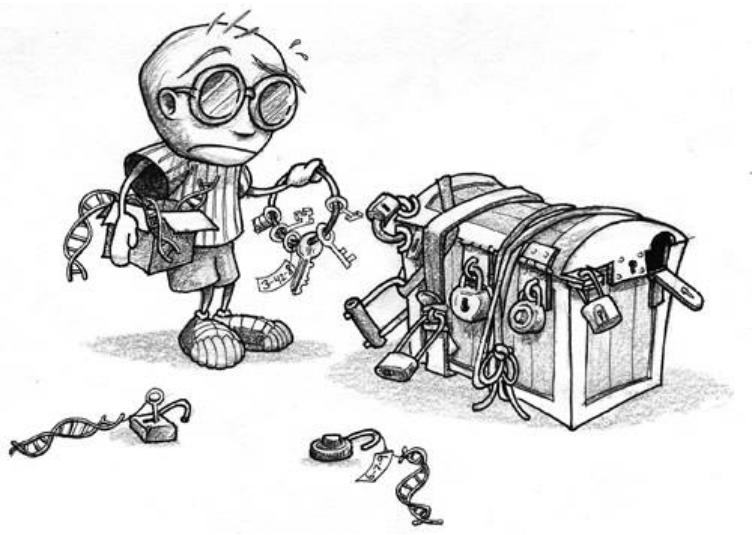


RNA Pol II modeling predicts active transcription factors.

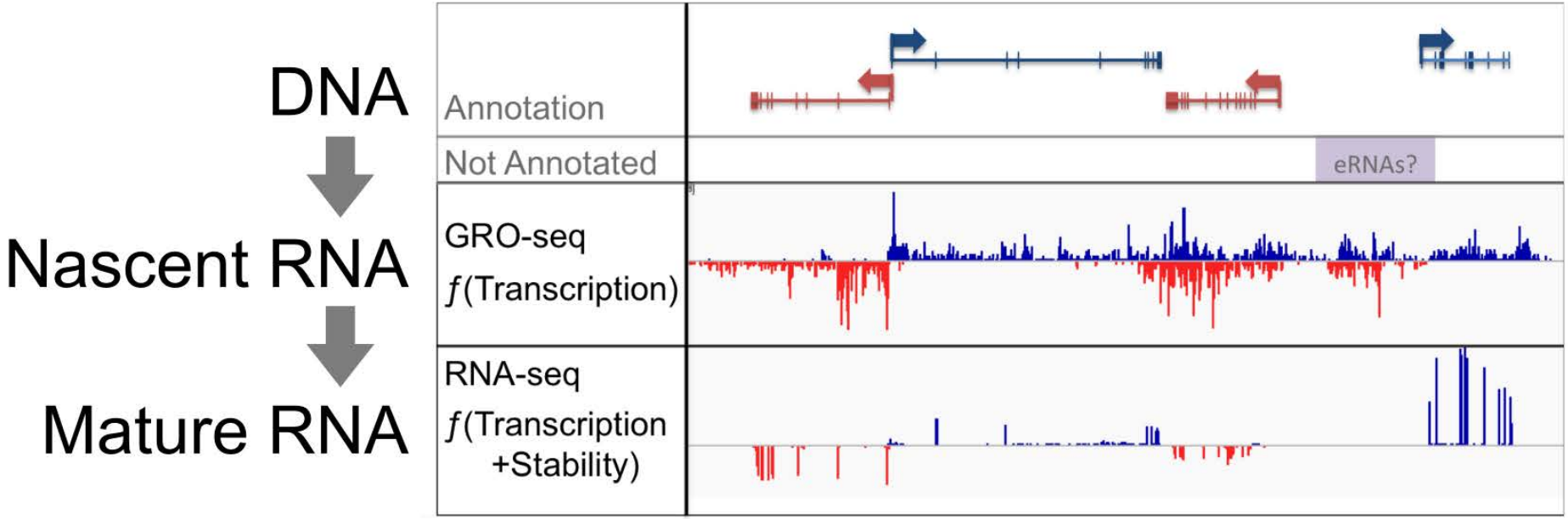
Robin Dowell

BioFrontiers Institute

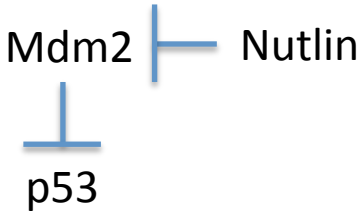
Molecular, Cellular and Developmental Biology

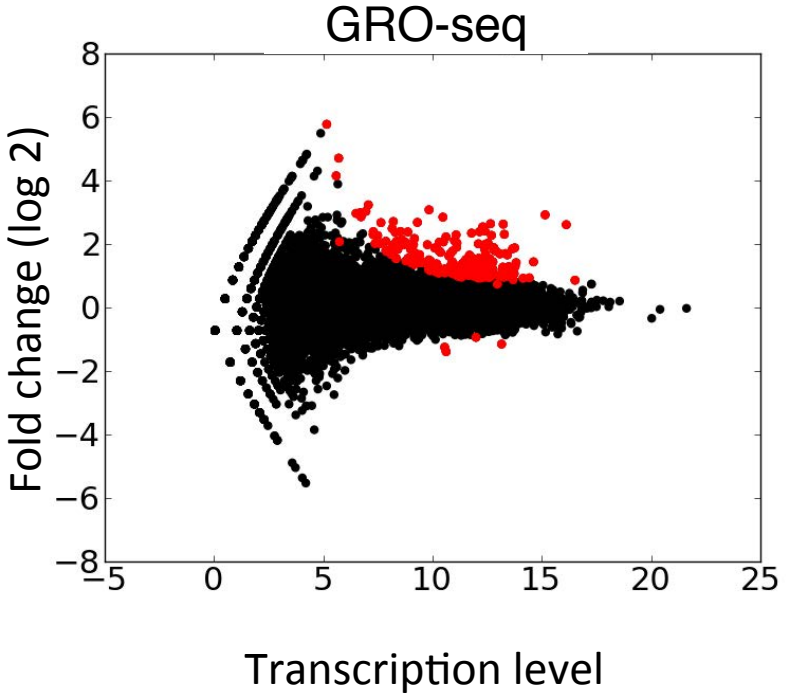
University of Colorado

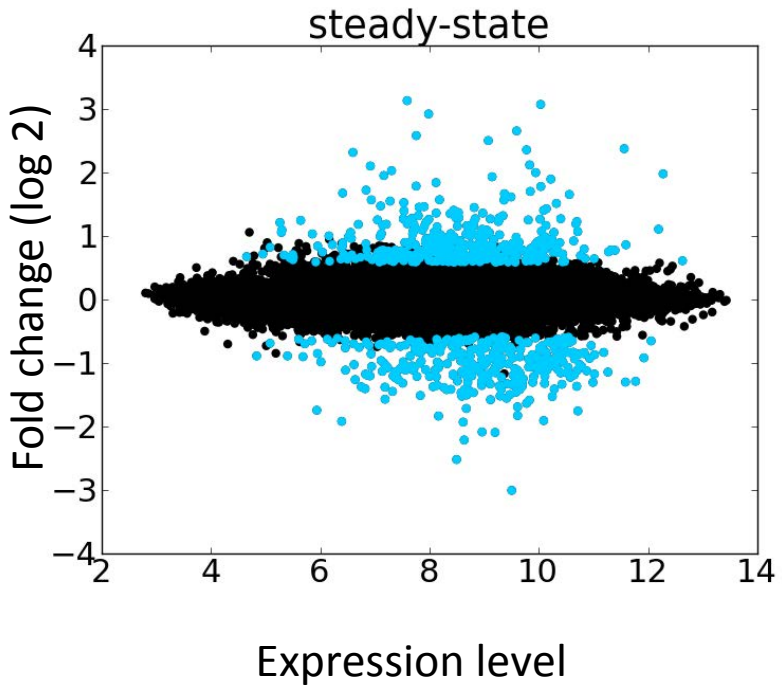
Nascent transcription is an excellent method for examining enhancer associated transcription.



Nascent transcription shows p53 is an activator

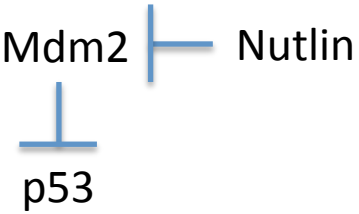


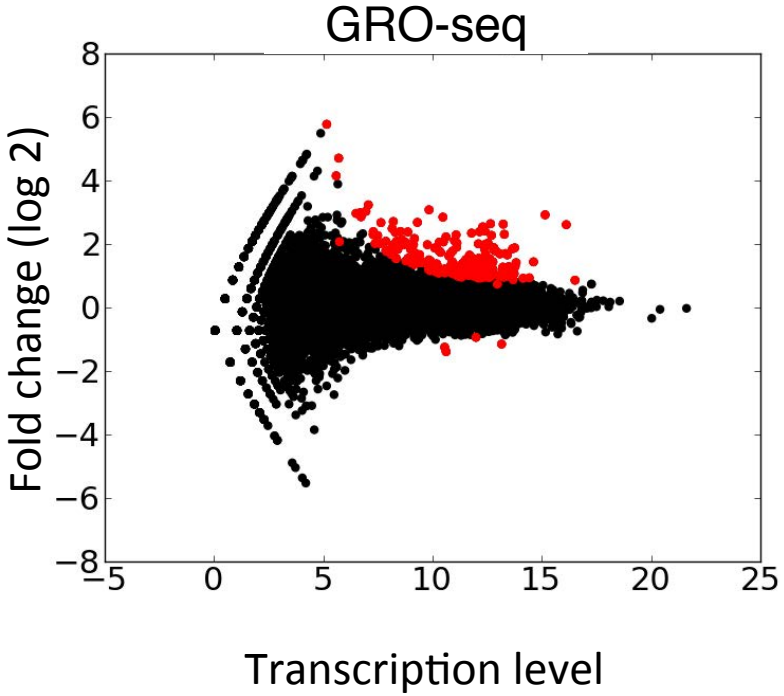


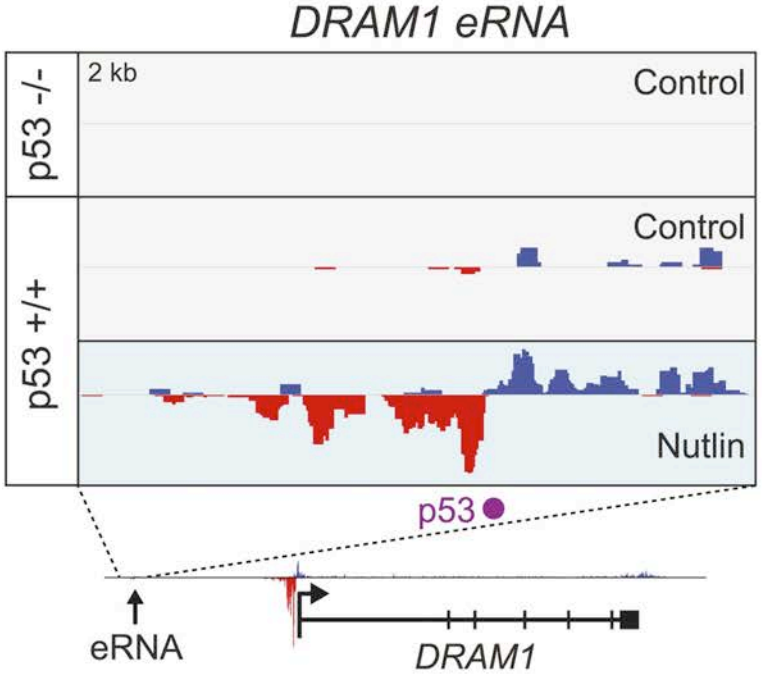


Allen et. al., Global analysis of p53-regulated transcription identifies its direct targets and unexpected regulatory mechanisms. eLife, 3, 2014

