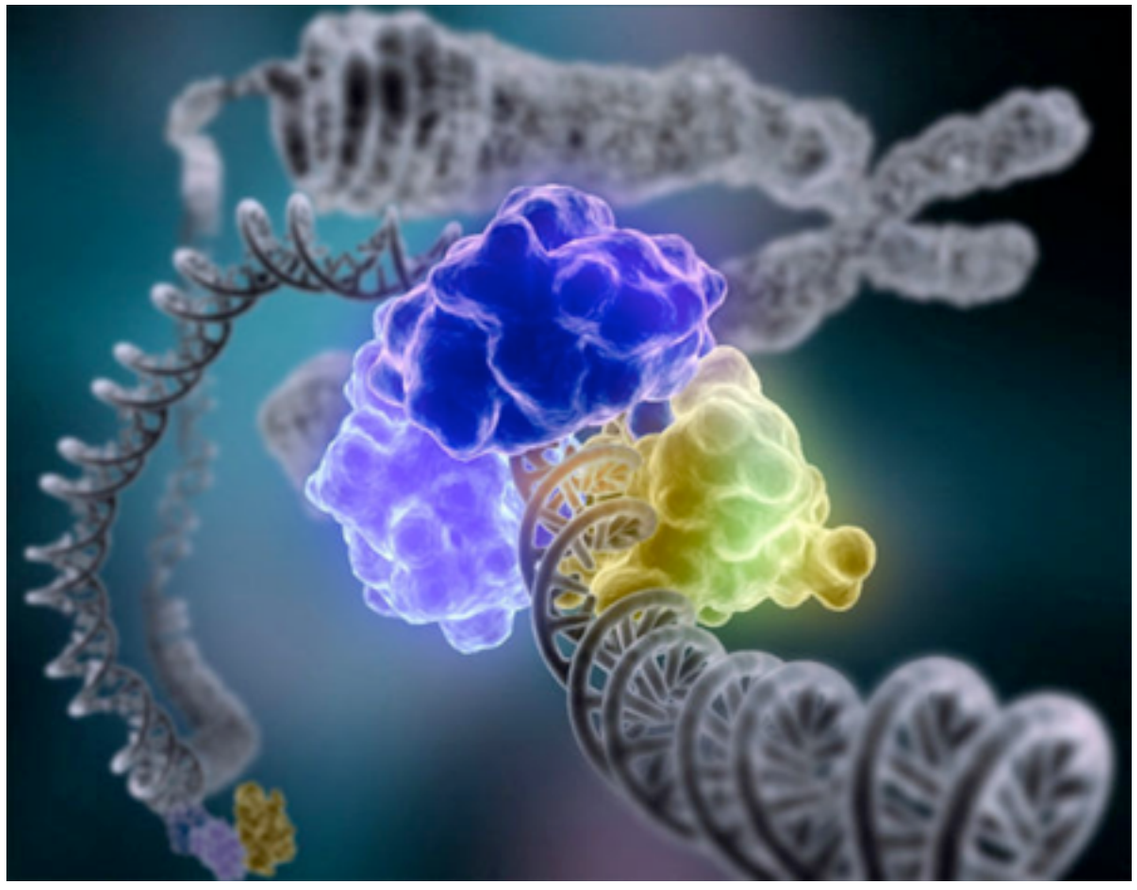
# Eukaryotic Genetics

## Expression & Regulation

# Molecular Basis of Inheritance

I.Main Idea: Bacteria often respond to environmental change by regulating transcription.



# Regulating Gene Expression

- A cell's genome consists of all its genes.
- NOT ALL genes need to be expressed at ALL times.
- ONLY certain genes are expressed at certain times.
- Bacteria turn on and turn off genes in response to the environmental conditions.
- As environmental conditions change so to does gene expression.

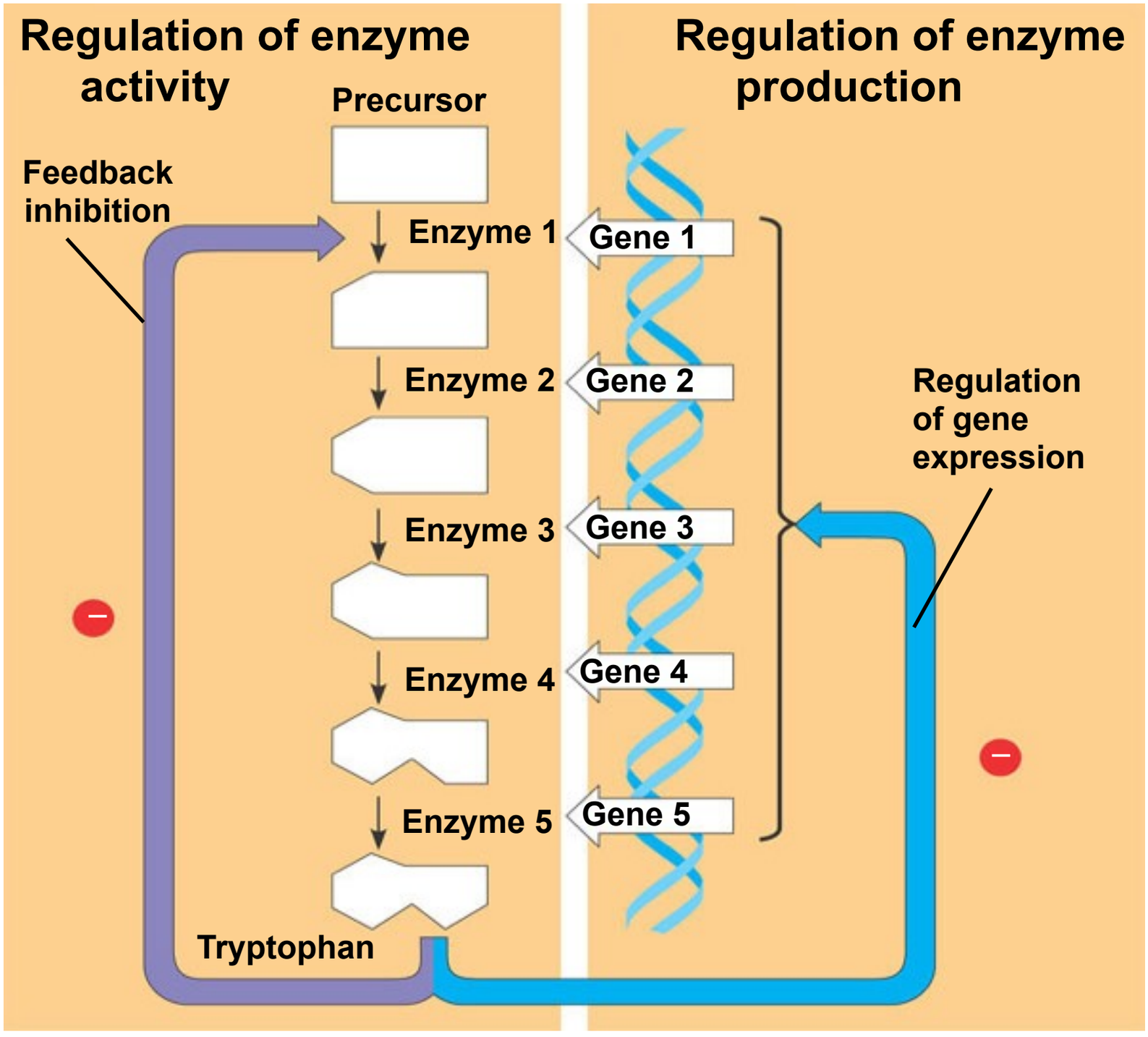
# Regulating Gene Expression

- A bacteria that can turn its genes on and off in response to environmental changes **will save both resources and energy over time**.
- Natural selection has favored these bacteria over those which have less control.

- *Consider E-coli that live in a human colon, if human meal includes a particular nutrient then they need not produce it (save energy ) BUT if human meal does not include a particular nutrient then they need produce it.*
- This fundamentally requires that the E-coli turn on/off certain genes depending on the presence/absence of a particular nutrient.

# How are bacterial genes controlled?

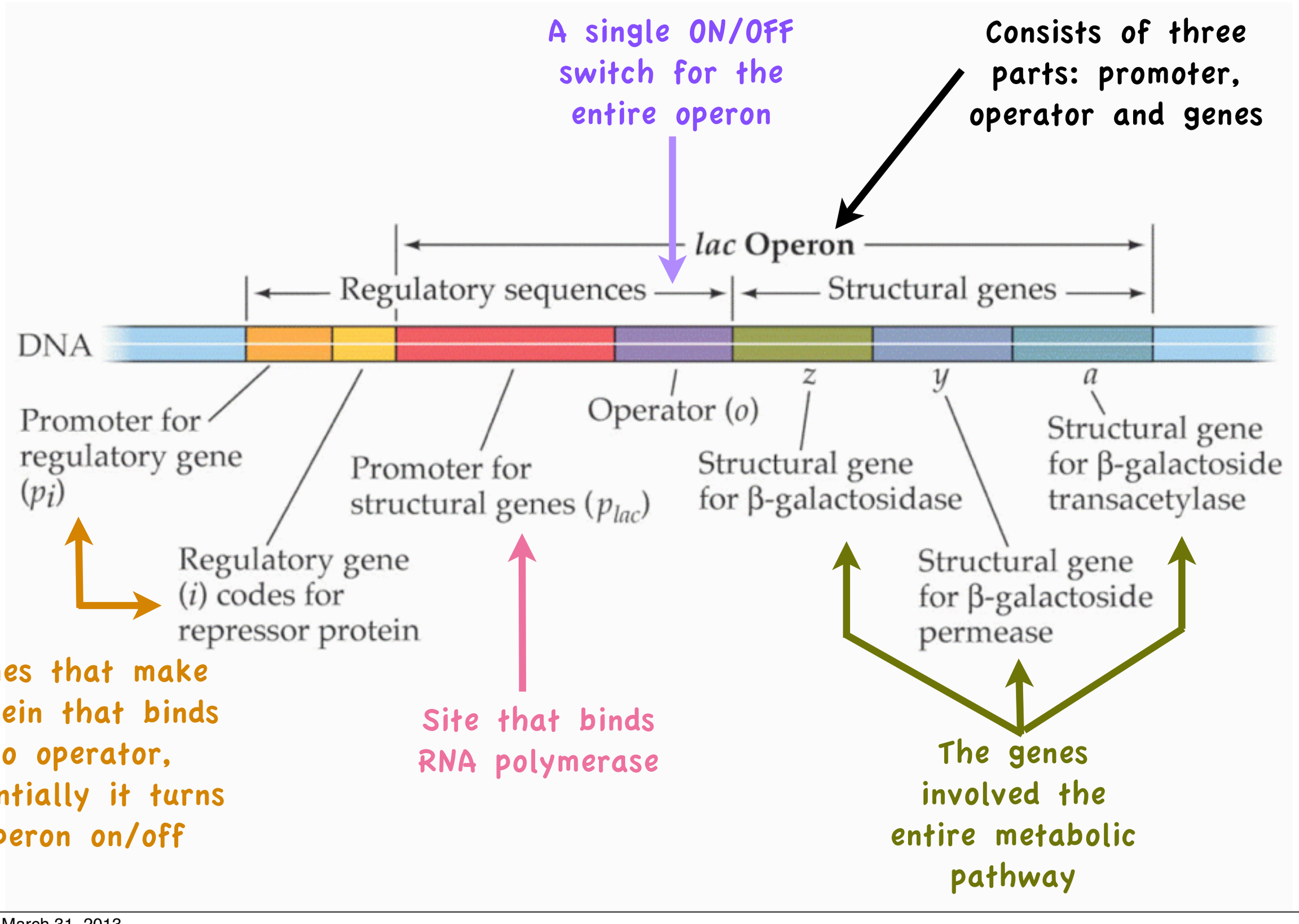
"order from the bakery case"




"place a custom order"

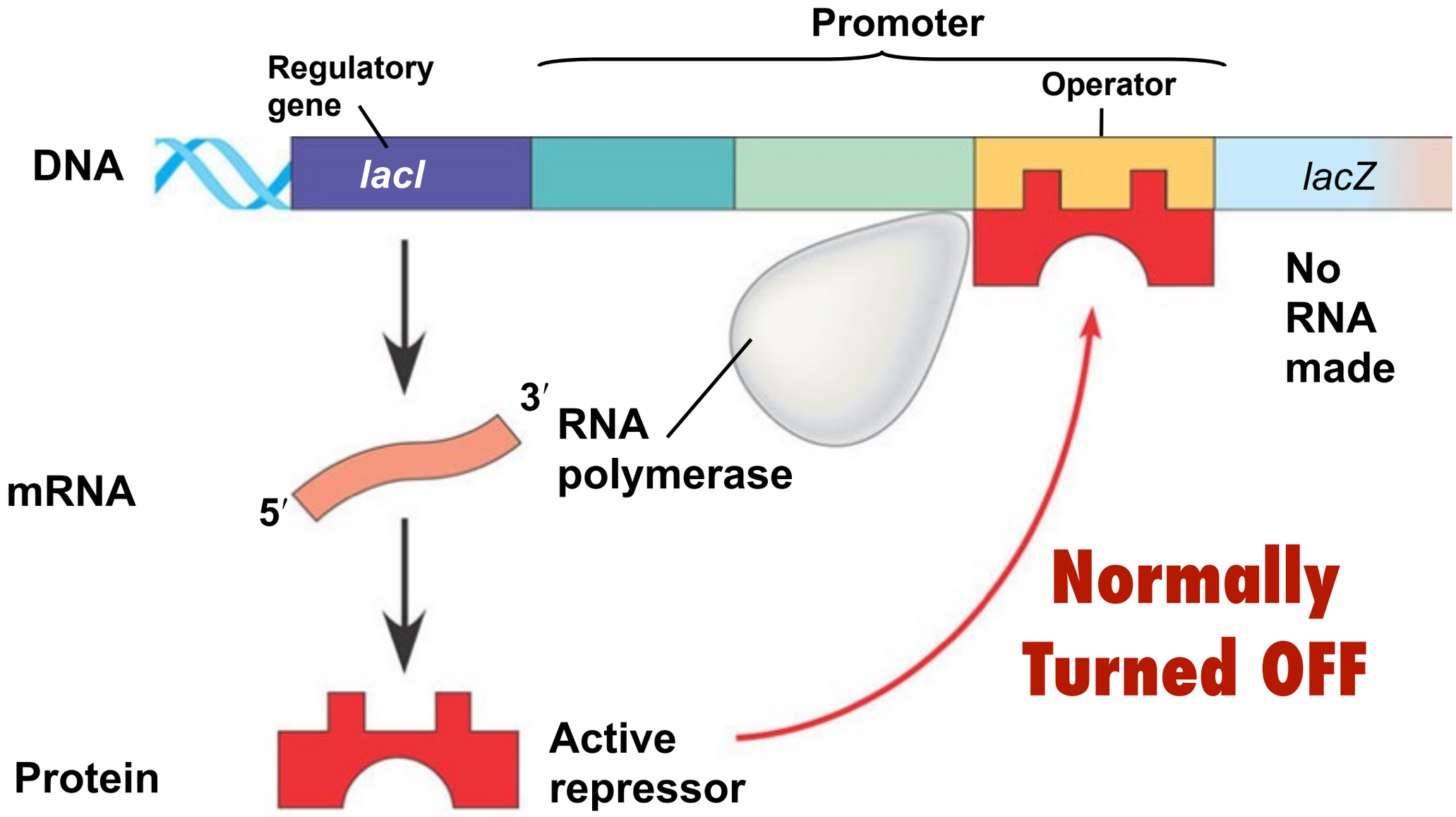
"two general ways of eliciting a cellular response"

# Bacterial Operon Concept



# Inducible Operons "turn-on-able"

## Negative Gene Regulation

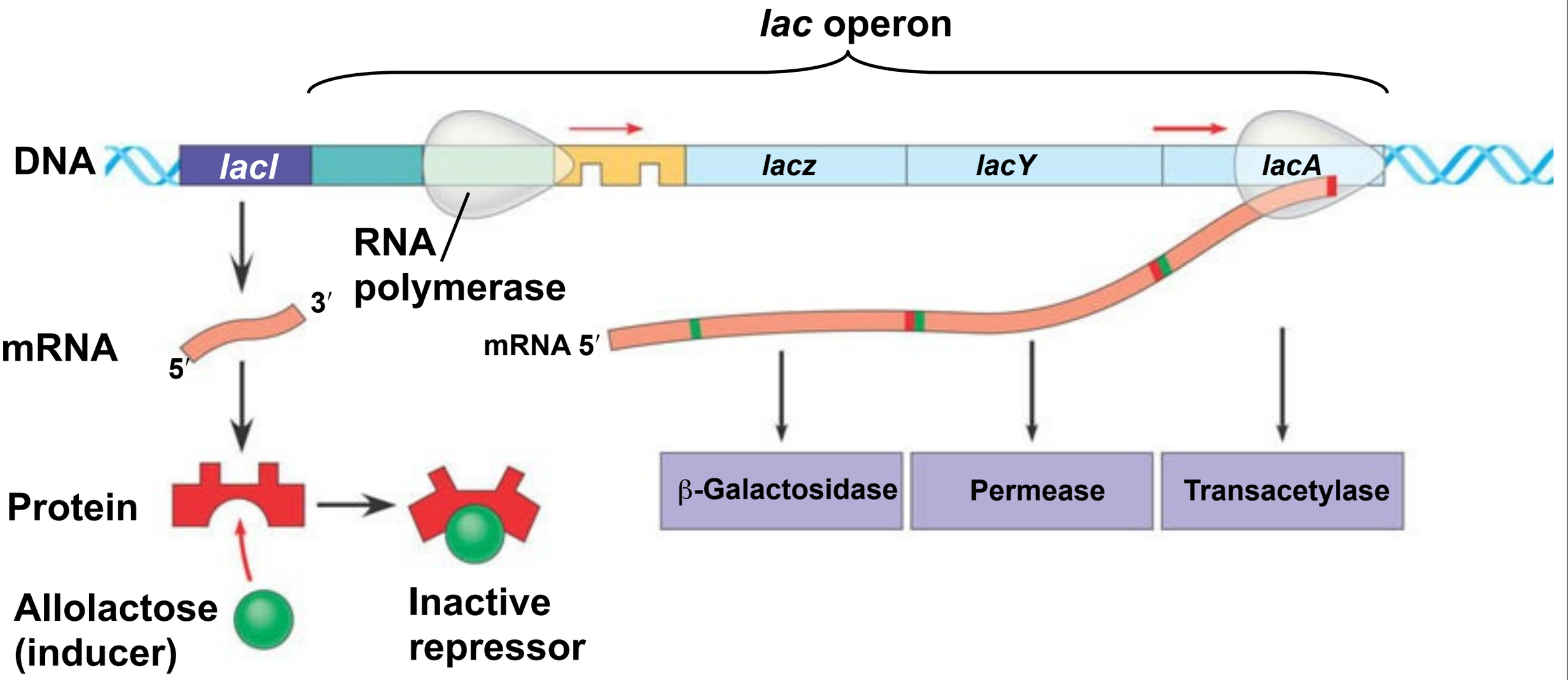



**Lactose absent, repressor active, operon off.** The *lac* repressor is innately active, and in the absence of lactose it switches off the operon by binding to the operator.

# Inducible Operons "turn-on-able"

## Negative Gene Regulation

**Normally Turned OFF**



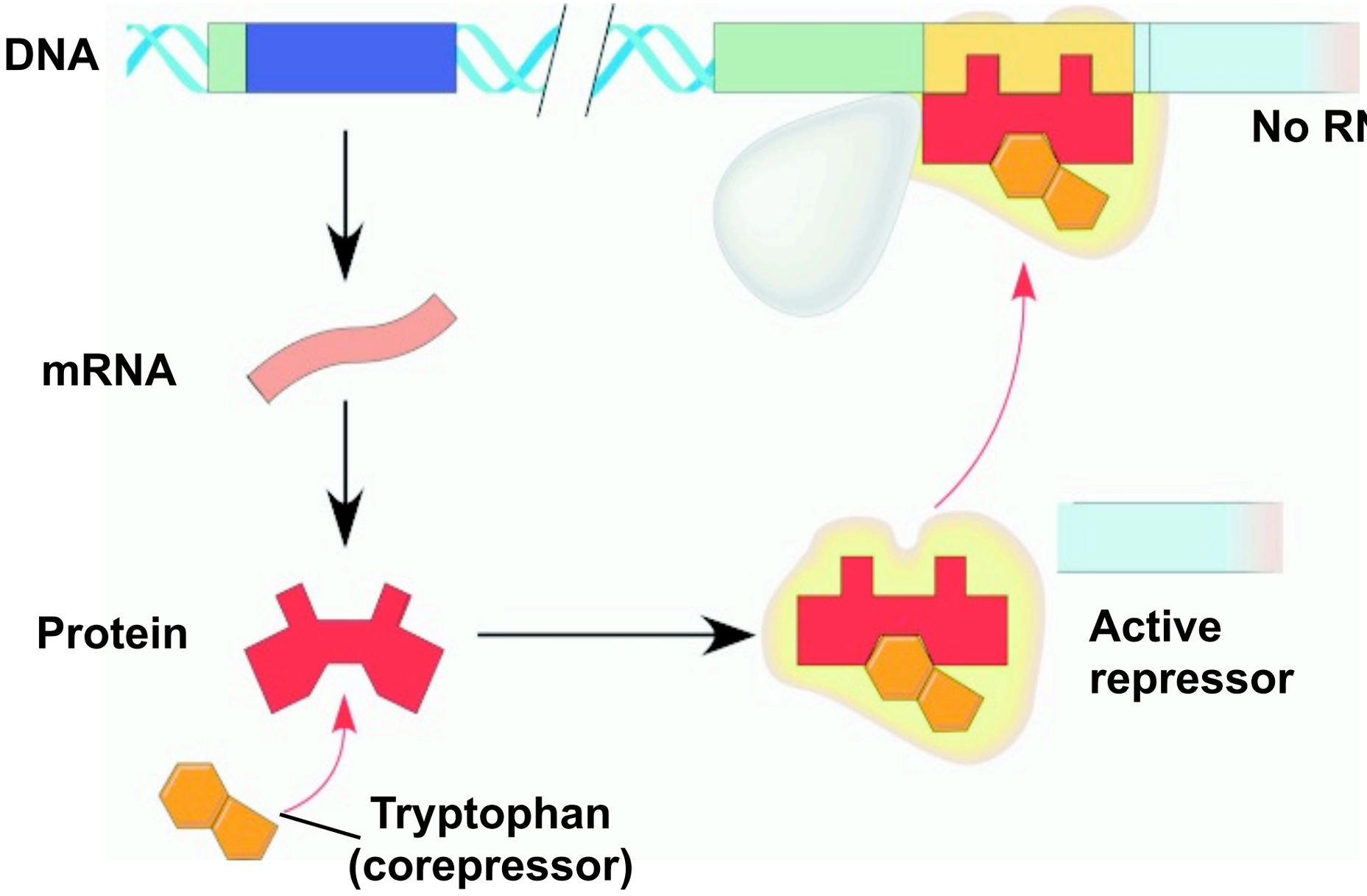

**Lactose present, repressor inactive, operon on.** Allolactose, an isomer of lactose, derepresses the operon by inactivating the repressor. In this way, the enzymes for lactose utilization are induced.

