DNA Tools & Gene Regulation

Lecture – 7

Prof. Sanjeeva Srivastava
BSBE, IIT Bombay

Outline

- I. COVID-19 Diagnostics
  - Genome sequencing
  - RT-PCR applications
  - Molecular diagnostics
- II. Regulation of Gene Expression
  - Transcription
  - Translation
- III. Gene Regulation Models
  - *Trp* operon
  - *Lac* operon

- I. COVID-19 Diagnostics
  - Genome sequencing
  - RT-PCR applications
  - Molecular diagnostics
- II. Regulation of Gene Expression
  - Transcription
  - Translation
- III. Gene Regulation Models
  - Trp operon
  - Lac operon

Molecular Biology Tools - COVID-19

A new coronavirus associated with human respiratory disease in China

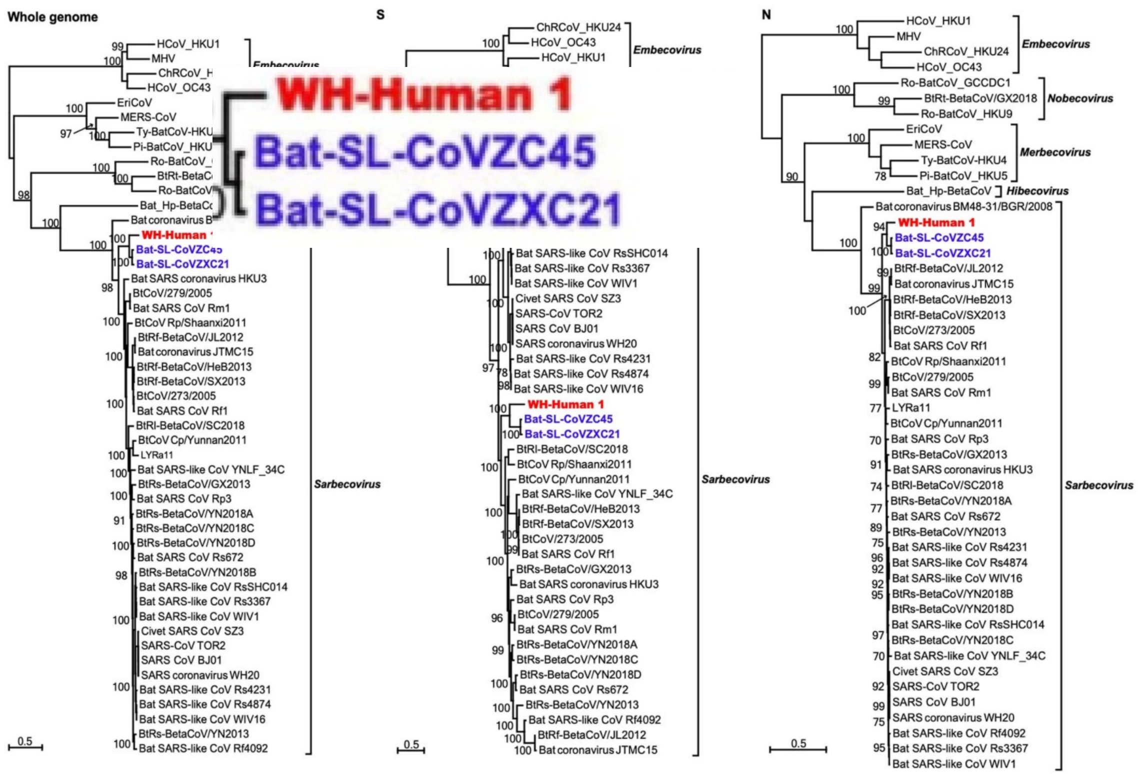
nature

Article | Open Access | Published: 03 February 2020

A new coronavirus associated with human respiratory disease in China

- Metagenomic RNA sequencing - bronchoalveolar lavage fluid from a patient identified new RNA virus strain (family *Coronaviridae)*, designated as 'WH-Human 1' coronavirus ('2019-nCoV').

- Phylogenetic analysis revealed that the virus was most closely related to a group of SARS-like coronaviruses previously been found in bats in China



Phylogenetic trees of the nucleotide sequences of the whole genome, and S and N genes of WHCV and related coronaviruses.

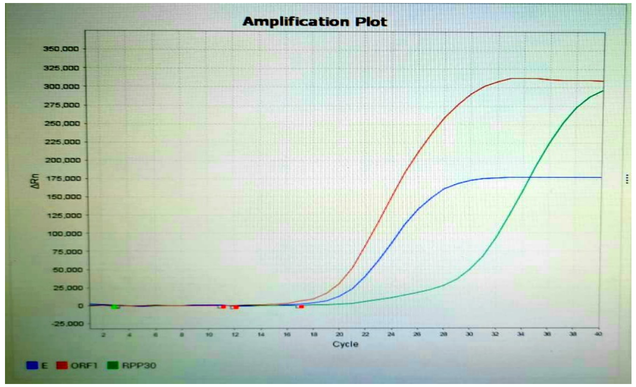
Schematic of Sequencing

1.

2.

Sample Collection (BALF)

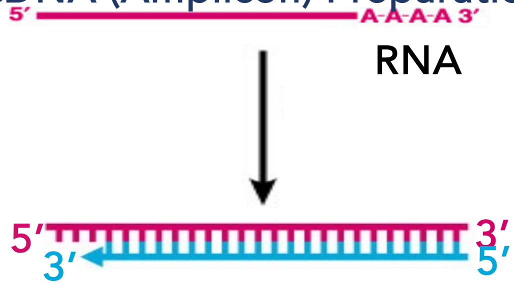
RNA Extraction



3.

4.

cDNA (Amplicon) Preparation

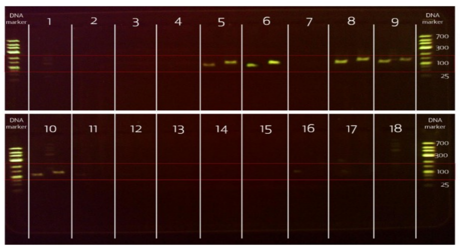


RNA

cDNA

5.

Amplicon Quantification



6.

Quality Check (Electrophoresis)

7.

Next Generation Sequencer

8.

Viral Genome sequence

RNA was extracted from BALF and a meta-transcriptomic library was constructed for pair-end sequencing using an Illumina MiniSeq

Sanger's Method for Sequencing DNA (1)

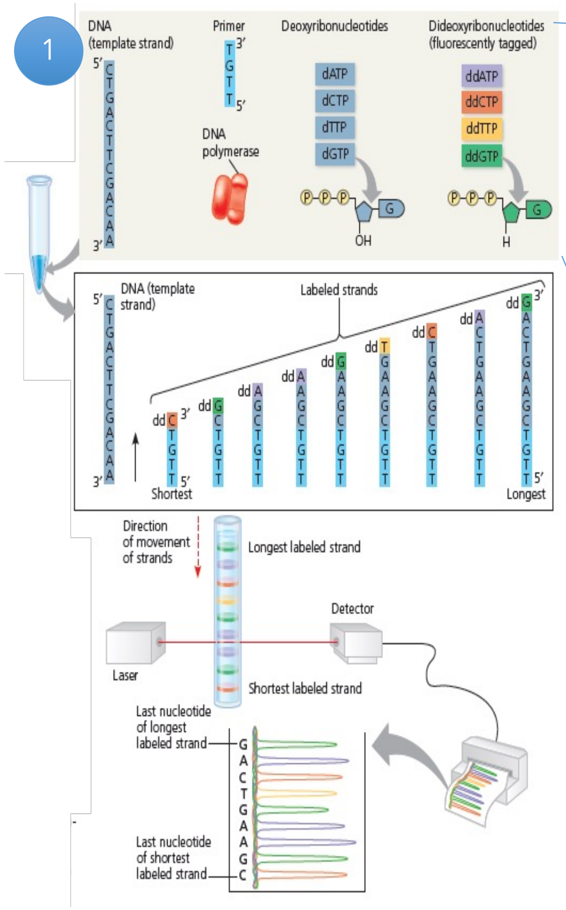
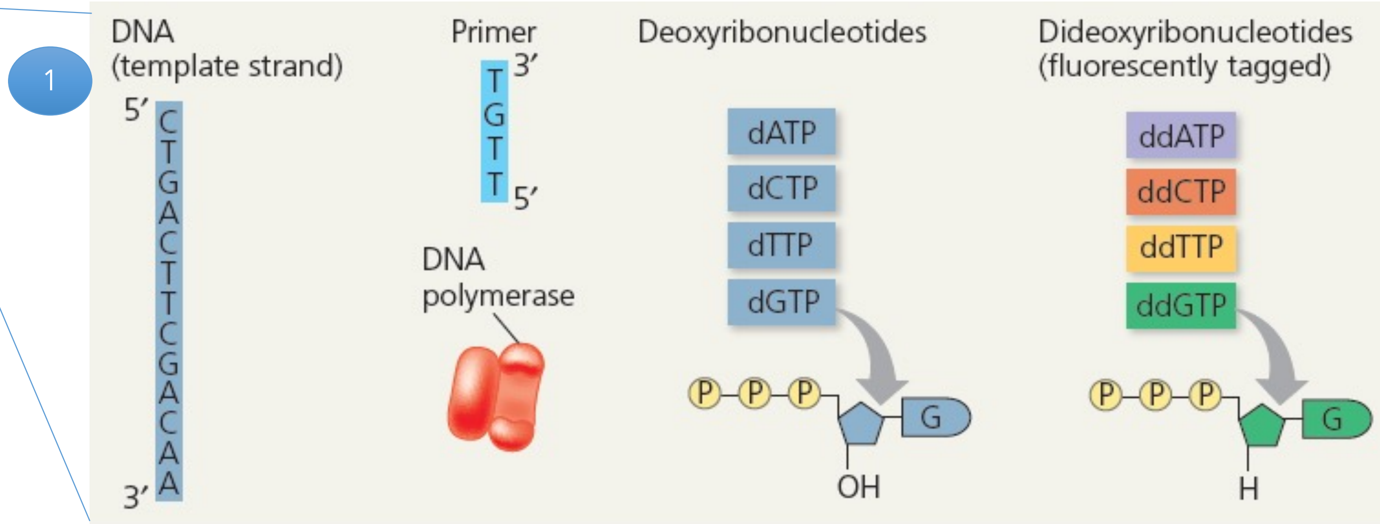
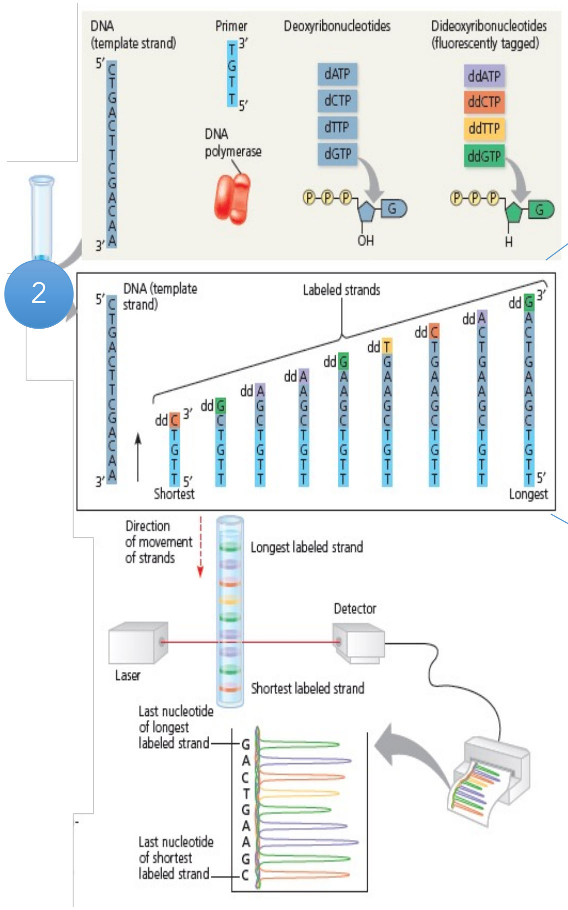


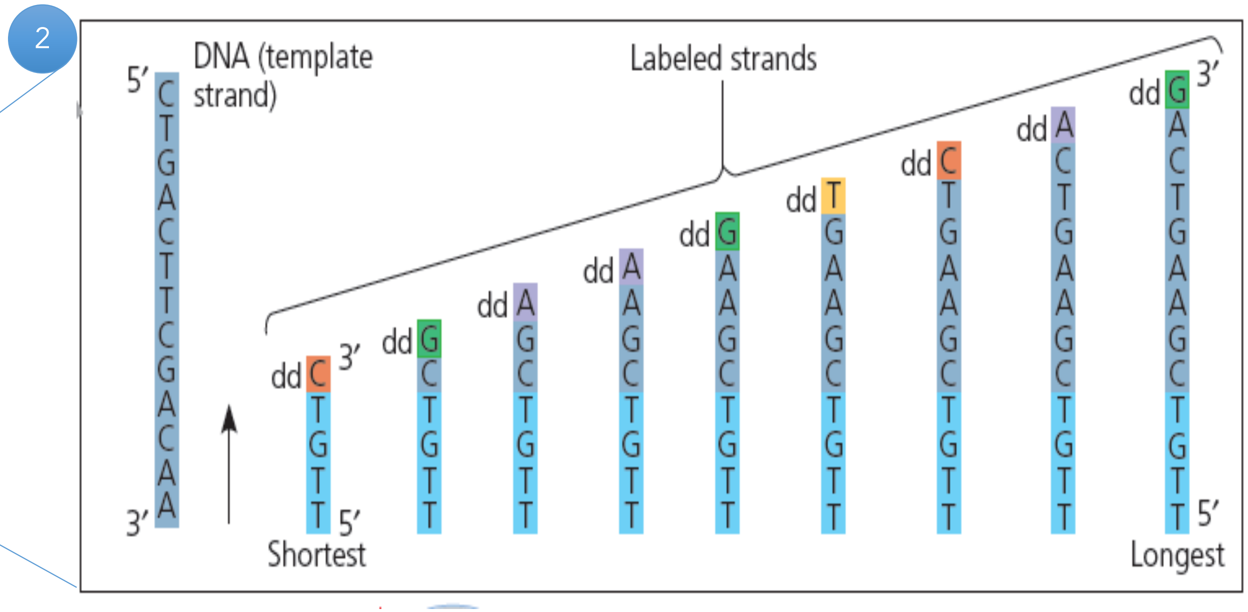
Figure 20.3



- DNA is denatured into single strands. Incubate a primer designed to base-pair with known 3′ end of the template strand along-with DNA polymerase, 4 dNTPs, & 4 ddNTPs, tagged with specific fluorescent molecule.
- Dideoxyribonucleotide (ddNTP), a modified deoxyribonucleotide (dNTP), its incorporation terminates a growing DNA strand because it lacks a 3′OH group

