

## Bayesian Probabilistic Modeling



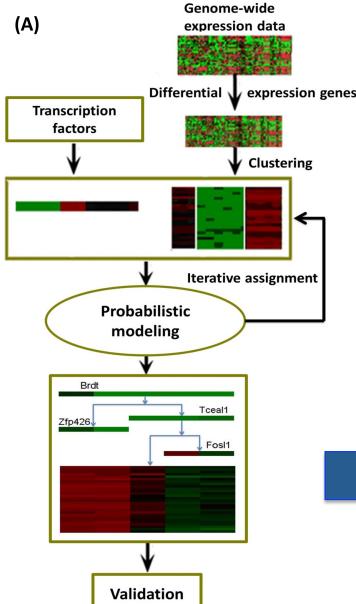
- Assign genes into co-regulated modules
- Construct regulatory relations of each module



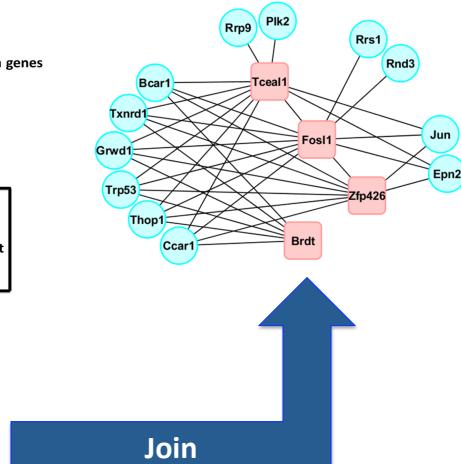
$$P(GRN | D) = \underset{GRN}{\operatorname{argmax}}(P(D|GRN) * P(GRN))$$

Posterior                  Likelihood                  Prior

## Gene Regulatory Network Modeling

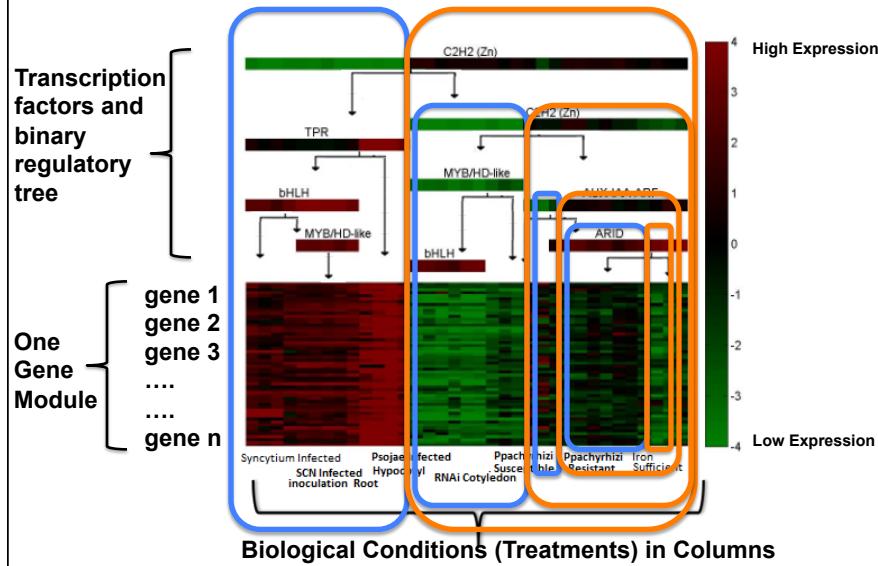


(B) A portion of a global gene regulatory network

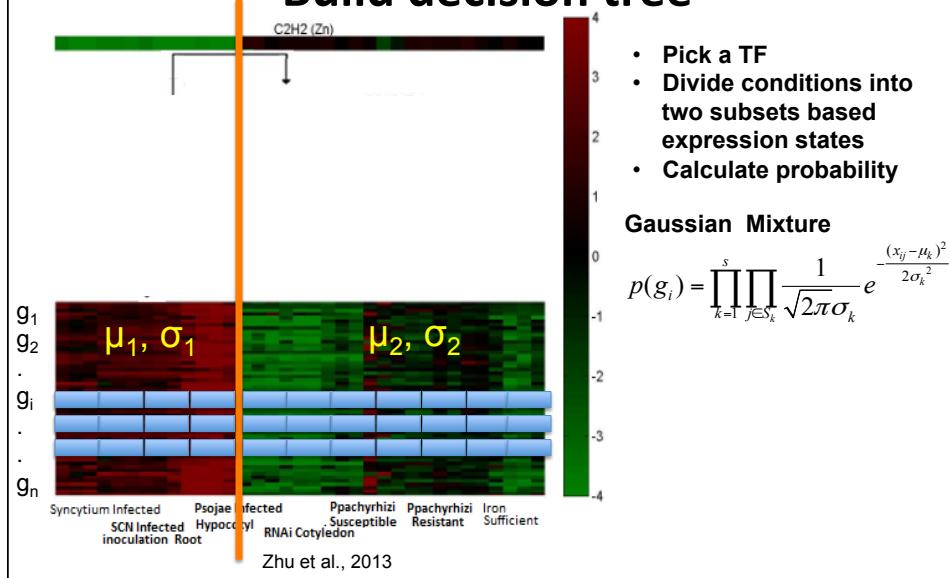


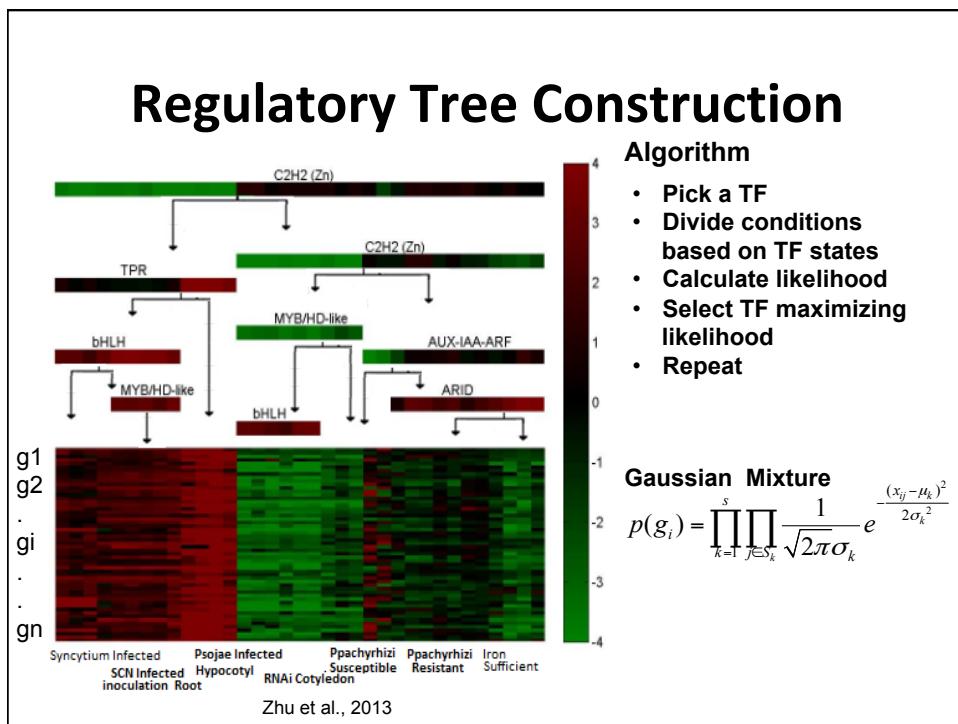
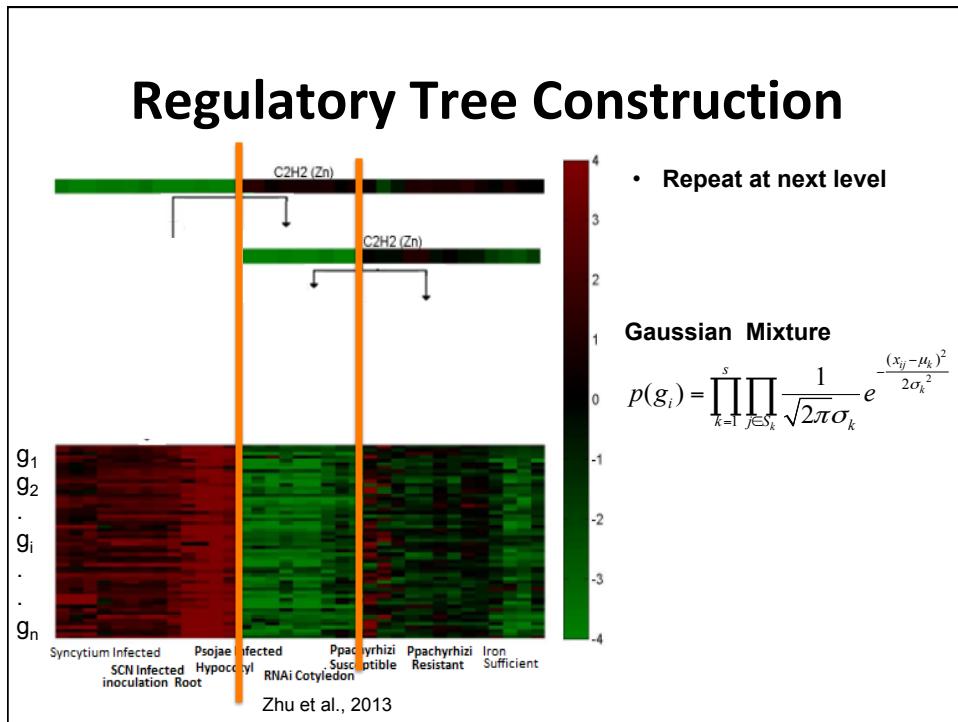
Zhu et al., 2013