# lac Operons "turn-on-able"

# Repressible Operons "turn-off-able"

## Negative Gene Regulation

**Normally Turned ON**

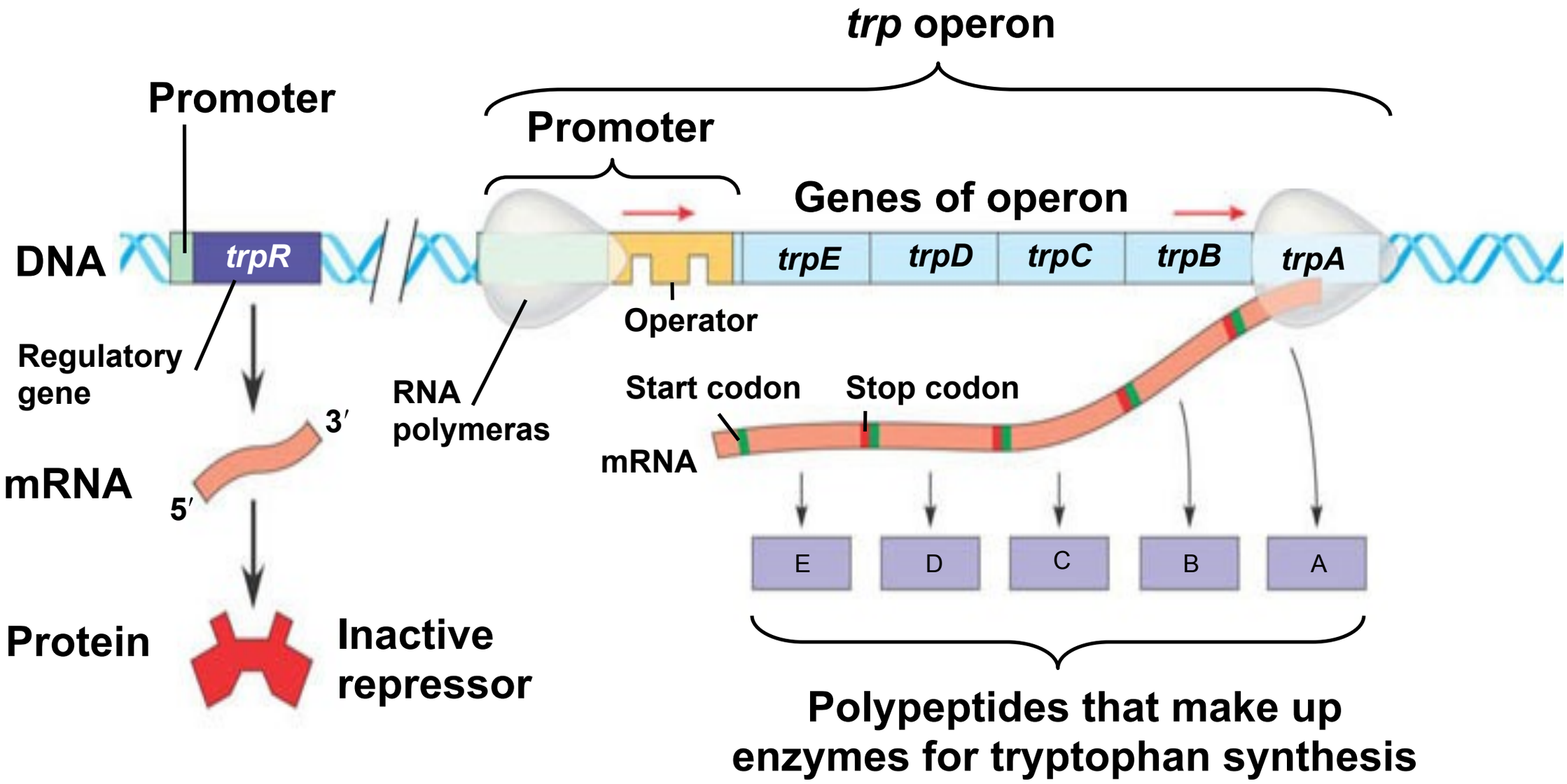



**Tryptophan present, repressor active, operon off. As tryptophan accumulates, it inhibits its own production by activating the repressor protein.**

# Repressible Operons "turn-off-able"

## Negative Gene Regulation

**Normally Turned ON**




**Tryptophan absent, repressor inactive, operon on. RNA polymerase attaches to the DNA at the promoter and transcribes the operon's genes.**

# Let's Review so far

- **Inducible Operons ("turn-on-able")**
  - inducer binds to innately active repressor there by inactivating the repressor and turning operon on
  - usually operates in catabolic pathways
- **Repressible Operons ("turn-off-able")**
  - repressor binds to innately inactive repressor there by activating the repressor and turning operon off
  - usually operates in anabolic pathways

# Let's Review so far

- **Negative Gene Control**
  - both repressible and inducible operons operate as negative control
    - because operons are switched off by an "active" form of the repressible protein
- **Positive Gene Control**
  - some operons operate under positive control
  - these use activator proteins that have a stimulatory effect

# Preface to Positive Gene Control

- **Negative Gene Control-** uses "repressors" to turn operons on/off
- **Positive Gene Control-** uses "activators" to turn the volume up/down
- Bacteria prefer glucose to produce ATP via glycolysis but they can use other sugars (ex. lactose)
  - Enzymes used to catabolize glucose are always present
  - How does cell "know when glucose is plentiful or scarce?
  - If glucose is plentiful then ATP>cAMP
  - But if glucose is scarce then ATP<cAMP

# Preface to Positive Gene Control

- **Here are scenarios bacteria might deal with**
  - **high glucose/high lactose**
    - **glucose enzymes active, volume low, lac operon turned on**
  - **high glucose/low lactose**
    - **glucose enzymes active, volume low, lac operon turned off**
  - **low glucose/high lactose**
    - **glucose enzymes inactive, volume high, lac operon turned on**
  - **low glucose/low lactose**
    - **glucose enzymes inactive, volume high, lac operon turned off**

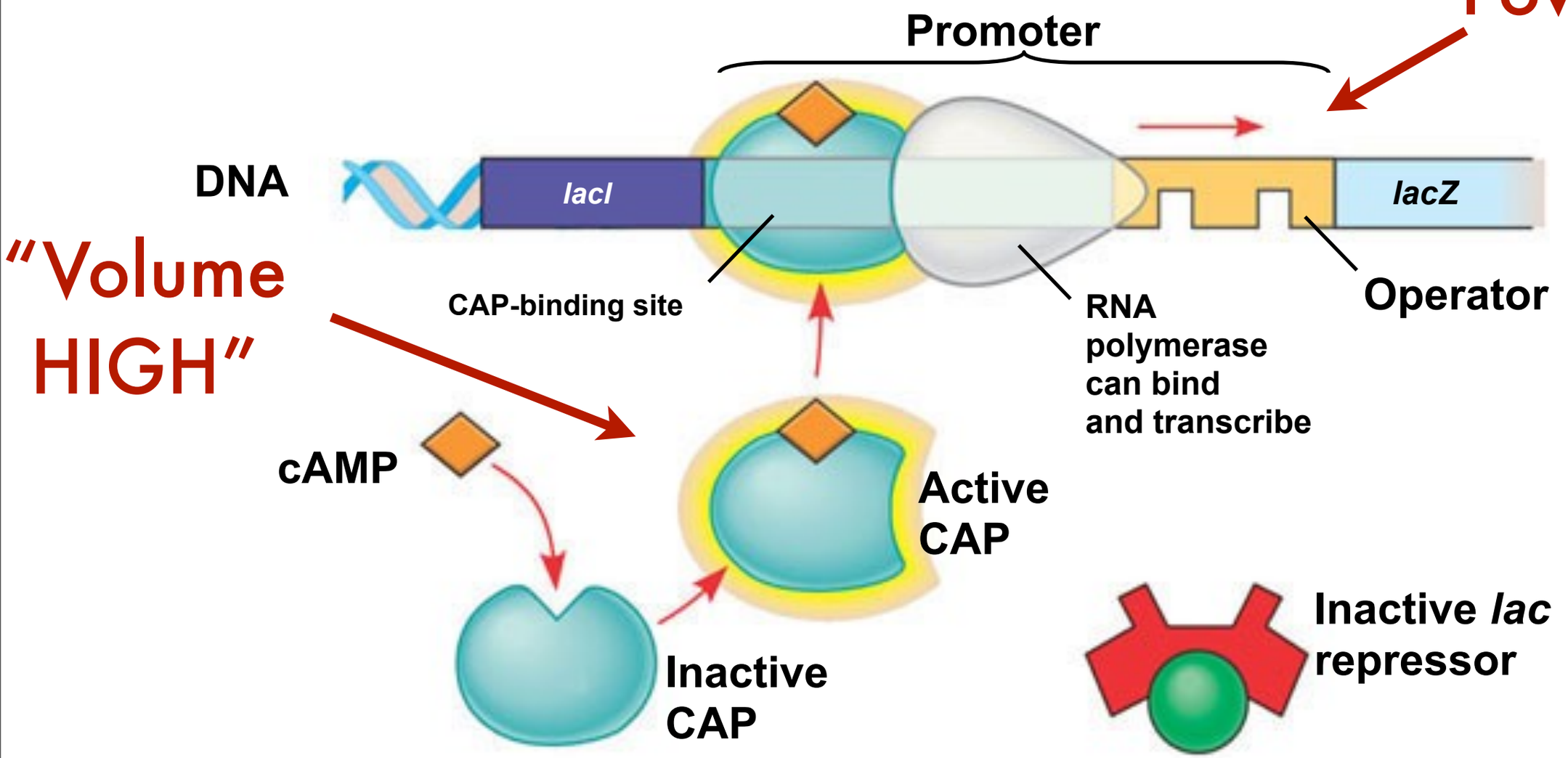
# Inducible "turn-on-able" *lac* operon

Negative Gene Regulation- on/off switch

Positive Gene Regulation- volume control

"Power ON"

"Volume HIGH"

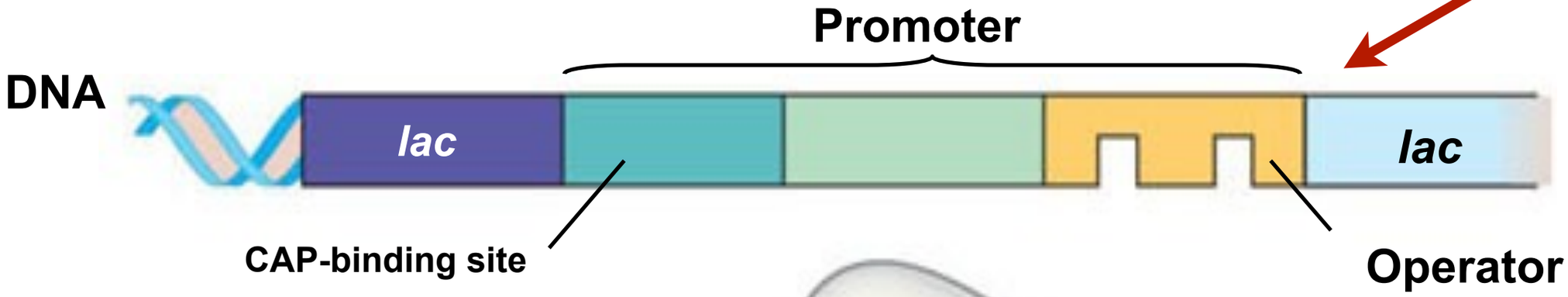



**Lactose present, glucose scarce (cAMP level high): abundant *lac* mRNA synthesized. If glucose is scarce, the high level of cAMP activates CAP, and the *lac* operon produces large amounts of mRNA for the lactose pathway.**

# Inducible "turn-on-able" *lac* operon

Negative Gene Regulation- on/off switch
Positive Gene Regulation- volume control

"Power ON"




RNA
polymerase
binds less effectively

"Volume LOW"

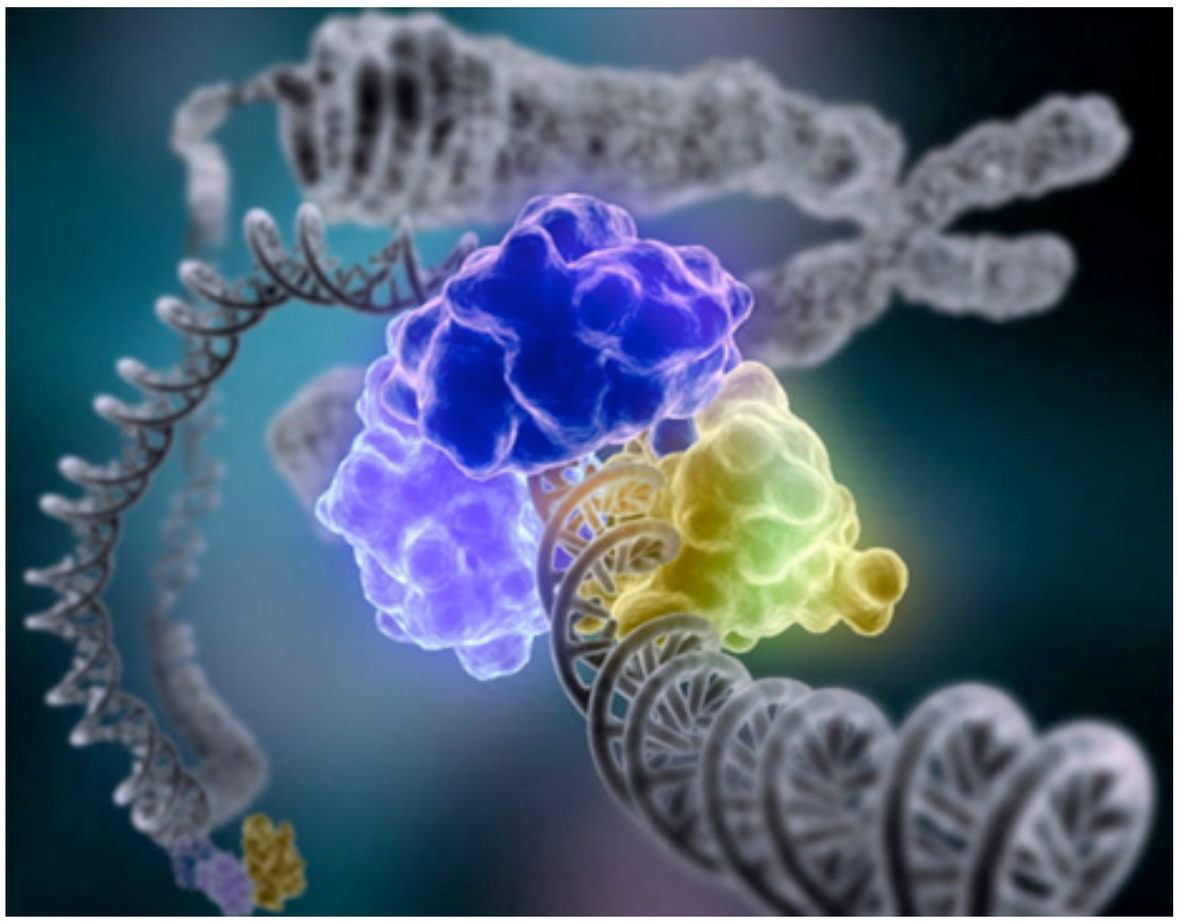
Inactive
CAP

Inactive
*lac*

**Lactose present, glucose present (cAMP level low): little *lac* mRNA synthesized. When glucose is present, cAMP is scarce, and CAP is unable to stimulate**

# Molecular Basis of Inheritance

II.Main Idea: Regulating gene expression is also important for eukaryotic cells, especially those that make up a multicellular organism.

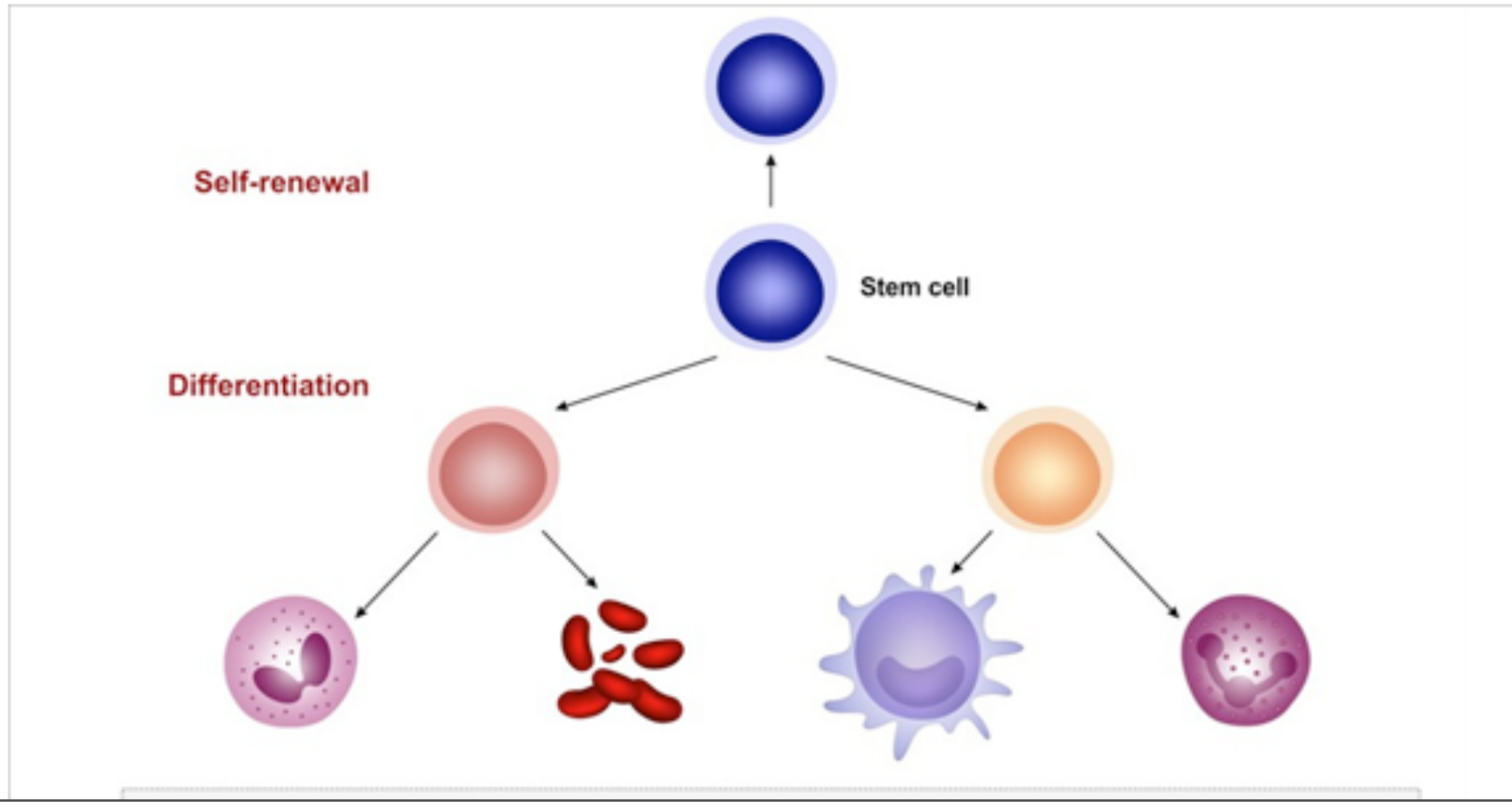


# Differential Gene Expression

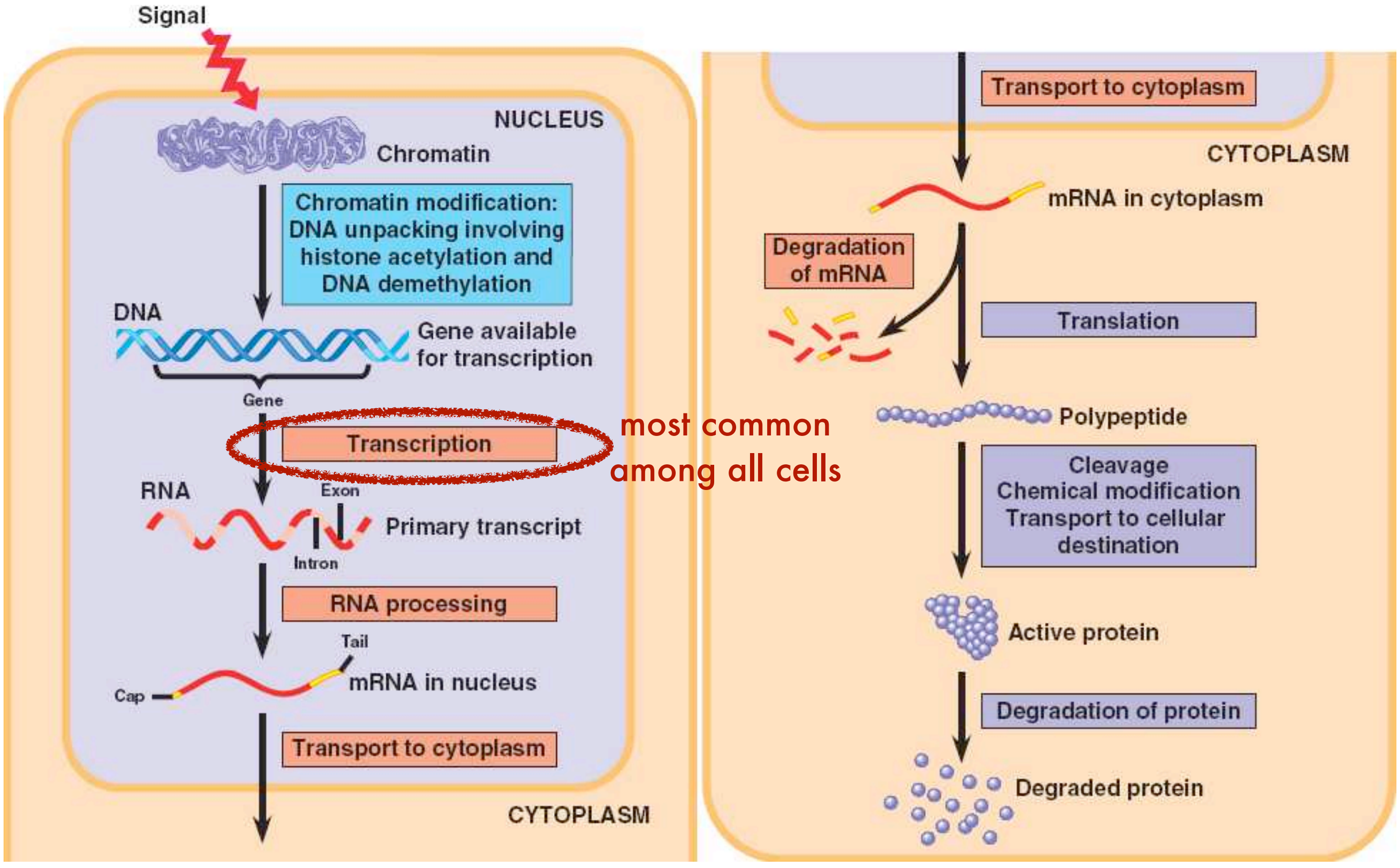
- Typical human cells expresses ~20%of its genes at any given time
  - highly differentiated cells like nerve cells even less
- Almost all cells have an identical genome
  - immune cells are an exception
- The difference between cell types are not due to different genes present but to differential gene expression, each cell using its own unique combination of genes.

