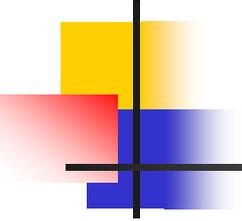


Transcription Factors

Chen-Hanson Ting

SVFIG

July 28, 2018

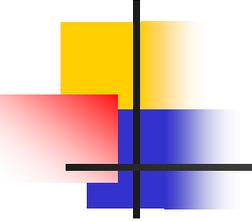


Summary

- **Cell computer postulates**
- **microRNA is the key**
- **Supporting evidences**
- **RNA polymerase**
- **Transcription factors**
- **Transcription factor binding sites**
- **My expectations**

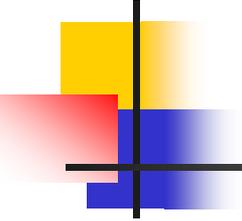
盲人摸象 《涅槃经 卷三十二》





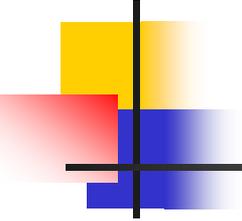
Cell Computer Postulates

- **A living cell functions by ordered release of proteins.**
- **A miRNA causes transcription of one RNA, which causes:**
 - **Transcribing a protein; or**
 - **Releasing a group of miRNAs; or**
 - **Other specific RNA function.**



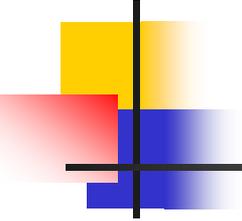
MicroRNA is the Key

- **Each cell function is represented by one miRNA.**
- **Each miRNA has a code of 20-24 bases.**
- **Each miRNA has a code field in genome, followed by a RNA field, which is transcribed to a functioning RNA.**



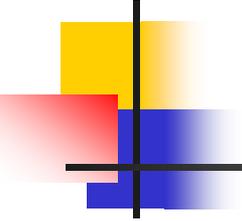
MicroRNA

- **miRNA functions by annealing to its code field in a genome.**
- **A RNA polymerase enzyme followed the annealed miRNA and transcribes a RNA whose sequence is stored in the RNA field.**



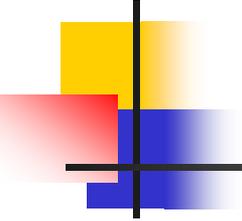
RNA Polymerase

- **A transcription factor binds to a promoter in front of a gene.**
- **RNA polymerase transcribes a RNA from DNA.**
- **The promoter generally has start code TATA.**
- **RNA polymerase stops at CCCTC.**



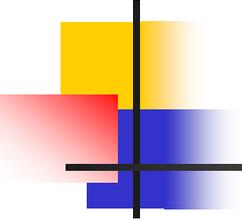
PCR & DNA Polymerase

- **PCR (polymerase Chain Reaction) is one of the greatest inventions in 20th century.**
- **It opened up the DNA in genomes and lead to genetics.**



Supporting Evidences

- **Many long non-coding lncRNAs contain clusters of miRNA.**
- **Many genomes contain clusters of miRNA.**
- **PCR (Polymerase Chain Reaction) shows that DNA polymerase can extend long DNA sequence from a primer of ~ 20 bases.**



Transcription Factors

- **Transcription Factor is a protein that controls the rate of transcription of a gene from DNA to messenger RNA, by binding to a specific DNA sequence.**
- **If a transcription factor contained a miRNA to bind to DNA, my postulate would be proven.**

Function of TF

transcription factors of eukaryotic cells

1 Activator proteins bind to pieces of DNA called enhancers. Their binding causes the DNA to bend, bringing them near a gene promoter, even though they may be thousands of base pairs away.

Enhancers

Activator proteins

Other transcription factor proteins

2 Other transcription factor proteins join the activator proteins, forming a protein complex which binds to the gene promoter.

Gene

Promoter

3 This protein complex makes it easier for RNA polymerase to attach to the promoter and start transcribing a gene.

RNA polymerase

note

This diagram simplifies the DNA greatly—promoters, enhancers, and insulators can be dozens or even hundreds of base pairs long.

