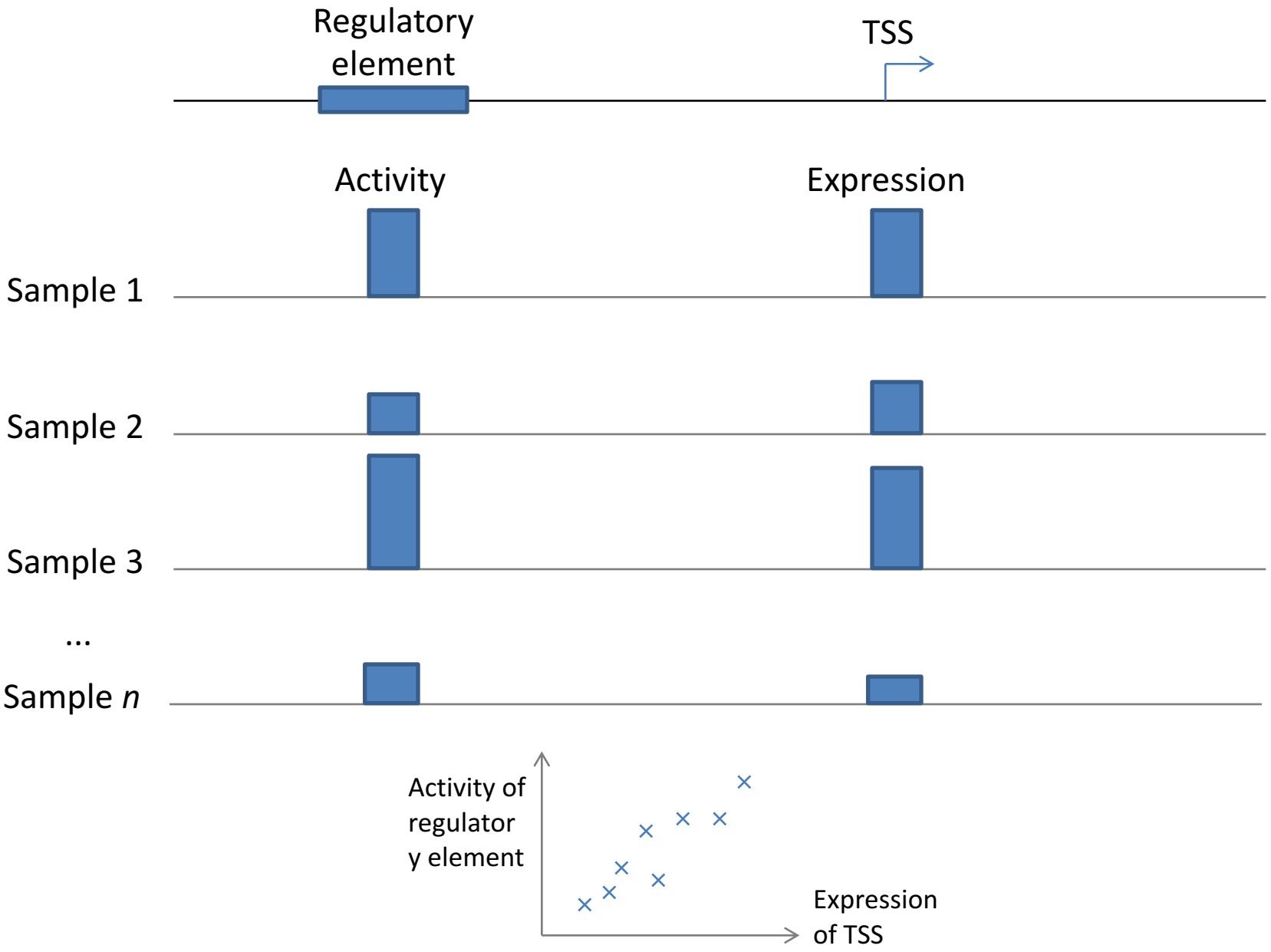
Reconstruction and analysis of enhancer-target networks

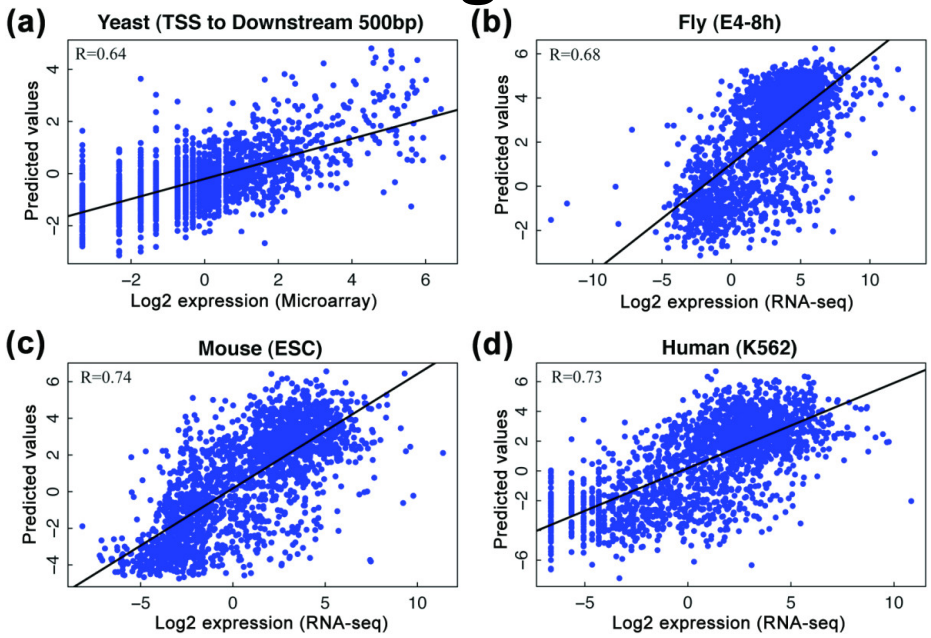
Kevin Yip (with Mark Gerstein)

