Lecture Series 8

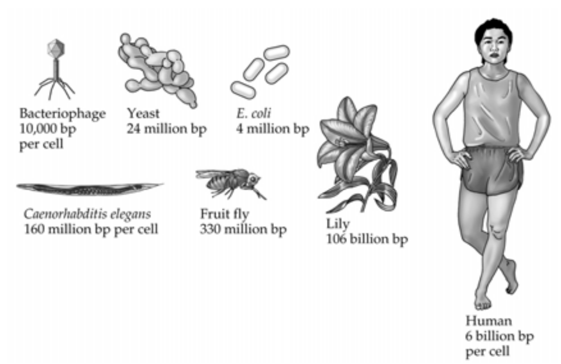
The Eucaryotic Genome and Its Expression

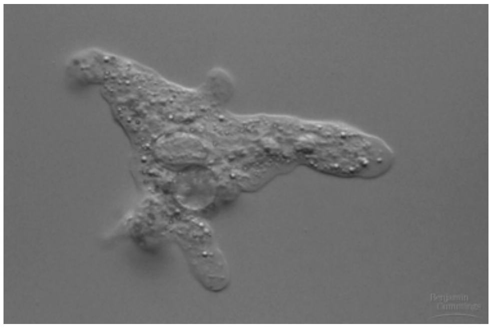
Reading Assignments

- Read Chapter 8
  Control of Gene Expression
- Skim Chapter 9
  How Genes and Genomes Evolve

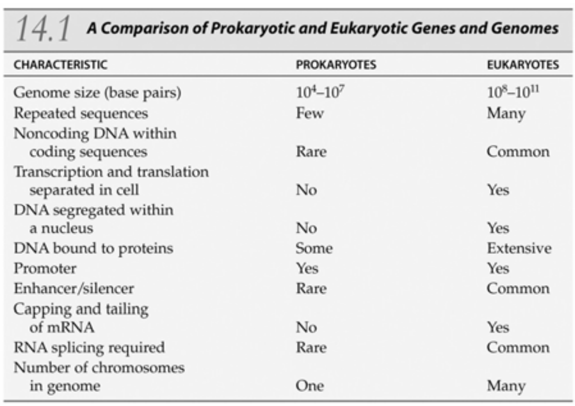
A. The Eucaryotic Genome

- Although eucaryotes have more DNA in their genomes than bacteria and archaea, in some cases there is NO apparent relationship between genome size and organism complexity.





Amoeba dubia is the big winner at **670 Billion** base pairs per cell and an uncertain phylogeny!

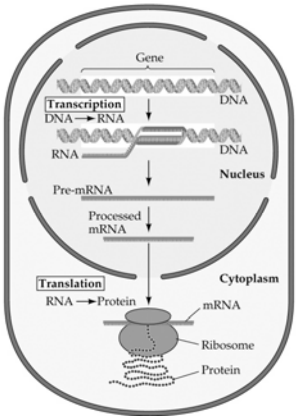


14.1 **A Comparison of Prokaryotic and Eukaryotic Genes and Genomes**

CHARACTERISTIC	PROKARYOTES	EUKARYOTES
Genome size (base pairs)	10^4–10^7	10^8–10^{11}
Repeated sequences	Few	Many
Noncoding DNA within coding sequences	Rare	Common
Transcription and translation separated in cell	No	Yes
DNA segregated within a nucleus	No	Yes
DNA bound to proteins	Some	Extensive
Promoter	Yes	Yes
Enhancer/silencer	Rare	Common
Capping and tailing of mRNA	No	Yes
RNA splicing required	Rare	Common
Number of chromosomes in genome	One	Many

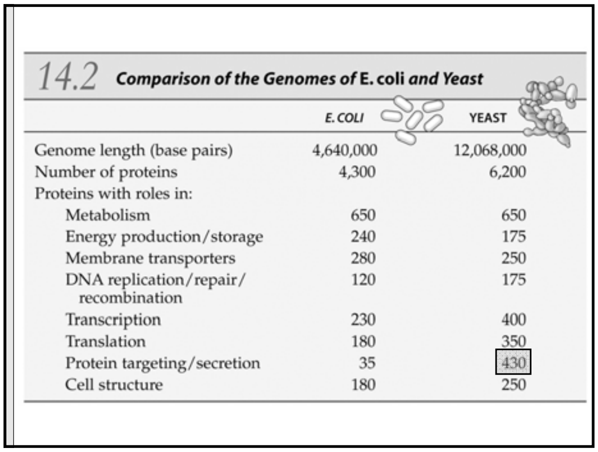
A. The Eucaryotic Genome

- Unlike bacterial or archaeal DNA, eukaryotic DNA is separated from the cytoplasm by being contained within a nucleus.
- The initial mRNA transcript of the DNA gets modified before it is exported to the cytoplasm.



A. The Eucaryotic Genome

- The genome of the single-celled budding yeast contains genes for the same metabolic machinery as bacteria, as well as genes for protein targeting in the cell.

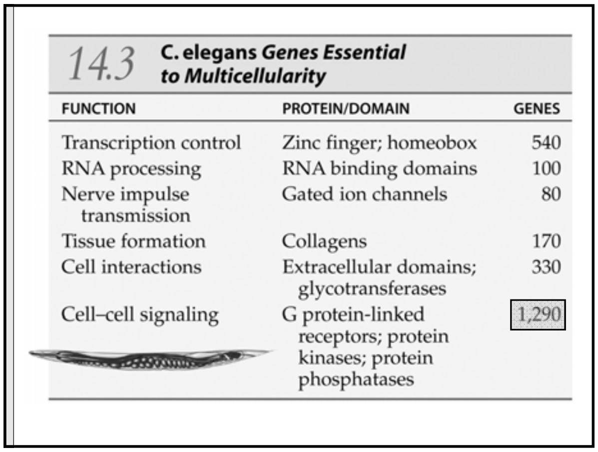


14.2 Comparison of the Genomes of E. coli and Yeast

	E. COLI	YEAST
Genome length (base pairs)	4,640,000	12,068,000
Number of proteins	4,300	6,200
Proteins with roles in:		
Metabolism	650	650
Energy production/storage	240	175
Membrane transporters	280	250
DNA replication/repair/recombination	120	175
Transcription	230	400
Translation	180	350
Protein targeting/secretion	35	430
Cell structure	180	250

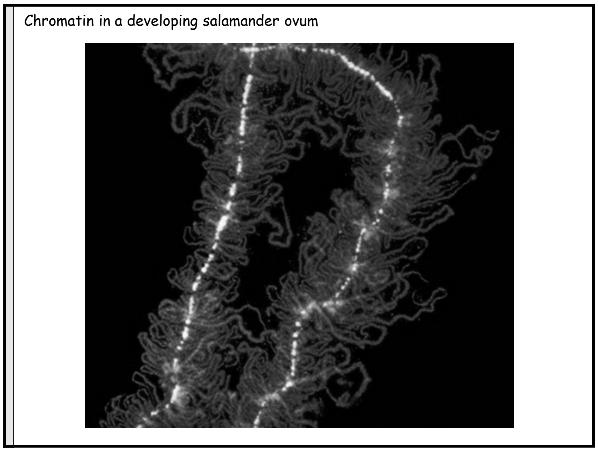
A. The Eucaryotic Genome

- The genome of the multicellular roundworm *Caenorhabditis elegans* contains genes required for intercellular interactions.
- The genome of the fruit fly has fewer genes than that of the roundworm. Many of its sequences are homologs of sequences on roundworm and mammalian genes.

