

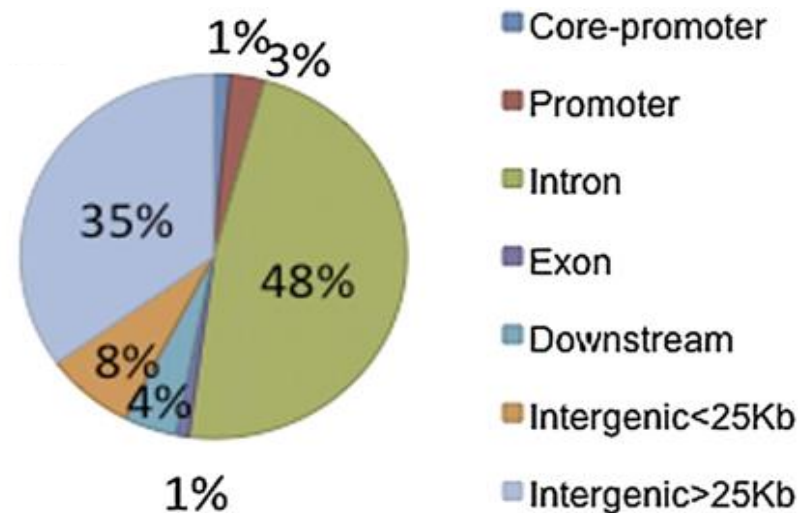
Determinants of Transcription Factor Regulatory Range

Cliff Meyer

GeCo Workshop in Como
5 March 2019

How to Assign Target Genes for Transcription Factor ChIP-seq?

- Initial goal of ChIP-chip/seq is to find target genes of a TF
- Typical TF ChIP-seq
 - So many peaks!
 - Small % in promoters
 - What are target genes?
- GWAS SNP
 - Which ones are causal, what TFs are involved, what genes do they target?

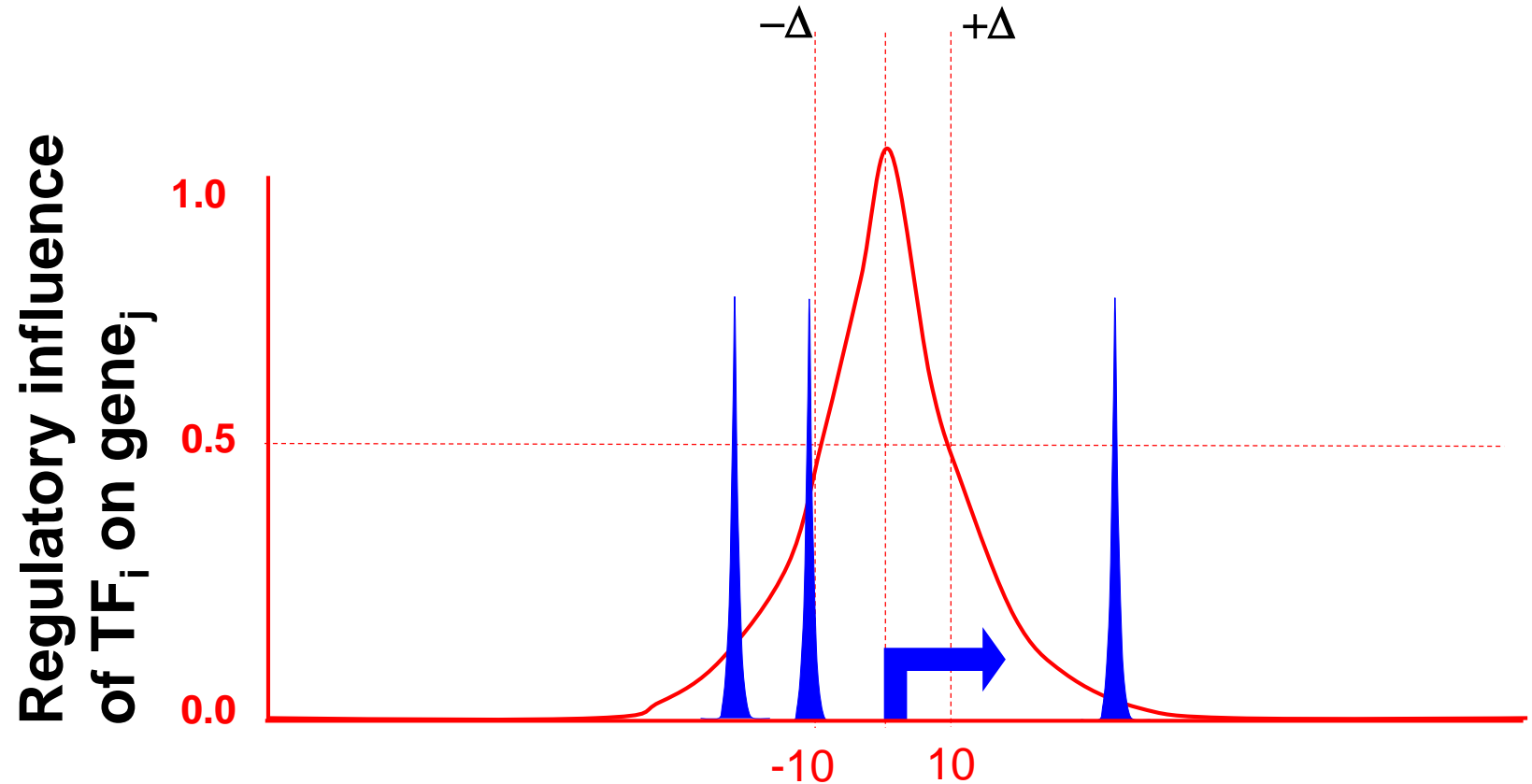


Assign TF Target Genes from ChIP-seq

- **Model #1:** Binding within **x** kb of transcription start sites (TSS)?
- **Model #2:** Nearest active gene for each binding site.
- **Model #3:** Number of binding sites from TSS with **decay distance**.
- **Model #4:** GWAS hits often assigned to the nearest active gene in the same topologically associating domain (TAD)
 - Discovery of TADs, Dixon et al, Nat 2012
 - Nearest active gene can be >> 100KB away from GWAS hits

1. What are the **regulatory
ranges of transcription factors?**

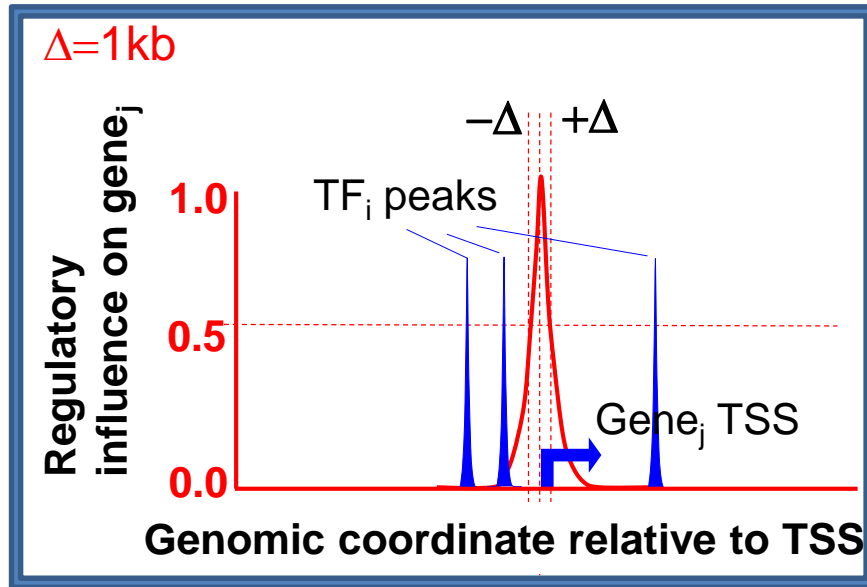
Regulatory Potential Model of Enhancer Activity



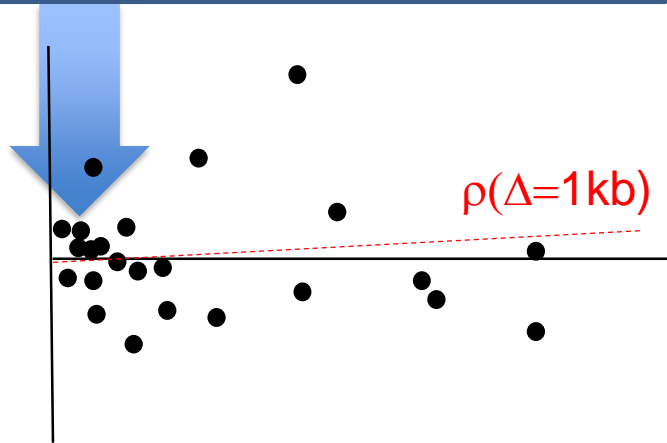
$$RP = \sum_{TF_{peaks}} 2^{-\frac{\text{distance to TSS}}{D}}$$

Genomic coordinate relative to TSS [kb]

