Unit 3

Genetics

This unit online, with its interactives:
science-playground.com/s/10a3

About one boy in every twelve in some countries can't tell red from green the way his friends do. A red apple among green leaves looks to him like part of the tree, and a traffic light is read by its place, top or bottom, not its color. Yet among girls in the same countries, only about one in two hundred is color-blind in that way. The same is true of hemophilia, the disease in which blood doesn't clot: almost everyone who has it is a boy.

Why should one sex suffer so much more than the other? The answer needs three ideas, and this unit builds all three: Mendel's rules for following two characters at once, the chemistry of DNA and RNA in which every gene is written, and the X chromosome, of which a boy has only one.

3.1 Mendel's law: law of independent assortment

In a **monohybrid cross**, two pure parents differ in one pair of characters, and the F_2 comes out 3 : 1. Mendel didn't stop there. He wanted to know what happens when parents differ in **two** characters at once. Do the two travel together, as a package from each grandparent, or does each go its own way?

A cross between two pure parents that differ in two pairs of contrasting characters is a **dihybrid cross**. Mendel's best-known one used two characters of the pea seed:

Character	Dominant	Recessive
Seed shape	round (R)	wrinkled (r)
Seed color	yellow (Y)	green (y)

Mendel's dihybrid cross

Parents (P). Mendel crossed plants pure for round yellow seeds (**RRYY**) with plants pure for wrinkled green seeds (**rryy**). A gamete gets one allele of each gene, so every gamete of the first parent is **RY** and every gamete of the second is **ry**.

First filial generation (F_1). Every seed was **RrYy**, and every one was round and yellow: both dominant characters showed, just as each would in a monohybrid cross.

Gametes of F_1. Here is the key step. An RrYy plant puts R or r into a gamete, and Y or y, and the two choices don't depend on each other. So it makes **four kinds of gametes**, **RY**, **Ry**, **rY** and **ry**, in equal numbers, a quarter each.

Second filial generation (F_2). Mendel let the F_1 plants self-pollinate. Each of the four kinds of pollen can meet each of the four kinds of egg, so there are $4 \times 4 = 16$ equally likely meetings. In his F_2, Mendel counted 556 seeds:

Phenotype	Seeds counted	Expected out of 556 (9 : 3 : 3 : 1)
Round yellow	315	312.75
Round green	108	104.25
Wrinkled yellow	101	104.25
Wrinkled green	32	34.75
Total	556	556

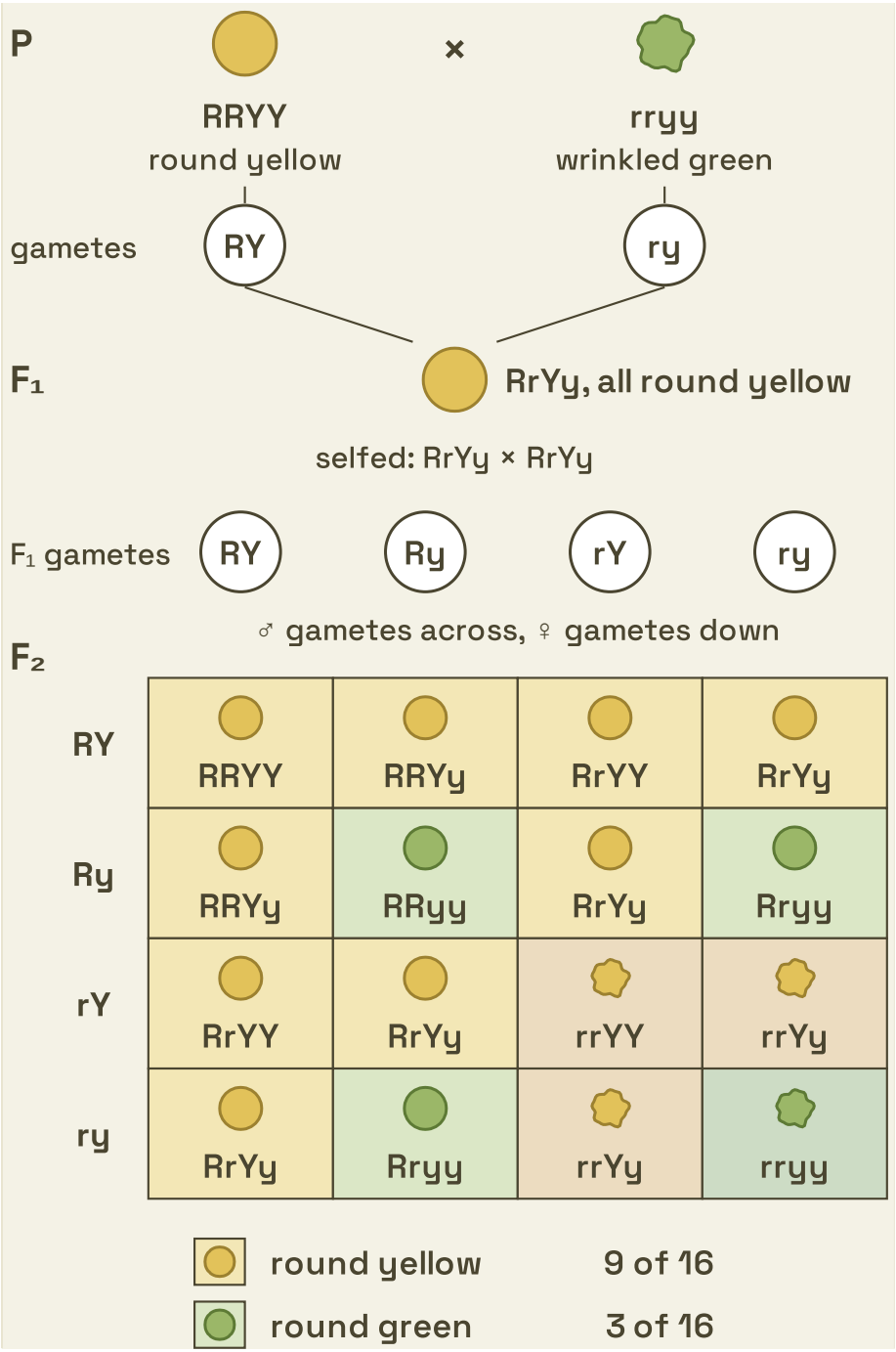
The expected numbers are $556 \times 9/16$, $556 \times 3/16$ and $556 \times 1/16$. Mendel's counts are close to them: the F_2 phenotypic ratio is **9 : 3 : 3 : 1**. Two of the four kinds, **round green** and **wrinkled yellow**, were in neither grandparent. They are **new combinations** (also called **recombinants**), and they could only appear if shape and color had separated.

Phenotypic chart of the dihybrid cross

The 16 meetings are set out in a **checkerboard** (a Punnett square): the F_1 male gametes across the top, the female gametes down the side, and in each box the genotype of the seed made where they meet. This chart, with the generations above it, is how the cross is drawn in an exam.

Read the checkerboard box by box and you find **nine different genotypes**:

Phenotype	Genotypes (boxes out of 16)	Total
Round yellow	RRYY (1), RRYy (2), RrYY (2), RrYy (4)	9
Round green	RRyy (1), Rryy (2)	3
Wrinkled yellow	rrYY (1), rrYy (2)	3
Wrinkled green	rryy (1)	1


P
RRYY
round yellow
×
rryy
wrinkled green
gametes
RY
ry
F₁
RrYy, all round yellow
selfed: RrYy × RrYy
F₁ gametes
RY
Ry
rY
ry
♂ gametes across, ♀ gametes down
F₂
RY
Ry
rY
ry
RRYY
RRYy
RrYY
RrYy
RRYy
RRyy
RrYy
Rryy
RrYY
RrYy
rrYY
rrYy
RrYy
Rryy
rrYy
rryy
round yellow
9 of 16
round green
3 of 16

So the F_2 **genotypic ratio** is RRYY : RRYy : RrYY : RrYy : RRyy : Rryy : rrYY : rrYy : rryy = **1 : 2 : 2 : 4 : 1 : 2 : 1 : 2 : 1**, sixteen boxes in all. Notice the pattern: the four genotypes pure for both genes (RRYY, RRyy, rrYY, rryy) appear once each; the four pure for one gene and hybrid for the other appear twice each; and RrYy, hybrid for both, appears four times. The four pure ones lie on one diagonal of the checkerboard, and RrYy on the other.

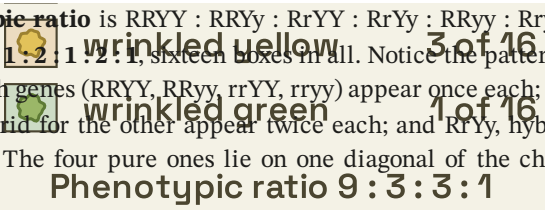


Figure 3.1 *The dihybrid cross from parents to F_2. Each box of the checkerboard is one of 16 equally likely seeds, shaded by its phenotype: round yellow (9 boxes), round green (3), wrinkled yellow (3) and wrinkled green (1).*

Statement of the law

Separate the two characters and look at each alone. Of Mendel's 556 seeds, 315 + 108 = 423 were round and 101 + 32 = 133 wrinkled: 423 ÷ 133 = 3.18, about 3 : 1. And 315 + 101 = 416 were yellow and 108 + 32 = 140 green: 416 ÷ 140 = 2.97, again about 3 : 1. Each character on its own behaves exactly as in a monohybrid cross, as if the other weren't there. Two independent 3 : 1 ratios together make

$$(3 \text{ round} : 1 \text{ wrinkled}) \times (3 \text{ yellow} : 1 \text{ green}) = 9 : 3 : 3 : 1$$

since $3 \times 3 = 9$ round yellow, $3 \times 1 = 3$ round green, $1 \times 3 = 3$ wrinkled yellow and $1 \times 1 = 1$ wrinkled green. This is **Mendel's law of independent assortment**:

When two pairs of contrasting characters are followed together, the alleles of one pair separate into the gametes independently of the alleles of the other pair. So an individual hybrid for two genes makes all four combinations of their alleles in equal numbers, and the characters can appear in new combinations in the offspring.

