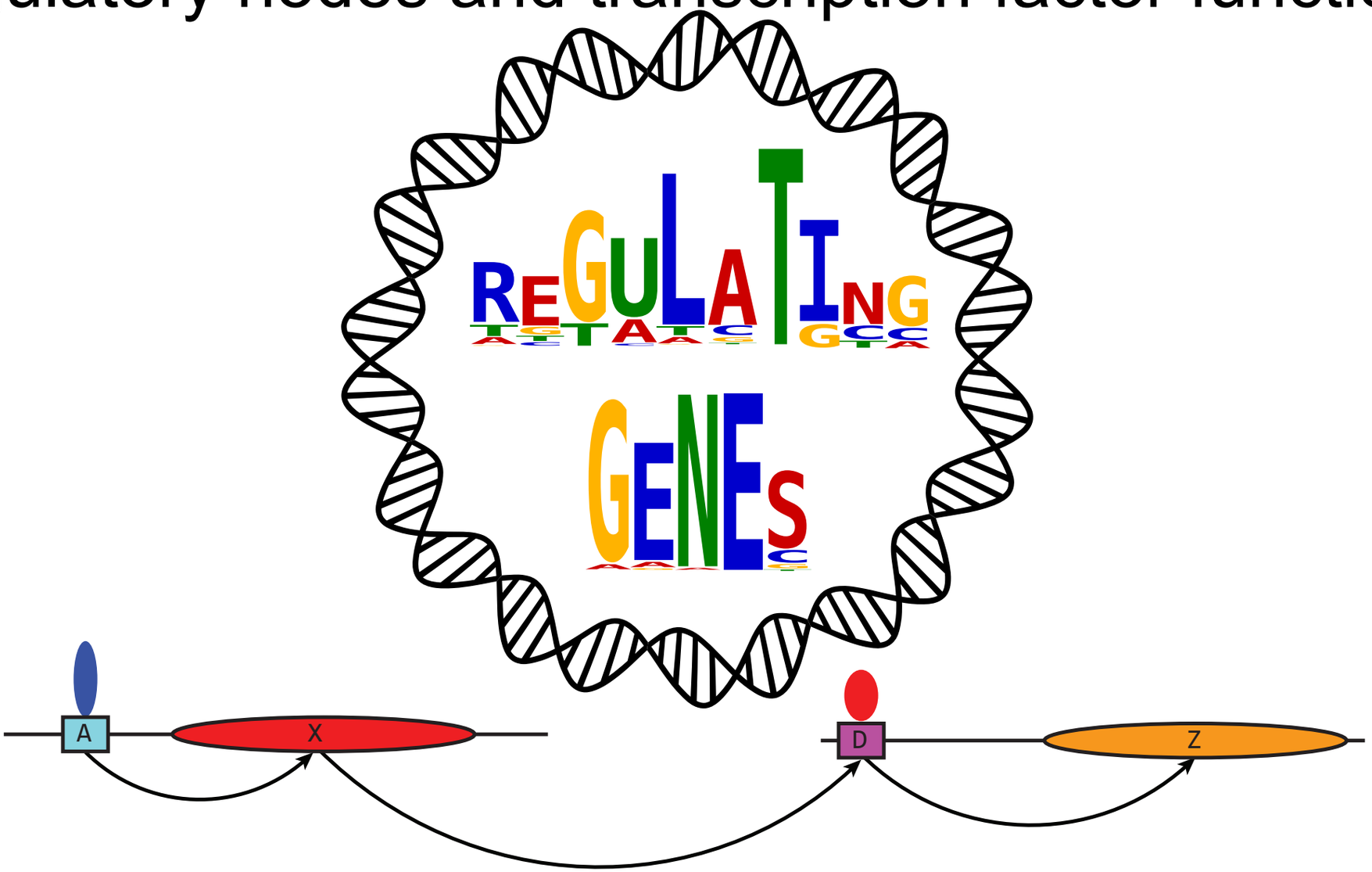
Kinetic networks of differentiation cascades identify key regulatory nodes and transcription factor functions



UCONN
HEALTH

Center for Cell Analysis & Modeling

Genetics and Genome Sciences Department

GUERTIN LAB

Michael Guertin

Transcription Factor Nobel Prizes
(Question-driven science and well-designed screens)

Q: Which Drosophila genes are important for developmental patterning?

1995: Classic genetics (perturb, observe, map) identifies proteins (>50% were TFs!) critical for *Drosophila* development

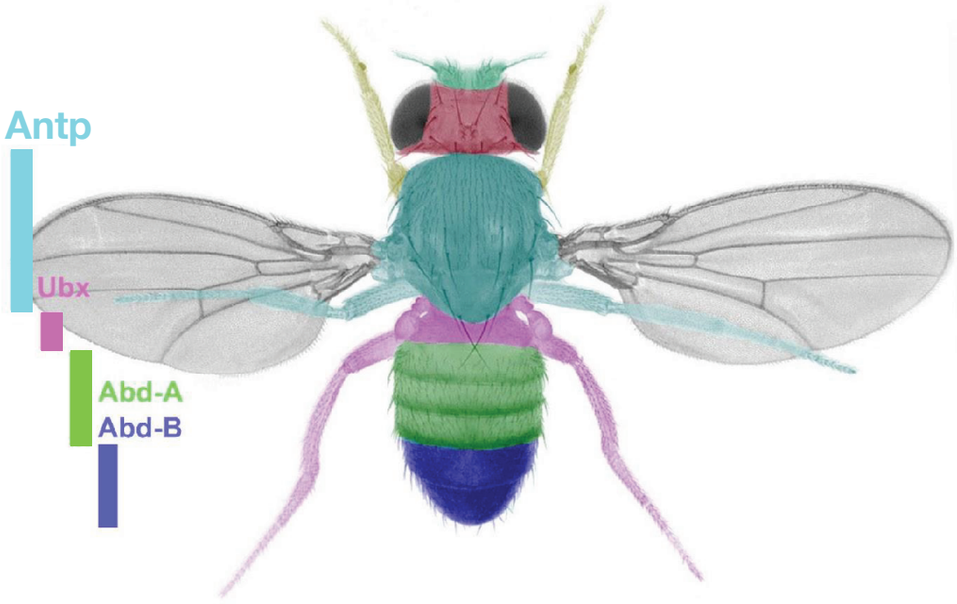
Q: Can we reprogram differentiated cells to a pluripotent state?

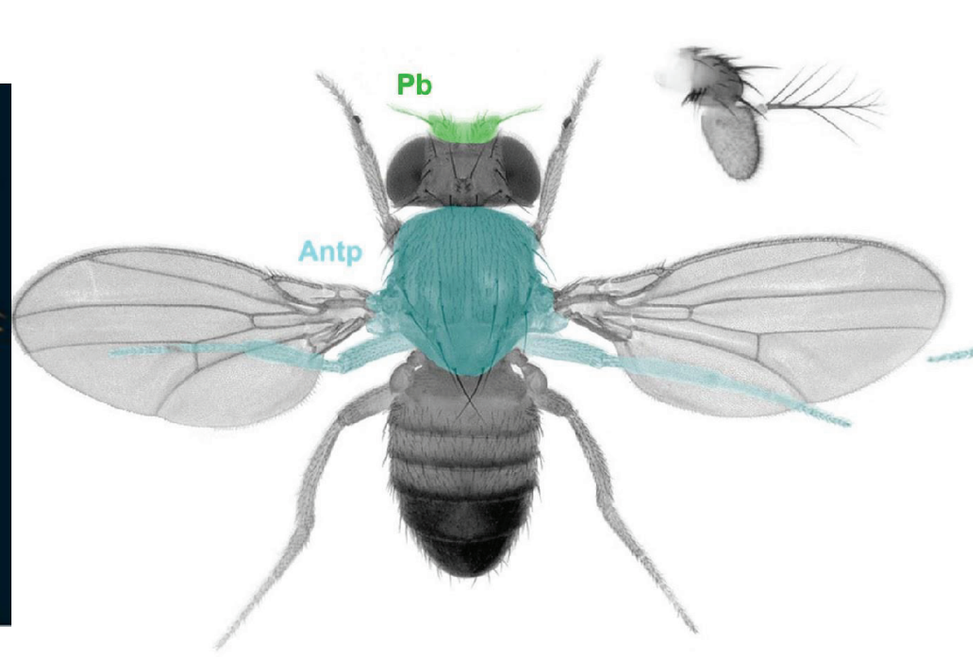
2012: Brute force candidate gene screen. Transduce 24 genes that are specifically expressed in embryonic stem cells into fibroblasts. Systematically narrow down the list: Oct3/4, Klf4, Sox2, and c-Myc (all TFs!)

Transcription dysregulation alters developmental patterning

pseudocolored flies: Justin Crocker, Ed Lewis, Nicolas Gompel, and Welcome Bender

Classic Genetics: Perturb and Map





pseudocolored flies: Justin Crocker, Ed Lewis, Nicolas Gompel, and Welcome Bender

pseudocolored SEM heads: Jürgen Berger

Classic genetics (**perturb, observe, map**) found that Transcription Factors control developmental patterning

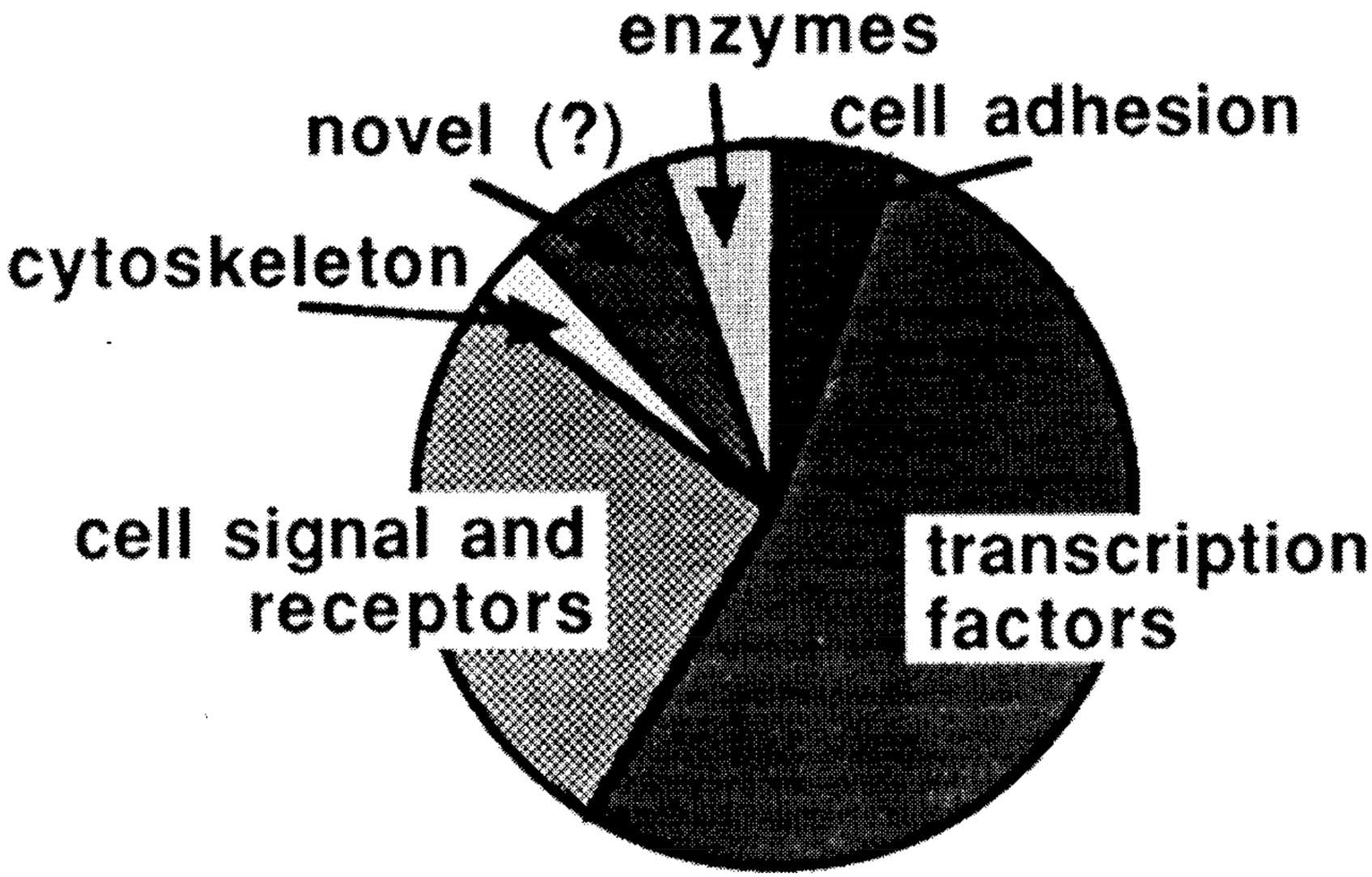
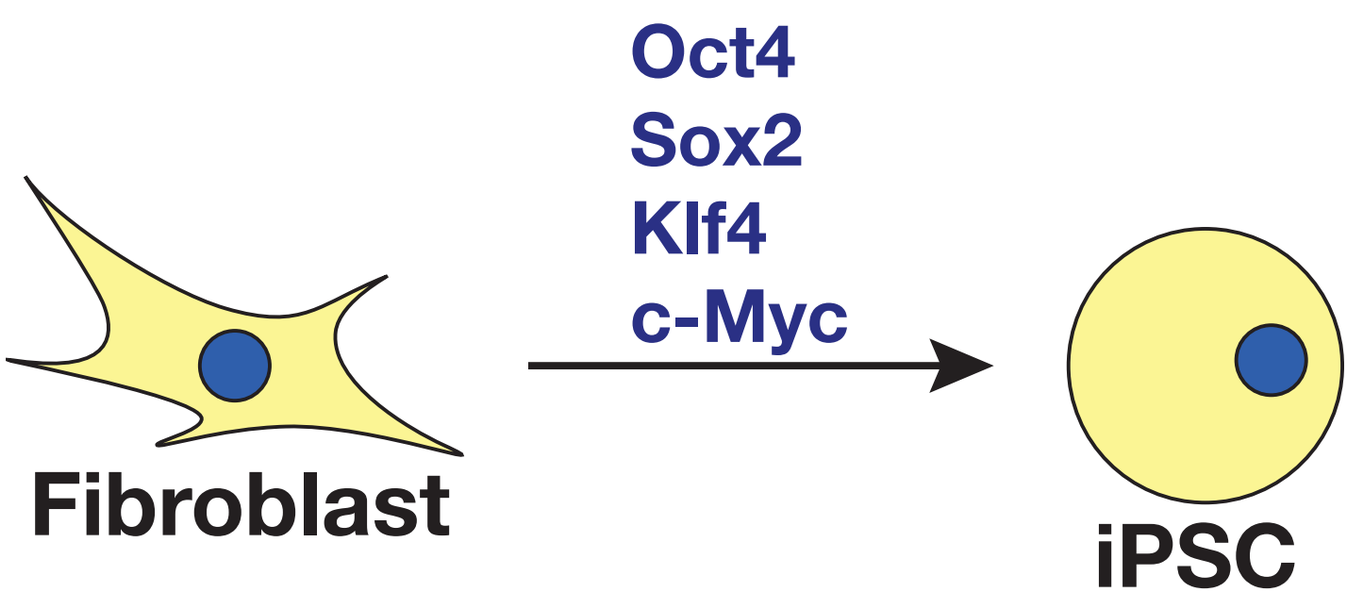


Figure 3. *Cellular Function of Heidelberg Mutations.* Based on the sequence of 75 cloned genes, most of the loci identified in Heidelberg encode transcription factors, or cell signals and receptors.

Eric Weischaus, Nobel Lecture 1995
Prize shared with Christiane Nüsslein-Volhard & Ed Lewis

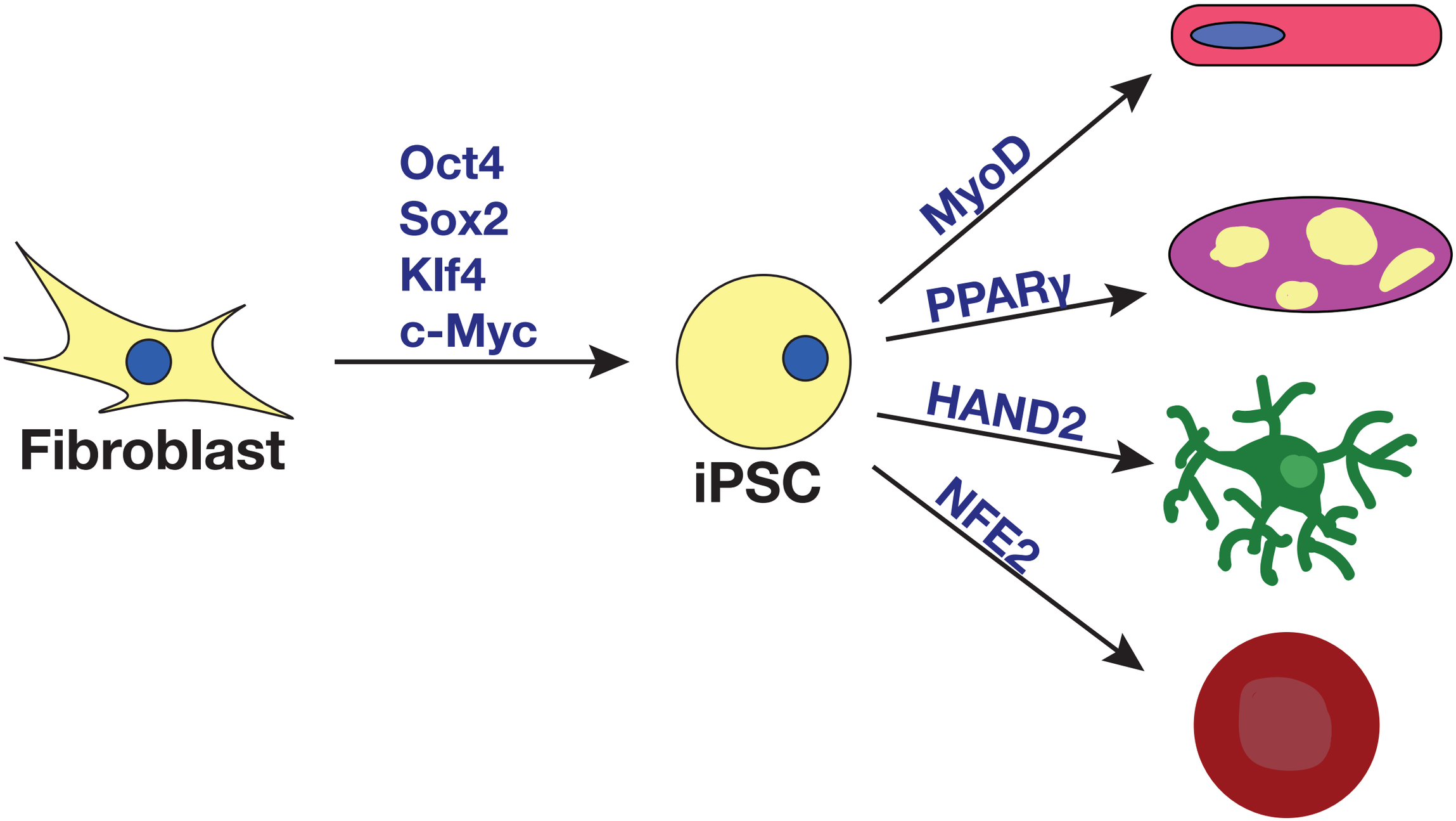
Transcription factors drive changes in cell identity



Brute force screening approach:
express many genes in combination until we get iPSCs

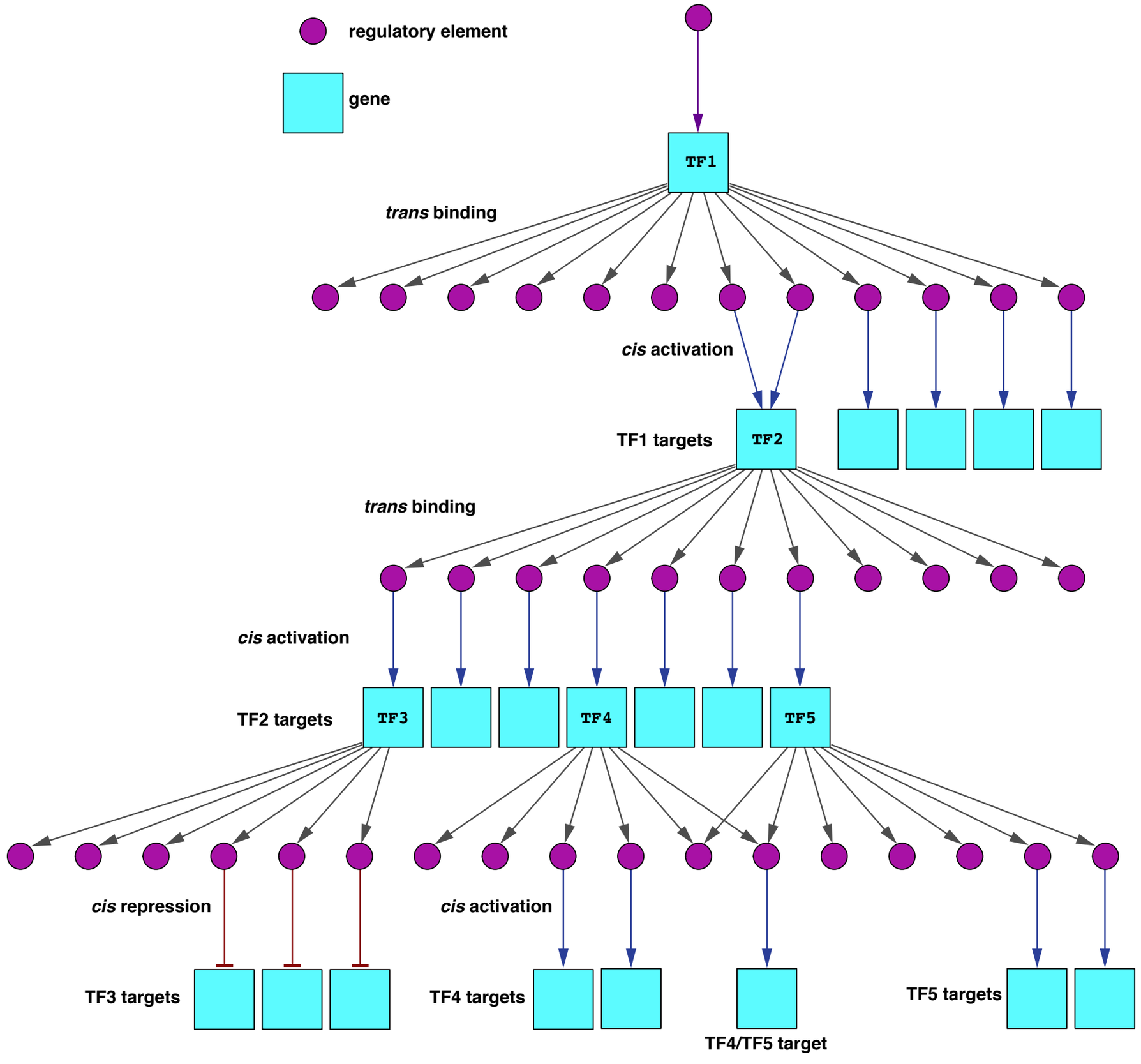
Takahashi & Yamanaka, Cell 2006
2012 Nobel Prize shared with John Gurdon

Transcription factors drive changes in cell identity

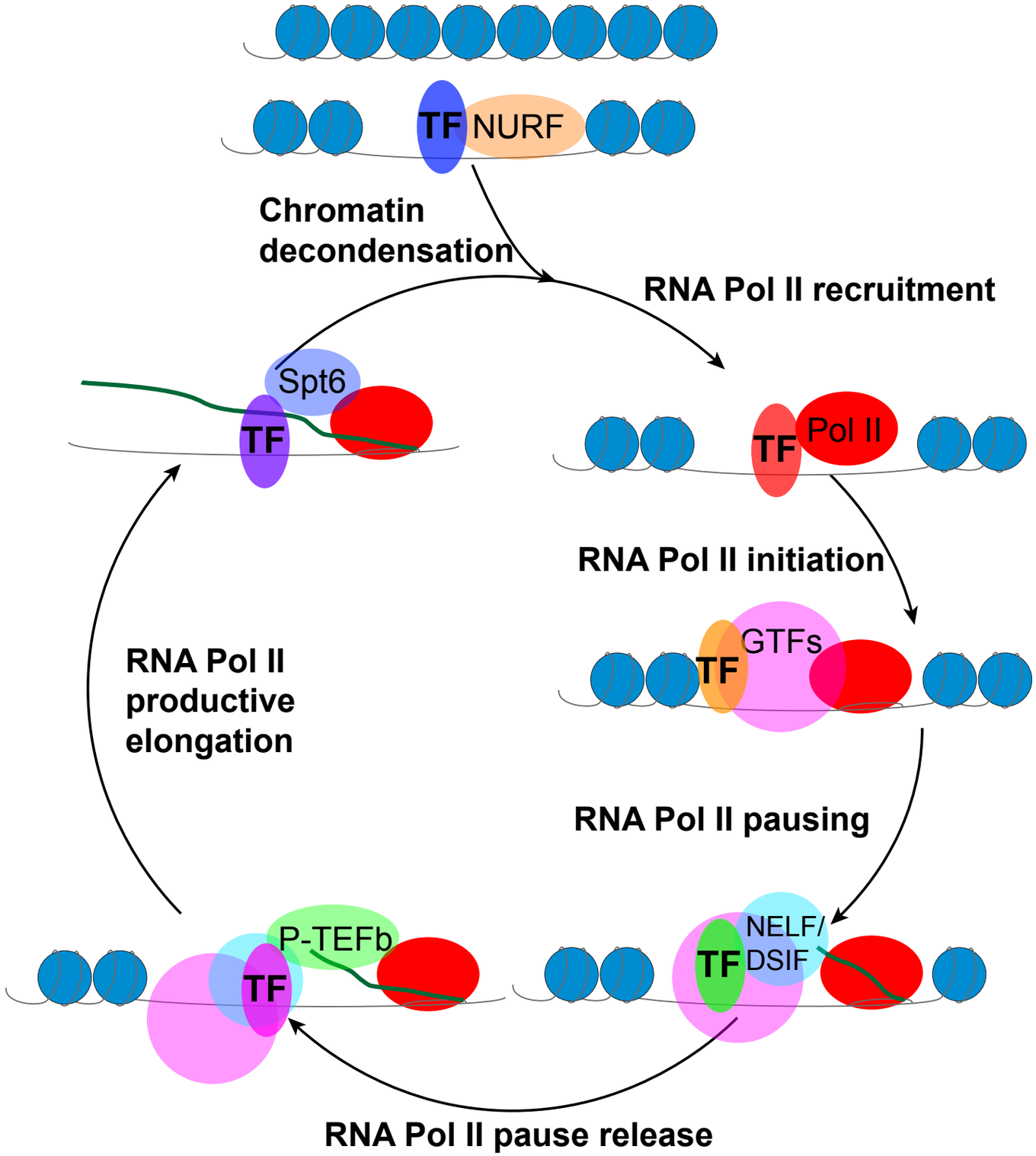


Takahashi & Yamanaka, Cell 2006
2012 Nobel Prize shared with John Gurdon

Which TFs and regulatory elements are important in regulatory cascades?



What are the step(s) in transcription that each key TF regulates?



Possible approaches and considerations

Possible approaches and considerations

- ChIP-seq all the TFs throughout the cascade, but which TFs?

Possible approaches and considerations

- ChIP-seq all the TFs throughout the cascade, but which TFs?
- ATAC-seq gives us putative regulatory elements, but which TFs are binding/dissociating during the cascade?

Possible approaches and considerations

- ChIP-seq all the TFs throughout the cascade, but which TFs?

- ATAC-seq gives us putative regulatory elements, but which TFs are binding/dissociating during the cascade?

- RNA-seq quantifies changes in gene expression, but detection is delayed depending on RNA turnover rate and accumulation rate over basal RNA levels.

Possible approaches and considerations

- ChIP-seq all the TFs throughout the cascade, but which TFs?

- ATAC-seq gives us putative regulatory elements, but which TFs are binding/dissociating during the cascade?

- RNA-seq quantifies changes in gene expression, but detection is delayed depending on RNA turnover rate and accumulation rate over basal RNA levels.

- Which time points do we choose?

Possible approaches and considerations

- ChIP-seq all the TFs throughout the cascade, but which TFs?

- ATAC-seq gives us putative regulatory elements, but which TFs are binding/dissociating during the cascade?

- RNA-seq quantifies changes in gene expression, but detection is delayed depending on RNA turnover rate and accumulation rate over basal RNA levels.

- Which time points do we choose?

- Which models can we genetically manipulate to validate candidate TFs and elements?

Outline

- Question-driven science and well-designed studies

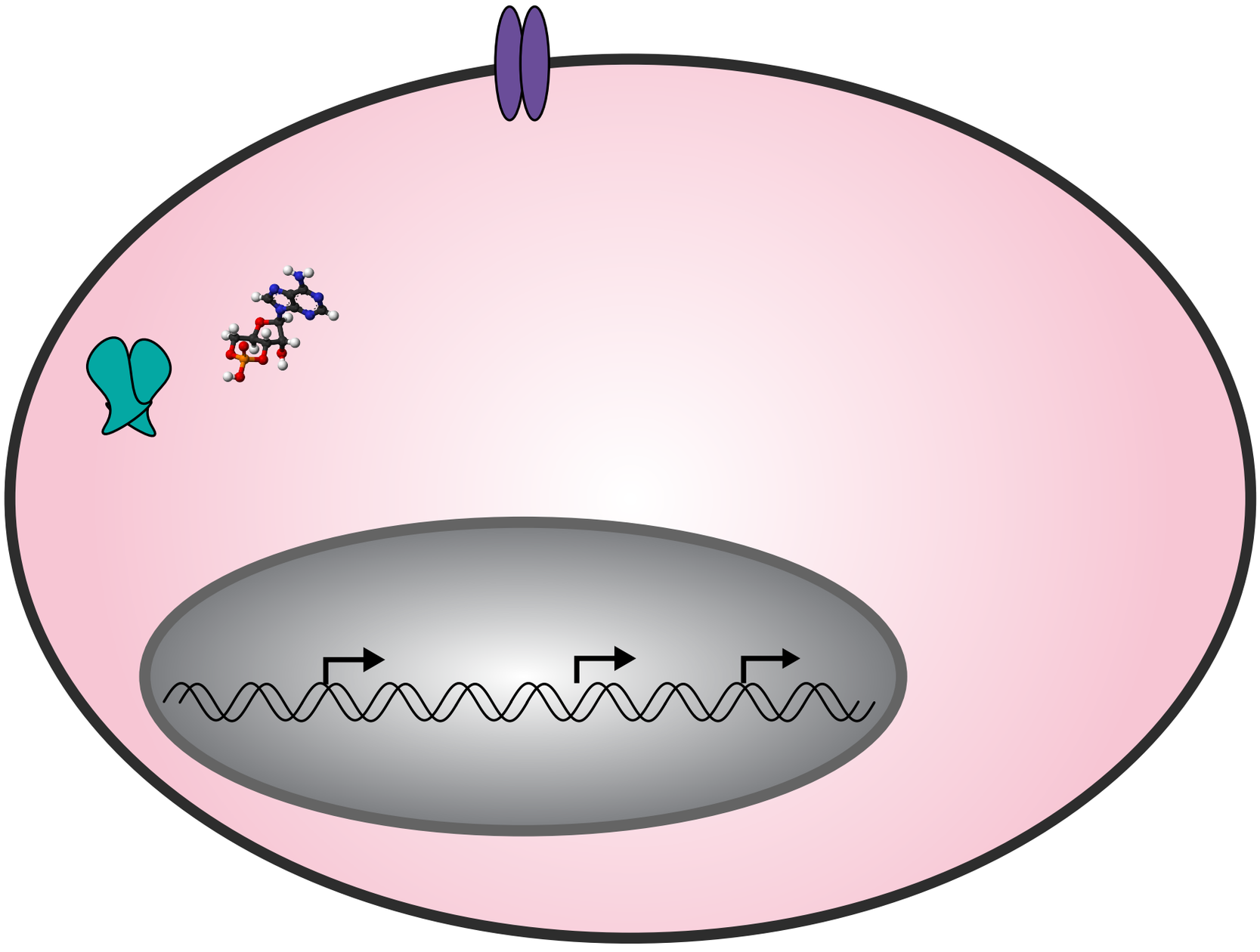
- Mechanistic insights from genomics-derived gene regulatory networks