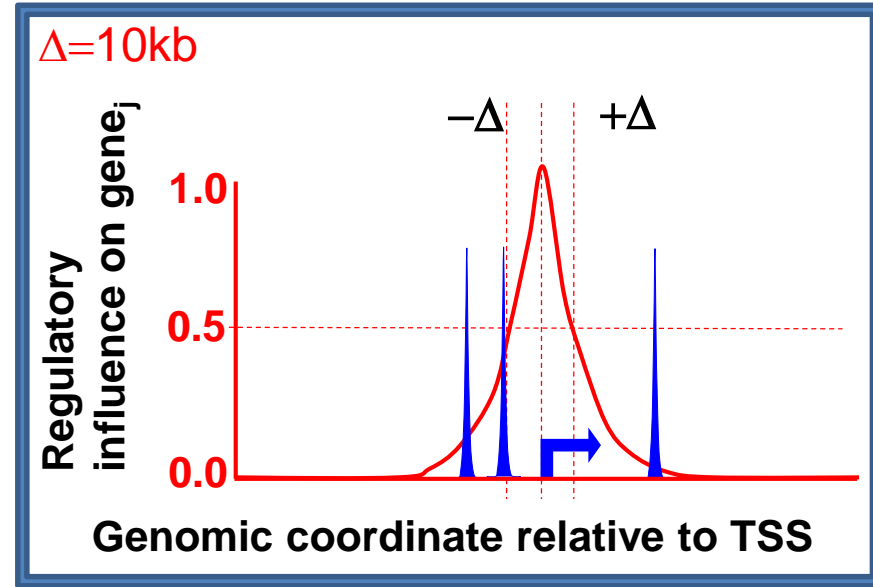
Distance decay effect on gene expression



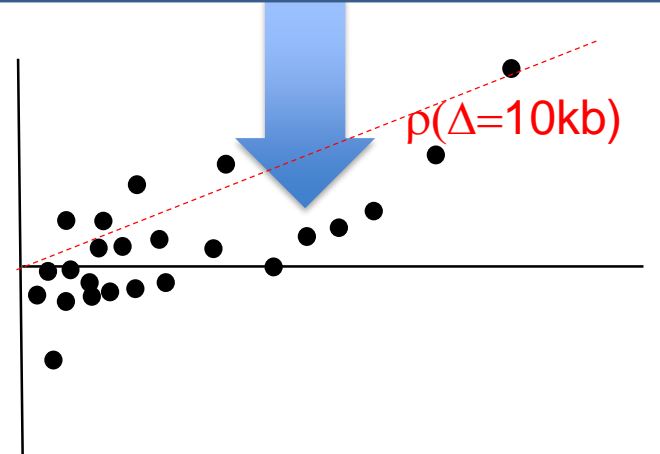
CCLF gene expression correlation Tf_i and gene_j



Regulatory potential ($\Delta=1\text{kb}$) of Tf_i on gene_j

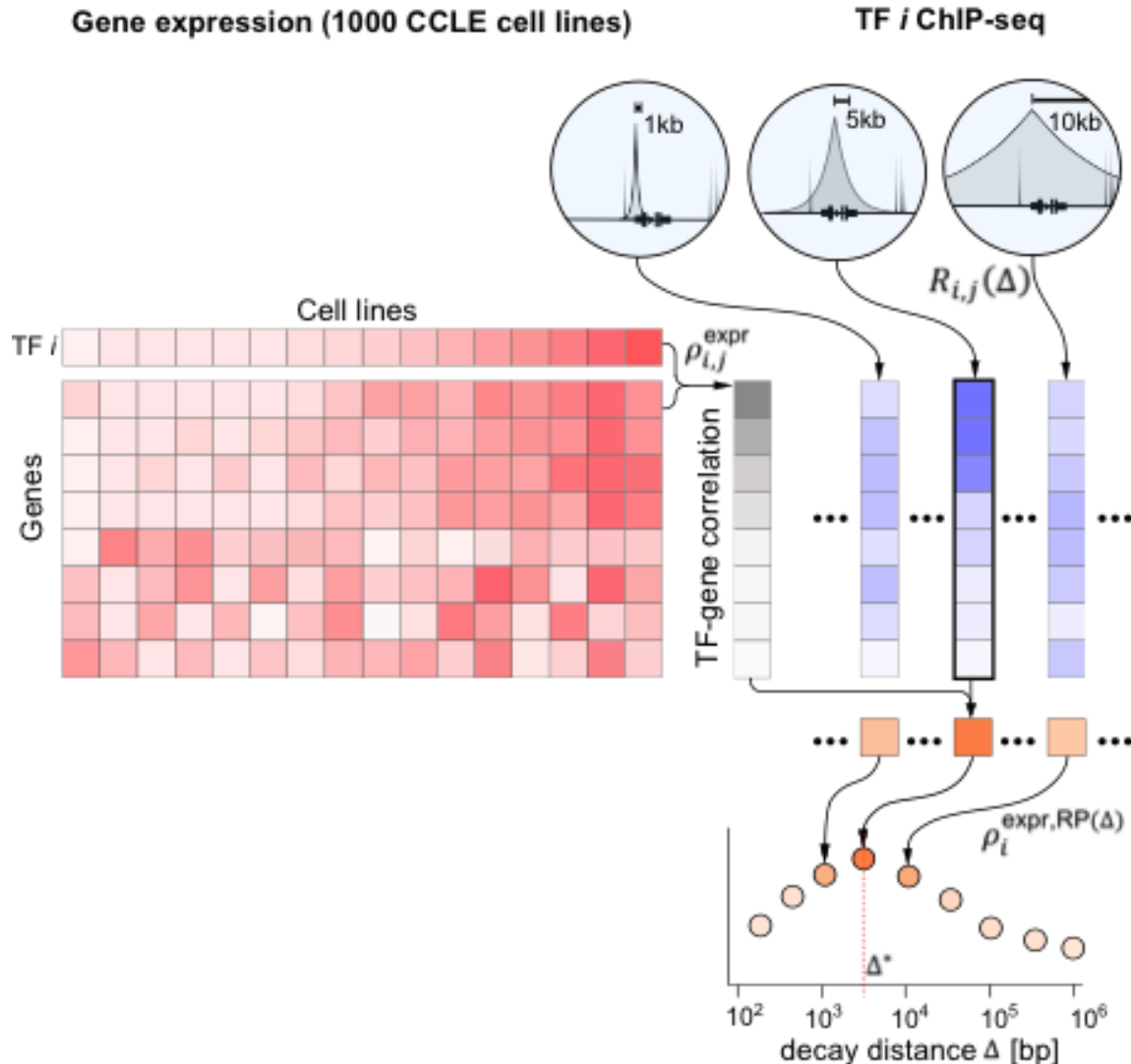


CCLF gene expression correlation Tf_i gene_j



Regulatory potential ($\Delta=10\text{kb}$) of Tf_i on gene_j

Estimating the Regulatory Distance of a Single TF





Cistrome DB

<http://cistrome.org/db>

Dataset Browser

Containing word(s):

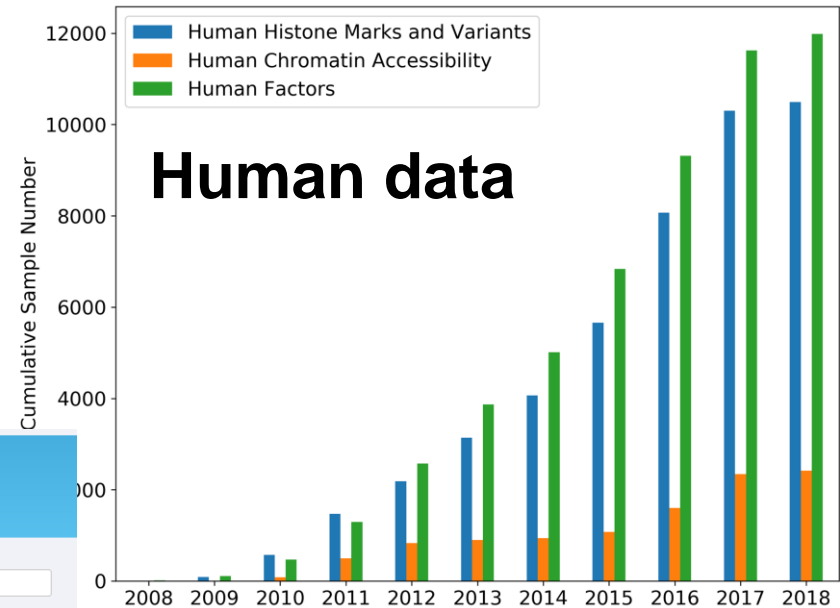
Species
All
Homo sapiens
Mus musculus

Biological Sources
All
22RV1
BicR
C4-2B
Colon
CWR22Pc

Factors
All
AR

Results

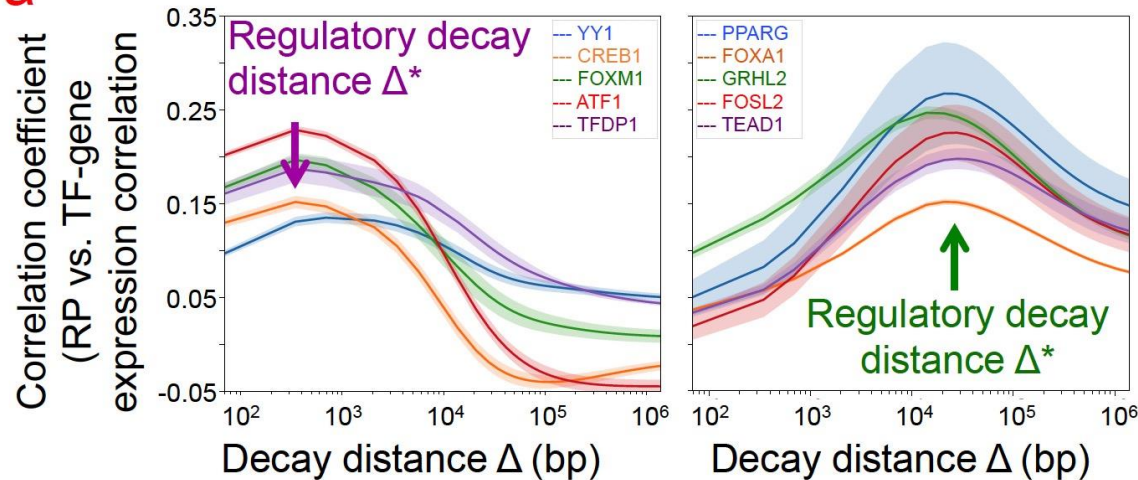
Batch	Species	Biological Source	Factor	Publication	Status
<input type="checkbox"/>	Homo sapiens	VCaP; Epithelium; Prostate	AR	Asangani IA, et al. Nature 2014	completed
<input type="checkbox"/>	Homo sapiens	DUCaP; Epithelium; Prostate	AR	Bu H, et al. Hum. Mutat. 2015	completed
<input type="checkbox"/>	Homo sapiens	R1-AD1; Epithelium; Prostate	AR	Chan SC, et al. Nucleic Acids Res. 2015	completed
<input type="checkbox"/>	Homo sapiens	Epithelium; Prostate	AR	Chen Z, et al. EMBO J. 2015	completed
<input type="checkbox"/>	Homo sapiens	LNCaP; Epithelium; Prostate	AR	Chen Z, et al. EMBO J. 2015	completed
<input type="checkbox"/>	Homo sapiens	VCaP; Epithelium; Prostate	AR	Chng KR, et al. EMBO J. 2012	completed
<input type="checkbox"/>	Homo sapiens	LNCaP; Epithelium; Prostate	AR	Decker KF, et al. Nucleic Acids Res. 2012	completed



