RAFFLES INSTITUTION

H2 BIOLOGY LECTURE NOTES

Syllabus 9744

2018 – 2019

2

CORE IDEA 2: GENETICS & INHERITANCE

<u>Topics</u>

1. Mitosis & Meiosis
2. DNA & Genomics
   a. Part I
   b. Part II
3. Viruses
4. Bacteria
5. Prokaryotic & Eukaryotic Genome
   a. Part I: Organisation
   b. Part II: Control
   c. Part III: Cancer

Name: **Class:**

CORE IDEA
(2) Genetics and Inheritance

MITOSIS & MEIOSIS

Content

- Replication and division of nuclei and cells
- Understanding of chromosome number and variation
- Effect of meiosis on chromosome number and variation

Learning Outcomes

2(l) Explain what is meant by the terms *gene mutation* and *chromosome aberration*.
For gene mutation, knowledge of how substitution, addition and deletion could change the amino acid sequence (e.g. frameshift) is required.
– to be covered in DNA & Genomics.

For chromosomal aberration, knowledge of numerical (e.g. aneuploidy, as in the case of trisomy 21, i.e. Down syndrome) and structural (e.g. translocation, duplication, inversion, deletion) aberration is required.

(n) Describe the events that occur during the mitotic cell cycle and the main stages of mitosis (including the behaviour of chromosomes, nuclear envelope, cell membrane and centrioles).

(o) Explain the significance of the mitotic cell cycle (including growth, repair and asexual reproduction) and the need to regulate it tightly. (Knowledge that dysregulation of checkpoints of cell division can result in uncontrolled cell division and cancer is required, but detail of the mechanism is not required.)

(s) Describe the events that occur during the meiotic cell cycle and the main stages of meiosis (including the behaviour of chromosomes, nuclear envelope, cell membrane and centrioles). (Names of the main stages are expected, but not the sub-divisions of prophase.)

(t) Explain the significance of the meiotic cell cycle (including how meiosis and random fertilisation can lead to variation).

References

- Campbell, N.A. and Reece, J.B. (2014). Biology, 10th edition. Pearson.
- Jones & Jones. (1997). Advanced Biology.
- Hoh Yin Kiong (2003). Longman A-Level Course in Biology, Vol. 1. Longman
- Green, Stout and Soper (1990). Biological Sciences, Vol 1. Cambridge
- Raven, Johnson, Losos, Mason & Singer (2008). Biology, 8th edition. McGraw-Hill.

*** This handout is the effort of several Biology teachers at RI. It has been and will continue to be updated.**
**** Any information given in a double-lined box is for your information only.**

Last updated by: Mrs S Nair, Mr Derek Tan and Mr Low CM

Table of Contents

(A) CELL DIVISION

The modern cell theory states that all new cells are derived from other cells. Many cells in a multi-cellular organism undergo a well-defined sequence of stages culminating in the division and formation of new cells. The processes involved are collectively grouped under the term **cell division.**

There are some types of cells that are continuously dividing throughout the lifetime of the organism (e.g. bone marrow cells, epithelial cells of the skin, etc.) whilst others stop dividing upon reaching maturity (e.g. neurones, muscle cells, etc.).

Cell division occurs in 2 main steps:

- **Nuclear division &**
- **Cytokinesis (cytoplasmic division)**

} **Cell division = Nuclear division+ Cytokinesis**

1) Nuclear Division (= division of the nucleus):

- There are **2 types of nuclear division: Mitosis & Meiosis**

2) Cytokinesis (= division of the cytoplasm)

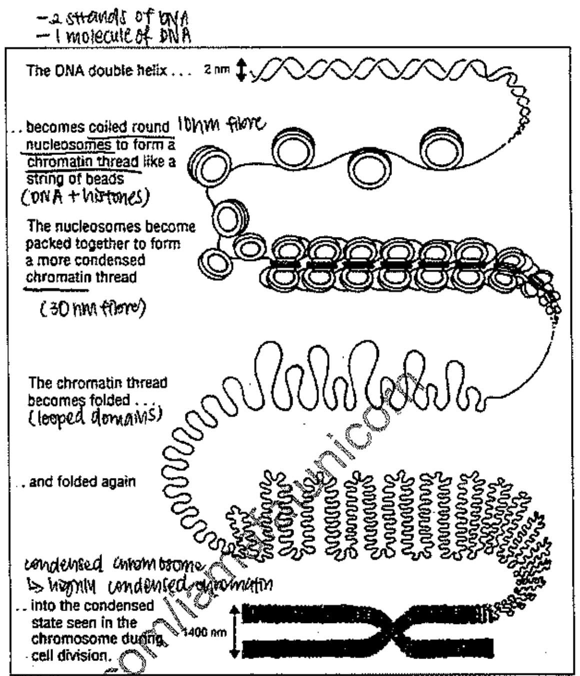
- Cytokinesis is the division of the cytoplasm to form 2 separate daughter cells immediately after mitosis, meiosis I or meiosis II.

⚠ Please note that mitosis and meiosis are 2 types of nuclear divisions and hence only occur in eukaryotic cells and not prokaryotic cells.

(B) CHROMOSOME STRUCTURE

- Chromosomes carry the hereditary material (DNA) in cells.
- The hereditary material is passed on to the next generation through cell division.
- The structure of the eukaryotic chromosome:
  - **DNA (deoxyribonucleic acid) is the double-stranded, helical** molecule found within the nucleus of each cell. DNA carries the **genetic information** that **codes for** the **proteins** which are necessary for cells to reproduce and to perform their functions.
  - When a cell is not dividing, **chromosomes** exist in their dispersed, **uncondensed** form, as a **mass of long, thin, thread-like fibres** known as **chromatin.**
  - **Chromatin** is a complex of **DNA** and **histone proteins**. The DNA winds around an **octamer** formed by **8 histone proteins** forming **nucleosomes.**
  - Chromatin that has been condensed by **coiling/folding** many times upon itself results in a **chromosome** which appears as a thicker, shorter and more visible structure.

Levels of packing of chromatin in a chromosome

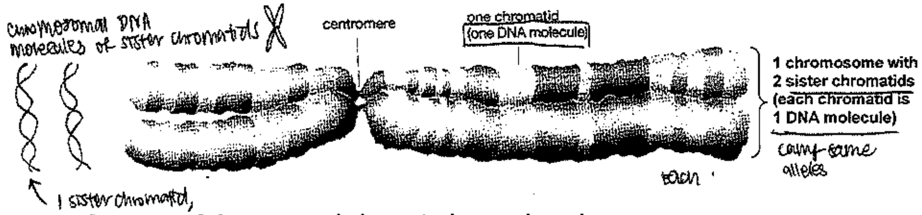


chromosome → uncondensed form → chromatin appears as a mass of long, thin, thread-like fibres

chromosome → condensed form → chromatin is tightly coiled & folded into a compact, thick & short structure

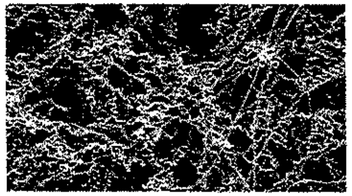
- **Condensed chromosomes** are found in dividing cells during some stages of mitosis and meiosis, such as late prophase, metaphase, anaphase, early telophase etc. During these stages, the **chromatin is condensed** and more tightly coiled and folded such that the chromosome appears thicker and shorter and more visible.
- **Sister chromatids** contain **identical DNA molecules** as they are **replicated from the same DNA molecule.**

  After DNA has replicated during interphase, the chromatin condenses during prophase to form two sister chromatids held together at the **centromere.**



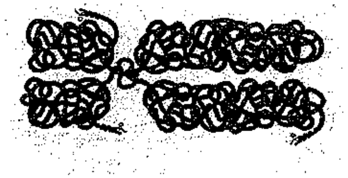
Structures of chromosome during metaphase and anaphase

(a) Chromatin fibers at interphase



(b) Chromosomes, each made up of 2 sister chromatids

(c) Diagrammatic picture of chromosome showing how chromatin is condensed

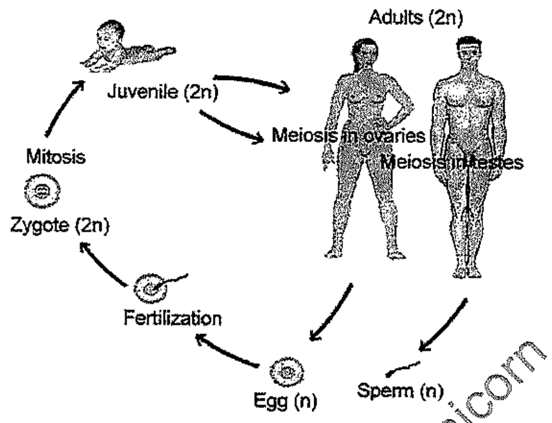


(d) Chromosome with 2 sister chromatids

(C) DIPLOID AND HAPLOID

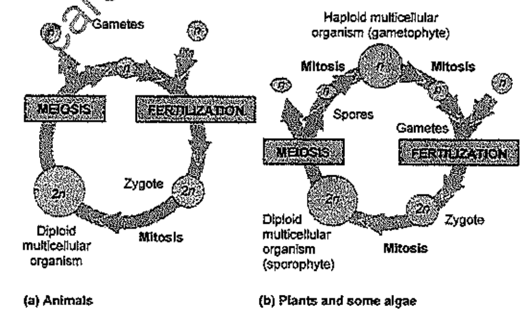
- The term **diploid** is used to describe a nucleus, cell or organism with **two complete sets** of chromosomes. The **chromosomes exist as homologous pairs where each chromosome of the pair comes from either parent** (maternal & paternal). Most organisms exist in a diploid condition.
- The diploid condition is represented as **2n**, where **n** represents **1 complete set of chromosomes.**
- **Somatic** (non-gametic) cells are diploid e.g. the diploid number of chromosomes for humans is 46.
- The term **haploid** is used to describe a nucleus, cell or organism that contains only **one complete set of chromosomes.** Thus it has **half the diploid number of chromosomes** and contains **one homologue of each homologous chromosome pair.**
- The haploid condition is represented as **n.** e.g. the haploid number of chromosomes for the gametes in human is 23.
- **Gametes** are haploid.

Organism	Diploid chromosome number	Haploid chromosome number
Fruit fly (*Drosophila melanogaster*)	8	4
Onion (*Allium cepa*)	6	3
Maize (*Zea mays*)	20	10
ocust (*Locusta migratoria*)	24	12
ily (*Lilium longiflorum*)	24	12
Tomato (*Solanum lycopersicum*)	4	2
Mouse (*Mus musculus*)	40	20
Human (*Homo sapiens*)	46	23
Potato (*Solanum tuberosum*)	48	24



Human life cycle

- The **fusion** of a **haploid** sperm and **haploid** egg during **fertilisation** results in the formation of a **diploid zygote**. This new cell is "diploid" because it contains the two haploid sets of chromosomes, **one set supplied by the mother and one by the father.** These chromosomes bear the ancestral genes that represent the maternal and paternal family lines.
- After fertilisation, the zygote undergoes a process of nuclear division called **mitosis.** This generates cells that are **genetically identical** to the original zygote. These cells are then stimulated to **differentiate** into **specialised** cells that form the organism.

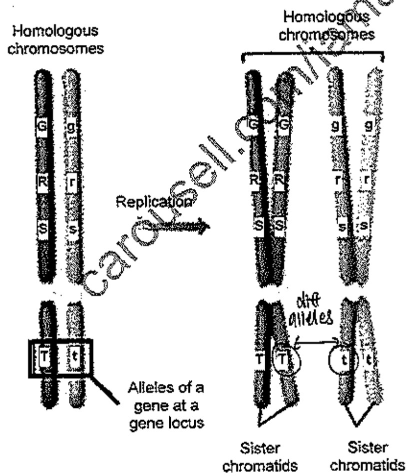


Animal and plant sexual life cycle

(D) HOMOLOGOUS CHROMOSOMES

- In a **diploid** cell, two chromosomes having the same **size, shape, centromere position, staining pattern**, and **position of genes** are known as **homologous chromosomes.**
- Each member of such a pair is called a **homologue.**
- Characteristics of homologous chromosomes:
    1. They have the **same genes** (that determine the **same characters**) at corresponding **loci**. e.g. blood group, hair colour.
       Note: Locus (singular)/ loci (plural) – is the fixed position of a gene in a chromosome
    2. One homologue originates from the **male** parent and the other from the **female** parent.
    3. They are similar in **size, shape, centromere position** and **staining pattern.**
    4. However they **may not be identical** in what they code for, thus they may have **different alleles** at the same locus.

       (**Alleles** are **alternative forms of a gene** and hence occupy the same loci on homologues. For example, a gene that controls hair colour, one homologue may carry an allele that codes for brown hair and the other homologue may carry the allele for blonde hair.)



e.g. T and t are alternative forms of the gene coding for the same trait of hair colour.

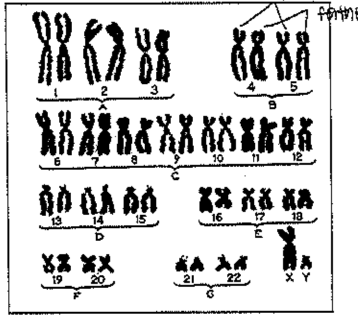
T allele codes for brown hair;
t allele codes for blonde hair

Homologous chromosomes

⚠ Sister chromatids, will ***always*** have the **same alleles** if no crossing over occurs between non-sister chromatids of homologous chromosomes.

⚠ Homologous chromosomes have the **same genes** but **may have different alleles** as one homologue is from the father while the other is from the mother.

- The number of chromosomes varies from one species to another but is always the same for normal individuals of one species. However, some species may share the same chromosome *number*, but the *types* of chromosomes are different. Humans have 46 chromosomes or **23 pairs** of chromosomes.



Human karyotype of a man with 23 pairs of chromosomes

Question: Are all the chromosome pairs homologous?

No. X and Y are not. (as they have different sizes, centromere position, etc)

Question: A chicken has 78 chromosomes in its somatic cells.
a) How many chromosomes did the chicken inherit from each parent? 39
b) How many chromosomes are in each of the chicken's gametes? 39
c) How many chromosomes are in each cell of the embryo? 78
d) How many chromosomes are in one "set" / haploid? 39

(E) FACTORS AFFECTING CELL DIVISION

	Factors	How they affect cell division
1.	**Surface area to volume ratio**	As a cell enlarges, its volume increases faster than its surface area. The cells in a given tissue are stimulated to divide when the ratio of surface are to volume exceeds a critical figure. Why? ***When the surface area is not great enough to accommodate the entry of food materials and oxygen at a rate sufficient to supply the cell's demands, the cell divides in order to restore a favourable surface area to volume ratio.***
2.	**Nucleo-cytoplasmic ratio**	When a cell increases beyond a certain size, the cell will divide. Why? ***As there is a limit to the amount of cytoplasm which a nucleus can control, the cell will divide so as to restore a more favourable nucleo-cytoplasmic ratio. By doing so, it will enable the nucleus to effectively direct and control the many activities in the cytoplasm.***
3	**Chemical substances**	Growth hormones in plants and animals are required to initiate cell divisions and determine the rate of cell division e.g. auxins in plants and thyroxine in vertebrates. In humans, how is cell growth stimulated in an injured area? ***Wounds usually causes the production of chemicals that stimulate cell division to grow over the injured area.***

Other factors influencing cell division includes age, temperature and mechanical stimuli. An example of mechanical factor is abrasion which removes cells from skin surfaces and this stimulates the cells beneath to divide more rapidly.

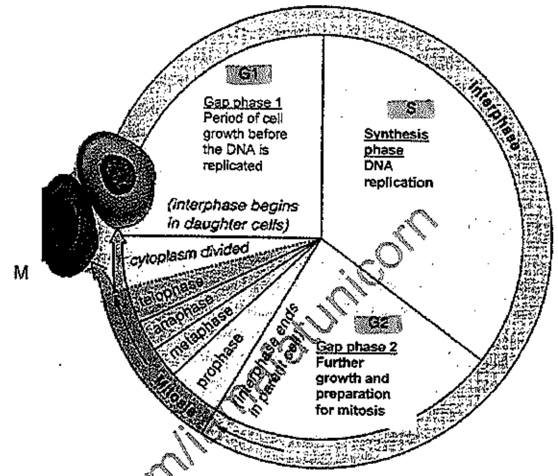
① SA:VR → surface area cannot accomodate entry of fd materials & O_2 fast enough to sufficiently supply cell's demands
② nucleo-cytoplasmic ratio → enable nucleus to better control & direct activities within cytoplasm (limit to amt of cytoplasm the nucleus can control)
③ Chemical substances

(F) THE CELL CYCLE

The cell cycle is the sequence of events which occurs between the formation of a cell and its division into daughter cells.

<u>Cell cycle</u> - A typical cycle.

The two gap phases (G_1 and G_2) are separated by the synthesis (S) phase, during which the DNA is replicated. Mitosis(M) phase, or nuclear division, is followed by cytoplasmic division (cytokinesis).



- The cell cycle consists of the following 3 main stages:

Stages in cell cycle	Main events
Interphase (<u>longest</u> phase, 90% of cell cycle)	- Period of **synthesis** and **growth** - Cell produces **many materials** and **organelles** required for carrying out all its functions - Cell **replicates** its **DNA** (during S phase of interphase) to prepare for nuclear division
Nuclear division	- Either mitosis or meiosis
Cytokinesis (cytoplasmic cleavage)	- Division of **cytoplasmic contents** into 2 daughter cells

(Note: Cells do not normally divide continuously i.e. a certain period will elapse between the two divisions.)

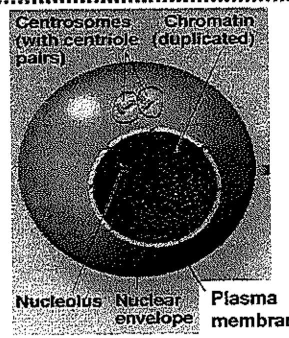
(G) INTERPHASE (Preparation for mitosis/meiosis; non-dividing phase; 90% of cell cycle)

Phase	Events within the cell
G_1 (Gap phase 1)	**Intensive cellular synthesis:** a) Organelle synthesis b) RNA synthesis c) Protein synthesis d) ATP synthesis
S (Synthesis phase)	**DNA replication occurs:** *(semi-conservative replication) → results in identical sister chromatids a) DNA molecules replicate → DNA content of the cell doubles
G_2 (Gap phase 2)	**Intensive cellular synthesis** (in preparation for mitosis) a) Organelle synthesis b) Synthesis of spindle proteins c) ATP synthesis

(H) MITOSIS

- **Mitosis** is a form of **nuclear division** in eukaryotic cells which produces **two daughter nuclei** containing **identical sets of chromosomes** as the parental cell nucleus.
- It occurs after interphase (if the conditions are right).
- It is usually followed immediately by **cytokinesis**, during which an equal division of the cytoplasm of the parent cell and formation of the cell membrane and cell wall (in plants) results in the formation of two daughter cells.
- Mitosis is made up of 4 main stages: **P**rophase, **M**etaphase, **A**naphase, **T**elophase.

1) INTERPHASE (not part of mitosis) *



- The longest part of the cell cycles
- Includes G_1, S and G_2 phases
- By the end of interphase:
  - **DNA duplicated**
  - Nucleus is bound by **nuclear envelope**
  - **Nucleolus** present
  - Single **centrosome replicated** to form two centrosomes, each containing a pair of centrioles in animals
  - Organelles duplicated

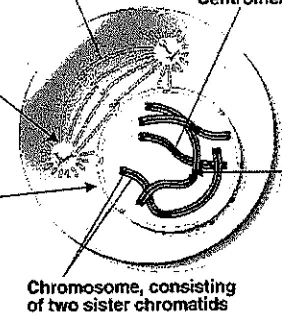
⚠ Centrioles are absent in higher plant cells. Use the term **microtubule organising center** instead.

2) PROPHASE (EARLY)

Spindle fibres begin to form.

Centrioles (in centrosome) move to opposite poles of the cell.

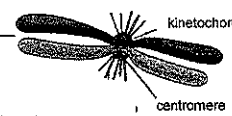
Nucleolus begins to disappear and **nuclear envelope** begins to disintegrate.



Chromatin in the process of **coiling**, and **shortening and thickening** into a condensed **chromosome**.

3) PROPHASE (LATE)

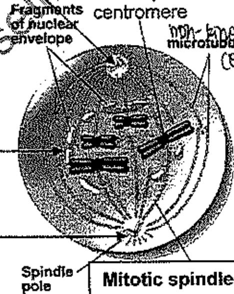
Centromere – consists of a sequence of **tandem repeats** on a region of the chromosome that holds sister chromatids together. A specialised protein complex, called the **kinetochore**, assembles at the centromere and serves as an attachment site for **spindle fibres**.



Centrioles (in centrosome) migrate to opposite poles of the cell.

Nuclear envelope disintegrates

Asters are radial arrays of shorter microtubules that extend from the centrosome.



Chromosomes are now even more condensed and visible and appear as **two sister chromatids** joined together at the **centromere**

Mitotic spindle continues to form

Spindle fibres extend from each pole toward i) kinetochores and ii) the middle region of the cell (metaphase plate).

Kinetochore microtubules **attach to centromeres of the chromosomes.**

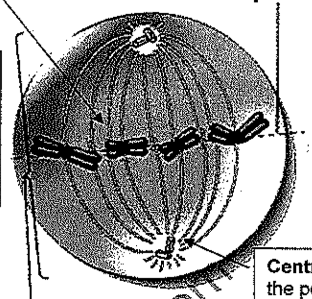
▸ No. of chromosomes per cell = 4
▸ No. of DNA molecules per cell = 8

4) METAPHASE – "meet"

Chromosomes are aligned at the equator / metaphase plate in a single row.

Spindle is **completely formed.**
Kinetochore microtubules are attached to centromeres of the chromosomes.
The attachment of 1 centromere to 2 kinetochore microtubules (1 from each pole of the cell) helps to position chromosomes along the metaphase plate.

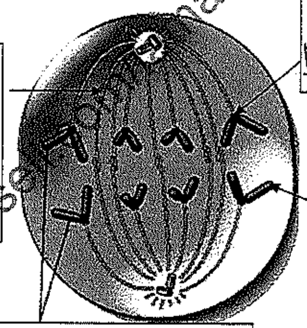
⚠ If the cell is diploid, homologous pairs are present. However they do not pair up. Instead they **line up singly in a row.**



Centrioles located at the poles of the cell.

5) ANAPHASE (shortest phase of mitosis) – "apart"

Spindle fibres (**non-kinetochore microtubules**) elongate and slide in opposite direction due to the actions of motor proteins. This cause the two poles to move further apart, **elongating the cell.**



Centromere of the chromosome **divides.** * chromosomes separate *

Daughter chromosomes are **led by their centromeres** resulting in the characteristic **'V' shape** of chromosomes seen in anaphase.

Separated sister chromatids, now called **daughter chromosomes** are pulled to **opposite poles** by **shortening** spindle fibres/**kinetochore microtubules.**

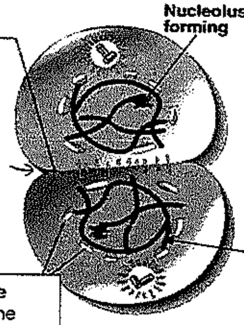
* use proper terminology!!

- No. of chromosomes per cell = 8
- No. of DNA molecules per cell = 8

6) TELOPHASE (considered as 1 cell forming 2 nuclei) – "the end"

Cleavage furrow (in animal cells) or **cell plate** (in plant cells) starts to form.

invagination of plasma membrane



Spindle fibres disintegrate.

- No. of chromosomes per cell = 8
- No. of DNA molecules per cell = 8

Nuclear envelope reforms around the chromatin at each pole and the **nucleolus reappears.**

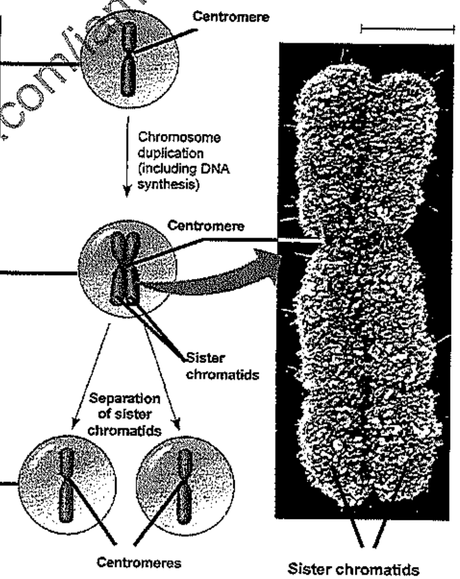
Daughter chromosomes reach the poles of the cell, and will **decondense** and lengthen into **chromatin**. The chromosomes will then appear diffused and are not clearly visible.

Review: Chromosome duplication and distribution during cell division

A eukaryotic cell has multiple chromosomes, one of which is represented here. Before replication, each chromosome is made up of a **single DNA molecule** with a single centromere.

Upon replication, a chromosome consists of two **genetically identical** sister chromatids connected at the centromere. Each sister chromatid comprise1 DNA molecule.

Division of the centromere and shortening of kinetochore microtubules separate the **sister chromatids** into **two daughter chromosomes** and distribute them to two daughter cells.



Note: all the chromosomes above are drawn as 'metaphase-stage' chromosomes for ease of viewing, not all of them should look like that – the topmost chromosome should appear in chromatin form because of the stage it is in i.e. before duplication (S phase of interphase)

Question: Fill in the blanks.
where n = no. of chromosomes and X = amount of DNA

S phase (semi-conservative DNA replication)

	In a cell		
	No. of chrs.	Amt. of DNA	No. of DNA molecules
G_1 phase	2n	X	4
Prophase	2n	2X	8
Metaphase	2n	2X	8
Anaphase	4n	2X	8
Telophase	4n	2X	8
Cytokinesis	2n	X	4

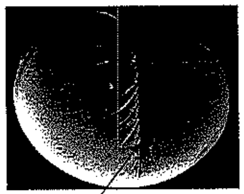
⚠ 1 chromatid = 1 DNA molecule; chromosome with 2 sister chromatids has 2 DNA molecules. After anaphase, 1 chromosome has 1 chromatid thus 1 DNA molecule. Each DNA molecule is made up of 2 strands of DNA.

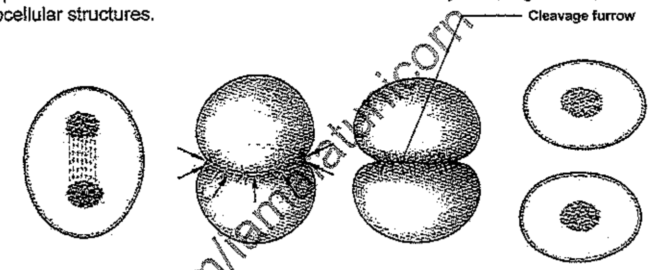
(I) CYTOKINESIS (CYTOPLASMIC DIVISION)

- Cytokinesis generally begins at **telophase** during which the cytoplasm and cell organelles of the parent cell are evenly distributed between the resulting daughter cells. (Although by definition cytokinesis and telophase are 2 separate stages, they may overlap temporally, ie the 2 processes may start at the same time.)

a) In animal cells

- The cell membrane begins to **invaginate** towards the region previously occupied by the equator / metaphase plate, forming a **cleavage furrow.**
- The cleavage furrow deepens until the parent cell is **pinched** into two, producing two completely separated cells each with its own nucleus and share of cytosol, organelles, and other subcellular structures.



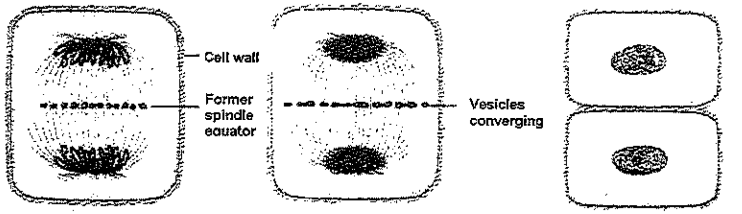


a. Mitosis is over, and the spindle is now disassembling
b. Just beneath the plasma membrane, the cell membrane begins to invaginate forming a cleavage furrow
c. The contractions continue, and cut the cell into two, forming two daughter cells

b) In plant cells

- In the cells of the higher plants, a series of **fluid-filled vesicles** (Golgi vesicles) (derived from the Golgi apparatus), move to the equator/metaphase plate of the cell and **coalesce** (i.e. fuse) to form **cell plate.**
- The contents of the vesicles will be converted to pectin and cellulose, which contribute to the middle lamella and cell wall matrix respectively, of the daughter cells. The membranes of the vesicles form the cell surface membranes of the daughter cells.
- The cell plate eventually fuses with the parent cell wall and cell membrane, separating the daughter cells.

Golgi vesicles:
- contents of Golgi vesicles → cell wall
- membranes → plasma membrane



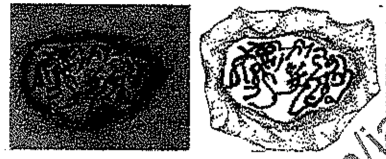
a. The Golgi apparatus produces a number of small fluid-filled vesicles, which first appear in the centre of the cell.

b. Guided by microtubules, the vesicles coalesce to form a cell plate which grows across the equatorial plane.

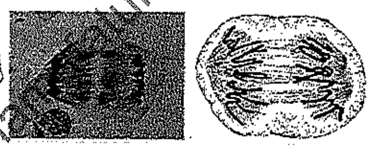
c. The contents of the Golgi vesicles contribute to the new cell walls of the daughter cells whilst their membranes form the new cell membranes. The spreading cell plate eventually fuses with the parent cell wall and separates the two daughter cells.

Review questions

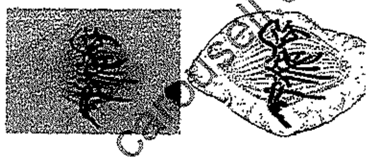
1. Identify the stages of mitosis shown in the pictures below:



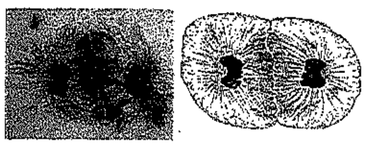
Phase: prophase



Phase: Late anaphase



Phase: metaphase



Phase: Telophase

Phase: Anaphase

Phase: metaphase

2. Name the chemical that inhibits spindle fibre formation.

Colchicine

(J) SIGNIFICANCE OF MITOSIS

a) Maintaining genetic stability

- Mitosis produces 2 daughter nuclei, each of which will have **the same number and same types of chromosomes** as the parent cell. (genetically identical to parent cell)
- Each daughter nucleus eventually becomes part of one of daughter cell.
- Since the chromosomes in both nuclei were derived from the parental chromosomes by **semi-conservative replication of their DNA** and **subsequently their even distribution**, the chromosomes in the two daughter cells are **genetically identical**.
- Thus mitosis produces daughter cells that are **genetically identical** to their parent cell.
- Mitosis **does not introduce genetic variation**, thus maintaining **genetic stability** within the populations of cells derived from the same parental cells.
- In animal, it **will not result in rejection by the body immune system** as the daughter cells produced are genetically identical to the parent cell, thus helping in recognizing **self- versus non-self cells**.

b) Growth

- Mitosis takes place during growth of an organism. Growth is defined as an increase in number of cells or size of cells.
- The **number of cells** within the organism **increases** and the new cells produced are **genetically identical to the existing cells**.

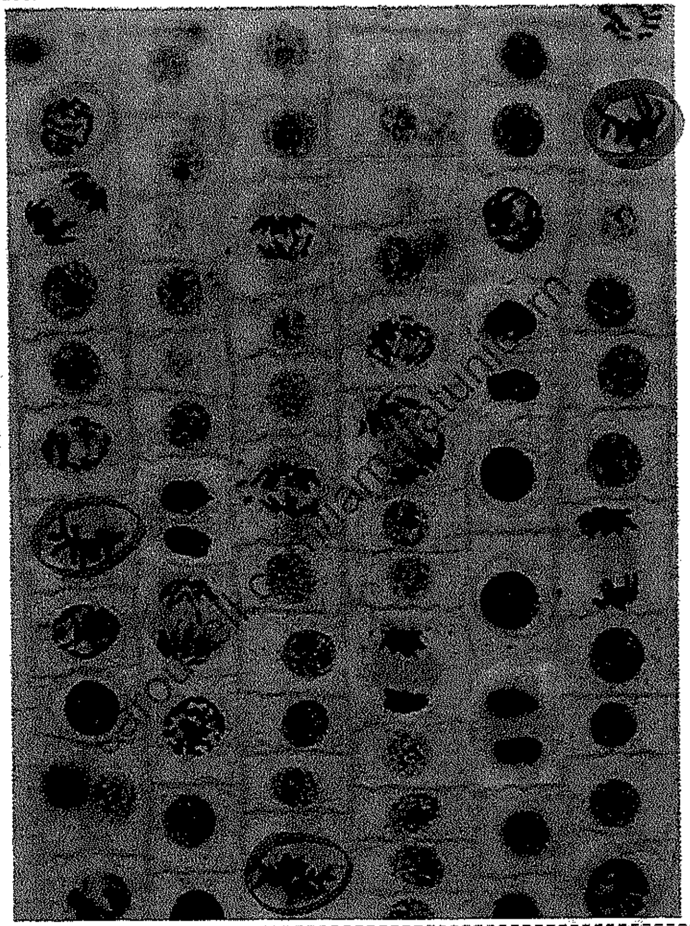
c) Regeneration and cell replacement

- Mitosis ensures that when damaged tissues are repaired, the **damaged cells** are replaced by cells that are **genetically identical** to the original cells.
- Mitosis helps in cell replacement and regeneration of missing parts, to varying degrees, in multi-cellular organisms.
- Examples of regeneration include regeneration of tails in lizards and arms in starfish.

d) Asexual reproduction

- A type of reproduction where an organism replicates itself without the production of eggs or without fertilisation. Thus asexual reproduction takes place when a single parent **produces offspring genetically identical** to itself.
- Many animal and plant species propagate by asexual means involving mitotic divisions of cells.
- Asexual reproduction is an advantage in stable (& favourable) environments where the offspring receive a set of genes from the parent who has survived and reproduced under the same conditions. With this set of genes, the offspring will be suitably adapted to the same conditions that have allowed the parent to thrive. In these ideal conditions, the population can reproduce very rapidly.
- e.g. vegetative reproduction in plants (e.g. strawberry)

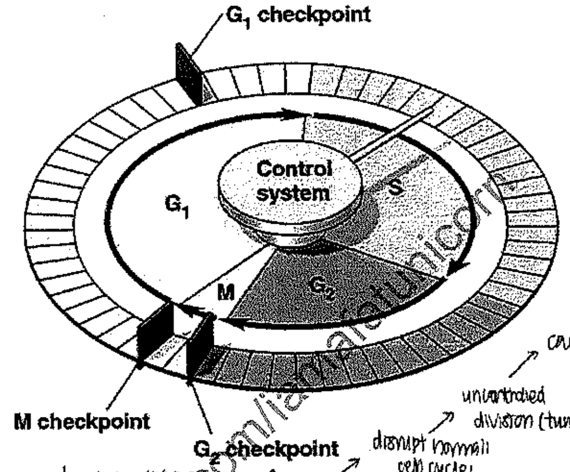
Review
Below is a micrograph of the root tip of an onion. Identify and label **1 cell undergoing interphase** and for each mitotic phase, i.e. **prophase**, **metaphase**, **anaphase** and **telophase.**



interphase
prophase
metaphase
Anaphase
Telophase

(K) UNCONTROLLED CELL DIVISION - CANCER

- The cell cycle is well regulated as it is important for normal growth and development. It is regulated at certain control points known as **checkpoints**. At these checkpoints, stop and go-ahead signals can determine whether or not the cell cycle can proceed.
- The main checkpoints are at G_1, G_2 and M phase.



due to mutations in the genes which control checkpoints of the cell cycle → disrupt normal cell cycle → uncontrolled division (tumour) → cancer

- Cancer occurs when the **dysregulation of checkpoints of cell division** occur or cells **escape the cell cycle control mechanism** that normally regulates their growth. This leads to **uncontrolled division** of cells.
- Eventually a mass of cells called a **tumour** can result. Tumour cells are **genetically identical** and **derived from a single, genetically altered cell (i.e. mutant cell).**

(L) MEIOSIS

- Most organisms produce offspring by a process of sexual reproduction. During this process, a haploid(n) gamete from one parent fuses with another haploid (n) gamete from the other parent to form a diploid (2n) **zygote** (or fertilized egg). This process results in offspring that have 2 sets of chromosomes, one set from each parent.

- Most organisms contain diploid (2n) chromosome numbers in their somatic (non-sex, non-gametic) cells. In order to ensure that the offspring have the same number of chromosomes (diploid) as the parents, the two **gametes** must have only **half the number of chromosomes (haploid)** as the parents.

- Hence, meiosis is also known as **"reduction division"** because resultant daughter cells have **half** as many chromosomes as their parent cells. The nucleus of each daughter cell will have one set of chromosomes i.e. no homologous pairs. The daughter cells would be **haploid (n)** while the parent cells would be **diploid (2n)**.

- **Meiosis** is a form of **nuclear division** in sexually reproducing organisms that produces **four haploid daughter nuclei**, each containing **half** the chromosome number of the parent cell.
  1 diploid parent cell → 4 haploid daughter cells

- Meiosis produces daughter cells that are **genetically different** from the parent and this is important because it contributes to **variation**.

- Meiosis involves **two nuclear divisions:**

Meiosis I (the first meiotic division)	**Meiosis II** (the second meiotic division)
Involves the **pairing of homologous chromosomes** and their subsequent **separation** into 2 daughter cells (which **reduces the chromosome number by half**).	Involves the **separation of the 2 sister chromatids**
prophase I metaphase I — goes straight to → anaphase I telophase I ✗ usually does not occur (to prevent wastage of resources)	prophase II metaphase II anaphase II telophase II

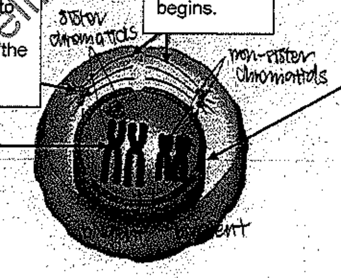
why is reduction division necessary?
▸ so that when haploid gametes fuse during fertilisation, the diploid no. of chromosomes is restored.

1) INTERPHASE

Phase	Events within the cell
G_1 (Gap phase 1)	**Intensive cellular synthesis:** a) Organelle synthesis b) RNA synthesis c) Protein synthesis d) ATP synthesis
S (Synthesis phase)	**DNA replication occurs:** a) DNA replicates → DNA content of the cell doubles (no. of chromosomes remain the same)
G_2 (Gap phase 2)	**Intensive cellular synthesis** (in preparation for mitosis): a) Organelle synthesis b) Synthesis of spindle proteins c) ATP synthesis

MEIOSIS I

2) PROPHASE I (EARLY)



Chromatin coils, shortens and thickens into a condensed chromosome.

Homologues (homologous chromosomes) **pair up** via the process of **synapsis** to form **bivalents** (this process is independent of spindle fibres)

In every pair, one homologue comes from the **father** and the other homologue from the **mother.**

Centrioles begin to migrate to opposite poles of the cell.

Spindle formation begins.

Nucleolus begins to disappear and **nuclear envelope** begins to disintegrate.

3) PROPHASE I (LATE)

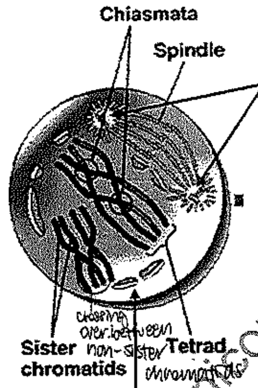
Crossing over occurs between the **non-sister chromatids of homologous chromosomes.**

Chiasmata (singular = chiasma)
→ sites where **non-sister chromatids** of **homologous chromosomes** break and rejoin with the other

Thus crossing over allows **exchange of corresponding alleles or genetic material** between **non-sister chromatids** of **homologous chromosomes.** Crossing over results in **new combination of alleles on the chromosome.** This contributes to **diversity and variation.**

During crossing over, **bivalents** are seen as **tetrads.** (each tetrad = 2 chromosomes each with 2 chromatids)

Recall: Each chromosome consists of **2 sister chromatids** joined at **centromere**

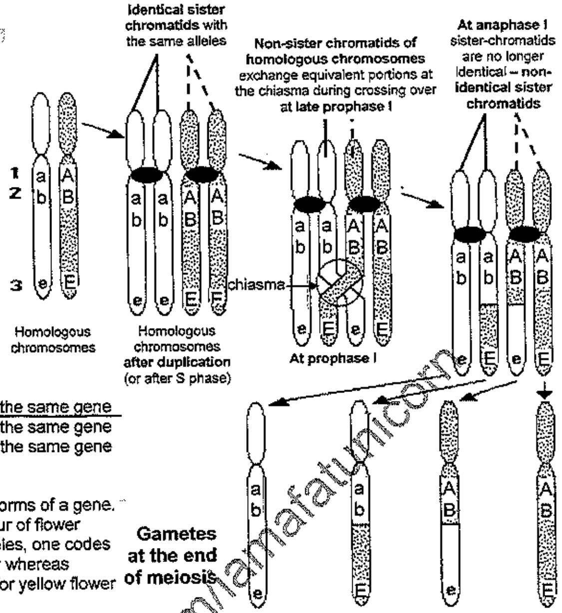


Centrioles
Continue to migrate to opposite poles of the cell.

Spindle fibres begin to **attach** to the **kinetochores** of each **homologue.**

Micrographs showing crossing over at late prophase I of meiosis I

CROSSING OVER

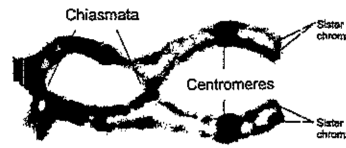


Note:
a and A are alleles of the same gene
b and B are alleles of the same gene
e and E are alleles of the same gene

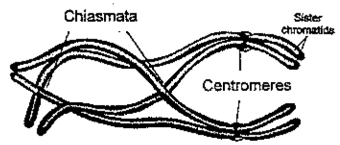
Recall:
Alleles – alternative forms of a gene.
e.g. the gene for colour of flower may have several alleles, one codes for purple flower color whereas another allele codes for yellow flower colour

- As a result of crossing over, there is an exchange of genetic material which is the **exchange of corresponding alleles** between **non-sister chromatids of homologous chromosomes.**
- **Genetic variation** in populations can arise as a result of crossing over.

Micrograph of homologous pairs at prophase I



Drawing interpretation of the micrograph



4) METAPHASE I

Spindle is completely formed

Spindle fibres (**kinetochore microtubules**) attach to the centromere of each homologue

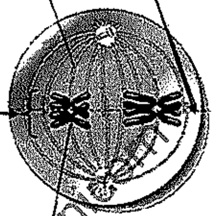
Tetrads /Homologous pair of chromosomes align along the **equator/ metaphase plate** (2 rows of chromosomes)

Homologous chromosomes move as a pair to the metaphase plate with the help of the **kinetochore microtubules**

Each homologue is **attached** to the **kinetochore microtubule** from the **pole** it faces

Independent assortment of homologous chromosome occurs at this stage. The arrangement of 1 pair of homologue at the metaphase plate is independent of other pairs. (refer to Mendel's Law of Independent Assortment on Pg 32)

arrangement of sets of homologous pairs



Question: Is there a difference between tetrads and bivalents?

No. Both terms refer to a pair of closely associated homologues, each of which is made up of 2 sister chromatids.

5) ANAPHASE I

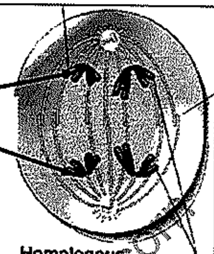
Note: **Sister chromatids** remain **attached** and move together towards the same pole (because the centromeres have not yet separated / divided, this pair of sister chromatids is considered **one chromosome**)

Homologues separate to opposite poles

Each homologue is pulled by a **shortening kinetochore microtubule** (that attaches to the centromere) towards one of the poles

*

Note: Centromeres do not divide / separate here.



Spindle fibres (**non-kinetochore microtubules**) elongate and slide in opposite direction due to the actions of motor proteins. This cause the two poles to move further apart.

6) TELOPHASE I

Chromosomes each consisting of **2 sister chromatids** reach opposite poles

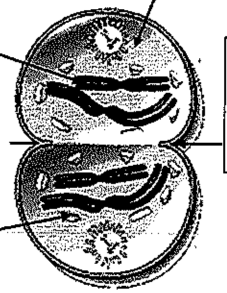
*

Each pole has a **haploid** set of chromosomes **(n)**

Chromosomes sometimes decondense into **chromatin** but no replication of DNA takes place

Nuclear envelope starts to reform around each group of chromosomes

Nucleolus reforms



Spindle fibres disintegrate

Cleavage furrow begins to form (i.e. cytokinesis begins at telophase II

⚠ At the end of meiosis I, the nuclei are **haploid**.

7) CYTOKINESIS

While cytokinesis occurs in some cells after telophase I, in many others, there is **no telophase I** and **no cytokinesis**. Such cells **enter prophase II directly from anaphase I**.

MEIOSIS II

- Meiosis II **begins with 2 haploid daughter cells** and involves the separation of sister chromatids, **forming 4 haploid cells**.

8) PROPHASE II
- Chromosomes condense
- Centrioles duplicate and move to opposite poles
- Spindle fibres begin to form
- Nuclear membrane disintegrates, nucleolus disappears

9) METAPHASE II
- Spindle is completely formed
- The centromere of each chromosome is attached to kinetochore microtubules
- Kinetochore microtubules align the chromosomes at the metaphase plate in a single file

10) ANAPHASE II
- Centromeres divide and sister chromatids separate
- **Each chromatid is now called a daughter chromosome**; these are pulled by **shortening** kinetochore microtubules to opposite poles centromeres first (i.e. led by the centromeres)
- Non-kinetochore microtubules lengthen and elongate the cell

11) TELOPHASE II
- Chromosomes reach the poles of the spindle where they decondense and become diffuse/indistinct
- Spindle fibres disintegrate
- Nuclear envelope reforms around each group of chromosomes (now existing in the form of chromatin) and nucleolus reappears in each daughter nucleus

12) CYTOKINESIS

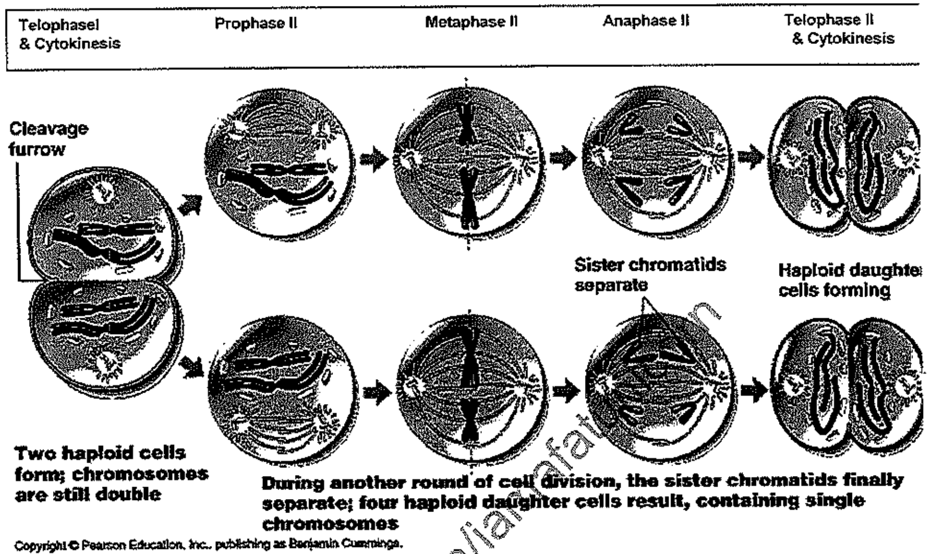
Cytokinesis follows meiosis II. During cytokinesis, the cells divide to give a total of 4 daughter cells with each daughter cell (**n**) possessing half the number of chromosomes as the parent cell (**2n**) and half the DNA amount as the parent cell *before S phase* of interphase (or a quarter the DNA amount as the parent cell *after S phase* of interphase).

Cytokinesis begins at telophase II.

⚠ Meiosis II is similar to mitosis but starts with haploid cells and sister chromatids may not be genetically identical due to crossing over.

⚠ There is no such thing as Interphase II.

Diagrammatic representation of events occurring in meiosis II



Question: Fill in the blanks.
Where n = no. of chromosomes and X = amount of DNA

	In a cell			In a cell	
	No. of chromosomes	Amount of DNA		No. of chromosomes	Amount of DNA
G₁ phase	2n	X	*Prophase II*	n	X
Prophase I	2n	2X	*Metaphase II*	n	X
Metaphase I	2n	2X	*Anaphase II*	2n	X
* *Anaphase I*	2n	2X	*Telophase II*	2n	X
Telophase I	2n	2X	*Cytokinesis*	n	½X
Cytokinesis	n	X			

(M) SIGNIFICANCE OF MEIOSIS

1. Formation of haploid gametes in sexual reproduction

- Meiosis produces **haploid gametes** (egg and sperm) for **sexual reproduction.** During **fertilisation**, the haploid nuclei of male and female gametes fuse to produce a **zygote** with a **diploid number of chromosomes**. Thus the **diploid condition** is **restored**, resulting in the restoration of ploidy level.

- If meiosis does not occur, fusion of male and female gametes by sexual reproduction will result in the **doubling of the number of chromosomes with each successive generation.** In order to ensure that the new adult organism has the same number of chromosomes as the parent, the two **gametes** must have only **half** the number of chromosomes as the parent.

- Meiosis ensures that the **chromosome number in each species is kept constant every generation.**

2. Genetic variation

- Meiosis allows for new combinations of alleles in the gametes which leads to genetic variation.

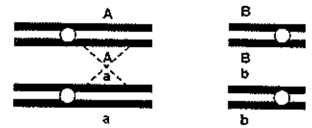
- The two important events in **meiosis** that create genetic variation are
  a) ① **crossing over** and (prophase I)
  b) ② **independent assortment of chromosomes.** (metaphase I) } during meiosis

- ③ **Random fusion of genetically different gametes** during **fertilisation** (**after meiosis**) also results in genetic variation. } after meiosis

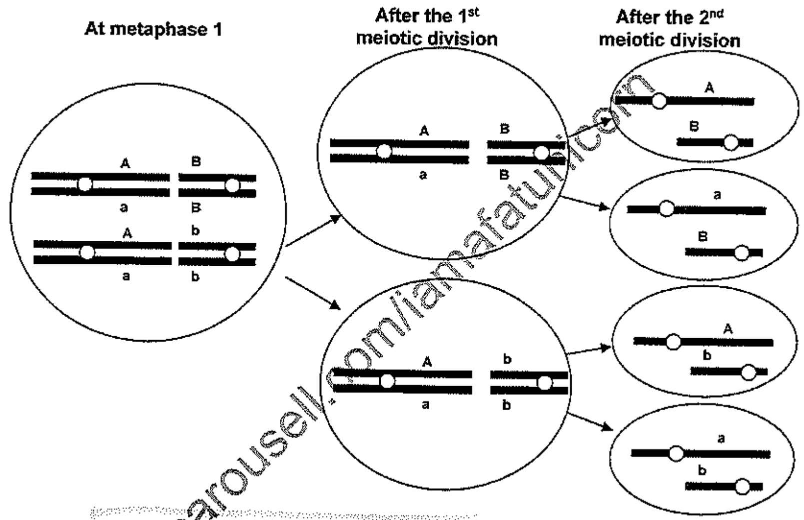
- Why is variation important?
  - Due to **genetic variation**, individuals in a **population** will have **different characteristics.**
  - When **environmental conditions change**, certain individuals in the population will be **better adapted** to the change than others. These individuals will be **selected for** as they have **favourable characteristics** that allow them to survive in the **new environment.** Individuals without the favourable characteristics will be **selected against** and will die off.
  - If there is no variation, when a catastrophic event occurs, the whole population (with same characteristics or without variation) maybe wiped out.

(a) Crossing over

- Crossing over of segments of non-sister chromatids of homologous chromosomes at prophase I of meiosis I.
- This leads to new combinations of alleles on chromosomes of the gametes. (see page 5)



Crossing over and Chiasma formation
These are 2 pairs of homologous chromosomes in a cell during meiosis. If crossing over occurs at point X, show the possible combinations of gametes that arise.

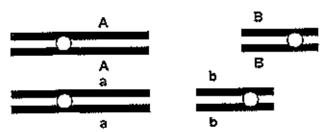


(b) Mendel's Law of Independent Assortment

- During metaphase I, **arrangement of one pair of homologues** at the **metaphase plate is independent of the arrangement of the other pairs of homologues.**
  e.g. In humans, there are 23 pairs of homologues. In every pair, there is one paternal and one maternal chromosome; where the paternal chromosome #1 aligns during metaphase I (e.g. on the left side of the metaphase plate) does not depend on where paternal chromosome #12 aligns (e.g. it can also be on the left side of the plate or on the right side of the plate)

- During anaphase I, the chromosomes of one homologous pair will **separate independently** of the other pairs to form daughter cells.

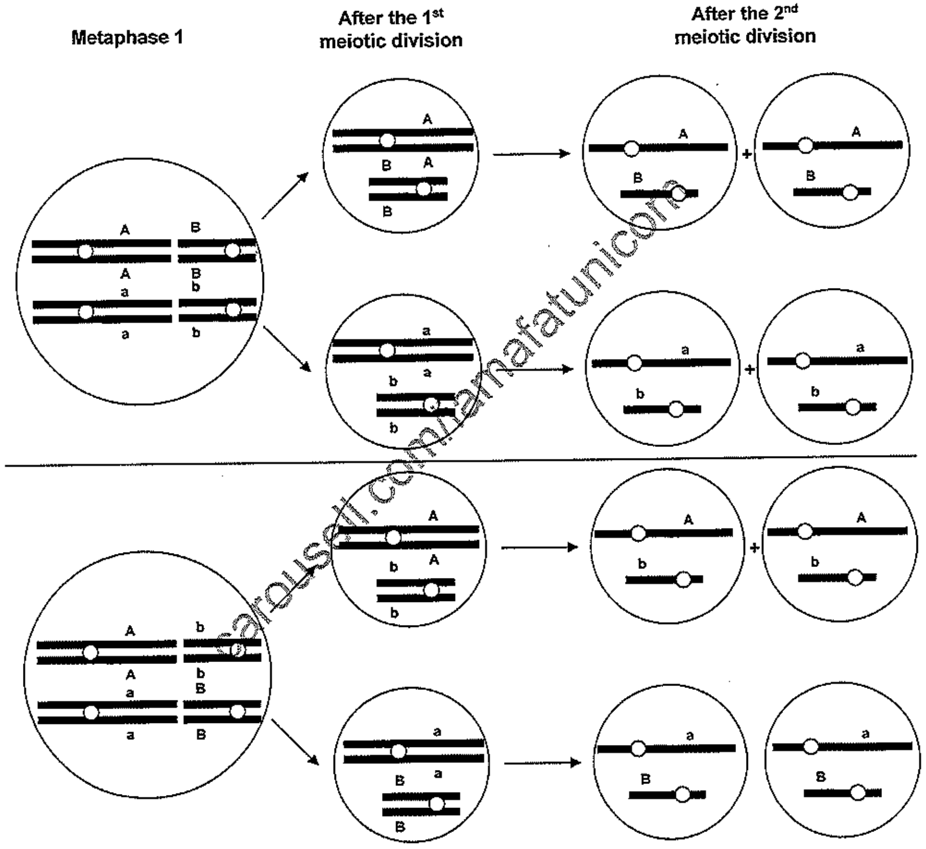
- This results in **different combinations of maternal and paternal chromosomes** in the daughter cells at the end of meiosis I and meiosis II.



Independent assortment of chromosomes

These are 2 pairs of homologous chromosomes in a cell during meiosis. Show how the chromosomes align at the metaphase plate in metaphase 1 and the possible combinations of gametes that arise at the end of meiosis II.

Assume that crossing over does not occur.



When there are two pairs of homologous chromosomes in a diploid cell, 4 *different types of gametes will result at the end of meiosis.*

- Due to independent assortment of chromosomes, there are **2^n possible combinations of gametes** where **n is number of homologous pairs.**

  Since 2 pairs of chromosomes results in **2^n** = 2^2 = 4 possible combinations of chromosomes in gametes, humans with 23 pairs of chromosomes will have 2^{23} = about 8 millions possible combinations of chromosomes in gametes.

(c) Random fusion of gametes during fertilisation

- Meiosis results in haploid gametes being formed.
- Random fusion of the gametes results in genetic variation.
- e.g. Humans = 2^{23} = 8,388,608 possible gametes
  Possible zygotes (after fertilization) = 8 million x 8 million = 64 trillion

Table of differences between mitosis and meiosis

(N) DIFFERENCES BETWEEN MITOSIS AND MEIOSIS

Feature	Mitosis	Meiosis
Location	**Somatic cells** in all parts of the body	**Precursor sex cells in reproductive organs** (that ultimately give rise to **gametes)**
Occurs in	**Haploid** or **diploid** cells	Only **diploid** cells
No of nuclear divisions	**One**	**Two**
Prophase	**No synapsis**/ Homologues do not pair up; **No chiasma** formation; **No crossing over** of corresponding segments of non-sister chromatids;	**PROPHASE I** **Synapsis occurs** / Homologues pair up to form **bivalents (tetrads)**; **Chiasma** formation; **Crossing over of corresponding segments of non-sister chromatids** (results in **non-identical sister chromatids** with **new combinations of alleles)**; **PROPHASE II** *No difference from prophase of mitosis*
Metaphase	Chromosomes **align individually** on equator/metaphase plate (i.e. form a single row); Each centromere attaches to **spindle fibres from both poles**;	**METAPHASE I** **Homologues align in pairs** along equator/metaphase plate (i.e. form 2 rows); Centromere of each chromosome attaches to **spindle fibre from only one pole** (Each member of a homologous pair attaches to spindle fibres from different poles); **Independent assortment of homologues** occurs (results in gametes with **new combinations of paternal and maternal chromosomes)** **METAPHASE II** *similar to metaphase of mitosis, except that:* chromosomes, consisting of **non-identical sister chromatids**, align in a random arrangement along the equator/metaphase plate

Anaphase	**Division of centromere;** **Separation of identical sister chromatids** to opposite poles; Once centromeres <u>separate</u> chromatids are called chromosomes;	<u>**ANAPHASE I**</u> **No separation of centromere;** **Separation of homologues** (i.e. pair of sister chromatids move to same pole); <u>**ANAPHASE II**</u> *similar to anaphase of mitosis, except that:* **non-identical sister chromatids** <u>**separate**</u> to opposite poles. They are now called chromosomes.
Telophase	**2 daughter nuclei** which are **genetically identical** & have the **same chromosome number as parental cells** (hence can be **2n or** n)	<u>**TELOPHASE I**</u> **2 daughter nuclei** which are **genetically different** & each has **half the chromosome number as parental cells (n)** (i.e. cells are **haploid** at the end of meiosis I) <u>**TELOPHASE II**</u> **4** daughter nuclei which are **genetically different** & each has **half the chromosome number as parental cells (n)**
Result of nuclear division	**2 genetically identical** daughter cells; No variation occurs (in the absence of mutation); Daughter cells have the **same number of chromosomes** as parent cells, hence mitosis is called **replicative division**;	**4 genetically different** daughter cells; **Genetic variation** has occurred (even in the absence of mutation); Daughter cells have **half the chromosome number** as parent cells, hence meiosis is called **reductive division**;

⚠ Non-identical sister chromatids are **not** the same as non-sister chromatids. The former refers to sister chromatids that have different combinations of alleles (due to crossing over) while the latter refers to chromatids from different homologues.

