Bi8 Final Review

3/10/2016

Disclaimer

This review is not necessarily comprehensive of everything that will be on the final exam. This review simply covers which topics and main points we think are most important based on what Professor Rothenberg has emphasized in class. Accordingly, this review will serve as a great starting point for your studying, but should not be your only reference when preparing for the exam.

Anything from all lectures is fair game.

Main topics from the first half of the course

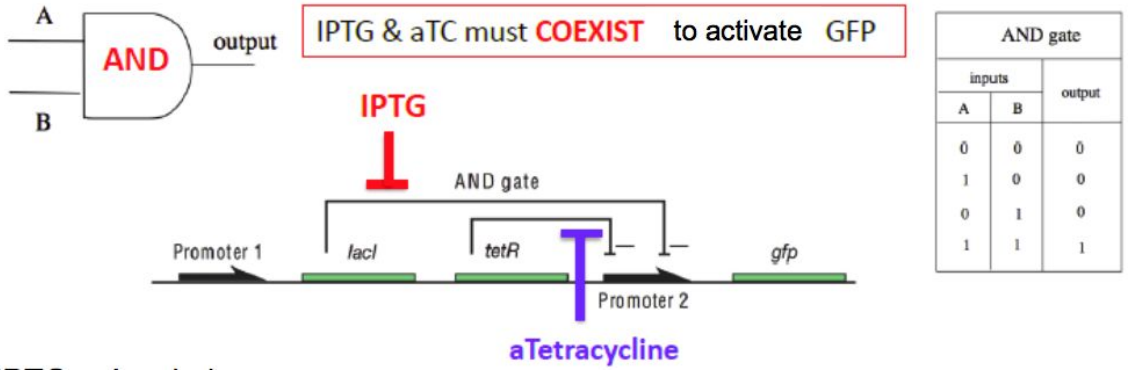
While the exam (and this review) will emphasize material from the second half of the course, these topics from the first half are still essential!

- Differences and similarities between DNA and RNA
- Transcription and (alternative) splicing
- Translation
- Protein structure and function
- Experimental techniques (probing, blotting, PCR, cloning, etc.)
- Basic bacterial regulation (e.g. lac and trp operons, lambda phage regulation)

Gene regulatory circuit design

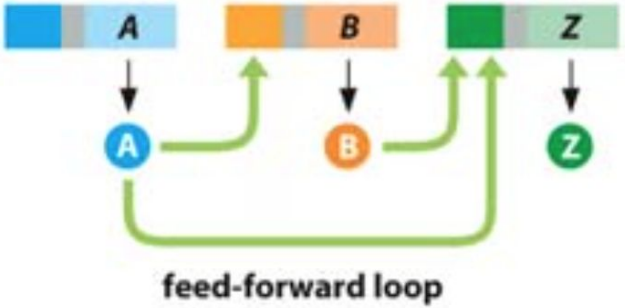
Main types of logic

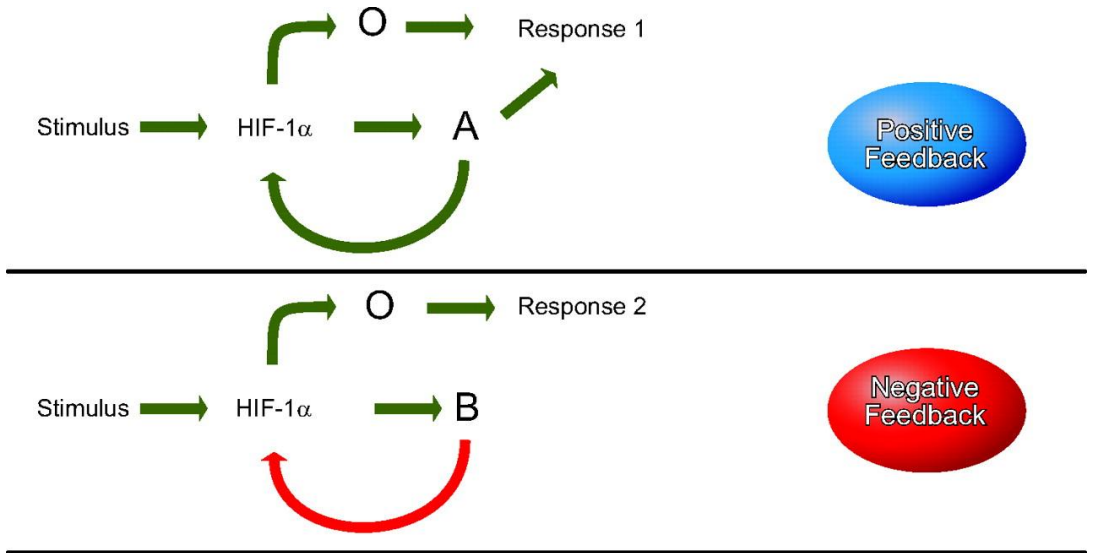
- AND
  - both inputs must be present to have output
- OR
  - either input must be present to have output
- Positive
  - activates function
- Negative
  - inhibits or represses function



AND gate		
inputs		output
A	B	
0	0	0
1	0	0
0	1	0
1	1	1

Feed forward vs Feed back

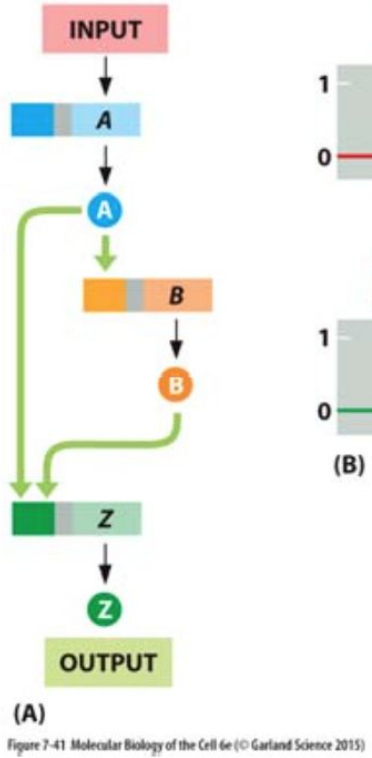




Positive Feed Forward Loop

Biological significance

- limit response to only when ‘real signal is detected



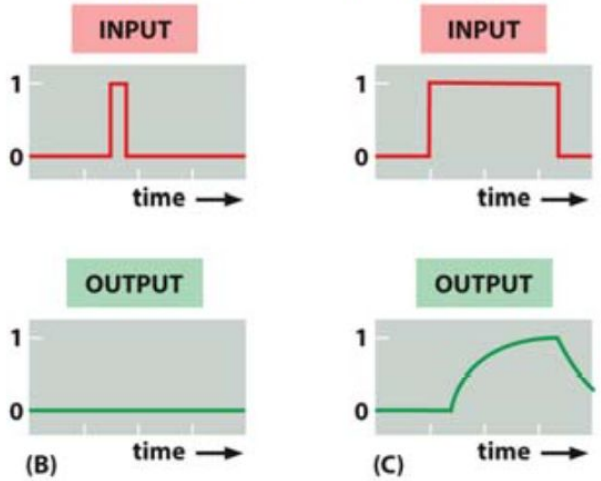
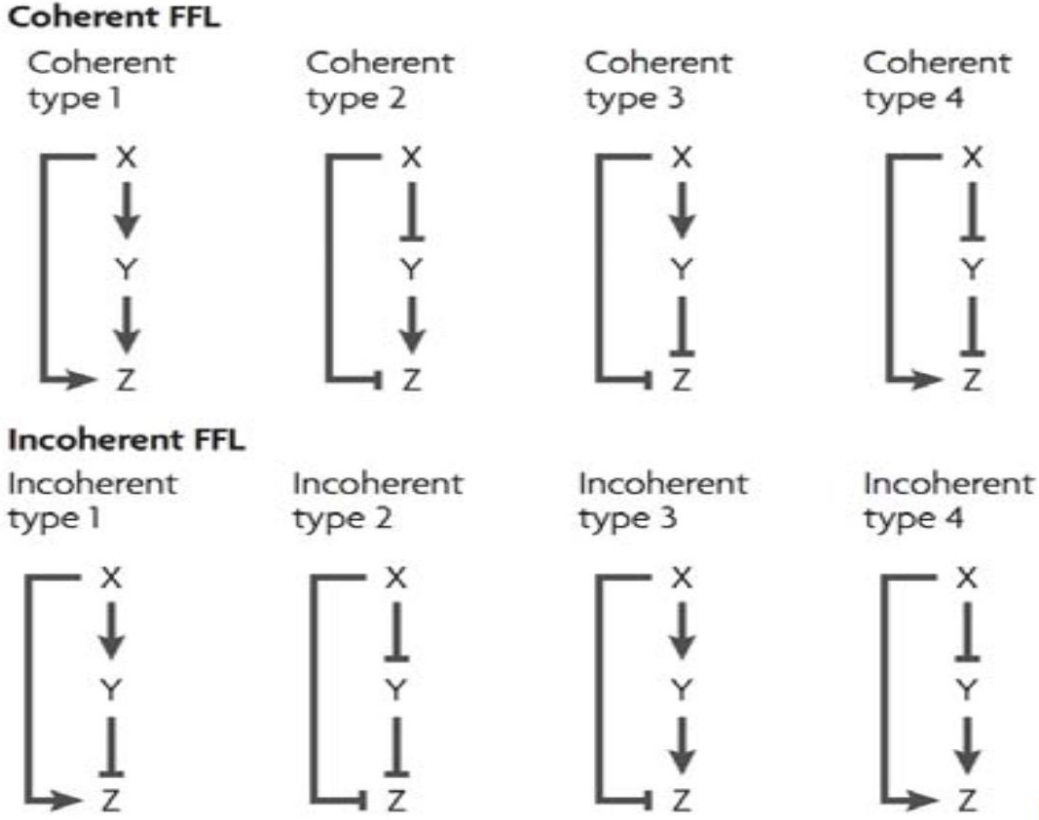


Figure 7-41 Molecular Biology of the Cell 6e (© Garland Science 2015)

Feed forward loop acting as a filter to ignore transient “noise”

