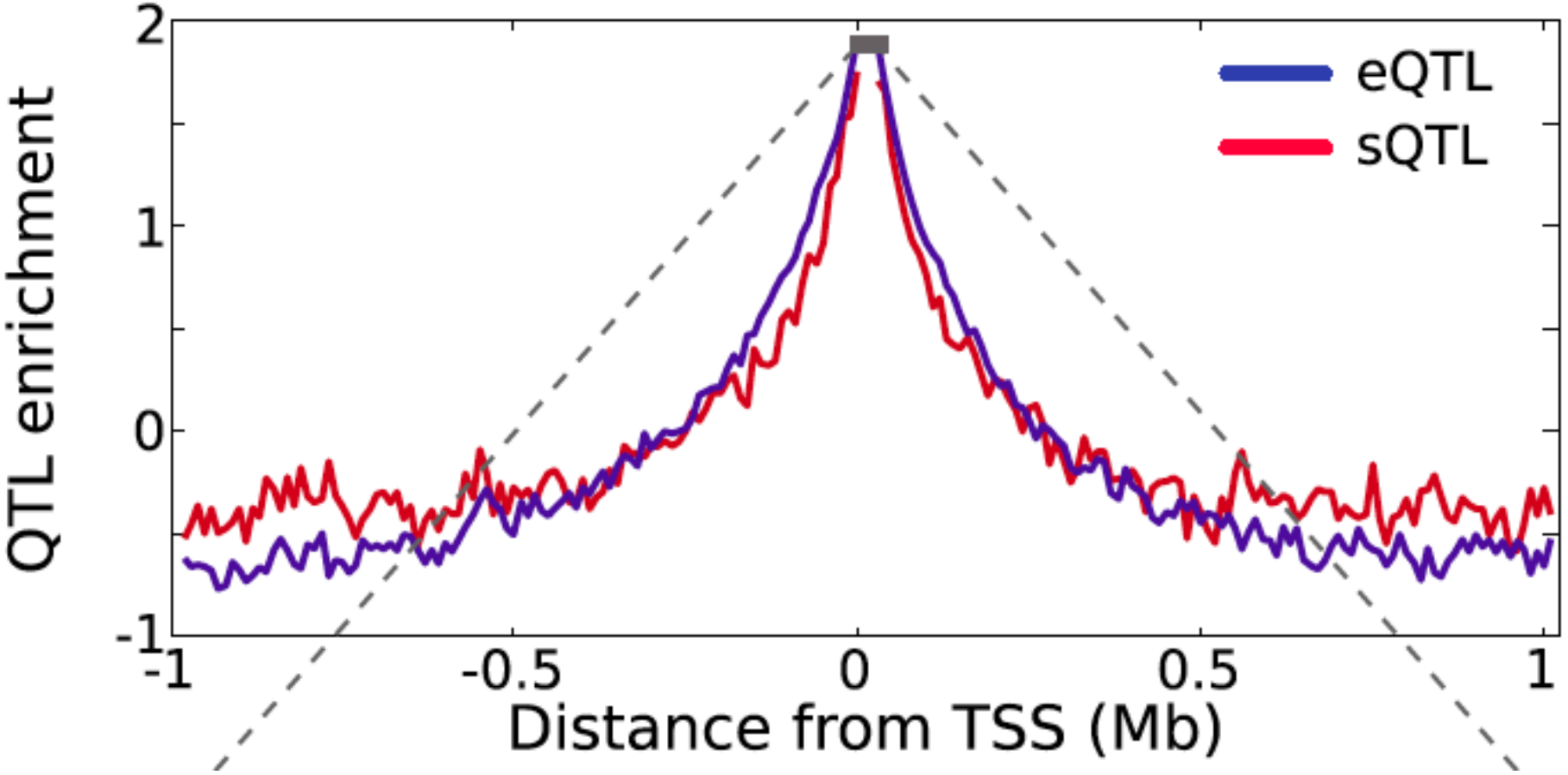
Pickrell et al. 2010

Splicing QTLs (sQTLs)



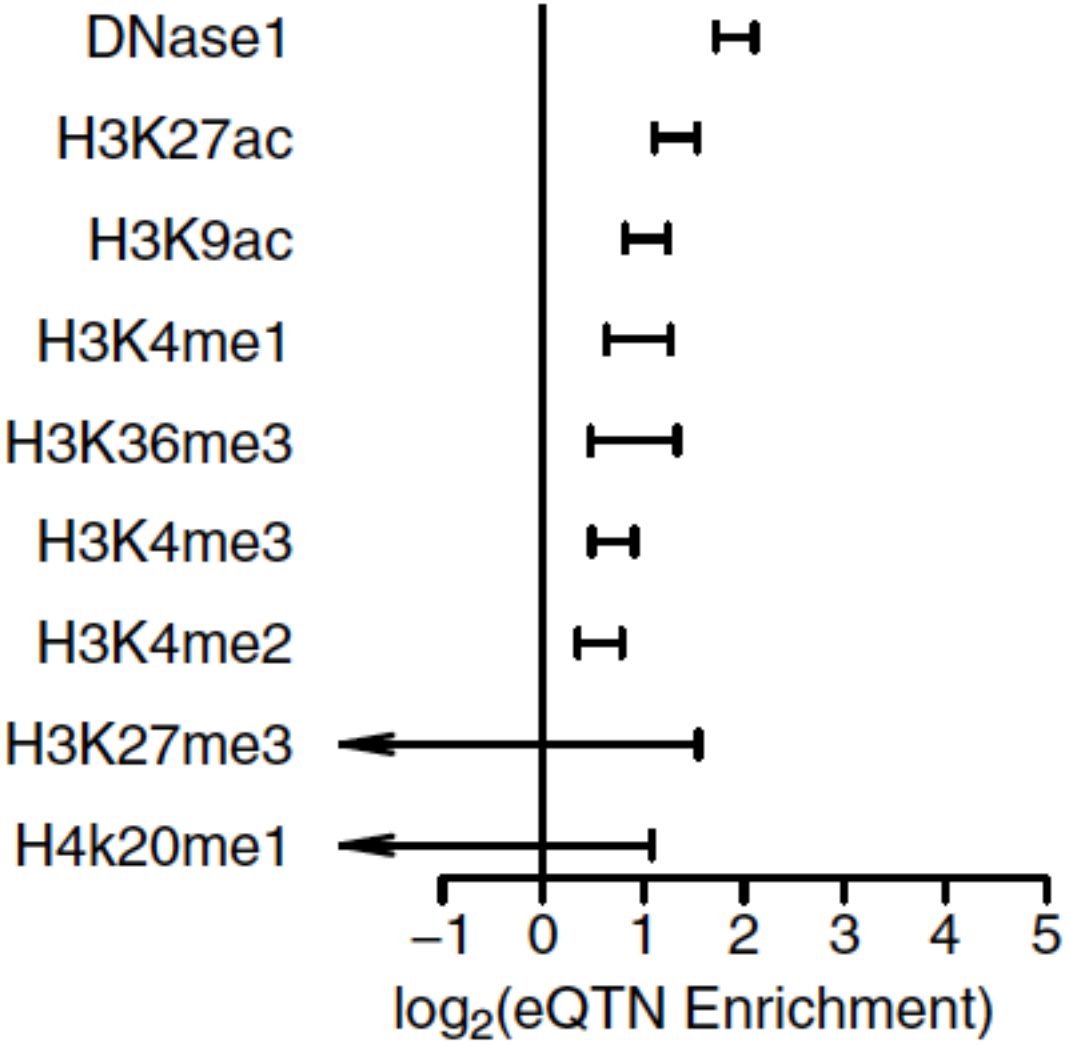
Pickrell_et al. 2010

eQTLs are enriched near transcription start sites



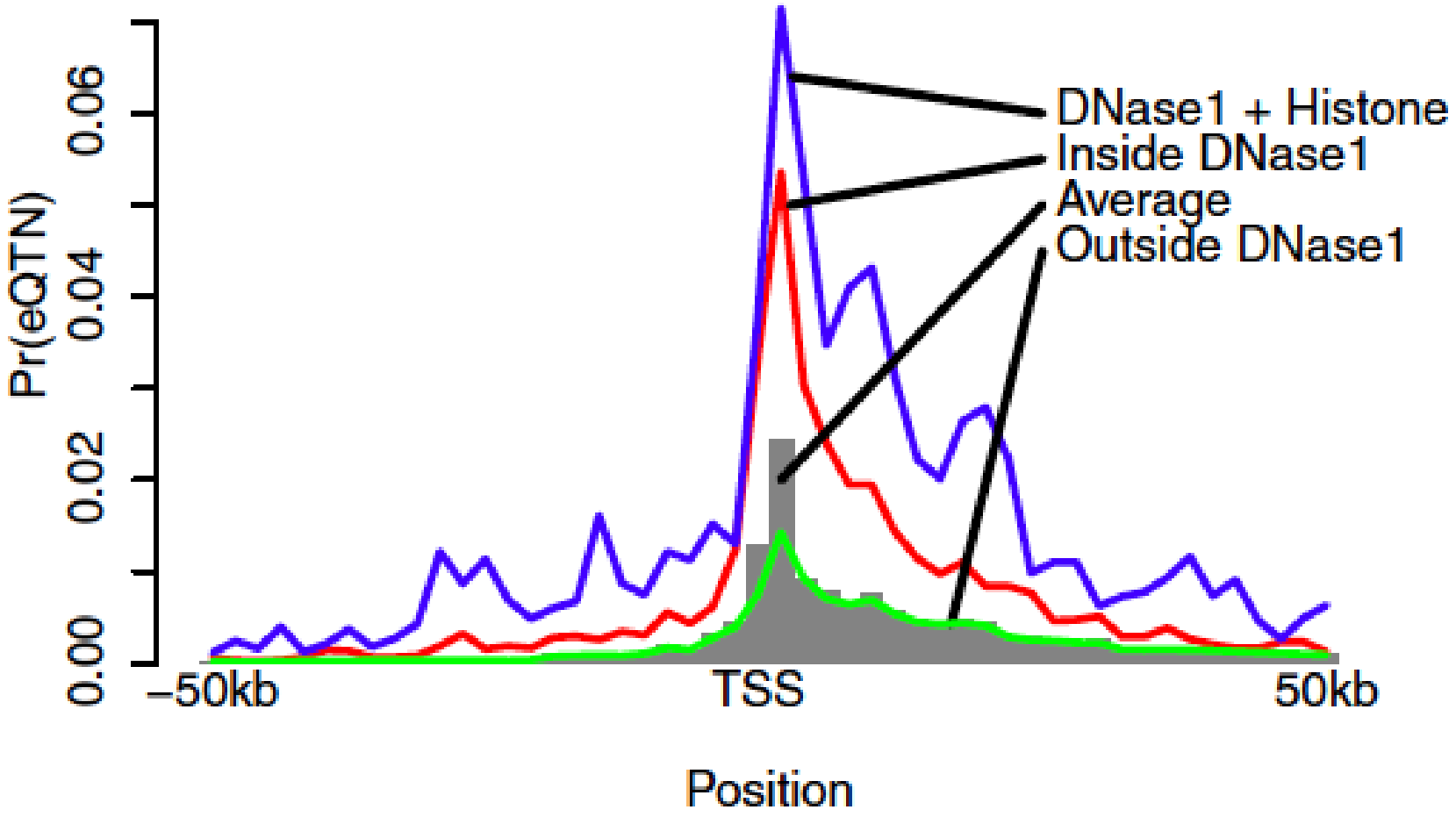
Battle et al. 2014

eQTLs are enriched in regions with open/active chromatin



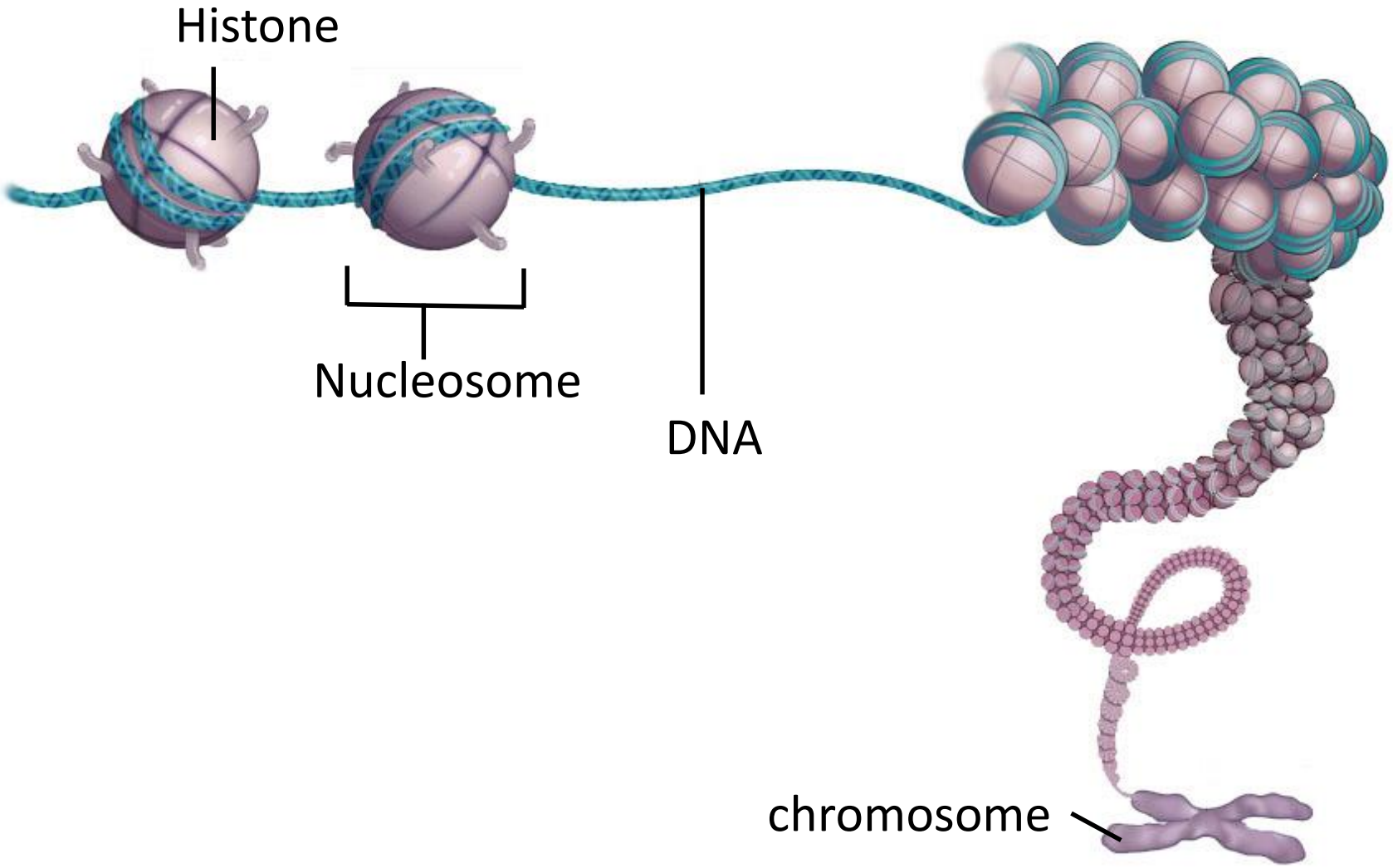
Gaffney et al. 2012

Distance from TSS and DNase sensitivity are predictive of eQTLs



Gaffney et al. 2012

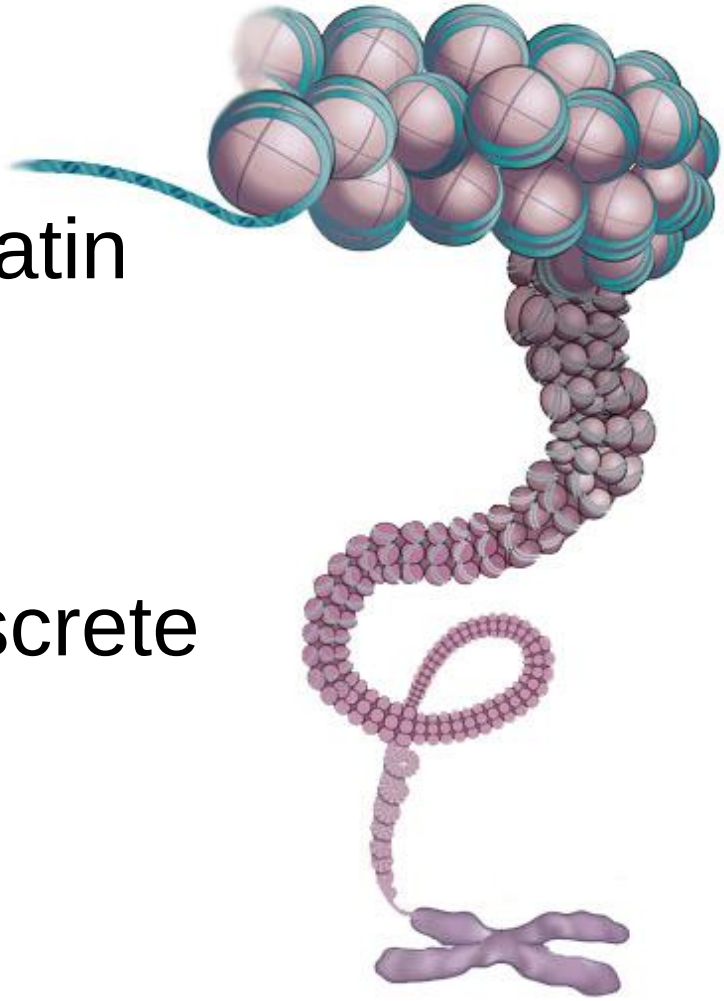
Chromatin as a molecular trait



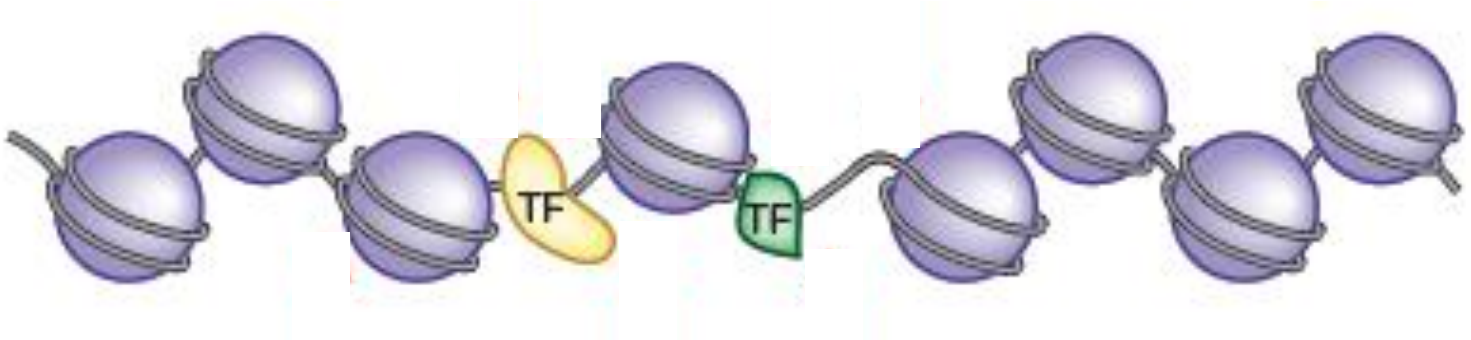
Ecker et al. 2012

Discussion

- What aspects of chromatin could be treated as a molecular trait?
- Do these traits have discrete or continuous values?