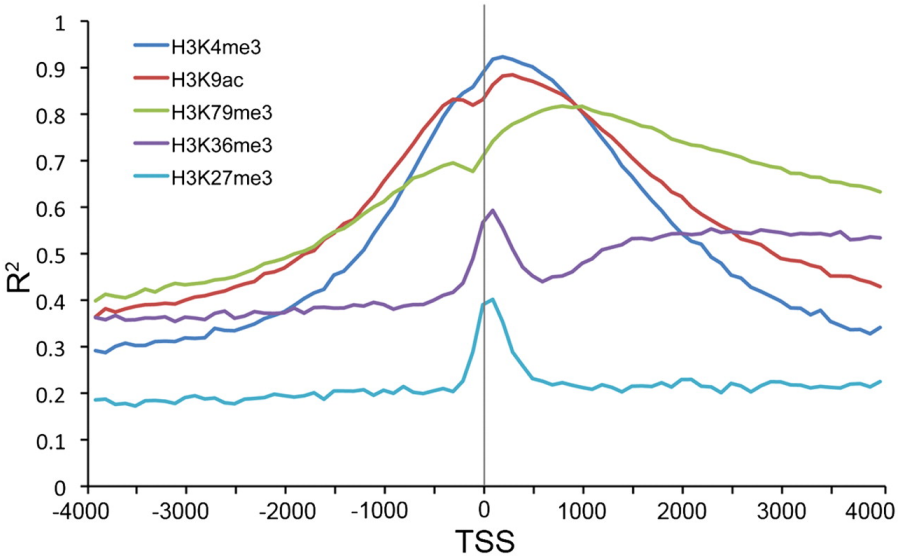
Expression modeling

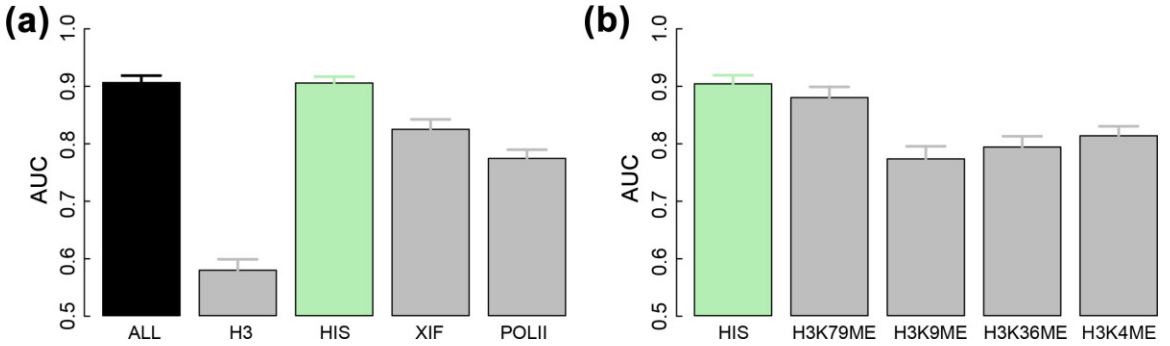


Evidence of quantitative relationships

- Promoters and gene bodies: histone mark and TF binding models



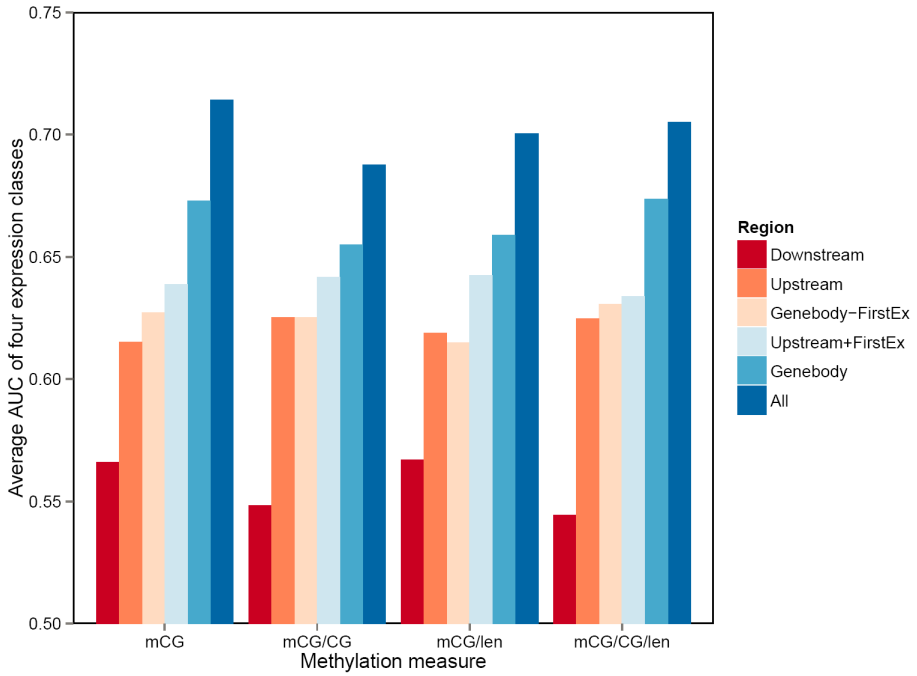


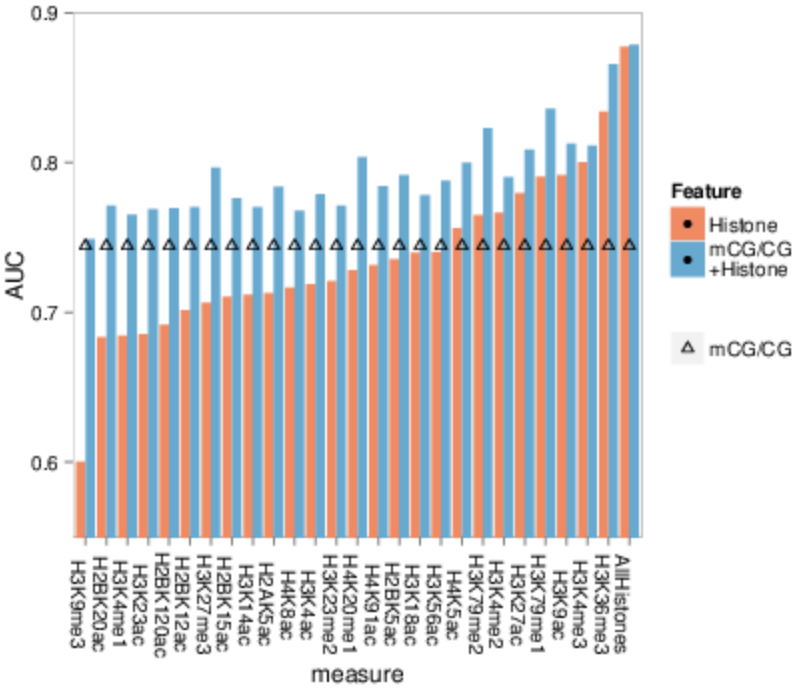


Cheng et al., *Genome Biology* 12(2):R15, (2011); Cheng et al., *Genome Research* 22(9):1658-1667, (2012)

Evidence of quantitative relationships

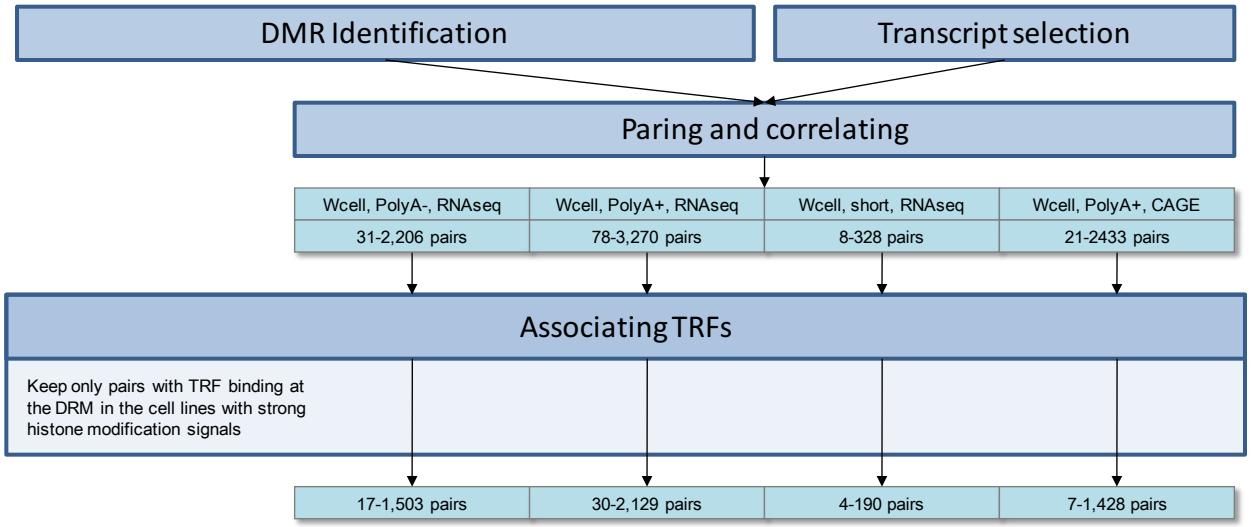
- Promoters and gene bodies: DNA methylation models

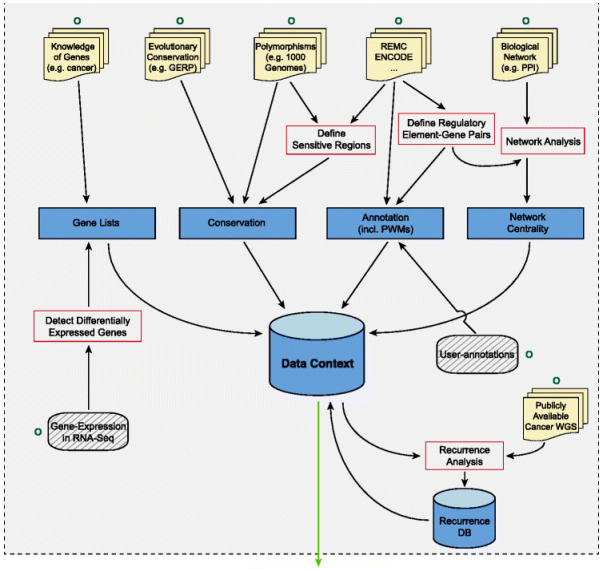


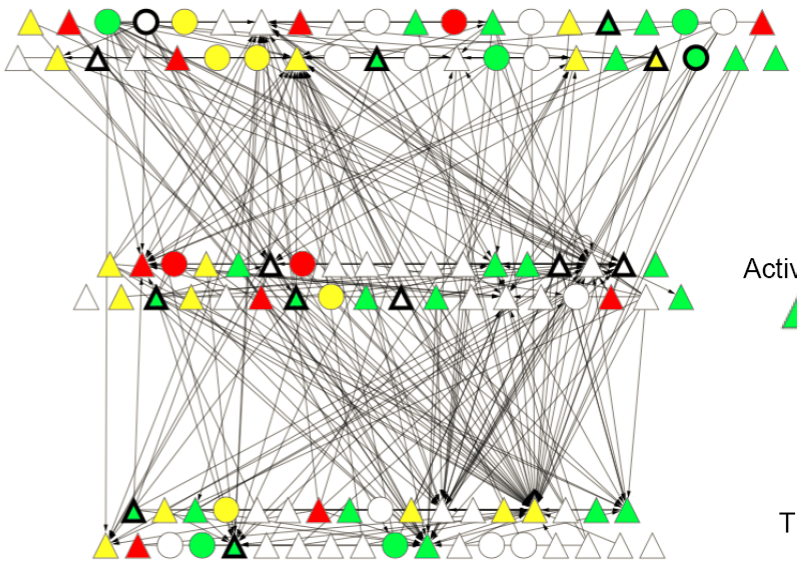


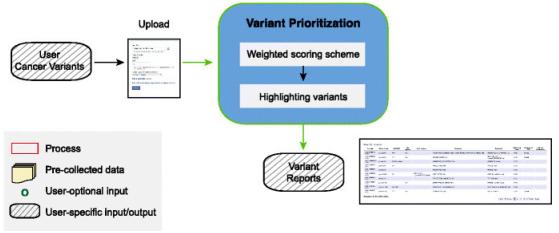
Lou, Lee et al., *Genome Biology* 15(7):408, (2014)

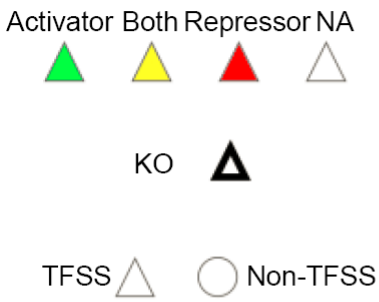
Correlation-based enhancer targets









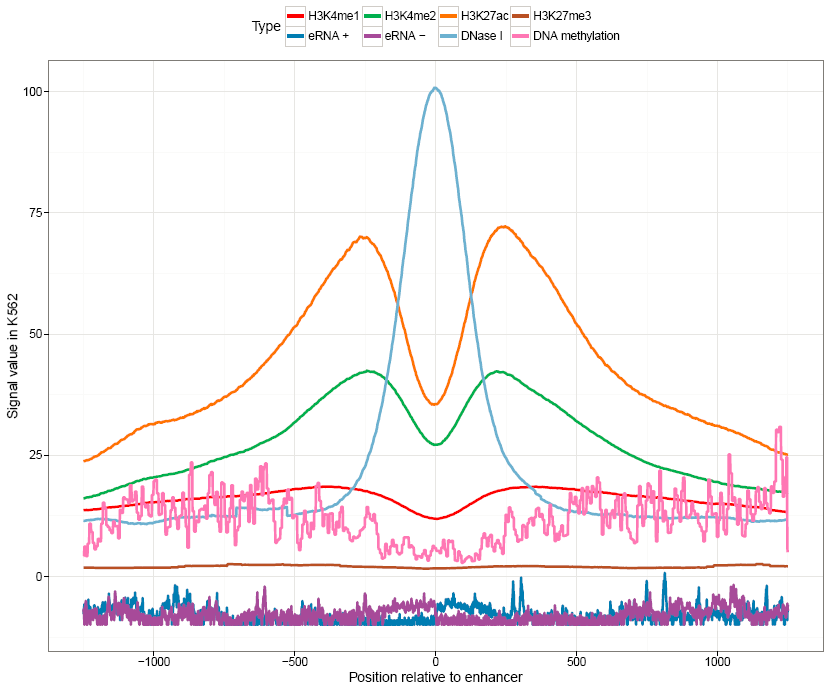


Yip et al., *Genome Biology* 13(9):R48, (2012) ; Gerstein et al., *Nature* 489(7414):91-100, (2012); Fu et al., *Genome Biology* 15(10):480, (2014)