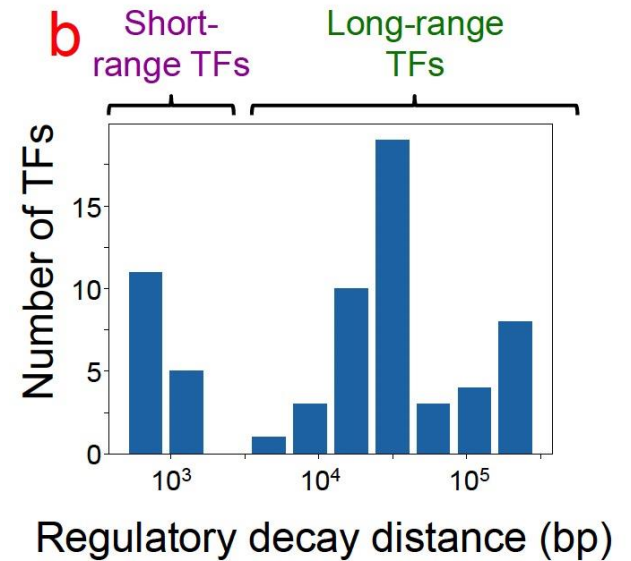
- > 4 samples
- > 20,000 peaks per sample
- > Maximum correlation > 0.1

Short-range and Long-range TFs

a

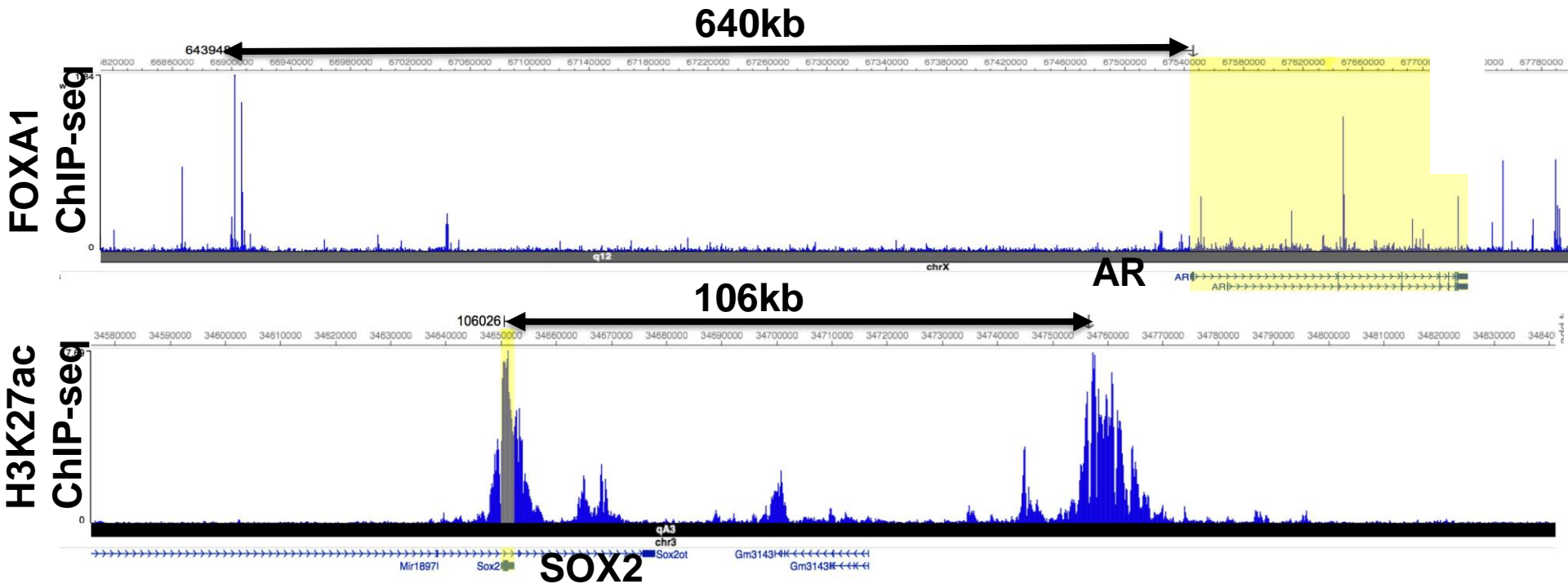


b



**2. Where in the genome are the
distant enhancers ?**

Very long range enhancers

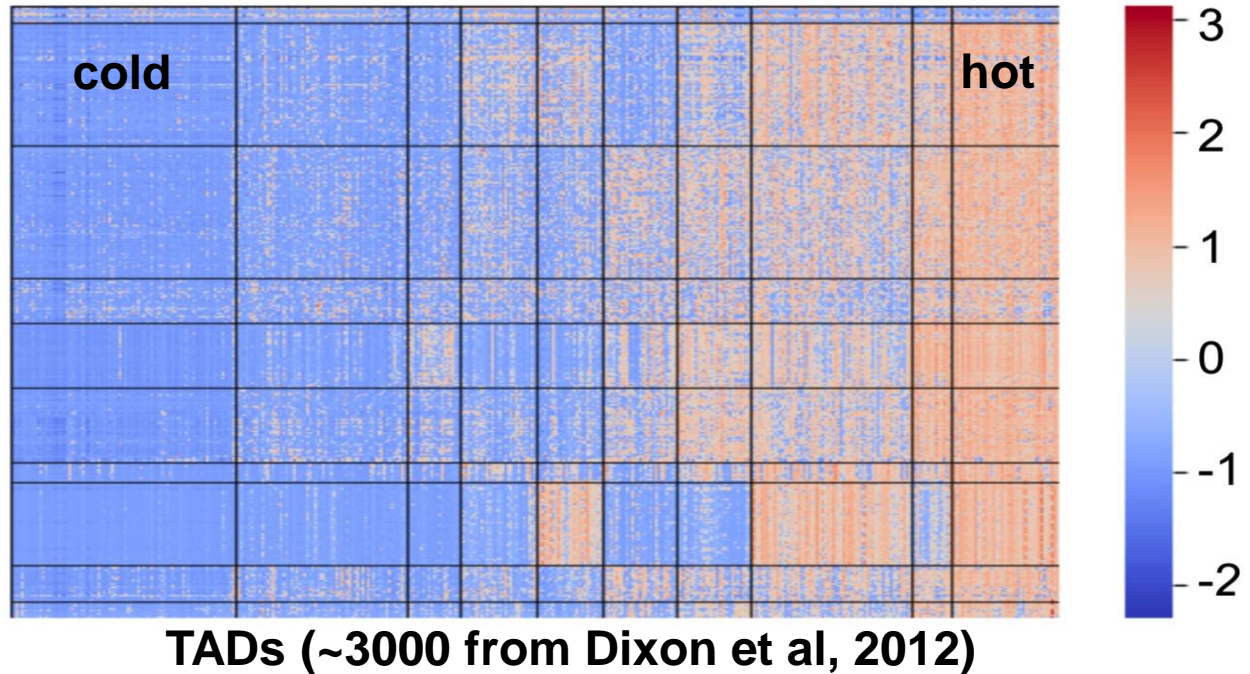


Clustering of TADs by H3K27ac

H3K27ac ChIP-seq samples
(~1,500 from Cistrome DB)