Check 1

How many kinds of gametes does a plant of genotype RrYy make, and in what proportion?

a two kinds, RY and ry, half each
b two kinds, Rr and Yy, half each
c four kinds, RY, Ry, rY and ry, a quarter each
d sixteen kinds, one for each box of the checkerboard

Why the law works: chromosomes in meiosis

Mendel had never seen a chromosome. We now know why his law holds. The gene for seed shape and the gene for seed color lie on **different chromosomes** of the pea. In meiosis I, each pair of homologous chromosomes lines up at the middle of the cell, and which member of a pair faces which pole is a matter of chance, decided separately for every pair. With two pairs there are two ways to line up, equally likely, and

each gives two kinds of gamete. (To watch a whole meiosis stage by stage, see 1.2 Meiosis cell division in Unit 1.)

This is the link Sutton and Boveri saw in 1902 to 1903 (Grade 9's chromosome theory of inheritance): Mendel's factors assort independently because the chromosomes that carry them do.

Verification of the law of independent assortment

A law needs a test. If the F_1 (RrYy) really makes four kinds of gametes in equal numbers, there are two ways to see them.

1. The test cross

Cross the F_1 with a plant that is **recessive for both genes** (rryy). Every gamete of rryy is **ry**, carrying only recessive alleles, so it hides nothing: each offspring shows the alleles that came from the F_1 parent's gamete.

	RY	Ry	rY	ry
ry	RrYy, round yellow	Rryy, round green	rrYy, wrinkled yellow	rryy, wrinkled green

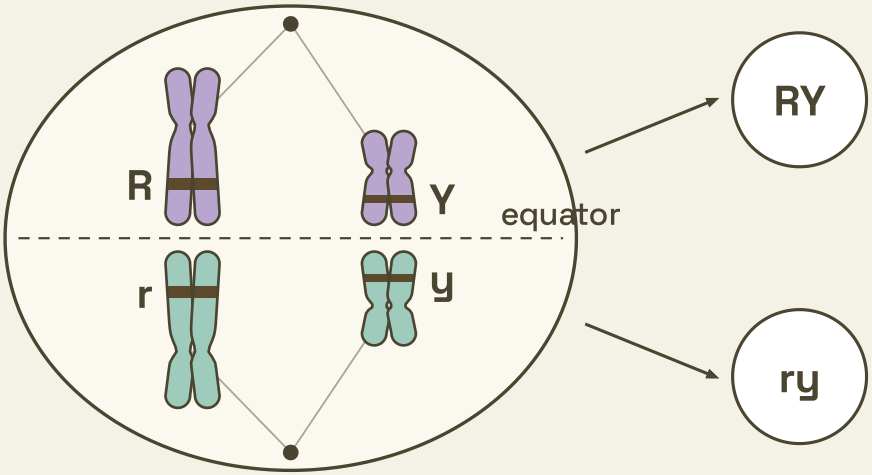
If the F_1 makes the four gametes equally, the offspring come out **1 : 1 : 1 : 1**. A cross of an individual with a parent recessive for the genes being studied is called a **test cross**; with two genes it is a **dihybrid test cross**.

Mendel did exactly this, both ways round. With the hybrid as the seed (female) parent and rryy as the pollen parent, he got 31 round yellow, 26 round green, 27 wrinkled yellow and 26 wrinkled green seeds. With rryy as the seed parent and the hybrid as the pollen parent, he got 24, 25, 22 and 27. Both are close to 1 : 1 : 1 : 1, so **both the eggs and the pollen of the hybrid were of four kinds in equal numbers**, as the law says.

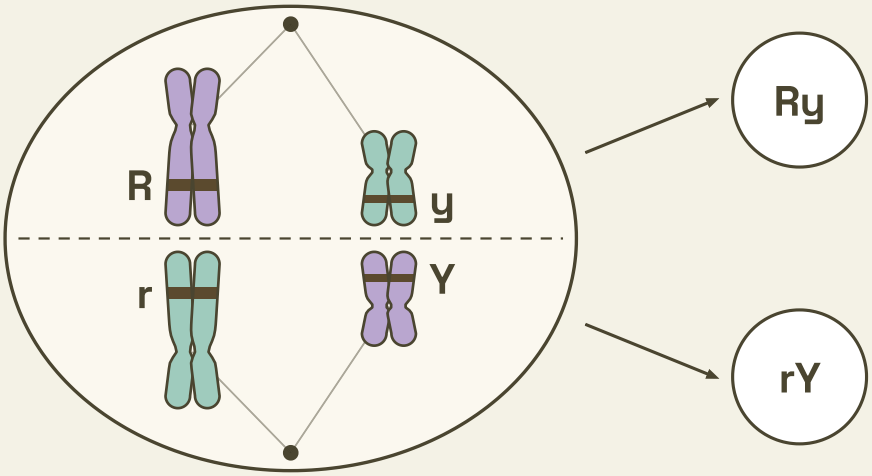
2. Growing the F_2 for one more generation

Mendel also sowed his F_2 seeds and let each plant self-pollinate. A seed pure for both genes (RRYY, say) gives seeds all of one kind; a seed hybrid for one gene gives two kinds; RrYy gives all four again. The checkerboard predicts that the pure-for-both, hybrid-for-one and hybrid-for-both seeds come in the ratio 1 : 2 : 4. From the round yellow seeds alone, Mendel found 38 plants that bred true (RRYY), 65 and 60 hybrid for one gene (RRYy and RrYY), and 138 hybrid for both (RrYy). Over all nine genotypes, the pure ones averaged 33 plants each, the half-hybrid ones 65 and the double

Way 1



Way 2



Purple: from the round yellow parent.

Teal: from the wrinkled green parent.

Half the cells line up each way: four gametes, ¼ each.

Figure 3.2 *The chromosomal basis of independent assortment. In a cell of an RrYy plant at metaphase I, the two pairs of homologous chromosomes can line up in two ways, equally often. Each way sends a different combination to each pole, so together they make the four kinds of gametes in equal numbers.*

CAREFUL

When the law doesn't hold

Independent assortment holds for genes on **different** chromosomes. Two genes close together on the **same** chromosome tend to go into the gametes together; they are **linked**, and their F_2 doesn't give 9 : 3 : 3 : 1. Morgan's fruit flies showed this: see Grade 9 Optional Science, Relation between gene and chromosome. Mendel's seven characters lie on only four of the pea's seven chromosomes, so some of them share a chromosome; but the characters he published crosses of behaved as independent, either because they are on different chromosomes or because they lie so far apart on one that crossing over separates them about as often as not. If an exam asks for the law, state it as above and add that it applies to genes on different chromosomes.

hybrid 138: close to 33 : 66 : 132, that is, 1 : 2 : 4. The genotypes were in the proportions the checkerboard gives.

A STORY FROM SCIENCE

The monk who counted

Mendel's F_2 in the dihybrid cross gave him 556 seeds, and the follow-up in the next year 529 plants, each to be grown, self-pollinated, harvested and sorted by hand. Few biologists of his time counted anything; Mendel had studied physics and mathematics at the University of Vienna, and he treated his peas as a physicist treats an experiment, predicting numbers and then checking them. In 1866 his paper went to libraries across Europe, and almost nobody understood it. It was found again only in 1900, by botanists who had come to the same rules with plants of their own.

Working out crosses quickly: the product rule

A checkerboard always works, but there is a faster way, which follows straight from the law. Because the two genes assort independently, **the chance of a combination is the chance of one character multiplied by the chance of the other**. This is the **product rule** of probability. Work out each gene on its own, as a monohybrid cross, then multiply.

WORKED EXAMPLE

A cross that isn't 9 : 3 : 3 : 1

A plant with round yellow seeds (RrYy) is crossed with one with round green seeds (Rryy). In what ratio do the four phenotypes appear?

Answer

Shape: Rr × Rr gives $\frac{3}{4}$ round, $\frac{1}{4}$ wrinkled.
Color: Yy × yy gives $\frac{1}{2}$ yellow, $\frac{1}{2}$ green.
Multiply each pair:

- round yellow: $\frac{3}{4} \times \frac{1}{2} = \frac{3}{8}$
- round green: $\frac{3}{4} \times \frac{1}{2} = \frac{3}{8}$
- wrinkled yellow: $\frac{1}{4} \times \frac{1}{2} = \frac{1}{8}$
- wrinkled green: $\frac{1}{4} \times \frac{1}{2} = \frac{1}{8}$

The ratio is **3 : 3 : 1 : 1**. (Set the widget above to RrYy × Rryy and count the boxes: 6, 6, 2 and 2 out of 16.)

WORKED EXAMPLE

How many seeds of each kind?

In a dihybrid F_2, a farmer harvests 1,280 seeds. How many does she expect to be wrinkled yellow, and how many of those will breed true?

Answer

Wrinkled yellow is 3 in 16:

$$\frac{3}{16} \times 1280 = 240$$

Of the 3 wrinkled yellow boxes, 1 is rrYY (pure) and 2 are rrYy, so a third of them breed true: $\frac{1}{3} \times 240 =$ **80 seeds**. About **240** wrinkled yellow seeds, of which about **80** breed true.

