River Valley High School
2025 JC1 H2 Biology

Lecture Topic 10: Control of Gene Expression - Eukaryotes

Name: _____ ( ) Class: 25J _____ Date: _____

References

Title	Authors
Biology (8th Edition)	Campbell and Reece
Biology: An Australian Focus (3rd Edition)	Knox, Ladiges, Evans and Saint
Molecular Cell Biology (6th Edition)	Lodish, Berk, Kaiser, Krieger, Scott, Bretscher, Ploegh and Matsudaira
Molecular Biology of the Cell (5th Edition)	Alberts, Johnson, Lewis, Raff, Roberts and Walter
Principles of Genetics (3rd Edition)	Snustad and Simmons

Websites

URL	Description
http://highered.mcgraw-hill.com/sites/9834092339/student_view0/chapter10/chromatin_remodeling.html	transcriptional activation and regulation.
https://highered.mheducation.com/sites/9834092339/student_view0/chapter15/processing_of_gene_information__prokaryotes_vs__eukaryotes.html	basic differences between prokaryotic and eukaryotic gene expression.
https://highered.mheducation.com/sites/9834092339/student_view0/chapter15/rna_splicing.html	summary of the splicing process.

H2 Biology Syllabus 9477 (2025)

Candidates should be able to use the knowledge gained in the following section(s) in new situations or to solve related problems.

Related Topics	**Content**
Organisation of Genomes	The structure of eukaryotic chromatin
Control of Gene Expression - Prokaryotes	Operons

Learning Outcomes

2C. Control of Gene Expression

b. Explain how differential (i.e. spatial and temporal) gene expression in eukaryotes can be regulated at different levels:
   i. Chromatin level (histone modification and DNA methylation);
   ii. Transcriptional level (control elements, such as promoters, silencers and enhancers, and proteins, such as transcription factors and repressors);
   iii. Post-transcriptional level (processing of pre-mRNA in terms of splicing, polyadenylation and 5' capping);
   iv. Translational level (half-life of RNA, 5' capping, initiation of translation); and
   v. Post-translational level (biochemical modification and protein degradation).

c. Describe the principles and procedures of these molecular techniques:
   i. Polymerase chain reaction (including its advantages and limitations);
   ii. Gel electrophoresis; and
   iii. Southern blotting and nucleic acid hybridisation.

Lecture Outline

I. Control of Gene Expression

A. Chromatin-level Control
- **A1.** Availability of genes for transcription
  - **A1.1.** Histone modification
  - **A1.2.** DNA methylation

B. Transcriptional Control
- **B1.** Initiation of transcription
  - **B1.1.** Control elements
  - **B1.2.** Transcription factors

C. Post-transcriptional Control via Processing of pre-mRNA
- **C1.** Alteration of mRNA ends
- **C2.** RNA splicing

D. Translational Control
- **D1.** mRNA stability
- **D2.** Initiation of translation

E. Post-translational Control
- **E1.** Biochemical modification
- **E2.** Protein degradation

F. Gene Amplification

II. Molecular Techniques
- **A.** Polymerase Chain Reaction
- **B.** Gel electrophoresis
- **C.** Southern blotting and nucleic acid hybridisation

I. CONTROL OF EUKARYOTIC GENE EXPRESSION

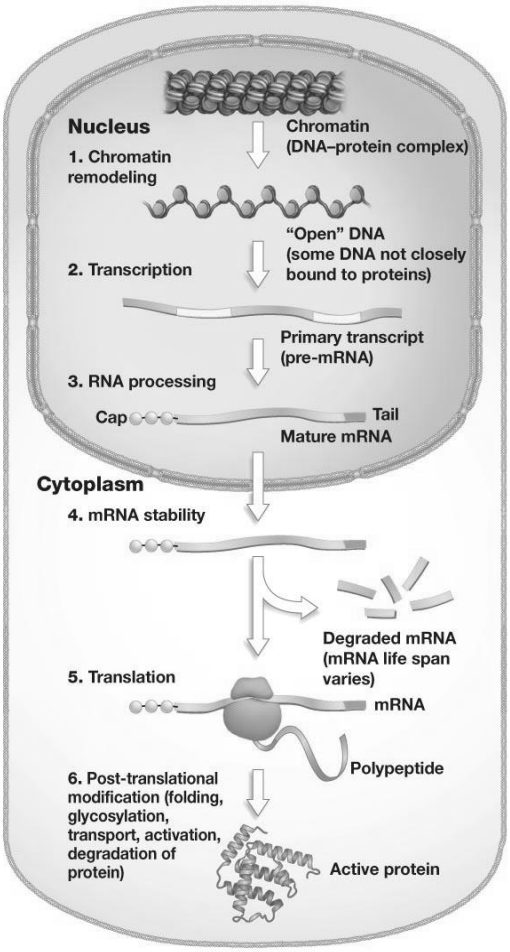
Gene expression, the generation of a protein or RNA product from a particular gene, is regulated by complex mechanisms in eukaryotes. Normally, only a fraction of the genes in a eukaryotic cell are expressed at any one place (**tissue specificity**) and at any one time (**temporal specificity**).

Gene expression and control in eukaryotes differ from those in prokaryotes for the following reasons:

1. Eukaryotic DNA is organised into nucleosomes.
   The genes must be in an active structure to be expressed.

2. Eukaryotic genes are not organised into operons.
   The genes encoding proteins that function together are usually located on different chromosomes. Each gene thus needs its own **regulatory sequences** (e.g. promoters, silencers and enhancers).

3. The processes of transcription and translation are separated in eukaryotes by the nuclear envelope. Eukaryotic pre-mRNA (which is heterogeneous nuclear RNA) must be processed and translocated out of the nucleus before it is translated.



Eukaryotic gene expression can be regulated at variuos levels:

A. Chromatin-level control
   - Availability of genes for transcription

B. Transcriptional control
   - Initiation of transcription

C. Post-transcriptional control
   - Processing of pre-mRNA

D. Translational control
   - mRNA stability
   - Initiation of translation

E. Post-translational control
   - Biochemical modification
   - Protein degradation

F. Gene amplification

A. Chromatin-level Control

The organisation of chromatin discussed earlier in Topic 9 serves a dual purpose. One function is to pack the DNA into a compact form that fits inside the nucleus of a cell. The other main function is regulatory: the physical state of DNA in or near a gene is important in helping control whether the gene is available for transcription.

Chemical modifications of the histone proteins and DNA of chromatin play a key role in chromatin structure and gene expression.

A1. Availability of Genes for Transcription

Once the haploid sperm and egg combine to form a diploid zygote, the number of genes in cells remains approximately the same. As cells differentiate, different genes are available for transcription.

A1.1. Histone modification

1. A typical nucleus contains chromatin that is diffused (euchromatin) and chromatin that is condensed (heterochromatin).

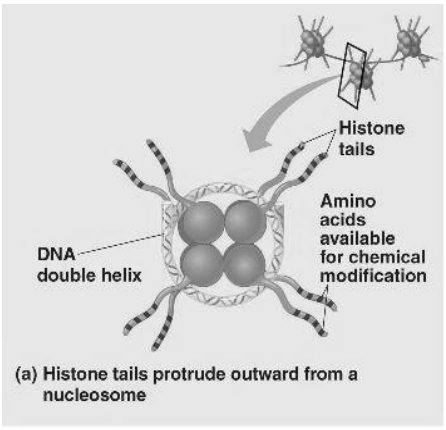
    - The genes found in euchromatin are active (available for transcription) whereas those found in heterochromatin are inactive.
    - Long-term changes in **gene activity** occur during development as chromatin goes from a diffused to a condensed state or vice versa.

2. The nuclear genome is packed together with histones into nucleosomes.

    - Transcription initiation is prevented if the **promoter** region is part of a nucleosome. Thus, activation of a gene for transcription requires changes in the state of the chromatin, called **chromatin remodelling**.
    - The N-terminus of a histone molecule in a nucleosome protrudes outward from the nucleosome. These **histone tails** are accessible to various modifying enzymes, which catalyse the addition or removal of specific chemical groups

3. **Histone acetylation**

    - In acetylation, acetyl groups are attached to lysines in histone tails. The enzyme catalysing the transfer is **histone acetyltransferase** (HAT).
    - When lysines are acetylated, their positive charges are neutralised. As such, histone tails no longer interact with the neighbouring nucleosomes.
    - Electrostatic interaction between neighbouring nucleosomes promotes folding of chromatin into more compact structure. When this interaction does not occur in an acetylated region, chromatin has a more diffused structure. **Transcription factors** have easier access to genes.
    - Histone acetylation does not only promote transcription initiation by remodelling the chromatin structure, the enzyme HAT also binds to and aids in the recruitment of the **transcription machinery**.