(O) MUTATIONS – FOCUS ON CHROMOSOME MUTATION

1. Gene mutations

- Definition: A gene mutation arises as a result of a <u>change in the sequence of nucleotide bases in the DNA of a gene</u>.
- For the types and relevant examples of gene mutations, please refer to "**DNA and Genomics**" lecture notes
- Sickle-cell anaemia and cystic fibrosis would be two examples that you will need to be very familiar with for your syllabus.

2. Chromosomal aberrations

There are two major forms of chromosomal aberrations:

(I) Variation in chromosomal structure and

(II) Variation in chromosomal number

(I) VARIATION IN CHROMOSOMAL STRUCTURE

Alterations in chromosomal structure (usually involves <u>several gene loci</u>) can be brought about by the following mechanisms:

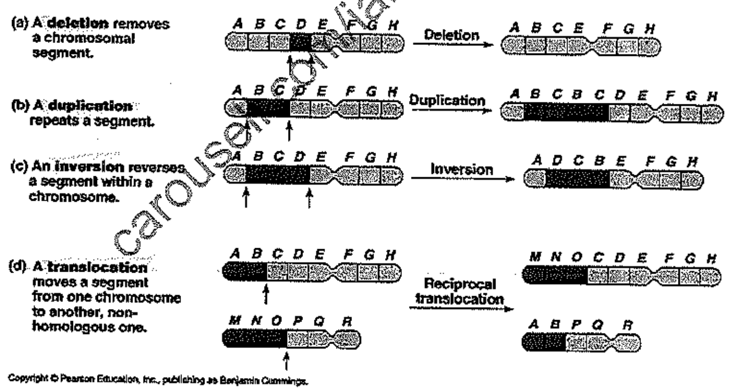


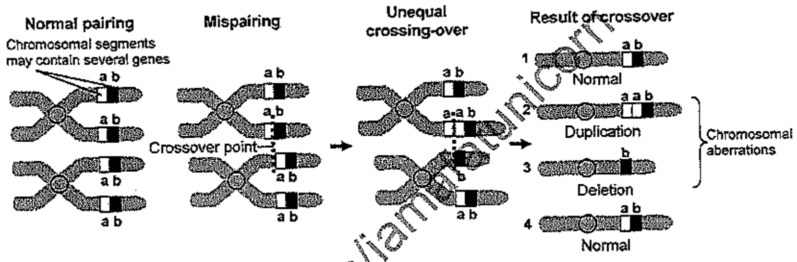
Figure 1: Mechanisms for changes in chromosome structure

- <u>Deletions</u> and <u>duplications</u> are especially likely to occur **during crossing over.**

  Non-sister chromatids of homologous chromosomes may break and rejoin at incorrect places such that one chromatid may give up more genes than it receives.

  The products of such an unequal crossover are one chromosome with a deletion mutation and one with a duplication mutation. (Figure 2)

- Variations in chromosome structures usually cause serious problems. Chromosome deletions frequently result in zygotic loss, stillbirths or infant deaths. Some survive a little longer.
- The phenotypic abnormalities that result is usually due to the reduced or additional genes reflected in **chromosomal deletions** and **duplications** respectively.
- You may then wonder how **chromosomal inversion** and **reciprocal translocations** may result in disease since the amount of genetic material remains the same.
- However, these chromosomal aberrations/mutations may still alter the phenotype because the expression of a gene can be influenced by its new location among neighbouring genes. e.g. juxtaposition (placement of things side by side) of genes next to regulatory elements such as enhancers, could up-regulate gene expression.



Crossing over typically occurs between homologous regions of chromosomes, so there is no net gain or loss of DNA in either chromosome. However, it can occur by error between non-homologous regions of homologous chromosomes. When that happens, there is a loss of a section of DNA in one chromosome, and a gain in the other.

In Figure 2, the end result is that chromosome 2 has gained an extra copy of chromosome section a and chromosome 3 has suffered a deletion, losing chromosome section a.

Figure 2 : A proposed mechanism where duplications and deletions of chromosomal segments may result.

Figure 3: Reciprocal translocation between Chr 9 and Chr 22 leads to the formation of the Philadelphia chromosome

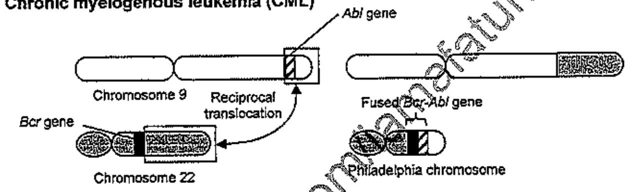
(Additional information)
Example 1: (Deletion)

Cri-du-chat ("Cat's cry" in French) syndrome

- Due to deletion in the short arm of chromosome 5.
- A child born with this deletion is physically and mentally retarded, has a small head, broad face and saddle nose, widely spaced eyes, unusual facial features and a cry that sounds like the mewing of a distressed cat. Such individuals usually die in infancy or early childhood.
- The signs and symptoms of *cri-du-chat* syndrome are probably related to the loss of multiple genes in this region. Most cases of *cri-du-chat* syndrome are not inherited. They result from a chromosomal deletion that occurs as a random event during the formation of gametes (eggs or sperm) or in early foetal development. Such people typically have no history of the disorder in their family.

Example 2: (translocation)

Chronic myelogenous leukemia (CML)



- When a genetic aberration occurs in somatic cells, cancer may result.
- In CML, most of the chromosome 22 has been translocated onto the long arm of chromosome 9. In addition, the small distal portion of chromosome 9 is translocated to chromosome 22. The resultant chromosome 22 is called the "Philadelphia chromosome." 95% of people with CML have this chromosome.
- The translocation brings two genes (*Abl* and *Bcr* genes) next to each other and genes are transcribed and translated as one protein. This protein causes increased cell proliferation and reduced apoptosis → cancer. It is unclear why this fusion product causes cancer.
- CML affects the stem cells that develop into white blood cells. These cells may not mature normally but proliferate rapidly.

Example 3: (translocation)

Burkitt's Lymphoma

- Burkitt's lymphoma is a form cancer involving B lymphocytes.
- This lymphoma results from chromosomal translocations that involve the *Myc* gene.
- The *Myc* gene is a proto-oncogene that is found on chromosome 8. The Myc protein acts as a signal for cell proliferation.
- In Burkitt's Lymphoma, the most common reciprocal chromosomal translocation results from the translocation of the *Myc* gene from chromosome 8 to chromosome 14.

- This brings the *Myc* gene under the influence of powerful regulatory sequences (***enhancers****) that normally control the production of large amounts of antibodies (IgH) from each cell. The gene for IgH is on chromosome 14.
- This results in excess transcription of the *Myc* gene instead of the usual *IgH* gene. Hence, mutant B lymphocytes proliferate and eventually form a tumour.

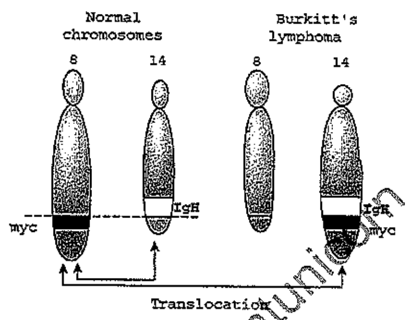


Figure 4 : Reciprocal translocation between Chr 8 and Chr 14 resulting in Burkitt's Lymphoma

After translocation, *Myc* gene originally from **Chr 8** is under the influence of the enhancer of *IgH* gene on **Chr 14**.

Example 4: (duplication)

Charcot-Marie-Tooth syndrome (CMT)

- In one form of CMT, a chromosomal duplication on chromosome 17 results in high gene dosage (3 instead of a normal 2) of a myelin sheath protein resulting in abnormal structure and function of the myelin sheath (an insulating sheath around nerve cells).
- This type of CMT is inherited in an autosomal dominant condition.
- Symptoms: Weakness of lower foot, loss of balance, poor motor skills and muscle atrophy. Not a fatal disease and sufferers have normal life expectancy.

(II) VARIATION IN CHROMOSOMAL NUMBER

ANEUPLOIDY

- ***Aneuploidy**** is a condition where the cell does not have a chromosome number that is a multiple of the haploid number. Chromosomes are present in either extra or fewer copies than the wild type.
  1. If chromosome is present in triplicate, the aneuploid cell is said to be trisomic e.g. 2n+1
  2. If the cell is missing a chromosome, it is said to be monosomic e.g. 2n-1
- Aneuploidy is a result of a ***non-disjunction**** event where:
  1. Homologous chromosomes do not move properly to opposite poles during ***meiosis I**** – failure of homologous chromosome to separate.

     OR
  2. When sister chromatids fail to separate properly to opposite poles during ***meiosis II****.
- So one gamete receives two of the same type of chromosome and another gamete receives no copy (see Figure 5).
- If either of the aberrant gametes unites with a normal gamete at fertilisation, offspring will have abnormal number of a particular chromosome = ***aneuploidy****
- Mitosis will subsequently transmit the anomaly to all embryonic cells.
- ***Non-disjunction**** can also occur during ***mitosis****. If such an error occurs early in embryonic development, then the aneuploid condition is passed on to a large number of cells where the severity of the effect is more pronounced.
- Aneuploidy is a genetic disorder.

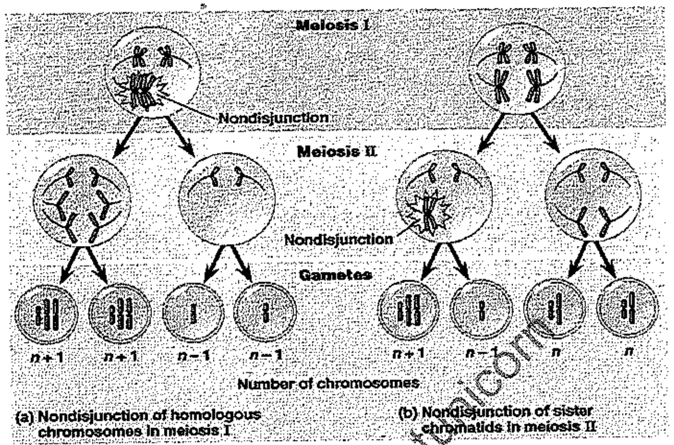


Figure 5 : Non-disjunction during meiosis I or meiosis II leading to aneuploidy

Example 1: Down syndrome (Trisomy 21)

- Down syndrome is result of an extra chromosome 21 (a total of 3 copies), so each body cell has a total of 47 chromosomes.
- Most cases result from non-disjunction during meiosis I.
- Down syndrome includes characteristic facial features, short stature, heart defects, susceptibility to respiratory infection and mental retardation. Most individuals are sexually underdeveloped and sterile.

(Additional information)

Example 2: Klinefelter syndrome (XXY)

- Males with an extra X chromosome suffer from Klinefelter syndrome. These individuals have male sex organs, but the testes are abnormally small and the man is sterile. Though extra X chromosome is inactivated, some breast enlargement and other female body characteristics are common. Affected individual is usually of normal intelligence

Example 3: Turner syndrome (monosomy X)

- Monosomy X, is the only known viable monosomy in humans. These XO individuals are phenotypically female, but are sterile and their sex organs do not mature. When provided with estrogen replacement therapy, girls with Turner syndrome do develop secondary sex characteristics.

Did you know?

- Although females have 2 X chromosomes, one X chromosome in each cell becomes almost completely inactivated during embryonic development. The choice of X chromosome to be inactivated is a random process. As a result, the cells of females and males have the same effective dose (one copy) of genes with loci on the X chromosome. This is called dosage compensation.
- Non-disjunction of sex chromosomes produces a variety of aneuploid conditions. Most of these aneuploid conditions upset genetic balance less than those involving autosomes. This may be because Y chromosome carries fewer genes and extra X chromosomes become inactivated in somatic cells.

(P) GLOSSARY OF TERMS

Definitions

Asters – the radial arrays of shorter microtubules that extend from centrosome.

Centrioles – a pair of cylindrical organelles located at the poles of the cell in animal cells.

Centromere – the constricted region of the chromosome which join 2 sister chromatids and where the spindle fibres attach during cell division.

Centrosome – A structure present in the cytoplasm of animal cells that function as the microtubule-organising centre and is important during cell division. A centrosome has two centrioles.

Chromatin – the complex of proteins and nucleic acids (DNA) that makes up chromosomes. When the cell is not dividing, chromatin exists in its dispersed form, as a mass of very long, thin fibres that are not visible with a light microscope.

Chromosomes – a structure carrying genetic material, found in the nucleus. Each chromosome consists of DNA and associated proteins. Chromosomes are most visible during mitosis and meiosis and are the condensed form of chromatin.

Kinetochore – a structure of proteins attached to the centromere that links each sister chromatid to the spindle.

Sister chromatids - either of two copies of a duplicated chromosome attached to each other by proteins at the centromere and eventually separated during anaphase of mitosis and meiosis II. The DNA molecules of sister chromatids are products of DNA replication using the same DNA molecule as a template. They are thus identical in terms of nucleotide sequence and combination of alleles.

Spindle – an organised system of microtubules that attaches to the centromere regions of chromosomes that draws them to opposite poles during cell division.

(Q) LINKS

The topic of meiosis is relevant to the following topics in the 'A' level Biology syllabus. The links also become clearer when you have gone through the other topics.

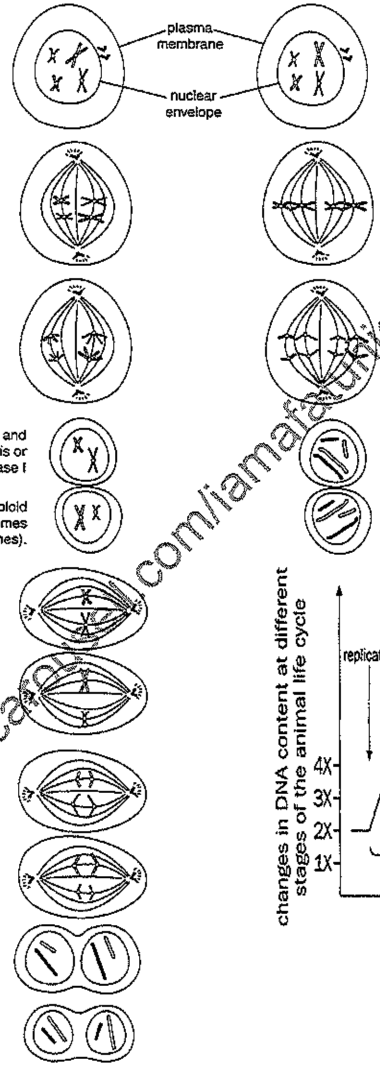
Topic No	Topic	Comments
1	Genetic Basis of Variation	Independent assortment and segregation of chromosomes which occurs during meiosis contributes towards genetic variation. More about how variation arises will be covered under Genetic Basis of Variation. The term homologous chromosome was introduced in the topic of mitosis and again in meiosis. It is important to be able to distinguish between the terms homologous, homozygous, heterozygous and hemizygous which will be covered under Genetic Basis of Variation.
2	Diversity and Evolution	How genetic variation contributes towards evolution will be covered under Diversity and Evolution.

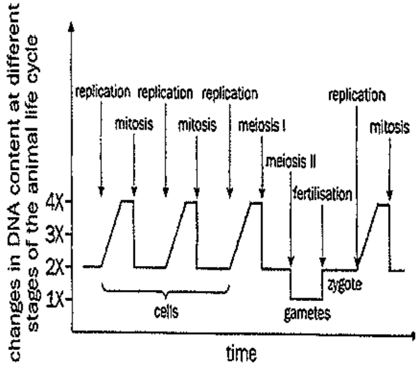
Keywords include

Names of all types of nuclear division and their stages:
- Mitosis – Prophase, Metaphase, Anaphase, Telophase
- Meiosis – Meiosis I: Prophase I, Metaphase I, Anaphase I, Telophase I
  – Meiosis II: Prophase II, Metaphase II, Anaphase II, Telophase II

Interphase: S phase
Cytokinesis
Chromatin
Chromosomes
Centromere divides
Kinetochore proteins
Kinetochore microtubules and non-kinetochore microtubules
Spindle fibers
Sister chromatids
Non-sister chromatids of homologous chromosomes
Non-identical sister chromatids
Homologous chromosomes / homologues
Crossing over
Independent assortment
Metaphase plate / equator of the cell
Opposite poles of cell
Semi-conservative replication
Genetically identical
Genes
Alleles
Synapsis
Chiasma
Bivalent
Tetrads
Centrioles
Microtubule Organising Center
Asters
Chromosomal abberations: Insertion / Deletion / Inversion / Translocation
Non-disjunction
Aneuploidy

Diagrammatic differences between mitosis and meiosis





CORE IDEA :
(2) GENETICS AND INHERITANCE

DNA & GENOMICS I

Content

- DNA Structure & Function

Learning Outcomes

Candidates should be able to:
(a) describe the structure and roles of DNA and RNA (tRNA, rRNA and mRNA). (knowledge of mitochondrial DNA is not required.)

(b) describe the process of DNA replication and how the end replication problem arises.

References

Biology a Global Approach, 10th Edition, by Campbell, N.A. and Reece J.B. (Red spot)
Biology, 8th edition Solomon, Berg & Martin (Red spot)
Biology, 8th Edition, Raven, Johnson, Losos, Mason and Singer, Mc Graw-Hill
Biological Science Vol 1 and 2, 3rd edition, Green N.P.O, Stout G.W and Taylor D.J, (2004) Cambridge
Biology, 9th Edition, Mader, Sylvia S (2007), Mc Graw-Hill
Molecular Biology of the Cell, 5th Edition, Alberts (2008)

Table of Contents

This handout is the effort of several Biology teachers at RI (Year 5-6). It has and will continue to be updated.

(1) STRUCTURE OF DNA AND RNA

(A) GENERAL INTRODUCTION TO NUCLEIC ACIDS

- Cells carry information that specify their structure, dictate their functions, and regulate their activities and these instructions can be passed on faithfully to daughter cells.
- This information is carried in the hereditary material, **DNA** (Deoxyribonucleic acid), a nucleic acid.
- In some cases of viruses, the hereditary material can be carried in another type of nucleic acid, **RNA** (Ribonucleic acid). → for some viruses
- The monomers of nucleic acids are called **nucleotides**.
- Before it was proven that DNA was the genetic material, many scientists thought that protein was the hereditary material because:
  - ➢ Proteins played a central role in all biochemical processes.
  - ➢ There were 20 different kinds of amino acids from which limitless combinations of proteins would be possible → a reflection of the complexity of life.
- The first concrete evidence that DNA is the genetic material was shown by the **Hershey and Chase experiment.** (Fig.1)
  - ➢ This experiment makes use of ***radioisotopes*** (radioactive isotopes) and ***phages*** (viruses that infect bacteria) because radioisotopes emits radioactivity that can be detected while phages are used to deliver the hereditary material.

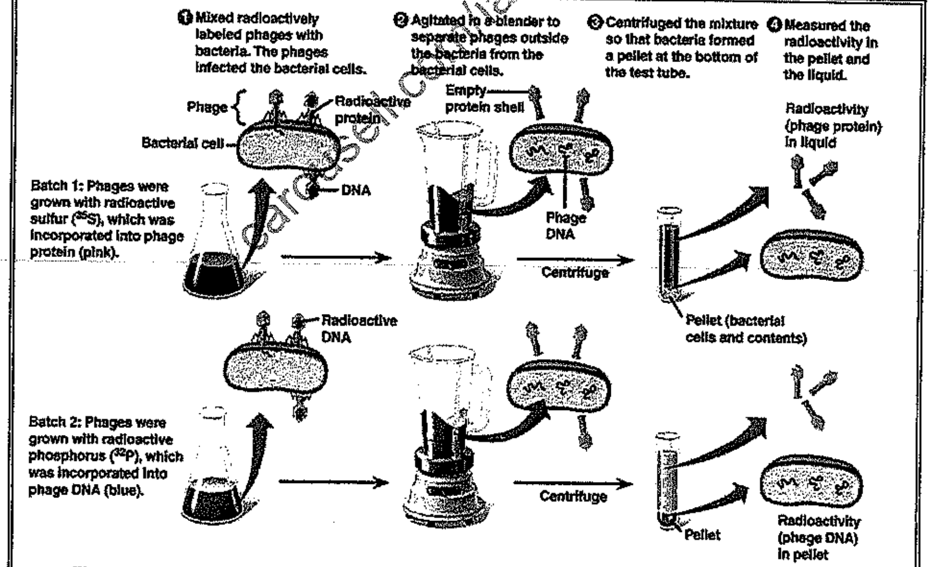
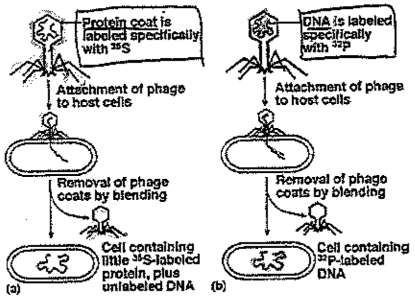


Figure 1. Experiment by Hershey & Chase



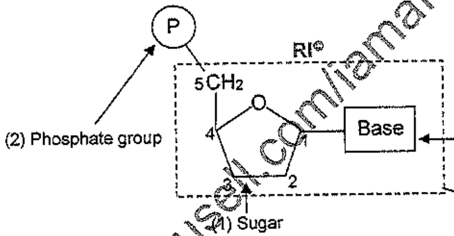
Conclusion:
Since only **radioactively labelled DNA** and not protein ended up **in the bacteria cell**, it was concluded that the DNA is the hereditary material.

Figure 2. Results of the experiment by Hershey & Chase

(B) STRUCTURE OF NUCLEOTIDES

Components of a nucleotide

- Nucleic acids are macromolecules that exist as polymers of nucleotides called polynucleotides.
- Each nucleotide is composed of 3 parts: **5-carbon sugar** (pentose sugar), **nitrogenous base** and **phosphate group(s)**. (Fig. 3)
- The number of phosphates can vary from 1 to 3. ATP (Adenosine triphosphate) is an example of a RNA nucleotide that has 3 phosphates.
- A **nucleoside** is composed of just a 5-carbon sugar and a nitrogenous base without the phosphate group.



Please be able to represent the structure of a nucleotide diagrammatically

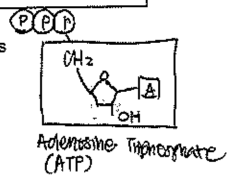


Figure 3. A nucleotide, a monomer of DNA & RNA

(1) Sugar

- The sugar component in a nucleotide has 5 carbon atoms. Therefore it is a **pentose**.
- The 2 types of nucleic acids (DNA and RNA) differ in the **type of pentose** they contain.
- The sugar **ribose** is present in RNA while the sugar **deoxyribose** is present in DNA. (Fig. 4)
- Deoxyribose differs from ribose in that the **hydroxyl group (-OH) at carbon 2, has O atom removed (hence 'deoxy')**.

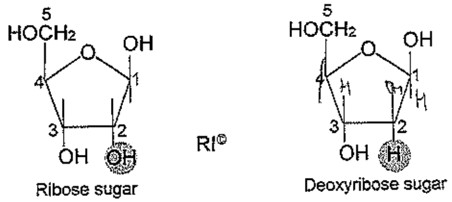
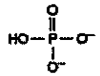


Figure 4. Pentose sugar in nucleic acids

(2) Phosphate

- Phosphate group is attached to **carbon 5** of the pentose sugar, and this gives nucleic acids their **negative charge** and **acidic** character.



Phosphate

(3) Nitrogenous base

- Each nucleotide contains one of five different nitrogenous bases on **carbon 1** of the pentose sugar.
- The nitrogenous bases can be categorised into **purines** and **pyrimidines**. (Fig.5)
- Purines have 2 rings whereas pyrimidines have 1 ring in their structures.
- In DNA → the purines : adenine **(A)** and guanine **(G)**
  → the pyrimidines : cytosine **(C)** and thymine **(T)**
- In RNA → the purines : adenine **(A)** and guanine **(G)**
  → the pyrimidines : cytosine **(C)** and uracil **(U)** (instead of thymine!!!!)

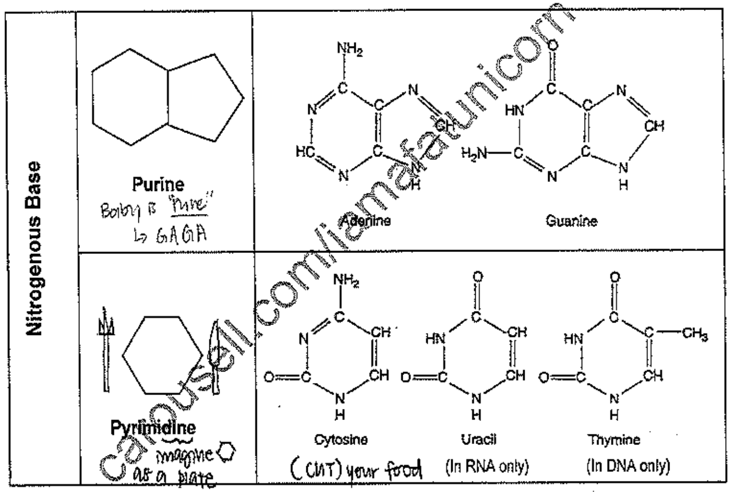


Figure 5. Nitrogenous bases

NOTE: Although the bases are commonly represented by their initial letters A, T, C, G and U, you need to spell out the name in full. i.e. Adenine, Thymine, Cytosine, Guanine and Uracil.

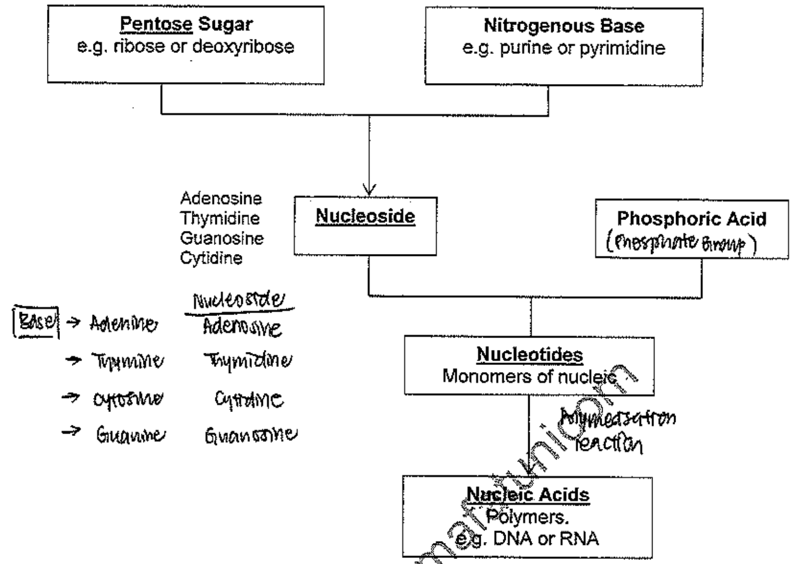


Figure 6. Nucleic Acids and their Components

(C) FORMATION OF NUCLEIC ACIDS FROM NUCLEOTIDES

- Nucleic acids are formed by combining **free nucleotides.**

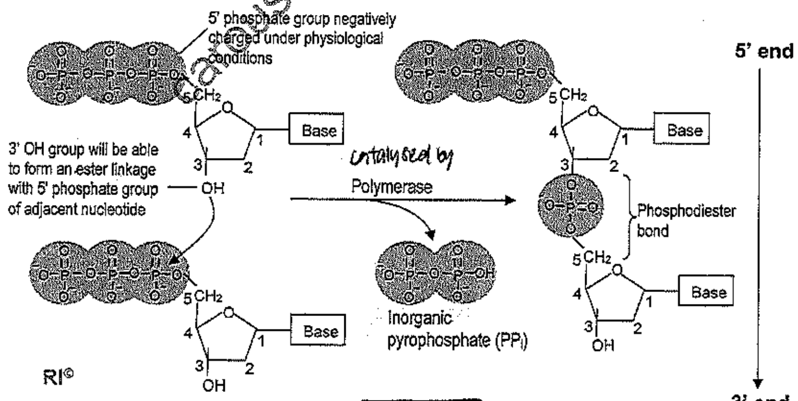


Figure 7. Formation of a dinucleotide by polymerisation

- This polymerisation reaction (Fig. 7) is catalysed by a **polymerase** e.g. DNA polymerase or RNA polymerase.

- The **covalent** bond linking two adjacent nucleotides is called a **phosphodiester bond.** It consists of phosphate group linked to two pentoses via two covalent ester linkages.
- The addition of further nucleotides produces a long **polynucleotide** chain whose backbone consists of alternating sugar and phosphate groups with the bases projecting sideways from the sugars.
  - ➢ The sugar and phosphate groups are identical all the way along a polynucleotide chain forming a **sugar-phosphate backbone.**
- In RNA, the sugars are all ribose. In DNA, the sugars are all deoxyribose.
- The **difference** between each polynucleotide lies in the **sequence in which the various nitrogenous bases** occur along their length.
- The **5'** (read as "five prime") end of a polynucleotide chain **ends with a free phosphate group** attached to **carbon 5 of a sugar residue.**

  The **3'** (read as "three prime") end of a polynucleotide chain **ends with the free hydroxyl (-OH) group** on **carbon 3 of a sugar residue.**

  ➢ Thus the nucleic acid strands are said to have **directionality.**

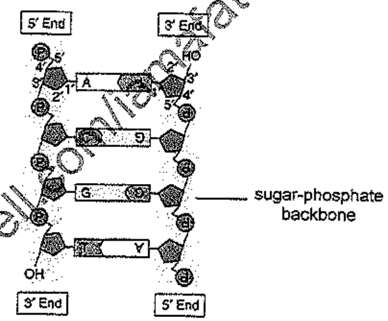


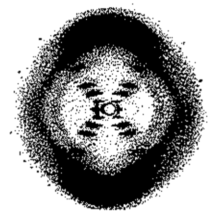
Figure 8. DNA structure with directionality.

(D) STRUCTURE OF DNA *(Elucidated by Watson and Crick, published in 1953)*

- The **basic unit** of a DNA molecule is the **deoxyribonucleotide**.
- The nitrogenous bases could be **adenine, thymine, guanine** or **cytosine.**

(1) In 1951-53 Rosalind Franklin took X-ray diffraction images of crystallised DNA which revealed that DNA was long and thin. There was some structural regularity every 3.4 nm and the molecule appeared helical.


(a) Rosalind Franklin


(b) Franklin's X-ray diffraction photograph of DNA

(2) In 1944-52 Erwin Chargaff studied the purine and pyrimidine bases present in DNA isolated from a variety of animals. He found out that in all organisms:

- Number of A = number of T
  Number of G = number of C
- Number of purines = number of pyrimidines

Chargaff's rules:
⇒ **ratio of A:T = 1:1; ratio of G:C = 1:1**
⇒ ratio of (A+G) : (T+C) = 1:1 (i.e. ratio of **purines : pyrimidines = 1:1**)

	Percentage of DNA Bases				Ratios	
Source of DNA	**A**	**T**	**G**	**C**	**A/T**	**G/C**
E. coli	26.1	23.9	24.9	25.1	1.09	0.99
Yeast	31.3	32.9	18.7	17.1	0.95	1.09
Sea urchin sperm	32.5	31.8	17.6	18.2	1.02	0.96
Herring sperm	27.8	27.5	22.2	22.6	1.01	0.98
Human liver	30.3	30.3	19.5	19.9	1.00	0.98
Corn (*Zea mays*)	25.6	25.3	24.5	24.6	1.01	1.00

(3) Based on these evidences, Watson and Crick, in 1953, worked out the molecular structure of DNA, using just cardboard representations of the nucleotide bases and a vivid imagination.

- To solve the riddle as to why the number of purines = number of pyrimidines, Watson started building a scale model. They figured that for two irregular sequences of bases to be regularly packed in the centre of the helix, a purine had to always pair with a pyrimidine.

(1) to (5) show evidences that support the proposed structure of DNA

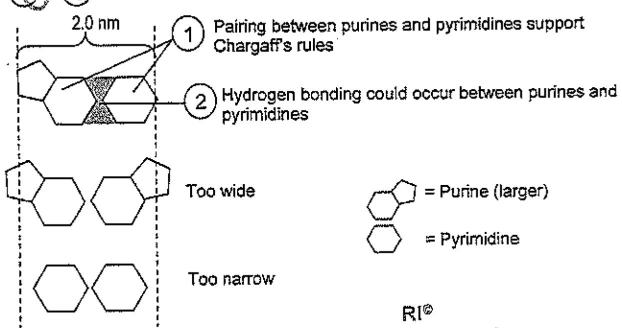


Figure 9. Showing constant width of base pairs

(3) • In fact, the **width between the 2 sugar phosphate backbones is constant (2.0 nm)** and equal to the combined width of **a purine and a pyrimidine.** (Fig. 9 & 12)

(4) • Stacking one base pair on top of another, they realised that **one complete turn** of the double helix has **10 base pairs,** and spans a distance of **3.4 nm.** (Fig. 11)

(5) • So the image of DNA now becomes clear. It consists of **two polynucleotide chains** (strands) twisted around each other to form a **double helix.**

Conclusion

- The two strands (or chains) run in **opposite directions** i.e. they are **antiparallel.** One strand runs in the 5' to 3' direction and the complementary strand runs in the 3' to 5' direction. (Fig. 12)
- The two strands are held together by **weak hydrogen bonds** that form between the **nitrogenous bases of opposite strands.**
- **Two antiparallel strands make up one DNA molecule.**
  - ➢ (purine) **A=T** (pyrimidine) base pairs forming **2 hydrogen bonds.**
  - ➢ (purine) **G≡C** (pyrimidine) base pairs forming **3 hydrogen bonds.** more proportion of G≡C, more stable
- This is known as **complementary base pairing** and Chargaff's rule provides strong supporting evidence for this.

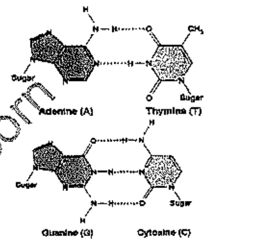


Figure 10. Complementary base-pairs

- The sugar-phosphate backbones of both strands lie on the outside of the molecule, with the nitrogen-containing bases on the interior.
- This double helix structure of DNA and complementary base pairing immediately suggests a means of replicating DNA. (To be covered in under DNA replication)
- The **helix is right-handed,** curving up to the right. If you make a 'thumbs up' sign with your right hand, the direction where your fingers point to is how the DNA strand spirals up. (Fig. 11)

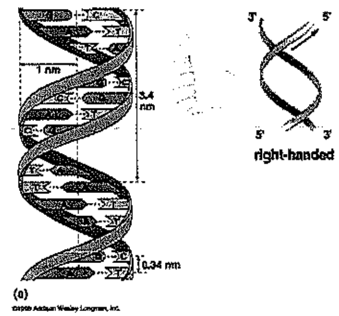


Figure 11. DNA double helix, a right handed helix

(E) STRUCTURE OF RNA

- The **basic unit** of a RNA molecule is the **ribonucleotide**.
- The nitrogenous bases could be adenine, **uracil**, guanine or cytosine. (A, U, G and C respectively)
- RNA are **single stranded** molecules (with the exception of the genomes of some RNA viruses).

Q. Identify whether the following is DNA or RNA in eukaryotes and state why?

% A	% C	% T	% G	% U	DNA or RNA	Why?
40	10	40	10	0	ds DNA	Ratio of A:T & G:C is 1:1 complementary base pairing
20	30	0	30	20	ds RNA	Uracil present, hence RNA; ratio of A:T & G:C is 1:1 complementary base pairing
20	40	0	40	0	ss RNA	Ratio A:T not 1:1, hence not DNA

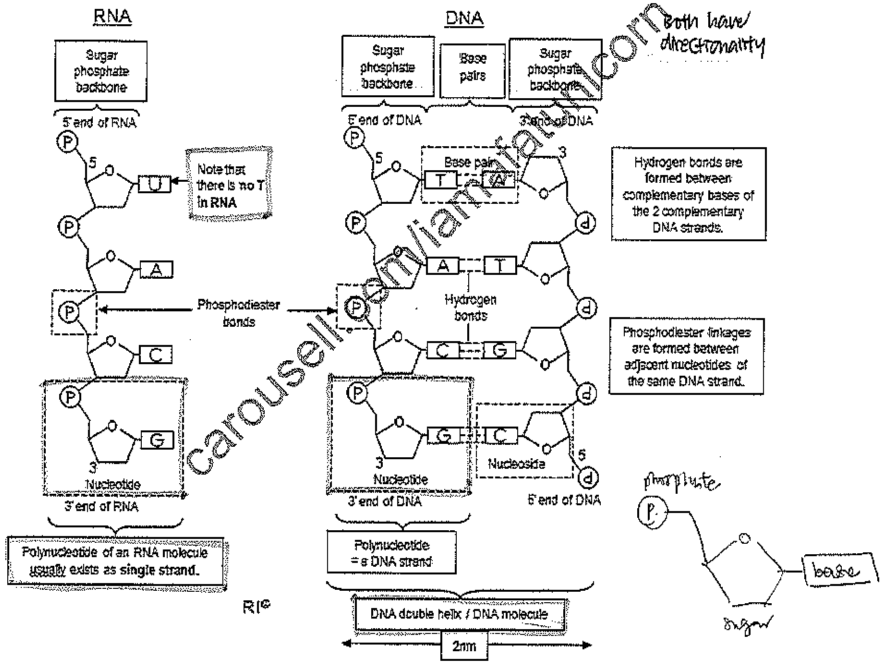
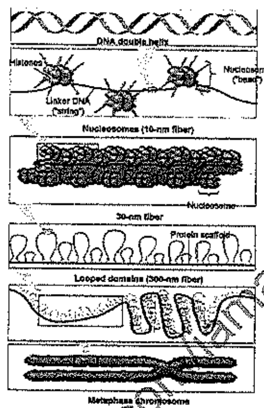
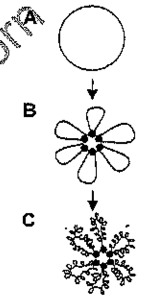


Figure 12. Structure of RNA and DNA

(F) COMPARING THE STRUCTURE OF PROKARYTIC AND EUKARYOTIC GENOMES

(Note: Genome - The complete DNA content of an organism, typically expressed in number of base pairs.)
(This topic will be covered in more detail under Organisation and Control of Prokaryotic and Eukaryotic Genomes Pt 1)

Feature	Structure of Eukaryotic Genome	Structure of Prokaryotic Genome
Size	Larger	Smaller
Appearance	Multiple, linear molecules	Generally a single, circular molecule
Molecule	Double helix DNA	Double helix DNA
Association with proteins	Yes – large amounts of it e.g. histones, scaffold proteins	Yes – relatively less
Level of DNA packing/coiling	High:  • The DNA double helix is associated with proteins called **histones**. DNA molecules are **negatively charged** whereas the histone proteins are **positively charged**. Thus the DNA molecule is held around the histones by **electrostatic interactions**. • Most of the DNA is wound around octamers of 8 histone proteins to form **nucleosomes, the 10nm fibre**. The remainder of the DNA, called the **linker**, joins adjacent nucleosomes. The 10nm fibre coils around itself to form a **30nm chromatin fiber** (or **solenoid**). • The 30nm fibre forms loops called **looped domains (a 300nm fibre)** which further coil & fold to produce the characteristic metaphase **chromosome**.	Relatively low: DNA double helix ↓ some looping, possibly around some proteins  • The circular double-stranded DNA has a diameter of about 430 μm when unfolded. • The **DNA is folded** into chromosomal domains by **protein-DNA associations**. Six domains are shown, but the actual number is about 50. • **Supercoiling** and other interactions cause **further compaction**, such that it fills an area of about 1 μm.
Location	Nucleus	Nucleoid region – not membrane-bound
Presence of extrachromosomal DNA	None (mitochondria and chloroplast have their own DNA)	Yes – plasmids (much smaller rings of DNA)

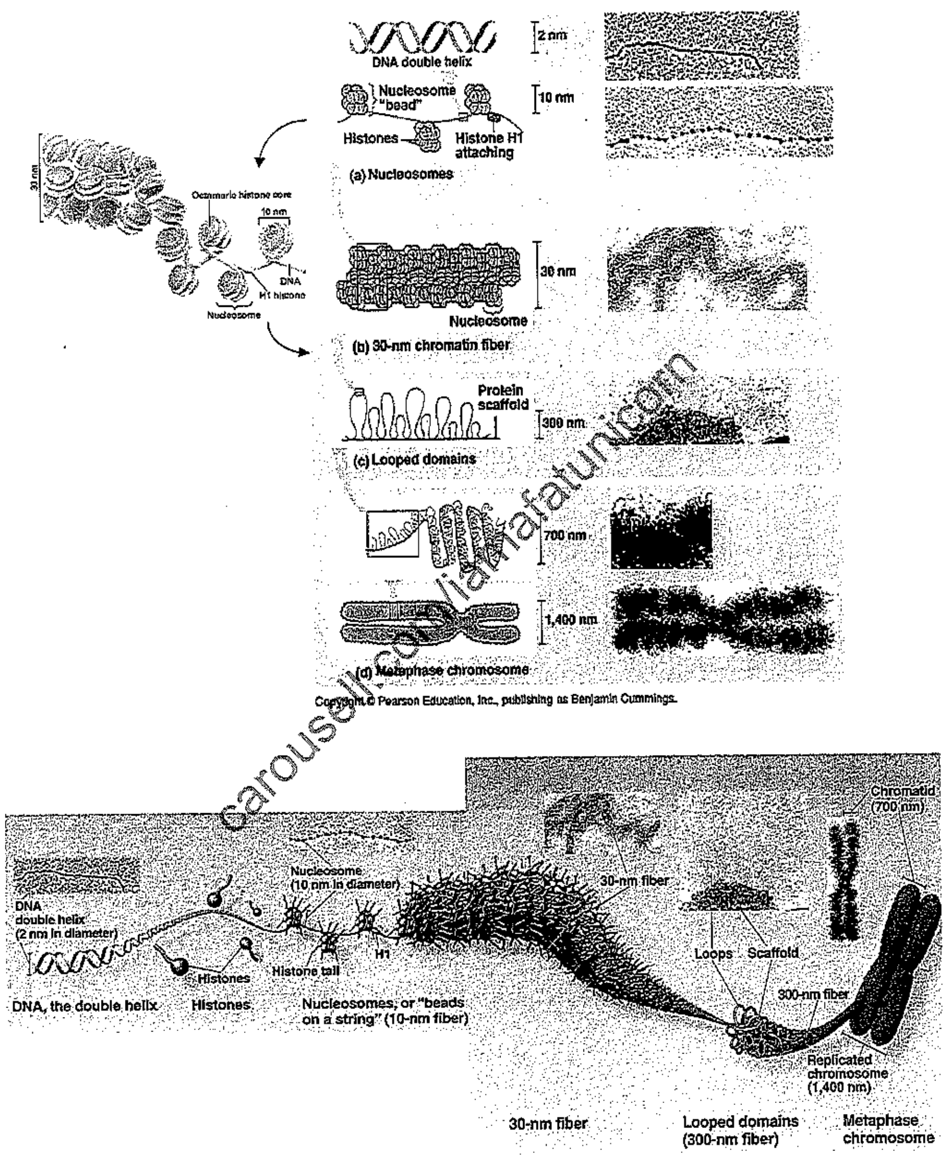
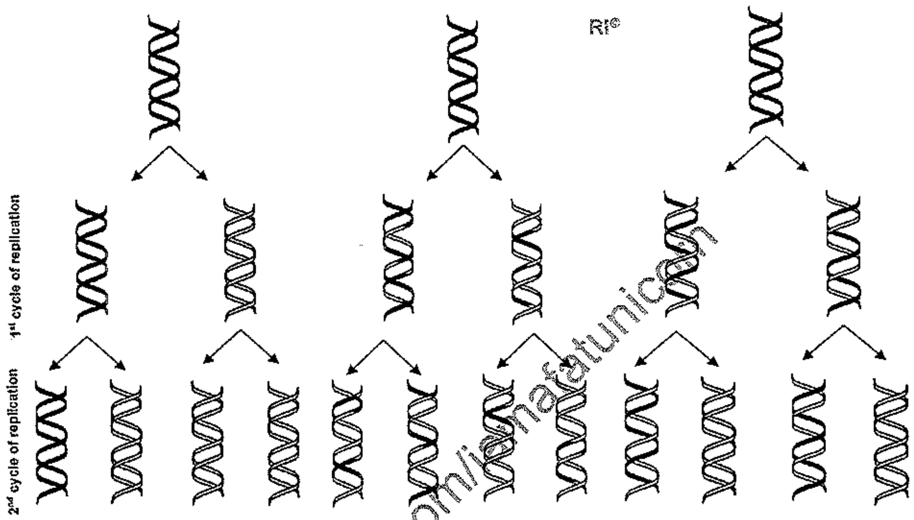


Figure 13. An overview of chromatin packing

(2) DNA REPLICATION

(G) HYPOTHESES FOR THE MECHANISM OF DNA REPLICATION



1) Conservative

- The two parental strands re-associate after acting as templates for new strand synthesis, thus **restoring the original double helix**
- The other daughter DNA molecule consists of **2 newly synthesised strands** making up a new DNA molecule.

2) Dispersive

- The parental DNA molecule is **fragmented and dispersed**
- Daughter molecules contains a **mixture of old and newly synthesised parts**

DNA molecule → double stranded

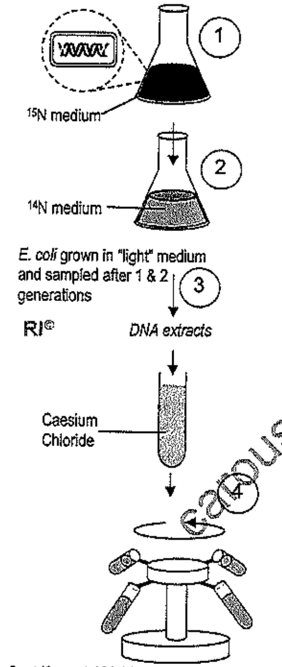
3) Semi-conservative

- Both parental strands of DNA **separate** through **breakage of hydrogen bonds** and **each act as a template** for the synthesis of a new strand through **complementary base pairing**.
- Each DNA molecule formed is a hybrid consisting of **one original strand and one newly synthesised strand (semi-conserved)**

Figure 14. The 3 possible hypotheses to explain DNA replication

(H) EVIDENCE FOR THE SEMI-CONSERVATIVE HYPOTHESIS

- Evidence for the semi-conservative hypothesis was provided by experiments performed by **Meselson and Stahl** in the late 1950s.
- To distinguish between the "new" and the "old" DNA strands, Meselson and Stahl used 2 different ***isotopes*** of nitrogen.
- ^{14}N, is the more common isotope and ^{15}N, is a less common, **heavier** isotope. (One more neutron in nuclei of atom.)
- N is found in the nitrogenous base of DNA and will be incorporated into DNA during replication.

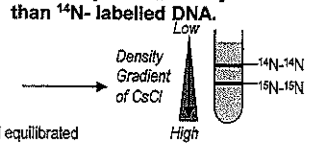


Step 1
A stock of parental ***E. coli**** were grown for many generations in a **medium containing only ^{15}N** so it is incorporated into the nitrogenous bases of all bacterial DNA. This resulted in DNA which was denser than normal DNA.

Step 2
The *E.coli* containing ^{15}N- labelled DNA were then **transferred into a medium containing only ^{14}N.**

Step 3
The transferred *E.coli* were allowed to **divide once and were then collected.** These were called the **"first generation cells"**. Some of these cells were then allowed to **divide once more**, to obtain **"second generation cells."**

Step 4
DNA was then extracted from these cells (from step 3) and then centrifuged at very high speeds for several days in a solution of CsCl.
Over time a **density gradient of CsCl** will be established as there is an increasing concentration of CsCl towards the bottom of the tubes due to sedimentation under centrifugal force. **DNA will settle at its respective density with ^{15}N- labelled DNA being denser than ^{14}N- labelled DNA.**



Step 5
UV absorption photographs reveal dark bands where DNA is present. (as DNA absorbs UV light)

Figure 15. Steps in the Meselson-Stahl experiment

* ***E. coli**, also known as **Escherichia coli*** is a Gram-negative, rod-shaped bacterium that is commonly found in the lower intestine of warm-blooded organisms (endotherms).

Given that;
1. ***Generation time for E. coli is 20 min***
2. ***^{15}N-^{15}N is denser than ^{14}N-^{14}N***

Can you predict the observations for the three proposed hypothesis by drawing bands on the tubes below?

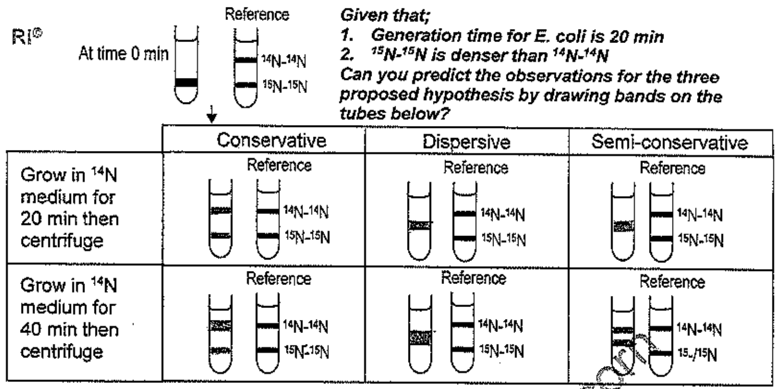


Figure 16. Possible results of the Meselson-Stahl experiment based on the 3 proposed hypotheses

Actual Results

- **Parental generation :** all heavy DNA: pure ^{15}N-DNA (i.e. ^{15}N-^{15}N)
- **1st generation** : **all hybrid** DNA, intermediate in density ('hybrid', containing one heavy ^{15}N chain and one light ^{14}N chain; i.e. ^{15}N-^{14}N).
  → this excludes conservative replication, in which no hybrids formed
- **2nd generation** : **50% hybrid** DNA(i.e. ^{15}N-^{14}N), and **50% light** DNA (i.e. ^{14}N-^{14}N).
  → this excludes dispersive replication, in which no pure ^{14}N DNA should be obtained.

This is consistent with <u>semi-conservative replication</u>.

(I) THE MECHANISM OF SEMI-CONSERVATIVE DNA REPLICATION

- DNA replication is a **complex** process involving **many enzymes** and other **proteins**.
- DNA replication is **rapid** → 6 billion base pairs are copied in a few hours in human cells
- DNA replication is **accurate** → Error rate of one in 10 billion nucleotides

When does DNA replication occur?

- DNA replication occurs during the **S-phase of interphase** of the cell cycle in eukaryotes (you have learned about this in the topic of cell division).

Before the start of DNA Replication

- **Free deoxyribonucleoside triphosphate (dNTP** where N = A, T, G or C**)** are manufactured in the cytoplasm and transported into the nucleoplasm via nuclear pores. E.g. dATP = deoxy**adenosine** triphosphate, dGTP = deoxy**guanosine** triphosphate, dTTP = deoxy**thymidine** triphosphate, dCTP = deoxy**cytidine** triphosphate.

Start of DNA Replication

- Replication begins at a specific site called the **origin of replication**, which has a specific sequence of nucleotides. (Fig. 17a) It is abbreviated as '**Ori**'.
- Specific enzymes such as **helicase** and other proteins are required to **initiate** replication. They recognise & bind to the origin of replication on the parental DNA molecule. (Fig. 17b)
- **Helicase unzips (use this term) and separates the two parental strands** of DNA double helix **by disrupting the hydrogen bonds** between complementary base pairs.
- **Replication forks** form and spread in both directions creating a **replication bubble** (Fig. 17c)
- **Single-strand binding proteins** bind to single DNA strands and **keep the strands apart, preventing** them **from reannealing** so that they can serve as templates for the synthesis of new complementary DNA strands. (Fig. 17c)
- **Topoisomerase** relieves "overwinding" strain ahead of replication forks by breaking, swivelling and rejoining DNA strands (not shown in diagrams).

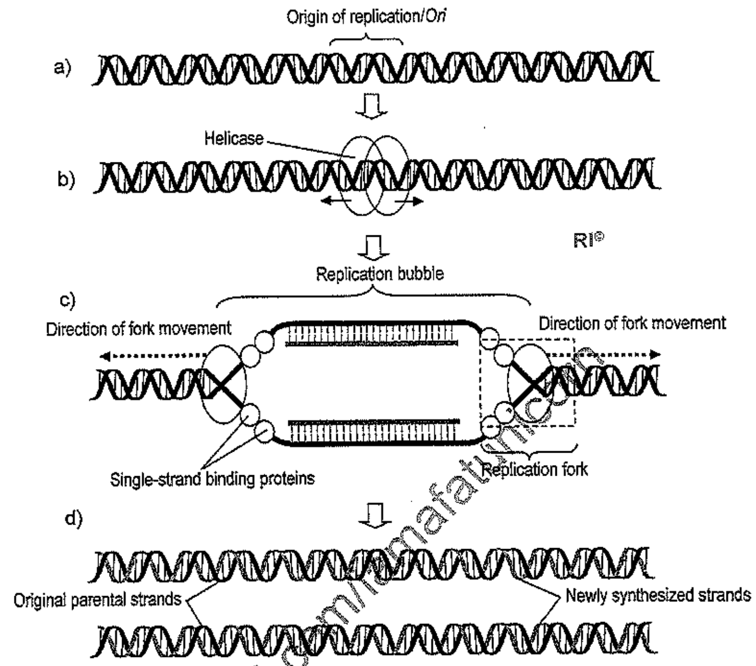


Figure 17: Summary of DNA replication.

Synthesis of the new DNA strands

- The synthesis of new DNA strands require enzymes called **DNA polymerases** that add nucleotides to a **pre-existing** chain which provides the **free 3' –OH group**.
- Therefore DNA polymerase can work only if it has:
  1. A pre-existing chain in the form of a **RNA primer.**
  2. A **template** in the form of the parental DNA strand.

(1) • On each of the parental DNA strands which were unzipped and separated, a short **RNA primer** is added by an enzyme called **primase.** (Fig. 18a)

- Once a RNA primer is synthesised, the enzyme **DNA polymerase** can begin to catalyse the elongation and synthesis of the new complementary strand.

- The end of the RNA primer provides a **free 3' OH** which **is required for DNA polymerase to initiate DNA synthesis.**

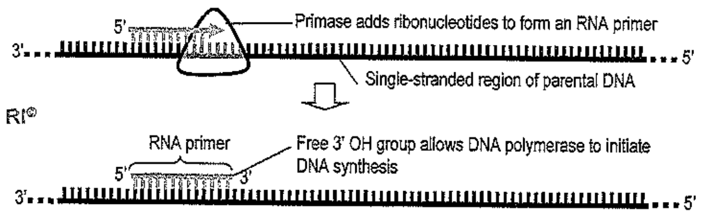


Figure 18a.

(2) • DNA polymerase adds DNA nucleotides to the growing new strand in the **5' to 3' direction.** New DNA nucleotides are added to the 3' hydroxyl end of the growing strand. (Fig. 18b)

- DNA polymerase uses the parental strand as a **template** and aligns the free **activated dNTPs** (deoxyribonucleoside triphosphates) in a sequence complementary to that of the parental strand.
    - **Adenine** base pairs with **thymine**, and vice versa
    - **Guanine** base pairs with **cytosine**, and vice versa

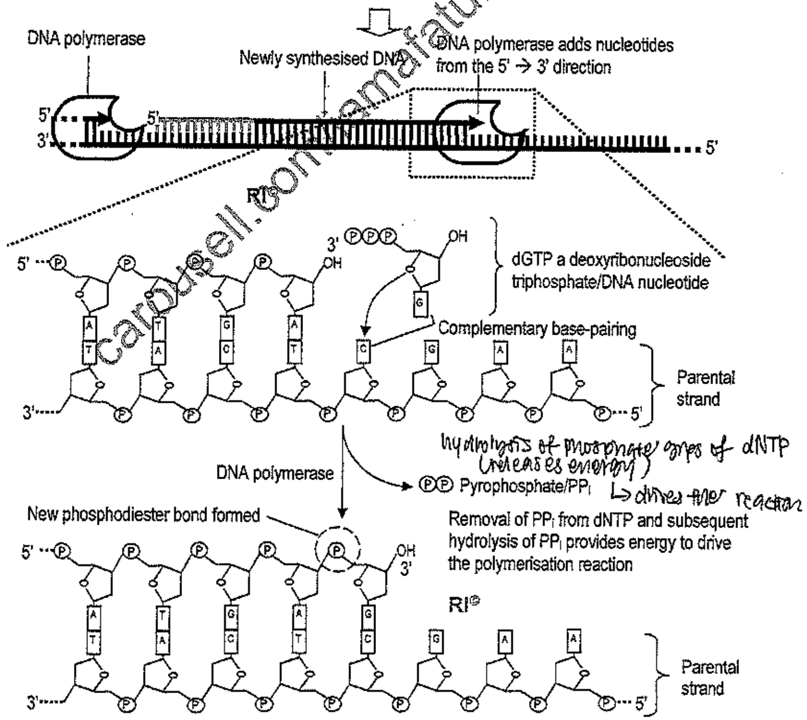


Figure 18b.

- DNA polymerase catalyses the formation of **phosphodiester bonds between adjacent DNA nucleotides** of the newly synthesised strand.

- As DNA polymerase moves along the parental strand, part of the enzyme "**proof-reads**" the previous region.
    - This proof-reading activity ensures that **proper base pairing** has taken place between the bases.
    - If an **incorrect DNA nucleotide is added**, it will be swiftly **removed** by the DNA polymerase and **replaced** with the correct one. This is to ensure the fidelity of the DNA.

(3) • A **different** DNA polymerase then removes the RNA primer and **replaces it with DNA nucleotides.** (Fig. 18c)

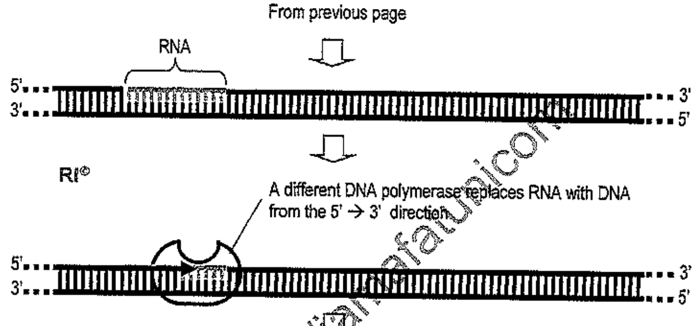


Figure 18c.