Nascent transcription shows p53 is an activator

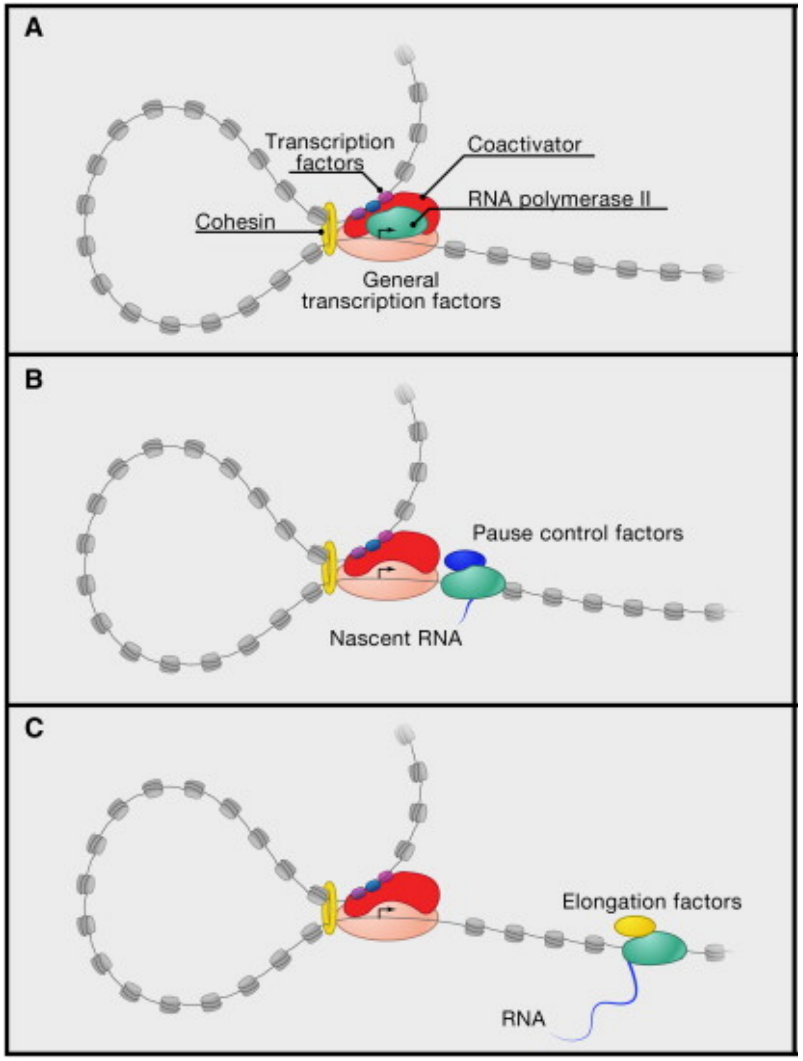




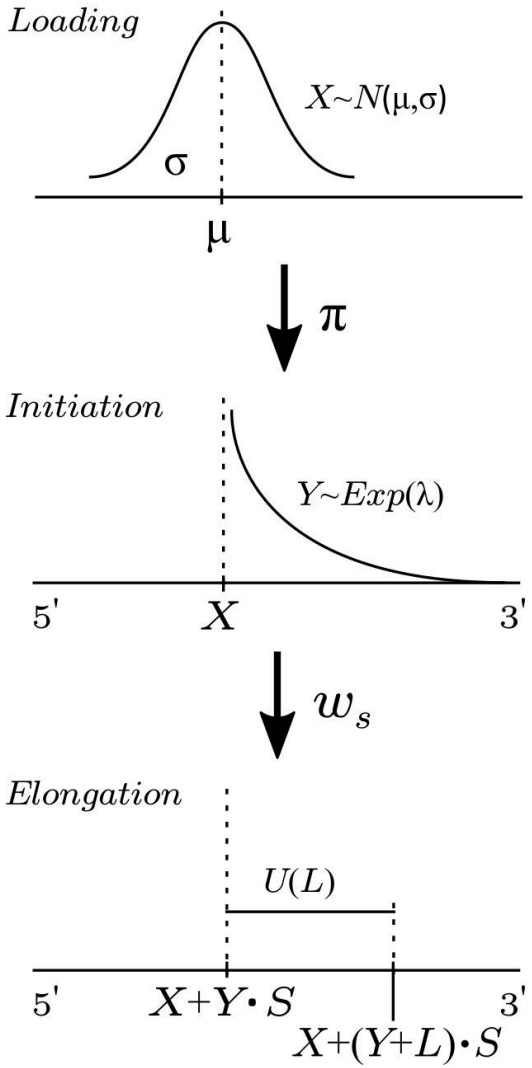


Allen et. al., Global analysis of p53-regulated transcription identifies its direct targets and unexpected regulatory mechanisms. eLife, 3, 2014

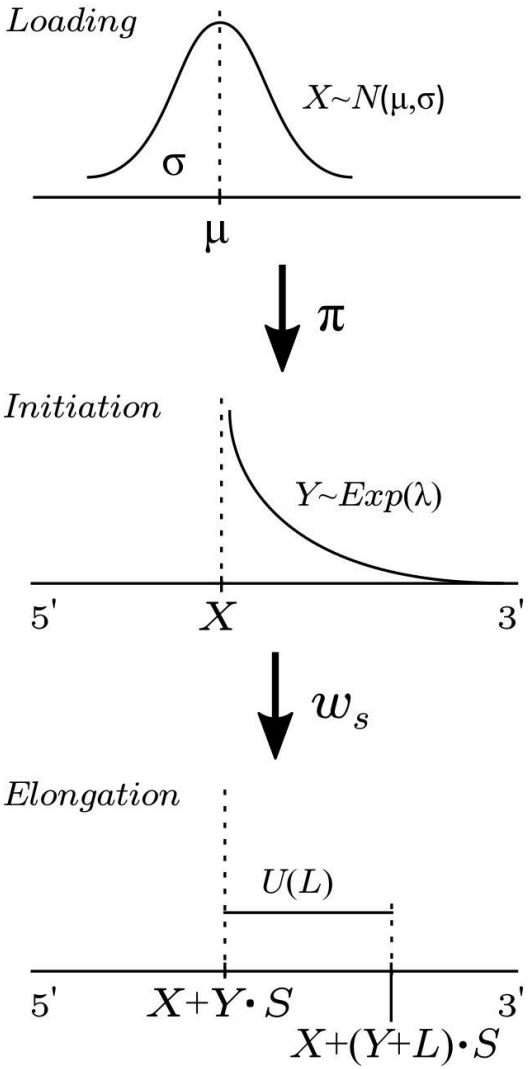
A mathematical model of polymerase behavior

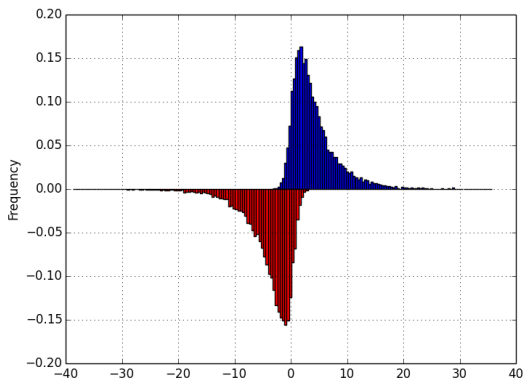


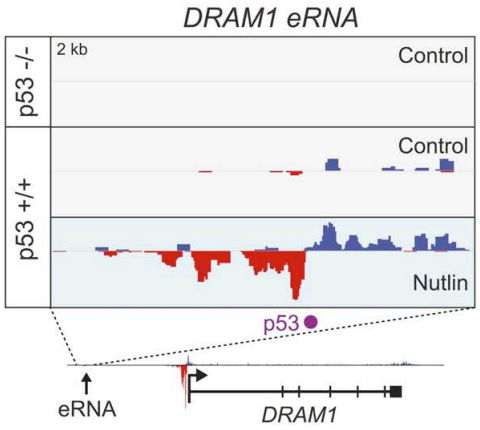
Lee and Young Cell 2013

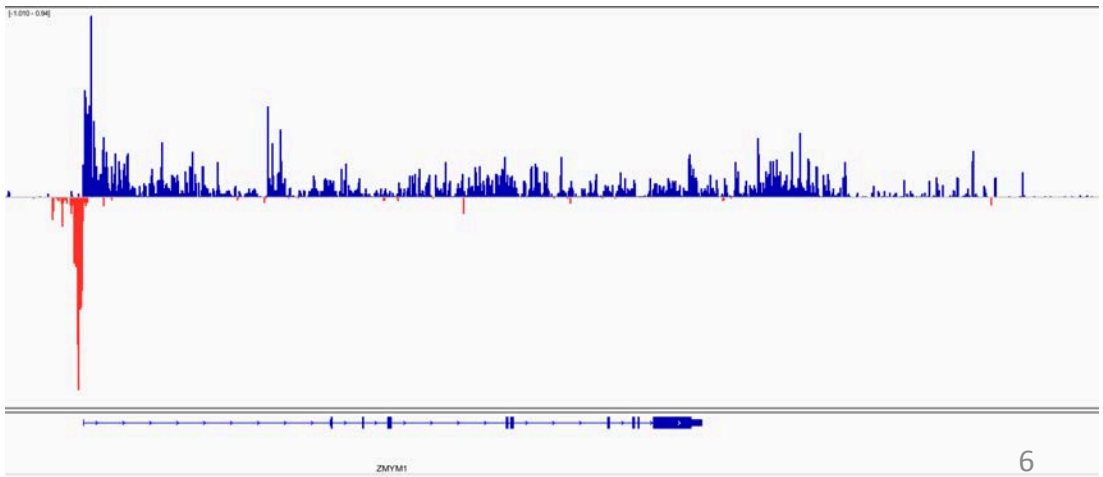


The model describes patterns observed within nascent transcription.

