

# Foundation of Biology II

**Google Classroom: 6yeub3ts**

Gaurav Ahuja Ph.D  
Associate Professor

Department of Computational Biology  
Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi)  
(A State University Established by Govt. of NCT of Delhi)

Okhla, Phase III  
New Delhi-110020, INDIA  
Email: gaurav.ahuja@iiitd.ac.in  
Website: <https://ahuja-lab.in/>  
Phone: (+91)11-26907475

# Attendance

**QR will be changed for every class  
so NO PHOTOGRAPH PLEASE**

**SCAN ME**



# Course Structure

Topic	Sub-topics
1	Cellular Signaling, Metabolism and Communication Signal Transduction Pathways, Hormonal Regulation and Second Messengers, Cell Metabolism
2	Molecular Genetics: DNA Repair and Recombination DNA Damage and Repair Mechanisms, Homologous Recombination and Genetic Recombination
3	Epigenetics and Gene Regulation DNA Methylation and Histone Modification, Epigenetic Inheritance and Phenotypic Variation
4	Biotechnology and Genetic Engineering Recombinant DNA Technology, CRISPR Cas9 and Genome Editing
5	Phylogenetics and Systematics Molecular Phylogenetics, Cladistics and Taxonomy
6	Population Genetics and Evolutionary Ecology Hardy Weinberg Equilibrium, Evolutionary Game Theory
7	Evolutionary Developmental Biology (EvoDevo) Evolutionary Basis of Developmental Processes, Genetic Regulation of Morphological Evolution
8	Immunology and Host Pathogen Interactions, Innate and Adaptive Immunity, Immunological Memory and Vaccines
9	Neurobiology and Behavior Neural Signaling and Synaptic Transmission, Neural Basis of Behavior and Cognitive Processes
10	Endocrinology and Hormonal Regulation Endocrine Glands and Hormone Function, Hormonal Regulation of Physiological Processes
11	Frontiers in Biology: Emerging Technologies and Research Areas Genomics and Big Data in Biology, Synthetic Biology and Bioinformatics, Biomedical Innovations and Future Directions

# Course Assessment

Type of Evaluation	% Contribution in Grade
--------------------	-------------------------

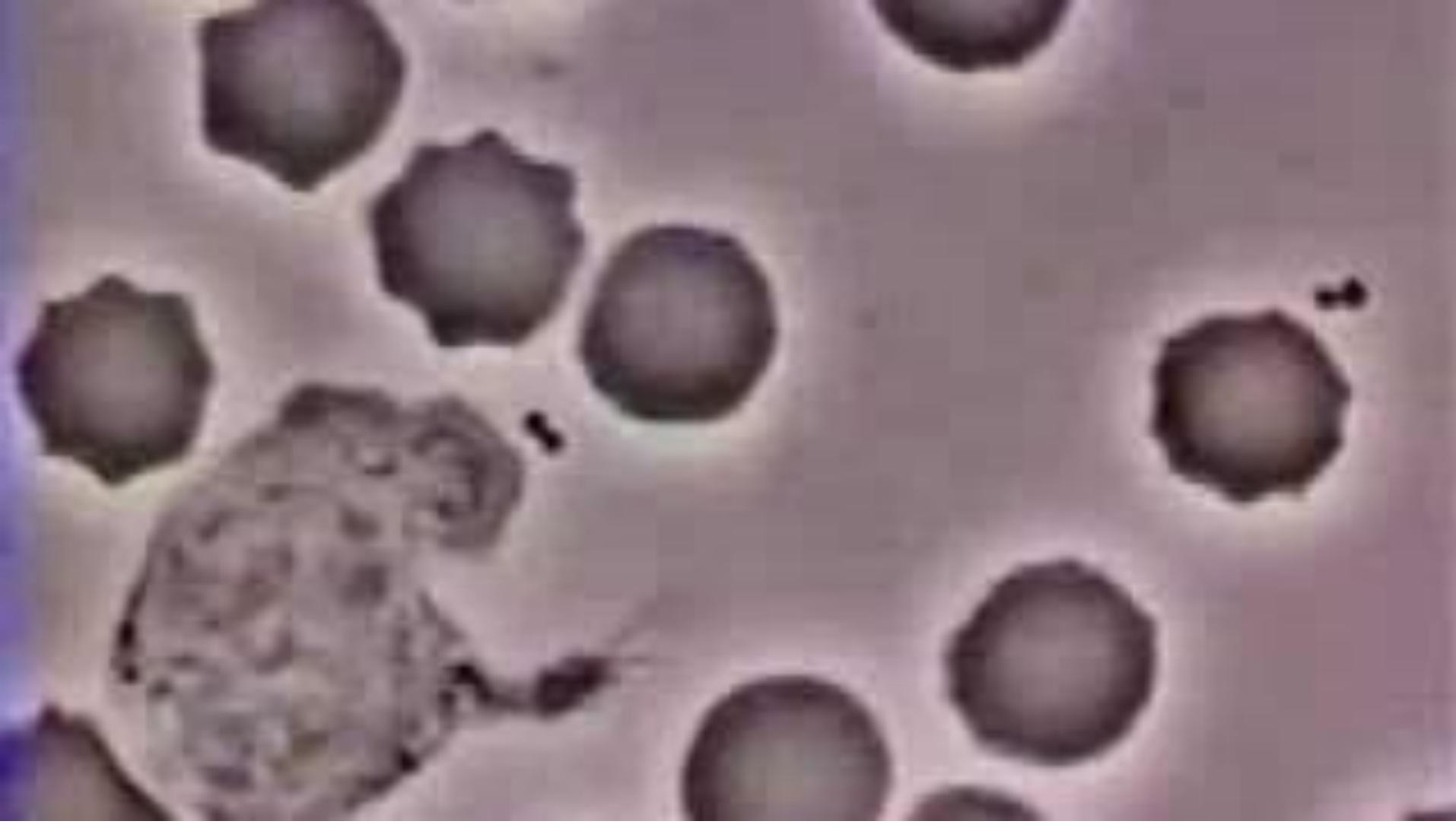
• Mid-sem	30
• End-sem	30
• Assignments (2)	10 (best one)
• Quiz (3 or 4)	10 (Best of two)
• In Class Assessment	10
• Attendance	10

- Lehninger Principles of Biochemistry, Nelson and Cox
- Essential cell biology 4th Ed. by Bruce Alberts, Dennis Bray, Karen Hopkin Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, Peter Walter
- APA. Voet, D., Voet, J. G., & Pratt, C. W. (2016). Fundamentals of biochemistry (5th ed.). John Wiley & Sons.

# **Expected Outcome**

- Explain core principles of cellular communication, metabolism, and molecular genetics, including DNA repair and gene regulation.
- Analyze and compare modern biotechnological techniques, such as CRISPR-Cas9, and their applications in genetic engineering.
- Describe fundamental concepts in evolutionary biology, including phylogenetics, population genetics, and developmental processes.
- Understand the mechanisms of immunological responses, neural signaling, and hormonal regulation within biological systems.

# **Cellular Signaling, Metabolism and Communication**



# Cell Signaling and Signal Transduction: Communication between Cells

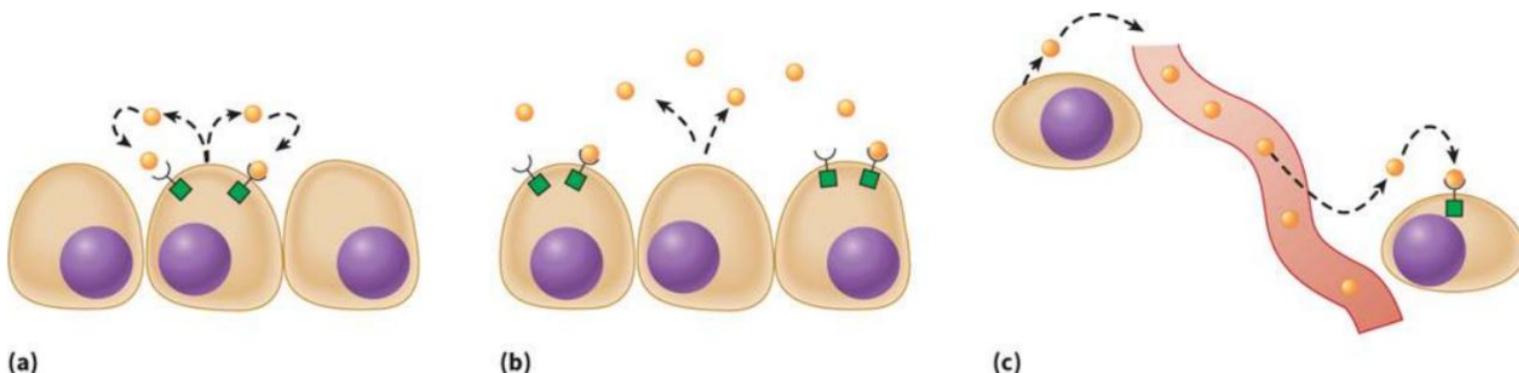


The Hawaiian bobtail squid maintains a symbiotic relationship with quorum-sensing bacteria. When present at high concentrations, the bacteria *Aliivibrio fischeri* produce light that allows the squid to hunt for prey without casting a shadow.

# Cell Signaling and Signal Transduction: Communication between Cells

*The English poet John Donne expressed his belief in the interdependence of humans in the phrase “No man is an island.”*

- Most cells in a plant or an animal are specialized to carry out one or more specific functions.
- Many biological processes require various cells to work together and coordinate their activities. To make this possible, cells have to communicate with each other, which is accomplished by a process called cell signaling.
- Cell signaling makes it possible for cells to respond in an appropriate manner to a specific environmental stimulus.



**FIGURE 15.1** Autocrine (a), paracrine (b), and endocrine (c) types of cell signaling.

### Autocrine Signaling:

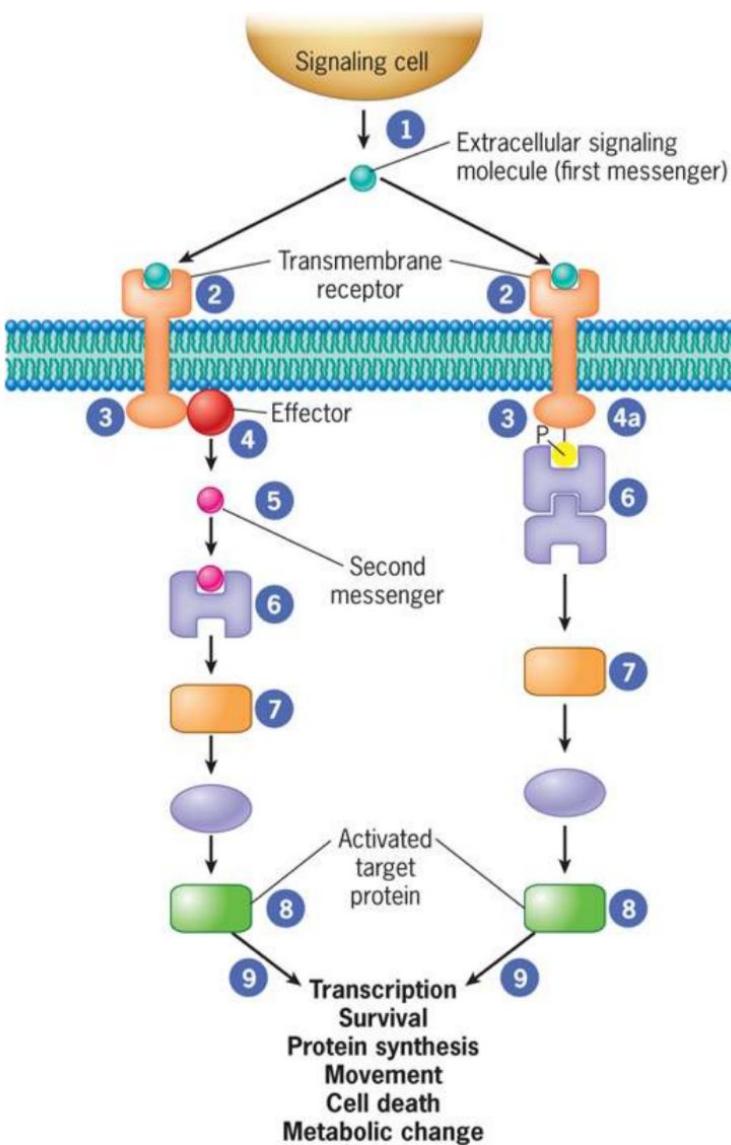
- A cell targets itself.
- The cell releases a signal and has its own receptors to receive it.

### Paracrine Signaling:

- A cell signals nearby cells.
- Messengers travel a short distance and are often unstable or quickly degraded.

### Endocrine Signaling:

- A cell signals distant cells throughout the body.
- Messengers, called hormones, travel through the bloodstream to reach their targets.



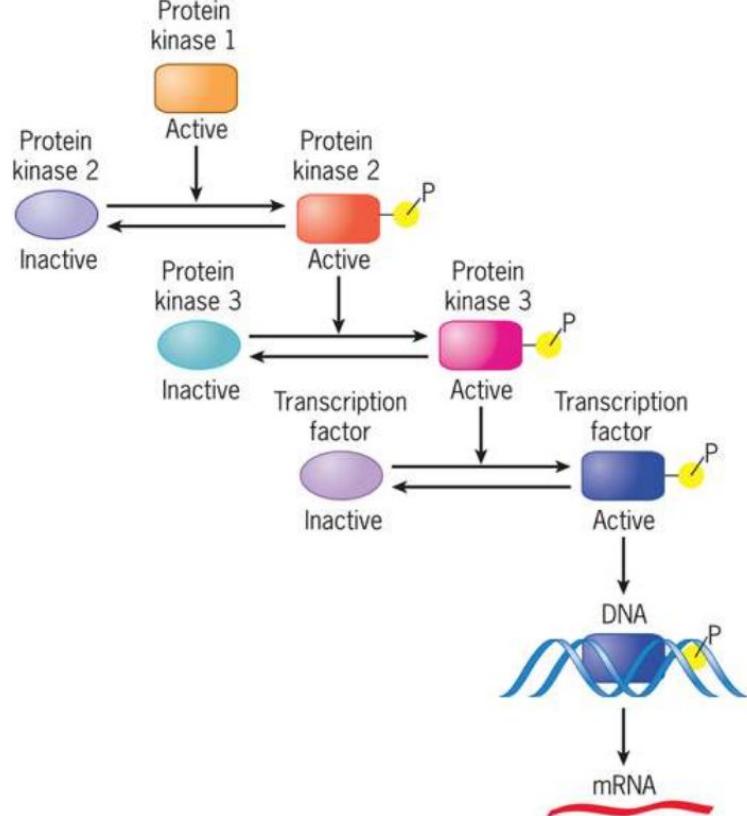
## How Signals Enter the Cell

- When an extracellular messenger binds to a **receptor** on a cell's surface, it causes the receptor to change shape. This relays the signal to the part of the receptor inside the cell. From there, the signal is typically transmitted further into the cell by one of two major routes.

### Two Main Signal Transmission Routes

- Route 1: Second Messengers** The receptor activates a nearby enzyme, called an **effector**. This effector then generates small molecules known as **second messengers**. These second messengers travel within the cell to activate or inactivate other proteins, carrying the signal forward.
- Route 2: Protein Recruitment** The receptor's internal domain becomes a recruiting station, directly attracting and activating various cellular **signaling proteins**.

## How Signaling Pathways Operate



## The Key "On/Off" Switch: Phosphorylation

This change in protein shape is typically accomplished by adding or removing phosphate groups.

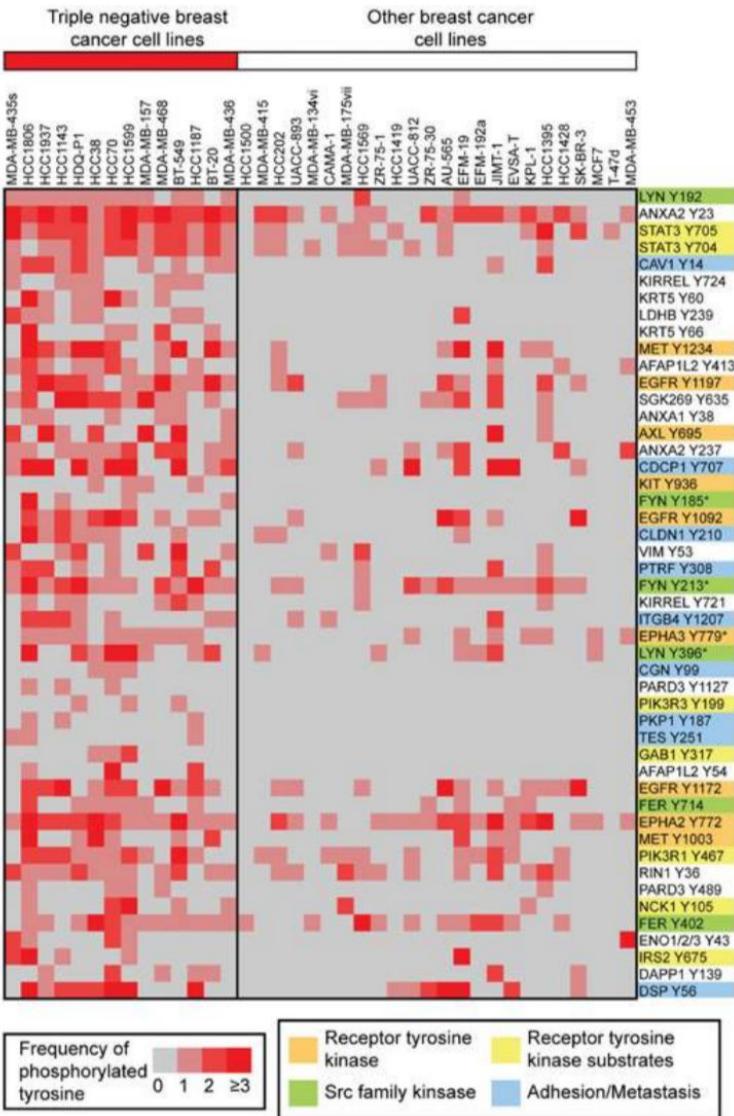
- **Protein Kinases:** These are enzymes that **add** phosphate groups to other proteins, usually activating them.
- **Protein Phosphatases:** These enzymes **remove** phosphate groups, often deactivating the protein.

Think of kinases as the "on" switch and phosphatases as the "off" switch for a protein's activity.

## Enzyme Structure and Specificity

While both are crucial, they are structured a bit differently:

- **Kinases** often work as single-unit enzymes.
- **Phosphatases** are typically more complex. They have a catalytic subunit that does the work (removing the phosphate) and a regulatory subunit that directs it to the correct target protein. This allows a single type of phosphatase to be highly specific and versatile.



# Kinase Targets and Specificity

Most protein kinases add phosphate groups to serine or threonine amino acids, while a smaller, important group targets tyrosine residues. Kinases can be found either free in the cytoplasm or as part of a cell membrane.

Despite there being thousands of potential phosphorylation sites in a cell, kinases and phosphatases are highly specific, acting only on their intended protein substrates. Some have many protein targets, while others are so specific they only modify a single site on a single protein.

## The Widespread Effects of Phosphorylation

Adding a phosphate group is a versatile molecular switch that can profoundly change a protein's function. Phosphorylation can:

- Activate or inactivate an enzyme.
- Increase or decrease protein-protein interactions.
- Trigger a protein to move to a different location within the cell.
- Act as a signal to mark the protein for degradation.

# Types of Cellular Messengers

Cells use a wide variety of molecules to send signals to one another. These include:

- **Amino Acids & Derivatives:** Small molecules like glutamate, acetylcholine, and epinephrine that often act as neurotransmitters and hormones.
- **Gases:** Simple gases like Nitric Oxide (NO) and Carbon Monoxide (CO).
- **Steroids:** Derived from cholesterol, these hormones regulate processes like sexual differentiation and metabolism.
- **Eicosanoids:** Molecules derived from fatty acids that regulate pain, inflammation, blood pressure, and clotting. (Drugs like aspirin and ibuprofen work by blocking their synthesis).
- **Proteins & Polypeptides:** A huge and diverse group that regulates cell division, differentiation, immune responses, and much more.

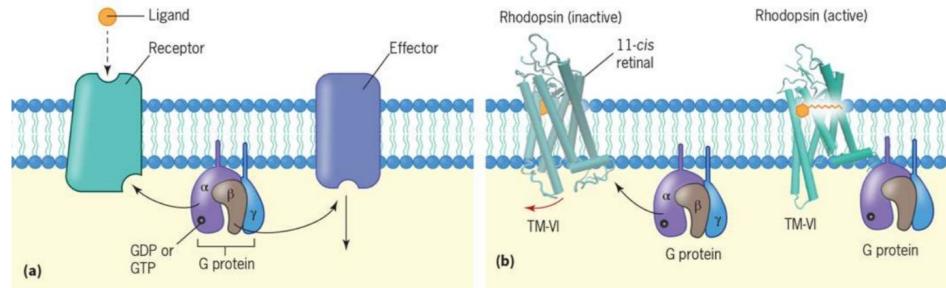
TABLE 15.1

## Examples of Physiologic Processes Mediated by GPCRs and Heterotrimeric G Proteins

Source: Adapted from L. Stryer and H. R. Bourne from the *Annual Review of Cell Biology*, Vol 2, 1986, by Annual Reviews Inc.

Stimulus	Receptor	Effector	Physiologic response
Epinephrine	β-Adrenergic receptor	Adenylyl cyclase	Glycogen breakdown
Serotonin	Serotonin receptor	Adenylyl cyclase	Behavioral sensitization and learning in <i>Aplysia</i>
Light	Rhodopsin	cGMP phosphodiesterase	Visual excitation
IgE-antigen complexes	Mast cell IgE receptor	Phospholipase C	Secretion
f-Met peptide	Chemotactic receptor	Phospholipase C	Chemotaxis
Acetylcholine	Muscarinic receptor	Potassium channel	Slowing of pacemaker activity

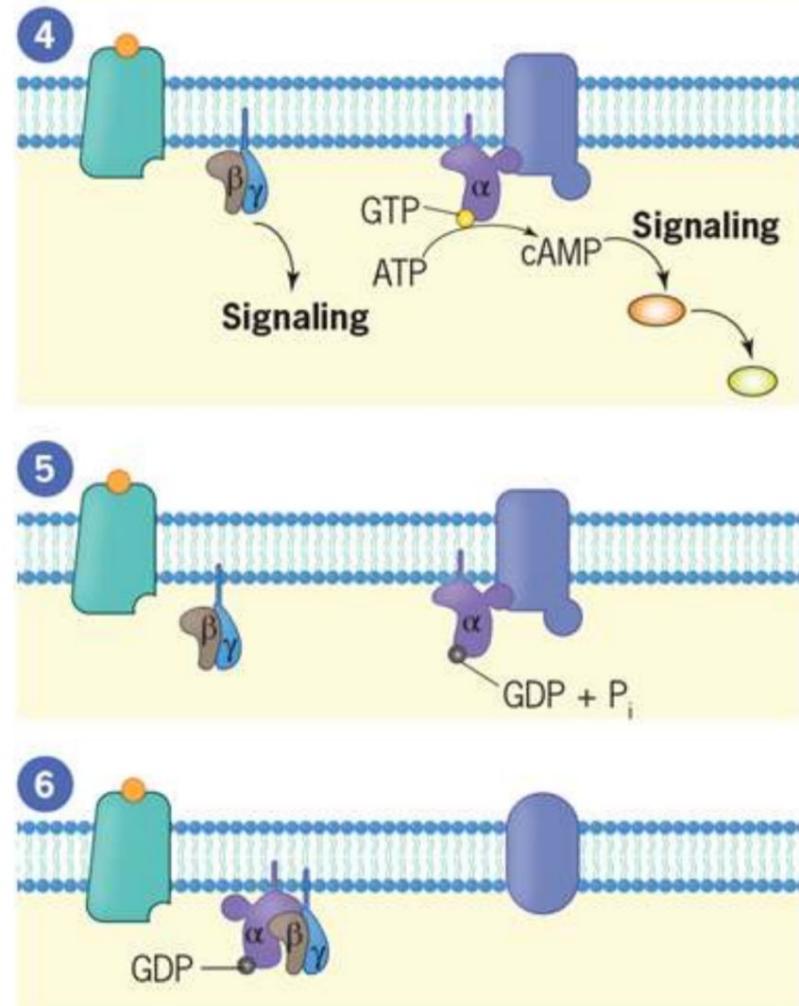
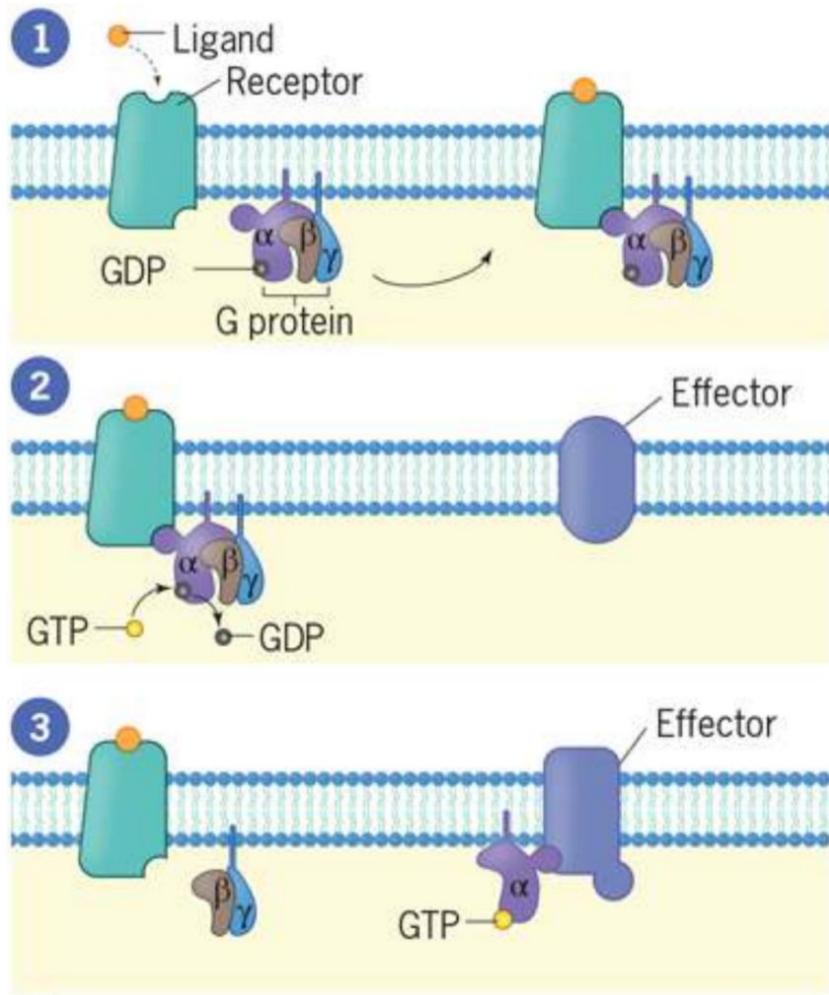
# The Receptors That Catch the Message



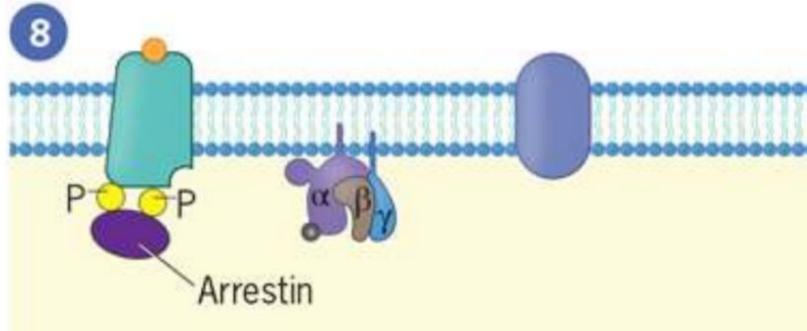
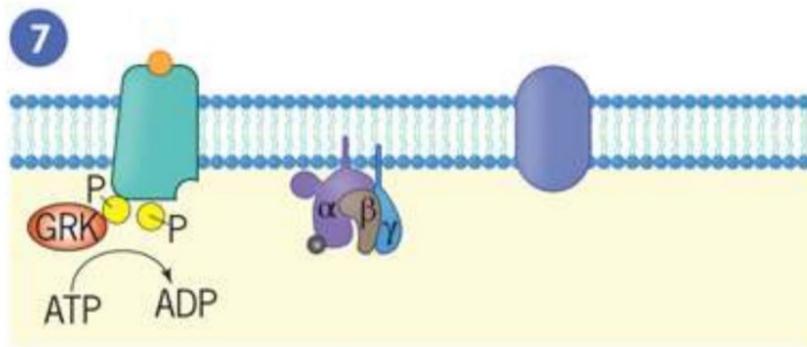
These messenger molecules are recognized by specific receptors, which translate the external signal into an internal cellular response. The major classes are:

- **G Protein-Coupled Receptors (GPCRs):** A massive family of receptors that weave through the membrane seven times. When a messenger binds, they activate an internal partner called a **G protein** to transmit the signal.
- **Receptor Protein-Tyrosine Kinases (RTKs):** When a ligand binds, these receptors pair up and activate their own internal kinase domain. They then add phosphate groups to tyrosine residues on other proteins, altering their function.
- **Ligand-Gated Channels:** These receptors are also ion channels. Binding of a ligand directly opens the channel, allowing ions to flow across the membrane. This is fundamental to nerve impulses.
- **Steroid Hormone Receptors:** These receptors are located *inside* the cell. Steroid hormones diffuse through the cell membrane, bind to their receptor in the cytoplasm, and the entire complex then moves into the nucleus to directly control gene transcription.

# The Receptors That Catch the Message



# Signal Transduction by G Protein-Coupled Receptors



# The Structure of G Protein-Coupled Receptors (GPCRs)



**G protein-coupled receptors (GPCRs)** have a distinct and predictable structure:

- Their amino-terminus is outside the cell.
- They feature **seven alpha-helices** that span the cell membrane.
- Their carboxyl-terminus is inside the cell.

The loops connecting the helices on the outside of the cell form a pocket that binds to the **ligand** (the extracellular messenger). The loops on the cytoplasmic side are structured to provide binding sites for intracellular partners, most importantly **G proteins**.

## How GPCRs Are Activated

The activation process is a physical chain reaction:

1. **Inactive State:** The receptor is held in an inactive conformation by noncovalent bonds between its helices.
2. **Ligand Binding:** A ligand binds to the receptor's outer pocket, disrupting these bonds.
3. **Conformational Change:** This causes the transmembrane helices to shift and rotate. This movement alters the shape of the internal, cytoplasmic loops.
4. **G Protein Binding:** The newly shaped cytoplasmic loops now have a high affinity for a G protein, causing the receptor and G protein to form a complex.

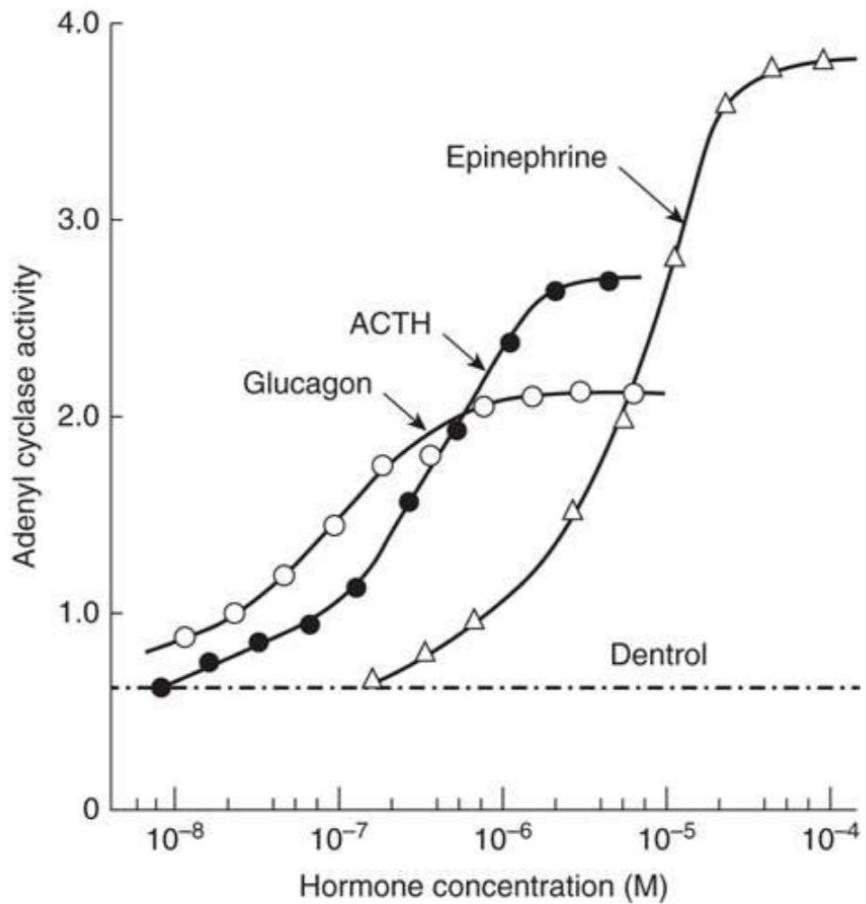
# Passing the Signal to the G Protein

Once the receptor-G protein complex is formed, the signal is officially transferred:

- The interaction with the activated receptor causes the G protein to change shape, making it release its bound **GDP**.
- The G protein then picks up a molecule of **GTP**, which switches the G protein into its "on" state, ready to carry the signal further into the cell.

Importantly, a single activated receptor can activate many G protein molecules, providing a powerful **amplification** of the original signal.

# The effects of Ligands on GPCR Signaling



The effects of varying concentrations of ACTH, epinephrine, and glucagon on adenylyl cyclase activity of fat cell ghosts. The dashed line represents the activity in the absence of added hormones.

Source: From L. Birnbaumer and M. Rodbell, J. Biol. Chem. 244:3478, 1969. Copyright 1969. Reproduced with permission of American Society for Biochemistry and Molecular Biology.

# The effects of Ligands on GPCR Signaling

TABLE 1

**Effects of Combinations of Hormones, at Supramaximal Concentrations,  
on Adenylyl Cyclase Activity in Fat Cell Ghosts\***

**Source:** From L. Birnbaumer and M. Rodbell, *J. Biol. Chem.* 244:3479, 1969. Copyright 1969.  
Reproduced with permission of American Society for Biochemistry and Molecular Biology.

Additions	Change in adenylyl cyclase activity due to hormones			
	Experiment 1 (37°C)		Experiment 2 (30°C)	
	Found	Calculated if additive	Found	Calculated if additive
ACTH	0.57 ± 0.02		1.19 ± 0.07	
Epinephrine	1.00 ± 0.06		1.79 ± 0.11	
Glucagon	0.32 ± 0.01		0.57 ± 0.04	
ACTH + epinephrine	0.80 ± 0.04	1.57	2.04 ± 0.12	2.98
Epinephrine + glucagon	0.99 ± 0.05	1.26	2.13 ± 0.10	2.36
ACTH + glucagon	0.64 ± 0.02	0.89	1.33 ± 0.06	1.76
ACTH + epinephrine + glucagon	0.85 ± 0.04	1.88	2.30 ± 0.10	3.55

\*Adenylyl cyclase activity was measured at either 37° in absence of, or at 30° in presence of, ATP-regenerating system. The following concentrations of the hormones were used, either individually or when combined: ACTH, 400 µg per ml; epinephrine, 400 µg per ml; glucagon, 60 µg per ml. Values are the mean of triplicate determination ± standard deviation.

# Turning Off the Signal: Desensitization



To prevent overstimulation, a cell must stop a signal even when the messenger molecule is still present. This process, called **desensitization**, involves two key steps to block the receptor from activating more G proteins:

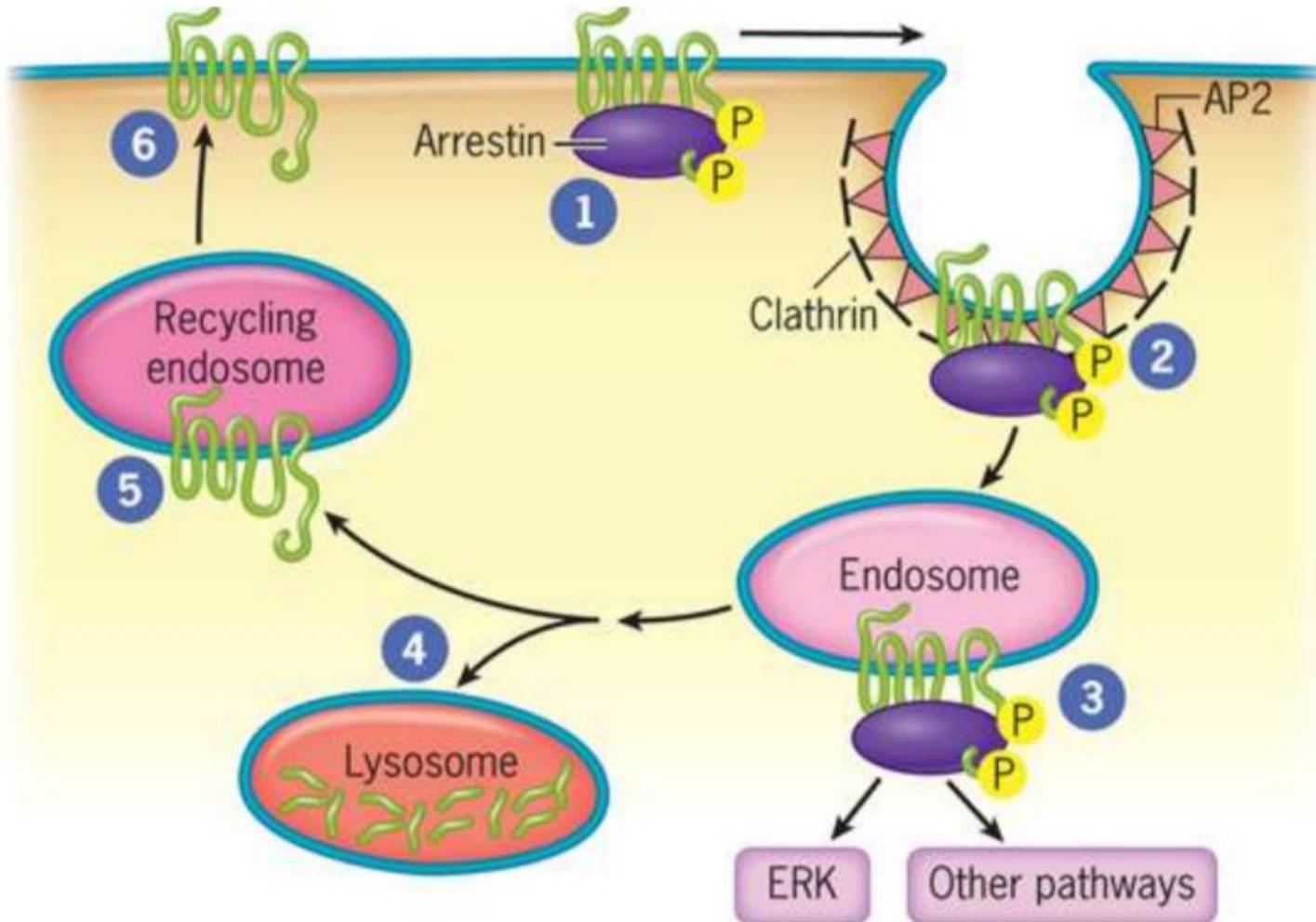
1. **Phosphorylation:** A specific enzyme, **G protein-coupled receptor kinase (GRK)**, recognizes the activated GPCR and adds phosphate groups to its internal domain.
2. **Arrestin Binding:** A protein named **arrestin** binds to the phosphorylated GPCR. This physically blocks the receptor from interacting with and activating any more G proteins, shutting down the initial signal.

## The Fate of the Receptor After Shutdown

Binding **arrestin** does more than just stop the signal; it also targets the receptor to be pulled into the cell via **endocytosis**. Once inside an endosome, the receptor has three possible fates:

- 1. **Signal from a New Location:** The internalized receptor-arrestin complex can act as a new signaling platform, activating different pathways (like the MAPK pathway) from inside the endosome.
- 2. **Degradation:** The receptor can be sent to the lysosome and destroyed. This makes the cell lose sensitivity to that specific ligand for a longer period.
- 3. **Recycling (Resensitization):** The receptor can be dephosphorylated (its phosphate groups removed) and returned to the cell surface. This restores the cell's sensitivity to the ligand, a process called **resensitization**.

# The Fate of the Receptor After Shutdown



# Human Diseases Linked to the G Protein Pathway

TABLE 1

## Human Diseases Linked to the G Protein Pathway

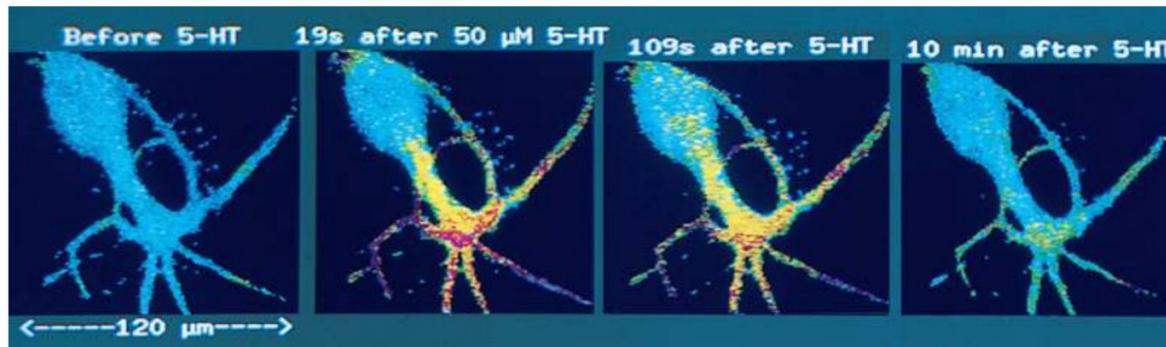
**Source:** Adapted from D. E. Clapham, *Nature*, Vol 371, p. 109, 1994, Nature by Nature Publishing Group.

Disease	Defective G protein*
Albright's hereditary osteodystrophy and pseudohypoparathyroidisms	$G_{sa}$
McCune–Albright syndrome	$G_{sa}$
Pituitary, thyroid tumors ( <i>gsp</i> oncogene)	$G_{sa}$
Adrenocortical, ovarian tumors ( <i>gip</i> oncogene)	$G_{ia}$
Combined precocious puberty and pseudohypoparathyroidism	$G_{sa}$

# Human Diseases Linked to the G Protein Pathway

Disease	Defective G protein-coupled receptor
Familial hypocalciuric hypercalcemia	Human analogue of BoPCAR1 receptor
Neonatal severe hyperparathyroidism	Human analogue of BoPCAR1 receptor (homozygous)
Hyperthyroidism (thyroid adenomas)	Thyrotropin receptor
Familial male precocious puberty	Luteinizing hormone receptor
X-linked nephrogenic diabetes insipidus	V2 vasopressin receptor
Retinitis pigmentosa	Rhodopsin receptor
Color blindness, spectral sensitivity variations	Cone opsin receptor
Familial glucocorticoid deficiency and isolated glucocorticoid deficiency	Adrenocorticotrophic hormone (ACTH) receptor
Short stature, obesity	Ghrelin receptor
Early-onset severe obesity	melanocortin-4 receptor (heterozygote)
Decreased fertility	Follicle stimulating hormone receptor

# What Are Second Messengers?



**Second messengers** are small, internal signaling molecules that are rapidly produced or released after a **first messenger** (e.g., a hormone) binds to a receptor on the cell surface.

A classic example is **cyclic AMP (cAMP)**, which can quickly diffuse throughout the cytoplasm to spread the signal.

## The Function: Amplifying and Coordinating a Response

Second messengers are crucial for amplifying and distributing a signal.

- A **first messenger** is specific and binds to only one type of receptor.
- The **second messenger**, in contrast, can stimulate a wide variety of different cellular activities.

This mechanism allows a single external signal to trigger a large-scale, coordinated response inside the cell.

# What Are Second Messengers?

## Other Important Second Messengers

Besides cAMP, other common second messengers include:

- Ca<sup>2+</sup> (calcium ions)
- Phosphoinositides
- Inositol trisphosphate (IP<sub>3</sub>)
- Diacylglycerol (DAG)
- Cyclic GMP (cGMP)
- Nitric Oxide (NO)

# The Discovery of cAMP



In the mid-1950s, a key question in biology was how a hormone binding to the *outside* of a cell could trigger changes to enzymes *inside* the cell. **Earl Sutherland** and his colleagues set out to solve this puzzle.

## The Crucial Experiment

Sutherland's team worked with a broken-cell preparation from the liver, which they separated into two components:

1. **A particulate fraction:** Containing the solid cell membranes.
2. **A soluble supernatant:** The liquid cytoplasm, containing the enzyme **glycogen phosphorylase**.

They observed that for a hormone like epinephrine to activate the enzyme, both fractions were required. They figured out the process happened in two distinct steps:

- First, incubating the hormone with the **membrane fraction** caused the membranes to release a specific, heat-stable substance.
- Second, when this substance was collected and added to the **supernatant**, it directly activated the glycogen phosphorylase enzyme.

# A New Role for Phospholipids



For a long time, **phospholipids** were seen simply as structural molecules that form the basic fabric of cell membranes. However, we now recognize that they are also crucial **precursors** for a variety of **second messengers**.

## How Phospholipids Become Messengers

When a cell receives an extracellular signal, it can activate specific enzymes that modify these membrane phospholipids to generate second messengers. These enzymes include:

- **Phospholipases** (which split lipids)
- **Phospholipid kinases** (which add phosphate groups to lipids)
- **Phospholipid phosphatases** (which remove phosphate groups from lipids)

The products generated by these enzymes act as internal signals to carry the message forward.

## Focus on Phosphatidylinositol

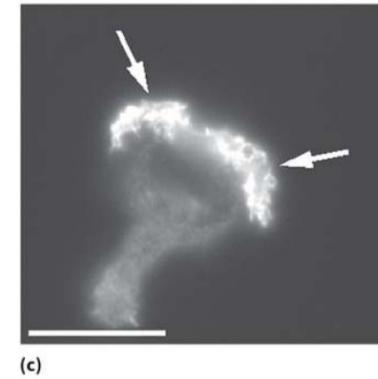
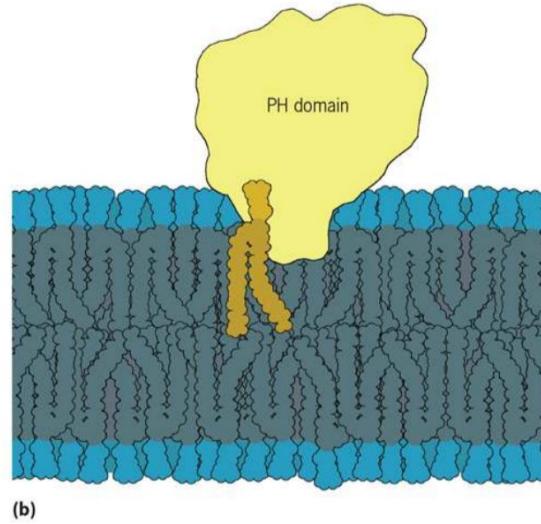
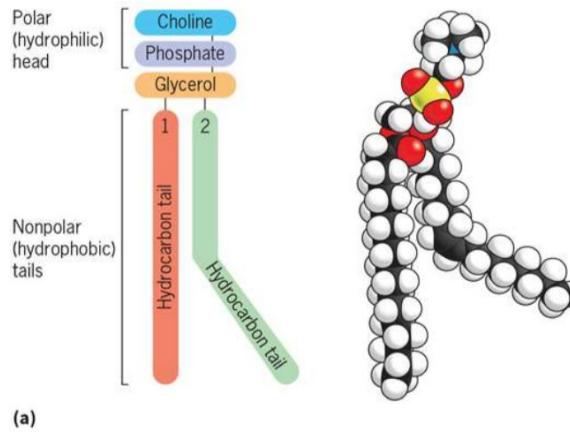
This section of the text focuses specifically on the best-studied lipid second messengers, which are all derived from a particular membrane phospholipid called **phosphatidylinositol**. The creation of these messengers is a common event following the activation of both **G protein-coupled receptors (GPCRs)** and **receptor protein-tyrosine kinases (RTKs)**.

# A New Role for Phospholipids



## Focus on Phosphatidylinositol

This section of the text focuses specifically on the best-studied lipid second messengers, which are all derived from a particular membrane phospholipid called **phosphatidylinositol**. The creation of these messengers is a common event following the activation of both **G protein-coupled receptors (GPCRs)** and **receptor protein-tyrosine kinases (RTKs)**.



# Regulating Blood Glucose



The body works to keep blood glucose levels within a narrow range to ensure a steady supply of energy for all cells. Excess glucose from the blood is stored as a large polymer called **glycogen**. This process is managed by three key hormones:

- **Glucagon:** Released by the pancreas when blood sugar is **low**. It stimulates the breakdown of glycogen to **raise** blood glucose.
- **Insulin:** Released by the pancreas when blood sugar is **high**. It stimulates the uptake and storage of glucose to **lower** blood glucose.
- **Epinephrine:** The "fight-or-flight" hormone released by the adrenal gland during stress. It **raises** blood glucose to provide a quick energy boost.