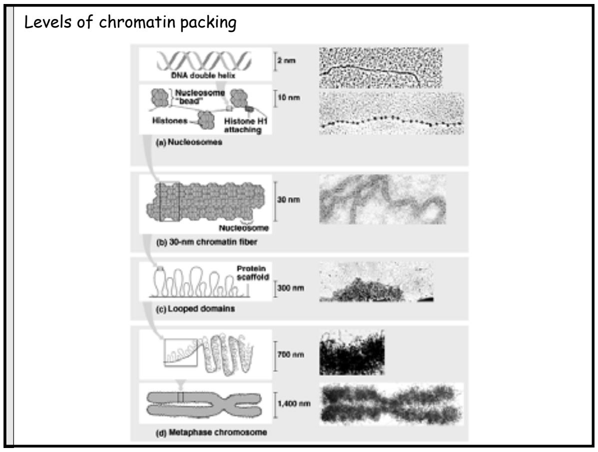


14.3 C. elegans Genes Essential to Multicellularity

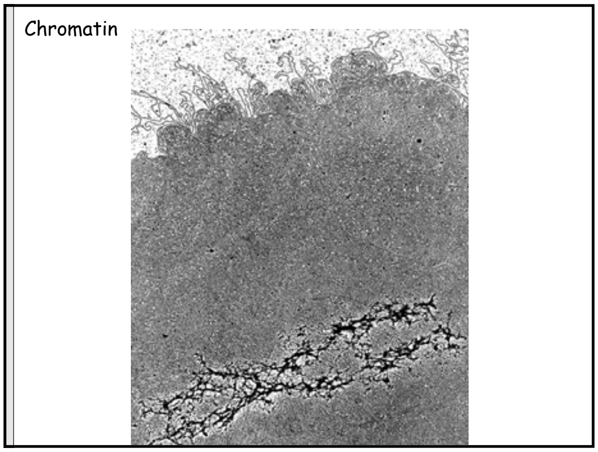
FUNCTION	PROTEIN/DOMAIN	GENES
Transcription control	Zinc finger; homeobox	540
RNA processing	RNA binding domains	100
Nerve impulse transmission	Gated ion channels	80
Tissue formation	Collagens	170
Cell interactions	Extracellular domains; glycotransferases	330
Cell–cell signaling	G protein-linked receptors; protein kinases; protein phosphatases	1,290



Chromatin in a developing salamander ovum

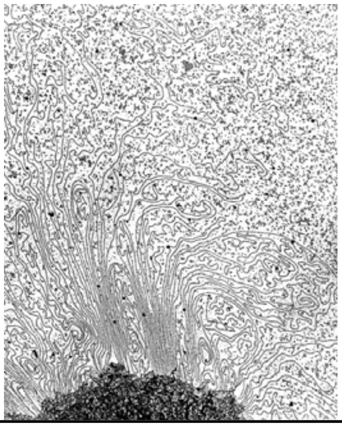


Levels of chromatin packing



Chromatin

Chromatin, detail

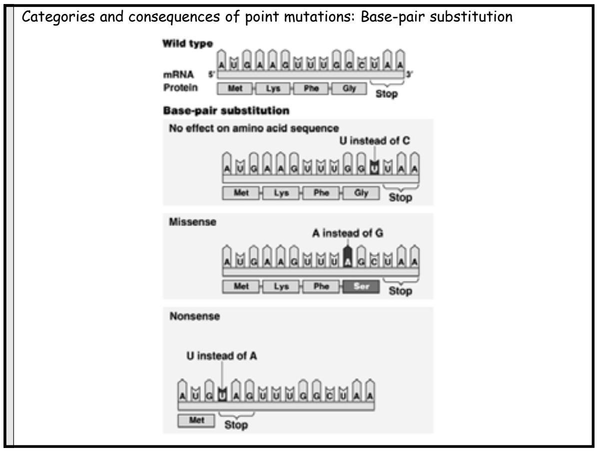


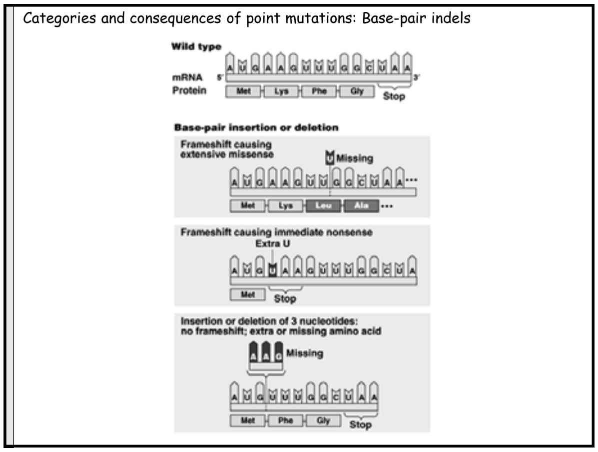
B. Mutations: Heritable Changes in Genes

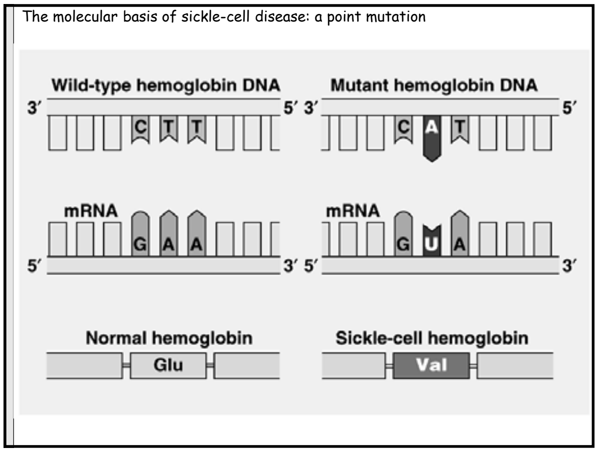
- Mutations in DNA are often expressed as abnormal proteins. However, the result may not be easily observable phenotypic changes.
- Raw materials for evolution to operate.
- Some mutations appear only under certain conditions, such as exposure to a certain environmental agent or condition.