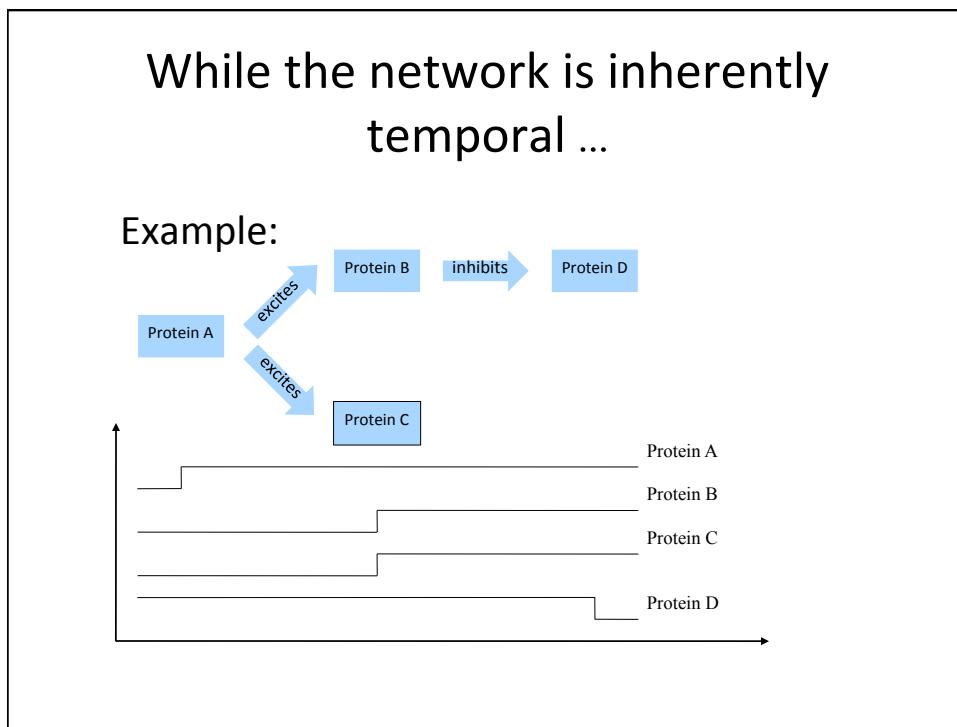
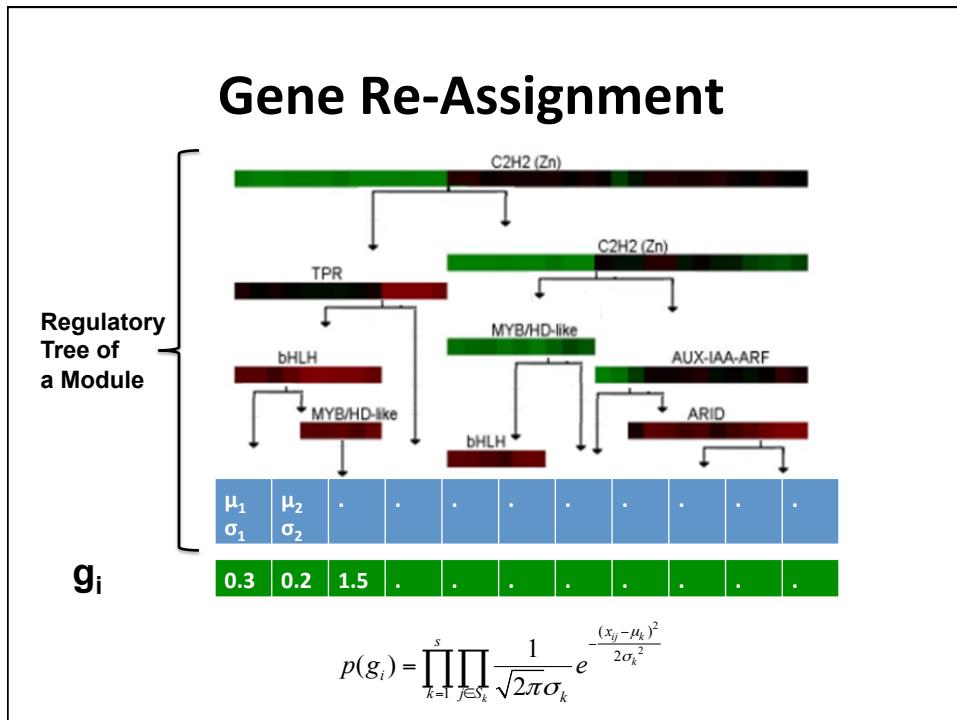
## Gene Regulatory Logic of a Gene Module as a Decision Tree



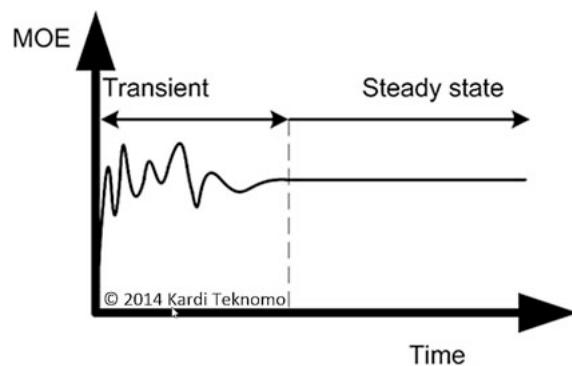
## Regulatory Tree Construction: Build decision tree



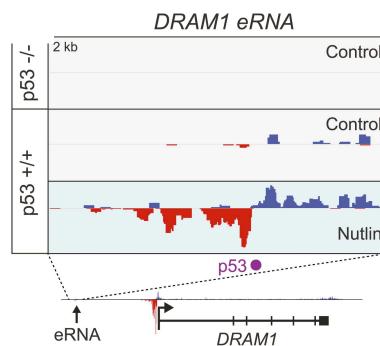


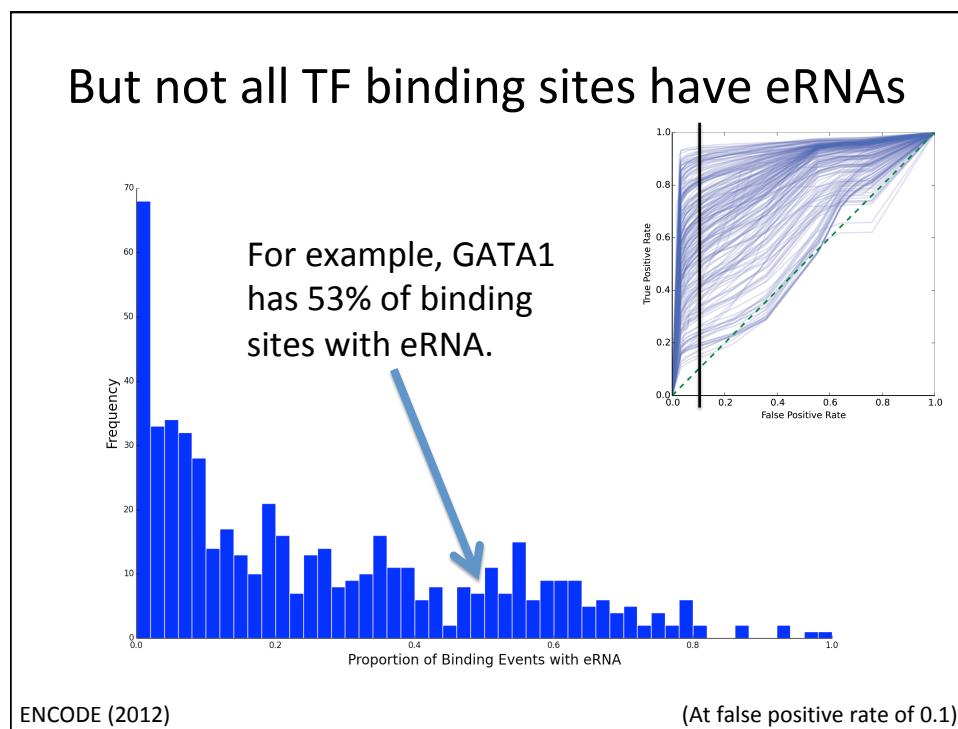
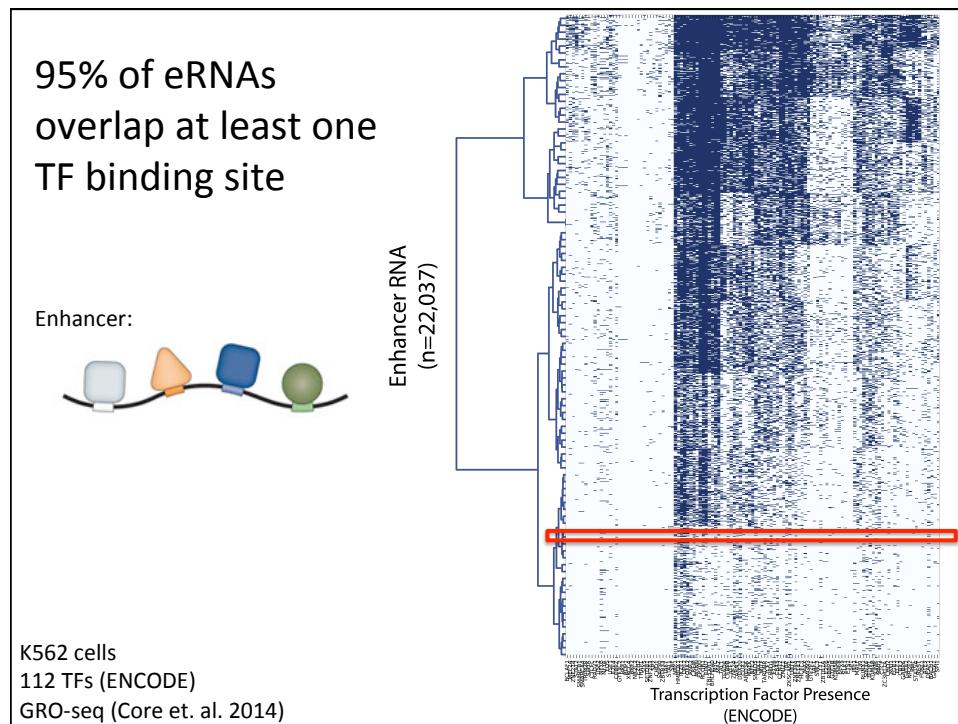


Expression studies are inherently problematic:  
steady state and temporal lag

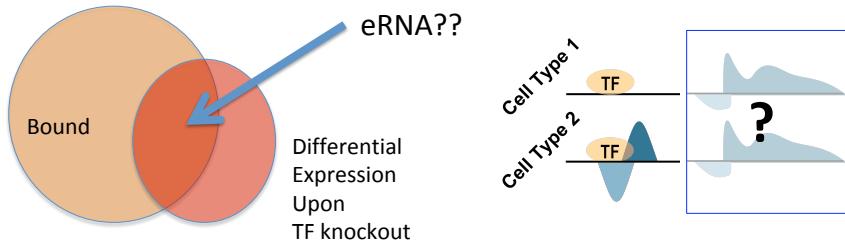


Nascent transcription is a more direct and immediate readout on transcription.