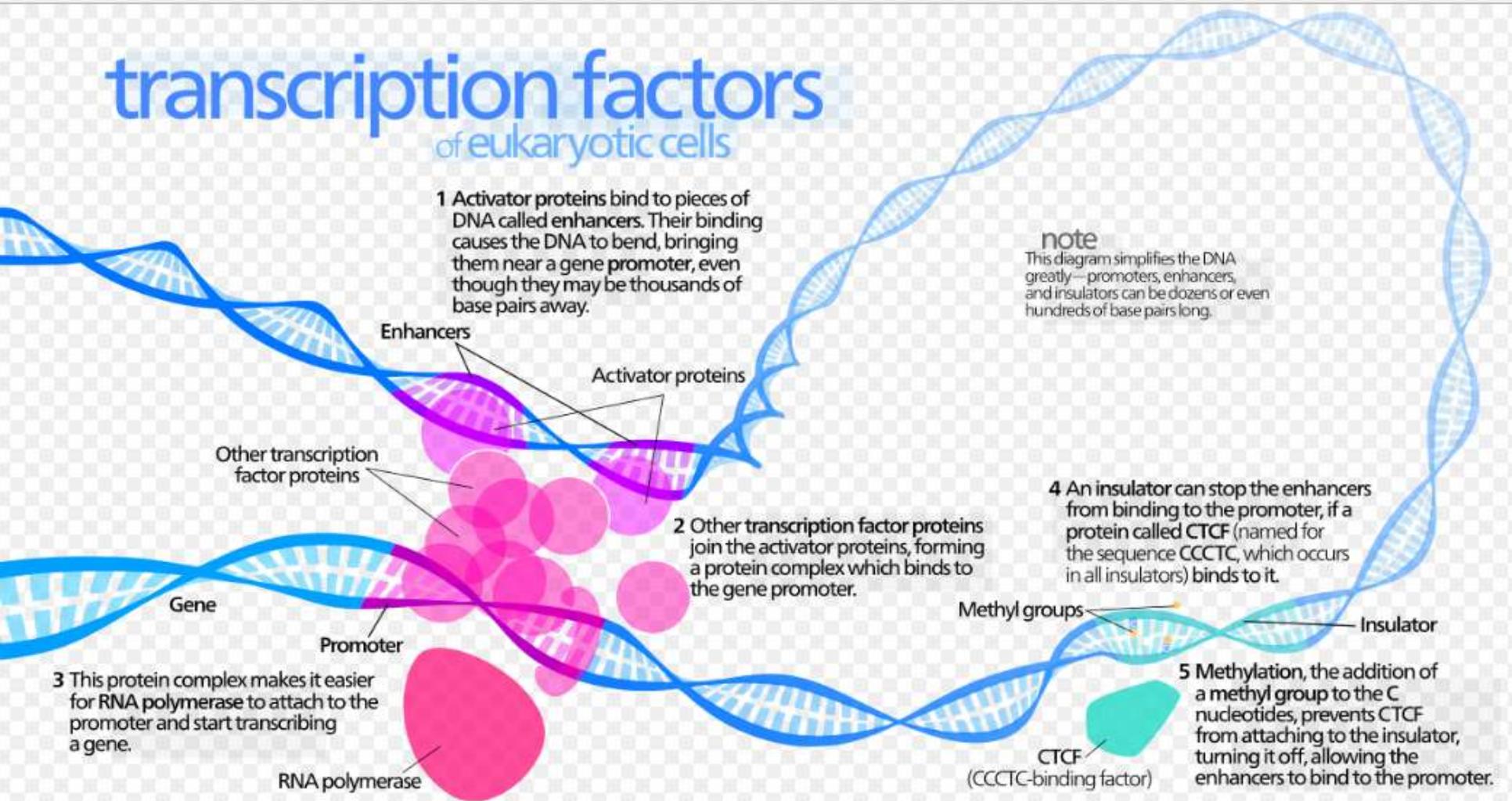
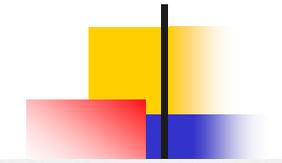
4 An insulator can stop the enhancers from binding to the promoter, if a protein called CTCF (named for the sequence CCCTC, which occurs in all insulators) binds to it.

