

A photograph of a network patch panel with a dense array of red and blue fiber optic cables. The cables are bundled and connected to various ports on the panel. The background is dark, making the bright colors of the cables stand out.

Gene regulatory elements and techniques to identify and characterize them

Nadav Ahituv

UCSF

AGGAAGGGAAAGCGCAAGAGAGAGCGCACACACACCCGCCGCGCGCACTCGCGCACGGAC
CCGCACGGGGACAGCTCGGAAGTCATCAGTTCCATCGGCGAATGCTGCTGCTGGCGAGATGTCTGC
TGCTAGTCCTCGTCTCCTCGCTGCGGTATGCTCGGGACTGGCGTGCGGACCGGGCAGGGGGTTCCG
GAAGAGGAGGCACCCCAAAAAGCTGACCCCTTTAGCCTACAAGCAGTTTATCCCCAATGTGGCCGA
GAAGACCCTAGGCGCCAGCGGAAGGTATGAAGGGAAGATCTCCAGAAACTCCGAGCGATTTAAGG
AACTCACCCCAATTACAACCCCGACATCATATTTAAGGATGAAGAAAACACCGGAGCGGACAGGCT
GATGACTCAGTAGGAACCCAGCGCCGGGGCGTGGAAATGTGTGGCTTTCCAGGGGGTTACGAGAAG
CCGAACACTTCCAGACTTAACTCTGTTTGCTCTTCGGGCAGATGAAGGTGATTTACCCGCTCCTTCC
CCACCCACCTGCCCGCCCCCATTCTTCTTCTGAGGAGAATGGAGGTCAAGGGTCCAGCTG
GAGAAGTTAGGGTGTGGTGGGGGTGAGGACGGTAACAGACGTGGTTCATTATGGCCTGATTTGATGA
GTCTTGCTACAATGGCCTTCCCCATCCTACCTCTGCCTGGCTTGTAACCTGGGGAGACCTTCACTTTG
GGGGCGTCGGCCCTTTCCAGTCAGGAGTGGAATGGAAGGAGAGGCTGGGAATCCCCCTCCCACA
AACATGAAGTGGTCTCCTGGTACTGTACGAACGAACGAACGTAGCCTTGGGCTTGGAGCTCAGAGC
CCCACGTTTCCCGTTGCCTCTGTGGTTTTCTTCCACCACTACCCCCACCCTGCACCTCCCACCA
AAGAATTCTCAACTGGAAAAGCCAGGAGGCGGTTCTGACAAAAGGCAGGGGGCTCCAGGGGAGACT
CCCCGTCCCTGGGTGGCTGGCTGTATCGCAGAGCTGGCTTTGCGATTGCGTGTCCGCAATTGTGCC
CATCAGAGTGTGAATGTATTGATATTTCTTTAAGGAAGCTCTTTTCGTTCTTCCAAGCCCGAGGTACCTT
AGGGGAGGGACTTAGAACTTATTGGCATTGCATCACTTTAGTTTTCAACCTGCTTGCATAAGAATTA
GAGCGAATAAATATTAGTGTGGGGGGAGGGGAAGCTAAGCAAATATGAATTCCTCTCTCTCTCCCC
ACCTCCTTTGAGATTTCTGAGCTGCCAATCTCCAGCCAATTCTAGACTTTCTGAAACTCCATGCACG
TATAACTGAAGCCAGAAATGGGTTTCCTTGCAAATATAGGTCAACATCCTTTTTATTGCCCTATTA
AATCAAGTCCTACCTTTAGGGCTAGGTGCGTACAGCGGCTGATGGAGTGGCGCTGGTGGGGCGCAA
GTGCAGGGGGAGGGTACTGACGGCAGAGAGAGAGGAGCTACCTCCGTGCCGCCCTGCTTCCCGAC
CCGATTCCAGGCTTGCTTGAGGCCGAGAAAGGCGAGGGGGCAGGCAAGGTAGCCTGCTCCAGCTG
TCGGAAGGGAGAGGAATGGGAAATGGTCCTGATTTCTTCTCCTCATCTGCTCCCGACCACCT
TAAATCTGGACCGCGAGTGTGGACGCGCGCCAGTGCCAGACAGCAGCGCGATCCACAATTA
ACTCTGCACGGGCCATGGGGTGCCCGTTGCGTGCAGCTGGCTGGAGGGAGTTCTCCGGCTAGCCCGAG
GCGCCCATCCTCTCGTCACCCTCACTCCCCGCGGAGGAGGGGGCCTTGCCAGGGTCCCTCGGAACC
CGAGAGGAGGGAGGCACTGCGGAGACAGCGGCGGGGGCGTGGATACCCGAGGTCCAGAGCCAG
AGTGGGTGAGCTTCTGACCTGCTCTGCGGGTGAACAATACCGCAGAAGGGGTCTGGGCTCGCACA
CCTTCCAGGGCTTGAGCCTTGACGCCCTGCTGCAATAACTACCCGTGATTTATTAGCTAGCGTACTA

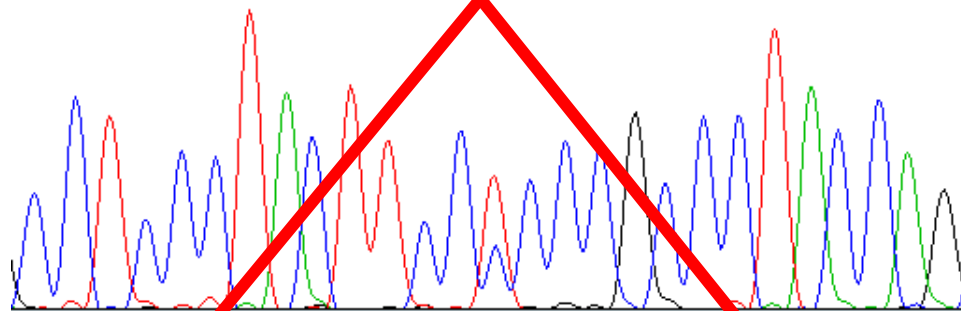
AGGAAGGGAAAGCGCAAGAGAGAGCGCACACGCACACACCCGCCGCGCGCACTCGCGCACGGAC
CCGCACGGGGACAGCTCGGAAGTCATCAGTTCCATGGGCGAATGCTGCTGCTGGCGAGATGTCTGC
TGCTAGTCCTCGTCTCCTCGCTGCGGTATGCTCGGGACTGGCGTGCGGACCGGGCAGGGGGTTCCG
GAAGAGGAGGCACCCCAAAAAGCTGACCCCTTTAGCCTACAAGCAGTTTATCCCCAATGTGGCCGA
GAAGACCCTAGGCGCCAGCGGAAGGTATGAAGGGAAGATCTCCAGAAACTCCGAGCGATTTAAGG
AACTCACCCCAATTACAACCCCGACATCATATTTAAGGATGAAGAAAACACCGGAGCGGACAGGCT
GATGACTCAGTAGGAACCCAGCGCCGGGGCGTGGAATGTGTGGCTTTCCAGGGGGTTACGAGAAG
CCGAACACTTCCAGACTTAACTCTGTTTGCTCTTCGGGCAGATGAAGGTGATTTACCCGCTCCTTCC
CCACCCACCTGCCCGCCCCCATTCTTCCTCTTCCTGGAGGAGAATGGAGGTCAAGGGTCCAGCTG
GAGAAGTTAGGGTGTGGTGGGGGTGAGGACGGTAACAGACGTGGTTCATTATGGCCTGATTTGATGA
GTCTTGCTACAATGGCCTTCCCCATCCTACCTCTGCCTGGCTTGTAACCTGGGGAGACCTTCACTTTG
GGGGCGTCGGCCCTTTCCAGTCAGGAGTGGAATGGAAGGAGAGGCTGGGAATCCCCCTCCCACA
AACATGAAGTGGTCTCCTGGTACTGTACGAACGAACGAACGTAGCCTTGGGCTTGGAGCTCAGAGC
CCCACGTTTCCCGTTGCCTCTGTCTTTTCTTCCACCCTACCCCCACCCTGCACCTCCCCACCA
AAGAATTCTCAACTGGAAAAGCGAGACGCGTCTCGCAAAAGGCAGGGGGCTCCAGGGGAGACT
CCCCGTCCCTGGGTGGCTGGCCTTCCCGAGCTGCTTTGCGATTGCGTGTCCGCAATTGTGCC
CATCAGAGTGTGAATGTATTGATATTTTAAAGAACTCTTCGTTCTTCCAAGCCCGAGGTACCTT
AGGGGAGGGACTTAGAACTTATTGGCATTGCATCACTTTAGTTTTCAACCTGCTTGCATAAGAATTA
GAGCGAATAAATATTAGTGTGGGGGGAGGGGAAGCTAAGCAAATATGAATTCCTCTCTCTCCCC
ACCTCCTTTGAGATTTCTGAGCTGCCAATCTCCAGCCAATTCTAGACTTTCTGAAACTCCATGCAG
TATAACTGAAGCCAGAAATGGGTTTCCTTGCAAATATAGGTCAACATCCTTTTTATTGCCCTATTA
AATCAAGTCCTACCTTTAGGGCTAGGTGCGTACAGCGGCTGATGGAGTGGCGCTGGTGGGGCGCAA
GTGCAGGGGGAGGGTACTGACGGCAGAGAGAGAGGAGCTACCTCCGTGCCGCCCTGCTTCCCGAC
CCGATTCCCAGGCTTGCTTGAGGCCGAGAAAGGCGAGGGGGCAGGCAAGGTAGCCTGCTCCAGCTG
TCGGAAGGGAGAGGAATGGGAAATGGTCCTGATTTCTTGCTCTCCCTCATCTGCTCCCGACCACCT
TAAATCTGGACCGCGAGTGTGGACGCGCGCCAGTGCCAGACAGCAGCGCGATCCACAATTA
CTGCACGGGCCATGGGGTGCCCGTTGCGTGCAGCTGGCTGGAGGGAGTTCTCCGGCTAGCCCGAG
GCGCCCATCCTCTCGTCACCCTCACTCCCCGCGGAGGAGGGGCCTTGCCAGGGTCCCTCGGAACC
CGAGAGGAGGGAGGCACTGCGGAGACAGCGGCGGGGGCGTGATACCCGAGGTCCCAGAGCCAG
AGTGGGTGAGCTTCTGACCTGCTCTGCGGGTGACCAATACCGCAGAAGGGGTCTGGGCTCGCACA
CCTTCCCAGGGCTTGAGCCTTGACGCCCTGCTGCAATAACTACCCGTGATTTATTAGCTAGCGTACTA

98%

Second base

		Second base					
		U	C	A	G		
First base	U	UUU } Phenylalanine F UUC } UUA } Leucine L UUG }	UCU } Serine S UCC } UCA } UCG }	UAU } Tyrosine Y UAC } UAA } Stop codon UAG } Stop codon	UGU } Cysteine C UGC } UGA } Stop codon UGG } Tryptophan W	U C A G	
	C	CUU } Leucine L CUC } CUA } CUG }	CCU } Proline P CCC } CCA } CCG }	CAU } Histidine H CAC } CAA } Glutamine Q CAG }	CGU } Arginine R CGC } CGA } CGG }	U C A G	
	A	AUU } Isoleucine I AUC } AUA } AUG } Methionine start codon M	ACU } Threonine T ACC } ACA } ACG }	AAU } Asparagine N AAC } AAA } Lysine K AAG }	AGU } Serine S AGC } AGA } Arginine R AGG }	U C A G	
	G	GUU } Valine V GUC } GUA } GUG }	GCU } Alanine A GCC } GCA } GCG }	GAU } Aspartic acid D GAC } GAA } Glutamic acid E GAG }	GGU } Glycine G GGC } GGA } GGG }	U C A G	

Third base



CCTCCTACTTCC^CCCGCCCTACCAG
T

ProProTyrPheProProProTyrGln

▼
Leu

Noncoding regulatory sequences

nucleotide
changes



Differences in gene expression

Human variation



Pharmacogenetics



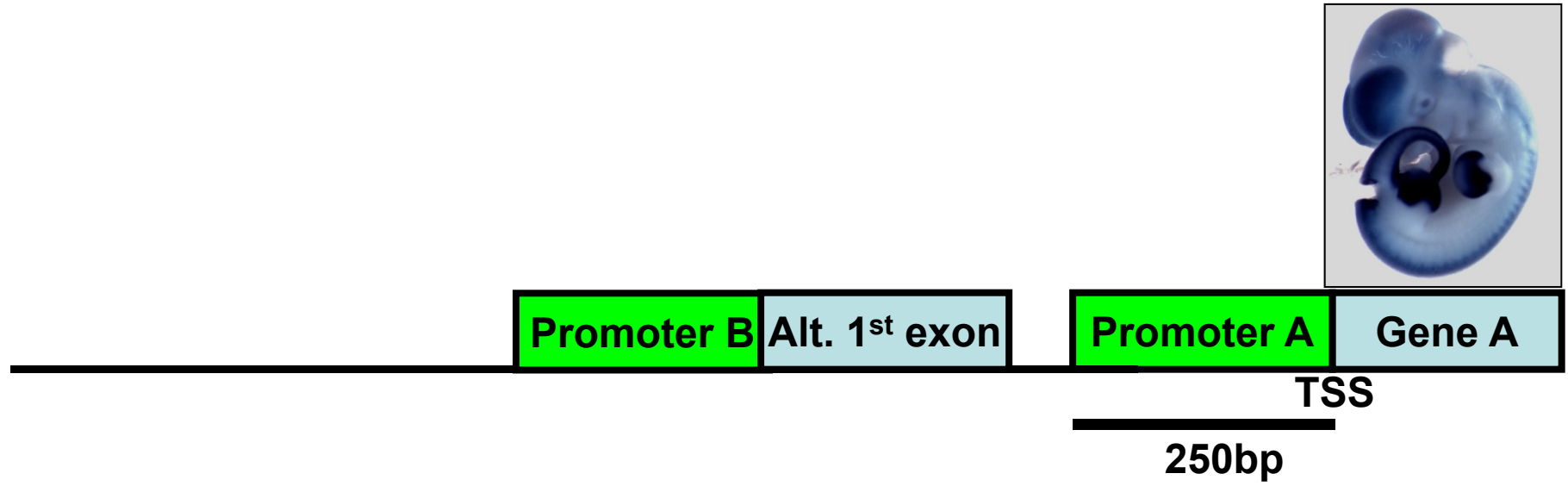
Human disease



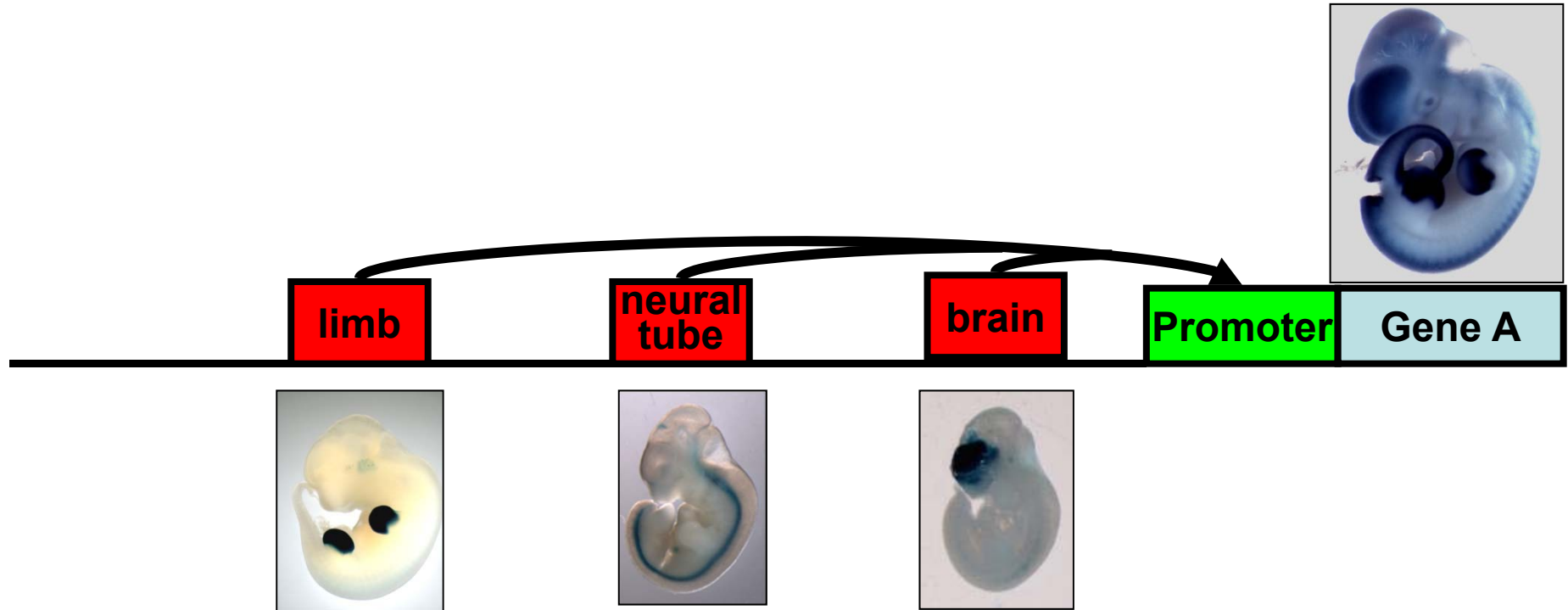
OUTLINE

- 1. The different kinds of gene regulatory elements**
- 2. How can we find them**
- 3. How far away to look for them**
- 4. Functional assays to test their function**
- 5. Regulation in 3D**
- 6. Example from our lab**

Promoters

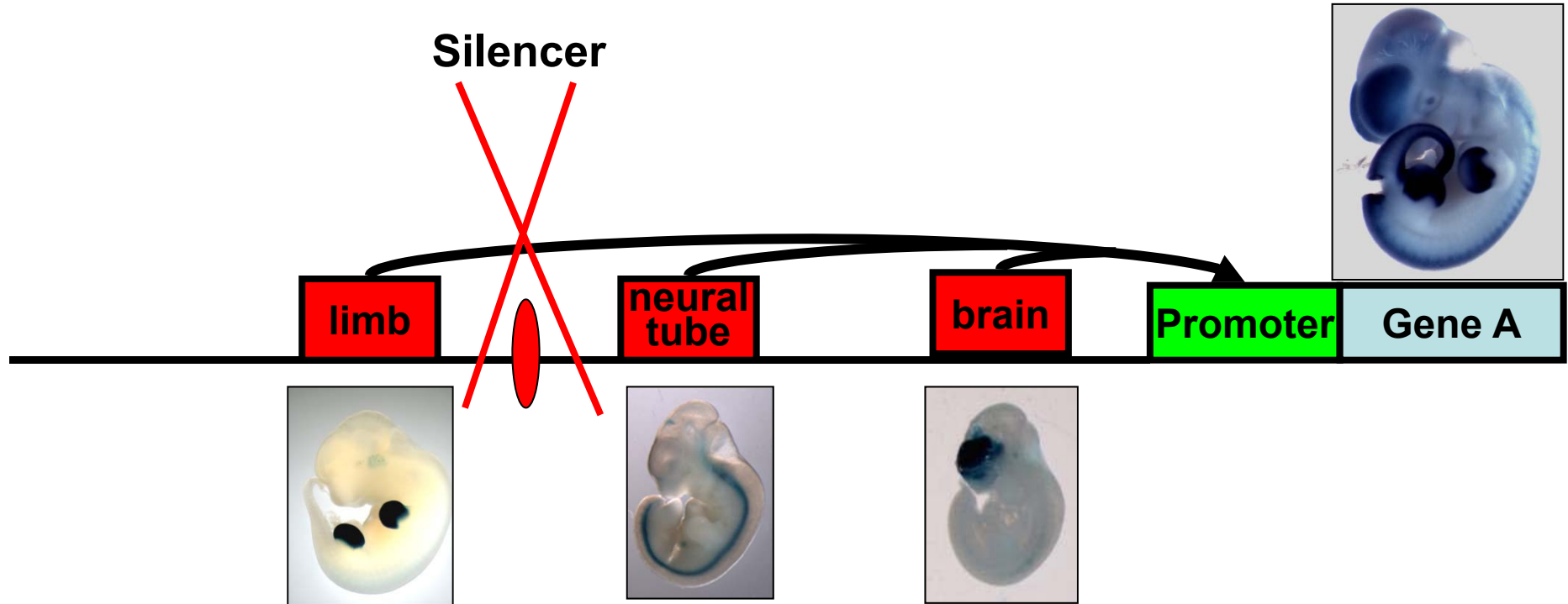


Enhancers



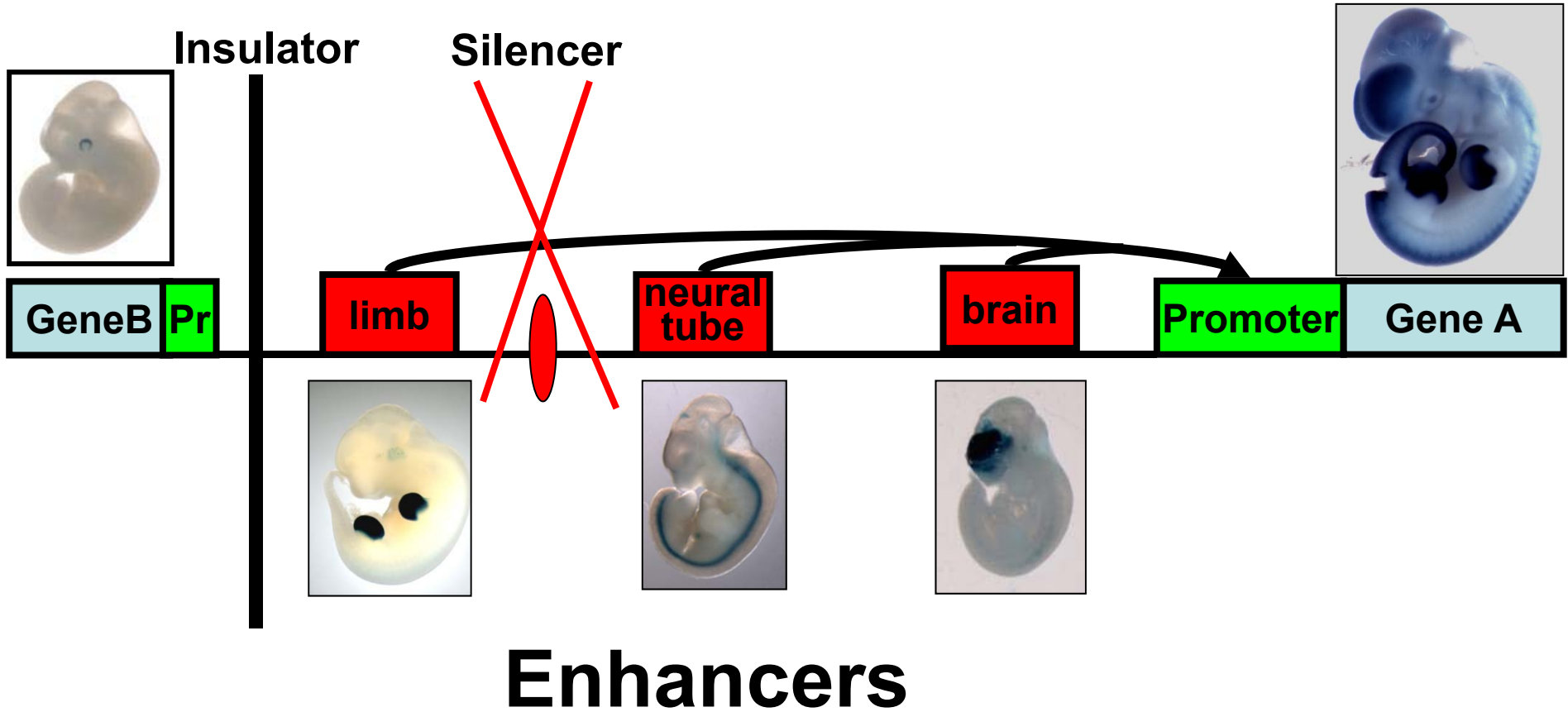
Enhancers

Silencers



Enhancers

Insulators



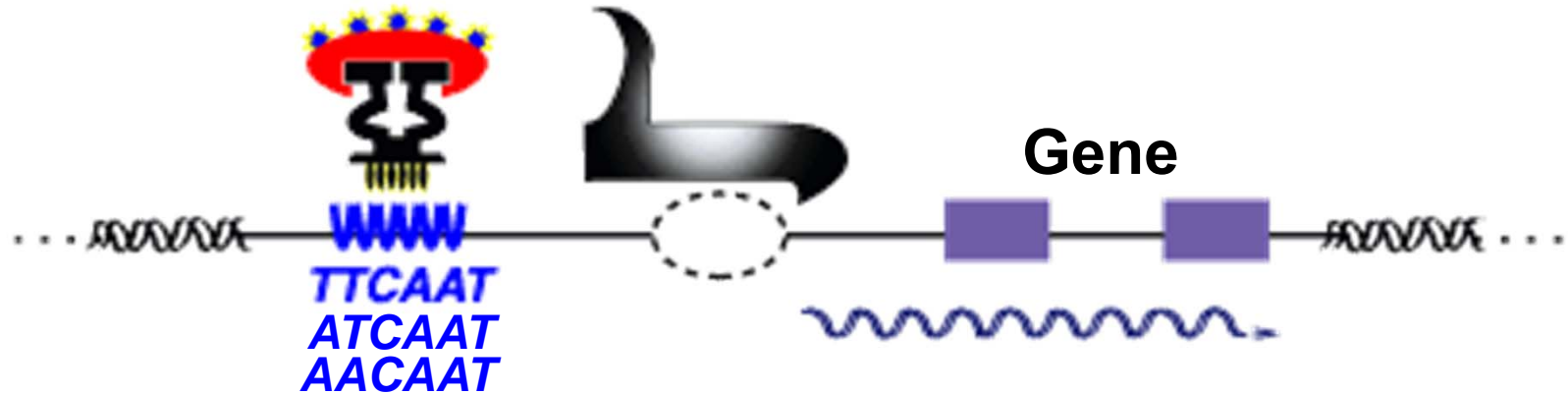
OUTLINE

- 1. The different kinds of gene regulatory elements**
- 2. How can we find them**
- 3. How far away to look for them**
- 4. Functional assays to test their function**
- 5. Regulation in 3D**
- 6. Example from our lab**

How do we find these regulatory elements?