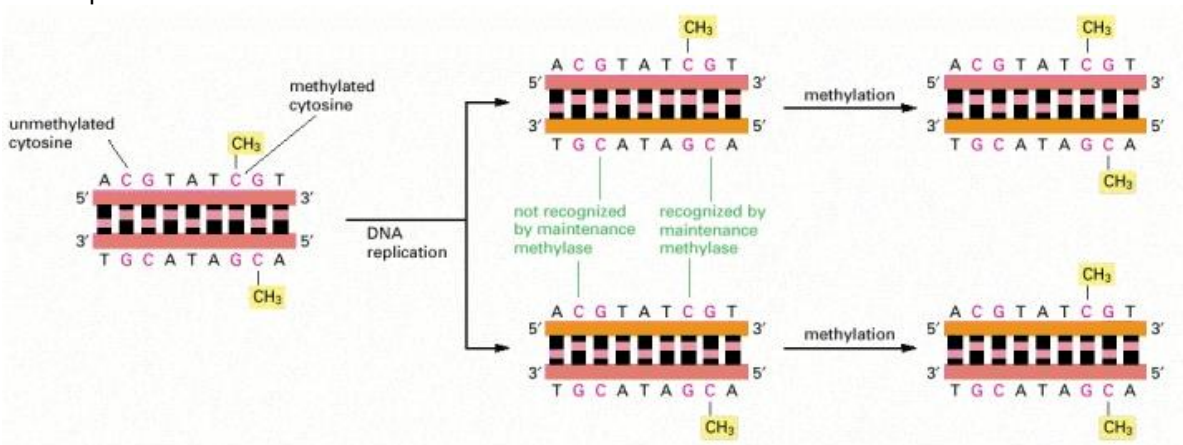


4. **Methylation**

    - Histone methylation is the modification of specific amino acids in a histone protein through the addition of methyl groups, catalysed by histone methyltransferase.
    - In most cases, histone methylation is associated with transcriptional repression, as the addition of methyl groups to histone tails promotes chromatin condensation.

A1.2. DNA Methylation

1. Cytosine residues in fungal, plant or animal DNA can be **methylated** to produce 5-methylcytosine. These methylated cytosine residues are located in GC-rich sequences (called **GC-islands**), which are often near or in the promoter region of a gene. This process is catalysed by enzymes known as DNA methyltransferases.
2. Comparison of the same genes in different tissues shows that genes are usually more heavily methylated in the cells in which they are not expressed. Removal of the extra methyl groups can turn on some of these genes.
   DNA methylation can either (1) inhibit the binding of general transcription factors, preventing the formation of the transcription initiation complex at the promoter, or (2) recruit proteins that bind to methylated DNA, which then attract histone deacetylases to condense chromatin and repress transcription.



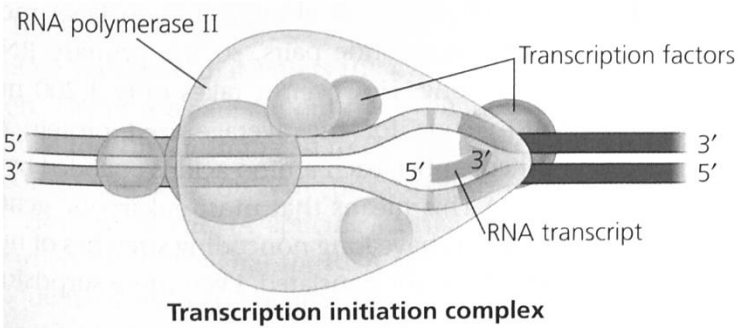
3. Application of DNA methylation in eukaryotes.

   - In some species, DNA methylation seems to be responsible for the long-term inactivation of the genes that occurs during normal cell differentiation in the embryo.
   - Once methylated, genes usually stay that way through successive cell divisions in a given individual.
   - At DNA sites where one strand is already methylated, methylation enzymes correctly methylate the daughter DNA strand after each round of DNA replication.
   - **Methylation patterns** are thus passed on, and cells forming specialised tissues keep a chemical record of what happened during embryonic development.

B. Transcriptional Control

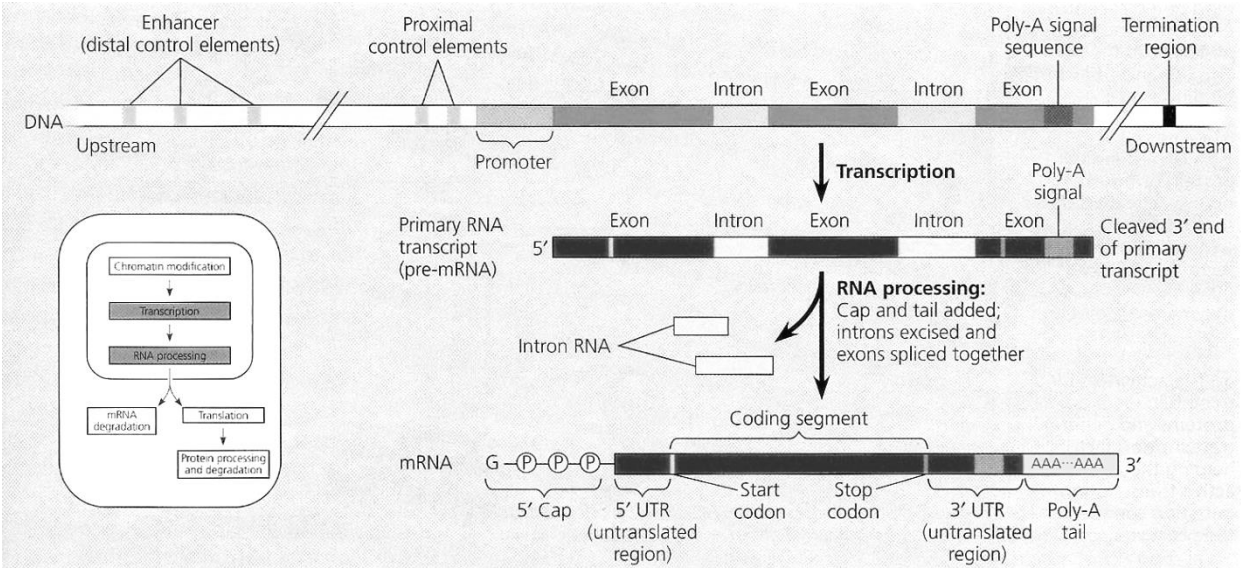
In all organisms, a common control point for gene expression is transcription. Regulation at this stage is often in response to signals coming from outside the cell, such as hormones or other signalling molecules.
(KIV: Communication & Equilibrium in Organisms)

B1. Initiation of Transcription



Transcription initiation complex

A cluster of proteins called a **transcriptional initiation complex** assembles on the promoter sequence found upstream of the gene. One of these proteins, **RNA Polymerase II**, then proceeds to transcribe the gene, synthesising a primary RNA transcript, known as **pre-mRNA**. Following this, **RNA processing** is carried out to form a **mature mRNA**.

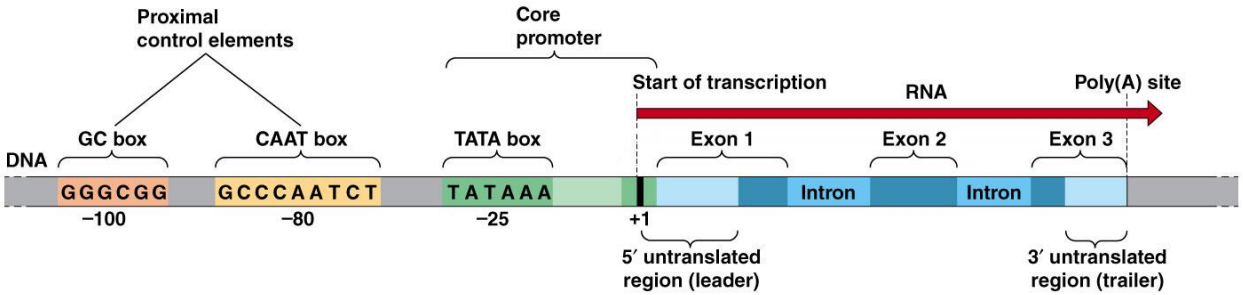


B1.1 Control Elements

1. Associated with most eukaryotic genes are **control elements**, which are segments of non-coding DNA that help regulate transcription by binding **transcription factors** [1]. These control elements and the transcription factors they bind are critical to the precise regulation of gene expression seen in many cell types.

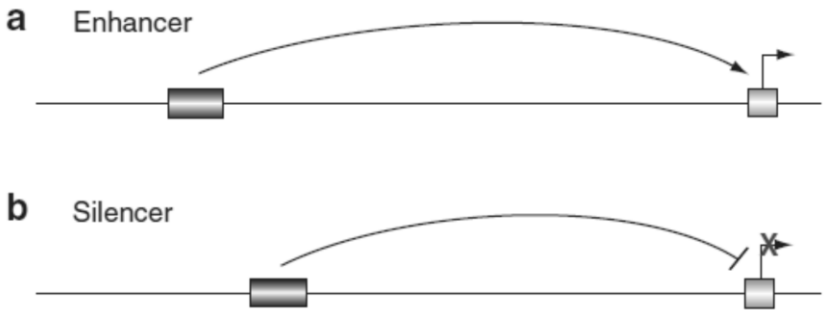
2. There are two distinct types of control elements:
   - **Promoter**, which comprises the **core promoter** and **proximal control elements**
   - **Distal control elements**, which comprises **enhancers** and **silencers**

[1] Transcription factor: A regulatory **protein** that binds to DNA and affects transcription of specific genes.



3. In both prokaryotes and eukaryotes, core promoter contains the transcription start site and a **TATA box**.
   - TATA box is an adenine- and thymine-rich consensus sequence found -30 bp upstream of the transcription start site.
   - In eukaryotic genes, TATA box serves as the binding site of **general transcription factors** that facilitate the binding of RNA polymerase II to the promoter, for transcription to occur.
   - Core promoter determines basal level of expression.

4. Eukaryotic genes also contain proximal control elements located immediately upstream from the core promoter (100 and 200 bp from transcription start site), which are sites that bind other gene regulatory proteins.
   - Two proximal control elements found in many eukaryotic genes are the **CAAT boxes** and **GC-rich sequences**. These two sequences determine how frequent the transcription event occurs.