Coherent vs Incoherent



U. Alon, 2007

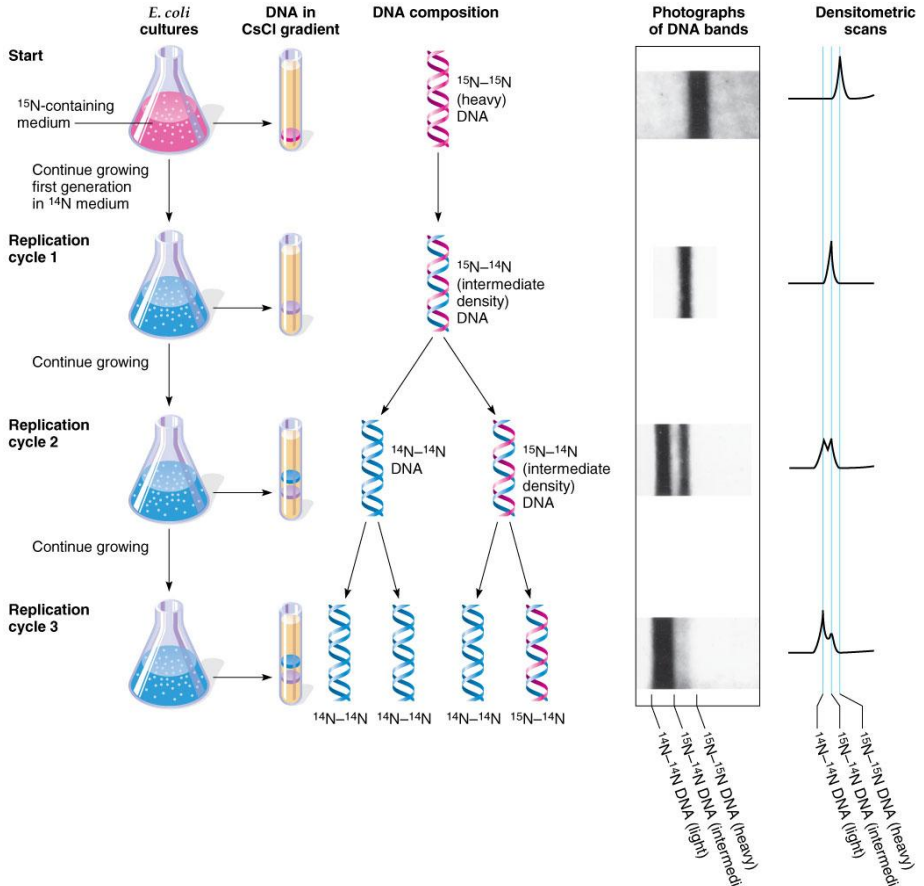
Importance of Genetic Circuits in Nature

Allows for transcriptional regulation

- Single celled organisms
  - Allows for tuning sensitivity and kinetics of responses to environment
  - Facilitates making major biological state changes
- Multicellular organisms
  - Allows for cellular differentiation (creation of various cell types)
  - Modulating responses
    - Of a cell’s own needs and to maintain the current physiology of the whole organsim

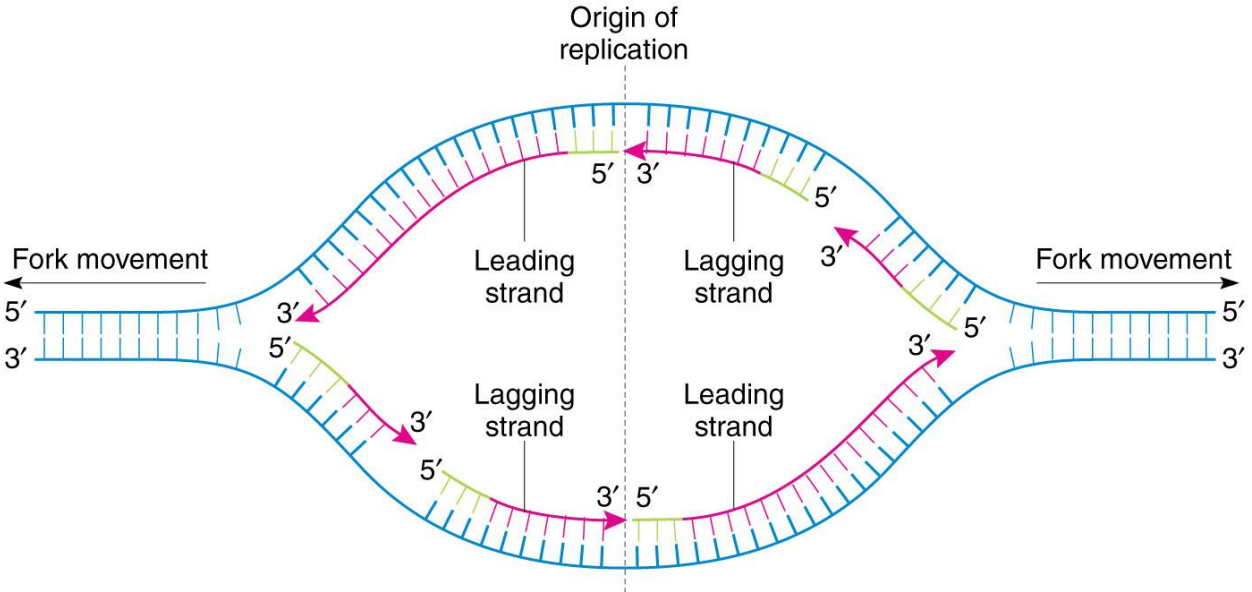
DNA replication

Meselson-Stahl Experiment: Genomic DNA Replication is Semi-Conservative



Defining Origins of DNA Replication

- Origin recognition complex identifies replication origin(s)
  - Multiple replication origins in eukaryotes
  - One replication origin in prokaryote (circular DNA)
- All DNA synthesis begins with primers generated by DNA primase



DNA Replication Overview

Key players:

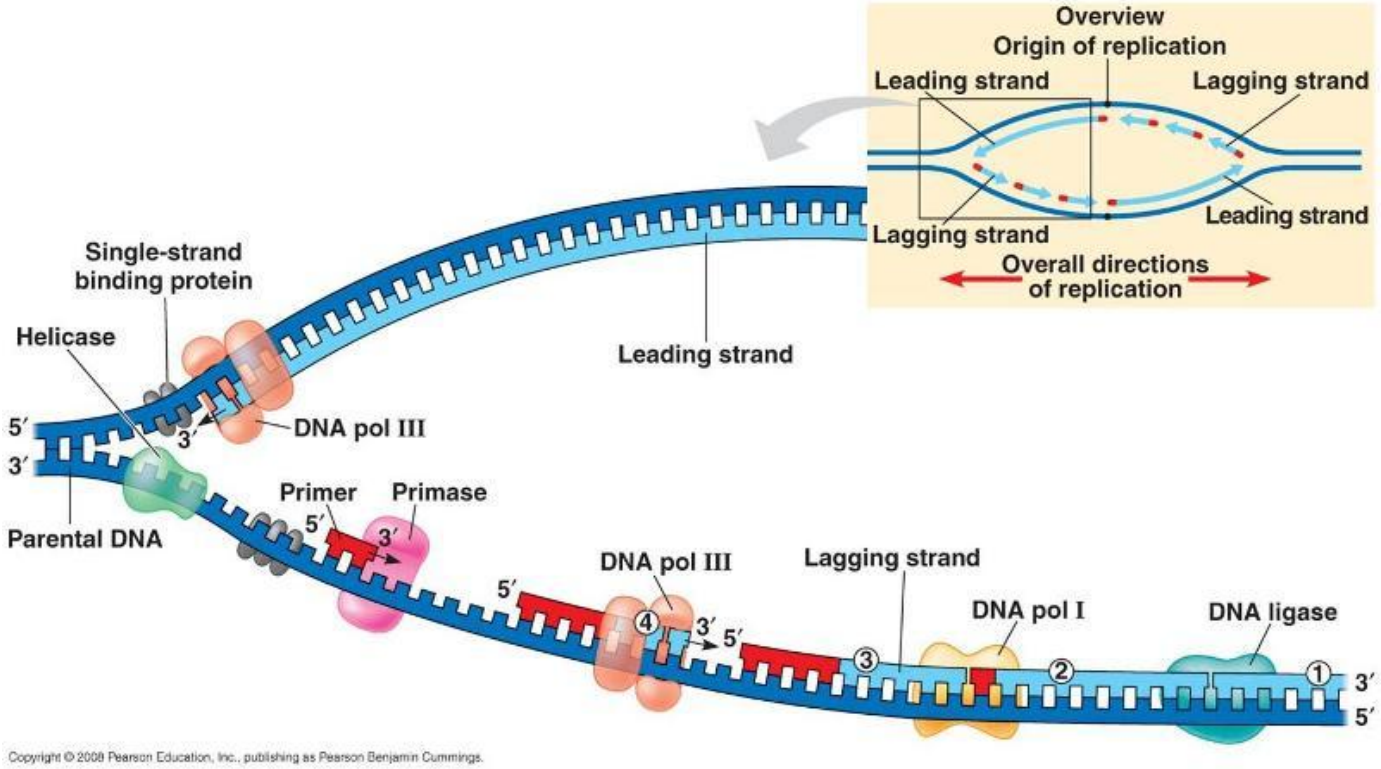
-DNA Polymerase

-Helicase

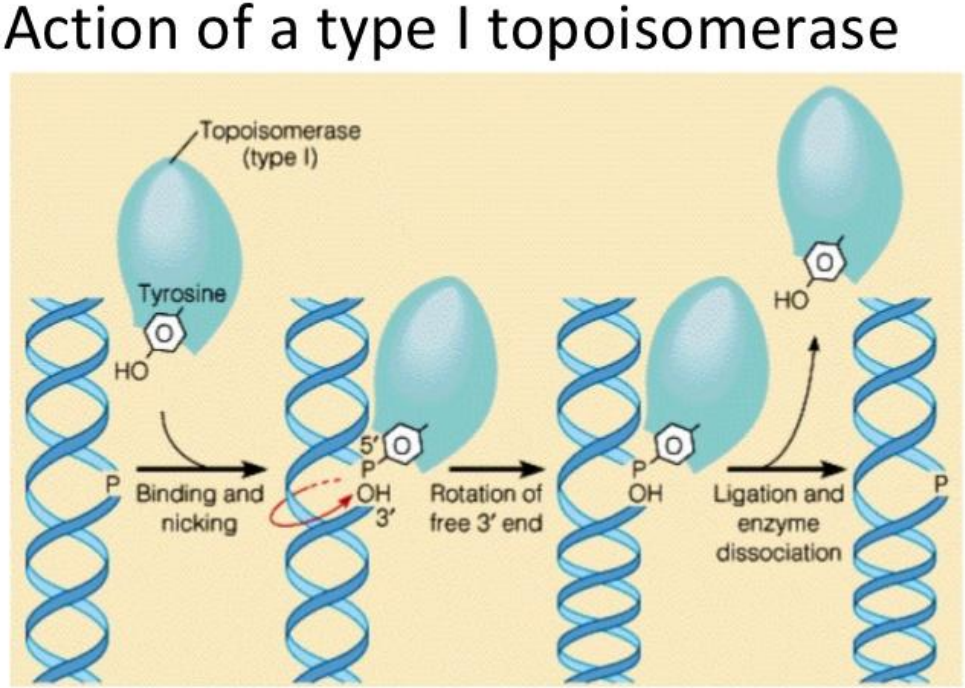
-Primase

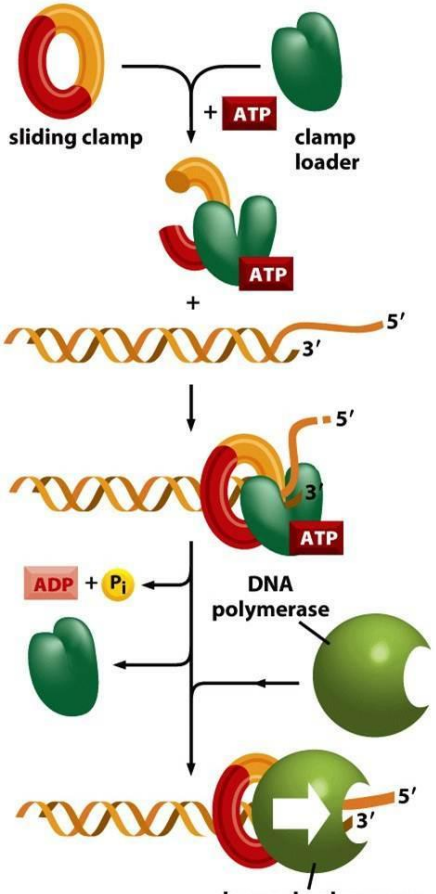
-DNA ligase

-Single stranded binding protein



Additional Players: Topoisomerase and Sliding Clamp

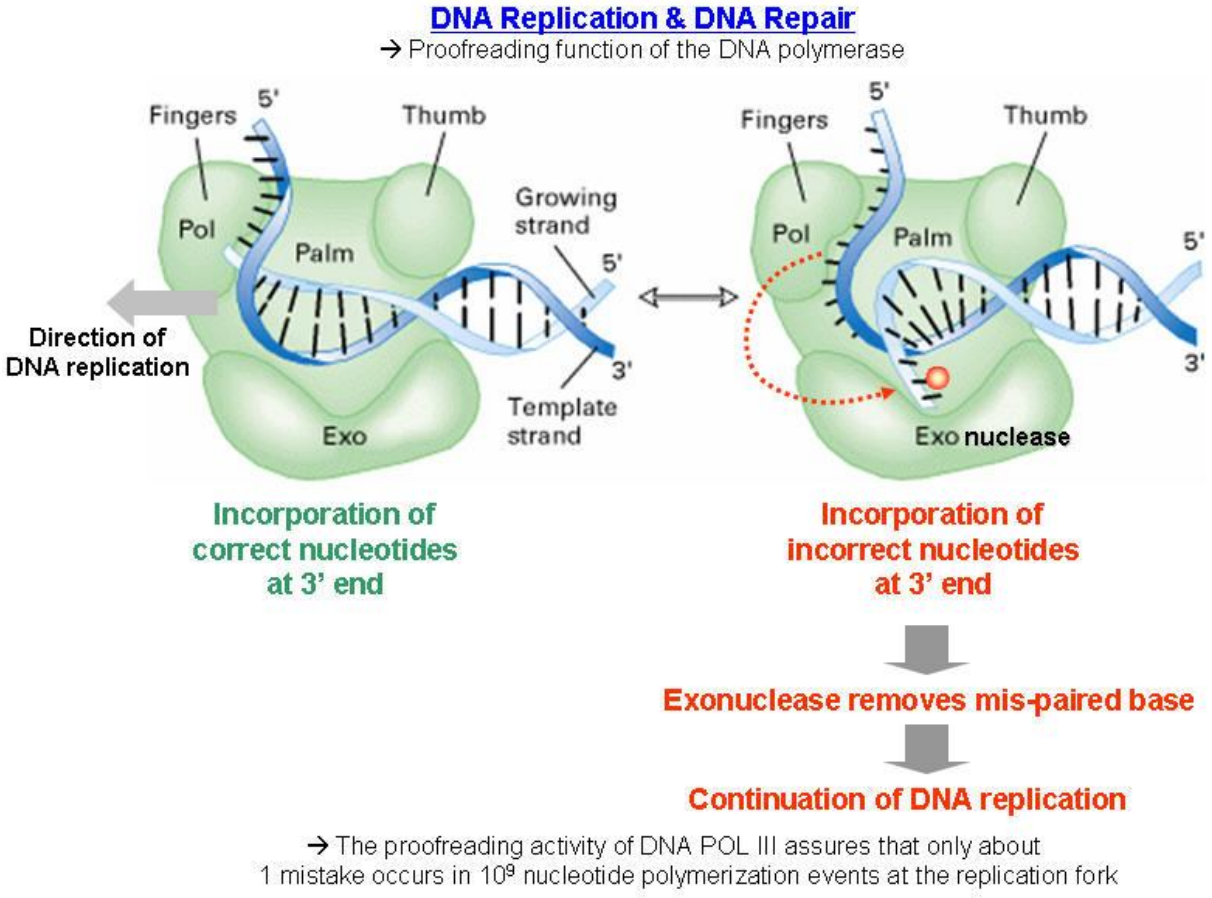




DNA Polymerase Exonuclease Activity

DNA Polymerase has two modes:

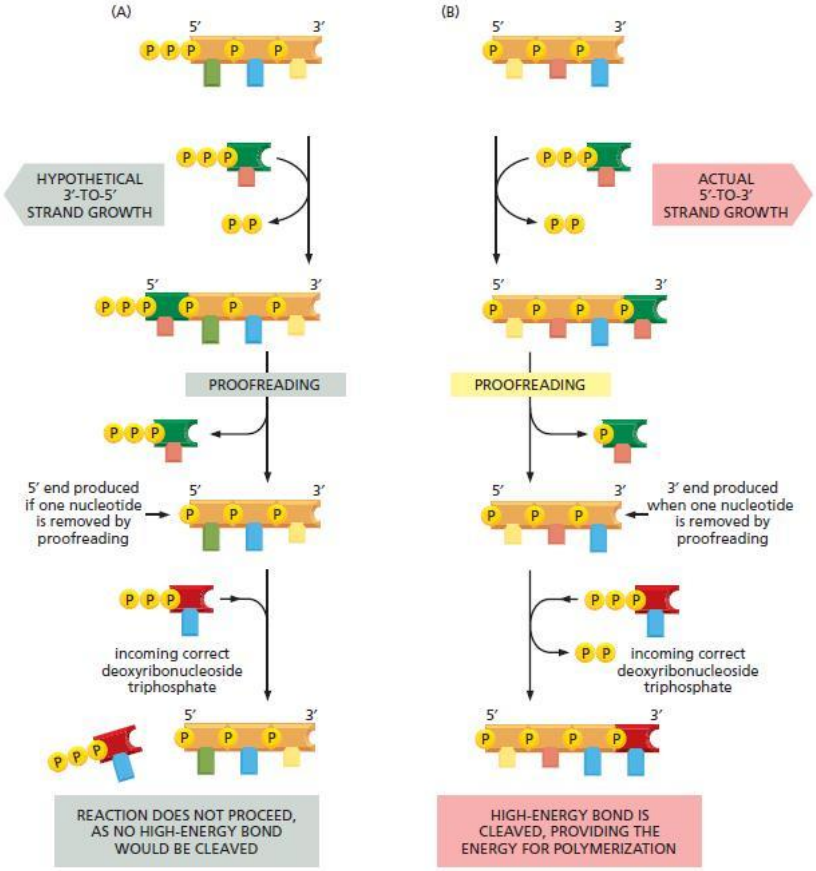
- Polymerization
- Exonucleolytic proofreading
  - in combination, result in only one error in every 10^9 bases



5’ to 3’ Addition of dNTPs

DNA polymerization occurs 5’ to 3’

- This is essential for DNA Polymeras
  s exonuclease activity



DNA mutation and repair

Terms

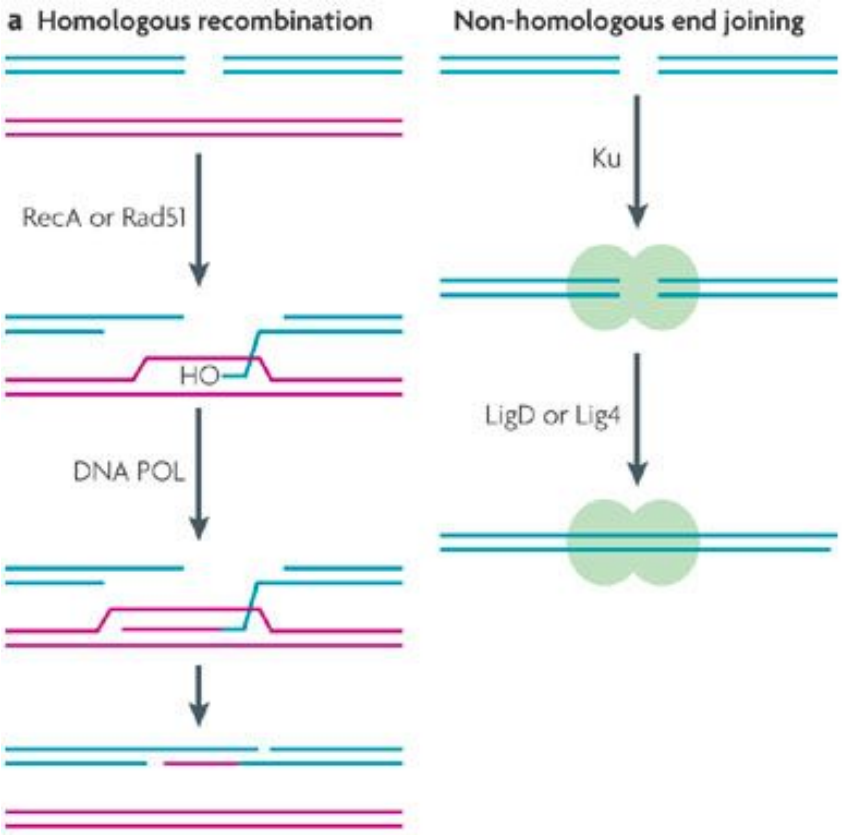
Damage: abnormal chemical structure

Examples: C deamination, C/U/T dimer, **dsDNA breaks**

Mutation: altered sequence

Which type of abnormality can be **inherited**?