- Rapid protein degradation to study transcription factor interaction

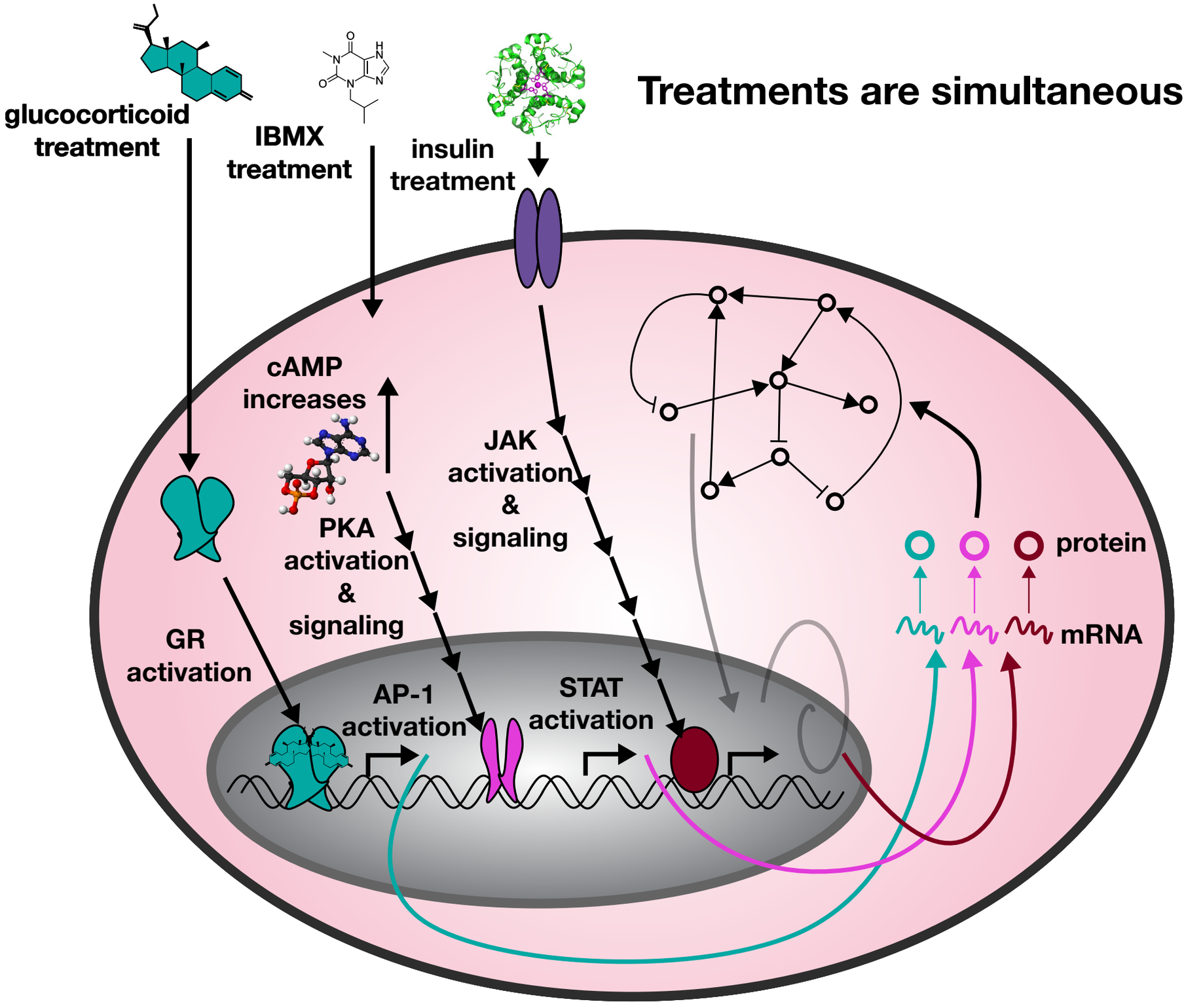
Kinetic networks identify key regulatory nodes and transcription factor functions in early adipogenesis

Arun Dutta

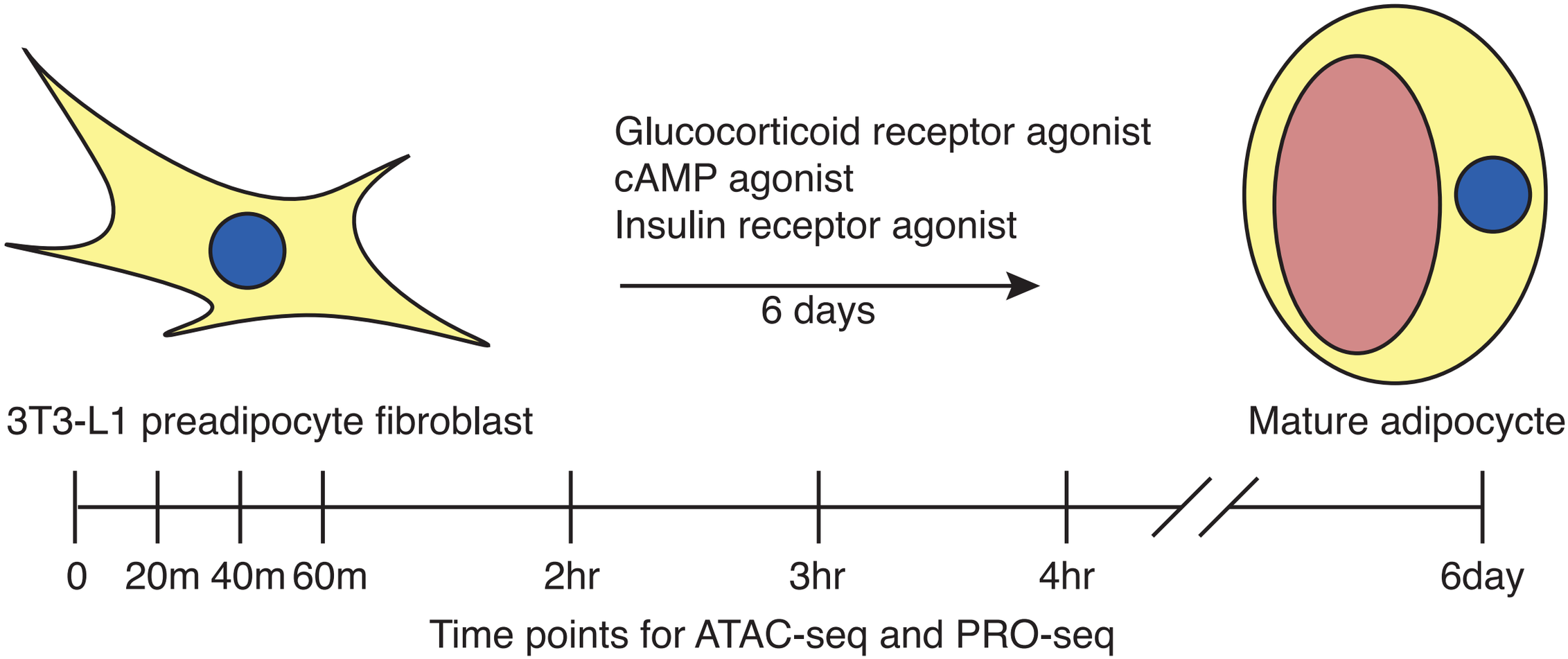
Adipogenesis of 3T3-L1 cells



Adipogenesis of 3T3-L1 cells



Experimental Design



Transposase Hypersensitivity (ATAC-seq)

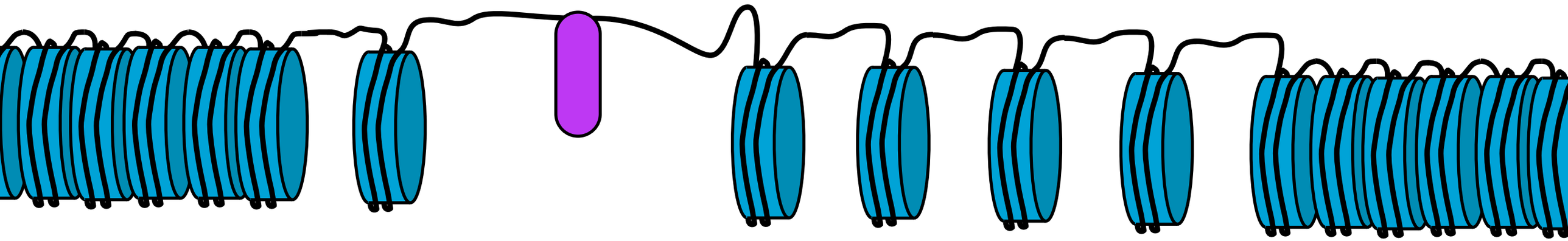
Transposase Hypersensitivity (ATAC-seq)

- A general measure of chromatin structure.

Transposase Hypersensitivity (ATAC-seq)

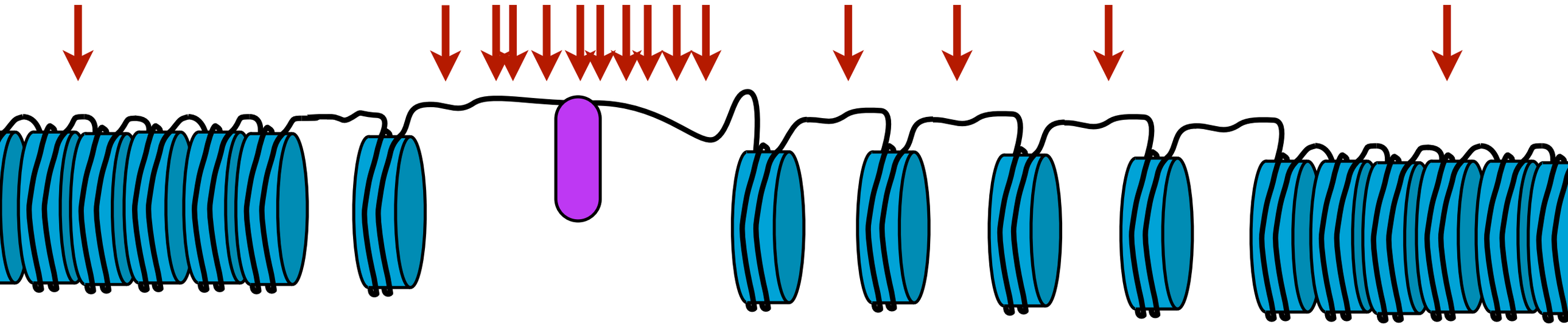
- A general measure of chromatin structure.
  - ATAC peaks and DNA motifs can be used to infer TF binding (i.e. if chromatin is accessible and contains a binding sequence, then a TF may be bound)

Transposase Hypersensitivity (ATAC-seq)



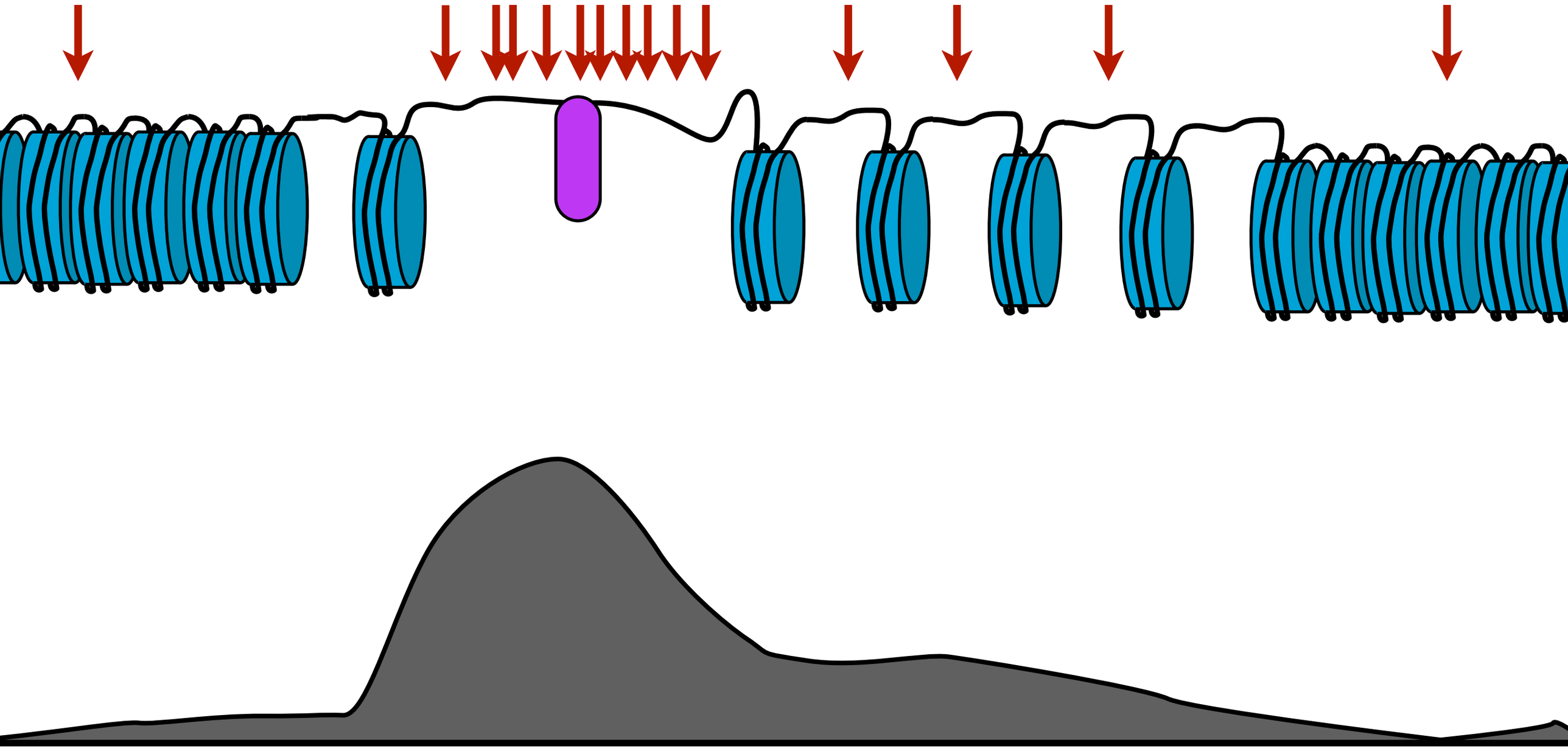
Buenrostro, et. al. Nature Methods, 2013

Transposase Hypersensitivity (ATAC-seq)



Buenrostro, et. al. Nature Methods, 2013

Transposase Hypersensitivity (ATAC-seq)

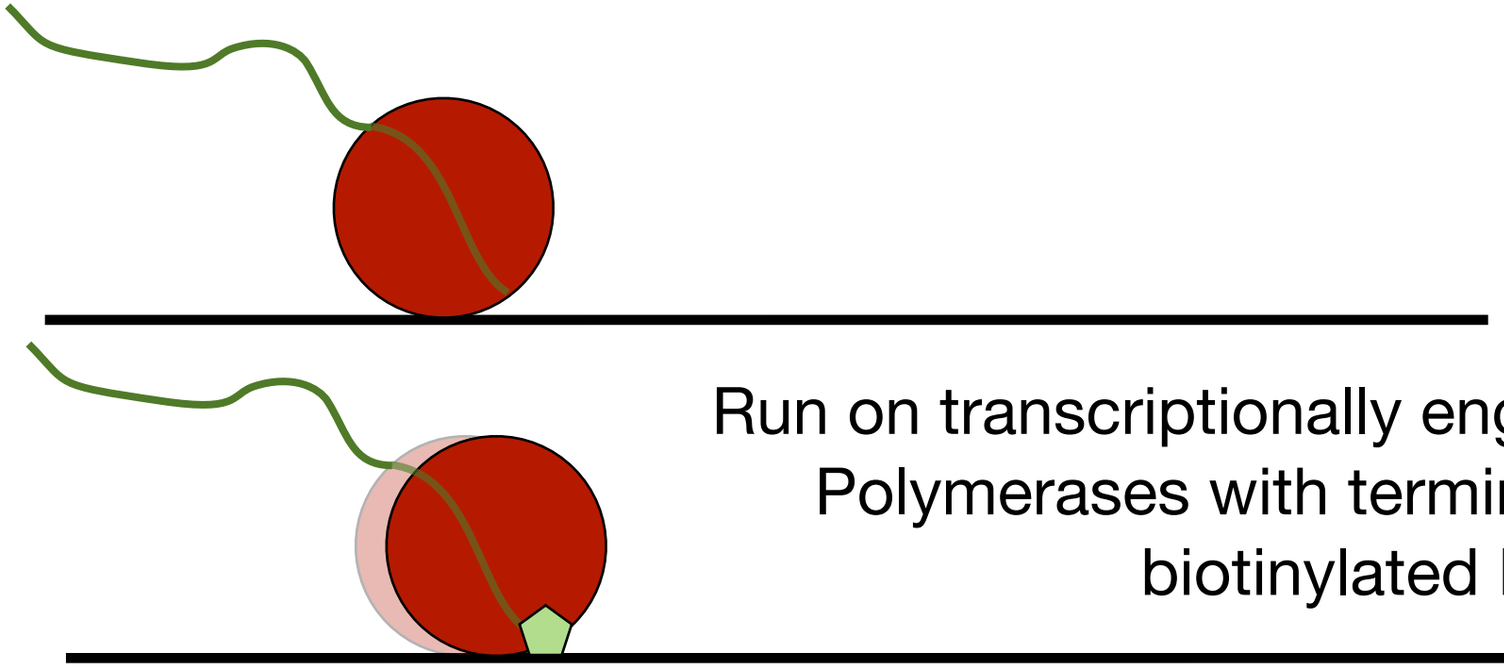


Buenrostro, et. al. Nature Methods, 2013

PRO-seq detects nascent RNA

Kwak, et. al. Science, 2013

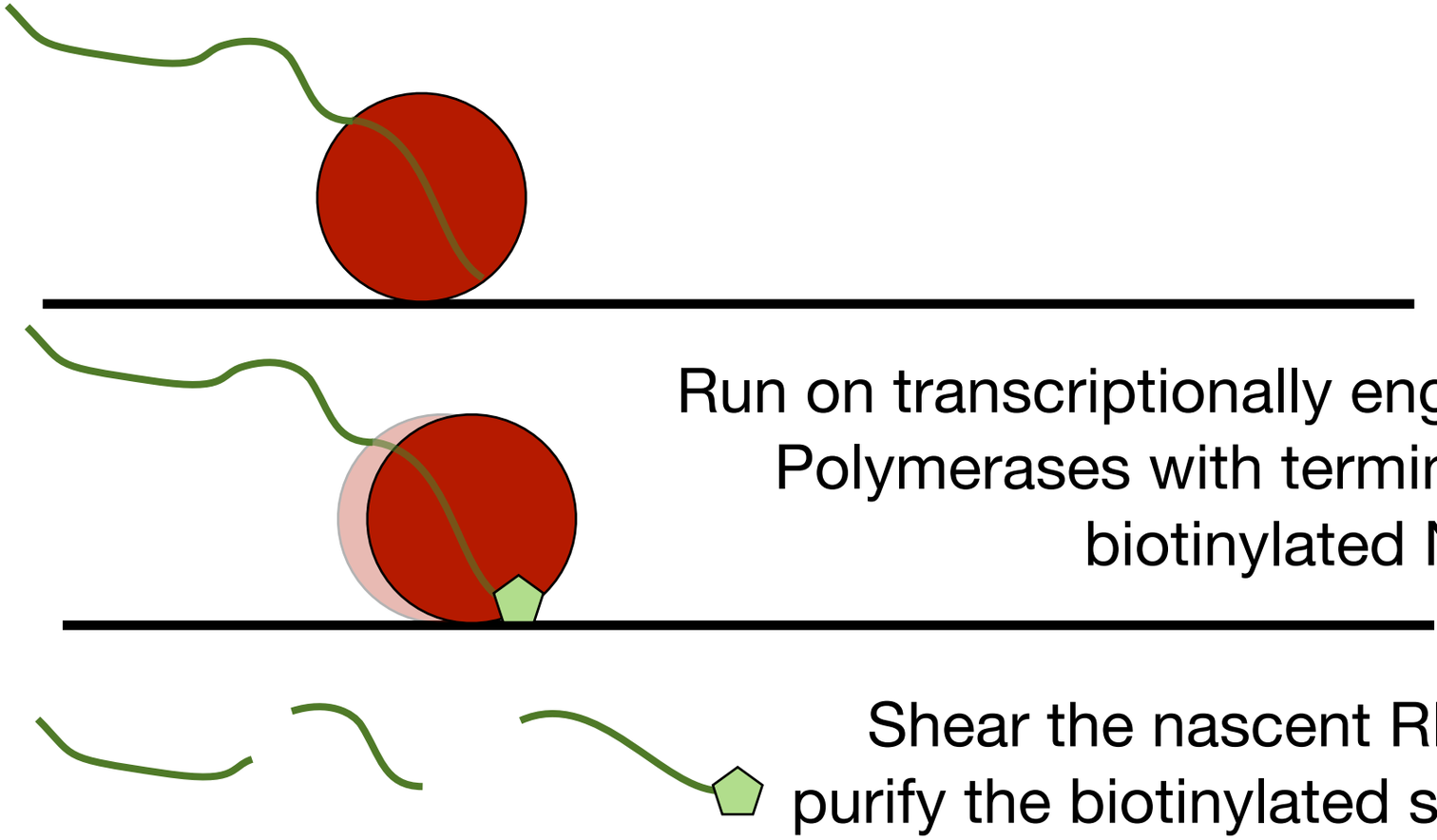
Precision Genomic-Run On (PRO-seq)



Run on transcriptionally engaged Polymerases with terminating biotinylated NTPs.

Kwak, et. al. Science, 2013

Precision Genomic-Run On (PRO-seq)

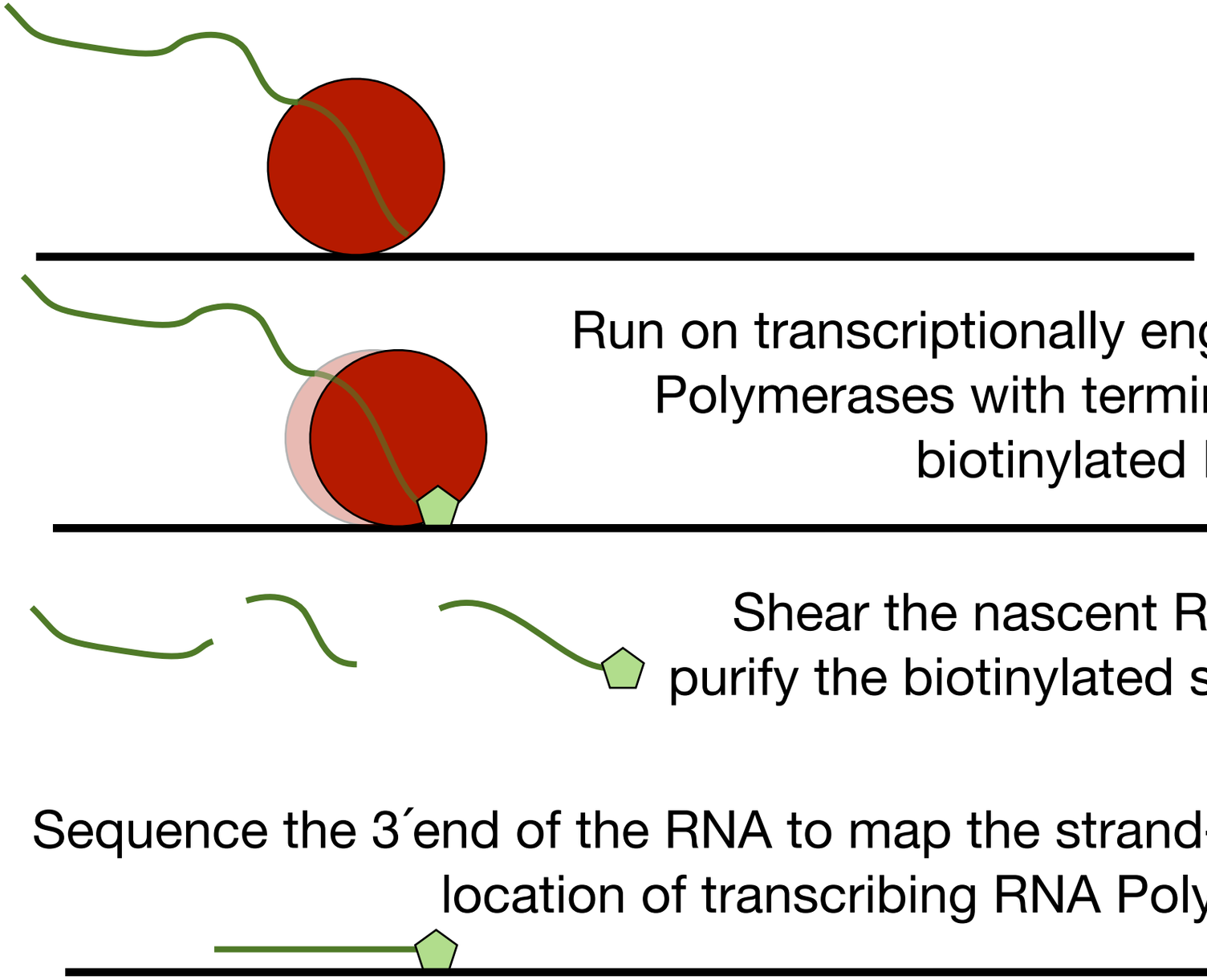


Run on transcriptionally engaged Polymerases with terminating biotinylated NTPs.

Shear the nascent RNA and purify the biotinylated species.

Kwak, et. al. Science, 2013

Precision Genomic-Run On (PRO-seq)



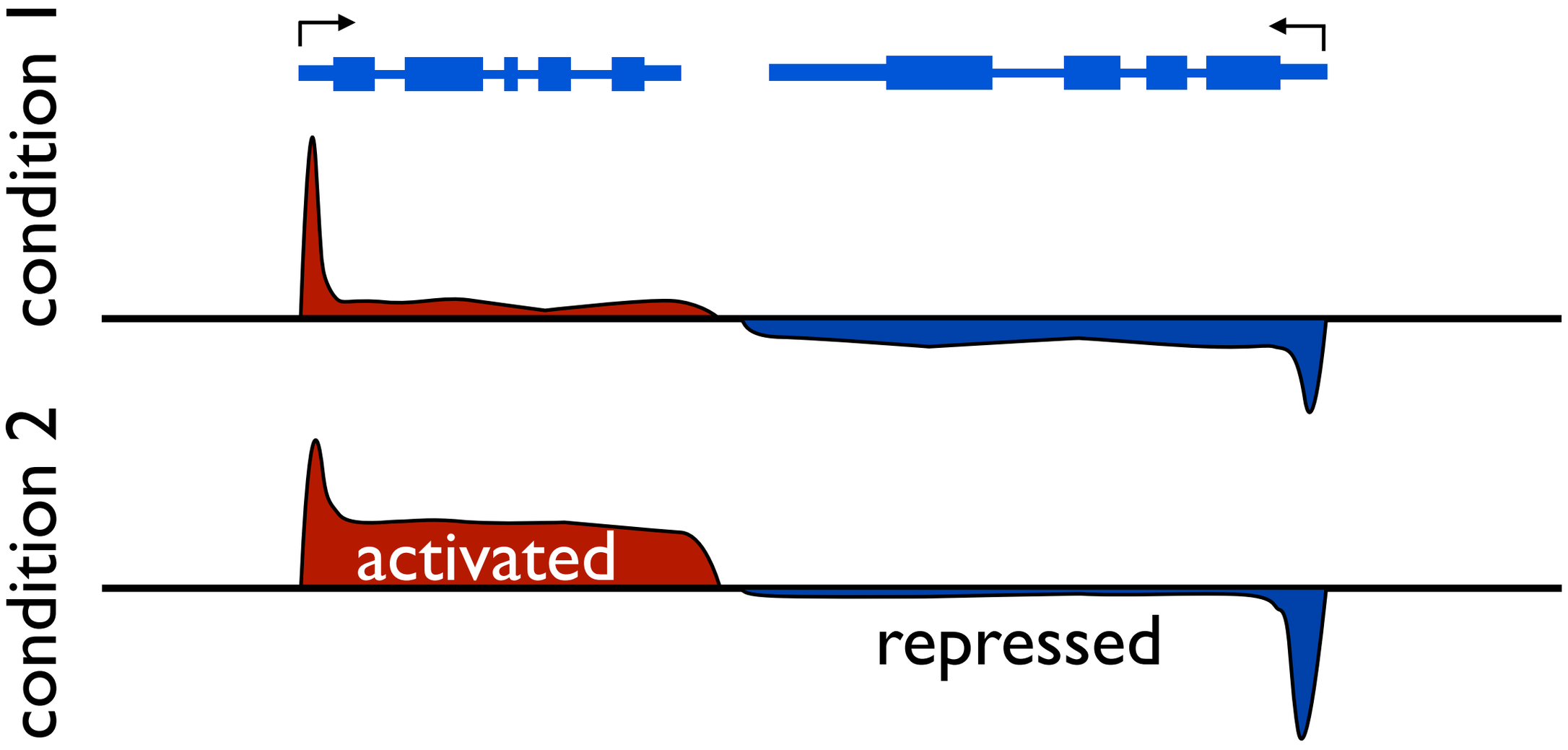
Run on transcriptionally engaged Polymerases with terminating biotinylated NTPs.

Shear the nascent RNA and purify the biotinylated species.

Sequence the 3´end of the RNA to map the strand-specific location of transcribing RNA Polymerase.

Kwak, et. al. Science, 2013

PRO-seq measures immediate responses



~200,000 peaks

ATAC signal

t0
20
40
60
120
180
240
6day

~30,000 genes

PRO signal

t0_plus
t0_minus
20min_plus
20min_minus
40min_plus
40min_minus
60min_plus
60min_minus
2hr_plus
2hr_minus
3hr_plus
3hr_minus
4hr_plus
4hr_minus
6day_plus
6day_minus

~200,000 peaks

ATAC signal

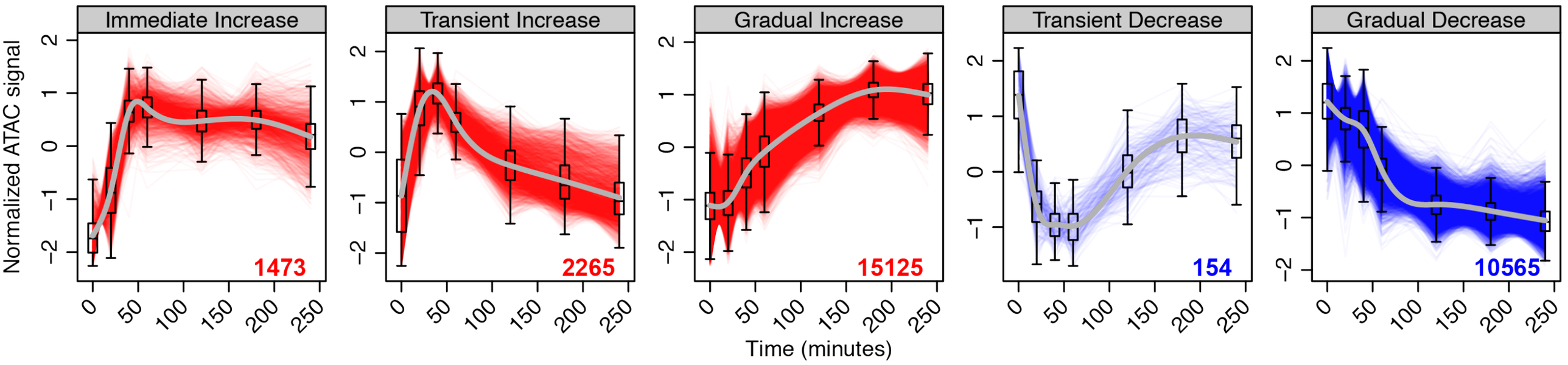
t0
20
40
60
120
180
240
6day

t0_plus
t0_minus
20min_plus
20min_minus
40min_plus
40min_minus
60min_plus
60min_minus
2hr_plus
2hr_minus
3hr_plus
3hr_minus
4hr_plus
4hr_minus
6day_plus
6day_minus

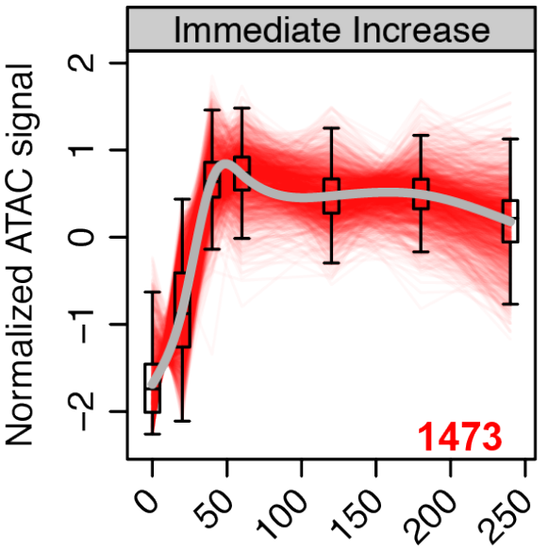
PRO signal

~30,000 genes

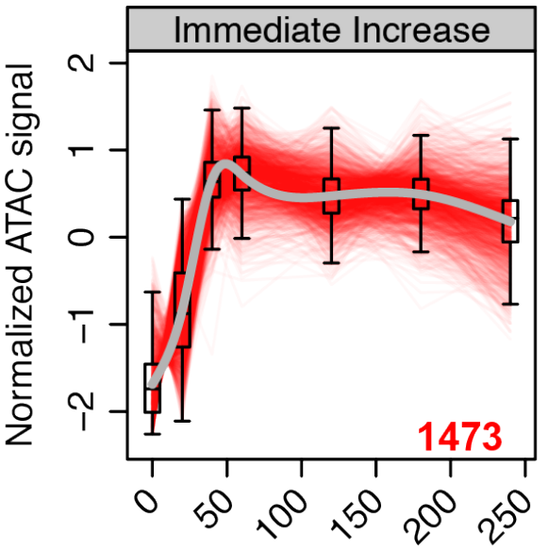
Interesting ATAC-seq peaks are dynamic over the time course

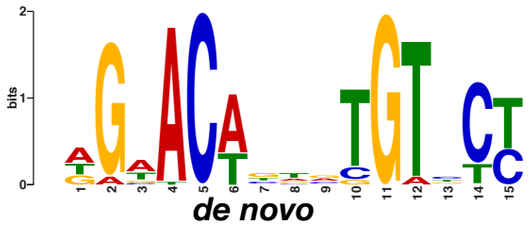


de novo motif analysis identifies enriched sequence elements within dynamic ATAC peaks

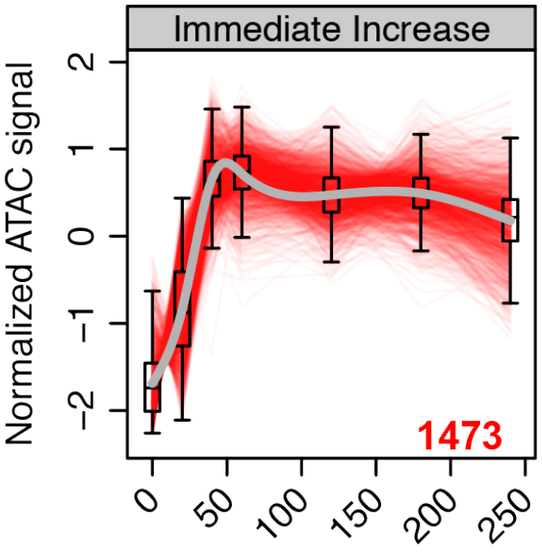


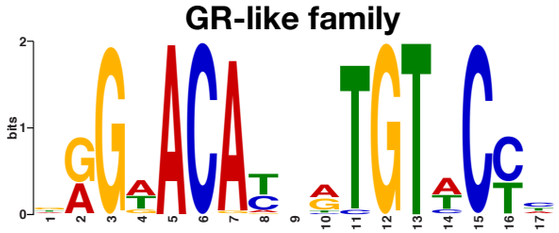
de novo motif analysis identifies enriched sequence elements within dynamic ATAC peaks