neuronal

lymphoblastoid



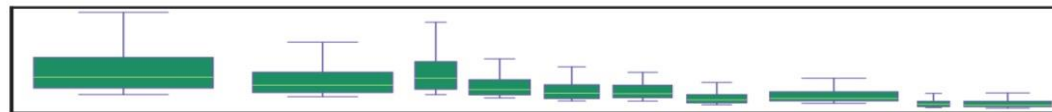
gene density



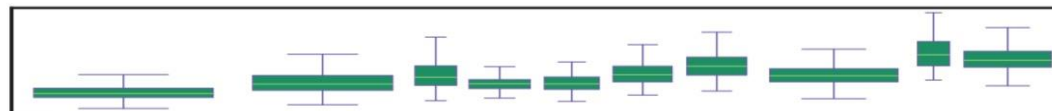
gene expression



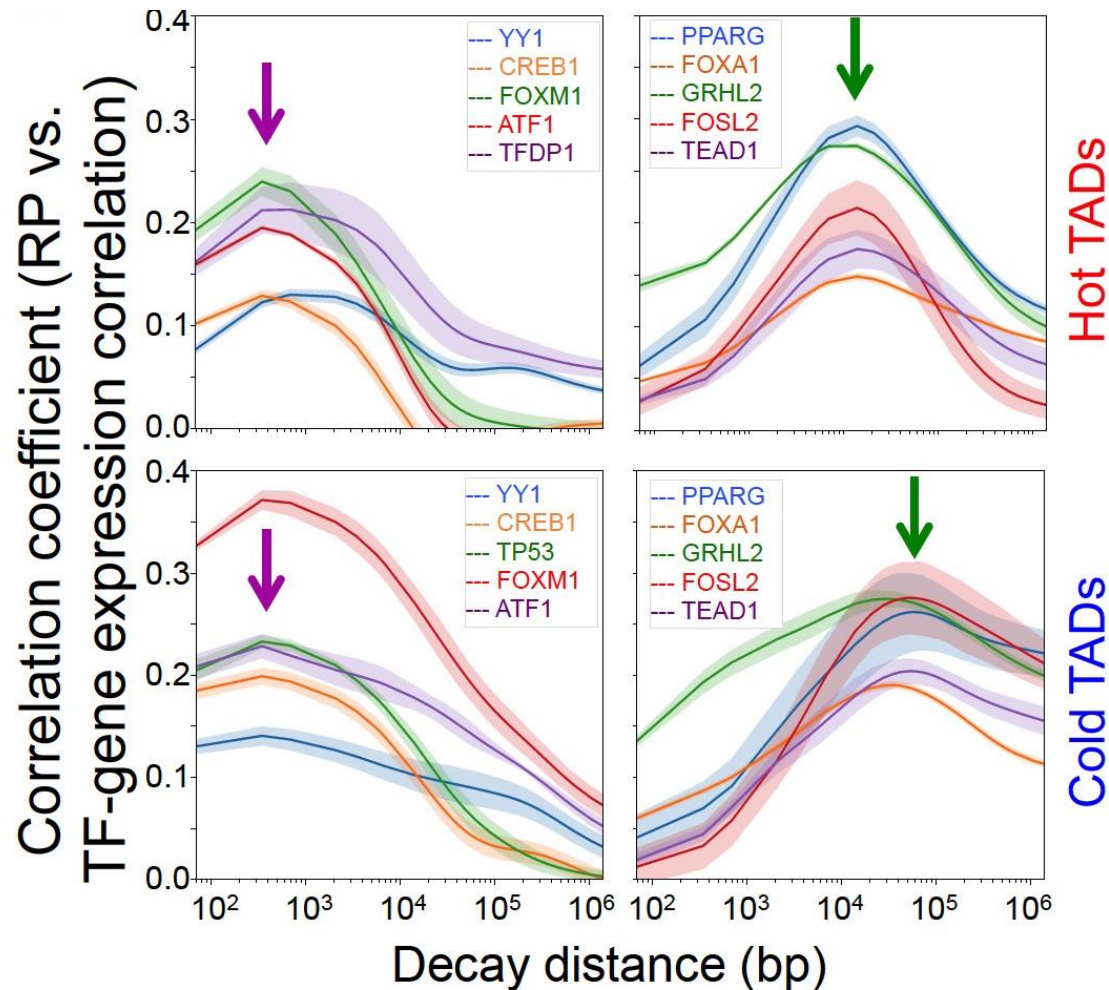
90th percentile
transcript length



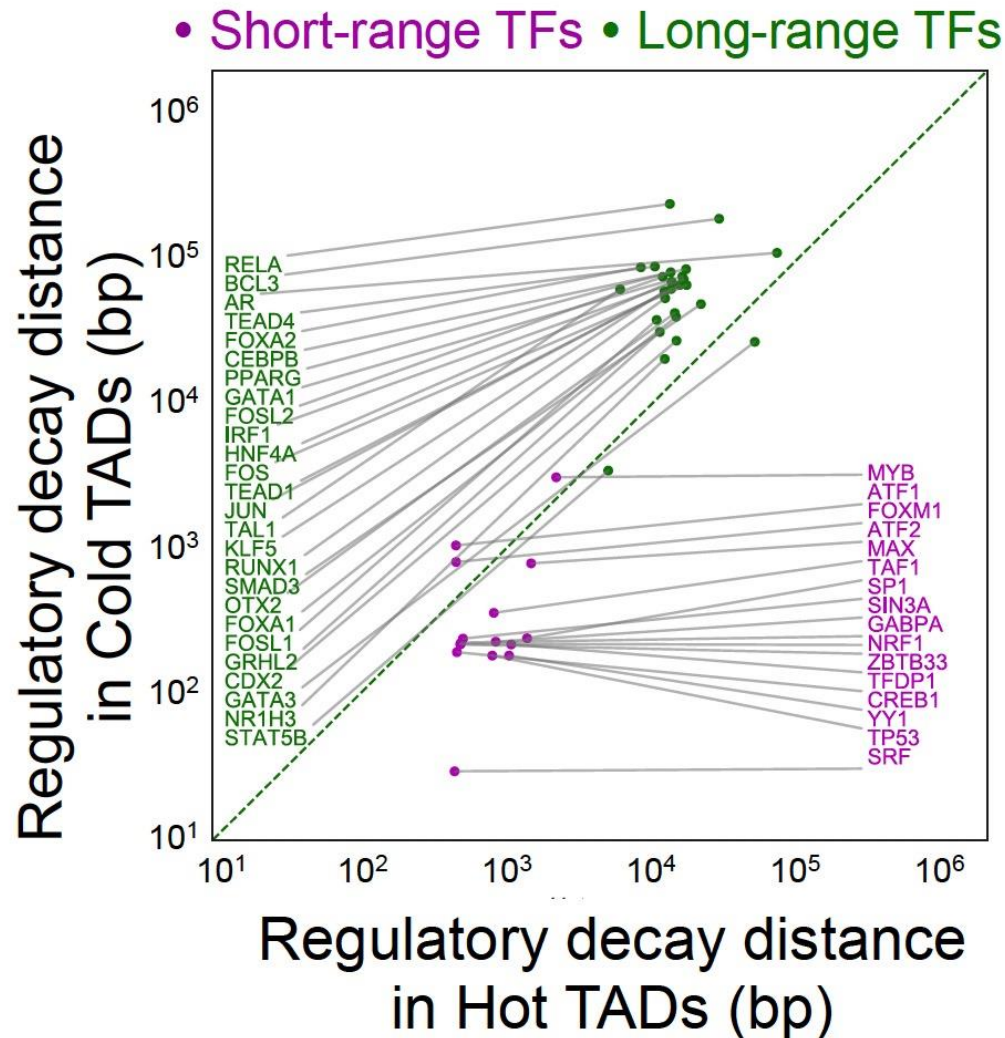
GC fraction



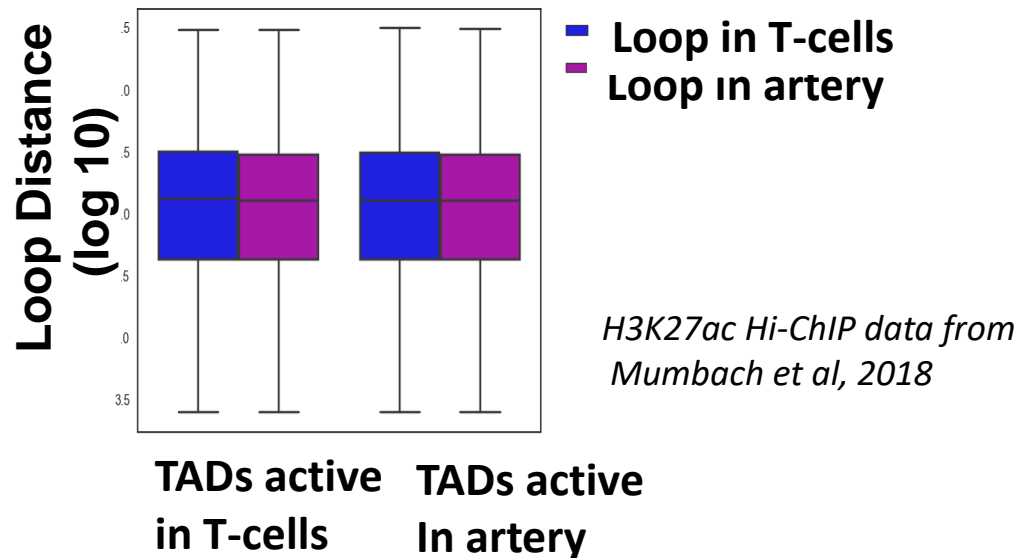
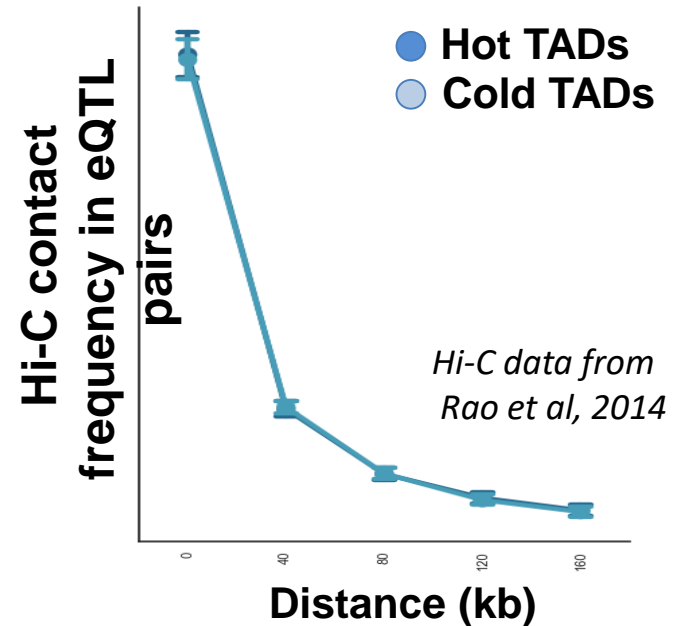
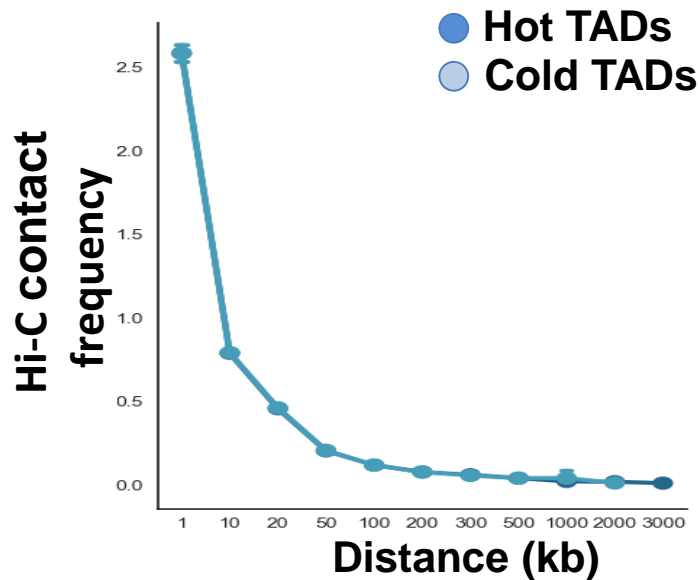
Regulatory decay rates vary between TFs and TAD types