(4) • **DNA ligase** forms a phosphodiester bond between two DNA fragments, sealing the nick. (Fig. 18d)

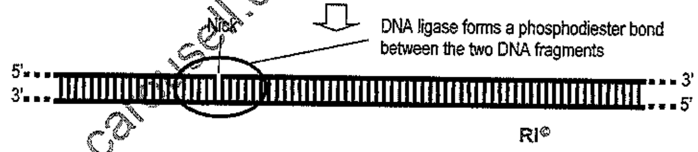


Figure 18d.

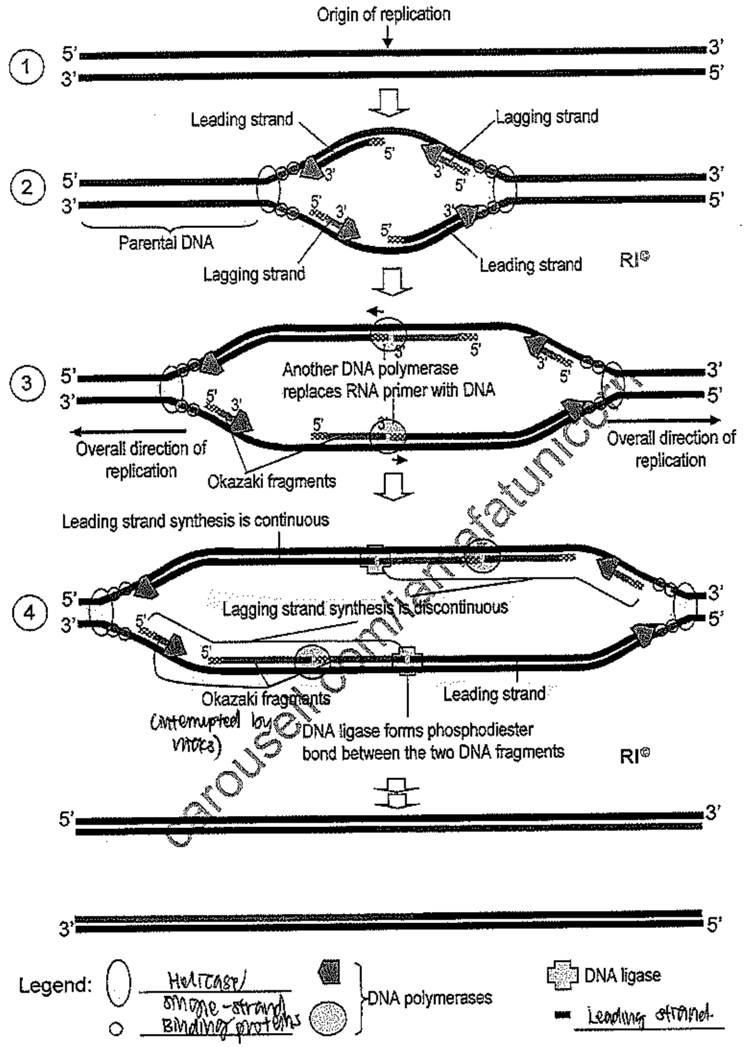


Figure 19. Synthesis of leading and lagging strand during DNA synthesis

(1) • Since the parental strands are anti-parallel, the **2 new strands are synthesised in opposite directions.**

(2) • If we consider just one replication fork in the replication bubble, the **leading strand** is **synthesised continuously** in the **5'→3'** direction. How about the **lagging strand**? **DNA polymerase cannot synthesise DNA in the 3'→5' direction.**

Q. Why can't DNA polymerase synthesise DNA in the opposite direction? (3'→5' direction.)

Recall →
- DNA polymerase is an enzyme & enzymes are specific
- The active site of DNA polymerase is complementary to the free 3'OH group attached at the end of a growing DNA strand
- It is not complementary to the free phosphate group at the 5' end of the DNA strand

- Hence the **lagging strand** is also synthesised in the **5'→3'** direction **but in short segments of** 100-200 nucleotides.

(3) • The fragments produced by this **discontinuous synthesis** are called **Okazaki fragments**. (you have to use this name) Synthesis of each Okazaki fragment is initiated by an RNA primer before the addition of DNA nucleotides.

- Each Okazaki fragment lengthens in the **5'→3'** direction and eventually joins up with the other fragments forming a continuous DNA strand.

(4) • The lagging strand is synthesised discontinuously through the addition of Okazaki fragments at the 5' end of the lagging strand.

- **To produce a continuous DNA strand from the joining of many Okazaki fragments, two steps are required :**
  i. **DNA polymerase** excises the RNA primer and replaces it with DNA.
  ii. Then a linking enzyme **DNA ligase** joins the 3' end of each new DNA fragment to the 5' end of the growing chain by forming a phosphodiester bond.

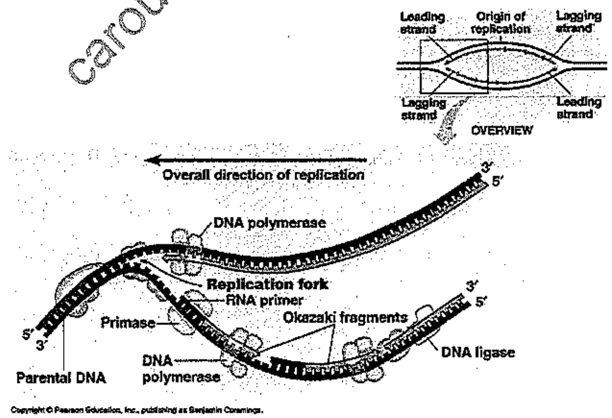


Figure 20. An overview of DNA replication

- Since each DNA polymerase can only add dNTPs at a certain maximum rate, to speed up the replication process, there are multiple origins of replication where many more DNA polymerases can work simultaneously. This occurs in eukaryotes. Prokaryotes such as bacteria have only a single origin of replication as their genomes are a lot smaller. (Fig. 21)

- The replication bubbles will extend in either direction until the bubbles meet and two separate DNA double helices form.
  There are hundreds or even a few thousand origins of replication. Multiple replication bubbles form and eventually fuse, thus speeding up the copying of the very long DNA molecules.

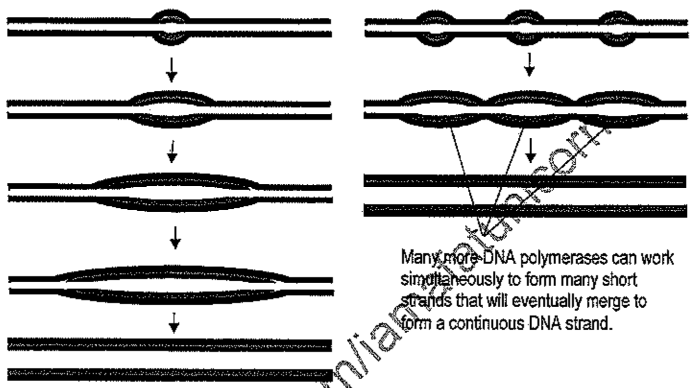


Figure 21. A comparison of the efficiency of a single vs multiple origins of replication

End of replication

- At the end of replication, the complementary parental and newly synthesised DNA strands **form a double helix.**

- The process is **semi-conservative** since each resultant double helix consists of one original strand and one newly synthesised strand.

Q. List the enzymes and proteins required for DNA replication and describe their roles briefly.

Enzyme/protein	Role
Helicase	
Single-strand binding protein	
Primase	
DNA polymerase	
DNA ligase	
Topoisomerase	

(J) ROLE OF DNA

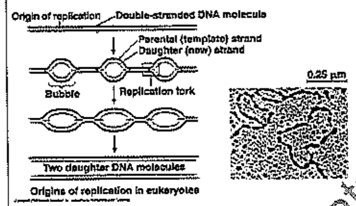
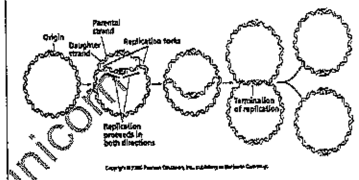
The main role of DNA is to store information and pass it on from one generation to the next.

Q. Why is DNA a suitable store of information? [10]

- Identify property/feature of DNA that makes it suitable as a store of information.
- How so? What are the structural feature(s) of DNA that give rise to the property identified?

Property of DNA that makes it suitable to store information	Structural feature of DNA that gives rise to this property
1. It can be **replicated accurately**. → daughter cells have **identical copies of DNA** as the parent cell	• Weak **hydrogen bonding** between the two strands allow them to separate and act as a **template** for new strand synthesis • (**Adenine** forms **2 hydrogen bonds** with **thymine** and **cytosine** forms **3 hydrogen bonds** with **guanine** through **complementary base pairing**)
2. It is a **stable** molecule. →can be passed on to the next generation without loss of the coded information	• Collectively, numerous **hydrogen bonds** hold the two strands of DNA together • Adjacent nucleotides in each strand are joined by strong covalent **phosphodiester bonds**
3. There are **backup of the code,** since there are two strands in a DNA molecule.	• DNA is **double stranded** • One strand serves as a **template** for the **repair** of the other. Mutations may occur spontaneously in either strand. ▪ *Note: This is different from proofreading which detects and fixes errors in incorporating free nucleotides in the polymerisation process.*
4. Coded information can be **readily utilised/ accessed.**	• weak **hydrogen bonding** allows the **template** strand to separate from the **non-template** strand allowing **transcription** to take place mRNA → proteins • **Complementary base pairing** allows the faithful transfer of info from DNA to RNA in transcription, which will be **translated** to protein subsequently

(K) DNA REPLICATION IN EUKARYOTES & PROKARYOTES

Point of Comparison	Eukaryotes	Prokaryotes
When it occurs	During S phase of interphase	DNA replication occurs prior to cell division by binary fission.
Where	Nucleus	Cytoplasm
Number of origins of replication	Multiple origins (hundreds or even a few thousand) of replication present per linear strand of DNA. 	A single origin of replication present per circular strand of DNA 
Replication ends at	The telomeres at the end of linear chromosomes	The terminus
Rate of elongation	50 nucleotides per sec (in humans)	500 nucleotides per sec (in bacteria)

(L) TELOMERES & THE END REPLICATION PROBLEM

- Telomeres are nucleotide sequences found **at both ends of eukaryotic chromosomes.**

Q. Do prokaryotes have telomeres? NO

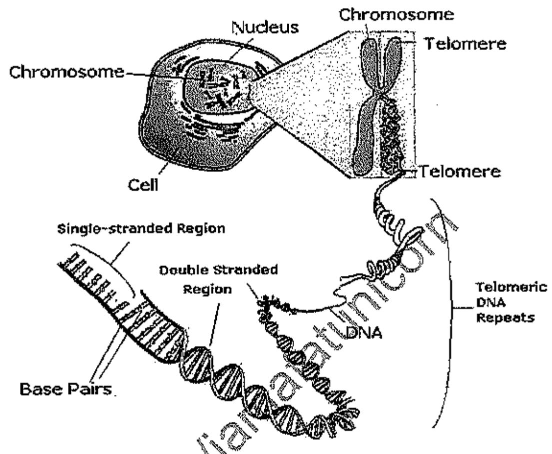


Figure 22. Diagram showing a telomere
http://www.chemsoc.org/ExemplarChem/entries/2003/imperial_Burgoine/origins.txt.html

- They are **non-coding regions of DNA** made up of **a series of tandem repeat sequences.** Each repeat is short (about 5-10 nucleotides) and the number of repeats can range from a 100 to a 1000. In humans, the 6 nucleotide repeat is TTAGGG.
  RNA primer removed → DNA polymerase cannot extend.
  Also, the telomere has a **single-stranded region** at the 3' end (the very end of the telomere) termed as the **3' overhang.** This region of DNA does not have a complementary region (Fig. 21).

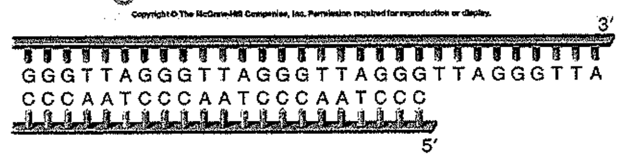


Figure 23. Close-up view of the end of the telomere & the single-stranded region called the overhang

- Telomeres **ensure genes are not lost/eroded with each round of DNA replication** due to the end replication problem. This prevents the **loss of vital genetic information with each replication cycle.**

What is the end-replication problem?

The end replication problem occurs during the replication of **linear** eukaryotic chromosomes.

Recall that during DNA replication, DNA polymerase needs a free 3' end to add free nucleotides to the growing DNA strand.

For one of the daughter chromosomes, this limitation leads to a 3' overhang because the 5' end of its newly-synthesized strand will have the RNA primer removed **without replacement** with DNA. The diagram below (Fig. 22), shows how the overhang comes about:

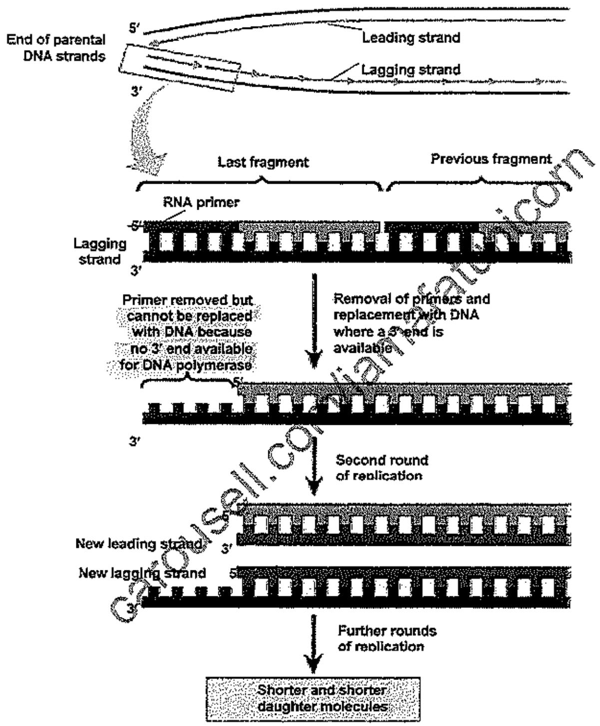


Figure 24. Sequence of events showing how the single-stranded region– the *overhang*, is formed

- Since telomeres are non-coding, the shortening of the chromosome ends leads to the shortening of the telomeres first, without much deleterious effects. The genes within the chromosome will thus be protected.
  In actual fact, research shows that when a shortening chromosome reaches a critical length (at which genes are not yet eroded), the cell will tend to undergo **apoptosis** (cell suicide) as a response to the apparent damage that has occurred to the chromosome and thus the cell as a whole dies even before genes get eroded.

- The enzyme telomerase, is responsible for extending the telomeres and their activity is detected in stem cells. This confers the ability of stem cells to divide indefinitely.

Key Words:

Deoxyribonucleic Acid (DNA)	Ribonucleic Acid (RNA)	Nucleotide
Nucleoside	Nitrogenous base	Adenine
Thymine	Guanine	Cytosine
3' OH group	5' Phosphate group	Purine
Pyrimidine	Sugar phosphate backbone	Phosphodiester bond
Antiparallel	DNA double helix	Semi-conservative replication
Origin of replication	Deoxyribonucleoside triphosphate (dNTP)	Free nucleotides
Complementary base-pairing	Hydrogen bonds	Template
Helicase	Unzipping	Primase
RNA primer	DNA polymerase	Replication bubble
Replication fork	Single-strand binding proteins	Topoisomerase
DNA ligase	Nick	Okazaki fragments
Continuous replication	Discontinuous replication	Leading and lagging strand
End-replication problem	Telomeres	Telomerase
Tandem repeats	3' overhang	Apoptosis

Links to other topics:

DNA replication linked to Nuclear division – S phase
DNA replication linked to Molecular techniques – Polymerase chain reaction
Structure of nucleic acids linked to Viruses and Bacteria – life cycle
End replication problem linked to Prokaryote and Eukaryotic Genome – telomeres and telomerase

CORE IDEA
(2) Genetics & Inheritance

DNA & GENOMICS II

Content
Central Dogma – DNA to RNA, RNA to protein

Learning Outcomes
Candidates should be able to:

2(c) Describe how the information on DNA is used to synthesise polypeptides in prokaryotes and eukaryotes. (Description of the processes of transcription, formation of RNA from pre-mRNA and translation is required.)

2(l) Explain what is meant by the terms gene mutation and chromosome aberration. For gene mutation, knowledge of how substitution, addition and deletion could change the amino acid sequence (e.g. frameshift) is required.
For chromosomal aberration, knowledge of numeral (e.g. aneuploidy, as in the case of trisomy 21, i.e. Down syndrome) and structure (e.g. Translocation, duplication, inversion, deletion) aberration is required (covered in Mitosis & Meiosis)

2(m) Explain how gene mutations can result in diseases (including sickle cell anaemia).

References
Biology, 9th Edition, by Campbell, N.A. and Reece J.B. (2010). (Red spot)
Biology, 8th edition Solomon, Berg & Martin (Red spot)
Biology, 8th Edition, Raven, Johnson, Losos, Mason and Singer, Mc Graw-Hill
Biological Science Vol 1 and 2 3rd edition, Green N.P.O, Stout G.W and Taylor D.J, (2004) Cambridge
Biology, 9th Edition, Mader, Sylvia S (2007), Mc Graw-Hill
Molecular Biology of the Cell, 5th Edition, Alberts (2008)

TABLE OF CONTENTS

This handout is the effort of several Biology teachers at RI(Yr 5-6). It has and will continue to be updated.

(1) THE GENE AND THE GENETIC CODE

(A) CONCEPT OF A GENE

- Every cell in your body contains hereditary instructions in the DNA that you receive from your parents. DNA carries the instructions ('blueprint") for making **proteins**. Information in DNA is copied to **messenger RNA (mRNA)**, which carries this information to ribosomes in the cytoplasm, where it is used to make the polypeptides that make up a protein.
- The information that is carried on DNA is found in stretches of nucleotides called **genes**.
- A gene is a **specific sequence of nucleotides in a DNA** molecule, which codes for a **specific sequence of amino acids** in one polypeptide chain
- Along the length of the DNA molecule, at specific locations, there are different genes each coding for a different gene product (e.g. polypeptide, rRNA etc).

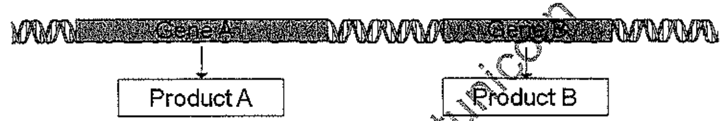


Figure 1. Genes and their products

- Different organisms differ in their sequence of nucleotides (genes) meaning that they each have distinct genetic information that make them distinctly different from each other.

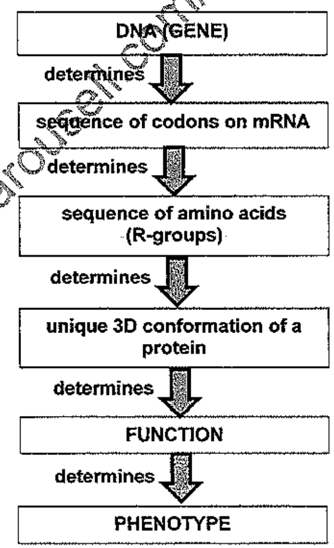


Figure 2. Genes determine phenotypic characteristics of organisms

(B) THE CENTRAL DOGMA

- Genetic information is stored as DNA in all cells, and in many viruses.
- The flow of genetic information from **DNA** to **RNA** to **protein** is the basic universal mechanism of **gene expression**. This is called the **central dogma** of molecular biology (a term coined by Francis Crick). ↳ transcription + translation

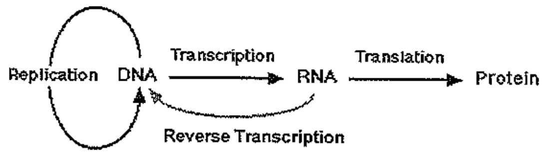


Figure 3. The central dogma of molecular biology

- The central dogma involves the following processes:
  1. **replication** of the genetic information which involves **DNA-directed DNA synthesis** (using DNA as template to synthesise DNA), allows the transmission of genetic information to daughter cells with exceptional fidelity
  2. **gene expression / protein synthesis**, which involves two stages:
     (a) **transcription**: information in DNA is transcribed from **DNA to RNA**, and
     (b) **translation**: where nucleotide sequences on **RNA** is translated into **amino acid sequence** of a polypeptide.

The linear order of bases in a gene specifies the amino acid sequence of a protein, which in turn specifies that protein's three-dimensional conformation and function in the cell.

```
5' ...A T G G C C T G G A C T T C A... 3'
3' ...T A C C G G A C C T G A A G T... 5'  ← Template strand
          ↓ Transcription of template strand
5' ...A U G G C C U G G A C U U C A... 3  ← mRNA
          ↓ Translation of mRNA
N— Met — Ala — Trp — Thr — Ser — C  ← Polypeptide
```

Figure 4. The central dogma

3. **reverse transcription:** With the discovery of the enzyme reverse transcriptase, the **reverse flow of information from RNA to DNA** is possible under certain circumstances.

- The processes have other differences:

Process	Key molecule	Location
replication	DNA polymerase	nucleus
transcription	RNA polymerase	nucleus
translation	ribosomes	rough endoplasmic reticulum and cytoplasm
reverse transcription	reverse transcriptase (commonly linked to viruses).	See 'Genetics of Viruses and Bacteria'

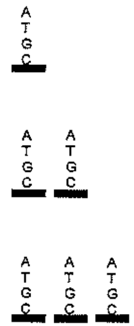
(C) THE GENETIC CODE AND ITS FEATURES

How do we make sense of the instructions encoded in DNA?

- In DNA, there are 4 different nitrogenous bases (adenine, thymine, guanine and cytosine) and hence, 4 different nucleotides dATP, dTTP, dGTP, dCTP). **Different genes have different sequences of nucleotides.** The sequence of nucleotides in the DNA eventually defines the sequence of amino acids in a protein.
- All proteins are made up of the permutations of **20 different amino acids**. Different proteins have different sequences of these 20 amino acids.

How many nucleotides code for 1 amino acid?

Conditions:
1. each nucleotide position (labelled as ▬) can be occupied by 4 different bases, and
2. 20 different amino acids must be produced.

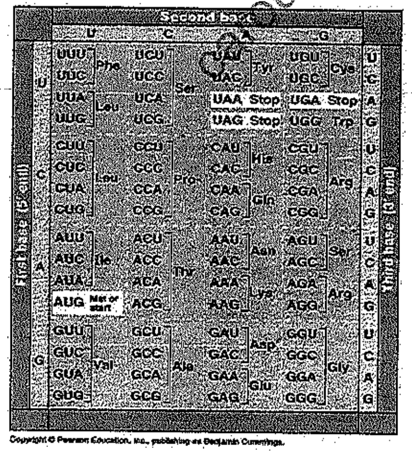


If 1 nucleotide is to code for 1 amino acid:
→ the 4 different bases will allow for only 4 codes, each coding for 1 amino acid.
→ insufficient to code for 20 amino acids

If 2 nucleotides are to code for 1 amino acid:
→ the protein can contain up to 4 x 4 = 16 different codes each coding for 1 amino acid
→ still insufficient to code for 20 amino acids

If 3 nucleotides code for 1 amino acid:
→ the protein can contain up to 4^3 = **64** different codes
→ more than sufficient to code for 20 amino acids
→ these **64 codons** form the set of instructions that tells a cell the order in which amino acids are to be joined to form a protein

Conclusion: **3 consecutive nucleotides along a DNA template code for 1 amino acid.**



First base (5' end)	U	C	A	G	Third base (3' end)
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
U	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
U	UUA Leu	UCA Ser	UAA Stop	UGA Stop	A
U	UUG Leu	UCG Ser	UAG Stop	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
C	CUC Leu	CCC Pro	CAC His	CGC Arg	C
C	CUA Leu	CCA Pro	CAA Gln	CGA Arg	A
C	CUG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
A	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
A	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
A	AUG Met or start	ACG Thr	AAG Lys	AGG Arg	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
G	GUC Val	GCC Ala	GAC Asp	GGC Gly	C
G	GUA Val	GCA Ala	GAA Glu	GGA Gly	A
G	GUG Val	GCG Ala	GAG Glu	GGG Gly	G

Figure 5. The mRNA codons for amino acids

	Amino acid
Ala	alanine
Arg	arginine
Asn	asparagine
Asp	aspartic acid
Cys	cysteine
Gln	glutamine
Glu	glutamic acid
Gly	glycine
His	histidine
Ile	isoleucine
Leu	leucine
Lys	lysine
Met	methionine
Phe	phenylalanine
Pro	proline
Ser	serine
Thr	threonine
Trp	tryptophan
Tyr	tyrosine
Val	valine

Figure 6. Three-letter letter abbreviations of amino acids

Features of the genetic code

1. The code is a **triplet code**
   - Every amino acid in a protein is coded for by a triplet code (sequence of 3 consecutive nucleotides/ nucleotide bases/ bases) on the DNA molecule.
   - The **triplet of nucleotides in the mRNA** are called **codons** that codes for an **amino acid.** (found on mRNA transcript)

2. The code is **universal**
   - The same triplet of nucleotides codes for the same amino acid in **all organisms.**
     e.g. GGU will always code for glycine, whether it is in *E. coli*, orchid or koala.
   - This is the basis of genetic engineering. Genetic material such as DNA / mRNA can be isolated from one species and be expressed by the protein-synthesising machineries of another species.

3. The code is **degenerate (redundant)**
   - For some amino acids out of the 20, the same amino acid may be coded for by several codons.
   - When an amino acid is encoded by more than 1 codon, i.e., the code is said to be **degenerate.**
     e.g. the amino acid glycine can be coded for by 4 different codons.
     ↗ First 2 of the 3 nucleotides in a codon are the same
   Question: What do you notice about the codons which code for the same amino acid?
   - more than 1 codon can code for same amino acid (degeneracy)
   - However, each codon only codes for 1 specific amino acid! (specific, no ambiguity)

4. The code is **non-overlapping**
   - The codons do not overlap, but read as **successive groups of 3 nucleotides.**
     i.e. each nucleotide in a triplet code is only used once.

Figure 7. The non-overlapping code

5. The code is **continuous (i.e. not punctuated)**
   - There are no nucleotides 'skipped' between the codons; the code is read as a continuous sequence of nucleotide bases.

6. The code includes 'stop' and 'start' sequences
   - **Start codon : AUG** (are you great?)
     - This functions as a 'start' signal for ribosomes to **begin translating** mRNA into a sequence of amino acids.
     - Also codes for the amino acid methionine. Hence, the first amino acid of a polypeptide is always methionine.
   - **Stop codons : UAG / UAA / UGA** (you are great | you are awesome | you're great & awesome)
     - These codons **do not code** for any amino acid, as there is no tRNA with an anticodon complementary to these 3 codons.
     - They act as 'stop signals' for the **termination** of polypeptide chain synthesis during translation.

Note:
- The way in which the codons are read during translation, in groups of three nucleotide bases beginning with the start codon is known as **reading frame.**

In summary, the genetic code:
- is stored in the sequence of nucleotides in DNA.
- codes for the amino acid sequence of polypeptides.

- is transcribed into a specific complementary sequence of nucleotides in mRNA, which is then translated into an amino acid sequence.

(2) PROTEIN SYNTHESIS

(A) GENE EXPRESSION/ PROTEIN SYNTHESIS

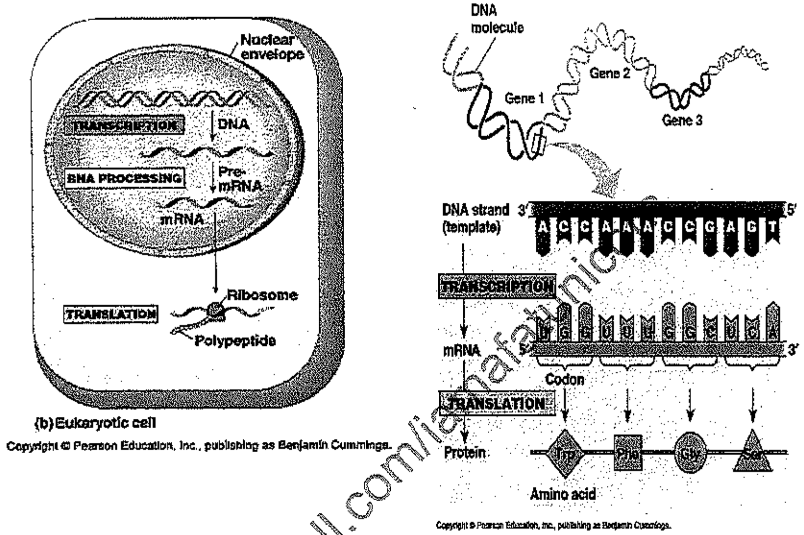


Figure 8. An overview of protein synthesis

- Genetic information carried in the **nucleotide sequence along a gene** (DNA) is found in the **nucleus**. This information specifies the **amino acid sequence** of the corresponding polypeptide chain.
- Since DNA molecules are too large to move through **nuclear pores in nuclear envelope**, part of the information on the DNA is **transcribed**/ copied into smaller **messenger RNA (mRNA)** molecules, which can pass through the nuclear envelope.
- In the cytoplasm, **ribosomes** interact with **mRNA** and amino-acid-carrying **transfer RNA (tRNA)** molecules, to translate the information in the mRNA molecule into a polypeptide, via a process called **translation**. Hence translation occurs in the **cytoplasm.**

(B) TRANSCRIPTION (DNA →RNA)

verb : transcribed.

Definition

Transcription is the process in which the nucleotide sequence of a gene (DNA) is used as a template to direct the synthesis of RNA (namely **mRNA, tRNA and rRNA**) made up of complementary base sequences.

1. Components of a Gene

- A gene is a **specific sequence of nucleotides in a DNA** molecule, which codes for a specific **sequence of amino acids** in one polypeptide chain. It is a unit of inheritance located in a fixed position (locus) on the chromosome which specifies a particular biological function, i.e. phenotype.

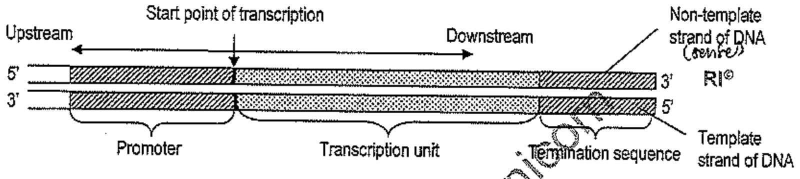


Figure 9. Components of a gene

- The **gene** includes:
  1) **promoters** are DNA sequences which serve as a recognition site for binding of **RNA polymerase** and regulatory proteins called **general transcription factors** to **initiate/begin** transcription.
     - It also determines which one of the two strands of the DNA molecule is used as the template for transcription.
     - Promoters are classified as non-coding DNA as they do not code for proteins or RNA products (i.e. rRNA and tRNA).
  2) **termination sequence**, at the end of a gene causes the synthesis of RNA to stop.
  3) **transcription unit**, is a sequence of DNA that is flanked by the promoter and termination sequence, and transcribed into RNA.
     - Only 1 of the 2 strands of the DNA, serves as a template for transcription, to direct synthesis of the RNA. This strand is known as **template strand** (template = non-coding strand, as mRNA sequences are complementary to the template strand. Sometimes called antisense strand).
     - The other strand is known as **non-template strand** (non-template = coding strand. mRNA sequences are identical to this strand except that mRNA has U instead of T. Sometimes called sense strand).

* Template B:
- synthesise from 5' to 3'
- Read from 3' to 5'

2. General Key Features of Transcription (in both prokaryotes and eukaryotes)

(1) Formation of single-stranded RNA	Transcription results in formation of **single-stranded RNA** molecules. • Only 1 of DNA strands, the template strand, is transcribed.
(2) Synthesis in 5' to 3' direction	Synthesis of RNA molecule occurs only in **5' to 3'** orientation. • 5' phosphate group of incoming ribonucleoside triphosphate reacts with 3' hydroxyl group (-OH) of the growing polynucleotide. • Template strand is read in 3' to 5' direction.
(3) Catalysed by RNA polymerase	Transcription requires presence of the enzyme **RNA polymerase** which catalyses assembly of **ribonucleotides** and formation of **phosphodiester bonds** between growing polynucleotide and incoming ribonucleotides.
(4) Complementary base-pairing	Newly synthesised RNA is assembled through **complementary base pairing** with DNA template, where adenine forms 2 hydrogen bonds with uracil, thymine forms 2 hydrogen bonds with adenine, while guanine forms 3 hydrogen bonds with cytosine.
(5) Recognition sequences	Transcription begins at **promoter** (recognised by RNA polymerase) and ceases when it encounters **termination sequence**.

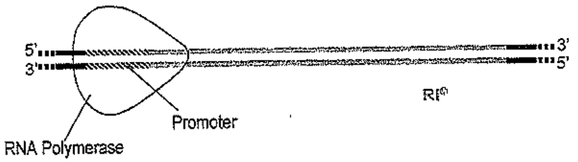
3. Transcription - The General Process

I) INITIATION

Step 1: Assembly of **RNA polymerase** and other proteins factors at the **promoter**.

(Note: differences between prokaryotic and eukaryotic systems will be discussed in 'Organisation and Control of Prokaryotic and Eukaryotic Genome').

Figure 10. Binding of RNA polymerase to promoter



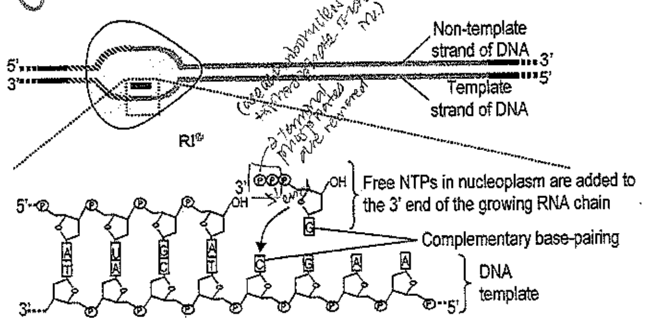
Step 2: RNA polymerase **unzips and separates the two strands** of the DNA double helix at the promoter by **breaking hydrogen bonds** between complementary base pairs.
DNA now has two exposed strands:
(1) one strand is used as a **template**.
(2) the other strand is not transcribed. (NB : It is not spelt as TRANSCRIPTED!!)

II) ELONGATION

Step 3: **Free ribonucleotides** from nucleoplasm are matched up with the DNA template by **complementary base pairing**

- Adenine forms 2 hydrogen bonds with uracil, thymine forms 2 hydrogen bonds with adenine, while guanine forms 3 hydrogen bonds with cytosine.
- As each **nucleoside triphosphate** is brought in, its two terminal phosphates are removed.
- 5' end of the nucleotide is added to 3' -OH end of the growing RNA chain via formation of a phosphodiester bond.
  i.e. mRNA is synthesised and elongated in **5' to 3' direction**.
- RNA polymerase catalyses the joining of adjacent ribonucleotides (polymerisation) through formation of covalent **phosphodiester bonds**. This forms the **sugar-phosphate backbone** of the growing mRNA transcript.

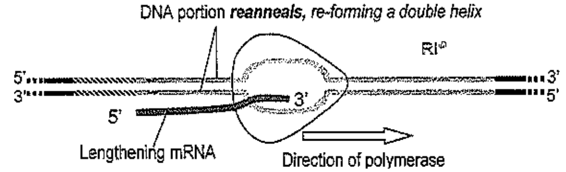
Figure 11. Elongation stage of transcription



Step 4: RNA polymerase continues to move along the **DNA template strand** from its 3' end towards its 5' end (i.e. **3' to 5' direction**), separating the 2 DNA strands. RNA polymerase continues to catalyse the assembly of ribonucleotides.

Step 5: As RNA polymerase continues down the template strand, region of DNA that has just been transcribed **reanneals.**

Figure 12. Elongation in 5' to 3' direction

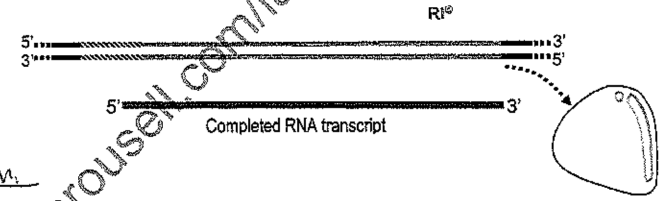


III) TERMINATION

Step 6: After RNA polymerase transcribed through specific DNA sequences known as **termination sequences**, the mRNA chain is released. RNA polymerase dissociates, terminating transcription.

(Note: Several molecules of RNA polymerase can simultaneously transcribe the same gene.)
- In eukaryotes, the completed mRNA undergoes some modifications (see 'Organisation and Control of Prokaryotic and Eukaryotic Genome') before it ready to be used for translation into a polypeptide.

Figure 13. Termination



Differences between DNA replication & transcription.
- Any RNA primer?
Transcription: no need RNA primer (RNA polymerase can start RNA chain from scratch).
Replication - primase synthe. RNA primer
- starting location?
Transcription - promoter of gene
Replication : ori

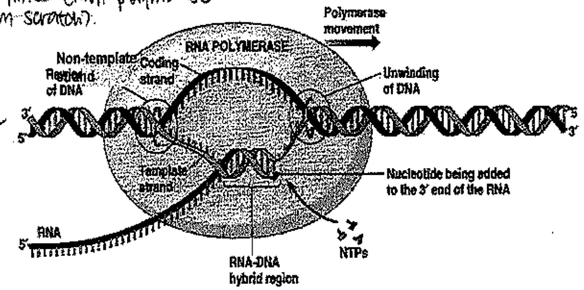


Figure 14. Overview of transcription

(C) RIBONUCLEIC ACID (RNA)

- DNA directs RNA synthesis and, through RNA, controls protein synthesis. RNA is a central player in gene expression. Its function depends critically upon base pairing, to (1) form its secondary structure, and (2) interact specifically with other RNA molecules.

Structural Differences between DNA and RNA

Differences	DNA	RNA
1) Double-stranded vs single-stranded	**Double-stranded** polynucleotide chain (except for some viral genomes).	**Single-stranded** polynucleotide chain (except for some viral genomes).
2) Size	Larger molecular mass (100 000- 150 000 000) and is much **longer.**	Smaller molecular mass (20 000 – 2 000 000) and is **shorter.**
3) Sugar residue	Pentose sugar is **deoxyribose**.	Pentose sugar is **ribose**.
4) Bases	Nitrogeneous bases are adenine (A), guanine (G), cytosine (C) & **thymine (T).**	Nitrogeneous bases are adenine (A), guanine (G), cytosine (C) & **uracil (U).**
5) Ratio of bases	Due to complementary base pairing, Ratio of A:T & C:G = **1:1**, and ratio of (A+G):(C+T) = **purine : pyrimidine = 1:1.**	**No fixed ratio** / cannot be predicted as single-stranded, hence no complementary base-pairing.
6) Structural variation	Always a double helix (except for some viral genomes).	Almost always occurs as a single-stranded molecule (except for some viral genomes). Different parts of a single RNA molecule can base pair with each other via complementary base pairing. This forms a complex structure.
7) Stability	Chemically stable as it is more resistant to spontaneous and enzymatic breakdown. Deoxyribose sugar contains a H atom instead of –OH at carbon 2.	Chemically unstable as it is more reactive partly because of the additionally reactive 2'OH groups of ribose. Being single-stranded, they are more prone to nuclease activity.
8) Forms	Only **one form.**	**Several kinds of** RNA, each with its own function. E.g. messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA)
9) Location (in eukaryotes)	Usually in nucleus (with exceptions in mitochondria, and chloroplast).	Manufactured in nucleus but found throughout cell including in nucleus.
10) Amount	Amount is constant for all somatic cells of a species (except for gametes and spores).	Amount varies from cell to cell (different cell type or the same cell at different stages of its life cycle or according to the cell's metabolic activity).

Table 1: Differences between DNA and RNA

- Only messenger RNA (mRNA) encodes polypeptides, so it called a **coding RNA.**. All the other classes of RNA are thus considered **non-coding RNA.**

- 3 key forms of RNA are involved in protein synthesis:
  1) messenger RNA (mRNA)
  2) ribosomal RNA (rRNA)
  3) transfer RNA (tRNA)

I. Messenger RNA (mRNA)

Structure:
- Exists in a single-stranded form.
- Varies in length, depending on the length of polypeptide that it codes for.
- In prokaryotes, the mRNA does not need to be modified and can be used immediately for translation.
- In eukaryotes, mRNA (more accurately distinguished as pre-mRNA) first undergoes **post-transcriptional modification** before it is transported to the cytoplasm. This includes the addition of 5' cap, 3' poly A tail, and splicing. The modification is necessary for translation to take place (see 'Organisation and Control of Prokaryotic and Eukaryotic Genome').
  As such, eukaryotic mRNA contains two regions:
  1) **coding region**: consists of many codons coding for the amino acid sequence of the protein, starting (usually) with start codon, AUG and ending with a stop codon.
  2) **untranslated regions (UTR)** :
     - 5' UTR: additional sequence at the 5' end upstream of the start codon, and
     - 3' UTR: sequence following the stop codon.

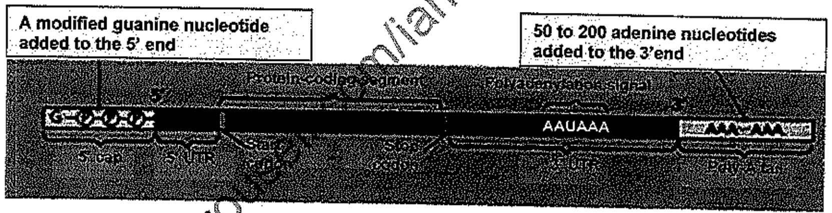


Figure 15. RNA processing: addition of the 5' cap and poly (A) tail. Both the 5' cap and poly (A) tail are not translated into protein, nor are the regions called the 5' untranslated region (5' UTR) and 3' untranslated region (3' UTR). *Do KIV this figure till you cover Prokaryote and Eukaryotic genome.*

Function :
- Messenger RNA (mRNA) serves as a **messenger** that is particularly important for eukaryotes because it is necessary to bring the information out of the nucleus via nuclear pore to cytoplasm where translation takes place in ribosomes.
- mRNA acts as a **template for translation**, i.e. guiding assembly of amino acids into a polypeptide chain.
- Each **codon** (read from 5' to 3') within the coding region of the mRNA represents an amino acid in a polypeptide. **Sequence of codons** on mRNA will determine **sequence of amino acids in corresponding polypeptide** chain.
- Depending on how many polypeptides need to be translated from it, RNA polynucleotide strand may exists for a relatively short-time. This helps the cell to control cellular activity.

(For your info only)

mRNA for HOUSEKEEPING PROTEINS

Many mRNAs are common to most cells, encoding "housekeeping" proteins needed by all cells (e.g. the enzymes of glycolysis). Other mRNAs are specific for only certain types of cells. These encode proteins needed for the function of that particular cell (e.g. the mRNA for haemoglobin in the precursors of red blood cells).

II. Transfer RNA (tRNA)

- Constitutes 15% of total RNA of the cell.

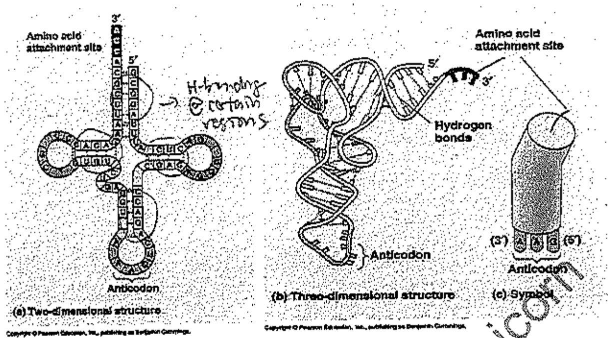


Figure 16. A simplified diagram of the structure of a tRNA molecule

Structure:

- Transfer RNA (tRNAs) are small (~4S, S = Svedberg, a measure of sedimentation rate of particles during centrifugation), containing only about 80 nucleotides.
- Exists in a single-stranded form but segments of a tRNA molecule can **fold** in such a way that they form complementary base pairs with each other.
- All tRNAs have a common structure:
  - It folds back upon itself and held in shape by **hydrogen bonding** between **complementary base pairs** at certain regions to form a 3D L-shaped structure (in diagrams it is simplified into a cloverleaf (2D) structure.)
    → it has 3 loops.
  - On one of the loops, 3 specific unpaired **triplet bases** form an **anticodon** that binds to a specific **mRNA codon** via **complementary base-pairing.**
  - ~~On another loop,~~ **3' end with CCA stem** is the attachment site for a **specific amino acid that corresponds to the anticodon.**
- The process of attaching one amino acid to the 3' CCA stem is called tRNA activation, which will be covered later.

Question: *What is the least no. of different tRNAs that should exist in any cell?*

Function:

- tRNAs bring in **specific amino acids** in a sequence corresponding to the **sequence of codon** in mRNA to the growing polypeptide.
- It can facilitate translation due to:
  (1) its ability to bind to **a specific single amino acid**, and
  (2) the ability of its **anticodon** to base-pair with the mRNA **codon.**

(For your info only)

Wobble base phenomenon

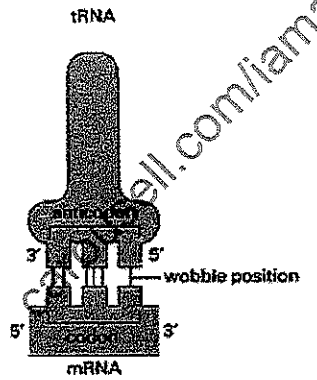
We have seen in the previous section that the genetic code is redundant; that is, several different codons can specify a single amino acid. This redundancy implies either (1) that there is more than one tRNA for many of the amino acids or (2) that some tRNA molecules can base-pair with more than one codon. In fact, both situations occur.

Some amino acids have more than one tRNA and some tRNAs are constructed so that they require accurate base-pairing only at the first two positions of the codon and can tolerate a mismatch (or ***wobble***) at the third position. In some cases, a single tRNA can recognize 2 or more of these degenerate codons, E.g. phenylalanine tRNA with the anticodon AAG recognizes not only UUC but also UUU.

This wobble base-pairing explains why so many of the alternative codons for an amino acid differ only in their third nucleotide.

The violation of the usual rules of the base pairing at the 3rd nucleotide of a codon is called **WOBBLE.** This suggests that a base-pairing at the 3rd base is not so specific; hence a change in the 3rd base by a mutation may still permit the correct incorporation of a given amino acid in a protein

In bacteria, wobble base-pairings make it possible to fit the 20 amino acids to their 61 codons with as few as 31 kinds of tRNA molecules. There are around 45 different tRNA in a typical eukaryotic cell, thus many amino acids have multiple tRNA. The exact number of different kinds of tRNAs, however, differs from one species to the next.



bacteria

wobble codon base	possible anticodon bases
U	A, G, or I
C	G or I
A	U or I
G	C or U

eucaryotes

wobble codon base	possible anticodon bases
U	G or I
C	G or I
A	U
G	C

Figure 17. Wobble base-pairing between codons and anticodons.
If the nucleotide listed in the first column is present at the third, or wobble, position of the codon, it can base-pair with any of the nucleotides listed in the second column.

III. rRNA in ribosomes

- rRNA is synthesised in the nucleus, within **nucleolus.**

Structure:
- Single-stranded chain made up of several thousands of nucleotides.
- Each chain is wound into a complex structure comprising of single and double helices.

Function:
- rRNA associates with a set of proteins to form ribosomes;
- rRNA is the main constituent of the
  1) interface between the large and small subunits of the ribosome
     → Thus the small ribosomal subunit can bind to the mRNA as **complementary base pairing** can occur between the rRNA in the mRNA binding site of the small ribosomal subunit and the mRNA.
  2) **Peptidyl site (P-site)** and **Aminoacyl site (A-site)** on the large ribosomal subunit
     → hence rRNA enables the binding of aminoacyl-tRNAs to the P site and A site (*KIV for translation*)

- Part of rRNA molecule on the large ribosomal subunit, acts as a ribozyme, **peptidyl transferase,** which **catalyses the formation of the peptide bond** between 2 amino acids.

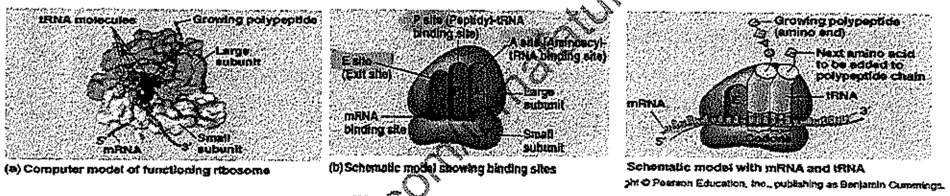


Figure 18. Structure of ribosome

Recall *(from Cell Structure notes)*:

Structure of Ribosome:
- Aggregates of **rRNA & proteins**
- Ribosomes consists of 2 subunits :
  1) **Small subunit** with a mRNA binding site.
  2) **Large subunit**: has 3 binding sites for tRNA, namely :
     (i) **Aminoacyl site** (A site) holds the incoming tRNA carrying next amino acid to be added
     (ii) **Peptidyl site** (P site) holds the tRNA carrying the growing polypeptide chain,
     (iii) **Exit site** (E site) from which the tRNA leaves the ribosome.

In Eukaryotes : 80S ribosome.

In Prokaryotes: 70S ribosome.
(S = Svedberg, a measure of sedimentation rate of particles during centrifugation)

Function of Ribosome:
- Ribosome is the organelle that synthesises polypeptide under the direction of mRNA by:
  - Holding the tRNA and mRNA in close proximity to allow interaction between the **codon of mRNA** and the **anticodon of tRNA.**
  - It positions the new amino acid for addition to the growing polypeptide.
  - **Peptidyl transferase** (the rRNA component of the ribosomal large subunit) catalyses the formation of **peptide bonds** between the two amino acids.

- A ribosome begins translation at the **5'end** of a coding region of mRNA and proceeds **towards the 3' end.**

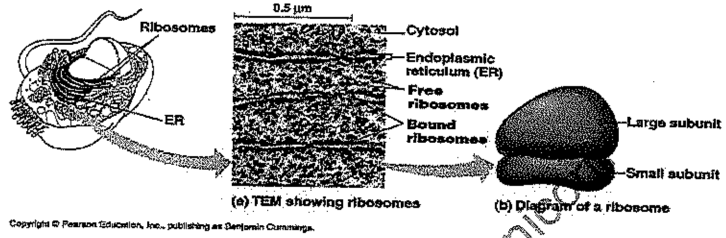


Figure 19. Ribosomes are bound to the endoplasmic reticulum or may be found freely floating in the cytoplasm

(D) TRANSLATION

Definition
Translation is the process by which the sequence of ribonucleotides in an **mRNA** molecule is converted into a sequence of **amino acids in a polypeptide** chain.

- Before translation can occurs, each **tRNA** must be attached to/ charged with the **correct amino acid**, a process referred to as **amino acid activation.**

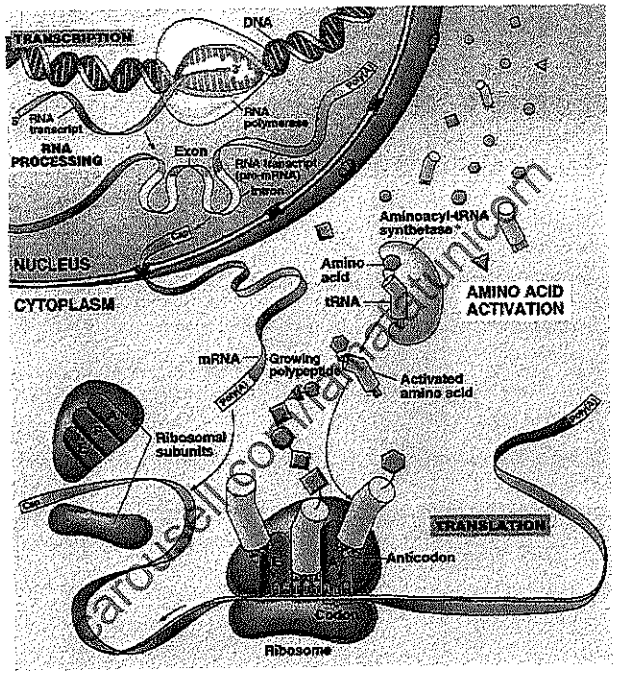
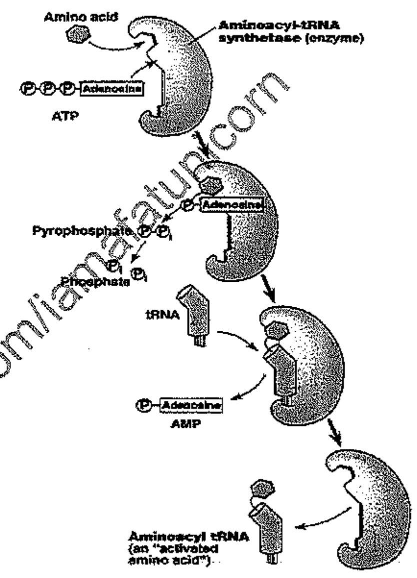


Figure 20. Overview of protein synthesis

1. Amino Acid Activation

- Each specific amino acid is covalently attached/bonded to its specific (set of) tRNA, with specific a anticodon, at its **3' CCA stem**, forming an **amino-acyl tRNA.**
- Catalysed by a group of about 20 different enzymes known as **aminoacyl-tRNA synthetases**, one for each amino acid.
- An aminoacyl-tRNA synthetase has active sites that are complementary to **conformation and charge** of
  (1) the **specific amino acid** to be attached to the tRNA, and
  (2) the **anticodon on the tRNA**.
  Through which, the enzyme achieves double specificity, ensuring that the correct amino acid is attached to the tRNA with the corresponding anticodon.
- This reaction requires ATP.

Figure 21. An example of amino acid activation. The attachment of amino acid to its specific tRNA molecule



- How is tRNA charged / activated with the correct amino acid?
  There are about **20 different amino-acyl tRNA synthetases**, one for each amino acid; i.e. one attaches glycine to all tRNAs that recognise codons for glycine, and so on.
  → These enzymes must recognise the different tRNAs through **unique identity sites** at the **acceptor stem and/or anticodon loop** of the tRNA molecule
  → and recognise **specific amino acids.**

2. Translation - The Process

I) INITIATION

STEP 1: Binding of ribosome to mRNA.

- Through the help of **initiation factors**, the small subunit of the ribosome assemble at the **start codon AUG** (downstream of 5' end of mRNA). This is where translation begins.

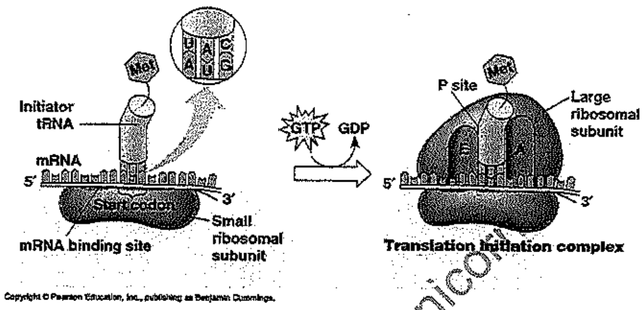


Figure 22. Binding of mRNA to ribosome and formation of translation initiation complex

- Next, **initiator tRNA** (carrying methionine) forms base pairs with complementary AUG codon on mRNA at the ribosome.

Question: *What is the anticodon on this initiator tRNA?*
Question: *What is the first amino acid on the polypeptide?*

- Binding of the large ribosomal subunit completes a ribosome, resulting in formation of **translation initiation complex**.
- The initiator aminoacyl-tRNA is now positioned in **P site**, leaving **A site** vacant for addition of the next aminoacyl-tRNA molecule.
- **GTP** required for the initiation stage.

II) ELONGATION AND TRANSLOCATION

- **GTP** provides energy for initiation and elongation.
- A polypeptide is always synthesised in one direction. It starts at **amino/ N terminal** and ends at **carboxylic/ C terminal.**

Step 2: Codon recognition

- A second tRNA carrying the next amino acid in the chain binds to **A site** by **complementary base-pairing**, forming hydrogen bonds, between its anticodon and the second codon on mRNA.

Step 3: Peptide bond formation

- A **peptide bond** is formed between the methionine and the second amino acid in the **A site**.
- This is catalysed by **peptidyl transferase**, a ribozyme present in the **large subunit** of ribosome.
- In order to form the peptide bond, the first amino acid dissociates from the initiator tRNA it was originally bound to.

Step 4: Translocation

- The ribosome translocates / shifts one codon down the mRNA in **5' to 3' direction**. As a result,
  - 1^{st} **tRNA** is now shifted to **E site** and then **released into cytosol**, where it can be recycled.
  - 2^{nd} **aminoacyl-tRNA** has now moved from **A site to P site.**

  - Empty A site is now ready to **receive 3^{rd} aminoacyl-tRNA**, with anticodon complementary to the third codon along mRNA.
- Thus,
  - P site holds tRNA carrying growing polypeptide chain,
  - A site holds tRNA carrying next amino acid and
  - E site holds tRNA which has donated its amino acid and is ready to leave
- The process is repeated until **stop / termination codon (UAA / UAG / UGA)** on mRNA is exposed at A site on the ribosome.

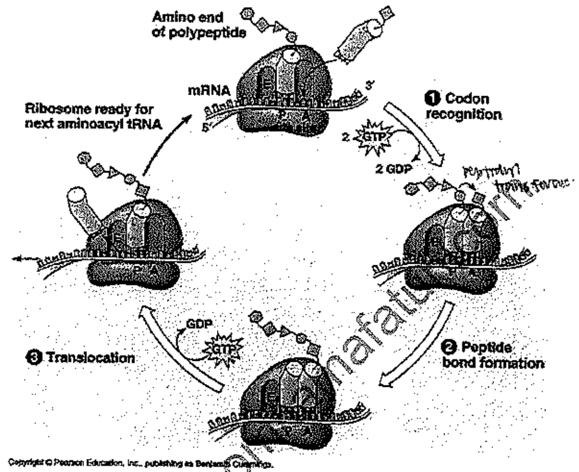


Figure 23. Translocation

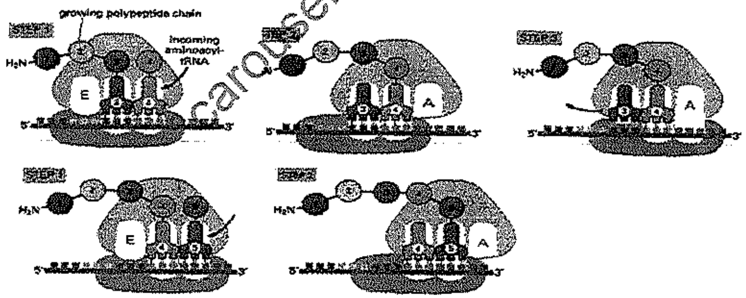


Figure 24. Translating an mRNA molecule. Each amino acid added to the growing end of a polypeptide chain is selected by complementary base-pairing between the anticodon on its attached tRNA molecule and the next codon on the mRNA chain.

Question: *How is fidelity in information transfer maintained?*

[refer to pg 21]

III) TERMINATION

STEP 5: Termination of translation

- Once stop codon reaches the A site, specific proteins called **release factors** enter A site (there are no aminoacyl-tRNA that is complementary to the stop codon).
- Binding of release factors causes **hydrolysis of the bond between the polypeptide chain and the tRNA** in the P site.
- The polypeptide is released from ribosome as it completes its folding to assume the necessary secondary and tertiary protein structures.
- The ribosome disassembles into its subunits.