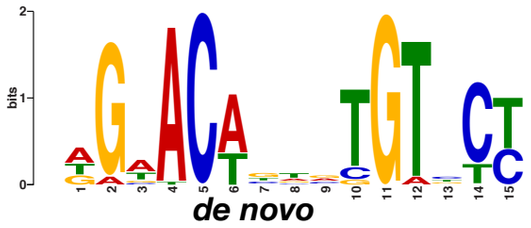


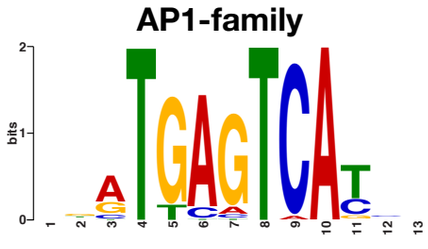


de novo motif analysis identifies enriched sequence elements within dynamic ATAC peaks

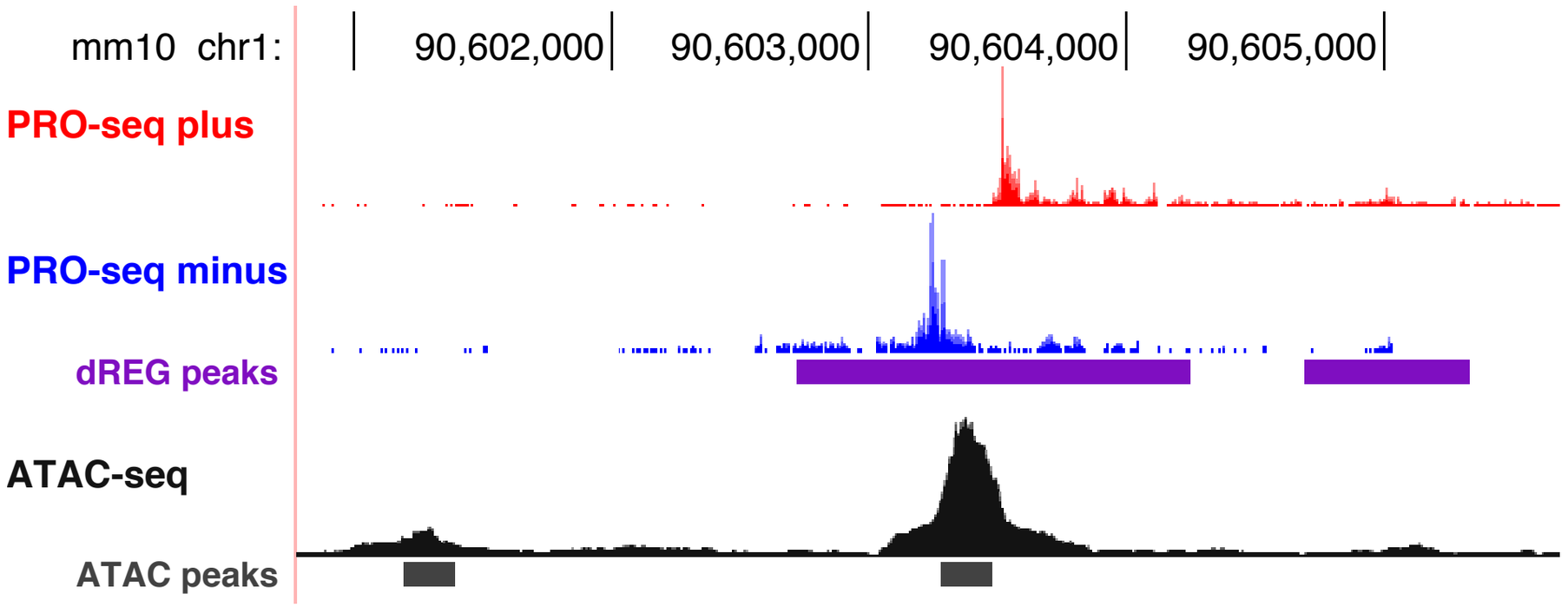


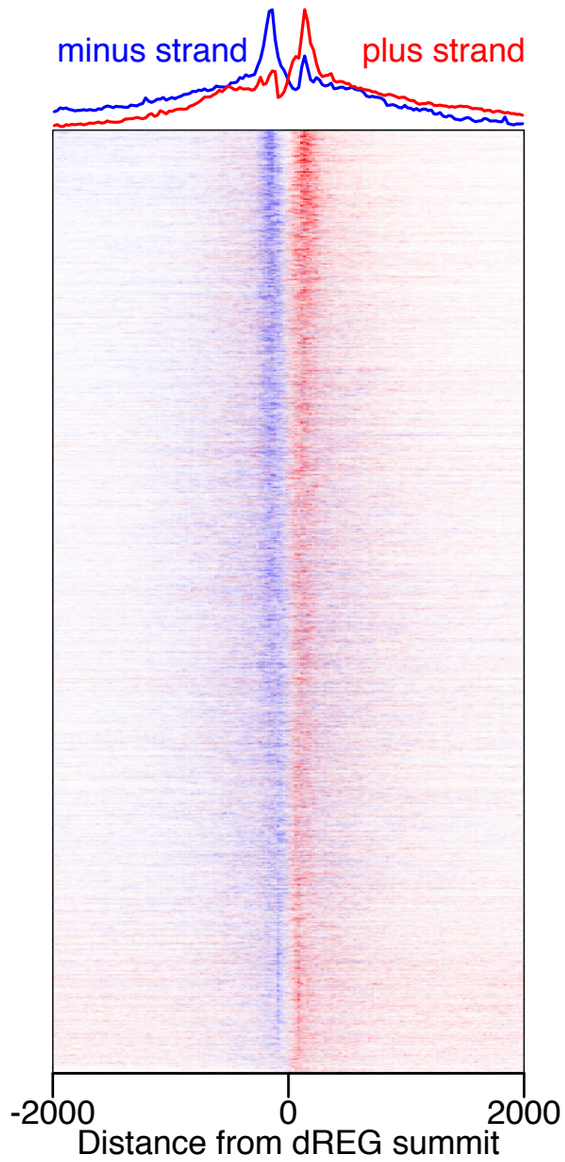




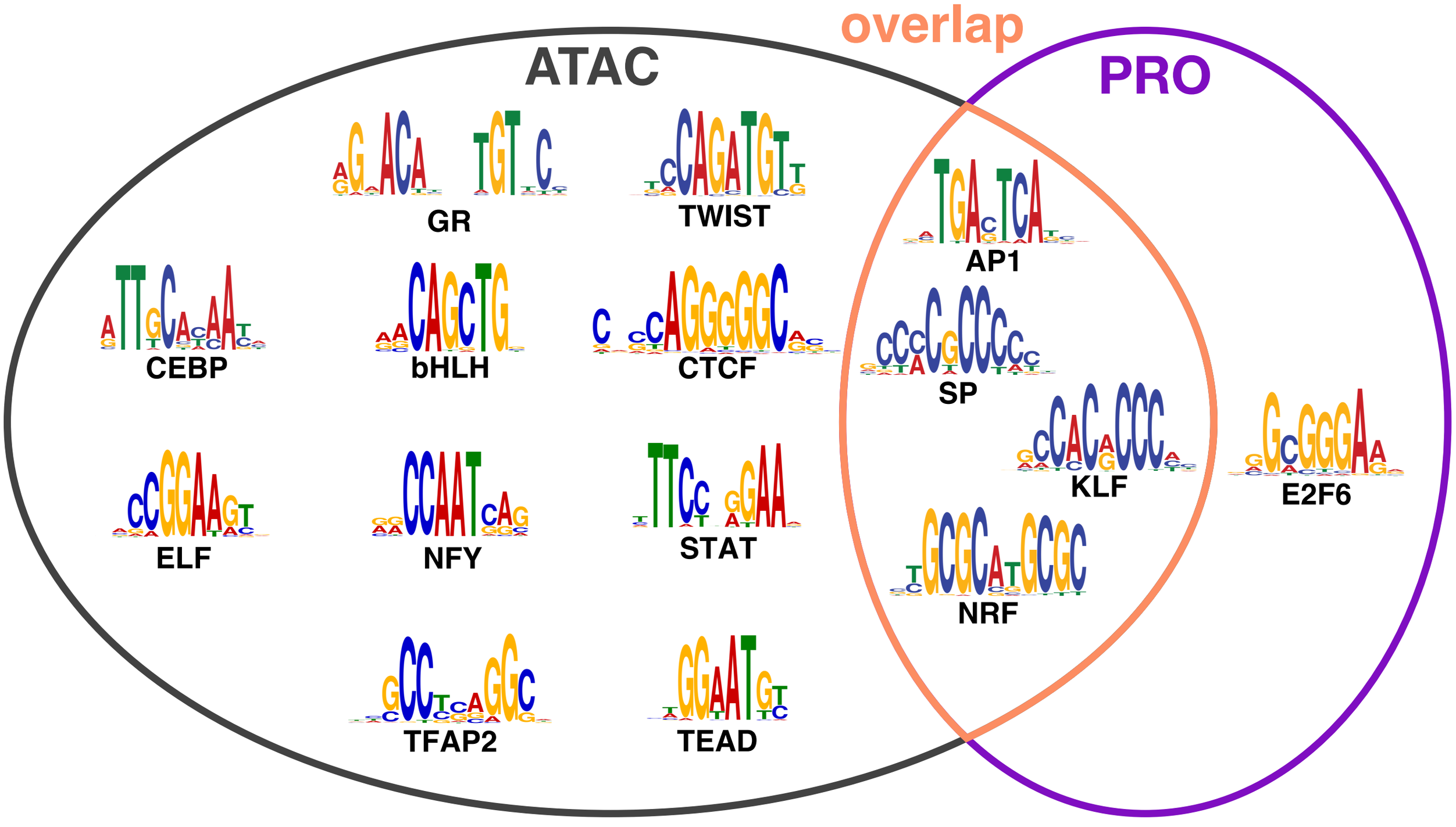


Bidirectional transcription signatures from PRO-seq independently identifies putative regulatory regions

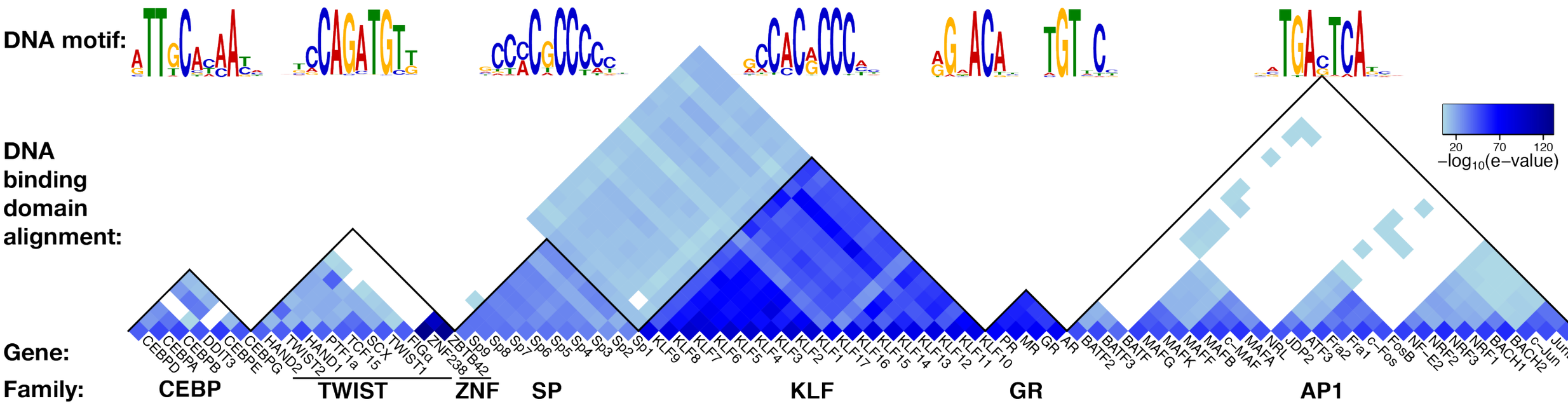




Motifs enriched within dynamic ATAC and bidirectional PRO peaks

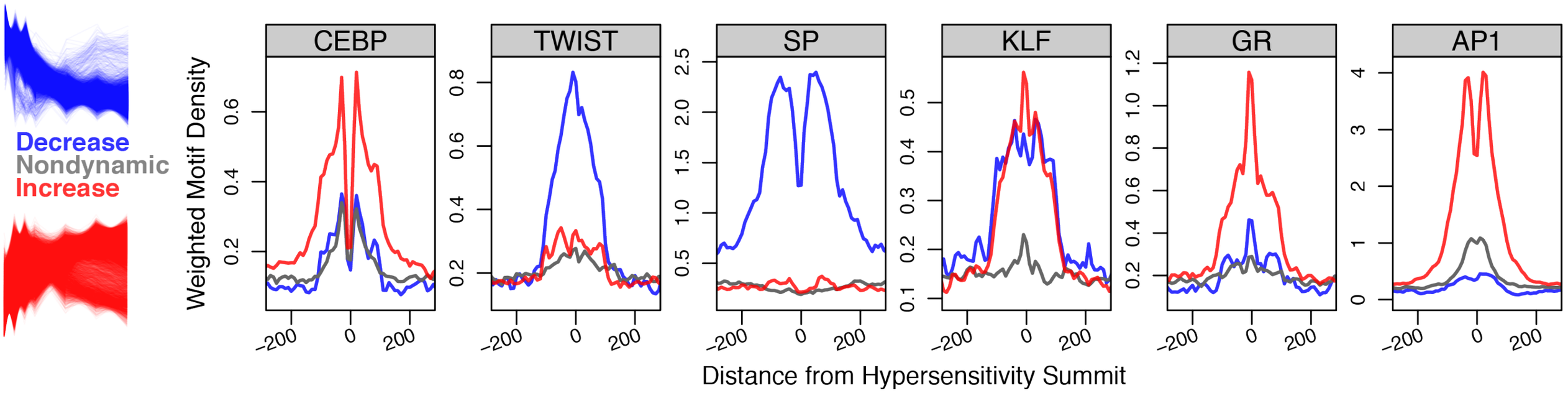


14 TF-family motifs (top 6 shown) drive early changes in chromatin and transcription

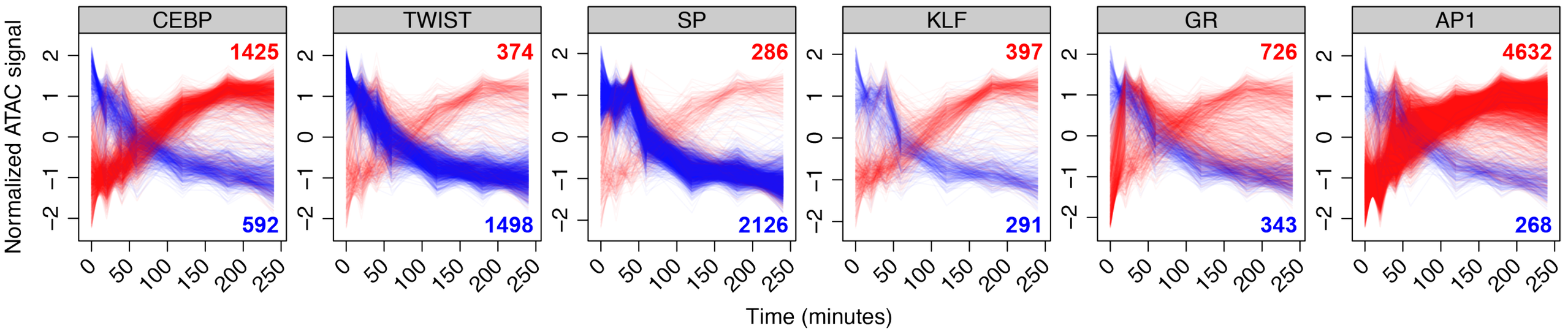


Paralogous TF DBD families that recognize each motif

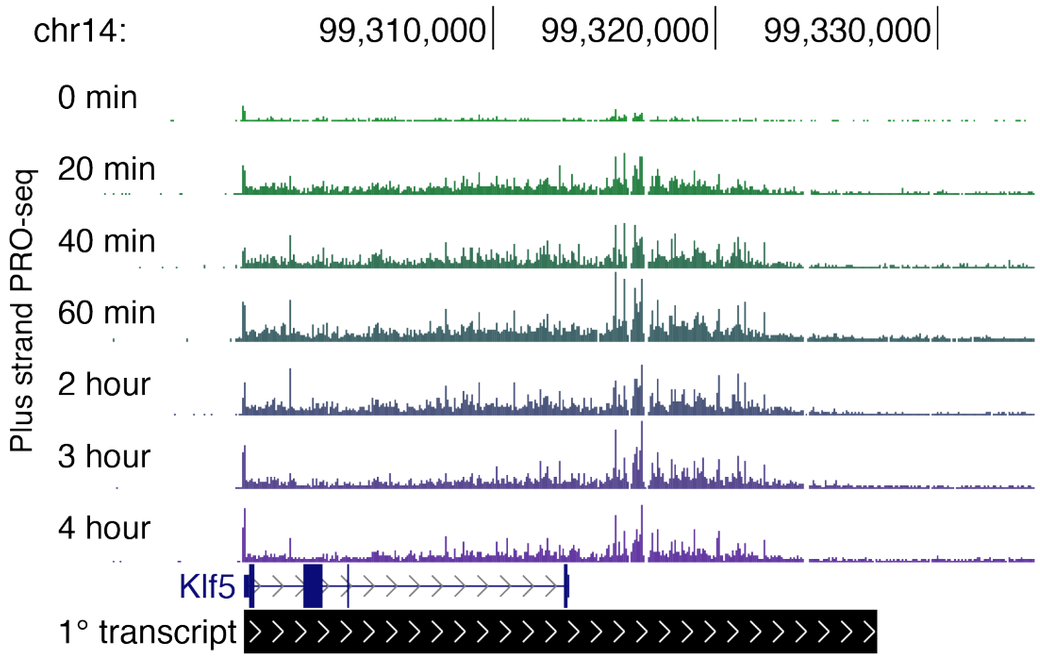
CEBP, KLF, GR, and AP1 motifs associate w/ increasing chromatin accessibility
TWIST and SP associate w/ decreasing chromatin accessibility

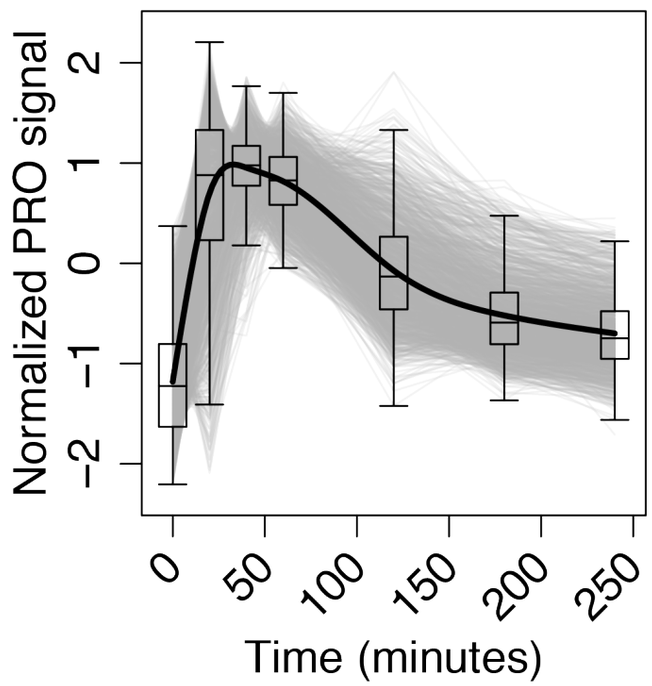


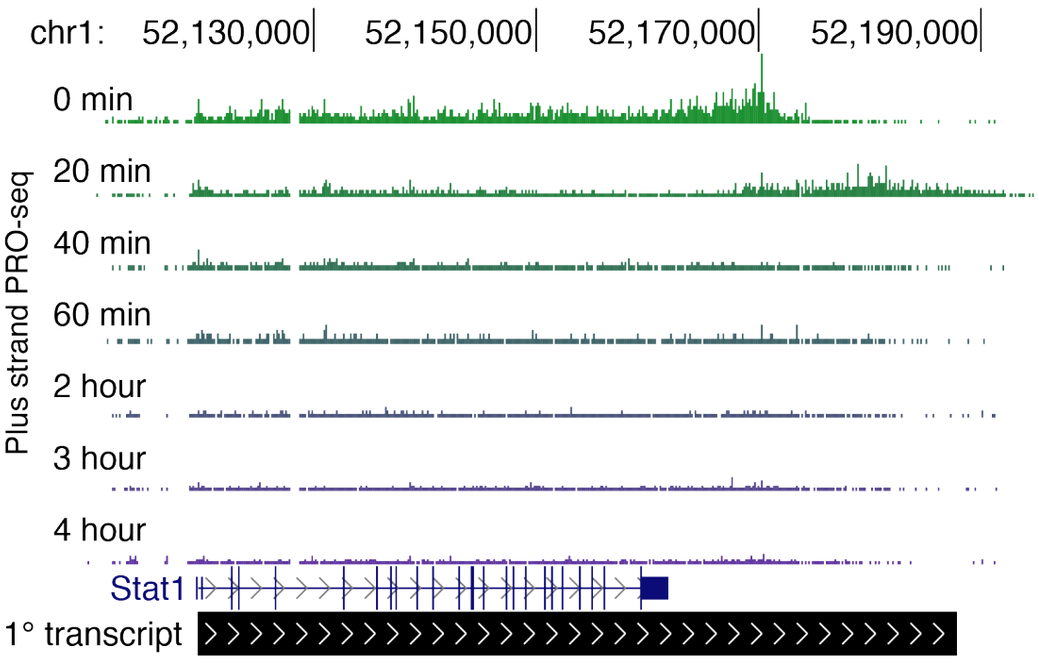
CEBP, KLF, GR, and AP1 motifs associate w/ increasing chromatin accessibility
TWIST and SP associate w/ decreasing chromatin accessibility

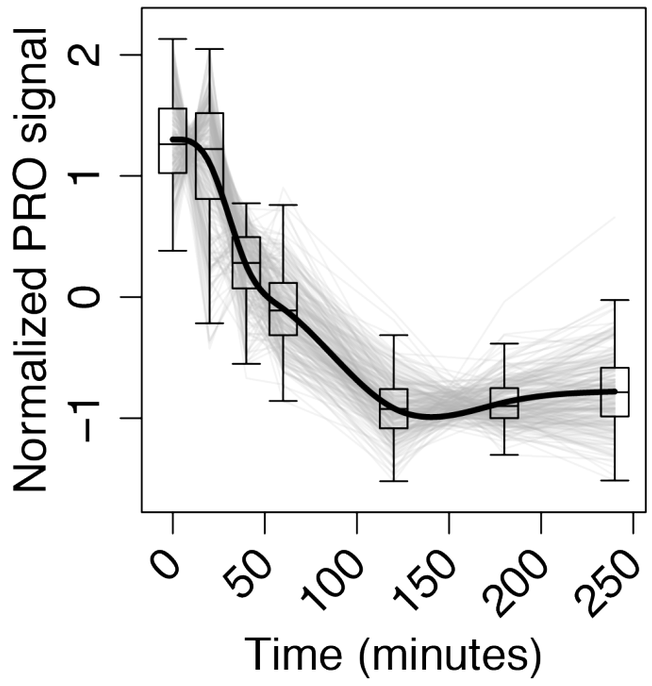


Genes also have distinct activation and repression kinetics



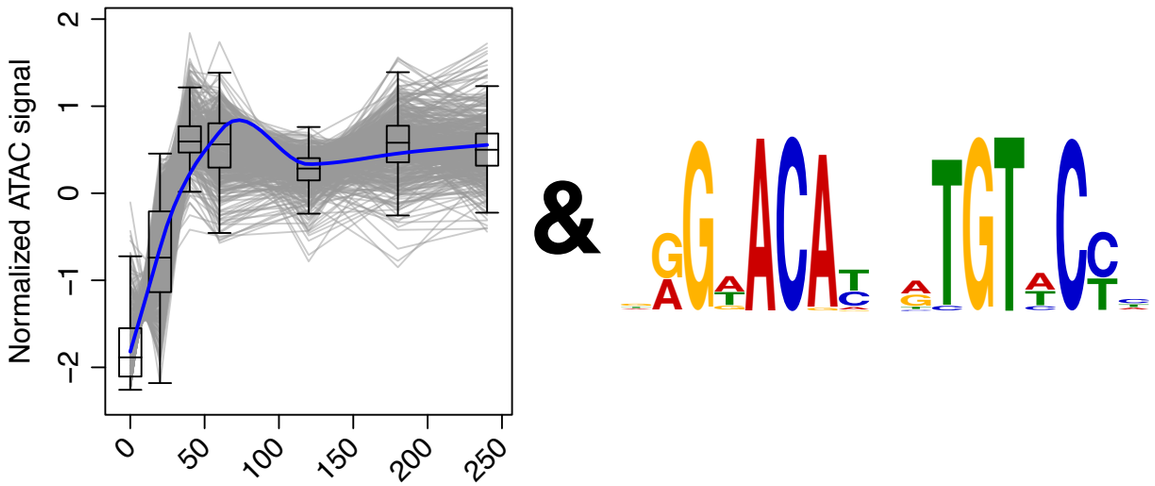






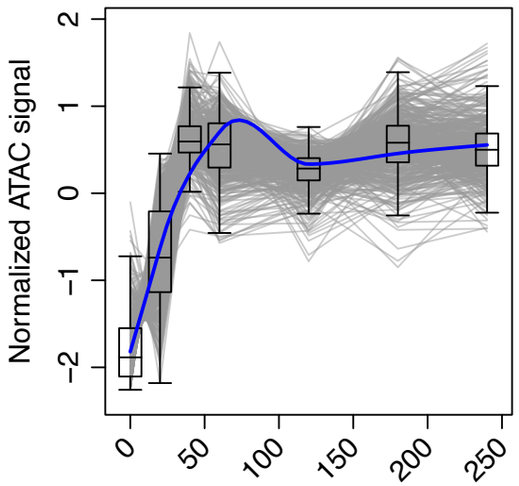
Linking direct functional interactions between Regulatory Elements (RE) and Genes (TU)

Linking direct functional interactions between Regulatory Elements (RE) and Genes (TU)

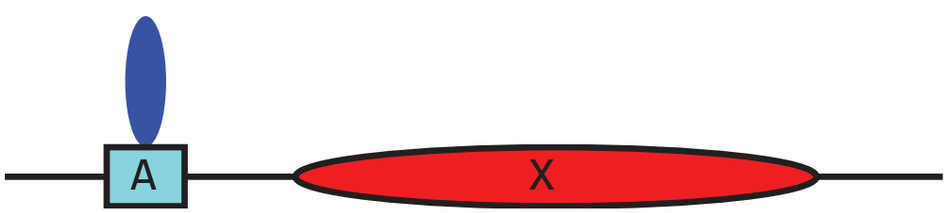


&

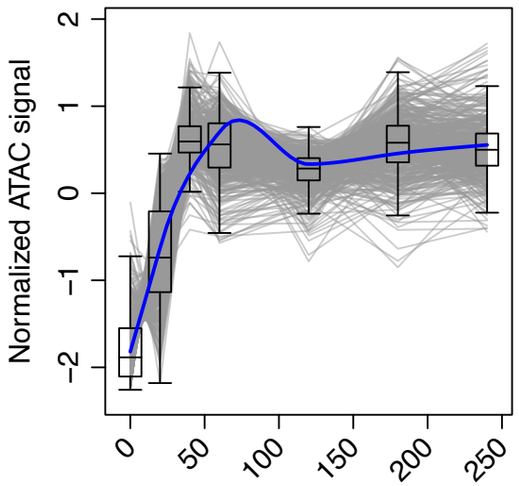
Linking direct functional interactions between Regulatory Elements (RE) and Genes (TU)

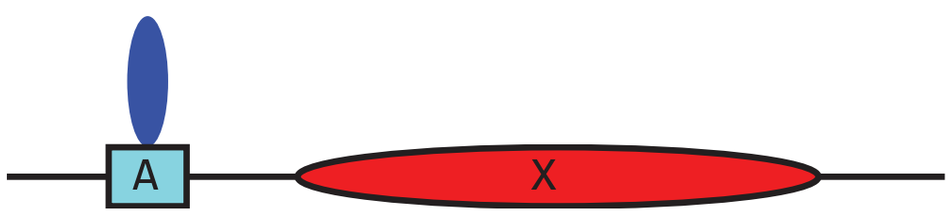


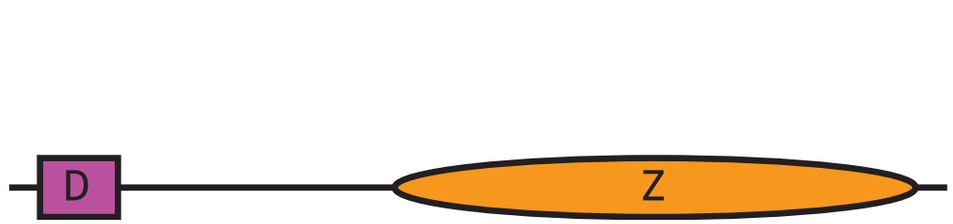
&



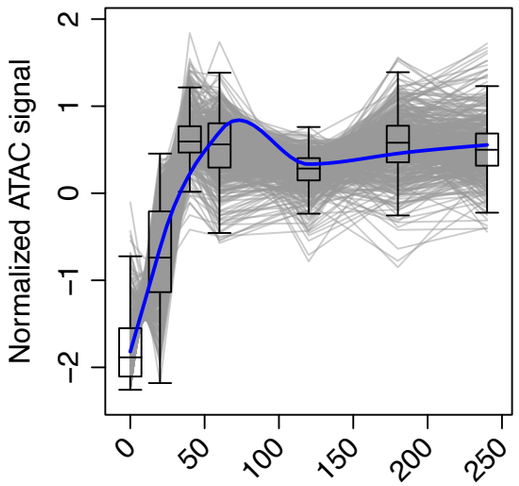
Linking direct functional interactions between Regulatory Elements (RE) and Genes (TU)

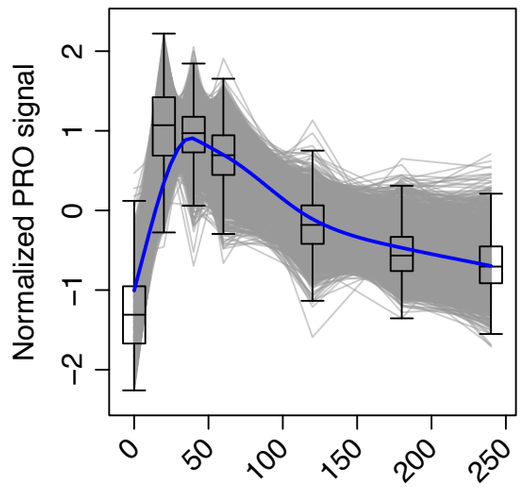


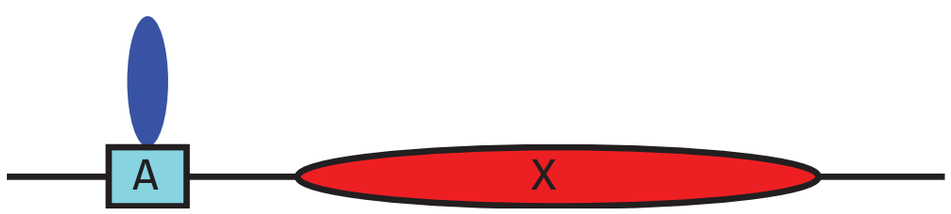




Linking direct functional interactions between Regulatory Elements (RE) and Genes (TU)





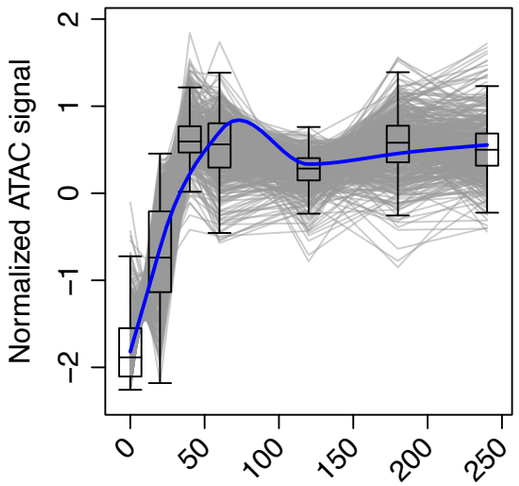


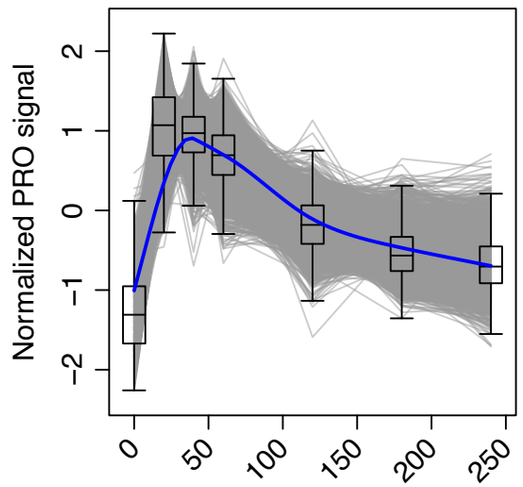
ATAC peak ↑
gene PRO ↑

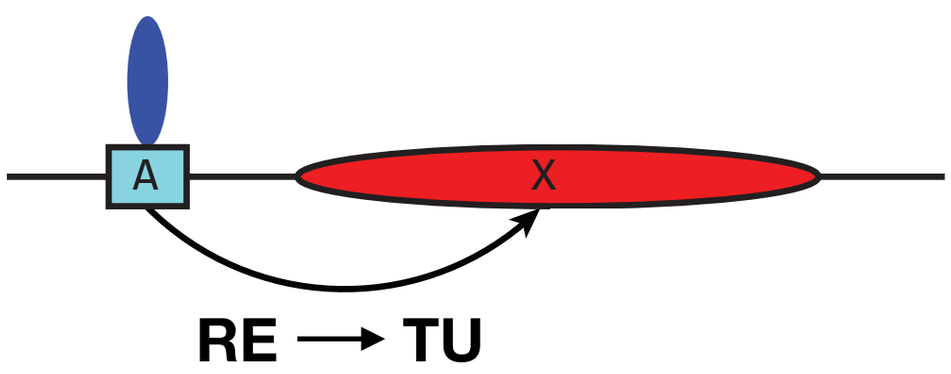
&

peak and gene are nearby

Linking direct functional interactions between Regulatory Elements (RE) and Genes (TU)