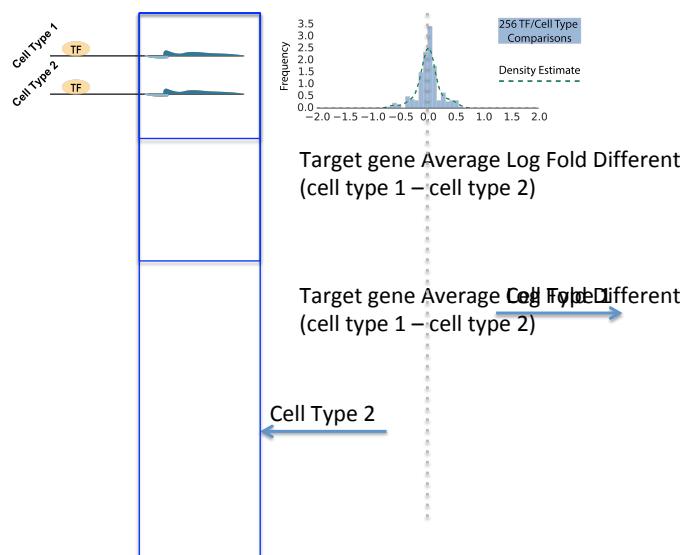


So what is the difference between a bound site with an eRNA and one without?

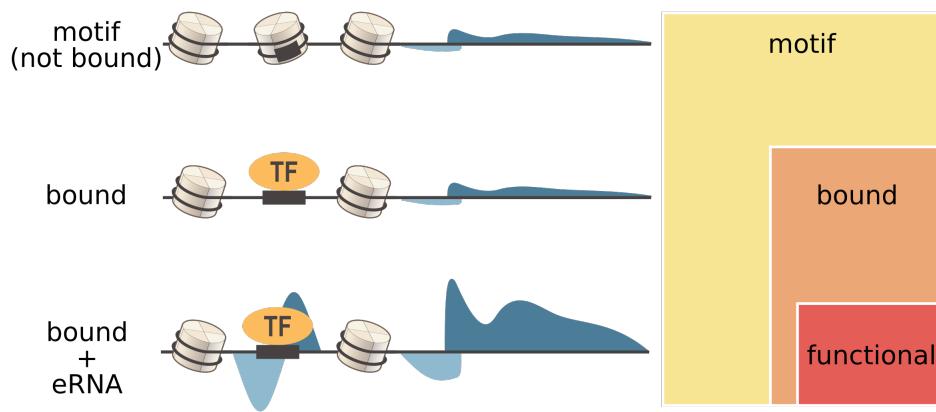


Perhaps eRNAs indicate sites of **functional** transcription factor binding?

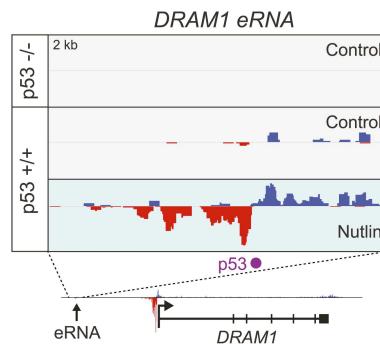
eRNAs indicate functional binding



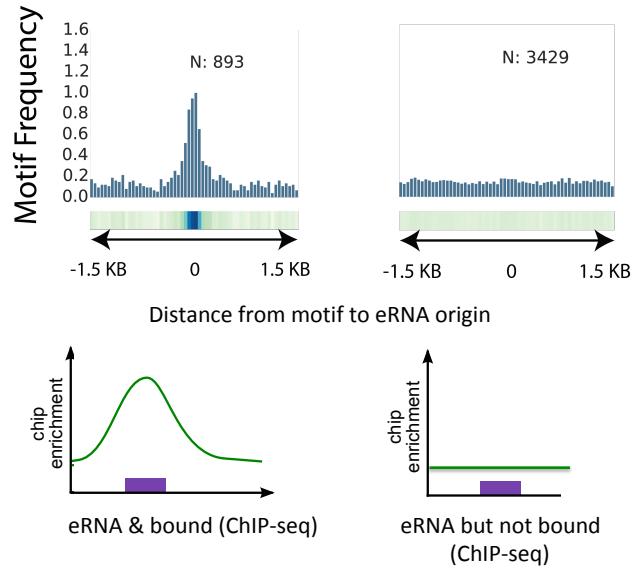
## Model: The functional motifs associate with eRNAs



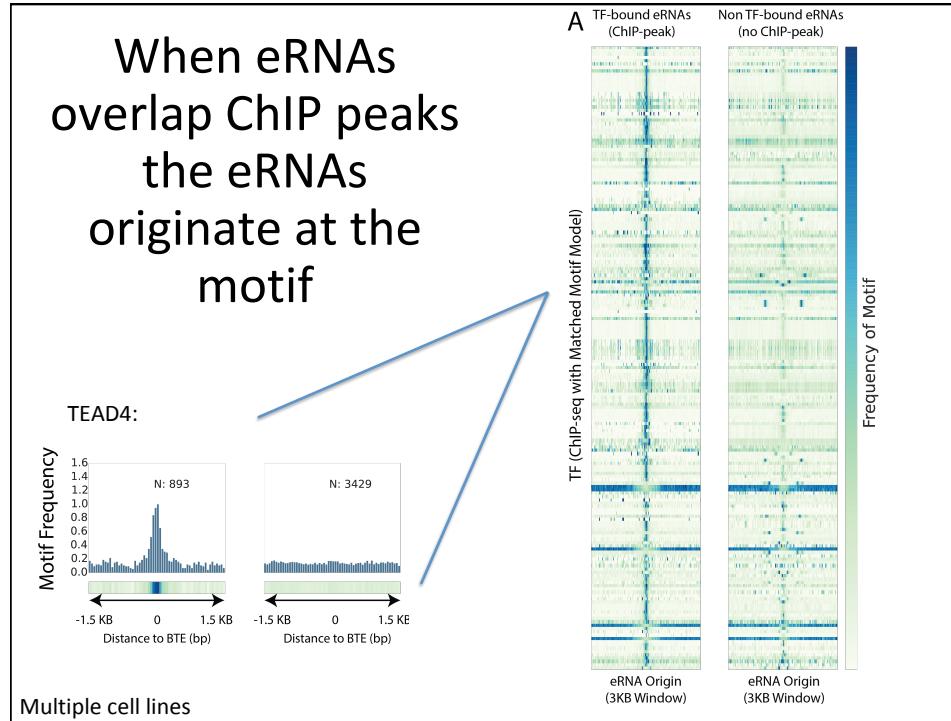
What is the spatial relationship  
between TF binding and eRNAs?