## How the Hormones Send Their Signals

These hormones use different types of cell surface receptors to transmit their message:

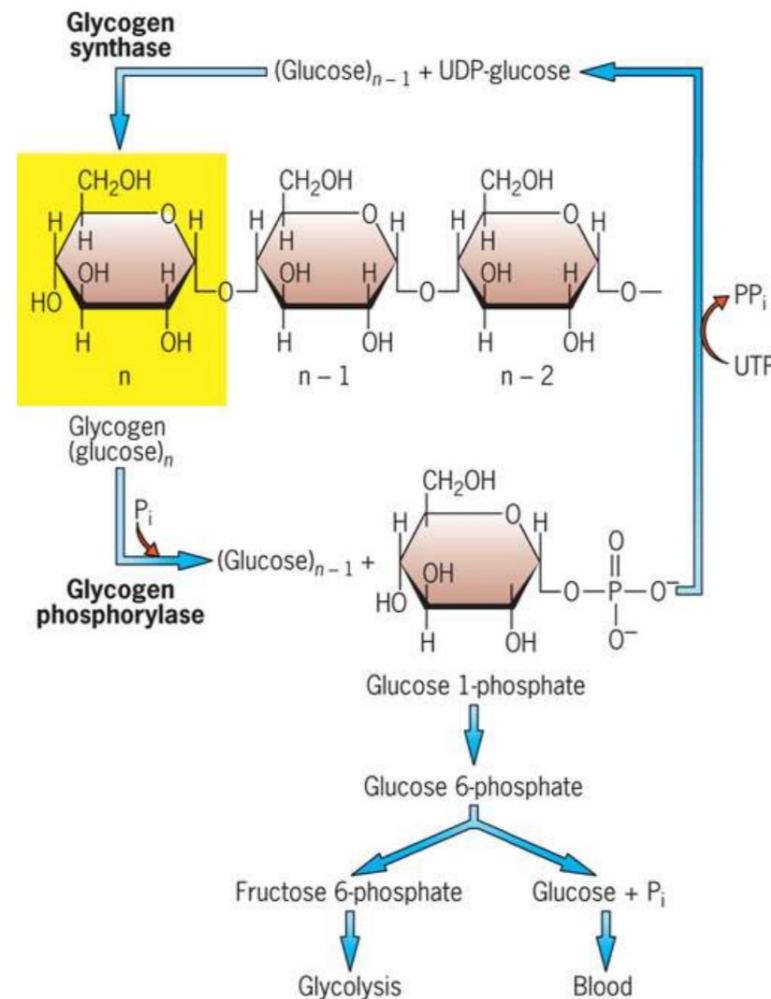
- **Insulin** signals through a **receptor protein-tyrosine kinase (RTK)**.
- Both **glucagon** and **epinephrine** signal through **G protein-coupled receptors (GPCRs)**.

# Convergent Signaling: One Response from Two Signals

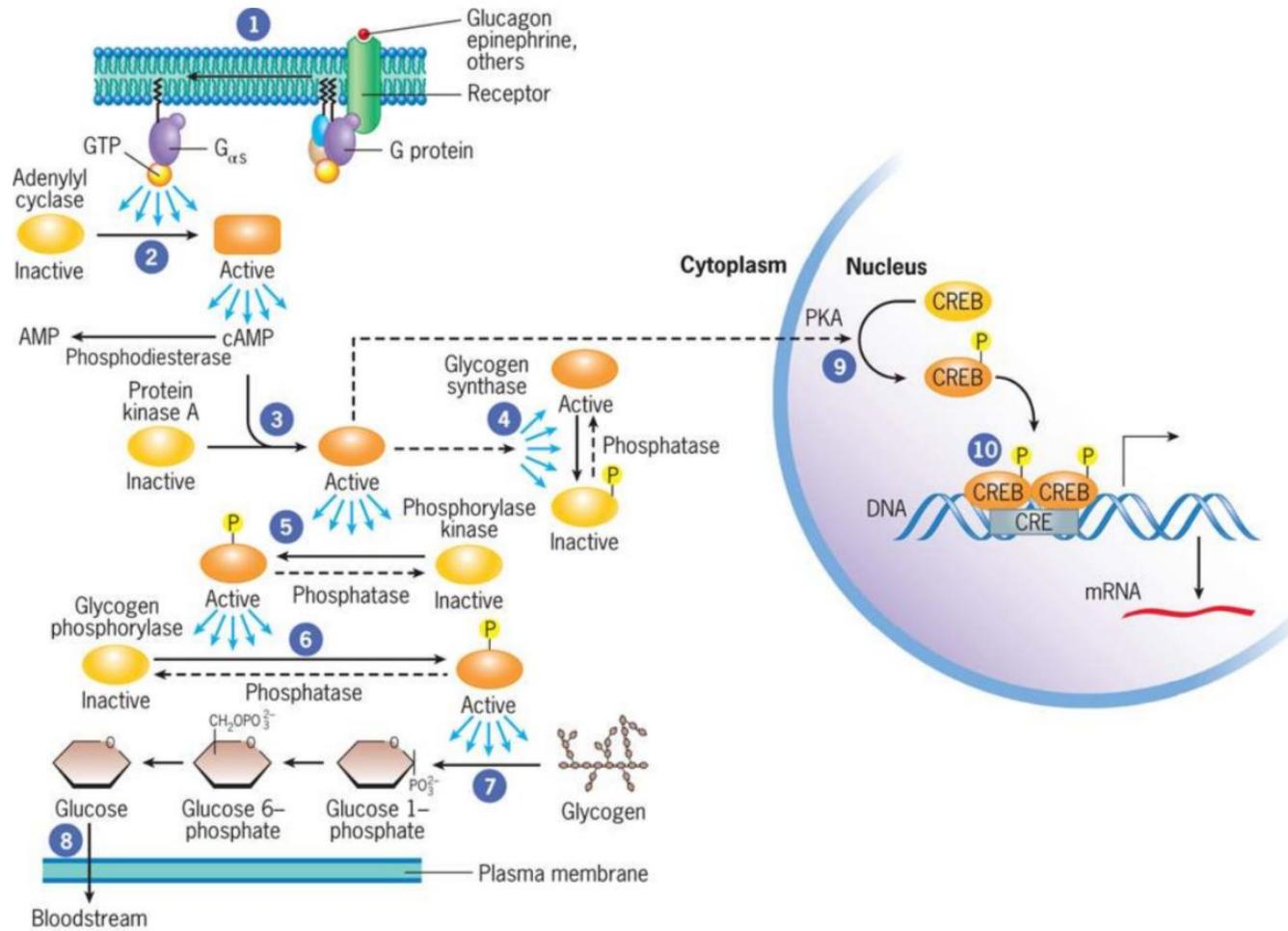
Glucagon and epinephrine are structurally very different molecules, yet they trigger the same response in a target cell (like a liver cell).

- **Shared Outcome:** Both hormones stimulate the breakdown of glycogen and, at the same time, inhibit the enzyme that synthesizes glycogen.
- **Shared Pathway:** They achieve this by binding to their own *separate and specific* GPCRs. However, both of these activated receptors then turn on the *same type* of G protein, which leads to an increase in the second messenger **cAMP**.

This is a classic example of **convergent signaling**, where two different external stimuli use different receptors to activate a common intracellular pathway and produce the same physiological result.



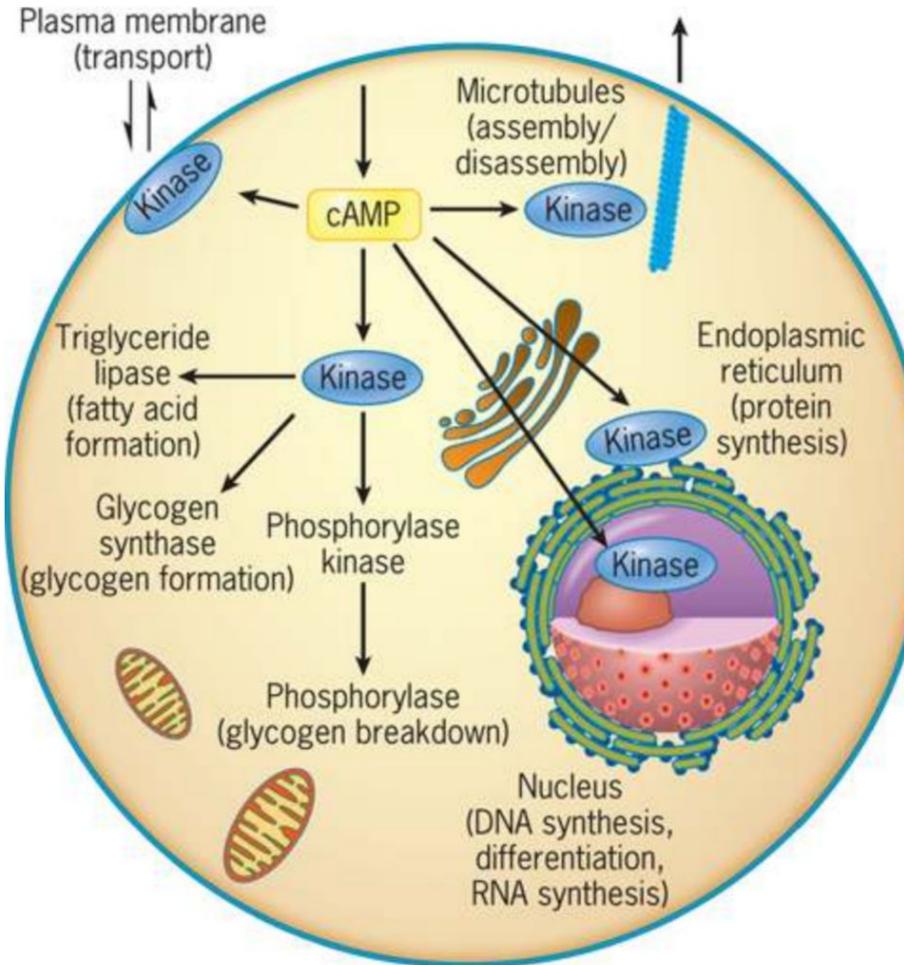
# Convergent Signaling: One Response from Two Signals



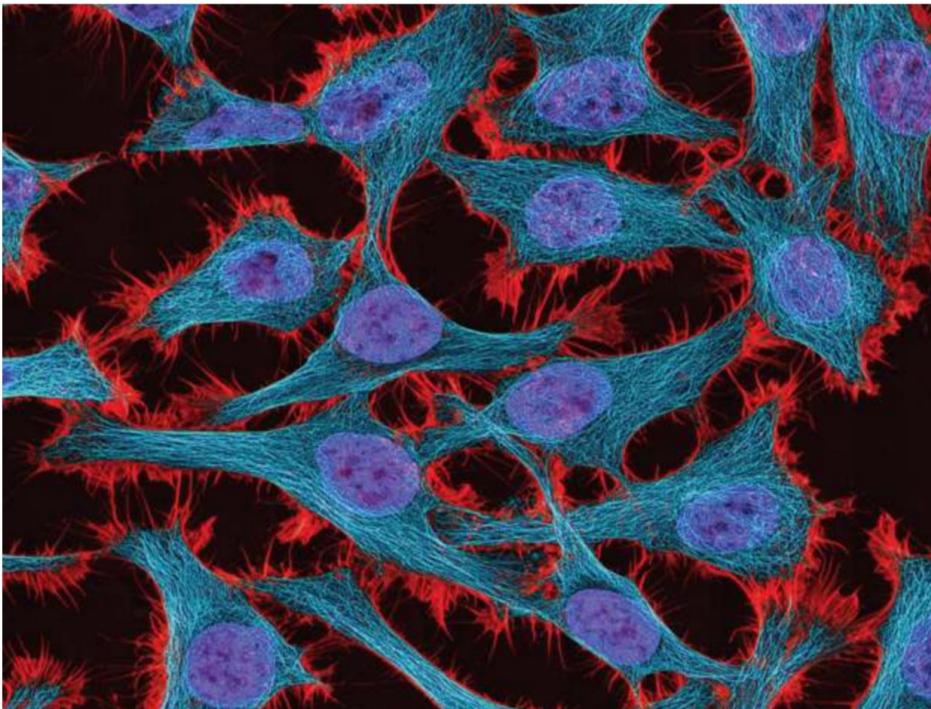
**TABLE 15.3****Examples of Hormone-Induced Responses Mediated by cAMP**

Tissue	Hormone	Response
Liver	Epinephrine and glucagon	Glycogen breakdown, glucose synthesis (gluconeogenesis), inhibition of glycogen synthesis
Skeletal muscle	Epinephrine	Glycogen breakdown, inhibition of glycogen synthesis
Cardiac muscle	Epinephrine	Increased contractility
Adipose	Epinephrine, ACTH, and glucagon	Triacylglycerol catabolism
Kidney	Vasopressin (ADH)	Increased permeability of epithelial cells to water
Thyroid	TSH	Secretion of thyroid hormones
Bone	Parathyroid hormone	Increased calcium resorption
Ovary	LH	Increased secretion of steroid hormones
Adrenal cortex	ACTH	Increased secretion of glucocorticoids

# Master Secondary Messenger: cAMP



# **Molecular Genetics: DNA Repair and Recombination**



## The Origin of HeLa Cells

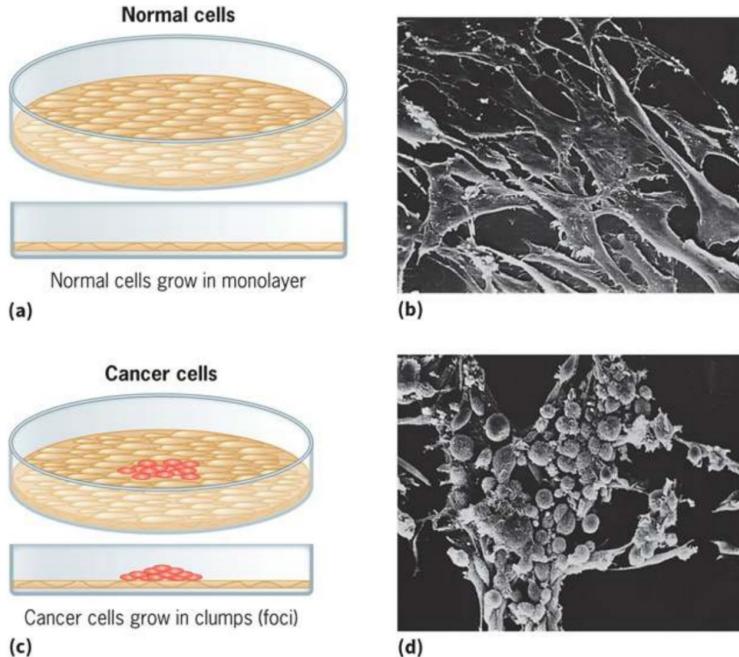


In 1951, a tissue sample was taken from **Henrietta Lacks**, a 31-year-old woman with aggressive cervical cancer. This was done without her or her family's consent, which was a common practice at the time. The cells were given to a researcher, **George Gey**, who had been trying for decades to grow a human cell line in his lab.

Unlike all previous samples, which died quickly, Lacks's cells grew and divided endlessly. Gey had successfully created the first **immortal human cell line**, which he named **HeLa** using the first two letters of her first and last names.

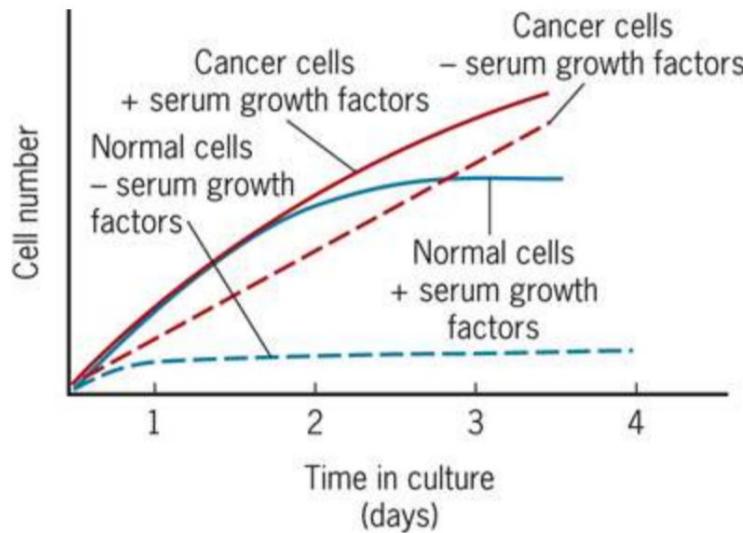
# The Defining Trait: Loss of Growth Control

The most fundamental characteristic of a cancer cell is its **loss of growth control**. This key difference is clearly visible when cancer cells are grown in culture compared to their normal counterparts:



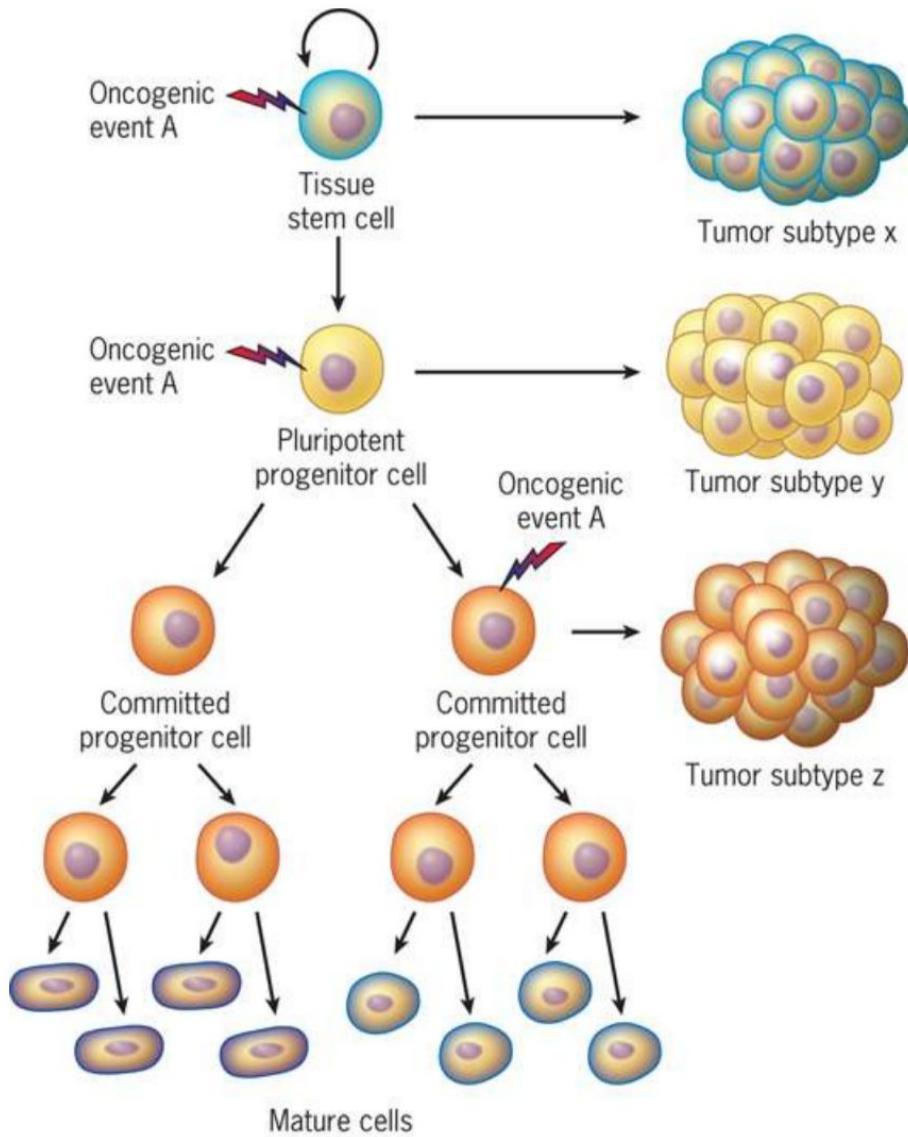
- **Normal Cells:**
  - They grow and divide until they cover the dish surface in a single, flat layer called a **monolayer**.
  - Their growth then stops because they respond to inhibitory signals, such as contact with neighboring cells (contact inhibition).
  - They depend on external **growth factors** (typically supplied by adding serum to the culture medium) to stimulate division.
- **Cancer Cells:**
  - They **ignore inhibitory signals**. When they cover the dish, they don't stop growing; instead, they pile on top of one another to form multilayered clumps known as **foci**.
  - They **don't need the same stimulatory signals**. They can often proliferate even in the absence of the growth factors that normal cells require.

# The effects of serum deprivation on the growth



## Genetic Instability & Evasion of Cell Death

- **Genetic Instability:** Cancer cells are genetically unstable, often having a highly abnormal number and structure of chromosomes—a condition known as **aneuploidy**. This is frequently caused by defects in cell cycle checkpoints.
- **Resistance to Apoptosis:** A normal cell with major genetic damage will typically trigger its own self-destruction via a process called **apoptosis**. A key hallmark of cancer cells is their ability to evade this process, allowing them to survive and multiply despite their genetic chaos.

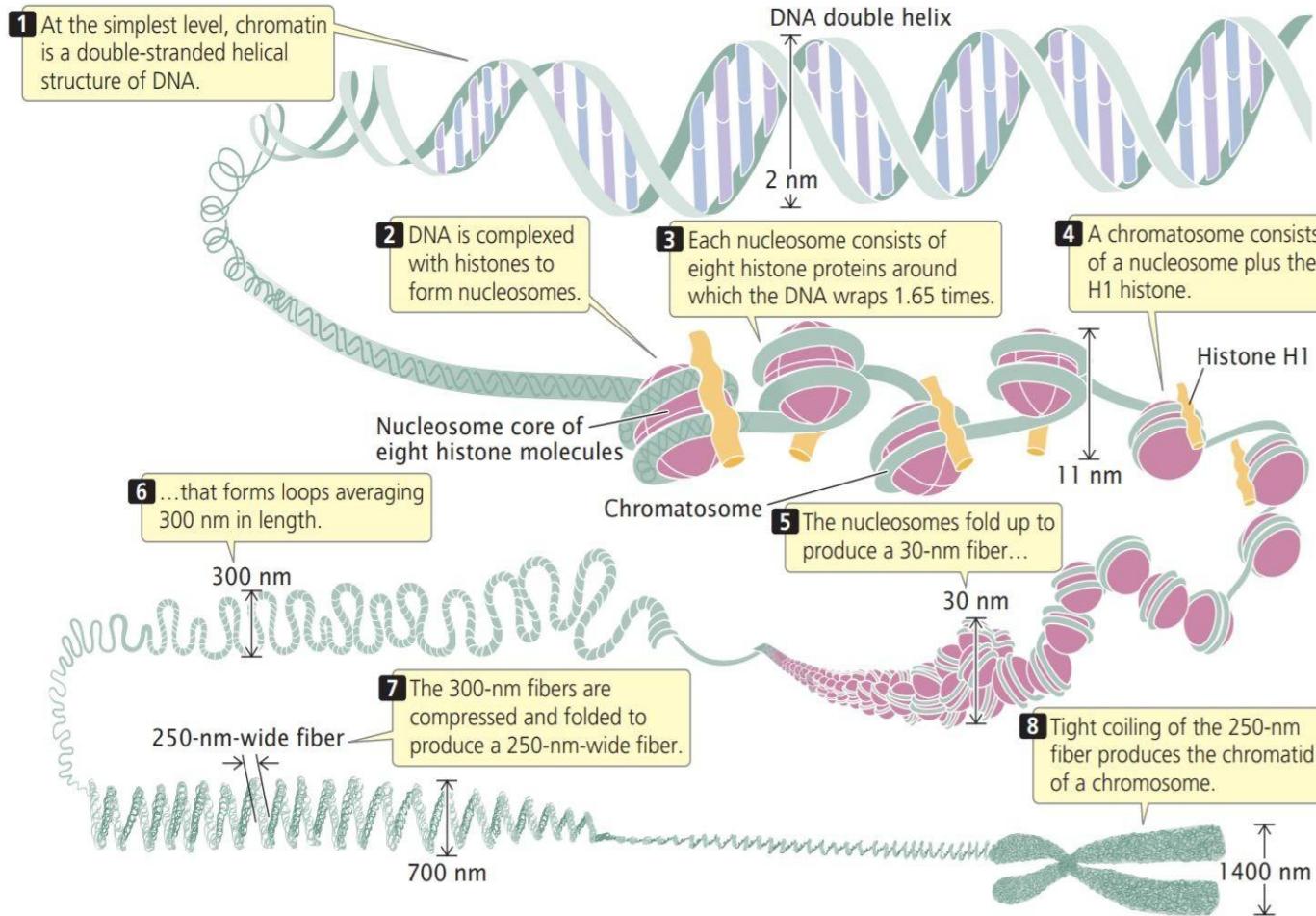


## The Cells of Origin

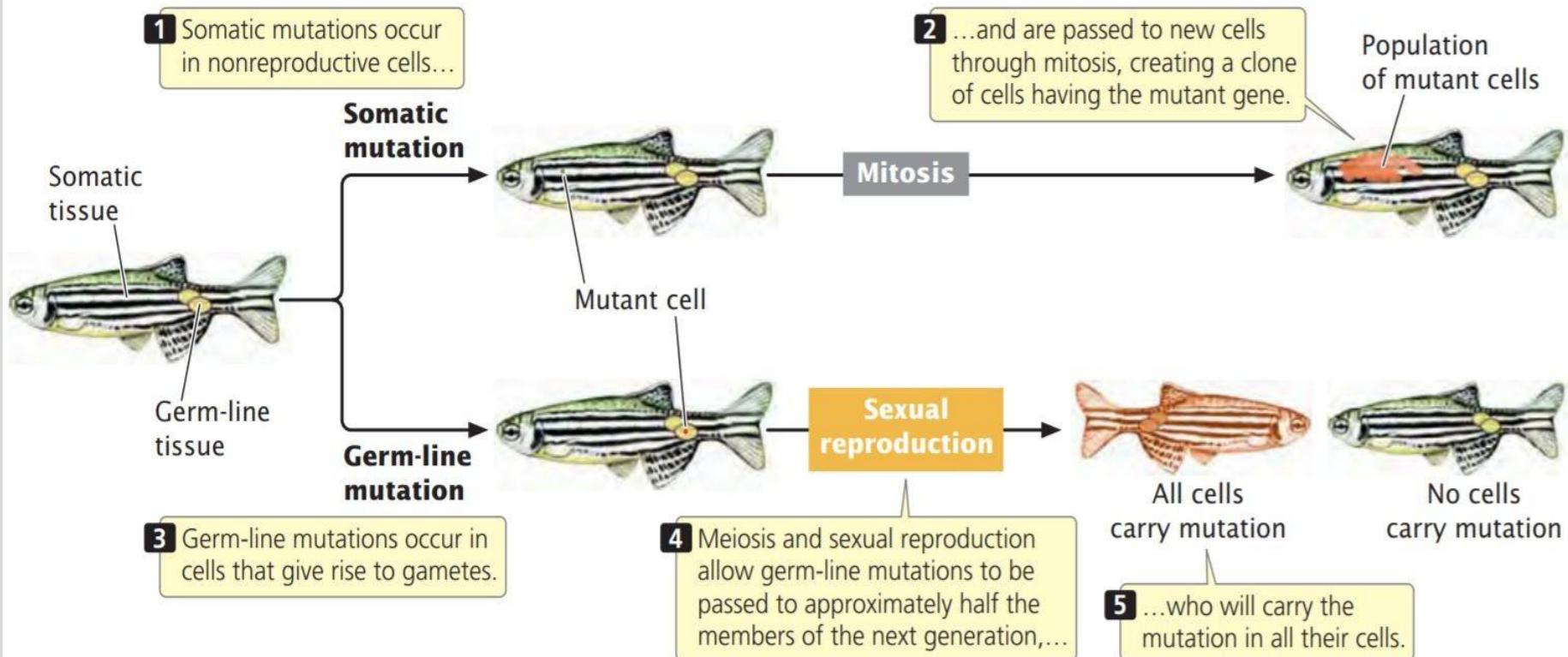
According to the model described, malignant tumors do not arise from fully differentiated cells. Instead, they are thought to originate from either **tissue stem cells** or the more specialized **progenitor cells**.

Crucially, the specific cell of origin can influence the type of cancer that forms. This means a tumor that begins from a **stem cell** may develop into a different kind of cancer than one that originates from a **progenitor cell**.

# **Chromatin has a highly complex structure with several levels of organization.**



# DNA Variation: Somatic or Germline Mutations



18.1 The two basic classes of mutations are somatic mutations and germ-line mutations.

# DNA Variation

- Understand the meaning of DNA sequence and amino acid polymorphisms.
- Recognize the different types of DNA sequence polymorphisms:
  - STR, SNP, CNV
- Know the different classes of DNA mutation:
  - Point mutations (silent, missense, nonsense, frameshift, splicing, regulatory) insertion/deletions, rearrangements
- Understand how to distinguish a disease-causing mutation from a neutral DNA sequence variation

# DNA Variation



- DNA Sequence Variation:
- Human to human: ~0.1% (1:1000 bp)
  - Human genome =  $3 \times 10^9$  bp X 0.1% =~ $3 \times 10^6$  DNA common variants
- Human to chimp: ~1-2%
- More common in “junk” DNA: introns, intergenic regions
- **poly·mor·phism**  
Pronunciation: "päl-i-'mor-“fiz-&m Function: *noun*  
: the quality or state of existing in or assuming different forms: as a (1) : existence of a species in several forms independent of the variations of sex (2) : existence of a gene in several allelic forms (3) : existence of a molecule (as an enzyme) in several forms in a single species

# Polymorphism

- A DNA polymorphism is any difference in the nucleotide sequence between individuals. These differences can be single base pair changes, deletions, insertions, or even changes in the number of copies of a given DNA sequence.
- SNPs (single nucleotide polymorphisms) are the most common type of DNA polymorphism in humans. An example of an SNP would be if a cytosine (C) nucleotide is present at a particular locus in one person's DNA but a thymine (T) nucleotide occurs at the same locus in another person's DNA.
- DNA polymorphisms are such a powerful tool for mapping human diseases that several projects are underway to identify DNA polymorphisms in humans and to make this data publicly available to scientists worldwide. Two groups that are involved in these massive efforts include the SNP consortium (TSC) and the DNA Polymorphism Discovery Resource.

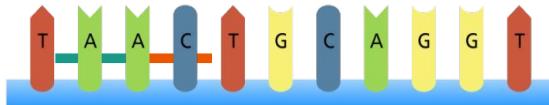
## MUTATION VERSUS POLYMORPHISM

Mutation is a permanent alteration of a nucleotide sequence of a gene	Polymorphism is the presence of more than one allele at a particular locus in a particular population
A physical event	A population attribute
A single base pair change in the nucleotide sequence of a gene is called a point mutation	A single base pair change in the nucleotide sequence is called a single nucleotide polymorphism
Sickle cell anemia, hemophilia, cystic fibrosis, Klinefelter syndrome, and Turner syndrome are results of mutations	Human gender, and ABO blood group are a result of polymorphism
Natural selection selects the mutations that are best suited for the environment	Natural selection does not affect alleles that brings polymorphism

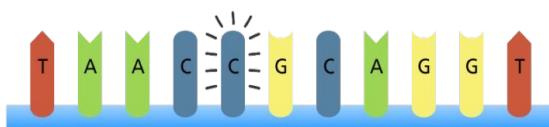
<https://pediaa.com/difference-between-mutation-and-polymorphism/>

# DNA Mutations

Original sequence



Point mutation



Original DNA sequence

GGG AGT GTA GAT CGT

(a)  
Base substitution

GGG AGT GCA GAT CGT  
One codon changed

A base substitution alters a single codon.

(b)  
Base insertion

GGG AGT GTT AGA TCG T

T

An insertion or a deletion alters the reading frame and may change many codons.

(c)  
Base deletion

GGG AGT GAG ATC GT

T

normal

AUG	GCC	TGC	AAA	CGC	TGG
met	ala	cys	lys	arg	trp

silent

AUG	GCT	TGC	AAA	CGC	TGG
met	ala	cys	lys	arg	trp

nonsense

AUG	GCC	TGA	AAA	CGC	TGG
met	ala	---	---	---	---

missense

AUG	GCC	GGC	AAA	CGC	TGG
met	ala	arg	lys	arg	trp

frameshift  
(deletion -1)

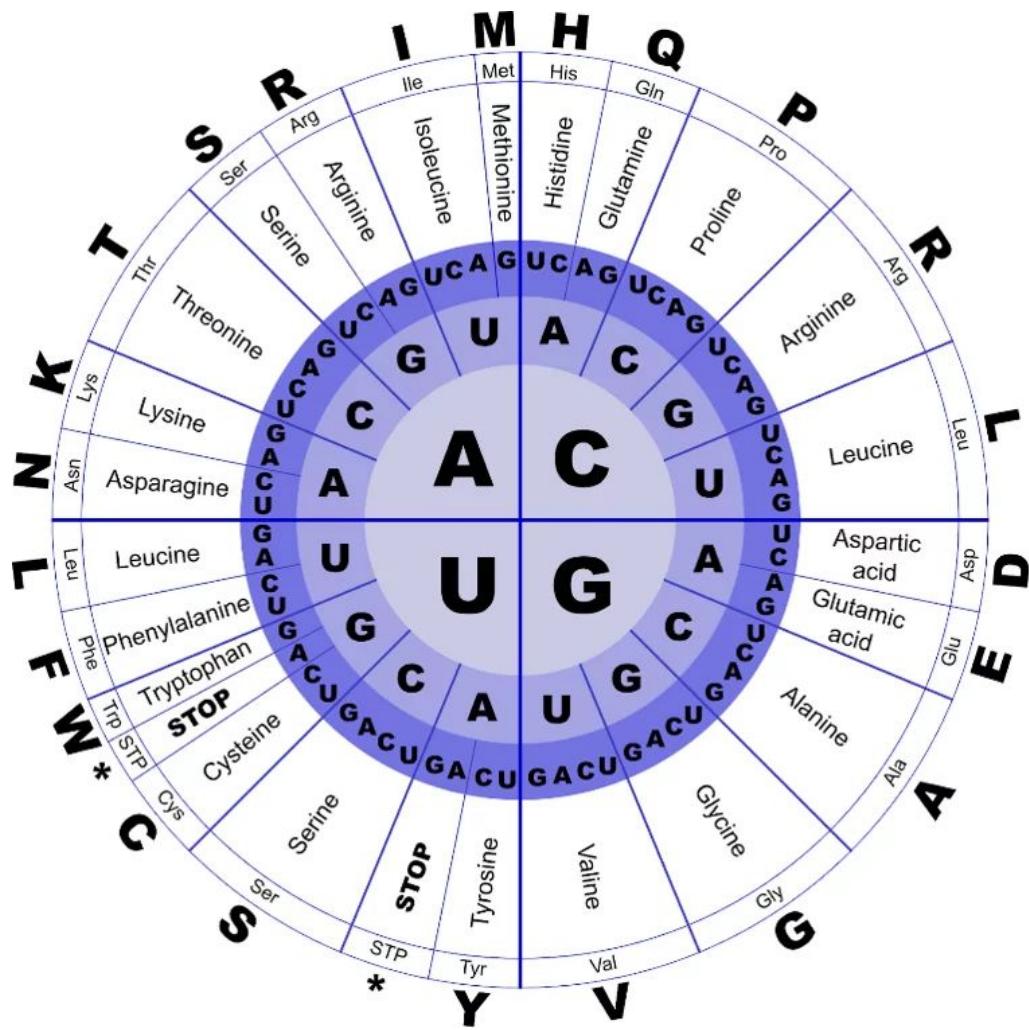
AUG	GC-	TGC	AAA	CGC	TGG
met	ala	glu	asn	ala	

frameshift  
(insertion +1)

AUG	GCC	C	TGC	AAA	CGC	TGG
met	ala	leu	gln	thr	leu	

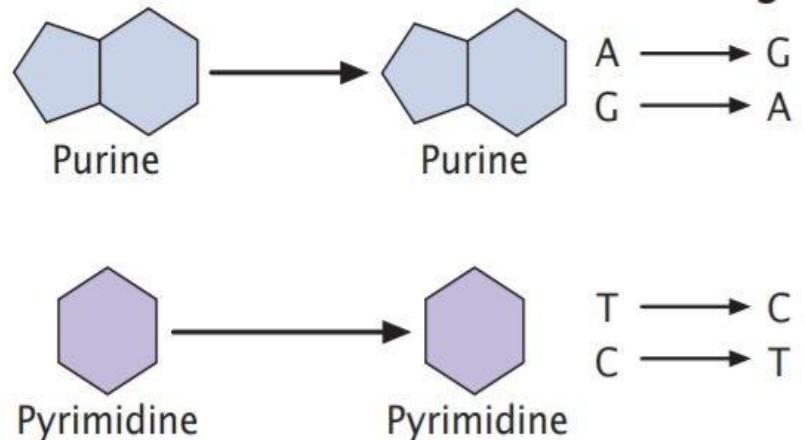
insertion +1,  
deletion -1

AUG	GCC	C	TGC	AAA	-GC	TGG
met	ala	leu	gln	thr		trp

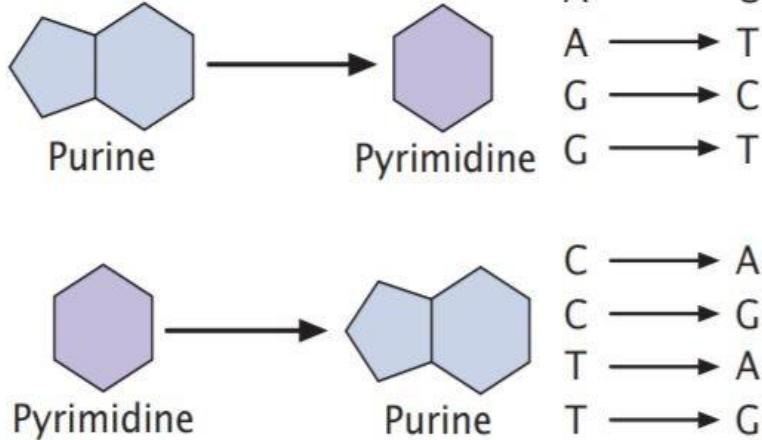


# Transitions vs Transversions

## Transitions



## Transversions



# Insertions and deletions

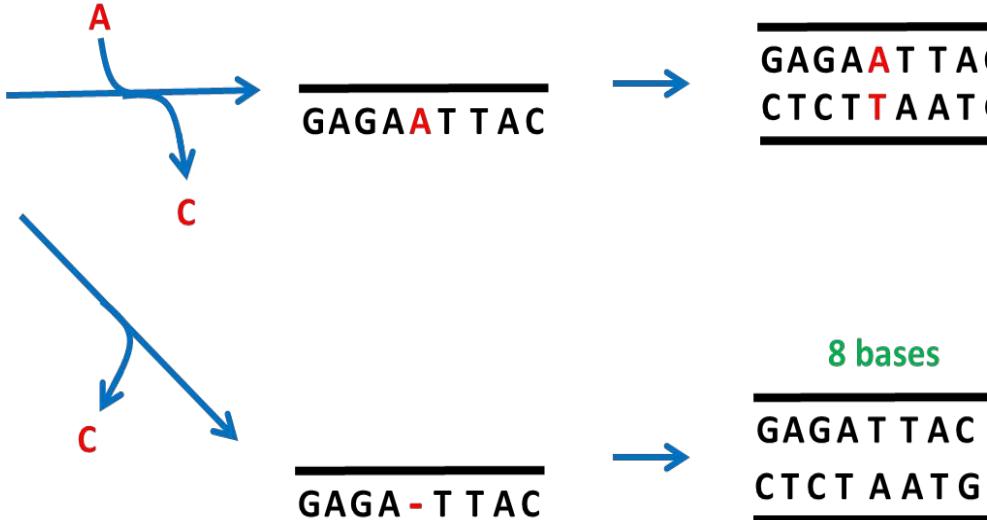
① Insertion



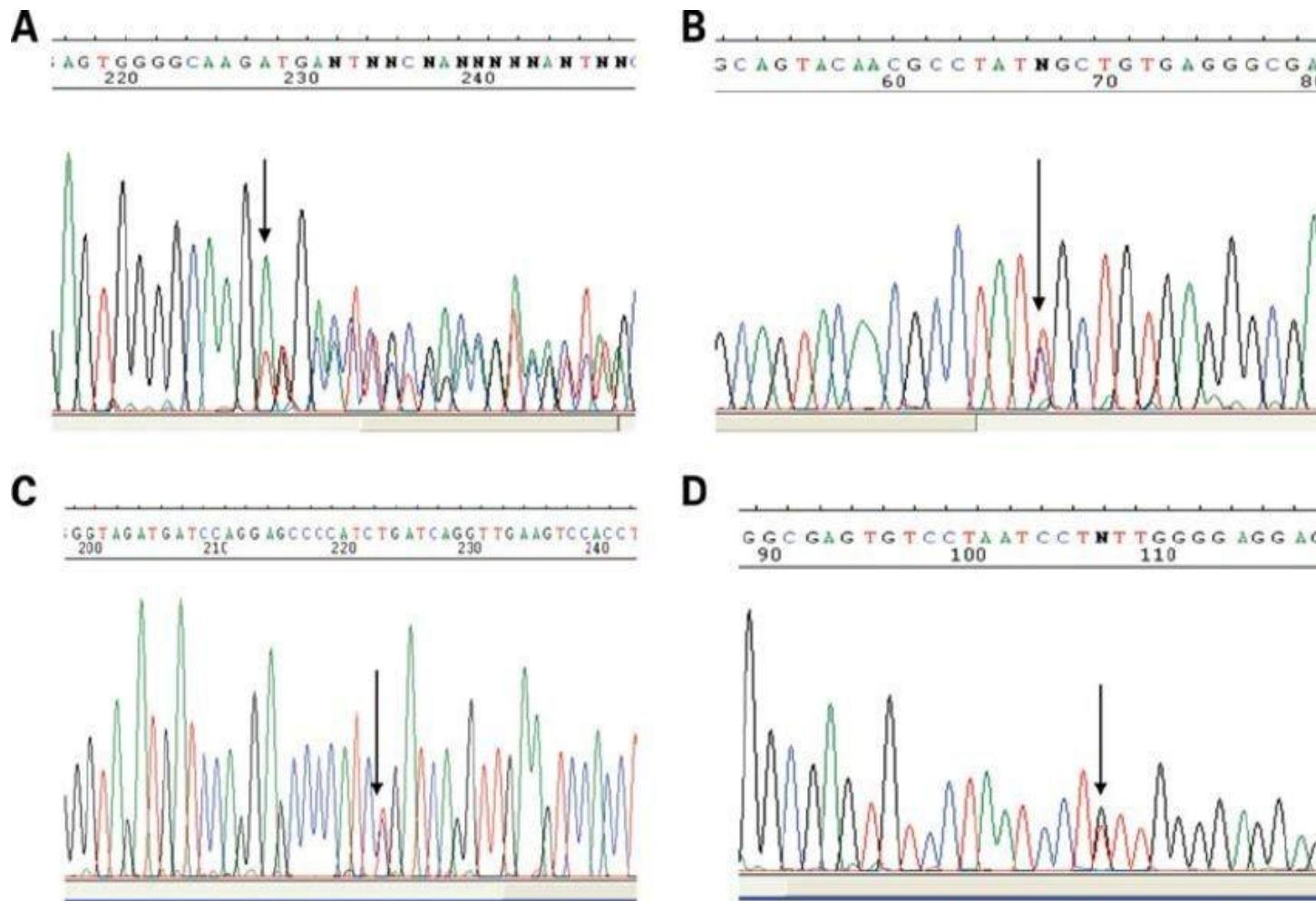
② Substitution



③ Deletion



# Understanding Chromatograms: Insertions and deletions



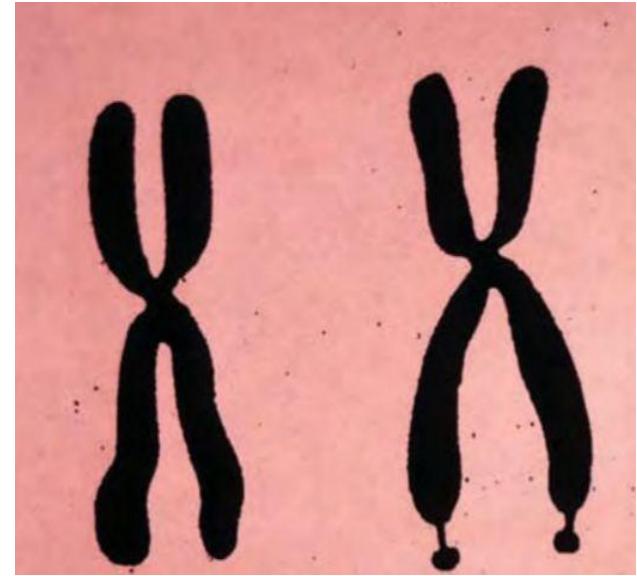
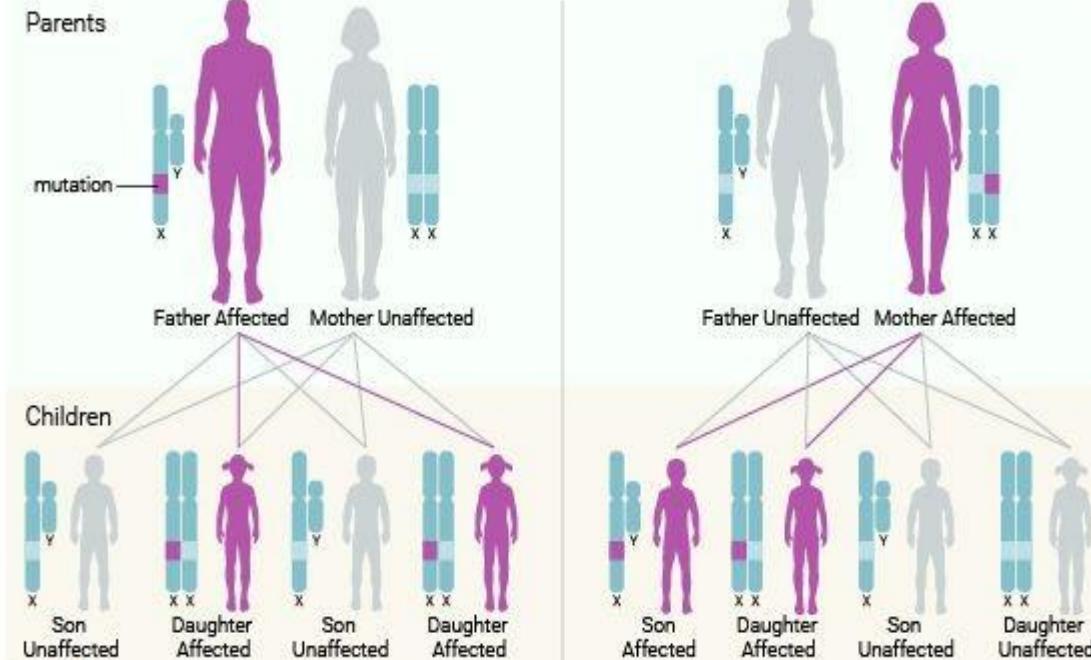
# Insertions and deletions

**Table 18.1** Examples of genetic diseases caused by expanding nucleotide repeats

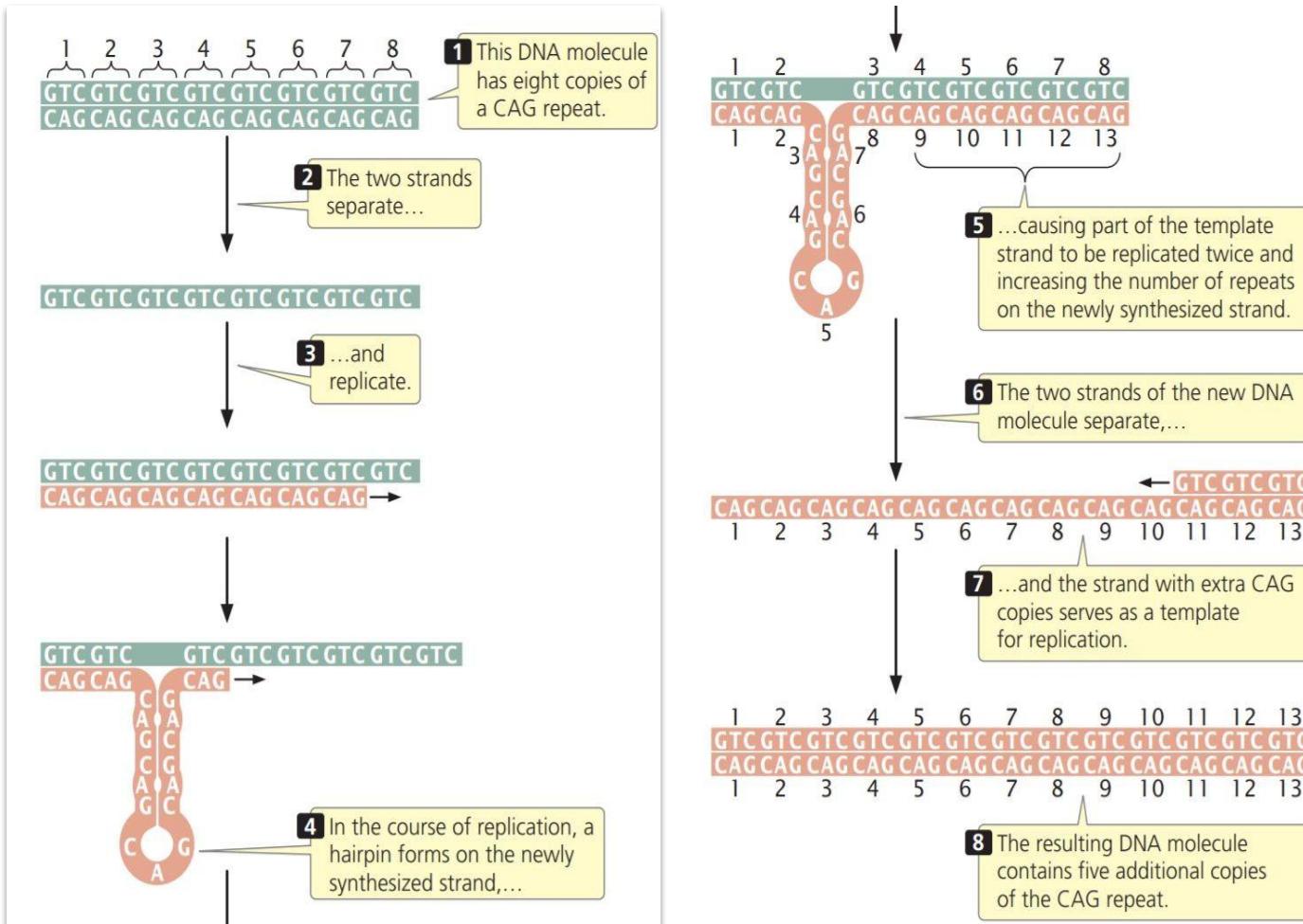
Disease	Repeated Sequence	Number of Copies of Repeat	
		Normal Range	Disease Range
Spinal and bulbar muscular atrophy	CAG	11–33	40–62
Fragile-X syndrome	CGG	6–54	50–1500
Jacobsen syndrome	CGG	11	100–1000
Spinocerebellar ataxia (several types)	CAG	4–44	21–130
Autosomal dominant cerebellar ataxia	CAG	7–19	37–220
Myotonic dystrophy	CTG	5–37	44–3000
Huntington disease	CAG	9–37	37–121
Friedreich ataxia	GAA	6–29	200–900
Dentatorubral-pallidoluysian atrophy	CAG	7–25	49–75
Myoclonus epilepsy of the Unverricht–Lundborg type	CCCCGCCCGCG	2–3	12–13

The fragile-X chromosome is associated with a characteristic constriction (fragile site) on the long arm.

