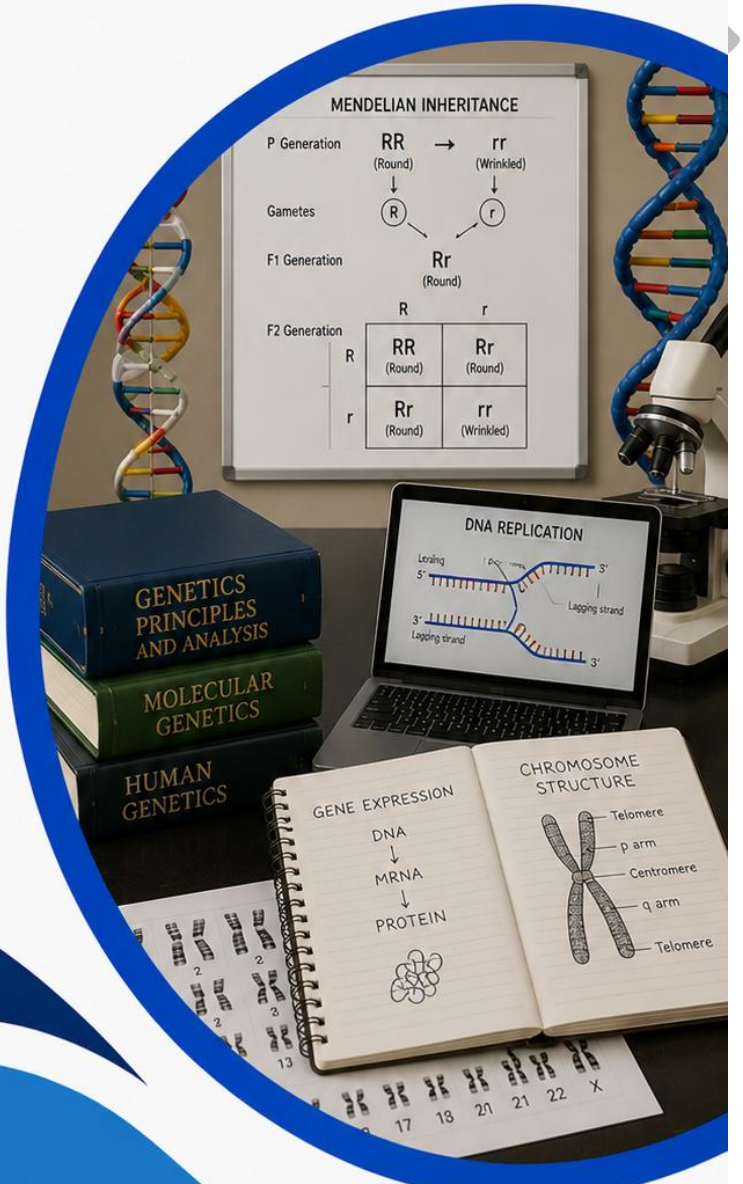
La-Net

BIO 301

GENETICS II



Downoad on the App Store

GET IT ON Google Play

Course video is available on our app

lanetonline.com

Course code: BIO 301

Course title: Genetics II

Course credit load: 2 Units

Series 1: Human Genetics and Pedigree Analysis

- **Part 1.1: Aspects of Human Genetics**
  - Sub Part 1.1.1: The human karyotype and chromosomal disorders
  - Sub Part 1.1.2: Autosomal inheritance patterns
  - Sub Part 1.1.3: Sex-linked inheritance
- **Part 1.2: Pedigree Analysis**
  - Sub Part 1.2.1: Pedigree symbols and conventions
  - Sub Part 1.2.2: Identifying inheritance patterns from pedigrees
  - Sub Part 1.2.3: Worked pedigree analysis examples
- **Part 1.3: Chromosomal Abnormalities**
  - Sub Part 1.3.1: Numerical chromosomal abnormalities
  - Sub Part 1.3.2: Structural chromosomal abnormalities
  - Sub Part 1.3.3: Sex chromosome abnormalities
- **Part 1.4: Genetic Counselling and Screening**
  - Sub Part 1.4.1: Principles of genetic counselling
  - Sub Part 1.4.2: Prenatal diagnosis techniques
  - Sub Part 1.4.3: Genetic screening programmes

Introduction

Human genetics applies the general principles of heredity to the human species, with the goal of understanding how traits and diseases are transmitted through families and populations. Unlike experimental organisms, humans cannot be bred to order; instead, genetic analysis in humans relies primarily on pedigree analysis, the systematic examination of family trees showing the presence or absence of traits across generations, combined with chromosomal and molecular investigation.

This Series introduces the human karyotype, autosomal and sex-linked inheritance patterns, the conventions and interpretation of pedigrees, the major categories of chromosomal abnormality, and the principles and techniques of genetic counselling and screening. This foundation supports the gene interaction, molecular genetics, and biotechnology topics addressed in the remaining Series of this course.

Series Learning Outcomes

Upon completing this Series, you should be able to:

- **SLO 1.1** describe the human karyotype and distinguish autosomal from sex-linked inheritance patterns,
- **SLO 1.2** use standard pedigree conventions to analyse family trees and identify inheritance patterns,
- **SLO 1.3** describe numerical and structural chromosomal abnormalities and their phenotypic consequences, and
- **SLO 1.4** explain the principles of genetic counselling, prenatal diagnosis, and genetic screening.

1.1 Aspects of Human Genetics

1.1.1 The human karyotype and chromosomal disorders

The normal human karyotype consists of 46 chromosomes (2n = 46) arranged as 23 pairs: 22 pairs of autosomes (non-sex chromosomes, numbered 1 to 22 from largest to smallest based on size and banding pattern) and one pair of sex chromosomes (XX in females, XY in males); a karyotype is prepared by arresting dividing cells at metaphase, staining the chromosomes, photographing and then arranging the image so chromosome pairs can be systematically examined for number and structural integrity.

The human genome contains approximately 3 billion base pairs and an estimated 20,000 to 25,000 protein-coding genes, distributed unevenly across the 23 chromosome pairs; the sex chromosomes differ markedly in gene content, with the X chromosome carrying approximately 800 to 900 genes (many unrelated to sex determination) while the Y chromosome is much smaller and carries far fewer genes, mostly related to male sex determination and fertility.

Fill in the blank: The normal human karyotype consists of _____ chromosomes arranged as 22 pairs of autosomes and one pair of sex chromosomes.

1.1.2 Autosomal inheritance patterns

Autosomal dominant inheritance is characterised by: the trait appearing in every generation (no skipping), affected individuals usually having at least one affected parent, approximately half the offspring of an affected parent being affected (on average), and equal frequency in males and females; examples include Huntington disease (caused by a dominant allele on chromosome 4 producing an abnormal huntingtin protein), familial hypercholesterolaemia, and polydactyly (extra digits).

Autosomal recessive inheritance is characterised by: the trait often skipping generations (carriers are phenotypically normal), both parents of an affected individual typically being unaffected carriers, approximately one quarter of the offspring of two carrier parents being affected, and equal frequency in males and females; examples include cystic fibrosis (defective CFTR chloride

channel protein), sickle cell disease (abnormal haemoglobin), phenylketonuria (PKU, deficient phenylalanine hydroxylase), and albinism (deficient melanin production).

Fill in the blank: Autosomal _____ inheritance often skips generations since carriers are phenotypically normal, with approximately one quarter of offspring from two carriers being affected.

1.1.3 Sex-linked inheritance

X-linked recessive inheritance is the most common form of sex-linkage in human genetics: because males have only one X chromosome, a single recessive allele on the X is sufficient to produce the affected phenotype (hemizygous males), while females require two copies of the recessive allele to be affected; carrier females (heterozygous, one normal and one recessive allele) are typically unaffected but can transmit the allele; affected males cannot transmit the trait to their sons (who receive the Y from their father) but all their daughters will be carriers.

Examples of X-linked recessive conditions include haemophilia A and B (clotting factor deficiencies), red-green colour blindness (defects in cone photoreceptor pigment genes), and Duchenne muscular dystrophy (deficiency of the dystrophin protein); X-linked dominant conditions (rarer) affect both males and females but are often more severe in males, while Y-linked (holandric) traits pass exclusively from father to all sons and are extremely rare, as the Y chromosome carries very few genes.

Fill in the blank: In X-linked recessive inheritance, _____ males are affected by a single recessive allele because they have only one X chromosome, while carrier females are typically unaffected.

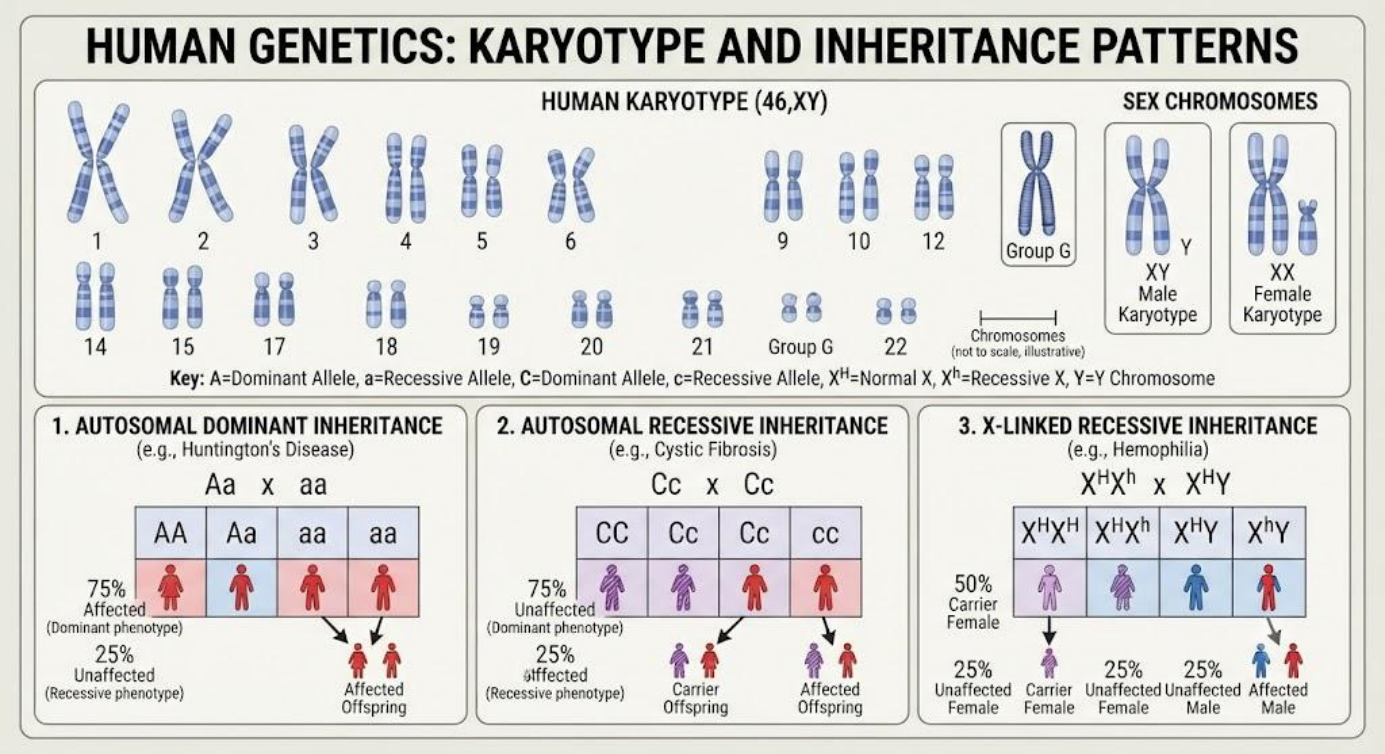


Figure 1.1: Human karyotype and inheritance patterns

Pilot questions 1.1

1. State the number of chromosomes in the normal human karyotype and how they are arranged.
2. Name three characteristics of autosomal dominant inheritance.
3. Name three characteristics of autosomal recessive inheritance.
4. Explain why X-linked recessive conditions are more common in males than females.
5. Give two examples each of autosomal dominant, autosomal recessive, and X-linked recessive conditions.

1.2 Pedigree Analysis

1.2.1 Pedigree symbols and conventions

Standard pedigree symbols allow family inheritance patterns to be represented diagrammatically across multiple generations; the conventions are: squares represent males, circles represent

females, a diamond or unspecified shape is used when sex is unknown; filled (shaded) symbols represent affected individuals, unfilled symbols represent unaffected individuals, and half-filled symbols may indicate carriers in some conventions; horizontal lines connect mating pairs, vertical lines connect parents to offspring, and a horizontal sibship line connects siblings.

Additional conventions include: a double horizontal line between mating pairs indicating a consanguineous union (between related individuals), diagonal lines through a symbol indicating a deceased individual, a dot within a circle or small filled dot indicating an obligate or confirmed carrier female, numbers within symbols indicating multiple individuals of the same type in the same generation, and Roman numerals labelling successive generations (I for the oldest generation shown, II for the next, and so on) with Arabic numerals identifying individuals within each generation.

Fill in the blank: In pedigrees, _____ represent males and circles represent females, with filled symbols indicating affected individuals.

1.2.2 Identifying inheritance patterns from pedigrees

Systematic pedigree interpretation follows a logical sequence of pattern recognition: if the trait appears in every generation without skipping, autosomal dominant is most likely; if the trait skips generations with unaffected carrier parents, autosomal recessive is most likely; if the trait is more common in males and transmitted through unaffected females, X-linked recessive is most likely; if the trait appears in females and males but more severely in males, X-linked dominant should be considered; if the trait passes exclusively from fathers to all sons, Y-linkage is suggested.

Key clues in pedigrees include: the sex ratio of affected individuals (equal for autosomal, male-biased for X-linked recessive), the presence of affected individuals with two unaffected parents (obligate carriers, diagnostic of recessive inheritance), consanguineous matings (increasing the probability of homozygosity for rare recessive alleles), and the proportion of affected offspring from different cross types, compared against expected Mendelian ratios.

Fill in the blank: If a trait appears in every generation without skipping, _____ dominant inheritance is most likely; if it skips generations with unaffected carrier parents, recessive is most likely.

1.2.3 Worked pedigree analysis examples

A three-generation pedigree showing a trait in Generation I father, Generation II son and daughter, and Generation III grandson (son of the II son) with no unaffected carrier parents at any point is consistent with autosomal dominant inheritance (trait in every generation, both sexes affected, approximately half offspring affected); a pedigree showing two unaffected parents (Generation II) with affected children of both sexes is consistent with autosomal recessive inheritance, making the parents obligate carriers.

A pedigree showing unaffected females in Generation I and II whose brothers and sons are affected, while daughters and granddaughters who are unaffected may be carriers, is consistent with X-linked recessive inheritance; the key diagnostic feature is that affected males receive the allele from their carrier mothers, and affected males cannot transmit the allele to their sons; working through a specific pedigree involves determining the genotype of each individual as far as possible and calculating the probability of different genotypes in individuals of interest.

Fill in the blank: In an X-linked recessive pedigree, affected males receive the allele from their _____ mothers, and affected males cannot transmit the allele to their sons.

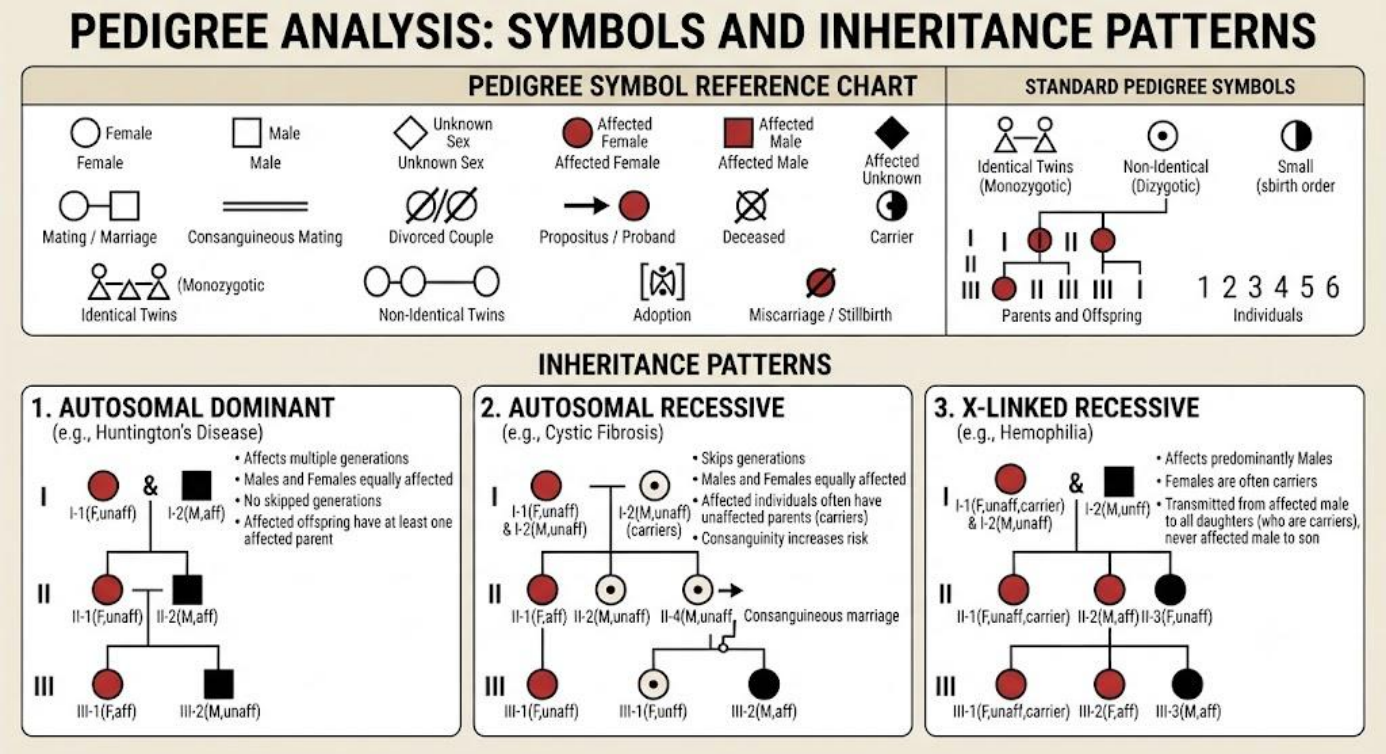


Figure 1.2: Pedigree symbols and analysis

Pilot questions 1.2

1. State what shapes represent males and females in a pedigree.
2. State what filled versus unfilled symbols indicate in a pedigree.
3. List three pattern features suggesting autosomal dominant inheritance.
4. List three pattern features suggesting X-linked recessive inheritance.
5. Explain what an obligate carrier is in pedigree analysis.

1.3 Chromosomal Abnormalities

1.3.1 Numerical chromosomal abnormalities

Numerical chromosomal abnormalities (aneuploidies) arise from errors in chromosome segregation during meiosis (non-disjunction), resulting in gametes with too many or too few chromosomes; monosomy (2n-1, one chromosome missing from a pair) and trisomy (2n+1, one

extra chromosome present) are the most common forms; trisomy 21 (Down syndrome) is the most frequent autosomal trisomy compatible with survival, characterised by intellectual disability, characteristic facial features, heart defects, and increased risk of leukaemia, with incidence increasing with maternal age.

Trisomy 18 (Edwards syndrome) and trisomy 13 (Patau syndrome) are also viable at birth but associated with severe abnormalities and typically short survival; polyploidy (having more than two complete chromosome sets) is lethal in humans at most stages but is naturally occurring and often beneficial in plants; the mechanism of non-disjunction (failure of homologous chromosomes or sister chromatids to separate at meiosis I or II respectively) produces eggs or sperm with abnormal chromosome numbers, and when fertilised by a normal gamete, produces an aneuploid zygote.

Fill in the blank: Trisomy 21, also called _____ syndrome, is the most frequent autosomal trisomy compatible with survival, with incidence increasing with maternal age.

1.3.2 Structural chromosomal abnormalities

Structural chromosomal abnormalities arise from chromosome breakage followed by abnormal repair or rejoining: deletions (loss of a chromosomal segment, causing haploinsufficiency for the genes lost), duplications (extra copy of a segment, causing gene dosage imbalance), inversions (a segment reversed in orientation within the chromosome, disrupting genes at the breakpoints and potentially affecting meiotic recombination in carriers), and translocations (movement of a segment from one chromosome to another, which may be balanced or unbalanced in terms of total gene content).

Balanced translocations (where no net loss or gain of genetic material occurs) typically produce a phenotypically normal carrier, since all gene sequences are present, though they carry a significant risk of producing unbalanced offspring; cri-du-chat syndrome (deletion of part of the short arm of chromosome 5, causing a distinctive cry in infants, intellectual disability, and other features) and DiGeorge syndrome (deletion at chromosome 22q11.2) are clinically significant examples of deletion syndromes.

Fill in the blank: A _____ translocation involves no net loss or gain of genetic material and typically produces a phenotypically normal carrier but a significant risk of unbalanced offspring.

1.3.3 Sex chromosome abnormalities

Sex chromosome aneuploidies are among the most common chromosomal abnormalities compatible with viability: Turner syndrome (45,X, a single X chromosome) presents in females with short stature, gonadal dysgenesis (streak ovaries), infertility, and characteristic physical features, with a surprisingly mild intellectual phenotype given monosomy; Klinefelter syndrome (47,XXY) presents in phenotypic males with small testes, infertility, reduced body hair, and often tall stature, with the extra X being Barr body-inactivated.

Other sex chromosome abnormalities include 47,XYY (typically normal phenotype with slightly increased stature), 47,XXX (triple X syndrome, typically normal or mildly affected phenotype in females), and mosaic conditions (where different cells in the same individual have different chromosomal constitutions, often producing milder phenotypes than non-mosaic aneuploidies); X-inactivation (the random inactivation of one X chromosome per cell in females, producing a Barr body visible in interphase nuclei) limits the pathological consequences of extra X chromosomes.

Fill in the blank: Turner syndrome is caused by the karyotype _____ , a single X chromosome, presenting as phenotypic females with short stature and gonadal dysgenesis.

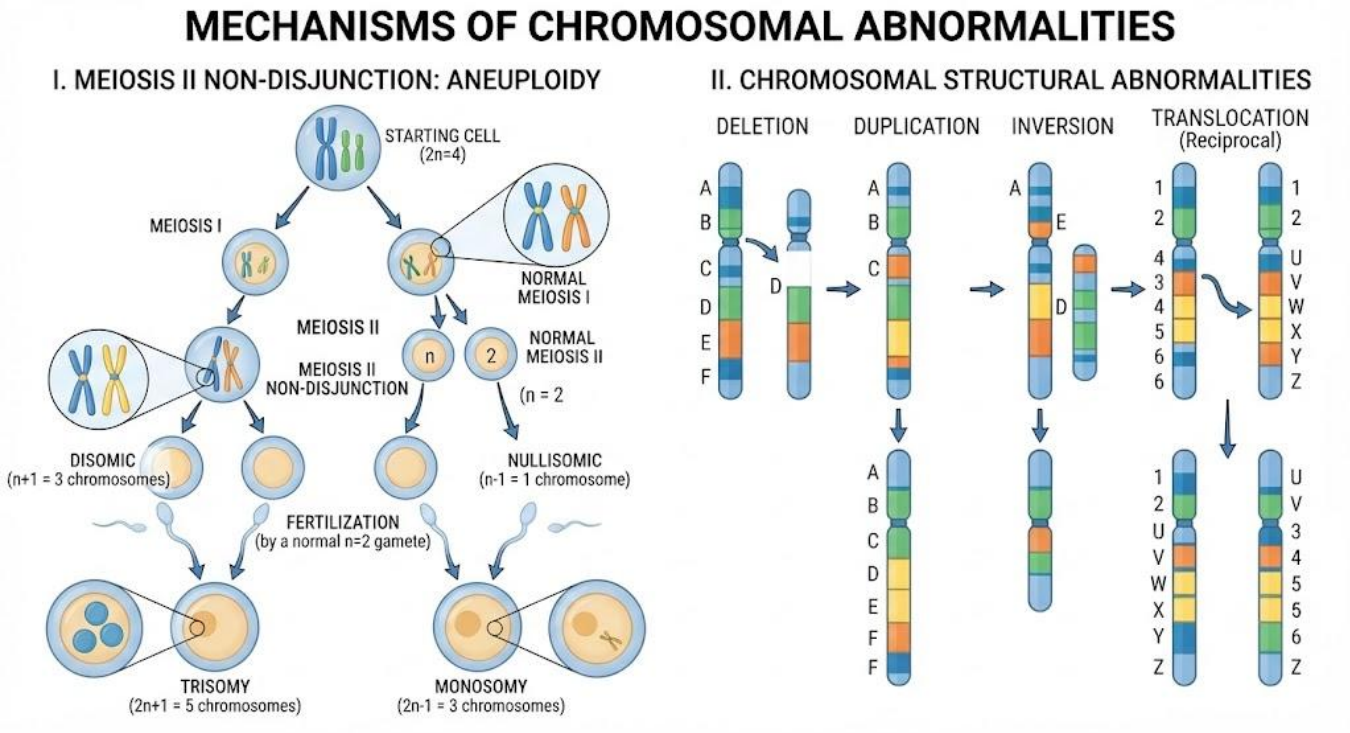


Figure 1.3: Chromosomal abnormalities

Pilot questions 1.3

1. Define aneuploidy and name the mechanism by which it arises.
2. Describe the clinical features of Down syndrome (trisomy 21).
3. Distinguish between a deletion and a duplication.
4. Explain why a balanced translocation carrier is typically phenotypically normal but at risk of abnormal offspring.
5. Distinguish between Turner syndrome and Klinefelter syndrome.

1.4 Genetic Counselling and Screening

1.4.1 Principles of genetic counselling

Genetic counselling is a communication process by which a trained genetic counsellor (or medical geneticist) helps individuals and families understand the nature of a genetic condition, the probability of its occurrence or recurrence in family members, the available options for

managing or preventing the condition, and the implications for family planning; it is a non-directive process, meaning the counsellor provides information and emotional support without directing the couple toward any particular decision.