B. Mutations: Heritable Changes in Genes

- Point mutations (silent, missense, nonsense, or frame-shift) result from alterations in single base pairs of DNA.


Categories and consequences of point mutations: Base-pair substitution

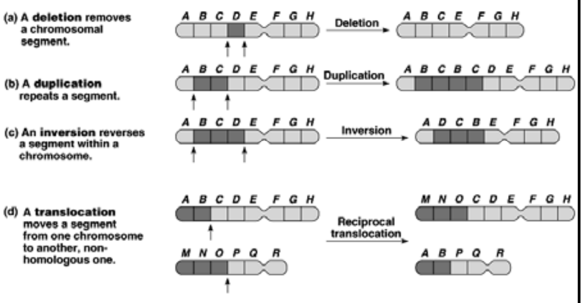

Categories and consequences of point mutations: Base-pair indels


The molecular basis of sickle-cell disease: a point mutation

B. Mutations: Heritable Changes in Genes

- Chromosomal mutations (deletions, duplications, inversions, or translocations) involve large regions of a chromosome.

Alterations of chromosome structure



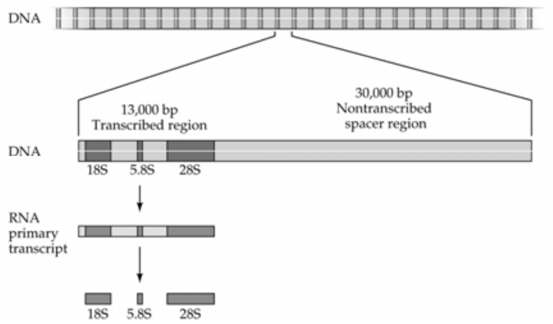
C. Repetitive Sequences

- Highly repetitive DNA is present in up to millions of copies of short sequences. It is not transcribed. Its role is unknown.
- Rem: Some moderately repetitive DNA sequences, such as telomeric DNA is found at the ends of chromosomes.

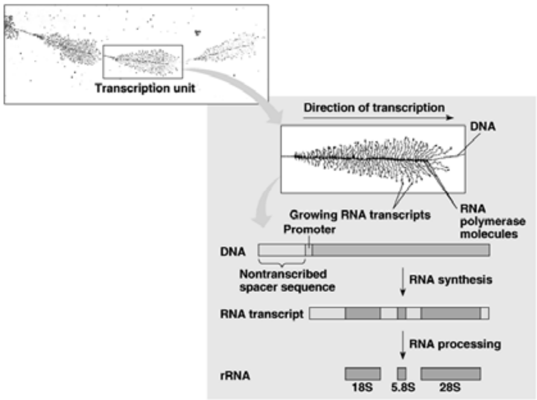
C. Repetitive Sequences

- Some moderately repetitive DNA sequences, such as those coding for ribosomal RNA's, are transcribed.
- Up to three rRNAs result, two go to the large subunit and one goes to the small subunit.

Moderately repetitive DNA sequences

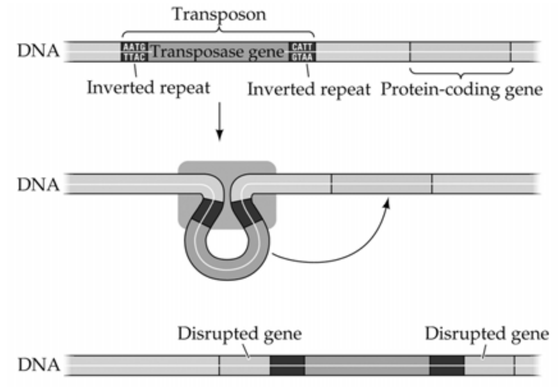


Part of a family of identical genes for ribosomal RNA

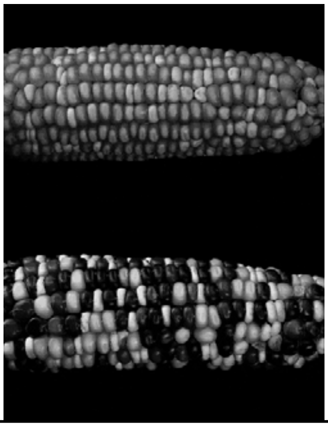


C. Repetitive Sequences

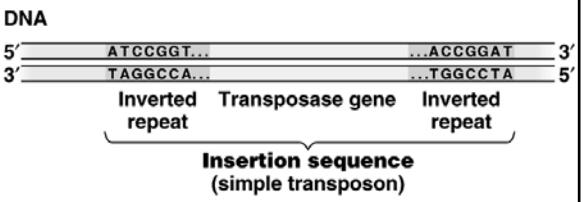
- Some moderately repetitive DNA sequences are transposable, or able to move about the genome. These are known as Transposons.
- Transposons can jump from place to place on the chromosome by actually moving or by making a new copy, inserted at a new location.



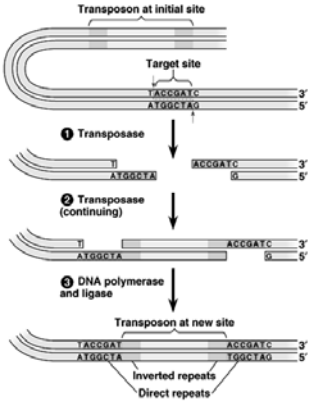
Transposons in corn



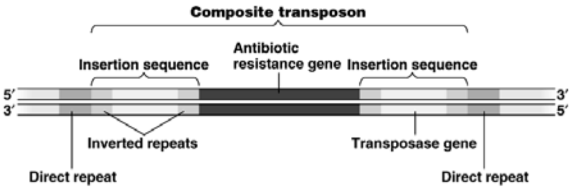
Insertion sequences, the simplest transposons



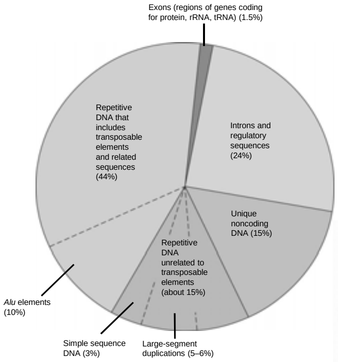
Insertion of a transposon and creation of direct repeats



