

Genetics

Andrey Nikitin

August 2023

Table of Contents

1	Introduction	3
2	Cell Cycle	3
2.1	Organization of DNA	3
2.2	Stages of the Cell Cycle	5
2.3	Mitosis	5
2.3.1	Additional Features in Plants	8
2.4	Meiosis	9
2.5	Regulation	11
2.5.1	Cyclins	11
2.5.2	Outside regulators	12
2.5.3	Cancer	13
2.5.4	Telomeres	13
2.5.5	Apoptosis	13
3	Inheritance	14
3.1	Mendel's Peas	14
3.1.1	Mendelian inheritance	14
3.1.2	Mendel's Laws	15
3.1.3	Punnett Squares	15
3.1.4	Pedigrees	16
3.1.5	Dihybrid Crosses	17
3.2	Morgan and Chromosomes	18
3.2.1	Sex-linked traits	18
3.2.2	Linked Genes	19
3.3	Types of inheritance	21
3.3.1	Dominant	21
3.3.2	Recessive	21
3.3.3	X-linked	22
3.3.4	Incomplete Dominance	27
3.3.5	Codominance	27
3.3.6	Mitochondrial	28
3.3.7	Aneuploidy	29

3.3.8	Epistasis	30
3.3.9	Duplicate	33
3.3.10	Other	34
3.4	Mutations	35
3.4.1	DNA mutations	35
3.4.2	Chromosomal mutations	37
4	Other Math	37
4.1	Bayes Rule	37
4.2	Coefficient of Relatedness	40
4.2.1	Haplo-Diplo systems	41
4.2.2	Coefficient of Inbreeding	42
4.3	Hamilton's Rule	42
5	Conclusion	43

1 Introduction

You are made of many parts, including your kidneys, toes, and the eyes you're using to read this page. Each of these parts was coded for by your DNA. A **gene** is a specific segment of DNA that codes for a specific protein. These proteins go on to make up your **phenotype**, which is all of the observable things that make you appear like you. Similarly, your **genotype** is the collection of all the genes, which you do not directly see, that codes for who you are. Genetics seeks to understand the connection between genotype and phenotype and how we pass on our traits to future generations.

2 Cell Cycle

Let's begin by understanding how genetics works on the smallest scale possible, the cell.

2.1 Organization of DNA

DNA is located in the **nucleus** of the cell in the form of **chromatin**, which is a loose spaghetti-like mass. Although this description seems messy, chromatin is actually highly organized. Loops of chromatin are called **TADs** (Topologically Associated Domains) and TADs that code for similar functions are often grouped together in special regions of the nucleus known as **transcription factories**.

Chromatin comes in two types: **euchromatin** (eu = true) is the looser version of chromatin, and **heterochromatin** is the denser form of chromatin. Since euchromatin is looser, it is easier for enzymes to access, which plays an important role in gene regulation. The packing of DNA into chromatin is due to special positively-charged proteins called **histones**. There are 5 types of histones: H1, H2a, H2b, H3, and H4. A **nucleosome** is a chunk of 8 histones (2 of each type except for H1). H1 is used to pack nucleosomes together into 30nm fibers. Since DNA is negatively charged, it is attracted to the positively charged histones and will wrap around the nucleosomes, forming chromatin.

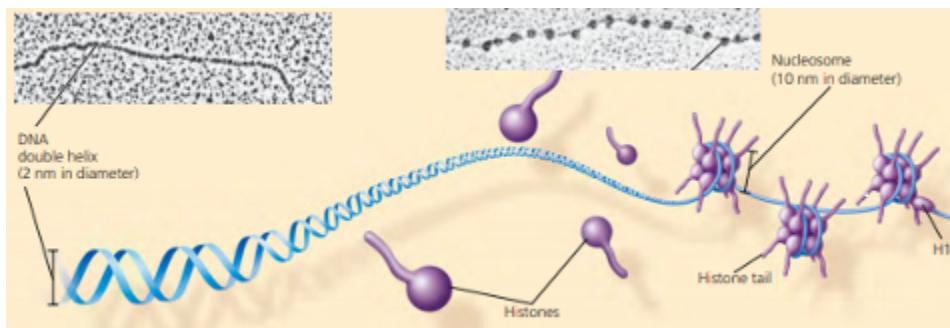


Figure 1: DNA is packed into chromatin by wrapping around nucleosomes.
(Source: Campbell's 9th Edition)

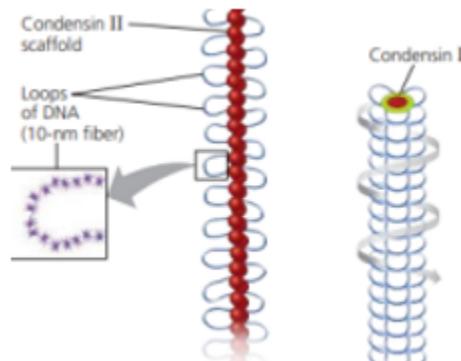


Figure 2: Chromatin is packed into chromosomes via condensins.

(Source: Campbell's 12th Edition)

When the cell prepares to divide, the DNA will pack into even tighter structures called **chromosomes**. This packing is done by special proteins called **condensin II** and then **condensin I**. These structures are so dense that they can be easily seen under a microscope! In fact, a **karyotype** is just a picture of all the chromosomes in an organism lined up. We usually consider genes on the chromosomal level: specific genes are located on certain chromosomes, and each species has their own number of chromosomes. A human, for example, has 23 sets of chromosomes, numbered 1 through 23. However, chromosome 23 is what's known as a **sex chromosome**, which is the chromosome that contains the genes for determining the sex of an individual. The remaining chromosomes are known as **autosomes**.

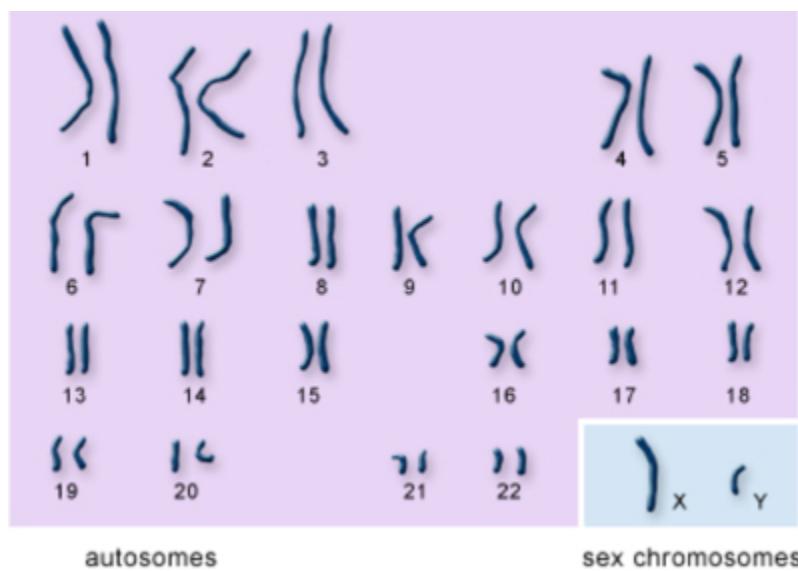


Figure 3: Karyotype of a human. (Source: Campbell's 9th Edition)

Notice that I said humans have 23 *sets* of chromosomes. That's because for each chromosome, we have two ($23 \times 2 = 46$ in humans): one from our mom and one from our dad. This is known as **diploid** and most of the cells in our body, known as **somatic cells**, are diploid. The other type of cell contains only 1 of each chromosome (23 in humans), which makes it **haploid**. Haploid cells in our body are called **gametes**. When two organisms reproduce, their gametes fuse together to form a single, diploid cell ($23 + 23 = 46$).

2.2 Stages of the Cell Cycle

The life of a cell is split into two phases: **interphase**, the stage when a cell is growing and performing its daily functions, and **mitosis**, which is when reproduces by splitting in two.

A cell will spend the majority of its life in interphase, which can be split into 3 phases:

- **G1** (stands for Gap 1, but you can think of it as Grow 1): During this phase, the cell is growing, producing proteins to carry out its normal functions, and storing energy to prepare for the next phase.
- **S** (stands for synthesis): During this phase, the cell replicates its DNA so that it now has 2 copies of every chromosome.
- **G2** (stands for Gap 2): During this phase, the cell is growing, producing proteins, and storing energy to prepare for mitosis.

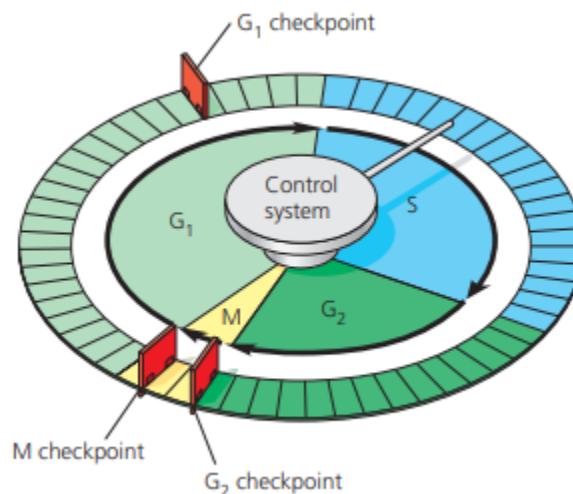


Figure 4: Diagram of the cell cycle. Note how the cell spends most of its time in interphase.
(Source: Campbell's 9th Edition)

In addition, some cells (such as neurons and muscle cells) that are never meant to divide can enter a special phase outside of the cell cycle known as **G0**, where the cell simply exists, performing its normal functions without ever having to worry about mitosis.

2.3 Mitosis

In mitosis, the single *parent* cell will divide to become two *daughter* cells that are equal in material. In a sense, these daughter cells are clones of the parent cell, since they contain the exact same genes.

Simple cells like bacteria practice **binary fission** (binary = 2, fission = split), where the cell elongates as its chromosome is replicated and literally just splits in half. For other organisms, mitosis is far more complicated.

Mitosis is also split up into phases. For these phases we will consider **n** (the number of chromosomes), and **c** (the amount of DNA) as a useful metric. A cell in G₁ has 2c and 2n since it has 2 copies of each chromosome (one from each parent). The phases of mitosis are:

- **Prophase:** The nucleolus disappears. Chromatin condenses into chromosomes. Since each chromosome was duplicated during S phase, the chromosomes are now made up of 2 identical **sister chromatids** joined together at the **centrosome**, forming an X-ish shape. In addition, special proteins called **cohesins** run along the sister chromatids, tying them together. Each sister chromatid is a copy of the DNA, but since they are attached to each other, 2 of them are considered a single chromosome. Therefore, the cell is at $2n$ and $4c$.

The cell has 2 **centrosomes**, which are responsible for producing the **mitotic spindles** (microtubules) that will move the chromosomes around during mitosis. Each centrosome is made of 2 **centrioles**. During this phase, short mitotic spindles extend radially from the centrosome. This structure is known as an **aster**.

- **Prometaphase:** This phase is often just lumped together with Prophase. The nuclear membrane disintegrates. **Kinetochores** form at the centromeres of the chromosomes. Mitotic spindles attach to the chromosomes at these kinetochores. At the same time, centrosomes move to opposite sides of the cell.

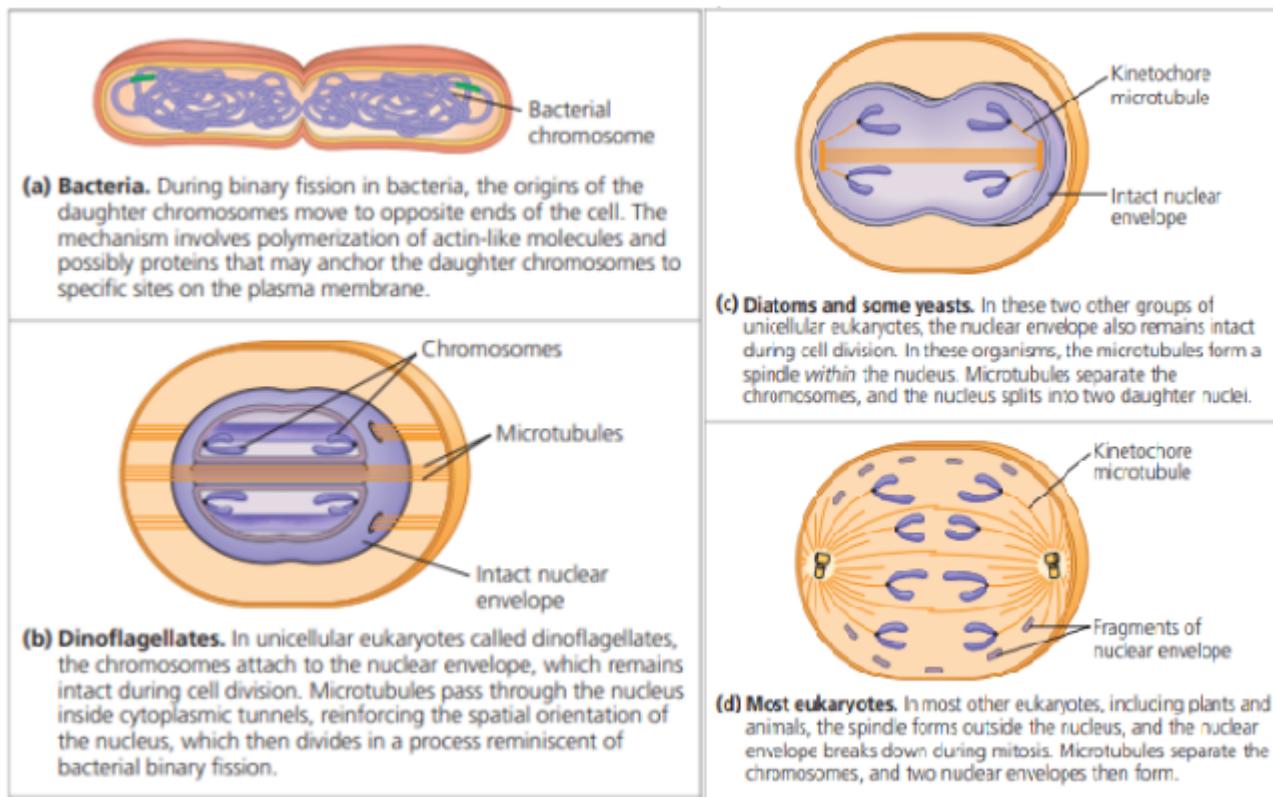


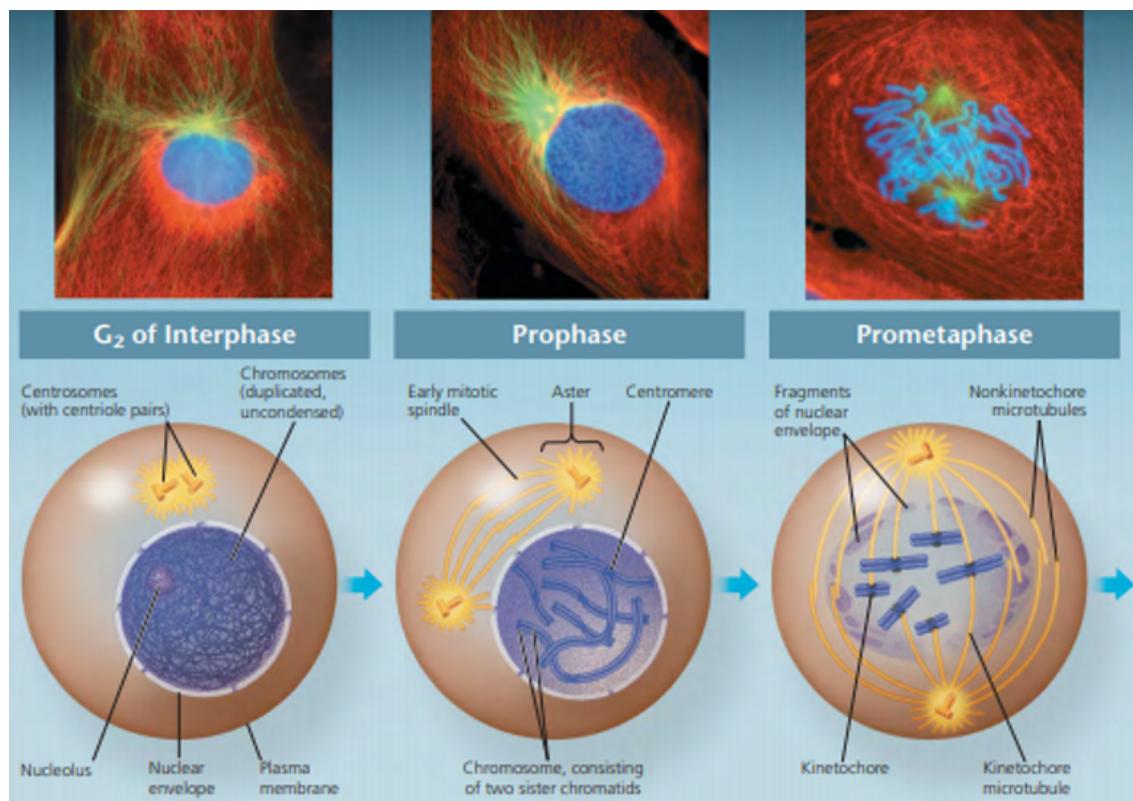
Figure 5: The nuclear membrane does not disintegrate in all organisms, as the diagram above shows. (Source: Campbell's 12th Edition)