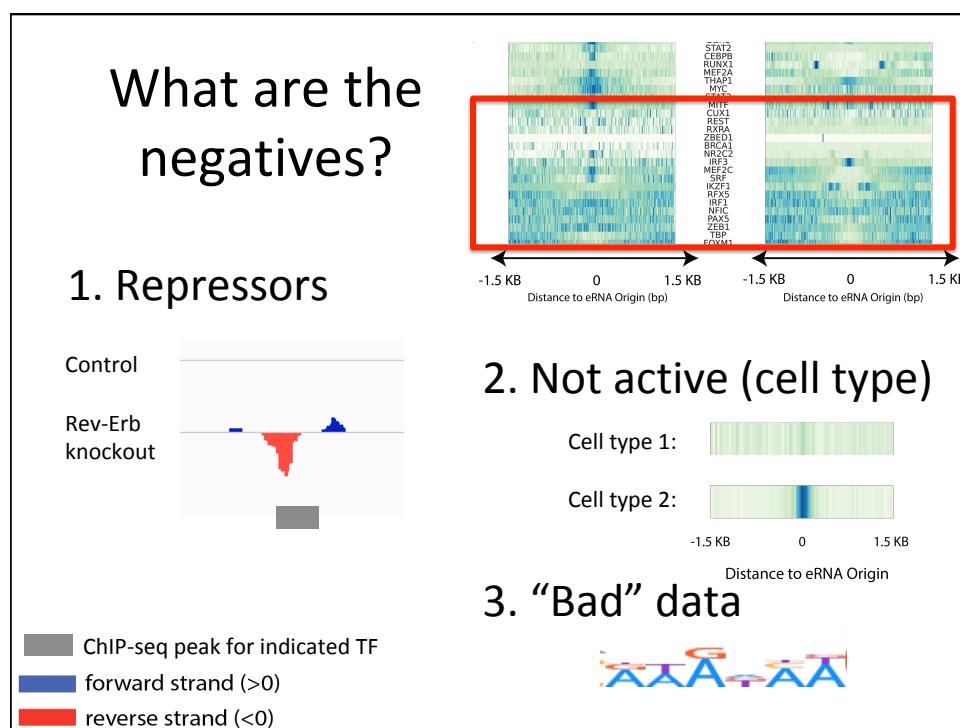
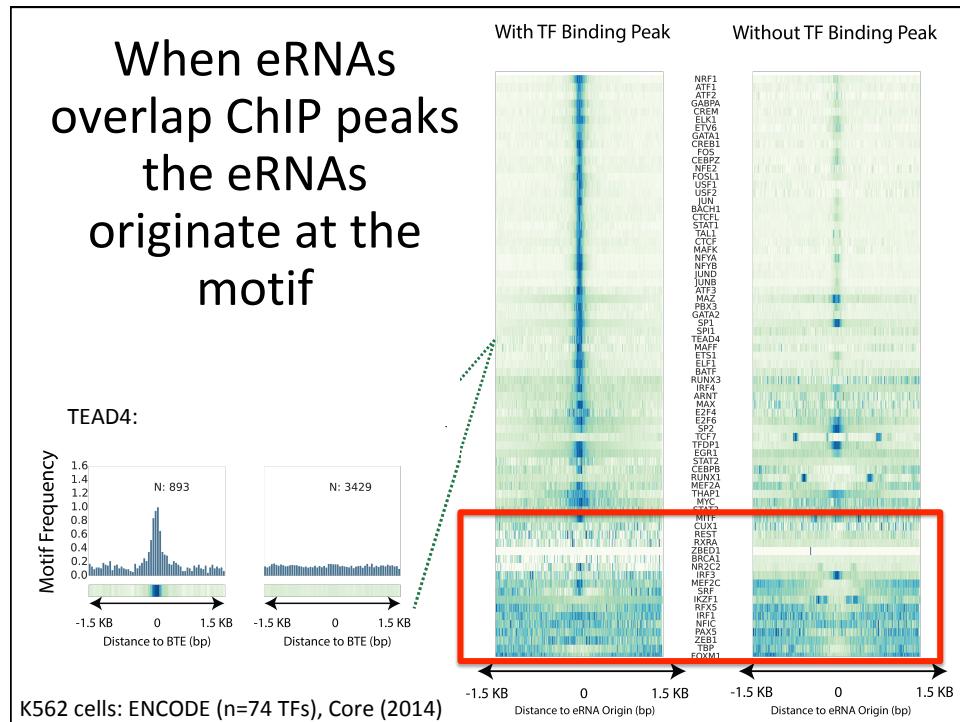


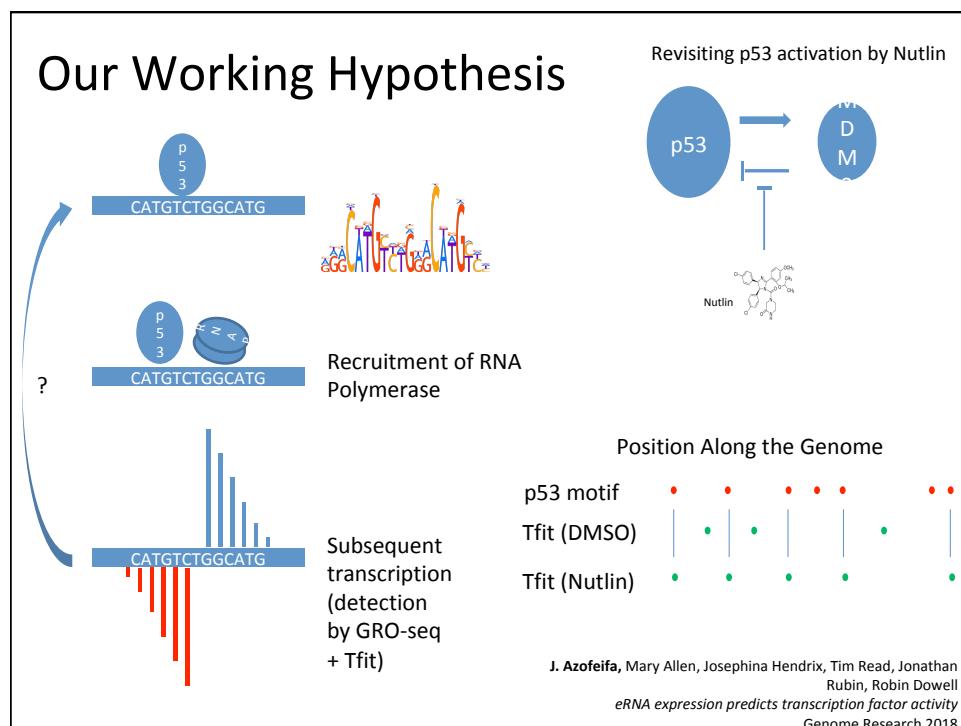
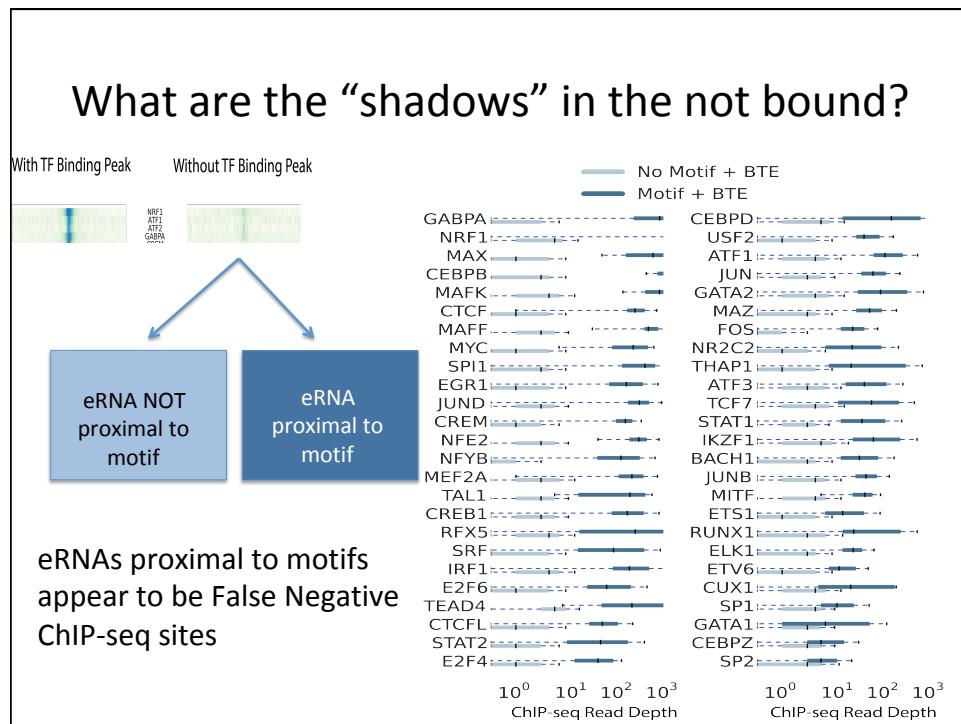
When eRNAs overlap TEAD4 binding sites, the eRNA origin is proximal to the TEAD4 motif

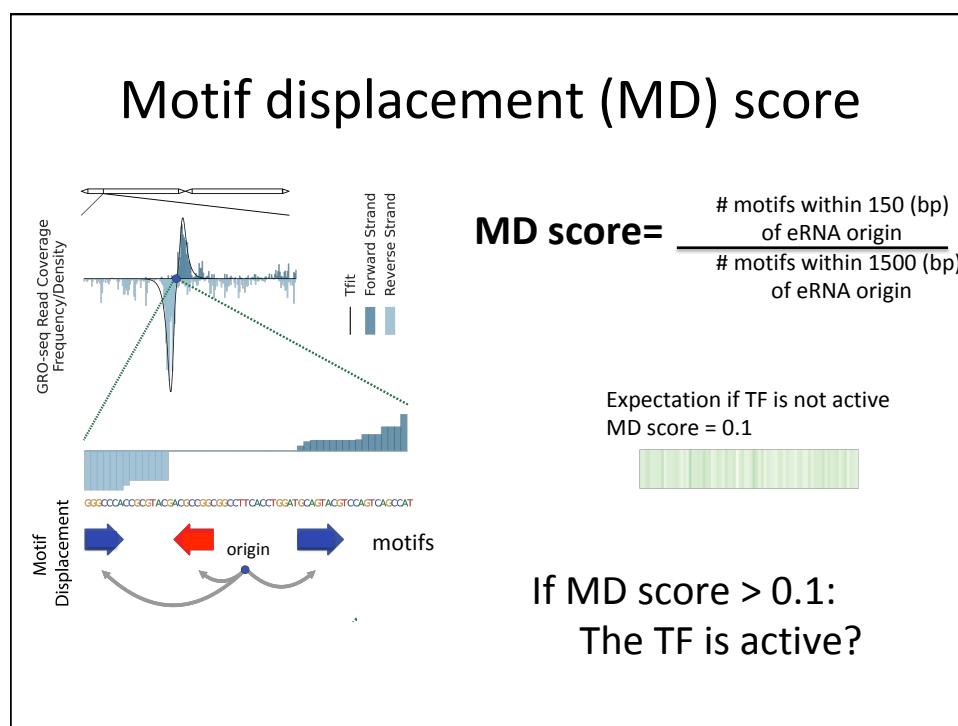
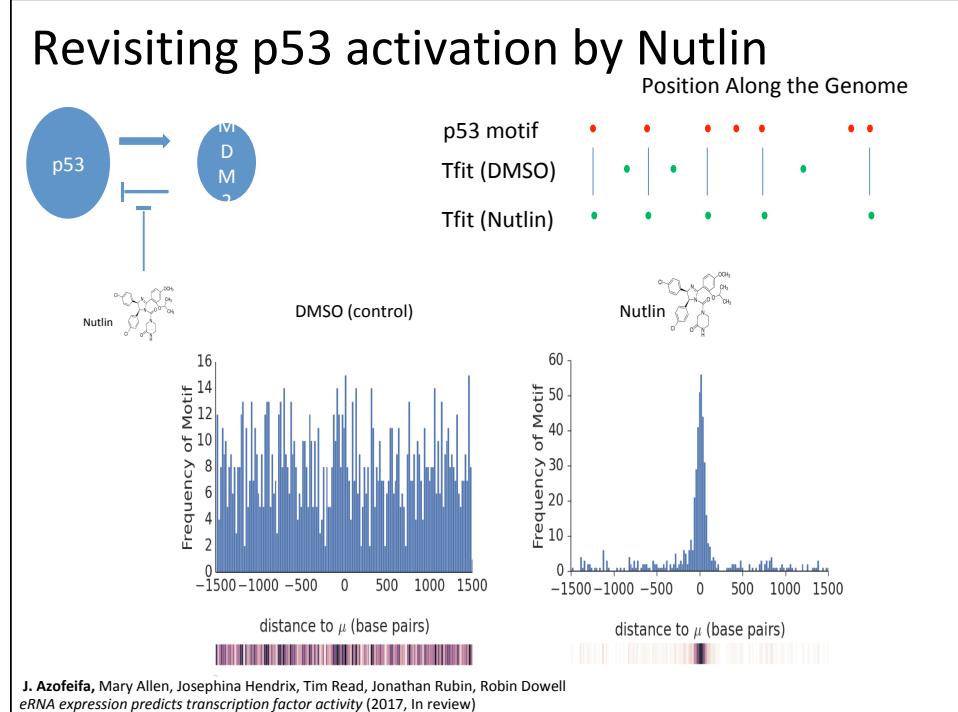