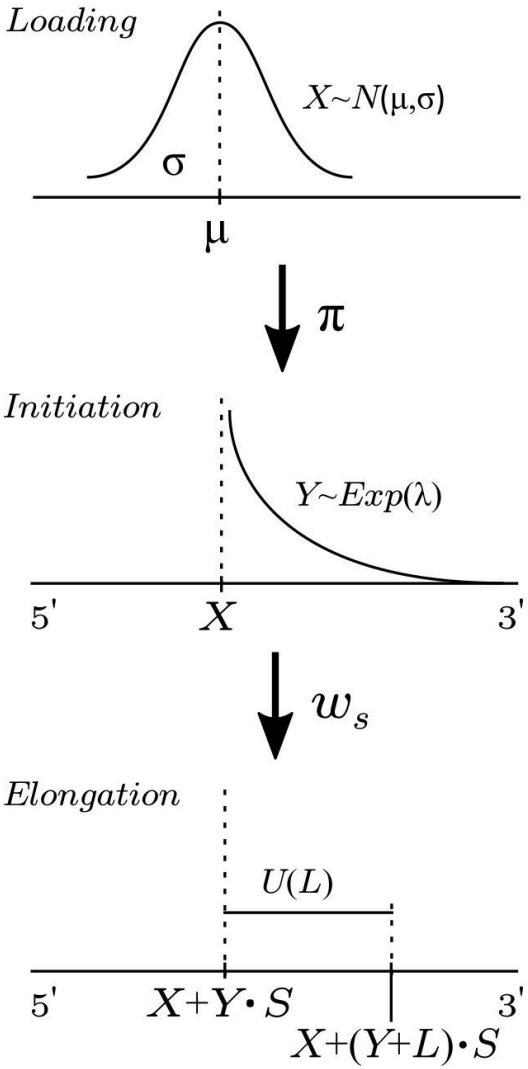


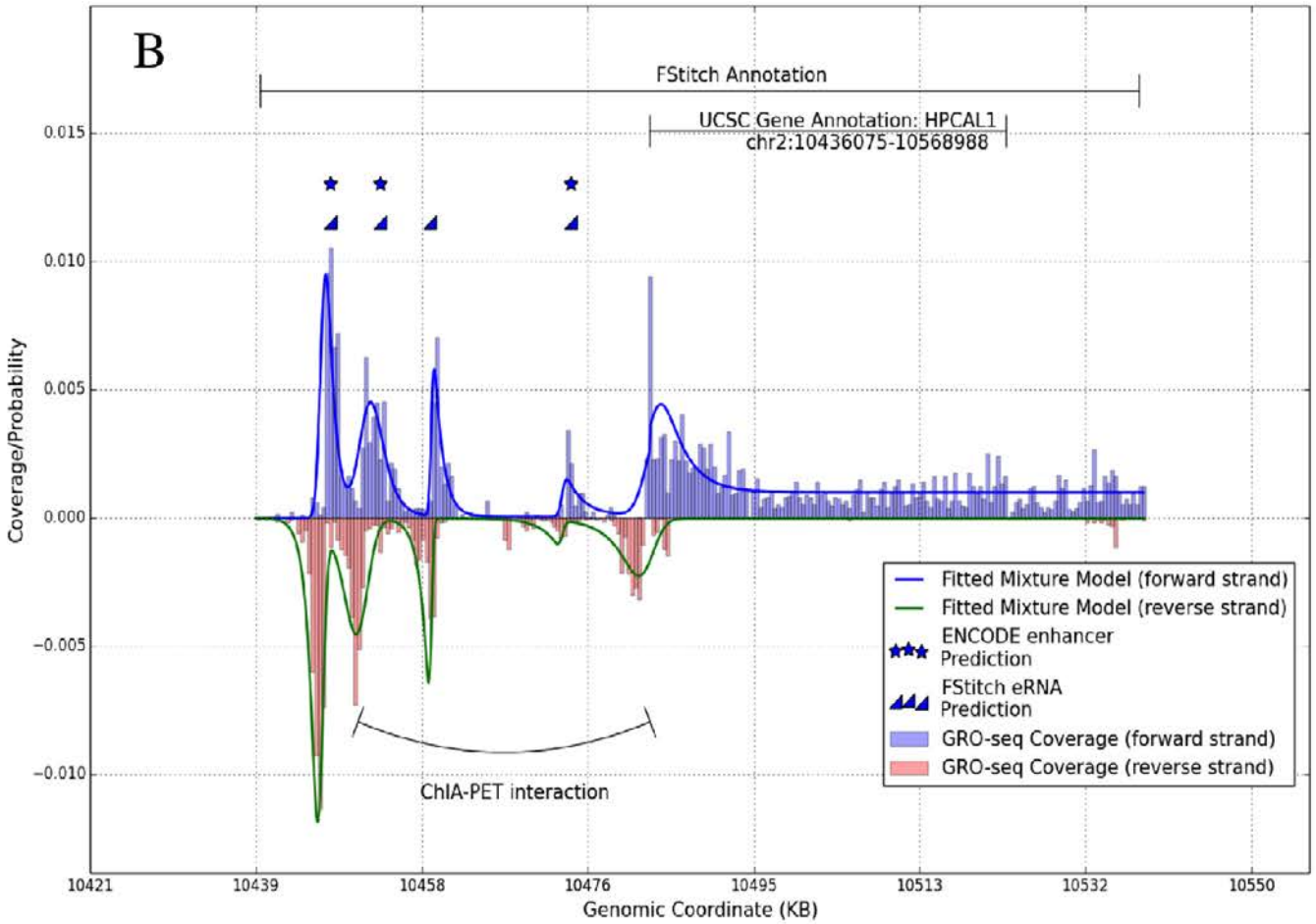






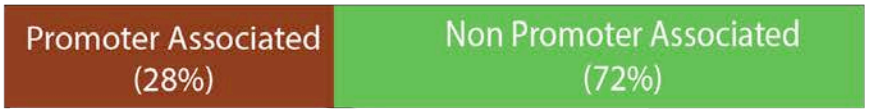
We can fit the model to nascent transcription data.





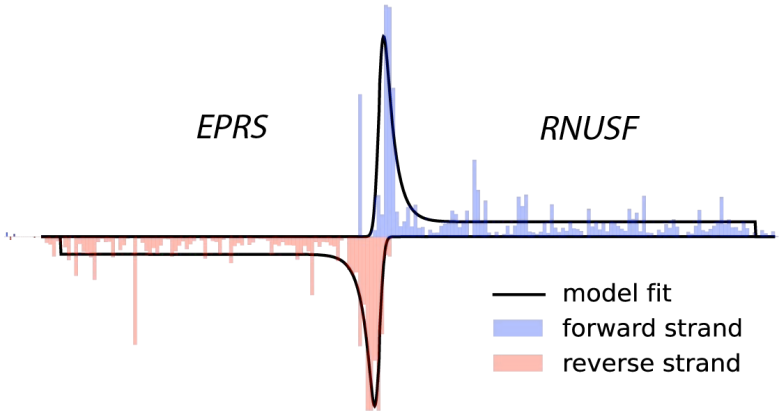
FStitch: Azofeifa et. al. 2016; ENCODE; ChIA: Li et al. 2013

Most bidirectionals are not at protein coding genes

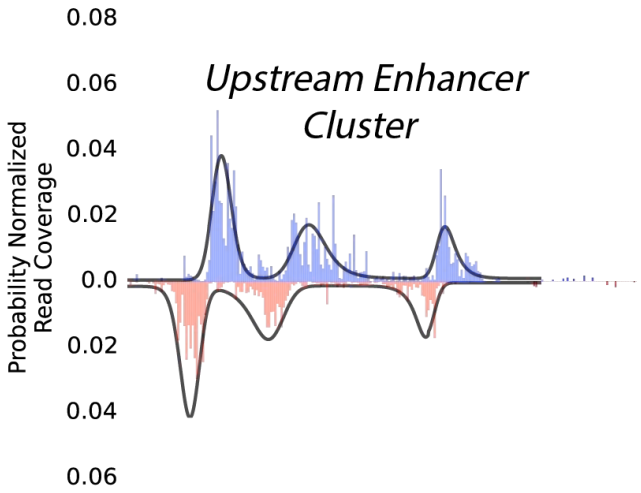


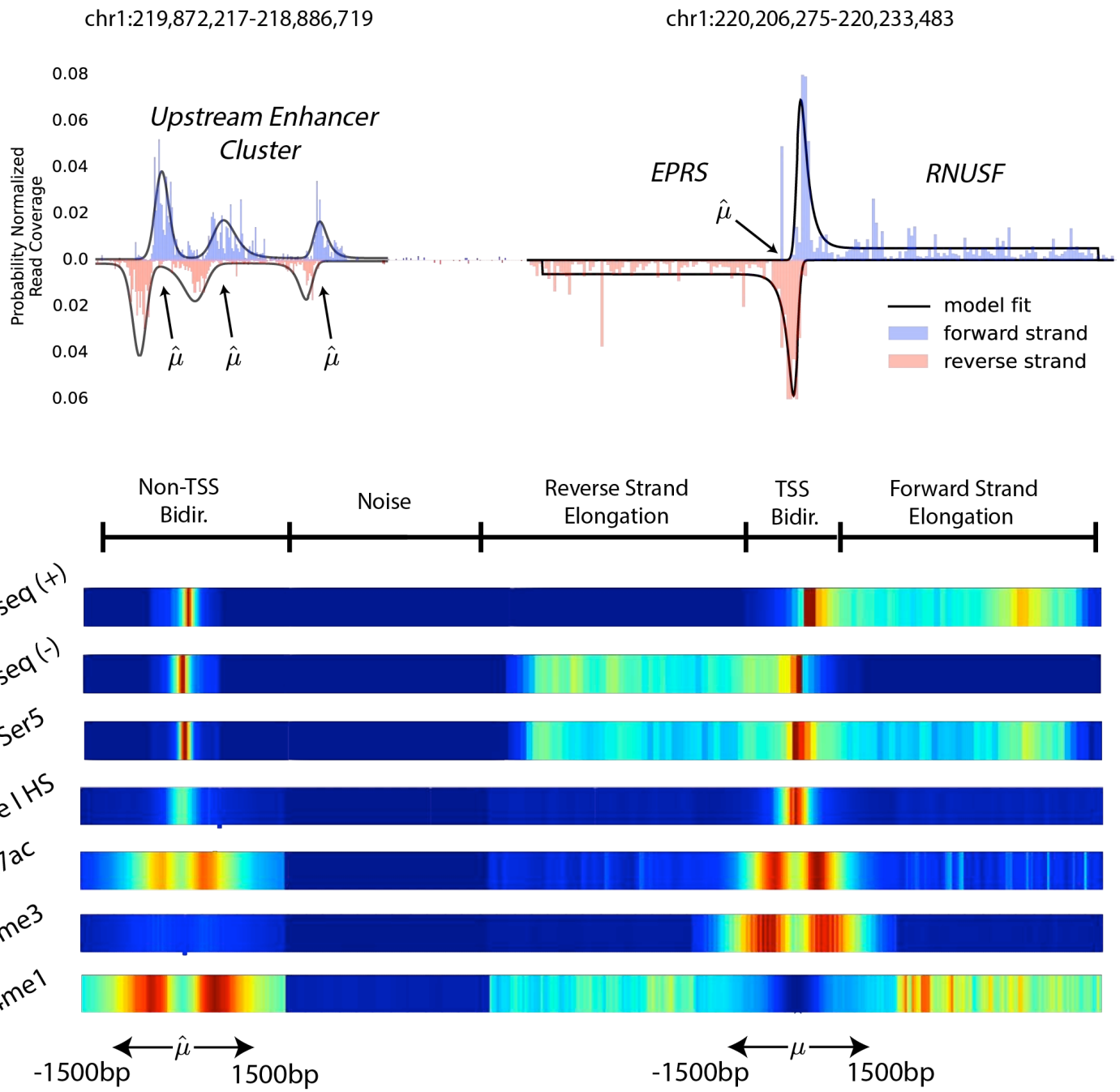
N=32,1298
(Total Bidirectional
Model Fits)