- **Metaphase:** The microtubules push and pull the chromosomes until they are all aligned in the **middle**. This imaginary line that they line up on is called the **metaphase plate**.
- **Anaphase:** Sister chromatids are pulled **apart** to opposite sides (poles) of the cell! Since the sister chromatids are no longer attached to one another they are each considered to be

their *own* chromosome, effectively doubling the number of chromosomes in the cell. This brings us to **4n** and **4c**.

There are two different models for how the microtubules shorten in length to pull the sister chromatids closer to the poles of the cell. They can either shorten at the centrosome (named the “reeled in” method), or at the sister chromatid (named the “pac man” method because it looks like the sister chromatid is moving along the microtubule, eating it).

- **Telophase:** The nuclear membrane reforms into *two* nuclei on opposite poles of the cell and the nucleoli reappear. The chromosomes begin to unravel into chromatin and any remaining microtubules are depolymerized. At this point, the genetic material has been divided into two nuclei, each of which is **2n** and **2c**, and mitosis is officially complete.
- **Cytokinesis:** The splitting of the cytoplasm of the cell in two. This step closely follows mitosis. In animal cells, a **cleavage furrow** forms, causing the cell membrane to pinch off in the middle.



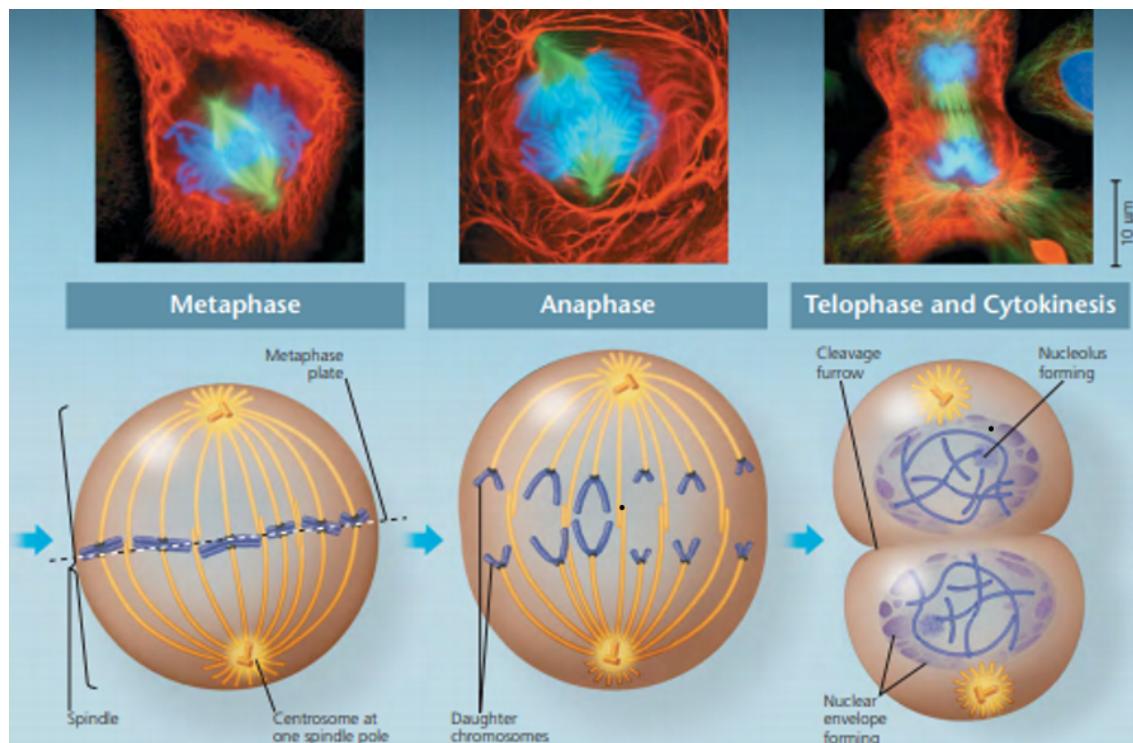


Figure 6: The phases of mitosis showing a diagram (bottom) and an actual image taken under a microscope (top) for each phase. (Campbell's 12th Edition)

2.3.1 Additional Features in Plants

Due to the rigidity of the cell wall in plants, plant cells have evolved unique features in mitosis:

- **Endoreduplication** is when a cell replicates its DNA (S phase) multiple times. This results in a condition known as **polyploidy**, where the cell has more than 2 copies of each chromosome, which is common in many plants. Some notable examples are that seedless watermelon is triploid and wheat is hexaploid.
- The **Phragmosome** anchors the nucleus to the midline of the cell, determining the future plane of division.
- The **Preprophase band** is a band of microtubules encircling the plane of division. It appears in G2, just before prophase.
- The **Phragmoplast** is like mitotic spindle fibers, except it holds the cell plate in place. The **cell plate** accomplishes the role that the cleavage furrow did in animal cells, splitting the cell in half by becoming the cell wall.

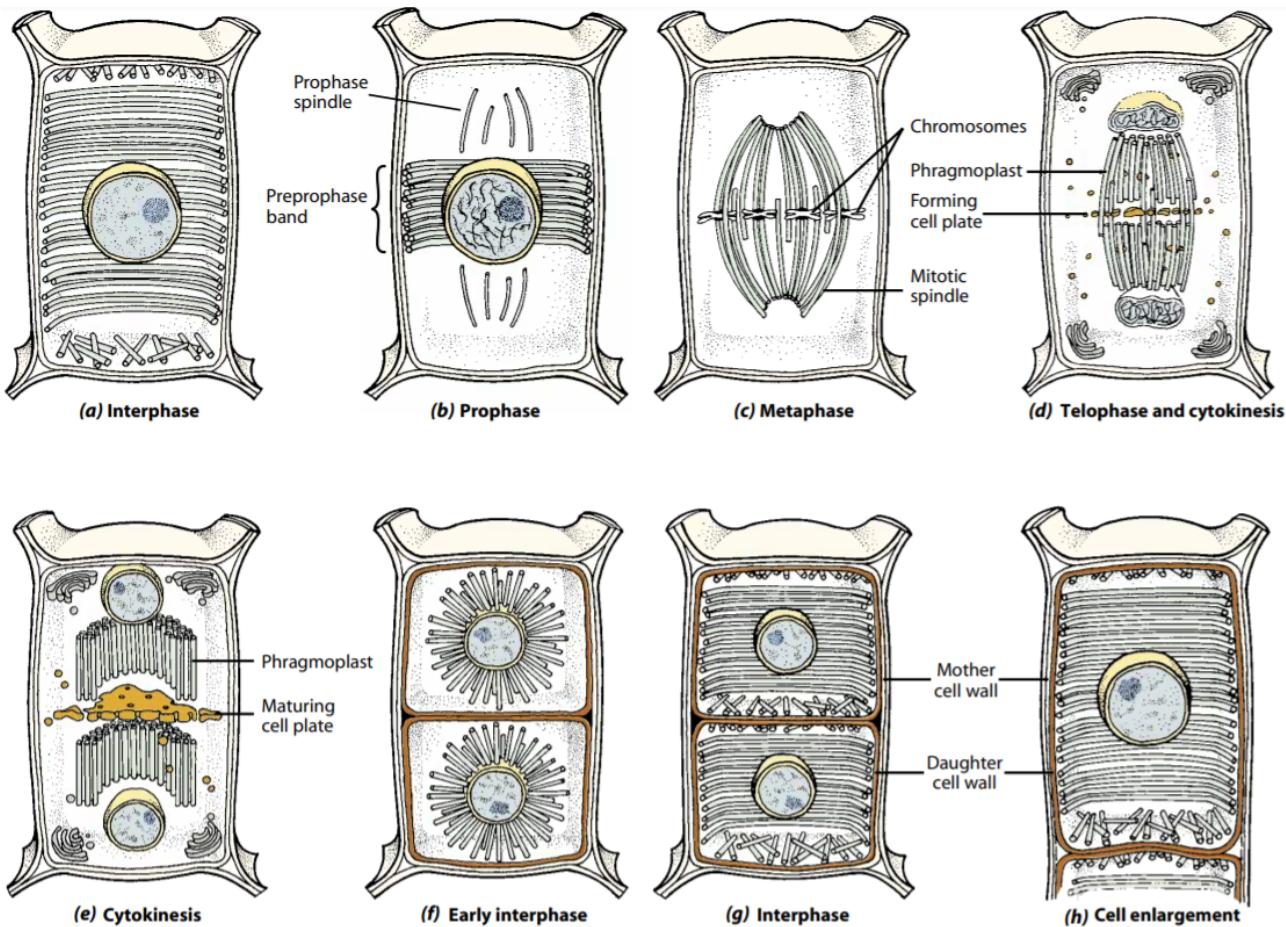


Figure 7: Stages of mitosis in plant cells. (Source: Raven)

Example 2.1: (USABO Opens 2017)

35. Retinoblastoma protein (pRb) prevents the G1-> S transition when it is active. You are analyzing the cell cycle of a somatic cell (2n). What is the DNA content of a mutant somatic cell that has a constitutively active pRb?

- A. n (analogous to the state in the gamete).
- B. 2n (analogous to somatic cells which have just completed cytokinesis).
- C. 4n (analogous to mid anaphase).
- D. 2n (analogous to early prophase).
- E. 4n (analogous to mid telophase).

Solution: If the cell can not enter S phase, then it never replicates its genome. As a result, it will be at 2n and 2c, so the **answer is B**. D would be 2n and 4c.

2.4 Meiosis

Mitosis brought our cells from 4c to 2c. In order to create haploid cells (with 1c), we need to perform another round of mitosis. **Meiosis** produces 4 unique haploid daughter cells from a single diploid parent cell.

Chromosomes come in sets, with one from each parent. After DNA replication, each chromosome is made up of two sister chromatids joined at the centromere. The two chromosomes in the set are known as **homologous chromosomes** (homo = same, logous = location) of each other since they contain the same genes in the same location. What differs between the homologous chromosomes is what value they have for each gene. Each gene typically comes in two forms, known as **alleles**. For example, a pea plant can be either green or yellow. These are the alleles of the color gene.

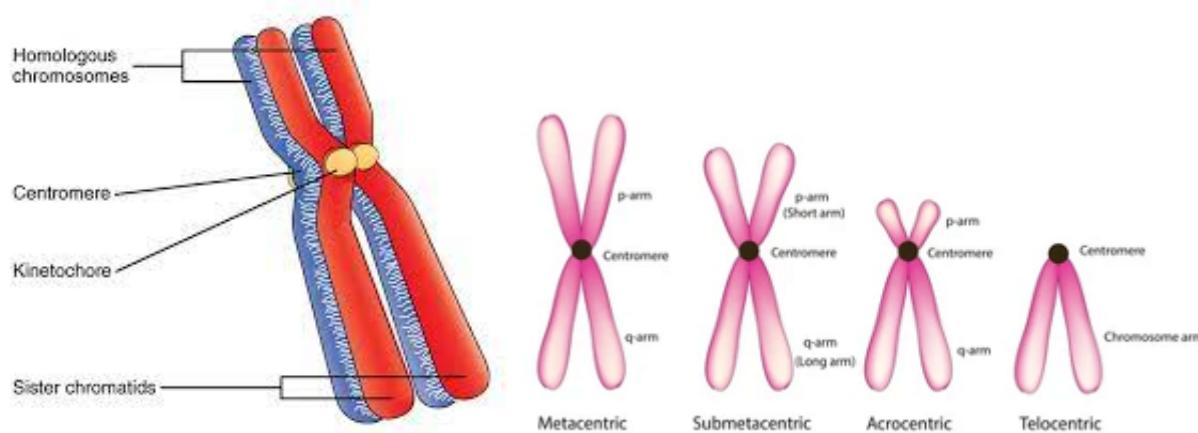


Figure 8: (left) Diagram of a tetrad. (right) Types of chromosomes based on location of centromere. (Source: Adobe Stock)

Meiosis is essentially just mitosis two times in a row. As a result, the steps of meiosis are named the same as the steps of mitosis. During Prophase I (prophase of the 1st round of mitosis), the homologous chromosomes are connected together. The pair of homologous chromosomes together is called a **tetrad** (tetra = 4. There are 4 sister chromatids because there are 2 for each chromosome). The homologous chromosomes fuse together at a point called the **chiasma** and the proteins that connect them are called the **synaptonemal complex**. This fusing is called **synapsis** and can cause the homologous chromosomes to exchange pieces, a process known as **crossing over**. Crossing over during Prophase I is important, as it can lead to genetic variation.

Prophase I is split into its own phases, but you don't need to know them.