3.2 RNA and DNA: their structure in detail

Grade 10 Science gave the outline: DNA is a double helix of nucleotides with the bases A, T, G and C, and RNA is a single strand with U in place of T. Now we look at what each part is made of and how the parts are joined.

The three parts of a nucleotide

Both DNA and RNA are **polynucleotides**: long chains of **nucleotides**. Every nucleotide has three parts.

1. Phosphoric acid. Phosphoric acid is H_3PO_4. In a nucleotide it is joined to the sugar, having lost water, and is present as a **phosphate group**. It carries a negative charge in the cell, which is why DNA and RNA are acids: nucleic acids.

2. A pentose sugar. The sugar has five carbon atoms, numbered **1′ to 5′** (read "one prime" to "five prime"; the primes keep the sugar's numbers apart from the base's). Four carbons and an oxygen make a five-sided ring; the 5′ carbon sticks out of it.

- In RNA the sugar is **ribose**, $C_5H_{10}O_5$, with an –OH group on its 2′ carbon.
- In DNA it is **deoxyribose**, $C_5H_{10}O_4$: the same, but with only H on the 2′ carbon. "Deoxy" means one oxygen less.

3. A nitrogen base. The bases are flat rings of carbon and nitrogen atoms, of two kinds.

- **Purines** have two rings, a six-sided ring joined to a five-sided ring: **adenine (A)** and **guanine (G)**.
- **Pyrimidines** have one six-sided ring: **cytosine (C)**, **thymine (T)** and **uracil (U)**.

Base	Kind	Formula	Found in
Adenine (A)	purine	$C_5H_5N_5$	DNA and RNA
Guanine (G)	purine	$C_5H_5N_5O$	DNA and RNA
Cytosine (C)	pyrimidine	$C_4H_5N_3O$	DNA and RNA
Thymine (T)	pyrimidine	$C_5H_6N_2O_2$	DNA only
Uracil (U)	pyrimidine	$C_4H_4N_2O_2$	RNA only

A sugar joined to a base is a **nucleoside**: adenine with ribose is **adenosine**, with deoxyribose **deoxyadenosine**. A nucleoside with a phosphate on its 5′ carbon is a **nucleotide**:

$$\text{nucleotide} = \text{phosphate} + \text{pentose sugar} + \text{nitrogen base}$$

The bonds that hold them together

CDC's course asks for the **bonding among** the parts. There are three kinds of bond, and each has its own job.

DID YOU KNOW?

A nucleotide you already know

ATP, the energy carrier of every cell, is a nucleotide: adenosine with three phosphates in a row instead of one. The same adenine and ribose that sit in RNA carry energy around the cell when they hold three phosphates.

1. The N-glycosidic bond (sugar to base). The base joins the **1′ carbon** of the sugar through one of its nitrogen atoms: nitrogen 9 of a purine, or nitrogen 1 of a pyrimidine. It is a strong covalent bond.

2. The phosphodiester bond (sugar to sugar, through a phosphate). The phosphate on the 5′ carbon of one nucleotide joins the 3′ carbon of the next. So the phosphate is joined to two sugars by two ester bonds, a **phosphodiester bond**. Sugar, phosphate, sugar, phosphate: these make the **sugar–phosphate backbone** of the strand, with the bases sticking out to one side. The bond is covalent and strong, so the order of the bases along a strand doesn't change by itself.

A strand has two different ends. At one end a sugar's 5′ carbon has a phosphate not joined to anything further: the **5′ end**. At the other end a sugar's 3′ carbon has a free –OH: the **3′ end**. A strand has a direction, like a written word, and its sequence is always read from 5′ to 3′.

3. Hydrogen bonds (base to base, across the strands). In DNA the bases of one strand pair with the bases of the other: **adenine with thymine** by **two** hydrogen bonds, and **guanine with cytosine** by **three**. A hydrogen bond is a weak attraction between a hydrogen atom joined to a nitrogen and a nitrogen or oxygen atom of the other base. One alone is about twenty times weaker than a covalent bond, but millions together hold the two strands firmly, and they can still be opened a few at a time when the DNA is copied or read.

Why does A pair only with T, and G only with C? Two reasons, both visible in Figure 3.4. A purine (two rings) always pairs with a pyrimidine (one ring), so every rung of the ladder is three rings wide and the helix is the same width all along; two purines would bulge, two pyrimidines would leave a gap. And of the possible purine–pyrimidine pairs, only A–T and G–C put their hydrogen atoms opposite atoms that can attract them. Pairing that follows these rules is **complementary base pairing**: knowing one strand, you know the other.

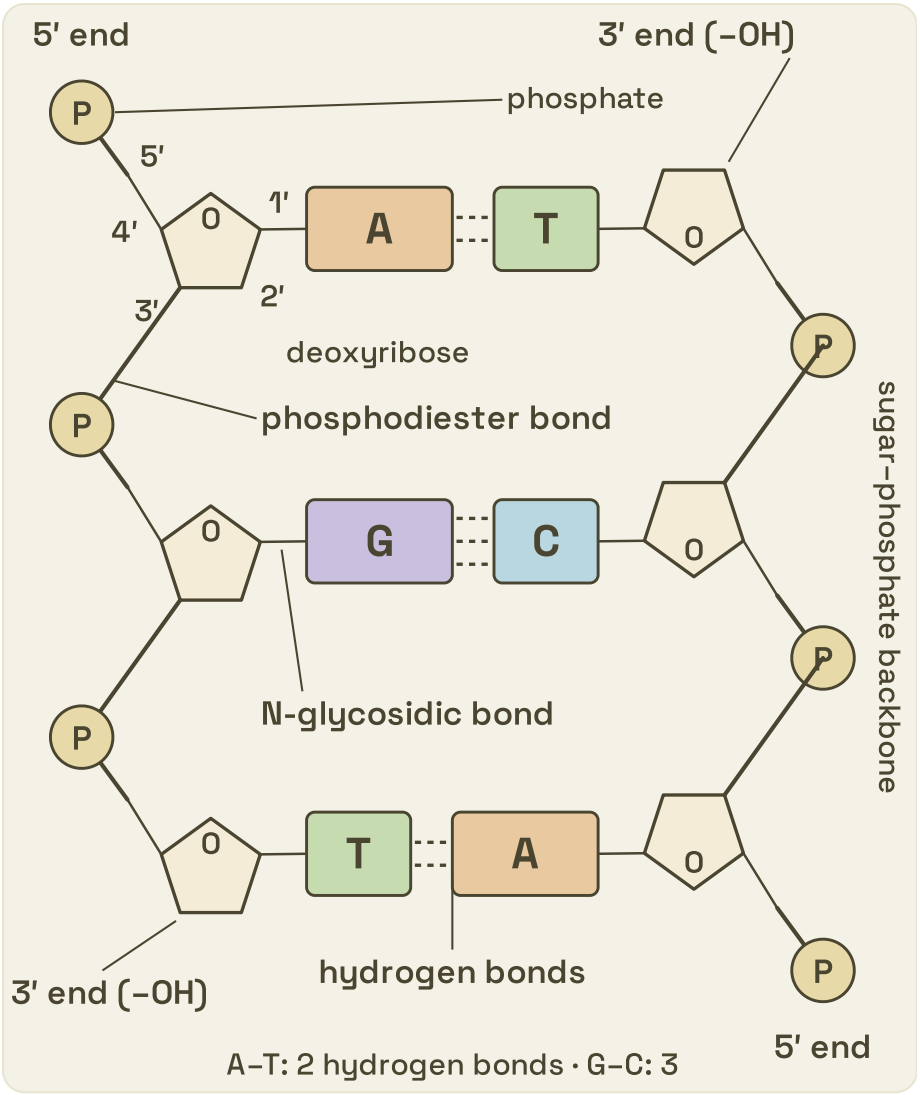
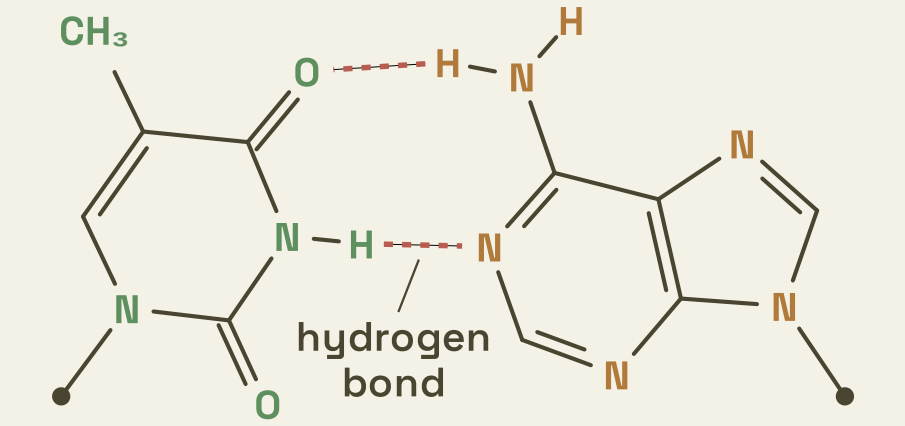


Figure 3.3 *Three base pairs of DNA, untwisted. Each strand is a chain of nucleotides; the phosphodiester bonds join the 3′ carbon of one sugar to the 5′ carbon of the next, and the N-glycosidic bonds join each base to its sugar's 1′ carbon. The strands run in opposite directions, and their bases pair by hydrogen bonds: two between A and T, three between G and C.*