chr1:220,206,275-220,233,483

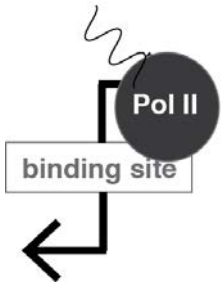


chr1:219,872,217-218,886,719

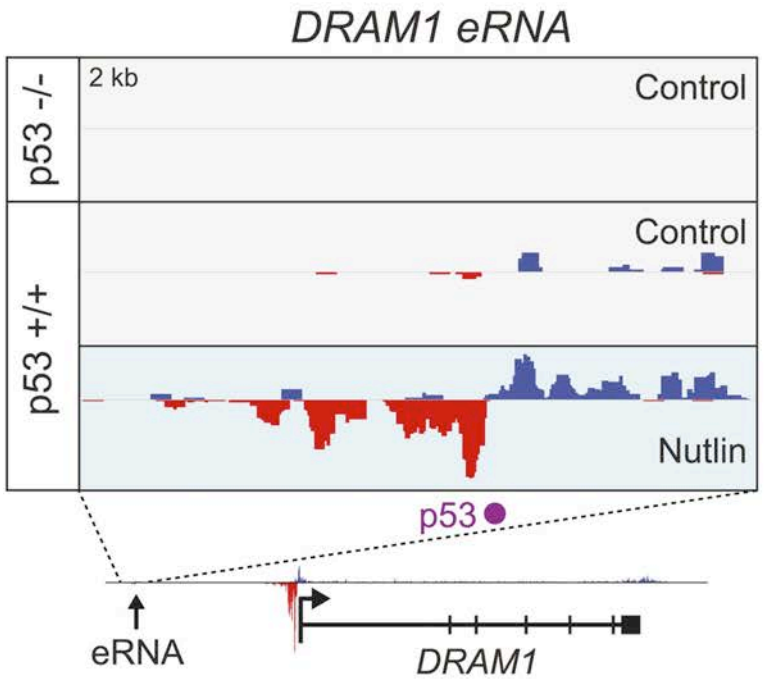



chr1:219,872,217-218,886,719
chr1:220,206,275-220,233,483
Probability Normalized Read Coverage
0.08
0.06
0.04
0.02
0.0
0.02
0.04
0.06
Upstream Enhancer Cluster
EPRS
RNUSF
$\hat{\mu}$
model fit
forward strand
reverse strand
Non-TSS Bidir.
Noise
Reverse Strand Elongation
TSS Bidir.
Forward Strand Elongation
GRO-seq (+)
GRO-seq (-)
Pol-II Ser5
DNase I HS
H3K27ac
H3K4me3
H3K4me1
-1500bp
1500bp
-1500bp
1500bp
Genome Wide Meta Signal
Normalized to max within each dataset
1.0
0.5
0.0

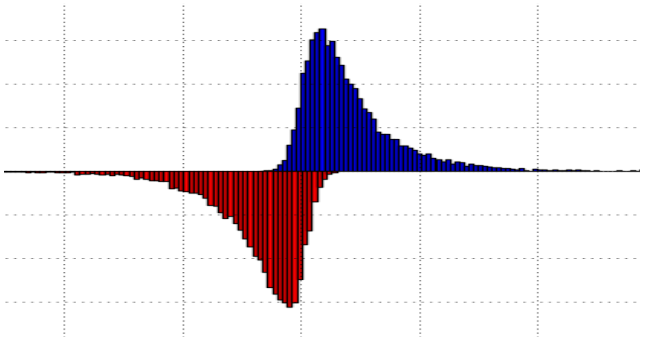
ENCODE

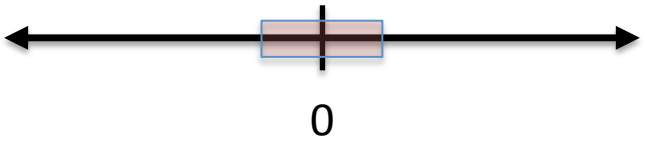


What is the spatial relationship between TF binding and eRNAs?



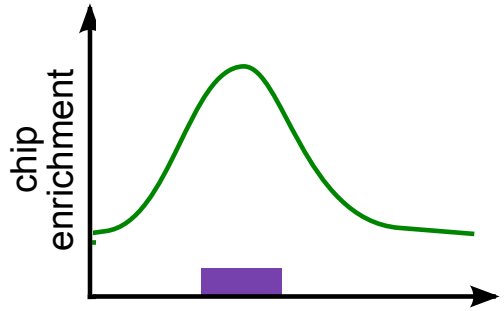
What is the spatial relationship between TF binding and bidirectionals?



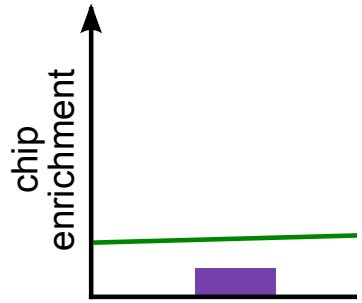


0

$\hat{\mu}$

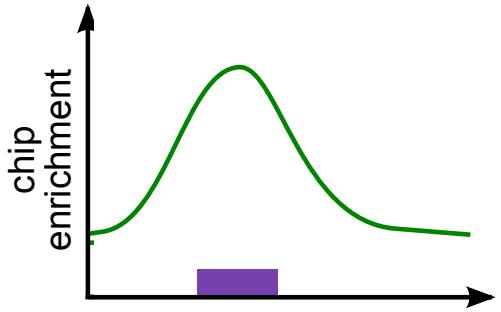


Bidirectional with bound motif
(ChIP-seq)

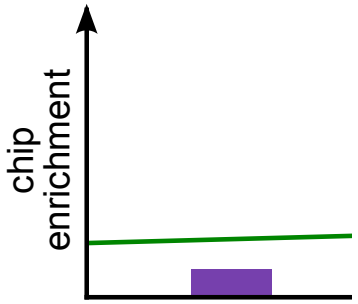


Bidirectionals not bound
(ChIP-seq)

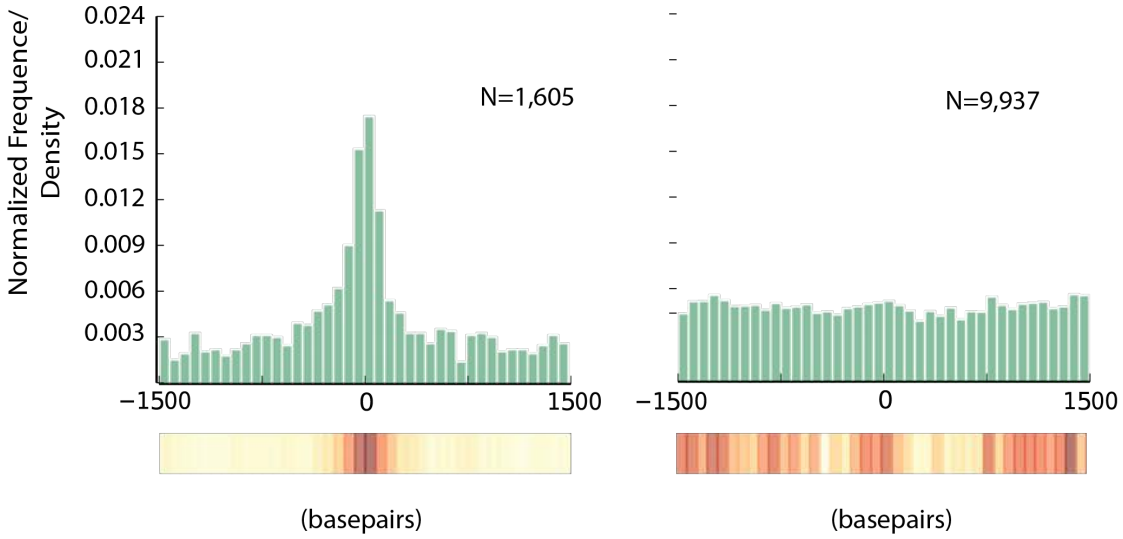
Bidirectionals that correspond to ChIP peaks appear to originate at the motif.



Bidirectional with bound motif (ChIP-seq)

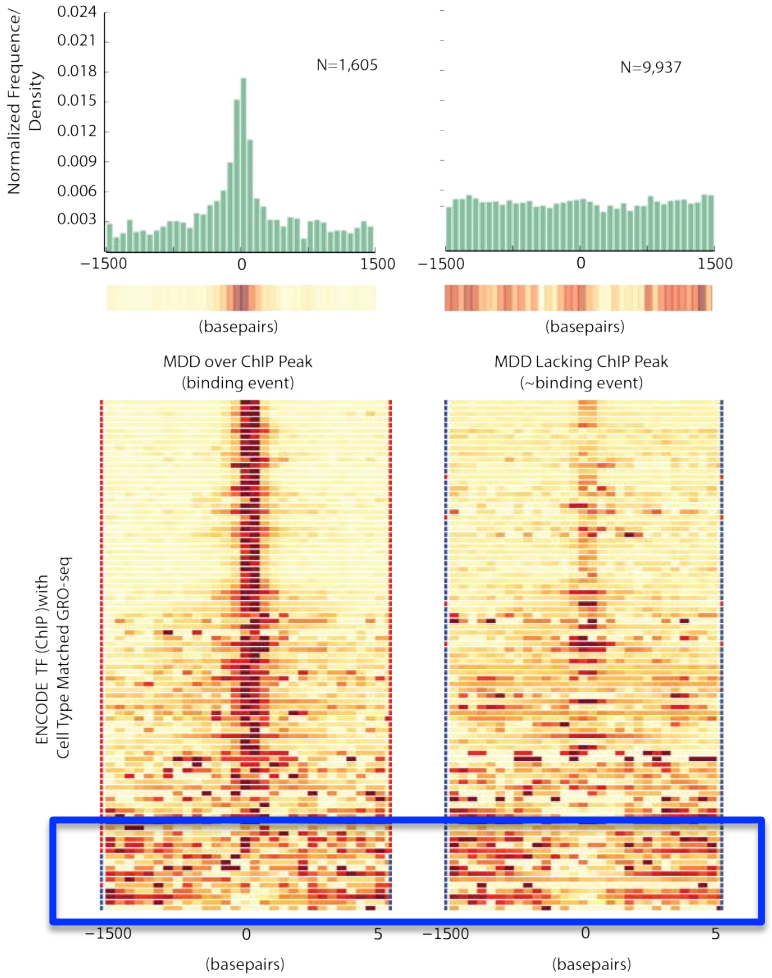


Bidirectionals not bound (ChIP-seq)



Displacement of the motif around the bidirectional.