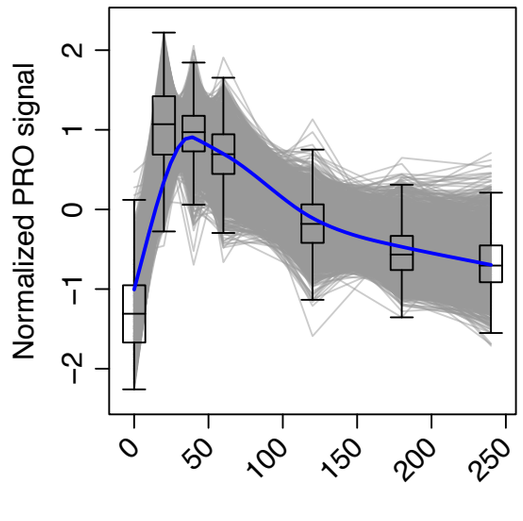


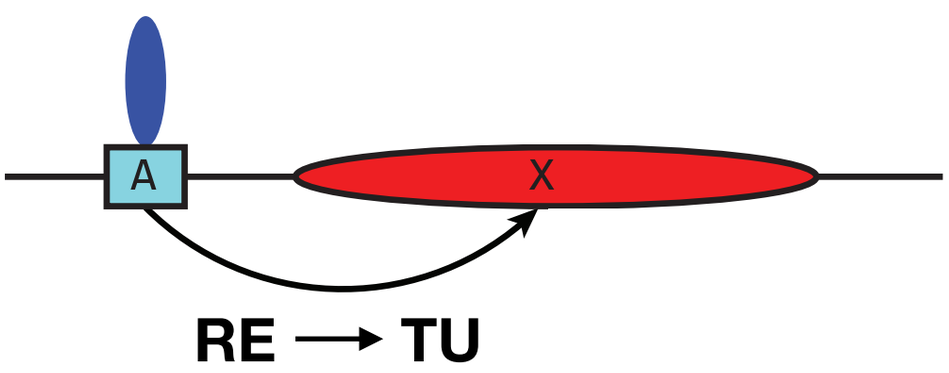
ATAC peak ↑
gene PRO ↑

&

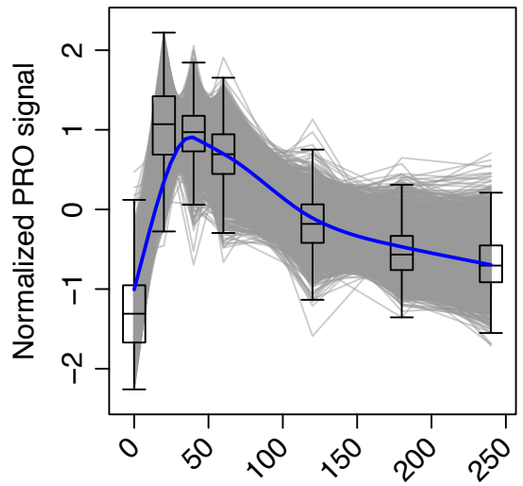
peak and gene are nearby

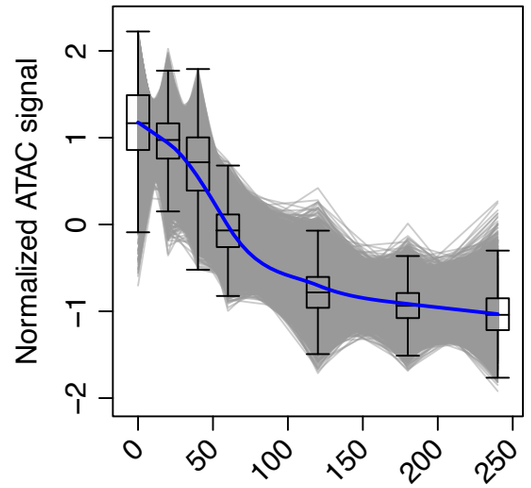
Linking direct functional interactions between Regulatory Elements (RE) and Genes (TU)

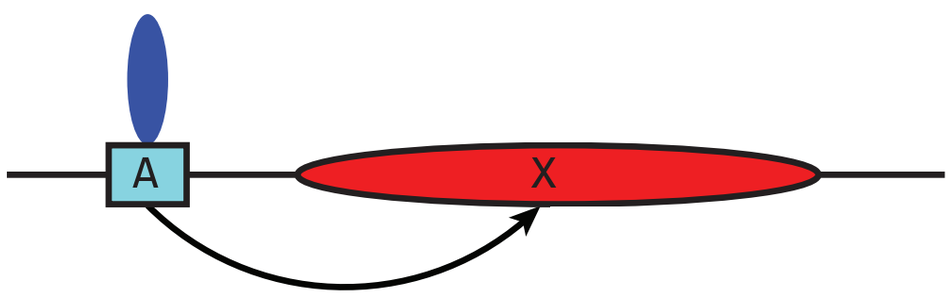




Linking direct functional interactions between Regulatory Elements (RE) and Genes (TU)







RE → TU

gene PRO ↑
ATAC peak ↓
(repressor binding)

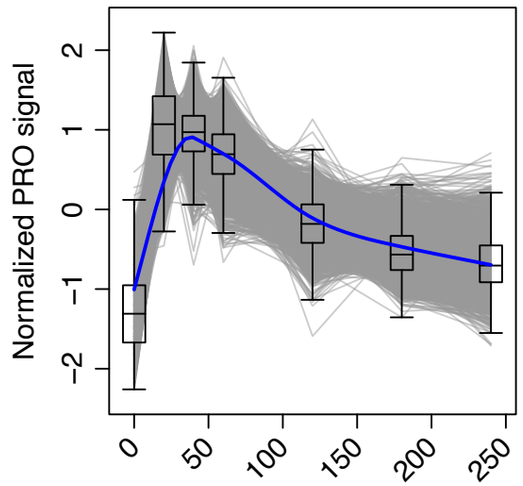
&

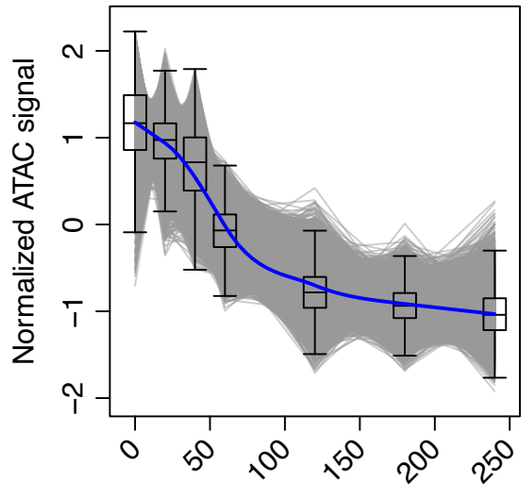
time delay
(PRO prior ATAC)

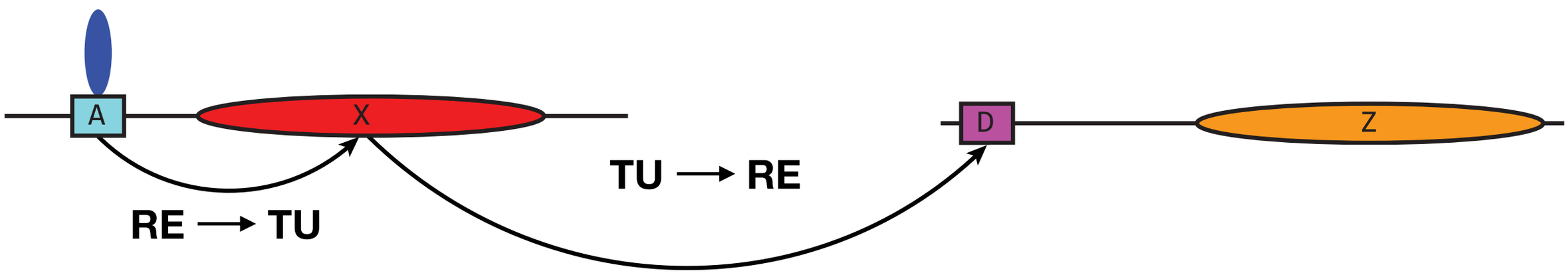
&

gene's motif within peak

Linking direct functional interactions between Regulatory Elements (RE) and Genes (TU)

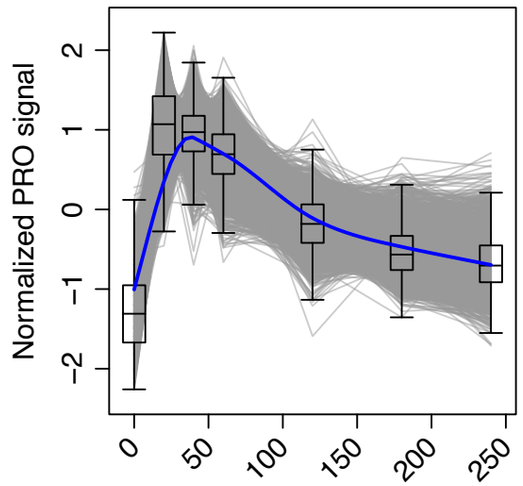


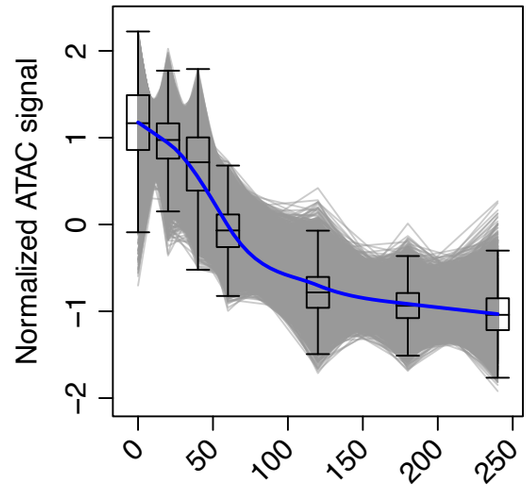


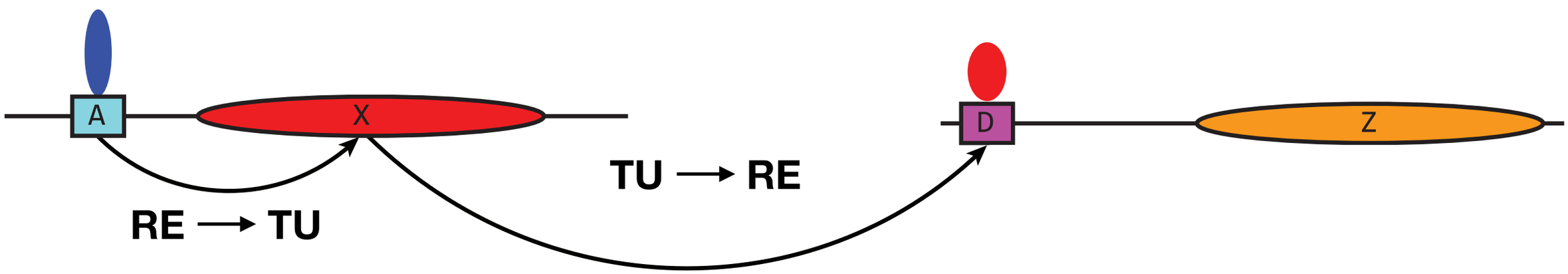


gene PRO ↕ ATAC peak (repressor binding) **&** **time delay (PRO prior ATAC)** **&** **gene's motif within peak**

Linking direct functional interactions between Regulatory Elements (RE) and Genes (TU)

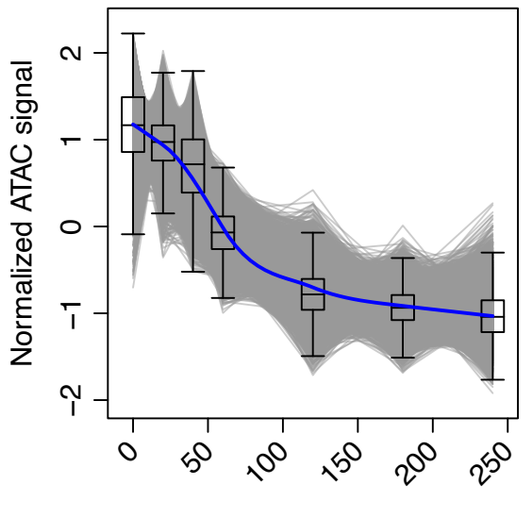


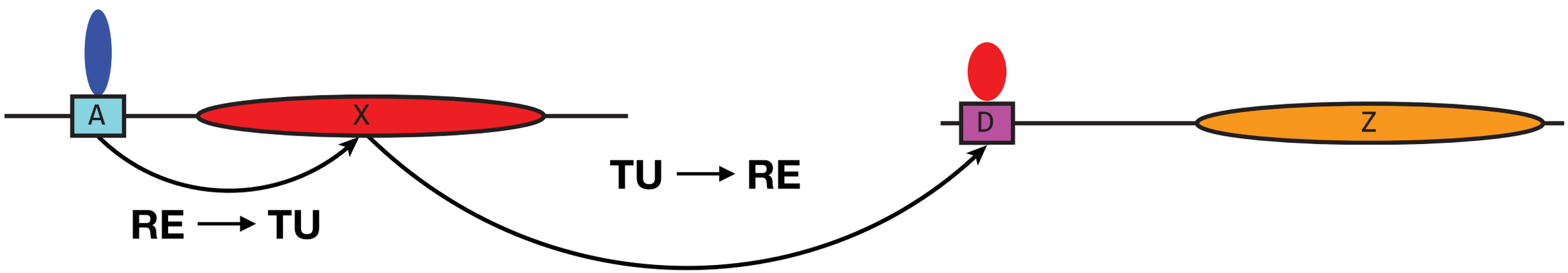




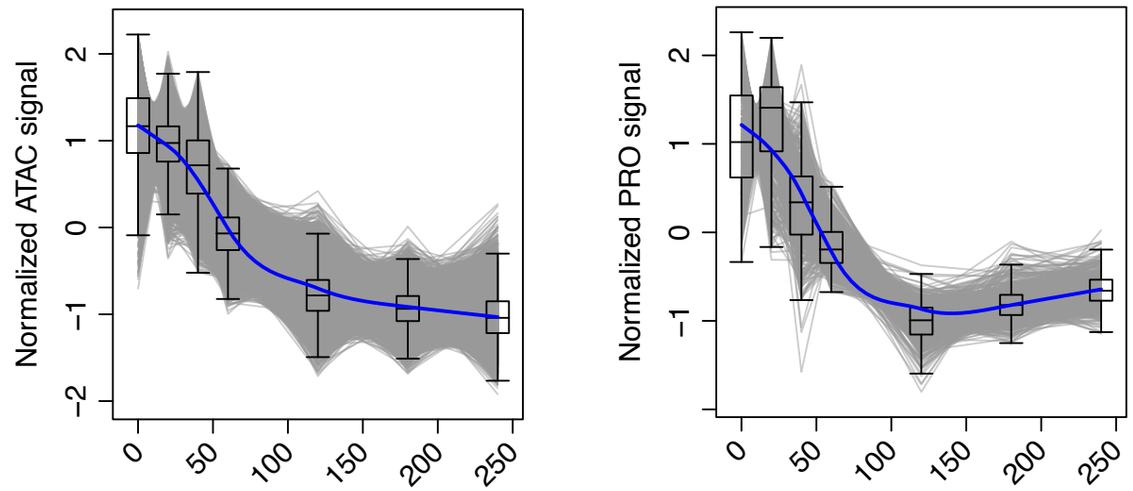
gene PRO ↕ ATAC peak (repressor binding) **&** **time delay (PRO prior ATAC)** **&** **gene's motif within peak**

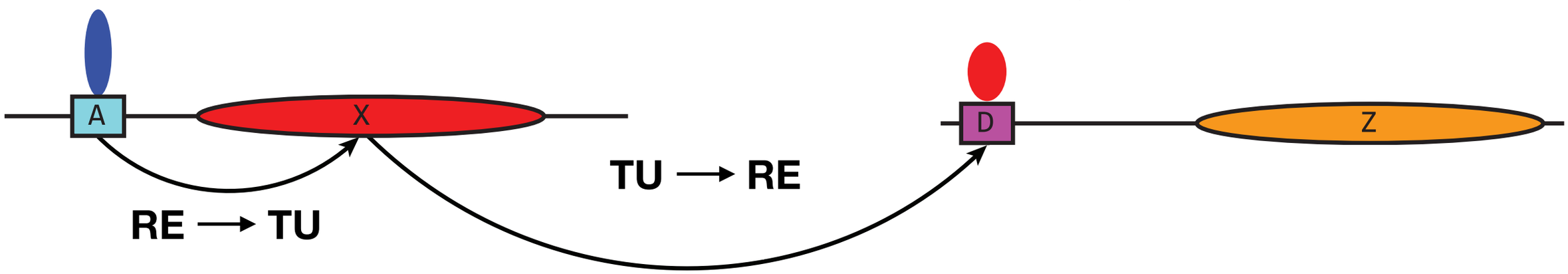
Linking direct functional interactions between Regulatory Elements (RE) and Genes (TU)





Linking direct functional interactions between Regulatory Elements (RE) and Genes (TU)



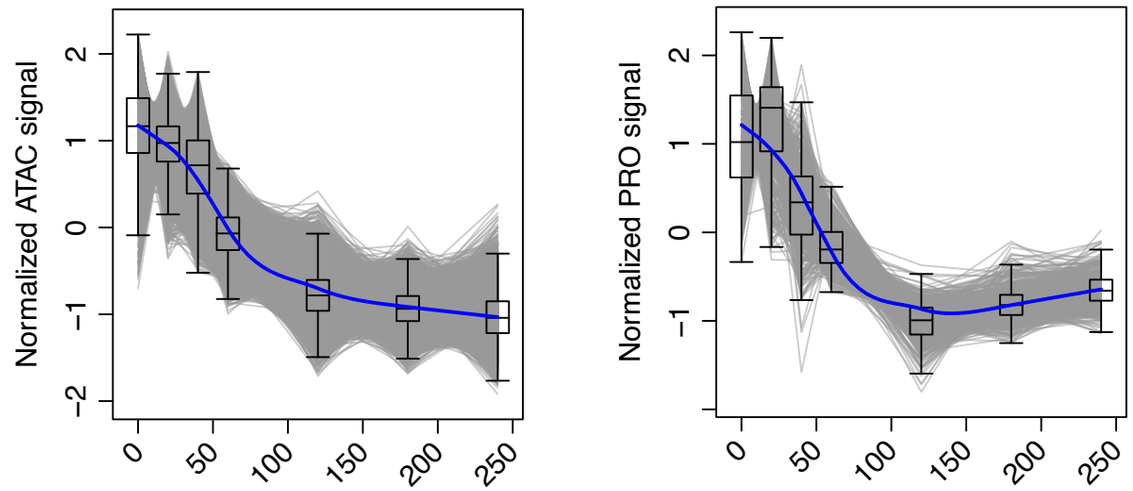


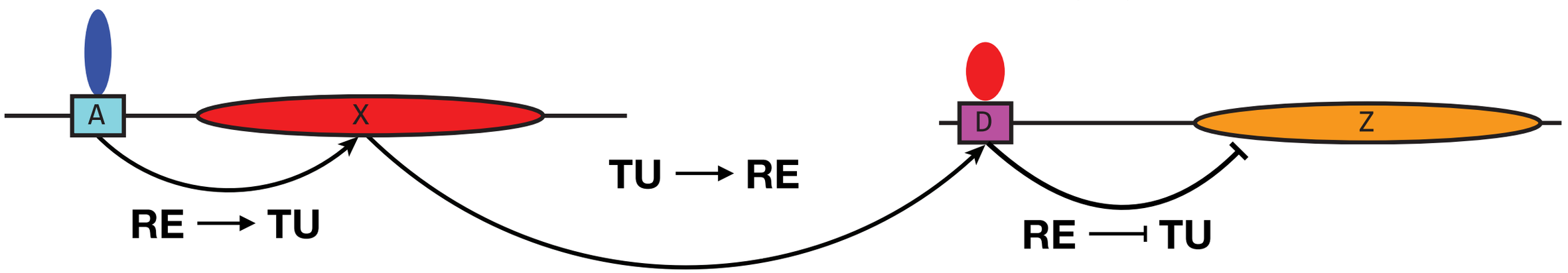
ATAC peak ↓
gene PRO ↓

&

peak and gene are nearby

Linking direct functional interactions between Regulatory Elements (RE) and Genes (TU)



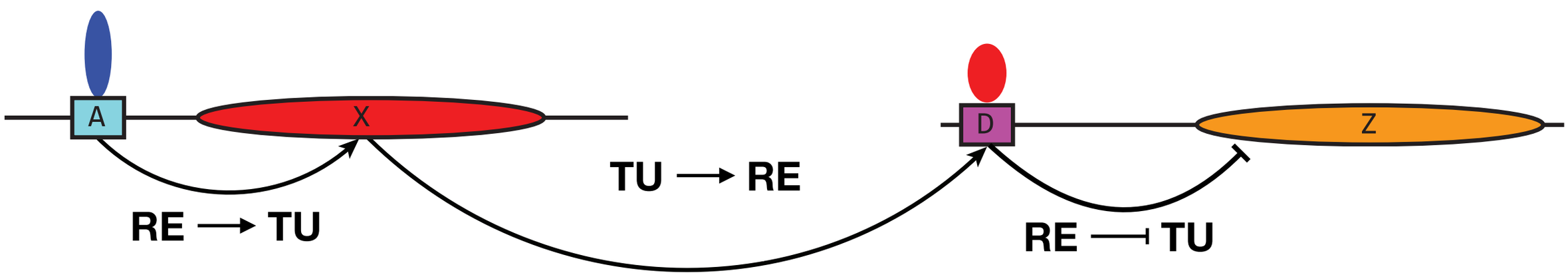


ATAC peak ↓
gene PRO ↓

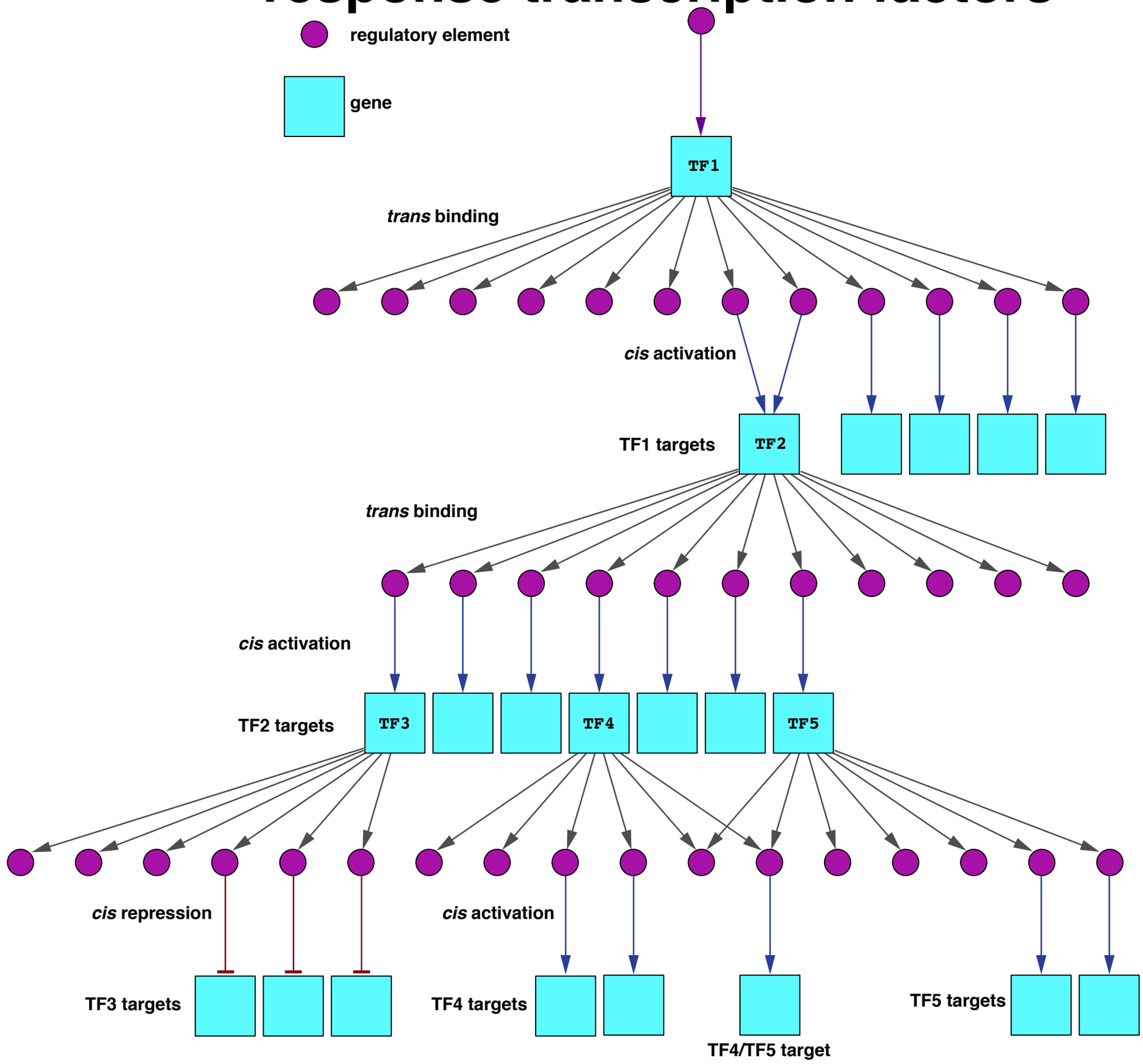
&

peak and gene are nearby

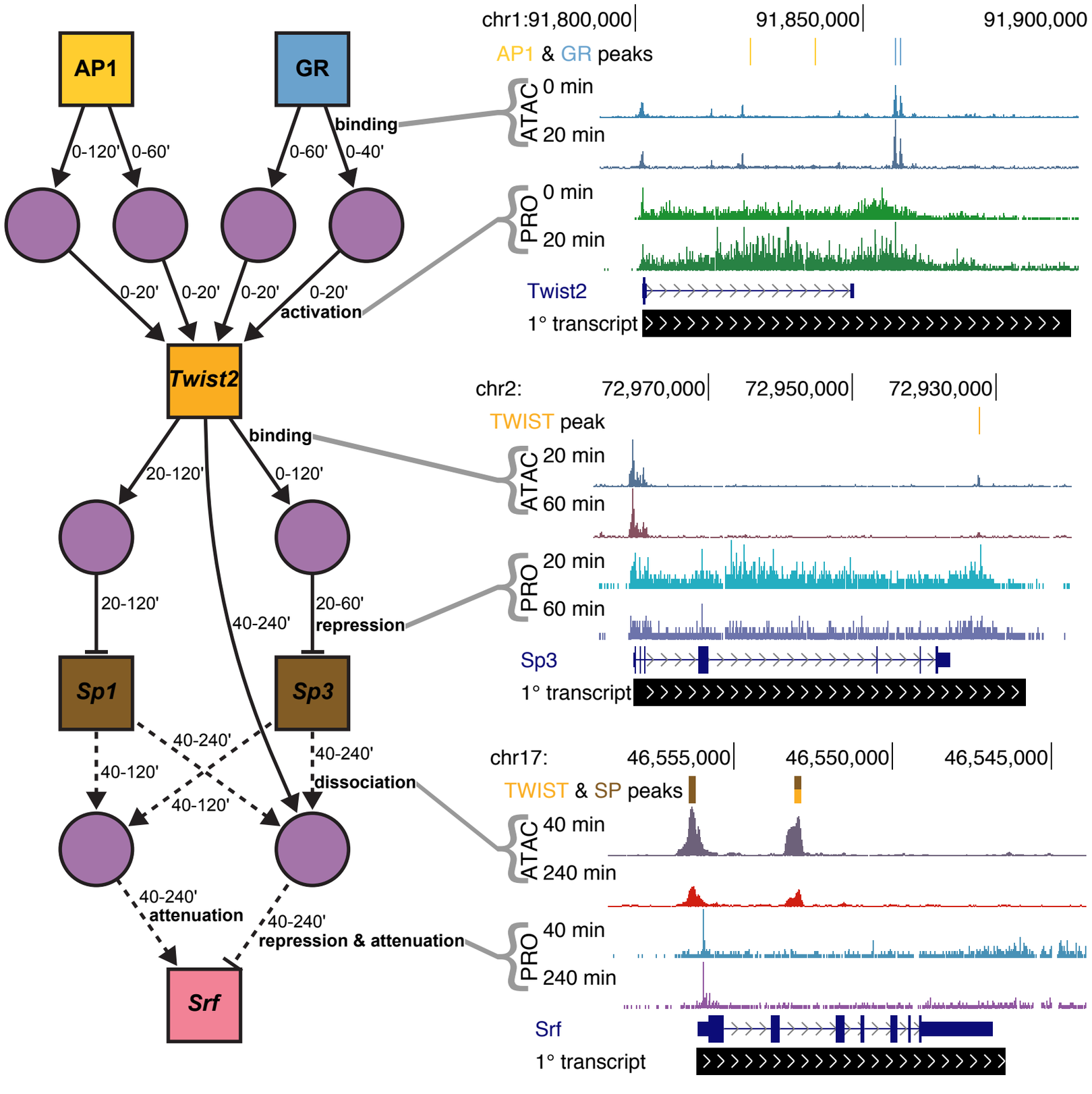
Linking direct functional interactions between Regulatory Elements (RE) and Genes (TU)



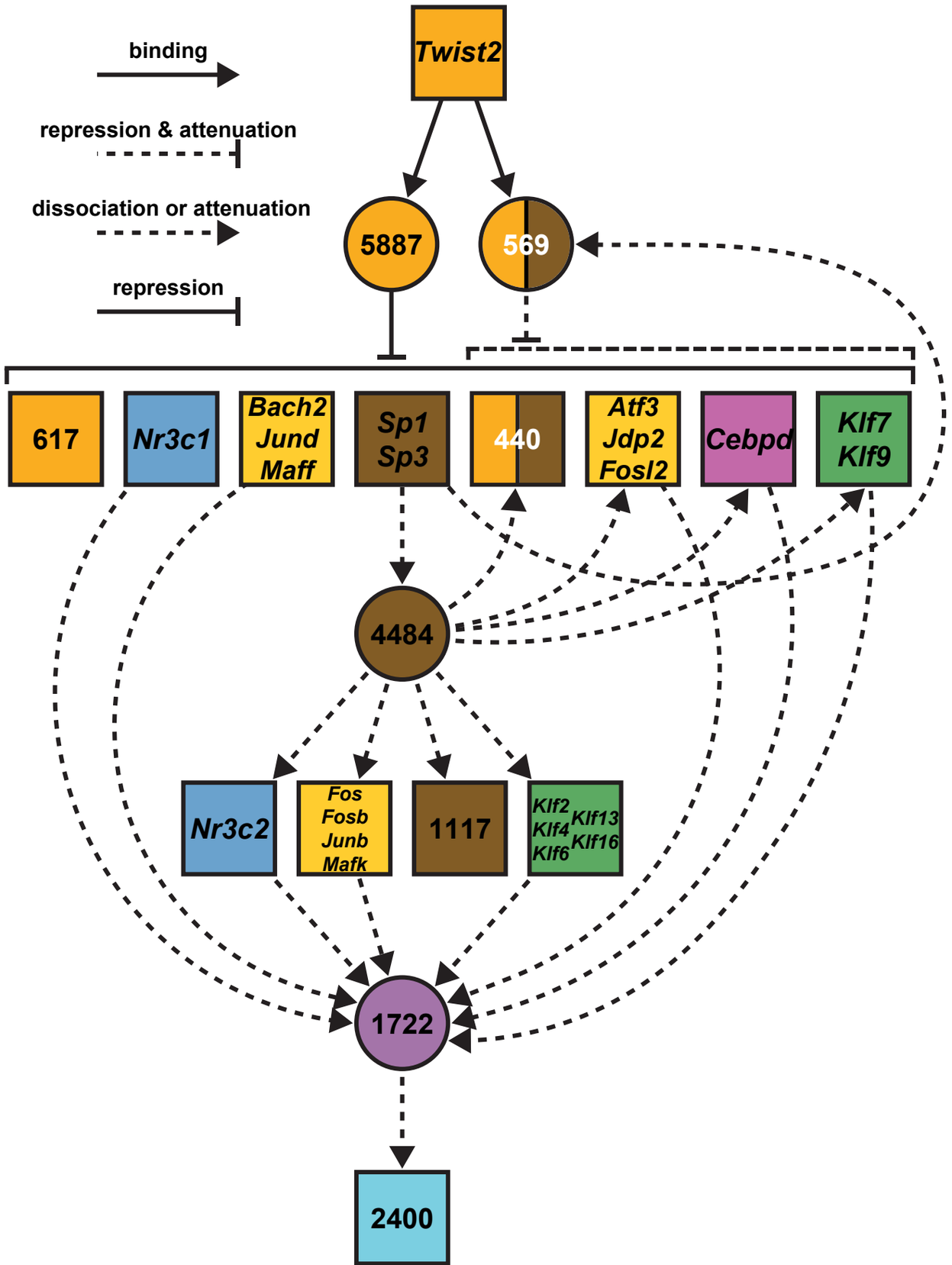
We are interested in highly connected early response transcription factors