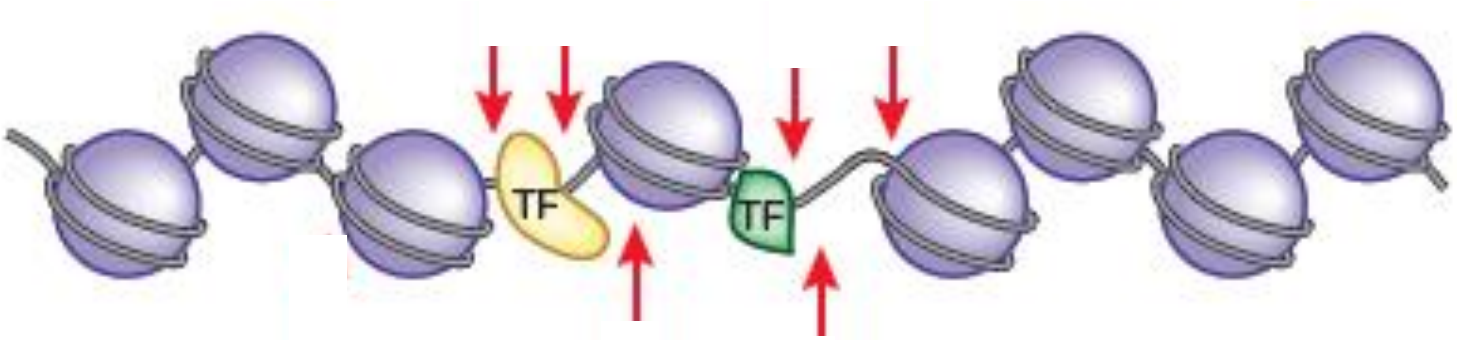


Nucleosomes are depleted in regulatory regions



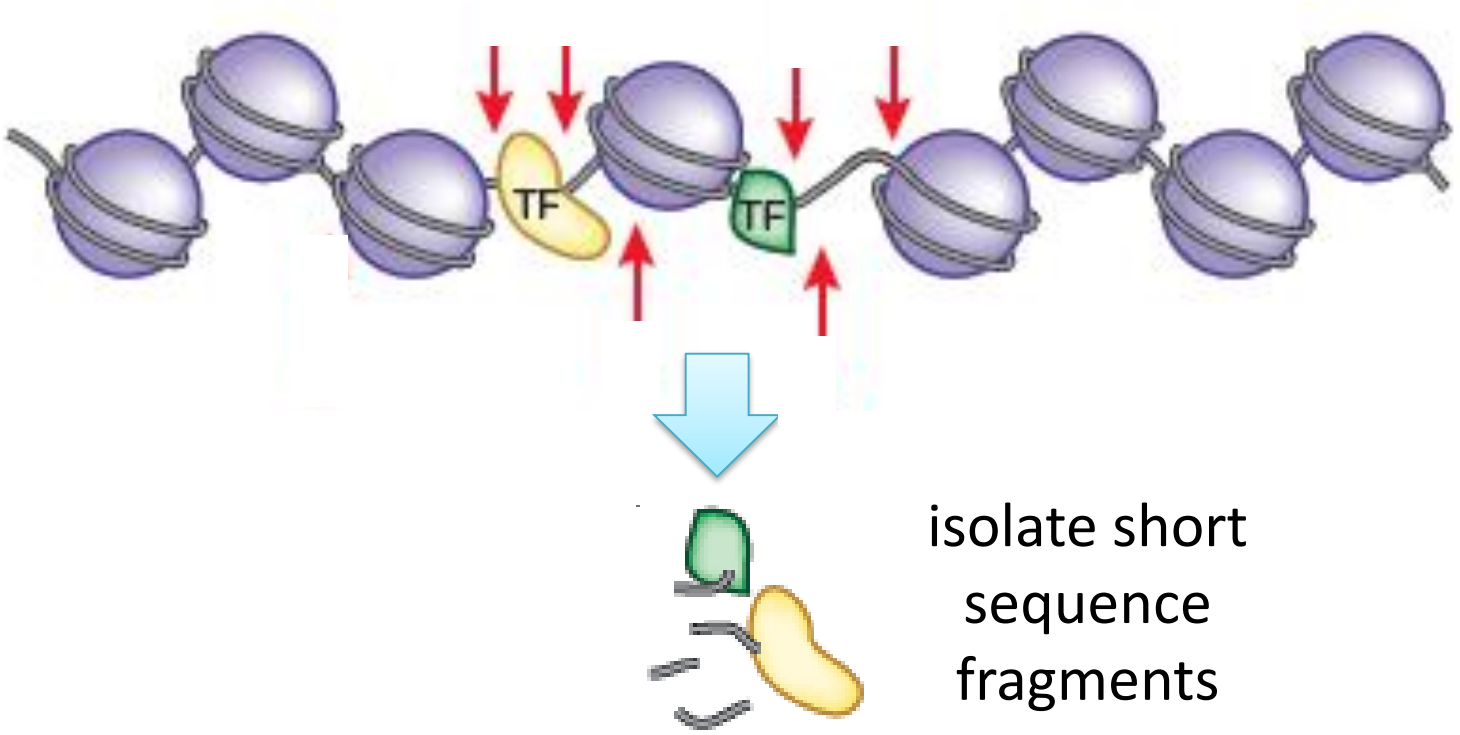
Raj & McVicker 2013

Measuring chromatin accessibility with DNase-seq



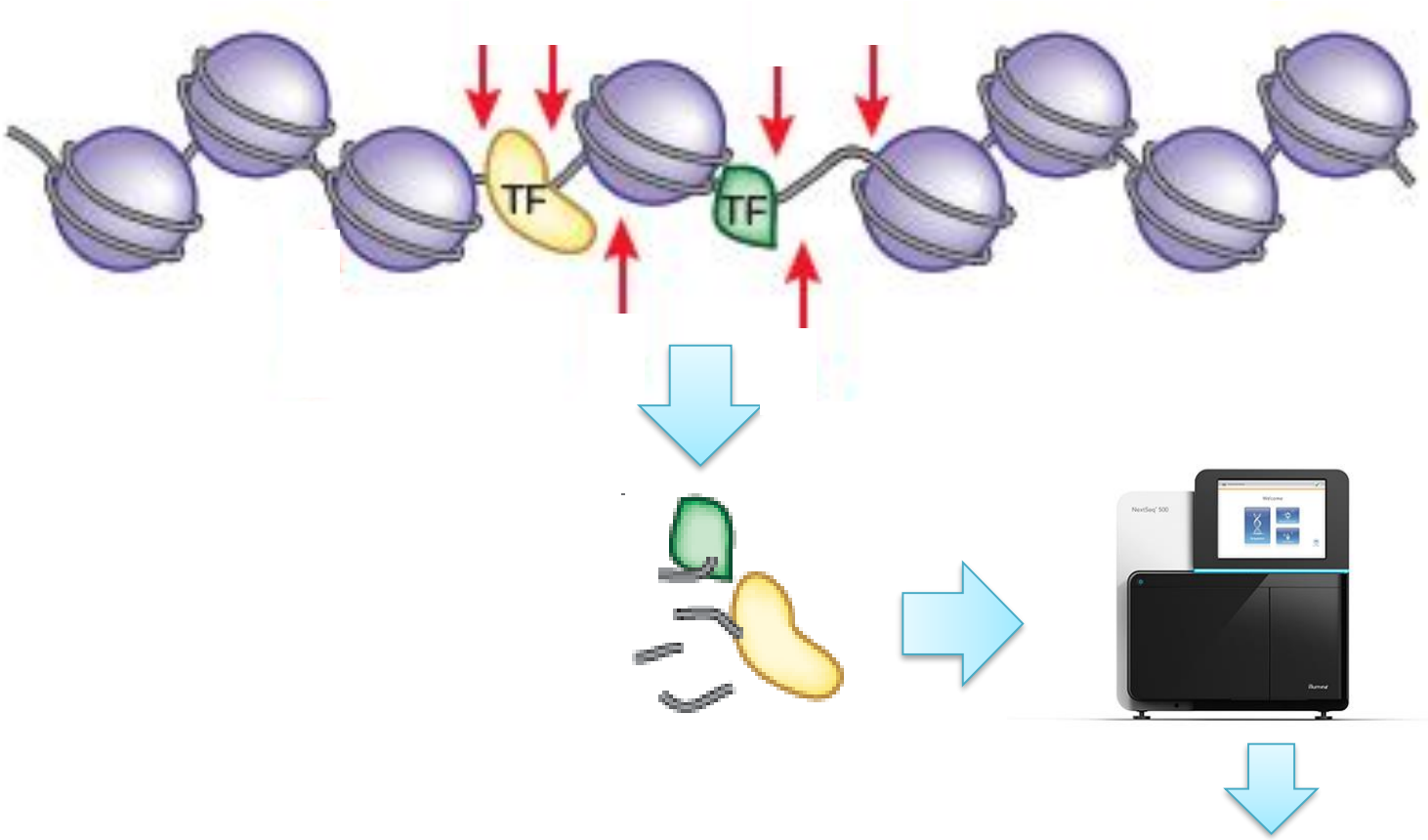
digest chromatin with Deoxyribonuclease 1 (DNase1)

Measuring chromatin accessibility with DNase-seq



isolate short sequence fragments

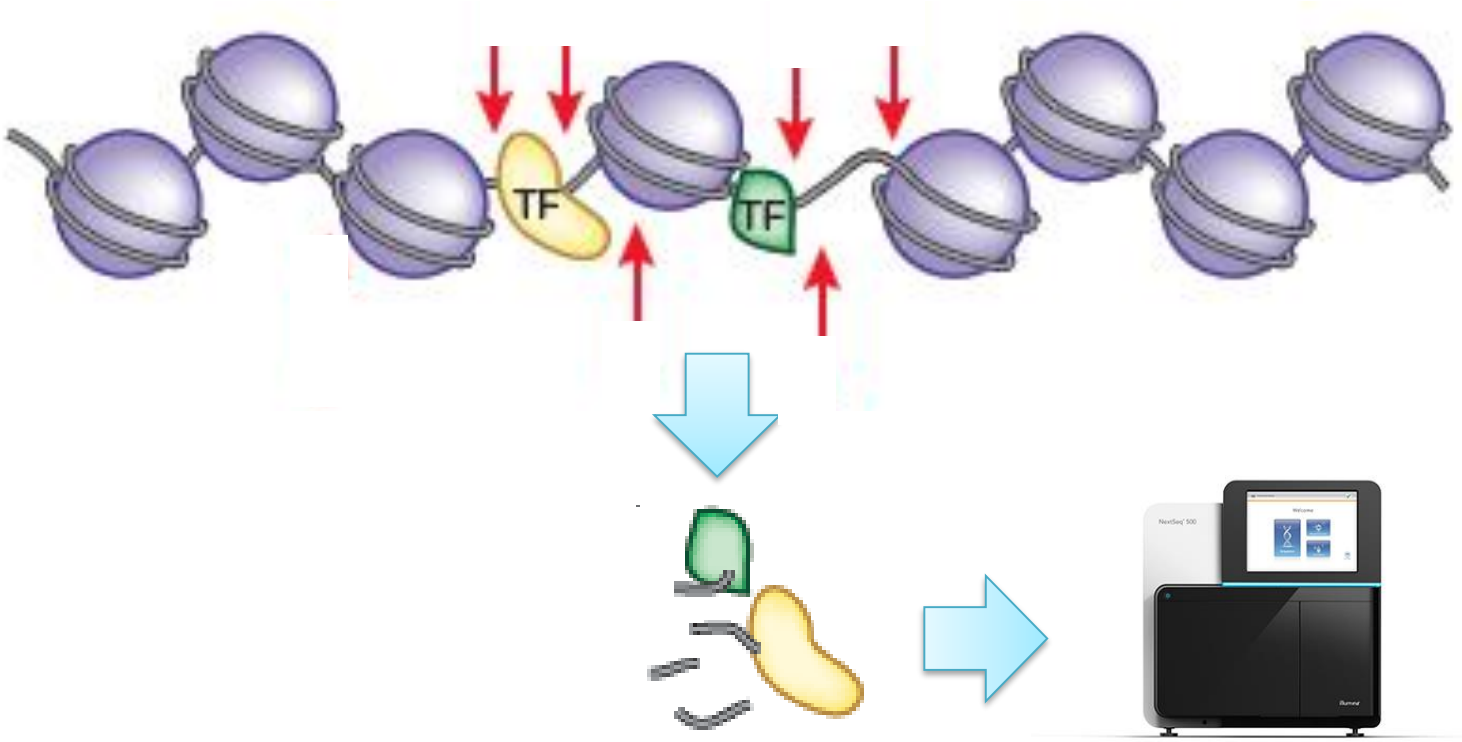
Measuring chromatin accessibility with DNase-seq



sequence
DNA fragments

TTGTTAAGGTTTATGTTCGTTTAATGGTG
TTTCTTACGACTGTACGATCAAAACGGGG
AGGCGGCAAGCAGGTGCAGCGTTTTTATA
GGGCTACAACACGTTGGTGCACCCAACAC

Measuring chromatin accessibility with DNase-seq

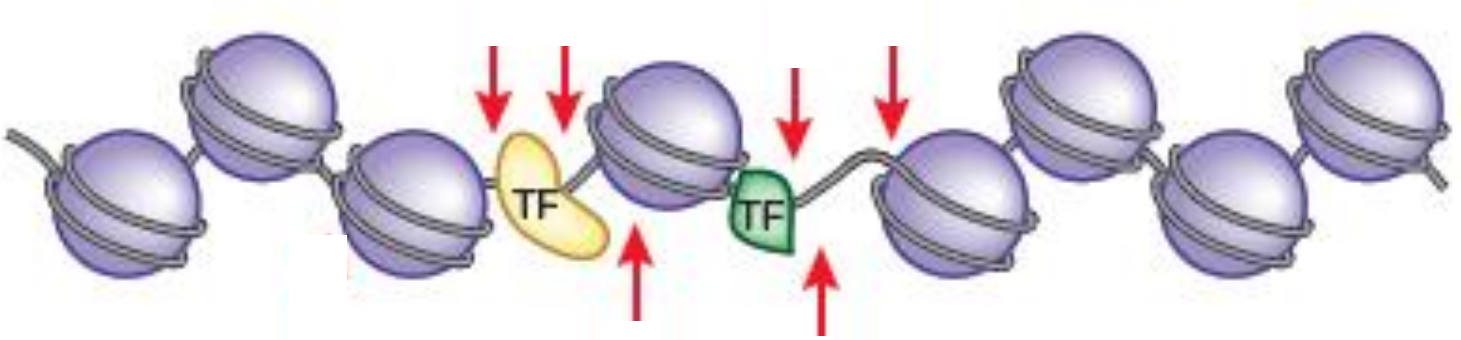


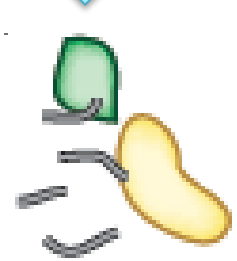
map to genome

```
                GTTCGTTTAATGGTGGCCGGAGGG
           GTTTATGTTCGTTTAATGGTGCCG
           GTTTATGTTCGTTTAATGGTGCCG
      AAGGTTTATGTTCGTTTAATGGTG
 TTGTTAAGGTTTATGTTCGTTTAA
...ACCATTGTTAAGGTTTATGTTCGTTTAATGGTGCCGGGAGGGATTTA...
```

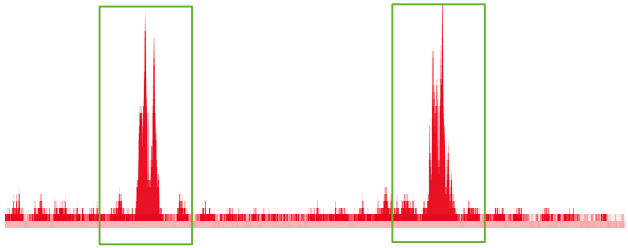
```
TTGTTAAGGTTTATGTTCGTTTAATGGTG
TTTCTTACGACTGTACGATCAAAACGGGG
AGGCGGCAAGCAGGTGCAGCGTTTTTATA
GGGCTACAACACGTTGGTGCACCCAACAC
```