Anatomy of a composite transposon

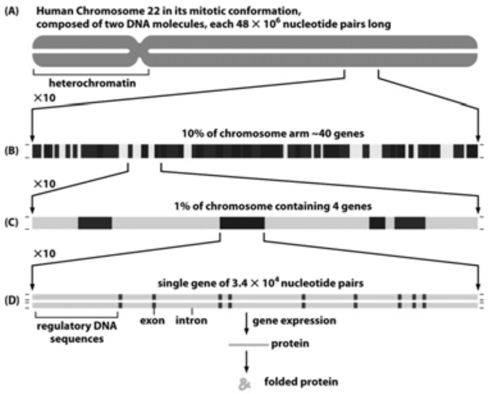


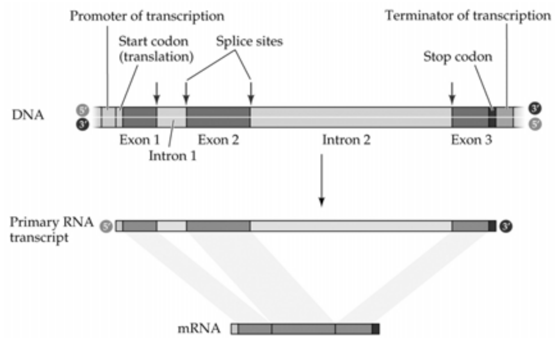
Types of DNA sequences in the human genome

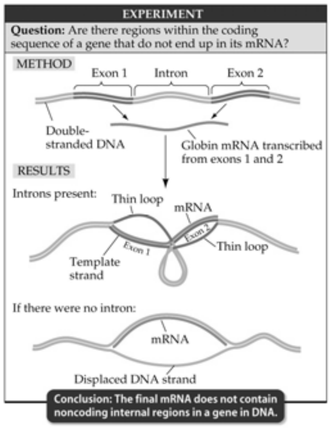


D. The Structures of Protein-Coding Genes

- A typical protein-coding gene has noncoding internal sequences (introns) as well as flanking sequences that are involved in the machinery of transcription and translation in addition to its exons or coding regions.
- These are usually single copy genes.






EXPERIMENT

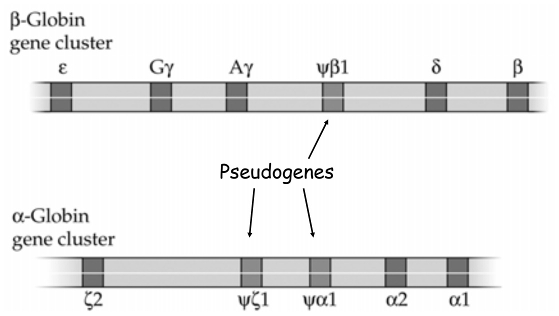
Question: Are there regions within the coding sequence of a gene that do not end up in its mRNA?

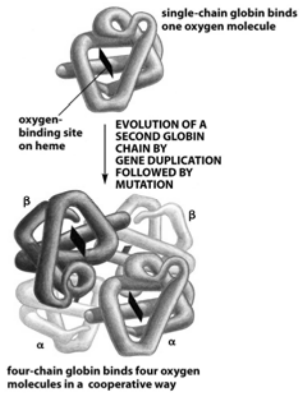
Conclusion: The final mRNA does not contain noncoding internal regions in a gene in DNA.

D. The Structures of Protein-Coding Genes

- Some eucaryotic genes form families of related genes that have similar sequences and code for similar proteins. These related proteins may be made at different times and in different tissues.
- Some sequences in gene families are pseudogenes, which code for nonfunctional mRNA's or proteins.

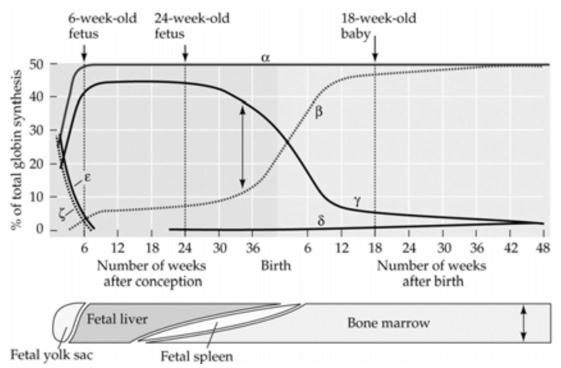
Gene Families



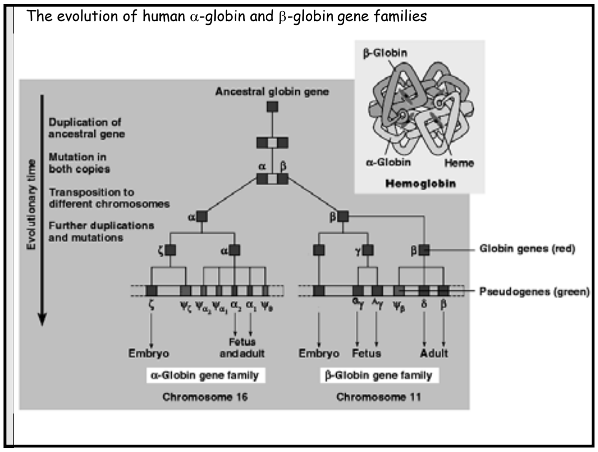


D. The Structures of Protein-Coding Genes

- Differential expression of different genes in the β-globin family ensures important physiological changes during human development.

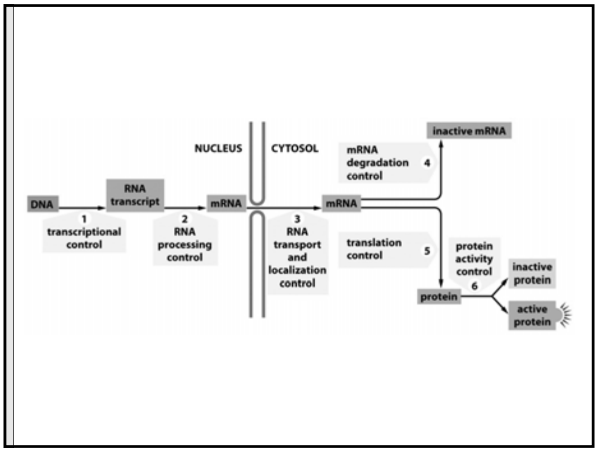


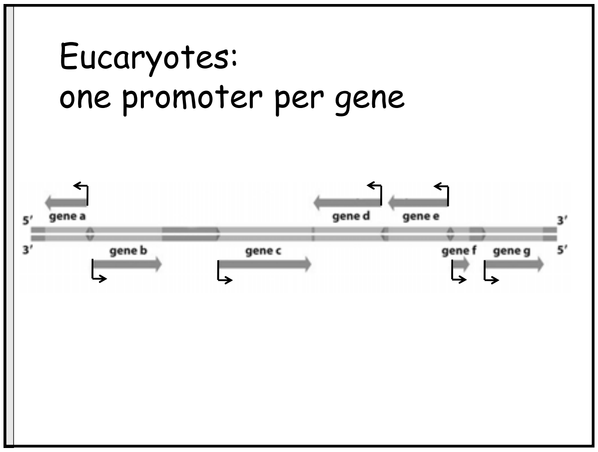
The evolution of human α-globin and β-globin gene families



