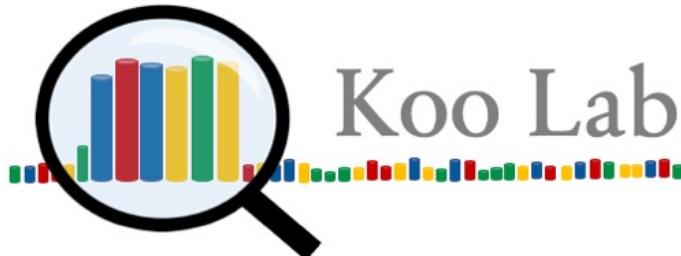


# Interpreting *cis*-regulatory mechanisms *learned by genomic deep learning*

Peter Koo

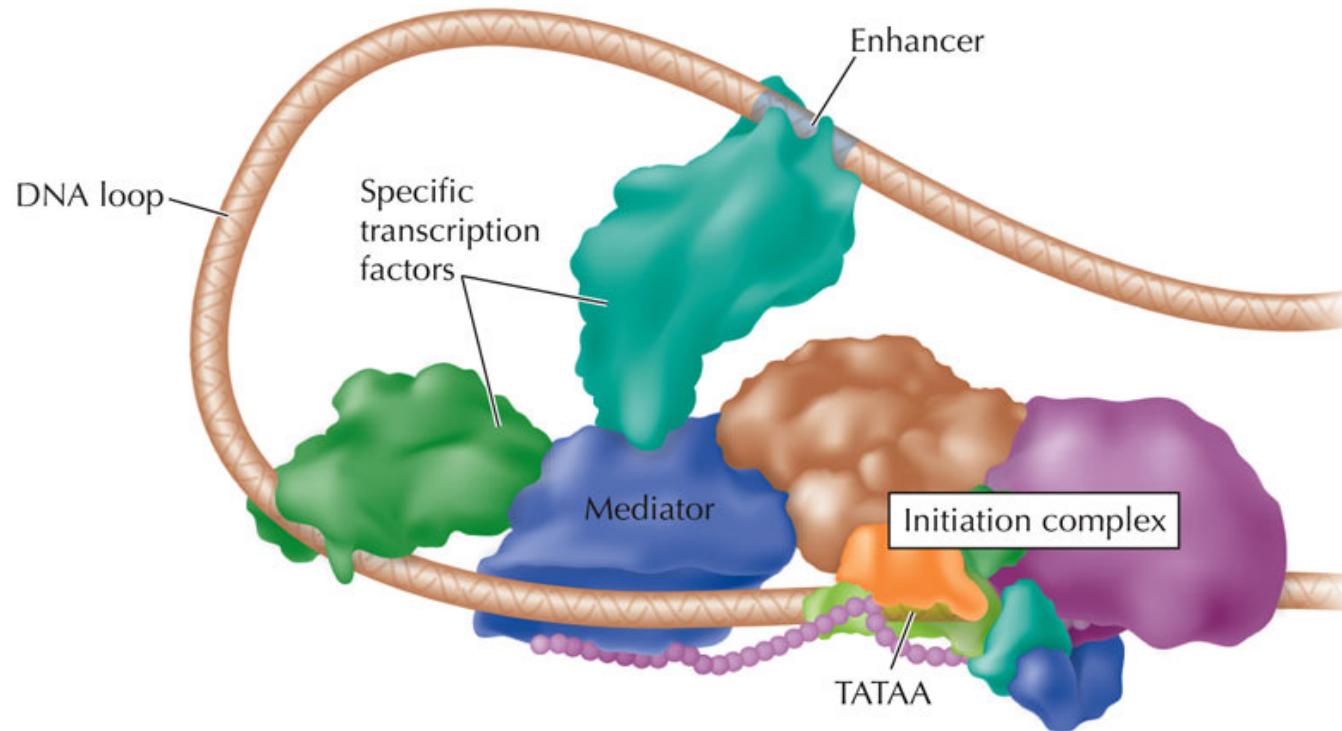
Assistant Professor

Simons Center for Quantitative Biology  
Cold Spring Harbor Laboratory



Cold  
Spring  
Harbor  
Laboratory

# Transcriptional regulation requires a complex coordination of many proteins

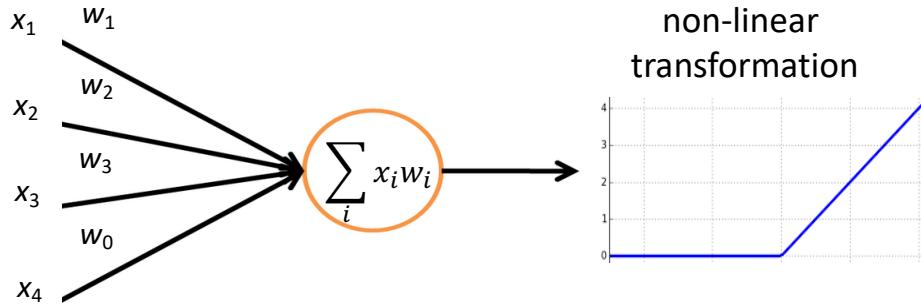


ACGTGGCTA**GCTGCC**TAAGCTGCTCGTACGTCTG**GTGCAC**CTAAGT**CCGGAA**CGTGC**TATAAC**A GCGCG CGGCC

**Uncovering cis-regulatory code is a major goal!**

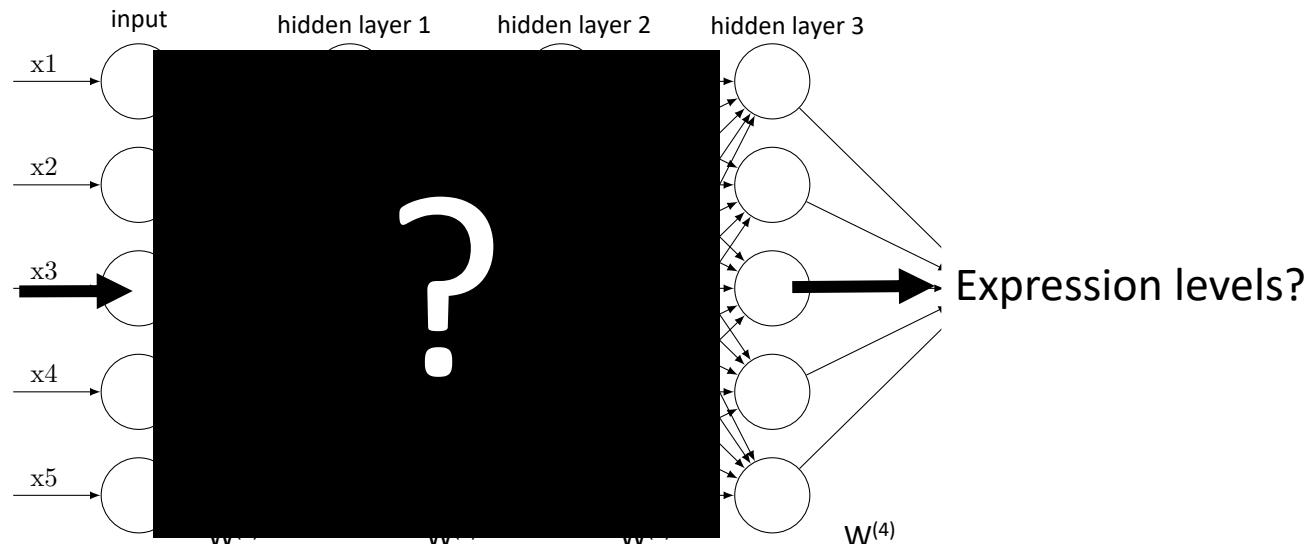
# Neural networks are black-box models that approximate functions