# Differential Gene Expression

- The function of any cell depends on the appropriate set of genes being expressed.
- Transcription factors must locate the “right genes” at the “right time”.
  - if this does not occur imbalances and disease can result



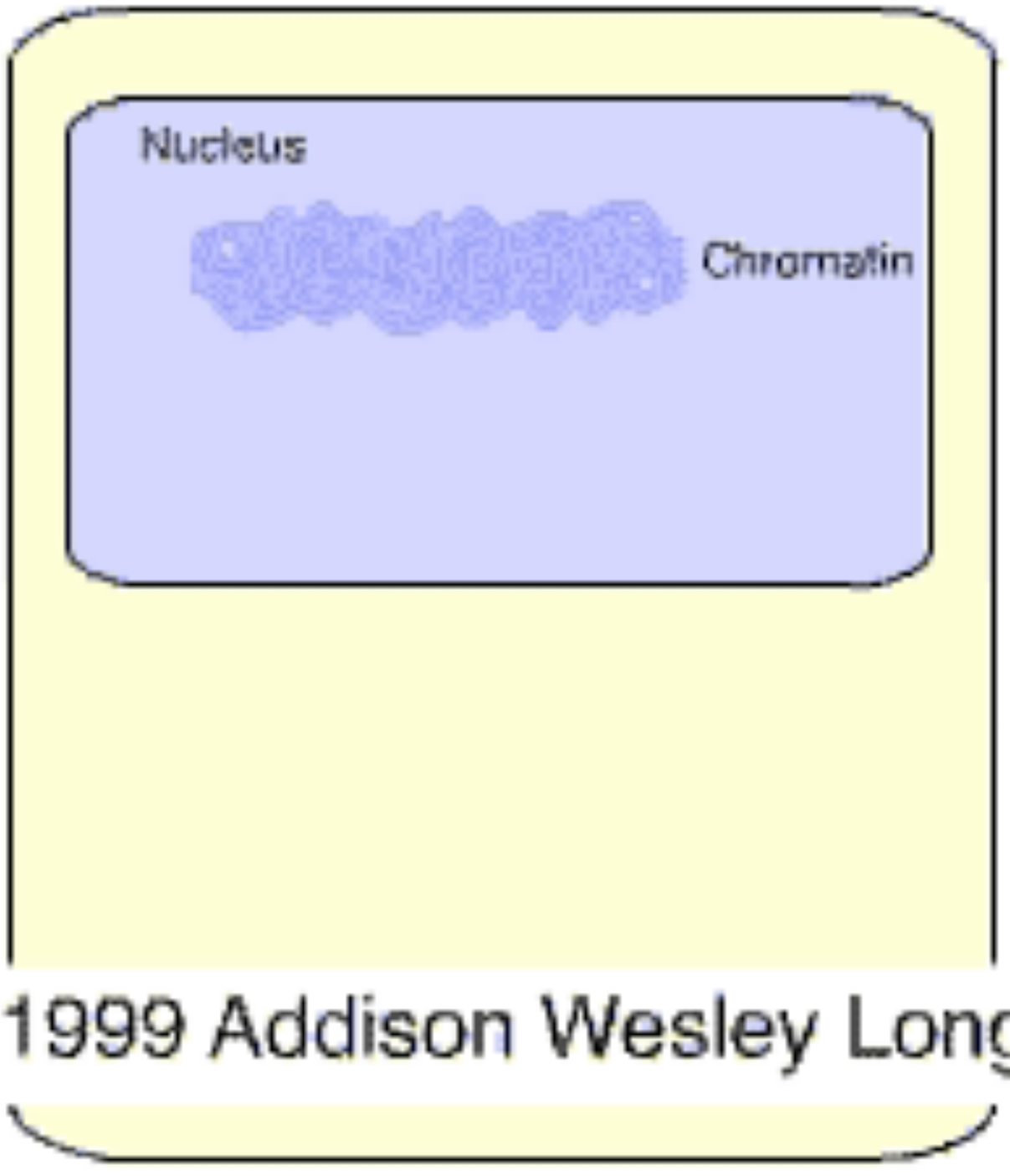
# Differential Gene Expression



# Regulation of Chromatin

- Chromatin not only packs a cell's DNA into a compact form to fit inside the nucleus, but also helps in gene regulation in several ways.
  - 1. The location of the genes promoter relative to the nucleosomes can affect a genes activity.
  - 2. Genes in the heterochromatin form (highly condensed) are usually not expressed.
  - 3. Certain chemical modifications to histone proteins and to DNA itself can affect chromatin structure and gene expression.

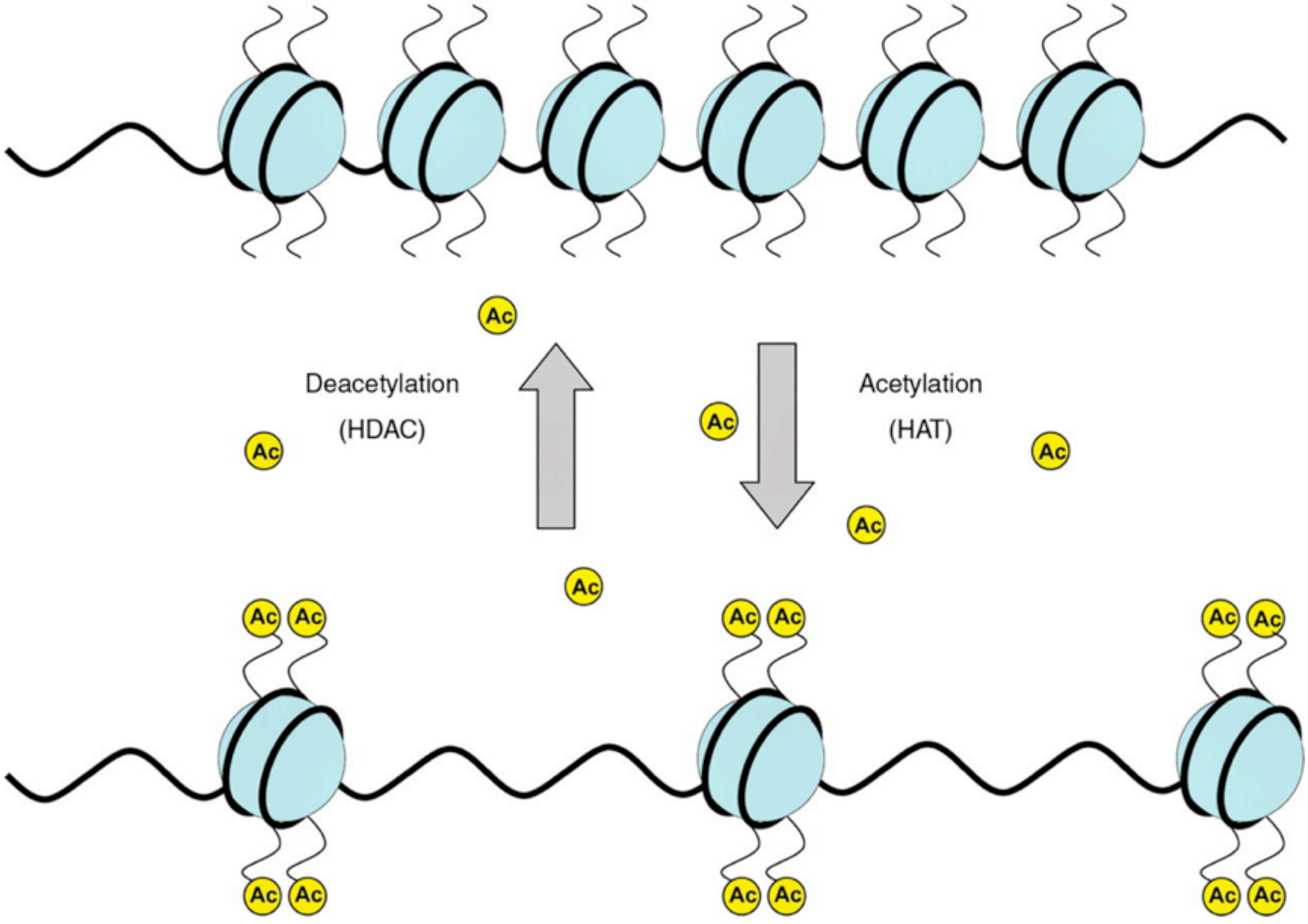
# Regulation of Chromatin



# Regulation of Chromatin

- <u>Histone Modifications</u>
  - Histone acetylation adds acetyl groups (-$COCH_3$) to histone tails, deacetylation removes them
  - When histones are acetylated they no longer bind to neighboring nucleosomes = looser chromatin = increase gene transcription!
  - When histones are not acetylated they bind to neighboring nucleosomes = dense chromatin = decrease gene transcription!

# Less accessible "genes turned off"



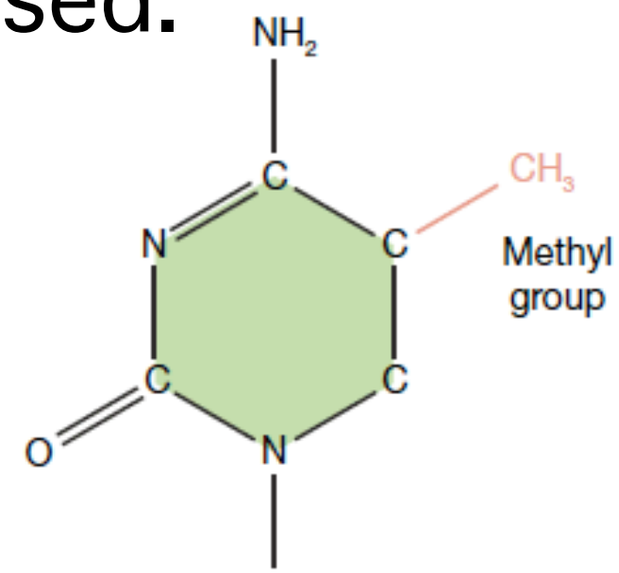

# Accessible "genes turned on"

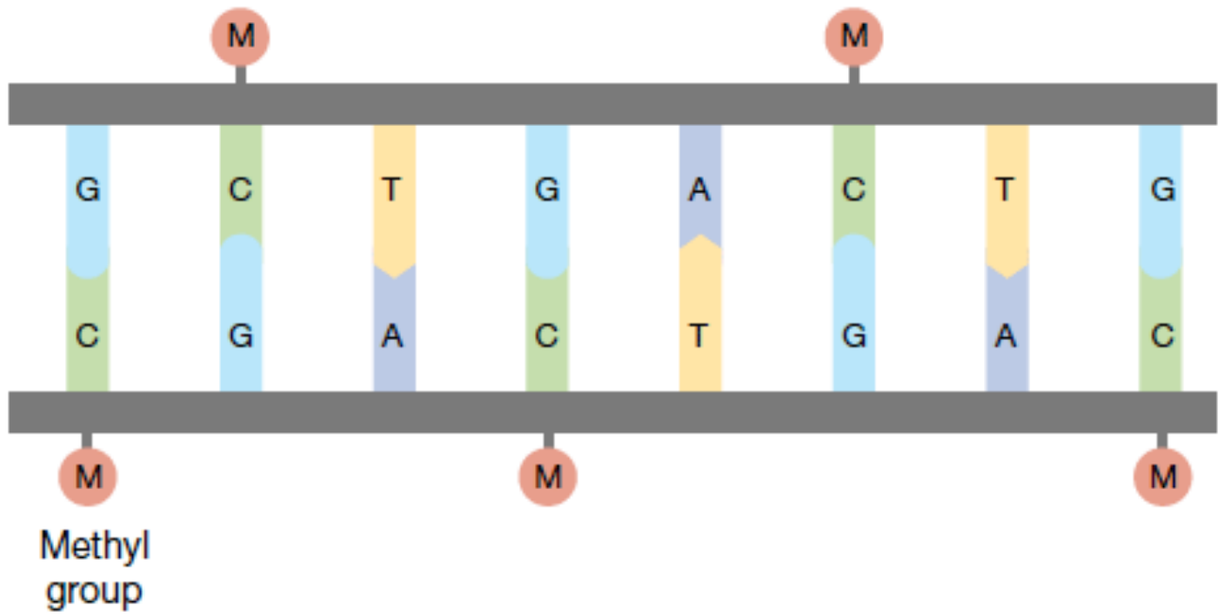
# Regulation of Chromatin

- Histone Modifications
  - Other groups can also attach to histone tails and affect chromatin and consequently transcription
  - Methyl groups (-$CH_3$) promote condensation (off switch)
  - Phosphate groups (-$PO_4$) have the opposite affect (on switch)
- Histone Code Hypothesis, specific combinations of these chemical groups and the order in which they occur help determine the chromatin configuration which in turn regulate transcription.

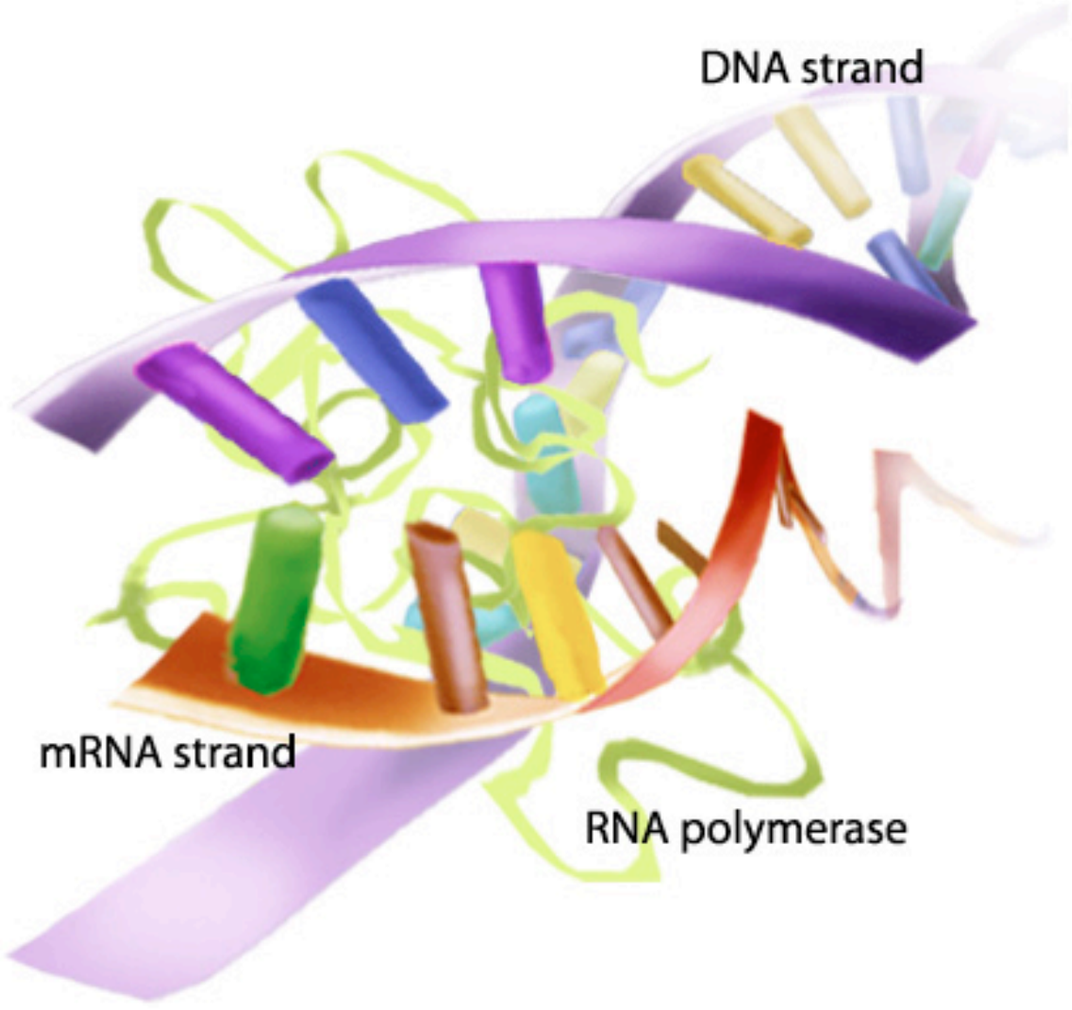
# Regulation of Chromatin

- DNA Methylation
  - The DNA itself can also be methylated.
    - occurs in plants, fungi and animals
    - responsible for X inactivation (Barr bodies)
  - Heavily methylated genes are not generally expressed.

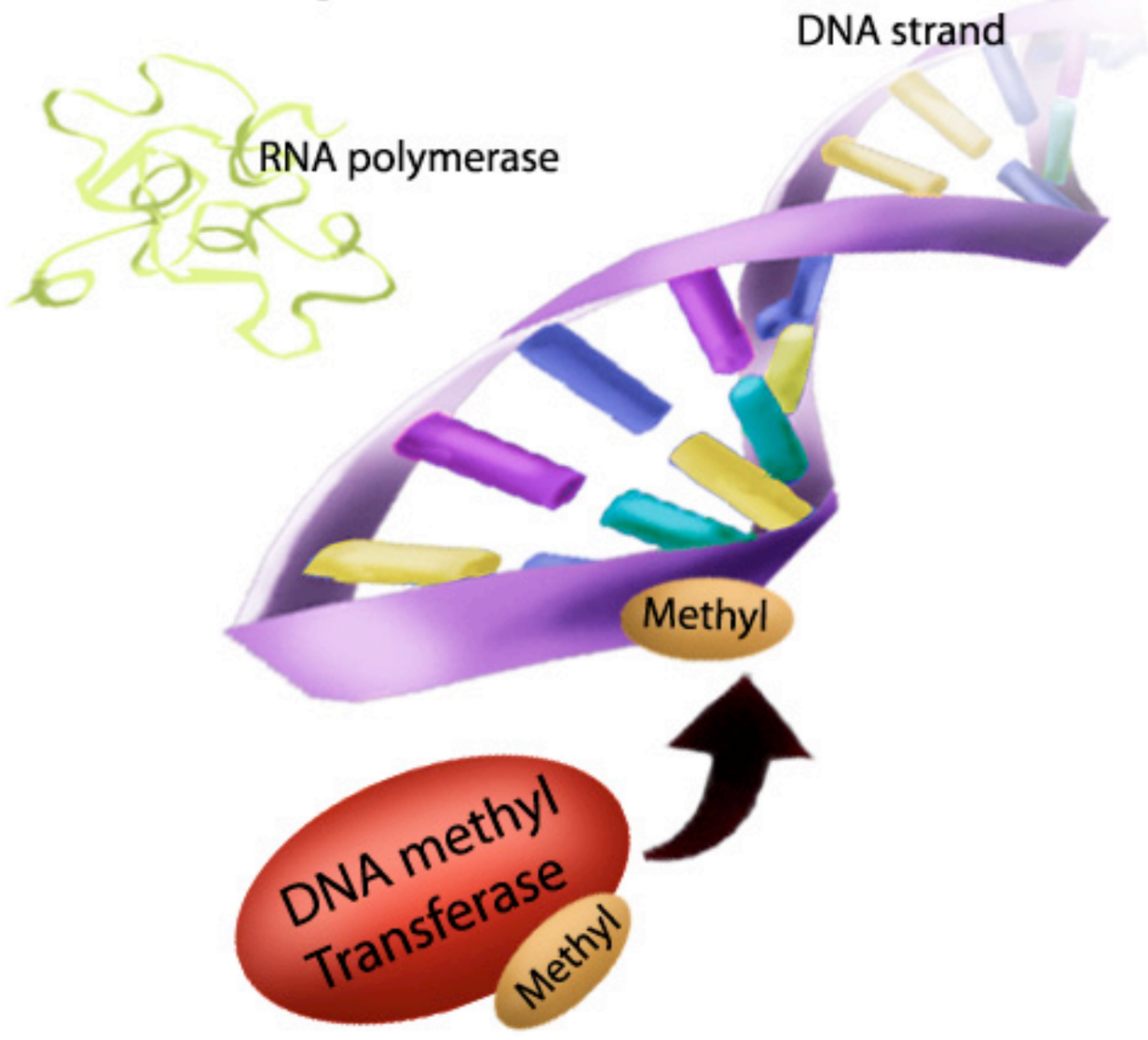


In DNA transcription, RNA polymerase unzips the DNA double helix and matches corresponding base pairs to synthesize an mRNA molecule.




DNA Methyl Transferase attaches methyl groups to the DNA, blocking transcription: a process called Methylation. In HD individuals, too much Methylation is occurring.