Measuring chromatin accessibility with DNase-seq





DNase hypersensitive sites



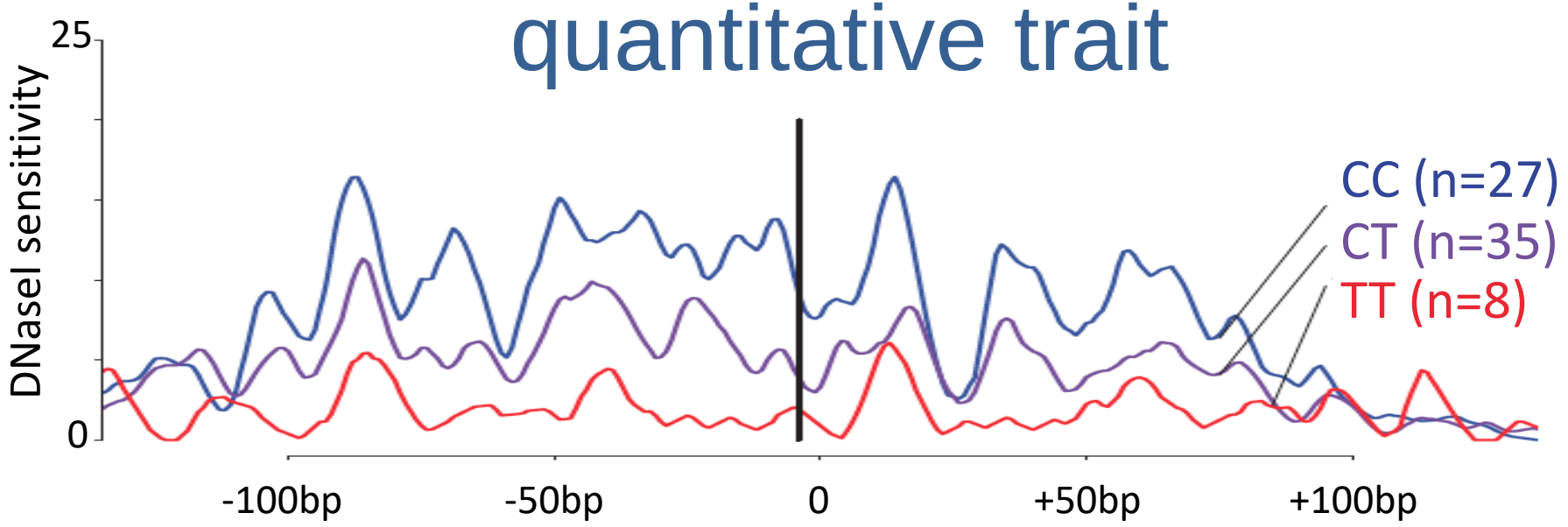
TTGTTAAGGTTTATGTTCGTTTAATGGTG

TTTCTTACGACTGTACGATCAAAACGGGG

AGGCGGCAAGCAGGTGCAGCGTTTTTATA

GGGCTACAACACGTTGGTGCACCCAACAC

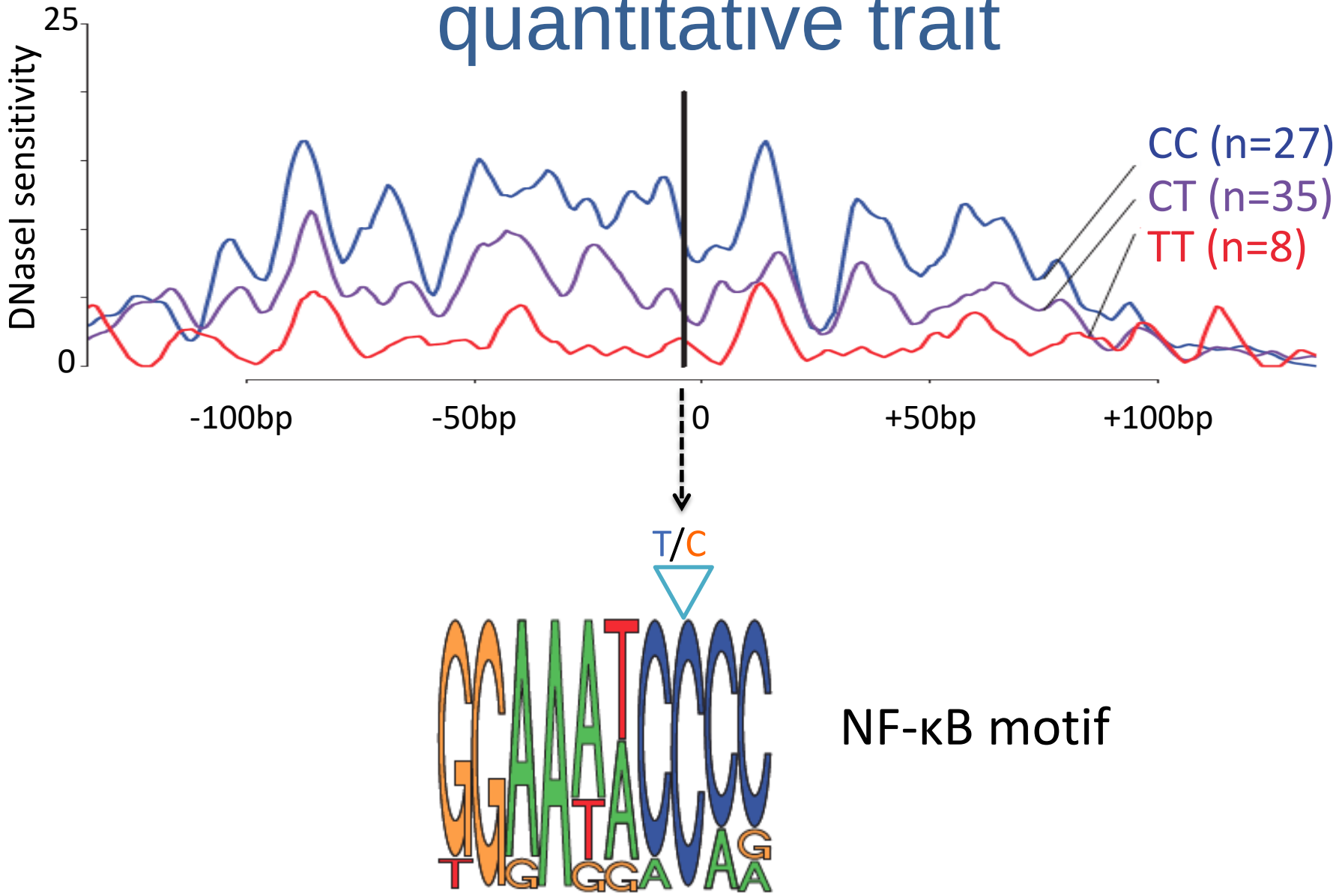
DNase sensitivity as a quantitative trait



distance from center of DNase peak

Degner et al. 2012

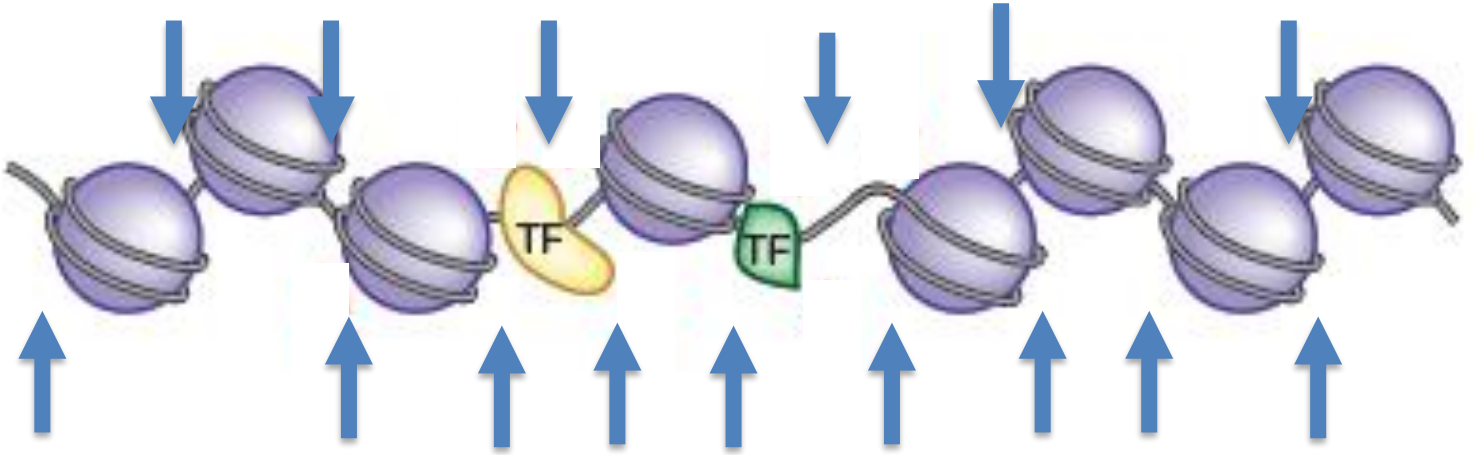
DNase sensitivity as a quantitative trait



NF-κB motif

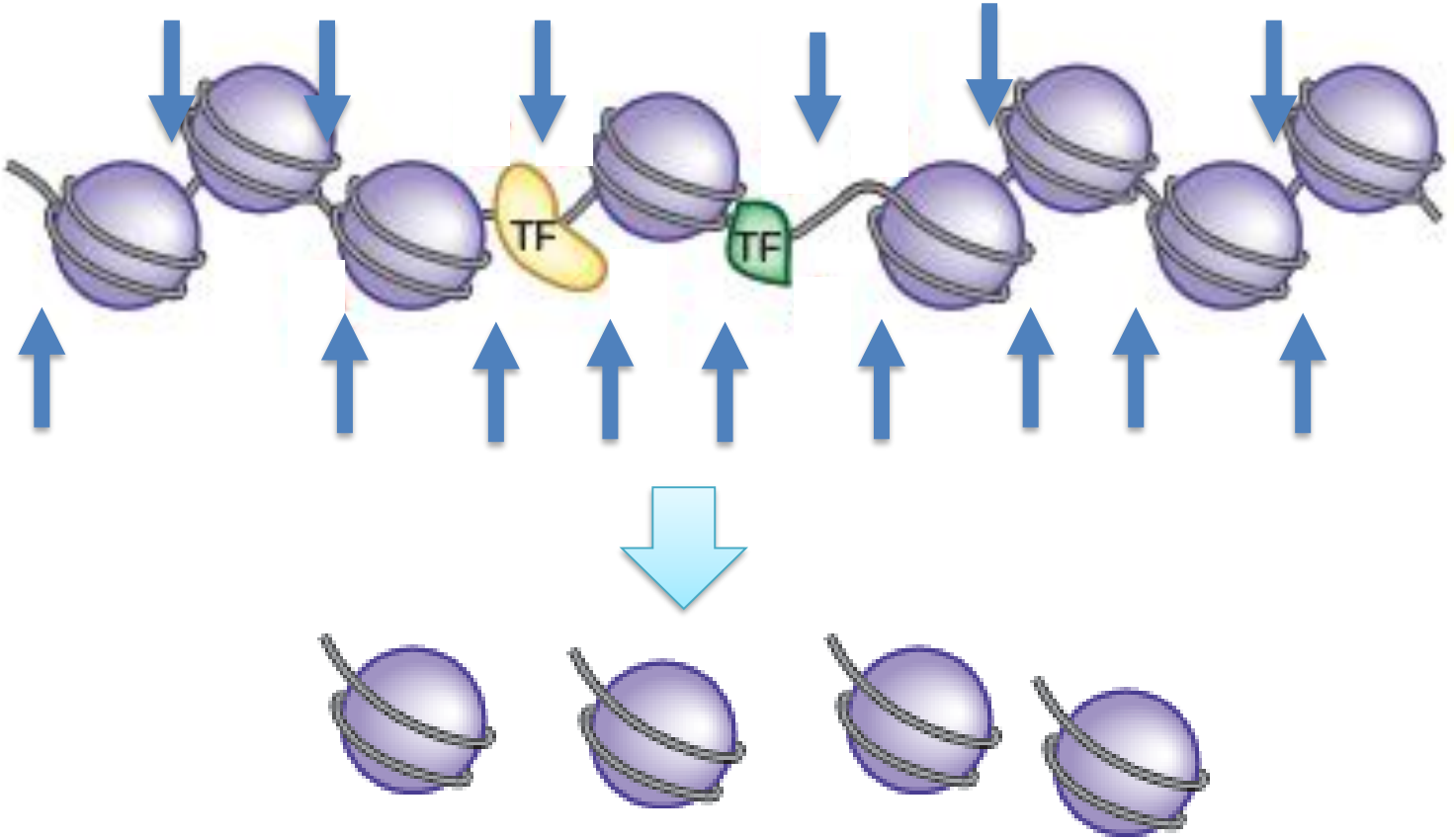
Degner et al. 2012

Determining nucleosome positions with MNase-seq



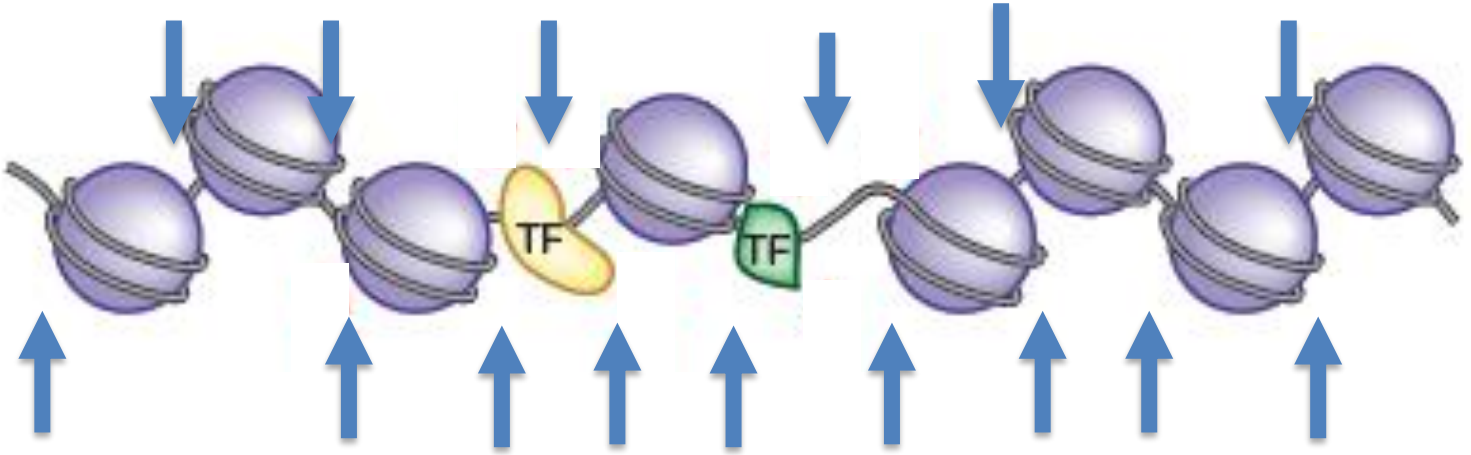
digest chromatin with micrococcal nuclease (MNase)

Determining nucleosome positions with MNase-seq

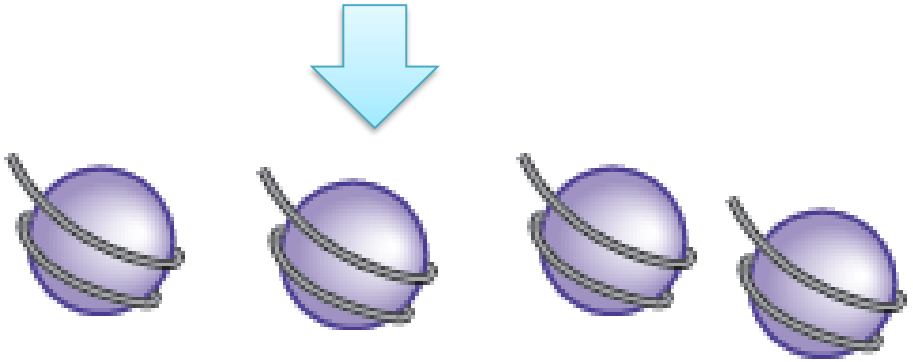


isolate nucleosome-sized fragments and sequence ends

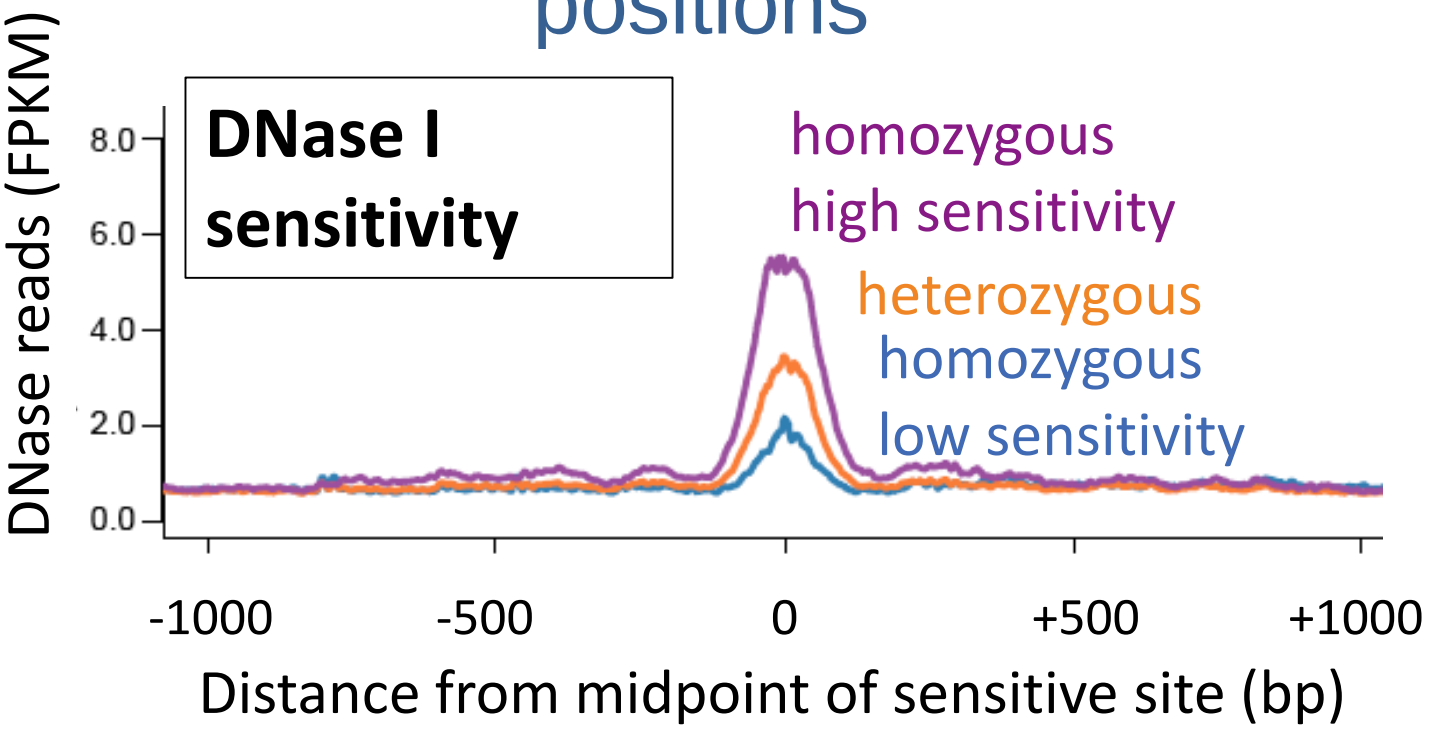
Determining nucleosome positions with MNase-seq