- 1. Transcription factor binding sites (TFBS)**
- 2. Comparative Genomics**
- 3. Chromatin Immunoprecipitation (ChIP)**

Position Weight Matrix



	1	2	3	4	5	6
A	60	20	0	90	80	0
C	0	0	80	0	0	0
G	0	0	0	0	0	20
T	40	80	20	10	20	80

Logo plot



Transcription factor binding site programs



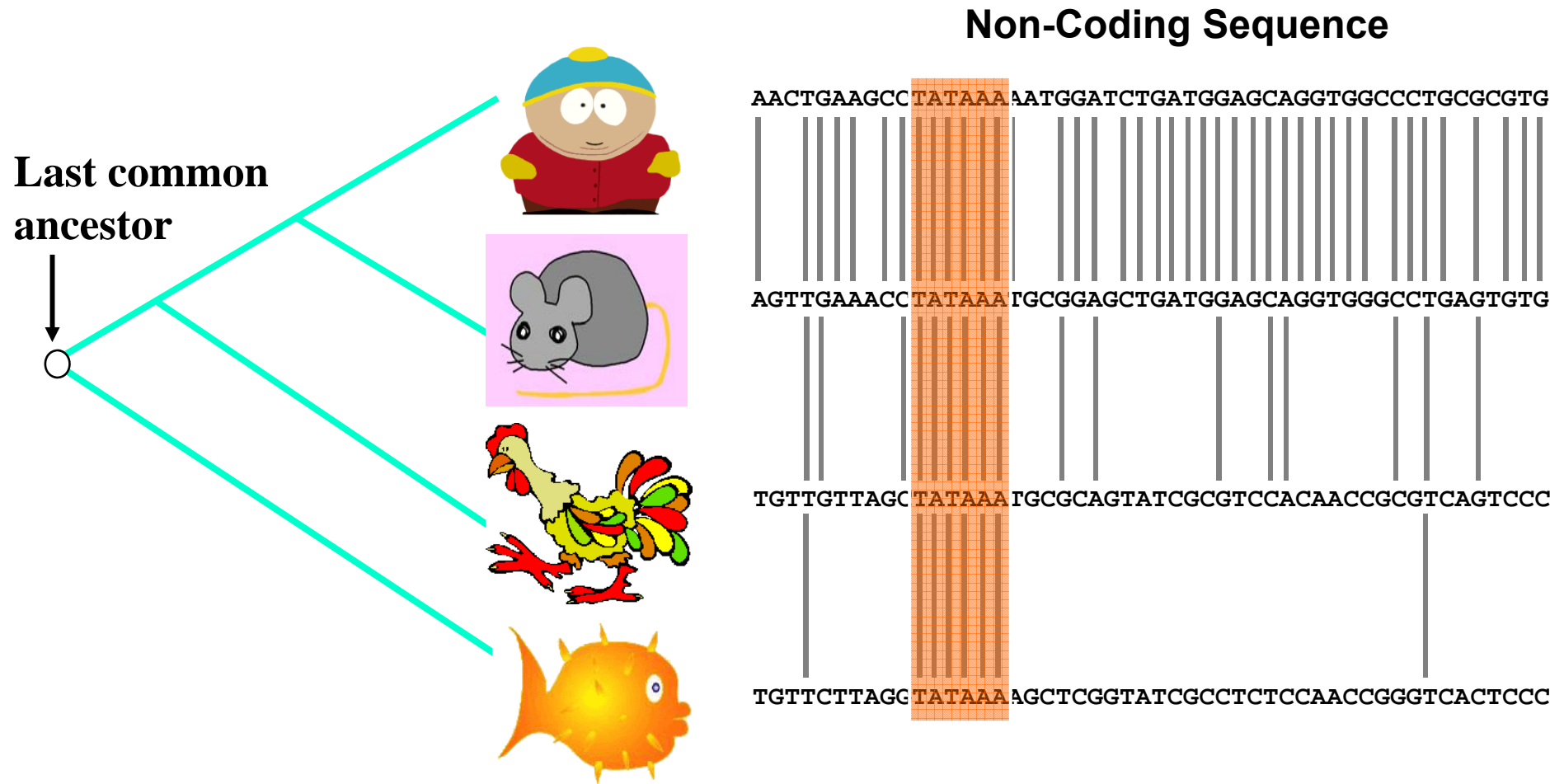
- **Match** (<http://www.gene-regulation.com/pub/programs.html#match>)
- **MatInspector** (http://www.genomatix.de/online_help/help_matinspector/matinspector_help.html)
- **rVISTA** (<http://rvista.dcode.org/>)
- **TF Search** (<http://www.cbrc.jp/research/db/TFSEARCH.html>)

ACAAACATGAAGTGGTCTCCGCGTGGTGGCAGACGAACGAACGTAGCCTTGGGCTTGGAGCTCAG
AGCCCCACGTTTCCCGTTGCCTCTGTGGTTTTCTTTCCACCACTACCCCCACCCTGCACCTCCCCA
CCAAAGAATTCTCAACTGGAAAAGCCAGGAGGCGGTTCTGACAAAAGGCAGGGGCTCCAGGGGAG
ACTCCCCCGTCCCTGGGTGGCTGGCTGTATCGCAGAGCTGGCTTTGCGATTGCGTGTCCGCAATTGT
GCCCATCAGAGTGTGAATGTATTGATATTTCTTTAAGGATGCTCTTTTCGTTCTTCCAAGCCCGAGGTAC
CTTAGGGGAGGGACTTAGAACTTATTGGCATTGCATCACTTTAGTTTTCAACCTGCTTGCATAAGAATT
AAGAGCGAATAAATATTAGTGTGGGGGGAGGGGAAGCTAAGCAAATATGAATTCCTCTCTCTCTCC
CCACCTCCTTTGAGATTTCTGAGCTGCCAATCAATCAGCCAATTCTAGACTTTCTGAAACTCCATGCA
CGTATAACTGAAGCCAGAAATGGGTTTCCTTGCAAATATAGGTCAACATCCTTTTTATTGCCCTATTAA
AATATTCAAGTCTACCTTTAGGGCTAGGTGCGTACAGCGGCTGATGGAGTGGCGCTGGTGGGGCGC
AAGTGCAGGGGGAGGGTACTGACGGCAGAGAGAGAGGAGCTACCTCCGTGCCGCCCTGCTTCCCG
ACCCGATTCCAGGCTTGCTTGAGGCCGAGAAAGGCGAGGGGCAGGCAAGGTAGCCTGCTCCAGC
TGTCGGAAGGGAGAGGAATGGGAAATGGTCCTGATTTCTTGCTCTCCCTCATCTGCTCCCGACCAC
CTTAAATCTGGACCGCGAGTGTGGACGCGCGCCAGTGCCAGACAGCAGCGCGATCCACAATTAA
CTCTGCACGGGCCATGGGGTGCCCGTTGCGTGCAGCTGGCTGGAGGGAGTTCTCCGGCTAGCCCG
AGGCGCCCATCCTCTCGTCACCCTCACTCCCCGCGGAGGAGGGGCCTTGCCATAGCTACGTTACGA

How do we find these regulatory elements?

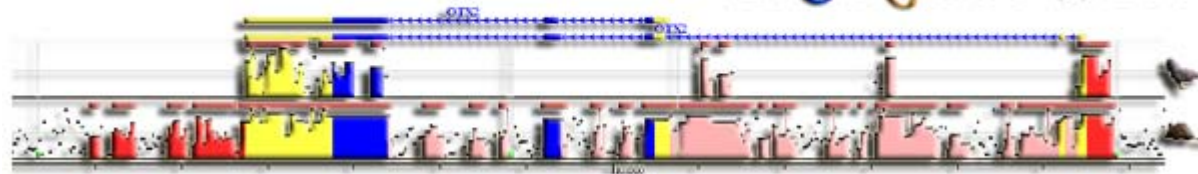
- 1. Transcription factor binding sites (TFBS)**
- 2. Comparative Genomics**
- 3. Chromatin Immunoprecipitation (ChIP)**

Evolutionary conservation can find regulatory elements



**Evolutionarily Conserved Sequences
are Functionally Important.**

ECR Browser



<http://www.dcode.org> - [NCBI DCODE.org](http://www.dcode.org) Comparative Genomics Developments

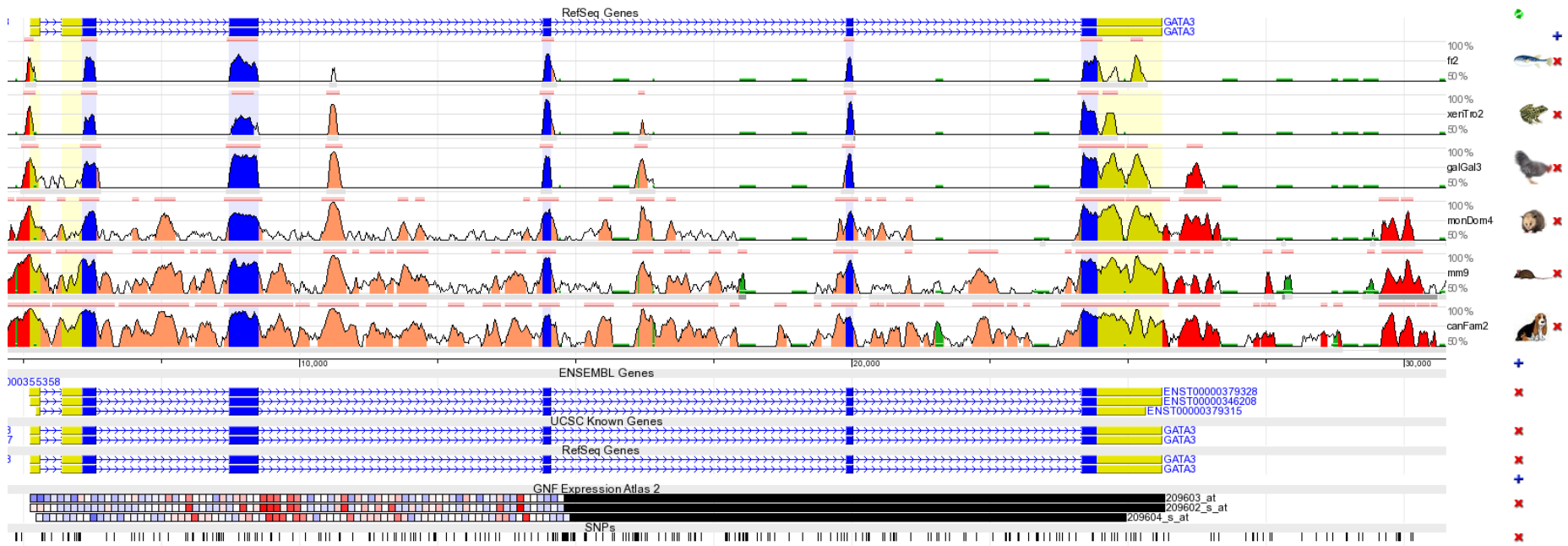
[ECRs](#) [Conserved SNPs](#) [Synteny/Alignments](#) [Custom annotation](#) [Core ECRs \[?\]](#) [Reset settings](#) [Instructions](#) [External tools](#)

ECR Browser on Human(hg18) <http://ecrbrowser.dcode.org> 30,744 bps

Parameters: Graph ECR ECR Layer Coordinate
[change] smooth length similarity height system
relative

gene or position (chrN:from-to)
chr10:8131549-8162292 **Submit**

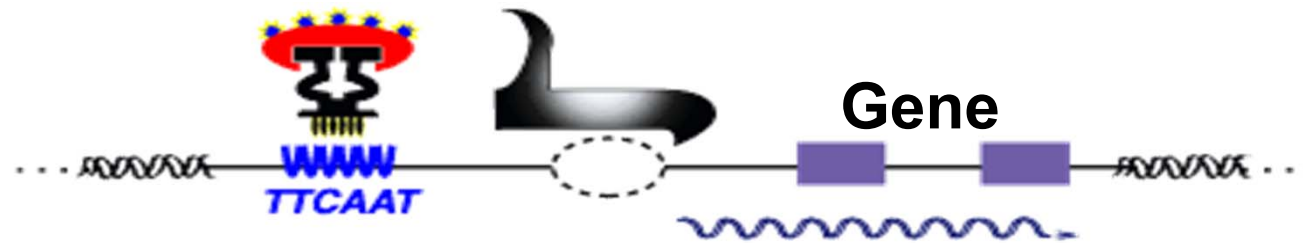
[GENOME ALIGNMENT](#): Align your sequence to a genome



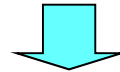
Flip move: Left Right zoom in: 1.5x 3x 10x zoom out: 1.5x 3x 10x

<http://ecrbrowser.dcode.org/>

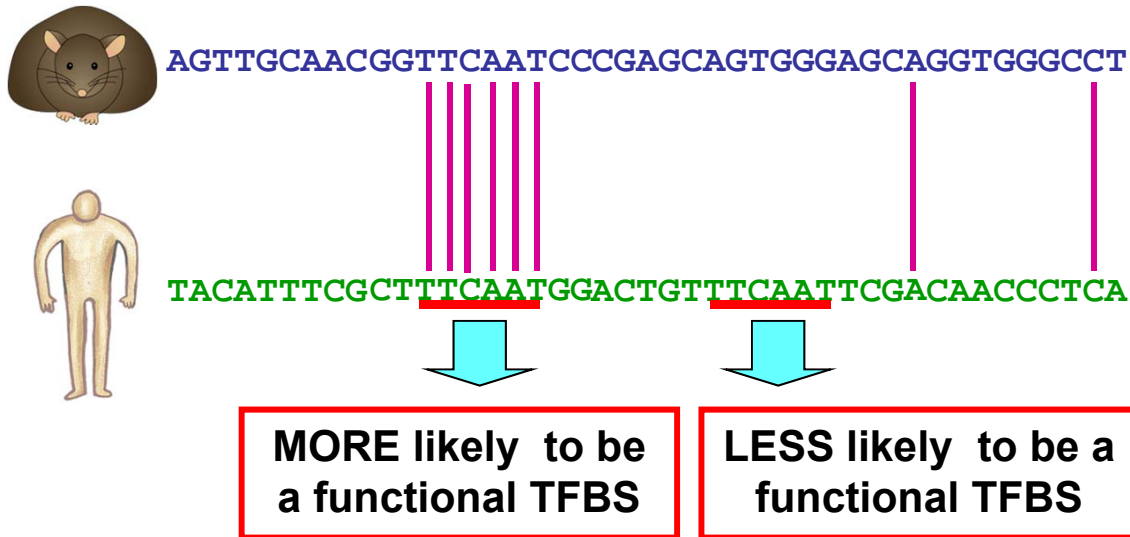
Combining comparative genomics & TFBS



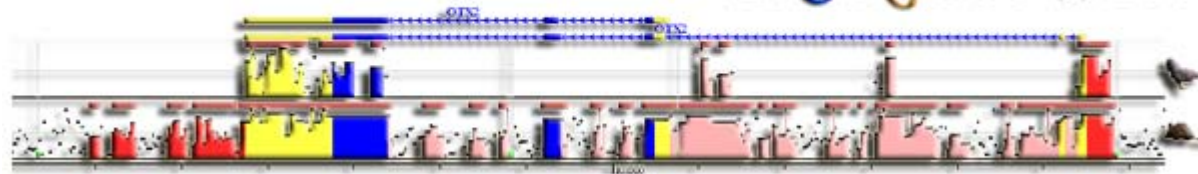
6bp sequences can appear ~734,421 times in the human genome by chance



Conservation can be used as a filter to screen TFBS more likely to be functional



ECR Browser



http://www
ECR :: Evolutionary Conserved Region
Close

ECRs
Conserved SNPs
Syntax/Alignmen

ECR Browser on Human(hg18) [http://](#)

Parameters: Graph ECR ECR
 [change] smooth length similarity
 100 70

ECR [Evolutionary Conserved Region]

location: chr10:8142040-8142264
 length: 225 bps
 identity: 72.9%

Alignment

overlapping alignment block:
 chr10:8142054-8142253 -vs- xenTro2:scaffold_11:4287054-4287257

	4287054	4287074	4287093
xenTro2	GAGATGTGATTTTCAGGGGGTTTTT-CCAGGGTTACATTTTCTTTGCCTGTCTGAGCTGGGA		
hg18	GAGATCAGTTTTTCAGGGTTTTTTTCCAGGGTGACATTCACCTCCGACTGCCTGAGC-AGGA		
	8142054	8142074	8142094

	4287113	4287133	4287153
xenTro2	CTGTCCTAATTTAGCTGATATGTTTGAGTTAATCCAATTAATTTTCAGACTGCTGCACGGC		
hg18	CTGTCCTAATTTAGTTGATAATGTTTTAGCTAATCCAATTAATTTTCAGACTGCTGCACGGC		
	8142113	8142133	8142153

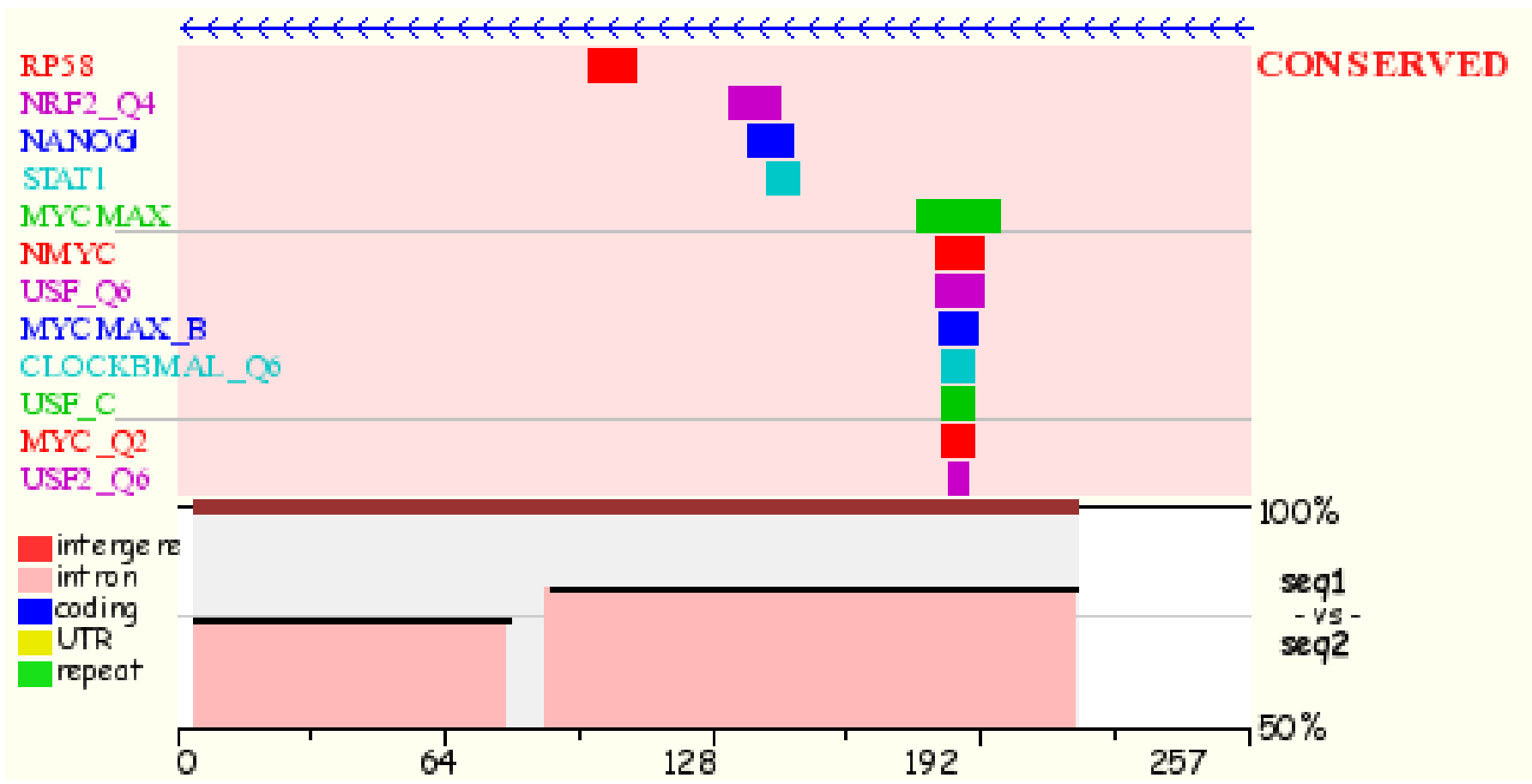
	4287173	4287193	4287213
xenTro2	ACAGTTTCATTGTTTCCCCAGCGCCAAAGGAGTTAATAGAGTCCGGATGTGGCCAGACTG		
hg18	ACAGTTGCATTGTTTATCCTGAGCCAGGAACITTAATAGAGTCCGGATGCGGTTAGGCT-		
	8142173	8142193	8142213

	4287233	4287253
xenTro2	GGGGGCAGCACAGAGCAGAGCTGCA	
hg18	---GACAGAAGAACCAGACCTGCA	
		8142249

Instructions

External tools

<http://ecrbrowser.dcode.org/>



Annotating new sequences/genomes



Make DNA



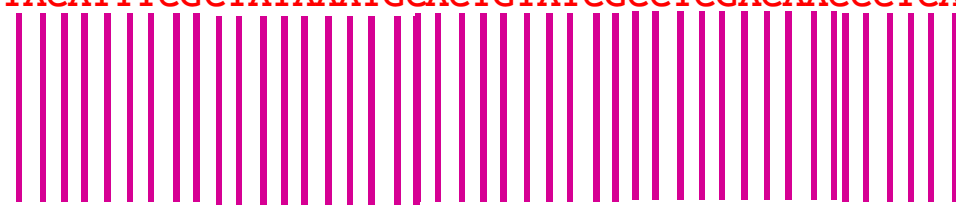
Compare DNA

Alien DNA

TACATTTGCTATAAATGCACTGTATCGCCTCGACAACCCTCA

Jellyfish GFP

TACATTTGCTATAAATGCACTGTATCGCCTCGACAACCCTCA



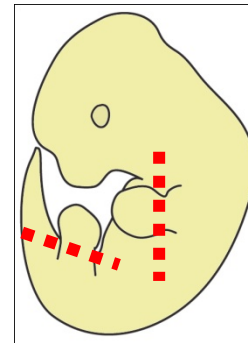
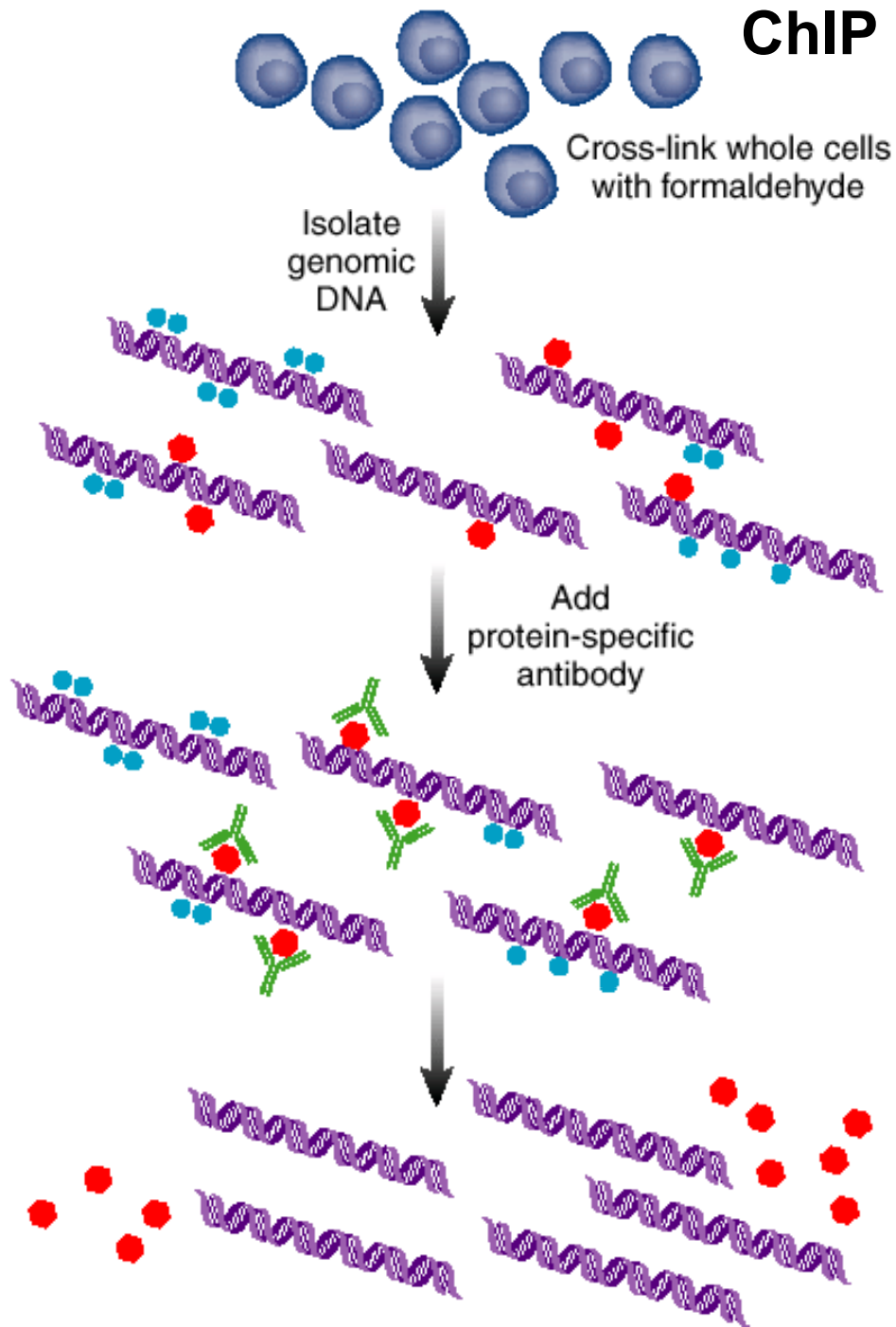
How do we find these regulatory elements?