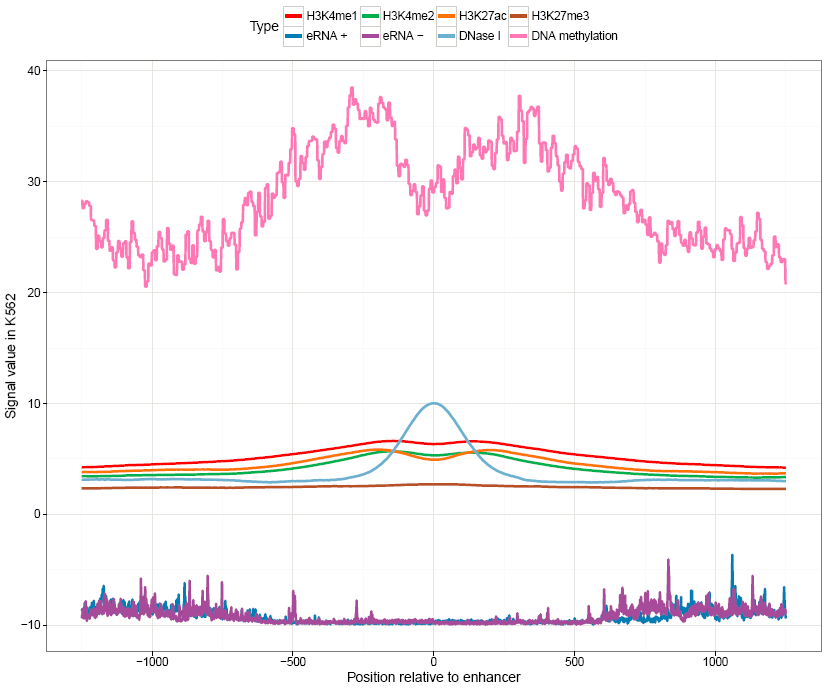
Some limitations

- Low statistical power
  - Need more samples
    - ENCODE+Roadmap: 127, ChromHMM enhancers
    - FANTOM5: 808, eRNA enhancers
  - Restricting to genomic neighborhood
- Not context-specific
  - Need to call targets in each specific context
- Considering each enhancer-gene pair separately
  - Need to consider the joint effect of multiple enhancers on the same gene
  - New method: Joint Expression Modeling of Enhancers (JEME)