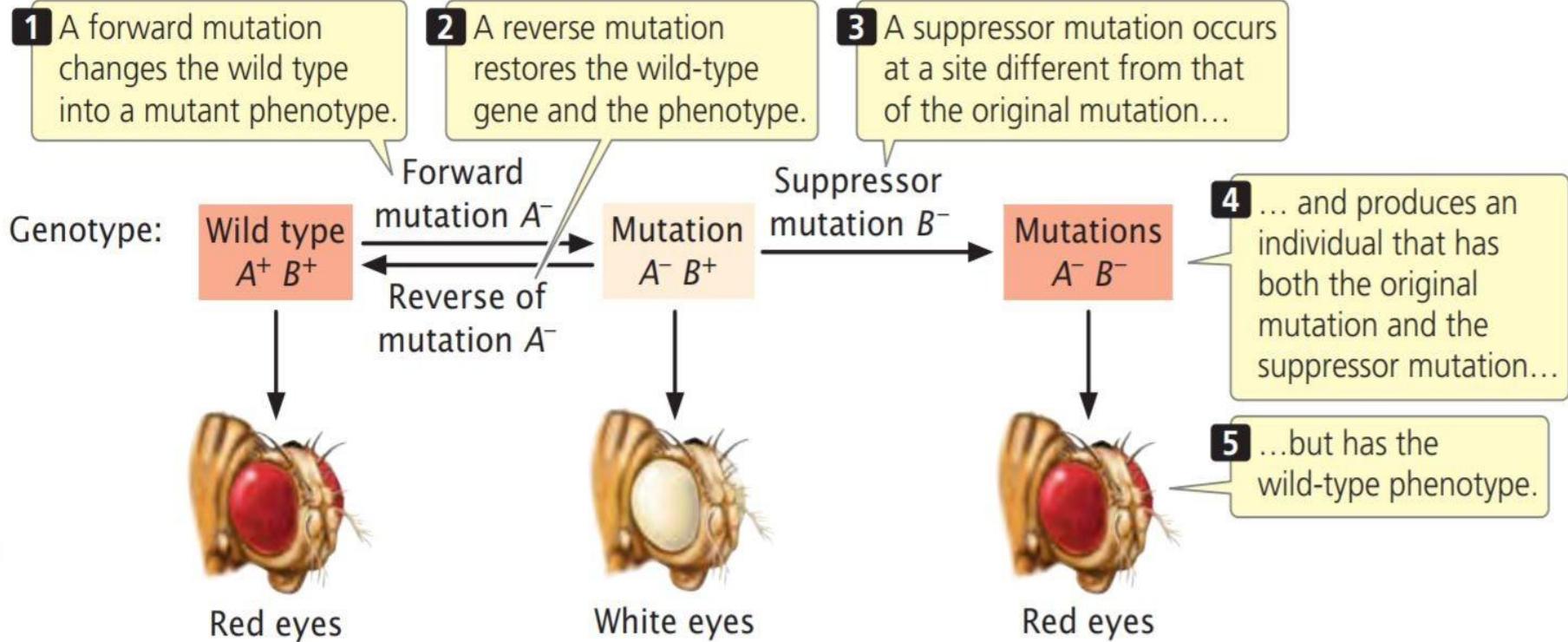
## X-Linked Dominant



# Number of copies of increase in replication



# Forward, reverse, and suppressor mutations.



# Summary: Types of Mutations

**Table 18.2** Characteristics of different types of mutations

Type of Mutation	Definition
Base substitution	Changes the base of a single DNA nucleotide
Transition	Base substitution in which a purine replaces a purine or a pyrimidine replaces a pyrimidine
Transversion	Base substitution in which a purine replaces a pyrimidine or a pyrimidine replaces a purine
Insertion	Addition of one or more nucleotides
Deletion	Deletion of one or more nucleotides
Frameshift mutation	Insertion or deletion that alters the reading frame of a gene
In-frame deletion or insertion	Deletion or insertion of a multiple of three nucleotides that does not alter the reading frame
Expanding nucleotide repeats	Repeated sequence of a set of nucleotides in which the number of copies of the sequence increases
Forward mutation	Changes the wild-type phenotype to a mutant phenotype
Reverse mutation	Changes a mutant phenotype back to the wild-type phenotype
Missense mutation	Changes a sense codon into a different sense codon, resulting in the incorporation of a different amino acid in the protein
Nonsense mutation	Changes a sense codon into a nonsense codon, causing premature termination of translation
Silent mutation	Changes a sense codon into a synonymous codon, leaving unchanged the amino acid sequence of the protein
Neutral mutation	Changes the amino acid sequence of a protein without altering its ability to function
Loss-of-function mutation	Causes a complete or partial loss of function
Gain-of-function mutation	Causes the appearance of a new trait or function or causes the appearance of a trait in inappropriate tissue or at an inappropriate time
Lethal mutation	Causes premature death
Suppressor mutation	Suppresses the effect of an earlier mutation at a different site
Intragenic suppressor mutation	Suppresses the effect of an earlier mutation within the same gene
Intergenic suppressor mutation	Suppresses the effect of an earlier mutation in another gene

**Question:**

A gene encodes a protein with the following amino acid sequence: Met-Arg-Cys-Ile-Lys-Arg

A mutation of a single nucleotide alters the amino acid sequence to:

Met-Asp-Ala-Tyr-Lys-Gly-Glu-Ala-Pro-Val

A second single-nucleotide mutation occurs in the same gene and suppresses the effects of the first mutation (an intragenic suppressor).

With the original mutation and the intragenic suppressor present, the protein has the following amino acid sequence: Met-Asp-Gly-Ile-Lys-Arg

***What is the nature and location of the first mutation and of the intragenic suppressor mutation?***

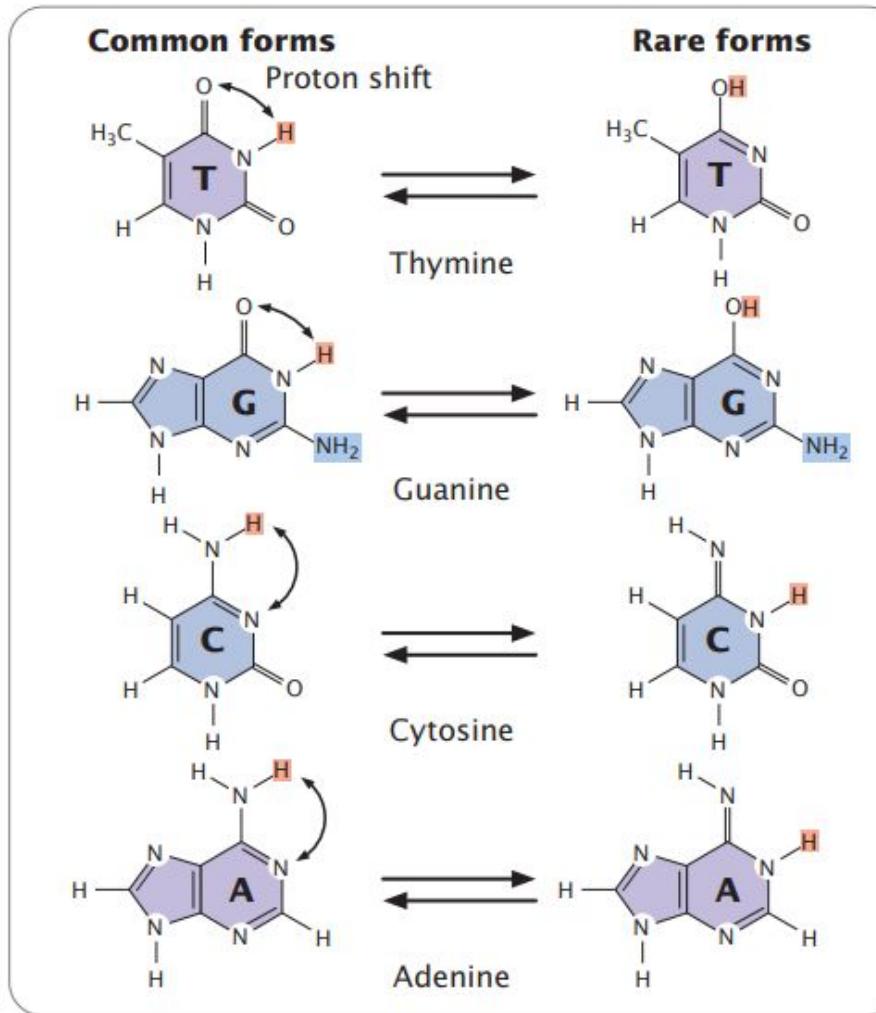
# Mutation Rate

- The frequency with which a wild-type allele at a locus changes into a mutant allele is referred to as the mutation rate and is generally expressed as the number of mutations per biological unit, which may be mutations per cell division, per gamete, or per round of replication.
- For example, achondroplasia is a type of hereditary dwarfism in humans that results from a dominant mutation. On average, about four achondroplasia mutations arise in every 100,000 gametes, and so the mutation rate is 4 /100,000, or 0.00004 mutations per gamete.
- The mutation rate provides information about how often a mutation arises

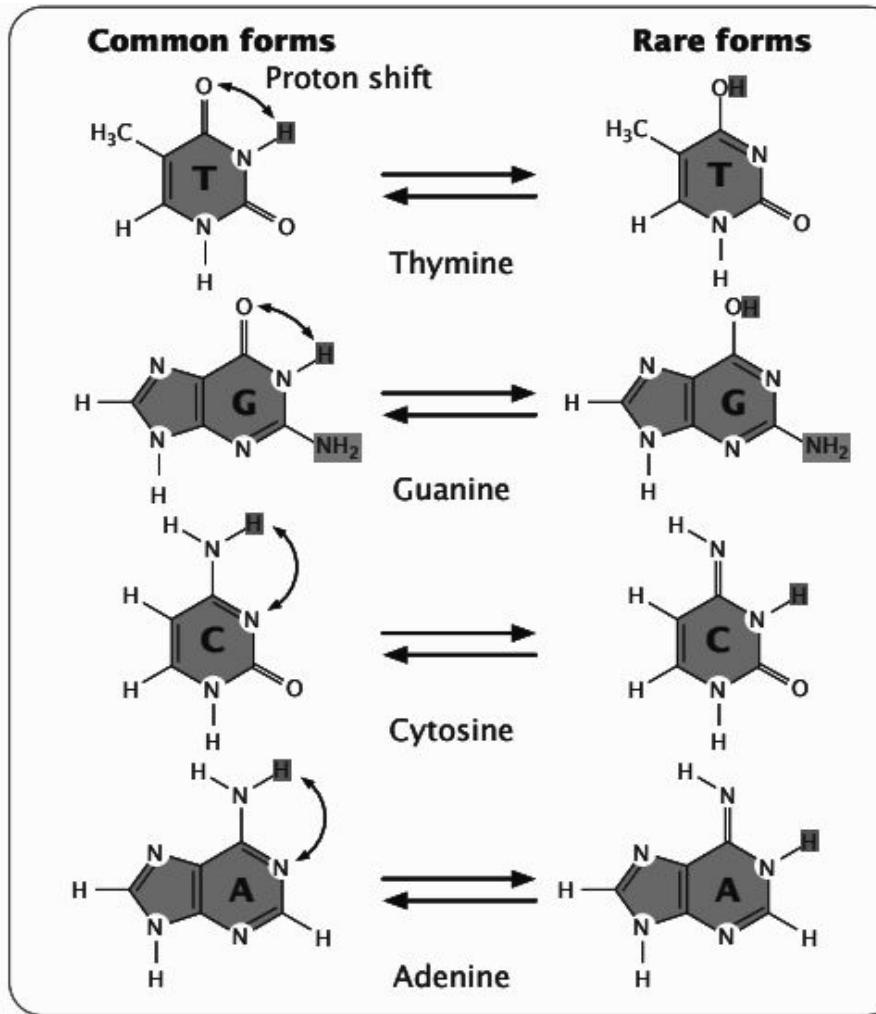
# Factors affecting Mutation Rate

- First, they depend on the frequency with which a change takes place in DNA. A change in the DNA can arise from spontaneous molecular changes in DNA or it can be induced by chemical, biological, or physical agents in the environment.
- The second factor influencing the mutation rate is the probability that, when a change takes place, it will be repaired. Most cells possess a number of mechanisms for repairing altered DNA, and so most alterations are corrected before they are replicated. If these repair systems are effective, mutation rates will be low; if they are faulty, mutation rates will be elevated. Some mutations increase the overall rate of mutation at other genes; these mutations usually occur in genes that encode components of the replication machinery or DNA-repair enzymes.
- The third factor is the probability that a mutation will be recognized and recorded.

Purine and pyrimidine bases exist in different forms called tautomers

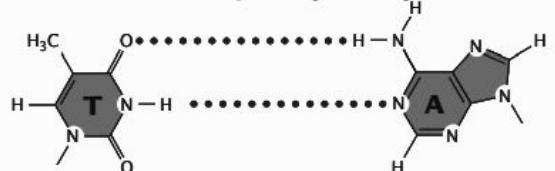


Purine and pyrimidine bases exist in different forms called tautomers

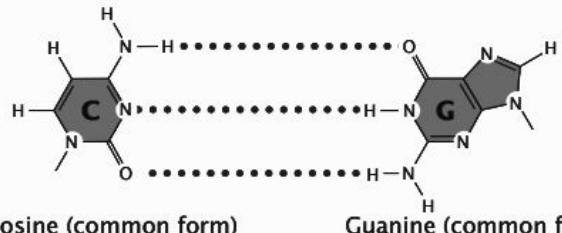


Purine and pyrimidine bases exist in different forms called tautomers

**Standard base-pairing arrangements**

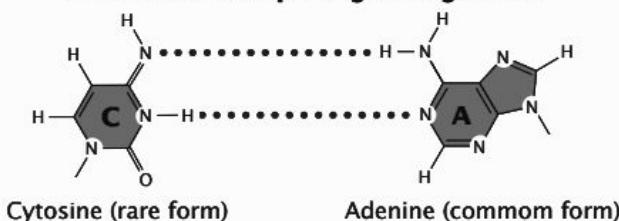


Adenine (common form)

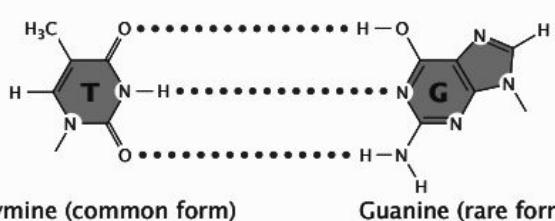


Guanine (common form)

**Anomalous base-pairing arrangements**



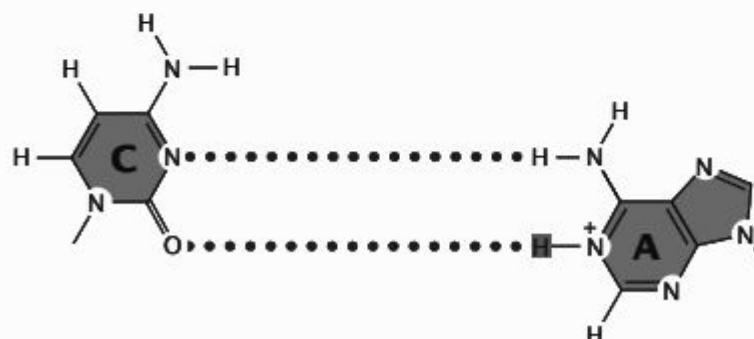
Adenine (common form)



**Non-Watson-and-Crick base pairing**

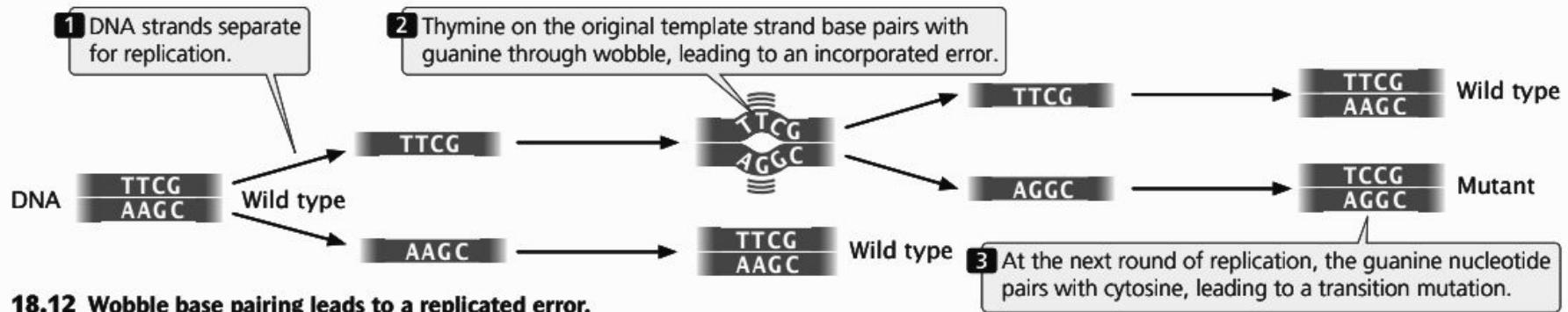


**Thymine-guanine wobble**

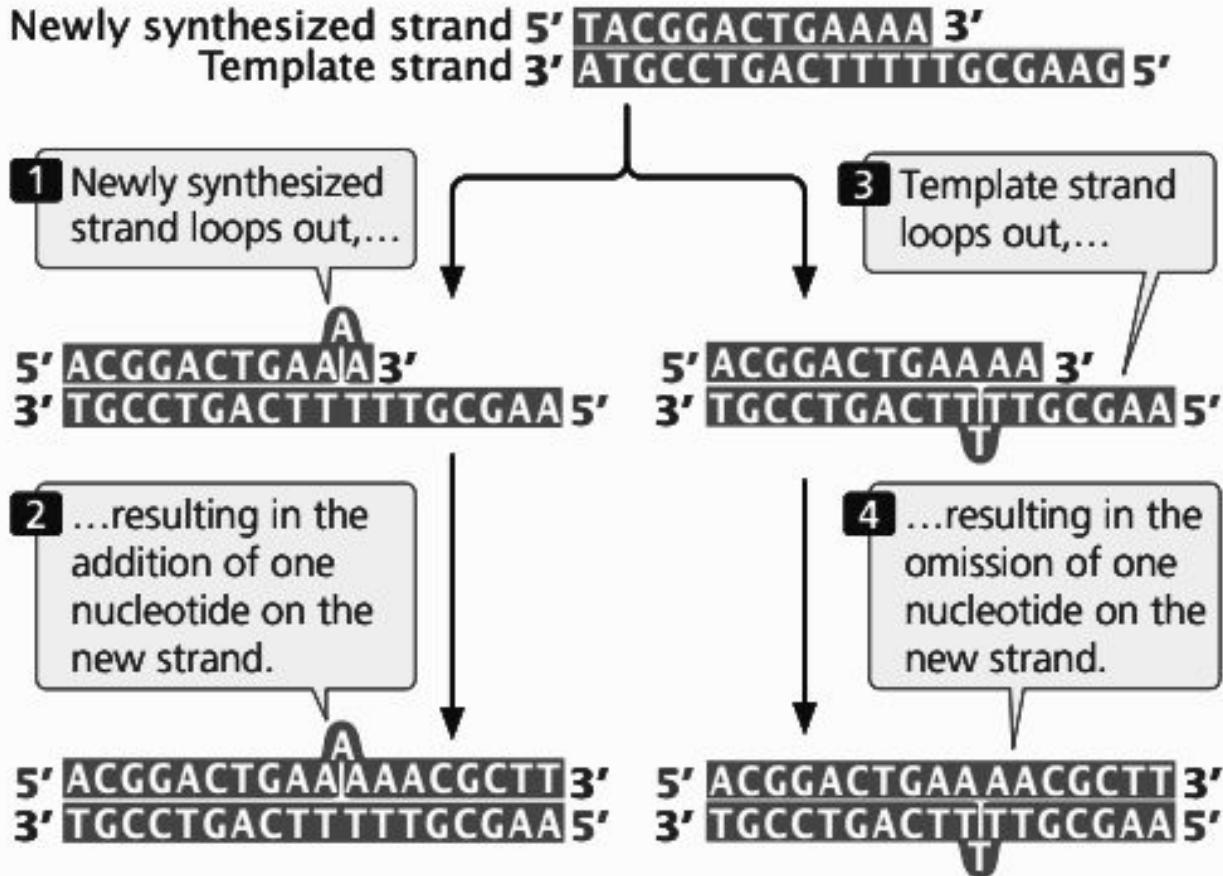


**Cytosine-adenine protonated wobble**

# Wobble base pairing leads to a replicated error.

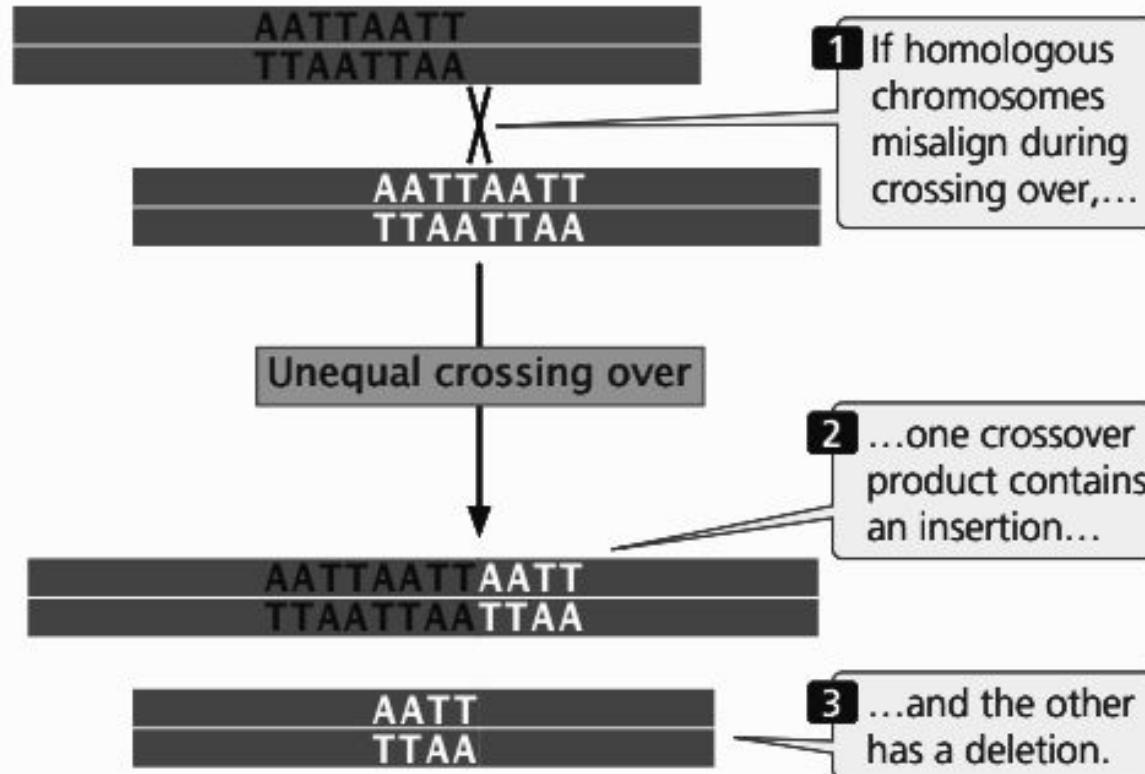


Insertions and deletions may result from strand slippage.



### 18.13 Insertions and deletions may result from strand slippage.

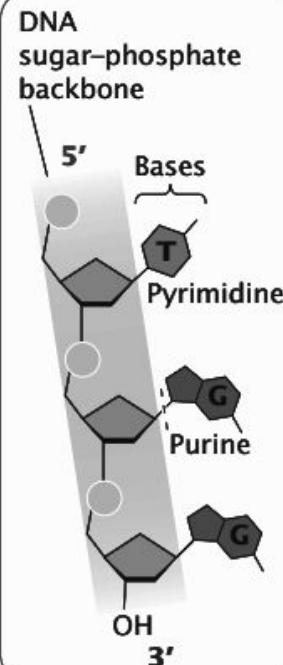
Unequal crossing over produces insertions and deletions



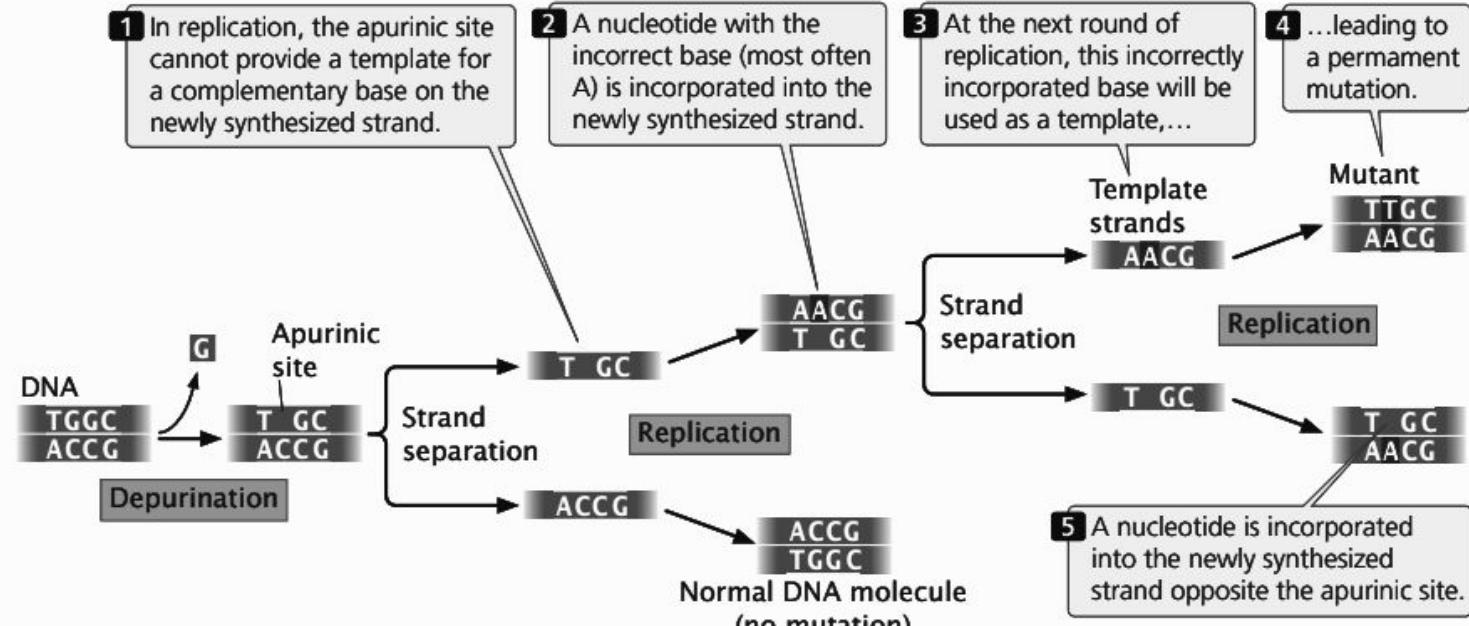
**18.14 Unequal crossing over produces insertions and deletions.**

Depurination (the loss of a purine base from a nucleotide) produces an apurinic site.

(a)

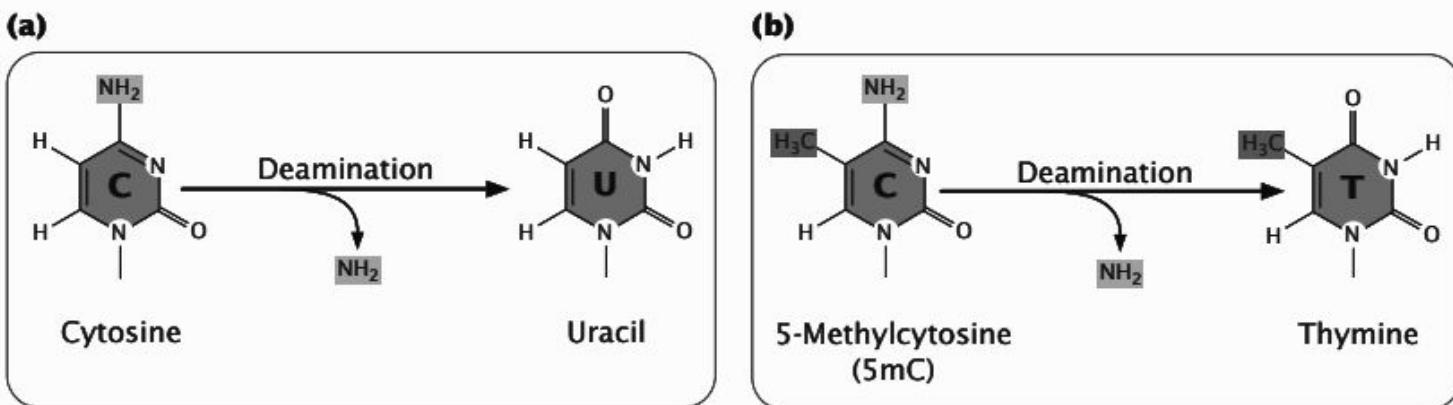


(b)



18.15 Depurination (the loss of a purine base from a nucleotide) produces an apurinic site.

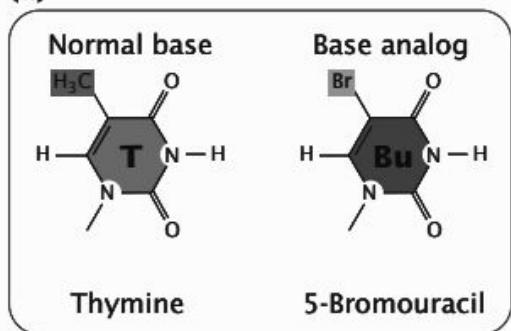
# Chemically Induced Mutations



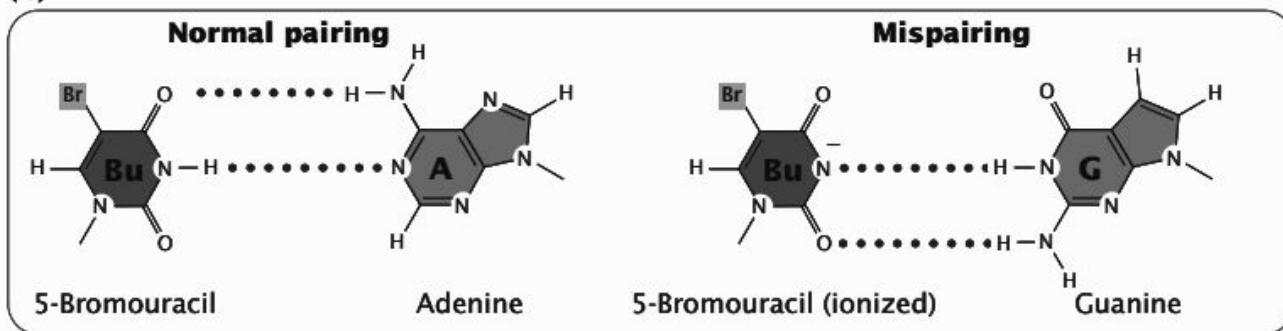
**18.16** Deamination alters DNA bases.

# Chemically Induced Mutations

(a)

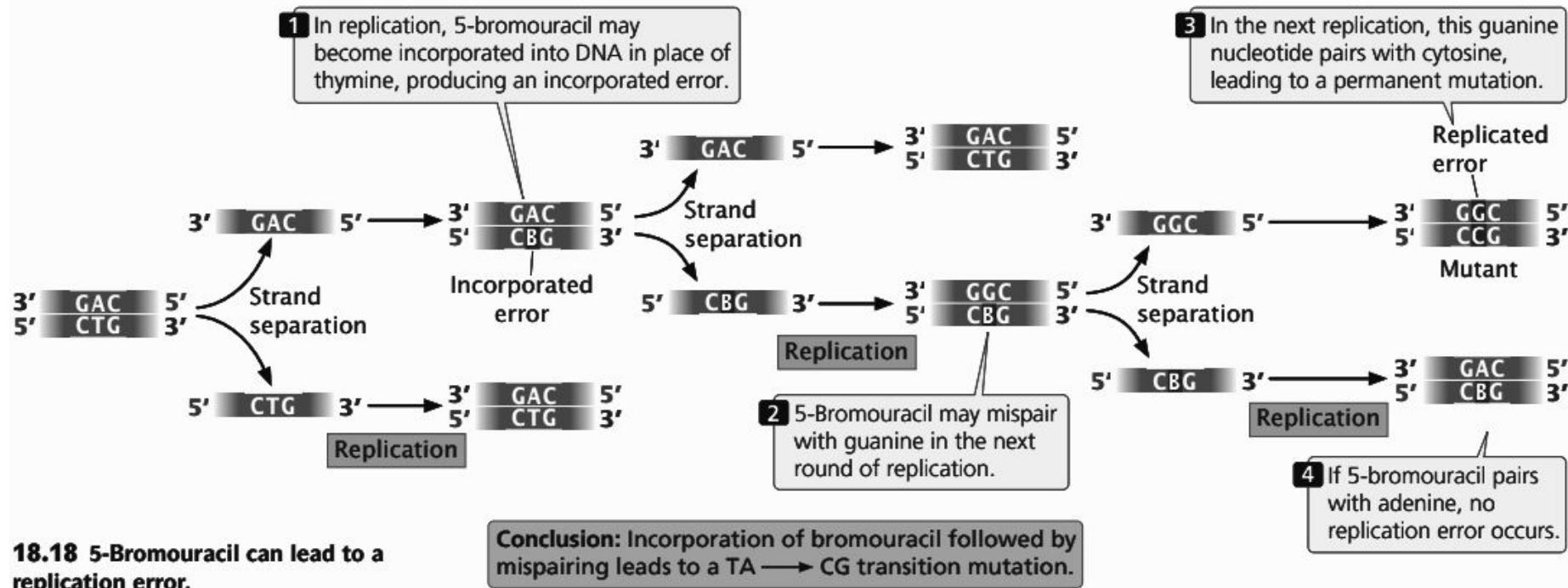


(b)



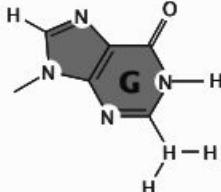
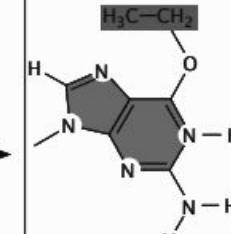
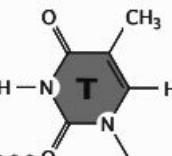
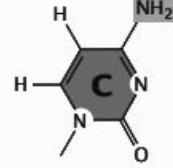
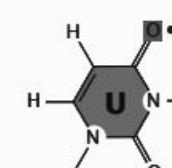
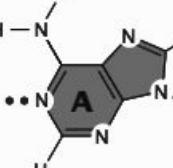
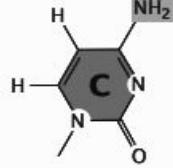
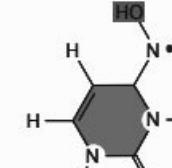
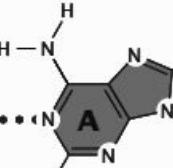
5-Bromouracil (a base analog) resembles thymine, except that it has a bromine atom in place of a methyl group on the 5-carbon atom. B

# 5-Bromouracil can lead to a replication error.

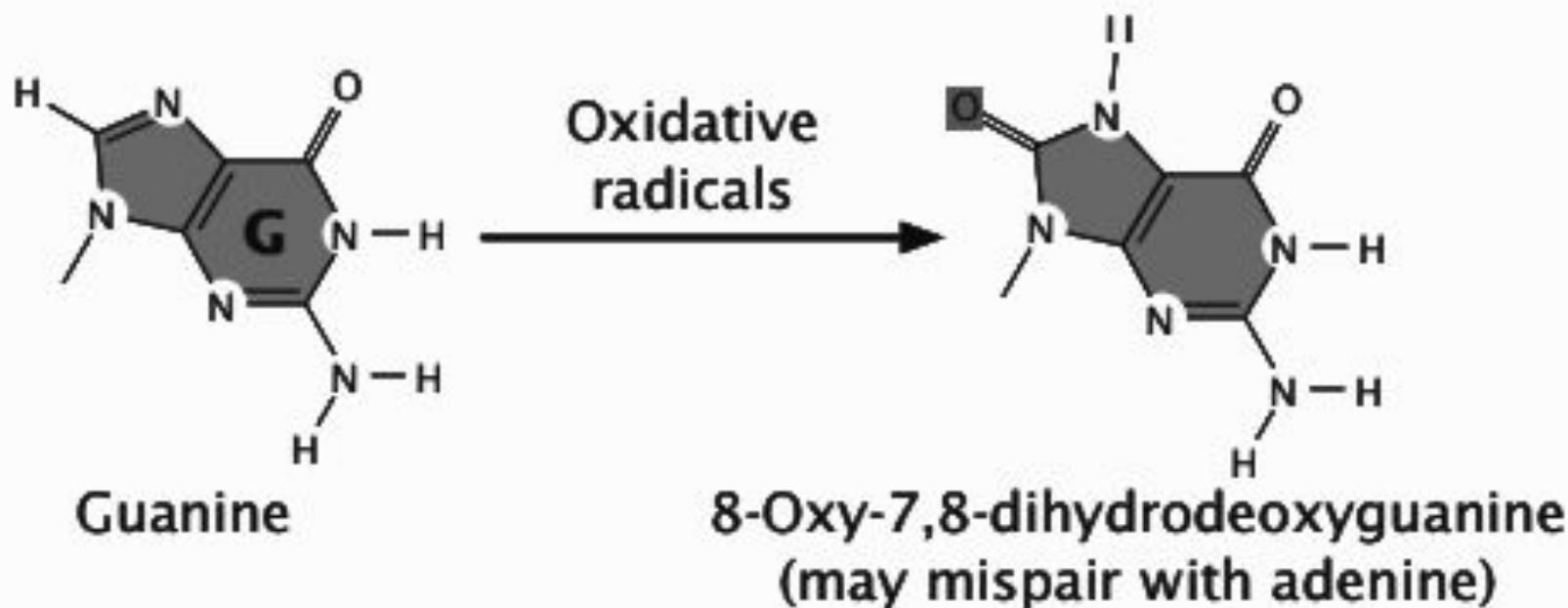


5-Bromouracil (a base analog) resembles thymine, except that it has a bromine atom in place of a methyl group on the 5-carbon atom. B

# Chemicals may alter DNA bases.

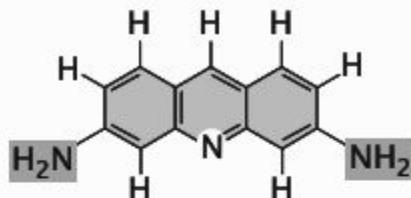
	<b>Original base</b>	<b>Mutagen</b>	<b>Modified base</b>	<b>Pairing partner</b>	<b>Type of mutation</b>
(a)	 Guanine	EMS Alkylation	 <i>O</i> <sup>6</sup> -Ethylguanine	 Thymine	CG → TA
(b)	 Cytosine	Nitrous acid (HNO <sub>2</sub> ) Deamination	 Uracil	 Adenine	CG → TA
(c)	 Cytosine	Hydroxylamine (NH <sub>2</sub> OH) Hydroxylation	 Hydroxylamino-cytosine	 Adenine	CG → TA

Oxidative radicals convert guanine into 8-oxy-7,8-dihydrodeoxyguanine.

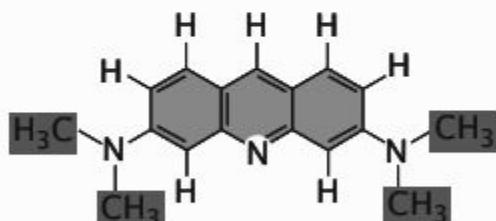


## Intercalating agents

(a)



Proflavin



Acridine orange

(b)

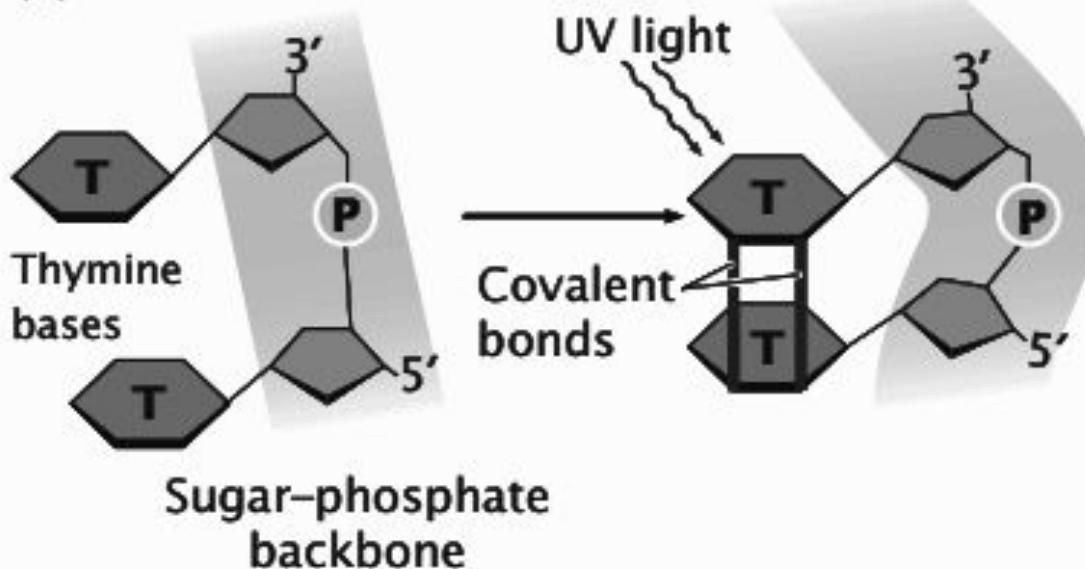


Nitrogenous bases  
Intercalated molecule

**18.21 Intercalating agents.** Proflavin and acridine orange (a) insert themselves between adjacent bases in DNA, distorting the three-dimensional structure of the helix (b).

Pyrimidine dimers result from ultraviolet light.

(a)



(b)



**18.22 Pyrimidine dimers result from ultraviolet light.**  
(a) Formation of thymine dimer. (b) Distorted DNA.

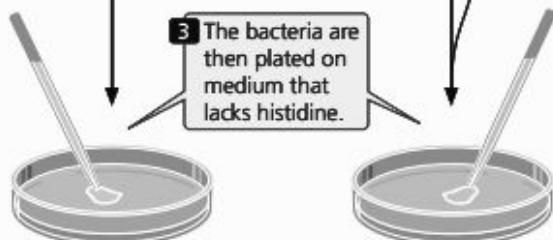
# The Ames test is used to identify chemical mutagens

## Methods

- 1 Bacterial  $his^{-}$  strains are mixed with liver enzymes (which have the ability to convert compounds into potential mutagens).
- 2 Some of the bacterial strains are also mixed with the chemical to be tested for mutagenic activity.



- 3 The bacteria are then plated on medium that lacks histidine.



## Results



- 4 Bacterial colonies that appear on the plates have undergone a  $his^{-} \rightarrow his^{+}$  mutation.