# Regulation of Chromatin

- DNA Methylation
  - DNA methylation appears to essential for the long term inactivation of genes that occurs in normal cell differentiation.
    - once methylated genes stay that way through successive cell divisions
    - these methylated patterns form the chemical record of cells forming specialized tissues
    - this methylation also accounts for the genomic imprinting in mammals, where maternal or paternal genes are turned on or off

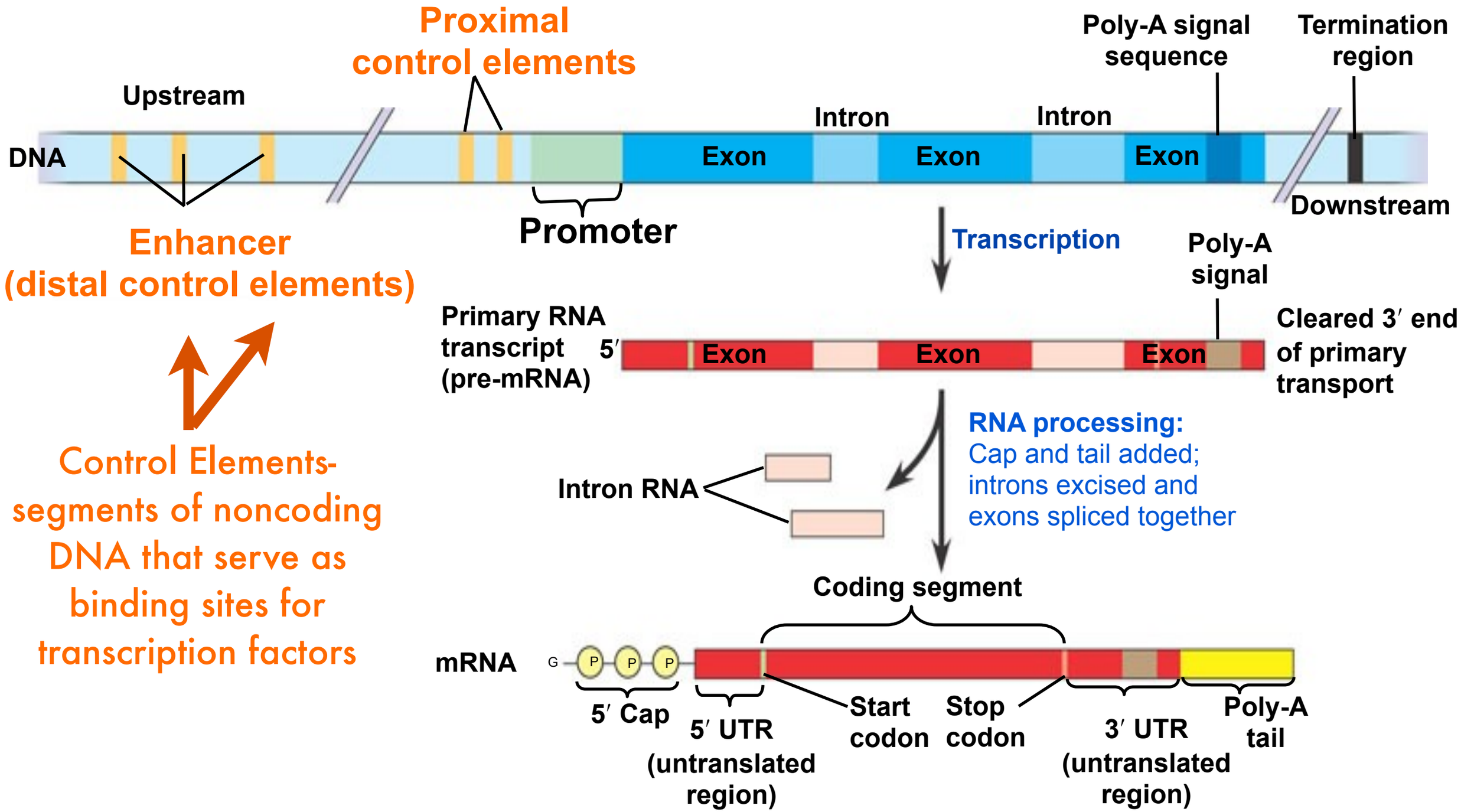
# Regulation of Chromatin

- Epigenetic Inheritance
  - These chromatin modifications may be passed done from generation to generation!
  - This means that the inheritance of traits may involve a mechanism(s) other than nucleotide sequences in DNA, this called epigenetic inheritance.
  - Unlike mutations that are not reversible these chemical modification pathways appear to be regulated and reversible.

# Regulation of Transcription

- The chromatin modifying enzymes made DNA more or less accessible,
- The next and most common place to regulate genes occurs at initiation step of transcription.
- This is true in both prokaryotes and eukaryotes!
- Since eukaryotic transcription is more complicated we should start by reviewing the structure of a eukaryotic gene.

- Organization of a Typical Eukaryotic Gene




Control Elements- segments of noncoding DNA that serve as binding sites for transcription factors

# Regulation of Transcription

- Roles of Transcription Factors
- Eukaryotes have a sort of volume control when comes to gene expression.
  - General Transcription Factors, bind to DNA, the promoter and the RNA polymerase (this what you saw in last powerpoint) = Low Volume
  - Specific Transcription Factors, bind to DNA, the promoter, the RNA polymerase and the control elements (seen on the last slide) = High Volume

# Regulation of Transcription

- Enhancers & Specific Transcription Factors
- Enhancers can greatly affect the rate of transcription depending upon the type of specific transcription factors.

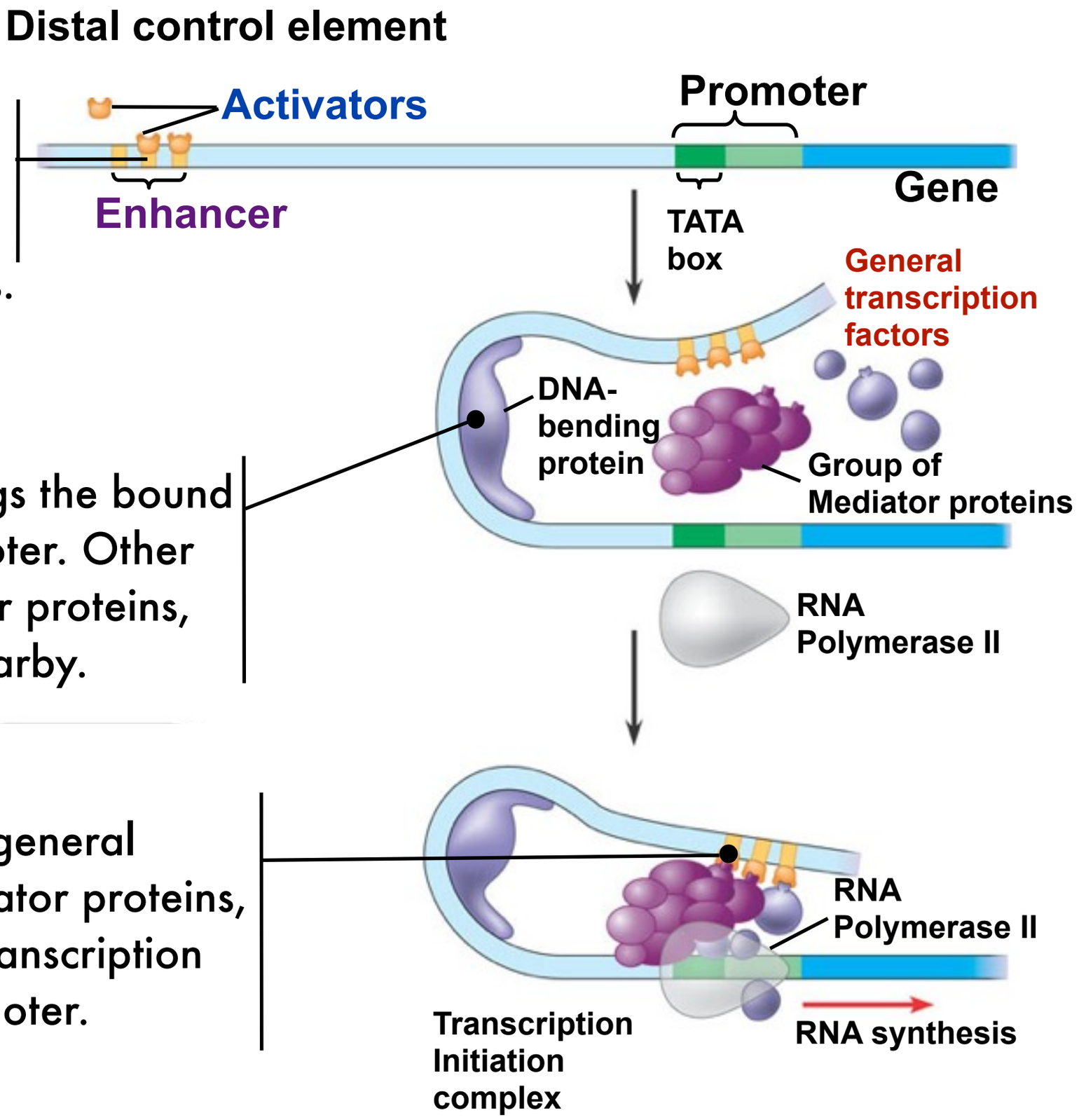
- Repressors, bind to the control elements of the enhancer = turn down the Volume (rate of transcription)
- Activators, bind to the control elements of the enhancer = turn up the Volume (rate of transcription)

- # Enhancers & Specific Transcription Factors

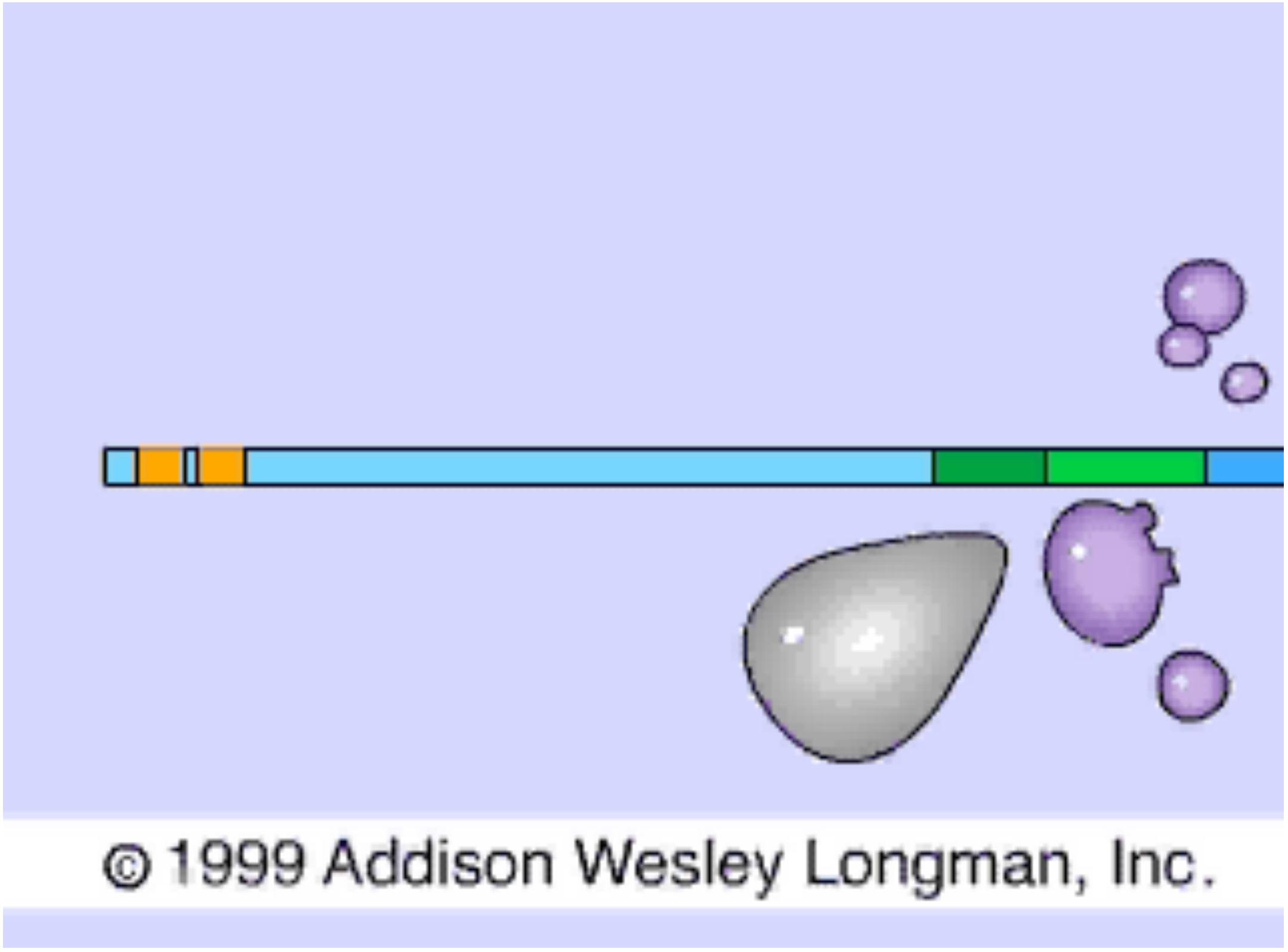
Activator proteins bind to distal control elements grouped as an enhancer in the DNA. This enhancer has three binding sites.

A DNA-bending protein brings the bound activators closer to the promoter. Other transcription factors, mediator proteins, and RNA polymerase are nearby.

The activators bind to certain general transcription factors and mediator proteins, helping them form an active transcription initiation complex on the promoter.



- Enhancers & Specific Transcription Factors

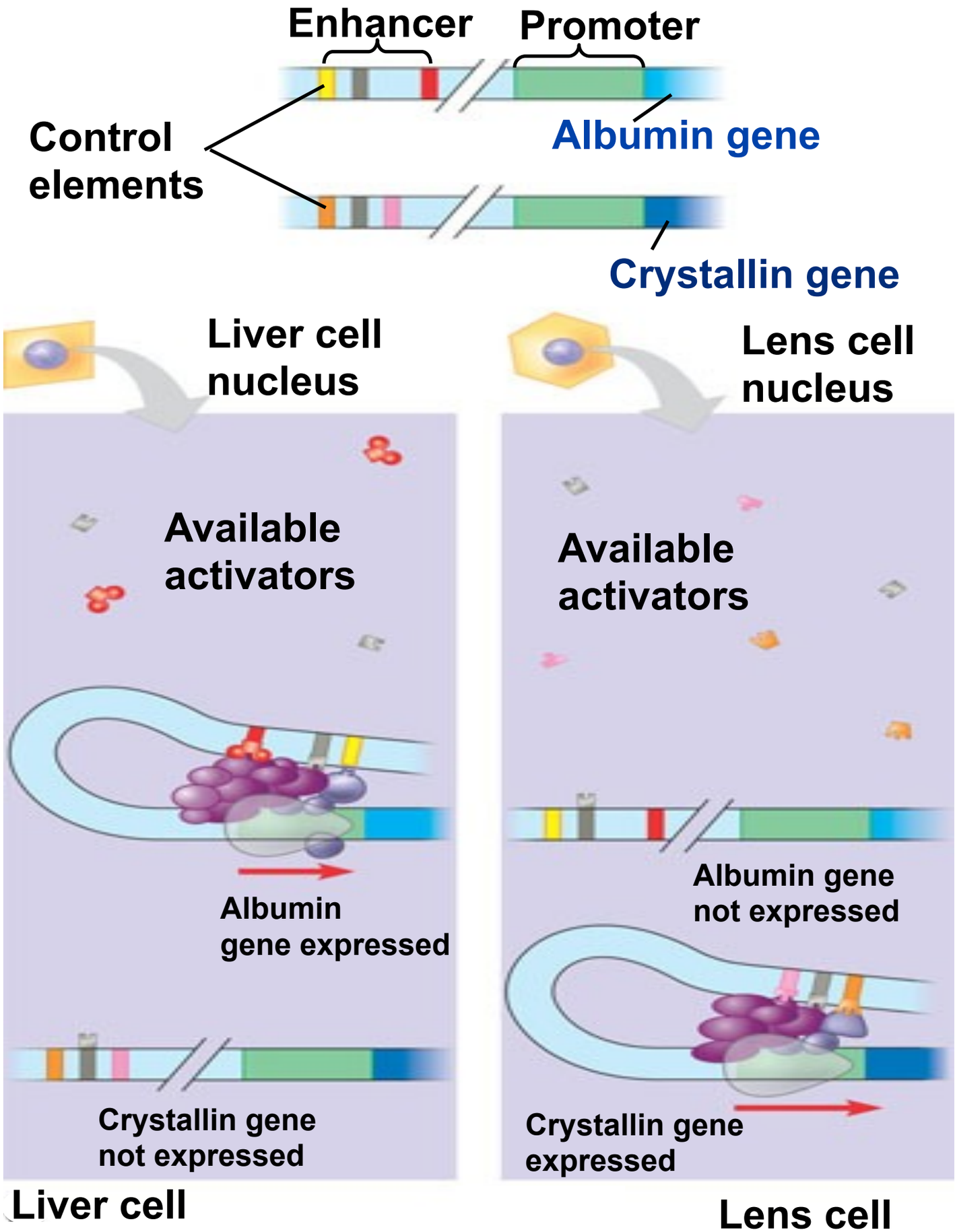


# Regulation of Transcription

- <u>Combinatorial Control of Gene Activation</u>
- In eukaryotes precise control of transcription depends largely on the binding of activators to DNA control elements.
  - With so many genes you might expect a lot of different activators, ironically the number is small.
  - These dozen or so sequences show up again and again in the control elements for different genes.
  - The combination of control elements is the important factor when regulating gene transcription.

- Combinatorial Control of Gene Activation

- Even with only a dozen control elements, a very large number of combinations are possible.



# Regulation of Transcription

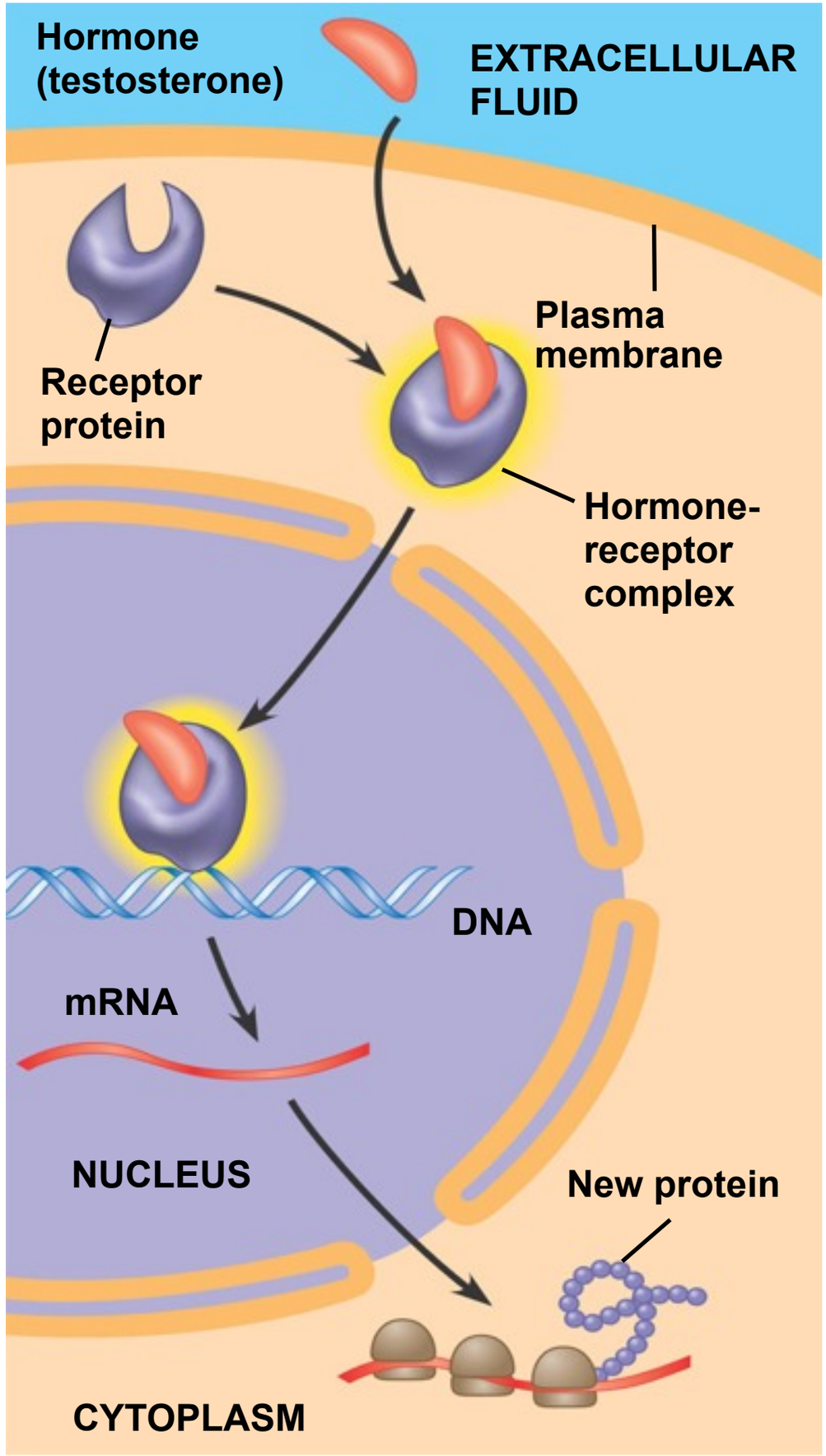
- <u>Coordinately Controlled Genes in Eukaryotes</u>
- Recall that prokaryotes clustered all the genes in a metabolic pathway together in one operon so that they might all be expressed at the same time.
- Eukaryotes need to express multiple genes at the same time as well BUT their genes are not placed together, in fact they may not even be found on the same chromosome.

SO, How do they do it? How do they transcribe all the genes in a metabolic pathway at the same time?

# Regulation of Transcription

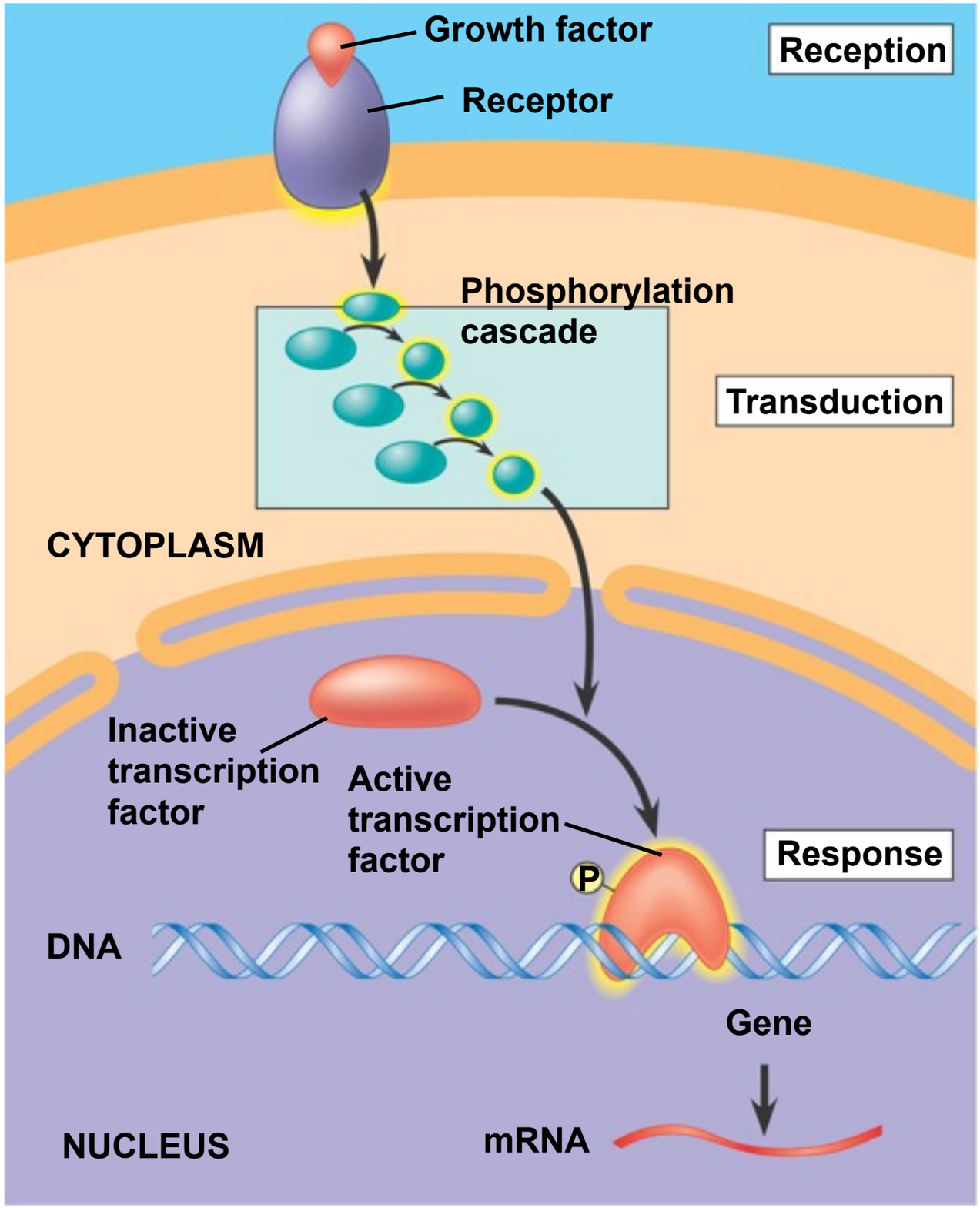
- Coordinately Controlled Genes in Eukaryotes
- Co-expressing genes coding for enzymes in a metabolic pathway depend on the association of a specific combination of control elements with every gene in the group.
- Copies of the activators that recognize the control elements bind to them, promoting the simultaneous transcription of genes regardless of their position in the genome.