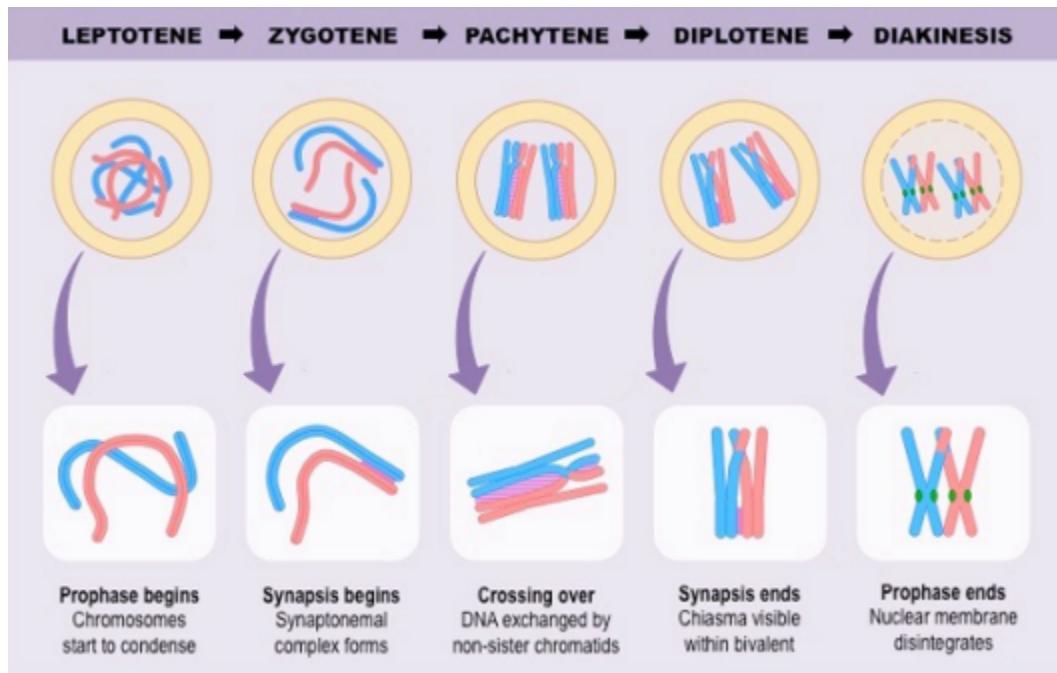


Figure 9: The stages of Prophase I. (Source: BioNinja)

Anaphase I involves the separation of homologous chromosomes while Anaphase II involves the separation of sister chromatids.

2.5 Regulation

Both S phase and mitosis are energy-intensive processes that disrupt a cell's normal function. Therefore, it is important that a cell has the necessary materials to carry out the phase. This is where regulation of the cell cycle is important.

2.5.1 Cyclins

The cell cycle has various checkpoints at crucial moments:

- The **G1 checkpoint** prevents cells from entering S phase until they have the sufficient materials to properly replicate DNA. This is also where cells can leave the cell cycle to enter G0.
- The **G2 checkpoint** prevents cells from entering mitosis until they have the sufficient materials to properly divide.
- The **M checkpoint** makes sure that the chromosomes are properly aligned at the metaphase plate before proceeding to anaphase. If something goes wrong during mitosis, the cell will be arrested at metaphase, preventing the cell from creating defective offspring.

So those are the checkpoints, but how does the cell determine when it is ready to move on? That is accomplished by special proteins called **cyclins**. Each checkpoint is associated with its own **cyclin-dependent kinase**. As a cell proceeds through its phases, it synthesizes cyclins, which slowly build up. This provides the cell time to prepare for the next phase. Once the cyclins reach a high enough concentration, most of the CDKs will be bound to a cyclin, giving the cell the signal to move past the checkpoint.

The most important CDK is associated with the G₂ checkpoint. When this CDK binds to its respective cyclin, it becomes **MPF** (maturation promoting factor), which allows the cell to advance into mitosis.

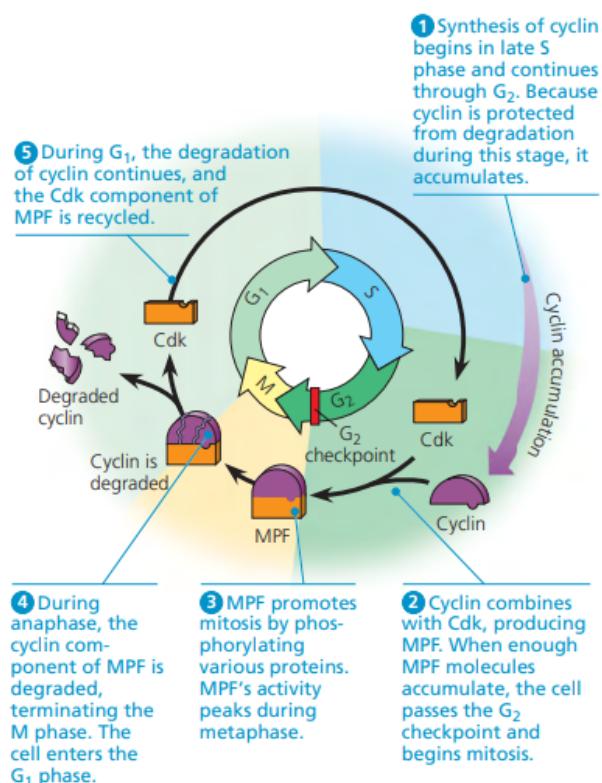


Figure 10: Diagram of how cyclins work.
(Source: Campbell's 9th Edition)

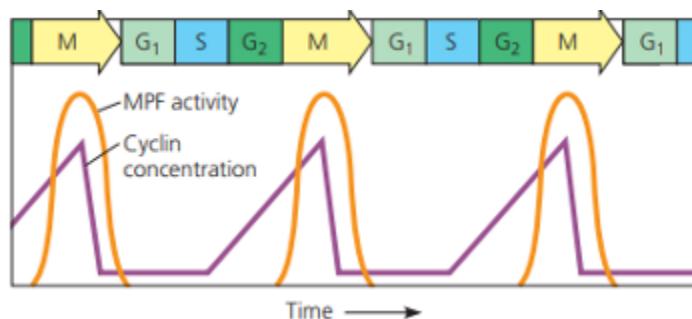


Figure 11: Chart showing concentration of MPF and its respective cyclin throughout the cell cycle. (Source: Campbell's 9th Edition)

2.5.2 Outside regulators

Cyclins are an process internal to the cell that regulates its life. However, cells can also receive cues from the outside world:

- **Growth factors** promote the cell to move quickly through the cell cycle, as more cells growing will result in more growth.
- **Anchorage dependence:** Some cells require to be attached to a surface before they can replicate. This is often seen in bacteria that will only proliferate after they've found a suitable environment (such as the lungs).

- **Density-dependent inhibition:** If a region is already crowded with cells, then it doesn't make any sense for a cell to proliferate as it would only be stealing resources from its neighbors. In this case, inhibition of the cell cycle increases as the local density of cells increase. This is often seen in bacterial cultures that use **quorum sensing** to get a feel of how many cells are nearby.

2.5.3 Cancer

Cancer is the unrestricted proliferation of cells, even at the expense of others. Cancer cells blast through their checkpoints, steal nutrients from neighboring cells, and can even activate genes that promote **angiogenesis** (blood vessel formation), to force the body to feed more nutrients to the cancer cells. A cancer cell that stays in place is **benign** and are generally regarded as less dangerous than **malignant** tumors that can spread throughout the body. The spread of cancer is known as **metastasis** and can occur if cancer cells infiltrate blood vessels.

Transformation is the process by which a normal cell becomes a cancer cell. This can be caused by mutations in genes that regulate the cell cycle:

- **Proto-oncogenes** are normal genes that promote the growth of a cell. However, when a proto-oncogene is mutated it can transform into an **oncogene**, which is like a proto-oncogene on overdrive, causing the cell to proliferate at a crazy rate. Since you only need 1 copy of the mutant gene for the cell to proliferate quicker, these mutations exhibit a *dominant* phenotype (more later). Examples include **Ras genes**.
- **Tumor-suppressor genes** are normal genes that inhibit the cell cycle, preventing the formation of tumors. However, when a tumor-suppressor gene is mutated, there is no longer anything stopping the cell from blasting through the cell cycle. So, it transforms into a cancer cell. Since you need both copies of the tumor-suppressor gene to be inactive (as long as one of them works it will inhibit the cell cycle), these mutations exhibit a **recessive** phenotype (more later). Examples include **p53**, which acts on p21 to inhibit the cell cycle.

2.5.4 Telomeres

Every time DNA is replicated, a portion at the end is lost to a primer (see DNA handout for more info on DNA replication). Therefore, to prevent valuable genetic information from being lost, chromosomes typically have **telomeres**, stretches of noncoding DNA at the end of chromosomes.

A cell can replicate only so many times until it loses all of its telomeres, a limit known as the **Hayflick limit**. However, the enzyme **telomerase** can extend this limit by adding new telomeres.

2.5.5 Apoptosis

Apoptosis is the controlled death of a cell. One pathway for apoptosis has **Bax channels** form in the mitochondria, allowing **cytochrome c** to be released into the cytoplasm, signaling apoptosis to begin. When a cell undergoes apoptosis, a process called **blebbing** occurs, where the cell breaks up into a bunch of tiny blobs and then lysosomes safely break-down those self contained blobs.

A cell can also die as a result of **necrosis**, or unintentional death due to injury to the cell. Necrosis is bad since it results in inflammation.

Example 2.2: (USABO Opens 2017)

10. The progression of cancer is often driven by genetic and epigenetic alterations. Which of the following is not likely to contribute to tumorigenesis?

- A. Loss-of-function mutation in the p53 gene.
- B. Upregulation of pro-apoptotic proteins.
- C. Duplication of the telomerase gene.
- D. Increased secretion of growth factors.
- E. Moderately high level of oxidative stress.

Solution: A does contribute since p53 is a tumor-suppressor gene. Cancer cells would want more telomerase so that they can divide past their Hayflick limit, so C is not true. D would cause more cell growth. E would increase mutation rate, making cancerous mutations more likely. The **answer is B**, as apoptosis results in cell death, which would prevent tumorigenesis.

3 Inheritance

3.1 Mendel's Peas

The first person to propose that traits are inherited through genes was **Gregor Mendel**, a German monk. In his experiments with **pea plants**, he took true-breeding (when you breed it, it'll always produce the same phenotype) pea plants of opposing traits and *crossed* (bred) them together to see what happened. For example, the **P generation** (parents) could be a true-breeding yellow pea plant and true-breeding green pea plant. Strangely, the offspring generation (**F₁**) consisted entirely of yellow peas! What was even weirder is that when Mendel crossed two F₁ yellow pea plants together, there were yellow *and* green pea plants in the offspring generation (**F₂**). Specifically, yellow and green plants appeared in a **3:1 ratio** in the F₂ generation. Where did the green plants go in F₁? How did they come back in F₂? And why was there always that 3:1 ratio???

Mendel repeated this experiment with a bunch of other traits (round vs wrinkled pea shape, purple vs white flower color, tall vs short plant height, etc.) and found this same 3:1 pattern showing up for every single trait. This evidence led Mendel to come up with the **Gene idea**, the theory that special segments of DNA called **genes** code for specific traits and that each gene can have multiple values, or **alleles**. For example, the pea color gene could have either a yellow allele or green allele. Remember that most cells are diploid, and therefore will have two copies of the same gene on homologous chromosomes, although they can have different alleles for that gene.

3.1.1 Mendelian inheritance

Alleles come in two types: dominant and recessive. A **dominant allele**, such as yellow peas, will *mask* the presence of a recessive allele. Therefore, you only need 1 copy of a dominant allele to display that phenotype. A **recessive allele**, such as green peas, gets masked by the dominant allele. Therefore, you need both copies of a recessive allele to display that phenotype.

In naming these alleles, we typically use a capital letter to represent the dominant allele ("Y" for yellow peas) and the lowercase version of that letter for the recessive allele ("y" for green

peas). When an individual has two copies of the same allele, they are **homozygous**. Such an individual can be *homozygous dominant* (**YY**) or *homozygous recessive* (**yy**). When an individual has different alleles, they are **heterozygous** (**Yy**).

3.1.2 Mendel's Laws

Through his experiments, Mendel came up with two laws that dictate how alleles are inherited:

- Law of segregation:** The two alleles for a heritable trait segregate (separate from each other) during gamete formation and end up in different gametes. This is because in Anaphase I, homologous chromosomes (each one containing one of the alleles) are pulled to opposite poles of the cell. Since it is random which side the chromosome is pulled to, offspring are equally likely to receive either of the parent's alleles.
- Law of independent assortment:** each pair of alleles segregates independently of each other pair of alleles during gamete formation. This basically means that you should treat each trait separately. This makes sense as genes are located on different chromosomes and one pair of homologous chromosomes has no affect on a different pair of homologous chromosomes during meiosis.

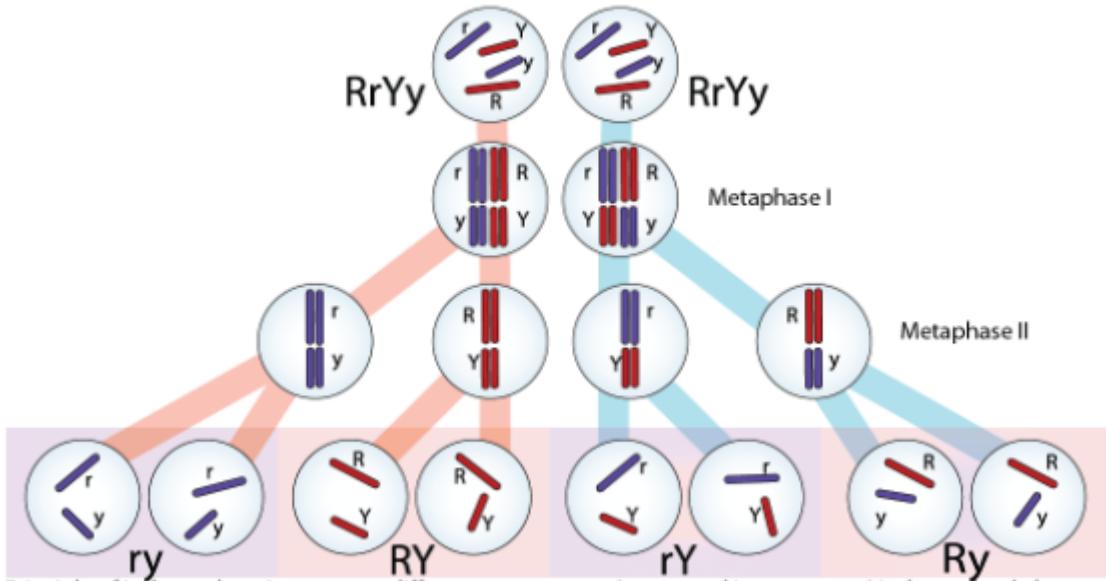


Figure 12: A dihybrid cross involving pea color (Y and y for yellow and green) and pea shape (R and r for round and wrinkled). Law of segregation is shown by the fact that in metaphase II, each cell has either one homologous chromosome or the other (y or Y). Law of independent assortment is shown by the fact that the two possible orientations (the red path and blue path) of the two sets of homologous chromosomes is equally likely. (Source: Biological Principles)

3.1.3 Punnett Squares

When trying to predict the results of a cross, we can use a helpful tool known as a **Punnett square**. By convention, the top row represents the alleles of the father and the left column represents the alleles of the mother. To find the genotype of each square within, simply take the allele from the respective row and column and put them together!