How to fix dsDNA breaks



1st choice: homologous recombination

1. Process broken strands to produce 3’ invader
2. Finds homologous site (template for synthesis) on another chromosome through sequence identity
3. Invader strand serves as primer, DNA polymerase extends

Consequences

1. Accurate (no loss of genetic information)
2. Conservative (both strands incorporate the homologous sequence)

Nonhomologous end-joining

1. dsDNA break is processed to give blunt ends (removed bases are never re-added)
2. Blunt ends are ligated together

Consequences

1. Lose genetic information
2. Still better than a broken chromosome

Recombination in other contexts

Meiosis

Transposons

Recombinases

Immune system

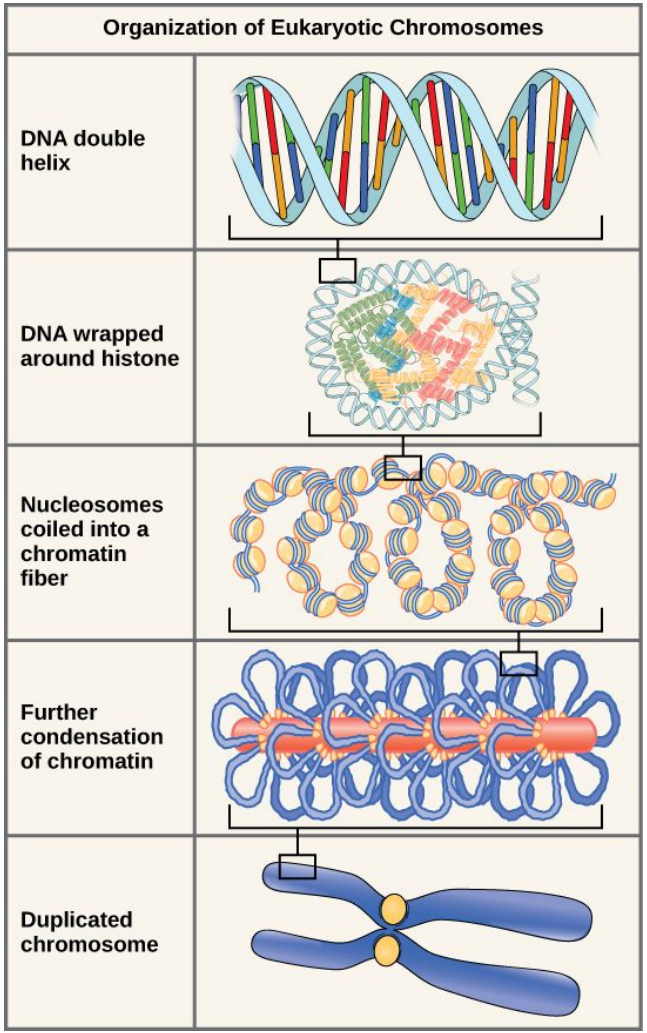
Transcriptional regulation in eukaryotes

Higher-order chromatin structure

DNA double helix: sequence information

Nucleosome: epigenetic information

Chromatin loops: facilitate long-range static & dynamic interactions between regulatory sequences and genes



Regulation in *cis* and *trans*

Cis-regulatory element: DNA sequence that acts to regulate another DNA sequence (*same molecule*)

Ex. promoter, enhancer, silencer

Trans-regulatory element: Protein that acts to regulate a DNA sequence (*different molecules*)

Ex. transcription factor

Cis & *trans* regulation in eukaryotes

Multiple transcription factors collaborate at each cis-reg element

- 2 kinds of logic?

Multiple cis-reg elements can control expression of the same gene

- Multiple promoters
- Enhancers v. silencers

Consequences for development

Development is: irreversible (~95% true), ordered increase in complexity, based on memory of past events.

Another 95% true statement: the most important genes in development are those that code for transcription factors.

Sequence of transcription factor action is crucial:

- Early TFs participate in chromatin remodeling and alters the regulatory state
- Later arriving TFs can recruit co-activators, polymerase, etc.

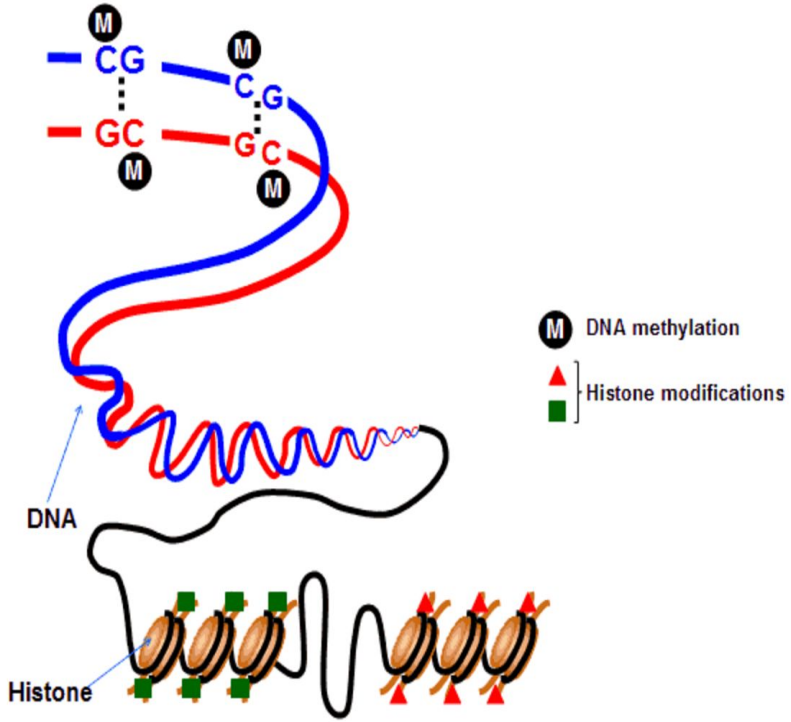
Epigenetics

Epigenetics:

Two major epigenetic modifications:

- DNA methylation - covalent (but reversible) methylation of DNA sequence itself on CpG dinucleotides.
  Cytosine is methylated

- Histone modification - covalent (but reversible) post-translational modification of histone protein tails in nucleosomes
  Methylation of lysine
  Acetylation of lysine

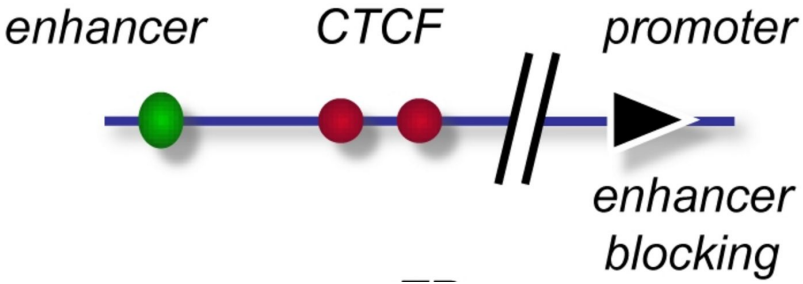


http://epigeek.com/functional-noncpg-methylation-mammalian-cells/

CpG methylation:

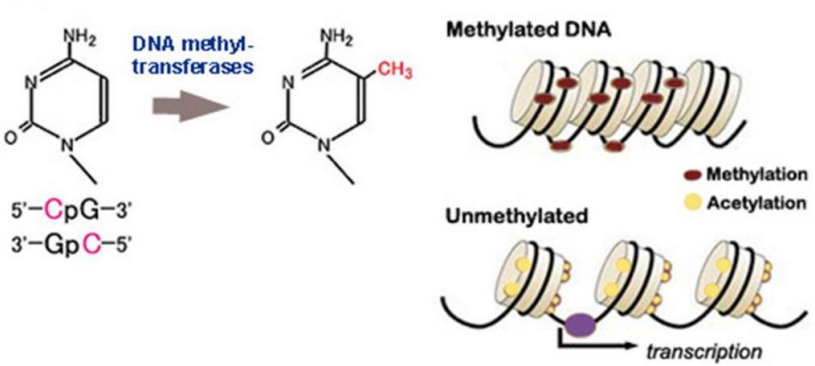
Effects of CpG methylation on gene regulation:

- If a TF normally binds to a site with a 5’-CG-3’ dinucleotide in it, methylation can block it from recognizing the newly methylated site.
- DNA with methylated CpG’s can be preferentially bundled into closed chromatin
- CpG methylation can block CTCF binding near a gene, causing the gene to end up in the wrong loop & under the regulation of a wrong regulatory elements



Without CpG methylation, CTCF can bind to specific DNA sequence motif and block the wrong enhancer to regulate its nearby gene.

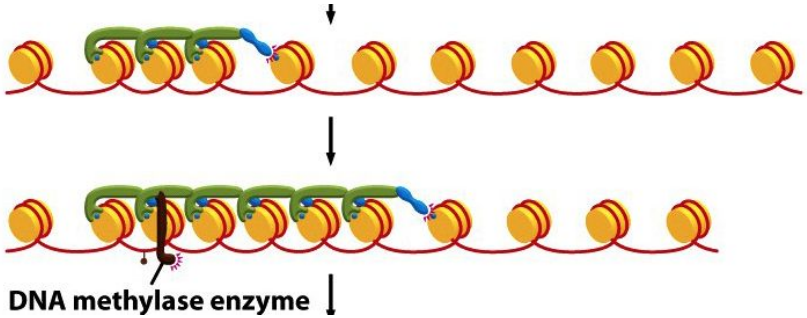
https://www.researchgate.net/figure/43202212_fig7_Enhancer-blocking-elements-may-require-multiple-CTCF-sites-for-optimal-function-as