Methyl groups

Insulator

5 Methylation, the addition of a methyl group to the C nucleotides, prevents CTCF from attaching to the insulator, turning it off, allowing the enhancers to bind to the promoter.

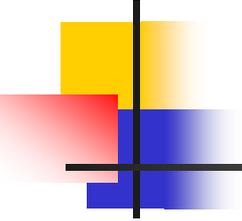
CTCF
(CCCTC-binding factor)



Structure of TF

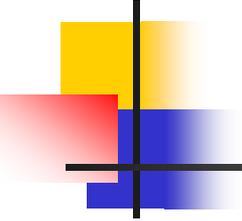


Schematic diagram of the amino acid sequence (amino terminus to the left and carboxylic acid terminus to the right) of a prototypical transcription factor that contains (1) a DNA-binding domain (DBD), (2) signal-sensing domain (SSD), and a transactivation domain (TAD). The order of placement and the number of domains may differ in various types of transcription factors. In addition, the transactivation and signal-sensing functions are frequently contained within the same domain. □



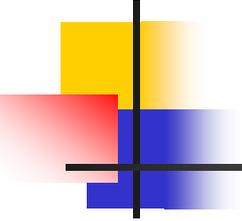
TF Specificity

- **DNA Binding Domain**
 - **How can a stretch of protein can accurately bind to specific DNA promoter?**
- **Response Element, Transcription Factor Binding Site.**
 - **No reliable experiment can accurately identify TF binding sites**



TF Specificity

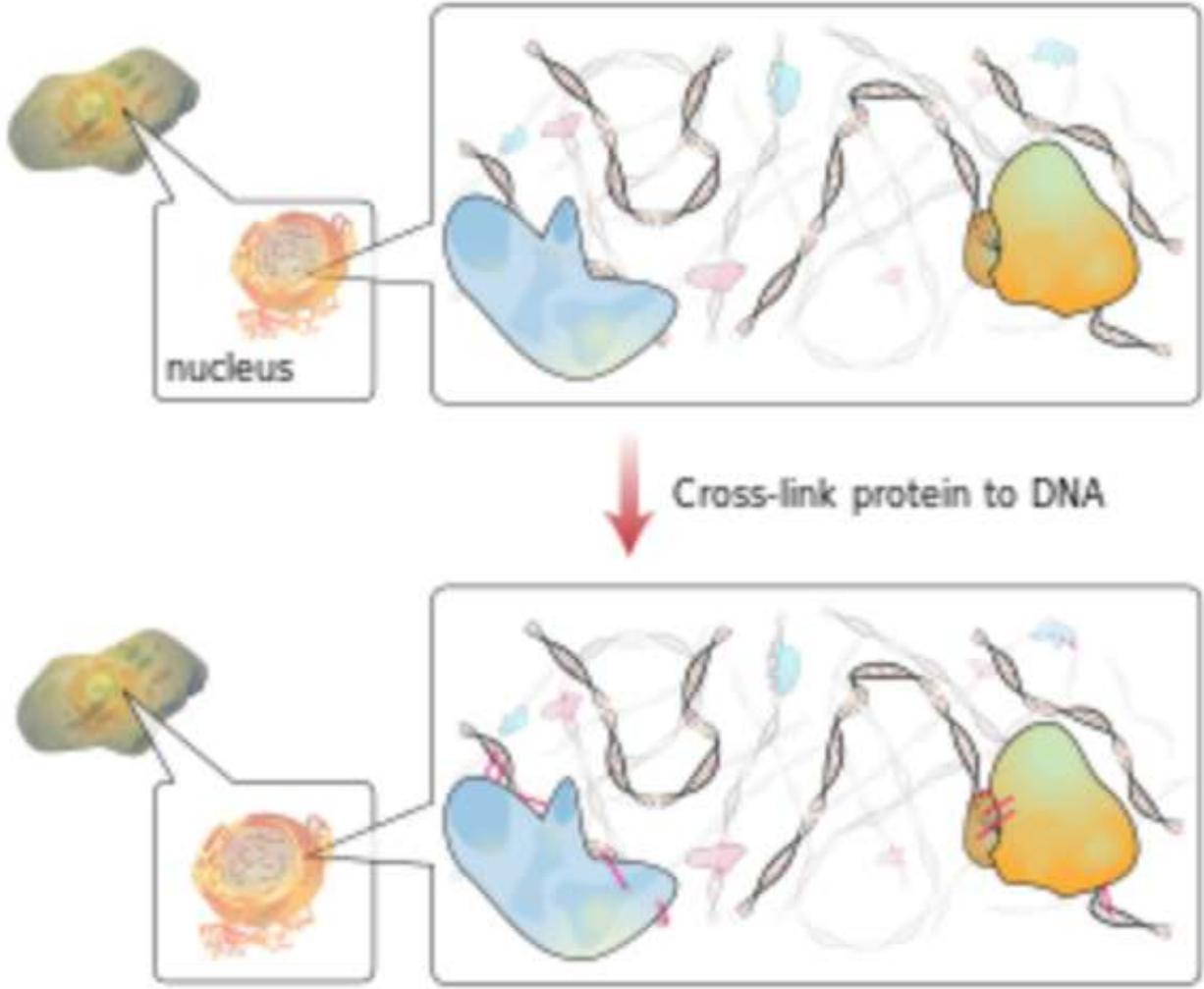
- **2600 transcription factors identified in human genome.**
- **The transcription factor binding sites are not reported so far.**