Risk assessment in genetic counselling uses Mendelian probability calculations (for single-gene conditions), empirical risk figures (derived from population data for conditions with complex or unknown inheritance), pedigree analysis, and, increasingly, molecular genetic test results; risk is typically expressed as a fraction (such as one in four) or percentage, with the distinction between recurrence risk (risk in future pregnancies given an affected child has already occurred) and prior risk (risk based on family history before any affected birth) being important for accurate counselling.

Fill in the blank: Genetic counselling is a _____ -directive process, meaning the counsellor provides information without directing the family toward any particular decision.

1.4.2 Prenatal diagnosis techniques

Amniocentesis (sampling of amniotic fluid at approximately 15 to 18 weeks of pregnancy) and chorionic villus sampling (CVS, sampling of placental tissue at approximately 10 to 13 weeks) provide fetal cells or DNA for chromosomal analysis (karyotyping), biochemical testing, and molecular genetic analysis; amniocentesis has a slightly lower procedural miscarriage risk than CVS, but CVS has the advantage of earlier testing allowing earlier decision-making; both are invasive procedures with a small but real risk to the pregnancy.

Non-invasive prenatal testing (NIPT), analysing cell-free fetal DNA circulating in maternal blood, allows screening for common aneuploidies (particularly trisomies 21, 18, and 13 and sex chromosome anomalies) from as early as 10 weeks of pregnancy without procedural risk to the fetus; ultrasound examination detects structural fetal abnormalities, and the combination of ultrasound markers (nuchal translucency) with maternal blood biochemical markers provides a probabilistic screening result that guides the decision whether to proceed to invasive diagnostic testing.

Fill in the blank: _____ villus sampling (CVS), performed at approximately 10 to 13 weeks, allows earlier prenatal diagnosis than amniocentesis but carries a slightly higher procedural risk.

1.4.3 Genetic screening programmes

Population-based genetic screening differs from diagnostic testing in that it is offered to individuals without prior family history of a condition, aimed at identifying carriers or affected individuals who are not yet aware of their genetic status; newborn screening (testing all newborns for a panel of treatable conditions, such as phenylketonuria and congenital hypothyroidism, where early treatment prevents or reduces disability) is the most established form of population genetic screening and has been systematically implemented in many countries.

Carrier screening (identifying heterozygous carriers of autosomal recessive or X-linked conditions) may be offered to specific ethnic groups with increased carrier frequency (such as sickle cell carrier screening in populations of sub-Saharan African descent, or Tay-Sachs screening in Ashkenazi Jewish populations) or more broadly as pan-ethnic expanded carrier screening; preimplantation genetic testing (PGT), performed on embryos created through in vitro fertilisation before implantation, allows selection of unaffected embryos for transfer in couples at high genetic risk.

Fill in the blank: Newborn screening identifies _____ genetic conditions where early intervention prevents or reduces disability, and is offered to all newborns regardless of family history.

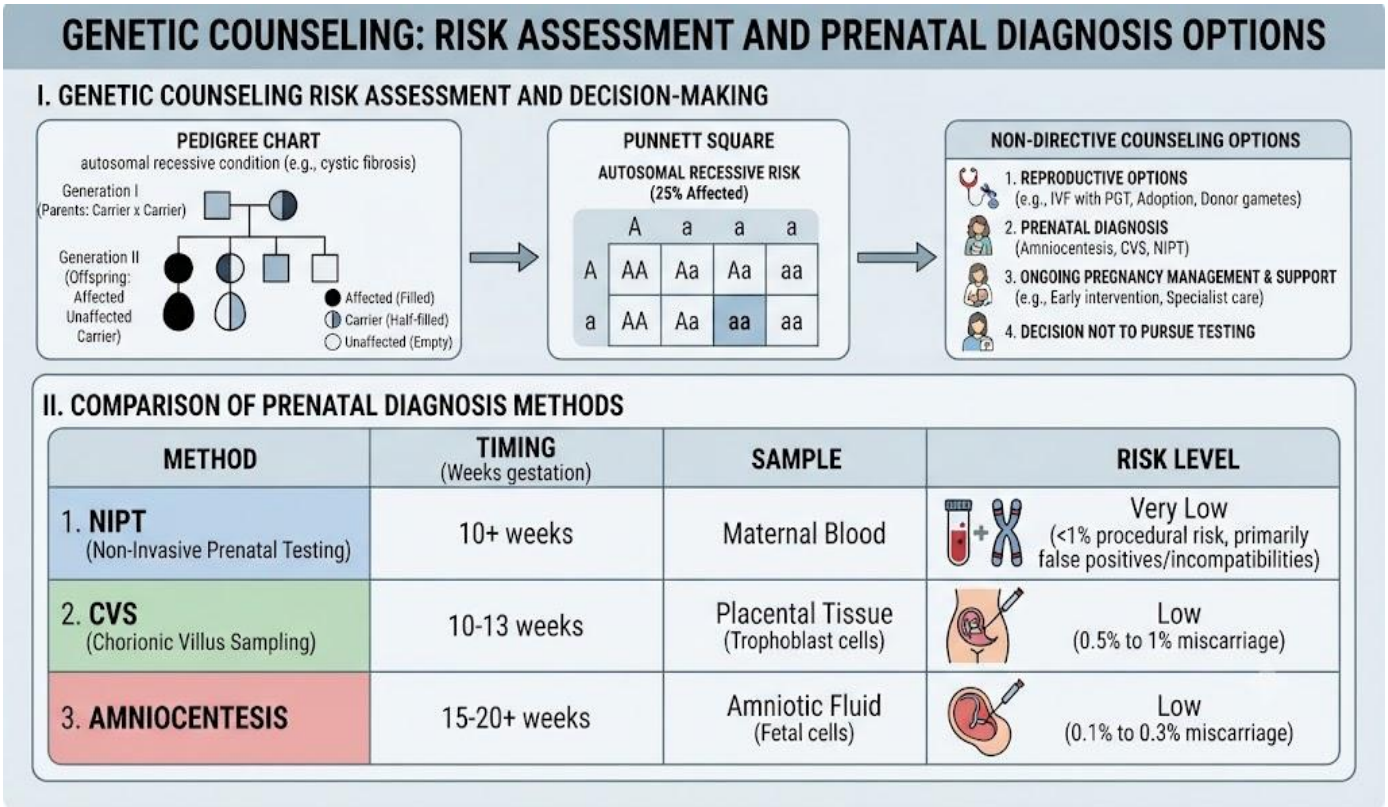


Figure 1.4: Genetic counselling and prenatal diagnosis

Pilot questions 1.4

1. Define genetic counselling and state its non-directive nature.
2. Explain how risk is expressed in genetic counselling.
3. Distinguish between amniocentesis and chorionic villus sampling.
4. Explain the advantage of NIPT over invasive prenatal testing.
5. Distinguish between genetic screening and genetic diagnostic testing.

Series Summary

This Series introduced the **principles of human genetics**, focusing on patterns of inheritance, pedigree analysis, and chromosomal abnormalities. You explored the normal human karyotype and the characteristics of autosomal dominant, autosomal recessive, and X-linked recessive inheritance, while learning how pedigree charts are used to identify inheritance patterns, recognise carriers, and estimate genetic risk.

The Series also examined numerical and structural chromosomal abnormalities, their causes and clinical examples, before introducing the principles of genetic counselling and modern genetic screening methods, including prenatal and population-based screening. By completing this Series, you should now be able to interpret common patterns of human inheritance, analyse pedigrees, explain the causes and consequences of chromosomal abnormalities, and evaluate the role of genetic counselling and screening in healthcare, thereby achieving all four **Series Learning Outcomes (SLOs)** and the **overall course outcomes**.

Glossary of Terms

- **Aneuploidy:** an abnormal chromosome number resulting from failure of chromosomes to segregate correctly during meiosis.
- **Autosomal dominant:** an inheritance pattern where one copy of a dominant allele on an autosome produces the trait in every generation.
- **Carrier:** a heterozygous individual who carries one copy of a recessive allele and is phenotypically unaffected.
- **Hemizygous:** having only one copy of a gene, as in males for X-linked genes.
- **Non-disjunction:** failure of homologous chromosomes or sister chromatids to separate during meiosis, producing aneuploid gametes.
- **Obligate carrier:** an individual who must be a carrier of a recessive allele based on their position in a pedigree, regardless of their own phenotype.

- **Pedigree:** a diagrammatic family tree showing the presence or absence of a genetic trait across generations.

Concept Check Questions [Series 1]

Objective Questions

1. The normal human karyotype consists of _____ chromosomes arranged as autosomes and sex chromosomes. (Linked to SLO 1.1)
2. In pedigrees, _____ represent males and circles represent females. (Linked to SLO 1.2)
3. Trisomy 21 is also called _____ syndrome, the most frequent autosomal trisomy compatible with survival. (Linked to SLO 1.3)
4. Genetic counselling is a non-_____ process in which the counsellor provides information without directing the family. (Linked to SLO 1.4)
5. X-linked recessive conditions are more common in _____ males because they have only one X chromosome. (Linked to SLO 1.1)

Short-Answer Questions

1. Name three characteristics of autosomal recessive inheritance. (Linked to SLO 1.1)
2. State three pedigree features suggesting X-linked recessive inheritance. (Linked to SLO 1.2)
3. Distinguish between Turner syndrome and Klinefelter syndrome. (Linked to SLO 1.3)
4. Distinguish between amniocentesis and CVS. (Linked to SLO 1.4)
5. Define obligate carrier. (Linked to SLO 1.2)

Long-Answer Questions

1. Describe autosomal dominant, autosomal recessive, and X-linked recessive inheritance with examples. (Linked to SLO 1.1)
2. Explain pedigree symbols and how to identify inheritance patterns from pedigrees. (Linked to SLO 1.2)
3. Describe chromosomal abnormalities, genetic counselling, and prenatal diagnosis. (Linked to SLOs 1.3, 1.4)

Answers to Concept Check Questions [Series 1]

Objective Questions

1. 46.
2. Squares.
3. Down.
4. Directive.
5. Hemizygous.

Short-Answer Questions

1. Skips generations, both parents of affected individuals are typically unaffected carriers, and approximately one quarter of offspring from two carriers are affected.
2. More common in males, transmitted through unaffected carrier females, and affected males do not transmit the trait to their sons.
3. Turner syndrome is 45,X (female, short stature, infertility); Klinefelter syndrome is 47,XXY (male, small testes, infertility).
4. Amniocentesis samples amniotic fluid at 15-18 weeks; CVS samples placental tissue at 10-13 weeks, allowing earlier diagnosis.
5. An obligate carrier is an individual who must be a heterozygous carrier based on pedigree position, regardless of phenotype.

Long-Answer Questions

1. Autosomal dominant: every generation, both sexes, approximately half offspring affected; examples Huntington disease, polydactyly. Autosomal recessive: skips generations, carriers unaffected, one in four offspring affected; examples cystic fibrosis, sickle cell disease. X-linked recessive: males affected more, carrier females, examples haemophilia, colour blindness.
2. Pedigree symbols: squares (males), circles (females), filled (affected), half-filled (carrier), double line (consanguineous). Pattern recognition: every generation suggests dominant; skipping suggests recessive; male-biased with carrier females suggests X-linked recessive. Obligate carriers identified from pedigree position.

3. Numerical abnormalities arise from non-disjunction producing trisomies (Down, Edwards, Patau) and monosomies (Turner). Structural abnormalities include deletions, duplications, inversions, and translocations. Genetic counselling is non-directive risk communication. Prenatal diagnosis uses amniocentesis (15-18 weeks), CVS (10-13 weeks), or non-invasive NIPT (from 10 weeks).

Answers to Pilot Questions [Series 1]

Part 1.1

1. 46 chromosomes: 22 autosome pairs (numbered 1-22) and one sex chromosome pair (XX female, XY male).
2. Trait in every generation, at least one affected parent, approximately half offspring affected.
3. Skips generations, both parents of affected individuals typically unaffected, approximately one quarter of offspring from two carriers affected.
4. Males have only one X chromosome (hemizygous), so a single recessive allele is sufficient to produce the phenotype.
5. Autosomal dominant: Huntington disease, polydactyly. Autosomal recessive: cystic fibrosis, sickle cell disease. X-linked recessive: haemophilia A, red-green colour blindness.

Part 1.2

1. Squares represent males, circles represent females.
2. Filled symbols indicate affected individuals; unfilled symbols indicate unaffected individuals.
3. Trait in every generation, both sexes affected, approximately half offspring affected.
4. More common in males, transmitted through unaffected carrier females, affected males do not transmit to sons.
5. An obligate carrier is an individual who must be a heterozygous carrier based on their pedigree position, regardless of their own phenotype.

Part 1.3

1. Aneuploidy is an abnormal chromosome number; it arises from non-disjunction, the failure of chromosomes to separate during meiosis.

2. Intellectual disability, characteristic facial features, heart defects, increased risk of leukaemia, and incidence increasing with maternal age.
3. A deletion removes a chromosomal segment causing haploinsufficiency; a duplication creates an extra copy of a segment causing gene dosage imbalance.
4. A balanced translocation has no net gene loss/gain so the carrier is normal, but during meiosis unbalanced gametes can be produced leading to abnormal offspring.
5. Turner (45,X): phenotypic female, short stature, gonadal dysgenesis, infertility. Klinefelter (47,XXY): phenotypic male, small testes, infertility, reduced body hair.

Part 1.4

1. Genetic counselling helps individuals/families understand a genetic condition, its probability of recurrence, and available options; it is non-directive, providing information without directing decisions.
2. Risk is expressed as a fraction (e.g. 1 in 4) or percentage, distinguishing recurrence risk from prior risk.
3. Amniocentesis samples amniotic fluid at 15-18 weeks; CVS samples placental/chorionic villus tissue at 10-13 weeks, allowing earlier results.
4. NIPT analyses cell-free fetal DNA in maternal blood from 10 weeks without any procedural risk to the fetus, unlike the small but real miscarriage risk of amniocentesis and CVS.
5. Genetic screening is offered to individuals without prior family history to identify carriers or affected individuals; diagnostic testing is for individuals already suspected to be at risk based on family history or symptoms.

References and Attribution

Textual References (Books and External Sources)

- Strachan, T., & Read, A. (2018). Human molecular genetics (5th ed.). Garland Science.
- Brooker, R. J. (2015). Genetics: Analysis and principles (5th ed.). McGraw-Hill.
- Nussbaum, R. L., McInnes, R. R., & Willard, H. F. (2015). Thompson and Thompson genetics in medicine (8th ed.). Elsevier.
- National Open University of Nigeria (NOUN). (2023). Genetics II course material. NOUN Press.

Figures, Plates, Tables, and Boxes

- Figure 1.1: Human karyotype and inheritance patterns Attribution: AI generated. Suggested OER search string: "human karyotype chromosomes autosomal dominant recessive X-linked inheritance Punnett square genetics OER".
- Figure 1.2: Pedigree symbols and analysis Attribution: AI generated. Suggested OER search string: "pedigree symbols conventions autosomal dominant recessive X-linked inheritance analysis diagram genetics OER".
- Figure 1.3: Chromosomal abnormalities Attribution: AI generated. Suggested OER search string: "chromosomal abnormalities non-disjunction trisomy monosomy deletion duplication inversion translocation diagram genetics OER".
- Figure 1.4: Genetic counselling and prenatal diagnosis Attribution: AI generated. Suggested OER search string: "genetic counselling prenatal diagnosis amniocentesis chorionic villus sampling NIPT screening genetics OER".

Course code: BIO 301

Course title: Genetics II

Course credit load: 2 Units

Series 2: Gene Interactions and Biochemical Mutants

- **Part 2.1: Deviations from Basic Mendelian Ratios**
  - Sub Part 2.1.1: Incomplete dominance and codominance
  - Sub Part 2.1.2: Multiple alleles
  - Sub Part 2.1.3: Lethal alleles
- **Part 2.2: Gene Interactions (Epistasis)**
  - Sub Part 2.2.1: Principles of epistasis
  - Sub Part 2.2.2: Types of epistasis and their phenotypic ratios
  - Sub Part 2.2.3: Complementary and duplicate gene action
- **Part 2.3: Polygenic Inheritance and Quantitative Traits**
  - Sub Part 2.3.1: Polygenic inheritance
  - Sub Part 2.3.2: Heritability
  - Sub Part 2.3.3: Gene-environment interactions
- **Part 2.4: Biochemical Mutants and Metabolic Pathways**
  - Sub Part 2.4.1: The one gene one enzyme hypothesis
  - Sub Part 2.4.2: Biochemical mutants and metabolic blocks
  - Sub Part 2.4.3: Inborn errors of metabolism

Introduction

Real inheritance patterns frequently deviate from the simple Mendelian ratios of 3:1 or 9:3:3:1, not because the laws of segregation and independent assortment are violated, but because the relationships between alleles at a single locus and between genes at different loci are more complex than simple dominance at one gene. This Series examines these deviations, from incomplete dominance and codominance through multiple alleles and epistasis, and then links gene interaction concepts to the biochemical understanding of how genes control metabolic pathways.

This Series covers deviations from basic Mendelian ratios; gene interactions including epistasis and complementary gene action; polygenic inheritance and quantitative traits; and the connection between genes and biochemical pathways, including the one gene one enzyme hypothesis, biochemical mutants, and inborn errors of metabolism.

Series Learning Outcomes

Upon completing this Series, you should be able to:

- **SLO 2.1** explain incomplete dominance, codominance, multiple alleles, and lethal alleles as deviations from basic Mendelian ratios,
- **SLO 2.2** explain epistasis and gene interaction, including complementary and duplicate gene action and their phenotypic ratios,
- **SLO 2.3** describe polygenic inheritance, heritability, and gene-environment interactions, and
- **SLO 2.4** explain the one gene one enzyme hypothesis, biochemical mutants, and inborn errors of metabolism.

2.1 Deviations from Basic Mendelian Ratios

2.1.1 Incomplete dominance and codominance

Incomplete dominance occurs when neither allele at a locus is fully dominant over the other, and the heterozygote displays a phenotype intermediate between the two homozygous phenotypes; the classic example is flower colour in snapdragons (Antirrhinum majus), where a cross between red-flowered (RR) and white-flowered (rr) plants produces pink-flowered (Rr) F1 plants, and the F2 generation produces a 1 red : 2 pink : 1 white ratio rather than the 3:1 expected for simple dominance; the pink colour reflects the reduced amount of red pigment produced by a single R allele compared to two.

Codominance differs from incomplete dominance in that both alleles are simultaneously and independently expressed in the heterozygote, rather than producing a blend; the ABO blood group system provides a human example of codominance between the IA and IB alleles, which independently specify different surface antigens on red blood cells, so that an IAIB individual has both A and B antigens and is blood type AB; the molecular basis of codominance is the simultaneous production of the products of both alleles in the same cell, without interference between them.

Fill in the blank: In _____ dominance, the heterozygote displays a phenotype intermediate between both homozygotes, such as pink flowers from a cross between red and white.

2.1.2 Multiple alleles

Although a diploid individual can carry at most two alleles at any given locus (one per homologue), a gene can exist in more than two allelic forms within a population; the ABO blood group system exemplifies multiple allelism, with three main alleles (IA, IB, and i) producing four blood group phenotypes (A, B, AB, and O) depending on the genotype; IA and IB are codominant with each other, and both are dominant over the recessive i allele, producing the observed pattern.

The coat colour gene in rabbits is another widely studied example of multiple allelism, with four alleles producing a dominance series: C (full colour, dominant) > cch (chinchilla, intermediate) > ch (Himalayan, light body with dark extremities) > c (albino, recessive); each allele encodes a version of the tyrosinase enzyme with different activity, from full activity (C) to zero activity (c), illustrating that multiple alleles often represent a series of mutations in the same gene producing a range of enzyme activities.

Fill in the blank: In the ABO blood group system, IA and IB alleles are _____ with each other, while both are dominant over the recessive i allele.

2.1.3 Lethal alleles

Some alleles are lethal when homozygous, causing death at some stage of development before the trait can be observed or counted; the yellow coat colour allele (AY) in mice is a well-studied example: AY/A mice have yellow coats (AY being dominant for coat colour), but AY/AY mice die as early embryos because the AY allele also disrupts expression of a gene essential for embryonic development, so yellow-coated mice are always heterozygous and a cross between two yellow mice gives a 2 yellow : 1 agouti ratio instead of the expected 3:1.

Lethal alleles illustrate the principle that a single gene can affect multiple phenotypic traits (pleiotropy), since the AY allele simultaneously affects coat colour and embryonic viability; recessive lethal alleles (only lethal when homozygous) are common in populations and are maintained by natural selection at low frequencies since heterozygous carriers are typically viable and may sometimes have a fitness advantage; dominant lethal alleles (lethal even in heterozygotes) are generally eliminated from populations within one generation unless they act after reproductive age, as in Huntington disease.

Fill in the blank: Lethal alleles illustrate the principle of _____ , where a single gene affects multiple phenotypic traits simultaneously, such as coat colour and embryonic viability.

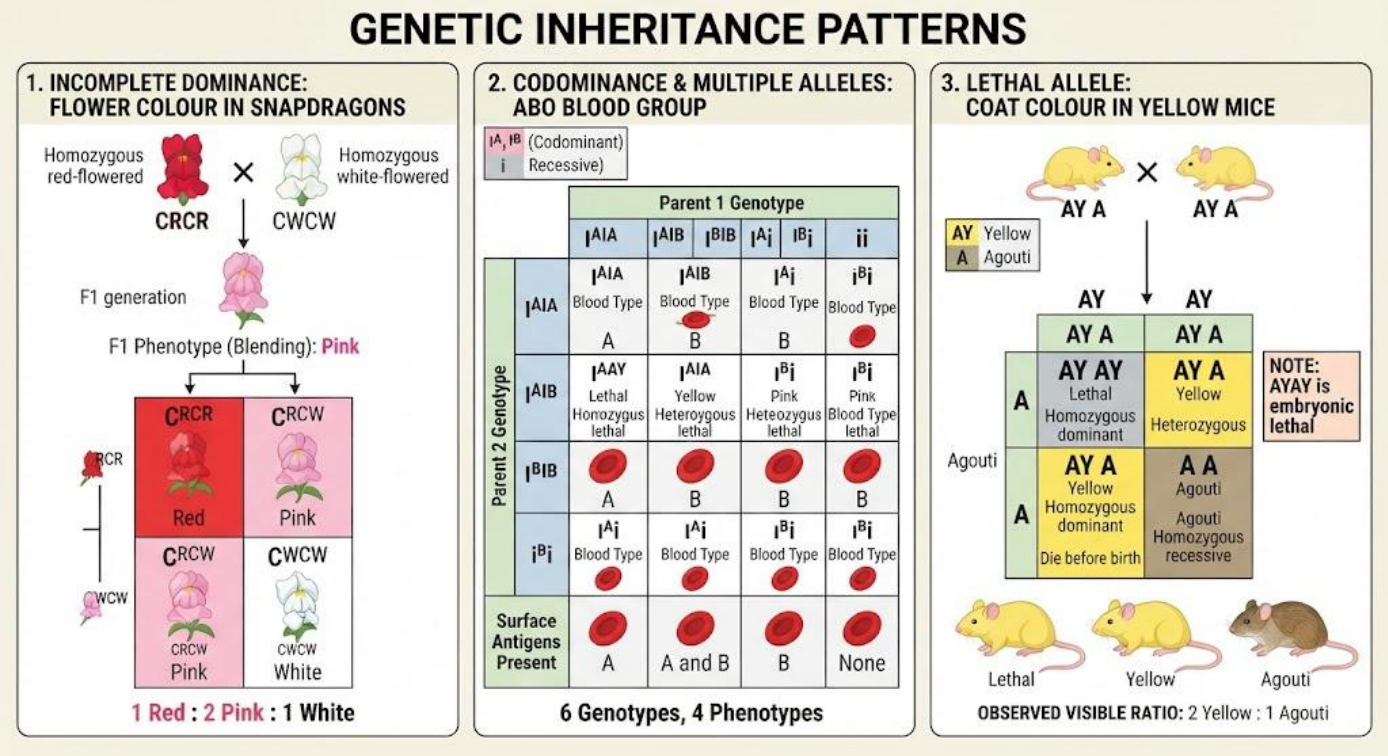


Figure 2.1: Deviations from Mendelian ratios

Pilot questions 2.1

1. Distinguish between incomplete dominance and codominance.
2. A cross between a red (RR) and white (rr) snapdragon gives pink F1. What F2 ratio is expected?
3. Define multiple allelism and give one example.
4. State the four possible ABO blood group phenotypes and the genotypes that produce each.
5. Explain why a cross between two yellow mice produces a 2:1 ratio of yellow to agouti offspring.

2.2 Gene Interactions (Epistasis)

2.2.1 Principles of epistasis

Epistasis occurs when the alleles at one gene locus mask or modify the expression of alleles at a different gene locus; the masking gene is said to be epistatic to the masked gene, and the masked gene is said to be hypostatic; epistasis differs from dominance (which describes allelic interactions at the same locus) in that it describes interactions between genes at different loci and typically alters the expected 9:3:3:1 dihybrid phenotypic ratio into a variety of modified ratios depending on the specific interaction.

The biochemical basis of epistasis is most clearly understood when two genes encode enzymes in the same metabolic pathway: if gene A encodes an enzyme required to produce a substrate for the enzyme encoded by gene B, then a homozygous recessive mutation at gene A will prevent the pathway from operating at all, regardless of the genotype at gene B, because the substrate for the gene B enzyme will be absent; this upstream epistasis results in the gene A alleles masking the expression of gene B alleles.

Fill in the blank: Epistasis differs from dominance in that it describes interactions between alleles at _____ gene loci rather than at the same locus.

2.2.2 Types of epistasis and their phenotypic ratios

Recessive epistasis (12:3:1 ratio): homozygous recessive at gene A (aa) masks expression of gene B, producing an epistatic phenotype regardless of B genotype; example: coat colour in Labrador retrievers (aa masks B_ to produce yellow regardless of chocolate or black B genotype). Dominant epistasis (12:3:1 or 13:3 ratio): a single dominant allele at gene A masks expression of gene B; example: summer squash fruit colour (W_ is white regardless of Y genotype). Recessive epistasis of gene B over gene A (9:3:4 ratio, duplicate recessive epistasis): homozygous recessive at either locus produces the same epistatic phenotype.