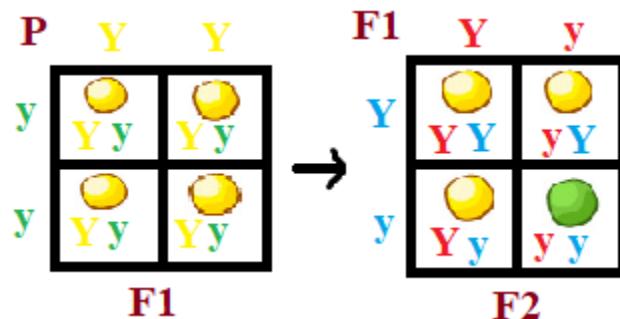


Figure 13: Punnett square showing the crosses in Mendel's experiment with pea color.
(Source: ANdrey)

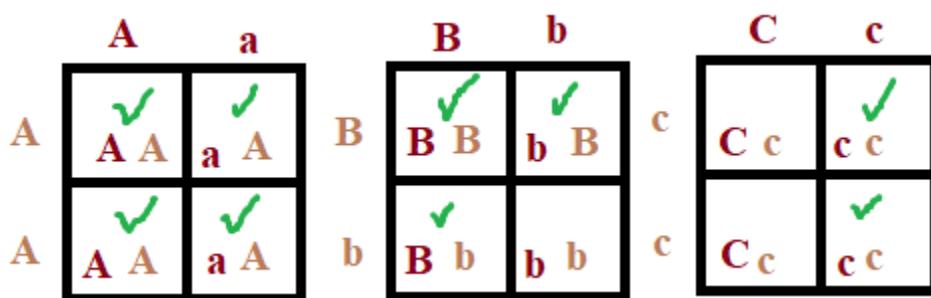
Through this Punnett square, we can explain the **3:1 phenotypic ratio** resulting from crossing two heterozygotes. In addition, we also see the **1:2:1 genotypic ratio** (homo dom : hetero : homo rec) resulting from this cross.

Example 3.1: (USABO Semifinals 2018)

31. What fraction of the offspring of **AaBbCc x AABbcc** will have the dominant **A**, dominant **B**, and recessive **c** phenotype, assume independence between all genes?

- A. 1/8
- B. 1/2
- C. 1/4
- D. 3/8
- E. 5/8

Solution: Due to the law of independent assortment, we can treat each gene separately and then multiply the probabilities together. $1\left(\frac{3}{4}\right)\left(\frac{1}{2}\right) = \frac{3}{8}$, the **answer is D**.



3.1.4 Pedigrees

Punnett squares are useful for understanding the genetics of a single cross. However, when you want to map the genetic history of an entire family, **pedigrees** are used instead. Like a family tree, a pedigree lines up individuals along rows according to generation. Females are represented as circles and males as squares. Shaded shapes mean that the individual has the trait while empty shapes mean that the individual doesn't have the trait. Sometimes, half-shaded shapes will be used to represent heterozygotes.

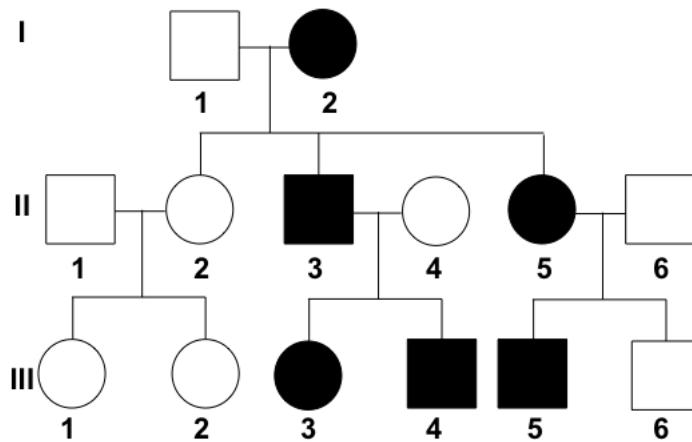


Figure 14: A pedigree. Individuals are typically identified by where they are along a generation (e.g. II-3 is a male with the trait). (Source: Khan Academy)

3.1.5 Dihybrid Crosses

Up until now, we have been dealing with only 1 gene at a time, a **monohybrid cross** (mono = 1, hybrid means that offspring aren't the same as parents). However, we often want to consider 2 genes at a time, a **dihybrid cross**.

We treat a dihybrid cross exactly the same way we did a monohybrid cross, by creating a Punnett square. The only difference is that now there are more combinations for the alleles that a parent can pass to their offspring. For a doubly heterozygote dihybrid cross, the **phenotypic ratio is 9:3:3:1**.

	Y R	Y r	y R	y r
Y R	YY RR Y ^W W ^W			
Y r	YY Rr Y ^W w ^W			
y R	yY RR y ^W W ^W			
y r	yY Rr y ^W w ^W			

Figure 15: The Punnett square for a dihybrid cross for pea color and shape. Each box is shaded to indicate what phenotype it has for each trait. (Source: ANdrey)

3.2 Morgan and Chromosomes

Thomas Hunt Morgan is credited with the **Chromosomal Theory of Inheritance**. In other words, he was the first to show that genes are located on chromosomes. He did this through his experiments where he bred **fruit flies** (*Drosophila melanogaster*).

3.2.1 Sex-linked traits

Genes are located on chromosomes. This means that some genes are located on the *sex chromosomes*. This is important because since those chromosomes also determine the gender of the organism, there *will* be a relationship between the gender of the offspring and their phenotype. These types of genes are known as **sex-linked genes**.

As Morgan bred his flies he noticed something interesting. White eye color in flies is recessive to red eye color. When Morgan bred a true-breeding red female and true-breeding white male, all of F₁ was red-eyed as expected. When he bred F₁ together, there were red eyes to white in a 3:1 ratio as expected. However, all of the white-eyed flies were males! This led Morgan to believe that the eye color gene was located on the X chromosome, making it **X-linked**.

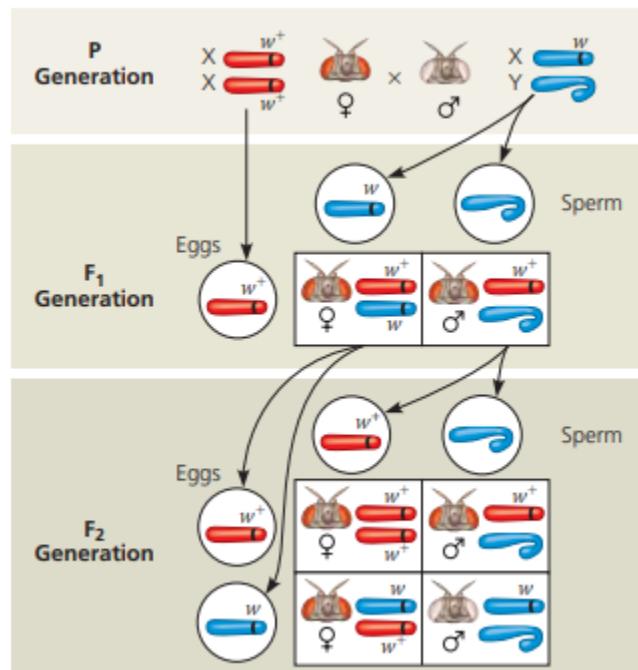


Figure 16: The X-linked inheritance of eye color in Morgan's fruit flies.
(Source: Campbell's 9th Edition)

As the Punnett squares above show, X-linked recessive traits are more likely to appear in males since they only have 1 X chromosome and therefore need only 1 copy of the recessive allele. Since the terms homozygous and heterozygous no longer make sense when you only have 1 allele of the trait, we use the term **hemizygous** to describe these males. Females, on the other hand, have 2 X chromosomes and need 2 copies of the recessive allele.

3.2.2 Linked Genes

As sex-linked genes show, you can have multiple genes on the *same* chromosome. This would violate Mendel's law of independent assortment as those genes would be inherited together. We call these **linked genes**. It is chromosomes that are independently assorted and those chromosomes contain multiple linked genes.

However, that does not mean that those two genes will always be inherited together! Due to **crossing over in Prophase I**, homologous chromosomes can exchange pieces. For the rest of this explanation, I will use the term "top" to refer to the part of the chromosome that is *not* exchanged and "bottom" to refer to the part that *is* exchanged. Even though two genes are on the same chromosome, if one is located on the bottom and the other is on the top, then they won't be inherited together! In fact, whenever crossing over occurs *between* the two genes, they won't be inherited together. We can assume that crossing over is equally likely to happen anywhere along a chromosome. Therefore, the farther apart the two genes are, the more likely this is to happen. When two genes are on opposite ends of the chromosome, then they behave the same as genes on two separate chromosomes.

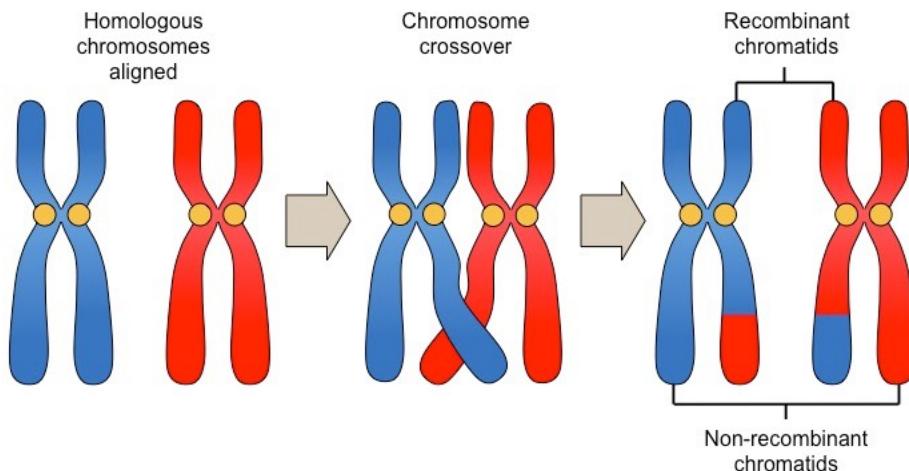


Figure 17: Crossing over in Prophase I. (Source: BioNinja)

To figure out how far apart two linked genes are, we can perform a **test-cross**, which involves breeding a doubly heterozygote with a doubly homozygous recessive. As a result, the phenotype depends entirely on what alleles are received from the heterozygote. The offspring's phenotype can be either a **parental phenotype** or a **recombinant phenotype**. As the name suggests, a parental phenotype matches that of one of the parents (either double dom or double rec) and indicates that the two traits were inherited together. A recombinant phenotype doesn't match that of either parent (dom in one but rec in the other) and shows that crossing over occurred between the two genes.

By figuring out what percent of phenotypes are recombinant, we can figure out how likely it is for crossing over to occur between the two genes. As explained earlier, it becomes more likely the farther apart the two genes are. Specifically, we use **linkage maps** as a linear representation of the chromosome where the distance between any two genes in **map units** reflects the **recombination**

frequency.

$$\frac{\text{recombinants}}{\text{total}} = \text{recombination frequency}$$

Example 3.2: (USABO Opens 2018)

43. The following are the results of a trihybrid cross.

<u>vcv⁺ct⁺</u>	<u>580</u>
<u>v⁺cvct</u>	<u>592</u>
<u>vcvct⁺</u>	<u>45</u>
<u>v⁺cv⁺ct</u>	<u>40</u>
<u>Vcvct</u>	<u>89</u>
<u>v⁺cv⁺ct⁺</u>	<u>94</u>
<u>vcv⁺ct</u>	<u>3</u>
<u>v⁺cvct⁺</u>	<u>5</u>

Which of the following correctly shows the correct chromosomal map and distance (in CM) between the three genes?

- A. v(13.2)ct(6.4)cv.
- B. cv(13.2)ct(6.4)v.
- C. ct(13.2)v(6.4)cv.
- D. ct(6.4)v(13.2)cv.
- E. v(6.4)cv(13.2)ct.

Solution: You will see many questions that require you to do linked genes for 3 genes at a time. For these questions, simply pick two genes at a time and find the distance between them. You will also notice that instead of doing dominant (V) and recessive (v) they do **wild-type** (v^+) and recessive (v). That's just convention for linked genes, so don't worry about it. Since the first two (vcv^+ct^+ and v^+cvct) are the most common, we'll assume that those are the parent phenotypes.

For v and cv we see that $\frac{\text{recombinant}}{\text{total}} = \frac{vcv^++v^+cv^+}{\text{total}} = \frac{(45+89)+(40+94)}{1448} = \frac{268}{1448} = .185 = 18.5\%$

For v and ct we see that $\frac{\text{recombinant}}{\text{total}} = \frac{vct+v^+ct^+}{\text{total}} = \frac{(89+3)+(94+5)}{1448} = \frac{191}{1448} = .132 = 13.2\%$

For cv and ct we see that $\frac{\text{recombinant}}{\text{total}} = \frac{cv^+ct^++cvct^+}{\text{total}} = \frac{(40+3)+(45+5)}{1448} = \frac{93}{1448} = .064 = 6.4\%$

Unfortunately, these numbers don't fit together nicely ($13.2 + 6.4 = 19.6 \sim 18.5$), as we would expect in a real experiment. Now that we know how far apart each pair of genes is, we can put the three of them together like a puzzle so that they all work out. From our results, our **answer is A.**

3.3 Types of inheritance

Not all traits follow Mendelian genetics. You need to be able to identify the different types of inheritances, what their pedigrees look like, and what diseases are associated with them.

3.3.1 Dominant

In a pedigree, you can tell when a disease is dominant if two parents with the disease can produce an unaffected offspring. This is possible when both parents are heterozygotes.

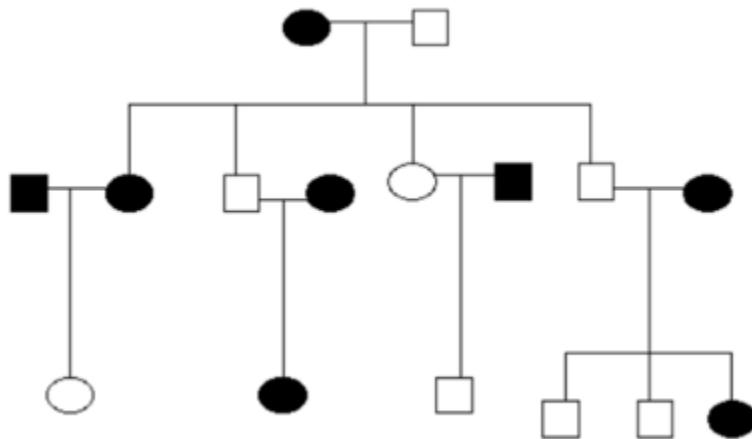


Figure 18: Pedigree of an autosomal dominant trait. Note how II-1 and II-2 both have the disease, but their offspring III-1 does not. (Source: ANdrey)

Most lethal diseases are not dominant, since anyone with the allele would develop the disease and die. However, in **Huntington's disease**, the symptoms (destruction of nervous system) show up later in life, giving people with the disease enough time to reproduce and pass on the allele. Huntington's disease is caused by CAG repeats in the *huntington* gene. In normal individuals, the gene will repeat about 20 times. Those with Huntington's have over 40 CAG repeats.

Achondroplasia, better known as dwarfism, is also dominant.

3.3.2 Recessive

In a pedigree, you can tell when a disease is recessive if two parents without the disease can produce an affected offspring. This is possible when both parents are heterozygotes.