http://helicase.pbworks.com/w/page/17605615/DNA%20Methylation

Some characteristics of CpG methylation

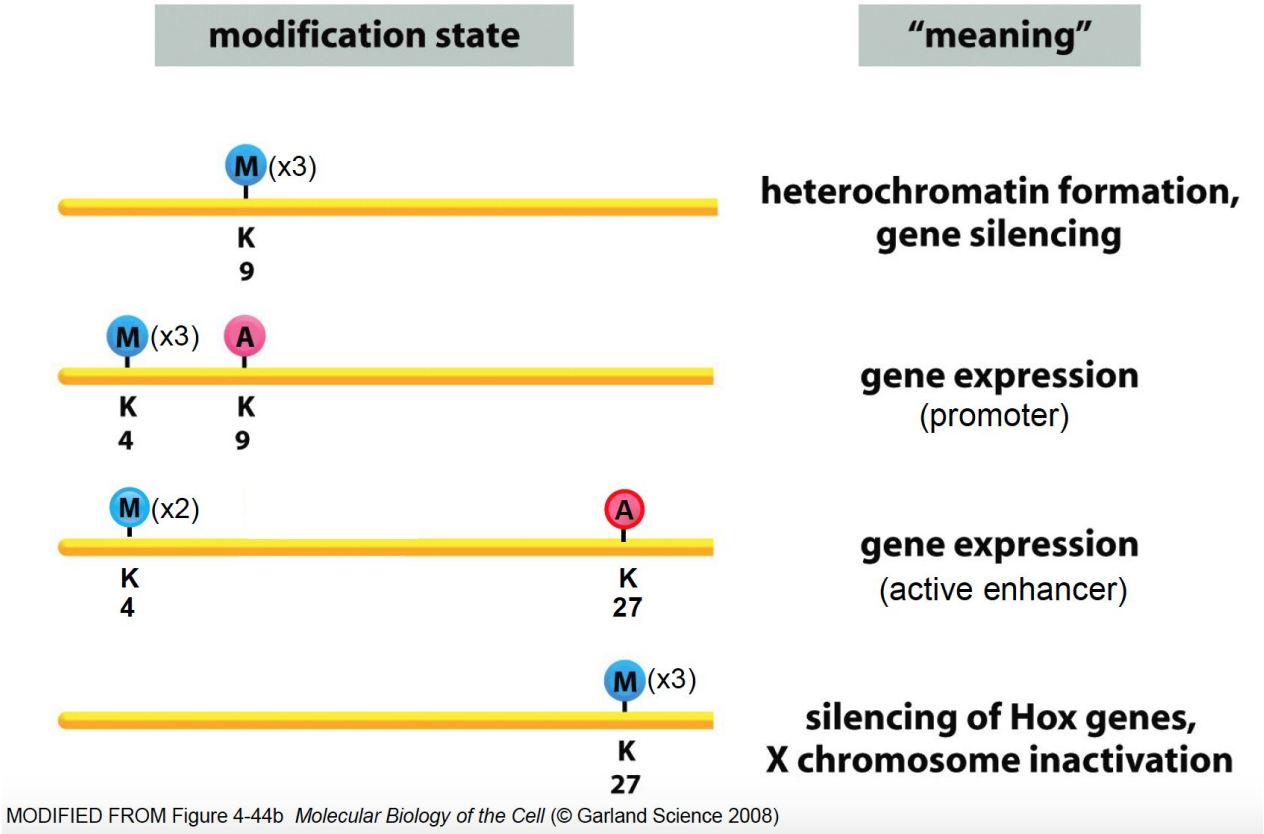
- Can be maintained over multiple cell generation
- Can Propagate and keep genes silent in future cells



- Can be erased (Key transcription factors can recruit DNA demethylation enzymes “TET (ten-eleven translocation) family enzymes” to demethylating local DNA)
- Occurs as a result of transcriptional silence, not as a cause of it.

Histone modification

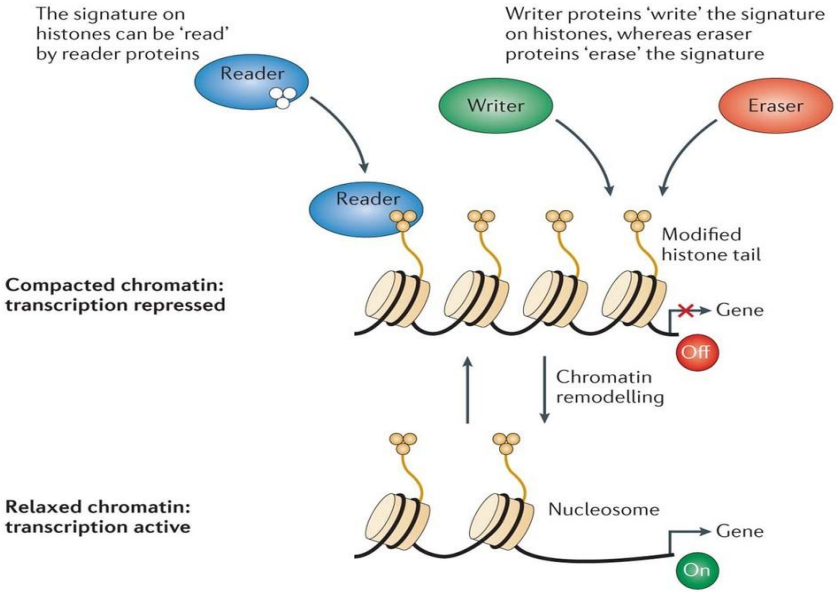
Please remember these common modifications!



MODIFIED FROM Figure 4-44b *Molecular Biology of the Cell* (© Garland Science 2008)

Histone Modification

- Modifications are on histone N-terminal tails
- Some histone modification markers in “silent chromatin” provide docking sites for chromatin-condensing factors (such as HP1 protein)
- Some modifications, especially “repressive” ones, can propagate themselves through a recruitment-relay mechanism
- Transcription factors can recruit histone modifying enzymes to install corresponding modification markers

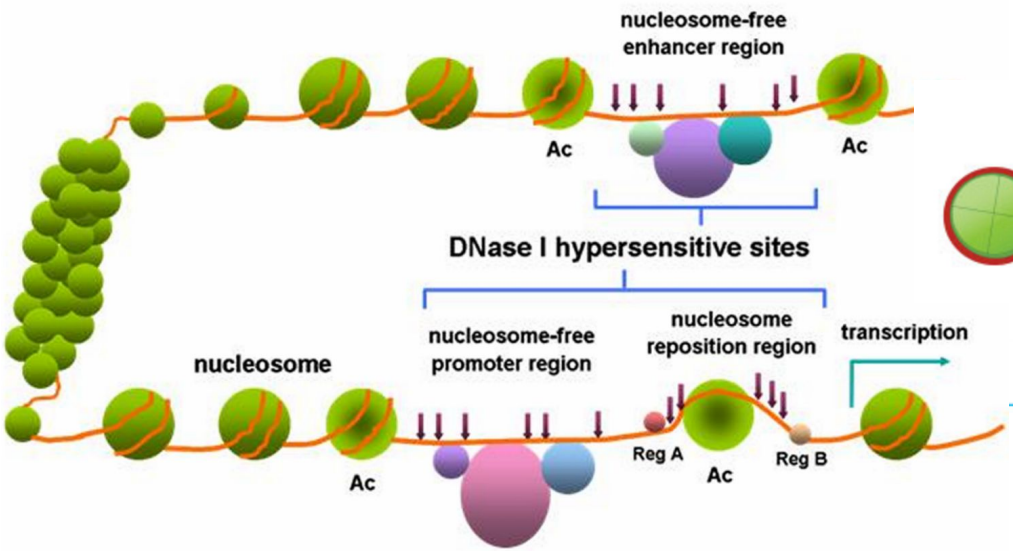


Dynamic changes of histone modification

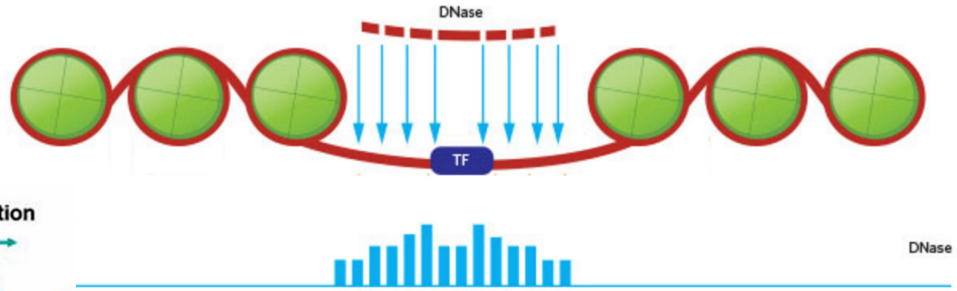
Jonas W. Højfeldt et al., 2013, Nature Reviews

DNase hypersensitivity regions may be especially active in cell-type specific gene regulation

- DNase I preferentially cuts nucleosome-free region
- Using modern high-throughput sequencing technology, we can locate this region



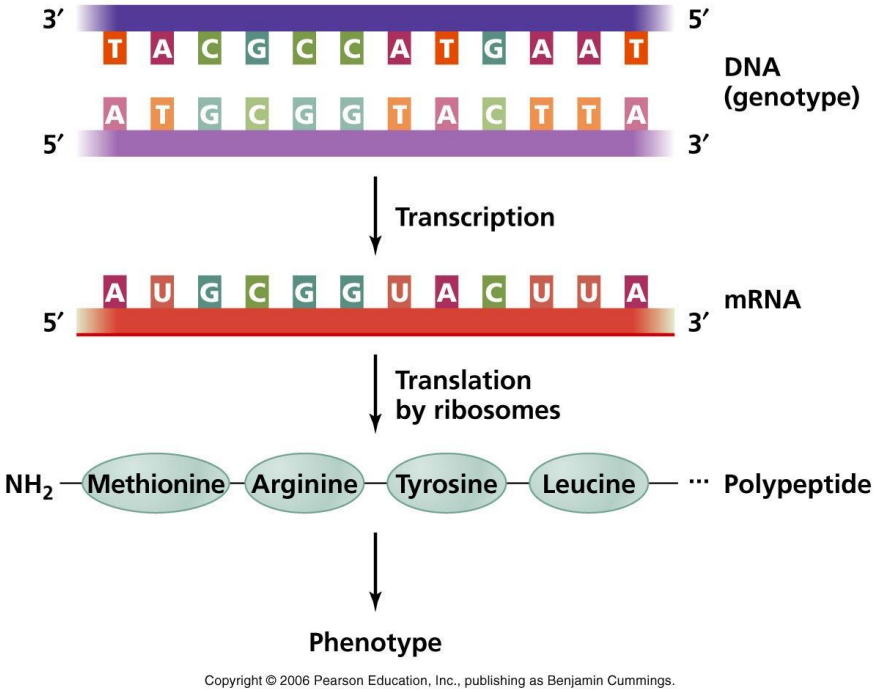
Ya-Mei Wang et al., 2012 PLoS One



Modified from http://www.the-scientist.com/?articles.view/articleNo/44772/title/Reveling-in-the-Revealed/

RNA World

~~The Central Dogma~~



But RNA can do so much more than just code for protein!

Other RNAs than just mRNA

miRNA/siRNA/shRNA/dsRNA: involved in RNAi

lncRNA: long non-coding RNA

crRNA: CRISPR RNA

snRNA: small nuclear RNA found in the spliceosome

TERC: telomerase RNA component

rRNA: ribosomal RNA

tRNA: transfer RNA in translation

miRNAs vs. siRNAs

miRNAs are **natural**, while siRNAs are **synthetic**

miRNAs bind with **lower specificity** while siRNAs require **exact base-pairing**

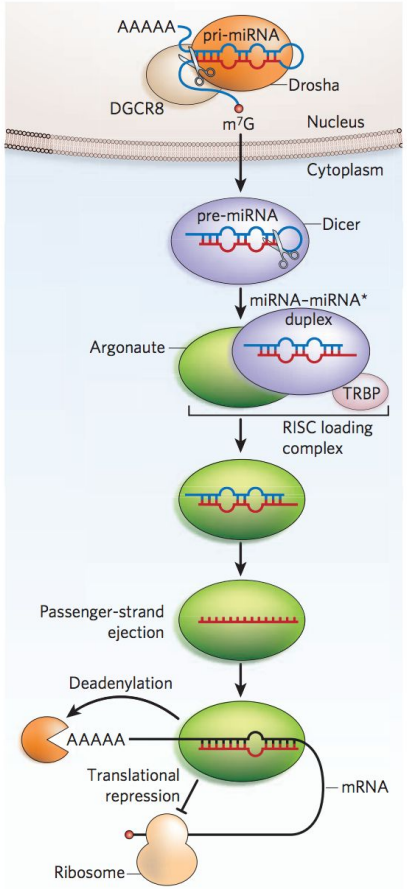
Thus miRNAs can **bind to many genes**, while siRNA **bind only once**

miRNAs generally act through **translation inhibition** while siRNAs work through **mRNA cleavage**

For both, can only cause **negative regulation** by inhibiting translation of an **already transcribed mRNA**

miRNA processing

1. “Cropping” of pri-miRNA by Drosha/Pasha to get hairpin substrate (pre-miRNA)
2. Exporting pre-miRNA out of nucleus
3. “Dicing” by Dicer to get double-stranded substrate
4. Loading of the mature miRNA onto Argonaute protein
5. Targeting of mRNA and inhibition of translation via:
    a. Deadenylation
    b. Decapping
    c. etc.