## artificial neuron

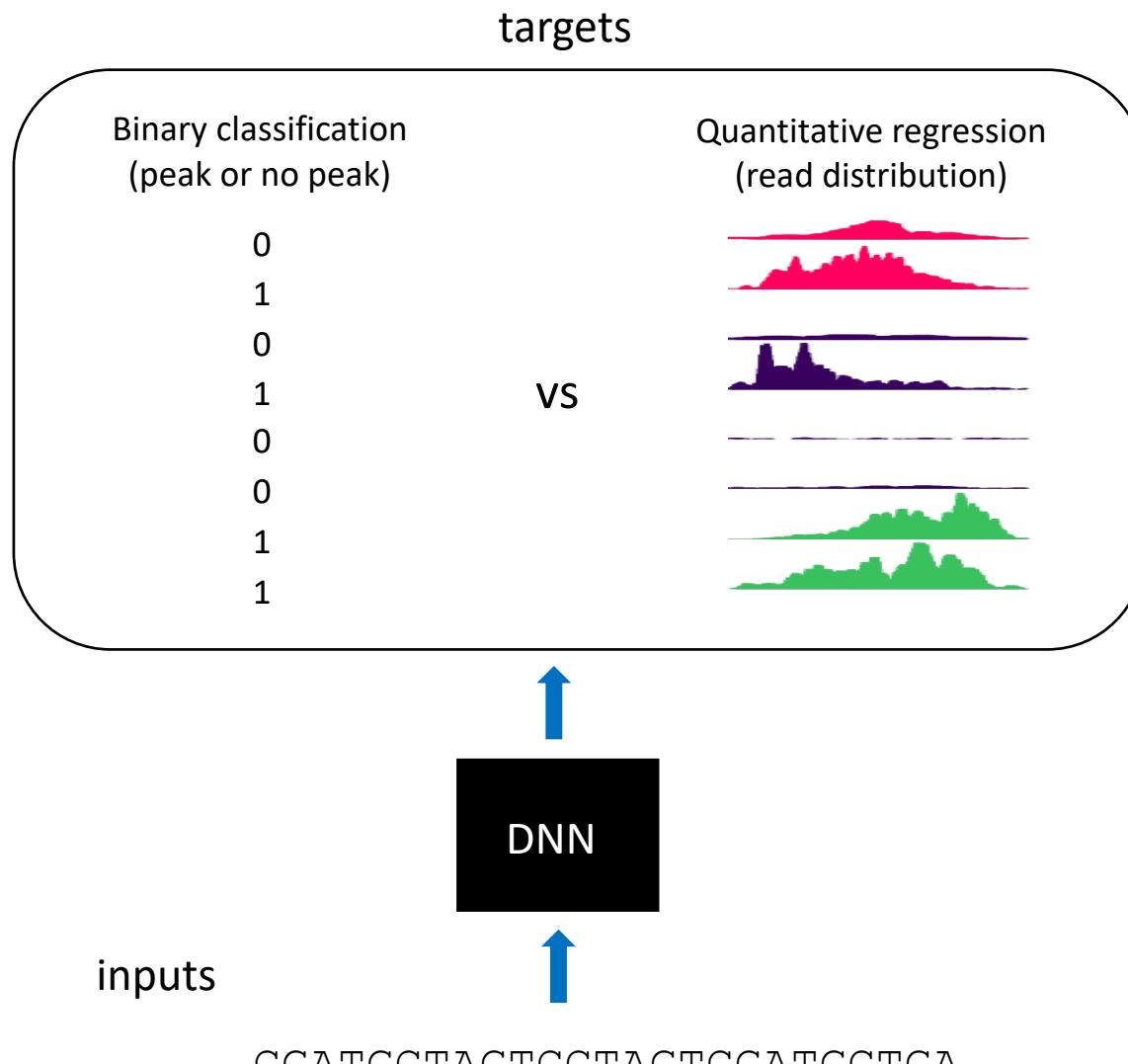


## artificial neural network

**ACTCTCAGTGTAG  
CGCGCATGCTCTA  
CTAGCCGCATAGT  
CGATCGTCGATTA  
ACTGATCGATCGG  
ACGATCGCTGACT  
AGCTGACTGTCAG**



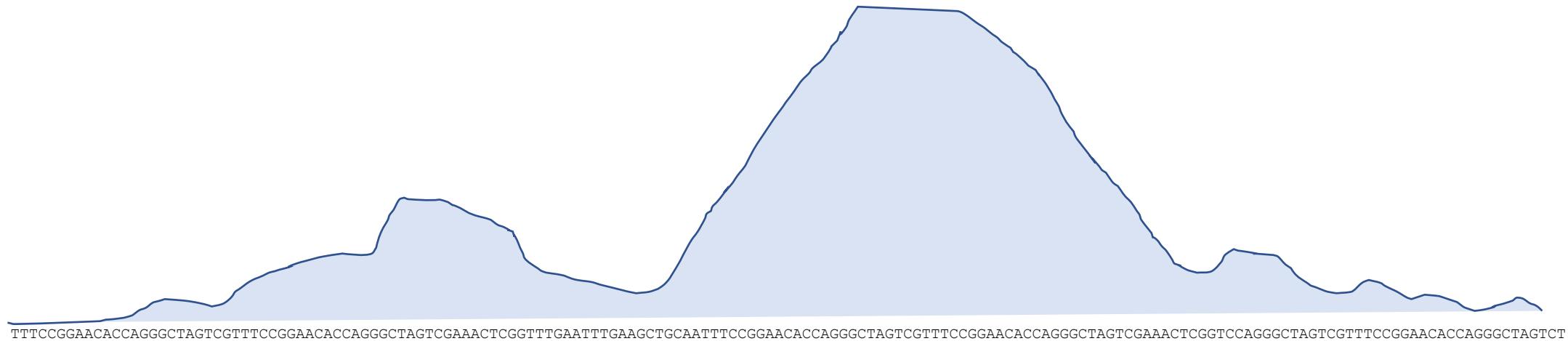
# Predicting regulatory function from sequence with deep learning



- **TF binding and histone marks:** ChIP-seq, CUT&RUN
- **Chromatin accessibility:** ATAC-seq/DNase-seq
- **Chromatin structure:** Hi-C, micro-C
- **mRNA expression:** RNA-seq, CAGE, PRO-seq,
- **Enhancer activity:** STARR-seq, MPRA

# Challenge –sequencing experiments provide low resolution view

- ChIP-seq peaks are often 150-1kb



**Computational task is to infer binding sites and their context dependencies**

# Challenge – DNA patterns are in different positions

DNA sequences (X)

ACTCGGTTGAATTGAAGCTGCAA  
TTTCCGGAAACACCAGGGCTAGTCGA  
AATTGGCGCTGCAGTAACCCATTAA  
CCCCGGAGTTAAGGCTGCCGGAAAT  
CCGATGCTGCCGGAAATGCGTTATG  
CGATGCTAGTCCTAGTCGATGCTGA

ChIP-seq peak (Y)

0  
1  
0  
1  
1  
0



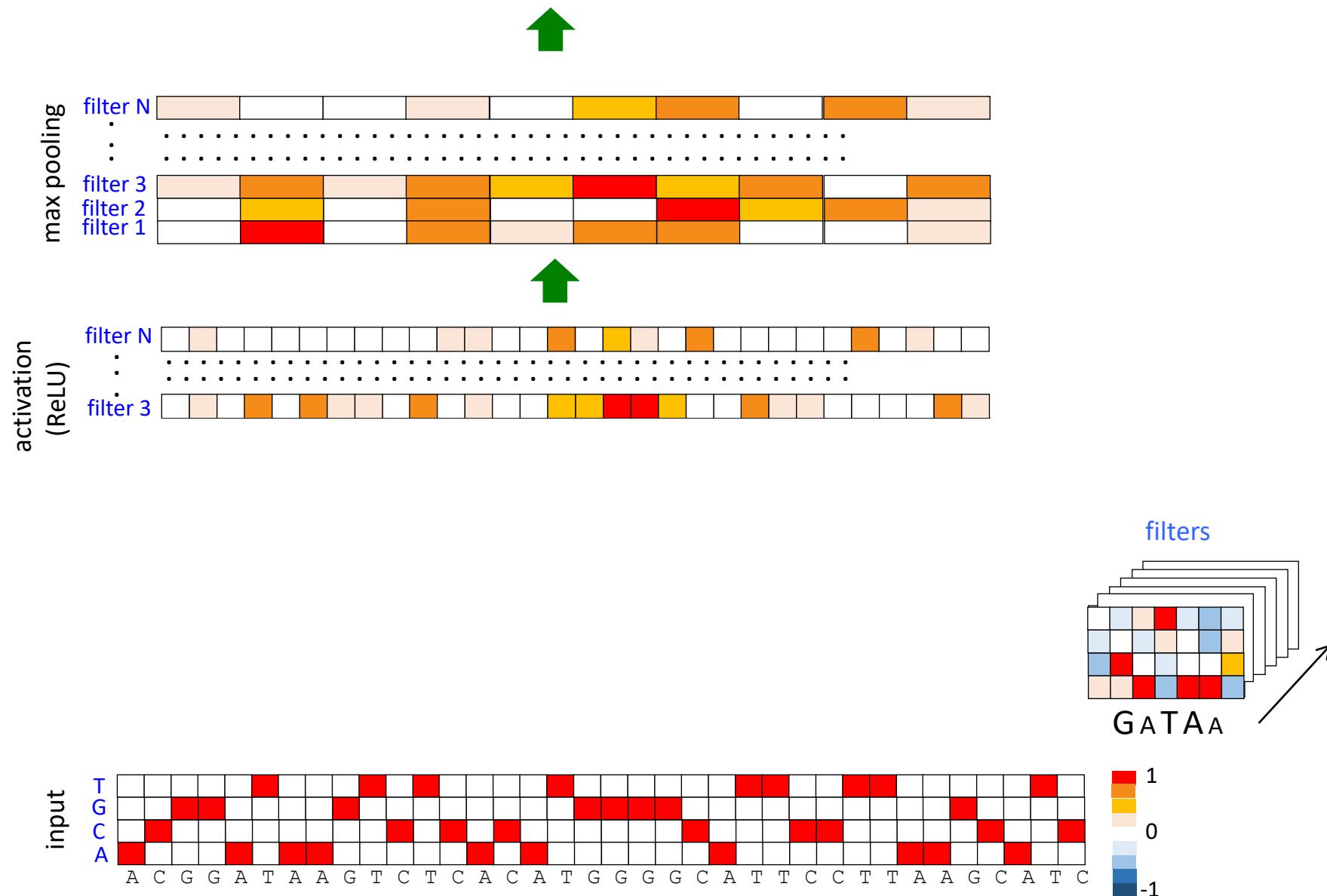
<http://waldo.wikia.com/wiki/Waldo>

## Convolution

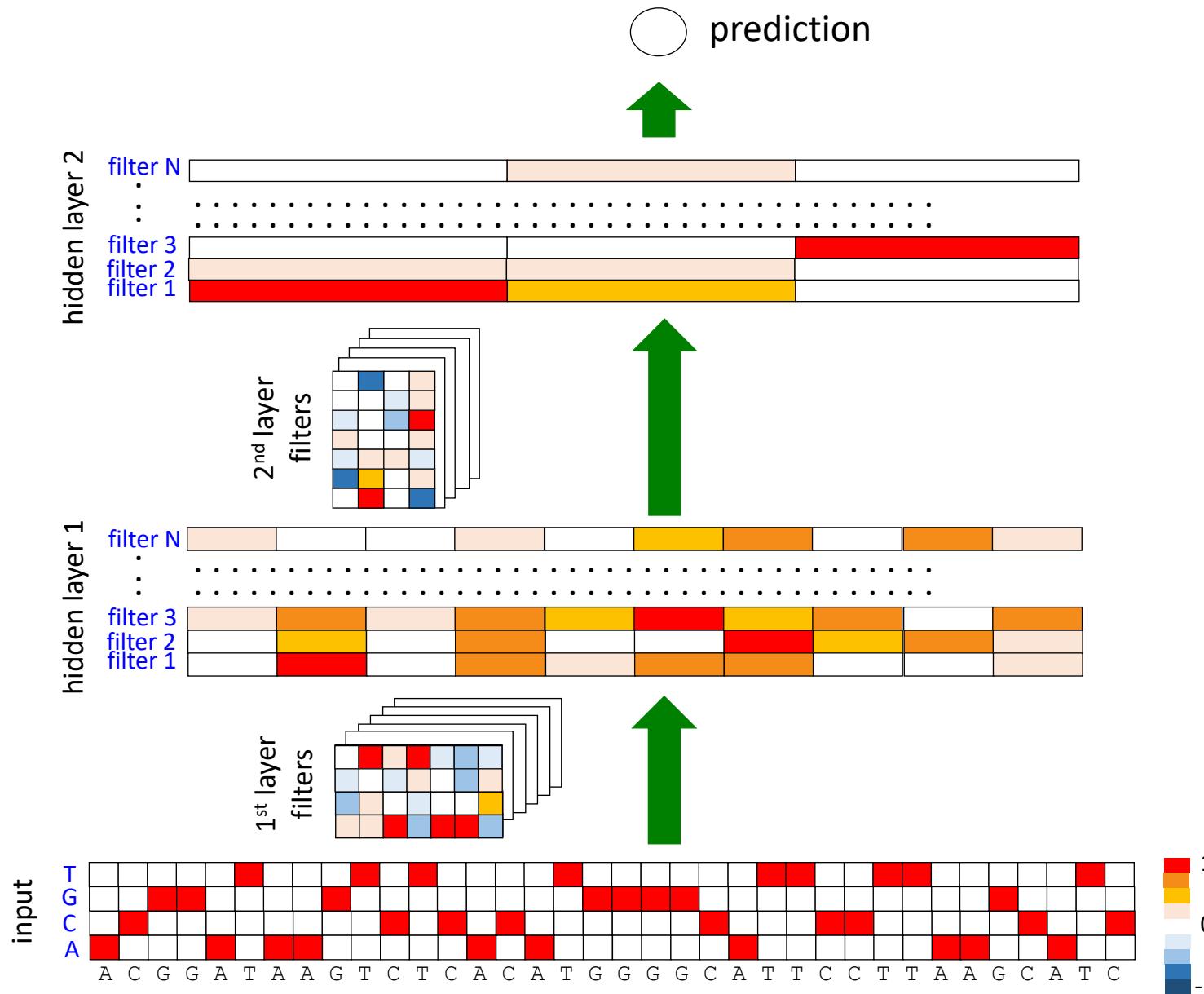
$$\begin{array}{c} \text{input (X)} \\ \begin{array}{|c|c|c|c|c|} \hline 1 & 1 & 1 & 0 & 0 \\ \hline 0 & 1 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 & 1 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 1 & 1 & 0 & 0 \\ \hline \end{array} \end{array} * \begin{array}{c} \text{filter (W)} \\ \begin{array}{|c|c|c|} \hline 1 & 0 & 1 \\ \hline 0 & 1 & 0 \\ \hline 1 & 0 & 1 \\ \hline \end{array} \end{array} = \begin{array}{|c|c|c|} \hline 4 & & \\ \hline & & \\ \hline & & \\ \hline \end{array}$$

The diagram illustrates a convolution operation. The input (X) is a 5x5 grid of binary values (0 or 1). The filter (W) is a 3x3 grid of binary values. The result of the convolution is a single value '4' in a 1x1 output (Z) grid. The asterisk (\*) indicates the convolution operation. Red 'x' marks indicate the elements of the filter that are multiplied by the corresponding input elements.