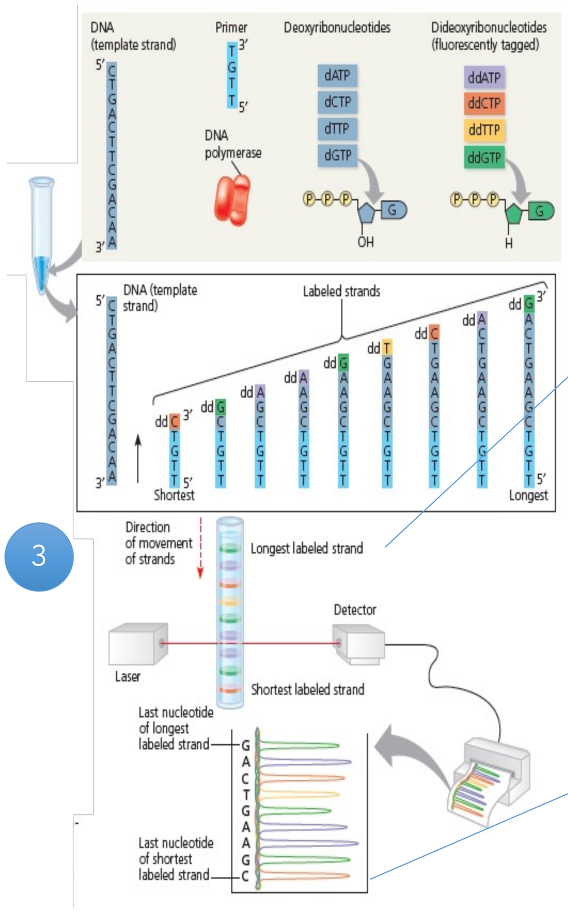
Sanger's Method for Sequencing DNA (2)

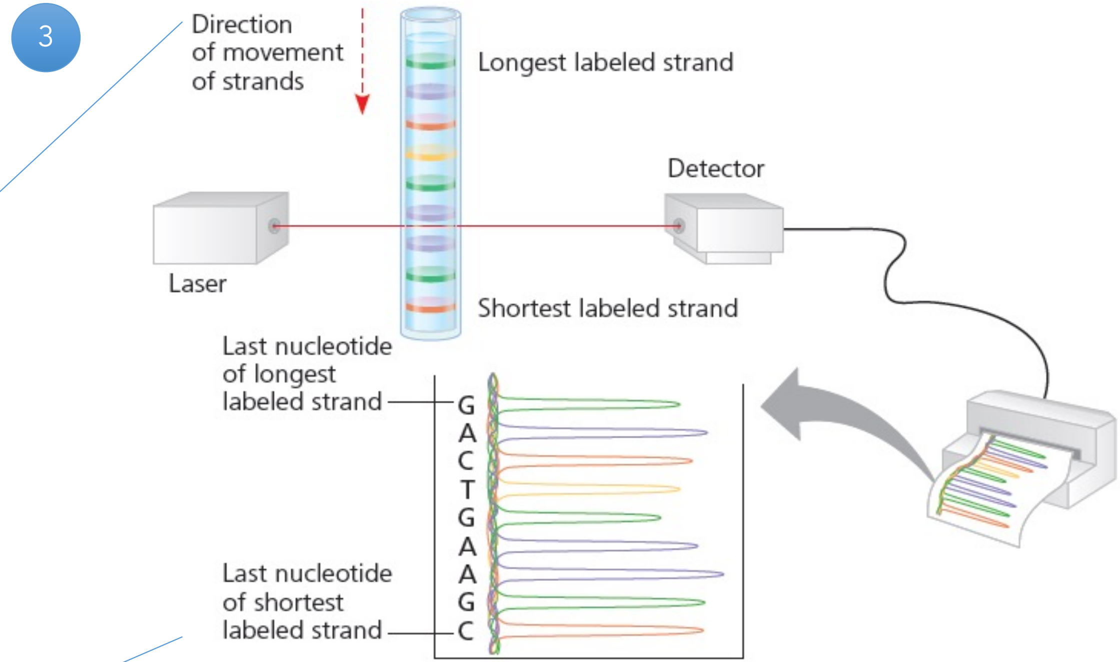




- Synthesis of each new strand starts at the 3′ end of the primer. The incorporated ddNTP prevents further elongation of the strand.

Sanger's Method for Sequencing DNA (3)





- The labeled strands in mixture are separated by passage through a gel. A fluorescence detector senses the color of each fluorescent tag as the strands come through.

Next Generation Sequencing

• Today's machines can sequence 700–900 million nucleotides in just few hours and used for large and fast sequencing.

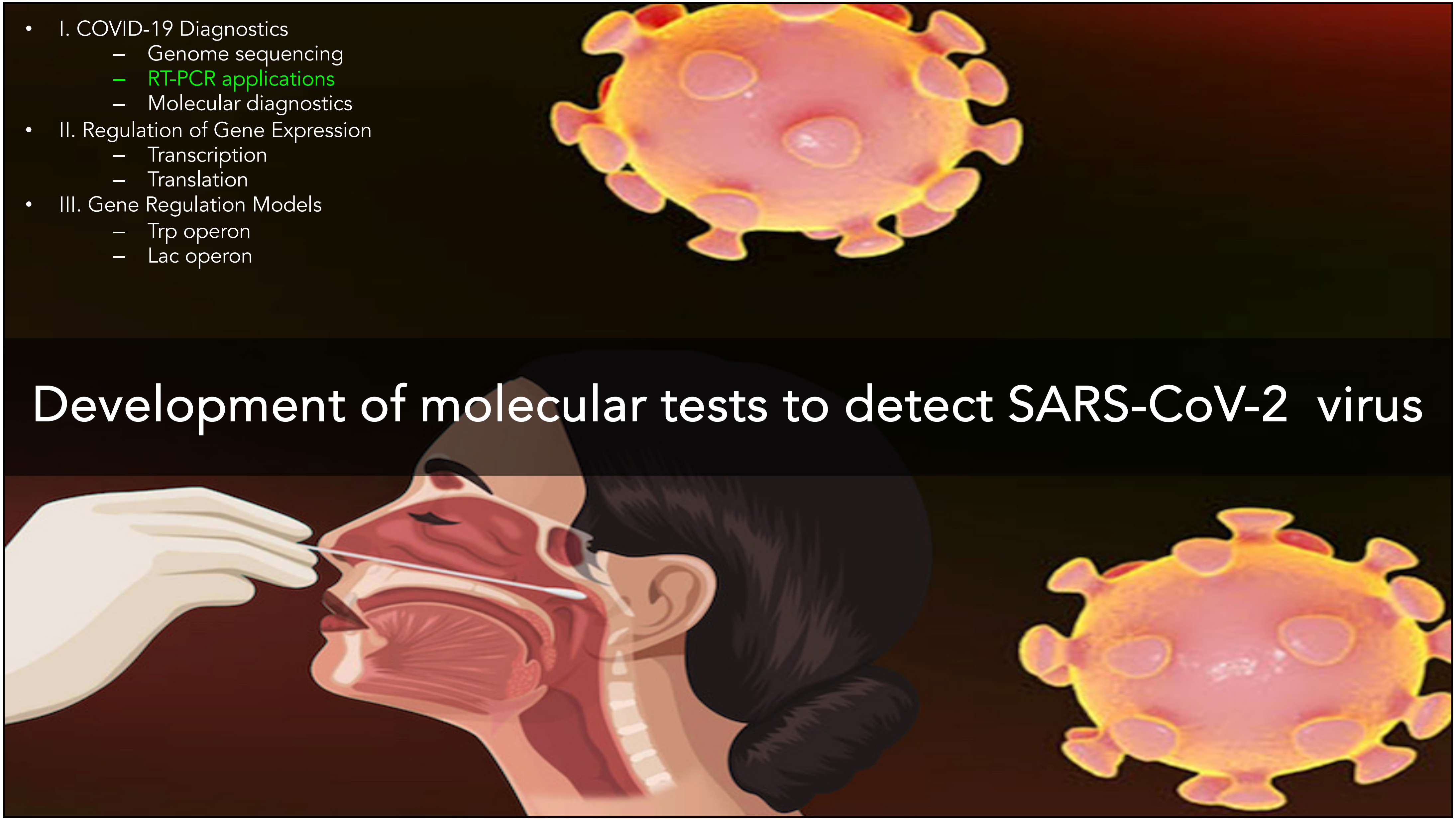
• First-generation sequencing - chain termination method or Chemical method

• Next generation or 2nd generation sequencing - by ligation or synthesis, including pyro-sequencing and reversible chain termination

• 3rd generation sequencing – to lower the cost, by using **scanning tunneling electron microscope**, **fluorescence resonance energy transfer, single-molecule detection and protein nanopores**

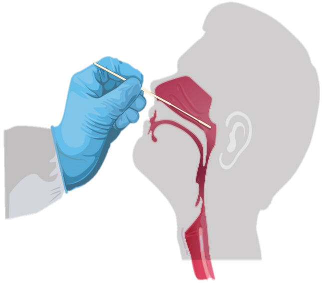
Next Generation Sequencing

- DNA fragments are amplified to yield an enormous number of identical fragments.
- A specific strand of each fragment is immobilized, & complementary strand is synthesized, one nucleotide at a time.
- It enables electronic monitors to identify in real time which of the four nucleotides is added; this method is thus called "sequencing by synthesis".

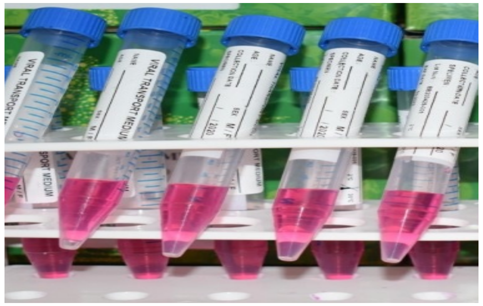


• I. COVID-19 Diagnostics
– Genome sequencing
– RT-PCR applications
– Molecular diagnostics
• II. Regulation of Gene Expression
– Transcription
– Translation
• III. Gene Regulation Models
– Trp operon
– Lac operon
Development of molecular tests to detect SARS-CoV-2 virus

COVID-19 Testing: RT-PCR



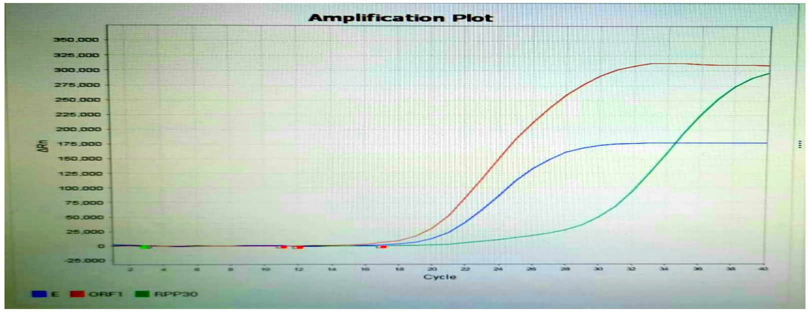
Sample
(Nasopharyngeal
Swab)



Transferred to VTM

RNA Extraction

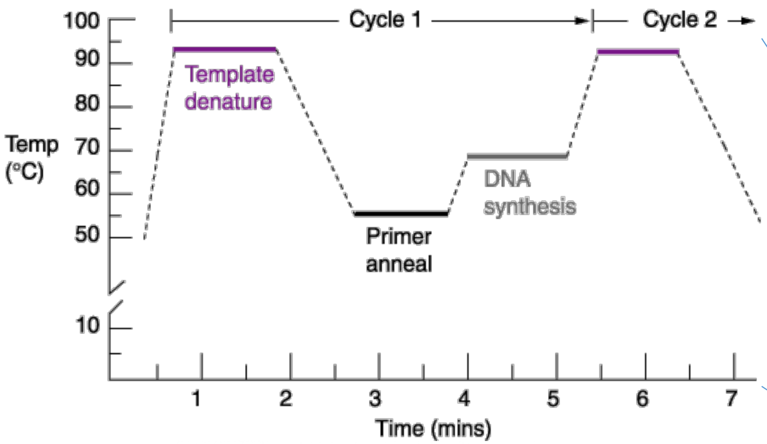
Quantitative PCR
(QuantStudio 12)

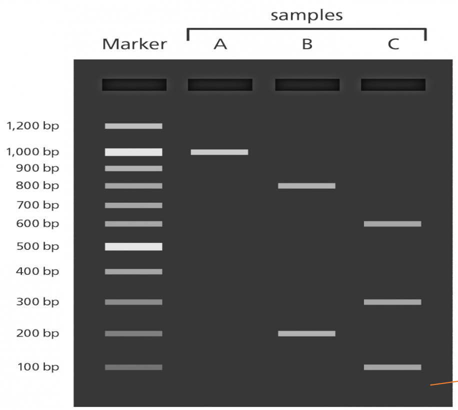


Q PCR Result (E and Orf gene Positive)

One method that is widely used is called reverse transcriptase polymerase chain reaction (RT-PCR) or Real-time PCR. It uses enzyme reverse transcriptase (RT) with PCR and gel electrophoresis

RT-PCR Analysis





1 **cDNA synthesis** is carried out by incubating the mRNAs with reverse transcriptase and other necessary components.

2 **PCR amplification** of the sample is performed using primers specific to the *Drosophila* gene of interest.

3 **Gel electrophoresis** will reveal amplified DNA products only in samples that contained mRNA transcribed from the specific *Drosophila* gene.