Regulatory Decay Rates of TF types and TAD types

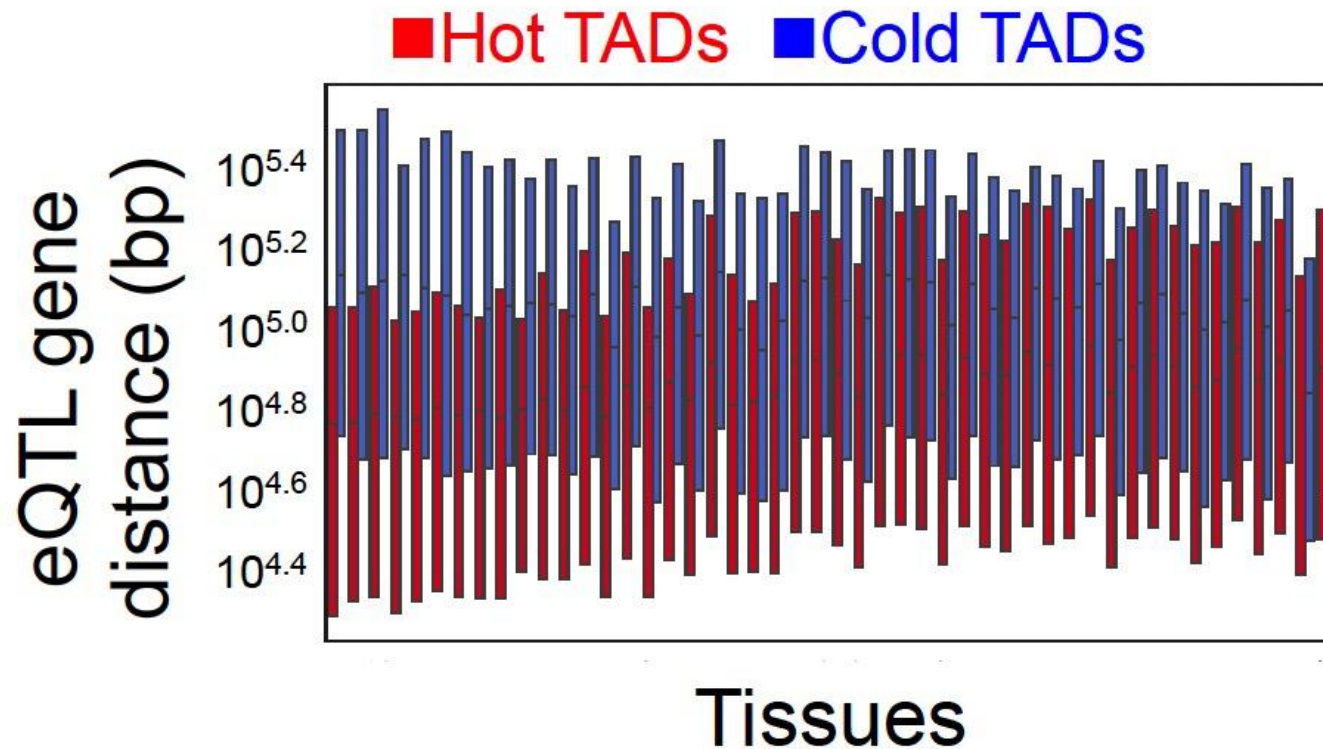


Regulatory decay rates are not explained by Hi-C or Hi-ChIP data

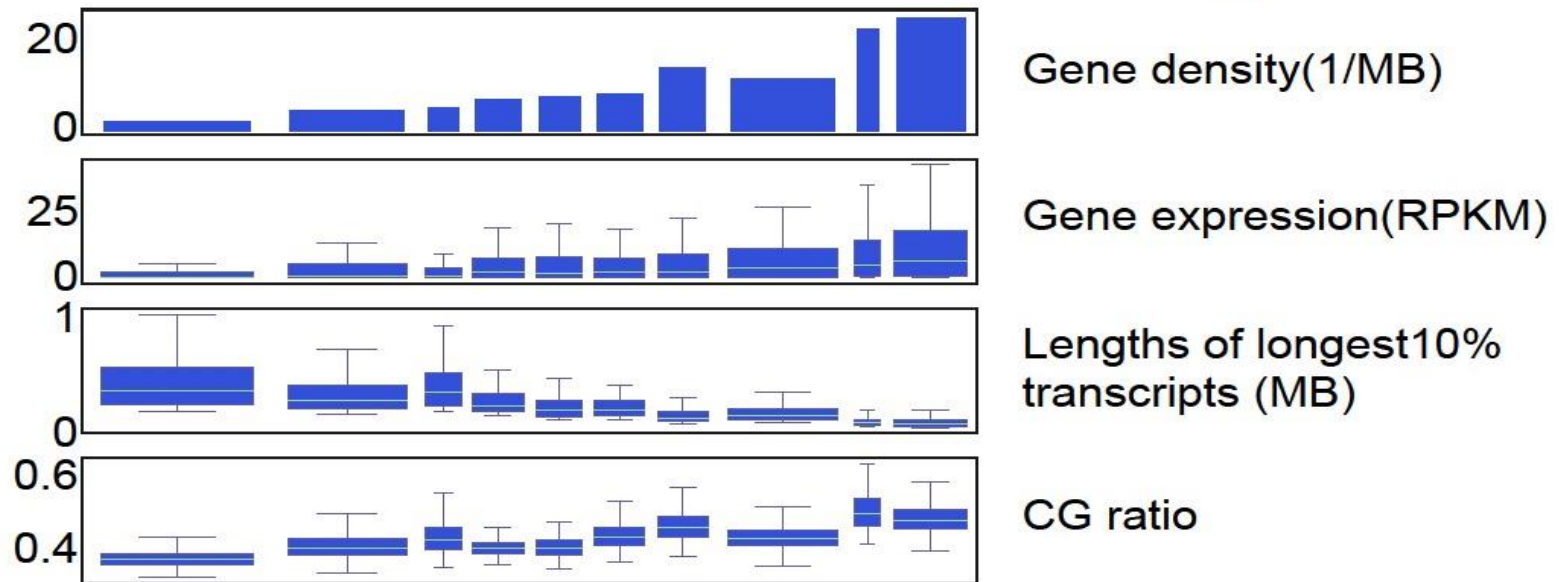


3. Is it **genetic or epigenetic effects that influence the regulatory range of transcription factors?**

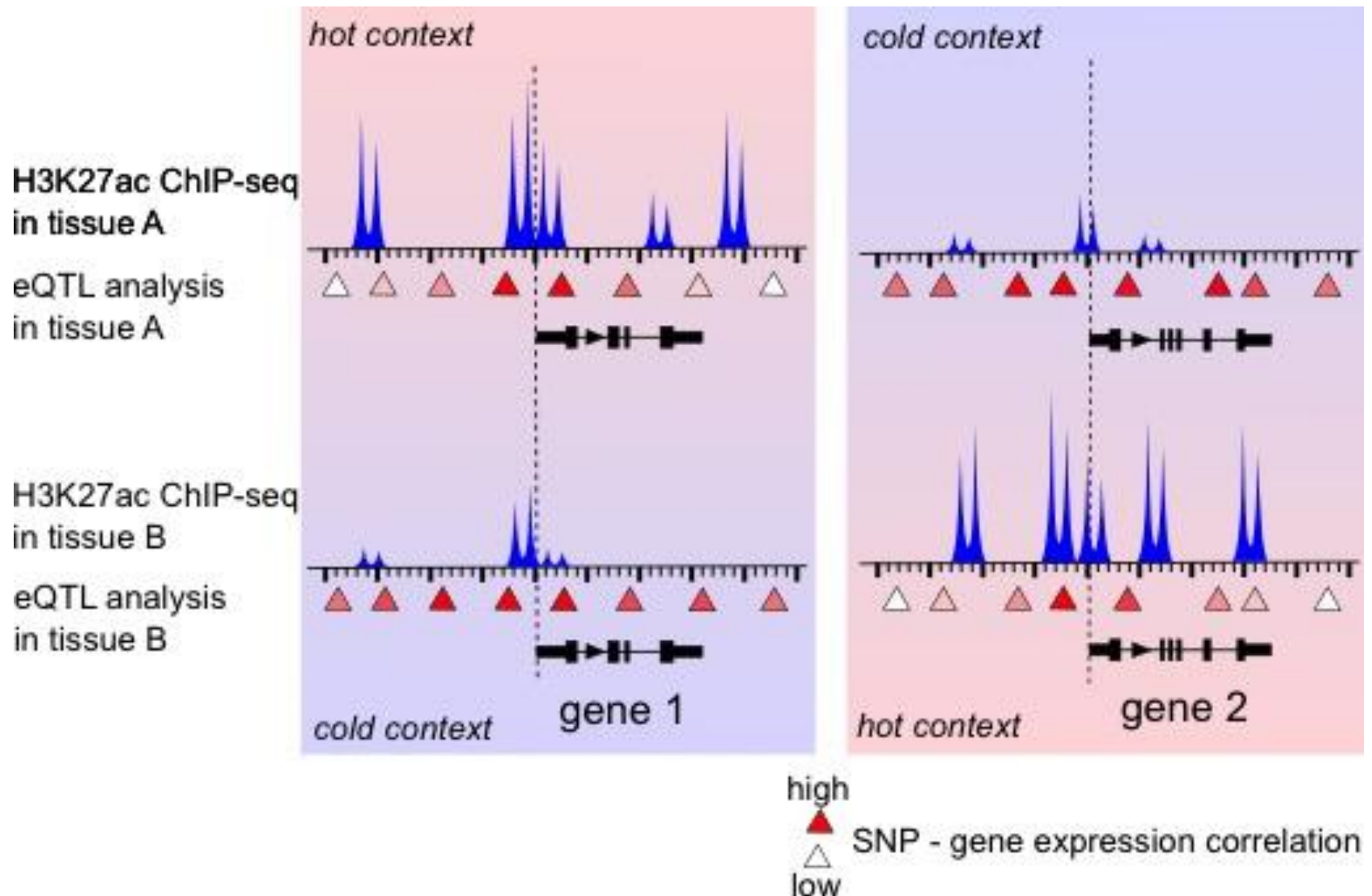
eQTL distances to TSS



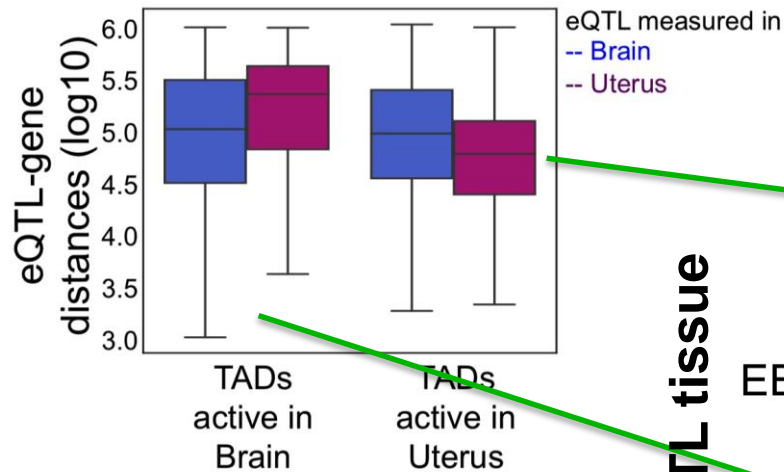
Genetic features of hot and cold TADs