Dominant suppression (13:3): a dominant allele at one locus suppresses the expression of a dominant allele at another locus; duplicate dominant epistasis (15:1): a dominant allele at either

of two loci produces the same phenotype, so only the double homozygous recessive (aabb) displays the alternative phenotype; each modified ratio reflects a specific pattern of how the two gene products interact in producing the final phenotype, and identifying which modified ratio is observed provides information about the biochemical relationships between the two genes.

Fill in the blank: In recessive epistasis, homozygous recessive at the epistatic locus masks expression of the hypostatic locus, producing a modified _____ :3:1 phenotypic ratio from a dihybrid cross.

2.2.3 Complementary and duplicate gene action

Complementary gene action (9:7 ratio) occurs when dominant alleles at two different loci are both required to produce a trait, so that homozygous recessive at either locus (or both) produces an alternative phenotype; the classic example is flower colour in sweet peas (Lathyrus odoratus), where two independently assorting genes (C and P) each encode an enzyme in the purple pigment pathway, and purple colour requires at least one dominant allele at both loci; ccPP, CCpp, and ccpp plants are all white, and a cross between two different white-flowered strains can produce all purple F1 (a phenomenon called complementation).

Duplicate gene action (15:1 ratio) occurs when dominant alleles at either of two loci produce the same phenotype, with only the doubly homozygous recessive (aabb) displaying the alternative phenotype; this is expected when two genes have redundant functions in the same pathway; the complementation test (crossing two mutant strains that appear to have the same phenotype: if F1 is wild type, the mutations complement and are in different genes; if F1 is mutant, the mutations fail to complement and are in the same gene) is a fundamental genetic tool for determining whether two mutations are in the same or different genes.

Fill in the blank: Complementary gene action produces a _____ :7 ratio from a dihybrid cross, because dominant alleles at both loci are required to produce the trait.

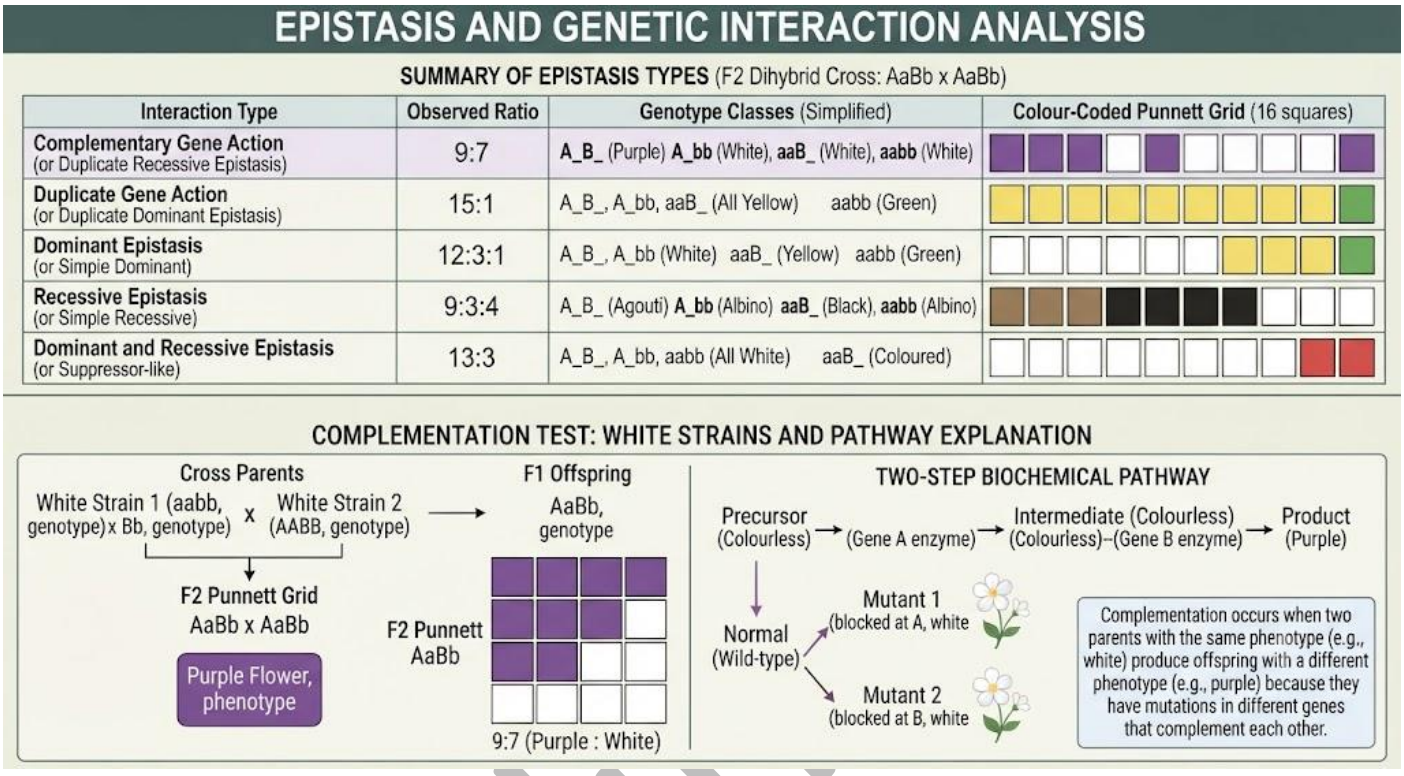


Figure 2.2: Epistasis and gene interactions

Pilot questions 2.2

1. Define epistasis and distinguish it from dominance.
2. Explain the biochemical basis of upstream recessive epistasis.
3. State the phenotypic ratio expected from complementary gene action and explain why.
4. Describe the complementation test and its purpose.
5. A dihybrid cross gives a 15:1 ratio. What type of gene interaction does this indicate?

2.3 Polygenic Inheritance and Quantitative Traits

2.3.1 Polygenic inheritance

Polygenic inheritance refers to traits whose phenotype is determined by the additive contributions of alleles at two or more gene loci, with each contributing locus typically having a small, equal, and additive effect on the phenotype; such traits show a continuous, normally

distributed range of phenotypic values in a population (rather than discrete Mendelian classes) because of the many possible genotypic combinations and environmental modifiers; examples include human height, skin pigmentation, body weight, and intelligence.

A simple two-locus model illustrates the principle: if alleles A and B each add one unit to a base phenotypic value and alleles a and b add zero, then an AaBb dihybrid cross produces nine genotypic classes (with anywhere from zero to four contributing alleles) and five phenotypic classes in a 1:4:6:4:1 ratio; as the number of contributing loci increases, the number of classes increases and the distribution approximates a normal curve more closely, matching the continuous variation observed for most polygenic traits.

Fill in the blank: Polygenic inheritance produces a _____ distribution of phenotypic values in a population because of the many possible genotypic combinations and environmental contributions.

2.3.2 Heritability

Heritability (H2 in the broad sense, h2 in the narrow sense) is a statistical measure of the proportion of total phenotypic variance in a population that is attributable to genetic variance; H2 = VG/VP, where VG is genetic variance and VP is total phenotypic variance (which includes both genetic and environmental components); heritability ranges from 0 (all phenotypic variation is environmental) to 1 (all phenotypic variation is genetic), and high heritability does not mean that a trait is unaffected by the environment, only that most of the variation within the measured population, under the measured conditions, reflects genetic differences.

Heritability estimates are population- and environment-specific and should not be extrapolated beyond the population and conditions in which they were measured; narrow-sense heritability (h2 = VA/VP, where VA is additive genetic variance only) is particularly important for predicting the response to selection in plant and animal breeding programmes; twin studies (comparing trait similarity in monozygotic/identical twins, who are genetically identical, to dizygotic/fraternal twins, who share approximately half their genes) are a major method for estimating heritability in human populations.

Fill in the blank: Heritability (H2) is the proportion of total phenotypic variance attributable to _____ variance, and ranges from 0 (all environmental) to 1 (all genetic).

2.3.3 Gene-environment interactions

The phenotype of an organism results from the interaction of its genotype with its environment; the norm of reaction describes the range of phenotypes that a given genotype can produce across a range of environmental conditions, and genotypes differ in their norms of reaction, meaning that the same environmental change can produce very different phenotypic responses in different genotypes; phenotypic plasticity (the capacity of a single genotype to produce different phenotypes in different environments) is an important adaptive property, particularly evident in plants.

Gene-environment interactions can complicate the interpretation of genetic experiments, particularly in humans where environmental conditions cannot be controlled; even highly heritable traits can be substantially modified by environmental intervention (phenylketonuria, highly heritable, is effectively managed by dietary phenylalanine restriction; height, highly heritable, shows substantial secular trends over generations due to improved nutrition), illustrating that heritability describes the extent of genetic determination of variation within current conditions but does not set a ceiling on the response to environmental change.

Fill in the blank: The norm of _____ describes the range of phenotypes that a given genotype can produce across a range of environmental conditions.

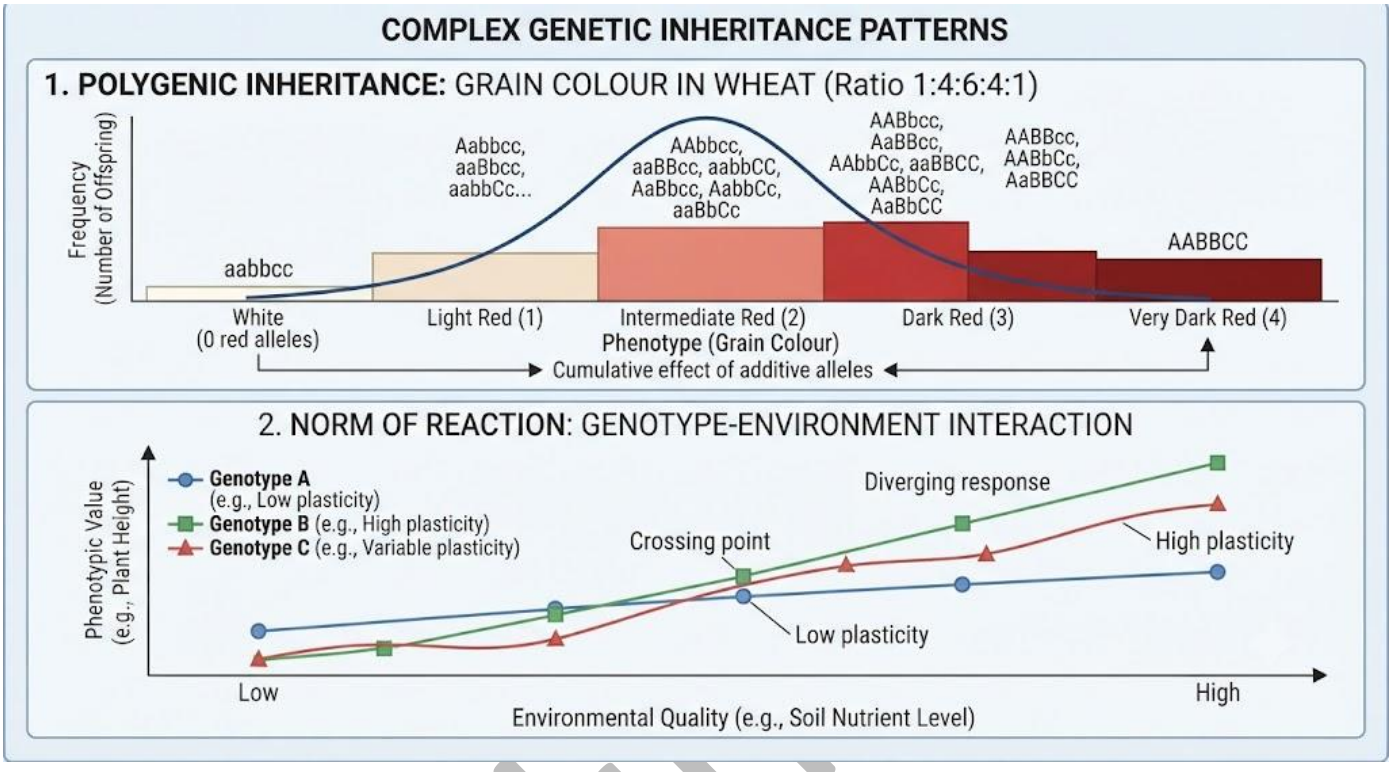


Figure 2.3: Polygenic inheritance and gene-environment interactions

Pilot questions 2.3

1. Define polygenic inheritance and explain why polygenic traits show continuous variation.
2. What phenotypic ratio is expected from a dihybrid cross for a trait controlled by two additive loci?
3. Define heritability (H2) and state the formula.
4. Explain why a high heritability does not mean a trait cannot be altered by environment.
5. Define norm of reaction.

2.4 Biochemical Mutants and Metabolic Pathways

2.4.1 The one gene one enzyme hypothesis

George Beadle and Edward Tatum proposed the one gene one enzyme hypothesis in 1941, based on experiments with the bread mould Neurospora crassa: they induced mutations by X-ray

irradiation and identified nutritional (auxotrophic) mutants unable to grow on minimal medium without specific nutrient supplements; by crossing mutants with wild-type and testing which nutritional supplements restored growth, they mapped each mutation to a specific biochemical step in a biosynthetic pathway, concluding that each gene controls the production of one enzyme.

The hypothesis has been refined to one gene one polypeptide (since some enzymes are multimeric, composed of polypeptide subunits encoded by different genes, and some genes encode structural or regulatory RNAs rather than polypeptides) but the core insight remains foundational: genes exert their effects on phenotype primarily through their protein (enzyme) products, and a mutation in a gene typically disrupts the function of the encoded protein; this biochemical gene-phenotype connection underpins the understanding of inborn errors of metabolism examined in Part 2.4.3.

Fill in the blank: The one gene one enzyme hypothesis, proposed by Beadle and Tatum using _____ crassa mutants, has been refined to one gene one polypeptide.

2.4.2 Biochemical mutants and metabolic blocks

A biochemical mutant (auxotroph) is an organism that has lost the ability to synthesise a required substance (such as an amino acid, vitamin, or nucleotide) due to a mutation in a gene encoding an enzyme in the biosynthetic pathway for that substance; such mutants require the substance (or a downstream product of the blocked step) to be supplied externally for growth, providing the nutritional supplementation test used by Beadle and Tatum to identify the blocked step.

The pattern of metabolic block in biochemical mutants provides information about pathway organisation: if supplying substance C (but not A or B) restores growth to a mutant, the block is between B and C in the pathway A to B to C; if supplying either B or C restores growth, the block is before B; this logic, called nutritional supplementation analysis, allows the order of steps in a metabolic pathway to be deduced from the growth responses of different mutants, complementing the direct biochemical analysis of pathway intermediates that accumulate upstream of the block.

Fill in the blank: An _____ mutant requires an externally supplied substance for growth because a mutation has blocked the biosynthetic pathway producing that substance.

2.4.3 Inborn errors of metabolism

Inborn errors of metabolism are human genetic conditions caused by mutations in genes encoding metabolic enzymes, resulting in the accumulation of toxic substrates upstream of the block and/or deficiency of essential products downstream; phenylketonuria (PKU) is caused by deficiency of phenylalanine hydroxylase, the enzyme converting phenylalanine to tyrosine, leading to accumulation of phenylalanine and its toxic metabolites causing intellectual disability if untreated; identified by newborn screening and managed by dietary phenylalanine restriction, allowing near-normal development.

Alkaptonuria (deficiency of homogentisic acid oxidase in the tyrosine degradation pathway, causing accumulation of homogentisic acid excreted in urine which turns black on exposure to air and deposited in connective tissue) was the first condition described as an inborn error of metabolism by Archibald Garrod in 1902, predating the one gene one enzyme hypothesis by nearly four decades; other examples include galactosaemia (galactose-1-phosphate uridyl transferase deficiency), maple syrup urine disease (branched-chain alpha-keto acid dehydrogenase deficiency), and the lysosomal storage disorders (accumulation of undegraded substrates in lysosomes).

Fill in the blank: Phenylketonuria is caused by deficiency of phenylalanine _____ , the enzyme converting phenylalanine to tyrosine, causing substrate accumulation and intellectual disability if untreated.

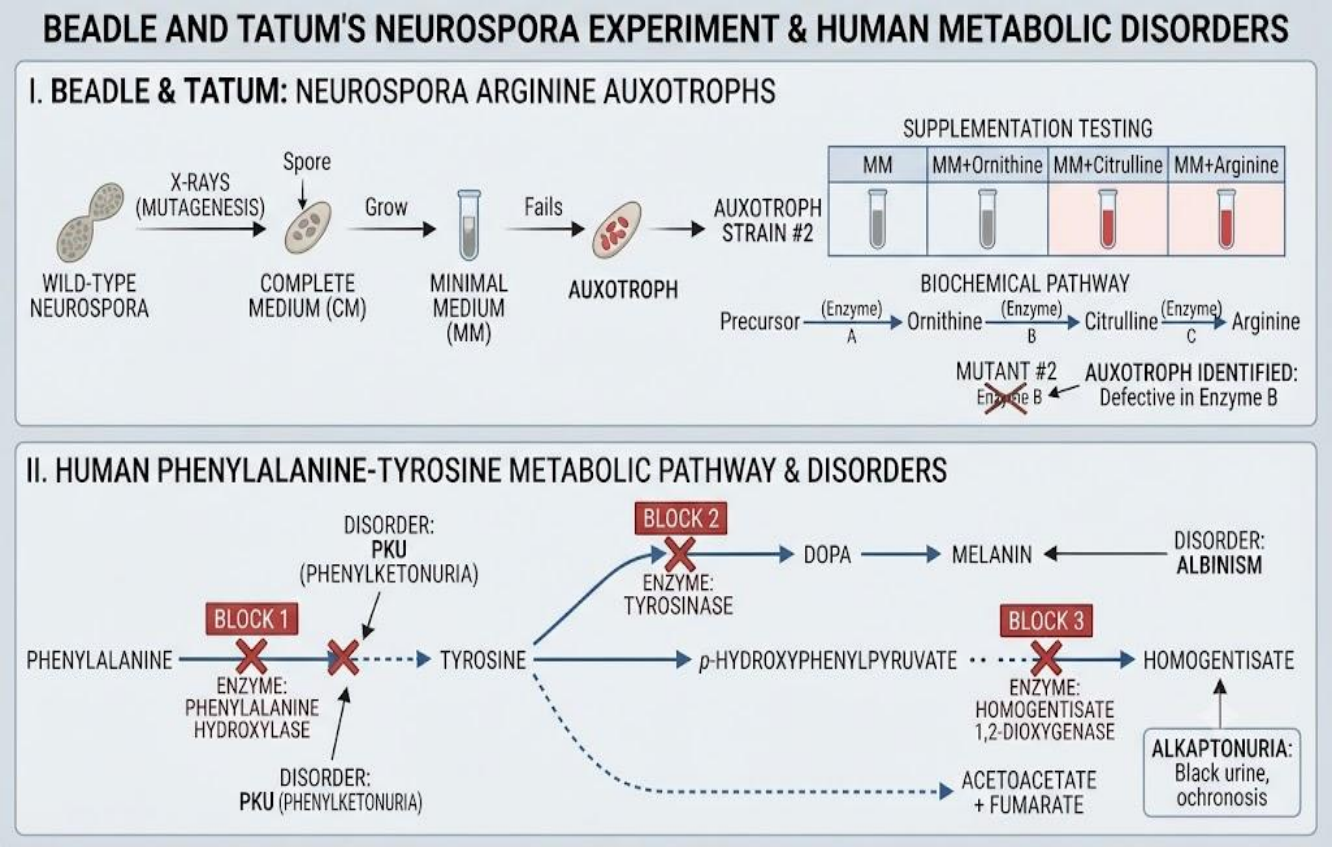


Figure 2.4: Biochemical mutants and metabolic pathways

Pilot questions 2.4

1. Describe the experimental approach Beadle and Tatum used to test the one gene one enzyme hypothesis.
2. Why has the hypothesis been refined to one gene one polypeptide?
3. Define auxotroph and explain how nutritional supplementation identifies the metabolic block.
4. Describe the metabolic basis of phenylketonuria.
5. What was the significance of Garrod description of alkaptonuria?

Series Summary

This Series explored **patterns of inheritance that extend beyond classical Mendelian genetics**, demonstrating how different genetic mechanisms produce modified inheritance ratios and phenotypes. You examined incomplete dominance, codominance, multiple alleles, lethal alleles, and various forms of gene interaction, including epistasis, complementary gene action, and complementation tests, to understand how genes interact to influence observable traits.

The Series also introduced the principles of polygenic inheritance and heritability, highlighting the combined effects of multiple genes and the environment on continuous traits. In addition, you explored the foundations of biochemical genetics through the one gene-one enzyme concept, metabolic pathways, and inherited metabolic disorders such as phenylketonuria, alkaptonuria, and galactosaemia. By completing this Series, you should now be able to explain non-Mendelian inheritance patterns, interpret gene interactions and quantitative traits, evaluate the influence of heredity and environment, and relate gene function to metabolic disorders, thereby achieving all four **Series Learning Outcomes (SLOs)** and the **overall course outcomes**.

Glossary of Terms

- **Auxotroph:** a mutant organism that requires an externally supplied nutrient because a biosynthetic enzyme is non-functional.
- **Complementation test:** a cross between two mutants to determine whether their mutations are in the same gene (fail to complement) or different genes (complement).
- **Epistasis:** gene interaction in which alleles at one locus mask or modify expression of alleles at a different locus.
- **Heritability:** the proportion of total phenotypic variance in a population attributable to genetic variance.
- **Inborn error of metabolism:** a genetic condition caused by a mutant enzyme in a metabolic pathway, causing substrate accumulation and/or product deficiency.

- **Norm of reaction:** the range of phenotypes a given genotype produces across different environmental conditions.
- **Polygenic inheritance:** inheritance of a trait determined by the additive contributions of alleles at two or more gene loci.

Concept Check Questions [Series 2]

Objective Questions

1. In incomplete dominance, the heterozygote displays a phenotype _____ between the two homozygous phenotypes. (Linked to SLO 2.1)
2. Epistasis differs from dominance in that it describes interactions between alleles at _____ gene loci. (Linked to SLO 2.2)
3. Polygenic traits show _____ variation in a population rather than discrete Mendelian classes. (Linked to SLO 2.3)
4. The one gene one enzyme hypothesis was tested using _____ crassa auxotrophic mutants. (Linked to SLO 2.4)
5. Complementary gene action produces a _____ :7 ratio from a dihybrid cross. (Linked to SLO 2.2)

Short-Answer Questions

1. Distinguish between incomplete dominance and codominance. (Linked to SLO 2.1)
2. Define epistasis and explain its biochemical basis. (Linked to SLO 2.2)
3. Define heritability and state the formula. (Linked to SLO 2.3)
4. Describe the metabolic basis of PKU. (Linked to SLO 2.4)
5. Explain the complementation test. (Linked to SLO 2.2)

Long-Answer Questions

1. Explain incomplete dominance, codominance, multiple alleles, and lethal alleles. (Linked to SLO 2.1)

2. Describe epistasis types, complementary gene action, and the complementation test. (Linked to SLO 2.2)
3. Explain polygenic inheritance, heritability, and the one gene one enzyme hypothesis. (Linked to SLOs 2.3, 2.4)

Answers to Concept Check Questions [Series 2]

Objective Questions

1. Intermediate.
2. Different.
3. Continuous.
4. Neurospora.
5. 9.

Short-Answer Questions

1. Incomplete dominance: heterozygote has a blended intermediate phenotype. Codominance: both alleles are simultaneously and independently expressed in the heterozygote.
2. Epistasis is gene interaction where alleles at one locus mask expression at another; biochemical basis is often sequential steps in the same pathway where the upstream gene is epistatic.
3. Heritability H2 = VG/VP, the proportion of total phenotypic variance attributable to genetic variance.
4. PKU is caused by deficiency of phenylalanine hydroxylase, blocking conversion of phenylalanine to tyrosine; phenylalanine and toxic metabolites accumulate causing intellectual disability.
5. Two mutants with the same phenotype are crossed; if F1 is wild type, the mutations are in different genes (complement); if F1 is mutant, mutations are in the same gene.

Long-Answer Questions

1. Incomplete dominance: heterozygote intermediate; 1:2:1 ratio; snapdragon example. Codominance: both alleles expressed; ABO blood groups. Multiple alleles: more than two alleles per locus in a population; ABO and rabbit coat. Lethal alleles: homozygous lethal; yellow mouse 2:1 ratio; illustrates pleiotropy.
2. Epistasis: alleles at one locus mask another; types include 12:3:1 (recessive), 9:3:4, 15:1 (duplicate dominant), 9:7 (complementary). Complementary gene action requires dominant alleles at both loci; 9:7 ratio; sweet pea example. Complementation test determines if two mutations are in one or two genes by crossing mutants and observing F1.
3. Polygenic inheritance: multiple additive loci produce continuous normal distribution; 1:4:6:4:1 ratio for two loci. Heritability H2 = VG/VP, population-specific, high H2 does not preclude environmental modification. One gene one enzyme: Beadle and Tatum Neurospora auxotrophs; each gene controls one enzyme step; refined to one gene one polypeptide; inborn errors of metabolism (PKU, alkaptonuria) result from enzyme-deficient mutations.

Answers to Pilot Questions [Series 2]

Part 2.1

1. Incomplete dominance: heterozygote shows blended intermediate phenotype. Codominance: both alleles independently and simultaneously expressed in the heterozygote.
2. F2 ratio of 1 red : 2 pink : 1 white.
3. Multiple allelism: more than two allelic forms of a gene in a population; example: ABO blood group (IA, IB, i) or rabbit coat colour (C, cch, ch, c).
4. A: IAIA or IAi; B: IBIB or IBi; AB: IAIB; O: ii.
5. AY/AY is lethal as early embryos; only AY/A and A/A survive, giving a 2:1 yellow to agouti visible ratio rather than 3:1.