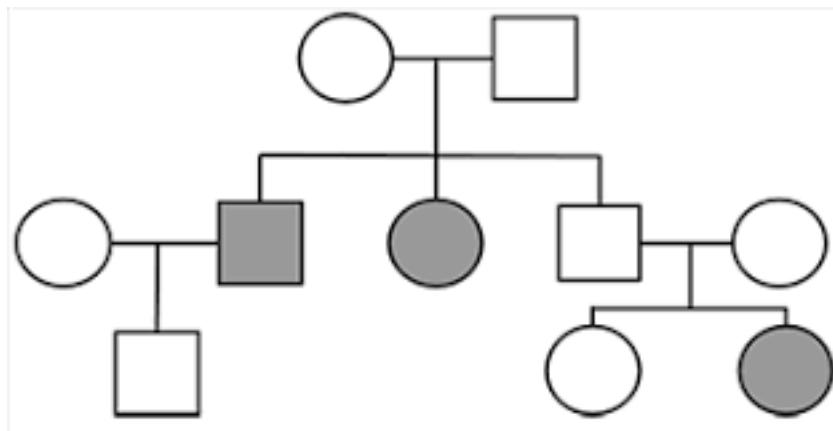


Figure 19: Pedigree of an autosomal recessive disease. Note how I-1 and I-2 both are unaffected, yet they produce affected offspring (II-1 and II-2). (Source: ANdrey)

Many diseases are recessive, due to the ability of the recessive allele to hide in **carriers** (heterozygotes). I remember the major ones as TSC (TS = Tay-Sachs, SC = Sickle cell, C = cystic fibrosis). Here are some common diseases:

- **Tay-Sachs disease:** The enzyme that breaks down lipids in the brain does not work properly, leading to destruction of the nervous system.
- **Sickle-Cell disease:** A point mutation in the 6th amino acid in hemoglobin causes red blood cells to have a sickle shape instead of their usual biconcave shape. The sickle shape is inconvenient for blood flow, leading to pain and blood clots. In addition, it is not able to transport as much oxygen as a normal blood cell.

However, people that are heterozygous for sickle cell have increased resistance to **malaria**, which is caused by the protist *Plasmodium* and transmitted by the **Anopheles mosquito**. This is why sickle cell is especially common in Africa, where malaria is prevalent.

- **Cystic Fibrosis:** This is the most common lethal genetic disease. Greater intracellular $[Cl^-]$ causes cells to absorb more water. This means that mucus secretions become thicker, leading to lung infections and problems with the digestive system.
- **Phenylketonuria (PKU):** Babies are unable to metabolize phenylalanine, but it can be treated with diet changes.

3.3.3 X-linked

In humans, there are two sex chromosomes: X and Y. Females are XX and Males are XY. Since females can only pass on an X chromosome, that means the father determines the gender of the child. The X chromosome is significantly larger than the Y chromosome and contains more genes. As a result, females with 2 X chromosomes sort of have too much genetic information. To deal with this, one of their X chromosomes will be inactivated at random and become a **Barr body**. This inactivation is a result of activation of the **XIST** gene in one of the X chromosomes, producing lncRNA that covers the chromosome. This is known as **dosage compensation**.

Tortoiseshell cats, with random spots of black and orange fur, are a result of Barr body inactivation. Fur color is a gene on the X chromosome and heterozygotes have an allele for orange fur

and an allele for black fur. Since barr body inactivation is random, each patch of fur will randomly be orange or black (due to the other color being inactivated).

Sex-linked diseases most often occur on the X chromosome since it is significantly larger than the Y chromosome and contains more genes. As a result, an individual can either have the sex alleles Y, X^A , or X^a .

In a pedigree, it's easy to determine whether a disease is dominant or recessive. However, figuring out whether it is **autosomal** or **X-linked** often comes down to what is more likely to happen. We can rarely be *certain* whether a disease is autosomal or X-linked, but we can be confident. If you see imbalances in the genders of affected offspring, such as all of the males but none of the females being affected, that is often a clue that the disease is X-linked.

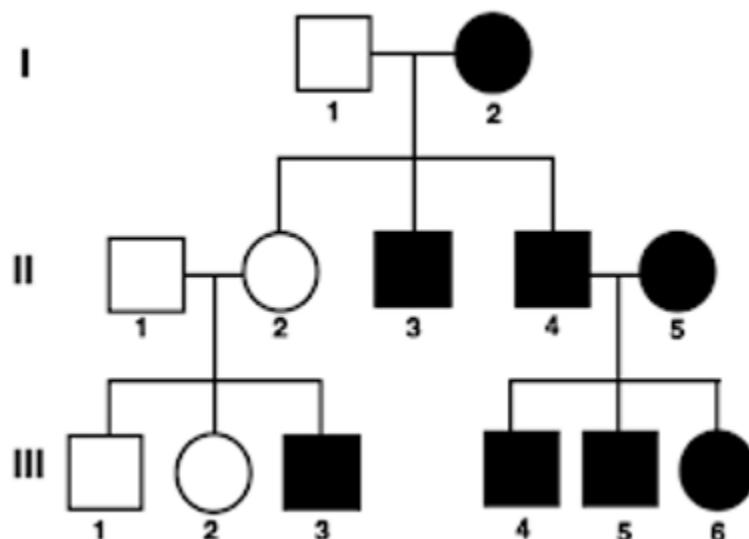


Figure 20: Pedigree of a X-linked recessive disease. First off, we know this is a recessive trait because II-1 and II-2 are both unaffected but produce affected offspring (III-3). Note how I-1 is unaffected ($X^A Y$) but I-2 is ($X^a X^a$), but their cross produces only affected males and no affected females since the mom will always pass on a recessive allele to her sons and the dad will always pass on a dominant allele to his daughters. (Source: Andrey)

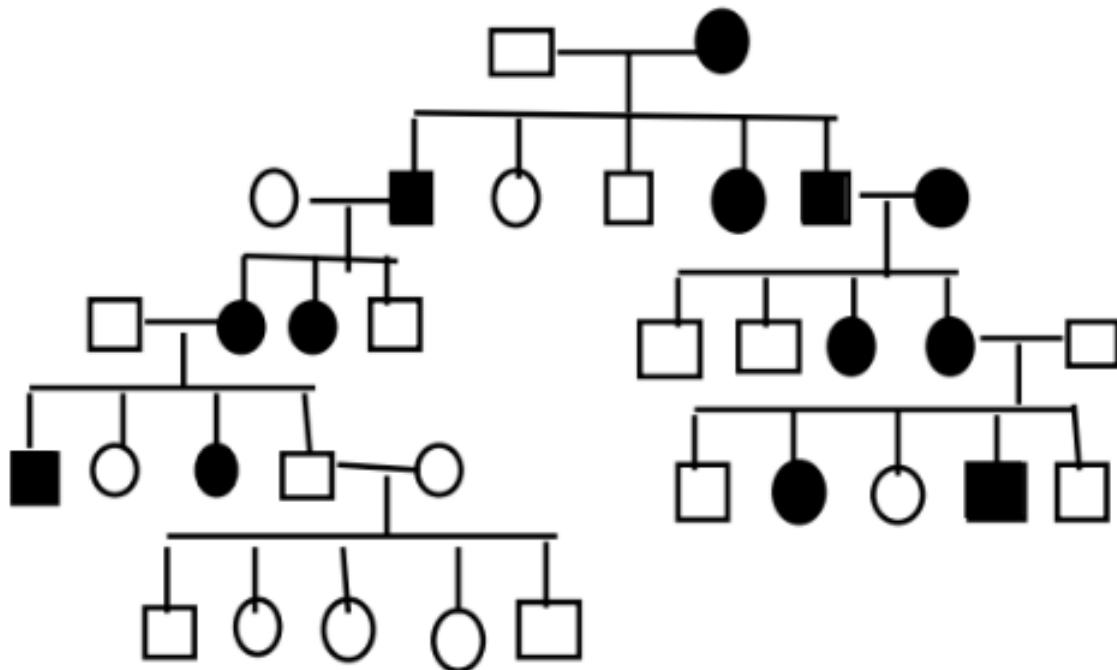


Figure 21: Pedigree of a X-linked dominant trait. First we know it is dominant because two II-6 and II-7 are affected but produce unaffected offspring (III-5 and III-6). As for whether it's X-linked, the fact that affected males always produce affected daughters should make us suspicious. (Source: ANdrey)

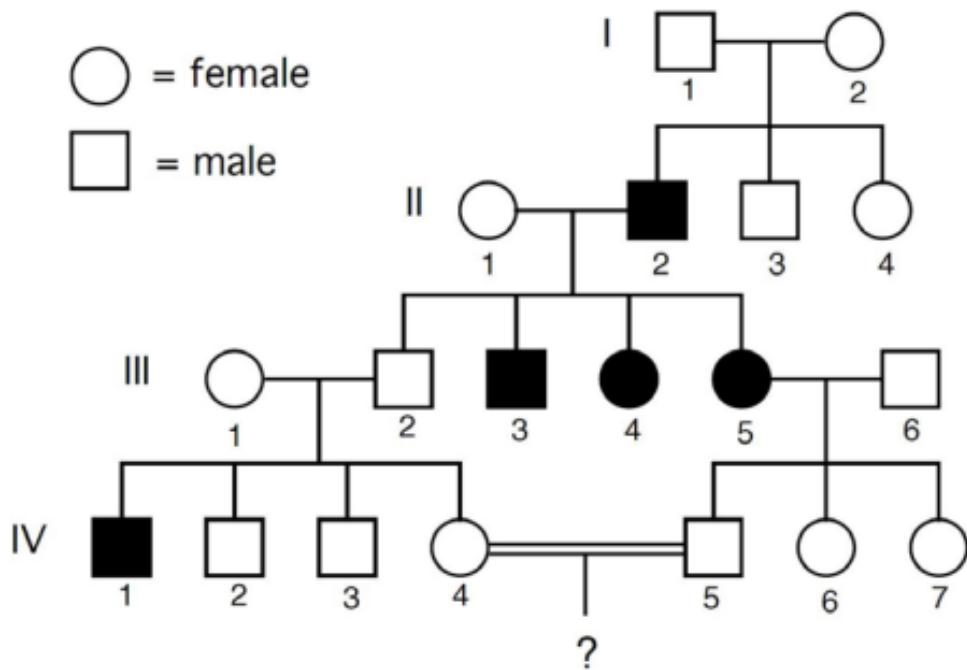
Here are some examples of X-linked recessive disorders:

- **Red-green color blindness**
- **Duchenne muscular dystrophy:** Muscles gradually weaken, causing most people with the disease to die before their 20's.
- **Hemophilia:** Some blood clotting proteins are missing so affected individuals bleed very easily and a lot. In fact, bruises that may seem harmless to us could potentially be life threatening to them.

Hemophilia A (Classic hemophilia) is missing factor VIII while Hemophilia B (Christmas disease) is missing factor IX.

Example 3.3: (USABO Semifinals 2017)

29. The image below depicts the inheritance of a rare skin disorder in four generations of a family; the disorder has 100% penetrance. Which of the following choices describes the most likely mode of inheritance of this skin disorder?



- A. Autosomal dominant.
- B. Autosomal recessive.
- C. Sex-linked dominant.
- D. Sex-linked recessive.
- E. Familial nondisjunction.

Solution: Penetrance will be explained later in mitochondrial disorders, so you can ignore that for now. First, note that the disease must be recessive since two unaffected individuals (I-1 and I-2) can produce affected offspring (II-2). As for whether it's sex-linked, let's assume for now that it is. In that case, the cross between III-5 (X^aX^a) and III-6 (X^AY) *must* produce affected sons, since the mother can only pass on recessive alleles. However, IV-5 is not affected, so the disease must be autosomal. Our answer is B.

While we are used to the sex chromosomes being X for females and Y for males, that is not always the case:

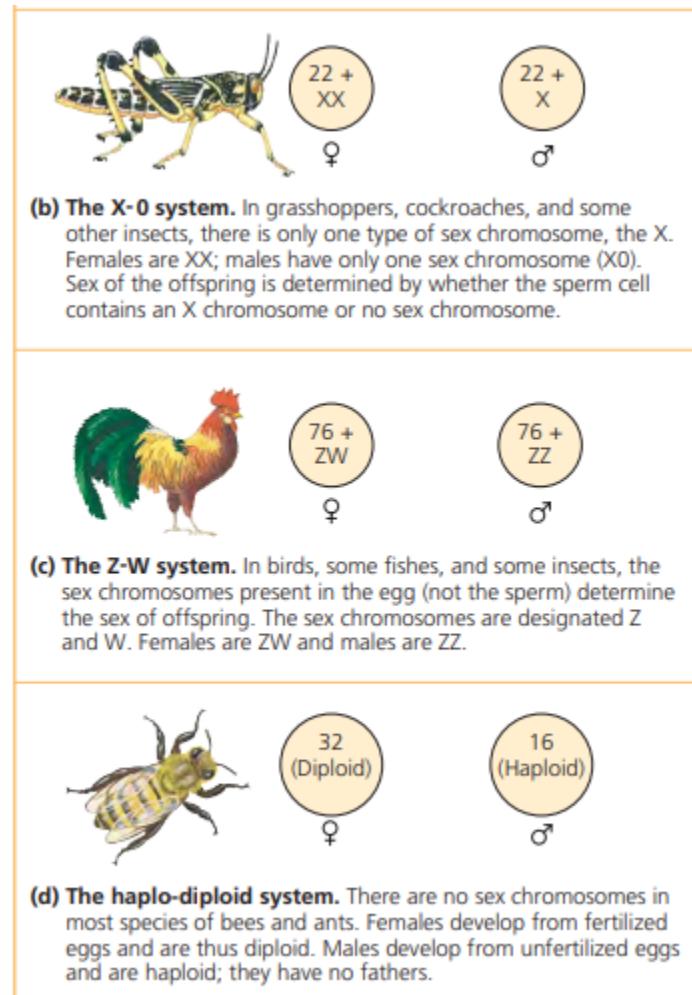


Figure 22: Modes of sex determination in other organisms. (Source: Campbell's 9th Edition)

Example 3.4: (USABO Semifinals 2018)

38. Which of these answer choices matches the organism with the correct chromosomal system of sex determination?

- A. Koala; Y-W
- B. Grasshopper; Haplo-diploid
- C. Chicken; Z-W
- D. Salmon; X-W
- E. Ant; X-0

Solution: By the chart above, the **answer is C.** A and D both have an X-Y system.

Example 3.5: (USABO Semifinals 2018)

34. Your best friend Dr. O is studying a population of *Ursus arctos*, a common mammal. She is interested in two phenotypes: black-haired *Ursus arctos* and blue-haired *Ursus arctos*. It is thought that the blue hair phenotype is the recessive phenotype. Interestingly enough, however, in a sample of 5000 male *Ursus arctos* and 5000 female *Ursus arctos*, Dr. O finds that there are 1600 male blue-haired *Ursus arctos* but only 512 female blue-haired *Ursus arctos*. Which of the following is the best explanation of this phenomenon?

- A. There are two genes that control hair color.
- B. Dr. O used too small of sample size to get a reliable result.
- C. The species *Ursus arctos* exhibits the temperature-dependent sex determination system.
- D. The allele responsible for blue hair is not autosomal.
- E. The blue hair phenotype is not actually the recessive phenotype.

Solution: The fact that a recessive disease is more common among males should make us think it is X-linked. To check, we see that the $1600/5000 = .32$ of males are blue haired while $512/5000 = .1024$. Conveniently, $.32^2 = .1024$, which is what we would expect. The **answer is D**.