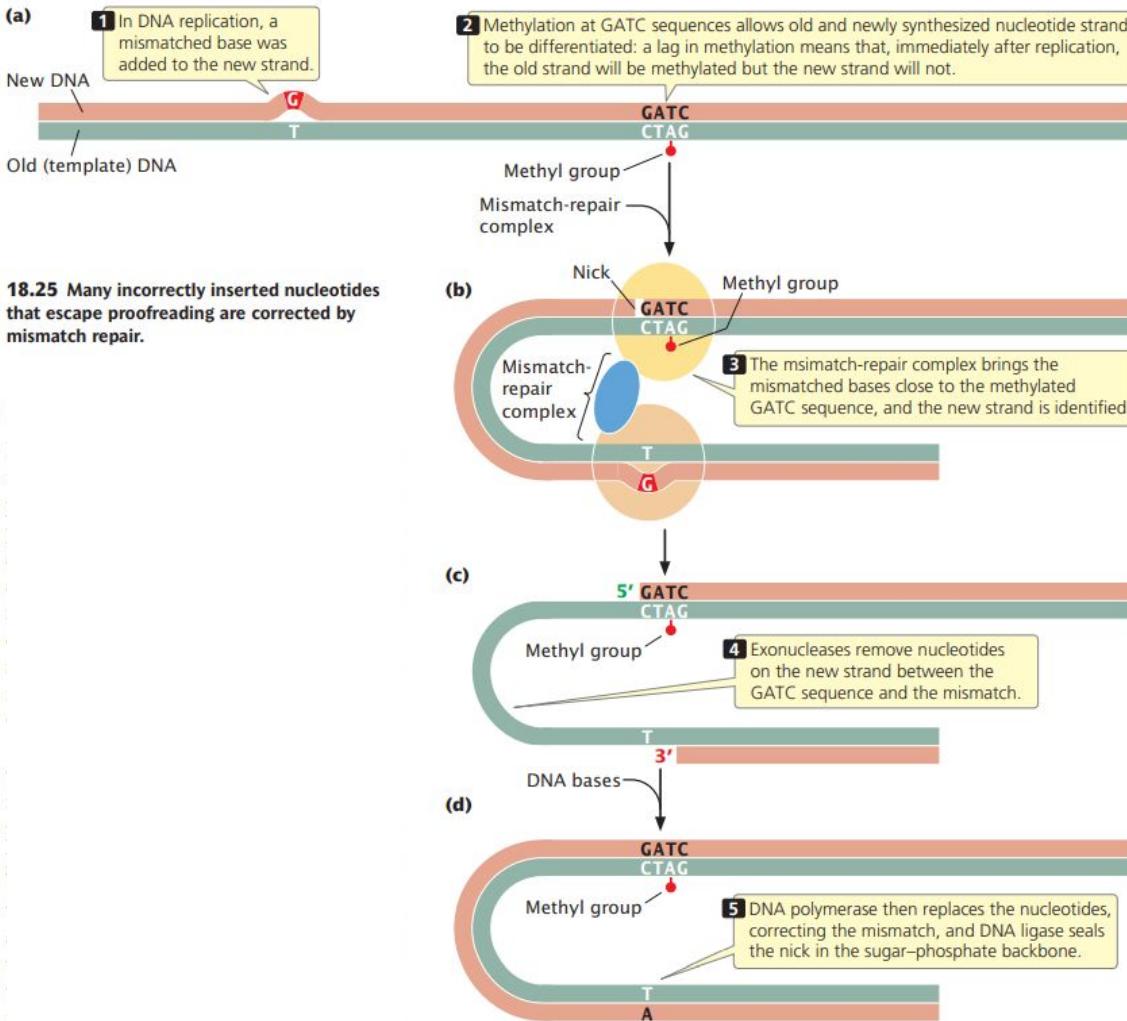
## Summary of common DNA repair

**Table 18.4**

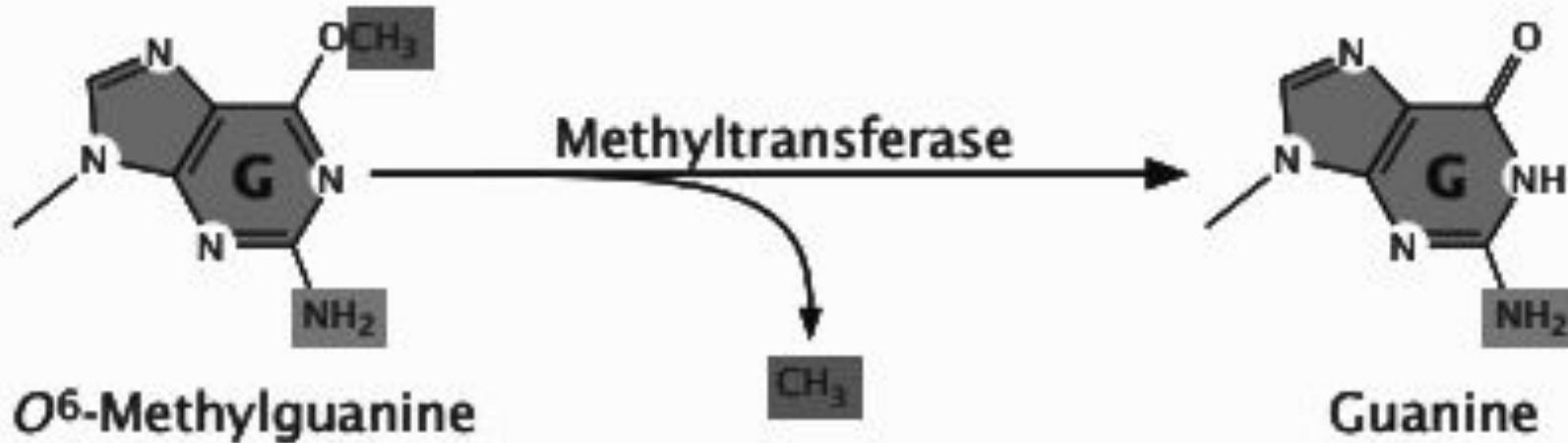
### Summary of common DNA repair mechanisms

Repair System	Type of Damage Repaired
Mismatch	Replication errors, including mispaired bases and strand slippage
Direct	Pyrimidine dimers; other specific types of alterations
Base excision	Abnormal bases, modified bases, and pyrimidine dimers
Nucleotide excision	DNA damage that distorts the double helix, including abnormal bases, modified bases, and pyrimidine dimers
Homologous recombination	Double-strand breaks
Nonhomologous end joining	Double-strand breaks

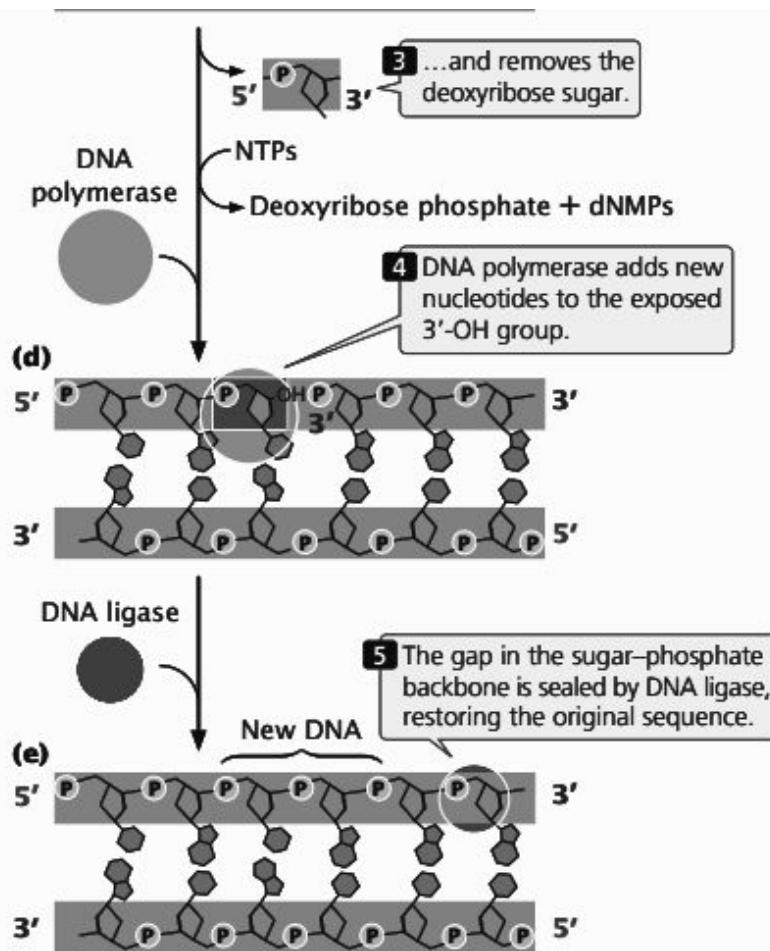
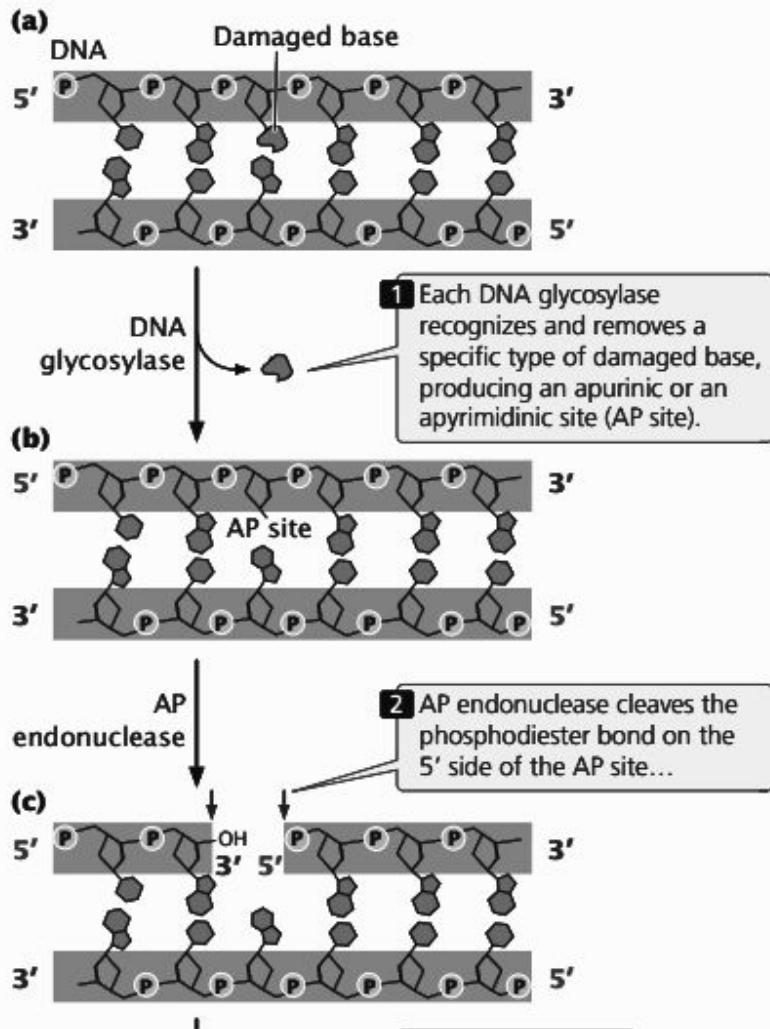
# Many incorrectly inserted nucleotides that escape proofreading are corrected by mismatch repair



## Direct Repair

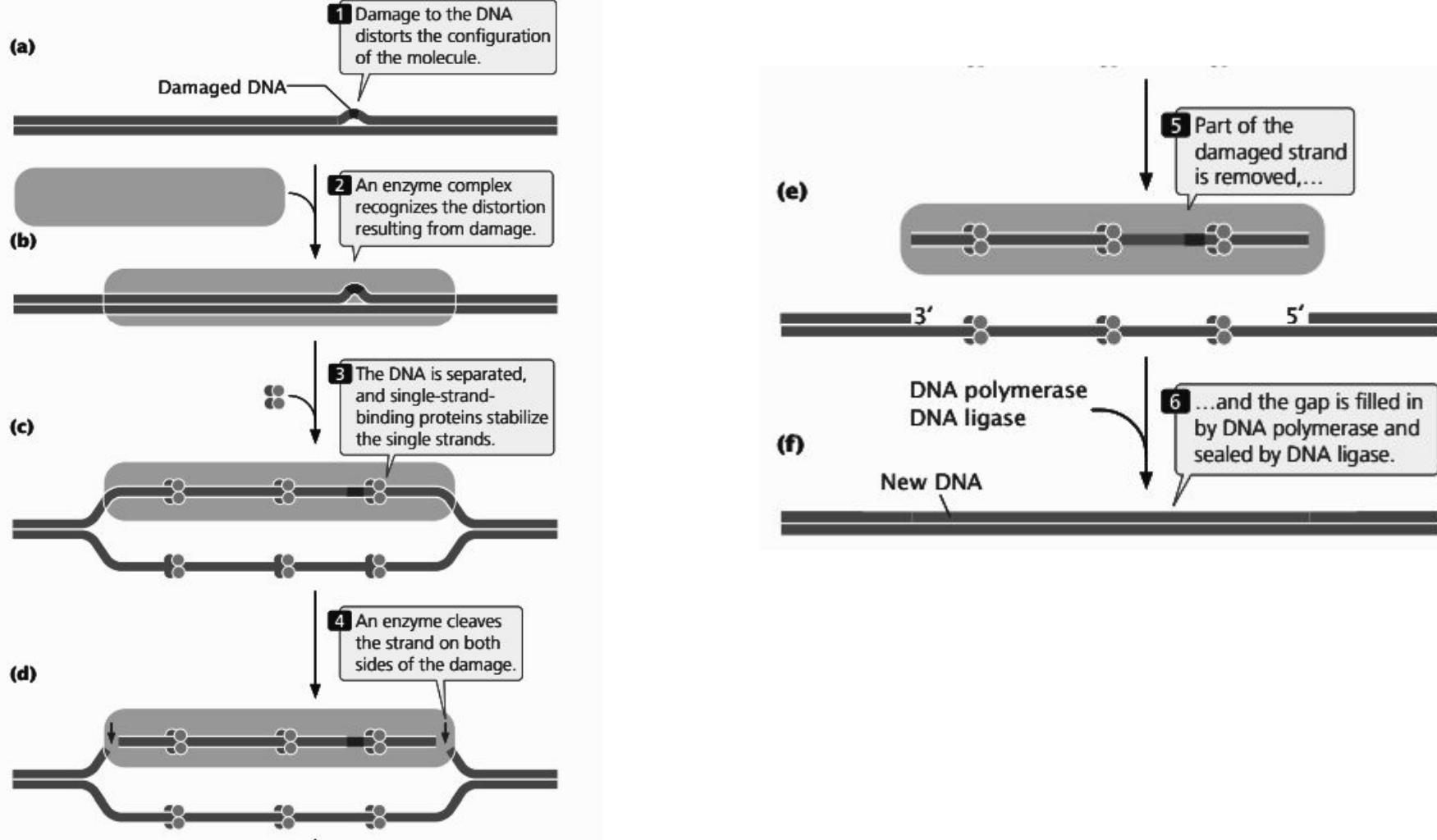


# Base-excision Repair

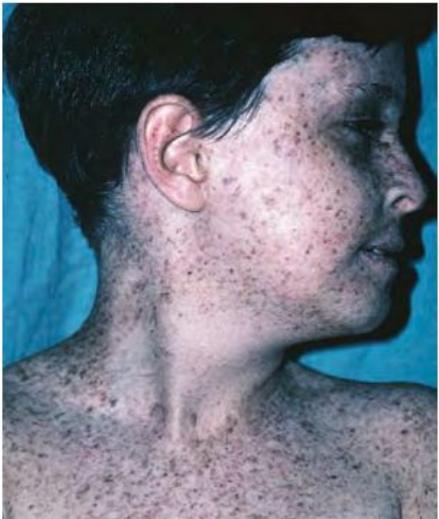


**18.27 Base-excision repair excises modified bases and then replaces one or more nucleotides.**

# Nucleotide-excision repair



# Nucleotide-excision repair

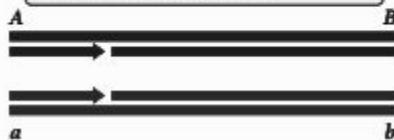


**Table 18.5** Genetic diseases associated with defects in DNA-repair systems

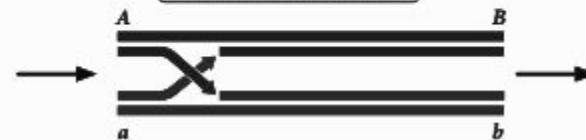
Disease	Symptoms	Genetic Defect
Xeroderma pigmentosum	Frecklelike spots on skin, sensitivity to sunlight, predisposition to skin cancer	Defects in nucleotide-excision repair
Cockayne syndrome	Dwarfism, sensitivity to sunlight, premature aging, deafness, mental retardation	Defects in nucleotide-excision repair
Trichothiodystrophy	Brittle hair, skin abnormalities, short stature, immature sexual development, characteristic facial features	Defects in nucleotide-excision repair
Hereditary nonpolyposis colon cancer	Predisposition to colon cancer	Defects in mismatch repair
Fanconi anemia	Increased skin pigmentation, abnormalities of skeleton, heart, and kidneys, predisposition to leukemia	Possibly defects in the repair of interstrand cross-links
Li-Fraumeni syndrome	Predisposition to cancer in many different tissues	Defects in DNA damage response
Werner syndrome	Premature aging, predisposition to cancer	Defect in homologous recombination

# The Holliday model of homologous recombination

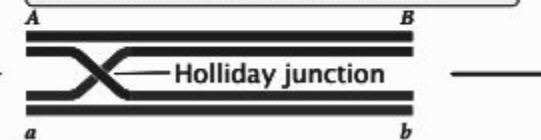
1 Homologous chromosomes align and single-strand breaks occur in the same position on both DNA molecules.



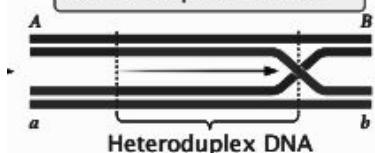
2 A free end of each broken strand migrates to the other DNA molecule.



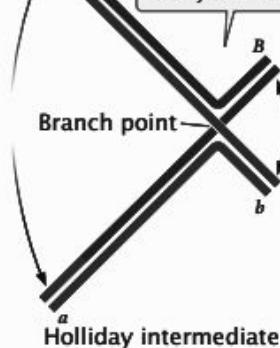
3 Each invading strand joins to the broken end of the other DNA molecule, creating a Holliday junction, and begins to displace the original complementary strand.



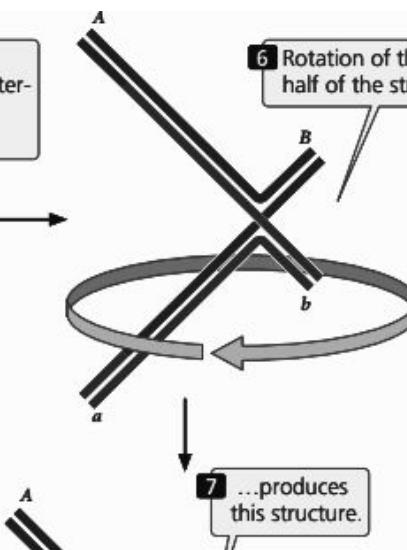
4 Branch migration takes place as the two nucleotide strands exchange positions, creating the two duplex molecules.



5 This view of the structure shows the ends of the two interconnected duplexes pulled away from one another.

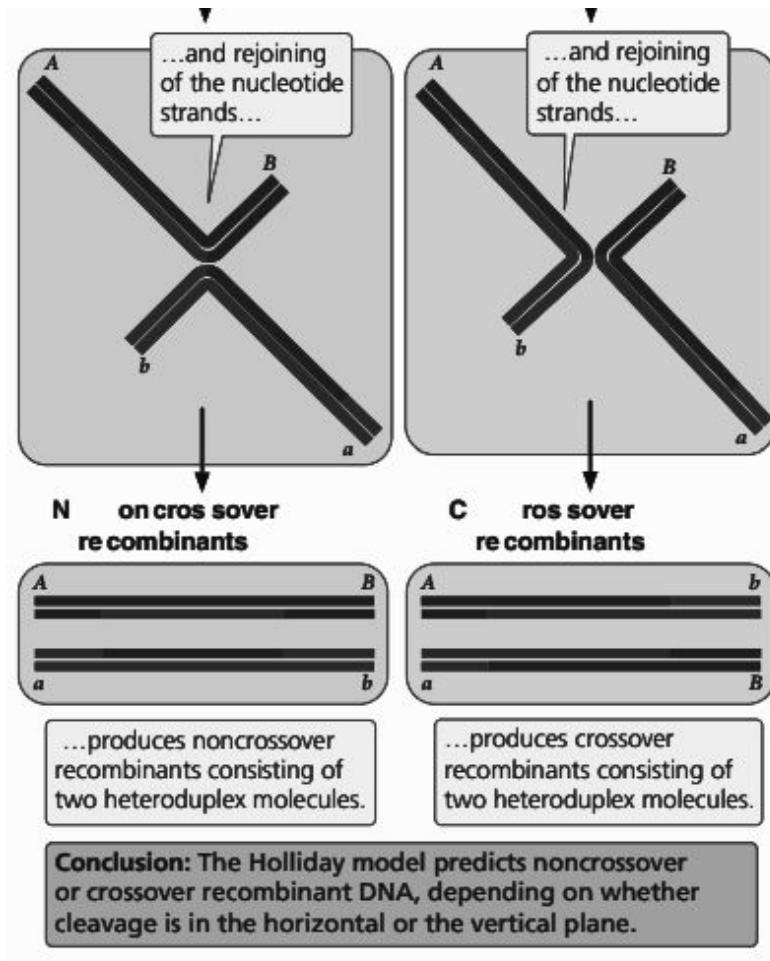
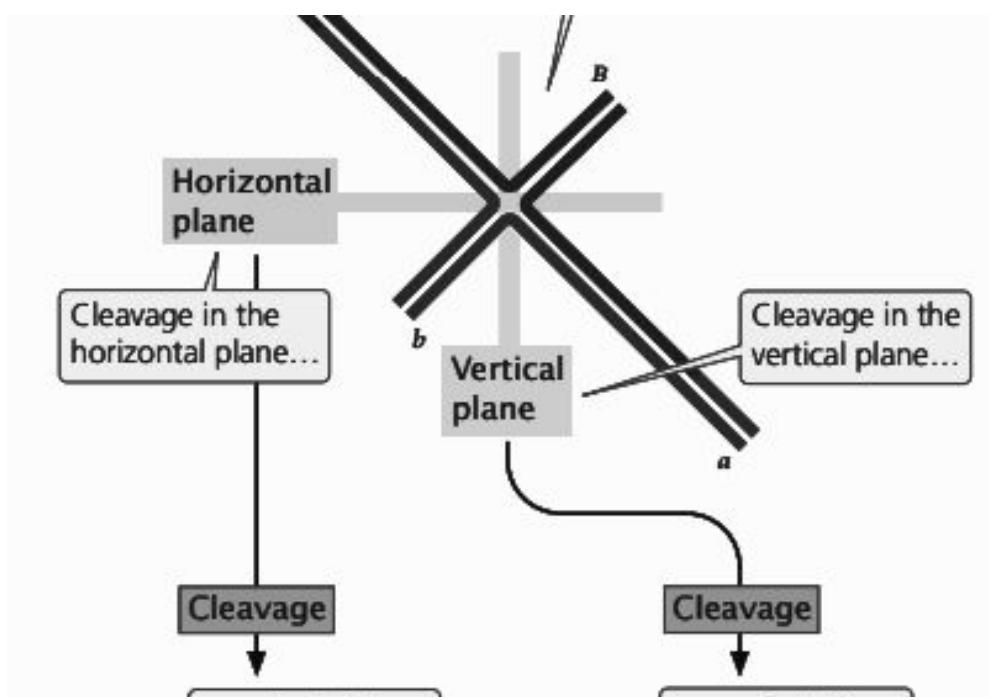


6 Rotation of the bottom half of the structure...

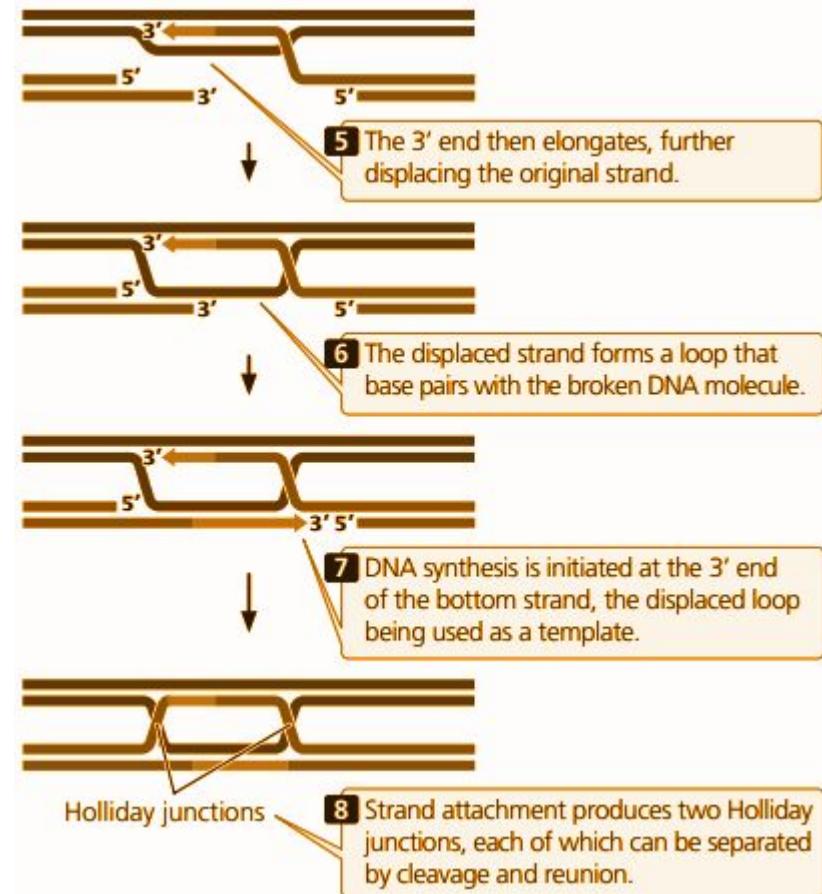
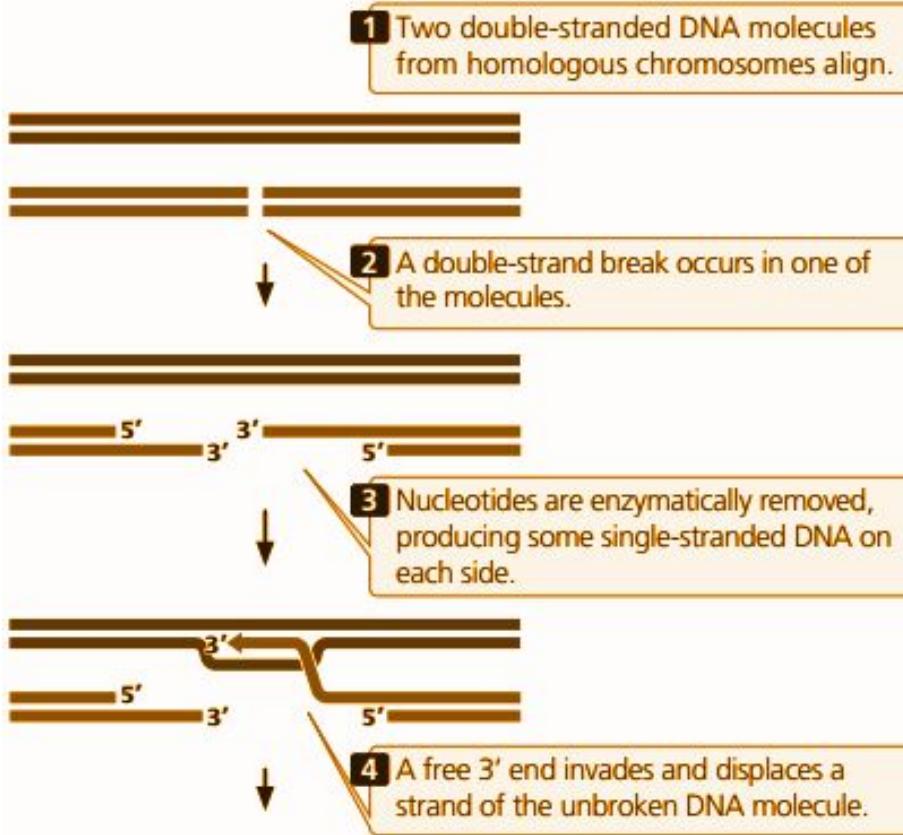


7 ...produces this structure.

# The Holliday model of homologous recombination



# The double-strand-break model

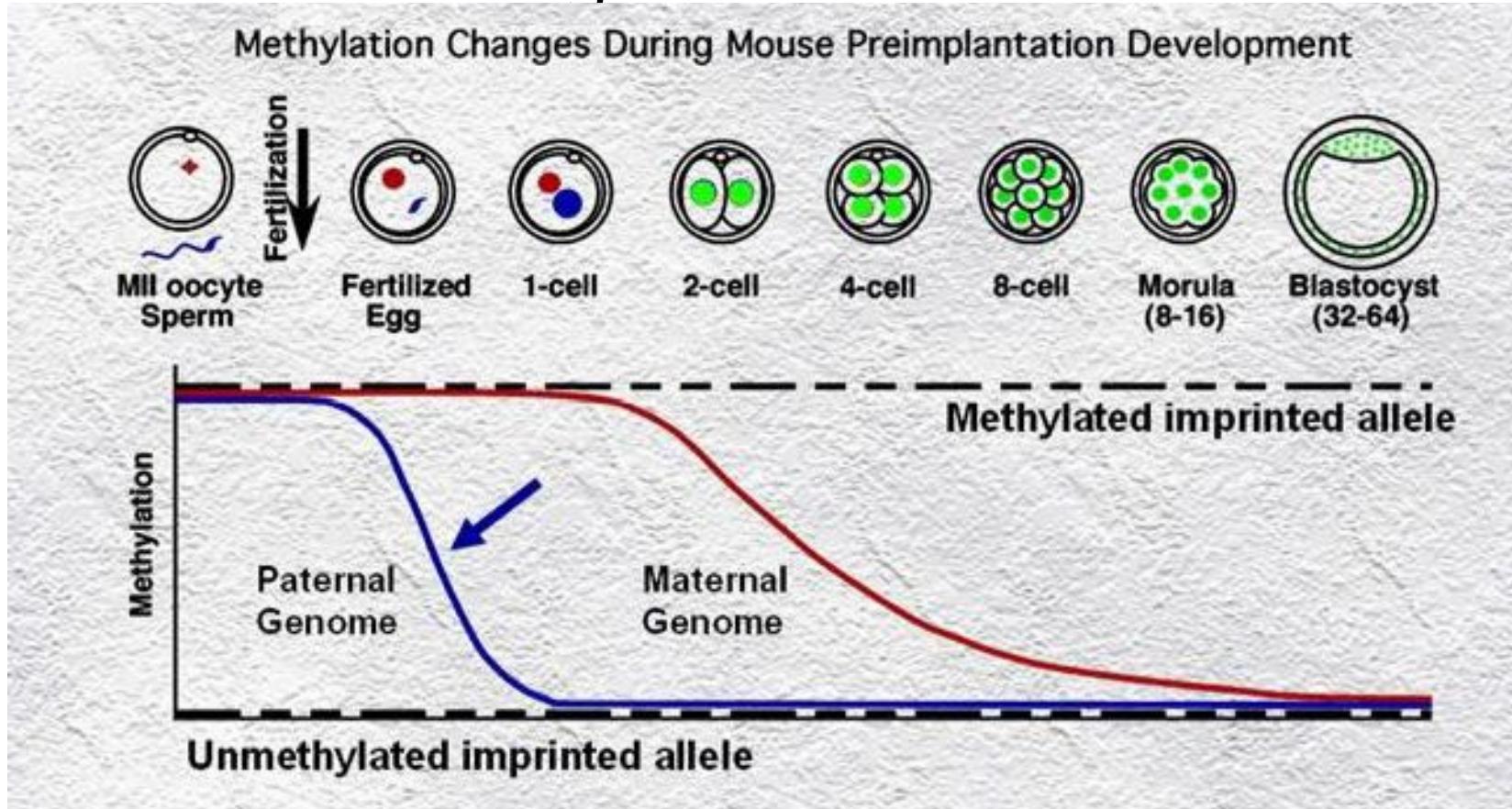


## Excellent Review Articles

- Friedberg, EC (2003) **DNA damage and repair.** *Nature* 421:436-440.
- Sancar A, Lindsey-Boltz LA, Unsal-Kacmaz K, Linn S (2004) **Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints.** *Annu Rev Biochem* 73: 39-85.

# EPIGENETIC

## S



Zygote



ACATAGACATACACACTGTTGATTAGGGAGATAGTGACAGATCCATTACAGCACCATAACCATGAT  
GTTTTTATTACCAAGGATGATCACCATGGGTACCATTTACCAAGGATTACACAGTTTAGATGACC  
AGTAGCTATTAGAGGATTAAATTATTTAGGATTATGGGATTGATAAAGGGAGATTTAACAA  
TAGACATACACACTGTTGATTAGGGAGATAGTGACAGATCCATTACAGCACCATAACCATGATGTT  
TTTATTACCAAGGATGATCACCATGGGTACCATTTACCAAGGATTACACAGTTTAGATGACCAGT  
AGCTATTAGAGGATTAAATTATTTAGGATTATGGGATTGATAAAGGGAGATTTTATTAT  
AGGACATAGACATACACACTGTTGATTAGGGAGATAGTGACAGATCCATTACAGCACCATAACCAT  
GATGTTTTATTACCAAGGATGATCACCATGGGTACCATTTACCAAGGATTACACAGTTTAGATG  
ACCAGTAGCTATTAGAGGATTAAATTATTTAGGATTATGGGATTGATAAAGGGAGATTTA  
ACATAGACATACACACTGTTGATTAGGGAGATAGTGACAGATCCATTACAGCACCATAACCATGAT

## How is the diversity of cell types created and maintained in multi-cellular organisms?



ACATAGACATACACACTGTTGATTAGGGAGATAGTGACAGATCCATTACAGCACCATAACCATGAT  
GTTTTTATTACCAAGGATGATCACCATGGGTACCATTTACCAAGGATTACACAGTTTAGATGACC  
AGTAGCTATTAGAGGATTAAATTATTTAGGATTATGGGATTGATAAAGGGAGATTTAACAA  
TAGACATACACACTGTTGATTAGGGAGATAGTGACAGATCCATTACAGCACCATAACCATGATGTT  
TTTATTACCAAGGATGATCACCATGGGTACCATTTACCAAGGATTACACAGTTTAGATGACCAGT  
AGCTATTAGAGGATTAAATTATTTAGGATTATGGGATTGATAAAGGGAGATTTTATTAT  
AGGACATAGACATACACACTGTTGATTAGGGAGATAGTGACAGATCCATTACAGCACCATAACCAT  
GATGTTTTATTACCAAGGATGATCACCATGGGTACCATTTACCAAGGATTACACAGTTTAGATG  
ACCAGTAGCTATTAGAGGATTAAATTATTTAGGATTATGGGATTGATAAAGGGAGATTTA  
ACATAGACATACACACTGTTGATTAGGGAGATAGTGACAGATCCATTACAGCACCATAACCATGAT

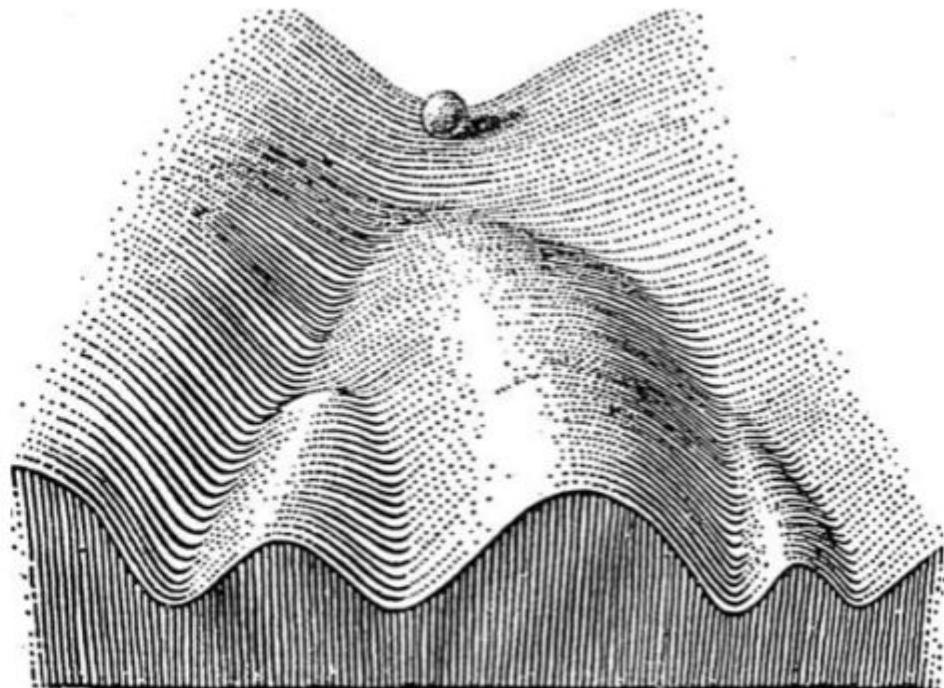
# Identical Twins with Different Hair Color



How can identical twin litter mates show different coat colors?

# What is Epigenetics?

1. C.H. Waddington coined the term epigenetics to mean above or in addition to genetics to explain differentiation.
2. How do different adult stem cells know their fate?
3. Myoblasts can only form muscle cells
4. Keratinocytes only form skin cells
5. Hematopoietic cells only become blood cells
6. But all have identical DNA sequences.
7. Modern definition is non-sequence dependent inheritance.
8. How can identical twins have different natural hair colors?



Waddington's Epigenetic Landscape

# Developmental potential

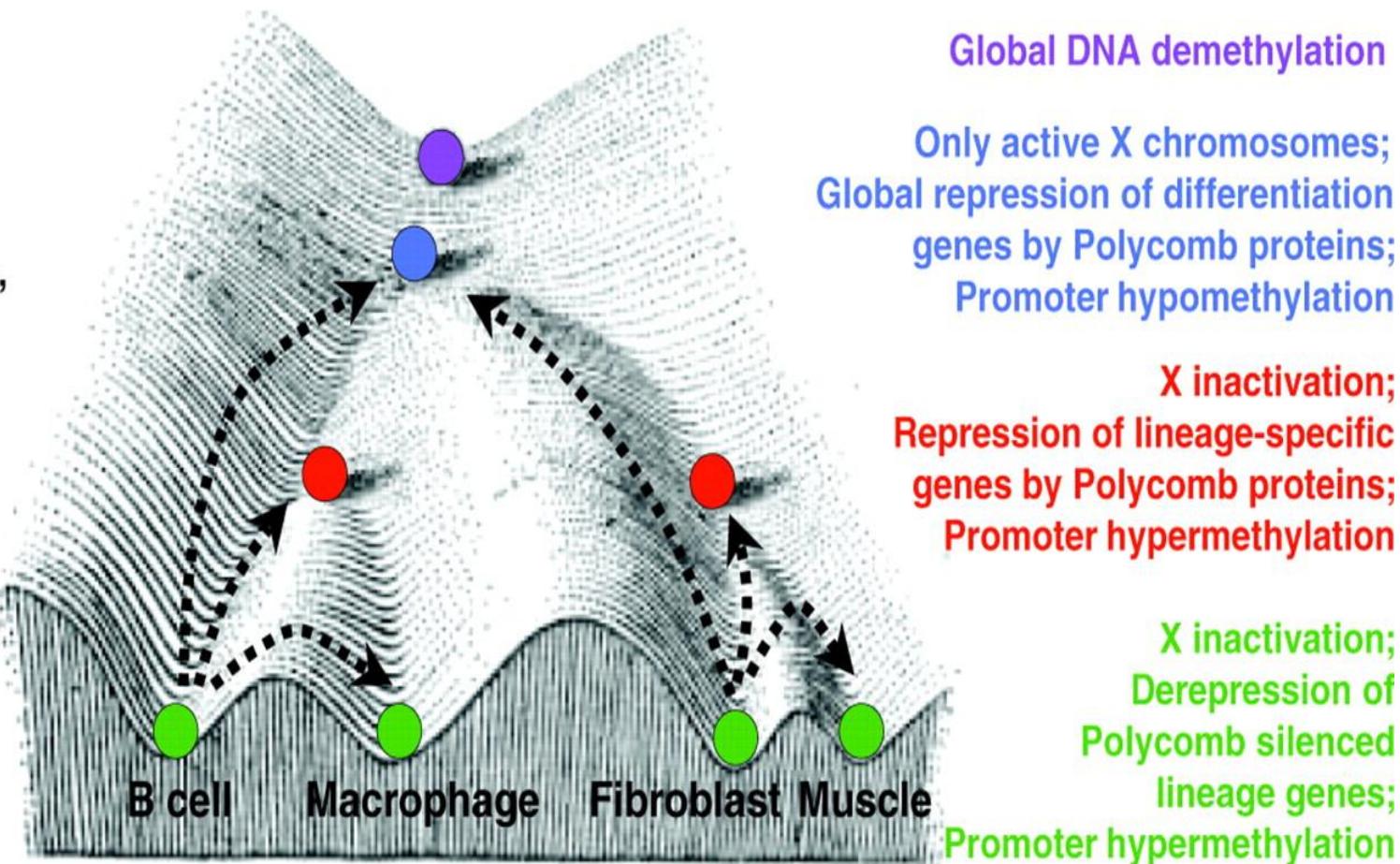
Totipotent  
Zygote

Pluripotent  
ICM/ES cells, EG cells,  
EC cells, mGS cells  
iPS cells

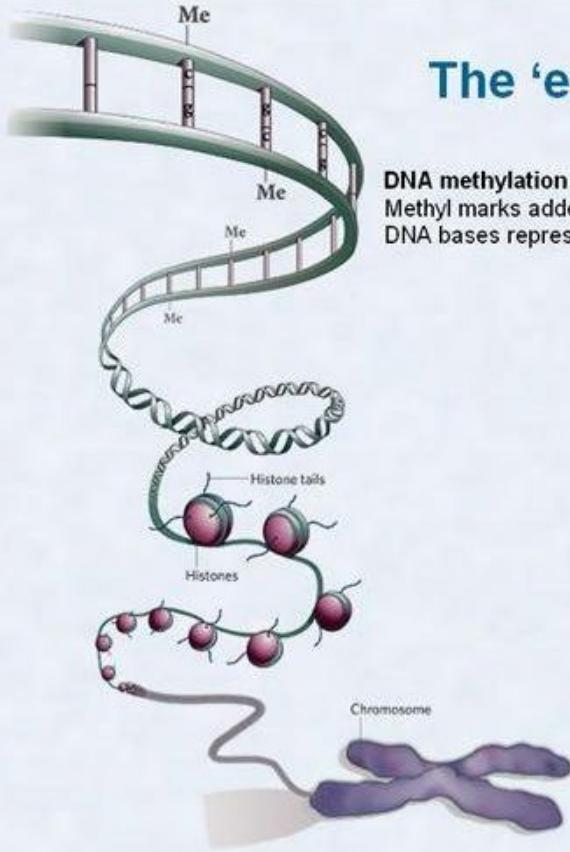
Multipotent  
Adult stem cells  
(partially  
reprogrammed cells?)

Unipotent  
Differentiated cell  
types

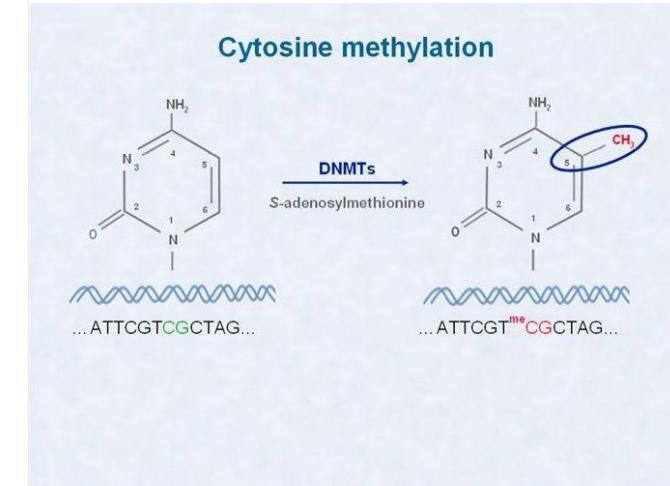
# Epigenetic status



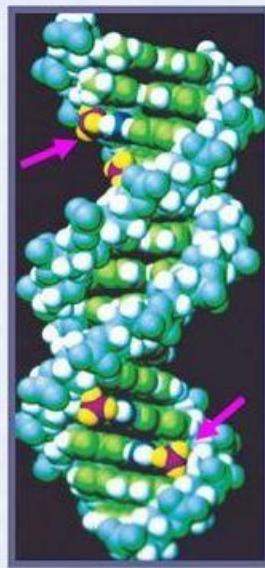
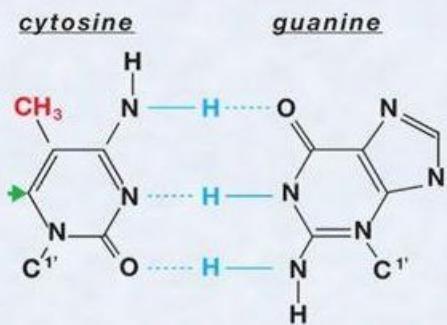
# DNA Methylation & Histone Modifications Form the Epigenetic Code



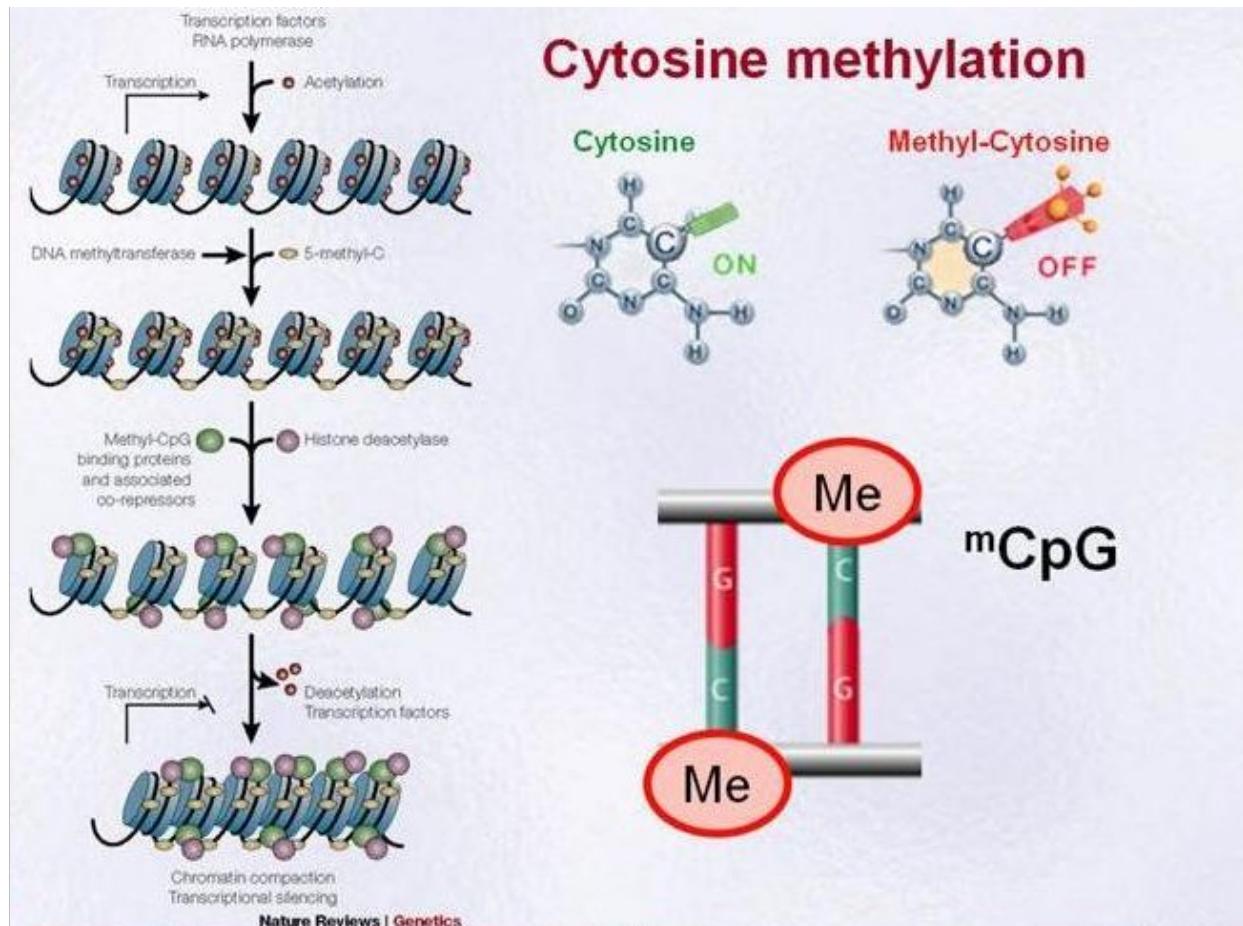
## The 'epigenetic' code



## Cytosine methylation

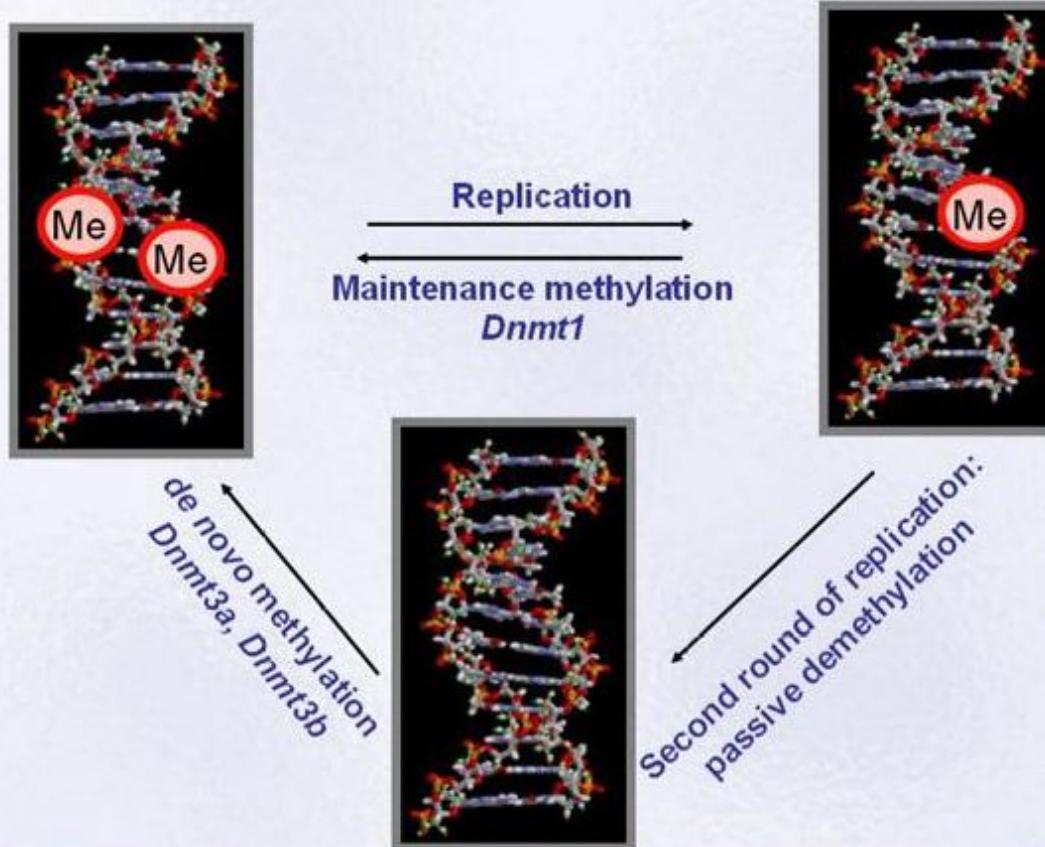


# Cytosine Methylation Maintains Inactive-Condensed Chromatin State



# Establishment and Maintenance of Cytosine Methylation

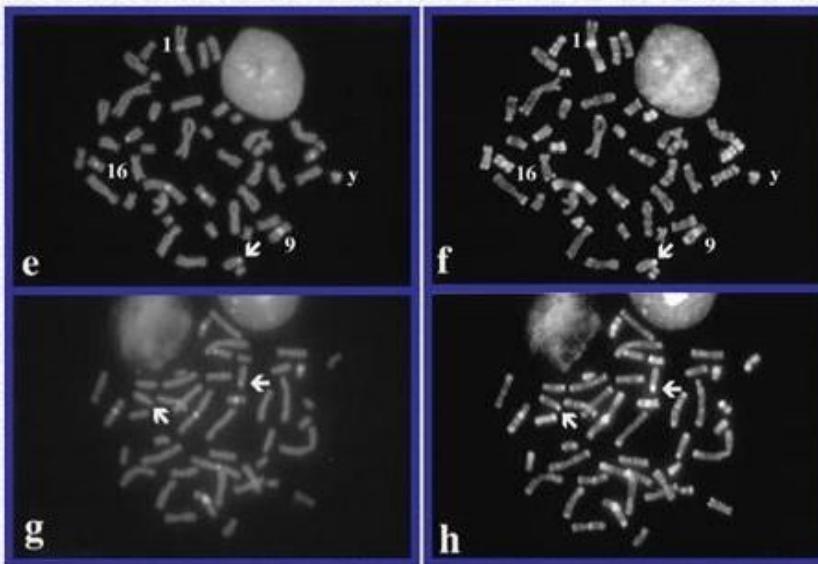
## Establishment and maintenance



# 5-Methyl Cytosine is Found in Heterochromatic Regions

## The distribution of cytosine methylation in mammals

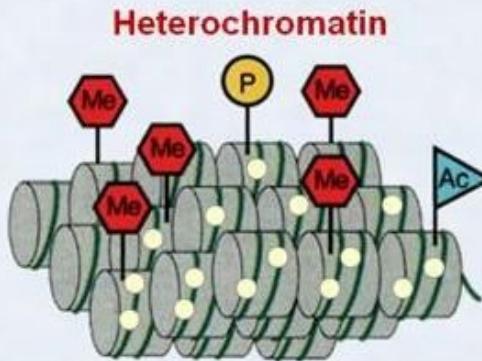
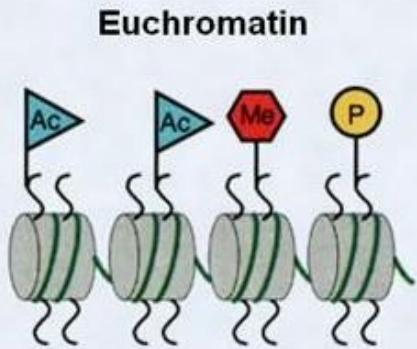
- Heterogeneity visible at cytogenetic scale
- Associated with heterochromatic regions



PMID: 9609658

# Structure & Epigenetics of Euchromatin versus Heterochromatin

DNA methylation and histone modifications  
help to compartmentalize the genome  
into domains of different transcriptional potentials



- High histone acetylation
  - Low DNA methylation
  - H3-K4 methylation
- Low histone acetylation
  - Dense DNA methylation
  - H3-K9 methylation

# Some DNA Methyl Transferases are Essential

## Mammalian Dnmts are essential

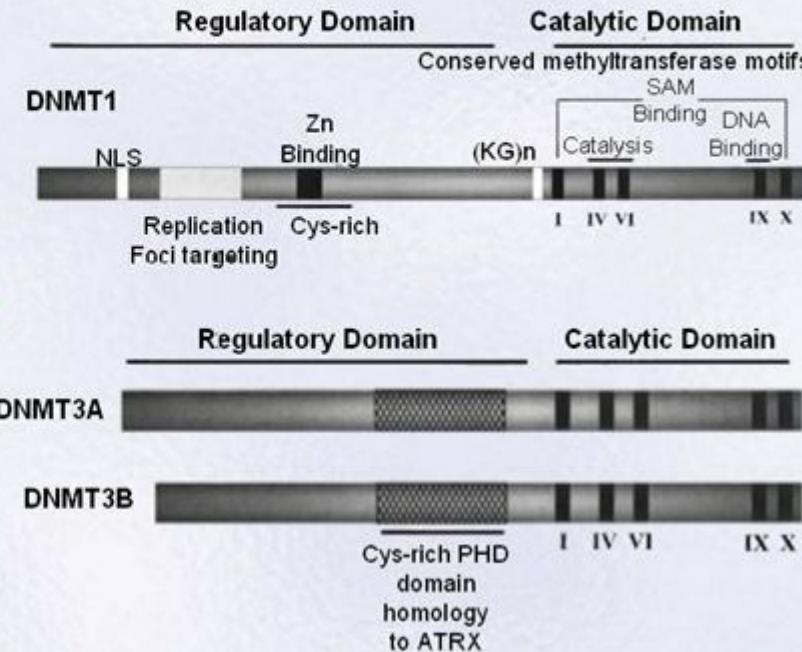
Dnmt1: embryonic lethal

Dnmt2: no obvious effect

Dnmt3a: perinatal death

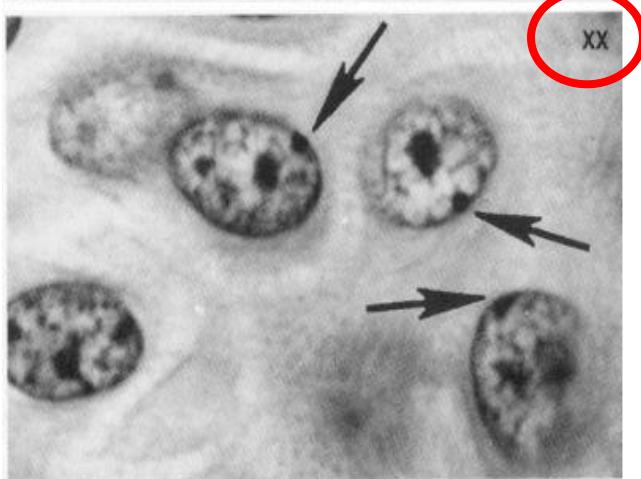
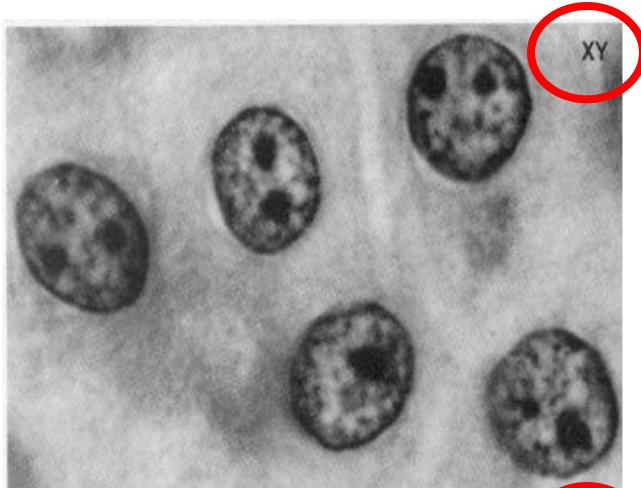
Dnmt3b: embryonic lethal

Dnmt3l: no imprints



Robertson, KD, *Oncogene* 2002

# Lyonization

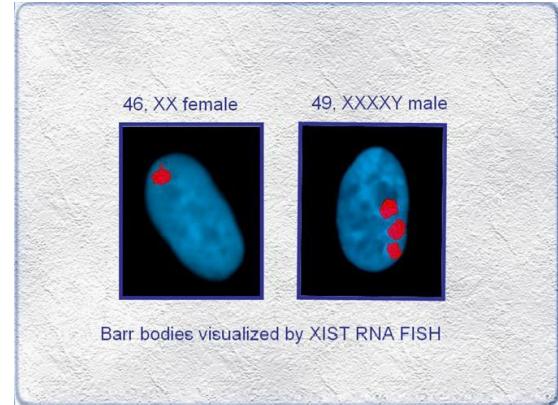


Barr Bodies = inactivated X chromosome

# Lyonization

## Mechanism

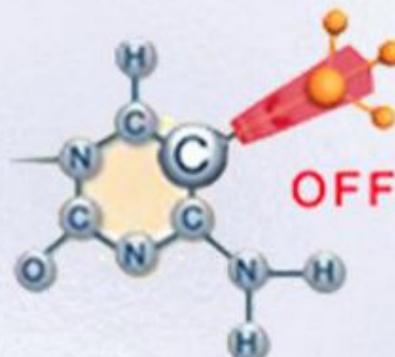
- **XIST** gene on the X chromosome turns on and produces **XIST** RNA.
- Molecules of **XIST** RNA accumulate along the chromosome with the active **XIST** gene.
- The binding of the **XIST** RNA with the DNA turns off the genes on that chromosome.