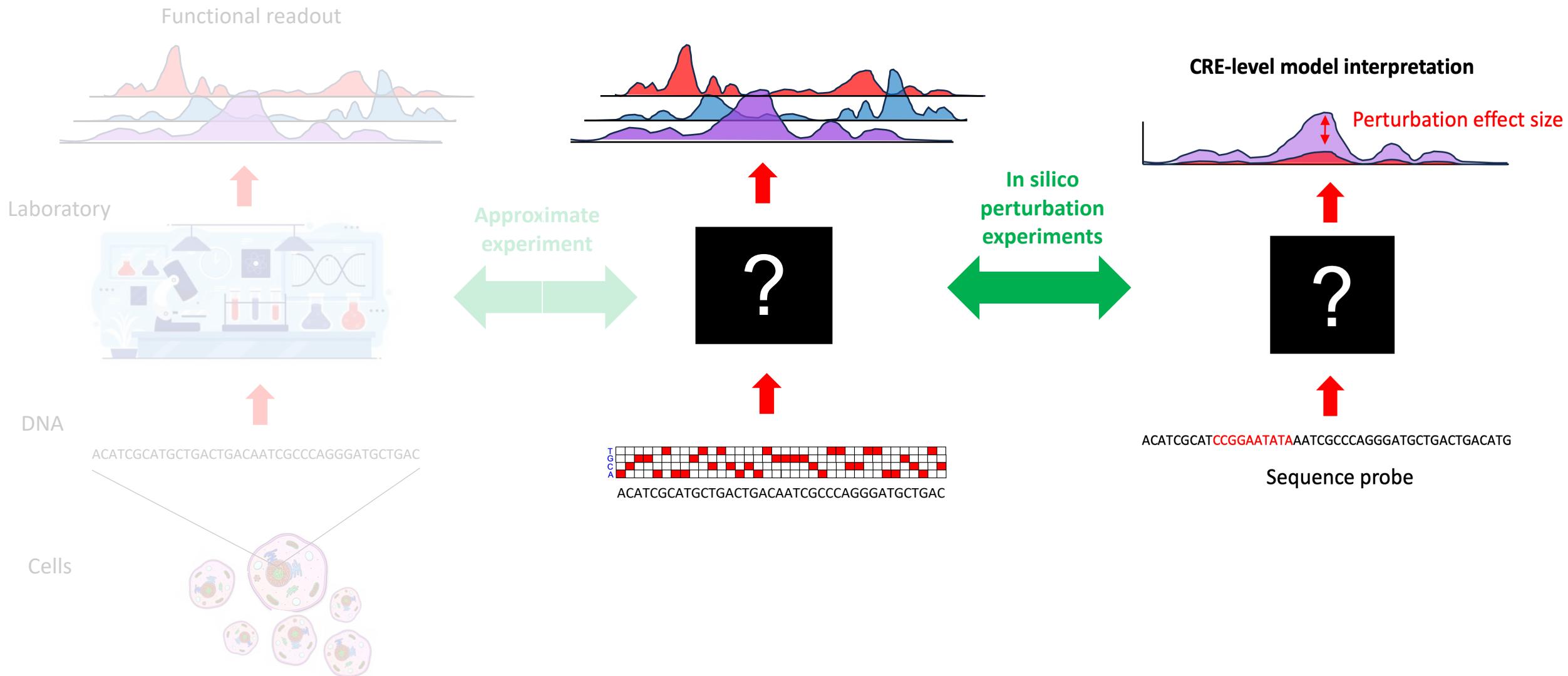
# classifier or another convolutional layer



# CNNs build representations hierarchically

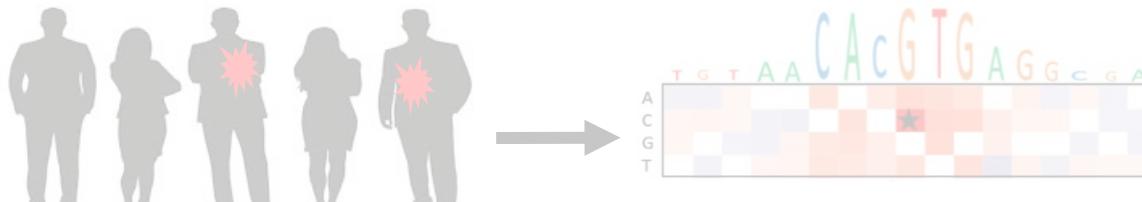


# DNN approximates function of experimental assay



# Downstream applications of sequence-function deep learning

## 1. score functional consequence of non-coding variants



**Should we trust deep learning predictions?**

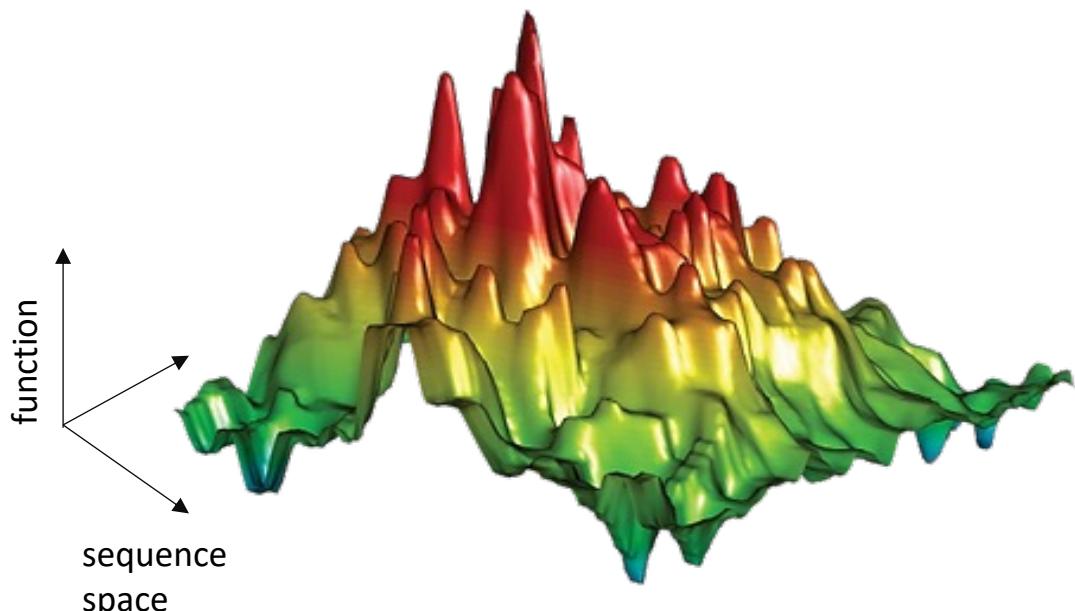
## 3. design novel sequences with desirable properties



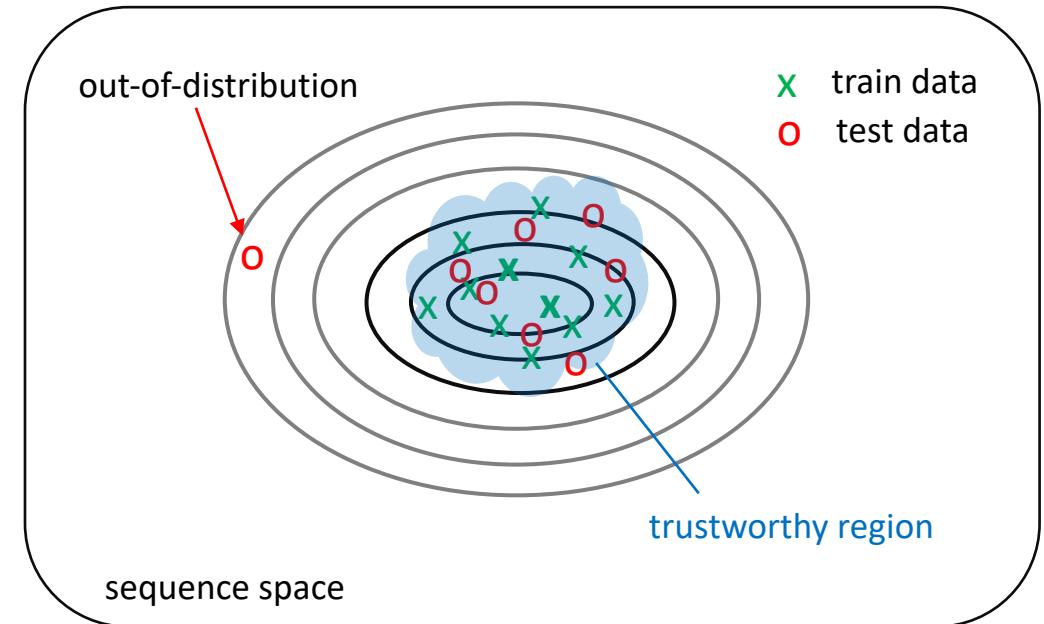
Stock photo: <https://cellculturedish.com>

# DNNs approximate sequence-function landscape of data

Toy sequence-function landscape



Example of out-of-distribution shift



**It is not clear to what extent can we trust deep learning predictions under a distribution shift?**

# Interpreting DNN helps to build trust in their predictions

## 1. Inherently interpretable design

(Alipanahi et al. 2015, Kelley et al. 2016, Koo & Eddy. 2019, Koo & Ploenzke. 2021, Ullah & Ben-Hur. 2021, Ghotra & Koo. 2022, Novakowvsky et al. 2023, + many more)

## 2. Feature attribution methods

(Alipahani et al. 2015, Shrikumar et al. 2017, Lundberg & Lee. 2017, Jha et al. 2020, Tseng et al. 2020, Majdandzic et al 2022, + many more)

## 3. Counterfactual *in silico* experiments

(Koo et al. 2018 (and 2021), Avsec et al. 2021, de Almeida et al. 2022, Gunsalus et al. 2022, Karollus et al. 2023, Kseniia et al. 2023, Toneyan & Koo. 2023)

# Interpreting DNN helps to build trust in their predictions

## 1. Inherently interpretable design

(Alipanahi et al. 2015, Kelley et al. 2016, Koo & Eddy. 2019, Koo & Ploenzke. 2021, Ullah & Ben-Hur. 2021, Ghotra & Koo. 2022, Novakowvsky et al. 2023, + many more)