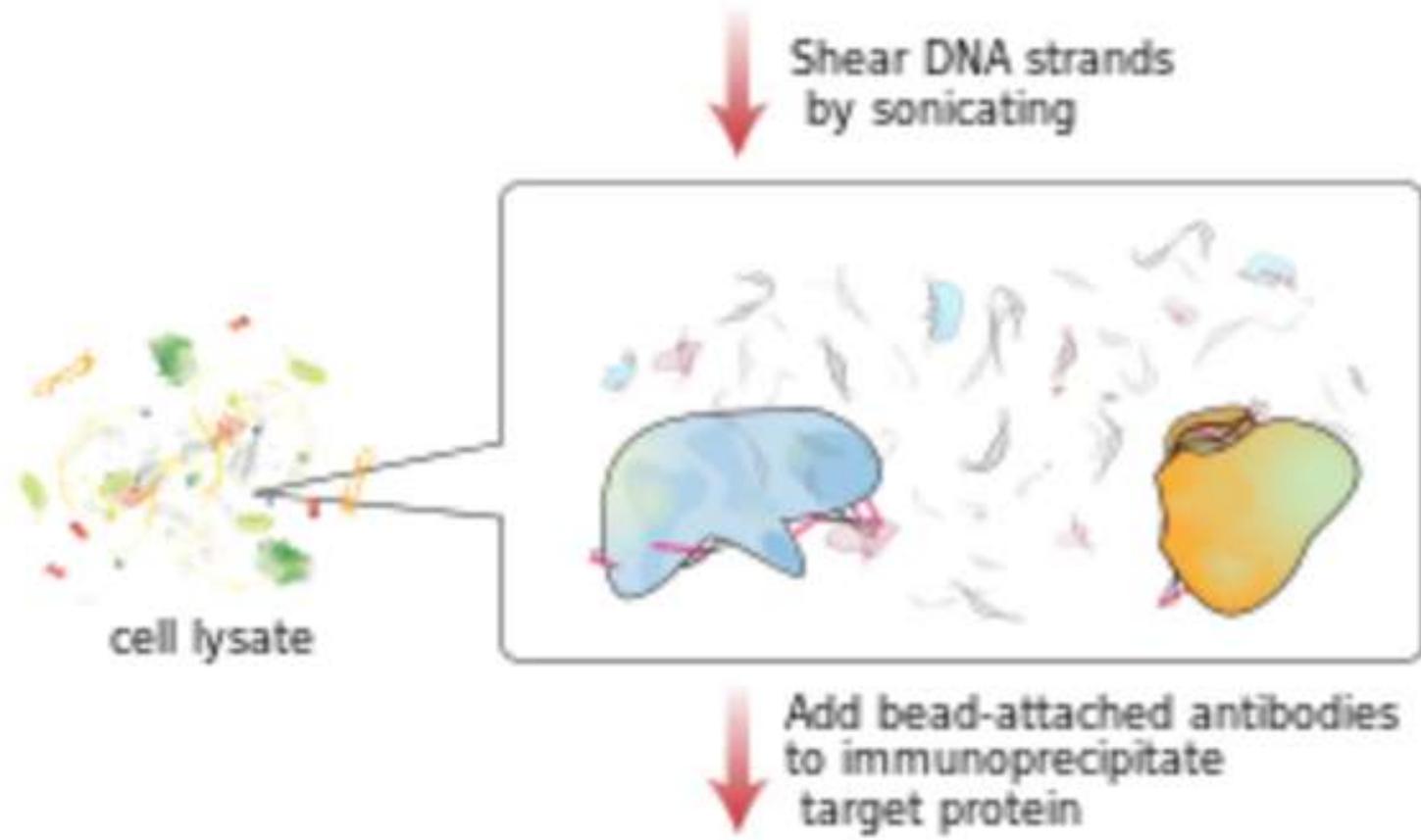
ChIP-Sequencing

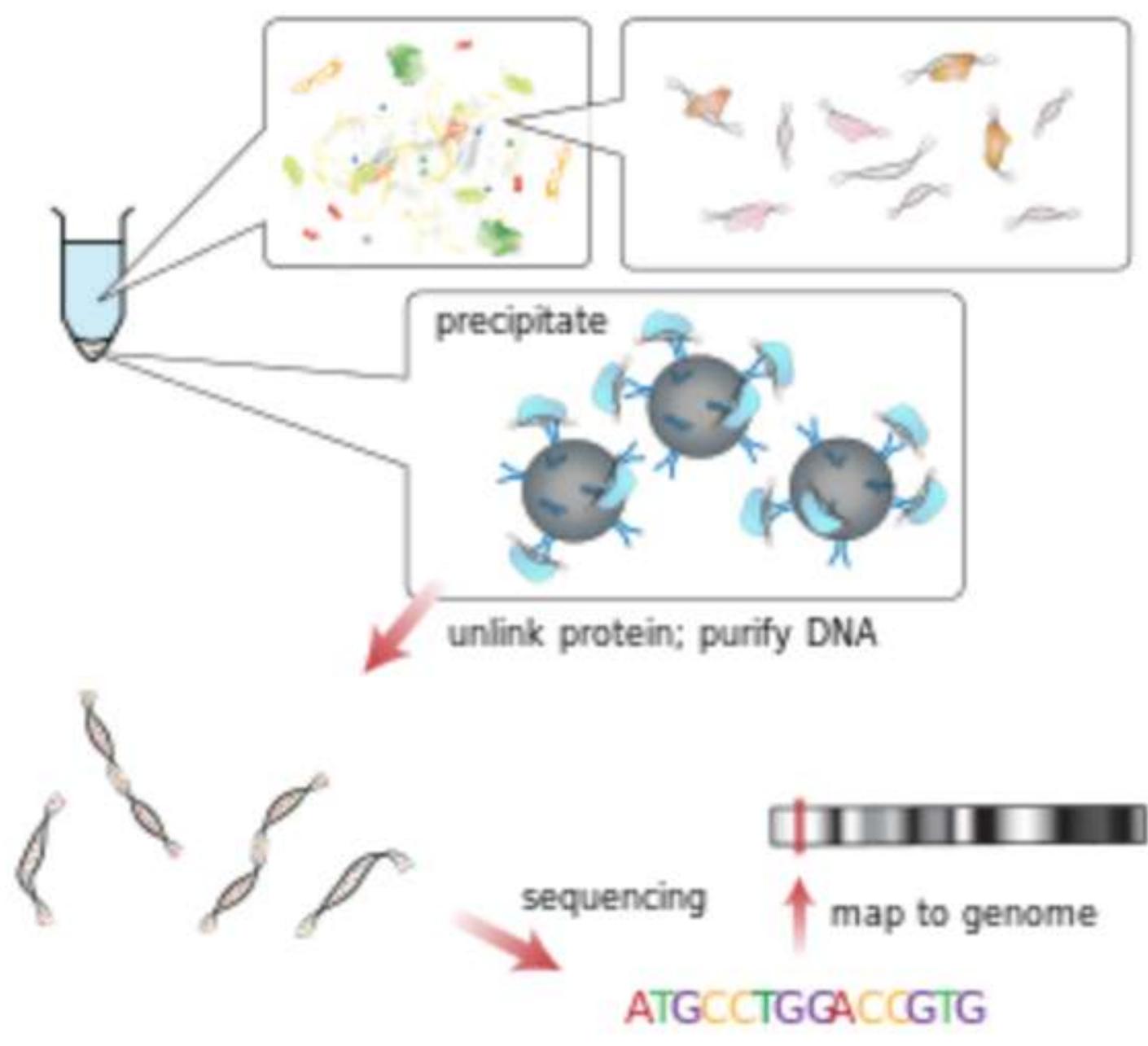
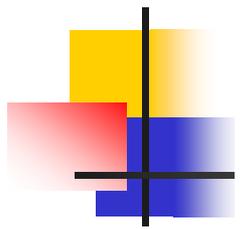
- **Chromatin Immunoprecipitation Sequencing**
- **Hundreds of ChIP sequencing experiments yielded tens of thousands sequences, but they are not transcription factor binding sites.**