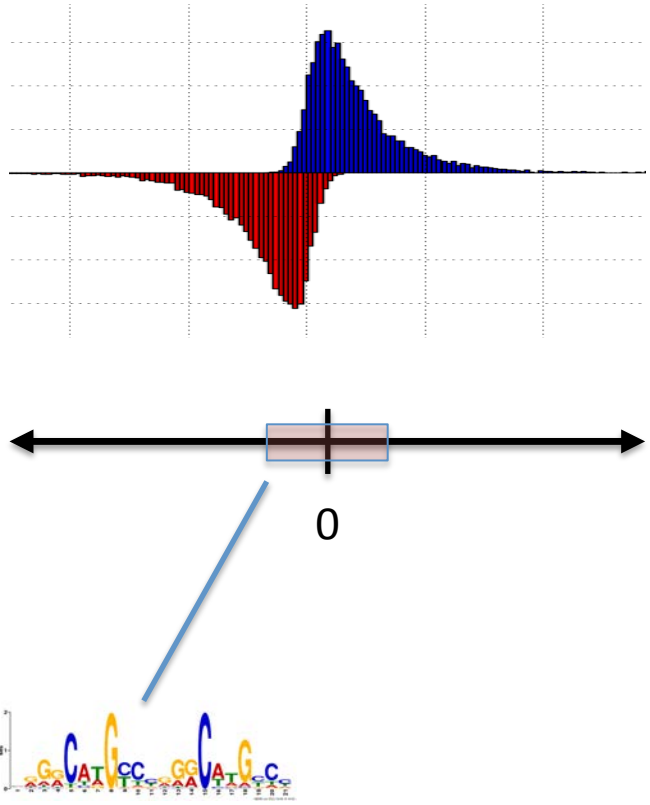
Bidirectionals that correspond to ChIP peaks appear to originate at the motif.

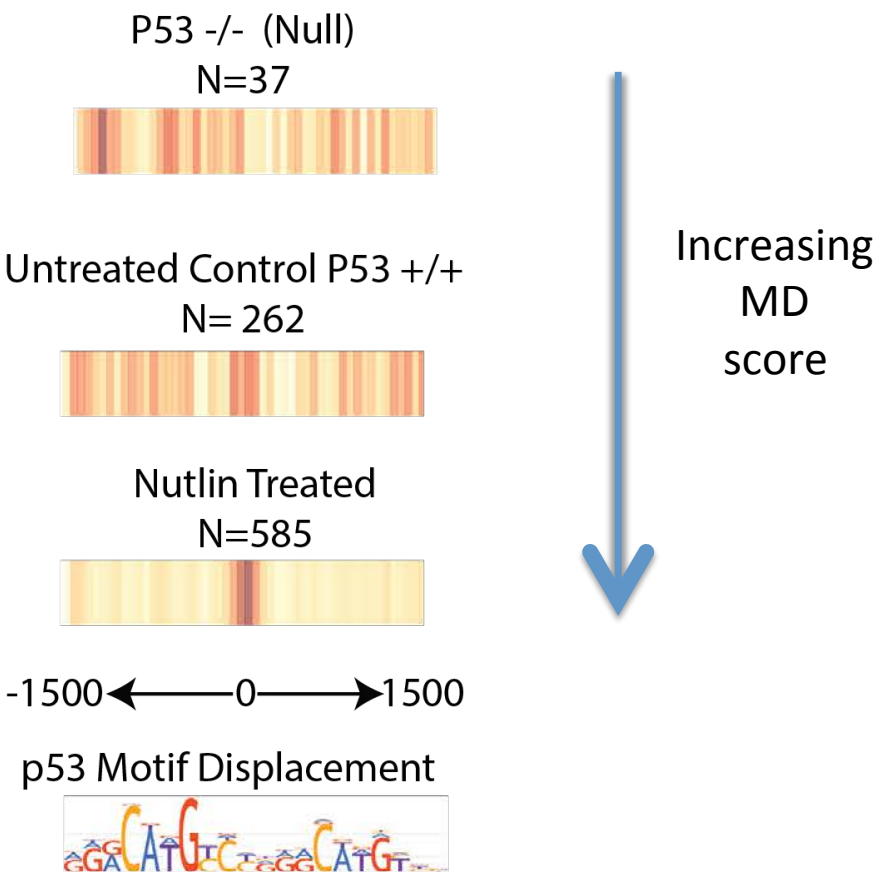


Repressors!

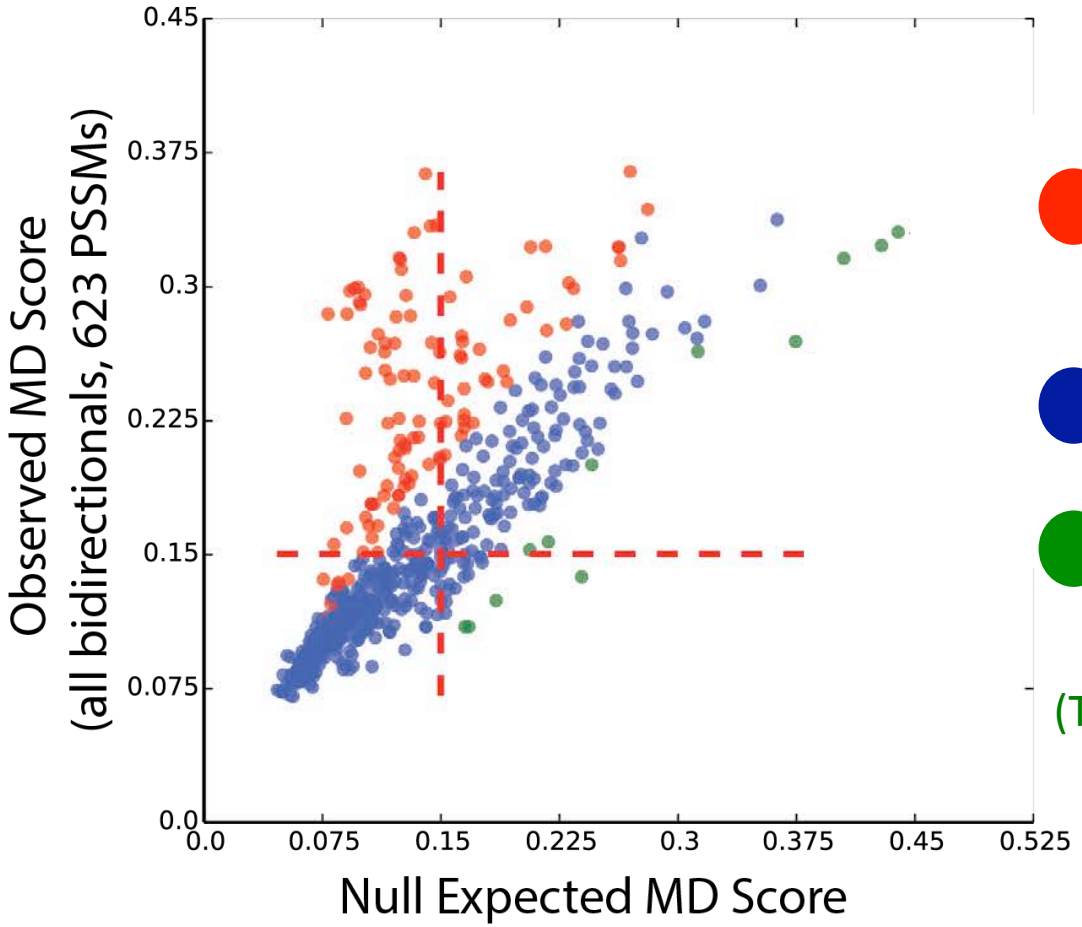
ENCODE, ~40 published human GRO-seq datasets

Our Motif Distribution (MD) score effectively measures co-occurrence of motif and bidirectional.





MD score predicts which TFs are active in a given cell type.



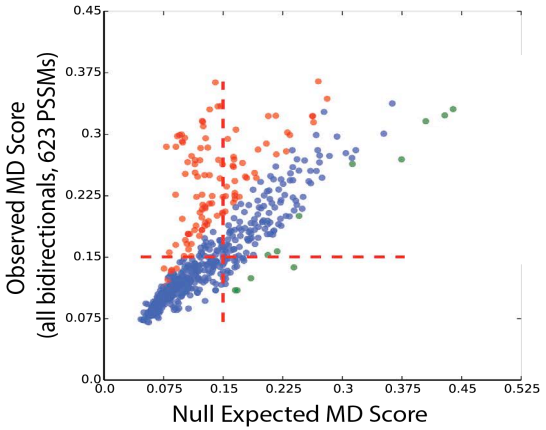
Active Transcription Factor (Statistically higher than expectation)

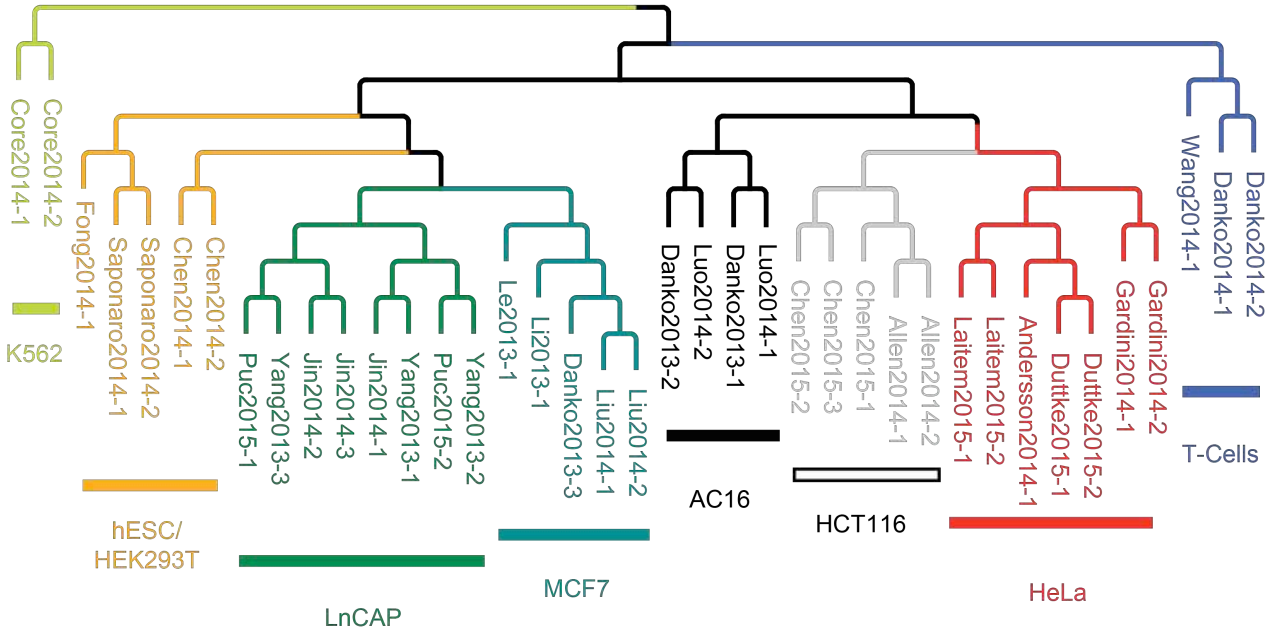
At Expectation

Depleted relative to expectation

(These turn out to be repressors!)