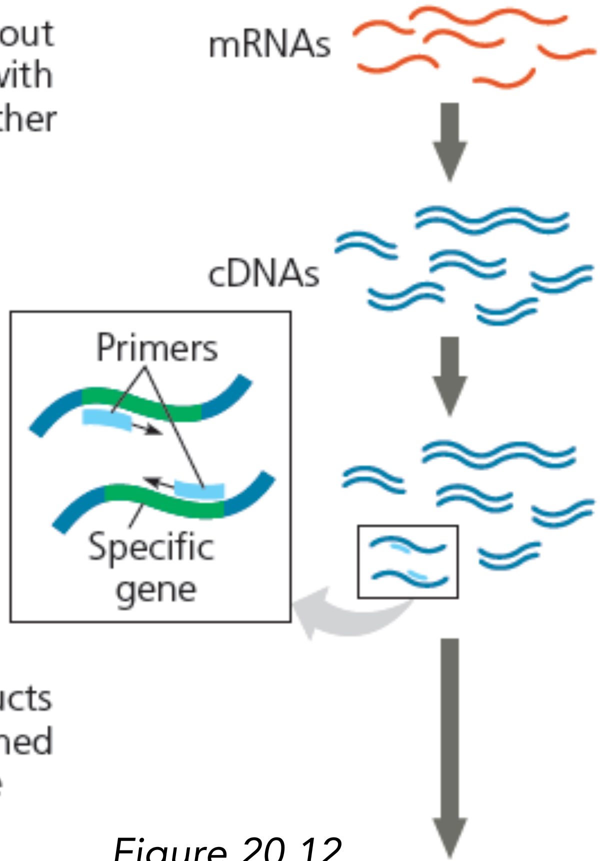


Figure 20.12

Complementary DNA (cDNA) Synthesis

1. mRNA & Reverse transcriptase added

3′ end of mRNA has a stretch of adenine - poly-A tail, which allows use of a short complementary strand of poly-dT primer for synthesis of DNA strand

2. RT makes first DNA strand (mRNA template; poly-dT DNA primer)

3. mRNA degraded by enzyme

4. DNA polymerase synthesizes 2nd DNA strand

5. cDNA is synthesized which carries complete coding sequence of gene but no introns

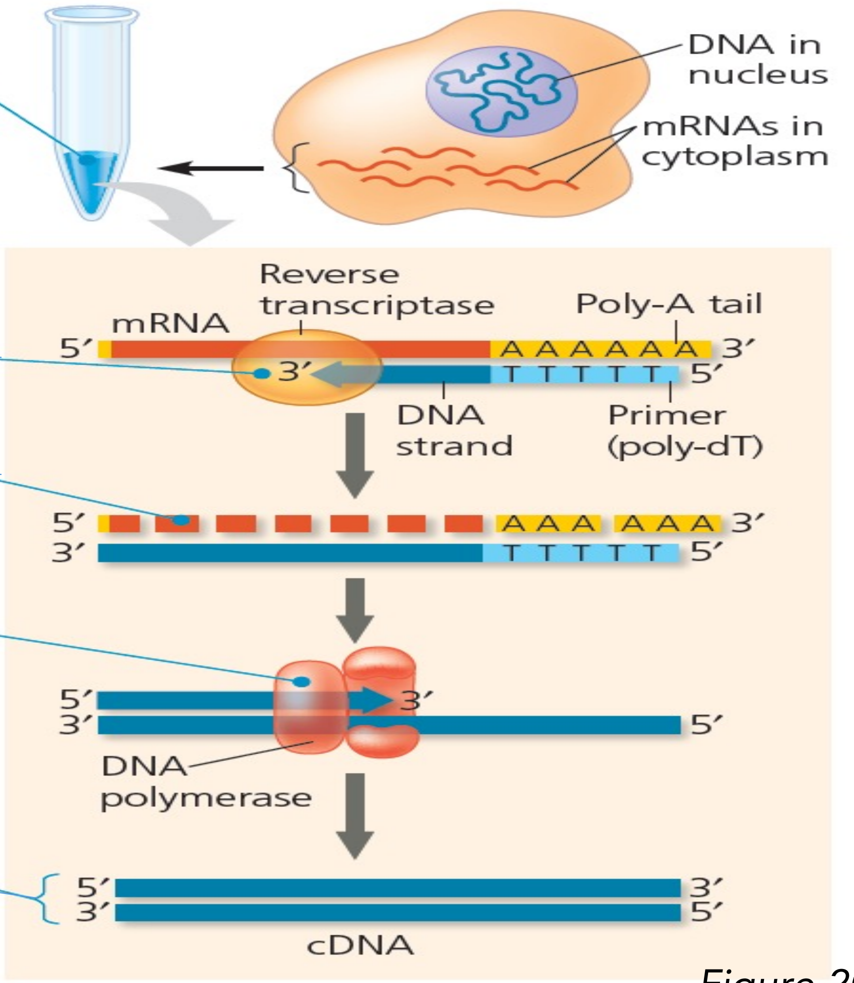
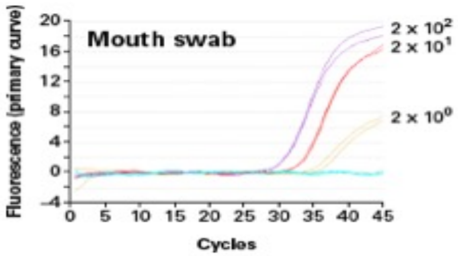
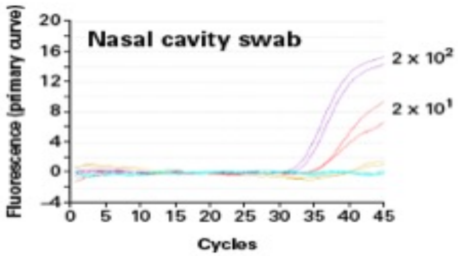
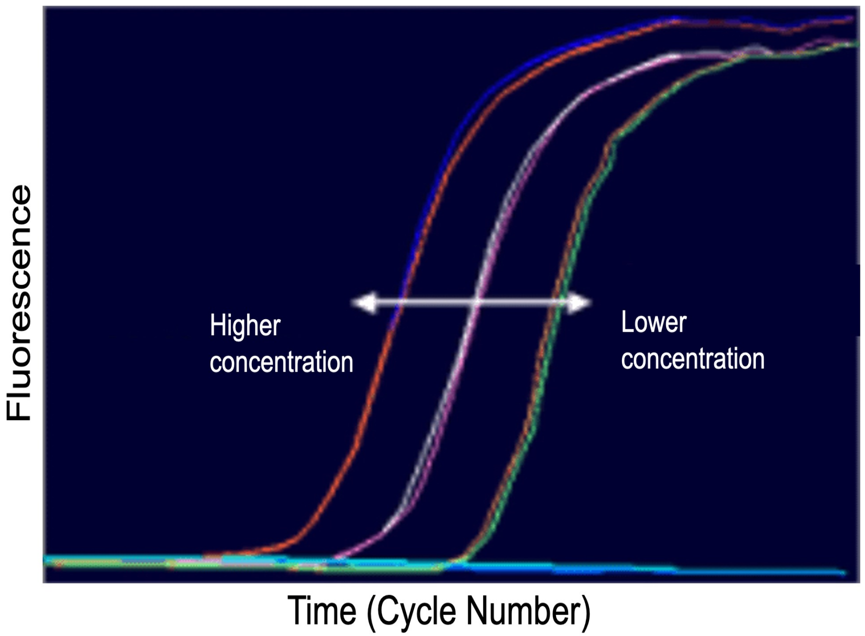


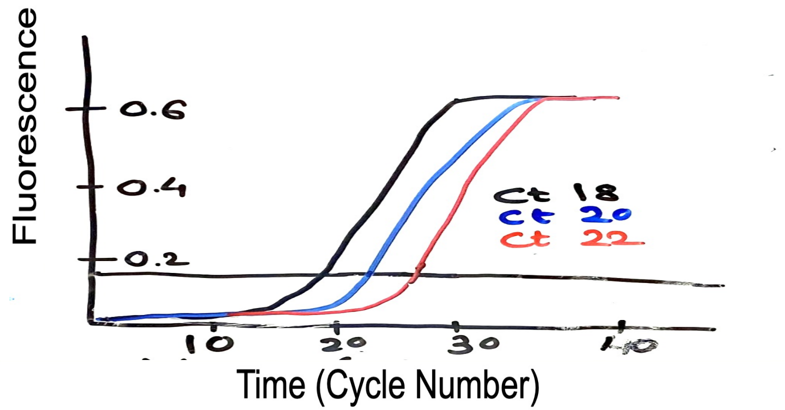
Figure 20.11

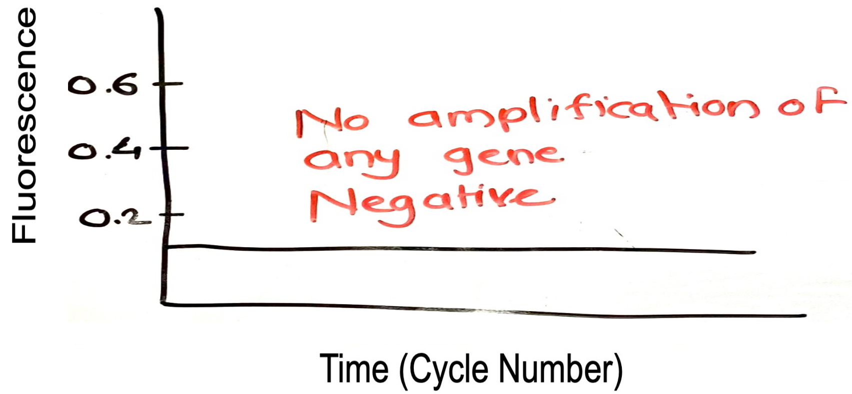
COVID-19 detection: Real-time PCR









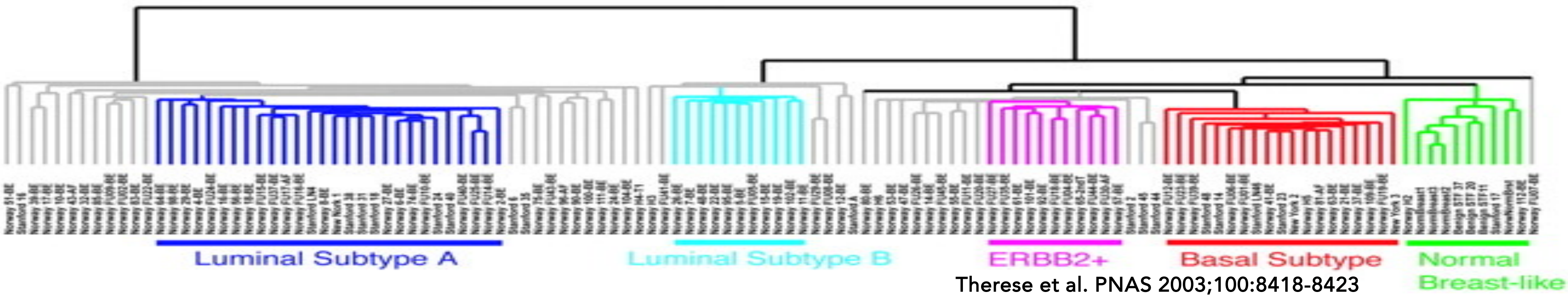


Quantitative expression analysis in real time manner

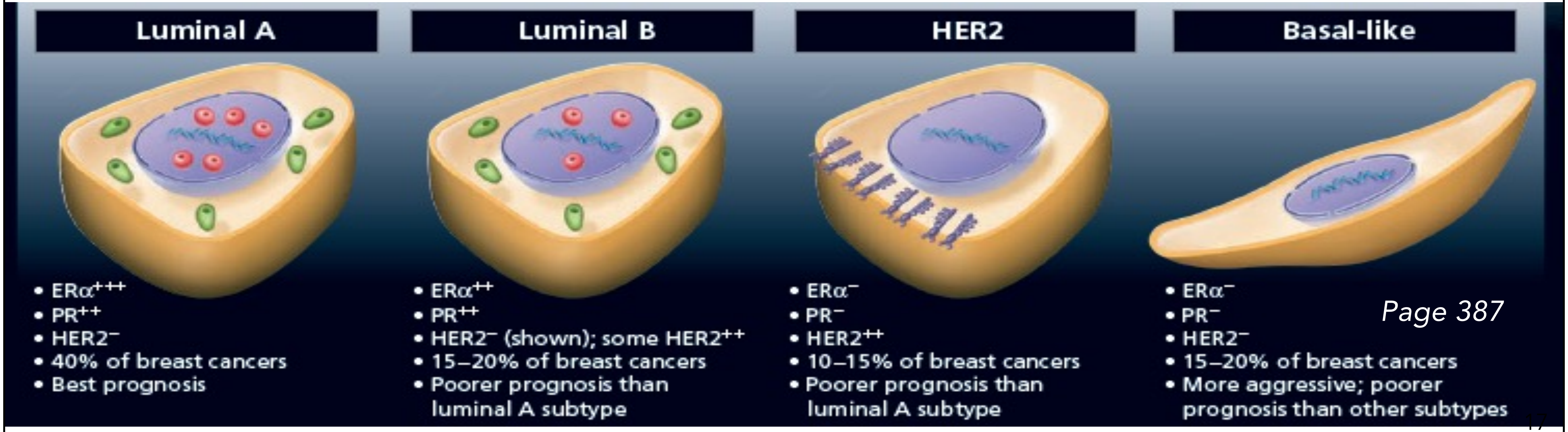
- I. COVID-19 Diagnostics
  - Genome sequencing
  - RT-PCR applications
  - Molecular diagnostics
- II. Regulation of Gene Expression
  - Transcription
  - Translation
- III. Gene Regulation Models
  - Trp operon
  - Lac operon

Molecular diagnostics

Gene Expression Analysis: Breast Tumor Subtypes



Therese et al. PNAS 2003;100:8418-8423



Page 387

Food for thought: A Modern Example of Genetic Analysis

- BRCA1 gene is tumor suppressor gene, which is responsible for DNA repair and maintaining the stability of cell's genetic information.

- BRCA1 mutations are dominant i.e. the presence of the mutations can give rise to breast cancer

  - Angelina Jolie had mutations in BRCA1 that predispose to breast cancer

  - **Assuming Angelina Jolie is heterozygous for BRCA1 mutations and Brad Pitt has no BRCA1 mutations, draw a Punnett square to predict the genotypes and phenotypes for breast cancer of their children?**

Personalized Diagnostics

- FLT3 (FMS-like tyrosine kinase 3) over-expression
- FLT3 inhibitor sunitinib (FDA - renal cell cancer)

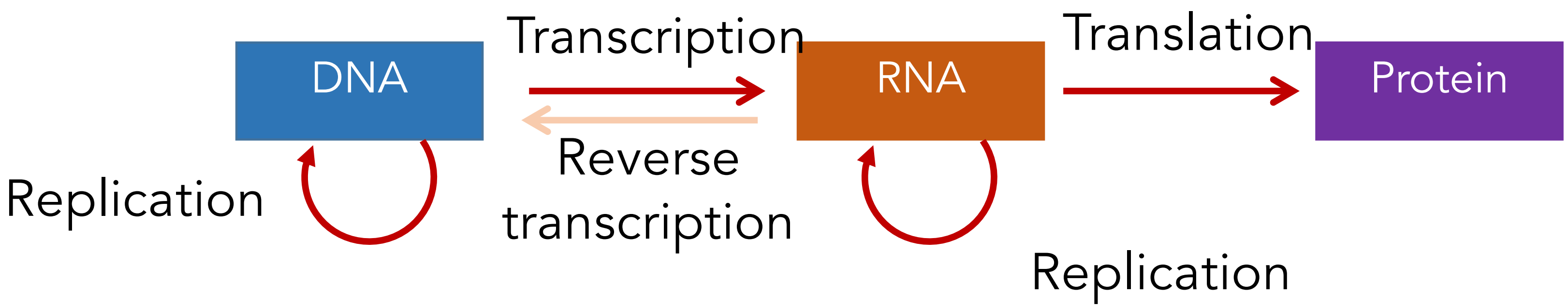
- I. COVID-19 Diagnostics
  - Genome sequencing
  - RT-PCR applications
  - Molecular diagnostics
- II. Regulation of Gene Expression
  - Transcription
  - Translation
- III. Gene Regulation Models
  - Trp operon
  - Lac operon

Regulation of gene expression

HOW does a cell decode only the relevant information contained in chromosomes?
By **regulating** the transcription of genes (i.e., flow of information)

Central dogma

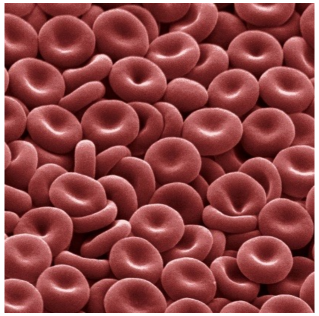
Central dogma is the concept that cells are governed by a cellular chain of command



Why might information reach a dead-end at proteins?
The "language" of DNA and RNA is the same (nucleic acids) while that of proteins is different (amino acids).