FANTOM5 enhancers



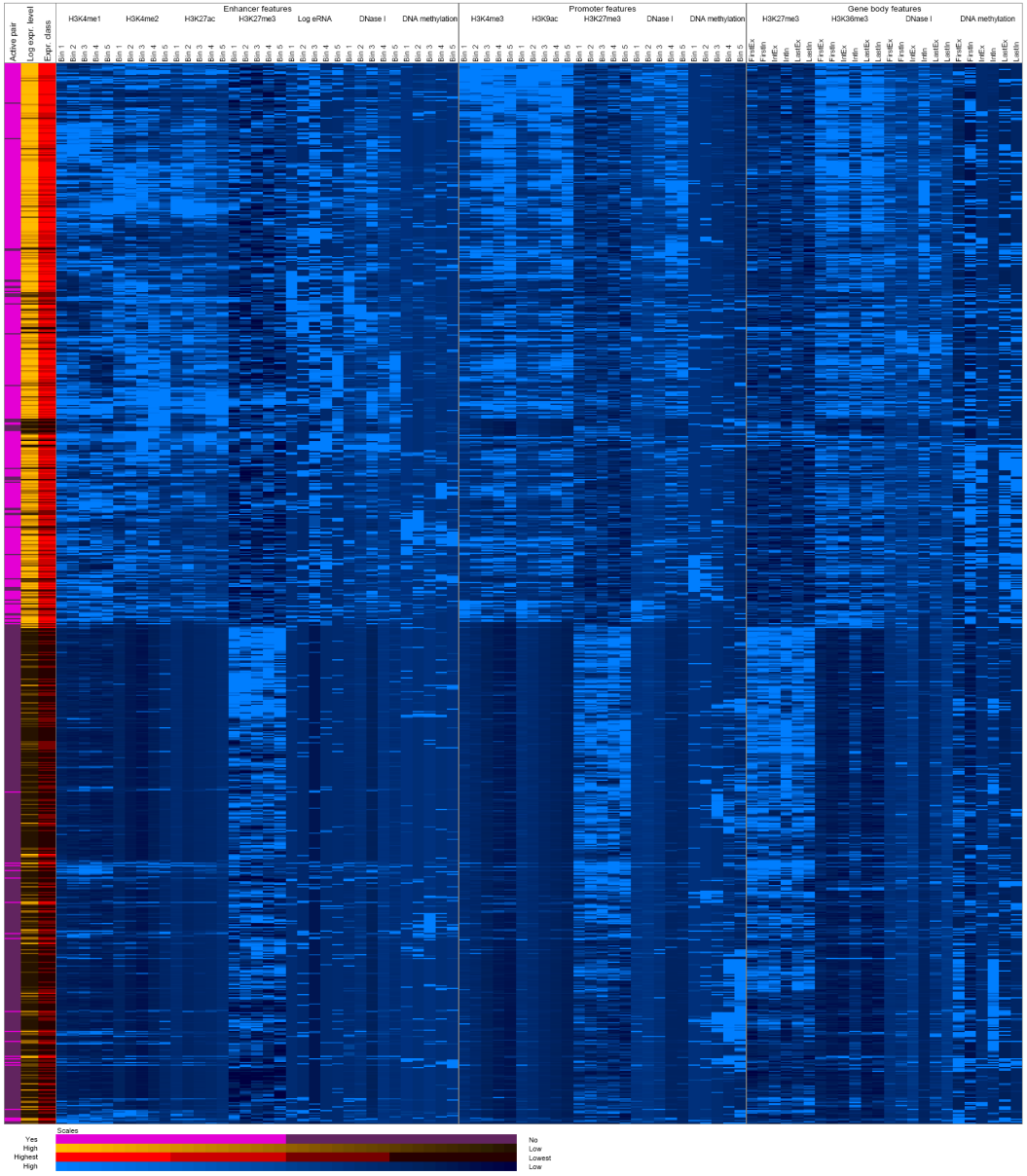
Active in K562



Inactive in K562

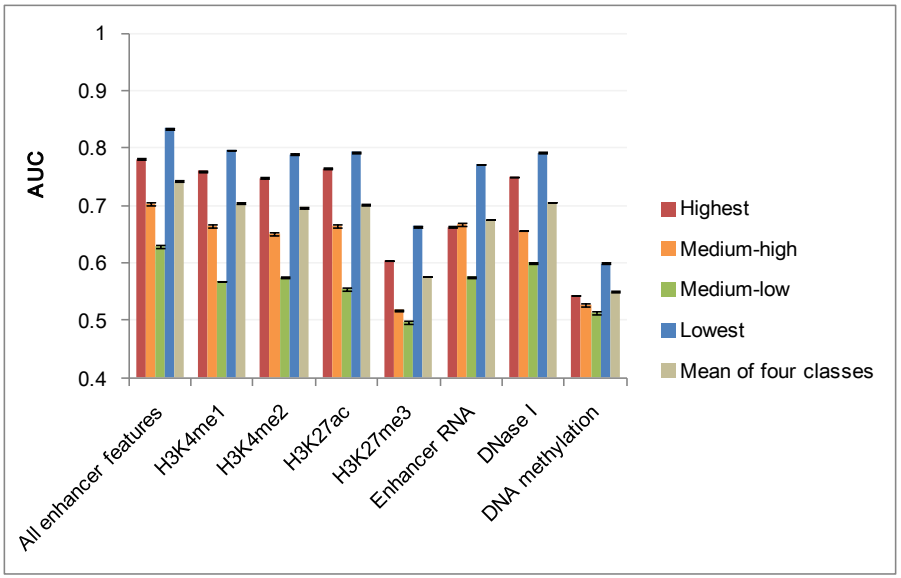
Cao et al., *(submitted)*

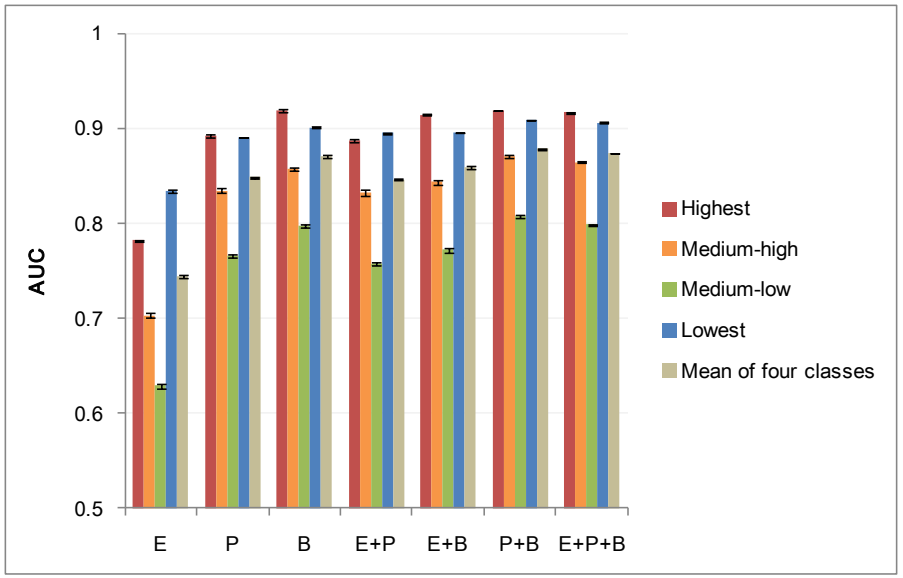
ChIA-PET connected one-to-one pairs

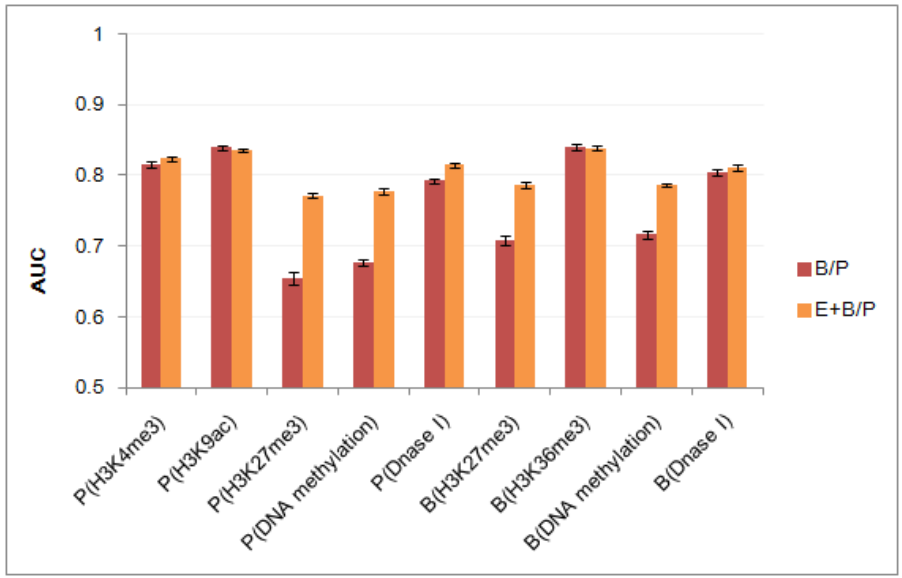


Cao et al., *(submitted)*

Accuracy of expression models



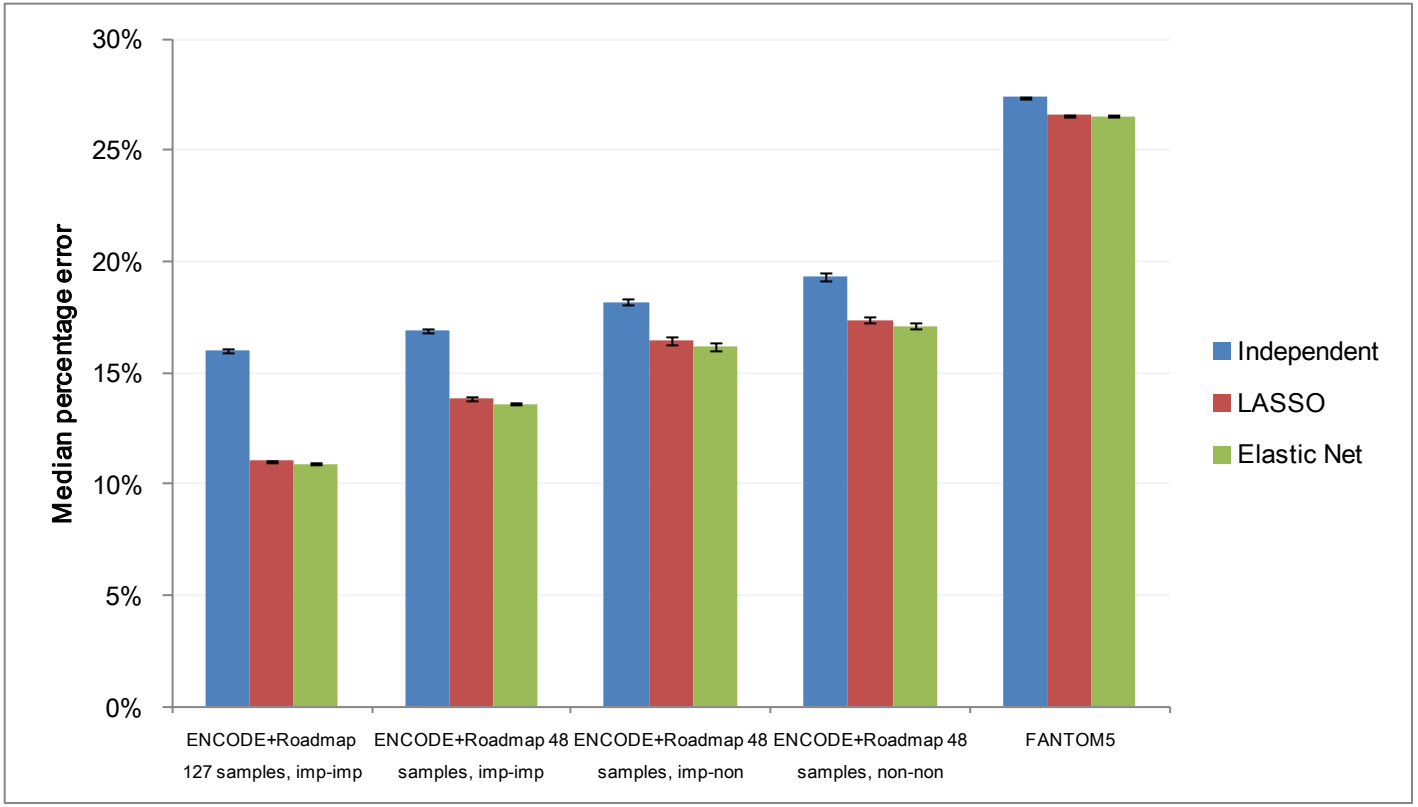




Cao et al., *(submitted)*

Modeling joint effect of multiple enhancers

- Considering all samples
- Considering all potential regulating enhancers within 1Mbp
- Modeling joint effect of multiple enhancers by LASSO or Elastic Net

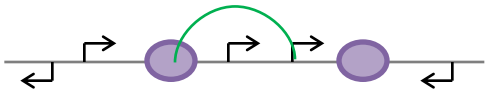


Cao et al., *(submitted)*

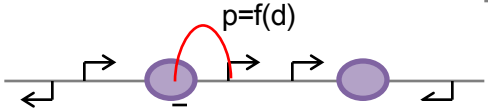
Calling sample-specific enhancer targets

1. Form global LASSO/Elastic Net model
   - E.g., $T = a_1E_1 + a_2E_2 + a_3E_3$
   - Coefficients describe general relationship between enhancer and TSSs
2. Examine amount of expression of a TSS explainable by an enhancer in a sample
3. Combine with genomic distance by a second-level model