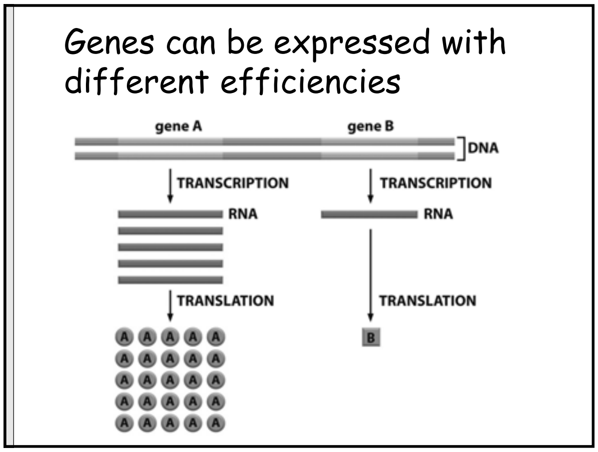
E. Differential Gene Expression

- Eucaryotic gene expression can be controlled at the transcriptional, posttranscriptional, translational, and posttranslational levels.

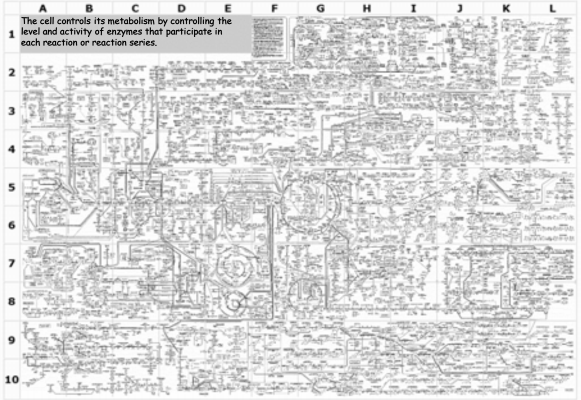




Eucaryotes:
one promoter per gene



Genes can be expressed with different efficiencies



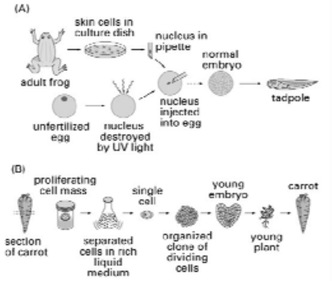
The cell controls its metabolism by controlling the level and activity of enzymes that participate in each reaction or reaction series.

E. Differential Gene Expression

?

Genes aren't lost during development, but rather each cell becomes more and more restricted in its fate, expressing ultimately a specific subset of genes responsible for defining its specific function.

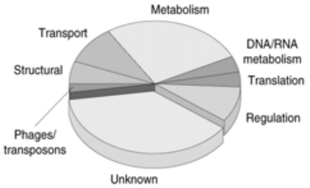
E. Differential Gene Expression



- We can show that all the information to make an organism resides in every cell.
- Theoretically, every cell could be used to regenerate a genetically identical adult (clone).
- Cells that are capable of regenrating a fully formed adult are called **totipotent**.

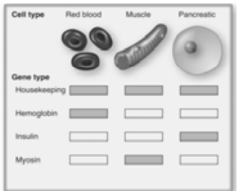
Why not synthesis all the genes all the time at a moderate level?

- Too expensive
- Levels need to be controlled
- Some products are incompatible
- Need change in response to signals
- Development



Principles of gene control

- Constitutive expression
  - A gene is expressed at approximately the same levels all the time: (for example: a housekeeping gene)
- Regulated expression
  - Gene expression in response to a signal



Cell type	Red blood	Muscle	Pancreatic
Gene type			
Housekeeping	▨	▨	▨
Hemoglobin	▨	☐	☐
Insulin	☐	☐	▨
Myosin	☐	▨	☐

F. Transcriptional Control

- The major method of control of eucaryotic gene expression is selective transcription, which results from specific proteins binding to regulatory regions on DNA.

RNA polymerases:
the more the merrier…

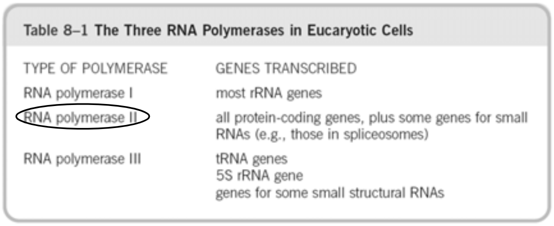


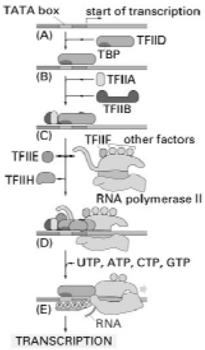
Table 8–1 The Three RNA Polymerases in Eucaryotic Cells

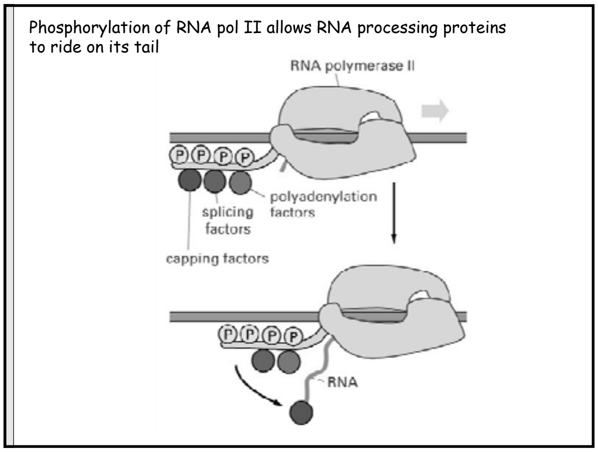
TYPE OF POLYMERASE	GENES TRANSCRIBED
RNA polymerase I	most rRNA genes
RNA polymerase II	all protein-coding genes, plus some genes for small RNAs (e.g., those in spliceosomes)
RNA polymerase III	tRNA genes 5S rRNA gene genes for some small structural RNAs