peaks are consistently-positioned nucleosomes

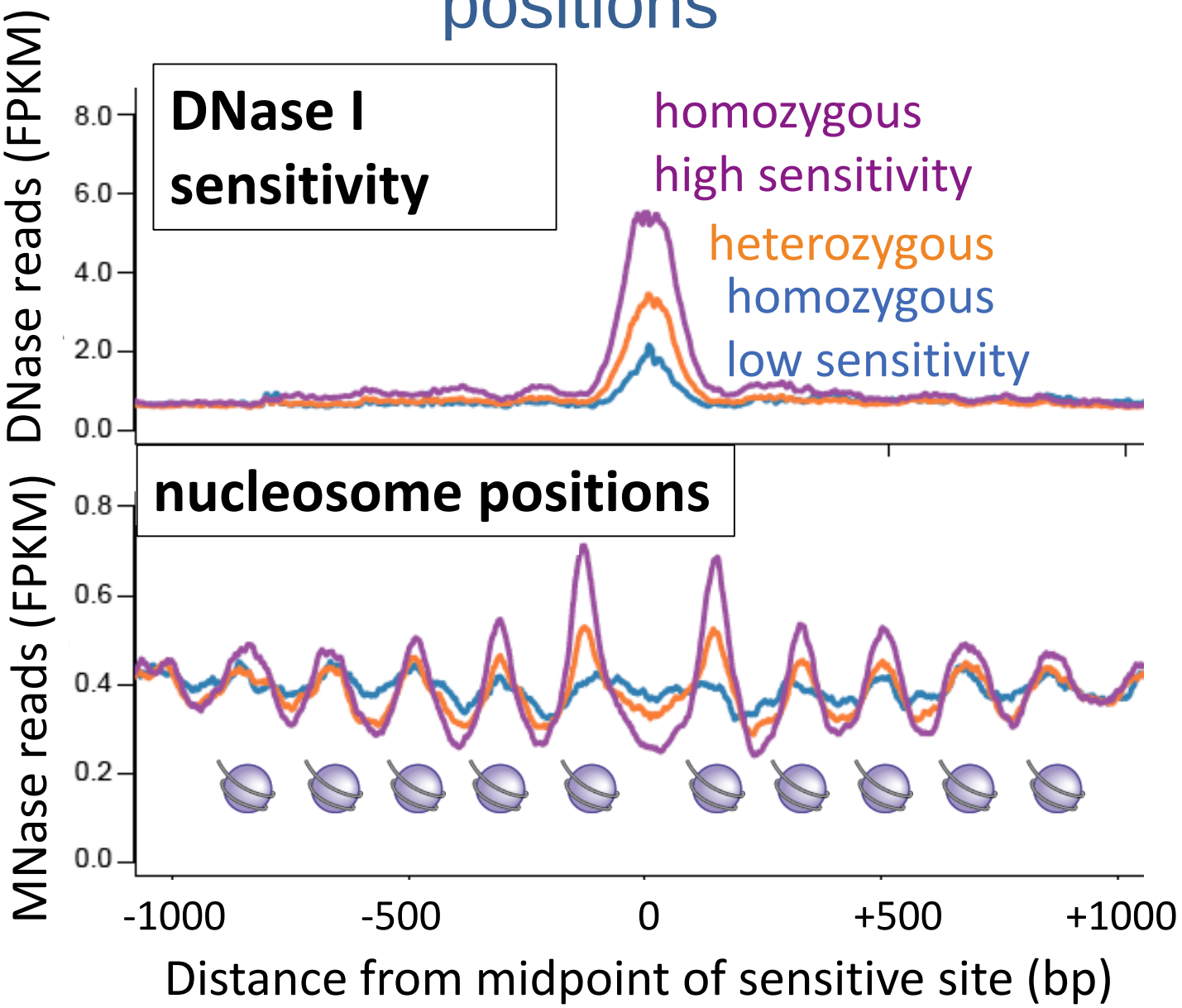


dsQTLs are associated with nucleosome positions



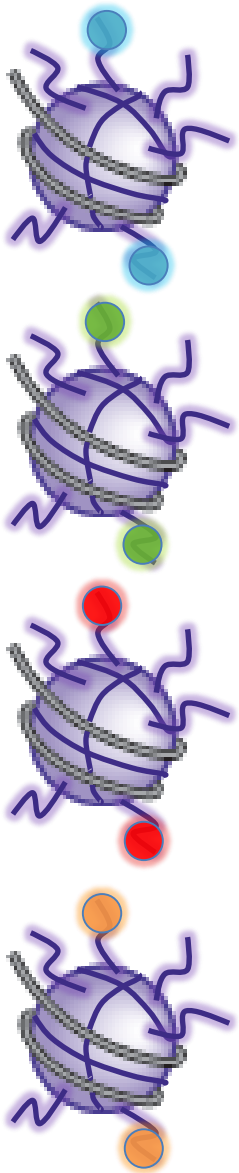
Gaffney*, McVicker* et al. 2012

dsQTLs are associated with nucleosome positions



Gaffney*, McVicker* et al. 2012

Histone modifications



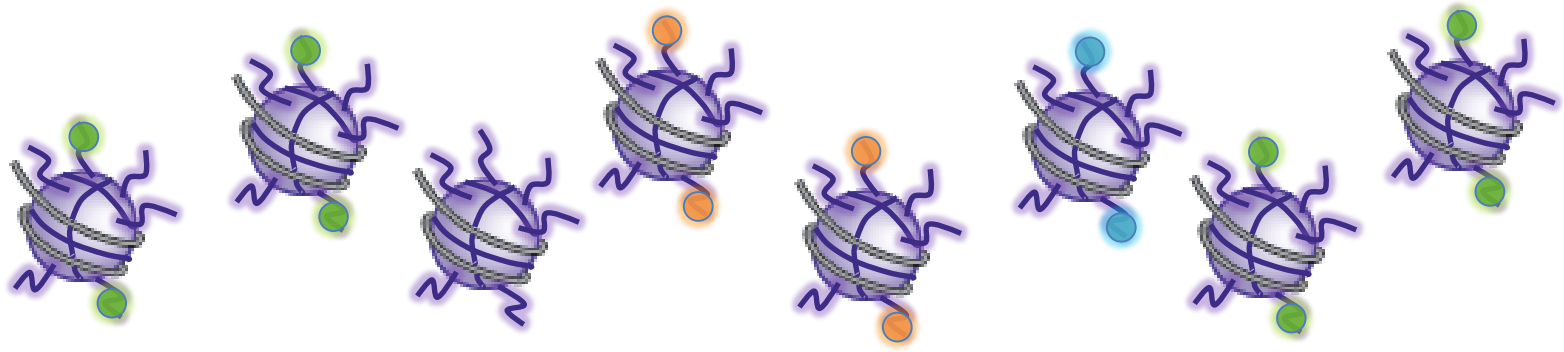
H3K4me1: active/open chromatin outside of promoters

H3K4me3: active promoters

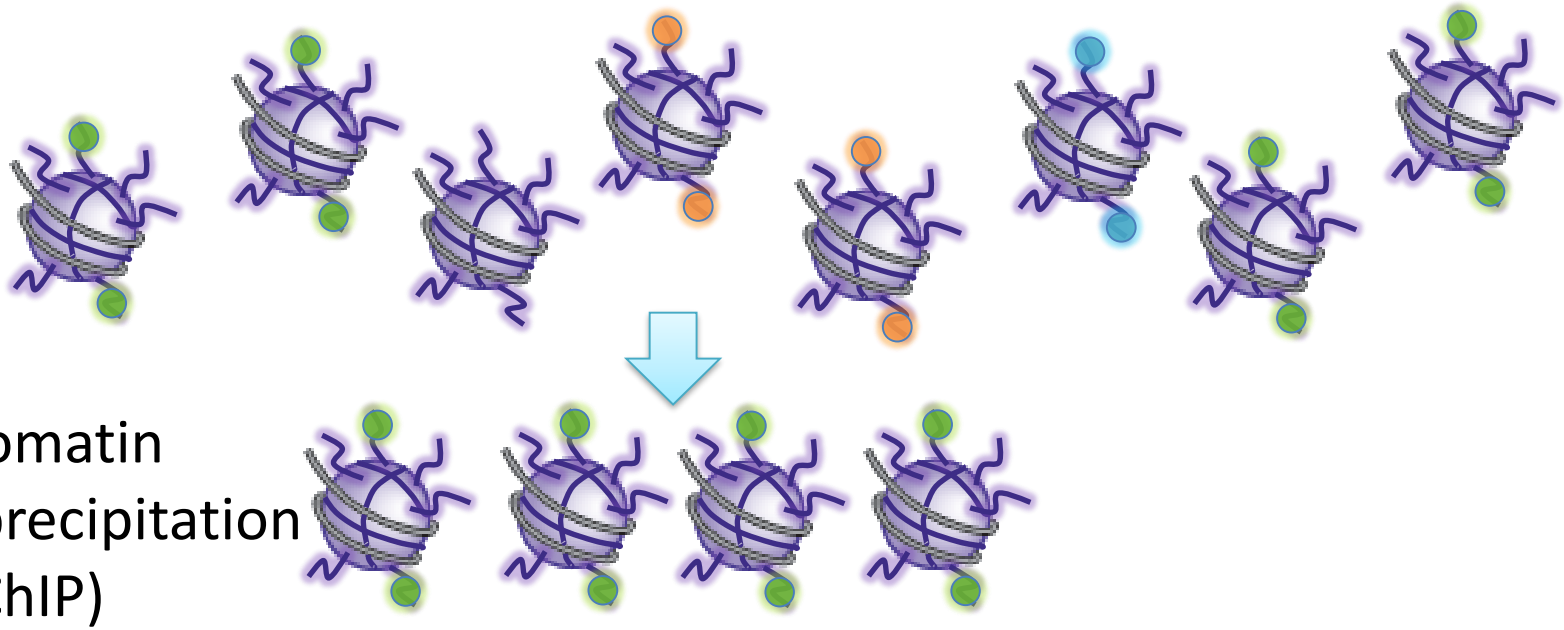
H3K27ac: active promoters & enhancers

H3K27me3: silenced genes

Measuring histone marks with ChIP-seq



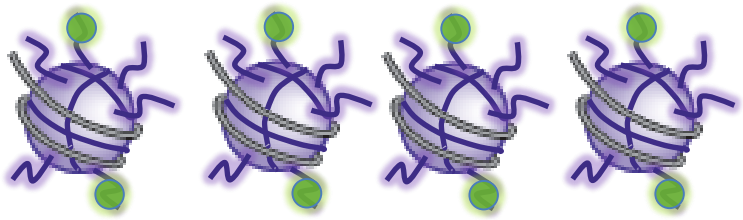
Measuring histone marks with ChIP-seq



Chromatin Immunoprecipitation (ChIP)

Measuring histone marks with ChIP-seq

High-throughput
DNA sequencing



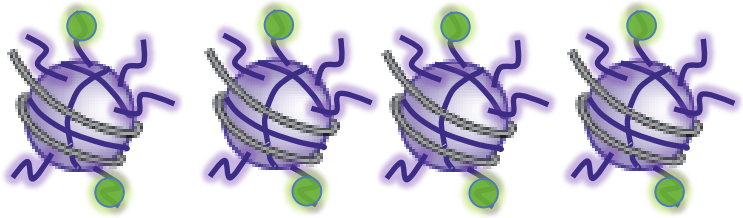
TTGTTAAGGTTTATGTTCGTTTAATGGTG

TTTCTTACGACTGTACGATCAAAACGGGG

AGGCGGCAAGCAGGTGCAGCGTTTTTATA

GGGCTACAACACGTTGGTGCACCCAACAC

Measuring histone marks with ChIP-seq



TTGTTAAGGTTTATGTTCGTTTAATGGTG

TTTCTTACGACTGTACGATCAAAACGGGG

AGGCGGCAAGCAGGTGCAGCGTTTTTATA

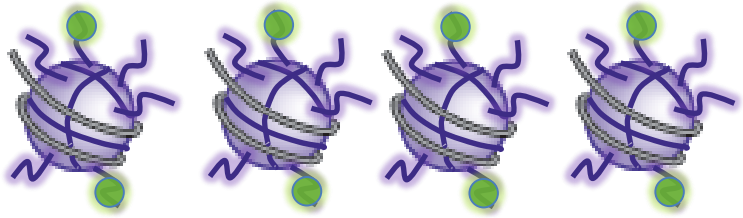
GGGCTACAACACGTTGGTGCACCCAACAC

Map reads to genome

GTTCGTTTAATGGTGGCCGGAGGG
GTTTATGTTCGTTTAATGGTGCCG
GTTTATGTTCGTTTAATGGTGCCG
AAGGTTTATGTTCGTTTAATGGTG
TTGTTAAGGTTTATGTTCGTTTAA

...ACCATTGTTAAGGTTTATGTTCGTTTAATGGTGCCGGGAGGGATTTA...

Measuring histone marks with ChIP-seq



TTGTTAAGGTTTATGTTCGTTTAATGGTG

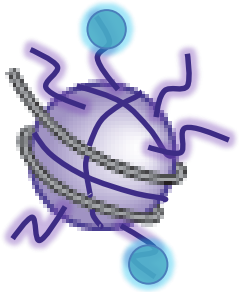
TTTCTTACGACTGTACGATCAAAACGGGG

AGGCGGCAAGCAGGTGCAGCGTTTTTATA

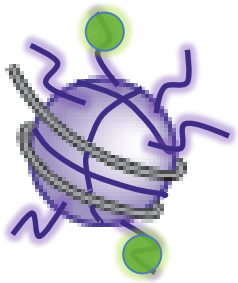
GGGCTACAACACGTTGGTGCACCCAACAC

Identify ChIP-seq peaks

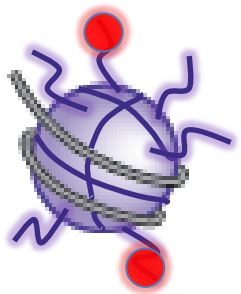
ChIP-seq data from 10 individuals



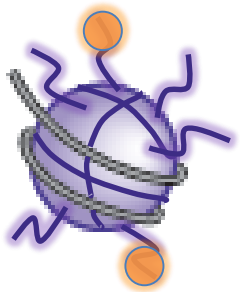
H3K4me1



H3K4me3

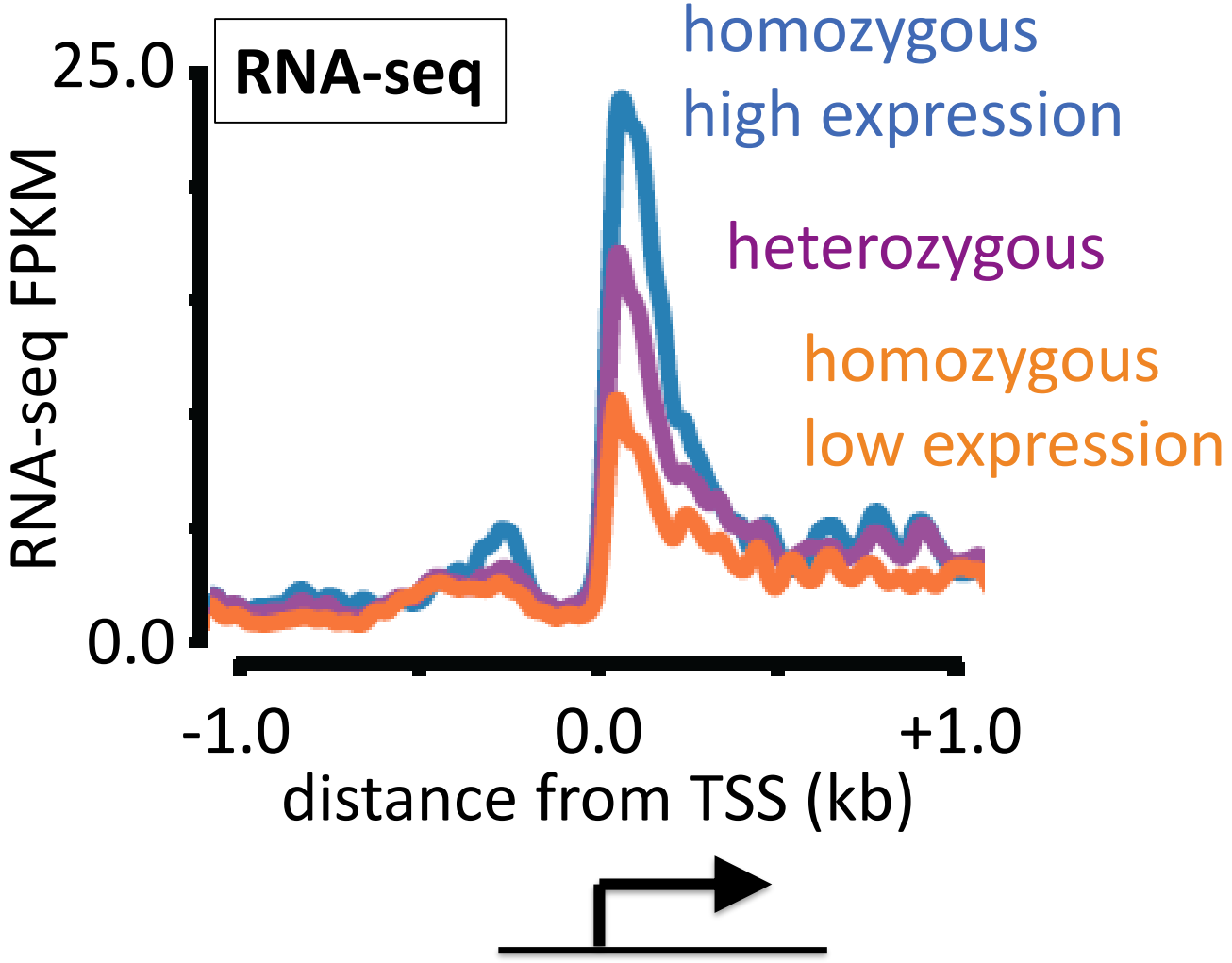


H3K27ac

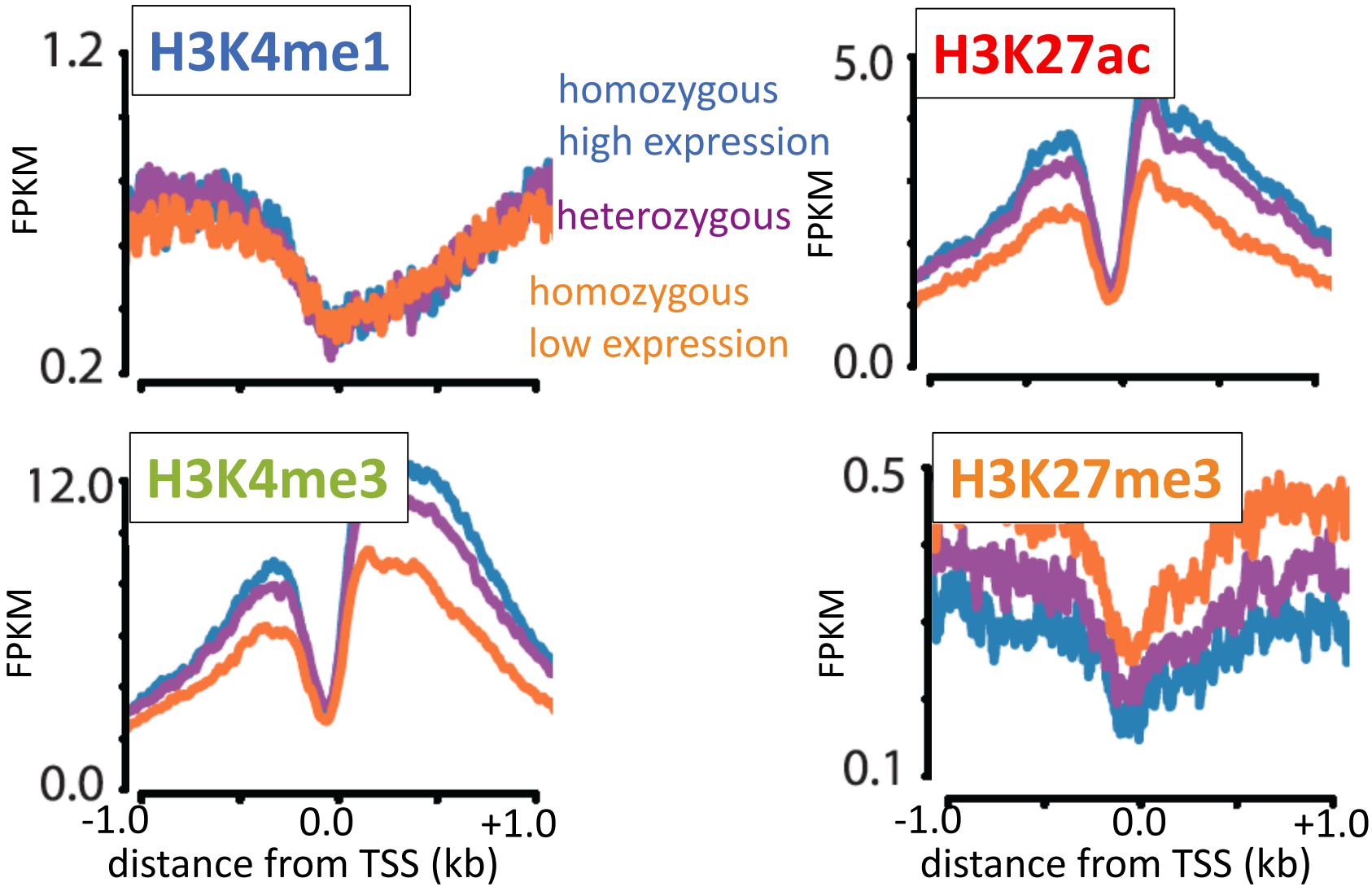


H3K27me3

Are eQTLs also associated with histone modifications?