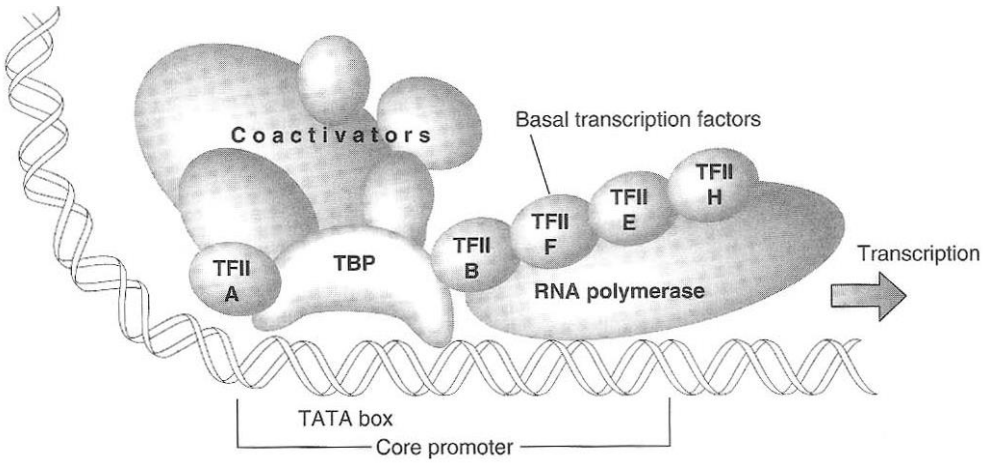
5. Distal control elements includes enhancers and silencers, which increase or decrease the rate of transcription initiation.
   - Enhancers / silencers can exert their effects when located hundreds or even thousands of bases away from the transcription units located on the chromosomes.
   - Enhancers / silencers have been found in a variety of locations both upstream and downstream of the transcription start site and even within the transcribed portions of some genes.
     Surprisingly, enhancers / silencers can function in an **orientation-independent** fashion.

6. Enhancer has a binding site for **specific transcription factors** called **activators**.
   - Attachment of activator to enhancer increases the rate of transcription of the gene.
   - A gene may have multiple enhancers, each active at a different time or in a different cell type or location in an organism.

7. Silencer has a binding sites for specific transcriptional factors called **repressors**
   - Attachment of repressor to silencer inhibits or decreases the rate of transcription of the gene.

B1.2 Transcription Factors

Eukaryotic RNA polymerase requires the assistance of proteins known as **transcription factors** to initiate transcription.

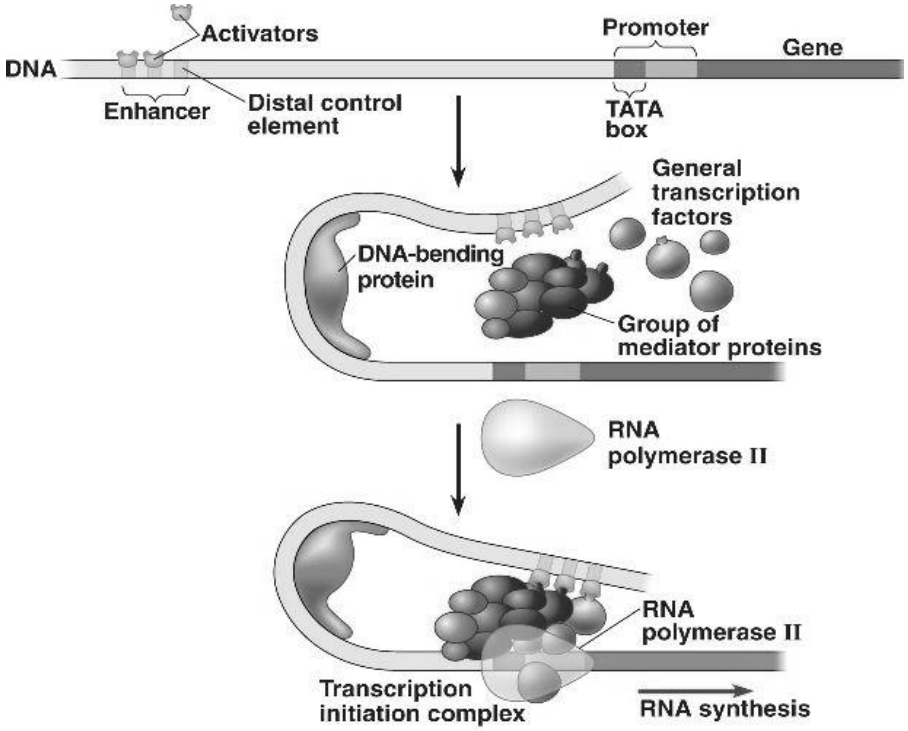


General transcription factors
These are essential for the transcription of all protein-coding genes.

- Only a few general transcription factors independently bind a DNA sequence, such as the TATA box within the promoter. The others primarily bind proteins, including each other, and RNA polymerase II.
- Protein-protein interactions are crucial to the initiation of eukaryotic transcription. Only when the complete **initiation complex** has assembled can the RNA polymerase begin to move along the DNA template strand, producing a complementary strand of RNA.
- The interaction of general transcription factors and RNA polymerase II usually leads to a low rate of initiation and production of few **RNA transcripts**.

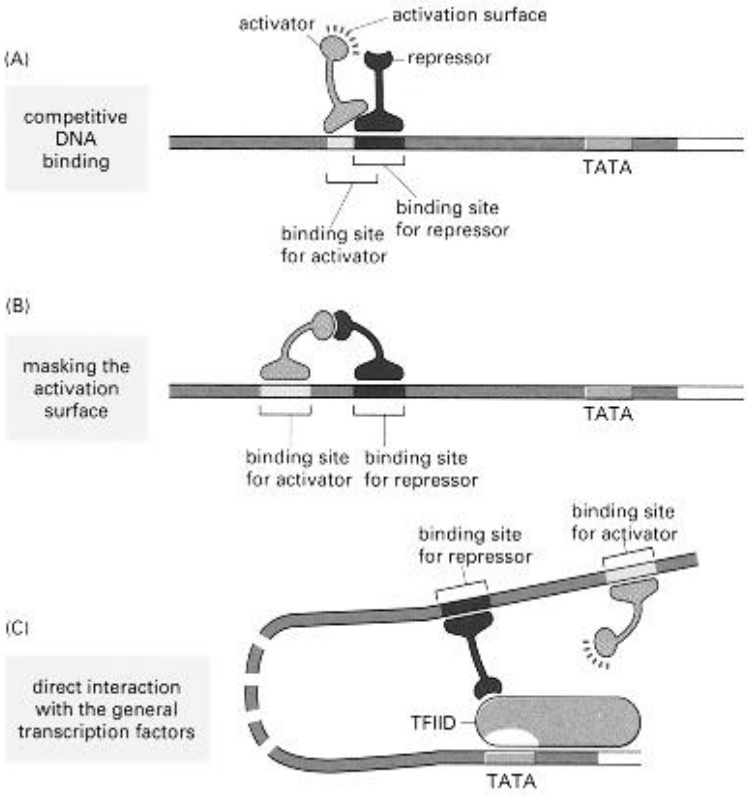
Specific transcription factors

In eukaryotes, high levels of transcription of particular genes at the appropriate time and place depend on the interaction of **control elements** with another set of proteins known as the **specific transcription factors**.



- Attachment of activator to enhancer increases the rate of transcription of the gene by
  - bending the DNA
  - helping RNA polymerase and transcription factors assemble at the promoter
  - facilitating the correct positioning of the transcription initiation complex on the promoter for transcription to occur.

- **Protein-mediated bending** of the DNA is thought to bring the bound activators in contact with a group of **mediator proteins** / **coactivators**, which in turn interact with proteins at the promoter. These multiple protein-protein interactions help assemble and position the initiation complex on the promoter.
  - Two common structures are present in activator proteins.
  - A DNA binding domain, a part of the protein's 3-dimensional structure that binds DNA
  - One or more **activation domains**, which binds other regulatory proteins or components of the transcription machinery, facilitating the protein-protein interactions that result in gene transcription.

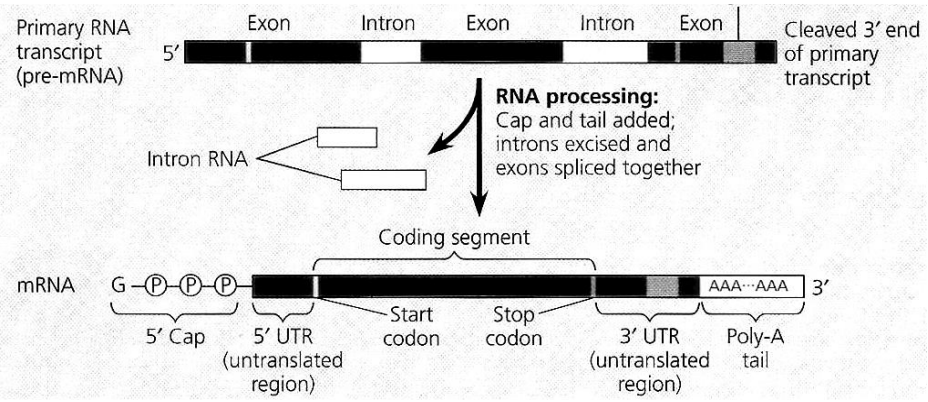


- Attachment of repressor to silencer inhibits or represses gene expression by either inhibiting activation or directly interferes with formation of the initiation complex.
  - Some repressors bind directly to distal control element (e.g. silencer), blocking the binding of activators, or turn off transcription even when activators are bound.