## Some DNA Methyl Transferases are Essential

### Cytosine methylation in mammals

- Gene expression
- Chromosomal stability
- Cell differentiation
- Imprinting
- X-Inactivation
- Carcinogenesis
- Aging

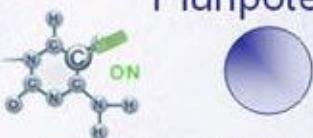


# DNA Methylation Differentiates Totipotent Embryonic Stem Cells from Unipotent Adult Stem Cells

## DNA methylation

Cytosine C

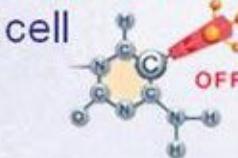
Pluripotent cell



ctggagggtcaatggctgtttgtcctggcatt  
ggacatgggctgaaataactgggttcaccatat  
ctaggactctaga~~gggtgggt~~aagcaagaact  
gaggagtggccccagaaataattggcacacgaa  
catcaatggatgttttaggctctccagaggat  
ggctgagtgggtgttaaggacaggcccggaggg  
tgcagtgccaacaggcttggtgatgggg  
catcgagcaactgggttgtgaggtgtccgggt  
acccaaggcaggggtgagaggaccttgaagggtt  
gaaaatgaaggcctctgggtccgtcctaag  
ggttgtcctgtccagactcccaacctcgtc  
tggaaagacacaggcagatagegtccctcaigt  
ttctcccaccccccacagctctgtctccaccc  
acccagggggg~~ggggcc~~cagaggtaaggctaga  
gggtgggattggggagggagaggtaaaact  
cttaggtgag~~gt~~tttccaccaggccccgg  
ctgggggtgcccac~~tt~~ccccatggctggacac

Methyl-Cytosine 5mC

Unipotent cell

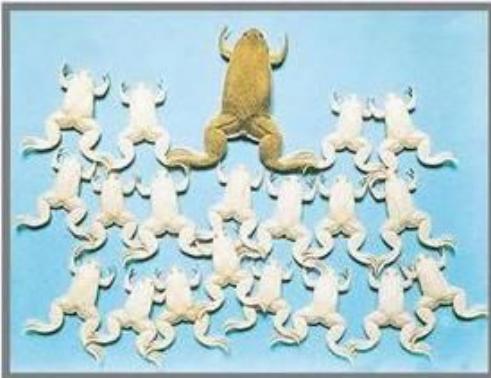


≠

Ctggagggtcaatggctgtttgtcctggcatt  
ggacatgggctgaaataactgggttcaccatat  
ctaggactctaga~~gggtgggt~~aagcaagaact  
gaggagtggccccagaaataattggcacacgaa  
catcaatggatgttttaggctctccagaggat  
ggctgagtgggtgttaaggacaggcccggaggg  
tgcagtgccaacaggcttggtgatgggg  
catcgagcaactgggttgtgaggtgtccgt  
acccaaggcaggggtgagaggaccttgaagggtt  
gaaaatgaaggcctctgggtccgtcctaag  
ggttgtcctgtccagactcccaacctcgtc  
tggaaagacacaggcagatagegtccctcaigt  
ttctcccaccccccacagctctgtctccaccc  
acccagggggg~~ggggcc~~cagaggtaaggctaga  
gggtgggattggggagggagaggtaaaact  
cttaggtgag~~gt~~tttccaccaggccccgg  
ctgggggtgcccac~~tt~~ccccatggctggacac

# Differentiated Cells can Become Totipotent

Nuclear transplantation demonstrates nuclear equivalence

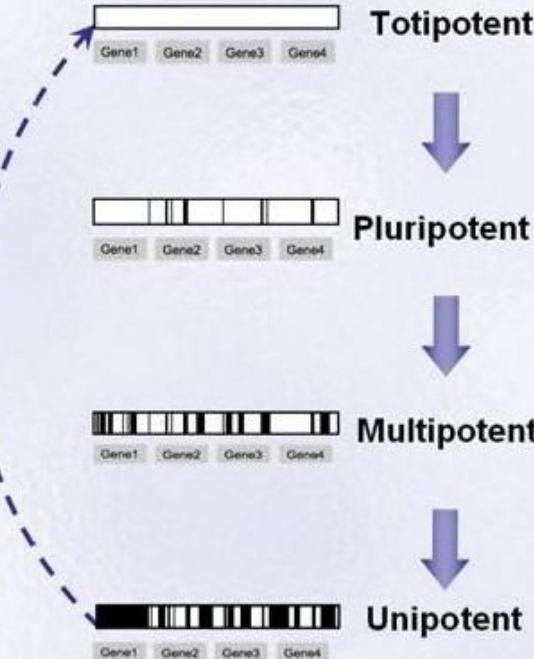


Briggs and King, 1952

Gurdon, 1960s

“Dolly”

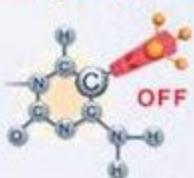
Differentiated cells maintain the potential to generate an entire organism



# Critical CpG Sequences in CpG Islands Near Promoters

## Genomic distribution of DNA methylation

### Methyl-Cytosine

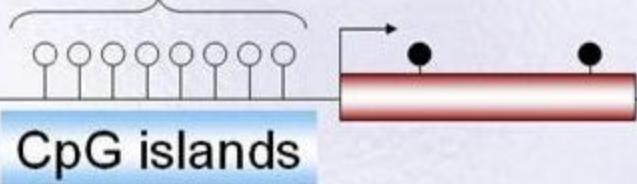


4% of all cytosines are methylated  
70-80% of all CpGs are methylated

98% of the genome  
1 CpG/100bp  
majority methylated



<2% of the genome  
1 CpG/10bp short stretches (~1000bp)  
majority unmethylated



## Recombinant DNA Technology

- Recombinant DNA Technology allows DNA to be produced via artificial means.
- The procedure has been used to change DNA in living organisms and may have even more practical uses in the future.

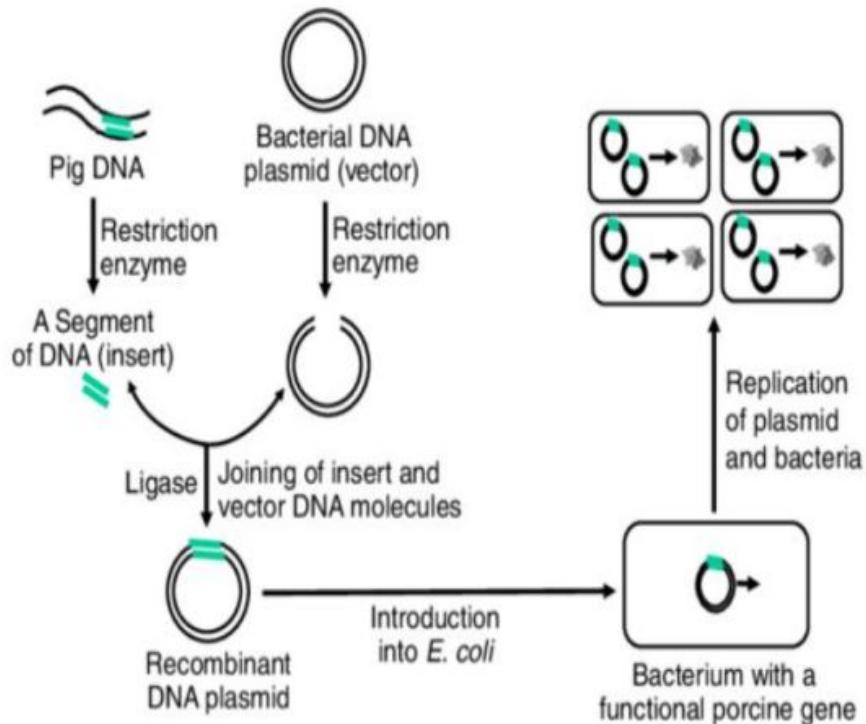
Recombinant DNA Technology works by taking DNA from two different sources and combining that DNA into single molecule. It was developed by Boyer and Cohen in 1973.

- This technology only becomes useful when that artificially created DNA is reproduced. That is known as DNA Cloning.



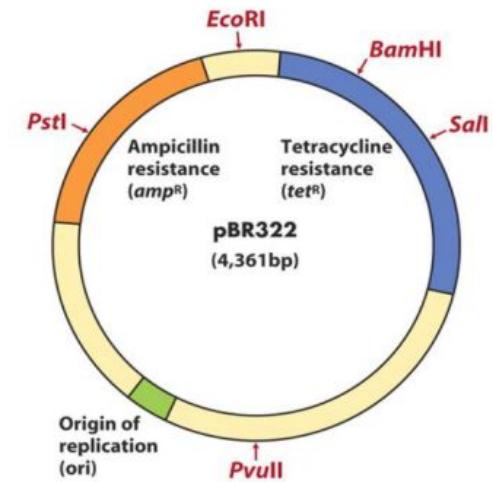
## BASIC PRINCIPLES OF RECOMBINANT DNA TECHNOLOGY

- Generation of DNA fragments and Selection of the desired piece of DNA.
- Insertion of the selected DNA into a cloning vector to create a rDNA.
- Introduction of the recombinant vectors into host cells.
- Multiplication and selection of clones containing the recombinant molecules.
- Expression of the gene to produce the desired product.



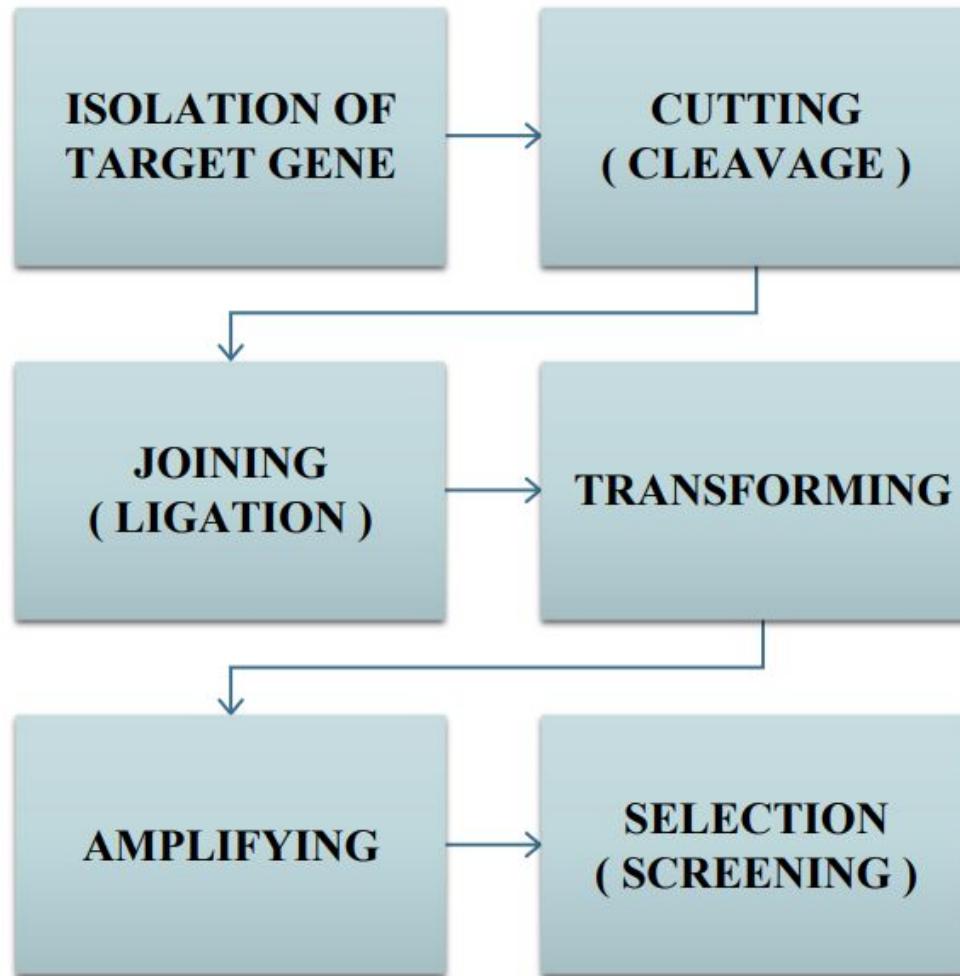
## BASIC PRINCIPLES OF RECOMBINANT DNA TECHNOLOGY

- A vector is an area of DNA that can join another DNA part without losing the limit for self-replication.
- Should be capable of replicating in host cell.
- Should have convenient RE sites for inserting DNA of interest.
- Should have a selectable marker to indicate which host cells received recombinant DNA molecule.
- Should be small and easy to isolate.
- One type of vector Plasmid has those character, so those are useful in RDT.
- pBR322 was one of the first versatile plasmid vector developed.
- pBR322 contains an origin of replication (ori) and a gene (rop) that helps regulate the no. of copies of plasmid DNA in a cell. There are also 2 marker genes: confers resistance to Ampicillin and Tetracycline. It also contain a number of Unique Restriction Sites that are useful for constructing



**Fig - Plasmid Vector**

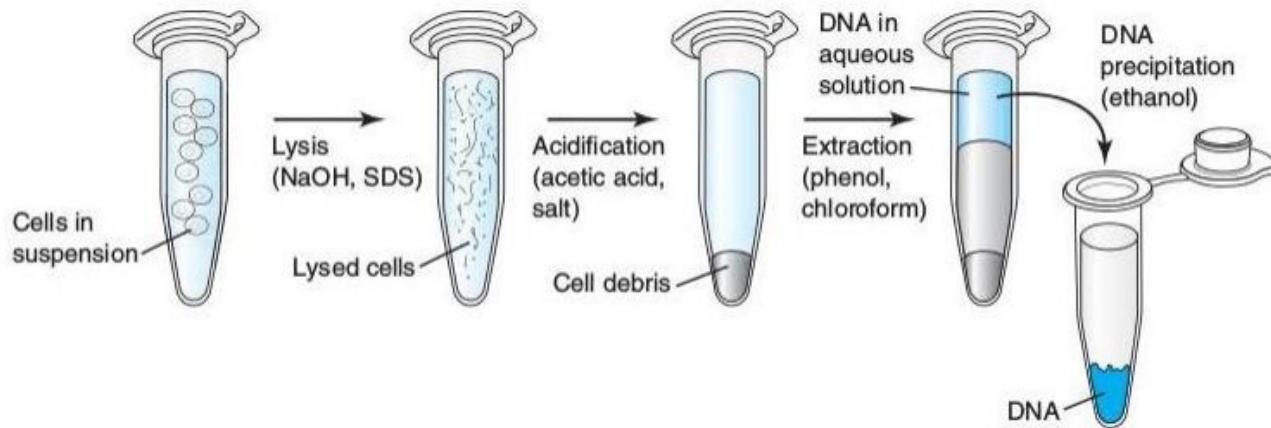
## SIX STEPS OF RECOMBINANT DNA



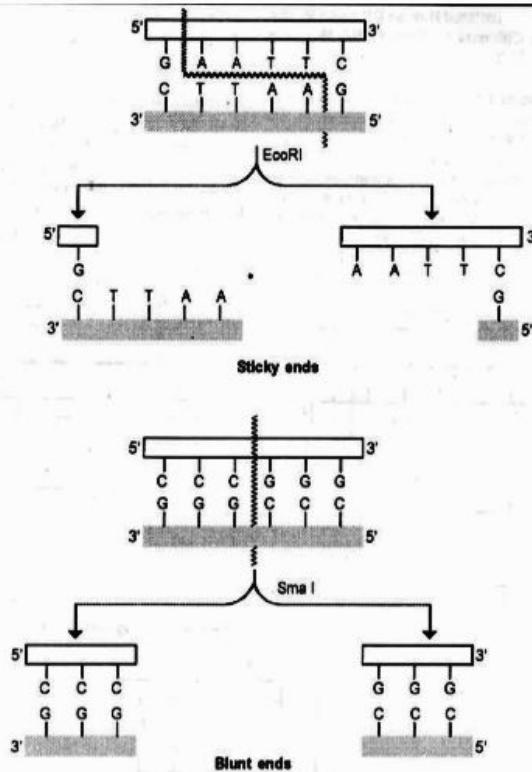
# ISOLATION OF DNA

- **BASIC STEPS :-**

- ✓ Cell lysis
- ✓ Removal of Proteins
  - Protease
  - Adsorption or Extraction
- ✓ DNA precipitation by Ethanol
- ✓ DNA dilution in Water or Buffer

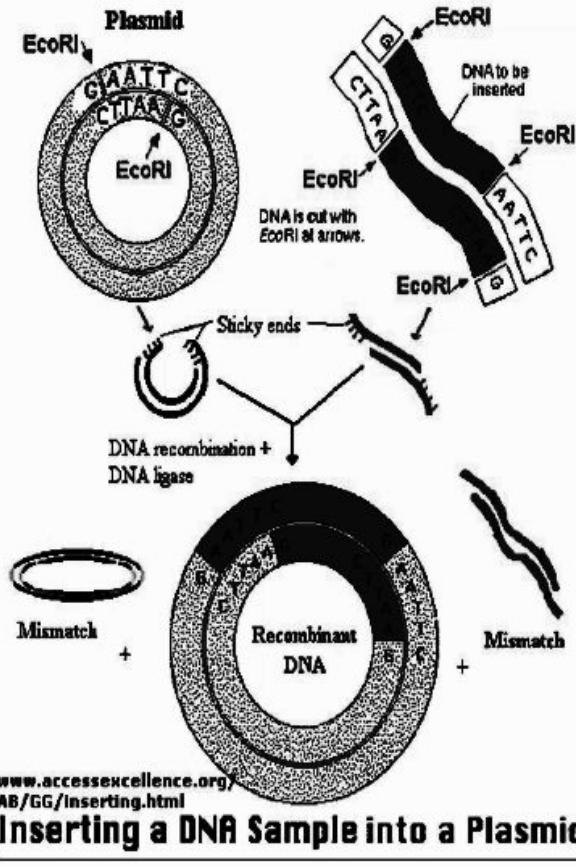


## RESTRICTION ENZYMES



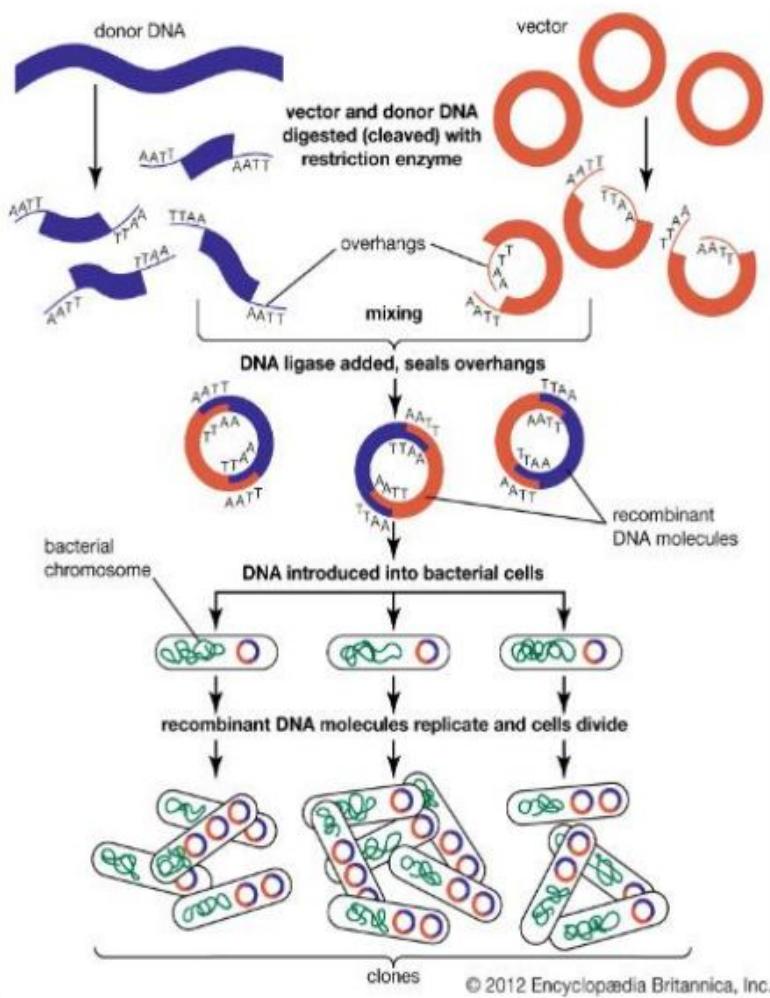
- These are the bacterial enzymes that can cut/split DNA at specific sites.
- These were first discovered in E.coli restricting the replication of bacteriophages, by cutting the viral DNA.
- **RECOGNITION SEQUENCE :**
  - ✓ Recognition sequence is the site where the DNA is cut by a REase, the restriction sites are short palindromic sequence.
  - ✓ REases can specifically recognise DNA with a particular sequence of 4-8 bp & cleave.
- **CLEAVAGE PATTERNS :**
  - ✓ The cut DNA fragments by REase may have mostly Sticky (the key to recombinant DNA) or Blunt ends e.g. Eco RI can make sticky ends and Sma I makes blunt ends.
  - ✓ DNA fragments with Sticky Ends are useful for RDT.

## LIGATION



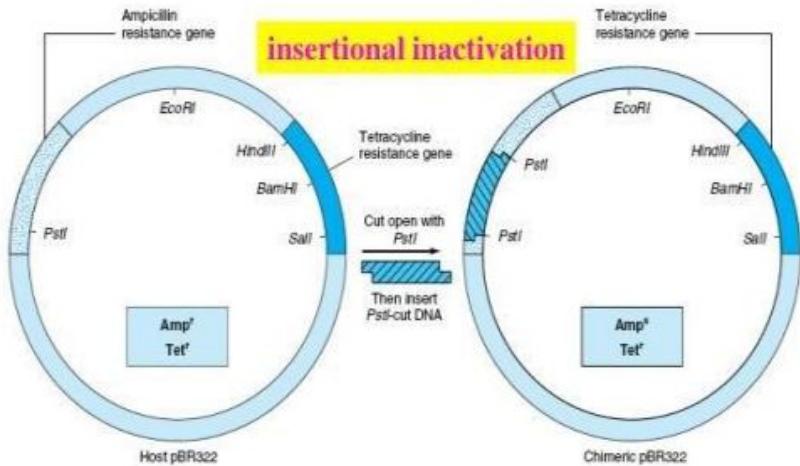
- DNA fragments with sticky ends can be inserted into vector DNA with the aid of **DNA Ligases**.
- For purposes of DNA cloning, purified DNA Ligase is used to covalently join the ends of a restriction fragment and vector DNA that have complementary ends. The vector DNA and restriction fragment are covalently ligated together through the strand 3' to 5' of DNA.
- DNA ligase “pastes” the DNA fragments together.

## TRANSFORMING AND CLONING



- Once a cloning vector and inserted DNA have been ligated, the rDNA molecule can be introduced into a bacterial host cell such as E.coli.
- The cells are treated with CaCl<sub>2</sub>.
- DNA is added.
- Cells are heat shocked at 42°C.
- Once in a cell, the recombinant DNA will be replicated.
- When the cell divides, the replicated recombinant molecules go to both daughter cells which themselves will divide later. Thus the DNA is amplified or cloned.

## SELECTION STRATEGIES USE MARKER GENES



- Many selection strategies involve selectable marker genes – genes whose presence can easily be detected. *ampR*
- Selection or screening can also be achieved using insertional inactivation.
- INSERTIONAL INACTIVATION :** Using the plasmid pBR322, a piece of DNA is inserted into the unique *Pst*I site. This insertion disrupts the gene coding for a protein that provides ampicillin resistance to the host bacterium. Hence the chimeric plasmid will no longer survive when plated on a substrate medium that contains this antibiotic. The differential sensitivity to tetracycline and ampicillin can therefore be used to distinguish clones of plasmid that contain an insert.

# Application of RDT

- **PHARMACEUTICAL PRODUCTS :**

- ✓ A no. of therapeutic gene products – insulin, the interleukins, interferons, growth hormones, etc. are now produced commercially from cloned genes.

- ✓ Production of Safer vaccines.

- **AGRICULTURAL USE :**

- ✓ Herbicide or pesticide resistant corn and soybean production.

- ✓ Development of “**Golden Rice**” with beta-carotene.

- ✓ Insect resistant tomato plants produced.

- ✓ Cold, Drought & Salinity tolerant crops production.

- **TRANSGENIC ANIMALS :**

- ✓ Many also made by RDT like many fluorescence fish, transgenic mouse, etc.

- **DIAGNOSIS OF MOLECULAR DISEASES :**

- ✓ Sickle cell anaemia, thalassaemia, familial hypercholesterolaemia, cystic fibrosis.

- **GENE THERAPY :**

- ✓ Adenosine deaminase (ADA) deficiency has been successfully treated.

- **INDUSTRIAL APPLICATION :**

- ✓ Enzymes – use to produce sugar, cheese, detergents.

- ✓ Protein Products – used as food additives, increase nutrient value, besides imparting flavour.



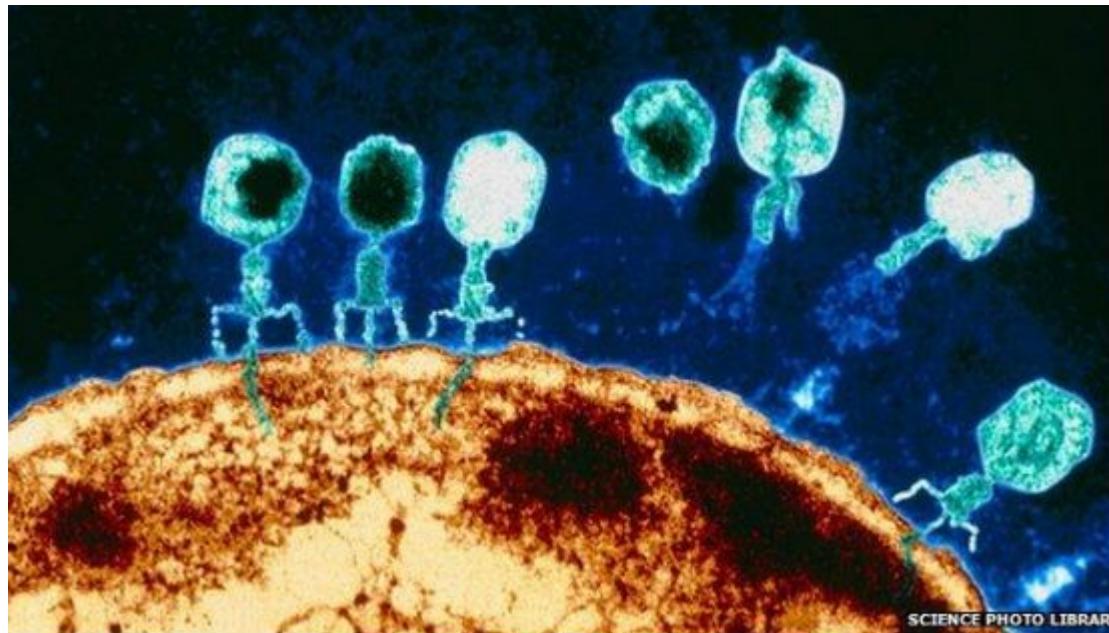
## CRISPR wins the Nobel Prize!

Emmanuelle Charpentier and Jennifer Doudna won the 2020 Nobel Prize in Chemistry “for the development of a method for genome editing.”



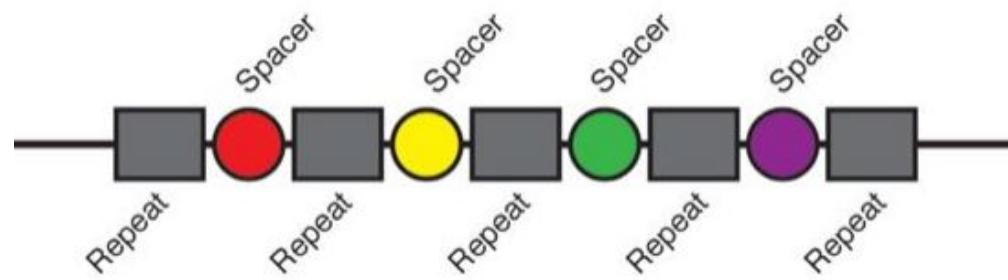
## What is CRISPR?

The CRISPR-Cas9 system evolved in bacteria as a defense against viral attacks. When foreign DNA is recognized, it is chopped into pieces to prevent replication.



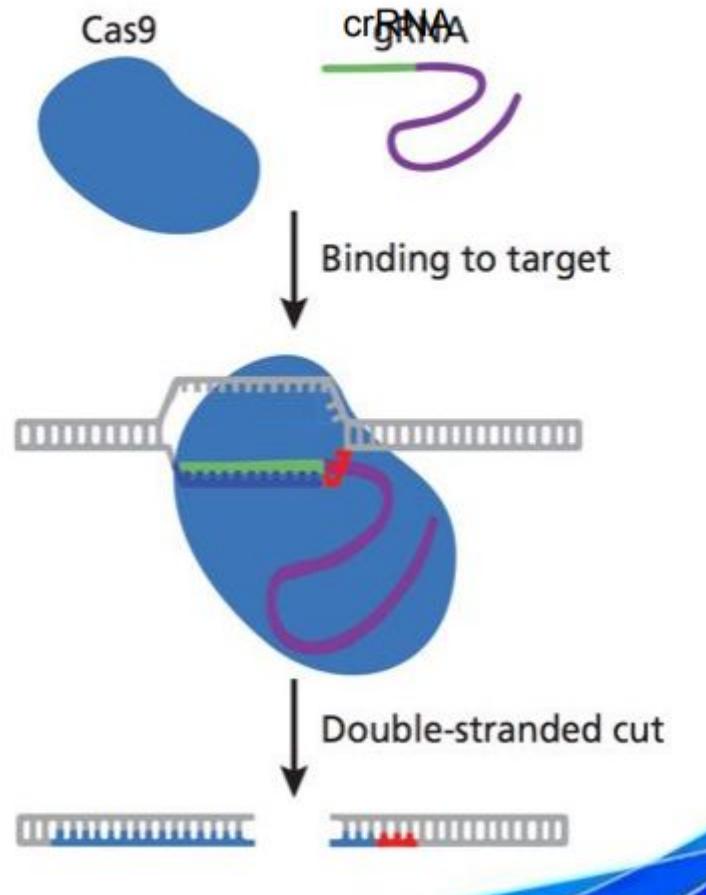
# What is CRISPR?

- Clustered Regularly Interspaced Short Palindromic Repeats.
- The bacterial genome contains stretches of repeating DNA segments separated by unique “spacer” sequences.
- Viral DNA sequences are incorporated into the spacer regions, becoming a database of previous viral infections.



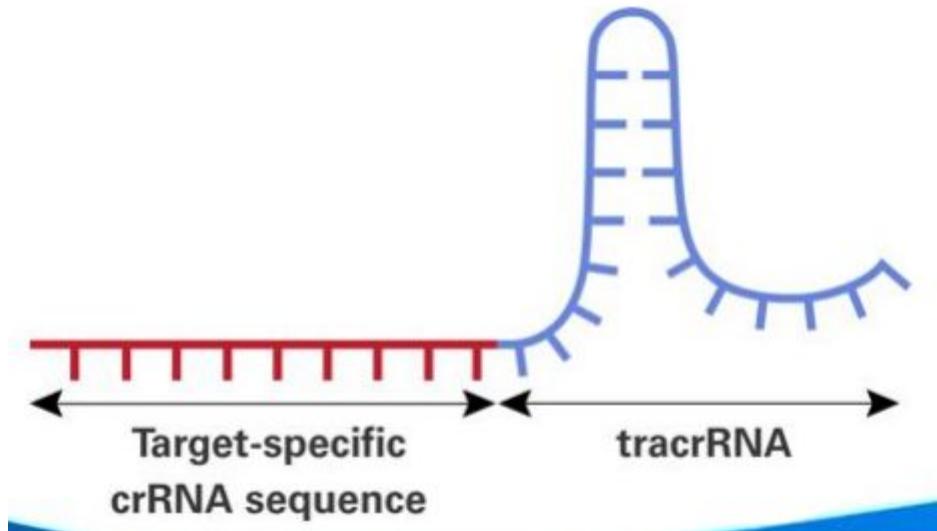
# What is CRISPR?

- The viral DNA from the ‘database’ are transcribed into short pieces of RNA called “CRISPR RNAs” or “crRNA”
- The crRNAs binds to the complementary sequence if viral DNA is present in the bacterial cell
- Once bound to the target, Cas digests the DNA



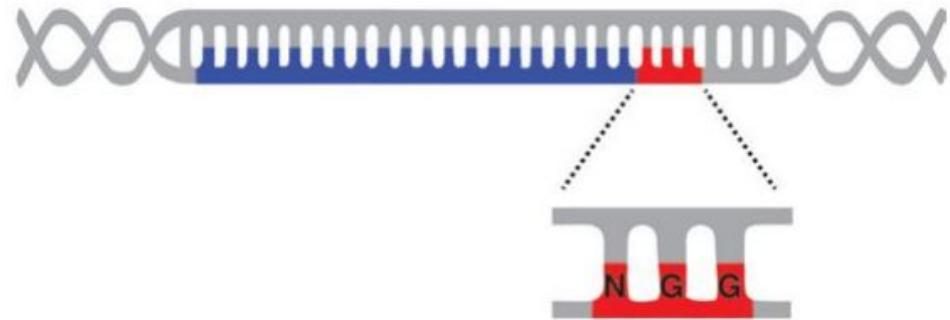
## Anatomy of the guide RNA

- The CRISPR guide RNA (gRNA) complex has two pieces: crispr RNA and tracrRNA
- crispr RNA (or crRNA) contains the 17-20nt sequence that is complementary to the target DNA
- Trans-activating crRNA (or tracrRNA) folds to take a secondary structure that recruits Cas
- “Single guide RNAs” (sgRNA) combine the crRNA and tracrRNA into a single hybrid RNA molecule



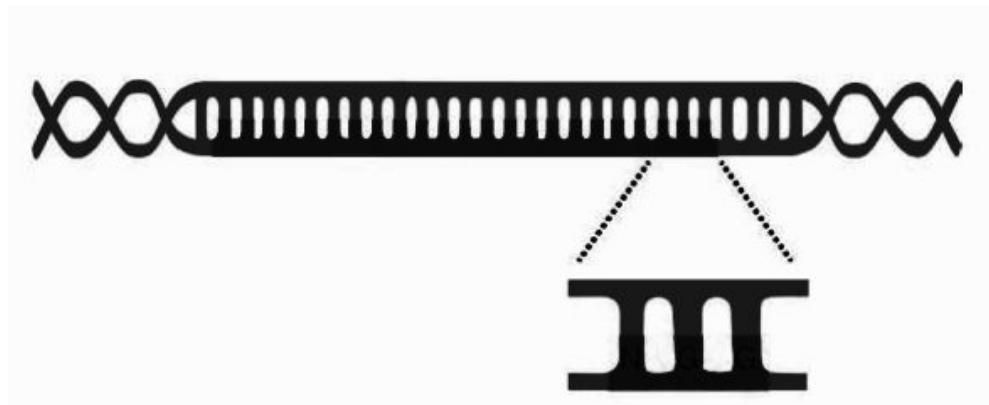
# Protospacer Adjacent Motif (PAM) Sequences

- PAM sequences are short DNA sequences that are recognized by specific Cas proteins.
- Each Cas enzyme recognizes a different PAM sequence
- The Cas9 the PAM is 5'-NGG-3', where "N" is any nucleotide



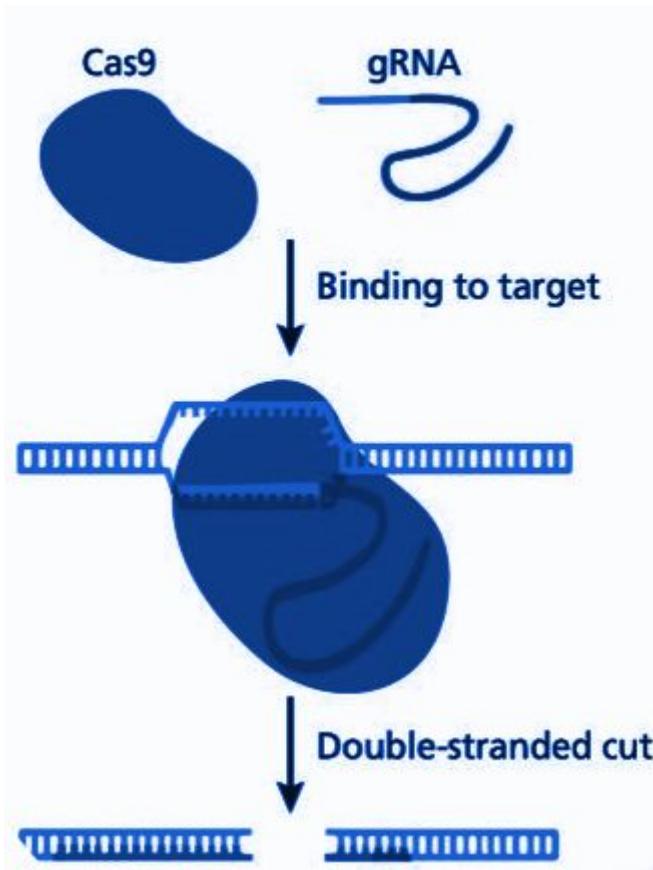
## Designing a sequence-specific gRNA

- The crRNA are designed using two criteria:
- There must be a PAM site immediately downstream (in the 3' direction) of the target DNA sequence.
- The sequence must be complementary to the target DNA sequence



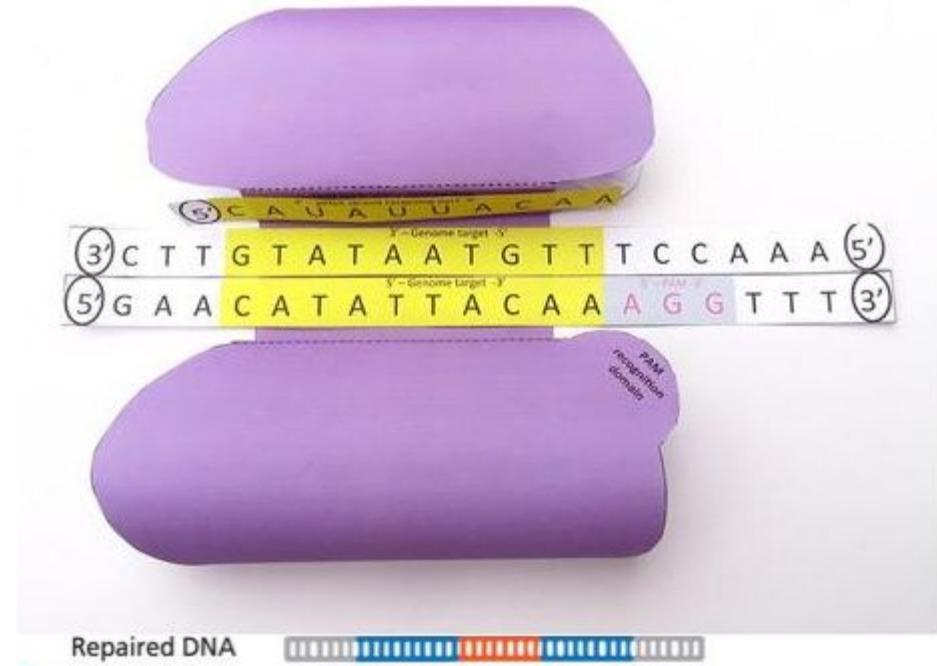
# Using CRISPR for gene knockout

- The CRISPR-Cas9 complex targets double-stranded DNA based on sequence
- Cas9 cuts through the strands of DNA, creating a double-stranded break in the target.
- The cell's DNA repair machinery fixes the break
- Non-homologous end joining (NHEJ) is highly efficient, but error prone
- This makes it great for knocking out a gene!



# Using CRISPR in Genetic Engineering

- Homology Directed Repair (HDR) allows scientists to create and provide a template for DNA repair at the double-stranded break.
- The “correct” DNA sequence is flanked by homologous DNA sequences.
- CRISPR-Cas9 digested DNA is then repaired using the template to introduce the correct sequence.



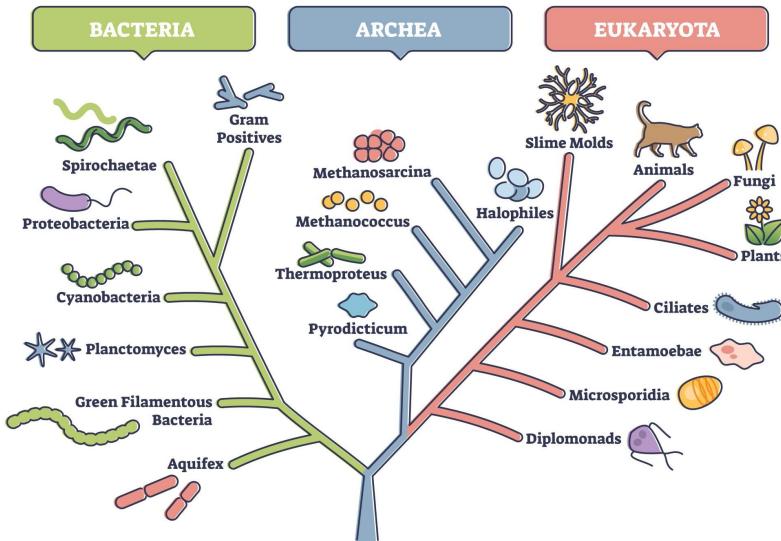
# What Is Phylogeny?

Phylogeny is the study of relationships among different groups of organisms and their evolutionary development. Phylogeny attempts to trace the evolutionary history of all life on the planet. It is based on the phylogenetic hypothesis that all living organisms share a common ancestry. The relationships among organisms are depicted in what is known as a phylogenetic tree. Relationships are determined by shared characteristics, as indicated through the comparison of genetic and anatomical similarities.

# What is Phylogeny?

- **Definition:** Phylogeny is the study of the evolutionary relationships among a group of organisms.
- **Analogy:** Think of it as creating a "family tree" for different species to see how they are related.
- These relationships are often visualized in a diagram called a phylogenetic tree.

## PHYLOGENETIC TREE



## The Challenge: Reconstructing the Past

- **Problem:** Most evolution happens over millions of years and cannot be observed directly.
- **Solution:** Biologists must infer or deduce these relationships by studying organisms that are alive today.
- This process is like being a detective, using clues from the present to reconstruct events from the past.

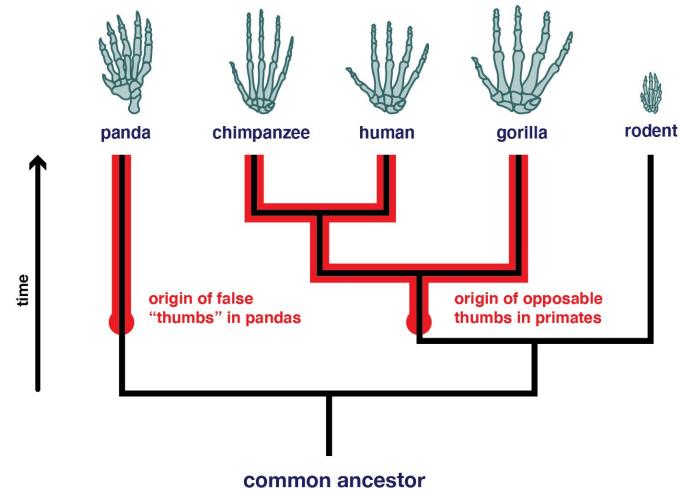
# Sources of Evidence for Phylogeny

Scientists use several types of data to build these evolutionary trees.

- **Fossil Record:**
  - **Pro:** Can provide direct evidence of ancestral organisms.
  - **Con:** The fossil record is often **incomplete** or poor, leaving many gaps.
- **Phenotypic Characteristics (The "Classic" Method):**
  - Based on observable traits, especially **anatomical structures**.
  - Example: Comparing the bone structure of different animals.
- **Molecular Data (The "Modern" Method):**
  - This is the most common method used today.
  - Involves comparing **protein** and **DNA sequences**.
  - The more similar the DNA or protein sequences are between two organisms, the more closely related they are considered to be.