- <u>Coordinately Controlled Genes in Eukaryotes</u>
- Steroid hormone enters cell, binds to intracellular receptor which serves as the transcription activator.
- Every gene in the same metabolic pathway is turned on by this same hormone receptor complex.
  - Estrogen activates all the genes that stimulate cell division in the uterus, in preparation for pregnancy.



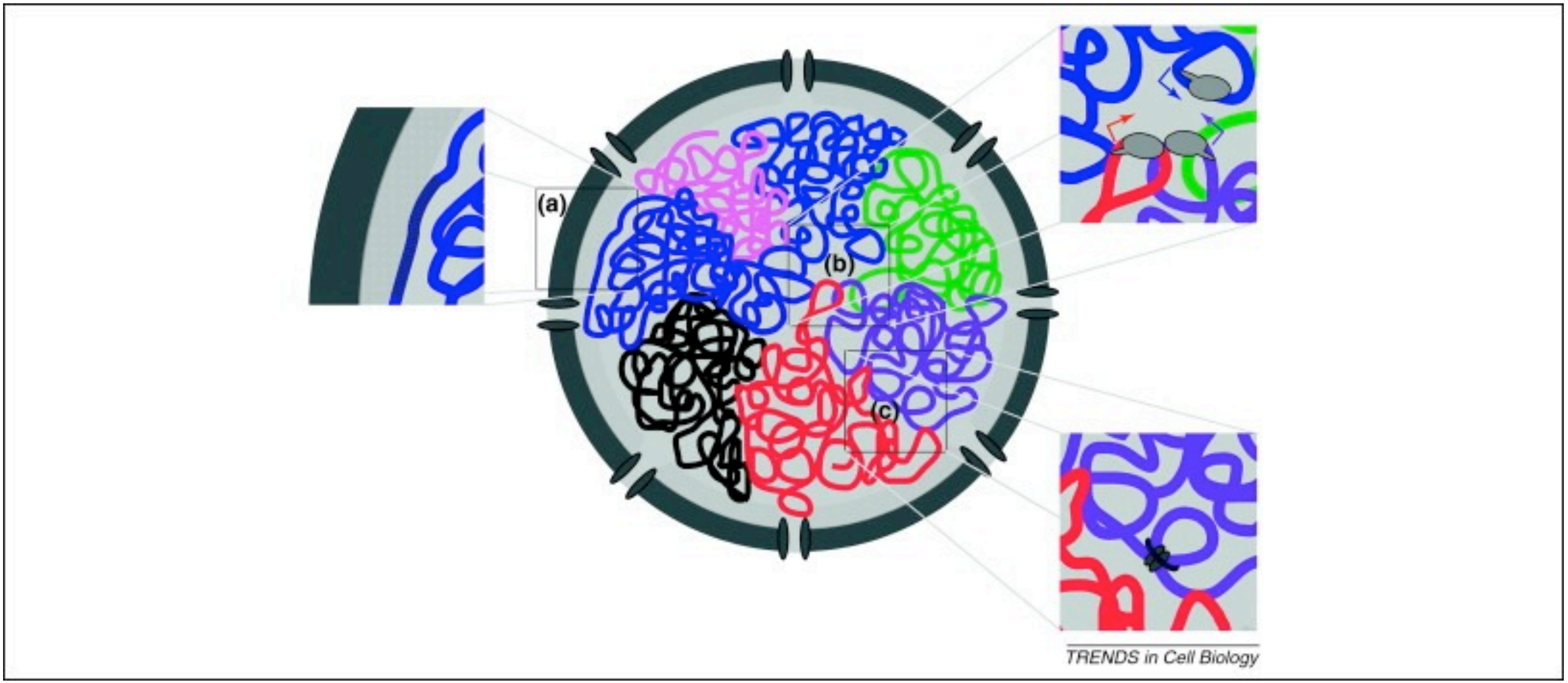
- <u>Coordinately Controlled Genes in Eukaryotes</u>

- Water soluble hormones need not enter the cell, instead they trigger a signal transduction pathway that leads to the activation of specific transcription activators or repressors.



- Nuclear Architecture and Gene Expression

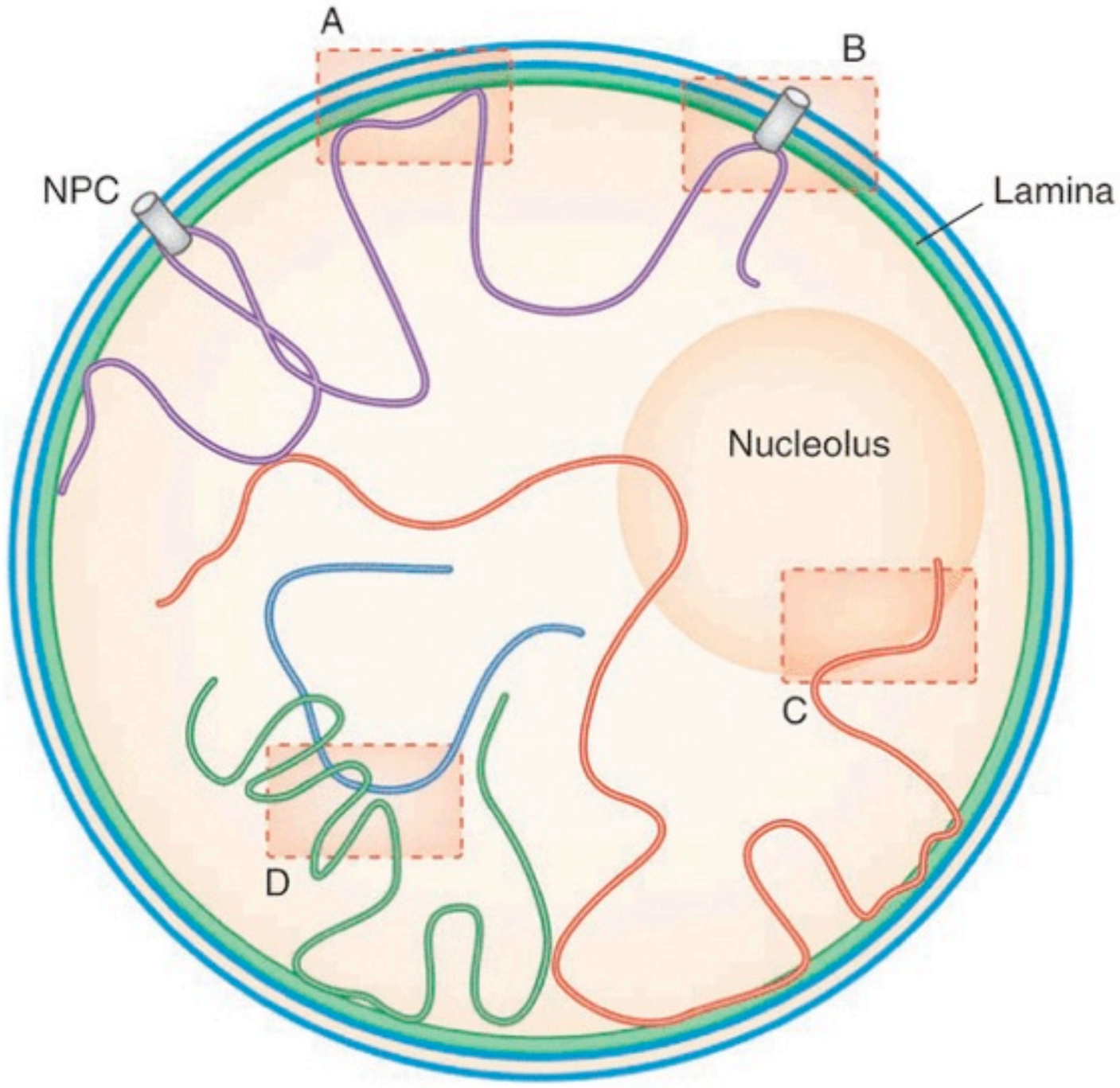
- The organization of chromosomes is not random, each chromosome has it own unique position, held in place by attaching to the nuclear envelope and overlapping certain other chromosomes.



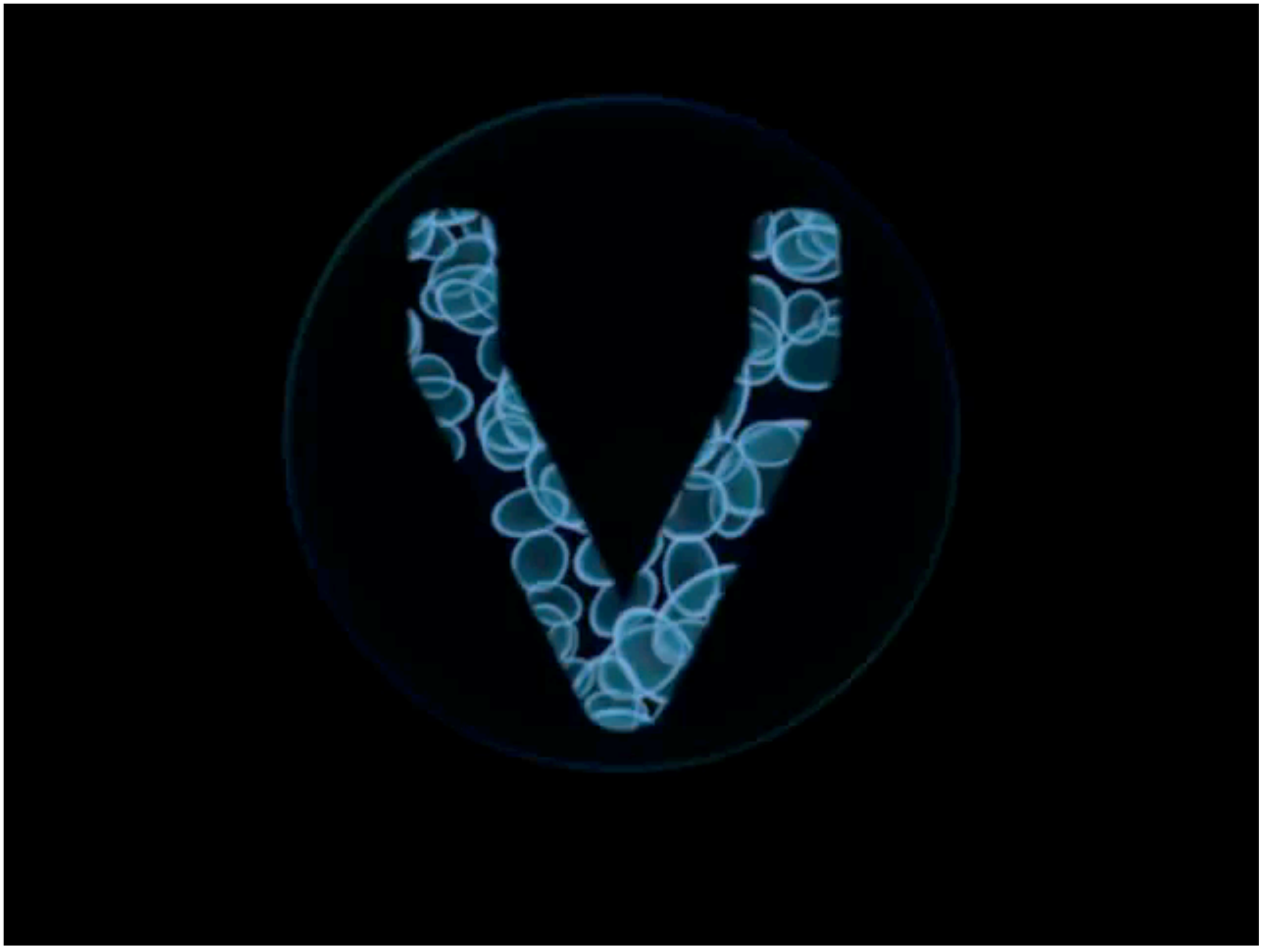
- Nuclear Architecture and Gene Expression

- The areas where chromosomes overlap (D) are rich in RNA polymerases and transcription factors.

- These overlapping regions are called transcription factories.



- Regulated Transcription



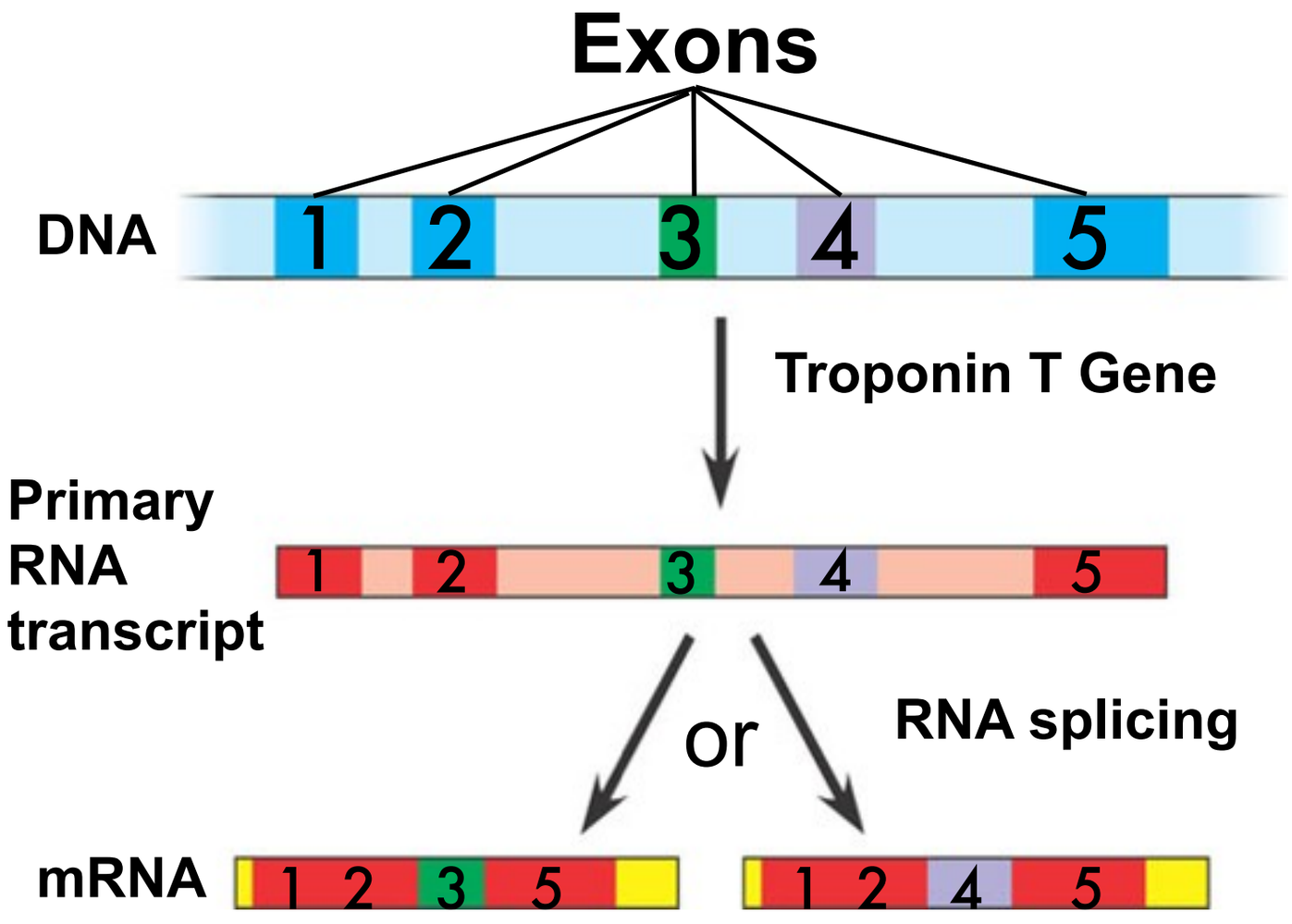
# Post Transcriptional Regulation

- Gene expression is not controlled at the level of transcription alone.
- Ultimately gene expression is measured by the amount of functional gene products.
- There are many opportunities after transcription for a cell to fine tune gene expression

# Post Transcriptional Regulation

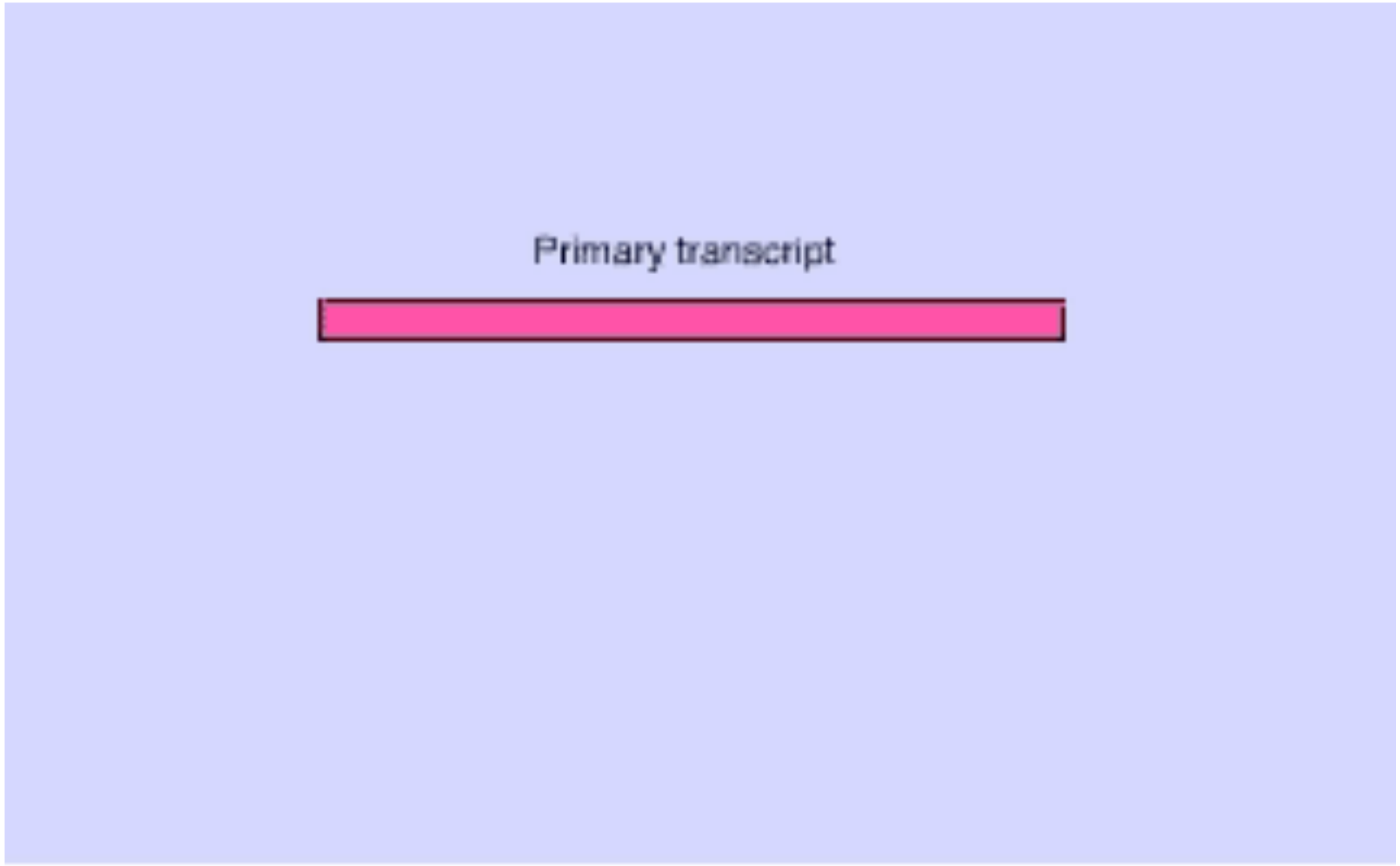
- <u>RNA Processing</u>
- Alternative RNA Splicing, in which different mRNA molecules are produced from the same primary transcript depending on which segments are treated as exons and which are introns.
- Regulatory proteins specific to cell type control these intron/exon choices

- RNA Processing
- This again might explain why humans have the same number genes as the mustard plant sea anemone and the nematode, splicing can generate thousands of different gene products from few genes.