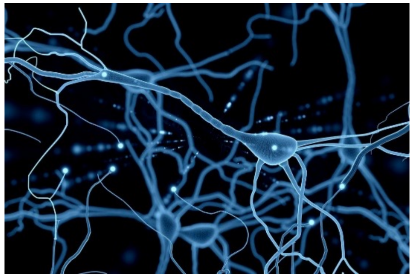
Different cells have different phenotypes

Red blood cells
Carry oxygen,
Easily deformable

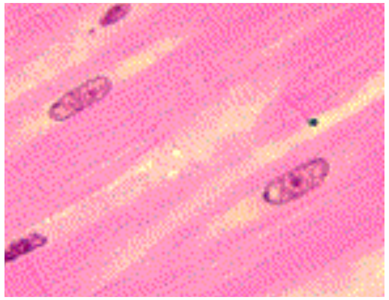


$\sim 10^{13}$ cells

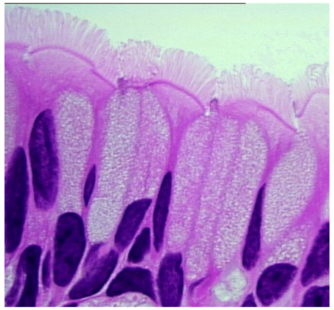
Nerve cells
Electrically active
Can be very long



Muscle cells
Contract and relax
High tensile strength



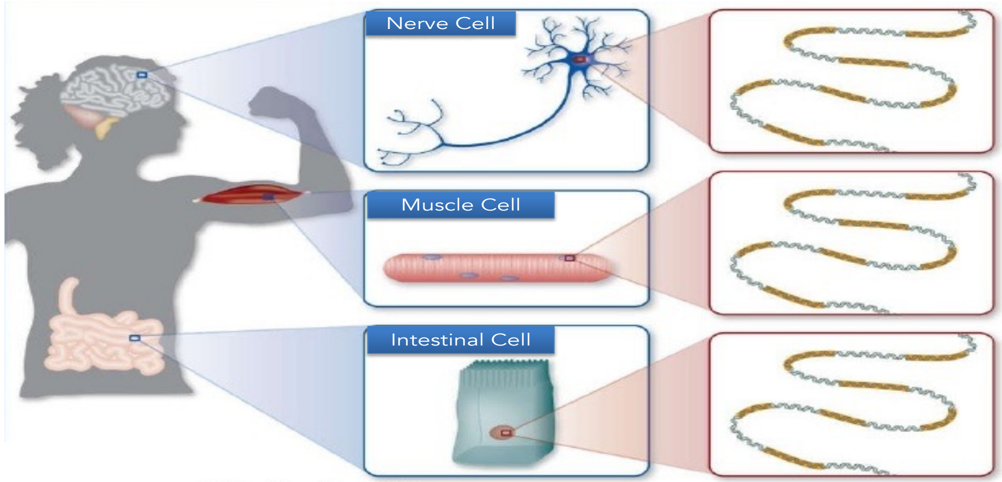
Respiratory cells
Exchange gases efficiently
Soft and spongy



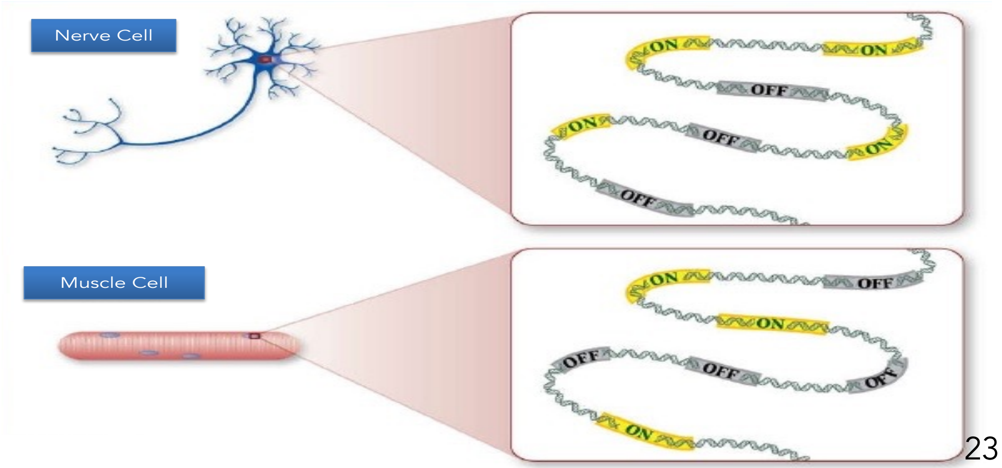
Gene Expression

- We have over 250 different types of cells
- All of them contain identical genome
  - But each cell type is specialized to perform different jobs

- <20% of the genes are expressed in a typical cell type
  - Differentiation of cells depends solely upon the set of genes expressed

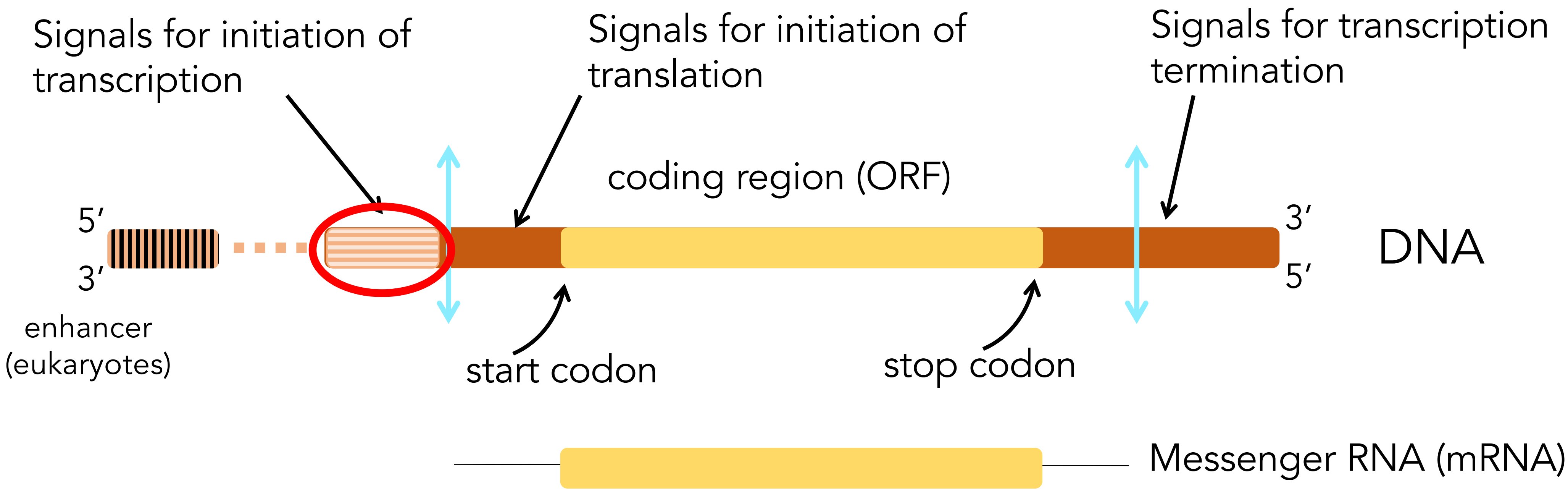


Each cell contains same genetic information!



Ref NIH National Human Genome Research Institute

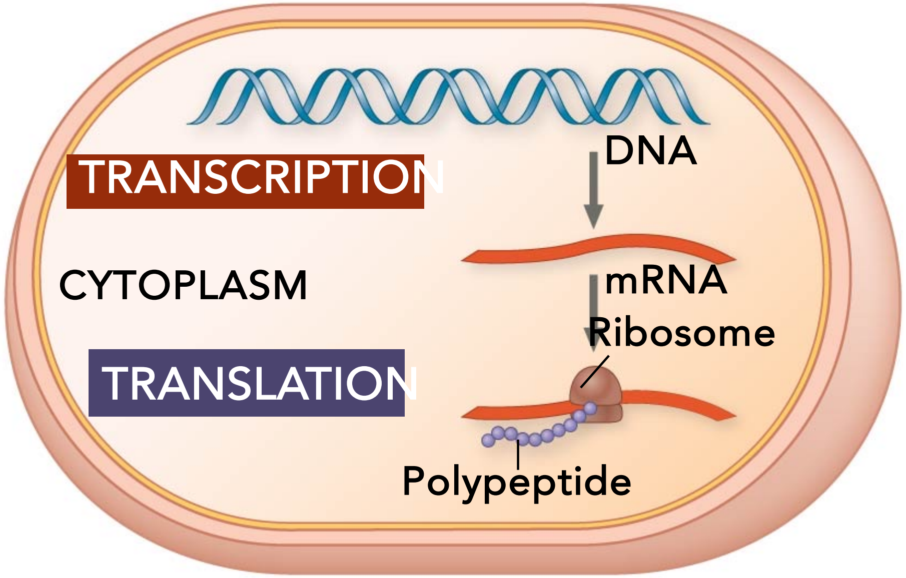
Gene: stretch of DNA that encodes a protein



A gene includes the entire sequence represented in mRNA

Gene expression

- "Gene expression" includes transcription & translation
- **Transcription** is synthesis of RNA using information in DNA.
- **Translation** is synthesis of a polypeptide using information in the mRNA.
- Transcription and translation occur in all organisms—those that lack a membrane-bounded nucleus (bacteria) and those that have one (eukaryotes).
- In a bacterial cell, which lacks a nucleus, *mRNA produced by transcription is immediately translated* without additional process



Bacterial cell

Biosynthesis of the polypeptide (protein) encoded by a gene

Figure 17.3

Transcription: Step 1 of 3

- RNA polymerase binds to promoter
- DNA unwinds
- Synthesis of RNA transcript begins

Promoter is a site where RNA polymerase can bind to DNA and begin transcription

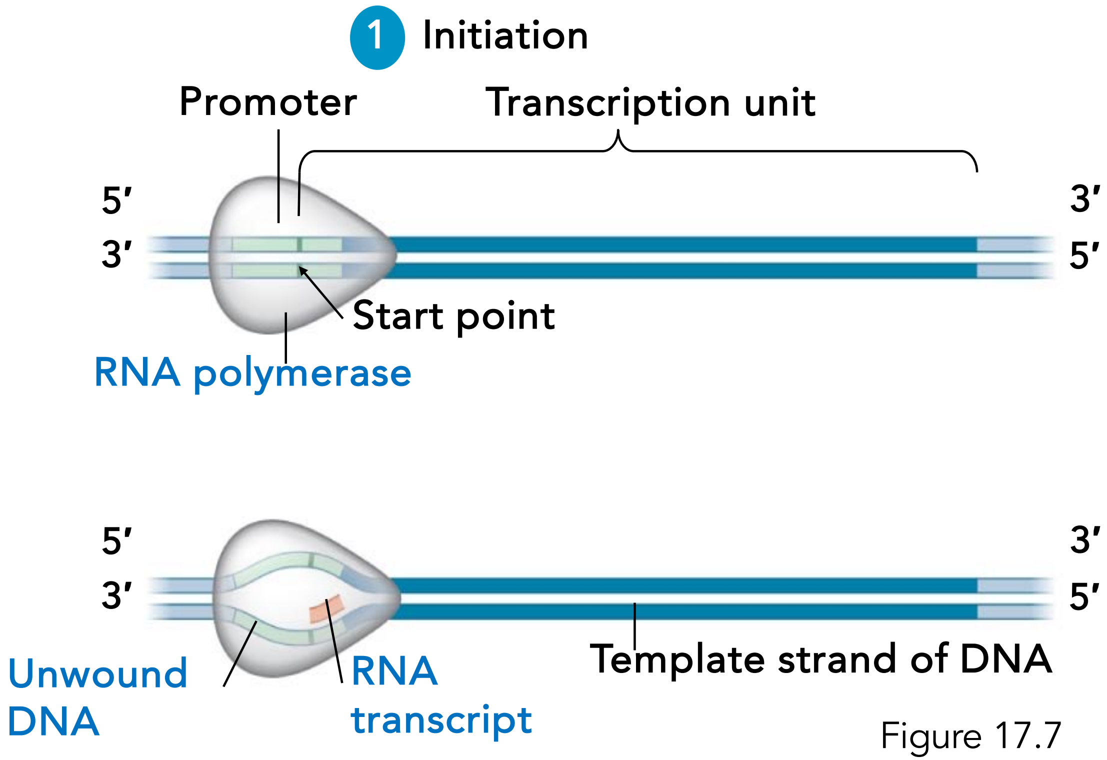
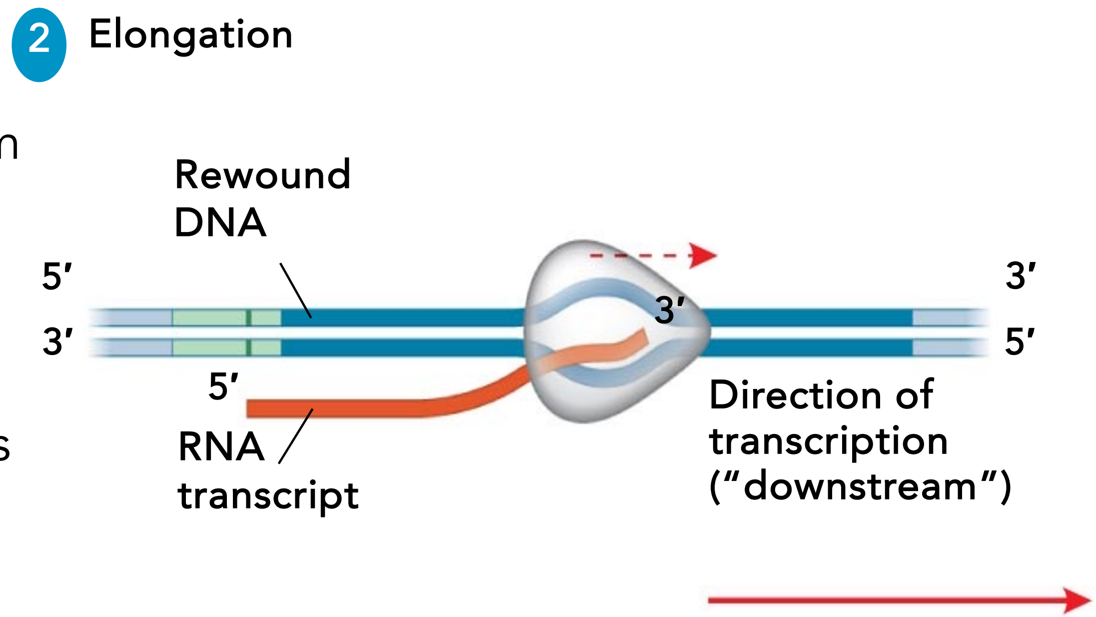


Figure 17.7

- **Initiation.** After RNA polymerase binds to promoter, DNA strands unwind, and polymerase initiates RNA synthesis at the start point on template strand.