# The Core Concept: Homologous Characteristics

- **Definition:** Phylogenies are reconstructed by analyzing **homologous characteristics**. These are traits shared by related species because they have been inherited from a **common ancestor**.
- **Key Idea:** The presence of homologous features is strong evidence of a shared evolutionary history.
- 



# Examples of Homology

## Anatomical Homology:

- **Classic Example:** The front leg of a mouse and the wing of a bat.
- **Analysis:**
  - **Different Function:** One is for walking, the other for flying.
  - **Different Appearance:** They look quite different on the surface.
  - **Shared Origin:** Despite these differences, their underlying bone structure and development show they both evolved from the **forelimb of an early mammal ancestor**. This shared ancestry is what makes them homologous.

## Molecular Homology:

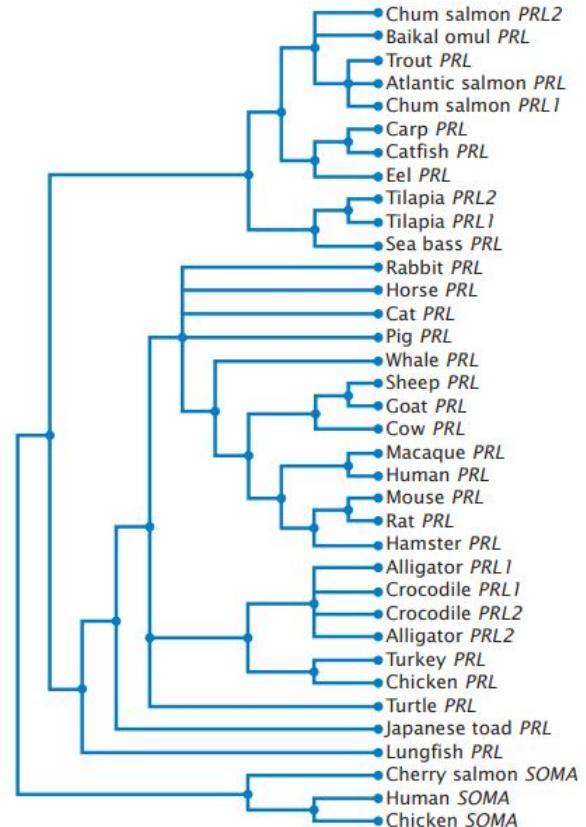
- The same principle applies to genes.
- **Example:** Two DNA sequences in different species are considered homologous if they both evolved from a single ancestral DNA sequence.

# Visualizing Phylogeny: The Phylogenetic Tree

**Definition:** The graphical representation of a phylogeny is called a **phylogenetic tree**. It is a diagram that visually depicts the evolutionary relationships among different organisms.

## Analogy to a Family Pedigree:

- A **pedigree** chart shows the genealogical relationships among family members (e.g., parents, children, cousins).
- Similarly, a **phylogenetic tree** shows the evolutionary relationships among species, indicating how they branched off from common ancestors over millions of years.



**26.14** A gene tree can be used to represent the evolutionary relationships among a group of genes. This gene tree is a rooted tree, in which *PRL* represents a prolactin gene; *PRL1* and *PRL2* are two different prolactin genes found in the same organism; and *SOMA* represents a somatotropin gene, which is related to prolactin genes. [After M. P. Simmons and J. V. Freudestein, Uninode coding vs. gene tree parsimony for phylogenetic reconstruction using duplicate genes, *Molecular Phylogenetics and Evolution* 23:488, 2002.]

## Building a Tree with DNA: Sequence Alignment

To build a tree from molecular data, we can't just look at the DNA; we have to compare it very carefully. This is done through a process called **sequence alignment**.

- **First Step:** The process begins by identifying **homologous genes** in the different organisms being compared.
- **The Goal:** The main goal of alignment is to line up the nucleotide sequences of these genes to see where they are similar and where they have changed.

# An Example of Alignment

Let's consider a homologous gene (Gene X) from two different species:

## Species A

5'-A T T G C G A A-3'

## Species B

5'-A T G C C A A C-3'

### Scenario 1: Assume only Base Substitutions

Here, we assume four individual mutations occurred.

Gene X (A): A T **T G C G A A** > Gene X (B): A T **G C C A A C**

This alignment suggests **4 evolutionary changes** (four base substitutions).

### Scenario 2: Assume Insertions/Deletions (Indels)

A more likely scenario might involve a single nucleotide being deleted from one sequence (or inserted into the other), plus one substitution. A gap, represented by a dash (-), is used to show this.

Gene X (A): A T T G C **G A A**

Gene X (B): A T - G C **C A A C**

This alignment suggests only **2 evolutionary changes** (one deletion/insertion and one base substitution).

## The Principle of Simplicity (Parsimony)

- Biologists often favor the second alignment because it requires **fewer evolutionary steps**. This idea is based on a principle called **parsimony**—the simplest explanation is often the correct one.
- **Computer Programs:** Because aligning even short sequences is complex, this work is done by powerful computer programs that can test many possible alignments to find the best fit.

Alignment is the critical foundation for constructing a phylogenetic tree. Once the alignment is set, scientists can quantify the differences and use that data to build the tree.

## From Changes to "Distance": Quantifying Evolution

The aligned sequences provide the raw data to calculate the evolutionary distance between species. The basic principle is simple: **the more differences there are between two sequences, the more distant their evolutionary relationship is.**

- **Counting the Differences:** Let's use our last example, which had 2 evolutionary changes between Species A and Species B. This number, "2," is a simple, direct measure of their genetic distance.
- **Creating a Distance Matrix:** Scientists do this for every pair of species in their study. The results are organized into a distance matrix, which is essentially a table that lists the "distance" score between each species.

## From Changes to "Distance": Quantifying Evolution

The aligned sequences provide the raw data to calculate the evolutionary distance between species.

The basic principle is simple: **the more differences there are between two sequences, the more distant their evolutionary relationship is.**

- **Counting the Differences:** Let's use our last example, which had 2 evolutionary changes between Species A and Species B. This number, "2," is a simple, direct measure of their genetic distance.
- **Creating a Distance Matrix:** Scientists do this for every pair of species in their study. The results are organized into a distance matrix, which is essentially a table that lists the "distance" score between each species.

	Species A	Species B	Species C
Species A	0	2	5
Species B	2	0	
Species C	5	2	0

# Synonymous vs. Nonsynonymous Substitutions

- **Nonsynonymous Substitutions:**

- Definition:** These are nucleotide changes in a gene that **alter the amino acid sequence** of the corresponding protein.
- Impact:** They create a different protein, which can affect the organism's traits.
- Rate of Change:** The rate varies dramatically between different genes. For example, the gene for **interferon γ** (part of the immune system) evolves almost 1000 times faster than the gene for **α-actin** (a structural protein).

- **Synonymous Substitutions:**

- Definition:** These are nucleotide changes that **do not alter the amino acid sequence** of the protein. This is possible because the genetic code is redundant (multiple codons can code for the same amino acid), and these changes often happen in the third position of a codon.
- Impact:** They are often called "silent" mutations because the protein remains the same.

**Table 26.5**

Rates of nonsynonymous and synonymous substitutions in mammalian genes based on human–rodent comparisons

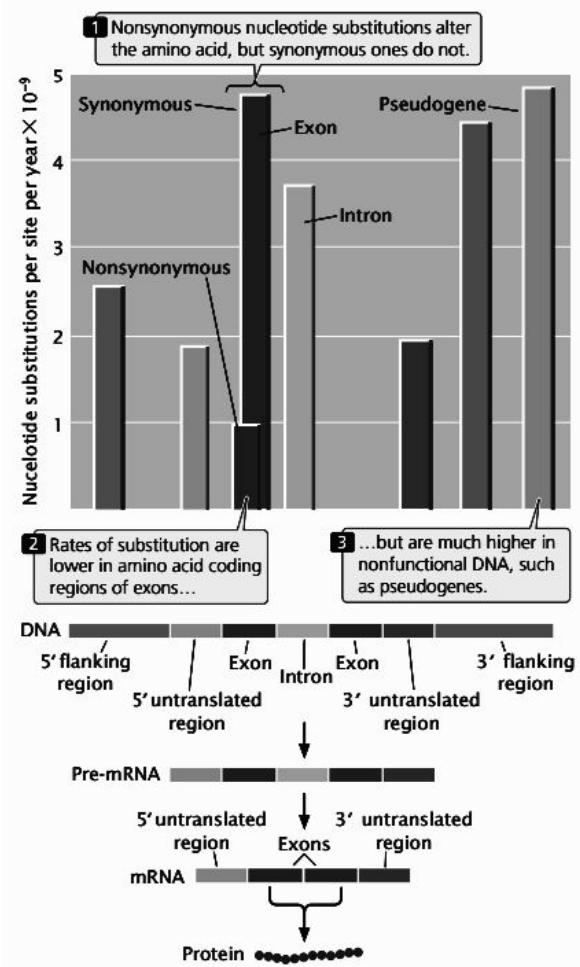
Gene	Nonsynonymous Rate (per site per 10 <sup>9</sup> years)	Synonymous Rate (per site per 10 <sup>9</sup> years)
α-Actin	0.01	3.68
β-Actin	0.03	3.13
Albumin	0.91	6.63
Aldolase A	0.07	3.59
Apoprotein E	0.98	4.04
Creatine kinase	0.15	3.08
Erythropoietin	0.72	4.34
α-Globin	0.55	5.14
β-Globin	0.80	3.05
Growth hormone	1.23	4.95
Histone 3	0.00	6.38
Immunoglobulin heavy chain (variable region)	1.07	5.66
Insulin	0.13	4.02
Interferon α 1	1.41	3.53
Interferon γ	2.79	8.59
Luteinizing hormone	1.02	3.29
Somatostatin-28	0.00	3.97

Source: After W. Li and D. Graur, *Fundamentals of Molecular Evolution* (Sunderland, Mass.: Sinauer, 1991), p. 69, Table 1.

# Synonymous vs. Nonsynonymous Substitutions

- This leads us to a fundamental rule in molecular evolution:  
Different parts of a gene evolve at different rates.
- The guiding principle is that the highest rates of nucleotide substitution are found in sequences that have the least effect on the organism's function.

Functional Importance	Level of Constraint	Rate of Evolution	Example
High	Strong	Slow	Nonsynonymous sites in a critical gene (e.g., actin)
Low	Weak	Fast	Synonymous ("silent") sites, Introns



*This principle is why scientists can use different genes to answer different evolutionary questions. A fast-evolving gene is useful for comparing closely related species, while a slow-evolving, highly conserved gene is used to study ancient, deep evolutionary relationships.*

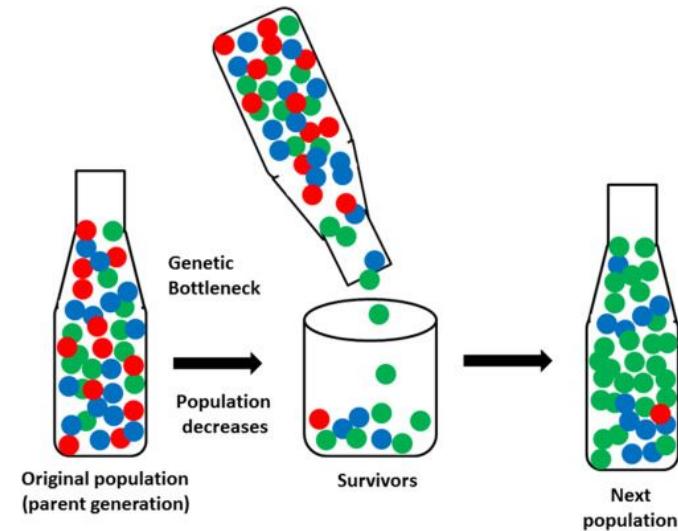
# Population Genetics

# Population Genetics



*Rocky Mountain bighorn sheep (*Ovis canadensis*)*

## Genetic Drift



## Genotypic and Allelic Frequencies Are Used to Describe the Gene Pool of a Population

- An obvious and pervasive feature of life is variability.
- Genetic Variation is more complex wrt to Phenotypic Variation e.g. two members of a population can produce the same protein even if their DNA sequences are different. DNA sequences between the genes and introns within genes do not encode proteins; much of the variation in these sequences also has little effect on the phenotype.



# Calculating Genotypic Frequencies

- A frequency is simply a proportion or a percentage, usually expressed as a decimal fraction. For example, if 20% of the alleles at a particular locus in a population are A, we would say that the frequency of the A allele in the population is 0.20.
- For large populations, for which a determination of the genes of all individual members is impractical, a sample of the population is usually taken and the genotypic and allelic frequencies are calculated for this sample
- The genotypic and allelic frequencies of the sample are then used to represent the gene pool of the population

## Calculating Genotypic Frequencies

- Genotypic frequency, we simply add up the number of individuals possessing the genotype and divide by the total number of individuals in the sample (N).
- For a locus with three genotypes AA, Aa, and aa, the frequency (f ) of each genotype is

$$f(AA) = \frac{\text{number of } AA \text{ individuals}}{N}$$

$$f(Aa) = \frac{\text{number of } Aa \text{ individuals}}{N}$$

$$f(aa) = \frac{\text{number of } aa \text{ individuals}}{N}$$

**The sum of all the genotypic frequencies always equals 1.**

# Calculating Allelic Frequencies

- The gene pool of a population can also be described in terms of the allelic frequencies.
- There are always fewer alleles than genotypes; so the gene pool of a population can be described in fewer terms when the allelic frequencies are used.

**Allelic frequencies can be calculated from (1) the numbers or (2) the frequencies of the genotypes.**

$$\text{frequency of an allele} = \frac{\text{number of copies of the allele}}{\text{number of copies of all alleles at the locus}}$$

For a locus with only two alleles (A and a), the frequencies of the alleles are usually represented by the symbols p and q, and can be calculated as follows:

$$p = f(A) = \frac{2n_{AA} + n_{Aa}}{2N}$$

$$q = f(a) = \frac{2n_{aa} + n_{Aa}}{2N}$$

where  $n_{AA}$ ,  $n_{Aa}$ , and  $n_{aa}$  represent the numbers of AA, Aa, and aa individuals, and N represents the total number of individuals in the sample. We divide by 2N because each diploid individual has two alleles at a locus. The sum of the allelic frequencies always equals 1 ( $p + q = 1$ ); so, after p has been obtained, q can be determined by subtraction:  $q = 1 - p$ .

## **Allelic frequencies can be calculated from the genotypic frequencies**

To calculate an allelic frequency from genotypic frequencies, we add the frequency of the homozygote for each allele to half the frequency of the heterozygote (because half of the heterozygote's alleles are of each type):

$$p = f(A) = f(AA) + \frac{1}{2}f(Aa)$$

$$q = f(a) = f(aa) + \frac{1}{2}f(Aa)$$

**We obtain the same values of p and q whether we calculate the allelic frequencies from the numbers of genotypes or from the genotypic frequencies.**

## **The Hardy–Weinberg Law Describes the Effect of Reproduction on Genotypic and Allelic Frequencies**

The primary goal of population genetics is to understand the processes that shape a population's gene pool.

First, we must ask what effects reproduction and Mendelian principles have on the genotypic and allelic frequencies: How do the segregation of alleles in gamete formation and the combining of alleles in fertilization influence the gene pool ?

The law is actually a mathematical model that evaluates the effect of reproduction on the genotypic and allelic frequencies of a population. It makes several simplifying assumptions about the population and provides two key predictions if these assumptions are met.

**Assumptions** If a population is large, randomly mating, and not affected by mutation, migration, or natural selection, then:

- **Prediction 1** the allelic frequencies of a population do not change; and
- **Prediction 2** the genotypic frequencies stabilize (will not change) after one generation in the proportions  $p^2$  (the frequency of AA),  $2pq$  (the frequency of Aa), and  $q^2$  (the frequency of aa), where p equals the frequency of allele A and q equals the frequency of allele a.

**The Hardy–Weinberg law indicates that, when the assumptions are met, reproduction alone does not alter allelic or genotypic frequencies and the allelic frequencies determine the frequencies of genotypes.**

		Sperm	
		$A$ $p$	$a$ $q$
Eggs	$A$ $p$	$AA$ $p \times p = p^2$	$Aa$ $q \times p = pq$
	$a$ $q$	$Aa$ $p \times q = pq$	$aa$ $q \times q = q^2$
$f(A) = p$ $f(a) = q$  $f(AA) = p^2$ $f(Aa) = 2pq$ $f(aa) = q^2$			

**Conclusion:** Random mating will produce genotypes of the next generation in proportions  $p^2(AA)$ ,  $2pq(Aa)$ , and  $q^2(aa)$

**25.2 Random mating produces genotypes in the proportions  $p^2$ ,  $2pq$ , and  $q^2$ .**

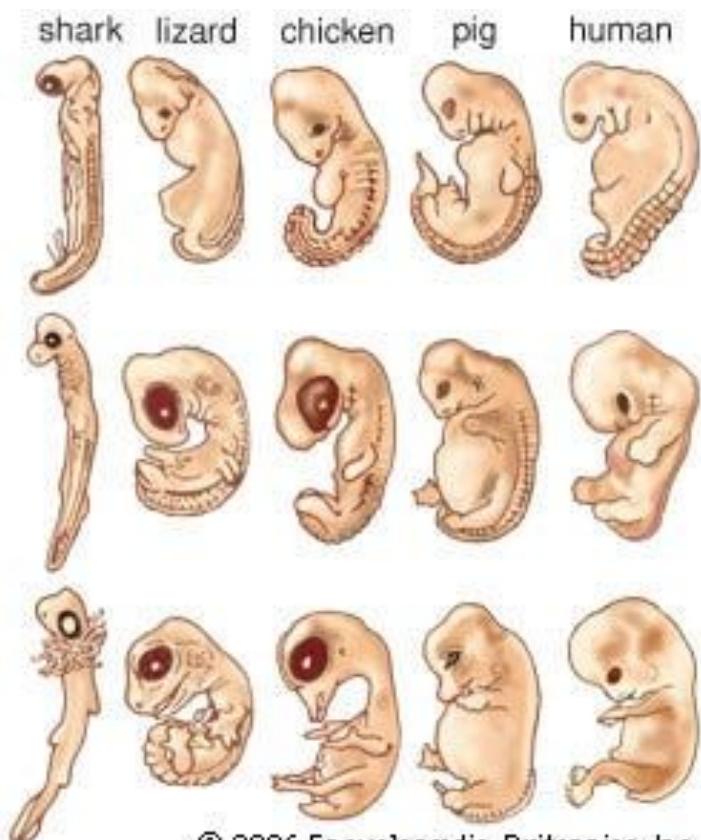
# What is Evo-Devo?

Evolutionary Developmental Biology (Evo-Devo) investigates the interplay between evolutionary processes and developmental mechanisms.

It asks two main questions:

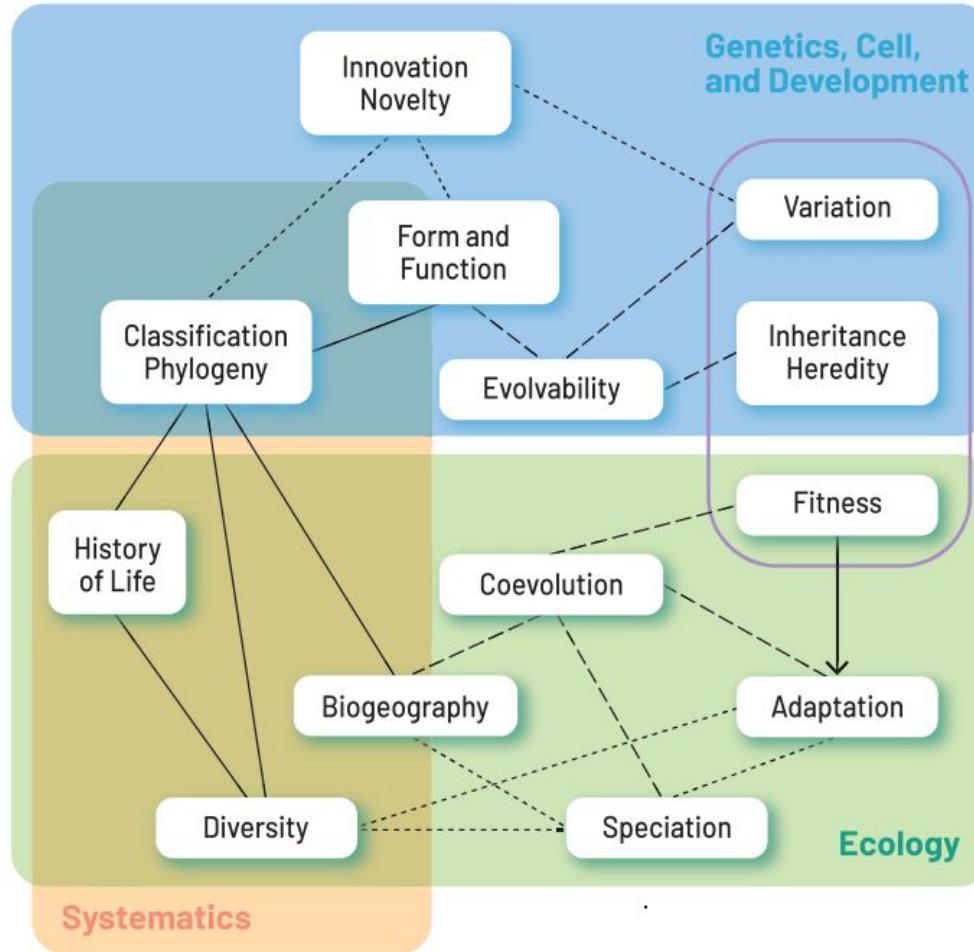
- The Evolution of Development: How do developmental processes themselves evolve and change over time?
- The Developmental Basis of Evolution: How do the processes of development (embryology) influence the path and possibilities of evolution?

It emerged as a field to address the neglect of development in the "Modern Synthesis" of evolutionary theory, which focused heavily on population genetics.

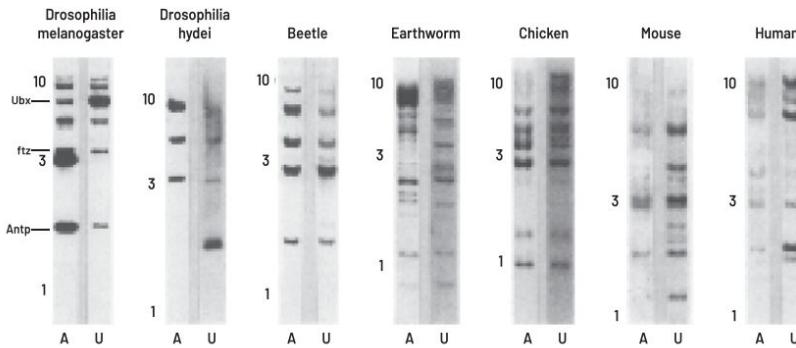


© 2006 Encyclopædia Britannica, Inc.

# What is Evo-Devo?



# The "Blot Seen Round the World"



**Figure 1** Conservation of the homeobox DNA sequence across metazoans. There are duplicate genomic blots for each species with two different probes containing the ~180 base pair (bp) homeobox sequence: A = 600 bp fragment from the *Antennapedia* homeobox gene of *Drosophila melanogaster* (the fruit fly); U = 450 bp fragment from the *Ultrabithorax* homeobox gene of *Drosophila melanogaster*. Radiolabeled hybridization fragments indicate a complementary DNA sequence and therefore the presence of the homeobox sequence in other species. Ten, three, and one kilobase labels are migration distance size standards. Abbreviations: Ubx: *Ultrabithorax*; ftz: *fushi tarazu*; Antp: *Antennapedia*.

Adapted from: McGinnis et al. (1984). Reproduced with permission from Elsevier.

A key moment in Evo-Devo's history was the discovery of the homeobox in the 1980s.

The homeobox is a ~180 base-pair DNA sequence that codes for a part of proteins (transcription factors) that control development.

This sequence was found to be highly conserved in animals as different as flies, mice, and humans.

This revealed a stunning fact: Diverse animals are built using a common, ancient "genetic toolkit."

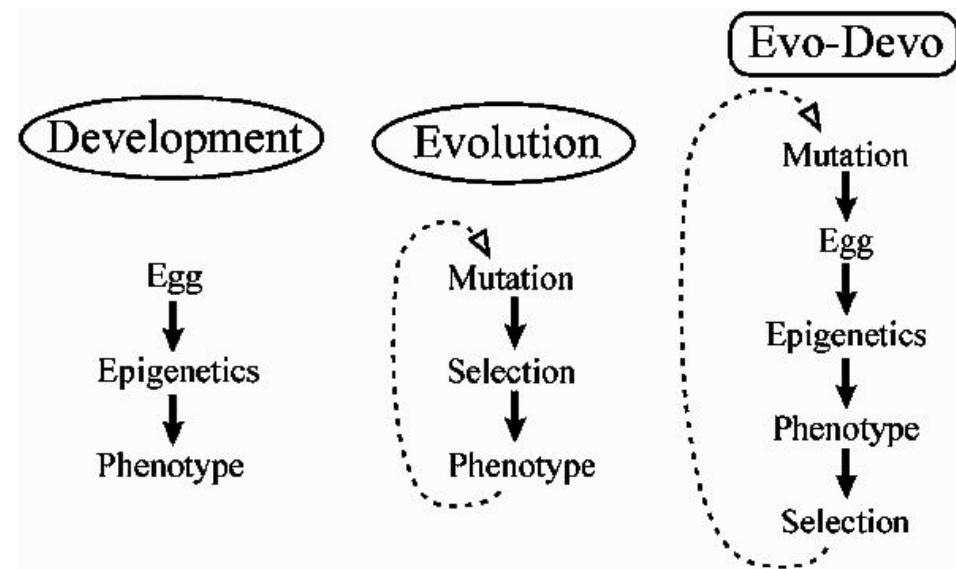
# Where Does Evo-Devo Fit in Evolutionary Theory?

Evo-Devo complements traditional evolutionary biology.

Traditional Focus (e.g., Population Genetics):

Explains the fate of variation in populations (e.g., via natural selection).

Evo-Devo Focus: Explains the origin and introduction of phenotypic variation through changes in development.



It asks: How do the biases and constraints of development channel evolution in certain directions?

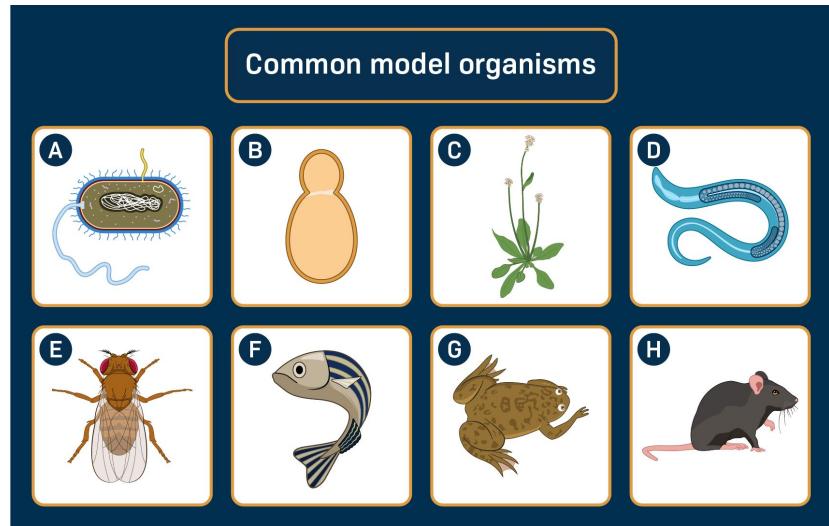
# Model Organisms in Evo-Devo

Scientists use model organisms to study development in depth.

They are chosen for:

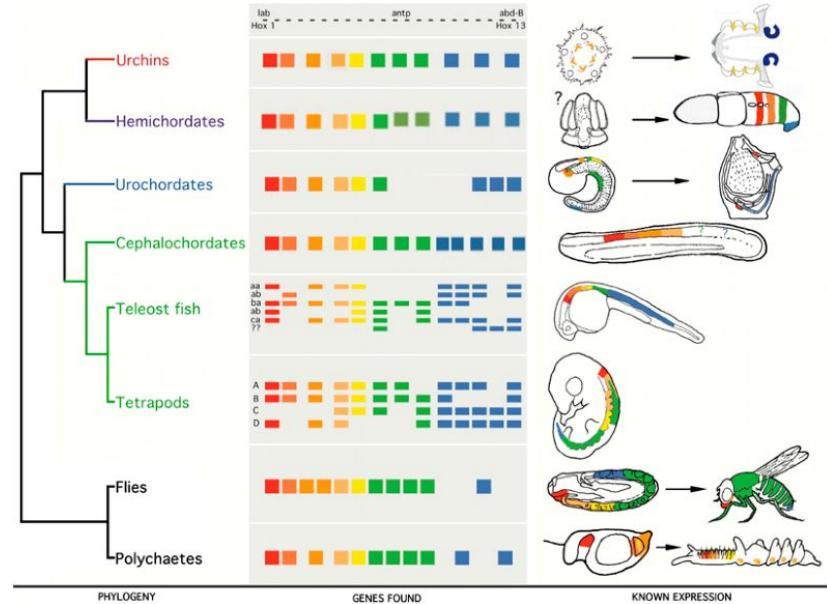
- Representation: How well they represent a broader group (e.g., mice for mammals).
- Manipulation: Ease of study (short generation time, available genetic tools).

Evo-Devo often uses "model systems" like the starlet sea anemone, chosen not just for convenience but for their key phylogenetic position to answer evolutionary questions.



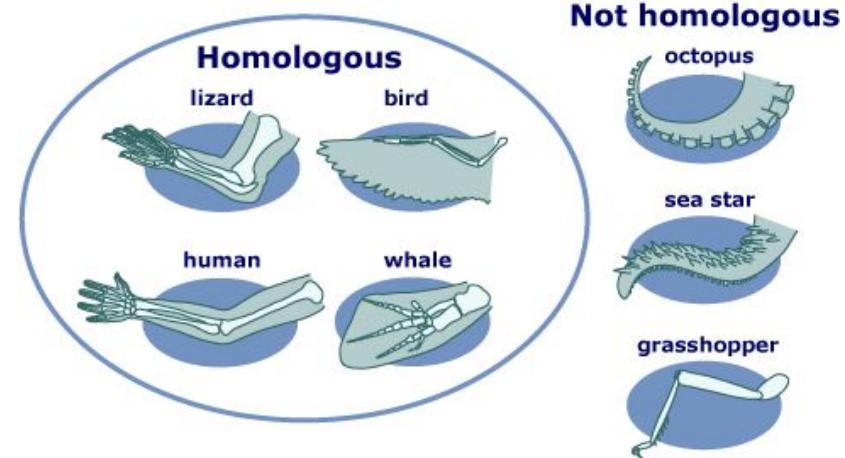
# The Puzzle of "Conserved Mechanisms"

- A core concept in Evo-Devo is the "conserved mechanism."
- These are organized sets of genes and proteins (e.g., signaling pathways like Wnt or Hedgehog) that perform a specific developmental task across many species.
- The Conceptual Puzzle: Homology (sameness due to common descent) is usually about structure, but mechanisms are defined by their function.
- Solution: A conserved mechanism is a shared, derived trait with specific parts, organization, and context that produces a stereotypical outcome, allowing for some variation.



# What is an Evolutionary Novelty?

- A central problem for Evo-Devo is explaining the origin of evolutionary novelties.
- A novelty is a new anatomical structure that is not homologous to any structure in the ancestral species.
- Examples: The turtle shell, insect wings, or the tetrapod limb.
- Explaining novelties requires more than just natural selection; it requires understanding how development generates new forms in the first place.

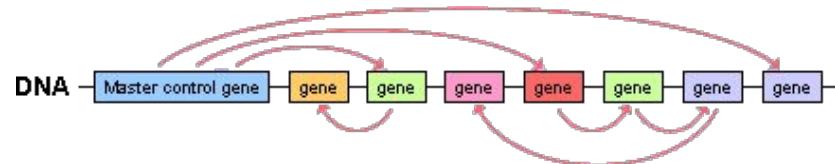


# How Do Novelties Evolve?

- Novelties rarely come from brand-new genes. Instead, they arise from tinkering with existing developmental tools.
- Key mechanisms include:
  - Co-option: The re-use of an existing gene or genetic network in a new place or time during development.
  - Changes in Gene Regulation: Altering how genes are switched on and off (e.g., via *cis*-regulatory elements), rather than changing the protein itself.
  - Developmental Plasticity: The environment triggers a new phenotype, which can later become genetically fixed.

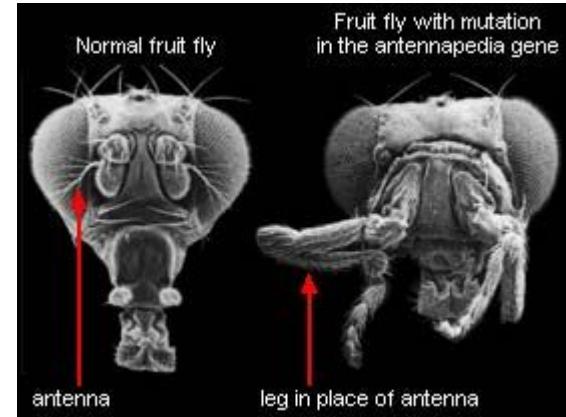
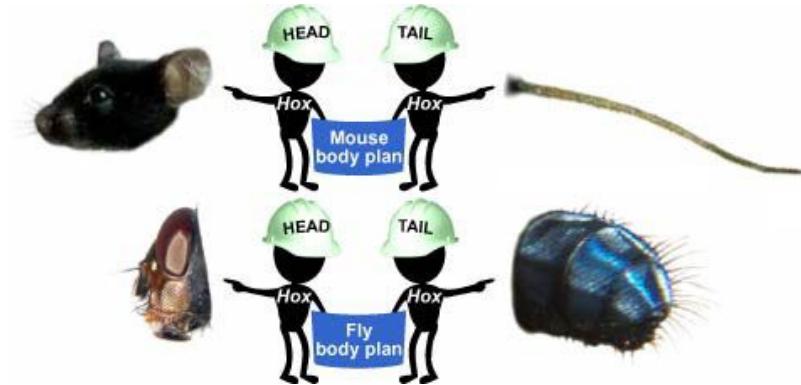
# The Genetic Toolkit: How a Single Egg Builds a Complex Body

- Key Concepts: The Blueprint for Complexity
  - The Central Mystery: A single fertilized egg contains a full set of DNA instructions. How do different cells use this same blueprint to become eyes, wings, or legs?
  - The Answer: Gene Regulation. Not all genes are active in all cells at all times. Regulatory genes act as master switches, controlling when and where other genes are turned on.
  - Chain Reactions: A single regulatory gene can trigger a cascade of genetic activity, turning on sets of genes that ultimately build complex structures like an entire leg or wing.

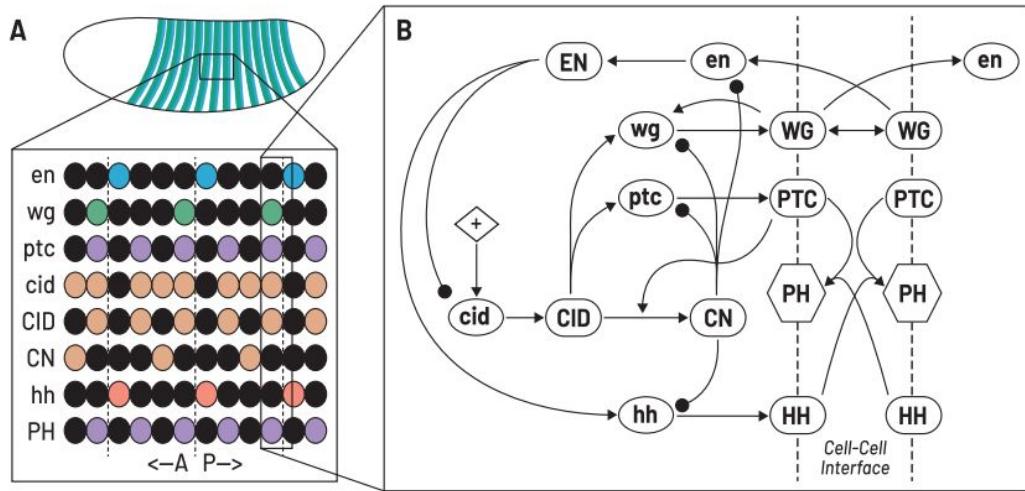
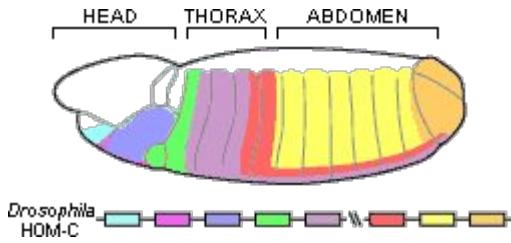


# The Genetic Toolkit: How a Single Egg Builds a Complex Body

- Hox genes are a famous set of regulatory genes that act as "master architects."
- They define the body plan from head to tail, providing instructions like "Put the head here! Legs go there!"
- Evidence from Mutations: Errors in Hox genes can cause dramatic changes, like growing a leg where an antenna should be, proving their powerful role in development.



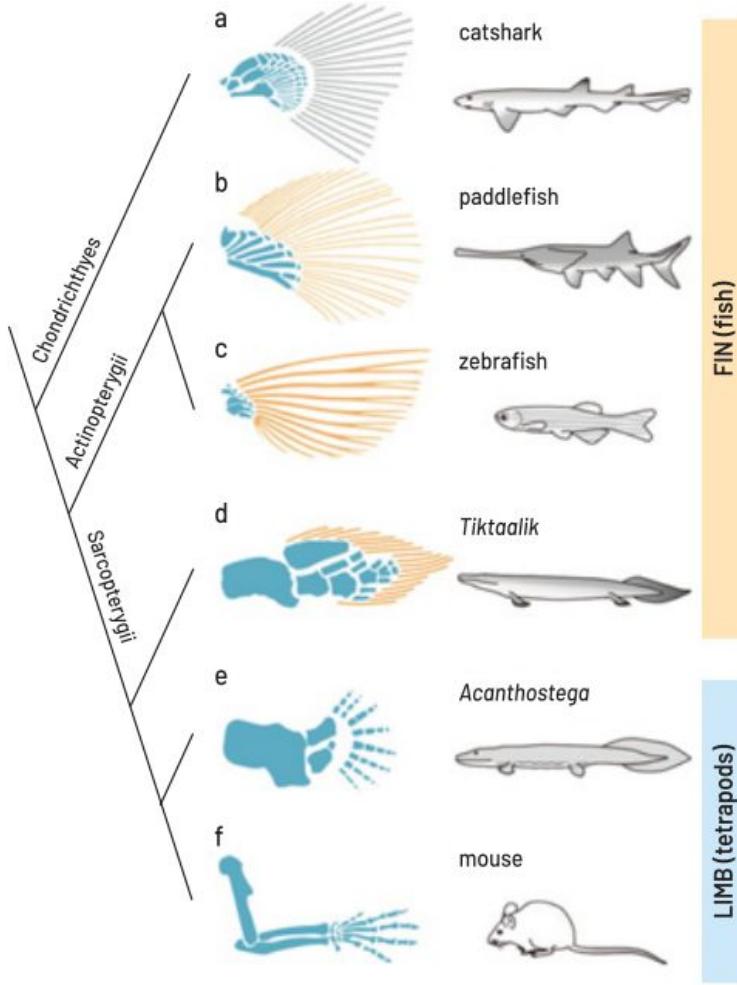
# The Genetic Toolkit: How a Single Egg Builds a Complex Body



**Figure 9** Segment polarity network details. (A) Cartoon representation of a *Drosophila* embryo with an inset of gene expression patterns in cells of the parasegments. Boundaries between parasegments (dashed vertical lines) are defined by the expression of *wingless* (*wg*) to the anterior and *engrailed* (*en*) to the posterior (see Figure 3.1 for more detail). (B) Inset depicting a cell–cell interface at the parasegment boundary, which shows the complex nature of interactions occurring in the basic segment polarity network. Abbreviations: CID/cid: cubitus interruptus (whole protein); CN: cubitus interruptus (*N*-terminal repressor); EN/en: engrailed; HH/hh: hedgehog; PH: patched-hedgehog complex; PTC/ptc: patched; Wg/wg: wingless. Redrawn.

# A Universal Toolkit: From Flies to Humans

- Key Concept: Deep Homology and Evolutionary Change
  - A Shared Inheritance: The same families of master control genes (like Hox genes) are found in insects, mice, and humans.
  - We All Use the Same Tools: This means a common ancestor, over 600 million years ago, evolved a core "genetic toolkit" for building bodies.
  - Proof of Commonality: A mouse version of a fly's eye-building gene can be inserted into a fly and still trigger the formation of a (fly) eye!
  - The Engine for Evolution: Evolution works by tinkering with these ancient, conserved toolkits. Small changes in these powerful regulatory genes can lead to major changes in body form.
- How Specific Cell Fates Are Determined
- Differential Gene Expression: Different cell types (e.g., eye cells vs. gut cells) are different because they express different sets of genes from the shared DNA blueprint.
- The Role of Signals: Chemical signals from a cell's environment and neighboring cells provide crucial instructions, telling it which genes to turn on to become a specific cell type (e.g., a lens cell in the eye).



**Figure 12** Phylogenetic representation of the comparative anatomy relevant to the fin–limb transition. *Tiktaalik* and *Acanthostega* are extinct taxa.

# Immunology and Host

# The Genetic Exception: Immune System Gene Rearrangement

## Core Concept: Controlled Genetic Alteration

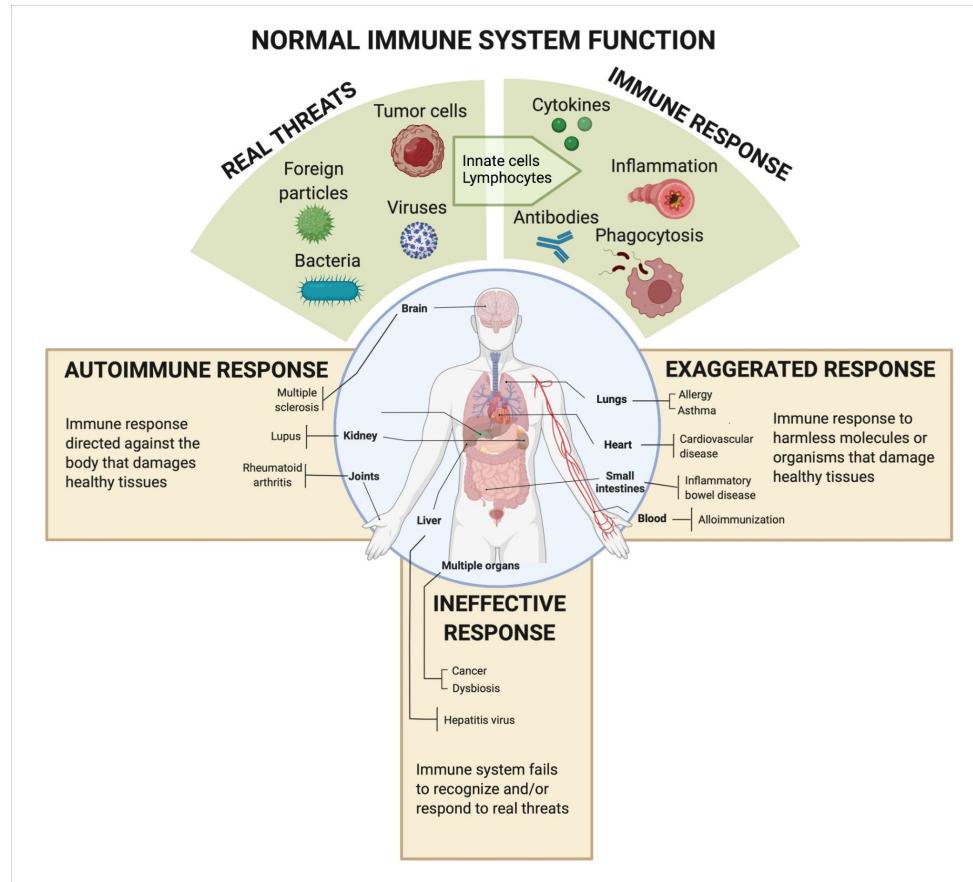
- A fundamental rule of biology is that all somatic cells have identical DNA.
- The immune system is a crucial exception to this rule.
- To fight a vast number of pathogens, immune cells deliberately rearrange, mutate, and lose specific gene segments.

## How It Works: Generating Diversity

- Process: Somatic DNA recombination shuffles gene segments in developing lymphocytes (B cells and T cells).
- Result: Creates a vast repertoire of unique receptors and antibodies, each capable of recognizing a different antigen (foreign molecule).
- Trade-off: This power comes from permanently altering the genome in these specific cells.

## The Critical Balance: Self vs. Non-Self

- This system must perfectly distinguish the body's own proteins (self-antigens) from foreign ones.
- Failure of this distinction leads to autoimmune disease, where the immune system attacks the body's own tissues.



# Two Branches of Adaptive Immunity

## Humoral Immunity (B Cells)

**Function:** Fights pathogens *outside* of cells in blood & body fluids ("humors").

**Key Player:** B Lymphocytes (mature in Bone marrow).

**Mechanism:** Produces **antibodies** that bind to free-floating antigens, marking them for destruction.

**Targets:** Bacteria, viruses in bloodstream.

## Cellular Immunity (T Cells)

**Function:** Fights infected, cancerous, or foreign cells *directly*.

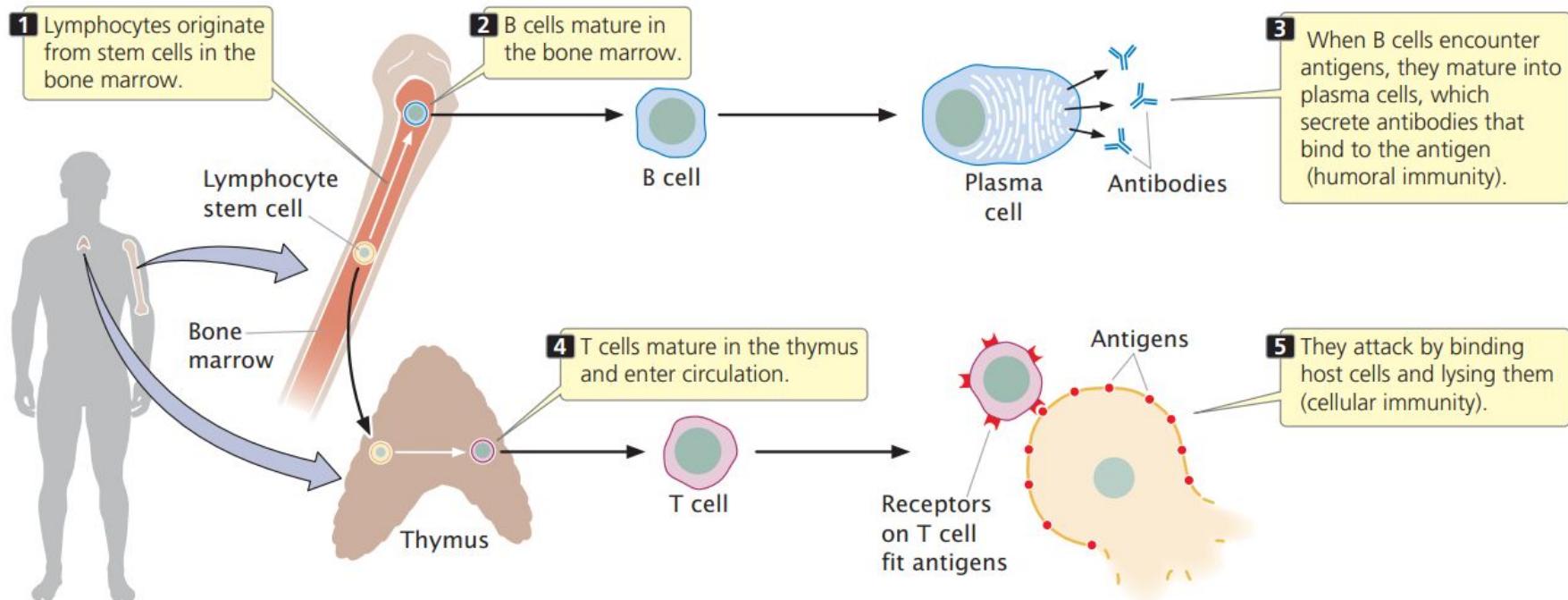
**Key Player:** T Lymphocytes (mature in Thymus).