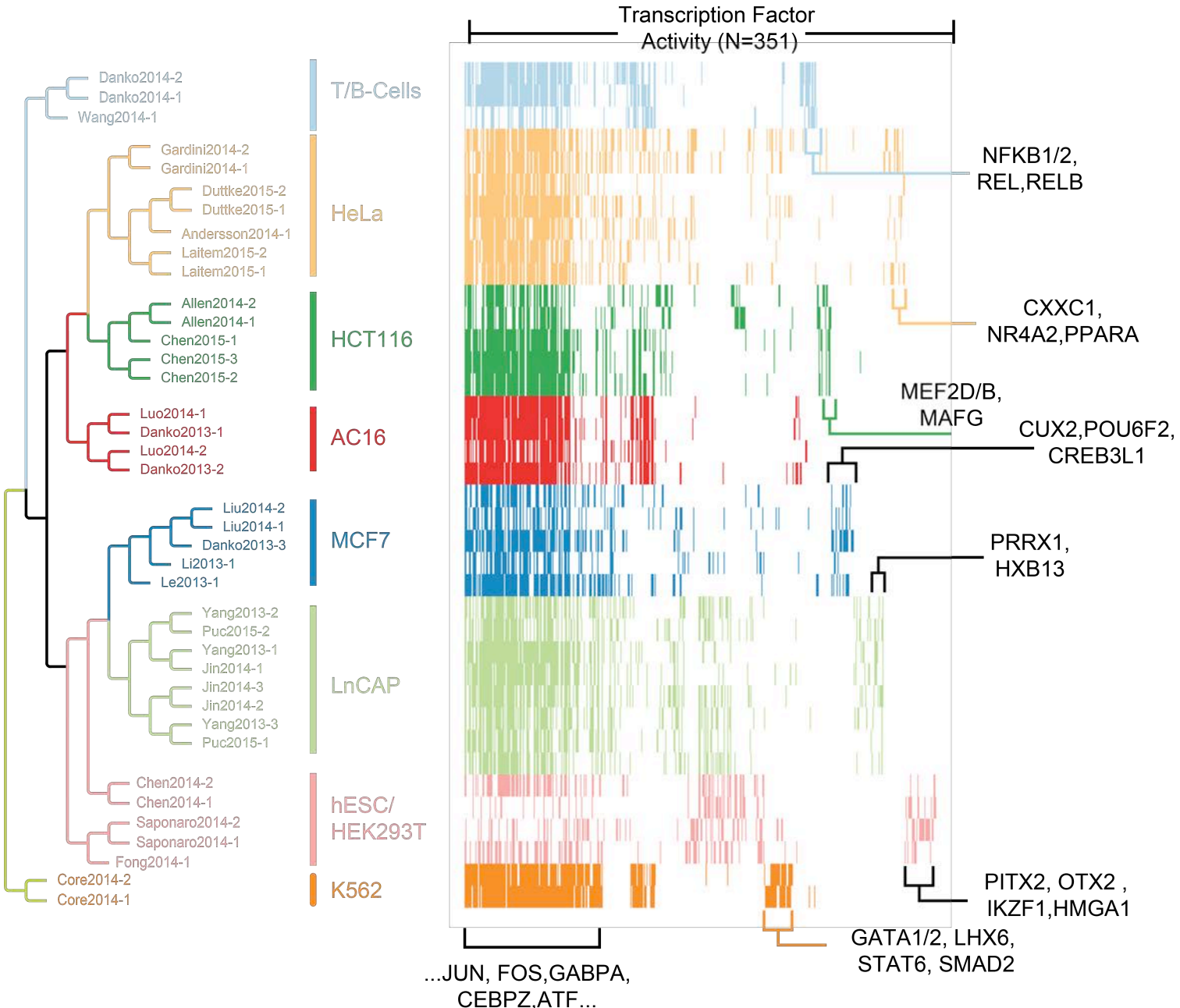
Clustering by “active” TFs separates datasets by cell type



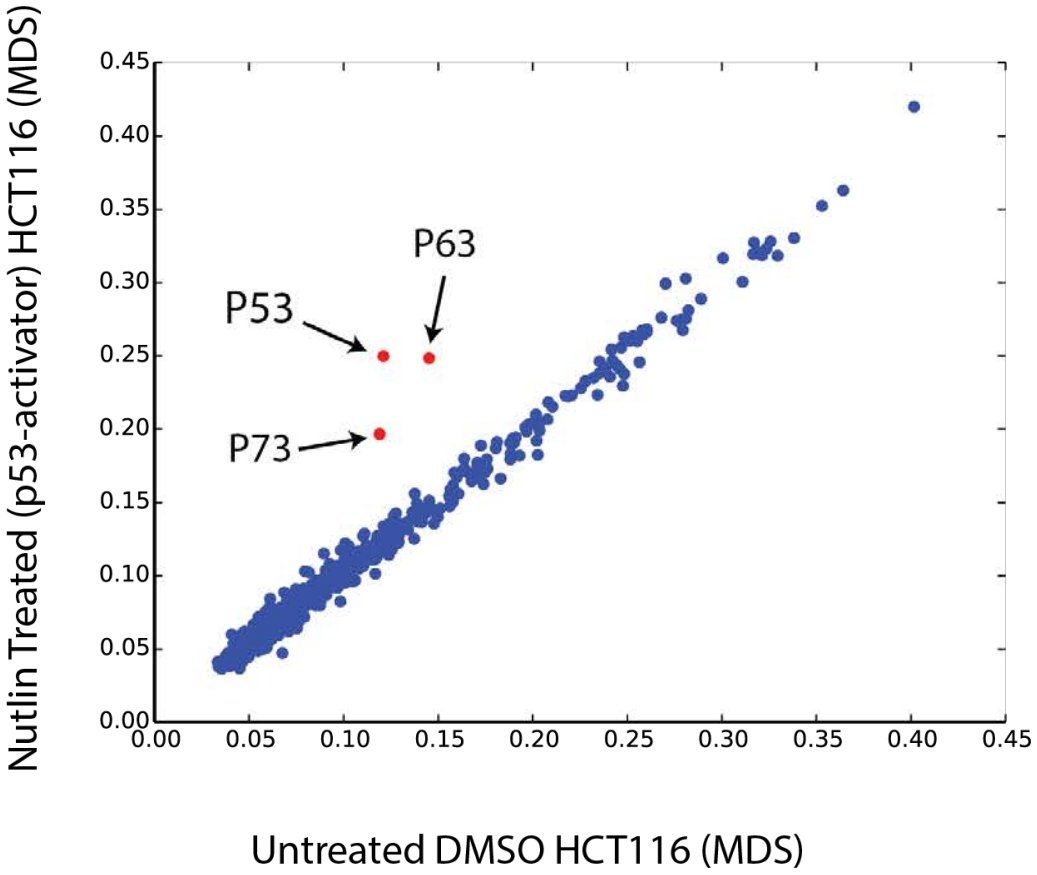


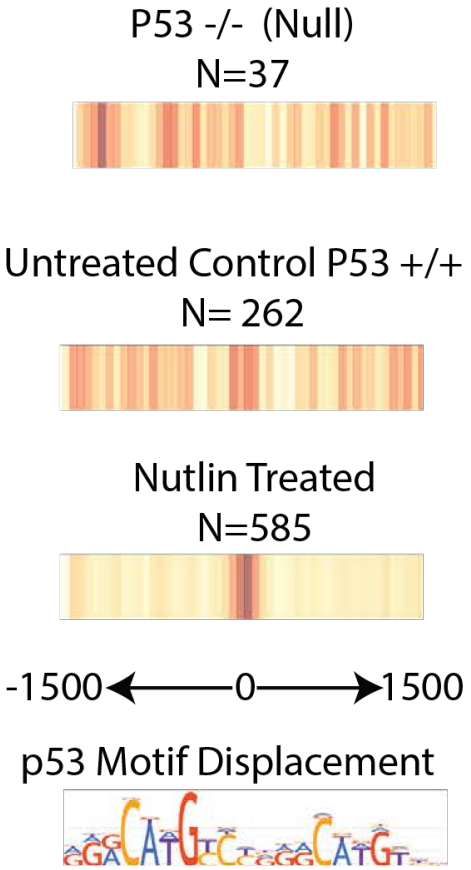
~40 published human GRO-seq datasets

And identifies cell type specific factors



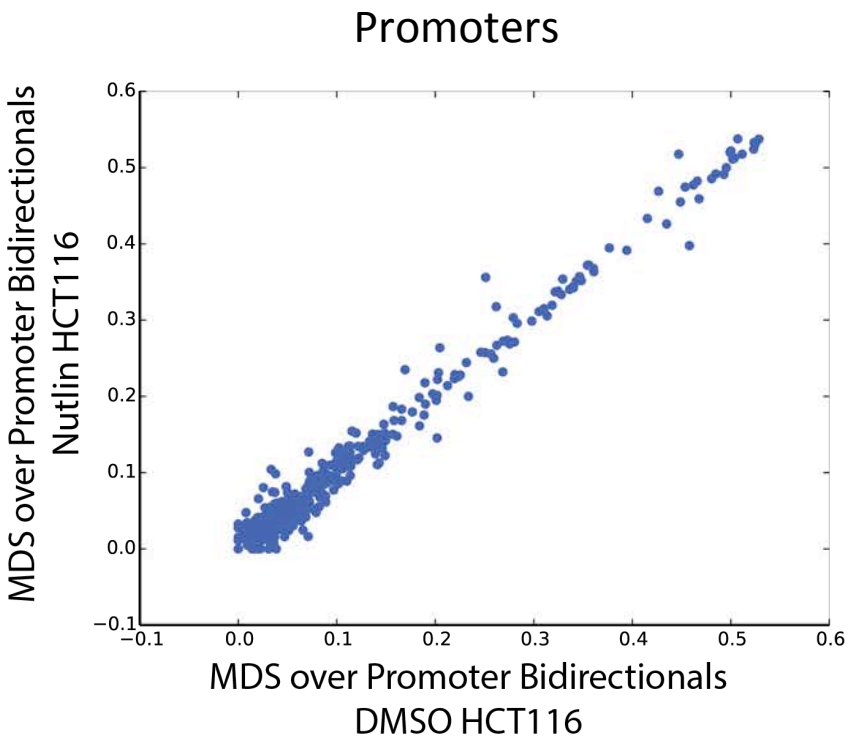
Differential MD score identifies which TF has changed after perturbation