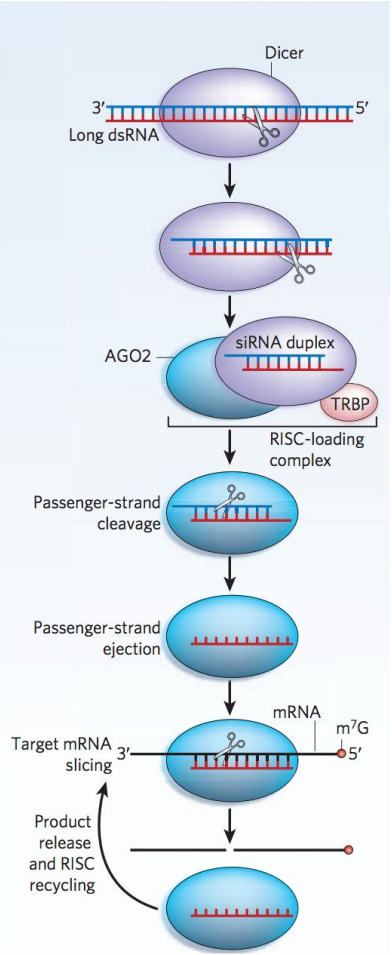


(Jinek and Doudna, 2009)

siRNA processing

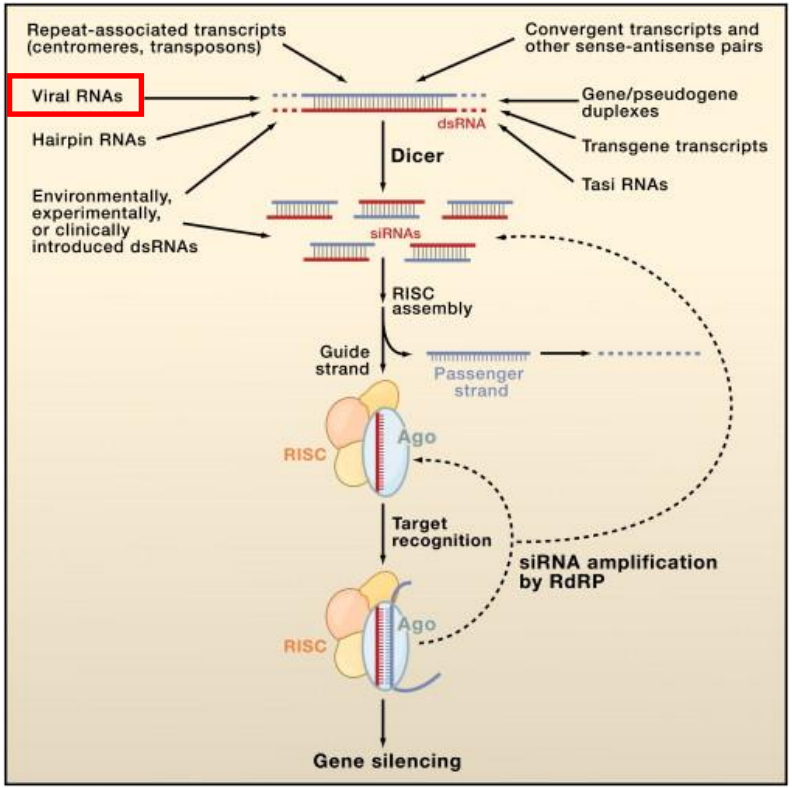
1. “Dicing” of synthetic long dsRNA to get smaller double-stranded duplex
2. Loading of siRNA onto Argonaute protein
3. Targeting of mRNA and “slicing”

Very similar to miRNA processing, but skips earlier steps because we don’t have to begin with a natural pri-mRNA



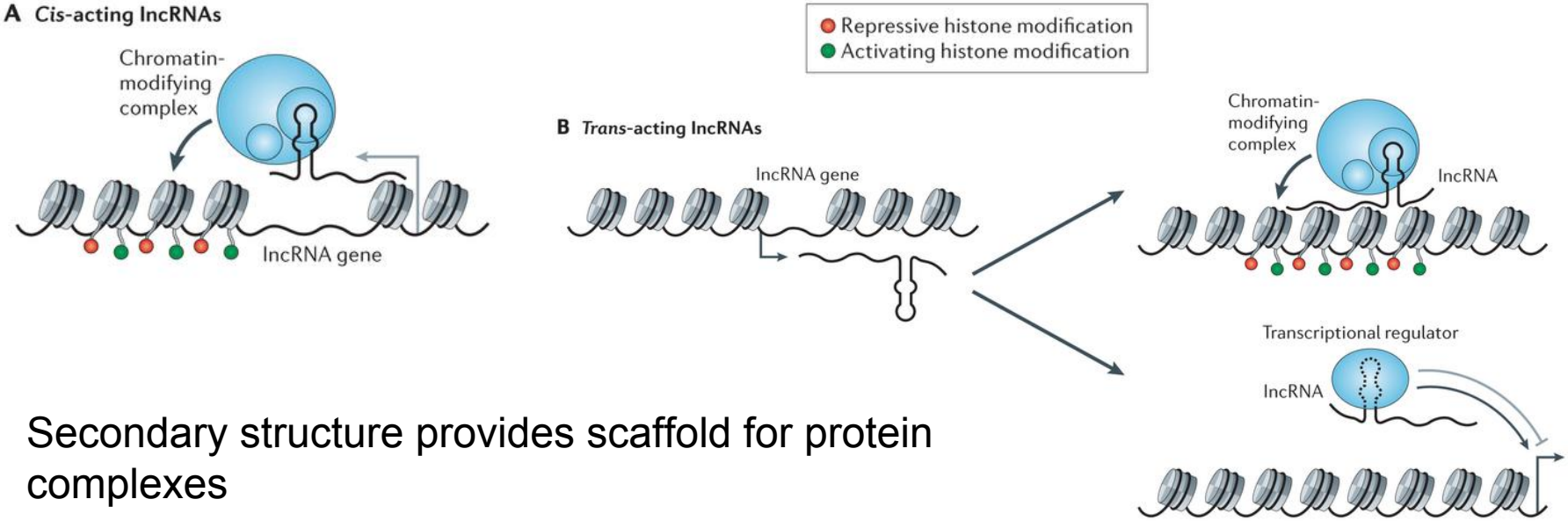
(Jinek and Doudna, 2009)

RNAi as a defense mechanism against dsRNA



(Carthew and Sontheimer, 2009)

lncRNAs (long noncoding RNAs)



Secondary structure provides scaffold for protein complexes

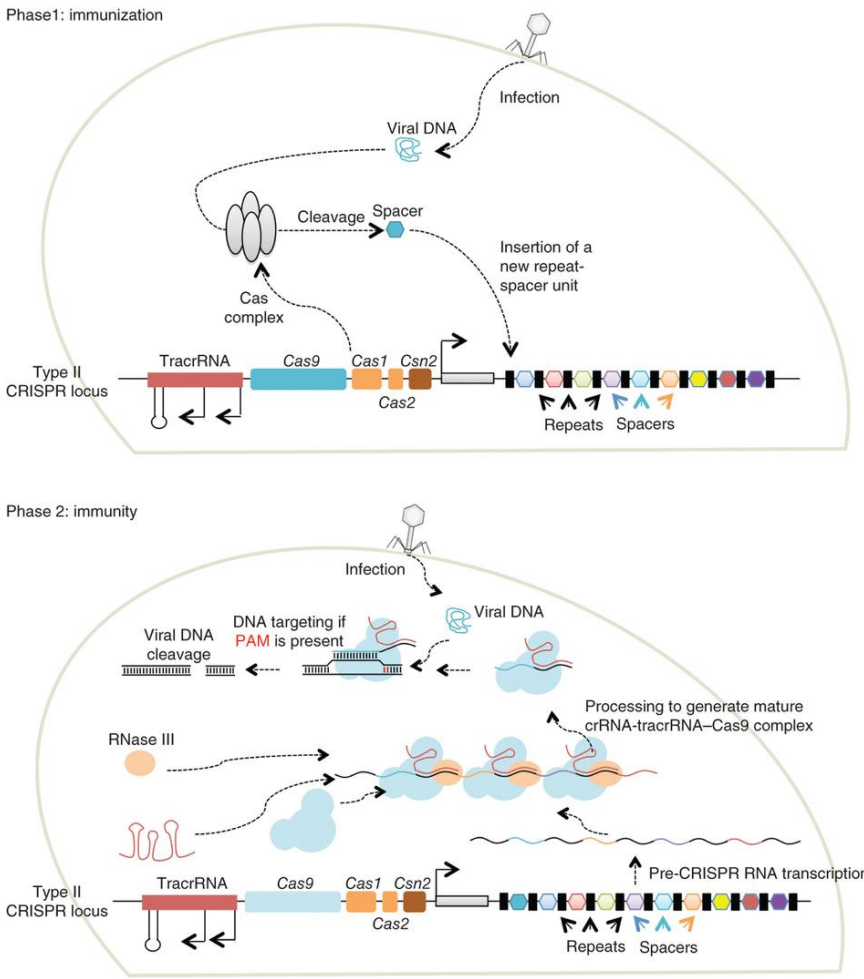
Functions of lncRNAs still being discovered