The double helix of DNA

In 1953 James Watson and Francis Crick, using the X-ray photographs of Rosalind Franklin and Maurice Wilkins (the story is in Grade 10 Science), proposed the struc-

adenine–thymine: 2 hydrogen bonds



guanine–cytosine: 3 hydrogen bonds

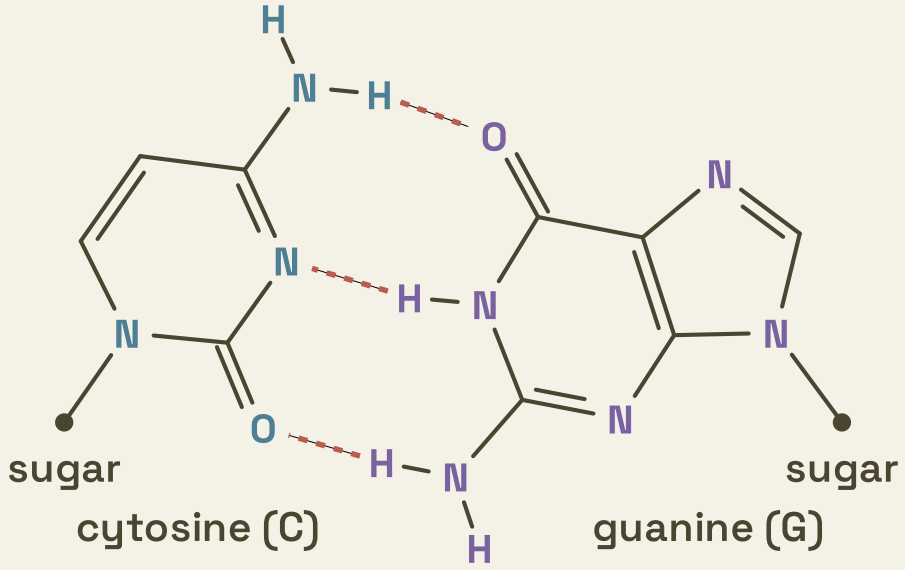


Figure 3.4 *The two base pairs of DNA, drawn from the positions of their atoms in a real DNA crystal. Hydrogen bonds (dashed) join adenine to thymine in two places and guanine to cytosine in three. Both pairs are the same width, from sugar to sugar.*

> CAREFUL
> **Hydrogen bonds, not double bonds**
>
> Some books write that A and T are joined by "a double bond" and G and C by "a triple bond". They are joined by **two** and **three hydrogen bonds**, weak attractions, not the covalent double and triple bonds of chemistry. Write "two hydrogen bonds" and "three hydrogen bonds" in an exam.

ture of DNA. Its main features:

1. DNA has **two polynucleotide strands** twisted round a common axis into a **right-handed double helix**, like a spiral staircase that turns clockwise as it rises away from you.
2. The sugar–phosphate backbones are on the **outside**; the bases are **inside**, stacked flat like the steps of the staircase.
3. The two strands are **antiparallel**: one runs 5′ → 3′ upward, the other 5′ → 3′ downward.
4. The strands are held together by **hydrogen bonds** between complementary bases, A with T (two) and G with C (three).
5. The helix is about **2 nm** wide. Neighboring base pairs are **0.34 nm** apart, and the helix makes one full turn in about **10 base pairs**, that is, about **3.4 nm**. (Measurements of DNA in water give about 10.5 base pairs a turn; the round 10 is the figure most books use.)
6. The backbones twist unevenly, leaving two grooves of different widths along the helix, the **major groove** and the **minor groove**, where proteins reach in to read the bases.

From Chargaff's rules, A = T and G = C, so A + G = T + C: **purines make up half the bases of DNA**, and given one base's share, you can find all four.

The structure of RNA

RNA is also a chain of nucleotides joined by phosphodiester bonds, with a 5′ end and a 3′ end, but:

- its sugar is **ribose**;

A STORY FROM SCIENCE

Chargaff's rules

Before the double helix was known, the Austrian-American chemist Erwin Chargaff measured how much of each base there was in the DNA of many living things. In 1950 he reported a pattern no one could yet explain: in every sample, **the amount of adenine equals the amount of thymine, and the amount of guanine equals that of cytosine**, though the A + T and G + C totals differ from one species to another. Human DNA is about 31% A, 29% T, 20% G and 20% C: A and T, and G and C, come out equal within the error of his measurements. Chargaff's rules were one of the clues Watson and Crick used: they are what you get if every A on one strand faces a T on the other, and every G a C.

WORKED EXAMPLE

From one base to all four

In a sample of DNA, 22% of the bases are adenine. What are the percentages of the other three?

Answer

A = T, so **thymine = 22%**.
A + T = 44%, so G + C = 100 − 44 = 56%. G = C, so **guanine = cytosine = 28%** each.

WORKED EXAMPLE

Length, turns and hydrogen bonds

A double-stranded DNA molecule has 1,000 base pairs, of which 300 are A–T pairs. Find a. its length, b. the number of turns of its helix, c. the number of hydrogen bonds between its strands.

Answer

a. Length = 1000×0.34 nm = **340 nm** (0.34 μm).

b. One turn is 10 base pairs, so $1000 \div 10$ = **100 turns**.

c. The other 1000 − 300 = 700 pairs are G–C. Hydrogen bonds = $300 \times 2 + 700 \times 3 = 600 + 2100$ = **2,700**.

- its bases are **A, G, C and U**: uracil takes the place of thymine, and pairs with adenine by two hydrogen bonds;

YOUR TURN

One strand of a stretch of DNA reads 5′-ATGCCTAG-3′. Write the other strand, with its ends, and count the hydrogen bonds between them.

Hint

Pair each base, and remember the strands are antiparallel.

Answer

The partner of A T G C C T A G is T A C G G A T C, running the other way: **3′-TACGGATC-5′** (read 5′ to 3′, that is 5′-CTAGGCAT-3′). The pairs are A–T, T–A, G–C, C–G, C–G, T–A, A–T, G–C: 4 A–T pairs and 4 G–C pairs, so $4 \times 2 + 4 \times 3 = \mathbf{20}$ **hydrogen bonds**.

DID YOU KNOW?

Why G–C-rich DNA melts at a higher temperature

Heat double-stranded DNA and the strands come apart, as the hydrogen bonds and the stacking of the bases give way. DNA rich in G–C pairs, with three hydrogen bonds each and stronger stacking, needs a higher temperature to separate than DNA rich in A–T. Bacteria that live in hot springs tend to have G–C-rich RNA in their ribosomes for the same reason. Every DNA test in a laboratory, from a COVID-19 PCR test to a paternity test, heats DNA to about 95 °C to separate its strands before copying them.

- it is usually a **single strand**. A single strand can still fold back on itself where its bases happen to be complementary, making short double-stranded stretches (stems) and loops. Its folded shape gives each kind of RNA its job.

RNA is made by copying one strand of DNA (transcription), so its sequence is complementary to that DNA strand, with U wherever the DNA strand has A.

Check 2

Which bond joins the phosphate of one nucleotide to the sugar of the next in a strand of DNA?

a. a hydrogen bond
b. a phosphodiester bond, from the 5′ carbon of one sugar to the 3′ carbon of the next
c. an N-glycosidic bond, from the 1′ carbon to the base
d. a peptide bond

3.3 RNA: types and function

A gene in DNA stays in the nucleus. To make the protein it codes for, the cell uses RNA, and three kinds of RNA share the work. How they work together, step by step, is told in How does a gene make a protein?; here we look at each kind.

Messenger RNA (mRNA)

Messenger RNA is a copy of one gene, made in the nucleus by transcription, which carries the gene's message out to the ribosomes in the cytoplasm.

- **Structure:** a single, unfolded strand, as long as its gene: a few hundred to many thousands of nucleotides. Its bases are read three at a time; each group of three, a **codon**, stands for one amino acid. In our cells an mRNA gets a protective cap at its 5′ end and a tail of a hundred or more A's at its 3′ end.
- **Function:** it carries the genetic code from DNA to the ribosome, where it is the template that sets the order of amino acids in the protein.
- It is short-lived: most mRNAs in a cell last minutes to hours before they are broken down, so the cell can change which proteins it makes. It makes up only a few percent of a cell's RNA.

Transfer RNA (tRNA)

Transfer RNA is a small RNA that brings amino acids to the ribosome and matches each one to its codon.

- **Structure:** a single chain, typically 76 to 90 nucleotides, the smallest of the three kinds. Its complementary stretches pair up into four short stems, and it folds into the shape of a **clover leaf**, which in three dimensions is folded again into an L.
  - The **acceptor arm**: the two ends of the chain, paired; the 3′ end always ends in the bases CCA, and the amino acid is attached to its last A.
  - The **anticodon arm**: its loop carries the **anticodon**, three bases that pair with a codon on the mRNA.
  - Two more arms with loops (the **D arm** and the **T arm**, named after unusual bases found in them) help the molecule fold and fit the ribosome.
- **Function:** each tRNA carries one particular amino acid; when its anticodon pairs with the matching codon on the mRNA, it leaves its amino acid in the right place in the growing protein. It is the **adapter** between the language of bases and the language of amino acids. Cells have dozens of kinds of tRNA, at least one for each of the 20 amino acids.