C. Post -Transcriptional Control via Processing of pre-mRNA

Transcription alone does not constitute gene expression. The expression of a protein-coding gene is ultimately measured by the amount of functional protein made by a cell. Researchers have discovered regulatory mechanisms that operate at various stages after transcription. These mechanisms allow a cell to fine-tune gene expression rapidly to environmental changes without altering its transcription patterns

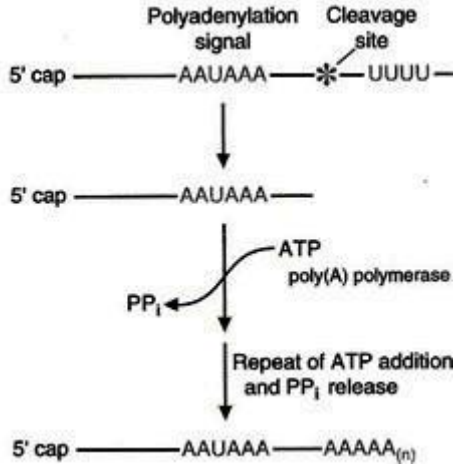
Processing of pre-mRNA in the nucleus and the export of mature mRNA to the cytoplasm provide several opportunities for regulating gene expression that are not available in prokaryotes. After transcription, **pre-mRNA** are modified by enzymes in various ways before the **mature mRNA** are exported to the cytoplasm via the nuclear pores.



During **post-transcriptional modification**, both ends of the pre-mRNA are usually altered. Certain interior sections (**introns**) of the mRNA are then cut out and the remaining pieces (**exons**) spliced together.

C1. Modification of mRNA Ends

Each end of a pre-mRNA molecule is modified in a particular way.



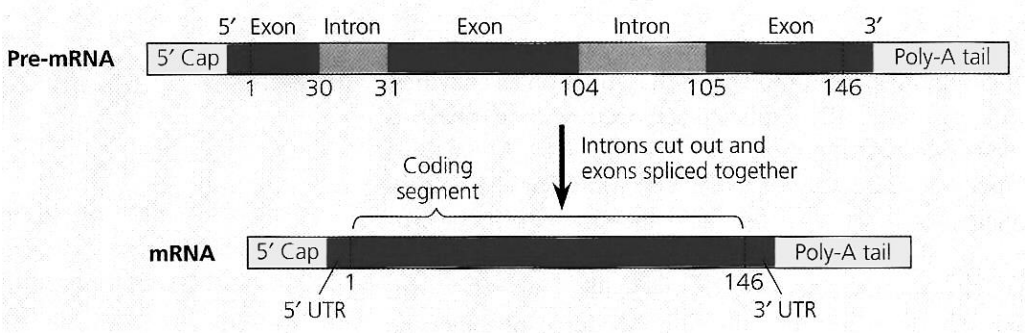
- **5' capping**
  A modified guanosine triphosphate residue is <u>added</u> to the 5' nucleotide of the pre-mRNA to create a 'capped' RNA.

- **3' polyadenylation**
  The 3' end of the pre-mRNA is <u>cleaved</u> and 50 to 250 adenine nucleotides are <u>added</u> to the new 3' end to create the poly(A)tail.

- The 5' cap and poly(A) tail share several important functions:
  - Facilitate the export of mature mRNA from the nucleus
  - Protect the mRNA from degradation by hydrolytic enzyme
  - Help ribosomes attach to the 5' end of the mRNA once the mRNA reaches the cytoplasm.

C2. RNA Splicing

The average length of a transcription unit along a eukaryotic DNA molecule is about 8000 nucleotides, however it takes only about 1200 nucleotides to code for an average sized protein of 400 amino acids. Most eukaryotic genes and their pre-mRNA have long, **non-coding** stretches of nucleotides that are not translated.
Even more surprising is the fact that most of these non-coding sequences are **interspersed** between **coding segments** of the gene and thus, those of the pre-mRNA. In other words, the sequence of DNA nucleotides that code for a eukaryotic polypeptide is <u>not continuous</u>.

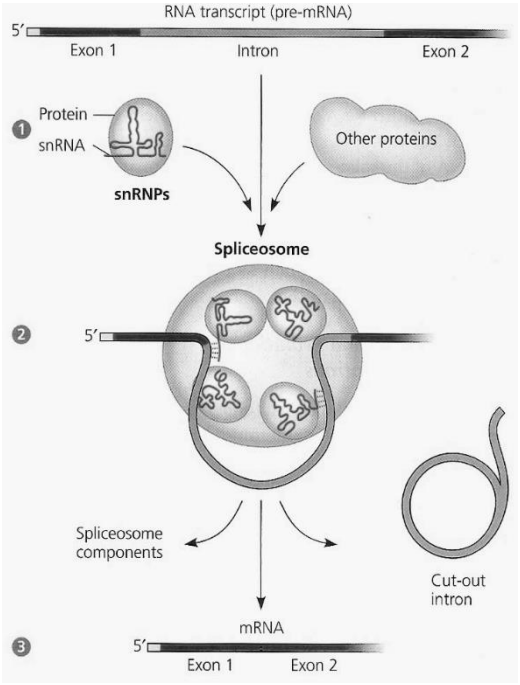


- The non-coding segments of nucleic acid that lie between coding segments are called **intervening sequences**, or **introns** for short.
- The coding segments are called **exons**, because they are eventually expressed, usually by being translated into amino acid sequences.

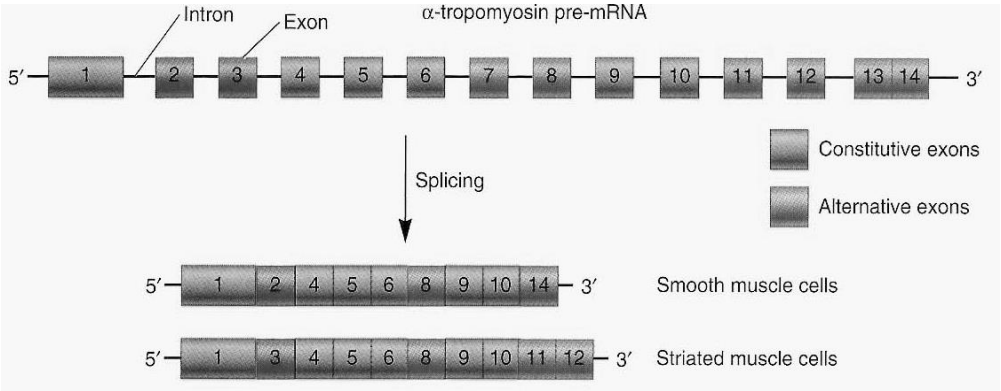
- Therefore, in making a primary transcript from a gene, RNA polymerase transcribes both introns and exons from the DNA. However, the mRNA molecule that enters the cytoplasm is an abridged version. The introns are cut out from the molecule and the exons joined together to form an mRNA molecule with a continuous coding sequence. Such is the process of **mRNA splicing**.

Mechanism of RNA Splicing

1. The signals for RNA splicing are short nucleotide sequences at the ends of introns known as **splice sites**. Particles called **small nuclear ribonucleoproteins**, or **snRNPs** recognise these splice sites.
2. As the name implies, snRNPs are located in the cell nucleus and are composed of snRNA and protein molecules.
3. Several different snRNPs join with additional proteins to form an even larger assembly called a **spliceosome**, which is almost as big as a ribosome.
4. The spliceosome interacts with the splice sites at the ends of an intron. It cuts at specific points to release the intron, then immediately joins together the exons that flank the intron.
5. There is strong evidence that snRNA plays a role in the catalytic processes as well as in spliceosome assembly and splice site recognition. The idea of a catalytic role for snRNA arose from the discovery of ribozymes, RNA molecules that function as enzymes.



Significance of Splicing

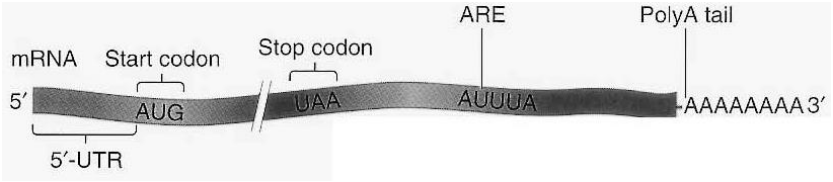


Energy is invested to synthesise intron sequences during transcription and to remove the introns via the spliceosome complex. Why is energy invested in such a 'wasteful process'?

The biological advantage lies in **alternative splicing**, which refers to the phenomenon that a pre-mRNA can be spliced in more than one way. Hence, alternative splicing produces two or more polypeptides with differences in their amino acid sequences, leading to possible variation in their functions.

D. Translational Control

D1. mRNA Stability



1. The **half-life** of mRNA (an indicator of mRNA stability) plays a central role in controlling the level of gene expression.

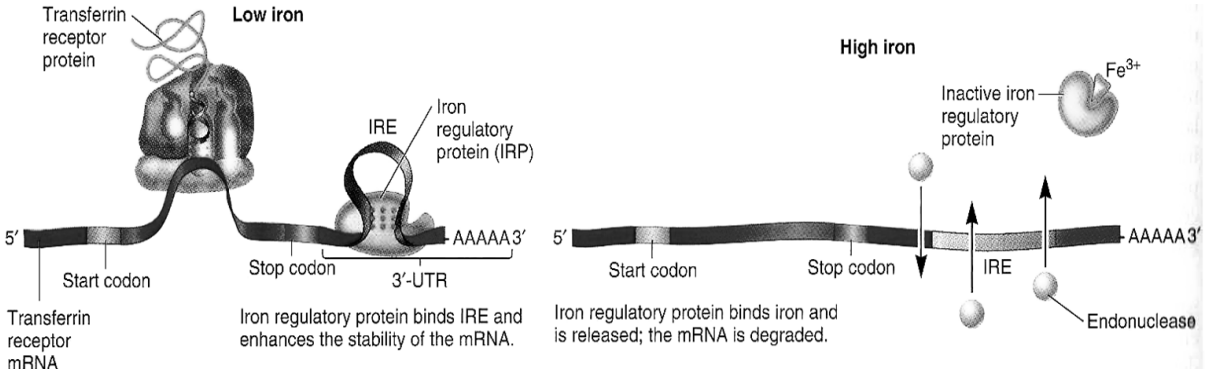
   - Bacterial mRNA molecules are degraded by enzymes within a few minutes of their synthesis. Hence, bacteria can change their patterns of protein synthesis so quickly in response to environmental changes.
   - mRNA molecules of multicelluar eukaryotes, in contrast, typically survive for hours, days or even weeks.
     E.g. mRNAs for haemoglobin polypeptides (α-globin and β-globin) in red blood cells are unusually stable, so that these long-lived mRNA are translated repeatedly.