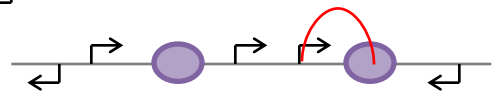
Validation of enhancer targets



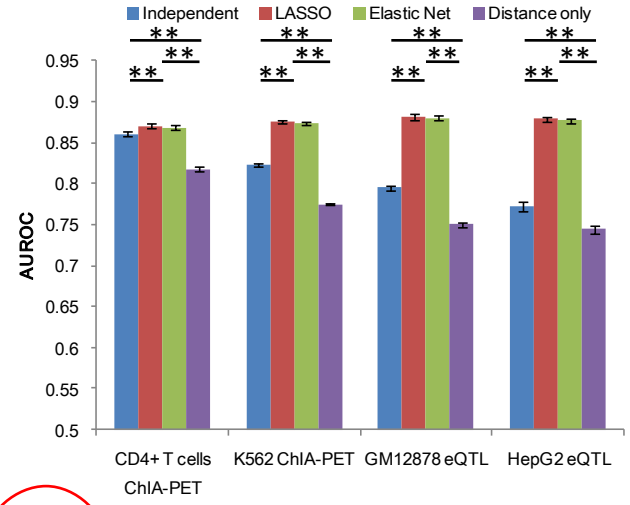
Inferred pairs

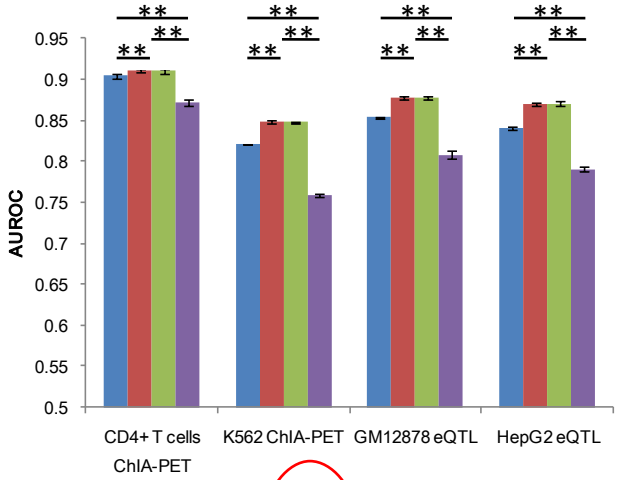


Control: Random contacts

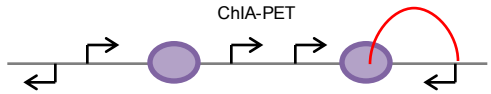


Control: Random enhancers

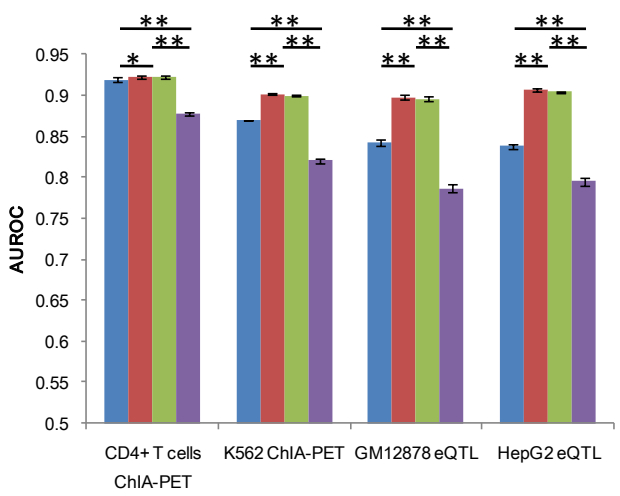


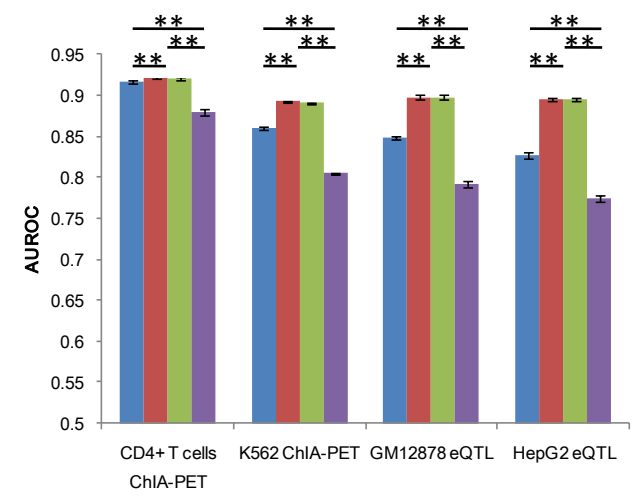


Control: Random targets



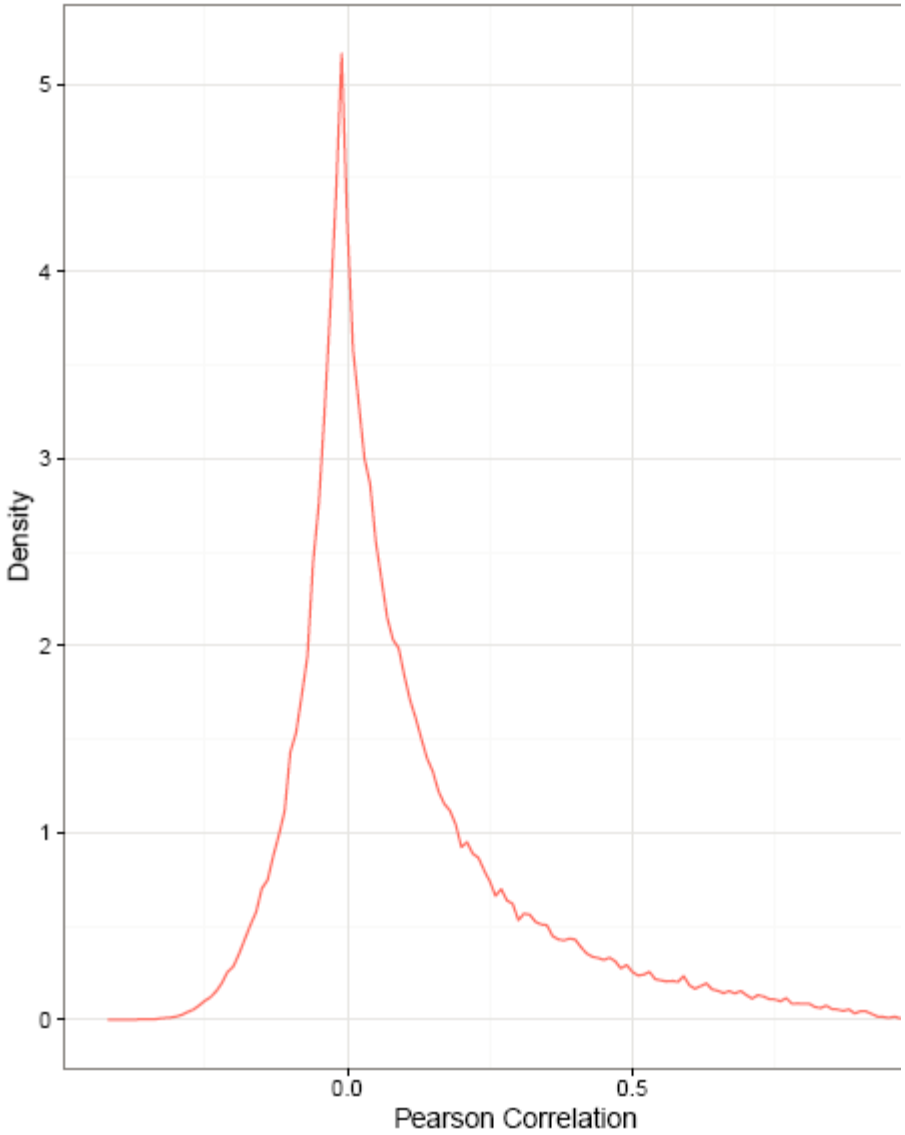
Control: Random pairs

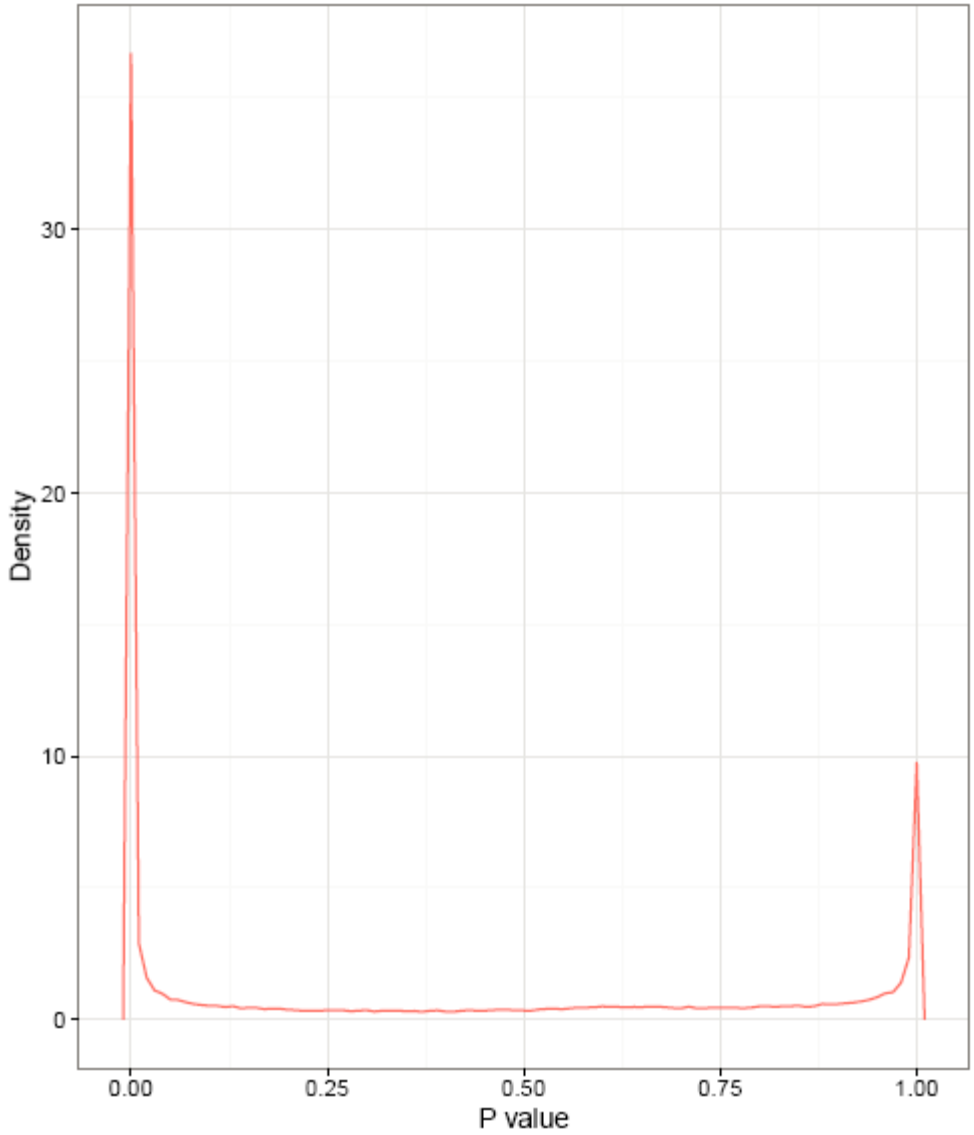




Cao et al., *(submitted)*

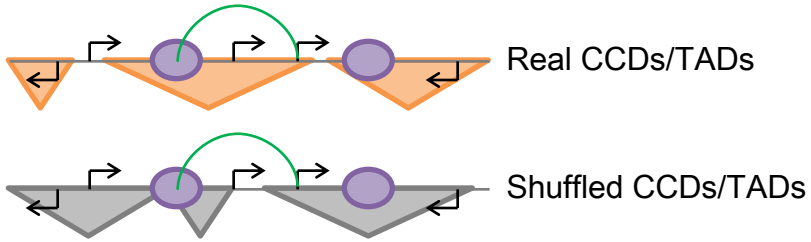
Activity correlations

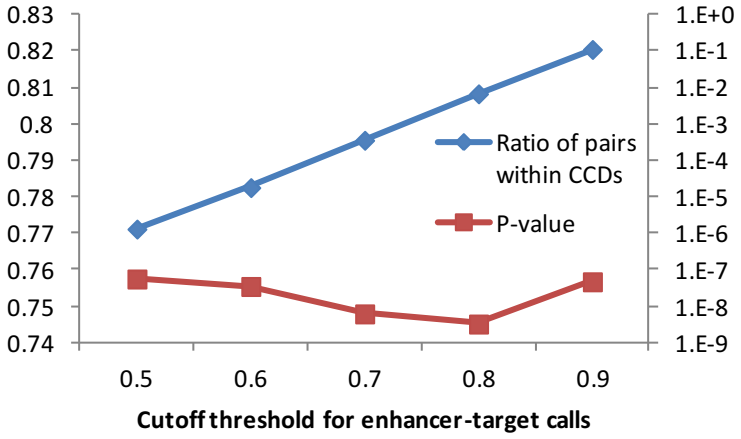


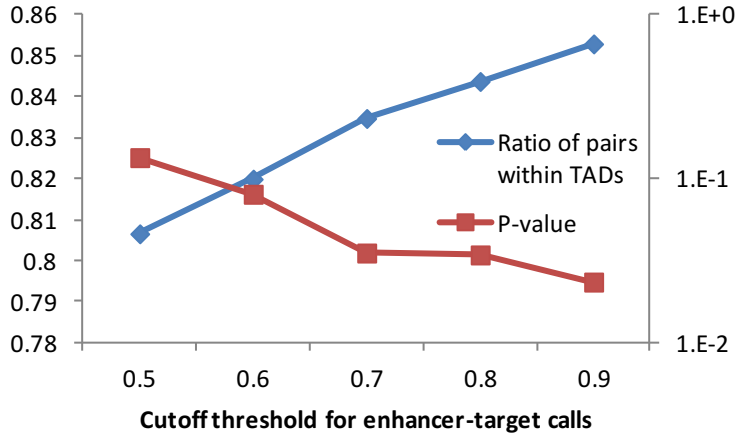


Consistency with topological domains

- Topologically associating domains (TADs) by Hi-C from Dixon et al., 2012
- Chromatin contact domains (CCDs) by CTCF ChIA-PET from Tang et al., 2015