- 1. Transcription factor binding sites (TFBS)**
- 2. Comparative Genomics**
- 3. Chromatin Immunoprecipitation (ChIP)**

ChIP-seq provides a major advancement

<u>Method</u>	<u>Enhancer Prediction Rate</u>	<u>Enhancer Tissue Specificity</u>
Comparative genomics	20-50%	<20%
Comparative genomics + TFBS	50-60%	20-50%
ChIP-seq p300	87%	84%

ChIP



Sonicate DNA to produce sheared, soluble chromatin

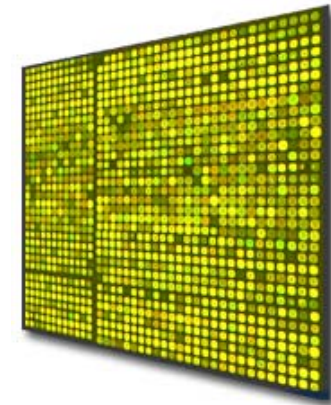
Immunoprecipitate and purify immunocomplexes

Reverse cross-links, purify DNA and prepare for sequencing

1. qPCR



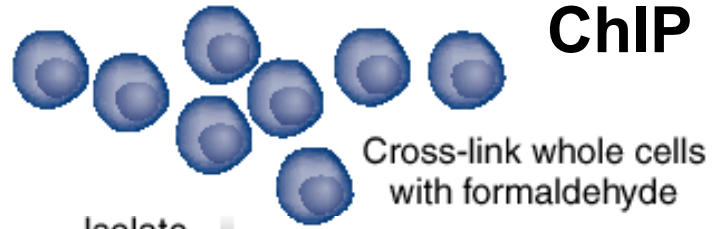
2. CHIP



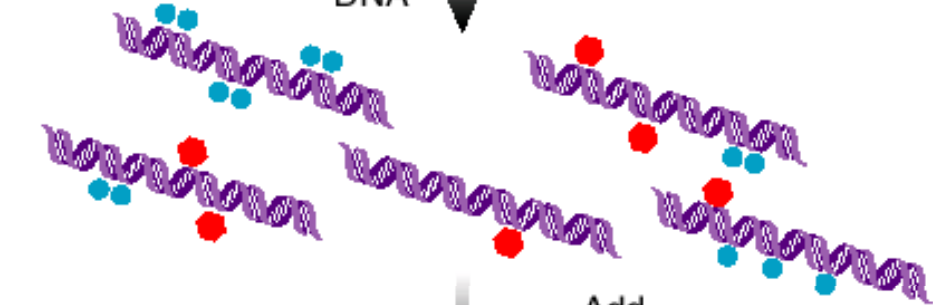
3. Seq



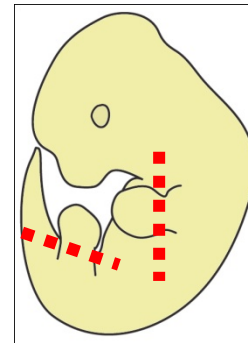
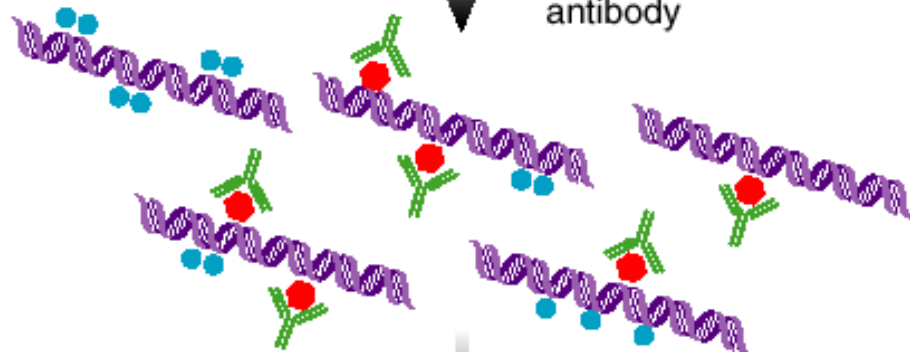
ChIP



Isolate genomic DNA



Add protein-specific antibody

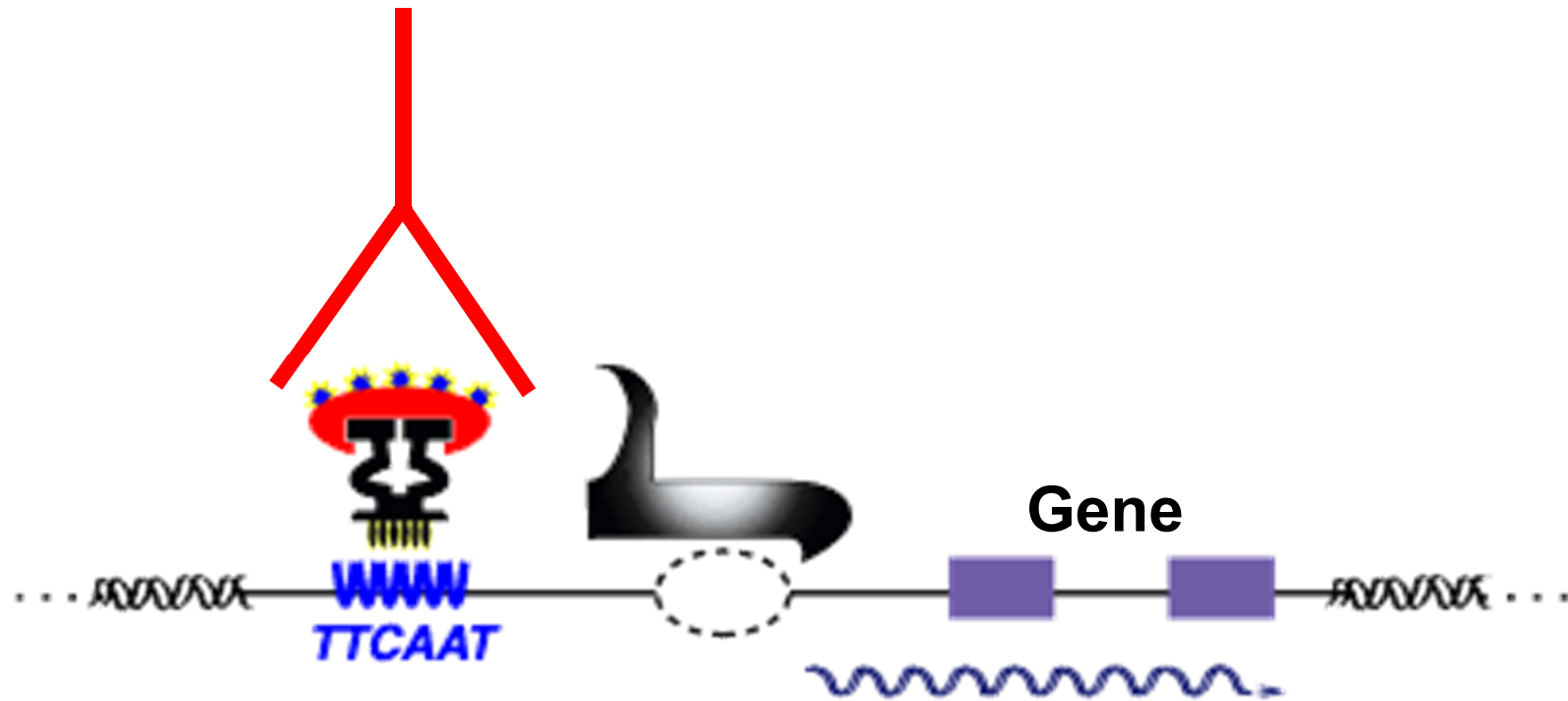


Sonicate DNA to produce sheared, soluble chromatin

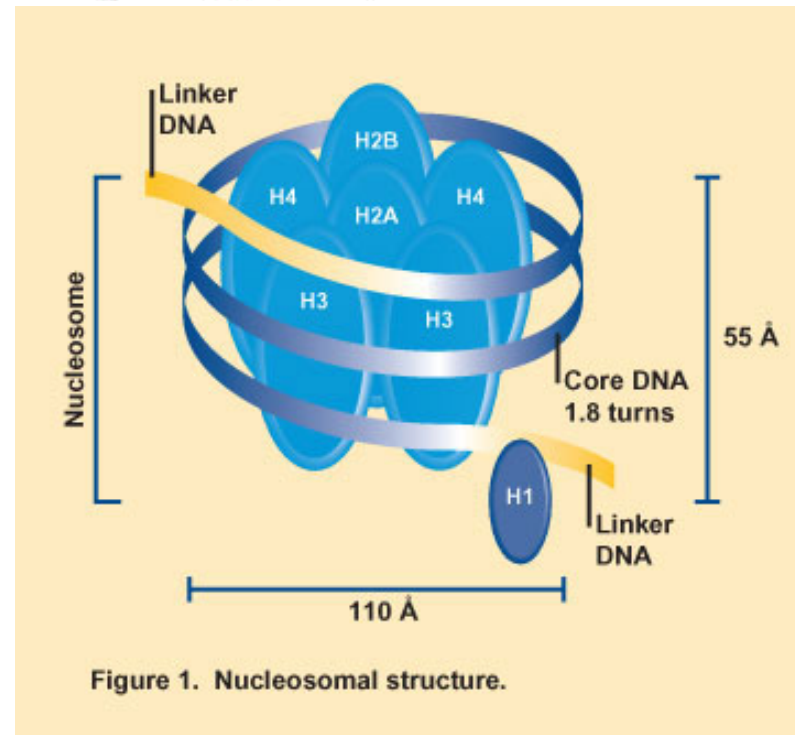
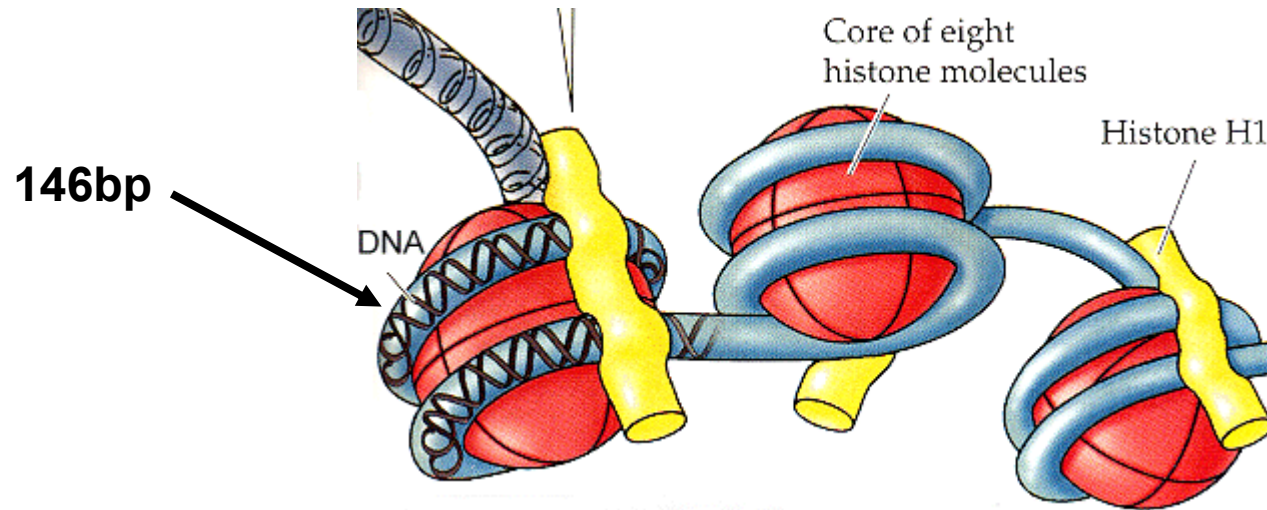
Immunoprecipitate and purify immunocomplexes

WHAT ANTIBODIES?

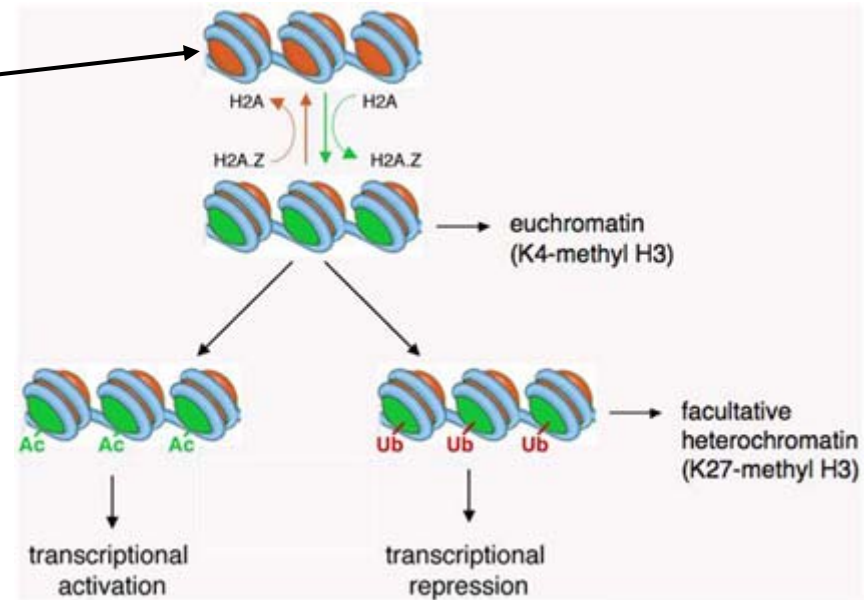
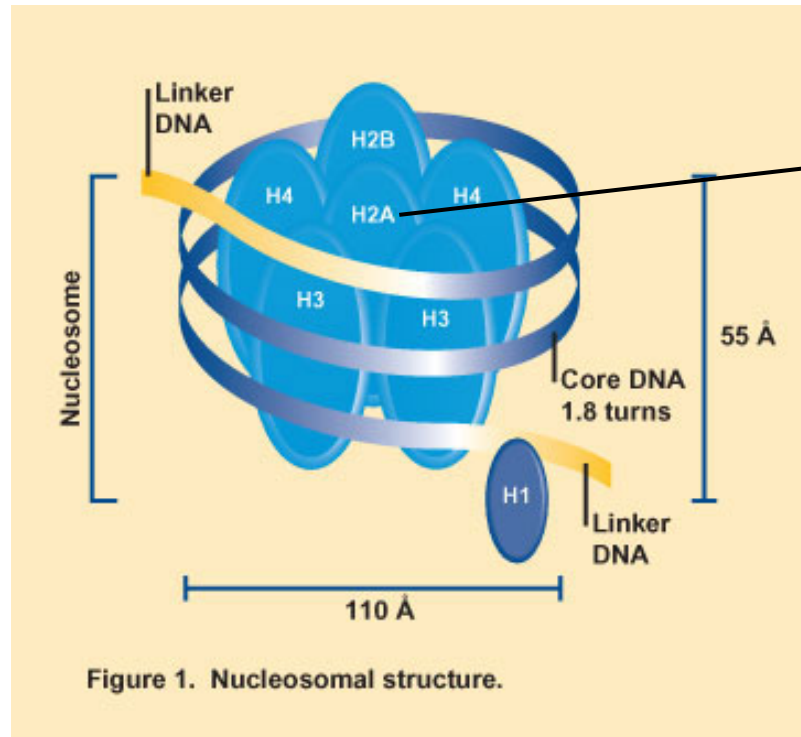
Transcription Factors



Histone Modifications

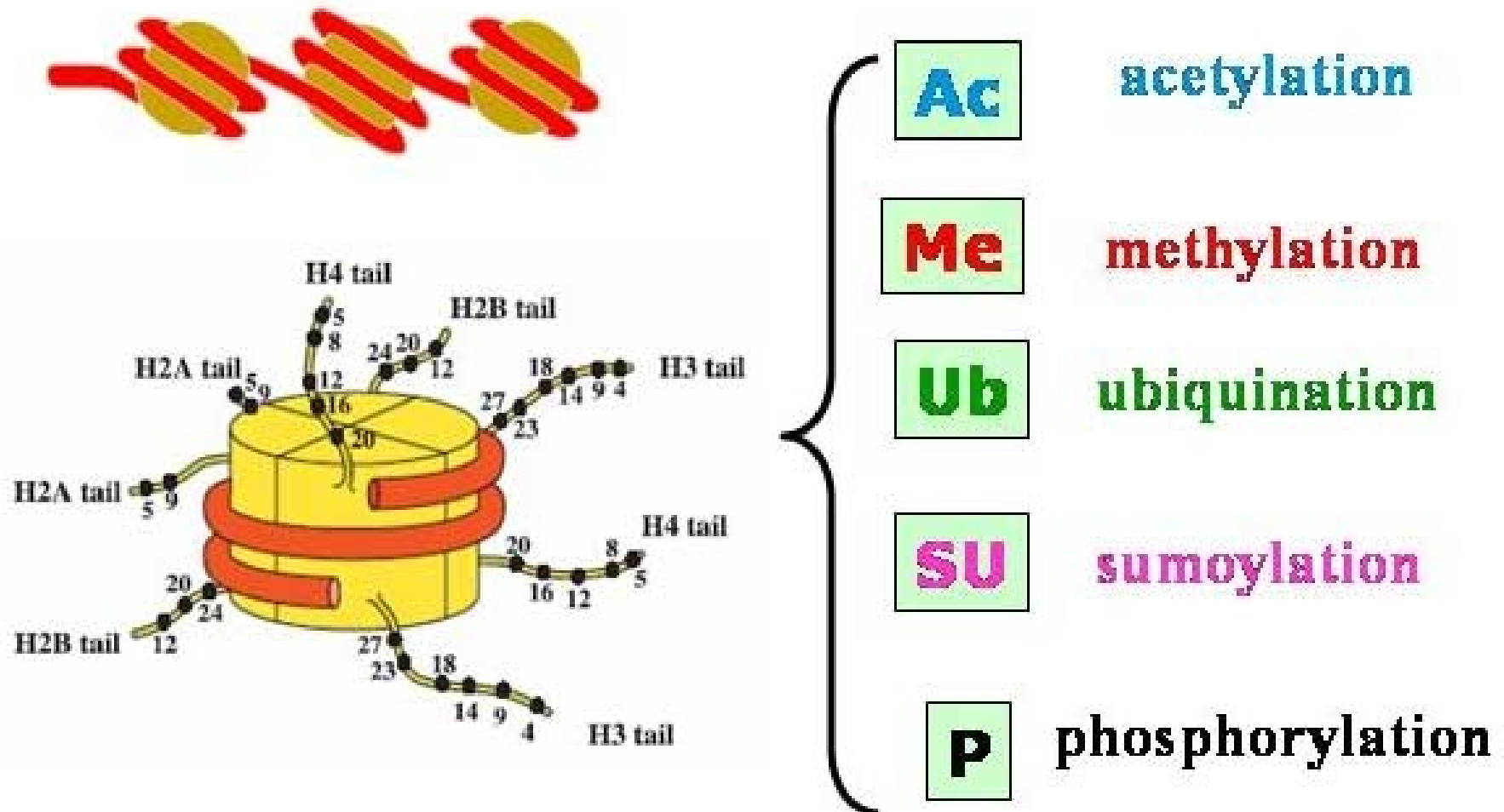
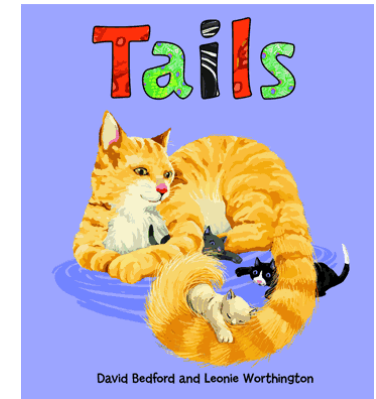