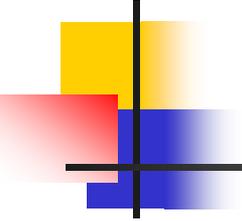
ChIP-seq



ChIP-seq

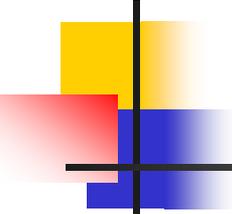






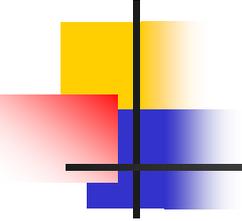
miRNA, TF Binding Sites

- **I picked one file:
ALX4_HUMAN.H11MO.0.D.words**
- **It reported 23663 sequences of experiments yielded tens of 12 bases.**
- **41 of these sequences are contained in miRNAs.**



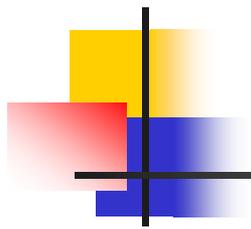
miRNA, TF Binding Sites

AAATTTAATTA	L15477	AGATTCAATTA	M12126
TGATTTAATTA	M05978	CTAGTTAATTA	M03991
TAATCCGATTA	L44635	TAATGCAATTG	L33476
TAATATAATTA	L33520	TAATACGATTA	M23422
TAAATCAATTA	L14755	GAAATCAATTA	M46405
AAACTTAATTA	L07386	TAAAGTGATTA	L28835
TAGTTTAATTA	L31182	TGCTTTAATTA	L33116
CAACCTAATTA	L45456	TATTTTAATTT	L07392
GAAGCTAATTA	M35659	TGAGTTAATTT	L23045
TTAGGTAATTA	L33620	AATTTCAATTA	M36383
TCAGTTAATTA	L21572	TAATCTCATT	M18736
TATTCCAATTA	M15471	TATGGTAATTA	M09815
TCAGCTAATTA	L48169	TAATGCAATCA	L39647
AAATTTAATTT	M09979	TAATACAATAA	M41031
TTATTTAATTG	M44455	GCATTCAATTA	M17564
TTAATCAATTA	L33669	AAGTCCAATTA	L30957
TTAGTTGATTA	M03759	AAAGTTAATTG	L34079
TAATTTGATAA	M33643	CAATCTAATTC	L41832
AAATCTAATTT	L12331	TATACCAATTA	L34088
TAATGCAATAA	L28102	AAGGTTAATTA	M35743

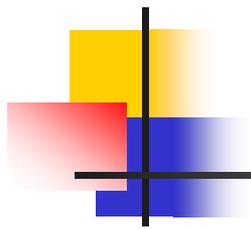


My Expectations

- **Each cell function has a miRNA.**
- **Each miRNA has a transcription factor, or something which guides RNA polymerase to transcribe one RNA.**
- **RNA polymerase has to act like DNA polymerase, starting from a simple miRNA primer.**



Questions?



Thank You!