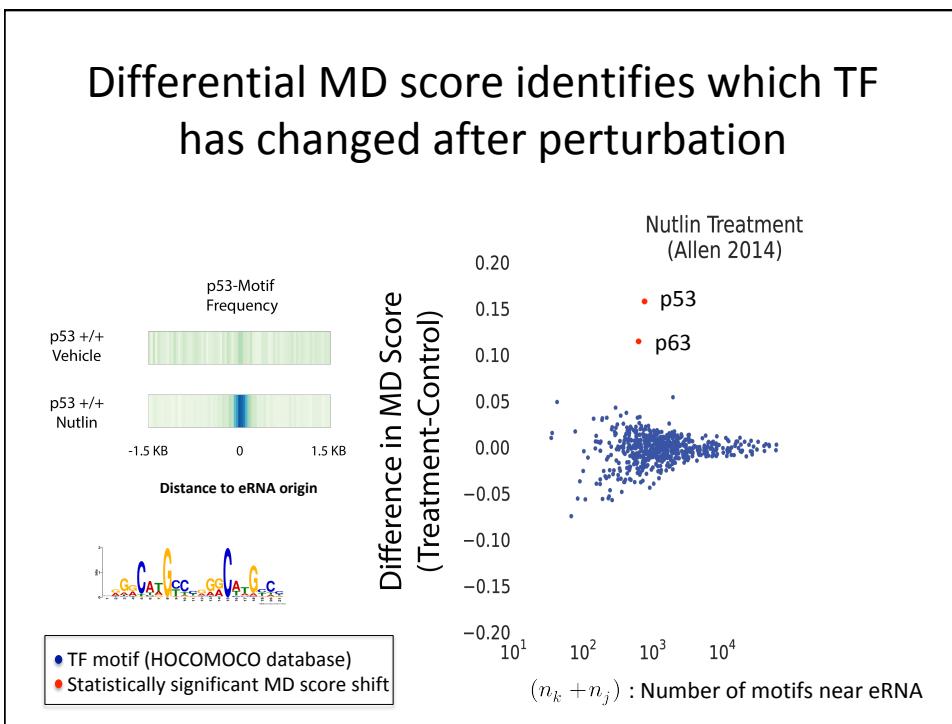
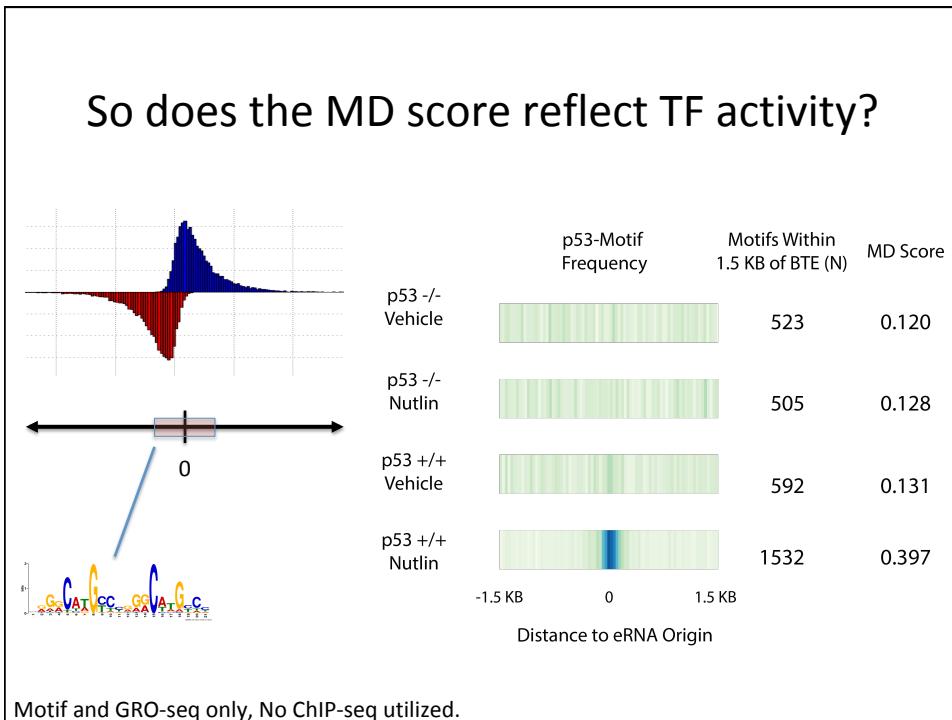
When eRNAs overlap ChIP peaks the eRNAs originate at the motif