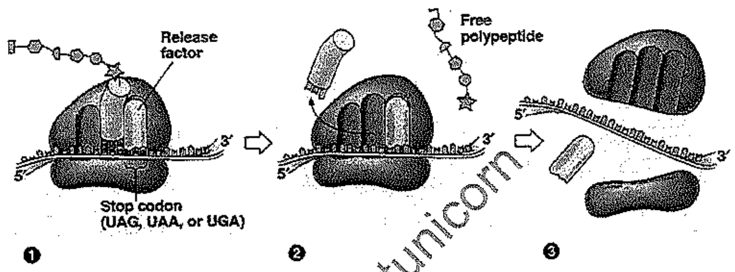
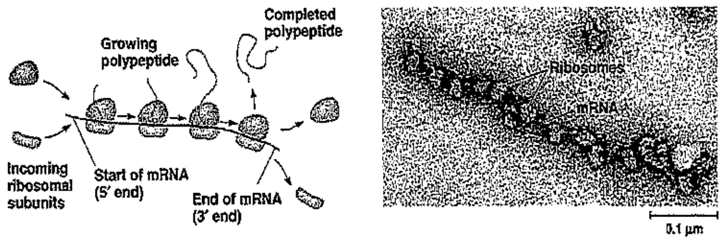


Figure 25. Termination

Two levels of specificity:
1. a tRNA (with a specific anticodon) is matched with a specific amino acid
→ aminoacyl-tRNA synthetase enzyme maintains this specificity

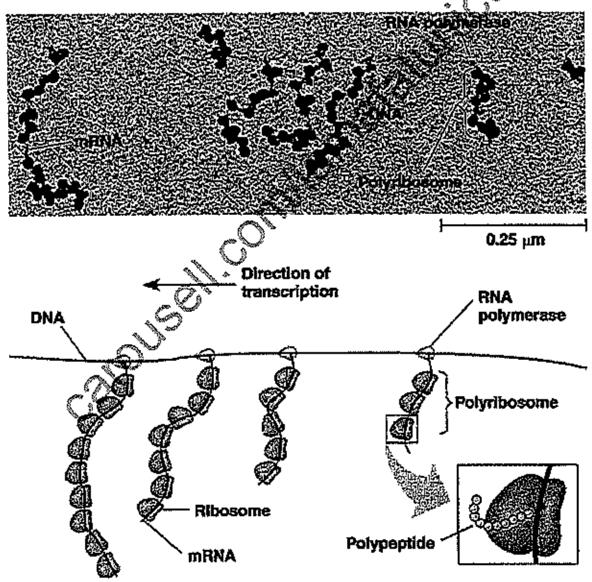
2. Anticodon is matched with a specific codon
→ complementary base pairing maintains this specificity



(a) An mRNA molecule is generally translated simultaneously by several ribosomes in clusters called polyribosomes.

(b) This micrograph shows a large polyribosome in a prokaryotic cell (TEM).

Figure 26a. Polyribosomes → a cluster of ribosomes simultaneously translating an mRNA molecule. Once a ribosome passes the initiation codon, the second ribosome can attach. What is the advantage of having polyribosomes?



Reprinted with permission from O. L. Miller, Jr., B. A. Hamkalo, and C. A. Thomas, Jr. Science 169 (1970):392. Copyright © 1970 American Association for the Advancement of Science. ©1999 Addison Wesley Longman, Inc.

Figure 26b. Simultaneous transcription and translation (of a prokaryotic DNA). The figure shows an electron micrograph and its interpretation. DNA and strings of individual ribosomes (polyribosome) attached to a single mRNA are visible. mRNA, RNA polymerase and polypeptides molecules are barely visible

From RI 2015 Prelim P1 Q10

Question: What protein would be coded by the following DNA molecule?

5' – CCTACTATGCGCCAGTATAAGTGACAATTA – 3'
3' – GGATGATACGCGGTCATATTCACTGTTAAT – 5'

1st letter	Second Letter: U	C	A	G	3rd letter
U	UUU Phe UUC Phe UUA Leu UUG Leu	UCU Ser UCC Ser UCA Ser UCG Ser	UAU Tyr UAC Tyr UAA Stop UAG Stop	UGU Cys UGC Cys UGA Stop UGG Trp	U C A G
C	CUU Leu CUC Leu CUA Leu CUG Leu	CCU Pro CCC Pro CCA Pro CCG Pro	CAU His CAC His CAA Gln CAG Gln	CGU Arg CGC Arg CGA Arg CGG Arg	U C A G
A	AUU Ile AUC Ile AUA Ile AUG Met	ACU Thr ACC Thr ACA Thr ACG Thr	AAU Asn AAC Asn AAA Lys AAG Lys	AGU Ser AGC Ser AGA Arg AGG Arg	U C A G
G	GUU Val GUC Val GUA Val GUG Val	GCU Ala GCC Ala GCA Ala GCG Ala	GAU Asp GAC Asp GAA Glu GAG Glu	GGU Gly GGC Gly GGA Gly GGG Gly	U C A G

Clue:
What is the first codon on the mRNA? 5'–AUG–3' → 3'–TAC–5'
What is the sequence of DNA that codes for the start codon? Determine the template strand

What are the stop codons and how does it look like on the DNA?

Stop codon on mRNA	5'-UAA-3'	5'-UAG-3'	5'-UGA-3'
DNA	3'-ATT-5'	3'-ATC-5'	3'-ACT-5'

3' – GGATGA <u>TAC</u> <u>GCG</u> <u>GTC</u> <u>ATA</u> <u>TTC</u> <u>ACT</u> GTTAAT – 5'

Since ACT is complementary to stop codon, the template strand will code for 5 amino acids.

Since the start codon is 5'-AUG-3' on the mRNA, look for 3'-TAC-5' on the template DNA strand. Once 3'-TAC-5' is located, read the template strand until the complementary triplet-base of the stop codon is located.

Similarities between transcription and replication

- Occur within the nucleus.
- Complementary base pairing occurs.
- Unzipping and rewinding of DNA double helix occurs.
- Separation of parental strands occurs progressively in short segments.
- Formation of phosphodiester bonds.
- Both involve proofreading mechanism for production of accurate nucleic acid chains.

Differences between replication and transcription

Feature	Replication	Transcription
Polymerase involved	DNA polymerase	RNA polymerase
Raw materials	Deoxyribonucleotides dNTP (for DNA), ribonucleotides NTP (for RNA primer)	Ribonucleotides NTP
Template	Both DNA strands serve as a template	Only 1 DNA strand acts as a template
Base pairing	Adenine with thymine, guanine with cytosine	Adenine with uracil, thymine with adenine, guanine with cytosine
Product(s)	2 double stranded DNA molecules synthesised. In each double stranded DNA molecule, one strand is the parental strand, while the other strand is newly synthesised.	1 single stranded RNA molecule is synthesised

State the differences between transcription and translation.

Feature	Transcription	Translation
Location	***Nucleus**** of eukaryotes	***Ribosomes**** in ***cytoplasm****/ at surface of ***rough endoplasmic reticulum****
Transfer of information	***DNA**** template strand to ***mRNA****	***mRNA* to polypeptide****
Enzyme	***RNA polymerase**** links <u>ribonucleotides</u>	***Peptidyl transferase**** on large ribosomal subunit to join <u>amino acids</u>, ***aminoacyl tRNA synthetase**** for <u>amino acid activation</u>
Type of bond between basic units	***Phosphodiester bond**** links ribonucleotides	***Peptide bond**** links amino acids
Direction in which genetic message is read	**<u>3' → 5' of DNA template</u>**	**<u>5' → 3' of mRNA</u>**

3. GENE MUTATIONS

(A) TYPES OF GENE MUTATIONS

- A gene mutation arises as a result of a **change in the sequence of nucleotide bases in the DNA of a gene**.
- A mutation could be caused by a **mutagen** – a chemical or physical agent that interacts with DNA causing a mutation.
  E.g. excessive UV, gamma radiation, carcinogens (cancer causing chemicals) such as tar in cigarettes, viruses. (See 'Cancer')
- A gene mutation may be harmful or beneficial to the organism. (See 'Evolution and Diversity')
- Gene mutation may involve a change in (one nucleotide) **one or a few nucleotide bases**. If it involves a change in just a **single base**, it is called a **point mutation.** (gene mutation ≠ point mutation)
- Types of gene mutation:
  1) **substitution** – occurs when one nucleotide is replaced by another nucleotide
     ATG GCC A → ACG GCC A
  2) **addition (insertion)** – occurs when one or several nucleotides are inserted in a sequence
     ATG GCC A → ATC TGG CCA
  3) **deletion** – occurs when one or several nucleotides are removed from a sequence
     ATG GCC A → ACC A
  4) **inversion** – a segment of nucleotide sequences separates from the allele and rejoins at the original position but it is inverted
     ATG GCC A → ATC CGG A

- **Insertion or deletion** (of 1 or 2 nucleotides)
  - often results in production of a **non-functional protein** as ribosomes begin to **read incorrect triplets** from the point of insertion or deletion. Original **codons** downstream of the point of mutation are not read correctly.
  - such mutations are known as **frame-shift mutations**, and tend to be more severe.
    E.g.:
    Original message (in-frame):
    THE CAT BIT THE RAT

    Insertion of one letter (frame shifted):
    THE CXA TBI TTH ERA T

    Deletion of one letter (frame shifted):
    THE CTB ITT HER AT
    ↳ A

    } Meaning lost beyond point of mutation

Question: *What is the effect of adding 3 consecutive letters to the reading frame here?*
THE **M**CA TBI TTH ERA T (severe disruption)
THE **MA**C ATB ITT HER AT (severe disruption)
THE **MAD** CAT BIT THE RAT → restoration of reading frame (most of it)

➤ Addition of 3 consecutive nucleotides may result in restoration of reading frame.
– Result in changes to primary sequence of polypeptide chain → may / may not result in a loss / modification of function of the polypeptide
– usually less severe than frameshift mutation

Q: which cells must the mutation take place for it to be passed on to the next generation
Germline cells, gametes, cells of gametes

Original mRNA :	AUG GGA AAU UUC CCU AAA UUU CUC CCU AUG	
Resultant amino sequence :	met – gly – asn – phe – pro – lys – phe – leu – pro - met	
Insertion of 1 nucleotide :	AUG GGA CAA UUU CCC UAA AUU UCU CCC UAU G (Insert)	Frame-shift may result in premature chain termination if a STOP codon is generated.
Resultant amino sequence :	met – gly – gln – phe – pro – stop	
Deletion of 1 nucleotide :	AUG GGA AUU UCC CUA AAU UUC UCC CUA UG (Delete)	Amino acid sequence of the polypeptide has been changed and the gene product may no longer be functional.
Resultant amino sequence :	met – gly – asn – phe – leu – asn – phe – ser – leu	

Figure 27. Frame-shift mutations

- **Substitution**
  - is the most common type of gene mutation.
  - is usually not as serious as deletion or addition as the replacement of a nucleotide may not necessarily alter the amino acid sequence in a polypeptide.

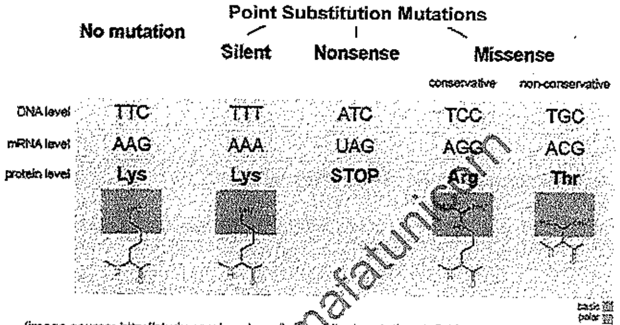
Question: *Can you think of two reasons why?*
Hint 1: 3rd nucleotide of codon….
Hint 2: Reading frame….

– Genetic code is degenerate. more than 1 codon can code for an amino acid.
– reading frame is not affected so severely. severity dependent on amino acid that is substituted. (e.g. when affected amino acid is not found in a critical location such as active site, effect may not be drastic)

Depending on the outcome of the mutation, it could be categorised as such:

- **Frame-shift mutation** → very serious :(
  - As mentioned above, a frame-shift mutation occurs due to mutation involving either insertion or deletion of a number of nucleotides not divisible by 3.
  - Due to the triplet code, these mutations would cause a disruption in the reading frame (the grouping of codons).
  - The result of this mutation would therefore be a completely different polypeptide chain, which is typically non-functional
- **Silent mutation**
  - A **point mutation** that **does NOT lead to a change in the amino acid sequence** in the polypeptide chain.
  - This is possible due to:
    1. Point mutation occurring within the **coding sequence** of a gene:
       ➤ Change in nucleotide base, led to a change in codon which codes for the **same amino acid**.
       ➤ Genetic code is **degenerate**, where more than one codon codes for the same amino acid.
    2. Point mutation occurring at non-coding regions (e.g. introns, outside of genes) therefore it does not affect the protein product.

- **Missense mutation**
  - A **point mutation**, in which a single nucleotide change results in a codon on the mRNA that codes for a **different amino acid**.
  - If the resultant amino acid has **similar biochemical properties** (e.g. charge, hydrophobicity and size) as the amino acid replaced, the mutation is considered a **conservative mutation**.
  - If the resultant amino acid has different biochemical properties as the amino acid replaced, the mutation is termed non-conservative mutation.
- **Nonsense mutation**
  - A **point mutation** which has resulted in a **premature stop codon (UAG, UAA, UGA)**, causing the polypeptide formed to be **truncated** and typically non functional.



(image source: http://study.com/academy/lesson/silent-mutation-definition-example-quiz.html)

Figure 28. Summary of the possible effects of point substitution mutations

(B) EXAMPLE: SICKLE-CELL ANAEMIA

- An example of a **point mutation**. In sickle-cell anaemia, it is a **substitution mutation**.
- In a normal adult:
  - Normal adult haemoglobin (Hb A) is a quaternary protein which is a **tetramer composed of 2 different types of polypeptide chains** (2 α-globin chains and 2 β-globin chains).
  - The α and β chains are coded for by **2 different genes** found on 2 different chromosomes.
  - Haemoglobin is abundant in **red blood cells** (RBC) and serves a role in transporting oxygen from lungs to tissue.

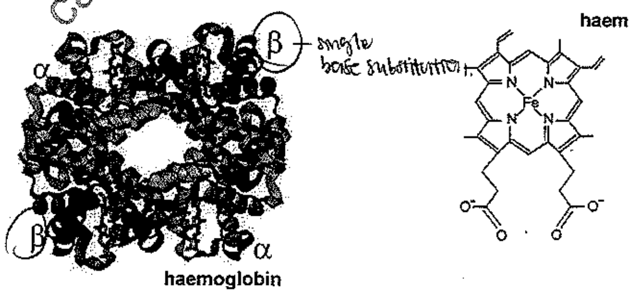


Figure 29. Haemoglobin and the haem group with Fe^{3+} that binds

- In **sickle cell anaemia**:

  Gene mutation - Change in sequence of DNA nucleotide
  - A **single base substitution** occurred in the gene which codes for **β-globin** chain
    → CTC becomes CAC (i.e. T is substituted by A) on template strand

  Results in change in sequence of amino acids in polypeptide chain
  - In the mRNA produced, at the **6th triplet codon**
    → GAG becomes GUG
    → previous codon coding for amino acid **glutamate** now codes for **valine**
    → forms a **sickle cell haemoglobin (Hb S)**

  Results in change in properties of haemoglobin and phenotype of RBC
  - Previous amino acid glutamate is **charged and hydrophilic** now valine is **non-polar and hydrophobic**.
  - Normal circular biconcave shape of red blood cell becomes sickle shape under low $[O_2]$
    → as the red blood cell moves to a region of **low oxygen concentrations** in actively respiring tissue, oxygen is released by abnormal Hb S and an unusual **conformational change** (the hydrophobic areas on different molecules would stick together) will occur.
    → This conformational change will cause Hb S to **polymerise/crystallise** into **abnormal rigid rod-like fibres** that distort biconcave shape of red blood cell, resulting in the characteristic **sickle shape** of the cell.
  - The polymers can be broken up by binding oxygen to Hb S. i.e. Hb S can still bind O_2! Sickling is reversible!

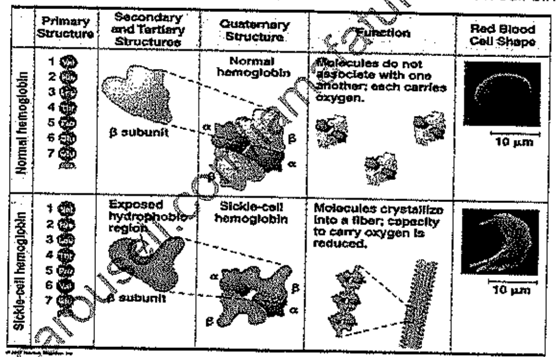


Figure 30. Differences in protein structure in Hb A and Hb S molecules.

In summary:

	Normal (Hb A)	Sickle-cell anaemia (Hb S)
Gene coding for β-globin chain	_____CTC_____	_____CAC_____
Resultant codon on mRNA	… … GAG…	… … GUG …
Resultant polypeptide chain	… **glutamate** …	… **valine** …
At low O_2 concentrations	Remain soluble	Crystallise into rigid rod-like fibres
Appearance of red blood cell at low $[O_2]$	Disc-shaped	Sickle-shaped

Effects of the disease:

(1) Sickle red blood cells are more **fragile** causing them to break up more easily. It is also actively destroyed in the spleen. This results in the **shortage of red blood cells** and **poor oxygen transport**.

- Patient suffers from anaemia, breathlessness and physical weakness
- Heart failure as the heart needs to work a lot harder
- Lack of energy in the form of ATP due to reduced cellular respiration

(2) Sickle-shaped red blood cells, being pointed and elongated, may also get **lodged in small blood vessels** (capillaries) and therefore interfere with blood circulation → sickle-cell crisis.
- Depriving organs e.g. brain of oxygen
- Severe pain due to many localised blockages resulting in death of surrounding tissue
- **Organ damage** occurs especially those with numerous fine capillaries such as spleen and lungs

Inheritance of the mutation:
- Sickle cell anaemia is a **homozygous recessive disorder.**
  - → sufferers would have **two copies** of the mutant form of the gene (HbS HbS). They are said to have the **sickle cell disease.**
  - → carriers/ **heterozygous individuals** have one copy of the mutant and **one copy of the normal form of the gene (HbA HbS).** They are said to have the **sickle cell trait.** (See 'Evolution and Diversity')

EXAMPLE: CYSTIC FIBROSIS (for your info only)

Genetics of cystic fibrosis
- Commonly afflicts Caucasians.
- **Cystic fibrosis gene** normally codes for a membrane ion channel called the **Cystic Fibrosis Transmembrane Conductance Regulator (CFTR).** CFTR controls movement of **chloride ions,** Cl^- into or out of cells (& influences Na^+ transport indirectly).
- More than 1000 cystic fibrosis mutations have been reported.
  - The most common mutation which occurs in 75% of cystic fibrosis patients, involves a **deletion of 3 nucleotides** on **exon 10 of chromosome 7, (autosomal)** resulting in the loss of an amino acid **phenylalanine** at position 508 of the CFTR polypeptide.
  - In affected patients, **cystic fibrosis transmembrane conductance regulator (CFTR)** are either **missing or defective.**

Symptoms of cystic fibrosis
1. In the lungs, defective/missing CFTR
   → Cl^- not transported out of epithelial cells into lumen of air cavity
   → Na^+ retained too (charge balance)
   → Ψ more –ve in cell → water retained in cell
   → mucus in lumen undiluted → thick → cannot flow → congestion
   → mucus remains too long in respiratory tract
   → bacteria growth → lung infection
   → **severe breathing difficulty**

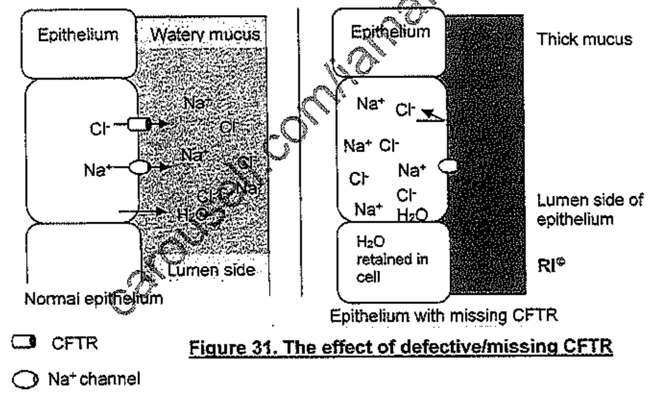


Figure 31. The effect of defective/missing CFTR

2. Pancreatic duct is choked by thick mucus preventing release of enzymes
   → **indigestion**
3. Thick mucus layer in intestines
   → **reduces absorption** of digested food
4. Sweat gland produces sweat
   → as it rises up the duct towards pore, upper duct reabsorbs Na^+ and Cl^- as opposed to secreting NaCl into lumen of lungs due to opposite orientation of CFTR in the sweat ducts.
   → defective CFTR results in no reabsorption of NaCl
   → **very salty and copious sweat** production
   → basis of diagnosis → measure [Cl^-] in sweat.
5. Death usually occurs by age 30

KEYWORDS include…

Transcription
RNA
Template strand
Hydrogen bonding
Complementary base pairing
RNA polymerase
Promoter
Phosphodiester bond

Translation
Ribosome
Large subunit of ribosome:
Aminoacyl site
Peptidyl site
Exit site
Peptidyl transferase
Peptide bonds
Polypeptide

Codons
Anticodons
3' CCA stem
Amino acid activation
Aminoacyl tRNA
Aminoacyl tRNA synthetase
Methionine

Mutations
Frameshift mutation
Silent mutation
Missense mutation
Conservative mutation
Nonsense mutation

Sickle cell anaemia
Genetic code
(details needed)

4. LINKS

The information from this lecture notes on protein synthesis and gene mutations will be relevant to the following topics:

S/N	Topic	Comments
1.	Proteins and enzymes	The type of protein/enzymes that an organism has depends on the presence of genes that code for them. The genes determine the primary structure of proteins / enzymes which determine their 3D conformation and hence their function.
2.	Mendelian Genetics	Genes and their alleles / genotype are located on DNA that codes for proteins that determine particular traits/characteristics. Gene products may interact with other DNA seq / gene product and affect their expression (epistasis)
3.	Photosynthesis & Respiration	DNA codes for the proteins (e.g. enzymes, electron carriers) involved in the processes. Any mutation in the proteins involved may greatly reduce the rate of the processes.
4.	Organisation and Control of the Prokaryotic and Eukaryotic Genome	Control of gene expression at a transcriptional and translational level in prokaryotes and eukaryotes.
5.	Genetics of Viruses	Expression of viral components in the host cell.
6.	Evolution and Diversity (Heterozygote advantage)	Gene mutations – Sickle cell anaemia/ trait

A. CORE SYLLABUS
(3) GENETICS OF VIRUSES AND BACTERIA

VIRUSES

Content
- The genetics of viruses

Learning Outcomes

Candidates should be able to:
1f. Discuss how viruses challenge the cell theory and concepts of what is considered living.
1e. Describe the structural components of viruses, including enveloped viruses and bacteriophages, and interpret drawings and photographs of them.

2d. Describe the structure and organisation of viral, prokaryotic and eukaryotic genomes (including DNA/RNA, single-/double-stranded, number of nucleotides, packing of DNA, linearity/circularity and presence/absence of introns). (Include parts relevant to viral genomes here)
e. Describe how the genomes of viruses are inherited through outlining the reproductive cycles of:
i. bacteriophages that reproduce via lytic cycle only, e.g. T4 phage;
ii. Bacteriophages that reproduce via lytic and lysogenic cycles, e.g. Lambda phage;
iii. enveloped viruses, e.g. influenza
iv. retroviruses, e.g. HIV.
f. Describe how variation in viral genomes arises, including antigenic shift and antigenic drift.

References
Main reference:
- Biology, 9th Edition, by Campbell, N.A. and Reece J.B. (2011).
- Microbiology – An Introduction, 9th Edition by Tortora, Funke and Case. (2007).
- Molecular Biology of the Cell, 3rd Edition by Alberts, Bray and Lewis. (2008).

TABLE OF CONTENTS

***This handout is the effort of several Biology teachers at RI. It has and will continue to be updated.**

(A) Introduction to Viruses

1. Brief History of Virology

1796:	Edward Jenner used cowpox to vaccinate against smallpox. Jenner was the first person to deliberately vaccinate against any infectious disease. Although Jenner is commonly given the credit for vaccination, variolation, the practice of deliberately infecting people with smallpox to protect them from the worst type of the disease, had been practiced in China at least two thousand years previously.
1885:	Louis Pasteur experimented with rabies vaccination, using the term "virus" (Latin, poison) to describe the agent. Although Pasteur did not discriminate between viruses and other infectious agents, he originated the terms "virus" and "vaccination" (in honour of Jenner) and developed the scientific basis for Jenner's experimental approach to vaccination.
1892:	Dimitri Iwanowski described the first "filterable" infectious agent - tobacco mosaic virus (TMV) - smaller than any known bacteria. Iwanowski was the first person to discriminate between viruses and other infectious agents, although he was not fully aware of the significance of this finding.
1898:	Martinus Beijerinick extended Iwanowski's work with TMV and formed the first clear concept of the virus "contagium vivum fluidum" - soluble living germ. Freidrich Loeffler and Paul Frosch demonstrated that foot and mouth disease is caused by such "filterable" agents. Loeffler and Frosch were the first to prove that viruses could infect animals as well as plants.
1915:	Frederick Twort discovered viruses infecting bacteria.
1917:	Felix d'Herelle independently discovered viruses of bacteria and coins the term bacteriophage. The discovery of bacteriophages provided an invaluable opportunity to study virus replication at a time prior to the development of tissue culture when the only way to study viruses was by infecting whole organisms.
1935:	Wendell Stanley crystallized TMV and showed that it remained infectious (Nobel Prize, 1946). Stanley's work was the first step towards describing the molecular structure of any virus and helped to further illuminate the nature of viruses.
1939:	Emory Ellis and Max Delbruck established the concept of the "one step virus growth cycle" essential to the understanding of virus replication (Nobel Prize, 1969). This work laid the basis for the understanding of virus replication - that viral particles do not "grow" but are instead assembled from preformed components.
1940:	Helmuth Ruska used an electron microscope to take the first pictures of viral particles. Along with other physical studies of viruses, direct visualization of virions was an important advance in understanding virus structure.
1961:	Sydney Brenner, Francois Jacob, and Matthew Meselson demonstrated that bacteriophage T4 uses host cell ribosomes to direct virus protein synthesis. This discovery revealed the fundamental molecular mechanism of protein translation.
1970:	Howard Temin and David Baltimore independently discovered reverse transcriptase in retroviruses (Nobel Prize, 1975). The discovery of reverse transcription established a pathway for genetic information flow from RNA to DNA, refuting the so-call "central dogma" of molecular biology.
1970s:	Advances in molecular biology techniques have led to the discovery of many new viruses.
1979:	Smallpox was officially declared to be eradicated by the World Health Organization (WHO). The last naturally occurring case of smallpox was seen in Somalia in 1977.
1983:	Luc Montaigner and Robert Gallo announced the discovery of human immunodeficiency virus (HIV), the causative agent of AIDS.
1990:	First (approved) human gene therapy procedure was carried out on a child with severe combined immune deficiency (SCID), using a retrovirus vector.
1999:	Number of confirmed cases of people living with HIV/AIDS worldwide reaches 33 million. The AIDS pandemic continues to grow. Nucleotide sequence of the largest known virus genome completed: *Paramecium bursaria* Chlorella virus 1. This 330,742 bp sequence represents the technical advances in sequencing which have occurred since the first genome sequence was completed in 1977.
2002-2004:	Outbreak of severe acute respiratory syndrome (SARS) caused by the SARS coronavirus.
2006:	Two vaccines protecting against several cervical cancer-causing strains of human papillomavirus (HPV) were released.
2009 - 2010:	Outbreak of H1N1 influenza pandemic. Two babies from Mexico die in the outbreak. By end of May 2010, steep decline in number of reported cases. WHO statistics report more than 18 000 cases of death from H1N1.
2013 - 2014:	Outbreak of Ebola epidemic. Started in West Africa. At present, 24 000 confirmed cases and the number of deaths have exceeded 9800.
2015 - 2016	Zika outbreak in South America linked to microcephaly.

LO 1f Discuss how viruses challenge the cell theory and concepts of what is considered living.

2. Are Viruses Living or Non-living?
Unlike bacteria, viruses lack the ability to replicate on their own. They absolutely require a living host, or a cell to support their metabolism and replication. They are able to enter a cell and then take over that cell, directing it to make more viral particles. In this sense, viruses are said to be obligate parasites.

Question:
What is an obligate parasite?
An organism that cannot live independently of its host & depends on its host to complete its life cycle.

2.1. Characteristics of living organisms based on the cell theory:

1. Have a cellular organization
   Regardless of whether they are unicellular or multi-cellular organisms, the smallest level of organization of living organisms is the cell.
2. Show metabolic activity
   All organisms need to acquire and use energy in order to maintain metabolic processes for survival.
3. Grow and develop
   Growth involves both the increase in size and number of cells. When organisms grow, they undergo changes known as development.
4. Reproduce
   Organisms produce offspring like themselves through sexual or asexual means.
5. Have a common hereditary molecule
   All living organisms have a common molecular inheritance based on the nucleic acid, which contains instructions for the structure and function of cells. (DNA → RNA)
6. Respond to stimuli
   Organisms have specialised receptors that detect environmental stimuli to allow their cells to adjust metabolism in response.
7. Adapt to the environment
   Adaptation is the accommodation of a living organism to its environment, which is fundamental to the process of evolution. in isolation: – have common hereditary material (5) in host cell: – metabolic activity

Question:
Of the above, viruses only possess one characteristic. Identify this characteristic.

There is great contention amongst scientists as to whether viruses should be classified as being living or non-living. They have been described as "organisms at the edge of life".

Q1 Explain why viruses may be regarded as living organisms.

1) Viruses **possess genetic material** and are capable of propagating their genetic information.
2) Once inside the host cell it **uses/directs host enzymes** to **carry out metabolic processes.**
3) Once inside the host cell, and **directs its own self-replication.** They **reproduce by creating multiple copies of themselves** through self-assembly.
4) Some viruses undergo **mutation and reassortment** of their genetic material during their replication giving rise to new viral strains.
5) Virus can react to its environment. They are able to **respond to stimuli** such as heat, radiation and chemical when inside the host cell.
6) They are able to **evolve** to adapt to new environment

Q2 Explain why viruses may be regarded as non-living organisms.

1. They are **acellular** and **lack cellular organelles**
2. Viruses **do not carry out metabolism** (e.g. respiration)
3. They **lack the ability to reproduce on their own independently** and can only undergo replication in living cell
4. They **do not grow** and undergo developmental changes and require a host cell to make new products such as coat protein and nucleic acids
5. They **do not respond to stimuli** when outside the host cell.
6. They can **only evolve** by natural selection **within a host cell.**

Conclusion:

Viruses are ***obligate parasites**** which requires a living host to support many of their functions.
Viruses can be crystallized like an ordinary chemical as the capsid coat is very regular.

(B) STRUCTURE OF VIRUSES

The sizes of viruses ranged from 10 to 300nm. An intact **infectious** viral particle is called a **virion.** A virion is a complete, fully developed viral particle composed of the **genome** and surrounded by a protein coat, the **capsid.**

Some types of viruses have an additional outer layer, an **envelope** in which viral **glycoproteins** are embedded in. → link to exocytosis, etc.

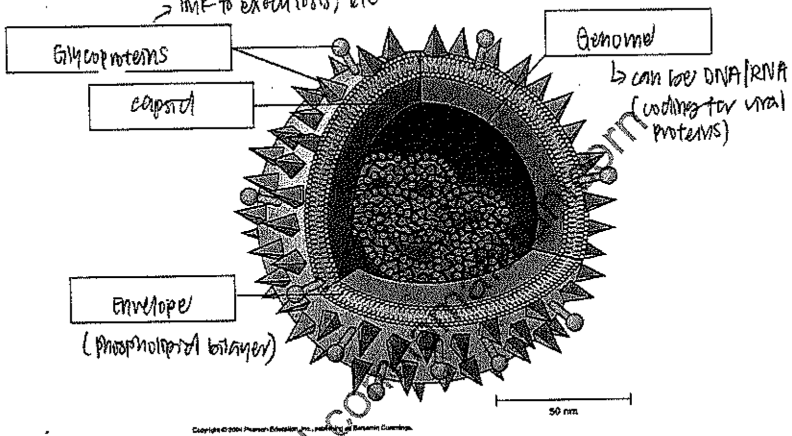


Fig. 1 Diagram of a generalized virus

LO 2d Describe the structure and organisation of viral (including DNA/RNA, single-/double-stranded, packing of DNA, linearity/circularity).

1e Describe the structural components of viruses, including enveloped viruses and bacteriophages, and interpret drawings and photographs of them.

1. Genome

- The viral genome is **single** or **several/segmented**, **circular** or **linear** molecules of nucleic acid that functions as the genetic material of the virus.
- In contrast to prokaryotic and eukaryotic cells in which DNA is usually the primary genetic material, a virus can have **either DNA** or **RNA** but **never both.**
- The nucleic acid can be **single-stranded** or **double-stranded**. Thus, in addition to the familiar double-stranded DNA, there are viruses with single-stranded DNA, double stranded RNA, or single-stranded RNA.
- The genome codes for the synthesis of **viral components** and **viral enzymes** for **replication** and **assembly** of a virion.

viral genome → RNA (ds, ss), DNA (ds, ss); linear / linear / circular; multiple / singular; linear, circular; singular, multiple

- The genes are few in number ranging from 3 – 100 depending on the class of virus. The number of genes present in the virus determines the degree of complexity displayed by the virus.

Table 19.1 Classes of Animal Viruses

Class/Family	Envelope	Examples That Cause Human Diseases
I. Double-Stranded DNA (dsDNA)		
Adenovirus (see Figure 19.3b)	No	Respiratory viruses; tumor-causing viruses
Papovavirus	No	Papillomavirus (warts, cervical cancer); polyomavirus (tumors)
Herpesvirus	Yes	Herpes simplex I and II (cold sores, genital sores); varicella zoster (shingles, chicken pox); Epstein-Barr virus (mononucleosis, Burkitt's lymphoma)
Poxvirus	Yes	Smallpox virus; cowpox virus
II. Single-Stranded DNA (ssDNA)		
Parvovirus	No	B19 parvovirus (mild rash)
III. Double-Stranded RNA (dsRNA)		
Reovirus	No	Rotavirus (diarrhea); Colorado tick fever virus
IV. Single-Stranded RNA (ssRNA); Serves as mRNA		
Picornavirus	No	Rhinovirus (common cold); poliovirus; hepatitis A virus; other enteric (intestinal) viruses
Coronavirus	Yes	Severe acute respiratory syndrome (SARS)
Flavivirus	Yes	Yellow fever virus; West Nile virus; hepatitis C virus
Togavirus	Yes	Rubella virus; equine encephalitis viruses
V. ssRNA; Template for mRNA Synthesis		
Filovirus	Yes	Ebola virus (hemorrhagic fever)
Orthomyxovirus (see Figures 19.3c and 19.9a)	Yes	Influenza virus
Paramyxovirus	Yes	Measles virus; mumps virus
Rhabdovirus	Yes	Rabies virus
VI. ssRNA; Template for DNA Synthesis		
Retrovirus (see Figure 19.8)	Yes	Human immunodeficiency virus (HIV/AIDS); RNA tumor viruses (leukemia)

Table 1 : Classification of animal viruses base on the types of genome

2. Capsid

- The genome is surrounded by a protein coat called a **capsid**. The structure of the **capsid** is ultimately determined by the viral genome and accounts for most of the mass of a virus, especially of the small ones.
- Each **capsid** is composed of protein subunits called **capsomeres.**
- The **capsid** serves to **protect**, **attach** and **introduce the genome into host cells.**
- Together, the **capsid** and the viral nucleic acid form the **nucleocapsid**. An infectious viral particle, a virion, will contain at minimum a nucleocapsid.

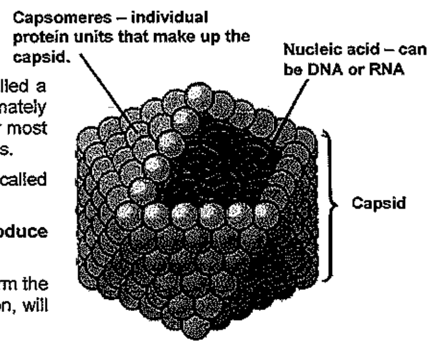


Fig. 2 Diagram of a viral capsid surrounding its genome (nucleocapsid).

3. Envelope

- Some types of viruses have an additional outer layer, an **envelope** that surrounds the nucleocapsid.
- The envelope is composed of **phospholipids** and **glycoproteins** that are arranged to form a **lipid bilayer**. For most viruses, it is derived from the **host cell membranes** by a process called **budding**. The envelope may come from the host cell's nuclear, vacuolar or plasma membranes.
- Although the envelope is usually of host cell origin, the virus does incorporate proteins of its own, often appearing as **glycoprotein spikes**, into the envelope. (made by ribosomes on RER)
- Most animal viruses have an envelope surrounding their nucleocapsid.
- Viruses that are composed of just the nucleocapsid are called **naked viruses** or **non-enveloped viruses** whilst those viruses whose nucleocapsids are covered by an envelope are termed **enveloped viruses**.

4. General Morphology

- Viruses may be classified into several different types based on the
  - shape of the viruses
  - type and structure of the genome,
  - the presence or absence of a viral envelope and
  - mode of replication.
- In general, four main morphological virus types can be identified: helical, icosahedral, enveloped and complex viruses

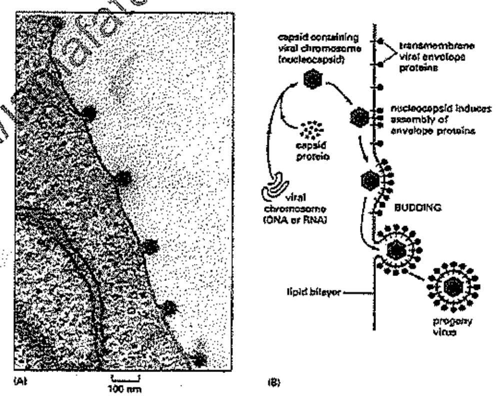


Fig. 3(A) Newly replicated viruses budding off from host cell.
Fig. 3(B) Assembly of viral components at the cell surface membrane of the host cell before budding off.

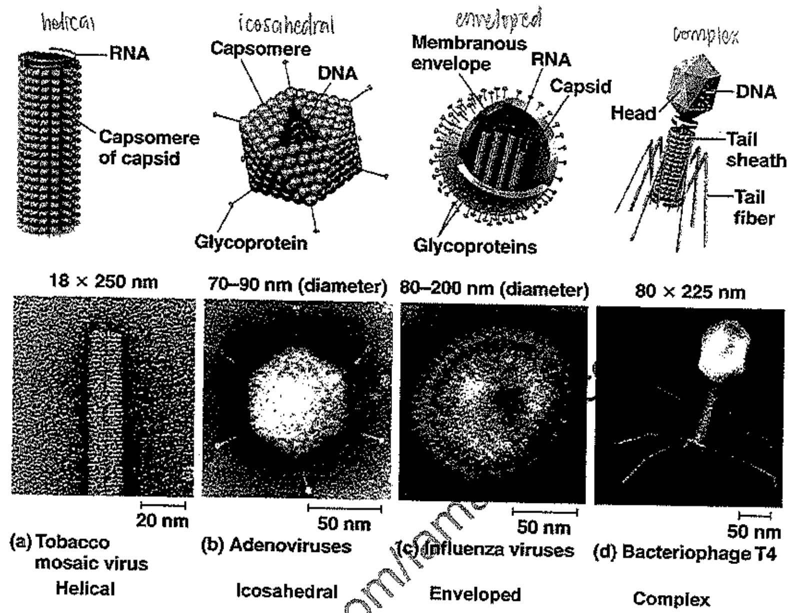


Fig. 4. Different morphological structures of viruses

Question:
List 3 definitive characteristics of a virus.

1. They are totally dependent on a host cell for replication (They are obligate parasites)
2. They contain only one type of nucleic acid: DNA/RNA but never both
3. Viral components must assemble into complete viruses (virions) to go from one cell to another.

(C) Viral Replication

How does a virus replicate

(1) Viral replication begins with the virus invading the host cell and taking over the host's metabolic machinery.
Virus has **specific** host that it infects and **recognises** its host cell via **host cell antigen** e.g. **glycoprotein** at cell surface membrane.
The **genetic material** is **injected** into the host cell or the **entire virus may enter** and disassemble inside the host cell to free the genetic material.

(2) Virus uses the host metabolism and machinery to **synthesise its nucleic acid** such as a DNA virus hydrolyses the DNA of host cell and uses the nucleotides and the host DNA polymerase to replicate its DNA.

(3) Virus uses the host cell RNA polymerase to **transcribe its genes and ribosome to translate its mRNA to viral coat proteins and enzymes**. Also supplied by the host cells are tRNA, nucleotides, amino acids and ATP.
It contains only a few genes which code for the viral structural components like the capsid proteins and viral enzymes that are involved in the viral life cycle.

(4) Once all the viral components are synthesised, it will **self-assemble** to form new viral particles or virion and **exit** the cell via **budding, exocytosis** or through **lysis** of the host cell.

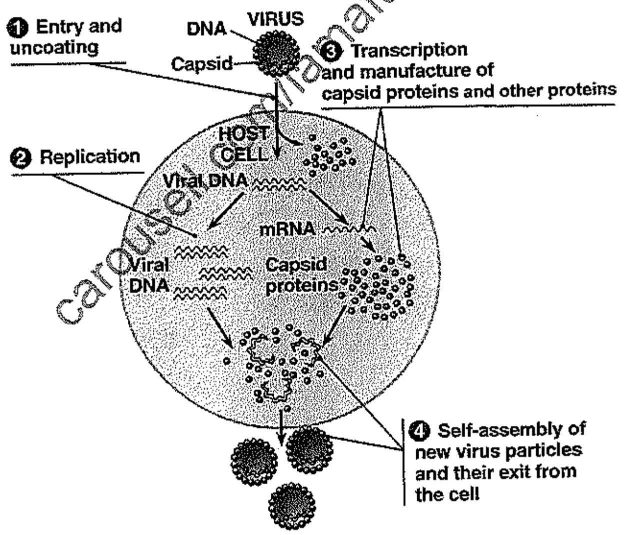


Fig. 5. A simplified viral reproductive cycle

1. Bacteriophages

- Bacteriophages are **viruses that only infect bacteria.**
- There are two primary types of bacteriophages: (cause cell to lyse)
  - **Lytic bacteriophages**; and e.g. T4 phage
  - **Temperate bacteriophages.** e.g. Lambda phage
- Bacteriophages that replicate through the **lytic life cycle** are called **lytic bacteriophages.** They lyse (*break up or disintegrate*) the host bacterium as a normal part of their life cycle.
- Bacteriophages capable of a **lysogenic life cycle** (does not kill host) are termed **temperate phages.** When a temperate phage infects a bacterium it can either replicate by means of the lytic life cycle (and cause **lysis** of the host bacterium) or it can incorporate its DNA into the bacterium's DNA and become a **prophage.**

LO 2e (i) Describe how the genomes of viruses are inherited through outlining the reproductive cycles of bacteriophages that reproduce via lytic cycle only, eg. T4 phage

1.1 T4 Bacteriophage

Structure

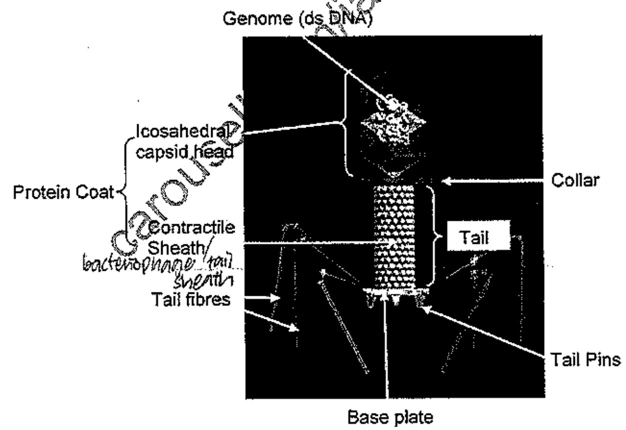


Fig. 5. Structure of T4 bacteriophage

Lytic Cycle

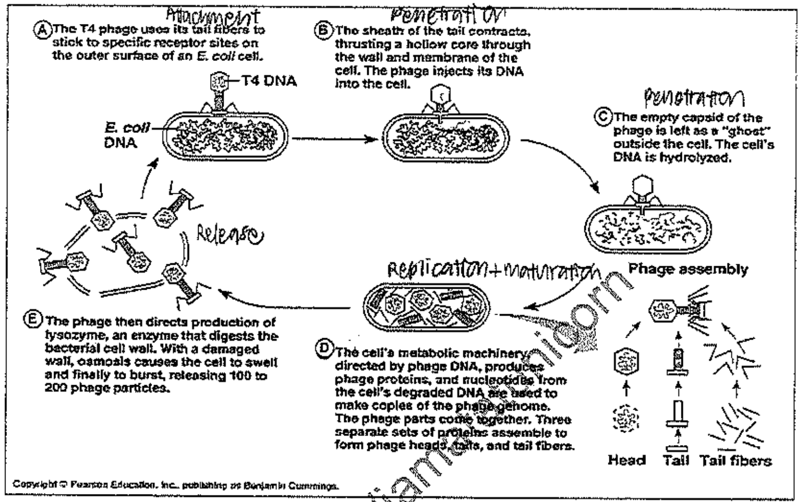
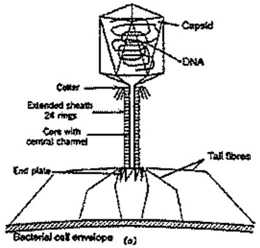


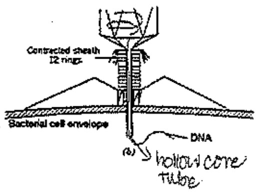
Fig. 6. Lytic life cycle of T4 bacteriophage

A. Attachment



- Attachment sites on the tail fibres recognise and attach or adsorb to complementary receptor sites on the bacterial surface. This attachment is a chemical interaction in which weak bonds are formed between the attachment and receptor sites.
- Specific strains of bacteriophages can only adsorb to specific strains of host bacteria. This is known as **viral specificity**.
- Although most bacteriophages attach to the bacterial cell wall, some are able to attach to the flagella or pili.

B. Penetration

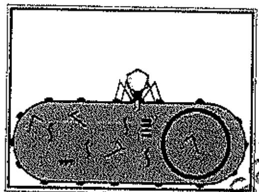


- The bacteriophage tail releases an enzyme, **phage lysozyme**, which **digest the bacterial cell wall**, allowing molecules to be released. When these molecules reach the virus, they trigger a change in the shape of the base plate. This initiates a **contraction of the bacteriophage tail sheath, thrusting the hollow core tube through the cell wall.**
- When the tip of the core reaches the plasma membrane, **DNA** from the bacteriophage is **injected** into the bacterial cell. The empty **capsid** remains outside the cell.

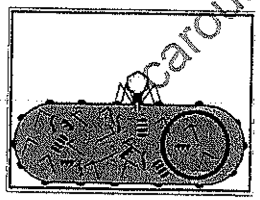
C. Replication

enzymes

- Inside the cell, the bacteriophage DNA is **immediately transcribed** to synthesise **messenger RNA** using the **host RNA polymerase.**
- Phages that are highly virulent (deadly) produce early proteins that completely take control from the host cell. For example, host cell DNA is degraded within minutes into nucleotides that are later reused to synthesise viral DNA. Viral DNA escapes degradation because of methylation of its DNA.
- Enzymes coded by the phage genome takes over the bacterium's macromolecular (protein, RNA, DNA) synthesising machinery for its own use.
- The phage uses the host cell's nucleotides and several of its own enzymes to synthesise many copies of phage DNA.
- Soon after, biosynthesis of viral proteins begins. It uses the bacterium's metabolic machinery to synthesise phage enzymes and phage structural components.
- For several minutes following infection, complete phages cannot be found in the host cell. Only separate components like the DNA and protein are present. In this period, infective virions are not yet present. This is known as the **eclipse period.** (infective virions are not yet present)



Early Replication
The bacteriophage genome replicates and bacteriophage components begin to be produced by way of the host bacterium's metabolic machinery.



Late Replication
The production of bacteriophage components and enzymes progresses.

D. Maturation

- Bacteriophage DNA and **capsid** are assembled into a DNA-filled head (nucleocapsid (?)). The head, tail and tail fibres are then assembled independently and joined with each other in a specific sequence. First, the tail fibres join with the tail, then a DNA-filled head attaches to the tail.

E. Release

- The final stage of viral multiplication is the release of virions from the host cell. The term **lysis** is generally used for this stage in the multiplication of T4 bacteriophages because the plasma membrane of the host cell actually breaks open (lyses).
- **Lysozyme**, which is **coded for by a phage gene**, is synthesised within the cell. This enzyme causes the **bacterial cell wall to break down.**
- The newly produced bacteriophages are released from the host cell. The mature phage particles will infect other susceptible cells in the vicinity and the viral multiplication cycle is repeated within those cells.

LO 2e (ii) Describe the reproductive cycle of bacteriophages that reproduce via a lysogenic cycle, e.g. lambda phage.

1.2 Lambda bacteriophage (λ bacteriophage)

Structure

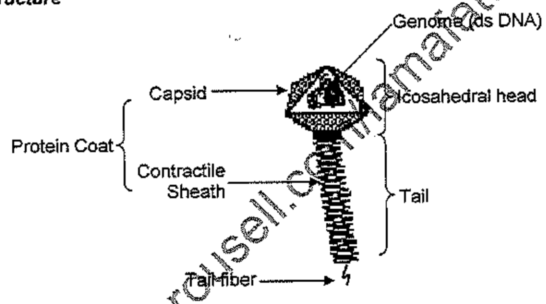


Fig. 7. Structure of lambda bacteriophage

The Lytic and Lysogenic Life Cycle of Lambda bacteriophage

Bacteriophages capable of a lysogenic life cycle are termed **temperate phages.** When a temperate phage infects a bacterium, it can either:

- **replicate by means of the lytic life cycle** and cause lysis of the host bacterium, or,
- it can **incorporate its DNA into the bacterium's DNA** and become a non-infectious **prophage.**

Lysogenic Life Cycle

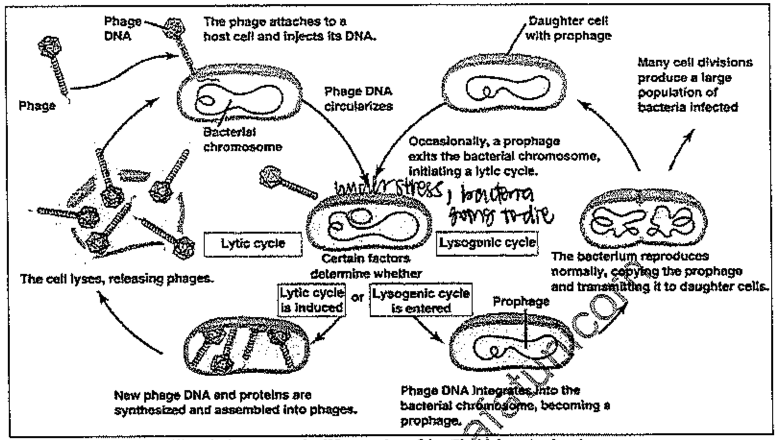


Fig. 8. Lysogenic life cycle of lambda bacteriophage

A. Attachment

- **Tail fiber adsorb** to **complementary receptor site** on host bacterium cell wall *(same as T4 phages).*

B. Penetration

- **Sheath of tail contracts** and drives a hollow tube through the bacteria cell wall.
- Phage **genome enters** the bacterium *(same as T4 phages).*

C. Replication

- Upon penetration into the bacteria cell, the originally linear phage DNA forms a **circle.**
- This **circular DNA can multiply** and be transcribed leading to the production of new phage and to cell lysis (via the lytic cycle).
- Alternatively, the **circular DNA can integrate** into and become part of the circular bacterial DNA (the lysogenic cycle).
- The inserted phage DNA is now called a **prophage.**

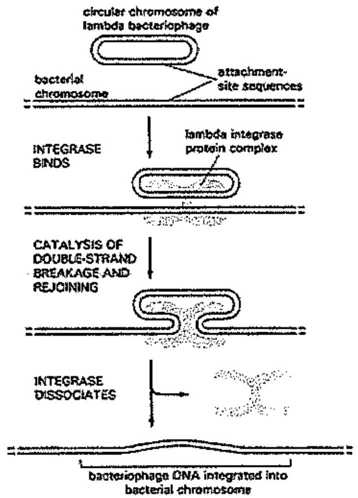


Fig. 9. Integration of lambda genome into host cell genome.

- Most of the **prophage genes are repressed** by **repressor proteins** that are the products of phage genes. These repressors **stop transcription of all the other phage genes**. Thus, the phage genes that would otherwise direct the synthesis and release of new virions are turned off.
- Every time the host cell's machinery replicates the bacterial chromosome, it also replicates the prophage DNA. The prophage will be found in all progeny cells, where it remains **latent**.

D. Spontaneous induction

- Occurs in one of every million to every billion bacteria containing a prophage.
- Induction occurs spontaneously but its frequency is enhanced by irradiation with ultraviolet light or exposure to agents that damage DNA. It activates the cellular proteases.
- Under these conditions, the repressor protein is destroyed by increased protease activity.
- The prophage is no longer repressed but is instead excised and **enters the lytic cycle**.

E. Maturation

- Since the phage genome is no longer repressed, phage components are produced using the host bacterium's metabolic machinery.
- More copies of viral genome are produced by DNA replication using host cell machinery.
- The bacteriophage components then assemble into complete virions. *(same as T4 phages)*

F. Release

- The complete virions are then released from the host cell in the same manner as the lytic cycle. *(same as T4 phages)*

Question:
What are the possible defense mechanisms of bacteria against phages?

1. mutant bacteria with receptor sites that are no longer complementing to phage attachment sites
2. develop restriction enzymes which recognise foreign phage viral DNA and cleave them. modify bacteria's own DNA to prevent attack by restriction enzymes. (can methylate its own DNA)
3. develop a lysogenic relationship with the phage.

- bacteria DNA → methylated
- viral DNA → not methylated

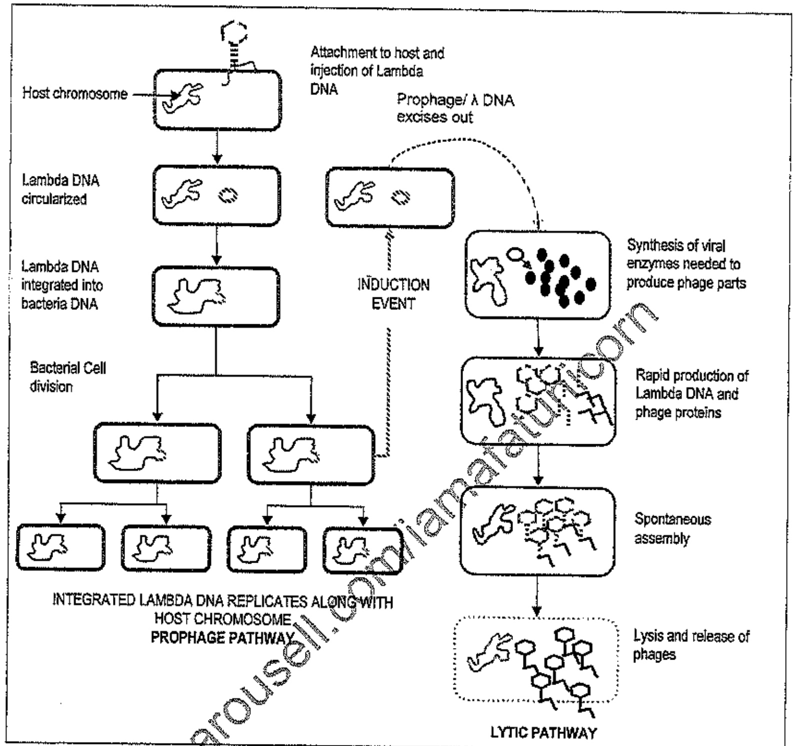


Fig. 10. Summary of the lambda bacteriophage reproductive pathway

Lysogenic host cells are immune to re-infection by the same bacteriophage. However, other types of bacteriophage may still infect the host cell.

Lysogenic bacterial cells may exhibit phage conversion. That is, the host cell may exhibit new properties following integration of the **prophage** into the host genome. For example, the bacterium *Corynebacterium diphtheriae*, which causes diphtheria by means of a toxin it produces can only produce this toxin when it is carrying a lysogenic phage because the **prophage** carries the gene coding for the toxin.

Lysogenic bacterial cells are capable of specialised transduction in which the lysogenic phage packages adjacent genes of bacterial DNA along with its own viral DNA in the same **capsid**. These genes are then transferred to a new bacterial cell along with the **prophage** when the virion infects a new cell. *(To be covered under Genetics of Bacteria and Viruses II – Bacteria).*

2. Animal Viruses

LO 2e (iii) Describe the reproductive cycle of an enveloped virus e.g influenza virus

2.1 Influenza virus

Structure

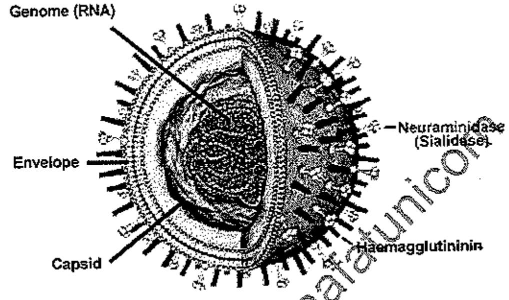


Fig. 11. Structure of the influenza virus

The structure of the influenza virus is somewhat variable, but the virion particles are usually spherical or ovoid in shape and 80 to 120 nm in diameter. Sometimes filamentous forms of the virus occur as well, and are more common among some influenza strains than others.

A. Genome

- The influenza genome is organised into **eight segments** of **single-stranded RNA.**
- The RNA genome is a **negative strand**, i.e. the sequence of the viral RNA genome is complementary to the sequence of the viral mRNA.
- The RNA is packaged with protein into a **helical nucleoprotein** form, with **three RNA segments** coding for **three different polymerases.** The three polymerases form an enzyme complex, **RNA-dependent RNA polymerase** or RNA replicase, which functions in both **replication** and **transcription** of the viral genome.
- The other **five RNA segments** code for **haemagglutinin, neuraminidase**, nucleoprotein, matrix protein M1 and non-structural proteins.