CRISPR-Cas9: immune system for bacteria

Can “hack” this system to do genome editing

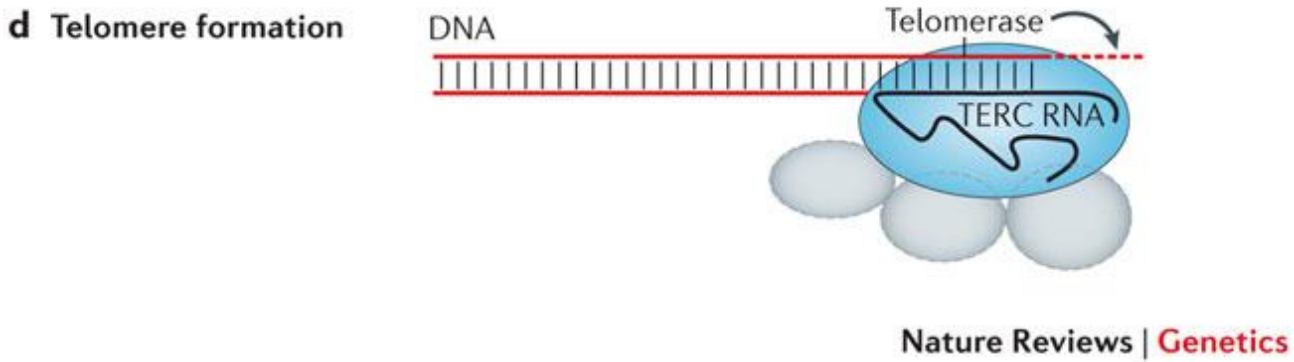
Design guide RNA construct to target and cleave anywhere in the genome

Provide desired construct to be integrated into the genome via homologous recombination



TERC: telomerase RNA component

Serves as template for the reverse transcription conducted by telomerase



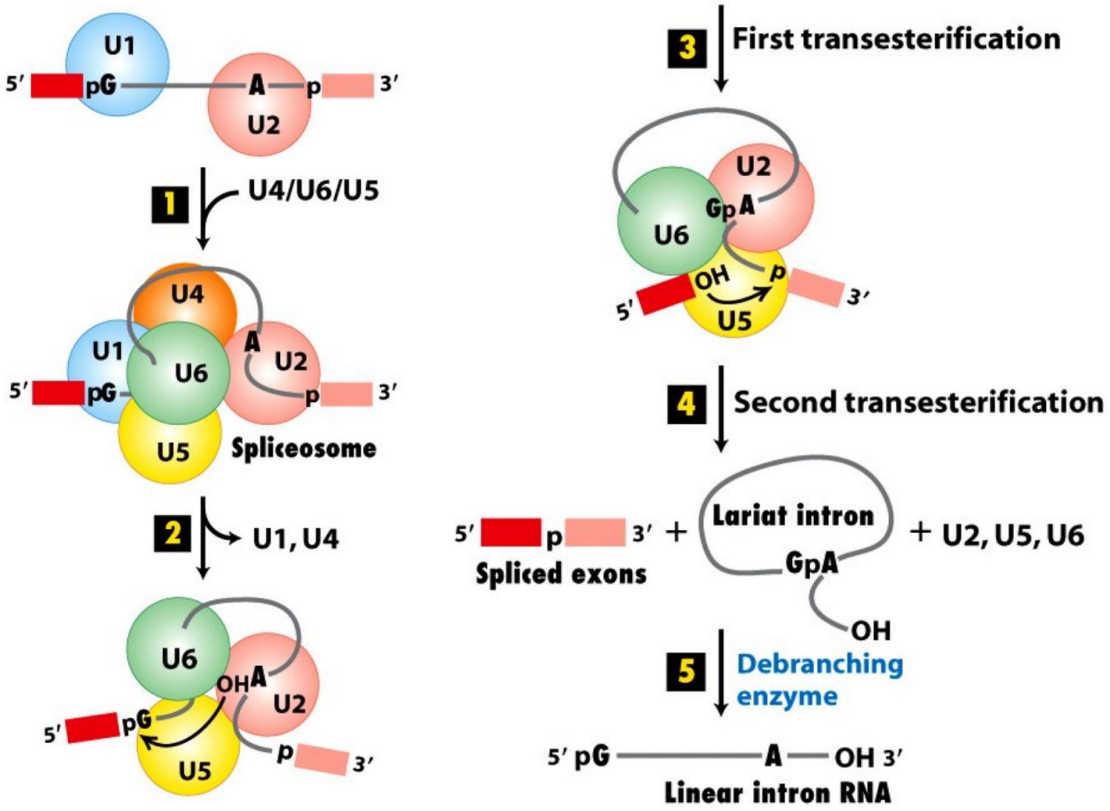
snRNAs found in snRPS

U1: 5’ for assembly

U2: at 3’ branch site

U5: for second attack

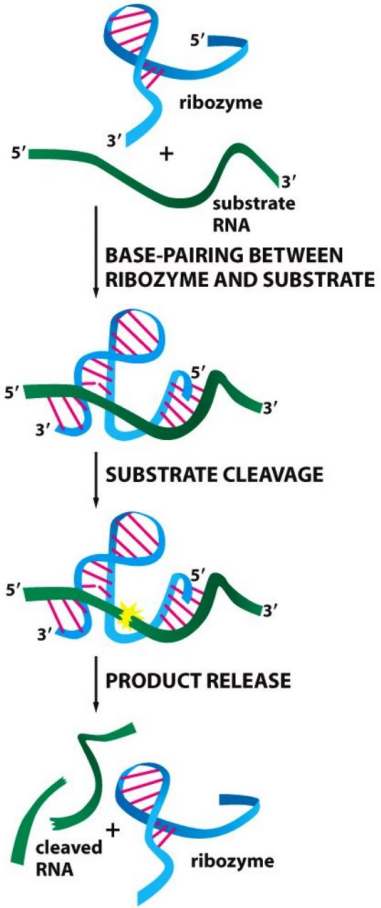
U6: 5’ for all catalytic events



RNA as enzymes (ribozymes)

2’OH reactivity makes this possible

Some RNAs can even cleave themselves given the right sequence/structure

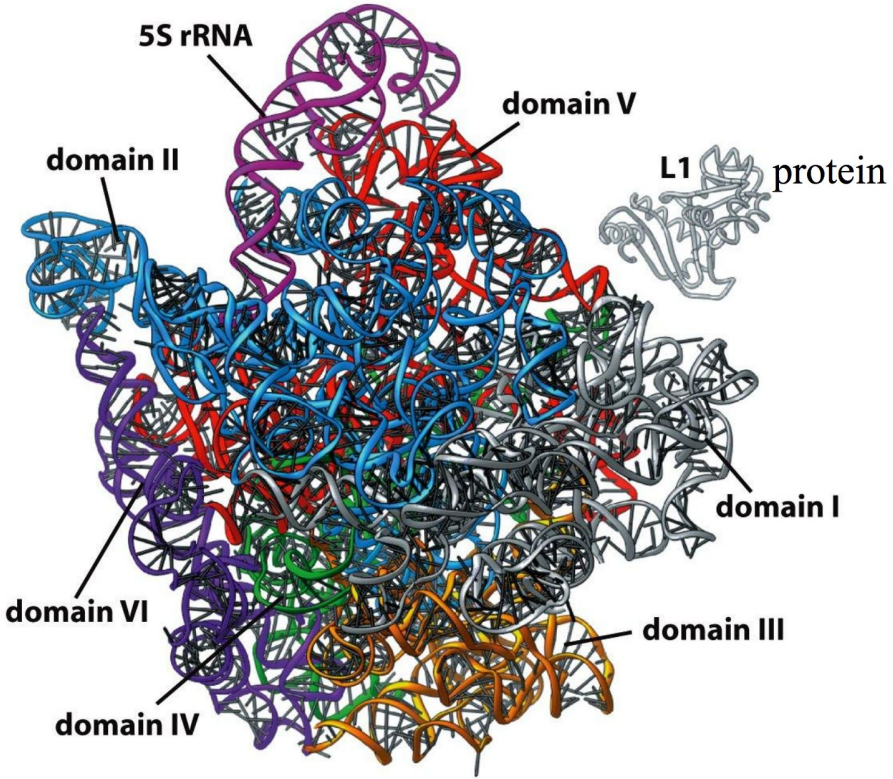


rRNA

Provides ribosome structure

Quality control for codon / anticodon pairing

Ribozyme activity responsible for addition of amino acid to peptide chain



RNA summary

Functions as:

- Regulators (RNAi, lncRNAs)
- Defense mechanisms against viral DNA (RNAi, crRNA)
- Enzymes (ribozymes, rRNA, snRNA)
- Templates (TERC, mRNA)

Techniques

DNA-protein interactions

Testing known protein and known DNA:

- EMSA or DNase Footprinting

Testing known protein for unknown DNA binding partner:

- Chip-seq or SELEX

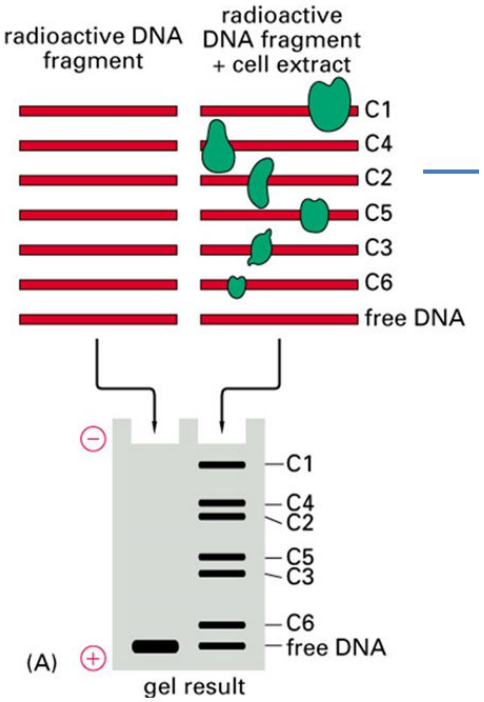
EMSA

Used to detect DNA-protein interactions

DNA and protein run out **together** on the gel

But only DNA is visualized

Tells you **if** the protein binds, but not where

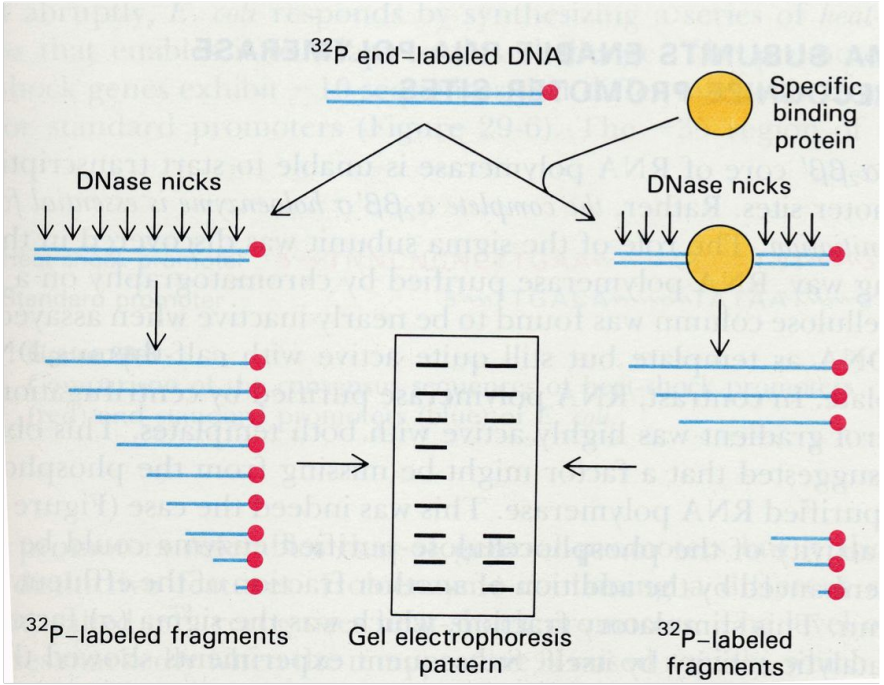


DNase Footprinting

Used to detect DNA-protein interactions

Only DNA is run out on the gel

Tells you **where** the protein binds.