Part 2.2

1. Epistasis is gene interaction where alleles at one locus mask expression at a different locus, unlike dominance which is allelic interaction at the same locus.

2. If gene A encodes an enzyme producing the substrate for the enzyme encoded by gene B, homozygous recessive aa prevents substrate formation, so gene B cannot function regardless of its genotype.
3. 9:7 ratio; dominant alleles at both loci are required; only A_B_ produces the trait, while aaB_, A_bb, and aabb all produce the alternative phenotype.
4. Cross two mutants with the same phenotype; if F1 is wild type, mutations are in different genes; if F1 is mutant, mutations are in the same gene.
5. Duplicate dominant epistasis: a dominant allele at either locus produces the trait; only aabb shows the alternative phenotype.

Part 2.3

1. Polygenic inheritance: a trait determined by additive contributions of alleles at two or more loci; continuous variation arises from many possible genotypic combinations and environmental modifiers.
2. 1:4:6:4:1 ratio (five phenotypic classes with 0, 1, 2, 3, and 4 contributing alleles respectively).
3. Heritability H2 = VG/VP, the proportion of total phenotypic variance attributable to genetic variance; ranges from 0 to 1.
4. Heritability describes how much of the variation within current conditions is genetic; even traits with high H2 can change dramatically with environmental intervention (e.g. PKU managed by diet).
5. The norm of reaction is the range of phenotypes a given genotype produces across a range of environmental conditions.

Part 2.4

1. They irradiated Neurospora with X-rays, grew colonies on minimal medium to identify auxotrophic mutants, then tested growth on minimal medium supplemented with individual amino acids to identify the specific biosynthetic step blocked.
2. Some enzymes are multimeric (composed of polypeptides encoded by different genes) and some genes encode RNAs rather than polypeptides, so the one-to-one gene-enzyme correspondence is not universal.
3. An auxotroph cannot synthesise a required substance; supplementing with substances downstream of the block restores growth, while substances upstream do not, allowing the block position in the pathway to be identified.

4. Deficiency of phenylalanine hydroxylase blocks conversion of phenylalanine to tyrosine; phenylalanine and toxic metabolites accumulate, causing intellectual disability; managed by dietary phenylalanine restriction.
5. Garrod described alkaptonuria as the first inborn error of metabolism in 1902, recognising that a specific chemical abnormality results from a hereditary deficiency of a specific enzyme, decades before the molecular basis of genes was understood.

References and Attributions

Textual References (Books and External Sources)

- Brooker, R. J. (2015). Genetics: Analysis and principles (5th ed.). McGraw-Hill.
- Griffiths, A. J. F., Wessler, S. R., Carroll, S. B., & Doebley, J. (2015). Introduction to genetic analysis (11th ed.). W. H. Freeman.
- Strachan, T., & Read, A. (2018). Human molecular genetics (5th ed.). Garland Science.
- National Open University of Nigeria (NOUN). (2023). Genetics II course material. NOUN Press.

Figures, Plates, Tables, and Boxes

- Figure 2.1: Deviations from Mendelian ratios Attribution: AI generated. Suggested OER search string: "incomplete dominance codominance ABO blood group multiple alleles lethal allele genetics diagram OER".
- Figure 2.2: Epistasis and gene interactions Attribution: AI generated. Suggested OER search string: "epistasis gene interaction complementary duplicate gene action complementation test phenotypic ratio genetics OER".
- Figure 2.3: Polygenic inheritance and gene-environment interactions Attribution: AI generated. Suggested OER search string: "polygenic inheritance quantitative traits heritability norm of reaction gene environment interaction diagram genetics OER".
- Figure 2.4: Biochemical mutants and metabolic pathways Attribution: AI generated. Suggested OER search string: "one gene one enzyme Neurospora auxotroph biochemical mutant PKU alkaptonuria metabolic pathway inborn errors diagram OER".

Course code: BIO 301

Course title: Genetics II

Course credit load: 2 Units

Series 3: Nucleic Acids and Nucleotides

- **Part 3.1: Nucleotide Structure**
  - Sub Part 3.1.1: Components of nucleotides
  - Sub Part 3.1.2: Purine and pyrimidine bases
  - Sub Part 3.1.3: Nucleoside and nucleotide formation
- **Part 3.2: DNA Structure**
  - Sub Part 3.2.1: The Watson-Crick double helix model
  - Sub Part 3.2.2: Base pairing and base stacking
  - Sub Part 3.2.3: DNA supercoiling and higher-order structure
- **Part 3.3: RNA Structure and Types**
  - Sub Part 3.3.1: Primary structure of RNA
  - Sub Part 3.3.2: Types of RNA and their functions
  - Sub Part 3.3.3: RNA secondary and tertiary structure
- **Part 3.4: Evidence that DNA is the Genetic Material**
  - Sub Part 3.4.1: The Griffith transformation experiment
  - Sub Part 3.4.2: The Avery, MacLeod, and McCarty experiment
  - Sub Part 3.4.3: The Hershey-Chase experiment

Introduction

The molecular basis of heredity rests on the structure and properties of nucleic acids. The demonstration that DNA, rather than protein, is the genetic material transformed biology, and the elucidation of DNA double helix structure provided the molecular explanation for how genetic information is stored, copied, and expressed. This Series examines the chemistry of nucleotides and nucleic acids in depth, alongside the key experimental evidence that established DNA as the carrier of genetic information.

This Series covers the structure of nucleotides, including purine and pyrimidine bases; the Watson-Crick double helix model of DNA and the forces that stabilise it; the types and structures of RNA; and the landmark experiments by Griffith, Avery, MacLeod and McCarty, and Hershey and Chase that proved DNA is the genetic material. This molecular foundation directly supports the DNA replication and mutation topics addressed in the subsequent Series.

Series Learning Outcomes

Upon completing this Series, you should be able to:

- **SLO 3.1** describe the structure of nucleotides, including purine and pyrimidine bases, nucleosides, and nucleotides,
- **SLO 3.2** describe the Watson-Crick double helix model of DNA, base pairing, and higher-order DNA structure,
- **SLO 3.3** describe the primary structure of RNA, the major types of RNA and their functions, and RNA secondary structure, and
- **SLO 3.4** describe the key experiments that established DNA as the genetic material.

3.1 Nucleotide Structure

3.1.1 Components of nucleotides

A nucleotide consists of three covalently linked components: a pentose (five-carbon) sugar, a phosphate group, and a nitrogenous base; the sugar in DNA is 2-deoxyribose (lacking a hydroxyl group at the 2-carbon position), while the sugar in RNA is ribose (with a hydroxyl group at the 2-carbon position); this single structural difference between the two sugars has significant consequences for the chemical stability and three-dimensional structure of DNA versus RNA, with the 2-OH group of RNA making it more susceptible to hydrolysis than DNA.

The phosphate group (derived from phosphoric acid, H3PO4) is linked to the 5-carbon of the sugar, and the nitrogenous base is linked to the 1-carbon of the sugar via an N-glycosidic bond; in a polynucleotide chain, adjacent nucleotides are joined by phosphodiester bonds connecting the 3-carbon of one sugar to the 5-carbon of the next via a shared phosphate group, forming a sugar-phosphate backbone with the bases projecting to the side; the backbone carries a net negative charge due to the phosphate groups, making nucleic acids polyanions.

Fill in the blank: The sugar in DNA is _____ -deoxyribose, which lacks a hydroxyl group at the 2-carbon, making DNA more chemically stable than RNA.

3.1.2 Purine and pyrimidine bases

The nitrogenous bases are divided into two structural classes: purines (adenine and guanine) have a bicyclic structure composed of a six-membered pyrimidine ring fused to a five-membered imidazole ring; pyrimidines (cytosine, thymine, and uracil) have a single six-membered ring; DNA contains adenine, guanine, cytosine, and thymine; RNA contains adenine, guanine, cytosine, and uracil (substituting uracil for thymine, which differs only in having a methyl group at position 5 of the pyrimidine ring).

The bases contain nitrogen atoms that can participate in hydrogen bonding: adenine has an amino group at C6 and a nitrogen at N1 that serve as hydrogen bond donor and acceptor respectively; guanine has a carbonyl at C6 and an amino at C2; cytosine has an amino at C4 and a carbonyl at

C2; thymine has methylated carbonyl groups at C2 and C4; these hydrogen bonding groups are precisely positioned to allow specific A-T and G-C base pairing in the double helix, as described in the Watson-Crick model.

Fill in the blank: Purines (adenine and guanine) have a _____ ring structure, while pyrimidines (cytosine, thymine, uracil) have a single six-membered ring.

3.1.3 Nucleoside and nucleotide formation

A nucleoside is a base linked to a sugar (without a phosphate group), formed by an N-glycosidic bond between the anomeric carbon of the sugar (C1) and a nitrogen of the base (N9 of purines, N1 of pyrimidines); the nucleoside names are: adenosine and deoxyadenosine, guanosine and deoxyguanosine, cytidine and deoxycytidine, thymidine (exclusively deoxy), and uridine (exclusively ribo).

A nucleotide is a nucleoside with one or more phosphate groups attached to the 5-carbon of the sugar; monophosphate nucleotides (such as AMP, dAMP, GMP, CMP, UMP, TMP) are the building blocks of nucleic acid chains; diphosphate and triphosphate nucleotides (such as ATP, GTP, dATP, dGTP) serve as energy carriers and activated substrates for nucleic acid synthesis; the energy released by hydrolysis of the terminal phosphate group of a triphosphate nucleotide drives many biosynthetic reactions, including DNA and RNA polymerisation.

Fill in the blank: A _____ is a base linked to a sugar without a phosphate group, while a nucleotide includes one or more phosphate groups attached to the 5-carbon of the sugar.

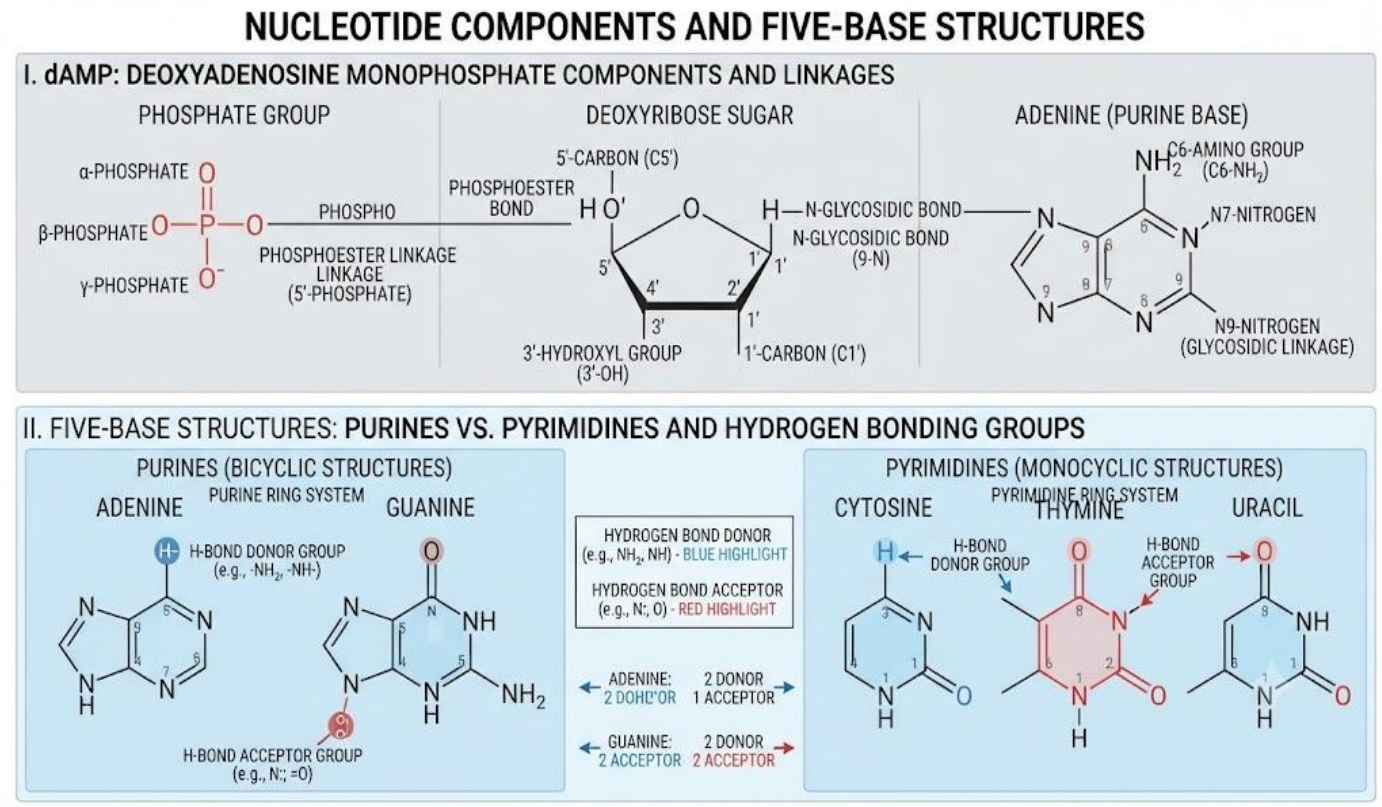


Figure 3.1: Nucleotide structure

Pilot questions 3.1

1. Name the three components of a nucleotide.
2. Distinguish between the sugars in DNA and RNA and explain the significance of the difference.
3. Name the two purines and three pyrimidines found in nucleic acids.
4. Which pyrimidine is found in DNA but not RNA, and which is found in RNA but not DNA?
5. Distinguish between a nucleoside and a nucleotide.

3.2 DNA Structure

3.2.1 The Watson-Crick double helix model

James Watson and Francis Crick, using X-ray diffraction data from Rosalind Franklin and Maurice Wilkins and base composition data from Erwin Chargaff (which showed that %A = %T and %G = %C in all DNA samples studied, now called Chargaff rules), proposed the double helix model of DNA in 1953; the model described DNA as two antiparallel (running in opposite 5 to 3 directions) polynucleotide chains wound around a common central axis with the sugar-phosphate backbones on the outside and the nitrogenous bases on the inside.

The model specified complementary base pairing between the two strands: adenine always pairs with thymine (A-T, connected by two hydrogen bonds) and guanine always pairs with cytosine (G-C, connected by three hydrogen bonds), an arrangement that both explains Chargaff rules and provides the molecular basis for the faithful copying of genetic information, since each strand can serve as a template for synthesis of a new complementary strand with predictable base sequence; the double helix has a diameter of approximately 2 nanometres and makes one complete turn every 10.5 base pairs (in the B-form helix predominant under physiological conditions).

Fill in the blank: Chargaff rules state that %A = %T and %G = %C, explained by _____ base pairing in the double helix where A pairs with T and G pairs with C.

3.2.2 Base pairing and base stacking

The double helix is stabilised by two types of non-covalent interactions: hydrogen bonds between complementary base pairs (two between each A-T pair and three between each G-C pair, making G-C pairs slightly stronger) and base stacking interactions (hydrophobic and van der Waals forces between the flat aromatic ring surfaces of adjacent bases stacked along the helix axis, perpendicular to the base pair plane); base stacking contributes more to helix stability than hydrogen bonding under physiological conditions, as the interior of the helix provides a hydrophobic environment that maximises the energetic benefit of stacking.

The major groove (wider, approximately 2.2 nm) and minor groove (narrower, approximately 1.2 nm) of the double helix provide distinct environments for protein-DNA interactions: regulatory proteins typically bind in the major groove because the pattern of hydrogen bond donors and acceptors projecting into the major groove is base-pair specific, allowing proteins to read the DNA sequence without needing to separate the strands; minor groove-binding proteins and small molecules (including some antibiotics) can also access the DNA sequence in the minor groove but with less sequence specificity.

Fill in the blank: The double helix is stabilised by hydrogen bonds between base pairs and _____ interactions between adjacent flat base ring surfaces stacked along the helix axis.

3.2.3 DNA supercoiling and higher-order structure

DNA supercoiling arises when additional turns are introduced into or removed from the double helix, causing the helix axis itself to coil; positive supercoiling (overwinding) occurs when extra turns are added in the direction of helix rotation, while negative supercoiling (underwinding) is the predominant form in most cells and arises when turns are removed; negative supercoiling facilitates strand separation (as required for replication and transcription) and compacts the DNA; topoisomerase enzymes (type I and type II) control supercoiling by transiently cutting and rejoining DNA strands.

In eukaryotes, the enormous length of chromosomal DNA (approximately 2 metres of DNA in each human cell) is compacted into nuclei of a few micrometres diameter through hierarchical levels of organisation: the first level is the nucleosome (approximately 147 base pairs of DNA wrapped around a histone octamer core of 8 histone proteins), which produces a beads-on-a-string appearance in chromatin; nucleosomes are further compacted by linker histones and additional folding into the 30 nm fibre and higher-order chromatin domains and loops.

Fill in the blank: In eukaryotes, DNA is wrapped around histone octamers to form _____ , the fundamental unit of chromatin organisation producing a beads-on-a-string appearance.

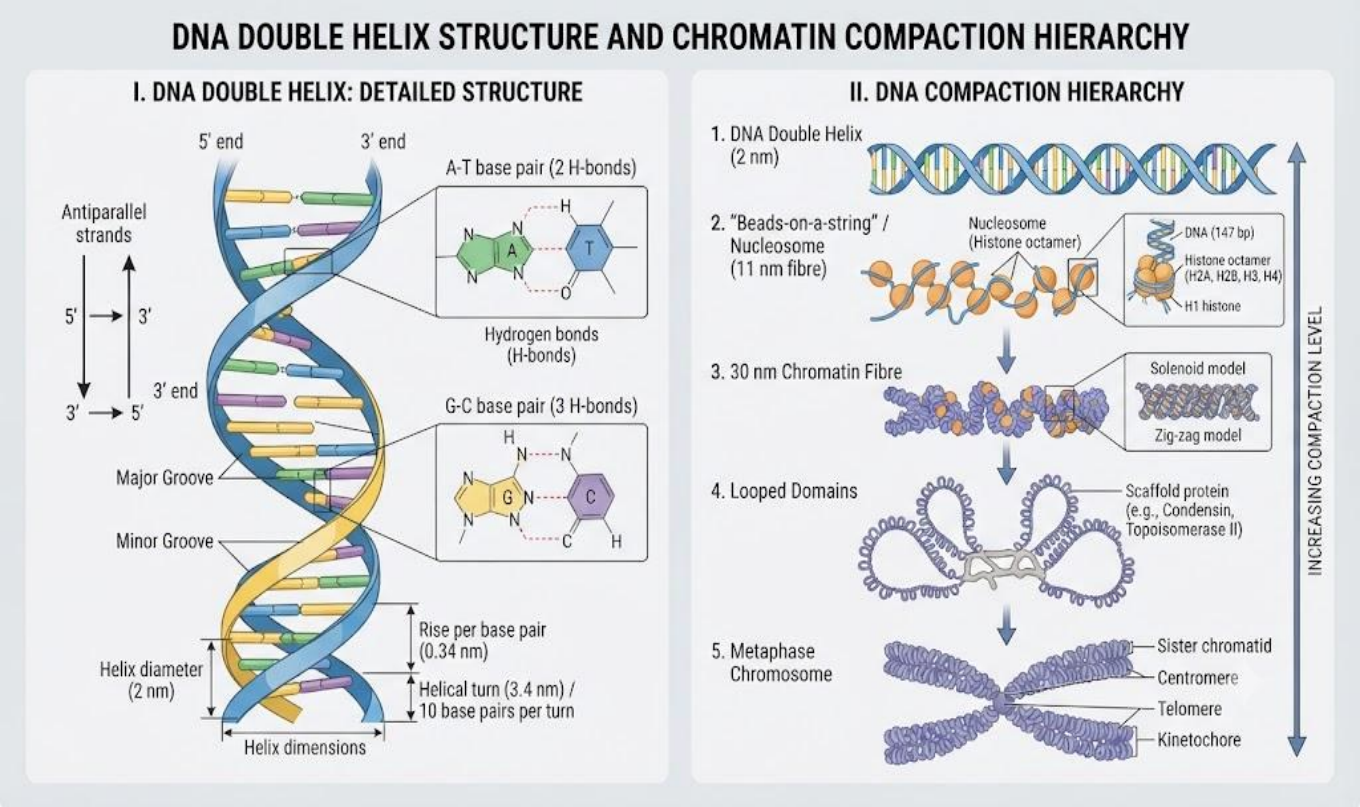


Figure 3.2: DNA structure

Pilot questions 3.2

1. State the key features of the Watson-Crick double helix model.
2. State Chargaff rules and explain how the double helix model accounts for them.
3. Distinguish between the two types of non-covalent forces stabilising the double helix.
4. Explain the significance of the major groove for protein-DNA interactions.
5. Describe the nucleosome and its role in DNA compaction.

3.3 RNA Structure and Types

3.3.1 Primary structure of RNA

RNA (ribonucleic acid) is generally a single-stranded polynucleotide in which the ribose sugar at each position carries a 2-OH group (unlike deoxyribose in DNA) and uracil substitutes for

thymine as the base pairing partner of adenine; the backbone is assembled in the same 5 to 3 direction as DNA using phosphodiester bonds, and the primary structure of RNA is determined by the DNA template from which it is transcribed, with each RNA base complementary to the corresponding DNA template strand base.

The presence of the 2-OH group makes RNA less stable than DNA (susceptible to alkaline hydrolysis, since the 2-OH can attack the adjacent phosphodiester bond) and provides a chemical handle that allows RNA molecules to adopt more diverse three-dimensional structures than DNA, contributing to the catalytic activity of ribozymes and the structural complexity of tRNA and rRNA; RNA is also typically much shorter-lived than DNA in the cell, with mRNA in particular having a relatively short and regulated half-life.

Fill in the blank: RNA contains _____ in place of thymine, and ribose in place of deoxyribose, with the 2-OH group making RNA more susceptible to hydrolysis than DNA.

3.3.2 Types of RNA and their functions

Messenger RNA (mRNA) is the information-carrying RNA molecule transcribed from a gene and translated by ribosomes to produce a specific protein; in eukaryotes, pre-mRNA is extensively processed before translation: it receives a 5-cap (7-methylguanosine cap protecting the 5-end and facilitating ribosome binding), a 3-poly(A) tail (added post-transcriptionally to protect the 3-end and facilitate nuclear export), and undergoes splicing (removal of non-coding intron sequences and joining of coding exon sequences by the spliceosome complex).

Transfer RNA (tRNA, approximately 73 to 93 nucleotides) carries specific amino acids to the ribosome during translation and recognises mRNA codons through its anticodon loop; ribosomal RNA (rRNA) is the major structural and catalytic component of ribosomes (both the large and small subunits), with the peptidyl transferase activity catalysing peptide bond formation being carried out by rRNA rather than by ribosomal proteins; small nuclear RNA (snRNA) is a component of the spliceosome; small interfering RNA (siRNA) and microRNA (miRNA) are short regulatory RNAs involved in post-transcriptional gene silencing.

Fill in the blank: Transfer RNA (tRNA) carries specific amino acids to the ribosome and recognises mRNA codons through its _____ loop.

3.3.3 RNA secondary and tertiary structure

Although RNA is single-stranded, intramolecular base pairing within a single RNA molecule allows regions of complementary sequence to fold back and pair, forming stem-loop (hairpin) structures, internal loops, bulges, and more complex motifs; the cloverleaf secondary structure of tRNA, for example, contains four stem-loop structures stabilised by Watson-Crick base pairs plus some non-Watson-Crick pairs; these secondary structure elements fold further into a compact L-shaped tertiary structure that is essential for tRNA function.

Ribozymes (catalytically active RNA molecules) depend on their precise three-dimensional structure to carry out specific chemical reactions; examples include the self-splicing group I and group II introns, ribonuclease P (cleaving tRNA precursors), and the peptidyl transferase activity of the large ribosomal subunit rRNA; the discovery of ribozymes by Thomas Cech and Sidney Altman (Nobel Prize 1989) challenged the assumption that only proteins could be enzymes and provided support for the RNA world hypothesis (that RNA molecules carrying both genetic information and catalytic activity preceded DNA and proteins in early life).

Fill in the blank: _____ are catalytically active RNA molecules; their discovery challenged the assumption that only proteins could function as enzymes.