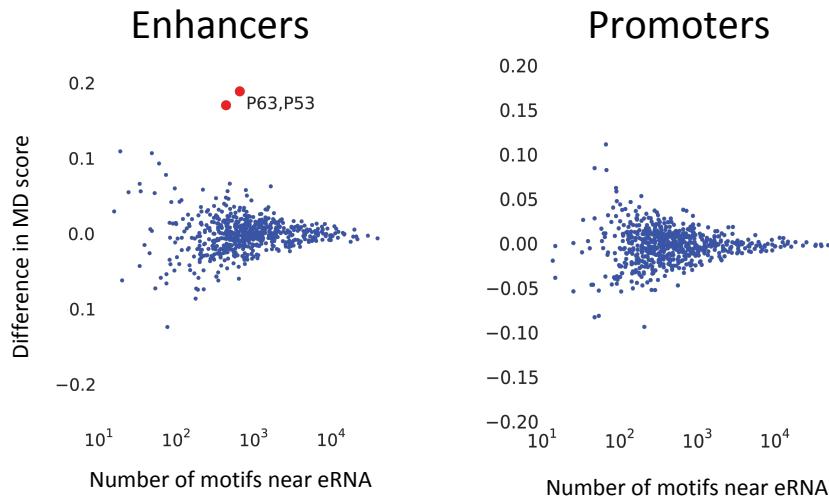




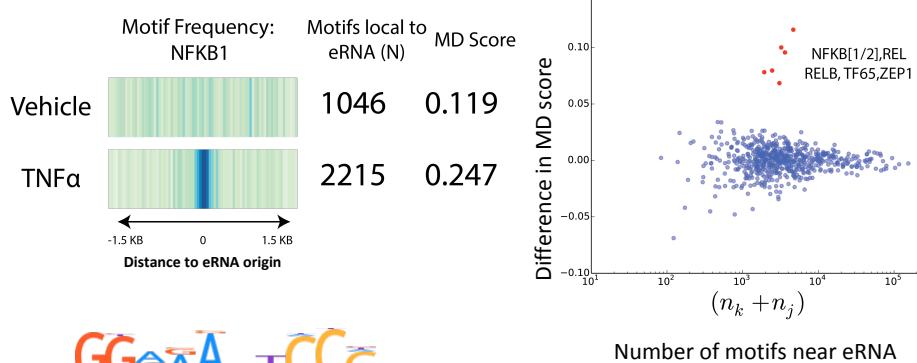


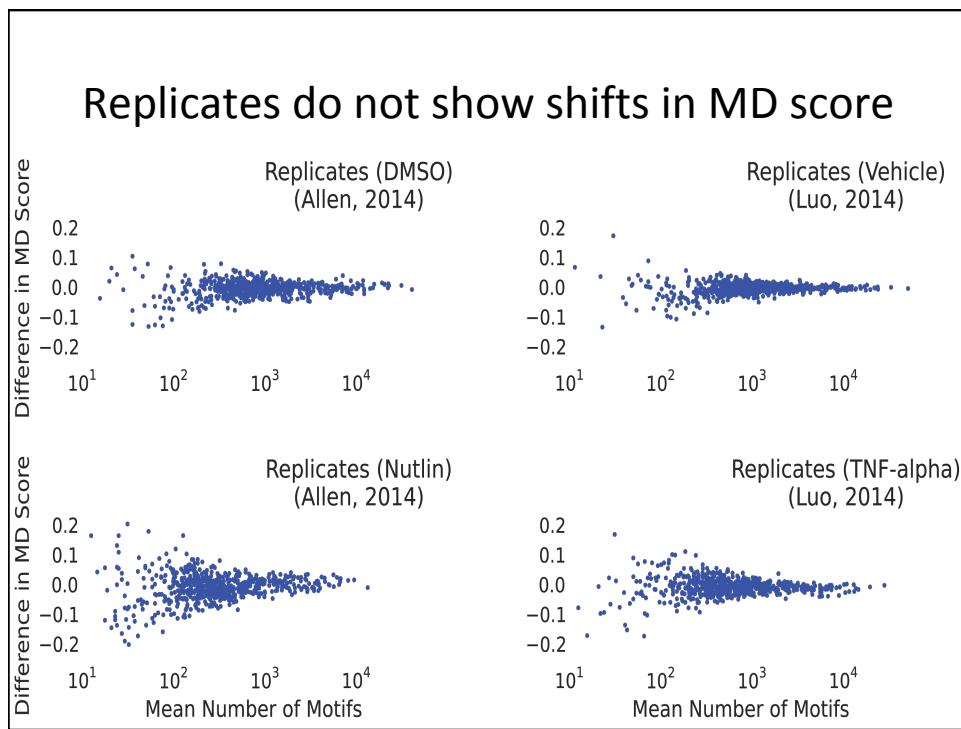
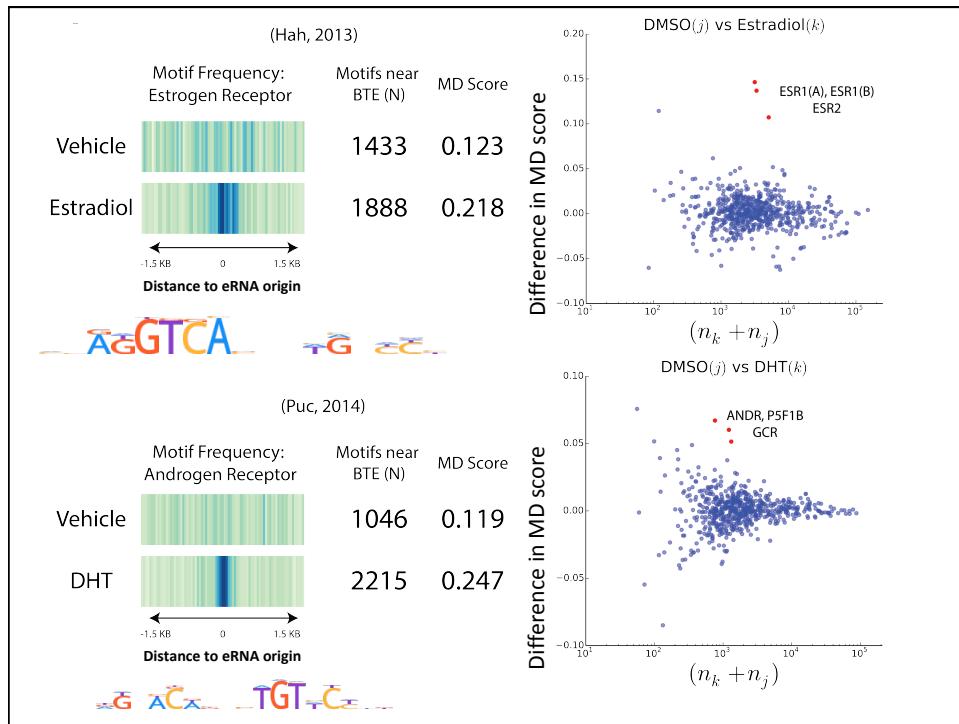


## Differential MD score is driven by eRNAs, not changes in promoters

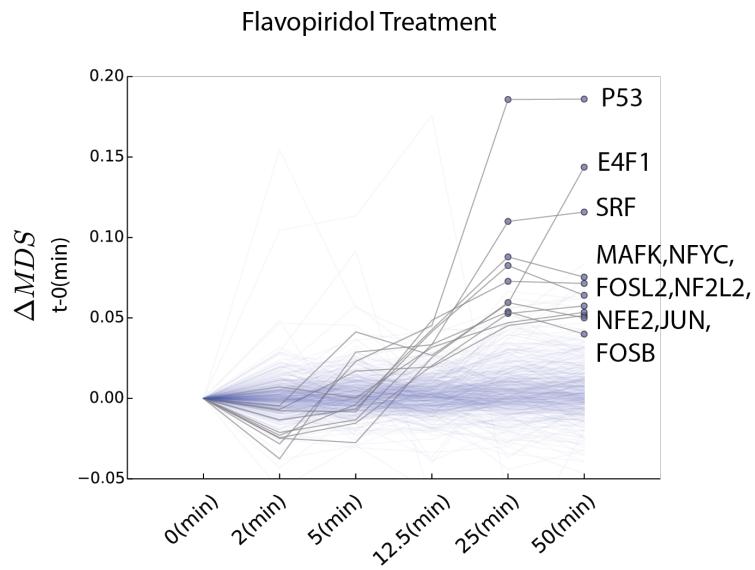


## Works for other stimuli beyond p53

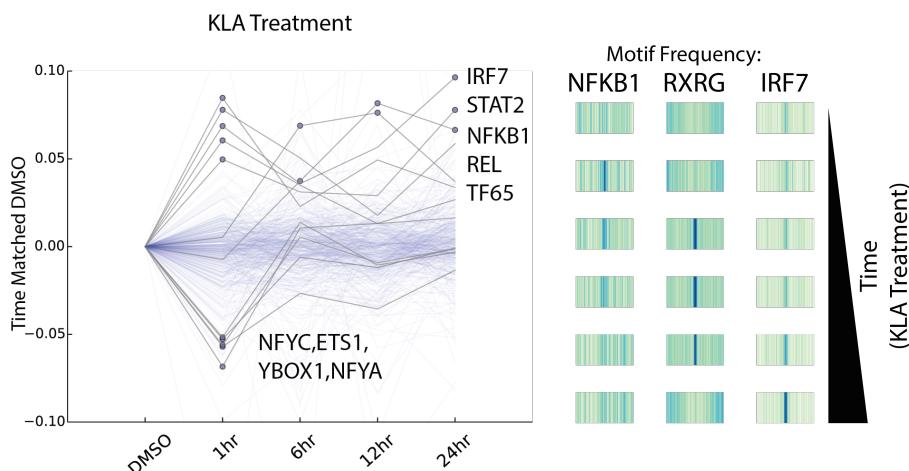




MD-scores can detect rapid changes.



And temporally order the regulatory network response.



# TF-binding motifs localize with eRNA (Tfit) predictions\*

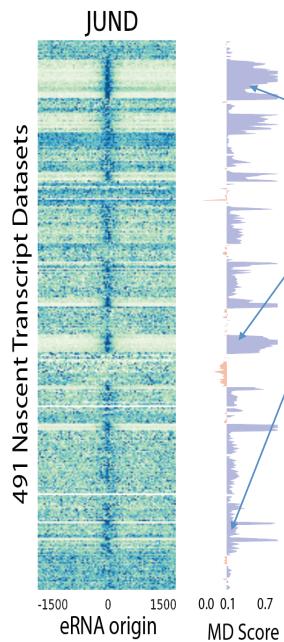
- Motif models strongly enriched with eRNAs
- Motif models lacking enrichment**
- Motif models depleted at eRNAs
- Motif models flank or are displaced from eRNA origins

Single Dataset (GRO-cap, K562)

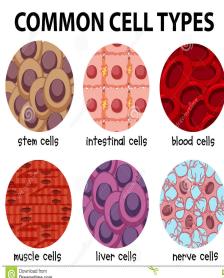
641 TF Motif Models  
(HOCOMOCO)

J. Azofeifa, Mary Allen, Josephina Hendrix, Tim Read, Jonathan Rubin, Robin Dowell  
eRNA expression predicts transcription factor activity (2017, In review)

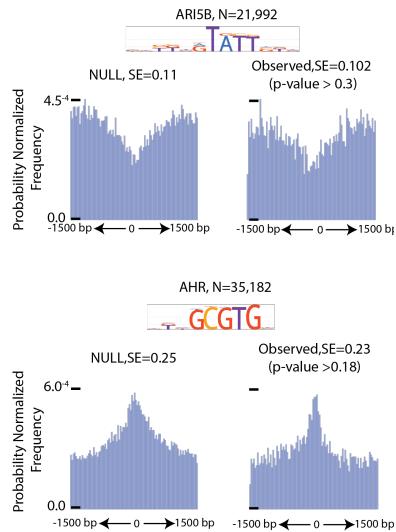
Motif to eRNA localization is extremely variable across datasets



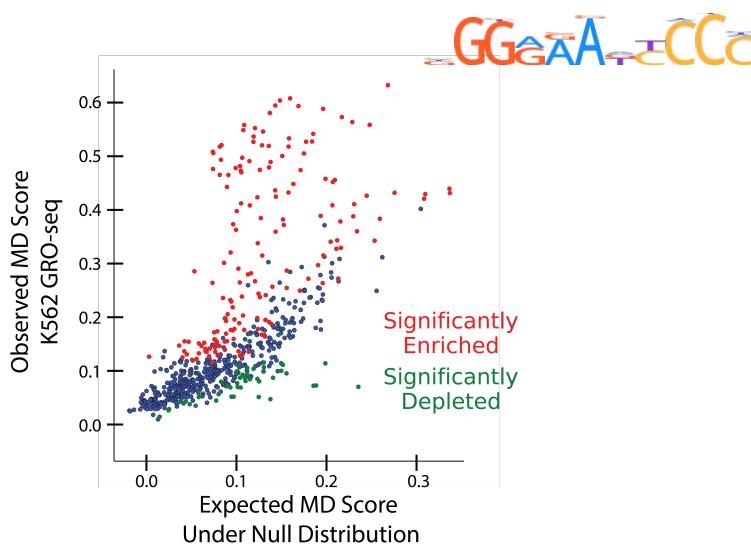
What do these datasets have in common?