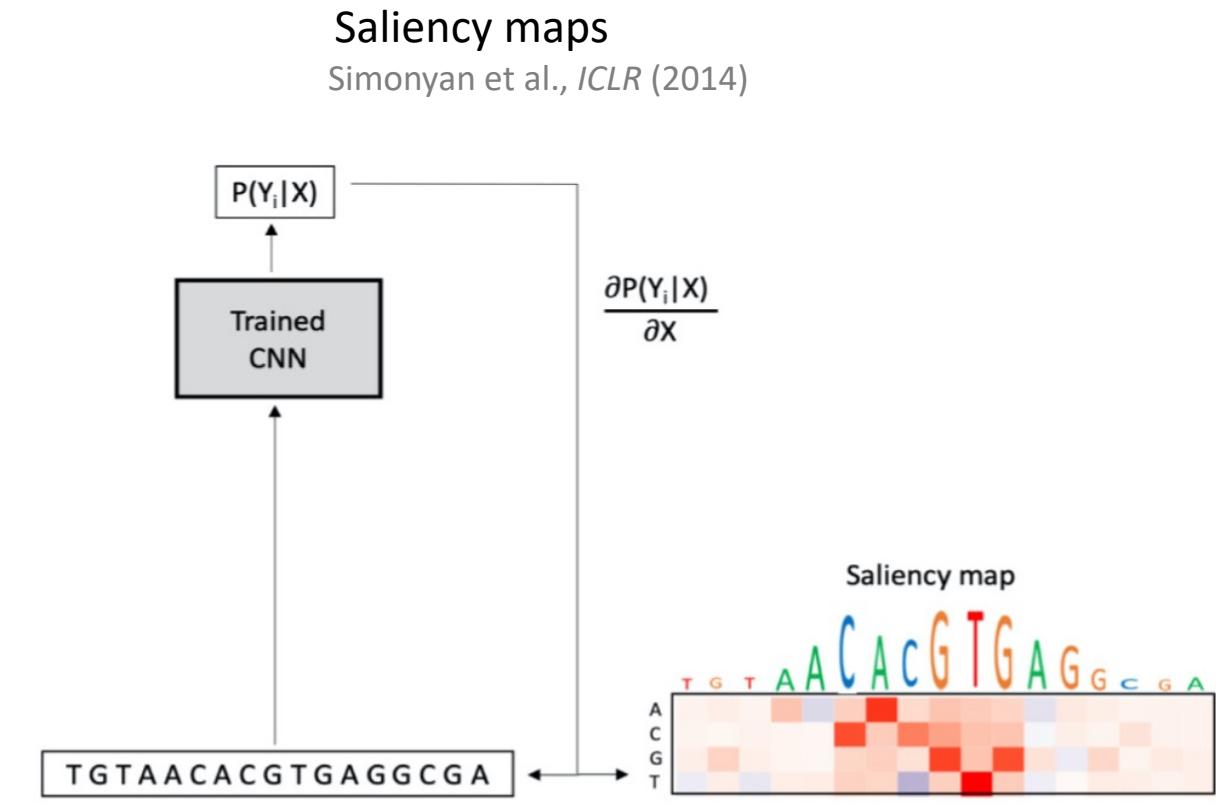
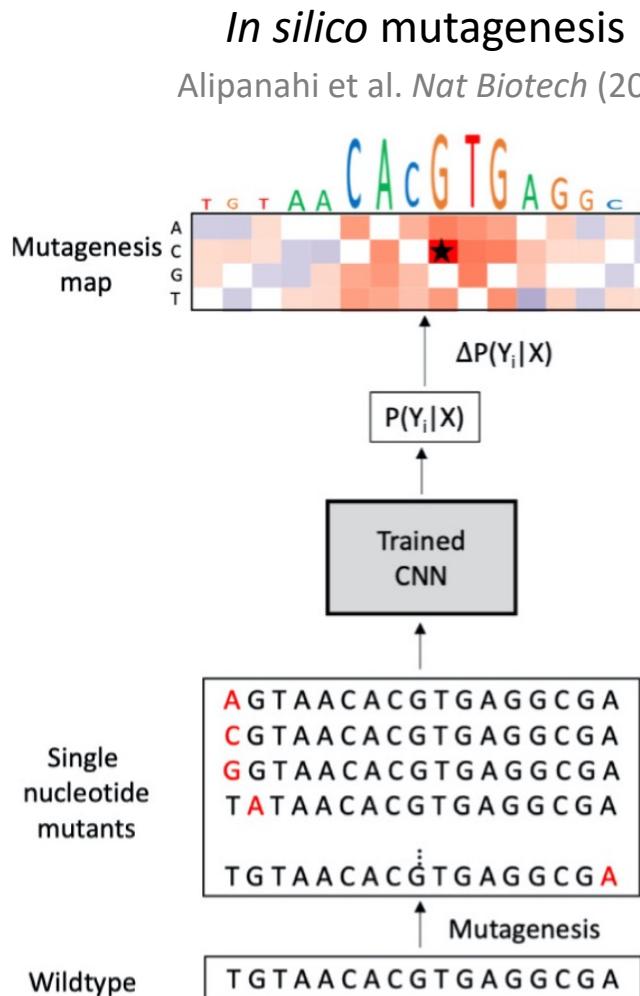
## 2. Feature attribution methods

(Alipahani et al. 2015, Shrikumar et al. 2017, Lundberg & Lee. 2017, Jha et al. 2020, Tseng et al. 2020, Majdandzic et al 2022, + many more)

## 3. Counterfactual *in silico* experiments

(Koo et al. 2018 (and 2021), Avsec et al. 2021, de Almeida et al. 2022, Gunsalus et al. 2022, Karollus et al. 2023, Kseniia et al. 2023, Toneyan & Koo. 2023)

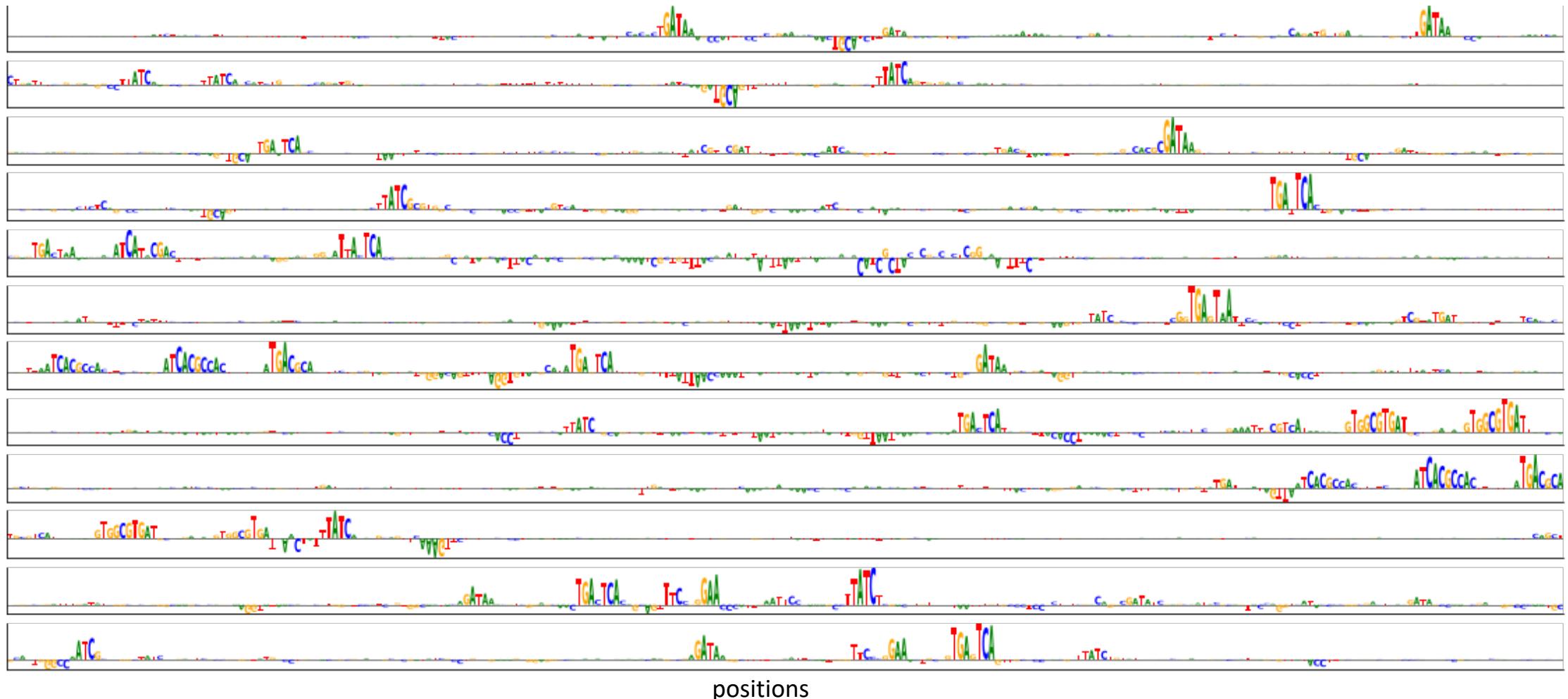
# Attribution methods: capture *sensitivity* of nucleotide variants on predictions



- Other attribution methods:**
- Lundberg et al. *NeurIPS* (2017)
  - Smilkov et al., *ICML* (2017)
  - Sundararajan et al. *ICML* (2017)
  - Shrikumar et al. *ICML* (2017)
  - Selvaraju et al. *ICCV* (2017)
  - Anupama et al. *Genome Biol* (2020)
  - + many more

# Attribution maps help to generate hypotheses of cis-mechanisms...

attribution score



**Problem: easy to over-interpret attribution maps!**

# Interpreting DNN helps to build trust in their predictions

## 1. Inherently interpretable design

(Alipanahi et al. 2015, Kelley et al. 2016, Koo & Eddy. 2019, Koo & Ploenzke. 2021, Ullah & Ben-Hur. 2021, Ghotra & Koo. 2022, Novakowvsky et al. 2023, + many more)

## 2. Feature attribution methods

(Alipahani et al. 2015, Shrikumar et al. 2017, Lundberg & Lee. 2017, Jha et al. 2020, Tseng et al. 2020, Majdandzic et al 2022, + many more)

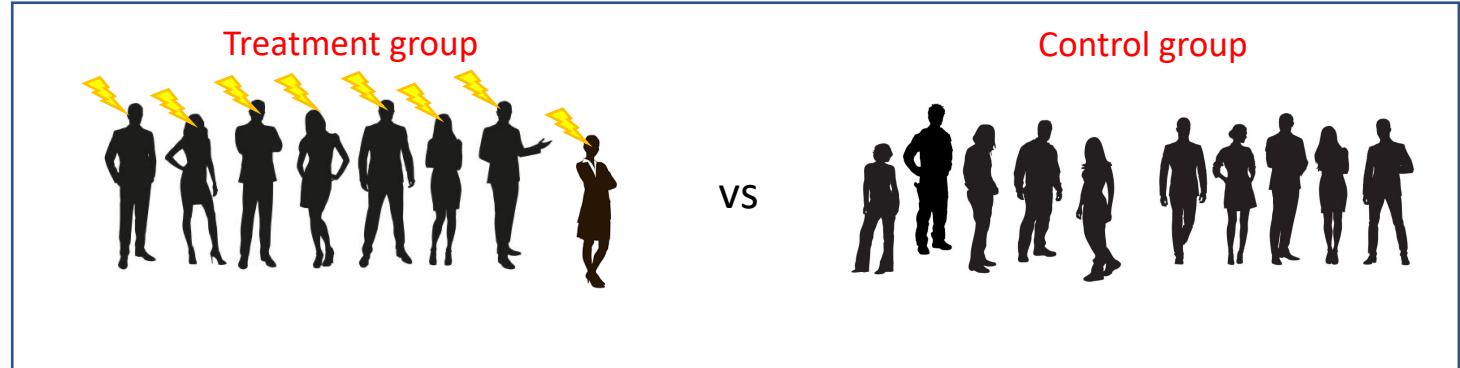
## 3. Counterfactual *in silico* experiments

(Koo et al. 2018 (and 2021), Avsec et al. 2021, de Almeida et al. 2022, Gunsalus et al. 2022, Karollus et al. 2023, Kseniia et al. 2023, Toneyan & Koo. 2023)