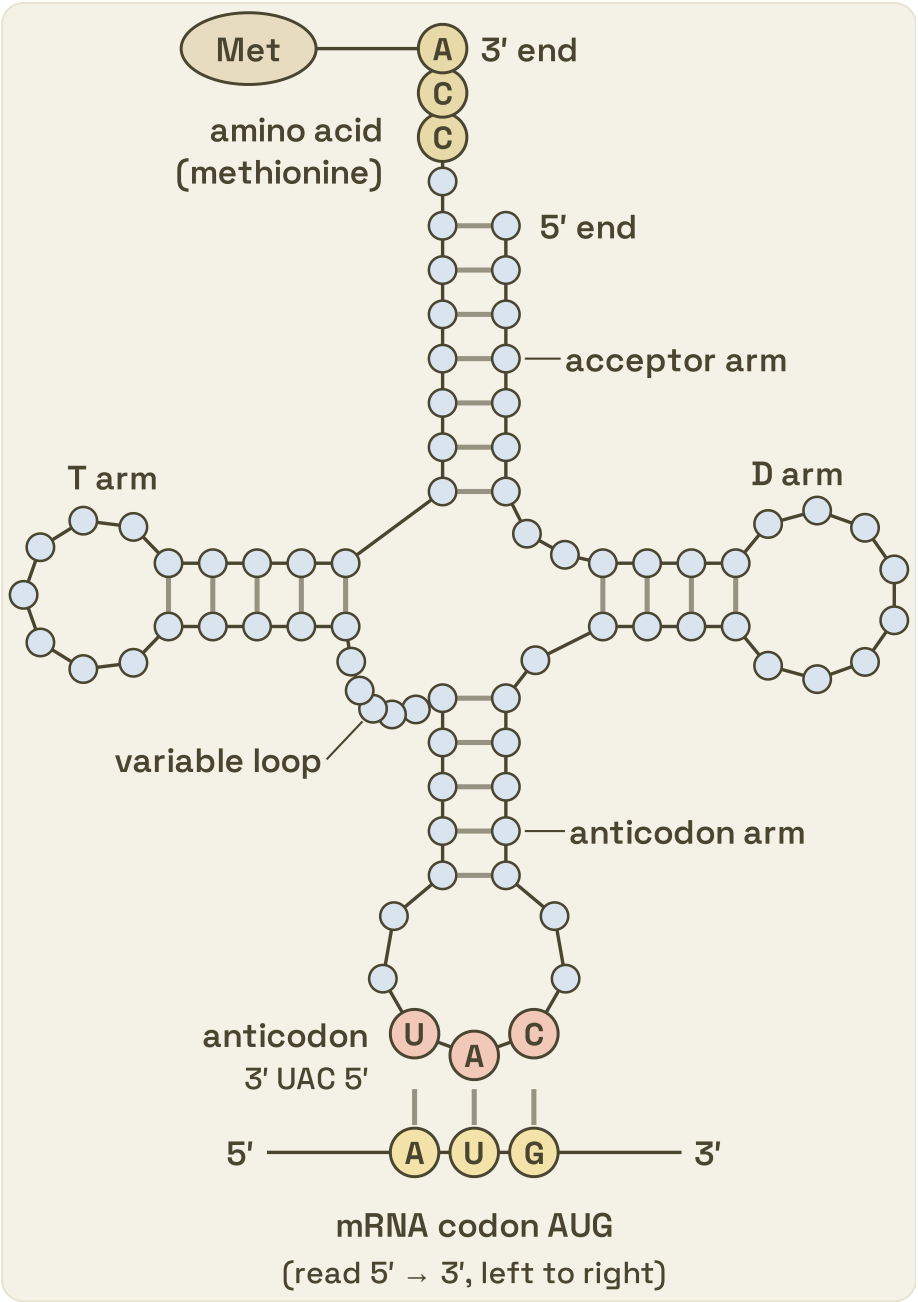


Figure 3.5 *A transfer RNA drawn as a clover leaf. Short stretches of its single chain pair up into four stems; the amino acid is carried at the 3′ end (…CCA), and the anticodon, at the far end, pairs with a codon of the mRNA.*

The mRNA is drawn 5′ to 3′ from left to right, so the anticodon under its codon AUG reads 3′-UAC-5′: the two run in opposite directions. This tRNA carries methionine.

Ribosomal RNA (rRNA)

Ribosomal RNA is the RNA of the ribosomes.

- **Structure:** long RNA chains folded tightly together with dozens of proteins into the two subunits of a ribosome. In our cells there are four kinds (named by how fast they settle in a centrifuge: 28S, 18S, 5.8S and 5S); bacteria have three (23S, 16S and 5S). It is made in the nucleolus (Grade 9 Optional Science, Ribosome).
- **Function:** it forms the structure of the ribosome, holds the mRNA and the tRNAs in place, and, remarkably, it is the rRNA of the large subunit, not a protein, that joins each new amino acid to the chain. So the machine that makes every protein is itself mostly RNA.
- It is the most plentiful kind: about four-fifths of all the RNA in a cell.

	mRNA	tRNA	rRNA
Full name	messenger RNA	transfer RNA	ribosomal RNA
Share of a cell's RNA	a few percent	about a sixth	about four-fifths
Shape	a straight single strand	a folded clover leaf (an L)	folded with proteins into ribosomes
Size	hundreds to thousands of nucleotides	about 76 to 90 nucleotides	about 120 to 5,000 nucleotides
Job	carries the code of one gene to the ribosome	brings amino acids and matches them to codons	builds the ribosome and joins the amino acids
Special part	codons	anticodon, CCA end	–

WORKED EXAMPLE

Reading with the three RNAs

A stretch of mRNA reads 5′-AUG GCU UGG-3′. a. How many amino acids does it code for? b. What are the anticodons of the three tRNAs that read it?

Answer

a. Each codon of three bases stands for one amino acid: 9 bases ÷ 3 = **3 amino acids** (here methionine, alanine and tryptophan).

b. Each anticodon pairs with its codon, A with U and G with C, running the other way: AUG is read by 3′-**UAC**-5′, GCU by 3′-**CGA**-5′ and UGG by 3′-**ACC**-5′.

A STORY FROM SCIENCE

The first RNA read letter by letter

In 1965 Robert Holley and his team at Cornell University worked out the whole sequence of one tRNA, the one that carries alanine in yeast. It took years, and a great deal of baker's yeast, to purify enough of that one kind of tRNA to work on. It was the first nucleic acid whose complete sequence was known, and from its sequence the team saw that it could fold into a clover leaf. He shared the 1968 Nobel Prize in Physiology or Medicine with Har Gobind Khorana and Marshall Nirenberg, who had worked out which codon stands for which amino acid.

Check 3

Which part of a tRNA pairs with the mRNA?

a its CCA end

b the anticodon, in the loop at the far end from the amino acid

c the amino acid it carries

d its D arm

3.4 Role of DNA in genetics

Everything genetics studies, from the color of a pea to a disease that runs in a family, comes back to DNA. Its roles are these.

1. It stores the genetic information

The order of the bases along DNA is the genetic information. A **gene** is a stretch of DNA whose base sequence codes for one protein or one RNA. With four bases, even a

short stretch can be written in an enormous number of ways: a stretch of just ten base pairs can be arranged in 4^{10}, about a million, different ways, and one set of your chromosomes holds about 3.1 billion base pairs and about 20,000 genes for proteins. DNA is also stable: its backbone is strong, and its two strands keep a copy of each other, so damage on one strand can be repaired from the other.

2. It copies itself: replication

Before a cell divides, its DNA is copied, so each daughter cell gets the full set of genes. This copying is **replication**.

1. An enzyme breaks the hydrogen bonds and **unzips** the two strands, starting at many places along the DNA.
2. Each old strand acts as a **template**: an enzyme, **DNA polymerase**, adds new nucleotides one by one, each pairing with the base it faces, A with T and G with C, and joins them into a new strand (always growing from its 5′ end toward its 3′ end).
3. The result is two DNA molecules, exactly alike and exactly like the first. Each has **one old strand and one new strand**.

Because each new molecule keeps one of the old strands, replication is called **semi-conservative** (“half kept”).

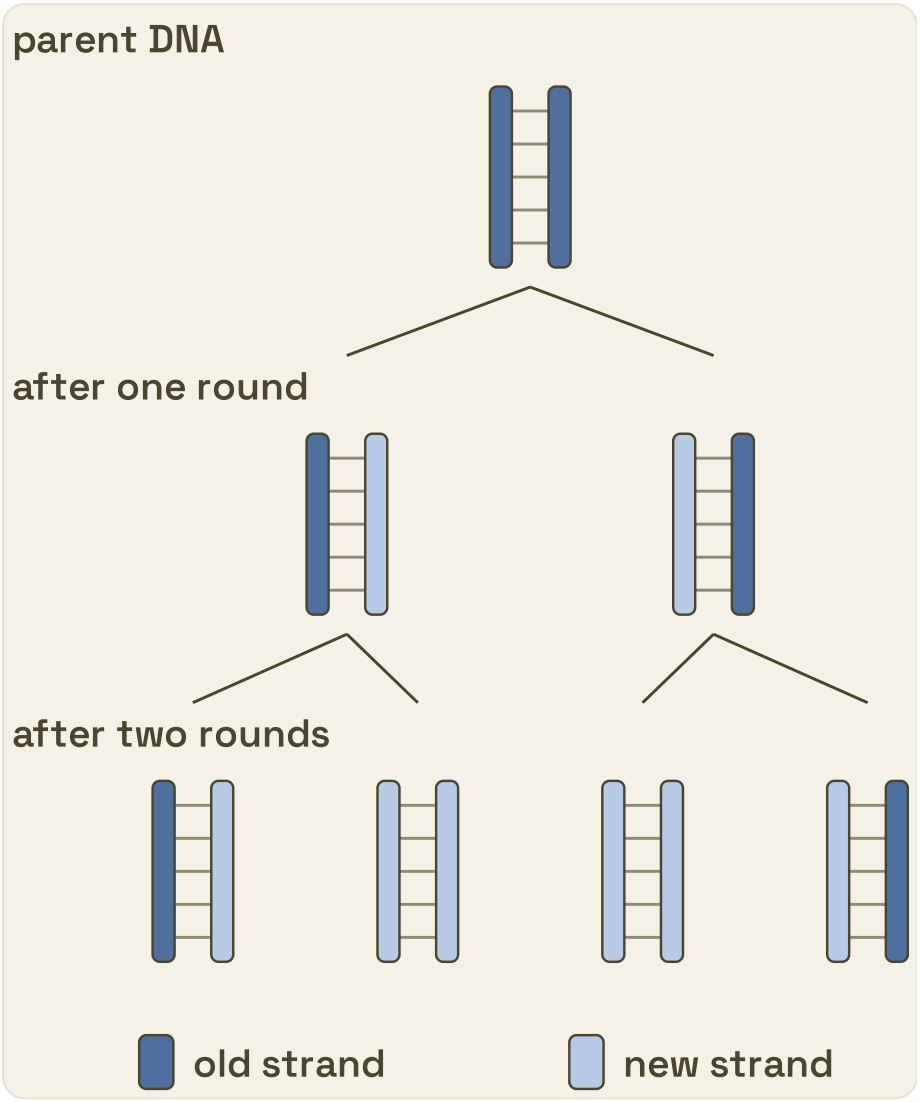


Figure 3.6 *Semi-conservative replication. Each strand of the parent DNA (dark) is the template for a new strand (light). After one round, every molecule has one old and one new strand; after two rounds, two of the four molecules still carry an old strand and two are all new.*