**Mechanism:** Uses **T-cell receptors** to recognize antigens presented on a cell's surface (with MHC).

**Targets:** Virus-infected cells, cancer cells, foreign tissues.

**Note:** These two systems are distinct but work together and communicate closely.

# Two Branches of Adaptive Immunity



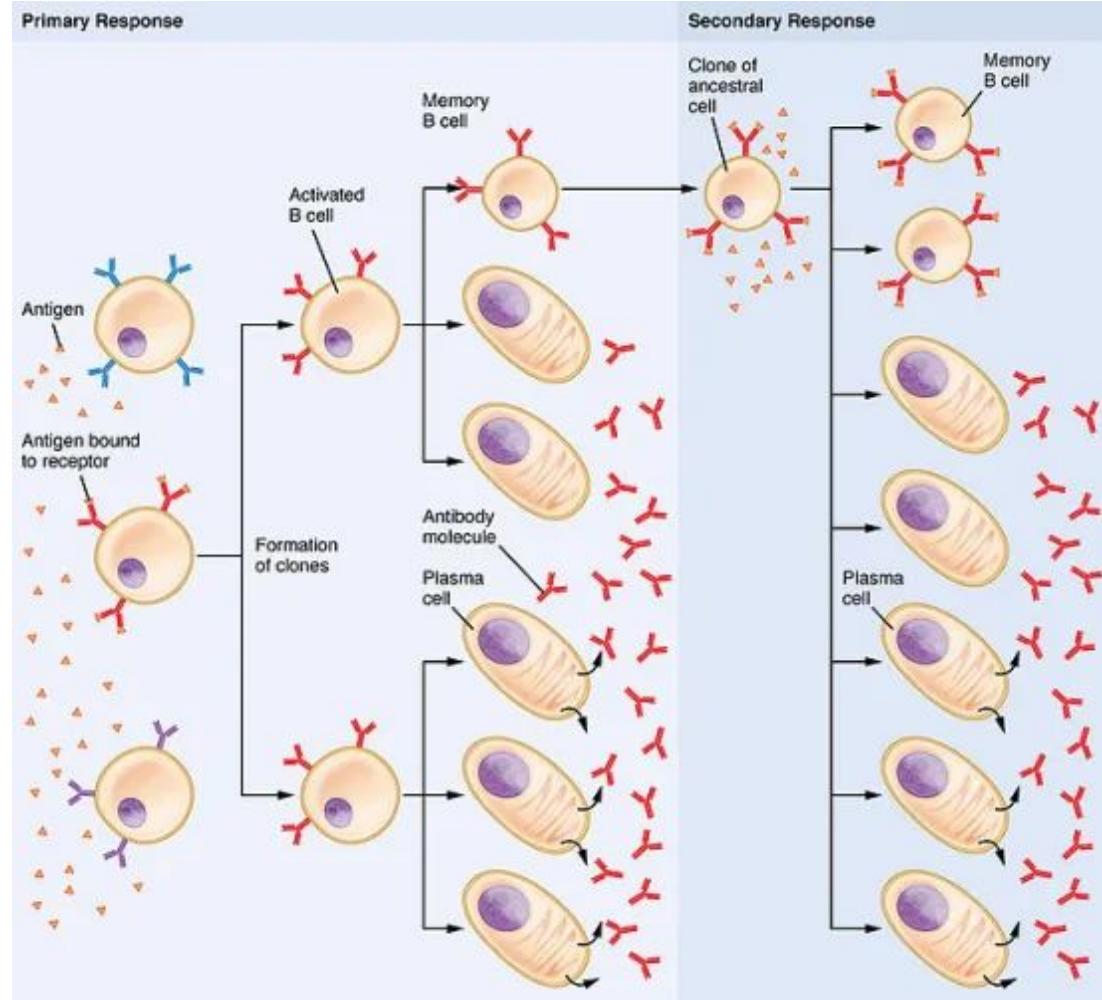
# The Unifying Principle: Clonal Selection

**Immense Pre-existing Diversity:** The body pre-makes a vast pool of lymphocytes, each genetically unique and programmed to recognize one specific antigen.

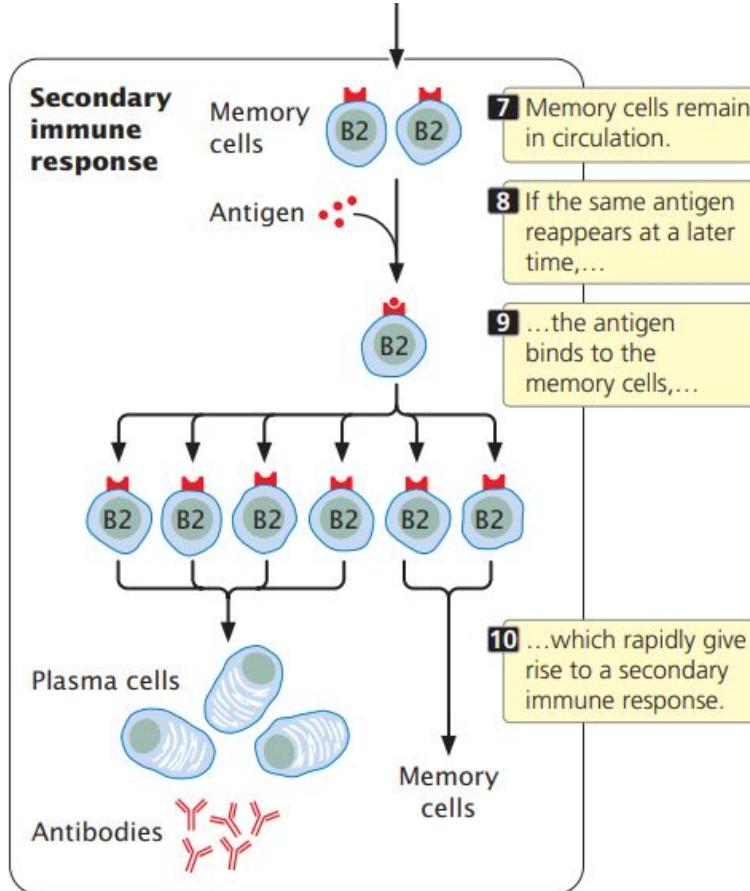
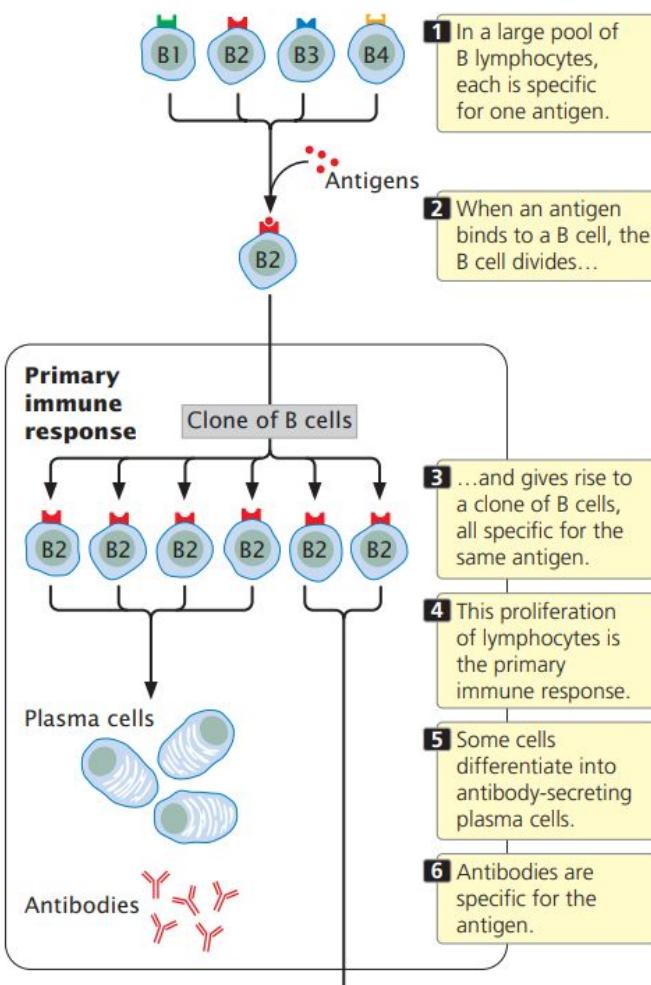
**"Select and Expand":** When an antigen enters the body, it selects the specific lymphocyte that recognizes it.

**Clonal Expansion:** That specific lymphocyte is activated and rapidly divides, creating a large clone of identical cells to attack the invader.

**Immunological Memory:** After the infection is cleared, memory cells remain, providing long-lasting immunity and a faster, stronger response upon re-exposure (the basis for vaccination).

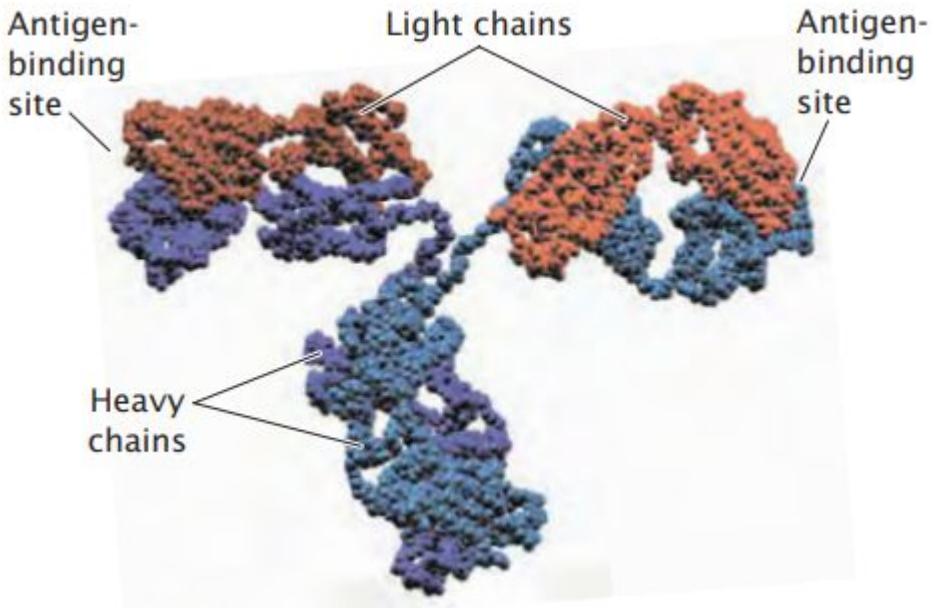
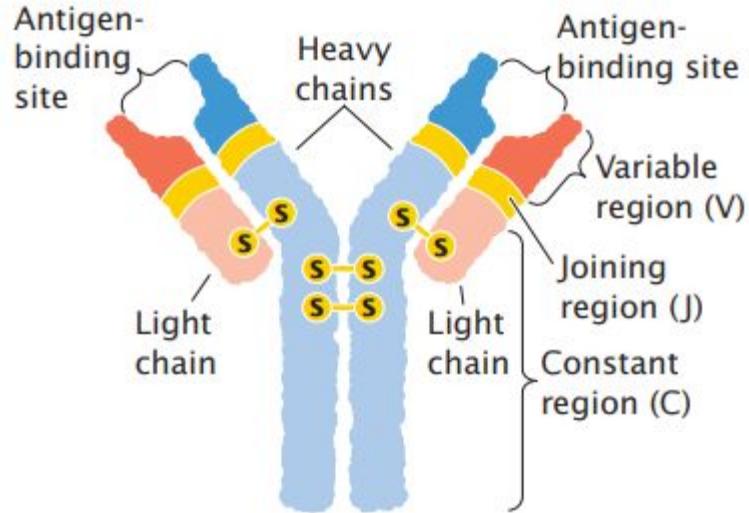


# The Unifying Principle: Clonal Selection



22.19 An immune response to a specific antigen is produced through clonal selection.

# The Unifying Principle: Clonal Selection



# Antibody Structure: The Key to Specificity

## Core Function:

Antibodies (Immunoglobulins, Ig) are the primary effector molecules of the humoral immune response, specifically binding to antigens to neutralize or mark them for destruction.

## Molecular Architecture: The Y-Shaped Molecule

- Composition: Made of four polypeptide chains.
  - Two identical Heavy Chains
  - Two identical Light Chains
- Bonding: Chains are held together by disulfide bonds.
- Overall Shape: A characteristic Y-shape.

## Light Chain Types

- There are two types: Kappa ( $\kappa$ ) and Lambda ( $\lambda$ ).
- A single antibody molecule will have two identical kappa chains OR two identical lambda chains (never one of each).

## Key Functional Regions

### 1. Variable Region (V)

- Location: Found at the tips of the "arms" of the Y.
- Composition: Created by the variable ends of both a light chain and a heavy chain.
- Function: Forms the Antigen-Binding Site. The unique amino acid sequence here determines which specific antigen the antibody can recognize and bind to.

### 2. Constant Region (C)

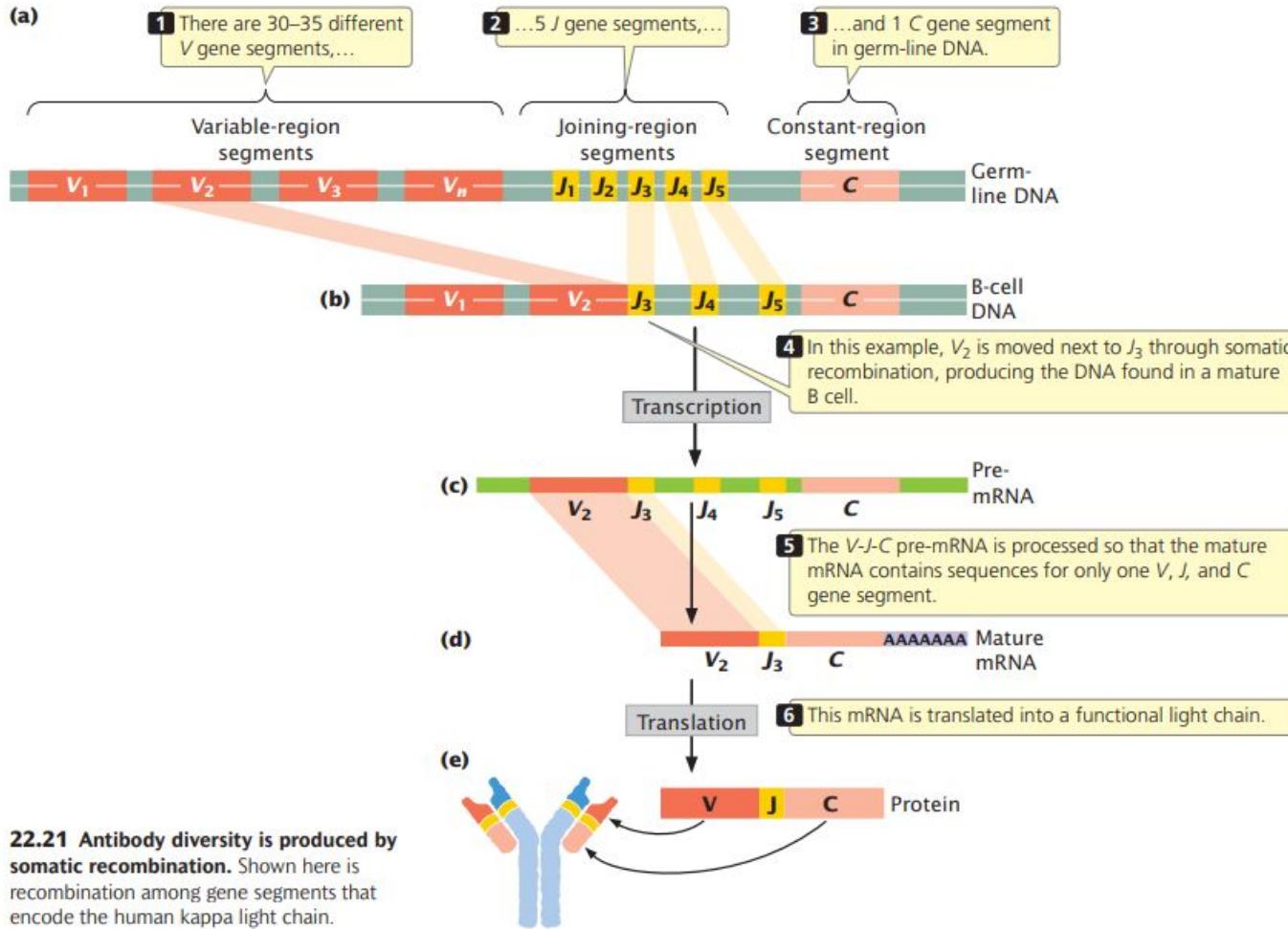
- Location: Forms the "stem" of the Y and the lower parts of the arms.
- Function: This region is relatively similar across antibodies of the same class. It determines the antibody's *effector function* (e.g., how it activates complement or interacts with immune cells).

# Generating Antibody Diversity: A Genetic Toolkit

## The Genetic Problem:

- The immune system must produce up to  $10^{15}$  different antibodies.
- The human genome only has about 20,000 genes.
- Solution: Antibody genes are not single units; they are assembled from a library of segments.

# Generating Antibody Diversity: A Genetic Toolkit



# Mechanism 1: Somatic Recombination ("Mix and Match")

The Segments:

- V (Variable) Segments: Encode most of the antigen-binding site.
- J (Joining) Segments: Short joining sequences.
- C (Constant) Segments: Encode the constant region that determines antibody class.
- (In Heavy Chains, there are also D (Diversity) segments)

The Assembly Process (e.g., Light Chain):

- Inheritance: A developing B cell inherits all V, J, and C segments in its germline DNA.
- Somatic Recombination: A random V segment is permanently joined to a random J segment via DNA rearrangement. The intervening DNA is lost.
- Transcription & Translation: The recombined DNA is transcribed and spliced into mRNA, which is translated into a unique light chain protein.

Example: Human Kappa Light Chain

**~35 V segments × 5 J segments × 1 C segment = ~175 possible combinations**

## Mechanism 2: Combinatorial Diversity

Heavy and Light Chain Pairing: Any unique heavy chain can pair with any unique light chain.

**Example:** If there are 5,000 different heavy chains and 1,000 different light chains, they can generate  $5,000 \times 1,000 = 5,000,000$  different antibodies.

## Mechanism 3: Junctional Diversity

Imprecise Joining: When V, J, and D segments are joined, the process is "sloppy."

Result: Random nucleotides are frequently deleted or inserted at the junctions.

Impact: Drastically increases the variation in the antigen-binding site's amino acid sequence.

## **Mechanism 4: Somatic Hypermutation**

Process: After exposure to an antigen, B cells undergo a high rate of point mutation in the antibody variable regions.

Result: Creates even more minor variations, allowing for "fine-tuning" and higher affinity antibodies.

# T-Cell Receptors: Cellular Immunity's Sensor

## Core Function:

Each mature T cell is genetically programmed to recognize one specific antigen presented on the surface of the body's own cells, initiating the cellular immune response.

## T-Cell Receptor (TCR) Structure

Composition: A cell-surface protein complex made of two chains:

- One Alpha ( $\alpha$ ) chain
- One Beta ( $\beta$ ) chain

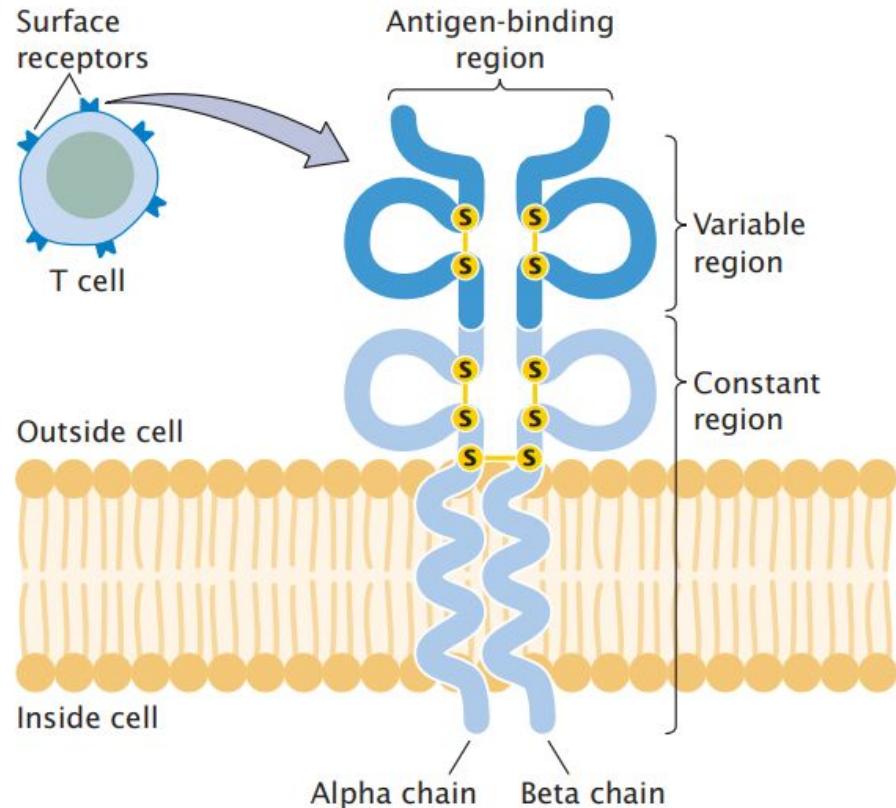
Bonding: Chains are held together by disulfide bonds.

## Key Regions:

Constant Region: Anchors the receptor in the cell membrane.

Variable Region: Forms the antigen-binding site at the tip of the receptor, determining specificity.

Analogy: Structurally very similar to an antibody fragment (Fab region), acting as the T cell's "antigen-sensing arm."



# T-Cell Receptors: Cellular Immunity's Sensor

The ability to recognize millions of different antigens is generated by the same genetic principles used for antibodies:

## 1. Somatic Recombination

- The genes for the  $\alpha$  and  $\beta$  chains are assembled from a library of segments (V, J, C for  $\alpha$ -chain; V, D, J, C for  $\beta$ -chain).
- Example: Human Alpha Chain Gene
  - $44\text{-}46 \text{ V segments} \times 50 \text{ J segments} \times 1 \text{ C segment} = \text{A vast number of possible combinations even before other mechanisms.}$

## 2. Combinatorial Diversity

- Any randomly generated  $\alpha$  chain can pair with any randomly generated  $\beta$  chain, multiplying the overall diversity.

## 3. Junctional Diversity

- The process of joining V, D, and J segments is imprecise, adding or deleting random nucleotides at the junctions, which dramatically increases sequence variation in the binding site.

## Key Difference from Antibodies

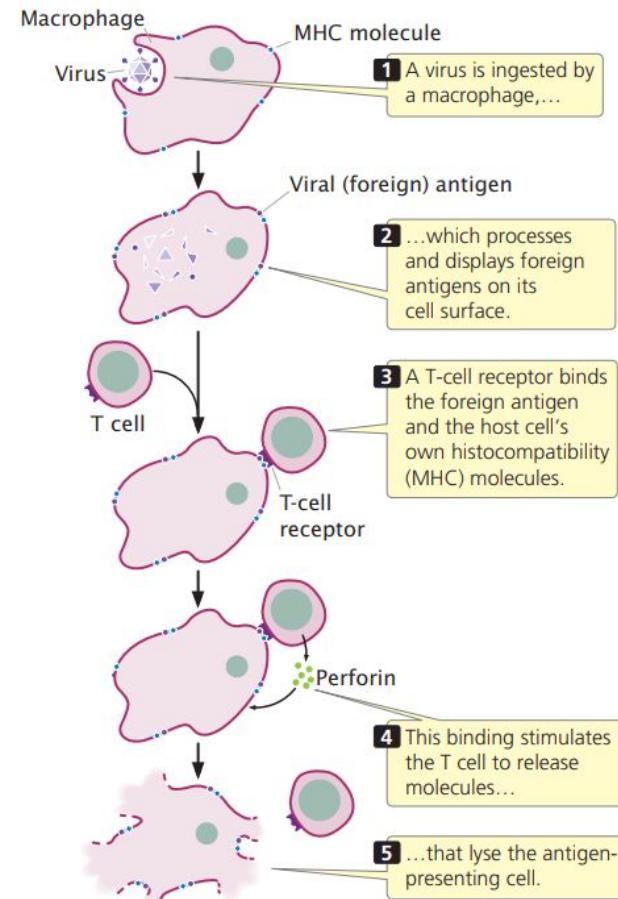
No Somatic Hypermutation: Unlike B cells, T cells do not undergo a high rate of mutation in their receptor genes after maturation. This is a critical safety mechanism to prevent T cells from developing reactivity against the body's own tissues.

# The Major Histocompatibility Complex (MHC) & Self-Recognition

## Core Concept: The "Self" ID Badge

The Major Histocompatibility Complex (MHC) is a cluster of genes that codes for histocompatibility antigens—proteins displayed on the surface of nearly all our own cells.

These proteins act as a unique "self" identification badge for every individual (except identical twins).



# The Dual-Signal Rule for T-Cell Activation

- A T cell will only become activated if its receptor binds BOTH signals simultaneously:
- A Foreign Antigen: A fragment (peptide) from a pathogen, like a virus or bacterium.
- A Self MHC Molecule: The host cell's own histocompatibility antigen, which "presents" the foreign fragment.
- This dual requirement ensures T cells only attack self cells that have been compromised (e.g., infected by a virus), not healthy cells or free-floating pathogens.
- Visual Analogy: The MHC is like a restaurant menu that displays samples (antigens) of what's inside the cell. The T cell is the inspector who reads the menu (MHC) and the sample (antigen) together.

# The Dual-Signal Rule for T-Cell Activation

- A T cell will only become activated if its receptor binds BOTH signals simultaneously:
- A Foreign Antigen: A fragment (peptide) from a pathogen, like a virus or bacterium.
- A Self MHC Molecule: The host cell's own histocompatibility antigen, which "presents" the foreign fragment.
- This dual requirement ensures T cells only attack self cells that have been compromised (e.g., infected by a virus), not healthy cells or free-floating pathogens.
- Visual Analogy: The MHC is like a restaurant menu that displays samples (antigens) of what's inside the cell. The T cell is the inspector who reads the menu (MHC) and the sample (antigen) together.

# The Transplant Challenge: Genetics & Immune Rejection

## **The Core Problem:**

Transplanted organs are recognized as "foreign" by the recipient's immune system, primarily due to mismatched Major Histocompatibility Complex (MHC) antigens, leading to organ rejection.

# The Genetic Keys to Compatibility

## Major Histocompatibility Complex (MHC)

Primary Culprit: MHC proteins on donor cells are the strongest triggers of immune rejection.

Goal: Maximize the number of matched MHC antigens between donor and recipient.

Evidence: Kidney transplant success rates jump from ~63% with 0-1 matches to ~90% with 4 matches.

## ABO Blood Group Antigens

Critical Factor: These red blood cell antigens also elicit a powerful immune reaction.

Requirement: Donor and recipient must have compatible ABO blood types.

# The Genetic Keys to Compatibility

## The Donor Match Hierarchy

Donor Type	Genetic Match	Advantage
Identical Twin	Perfect MHC & ABO match	Ideal; minimal risk of rejection.
Sibling	High probability of MHC & ABO match	Excellent alternative.
General Population	Variable MHC match, ABO compatible	Requires careful screening and drugs.

# **The Genetic Keys to Compatibility**

## **Managing the Immune Response**

### **Strategy 1: Immunosuppressive Drugs**

- Function: Suppress the patient's entire immune system to prevent attack on the new organ.
- Major Drawback: Leaves the patient vulnerable to common infections and other diseases.

### **Strategy 2: Meticulous Genetic Matching**

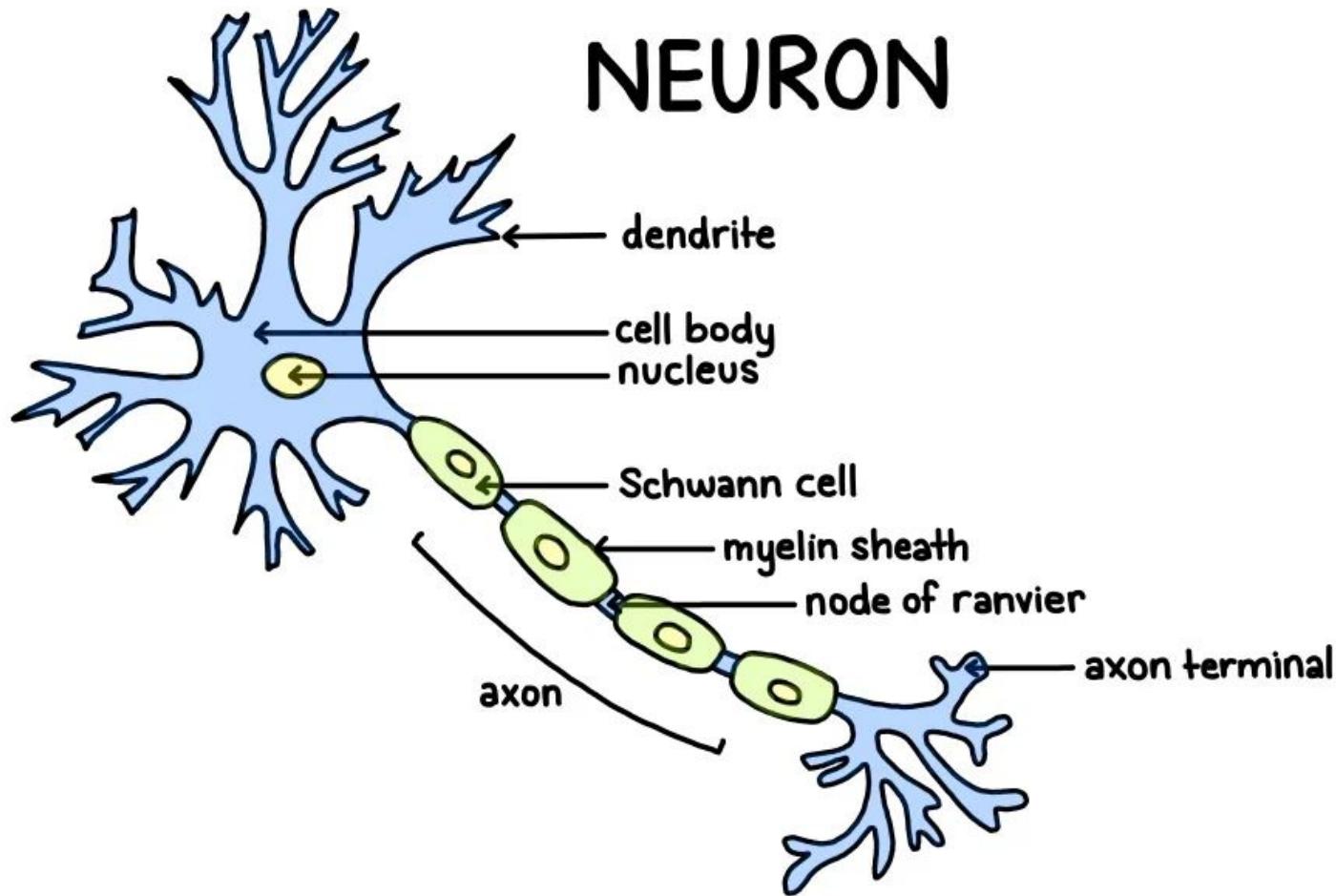
- Function: Reduces the inherent "foreignness" of the organ, lessening the need for high drug doses.
- The best clinical outcome is achieved by combining both strategies.

# The Body's Command & Control: Neurobiology and Endocrinology

## Key Topics:

- Neural Signaling & Synaptic Transmission
- The Neural Basis of Behavior & Cognition
- Endocrine Glands & Hormone Function
- Hormonal Regulation of Physiological Processes

# NEURON



# The Neuron: The Basic Unit of Communication

A neuron is a nerve cell that processes and transmits information through electrical and chemical signals in the nervous system.

Neurons consist of a cell body, dendrites (which receive signals), and an axon (which sends signals).

Synaptic connections allow communication between neurons, facilitating the relay of information throughout the body.

# The Neuron: The Basic Unit of Communication

## Dendrites

Dendrites are the tree-root-shaped part of the neuron which are usually shorter and more numerous than axons.

Their purpose is to receive information from other neurons and to transmit electrical signals to the cell body.

**Dendrites are covered in synapses**, which allow them to receive signals from other neurons. Some neurons have short dendrites, whilst others have longer ones.

For instance, cells called Purkinje cells, which are **found in the cerebellum**, have highly developed dendrites to receive signals from thousands of other cells.

# The Neuron: The Basic Unit of Communication

## Soma (Cell Body)

The soma, or cell body, is essentially the core of the neuron.

The soma's function is **to maintain the cell and to keep the neuron functioning efficiently** (Luengo-Sanchez et al., 2015).

The soma is enclosed by a membrane that protects it and allows it to interact with its immediate surroundings.

The soma contains a cell nucleus that produces genetic information and directs the synthesis of proteins.

These proteins are vital for other parts of the neuron to function.

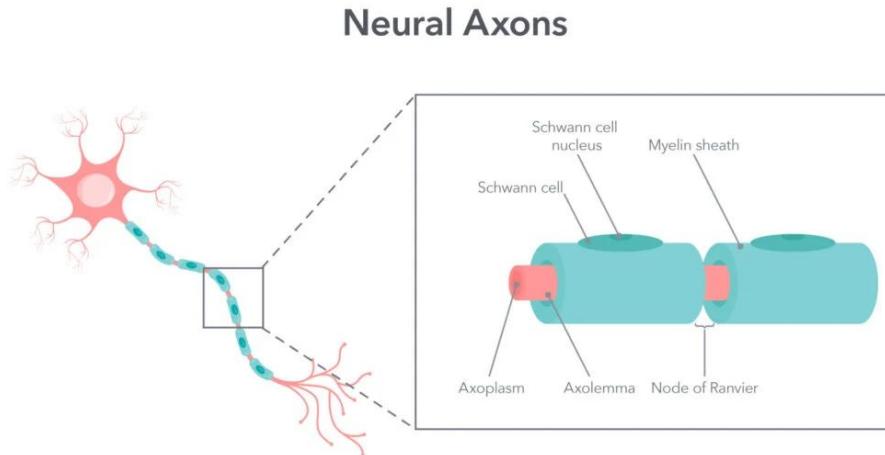
# The Neuron: The Basic Unit of Communication

## Axon

The axon, also called a nerve fiber, is a tail-like structure of the neuron that joins the cell body at a junction called the axon hillock.

The function of the axon is to **carry signals away from the cell body** to the terminal buttons to transmit electrical signals to other neurons.

Acting as a conduit, the axon carries these signals to other neurons, muscles, or glands.



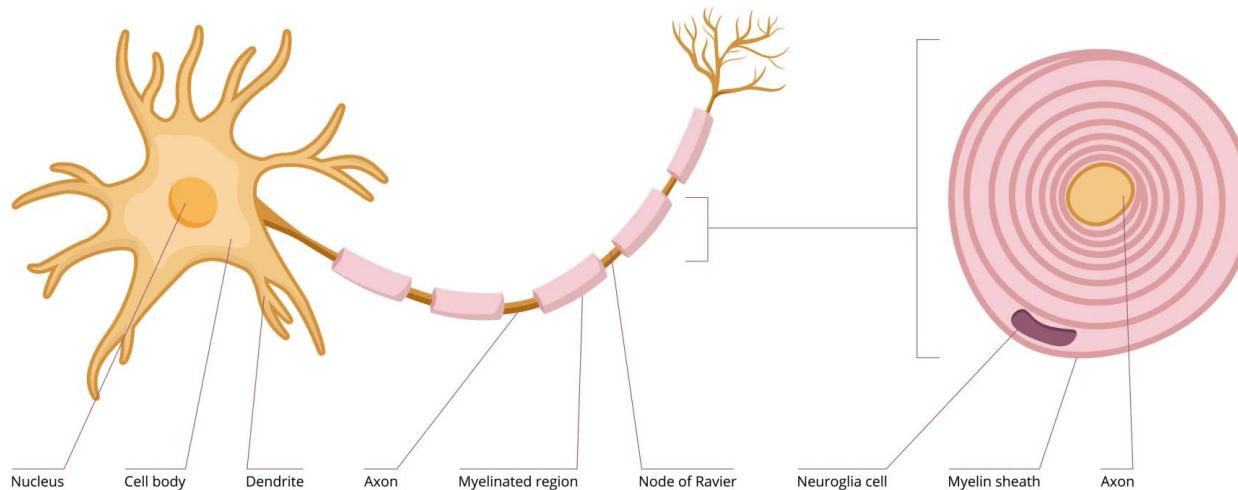
# The Neuron: The Basic Unit of Communication

## Myelin Sheath

The myelin sheath is a layer of fatty material that covers the axons of neurons.

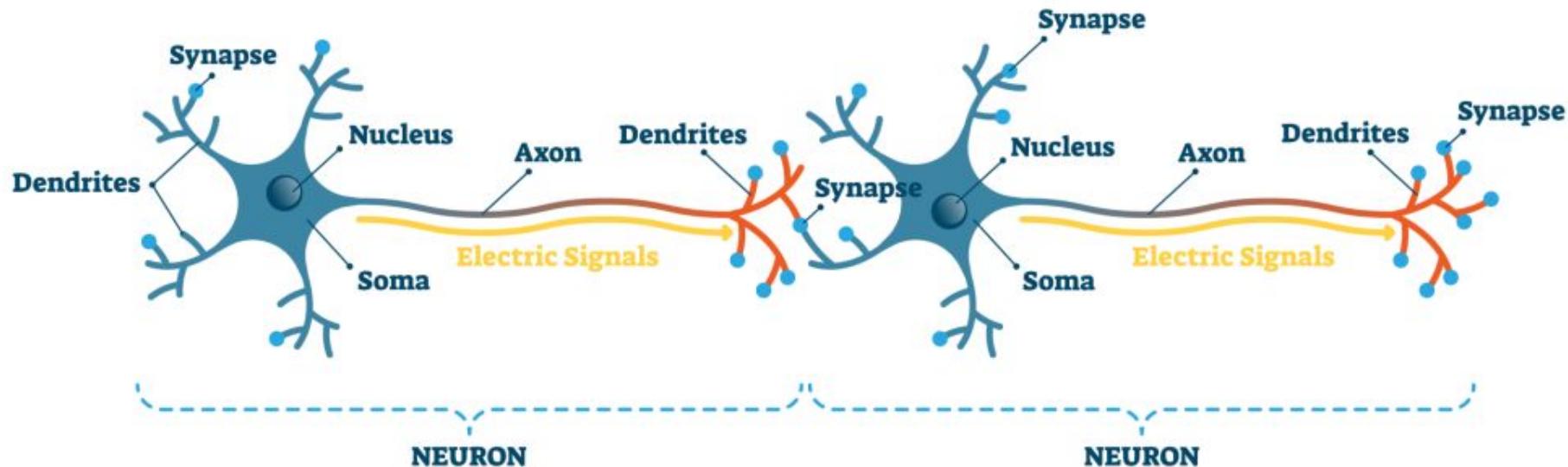
Its purpose is **to insulate** one nerve cell from another and to **prevent the impulse** from one neuron from interfering with the impulse from another.

The second function of the myelin sheath is to speed up the conduction of nerve impulses along the axon.



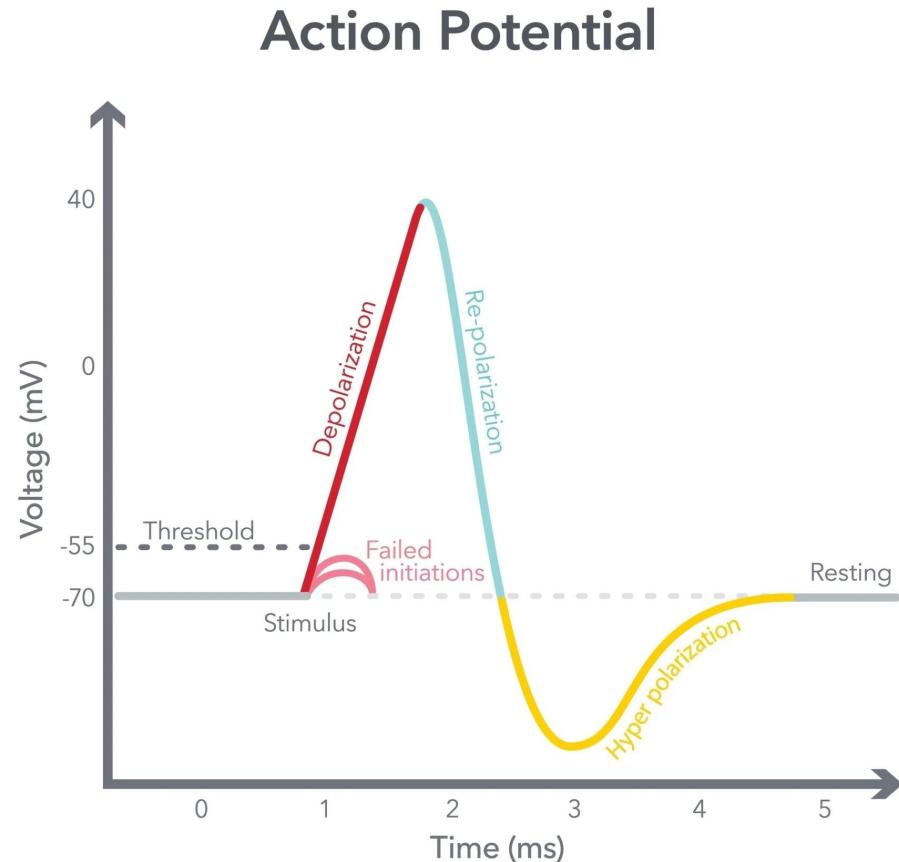
# Neural Signaling: The Action Potential

- Neurons lie adjacent to each other but are not connected.
- There is a tiny gap between neurons called a synapse.
- The function of a neuron is to transmit nerve impulses along the length of an individual neuron and across the synapse into the next neuron.
- The electrical signals transmitted by neurons are called action potentials.



# What Is Action Potential?

An action potential is an electrical nerve impulse that travels along a neuron's axon. It's a transient, all-or-nothing electrical current that is conducted down the axon when the neuron's membrane potential reaches a specific "threshold of excitation".



# Why are action potentials important?

Action potentials are fundamental electrical impulses crucial for all nerve functions and the nervous system's ability to process and transmit information.

They enable neurons to communicate by acting as the electrical signal that travels down an axon.

When this impulse reaches the axon terminal, it triggers the release of chemical neurotransmitters into the synapse, transmitting the message to adjacent neurons.

This electrochemical communication is the basis for all nerve functions, allowing processes like thought, memory, and emotion.

Action potentials are vital for muscle contraction, as motor neurons carry these impulses from the central nervous system to muscles, enabling voluntary movement.

They are critical for brain signaling, forming the rapid electrical basis of sensory detection (e.g., light, touch) and the perception of sensations like pain.

In reflexes, action potentials allow for immediate, involuntary responses, like withdrawing a hand from a hot object, without direct brain input, saving crucial seconds for survival.

Overall, they are essential for coordinated bodily control and interaction with the environment.

# Where do action potentials occur?

Action potentials are typically generated in a neuron's axon hillock, the specialized region where the axon extends from the cell body, once the "threshold of excitation" is reached.

From there, this "all-or-none" electrical signal rapidly travels down the entire length of the axon to the axon terminals (or terminal buttons) at its end.

Beyond neurons, action potentials are fundamental to other excitable cells.

For instance, motor neurons use action potentials to transmit commands from the nervous system to muscle fibers, initiating muscle contraction.

Similarly, sensory neurons convert external stimuli into action potentials to relay information about touch, sight, hearing, and pain to the central nervous system.

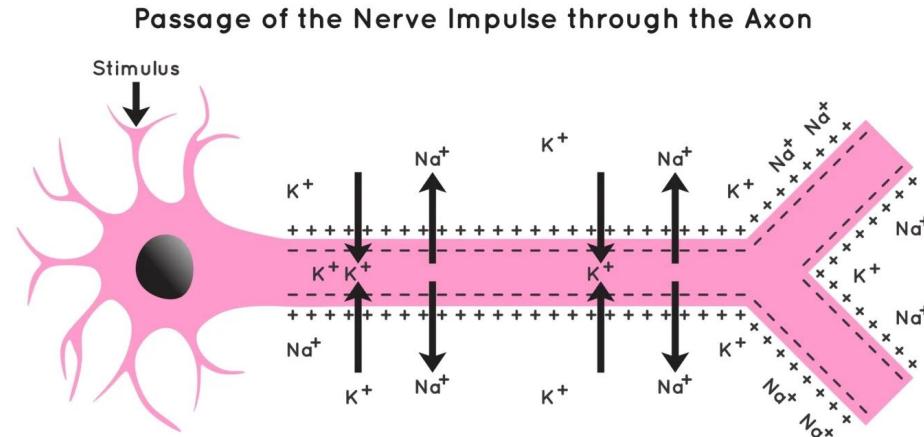
# The 3 phases of an action potential

## Depolarization (Rising Phase)

A neuron at resting potential maintains a negative charge inside, usually around -70mV, with higher sodium ( $\text{Na}^+$ ) ion concentrations outside and potassium ( $\text{K}^+$ ) ions inside.

When sufficiently stimulated, voltage-gated sodium ( $\text{Na}^+$ ) channels open, **allowing positively charged  $\text{Na}^+$  ions to rapidly rush into the cell**. This influx causes the inside of the cell to become more positive.

If this charge reaches the threshold of excitation (around -55mV), a massive influx of  $\text{Na}^+$  occurs, creating a positive spike in the membrane potential.



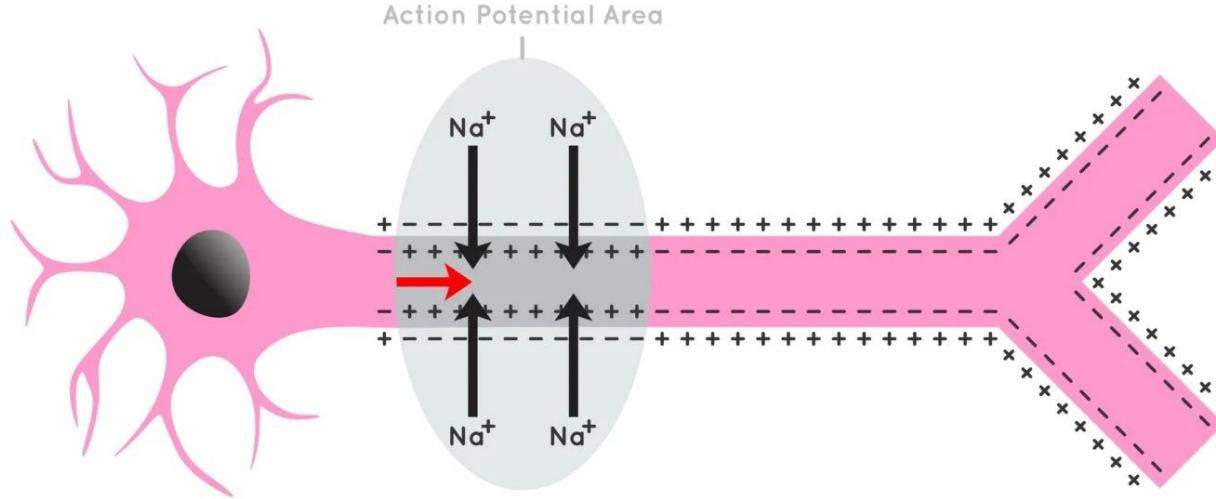
# The 3 phases of an action potential

## Repolarization (Falling Phase)

Immediately following the peak of depolarization, sodium ( $\text{Na}^+$ ) channels close, halting the influx of positive ions.

Simultaneously, potassium ( $\text{K}^+$ ) channels open, leading to a rapid outward flow of positively charged  $\text{K}^+$  ions from the neuron.

This efflux of positive charge quickly **restores the membrane potential to a negative charge** inside the cell.



# The 3 phases of an action potential

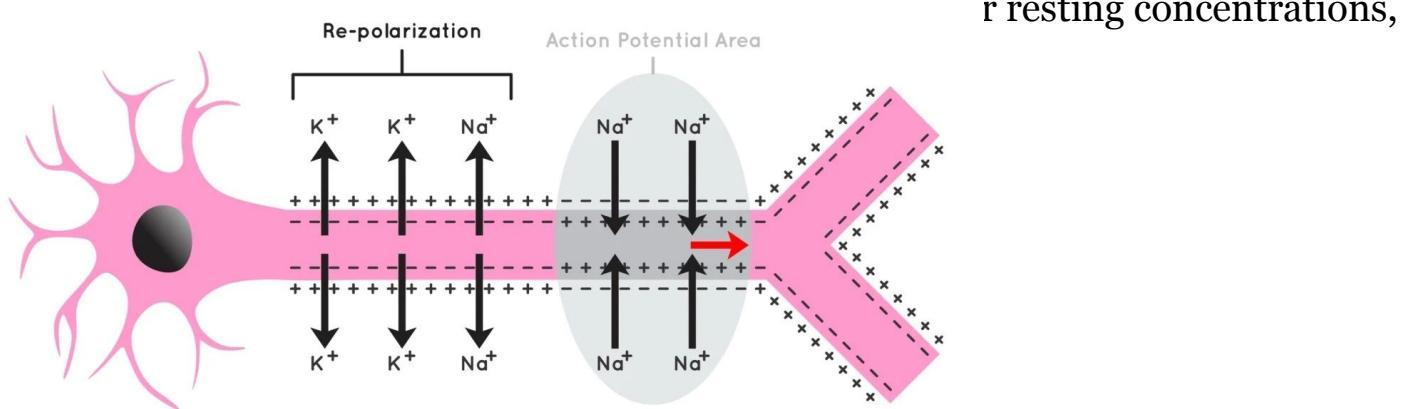
## Hyperpolarization (Undershoot)

This brief phase, also known as the undershoot, occurs because the potassium ( $K^+$ ) channels close relatively slowly.

This allows a slight excess of  $K^+$  ions to leave the cell, causing the membrane potential to **temporarily become even more negative than its normal resting potential** (e.g., dropping to -80mV).

This state contributes to the relative refractory period, making it harder for the neuron to fire again immediately.

Finally, the sodium-potassium pump returns the neuron to its resting concentrations,



# Synaptic Transmission

Synaptic transmission is the process by which one neuron communicates with another.

Information is passed down the axon of the neuron as an electrical impulse known as action potential.

The electrical signal needs to cross the synaptic gap to continue on its journey to or from the CNS.

This is done using chemicals that diffuse across the gap between the two neurons. These chemicals are called neurotransmitters.

During **synaptic transmission**, the action potential (an electrical impulse) triggers the synaptic vesicles of the pre-synaptic neuron to release neurotransmitters.

These neurotransmitters diffuse across the synaptic gap and bind to specialized receptor sites on the post-synaptic neuron.

This will then trigger an electrical impulse in the adjacent cell.

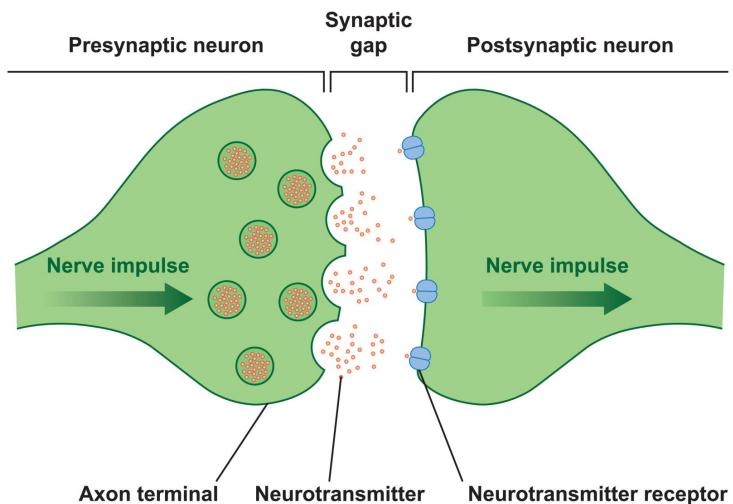
The **central nervous system**, which comprises the brain and spinal cord, and the peripheral nervous system, which consists of sensory and motor nerve cells, all contain these information-processing neurons.