Histone Variant: H2A.Z

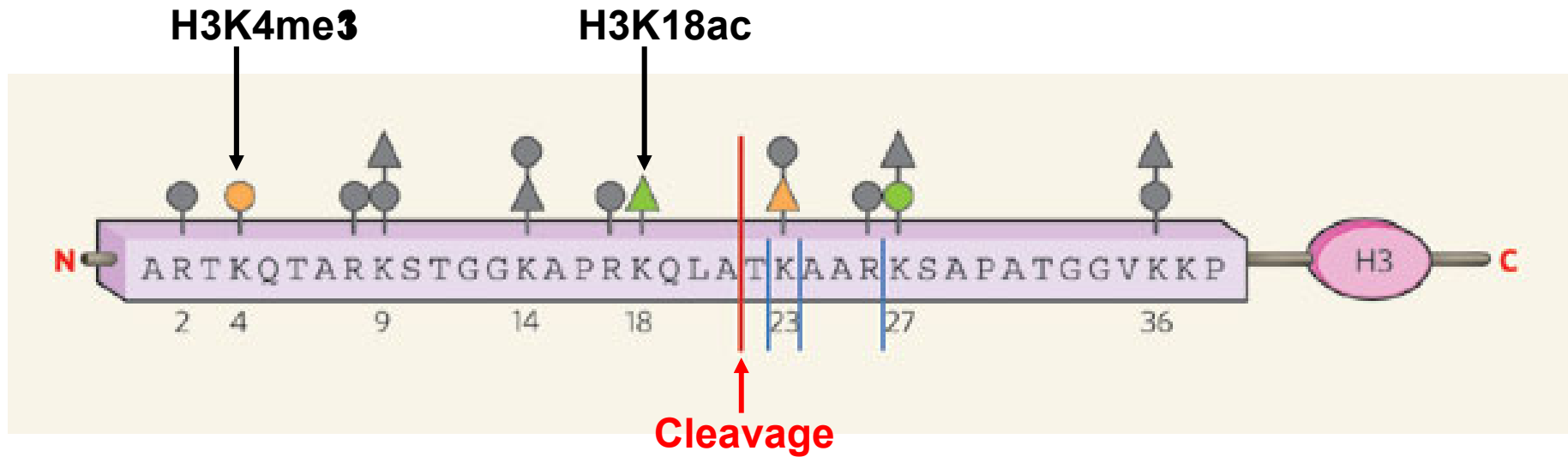


ACTIVE regions (promoters, enhancers) are thought to have histone H2A.Z

Histones have tails



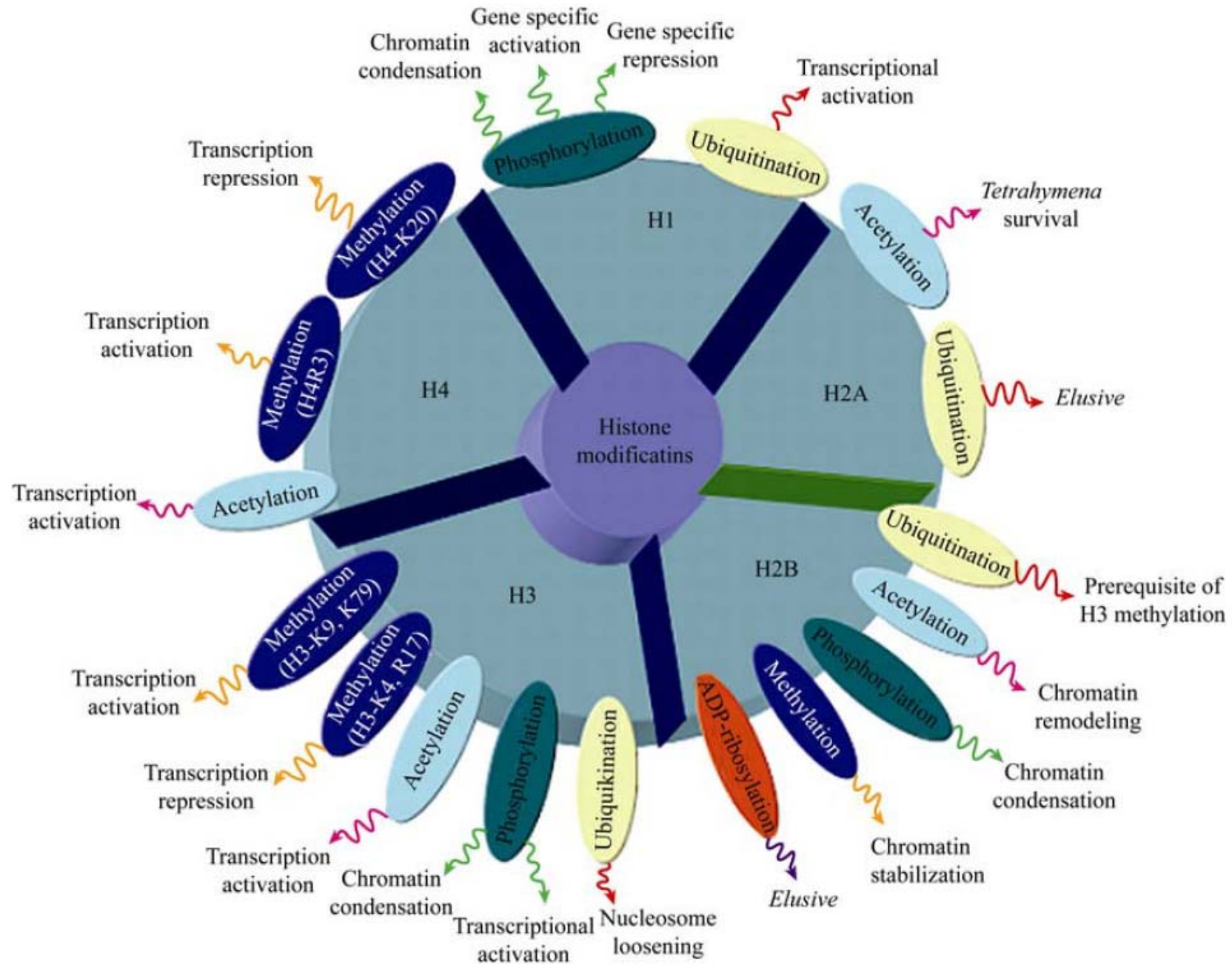
Histone H3: Example



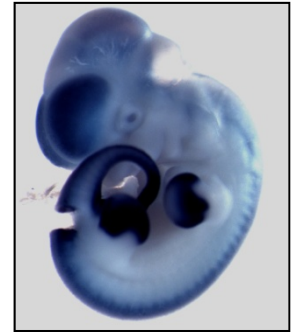
- | | |
|----------------------|----------------------|
| <u>● Methylation</u> | <u>▲ Acetylation</u> |
| ● Unknown | ▲ Unknown |
| ● Repression | ▲ Repression |
| ● Activation | ▲ Activation |