# *In silico* experiments provide counterfactual explanations

- Randomized controlled trials (RCTs) are gold standard for measuring causal effect size

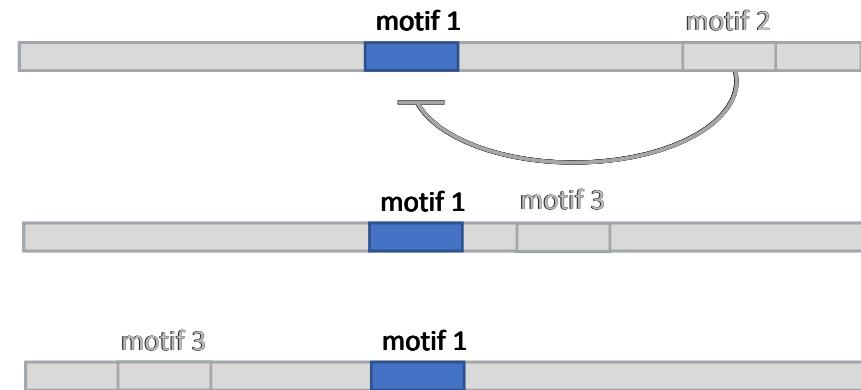
1. Random assignment to group
2. Treatment group is given intervention
3. Effect is measured across each population



- Testing hypotheses of feature importance with Global Importance Analysis (GIA)

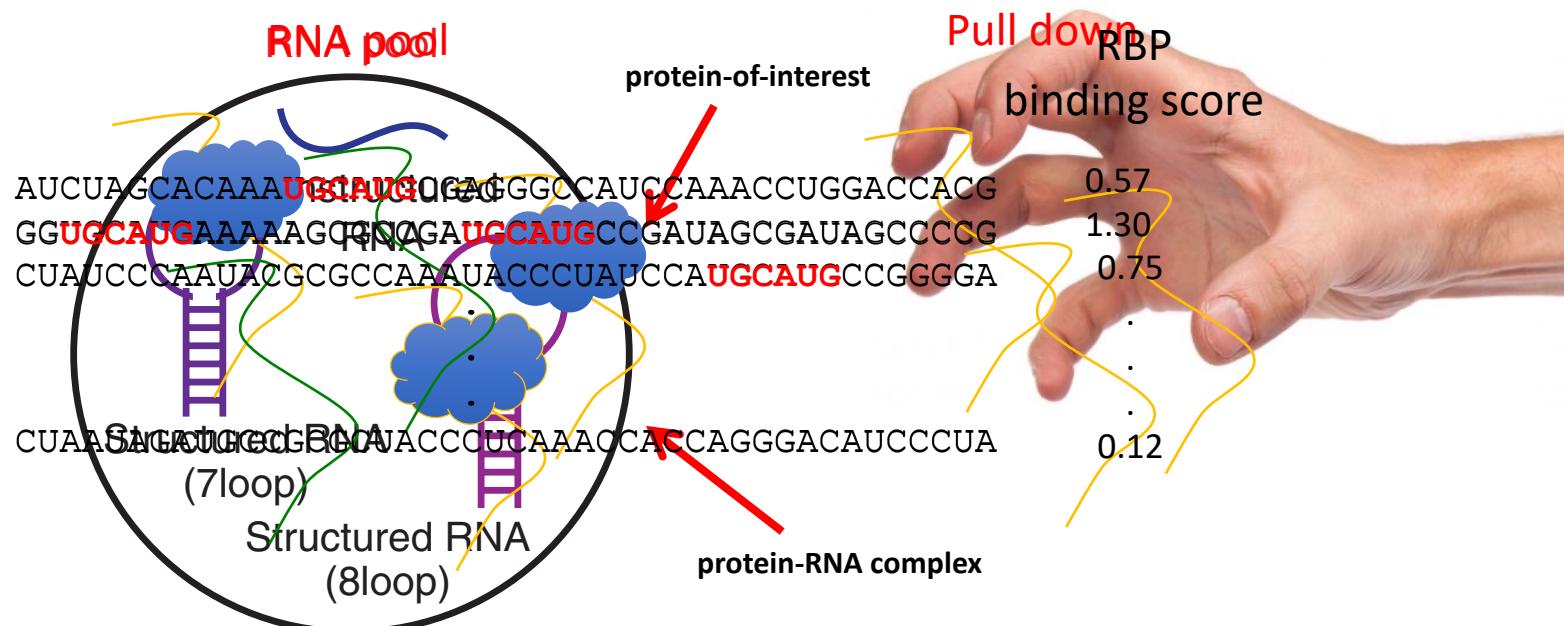
$$\begin{aligned} \mathcal{I}^{\text{global}} &= \mathbb{E}_{\mathbf{x}^{\phi_i} \sim \mathcal{D}} [\mathbf{y} | \mathbf{x}] - \mathbb{E}_{\mathbf{x} \sim \mathcal{D}} [\mathbf{y} | \mathbf{x}] \\ &\approx \frac{1}{N} \sum_n^N (f(\mathbf{x}_n^{\phi_i}) - f(\mathbf{x}_n)) p(x_n) \end{aligned}$$

intervention  
NN predictions



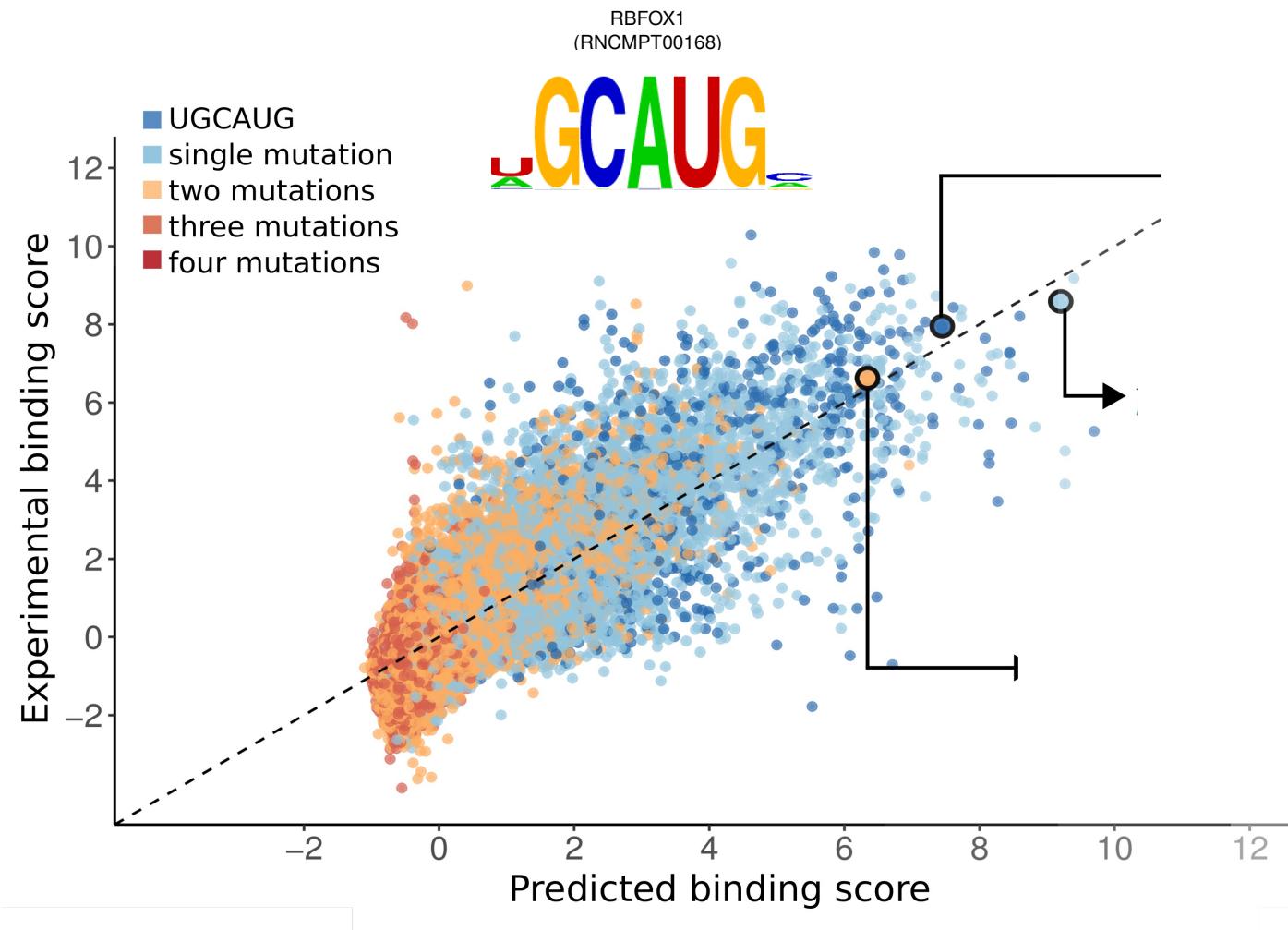
# Case study: RNAcompete *in vitro* affinity-selection for RBPs

1. create diverse pool of short RNA sequences (~240,000 sequences)
2. let protein-of-interest interact with RNA pool
3. pull down protein-of-interest
4. quantify binding score



RNAcompete dataset: 244 experiments with over 200 unique RBPs

# *In silico* mutagenesis footprints motif features in individual sequences



# Uncovering **global importance** with *in silico* experiments

## 1. Generate synthetic sequences ( $\mathbf{x}$ )

NNNNNNNNNNNN**UGCAUG**NNNNNNNNNNNN

NNNUGCAUGNNNUGCAUGNNNNNNNNNN

NNNUGCAUGNNNUGCAUGNNNUGCAUGNNN

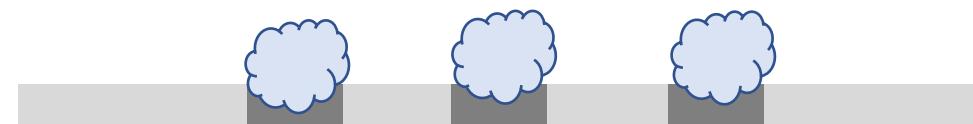
## 2. Get network prediction ( $y$ )