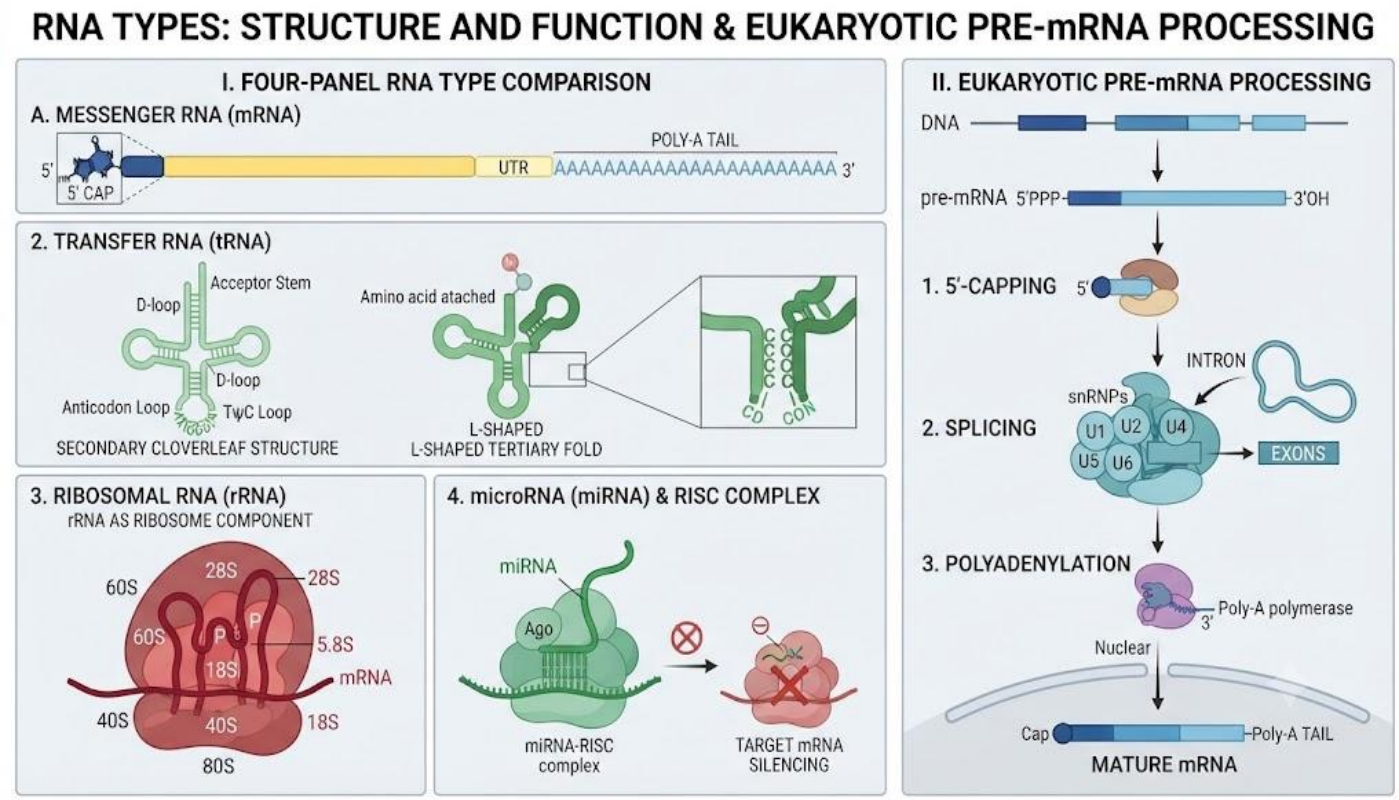


Figure 3.3: RNA structure and types

Pilot questions 3.3

1. Name two chemical differences between RNA and DNA.
2. Name four types of RNA and briefly describe the function of each.
3. Describe the three processing steps that convert pre-mRNA to mature mRNA in eukaryotes.
4. Describe the cloverleaf secondary structure of tRNA.
5. Define ribozyme and give one example.

3.4 Evidence that DNA is the Genetic Material

3.4.1 The Griffith transformation experiment

Frederick Griffith in 1928 demonstrated that a non-virulent (rough, R) strain of Streptococcus pneumoniae could be transformed into a virulent (smooth, S) form by mixing living R cells with

heat-killed S cells; the transformed bacteria acquired the ability to produce a polysaccharide capsule (giving the smooth colony appearance and protecting against phagocytosis), and this acquired virulence was heritable through subsequent generations; Griffith called the substance responsible the transforming principle, but did not identify its chemical nature.

The Griffith experiment established that genetic information could be transferred from killed cells to living cells in a stable, heritable form, setting the stage for the identification of the chemical nature of the transforming principle; the experiment demonstrated that something in the dead S cells was capable of permanently altering the hereditary properties of living R cells, but the available molecular understanding at the time was insufficient to identify the transforming principle, and many scientists assumed it was protein rather than nucleic acid.

Fill in the blank: Griffith called the substance that transformed non-virulent R bacteria into virulent S bacteria the _____ principle, but did not identify its chemical nature.

3.4.2 The Avery, MacLeod, and McCarty experiment

Oswald Avery, Colin MacLeod, and Maclyn McCarty published in 1944 the identification of the transforming principle as DNA; they prepared extracts from heat-killed S cells and systematically treated them with enzymes that degrade specific macromolecules: protease (destroying protein), RNase (destroying RNA), lipase (destroying lipids), and DNase (destroying DNA); only DNase treatment abolished transforming activity, while protease and RNase treatment left transforming activity intact, demonstrating that DNA, not protein or RNA, was the transforming principle.

Despite the rigour and clarity of the Avery-MacLeod-McCarty experiment, its conclusion was initially met with scepticism, partly because the dominant paradigm at the time held that proteins (with their 20 amino acid building blocks) were more informationally complex than nucleic acids (with only four bases) and therefore more likely to be the genetic material; it was not until the Hershey-Chase experiment of 1952 and the Watson-Crick structure of 1953 that DNA was widely accepted as the genetic material.

Fill in the blank: Avery, MacLeod, and McCarty showed that only _____ treatment abolished transforming activity, identifying DNA as the transforming principle.

3.4.3 The Hershey-Chase experiment

Alfred Hershey and Martha Chase in 1952 used bacteriophage T4 to definitively demonstrate that DNA, not protein, carries genetic information into a cell during infection; they exploited the fact that T4 phage DNA contains phosphorus but no sulfur, while T4 phage protein coat contains sulfur but no phosphorus; they labelled one batch of phage with radioactive 35S (labelling only the protein coat) and another with radioactive 32P (labelling only the DNA), then allowed each to infect bacteria.

After infection, they used a blender to separate phage coats from bacteria and a centrifuge to pellet the bacteria; the 35S label (protein) was found predominantly in the supernatant (phage coats outside bacteria) while the 32P label (DNA) was found predominantly inside the bacteria (and in phage offspring produced by infected bacteria); this demonstrated that DNA, not protein, was injected into the bacteria and directed the production of new phage particles, providing strong evidence that DNA is the genetic material.

Fill in the blank: In the Hershey-Chase experiment, radioactive _____ (labelling phage protein) remained outside the bacteria while 32P (labelling phage DNA) entered the bacteria, proving DNA is the genetic material.

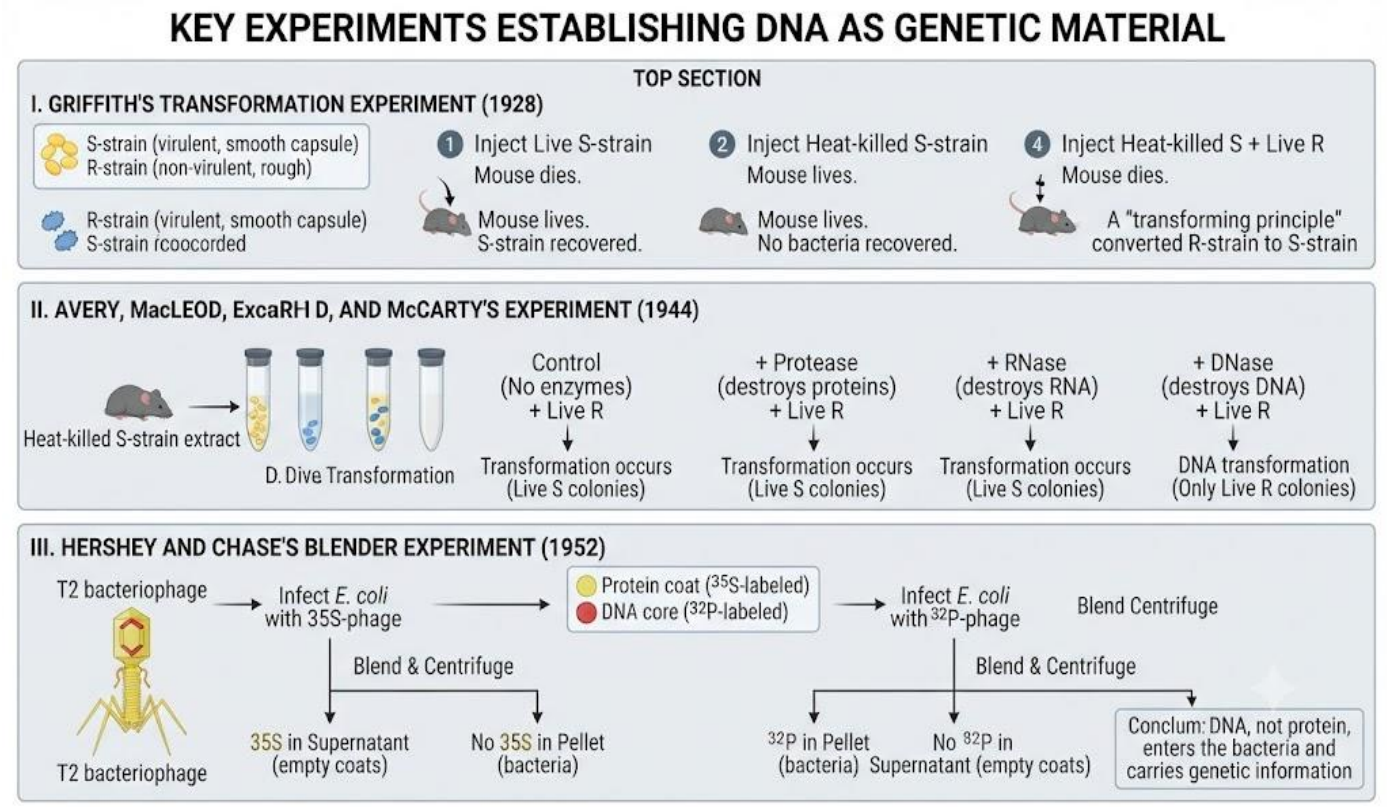


Figure 3.4: Evidence for DNA as the genetic material

Pilot questions 3.4

1. Describe the Griffith transformation experiment and its conclusion.
2. What did Avery, MacLeod, and McCarty do to identify the transforming principle?
3. Why did the Avery-MacLeod-McCarty result face initial scepticism?
4. Explain how Hershey and Chase used radioactive labels to distinguish DNA from protein.
5. What did the Hershey-Chase experiment conclude, and why was it more convincing than Avery-MacLeod-McCarty?

Series Summary

This Series examined the **structure and organisation of nucleic acids**, explaining how DNA and RNA store, transmit, and express genetic information. You explored the components of nucleotides, the structural differences between DNA and RNA, the Watson-Crick double helix, complementary base pairing, chromatin organisation, and the formation of phosphodiester bonds that create nucleic acid polymers.

The Series also examined the diverse structures and functions of RNA molecules, including mRNA, tRNA, rRNA, snRNA, and regulatory RNAs, as well as the catalytic role of ribozymes. In addition, you explored the landmark experiments of Griffith, Avery-MacLeod-McCarty, and Hershey-Chase, which established DNA as the hereditary material. By completing this Series, you should now be able to explain the molecular structure of DNA and RNA, distinguish the functions of different RNA types, describe the organisation of genetic material, and evaluate the experimental evidence supporting DNA as the genetic molecule, thereby achieving all four **Series Learning Outcomes (SLOs)** and the **overall course outcomes**.

Glossary of Terms

- **Anticodon:** the three-nucleotide sequence on tRNA that is complementary to the mRNA codon.
- **Chargaff rules:** the empirical observation that in DNA, %A = %T and %G = %C.
- **Nucleosome:** the fundamental unit of chromatin, consisting of approximately 147 bp of DNA wrapped around a histone octamer.
- **Phosphodiester bond:** the covalent bond linking adjacent nucleotides via a shared phosphate group connecting the 3-carbon of one sugar to the 5-carbon of the next.
- **Ribozyme:** a catalytically active RNA molecule capable of carrying out specific chemical reactions.
- **Spliceosome:** the ribonucleoprotein complex that removes introns from pre-mRNA.

- **Transforming principle:** the term used by Griffith for the substance that transferred virulence between pneumococcal strains, later identified as DNA by Avery, MacLeod, and McCarty.

Concept Check Questions [Series 3]

Objective Questions

1. The sugar in DNA is 2-_____ -deoxyribose, lacking the 2-OH group present in RNA ribose. (Linked to SLO 3.1)
2. In the DNA double helix, A pairs with T via _____ hydrogen bonds. (Linked to SLO 3.2)
3. Transfer RNA recognises mRNA codons through its _____ loop. (Linked to SLO 3.3)
4. Avery, MacLeod, and McCarty showed that only _____ treatment abolished transforming activity. (Linked to SLO 3.4)
5. _____ are catalytically active RNA molecules whose discovery challenged the protein-only enzyme paradigm. (Linked to SLO 3.3)

Short-Answer Questions

1. Distinguish between purines and pyrimidines. (Linked to SLO 3.1)
2. State the key features of the Watson-Crick double helix model. (Linked to SLO 3.2)
3. Name three types of RNA and state the function of each. (Linked to SLO 3.3)
4. Describe the Hershey-Chase experiment. (Linked to SLO 3.4)
5. Describe the three processing steps in eukaryotic pre-mRNA maturation. (Linked to SLO 3.3)

Long-Answer Questions

1. Describe the structure of nucleotides and the DNA double helix. (Linked to SLOs 3.1, 3.2)
2. Describe RNA types, their structures, and functions. (Linked to SLO 3.3)
3. Describe the three key experiments establishing DNA as the genetic material. (Linked to SLO 3.4)

Answers to Concept Check Questions [Series 3]

Objective Questions

1. Deoxy.
2. Two.
3. Anticodon.
4. DNase.
5. Ribozymes.

Short-Answer Questions

1. Purines (adenine, guanine) have a bicyclic structure; pyrimidines (cytosine, thymine, uracil) have a single six-membered ring.
2. Two antiparallel strands; sugar-phosphate backbone outside; bases inside; complementary base pairing (A-T, G-C); stabilised by hydrogen bonds and base stacking; diameter 2 nm.
3. mRNA carries genetic information to ribosomes; tRNA carries amino acids and recognises codons; rRNA is structural and catalytic component of ribosomes.
4. Phage labelled with 35S (protein) or 32P (DNA) were used to infect bacteria; after blending and centrifugation, 35S remained outside and 32P entered bacteria, proving DNA is the genetic material.
5. 5-cap addition, 3-poly(A) tail addition, and splicing (intron removal and exon joining by the spliceosome).

Long-Answer Questions

1. Nucleotides consist of a pentose sugar (deoxyribose in DNA, ribose in RNA), phosphate group, and nitrogenous base (purines A/G, pyrimidines C/T/U). The Watson-Crick double helix has antiparallel strands, complementary A-T and G-C pairing, sugar-phosphate backbone outside, bases inside, stabilised by hydrogen bonds and base stacking, with major and minor grooves. DNA is compacted into nucleosomes and higher chromatin structures.
2. mRNA: information carrier, processed in eukaryotes (5-cap, poly-A tail, splicing). tRNA: cloverleaf/L-shape, anticodon loop, carries amino acids. rRNA: ribosome structure and peptidyl transferase activity. snRNA: spliceosome component. siRNA/miRNA: gene

silencing. Ribozymes are catalytically active RNA molecules, exemplified by rRNA peptidyl transferase, self-splicing introns, and ribonuclease P.

3. Griffith (1928): heat-killed S cells plus live R cells transformed R to virulent S, defining the transforming principle. Avery-MacLeod-McCarty (1944): only DNase abolished transformation, identifying DNA as the transforming principle. Hershey-Chase (1952): 35S protein stayed outside bacteria while 32P DNA entered and directed phage production, definitively proving DNA is the genetic material.

Answers to Pilot Questions [Series 3]

Part 3.1

1. Pentose sugar, phosphate group, and nitrogenous base.
2. DNA has deoxyribose (no 2-OH); RNA has ribose (2-OH present); the 2-OH makes RNA more susceptible to hydrolysis and allows more structural diversity.
3. Purines: adenine and guanine. Pyrimidines: cytosine, thymine, and uracil.
4. Thymine is in DNA but not RNA; uracil is in RNA but not DNA.
5. A nucleoside is a base linked to a sugar; a nucleotide also includes one or more phosphate groups at the 5-carbon.

Part 3.2

1. Two antiparallel polynucleotide strands; sugar-phosphate backbone outside; bases inside; complementary A-T and G-C pairing; held by hydrogen bonds and base stacking; 2 nm diameter; 10.5 bp per turn.
2. Chargaff rules: %A = %T and %G = %C; explained by complementary base pairing where each A on one strand is paired with a T on the other, and each G with a C.
3. Hydrogen bonds between complementary base pairs (specific, between strands); base stacking interactions (hydrophobic/van der Waals, between adjacent bases within each strand).

4. The major groove exposes base-specific hydrogen bond donors and acceptors, allowing sequence-reading proteins to identify the base sequence without strand separation.
5. A nucleosome is approximately 147 bp of DNA wrapped around a histone octamer (8 histone proteins); it is the fundamental unit of chromatin compaction.

Part 3.3

1. RNA has ribose instead of deoxyribose, and uracil instead of thymine.
2. mRNA (carries genetic information from DNA to ribosome), tRNA (carries amino acids and recognises codons), rRNA (structural and catalytic ribosome component), miRNA/siRNA (post-transcriptional gene silencing).
3. 5-cap addition (7-methylguanosine), 3-poly(A) tail addition, and splicing by the spliceosome (intron removal and exon joining).
4. tRNA has four stem-loop structures forming a cloverleaf secondary structure, with an acceptor stem at the 3-end bearing the amino acid attachment site and an anticodon loop at the bottom.
5. A ribozyme is a catalytically active RNA; example: the rRNA peptidyl transferase activity in the large ribosomal subunit (or self-splicing introns, ribonuclease P).

Part 3.4

1. Griffith mixed living R (non-virulent) bacteria with heat-killed S (virulent) bacteria and injected them into mice; the mice died and living virulent S bacteria were recovered, showing that a heritable transforming principle from killed S cells had altered the R bacteria.
2. They treated S-cell extracts with protease, RNase, lipase, and DNase, and tested each treated extract for transforming activity; only DNase treatment abolished transformation, identifying DNA as the transforming principle.
3. The prevailing assumption was that proteins, with 20 amino acids, were more complex and informationally richer than nucleic acids with only 4 bases, making proteins the expected genetic material.
4. They labelled one phage batch with 35S (marks protein only) and another with 32P (marks DNA only); after infection and blender separation, 35S remained outside bacteria while 32P was found inside bacteria and in progeny phage.
5. The Hershey-Chase experiment concluded that DNA, not protein, is the genetic material; it was more convincing because it directly traced the molecule injected and used to direct phage

reproduction, whereas the Avery-MacLeod-McCarty result could still be questioned over possible protein contamination.

References and Attribution

Textual References (Books and External Sources)

- Watson, J. D., Baker, T. A., Bell, S. P., Gann, A., Levine, M., & Losick, R. (2014). Molecular biology of the gene (7th ed.). Cold Spring Harbor Laboratory Press.
- Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., & Walter, P. (2014). Molecular biology of the cell (6th ed.). Garland Science.
- Griffiths, A. J. F., Wessler, S. R., Carroll, S. B., & Doebley, J. (2015). Introduction to genetic analysis (11th ed.). W. H. Freeman.
- National Open University of Nigeria (NOUN). (2023). Genetics II course material. NOUN Press.

Figures, Plates, Tables, and Boxes

- Figure 3.1: Nucleotide structure Attribution: AI generated. Suggested OER search string: "nucleotide nucleoside purine pyrimidine adenine guanine cytosine thymine uracil structure diagram genetics OER".
- Figure 3.2: DNA structure Attribution: AI generated. Suggested OER search string: "DNA double helix base pairing major minor groove nucleosome chromatin compaction structure diagram genetics OER".
- Figure 3.3: RNA structure and types Attribution: AI generated. Suggested OER search string: "mRNA tRNA rRNA structure function types pre-mRNA processing splicing 5-cap poly-A tail ribozyme diagram genetics OER".
- Figure 3.4: Evidence for DNA as the genetic material Attribution: AI generated. Suggested OER search string: "Griffith transformation Avery MacLeod McCarty Hershey Chase DNA genetic material experiment diagram genetics OER".

Course code: BIO 301

Course title: Genetics II

Course credit load: 2 Units

Series 4: DNA Replication and Mutation

- **Part 4.1: Mechanisms of DNA Replication**
  - Sub Part 4.1.1: The semi-conservative model
  - Sub Part 4.1.2: Enzymes and proteins of replication
  - Sub Part 4.1.3: Replication in prokaryotes and eukaryotes
- **Part 4.2: Fidelity of DNA Replication**
  - Sub Part 4.2.1: Proofreading by DNA polymerase
  - Sub Part 4.2.2: Mismatch repair
  - Sub Part 4.2.3: Sources of replication errors
- **Part 4.3: Types of DNA Mutation**
  - Sub Part 4.3.1: Point mutations
  - Sub Part 4.3.2: Insertions, deletions, and frameshift mutations
  - Sub Part 4.3.3: Chromosomal mutations
- **Part 4.4: Causes, Consequences, and Repair of Mutations**
  - Sub Part 4.4.1: Spontaneous and induced mutations
  - Sub Part 4.4.2: Consequences of mutations
  - Sub Part 4.4.3: DNA repair mechanisms

Introduction

DNA replication must be both accurate and complete: every cell division requires the faithful duplication of the entire genome, with errors that escape correction becoming permanent mutations. Understanding how DNA is replicated with extraordinary fidelity, the nature of mutations that arise despite this fidelity, and the mechanisms by which cells detect and repair DNA damage are central topics in molecular genetics with direct relevance to cancer biology, human genetic disease, and evolutionary change.

This Series covers the semi-conservative mechanism of DNA replication and the enzymes and proteins involved; the proofreading and repair mechanisms that maintain replication fidelity; the classification of mutations by type and molecular mechanism; and the causes, consequences, and cellular responses to DNA damage, including the major DNA repair pathways. This directly supports the gene regulation and biotechnology topics addressed in the final Series of this course.

Series Learning Outcomes

Upon completing this Series, you should be able to:

- **SLO 4.1** describe the semi-conservative model of DNA replication and the enzymes and proteins required,
- **SLO 4.2** explain the mechanisms that maintain replication fidelity, including proofreading and mismatch repair,
- **SLO 4.3** classify and describe the major types of DNA mutation, and
- **SLO 4.4** explain the causes, consequences, and repair mechanisms for DNA mutations.

4.1 Mechanisms of DNA Replication

4.1.1 The semi-conservative model

Three models were proposed for DNA replication: conservative (the parental double helix remains intact and a completely new double helix is synthesised), semi-conservative (each strand of the parental double helix serves as a template for a new complementary strand, producing two daughter duplexes each containing one parental and one new strand), and dispersive (parental and new DNA are interspersed throughout both daughter molecules); the Meselson-Stahl experiment (1958) definitively demonstrated that replication is semi-conservative.

Meselson and Stahl grew E. coli in medium containing heavy nitrogen (15N) until all DNA was fully 15N-labelled (heavy), then transferred cells to medium with light nitrogen (14N) and sampled at each generation; after one generation, all DNA centrifuged as a single intermediate-density band (one 15N and one 14N strand per duplex, consistent with semi-conservative); after two generations, bands at intermediate and light density appeared in equal amounts (consistent only with semi-conservative, not conservative or dispersive); these results ruled out the conservative and dispersive models.

Fill in the blank: The Meselson-Stahl experiment demonstrated that DNA replication is _____ -conservative, with each daughter duplex containing one parental and one newly synthesised strand.

4.1.2 Enzymes and proteins of replication

DNA replication requires a coordinated assembly of multiple proteins: helicase unwinds the double helix by breaking hydrogen bonds between base pairs ahead of the replication fork; single-strand binding proteins (SSBPs) stabilise the separated strands and prevent re-annealing; primase synthesises a short RNA primer (8 to 12 nucleotides) providing a 3-OH terminus from which DNA polymerase can extend (because DNA polymerase cannot initiate a new chain and can only add nucleotides to a pre-existing 3-OH end); and topoisomerases relieve the torsional stress (positive supercoiling) that builds up ahead of the moving helicase.

DNA polymerase III (in prokaryotes) is the primary replicative polymerase, synthesising new DNA by adding deoxyribonucleoside triphosphates to the 3-OH end of the primer in the 5 to 3 direction, using each parental strand as a template; because both new strands must be synthesised 5 to 3 but the two template strands run antiparallel, one strand (the leading strand) is synthesised continuously toward the replication fork, while the other (the lagging strand) is synthesised discontinuously as a series of short Okazaki fragments (each initiated by a new RNA primer), which are later joined by DNA ligase after RNA primer removal and gap filling by DNA polymerase I.

Fill in the blank: DNA polymerase can only add nucleotides to a pre-existing 3-OH terminus, which is provided by an RNA _____ synthesised by primase before polymerisation begins.

4.1.3 Replication in prokaryotes and eukaryotes

Prokaryotic chromosomes (circular) have a single origin of replication (oriC in E. coli), at which the replication bubble initiates and two replication forks proceed bidirectionally around the chromosome until they meet at the terminus; the entire E. coli chromosome (approximately 4.6 million base pairs) is replicated in approximately 40 minutes; the relatively simple replication machinery of prokaryotes and the single origin make prokaryotic replication mechanistically simpler than eukaryotic replication.

Eukaryotic chromosomes are linear, much larger (hundreds of millions of base pairs per chromosome), and have multiple origins of replication (hundreds to thousands per chromosome) that fire in a regulated temporal sequence during S phase; the multiple origins allow the entire eukaryotic genome to be replicated in a few hours despite the greater total length; eukaryotic replication also involves the assembly of newly replicated DNA into nucleosomes, ensuring that chromatin structure is restored as well as DNA sequence; telomere replication requires a specialised enzyme, telomerase, to extend the lagging strand template at chromosome ends.

Fill in the blank: Eukaryotic chromosomes have _____ origins of replication that fire in a regulated sequence during S phase, allowing replication of the large eukaryotic genome in a few hours.