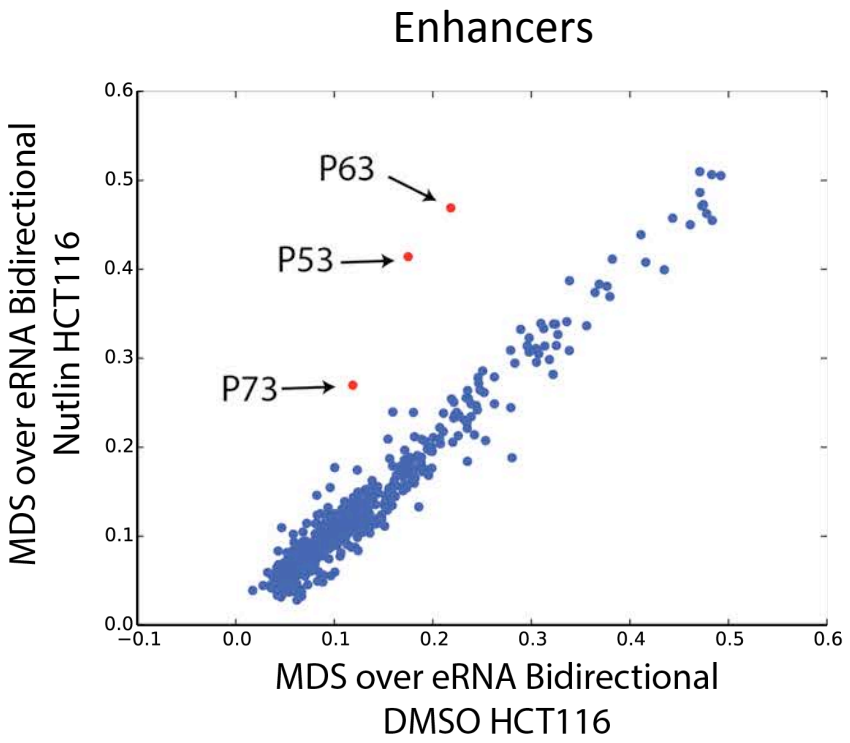




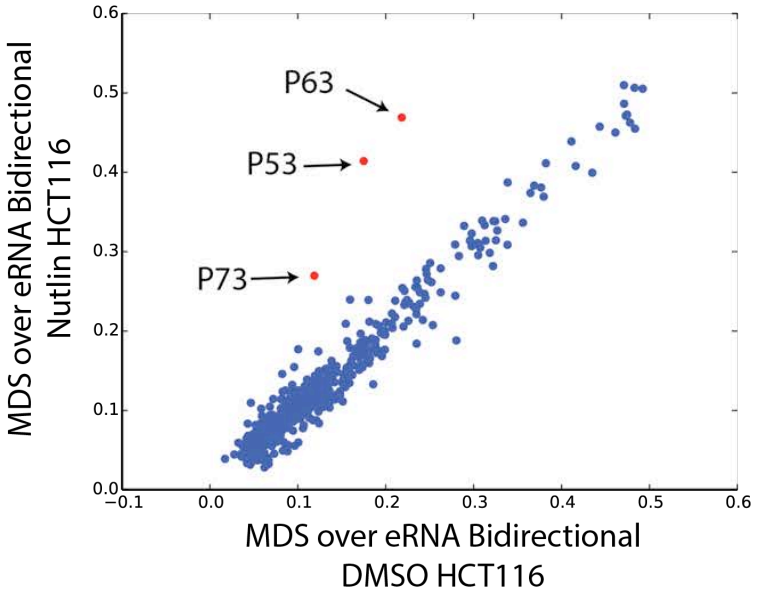
Allen et. al. 2014

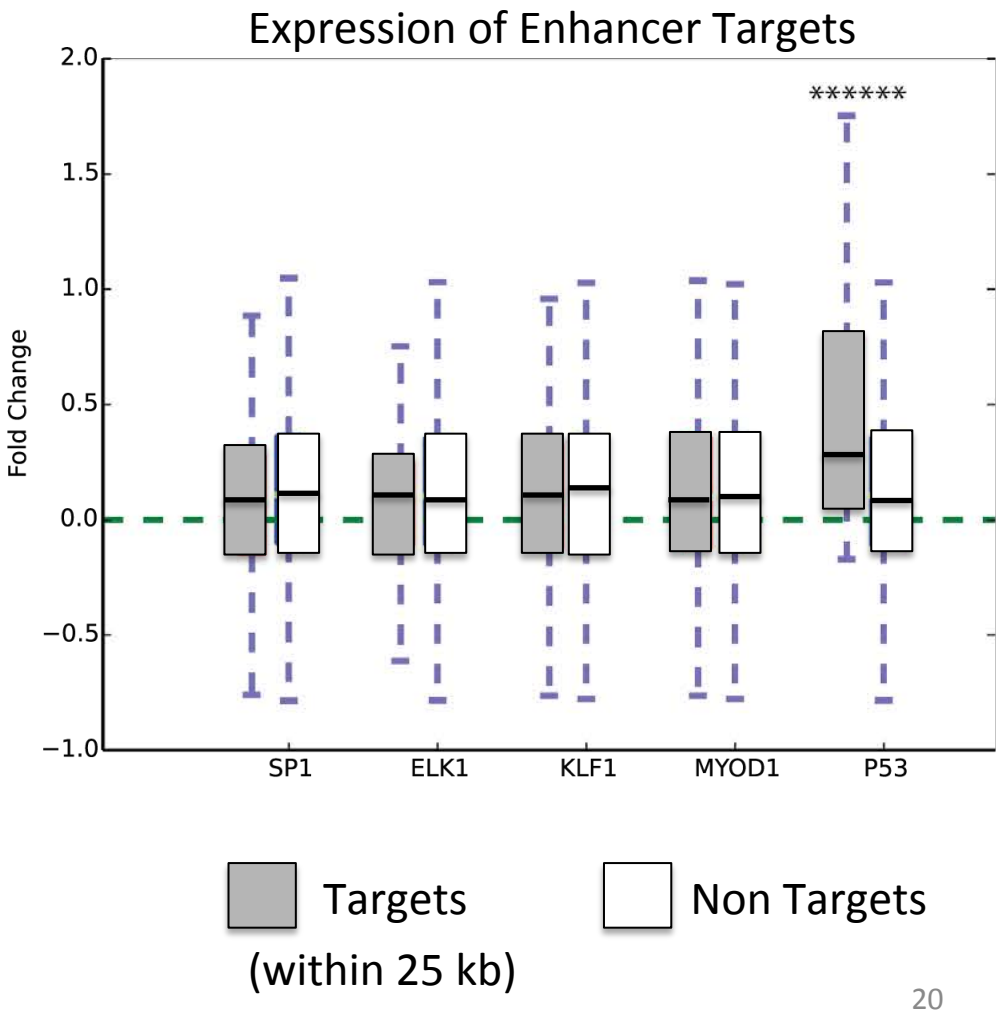
The activation signal is predominantly at enhancers.



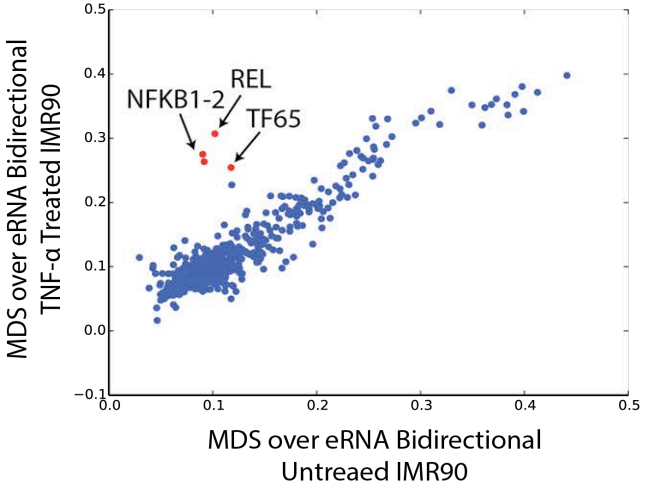


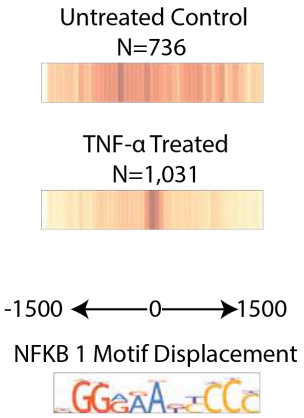
Furthermore, the changed transcription factor drives activity at target genes.

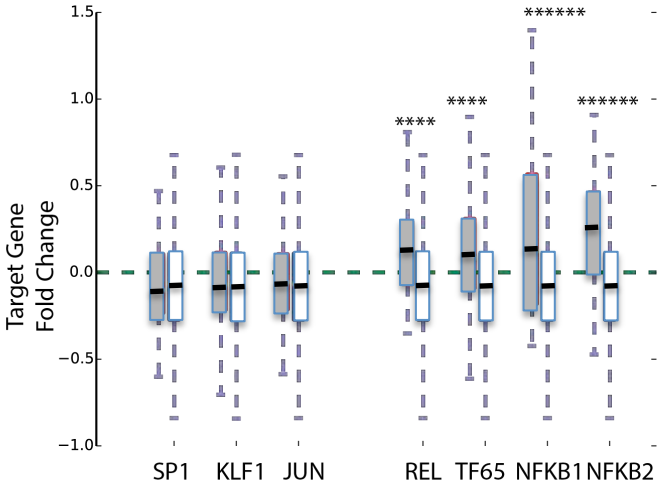


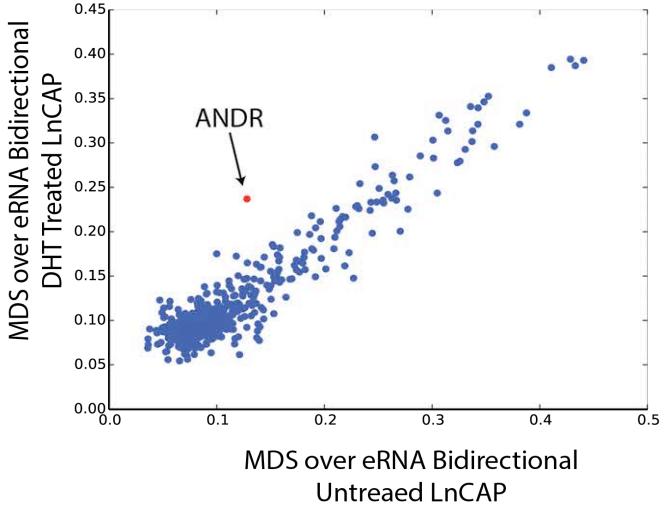


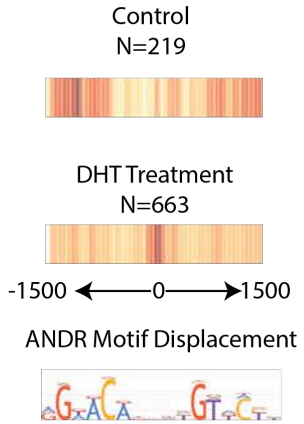
Enhancers are major responder to multiple stimuli