Researchers have discovered a Drosophila gene with enough exons to create 19,000 different membrane proteins, at least 17,400 are actually made

- RNA Processing



# Post Transcriptional Regulation

- <u>mRNA Degradation</u>
- The lifespan of the mRNA product also determines the pattern of protein synthesis.
  - A short lived mRNA makes less functional proteins, while a long lived mRNA makes much more.
  - Typically bacterial mRNA's last for minutes while eukaryotic mRNA's last for hours, days or even weeks.
  - Degradative enzymes determine the lifespan of these mRNAs
  - Nucleotide sequences in the UTR region of the mRNA itself affect its lifespan

- mRNA Degradation


Longer-lived mRNA

# Post Transcriptional Regulation

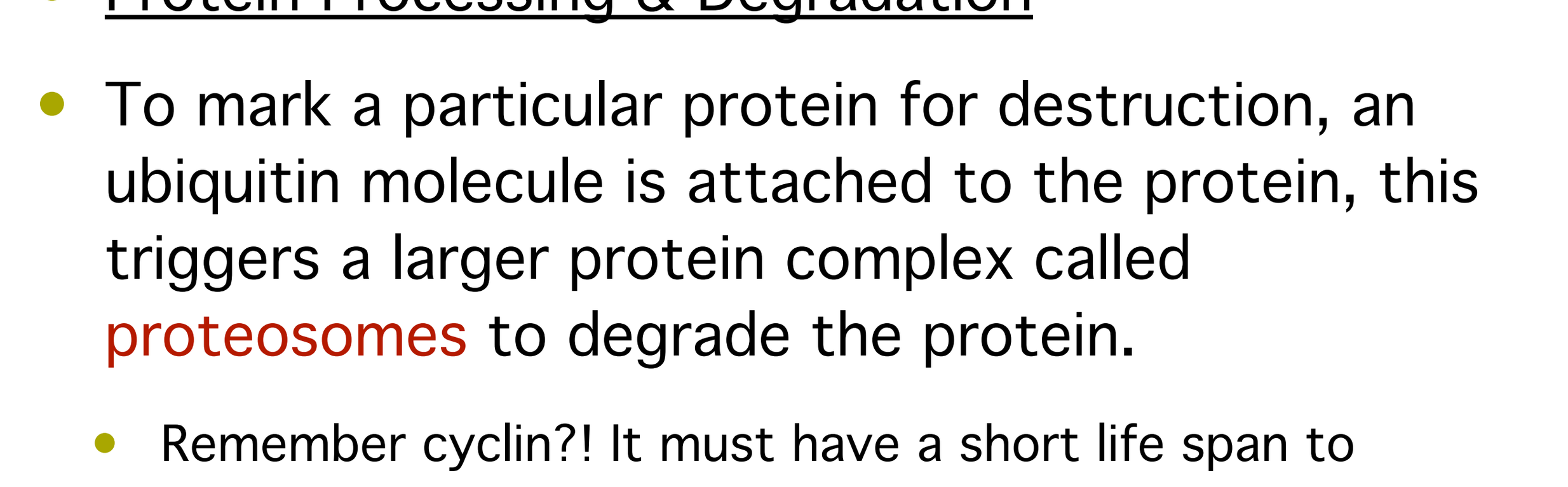
- Initiation of Translation
- Here is yet another opportunity for gene regulation, some regulatory proteins operate by blocking the attachment of ribosomes.
- Usually this blocking occurs at either the 5' cap or the poly-A tail.
  - In fact some mRNAs are made intentionally with short poly-A tails, and only when the protein is needed are the tails lengthen and translation begins.

# Post Transcriptional Regulation

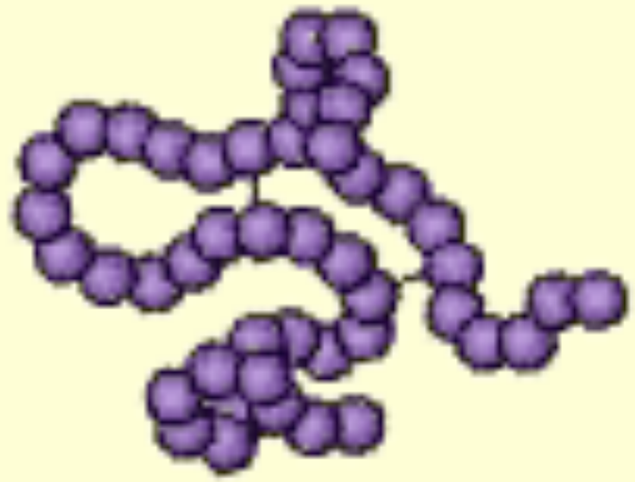
- <u>Protein Processing & Degradation</u>
- The final opportunity for gene regulation occurs after translation, often polypeptides must be processed before they become functional.

- Regulation can therefore before at any of the steps that might involve modifying or transporting the protein.
- Finally, the lifespan of the protein product itself might be regulated by means of selective degradation.

- <u>Protein Processing & Degradation</u>
- To mark a particular protein for destruction, an ubiquitin molecule is attached to the protein, this triggers a larger protein complex called proteosomes to degrade the protein.
  - Remember cyclin?! It must have a short life span to function properly, it is degraded in this way.



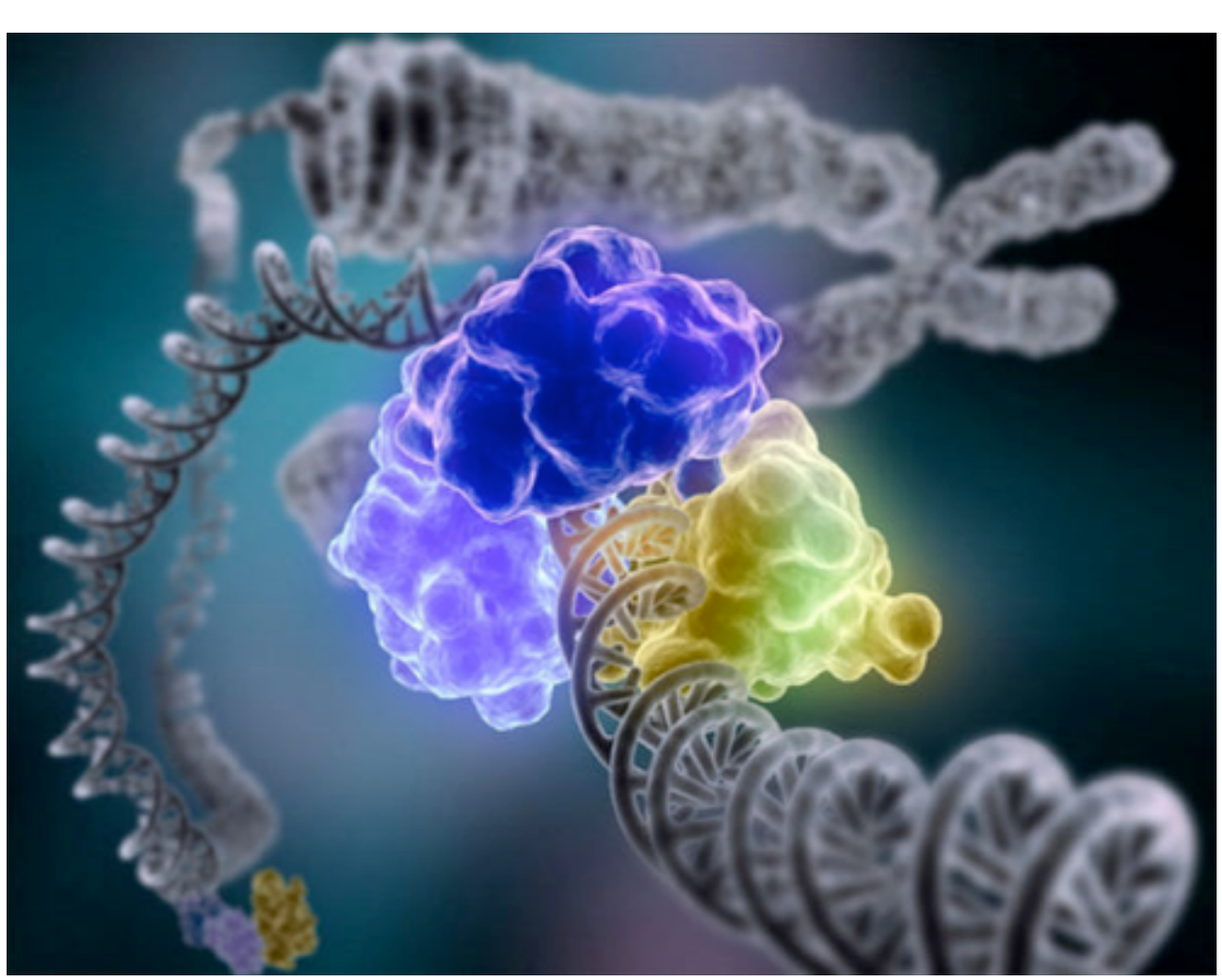
- Protein Processing & Degradation



Polypeptide

# Molecular Basis of Inheritance

III.Main Idea: Much of the DNA does not code for proteins and we thought that this DNA did not play a role or was simply evolutionary baggage but today we are learning that is far from the case.

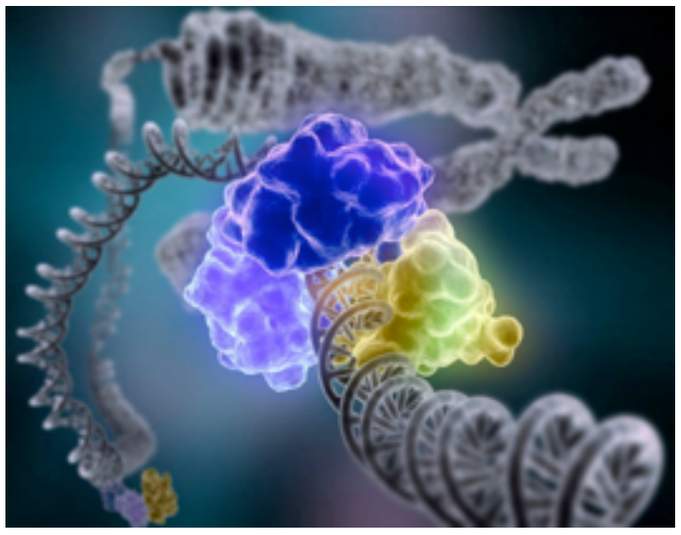


# Roles of Non-Coding RNAs

- Less than 2% of the total genome codes for polypeptides.
- A similarly small percentage codes for RNAs that are actually translated like tRNA or rRNA.
- Until recently, the remaining DNA was assumed to be untranscribed.
- Recent evidence contradicts this idea and reveals that most of the DNA is transcribed into ncRNA that is never translated into anything.

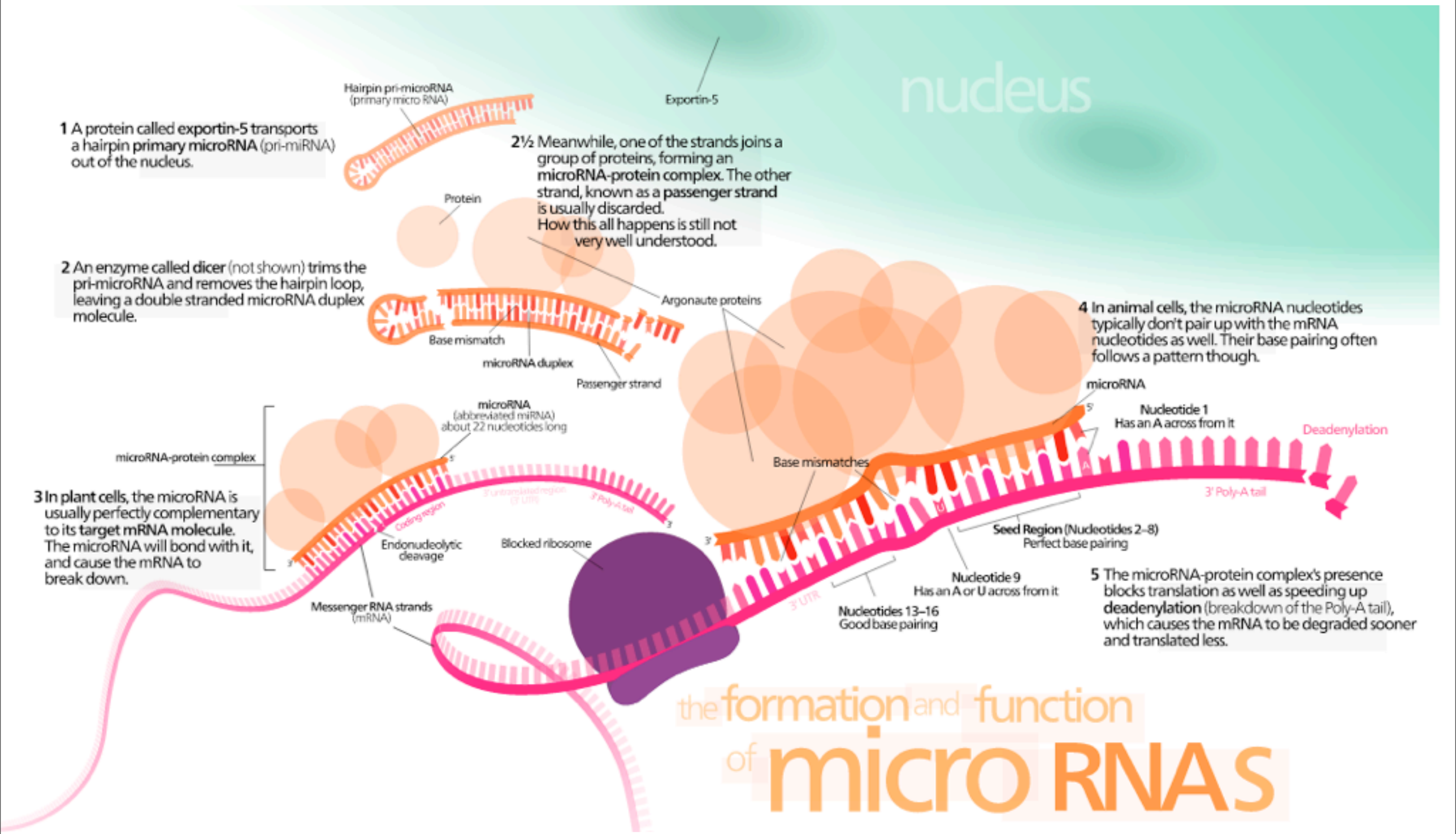
# Roles of Non-Coding RNAs

- More research is revealing that these ncRNA molecules play crucial roles in gene regulation.
- It is as if we have been focusing on the lead singer of a rock band, while ignoring the musicians and back up singers that have been there all along.
- Two of these ncRNAs have been studied extensively in the last couple decades, siRNA and miRNA.



This picture you have seen on the main idea slides shows ncRNA interacting with DNA!

# Non-Coding miRNAs



# Chromatin Remodeling & ncRNAs

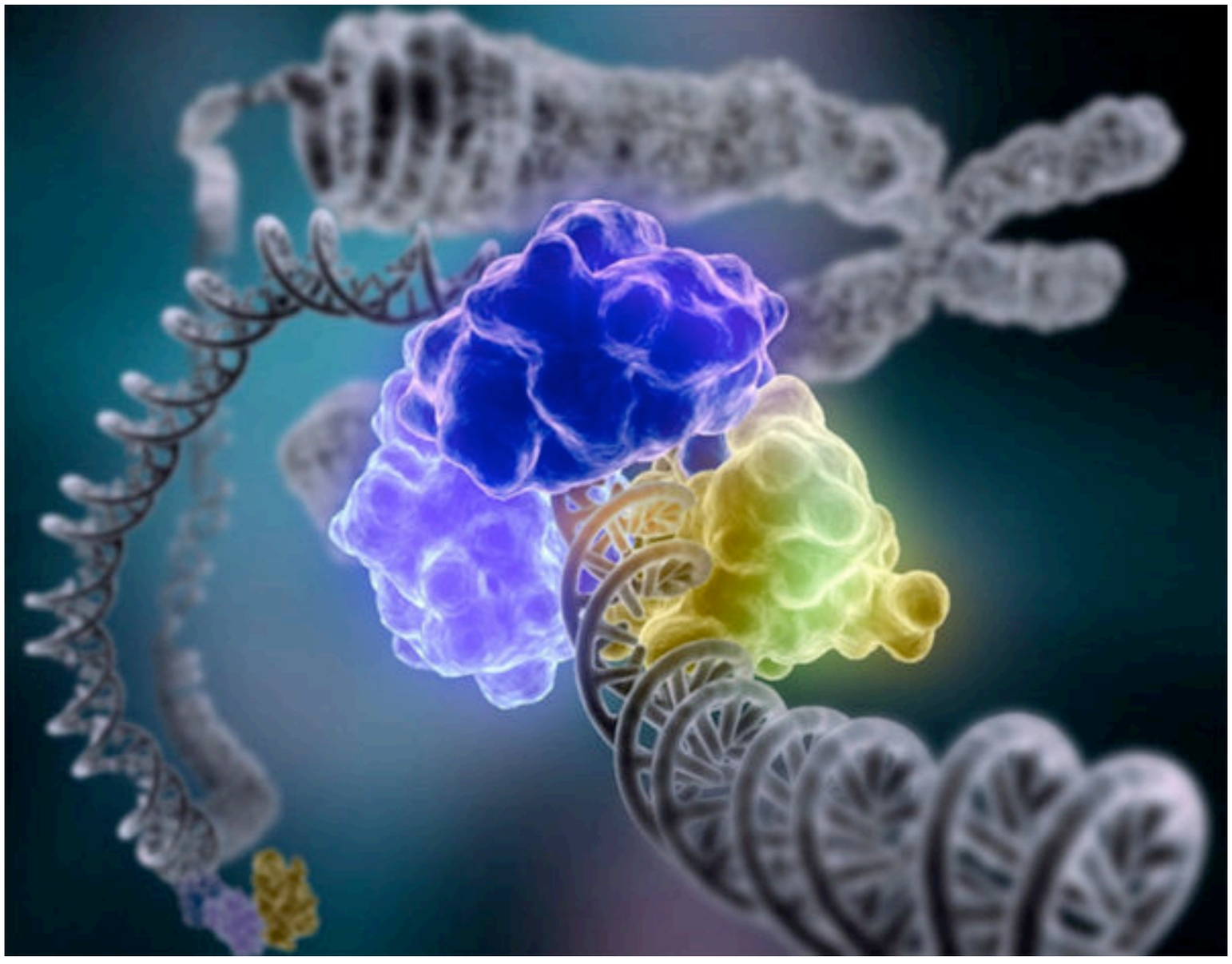
- In addition to affecting mRNA, small RNAs can also affect chromatin structure.
  - These small RNAs recruit proteins that modify the chromatin into the highly condensed form (heterochromatin)
  - Newly discovered piwi associated RNAs (piRNAs) have been found to induce heterochromatin formation and to block parasitic DNA elements called transposons discussed later.
  - They play an indispensable role in germ cells where they re-establish methylation patterns in the genome during gamete formation.

# Evolutionary Significance (ncRNA)

- Small ncRNA can regulate gene expression at multiple steps and in multiple ways.
- Scientists hypothesize that extra levels of regulation might allow a higher degree of complexity of form.
- Genome sequencing supports the idea that siRNA evolved first followed by the miRNA and then the piRNA (found only in animals)

# Molecular Basis of Inheritance

## IV.Main Idea: Gene regulation and expression plays a critical role in the orchestration of development in animals.



# Embryonic Development Genes

- The transformation of a zygote into a multicellular organism is amazing and relies on three processes: cell division, cell differentiation & morphogenesis.
  - Cell division generates the vast number of cells required to build say a human.
  - Cell differentiation is the process where cells become specialized in structure and function.
  - Morphogenesis organizes and arranges these specialized cells into a particular 3-dimensional arrangement that gives it creative form.

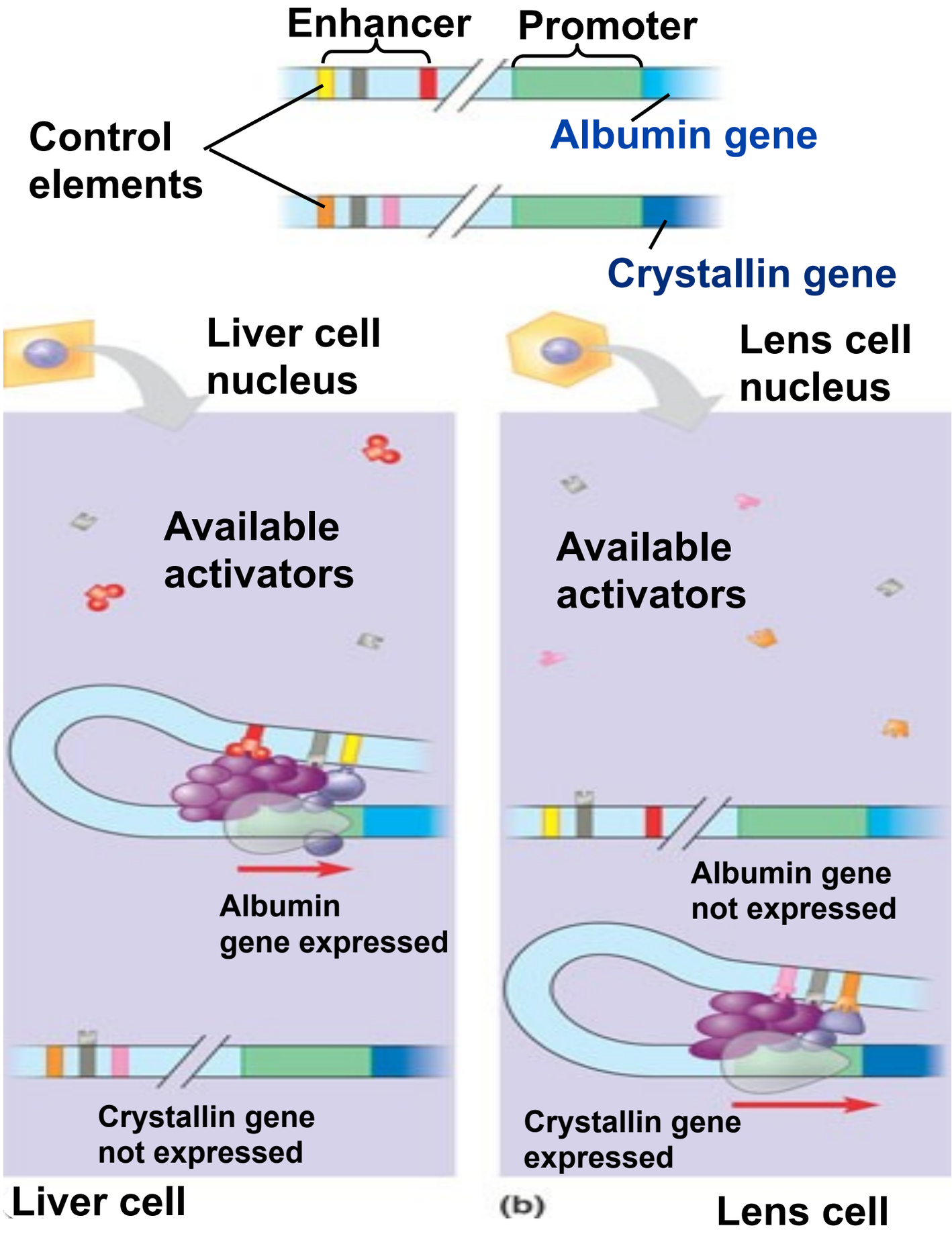
# Embryonic Development Genes

- All three processes: cell division, cell differentiation & morphogenesis have their basis in cellular behavior.
- Cellular form and behavior depend on the genes it expresses and the proteins it produces.
- Remember almost all cells in an organism have the same genome, thus differential gene expression results when cells regulate the genes differently from other cell types.