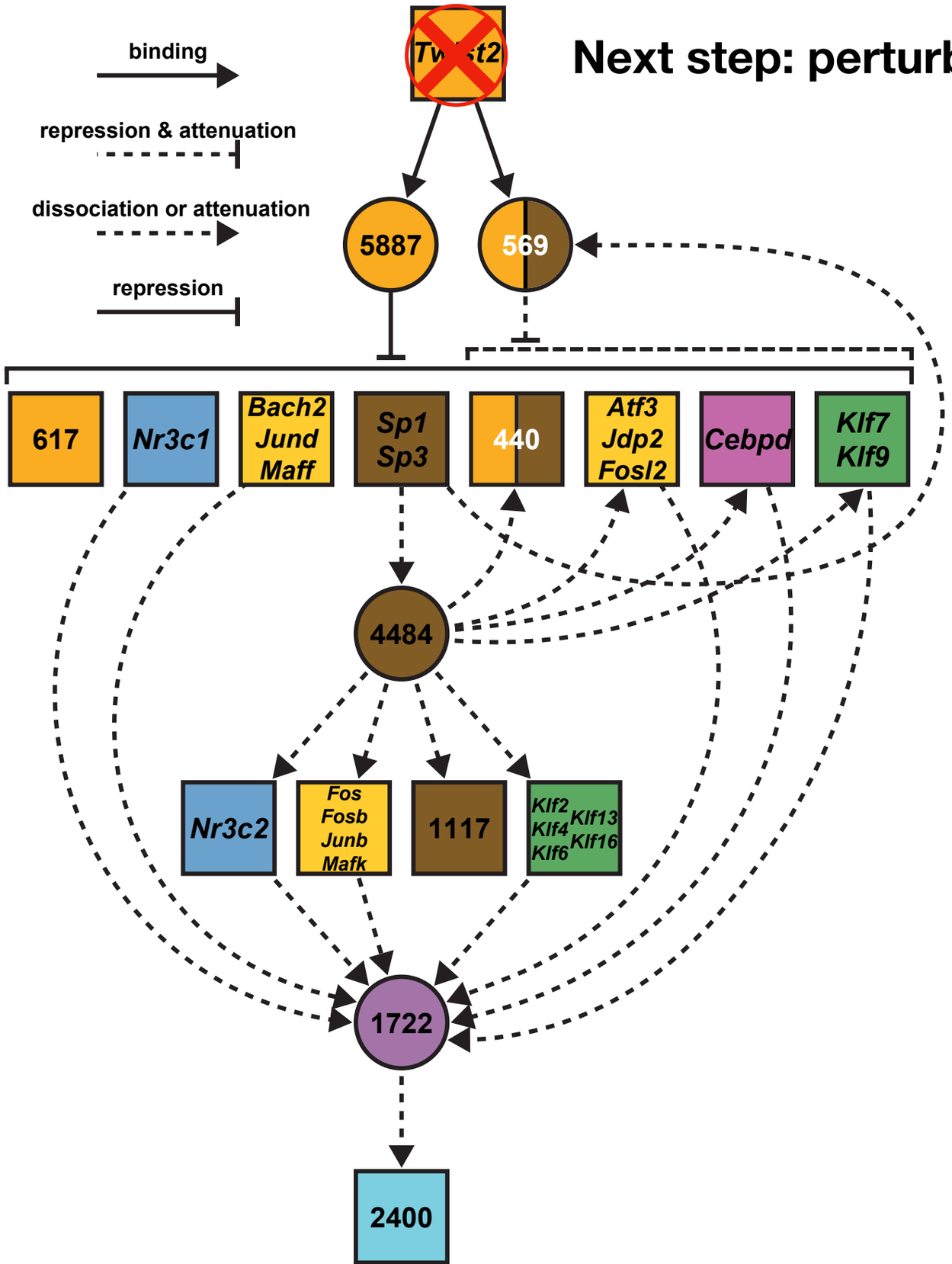
TWIST2 is active early and transiently



TWIST2 is a highly connected early response gene

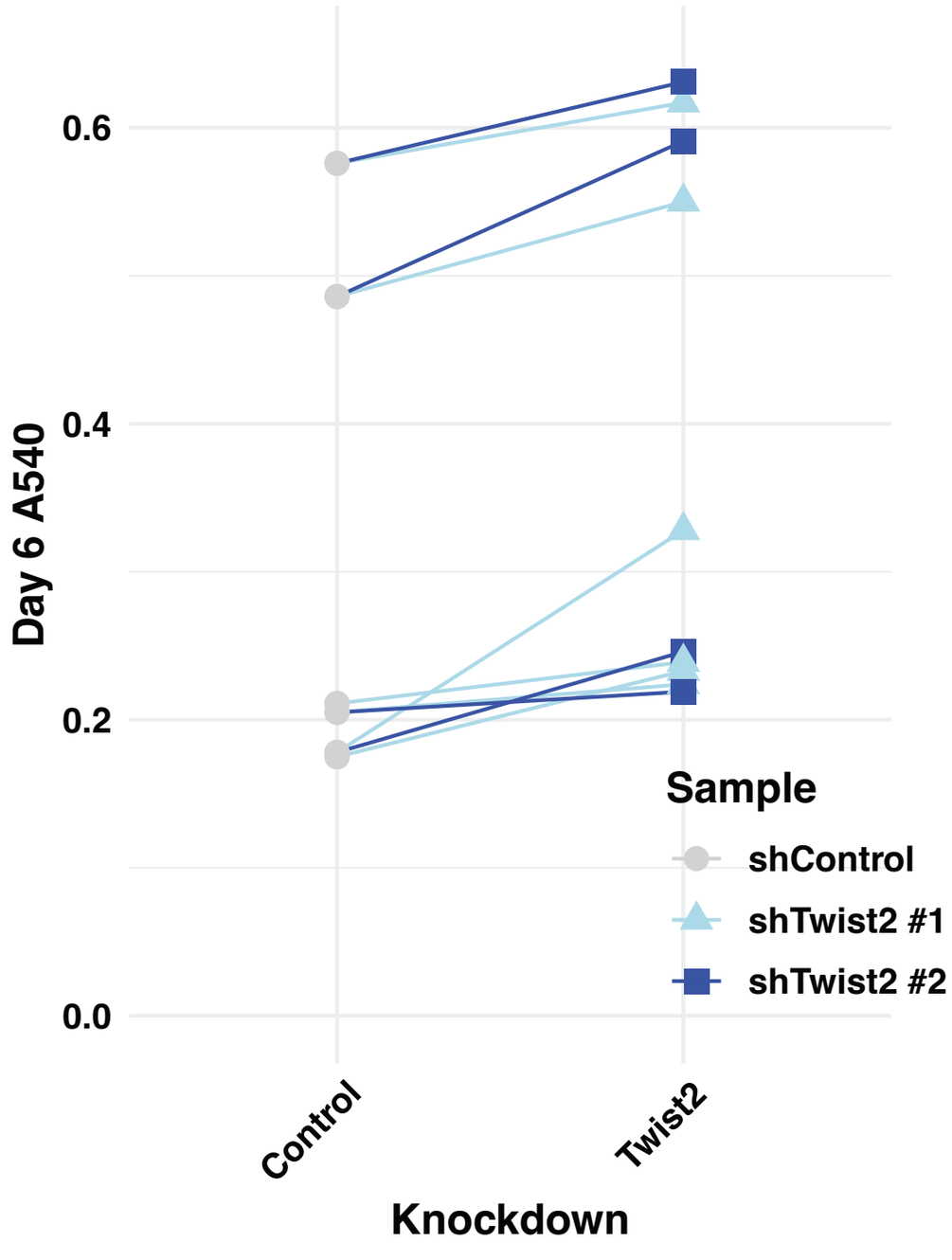


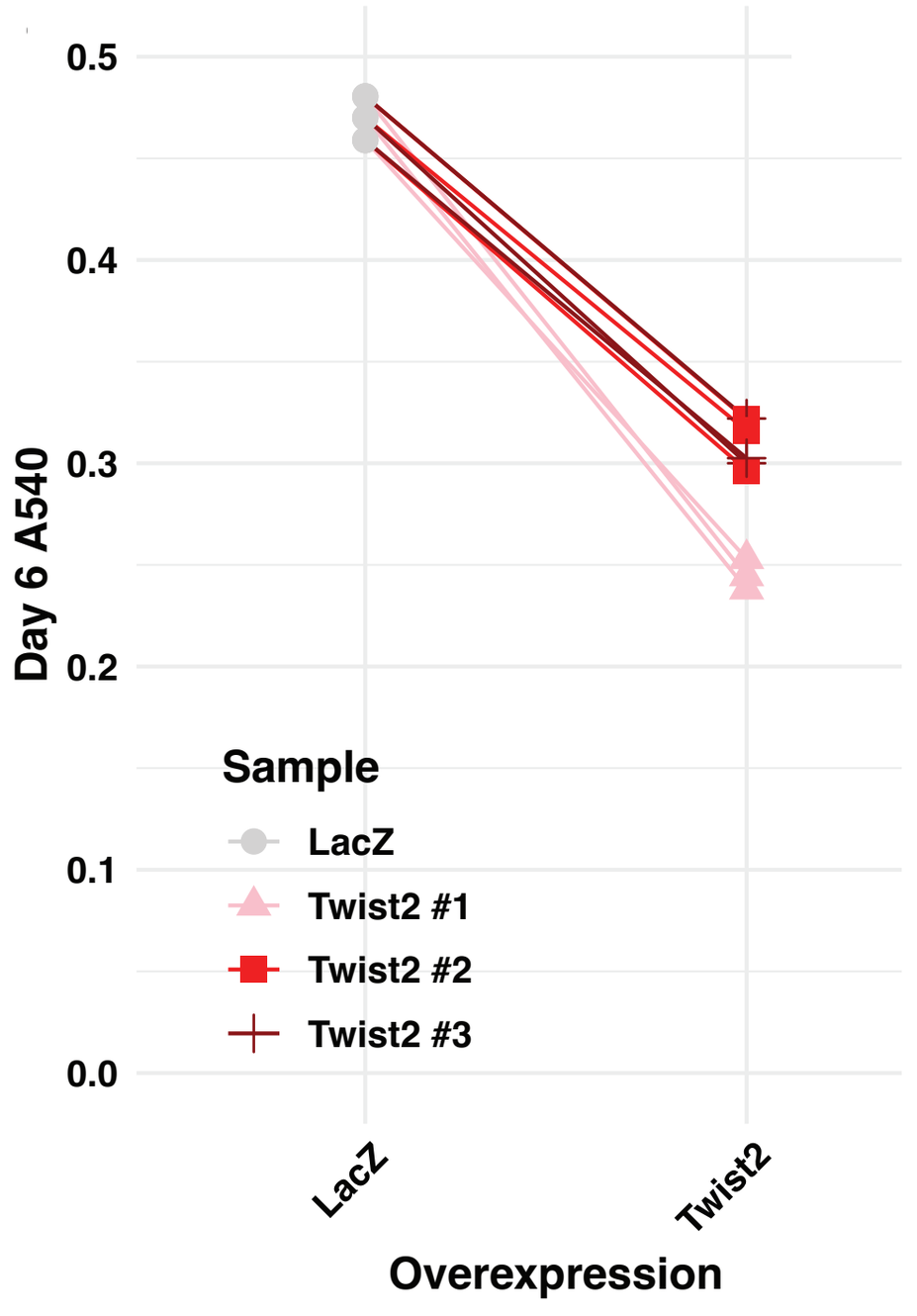
TWIST2 is a highly connected early response gene



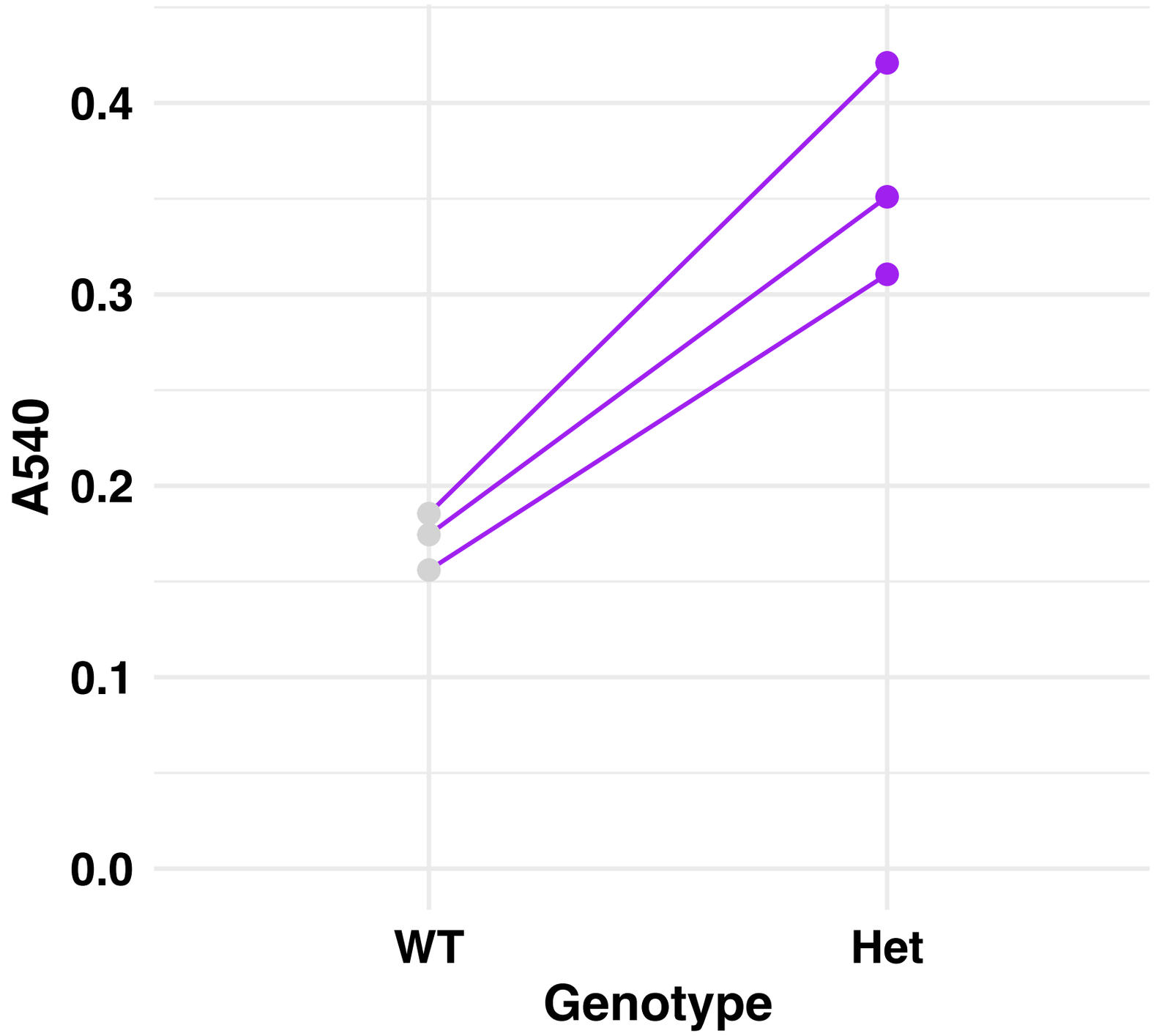
Next step: perturb and observe

Twist2 depletion and over expression in 3T3-L1 cells modulate adipogenesis

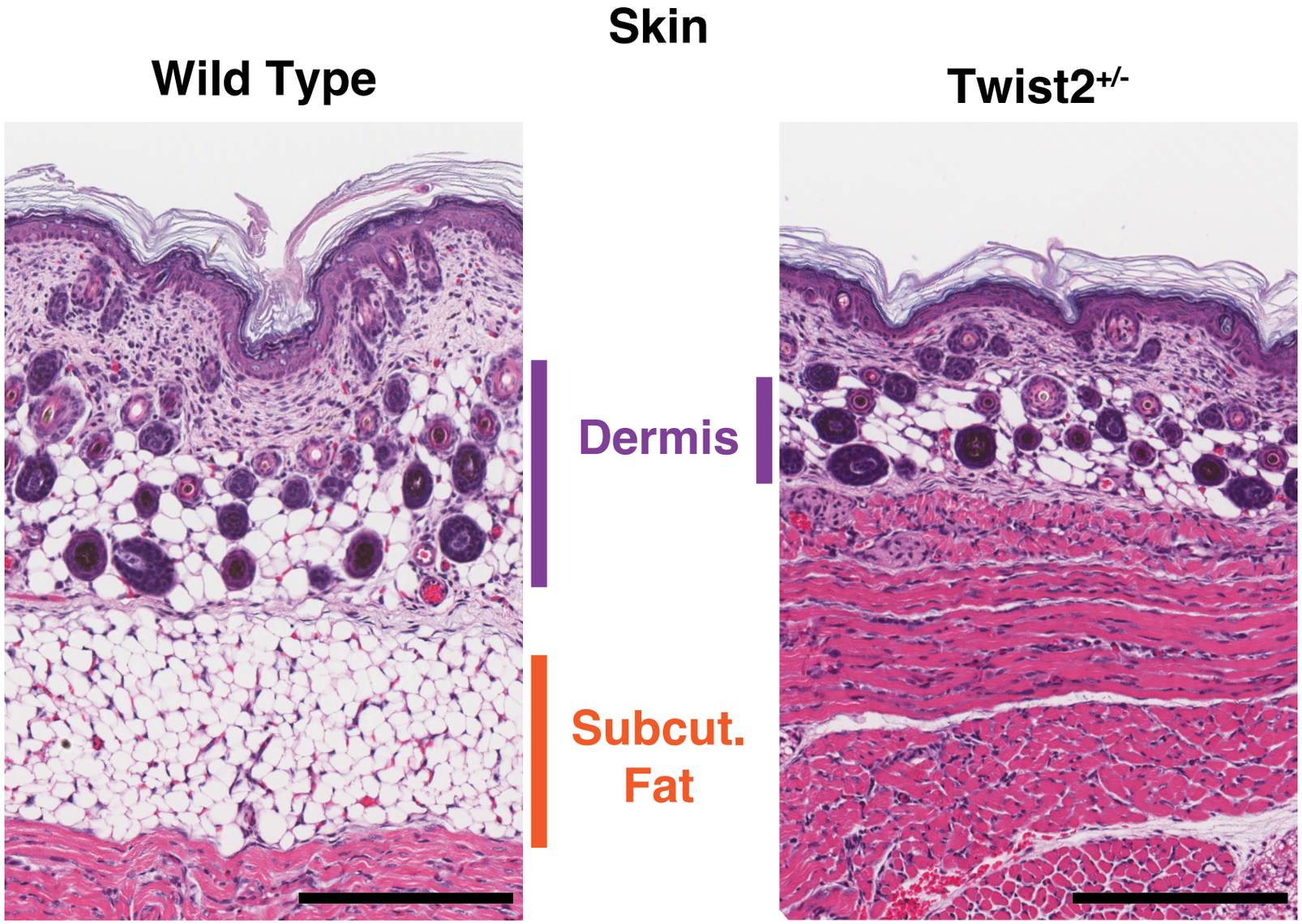




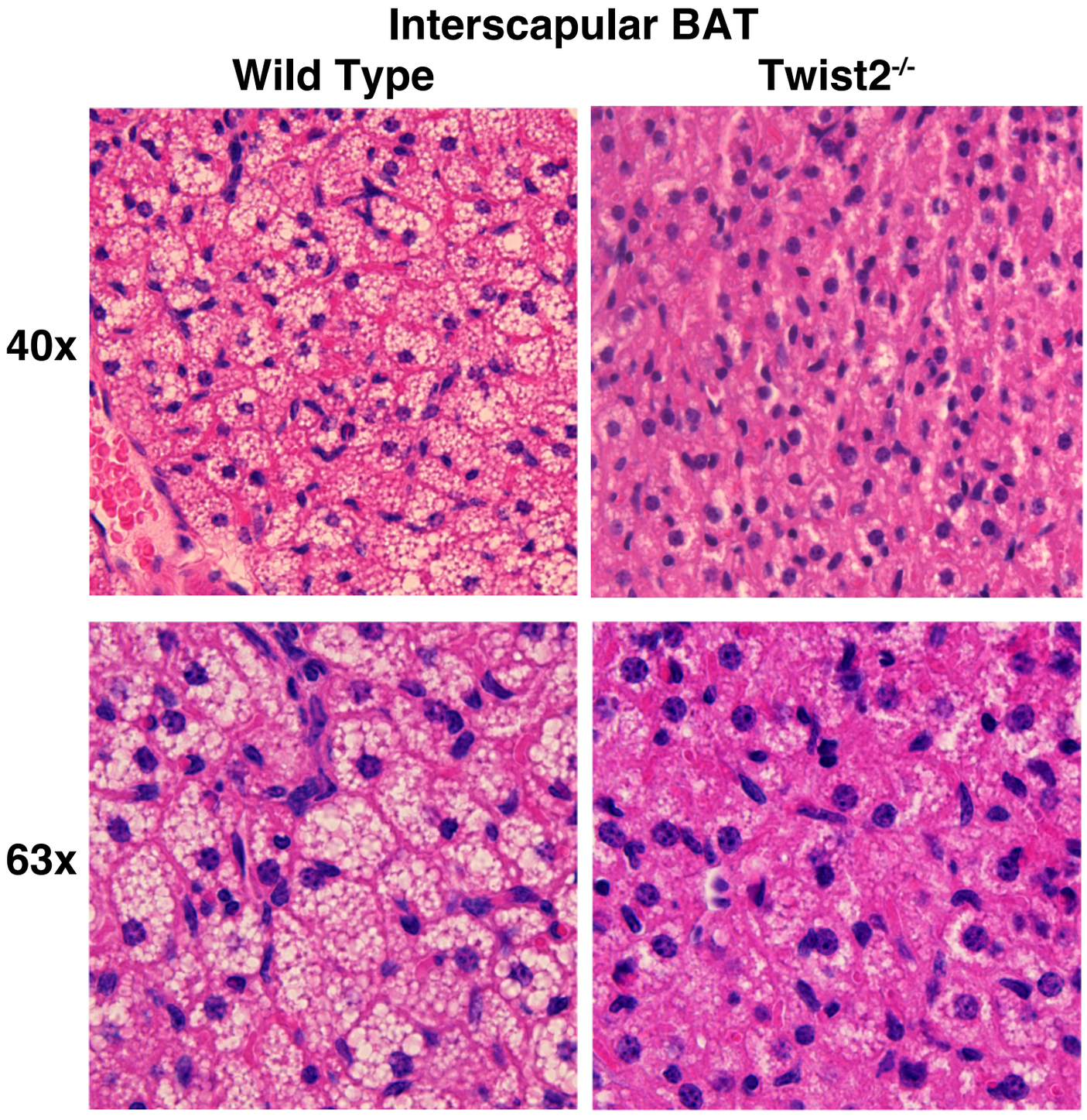
Twist2 heterozygote mice have increased differentiation of ex vivo cultured white adipocytes



Twist2$^{+/-}$ mice have a near absence of subcutaneous fat



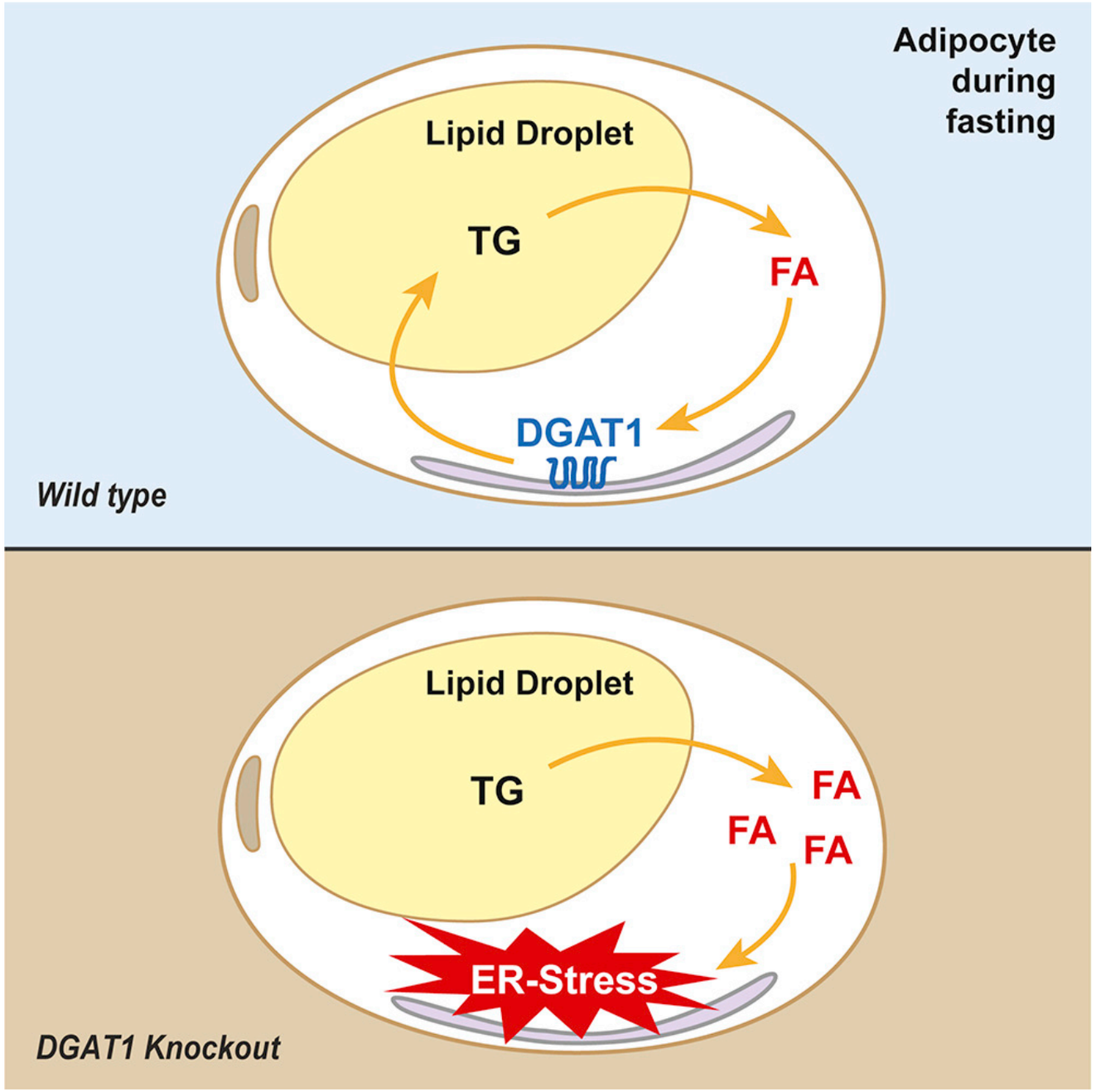
Twist2$^{-/-}$ mice have reduced lipid droplets in brown fat



Can we reconcile the *in vitro* and *in vivo* results?

Recall: *in vitro* Twist2 restricts adipogenesis/lipid droplet formation
**
in vivo Twist2 promotes adipogenesis/lipid droplet formation**

Build up of fatty acids cause ER stress, which can lead to cell exhaustion and death

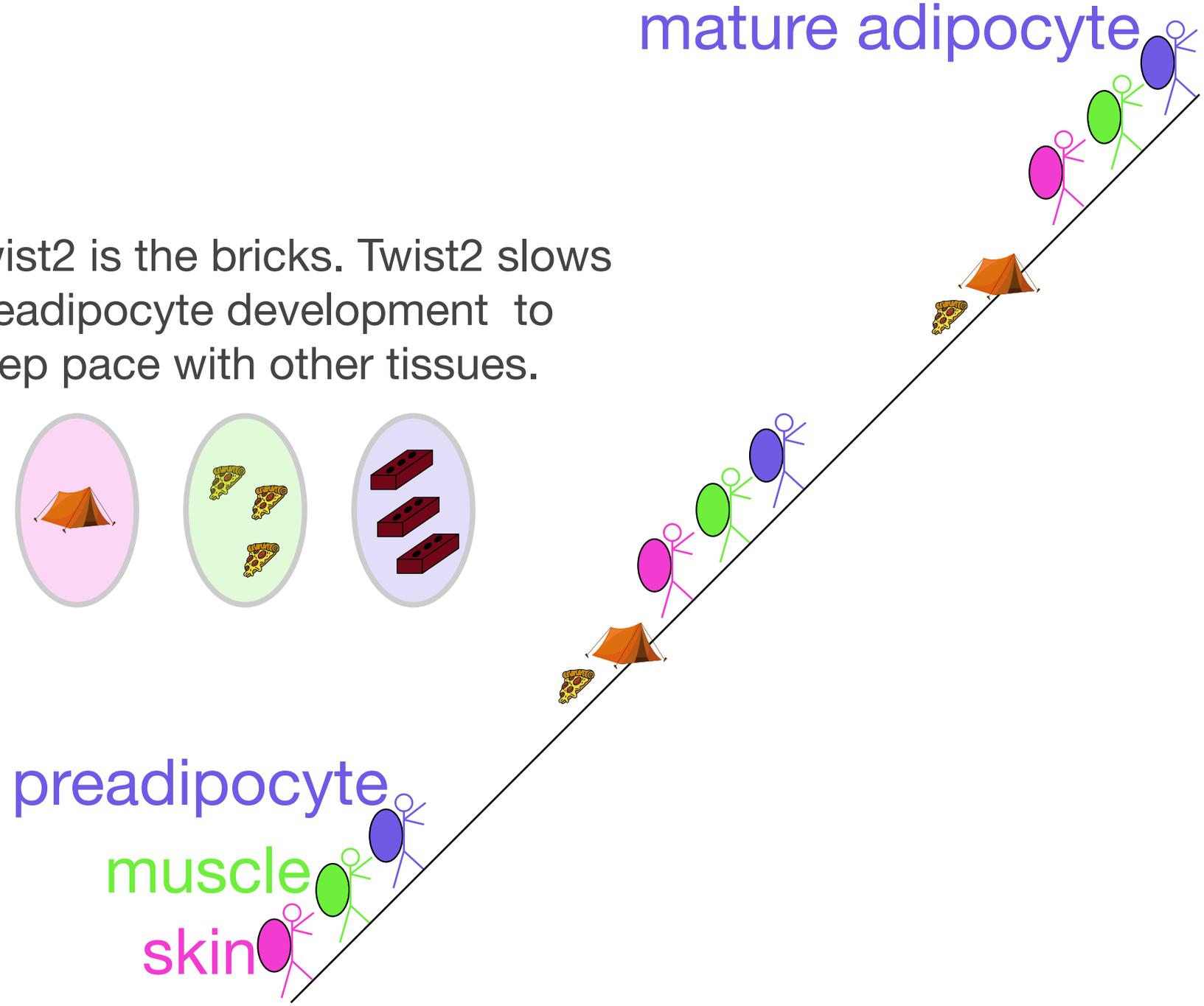


DGAT1 activation during lipolysis leads to re-esterification of fatty acids to prevent ER damage

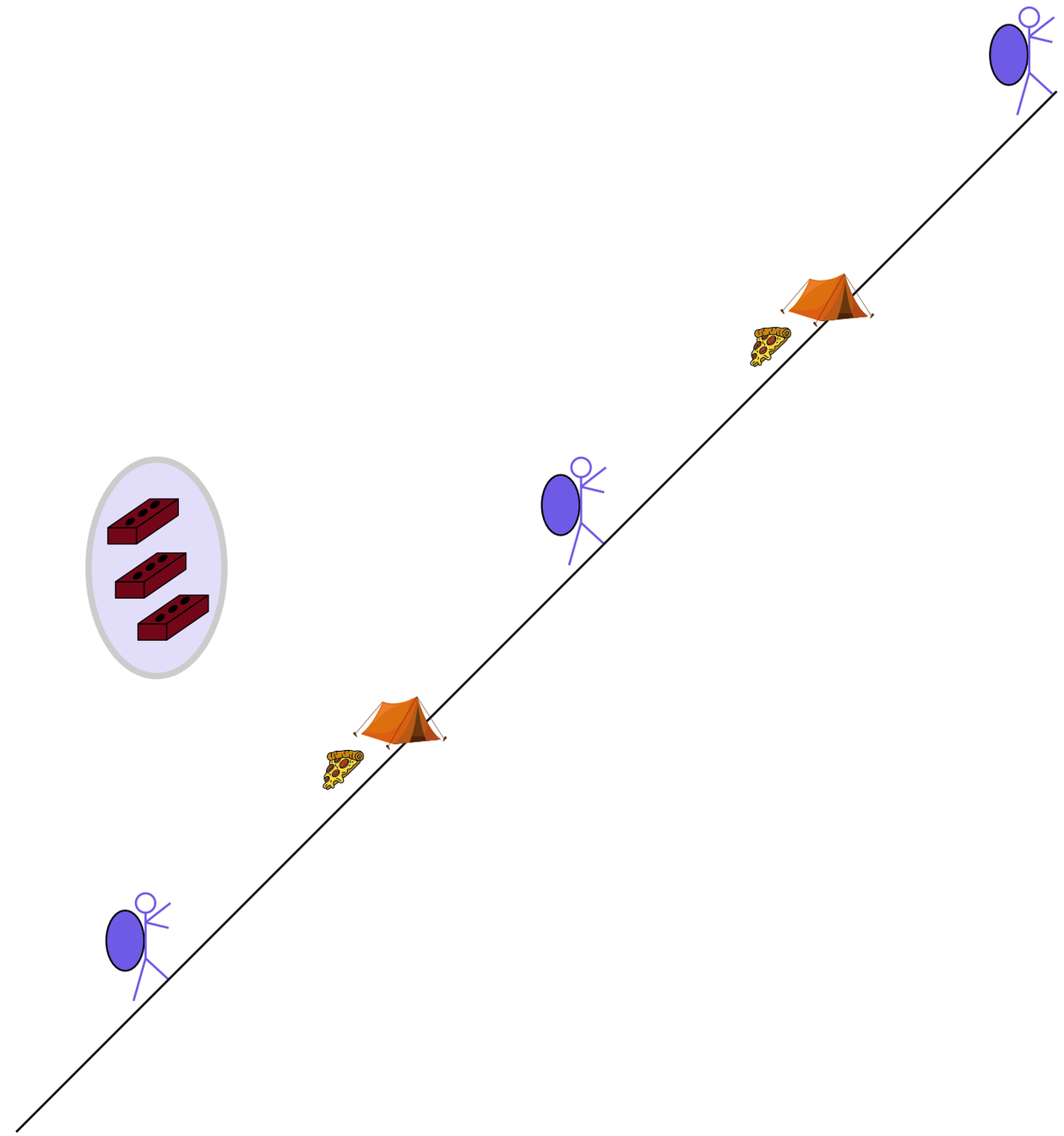
Chitraju, et al. Cell Metabolism, 2017

A metaphor to communicate the *in vitro* and *in vivo* results: climbing Mount Development to become a mature adipocyte -/+ Twist2

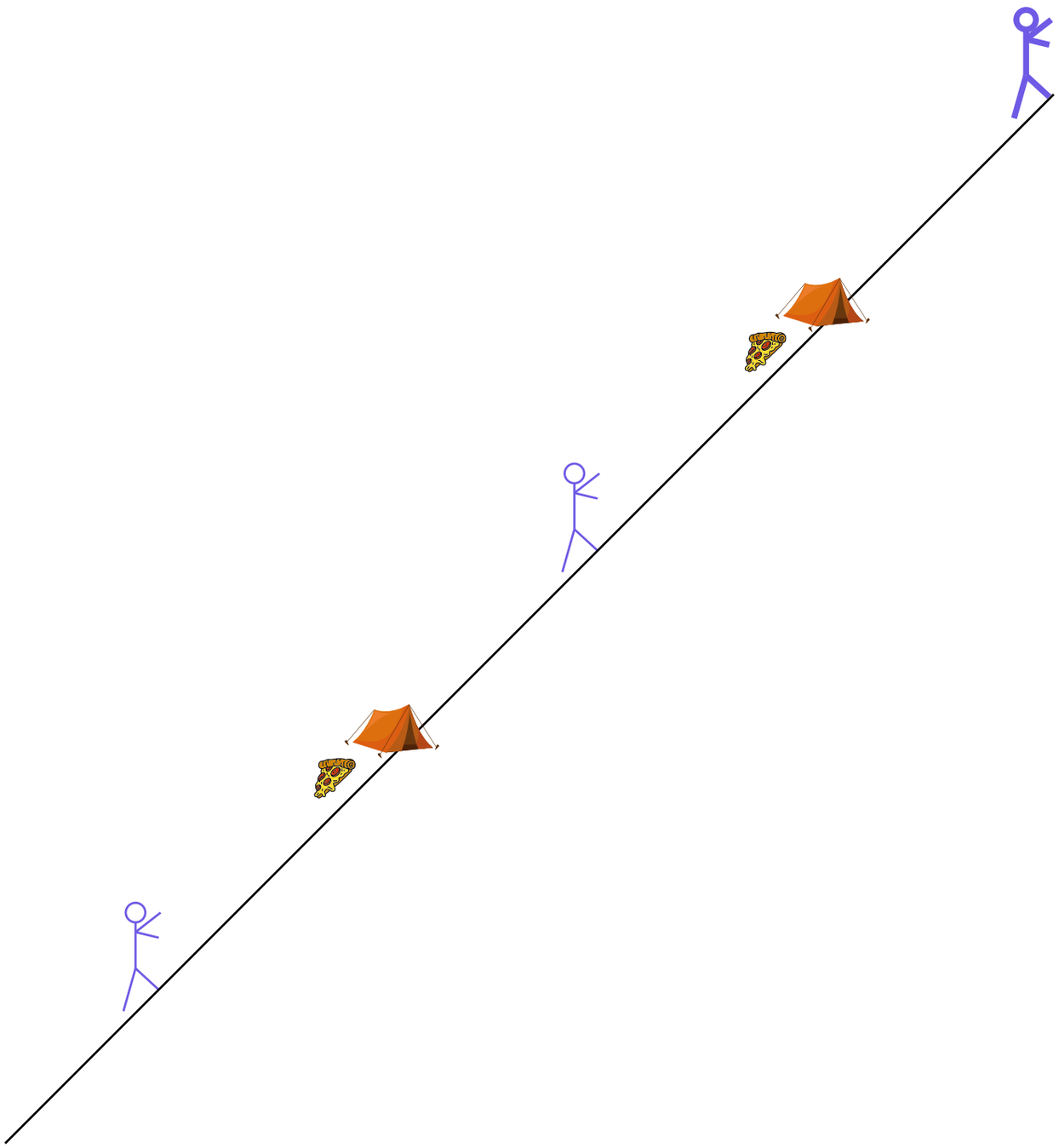
Twist2 is the bricks. Twist2 slows preadipocyte development to keep pace with other tissues.