2. Various factors play a role in determination of mRNA stability.

   Length of poly(A)tail
   - Most newly made mRNAs contain a poly(A)tail that averages 200 nucleotides in length. The poly(A)tail is recognised by **poly(A)-binding protein**. The binding of this protein enhances RNA stability.
   - As an mRNA ages, its poly(A)tail tends to be shortened by the action of **cellular exonucleases**.
   - Once the poly(A)tail becomes less than 30 adenosine residues in length, the poly(A)-binding protein can no longer bind and the mRNA is quickly degraded by exo- and endo-nucleases.
   - **Enzymatic shortening** of the poly(A)tail also helps trigger the action of enzymes that remove the 5' cap. Once the 5' cap is removed, the nucleases quickly degrade the mRNA.

   Destabilising sequences
   - Certain mRNAs, particularly those with short half-lives, contain sequences that act as **destabilising elements**.
   - Whilst these destabilising elements can be located anywhere within the mRNA, they are most commonly found at the 3' UTR.
     E.g. **AU-rich element** (ARE) are found in many short-lived mRNAs.
   - The consensus sequence AUUUA is recognised by cellular proteins that bind to the ARE, hence influencing whether or not the mRNA is rapidly degraded.

Binding of regulatory protein to 3'UTR



- Binding of certain regulatory protein to 3'UTR may also stabilise the mRNA by preventing degradation action of endonucleases
  E.g. Iron regulatory protein (IRP) binds at iron response element (IRE).

Assimilation of Iron in Mammalian Cells

1. Ingested iron (Fe^{3+}) is absorbed into the bloodstream and becomes bound to transferrin.

⇨ The transferrin-Fe^{3+} complex is recognised by transferrin receptor on the cell surface ⇨ the complex binds to the receptor and is then transported to the cytosol by endocytosis.

⇨ Fe^{3+} is released from transferrin and is used.

2. If there is an overabundance, the excess Fe^{3+} is stored in ferritin to prevent the toxic buildup.

Source: Genetics: Analysis and Principles (2nd Edition) pp. 425

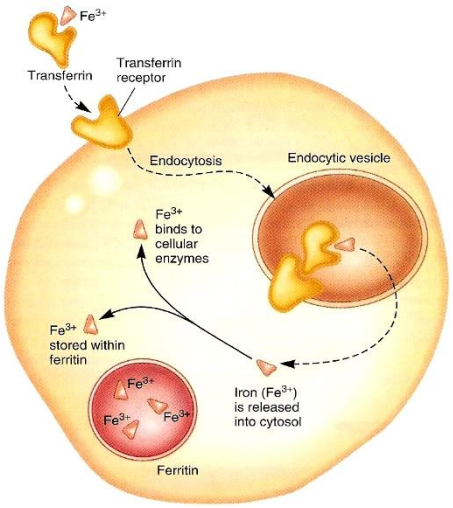
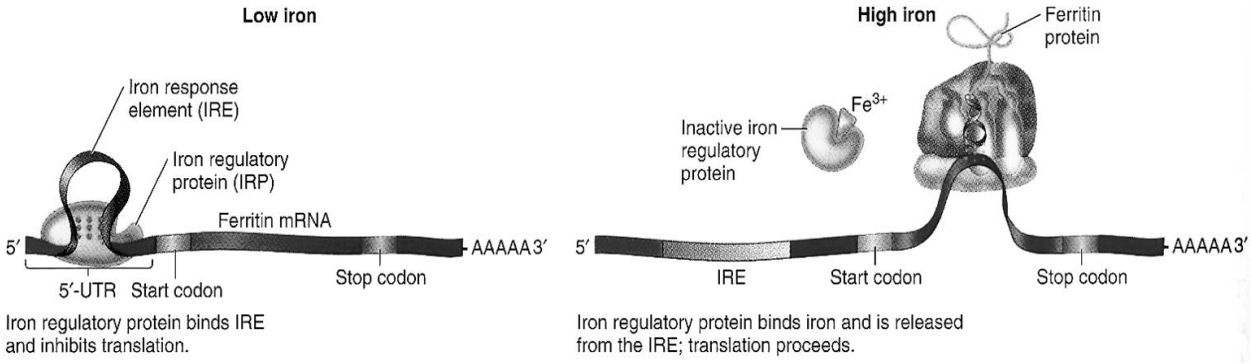


FIGURE 15.26 The uptake of iron (Fe^{3+}) into mammalian cells.

D2. Initiation of Translation

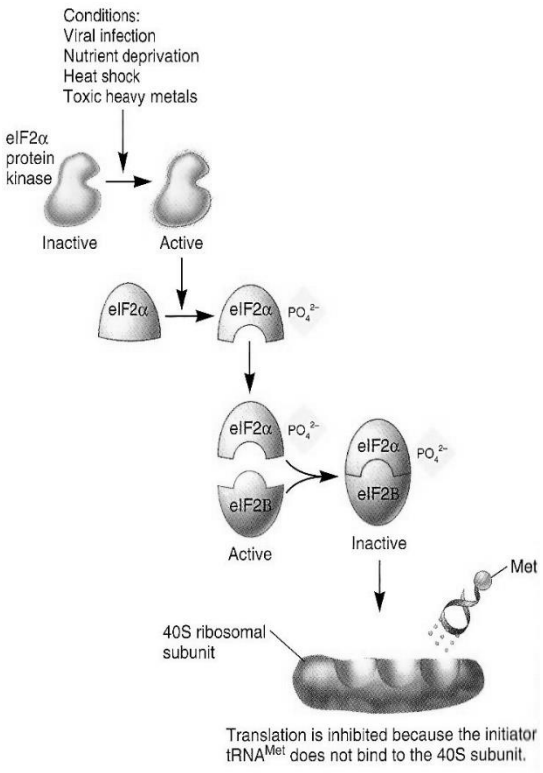


1. The initiation of translation of some mRNAs can be inhibited by **regulatory proteins** that bind to specific sequences within the 5' UTR, thus inhibiting the attachment of ribosomes.
   E.g. Iron regulatory protein (IRP) binds at iron response element (IRE).

2. Both the 5' cap and 3' poly(A tail are important for ribosome binding.
   - In a variety of mRNAs present in the eggs of many organisms, the stored mRNAs initially lack poly(A)tails of sufficient length to allow translation initiation.
   - At the appropriate time during embryonic development, a cytoplasmic enzyme adds more adenine nucleotides, prompting translation to begin.

3. Translation of all mRNAs in a cell may also be regulated simultaneously.

   - In a eukaryotic cell, such global control usually involves the inactivation or activation of one or more **protein factors** required to initiate translation.



Inactivation of eukaryotic initiation factors (eIF)
- Under certain conditions, it is advantageous for a cell to stop protein synthesis.
  E.g. If a virus infects a cell, it is vital to stop protein synthesis so that the virus cannot manufacture viral proteins.
- **Initiation factors** are required to begin translation.
  The phosphorylation of many different initiation factors (e.g. eIF2) have been found to affect translation in eukaryotic cells.
  When the α subunit of eIF2 (eIF2 α) is phosphorylated, translation is inhibited.

Activation of eukaryotic initiation factors
- Observed in the starting of translation of mRNAs stored in eggs.
- Just after fertilisation, translation is triggered by sudden activation of translation initiation factors. The response is a burst of synthesis of proteins that are encoded by the stored mRNAs.

E. Post-translational Control

The final opportunity for control of gene expression occur after translation

E1. Biochemical Modification

Often, eukaryotic polypeptides must be processed to yield functional protein molecules.

1. For example, **cleavage** of the initial insulin polypeptide (pro-insulin) forms the active hormone.

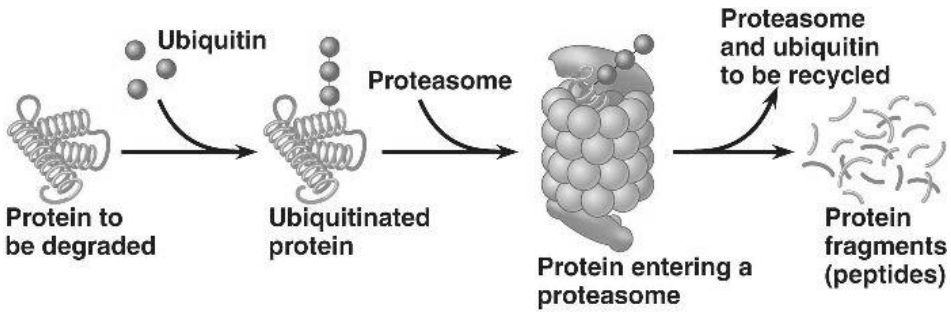
2. In addition, many proteins undergo **chemical modifications** that make them functional, such as
   - **Phosphorylation**
     Regulatory proteins are often activated or inactivated by the reversible addition of phosphate groups.
   - **Glycosylation**
     Proteins destined for the surface of animal cells acquire sugars.