A STORY FROM SCIENCE

The most beautiful experiment in biology

How could anyone see which strand is old? In 1958 Matthew Meselson and Franklin Stahl grew the bacterium *E. coli* for many generations on food containing a heavy kind of nitrogen, ^{15}N, until all its DNA was heavy. Then they moved it to food with ordinary ^{14}N and took samples as it divided. Spun very fast in a dense salt solution, DNA settles at a height set by its density. After one generation, all the DNA was halfway between heavy and light: each molecule was half old and half new. After two generations, half was halfway and half was light, just as Figure 3.6 shows. The experiment is so clean and simple that biologists still call it the most beautiful one in their field.

3. It controls the characters: from gene to protein

DNA itself doesn't make a pea round or an eye brown. It controls characters through the **proteins** it codes for: enzymes, pigments, hormones, the proteins of muscle and blood. The flow of information goes **DNA → RNA → protein**, which Crick called the **central dogma** of molecular biology.

- **Transcription:** in the nucleus, a gene's DNA is copied into mRNA.
- **Translation:** at a ribosome, the mRNA is read codon by codon; tRNAs bring the amino acids, and rRNA joins them into a protein.

The code is read in **codons** of three bases. With four bases there are $4^3 = 64$ codons: 61 stand for the 20 amino acids (most amino acids have more than one codon), and 3 (UAA, UAG and UGA) mean "stop". AUG both means methionine and marks where a protein starts. The same code is used by almost every living thing, from bacteria to people, which is why a human gene put into a bacterium can make a human protein: this is how insulin for diabetes is made.

Mendel's round and wrinkled peas are an example. The R allele codes for a starch-branching enzyme, which the seed needs to build its branched starch. The r allele carries an extra stretch of about 800 base pairs in the middle of the gene, so the enzyme isn't made. A wrinkled (rr) seed keeps more sugar, takes up more water while it grows, and wrinkles when it dries.

4. It brings variation

Children differ from their parents and from each other because of DNA.

- **Mutation:** a change in the base sequence (one base changed, added or lost, or a larger stretch). A mutation in a gene may change its protein; most mutations do

little or harm, a few help. Mutations are the source of all new alleles; the r allele of the pea and the allele for color blindness both began as mutations.

- **Recombination:** in meiosis, independent assortment (3.1) and crossing over shuffle the alleles a person has into new combinations in each gamete, so every child of the same parents gets a different mix.

Over many generations, variation is what natural selection acts on, so DNA is also the basis of evolution.

5. It is the material passed from generation to generation

Through gametes, DNA carries the genes from parents to offspring: half from each. And its sequence can be read: DNA fingerprinting, in Unit 5, uses the differences in DNA between people to identify them.

How we know genes are made of DNA

Until the 1940s, most scientists thought genes were made of protein, which seemed more varied than a molecule of only four kinds of nucleotide. Three experiments changed their minds.

1. **Griffith, 1928.** Frederick Griffith found that a harmless strain of the pneumonia bacterium became deadly when mixed with dead cells of a deadly strain: something from the dead cells **transformed** the living ones.
2. **Avery, MacLeod and McCarty, 1944.** They purified the dead cells' substances one at a time. Only DNA transformed the harmless bacteria; destroying the DNA with an enzyme stopped it, and destroying the proteins didn't.
3. **Hershey and Chase, 1952.** Alfred Hershey and Martha Chase labeled bacteriophages either in their protein (with radioactive sulphur) or in their DNA (with radioactive phosphorus). Only the DNA went into the bacteria, and new phages were made from it, as you saw in the lytic cycle in Grade 9 Optional Science.

In a few viruses, such as those of flu, measles and COVID-19, the genes are RNA, not DNA. Everywhere else, the genetic material is DNA.

> **Check 4**
> How many codons are there, and how many of them stand for amino acids?
>
> a 20 codons, all for amino acids
> b 16 codons, 15 for amino acids
> c 64 codons, 61 for amino acids
> d 64 codons, all for amino acids

3.5 Sex-linked inheritance

A human body cell has 22 pairs of autosomes and one pair of **sex chromosomes**: XX in a woman, XY in a man (see Grade 9 Optional Science). The sex chromosomes don't only decide sex. The X chromosome is large, with about 800 genes for proteins, most of which have nothing to do with being male or female: genes for seeing colors, for clotting blood, for the muscles. The Y is small, with only a few dozen different genes for proteins, several of them in many copies.

Genes on the sex chromosomes are **sex-linked genes**, and the inheritance of the characters they control is **sex-linked inheritance**. Such characters are passed on differently to sons and daughters, because a son gets his X from his mother and his Y from his father, while a daughter gets one X from each.

Types of sex-linked genes on the basis of location

The X and Y share only small stretches at their tips; the rest of each is its own. Where a gene lies decides how it is inherited.

1. X-linked genes. These lie on the part of the X that has no partner on the Y (the non-homologous part of the X). A woman has two copies of each; a man has only one, and no second copy on his Y to hide a recessive allele. Examples: the genes for **red–green color vision**, for the clotting factors whose loss causes **hemophilia**, for the muscle protein dystrophin (Duchenne muscular dystrophy), and for the enzyme G6PD. Almost all sex-linked inheritance in humans is X-linked.

2. Y-linked genes (holandric genes). These lie on the part of the Y that has no partner on the X. They pass **only from father to son**, to every son, and never to a daughter: *holandric* means "wholly in males". The best known is **SRY**, the gene that makes an embryo develop as a male. Most of the Y's other genes are needed for making sperm.

3. XY-linked genes (incompletely sex-linked). These lie in the **pseudoautosomal regions**, the short stretches the X and Y share at their tips, where the two pair up and cross over in meiosis. A gene here has a copy on both the X and the Y, so it is inherited almost like a gene on an autosome. Example: **SHOX**, a gene needed for the growth of the bones of the arms and legs.