Transcription: Step 2 of 3

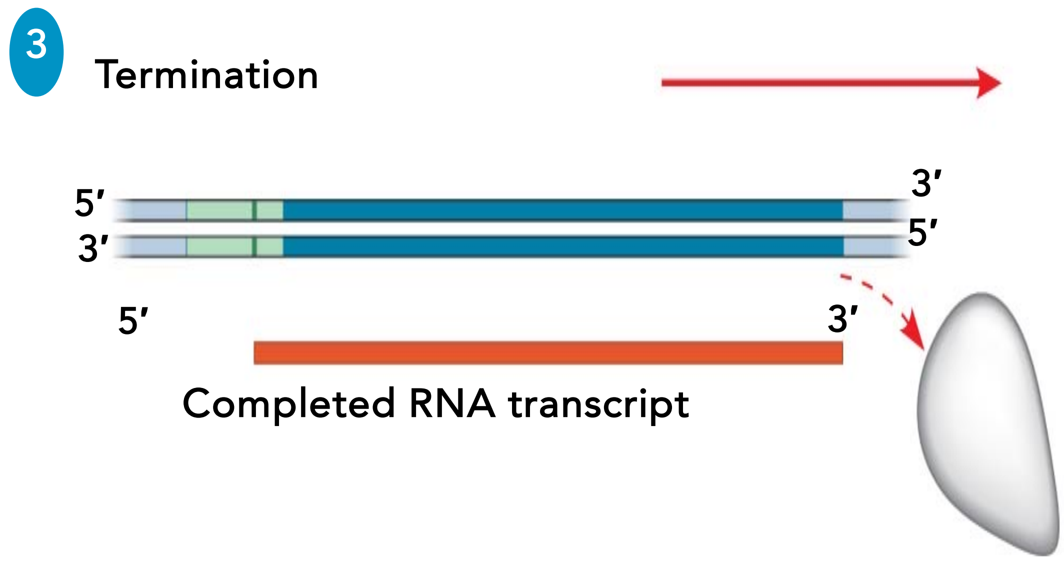
- RNA polymerase moves downstream
- DNA keeps unwinding
- RNA transcript grows
- Unwound DNA, after the message is transcribed, winds back



- **Elongation.** The polymerase moves downstream, unwinding DNA and elongating RNA transcript 5′ → 3′. In the wake of transcription, the DNA strands re-form a double helix.

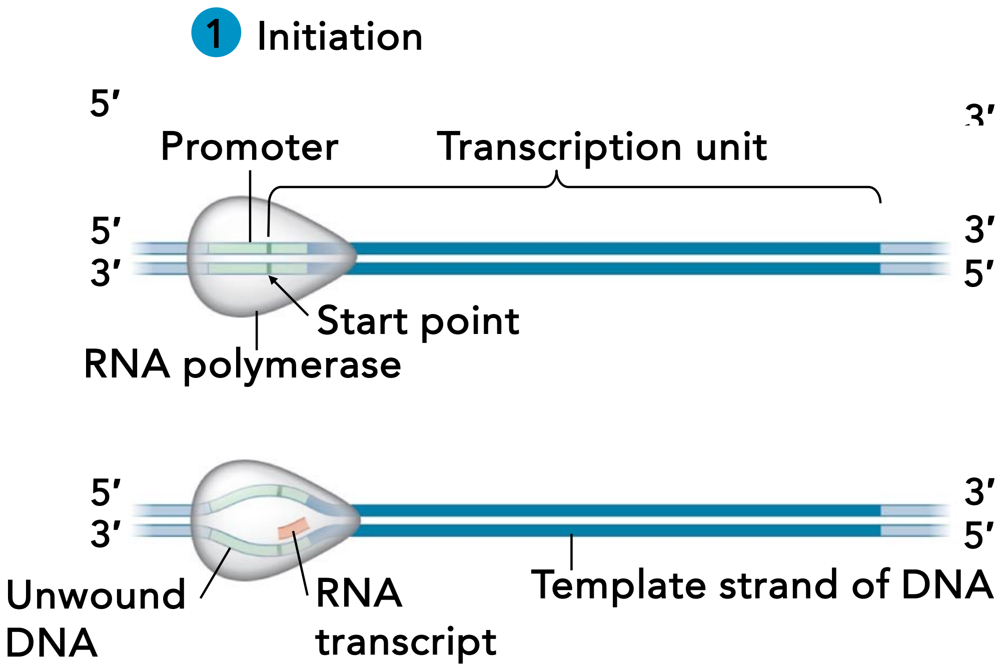
Transcription: Step 3 of 3

- RNA polymerase reaches the termination site
- Synthesis of RNA transcript ends
- RNA polymerase "falls off"
- Completed RNA transcript is also released



- **Termination.** RNA transcript is released & polymerase detaches from DNA.

How is transcription regulated?



"Go" or "no go"

- A major strategy for regulating transcription is to either make mRNA or not
- This decision is made at the step of initiation

Translation

- A molecule of **mRNA is moved through a ribosome**, *codons* (Triplets of nucleotides) are translated into amino acids, one by one.
- Interpreters are **tRNA** molecules, each type with a **specific nucleotide triplet called an *anticodon*** at one end and a corresponding amino acid at the other end.
- A tRNA adds its amino acid cargo to a growing **polypeptide chain** when anticodon hydrogen-bonds to complementary codon on mRNA.

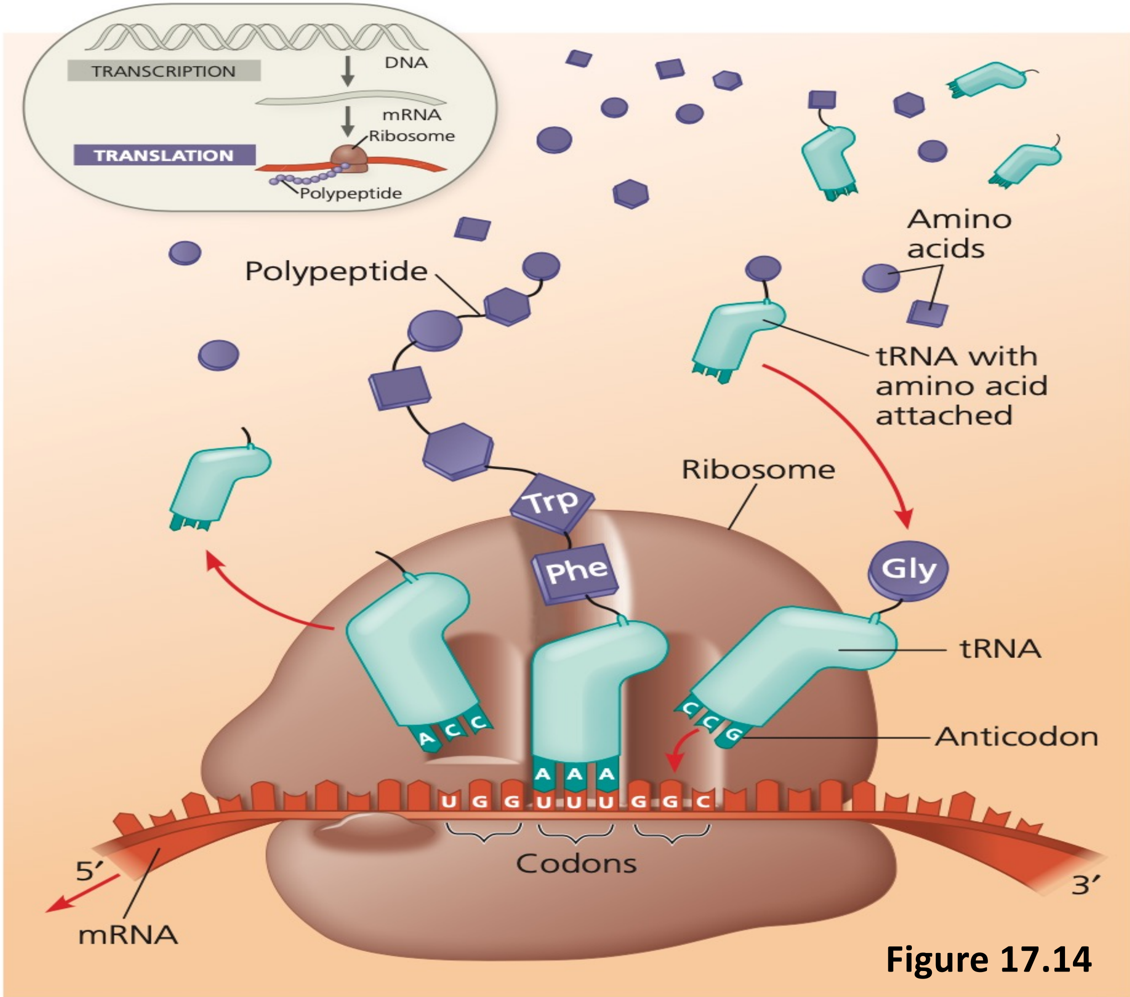
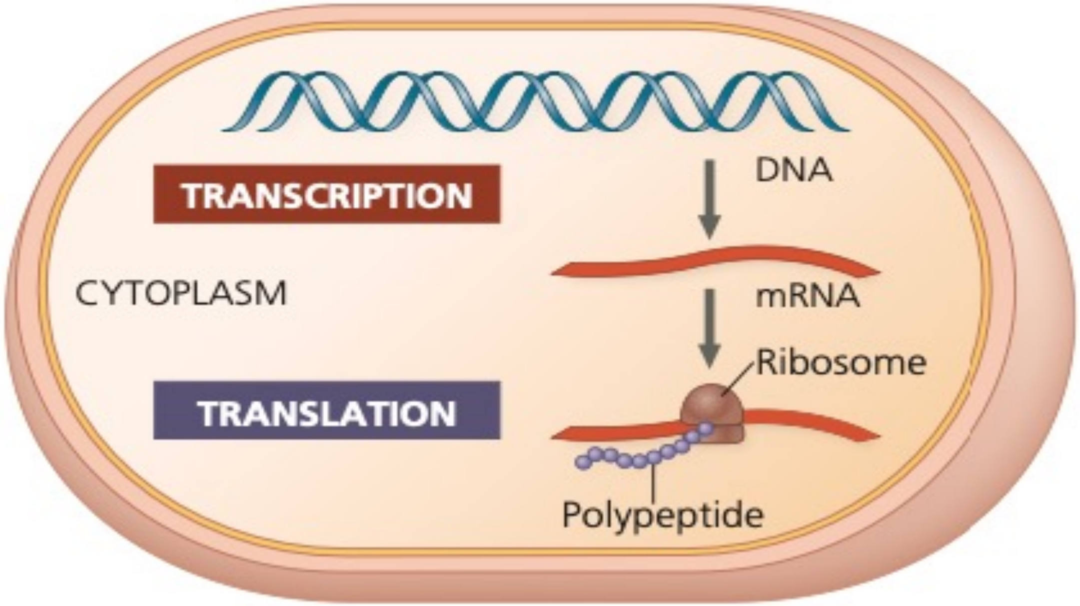
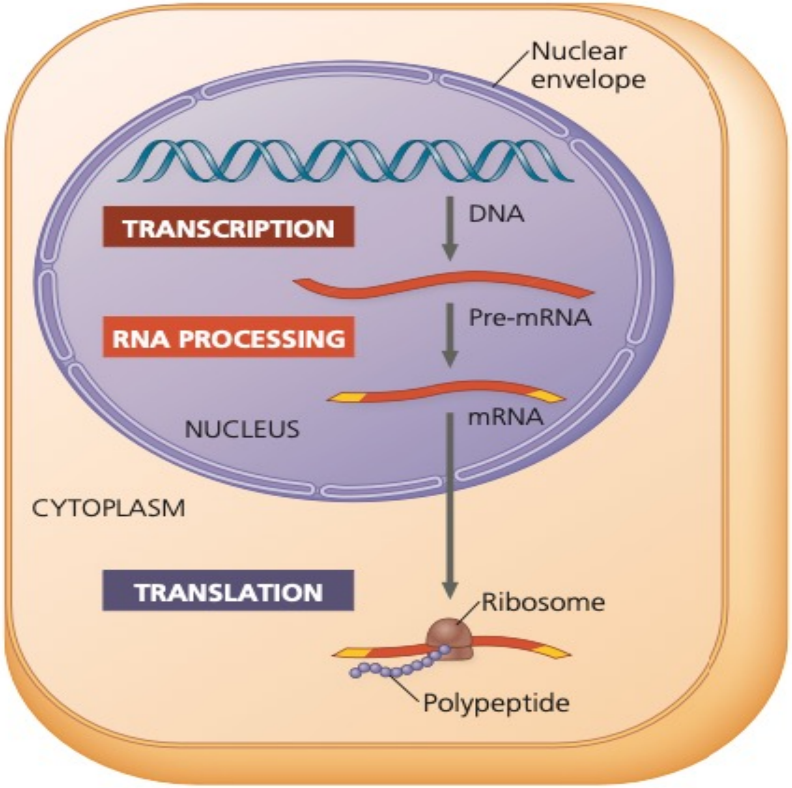


Figure 17.14

Comparison of Transcription & Translation in Prokaryotes & Eukaryotes



- Bacteria do not have nuclei, their DNA is not separated by nuclear membranes from ribosomes & other protein-synthesizing machinery.
- Lack of compartmentalization allows **translation of an mRNA to begin while its transcription is still in progress**



In a eukaryotic cell, by contrast, the **nuclear envelope separates transcription from translation** in space and time

- I. COVID-19 Diagnostics
  - Genome sequencing
  - RT-PCR applications
  - Molecular diagnostics
- II. Regulation of Gene Expression
  - Transcription
  - Translation
- III. Gene Regulation Models
  - Trp operon
  - Lac operon

Gene Regulation

A hallmark of prokaryotic and eukaryotic cells alike—from a bacterium to the cells of a fish - is their intricate and precise regulation of gene expression.