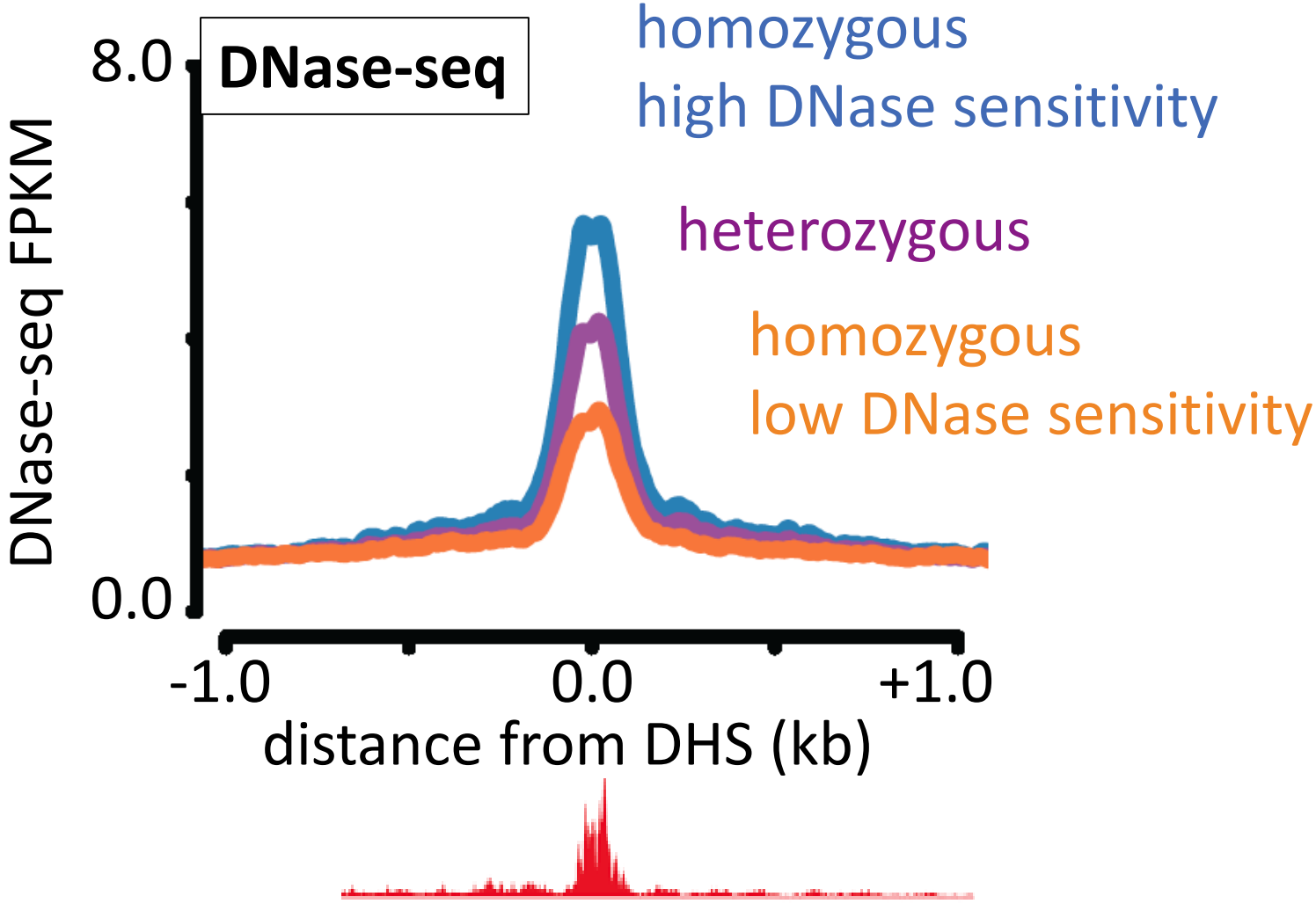


Are eQTLs also associated with histone modifications?

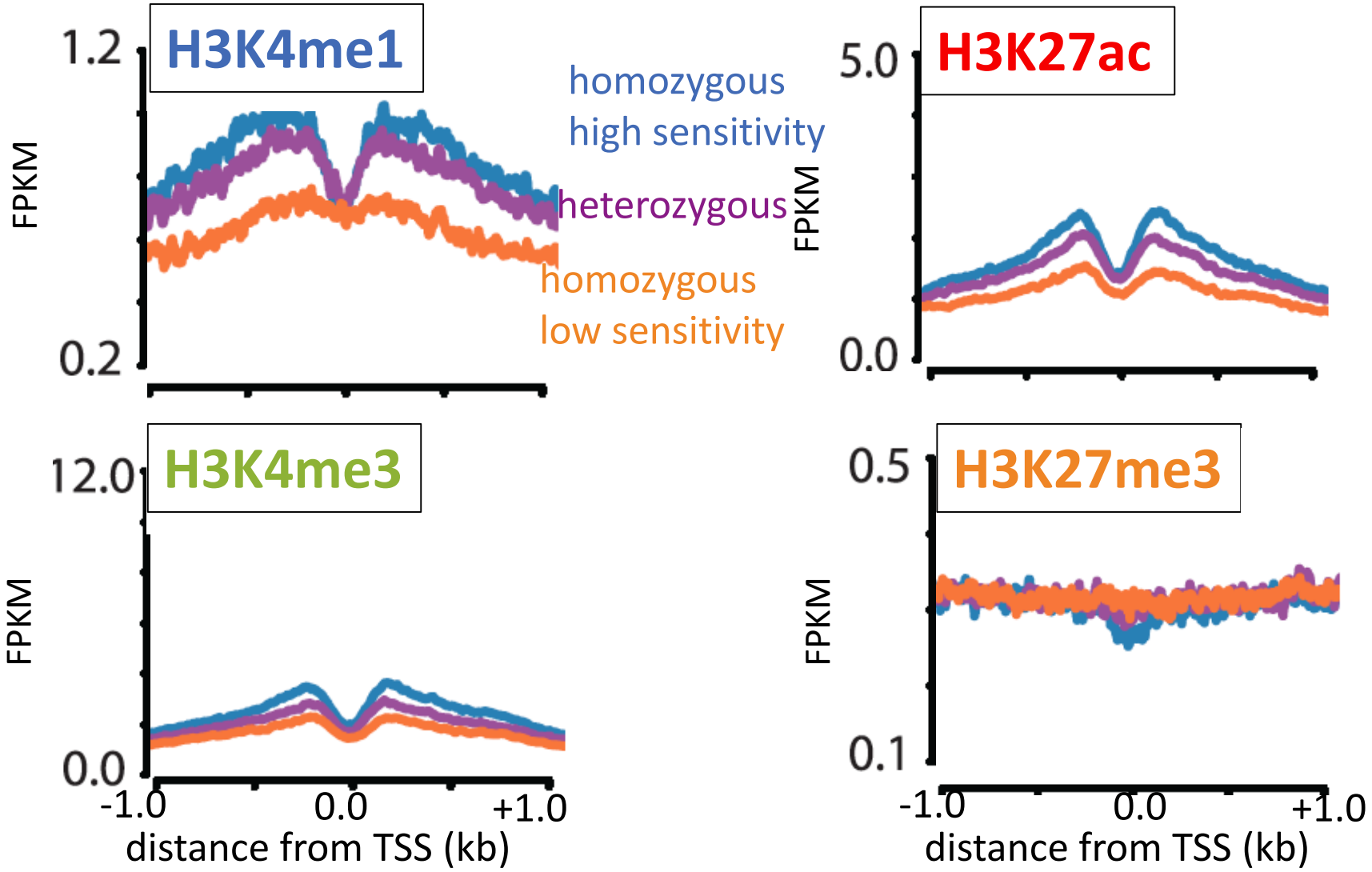


McVicker*, van de Geijn* et al. 2013

Are dsQTLs also associated with histone modifications?

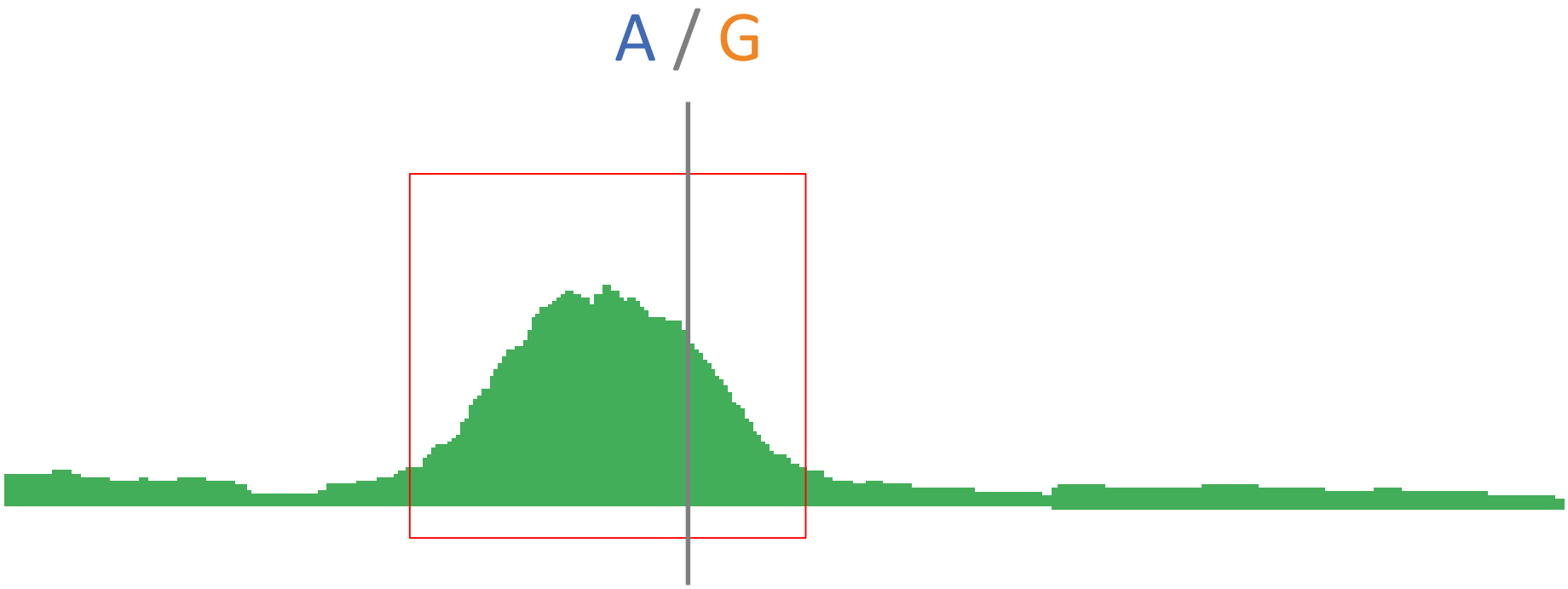


Are dsQTLs also associated with histone modifications?

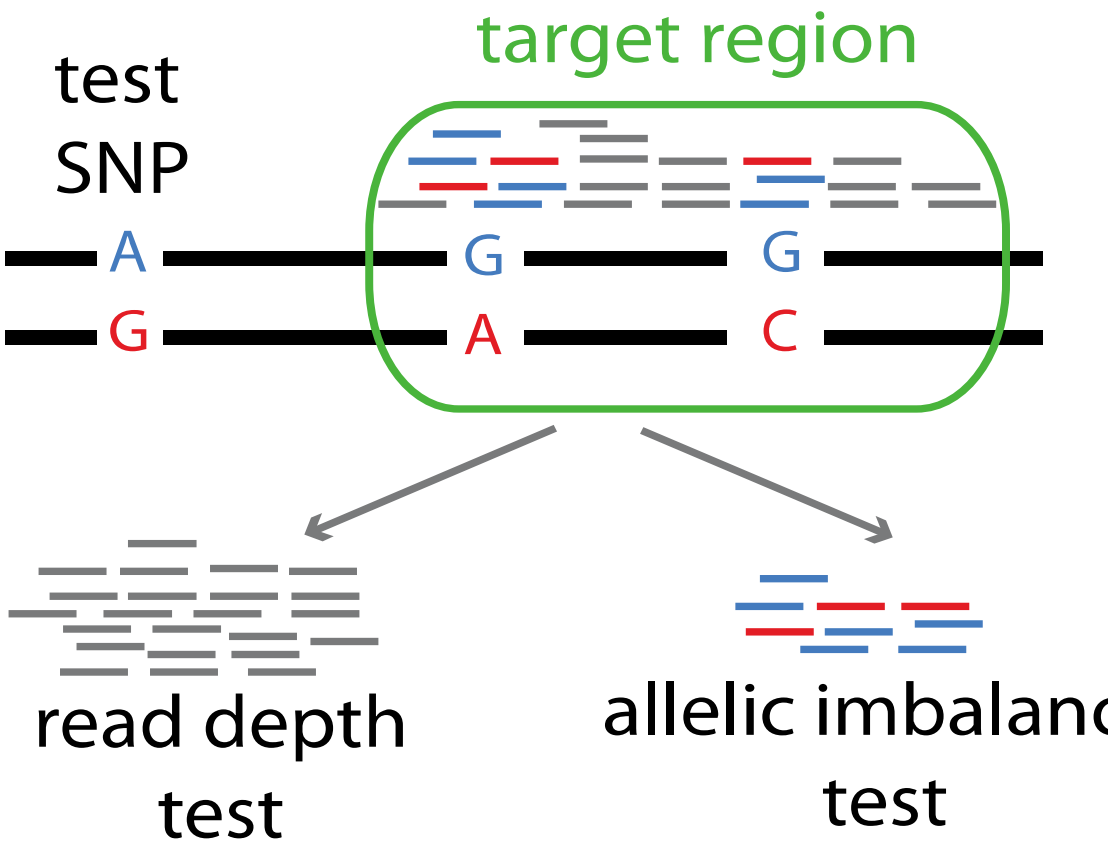


McVicker*, van de Geijn* et al. 2013

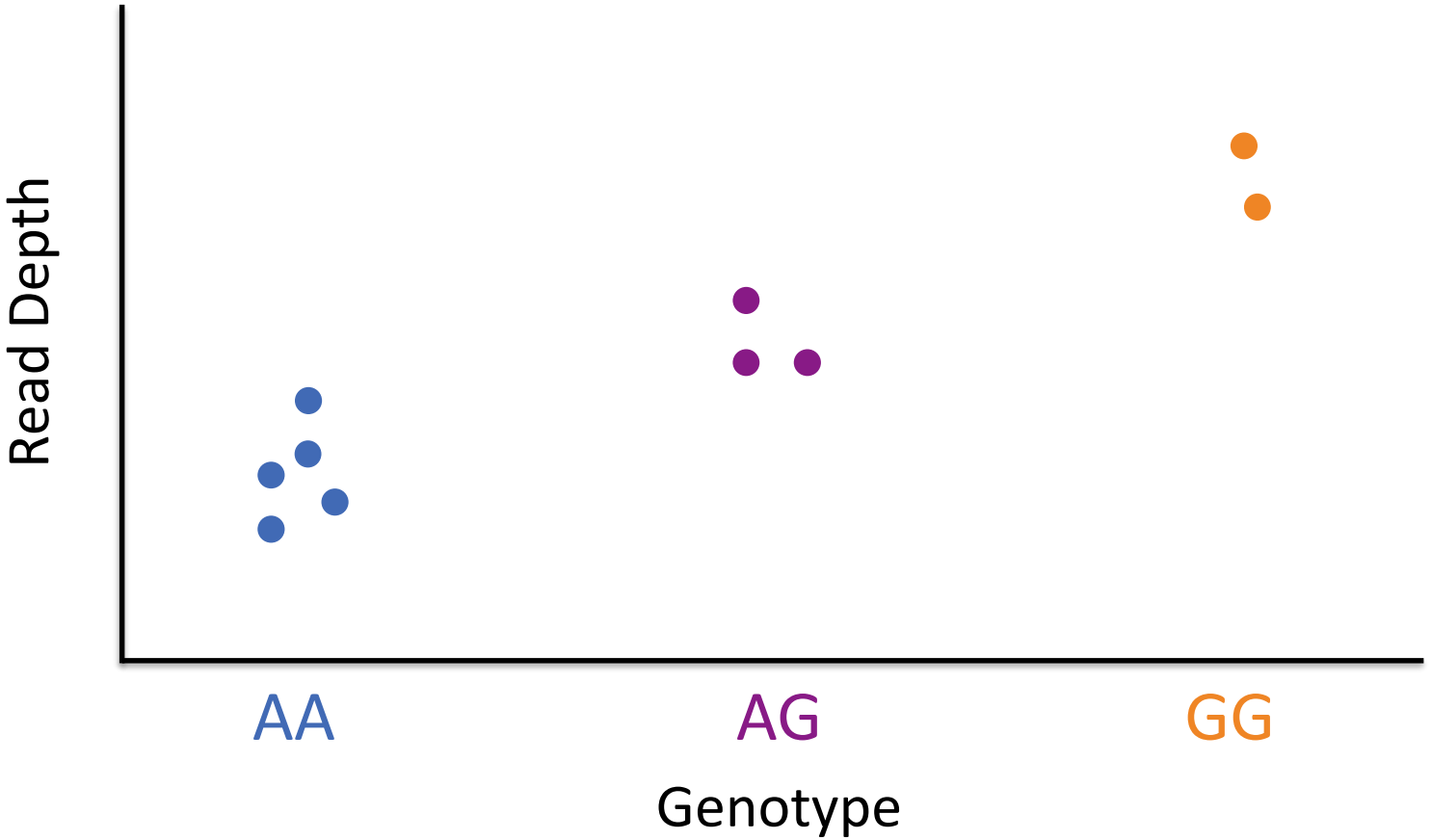
Can we identify Histone Mark QTLs?