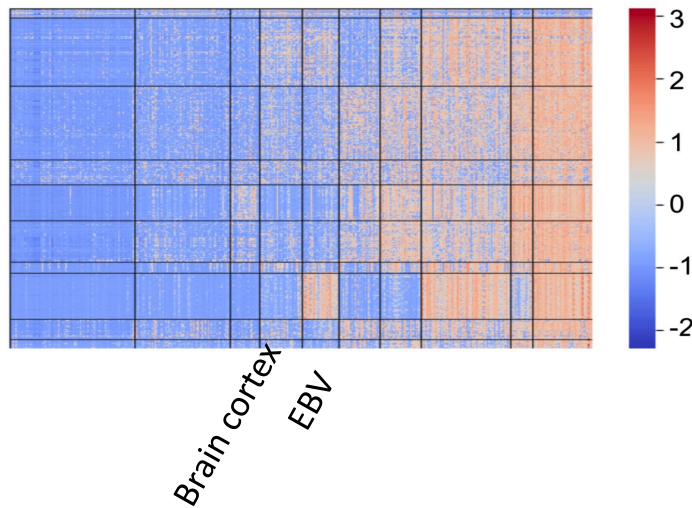
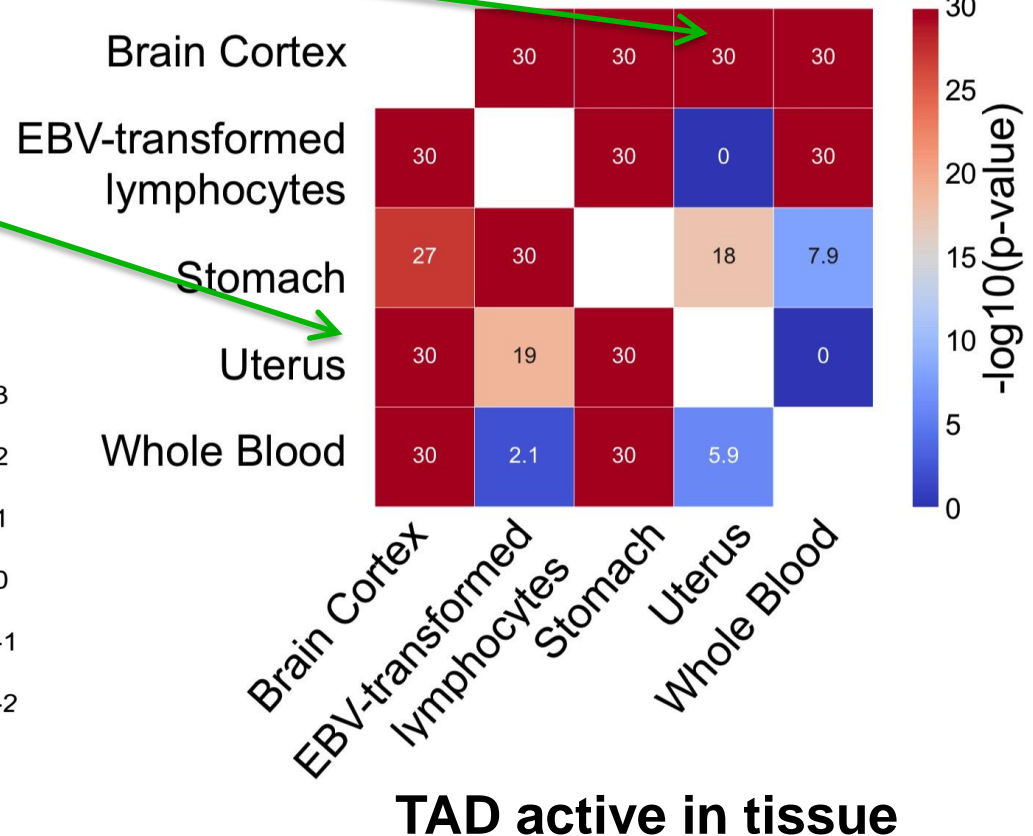
Does chromatin state in tissue influence eQTL range of effect?



Tissue restricted eQTLs in TADs with tissue restricted activity



eQTL tissue

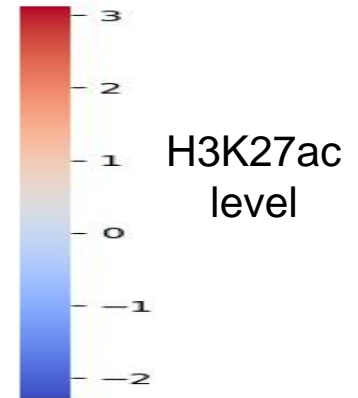
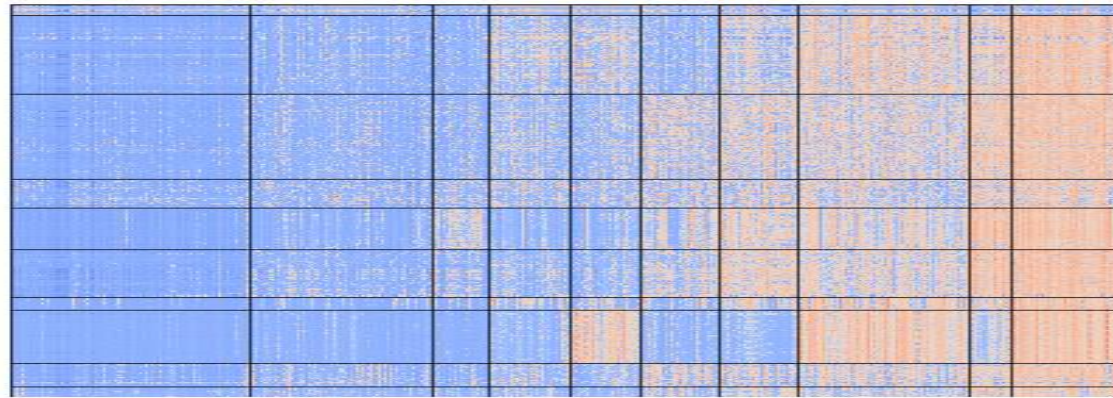


4. Do specific transcription factors dominate the regulation of specific TADs?

What are the dominant TFs in each TAD cluster?