Twist2 slows you down *in vitro,* but you form adipocytes;
resources are provided, just as we change out the media and split cells

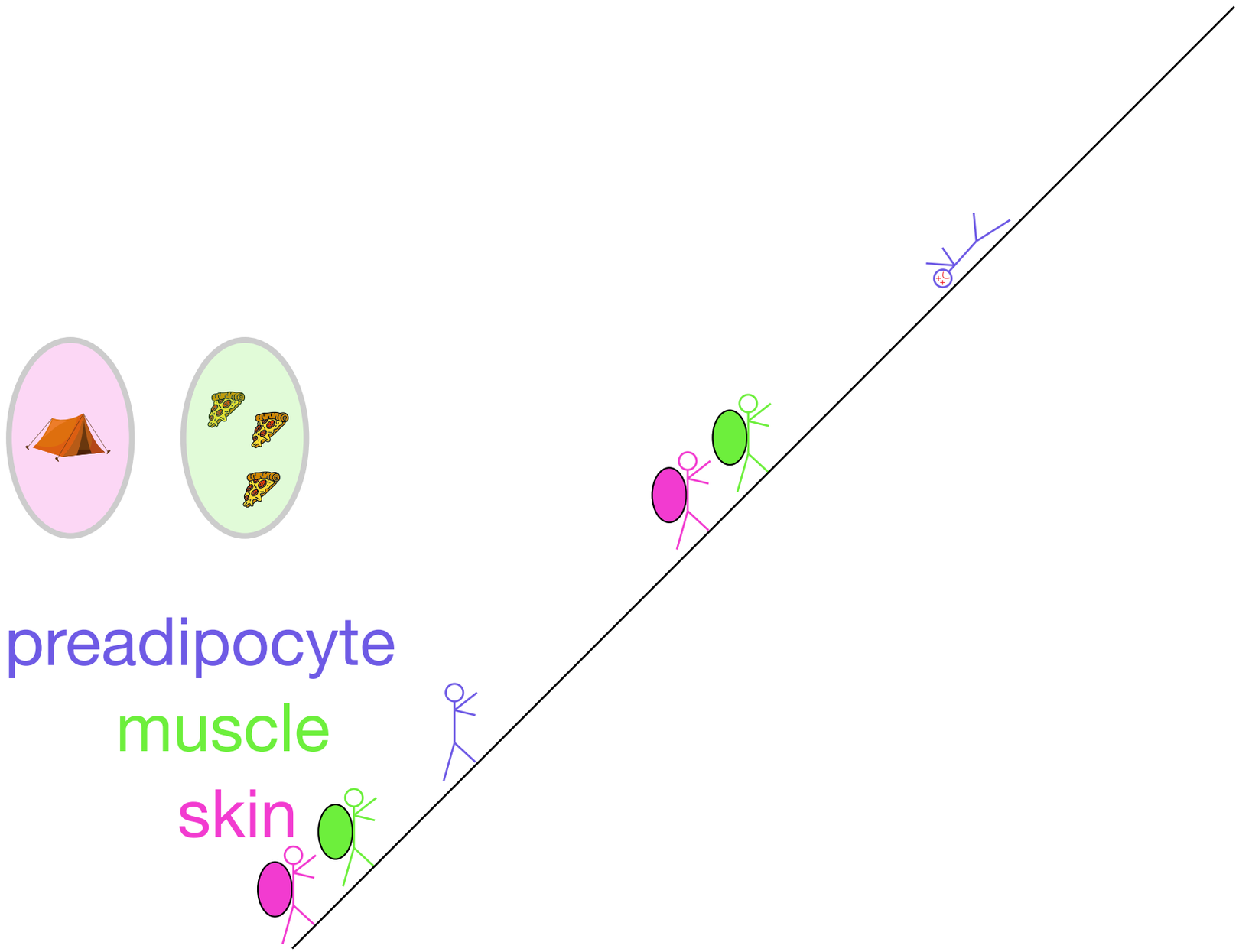


Preadipocytes thrive without Twist2 slowing them down *in vitro;* adipogenesis and fat deposition are more efficient

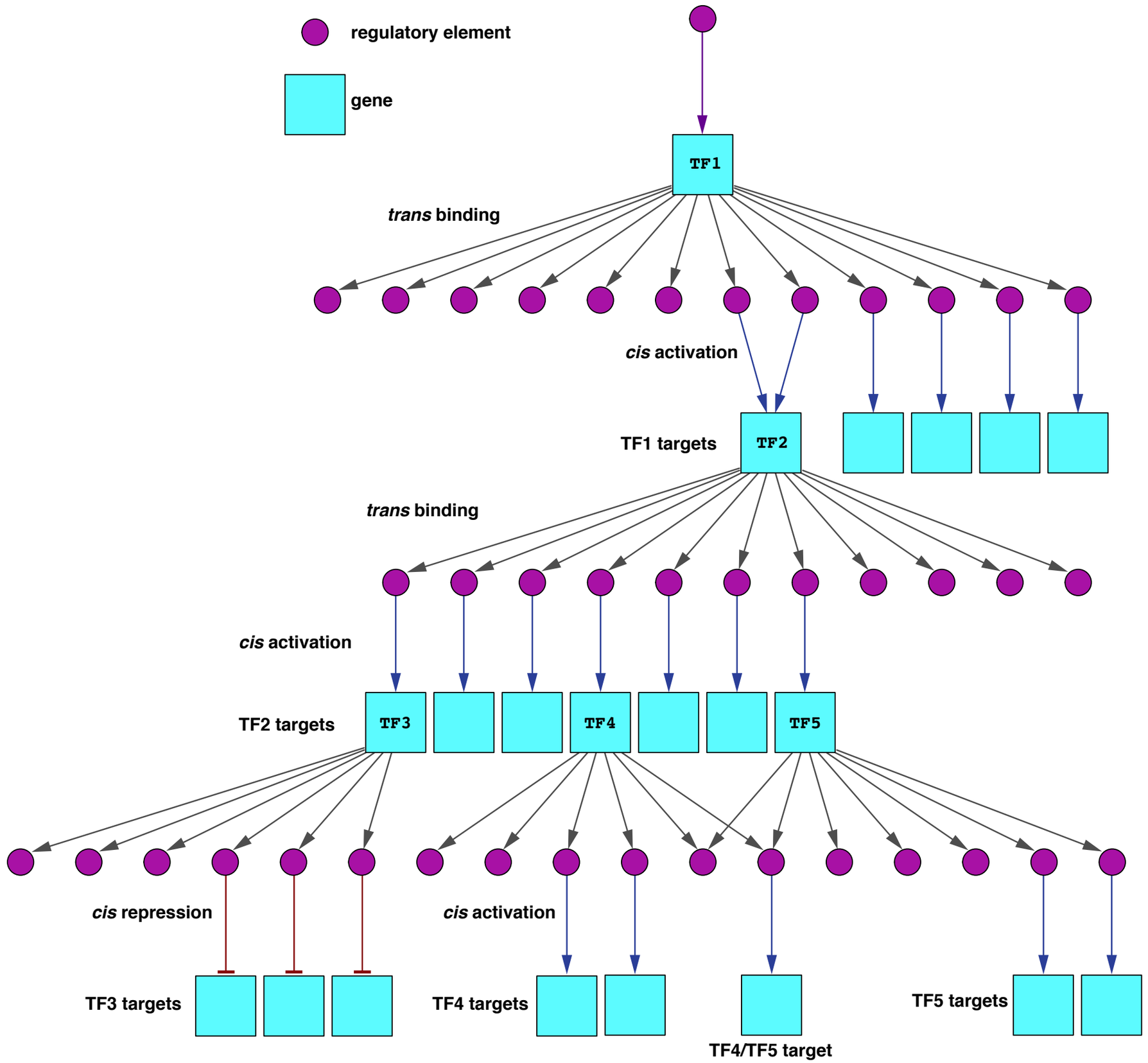


Supporting tissues have resources to ascend Mount Development *in vivo*;
Twist2 forces you to pace with other tissues and you form adipocytes

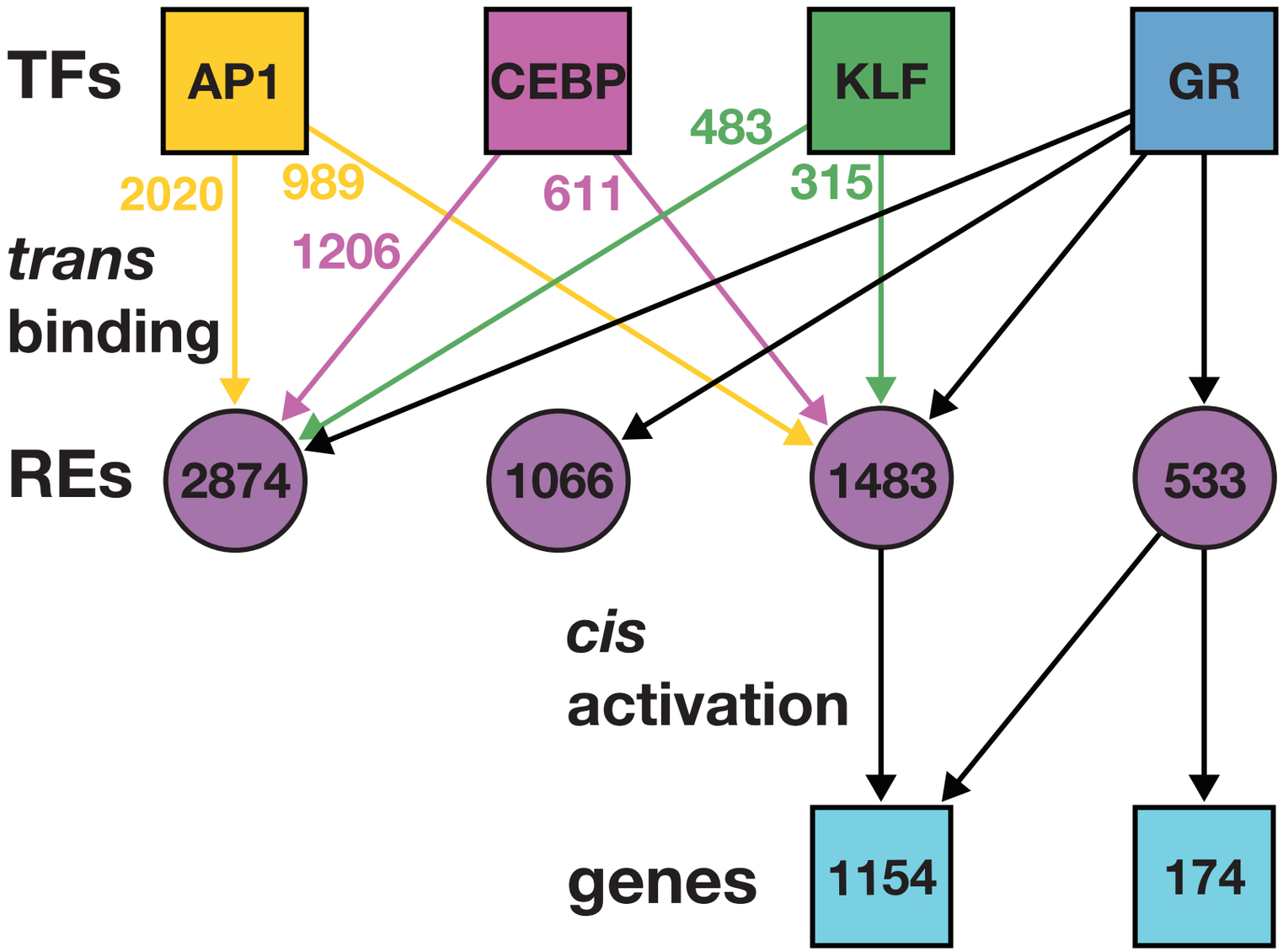
Healthy adipocytes do not develop *in vivo* without Twist2;
preadipocytes ascend too quickly and tissues can not provide support



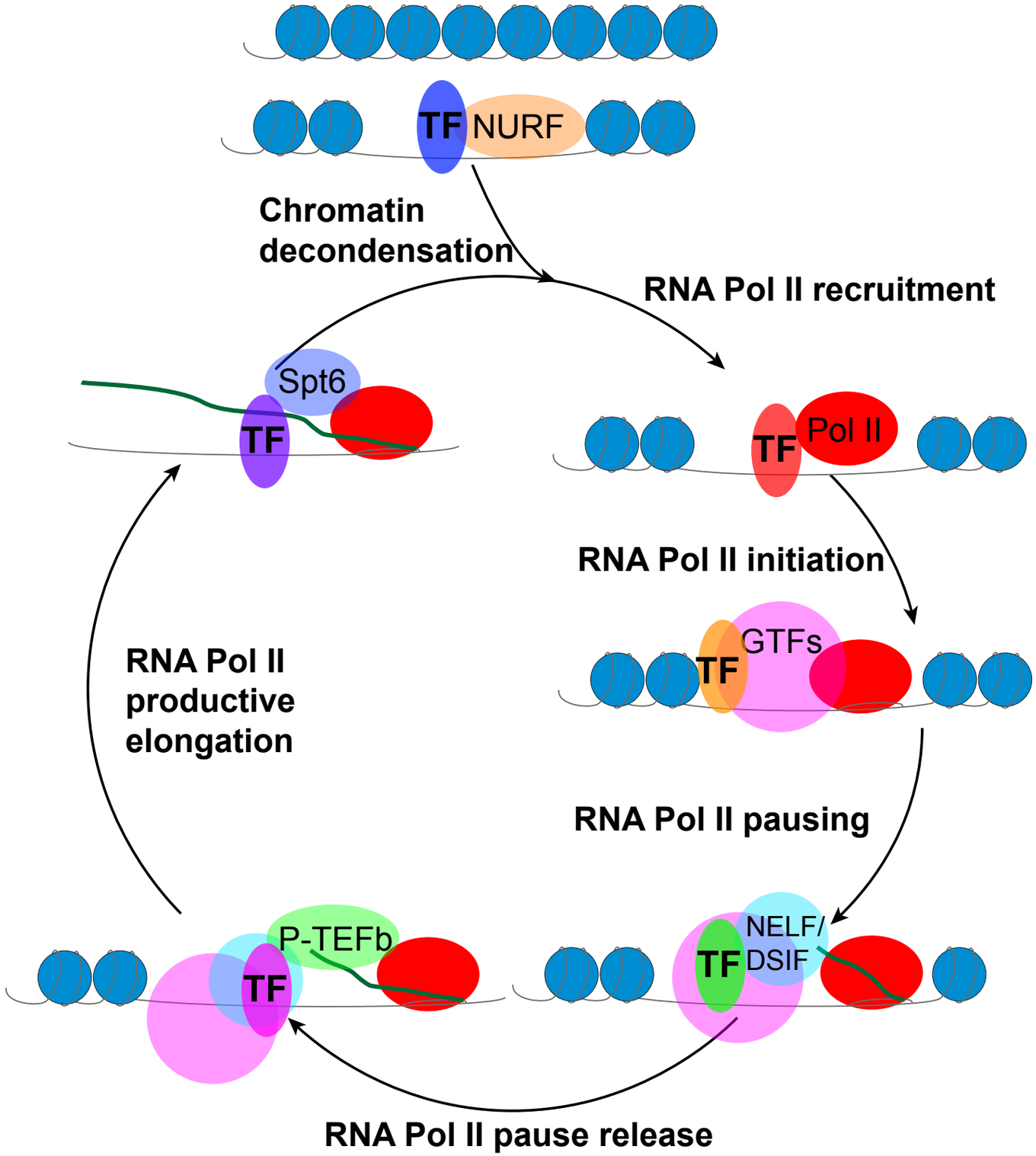
What are the molecular functions of these key TFs?



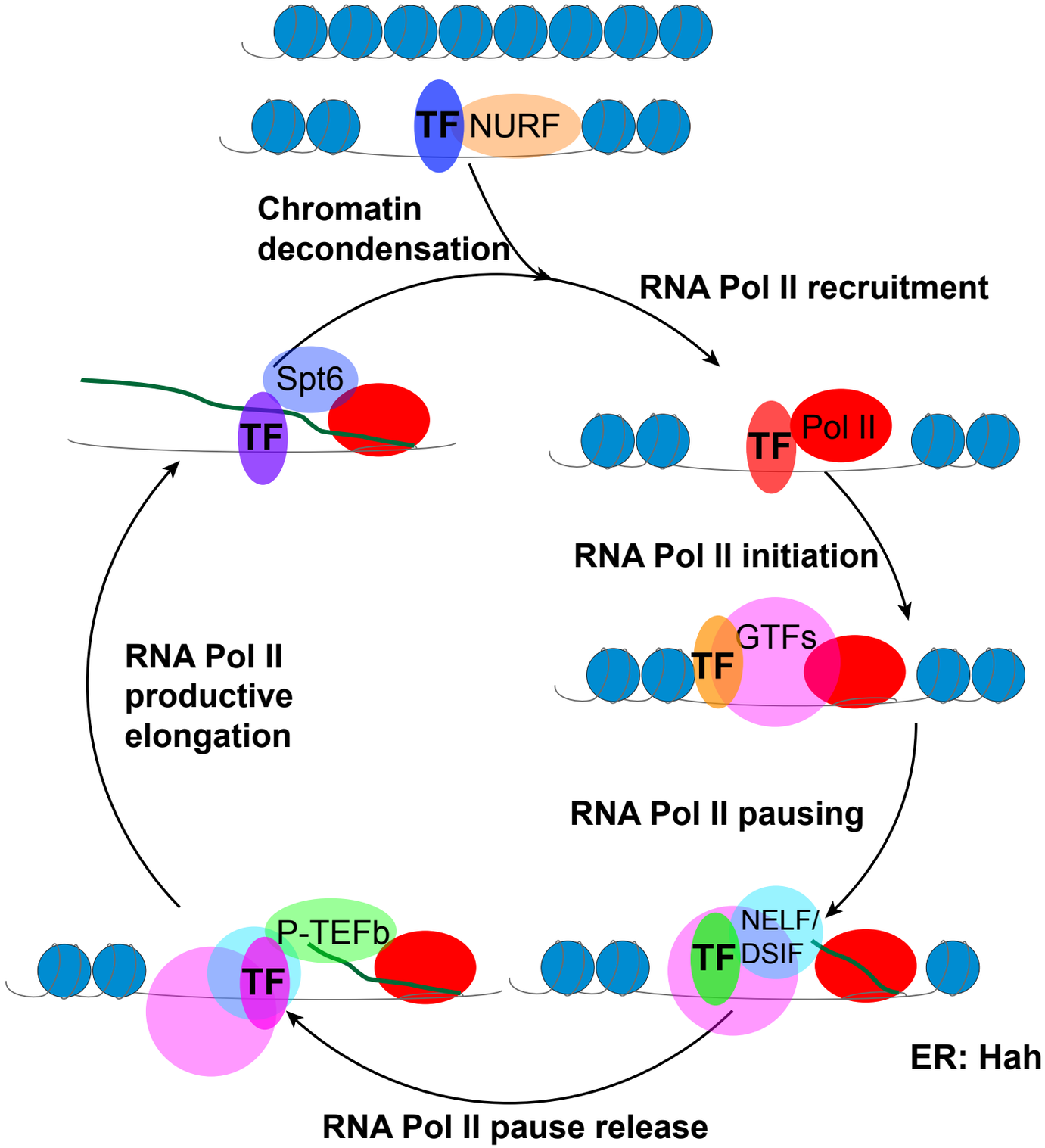
Simplified networks identify genes that are primarily regulated by a single factor

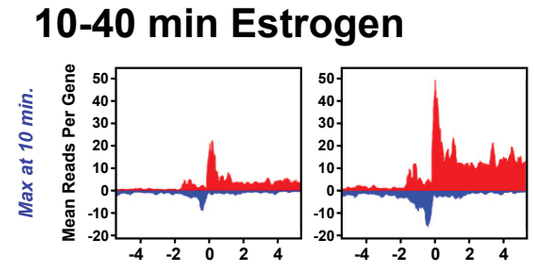


Coupling rapidly inducible systems to nascent transcript profiling informs on TF function



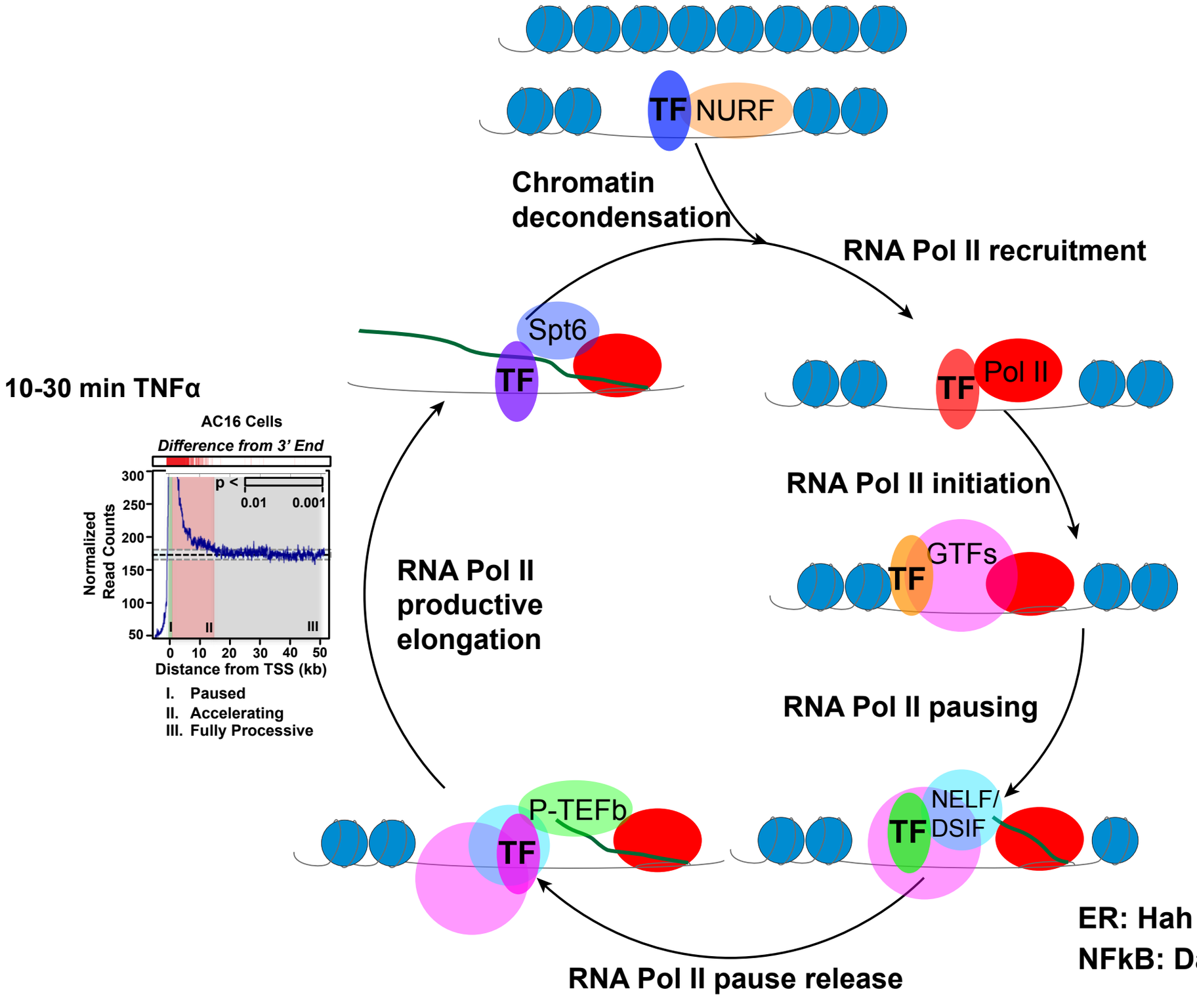
Coupling rapidly inducible systems to nascent transcript profiling informs on TF function





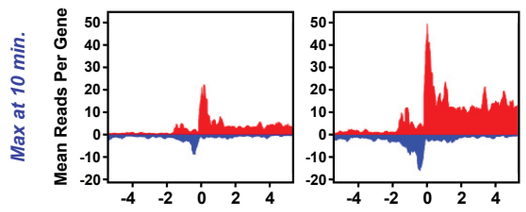
ER: Hah et. al., Cell, 2011

Coupling rapidly inducible systems to nascent transcript profiling informs on TF function



10-30 min TNFα

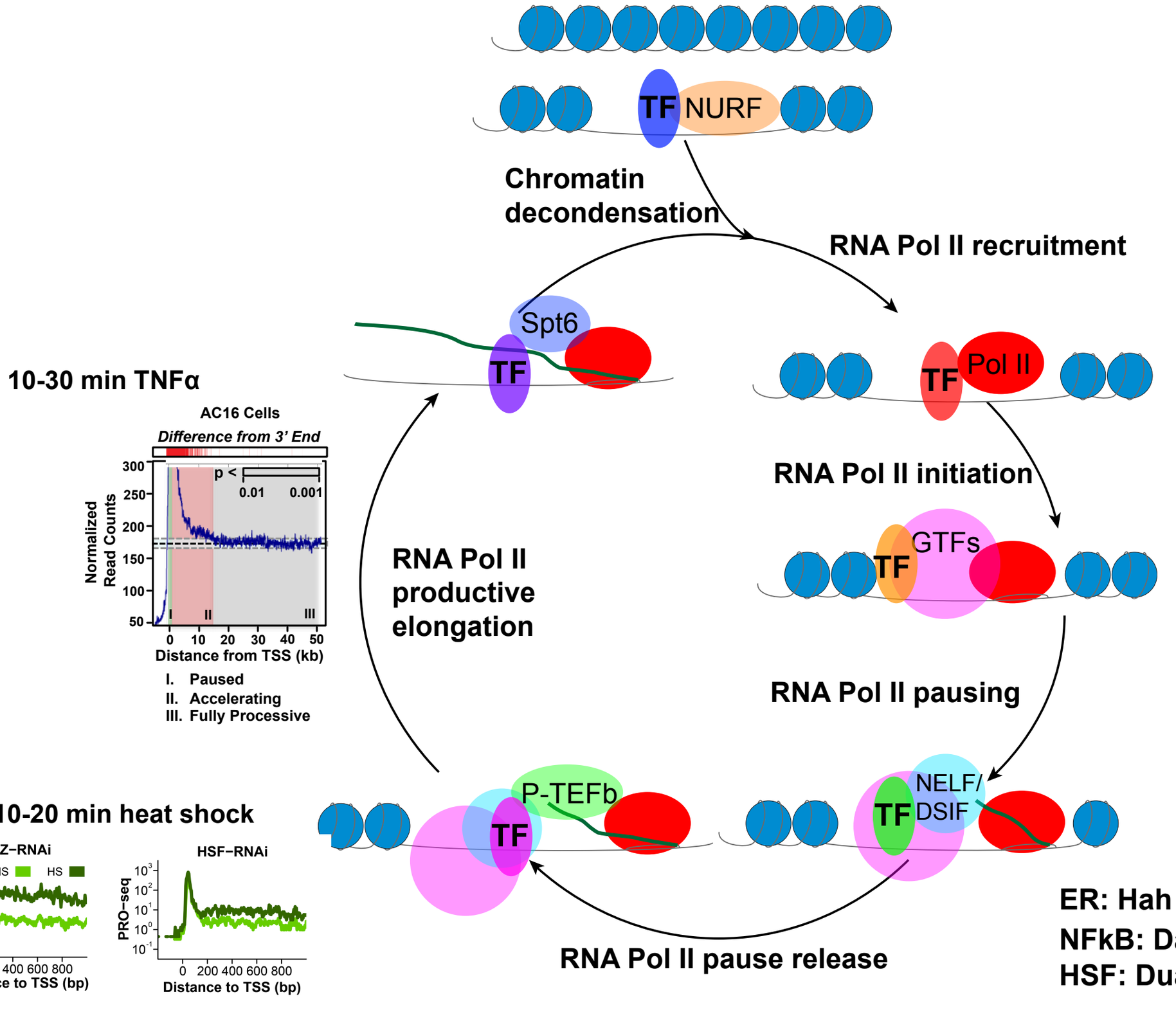
10-40 min Estrogen



ER: Hah et. al., Cell, 2011

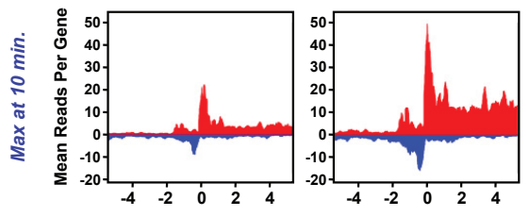
NFkB: Danko, et. al. Mol. Cell, 2014

Coupling rapidly inducible systems to nascent transcript profiling informs on TF function

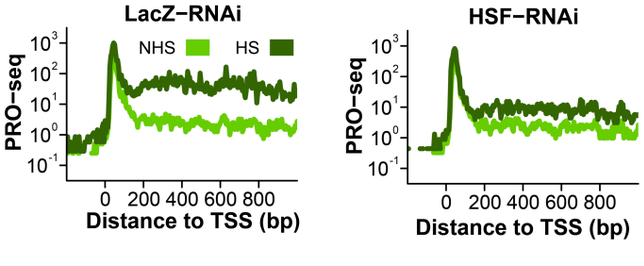


10-30 min TNFα

10-40 min Estrogen



10-20 min heat shock

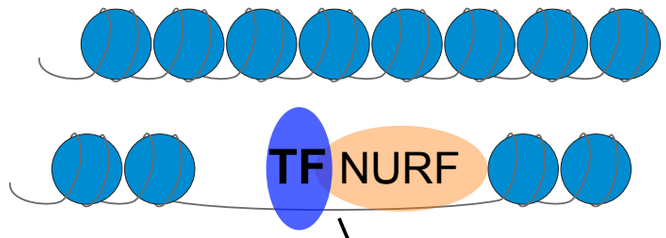


ER: Hah et. al., Cell, 2011

NFkB: Danko, et. al. Mol. Cell, 2014

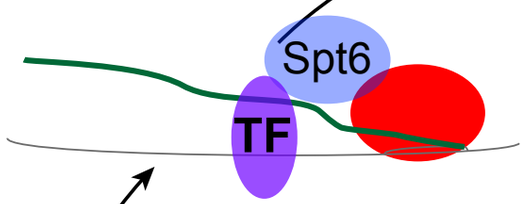
HSF: Duarte et. al., G&D, 2016

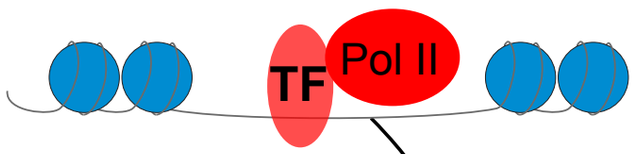
Coupling rapidly inducible systems to nascent transcript profiling informs on TF function