3.3.4 Incomplete Dominance

Sometimes there is no dominant trait. Instead, both traits are recessive and heterozygotes display a *mix* of the two traits. For example, in snapdragons, when a red flower (RR) and a white flower (rr) are crossed, they will produce pink flowers (Rr).

3.3.5 Codominance

Sometimes both traits are dominant. In this case, heterozygotes will display *both* traits at the same time. For example, crossing black cows with white cows produces the classic Holstein cow with black and white spots.

An example that displays both regular dominance and codominance is blood types in humans. Humans have 3 alleles for blood type: A (I^A), B (I^B), and O (i). Both A and B are dominant over O. As a result, $I^A I^A$ and $I^A i$ will both be type A, $I^B I^B$ and $I^B i$ will both be type B, and ii will be type O. However, A and B are codominant, so $I^A I^B$ will be type AB and display both type A and type B antigens. Type O is known as a *universal donor* because it lacks A and B antigens, so no blood type will recognize type O blood as foreign. Similarly, Type AB is known as a *universal acceptor* because it has both A and B antigens, so it will recognize type AB, A, B, or O blood as part of the body.

Example 3.6: (USABO Opens 2016)

35. At the end of a perfect week, Barney is suddenly confronted by the horrible news that he is the father of a baby girl. After Barney disappears to Bermuda, the mother orders a paternity test to confirm that Barney truly is the father. If Barney's blood type is AB and Rh+, which of the following blood types could his daughter NOT have?

- A. A and Rh+.
- B. B and Rh-.
- C. AB and Rh-.
- D. AB and Rh+.
- E. O and Rh-.

Solution: We did not talk about Rh factor, but Rh⁺ is dominant to Rh⁻. However, that is irrelevant for this question because an AB father ($I^A I^B$) can not have a type O (ii) child. The answer is E.

3.3.6 Mitochondrial

Besides the genes in the nucleus, there are also **extraplastic or cellular genes**, which are a result of mitochondria and chloroplast having their own DNA. Since it is the egg cell that gives its cytoplasmic contents to the zygote (the sperm cell contributes only the nuclear DNA), **mitochondrial** diseases are always passed through the mother.

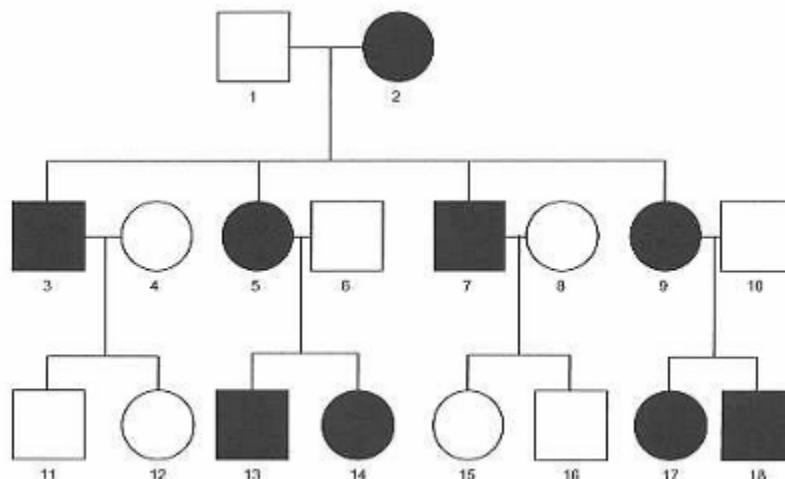


Figure 23: Pedigree of a mitochondrial disease. Note that whenever the mom is affected, *all* of the offspring are affected. (Source: Andrey)

Up until now, we have assumed that alleles *always* express themselves. However, this is not true. There is a phenomenon known as **penetrance**, which is basically a measure of how likely an allele is to be expressed in the phenotype.

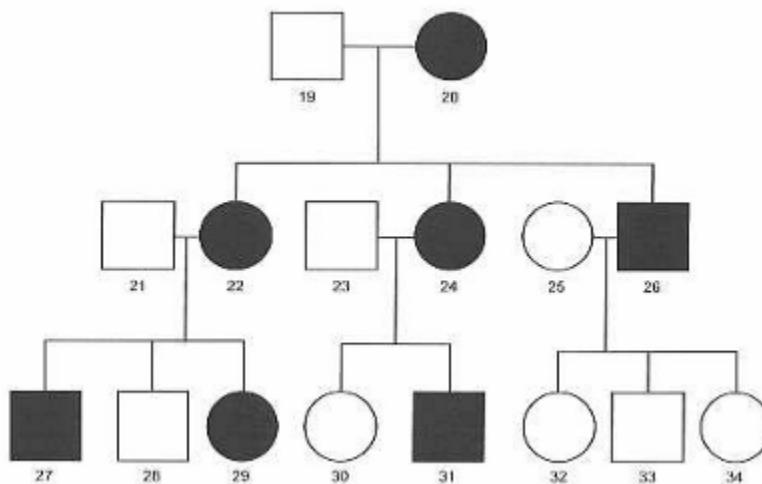


Figure 24: Pedigree of a mitochondrial disease with 75% penetrance. (Source: Andrey)

Mitochondrial myopathy and Leber's hereditary optic neuropathy are examples of mitochondrial diseases.

3.3.7 Aneuploidy

Sometimes during meiosis, the homologous chromosomes or sister chromatids don't separate properly. This can result in a cell having an improper number of chromosomes, which is known as **aneuploidy**. Here are some common diseases that result from aneuploidy:

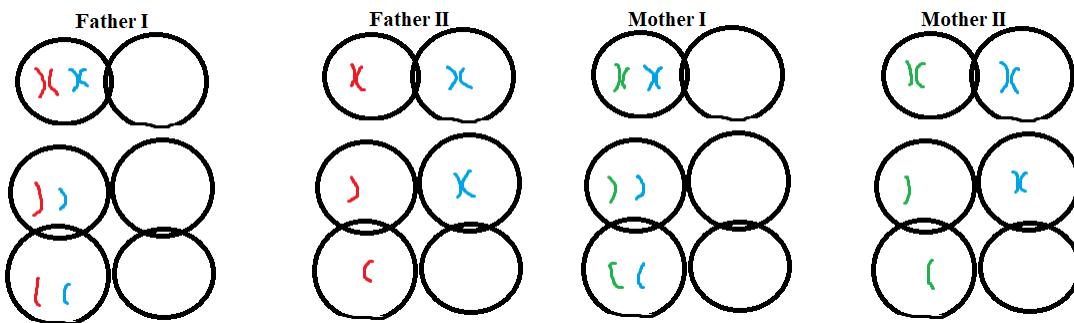
- **Down Syndrome (Trisomy 21):** When you have 3 copies of chromosome 21. Usually not lethal, but results in learning disorders, sterility, and more issues.
- **Klinefelter Syndrome (XXY):** The individual is male, but will have feminine features and be sterile due to the extra X chromosome.
- **Jacob Syndrome (XYY):** Mostly normal male, except usually taller.
- **Trisomy X (XXX):** Mostly normal female, except usually taller.
- **Turner Syndrome (Monosomy X):** When a female lacks an X chromosome (X0). They are mostly normal, but sterile.

Example 3.7: (USABO Opens 2017)

43. John has Downs Syndrome. At a given locus on chromosome 21 in a region with low crossover, John's genotype is ACC. At this same locus, his father has genotype AC and his mother BC. In which stage of his parents' meiosis could nondisjunction have occurred (**Select ALL that apply**)?

- Meiosis I of the father.
- Meiosis II of the father.
- Meiosis I of the mother.
- Meiosis II of the mother.
- None of the above.

Solution: Some questions test your understanding of meiosis by asking you if you can predict where nondisjunction happens to create weird combinations. John needs ACC from his parents. He can receive an A only from his father (AC), while he can receive a C from either his father or mother (BC). The simplest way to solve this is to draw it out. In the diagram below, red is A, green is B, and blue is C.



As you can see, **A works** since we receive an A+C from the father, and we can receive a C from the mother. In addition, **D works** since we receive an A from the father, and we can receive C+C from the mother. There is no other way to get an A and 2 Cs. The **answer is AD**.

3.3.8 Epistasis

Sometimes the expression of one gene can affect the expression of a different gene. This is known as **epistasis**. It comes in recessive and dominant forms, although they are fairly obscure.

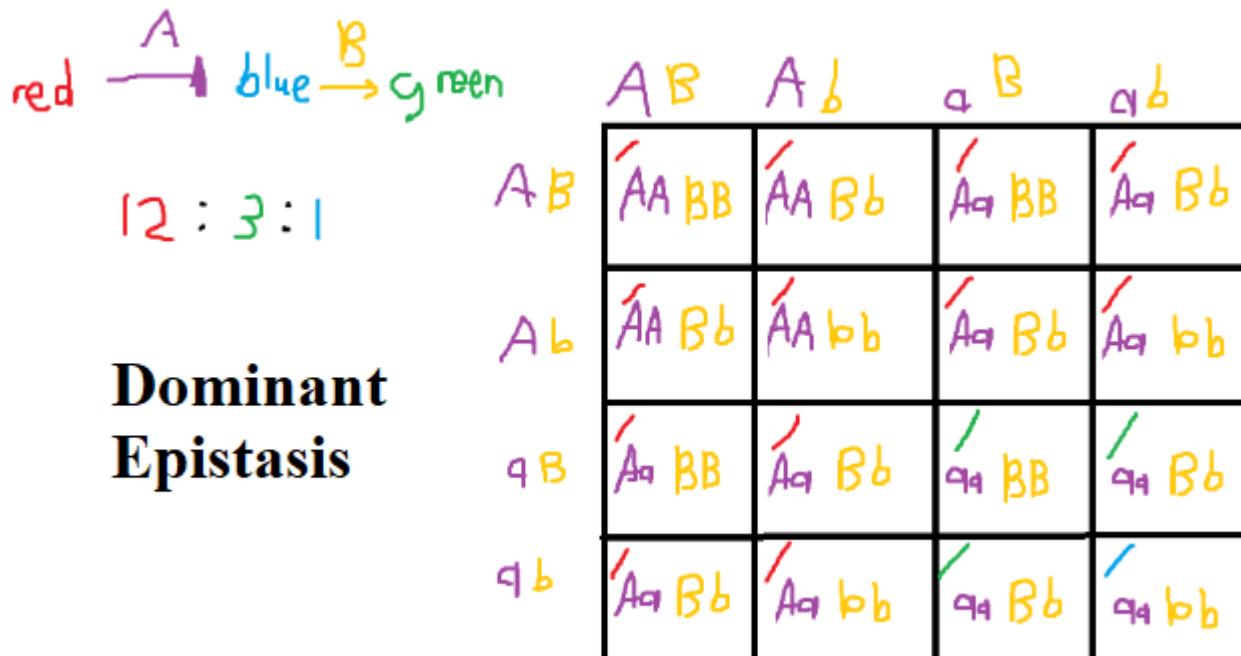


Figure 25: A epistatic dominant disease occurs when the dominant form of one gene (A) masks the effect of the other gene (B). In this diagram, a pathway for a pigment shows how such a setup occurs. As the Punnett square shows, they exhibit a 12:3:1 phenotypic ratio. (Source: ANdrey)

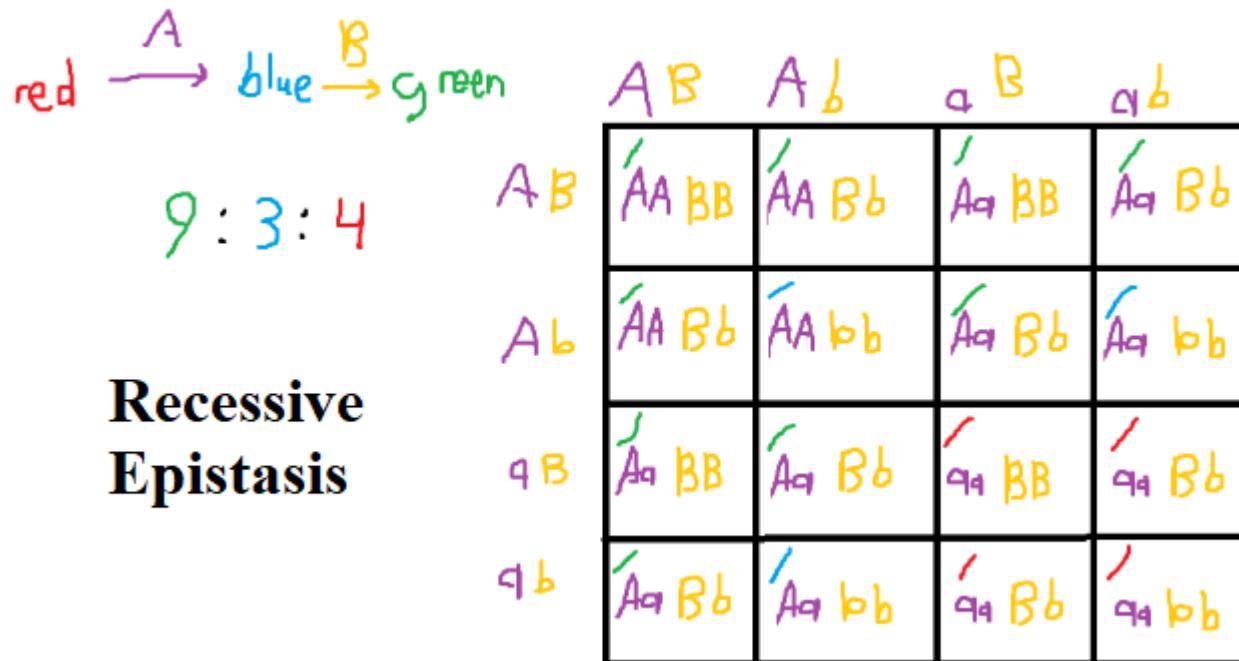


Figure 26: A epistatic dominant disease occurs when the recessive form of one gene (a) masks the effect of the other gene (B). In this diagram, a pathway for a pigment shows how such a setup occurs. As the Punnett square shows, they exhibit a 9:3:4 phenotypic ratio. (Source: ANdrey)

Example 3.8: (USABO Semifinals 2017)

30. In domestic cats, hair color is determined by multiple gene loci. The Brown/eumelanin gene controls the amount of pigment produced in the hair shaft. Cats can have B, b, or c alleles at this locus. Cats with at least one dominant allele (B) are brown. bb and bc cats are chocolate colored, and cc cats are cinnamon colored. The white masking gene is epistatic to other pigmentation genes. WW or Ww cats will have a reduced number of pigment producing cells, and will be white. The color of ww cats will not be affected by the white masking gene. If a heterozygous brown (Bb) cat is crossed with a bcWw cat, what will the expected phenotypic ratio of the kittens be?

- A. 4 White: 2 Brown: 1 Cinnamon: 1 Chocolate.
- B. 2 White: 1 Brown: 1 Chocolate.
- C. 2 Brown: 1 White: 1 Chocolate.
- D. 9 Brown: 3 Chocolate: 3 Cinnamon: 1 White.
- E. 9 White: 3 Brown: 3 Chocolate: 1 Cinnamon.

Solution: This is an example of epistasis. If you draw out the Punnett square, you will see the answer is B.