Combined Haplotype Test



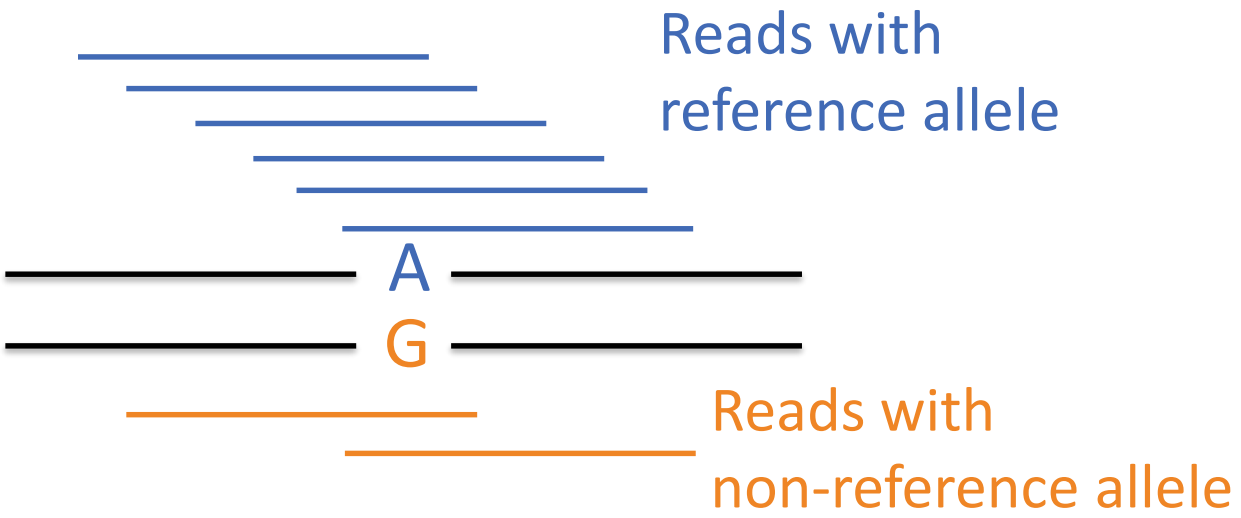
Read Depth Association Test



Model read counts with Poisson distribution:

$\lambda = 2\alpha \quad \alpha+\beta \quad 2\beta$

Allelic Imbalance Test



Model reference proportion with binomial distribution:

$$p = \frac{\alpha}{\alpha+\beta}$$

Haplotype Imbalance Test

- Phase SNPs
- Test allelic imbalance across entire haplotype

Combined Haplotype Test

Read Depth Association Test

$$\lambda = \begin{cases} 2\alpha & \text{if homozygous AA} \\ \alpha+\beta & \text{if heterozygous AB} \\ 2\beta & \text{if homozygous BB} \end{cases}$$

+

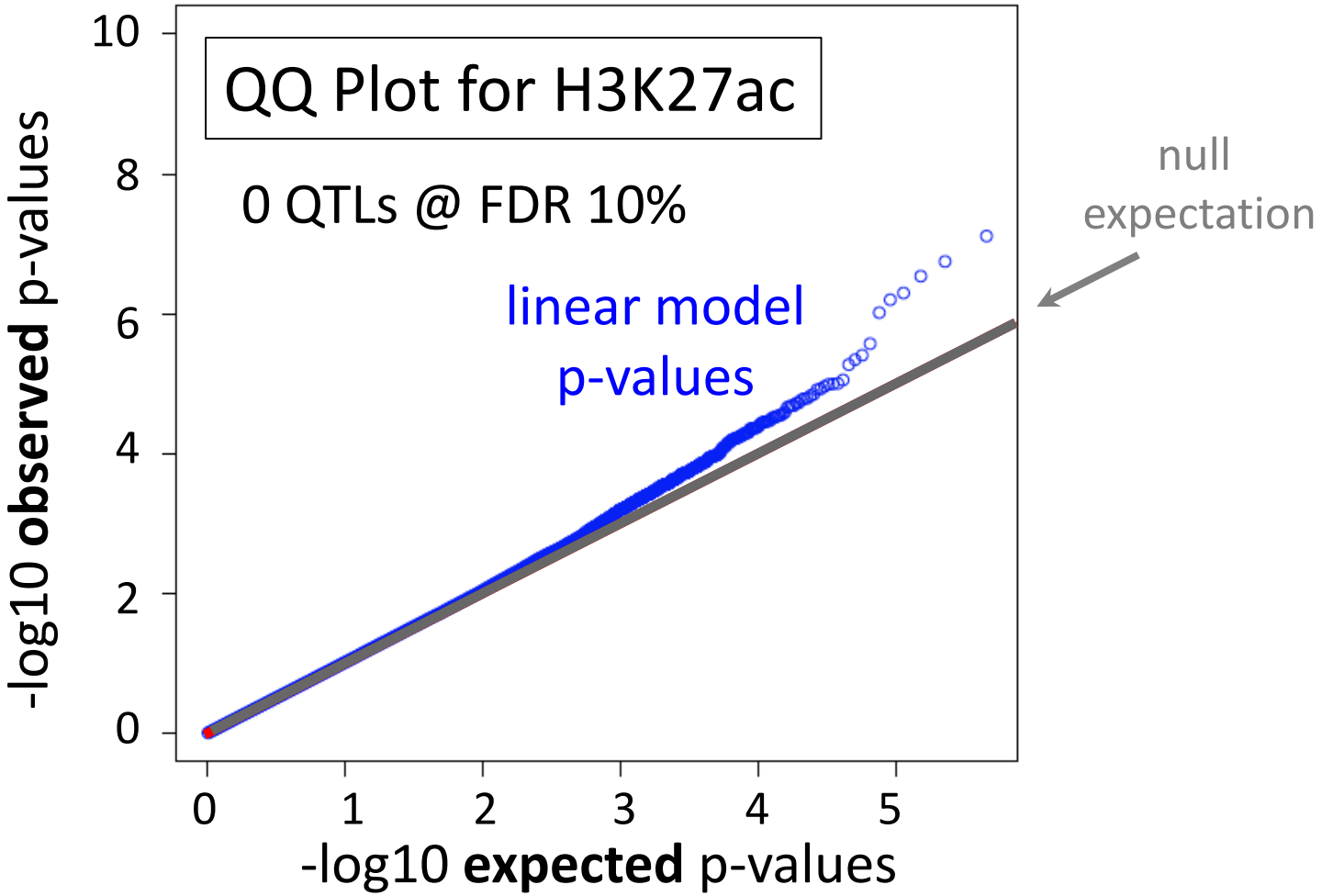
Haplotype Imbalance Test

$$p = \frac{\alpha}{\alpha+\beta}$$

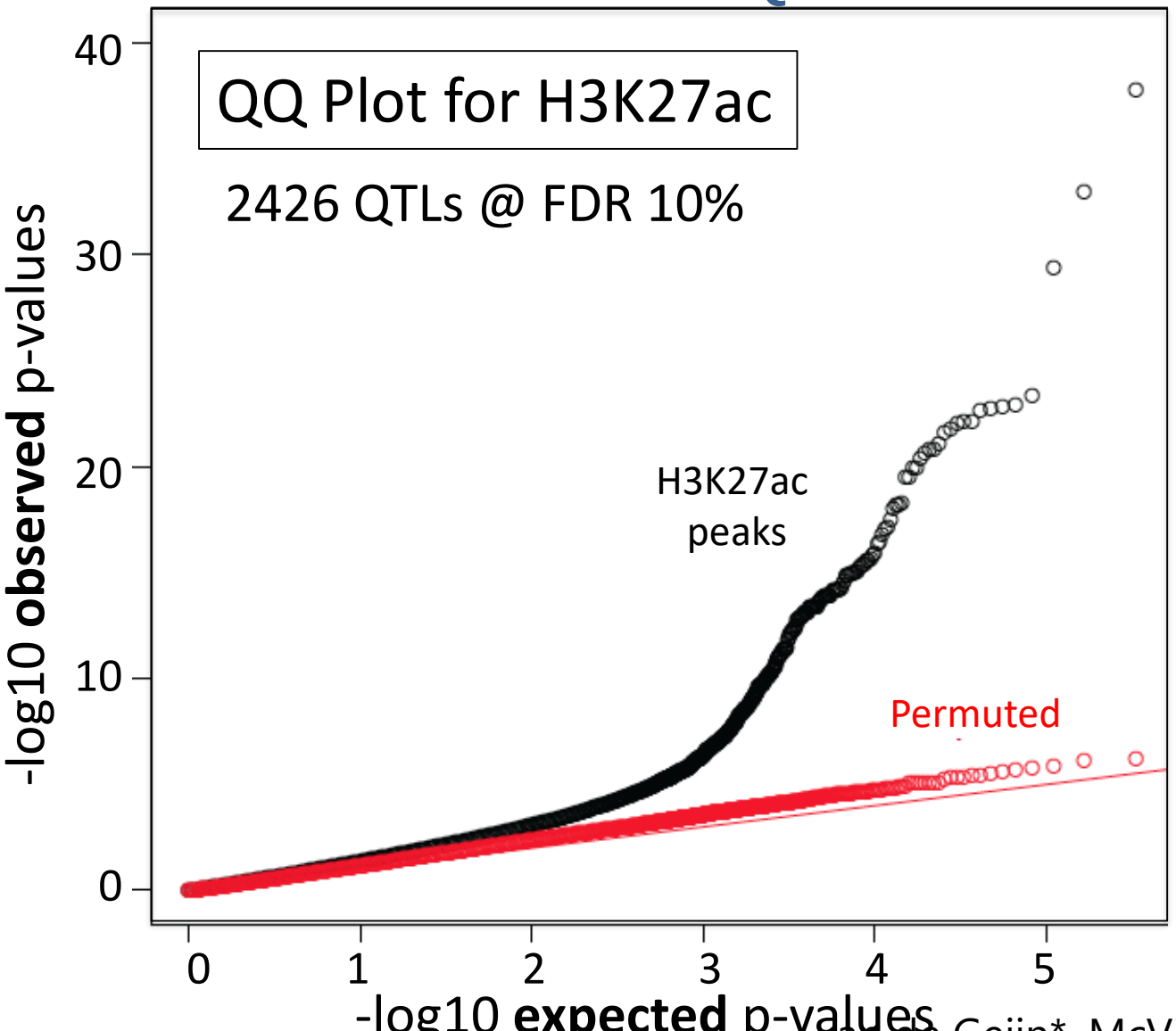
null hypothesis, $\mathbf{H_0}$: $\alpha = \beta$