K=Lysine: Methylated or Acetylated

R=Arginine: Methylated



ChIP targets for expressed genes



Promoter A

Gene A

H3K36me3

ChIP targets for active promoters



Promoter A

Gene A

RNAPII

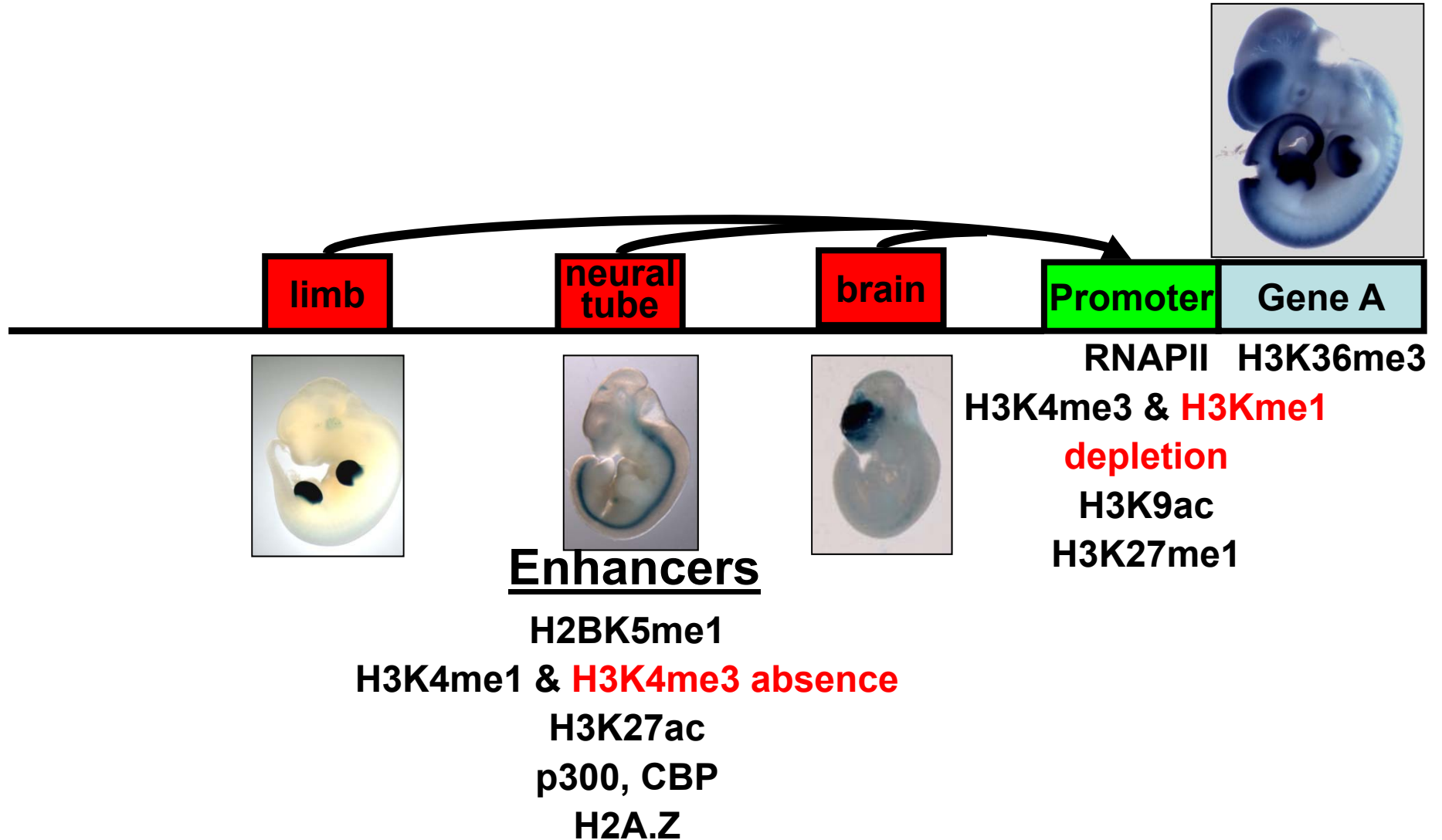
H3K36me3

H3K4me3 & **H3K4me1**
depletion

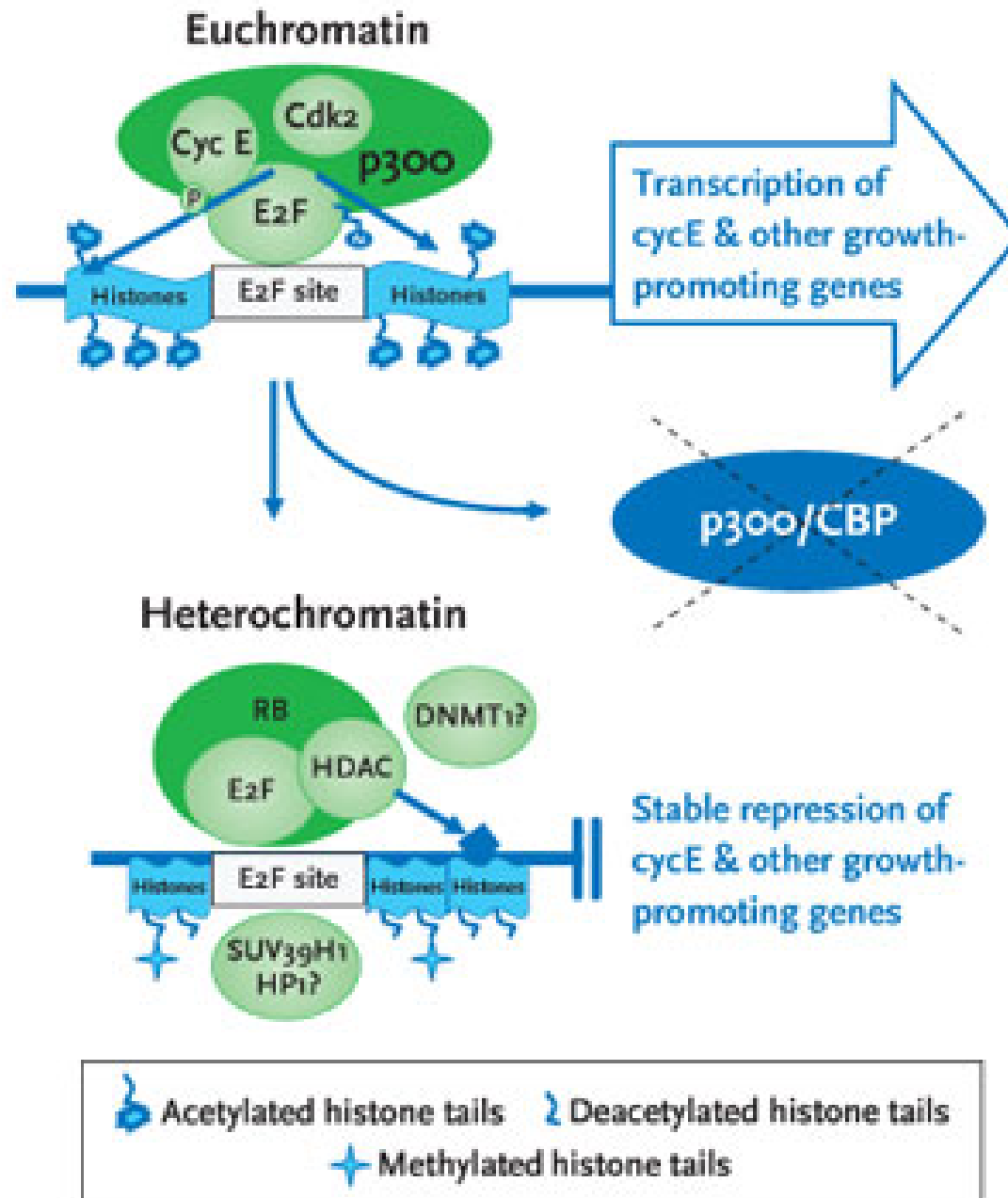
H3K9ac

H3K27me1

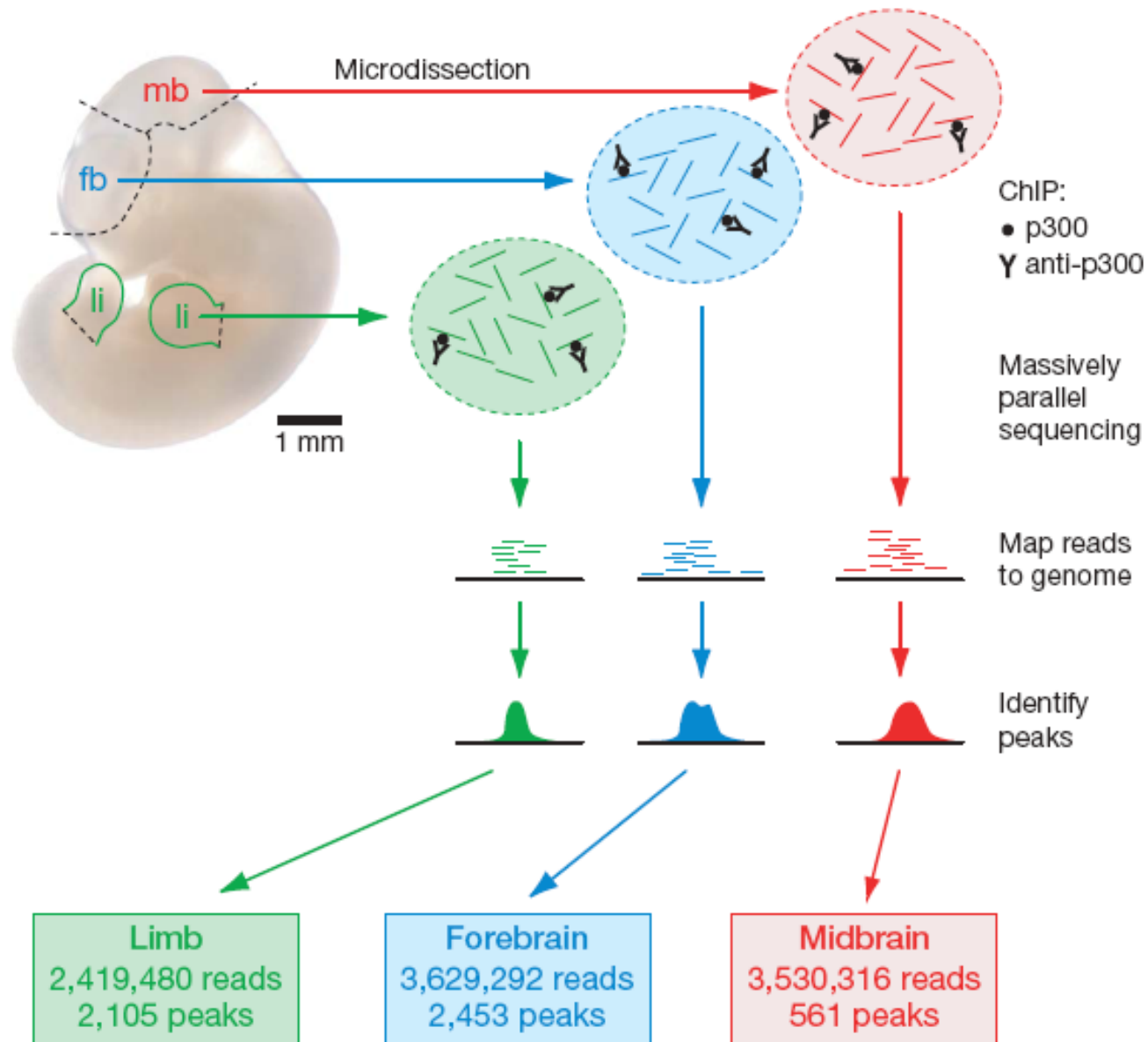
ChIP targets for active enhancers



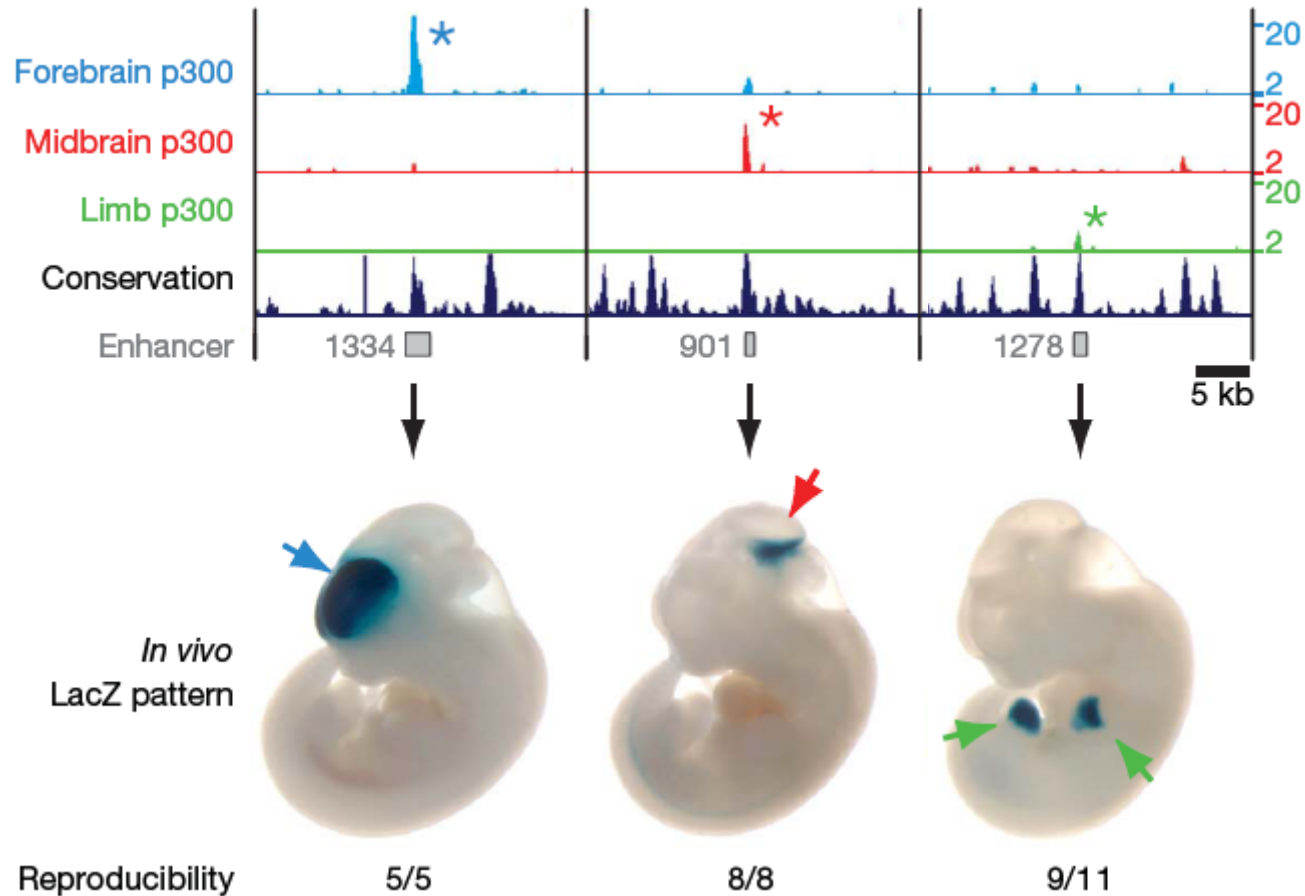
Acetylation: p300



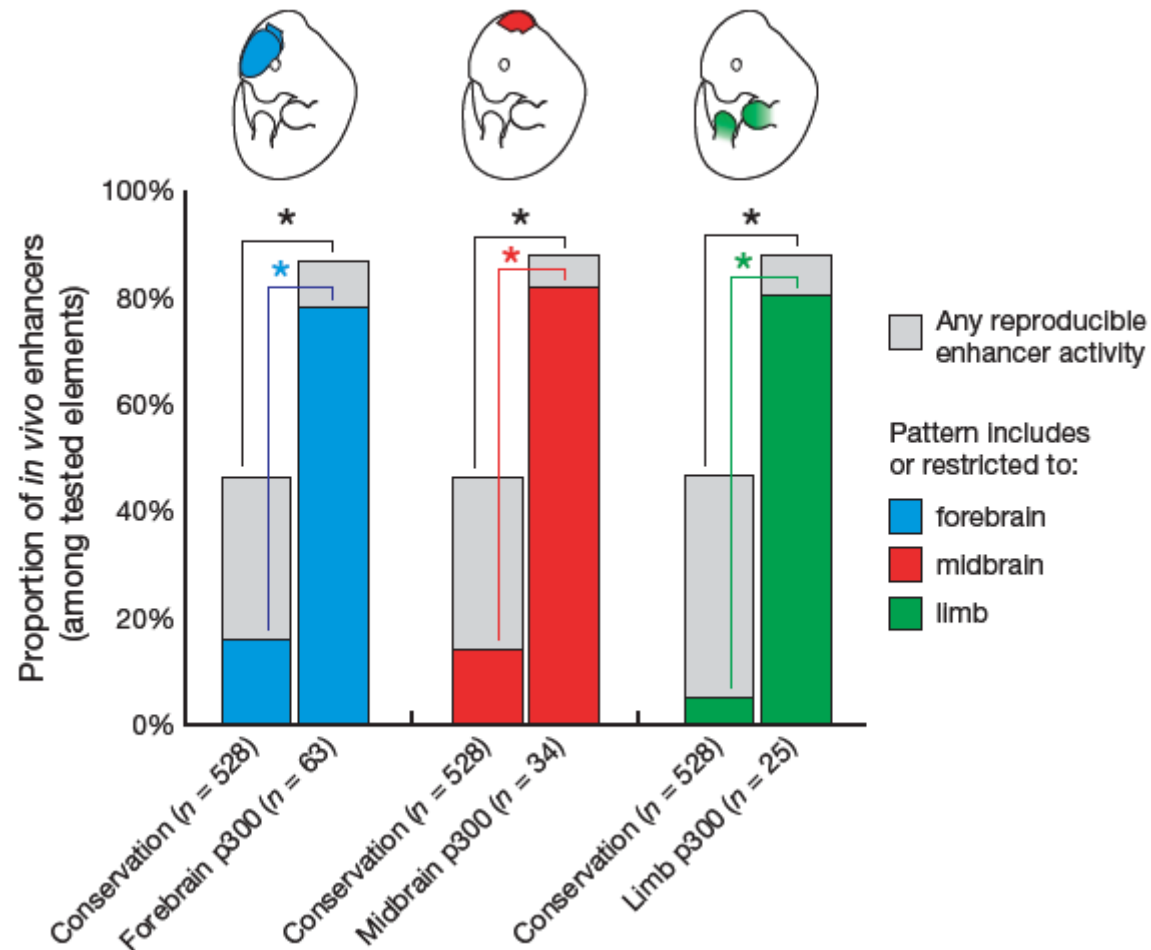
p300 is a good predictor of enhancers



p300 is a good predictor of tissue specificity



p300 is a good predictor of enhancers & tissue specificity



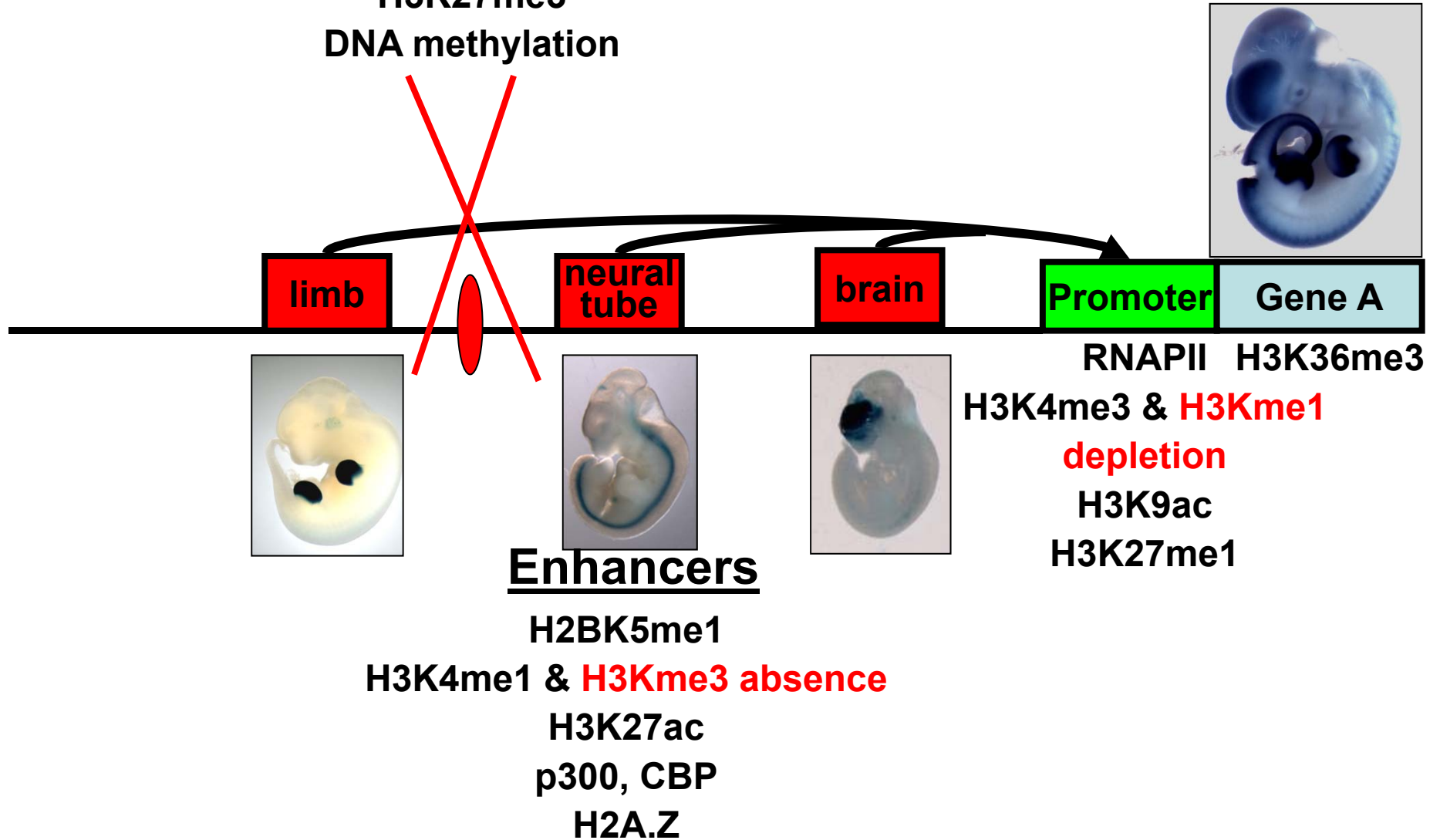
ChIP targets for silencers

Silencers

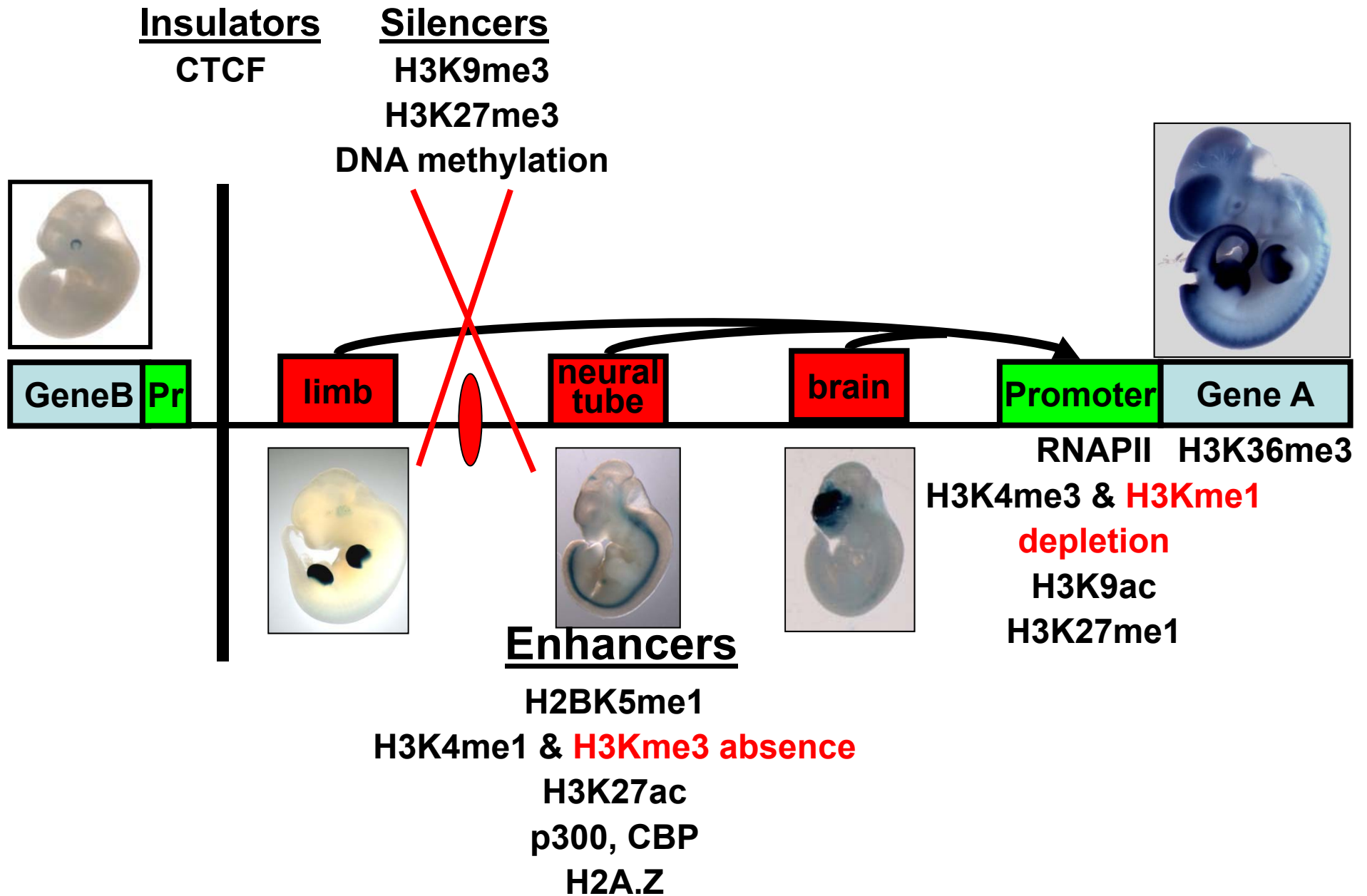
H3K9me3

H3K27me3

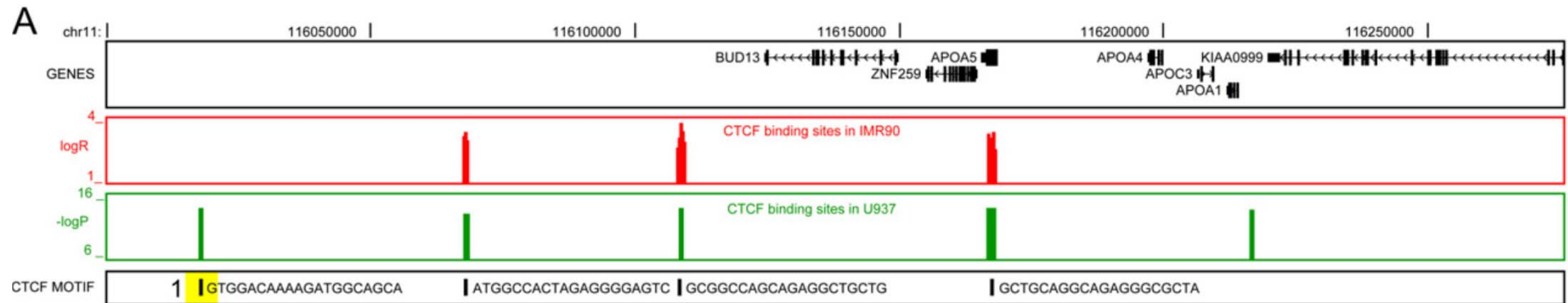
DNA methylation



ChIP targets for insulators



CTCF ChIP-seq

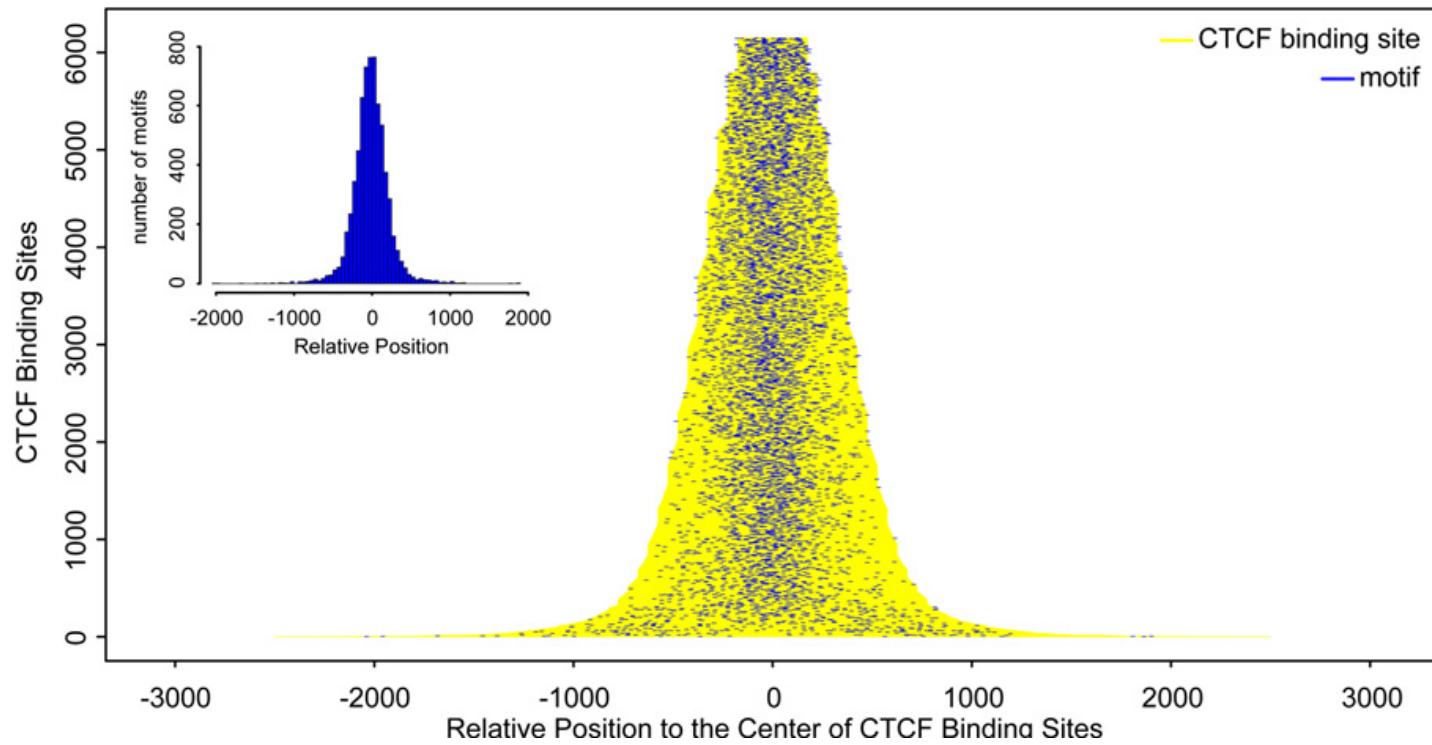


http://bioinformatics-renlab.ucsd.edu/retrac/wiki/CTCF_Project

A



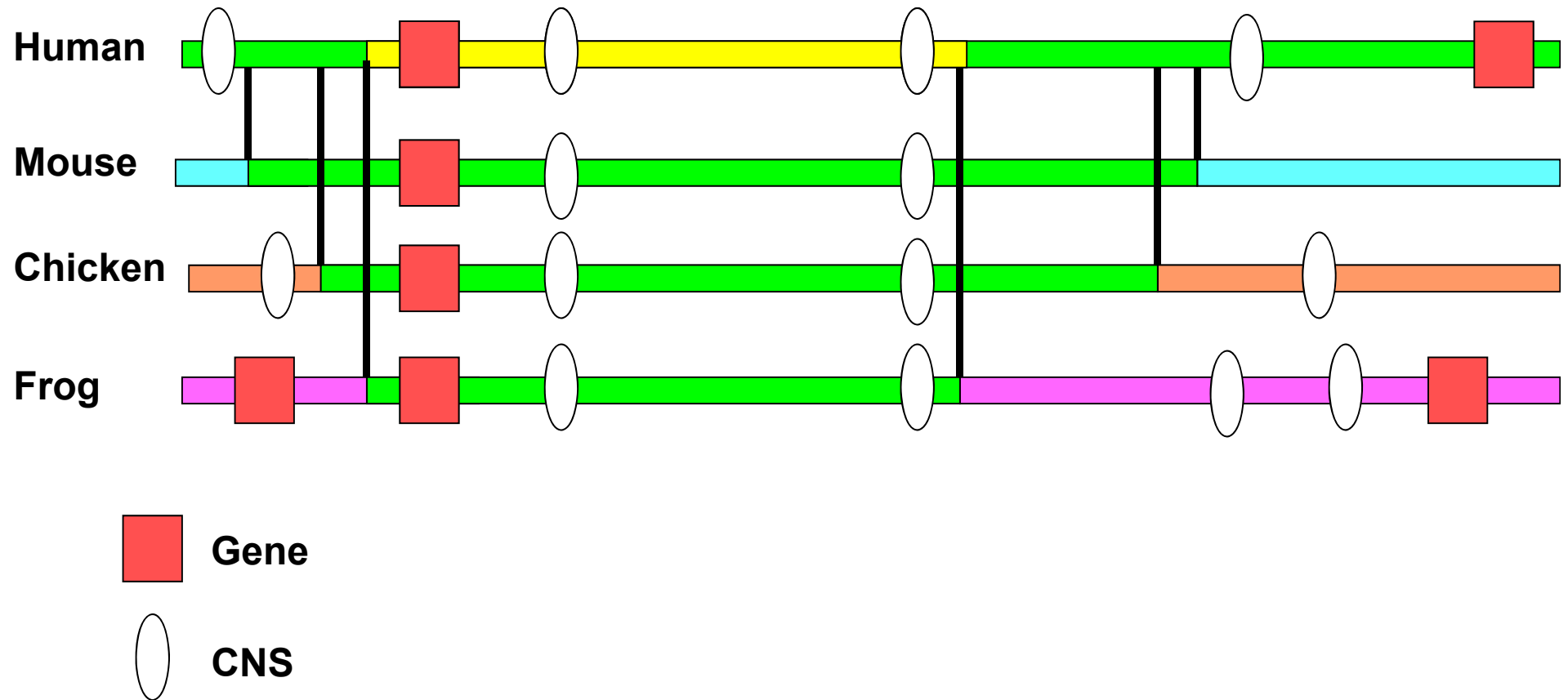
B



OUTLINE

- 1. The different kinds of gene regulatory elements**
- 2. How can we find them**
- 3. How far away to look for them**
- 4. Functional assays to test their function**
- 5. Regulation in 3D**
- 6. Example from our lab**

Synteny blocks can define cis-regulatory domains



Synteny blocks can assign cis-regulatory borders

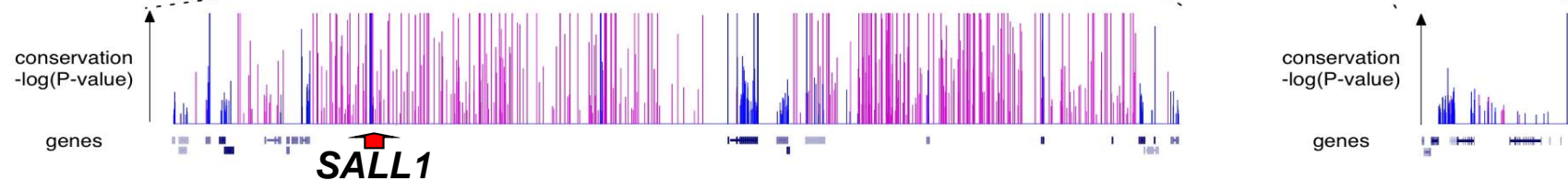
a. Human-mouse-chicken synteny



b. CNS density



c.



	HMC	HMF
Number of segments	2116	1942
Length on human (Gb)	1.53	0.86
N50 length (Mb)	1.02	0.48
Maximum length (Mb)	5.68	2.93

Fixed distance



Neighboring genes

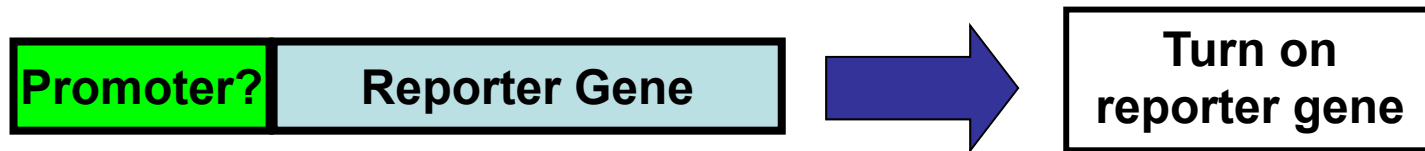


OUTLINE

- 1. The different kinds of gene regulatory elements**
- 2. How can we find them**
- 3. How far away to look for them**
- 4. Functional assays to test their function**
- 5. Regulation in 3D**
- 6. Example from our lab**

Promoter & Enhancer Assays

Promoter



Enhancer



Cell Culture

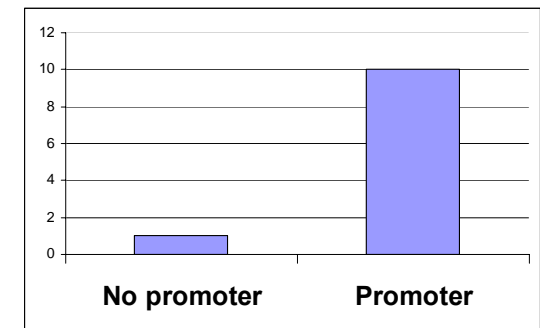


Advantages:

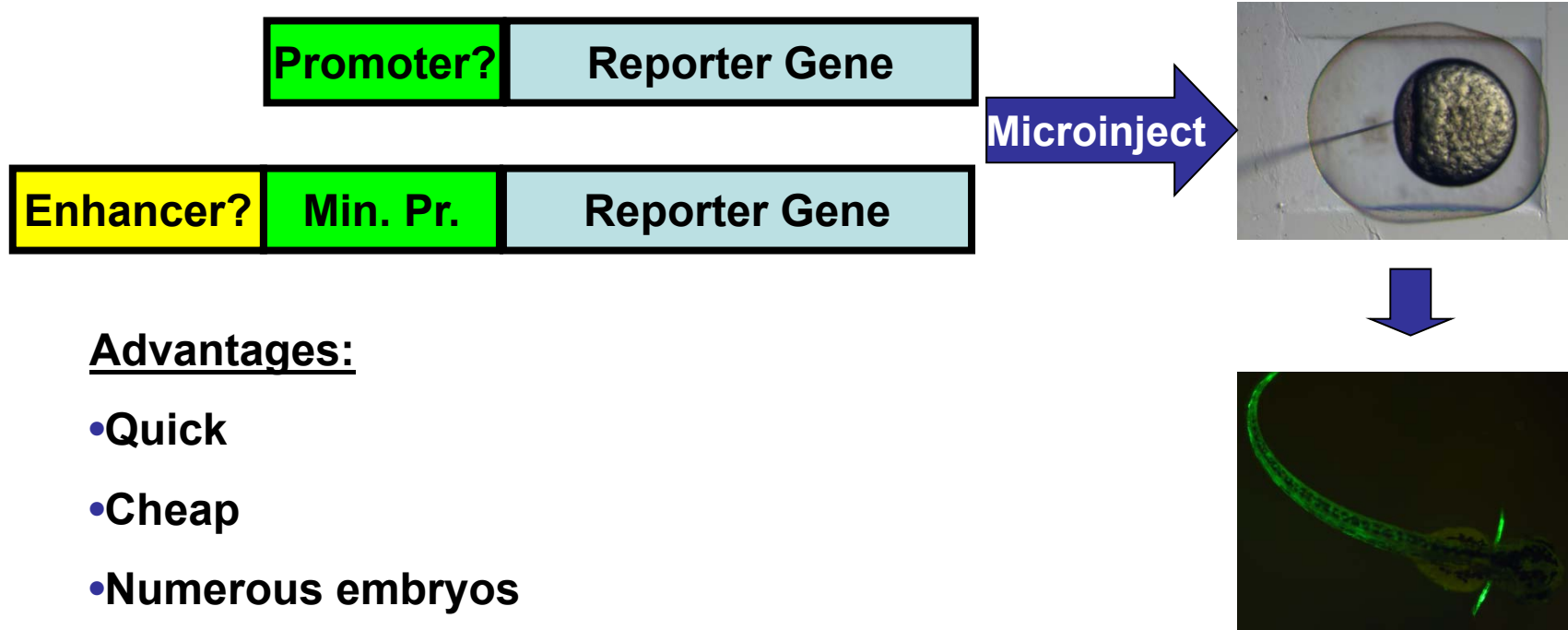
- Quick
- Quantitative
- Cheap

Disadvantages:

- Not *in vivo*, whole organism
- Cultured cells lose a lot of their tissue characteristics
- A lot of variability (different prep, cell passage etc.)



Zebrafish Transgenics



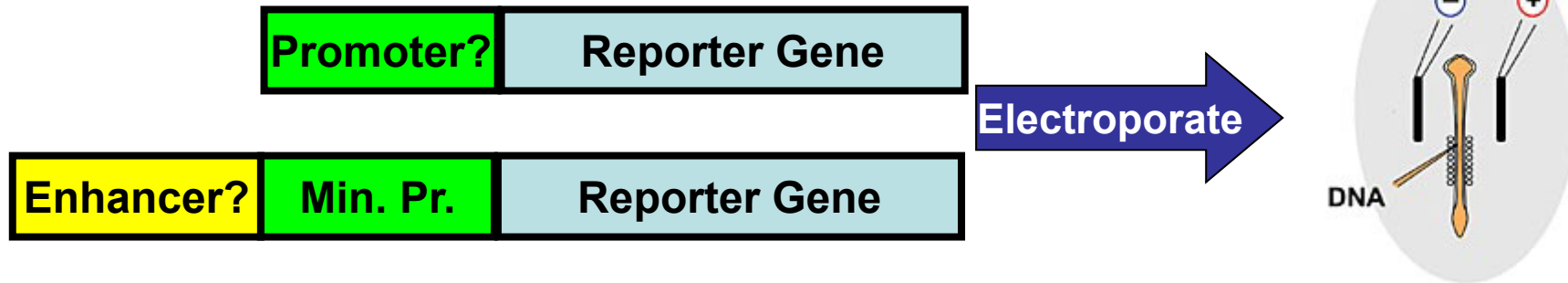
Advantages:

- Quick
- Cheap
- Numerous embryos
- Can track fish over time

Disadvantages:

- Mosaicism and average of 5 integrations can lead to position effects
- Distant to human

Chicken Electroporation



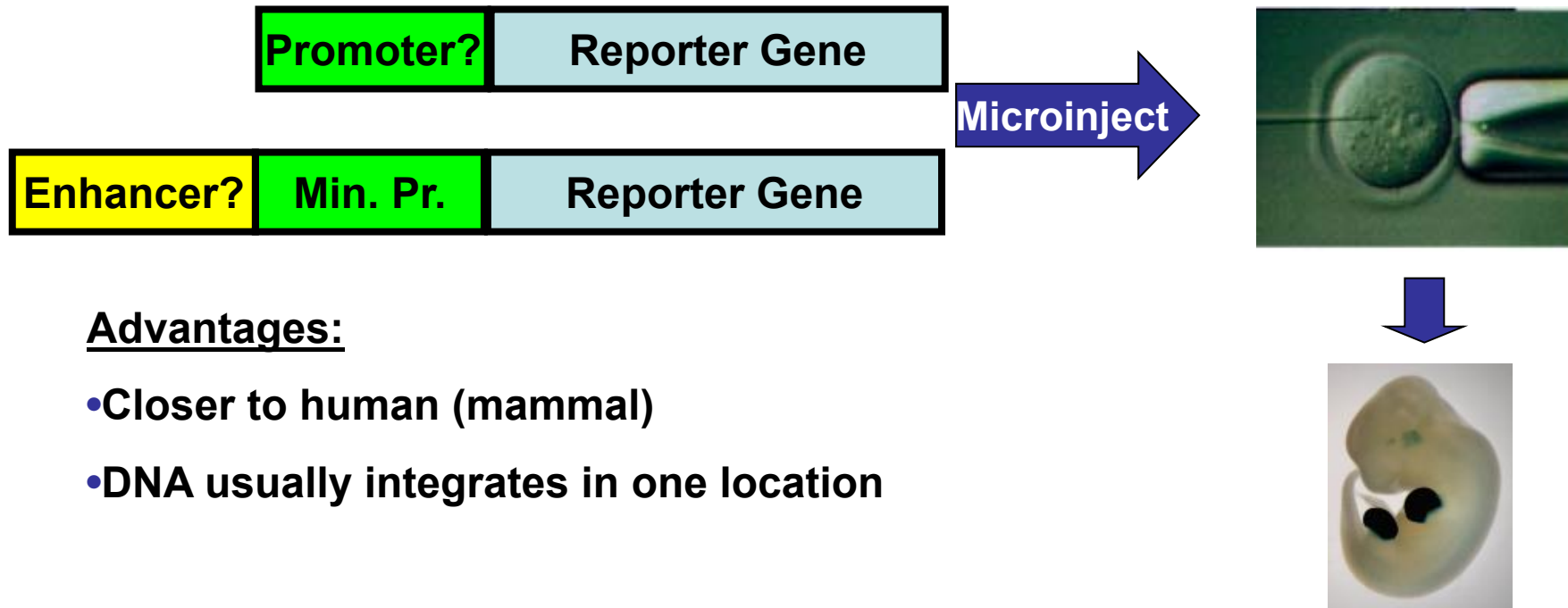
Advantages:

- Quick
- Cheap

Disadvantages:

- Only accessible regions can be electroporated
- Not expressed in all the cells
- Variable expression due to the DNA not integrating into the genome
- Have to take out embryos for each time point
- Distant to humans

Mouse Transgenics



Advantages:

- Closer to human (mammal)
- DNA usually integrates in one location

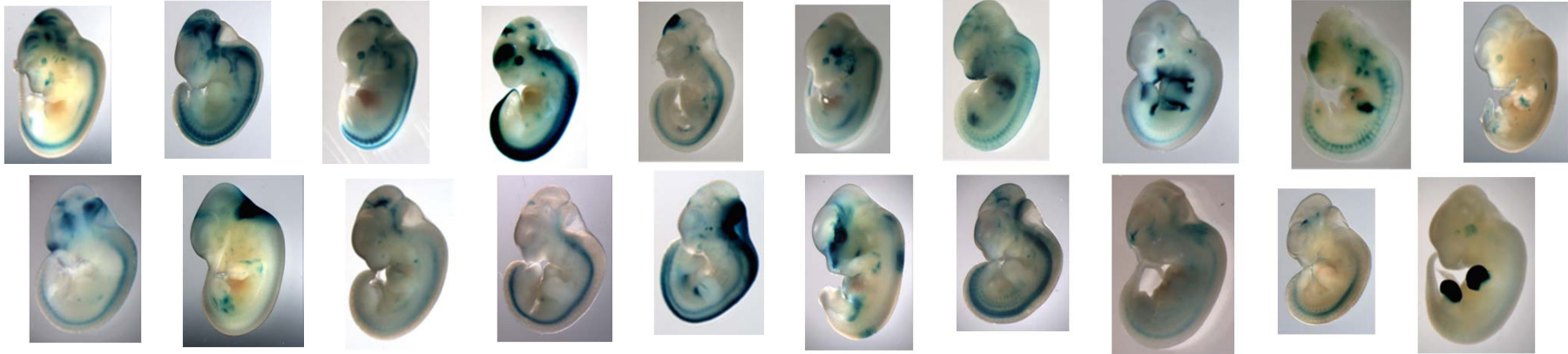
Disadvantages:

- Not high-throughput
- Expensive
- Embryos need to be taken out for each time point

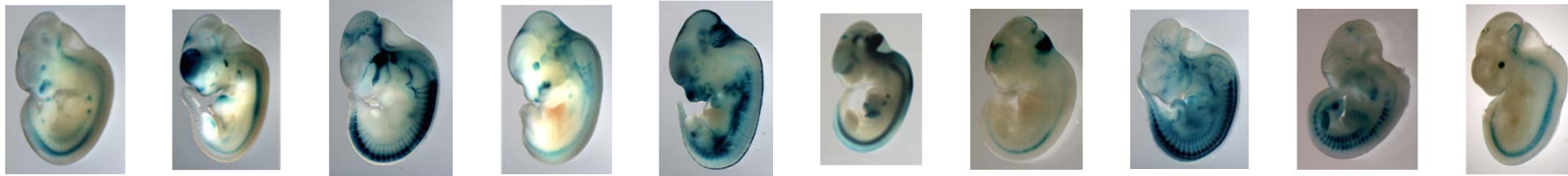


VISTA Enhancer Browser

whole genome enhancer browser



<http://enhancer.lbl.gov>



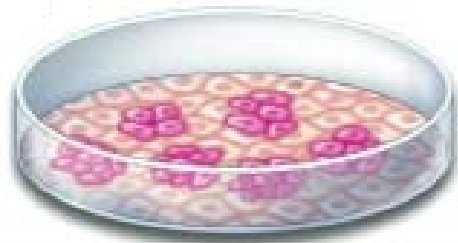
746 human enhancers out of 1,503 tested

Silencer & Insulator Assays

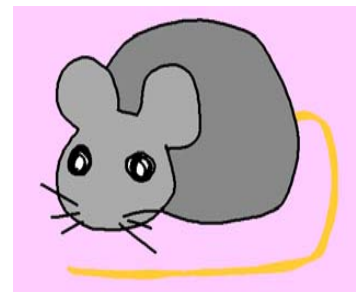
Silencer



Insulator



Cell Culture



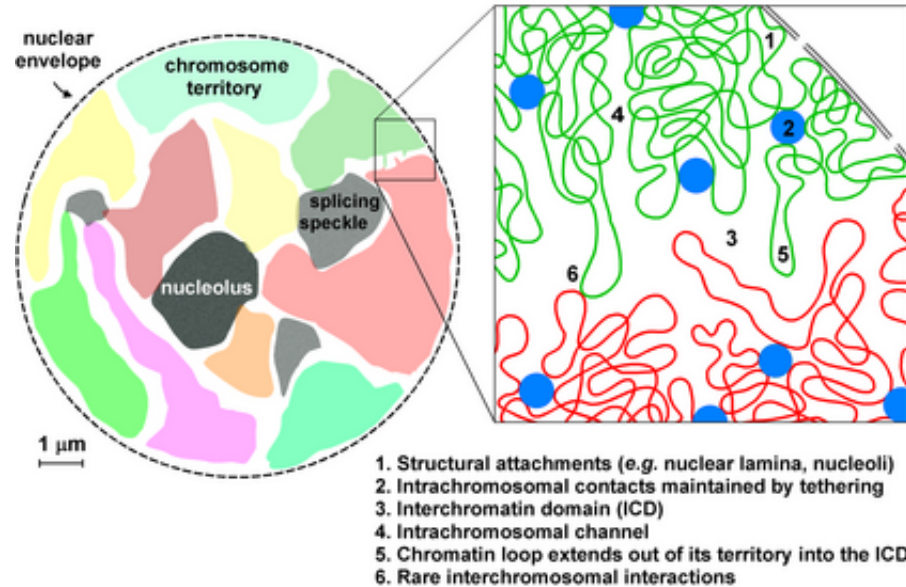
Mice

OUTLINE

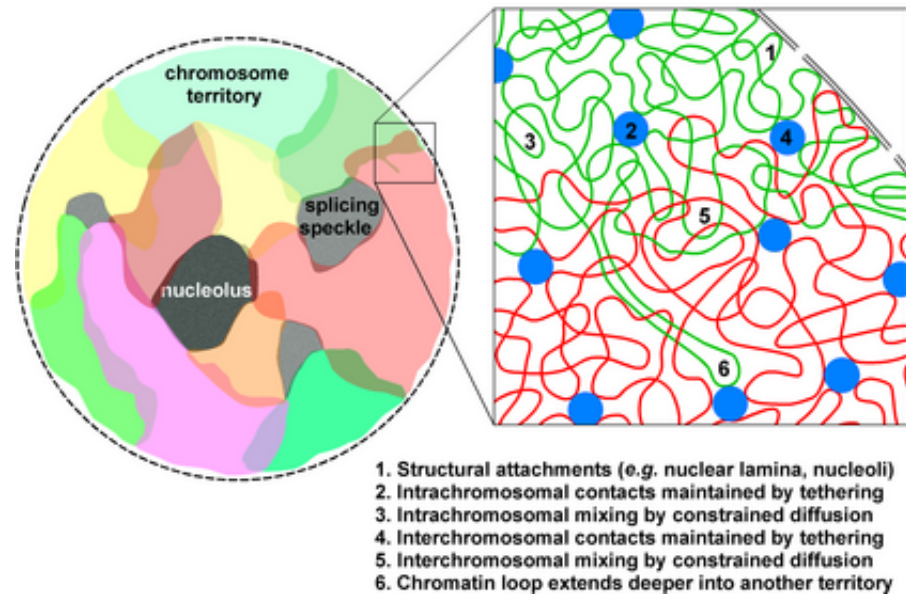
- 1. The different kinds of gene regulatory elements**
- 2. How can we find them**
- 3. How far away to look for them**
- 4. Functional assays to test their function**
- 5. Regulation in 3D**
- 6. Example from our lab**