# Synaptic Transmission

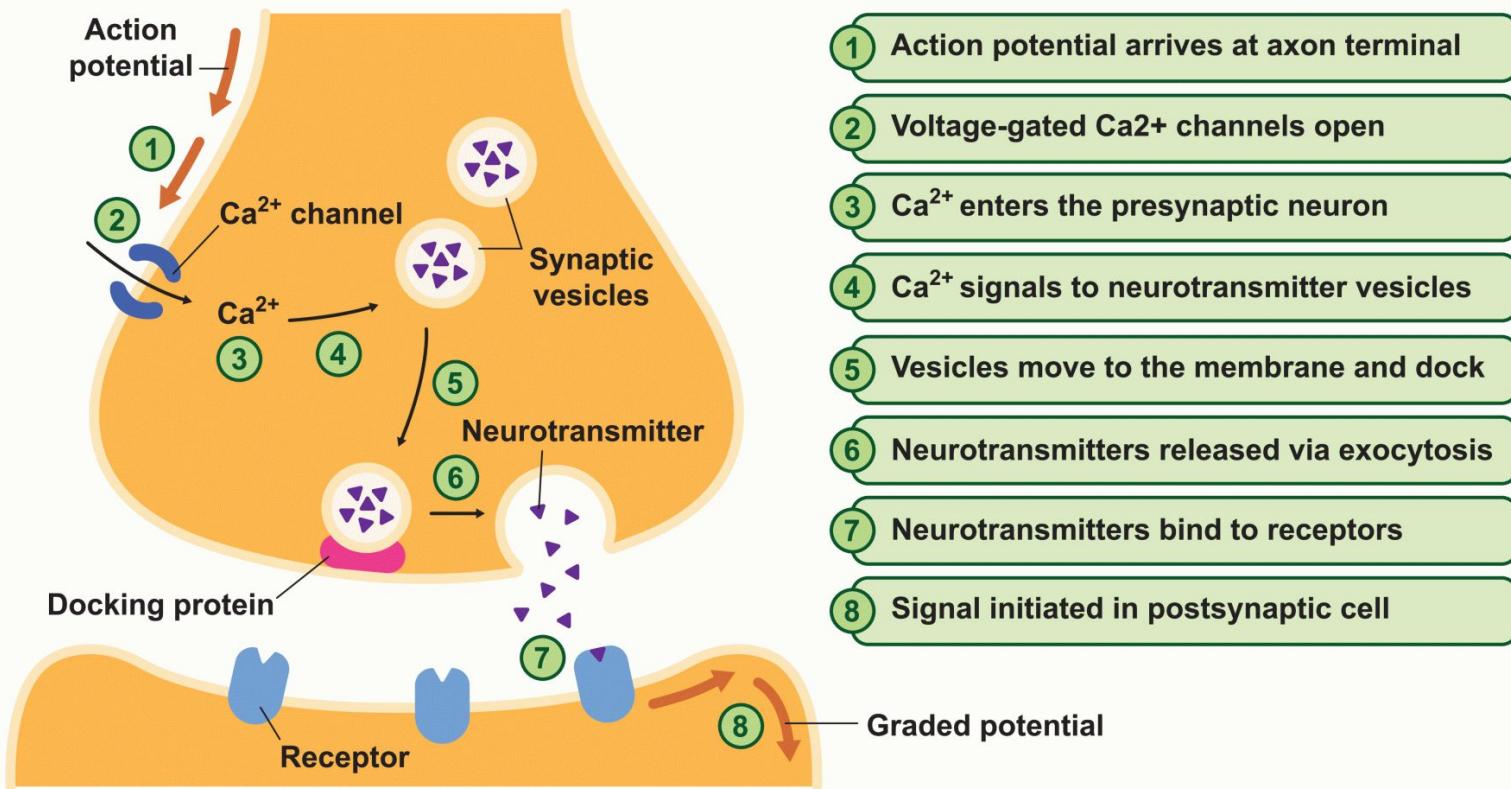
The structures include:

- **Presynaptic membrane:** The nerve ending of the transmitting neuron.
- **Postsynaptic membrane:** The membrane of the receiving neuron.
- **Synaptic cleft:** The tiny gap between the presynaptic and postsynaptic membranes.
- **Neurotransmitters:** Chemical messengers stored in packets called vesicles within the presynaptic terminal (axon terminals).
- **Postsynaptic receptor sites:** These are specific molecules on the membrane of the postsynaptic neuron that neurotransmitters bind to.

## Synaptic Transmission



# Synaptic Transmission



# **Neurotransmitters: Chemical Messengers**

**Neurotransmitter**s are brain chemicals that communicate information throughout our brain and body. They are released from the axon terminals.

The brain uses them for vital functions like telling your heart to beat or lungs to breathe, and they can also affect mood, sleep, and concentration.

Each neurotransmitter has a specific function. For example, acetylcholine is found where a motor neuron meets a muscle and causes the muscle to contract when released.

The action of **dopamine** at the synapse is linked to explanations for schizophrenia. Imbalances in neurotransmitters like serotonin are associated with conditions such as depression

Synaptic connections can be defined by the specific neurotransmitter they release, such as serotonin, dopamine, adrenaline, or GABA.

# The Body's Command & Control: Neurobiology and Endocrinology

# Introduction

- The nervous and endocrine systems act together to coordinate functions of all body systems.
- The nervous system acts through nerve impulses (action potentials) conducted along axons of neurons. At synapses, nerve impulses trigger the release of mediator (messenger) molecules called neurotransmitters.
- The endocrine system also controls body activities by releasing mediators, called hormones
- The means of control of the two systems are very different.

# The Hormones

Hormones are chemical messengers created by the body.

They transfer information from one set of cells to another

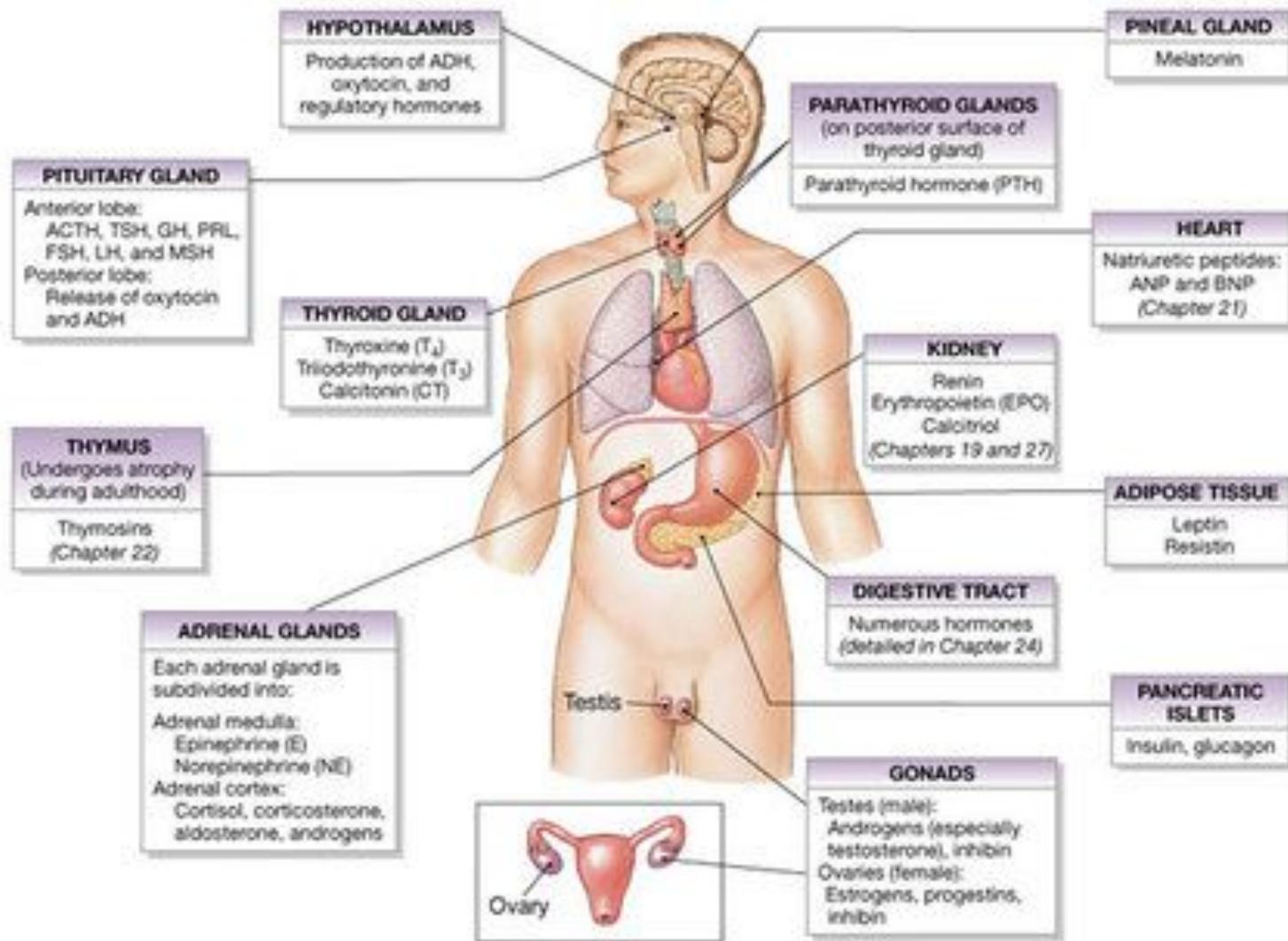
to coordinate the functions of different parts of the body.

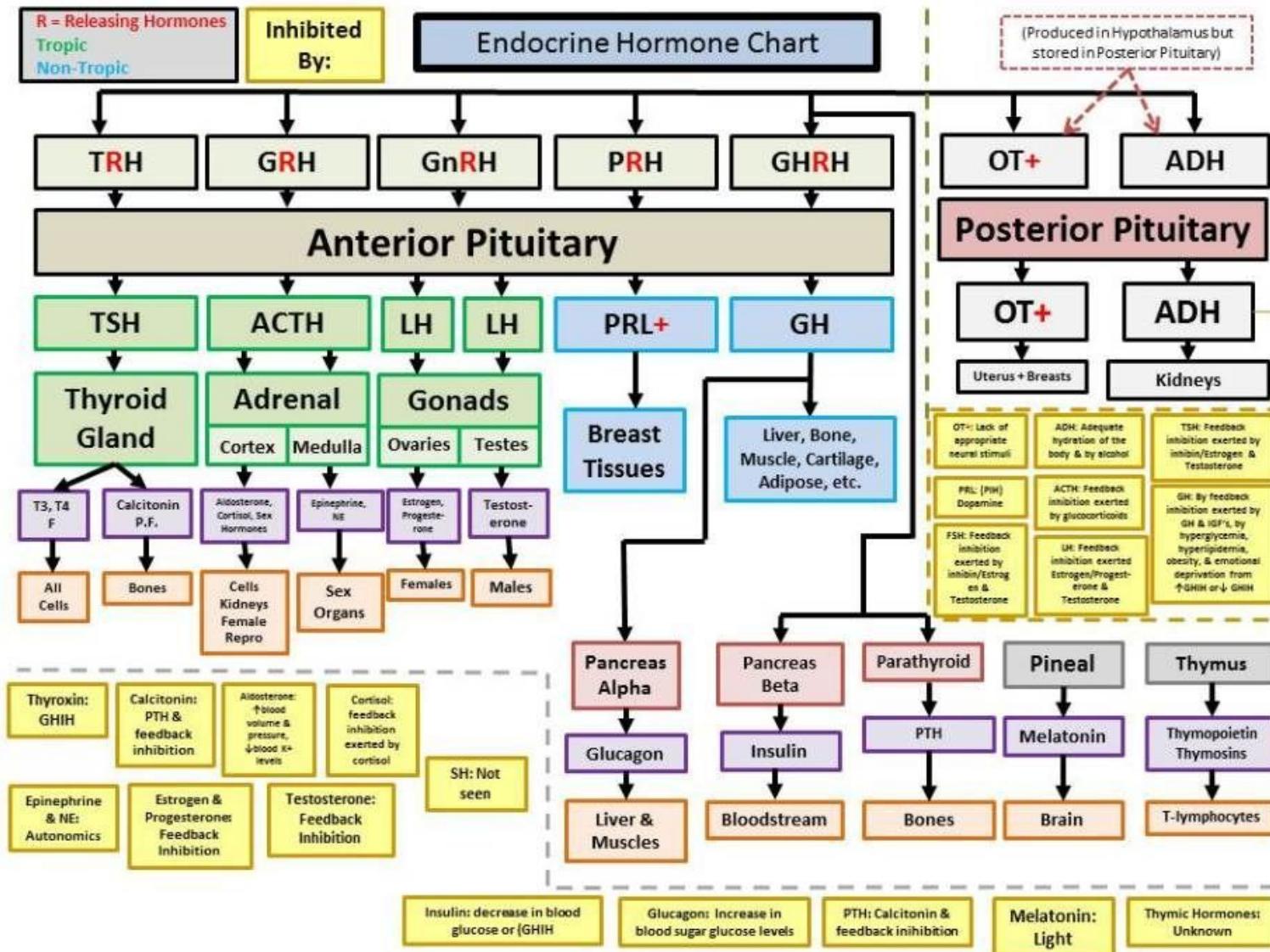
# **How the system communicates**

The organs of the endocrine system communicate with each other and the rest of the body using hormones.

Hormones are made and stored in their originating organ. The place they are sent is called the target.

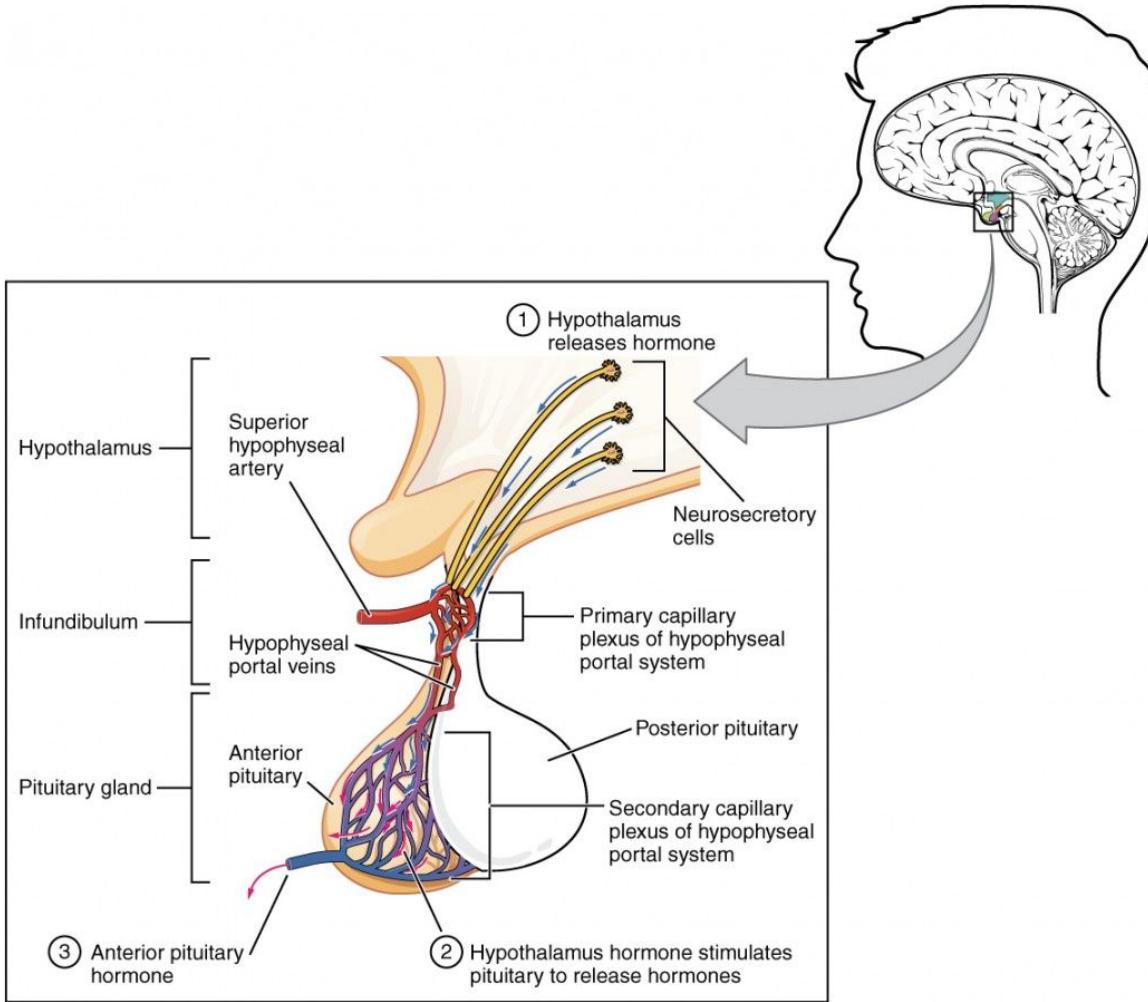
Targets can be anything from all the bones in the body to another endocrine organ. When the hormones reach their target, they trigger events in that location.

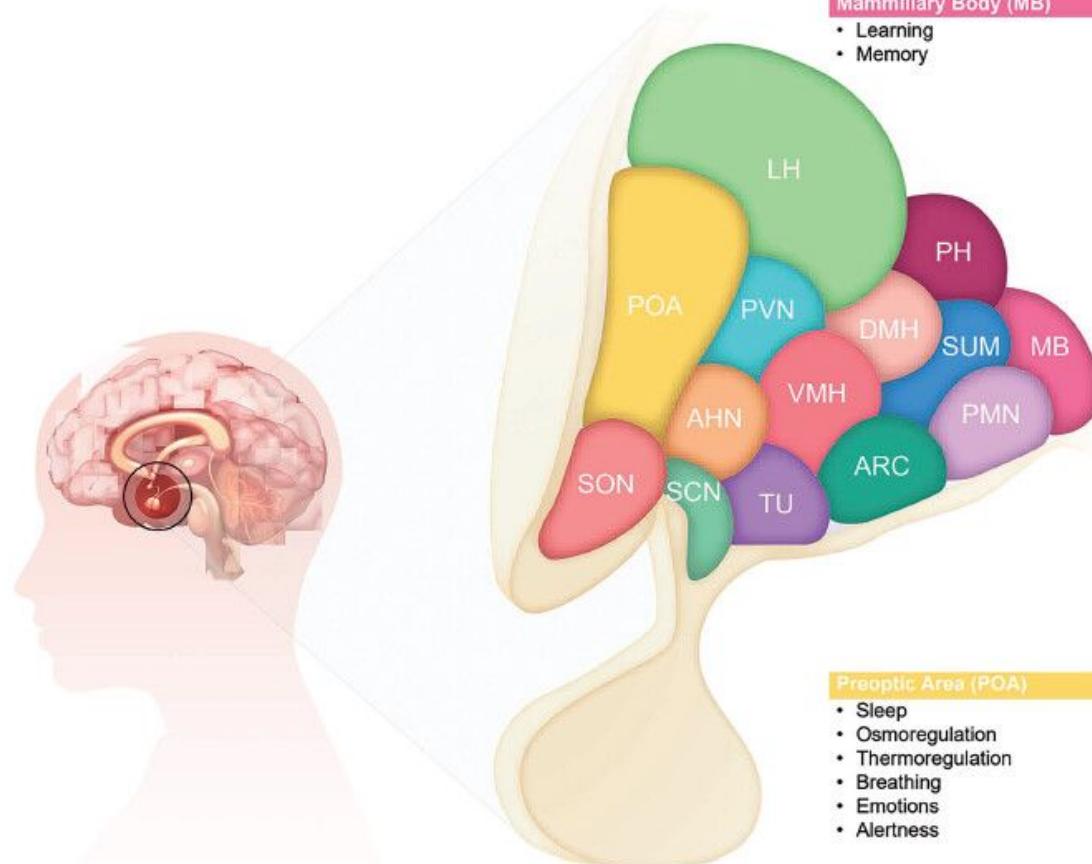




GLAND	HORMONES PRODUCED	EFFECTS
<b>Hypothalamus</b>	<ul style="list-style-type: none"> <li>Gonadotropin-releasing hormone (GnRH)</li> <li>Prolactin-releasing hormone</li> <li>Relaxin</li> <li>Growth hormone</li> </ul>	<ul style="list-style-type: none"> <li>Body temperature</li> <li>Hunger</li> <li>Moods</li> <li>Thirst</li> <li>Sleep</li> <li>Directs the release of hormones from other glands.</li> </ul>
<b>Pineal gland</b>	<ul style="list-style-type: none"> <li>Melatonin</li> </ul>	<ul style="list-style-type: none"> <li>Sleep</li> </ul>
<b>Pituitary gland</b> (aka the 'master control gland') • Has two parts: anterior and posterior; • Connects to the hypothalamus by a stalk made of blood vessels and nerve fibers	<ul style="list-style-type: none"> <li>Follicle-stimulating hormone (FSH)</li> <li>Luteinizing hormone (LH)</li> <li>Prolactin</li> <li>Growth hormone</li> <li>Adrenocorticotropic hormone (ACTH)</li> <li>Oxytocin</li> </ul>	<ul style="list-style-type: none"> <li>Regulates most other endocrine glands (including ovaries)</li> <li>Controls function of some organs</li> <li>Growth</li> <li>Milk production</li> <li>Development of breast tissue</li> <li>Autonomic nervous system (including heart rate, body temperature, and urination)</li> <li>Cortisol production (maintains blood pressure and blood sugar levels)</li> <li>Progresses labor</li> </ul>
<b>Thyroid gland</b>	<p>Uses iodine from food to make</p> <ul style="list-style-type: none"> <li>Triiodothyronine (T3)</li> <li>Thyroxine (T4)</li> </ul>	<ul style="list-style-type: none"> <li>Metabolism</li> <li>Breathing</li> <li>Heart rate</li> <li>Cholesterol levels</li> </ul>
<b>Parathyroid glands</b>	<ul style="list-style-type: none"> <li>Parathyroid hormone (PTH)</li> </ul>	<ul style="list-style-type: none"> <li>Regulates the amount of calcium, phosphorus and vitamin D in the body</li> </ul>
<b>Thymus gland</b> • Only active till puberty • Also part of the immune system	<ul style="list-style-type: none"> <li>Thymosin</li> </ul>	<ul style="list-style-type: none"> <li>Production and maturation of T-lymphocytes or T cells (note that all T cells in the body are produced by puberty)</li> </ul>
<b>Adrenal glands</b>	<ul style="list-style-type: none"> <li>Cortisol</li> <li>Aldosterone</li> <li>Androgenic steroids (converted to estrogens in the ovaries)</li> <li>Epinephrine (Adrenaline) Norepinephrine (Noradrenaline)</li> </ul>	<ul style="list-style-type: none"> <li>Response to stress</li> <li>Metabolism</li> <li>Immune system</li> <li>Blood pressure</li> <li>Heart rate</li> </ul>
<b>Pancreas</b>	<ul style="list-style-type: none"> <li>Insulin</li> </ul>	<ul style="list-style-type: none"> <li>Maintains blood sugar levels</li> </ul>
<b>Ovaries</b>	<ul style="list-style-type: none"> <li>Progesterone</li> <li>Estradiol</li> <li>Estrone</li> <li>Estriol</li> </ul>	<ul style="list-style-type: none"> <li>Development of female sex characteristics menstrual cycle</li> <li>Reproductive system</li> </ul>

# Pituitary and Hypothalamus





#### Arcuate Nucleus (ARC)

- Metabolic homeostasis
- Growth hormone release
- Reproduction
- Prolactin release

#### Premammillary Nucleus (PMN)

- Social behavior
- Reproduction

#### Paraventricular Nucleus (PVN)

- Stress response
- Thyroid regulation
- Osmoregulation
- Food intake
- Breathing
- Cardiac rhythm
- Blood Pressure

#### Mammillary Body (MB)

- Learning
- Memory

#### Posterior Hypothalamus (PH)

- Stress response
- Blood pressure

#### Dorsomedial Hypothalamus (DMH)

- Energy Expenditure
- Circadian behavioral patterns
- Thermoregulation
- Food intake
- Body weight regulation

#### Ventromedial Hypothalamus (VMH)

- Energy balance
- Glucose metabolism
- Sex-specific social behavior

#### Supramammillary Nucleus (SUM)

- Spatial memory
- Arousal

#### Lateral Hypothalamus (LH)

- Arousal
- Food intake
- Reward
- Decision making
- Cost benefit
- Predation/Evasion

#### Anterior Hypothalamic Nucleus (AHN)

- Thermoregulation
- Defensive behavior

#### Supraoptic Nucleus (SON)

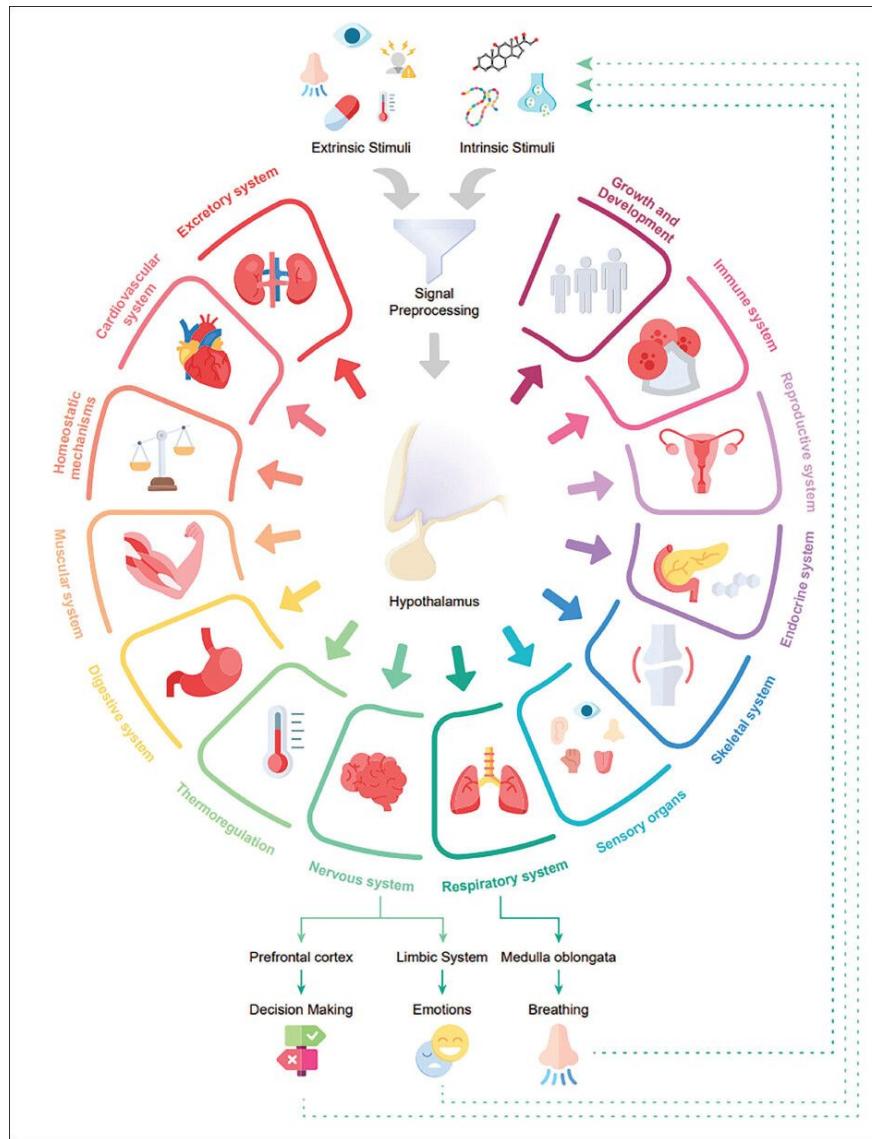
- Osmoregulation
- Blood pressure
- Lactation
- Parturition
- Metabolic regulation
- Energy balance

#### Tuberal Nucleus (TU)

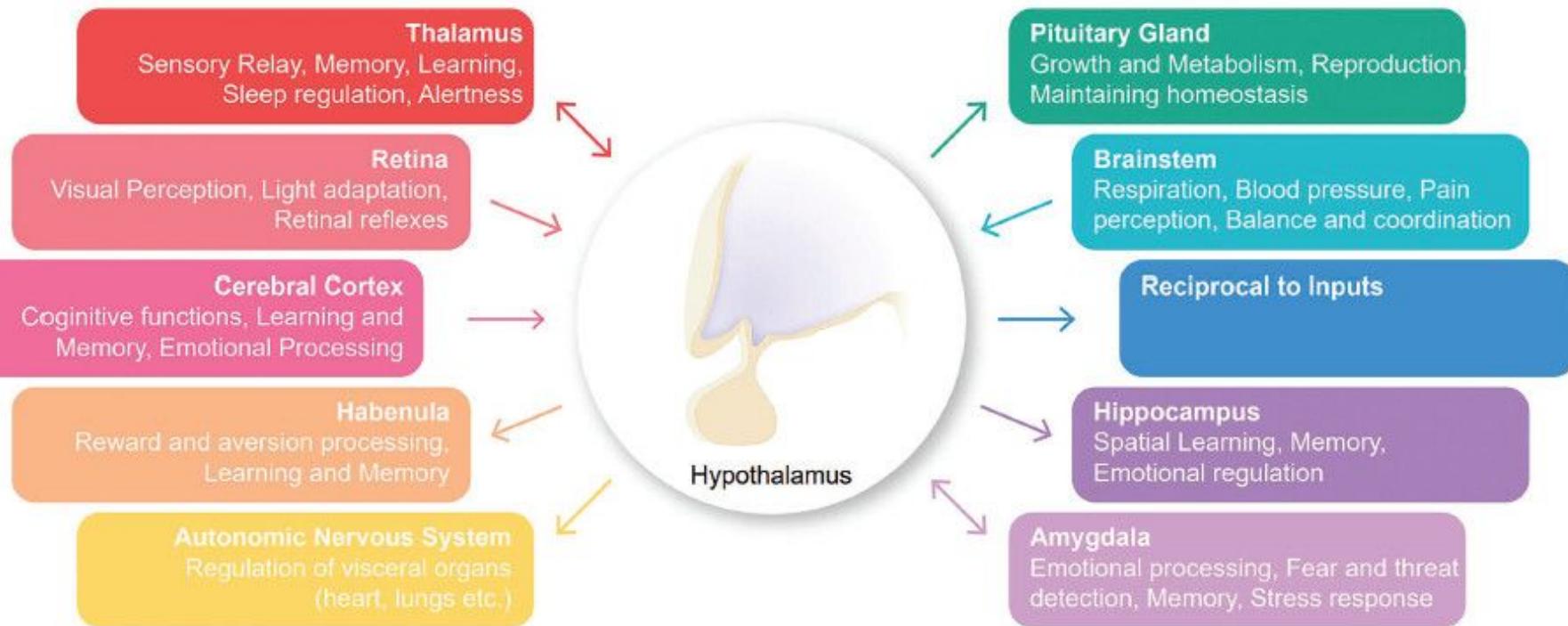
- Food intake

#### Suprachiasmatic Nucleus (SCN)

- Sleep
- Circadian rhythm

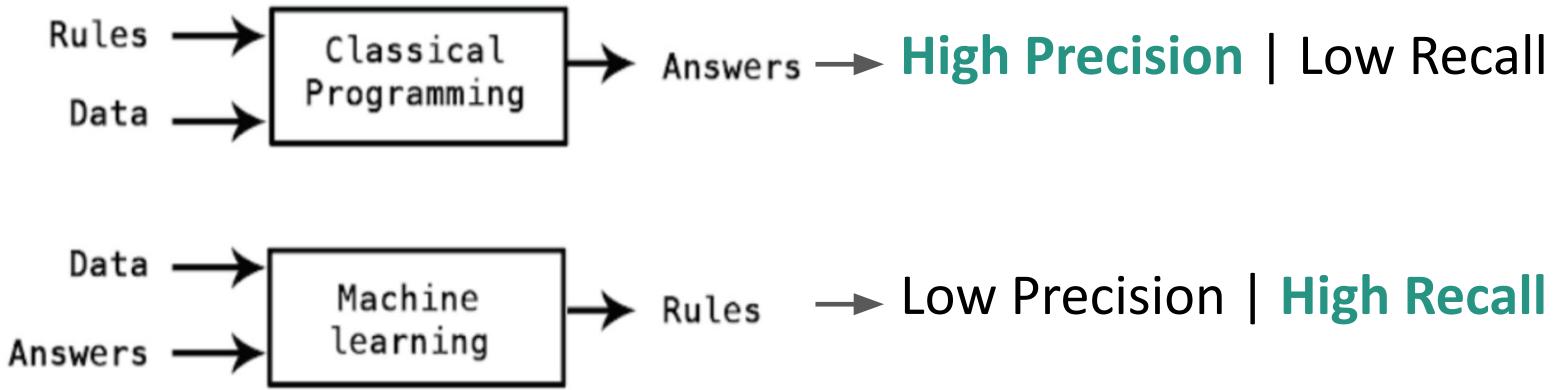
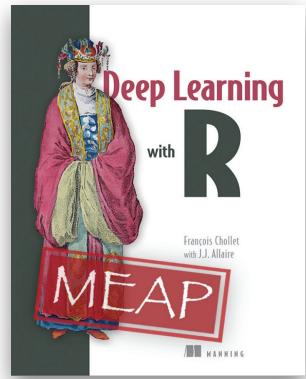


## Hypothalamic Connections



# **Endocrine System**

# Combining the Best of Both Worlds: Modeling Mechanisms



- Precision can be seen as a measure of quality
- Recall as a measure of quantity

Is Hybrid Approach is Good for Chemical Space Exploration?

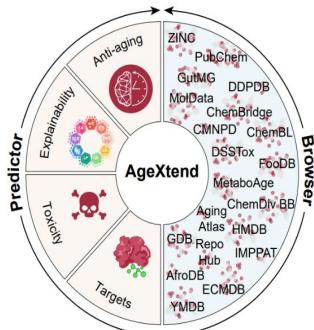
# Some Technologies From Our Lab



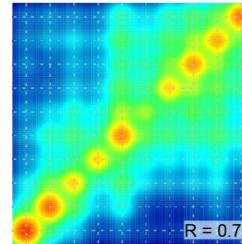
Oncogen AI  
*In review*



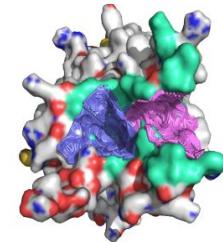
MetaboKiller  
*Nature Chem Bio*



AgeXtend  
*Nature Aging*



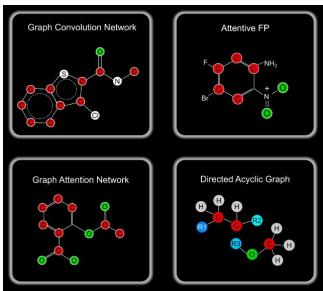
scCamAge  
*Cell Reports*



Gcoupler  
*ELife*



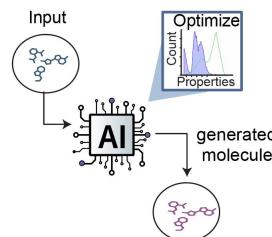
OdoriFy  
*J. Bio Chem (JBC)*



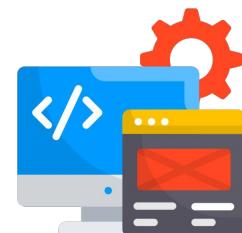
DeepGraphh  
*Brief. In BioInfo.*



EvoKG  
*Unpublished*



MetaboGlue  
*In review*



Many Others

# Metabolites as Endogenous Carcinogens

nature chemical biology

Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

nature > nature chemical biology > articles > article

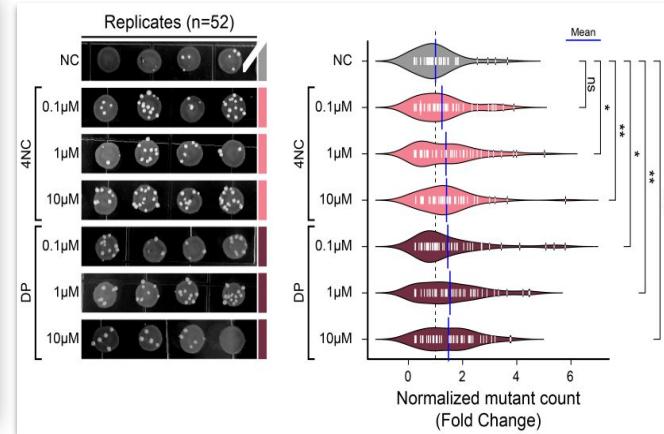
Article | Published: 11 August 2022

## Artificial intelligence uncovers carcinogenic human metabolites

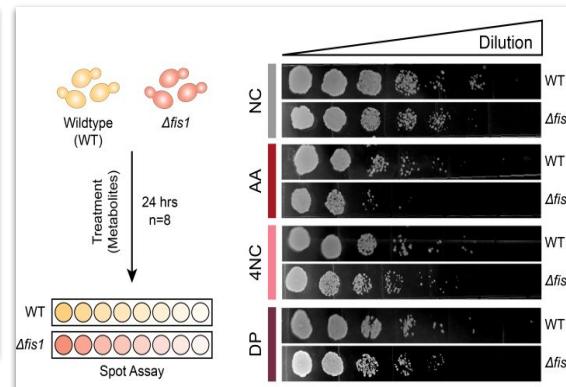
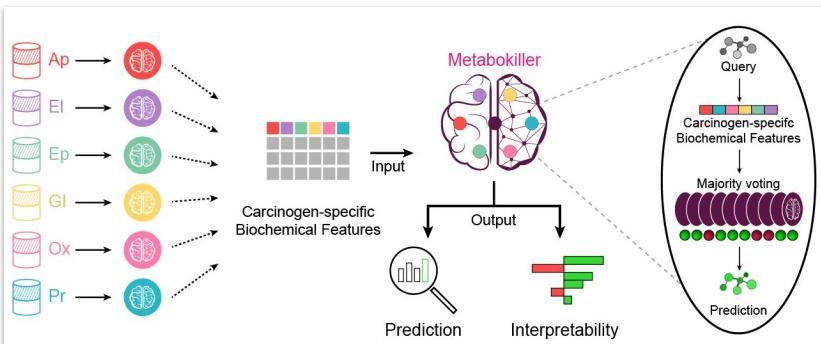
Aayushi Mittal, Sanjay Kumar Mohanty, Vishakha Gautam, Sakshi Arora, Sheetanshu Saproo, Ria Gupta, Roshan Sivakumar, Prakriti Garg, Anmol Aggarwal, Padmasini Raghavachary, Nilesh Kumar Dixit, Vijay Pal Singh, Anurag Mehta, Juhu Tayal, Srivatsava Naidu, Debarka Sengupta & Gaurav Ahuja

Nature Chemical Biology 18, 1204–1213 (2022) | Cite this article

3606 Accesses | 4 Citations | 98 Altmetric | Metrics



Aayushi



# Hunting Geroprotectors using AgeXtend

nature aging

Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

nature > nature aging > technical reports > article

Technical Report | Published: 03 December 2024

## Discovering geroprotectors through the explainable artificial intelligence-based platform AgeXtend

Sakshi Arora, Aayushi Mittal, Subhadeep Duari, Sonam Chauhan, Nilesh Kumar Dixit, Sanjay Kumar

Mohanty, Arushi Sharma, Saveena Solanki, Anmol Kumar Sharma, Vishakha Gautam, Pushpendra Singh

Gahlot, Shiva Satija, Jeet Nanshi, Nikita Kapoor, Lavanya CB, Debarka Sengupta, Parul Mehrotra, Tarini

Shankar Ghosh & Gaurav Ahuja ☐

Nature Aging (2024) | Cite this article

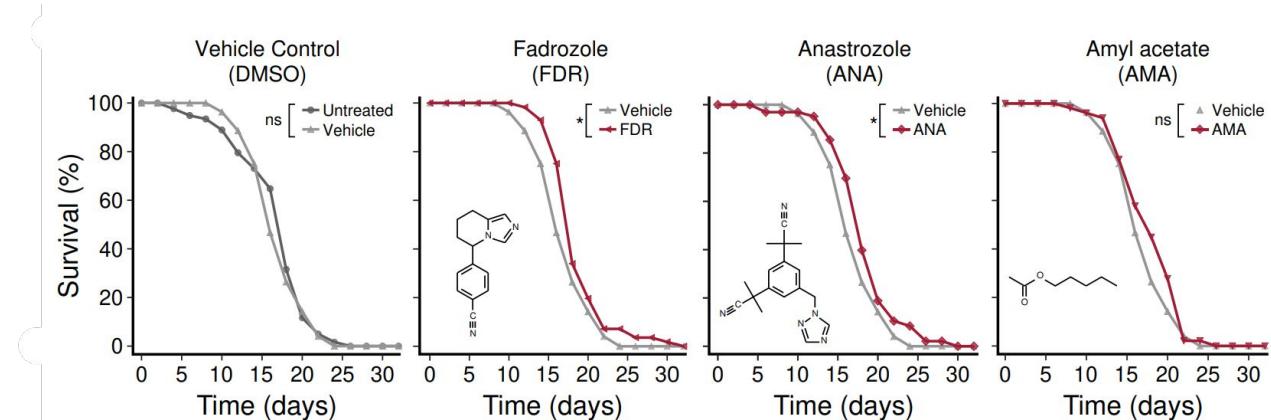
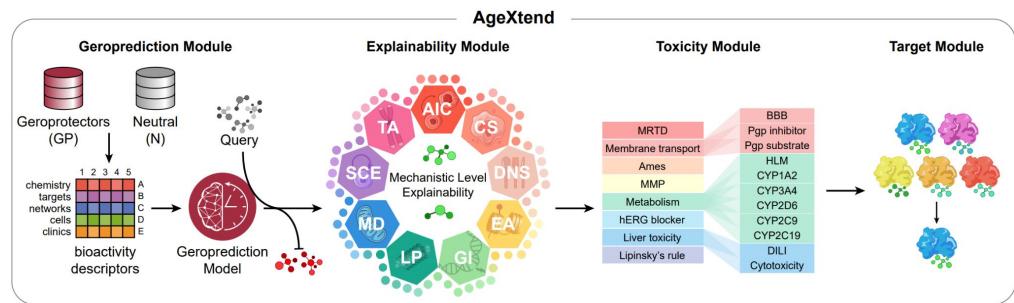
380 Accesses | 143 Altmetric | Metrics



Sakshi



Aayushi



# Deep Learning Unveiled Evolutionary Conservation of Aging-Related Cellular Phenotypes

Cell Reports



Resource

## scCamAge: A context-aware prediction engine for cellular age, aging-associated bioactivities, and morphometrics

Vishakha Gautam,<sup>1,4,\*</sup> Subhadeep Duari,<sup>1,4</sup> Saveena Solanki,<sup>1</sup> Mudit Gupta,<sup>1</sup> Aayushi Mittal,<sup>1</sup> Sakshi Arora,<sup>1</sup> Anmol Aggarwal,<sup>1</sup> Anmol Kumar Sharma,<sup>1</sup> Sarthak Tyagi,<sup>1</sup> Rathod Kunal Pankajbhai,<sup>1</sup> Arushi Sharma,<sup>1</sup> Sonam Chauhan,<sup>1</sup> Shiva Satija,<sup>1</sup> Suvendu Kumar,<sup>1</sup> Sanjay Kumar Mohanty,<sup>1</sup> Juhi Tayal,<sup>3</sup> Nilesh Kumar Dixit,<sup>1</sup> Debarka Sengupta,<sup>1,2</sup> Anurag Mehta,<sup>3</sup> and Gaurav Ahuja<sup>1,2,5,\*</sup>

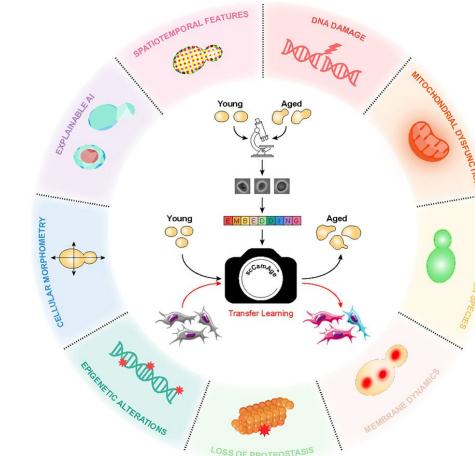
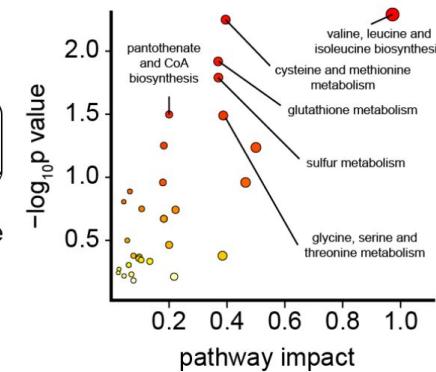
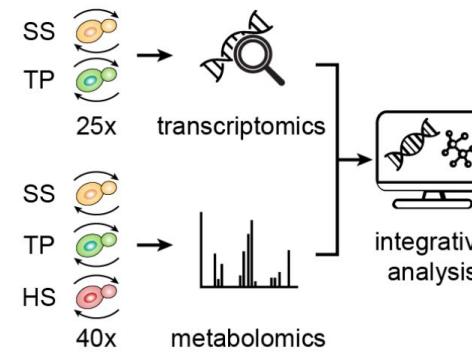
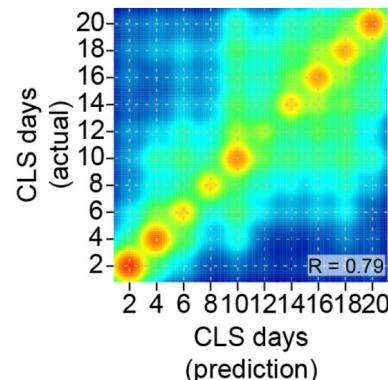
<sup>1</sup>Department of Computational Biology, Indraprastha Institute of Information Technology - Delhi (IIIT-Delhi), Okhla, Phase III, New Delhi 110020, India



Vishakha



Subhadeep



# Sex, Suicide or Survival - A decision governs by endogenous Sterols

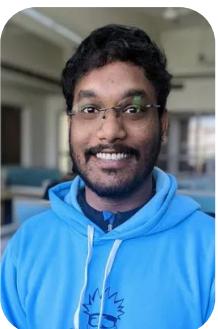
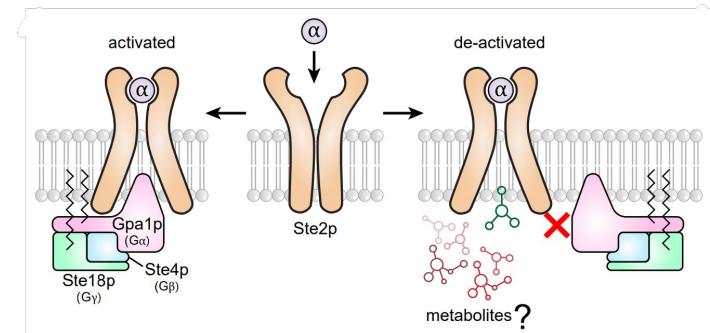


Biochemistry and Chemical Biology

## Deep Learning Reveals Endogenous Sterols as Allosteric Modulators of the GPCR-G<sub>a</sub> Interface

Sanjay Kumar Mohanty, Aayushi Mittal, Namra, Aakash Gaur, Subhadeep Duari, Saveena Solanki, Anmol Kumar Sharma, Sakshi Arora, Suvenu Kumar, Vishakha Gautam, Nilesh Kumar Dixit, Karthika Subramanian, Tarini Shankar Ghosh, Debarka Sengupta, Shashi Kumar Gupta, Natarajan Arul Murugan, Deepak Sharma, Gaurav Ahuja

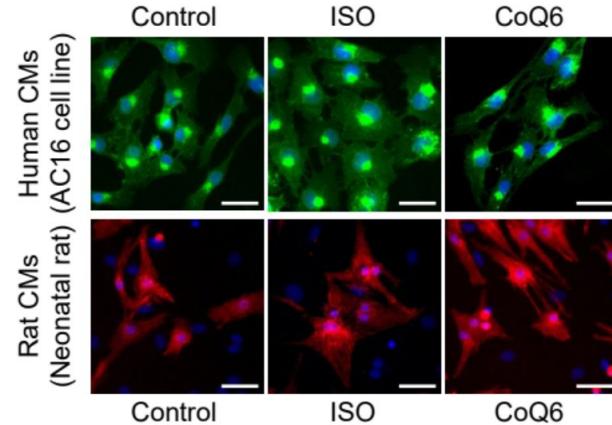
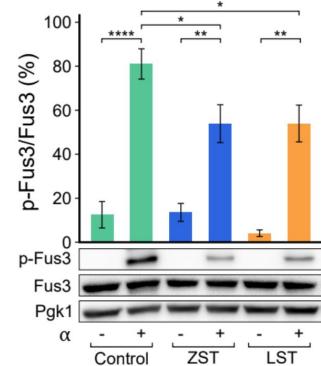
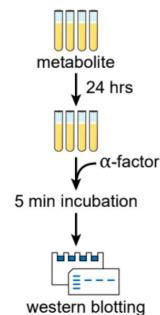
Department of Computational Biology, Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi), New Delhi, India • Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, India • CSIR-Institute of Microbial Technology, Chandigarh, India • Pharmacology Division, CSIR-Central Drug Research Institute, Lucknow, India • Infosys Centre for AI, Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi), New Delhi, India



Sanjay



Aayushi



# Thank You