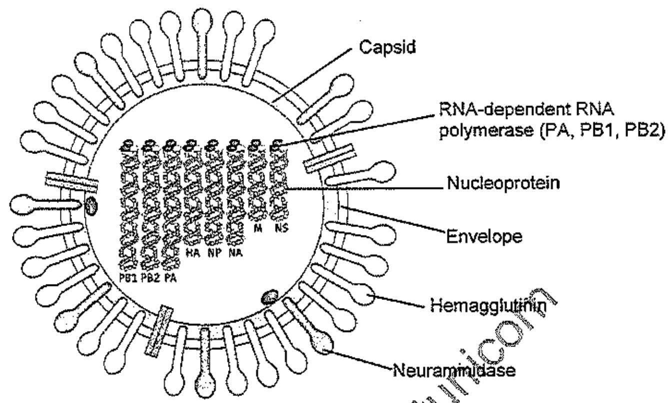


Fig. 12. Structure of influenza virus with RNA-dependent RNA polymerase complex attached to each segment of the RNA genome.

B. Capsid

- The **capsid** is an antigenic protein lining on the inner side of the envelope.

C. Envelope

- The influenza virion is an enveloped virus that derives its lipid bilayer from the plasma membrane of a host cell.
- **Haemagglutinin (a glycoprotein)** and **neuraminidase (an enzyme)**, are embedded in the envelope. Different types of **haemagglutinin** and **neuraminidase** glycoproteins give rise to different strains of influenza virus.

Exocytosis vs. Budding

- release of vesicle content to the exterior
- fusion of the vesicle membrane with the cell membrane (lengthening of cell membrane)

- enveloped viruses acquire their external envelope from the host cell membrane
- membrane evaginates, encloses virion & pinches off. (shortening of cell membrane)

Endocytosis vs. Fusion

Life Cycle of Influenza virus

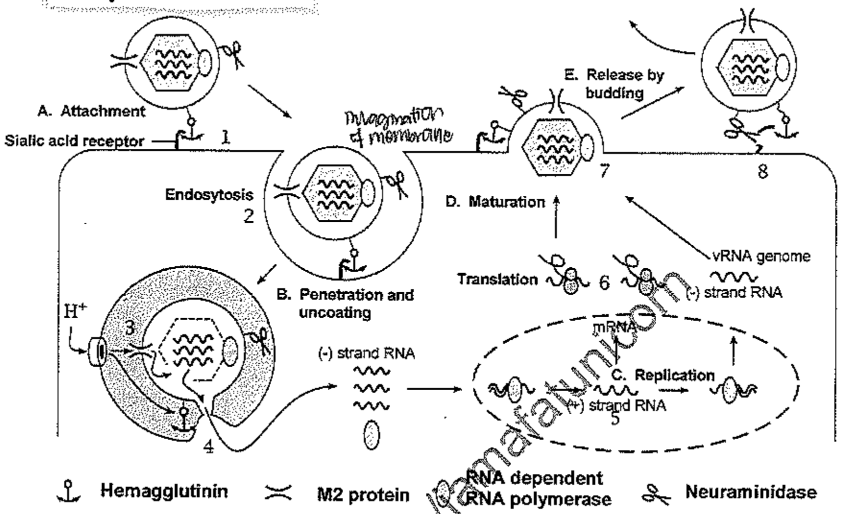


Fig. 13. Life cycle of an influenza virus

A. Attachment

- Protruding glycoproteins bind to specific receptor molecules on surface of host cell. In humans, **haemagglutinin** on the influenza virus binds to the **sialic acid receptor** on the host cell membrane. *(See pg 21 for structure & location of sialic acid).*

B. Penetration and Uncoating

- The virus usually enters by **endocytosis**. The host plasma membrane invaginates and pinches off, placing the virus in an endocytic vesicle/endosome.
- The **vesicle will then fuse with a lysosome** causing its pH to drop. Within the vesicle, the low pH environment will stimulate the viral envelope to fuse with lipid bilayer of the vesicle membrane and nucleocapsid is released into the cytoplasm.
- The **capsid** is then degraded by cellular enzymes leaving behind the helical nucleoprotein *(see Fig. 12, pg 18)*. The helical nucleoprotein then enters the nucleus of the cell.

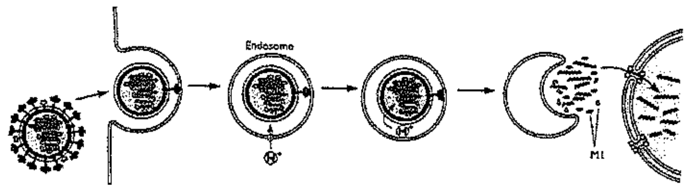


Fig. 14. Entry and uncoating of influenza virus in the host cell

C. Replication

- The viral genome is used as a template to synthesise the viral mRNA/(+) strand RNA catalysed by the viral **RNA-dependent RNA polymerase.** The mRNA produced in turn acts as a template for the synthesis of new viral RNA genome.
- The mRNA strands then exit the nucleus to the cytosol and RER where they are **translated** into viral structural components such as the glycoproteins to be incorporated into the viral envelope (at the ER) and capsid proteins (in the cytosol).

D. Maturation

- Viral glycoproteins are transported by the vesicles from the ER. They are incorporated into the plasma membrane.
- Capsid proteins then associate with these glycoproteins at the plasma membrane.
- The viral genome associates with proteins to form the helical nucleoprotein which then interacts with the capsid proteins at the plasma membrane of the host cell.
- Interaction of the **capsid** with the nucleoprotein will initiate the **budding** process.

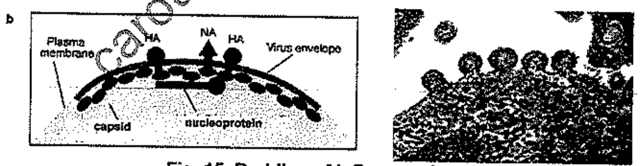
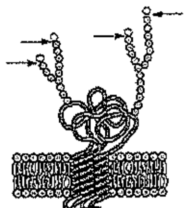


Fig. 15. Budding of influenza virus

E. Release by budding

- Each new virus **buds** from the cell (**evagination**) *(Fig.13 and 15).*
- It will acquire the host membrane with viral glycoproteins embedded.
- With enveloped viruses, host cells may or may not be lysed
- The release is facilitated by neuraminidase *(see pg 21 for location of bond cleaved by neuraminidase).* **Neuraminidase cleaves sialic acid** from cell surface and progeny virions facilitating virus release from infected cells.

Host receptor glycoprotein & its interaction with influenza virus.



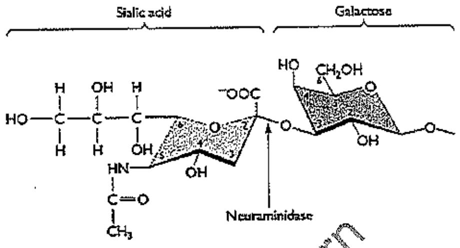
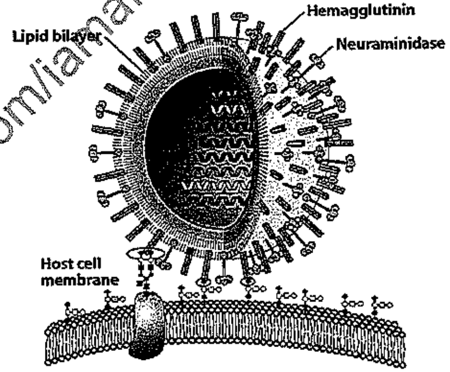


Fig (A) Host cell glycoprotein containing terminal sialic acid. (B) Molecular structure of sialic acid.

The diagram on the left (Fig. A) represents an integral membrane glycoprotein found on the host cell; the arrows point to terminal sialic acid units that are attachment sites for influenza virus.



The shape of the sialic acid is complementary to the binding site on the influenza haemagglutinin glycoprotein. The binding between sialic acid and haemagglutinin will facilitate the attachment of the virus to its host cell and its subsequent entry into the host cell.

In the diagram above, (Fig. B), the molecular structure of a terminal sialic acid moiety can be seen. Sialic acid is attached to galactose by an α(2,3) linkage in the example shown above; certain HA subtypes preferentially bind to molecules with an α(2,6) linkage. The site of cleavage by the influenza virus envelope protein neuraminidase is indicated.

The sialic acid shown is N-acetylneuraminic acid, which is the preferred receptor for influenza A and B viruses. These viruses do not bind to 9-O-acetyl-N-neuraminic acid, the receptor for influenza C viruses.

Adapted from E. Wimmer, Cellular Receptors for Animal Viruses (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1994).

LO 2e (iv) Describe the reproductive cycle of retroviruses, e.g. HIV

2.2 Retrovirus - HIV

Structure

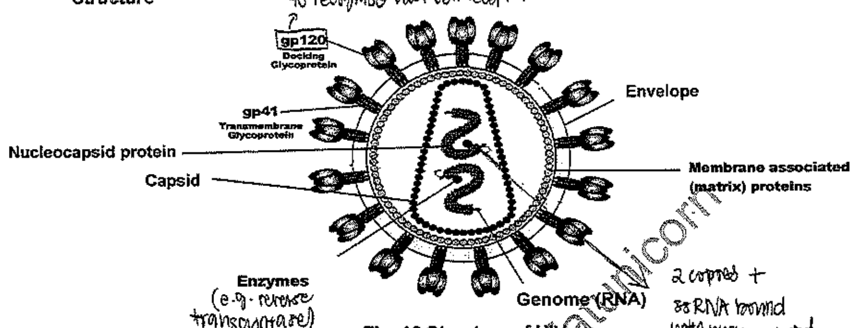


Fig. 16 Structure of HIV

The HIV virus is around 120 nm in diameter and roughly spherical.

A. Genome

- HIV-1 has **two copies of single-stranded RNA.**
- The two copies of single-stranded RNA are **positive strands**, i.e. the viral genome has same sequence as the viral mRNA.
- RNA is tightly bound to proteins, known as the **nucleocapsid proteins.** (Please note that the nucleoprotein in HIV is known as the nucleocapsid protein.)
- The HIV genome contains three major genes, 5'gag-pol-env-3', encoding major structural proteins as well as essential enzymes.
  These are synthesized as polyproteins which produce proteins for the virion:
  - *Gag* codes for structural proteins (capsid, matrix and nucleocapsid protein)
  - *Pol* codes for the viral enzymes (reverse transcriptase, integrase and HIV protease)
  - *Env codes for* the glycoproteins gp120 and gp41

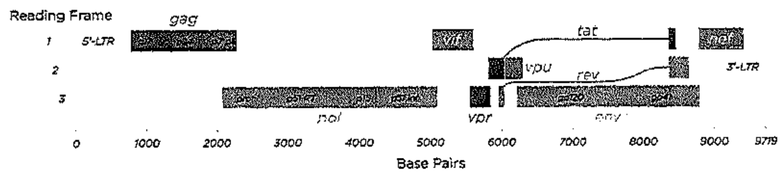


Fig. 17 Structure of the RNA genome in HIV

B. Capsid

- The **capsid** is usually **conical-shaped** and made of another type of proteins different from the nucleocapsid proteins.
- Within the **capsid** are two molecules of enzyme **reverse transcriptase.** The reverse transcriptase transcribes RNA (as template) into DNA. Two other enzymes contained within the **capsid** are **integrase** and **protease.**
- The **capsid** together with the viral genome forms the virus **core**.

C. Envelope

- The **capsid** is in turn surrounded by an envelope that is formed from part of the host cell plasma membrane.
- Through the envelope, glycoproteins protrude. These glycoproteins, **gp120 and gp41**, have a specific conformation that allows the virus to bind to certain **receptors on T_4 helper cells.**

Life Cycle of HIV

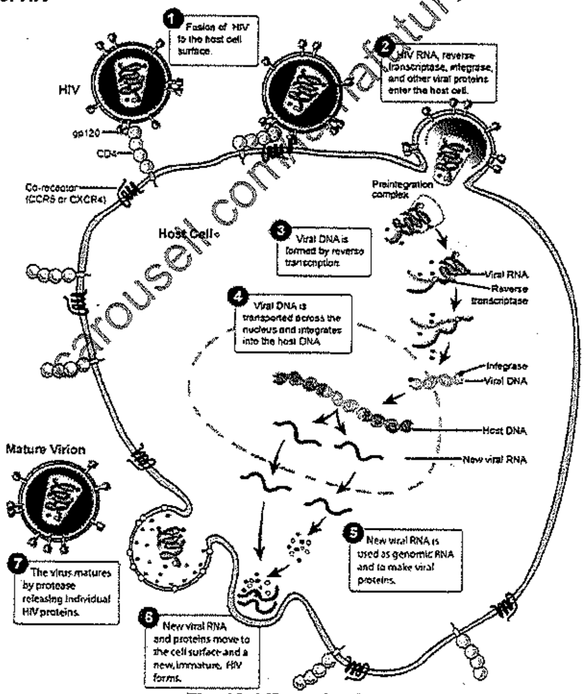


Fig. 18 Life cycle of HIV

A. Attachment

- The process typically begins when a viral particle comes into contact with a cell that carries on its surface a special protein called **CD4.** The glycoprotein, **gp120** on the surface of the viral particle interacts with the CD4 on the target cell (**T lymphocytes, macrophages**), with the help of a co-receptor. (**Step 1**).

B. Penetration and Uncoating

- (**Step 2**) With the help of gp41, the viral envelope will fuse with the host cell membrane and the **capsid** is then released into the cell, leaving the envelope behind.
- The **capsid** and the **nucleocapsid protein** are then degraded, releasing viral enzymes and the RNA into the cytoplasm.

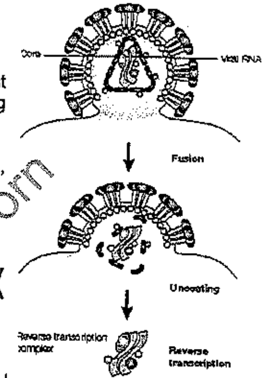


Fig. 19. Entry of HIV by fusion with host cell membrane followed by uncoating of the capsid and nucleocapsid to release the RNA genome.

C. Replication

- (**Step 3**) The viral **reverse transcriptase** enzyme (single-stranded.) will then catalyse the conversion of the viral RNA into DNA. Reverse transcriptase will first catalyse synthesis of a single DNA strand complementery to the viral RNA strand → form RNA-DNA hybrid.
- The RNA strand is degraded and a second DNA strand complementary to the first is synthesised to form a double-stranded DNA molecule.
- (**Step 4**) Double stranded viral DNA enters the host cell nucleus where it is integrated into the genetic material of the host. It is now known as a **provirus.** The enzyme **integrase** catalyses this process. Once viral DNA is integrated into the host genetic material, it may persist in its **latent** state for many years.
- Activation of the host cell will result in transcription of viral DNA into viral RNA which serves as the mRNA.

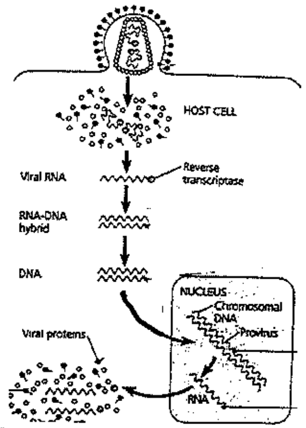


Fig. 20. Reverse transcription and integration into host genome by HIV virus.

- (**Step 5**) The mRNA exits the nucleus into the cytoplasm where it is translated into viral **polyproteins.** (polypeptides)
  ↳ many proteins that are stuck tgt (not active) → have to be cleaved by protease

- The envelope glycoproteins **gp 120** and **gp 41** are made in the ER and vesicles will transport them to the cell membrane. (The *env* polyprotein is cleaved by host cell protease in the ER) produced via RER
- The viral RNA also forms the genetic material for the next generation of viruses.

D. Maturation
- **(Step 6)** Polyproteins and HIV genomic RNA assemble at the inner surface of the plasma membrane of the host cell.

E. Release
- After assembly at the plasma membrane, the virus **buds off/evaginates** from the cell is released.
- The viral envelope is derived from the host cell membrane containing gp41 and gp120.
- Polyproteins will be cleaved into the functional proteins by **HIV protease.**
- The functional proteins include structural proteins (matrix, capsid, nucleocapsid proteins) and viral enzymes (reverse transcriptase, integrase, HIV protease).
- The virion is now considered mature and ready to infect another cell.

Different mode of penetration for different enveloped viruses.

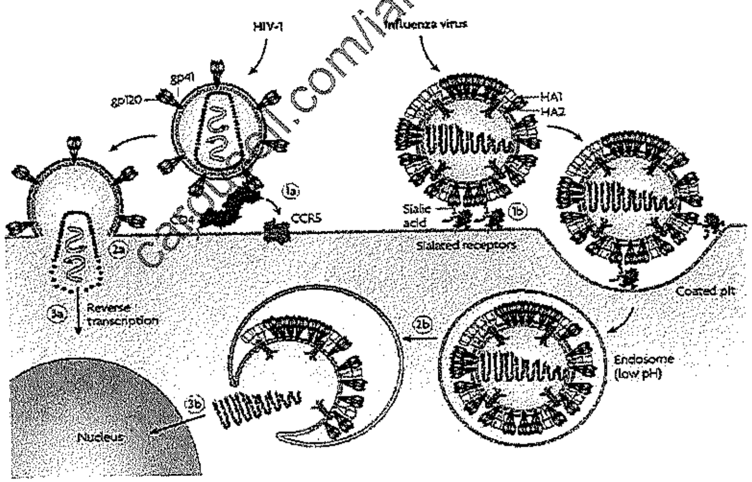


Fig. 20. Diagram showing entry of HIV virus via fusion of lipid membranes and entry of influenza virus via endocytosis.

Question:
What are the differences between multiplication of animal viruses and that of bacterial viruses?

Stage	*Bacteriophage*	*Animal Viruses*
Adsorption	Tail fibres attach to cell wall proteins	Attachment sites are plasma membrane and glycoproteins proteins
Penetration	Viral DNA injected into host cell	Capsid enters by endocytosis or fusion
Uncoating	Not required	Enzymatic removal of capsid proteins
Genome replication	In cytoplasm	In nucleus (DNA viruses) or cytoplasm (RNA viruses with the exception of HIV and Influenza)
Release	Host cell lysed	Enveloped viruses bud off/ non-enveloped viruses rupture the plasma membrane

(D) ANTIGENIC DRIFT AND ANTIGENIC SHIFT

LO 2f Describe how variation in viral genomes arises, including antigenic shift and antigenic drift

- Antigens are the specific molecular structures that antibodies and other receptors in our immune systems recognize.
- The immune system recognizes viruses when antigens on the surfaces of virus particles bind to immune receptors that are specific for these antigens.
- After an infection, the body produces many more of these virus-specific immune receptors, which prevent re-infection by this particular strain of the virus and produce acquired immunity.
- Similarly, a vaccine against a virus works by enabling the immune system to recognize the antigens exhibited by this virus.
- However, viral genomes are constantly mutating, producing new forms of these antigens. If one of these new forms of an antigen is different from the old antigen, it will no longer bind to the receptors. Viruses with these new antigens can evade immunity to the original strain of the virus. When such a change occurs, people who have had the infection in the past do not have immunity to the new virus and vaccines against the original virus will also become less effective.
- Two processes drive the antigens to change: **antigenic drift** and **antigenic shift**, with antigenic drift being the more common.

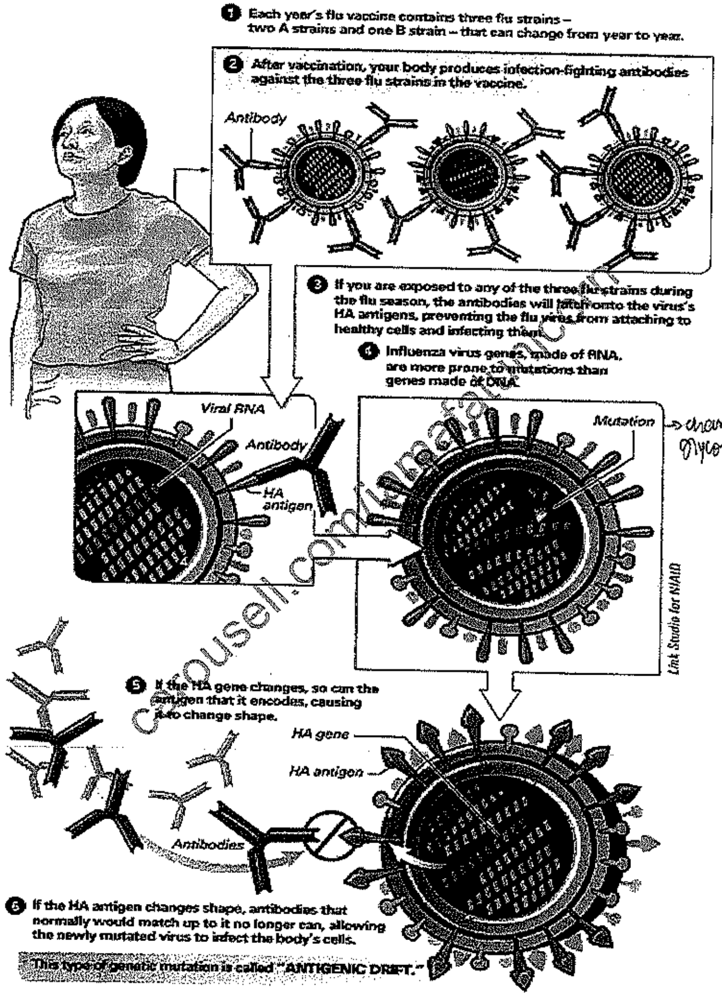
Antigenic drift → minor changes in the structure & surface antigens

- A mechanism of variation by viruses that involves the accumulation of **mutations** in the genes encoding the surface glycoproteins of the virus. The resulting viruses **have surface antigens or glycoproteins** that have **a different conformation** to the previous virus strain.
- Hence, the new virus strain cannot be recognised by antibodies against previous strains making it easier for them to infect the host and spread throughout a partially immune population.
- Antigenic drift occurs in both influenza A and influenza B viruses. In the influenza virus, the two relevant antigens are the surface proteins, **haemagglutinin** and neuraminidase.
- Sites recognized on the **haemagglutinin** and neuraminidase proteins by host immune systems are under constant selective pressure. Antigenic drift allows for evasion of these host immune systems by small mutations in the **haemagglutinin** and **neuraminidase** genes that make the protein unrecognizable to pre-existing host immunity.
- Antigenic drift is this continuous process of genetic and antigenic change among influenza strains as a result of the **lack of proof reading** ability of **RNA-dependent RNA polymerase** and the **fast/high rate of replication** of the virus.

Why Antigenic drift occurs frequently in influenza?

From: http://en.wikipedia.org/wiki/Antigenic_drift

Fig. 21 Antigenic Drift



Antigenic shift

- A process whereby there is a sudden and **major change** in the **surface antigens** of a virus. The genetic change that enables a flu strain to **jump from one animal species to another**, including humans, is called antigenic shift.
- This occurs when **two or more different strains** of a virus, or strains of two or more different viruses, combine to form a **new subtype** having a mixture of the surface antigens of the two or more original strains (refer to method 3 in Fig. 22 on pg 30)
- Antigenic shift is a specific case of **genetic reassortment** that confers a phenotypic change.
- The term is often applied specifically to influenza, as that is the best-known example. Antigenic shift occurs mainly in influenza A. This is because influenza A viruses are found in many different animals, including ducks, chickens, pigs, humans, whales, horses, and seals.
- Antigenic shift occurs because the genome of the virus is segmented, allowing for major genetic changes of type by re-assortment of its segmented RNA genome.
- Flu strains are named after their types of **haemagglutinin** and neuraminidase surface proteins, so they will be called, for example, H3N2 for type-3 **haemagglutinin** and type-2 neuraminidase.
- When two different strains of influenza infect the **same cell** simultaneously, their protein capsids and lipid envelopes are removed, exposing their RNA, which is then transcribed to mRNA. The host cell then forms new viruses that combine their antigens; for example, H3N2 and H5N1 can form H5N2 this way.
- Because the human immune system has difficulty recognizing the new influenza strain, most people do not have pre-existing antibody protection to these novel viruses.
- The new strain may further **evolve** to spread from person to person. This could cause the formation of a highly virulent virus. If so, a flu pandemic could arise.

Note:

1. Refer to Fig. 22 on pg 30. Method 1 and 2 where infection jump from one species to another are also considered as antigenic shift.
2. Influenza viruses which have undergone antigenic shift have caused the Asian Flu pandemic of 1957, the Hong Kong Flu pandemic of 1968, and the Swine Flu scare of 1976.

From: http://en.wikipedia.org/wiki/Antigenic_shift

Fig. 22 Antigenic Shift

Question:
What are the differences between antigenic shift and antigenic drift?

Feature	*Antigenic shift*	*Antigenic drift*
Change in antigen	Major antigenic change An antigenic change which results in drastic alteration in haemagglutinin / neuraminidase subtypes	Minor Antigenic Change An antigenic change can alter antigenic sites on the molecule such that a virion can escape recognition by the host's immune system.
New strain or subtype	Forming a new subtype (subtype A + subtype B → new subtype)	Forming new strain of virus
No. of type of virus involved	one or two viruses are involved	Only one virus is involved
Host species	May jump from one species to another (animal-human)	infect animals of the same species.
Change in genome	Large change in nucleotides of RNA	small mutation of RNA.
Process that leads to change in genome	occurs as a result of genome reassortment between different subtypes	Occurs as a result of the accumulation of point mutations in the gene.
Frequency of occurrence	Occurs once in a time	Occurs frequently
Consequences	Give rise to pandemics, which occurs irregularly and unpredictably.	Usually responsible for epidemics in between pandemics.

(E) PATHOGENICITY OF ANIMAL VIRUSES

Will be revisited in Infectious disease

Pathogenicity of a virus refers to its ability to cause disease.

Most pathogenic viruses produce acute or asymptomatic infections that rapidly run their course and stimulate permanent immunity in survivors. This is due to the production of immune cells and antibodies that specifically recognise and inhibit subsequent infection by the same types of viruses. E.g. chickenpox, measles and mumps.

On the other hand, common colds may be caused by more than 100 distinct strains of the *rhinoviruses* and because immunity is specific, infection with one strain fails to induce host immunity to the other strains.

Some viruses escape elimination by the immune response by establishing **latent** (hidden) **infection.** These viruses remain in the host even after disease symptoms disappear and are generally undetectable during the latent periods. The disease may be periodically reactivated by various stimuli.

An example (for your interest only) are the recurrent episodes of Type 1 (oral) herpes or Type 2 (genital) herpes, that may be triggered by emotional stress, sunburn, menstruation, pregnancy, common colds or diseases that cause fever. Between episodes, the virus silently resides within the regional nerve cells.
Persistent infection may also be established in brain tissue following measles infection, leading to the development of a slowly progressive, ultimately fatal neurological disease called subacute sclerosing panencephalitis (SSPE).

1. Influenza virus

- ***Pathogen:***
  - A type of myxovirus
- ***Target organ:***
  - Epithelial cells of the respiratory tract; virus binds to the sialic acid receptor found on the cell membrane of epithelial cells.
- ***Symptoms:***
  - Body aches, headache, chills and fever, running nose. In more serious cases, influenza can cause pneumonia which can be fatal especially in young children and the elderly.
- ***The disease:***
  - Influenza is a respiratory disease in humans, with the infection usually localised in the respiratory tract.
  - Once the virus settles on the mucous membrane lining the nose, pharynx, trachea and bronchi, the neuraminidase enzyme on the surface of the viruses helps them to penetrate the mucoproteins in the mucus layer. The mucoproteins are glycoproteins in nature.
  - The haemagglutinin, a glycoprotein on the viral envelope, then helps the virus bind to specific receptors on cell membrane of the epithelial cell lining the respiratory tract.

Eventually, the virus penetrates into these host cells. Once inside, the virus replicates within them.

- The incubation period is around 24 to 48 hours, after which the infected epithelial cells are destroyed. These lead to inflammation and the buildup of dead epithelial cells in the airways cause the symptoms of influenza like running nose and scratchy throat to appear.
- Weakening of the epithelial layer caused by viral replication can make the respiratory passage more susceptible to secondary bacterial infections leading to diseases like pneumonia which can be fatal.

- ***Mode of transmission:***
  - Droplets of moisture from lungs of infected persons or from infected bird droppings etc.

- ***Treatment***
  - There is no treatment for most people who develop influenza. However, bed rest, and perhaps the administration of aspirin or paracetamol to alleviate headaches and fever; is the only helpful means towards recovery.
  - Antibiotics are administered to prevent secondary bacterial infection like pneumonia.
  - Vaccinations against influenza are also sometimes administered. The influenza vaccine contains purified and inactivated material from the three common influenza viral strains.
  - Antiviral drugs such as oseltamivir (trade name Tamiflu) and zanamivir (trade name Relenza) are neuraminidase inhibitors that are designed to halt the spread of the virus in the body. However, these drugs are more suited towards the Influenza A and Influenza B strains. The antiviral drugs amantadine and rimantadine are designed to block a viral ion channel (M2 protein) and prevent the virus from infecting cells. These drugs are effective against Influenza A but not against Influenza B.

2. HIV

- ***Pathogen :***
  - Human immunodeficiency virus (HIV), which is a **retrovirus**

- ***Target organ :***
  - The immune system specifically cells with CD4 receptors such as macrophages and T helper cells.

- ***Mode of transmission***:
  - Transmitted primarily through unprotected sexual contact or by exposure to infected blood and blood products.
  - It can also be transmitted from mother to child either via the placenta, childbirth or during breastfeeding.

- ***The disease***

  - Once the virus enters the blood stream, its primary targets are macrophages (a large phagocytic cell) and T helper cells (a type of lymphocyte responsible for the coordination of an immune response to infection).
  - The HIV virus has a very strong affinity for and binds to CD4, a surface protein of T helper cells. As the HIV infects more and more T helper cells, levels of T helper cells lower as the infected cells are destroyed.
  - The macrophages may survive HIV infection (because they are not lysed by the virus) and may thus act as reservoirs.
  - HIV may be passed from cell to cell in an infected individual, or it may be transmitted via body fluids to another person, while still remaining undetected.
  - The virus mutates at a very high rate during replication resulting in altered proteins on the surface of the virus. In this way, the virus prevents recognition and elimination by the immune system allowing it to evolve rapidly within the body.
  - The increasing loss of T helper cells leads to impaired immune responses in the affected individual who then becomes increasingly susceptible to opportunistic diseases. 'Full-blown AIDS' occurs when the infections become unmanageable. The immuno-suppression becomes worse, usually with fatal results. Death usually results from secondary infections.

- ***Treatment***
  - Due to the high rates of virus productions and mutation rate of the virus, treatment of HIV infection generally includes administration of 3 agents in combination. Sustained treatment results in suppression of viral replication, dramatically increasing life expectancy of HIV-infected individuals.
  - Currently, there are 24 approved retroviral drugs which can be used to treat HIV infection. They include reverse transcriptase inhibitors, protease inhibitors, integrase inhibitors and entry inhibitors. The first three enzyme inhibitors work by inhibiting their respective enzymes. Entry inhibitors work by blocking interaction between the HIV envelope and CD4 or the co-receptor, or by preventing fusion of the viral and host cell membranes, thus blocking entry of HIV into the cells.

3. Pathogenesis of other viruses

Infection of cells by a virus can alter the host cell in various ways resulting in disease in the affected individual. Here are some of the ways in which viruses cause disease:

a. Death of host cell
When a virus enters a cell, it can incur temporary or permanent changes to the host cell. In bacteria viruses, it usually ends with lysis of the bacterium. Cell respond to the presence of viral DNA in animal cells may result in apoptosis of the cell.

b. Production of toxic substances

During the course of virus replication, many viral components as well as by-products of viral replication accumulate in the cell. These are often cytotoxic. The molecular mechanism of these toxins is not known in most cases.

Some examples of toxin effects include:

i. Herpesvirus components produce syncytia – a multi-nucleated protoplasmic mass that is not viable and eventually dies.
ii. Cytotoxicity of preformed viral parts. e.g., Sendoi virus, Newcastle disease virus, measles virus and SV5 produce rapid polykaryocytosis (fusion of chromosomes).
iii. The antigen of the adenovirus **capsid** inhibits RNA, DNA and protein synthesis.

c. Cell transformation

Certain viruses have the ability to enter a cell and follow one of two alternative courses. They either multiply in a normal manner and are eventually released from the cell, or they may be dormant in the cell and eventually transform the cell into a cancerous cell. The growth and division of normal cells is regulated by proto-oncogenes and tumour suppressor genes. Changes in expression of either gene will lead to the uncontrolled cell division and growth causing tumour formation. Tumours can be either benign or malignant. Malignant tumours are able to travel to other parts of the body and invade tissues and organs. Viruses usually cause genetic changes which affect the proto-oncogenes or the tumour suppressor genes resulting in cancer. For example, some retroviruses activate the proto-oncogenes to oncogenes (a gene that causes tumour formation).

d. Suppression of immune mechanisms

Since many viruses are known to replicate in cells of the lymphatic system, it is possible that these viruses can affect the immune system. The nature and extent of the immunologic alteration depends on the organ or cell type infected and the species of virus causing the infection e.g. HIV.

e. Induction of non-normal host-specified products

Virus-infected cells, at times, will produce compounds coded for by the host DNA, but which are not normally produced by the host. These are often cytotoxic at relatively high concentrations. Other host compounds which are normally found in low concentration may be produced in higher concentration during a virus infection. Again, this high concentration may be cytotoxic. Some virus-induced products release autolytic enzymes from the cells own lysosomes.

f. Induction of structural alterations to the host cell

Viruses can induce structural alterations in the host cell's cytoplasm and nucleus.

1. Cytoplasmic changes
   (a) Small non-enveloped RNA viruses produce a large eosinophilic mass which displaces the nucleus. There is a generalized increase in basophilia. The cytoplasm appears to bubble at the cell periphery.
   (b) Herpesvirus causes vacuolisation.

2. Nuclear changes
   (a) Nuclear inclusion (bodies in the nucleus); e.g., herpesvirus, adenovirus.
   (b) Margination and coarsening of chromatin; e.g. herpesvirus, poxvirus.
   (c) Formation of chromosomal breaks.

3. Membrane changes
   (a) The human cell membrane is a dynamic structure continually changing in lipid and protein content during normal cellular growth and division. Viral infection of the cell often results in viral protein being incorporated into this membrane. These changes can lead to production of antibodies against the cell membrane and lysis of this membrane.

In summary, we get sick from viral infections because:

1) The infected cells die as a direct or indirect result of the virus,
2) The viral products or components can cause an immune response e.g. fever, allergy, etc.
3) The virus can suppress the immune system increasing the susceptibility to secondary infections.
4) Viruses can cause cancer.

4. Antiviral Drugs and how they work

- Chemically resembles nucleosides (e.g. AZT) and interferes with the viral nucleic acid synthesis
- e.g. acyclovir – inhibits viral polymerase (disrupts viral DNA synthesis) in herpes virus
- e.g. azidothymidine (AZT) - inhibits the enzyme reverse transcriptase in HIV, preventing the production of a DNA copy of the virus RNA genome.

(F) LINKS

Topics	Link to
T4 life cycle lambda bacteriophage life cycle	Transduction in Bacteria
Retrovirus or other virus that integrate into host cell genome	Cancer
HIV, influenza	Infectious diseases
Antigenic shift and drift	Evolution, Mutation
Enveloped virus glycoprotein synthesis	Protein synthesis and the endomembrane system
Enveloped virus glycoprotein	Protein structure

(G) GLOSSARY

Antigen – any substance capable of inducing a specific immune response and of reacting with the products of that response, i.e., with specific antibody or specifically sensitized T lymphocytes, or both

Antigenic drift – Minor changes in the structure and immunogenicity of antigens, specifically surface proteins of influenza virus, caused by mutation.

Antigenic shift – major changes in the structure and immunogenicity of the surface proteins of influenza virus caused by gene exchange with related viruses.

Bacteriophage - A virus that parasitizes a bacterium by infecting it and reproducing inside it.

Budding - is a form of **viral** shedding by which enveloped **viruses** acquire their external envelope from the host cell membrane, which bulges outwards and encloses the virion

Capsid - the protein coat or shell of a viral particle, surrounding the nucleic acid or nucleoprotein core

Capsomeres – the protein components of **capsid**

Genome - the total genetic content contained in a haploid set of chromosomes in eukaryotes, in a single chromosome in bacteria, or in the DNA or RNA of viruses

Integrase - a viral enzyme that enables the integration of viral genetic material into a host cell's DNA. Usually refer to retroviral integrase (IN) in HIV but should not confuse with phage integrases, such as λ phage integrase (Int).

Lysogenic cycle – a type of phage replicative cycle in which the viral genome becomes incorporated into the bacterial host chromosome as a **prophage**, is replicated along the chromosome, and dose not kill the host.

Lytic cycle - a type of phage replicative cycle resulting in the release of new phages by lysis of the host cell.

Nucleocapsid – nucleic acid plus capsid. A complex of proteins and the viral genomic nucleic acid.

Nucleoprotein - A complex consisting of a nucleic acid bonded to a protein

Prophage - The genetic material of a bacteriophage, incorporated into the genome of a bacterium and able to produce phages if specifically activated

Provirus - The viral genome that is incorporated into, and able to replicate with, the genome of a host cell

Retrovirus - A retrovirus is an RNA virus that is duplicated in a host cell using the reverse transcriptase enzyme to produce DNA from its RNA genome

Reverse transcription - DNA synthesis from RNA templates, catalysed by the enzyme reverse transcriptase; the opposite of transcription. Occurs naturally in retroviruses.

Reverse transcriptase - a **reverse transcriptase**, also known as **RNA-dependent DNA polymerase**, is a DNA polymerase enzyme that transcribes single-stranded RNA into single-stranded DNA.

Virion - The complete, infective form of a virus outside a host cell, with a core of RNA or DNA and a capsid

Vaccine – a harmless variant or derivative of a pathogen that stimulates the host's immune system to mount defenses against the pathogen.

(H) SOME KEY WORDS:

T4 bacteriophage	Antigenic shift	Haemagglutinin
Lambda bacteriophage	Antigenic drift	Neuraminidase
Capsid	Mutation	Lysosome
Capsomere	8 single-stranded RNA segments	Evaginates
Nucleocapid	endocytosis	Bud off/budding
Nucleoprotein	Reverse transcriptase	
Tail fiber	Integrase	
Tail sheath	Protease	
Base plate	Polyprotein	
Lysogenic cycle	gp41 and gp120	
Lytic cycle	CD4 receptor	
Prophage	RNA dependent RNA polymerase	
Lysozyme	T4 helper cell	
Repressor protein	2 single-stranded mRNA	
Virion	Provirus	
Latent stage	Acellular	
Lysis	Fusion	
Obligate parasite	Retrovirus	
Temperate phage	Genetic reassortment	

CORE IDEA
(1) The Cell and Biomolecules of Life
(2) Genetics and Inheritance

BACTERIA

Content

- The structure of a typical bacterial cell
- The genetics of bacteria

Learning Outcomes

Candidates should be able to:

1d Describe the structure of a typical bacterial cell (small and unicellular, peptidoglycan cell wall, circular DNA, 70S ribosomes and lack of membrane-bound organelles).

2d Describe the structure and organisation of prokaryotic genome (including DNA, double-stranded, number of nucleotides, packing of DNA, circularity and absence of introns)

2g Outline the mechanism of asexual reproduction by binary fission in a typical prokaryote and describe how transformation, transduction and conjugation (including the role of F plasmids but not Hfr) give rise to variation in prokaryotic genomes.

2i Explain how gene expression in prokaryotes can be regulated, through the concept of simple operons (including lac and trp operons), including the role of regulatory genes; and distinguish between inducible and repressible systems. (Attenuation of trp operon is not required).

Use the knowledge gained in this section in new situations or to solve related problems.

References
Campbell and Reece, BIOLOGY, 9th Ed.
Brooker, Widmaier, Graham and Stiling, BIOLOGY
Green, Stout and Taylor, BIOLOGICAL SCIENCE Vol 1 and 2, 3rd Ed.
Soloman, Berg and Martin BIOLOGY 7th Ed.
Tortora, Funke and Case, MICROBIOLOGY, 6th Ed.
Alberts, Bray, Lewis, Raff, Roberts and Watson, MOLECULAR BIOLOGY OF THE CELL, 3rd Ed.
Raven, Johnson, Losos, Mason and Singer, Mc Graw-Hill, Biology, 8th Edition

This handout is the effort of several Biology teachers at RI. It has and will continue to be updated.

TABLE OF CONTENTS

PAGE

LO 1d Describe the structure of a typical bacterial cell (small and unicellular, peptidoglycan cell wall, circular DNA, 70S ribosomes and lack of membrane-bound organelles).

2d Describe the structure and organisation of prokaryotic genome (including DNA, double-stranded, number of nucleotides, packing of DNA, circularity and absence of introns)

(A) General Bacterial Structure

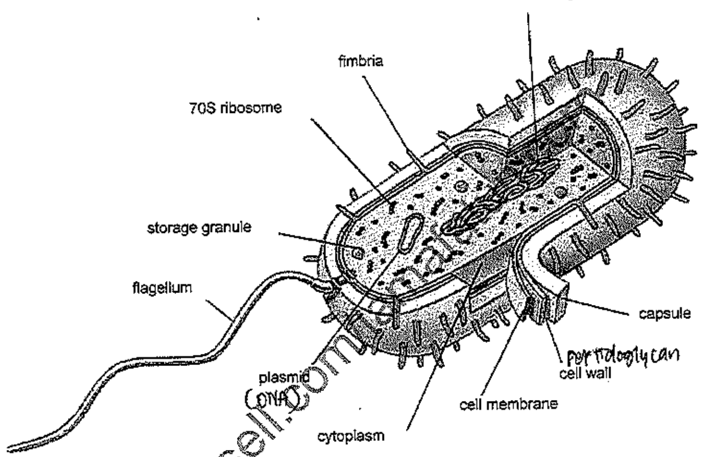


Fig. 1 - Diagram of a bacterial cell

(1) Internal Structure

Bacteria are prokaryotes and have a very simple internal structure with **no membrane-bound organelles.** Structures present include:

- **The Bacterial Chromosome** *(found in nucleoid → not bound by membrane)*
  - The main component of the genome in most bacteria is one **double-stranded, circular DNA molecule** that is **associated with proteins** (they are not called histones!). *→ they are "histone-like proteins"*
  - The DNA forms **loop domains** with the proteins, followed by further **supercoiling,** forming highly condensed DNA

    (NB: stretched out, the DNA of *E. coli* would measure about 1mm in length, 500 times longer than the cell).
  - The entire structure is referred to as the **bacterial chromosome.** Bacterium being prokaryote has **no intron** in its chromosome (non-coding sequences within the gene). *↳ non-coding sequences*
  - The bacterial chromosome makes up a dense region within the cell called the **nucleoid,** which is not bound by a membrane.
  - In addition to the chromosome, some bacteria may also have **plasmids,** which are much smaller rings of **autonomously replicating circular DNA.**

a)

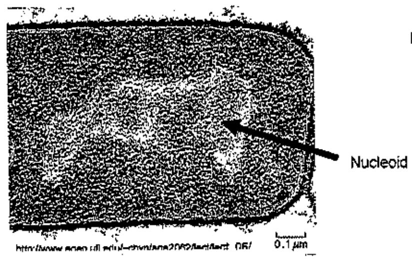
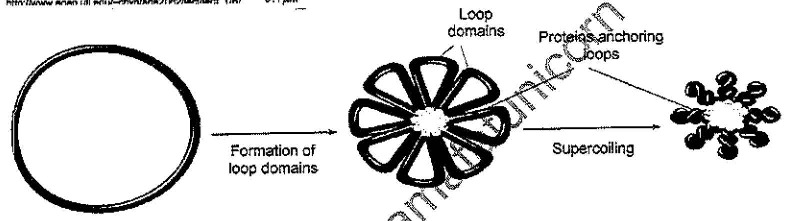


Fig. 2 **a)** The nucleoid region of a bacterial cell is stained less densely than the surrounding cytoplasm

b) Condensation of the bacterial DNA is made possible through its association with proteins

b)



- **Nucleoid:**
  - Region in the bacterial cell where chromosomal DNA is generally confined to.
  - It is not bound by a membrane but is visibly distinct from the rest of the cell interior.
- **Ribosomes:**
  - **70S** (vs. 80S in eukaryotes). They are needed for protein synthesis.
  - The ribosomes give the cytoplasm of bacteria a granular appearance in electron micrographs.
- **Storage granules:**
  - Nutrients and chemical reserves may be stored in the cytoplasm in the form of granules e.g. granules of glycogen, lipids and ions like phosphorous and magnesium.
- **Plasmid(s) (may be present)** *extrachromosomal, small circular DNA.*
  - A small, **circular autonomously replicating DNA molecule.** (This is not the bacterial chromosome.) *↳ independent of bacterial chromosome, multiple copies*
  - The plasmid contains genes which may confer advantages on bacteria living in stressful environments e.g. antibiotic resistance genes. *↳ e.g. F plasmid; R plasmid*
  - Multiple copies are usually present in a cell.
  - Plasmids are used extensively in genetic engineering as vectors for carrying and expressing foreign DNA in bacterial cells.
  - Different bacteria can have different plasmids.
  - Bacterial genome includes both the bacterial chromosome and the plasmid (if present).

– in bacteria, unlike eukaryotic cells, bacteria do not have nuclear envelope that separates DNA from ribosomes
– mRNA translated into protein as soon as it is transcribed as there is no barrier to immediate translation.

(2) Surface Structure

- **Cell Membrane**
  - A phospholipid bilayer similar the cell membrane of other cells.
  - In addition to the roles of a cell membrane which you have learned (See Notes on *Cell Membranes*), the membrane of a bacteria is also where the electron transport chains, as well as the enzyme ATP synthase are embedded to produce ATP during photosynthesis and/or respiration. (How is this different from an eukaryotic cell?)

- **Cell Wall** *→ a polymer of carbohydrates & proteins*
  - Consists of a polymer called **peptidoglycan** – long chains of sugars cross-linked by short peptide chains. (compare: what is the cell wall in plants made up of?)
  - It protects the cell from osmotic lysis.
  - Bacteria may be classified as gram-positive or gram-negative bacteria (depending on whether they get stained by Gram stain which indicates the nature of the cell wall). *▷ mycoplasma has no peptidoglycan cell wall*

(For info)

Peptidoglycan *→ enzymes that form peptidoglycan cell wall are the target of common antibiotics e.g. fleming's penicillin → binding to bacterial enzymes; weaken cells → lyse easily*

> the sugar component is made up of alternating units of NAG (N-acetylglucosamine) and NAM (N-acetylmuremic acid).

> the sugar is linked by short peptide chains as shown on the right

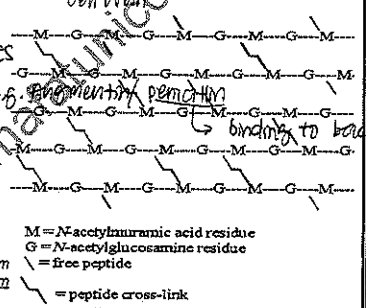
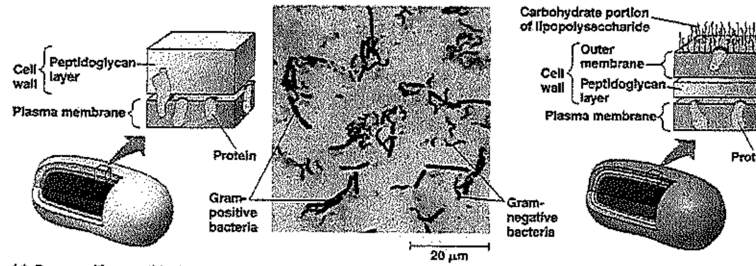


Diagram from http://www.mikeblaber.org/oldwine/bch5425/lect10/lect10.htm

> In gram-positive bacteria, the cell wall is a thick peptidoglycan layer.

> In gram-negative bacteria, the cell wall includes a thin peptidoglycan layer, followed by an **additional outer membrane.**

(For the purpose of the syllabus, only peptidoglycan is considered as the constituent of bacterial cell wall unless the bacteria is specified.)



(a) Gram-positive: peptidoglycan traps crystal violet.

(b) Gram-negative: crystal violet is easily rinsed away, revealing red dye.

- **Capsule** (may be present in some bacteria)
  - Some bacteria have a layer of polysaccharides known as glycocalyx (= sugar coat) to the exterior of the cell wall.
  - The glycocalyx can be a distinct layer, referred to as the **capsule** seen in Fig. 2, or exists it is a diffused mass known as the **slime layer.**
  - The capsule may also contain proteins. *- distinct & gelatinous → capsule - irregular diffused-layer → slime layer*
  - Functions:
    > Often, the capsule protects the bacteria from being taken in via phagocytosis by the white blood cells which are unable to recognize the bacteria due to the capsule

    > It also enables bacteria to adhere to one another or to particular surfaces e.g. mucous membrane.

(3) Appendages (may be present in some bacteria) *< fimbriae, pili, flagella*

Both fimbriae and pili are hollow, hair-like structure composed of protein.

- **Fimbriae** (singular: fimbria)
  - These are short, bristle-like fibres extending from the cell surface and are usually evenly distributed over the entire cell surface or at poles of cells.
  - Function: for attachment to surfaces or other bacteria/organisms

- **Pili** (singular: pilus)
  - pili are longer and fewer in numbers than fimbriae.
  - Function: involves in motility and DNA transfer
    > motility: a pilus makes contact with a surface and retract to pull the bacteria forward in a jerky, intermittent movement

    > DNA transfer: a specialised pilus like the sex pilus, allows two bacterial cells to be drawn close to each other so that a mating bridge can be formed for the transfer of genetic material.

- **Flagella** (singular: flagellum) *(bacteria can have more than 1 :D)*
  - Long appendages for **motility.**
  - The bacterial flagellum is a hollow cylindrical protein thread that **propels** the bacterium by rotation. *(propulsion)*
  - Some bacteria possess more than one flagellum and they may be found distributed all over the cell, at one pole or at opposite poles of a cell.

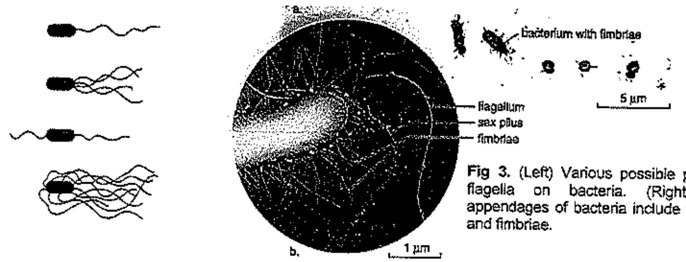


Fig 3. (Left) Various possible positions of flagella on bacteria. (Right) Surface appendages of bacteria include flagella, pili and fimbriae.

Structural features of Bacteria

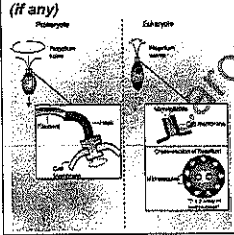
Structural feature	Information
Cell wall	Prevents osmotic lysis of cell protoplast and confers rigidity and shape to cells - composed of peptidoglycan
Location of chromosome	Found within nucleoid region; no true nucleus
Chromosome	A single, circular, double helix DNA – supercoiled
DNA-associated proteins	Proteins anchoring loop domains present
Plasmids	Extra-chromosomal DNA that replicates autonomously; quantity can range from 5 - 100s!
Organelles	No membrane-bound organelles
Ribosomes	70S (vs 80S in eukaryotes)
Appendages:	
a) Fimbriae	Attachment to surfaces and to other bacteria/organisms
b) Pili	(i) Mediates DNA transfer during conjugation (sex pilus) (ii) Motility by retraction
c) Flagella	Swimming movement; propulsion (motility)
Capsules (organised mass of glycocalyx)	protection against phagocytic engulfment; attachment to surfaces; contains water to prevent desiccation → dehydration. - composed of polysaccharides and sometimes polypeptides
Slime layers (diffused mass of glycocalyx)	Attachment to surfaces; to form biofilm - composed of polysaccharides and sometimes polypeptides

Question: List 4 differences between a bacterial and a eukaryotic chromosome [4]

	Point of comparison	Eukaryotic Chromosome	Bacterial Chromosome
①	Location	located within membrane-bound nucleus	Located within nucleoid region; not enclosed by a membrane
②	Structure of DNA	Linear DNA	circular DNA
③	No. of chromosomes in a cell	Have several different chromosomes	only one chromosome
④	Intron	Presence of introns within the genes	absence of introns.
⑤	Associated proteins	Associated with large amount of histone proteins	Associated with small amount of proteins

Structural Features that Distinguish a Prokaryotic Cell from a Eukaryotic Cell

(This will be covered in greater detail in topic 4: Prokaryotic and Eukaryotic genome)

Feature	Prokaryotic cell	Eukaryotic cell
Cell size	Smaller	Larger
Nucleus	No true nucleus / No nuclear envelope	Nucleus present / Nucleus with nuclear envelope present
Genetic material	Circular DNA lying naked in a region in the cytoplasm known as the nucleoid; DNA is associated with small amount of histone-like proteins	Linear DNA found within membrane-bound nucleus; DNA is associated with large amounts of histones proteins
Ribosome for protein synthesis	70S; No ER present for ribosomes to attach	80S; Ribosomes may be attached to ER or may be free in the cytosol
Organelles	Few e.g. ribosomes No membrane bound organelles	Many; Membrane bound organelles present e.g. nucleus, mitochondria; → double membrane e.g. Golgi apparatus, lysosomes, vacuoles, endoplasmic reticulum → single membrane
Cell walls	Composed of peptidoglycan (murein)	Composed of cellulose in plants Composed of chitin in fungi
Flagella (if any) 	Simple; No microtubules; composed of protein flagellin instead; each is a single strand of protein; Extracellular (not enclosed by plasma membrane);	Complex; composed of tubulin; each is made up of several strands with a 9+2 arrangement of microtubules (read Campbell if interested) Intracellular (surrounded by plasma membrane);
Photosynthesis	Involves plasma membrane of bacteria	Involves chloroplast
Respiration	Involves plasma membrane of bacteria	Involves mitochondrion and cytoplasm

DNA Replication (prok v.s. euk)

prok | euk

similarities: semi-conservative, helicase, primer, DNA polymerase, direction, leading & lagging strands, okazaki fragments, etc.

	prok	euk
Points of origin	one	multiple
	short	long
Location	cytoplasm	nucleus.

Q: Does the end replication problem exist for bacteria?
↳ No, because bacterial DNA is circular. No 3' overhang after removal of RNA primer

LO 2g Outline the mechanism of asexual reproduction by binary fission in a typical prokaryote and describe how transformation, transduction and conjugation (including the role of F plasmids but not Hfr) give rise to variation in prokaryotic genomes.

(B) Binary Fission — asexual reproduction

- Binary fission is the means by which bacteria replicate and produce offspring. → no gametes involved (NB: this process is different from division through mitosis). same genetic content in daughter cells, no genetic variation
- Before one bacterial cell splits into two independent cells, the bacterial chromosome must first replicate.

* DNA replication occurs @ same time as the cell is dividing (no nuclear division) ↳ fast & efficient

Bacteria chromosome replication

1. DNA replication begins at the **origin of replication (*ori*)**, made up of a specific sequence of nucleotide bases (Fig. 4).
2. There, the double helix separates to form a **replication bubble** made up of two single DNA strands. Replication takes place outward from the origin in both directions forming 2 replication forks.
   Each replication fork will have both a leading strand and a lagging strand being synthesized, just as in eukaryotes.
3. As the chromosome replicates, the 2 newly formed *ori* move to **opposite poles** of the cell and attach to the plasma membrane.
4. The cell also **elongates** to prepare for division.
5. But because the DNA is circular with no free ends, an interlocking structure made up of the 2 daughter DNA molecules will be formed with the completion of replication.
   Enzyme **topoisomerase** is needed to cut, separate and reseal the two DNA molecules (Fig. 5).
6. When the daughter DNA molecules are separated, the bacterium will have reached twice its initial size. **Invagination** of the plasma membrane and the deposition of new cell wall (also known as division septum) eventually divide the parent cell into two daughter cells, with each inheriting a complete genome (genetically identical).

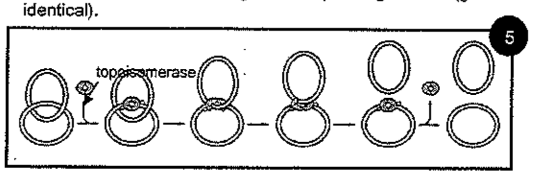


Fig. 5 - Topoisomerase helps to separate 2 entangled DNA molecules.

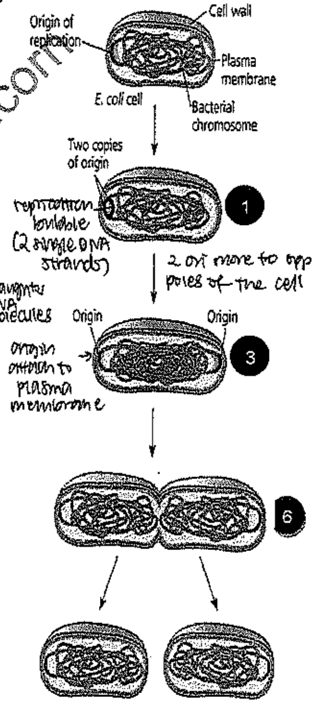


Fig. 4 - The process of binary fission

Binary fission is the asexual means by which bacterial cells produce **genetically identical offspring**.
→This can be a selective advantage in a stable, favourable environment as it allows successful genotypes to rapidly reproduce and colonise a habitat.

Sexual reproduction
↳ changing environment
↳ genetic variation
↳ at least some individuals may survive to reproduce.

Binary fission
- Replication & cell division are occurring simultaneously
- No nuclear division
- No spindle fibres involved
- No specific positioning of chromosomes in the cell that characterise the diff stages in mitosis.

Question: Can you describe how binary fission is different from the process of mitosis?

No spindle fibres are involved; no specific positioning of linear chromosomes in the cell that characterizes the different stages of mitosis, no nuclear division.

(C) Methods by which new DNA is introduced into bacteria to give rise to genetic variation in bacteria genomes

mutation → spontaneous | environment induced

In a rapidly changing environment, **generating genetic variation** through forming **new combination of new alleles** becomes crucial for **enhancing reproductive success** i.e. at least some individuals may be selected for and survive to reproduce (Theory of Natural Selection to be covered under *Evolution*). In eukaryotes, sexual reproduction increases genetic variable within a population.

In bacterial cells which undergo asexual reproduction, 3 processes, **transformation, transduction** and **conjugation** help to increase genetic variation by bringing together DNA from different individuals. ↳ result in exchange of alleles.