H3K27ac
ChIP-seq samples

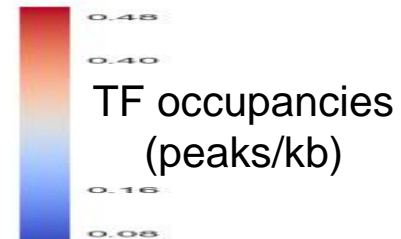
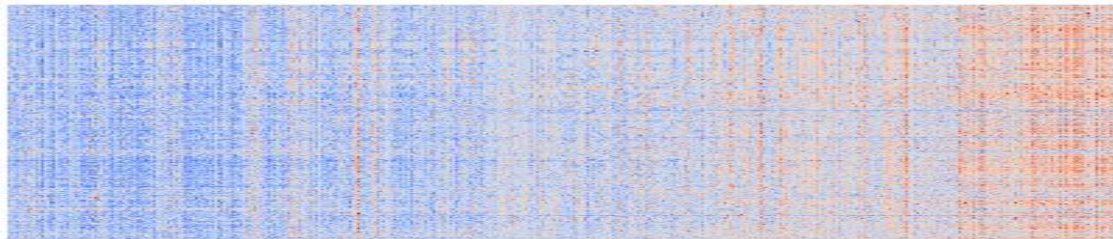
TF ChIP-seq
(~3,000 samples
from Cistrome DB)



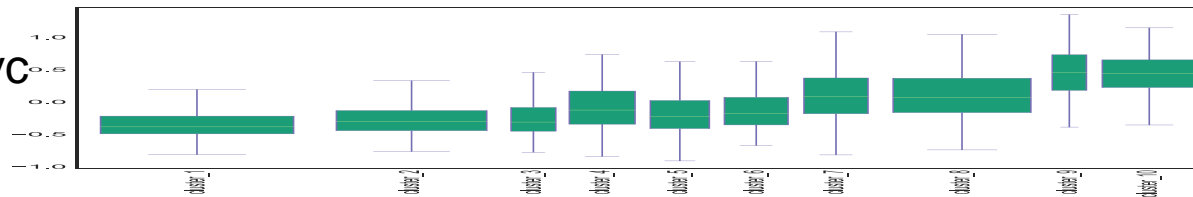
Cold

TADs

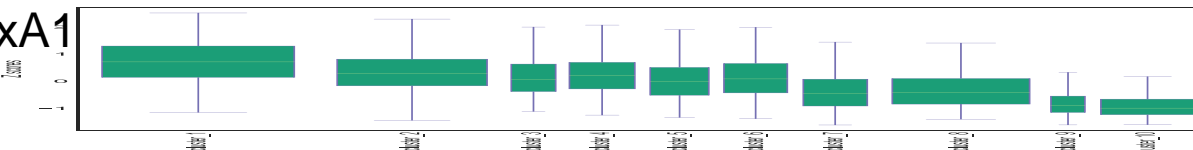
Hot



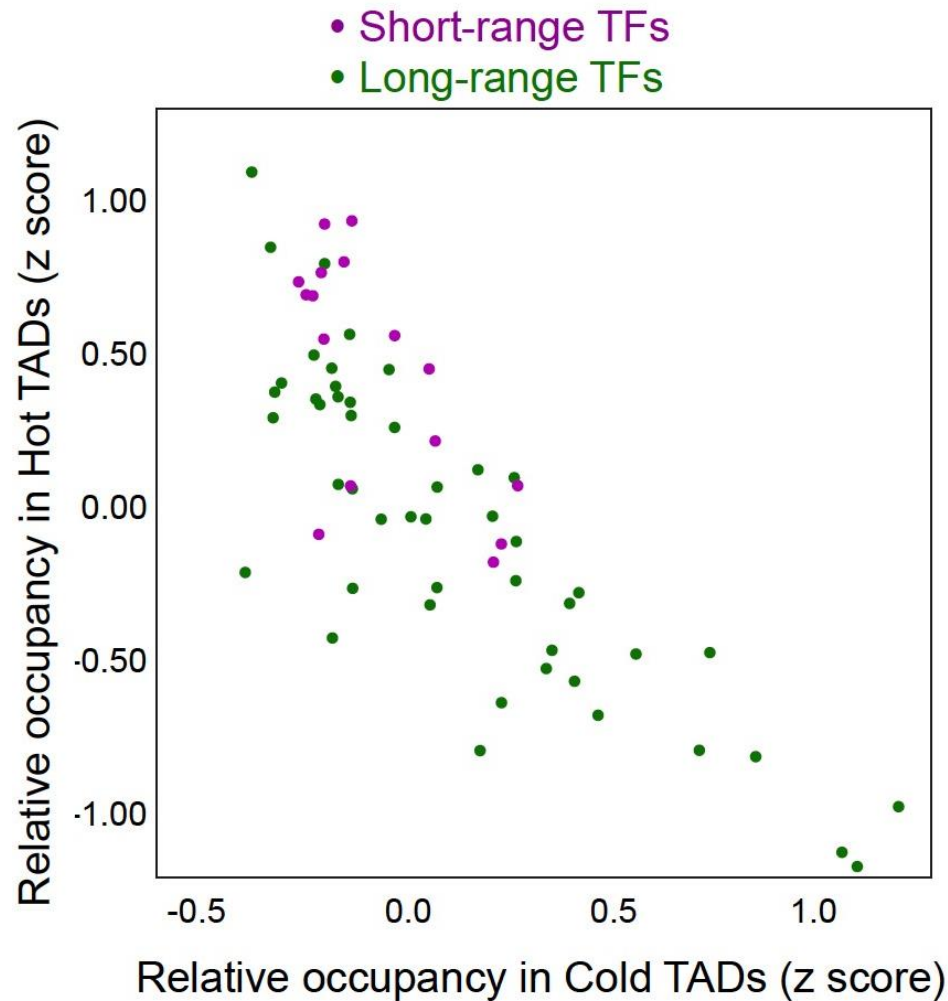
Relative Myc
occupancy



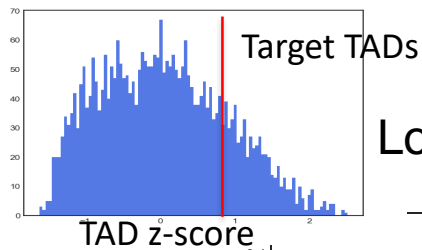
Relative FoxA1
occupancy



What are the dominant TFs in hot and cold TAD clusters?



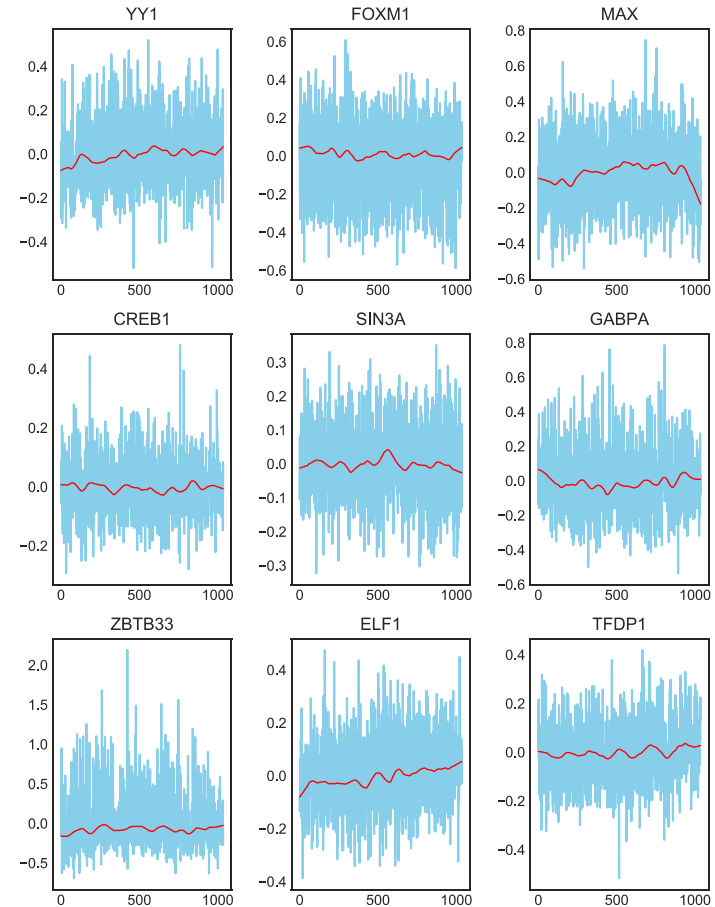
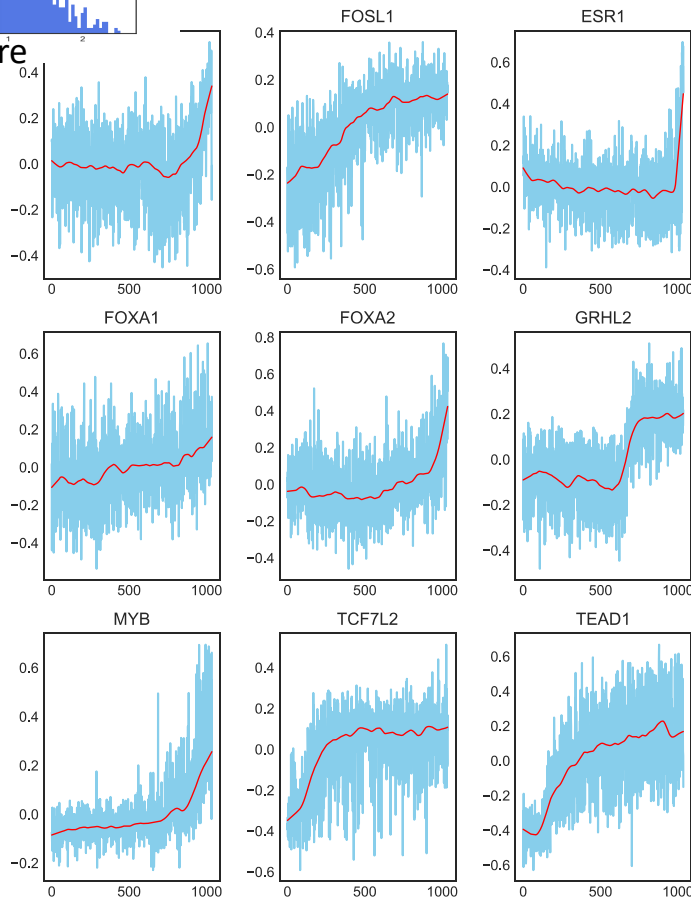
TF regulation of genes in target TADs