Chromosomal interactions in the nucleus

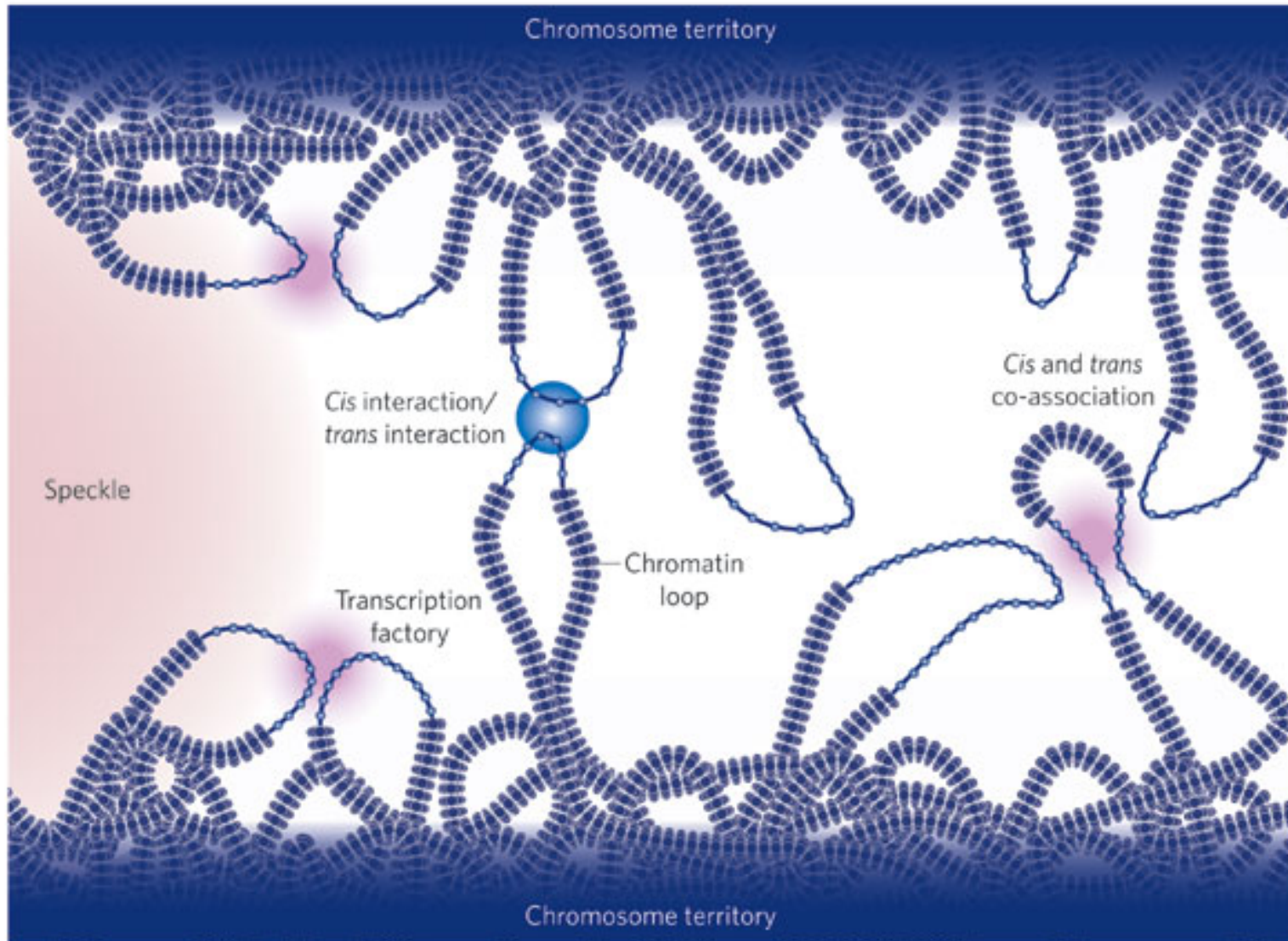
A. Interchromatin domain model



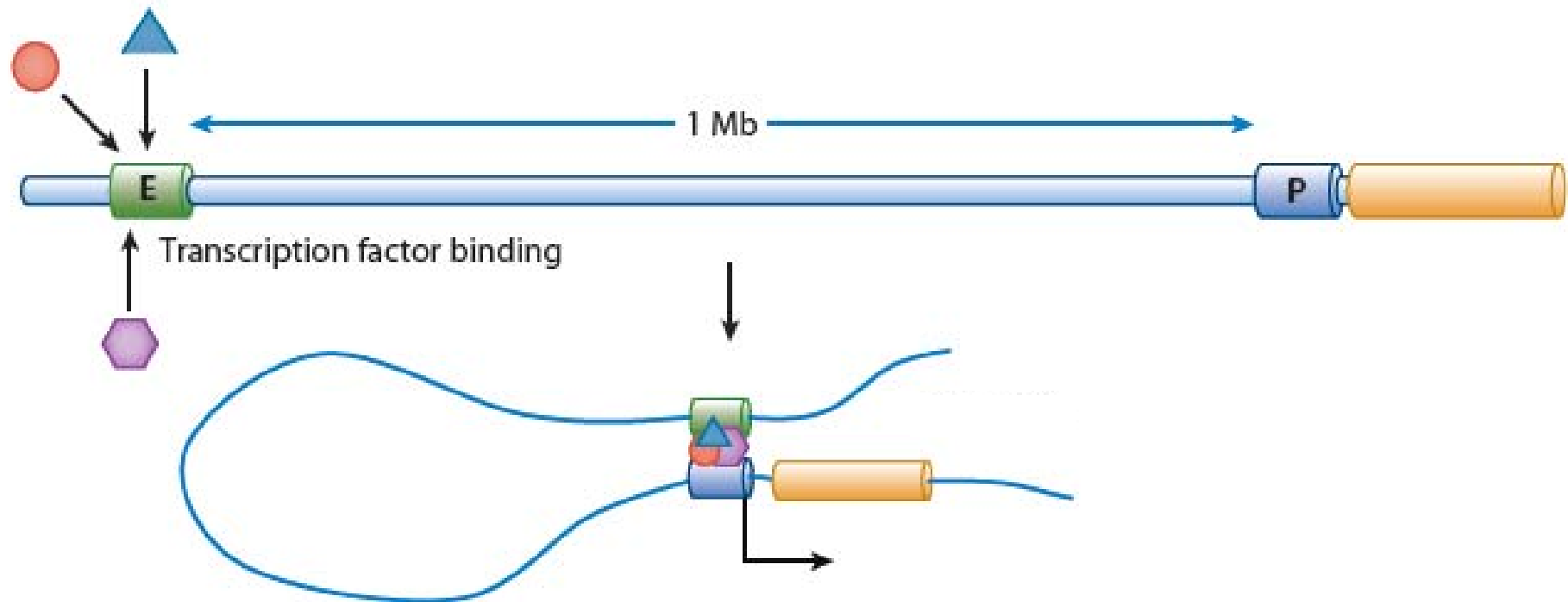
B. Interchromosomal network model



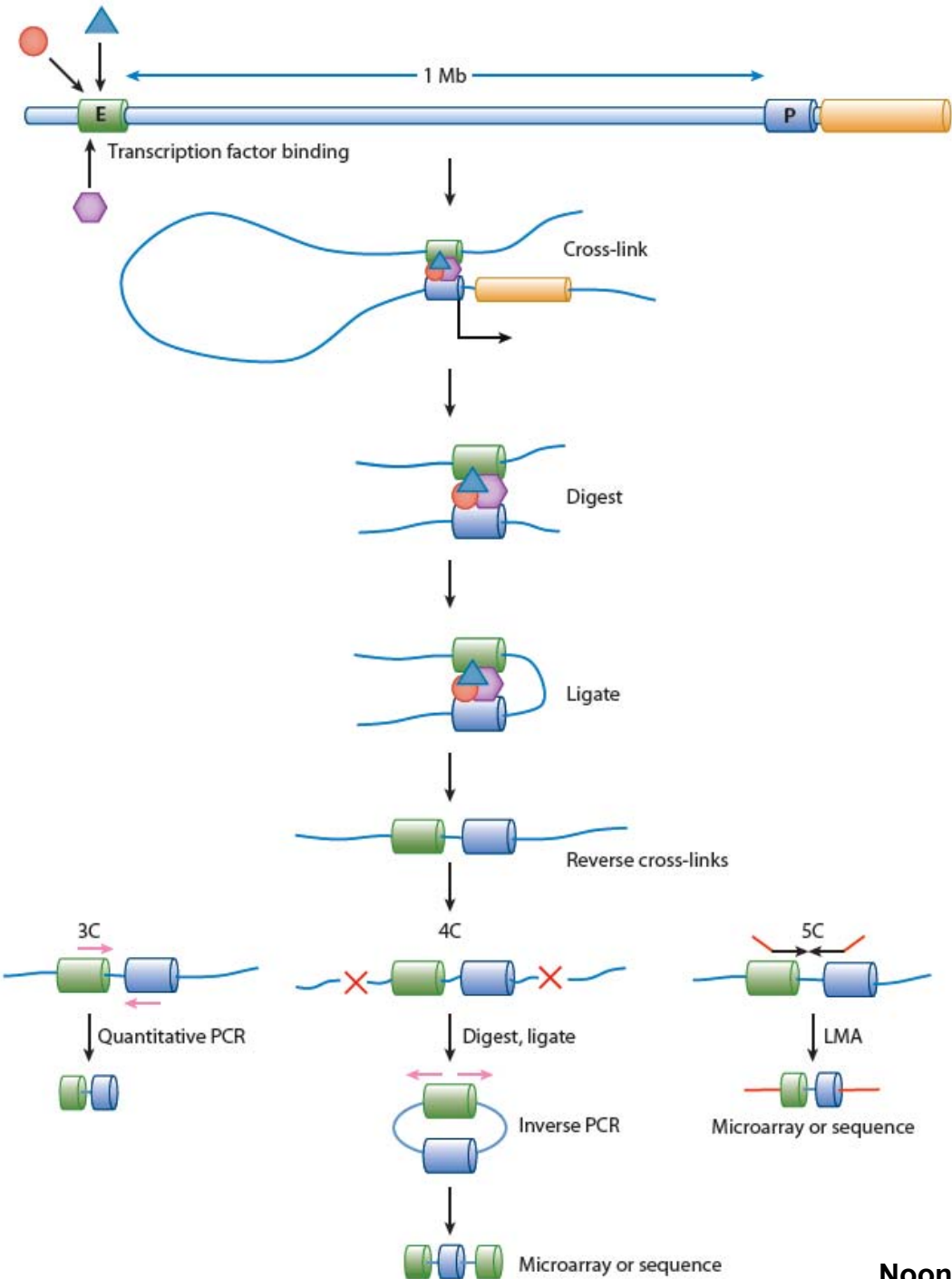
3D and gene regulation



Enhancers loop to bind to promoters

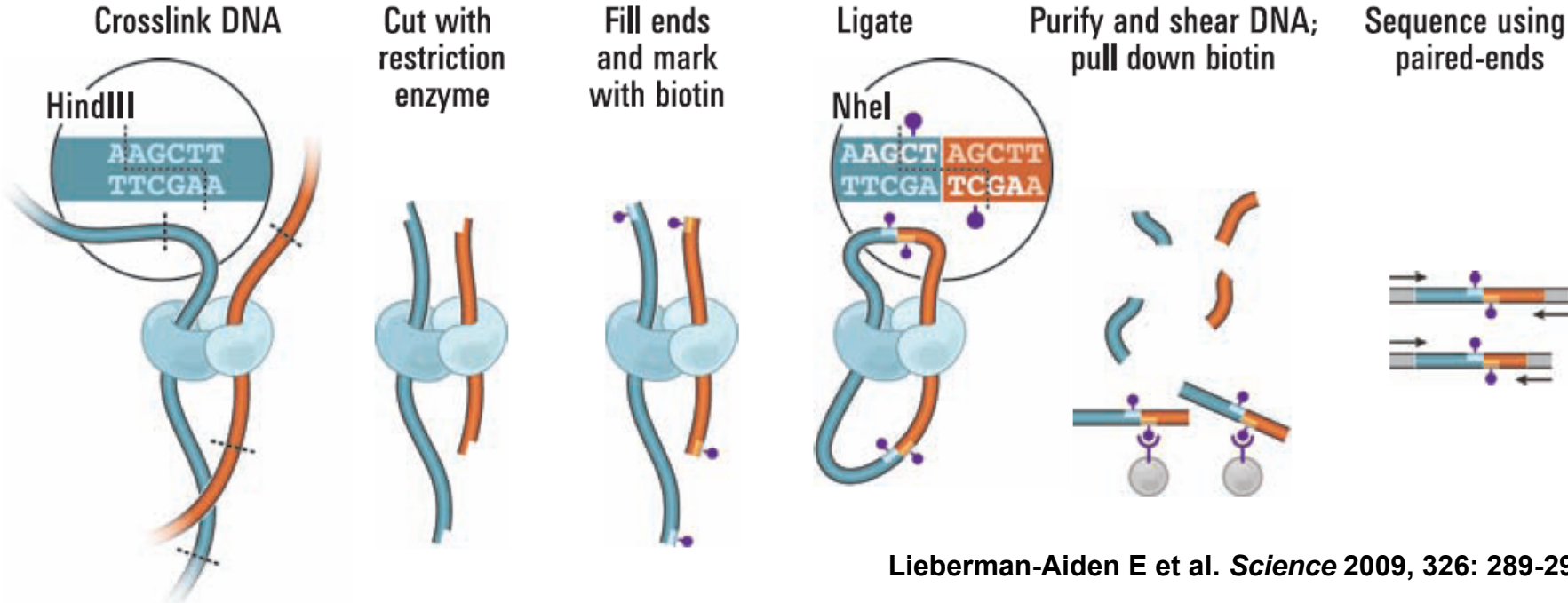


3C, 4C and 5C

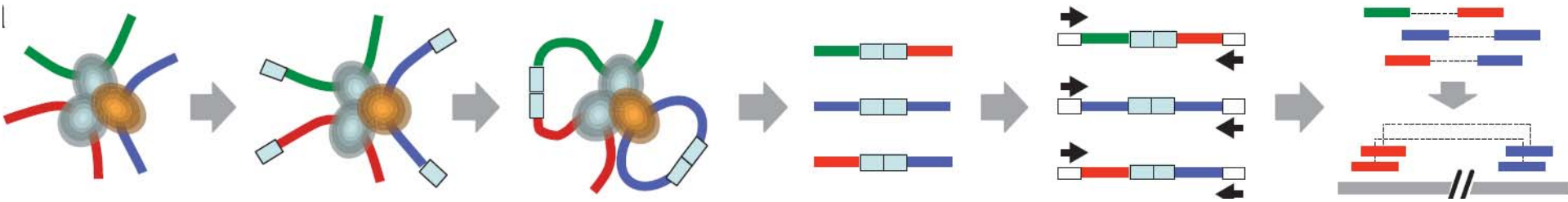


Hi-C & ChIA-PET

Hi-C

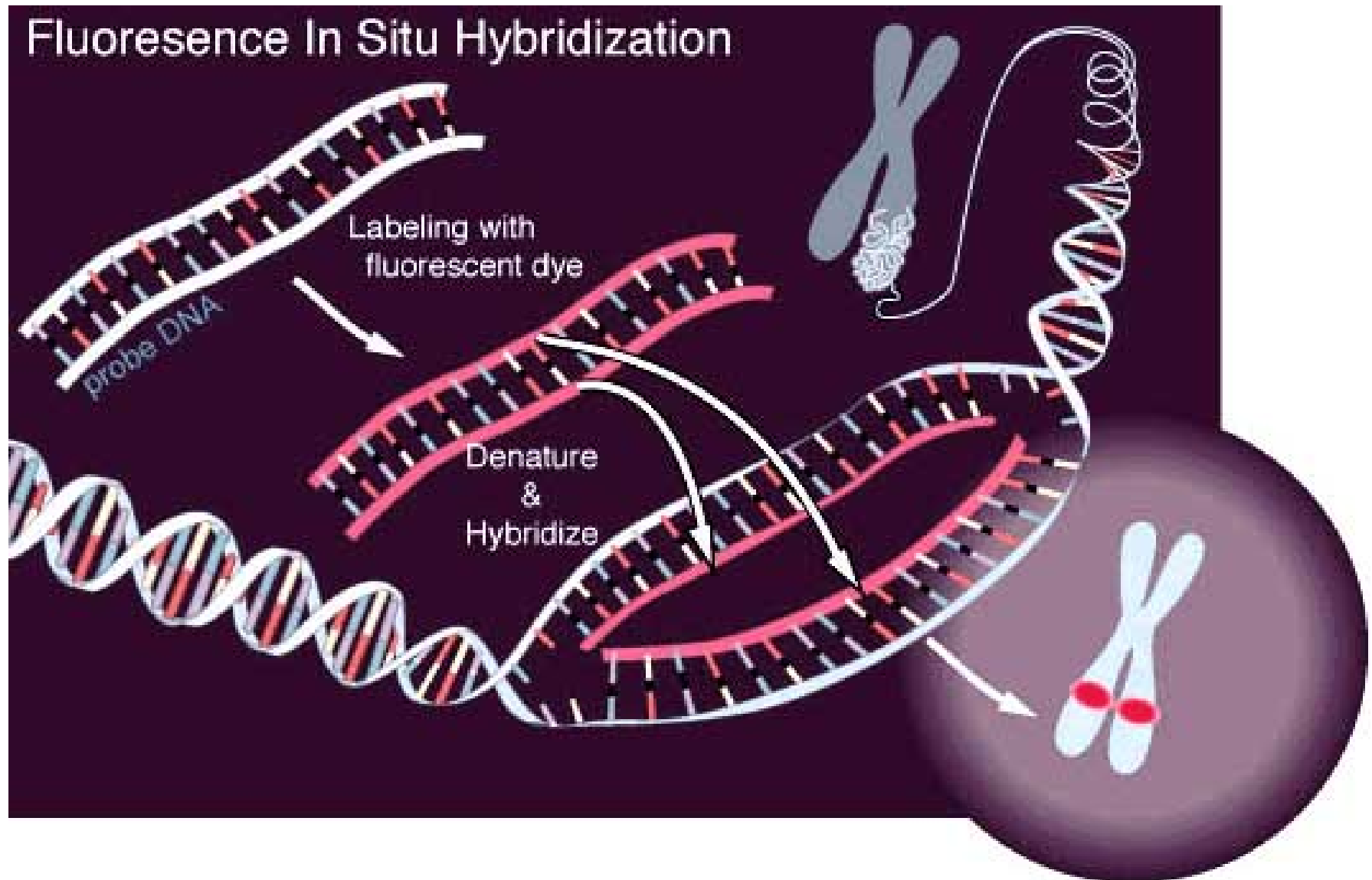


ChIA-PET

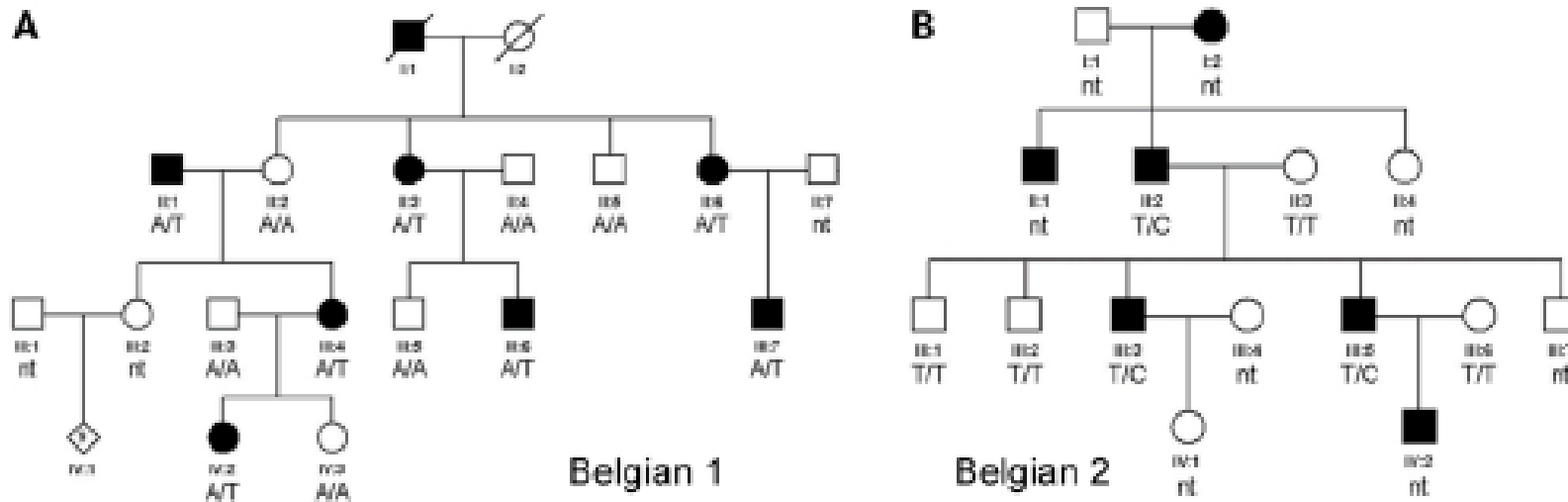
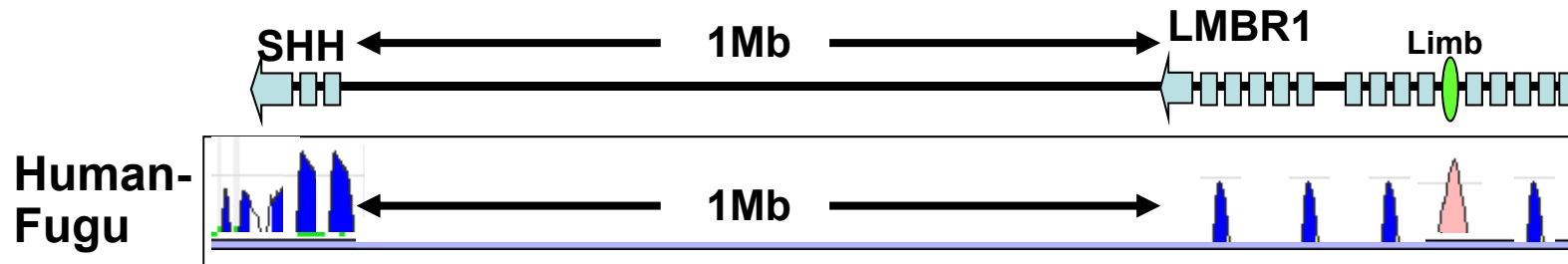


Fullwood MJ et al. *Nature* 2009, 462: 58-64

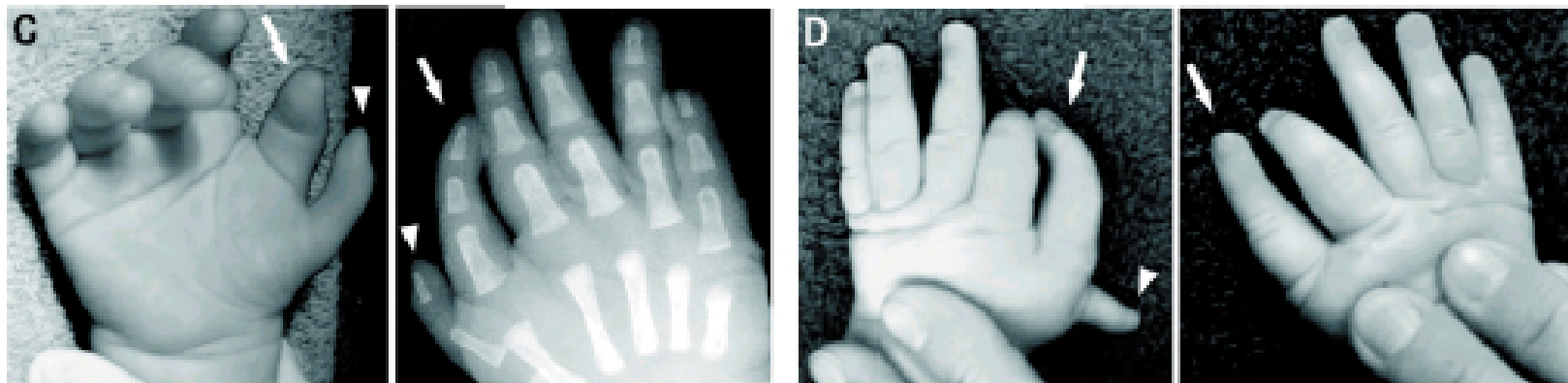
Fluorescence In Situ Hybridization (FISH)



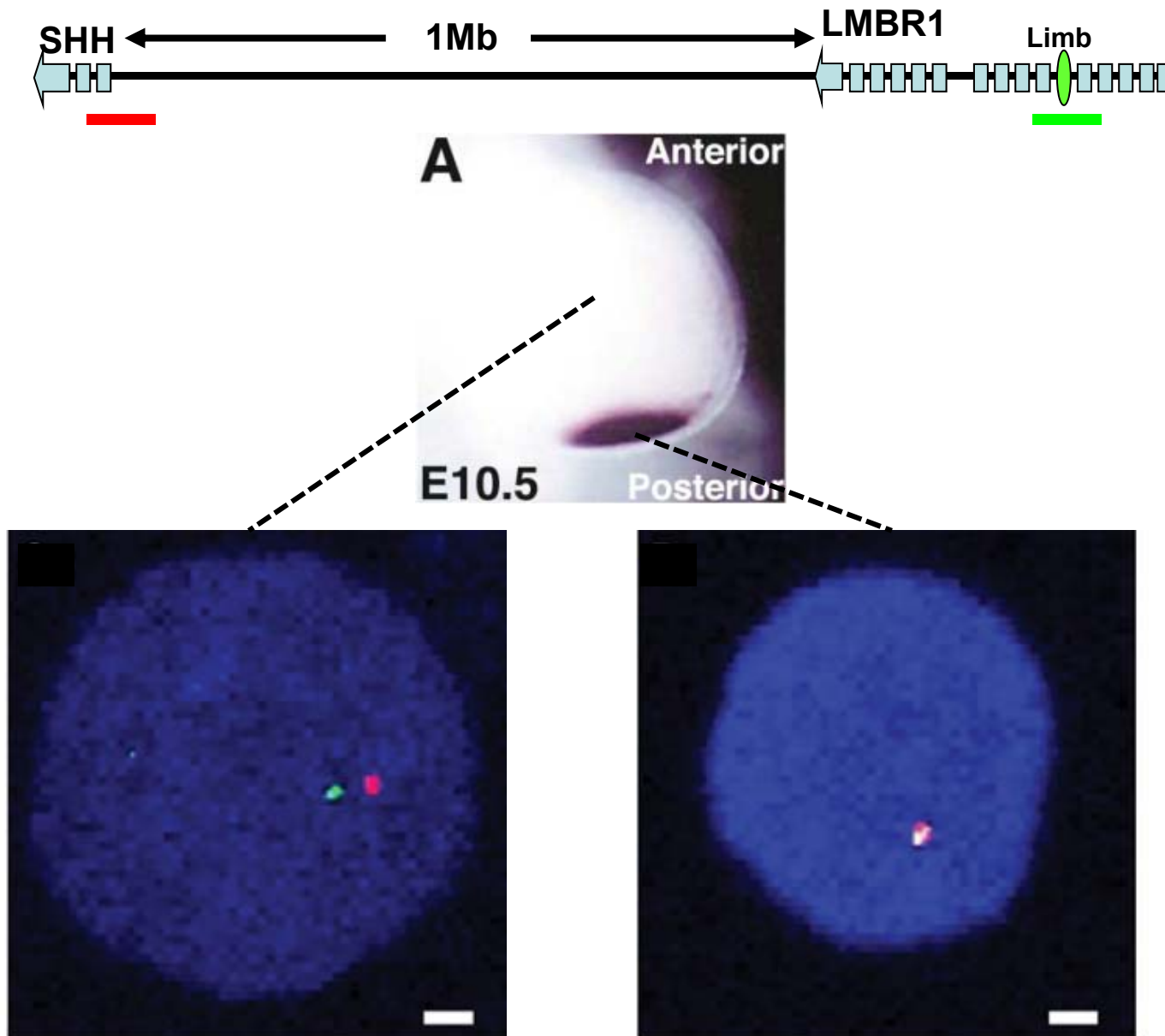
The *SHH* limb enhancer as an example



Lettice et al. *HMG* 2003 12: 1725-35



FISH can detect long range chromosomal interactions



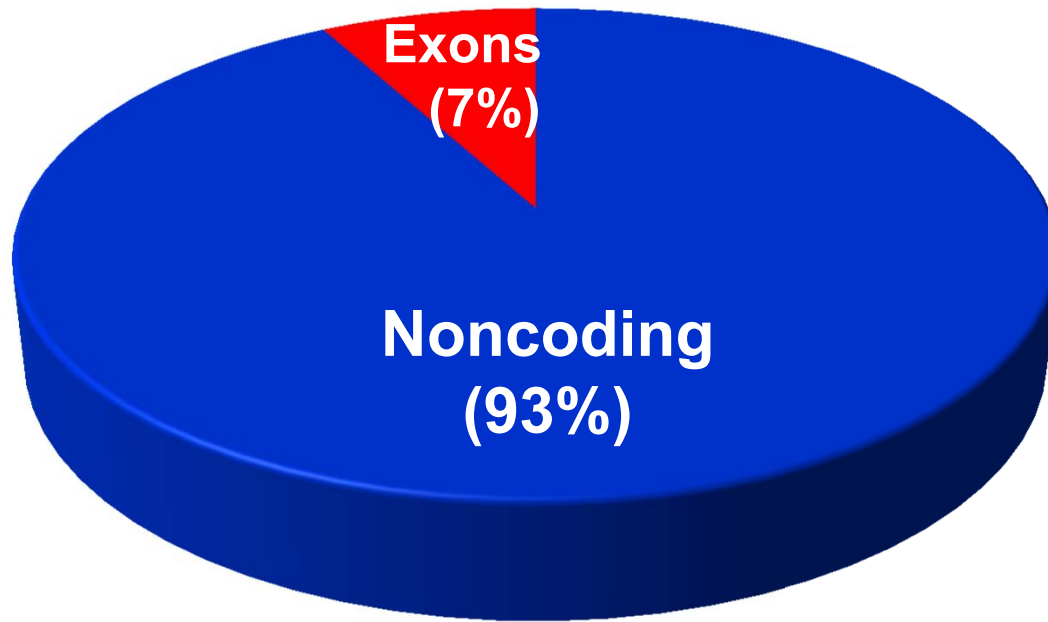
OUTLINE

- 1. The different kinds of gene regulatory elements**
- 2. How can we find them**
- 3. How far away to look for them**
- 4. Functional assays to test their function**
- 5. Regulation in 3D**
- 6. Example from our lab**



Ramon Y. Birnbaum

Selection of potential exonic enhancers of nearby genes



7 candidates near important limb genes

- 2 near *Dlx5/6*
- 2 near *Twist1*
- 1 near *Dicer1*
- 1 near *Ptch1*
- 1 near *Msx1*

Coding exons (7%)



2,105 limb p300 ChIP-Seq peaks

152 overlap coding exons

Gene containing exon is expressed in the limb

Yes ~~X~~
18 exons

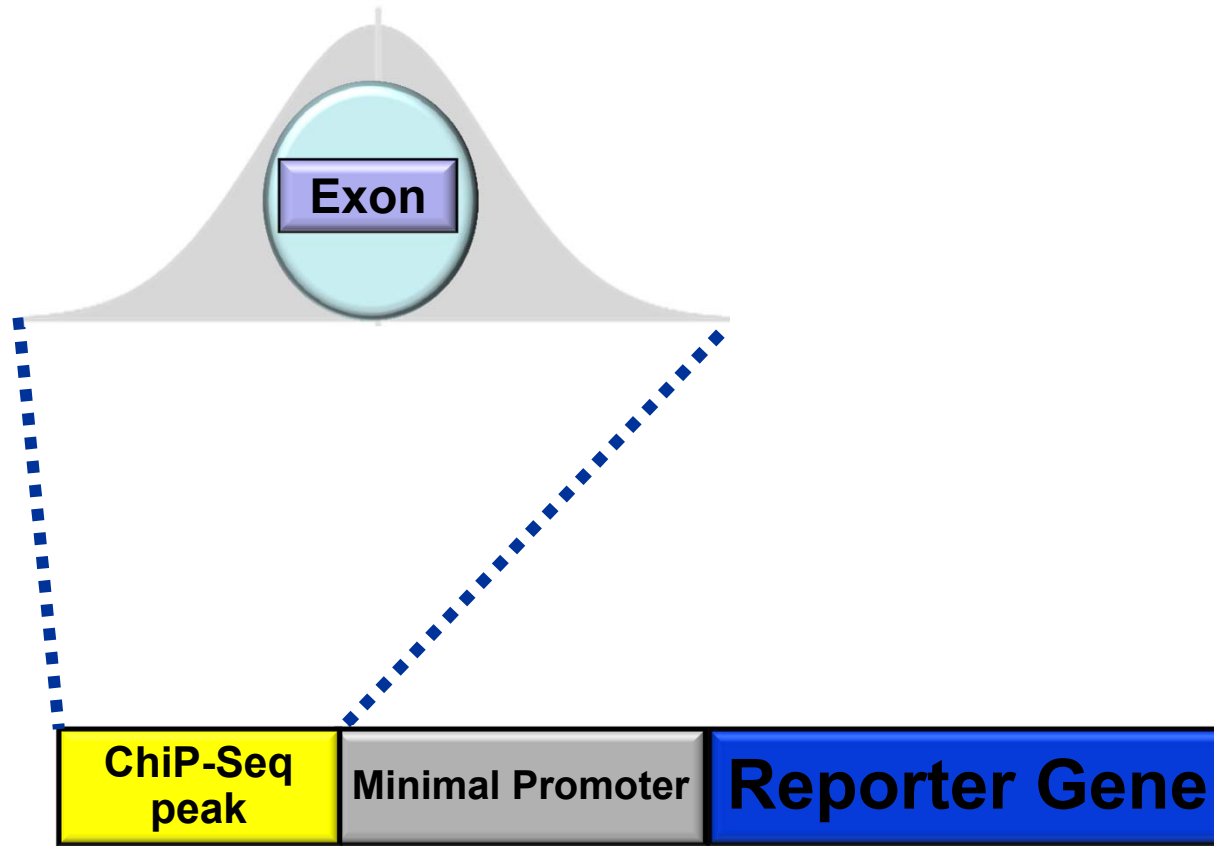
No [✓]
134 exons

Limb expressed genes nearby (<2 Mb)

No ~~X~~
44 exons

Yes [✓]
90 exons

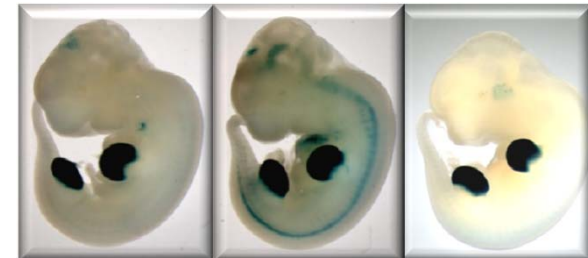
Enhancer Assay



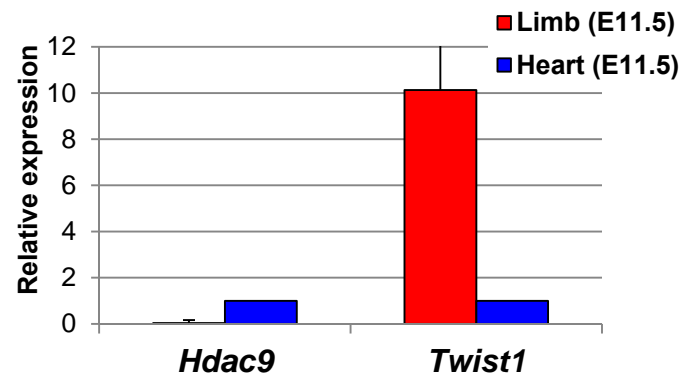
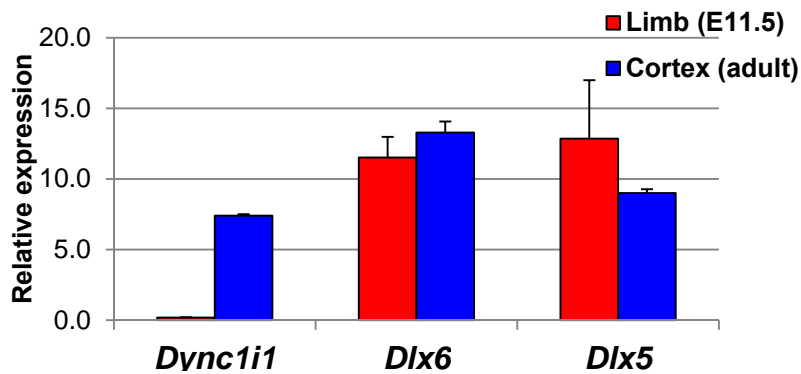
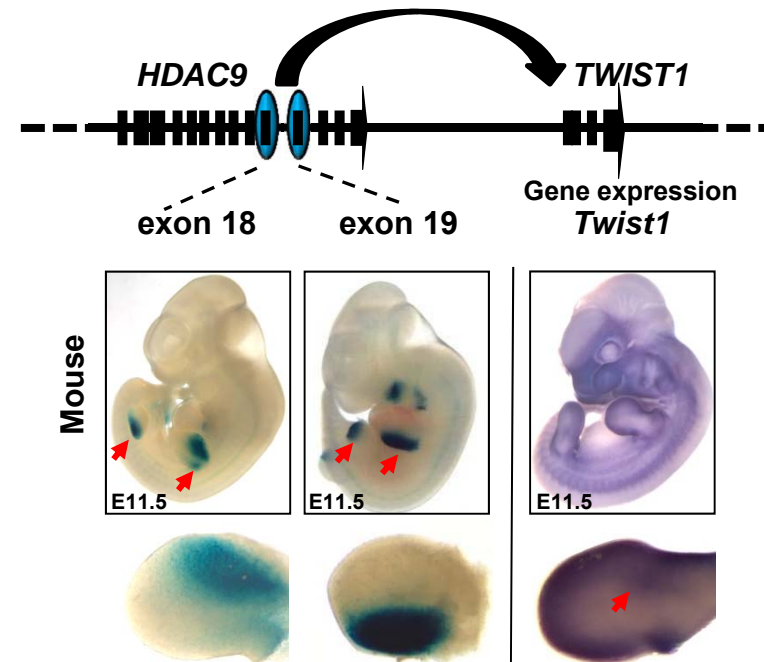
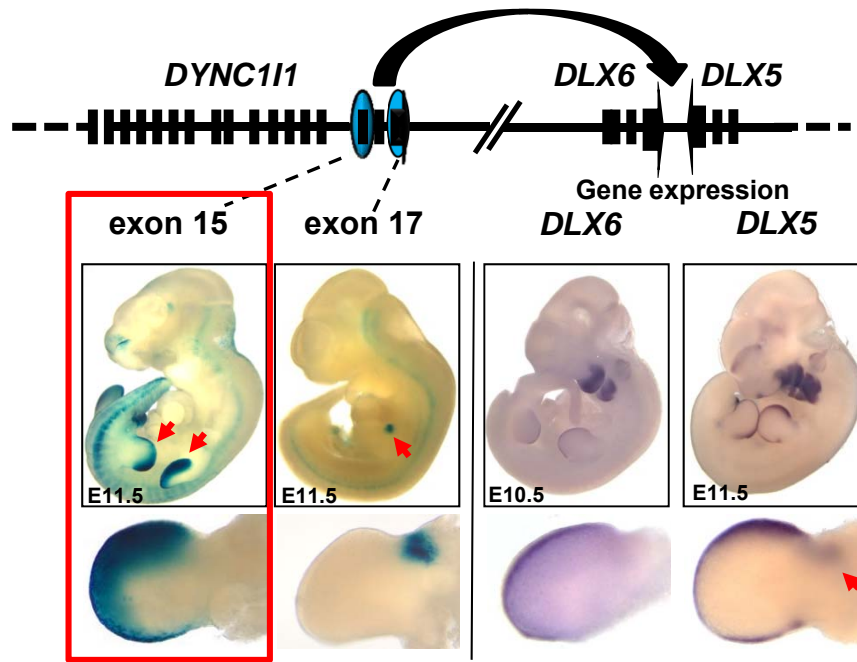
Injection into mouse fertilized eggs