	B w	B w	b w	b w
b W	 B bWw	 B bWw	 b bWw	 b bWw
b w	 B bww	 B bww	 b bww	 b bww
c W	 B cWw	 B cWw	 b cWw	 b cWw
c w	 B cwW	 B cwW	 b cwW	 b cwW

Example 3.9: (USABO Semifinals 2018)

32. A good friend of yours, Dr. A, is breeding his favorite species of cats. For this specific experiment, he has a true breeding strand of red cats, and a true breeding strand of silver cats. When he crosses these two strands, he only gets golden cats. When he breeds these golden cats together, he gets 100 offspring: 56 golden cats, 19 silver cats, and 25 red cats. Which of the following statements must be true?

- A. This is an example of dominant epistasis
- B. This is not an example of recessive epistasis.
- C. The enzyme creating golden pigment comes before the enzyme creating silver pigment.
- D. The enzyme creating red pigment is mutated in the true breeding strand of silver cats.
- E. The enzyme creating red pigment is not mutated in the true breeding strand of gold cats.

Solution: We see that the ratio of the phenotypes is approximately gold:silver:red = 54:18:24 = 9:3:4, which is an example of recessive epistasis! Therefore, the pathway is red → silver → gold, which means that gold cats have to be able to produce red pigment. The answer is E.

Example 3.10: (USABO Semifinals 2018)

33. A well-established geneticist, Dr. L is famous for experiments on the common plant *Camellia sinensis*. In particular, Dr. L has a few mutant, true-breeding lines of *Camellia sinensis*. Line A gives off a sweet odor, line B gives off a musty odor, and line C gives off no odor at all. When he crosses lines A and B, he gets only sweet odor *C. sinensis*. When he crosses lines B and C, he gets only musty odor *C. sinensis*. Moreover, when he self-crosses the offspring of lines of lines A and B, he gets mostly sweet odor *C. sinensis*, some musty odor *C. sinensis*, and few no odor *C. sinensis*. Which of the following statements must be not true?
- A. This is an example of dominant epistasis.
 - B. This is not an example of pleiotropy.
 - C. If he self-crossed the no odor *C. sinensis* from the last cross, Dr. L would get some sweet odor *C. sinensis* offspring.
 - D. If he self-crossed the musty odor *C. sinensis* from the last cross, Dr. L would get some no odor *C. sinensis* offspring.
 - E. The enzyme responsible for creating the intermediate that gives off the sweet odor is not in the same pathway as the enzyme responsible for creating the intermediate that gives off the musty odor.

Solution: The various crosses here show us that these phenotypes exhibit dominant epistasis, and the ratio (12 : 3 : 1) is implicit in the final cross. The genotype of line A is AAbb (where A gives sweet odor, and is dominantly epistatic to B), the genotype of line B is aaBB, and the genotype of line C is aabb. The musty odor plants of the last cross give us aaB?, where the last allele is either B or b. Some will then be aaBb, so these can produce aabb offspring, which have no odor. Moreover, as this is dominant epistasis, the sweet odor “masks” the musty odor, so they can’t be in the same pathway, as one does not depend on the other. Therefore, the **answer is C** as no odor (aabb) can only produce more no odor.

3.3.9 Duplicate

Duplicate dominant/recessive diseases occur when multiple genes determine a single pathway, although they are fairly obscure.

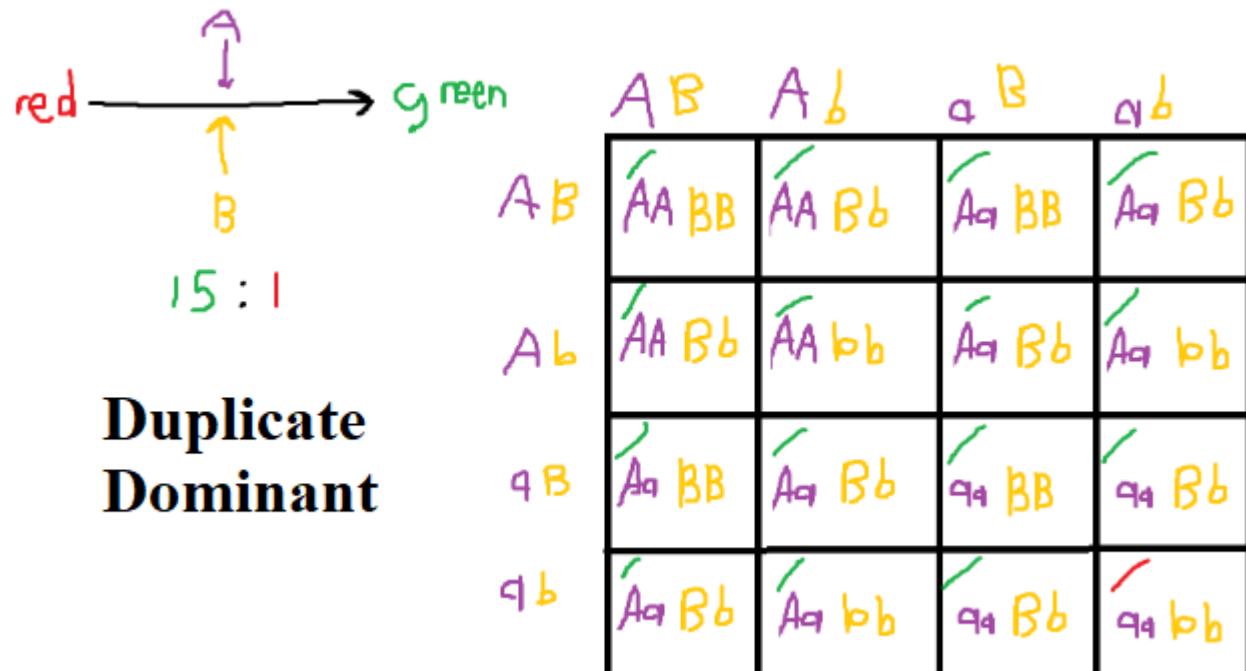


Figure 27: A duplicate dominant disease occurs when the dominant form of either gene (A or B) causes the trait. In this diagram, a pathway for a pigment shows how such a setup occurs. As the Punnett square shows, they exhibit a 15:1 phenotypic ratio. (Source: ANDrey)

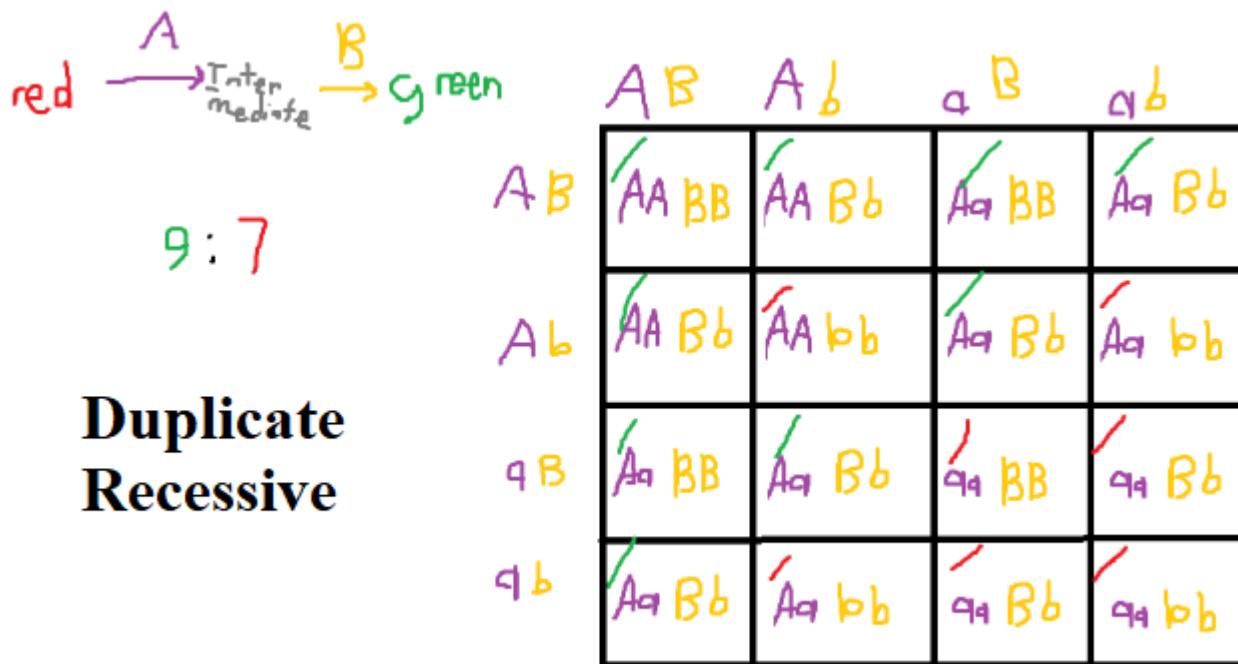


Figure 28: A duplicate recessive disease occurs when the recessive form of both genes (A and B) causes the trait and the intermediate has no effect on the phenotype. In this diagram, a pathway for a pigment with a colorless intermediate shows how such a setup occurs. As the Punnett square shows, they exhibit a 9:7 phenotypic ratio. (Source: ANDrey)

3.3.10 Other

Sometimes the relation between gene and effect is not 1-to-1:

- **Pleiotropy:** One gene affects multiple phenotypes.
- **Polygenic inheritance:** Multiple genes determine one phenotype. In these cases, the trait is often a **qualitative character** that exists across a spectrum, such as skin color or height in humans.
- **Multifactorial:** The phenotype is influenced by both genotype and environment. In this case, the phenotype exists in a **norm of reaction**, which is a spectrum of features based on the environment. For example, hydrangea plants change their color depending on the acidity of the soil they grow in (blue = acidic, pink = basic).
- **Imprinting:** The gene is on an autosome, yet its expression is still affected by which parent it is inherited from. This is often due to **epigenetic** changes in DNA, such as *methylation*. This was first discovered in the *igf-2* (insulin growth factor) gene in mice.

Example 3.11: (USABO Opens 2018)

35. Prader-Willi syndrome and Angelman Syndrome can each occur due to a deletion in the same part of chromosome 15. However, Prader-Willi syndrome will result if the defective chromosome was inherited from the father, while Angelman syndrome results if it was inherited from the mother. Identify the answer choice below that also exemplifies a similar inheritance pattern.

Choice	
A	Inheritance of Flower Color in Peas
B	Inheritance of Flower Color in Snapdragons
C	Inheritance of the Igf-2 gene in Mice
D	Inheritance of X-linked Color Blindness in Humans
E	Inheritance of Height in Humans

Solution: This question is describing imprinting, so the **answer is C**. A is dominant, B is incomplete dominance, D is X-linked recessive, and E is polygenic.

3.4 Mutations

3.4.1 DNA mutations

Mutations can occur in DNA when one base takes the place of another. To fully understand this, you should first check the DNA handout. Mutations can be classified based on how it changes the bases:

- **Point mutation:** When a single base is changed to another. An example of this is sickle-cell anemia.
- **Insertion/Deletion:** When a nucleotide is inserted to/removed from the sequence. If you insert/remove 3 nucleotides at a time, then you simply added/removed a codon. However, if

you insert/remove something that is not a multiple of 3, that causes a **frameshift mutation**. Basically, since codons are made of 3 nucleotides, if you add/remove something that is not a multiple of three, that means you're going to have to shift over all subsequent bases, changing your *reading frame*. Frameshift mutations are serious issue because they affect *all* codons that come afterwards instead of just 1.

We can also classify mutations based on their effect:

- **Silent mutation:** Some amino acids have multiple codons that code for them. As a result, you can change a codon, but not change the amino acid it codes for. Silent mutations produce no observable effect on the phenotype.
- **Missense mutation:** The new codon codes for a different amino acid.
- **Nonsense mutation:** The new codon is a stop codon, terminating translation early.

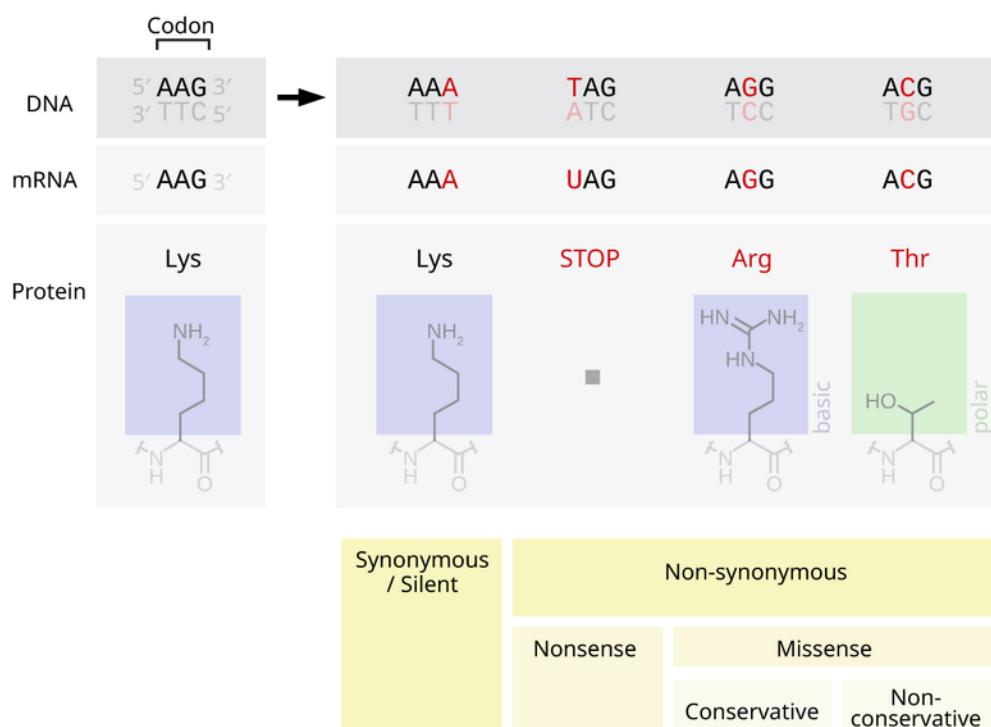


Figure 29: A diagram showing all the possible results of a point mutation. We did not discuss synonymous, conservative, or non-conservative mutation, but feel free to search them up yourself.
(Source: Wikipedia)

3.4.2 Chromosomal mutations

There are various ways chromosomes can get mutated, resulting in some notable diseases:

- **Cri du chat:** Caused by a deletion in chromosome 5. The name means “cry of the cat” because that is the sound that babies with this disease make. They usually die in infancy or early childhood.
- **Chronic myelogenous leukemia:** This occurs during the mitosis of white blood cells and can lead to cancer formation. What happens is that there is a translocation between the tip of chromosome 9 and 22. As a result, the chromosome 22 is shorter than usual and nicknamed the **Philadelphia chromosome**.

This is an example of a **Robertsonian translocation**, which is essentially when the translocation causes one chromosome to effectively become both chromosomes attached end to end, while the other chromosome is the short tips connected together. The Philadelphia chromosome is a result of a Robertsonian translocation.