alternative hypothesis, $\mathbf{H_1}$: $\alpha \neq \beta$

Standard mapping of H3K27ac QTLs with 10 Individuals

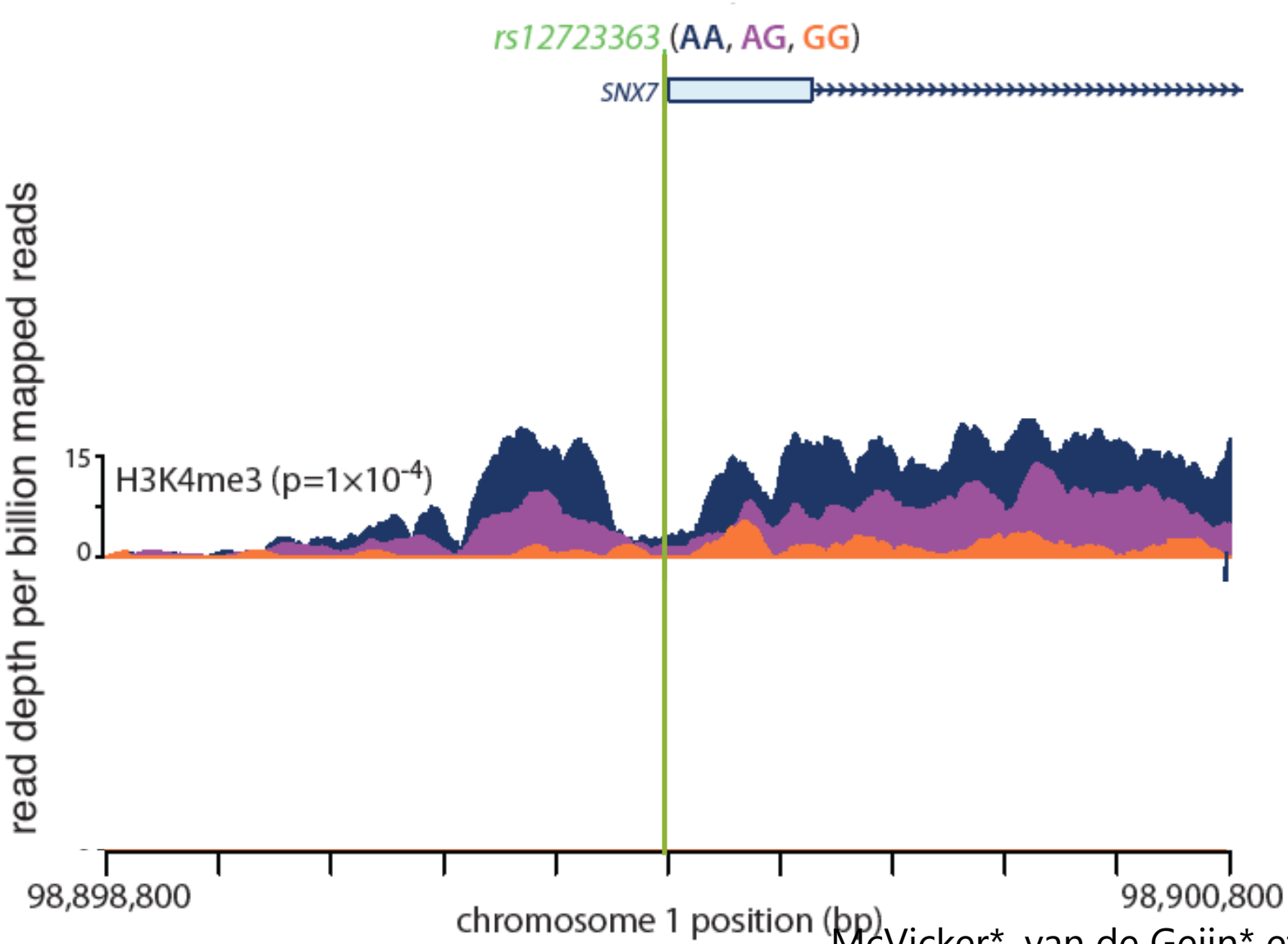


Combined Haplotype Test Mapping of H3K27ac QTLs



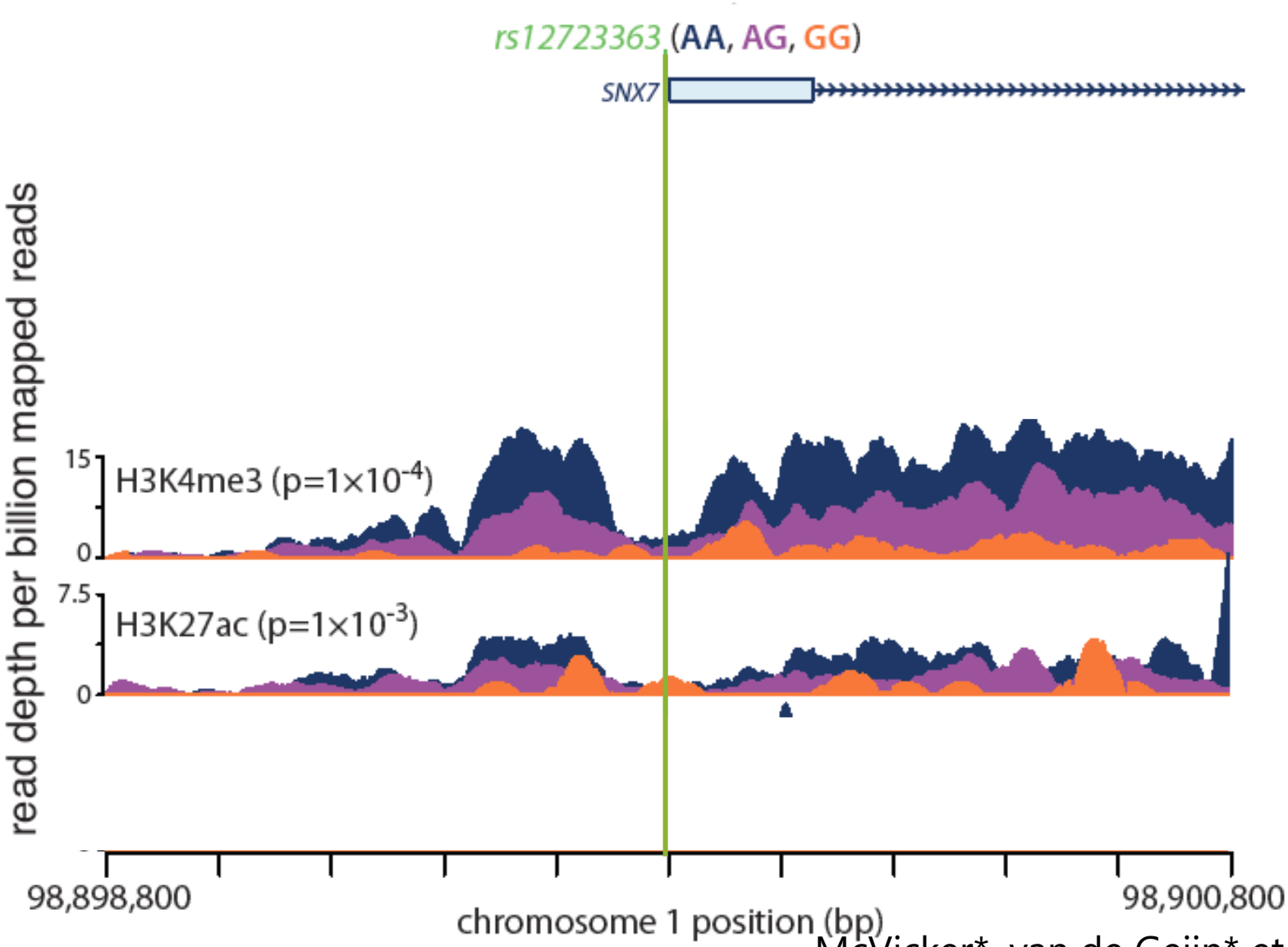
van de Geijn*, McVicker*, et al. 2015

Example of a histone mark QTL



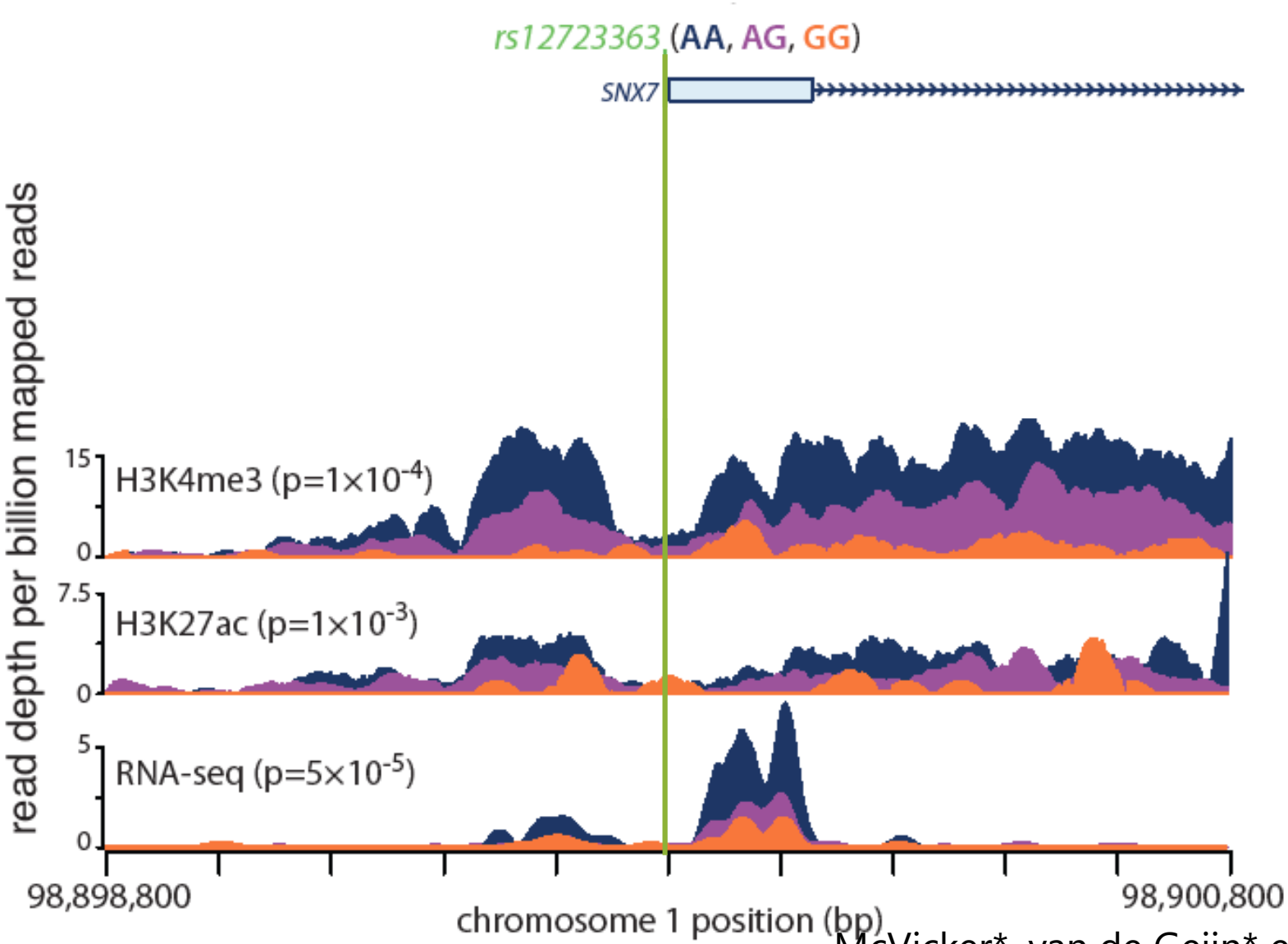
McVicker*, van de Geijn* et al. 2013

Example of a histone mark QTL



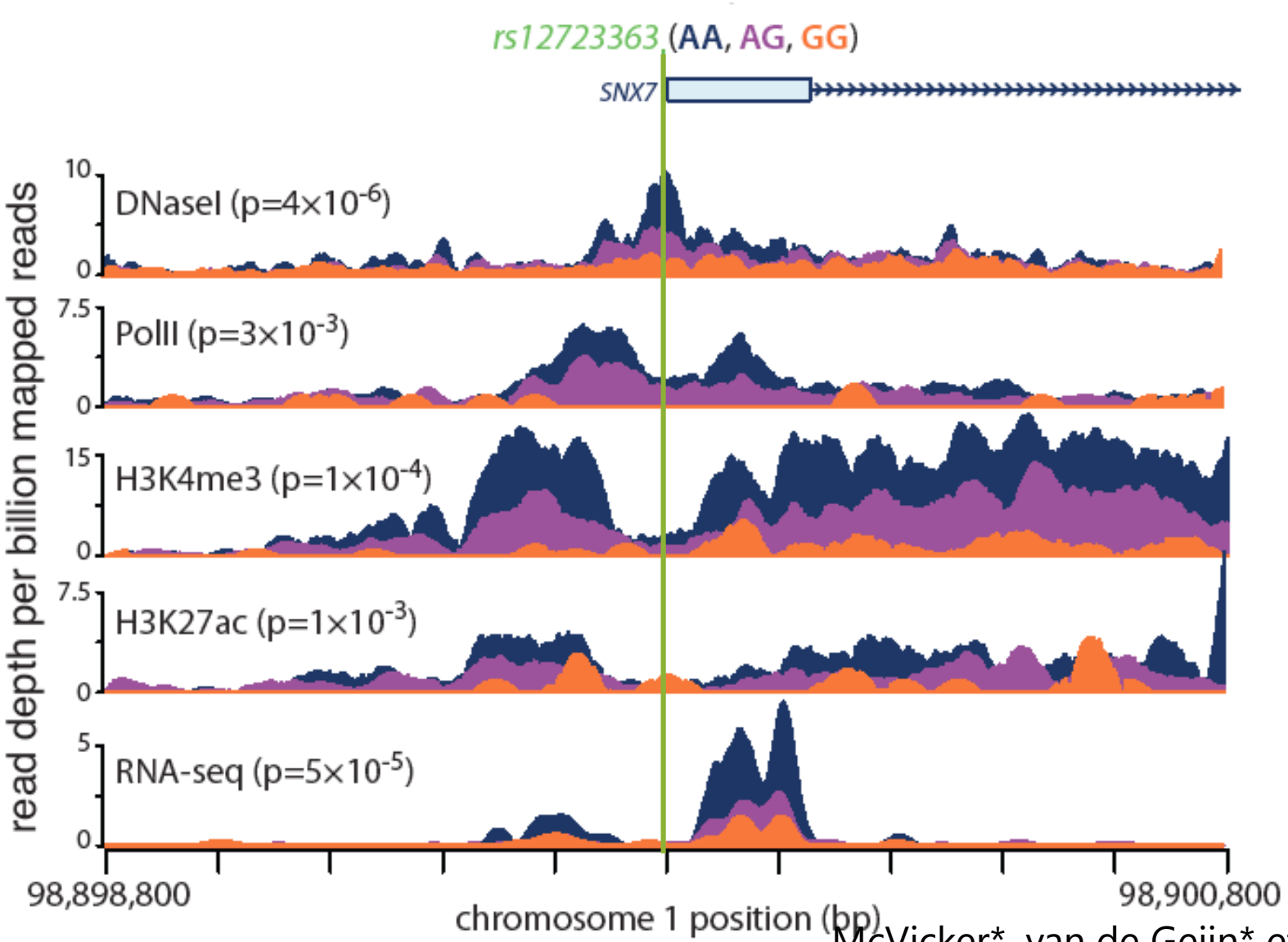
McVicker*, van de Geijn* et al. 2013

Example of a histone mark QTL



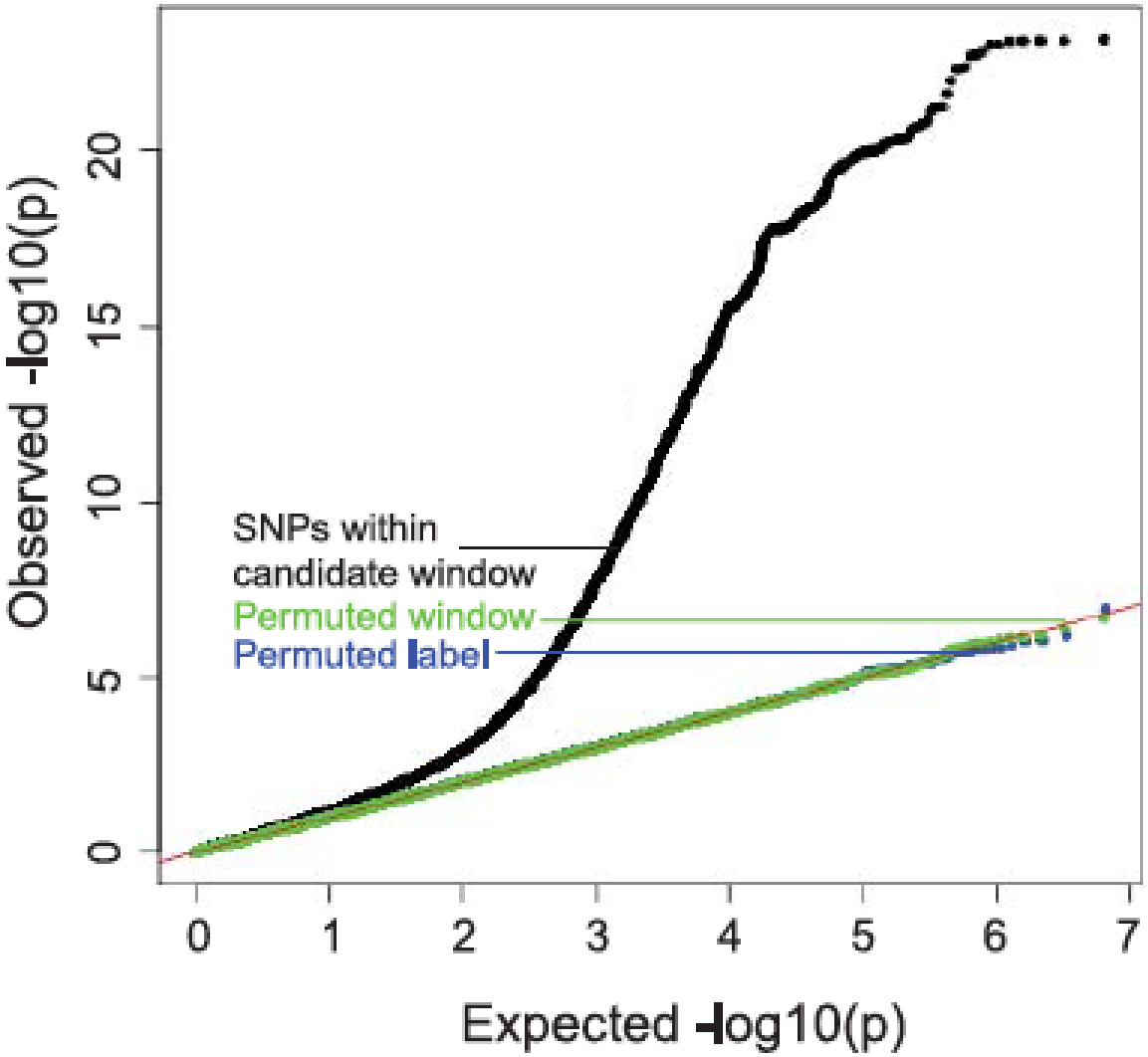
McVicker*, van de Geijn* et al. 2013

Example of a histone mark QTL



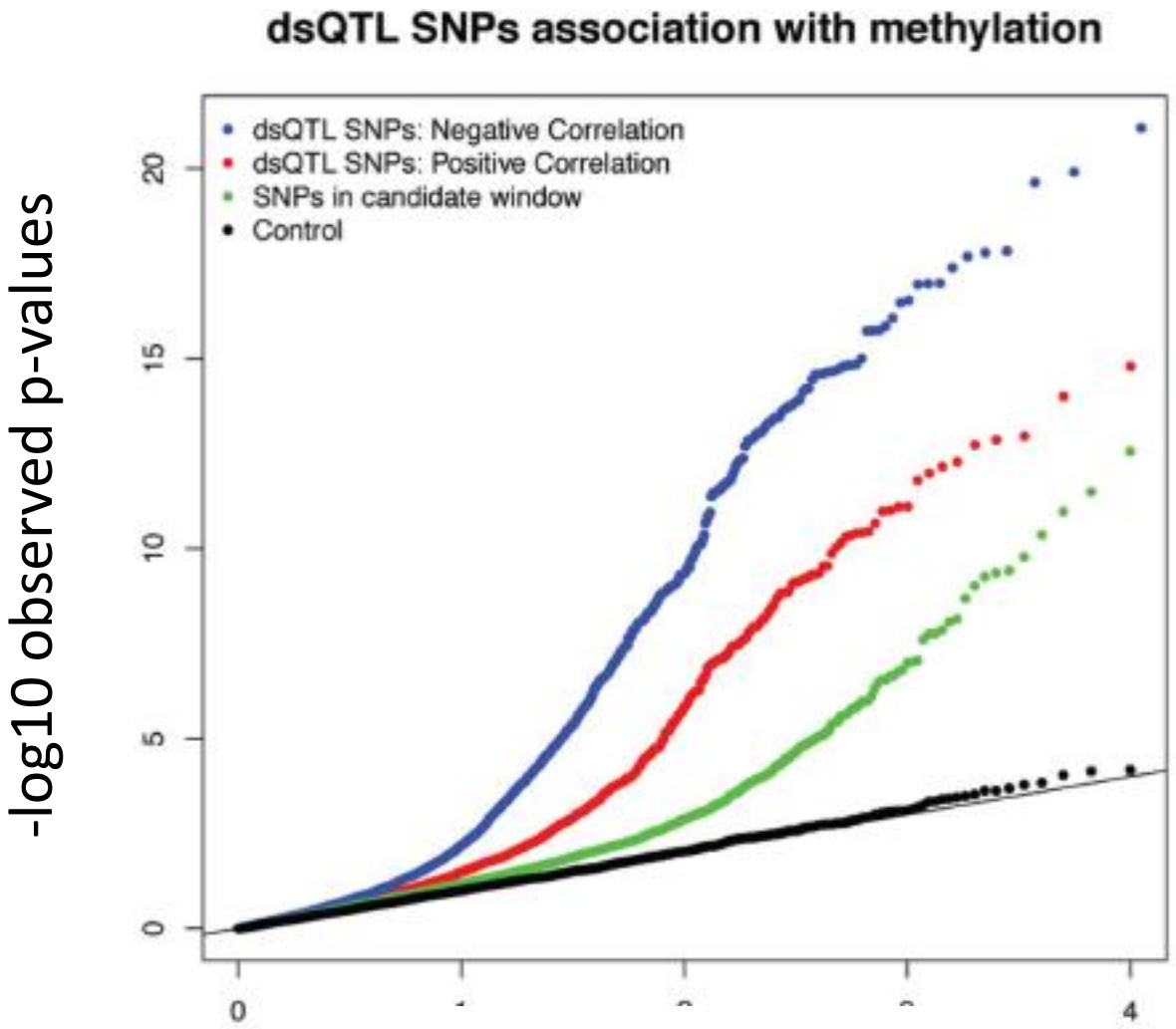
McVicker*, van de Geijn* et al. 2013

DNA methylation QTLs



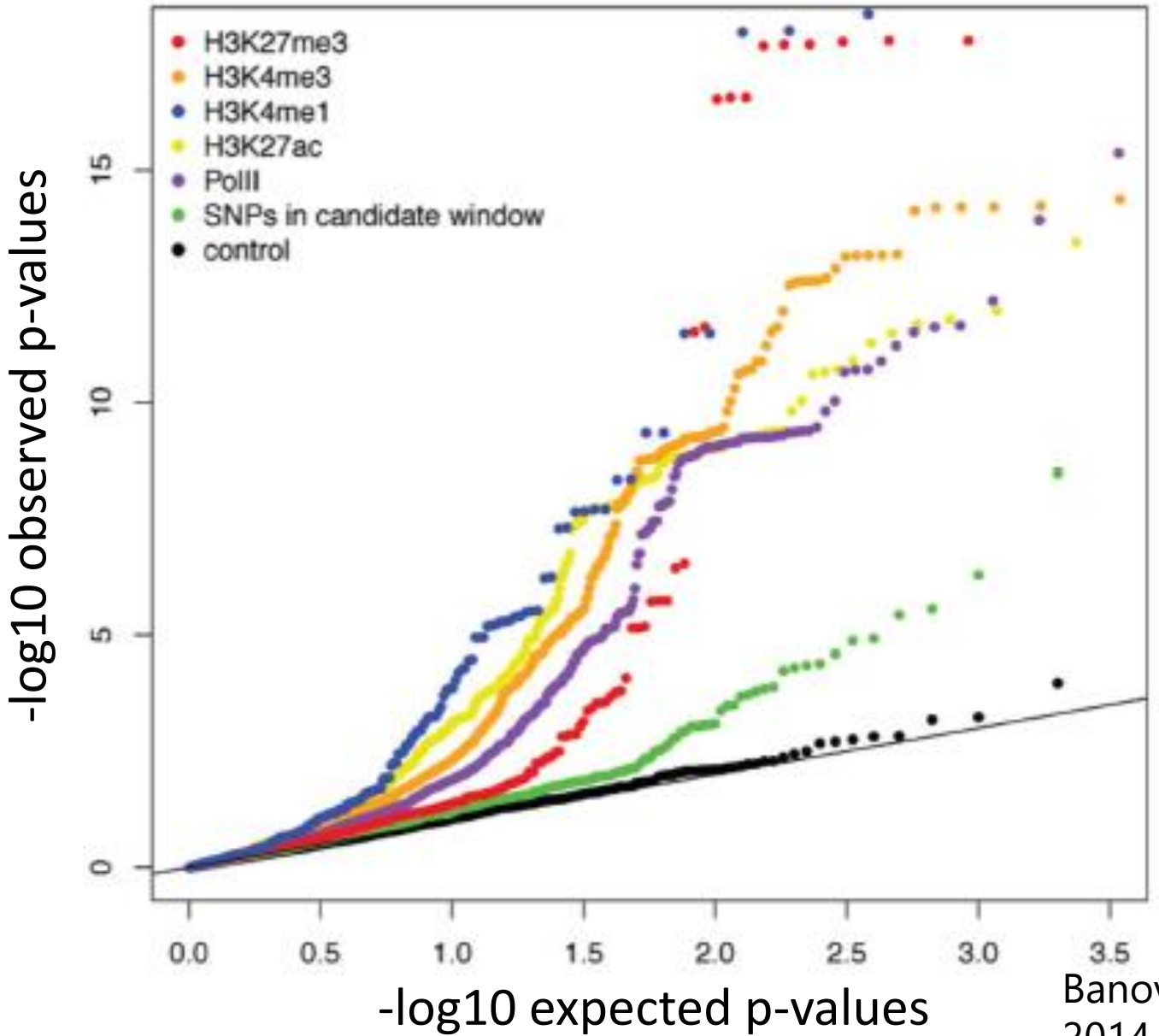
Banovich et al. 2014

DNA Methylation QTLs are often also DNase sensitivity QTLs

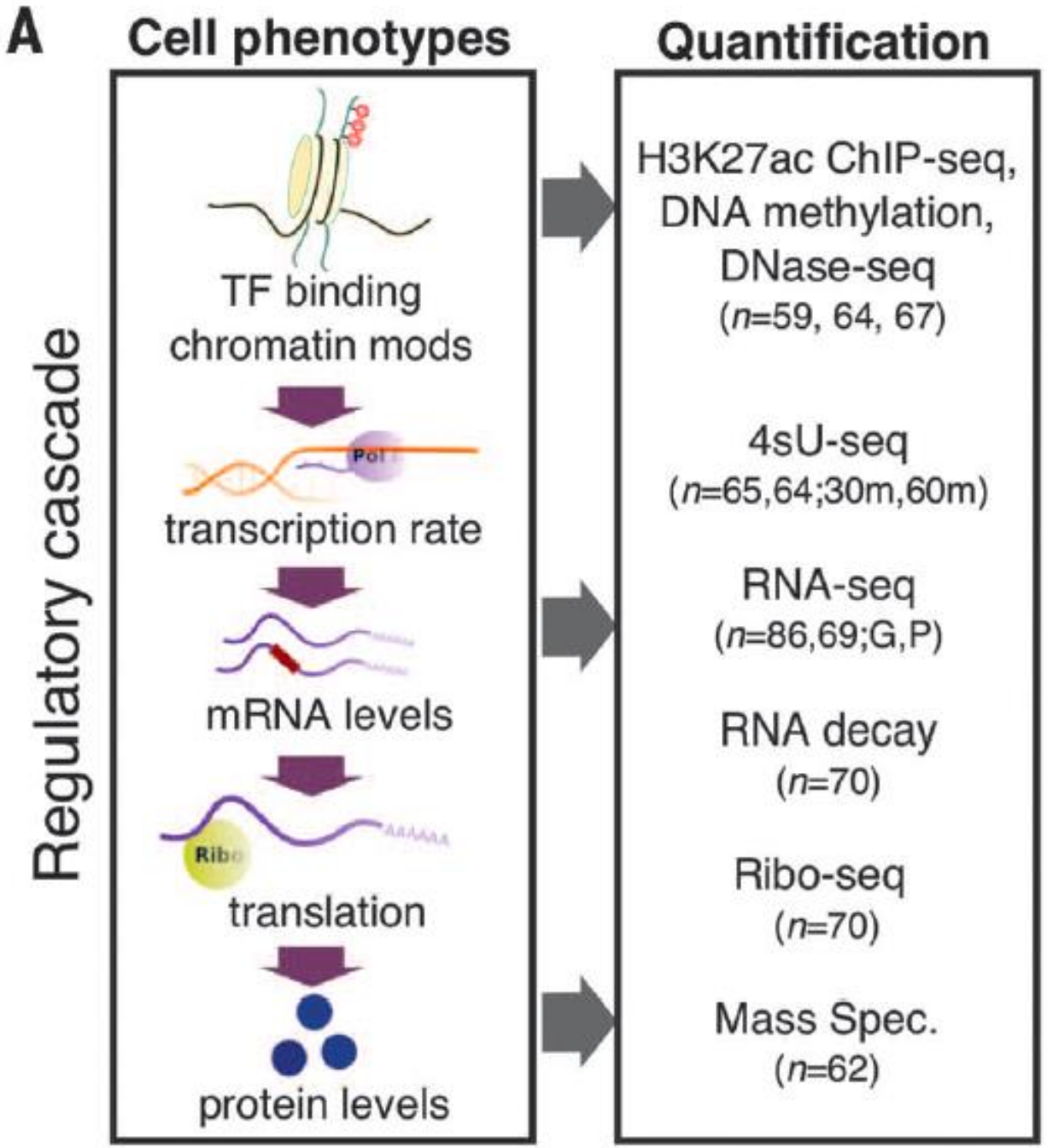


Banovich et al. 2014

Histone mark QTLs are often DNA methylation QTLs