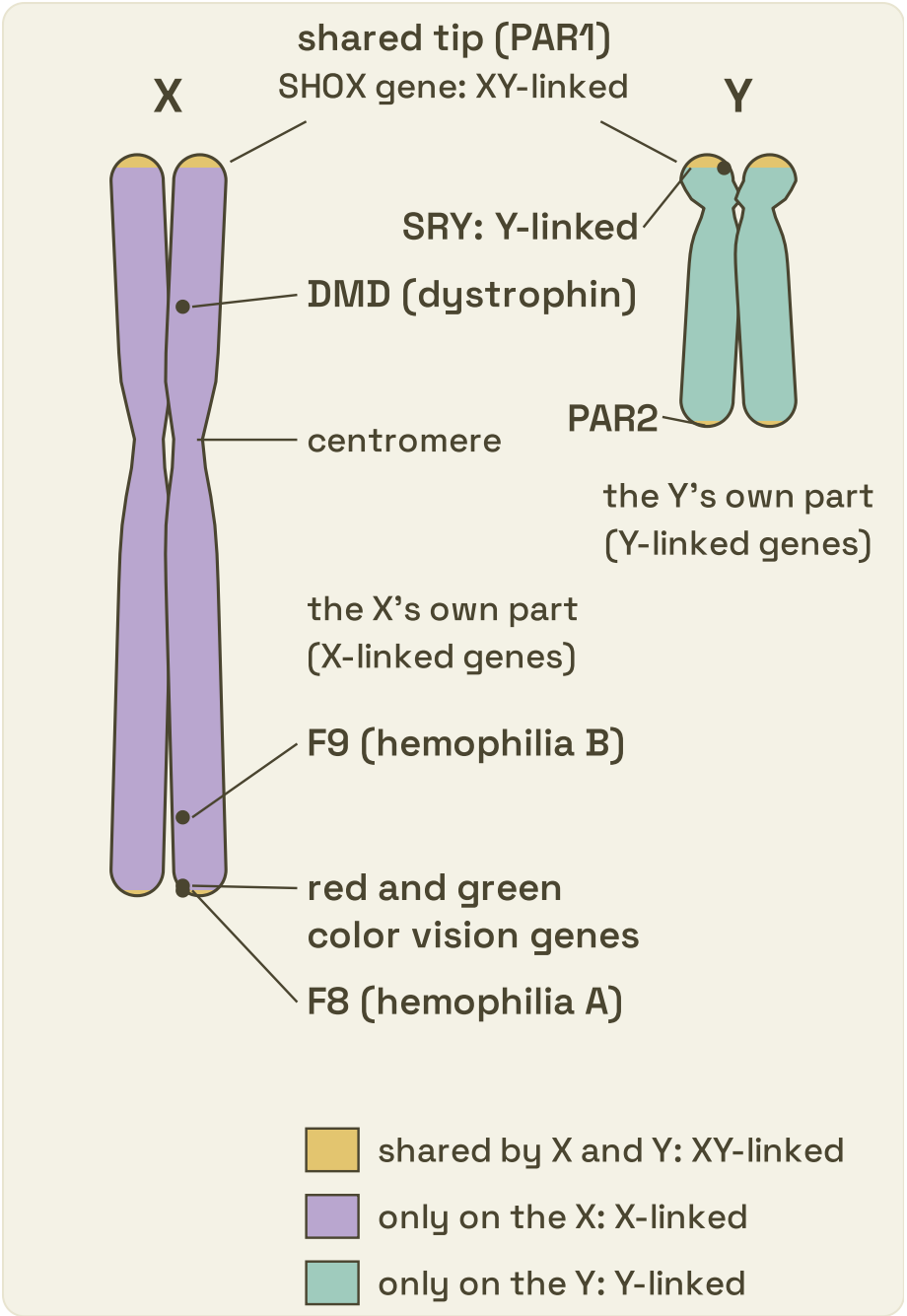

shared tip (PAR1)
SHOX gene: XY-linked
X
Y
SRY: Y-linked
DMD (dystrophin)
PAR2
centromere
the Y's own part
(Y-linked genes)
the X's own part
(X-linked genes)
F9 (hemophilia B)
red and green
color vision genes
F8 (hemophilia A)
shared by X and Y: XY-linked
only on the X: X-linked
only on the Y: Y-linked

Figure 3.7 *The human X and Y chromosomes, drawn to scale (X about 156 million base pairs, Y about 57 million). The X and Y share two short stretches at their tips (the pseudoautosomal regions); the one at the bottom is drawn a little wider so it shows. Genes in the rest of the X are X-linked; genes in the rest of the Y are Y-linked.*

CAREFUL

Old examples that are wrong

Some books give **hairy ears** (hypertrichosis of the pinna) as a Y-linked character and **total color blindness** or **xeroderma pigmentosum** as XY-linked. Hairy ears were suggested as Y-linked from family studies in the 1950s and 1960s, but no gene for them has ever been found on the Y, and the evidence is now doubted; total color blindness and xeroderma pigmentosum are caused by genes on autosomes.

In the exam: give **SRY** as the example of a Y-linked gene and **SHOX** of an XY-linked one. Many books and marking schemes still give hairy ears as the Y-linked example; if you name it, call it a character once *thought* to be Y-linked.

What is special about X-linked inheritance

Most X-linked disorders are caused by a **recessive** allele. Write the normal X as X and the X with the recessive allele as X^c (for color blindness) or X^h (for hemophilia). Then:

	Normal	Carrier	Affected
Woman	XX	X^cX	X^cX^c
Man	XY	none	X^cY

A **carrier** is a person who has one recessive allele, shows no sign of the disorder, and can pass it on. A woman with one X^c is a carrier: her other X hides it. A man with one X^c has no other X, so he is affected. A man has only one copy of each X-linked gene; he is **hemizygous** for it. So:

1. **More men than women are affected.** A man needs one recessive allele, a woman two. If one X in twelve carries the allele, about 1 man in 12 is affected, but only about 1 woman in 144 ($1/12 \times 1/12$). (The real figure for women, about 1 in 200, is lower still, because red–green color blindness comes from two different genes: a woman with a faulty red gene on one X and a faulty green gene on the other sees normally.)
2. **A father never passes an X-linked character to his son**, because he gives his son his Y, not his X.

3. **An affected father passes the allele to every daughter**, who becomes a carrier (or affected, if her mother gave one too).
4. **Criss-cross inheritance:** the character typically passes from an affected man, through his carrier daughter, to her son, his grandson, skipping a generation.
5. An affected woman must have an affected father and a carrier or affected mother; all her sons are affected.

Color blindness

In the retina, light is sensed by **cones** of three kinds, each with a pigment most sensitive to red, green or blue light. The genes for the red and green pigments sit side by side near the tip of the long arm of the X chromosome. An allele that makes either pigment faulty or missing gives **red–green color blindness**: red and green, and colors mixed from them, look alike. (The blue pigment's gene is on chromosome 7.) Among people of Northern European ancestry, red–green color blindness affects about **1 man in 12** and **1 woman in 200**.

Here is how color blindness passes in the four kinds of marriage.

Cross 1: a normal woman and a color-blind man.

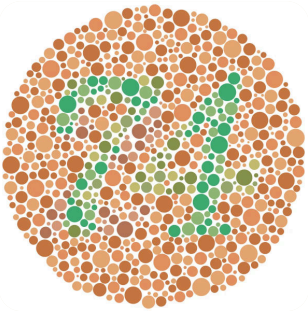


Figure 3.8 *A plate of the kind used in the Ishihara test, designed by the Japanese eye doctor Shinobu Ishihara in 1917. A person with normal color vision reads the number 74; a person with red–green color blindness sees 21, or no number.*
Photo: Shinobu Ishihara, Public domain

Gametes	X (from mother)	X (from mother)
X^c (from father)	X^cX, carrier daughter	X^cX, carrier daughter
Y (from father)	XY, normal son	XY, normal son

All the daughters are carriers with normal vision; all the sons are normal.

Cross 2: a carrier woman and a normal man.

Gametes	X^c (from mother)	X (from mother)
X (from father)	X^cX, carrier daughter	XX, normal daughter
Y (from father)	X^cY, color-blind son	XY, normal son

Half the sons are color-blind; half the daughters are carriers; no daughter is color-blind. Put Crosses 1 and 2 together and you see criss-cross inheritance: a color-blind grandfather, a carrier daughter, a color-blind grandson.

Cross 3: a carrier woman and a color-blind man. The daughters are X^cX^c (color-blind) or X^cX (carrier), half each; the sons are X^cY (color-blind) or XY (normal), half each. This is the only way a daughter can be color-blind: her father must be color-blind and her mother must carry the allele.

Cross 4: a color-blind woman and a normal man. Every daughter is X^cX, a carrier; every son is X^cY, color-blind.

A STORY FROM SCIENCE

Dalton's pink geranium

In 1794 the English chemist John Dalton, who later gave chemistry its atomic theory, described to the Manchester Literary and Philosophical Society how a geranium that everyone called pink looked almost sky-blue to him by day, and red by candlelight, while his friends saw hardly any change. Only his brother saw it as he did. Dalton guessed that the jelly inside his eyes was tinted blue, and asked for his eyes to be examined after his death. They were, in 1844, and the jelly was clear. In 1995 scientists read the DNA in his preserved eye tissue: he lacked the green pigment gene. Red–green color blindness is still called **Daltonism** in many languages.

WORKED EXAMPLE

A family puzzle

A couple, both with normal color vision, have a color-blind son. Which parent passed him the allele, and what is the chance that their next child is a color-blind son?

Answer

A son gets his X from his mother; the father gave him a Y. So the **mother** passed the allele: she has normal vision, so she is a carrier, X^cX. The father is XY.
The cross is $X^cX \times XY$ (Cross 2): the four equally likely children are a carrier daughter, a normal daughter, a color-blind son and a normal son. The chance that the next child is a color-blind son is **1 in 4** (25%). The chance that a son is color-blind is 1 in 2.

Hemophilia

Hemophilia is a disorder in which the blood clots very slowly, because one of the proteins needed for clotting, a **clotting factor**, is missing or faulty. A cut bleeds for a long time, and bleeding inside joints and muscles, even from a bump, is painful and damaging. There are two main kinds, both X-linked recessive:

- **Hemophilia A**, the commoner, lacks clotting factor VIII; it affects about 1 in 4,000 to 5,000 boys born.
- **Hemophilia B** (also called **Christmas disease**, after Stephen Christmas, the first patient described with it, in 1952) lacks factor IX; it affects about 1 in 20,000 boys born.

It passes exactly as color blindness does. With X^h for the hemophilia allele:

A carrier woman (X^hX) and a normal man (XY) have the four kinds of child of Cross 2: a carrier daughter (X^hX), a normal daughter (XX), a hemophilic son (X^hY) and a normal son (XY), each with a chance of 1 in 4.

A normal woman (XX) and a hemophilic man (X^hY): every daughter is a carrier, X^hX, and every son normal, XY.

A hemophilic woman, X^hX^h, needs a hemophilic father and a carrier mother, and both alleles are rare, so she is very rare; she is not, as older books say, impossible. Today hemophilia is treated by injecting the missing clotting factor, made from donated blood or by genetically engineered cells, and people with it can live a nearly normal life.