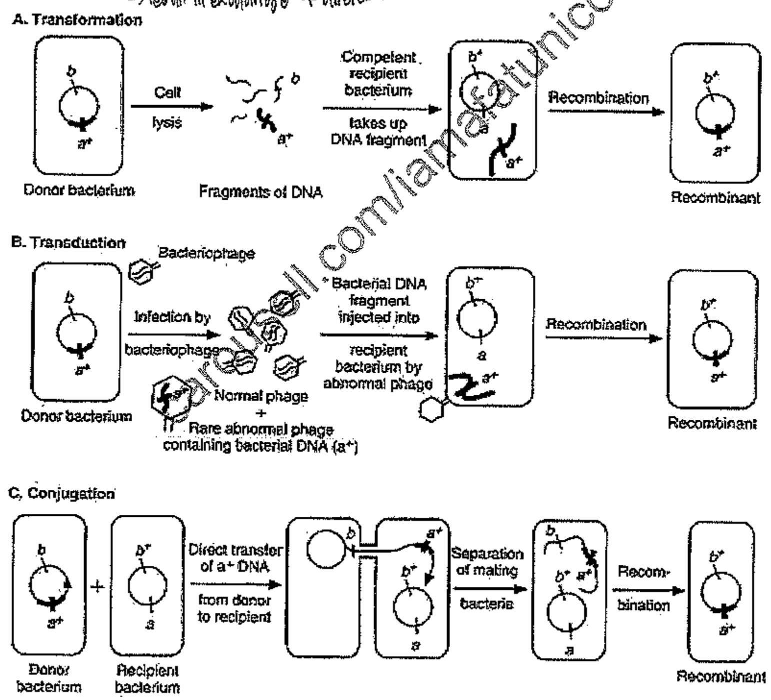


Fig. 6 - Methods by which DNA are introduced into bacteria:

(1) Transformation

- Transformation refers to the **uptake of naked, foreign DNA** from the surrounding environment, resulting in a **change of the bacterial cell's genotype and phenotype.**
- The foreign DNA may have come from **dead lysed neighbouring** cells in the medium.
- Some bacteria possess **cell-surface proteins** that can **bind to and transport DNA** into the cell. Such cells with the natural ability to take up foreign DNA are described as **competent cells.** (naturally competent)
- Bacterial cells that lack these surface proteins can be made artificially competent through immersion in a culture medium with high concentrations of $CaCl_2$ followed by a heat shock treatment. This technique is used in genetic engineering to introduce foreign genes into the *E. coli* genome. (Bacteria has to be induced (artificially competent); transient pores or electroporation)

> *For info*
>
> Bacterial cells are permeable to chloride ions but not the calcium ions. The chloride intake is accompanied by an influx of water into the cells, causing the cells to swell and priming it for the heat shock treatment. The heat shock treatment induces the formation of transient pores which allows for the uptake of DNA from the surrounding medium.
>
> Also as a cation, calcium can bind to both the negatively charged DNA and the cell membrane, which also has a negative charge. This neutralization of charges enhances the ability of the cell to take up the DNA.

- The foreign DNA can then be incorporated into the chromosome through **crossing over at 2 homologous regions** found on the bacterial chromosome (i.e. **homologous recombination**). (Note: If no crossing over occurs, the foreign DNA will not be incorporated into the bacterial chromosome. It will be degraded.)
- The resultant cell is a **recombinant cell.** If different alleles for a gene were exchanged, there will be a permanent change in the organism's **phenotype as the new allele is expressed.** The bacterium's recombinant genome will be passed on to all subsequent offspring through binary fission.

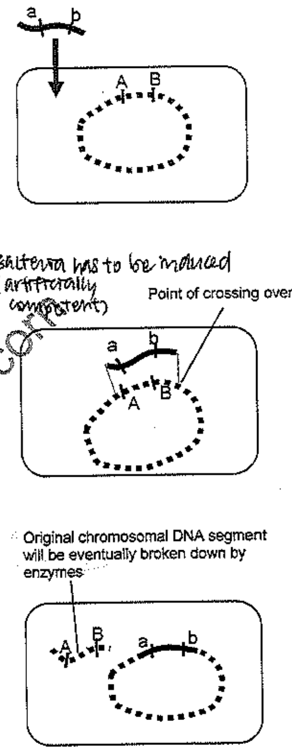


Fig. 7 - Bacterial transformation, including crossing over at a region containing alleles A / a, and alleles B / b

(2) Transduction

In this process, **phages** (viruses that infect bacteria) randomly **carry bacterial genes from one host cell (donor cell) to a recipient cell** as a result of aberrations in the phage reproductive cycle.

2.1 Generalised Transduction

(- random DNA fragments; - lytic cycle (T4 bacteriophage))

- When a phage undergoes the **lytic cycle,** phage enzymes may **hydrolyse the bacterial chromosome** into small pieces of DNA.
- During assembly of the phage genome within the phage capsid, **a small piece of the host cell's degraded DNA is randomly packaged** within a capsid (instead of the phage's own genetic material).
- Following lysis of the host cell (donor cell), the **defective phage** is released and can infect another bacterium (the recipient). The piece of bacterial DNA acquired from the host cell (donor cell) injected into the recipient cell.
- Since viral genes have been replaced by bacterial genes in the defective phage, no new phages can be synthesized (sometimes) in the recipient cell.
- The foreign bacterial DNA can subsequently replace the homologous region of the recipient cell's chromosome if **crossing over** and **homologous recombination** takes place.
- The recipient cell with the new alleles integrated into its genome becomes a **recombinant cell**, which expresses new characteristics.
- As any **random portion of the bacterial DNA** may be transferred, this process is called **generalized transduction.**

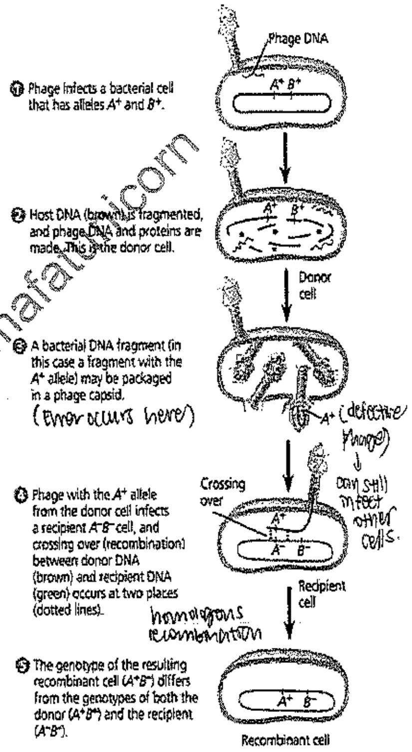


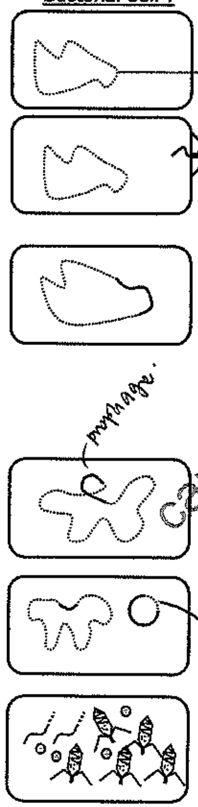
Fig. 8 - The process of generalised transduction

2.2 Specialised Transduction

- This process is carried out by **temperate phages.** (undergoes lysogenic cycle to integrate their genome into the bacterial chromosome, forming a **prophage** ...refer back to your *Virus* notes)
- During specialised transduction, the bacterial DNA that is transferred is **restricted to bacterial genes adjacent to the integrated prophage.** Thus, it is called **specialized transduction.**

Process (including lysogenic cycle):

Bacterial Cell 1



RI©

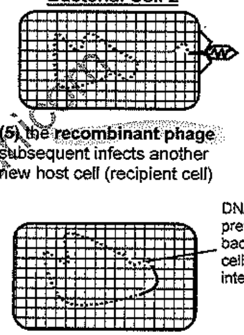
(1) Phage attaches to host cell and injects viral genome into host (donor cell).

(2) Phage DNA integrates into the bacterial chromosome becoming a **prophage** → phage enters the **lysogenic pathway**. As the cell divides by binary fission, a single infected cell can give rise to a large population of bacteria carrying the prophage.

(3) Upon **induction**, the prophage exits the bacterial chromosome, i.e. switches from the lysogenic cycle to the lytic cycle, but the viral DNA **may be improperly excised**, such that **bacterial DNA** adjacent to the prophage is **excised** as well.

(4) The **phage-host hybrid DNA is replicated**, and packaged in new viral progeny. All the viral progeny, which are recombinant phages are capable of infecting other bacteria (transducing particles).

Bacterial Cell 2



(5) **the recombinant phage** subsequent infects another new host cell (recipient cell)

(6) New alleles can be incorporated into the bacterial genome of recipient cell by:

i) crossing over and homologous recombination at homologous regions – refer to Fig. 7.

OR

ii) integration of the phage-bacterial hybrid DNA into genome of recipient cell, as defective phage enters the lysogenic cycle (shown in the figure above)

Fig. 9 - The process of specialised transduction

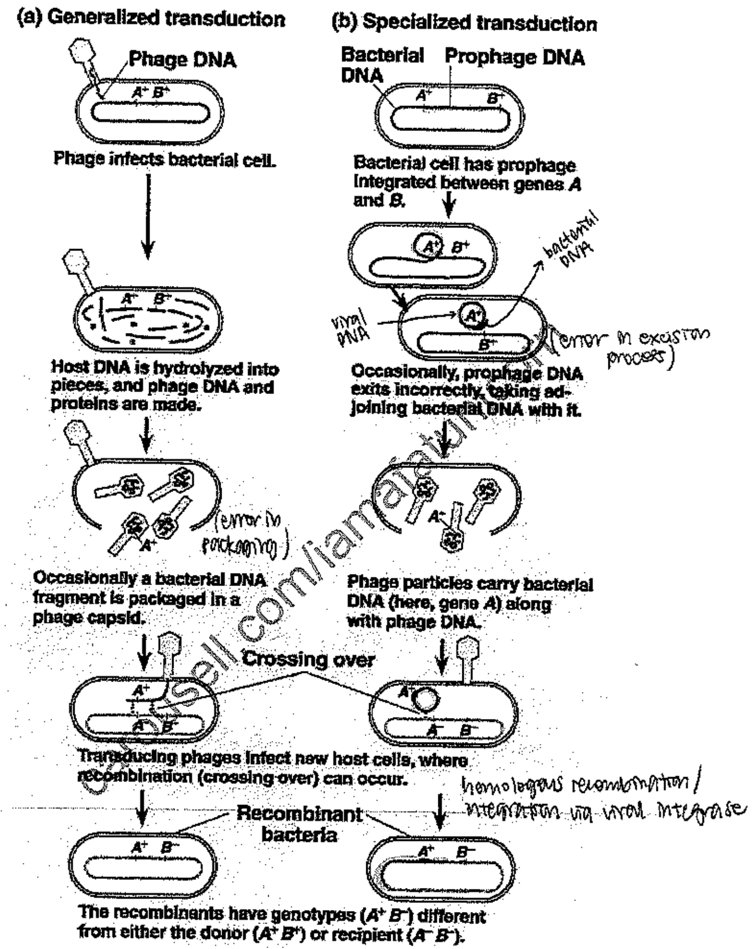
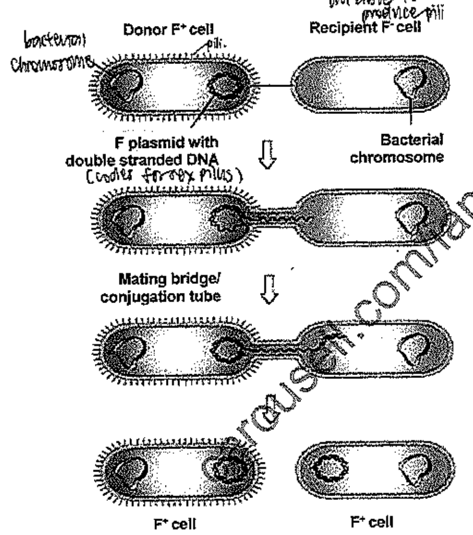


Fig. 10 - The process of generalized and specialised transduction

(3) Conjugation

- Conjugation refers to the **direct transfer** of genetic material from **one bacterial cell to another**, through a **temporary link** between the two cells (refer to Fig. 11).
- The transfer of DNA is always one-way → from **donor cell** (called an F⁺ cell) to **recipient cell** (called an F⁻ cell) in that it possesses an **F plasmid**.
  - → On the F plasmid is a segment of DNA called an **F factor** (F = fertility) that carries **genes coding for sex pili.**
  - → Due to the presence of F factor, the donor cell is able to produce appendages called **sex pili** to attach itself to the recipient cell.

Process :



(1) Conjugation begins when **sex pilus** on the surface of an **F⁺ cell** make contact with an **F⁻ cell.**

(2) After the sex pilus attaches, it retracts, pulling the two cells closer. The hollow pilus then forms a temporary **mating bridge/conjugation tube.** A **single strand of plasmid DNA** passes through the bridge and is replicated via the **rolling circle DNA replication** (see below)

(3) The single strand of plasmid DNA in the recipient cell recircularises and serves as **template** ~~serves as~~ for the synthesis of a complementary daughter stand

(4) Through semi-conservative replication, both cells now contain a double-stranded F plasmid. **Both cells are now F⁺.**

Fig. 12 - Conjugation involving an F⁺ and F⁻ cell; recipient cell becomes F⁺

- Bacterial conjugation can occur between bacteria of the same species or of different species. Success rate decreases with decreasing relatedness between species.

Rolling circle DNA replication

- For the transfer to take place, **one strand of the double-stranded F plasmid is nicked by a nuclease.**
- The free **3' end** of the nick is **extended by DNA polymerase** for the synthesis of a **new complementary strand** using the intact strand as the template.
- The **newly synthesized strand displaces the nicked strand which is transferred concurrently**, via the 5' end, across the mating bridge into the recipient cell.

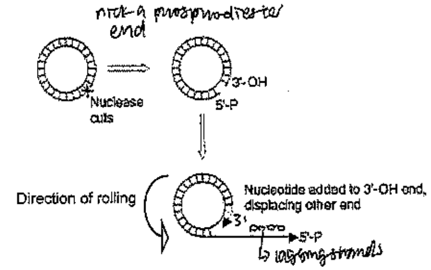


Fig. 11 - Rolling circle DNA replication

- Upon completion of a unit length of the plasmid DNA (after 1 round), **another nick occurs to release the original strand** and end the replication of the newly synthesized strand;
- This form of DNA replication for the plasmid is known as **rolling circle DNA replication** and is also found in some phage genomes. In this case, it is coupled to the transfer of DNA to another cell.
- In the **recipient cell**, the **single strand of F plasmid DNA re-circularises and serves as a template** for the synthesis of a complementary daughter stand

Points of comparison	Transformation	Transduction (Generalised)	Conjugation
Source of DNA	Foreign naked DNA	Bacterial cell infected by virulent phage	F⁺ cell containing F plasmid
Agent mediating DNA transfer	Cell surface proteins or treatment to make artificially competent	Bacteriophage	F factor on F plasmid codes for formation of sex pili & subsequent cytoplasmic mating bridge
Type of DNA transferred	Random fragments of the bacterial genome, usually closely related species	Random fragments of the bacterial genome small enough to fit into phage capsid; usually closely related species	F plasmid
Homologous recombination needed for permanent expression of foreign genes?	Yes	Yes	No

For info (but it is recommended that you have a brief understanding of Hfr cells)

Conjugation involving an Hfr cell (refer to Fig. 13):

- Sometimes, a donor cell's **F plasmid** can be **integrated** into the **bacterial chromosome**.
- Such a cell is called an ***Hfr* cell** (for *High frequency of recombination*); it will similarly function as the donor during conjugation.
- During transfer, DNA replication is initiated at the Ori within the F factor DNA.
- A single strand of DNA moves into the F⁻ cell.
- Usually, the fragile mating bridge is broken before an entire strand of bacterial chromosome and the rest of the F factor can be transferred to the recipient cell.
- The single strand of DNA serves as the template for the synthesis of a new complementary DNA strand in both cells.
- If part of the newly acquired DNA crosses over with a homologous region on the F⁻ chromosome, a **recombinant F⁻ cell** will result.

Question: Conjugation has sometimes been called 'bacterial sex'. However, this term is misleading. Can you explain how conjugation is different from sexual reproduction?

Process does not involve equal contributions of genetic material from 2 gametes. No offspring are produced. Rather, it is a form of genetic transfer where the genetic composition of the recipient cell is altered.

Question: Describe the differences which occur during conjugation when the donor cell is an F⁺ cell, compared to an Hfr cell.

Point of Comparison	F⁺ Cell	Hfr Cell
Type of DNA transferred to recipient	The entire strand of F plasmid DNA is transferred across.	Part of the F plasmid DNA and some neighbouring bacterial chromosomal DNA is transferred across.
Subsequent change in genotype of recipient cell	The recipient changes from F⁻ to F⁺	The recipient remains F⁻ but can still become a recombinant.

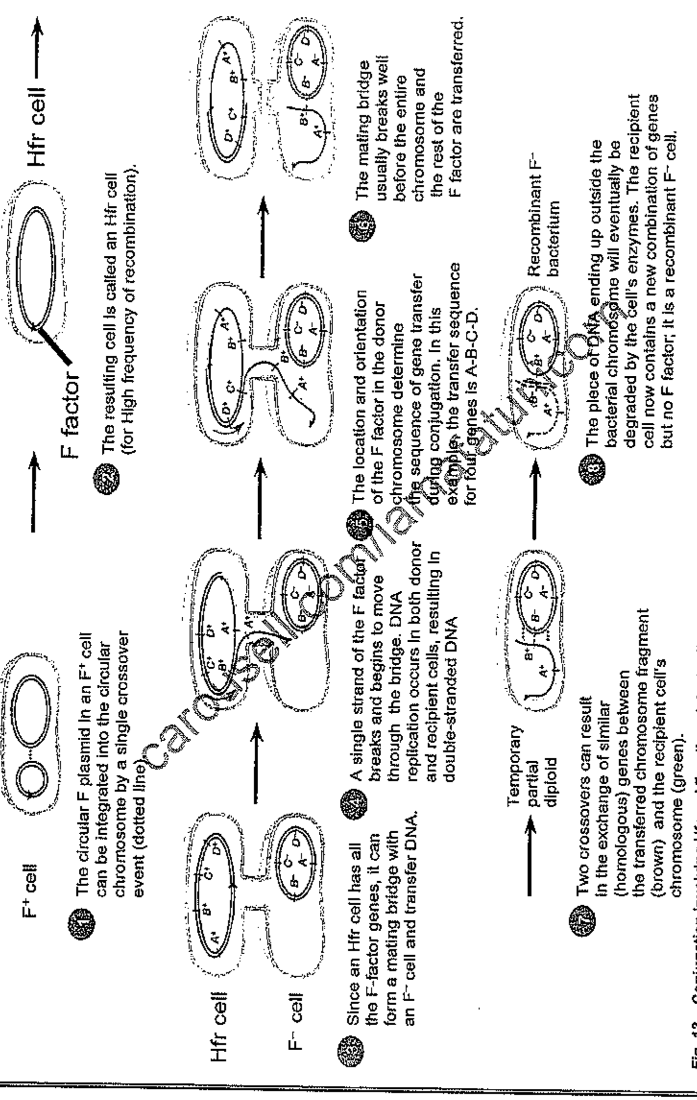


Fig. 13 – Conjugation involving Hfr and F⁻ cell; recipient cell remains F⁻ at the end of the conjugation process

LO 2i Explain how gene expression in prokaryotes can be regulated, through the concept of simple operons (including lac and trp operons), including the role of regulatory genes; and distinguish between inducible and repressible systems. (Attenuation of trp operon is not required).

D) Introduction to Gene Regulation in Bacteria

- All somatic cells of an organism carry **identical genes**. Despite this, cells in a **multicellular organism show a wide variation in structure and function.** For example, a liver cell have different structure and function from a white blood cell even though they share the same genetic makeup.
- Even within a single cell, the rate at which certain protein molecules are synthesized varies according to **circumstances and demand.**

Question: Why do different cell types have different structure and function?

- *This is a result of regulation of gene expression.*
- *In each specialized cell type, **certain sets of genes are expressed**, hence **certain set of tissue-specific proteins are synthesized**, which determine the **specialized function** of the cell.*
- *Hence certain subsets of the total genetic information are expressed in any given cell, allowing cells to **specialise.***
- *Gene regulation can also be **influenced by the environment,** allowing the cell to be **responsive** to changes in the environment.*

- Some proteins are **synthesized continuously at a constant rate** and genes coding for such proteins are said to be **constitutively** expressed.
  e.g. genes that make ribosomes, etc. ↳ expressed all the time (e.g. housekeeping genes)
- However not all proteins are constitutively expressed. The expression of other genes are **regulated** and there are several mechanisms by which this is done:
  ↳ Regulating gene expression in order to survive → cope with changing environment.

Legend: ↑ = *point of control*

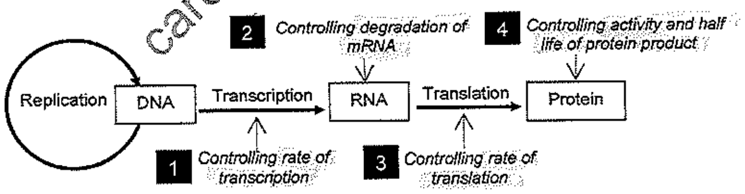


Fig. 14 - The Central Dogma and various points of gene regulation

Q: Can you suggest which level of gene regulationcontrol predominates? Why?

Transcriptional-level
↳ efficient with minimal wastage

*[For this topic we will be focusing on **regulation at the level of transcription.**]*

- in Eukaryotes, each gene is transcribed separately
- In prokaryotes | bacteria → all the genes that will collectively serve a function are transcribed together.

(E) The *lac* Operon in *E. coli*

(1) Background

- *Escherichia coli* (***E. coli***) is a bacterium commonly found in the **intestines** of humans and other mammals.
- *E.coli* living in the colon of an adult cow is NOT normally exposed to the milk sugar lactose, a disaccharide. However, the *E. coli* living in a calf will be exposed to lactose from the mother's milk.
- This situation presents a dilemma.
  - Should a bacterial cell invest energy and materials to produce lactose-metabolizing enzymes *just in case* it ends up in the digestive system of a calf?
  - Given that the average life span of an actively growing *E. coli* cell is about 30 minutes, such an evolutionary strategy appears wasteful. And yet, if *E. coli* cells cannot produce those enzymes, they might starve in the middle of an abundant food supply.
- This dilemma is overcome because *E. coli* is able to **regulate the expression of genes coding for various enzymes,** thereby allowing them to make use of available organic molecules efficiently.

Question: What benefit is there for the bacteria to regulate its genes?

- *Enables bacteria to **respond appropriately and rapidly to changes in the environment.***
- *Ability to do the above confers a **selective advantage** to such bacteria over those that can't regulate gene expression*
- economical use of energy & resources

(2) Organisation of an operon

- The basic mechanism for this type of control of gene expression in bacteria, described as the operon model, was discovered in 1961 by Nobel prize winners, Francois Jacob and Jacques Monod at the Pasteur Institute in Paris.

  Their work involved the use of bacteria ***E. coli.* mutants**. Through their investigations, they found out that genes involved in **lactose metabolism** in *E. coli* were clustered together in a region of the bacterial chromosome known as the **operon.**

- An **operon** is a **cluster of genes with related functions**, regulated in such a way that all the genes in the cluster are turned on and off together (see Fig. 14). It includes a common **promoter**, an **operator**, and **more than one structural genes** that are **controlled as a unit** to produce a **single polycistronic messenger RNA** (mRNA).

  → **Operator** is a site on DNA at which a repressor protein binds to prevent transcription from initiating at the adjacent promoter.

  → **Regulatory gene** codes for a protein that is involved in the regulation of the expression of other genes eg. repressor, CAP.

  → **Structural gene** is any gene that codes for a protein (or RNA) product that forms part of a structure or has a metabolic function, e.g.enzyme.

- Operons occur primarily in **prokaryotes** such as *E. coli* and certain simple eukaryotes e.g. nematodes.

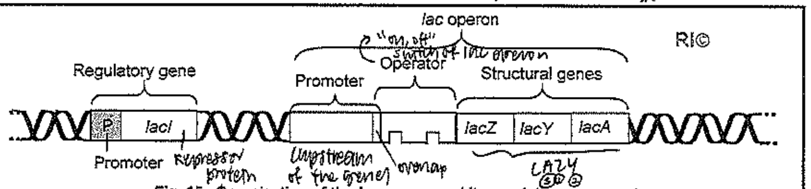


Fig. 15 - Organisation of the *lac* operon and its regulatory gene upstream

Structural Gene	*lacZ*	*lacY*	*lacA*
Enzyme coded for	β-galactosidase	Permease	Transacetylase
Function	Hydrolyses lactose to glucose and galactose	A membrane transport protein that enables cells to take up lactose efficiently	Function not well known. Metabolises certain disaccharides.

Table 1. Genes found within the *lac* operon

Organisation of the *lac* operon (an inducible operon) in *E.coli*:

i) Within the *lac* operon are three structural genes arranged in the sequence ***lac Z, lac Y* and *lac A*** as seen in Figure 14. The enzymes which they encode are given in Table 1.

ii) The three structural genes are under the control of one common **promoter** sequence upstream **for RNA polymerase to bind and initiate transcription.**

iii) The **operator** is located between the promoter and the structural genes to **control the transcription of the structural genes by controlling access of RNA polymerase to the genes**, therefore turning the genes "on" or "off" like a switch.

- A short distance upstream of the operon is its **regulatory gene, *lacI*** (pronounced "lac-i") which has its own promoter and terminator sequences.
  → The ***lacI* gene** codes for the ***lac* repressor protein.**
  → Note that the regulatory gene is normally **NOT** considered to be part of the operon.

- The *lac* operon is an example of an **inducible operon** because **the expression of the 3 genes** (***lac Z, lac Y* and *lac A***) is usually "off" but can be **induced** and hence is turned "on" in the presence of an **inducer** molecule. In fact, ***lacI*** is named as such because "*I*" stands for "inducibility".

  → The **inducer** molecule for the *lac* operon is **lactose**, or more accurately, its isomer **allolactose**.
  → Lactose that is transported into the cells and converted into the inducer, allolactose.

(3) Regulation of the expression of the *Lac* operon

Since the Lac operon produces enzymes involved in the metabolism / hydrolysis of lactose to glucose (to be used as respiratory substrate) and galactose, it makes sense that Lac operon is expressed in the (i) presence of lactose & (ii) absence of glucose.

Hence the expression of the *Lac* operon is regulated by :
(i) **Negative regulation**: turned off by **repressor protein** in the absence of lactose.
(ii) **Positive regulation**: upregulated by **the CAP protein** in the absence of glucose.

(i) NEGATIVE REGULATION

(a) Default mode of *lac* operon (absence of lactose & glucose)

- By default, the *lac* operon is considered **repressed** i.e."off"

1. The **regulatory gene *lacI***, is **constitutively transcribed**, resulting in continued production of small amounts of the **lac repressor protein** (see Fig. 16).

2. The **repressor protein** is produced in the **active form and binds specifically to the *lac* operator** sequence via its **DNA-binding site.**

3. **In the absence of lactose,** the **repressor** binds to the **operator site, denying RNA polymerase access to the promoter.**

4. Transcription of the structural genes of the *lac* operon is hence blocked. This has the effect of switching the ***lac* operon off** (i.e. the operon **is repressed**).

- However, the binding of the repressor to the operator is mediated by weak interactions. As such, the repressor sometimes dissociates from the operator, resulting in a basal level of lac operon products i.e. galactosidase, permease and transacetylase within the cell, and whose presence is equally important and necessary for the regulation of *lac* operon as shall be discussed later.

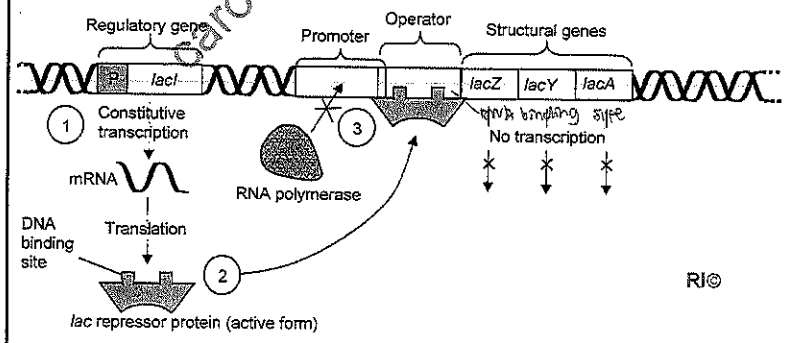
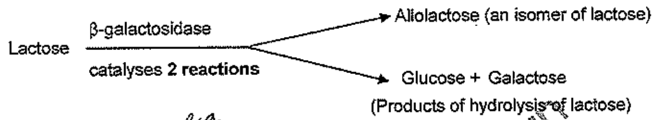


Fig. 16 - Repression of the *lac* operon by the repressor molecule, in the absence of lactose

(b) *lac* operon in the presence of lactose

- The *lac* repressor protein contains another functional region apart from its DNA-binding site. Known as the **allosteric site**, for the specific binding of **allolactose**, a structural isomer from lactose.

(1.) In the presence of **lactose**, a few molecules of lactose will enter the cell with the help of permease and are converted to allolactose by the few β-galactosidase molecules present. (Lactose is sometimes called an **inducer / effector** molecule as it causes a response.)



Even though the lac operon may be repressed, the repression is somewhat leaky/not 100% efficient, resulting in a low basal level of lac operon products i.e. permease, β-galactosidase (and transacetylase) within the cell.

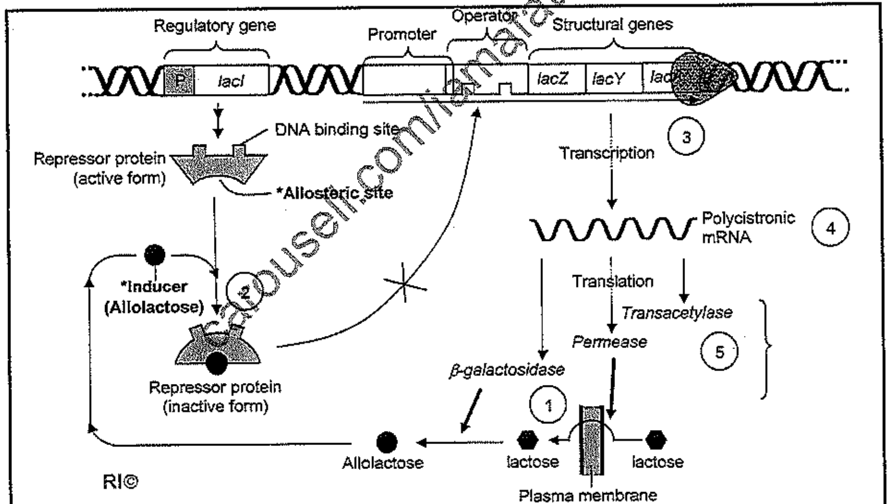


Fig. 17- The action of inducer relieving the repression of *lac* operon in the presence of lactose

(2.) Binding of allolactose to the **allosteric site inactivates the repressor** by altering the tertiary structure of the repressor so that its **DNA-binding site is no longer complementary to and cannot bind to the operator.**

(3.) **RNA polymerase** can **access and bind** to the **promoter** to initiate the transcription of the structural genes of the operon.

(4.) The structural genes are transcribed as a single **polycistronic mRNA**.
Polycistronic mRNA is a messenger RNA that contains the base sequence coding for the amino acids sequence of several proteins.

(5.) All three enzymes are translated from a single mRNA molecule. Thus, all the genes in an operon are always expressed (or not expressed) in unison.
(The enzymes are translated separately because each has its own start and stop codon on the mRNA.)

- The *lac* operon is thus an **inducible operon** that exhibited **negative gene regulation** (by default), a regulatory mechanism in which the **DNA-binding regulatory protein is a repressor that turns off transcription of the gene(s).**

- Inducible genes or operons usually code for enzymes that are part of **catabolic pathways**, which **break down** molecules. Hence the enzymes are expressed on only when lactose if present

(ii) POSITIVE REGULATION

- So far we studied how lactose regulates the expression of the *lac* operon but that is only part of a big picture. A second metabolite, **glucose**, is also **involved in the regulation of the *lac* operon.**

- If you look at the metabolic pathway in Fig.18, all sugars are converted to glucose before they enter the respiratory pathway to yield energy in the form of ATP.

-

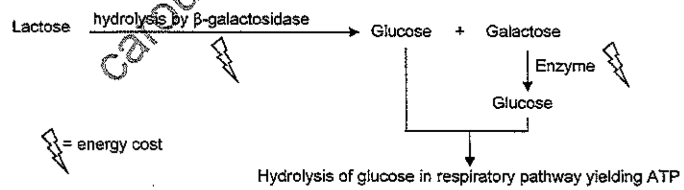


Fig. 18 - The lactose metabolic pathway

- Since considerable energy expenditure is required to synthesize additional lactose-metabolising enzymes such as *β*-galactosidase, it makes more sense for *E. coli* cells to **utilise all available supplies of glucose first, before they start metabolising lactose.**
- So what regulatory mechanism ensures that the *lac* operon is "OFF" when glucose is present?

(c) *lac* operon in the presence of lactose and glucose

1. In reality, the **_lac_ operon promoter** has a **low affinity for RNA polymerase.** Therefore, even in the presence of lactose, which inactivates the repressor, the *lac* operon is not fully activated on its own is unable to fully activate the *lac* operon. (Fig. 19)

2. A **second regulatory mechanism** that is **sensitive to the presence of glucose** is involved in the regulation of the *lac* operon.

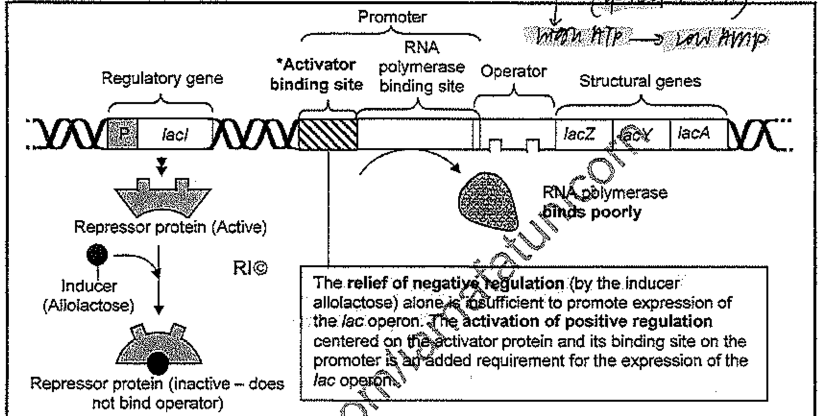


Fig. 19 - Efficient transcription of *lac* operon requires more than the presence of lactose

- **Positive gene regulation** of *Lac* operon therefore involves up-regulation by activator protein, **catabolic activator protein**, **CAP**, which binds to DNA and stimulate the transcription of the gene(s).

(d) *lac* operon in the presence of lactose but absence of glucose

- Similar to the repressor protein, the **activator protein** has a **DNA-binding site** and an **allosteric site.**
- The activator protein is the **catabolite activator protein (CAP).**
  - Its **DNA-binding site** allows it to bind to the **activator / CAP-binding site** situated within the **promoter.**
  - Its **allosteric site** is specific for binding of **cAMP, or cyclic AMP**, an alternative form of AMP (adenosine monophosphate).
  (CAP is thus sometimes referred to as cAMP receptor protein (CRP).

1. In the absence of glucose, cAMP levels will increase.

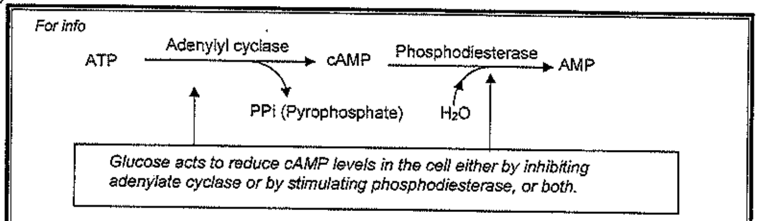
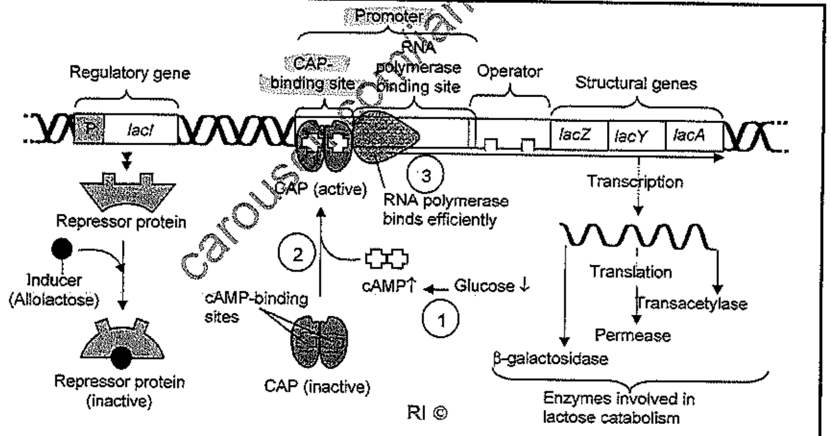


Fig. 20 - Relationship between AMP, glucose and cAMP

** (*You just need to know that ↑**glucose** → ↓**cAMP** & ↓**glucose** → ↑**cAMP***)

2. The **cAMP** binds to the **allosteric site** of CAP, **activating CAP** which binds to the **CAP-binding site within the promoter.**

3. This binding of activated CAP **increases the affinity of the promoter region for RNA polymerase.** Thus, **the increasing the rate of transcriptional initiation** of the *lac* operon.



Scenario: **Lactose high, glucose low→cAMP high.** When glucose levels are low, each CAP polypeptide has cAMP bound to its allosteric site. The active form of CAP binds to the DNA at the CAP-binding site, activating transcription.

Fig. 21 -Transcription of *lac* operon requires both lactose as well as the absence of glucose

Summary

lactose (+/-)	Glucose (+/-)	Lac repressor (active/inactive)	CAP activator (active/inactive)	Lac operon (on/off)
–		ACTIVE		OFF
+	+	INACTIVE	INACTIVE	on
+	–	INACTIVE	ACTIVE	ON

- So for **lactose-metabolising enzymes** to be produced in **appreciable quantity**, it is not sufficient for **lactose** to be **present** in the bacterial cell. The other requirement is that **glucose** must be in **short supply**.
- This ensures that the preferred carbon source, glucose, is used before other alternative carbon sources are used i.e. if both glucose and lactose are present, rate of transcription of the operon genes will still be low.
- Thus the *lac* operon is under **dual control: negative regulation by the *lac* repressor and positive regulation by CAP.**
- Hence, the state of the *lac* repressor (with or without bound allolactose) determines whether the *lac* operon's genes undergo transcription or not; the state of CAP (with or without bound cAMP) controls the rate of transcription when the operon is repressor-free.
(It is as though the operon has both an on-off switch and a volume control.)

(F) The *trp* Operon (as an example of a repressible operon)

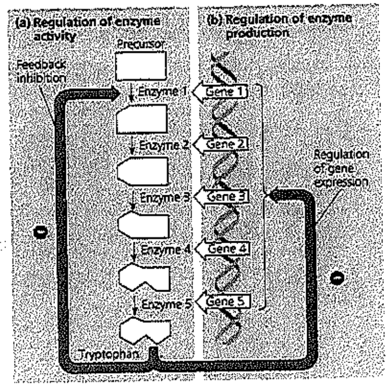
- The **tryptophan (*trp*) operon** (*trp* is pronounced "trip") is an example of a **repressible operon.**

Repressible operons in bacteria are associated mainly with **anabolic** pathways which involve the synthesis of amino acids, nucleotides etc. from simpler materials.

Repressible operons are normally turned "on" by default and they are turned "off" in the presence of an effector molecule. In most cases, this effector molecule is the **end product of the anabolic pathway** so as to avoid devoting resources to unnecessary synthetic activities once the end product has accumulated to sufficient amounts.

- Gene products of the *trp* operon in *E. coli* are involved in the **synthesis of the amino acid tryptophan** which is essential for protein synthesis. Hence the operon is usually "ON". (Fig. 22)

However, when concentration of tryptophan rises e.g. when the host mammal consumes a protein rich diet, enzyme synthesis will be repressed. Since the ***trp* operon is turned 'off' in the presence of tryptophan,** tryptophan is thus the **effector molecule.**



- This process of regulation is called **end-product repression** or often just **repression.** This should be clearly **distinguished from end-product inhibition.** (Repression always occurs at the **level of transcription** of the enzyme while inhibition often involves **inhibition of enzyme activity.**)

- Tryptophan does not act directly on the operator but works together with a repressor molecule to repress the transcription of the *trp* operon.

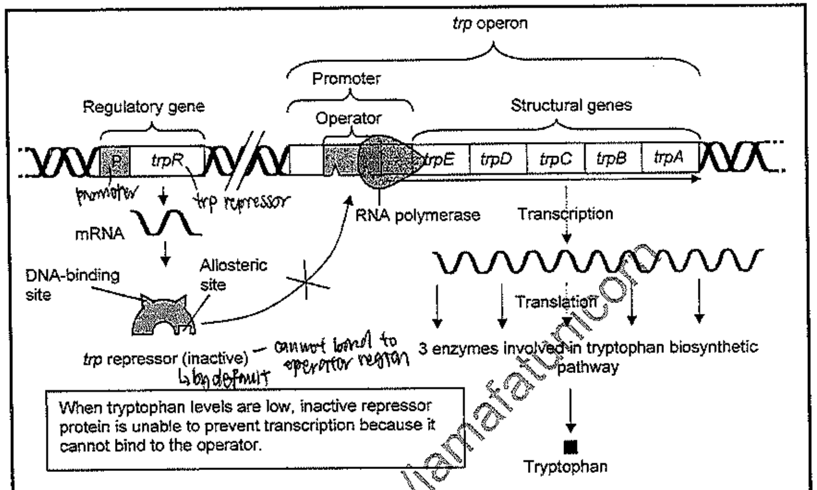


Fig. 22 - Organisation of the repressible *trp* operon; Default 'On'

1. The tryptophan repressor is **synthesised in its inactive form** with little affinity for the *trp* operator, which lies within the *trp* promoter. (See Fig. 21)
2. As tryptophan accumulates, it binds to the **allosteric site of the *trp* repressor, activating** the repressor.
3. **The activated repressor protein** binds to the operator at its DNA-binding site.
   - Tryptophan therefore serves as a **corepressor,** which works together with a repressor protein (by activating it) to switch an operon off.
4. With repressor bound to operator, RNA polymerase cannot bind to promoter and transcription cannot proceed, hence turning the operon off.

Qn: Which of these contain nucleotide bases?
1. operator } found in DNA.
2. promoter
3. repressor – protein in nature

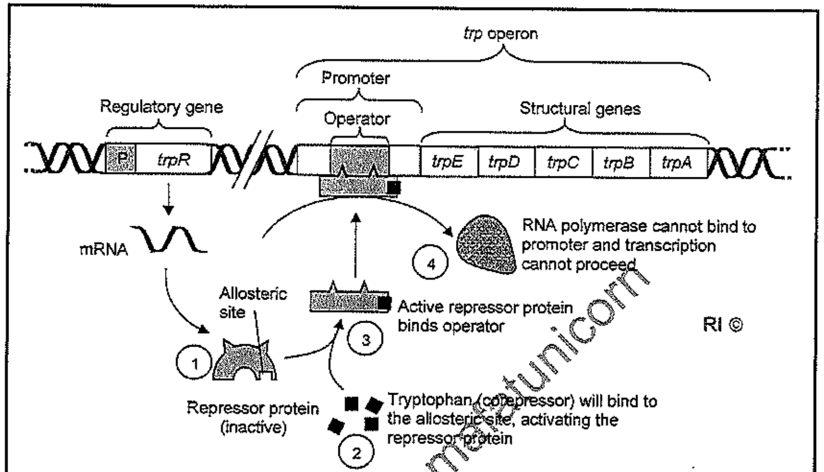
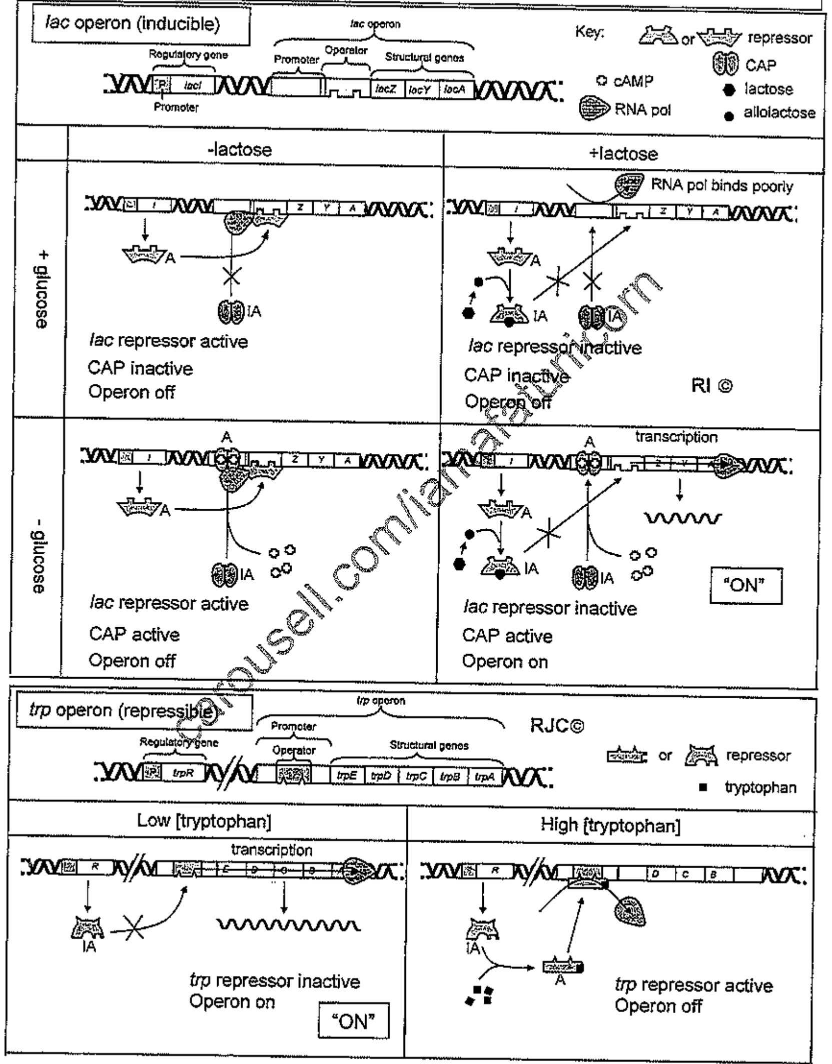


Fig. 23 - End-product repression by tryptophan

	Repressible operon	Inducible operon (consider just negative gene regulation)
Named example	trp operon	*lac* operon
Type of metabolic pathway	Anabolic pathways	Catabolic pathways
Effector molecule	Corepressor tryptophan, an end product	Inducer, lactose, the substrate
Effect of effector on operon	Turns off structural genes	Turns on structural genes
Repressor synthesised in active/inactive form	Inactive form	Active form
When does repressor bind operator?	When complexed with corepressor	On its own
When does repressor not bind operator?	On its own	When complexed with inducer
Default operon expression	ON	OFF
Operon expression when effector molecule is present	OFF	ON

Table 2. Comparison between a repressible operon and an inducible operon

(G) Summary of *lac* and *trp* Operons



Note: A = active; IA = inactive

(H) Glossary

Anabolic pathway: *Series of reactions that results in the synthesis of one of more specific cellular components.*

Catabolic pathway: *Series of reactions that results in the degradation of one or more specific cellular components.*

Colon: *Lower and larger part of the intestine.*

Constitutive gene: *Refers to a gene that encodes a product required in the maintenance of basic cellular processes or cell architecture. Also known as housekeeping genes. They are expressed all the time.*

Inducible enzymes: *Enzymes for which synthesis is regulated by the presence or absence of its substrate.*

Metabolic pathway: *Series of enzymatic reactions that convert one molecule to another via a series of intermediates. It is the sum of* ***catabolic**** *and* ***anabolic**** *pathways.*

Motility: *The ability of an organism to move by itself, different from mobility.*

Operator: *A site on DNA at which a repressor protein binds to prevent transcription from initiating at the adjacent promoter. (Genes VII, Lewin)*

Polycistronic mRNA: *A messenger RNA that contains the base sequence coding for the amino acids sequence of several proteins.*

Recombination: *The formation of a new combination of genes on a chromosome as a result of crossing over.*

Regulatory gene: *Any of several kinds of nucleotide sequence involved in the control of the expression of structural genes. It codes for a protein involved in regulating the expression of other genes eg. repressor,CAP. (Genes VII, Lewin)*

Repressible enzymes: *An enzyme whose synthesis is regulated by the presence or absence of a specific metabolite.*

Structural gene: *Any gene that codes for a protein (or RNA) product that forms part of a structure or has an enzymatic function. (Genes VII, Lewin. Molc bio of the Cell, Bruce Alberts)*

Terminator: *A regulatory sequence that signals the end of transcription*

(I) LINKS

The topic of cell structure is relevant to the following topics and learning outcomes in the A level Biology syllabus. The links also become clearer when you have gone through the other topics.

Topic	Topic	How it is linked to Bacteria
1	Cell division	Binary fission is a means by which bacterial DNA replicates and cell divides. This is not to be confused with mitosis and cytokinesis.
2	DNA and genomics	DNA replication occurs during binary fission.
3	Virus and bacteria	Generalised and specialized transduction.
4	Prokaryotic and eukaryotic genome / Control of gene expressions	Bacteria are prokaryotes. Gene expression in bacteria using the operon system is regulated at the level of transcription.
5	Immunology	The mode of transmission and infection of bacterial pathogen. The modes of action of antibiotics, including penicillin, on bacteria.

Some Keywords

Structural gene	Generalised transduction	*lac* Z	*lac* repressor
Sex pilus	Specialised transduction	*lac* A	Negative gene regulation
Mating bridge/ conjugation tube	Prophage	*lac* Y	Positive gene regulation
F⁺ plasmid	Template	*lac I*	Permease
Regulatory gene	Conjugation	*lac* operon	Transacetylase
Constitutive	Homologous recombination	*trp* operon	β-galactosidase
Peptidoglycan cell wall	Recombinant	Repressible operon and enzyme	Allolactose/lactose
Circular DNA	Lytic cycle	Inducible operon and enzyme/s	Allosteric site
70S ribosomes	Lysogenic cycle	Operator	Catabolite activator protein (CAP)
Binary fission	Transformation	Promoter	CAP binding site
Regulatory gene	RNA polymerase	Polycistronic mRNA	cAMP
Inducer/ Co-inducer	Co-repressor	Catabolic	Anabolic

CORE IDEA:
(2) GENETICS AND INHERITANCE
ORGANISATION AND CONTROL OF PROKARYOTIC AND EUKARYOTIC GENOME (I)

ORGANISATION OF PROKARYOTIC AND EUKARYOTIC GENOME

Content
Genome organisation at the DNA level

Learning Outcomes
Candidates should be able to:

2 (d) Describe the structure and organisation of viral, prokaryotic and eukaryotic genomes (including DNA/RNA, single-/double-stranded, number of nucleotides, packing of DNA, linearity/circularity and presence/absence of introns).

2 (h) Describe the structure and function of non-coding DNA in eukaryotes (ie. portions that do not encode protein or RNA, including introns, centromeres, telomeres, promoters, enhancers and silencers). (Knowledge of transposons, satellite DNA, pseudo-genes and duplication of segments is not required.)

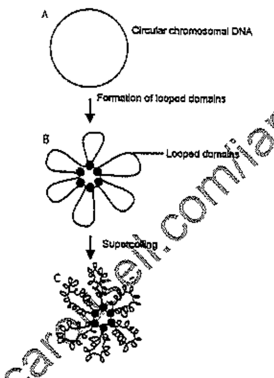
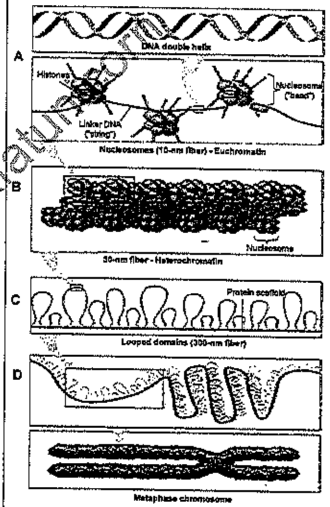
References
Campbell N.A. and Reece J.B. (2008) Biology, 8th Edition, Pearson Education, Inc.
Brooker R.J. (2005) Genetics: Analysis and Principles, 2nd Edition, McGraw-Hill.
Brooker R.J., Widmaier E.P., Graham L.E. and Stiling P.D. (2008) Biology, McGraw-Hill.
Lewin, B. (2008) Genes IX. Pearson Education, Inc.
Raven P., Johnson G., Losos K., Mason L. and Singer S. (2008). Biology, 8th edition. McGraw-Hill.
Sadava D., Heller H.C., Orians G.H., Purves W.K. and Hillis D.M. (2008) Life: The Science of Biology. Sinauer Associates, Inc.

This handout is the effort of Biology teachers at RI. It has been and will continue to be updated.

<u>Table of Contents</u>

(A) Comparing Structure & Organisation of Prokaryotic & Eukaryotic Genomes

Note: ***Genome*** *refers to a complete set of genetic material in a particular cellular component.*

Feature	Prokaryotic genome	Eukaryotic genome
Size	Smaller (10^5-10^7 base pairs)	Larger (10^7-10^{11} base pairs)
Appearance	Generally **a single, circular** molecule	**Multiple, linear** molecules
Molecule	Double Helix DNA	
Association with proteins	Yes – relatively less e.g. **histone-like proteins**	Yes – large amounts e.g. **histones**, scaffold proteins
Level of DNA packing/coiling	Relatively low: DNA double helix ↓ some looping around histone-like proteins  (A) Unfolded chromosome from *E. coli* has a diameter of 430μm. (B) DNA is folded into **chromosomal looped domains by protein-DNA associations**. Six domains are shown, but actual number is about 50. (C) **Supercoiling** cause further compaction, such that it fills an area of about 1 μm.	High:  (A) DNA double helix is associated with proteins called **histones**. DNA molecules are **negatively-charged**, histone are **positively-charged**. Thus, the DNA molecule is held around histones by **electrostatic interactions**. Most of DNA is wound around octamers of 8 histone proteins to form **nucleosomes**, the 10 nm fibre. Remainder of DNA, called **linker**, joins adjacent nucleosomes. (B) The 10-nm fibre coils around itself to form a **30 nm chromatin fiber** (or **solenoid**).

		(C) The 30-nm fibre forms loops called **looped domains (a 300-nm fibre)** when associated with scaffold proteins. (D) **Supercoiling** present. The loops further coil and fold to produce characteristic metaphase **chromosome.**
Location	**Nucleoid** region, not membrane-bound	**Nucleus** (surrounded by nuclear envelope)
Extrachromosomal DNA	Present – **plasmids** (much smaller rings of DNA compared with bacteria chromosome)	**No plasmids** (however mitochondria and chloroplast have their own DNA)
Number of genes	Fewer genes (e.g 4,500)	Many genes (e.g. 25,000)
Non-coding regions (between and within genes)	Not common – typically less than 15%	Common – about 98%
i. **introns**	None present (rare; only in some genes)	Many present
ii. **promoters**	Present	
iii. **enhancers/ silencers**	Rarely present	Present
iv. repeated sequences	Few	Many
v. **operons**	Many	Few known ones (e.g. in nematodes)
Origin of replication (per chromosome)	One	Many

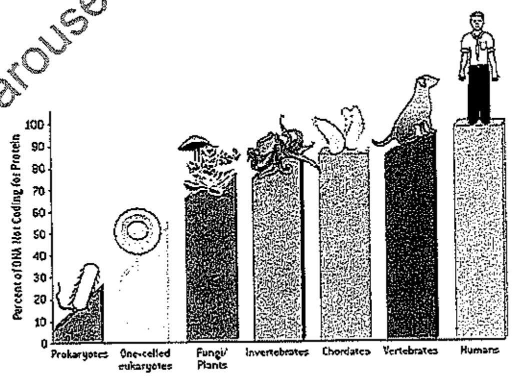


Figure 1. Non-coding sequences across species.
[http://fig.cox.miami.edu/~cmallery/150/gene/non-coding.genes.jpg]

(B) Non-coding DNA in the Eukaryotic Genome

- Includes any part of a genome that **does not code for proteins or RNA products** (i.e. rRNA, tRNA etc).
- Was originally referred to as 'junk DNA'. However, we now know that non-coding DNA can play important roles in the eukaryotic cell (e.g. regulatory functions). Also, non-coding DNA may have as-yet undetermined functions that are possibly important. This explains their presence in genomes over hundreds of generations. Hence, you should avoid using the term 'junk DNA' and use the term non-coding DNA instead.
- Forms **a large component of the eukaryotic genome** (Figs. 1 and 2). Note, however, that the size of a genome (i.e. amount of DNA) of an organism is not always proportional to its complexity (i.e. number of genes). e.g., the lily plant has 18 times more DNA than humans, but produces fewer proteins than us.

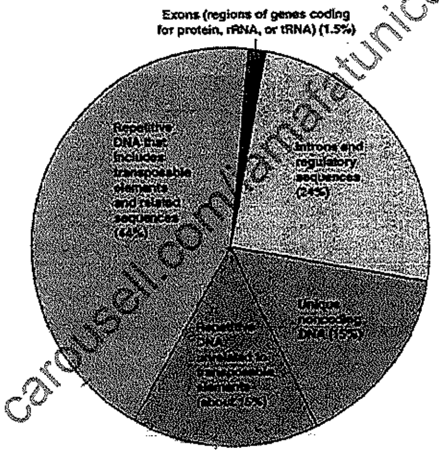


Figure 2. Types of DNA sequences in the human genome.
A staggering 98.5% of the human genome comprises of non-coding DNA.

- Most of the non-coding sequences consist of **repeated sequences.** That means they can be found **a few hundred to a few million times all over the genome.**
- Repetitive sequences can be found grouped together in a tandem array/tandem repeats. **A typical tandem repeat consists of a short nucleotide sequence repeated consecutively many times in a row.**

  e.g.: In *Drosophila*, tandem repeats forms 19% of the genome. An example of a tandem repeat, which consists of sequence AATAT is: **AATAT**AATAT**AATAT**AATAT**AATAT**

1. Introns

Structure

- Are **non-coding sequences found within a gene.**
- Are found **between exons in a molecule.**
  - Exons are the coding regions of DNA in a gene that codes for proteins or RNAs.

Function

- **Introns** have **no involvement in the translation** of an mRNA and hence are **excised**.
- In eukaryotes, after transcription, RNA processing takes place on the pre-mRNA. **Introns are excised** and the subsequent **joining of exons** form a mature RNA. **Splicing refers to both the excising of introns and the joining of exons** (Fig. 3).
  - Involves a **spliceosome**, a snRNA (small nuclear RNA)-protein complex. Points of excision are very precise and are determined by the sequence of nucleotides at the intron-exon boundaries.

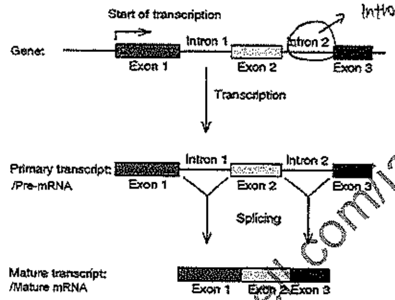


Figure 3. Outline of splicing. The introns in a gene are transcribed along with the exons (boxes) in the primary transcript/pre-mRNA. Then, introns are excised and exons are joined together.