F. Transcriptional Control

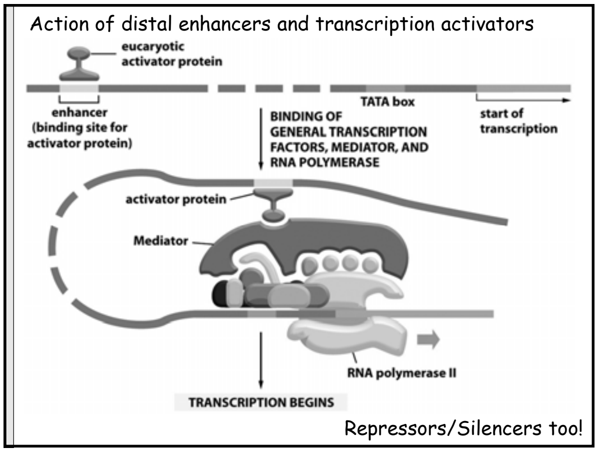
- A series of "general" transcription factors must bind to the promoter before RNA polymerase can bind.
- Whether RNA polymerase will initiate transcription also depends on the binding of regulatory proteins, activator proteins, and repressor proteins.

RNA pol II requires many "general" transcription factors



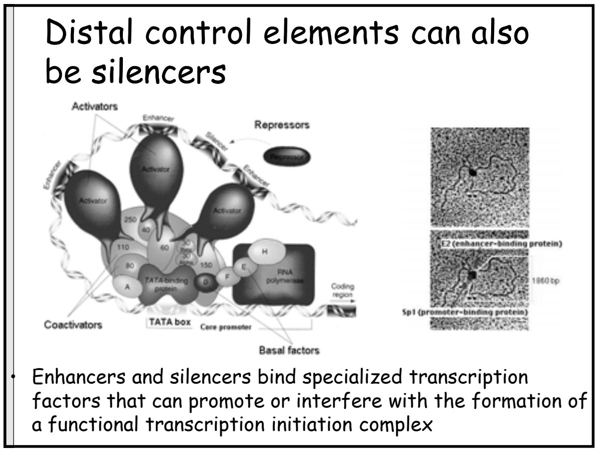


Phosphorylation of RNA pol II allows RNA processing proteins to ride on its tail



Action of distal enhancers and transcription activators

Repressors/Silencers too!



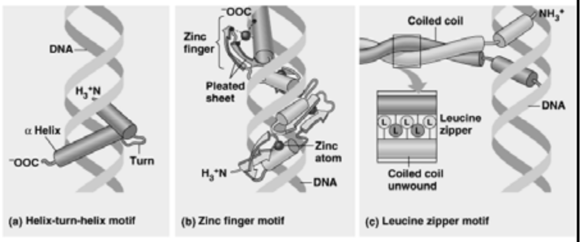
Distal control elements can also be silencers

- Enhancers and silencers bind specialized transcription factors that can promote or interfere with the formation of a functional transcription initiation complex

F. Transcriptional Control

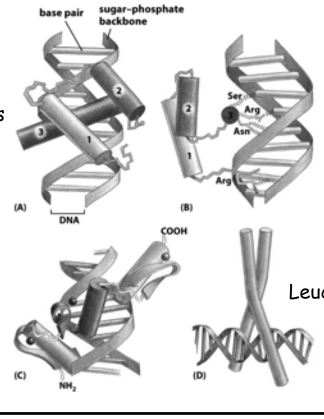
- The DNA-binding domains of most DNA-binding proteins have one of four structural motifs: helix-turn-helix, zinc finger, leucine zipper, or homeodomain.

Three of the major types of DNA-binding domains in transcription factors



(a) Helix-turn-helix motif (b) Zinc finger motif (c) Leucine zipper motif

Homeodomains



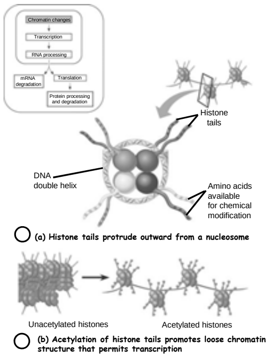
Leucine Zipper

Zinc Finger

F. Transcriptional Control

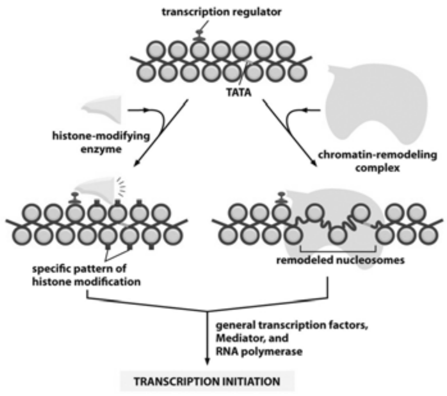
- Acetylation of histone tails promotes loose chromatin structure that permits transcription to more readily occur.

A simple model of histone tails and the effect of histone acetylation

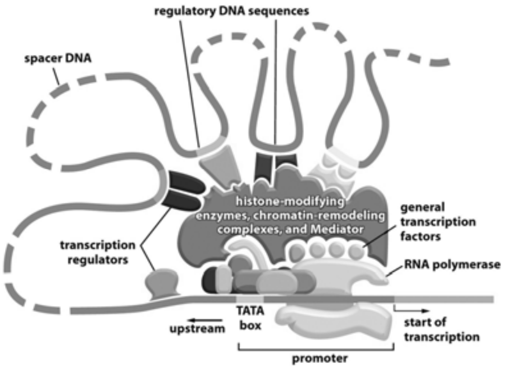


(a) Histone tails protrude outward from a nucleosome

(b) Acetylation of histone tails promotes loose chromatin structure that permits transcription



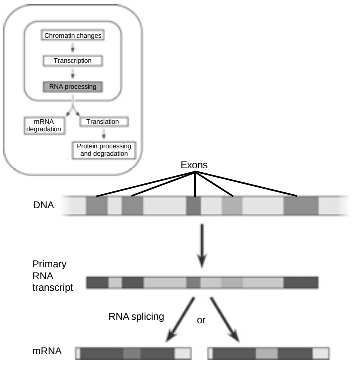
Combinatorial control regulation concept



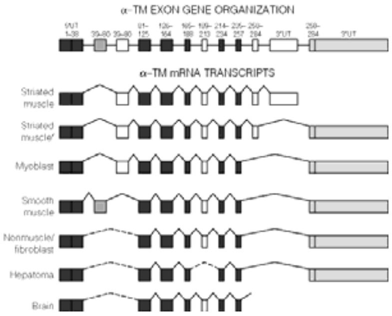
G. Posttranscriptional Control

- Because eucaryotic genes have several exons, alterative mRNAs can be generated from the same RNA transcript.
- This alternate splicing can be used to produce different proteins.
- The stability of mRNA in the cytoplasm can be regulated by the binding of proteins.