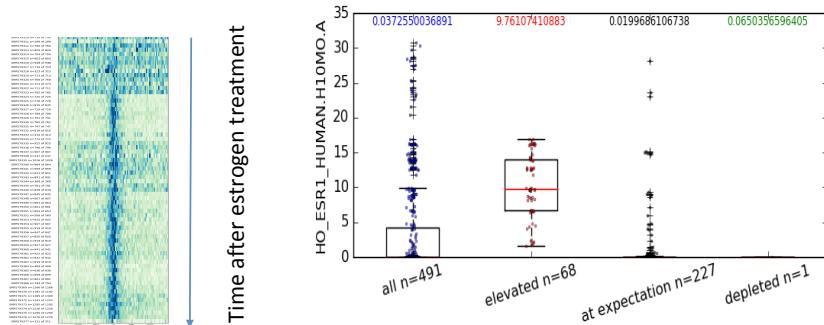
## Our expected MD-score is a function of the motif content.



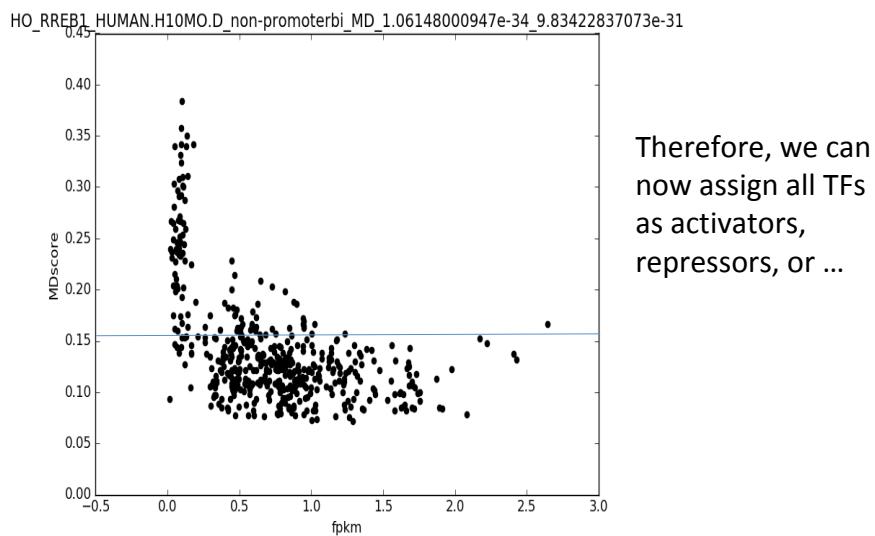
## MD score predicts which TFs are active in a given cell type



For activators, an elevated MD-score tends to correlate with elevated transcription.



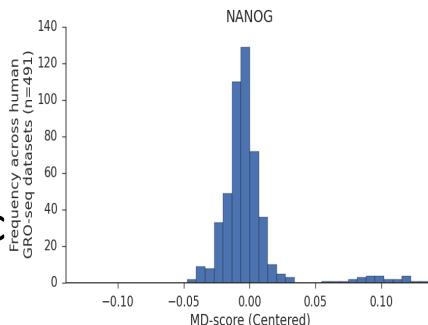
Transcriptional Repressors show high expression when MD score is depleted

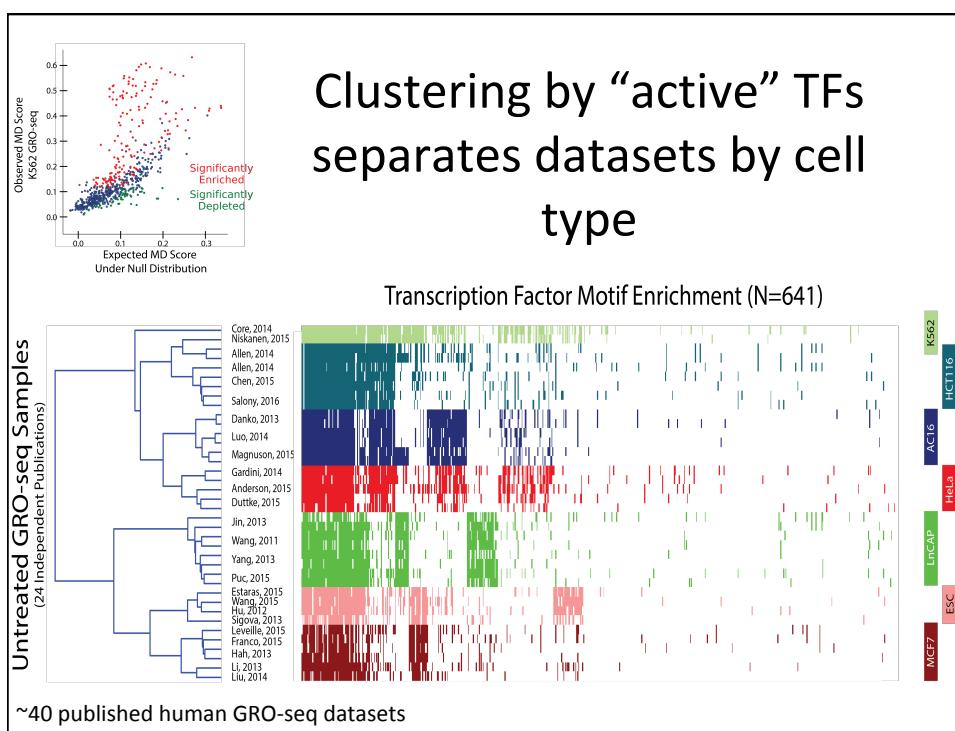
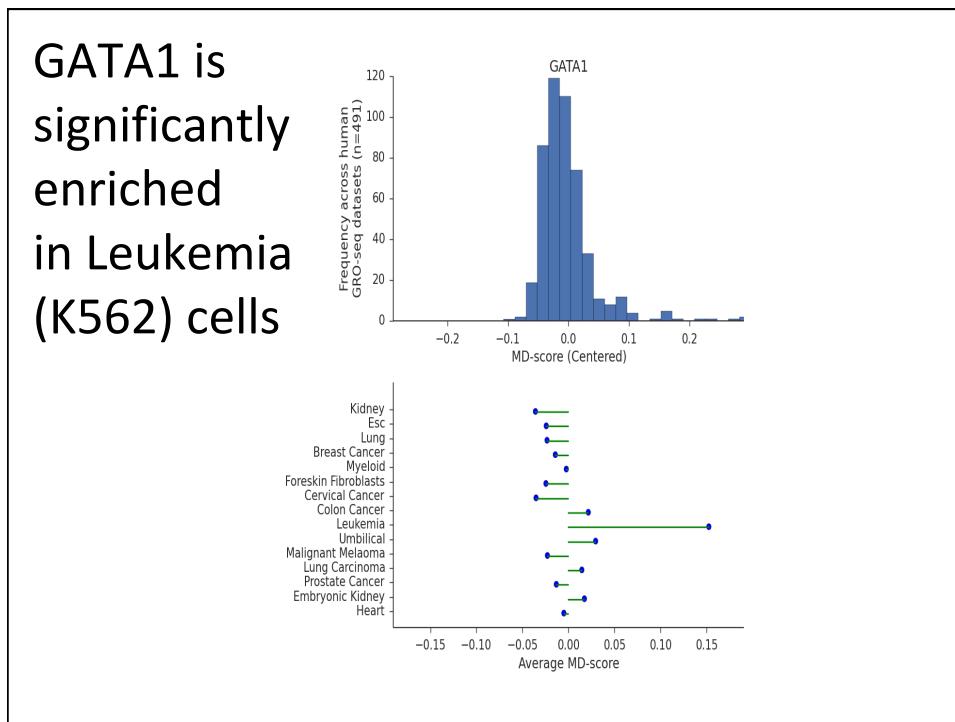


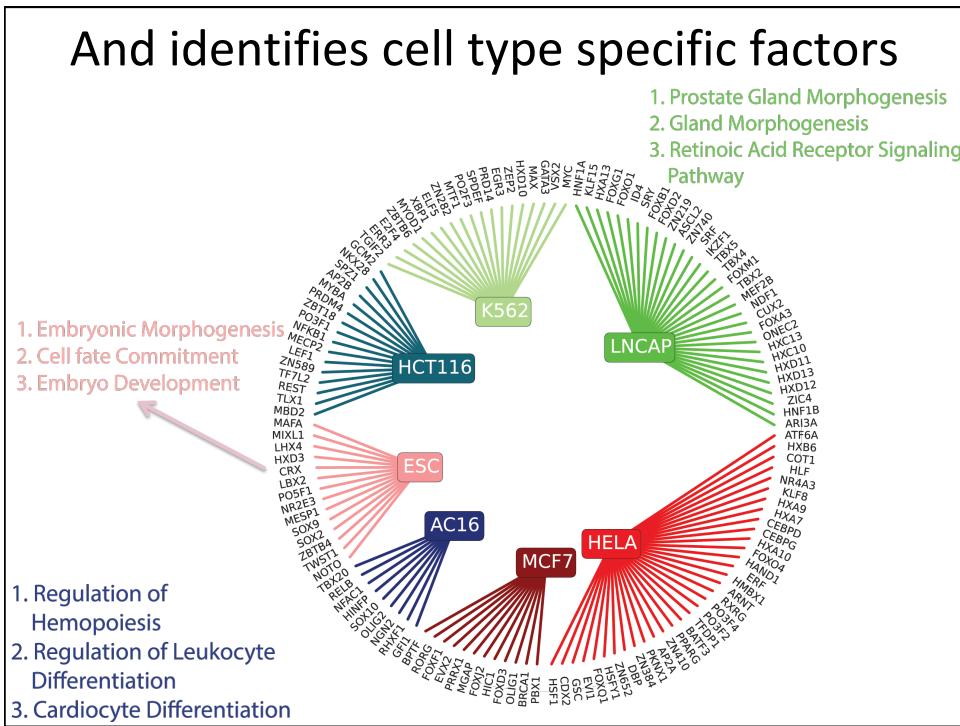
Interestingly ....