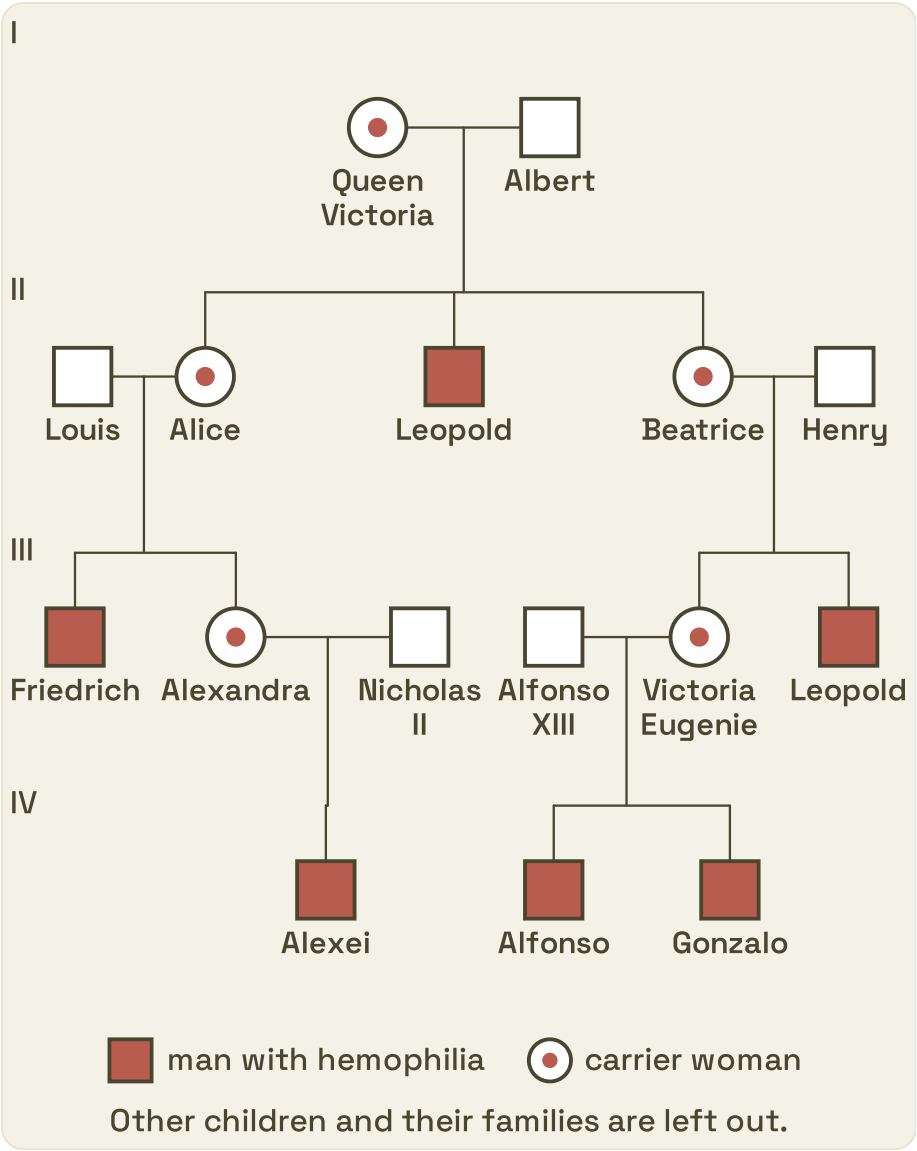


Figure 3.9 *Hemophilia in Queen Victoria's descendants (a part of the family). Squares are males, circles females; a filled square is a man with hemophilia, and a circle with a dot a woman known to be a carrier from her children. The allele passed from Victoria through two of her daughters to the royal families of Russia and Spain.*

A STORY FROM SCIENCE

The royal disease

Queen Victoria of Britain (1819–1901) was a carrier of hemophilia, though no one in her family before her is known to have had it: the allele probably arose as a new mutation in her or in one of her parents' gametes. Her son Leopold had hemophilia and died at 30 after a fall. Two of her daughters, Alice and Beatrice, were carriers, and through their daughters, who married into the royal families of Russia, Prussia and Spain, the allele reached the thrones of Europe. Alice's grandson **Alexei**, the heir to the Russian throne, had hemophilia, and his mother's fear for him drew the mystic Rasputin into the Russian court. In 2009, scientists read DNA from the bones of the Russian royal family, who were killed in 1918, and found the mutation: it is in the gene for factor IX, so the royal disease was **hemophilia B**.



Figure 3.10 *Alexei Nikolaevich (1904–1918), heir to the Russian throne, Queen Victoria's great-grandson. He had hemophilia B, passed to him through his mother and grandmother.*

Photo: Boasson and Eggler St. Petersburg Nevsky 24., Public domain

Check 5

A man with hemophilia and a woman who is not a carrier have a son. What is the chance that the son has hemophilia?

a 0, because the son gets his X from his mother

b 1 in 4

c 1 in 2

d 1, because his father has it

IN SHORT

- **Dihybrid cross:** pure RRYY × rryy gives F_1 RrYy (round yellow), which makes four kinds of gametes, RY, Ry, rY, ry, equally. F_2: phenotypes **9 round yellow : 3 round green : 3 wrinkled yellow : 1 wrinkled green**; nine genotypes in the ratio **1 : 2 : 2 : 4 : 1 : 2 : 1 : 2 : 1**.
- **Law of independent assortment:** the alleles of one pair separate into gametes independently of those of another pair. It holds for genes on different chromosomes, because each pair of homologous chromosomes lines up independently in meiosis I.
- **Verification:** the dihybrid **test cross** RrYy × rryy gives **1 : 1 : 1 : 1** (Mendel: 31 : 26 : 27 : 26); and selfing the F_2 shows its genotypes in the expected proportions.
- **Product rule:** the chance of a combination of independent characters is the product of their chances.
- **Nucleotide** = phosphate (from phosphoric acid) + pentose sugar (deoxyribose in DNA, ribose in RNA) + nitrogen base (purines A, G; pyrimidines C, T, U).
- **Bonds:** N-glycosidic (base to the sugar's 1′ carbon), phosphodiester (3′ carbon to 5′ carbon through a phosphate: the backbone), hydrogen bonds (two between A and T, three between G and C).
- **DNA:** a right-handed double helix of two antiparallel strands, 2 nm wide, 0.34 nm per base pair, about 10 base pairs (3.4 nm) a turn. **Chargaff's rules:** A = T, G = C.
- **RNA:** single-stranded, ribose, uracil. **mRNA** carries the code (codons) to the ribosome; **tRNA** (clover leaf, anticodon, CCA end) brings amino acids; **rRNA** builds ribosomes and joins amino acids, about 80% of a cell's RNA.
- **Role of DNA:** stores information; replicates semi-conservatively; controls characters through proteins (DNA → RNA → protein, 64 codons); brings variation by mutation and recombination; carries genes between generations.
- **Sex-linked genes:** X-linked (non-homologous part of X: color blindness, hemophilia), Y-linked or holandric (non-homologous part of Y: SRY, father to every son), XY-linked (pseudoautosomal: SHOX).
- **X-linked recessive inheritance:** more men affected; no father-to-son passing; affected father → carrier daughters; criss-cross from grandfather to grandson. Carrier mother × normal father: half the sons affected, half the daughters carriers.

Exercise

1. Choose the correct option

a. In a dihybrid cross, the F_2 phenotypic ratio is:

- **a** 3 : 1
- **b** 1 : 2 : 1
- **c** 9 : 3 : 3 : 1
- **d** 1 : 1 : 1 : 1

b. How many different genotypes appear in the F_2 of a dihybrid cross?

- **a** 4
- **b** 9
- **c** 16
- **d** 3

c. Which of these is a pyrimidine?

- **a** adenine
- **b** guanine
- **c** cytosine
- **d** all of these

d. The RNA that carries amino acids to the ribosome is:

- **a** mRNA
- **b** tRNA
- **c** rRNA
- **d** all three

e. A color-blind man marries a woman who is not a carrier. Their daughters will be:

- **a** all color-blind
- **b** all carriers with normal vision
- **c** half color-blind, half normal
- **d** all normal, not carriers

f. Genes found on the part of the Y chromosome that has no partner on the X are called:

- **a** X-linked genes
- **b** holandric genes
- **c** pseudoautosomal genes
- **d** autosomal genes

2. Give reasons

a. New combinations of characters appear in the F_2 of a dihybrid cross.

b. The two strands of DNA are called complementary.

c. DNA replication is called semi-conservative.

d. Color blindness and hemophilia are far commoner in men than in women.

3. Differentiate between

i. DNA and RNA

ii. Purines and pyrimidines

iii. X-linked genes and Y-linked genes

4. Answer the following questions

a. State Mendel's law of independent assortment. Explain it with a dihybrid cross between pure round yellow and pure wrinkled green peas, drawing the checkerboard and giving the phenotypic and genotypic ratios.

b. How did Mendel verify the law of independent assortment? Show the test cross.

c. Describe the structure of DNA according to Watson and Crick, with a labeled diagram.

d. What is a nucleotide? Name the bonds that join a nucleotide's parts and join nucleotides to each other.

e. Name the three types of RNA and give the function of each. Draw the clover-leaf structure of tRNA.

f. Explain the role of DNA in genetics.

g. What is sex-linked inheritance? Classify sex-linked genes on the basis of their location, with an example of each.

h. Show, with a checkerboard, the children of a woman who carries the allele for hemophilia and a normal man. What is criss-cross inheritance?

5. Numerical problems

a. In the F_2 of a dihybrid cross between pure tall pea plants with purple flowers (TTPP) and pure dwarf plants with white flowers (ttpp), 800 plants were raised. How many are expected to be a. tall with white flowers, b. dwarf with white flowers, c. pure for both characters?

b. A pea plant with round yellow seeds is test-crossed and gives 52 round yellow and 49 round green seeds. What is its genotype? Show why.

c. A DNA molecule is 6.8 μm long. How many base pairs and how many complete turns does it have?

d. A stretch of DNA has 500 base pairs and 1,300 hydrogen bonds between its strands. How many A–T pairs and how many G–C pairs does it have?

e. A protein is made of 250 amino acids. How many nucleotides of mRNA code for it, counting the stop codon? How many base pairs of DNA are needed for that much coding sequence?

f. A woman with normal vision, whose father was color-blind, marries a color-blind man. a. What are the genotypes of the woman and the man? b. What fraction of their sons, and of their daughters, are expected to be color-blind?

g. In a population, 1 man in 20 is color-blind. Estimate how many women in 10,000 are color-blind and how many are carriers.

Found a mistake? Please tell us.