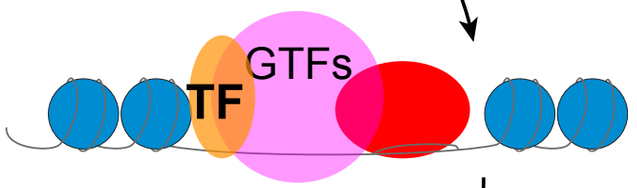
Chromatin decondensation

RNA Pol II recruitment

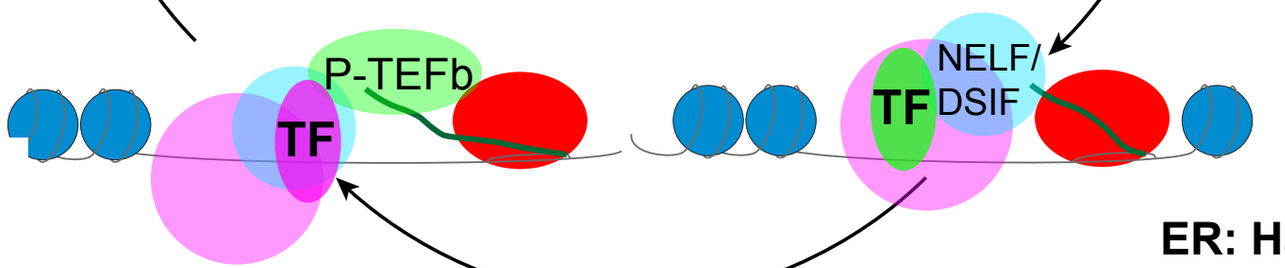




RNA Pol II initiation



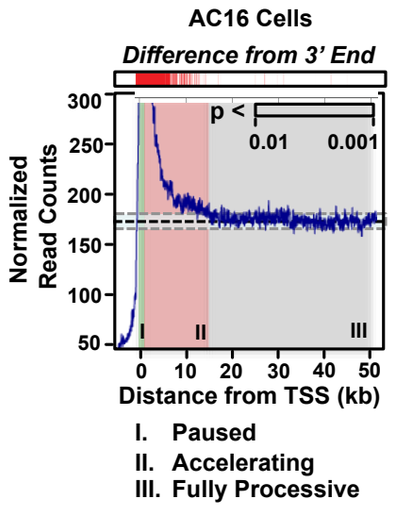
RNA Pol II pausing



RNA Pol II pause release

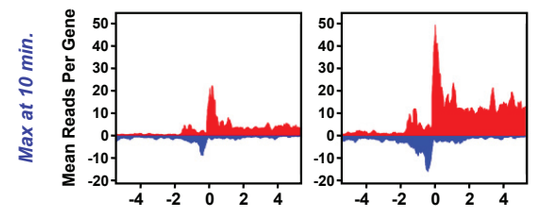
RNA Pol II productive elongation

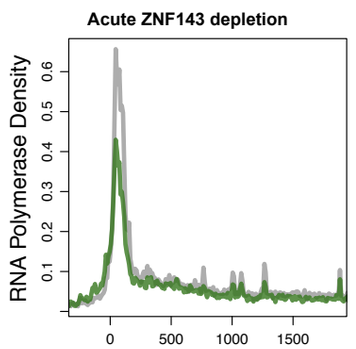
10-30 min TNFα



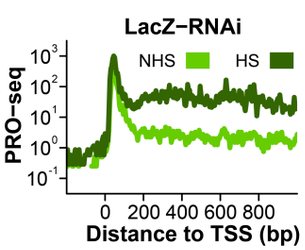
I. Paused
II. Accelerating
III. Fully Processive

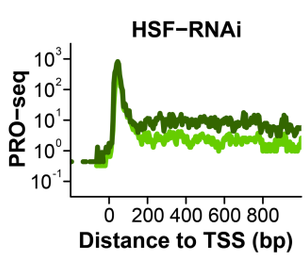
10-40 min Estrogen





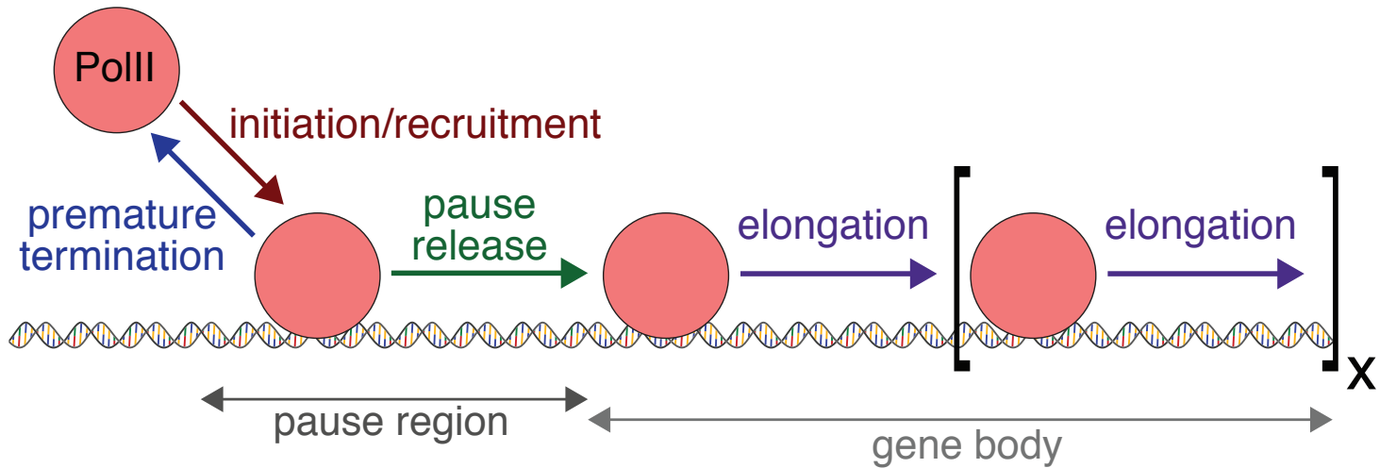
10-20 min heat shock

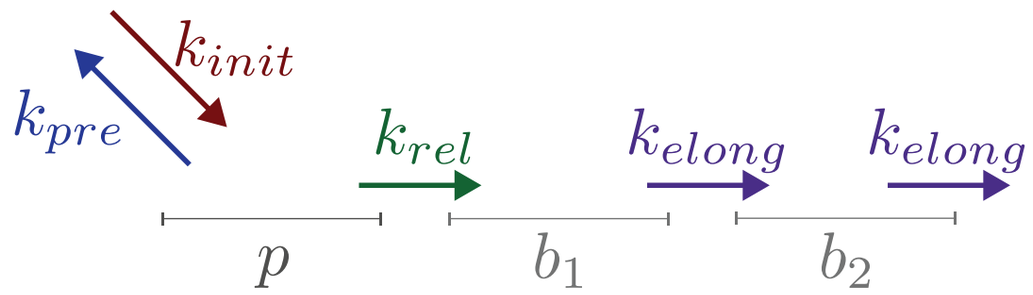


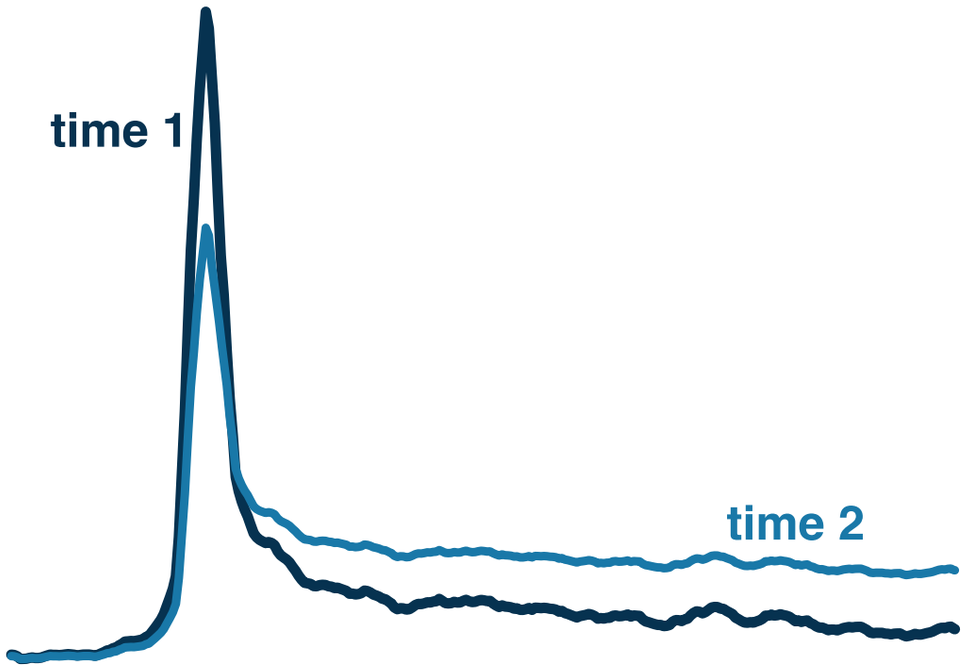


ER: Hah et. al., Cell, 2011
NFkB: Danko, et. al. Mol. Cell, 2014
HSF: Duarte et. al., G&D, 2016
ZNF143: Mattada et. al., G&D, 2019

Coupling rapidly inducible systems to nascent transcript profiling informs on TF function

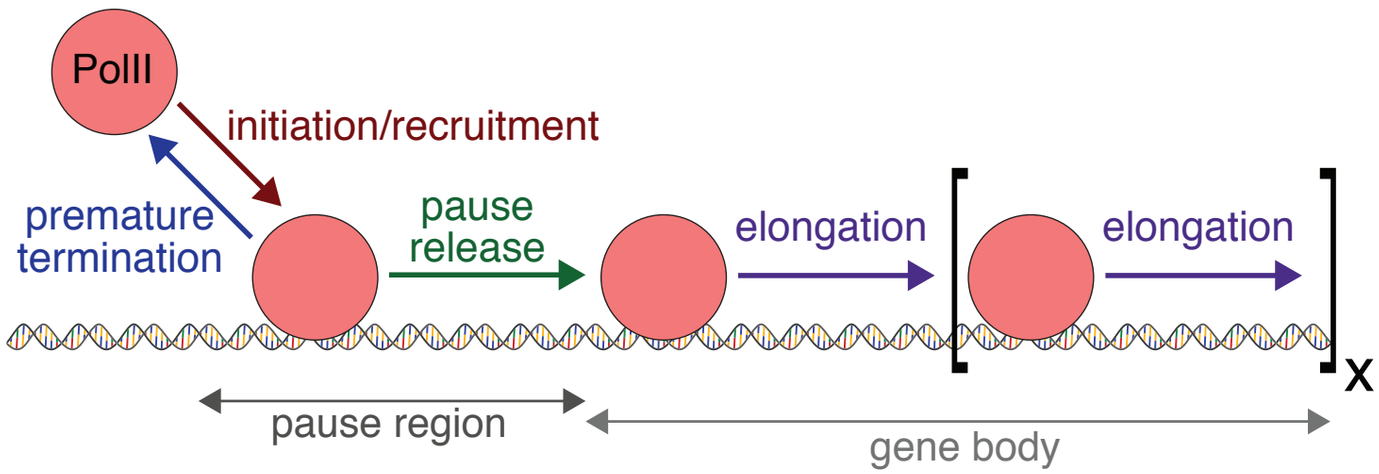


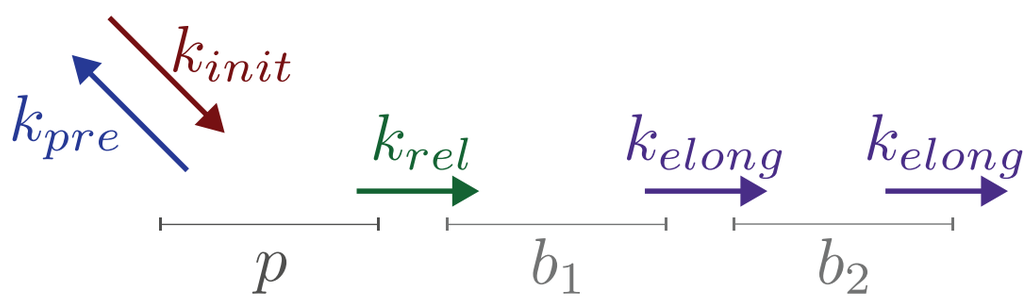




RNA polymerase composite profile

We can determine the step(s) that a TF regulates by quantifying RNA polymerase density changes in genic regions





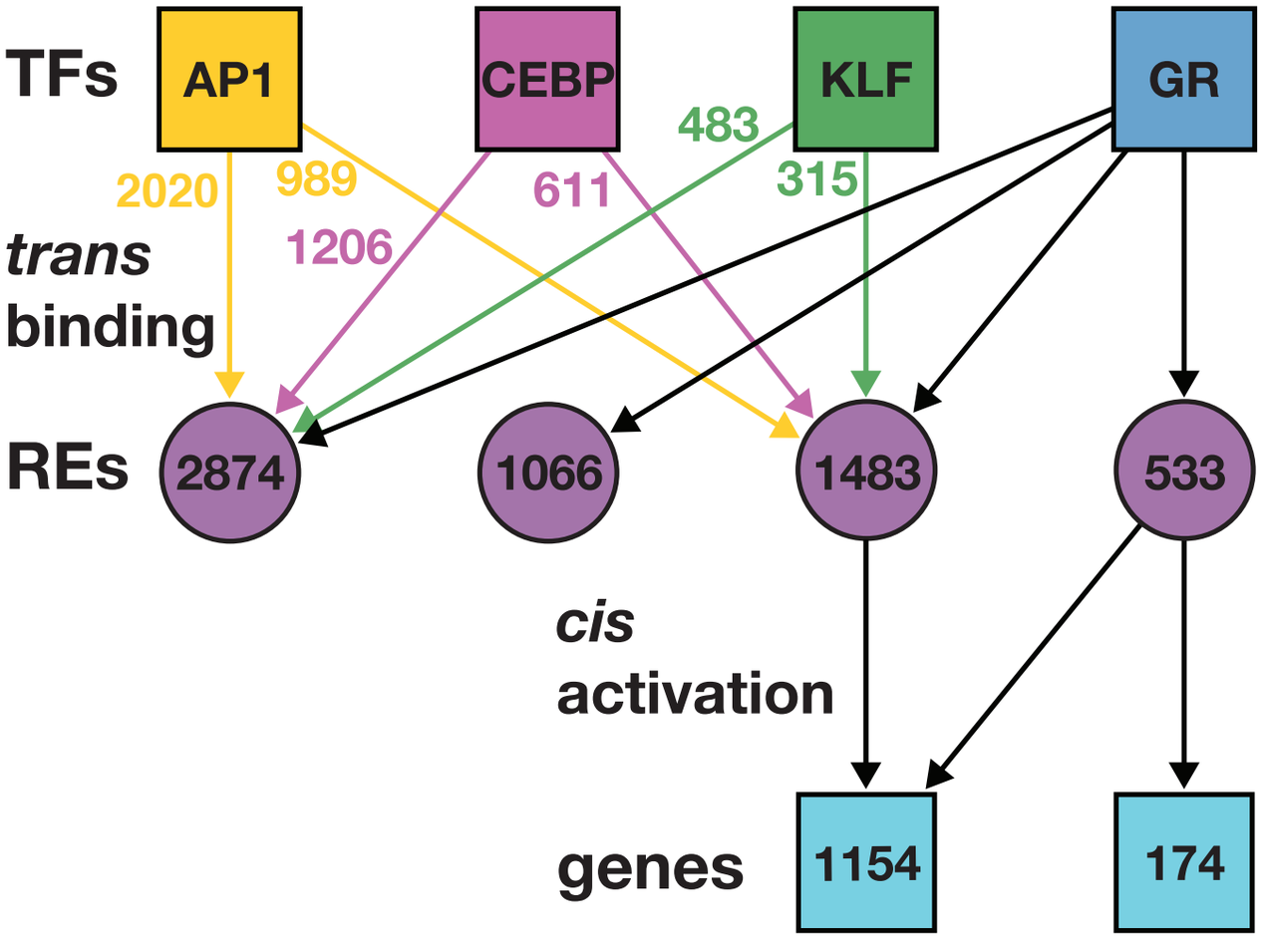
$$\frac{dp}{dt} = k_{init} - (k_{pre} + k_{rel})\, p$$

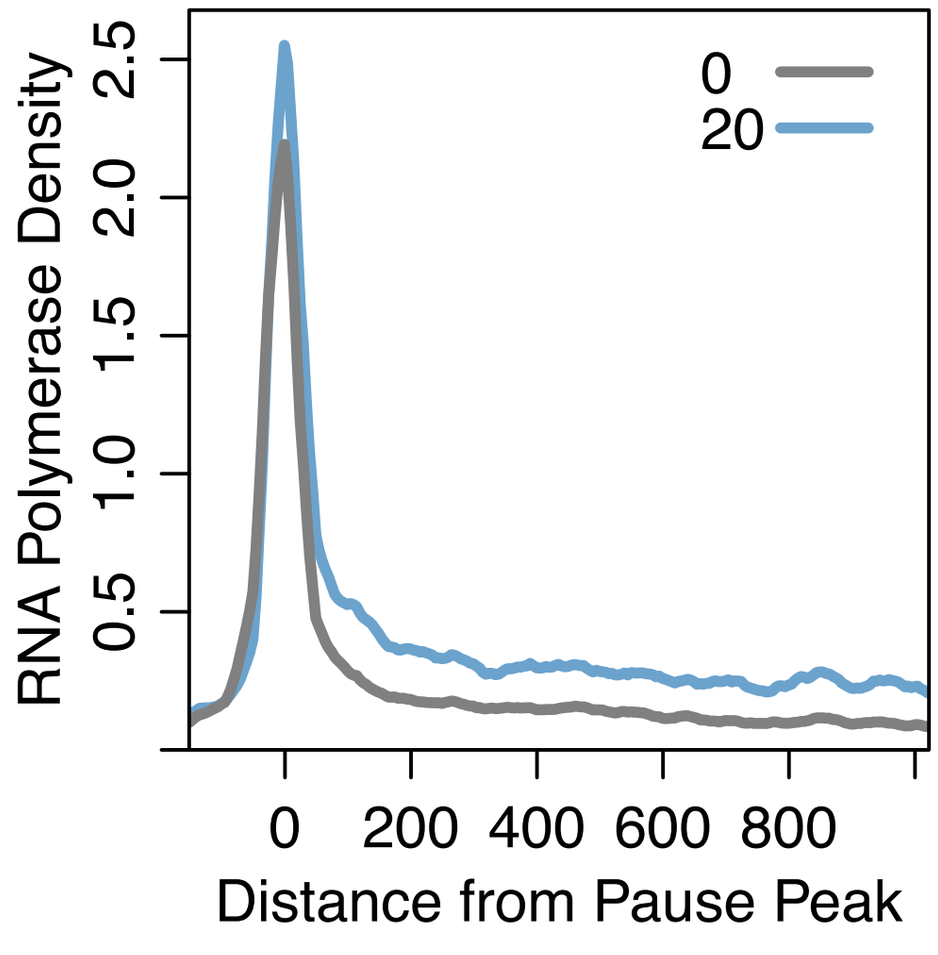
$$\frac{db_1}{dt} = k_{rel}\, p - k_{elong}\, b_1$$

$$\frac{db_2}{dt} = k_{elong}\, b_1 - k_{elong}\, b_2$$

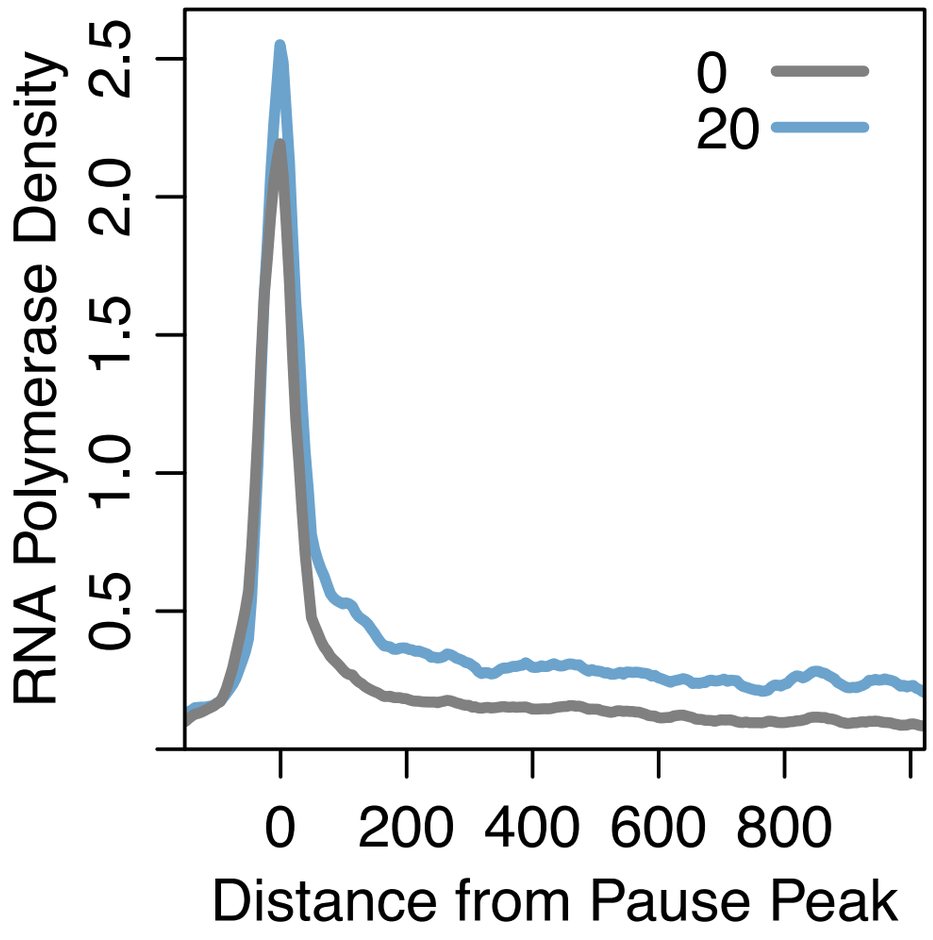
$$\frac{db_x}{dt} = k_{elong}\, b_{x-1} - k_{elong}\, b_x$$

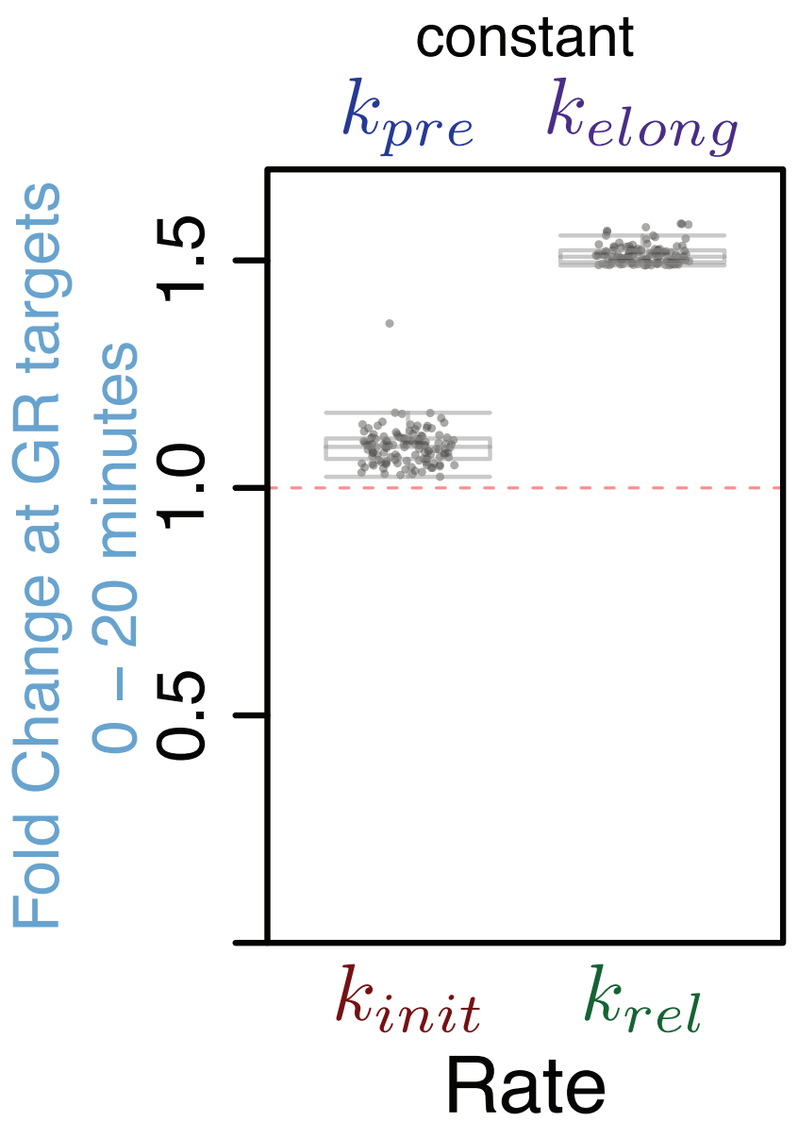
Simplified networks identify genes that are primarily regulated by a single factor



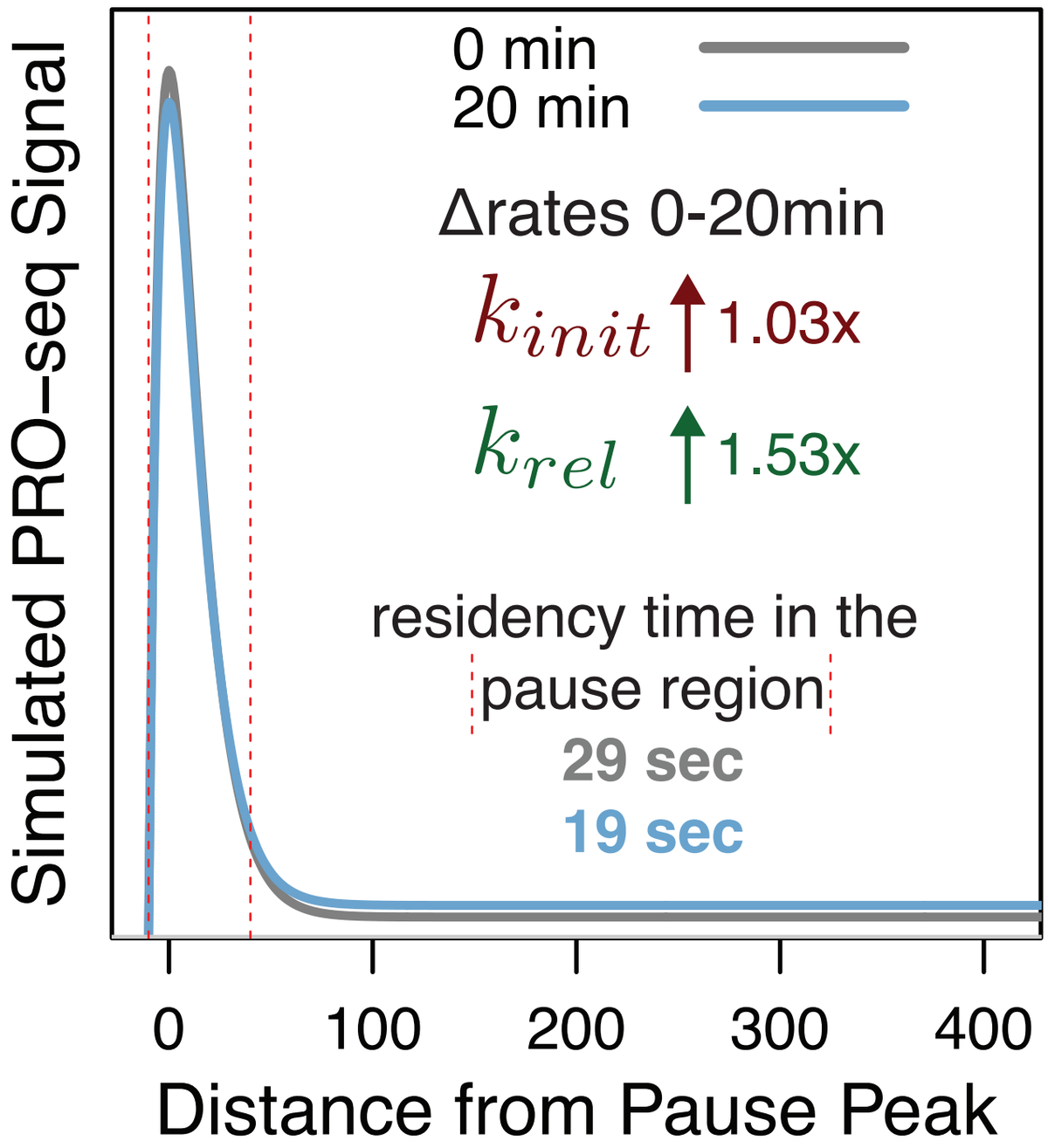


GR preferentially regulates pause release

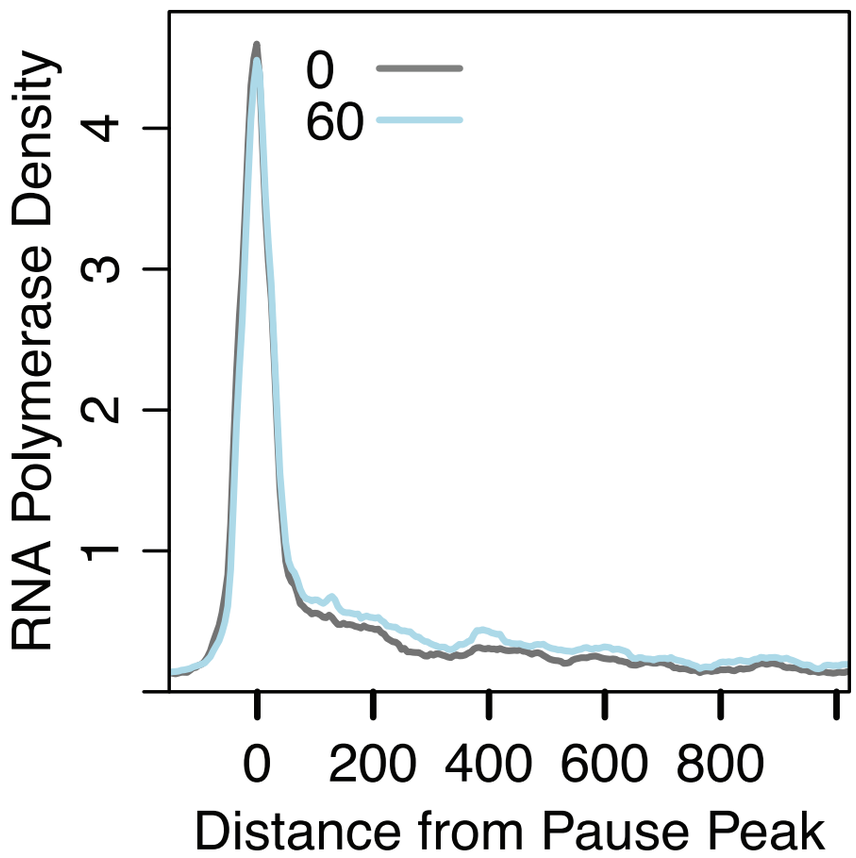


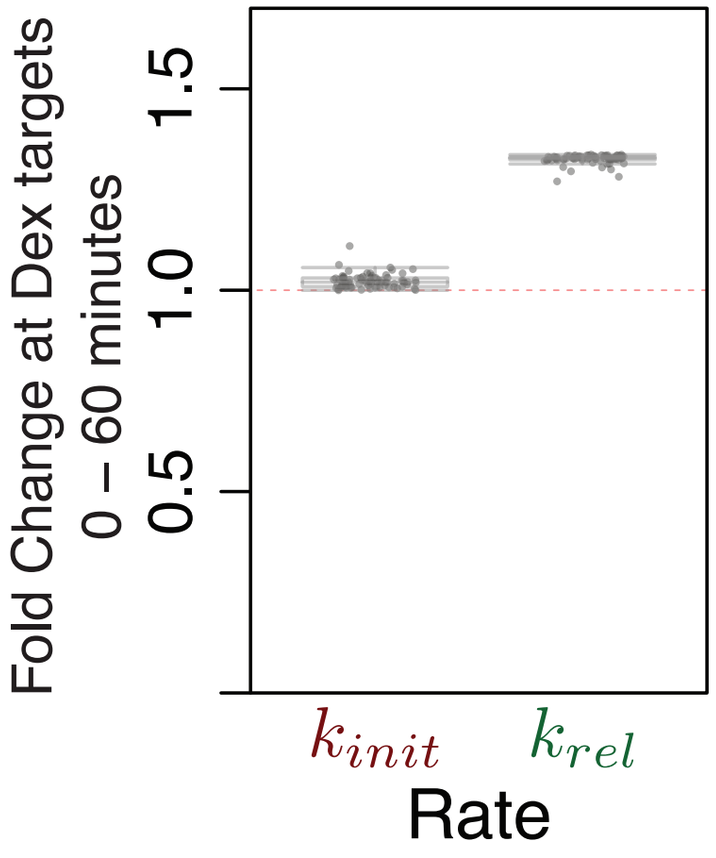


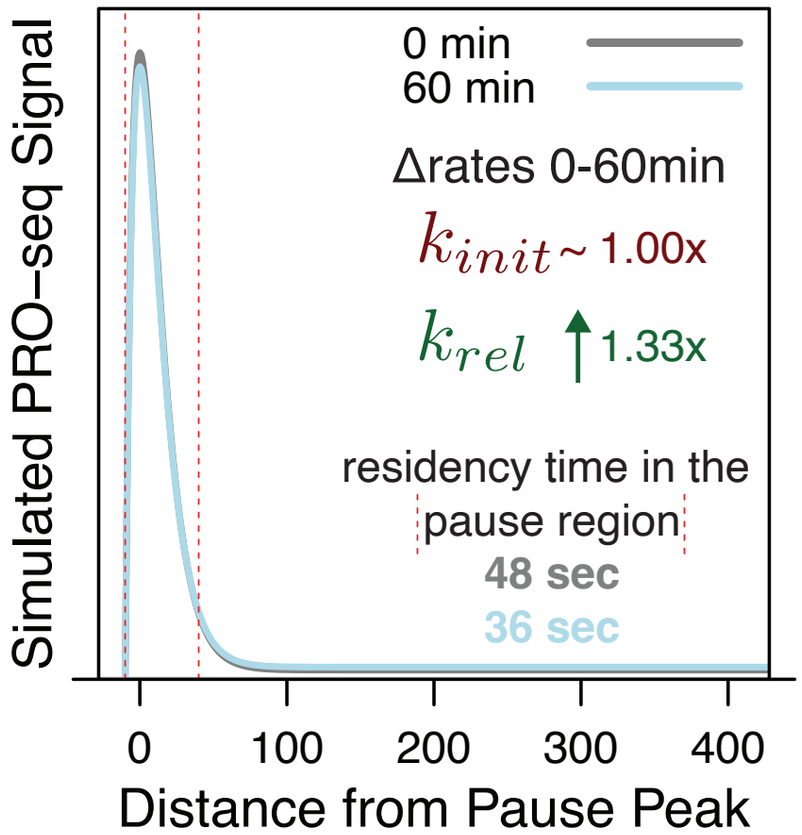
GR preferentially regulates pause release