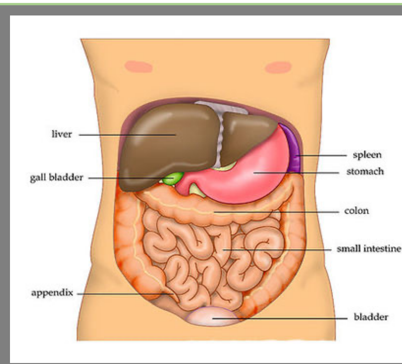
- I. COVID-19 Diagnostics
  - Genome sequencing
  - RT-PCR applications
  - Molecular diagnostics
- II. Regulation of Gene Expression
  - Transcription
  - Translation
- III. Gene Regulation Models
  - Trp operon
  - Lac operon

Operon Model

*In bacteria, genes are often clustered into **operons**, with one promoter serving several adjacent genes. An **operator** site on the DNA switches the operon ON or OFF resulting in coordinate regulation of the genes.*

Bacteria often respond to environmental change by regulating transcription

- Example: *Escherichia coli* cell in human colon
- Dependency on host's eating habits for nutrients
- Response to amino acid **tryptophan deficiency**: Activation of metabolic pathway to synthesize tryptophan
- Subsequent availability of tryptophan in host's meal leads to cessation of bacterial tryptophan production
- Efficient **resource utilization to avoid unnecessary production** when the substance is readily available in surrounding solution



Gut *E. coli*

Tryptophan available → Absorb and utilize

Tryptophan not available → Express requisite genes → Biosynthesize

Regulating product formation

- *E. coli* produce only proteins needed by that cell
    - *idea is similar to the expression of muscle genes only in muscle cells*
- *E. coli* can regulate production of tryptophan by
    - *(A) feedback inhibition and*
    - *(B) regulating gene expression*

Regulation of a metabolic pathway

1. Tryptophan accumulation in a cell leads to shutdown of further tryptophan synthesis

   - *Mechanism involves inhibition of enzyme activity through **feedback inhibition***

2. Regulatory role in gene expression: Control of genes encoding enzymes

- Adequate environmental tryptophan stops enzyme production

   - *Control at transcription level: Synthesis of mRNA for enzyme-coding genes*

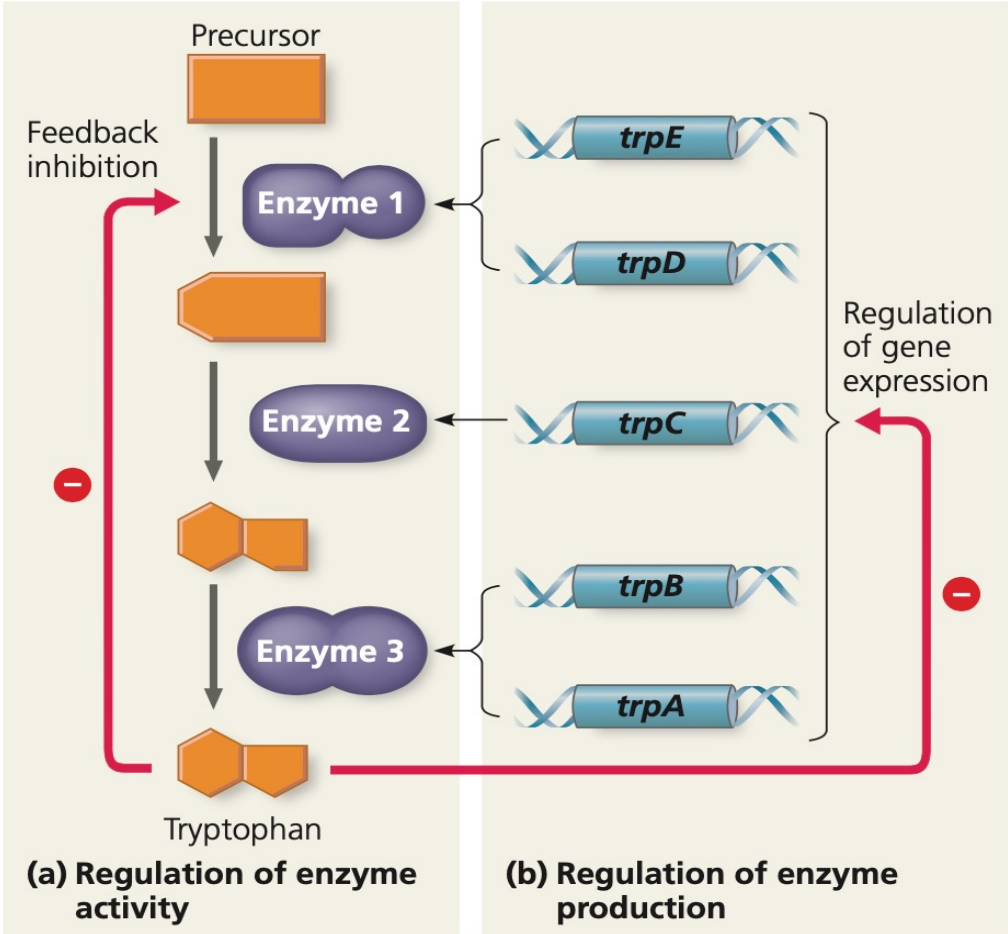
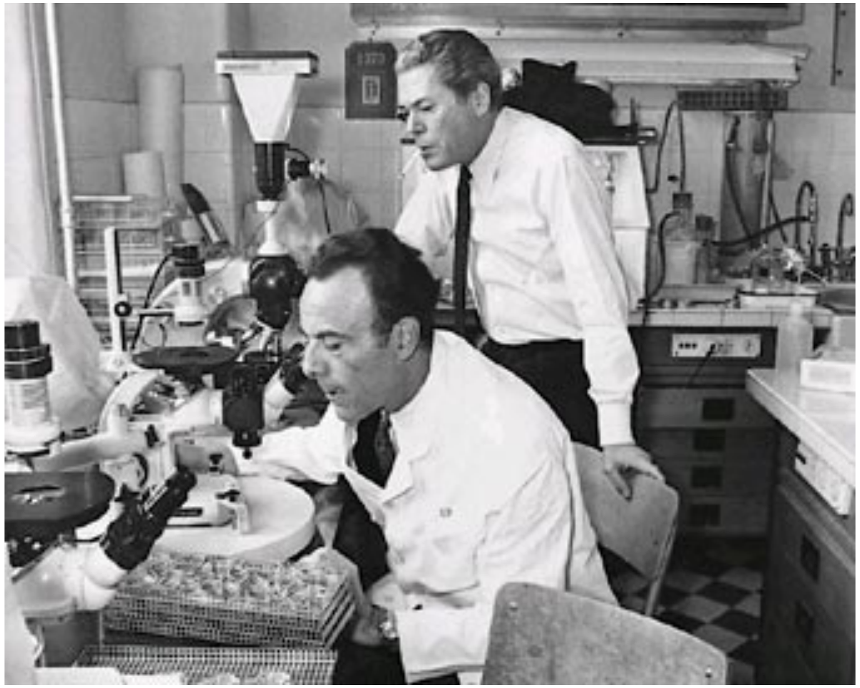


Figure 18.2

Operons were studied by Jacob & Monod

- Gene expression in bacteria is often influenced by **metabolic status** of cell.
- Numerous genes in bacterial genome can be activated or deactivated based on cellular **metabolic changes**.
- The **operon model** is a fundamental mechanism that explains this control of gene expression in bacteria.

Jacques Monod



Francois Jacob

1965 Nobel Prize in Physiology or Medicine

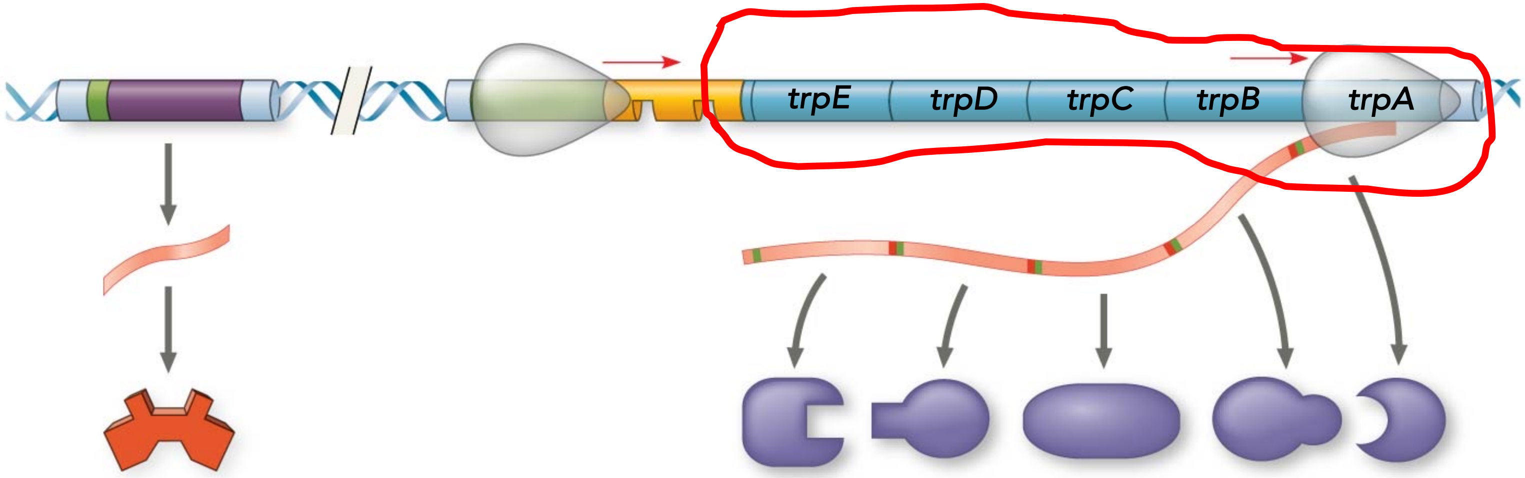
I. Repressible Operons

The *trp* operon is *repressible operon* because its transcription is usually ON but can be inhibited (repressed) when a specific small molecule (e.g. tryptophan) binds allosterically to a regulatory protein.

The trp operon is one example of how gene expression can respond to changes in the cell's internal and external environment.

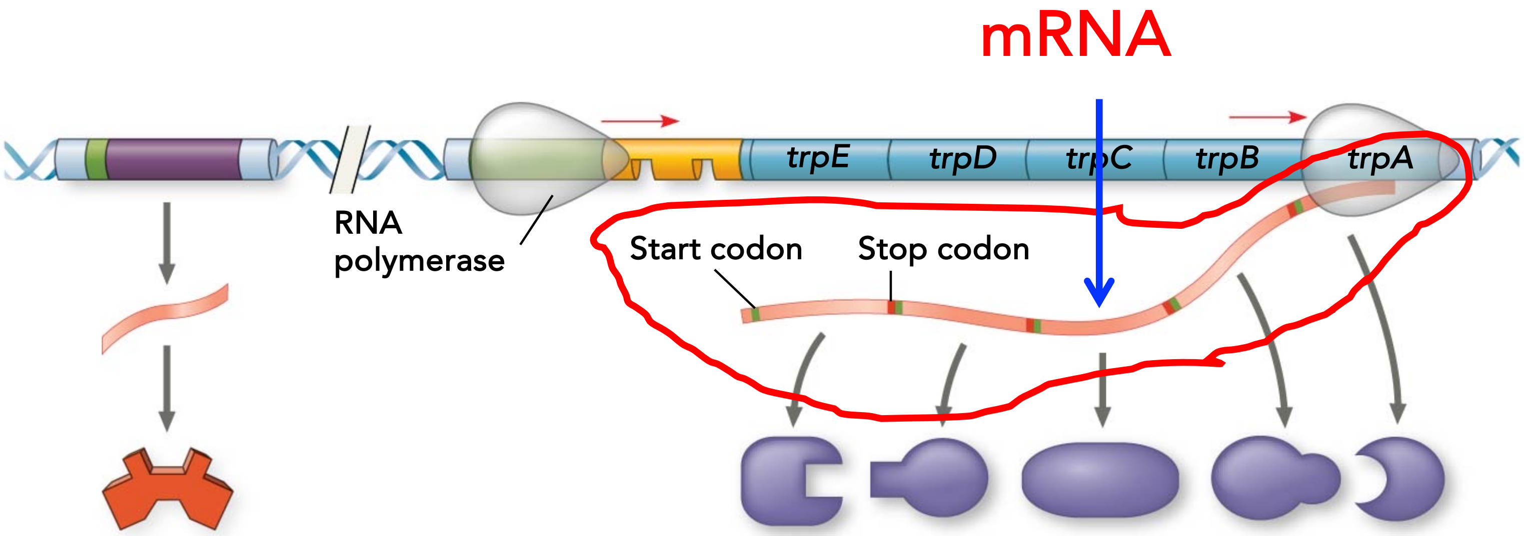
Genes of Operon: Coordinated switching on-off

Cluster of 5 genes whose protein products are tryptophan biosynthesis enzymes



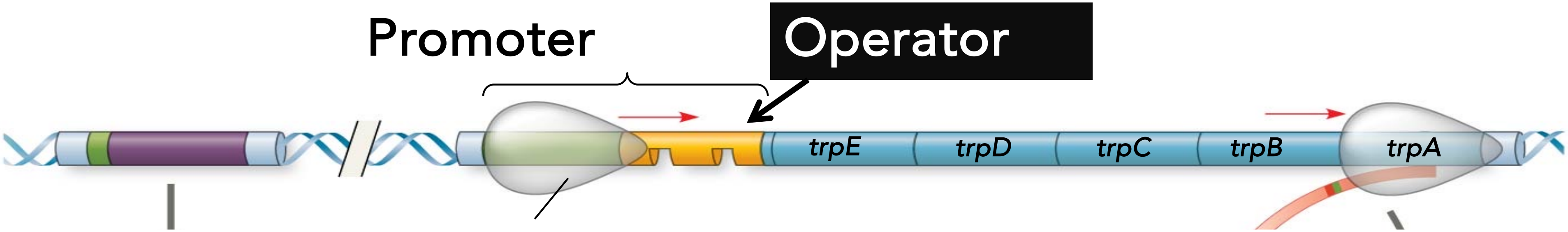
- *E. coli* synthesizes tryptophan from a precursor molecule in 3-step pathway
- Each reaction is catalyzed by a specific enzyme, and 5 genes that code for subunits of these enzymes are clustered together on the bacterial chromosome
- A single promoter serves all 5 genes, together it constitutes a **transcription unit**

Coupled expression of genes in bacteria: Transcription linked to translation



- Cell can translate one **mRNA into five separate polypeptides** because the mRNA is punctuated with start and stop codons that signal where the coding sequence for each polypeptide begins and ends.
- Switch on the expression of all 5 genes (trp A, B, C, D, E); when sufficient tryptophan is made turn off all five genes

Operator: the single on-off switch



Operator is a stretch of DNA sequence, located between promoter and the *trp* genes

Promoter + operator + genes = operon

Role of operator in regulating gene expression?