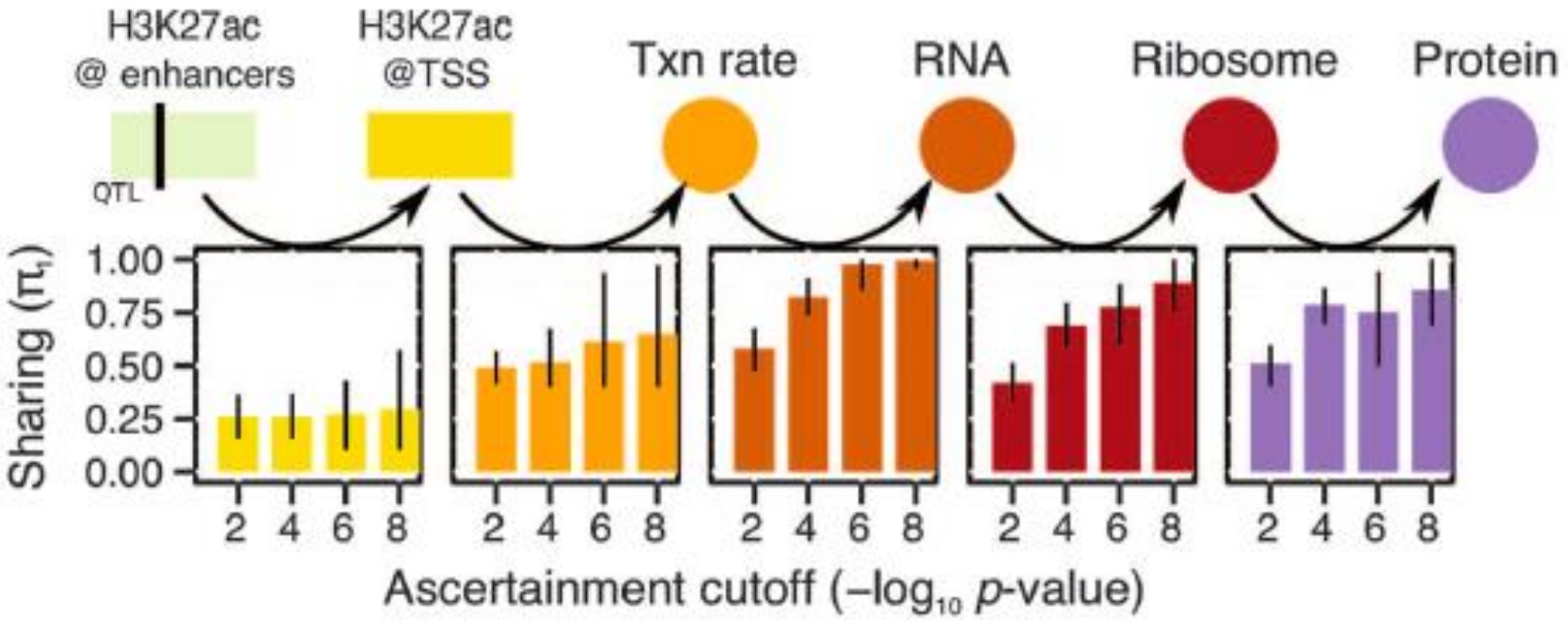


Banovich et al. 2014


A
Cell phenotypes
Quantification
Regulatory cascade
TF binding
chromatin mods
transcription rate
mRNA levels
translation
protein levels
H3K27ac ChIP-seq,
DNA methylation,
DNase-seq
(n=59, 64, 67)
4sU-seq
(n=65,64;30m,60m)
RNA-seq
(n=86,69;G,P)
RNA decay
(n=70)
Ribo-seq
(n=70)
Mass Spec.
(n=62)

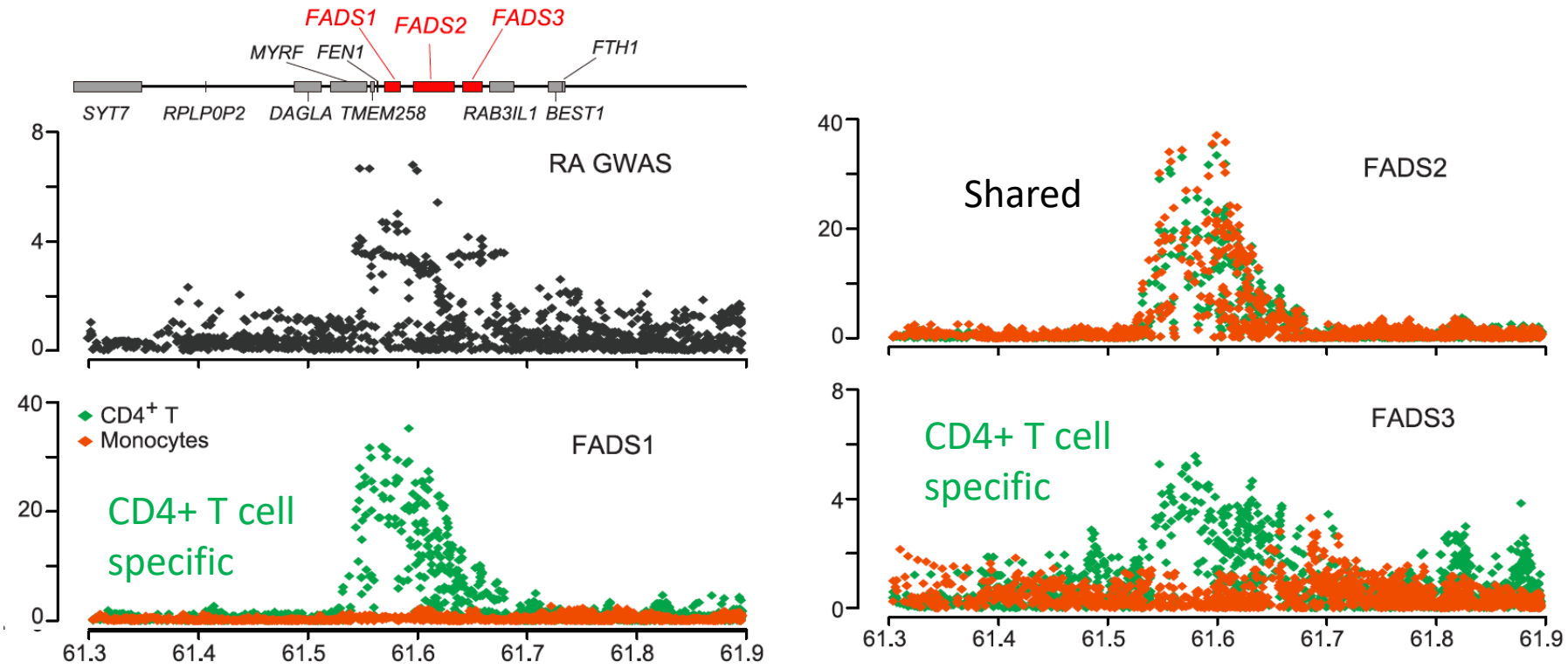
Li et al. 2016

Sharing of regulatory QTLs



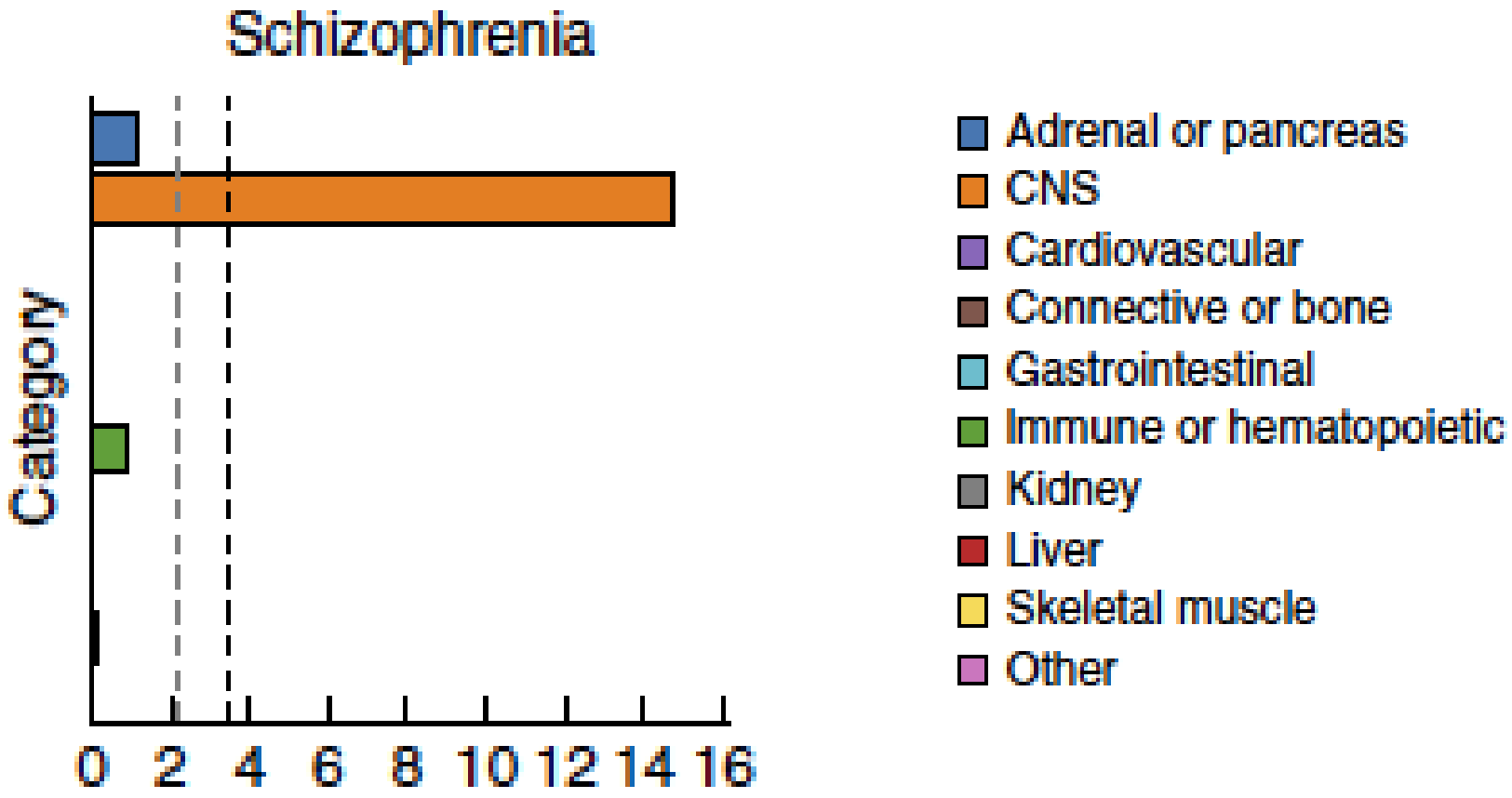
Li et al. 2016

Intersecting eQTLs with GWAS

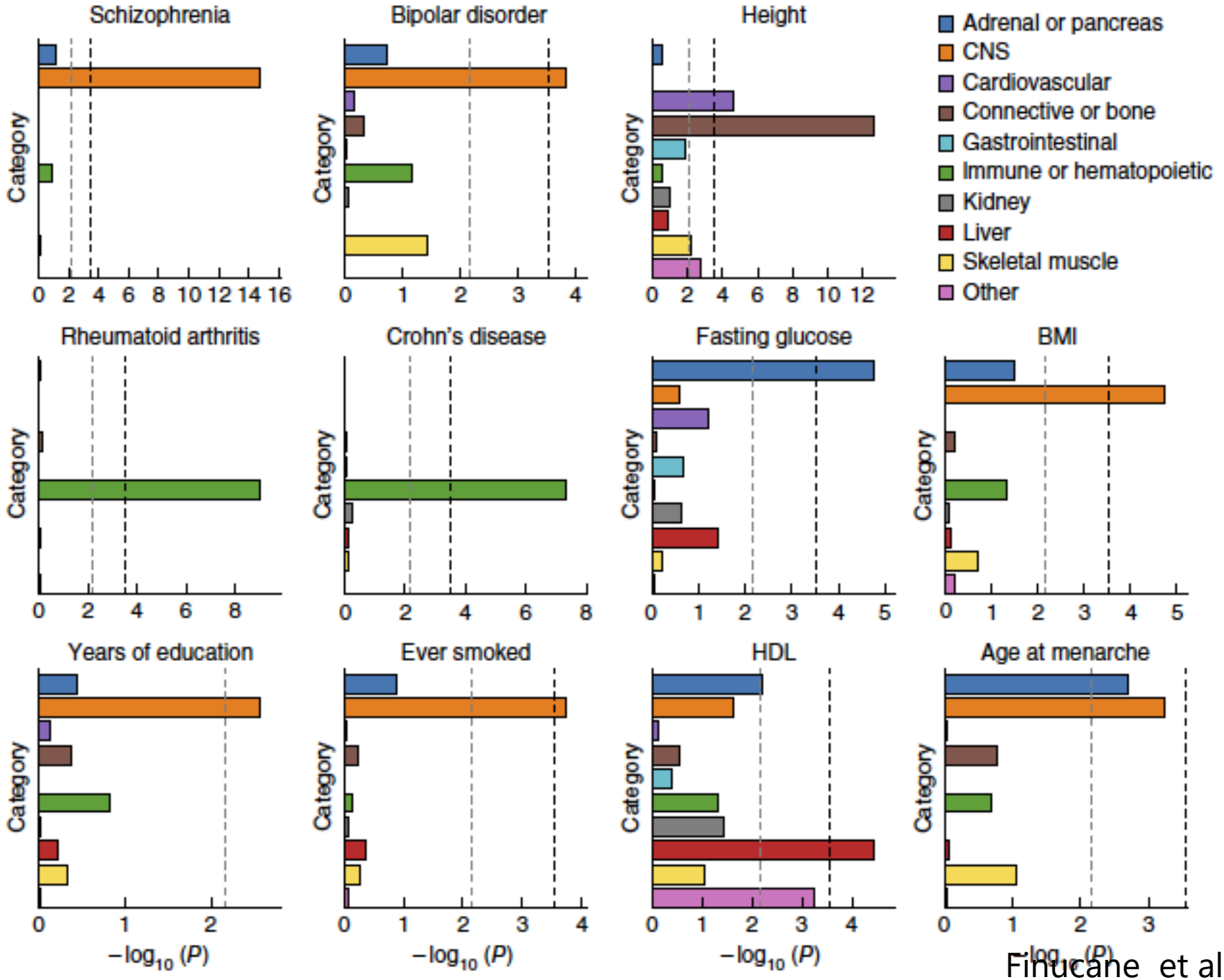


Raj et al. 2014

GWAS enrichment in cell-type specific annotations



Finucane et al. 2015



Finucane et al. 2015

Summary

- Many molecular traits can be associated with genetic variants
- Molecular QTLs can reveal mechanism underlying organismal traits
- Smaller samples are needed to map molecular QTLs that organismal traits
- Genetic associations are challenging to interpret