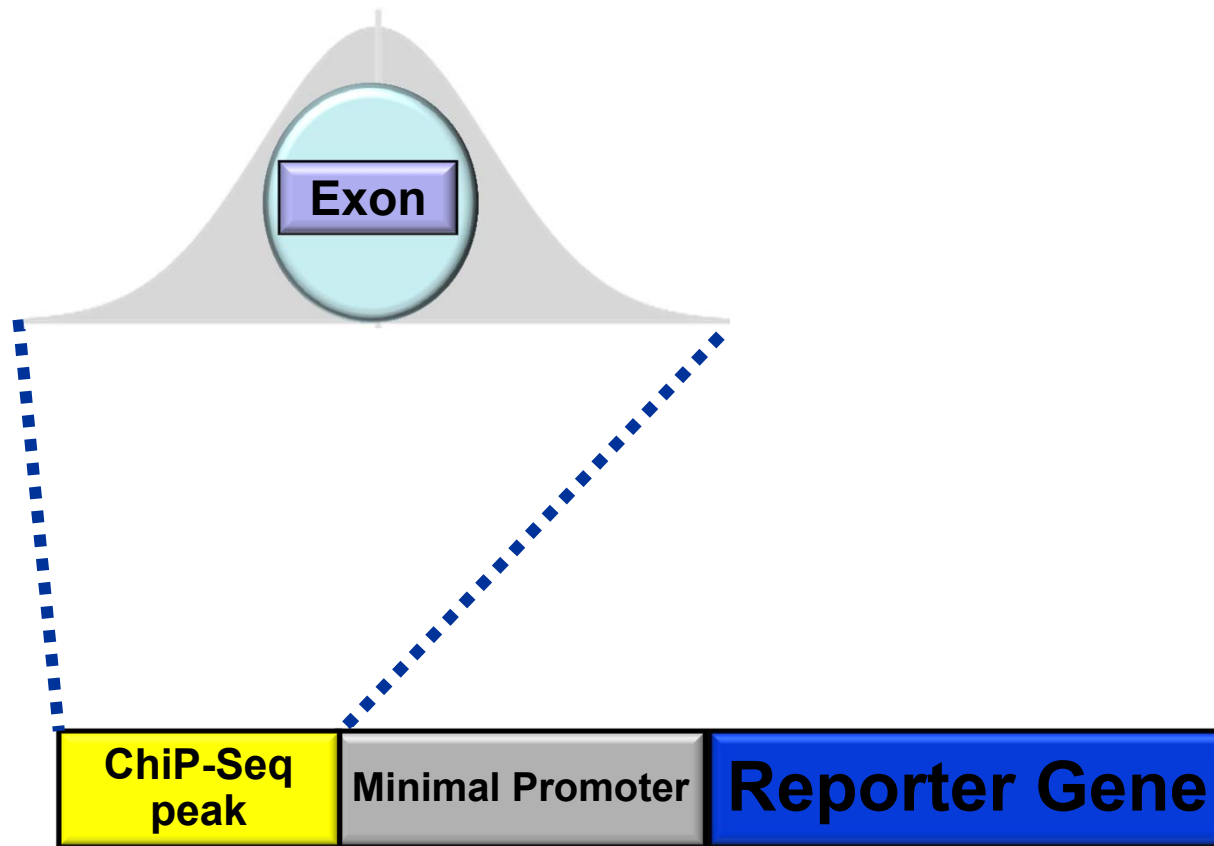
Harvest E11.5 embryos, LacZ stain



4/7 exons were positive limb enhancers in mouse



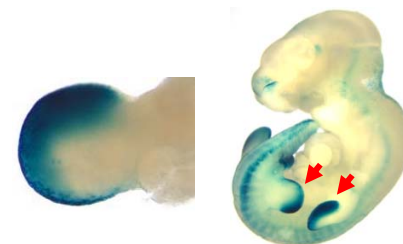
The enhancer assay tested the whole ChIP peak



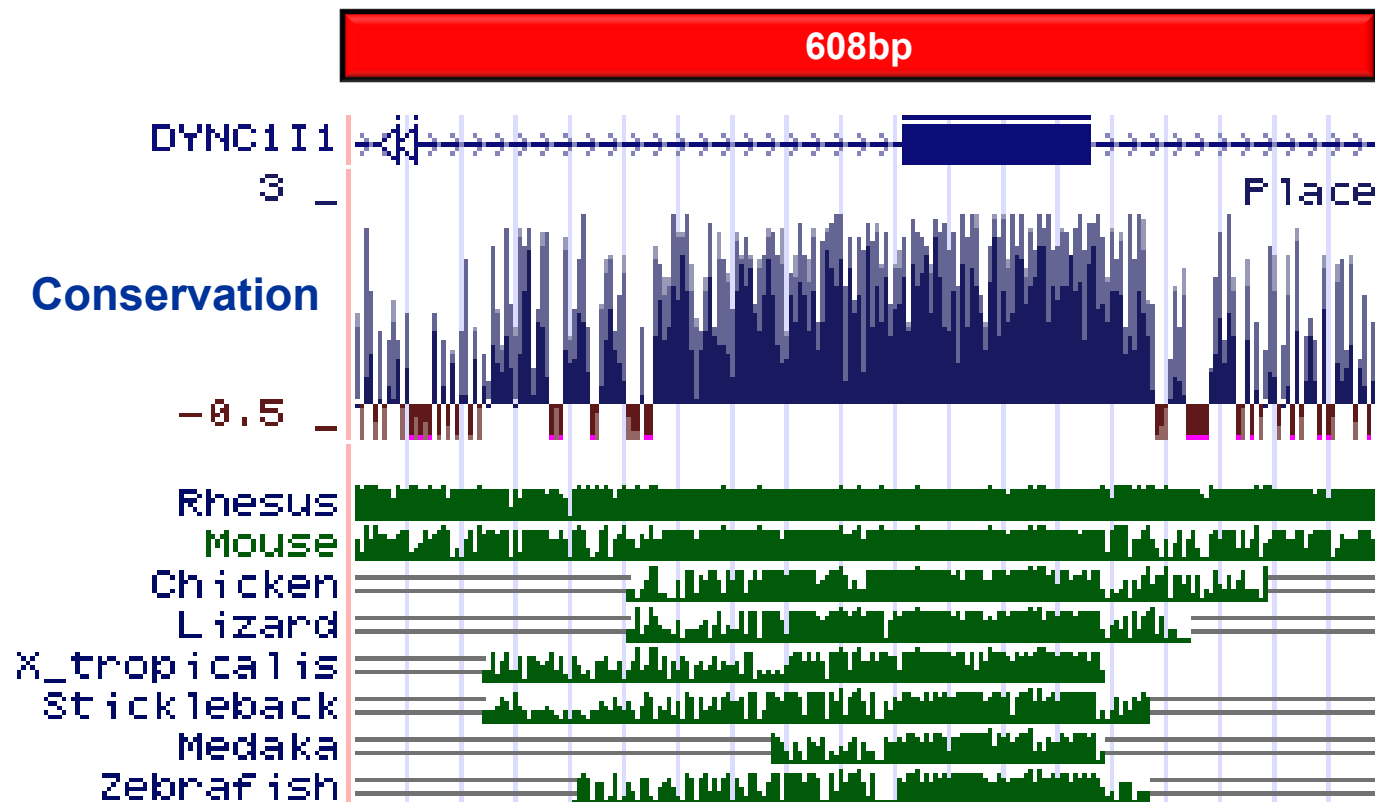
Injection into mouse fertilized eggs



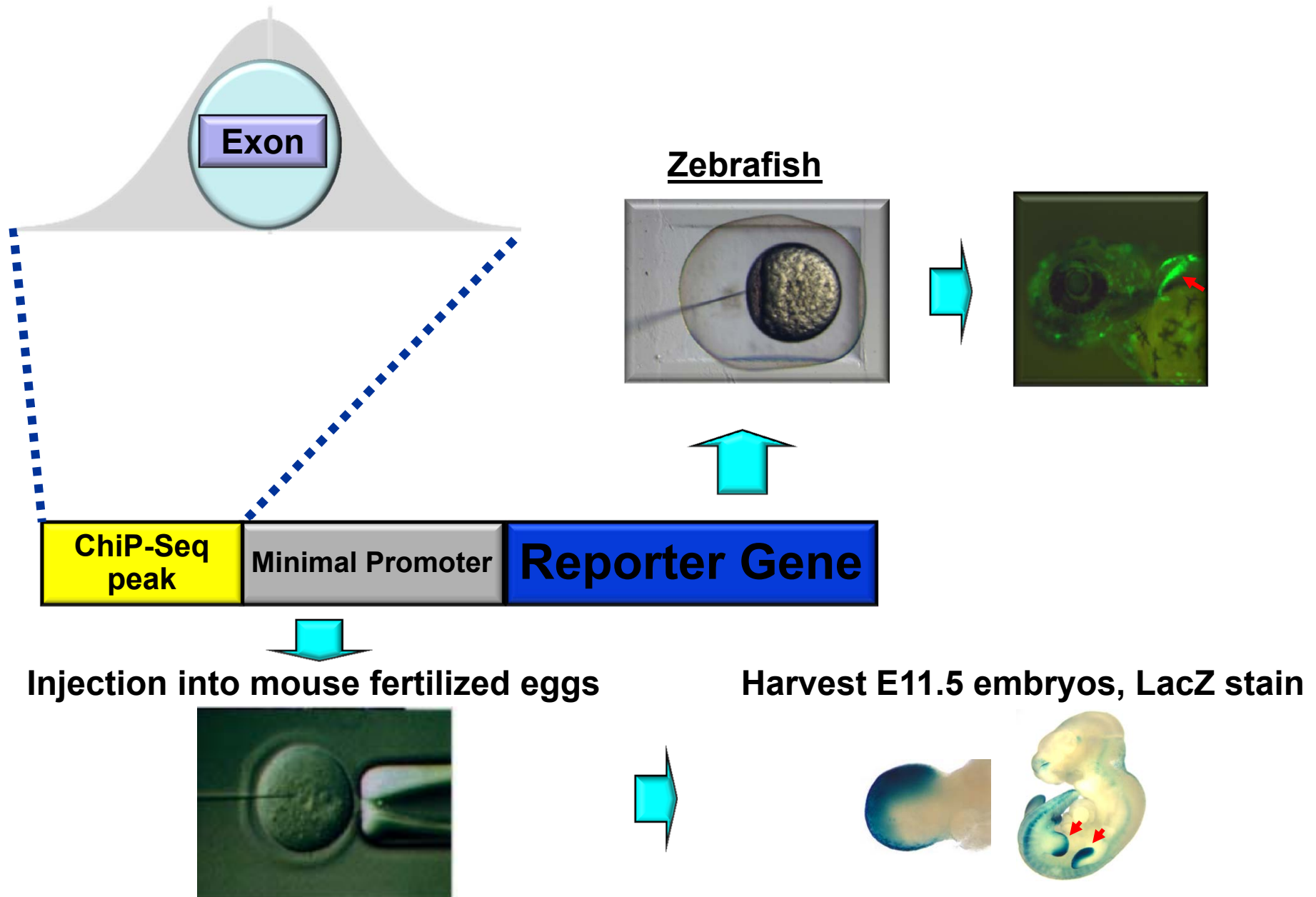
Harvest E11.5 embryos, LacZ stain



The intronic sequence could be responsible for the enhancer activity



DYNC111 exon 15 is a fin enhancer in zebrafish



DYNC111 exon 15 is necessary for enhancer activity

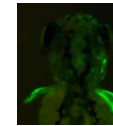
5' intron + Exon

3' intron

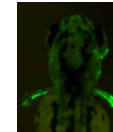
Exon

5' intron

608bp

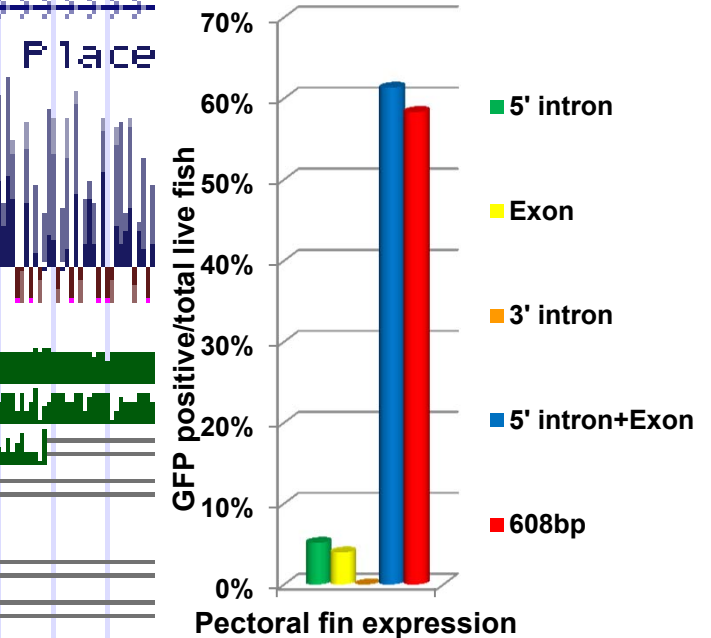
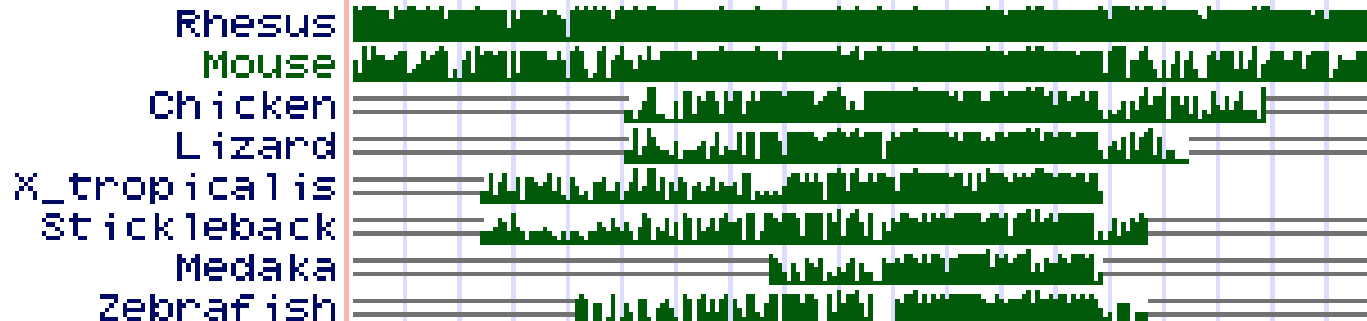


Neg.

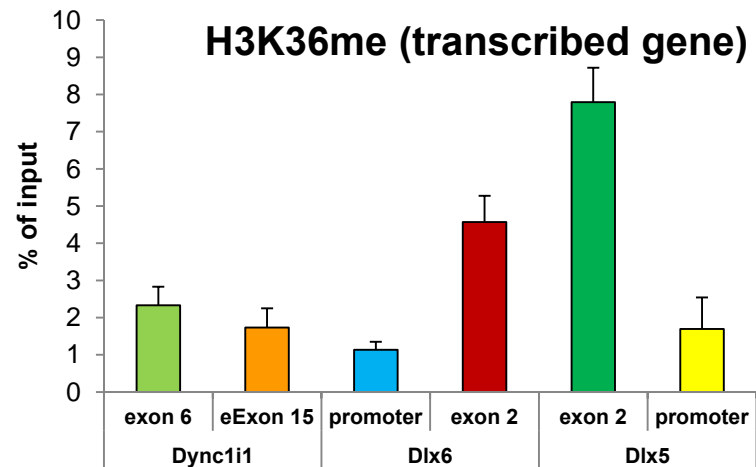
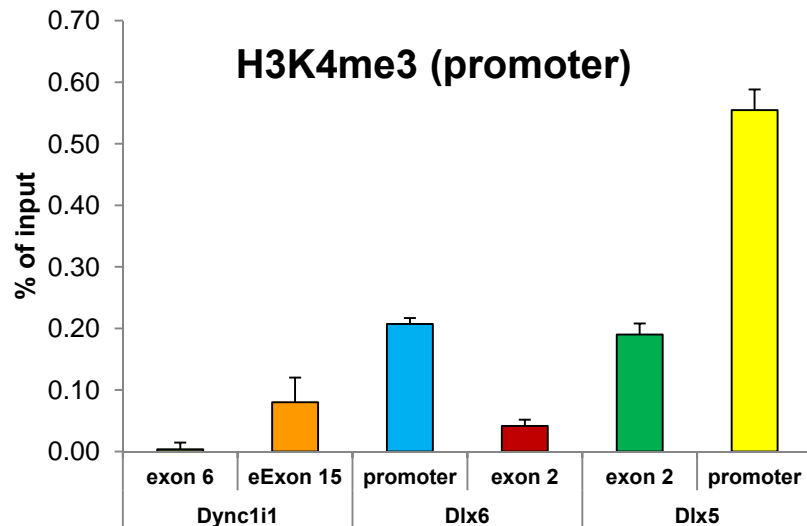
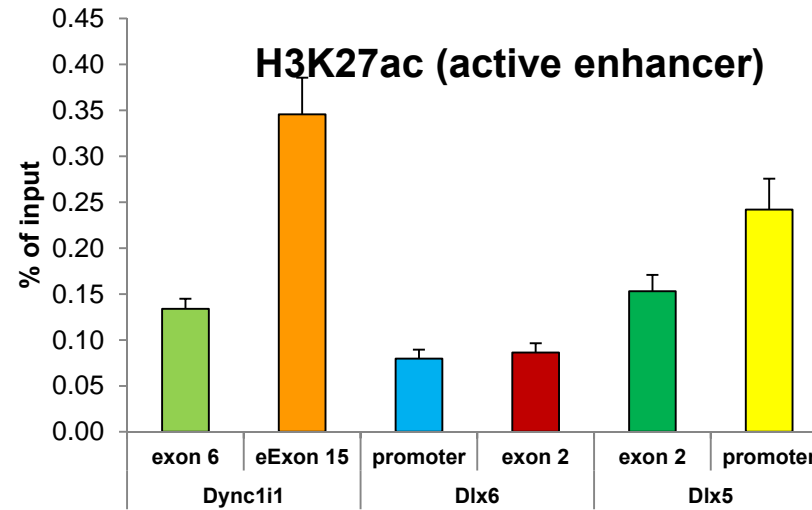
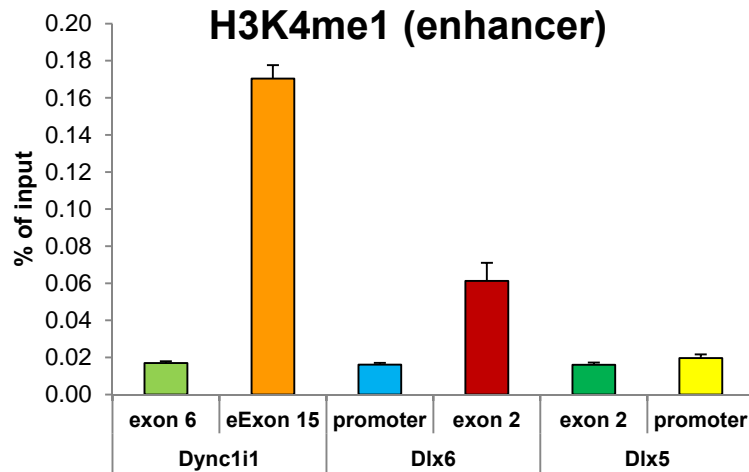
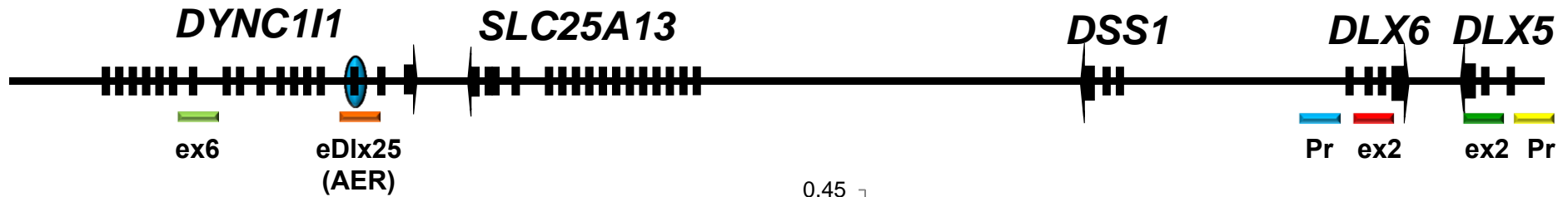


Conservation

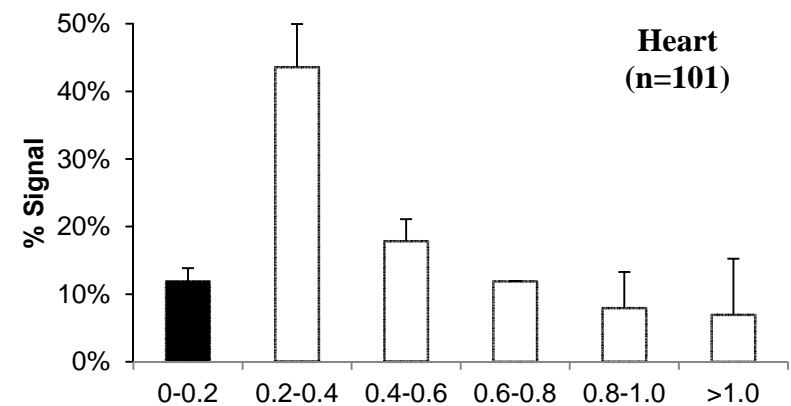
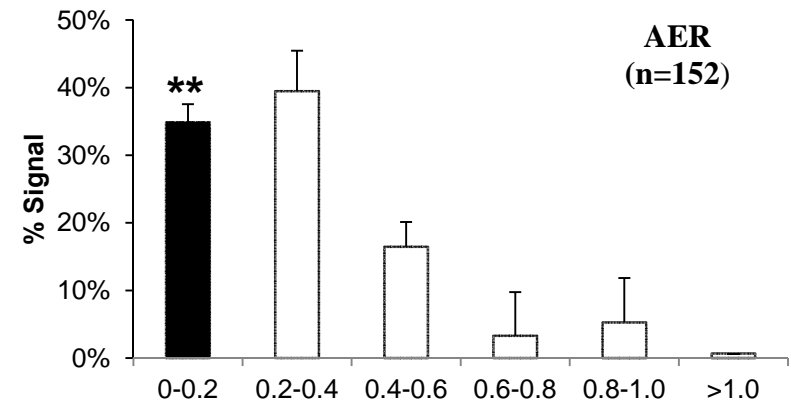
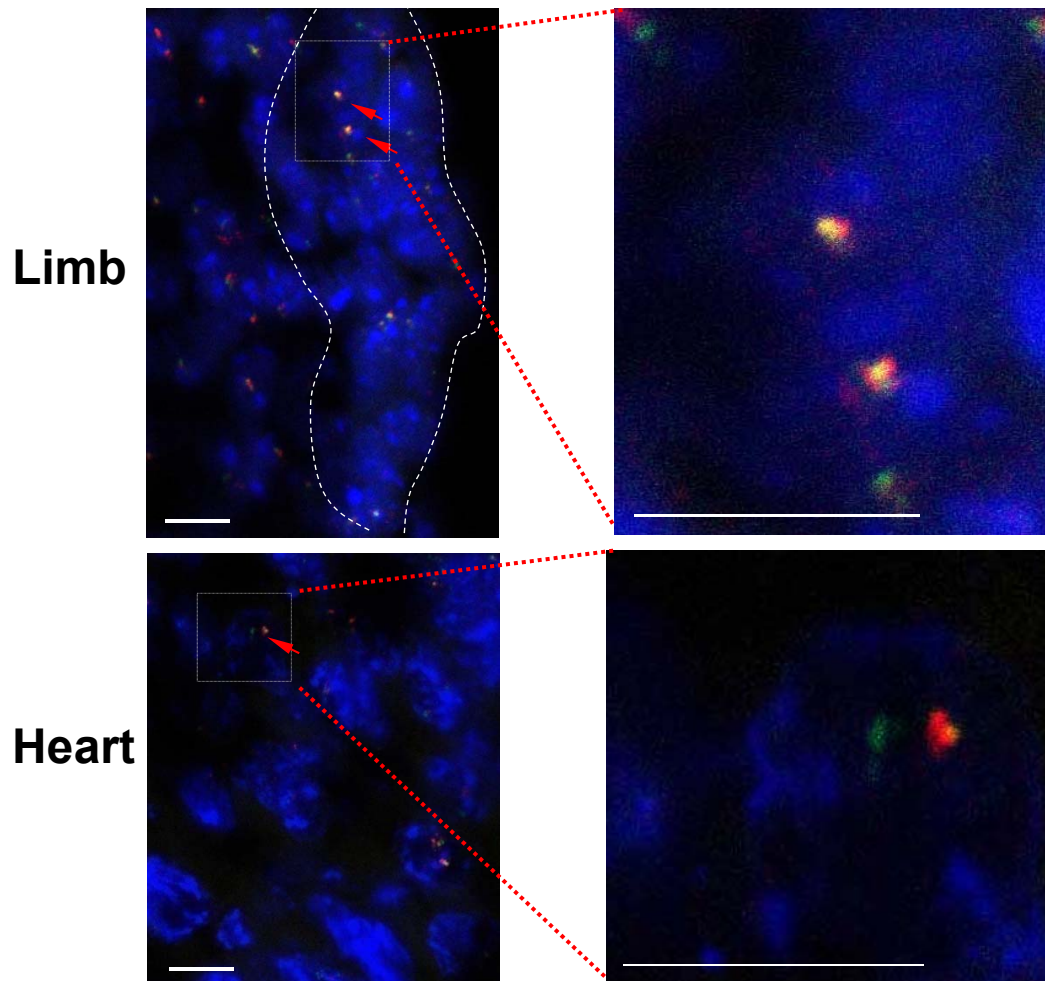
-0.5



ChIP: *Dync1i1* exon 15 has enhancer marks in the limb



DNA FISH shows interaction between *Dync1i1* exon 15 & *Dlx5/6* promoters



Chromosome Conformation Capture (3C) also shows interaction between *Dync1i1* exon 15 and *Dlx5/6* promoters