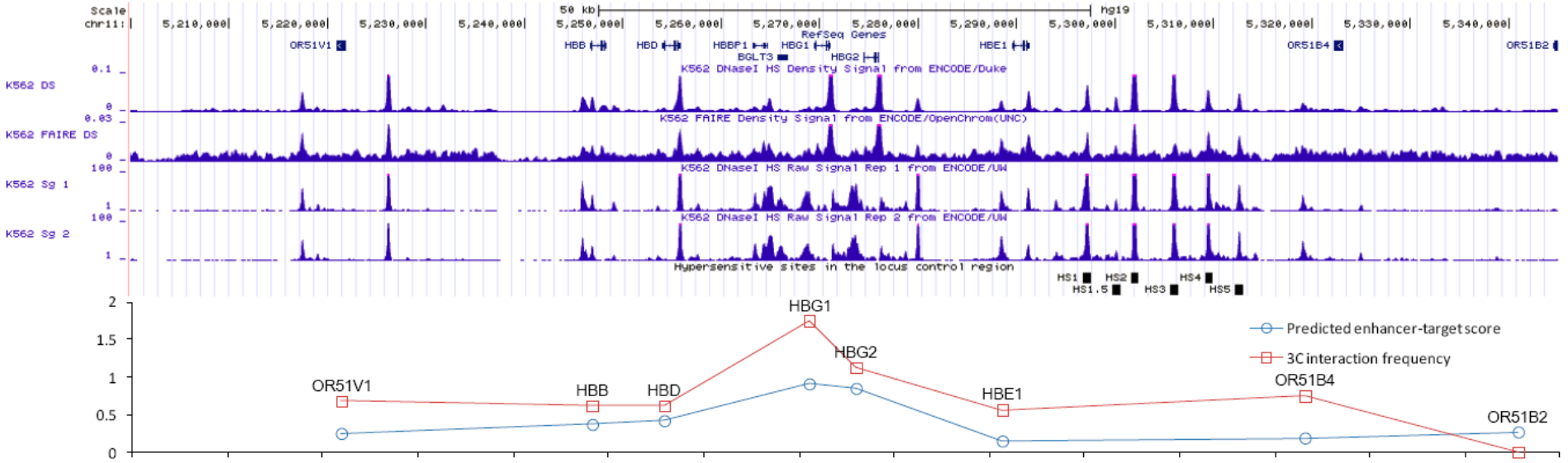




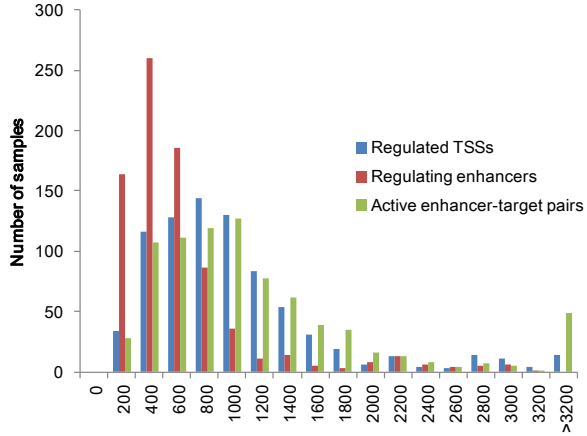


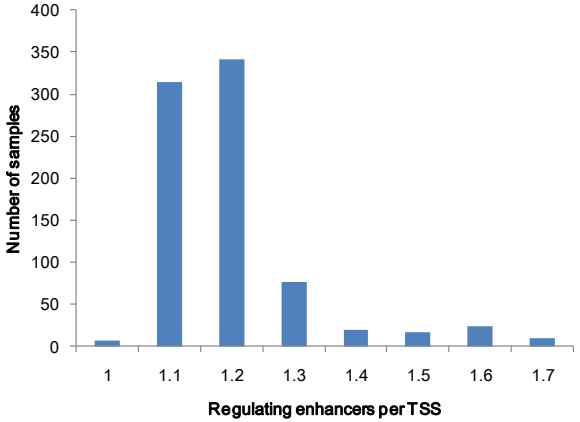
Cao et al., *(submitted)*

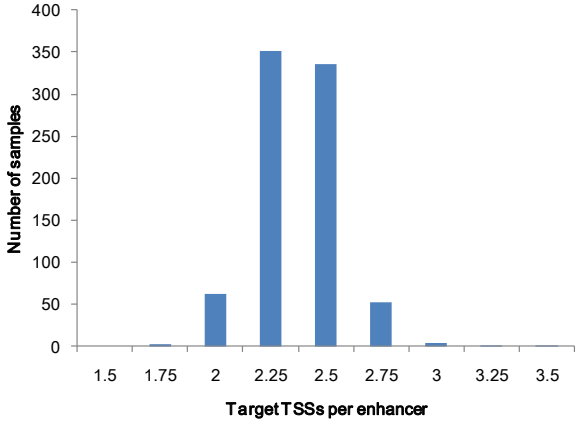
LCR of beta-globin locus

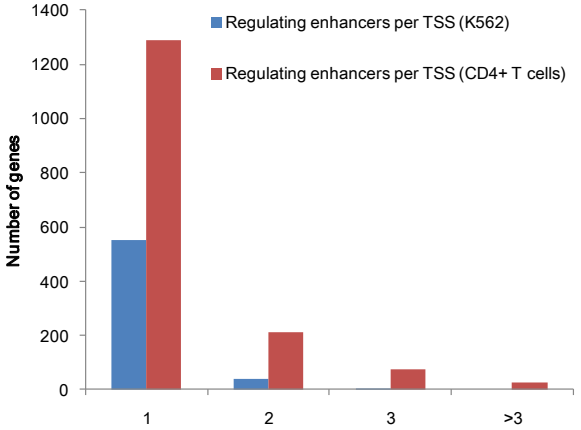


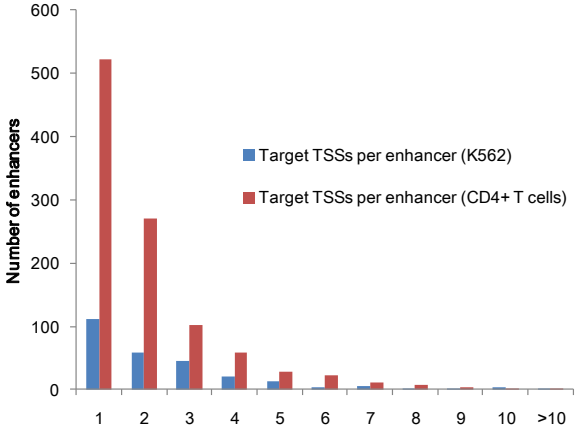
Cao et al., *(submitted)*

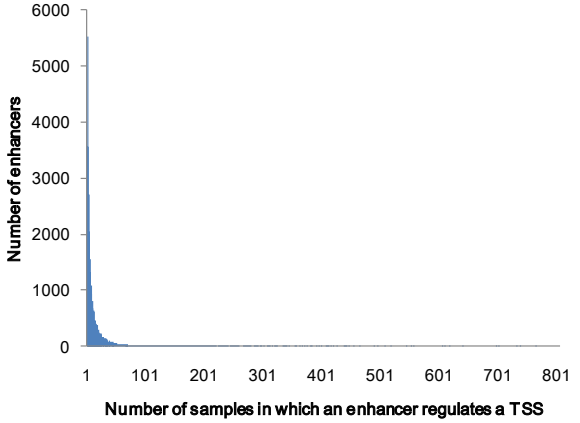


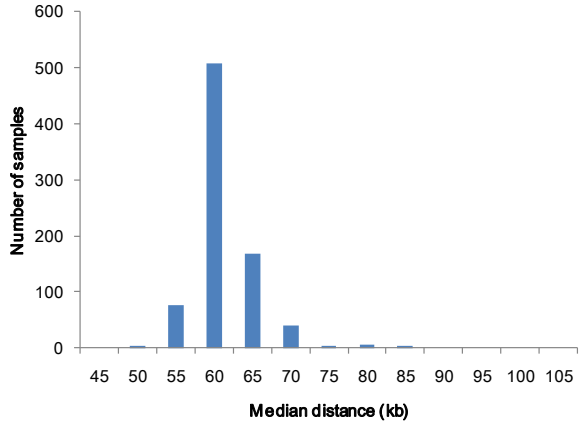


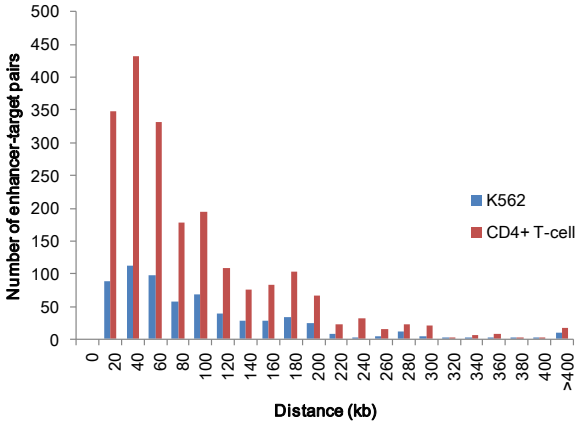


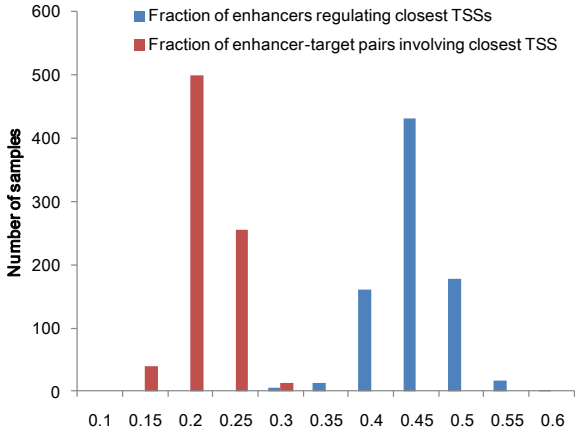






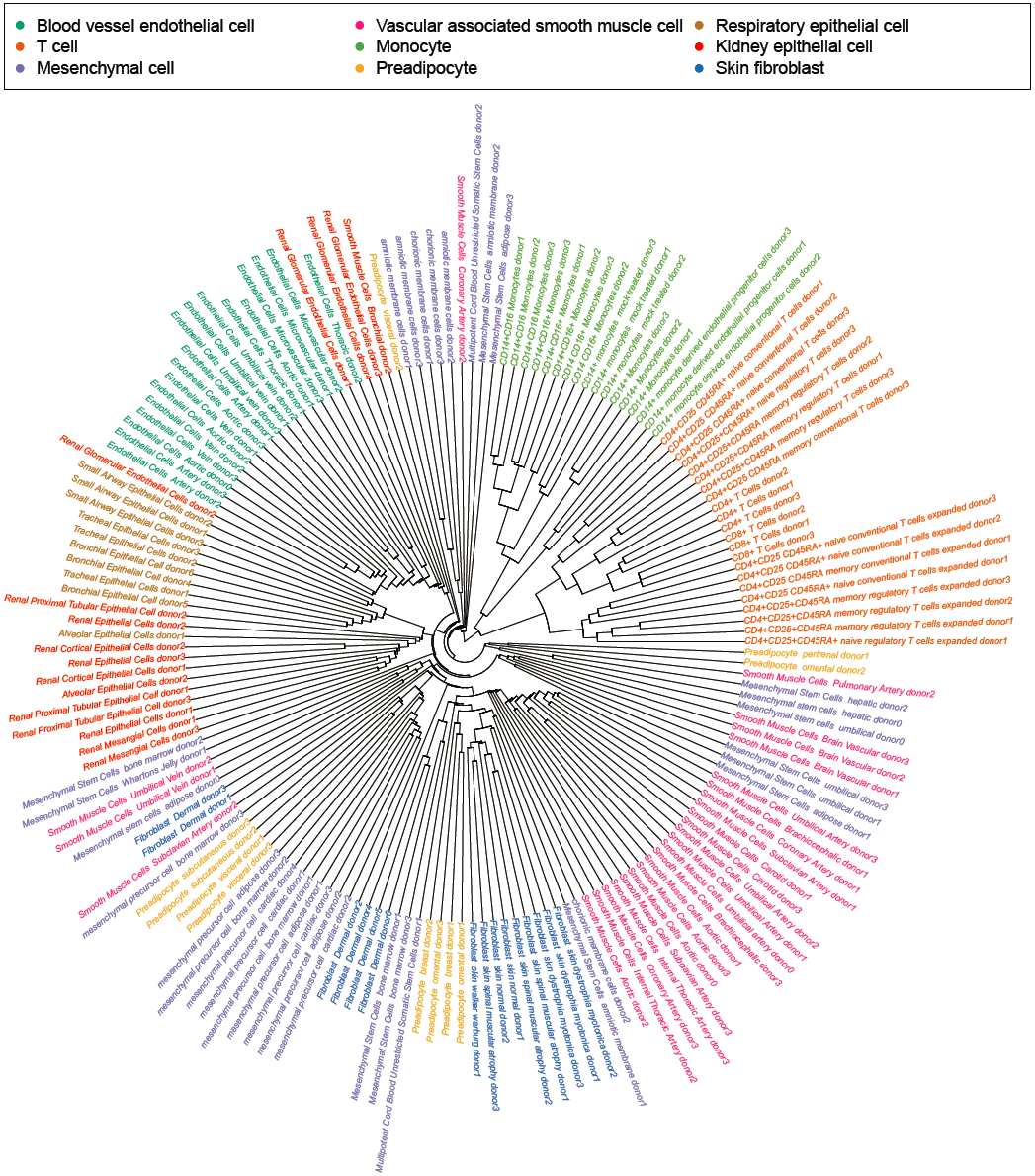