- Mutual Information appears to show stronger signals in comparisons of MD-score to transcription of the gene encoding the TF.
- This suggest that some TFs have a non-binary relationship between their transcriptional regulation and activity. (i.e. they could actually be both).

NANOG is significantly enriched in embryonic stem cells

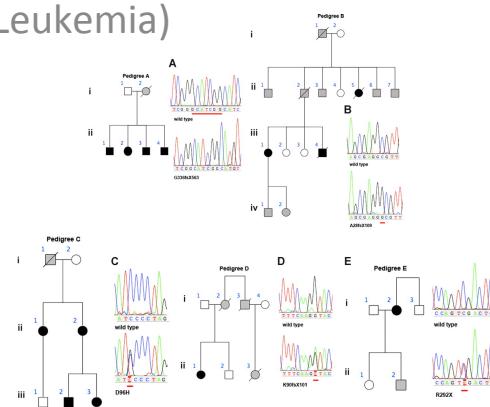


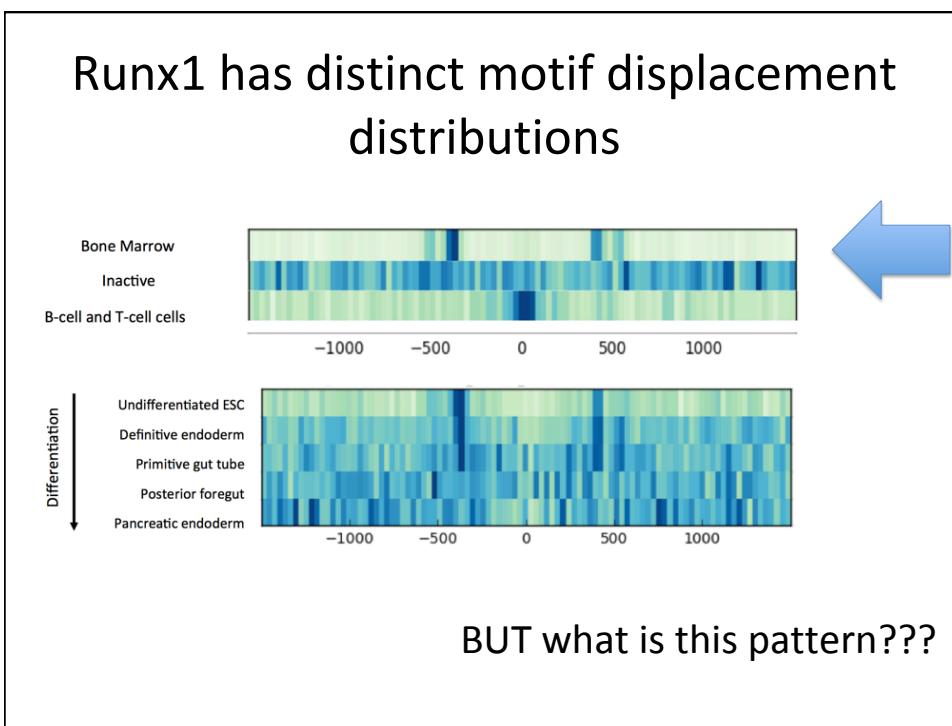
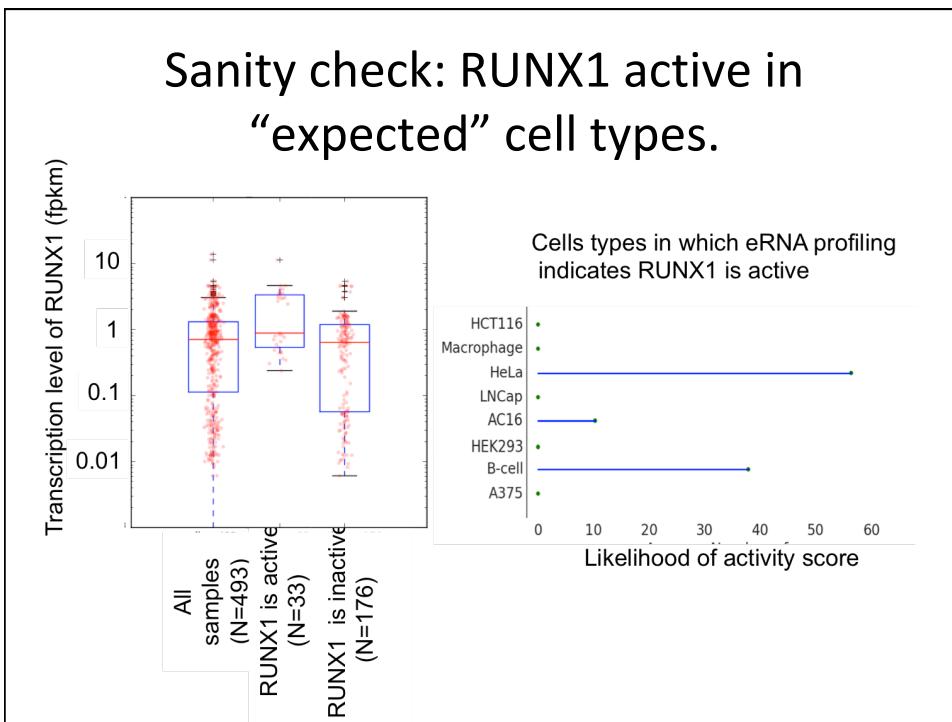




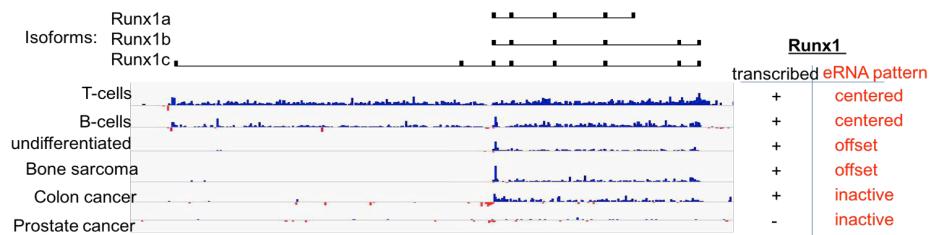
**Assume we can tell when a TF is active:  
Why does that matter?**

Subplot: Runx1 a gene mutated in  
Familial Platelet disorder and  
AML(Leukemia)

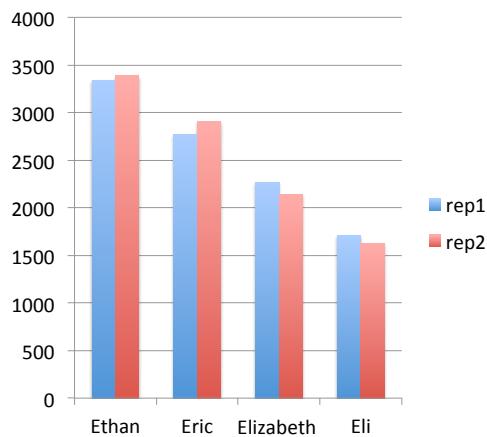




## Distinct distributions correspond to distinct isoform (5' end) usage



## Another interesting subplot: Individual with Down syndrome have more eRNAs



Indicates that the total eRNA count has increased in Ethan, an individual with Down syndrome.

## So are TF changing in Down syndrome?

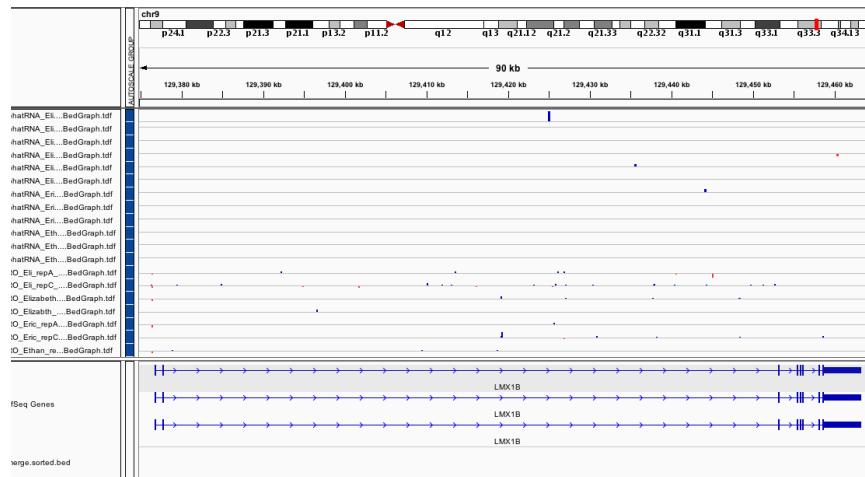
	Eli	Eliz	Eric	Ethan
Eli	0	2	6	10
Eliz	2	0	2	9
Eric	6	2	0	22
Ethan	10	9	22	0

By MD-score analysis, it looks like a LOT of TFs change in response to Down syndrome.

## Several TF are consistently up-regulated in Ethan

- RAX2-mutations in RAX2 are linked to macular degeneration
- GBX1
- LHX3
- LHX4
- PAX4 -has the potential to function as a tumor suppressor in human melanoma
- LMX1B

But these TFs are not expressed



However, they have a similar core motif.

- Protein expression without RNA expression?



- Mystery TF? (Not in Hocomoco)

Caveat: This could be something specific to Ethan