- Recall

Both cells arose from the same zygote through mitosis. So…

How do different sets of activators come to be present in the two cells?

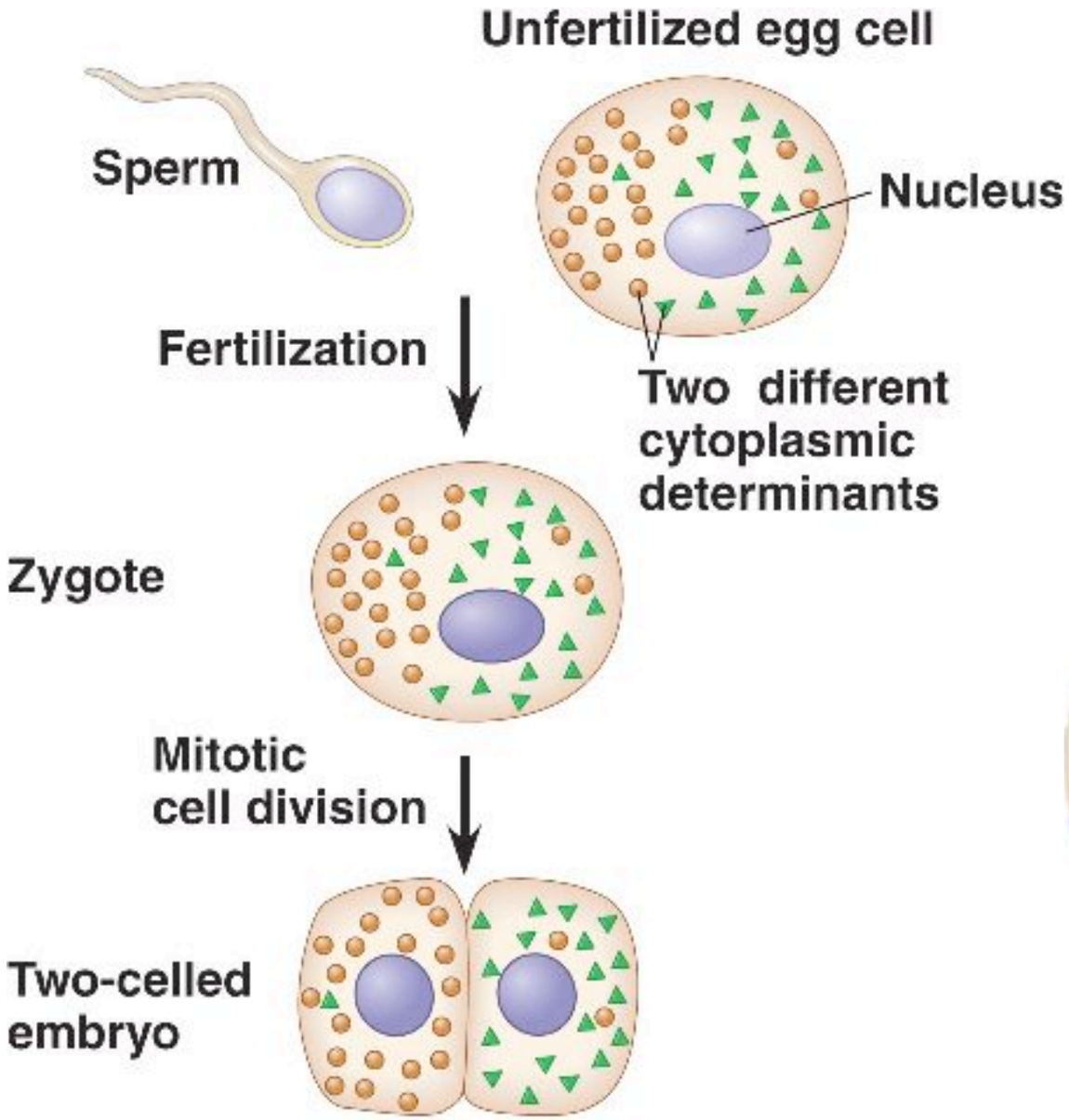


# Cytoplasmic Determinants & Inductive Signals

- Two sources of information used by early cells of the embryo tell a the cell which genes to express at any given time in development:

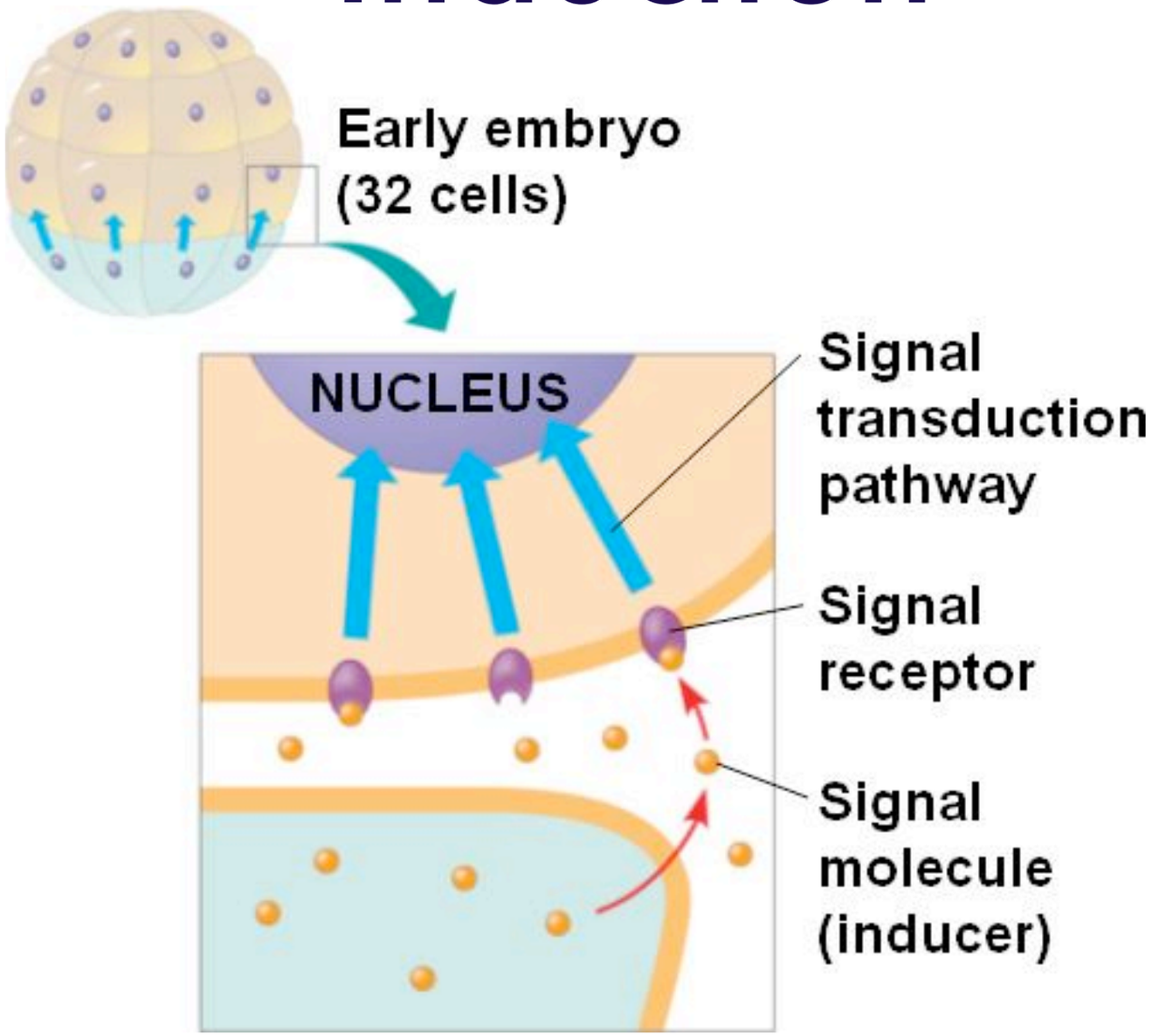
- 1. Cytoplasmic determinants- maternal substances in the egg, RNAs proteins and other substances play a profound role.

- 2. The environment around the developing cells, particularly inductive signals that come from other cells in the vicinity.



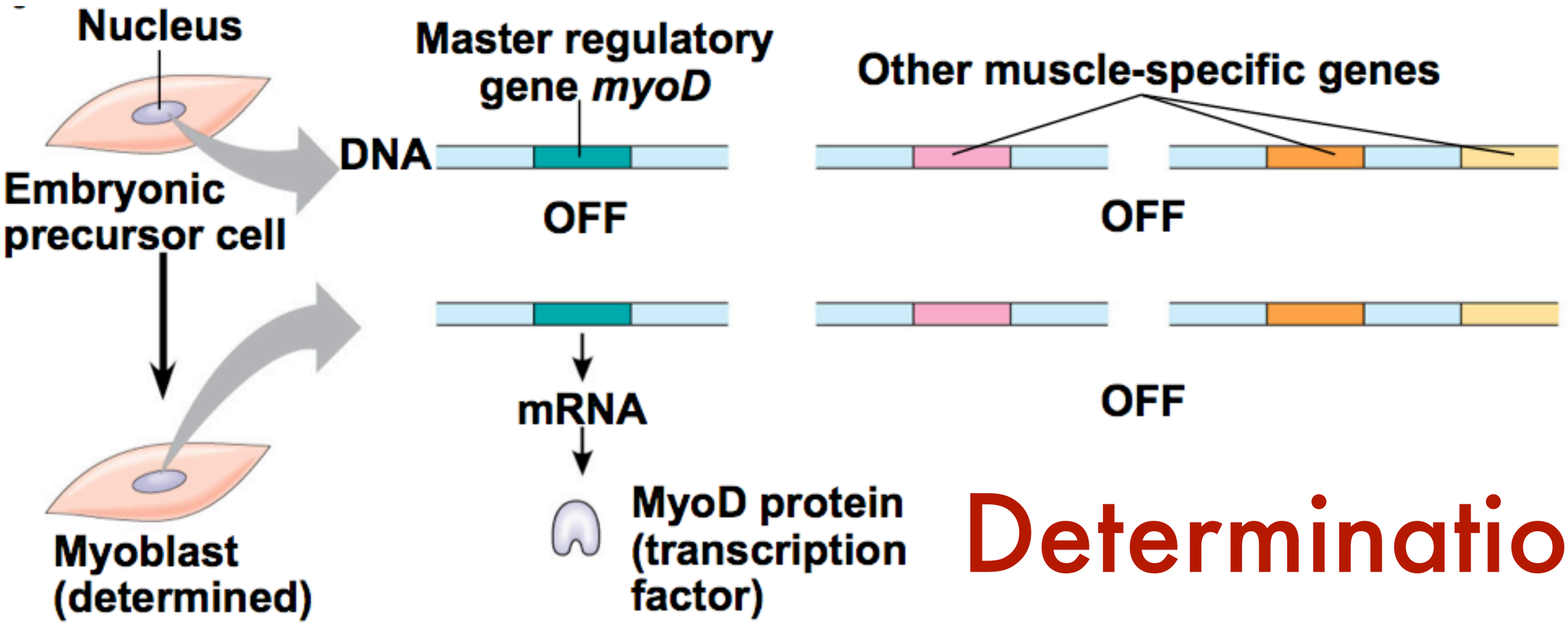

# Cytoplasmic Determinants

# Induction



# Sequential Regulation of Gene Expression During Differentiation

- Determination commits a cell to its final
- Determination precedes differentiation
- Cell differentiation is marked by the production of tissue-specific proteins

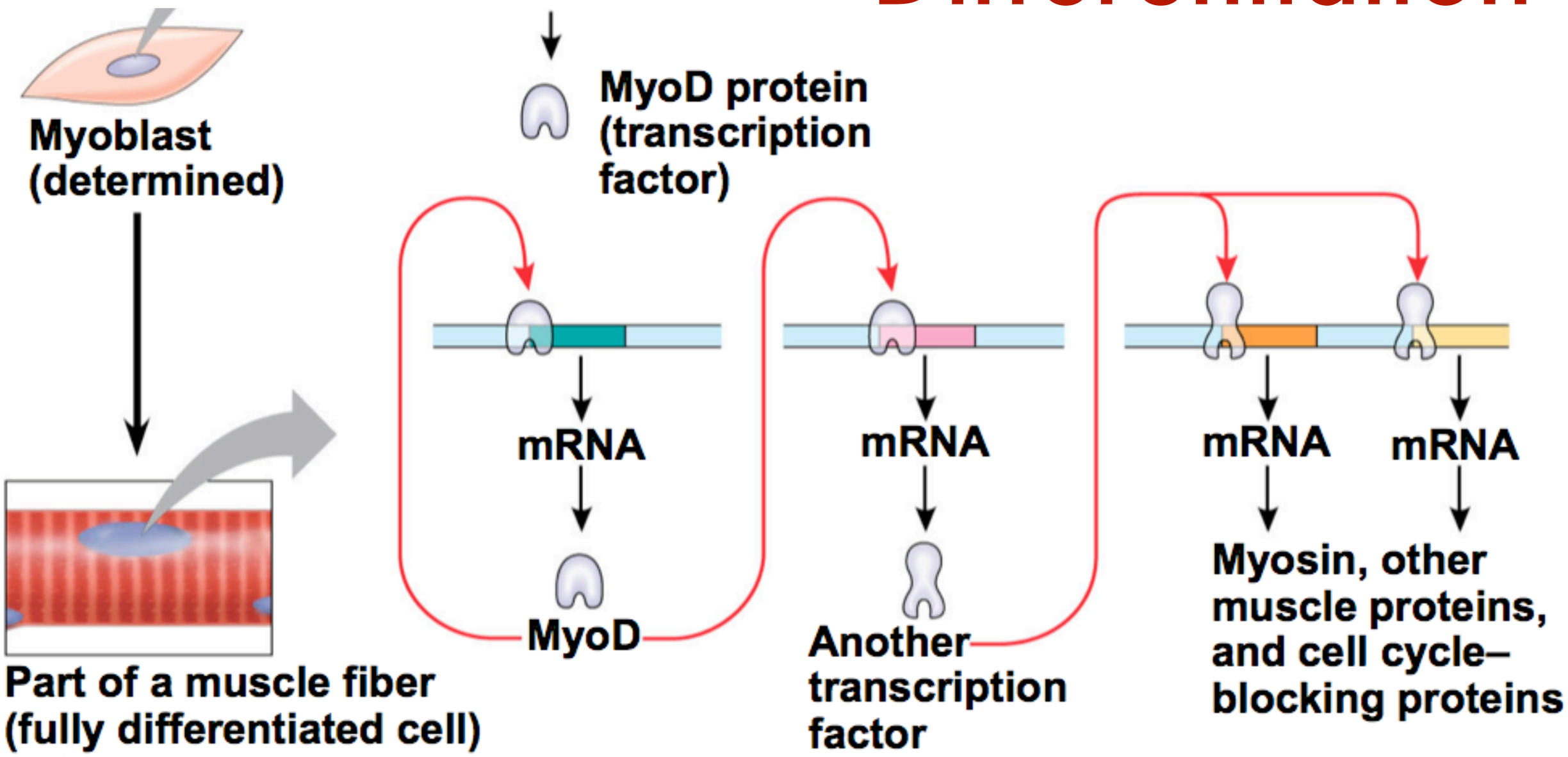



# Determination

- Myoblasts produce muscle-specific proteins and form skeletal muscle cells
- *MyoD* is one of several “master regulatory genes” that produce proteins that commit the cell to becoming skeletal muscle

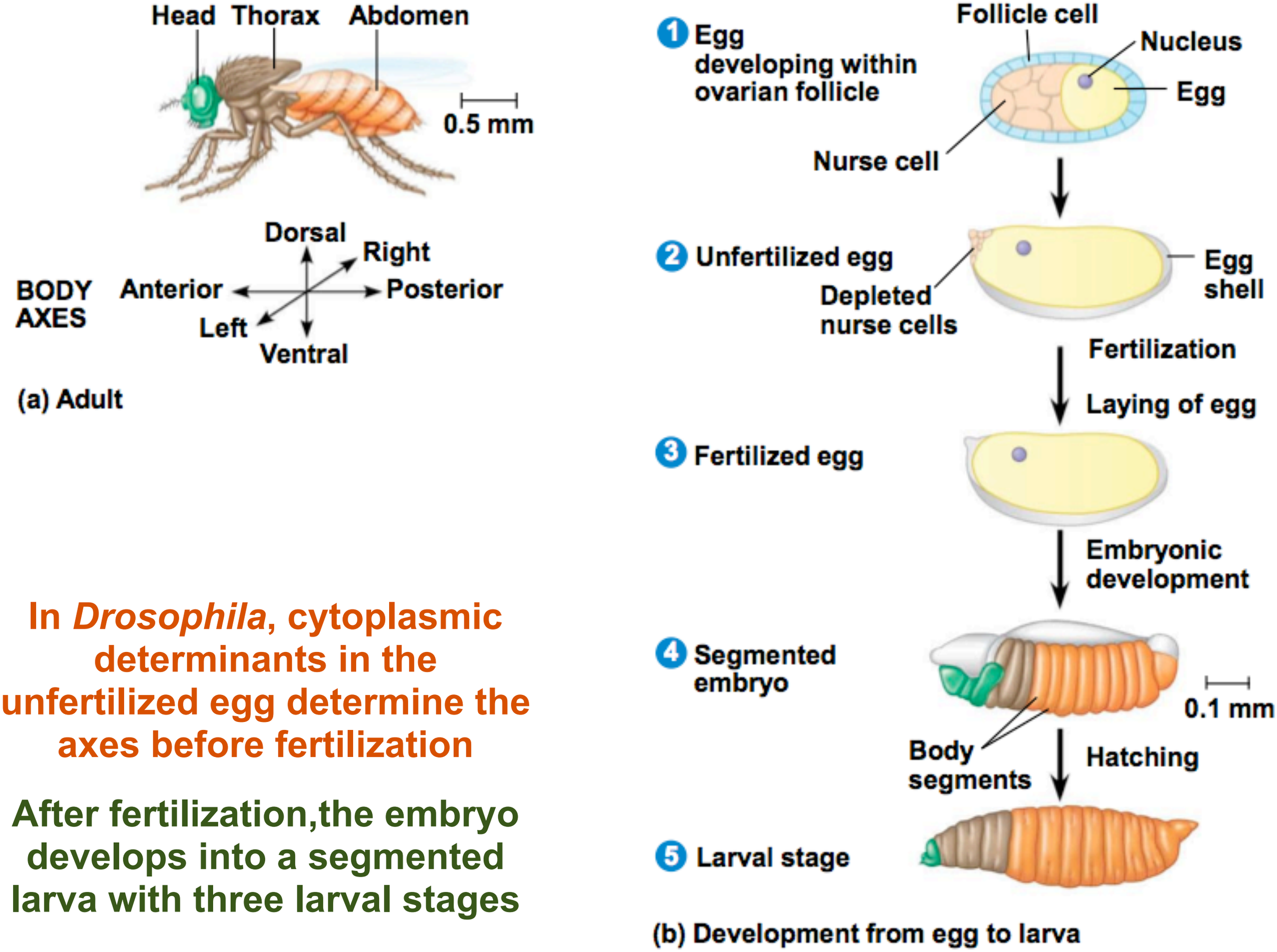
- The MyoD protein is a transcription factor that binds to enhancers of various target genes

# Differentiation



# Pattern Formation- The Body Plan

- Pattern formation is the development of a spatial organization of tissues and organs
- In animals,pattern formation begins with the establishment of the major axes
- Positional information, the molecular cues that control pattern formation, tells a cell its location relative to the body axes and to neighboring cells
- Pattern formation has been extensively studied in the fruit fly *Drosophila melanogaster*