3. Cell-surface proteins and secretory proteins must be transported to target destinations inside / outside the cells in order to function.

4. Regulation may occur at any of the steps involved in modifying or transporting a protein.

E2. Protein Degradation

Finally, the length of time each protein functions in a cell is strictly controlled by means of **selective degradation**.



- For example, proteins such as cyclins involved cell cycle regulation must be relatively short-lived if the cell were to function properly.
- To mark a protein for destruction, a cell commonly attaches molecules of a small protein called **ubiquitin** to the target protein.
- Giant protein complexes called **proteasomes** then recognise the ubiquitin-tagged proteins and degrade them.
- The importance of proteasomes is underscored by the finding that mutations making specific cell cycle proteins impervious to proteasomes degradation can lead to cancer.

F. Gene Amplification

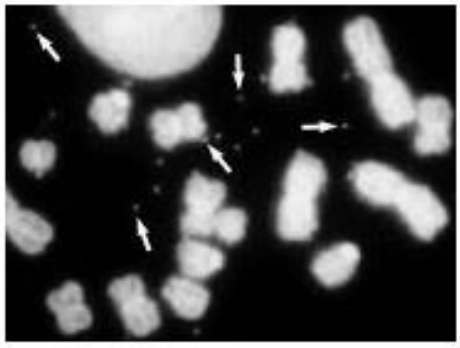
Gene amplification is not the usual physiological means of regulating gene expression, but it does occur in response to certain stimuli if the cell can obtain growth advantage by producing large amounts of a RNA / protein.

Objective of gene amplification: To increase the number of **DNA templates** for transcription.

Expression-related gene amplification
Ribosomal RNA genes in *Xenopus laevis* (South African toad) oocytes.

In eukaryotes, there are four main types of rRNA genes required to form the structural components of ribosomes:

5S gene, which encodes a 120-base rRNA
5.8S gene, which encodes a 160-base rRNA
18S gene, which encodes a 1.8-kb rRNA
28S gene, which encodes a 4.7-kb rRNA



- As each ribosome is made up of one molecule of each type of rRNA (5S, 5.8S, 18S and 28S), all 4 types of rRNAs must be synthesised.
- In oocytes, this transcription must be especially vigorous because the egg needs to accumulate as many as 10^{12} ribosomes to support the early development of the embryo after fertilisation.
- In the genome of *X. laevis*, thousands of 5S genes are scattered all over the chromosomes, but there are far fewer copies of 5.8S, 18S and 28S genes.
- *X. laevis* has 24,000 5S genes, including many that are specifically activated in the oocytes. This enormous set of genes is therefore able to generate the 5S rRNAs needed for ribosome production in the egg.

- The smaller number of 5.8S, 18S and 28S genes (800 to 1000 per diploid cell) is specifically **amplified** in the oocytes by creating **extrachromosomal copies** of these genes.

- Small circular DNAs carrying the 5.8S, 18S and *28S* genes are formed. These replicate by a **rolling-circle mechanism** *(KIV: Organisation of Genome - Prokaryotes)*, producing many copies to form supernumerary nucleoli within the oocytes.

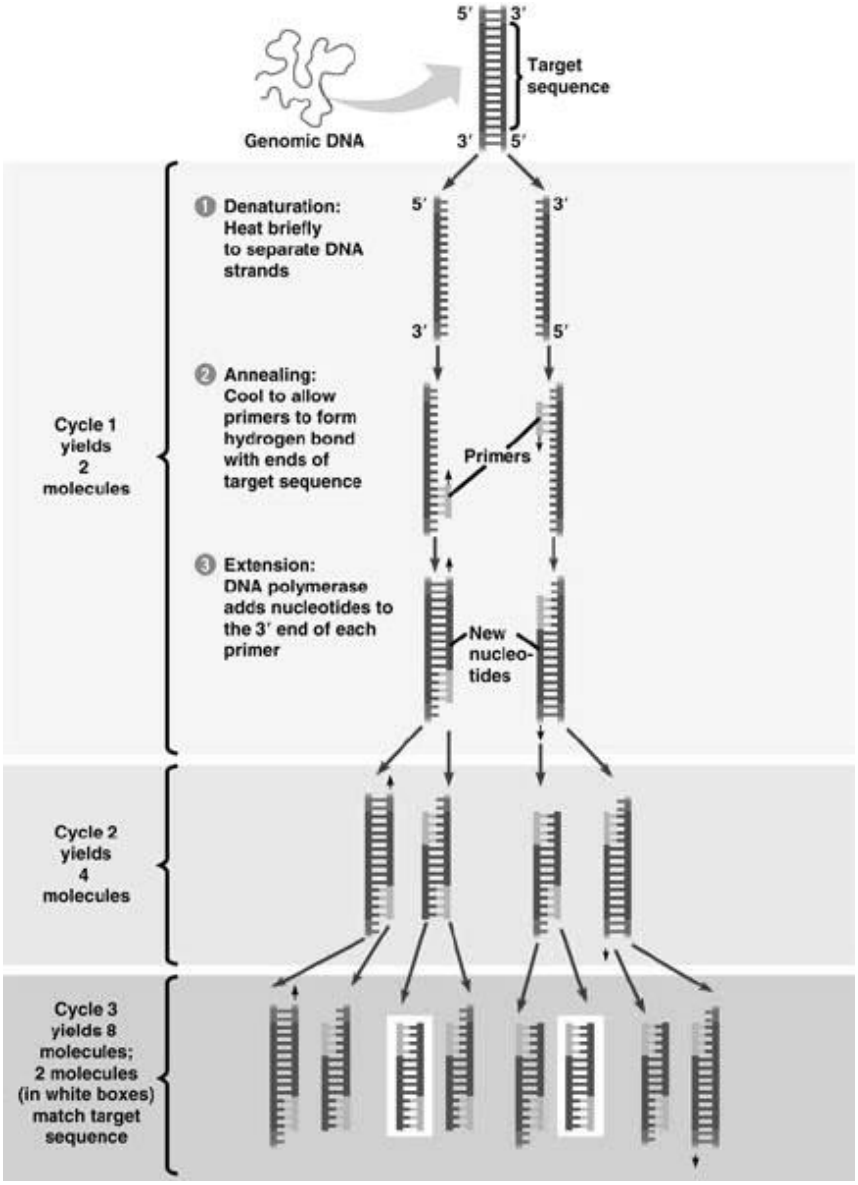
- Transcription from these circular DNAs provides much of the 5.8S, 18S and 28S rRNA required for ribosome assembly in the egg.

II. Molecular Techniques

The investigation of changes in genome structure and activity became possible with the development of various molecular techniques. This section illustrates the three most common laboratory techniques used to study DNA:

- **Polymerase Chain Reaction** (PCR) is used to amplify specific DNA fragment.
- **Gel electrophoresis** is used to separate DNA fragments accordingly to size.
- **Southern blotting and nucleic acid hybridisation** is used to identify specific DNA sequence.

A. Polymerase Chain Reaction



Application

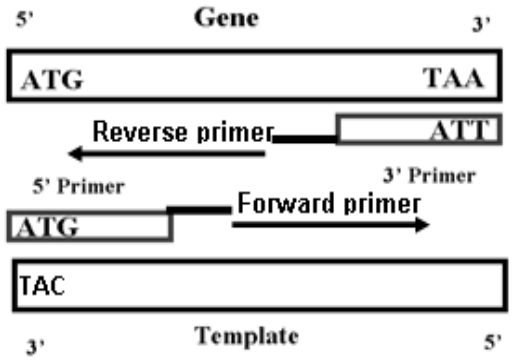
- The **polymerase chain reaction** (PCR) technique, developed by U.S. biochemist Kary Mullis in 1985, is a method for amplifying a specific DNA sequences *in vitro*.
- The method is designed to permit **selective amplification** of a specific target DNA sequence(s) within a heterogeneous collection of DNA sequences.
- Possible applications of PCR include:
  - Amplify specific gene or DNA fragment.
  - Tests for detection of bacteria / virus.
  - DNA fingerprinting.
  - Studying human ancestry.
  - Paternity testing.
  - Cloning genes for use in genetic manipulation.

Procedure

A PCR reaction requires the following elements:

1. Template DNA
   - It is the DNA to be copied.

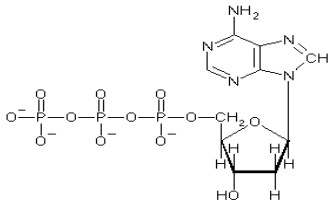
2. Oligonucleotide primers

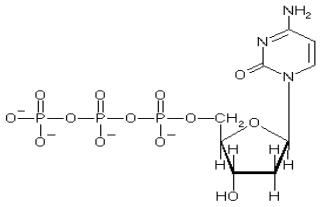


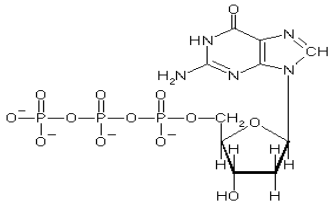
   - The primers are synthetic single-stranded DNA of about 15 - 20 nucleotides long.
   - The primers (forward and reverse primers) are complementary to sequences at the end of the DNA fragment to be amplified.
   - The forward primer will hybridise to the DNA (3' to 5', the coding strand).
   - The reverse primer will hybridise to the DNA (5' to 3', the non-coding strand).