The *trp* operon in *E. coli*: regulated synthesis of repressible enzymes

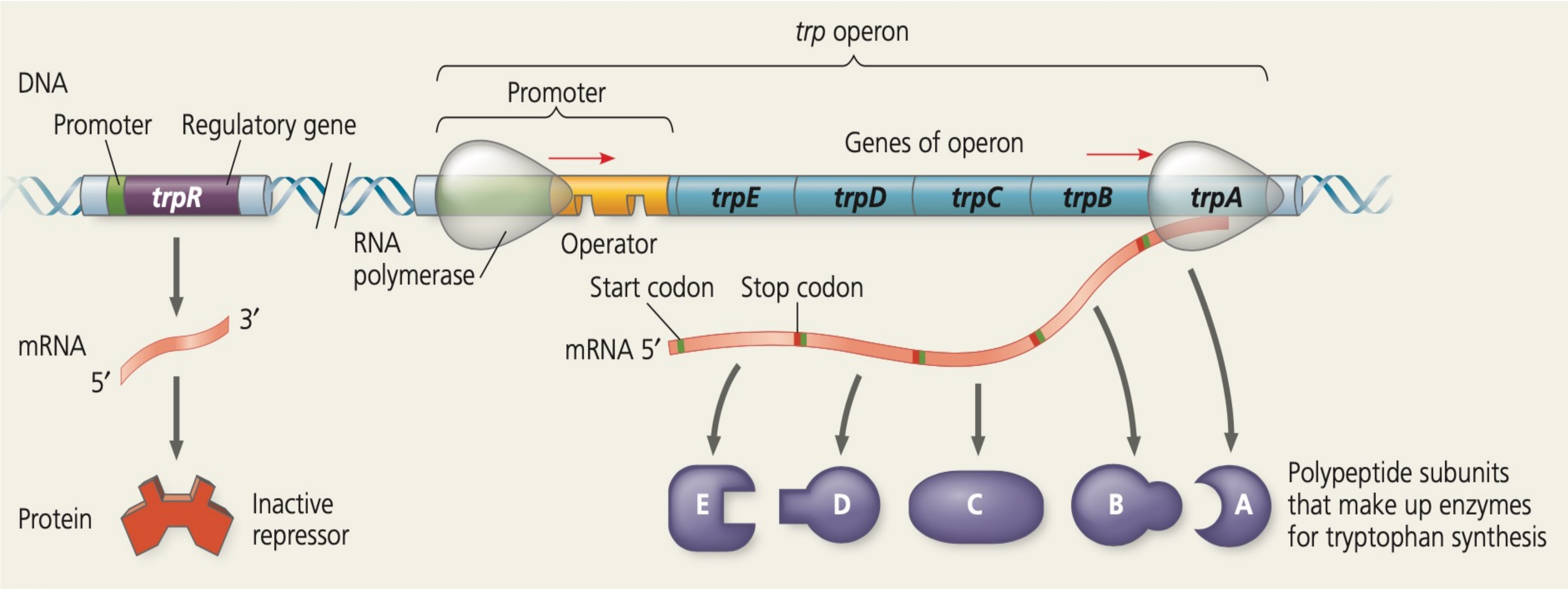


Figure 18.3

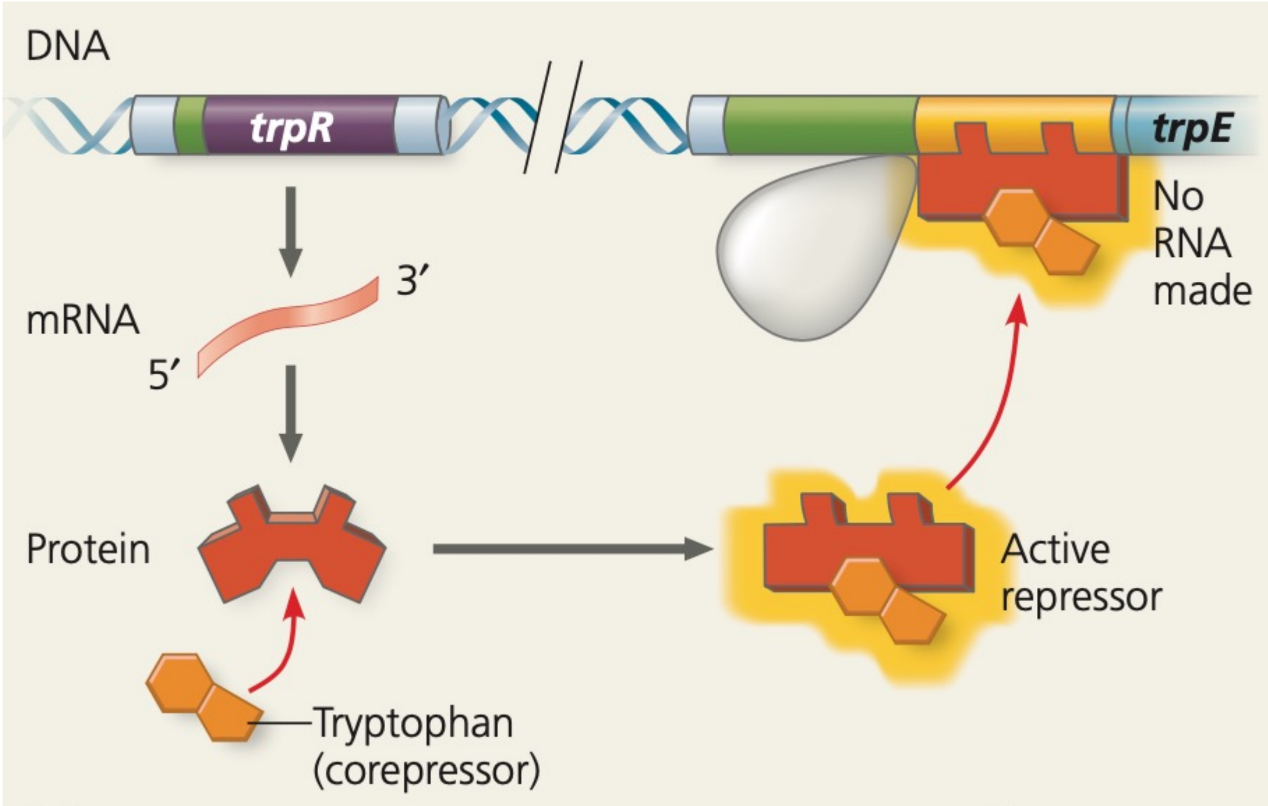
Tryptophan absent, repressor inactive, operon on. RNA polymerase attaches to the DNA at the operon's promoter and transcribes the operon's genes.

The *trp* operon in *E. coli*: regulated synthesis of repressible enzymes

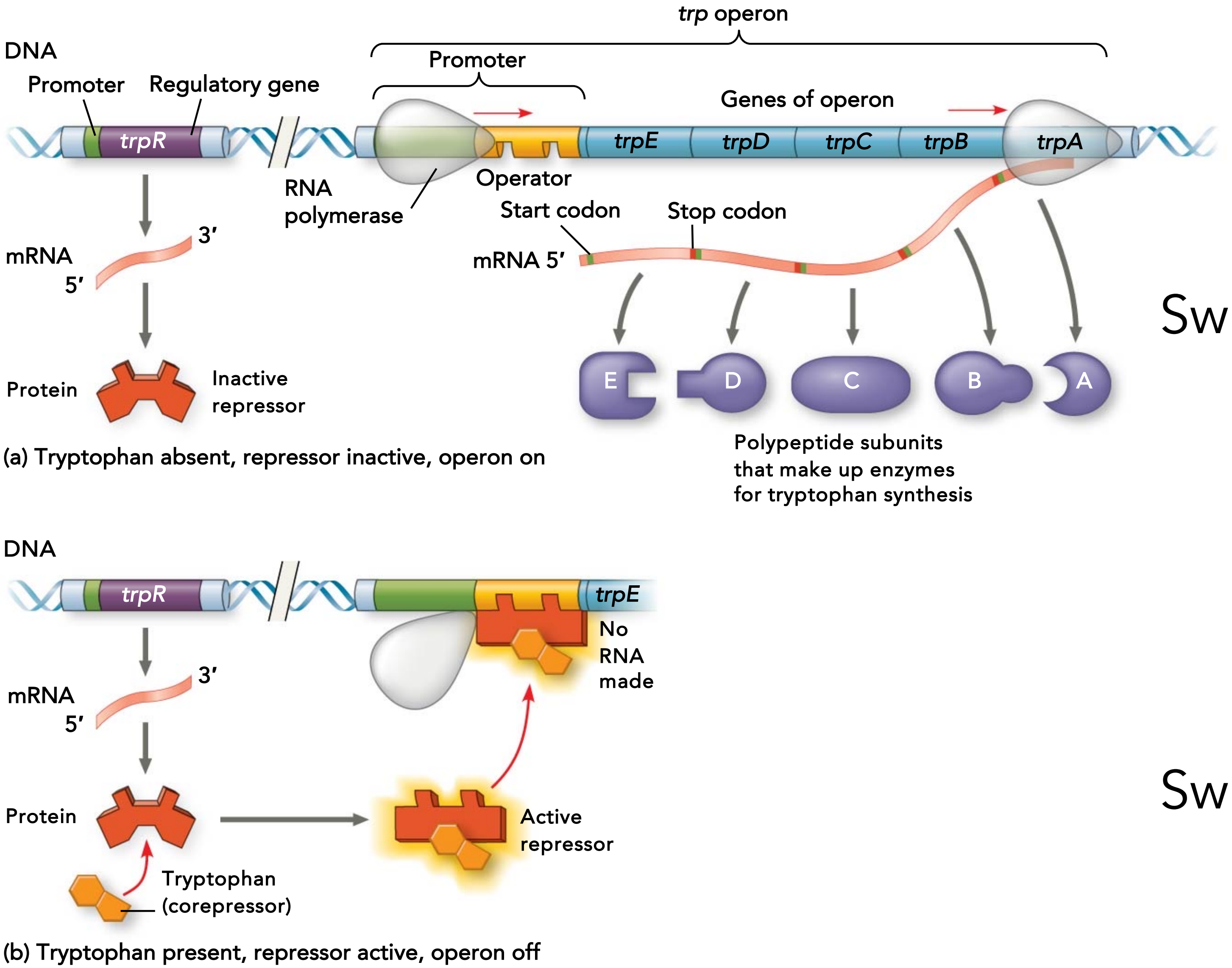
Tryptophan present, repressor active, operon off.

Accumulation of tryptophan, the end product of the pathway, represses transcription of the trp operon, thus blocking synthesis of all the enzymes in the pathway and shutting down tryptophan production

Operon can be switched off by a protein that is called the *trp* repressor.



An Overview of *trp* operon in *E. coli*: regulated synthesis of repressible enzymes



Switch is ON

Switch is OFF

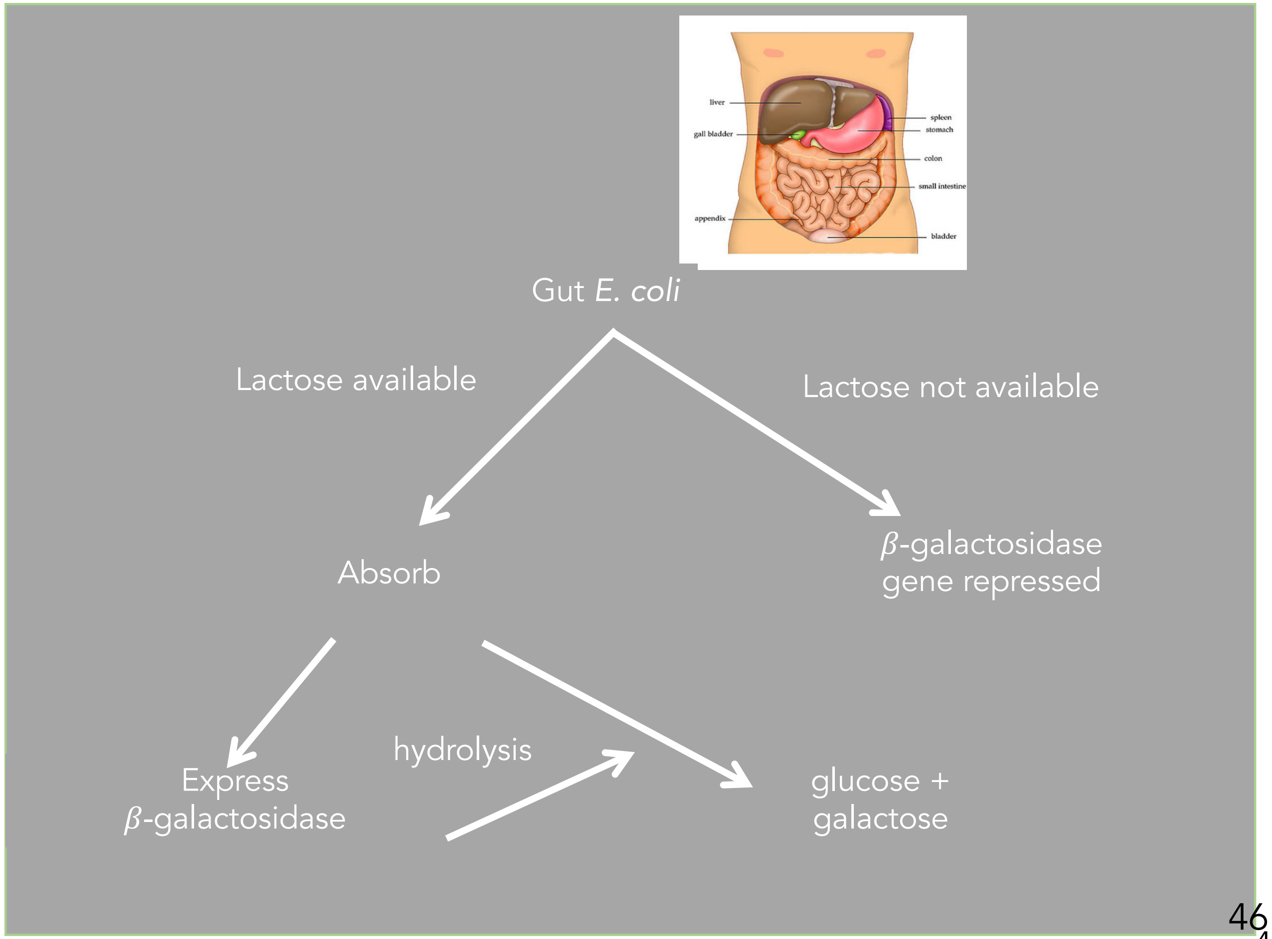
- I. COVID-19 Diagnostics
  - Genome sequencing
  - RT-PCR applications
  - Molecular diagnostics
- II. Regulation of Gene Expression
  - Transcription
  - Translation
- III. Gene Regulation Models
  - Trp operon
  - Lac operon

II. Inducible Operons

An inducible operon is usually off but can be stimulated (induced) when a specific small molecule interacts with a regulatory protein. Example of an inducible operon is the *lac* operon (*lac* for lactose).