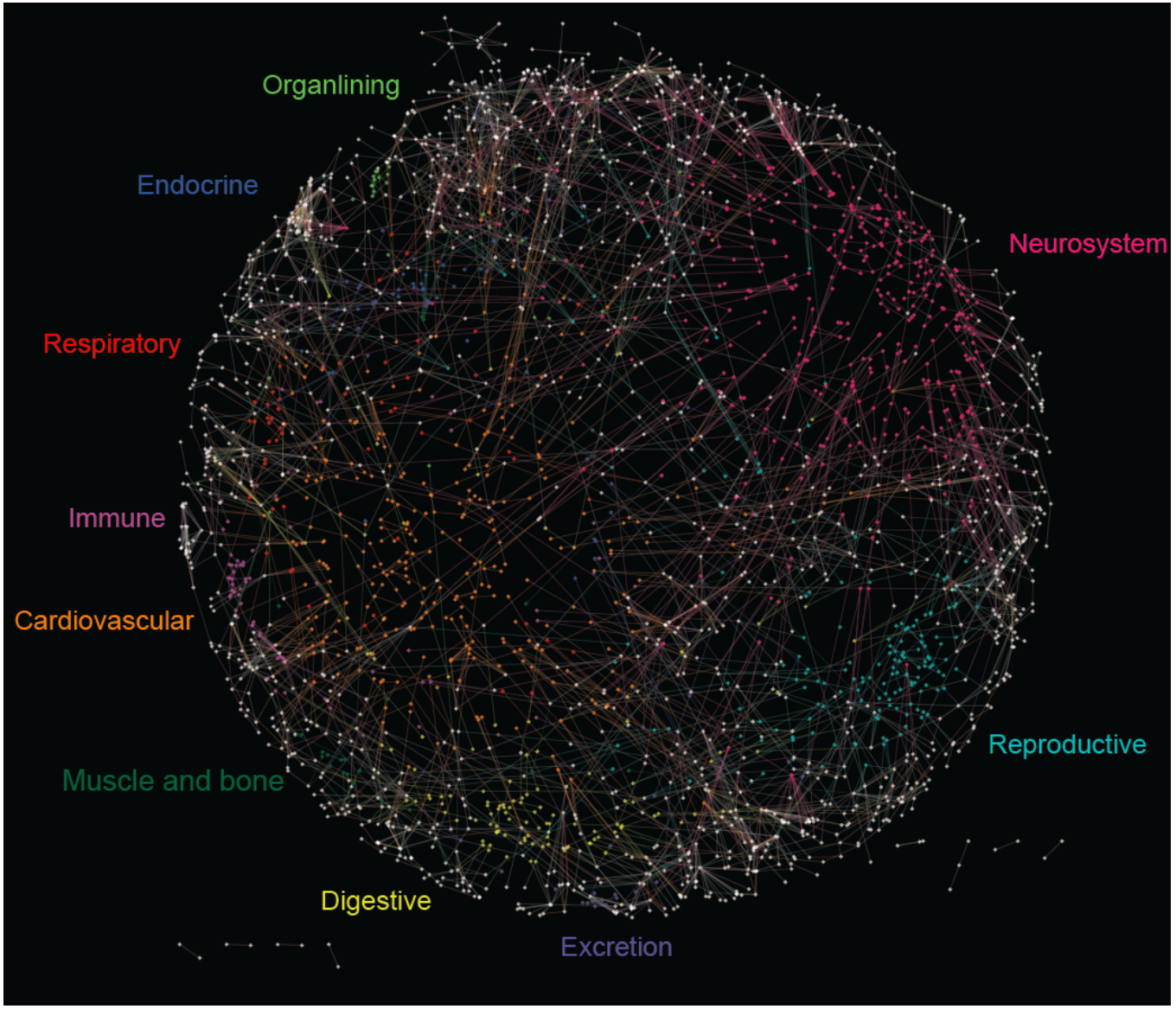
Cao et al., *(submitted)*

Enhancer network as sample signature



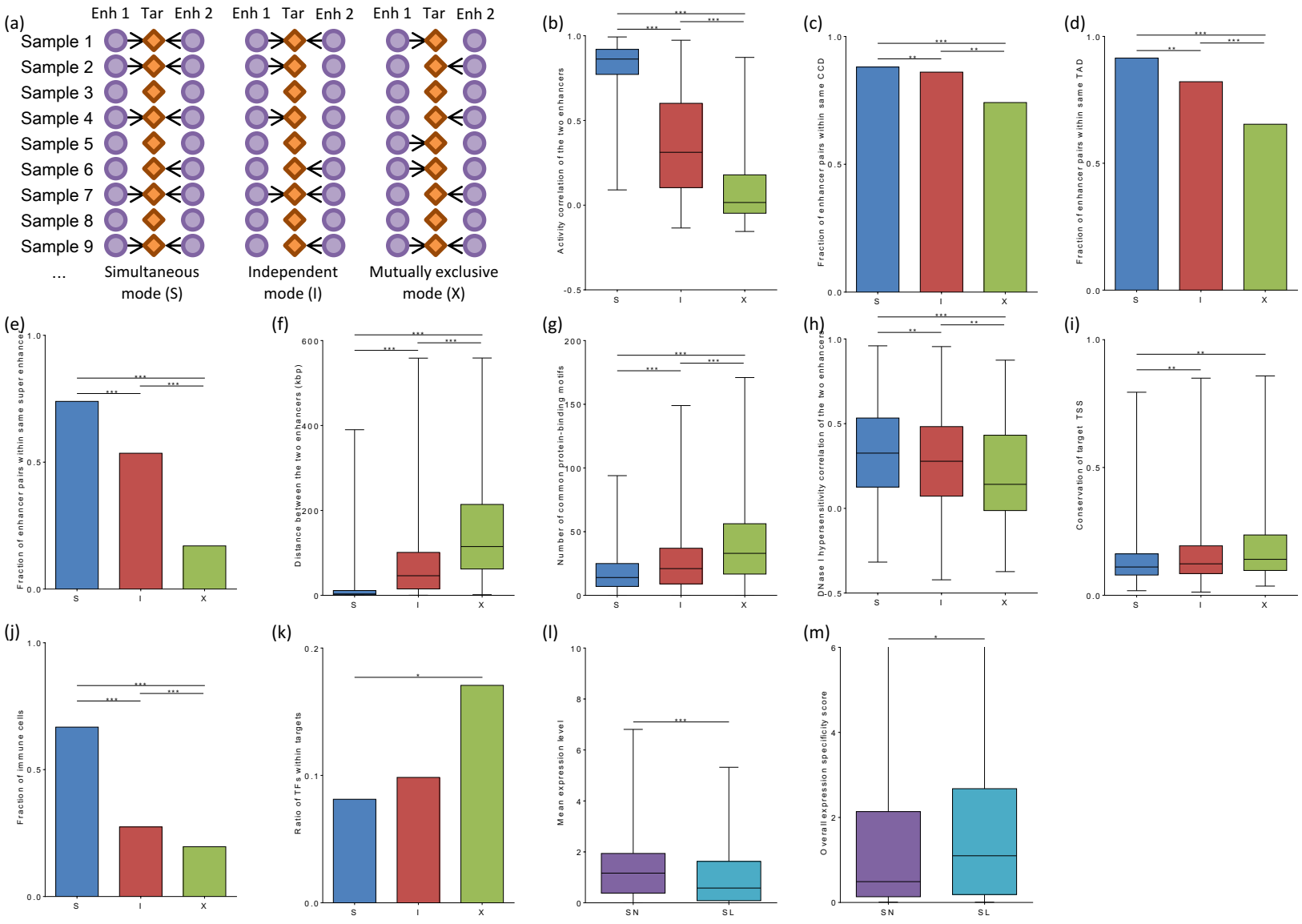
Cao et al., *(submitted)*

Sample-group specificity



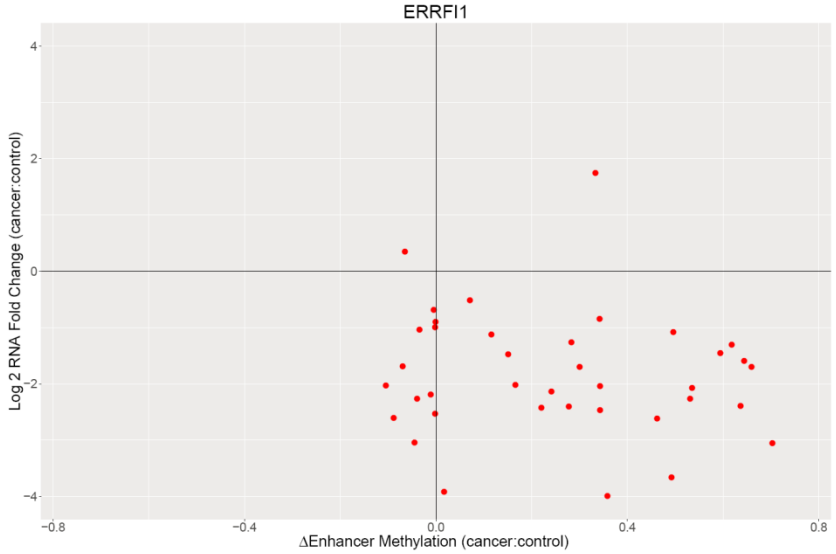
Cao et al., *(submitted)*

Co-regulation modes



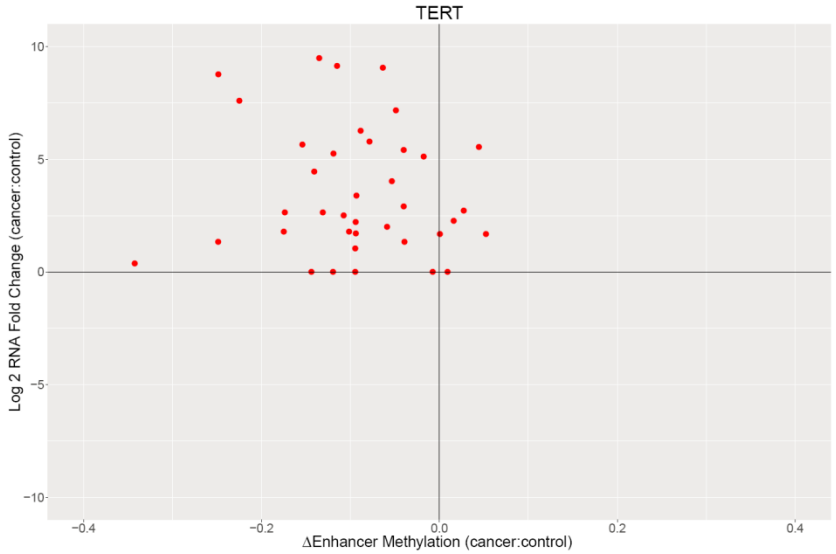
Cao et al., *(submitted)*

Differential enhancer methylation in HCC

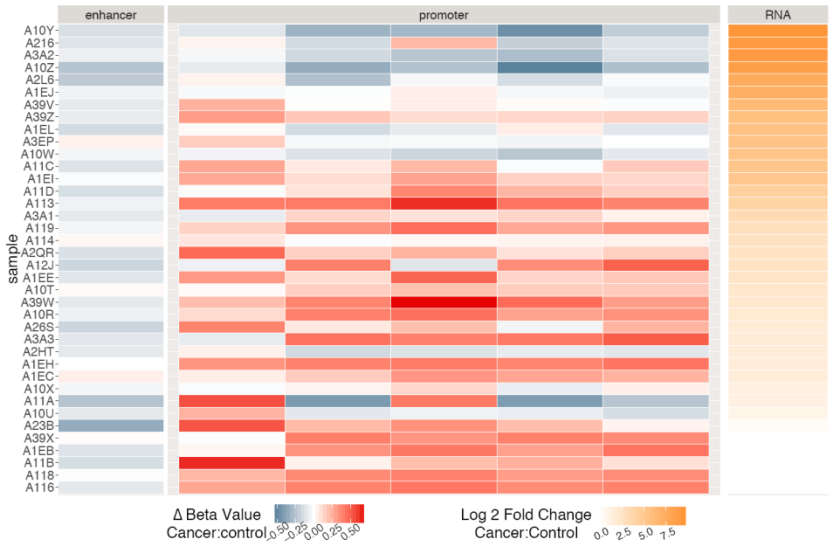


(a)

(b)



(c)



(d)

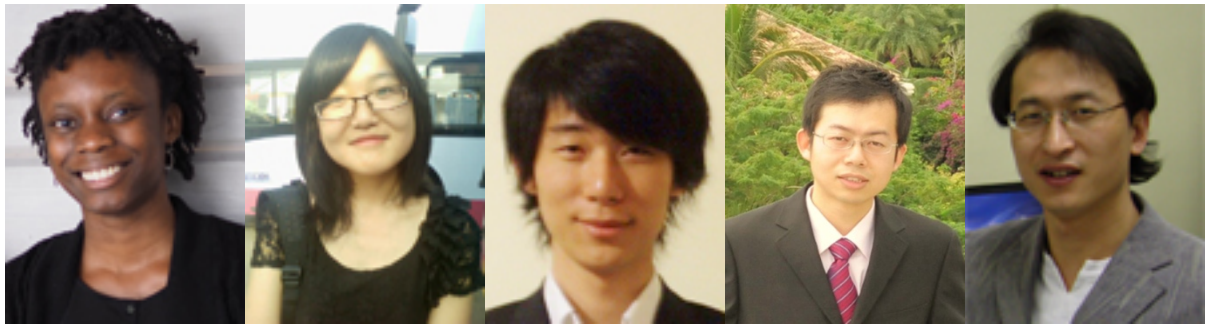
Cao et al., *(submitted)*

Data availability

- http://yiplab.cse.cuhk.edu.hk/enhancernetworks/
  - Version 2 involving more features coming
- Fields provided in the file of each sample:
  1. Enhancer location
  2. Regulated TSS and transcripts
  3. Confidence score (0-1)
  4. Activity correlations (with enhancer activity quantified by H3K4me1, H3K27me3, H3K27ac, eRNA and combined value)

Acknowledgments

- ENCODE Consortium
- Chao Cheng, Yao Fu, Ekta Khurana
- Members involved:



Christine Anyansi (Intern) | Qin Cao (PhD student) | Landon Chan (RA alumnus) | Xihao Hu (PhD alumnus) | Shaoke Lou (Postdoc alumnus)