CHIP-seq

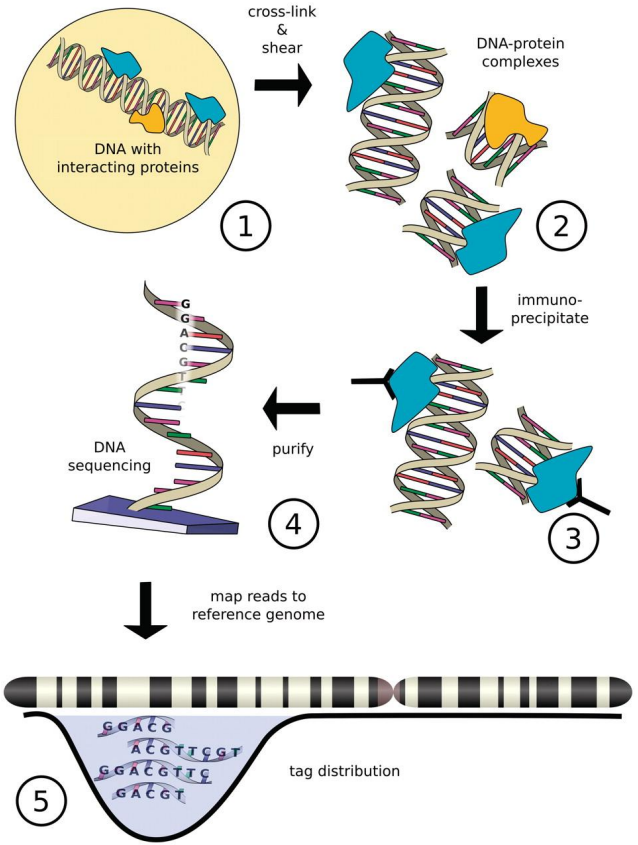
Used to detect protein-DNA interactions

Pull down protein and **sequence DNA** that was bound

Whole-genome approach

Good as a screen for where a protein binds and what sequence it prefers

Can also be used to detect where certain chromatin marks are found



Limitations of ChIP-seq

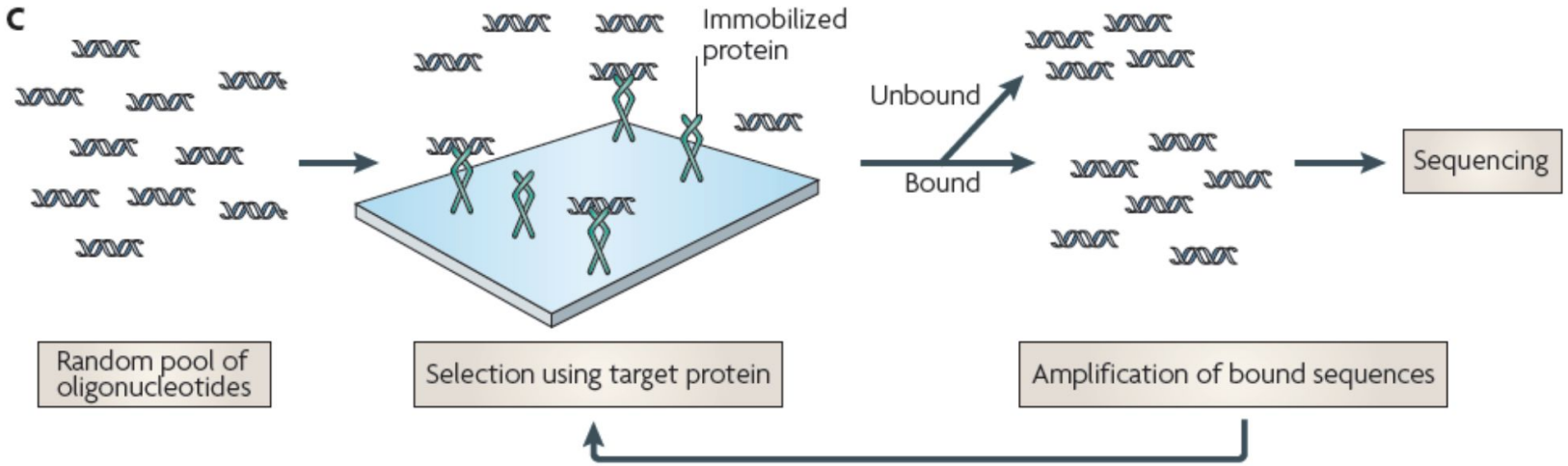
Although ChIP-seq data show that a specific transcription factor binds to cis-regulatory region of a gene, it DOES NOT mean that the gene is expressed at that moment. (Commonly, activation of a gene needs cooperativeness of several TFs.)

Only for detecting where a **known** protein binds. Can’t screen in the opposite direction, where DNA is known and we want to discover the protein(s).

SELEX

Test any sequence you would like, not just binding sites in the genome

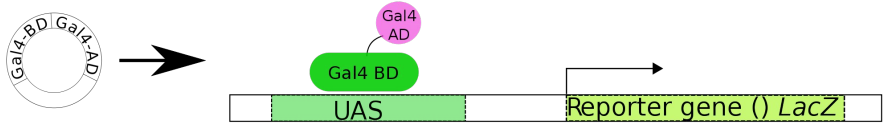
Once again, good as a screen for what sequences a given protein binds



Yeast 2 hybrid assay

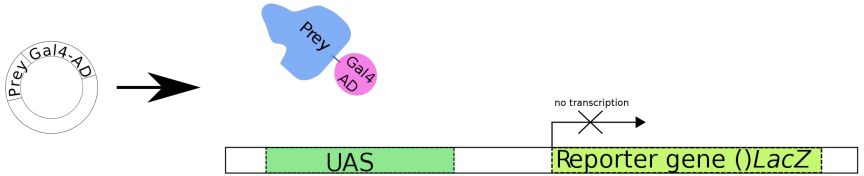
Used to detect known protein (bait) binding to unknown protein (prey)

Gal4 binding domain / bait protein --- prey protein / Gal4 activating domain

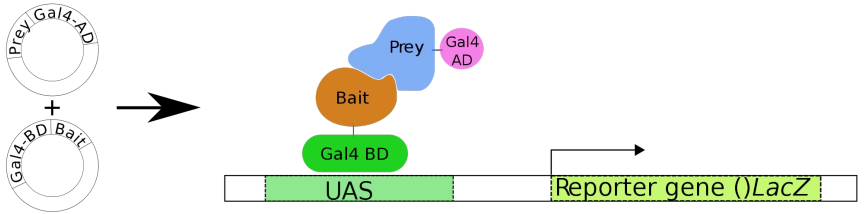


A. Regular transcription of the reporter gene

B. One fusion protein only (Gal4-BD + Bait) - no transcription



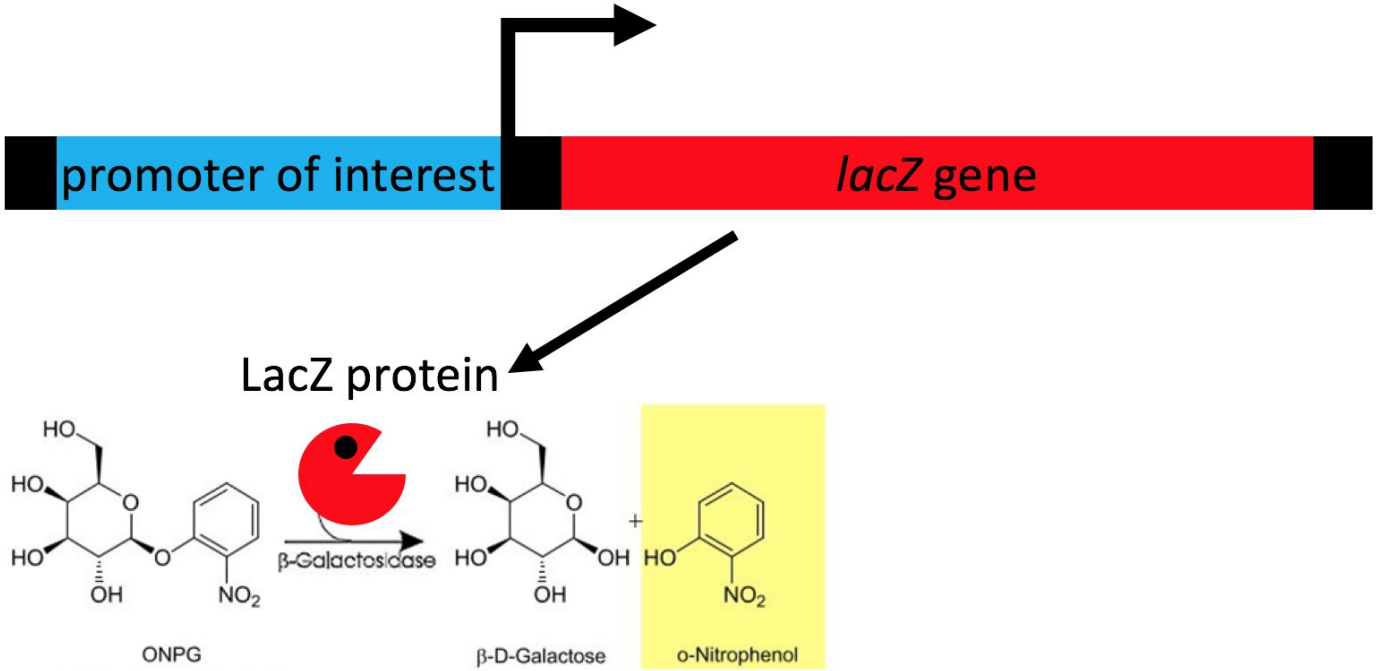
C. One fusion protein only (Gal4-AD + Prey) - no transcription



D. Two fusion proteins with interacting Bait and Prey

β-galactosidase assay

Used to detect **activity of a promoter of interest**



Actual β-galactosidase assay results!

More yellow →

more o-nitrophenol →

more LacZ activity →

more *lacZ* transcription →

stronger promoter