Alternative RNA splicing



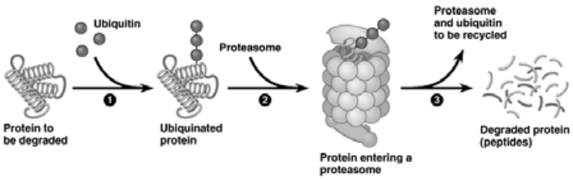
Alternative RNA splicing



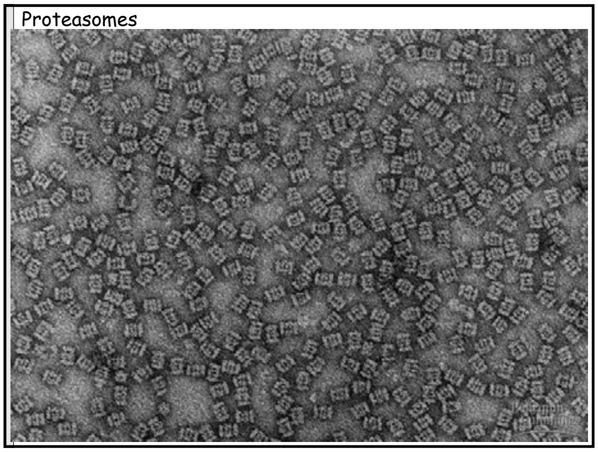
G. Posttranslational Control

- Proteasomes degrade proteins targeted for breakdown.

Degradation of a protein by a proteasome



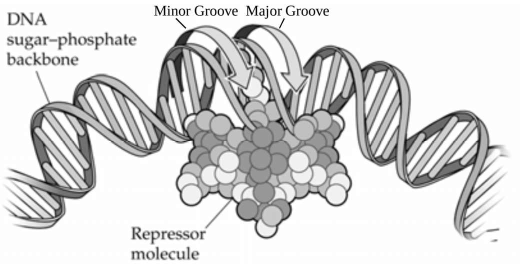
Proteasomes



H. Regulation of Gene Expression in Bacteria

- An operon consists of a promoter, an operator, and structural genes. Promoters and operators do not code for proteins, but serve as binding sites for regulatory proteins.
- When a repressor protein binds to the operator, transcription of the structural genes is inhibited.

Repressor Bound to an Operator Blocks Transcription

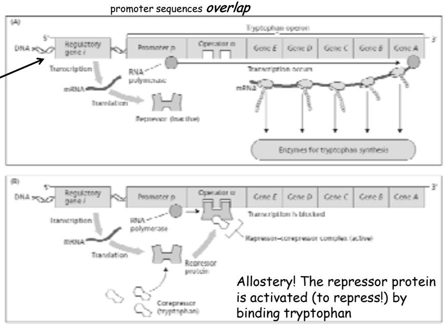


H. Regulation of Gene Expression in Bacteria

- The expression of bacterial genes is regulated by: inducible operator–repressor systems, repressible operator–repressor systems (e.g., both negative control), and systems that increase the efficiency of a promoter (e.g., positive control).
- Repressor proteins are coded by constitutive regulatory genes.

The tryptophan operon:

a biosynthetic operon controlled by a **repressor**

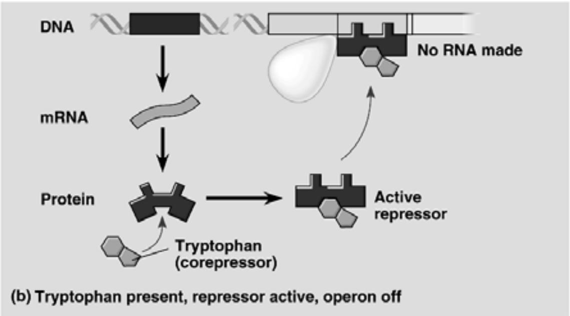


NOTE! The operator and promoter sequences *overlap*

This gene has its own promoter and is constitutively expressed at a low levels.

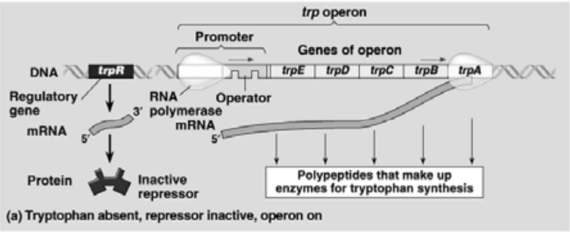
Allostery! The repressor protein is activated (to repress!) by binding tryptophan

The *trp* operon: regulated synthesis of repressible enzymes



(b) Tryptophan present, repressor active, operon off

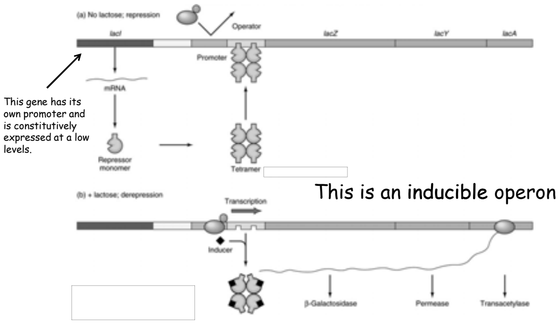
The *trp* operon: regulated synthesis of repressible enzymes



(a) Tryptophan absent, repressor inactive, operon on

The lac operon:

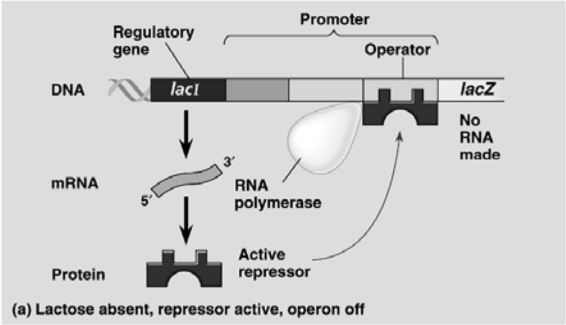
a catabolic operon controlled by a repressor



This gene has its own promoter and is constitutively expressed at a low levels.