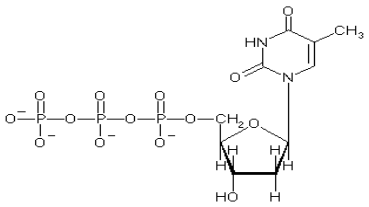
3. DNA polymerase
   - A heat-stable type of DNA polymerase called ***Taq* polymerase**.
   - *Taq* is isolated from *Thermus aquaticus*, a bacterium that lives in hot springs. As the water in this environment is close to the boiling point, all enzymes in *T. aquaticus* have evolved to be stable at high temperatures.
   - The polymerase adds dNTPs to the 3' end of the primer (from 5' to 3' direction) and bases added are complementary to the DNA template.

4. Deoxynucleoside triphosphates (dNTPs)


Deoxyadenosine 5'-triphosphate (dATP)


Deoxycytidine 5'-triphosphate (dCTP)


Deoxyguanosine 5'-triphosphate (dGTP)


Deoxythymidine 5'-triphosphate (dTTP)

- The four dNTPs are dATP, dCTP, dGTP and dTTP.

There are three major steps in a PCR, which are repeated for 30 or 40 cycles.

Step 1: Denaturation.
- An excess of primer is mixed with the DNA fragment to be **amplified**.
- The **thermal cycler** heats the mixture of primer and DNA fragment to 94˚C.
- At this temperature, hydrogen bonds holding the double-stranded DNA fragment break and the double-stranded DNA dissociates into single strands.

Step 2: Annealing of primers.
- The solution is allowed to cool to 65˚C.
- As it cools, the single strands of DNA tend to reassociate into double strands.
- Because of the large excess of both the **forward and reverse primers**, primers base pair with complementary sequences in the single-stranded DNA.

Step 3: Primer extension.
- The thermal cycler then raises the temperature to 72˚C, the temperature at which the *Taq* polymerase functions best.
- Using the primer, the heat stable polymerase copies the rest of the fragment as if it were replicating DNA (nucleotides are added to the 3' ends on both primers).
- The primer is then lengthened into a complementary copy of the entire single-stranded fragment.
- Because both DNA strands are replicated, there are now two copies of the original fragment.

- Steps 1 to 3 are repeated and the two copies become four.
- It is not necessary to add any more polymerase, because the heating step does not denature this particular enzyme.
- Each heating and cooling cycle, which can be as short as 3 or 5 minutes, doubles the number of DNA molecules.
- After 20 cycles, a single fragment produces more than one million (2^{20}) copies. In a few hours, 100 billion copies of the fragment can be manufactured.
  - Number of cycle = n
  - Single fragments produced after n cycle = 2^n
- PCR is carried out in a themal cycler, which can automatically heat and cool the tubes with reaction mixture in a very short time.

Advantages

1. Sensitivity
   - It allows the investigation of minute samples of DNA.
     - In criminal investigation, DNA fingerprints are prepared from the cells in a tiny speck of dried blood or at the base of a single human hair.
     - In pediatric medicine, physicians can detect genetic defects in very early embryos by collecting a few sloughed-off cells and amplifying their DNA.
     - PCR could also be used to examine the DNA of historical figures, such as Abraham Lincoln, or that of now-extinct species, as long as even a minuscule amount of their DNA remains intact.

2. Speed and ease of use
   - DNA cloning by PCR can be performed in a few hours, using relatively unsophisticated equipment.
   - This compares favourably with the time required for cell-based DNA cloning, which may take weeks.

3. Robustness
   - PCR can permit amplification of specific sequences from material in which the DNA is badly degraded or embedded in a medium from which conventional DNA isolation is problematic.
   - As a result, it is suitable for molecular anthropology and palaeontology studies.
     E.g. the analysis of DNA recovered from archaeological remains.
     It has also been used successfully to amplify DNA from formalin-fixed tissue samples, which has important applications in molecular pathology and genetic linkage studies.

Disadvantages

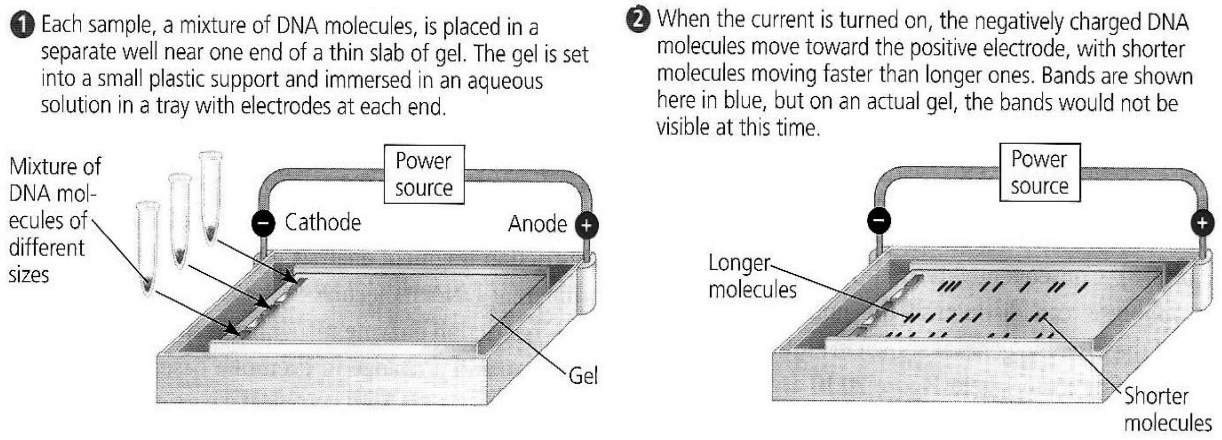
1. Too sensitive
   - Even a tiny amount of contaminant DNA in a sample may become amplified if it includes a DNA sequence complementary to the primers, potentially leading to an erroneous conclusion.

2. Need for target DNA sequence information
   - In order to construct **specific oligonucleotide primers** that permit selective amplification of a particular DNA sequence, some prior sequence information is necessary.

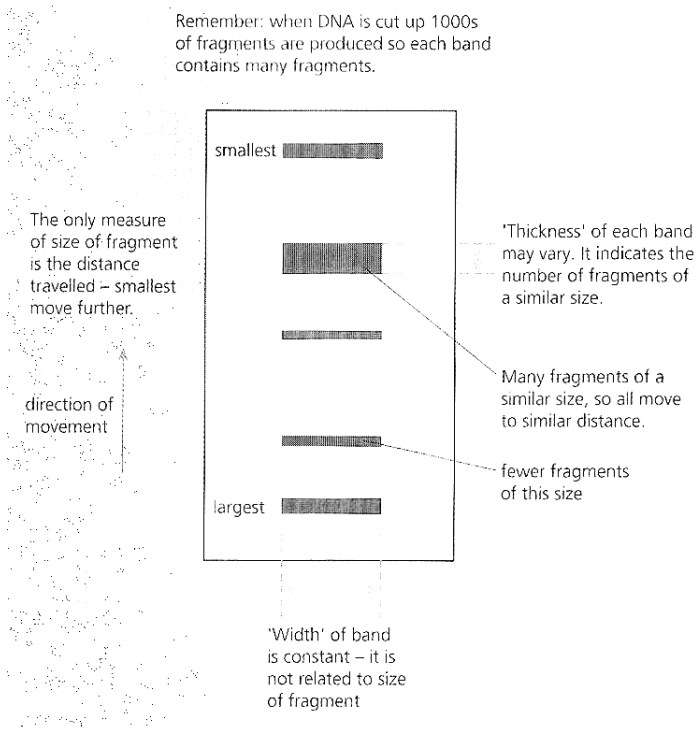
3. Short size and limiting amounts of PCR product
   - A clear disadvantage of PCR as a DNA cloning method has been the size range of the DNA sequences that can be cloned.
   - Unlike cell - based DNA cloning where the size of cloned DNA sequences can approach 2 Mb, reported DNA sequences cloned by PCR have typically been in the 0.1 - 5 kb size range, often at the lower end of this scale.

4. Infidelity of DNA replication
   - Occasional errors during PCR replication limit the number of good copies that can be made by this method.
   - *Taq* DNA polymerase has no associated 3'→5' exonuclease to confer a **proofreading function**, and the error rate due to base misincorporation during DNA replication is rather high.

B. Gel electrophoresis



- **Gel electrophoresis** is a procedure for separating a mixture of molecules on the basis of their rate of movement through a polymeric gel in an electric field.
- The gel acts a molecular sieve to separate nucleic acids or proteins on the basis of size, electrical charge and other physical properties.
- A gel material consisting of agarose is poured as a thin slab on a glass or Plexiglas holder.
- After the gel has solidified, the researcher loads samples containing a mixture of macromolecules of different sizes in wells formed at one end of the gel.
- The gel is bathed in an aqueous solution and electrodes are attached to both ends and an electrical current is applied.



- The DNA molecules, which are negatively charged due to their phosphate groups, migrate toward the positive electrode, the anode.
- The distance a DNA molecule travels is inversely proportional to its length. Larger molecules travel more slowly through the pores in the gel, thus the smallest DNA fragments travel the longest distance.
- The gel slows down the large molecules more than the smaller molecules because the large molecules have more difficulty in moving through the pores in the gel.

- When the current is turned off, the DNA molecules in each sample are arrayed in bands along a "lane" according to their size. The shortest molecules, having travelled the furthest, are in bands at the bottom of the gel.
- These bands of separated molecules can be visualised with stains. E.g. ethidium bromide is a stain that binds to DNA and RNA; and can be seen under ultraviolet light.
- The actual size of the DNA fragments is determined by comparison to migration distances of known **marker fragments** (e.g. size markers) that are subjected to electrophoresis in an adjacent lane of the gel.