### 3. Calculate importance

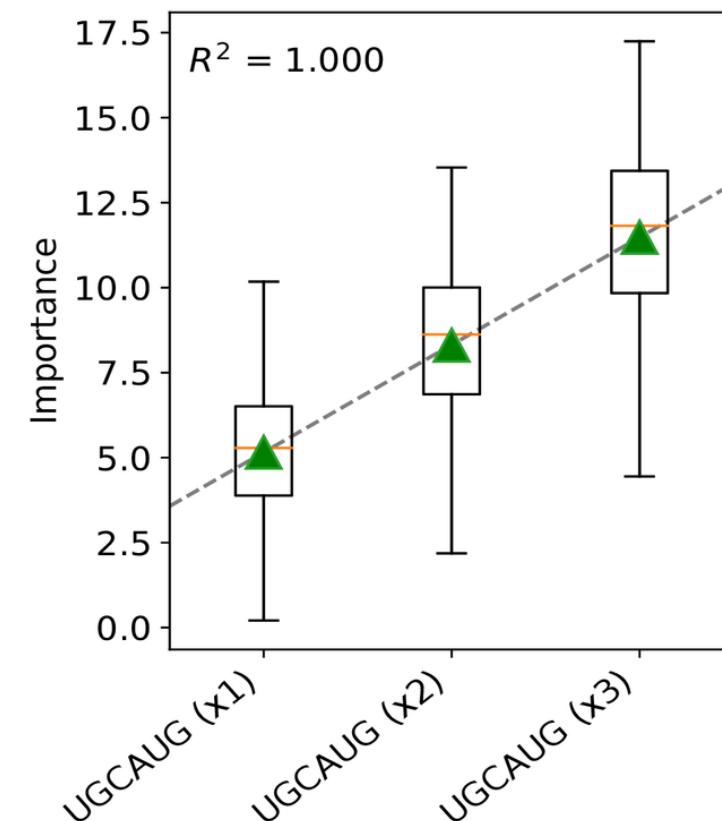
$$\begin{aligned} \mathcal{I}^{\text{global}} &= \mathbb{E}_{\mathbf{x}^{\phi_i} \sim \mathcal{D}} [\mathbf{y} | \mathbf{x}] - \mathbb{E}_{\mathbf{x} \sim \mathcal{D}} [\mathbf{y} | \mathbf{x}] \\ &\approx \frac{1}{N} \sum_n^N (f(\mathbf{x}_n^{\phi_i}) - f(\mathbf{x}_n)) p(x_n) \end{aligned}$$

intervention

NN predictions



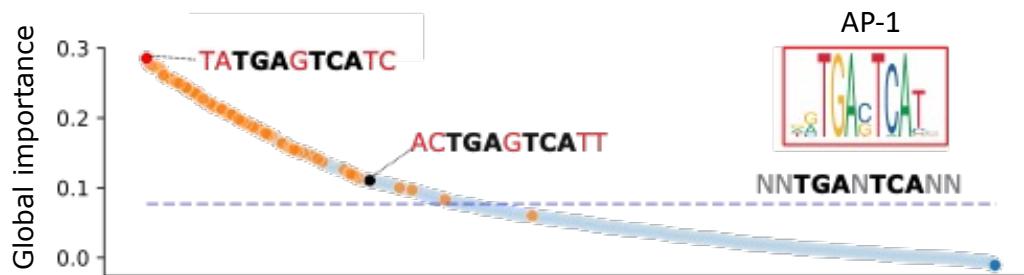
RNAcompete – varying RBFOX1 motif



# GIA can provide deeper insights into TF binding

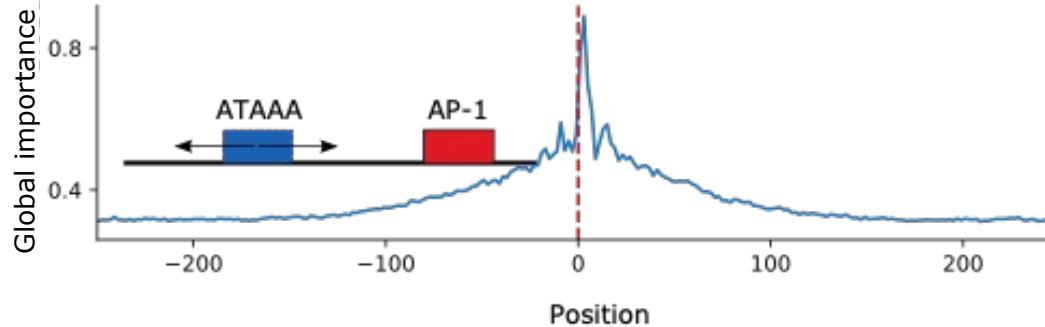
## Exploring the importance of motif flanks

de Almeida et al. *Nat Genetics* 2022  
Le et al. *PNAS* 2018



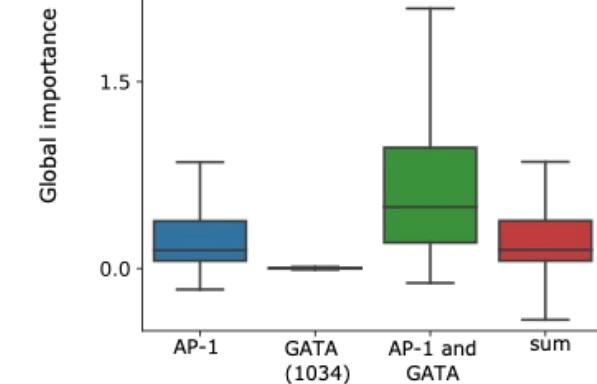
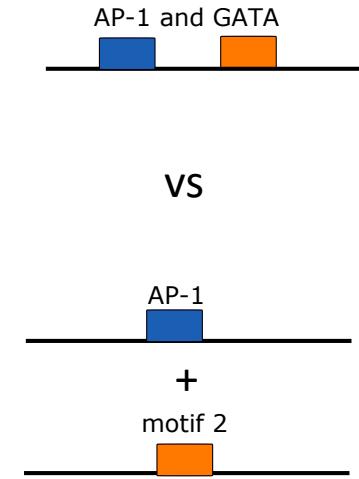
## Exploring motif distance dependence

Avsec et al. *Nat Genetics* 2021



## Exploring motif cooperativity

Avsec et al. *Nat Genetics* 2021  
de Almeida et al. *Nat Genetics* 2022  
Kim et al. *Nat Genetics* 2021



Shush Toneyan  
CSHL grad student



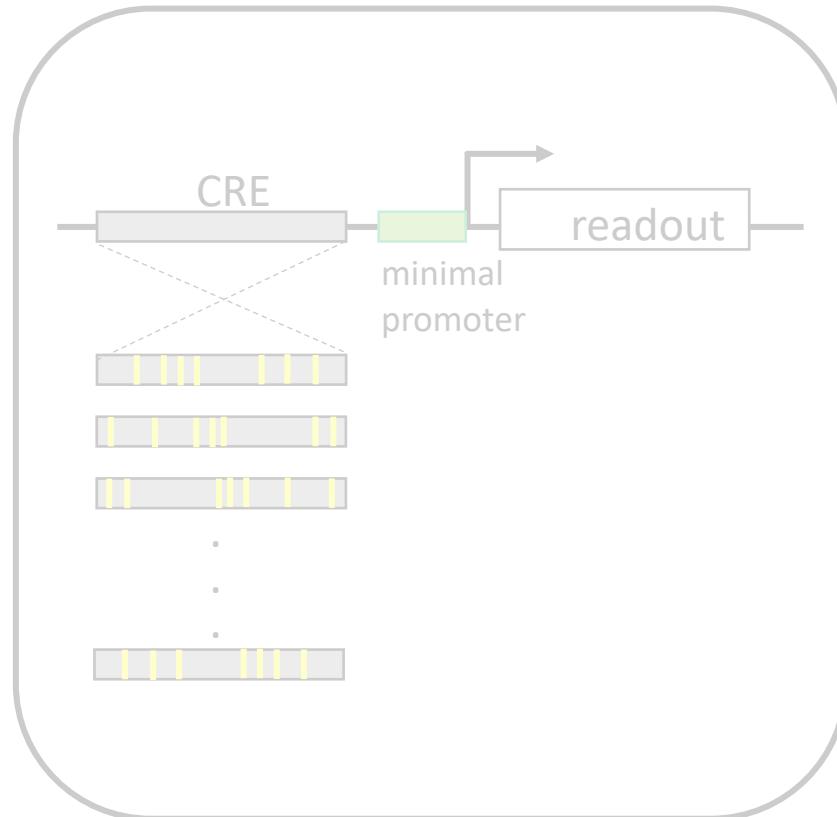
Amber Tang  
CSHL grad student

# Leveraging multiplexed perturbations to characterize functional elements

**MAVE** – probe causal relationships  
within cis-regulatory elements (CREs)



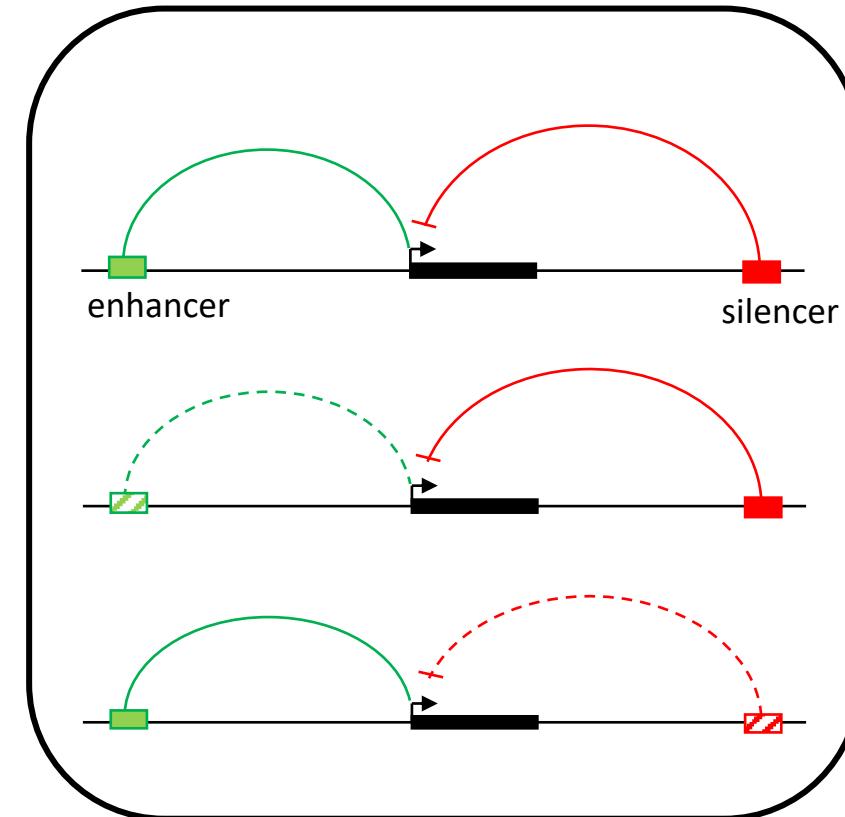
Evan Seitz  
Postdoc



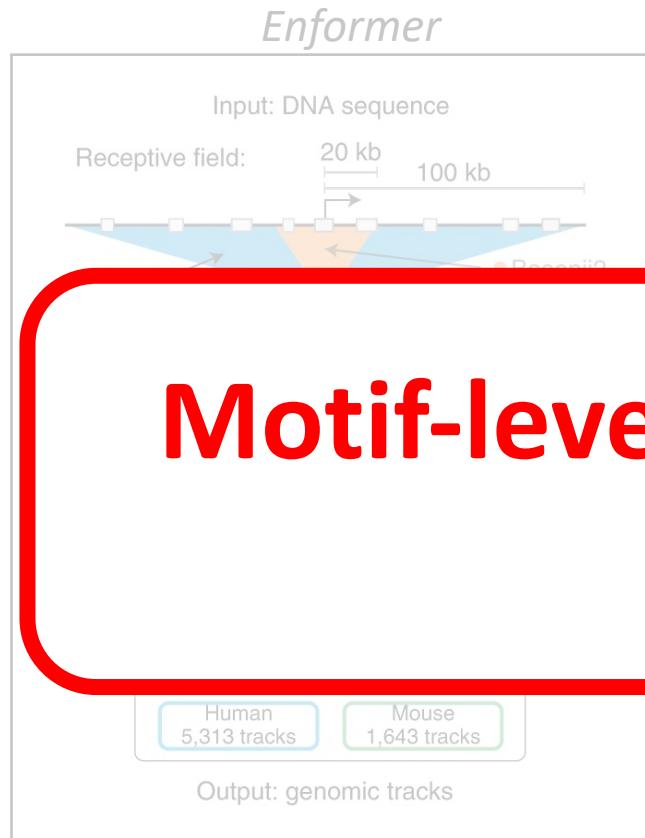
**CRISPR** – characterize role of CREs  
on gene expression



Shush Toneyan  
Grad student



# Rise of large-scale deep learning models in genomics



Avsec et al. 2021

ARTICLES  
<https://doi.org/10.1038/s41592-021-01252-x>

nature methods

Check for updates

OPEN  
Effective gene expression prediction from sequence by integrating long-range interactions

Žiga Avsec<sup>1</sup>✉, Vikram Agarwal<sup>2,4</sup>, Daniel Visentin<sup>1,4</sup>, Joseph R. Ledsam<sup>1,3</sup>,  
Agnieszka Grabska-Barwinska<sup>1</sup>, Kyle R. Taylor<sup>1</sup>, Yannis Assael<sup>1</sup>, John Jumper<sup>1</sup>, Pushmeet Kohli<sup>1</sup>✉

Motif-level explanations are too complex  
for long sequences !