- **Alternative RNA splicing of a single pre-mRNA can produce different mature mRNA,** depending on which combination of exons is spliced together. This means that **one gene can code for more than one type of polypeptide.**
- An advantage of alternative splicing is that it enables a larger number of proteins to be produced relative to the number of genes. Alternative splicing produces **different protein isoforms from one gene** (Fig. 4). Isoforms are alternative forms of the same protein. About 40-60% of human genes produce multiple proteins in this way.

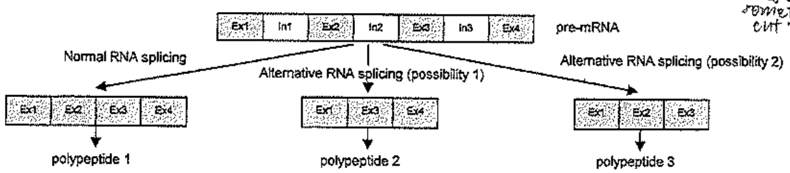


Figure 4. Three of the various ways of alternative RNA splicing. The different mRNAs are translated into different proteins.

- E.g. Alternative processing of the calcitonin gene transcript in rats yields two different protein products, calcitonin-gene-related peptide (CGRP) in the brain and calcitonin in the thyroid (Fig. 5). The pre-mRNA transcript has two poly(A) sites; one predominates in the brain, the other in the thyroid. In the brain, splicing eliminates the exon 4 while in the thyroid, exon 4 is retained.

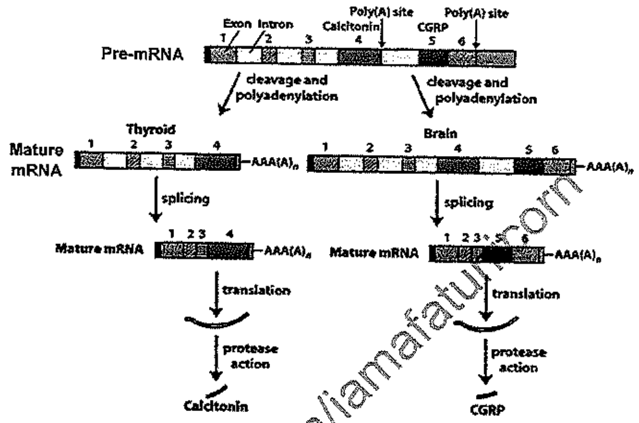


Figure 5. Alternative splicing of pre-mRNA to produce different protein isoforms.

Distribution

- In eukaryotes only

Why do eukaryotes bother to have introns?
↳ Evolutionary Advantage : Alternative splicing

2. PROMOTERS – control element (regulatory sequence)

Structure
- Are **non-coding** DNA sequences located **just upstream of the transcription start site** of a gene.

Function
- Promoters are DNA sequences which serve as a **recognition site** for the **binding of general transcription factors and RNA polymerase to start / initiate transcription.** Sequences within the promoter determine the **strength of the promoter** which in turn, determines the binding efficiency of RNA polymerase and hence the **frequency of transcription**.

- How do the sequences within a promoter regulate gene expression?
  - There are similar sequences known as **critical elements/short sequences** (e.g. at the **-10 and -35 sites** in bacterial promoter) found in promoters of different genes.
  - The **consensus sequence** was derived from the most commonly occurring bases within critical elements/short sequences of different promoters.
  - The **more the critical elements/short sequences in a given promoter resemble the consensus sequences, the greater the binding efficiency between RNA polymerase and the promoter**. This will lead to an **increase in the frequency of transcription.**

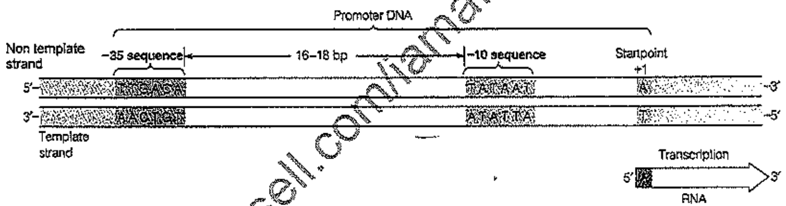


Figure 6. Conventional numbering system of promoters. Positions of -10 and -35 sequences are shown relative to the transcription start site for a typical bacterial promoter. The first nucleotide where transcription begins is designated +1 and the preceding nucleotide is -1. There is no zero.

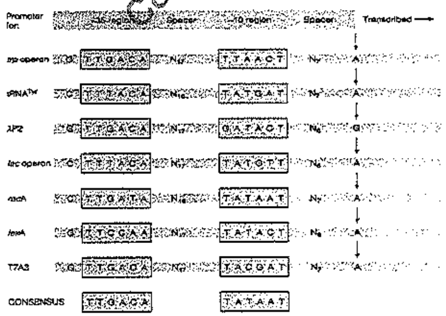


Figure 7. -35 and -10 sequences within a variety of bacterial promoters. The consensus sequence is shown at the bottom. Spacer regions are shown – for example, N_{17} means 17 nucleotides between the end of -35 region and start of -10 region.

Distribution
- In prokaryotes and eukaryotes

3. ENHANCERS & SILENCERS – Control element (regulatory sequence)

Structure
- **Enhancers** and **silencers** are short non-coding DNA sequences which are **usually located far away** (up to thousands of nucleotides upstream or downstream) from the promoter, but can also be found within an intron or near the gene to be controlled.
- They are regulatory elements.

Function

Enhancers
- Enhancers are DNA sequences that help to **increase frequency of transcription.**
- Enhancers, when bound with **specific transcription factors** known as **activators**, **increase frequency of transcription** by **promoting assembly of transcription initiation complex.**

Silencers
- Silencers are DNA sequences that help to **reduce efficiency of transcription.**
- **Silencers,** when bound with **specific transcription factors** known as **repressors, decrease frequency of transcription** by **inhibiting assembly of transcription initiation complex.**

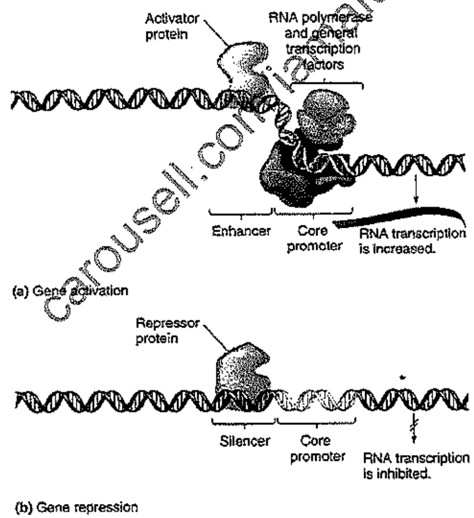


Figure 8. Overview of transcriptional regulation at (a) enhancer and (b) silencer.

Distribution
- In eukaryotes only

4. Telomeres and Telomerase

- All eukaryotic chromosomes must have 3 types of DNA sequences for it to be functional, to maintain its structural integrity and be passed on to subsequent generations.
  They are the **1) origins of replication** (to initiate DNA replication)
  2) telomeres, and
  3) a *centromere** (Fig. 14).

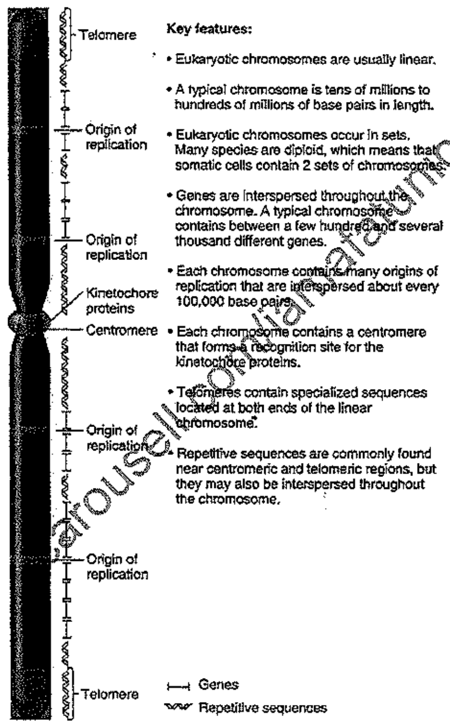
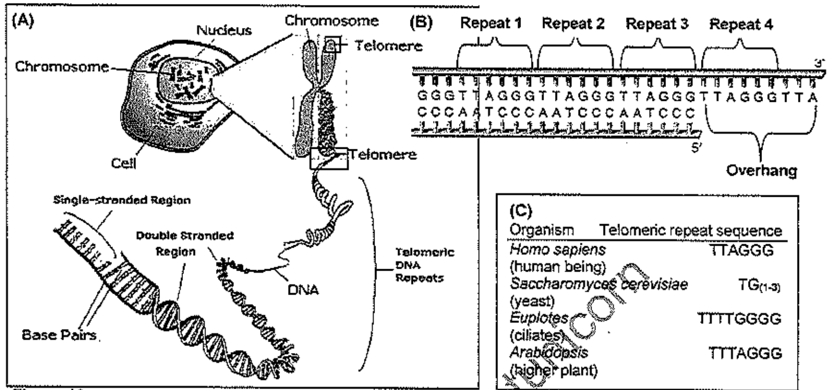


Figure 9. Eukaryotic chromosome.

Telomeres



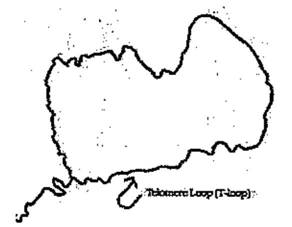
(C)

Organism	Telomeric repeat sequence
Homo sapiens (human being)	TTAGGG
Saccharomyces cerevisiae (yeast)	$TG_{(1-3)}$
Euplotes (ciliates)	TTTTGGGG
Arabidopsis (higher plant)	TTTAGGG

Figure 10.
(A) Telomeres on a chromosome.
[http://www.chemsoc.org/ExemplarChem/entries/2003/imperial_Burgoine/origins.txt.html]
(B) End of the telomere, and the single-stranded region called the 3' overhang. Four repeats of the sequence TTAGGG are shown. In reality, human telomeres are thousands of nucleotides long.
(C) Telomeric repeat sequences of different species.

Structure (Fig. 10)

- **Telomeres** are nucleotide sequences found **at both ends of eukaryotic chromosomes**.
- Telomeres are **non-coding** regions of DNA made up of a series of **tandem repeat sequences** (multiple repetitions of one short nucleotide sequence) where each repeat is about 5-10 nucleotides long.
- Telomeres have a **single stranded region** of DNA at their 3' ends known as the **3' overhang**. This region of DNA does not have a complementary strand.
  - The 3' single-stranded end loops back and displaces the same sequence in the upstream region of the telomere and binds to the complementary sequence of the other strand. This process is brought about by special telomere-binding proteins.



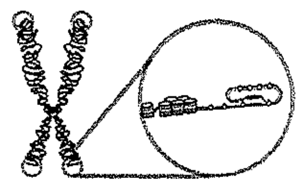


Figure 11 and 12. Formation of loop stabilizes terminal ends of chromosomes

Function

- **Role 1: Telomeres ensure genes are not lost/eroded with each round of DNA replication due to the end replication problem. This prevents loss of vital genetic information.**
  - What is the **end-replication problem?** (Fig. 13)
    - The end replication problem occurs during the replication of linear eukaryotic chromosomes.
    - During DNA replication, DNA polymerase requires a free 3'OH of a pre-existing strand to add free nucleotides.
    - An RNA primer is synthesised to provide a free 3'OH end for the addition of free nucleotides.
    - However, the RNA primer (at the end of the DNA strand) is removed and cannot be replaced with nucleotides. This creates a 3' overhang at the end of the chromosome.

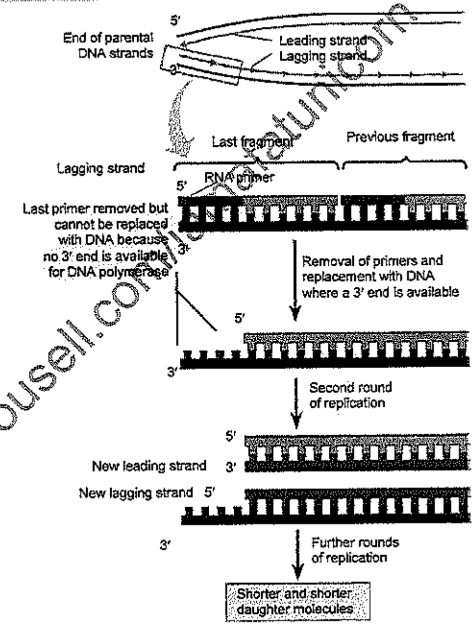


Figure 13. End-replication problem.

- Due to the end-replication problem, the ends of chromosomes shorten with every round of DNA replication.
- Since telomeres are **non-coding, shortening of chromosomal ends leads to shortening of the telomeres** without any deleterious effects.
- The genes within the chromosome will thus **not be eroded/lost** with each round to DNA replication, **preventing loss of vital genetic information.**
  - Eventually, telomeres in cells which have divided many times tend to be shorter.

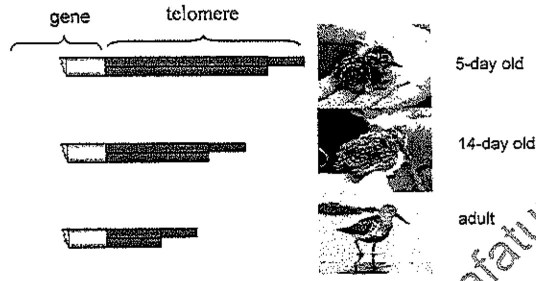


Figure 14. Shortening of telomeres, but not genes, at the end of linear chromosomes in dunlins. [http://www.zoologi.gu.se/]

- **Role 2: Telomeres protect and stabilise the terminal ends of chromosomes.**
  - The single stranded 3' overhang at the ends of linear chromosomes poses some problems. Without telomeres, chromosome ends look like broken chromosomes, and lead to:
    (i) The single stranded 3' single-stranded overhang of one terminal end of a chromosome may anneal to a complementary single-stranded region of the terminal end of another chromosome. This causes joining of different chromosomes.
    (ii) Such ends are similar to DNA damage formed due to double stranded breaks, and sends signals to trigger cell arrest and cell death (apoptosis).
  - Thus by forming a loop, the telomeres **stabilize the ends** of the chromosomes by
    - **preventing them from fusing with other chromosomes**, and
    - prevent DNA repair machinery from recognising the ends of chromosomes as DNA breaks, **protecting the chromosome**, hence **preventing apoptosis.**

Role 3: Telomeres allow their own extension, by providing an attachment point for the correct positioning of the enzyme telomerase.

- Although telomeres shorten with every round of DNA replication, **telomerase** activity in **germ cells, embryonic stem cells** and **cancer cells** can **maintain telomere length.**
- Telomerase prevents shortening of telomeres by lengthening them. ↳ never dies so they continue replicating

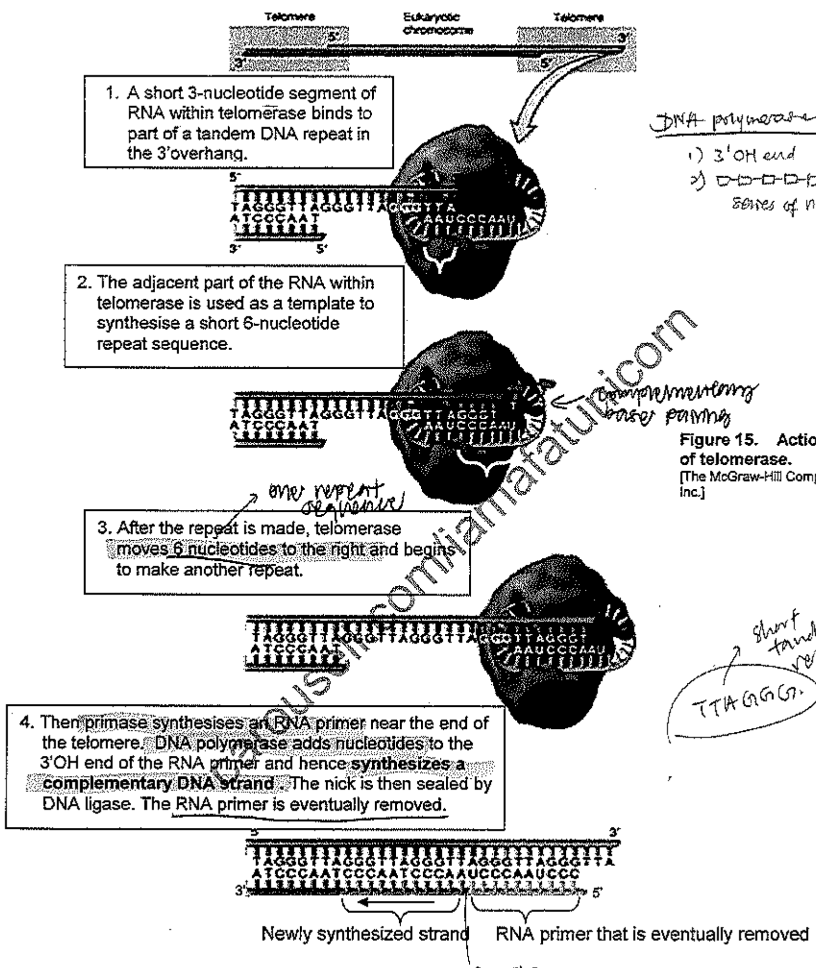


Figure 15. Action of telomerase.
[The McGraw-Hill Companies, Inc.]

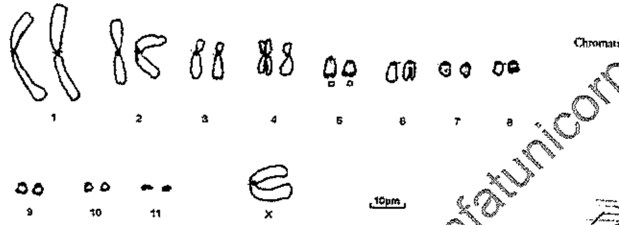
Distribution

- Telomeres are found at both ends of **linear chromosomes** – hence in eukaryotes.

5. CENTROMERE

Structure (Fig. 16)

- **Centromeres** are constricted **regions on chromosomes** where spindle fibres attach to during nuclear division.
- Centromeres can be located anywhere along the length of a chromosome.
- Most functional centromeres have a long stretch of DNA, comprising largely or entirely of repetitive **non-coding** DNA. In humans, the centromere region is made up of **tandemly repeating** units of 170 base pairs.
- Each sister chromatid has its own centromeric DNA sequences.



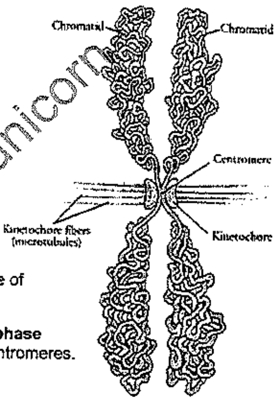


Figure 16.
(Above) Banding pattern of chromosomes of an insect. A special type of staining show that centromeres are darkly-staining regions.
[http://www.insectscience.org/2.24/]
(Right) A chromosome, highly condensed, at late prophase or metaphase stage of nuclear division. Two chromatids are seen 'touching' at the centromeres.

Function

- Centromeres are pivotal in ensuring proper nuclear division.
  1. They allow **sister chromatids to adhere to each other.**
  2. They allow proteins called **kinetochore proteins**, and subsequently **spindle fibres, to attach so that sister chromatids/homologous chromosomes can be separated to opposite poles**. Since kinetochores bind specifically to centromeric regions, special DNA sequence in centromeric region must be important for kinetochore recognition.
- Without centromeres, improper chromosomal alignment and segregation will result, leading to aneuploidy and conditions such as Down's Syndrome.

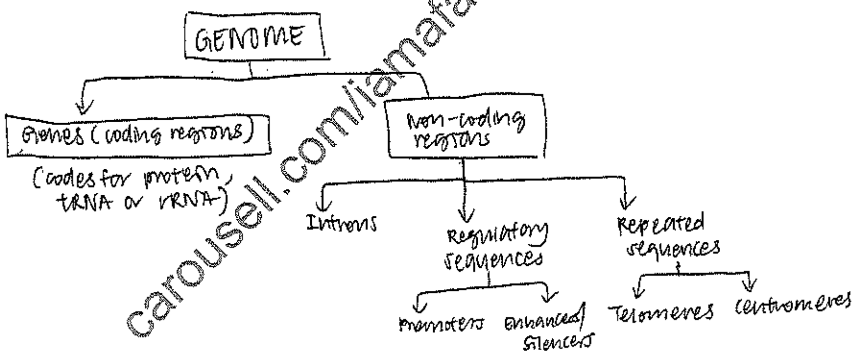
Distribution

- Centromeres are found in **linear chromosomes** only – hence in eukaryotes.

Keywords include:

Circular DNA	Exons	Spliceosome	RNA polymerase
Linear DNA	Genome	Alternative RNA splicing	Consensus sequence
Plasmid	Promoter	Tandem repeats	Specific transcription factors
Histones	Enhancer	Telomeres	Activators
Chromatin fiber	Silencer	telomerase	Repressors
Nucleosomes	Non-coding	Centromere	Transcription initiation complex
Introns	Splice/ Splicing	Kinetochore proteins	3' overhang
Supercoiling	Operon	General transcription factors	

Summary



carousell.com/iamafatunicorn

NAME: CT GROUP:

CORE IDEA 2 : GENETICS AND INHERITANCE

ORGANISATION AND CONTROL OF PROKARYOTIC AND EUKARYOTIC GENOME (II)

CONTROL OF PROKARYOTIC AND EUKARYOTIC GENOME

Content

Control of Gene Expression

Learning Outcomes

Candidates should be able to:

2 (i) Explain how gene expression in prokaryotes can be regulated, through the concept of simple operons (including *lac* and *trp* operons), including the role of regulatory genes; and distinguish between inducible and repressible systems (knowledge of attenuation of *trp* operon is not required). *(Covered in Bacteria notes)*

2 (j) Explain how differential (ie. 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).

References

Campbell, N.A. and Reece J.B. (2008). Biology (8th Ed), Pearson Education, Inc.
Brooker, R.J. (2005) Genetics: Analysis and Principles (2nd Ed), McGraw-Hill.
Alberts, B. *et al.* (2008) Molecular Biology of The Cell (5th Ed), Garland Science.
Lewin, B. (2000) Genes VII (7th Ed), Oxford University Press.
Green, Stout and Taylor (1997) Biological Science 2 (3rd Ed), Cambridge University Press.

This handout is the effort of Biology teachers at RI. It has and will continue to be updated.

<u>**Table of Contents**</u>

✱ Every somatic cell contains the same genome, but different genes are expressed in different types of cell

(A) Introduction to Gene Regulation

- Regulation of gene expression includes a wide range of mechanisms that are used by cells to increase or decrease the production of specific gene products (protein or RNA).
- Some proteins are **synthesized continuously at a fixed rate** and genes coding for such proteins are said to be **constitutively** expressed.
- However, not all proteins are constitutively expressed. Any step of gene expression may be modulated. The expression of other genes are **regulated** and there are several mechanisms by which this is done:

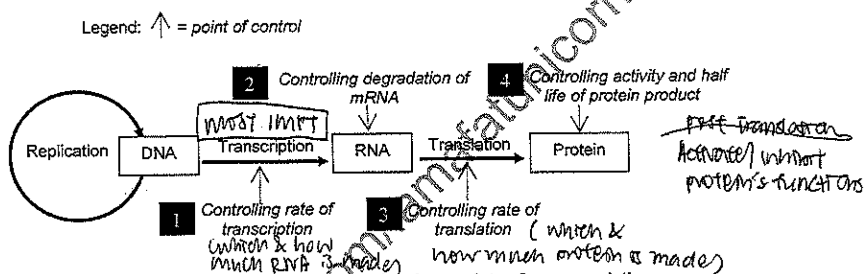


Figure 1. The Central Dogma and various points of gene regulation.

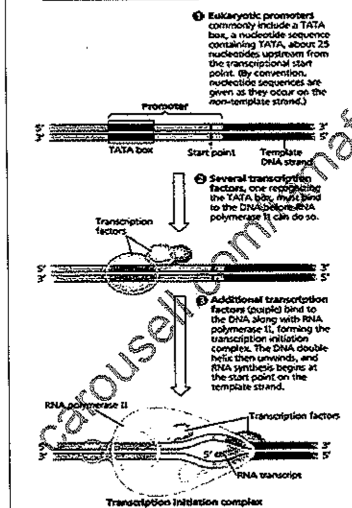
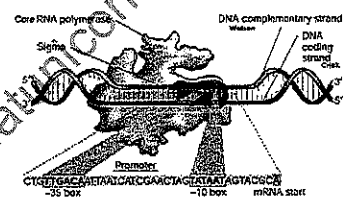
Q1: Can you suggest which point of control of gene regulation predominates? Why?

Transcription-level control.
It is the most efficient mechanism with minimal wastage.

Diff gene expression in diff cells
↓
diff proteins produced in diff cells
↓
diff cells have diff properties / structural features
↓
diff cells have diff roles & functions.

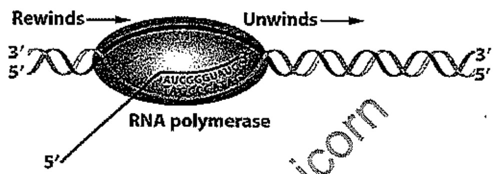
Why regulate gene expression?
- differentiation
- adapt to changes
- conserve resources
- more varied proteome (entire types of protein) despite limited genome size

(B) Brief Overview of Transcription and Translation in Prokaryotes and Eukaryotes

Transcription	Eukaryotes	Prokaryotes
Initiation	Step 1: Promoter recognized by general transcription factors 1) ***General transcription factors**** assemble at the **TATA box** in the promoter region. 2) Recruitment of ***RNA polymerase****. 3) The whole complex of transcription factors and RNA polymerase bound to the promoter is called a ***transcription initiation complex****. 	Step 1: Promoter recognized by sigma factor 1) **Sigma factor** binds to core RNA polymerase to form RNA polymerase holoenzyme. 2) RNA polymerase holoenzyme scans along the DNA, and its sigma factor recognizes and binds to the promoter elements at both the -10 and -35 sequences. 
	Step 2: RNA polymerase ***unzips and separates**** **the two strands** of the DNA double helix at the promoter by ***breaking hydrogen bonds**** that link base pairs. DNA now has two exposed strands: (1) one strand is used as the ***template****; (2) the other strand is not transcribed. (NB: It is not spelt as TRANSCRIPTED!!)	
Elongation	Step 3: 1) **Free ribonucleotides** are matched up with the DNA template by ***complementary base pairing****. 2) mRNA is synthesized and elongated in the ***5' to 3' direction****. 3) RNA polymerase catalyses the joining of adjacent ribonucleotides through the formation of covalent ***phosphodiester bonds****.	

Step 4:
RNA polymerase moves along the DNA template strand from the 3' end towards the 5' end (**3' to 5' direction**) separating the 2 DNA strands and catalyzing the assembly of the ribonucleotides.

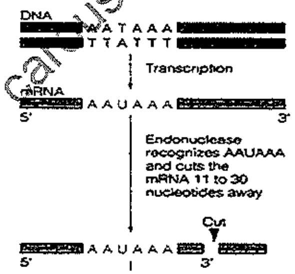
Step 5:
As RNA polymerase continues down the template strand, the region of DNA that has just been transcribed **reanneals**.



Termination

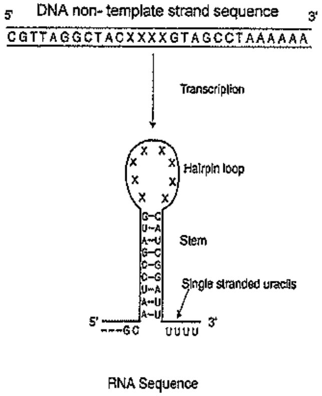
Eukaryotes

Step 6:
1) RNA polymerase transcribes a sequence in the DNA called the **polyadenylation signal (AAUAAA)** in the pre-mRNA.
2) At a point about 10-35 nucleotides downstream from the polydenylation signal, proteins associate with growing RNA transcript and cleave it from the RNA polymerase, causing it to be released.



Prokaryotes

Step 6:
1) Transcription proceeds through a terminator sequence in the DNA.
2) The transcribed **terminator** (an RNA sequence) folds back to form a **hairpin loop** and functions as the **termination signal**, causing the RNA polymerase to detach from the DNA and release the RNA transcript.

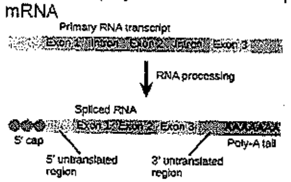


Post transcriptional modification

Process

Eukaryotes

1) Addition of 5'cap
2) Splicing - the process where introns are excised and exons are joined together.
3) Addition of poly A tail at 3'end of pre-mRNA



Prokaryotes

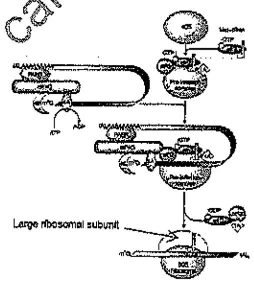
None

Translation

Initiation

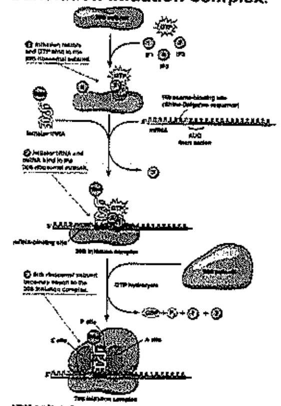
Eukaryotes

Step 1:
1) **Small ribosomal subunit**, **eukaryotic translation initiation factors (eIFs)** and **initiator-tRNA** form a complex. (There is interaction between other eukaryotic translation initiation factors, the 5' cap and the poly-A tail of the mRNA, circularising the mRNA.)
2) Complex from 1) binds to the **5' cap** of an mRNA and moves in the **5' to 3' direction** along the mRNA to locate the **start codon, AUG**.
3) Binding of the large ribosomal subunit completes the **translation initiation complex**.

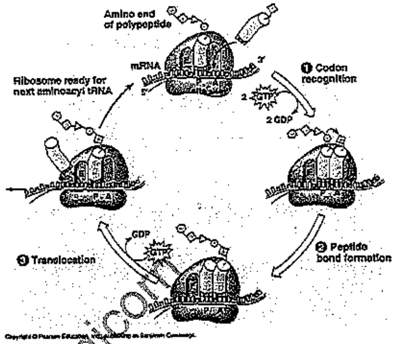
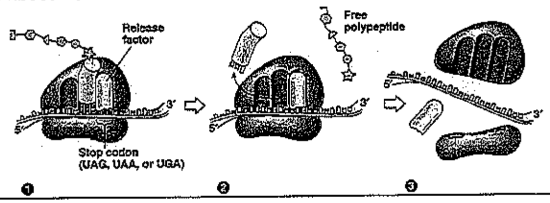


Prokaryotes

Step 1:
1) Translation initiation factors bind to the small ribosomal subunit and facilitate the binding of the small ribosomal subunit to **Shine-Dalgarno sequence** on the mRNA.
2) Initiator tRNA then binds to start codon, AUG, on mRNA
3) Large ribosomal subunit binds to complete the formation of the **translation initiation complex**.



4) The initiator-tRNA is now positioned at the **P site** leaving the **A site** vacant for the addition of the next aminoacyl-tRNA molecule.
5) **GTP** is required for the initiation stage.

Elongation	Step 2: Codon recognition A second tRNA carrying the next amino acid binds to the **A site** by forming **hydrogen bonds** between the anticodon of the tRNA and the second codon of the mRNA via ***complementary base pairing****.  Step 3: Peptide bond formation A ***peptide bond**** is formed between the methionine and the second amino acid in the A site. This is catalyzed by **peptidyl transferase** present in the large ribosomal subunit. Step 4: Translocation 1) The ribosome shifts one codon down the mRNA in a **5' to 3' direction.** ▪ 1st tRNA shifted to **E site** and released into cytosol. ▪ 2nd aminoacyl-tRNA shifted from A site to **P site.** ▪ The **A site** empty and ready to receive 3rd aminoacyl-tRNA, with anticodon complementary to third codon of mRNA. 2) Thus, ▪ P site holds tRNA carrying the growing polypeptide chain ▪ A site holds tRNA carrying the next amino acid ▪ E site hold tRNA which has donated its amino acid and is ready to leave. 3) The ribosome continues to translate the remaining codons on the mRNA until the ribosome reaches a ***stop codon**** on mRNA (**UAA, UGA, or UAG**).
Termination	Step 5: 1) Termination occurs when one of the stop codons enters the A site. 2) These codons are not recognized by any tRNAs. Instead, they are recognized by release factors which trigger the eventual hydrolysis of the bond between the polypeptide and the tRNA in the P site. 3) The polypeptide is released from the ribosome as it completes its folding to assume the necessary secondary and protein structures. 4) The ribosome dissembles into its subunits. 

PG8-11 can ignore (can be found elsewhere!!)

Based on our understanding of the process of transcription and translation, let us now look at how gene expression in prokaryotes and eukaryotes can be regulated.

(C) Prokaryotic Gene Expression

(1) Control @ transcriptional level (Pro)

[Controlling when and how often a gene is transcribed]

Note: Transcriptional control is the most important level of regulation in prokaryotes, followed by translational control.

(i) Promoter [Refer to 'Prok and Euk Part I']

- The promoter is a DNA sequence where ***RNA polymerase holoenzyme**** (with its sigma factor) binds to and **starts/initiates** transcription.

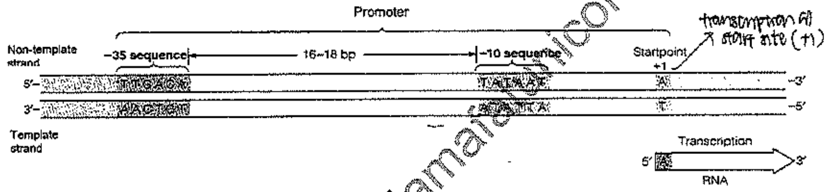


Figure. 2. General structure of prokaryotic promoter. By convention, nucleotide sequences are given as that on the non-template DNA strand. The promoter actually includes transcriptional start site (+1), where transcription of a gene actually begins. Direction of transcription is referred to as "downstream" and the other direction as "upstream".

- Bacterial promoters contain two important **critical elements**/short sequences.
- Sequences at the
  (a) -10 site are known as ***-10 sequence / Pribnow box****
  (b) -35 site are known as ***-35 sequence***
- Many promoters have -10 and -35 sequences that are similar to the 'ideal' / consensus sequence. The ***consensus sequence**** at :
  (a) -10 site is 5'-TATAAT-3'
  (b) -35 site is 5'-TTGACA-3'
- Bacterial ***RNA polymerase holoenzyme**** is made up of a ***core polymerase**** and a ***sigma factor****. The **sigma factor** on the holoenzyme **recognizes and binds** to **both** the **critical elements at the promoter** (Fig. 2a). The **shape of the RNA polymerase and sigma factor is complementary** to the **nucleotide sequence** that they bind to.

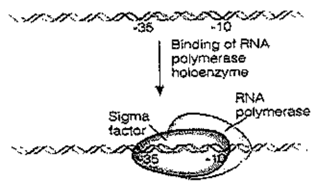


Figure 2a. Sigma factor of RNA polymerase holoenzyme recognises a prokaryotic promoter

- Regulation **Strength of any given promoter** is determined by **how similar the -10 and -35 sequences are to the consensus sequences**. The more similar the sequences are to the consensus sequences, the stronger the promoter. Strong promoters have a **higher frequency of transcription**. (Note: Do not use term 'rate of transcription'.)

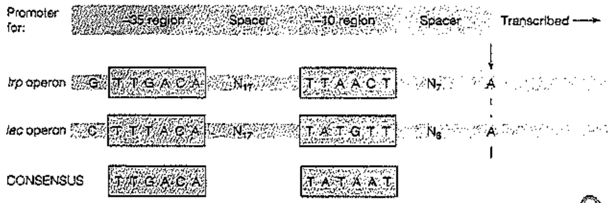
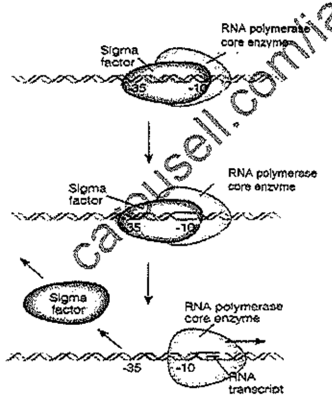


Figure 3. -35 and -10 sequences within two bacterial promoters. The consensus sequence is obtained after comparing hundreds of promoters.

(ii) Sigma Factor = a subunit of RNA polymerase

- RNA polymerase needs to recognise and bind to the promoter to start transcription. In prokaryotes, this recognition role is specifically undertaken by the ***sigma factor****, a subunit of RNA polymerase. There are **different sigma factors, each recognize and bind to a different promoter**. The sigma factor is a protein.



RNA polymerase holoenzyme comprises a core polymerase and sigma factor. As the assembled RNA polymerase scans along the DNA, its sigma factor recognises the critical elements at both -10 and -35 positions.

RNA polymerase holoenzyme unzips and separates part of the double-stranded DNA for transcription. Only one of the two DNA strands is used as a template.

After a short distance, the sigma factor dissociates from the polymerase. Synthesis of RNA continues.

Figure 4. Sigma factor helps RNA polymerase core enzyme recognise a prokaryotic promoter. Diagram shows transcription initiation process.

↑ sigma factors → ↑ rate of transcription, vice versa

- Regulation **Controlling the availability of different sigma factors** determines which genes can be transcribed. A simple change of sigma factors can allow different sets of genes / operons to be transcribed by the same RNA polymerase core enzyme.

- ~~The sigma factor and RNA polymerase core enzyme being recruited to the promoter on the DNA forms the ***transcription initiation complex**** which can then result in transcription.~~ → this is only in eukaryotes

(iii) Operon [Refer to 'Genetics of Bacteria']

- In prokaryotes, genes with related functions are grouped into an operon. Expression of these genes are controlled by a single promoter and transcribed into a **single, *polycistronic**** mRNA.

- An operon consists of an operator, a common promoter and two or more structural genes. The constitutively expressed regulatory gene is usually located upstream of the operon, and can code for either an active or an inactive repressor.

- Regulation
  - An operon allows the bacteria to **coordinately regulate a group of genes that encode gene products with related functions**. As operons consist of a set of structural genes that lie near one another on the chromosome and they are coordinately transcribed, these structural genes are turned on or turned off together. The expression of the structural genes occurs as a single unit. In this way, prokaryotes can, for example, turn on all of the structural genes of the *lac* operon so that they can take up and use lactose, when lactose is needed and available in their environment. By grouping the genes together, control of gene expression becomes more efficient.
  - **Small molecules** help the operon fine tune control of gene expression.
    - ***Repressor****, when bound to the operator, will block binding of RNA polymerase to the promoter and prevent gene transcription of the group of structural genes.
    - Small effector molecules such as an ***inducer**** (e.g. ***allolactose****, *lac* operon) can further control **transcription frequency**.

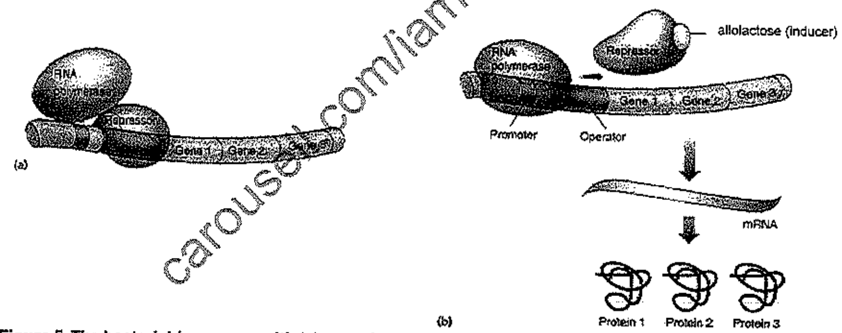


Figure 5. The bacterial *lac* operon with (a) an active repressor and (b) a repressor inactivated by allolactose.

(2) Control @ post-transcriptional level (Pro)

[Controlling after mRNA is synthesised]

- **None** meiyou
  - RNA splicing does not occur; bacterial mRNA is used for translation as it is being produced.

Why prokaryote mRNAs have short lives?

(3) Control @ translational level (Pro)
[Controlling how many polypeptides can be made]

(i) mRNA stability/ half-life

- Stability / half-life of an mRNA molecule in the cytoplasm will determine the number of polypeptides that can be translated from it.

- **Regulation** Prokaryotic mRNAs, unlike in eukaryotes, have a relatively **short half-life**. Short mRNA half-lives allow **rapid degradation**.
  - A bacterial cell can then control gene expression by rapidly adjusting synthesis of proteins in response to environmental changes. mRNAs are degraded by **ribonucleases** (RNases) into constituent ribonucleotides usually a few minutes after they are synthesised.

- **Regulation** When induced by certain changes in conditions (e.g. water potential changes), an antisense RNA complementary to a particular mRNA will be quickly synthesized by the bacteria.
  - Binding of an **anti-sense RNA** complementary to the mRNA will reduce half-life of mRNA. This further helps gene regulation because the **anti-sense RNA will block translation / target this RNA for degradation** (details not needed).

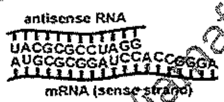


Figure 6. Anti-sense RNA forms a double-stranded and reduces half-life of mRNA.

(ii) Binding of small ribosomal subunit*

- In prokaryotes, the ***Shine-Dalgarno sequence**** (5'-AGGAGG-3') is found a few nucleotides upstream of each AUG start codon in a polycistronic mRNA. The small ribosomal subunit will bind to this sequence so that the start codon can be correctly positioned in the subunit before initiator tRNA and large ribosomal subunit come along for translation.

- **Regulation** Translation initiation can be blocked by:
  (a) **binding of a translational repressor** protein ('R' in Fig. 7) at/near to Shine-Dalgarno sequence
  - prevents binding of small ribosomal subunit to Shine-Dalgarno sequence. Since ribosomes cannot assemble properly, translation fails.

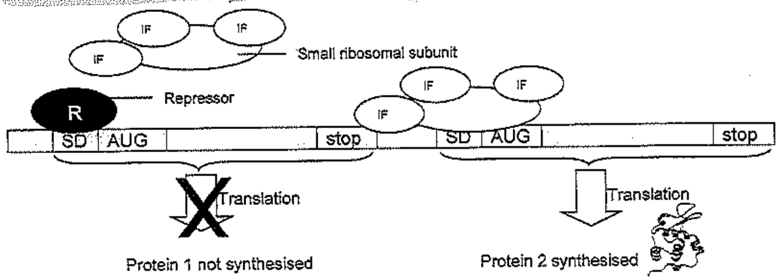


Figure 7. Blocking translation initiation. A polycistronic mRNA coding for two proteins is shown. SD: Shine-Dalgarno sequences. IF: Initiation factors [refer to next section].

(b) **binding of an anti-sense RNA** complementary to the mRNA at/near Shine-Dalgarno sequence.
- prevents binding of small ribosomal subunit to Shine-Dalgarno sequence as it recognises only single-stranded mRNA. When ribosomes cannot assemble properly, translation cannot occur and polypeptides cannot be synthesised.

(iii) Initiation factors

- Initiation factors are required for proper positioning of the small ribosomal subunit together with initiator tRNA on the mRNA, and the subsequent recruitment of the large ribosomal subunit which together forms the ***translation initiation complex****.

- **Regulation** Availability of initiation factors controls initiation of translation.

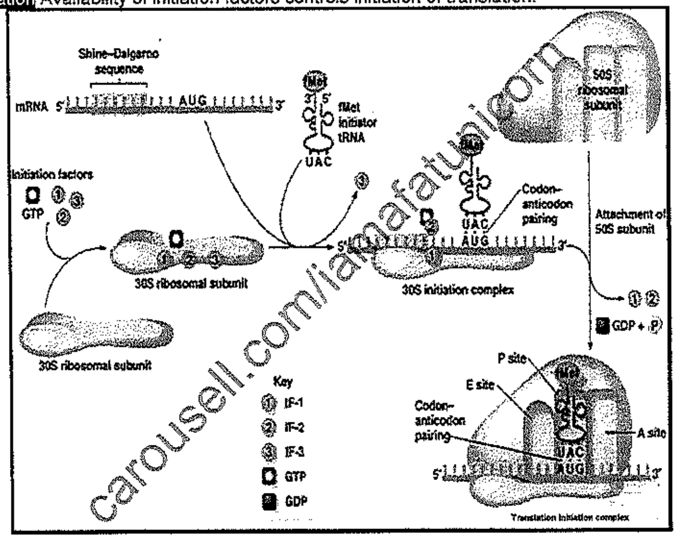


Figure 8. Initiation factors ('IF' in diagram) are needed for translation initiation.

(4) Control @ post-translational level (Pro)
[Controlling proteins that are already present in the cell by activating or inhibiting their functions]

- Not significant, however,

1. Covalent modification.
2. Phosphorylation/ dephosphorylation
3. Protein degradation

do occur to regulate protein activity

(D) Eukaryotic Gene Expression

(1) Control @ genomic level (Euk)

[Controlling by changing chromosomal structure]

Organisation of chromatin structure

- Organisation of DNA into chromatin helps to:
  (a) pack DNA into a compact form that fits inside the nucleus of a cell.
  (b) regulate gene expression - physical state of DNA in / near a gene determines if the gene is **accessible** for transcription.

- Some sections within a chromosome are more compact than others:
  (a) ***Heterochromatin**** = highly compacted DNA where DNA winds **more tightly** around histones.
  – formation of heterochromatin results in **silencing of genes/ inactive gene expression** as it **limits access of RNA polymerase and general transcription factors to promoters** of genes and thus **prevents formation** of the ***transcription initiation complex****.
  (b) ***Euchromatin**** = less compacted DNA where DNA winds **less tightly** around histones.
  – formation of euchromatin **promotes access of RNA polymerase and general transcription factors** to **promoters** of genes hence allowing the formation of the ***transcription initiation complex****.

- Chromatin is a dynamic structure that alternates between heterochromatin and euchromatin.

(i) Chromatin remodeling complex

- Gene expression can be controlled by modifying chromatin structure.

- **Regulation** Chromatin remodeling complex are protein complexes that **alter structure of nucleosomes** temporarily.
  (a) results in **DNA being less tightly bound to histones**
  – allows RNA polymerase and general transcription factors involved in gene expression to access the promoter to **initiate transcription**.
  Euchromatin
  (b) can also result in **DNA being more tightly coiled around histones**
  – prevents RNA polymerase and general transcription factors involved in gene expression from accessing the promoter, thereby **blocking transcription**.
  Heterochromatin

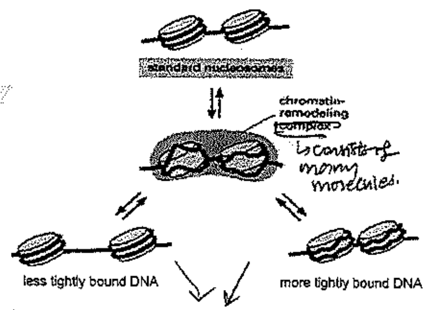


Figure 9. Chromatin remodeling complex on nucleosomes. (depends on the type of chromatin remodelling complex)

Alternates between heterochromatin & Euchromatin.

(ii) DNA Methylation

- ***DNA methylation**** involves addition of a methyl group to selected cytosine (C) nucleotides located in the **sequence CG.** (CG rich sequences DNA)
- DNA methylation can be found extensively within heterochromatin.

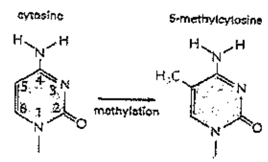


Figure 10. Methylation of a cytosine base.

- **Regulation** DNA methylation usually **prevents transcription** by:
  (a) **blocking binding of general transcription factors** and hence preventing assembly of transcription initiation complex at the promoter.
  (b) **recruiting DNA-binding proteins** (i.e. transcriptional ***repressors***, ***histone deacetylases**** and repressive chromatin remodeling complexes) to the methylated DNA to condense chromatin
  - results in gene silencing/ no gene expression.

(iii) Histone acetylation / deacetylation

- ***Acetylation**** and ***deacetylation**** of **histones allows chromatin to decondense and condense**, respectively, and alternate between loose and tightly condensed states. Gene expression can then be manipulated.

- **Regulation** **Acetylation of histones** is catalysed by ***histone acetylase****: histone acetyltransferase (HAT)
  – **addition of acetyl groups** (-$COCH_3$) to lysine residues **removes positive charges** on histones, decreasing the electrostatic interactions between the negatively-charged DNA and the histones. Tight binding between DNA and histones is loosened, making the **promoter region more accessible to RNA polymerase and general transcription factors**. Acetylation works in concert with chromatin remodeling complex, **allowing formation of the transcription initiation complex resulting in transcription**.

- **Regulation** **Deacetylation of histones is** catalysed by ***histone deacetylase****:
  – **removal of the acetyl groups**. This restores a tighter interaction between DNA and histones, **inhibiting transcription**.

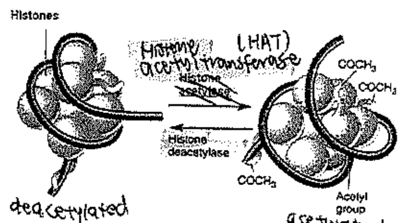


Figure 11. Histone acetylation and deacetylation.

This is an example of covalent modification of histones.

Additional info: (but that doesn't mean it won't be tested as info here may help you answer questions)

X chromosome inactivation

A special case of gene silencing is X chromosome inactivation:

- there are 2 X chromosomes in females and 1 in male (mammals like humans, and flies). If genes are expressed in the same amount, females will obtain double the amount of gene products.
- 1 of the 2 X chromosomes in diploid cells of such organisms is therefore inactivated. This is achieved by compacting its chromosomal DNA via extensive formation of heterochromatin, histone modifications, DNA methylation etc.
- the highly compacted chromosome is called a **Barr body**. Most of its genes cannot be expressed, only several can. As a result of inactivation, only genes on 1 of the 2 X chromosome can be transcribed readily. This prevents excessive amount of gene products from accumulating.

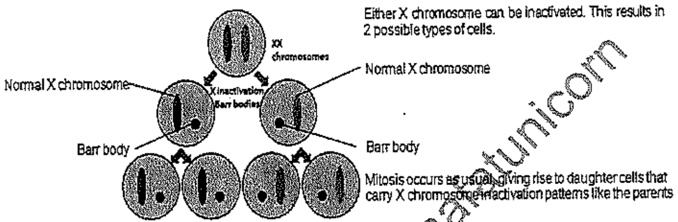


Figure 12. Inactivation of either X chromosome in some animals.

(iv) Gene amplification

- Gene amplification refers to the replication of a specific gene multiple times to create <u>more copies of that gene</u>.
- Regulation Gene amplification results in a cell which has a normal number of all other genes except for the **gene of interest which exist in high copy number.** During transcription and translation, <u>increased copies of mRNA</u> and <u>increased copies of the required protein</u> will be obtained.

(2) Control @ transcriptional level (Euk)

- Control elements are non-coding DNA segments that ***transcription factors**** **bind** to in order to regulate transcription.
- Control elements can be divided into:
  (a) Proximal control elements (**e.g. *promoters****)
  - non-coding DNA that usually lie directly <u>upstream of the transcription start site</u>. They are bound by ***general transcription factors****.

  (b) Distal control elements (**e.g. *enhancers**** **and *silencers****)
  - non-coding DNA that can be <u>located thousands of nucleotides upstream or downstream of the gene, or even within an intron.</u> They bind to other types of transcription factors known as ***specific transcription factors****.

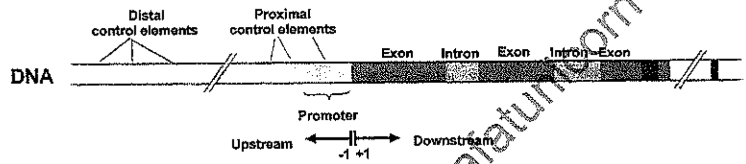


Figure 13. Types of eukaryotic control elements.

- Transcription factors are **gene regulatory proteins** needed for transcription. They <u>bind to control elements</u> as well as <u>other transcription factors/proteins</u>.

(i) Promoter

- The ***promoter**** is located <u>just upstream of the transcription start site</u> of a gene.
- As in prokaryotes, they function <u>as recognition site</u> for the <u>binding of general transcription factors</u> and <u>RNA polymerase</u> to start/initiate transcription.
- There are critical elements/short sequences in the promoter that improve the efficiency of the promoter by helping to <u>recruit RNA polymerase and general transcription factors</u> to the promoter.
- Regulation In eukaryotes, the similarity of critical elements to consensus sequences is <u>not that crucial in controlling gene expression</u>. This is markedly different from gene regulation in prokaryotes. Eukaryotic cells rely heavily on other factors such as presence of enhancers/ silencers.

- <u>Critical elements / short sequences</u> within the eukaryotic promoter:
  1) *TATA box** at -25 site
  - at a relatively fixed position.
  - Has a consensus sequence of 5'-TATAA-3'
  - Regulation Important in determining <u>precise location of the transcription start site</u>. Its deletion will result in transcription starting at a variety of locations, resulting in a truncated/ non-functional protein. It is similar in function to Pribnow box (-10 element) in prokaryotic promoters.

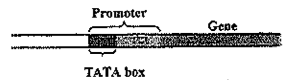
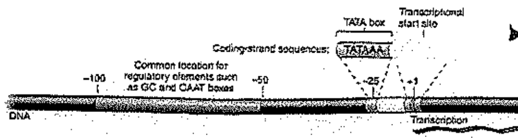


Figure 13a. Position of a eukaryotic promoter

2) CAAT and GC boxes:
- Can be found at varying locations within the promoter, although most commonly found at the locations of -75 and -90, respectively. They are not always present; GC can be found in multiple copies.
- CAAT box has a consensus sequence of 5'-GGCCAATCT-3' whereas GC box has a consensus sequence of 5'-GGGCGG-3'.
- Regulation Improve efficiency of promoter by helping to **recruit general transcription factors and RNA polymerase** to the promoter.



▶ At promoters only get BASAL level of transcription. To further control @ transcription, need enhancers & silencers

Figure 13b. Critical elements in the eukaryotic promoter

(ii) Enhancers & Silencers

↳ determine frequency of transcription.

Structure
- ***Enhancers**** and ***silencers**** are short non-coding sequences which are usually located far away (up to thousands of nucleotides upstream or downstream) from the promoter, but can also be found within an intron or near the gene to be controlled.
- They are regulatory elements.

- Regulation Enhancers and silencers are DNA sequences that help to increase and reduce frequency of transcription, respectively:
  (a) ***Enhancers****, when bound by ***specific transcription factors**** known as ***activators****, increase frequency of transcription by promoting assembly of transcription initiation complex.

  (b) ***Silencers****, when bound by ***specific transcription factors**** known as ***repressors****, decrease frequency of transcription by inhibiting assembly of transcription initiation complex.

ENHANCERS
- In the absence of enhancer sequences, many eukaryotic genes have very low levels of basal transcription. Enhancer sequences can increase rate of transcription 10 to 1000 fold.

Function
- **Enhancer sequences** allow the binding of ***specific transcription factors**** called ***activators****. This **increases the frequency of transcription** of the genes they control. → gene activation

- Enhancers are positive regulatory elements involved in the upregulation of transcription as they promote the assembly of the ***transcription initiation complex**** via their interaction with **activators**.

- Activators bind to enhancers to increase frequency of transcription by:

1. Promoting assembly of transcription initiation complex
- Upon binding of **activators to enhancers**, spacer DNA (regions of non-transcribed DNA between genes) bends.
- **Bending of spacer DNA** allows direct interaction of activators with RNA polymerase and/or general transcription factors at the promoter, promoting the assembly of transcription initiation complex.

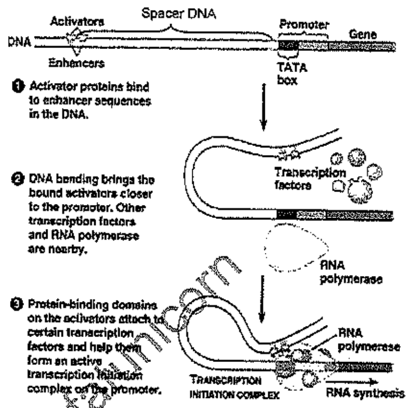


Figure 14. Looping mechanism that allows for interaction between activators and RNA polymerase / transcription factors at the promoter.

2. Increasing accessibility to promoter DNA
- Bound activator may recruit histone acetylase and chromatin remodeling complex to decondense chromatin.
- This allows greater accessibility of general transcription factors and RNA polymerase to the promoter.

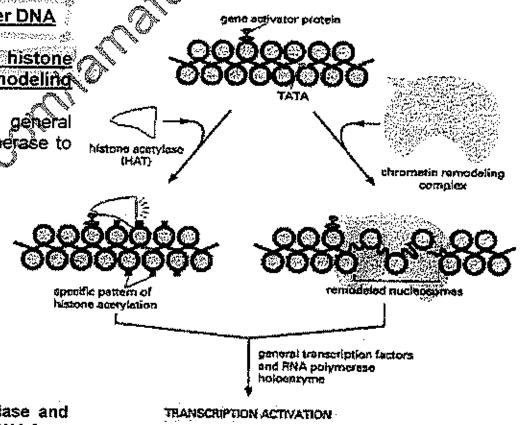


Figure 15. Recruitment of histone acetylase and chromatin remodeling complex to loosen DNA from histones under the influence of an activator bound to its enhancer.

SILENCERS

- *Silencer** **sequences** allow binding of ***specific transcription factors**** called **repressors*** which represses / prevents / **decreases the frequency of transcription** of the genes they control. → **gene silencing / gene repression**.
- Silencers are **negative regulatory elements** involved in the **downregulation** of transcription as they **prevent the assembly of the *transcription initiation complex**** via their interaction with **repressors***.

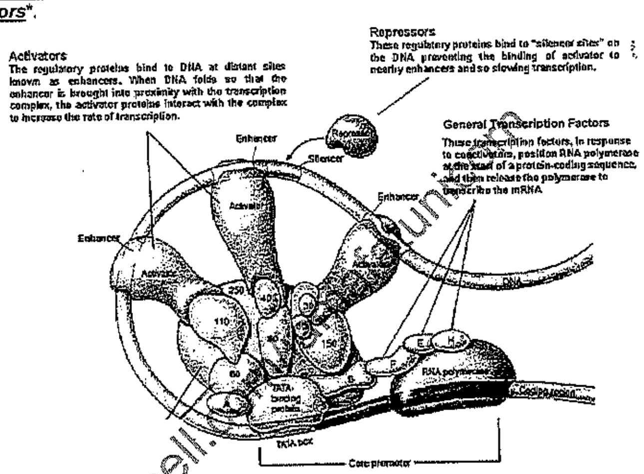
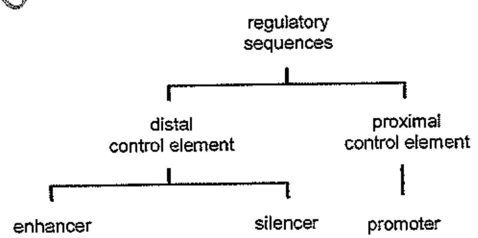


Figure 16. Overview of how enhancers and silencers act in concert to regulate efficiency of transcription.