Application of Gel Electrophoresis

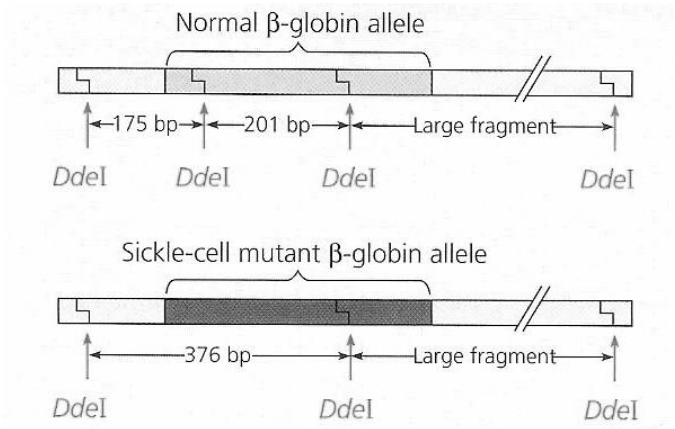
Identification of normal / mutant allele

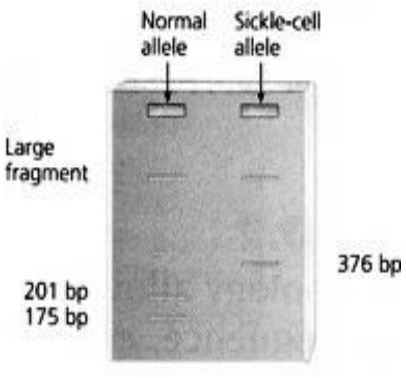
- Gel electrophoresis may be used to compare two alleles of the same gene.
  - A restriction enzyme recognises and cuts DNA molecules at a specific base sequence (restriction site).
  - A change in even 1 bp will prevent the restriction enzyme from cutting that particular restriction site.
  - Hence, if the base sequence difference between two alleles occurs within a restriction site, cutting with the restriction enzyme that recognises the site will produce a different mixture of fragments from each allele. Each mixture will give its own **band pattern** in gel electrophoresis.

Figure shows the restriction site for the restriction enzyme *Eco*RI. Restriction enzymes always cut DNA at a specific sequence of DNA.

Source: https://www.labxchange.org/library/items/lb:LabXchange:adb21ebe:lx_image:1

Analysis of the normal and sickle-cell alleles of the β-globin gene.





- Sickle-cell anaemia is caused by a point mutation (base substitution) within a restriction site in the human β-globin gene.
- The point mutation destroys one of the *Dde*I restriction sites at the 5' end of the β-globin gene.
- Gel electrophoresis may be used to distinguish the normal and sickle-cell alleles of the β-globin gene.
- As such, this technique may also be used to determine the genotype of a person for the β-globin gene.

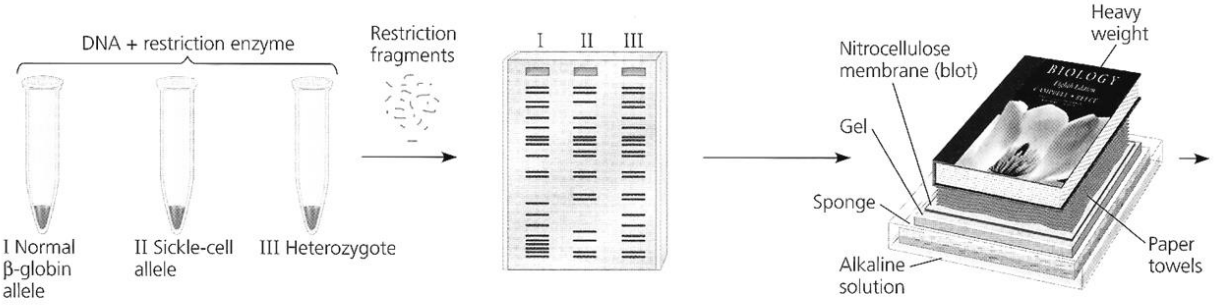
C. Southern Blot and Nucleic Acid Hybridisation

- Suppose that the DNA of the whole genome (genomic DNA) is cleaved with the restriction enzyme and the investigator wants to know the size of the fragment that contains the gene of interest.
- When stained with ethidium bromide, the thousands of DNA fragments generated by cutting the genome are too numerous to resolved into discretely visible bands, and they look like a smear.
- The technique of **Southern blot** (a method which combines gel electrophoresis and nucleic acid hybridisation) will identify within the smear the size of the particular fragment containing the gene of interest.

Southern Blotting of DNA Fragments

APPLICATION Researchers can detect specific nucleotide sequences within a complex DNA sample with this method. In particular, Southern blotting is useful for comparing the restriction fragments produced from different samples of genomic DNA.

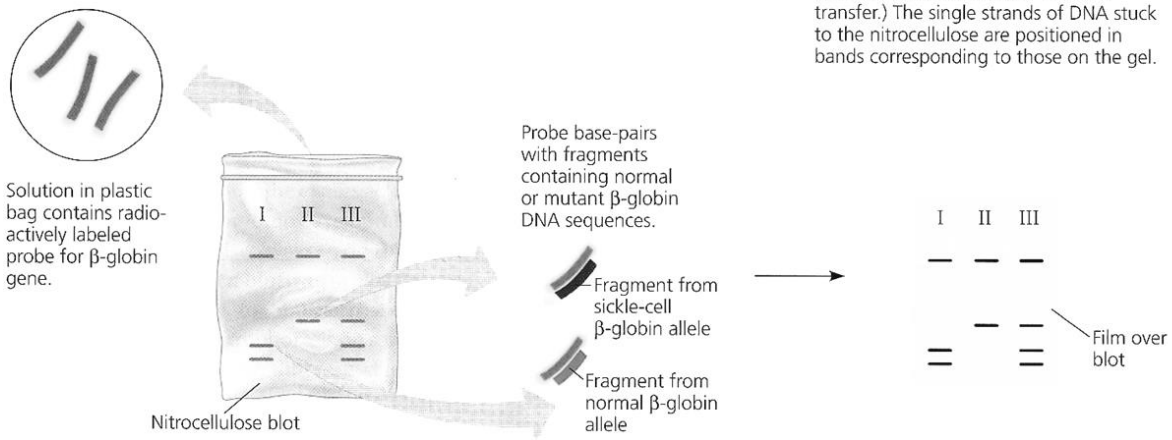
TECHNIQUE In this example, we compare genomic DNA samples from three individuals: a homozygote for the normal β-globin allele (I), a homozygote for the mutant sickle-cell allele (II), and a heterozygote (III). As in Figure 20.7, we show a radioactively labeled probe, but other methods of probe labeling and detection are also used.



1 Preparation of restriction fragments. Each DNA sample is mixed with the same restriction enzyme, in this case *Dde*I. Digestion of each sample yields a mixture of thousands of restriction fragments.

2 Gel electrophoresis. The restriction fragments in each sample are separated by electrophoresis, forming a characteristic pattern of bands. (In reality, there would be many more bands than shown here, and they would be invisible unless stained.)

3 DNA transfer (blotting). With the gel arranged as shown above, capillary action pulls the alkaline solution upward through the gel, transferring the DNA to a nitrocellulose membrane, producing the blot; the DNA is denatured in the process. (Another version of this method uses an electrical current to hasten the DNA transfer.) The single strands of DNA stuck to the nitrocellulose are positioned in bands corresponding to those on the gel.



4 Hybridization with radioactive probe. The nitrocellulose blot is exposed to a solution containing a radioactively labeled probe. In this example, the probe is single-stranded DNA complementary to the β-globin gene. Probe molecules attach by base-pairing to any restriction fragments containing a part of the β-globin gene. (The bands would not be visible yet.)

5 Probe detection. A sheet of photographic film is laid over the blot. The radioactivity in the bound probe exposes the film to form an image corresponding to those bands containing DNA that base-paired with the probe.

RESULTS Because the band patterns for the three samples are clearly different, this method can be used to identify heterozygous carriers of the sickle-cell allele (III), as well as those with the disease, who have two mutant alleles (II), and unaffected individuals, who have two normal alleles (I). The band patterns for samples I and II resemble those observed for the purified normal and mutant alleles, respectively, seen in Figure 20.10b. The band pattern for the sample from the heterozygote (III) is a combination of the patterns for the two homozygotes (I and II).

Procedure

1. Restriction fragment separation.
   - Genomic DNA samples are isolated from three individuals.
   - A restriction enzyme is added to the three samples of DNA to produce restriction fragments.

2. Gel electrophoresis.
   - The restriction fragments from each sample are then separated by electrophoresis.
   - Each sample forms a characteristic pattern of bands.

3. DNA transfer (**blotting**).
   - These fragments will then be transferred by capillary action (blotting) from the gel to a sheet of **nitrocellulose membrane**.
   - **Capillary action** pulls an alkaline solution upward through the gel and through a sheet of nitrocellulose membrane laid on top of it, transferring the DNA to the membrane and denaturing it in the process.
   - The single strands of DNA stick to the membrane, positioned in bands exactly as on the gel.

4. **Hybridisation** with radioactive probe.
   - The blot is exposed to a solution containing radioactively labelled probe.
   - The probe is single-stranded DNA complementary to the DNA sequence of interest, and it attaches by base-pairing to restriction fragments of complementary sequence.

5. **Autoradiography**.
   - A sheet of photographic film (X-ray film) is laid over the blot.
   - The radioactivity in the bound probe exposes the film to form an image corresponding to specific DNA bands containing the DNA that base-paired with the probe.