Enformer  
Ref-alt  
Prediction



# Cis-Regulatory Element Model Explanations (CRÈME)

*in silico* perturbation toolkit to interpret large-scale DNNs



Shushan Toneyan  
CSHL grad student

The diagram illustrates the CRÈME pipeline. It starts with 'CAGE-seq' data on the left, which feeds into a 'CRÈME: in silico perturbation assay' block on the right. A red box highlights a 'Case study: interpret Enformer focusing on TSS activity in K562 (using Cage-seq)' section. Below this box, a list of features includes 'Motif Content test' and '+ more'.

CAGE-seq

CRÈME: *in silico* perturbation assay

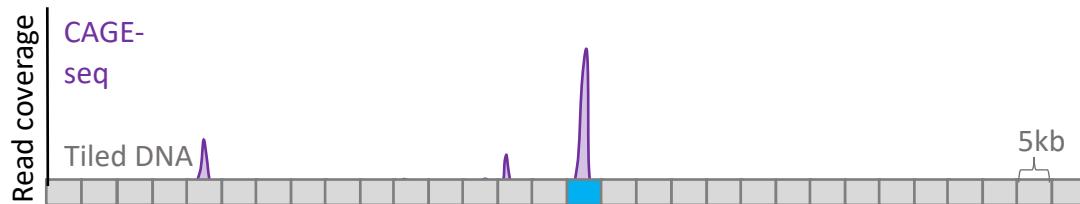
Case study: interpret Enformer focusing on TSS activity in K562 (using Cage-seq)

- Motif Content test
- + more

# What is dependence of TSS activity on distal sequence context?

## TSS Context Dependence Test

### 1. Tile sequence into 5kb bins

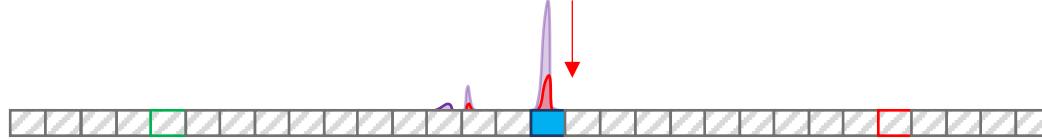


### Annotation

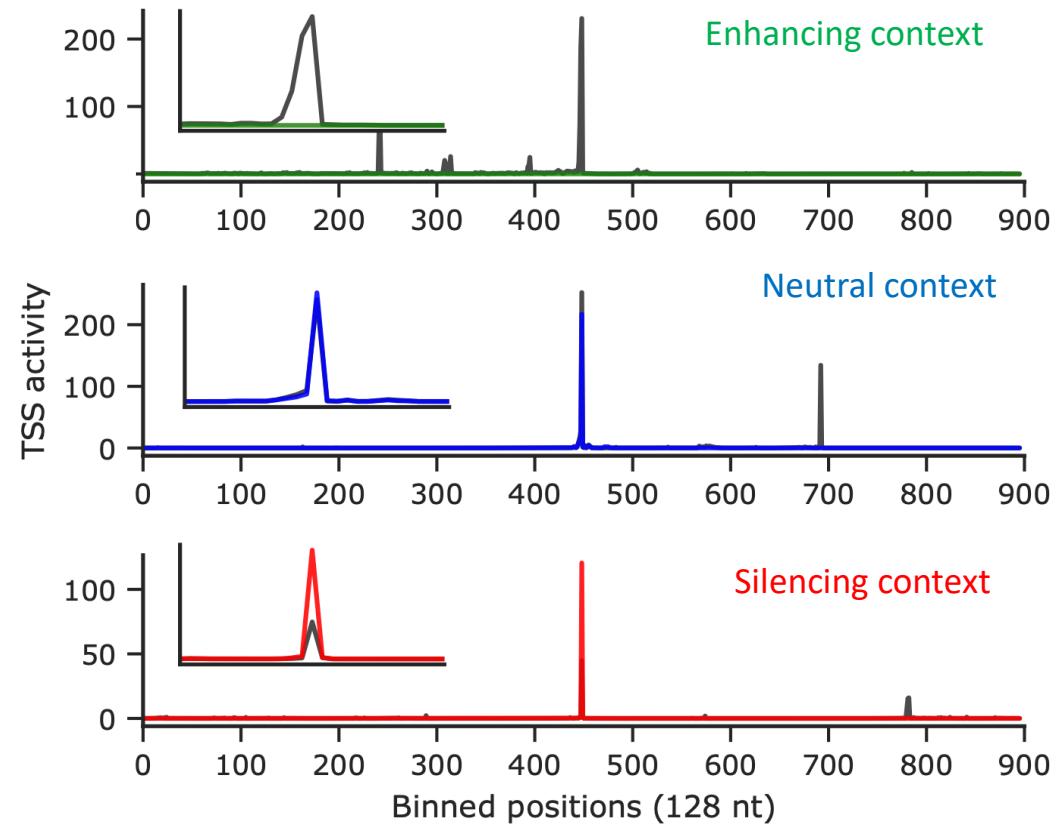
### 2. Dinuc shuffle whole sequence



### 3. measure effect of shuffling context

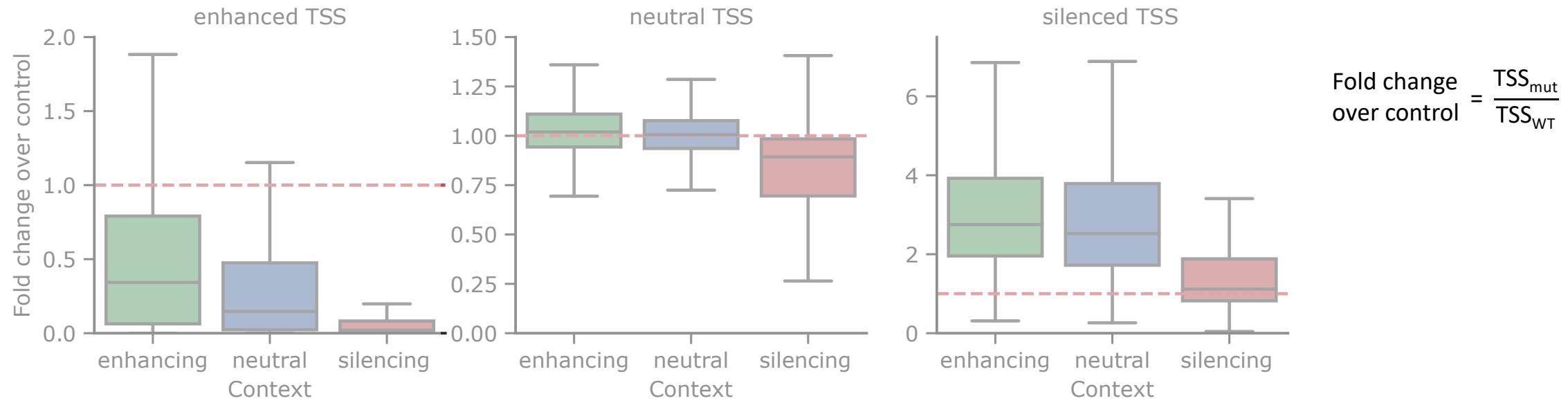


## Examples



# How compatible are TSSs with different sequence contexts?

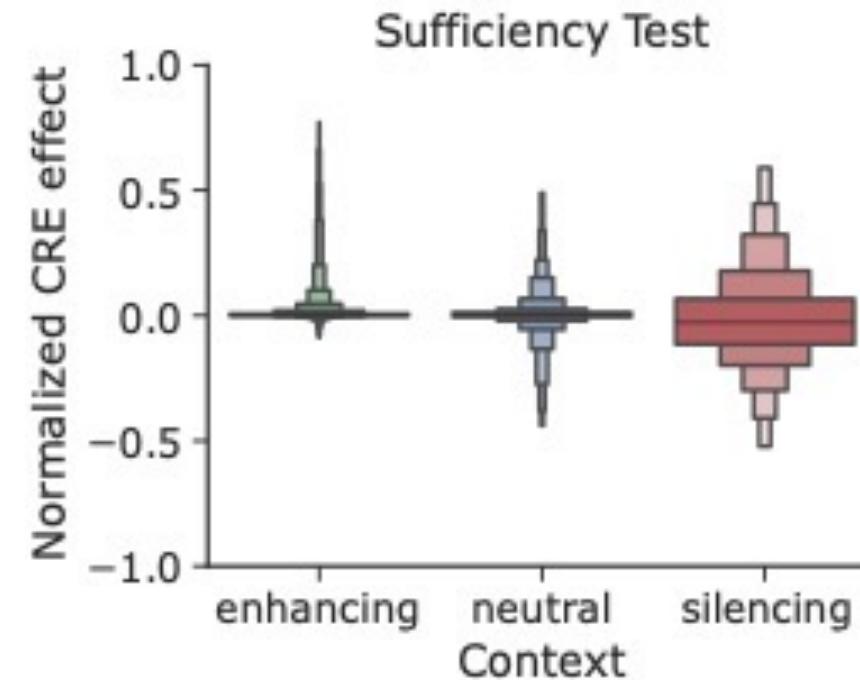
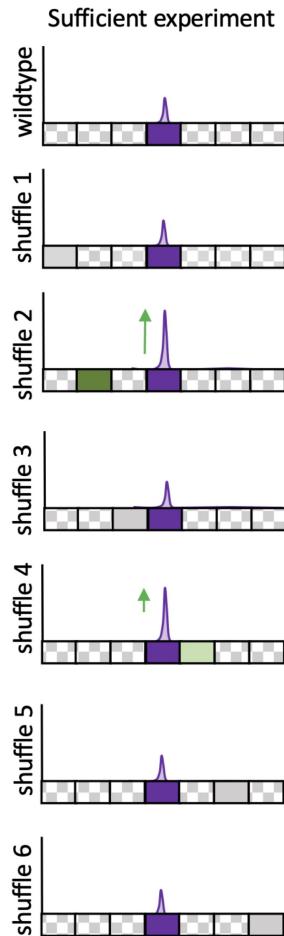
**TSS Context Swap Test:** Swap TSS tile into different genomic sequences and measure how the TSS activity changes



- **Enhanced TSS:** enhancing context is weakly enhancer to other TSSs but is more tuned to its matched TSS (compatibility?)
- **Neutral TSS:** neutral TSS is strong in many contexts (i.e. housekeeping genes over 50%)
- **(weakly) Silenced TSS:** silencing context is generic (actively silences any TSS)