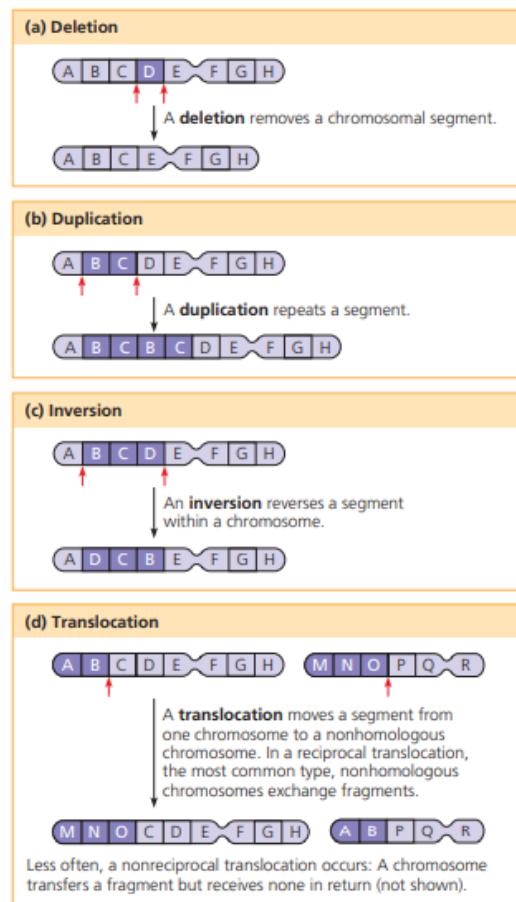


Figure 30: Types of chromosomal mutations. (Source: Campbell's 9th Edition)

4 Other Math

4.1 Bayes Rule

If I asked you what the probability is that a coin comes up heads when I flip it, you would probably say $\frac{1}{2}$. However, if asked you what the probability is that the coin comes up heads *given* that the past 100 flips all came up heads, you may start to rethink your answer. This phenomenon is known as *conditional probability*, where having additional information changes how likely it is for something to happen. People often use **Bayes rule** in these situations:

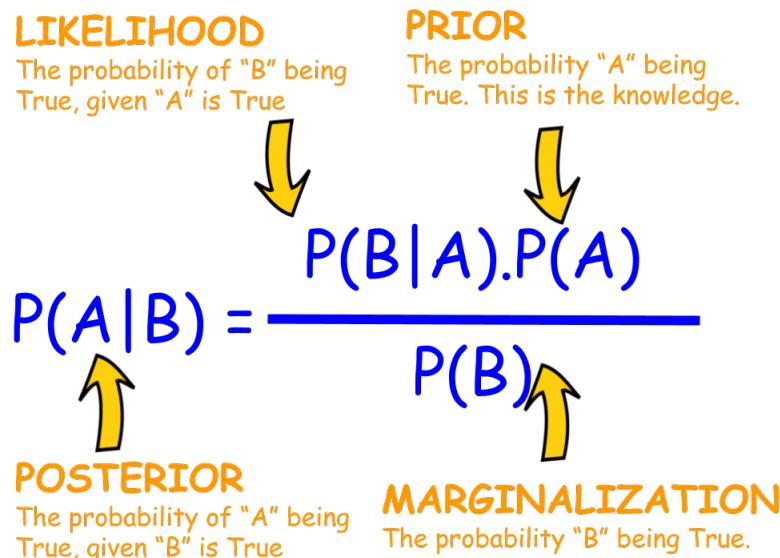


Figure 31: Bayes rule. (Source: Towards Data Science)

Note that $P(B)$ can be expanded to $P(B|A) * P(A) + P(B|\tilde{A}) * P(\tilde{A})$ where \tilde{A} is the conjugate of A . For example if you were trying to find the probability of a person being a carrier A would be the probability they are not the carrier.

However, that seems very mathy. Instead, I use a different approach that involves considering the probability that something happens in each scenario and then “averaging” those probabilities to get an answer. It is basically just Bayes rule, but it relies on logic instead of memorizing a formula.

For example, let’s go back to our coin. If I flip a coin and it comes up heads 3 times and I ask you what’s the probability that on the 4th flip it comes up heads, what would you say? Well, let’s consider each scenario:

1. The coin is a fair coin, in which case there is a $\frac{1}{2}$ chance each flip is heads. Therefore, the chance all three flips are heads is $(\frac{1}{2})^3 = \frac{1}{8}$.
2. Both sides of the coin are heads, in which case it will always come up heads and the probability is 1. We can ignore the scenario where both sides are tails since that will never come up heads.
3. Now we “average” these two probabilities. To find the actual probability that the coin is fair we do:

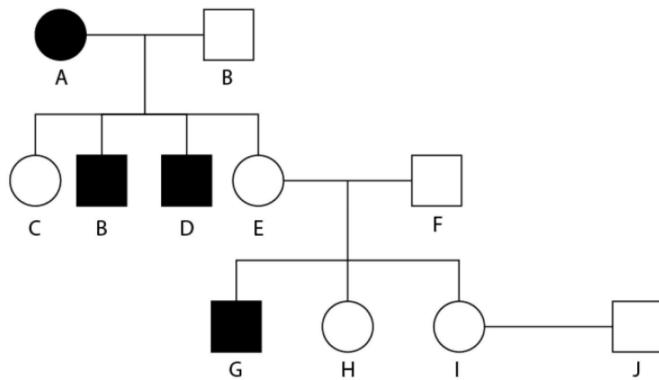
$$\frac{\frac{1}{8}}{\frac{1}{8} + 1} = \frac{1}{9}$$

Similarly, there is a $\frac{8}{9}$ chance the coin has both sides heads.

4. Now that we know how likely each scenario is, we can then find the chance that each scenario comes up heads on the 4th flip and add them together.

$$\frac{1}{9}(\frac{1}{2}) + \frac{8}{9}(1) = \frac{17}{18}$$

We can use this same approach of “averaging” scenarios to solve many conditional genetics problems...

Example 4.1: (USABO Semifinals 2020)

47. [1.5 pts] Individuals I and J return several years later, and excitedly inform you that they have had 3 boys and 1 girl, none of whom have had the disease. If they try for a 5th child, what is the probability that the fifth child will have the disease, based on the most probable inheritance pattern? Select ONE.

- a. 1/24
- b. 1/36
- c. 3/32
- d. 1/48
- e. 5/128

Solution: First we need to figure out what type of inheritance this disease has. We see that unaffected individuals (E and F) can produce affected offspring (G), so we know it is recessive. In addition, the fact that all of the males, but none of the females have the disease should make us think that it is X-linked. We can check this by assigning genotypes to all of the individuals and see that it works out! Notice in this problem you are given information about what has happened and we need to use that information to adjust our probability in real time. This is a marker that we need to use Bayes. Since it is X-linked recessive, the only way I ($X^A?$) and J ($X^A Y$) can produce an affected offspring is if I is a carrier ($X^A X^a$) and they have a son ($X^a Y$). We need to find the probability that I is a carrier given she has had 3 boys and 1 girl without the disease.

First, We can use the formulaic route. This can be written as:

$$P(\text{Carrier} \mid 3\text{B}1\text{G}) = \frac{P(4 \mid C) \times P(C)}{P(4 \mid C) \times P(C) + P(4 \mid \neg C) \times P(\neg C)}$$

where $P(C) =$ probability of carrier and $P(4) =$ probability of 4 unaffected children. Writing in the numbers we get:

$$P(\text{Carrier} \mid 3B1G) = \frac{\frac{1}{8}(\frac{1}{2})}{\frac{1}{8}\frac{1}{2} + \frac{1}{2}} = \frac{1}{9}$$

Therefore the probability of I being a carrier is 1/9 and the probability of them having an effected child is just 1/9 * 1/4 or 1/36.

Using the “approximation” route:

1. If I is not a carrier, then all of her children will be unaffected, so a chance of 1.
2. If I is a carrier, then all of her daughters will still be unaffected (since J gives a X^A). There is a $\frac{1}{2}$ chance that I gives the X^A to her son, making him unaffected. Therefore, the chance she has 3 unaffected sons is $(\frac{1}{2})^3 = \frac{1}{8}$
3. The probability I is a carrier is $\frac{\frac{1}{8}}{1+\frac{1}{8}} = \frac{1}{9}$
4. If I is a carrier, the fifth child will have the disease only if it is a boy (X^aY). There is a $\frac{1}{2}$ chance I passes on the recessive allele (X^a) and a $\frac{1}{2}$ chance J passes on the Y. Therefore, the answer is $\frac{1}{9}(\frac{1}{2})(\frac{1}{2}) = \frac{1}{36}$

The **answer is B.**

4.2 Coefficient of Relatedness

People have relatives and we all know that your brother is more closely related to you than your 3rd cousin twice removed, but what exactly does relatedness mean? The **coefficient of relatedness** is a measure of if an individual has a gene, what is the probability that their relative will also have that gene. Although that seems like a complicated question, I have a convenient trick for solving it:

1. Draw out the family tree.
2. Find a path where you go from person 1, to the common ancestor, back to person 2. If n is the number of “steps” it takes to complete the path, then find 2^{-n} .

The reason this works is that a person is equally likely to have received a gene from either parent. Therefore, each step “up” the family tree has a $\frac{1}{2}$ chance of coming from that parent. Similarly, a parent will randomly pass on one of their two alleles for that gene to their offspring. Therefore, each step “down” the family tree has a $\frac{1}{2}$ chance of passing the allele in question.

3. Repeat step 2 for all possible paths and sum up your values together. This is your coefficient of relatedness.

4.2.1 Haplo-Diplo systems

In ants, bees, and wasps, males are haploid while females are diploid. Males are born from unfertilized eggs while females are born from fertilized eggs. To solve for coefficient of relatedness in this scenario, we do virtually the same thing except a step “up” from a male bee to their mother will have a 1 chance of happening instead of $\frac{1}{2}$, since the male receives *all* of their genes from the mother. Similarly, a step “down” from a male bee to their daughter will also have a 1 chance of happening instead of $\frac{1}{2}$, since the male is haploid and has only 1 copy of each gene, which he *must* give to his offspring.

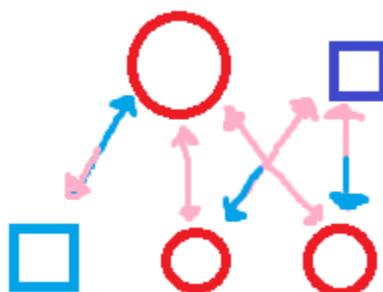
Example 4.2: (USABO Semifinals 2017)

Questions 31 to 33. Honeybees are haplodiploid and reproduce sexually, meaning that females are diploid and males are haploid. In a given hive, only the queen bee reproduces – the other females are worker bees and do not mate even though they may be fertile. In one hive, the queen bee has just mated with a single unrelated drone (a male bee). Their offspring includes a number of females and males. Use the following choices to answer Questions 32 to 34:

- A. 0.125.
- B. 0.25.
- C. 0.50.
- D. 0.75.
- E. 1.00.

31. What is the coefficient of relatedness between two of the female offspring?
32. What is the coefficient of relatedness between a female offspring and her brother?
33. What is the coefficient of relatedness between a male offspring and his sister?

Solution: In the diagram below of the bee family tree, pink arrows represent a relatedness of $\frac{1}{2}$ and blue arrows represent a relatedness of 1. For 31, we see that through the mother, the relatedness is $\frac{1}{2}(\frac{1}{2}) = \frac{1}{4}$ and through the father, the relatedness is $\frac{1}{2}(1) = \frac{1}{2}$. So, $\frac{1}{4} + \frac{1}{2} = .75$, **D is the answer**. For 32, we see that the only path is through the mother, which is $\frac{1}{2}(\frac{1}{2}) = .25$, **B is the answer**. For 33, we see that the only path is through the mother, which is $1(\frac{1}{2}) = .5$, **C is the answer**. Note that for each question that asked for “the coefficient of relatedness between A and B”, we started on A and found paths to B. The fact that we got different answers for 32 and 33 shows that for bees, order does matter.



4.2.2 Coefficient of Inbreeding

Incest occurs when you cross two members of the same family. As a result, there is the possibility that the offspring could receive two copies of the *same* allele. This can make it easier for the offspring to get two copies of a recessive allele, leading to many unique genetic deformities (such as the Habsburg jaw). The **coefficient of inbreeding** is defined as the probability that the offspring will receive two copies of the same allele. As a result it is just

$$\frac{1}{2} \times (\text{coefficient of relatedness between the two parents})$$

This is true because:

1. The offspring has to receive the allele from at least one of their parents, so we can pick a parent to start from.
2. From there, the probability that the other parent has the allele is just the coefficient of relatedness.
3. Then, that other parent has a $\frac{1}{2}$ chance of passing that allele to the offspring. Again, haploid-diplo systems complicate things.

4.3 Hamilton's Rule

From an evolutionary standpoint, the meaning of life is to pass on your genes to the next generation. Assuming that an organism wants to optimize how many of its genes are passed on, there are some cases where an organism may sacrifice itself to let a relative survive. For these cases, we use **Hamilton's Rule** to figure out whether an organism will sacrifice itself. It is often written as:

$$rB > C$$

where r is the coefficient of relatedness, B is the benefit, and C is the cost.

According to Hamilton's rule, an organism will sacrifice itself when the genes "gained" is greater than the genes "lost" due to the sacrifice. How this is actually calculated is best shown through examples...

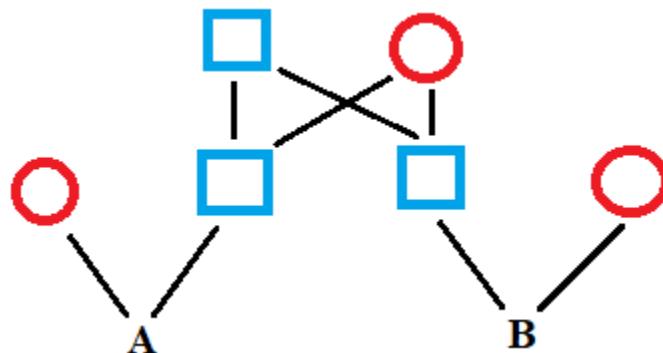
Example 4.3: (USABO Opens 2016)

32. Person A and Person B, who are cousins, are both childless. Their fathers were brothers, while their mothers are unrelated. Person B is engaged and intends to have nine children. Person A is married expects to have only one child. One day, Person B falls into a river, from which he cannot swim out alone and would die otherwise. Person A has a 40% chance of saving Person B if he enters the river. However, entering the river also means certain death for Person A, regardless of whether or not he saves Person B.

Calculate the coefficient of relatedness between Person A and Person B and decide whether, based on Hamilton's rule, Person A should try to save Person B.

- A. 0.25; Person A should try to save B.
- B. 0.125; Person A should NOT try to save B.
- C. 0.25; Person A should try to save B.
- D. 0.0625; Person A should NOT try to save B.
- E. 0.125; Person A should try to save B.

Solution: Let's draw our family tree. We see that there are two paths between A and B, through paternal grandfather and grandmother. Each path has 4 steps, so the coefficient of relatedness is $\frac{1}{16} + \frac{1}{16} = \frac{1}{8} = .125$. Now for Hamilton's rule, the cost is 1, as A will definitely die and his child will not be born. The gain is $.125(\frac{2}{5} \text{ chance of saving B})(9 \text{ children}) = \frac{9}{20} = .45$. Since $.45 < 1$, A should NOT save B. The **answer is B**.



5 Conclusion

That's basically everything you need to know for genetics! It is definitely the most mathy part of biology, but through all of the examples in this handout, hopefully you've developed an intuition on how to solve these problems.