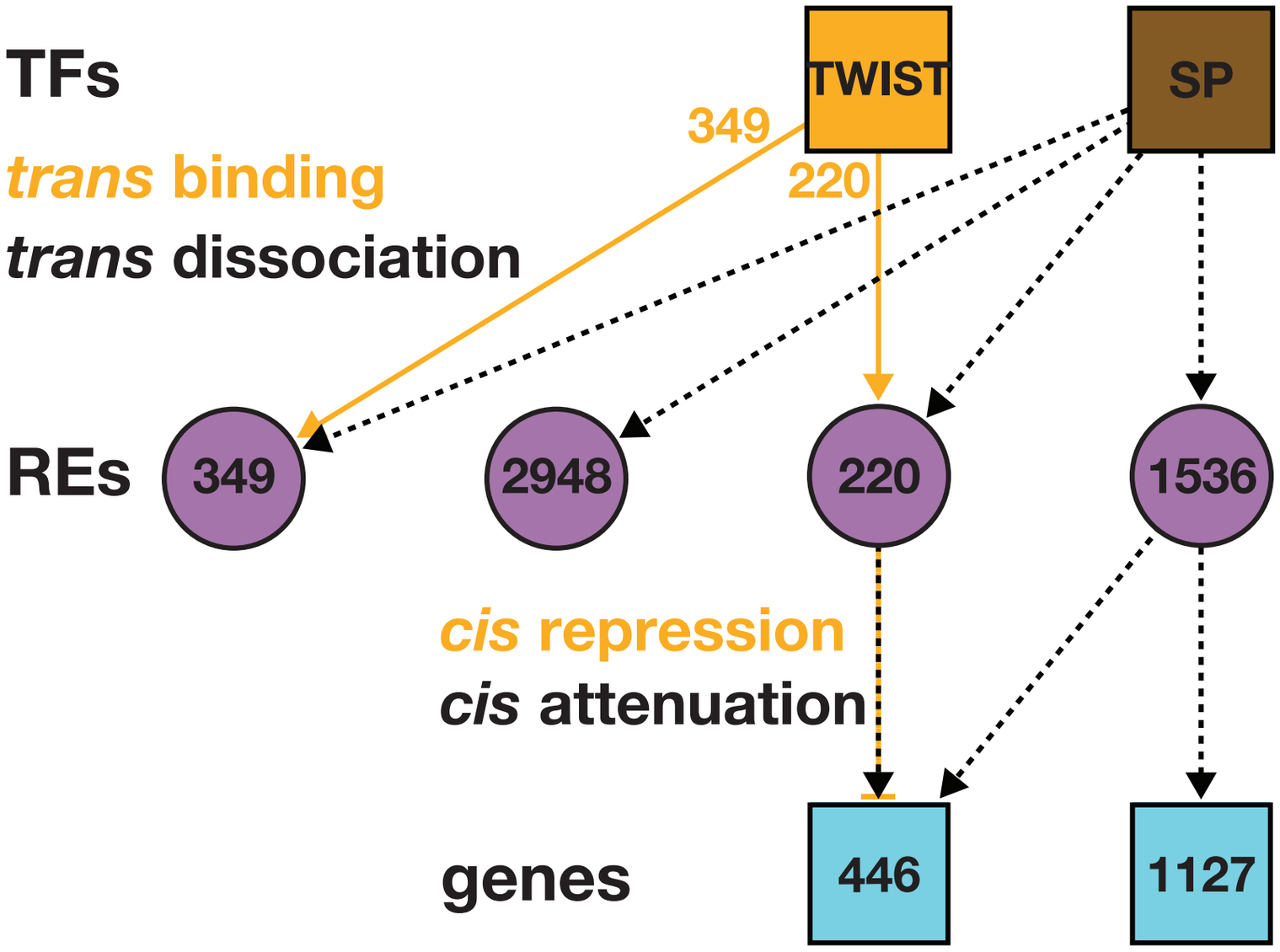
GR preferentially regulates pause release in another system: Leukemia cells treated with dexamethasone for 1 hour



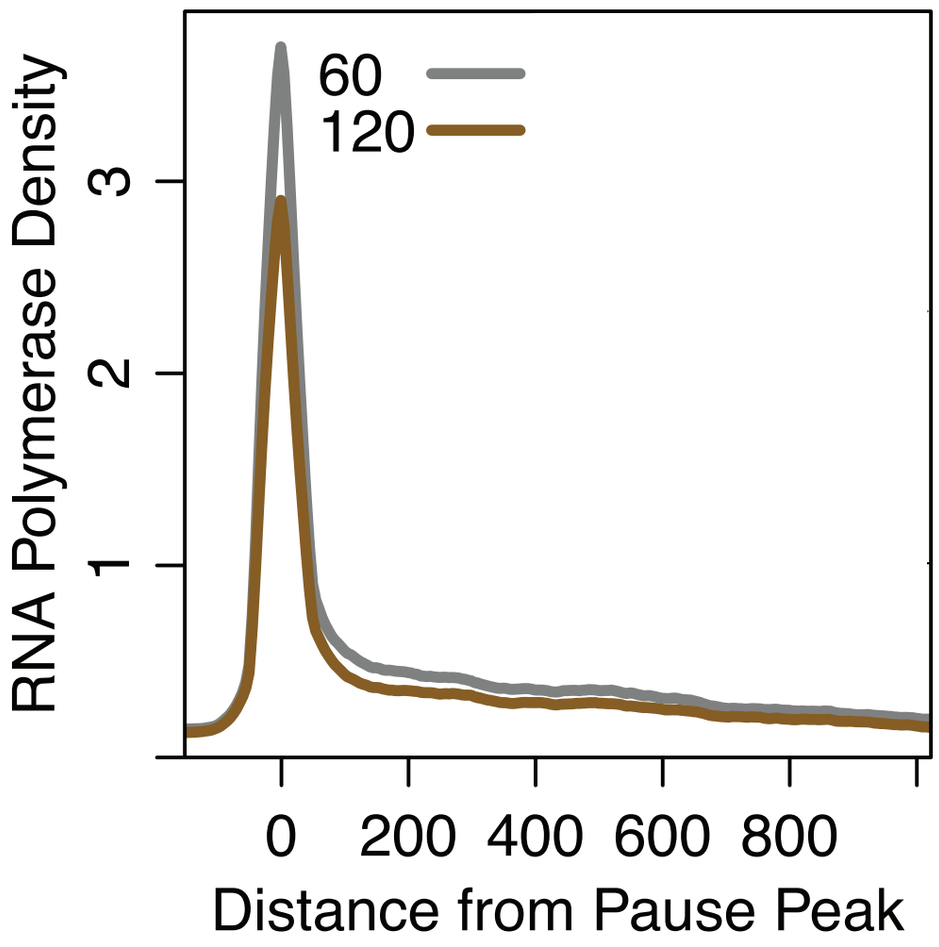


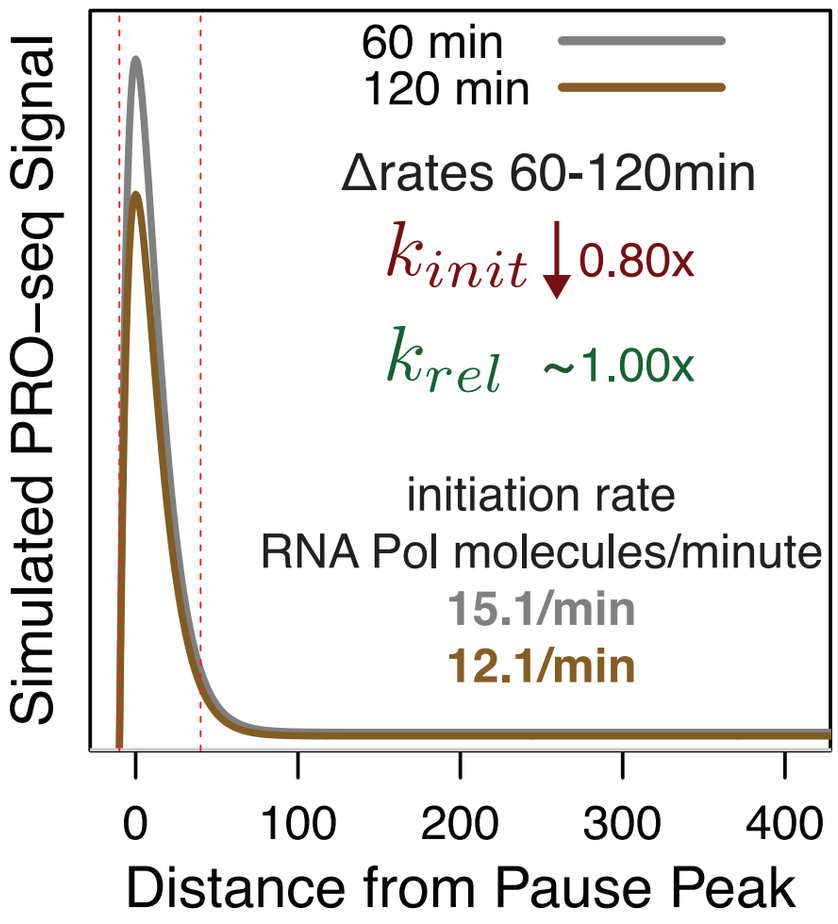


Simplified networks identify genes that are primarily regulated by a single factor



SP preferentially regulates initiation rate





Adipogenesis Network Conclusions

Adipogenesis Network Conclusions

- Kinetic ATAC and PRO data can be used to infer key functional interactions between regulatory elements and genes.

Adipogenesis Network Conclusions

- Kinetic ATAC and PRO data can be used to infer key functional interactions between regulatory elements and genes.

- TWIST2 is a highly connected node in the adipogenesis network.

Adipogenesis Network Conclusions

- Kinetic ATAC and PRO data can be used to infer key functional interactions between regulatory elements and genes.

- TWIST2 is a highly connected node in the adipogenesis network.

- TWIST2 deletion cause an absence of subcutaneous fat deposits.

Adipogenesis Network Conclusions

- Kinetic ATAC and PRO data can be used to infer key functional interactions between regulatory elements and genes.

- TWIST2 is a highly connected node in the adipogenesis network.

- TWIST2 deletion cause an absence of subcutaneous fat deposits.

- GR regulates pause release

Adipogenesis Network Conclusions

- Kinetic ATAC and PRO data can be used to infer key functional interactions between regulatory elements and genes.

- TWIST2 is a highly connected node in the adipogenesis network.

- TWIST2 deletion cause an absence of subcutaneous fat deposits.

- GR regulates pause release

- SP regulates initiation

Adipogenesis Network Conclusions

- Kinetic ATAC and PRO data can be used to infer key functional interactions between regulatory elements and genes.

- TWIST2 is a highly connected node in the adipogenesis network.

- TWIST2 deletion cause an absence of subcutaneous fat deposits.

- GR regulates pause release

- SP regulates initiation

- Rapidly inducilbe systems are necessary to provide these mechanistic insights…

Challenges with RNAi depletion and knock outs

Challenges with RNAi depletion and knock outs

- Chronic depletion can result in compensatory feedback.

Challenges with RNAi depletion and knock outs

- Chronic depletion can result in compensatory feedback.

- Secondary (and post secondary) effects can dominate after extended depletion.

Challenges with RNAi depletion and knock outs

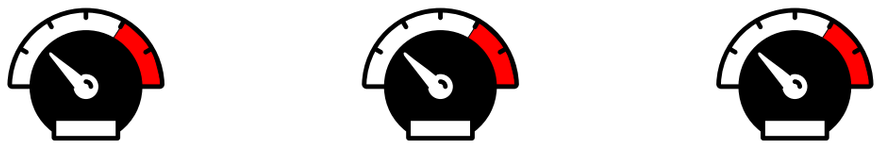
- Chronic depletion can result in compensatory feedback.
- Secondary (and post secondary) effects can dominate after extended depletion.
- Extended depletion can be lethal.

Challenges with RNAi depletion and knock outs

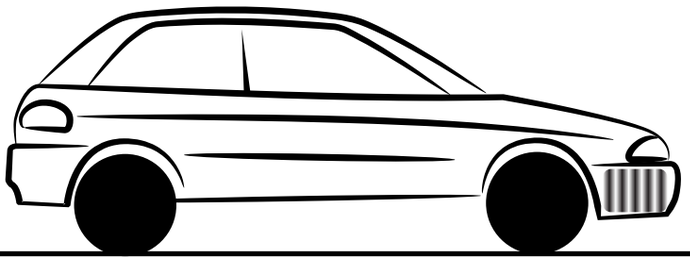
- Chronic depletion can result in compensatory feedback.
- Secondary (and post secondary) effects can dominate after extended depletion.
- Extended depletion can be lethal.
- Degron tagging is an attractive alternative

What does a car’s radiator do?

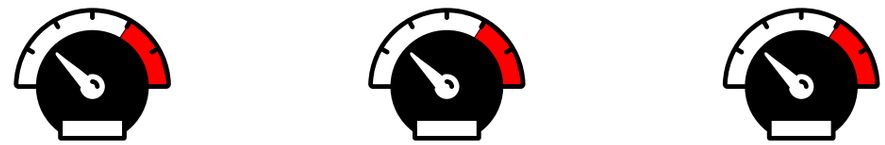
Speed **Tire Pressure** **Temperature**



Poke small holes in radiator (3 days to fully deplete fluid)

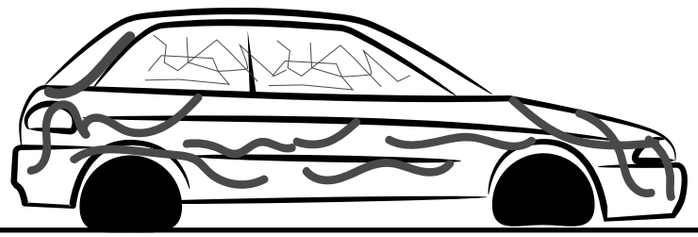


Speed **Tire Pressure** **Temperature**

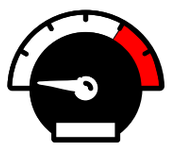


Control

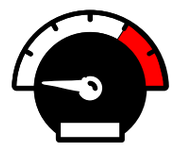
What does a car’s radiator do?



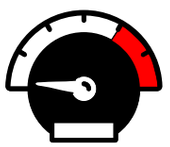
Speed



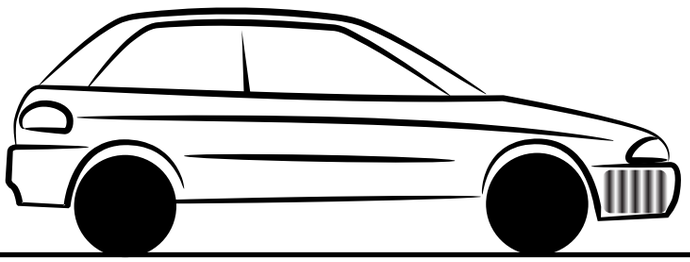
Tire Pressure



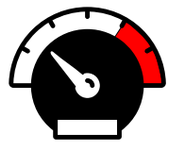
Temperature



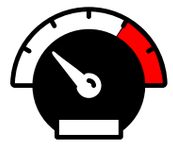
At some point on day 3



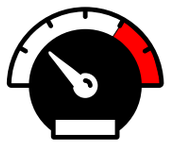
Speed



Tire Pressure

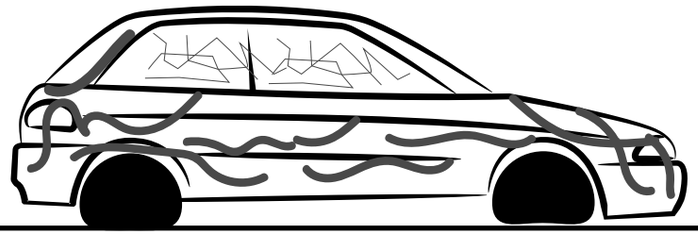


Temperature

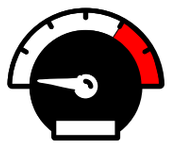


Control

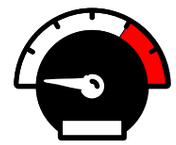
What does a car’s radiator do?



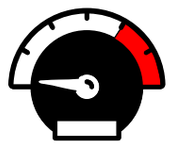
Speed



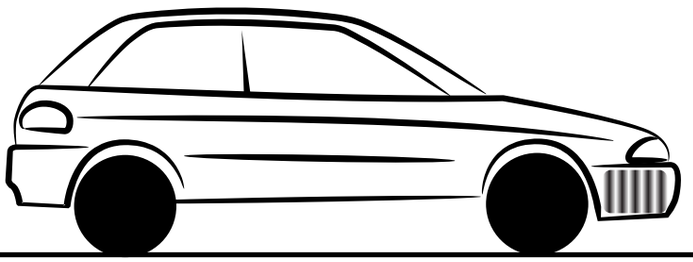
Tire Pressure



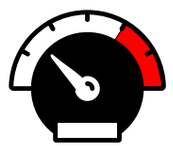
Temperature



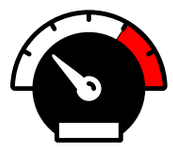
Assess gauges and car phenotypes on day 4



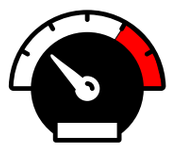
Speed



Tire Pressure



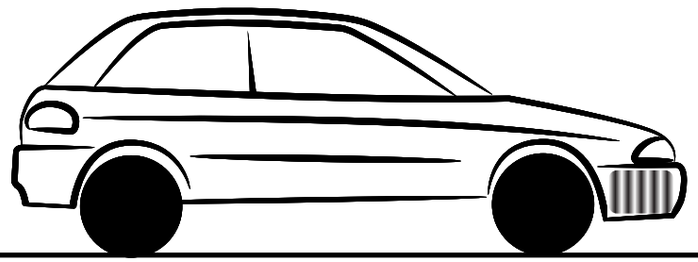
Temperature



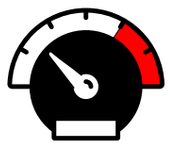
Control

The radiator keeps the car healthy;
The radiator affects the starter, speed, tire pressure, and temperature

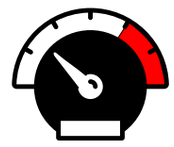
What does a car’s radiator do?



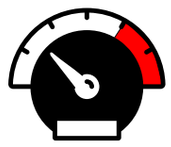
Speed



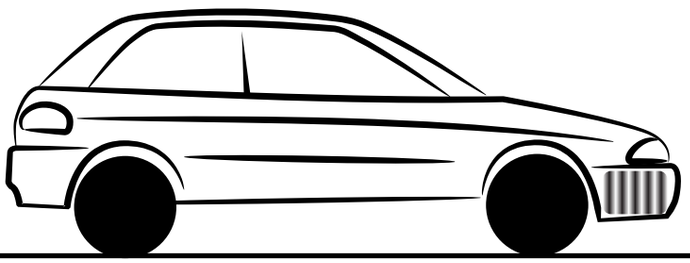
Tire Pressure



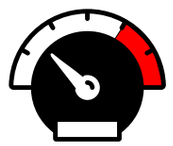
Temperature



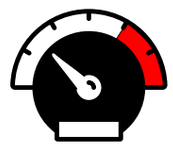
Rip out the radiator while driving



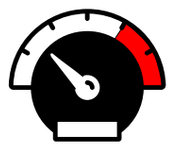
Speed



Tire Pressure

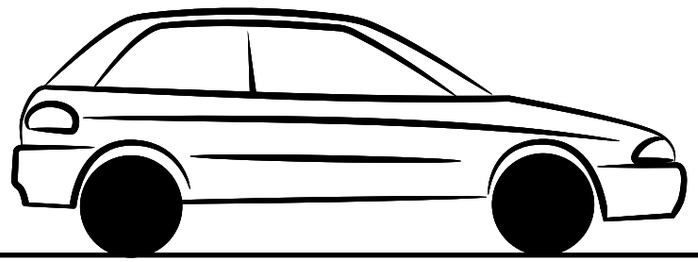


Temperature

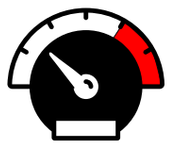


Control

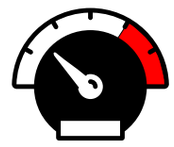
What does a car’s radiator do?



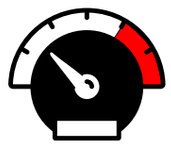
Speed



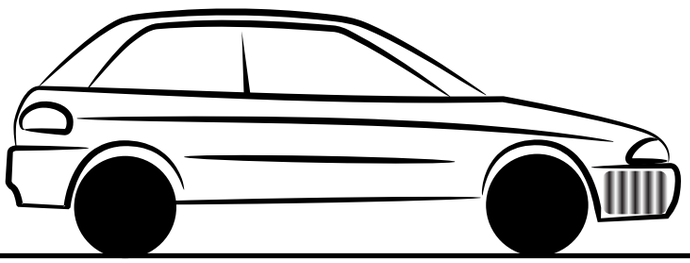
Tire Pressure



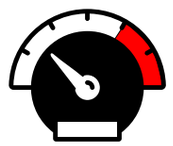
Temperature



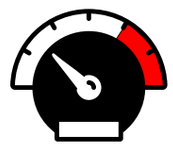
Rip out the radiator while driving



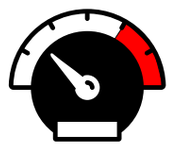
Speed



Tire Pressure

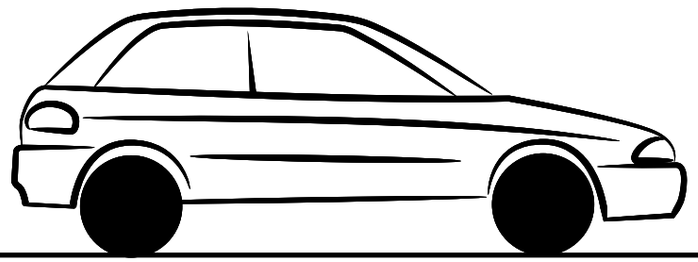


Temperature

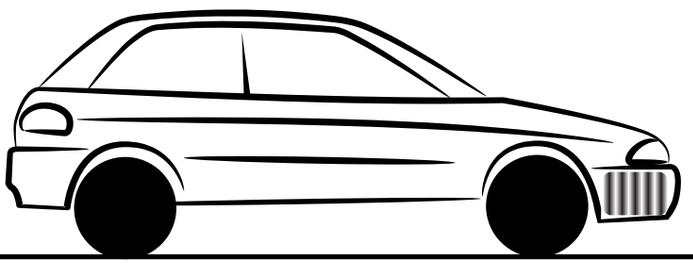


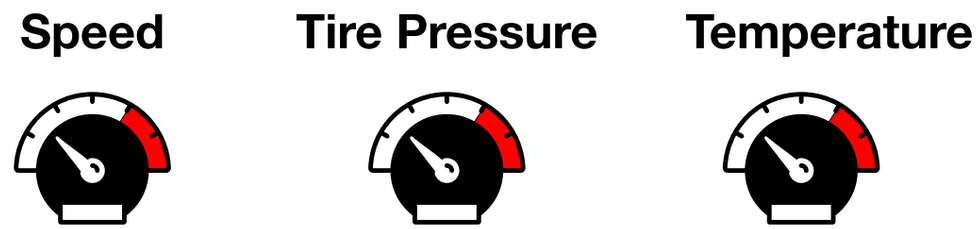
Control

What does a car’s radiator do?



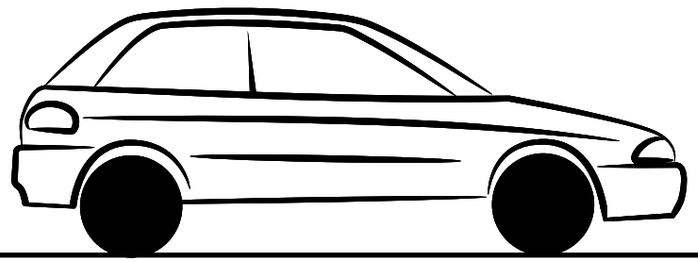
Monitor gauges and car phenotypes continually

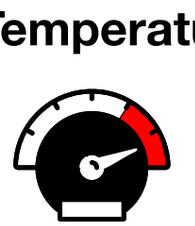




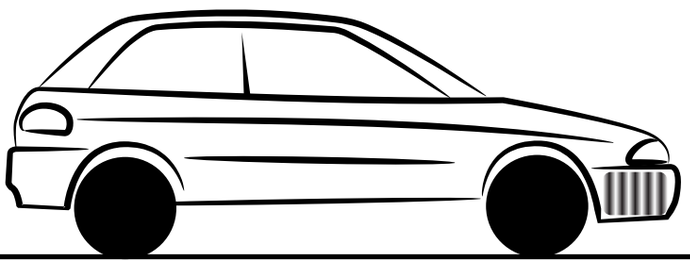
Control

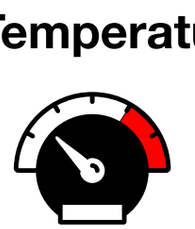
What does a car’s radiator do?





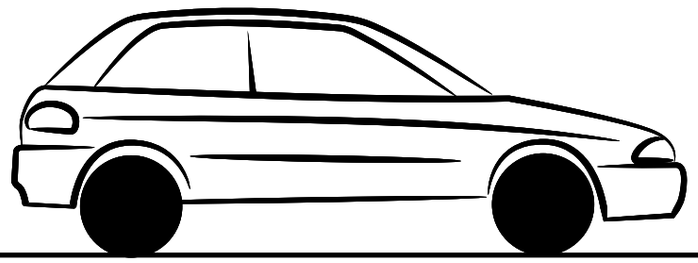
Monitor gauges and car phenotypes continually

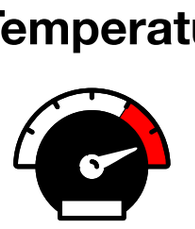




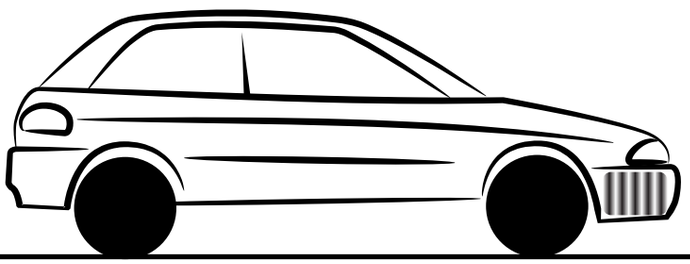
Control

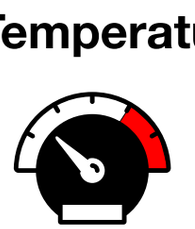
What does a car's radiator do?





Monitor gauges and car phenotypes continually

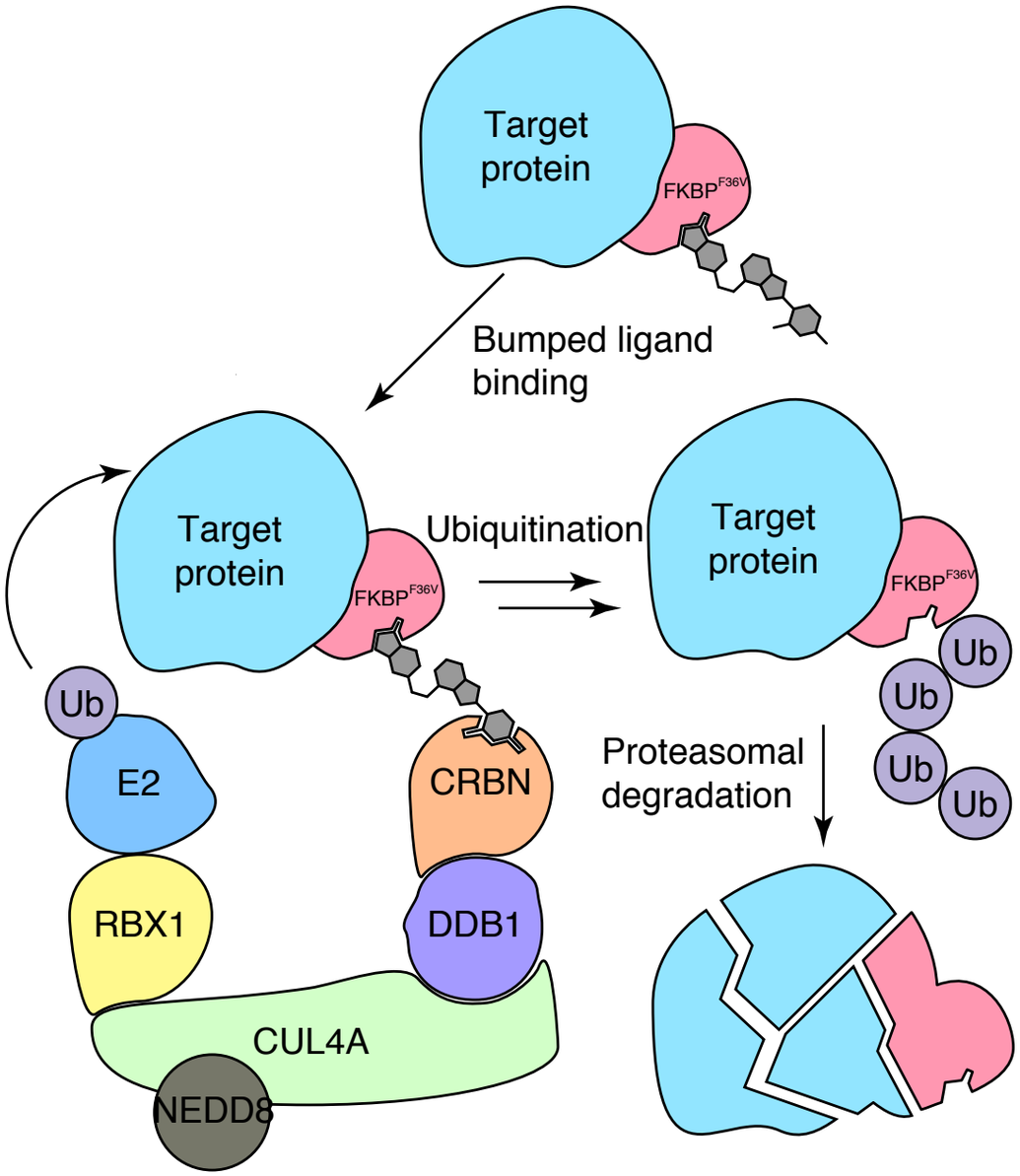




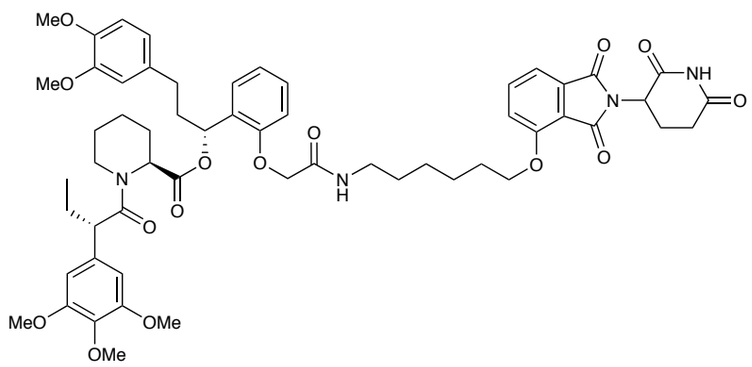
Control

The radiator directly regulates temperature;
all other effects are indirect effects of the car catching fire.

dTAG system



dTAG-13



(dTAG-V1 is another molecule that targets a distinct ubiquitination complex)

Nabet, et. al., Nat Chem Biol. 2018

GUERTIN LAB

www.guertinlab.org
@GuertinLab
guertin@uchc.edu

Dr. Sathyan Mattada (ARF-AID/parallel factor ChIP/degrons)

Dr. Arun Dutta (adipogenesis)

Jinhong Dong (TSSInference & SPDEF degron)

Rudradeep Mukherjee (compartment modeling & Twist1/2)

Tommy Scott (Estrogen response & TF degradation)

Siyu Sun (TRPS1 degron & Estrogen response)

we are recruiting students or postdocs!

UCONN SCHOOL OF MEDICINE

National Institute of General Medical Sciences

National Heart Lung and Blood Institute