# Is a single CRE sufficient to drive TSS activity?

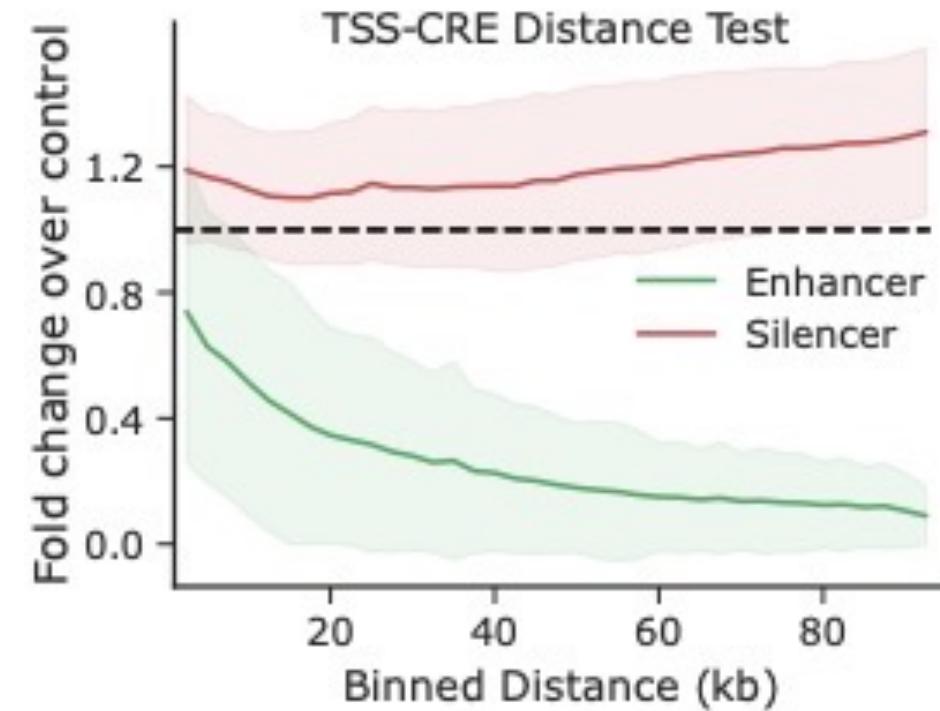
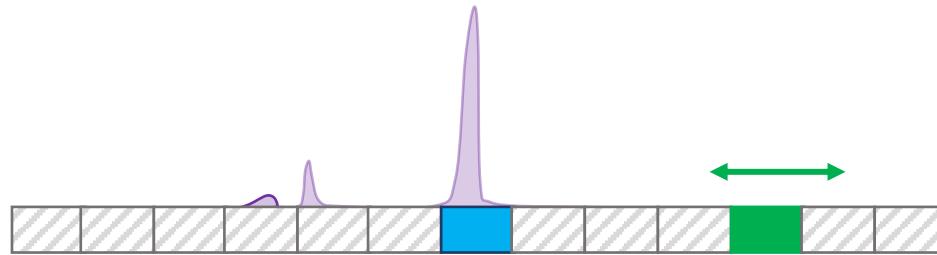
**CRE Sufficiency Test:** probe TSS-CRE pairs in shuffled sequences and measured effect on TSS activity



- Enhancers maintain positive effect on their target TSS
- Weak effect size of individual CREs suggests TSS often relies on multiple CREs

# Do CREs have a distance-dependent effect on TSS activity?

**TSS-CRE Distance Dependence Test:** systematically varied distance between TSS-CRE pairs in shuffled sequences

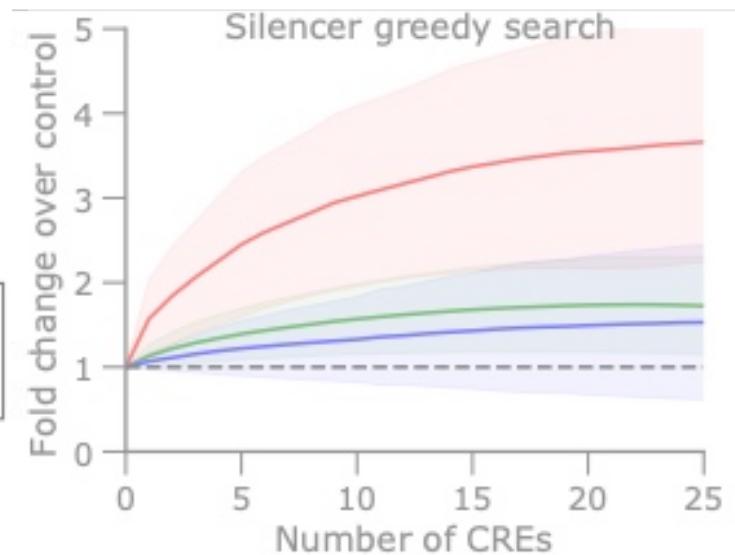
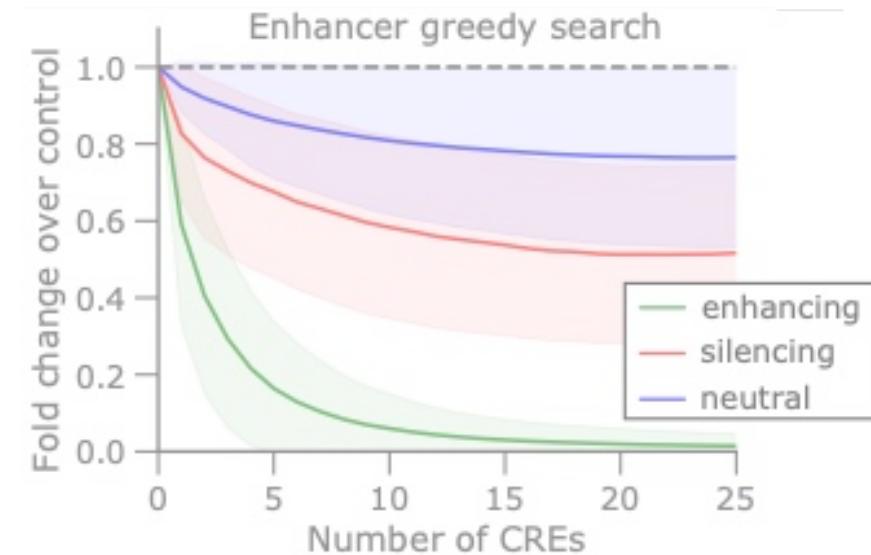
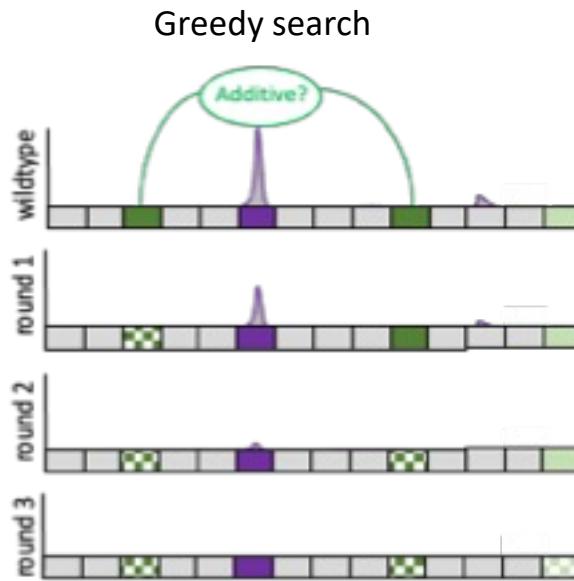


- **Enhancers:** effect size decays with distance to TSS
- **Silencers:** effect size is largely independent of distance

$$\text{Fold change over control} = \frac{\text{TSS}_{\text{mut}}}{\text{TSS}_{\text{WT}}}$$

# What set of CREs is necessary for maximally enhancing/silencing TSS activity?

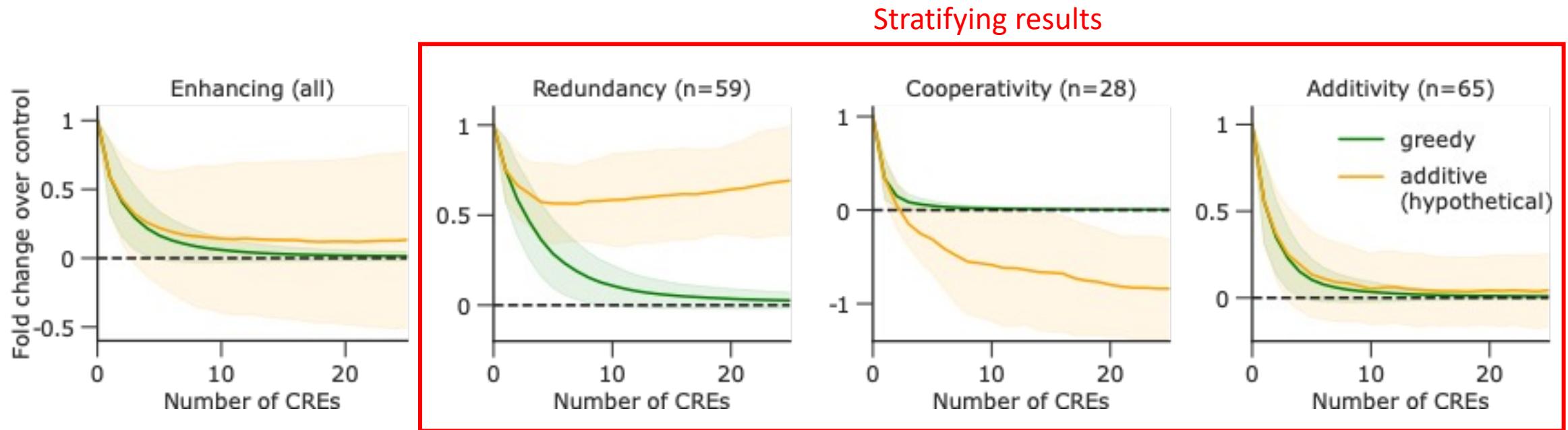
**Higher-order CRE Interaction Test:** CRE-level evolution – systematically shuffle tile with largest effect size on TSS



- Enhancers: removing 4-5 CREs reduces TSS activity by greater than 80%
- Silencer: weak effect size and work in larger numbers
- Net effect of enhancers and silencers drives observed TSS activity

# How do CREs interact to regulate TSS activity?

**Higher-order CRE Interaction Test** : Look into enhancing contexts and stratify results – compare with (hypothetical) additive effects



- Enhancers behave additively, on average
- Sometimes enhancers show high redundancy
- Sometimes enhancers exhibit cooperativity, albeit a smaller percentage

# Conclusion

- DNNs are powerful tools to study genomic features predictive of sequencing experiments
- Although they are black boxes, they can be probed with systematic perturbation experiments to better understand their biological world-view
- Are we there yet? No!
- Beware: many papers are published – lots of hype!
  - shortage of qualified reviewers means many low quality papers in high profile journals
- Suggestion: don't believe deep learning claims unless strong evidence is provided.

# Koo lab at CSHL!

## Postdocs

- Evan Seitz
- Alessandro Crnjar
- Anirban Sarkar
- Jessica Zhou

## Grad Students

- Shushan Toneyan
- Amber Tang
- Jakub Kaczmarek
- Chandana Rajesh
- Kaeli Rizzo

## Postbac

- Pretty Garcia (*PREP Scholar*)

## High School Researchers

- Aayush Prakash (*Half Hollow Hills High*)
- Katie Engel (*Cold Spring Harbor High*)
- Shivani Muthukumar (*Commack high*)



Funding:



National Institutes  
of Health