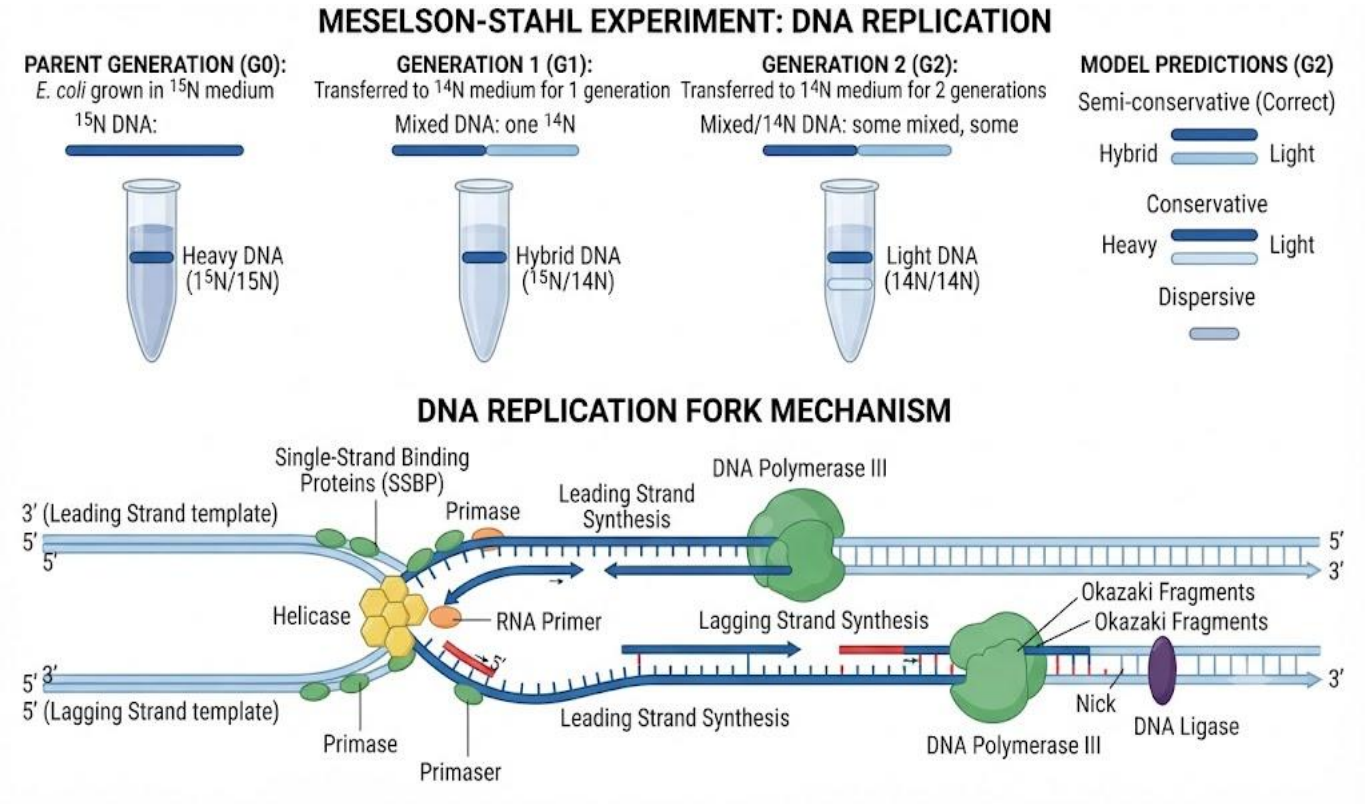


Figure 4.1: DNA replication mechanisms

Pilot questions 4.1

1. Describe the three models of DNA replication and state which was proved correct.
2. Describe the Meselson-Stahl experiment and its results after one and two generations.
3. Explain why DNA polymerase requires an RNA primer.
4. Explain why the lagging strand is synthesised as Okazaki fragments.
5. State two differences between prokaryotic and eukaryotic DNA replication.

4.2 Fidelity of DNA Replication

4.2.1 Proofreading by DNA polymerase

DNA replication is extraordinarily accurate, with an error rate of approximately one mistake per 10 to the power of 9 to 10 to the power of 10 base pairs replicated; this fidelity is achieved through multiple mechanisms working in concert; the intrinsic base selection accuracy of DNA

polymerase (approximately one error per 10 to the power of 5) is greatly enhanced by the 3 to 5 exonuclease proofreading activity built into the polymerase: immediately after incorporating each new nucleotide, the polymerase checks whether the correct base pair has been formed; if a mismatched base is detected, the 3 to 5 exonuclease activity excises the incorrect nucleotide and the polymerase tries again.

Proofreading reduces the error rate by approximately 100-fold, from about one in 10 to the power of 5 to about one in 10 to the power of 7; the structural mechanism relies on the fact that a correctly paired base pair fits precisely into the active site geometry of the polymerase, while a mismatch introduces a geometric distortion that is detected by the proofreading domain; the high energetic cost of proofreading (hydrolysis of the phosphodiester bond of the just-incorporated mismatched nucleotide) is justified by the critical importance of replication fidelity for genome stability.

Fill in the blank: DNA polymerase proofreading uses a 3 to 5 _____ activity to excise misincorporated nucleotides immediately after their incorporation.

4.2.2 Mismatch repair

Post-replication mismatch repair (MMR) provides a third line of defence against replication errors, correcting mismatches that escape the proofreading activity of DNA polymerase; the MMR system must be able to distinguish the newly synthesised (potentially erroneous) strand from the parental (template) strand so that it corrects the error on the new strand rather than altering the parental template; in E. coli, the parental strand is distinguished by adenine methylation at GATC sequences (the newly replicated strand is transiently unmethylated, allowing the MMR system to identify it).

In humans and other eukaryotes, the strand discrimination mechanism for MMR is less completely understood but likely involves strand breaks and other signals at the replication fork; key MMR proteins in humans include MSH2, MSH6, MLH1, and PMS2; defects in the human MMR genes cause Lynch syndrome (hereditary non-polyposis colorectal cancer, HNPCC), an autosomal dominant cancer predisposition syndrome illustrating the critical role of MMR in preventing cancer-causing mutations; MMR deficiency produces microsatellite instability (MSI),

a characteristic pattern of length variation at repetitive DNA sequences used as a diagnostic marker.

Fill in the blank: Mismatch repair (MMR) must distinguish the newly synthesised strand from the parental strand to ensure _____ are corrected on the new strand rather than the template.

4.2.3 Sources of replication errors

Replication errors arise from several sources: misincorporation of incorrect nucleotides (the most common source, reduced by polymerase base selection accuracy and proofreading); template misalignment, in which the template or primer strand slips during replication of repetitive sequences, causing insertions or deletions (particularly at simple sequence repeats); and lesions in the template DNA (including oxidative damage, deamination of bases, and UV photoproducts) that cause the replicative polymerase to stall or misincorporate opposite the lesion.

The presence of DNA lesions that block the replicative polymerase is handled by translesion synthesis (TLS) polymerases, a specialised family of lower-fidelity DNA polymerases that can bypass lesions that block the main replicative polymerase; TLS polymerases are less accurate than the main replicative polymerase because their active sites are more accommodating to accommodate the distorted geometry of lesion-containing template bases, but they allow replication to continue through otherwise blocking lesions, often at the cost of some increased mutagenesis.

Fill in the blank: _____ synthesis polymerases bypass DNA lesions that block the main replicative polymerase, allowing replication to continue at the cost of reduced accuracy.

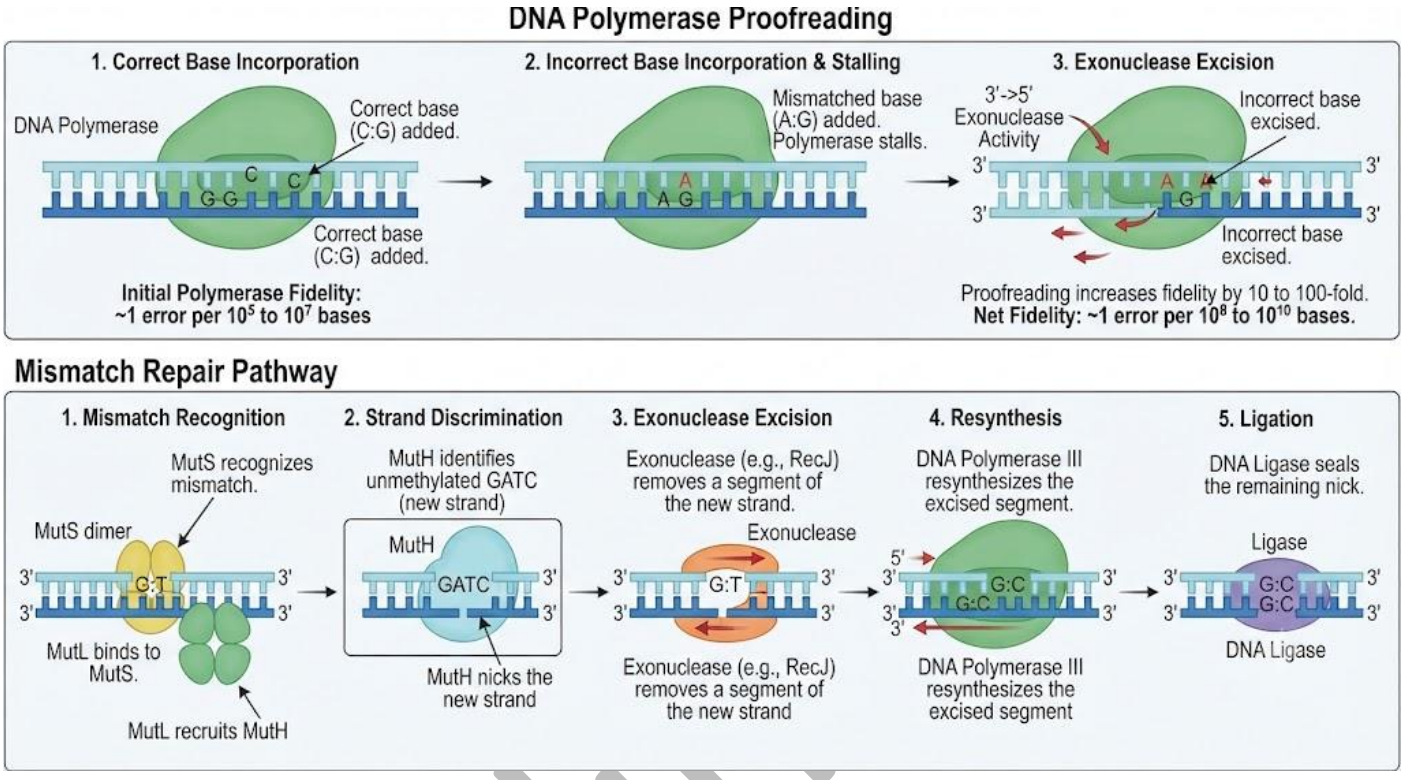


Figure 4.2: Replication fidelity and mismatch repair

Pilot questions 4.2

1. Explain how DNA polymerase proofreading improves replication accuracy.
2. Explain how the mismatch repair system identifies which strand to correct in E. coli.
3. Name two human diseases associated with defects in DNA repair genes.
4. Explain what translesion synthesis polymerases do and why they are less accurate.
5. What is microsatellite instability and what causes it?

4.3 Types of DNA Mutation

4.3.1 Point mutations

Point mutations are changes affecting a single base pair and are divided into two main classes: transitions (substitution of a purine for a purine or a pyrimidine for a pyrimidine, for example A-T to G-C or G-C to A-T) and transversions (substitution of a purine for a pyrimidine or vice

versa, for example A-T to C-G or G-C to T-A); transitions are generally more common than transversions because the geometric change between two purines or two pyrimidines is smaller and less disruptive.

The consequence of a point mutation in a coding sequence depends on its position and the degeneracy of the genetic code: a synonymous (silent) mutation changes the codon to another codon specifying the same amino acid (due to the redundancy of the genetic code, particularly at the third codon position); a missense mutation changes the codon to one specifying a different amino acid, potentially altering protein function (the severity depending on how chemically different the new amino acid is from the original); and a nonsense (stop) mutation changes a sense codon to a stop codon, prematurely truncating the protein and usually severely disrupting function.

Fill in the blank: A _____ mutation changes a codon to one specifying a different amino acid, while a nonsense mutation changes a sense codon to a stop codon.

4.3.2 Insertions, deletions, and frameshift mutations

Insertions and deletions (collectively indels) involve the addition or removal of one or more base pairs; when the number of inserted or deleted bases is not a multiple of three, a frameshift mutation results, altering the reading frame of the mRNA from the point of the insertion or deletion onward; frameshifts typically produce a completely different and usually non-functional protein sequence downstream of the mutation, followed by an early stop codon arising by chance in the altered reading frame.

Insertions or deletions of a multiple of three base pairs (in-frame indels) add or remove one or more amino acids from the protein without disrupting the reading frame, and may be less severely disruptive depending on the structural and functional importance of the affected protein region; repeat expansion mutations (expansion of tandem repeat sequences such as trinucleotide repeats) are a specific category of insertion responsible for several human neurodegenerative diseases including Huntington disease (CAG repeat expansion in the huntingtin gene) and fragile X syndrome (CGG repeat expansion in the FMR1 gene).

Fill in the blank: A frameshift mutation results from an insertion or deletion of a number of bases that is _____ a multiple of three, shifting the reading frame of all downstream codons.

4.3.3 Chromosomal mutations

Chromosomal mutations (also called chromosomal aberrations) affect larger regions of chromosomal DNA or involve changes in chromosome number or arrangement; the major types of chromosomal structural mutations are the same as the chromosomal structural abnormalities discussed in the context of human genetic disorders: deletions (loss of a chromosomal segment), duplications (extra copy of a segment), inversions (reversal of a segment within a chromosome), and translocations (movement of a segment between chromosomes, either reciprocal or non-reciprocal).

Chromosomal numerical changes (aneuploidy) involve gain or loss of individual chromosomes (such as trisomy or monosomy), while polyploidy involves the addition of complete chromosome sets; these numerical changes can arise from non-disjunction at meiosis or mitosis, as described for human chromosomal disorders; in evolutionary terms, chromosomal rearrangements (particularly translocations and inversions) play important roles in reproductive isolation and speciation, since organisms with different chromosomal arrangements may produce chromosomally unbalanced offspring that are infertile, reducing gene flow between populations.

Fill in the blank: Chromosomal numerical changes such as _____ can arise from non-disjunction at meiosis, resulting in gametes with abnormal chromosome numbers.

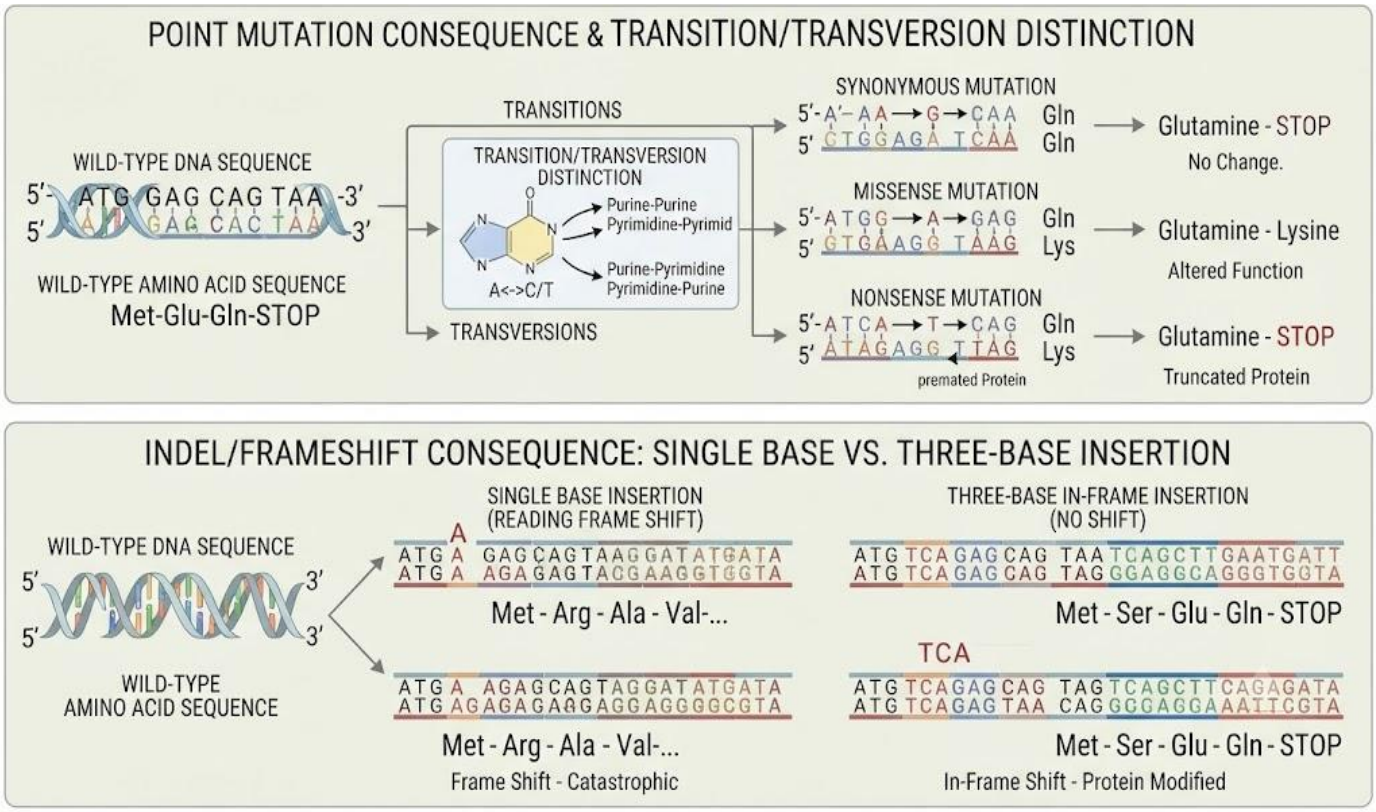


Figure 4.3: Types of DNA mutation

Pilot questions 4.3

1. Distinguish between transitions and transversions.
2. Distinguish between synonymous, missense, and nonsense mutations.
3. Explain why a frameshift mutation is typically more damaging than a missense mutation.
4. Name two human diseases caused by trinucleotide repeat expansions.
5. Name four types of chromosomal structural mutation.

4.4 Causes, Consequences, and Repair of Mutations

4.4.1 Spontaneous and induced mutations

Spontaneous mutations arise in the absence of any known mutagenic agent, from inherent chemical instabilities of DNA: tautomeric shifts (rare but transient alternative hydrogen bonding

configurations of bases that cause mispairing; the enol form of thymine can pair with guanine rather than adenine), depurination (spontaneous hydrolysis of the N-glycosidic bond releasing a purine base and leaving an abasic site), deamination of cytosine to uracil (particularly common, producing a C-to-T transition if not repaired before replication), and replication slippage at repetitive sequences.

Induced mutations are caused by external mutagens, classified as chemical or physical: chemical mutagens include base analogues (such as 5-bromouracil, which substitutes for thymine but can pair with guanine, causing A-T to G-C transitions), intercalating agents (such as ethidium bromide and acridine dyes, inserting between base pairs and causing frameshift mutations), and alkylating agents (such as ethyl methanesulfonate/EMS and nitrogen mustards, adding alkyl groups to bases that cause mispairing or block replication); physical mutagens include UV radiation (producing pyrimidine dimers, particularly thymine dimers) and ionising radiation (X-rays and gamma rays, causing double-strand breaks and base modifications).

Fill in the blank: UV radiation causes the formation of pyrimidine _____ between adjacent pyrimidines on the same strand, particularly thymine dimers, which block normal replication.

4.4.2 Consequences of mutations

The consequences of a mutation depend on its location (coding versus non-coding regions), its nature (synonymous, missense, nonsense, or frameshift), and whether the affected gene product has redundant functions; mutations in critical structural or catalytic residues of a protein are more likely to be functionally significant than mutations in less constrained regions; most new mutations in coding regions are slightly deleterious rather than either neutral or strongly harmful, consistent with the distribution of their effects on protein function.

Somatic mutations (arising in non-reproductive cells) affect only the individual in which they arise and cannot be inherited by offspring; they can, however, contribute to cancer (a somatic genetic disease requiring accumulation of multiple driver mutations in oncogenes and tumour suppressor genes within a single cell lineage); germline mutations (arising in reproductive cells or early embryos) are heritable and can be transmitted to all cells of offspring, where they may

produce inherited genetic diseases or, if phenotypically neutral, contribute to polymorphism and evolutionary change.

Fill in the blank: _____ mutations arise in non-reproductive cells and cannot be inherited by offspring, but can contribute to cancer through accumulation of driver mutations.

4.4.3 DNA repair mechanisms

Cells have evolved multiple DNA repair pathways to correct different types of DNA damage: base excision repair (BER) removes chemically altered or inappropriate bases (such as uracil from cytosine deamination, oxidised bases, and alkylated bases) by DNA glycosylases that cleave the N-glycosidic bond, followed by AP endonuclease cleavage, gap filling, and ligation; nucleotide excision repair (NER) removes bulky helix-distorting lesions such as UV-induced thymine dimers by incising the damaged strand on both sides of the lesion, removing an oligonucleotide fragment of approximately 25 to 30 nucleotides, and filling and ligating the gap.

Double-strand breaks (the most dangerous type of DNA damage, capable of causing chromosomal rearrangements or cell death if unrepaired) are repaired by homologous recombination (HR, using an intact sister chromatid as a template for accurate repair, preferentially in S and G2 phases when a sister chromatid is available) or non-homologous end joining (NHEJ, directly joining broken ends without a template, which is error-prone but available throughout the cell cycle); SOS response in bacteria (global induction of DNA repair genes when DNA damage is extensive); and direct reversal repair, which includes photoreactivation (light-activated cleavage of thymine dimers by photolyase) and repair of O6-methylguanine by O6-methylguanine-DNA methyltransferase (MGMT).

Fill in the blank: Nucleotide excision repair removes bulky helix-distorting lesions such as _____ dimers by excising an approximately 25 to 30 nucleotide fragment around the damage.

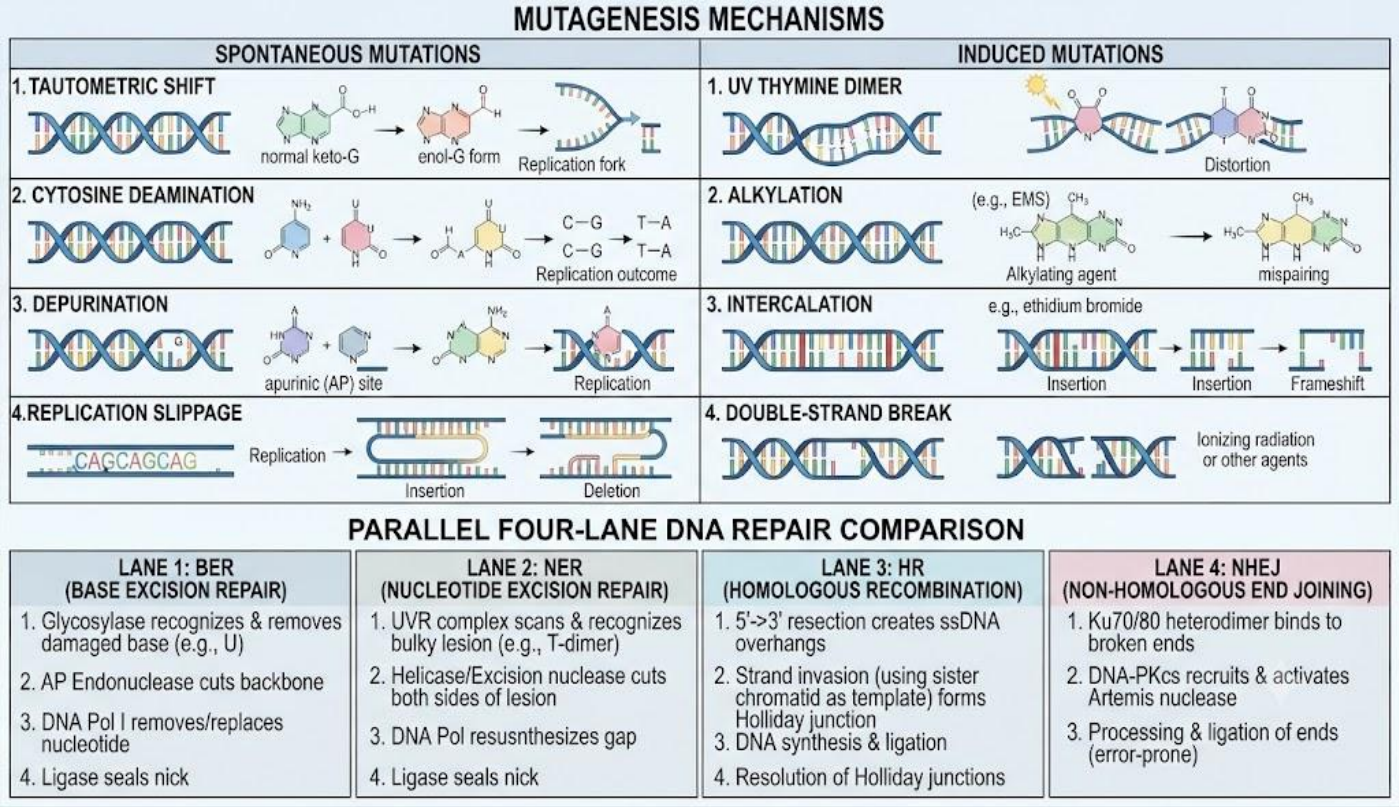


Figure 4.4: Mutation causes and DNA repair pathways

Pilot questions 4.4

1. Name three sources of spontaneous mutations.
2. Distinguish between chemical and physical mutagens and give one example of each.
3. Distinguish between somatic and germline mutations in terms of inheritance and consequences.
4. Describe base excision repair and state the type of damage it corrects.
5. Distinguish between homologous recombination and non-homologous end joining as double-strand break repair mechanisms.

Series Summary

This Series explored **DNA replication and mutation**, explaining how genetic information is accurately copied and maintained across generations. You examined the semi-conservative model of DNA replication, the roles of key replication enzymes, the synthesis of leading and lagging strands, the significance of multiple replication origins and telomerase, and the mechanisms that ensure high replication fidelity through proofreading and mismatch repair.

The Series also examined the different types, causes, and consequences of genetic mutations, ranging from point mutations and insertions or deletions to chromosomal abnormalities. In addition, you explored spontaneous and induced mutagens, the effects of mutations on health and disease, and the major DNA repair pathways that preserve genome integrity. By completing this Series, you should now be able to explain the molecular mechanisms of DNA replication, classify and evaluate different types of mutations, relate mutations to inherited disorders and cancer, and describe the cellular systems that repair DNA damage, thereby achieving all four **Series Learning Outcomes (SLOs)** and the **overall course outcomes**.