- Repressors bind to silencers to decrease the frequency of transcription by several different mechanisms:

1. Interfering with action of activator

A. **Competitive DNA binding**
- Enhancer region overlaps with silencer region. Therefore, binding of repressor prevents binding of activator.

B. **Masking activation surface**
- Repressor binds to activator to prevent it from interacting with general transcription factors.

C. **Direct interaction with general transcription factors**
- Repressor interacts with general transcription factors to prevent assembly of transcription initiation complex.

2. Changing the chromatin structure
- Bound repressor may **recruit histone deacetylase and repressible chromatin remodeling complex** to condense chromatin. This decreases the accessibility of general transcription factors and RNA polymerase to the promoter.

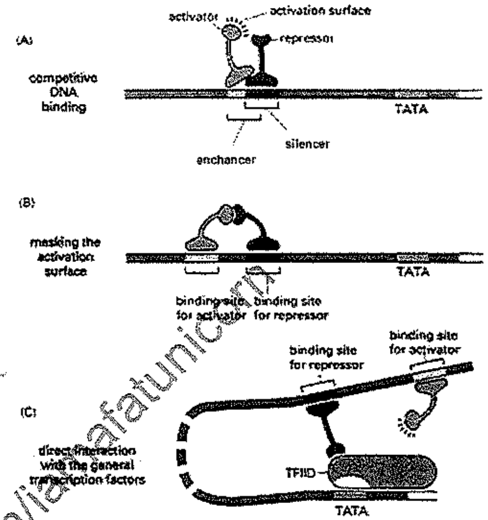


Figure 17. Mechanisms of repressors in downregulating transcription

(3) Control @ post-transcriptional level (Euk)

- RNA processing takes place after transcription. Majority of post-transcriptional modifications occur on mRNA.
- RNA processing occurs only in **eukaryotes**, where it takes place in **nucleus** of the cell. **In prokaryotes, mRNA undergoes little or no modification following synthesis** by RNA polymerase. Indeed, many of them are being translated simultaneously while they are being transcribed.
- RNA transcript formed **immediately after transcription** is known as **pre-mRNA**/ primary mRNA transcript / nascent mRNA / precursor mRNA. This has to be modified to become a **mature mRNA** that can be exported out of the nucleus for translation.
- **Regulation** RNA processing is crucial to regulate gene expression. For instance, (1) translation of incompletely processed pre-mRNAs containing introns produces defective proteins that might interfere with functioning of the cell. (2) By altering mRNA stability, it affects the amount of protein that can be made.

- Post-transcriptional modification of mRNA is carried out in this order:
  (a) capping at the 5' end
  (b) splicing of pre-mRNA
  (c) adding a poly-A tail to the 3' end (polyadenylation)

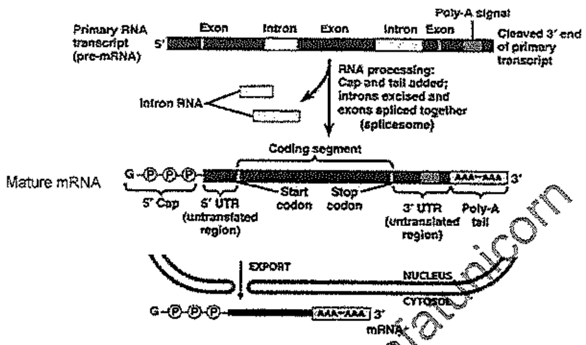


Figure 18. Modifying pre-mRNA into a mature, functional mRNA in eukaryotes. Only mature mRNA exits through the nuclear pores and are translated in cytoplasm.

(i) Capping at 5' end of mRNA

- Capping involves the addition of a **7-methylguanosine** nucleotide to 5'end of pre-mRNA.

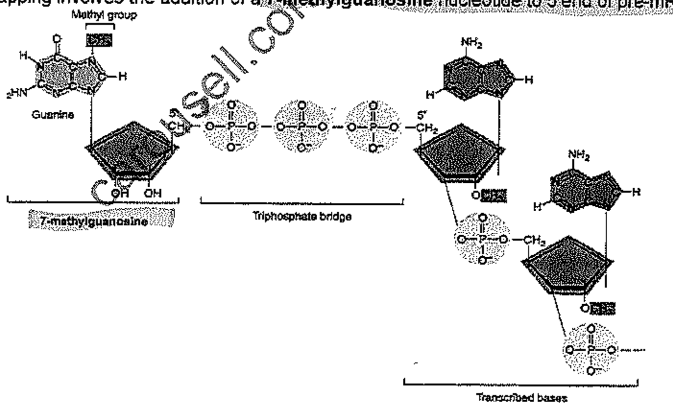


Figure 19. Addition of a 7-methylguanosine cap to 5'end of pre-mRNA.

- Regulation Importance of 7-methylguanosine cap at 5' end of mRNA:
  (i) **processing (mRNA splicing and polyadenylation):**
  – ***5' cap**** helps the cell to recognise mRNA (amongst all other RNA molecules in the cell). This ensures subsequent steps such as splicing occurs on the correct RNA molecule.
  (ii) **export out of nucleus**
  – 5' cap is recognised by certain proteins, which are required for the mRNA to exit from nucleus via nuclear pores.
  (iii) **half-life / stability**
  – 5' cap stabilises mRNA by protecting the growing pre-mRNA from rapid degradation by cellular ribonucleases.
  (iv) **translation**
  – 5' cap helps promote translation initiation. The cap is recognised by eukaryotic initiation factors. Subsequent binding of initiation factors to the cap helps recruitment of mRNA to small ribosomal subunit.
  [Note: 5' cap is added as mRNA is being synthesised]

(ii) Splicing of pre-mRNA

- Splicing is a process during which introns are **excised** and exons are **joined together/ spliced together** by a ***spliceosome****.
- Involves ***spliceosome****, a snRNA (small nuclear RNA)-protein complex. Points of excision are very precise and are determined by sequence of nucleotides at intron-exon boundaries.

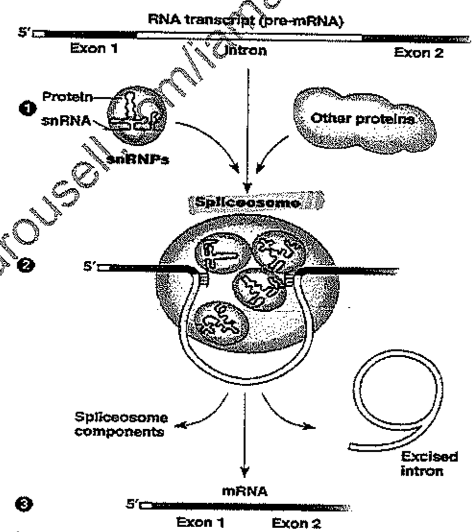


Figure 20. Splicing of pre-mRNA. Involves spliceosome, a snRNA (small nuclear RNA)-protein complex. Points of excision are very precise and are determined by sequence of nucleotides at intron-exon boundaries.

proteome – entire set of proteins that can be expressed by a genome.

- Regulation **Importance of alternative splicing** [Refer to 'Org and Control Part I']:
  - a pre-mRNA transcript can be spliced in different ways. This allows 1 gene to generate different mature RNA and subsequently different polypeptides. Each polypeptide has its own unique properties. By having different polypeptides at any one time, different metabolic activities can be easily controlled.

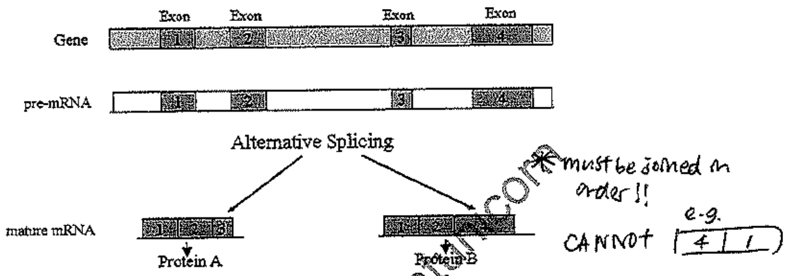


Figure 21. In alternative splicing, different mRNA transcripts are obtained from one gene. Different combinations of exons are found on the mature mRNA. Exons from one mRNA molecule will not be joined with those from another mRNA molecule. Alternative splicing only occurs for exons found on the same mRNA. 2 of the possible combinations are shown here.

(iii) Adding a poly-A tail to 3' end of mRNA

- Polyadenylation begins with 3' end of pre-mRNA being cleaved enzymatically to make it shorter. Adenosine monophosphates (a ribonucleotide), are added one at a time to form a ***poly-A tail*** at the 3' end of the mRNA. Length of the poly-A tail varies among different mRNAs.
- The ribonucleotides are added by the enzyme poly-A polymerase.

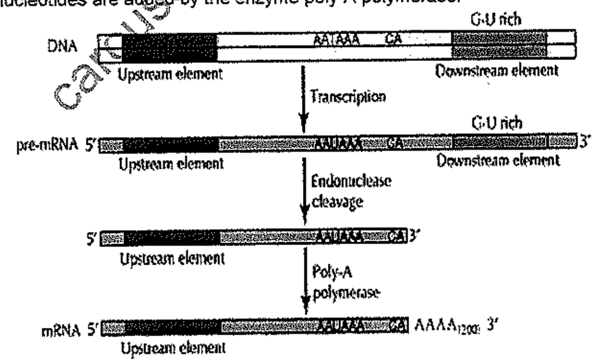


Figure 22. Addition of a poly-A tail to the 3' end of pre-mRNA. Many adenine-containing nucleotides are added.

- Regulation **Importance of poly-A tail at 3' end of mRNA:**

(i) enhances the half-life/stability of the mRNA transcript by slowing down its degradation by ribonucleases in nucleus and cytoplasm.
(ii) serves as a signal to direct export of mature mRNA from nucleus to cytoplasm.
(iii) interacts with the 5' cap for translation by recruiting initiation factors (proteins) to form the translation initiation complex.

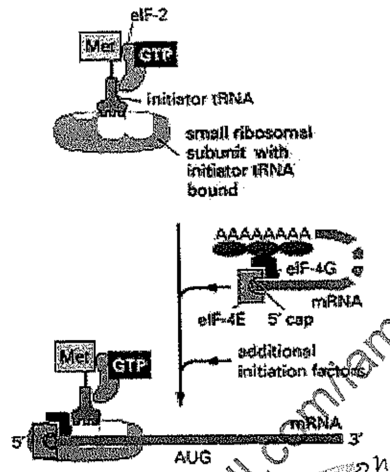


Figure 23. Interaction between eIF (eukaryotic initiation factor) at 5' cap and the poly-A tail is necessary for translation initiation. Translation initiation factors will bind to the 5' cap and poly A tail causing the mRNA to circularise. At the same time, other translation initiation factors will bind to initiator tRNA and small ribosomal subunit to form a complex. This complex will then scan the mRNA to locate the AUG start site.

(4) Control @ translational level (Euk)

→ had longer half life than Prok.

(i) mRNA stability/ half-life

- Regulation Stability/half-life of an mRNA molecule in eukaryotes is influenced by factors such as length of poly-A tail. Like in prokaryotes, controlling half-lives of mRNAs ensures proteins that are not needed in large amounts are not frequently translated from their mRNAs.
- The more stable the mRNA, the longer its half-life, hence the longer it can be used as a template for translation.
- In mRNA degradation, the poly-A tail is steadily removed by a ribonuclease in the 3' to 5' direction in cytoplasm. Once a critical length is reached (~ 30 remaining A), it will trigger the removal of 5' cap by a different ribonuclease.

(ii) Binding of small ribosomal subunit

- During translation initiation, small ribosomal subunit binds to 5' cap of mRNA.

- **Regulation** Translation initiation can be blocked by a translational repressor protein that binds to:
  (a) 5' cap and/or its vicinity i.e. 5' untranslated region (UTR).
  (b) 3' untranslated region to interfere with the interaction between 3' poly-A tail, initiation factors and 5' cap which is needed for translation.

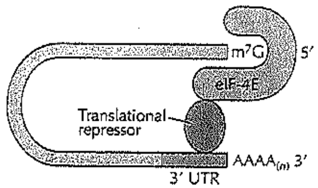


Figure 24. Translational repressors can block translation initiation.
Interaction between eIF (eukaryotic initiation factor) at 5' cap and the poly-A tail is necessary for translation initiation. Typically translation initiation factors will bind to the 5' cap and poly A tail causing the mRNA to circularise. If the 5' UTR or 3' UTR is blocked by a translation repressor protein, translation will not occur.

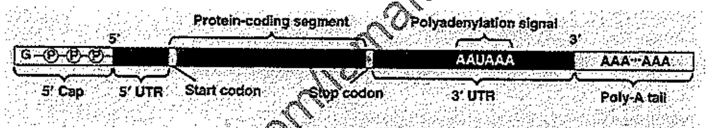


Figure 25. The final product - a mature mRNA ready for translation in cytoplasm. UTR: untranslated regions.

(iii) Initiation factors

- Initiation factors are needed to begin protein synthesis. They are required for proper positioning of the small ribosomal subunit together with initiator tRNA on mRNA, and the subsequent recruitment of large ribosomal subunit. Eukaryotic initiation factors are different from those used by prokaryotes and are greater in quantity.

- **Regulation** Availability of initiation factors is determined by whether they are phosphorylated/ dephosphorylated.

Additional info: (but that doesn't mean it won't be tested as info here may help you answer questions)

e.g. eIF2 (eukaryotic initiation factor 2) is inactivated when phosphorylated; conversely eIF4 is activated when phosphorylated.

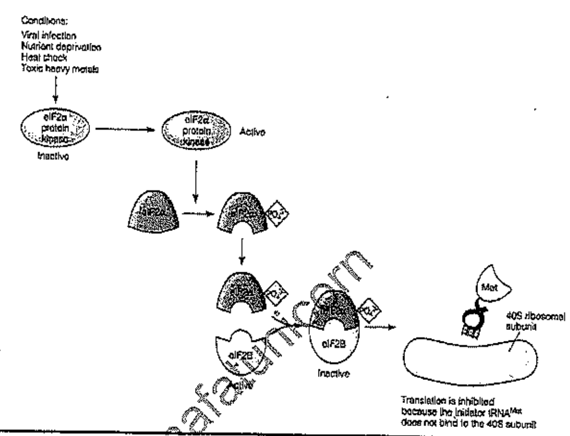


Figure 26. Phosphorylation of eIF2 (caused by certain stress conditions) leads to inhibition of translation initiation.

(5) Control @ post-translational level (Euk)

Controlling proteins that are already present in the cell by activating or inhibiting function of the protein.

(i) Covalent modifications to form functional proteins

- Many newly synthesised polypeptides cannot immediately serve as functional proteins.

- **Regulation** Further processing by cleavage and/or covalent modification such as glycosylation, disulfide bond formation, and attachment of prosthetic groups etc. is required.
  - For eukaryotes, these modifications occur when polypeptides pass through rough endoplasmic reticulum (RER) and Golgi apparatus.

(ii) Phosphorylation/ dephosphorylation to regulate protein activity

- **Regulation** Activity of proteins within a cell is commonly regulated through **phosphorylation** (addition of phosphate group) or **dephosphorylation** (removal of phosphate group) of the proteins.

- As mentioned earlier, for some eukaryotic initiation factors, addition of a phosphate group to a protein can render it active or inactive, while dephosphorylation has the opposite effects.

- Phosphorylation and activation of kinases in phosphorylation cascade can be used to transduce a signal from the extracellular environment to the intracellular environment such that an appropriate cellular response is produced. (*To be covered in more detail under the topic of Cell Signaling)*

(iii) Protein degradation

- Regulation: This mechanism determines how long the protein will remain in the cell to carry out its function. Proteins which are no longer needed will be degraded immediately.

- To tag a protein for degradation, an enzyme, ubiquitin ligase, will catalyse addition of a protein, ubiquitin, to the target protein. The ubiquitin-tagged protein is then recognised by a proteasome which can cleave this protein into smaller peptides that can be further degraded by enzymes in the cytoplasm.

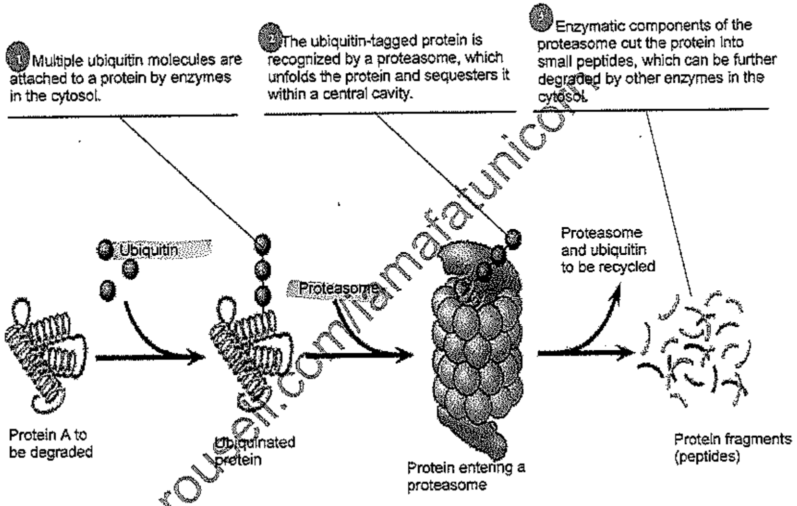
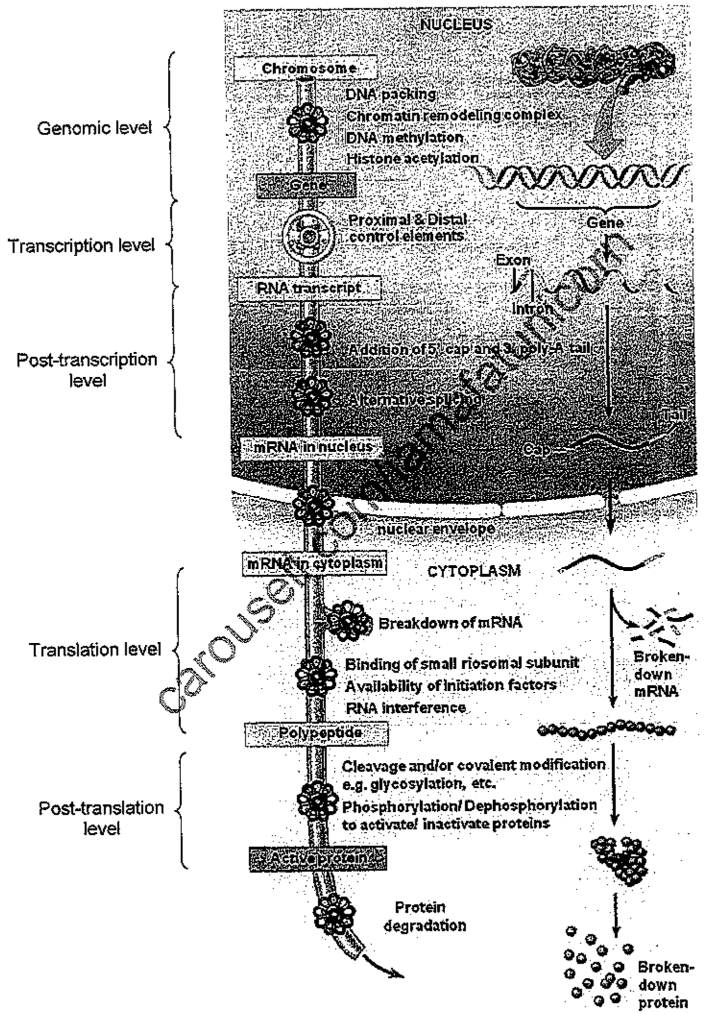


Figure 27. Protein degradation by proteasome with the help of ubiquitins.

Summary (for eukaryotes)



Keywords include:

Promoter	Consensus sequence	Translation initiation complex	Silencers
RNA polymerase	Frequency of transcription	mRNA	Specific transcription factors
General transcription factor	Shine Dalgarno sequence	Chromatin remodeling complex	TATA box
Transcription initiation complex	Translation initiation factors	DNA methylation	Activators
-10 sequence/ Pribnow box	AUG, start codon	Histone acetylase	Repressors
-35 sequence	UAG, UAA, UGA, Stop codon	Histone deacetylase	5' cap
RNA polymerase holoenzyme	Small ribosomal subunit	Negatively charged DNA	Poly A tail
Sigma factor	Large ribosomal subunit	Enhancers	

CORE IDEA
(2) Genetics & Inheritance

ORGANISATION AND CONTROL OF PROKARYOTIC AND EUKARYOTIC GENOME (III)

CANCER

Content
The Molecular Biology of Cancer

Learning Outcomes
Candidates should be able to:
2 (o) Explain the significance of the mitotic cell cycle (including growth, repair and asexual reproduction) and the need to regulate it tightly (This was covered under mitosis). (Knowledge that dysregulation of checkpoints of cell division can result in uncontrolled cell division and cancer is required, but detail of the mechanism is not required.)
2 (p) Identify the causative factors, including genetic, chemical carcinogens, ionising radiation and loss of immunity, which may increase the chances of cancerous growth.
2 (q) Explain how the loss of function mutation of tumour suppressor genes, including p53, and gain in function mutation of proto-oncogenes, including ras, results in uncontrolled cell division.
2 (r) Describe the development of cancer as a multi-step process that includes accumulation of mutations, angiogenesis and metastasis.

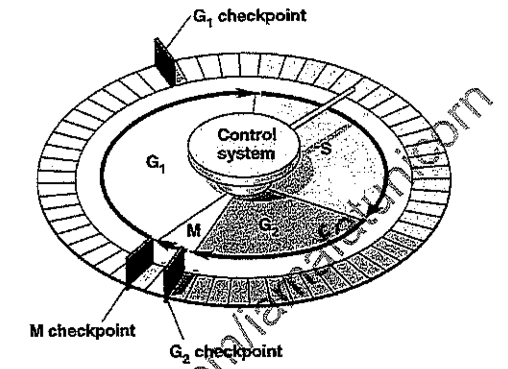
References
Campbell, N.A. and Reece J.B. (2008). Biology (8th Ed), Pearson Education, Inc.
Brooker, R.J. (2005) Genetics: Analysis and Principles (2nd Ed), McGraw-Hill.
Alberts, B. *et al.* (2008) Molecular Biology of The Cell (5th Ed), Garland Science.
Raven et al. (2008). Biology (8th Ed), McGraw-Hill.

Table of Contents

* **This handout is the effort of several Biology teachers at RI. It has and will continue to be updated.**
** **Any information given in a double-lined box is for your information only.**

(A) Introduction

- Cancer is a group of diseases characterised by **uncontrolled cell division** and **spread** of abnormal cells.
- The cell cycle is well regulated as it is important for normal growth and development. It is regulated at certain control points known as **checkpoints**. At these checkpoints, stop and go-ahead signals can determine whether or not the cell cycle can proceed.
- The main checkpoints are at G_1, G_2 and M phase.



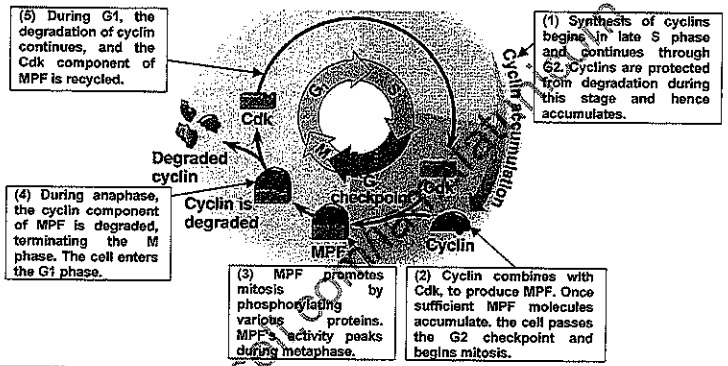
- Cancer occurs when the **dysregulation of checkpoints of cell division** occur or cells **escape the cell cycle control mechanism** that normally regulates their growth. This leads to **uncontrolled division** of cells.
- Eventually a mass of cells called a **tumour** can result. Tumour cells are **genetically identical** and **derived from a single, genetically altered cell (i.e. mutant cell).**
- Cancerous cells carry genetic mutations. Unlike normal cells, cancerous cells escape control mechanisms that normally limit their growth and division. Mutations in two groups of genes, ***proto-oncogenes***** and ***tumour suppressor genes*****, contribute to many kinds of cancer.

What do gene mutations do?

A gene mutation can instruct a healthy cell to:
- **continue to divide.** ***Proto-oncogenes*****, when they become ***oncogenes*****, can tell a cell to grow and divide more rapidly. This creates many new cells that all have that same mutation.
- **fail to stop uncontrolled cell growth.** Normal cells know when to stop dividing so that you have just the right number of each type of cell. Cancer cells lose the controls due to mutations in the ***tumor suppressor genes***** that tell them when to stop dividing.
- **make mistakes when repairing DNA errors.** DNA repair genes look for errors in a cell's DNA and make corrections. A mutation in a DNA repair gene may mean that other errors are not corrected, causing cells to accumulate mutations and become cancerous.

Check points	Checkpoint is passed if
G_1	• DNA is not damaged
G_2	• DNA is not damaged • chromosomes have replicated successfully • accumulation of sufficient cyclins to bind with Cdks to form MPF
M	• all chromosomes are attached to spindle fibres from both poles

- One way in which regulation occurs at the G_2 checkpoint is as follows.
  - Two types of regulatory proteins are involved in the control of the cell cycle: **cyclins** and **cyclin-dependent kinases (Cdks)**.
  - Cyclins need to bind to Cdks to activate the Cdks.
  - Cdk activity levels hence depend on the levels of its cyclin partner.
  - **MPF (maturation-promoting factor)** is a **cyclin-Cdk complex** that triggers a cell's passage past the G_2 checkpoint into the M phase.



(B) Causative factors of cancer

What causes gene mutations?

- **Gene mutations you're born with.** Gene mutations that increase cancer risk can be inherited from our parents if the mutations are present in germ cells, which are the reproductive cells of the body (eggs and sperm).
- **Gene mutations that occur after birth.** Cancer-causing mutations can also be acquired during one's lifetime, as the result of errors that occur as cells divide during a person's lifetime. Other factors can also cause gene mutations such as exposure to substances (chemicals in tobacco smoke) or exposure to radiation, such as ultraviolet rays from the sun, that can damage DNA.

1. Heredity causes

Example: Mutation in the *BRCA1* gene in humans

- *BRCA1* is a tumor suppressor gene found in all humans and located on the long arm of chromosome 17.
- *BRCA1* produces a normal protein product that is responsible for repairing DNA. If *BRCA1* is mutated, defective BRCA1 protein is unable to fix DNA damage leading to mutations in other genes, increasing the risk for cancer. The mutations can accumulate within a single cell and lead to ***uncontrolled cell division***.
- BRCA1 has been found to have roles in DNA repair, ubiquitination and transcriptional regulation.
- Researchers have identified hundreds of mutations in the *BRCA1* gene, many of which are associated with an increased risk of cancer. Women with an abnormal *BRCA1* gene have up to an 80% risk of developing breast cancer by age 90 and a 55% risk of developing ovarian cancer. In addition to breast and ovarian cancer, mutations in the *BRCA1* gene also increase the risk of prostate cancers.

2. Chemical carcinogens

Example 1: Smoking (Fig. 1)

- Smoking has been correlated to an increased risk in developing lung cancer. The carcinogens in cigarette smoke belong to multiple chemical classes, including polycyclic aromatic hydrocarbons, N-nitrosamines, aromatic amines, aldehydes, volatile organic hydrocarbons, and metals.

  Most carcinogens in cigarette smoke require a metabolic activation process to convert the carcinogens to forms that can covalently bind to DNA forming DNA adducts.

  Competing with the activation process is metabolic detoxification, which excretes carcinogen metabolites in generally harmless forms. If we are less able to detoxify the carcinogens, then we could be more susceptible to cancer.

  Normally there are ample cellular repair systems that can remove DNA adducts and maintain normal DNA structure. However if repair enzymes are overwhelmed by DNA damage or for other reasons cannot function efficiently, DNA adducts may persist and increase the likelihood of developing somatic mutations.

  Persistent DNA adducts can cause miscoding (e.g., insertion of the wrong base) during DNA replication when DNA polymerase enzymes process the adducts incorrectly.

- Another cause of cancer is due to nicotine and tobacco-specific nitrosamines that bind to nicotinic receptors and other cellular receptors. This binding then leads to the activation of protein kinase B and protein kinase A which in turn could activate signaling pathways in cells resulting in uncontrolled cell division.
- When DNA damage occurs due to the chemical carcinogens in cigarette smoke, DNA methyltransferases will be recruited. These enzymes may cause hypermethylation of promoter regions of genes, which can result in gene silencing. If this occurs in tumor-suppressor genes, the result can be uncontrolled cell division.

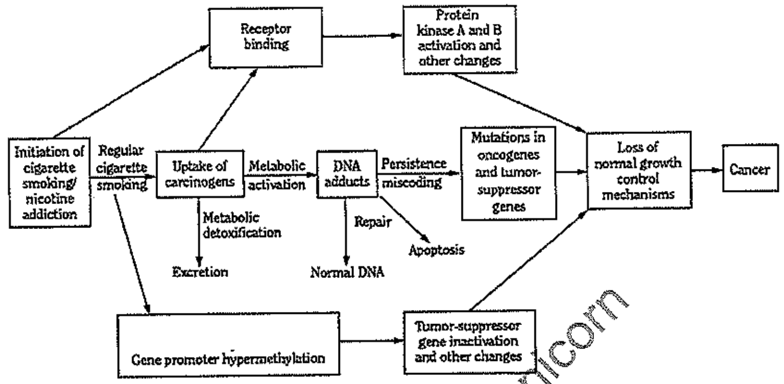


Figure 1. Link between cigarette smoking and cancer through carcinogens in tobacco smoke

Example 2: Asbestos

- Asbestos can be found in some construction materials, such as in vermiculite insulation. Since the early 1970s, use of asbestos has declined substantially due to health concerns.
- Asbestos fibers can be released into the air when asbestos-containing materials are disturbed or degraded.
- People exposed to asbestos are at risk of developing mesothelioma, a malignant cancer of the membranes that cover the lungs and abdomen.
- When asbestos fibers kill cells, they do so by inducing a process that leads to the release of a protein that starts a particular type of inflammatory reaction. This causes the release of mutagens and factors that promote tumor growth.

3. Ionising radiation

Example: Ultraviolet light

- Wavelengths of both ultraviolet A (UVA 320-400nm) and ultraviolet B (UVB 280-320nm) radiation have been implicated as carcinogens. The two wavelengths of radiation are able to penetrate to different depths of the skin and hence affect different cells in the epidermis and dermis.
- UVB radiation is mainly absorbed by epidermal components such as proteins or DNA, whereas UVA radiation penetrates deeply into the skin and reaches the lower epidermis and dermal fibroblasts.
- UVA radiation's toxicity mainly comes from oxidative damage to skin cell components.
- UVB radiation directly damages the DNA within skin cells. DNA damage occurs when the chemical bonds within a DNA molecule are altered. A photon of UVB radiation penetrates the cell and could cause the hydrogen bond between the nitrogenous bonds to break. The unbonded

base then interacts with adjacent bases on the same DNA strand to create new bonds to form dimers.

- The effects of UVB on DNA are mostly caused by the formation of these dimers between two adjacent pyrimidines (cytosine and/or thymine) on the same DNA strand. (Fig. 2).

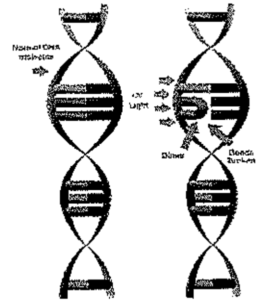


Figure 2. DNA damaged by UVB radiation forms a dimer

- These reactions occur hundreds of times within seconds of exposing skin to sunlight (approximately 80,000 dimers per cell within one hour of sunlight exposure), though most of this damage is temporary. The human body has inbuilt systems of DNA repair, to recognise and eliminate such changes. Almost immediately, the section surrounding the flawed segment is excised and replaced by the correct sequence. Occasionally such errors are left unrepaired or incorrectly repaired and the damage becomes a permanent mutation which can lead to the development of skin cancer.
- DNA mutations resulting from unrepaired or misrepaired pyrimidine dimers frequently arise in the *p53* tumor suppressor gene in skin cancers (see section C2). The protein produced from the *p53* gene in a healthy cell stops the cell cycle so that DNA damage can be repaired prior to the cell's replication. Failing this, it has an important role in the pathway leading to apoptosis. Mutated *p53* genes will not be able to stop the cell cycle or cause apoptosis in cells with significant mutations, esulting in accumulation of mutations in these cells which may then divide uncontrollably.
- UV-specific *p53* mutations have been reported in 50% of human basal cell carcinoma and in over 90% of squamous cell carcinoma (the most common types of skin cancer), making them the mutations most frequently found in skin cancer patients.

4. Loss of immunity

- People infected with HIV are several thousand times more likely than uninfected people to be diagnosed with Kaposi sarcoma, at least 70 times more likely to be diagnosed with non-Hodgkin's lymphoma, and, among women, at least 5 times more likely to be diagnosed with cervical cancer
- The connection between HIV and certain cancers is not completely understood, but the link likely depends on a weakened immune system. **Infection with HIV weakens the immune system and reduces the body's ability to fight infections that may lead to cancer.** Many people infected with HIV are also infected with other viruses that cause certain cancers. The following are the most important of these cancer-related viruses:
    - Human herpesvirus 8, also known as Kaposi sarcoma-associated herpes virus, is the cause of

Kaposi sarcoma.

- Epstein Barr virus causes some subtypes of non-Hodgkin's and Hodgkin's lymphoma.
- Human papillomavirus causes cervical cancer.
- Hepatitis B virus and hepatitis C virus both can cause liver cancer.

- Infection with most of these viruses is more common among people infected with HIV than among uninfected people.

How does Kaposi Sarcoma-associated Herpes Virus cause cancer

- Kaposi's sarcoma is a type of skin cancer, which has traditionally occurred in older men of Jewish or Mediterranean descent, young men in Africa, or people who have received organ transplantation.
- Kaposi's sarcoma causes lesions to arise in multiple sites in the body, including the skin, lymph nodes, and organs such as the liver, spleen, lungs, and digestive tract.
- When a person is infected with this DNA virus, the viral genome is integrated into that of the host cell. Within the viral genome, there are genes that control that DNA synthesis, affect cell division and also affect the tumour suppression pathways which could result in tumour formation.

How does the Epstein Barr Virus cause lymphoma

- Lymphoma is a group of blood cell tumors that develop from lymphatic cells. Hodgkin's lymphoma is a type cancer which originates from a specific type of white blood cells called lymphocytes. Non-Hodgkin lymphomas are diverse group of blood cancers that include any kind of lymphoma except Hodgkin's lymphomas.
- The Epstein-Barr virus (EBV) is the first human virus identified with a proven association with the pathogenesis of cancer. EBV preferentially infects B lymphocytes.
- According to current knowledge, antigens encoded by EBV interfere with a number of important cellular pathways, thereby leading to tumour formation. These EBV antigens have been found to immortalise B cells by facilitating p53 degradation, enhancing transcription of certain host and viral genes, blocking apoptosis and affecting chromatin remodeling processes.

(C) Molecular basis of cancer

1. Proto-oncogenes

- ***Proto-oncogenes**** are **normal cellular genes** which codes for proteins that are involved in regulation of cell growth and proliferation.
- Proto-oncogene products
  - are proteins derived from proto-oncogenes.
  - are involved in **stimulating normal cell growth and division**.
  - e.g. growth factors, growth factor receptors, transcription factors etc.
- Mutation : proto-oncogenes have the potential to become oncogenic
  - when proto-oncogenes mutate, they are known as ***oncogenes****.
  - the mutation results in an **increase** in
    (i) **amount** of a proto-oncogene's protein product or
    (ii) **intrinsic activity** of that protein product.
  - proto-oncogenes are converted to oncogenes through ***gain-in-function mutations**** (see Page 8).

Mutation mechanisms that leads to the formation of oncogene

<table>
<tr><td>A. Gene mutations
(Fig. 3)</td><td>
▪ Point mutation within the proto-oncogene changes amino acid sequence of the proto-oncogene protein.
e.g.: a regulatory protein (such as a growth factor encoded by proto-oncogene)
→ becomes hyperactive or more resistant to degradation.

▪ Point mutation in base sequences of regulatory elements (e.g. promoters of proto-oncogenes)
→ leads to increased transcription, and excess production of the growth-stimulating protein.

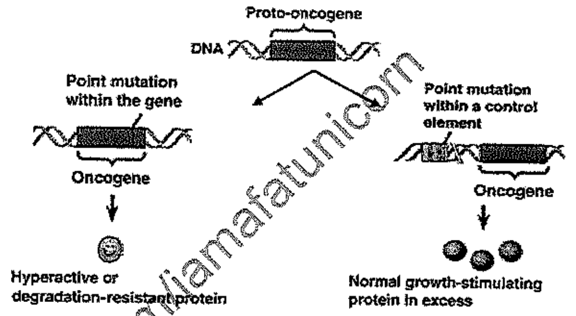


Figure 3. Consequence of point mutation(s) within a control element of a gene, or the gene itself
</td></tr>
<tr><td>B. Gene amplification
(Fig. 4)</td><td>
▪ The normal DNA replication process is flawed. Instead of making a single copy of a region of a chromosome, many copies are produced.
→ leads to the production of many copies of the genes that are located on that selected region of the chromosome.

▪ If a proto-oncogene is included in the amplified region, the number of proto-oncogenes in a cell is amplified
→ leads to excessive production of proto-oncogene protein, which can de-regulate cell growth, promoting excessive cell division.

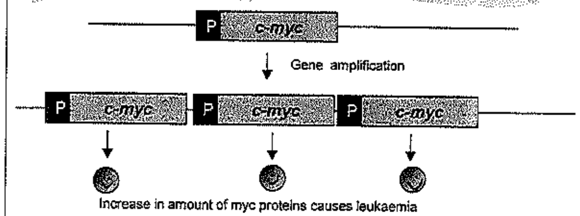


Figure 4. Consequence of gene amplification of the myc gene – an increase in the amount of proto-oncogene's protein product (key: c means cellular).
</td></tr>
</table>

C. Movement of DNA

(a) Chromosomal translocation (Fig. 5)

- Involves unusual exchange between chromosomes, e.g. chromosomes break and join at another location, crossing over between non-homologous chromosomes.
- If a proto-oncogene ends up under the control of an enhancer
  → leads to increased transcription and production of more gene products.

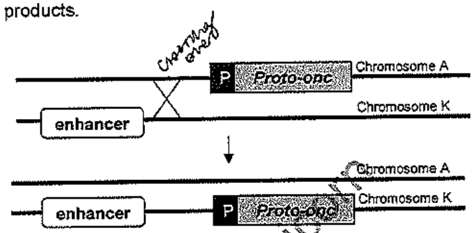


Figure 5. Abnormal crossover between non-homologous chromosomes can cause a common proto-oncogene to become oncogenic. The proto-oncogene now falls under the influence of a highly active enhancer, upregulating its gene expression.

(b) Retroviral integration (Fig. 6) e.g. HIV

- Viruses can cause cancer in animals and humans through several mechanisms.

↑ amt of protein
↑ intrinsic activity of protein

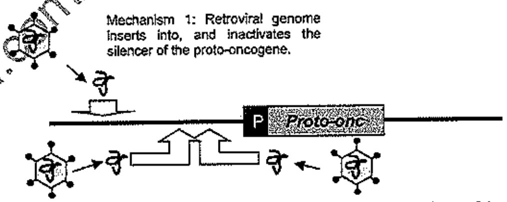
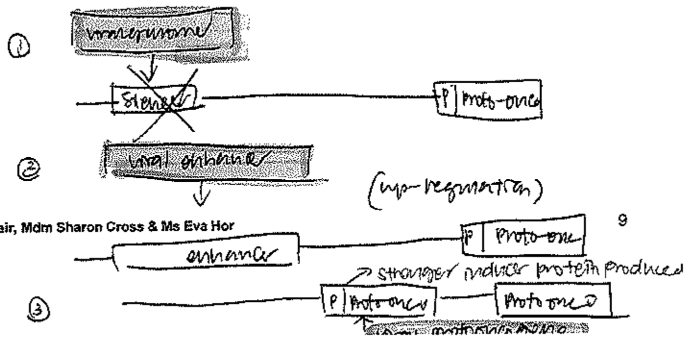


Figure 6. Viral genome integrating into host cell genome mutates a common proto-oncogene into an oncogene.

gain-of-function mutation



Example of proto-oncogene: *ras* gene ON by default

- The *ras* gene codes for Ras proteins which are involved in signal transduction pathways (Fig. 7).
  - the binding of a growth factor to the receptor triggers a series of reactions inside a cell. This results in binding of GTP to an inactive Ras protein, thereby activating the Ras protein.
  - the activated Ras proteins **transduce signals from the growth factor** to downstream signaling processes, thereby leading to cell division.
    (Details of cell signaling will be covered in the future).
  - the Ras protein cannot stay active for too long, since cell division cannot go on indefinitely. At the appropriate time, the Ras protein becomes inactive again. When there is a need for the cell to divide/when growth factor binds again, the Ras protein will be activated.

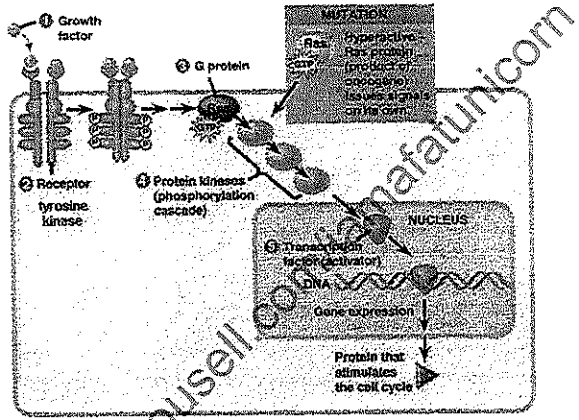


Figure 7. Ras protein in transducing signals from growth receptors.

- Mutation in *ras* gene
  - can be caused by any of the mechanisms in Fig. 3 to 6.
  - results in a constitutively active Ras protein. This leads to **increased cell division** even when the growth factor does not bind to the receptor.

2. Tumour suppressor genes

- *Tumour suppressor genes** are normal cellular genes that code for products which **inhibit cell division** and **helps to prevent *uncontrolled cell division****.
- Products of tumour suppressor genes activate cell cycle arrest, DNA repair and/or apoptosis (programmed cell death).
- Tumour suppressor genes contribute to cancer when there is a ***loss-of-function mutation****.

Example of tumour suppressor gene: *p53* gene

- The ***p53* gene*** is the most commonly mutated gene in human cancers. About 50% of all human cancers are associated with mutation in this gene.
- The *p53* gene codes for a **specific transcription factor (p53)** that binds to DNA to promote synthesis of cell cycle-inhibiting proteins.

The normal role of p53

- When there is DNA damage, the *p53* gene is activated to produce p53 proteins. As a **specific transcription factor**, the p53 protein can activate genes that are involved in:
  (a) cell cycle arrest
    - cell cycle is halted so that there is enough time for the cell to repair its damaged DNA and prevent production of mutant daughter cells
  (b) DNA repair:
    - preserves genomic integrity and prevents mutations that may lead to formation of oncogenes or inactivate other tumor-suppressor genes
  And when DNA damage is beyond repair:
  (c) initiate apoptosis (cell death)
    - Apoptosis is a process where cell shrinks, DNA and cellular structures get degraded. This is important as it removes cell with potential to cause cancer.

Mutation in p53 gene (Fig. 8)

- a defective p53 protein will not restrict cell growth and proliferation. Instead, cell cycle continues without repairing DNA/ cell does not undergo apoptosis.
- When mutations accumulate over time, organisms run an increased risk of developing cancer cells.
- Mutations in p53 could be due to point mutation or chromosomal translocation or retroviral integration.

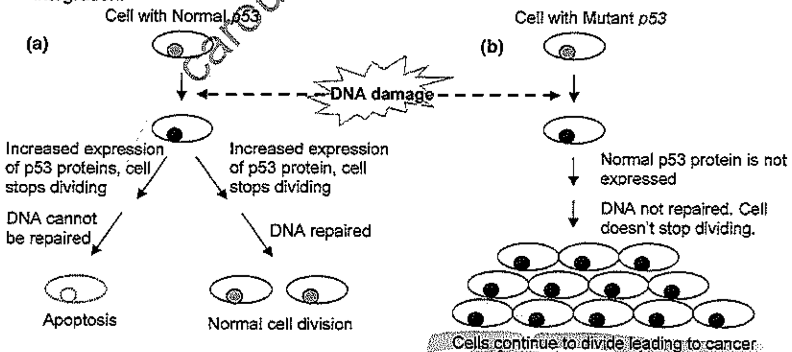
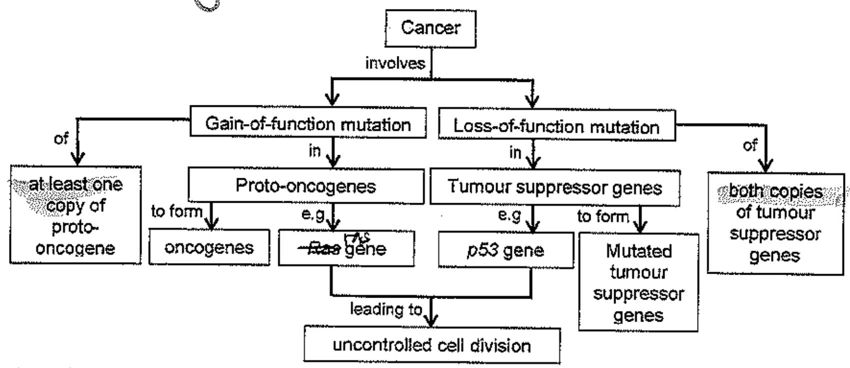


Figure 8. Mutations in p53 can lead to tumour formation.
(a) In a normal cell, if the DNA is damaged, the levels of p53 protein are increased. p53 protein activates genes to halt the cell cycle, repair damaged DNA or bring about apoptosis. After the DNA is repaired, the cell continues to divide normally. (b) However, if the damaged DNA is found in a cell with a mutant *p53* gene, the cell will continue to divide, leading to perpetuation of genetic mutations and eventually to cancer.

3. Gain-in-function mutation vs loss-in-function mutation

	Gain-in-function mutation	Loss-in-function mutation
Definition	- A mutation that causes a gene to be expressed in a cell/ at a time when it is not normally expressed - Gene products of proto-oncogenes become hyperactive/ resistant to degradation/ are produced in excessive amounts	- A mutation that causes a gene product to be non-functional. - Gene products of tumour-suppressor genes (such as p53 protein) are defective and cannot activate other genes.
Type of genes the mutation affects	Affects proto-oncogene, e.g. *ras* gene	Affects tumour suppressor gene, e.g. *p53* gene
Number of alleles that has to be mutated in order to be cancerous	Only one copy of the allele need to be mutated. This is known as a dominant mutation. - mutation in just one copy is enough to give extra gene products (such as growth factor). - A normal cell is sensitive to the amount of such gene products. Any extra means cell cycle escapes normal control.	Both copies of allele need to be mutated. This is known as a recessive mutation. - if one is mutated, there is still another copy of the normal allele that can function normally (e.g. its products can still activate other genes for DNA repair).
Effect on cell cycle (in cancer cells)	Overstimulate cell cycle	Can't stop cell cycle to repair damages to DNA. Cells with accumulated mutations keep dividing.
Type of mutation	Dominant mutation	Recessive mutation

Table 1. Two kinds of mutations that affect proto-oncogenes and oncogenes.



(D) Physiological basis of cancer

- Cancer can be caused by spontaneous genetic mutations, viruses, environmental factors (including carcinogens such as UV rays) and genetic predisposition.

Cancer and tumours

- cancer has often been closely associated with the formation of tumours, but having a tumour ≠ cancer.
- there are 2 types of tumours
  1. ***benign**** **tumour** - mutations in several genes that are involved in the regulation of cell growth, division and cell death can result in a big mass of cells that keeps on dividing and growing, and does not die off readily. However this tumour is localized and does not spread to other regions. This is known as a benign tumour and is **non-cancerous**.
  2. ***malignant* tumour**** - a benign tumour can transform into a **cancerous** one i.e. a malignant tumour. A malignant tumour results when more **mutations accumulate** from the **continual cell divisions** to allow the tumour to acquire two related traits:
     - **invasiveness** (erodes normal surrounding tissue) and
     - **ability to undergo metastasis** (can spread to other parts of the body): tumours originate from a single aberrant cell that proliferates to give rise to a **primary tumour** whose cells eventually metastasise to other parts of the body to form **secondary tumours** (Fig. 9).

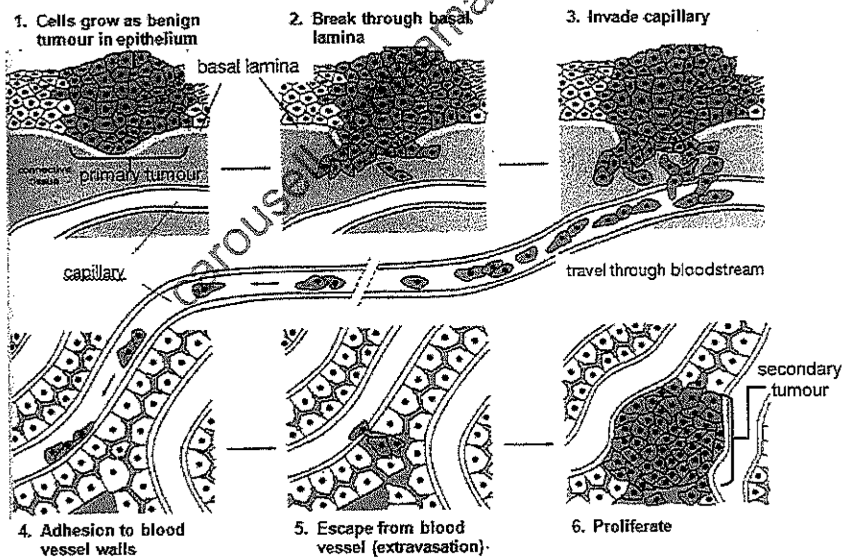


Figure 9. Spread of a malignant tumour.

Differences between normal and cancer cells

	Normal cells	Cancerous cells
1. Proto-oncogene	Normal cells have **proto-oncogenes**, whose functions are to promote the normal growth and division of cells. Rate of cell division is regulated.	Cancerous cells have **oncogenes**, which are the mutated form of proto-oncogenes. Gain of function mutation result in excessive cell division.
2. Cell division	Show **controlled cell division.**	Oncogenes cause **uncontrolled cell division.**
3. Nuclei / mutation	**Normal nuclei** present.	Cancer cells have **abnormal nuclei.** The **chromosomes have mutated**; some parts of the chromosome may be **duplicated** and some **deleted**.
4. Apoptosis	Normal cells show **programmed cell death/apoptosis.** They divide for a certain number of times and then stop dividing.	Cancerous cells **do not show apoptosis.** They can divide indefinitely.
5. Contact inhibition	Normal cells show **contact inhibition.** This means that they do not divide further when in contact with other cells.	Cancerous cells **do not show contact inhibition.**
6. Differentiation	Normal **cells differentiate** to become **specialised cells.** Specialised cells, such as nerve cells, do not divide.	Cancerous **cells fail to differentiate** properly.
7. Tumour suppressor gene	Presence of **tumour suppressor genes.**	**Tumour suppressor genes are absent** or have been **mutated** (resulting in loss-in- function of tumour suppressor genes).
8. Cell adhesion / metastasis	Cell adhesion →Formation of tissue and organs	Can detach from surrounding cells
9. Angiogenesis – formation of blood vessel	Does not stimulate new blood vessels	Stimulates growth of new blood vessels within tumours.
10. Nucleo-cytoplasmic ratio	High nucleo-cytoplasmic ratio	Low nucleo-cytoplasmic ratio

Figure 10. Main differences between normal cells and cancerous cells

Why is development of cancer a multi-step process?

- The development of cancer requires the **accumulation of mutations** in the genes which **control regulatory checkpoints** of the cell cycle in a **single cell**.
- This will **disrupt** the **normal cell cycle**, thus causing the cell to undergo **excessive cell growth and proliferation** resulting in ***uncontrolled cell division****.
- A ***gain-in-function mutation**** is a ***dominant**** mutation where mutation in just **one copy/allele of a proto-oncogene** will result in its **overexpression** which will result in the production of **excessive amounts of growth factors, or production of hyperactive/degradation resistant growth factors**, leading to cell proliferation
- ***Loss-of- function mutation**** is a ***recessive**** mutation where mutation in **both copies/alleles of a tumour suppressor gene** will **disrupt their ability to inhibit cell cycle, enable DNA repair and promote apoptosis**.
- **Upregulation/activation** of the genes coding for **telomerase** result in telomeres being lengthened and the cell can thus **divide indefinitely** as the chromosomes are prevented from shortening with each DNA replication cycle.
- Multiple mutations that are unrepaired in a single cell that is immortal and continues to divide will allow the cells to accumulate more and more mutations.
- **Loss of contact inhibition** will enable the cells to **grow into a tumour/mass of cells**.
- **Angiogenesis** is the growth of new blood vessels and must occur within the tumour so that the **blood vessels** formed can **transport oxygen and nutrients to it for its growth and remove the metabolic waste products and carbon dioxide**. Without a blood supply, tumours cannot grow as diffusion alone is insufficient to provide for the tumour. New growth in the blood vessels also contributes to metastatic spread of cancer.

  In adults, angiogenesis is only 'switched on' during physiological processes such as wound healing. However, in cancer cells, angiogenesis is always 'switched on'.

  Lack of oxygen (hypoxia) could result due to rapid growth of the tumour, thus stimulating the production of angiogenic growth factors by cancer cells. These growth factors will bind to receptors on endothelial cells (cells that line the interior of blood vessels). Endothelial cells will then detach from its surrounding tissues and move towards the tumour resulting in development of blood vessels.
- Finally the cancer cells must **metastasise**. i.e. **leave the primary site** and **spread to other tissues in different parts of the body** via the **blood stream** and **form secondary tumours** there.
- The above steps should occur for cancer to develop.
- As it takes years to **accumulate** these mutations, developing cancer **increases with age**.

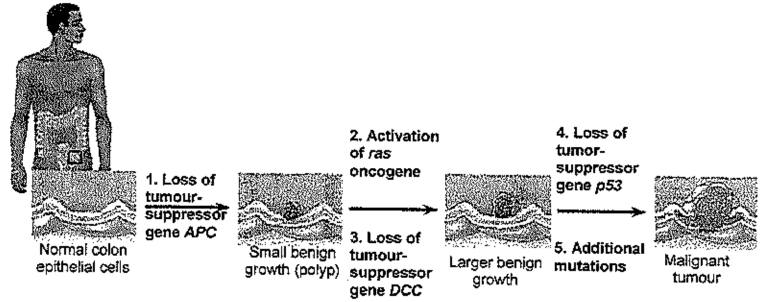


Fig. 11. Multi-step model of cancer development – colorectal cancer. The first sign is often a polyp in the colon lining. Through gradual accumulation of mutations that mutate proto-oncogenes to oncogenes and knock out of tumour suppressor genes, the polyp can develop into a malignant tumour. A *ras* oncogene and a mutated *p53* tumour suppressor gene are often involved.

The End

Additional info: (but that doesn't mean it won't be tested as info here may help you answer questions)

(E) Treatment of cancer

Based on certain hallmarks of cancer, various methods have been developed to treat cancer:

Cancer hallmark	Treatment
Uncontrolled proliferation due to: - self-sufficiency in growth signals - insensitivity to anti-growth signals	Chemotherapy: mitotic inhibitors that prevent assembly of microtubules (e.g. vinblastine & vincristine from *Catharanthus roseus*) arrests mitosis at metaphase, eventually resulting in cell death.
Evasion of apoptosis	The evasion of apoptosis renders tumour cells immortal. To kill tumour cells, classical cancer treatments usually rely on chemo- and/or radiation therapy to induce cell death.
Limitless cell division	Mitotic inhibitors which prevent assembly of microtubules would also circumvent the problem of limitless cell division. Since telomerase activity is necessary for the limitless proliferation of cancer cells, telomerase inhibitors might provide an effective therapy to limit the growth of cancer cells by triggering telomere shortening and cell death.
Angiogenesis	The formation of blood vessels can be inhibited using angiogenesis inhibitors. This limits availability of nutrients and oxygen to tumour cells, thereby limiting tumour growth.
Metastasis	Angiogenesis is an essential component of tumour metastasis, as it provides an efficient route of exit for tumour cells to leave the primary site and enter the bloodstream. Therefore, angiogenesis inhibitors also serve to block metastasis of tumour cells from the primary site, as well as the growth of metastases at secondary sites.

(F) Links

The topic of organization and control of prokaryotic and eukaryotic genome and cancer is relevant to the other topics and learning outcomes in the A level Biology syllabus.

Topic No.	Topic	Comments
1	Cellular Functions (Cell Division)	Cancer cells do not heed the normal signals that regulate the cell cycle. They escape cell cycle checkpoints, undergoing uncontrolled proliferation and cell divisions despite presence of damaged DNA.
3	Genetics of Viruses	Viruses are involved in 10-20% of all cancers. The infection of a cell by such viruses (e.g. retroviruses) leads to the expression of viral proteins that enhance the growth potential/survival of that cell; these viruses either carry a copy of an oncogene or the integration of viral genome leads to altered expression of proto-oncogenes or tumour suppressor genes. Often, the viral infection alone is not sufficient to cause cancer. Over time, coupled with the accumulation of mutations that enhance growth of the cell, cancer may develop.
4	Organisation and Control of Prokaryotic and Eukaryotic Genome (I and II)	In many malignant tumours, the gene for telomerase is activated. Telomerase reverses the shortening of chromosome ends during DNA replication. Production of telomerase in cancer cells removes the natural limit on the number of times the cell can divide.
5	Cellular Physiology and Biochemistry (Cell Signalling & Communication)	Cell signaling pathways can lead to inactivation/activation of protein factors that regulate gene expression.
6	Cellular Physiology and Biochemistry (Cell Signalling & Communication)	The proteins encoded by many proto-oncogenes and tumour suppressor genes are components of cell-signaling pathways, e.g. ras protein and p53 respectively. These proteins convey external signals to the DNA in the cell's nucleus, leading to cell division or inhibition of cell division respectively. The ras protein is a G-protein that relays a signal from a growth factor receptor (receptor tyrosine kinase) on plasma membrane to a cascade of protein kinases. The cellular response at the end of the pathway is the synthesis of a protein that stimulates the cell cycle.

Keywords include:

Carcinogen	constitutive	Tumour
Uncontrolled cell division	*p53* gene	Angiogenesis
Proto-oncogene	Dominant	P53 protein
Tumour suppressor gene	Recessive	Contact inhibition
Gain in function mutation	Benign	Apoptosis, programmed cell death
Loss of function mutation	Malignant	Cell-cell adhesion
Ras gene	Metastasis	
Ras protein		