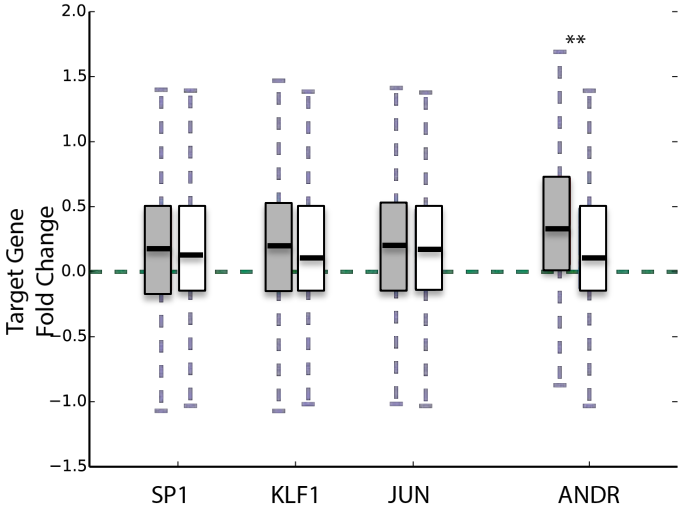








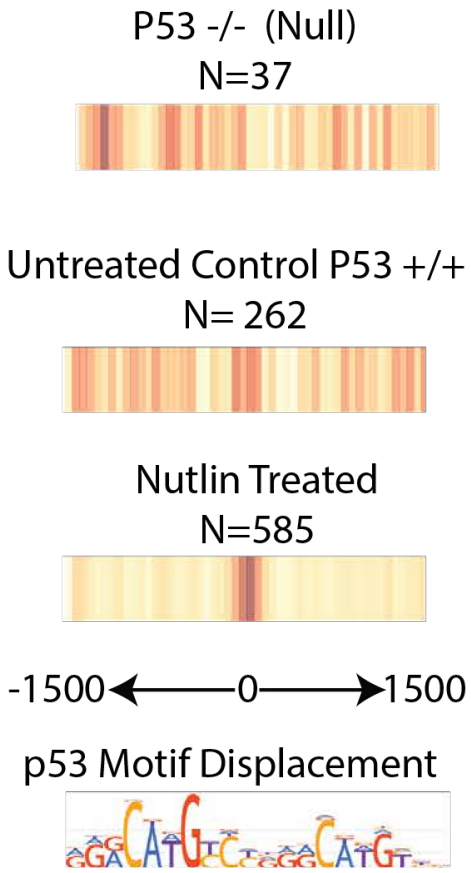




Jin et. al. 2013; Puc et. al. 2015

Summary

- We have a principled mathematical model of nascent transcription.
- We can infer from patterns of bidirectionals which TFs are active in the cell type.
- We can infer which TFs respond to a perturbation by changes in bidirectional usage.



Acknowledgements

Dr. Mary A. Allen
(MCD Biology, BioFrontiers)

Joseph G. Azofeifa
(Computer Science)

BOETTCHER®
FOUNDATION

http://dowell.colorado.edu

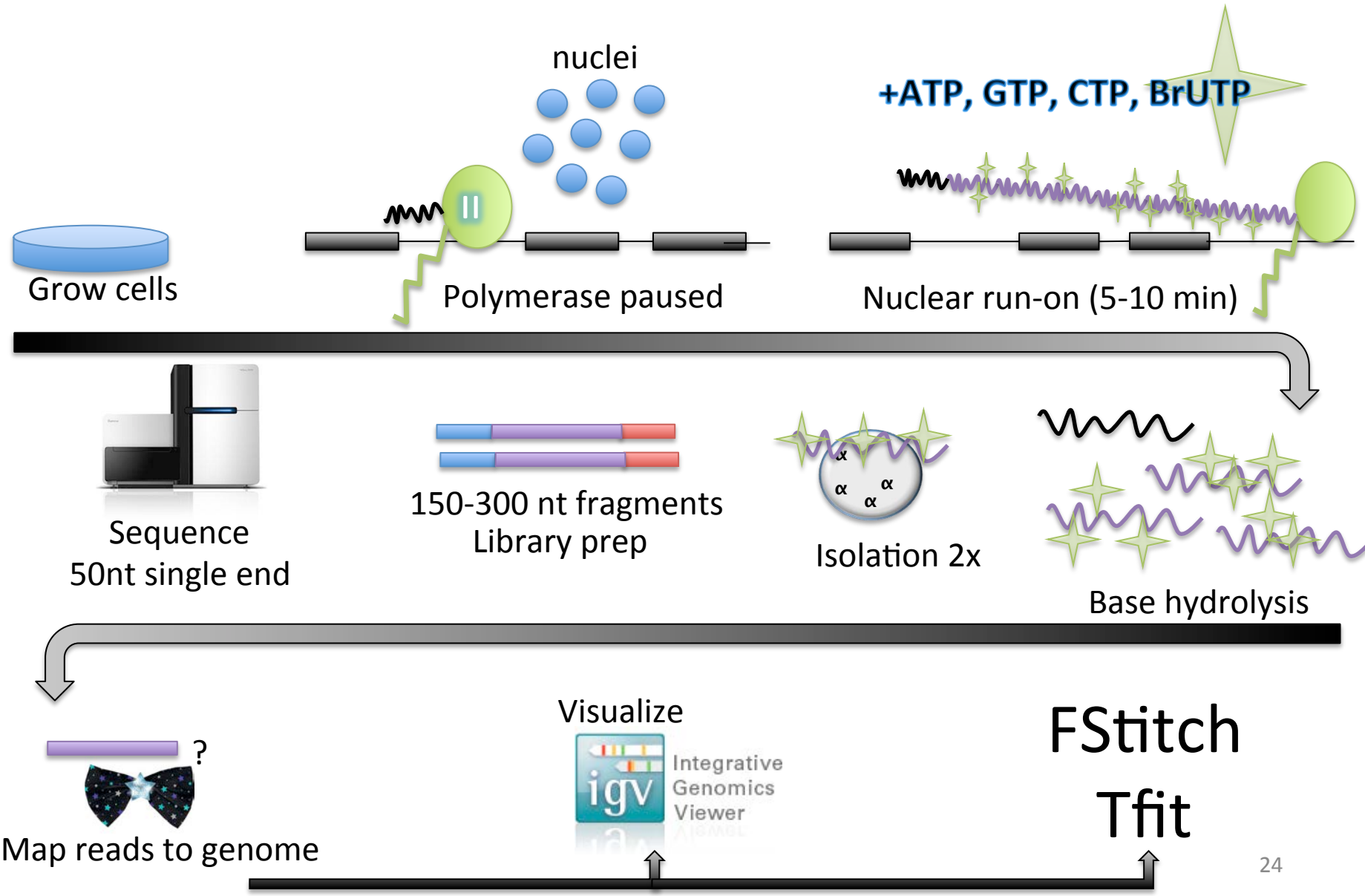
Josephina Hendrix
(MCD Biology)

Dr. Tim Read
(MCD Biology, BioFrontiers)

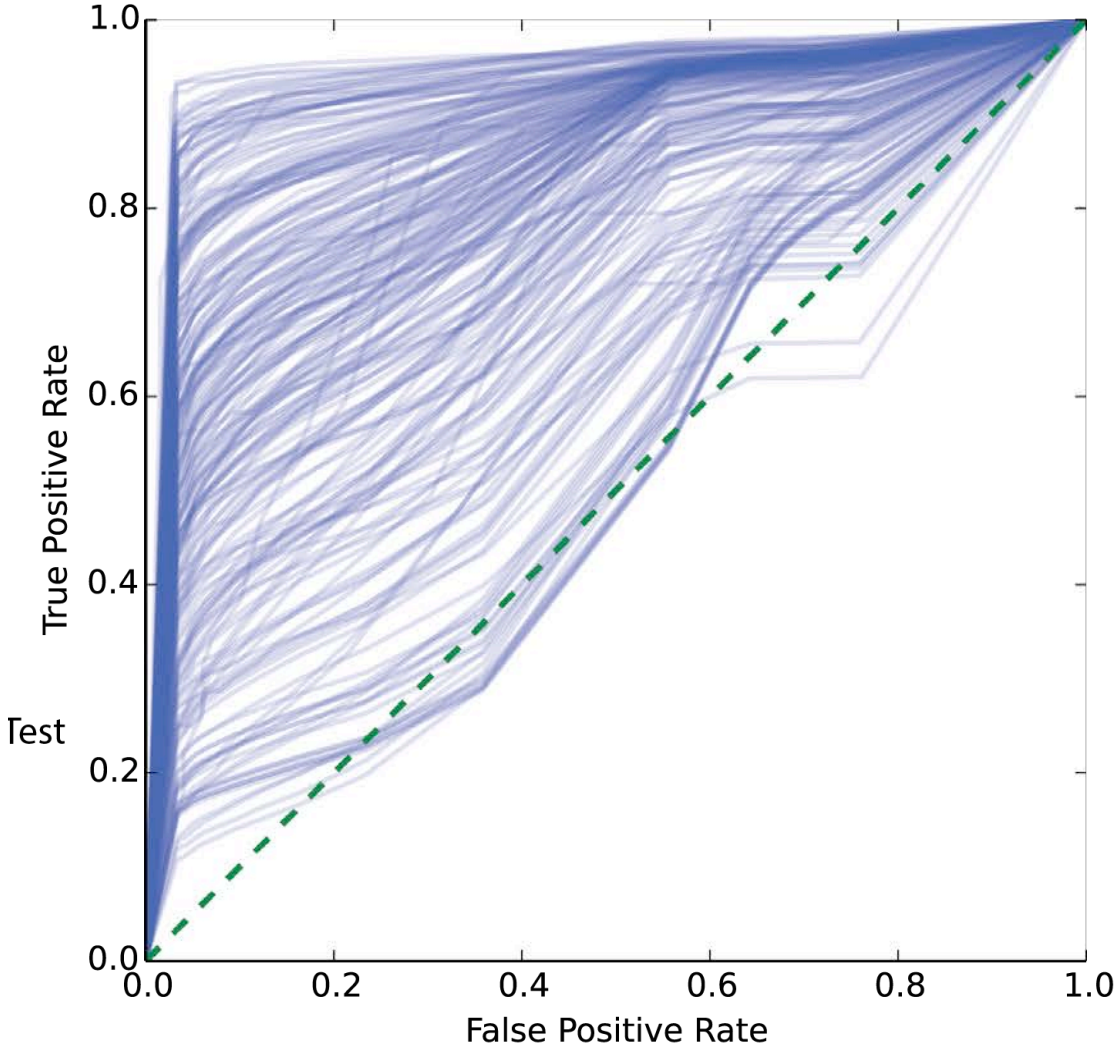
Jonathan Rubin
(Biochemistry)

Global Nuclear Run On Sequencing

GRO-seq: isolating and sequencing nascent RNA



Fraction of ChIP signal predicted by bidirectional varies by TF.



Enrichment over background for MD scores