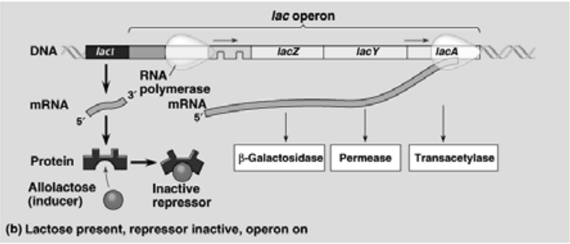
This is an inducible operon

The *lac* operon: regulated synthesis of inducible enzymes



(a) Lactose absent, repressor active, operon off

The *lac* operon: regulated synthesis of inducible enzymes

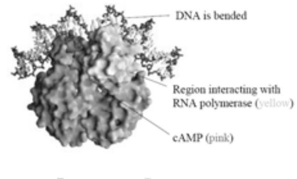


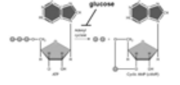
(b) Lactose present, repressor inactive, operon on

H. Regulation of Gene Expression in Bacteria

- The efficiency of RNA polymerase can be increased by regulation of the level of cyclic AMP, which binds to CAP (cAMP activator protein).
- The CAP–cAMP complex then binds to a site near the promoter of a target gene, enhancing the binding of RNA polymerase and hence transcription.

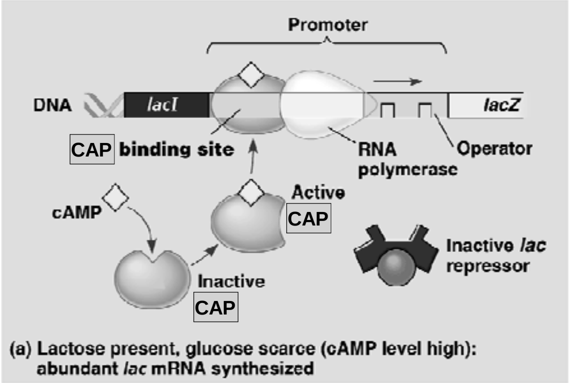
The lac operon: positive control





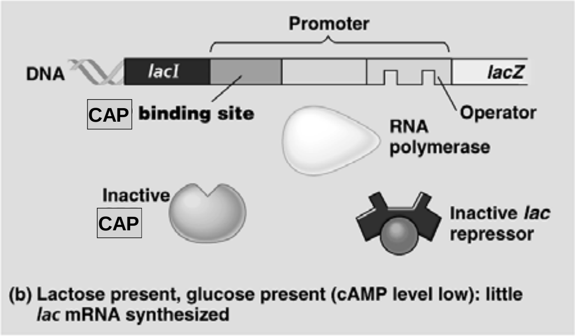
The presence of glucose prevents the transcription of the lac operon.

Positive control: cAMP activator protein

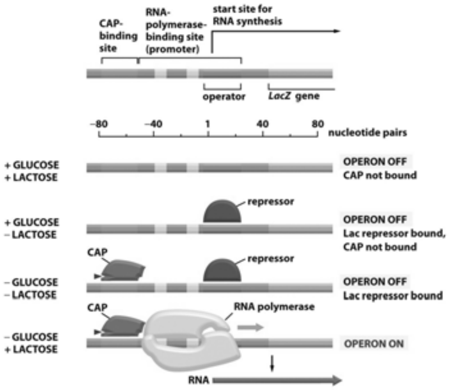


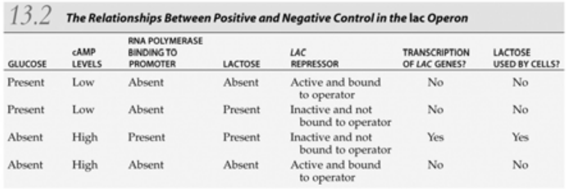
(a) Lactose present, glucose scarce (cAMP level high): abundant *lac* mRNA synthesized

Positive control: cAMP activator protein



(b) Lactose present, glucose present (cAMP level low): little *lac* mRNA synthesized





13.2 *The Relationships Between Positive and Negative Control in the lac Operon*

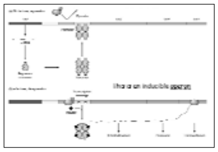
GLUCOSE	cAMP LEVELS	RNA POLYMERASE BINDING TO PROMOTER	LACTOSE	*LAC* REPRESSOR	TRANSCRIPTION OF *LAC* GENES?	LACTOSE USED BY CELLS?
Present	Low	Absent	Absent	Active and bound to operator	No	No
Present	Low	Absent	Present	Inactive and not bound to operator	No	No
Absent	High	Present	Present	Inactive and not bound to operator	Yes	Yes
Absent	High	Absent	Absent	Active and bound to operator	No	No

Operons: Review

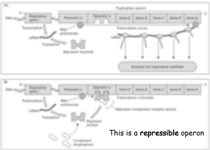
Inducible vs. repressible operons

Defined by response of operon to a metabolite (small molecule).

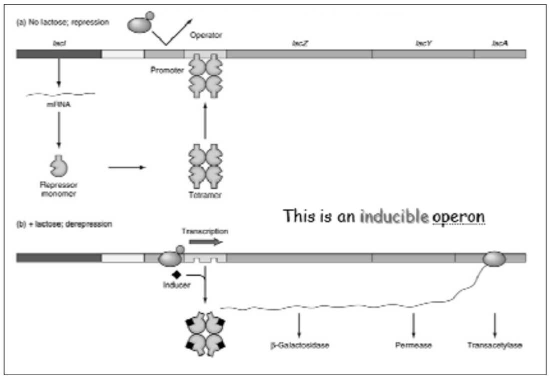
Type of operon	Presence of	Effect	Examples Metabolite	 Operon
Inducible	metabolite	ON	lactose	*lac*
Repressible	metabolite	OFF	Trp	*trp*



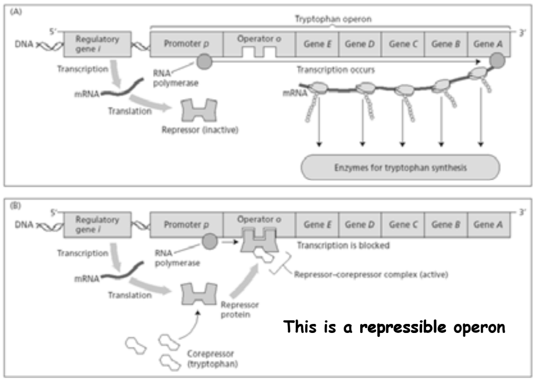
This is an inducible operon



This is a repressible operon



This is an inducible operon



This is a repressible operon

I. Comparison of Control Features in Bacteria & Eucarya

- Bacteria have multiple genes under single control: operons
- Eucarya have multiple RNA polymerases
- Simple vs. Complex Transcription Factors
- Local vs. Distal Control: Enhancers/Silencers
- Eucarya must contend with Chromatin

What are eucaryotic-specific control issues?

• Distal control elements • Chromatin	→ Transcriptional control
• Splicing • mRNA transport	→ Post-transcriptional control
• Protein transport • Protein modifications	→ Post-translational control
• Multicellularity	→ Signal transduction