Glossary of Terms

- **Base excision repair (BER):** a repair pathway that removes chemically altered individual bases using DNA glycosylases.
- **Frameshift mutation:** a mutation caused by an insertion or deletion not a multiple of three, shifting the reading frame.
- **Missense mutation:** a point mutation that changes a codon to one specifying a different amino acid.
- **Non-homologous end joining (NHEJ):** error-prone repair of double-strand breaks by direct ligation of broken ends.
- **Nonsense mutation:** a point mutation that changes a sense codon to a stop codon, prematurely terminating translation.

- **Okazaki fragment:** a short DNA fragment synthesised on the lagging strand template at a replication fork.
- **Semi-conservative replication:** the mode of DNA replication in which each daughter duplex contains one parental and one newly synthesised strand.

Concept Check Questions [Series 4]

Objective Questions

1. The Meselson-Stahl experiment demonstrated that DNA replication is _____ -conservative. (Linked to SLO 4.1)
2. DNA polymerase proofreading uses a 3 to 5 _____ activity to excise misincorporated nucleotides. (Linked to SLO 4.2)
3. A _____ mutation changes a sense codon to a stop codon, prematurely truncating the protein. (Linked to SLO 4.3)
4. UV radiation causes pyrimidine _____ between adjacent bases, particularly thymine dimers. (Linked to SLO 4.4)
5. Nucleotide excision repair removes bulky lesions such as _____ dimers by excising a 25-30 nt fragment. (Linked to SLO 4.4)

Short-Answer Questions

1. Explain why the lagging strand is synthesised discontinuously as Okazaki fragments. (Linked to SLO 4.1)
2. Explain how MMR distinguishes the new strand from the parental strand in E. coli. (Linked to SLO 4.2)
3. Distinguish between a missense and a nonsense mutation. (Linked to SLO 4.3)
4. Distinguish between somatic and germline mutations. (Linked to SLO 4.4)
5. Distinguish between homologous recombination and NHEJ. (Linked to SLO 4.4)

Long-Answer Questions

1. Describe the semi-conservative model and the key enzymes of DNA replication. (Linked to SLO 4.1)
2. Describe the mechanisms that maintain replication fidelity. (Linked to SLO 4.2)
3. Describe the types of DNA mutation and their causes and repair. (Linked to SLOs 4.3, 4.4)

Answers to Concept Check Questions [Series 4]

Objective Questions

1. Semi.
2. Exonuclease.
3. Nonsense.
4. Dimers.
5. Thymine.

Short-Answer Questions

1. DNA polymerase synthesises only 5 to 3, but the lagging strand template runs 3 to 5 away from the fork; new fragments must be initiated repeatedly in the direction toward the fork.
2. The parental strand is methylated at GATC sequences; the newly synthesised strand is transiently unmethylated, allowing the MMR system to identify and correct it.
3. A missense mutation changes a codon to one for a different amino acid; a nonsense mutation changes a sense codon to a stop codon, truncating the protein.
4. Somatic mutations occur in non-reproductive cells and affect only the individual; germline mutations occur in reproductive cells and are heritable by offspring.
5. HR uses a sister chromatid as template for accurate repair; NHEJ directly joins broken ends without a template and is error-prone.

Long-Answer Questions

1. Semi-conservative replication: each daughter duplex has one parental and one new strand (Meselson-Stahl). Key enzymes: helicase (unwinds), SSBP (stabilises), primase (RNA primer), DNA polymerase III (5-to-3 synthesis), topoisomerase (relieves supercoiling), DNA polymerase I (removes primers), ligase (joins fragments). Leading strand continuous; lagging strand as Okazaki fragments.
2. Three mechanisms: base selection accuracy (~1 in 10^5); 3-to-5 proofreading exonuclease (reduces to ~1 in 10^7); mismatch repair (distinguishes new strand by methylation in E. coli; corrects mismatches post-replication). Translesion synthesis polymerases bypass blocking lesions at cost of accuracy.

3. Point mutations: transitions/transversions; synonymous, missense, or nonsense effects. Indels: frameshift if not multiple of three; repeat expansions (Huntington, fragile X). Chromosomal: deletions, duplications, inversions, translocations, aneuploidy. Causes: spontaneous (deamination, depurination, tautomeric shifts); induced (UV dimers, alkylating agents, ionising radiation). Repair: BER (altered bases), NER (bulky lesions), HR (accurate DSB repair), NHEJ (error-prone DSB repair).

Answers to Pilot Questions [Series 4]

Part 4.1

1. Conservative: parental double helix intact, new double helix synthesised. Semi-conservative: each strand templates a new strand. Dispersive: parental/new DNA interspersed. Semi-conservative was proved correct by Meselson-Stahl.
2. E. coli grown in 15N transferred to 14N: after 1 generation, single intermediate-density band (one 15N/one 14N strand per duplex); after 2 generations, equal intermediate and light bands. Only semi-conservative model is consistent.
3. DNA polymerase can only extend an existing strand from a 3-OH end; it cannot initiate a new chain, so a short RNA primer must first be synthesised by primase.
4. DNA polymerase synthesises only 5 to 3; the lagging strand template runs antiparallel to the fork direction, so new fragments must be initiated repeatedly in the opposing direction.
5. Prokaryotes have one circular chromosome with one origin; eukaryotes have linear chromosomes with multiple origins. Eukaryotic replication also involves nucleosome reassembly and requires telomerase for chromosome end replication.

Part 4.2

1. If an incorrect nucleotide is incorporated, a geometric mismatch is detected; the 3-to-5 exonuclease domain excises the wrong nucleotide, and the correct one is inserted; this reduces error rate approximately 100-fold.
2. The parental strand is methylated at GATC; the newly synthesised strand is transiently unmethylated; MMR proteins recognise the mismatch and identify the unmethylated strand as the one to correct.
3. Lynch syndrome (hereditary non-polyposis colorectal cancer, HNPCC) from MMR defects; xeroderma pigmentosum from NER defects.
4. TLS polymerases have more accommodating active sites that tolerate damaged templates, allowing bypass; they are less accurate because their looser geometry also tolerates incorrect base pairing.
5. Microsatellite instability (MSI) is the characteristic length variation at simple sequence repeats caused by MMR deficiency failing to correct replication slippage at these sequences.

Part 4.3

1. Transitions: purine for purine or pyrimidine for pyrimidine substitution. Transversions: purine for pyrimidine or vice versa.
2. Synonymous: same amino acid encoded (silent). Missense: different amino acid encoded. Nonsense: sense codon changed to stop codon.
3. A frameshift alters all downstream codons from the mutation point onward, producing an entirely different protein sequence and usually an early stop codon, making frameshifts almost always severely disruptive.
4. Huntington disease (CAG repeat expansion in huntingtin) and fragile X syndrome (CGG repeat expansion in FMR1).
5. Deletion, duplication, inversion, and translocation.

Part 4.4

1. Tautomeric shifts, cytosine deamination to uracil, and depurination (or replication slippage at repetitive sequences).
2. Chemical mutagen: alkylating agent (e.g. EMS) adds alkyl groups to bases causing mispairing. Physical mutagen: UV radiation produces thymine dimers.
3. Somatic mutations affect only the individual and cannot be inherited; they can contribute to cancer. Germline mutations are heritable and can produce inherited genetic disease or polymorphism.
4. BER uses DNA glycosylases to remove damaged or inappropriate individual bases (e.g. uracil, oxidised bases), followed by AP endonuclease, gap filling, and ligation.
5. HR uses a sister chromatid as a template for accurate repair of double-strand breaks (preferred in S/G2 phase). NHEJ directly joins broken ends without a template and is error-prone, available throughout the cell cycle.

References and Attribution

Textual References (Books and External Sources)

- Watson, J. D., Baker, T. A., Bell, S. P., Gann, A., Levine, M., & Losick, R. (2014). Molecular biology of the gene (7th ed.). Cold Spring Harbor Laboratory Press.
- Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., & Walter, P. (2014). Molecular biology of the cell (6th ed.). Garland Science.
- Griffiths, A. J. F., Wessler, S. R., Carroll, S. B., & Doebley, J. (2015). Introduction to genetic analysis (11th ed.). W. H. Freeman.
- National Open University of Nigeria (NOUN). (2023). Genetics II course material. NOUN Press.

Figures, Plates, Tables, and Boxes

- Figure 4.1: DNA replication mechanisms Attribution: AI generated. Suggested OER search string: "Meselson Stahl semi-conservative replication DNA replication fork helicase primase Okazaki fragment ligase diagram genetics OER".
- Figure 4.2: Replication fidelity and mismatch repair Attribution: AI generated. Suggested OER search string: "DNA polymerase proofreading mismatch repair MMR fidelity exonuclease strand discrimination diagram genetics OER".
- Figure 4.3: Types of DNA mutation Attribution: AI generated. Suggested OER search string: "point mutation synonymous missense nonsense frameshift insertion deletion transition transversion genetics diagram OER".
- Figure 4.4: Mutation causes and DNA repair pathways Attribution: AI generated. Suggested OER search string: "DNA repair base excision nucleotide excision homologous recombination NHEJ UV thymine dimer mutation diagram genetics OER".

Course code: BIO 301

Course title: Genetics II

Course credit load: 2 Units

Series 5: Proteins, Gene Regulation, and DNA Technology

- **Part 5.1: Protein Synthesis**
  - Sub Part 5.1.1: Transcription
  - Sub Part 5.1.2: The genetic code
  - Sub Part 5.1.3: Translation
- **Part 5.2: Regulation of Gene Expression**
  - Sub Part 5.2.1: The lac operon in prokaryotes
  - Sub Part 5.2.2: Eukaryotic gene regulation
  - Sub Part 5.2.3: Post-transcriptional and post-translational regulation
- **Part 5.3: Recombinant DNA Technology**
  - Sub Part 5.3.1: Restriction enzymes and DNA cloning
  - Sub Part 5.3.2: PCR and gel electrophoresis
  - Sub Part 5.3.3: DNA sequencing
- **Part 5.4: Genetic Engineering and Applications**
  - Sub Part 5.4.1: Gene cloning and expression systems
  - Sub Part 5.4.2: CRISPR-Cas9 genome editing
  - Sub Part 5.4.3: Applications of DNA technology

Introduction

The central dogma of molecular biology describes the flow of genetic information from DNA through RNA to protein, and the regulation of this flow at every step determines which proteins a cell produces, when, and in what quantity. Understanding how genes are expressed and how that expression is regulated is central to understanding development, physiology, and disease. DNA technology builds on this molecular understanding to enable the deliberate manipulation of genetic information for research, medicine, and agriculture.

This final Series covers transcription and translation (the two steps of gene expression producing proteins); prokaryotic and eukaryotic gene regulation; and the principal tools of recombinant DNA technology, including restriction enzymes, cloning, PCR, DNA sequencing, CRISPR-Cas9 genome editing, and the major applications of genetic engineering. This Series completes the course by connecting molecular genetics to its biotechnological applications.

Series Learning Outcomes

Upon completing this Series, you should be able to:

- **SLO 5.1** describe transcription, the genetic code, and translation as the steps of protein synthesis,
- **SLO 5.2** explain gene regulation in prokaryotes (lac operon) and eukaryotes, including post-transcriptional regulation,
- **SLO 5.3** describe restriction enzymes, DNA cloning, PCR, gel electrophoresis, and DNA sequencing, and
- **SLO 5.4** explain gene cloning and expression systems, CRISPR-Cas9 genome editing, and the applications of DNA technology.

5.1 Protein Synthesis

5.1.1 Transcription

Transcription is the synthesis of an RNA molecule using a DNA strand as a template, catalysed by RNA polymerase; the process begins when RNA polymerase binds to a specific DNA sequence called the promoter, typically located upstream (5) of the gene; in prokaryotes, the sigma factor subunit of RNA polymerase recognises the promoter consensus sequences at -10 and -35 positions; in eukaryotes, transcription factors bind to promoter elements (including the TATA box at approximately -25) and recruit RNA polymerase II (for mRNA synthesis) to the transcription start site.

RNA polymerase unwinds the DNA double helix, reads the template strand in the 3 to 5 direction, and synthesises the RNA transcript in the 5 to 3 direction, incorporating ribonucleoside triphosphates complementary to the template (with uracil pairing with adenine in the template); transcription continues until RNA polymerase reaches a termination signal; in eukaryotes, the primary transcript (pre-mRNA) undergoes processing (5-cap addition, 3-poly(A) tail addition, and splicing) before export to the cytoplasm for translation.

Fill in the blank: RNA polymerase reads the DNA template strand in the 3 to 5 direction and synthesises the RNA transcript in the _____ to 3 direction.

5.1.2 The genetic code

The genetic code is the set of rules by which nucleotide triplets (codons) in mRNA specify the amino acids incorporated into a protein; there are 64 possible codons (4 possible bases at each of 3 positions, giving 4^3 = 64 combinations) but only 20 amino acids, so most amino acids are specified by more than one codon (the code is degenerate); the degeneracy is largely concentrated at the third codon position (wobble position), where alternative bases often specify the same amino acid.

Three codons (UAA, UAG, and UGA) are stop codons that signal termination of translation rather than specifying an amino acid; AUG (methionine) serves as the universal start codon and

also codes for internal methionine residues; the genetic code is essentially universal (the same codons specify the same amino acids in virtually all organisms, from bacteria to humans, with minor exceptions in mitochondria and some unicellular organisms), reflecting the common ancestry of all life and enabling cross-species recombinant DNA expression.

Fill in the blank: The genetic code is _____ , meaning most amino acids are specified by more than one codon, with variation concentrated at the third (wobble) codon position.

5.1.3 Translation

Translation is the synthesis of a polypeptide chain by ribosomes using the codon sequence of an mRNA as a template; ribosomes consist of a small subunit (30S in prokaryotes, 40S in eukaryotes) and a large subunit (50S prokaryotic, 60S eukaryotic), with the small subunit binding the mRNA and the large subunit catalysing peptide bond formation; translation proceeds in three stages: initiation (assembly of the ribosome at the start codon AUG with the initiator Met-tRNA in the P site), elongation (sequential addition of amino acids as tRNAs with matching anticodons enter the A site, peptide bond is formed, and the ribosome translocates one codon), and termination (stop codon in A site recognised by release factors, polypeptide released).

The ribosome has three tRNA-binding sites: the A (aminoacyl) site receives new aminoacyl-tRNAs, the P (peptidyl) site holds the growing polypeptide chain attached to a tRNA, and the E (exit) site releases the uncharged tRNA; the peptidyl transferase activity (carried out by the 23S/28S rRNA of the large subunit) catalyses formation of the peptide bond between the amino acid in the A site and the growing chain in the P site; multiple ribosomes can translate the same mRNA simultaneously, forming a polysome.

Fill in the blank: The ribosome A site receives new aminoacyl-tRNAs, the P site holds the _____ -tRNA complex, and the E site releases the uncharged tRNA after translocation.

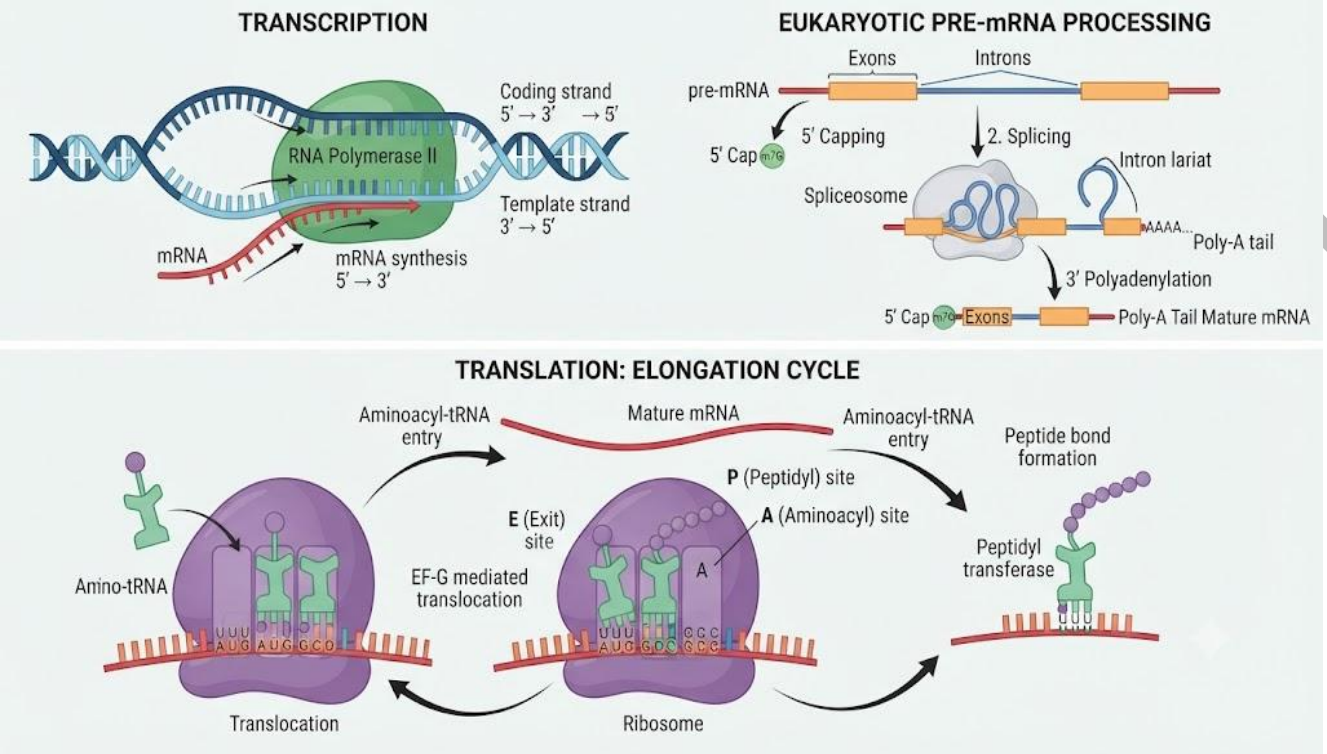


Figure 5.1: Transcription and translation

Pilot questions 5.1

1. Describe the role of the promoter in transcription.
2. Explain what is meant by the degeneracy of the genetic code.
3. State the three stop codons and the universal start codon.
4. Describe the three stages of translation.
5. Name the three tRNA-binding sites of the ribosome and state the function of each.

5.2 Regulation of Gene Expression

5.2.1 The lac operon in prokaryotes

The lac operon (Jacob and Monod, 1961) is the paradigm model of prokaryotic gene regulation; the operon consists of three structural genes (lacZ encoding beta-galactosidase, lacY encoding lactose permease, and lacA encoding thiogalactoside transacetylase) transcribed as a single

polycistronic mRNA, along with a promoter and an operator (a sequence overlapping the promoter to which the lac repressor protein binds to block transcription); the lac repressor is encoded by the regulatory gene lacI, expressed constitutively from its own promoter.

The lac operon is regulated by both negative control (repression by the lac repressor) and positive control (catabolite repression/catabolite activator protein/CAP regulation): in the absence of lactose, the repressor binds the operator and blocks RNA polymerase; when lactose is present, allolactose (a metabolic derivative) binds the repressor as an inducer, causing conformational change and release from the operator, allowing transcription; when glucose is absent and cAMP levels are high, cAMP binds the CAP protein which then binds the CAP site upstream of the promoter, stimulating RNA polymerase binding and transcription; the lac operon is thus fully active only when lactose is present and glucose is absent.

Fill in the blank: In the lac operon, _____ acts as an inducer by binding the lac repressor and causing it to release from the operator, allowing transcription.

5.2.2 Eukaryotic gene regulation

Eukaryotic gene regulation is considerably more complex than prokaryotic regulation, reflecting the greater complexity of eukaryotic cells, the separation of transcription (in the nucleus) from translation (in the cytoplasm), and the organisation of eukaryotic DNA into chromatin; gene expression in eukaryotes can be regulated at multiple levels: chromatin remodelling (altering nucleosome positioning and histone modification to make DNA more or less accessible to transcription factors and RNA polymerase), DNA methylation (methylation of cytosine in CpG dinucleotides, generally associated with gene silencing), and transcription factor binding.

Transcription factors are proteins that bind to specific DNA sequences in promoters and enhancers to activate or repress transcription; enhancers (cis-regulatory elements that can act at great distances from the gene they regulate, interacting with the promoter through DNA looping) allow tissue-specific and developmental stage-specific gene expression; combinatorial regulation (the combined action of multiple transcription factors, some activating and some repressing) allows a relatively small number of transcription factors to generate the enormous diversity of cell-type-specific gene expression patterns seen in complex organisms.

Fill in the blank: _____ are cis-regulatory DNA elements that can act at great distances from the gene they regulate, interacting with the promoter through DNA looping to activate transcription.

5.2.3 Post-transcriptional and post-translational regulation

Post-transcriptional regulation controls which mRNAs are translated, at what rate, and for how long; RNA splicing (alternative splicing of pre-mRNA to produce different mRNA isoforms from the same gene, greatly expanding proteome diversity without expanding gene number) is a major mechanism; mRNA stability control (sequences in the 5-UTR and 3-UTR of mRNA affect its half-life, allowing rapid changes in protein levels in response to signals without changing transcription rates); and microRNA (miRNA) and small interfering RNA (siRNA) mechanisms post-transcriptionally silence specific target mRNAs by degradation or translational repression via the RNA-induced silencing complex (RISC).

Post-translational regulation modifies protein activity, stability, or localisation after synthesis: phosphorylation (addition of phosphate groups by kinases, often activating or inactivating enzymes), ubiquitination (attachment of ubiquitin chains targeting proteins for proteasomal degradation), glycosylation (addition of carbohydrate groups affecting protein folding and cell-surface recognition), and proteolytic cleavage (removal of signal sequences or activation of zymogens) are all important mechanisms; the combination of transcriptional, post-transcriptional, and post-translational regulation allows cells to respond rapidly and precisely to changing conditions.

Fill in the blank: Alternative _____ of pre-mRNA produces different mRNA isoforms from the same gene, greatly expanding proteome diversity without requiring additional genes.

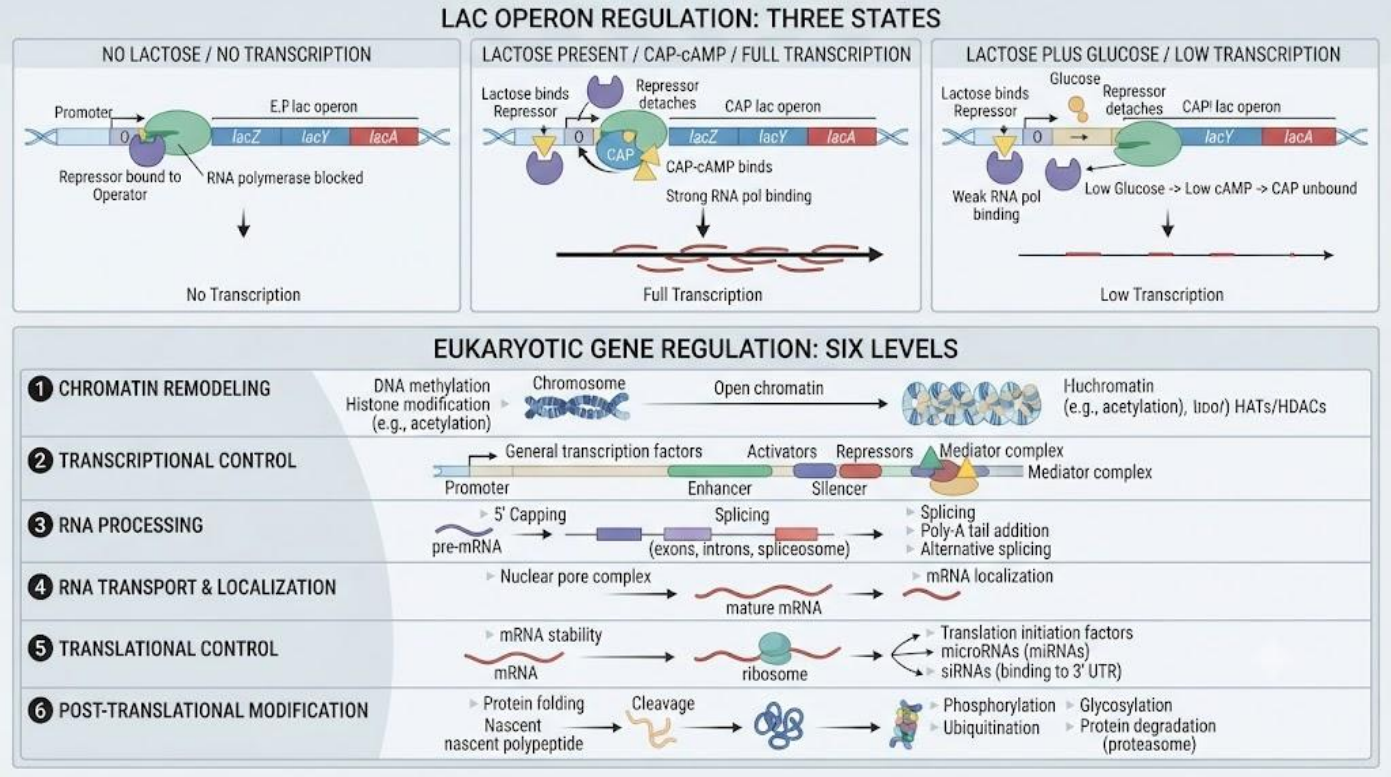


Figure 5.2: Gene regulation in prokaryotes and eukaryotes

Pilot questions 5.2

1. Describe the role of the lac repressor in the lac operon.
2. Explain the role of CAP and cAMP in lac operon regulation.
3. Name three levels at which eukaryotic gene expression can be regulated.
4. Explain how enhancers regulate gene expression despite being distant from the gene.
5. Explain how miRNA regulates gene expression.

5.3 Recombinant DNA Technology

5.3.1 Restriction enzymes and DNA cloning

Restriction endonucleases (restriction enzymes) are bacterial enzymes that cut double-stranded DNA at specific short palindromic recognition sequences (typically 4 to 8 base pairs); Type II restriction enzymes cut within or near the recognition sequence and are the principal tools of

recombinant DNA technology; many restriction enzymes make staggered cuts producing short single-stranded overhangs called sticky ends (cohesive ends) that can base-pair with complementary sticky ends from other DNA fragments cut with the same enzyme, facilitating directional ligation.