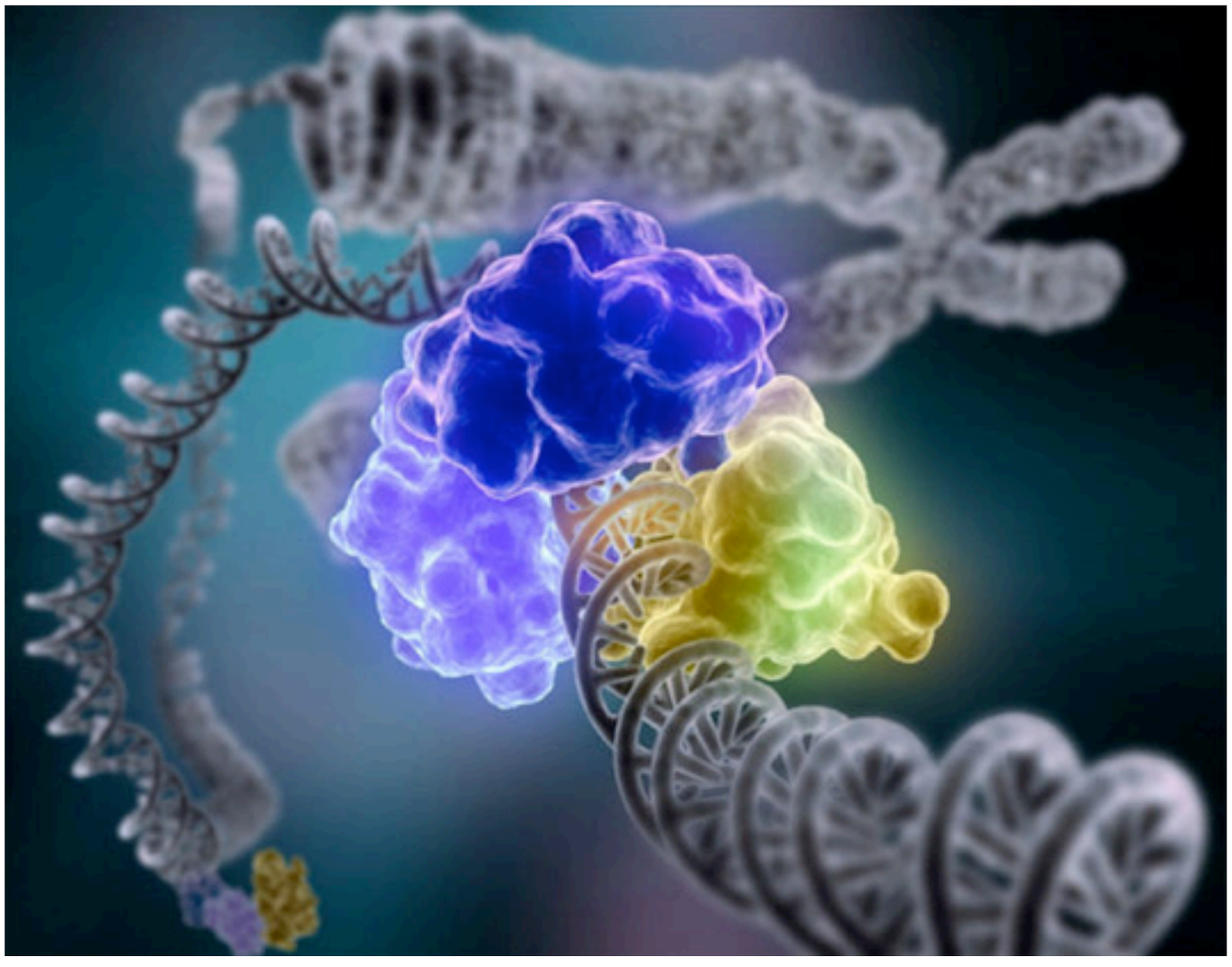



**In *Drosophila*, cytoplasmic determinants in the unfertilized egg determine the axes before fertilization**

**After fertilization,the embryo develops into a segmented larva with three larval stages**

# Molecular Basis of Inheritance

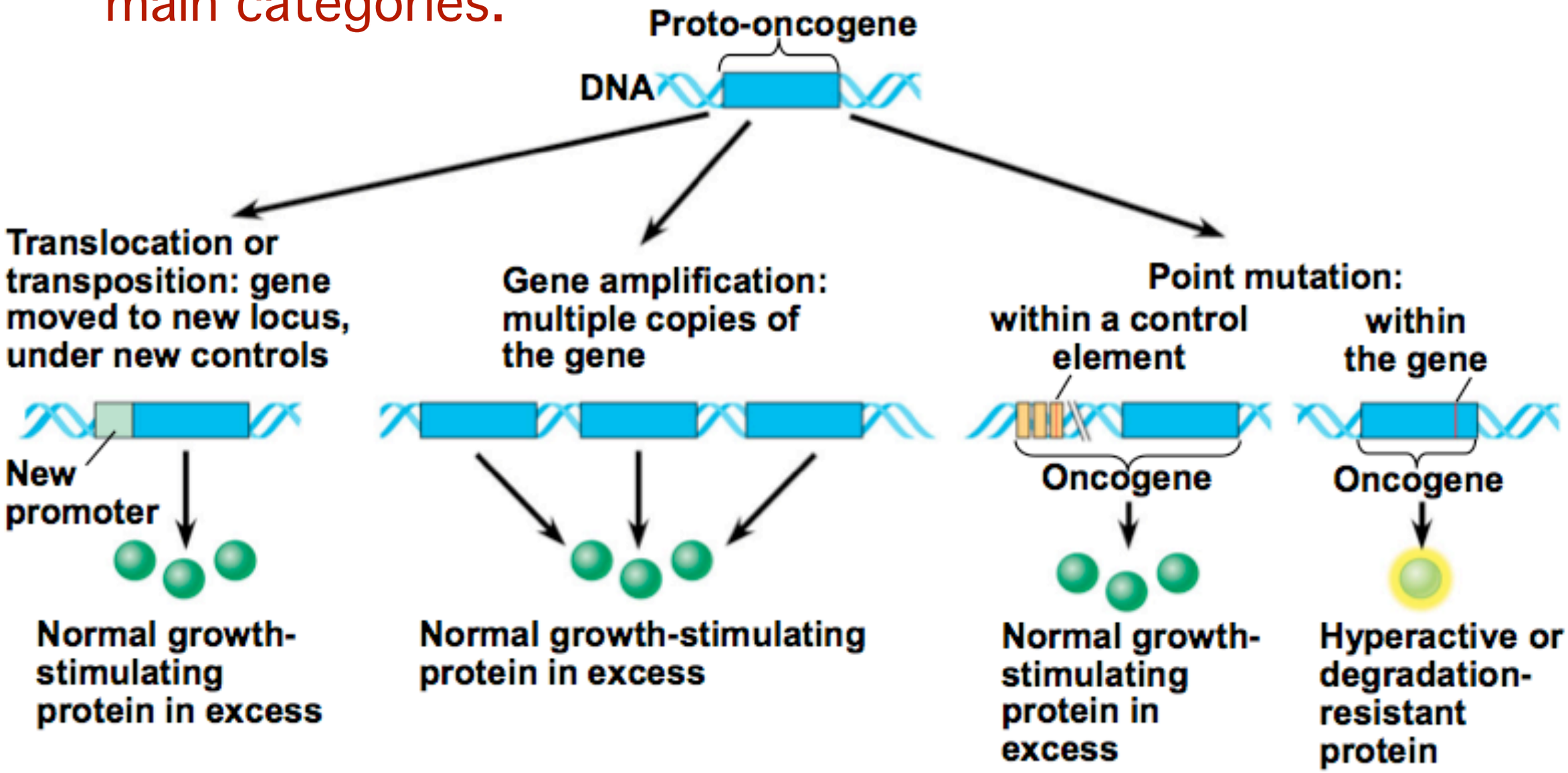
## V.Main Idea: Cancer has a molecular basis. Many cancers result from a loss of gene regulation.



# Genes Associated with Cancer

- Cancer may result from mutations in the genes that regulate cell growth and cell division.
  - These genes may code for growth factors, their receptors, intracellular molecules and intracellular receptors involved in cell division pathways.
- Normal genes that are involved in cell growth and division are called collectively proto-oncogenes.
- When these genes mutate, they become oncogenes or causing genes.

- In general, oncogenes lead to an increase in the gene product (too much of the protein) or the intrinsic activity of the protein (can not be turned off).
  - The genetic pathways that cause oncogenes fall into three main categories.



# Genes Associated with Cancer

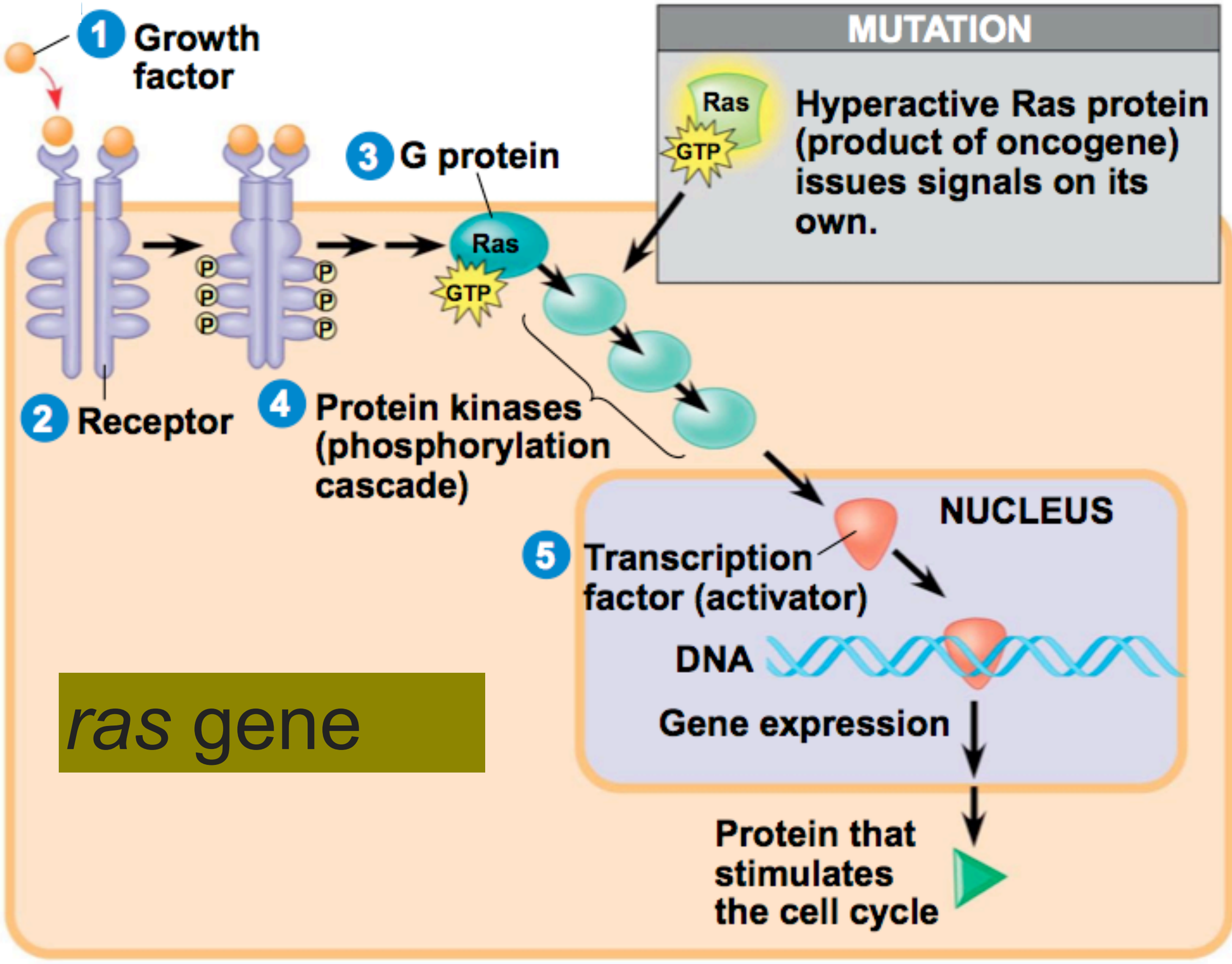
- Oncogenes promote cell growth and division.
  - Imagine a car were the gas pedal is stuck to the floor.
  - Thankfully the car's breaks might save us in this scenario
- Another class of cancer causing genes called tumor-suppressor genes are involved in cell growth inhibition pathways.
- When these genes mutate, the cell loses its ability to stop cell growth and division.
  - Now imagine a car were the gas pedal is stuck to the floor AND the brakes do not work!

# Genes Associated with Cancer

- The products of tumor suppressor genes have various functions.
  - Some produce proteins that repair damaged DNA.
    - without them mutations accumulate
  - Some control the adhesion of cells to each other
    - normal cell division requires cells to adhere to some substrate
  - While others are involved with the direct inhibition of the cell cycle and growth

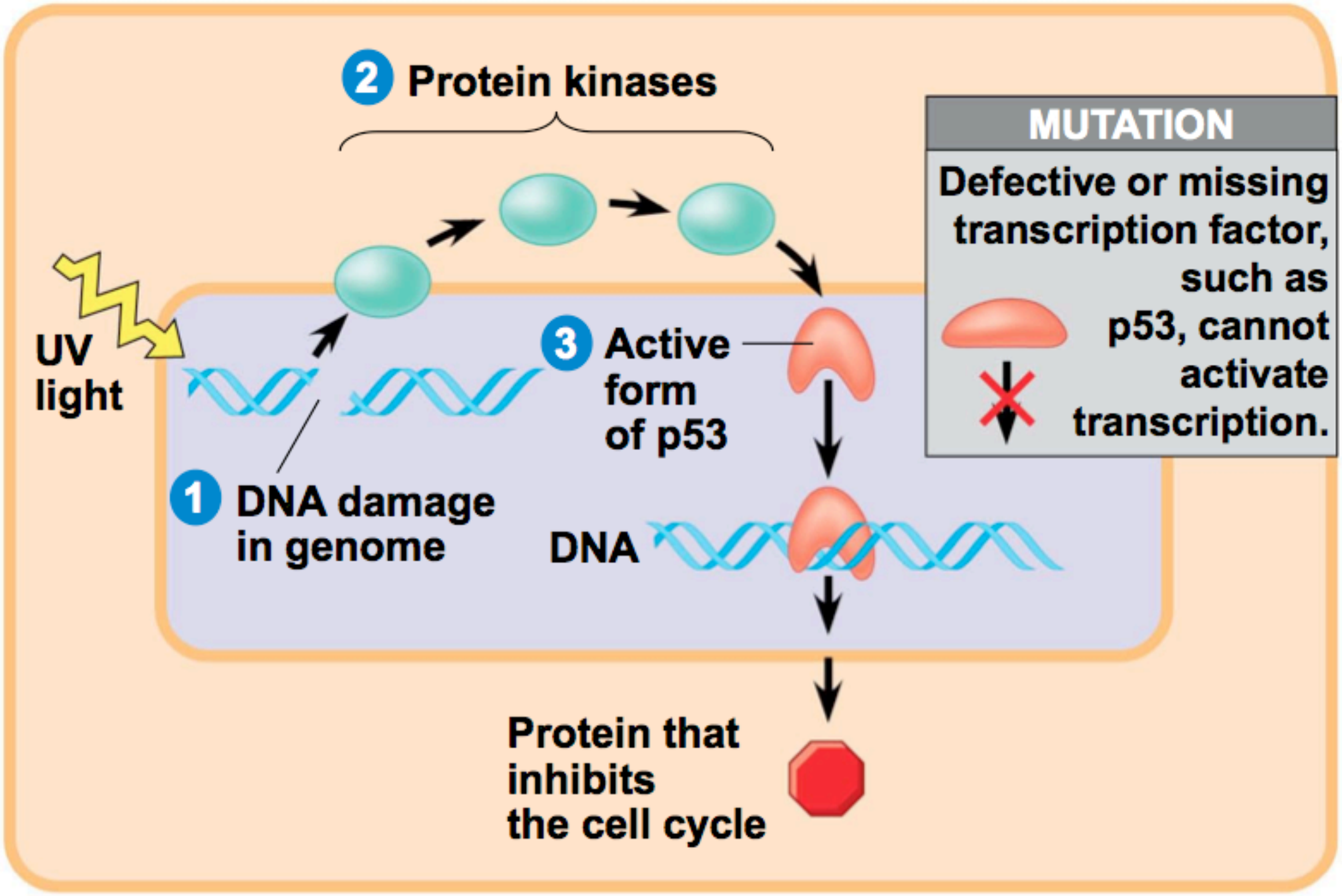
# Interference of Cell Signaling

- The proteins encoded by most oncogenes and tumor-suppressor genes are components of cell signaling pathways.
- To understand how these genes function in cancer we will look at two common causing genes:
  - The ras proto-oncogene, mutations in this gene occur in ~30% of all cancers.
  - The p53 tumor-suppressor gene, mutations in this gene occur in more than 50% of all cancers.




(a) Cell cycle–stimulating pathway

# *p53* gene



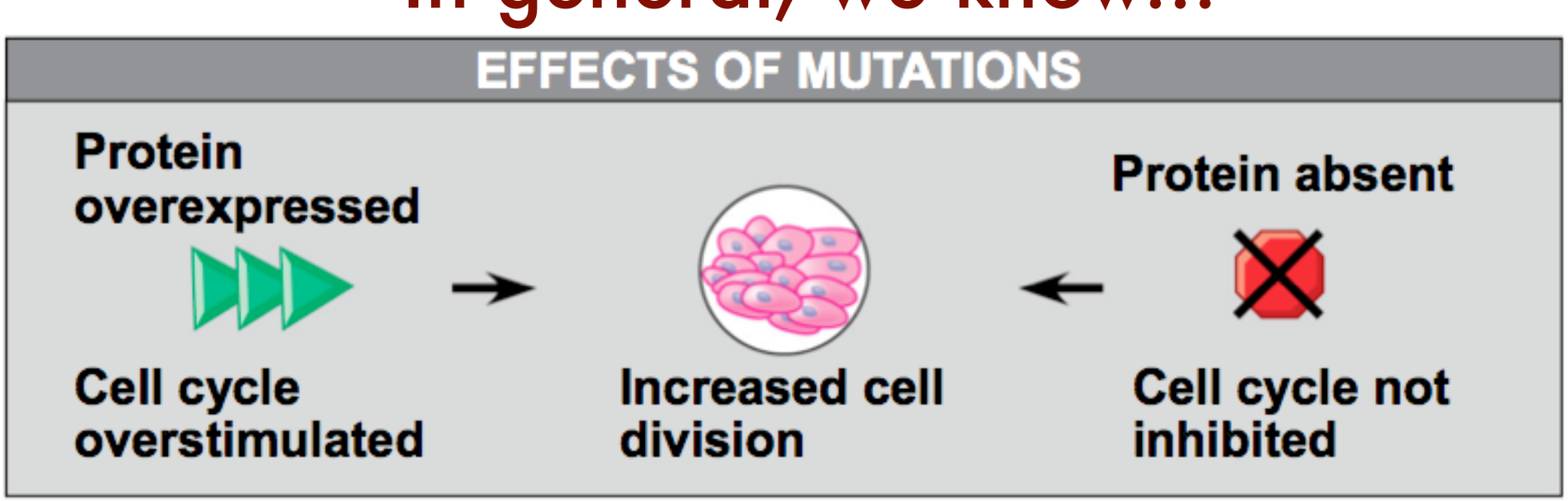

(b) Cell cycle–inhibiting pathway

# Interference of Cell Signaling

- The p53 tumor-suppressor gene, has been called the “guardian angel” of the genome, it functions as an activator for many genes.
  - p53, activates a gene that stops the cell cycle
  - p53, activates miRNAs that stop the cell cycle
  - p53, activates genes directly involved in DNA repairing damaged DNA
  - p53, activates “suicide genes” that cause cell apoptosis when damage is irreparable.

- The p53 tumor-suppressor gene, ultimately tries prevent cells from passing on mutations to daughter cells .
  - The entire story of cancer development is complicated and still unknown.

## In general, we know...




## Unfortunately,
## ...we do not fully understand all the details

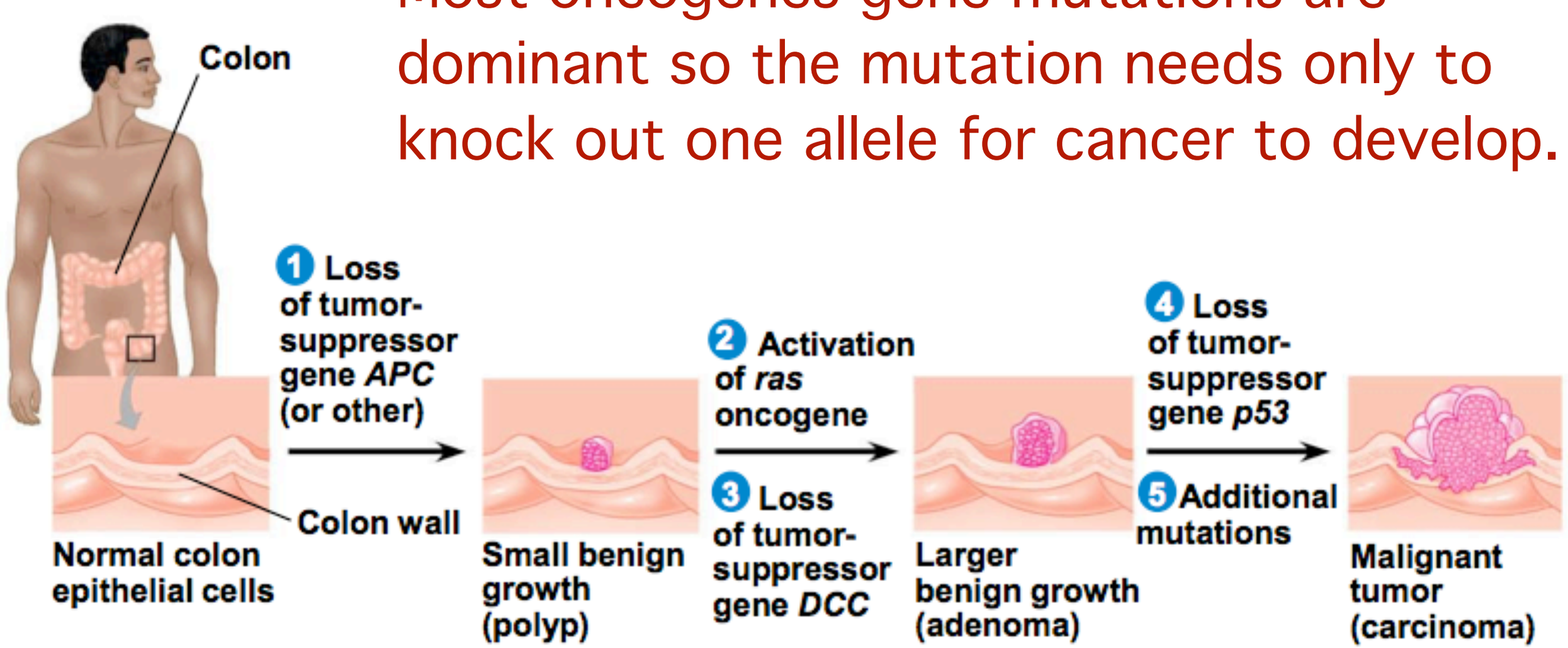
# Multistep Model of Cancer

- Multiple mutations are necessary to produce full-fledged cancer.
- This explains why most cancers occur late in life, after mutations have had time to accumulate.

- The idea that many mutations (~6 or so) are responsible for cancer is fairly well supported.
  - In addition, we know that mutations must occur in both proto-oncogenes and tumor suppressor genes.

# Multistep Model of Cancer

- Most tumor suppressor genes mutations are recessive so the mutations must knock out both alleles for cancer to develop.
- Most oncogenes gene mutations are dominant so the mutation needs only to knock out one allele for cancer to develop.



# Other Factors Contributing to Cancer

- The fact that multiple mutations are necessary for cancer, explains why cancers run in families.
- Individuals who inherit one or more mutations in these critical genes are one step closer to cancer.
- Geneticists are working hard to to identify cancer alleles so that those with predispositions to certain cancers can be detected early in life.

- Colorectal Cancers
- 15% involve inherited mutations.
- Many of these mutations effect a tumor suppressor gene called the APC gene, which plays many important roles in cell growth and division.
- 60% of colorectal cancer patients, even those with no family history, have one mutated APC allele
  - remember to fully knock out the function of this gene the allele must mutate as well.
  - today researchers try find factors that might contribute to the other allele mutating.

- Breast Cancer
- 5-10 % of patients have genetic predisposition.
- For instance, 1990 Claire King demonstrated that mutations in the BRCA1 or the BRCA2 gene were associated with breast cancer.

- One woman who inherits one mutated BRCA1 allele has 60% chance of developing breast cancer before the age of 50, compared to 2% when the woman is homozygous with two normal alleles.
  - both BRCA1 and BRCA2 genes are tumor suppressor genes

# Other Factors Contributing to Cancer

- We know that damaging DNA contributes to cancer.
- As a result we know that avoiding "mutagens" like UV light or certain chemical agents can lower the risk of developing cancer.

# Other Factors Contributing to Cancer

- We know that damaging DNA contributes to cancer.
- As a result we also know that viruses can cause cancer.
  - The Epstein-Barr virus (mononucleosis) is associated with Burkitt's Lymphoma
  - Papillomaviruses are associated with cervical cancers.
  - HTVL-1 virus causes a type of adult leukemia
- Worldwide viruses seem to play a role in about 15% of human cancers.