The *lac* Operon

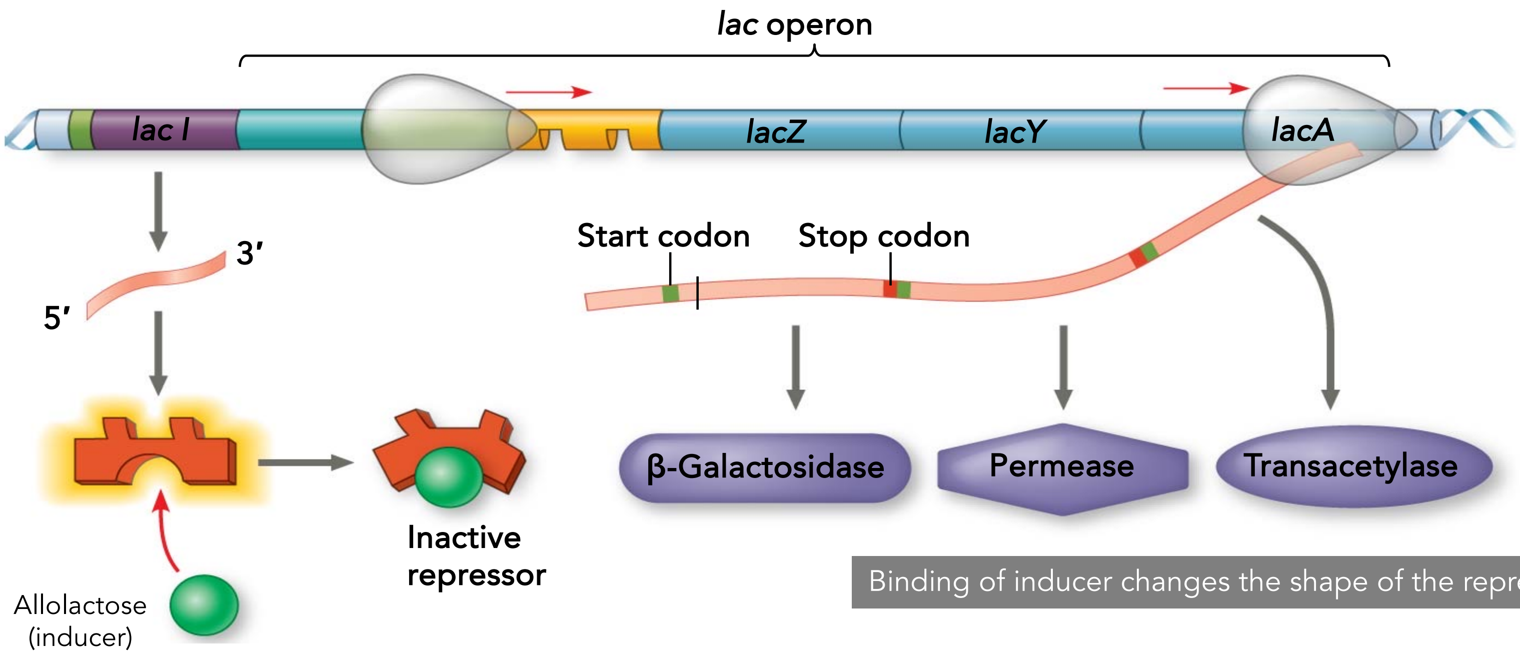
- *E. coli* inhabits human gut
- Disaccharide lactose is available from milk



46

Induction of the *lac* operon

T



Binding of inducer changes the shape of the repressor

In an inducible operon, binding of an **inducer** to an innately active repressor inactivates repressor and turns on transcription

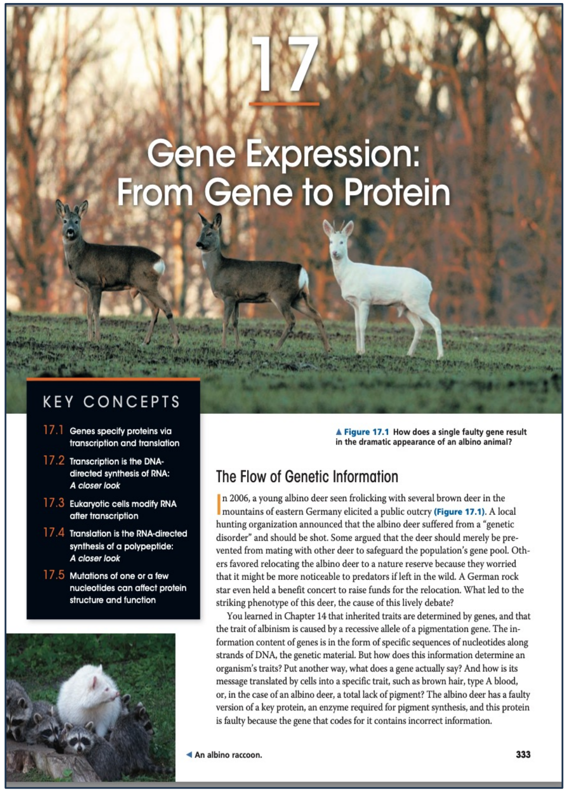
E. coli uses 3 enzymes to metabolize lactose, the genes for which are clustered in ***lac* operon**. *lacZ* (**β**-galactosidase to hydrolyzes lactose to glucose and galactose), *lacy* (permease, membrane protein that transports lactose into cell), *lacA* (transacetylase, unclear role). Unusually, the gene for the lac repressor, *lacI*, is adjacent to the lac operon

Overall Summary

- We discussed advanced techniques such as genome sequencing and RT-PCR within the context of COVID-19 diagnostics.
- Investigated the intricacies of gene expression, specifically emphasizing transcription and translation processes in bacterial systems.
- Examined models of gene regulation, with a particular focus on the trp and lac operons.

References

- Campbell Biology - Reece, Urry, Cain, Wasserman, Minorsky, Jackson 10th Edition, Cummings
- *Acknowledgment*
    - Cover images – getty images



17

Gene Expression: From Gene to Protein

KEY CONCEPTS

17.1 Genes specify proteins via transcription and translation

17.2 Transcription is the DNA-directed synthesis of RNA: *A closer look*

17.3 Eukaryotic cells modify RNA after transcription

17.4 Translation is the RNA-directed synthesis of a polypeptide: *A closer look*

17.5 Mutations of one or a few nucleotides can affect protein structure and function

▲ **Figure 17.1** **How does a single faulty gene result in the dramatic appearance of an albino animal?**

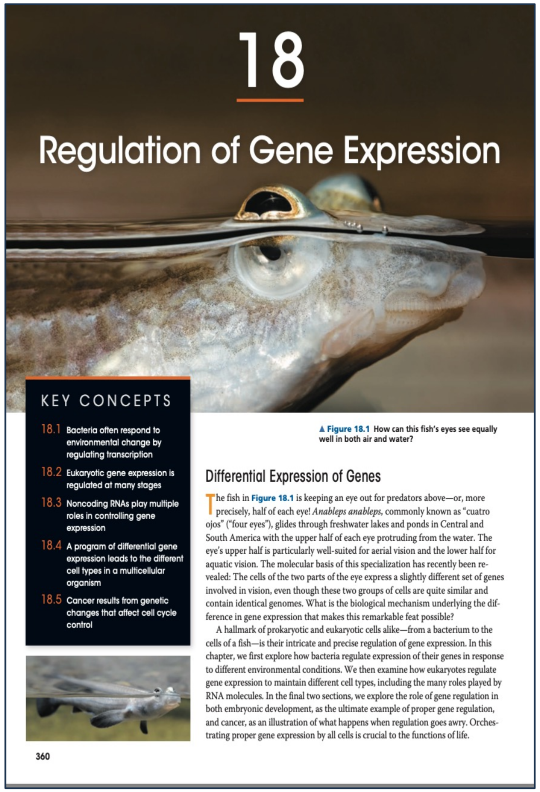
The Flow of Genetic Information

In 2006, a young albino deer seen frolicking with several brown deer in the mountains of eastern Germany elicited a public outcry (**Figure 17.1**). A local hunting organization announced that the albino deer suffered from a "genetic disorder" and should be shot. Some argued that the deer should merely be prevented from mating with other deer to safeguard the population's gene pool. Others favored relocating the albino deer to a nature reserve because they worried that it might be more noticeable to predators if left in the wild. A German rock star even held a benefit concert to raise funds for the relocation. What led to the striking phenotype of this deer, the cause of this lively debate?

You learned in Chapter 14 that inherited traits are determined by genes, and that the trait of albinism is caused by a recessive allele of a pigmentation gene. The information content of genes is in the form of specific sequences of nucleotides along strands of DNA, the genetic material. But how does this information determine an organism's traits? Put another way, what does a gene actually say? And how is its message translated by cells into a specific trait, such as brown hair, type A blood, or, in the case of an albino deer, a total lack of pigment? The albino deer has a faulty version of a key protein, an enzyme required for pigment synthesis, and this protein is faulty because the gene that codes for it contains incorrect information.

◀ **An albino raccoon.**

333



18

Regulation of Gene Expression

KEY CONCEPTS

18.1 Bacteria often respond to environmental change by regulating transcription

18.2 Eukaryotic gene expression is regulated at many stages

18.3 Noncoding RNAs play multiple roles in controlling gene expression

18.4 A program of differential gene expression leads to the different cell types in a multicellular organism

18.5 Cancer results from genetic changes that affect cell cycle control

▲ **Figure 18.1** **How can this fish's eyes see equally well in both air and water?**

Differential Expression of Genes

The fish in **Figure 18.1** is keeping an eye out for predators above—or, more precisely, half of each eye! *Anableps anableps*, commonly known as "cuatro ojos" ("four eyes"), glides through freshwater lakes and ponds in Central and South America with the upper half of each eye protruding from the water. The eye's upper half is particularly well-suited for aerial vision and the lower half for aquatic vision. The molecular basis of this specialization has recently been revealed: The cells of the two parts of the eye express a slightly different set of genes involved in vision, even though these two groups of cells are quite similar and contain identical genomes. What is the biological mechanism underlying the difference in gene expression that makes this remarkable feat possible?

A hallmark of prokaryotic and eukaryotic cells alike—from a bacterium to the cells of a fish—is their intricate and precise regulation of gene expression. In this chapter, we first explore how bacteria regulate expression of their genes in response to different environmental conditions. We then examine how eukaryotes regulate gene expression to maintain different cell types, including the many roles played by RNA molecules. In the final two sections, we explore the role of gene regulation in both embryonic development, as the ultimate example of proper gene regulation, and cancer, as an illustration of what happens when regulation goes awry. Orchestrating proper gene expression by all cells is crucial to the functions of life.

360

Next Lecture…

Cell Communication

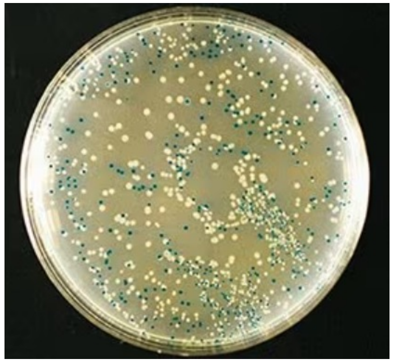
Experimental demonstrations

Blue White Screening

Two possibilities in a cloning experiment

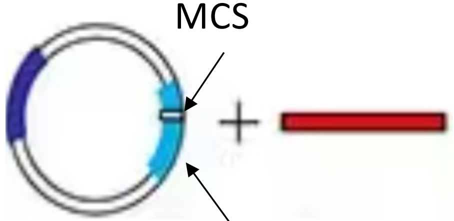
Gene of interest successfully integrated in plasmid (recombinant plasmid)

Plasmid gets self ligated



Q. How to screen for bacterial colonies which have the recombinant plasmid?

MCS

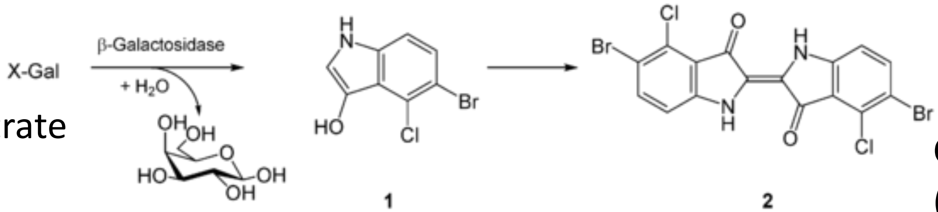


Lac Z gene

Recombinant plasmid causes insertional inactivation of lacZ gene → No β-galactosidase → White colonies

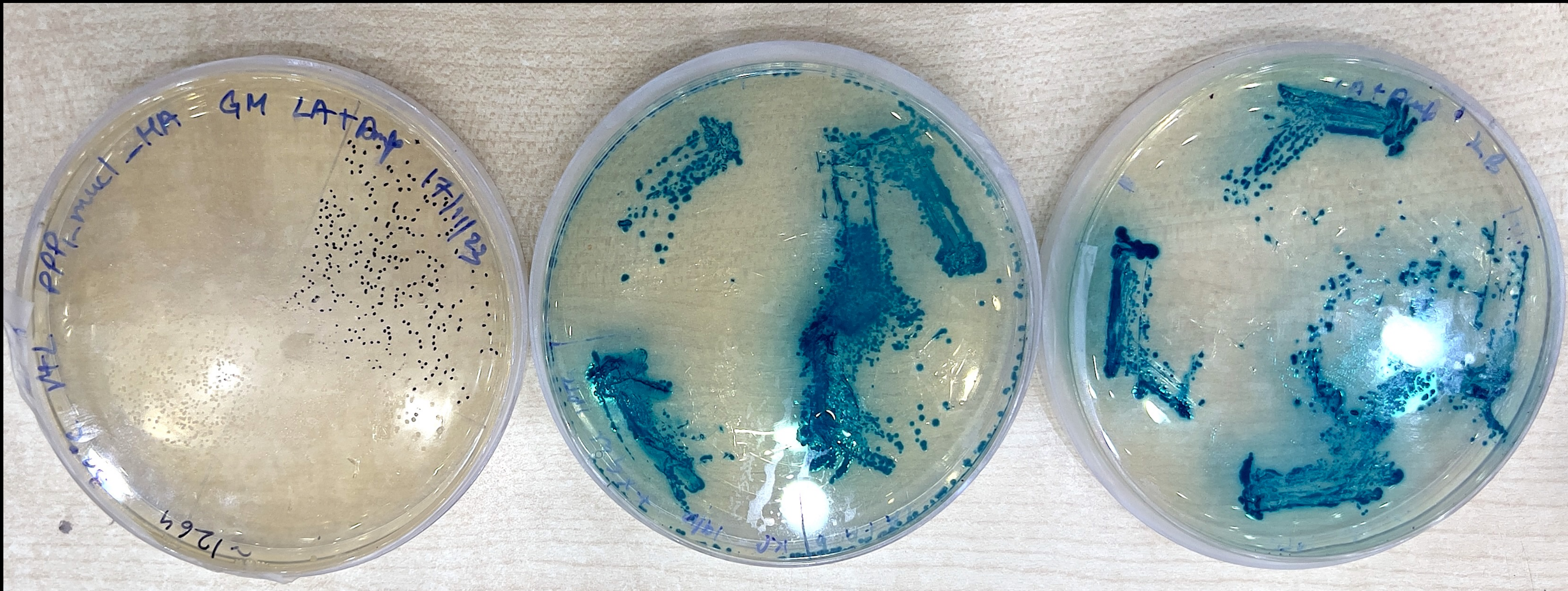
Self-ligated plasmid → β-galactosidase enzyme produced, cleaves X-gal → Blue colonies

Substrate



Colored product (blue)

Blue White Screening



Agarose gel electrophoresis