DNA cloning involves inserting a DNA fragment of interest into a vector (typically a plasmid, bacteriophage, or artificial chromosome), which is then introduced into a host cell (usually E. coli) that replicates the vector and insert together; the key steps are: cutting both the insert DNA and the vector with the same restriction enzyme(s); ligating the compatible ends using DNA ligase; transforming the recombinant vector into competent host cells; and selecting for cells that have taken up the vector using an antibiotic resistance marker or other selectable marker on the vector.

Fill in the blank: Restriction enzymes that make staggered cuts produce single-stranded overhangs called _____ ends that can base-pair with complementary ends from other fragments cut with the same enzyme.

5.3.2 PCR and gel electrophoresis

The polymerase chain reaction (PCR, developed by Kary Mullis, 1983) is an in vitro technique for exponential amplification of a specific DNA sequence using repeated cycles of denaturation (heating to approximately 95 degrees C to separate the double helix), annealing (cooling to approximately 50 to 65 degrees C allowing short synthetic oligonucleotide primers to bind specifically to complementary sequences flanking the target), and extension (heating to approximately 72 degrees C for thermostable DNA polymerase such as Taq polymerase to synthesise new strands from the primers); each cycle approximately doubles the number of target DNA copies, producing millions of copies after 25 to 35 cycles.

Gel electrophoresis separates DNA fragments by size: DNA is loaded into wells in an agarose gel and an electric current applied; negatively charged DNA molecules migrate toward the positive electrode, with smaller fragments moving faster through the gel matrix and thus travelling further in a given time; fragments are visualised by staining with a fluorescent dye (ethidium bromide or safer alternatives such as SYBR Green) and their sizes estimated by

comparison with a DNA ladder (a set of fragments of known size run alongside the samples); gel electrophoresis is used to verify PCR products, analyse restriction digestion patterns, and separate DNA fragments for further analysis.

Fill in the blank: In PCR, each cycle of denaturation, annealing, and extension approximately _____ the number of target DNA copies, producing millions of copies after 25 to 35 cycles.

5.3.3 DNA sequencing

Sanger (chain termination) sequencing (developed by Frederick Sanger, 1977) uses dideoxynucleoside triphosphates (ddNTPs), which lack the 3-OH group required for chain elongation, to randomly terminate DNA synthesis at each position; by including a small proportion of a specific ddNTP in a synthesis reaction, a population of fragments terminating at every occurrence of that base is produced; the size of each terminated fragment indicates the position of that base in the sequence, and reading the ladder of fragments (originally by gel electrophoresis, now by capillary electrophoresis with fluorescently labelled ddNTPs) gives the sequence.

Next-generation sequencing (NGS) platforms have massively increased sequencing throughput and reduced cost, enabling whole-genome sequencing of complex organisms routinely; NGS approaches (including Illumina sequencing by synthesis) generate millions of short reads in parallel, which are computationally assembled into a complete genome sequence; long-read technologies (such as Oxford Nanopore and PacBio) sequence single DNA molecules directly, producing long reads that better resolve complex repetitive regions; genome sequencing has transformed genetics, enabling genome-wide association studies (GWAS), personal genomics, and comprehensive mutation analysis.

Fill in the blank: Sanger sequencing uses _____ triphosphates (ddNTPs) that lack the 3-OH group to randomly terminate chain elongation at each occurrence of a specific base.

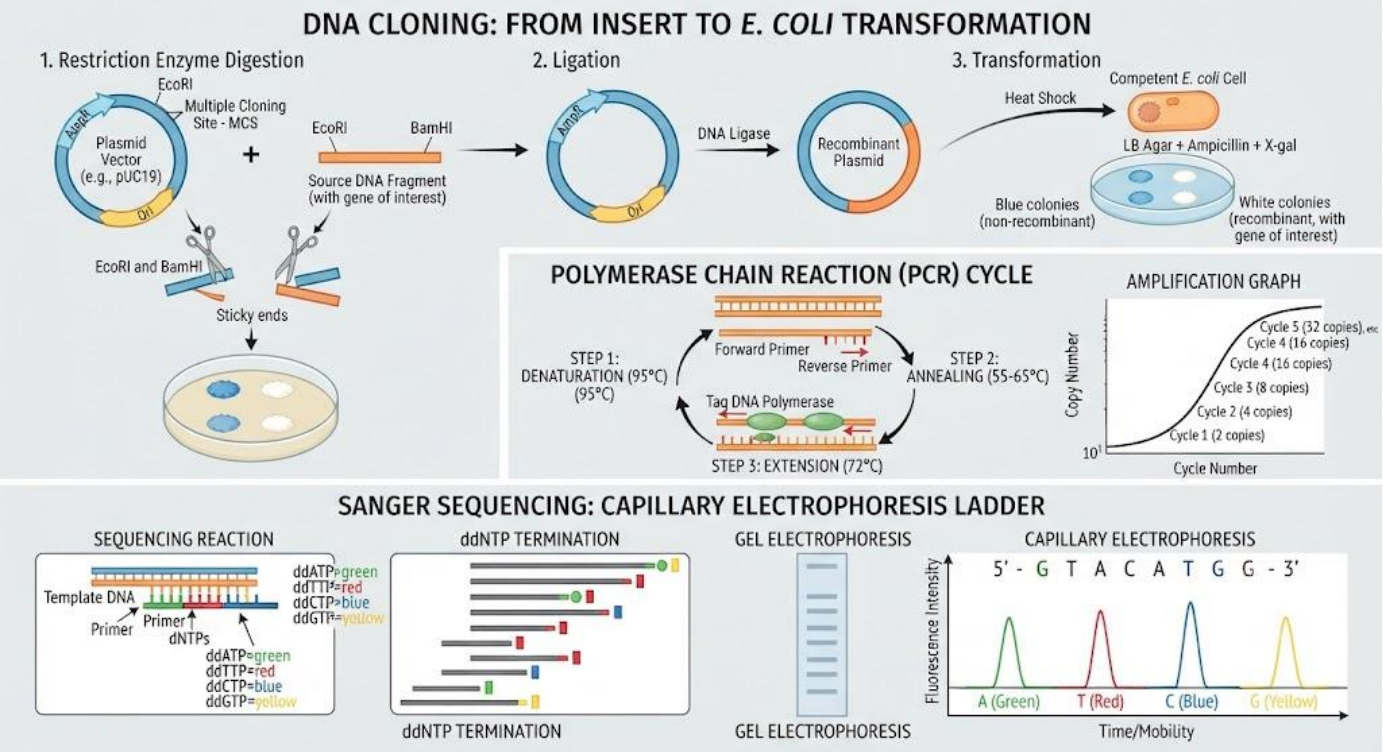


Figure 5.3: Recombinant DNA technology

Pilot questions 5.3

1. Explain what a restriction enzyme does and what sticky ends are.
2. Describe the steps of DNA cloning into a plasmid vector.
3. Describe the three steps of one PCR cycle and explain why the target sequence is amplified exponentially.
4. Explain how gel electrophoresis separates DNA fragments.
5. Explain how Sanger sequencing uses ddNTPs to determine a DNA sequence.

5.4 Genetic Engineering and Applications

5.4.1 Gene cloning and expression systems

Gene cloning (molecular cloning) involves isolating a gene of interest and inserting it into an expression vector designed to allow high-level transcription and translation of the cloned gene in

a suitable host cell; expression vectors contain strong promoters, ribosome binding sites (in prokaryotes), Kozak sequences (in eukaryotes), and often sequences encoding purification tags (such as a histidine tag allowing purification by nickel-affinity chromatography) fused to the gene of interest; prokaryotic expression systems (E. coli) are rapid and inexpensive but cannot perform the post-translational modifications (glycosylation, disulfide bond formation) required by many mammalian proteins.

Eukaryotic expression systems (yeast, insect cells via baculovirus, and mammalian cell culture) allow post-translational modifications and correct protein folding for complex eukaryotic proteins; transgenic organisms (bacteria, plants, or animals engineered to carry and express foreign genes) are produced by introducing recombinant expression constructs into the organism; recombinant DNA technology has enabled the production of important medical proteins including insulin (produced in E. coli since 1982), growth hormone, erythropoietin, and blood clotting factors, replacing scarce or potentially contaminated human-derived products.

Fill in the blank: Prokaryotic expression systems are rapid but cannot perform _____ modifications such as glycosylation required by many mammalian proteins.

5.4.2 CRISPR-Cas9 genome editing

CRISPR-Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats, associated protein 9) is a bacterial adaptive immune system repurposed as a precise genome editing tool; the system requires two components: the Cas9 nuclease (a DNA-cutting enzyme) and a single guide RNA (sgRNA) that directs Cas9 to a specific genomic target complementary to a 20-nucleotide sequence in the sgRNA followed by a short NGG protospacer adjacent motif (PAM) sequence; Cas9 binds the sgRNA and cuts both strands of the target DNA, producing a double-strand break at the desired genomic location.

The double-strand break introduced by CRISPR-Cas9 is repaired by the cell using either NHEJ (which introduces small insertions or deletions that can disrupt gene function, useful for gene knockout) or homology-directed repair (HDR, using a supplied donor template to introduce a specific sequence change, enabling precise gene correction or insertion); CRISPR-Cas9 has transformed biological research and holds promise for therapeutic genome editing of inherited

diseases including sickle cell disease, beta-thalassaemia, and some forms of blindness; it was recognised with the 2020 Nobel Prize in Chemistry awarded to Jennifer Doudna and Emmanuelle Charpentier.

Fill in the blank: CRISPR-Cas9 requires a single guide RNA (sgRNA) that directs the Cas9 _____ to cut both strands of DNA at a target sequence complementary to the 20-nucleotide guide sequence.

5.4.3 Applications of DNA technology

DNA technology has transformed medicine, agriculture, and forensic science: in medicine, recombinant proteins (insulin, clotting factors, vaccines including hepatitis B subunit vaccine produced in yeast), gene therapy (delivery of functional gene copies to correct hereditary deficiency), pharmacogenomics (tailoring drug treatment to an individual genomic profile), and cancer genomics (identifying driver mutations to guide targeted therapy) are all direct products of recombinant DNA and sequencing technology; DNA diagnostics (PCR-based detection of pathogens and genetic mutations) enable rapid, sensitive, and specific disease diagnosis.

In agriculture, genetically modified organisms (GMOs) have been developed with traits including herbicide tolerance (glyphosate-resistant crops carrying a bacterial EPSPS gene), insect resistance (Bt crops expressing the Bacillus thuringiensis delta-endotoxin), delayed fruit ripening, and improved nutritional profiles (Golden Rice engineered to produce beta-carotene); in forensics, DNA profiling (using variable number tandem repeats/microsatellites or single nucleotide polymorphisms) provides highly discriminatory individual identification for criminal investigation, paternity testing, and identification of disaster victims; ethical debates around GMOs, gene therapy, and personal genomics accompany the rapid expansion of DNA technology applications.

Fill in the blank: DNA _____ uses variable number tandem repeats or single nucleotide polymorphisms to provide individual identification for forensic investigation and paternity testing.

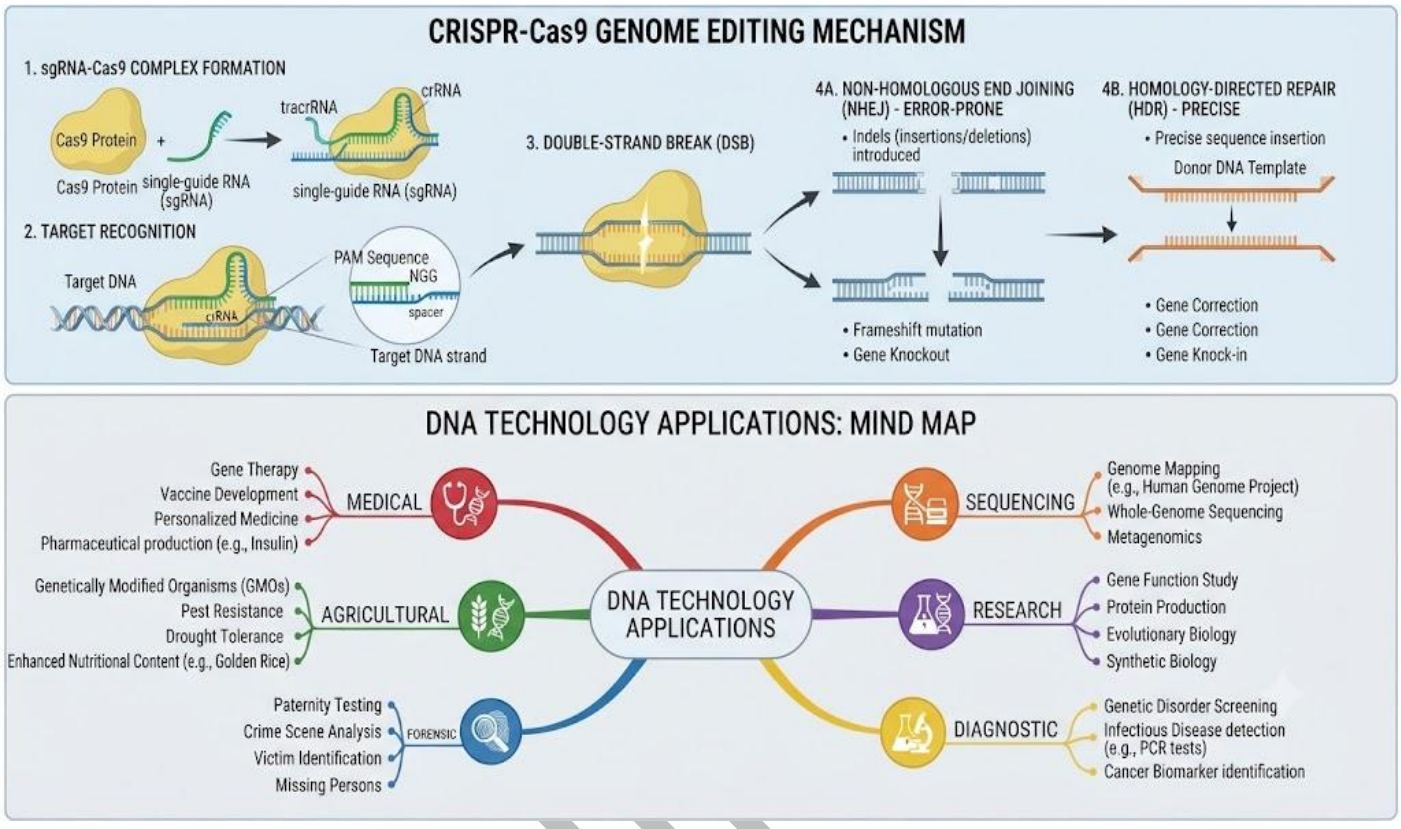


Figure 5.4: Genetic engineering and CRISPR-Cas9

Pilot questions 5.4

1. Explain what an expression vector is and name two features that distinguish it from a standard cloning vector.
2. Give two reasons why eukaryotic expression systems may be preferred over E. coli for producing some proteins.
3. Describe the two components of the CRISPR-Cas9 system and explain how they work together.
4. Explain how CRISPR-Cas9 can be used to knock out a gene.
5. Name three fields in which DNA technology has been applied and give one specific application in each.

Series Summary

This Series explored **gene expression and modern genetic technologies**, explaining how genetic information is transcribed, translated, and regulated to produce functional proteins. You examined the stages of transcription and translation, the features of the genetic code, and the mechanisms that regulate gene expression in both prokaryotes and eukaryotes, including chromatin modification, transcriptional control, post-transcriptional processing, and post-translational modification.

The Series also introduced the principles and applications of recombinant DNA technology, including restriction enzymes, DNA cloning, PCR, gel electrophoresis, DNA sequencing, and next-generation sequencing. In addition, you explored genetic engineering tools such as CRISPR-Cas9 and their applications in medicine, agriculture, and forensic science. By completing this Series, you should now be able to explain how genes are expressed and regulated, apply the principles of molecular genetics techniques, evaluate modern genetic engineering methods, and relate these technologies to real-world biological applications, thereby achieving all four **Series Learning Outcomes (SLOs)** and the **overall course outcomes**.

Glossary of Terms

- **CRISPR-Cas9:** a genome editing system using a single guide RNA to direct the Cas9 nuclease to cut a specific DNA sequence.
- **Expression vector:** a cloning vector designed to drive high-level transcription and translation of an inserted gene in a host cell.
- **Genetic code:** the set of rules by which mRNA codons specify amino acids during translation.
- **Lac operon:** the paradigm prokaryotic gene regulation system encoding lactose metabolism genes controlled by a repressor and catabolite activator protein.
- **Polymerase chain reaction (PCR):** an in vitro technique for exponential amplification of a specific DNA sequence.

- **Restriction endonuclease:** a bacterial enzyme that cuts double-stranded DNA at a specific palindromic recognition sequence.
- **sgRNA (single guide RNA):** the guide RNA component of CRISPR-Cas9 containing a 20-nucleotide sequence that directs Cas9 to its genomic target.

Concept Check Questions [Series 5]

Objective Questions

1. RNA polymerase reads the DNA template strand 3 to 5 and synthesises RNA in the _____ to 3 direction. (Linked to SLO 5.1)
2. The genetic code is _____ , meaning most amino acids are specified by more than one codon. (Linked to SLO 5.1)
3. In the lac operon, allolactose binds the repressor and causes it to _____ from the operator. (Linked to SLO 5.2)
4. Restriction enzymes that make staggered cuts produce single-stranded _____ ends. (Linked to SLO 5.3)
5. CRISPR-Cas9 requires a single guide RNA and the Cas9 _____ to cut both DNA strands at the target. (Linked to SLO 5.4)

Short-Answer Questions

1. Name the three stages of translation. (Linked to SLO 5.1)
2. Explain the role of CAP and cAMP in lac operon positive control. (Linked to SLO 5.2)
3. Describe the steps of DNA cloning into a plasmid. (Linked to SLO 5.3)
4. Describe the three steps of one PCR cycle. (Linked to SLO 5.3)
5. Explain how CRISPR-Cas9 can be used to knock out a gene. (Linked to SLO 5.4)

Long-Answer Questions

1. Describe transcription, the genetic code, and translation. (Linked to SLO 5.1)
2. Explain gene regulation in prokaryotes and eukaryotes. (Linked to SLO 5.2)
3. Describe restriction enzymes, PCR, Sanger sequencing, CRISPR-Cas9, and applications of DNA technology. (Linked to SLOs 5.3, 5.4)

Answers to Concept Check Questions [Series 5]

Objective Questions

1. 5.
2. Degenerate.
3. Dissociate (release).
4. Sticky.
5. Nuclease.

Short-Answer Questions

1. Initiation, elongation, and termination.
2. When glucose is absent, cAMP levels rise; cAMP binds the CAP protein which then binds the CAP site upstream of the lac promoter, stimulating RNA polymerase binding and increasing transcription.
3. Cut vector and insert with same restriction enzyme; ligate with DNA ligase; transform into E. coli; select on antibiotic plates for cells carrying the vector.
4. Denaturation (95 degrees C, strands separate), annealing (50-65 degrees C, primers bind flanking sequences), extension (72 degrees C, Taq polymerase synthesises new strand).
5. Design an sgRNA targeting the gene; deliver sgRNA and Cas9 to the cell; Cas9 cuts both strands at the target; NHEJ repair introduces small indels that disrupt the reading frame and knock out gene function.

Long-Answer Questions

1. Transcription: RNA polymerase binds promoter, reads template 3-to-5, synthesises RNA 5-to-3; eukaryotic pre-mRNA is processed (5-cap, poly-A tail, splicing). Genetic code: 64 codons for 20 amino acids, degenerate (wobble position), universal, AUG start, UAA/UAG/UGA stop. Translation: initiation (ribosome assembles at AUG with Met-tRNA), elongation (aminoacyl-tRNA enters A site, peptide bond formed, ribosome translocates), termination (stop codon, release factors, polypeptide released).
2. Lac operon: negative control (repressor binds operator blocking transcription, allolactose inducer releases repressor); positive control (CAP-cAMP complex binds CAP site, stimulates transcription when glucose is absent). Eukaryotic: chromatin remodelling, DNA methylation,

transcription factors, enhancers (DNA looping); post-transcriptional (alternative splicing, mRNA stability, miRNA/RISC); post-translational (phosphorylation, ubiquitination).

3. Restriction enzymes cut at palindromic sequences producing sticky ends for ligation. PCR amplifies specific sequences exponentially using denaturation/annealing/extension cycles and Taq polymerase. Sanger sequencing uses ddNTPs for chain termination to determine sequence. CRISPR-Cas9 (sgRNA + Cas9 nuclease) cuts DNA at a specified target; NHEJ causes gene knockout; HDR enables gene correction. Applications: medicine (recombinant insulin, gene therapy), agriculture (Bt crops, herbicide tolerance), forensics (DNA profiling).

Answers to Pilot Questions [Series 5]

Part 5.1

1. The promoter is a specific DNA sequence where RNA polymerase binds to initiate transcription, determining where and when transcription begins.
2. The code is degenerate because 64 codons specify only 20 amino acids, so most amino acids are encoded by more than one codon.
3. Stop codons: UAA, UAG, UGA. Start codon: AUG (methionine).
4. Initiation: ribosome assembles at AUG with initiator Met-tRNA in the P site. Elongation: aminoacyl-tRNAs enter A site, peptide bond formed, ribosome translocates one codon. Termination: stop codon in A site recognised by release factors, polypeptide released.
5. A site: receives incoming aminoacyl-tRNA. P site: holds the growing peptidyl-tRNA. E site: releases the uncharged tRNA after translocation.

Part 5.2

1. In the absence of lactose, the lac repressor binds the operator, blocking RNA polymerase from transcribing the structural genes.
2. When glucose is absent, cAMP levels rise; cAMP binds CAP; CAP-cAMP complex binds the CAP site upstream of the promoter, stimulating RNA polymerase binding and increasing transcription of the lac operon.
3. Chromatin remodelling, DNA methylation, and transcription factor binding (or enhancer action, alternative splicing, mRNA stability, miRNA, phosphorylation/ubiquitination).

4. Enhancers loop the DNA to bring themselves into contact with the promoter; transcription factors bound to the enhancer then interact with the transcription initiation complex at the promoter to stimulate transcription.
5. miRNA is incorporated into the RISC complex; RISC binds to complementary sequences in target mRNA 3-UTR; binding leads to target mRNA degradation or translational repression.

Part 5.3

1. A restriction enzyme recognises a specific palindromic sequence and cuts both DNA strands; staggered cuts produce sticky ends, single-stranded overhangs that can base-pair with complementary sticky ends from another fragment cut with the same enzyme.
2. Cut vector and insert with the same restriction enzyme; ligate compatible ends with DNA ligase; transform recombinant plasmid into E. coli; select colonies on antibiotic medium; screen for correct insert.
3. Denaturation (95 degrees C, strands separate), annealing (55 degrees C, primers bind specific target sequences), extension (72 degrees C, Taq polymerase synthesises new strands). Each cycle approximately doubles the target, giving exponential amplification.
4. DNA is negatively charged; applied electric field drives it toward the positive electrode through agarose gel; smaller fragments migrate faster and further; sizes are estimated by comparison with a DNA ladder of known sizes.
5. ddNTPs lack a 3-OH group so chain extension stops whenever a ddNTP is incorporated; by using a small proportion of one fluorescently labelled ddNTP, all fragments terminating at that base are produced; the length of each fragment corresponds to the position of that base in the sequence.

Part 5.4

1. An expression vector is designed to drive high-level transcription and translation of an inserted gene; it contains a strong promoter and, in many cases, a purification tag sequence, distinguishing it from simple cloning vectors.
2. Eukaryotic systems perform post-translational modifications such as glycosylation and correct disulfide bond formation needed by many mammalian proteins, and allow correct folding of complex proteins.

3. CRISPR-Cas9 requires: (1) the Cas9 nuclease and (2) a single guide RNA (sgRNA) with a 20-nt sequence complementary to the target; the sgRNA directs Cas9 to the genomic target where it cuts both DNA strands three bases upstream of the PAM sequence.
4. Design an sgRNA targeting the gene of interest; deliver Cas9 and sgRNA to cells; Cas9 creates a double-strand break; NHEJ repair introduces insertions or deletions (indels) that disrupt the reading frame and eliminate gene function.
5. Medicine: recombinant insulin production in E. coli. Agriculture: Bt crops expressing insecticidal toxin. Forensics: DNA profiling using STR loci for individual identification.

References and Attribution

Textual References (Books and External Sources)

- Watson, J. D., Baker, T. A., Bell, S. P., Gann, A., Levine, M., & Losick, R. (2014). Molecular biology of the gene (7th ed.). Cold Spring Harbor Laboratory Press.
- Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., & Walter, P. (2014). Molecular biology of the cell (6th ed.). Garland Science.
- Griffiths, A. J. F., Wessler, S. R., Carroll, S. B., & Doebley, J. (2015). Introduction to genetic analysis (11th ed.). W. H. Freeman.
- National Open University of Nigeria (NOUN). (2023). Genetics II course material. NOUN Press.

Figures, Plates, Tables, and Boxes

- Figure 5.1: Transcription and translation Attribution: AI generated. Suggested OER search string: "transcription RNA polymerase promoter translation ribosome A P E site elongation genetic code codon anticodon diagram OER".
- Figure 5.2: Gene regulation in prokaryotes and eukaryotes Attribution: AI generated. Suggested OER search string: "lac operon gene regulation prokaryote eukaryote transcription factor enhancer alternative splicing miRNA post-translational diagram genetics OER".
- Figure 5.3: Recombinant DNA technology Attribution: AI generated. Suggested OER search string: "DNA cloning restriction enzyme plasmid PCR gel electrophoresis Sanger sequencing dideoxy ddNTP diagram genetics OER".
- Figure 5.4: Genetic engineering and CRISPR-Cas9 Attribution: AI generated. Suggested OER search string: "CRISPR Cas9 genome editing sgRNA PAM DNA technology applications GMO forensics gene therapy medicine diagram genetics OER".