Long-range TFs

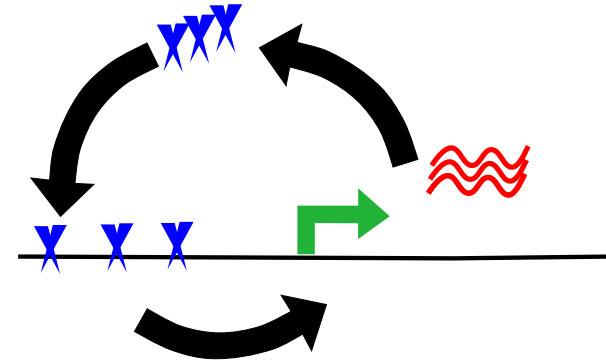
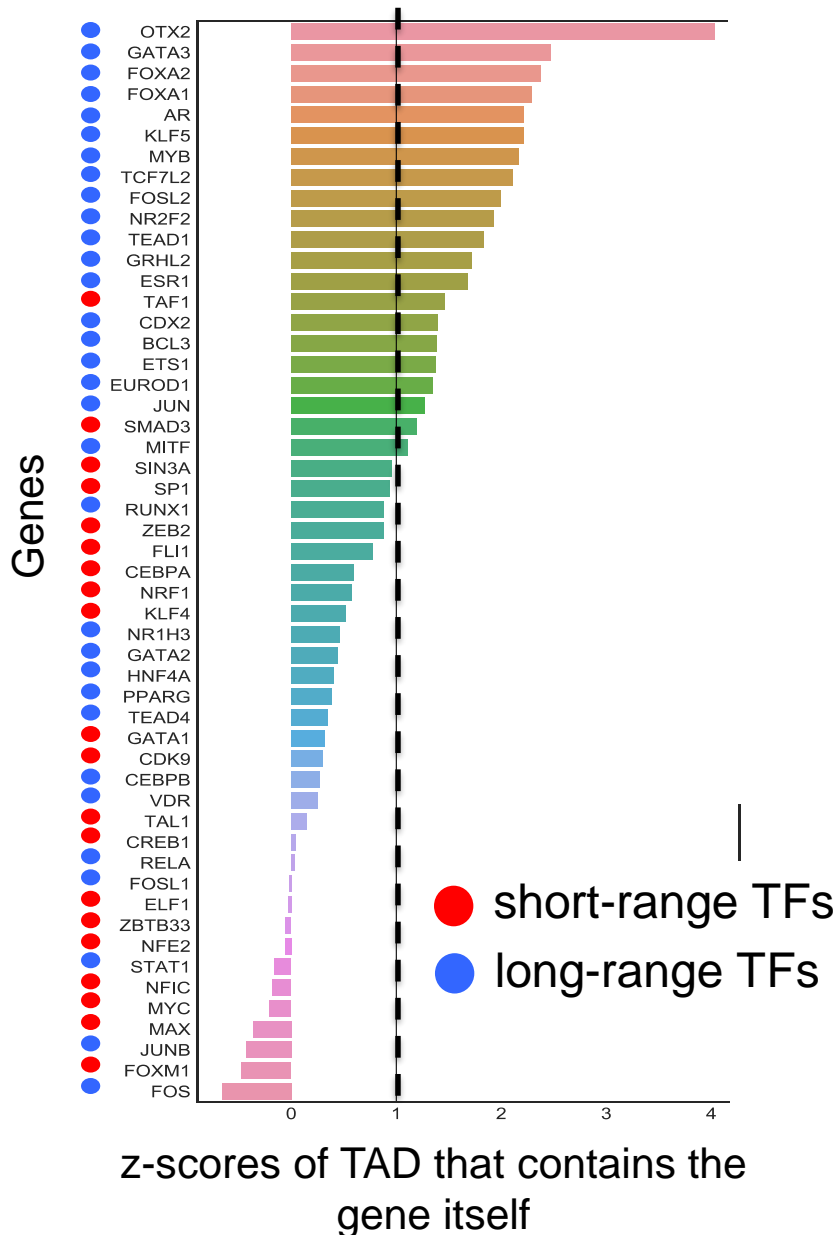
Short-range TFs

Average expression of genes
in target TADs

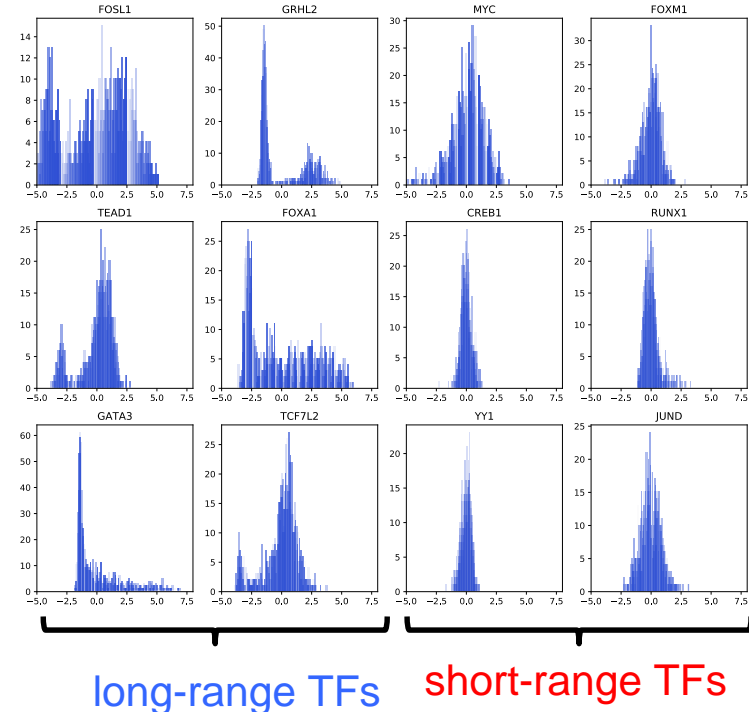


Rank of TF Expression in CCLE cell lines

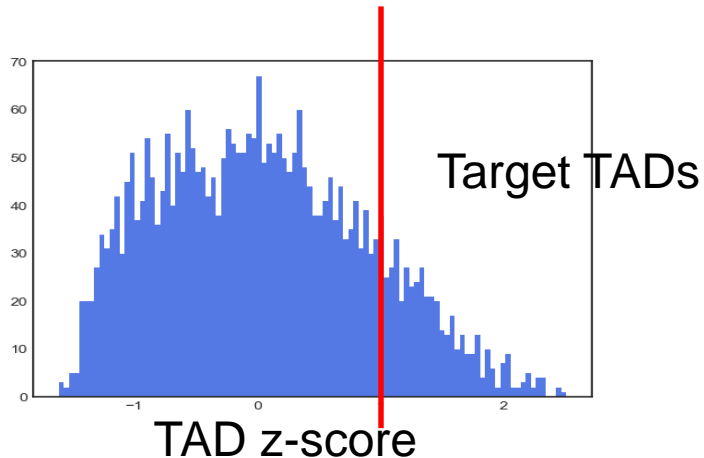
TF TAD positive feedback



Distribution of gene expression in CCLE



Gene function in targeted TADs



- **FOXA1**: system development ($6e-13$)
- **GRHL2**: cornification ($1e-30$)
- **SPIB**: B cell activation ($4.6e-23$)
- **TEAD1**: regulation of cell migration ($5.7e-13$)
- **FOSL2**: anatomical morphogenesis ($6.1e-14$)
- **GATA3**: regulation of multicellular organismal process ($1e-13$)

Summary

- ① What are the **regulatory ranges** of transcription factors?
There are two types of TF, short- and long-range.
Short-range: <1kb.
Long-range: ~30kb
- ② Are very **distant enhancers** a genomic context dependent phenomenon?
Enhancers influence genes over longer genomic distances in cold TADs.
- ③ Do **genetic or epigenetic** factors influence the regulatory range of TFs?
Epigenetic factors are involved although hot and cold TADs have very different genetic properties.
- ④ Do specific **TFs** dominate the regulation of specific TADs?
Lineage specific factors predominate the regulation of

Acknowledgements

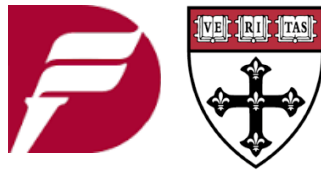
Chen-Hao Chen

X. Shirley Liu
Myles Brown

Peng Jiang
Shengqing Gu
Jingyu Peng
Xihao Hu
Qin Tang
Avinash Sahu

Rongbin Zheng
Changxin Wan
Shenglin Mei

Qiu Wu
Hanfei Sun
Muyuan Zhu
Jiaxin Wu
Xiaohui Shi



Questions

- ① What are the **regulatory ranges** of transcription factors?
- ② Where in the genome are the **distant enhancers**?
- ③ Do **genetic or epigenetic** factors influence the regulatory range of TFs?
- ④ Do specific **transcription factors** dominate the regulation of specific TADs?
- ⑤ Should we consider TAD types in the interpretation of **non-coding GWAS** hits?